Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **537,** 3396–3420 (2025) https://doi.org/10.1093/mnras/staf237
Advance Access publication 2025 February 13

Time-dependent models of AGN discs with radiation from embedded stellar-mass black holes

Marguerite Epstein-Martin [1]★ Hiromichi Tagawa,[2] Zoltán Haiman [1,3] and Rosalba Perna [4,5]

[1]*Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY 10027, USA*
[2]*Shanghai Astronomical Observatory, Shanghai 200030, China*
[3]*Department of Physics, Columbia University, 550 West 120th Street, New York, NY 10027, USA*
[4]*Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794-3800, USA*
[5]*Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA*

Accepted 2025 January 29. Received 2025 January 28; in original form 2024 May 14

ABSTRACT

The brightest steady sources of radiation in the universe, active galactic nuclei (AGNs), are powered by gas accretion on to a central supermassive black hole (SMBH). The large sizes and accretion rates implicated in AGN accretion discs are expected to lead to gravitational instability and fragmentation, effectively cutting off mass inflow to the SMBH. Radiative feedback from disc-embedded stars has been invoked to yield marginally stable, steady-state solutions in the outer discs. Here, we examine the consequences of this star formation with a semi-analytical model in which stellar-mass black hole (sBH) remnants in the disc provide an additional source of stabilizing radiative feedback. Assuming star formation seeds the embedded sBH population, we model the time-evolving feedback from both stars and the growing population of accreting sBHs. We find that in the outer disc, the luminosity of the sBHs quickly dominates that of their parent stars. However, because sBHs consume less gas than stars to stabilize the disc, the presence of the sBHs enhances the mass flux to the inner disc. As a result, star formation persists over the lifetime of the AGN, damped in the outer disc, but amplified in a narrow ring in the inner disc. Heating from the embedded sBHs significantly modifies the disc's temperature profile and hardens its spectral energy distribution, and direct emission from the sBHs adds a new hard X-ray component.

Key words: accretion, accretion discs – stars: black holes – galaxies: active.

1 INTRODUCTION

Active galactic nuclei (AGNs) are widely understood to be the result of viscous accretion on to a central supermassive black hole (SMBH; $\gtrsim 10^6$ $M_\odot$). Continuum spectra suggest that bright AGNs are geometrically thin and optically thick (Shields 1978; Malkan & Sargent 1982), conventionally modelled using the Shakura−Sunyaev α prescription (Shakura & Sunyaev 1973). These 'α-disc' models are, however, well known to break down at large radii, becoming unstable to self-gravity beyond ~ 0.1 parsec (pc) and failing to transport mass quickly enough to maintain the disc beyond ~ 1 pc. In particular, the latter is expected to give rise to cascading fragmentation and star formation on a dynamical time-scale – rapidly transforming the disc into a flat stellar system and cutting off accretion on to the central SMBH (Shlosman & Begelman 1989; Shlosman, Begelman & Frank 1990; Goodman 2003).

In an effort to resolve this apparent contradiction, Collin & Zahn (1999) suggested that instability due to self-gravity may be mitigated by feedback from star formation including via stellar winds, supernovae explosions, and stellar accretion. Since then several models have been proposed which hinge on the assumption that fragmentation is self-limiting. In particular, Sirko & Goodman (2003) and Thompson, Quataert & Murray (2005, hereafter TQM05) have constructed modified viscous disc models in which the disc is assumed to maintain marginal stability at large radii.

Though similar in their approach, these models are distinguished in their treatment of mass inflow through the AGN, resulting in significant discrepancies in predicted disc density and scale height profiles. Sirko & Goodman (2003) are strongly motivated by modelling the AGN infrared (IR) spectral energy distributions (SEDs) and assume a constant rate of mass inflow to find the radiation pressure necessary to support the disc. In contrast, TQM05 allow for changes in mass inflow, accounting for the star formation required to support the disc. Given the explicit relationship with star formation and more physically motivated mass dependencies, in the work that follows we will rely heavily on the original work of TQM05 [see section 2.2 of Fabj et al. (2020) and Gangardt et al. (2024) for detailed comparisons].

The recent discoveries of gravitational waves (GWs) by LIGO-Virgo-KAGRA have focused renewed attention on AGN discs, whose high density of embedded stars and compact objects make them an intriguing site for mergers and subsequent GW detections. Whether captured from the nuclear population (Syer, Clarke & Rees 1991; Bartos et al. 2017; Panamarev et al. 2018; Fabj et al. 2020; MacLeod & Lin 2020) or formed *in situ* (Stone, Metzger & Haiman

★ E-mail: mae2153@columbia.edu

2017), once aligned with the disc, embedded objects are expected to exchange torque with the disc gas and migrate through the AGN. Migration can enhance the formation of new binaries via low-velocity encounters (McKernan et al. 2012; Leigh et al. 2018; Tagawa, Haiman & Kocsis 2020) or upon entering annular gaps (Tagawa et al. 2020) or migration traps (Bellovary et al. 2016; Secunda et al. 2019; Yang et al. 2019; Secunda et al. 2020; Grishin, Gilbaum & Stone 2024), whereupon gas dynamical friction and binary-single encounters act to harden binaries and promote mergers (Baruteau, Cuadra & Lin 2011; Tagawa et al. 2020; Li et al. 2021; Tagawa et al. 2021a, b; Li & Lai 2024; Boekholt, Rowan & Kocsis 2023; DeLaurentiis, Epstein-Martin & Haiman 2023; Li, Rodet & Lai 2024).

Having shown that AGN discs can significantly alter the dynamics of stellar and compact object populations, it is pertinent to ask how such discs in turn may be affected by the objects they harbour. Levin (2003, 2007) concluded that if BHs accreted at super-Eddington Bondi-limited rates, their luminosity would be sufficient to stabilize the AGN. However, embedded black holes accreting at highly super-Eddington rates for extended periods will quickly grow to intermediate masses, disrupting and depleting the AGN disc interior (Goodman & Tan 2004; McKernan et al. 2012; Stone et al. 2017). Several feedback mechanisms have been suggested that could slow this accretion process (Inayoshi, Haiman & Ostriker 2016; Tagawa et al. 2022). Gilbaum & Stone (2022) have recently developed a detailed model in which they assume pressure support from remnant stellar mass black holes (sBHs) fully overtakes the stellar component within one generation of massive star formation. They then calculate the effect of sBH radiation pressure support on a marginally stable, steady-state disc.

In this paper, we develop a semi-analytic model for time-evolving AGN discs under the concurrent effects of star formation and sBH accretion, with the goal of understanding how a growing population of remnants impacts the disc structure, stellar distribution and electromagnetic spectrum of the disc over time. Assuming star formation seeds the population of sBHs, in Section 2 we introduce a set of equations to describe the evolution of the disc together with the embedded stars and sBHs. We use these equations in Section 3 to discuss several time-scales relevant to this model and define the parameters determining the number and mass distribution of disc-born stars. In Section 4, we present order-of-magnitude estimates for the relation between mass accretion and radiation pressure from the two distinct populations of stars and sBHs. The numerical approach we used is explained in Section 5. Our models build in complexity, first comparing AGN discs supported either by star formation or by sBHs alone (Section 6) and their respective spectral signatures. Then, in Section 7, we model the response of the AGN disc to the combined effects of star formation and a growing populations of sBHs, approximating the evolution as a sequence of steady states. We use these disc models to construct time evolving SEDs. In Section 8, we examine the dependence of our models on disc viscosity and SMBH mass. In Section 9, we summarize our main conclusions and note avenues for future study.

2 NUMERICAL MODEL OVERVIEW

We constructed our model in the steady-state Shakura–Sunyaev mould (Shakura & Sunyaev 1973), assuming a geometrically thin, axisymmetric disc in local thermal equilibrium (LTE), undergoing quasi-viscous transport. As in TQM05, our model is distinguished from the standard α-disc by requiring gravitational stability in the outer disc. The stability of a disc against small overdensities is typically represented by the dimensionless Toomre parameter Q_T (Toomre 1964), where $Q_{\rm T} \lesssim 1$ indicates instability. Maintaining marginal stability requires

$$Q_{\rm T} \simeq \frac{c_{\rm s}\Omega}{\pi G \Sigma} = \frac{\Omega^2}{\sqrt{2}\pi G \rho} \geq 1, \tag{1}$$

where $c_{\rm s}$ is the sound speed related to the scale height h by $c_{\rm s} = h\Omega$. $\Sigma = 2\rho h$ is the surface density, ρ is the disc density, G is the gravitational constant, and Ω is the orbital frequency, given by

$$\Omega = \left(\frac{G M_{\rm sBH}}{r^3} + \frac{2\sigma^2}{r^2} \right)^{1/2}. \tag{2}$$

Here, r is the distance from the SMBH, and σ is the velocity dispersion characterizing the gravitational potential on galactic scales, beyond the central SMBH's sphere of influence. In general, σ scales with the mass of the central SMBH as $(\sigma/200\ {\rm km\,s^{-1}})^4 \sim M_{\rm sBH}/(2\times10^8\ {\rm M_\odot})$ (Kormendy & Ho 2013). Note that on parsec to tens of parsec scales, and for SMBH masses 10^6–$10^9\ {\rm M_\odot}$, the second term in equation (2) is of the same order of magnitude as the standard Keplerian frequency. But in the interior disc the effect of the velocity dispersion term is negligible.

From equation (1), it follows that the density profile of the outer disc, where $Q = 1$ is imposed, is solely determined by the orbital frequency. Moreover, from the Shakura–Sunyaev α viscosity parameter $\nu = \alpha c_s h$ and the continuity equation, the relationship between density and scale height may be calculated in terms of the mass accretion rate, i.e.

$$\dot{M} = 2\pi\nu\Sigma \left| \frac{{\rm d}\ln\Omega}{{\rm d}\ln r} \right| = 4\pi G\alpha\Omega\rho h^3 \left| \frac{{\rm d}\ln\Omega}{{\rm d}\ln r} \right|. \tag{3}$$

Equations (1)–(3) are identical to those presented in appendix C of TQM05. Our model for the disc differs by the addition of new heating and mass accretion terms representing feedback from *both* progenitor stars and their remnant black holes. In this disc, gas is consumed by star formation ($\dot{\Sigma}_\star$) and sBH accretion ($\dot{\Sigma}_{\rm BH}$), so that the disc accretion rate decreases inwards as

$$\dot{M}(r) = \dot{M}_{\rm out} - \int_r^{R_{\rm out}} 2\pi r(\dot{\Sigma}_\star + \dot{\Sigma}_{\rm BH}){\rm d}r, \tag{4}$$

where $\dot{M}_{\rm out}$ is the mass supplied at the outermost radius of the disc ($R_{\rm out}$).

In our models, marginal stability is maintained by vertical pressure support from star formation and sBH accretion. Following TQM05, we calculate the pressure associated with star formation as

$$p_\star = \dot{\Sigma}_\star c \epsilon_\star \left(\frac{\tau}{2} + \xi \right), \tag{5}$$

representing two distinct components: the radiation pressure on dust grains in the optically thick limit and the UV radiation pressure and turbulent support by supernovae in the optically thin limit. Here, $\epsilon_\star$ is the matter–radiation conversion efficiency of stars in the disc, discussed in more detail in Section 3.2. The kinetic pressure associated with star formation is parametrized by the non-dimensional ξ, the ratio of star formation to supernovae pressure, which we set to $\simeq 1$ as in TQM05.

Radiation pressure scales with the optical depth ($\tau = \kappa\rho h$) where κ is the opacity. We calculate opacity according to Semenov et al. (2003).[1] Because these opacity tables only extend to 10^4 K, we follow the approach of TQM05, and smoothly connect the opacity to power

[1]The Semenov et al. (2003) opacity tables do not include graphite grain opacity as discussed in Baskin & Laor (2018), which dominates for hot dust.

laws given by Bell & Lin (1994) for temperatures exceeding 10^4 K. Note that at temperatures of ≈ 1500 K dust sublimation results in a steep opacity drop. The opacity rises again at $\approx 10^4$ K, with the ionization of hydrogen, creating a feature known as the 'opacity gap' between 10^3 and 10^4 K. This drop in opacity requires us to consider both the optically thick ($\tau \gg 1$) and thin ($\tau \ll 1$) regimes, in which the temperature varies with opacity as $T^4 \sim \tau T_{\rm eff}^4$ and $T^4 \sim \tau^{-1} T_{\rm eff}^4$, respectively. Interpolating between these two regimes, the temperature is

$$T^4 = T_{\rm eff}^4 \left(\frac{3}{4}\tau + \frac{1}{2\tau} + 1 \right) . \tag{6}$$

Here, the effective temperature $T_{\rm eff}$ is calculated assuming thermal equilibrium in the disc,

$$\sigma_{\rm SB} T_{\rm eff}^4 = \frac{3}{8\pi}\dot{M}\left(1 - \sqrt{\frac{R_{\rm in}}{r}}\right)\Omega^2 + \frac{\dot{\Sigma}_\star \epsilon_\star c^2}{2} + \frac{Q_{\rm sBH}}{2}, \tag{7}$$

where $\sigma_{\rm SB}$ is the Stefan–Boltzmann constant, $Q_{\rm sBH}$ represents the heating rate associated with sBH accretion, and $R_{\rm in}$ is the innermost radius of the disc. Using equations (5) and (7), the total disc pressure can be written as

$$p_{\rm tot} = \rho c_{\rm s}^2 = p_{\rm gas} + \frac{\tau}{c}\sigma_{\rm SB} T_{\rm eff}^4 + \dot{\Sigma}_\star \epsilon_\star c \xi, \tag{8}$$

where $p_{\rm gas} = \rho k_{\rm B} T/(\mu m_{\rm p})$, $k_{\rm B}$ is the Boltzmann constant, μ is the gas mean molecular weight, and m_p is the proton mass. Here, we set $\mu = 1.23$, appropriate for neutral gas with a primordial He/H mass ratio.

Having introduced $Q_{\rm sBH}$, the sBH heating term, we require at least one more equation to properly close our model. If we assume the number density of sBHs in the disc at a given time and radius is determined by the accumulated star formation, a self-consistent solution can be reached. We begin with an initial mass function (IMF) of the form $\mathrm{d}N_\star/\mathrm{d}m_\star = A_{\rm x} m_\star^{-\delta}$. Taking the mass of stars per unit disc surface area formed up to time t to be $\int \dot{\Sigma}_\star(t)\mathrm{d}t$ we can calculate the proportionality constant $A_{\rm x}$:

$$A_{\rm x}(t) = \frac{\int_0^t \dot{\Sigma}_\star(t)\mathrm{d}t}{\int_{m_{\rm min}}^{m_{\rm max}} m_\star^{1-\delta}\mathrm{d}m_\star} . \tag{9}$$

The minimum ($m_{\rm min}$) and maximum ($m_{\rm max}$) stellar mass are free parameters that we take to be 0.1 and 120 $\mathrm{M}_\odot$, respectively.

The number of remnant black holes can be found by taking the lower bound of the IMF and setting it to the turn-off mass ($m_{\rm TO}$) or the time-dependent mass at which stars are expected to evolve off the main sequence (MS). For ease of use we approximate the turn-off mass by a piecewise function,

$$m_{\rm TO} = \begin{cases} m_{\rm max} & t \le t_1 \\ 120 \times 10^{-\sqrt{(\log_{10}(t/\mathrm{yr}) - 6.43)/0.825}} & t_1 < t < t_2 \\ m_{\rm trans} & t \ge t_2 \end{cases} . \tag{10}$$

Here, $m_{\rm trans} \simeq 20\,\mathrm{M}_\odot$ is the transition progenitor mass between neutron stars (NSs) and sBHs. t_1 and t_2 represent the time for the first and last sBH to be produced. That is, $t_2 \simeq 10.0$ Myr is the time at which the smallest sBH progenitors evolve off the MS, and $t_1 \simeq 2.7$ Myr is the minimum time required for the first sBHs to form. Between t_1 and t_2, we use the fitting function for $m_{\rm TO}$ given by equation (3) of Buzzoni (2002). The number of sBHs per unit disc area at time t can thus be written as

$$S_{\rm sBH}(t) = \int_{t_0}^{t} A_{\rm x}(t') \int_{m_{\rm TO}(t-t')}^{m_{\rm max}} m_\star^{-\delta}\mathrm{d}m_\star \mathrm{d}t' . \tag{11}$$

With the number of sBHs in hand, calculating the corresponding heating and mass accretion terms is relatively straightforward. We assume an accretion rate on to embedded sBHs given by

$$\dot{m}_{\rm sBH} = \min\left(\dot{M}_{\rm Edd}, \dot{m}_{\rm B}\right), \tag{12}$$

i.e. capped at the Eddington rate:

$$\dot{M}_{\rm Edd} = \frac{4\pi G m_{\rm sBH} \mu_{\rm e} m_{\rm p}}{c\eta\sigma_{\rm T}} . \tag{13}$$

Here, $\mu_{\rm e} \sim 1.15$ is the mean weight per electron, $\sigma_{\rm T}$ is the Thompson cross-section, $m_{\rm p}$ is the proton weight, and $\eta = 0.1$ is the radiative efficiency.

At sufficiently low density or scale height, the gravitational sphere of influence of an embedded sBH will not contain enough gas to sustain Eddington accretion, and the accretion will instead be set by the Bondi rate:

$$\dot{m}_{\rm B} = \pi r_{\rm w} r_{\rm h} \rho c_{\rm s}, \tag{14}$$

where $r_{\rm w}$ and $r_{\rm h}$ are the width and height of the cross-section of accretion. Oriented parallel to the plane of the disc, the width is calculated as

$$r_{\rm w} = \min\left(R_{\rm B}, R_{\rm Hill}\right), \tag{15}$$

where $R_{\rm Hill} = r(m_{\rm sBH}/M)^{1/3}$ is the Hill radius (Murray & Dermott 1999) and $R_{\rm B} = G m_{\rm sBH}/c_{\rm s}^2$ is the Bondi radius (Bondi 1952). Perpendicular to the plane of the disc, $r_{\rm h} = \min(r_{\rm w}, h)$, ensuring that the cross-sectional area does not extend above the vertical height of the disc (Stone et al. 2017; Rosenthal et al. 2020; Dittmann, Cantiello & Jermyn 2021).

Within the SMBH's sphere of influence $\Omega \simeq \sqrt{GM/r^3}$, and with some algebraic manipulation, we can re-define the three length-scales – $R_{\rm B}$, $R_{\rm Hill}$, and h – relative to one another: $R_{\rm B} > R_{\rm Hill}$ if $h < \sqrt{3} R_{\rm Hill}$. In terms of units relevant to this work, for an SMBH mass of $4 \times 10^6 \mathrm{M}_\odot$ and sBH mass of $10\,\mathrm{M}_\odot$, $\dot{m}_{\rm B} \propto R_{\rm B}^2$ where $h/r > 1.6 \times 10^{-2}$. In practice, this condition is nearly always satisfied.

Having specified the accretion rate for individual sBHs, the total sBH accretion per unit disc area can be written as

$$\dot{\Sigma}_{\rm sBH} = \sum_i S_{\rm sBH}(m_{\rm sBH_i})\dot{m}_{\rm sBH}, (m_{\rm sBH_i}), \tag{16}$$

where we have summed the accretion rate per unit area for sBHs of mass $m_{\rm sBH_i}$. The heating term is similarly

$$Q_{\rm sBH} = \sum_i S_{\rm sBH}(m_{\rm sBH_i})\dot{m}_{\rm sBH}(m_{\rm sBH_i})\eta(m_{\rm sBH_i})c^2 = \dot{\Sigma}_{\rm sBH}\eta c^2 . \tag{17}$$

Here, we assume that the radiative efficiencies during Bondi-limited and Eddington-limited accretion phases are the same, and set it to $\eta(m_{\rm sBH_i}) = \eta = 0.1$, for simplicity.[2]

In the high-density environment expected in AGN, the Bondi rate generally results in super-Eddington accretion. In our models, this means accretion is nearly always Eddington capped. By limiting accretion to the Eddington rate, as in previous works (i.e. Tagawa

This additional component increases the opacity below 2000 K, more tightly constraining the opacity gap. However, we do not expect this adjustment to qualitatively change the conclusions made here and defer its inclusion to future work.

[2]In our models, sBH accretion does not fall below $\sim 0.01\,\dot{M}_{\rm Edd}$, justifying this assumption. If this were not the case, our models would need to be amended to account for the lower radiative efficiencies expected in the advection-dominated regime.

et al. 2020; Gilbaum & Stone 2022), the sBHs are prevented from quickly growing to intermediate sizes and depleting the disc interior (Goodman & Tan 2004; Stone et al. 2017). On the micro-scale, Eddington-limited accretion is motivated by significant radiation pressure acting on dust grains in high-metallicity gas (Toyouchi et al. 2019) or by wind mass-loss (Blandford & Begelman 1999). Note that the latter can further reduce the mass flux through the disc.

3 PARAMETER CHOICES AND REGIMES OF INTEREST

Our fiducial model assumes parameters appropriate to the Galactic Centre, namely an SMBH mass $M_{\rm BH} = 4 \times 10^6$ M$_\odot$ and velocity dispersion $\sigma = 75$ km s^{-1}. The outer edge of the disc is defined as the radius where the velocity dispersion equals the Keplerian velocity or $R_{\rm out} = GM_{\rm BH}\sigma^{-2} \simeq 3$ pc. The inner edge of the disc is assumed to be at the innermost stable circular orbit (ISCO) for a non-rotating black hole or $R_{\rm in} = 3\left(2GM_{\rm BH}c^{-2}\right) \simeq 10^{-6}$ pc. Following TQM05, we increase opacity by a factor of 3 to account for super solar metallicity in the Galactic Centre. We expand our discussion to a higher mass AGN in Section 8, setting $M_{\rm BH} = 10^9$ M$_\odot$. Details of other relevant parameter choices are discussed below.

3.1 Relevant time-scales

As mentioned in the Introduction, one problem for AGN disc models is the long viscous time-scales predicted by the 'α'-viscosity parametrization, which suggests outer viscous time-scales at least an order of magnitude longer than anticipated AGN lifetimes (Shlosman & Begelman 1989). Most proposed resolutions of this problem invoke increased angular momentum transport efficiency via non-axisymmetric global-torques, which could be provided by stellar bars, spiral waves, or magnetic stresses (Shlosman et al. 1990; Hopkins & Quataert 2011). Gilbaum & Stone (2022) showed, however, that the luminosity from a growing population of sBHs may serve to increase the disc scale height, which scales inversely with the viscous time-scale, partially alleviating the inflow problem. In order to similarly examine the relative change in viscous time-scale we also assume a standard local α-viscosity prescription. The kinematic viscosity may then be written as $\nu = \alpha c_s h$, with the dimensionless parameter α set to 0.3 in this work.

The viscous time-scale t_ν is also important in determining the critical mass supply rate ($\dot{M}_{\rm c}$), identified as the minimum $\dot{M}_{\rm out}$ below which star formation in the outer disc significantly depletes the disc mass. This limit corresponds to when the viscous time-scale (t_ν) is shorter than the star formation time-scale or $t_{\rm SF} < t_\nu$ where $t_{\rm SF} = \Sigma/\dot{\Sigma}_\star$. By making a few physically motivated approximations at the outer disc we can use this inequality to calculate $\dot{M}_{\rm c}$. Namely, we assume $Q_{\rm T} = 1$, and that the disc is optically thick and radiation-pressure dominated. In the outer disc dust opacity dominates which has the temperature dependence $\kappa = \kappa_0 T^2$. This yields a star-formation time-scale of $t_{\rm SF} \propto \epsilon_\star \kappa_0 T^2$ and a viscous time-scale $t_\nu \propto r^2 T^2/\dot{M}$. The critical mass supply rate is then expressed as (see also equation 44 in TQM05 and the appendix of Inayoshi & Haiman 2016),

$$\dot{M}_{\rm c} = 0.26 \left(\frac{r}{3\ {\rm pc}}\right)^2 \left(\frac{\kappa_0}{7.2\times10^{-4}{\rm cm^2\,g^{-1}\,K^{-2}}}\right)^{-1} \times \left(\frac{\epsilon_\star}{8.3\times10^{-4}}\right)^{-1} \left(\frac{M_{\rm BH}}{4\times10^6\ {\rm M_\odot}}\right)^{-1} \dot{M}_{\rm Edd}\,. \quad (18)$$

Note that to find $\dot{M}_{\rm c}$ for a disc supported by sBHs, we replace $\epsilon_\star$ with $\eta = 0.1$ and find $\dot{M}_{\rm c} \simeq 10^{-3}\ \dot{M}_{\rm Edd}$ – far below the parameter space with which we are concerned here. In our disc models, it can thus be safely assumed that $\Sigma/\dot{\Sigma}_{\rm sBH} \gg t_\nu$.

Another important time-scale is the lifetime of the AGN disc ($t_{\rm AGN}$). The AGN duty cycle is still somewhat controversial, with expected lifetimes ranging from 10 to 10^3 Myr (Martini 2004). In the simplest case, these lifetimes are consistent with the mass-independent, Eddington limited e-folding time for the SMBH, i.e. $t_{\rm AGN} = \eta\sigma_{\rm T}c/(4\pi G m_p) \sim 50(\eta/0.1)$ Myr (Salpeter 1964). In the shortest AGN lifetime limit, massive sBH progenitors will not have time to form and leave the MS. Therefore, we discount these regimes focusing on AGN lifetimes exceeding 10^7 years.

There are two additional physically relevant time-scales: the thermal ($t_{\rm th}$) and the dynamical ($t_{\rm dyn}$) time. The latter scales as Ω^{-1} and the thermal time-scale, or the disc diffusion time-scale, is given by $\Sigma c_{\rm s}^2/(\sigma_{\rm b} T_{\rm eff}^4) \simeq \Sigma c_{\rm s}^2/(\epsilon_\star \dot{\Sigma}_\star c^2/2)$. In the optically thick, radiation-pressure dominated limit, the ratio of the thermal and dynamical time-scales is $t_{\rm th}/t_{\rm dyn} \sim \tau c_{\rm s}/c$. Because τ never exceeds 1000 in our models, we conclude that $t_{\rm th}/t_{\rm dyn} \ll 1$. Where the disc is optically thin, $t_{\rm SF} \simeq t_\nu$. Substituting this into $t_{\rm th}$, we find $t_{\rm th}/t_{\rm dyn} \simeq 2(\Omega r)^2/(\epsilon_\star \alpha c^2) \ll 1$. This inequality implies that the disc should self-regulate, such that $Q \sim 1$ as our model assumes.[3]

3.2 Star formation

In determining the number and mass distribution of disc-born stars, we must define several parameters including the minimum ($m_{\rm min}$) and maximum ($m_{\rm max}$) stellar masses, the IMF power-law index δ, and the efficiency of converting stellar mass to radiation ($\epsilon_\star$). In this work, we assume a minimum stellar mass of 0.1 M$_\odot$, the approximate mass required to ignite thermonuclear reactions in the stellar interior. The maximum mass of 120 M$_\odot$ is less rigorously defined, and is likely a lower limit (Cantiello, Jermyn & Lin 2021; Dittmann et al. 2021; Jermyn et al. 2022), but given $m_{\rm min} \ll m_{\rm max}$ and therefore $N_\star \sim A_x m_{\rm min}^{1-\delta}$, its value is unlikely to strongly affect the outcome of our models.

We consider two different IMF slopes: the standard Salpeter IMF $\delta = 2.35$ determined using the luminosity function of stars in the solar neighbourhood (Salpeter 1955) and a top-heavy IMF $\delta = 1.7$, an observational estimate based on the Milky Way's nuclear stellar disc (Lu et al. 2013). These δ values also encompass the range of IMF slopes investigated in the AGN models of Tagawa et al. (2020).

The stellar mass conversion efficiency has a range of values in the literature, likely because the calculation requires approximations of the lifetime of the AGN as well as stellar lifetimes and luminosities across the stellar mass spectrum. In the original work of TQM05, an efficiency of $\epsilon_\star = 10^{-3}$ was assumed for an IMF slope $\delta = 2.35$, although the stellar lifetime was not considered in their calculation. More recently, Tagawa et al. (2020) recalculated the efficiency including the effect of stellar and AGN lifetimes, calculating an efficiency of 1.5×10^{-4} and 7.7×10^{-4} for an IMF exponent $\delta = 2.35$ and 1.7, respectively, an AGN lifetime of $t_{\rm AGN} = 10^8$ yr,

[3]We can similarly compare $t_{\rm dyn}$ and $t_{\rm SF}$, finding that $t_{\rm SF}/t_{\rm dyn} \sim \tau\epsilon_\star c_{\rm s}/c$ when $p_{\rm rad} \gg p_{\rm gas}$ and $\tau \gg 1$. For $\epsilon_\star \ll 1$, $t_{\rm dyn} \ll t_{\rm SF}$. In the optically thin limit, $t_{\rm th}/t_{\rm SF} = 2c_{\rm s}^2/(\epsilon_\star \alpha c^2)$, dividing this ratio by $t_{\rm th}/t_{\rm dyn}$ we find $t_{\rm dyn}/t_{\rm SF} \simeq (h/r)^2 \ll 1$. Thus, we can express the relationship between time-scales as $t_{\rm th} \ll t_{\rm dyn} \ll t_{\rm SF}$, and expect it to hold for all relevant disc parameters. Because $t_{\rm th} \ll t_{\rm SF}$, $\dot{\Sigma}_\star$ can be assumed constant on time-scales of $t_{\rm th}$. This has implications for thermal disc stability, as noted in Appendix B.2 of TQM05.

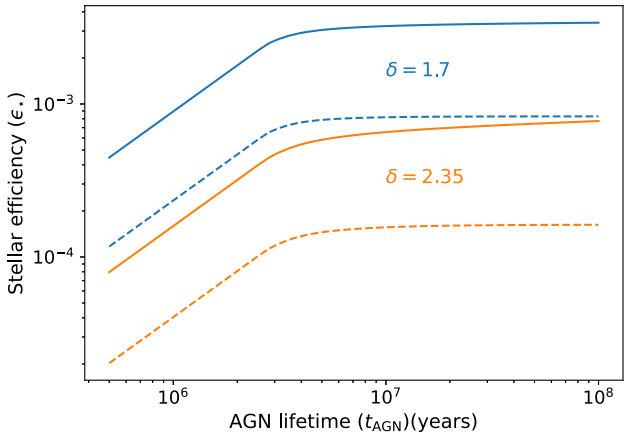

Figure 1. The efficiency of converting stellar mass to radiation calculated for the stellar IMF exponent $\delta_{\rm IMF} = 1.7$ (blue) and $\delta_{\rm IMF} = 2.35$ (orange). The solid lines illustrate the efficiency assuming the disc is optically thick to all wavelengths, while the dashed lines include only wavelengths in or above the UV or $\lambda \geq 100$ nm. Throughout this paper, we assume a top-heavy IMF ($\delta = 1.7$), an AGN lifetime of 10^8 yr, and a UV-limited stellar radiative efficiency of $\epsilon_\star = 8.3 \times 10^{-4}$ as defined by the dashed blue line.

and stellar lifetime of $t_{\rm star} = 10$ Gyr $(m_\star/{\rm M}_\odot)(L_\star/{\rm L}_\odot)^{-1}$. Here, we calculate the efficiency for the two adopted stellar IMFs scaling by stellar lifetime and an efficiency assuming UV radiation alone supports the AGN disc. This is likely to be the case in the outer disc where gas-coupled dust grains make the disc optically thin to re-radiated IR radiation but optically thick to UV radiation. Here, we compute the UV-weighted case and compare with the non-UV weighted efficiency assumed in previous works. Efficiency is calculated according to equation A4 from Tagawa et al. (2020),

$$\epsilon_\star = \frac{\int_{m_{\rm min}}^{m_{\rm max}} L(m_\star) \min[t_{\rm AGN}, t_\star(m_\star)] m_\star^{-\delta} {\rm d}m_\star}{\int_{m_{\rm min}}^{m_{\rm max}} m_\star c^2 m_\star^{-\delta} {\rm d}m_\star}, \tag{19}$$

assuming the same mass-stellar lifetime relation

$$t_\star = 10^{0.825 \log^2(m_\star/120{\rm M}_\odot)+6.43}\,{\rm yr} \tag{20}$$

as in equation (10). The stellar mass–luminosity relation was taken to be a piecewise function (e.g. equation A2 in Tagawa et al. 2020 and Salaris & Cassisi 2005):

$$L(m_\star) = \begin{cases} 0.27\,{\rm L}_\odot (m_\star/{\rm M}_\odot)^{2.6} & m_\star < 0.5{\rm M}_\odot \\ {\rm L}_\odot (m_\star/{\rm M}_\odot)^{4.5} & 0.5\,{\rm M}_\odot < m_\star < 2\,{\rm M}_\odot \\ 1.9\,{\rm L}_\odot (m_\star/{\rm M}_\odot)^{3.6} & 2\,{\rm M}_\odot < m_\star < 42\,{\rm M}_\odot \\ 32000\,{\rm L}_\odot (m_\star/{\rm M}_\odot) & 42\,{\rm M}_\odot < m_\star < 120\,{\rm M}_\odot \end{cases}. \tag{21}$$

Together, equations (19) and (21) allow us to calculate the non-UV weighted stellar efficiency as a function of AGN lifetime, illustrated by the solid lines in Fig. 1. In order to determine the UV-weighted efficiency, we assumed a stellar radius–mass relation $R_\star \propto m_\star^{1/2}$ and effective temperature scaling $T_{\rm eff} \propto (L_\star/R_\star^2)^{1/4}$. Assuming a black-body spectrum for stars, we then calculated the stellar UV luminosity as

$$L_{\star_{\rm UV}}(m_\star) = \int_0^{100{\rm nm}} \frac{2h_{\rm p}c^2/\lambda^5}{\exp\left[\frac{h_{\rm p}c}{\lambda k_{\rm b} T_{\rm eff}(m_\star)}\right] - 1} {\rm d}\lambda, \tag{22}$$

where $h_{\rm p}$ is the Plank constant, and our upper bound of 100 nm limits the computed luminosity to the UV range and above. Substituting equation (22) into (19), we calculated the UV-weighted stellar efficiency as a function of AGN lifetime, described by the dashed curves in Fig. 1. For a Salpeter IMF, the UV and non-UV weighted efficiencies are $\epsilon_\star = 1.6 \times 10^{-4}$ and 7.7×10^{-4}, respectively, while for a top heavy IMF they are 8.3×10^{-4} (UV-weighted) and 3.4×10^{-3} (non-UV weighted). The discrepancy between the non-UV weighted efficiencies calculated here and those calculated in Tagawa et al. (2020) come from the difference in approach to stellar lifetime. For the remainder of this paper, we assume a top-heavy IMF and a UV-weighted efficiency of $\epsilon_\star = 8.3 \times 10^{-4}$.

4 SOURCES OF FEEDBACK

It is illustrative to compare the anticipated values of luminosity and accretion rates for a single generation of star formation and a remnant sBH population. We follow the procedure laid out in equations (9) and (11), substituting $\dot{\Sigma}_\star \delta t$ into our equation for $A_{\rm x}$. We can then solve for the number of black holes per unit disc area, assuming the boundary of the integral taken at the transition mass between sBHs and NSs ($m_{\rm trans} \simeq 20$ M$_\odot$) to include all sBHs produced from the mass of stars $\dot{\Sigma}_\star \delta t$. Solving for the relative stellar and sBH heating components, we find

$$\frac{Q_{\rm sBH}}{Q_\star} = \frac{S_{\rm sBH} \frac{4\pi G \mu_{\rm e} m_{\rm p} m_{\rm sBH} c}{\sigma_{\rm T}}}{\dot{\Sigma}_\star \frac{\epsilon_\star c^2}{2}} \simeq 1.9 \left(\frac{m_{\rm sBH}}{10\,{\rm M}_\odot}\right) \left(\frac{\delta t}{3\,{\rm Myr}}\right) \left(\frac{\epsilon_\star}{8.3 \times 10^{-4}}\right)^{-1}, \tag{23}$$

where we have assumed a top-heavy IMF of $\delta = 1.7$. This suggests that heating from sBHs should overtake stars within a single generation or approximately the time it takes for the most massive stars to evolve off the MS ($\simeq 3$ Myr).

Under the assumption that the sBH luminosity will quickly surpass stellar luminosity in AGN, Gilbaum & Stone (2022) did not include radiation pressure or disc mass depletion due to star formation in their models. However, in comparing the mass accretion rates of these two populations, we found

$$\frac{\dot{\Sigma}_{\rm sBH}}{\dot{\Sigma}_\star} = \frac{\epsilon_\star}{\eta} \frac{Q_{\rm sBH}}{Q_\star}, \tag{24}$$

indicating that $\dot{\Sigma}_{\rm sBH} \ll \dot{\Sigma}_\star$ for expected sBH radiative efficiencies between $0.1 \leq \eta \leq 1$. The significant disparity in the mass accretion rates between stars and sBHs suggests that as radiation pressure shifts from being star formation to sBH dominated, the mass flow through the disc should increase, since less mass is removed by sBH accretion than by star formation, to achieve the same heating rate. The luminosity from the sBHs in the disc interior is set by the preceding generation of star formation. If this sBH luminosity cannot sustain disc stability in the face of increased mass flux, additional star formation will be necessary to stabilize the disc interior.

Note that in our discussion of feedback mechanisms we have neglected feedback from NSs and white dwarfs (WDs). We discount these sources on the basis that NS masses and therefore Eddington-limited accretion rates are expected to be an order of magnitude lower than their sBH counterparts. WD feedback is also ignored under the assumption that a significant population would not have time to form within a typical AGN lifetime.

4.1 X-ray opacity

In TQM05, it is argued that UV radiation from massive stars is absorbed and scattered by dust grains, which reprocess the UV emission into the IR. As applied in TQM05, τ refers to the optical

depth of the disc to IR photons and not to the stellar UV emission to which the disc is always optically thick. Moreover, the temperature profiles presented in TQM05 do not exceed 10^4 K in the star-forming, outer disc. At such low temperatures, we expect the disc gas to be neutral, interacting with the emitted UV radiation via photoionization. This set-up is not unlike models for feedback in the inter-stellar medium (ISM; e.g. McKee & Ostriker 1977), in which stellar radiation creates hot ionized bubbles in a cold inter-bubble medium. These bubbles emit radiation from recombination in the IR. In the AGN disc context, radiative diffusion ensures that the momentum from these re-radiated IR photons is efficiently coupled to the gas and uniformly heats the disc.

We expect the X-ray flux emitted from accreting sBHs to interact with dust and gas in the disc in a way similar to their parent stars. That is, dust should scatter and absorb X-rays while neutral gas is photoionized. Additionally, high-energy X-rays (>1 keV) undergo Compton scattering by the electrons bound in neutral hydrogen and helium atoms (Sunyaev & Churazov 1996). Still, because the mean free path of X-ray photons is large relative to UV photons, we cannot assume the disc is always opaque to X-rays. In this section, we calculate the optical depth of X-rays in the disc and determine $f_{\rm esc}$, the fraction of sBH emission we anticipate escaping the disc.

We begin by calculating an effective cross-section for absorption and scattering of X-ray photons according to (Rybicki & Lightman 1986),

$$\sigma_{\rm eff} = \sigma_{\rm abs}\sqrt{1+\frac{\sigma_{\rm scat}}{\sigma_{\rm abs}}}\,. \tag{25}$$

Here, $\sigma_{\rm abs}$ and $\sigma_{\rm scat}$ are the cross-sections for X-ray absorption and scattering, respectively. Included in $\sigma_{\rm abs}$ is dust absorption and photoionization of hydrogen and helium,

$$\sigma_{\rm abs} = X\left(\sigma_{\rm HI}+\sigma_{\rm DA}\right)+\frac{Y}{4}\sigma_{\rm HeI}, \tag{26}$$

where $\sigma_{\rm DA}$ is the dust absorption taken from Draine (2003) and $X = 0.75$ and $Y = 0.25$ are the hydrogen and helium mass fractions, respectively, assuming a primordial gas composition. The equation for the photoionization cross-section of hydrogen $\sigma_{\rm H\,I}$ is adopted from equation 2.4 of Osterbrock & Ferland (2006),

$$\sigma_{\rm HI} = A_0\left(\frac{E_1}{E}\right)^4\frac{\exp\left(4-4\tan^{-1}\left(\mathcal{E}\right)/\mathcal{E}\right)}{1-\exp\left(-2\pi\mathcal{E}\right)}, \tag{27}$$

where $A_0 = 6.30\times10^{-18}$ cm^2, $\mathcal{E}\equiv(E/E_1-1)^{1/2}$, and $E_1 = h_P\nu_1 = 13.6$ eV is the hydrogen ionization threshold energy, and h_P is Planck's constant. From Haardt & Madau (1996), we calculate the cross-section for singly ionized helium as

$$\sigma_{\rm HeI} \simeq \frac{0.694\times10^{-18}}{(E/10^2\ {\rm eV})^{1.82}+(E/10^2\ {\rm eV})^{3.23}}\ {\rm cm}^2\,. \tag{28}$$

The scattering cross-section $\sigma_{\rm scat}$ is taken to be the sum of dust and Compton scattering from neutral atoms. The dust scattering cross-section per hydrogen atom $\sigma_{\rm DS}$ is taken from Draine (2003) for Milky-Way dust. The Compton cross-section for neutral hydrogen $\sigma_{\rm C}$ is set equal to the Thompson cross-section ($\sigma_T = 6.65\times10^{-25}$ cm^2) for photon energy $E > 1$ keV, and zero everywhere else. Following Sunyaev & Churazov (1996), we multiply $\sigma_{\rm C}$ by 1.5 to account for additional scattering by helium and other metals, allowing us to write the scattering cross-section as

$$\sigma_{\rm scat} = X\left(1.5\sigma_{\rm C}+\sigma_{\rm DS}\right)\,. \tag{29}$$

The effective cross-section $\sigma_{\rm eff}$ and component cross-sections are plotted in Fig. 2.

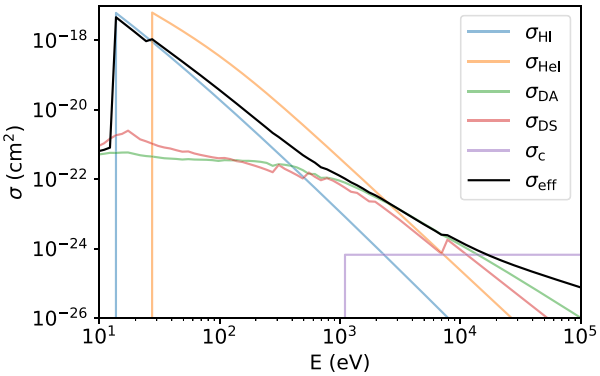

Figure 2. Cross-sections as a function of photon energy for photoionization of hydrogen (blue) $\sigma_{\rm H\,I}$ and singly ionized helium (orange) $\sigma_{\rm He\,I}$. The cross-section for dust absorption $\sigma_{\rm DA}$ and scattering $\sigma_{\rm DS}$ are plotted in green and red, respectively, while in purple we display the Compton cross-section for neutral hydrogen, $\sigma_{\rm C} = \sigma_{\rm T}$ when $E > 1$ keV. The effective cross-section $\sigma_{\rm eff}$ (black) is defined in equation (25) in the text. This parameter is used to determine the disc's optical depth to X-rays and the relative fraction of sBH X-ray emission we expect to escape the disc.

An X-ray photon will escape the gas disc if the effective optical depth of the disc to X-rays is $\tau_{\rm X} < 1$. Writing $\tau_{\rm X}$ as the ratio of the half-thickness h and the X-ray mean free path $\Lambda_{\rm X}$, the fraction of radiation from a single sBH to escape the disc is calculated as

$$f_{\rm esc} = \frac{\int_{\nu_1}^{\infty} L_{{\rm sBH}_\nu}\exp\left(-\frac{h}{\Lambda_{\rm X}}\right)d\nu}{\int_{\nu_1}^{\infty} L_{{\rm sBH}_\nu}d\nu}\,. \tag{30}$$

Here, the mean free path is defined in terms of the effective cross-section and disc density as $\Lambda_{\rm X} = (\rho/m_{\rm p}\sigma_{\rm eff})^{-1}$ and emission from an accreting sBH ($L_{{\rm sBH}_\nu}$) is modelled as a Shakura–Sunyaev accretion disc, with a mass flux $\dot{m}_{\rm sBH}$, inner radius equal to the ISCO ($r_{\rm ISCO} = 3R_{\rm s}$), and outer radius of $r_{\rm h}$.

The top panel of Fig. 3 shows $f_{\rm esc}$ as a function of Σ for a 10 $M_\odot$ sBH embedded in the disc at a radius of 1 pc (solid black line), 0.3 pc (dot–dashed line), and 0.1 pc (dotted line). We assume a marginally stable disc with volumetric density $\rho = \Omega^2/(\sqrt{2}\pi G)$, such that $\rho = 1.0\times10^{-16}$, 2.7×10^{-15}, and 6.5×10^{-14} g cm^{-3} at $r = 1$, 0.3, and 0.1 pc, respectively.

When sBH accretion is Eddington capped, $f_{\rm esc}$ is independent of the location of the embedded sBH in the AGN disc and follows the solid $r = 1$ pc curve. In the top panel of Fig. 3, red, blue, and green circles plotted along the solid line correspond to the same colour sample spectra in the middle panel, whose shaded regions illustrate the radiant energy that escapes the disc. Note that the Eddington-capped intrinsic sBH spectra – shown as solid curves in the middle panel of Fig. 3 – are largely independent of disc parameters. Small differences between spectra at the lowest energies arise because the sBH accretion discs are truncated at $r_{\rm h} = h$. In these cases, the dependence of $f_{\rm esc}$ on the local disc parameters is limited to the exponential term in equation (30), which we can re-write in terms of Σ explicitly: $f_{\rm esc}\propto\exp(-\Sigma/m_{\rm p}\sigma_{\rm eff})$.

Closer to the SMBH, at $r = 0.1$ pc (0.3 pc), $f_{\rm esc}$ diverges from the solid curve and flattens for $\Sigma\lesssim1$ g cm^{-2} (0.3 g cm^{-2}) dropping to an approximately constant value of $f_{\rm esc}\simeq0.1$ (0.3) as indicated by the dotted (dash–dotted) line in the top panel of Fig. 3. Red, blue, and green crosses on the $r = 0.1$ pc curve correspond to the same colour spectra in the bottom panel of Fig. 3. The escape fraction drops between $r = 1$ and 0.1 pc for $\Sigma\lesssim1$ g cm^{-2} because $\dot{m}_{\rm sBH}$

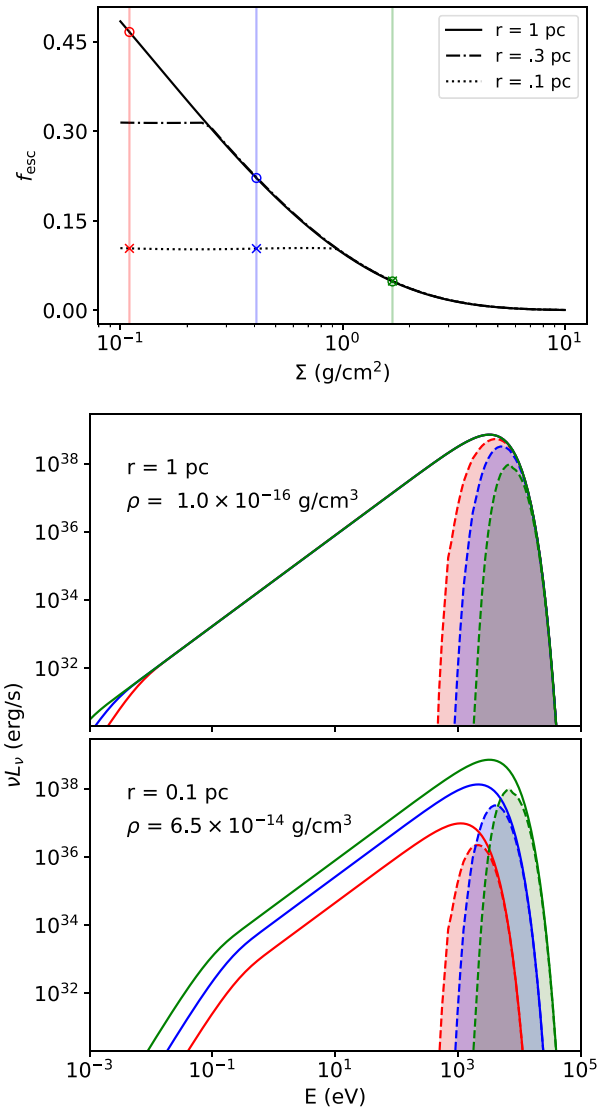

Figure 3. Top panel: The escape fraction of ionizing radiation as a function of Σ, for a 10 $M_\odot$ sBH embedded at a distance of $r = 1$ (solid black line), $r = 0.3$ pc (dot–dashed line) and $r = 0.1$ pc (dotted line) from the central SMBH. We adopt the volumetric density of a marginally stable disc, i.e. $\rho = \Omega^2/(\sqrt{2}\pi G)$. A red, blue, and green circles with coordinates $\Sigma(\mathrm{g/cm^2}) = 0.11, 0.41$, and 1.68 and $f_{esc} = 0.47, 0.22$, and 0.05, respectively, are plotted along the $r = 1$ pc curve. Blue, red, and green crosses mark the same Σ values along the $r = 0.1$ pc curve, with $f_{esc} = 0.10, 0.10$, and 0.05. The circles and crosses correspond to the spectra in the same colours in the middle and bottom panels, respectively. Note that at 1 pc, sBH accretion is Eddington-limited across plotted values of Σ. In the Eddington-limited regime, f_{esc} scales inversely with Σ, otherwise f_{esc} is approximately constant. Middle panel: Three example spectra for a 10 $M_\odot$ sBH embedded at $r = 1$ pc. Solid lines indicate the intrinsic (unabsorbed) sBH emission spectra. Filled in regions illustrate the radiation that escapes from the disc. *Bottom panel*: Same as the middle panel but for $r = 0.1$ pc.

is not Eddington capped; instead, $\dot{m}_{sBH} = \dot{m}_B$, determined by $r_w = R_{Hill}$ and $r_h = h$, with $\dot{m}_{sBH} \propto \Sigma^2$ (equation 14). The decrease in $\dot{m}_{sBH}$ lowers both the energy and amplitude of the intrinsic sBH spectra, as seen in the red and blue solid curves plotted in the bottom panel. Despite the reduced sBH emission, f_{esc} remains approximately constant for $\Sigma \lesssim 1\ \mathrm{g\,cm^{-2}}$ because $\tau_X \propto \Sigma$. For $\Sigma = 1.68\ \mathrm{g\,cm^{-2}}$ (green), $f_{esc} = 0.05$ at $r = 0.1$ and 1 pc, as $\dot{m}_{sBH} = \dot{M}_{Edd}$ in both cases, resulting in identical spectra above 1 eV.

Note that in a marginally stable disc the volumetric density scales with radius as $\rho \propto r^{-3}$, and the scale height goes as $h \propto (\dot{M}(\rho\Omega)^{-1})^{1/3} = \dot{M}^{1/3} r^{3/2}$. The column density is therefore proportional to $\Sigma \propto \dot{M}^{1/3} r^{-3/2}$ and we expect Σ to scale inversely with r. We verify this general trend in our steady-state models, discussed in Section 6, plotting f_{esc} as a function of distance from the SMBH in the top panel of Fig. 9. Note the second row of Fig. 5, where we plot Σ. We find that where the disc is marginally stable, and changes in $\dot{M}$ are small, our scaling relation holds. This plot also shows that in general $\Sigma \sim 10\ (\mathrm{g\,cm^{-3}})$ at $r \sim 0.1$ pc, while $\Sigma \sim 0.1\ (\mathrm{g\,cm^{-3}})$ at $r \sim 1$ pc. We therefore expect sBH accretion to be Eddington-limited across these models.

Finally, in order to incorporate the reduction in the heating rate due to some of the radiation escaping from the disc, we amend equation (17) to

$$Q_{sBH} = S_{sBH} \dot{\Sigma}_{sBH} \eta c^2 (1 - f_{esc}). \tag{31}$$

Note that the loss of sBH radiation is equivalent to a decrease in the effective radiative efficiency as it relates to disc heating, i.e. $(1 - f_{esc})$ is degenerate with η. The radiation loss decreases the effective heating rate of sBHs relative to the stellar disc component such that the ratio $Q_{sBH}/Q_\star$ (calculated in equation 23) decreases and the time required until sBH radiation dominates stellar radiation increases.

5 NUMERICAL IMPLEMENTATION

The equations outlined in Section 2 represent a set of differential equations that may in principle be solved to give the full time evolution of the disc, under the assumption that this evolution is well approximated as a sequence of steady-state solutions (see discussion of this assumption below). We consider two models: a single-component disc supported either by stars or by sBHs only (to build intuition), and then an evolving model in which a growing population of sBHs seeded by the progenitor stellar population changes the steady-state structure of the disc over time.

Our procedure for solving the sBH-only disc model is similar to the one laid out in TQM05 and more recently developed in Gangardt et al. (2024). Beginning at the outer disc boundary with $r = R_{out}$ and $\dot{M} = \dot{M}_{out}$, and assuming a marginally stable disc (i.e. $\rho \propto \Omega^2$), we can solve for τ and T_{eff} as a function of T. The total pressure is uniquely determined by ρ and $\dot{M}$, while the right-hand side of equation (3) yields $p_{tot} = p_{tot}(T)$, allowing us to solve for T. Note that at some radii multiple disc solutions exist. This was dealt with in TQM05 by assuming the lowest T solution to be the most physically motivated. Near R_{out}, we follow TQM05, but at interior annuli, we choose the solution with opacity κ closest to that of the adjacent exterior radius. Our motivation for this choice is discussed in more detail in Section 6.3.

By assuming the disc is in LTE, the disc flux $\sigma_b T_{eff}^4(T)$ can be used to calculate the total auxiliary heating necessary to support the disc via equation (7), where we set $\dot{\Sigma}_\star = 0$ and solve for Q_{sBH}. Equation (12) gives $\dot{m}_{sBH} = \dot{m}_{sBH}(\rho, H)$ from which we can calculate the total emitted radiation from an sBH. Using equation (30), we determine f_{esc} and solve equation (31) for S_{sBH} or the total number of sBHs per unit area. The final step is calculating the mass accretion from the sBHs as $\dot{\Sigma}_{sBH} = S_{sBH} \dot{m}_{sBH}$, multiplying by the area of the outermost radial annulus, and subtracting from $\dot{M}_{out}$ as in equation (4), to obtain the mass flux at the adjacent interior radius where we apply the same algorithm.

This procedure is applied at each radial annulus moving inwards in the disc, until solutions maintaining a marginally stable disc require non-physical model parameters, namely $Q_{\rm sBH} < 0$. At this point, viscous heating alone provides sufficient support and we revert to a standard Shakura–Sunyaev disc model. Note that for the Shakura–Sunyaev disc, $T_{\rm eff}$ is given and ρ is not. To solve for ρ and T in this case, we balance equations (8) and (6) using SCIPY's (Virtanen et al. 2020) optimization routine 'root.' Values for ρ and T from the adjacent exterior radii are used as input guesses. We use 500 equally spaced logarithmic radial bins to model the discs. At this resolution, we found that these steady-state disc models converge.

Our second model assumes the AGN disc's evolution over time can be treated as a sequence of steady states that change in response to the accumulating population of remnant sBHs. We begin at $t = 0$ with a TQM05 disc, fully supported by star formation. We construct this model in much the same was as the sBH only model, but solve for $\dot{\Sigma}_\star$ instead of $Q_{\rm sBH}$. Given $\dot{\Sigma}_\star$ and a time-step Δt and assuming $\dot{\Sigma}_\star$ is constant over Δt, we solve for the number of sBHs in each radial annulus according to equations (9)–(11).

Having populated the disc with remnants, we solve for two solutions sets: (1) a marginally stable ($Q_{\rm T} = 1$) disc solution supported by both star formation and radiative emission from embedded sBHs and (2) a stable disc supported only by sBHs ($Q_{\rm T} \geq 1$ and $\dot{\Sigma}_\star = 0$). If the former requires negative pressure support from stars ($\dot{\Sigma}_\star < 0$) or the latter requires a volume density exceeding $\Omega^2/(\sqrt{(}2)\pi G$ the solutions are presumed to be not physical and the alternate solution set is chosen. In some cases, both sets of solutions yield physical disc solutions. This can occur where degenerate, marginally stable starburst solutions exist or where stable solutions require unreasonably low Σ such that the sBHs can support the disc at sub-Eddington accretion rates. Where both solutions exist, we choose the stable (sBH-only) solution set when the number density of sBHs exceeds the minimum number required to support the disc as found in our initial, sBH-only models. This minimum sBH number is referred to as $S_{\rm sBH_{min}}$ in the schematic shown in Fig. 4.

Once solutions are found across all radii, we jump to the next time-step and calculate the distribution of sBHs. Note that after the first time-step we must discretize equation (11) according to

$$S_{\rm sBH}(t_n) = \sum_{i=0}^{n} A_{\rm x}(t_i) \int_{m_{\rm TO}(t_n - t_i)}^{m_{\rm max}} m_\star^{-\delta} {\rm d}m_\star \tag{32}$$

where

$$A_{\rm x}(t_i) = \frac{\dot{\Sigma}_\star(t_i)\Delta t_i}{\int_{m_{\rm min}}^{m_{\rm max}} m_\star^{1-\delta} {\rm d}m_\star}\,. \tag{33}$$

This procedure is laid out explicitly in Fig. 4. Note that in this figure we distinguish between marginally stable solutions including both star formation and sBH accretion and those supported by sBHs only with subscripts 1 and 2, respectively.

In our models, we take the first time-step to be 2.5 Myr, approximately the delay between the start of star formation and the most massive stars evolving off the MS. The second and subsequent time-steps are set to $\Delta t = 1$ Myr, although we find that time resolution has limited effect on our models so long as $\Delta t \leq 2.5$ Myr. We also enhance our radial resolution in these models, increasing Δr in regions where enhanced star formation leads to steep drops in mass flux. We discuss these resolutions tests in more detail in Section 7.2.

We opt to simplify our models significantly by assuming all sBHs have a mass of 10 $M_\odot$. The initial *versus* final mass relation (IFMR), linking remnant sBHs to their respective progenitor stars, can vary significantly depending on assumptions made about stellar winds, supernovae, and chosen metallicity (Spera, Mapelli & Bressan 2015; Spera & Mapelli 2017; Raithel, Sukhbold & Özel 2018). Assuming solar metallicity, Raithel et al. (2018) predicts maximum sBH masses of 16 $M_\odot$, while the models of Spera et al. (2015) suggest sBH masses can reach 25 $M_\odot$ for high-metallicity systems ($Z = 2 \times 10^{-2}$). Despite discrepancies, we expect $m_{\rm sBH} \sim 10 M_\odot$ to be a reasonable estimate for our purposes.[4]

Although we incorporate the feedback from sBHs in our disc models, we do not account for the accompanying sBH growth. If we were to account for sBH growth, the number ($S_{\rm sBH}$), accretion rate ($\dot{m}_{\rm sbh}$), and escape fraction ($f_{\rm esc}$) would be evaluated independently for sBHs of differing mass and equations (16) and (31) would need to re-written as a sum of products. This added a layer of complication that slowed our computation but would not significantly enhance our results. For sBH accretion rate set to the maximum, Eddington rate, an AGN lifetime of $t_{\rm AGN} = 10^8$ yr represents ~ 2.5 e-folding times, or about an order of magnitude increase in mass (and accretion rate). While this enhances the importance of sBHs, we do not expect this to change our qualitative results, a point we discuss in more detail in Section 6.5.

6 STEADY-STATE DISCS: SINGLE SOURCE MODELS

Here, we analyse steady-state AGN discs supported entirely by accretion onto sBHs. As noted in Section 4, we do not anticipate that sBH heating will quickly overcome stellar heating across the disc, nor do we expect the sBH population to grow to precisely the size necessary to maintain a marginally stable disc. However, by investigating the 'sBH-only' case we can clarify expected disc profiles given a minimum number of sBHs, comparing directly with the TQM05 starburst-only models with equivalent boundary conditions.

Our steady-state models are also analogous to the 'pile-up' solutions found by Gilbaum & Stone (2022). Note that the 'pile-up' solution assumes that the sBHs do not migrate and that mass flux is constant across the disc. The latter constraint is not assumed in the models presented here (we account for the reduction in the inflow rate as gas is consumed by BH accretion). Additionally, Gilbaum & Stone (2022) take the disc to be largely ionized such that electron scattering is the dominant source of X-ray opacity. Our models assume a neutral disc, opaque to X-ray flux emitted by sBHs when $\tau_{\rm X} > 1$.

As described in Section 4.1, $\tau_{\rm X}$ includes contributions from scattering and absorption by dust, neutral atoms, and photoionization. We refer to disc models including all of these X-ray/gas interactions as the 'sBH-only' models. Here, we also present an alternative prescription for optical depth that includes only the effects of photoionization of hydrogen and singly ionized helium. These models are intended to address the possibility that the embedded BHs launch jets and evacuate the surrounding gas, creating a low-density 'chimney' out of the disc (Tagawa et al. 2022). As a result, the effective Σ encountered by the BH radiation may be lower than that of the average background disc. Below $\Sigma \sim 10^{-2}$ g cm^{-2}, photoionization is the dominant contributor to $\sigma_{\rm eff}$ and other processes can be ignored.

[4]Although we expect our results to be robust against small changes in $m_{\rm sBH}$, it is useful to note that increasing the mass would have the dual effect of increasing the accretion rate and softening the spectrum of the individual sBH. The latter would lead to a slight decrease in $f_{\rm esc}$, slightly decreasing the effective radiative efficiency of embedded sBHs.

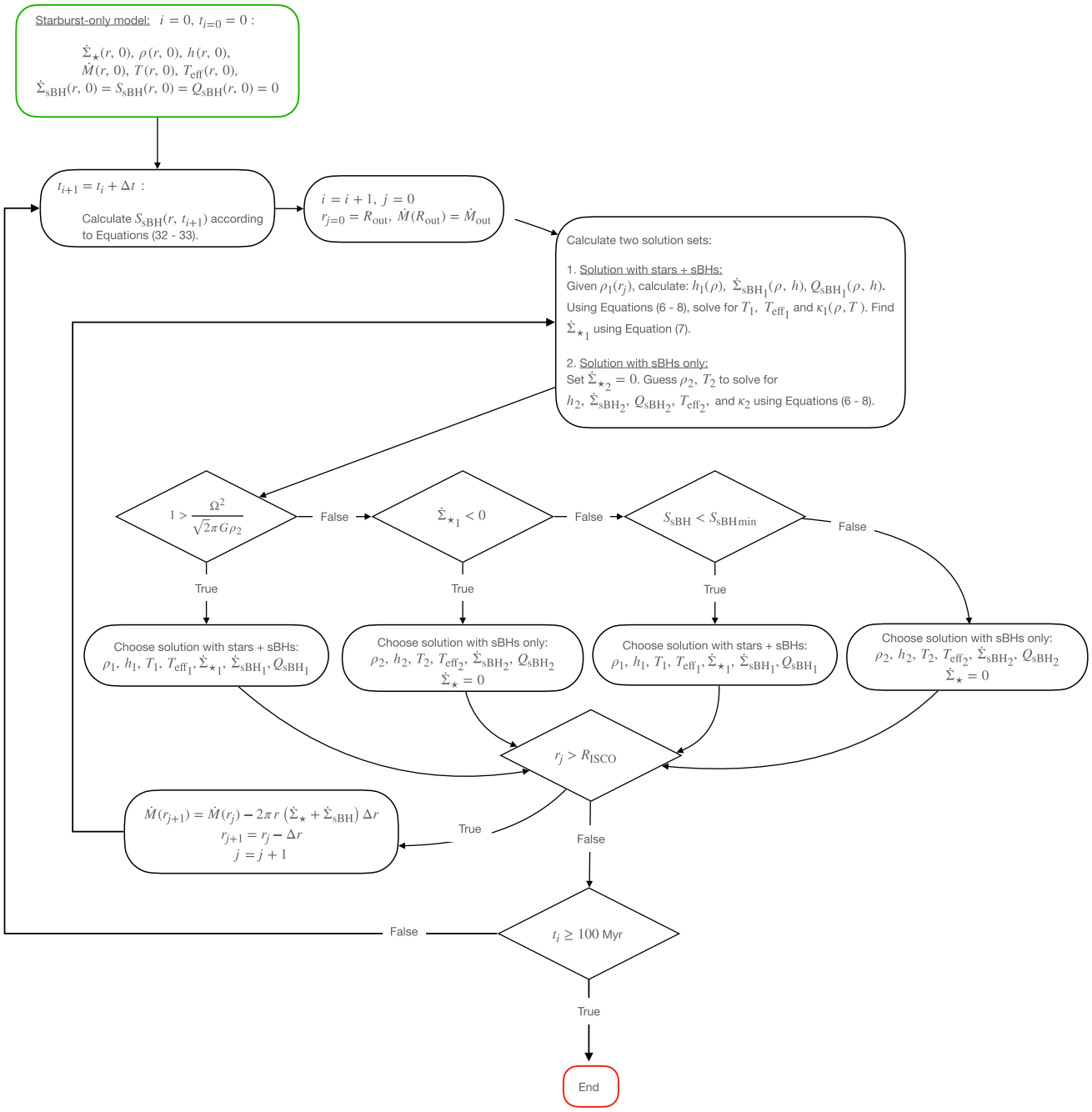

Figure 4. Schematic detailing our solution strategy in our disc evolution models. Beginning with a starburst only model at $t = 0$, we calculate the distribution of sBHs across the disc according to the preceding star formation rates. We then choose between marginally stable solutions including pressure support from star formation + sBH accretion and disc solutions supported by sBHs only (differentiated by subscript 1 or 2 in the schematic). The chosen solution gives ρ, h, T, $T_{\rm eff}$, $\dot{\Sigma}_\star$, and $\dot{Q}_{\rm sBH}$ for each radius starting from $R_{\rm out}$ and moving inwards. This procedure is repeated at each time-step until $t = 100$ Myr.

We refer to discs in which we used this amended $\tau_{\rm X}$ as 'sBH-only-chimney' models and note that they have a larger $f_{\rm esc}$ than the 'sBH-only' models.

In the following subsections, we describe these two classes of models, beginning with a description of relevant disc parameters as a function of distance from the central SMBH in Section 6.1. We justify our assumption of a neutral gas disc in Section 6.2 and address degenerate steady-state solutions in Section 6.3. In Section 6.4, we calculate the anticipated spectra for our steady-state models, and in Section 6.5, we estimate the time-scale over which an AGN disc is fully supported by sBHs.

6.1 Star- versus sBH-driven disc structures

In Fig. 5, we show the radial profiles of different disc variables for starburst-only and sBH-only. Starburst-only profiles are on the leftmost (first and third) columns, while sBH-only profiles are in the rightmost (second and fourth) columns.

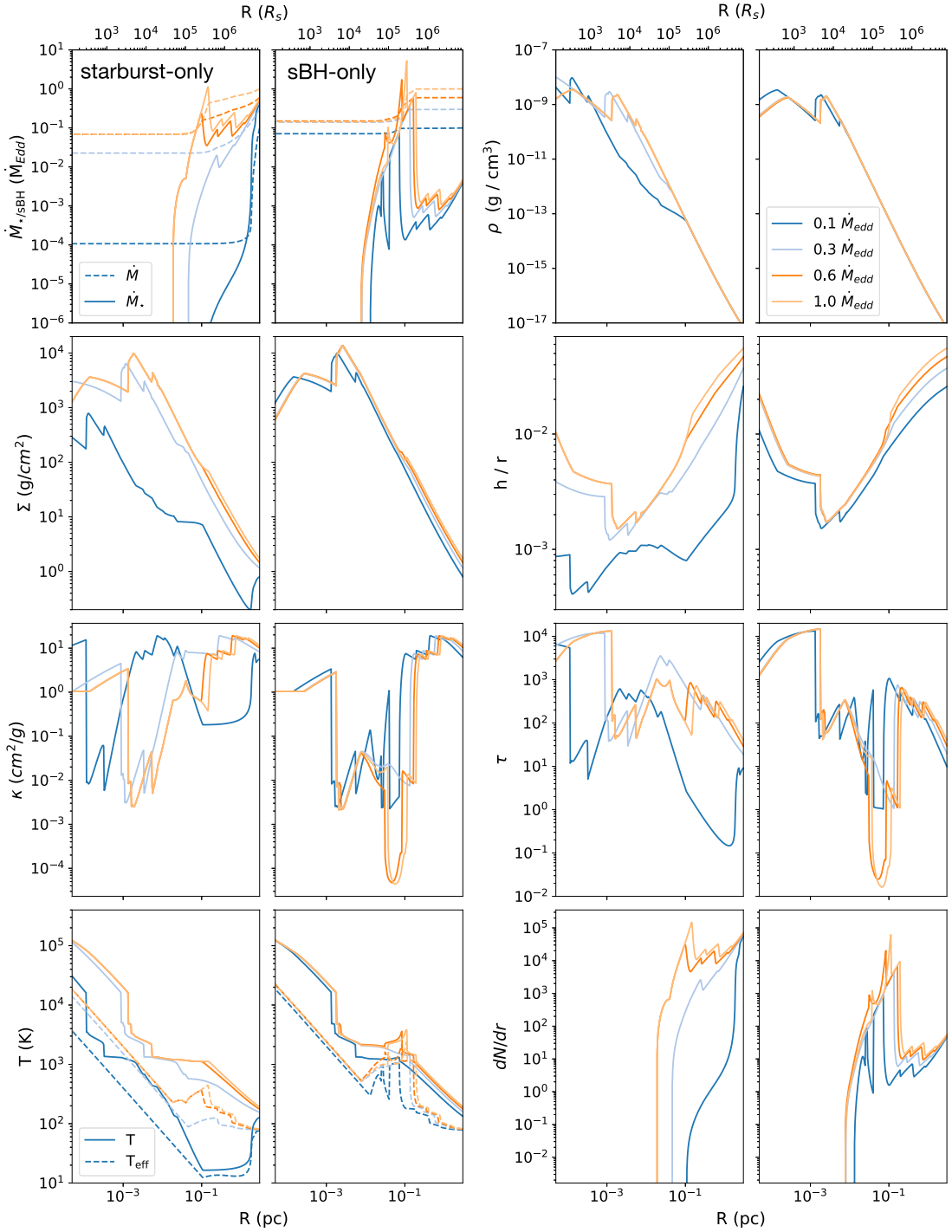

Figure 5. Disc parameters as a function of radius for models with $M_{\rm SMBH} = 4 \times 10^6 {\rm M}_\odot$, $\sigma = 75$ km s^{-1}, and $R_{\rm out} = 3$ pc. The outer boundary conditions $\dot{M}_{\rm out} = 0.1$, 0.3, 0.6, and 1 $\dot{\rm M}_{\rm Edd}$ are differentiated by line colour and correspond to dark blue, light blue, dark orange, and light orange, respectively. The leftmost columns in each panel (i.e. the first and third column from left to right) show starburst-only solutions and the rightmost (second and fourth) columns show sBH-only disc solutions. From top to bottom, the panels on the left show mass flux through the disc ($\dot{M}$) and mass consumption by the auxiliary heat source, column density (Σ), opacity (κ), and temperature (solid) and effective temperature (dashed). From top to bottom, the right panels show volume density (ρ), dimensionless scale height (h/r), optical depth (τ), and the number of stars/sBHs per radial annulus. The total number and mass of stars/sBHs is further documented in Table 1.

The mass-feeding rates span $0.1\,\dot{M}_{\rm Edd}$ (dark blue) to $1\,\dot{M}_{\rm Edd}$ (light orange). This range encompasses the critical value $\dot{M}_{\rm c} = 0.2\,\dot{M}_{\rm Edd}$ defined in equation (18), below and above which star-supported disc models exhibit fundamentally different behaviour. By examining mass fluxes on either side of this bifurcation, we can understand how the behaviour changes when discs with the same parameters are instead supported entirely by sBHs.

Table 1. The total number and mass in stars or sBHs accumulated over 10^8 yr in the steady-state models illustrated in Fig. 5.

	$\dot{M}_{\rm out}(\dot{M}_{\rm Edd})$	$N_{\rm tot}$	$M_{\rm tot}$ ($M_\odot$)
Stars only	0.1	4.5×10^5	1.0×10^6
	0.3	1.3×10^6	2.7×10^6
	0.6	2.4×10^6	5.2×10^6
	1	4.2×10^6	9.2×10^6
sBHs only	0.1	1.2×10^4	1.2×10^5
Chimney	0.3	6.0×10^4	6.0×10^5
	0.6	1.7×10^5	1.7×10^6
	1	3.4×10^5	3.4×10^6
sBHs only	0.1	1.2×10^4	1.2×10^5
	0.3	6.4×10^4	6.4×10^5
	0.6	1.8×10^5	1.8×10^6
	1	3.4×10^5	3.4×10^6

The upper left panels in Fig. 5 plot the mass flux through the disc (dashed) and mass consumption rates by stars (solid, left panel) and by sBHs (second and third-to-left panels). In the starburst-only case, there are two distinct disc solutions represented, differentiated by low (0.1 and 0.3 $\dot{M}_{\rm Edd}$) and high (0.6 and 1 $\dot{M}_{\rm Edd}$) $\dot{M}_{\rm out}$. In the smaller $\dot{M}_{\rm out}$ cases, the star formation rate peaks at the outer boundary leading to a steep drop in mass flux. Between 0.1 and 1 pc, $\dot{M}$ levels out, having dropped to 2×10^{-2} and $1 \times 10^{-4} \dot{M}_{\rm Edd}$ in the $\dot{M}_{\rm out} = 0.3 \dot{M}_{\rm Edd}$ and 0.1 $\dot{M}_{\rm Edd}$ cases, respectively. This behaviour is characteristic of discs whose $\dot{M}_{\rm out}$ is approximately equal to or below the critical mass supply rate $\dot{M}_{\rm c}$, defined in equation (18). In contrast, the larger $\dot{M}_{\rm out}$ disc models maintain star formation down to $r = 2 \times 10^{-2}$ pc, at which point star formation peaks and then plummets, and the mass supply rate drops to a constant value of $\simeq 7 \times 10^{-2}\ \dot{M}_{\rm Edd}$.

In the sBH-only models, $\dot{M}_{\rm out}$ exceeds the critical mass supply rate in all of the cases considered, and there is no significant depletion in the outer disc. The change in mass flux across these discs is smaller than the starburst-only solutions in all cases, due to the higher radiative efficiency of sBHs relative to stars. Despite distinct $\tau_{\rm x}$ and $f_{\rm esc}$, the sBH-only and sBH-only-chimney models have very similar radial profiles. This similarity is reflected in the intrinsic disc spectra shown in Fig. 8 and the similar numbers of sBHs necessary to support the discs (Table 1). In both sets of models, the mass flux through the disc drops to $\simeq 0.15\ \dot{M}_{\rm Edd}$ in the three highest $\dot{M}_{\rm out}$ cases, while in the lowest $\dot{M}_{\rm out}$ case the mass flux reaching the SMBH is $\simeq 7 \times 10^{-2}\ \dot{M}_{\rm Edd}$. Because the disc is not significantly depleted in either of the sBH cases, the volume density (ρ), column density (Σ), and scale height (h) all remain comparable to the higher $\dot{M}_{\rm out}$, starburst-only cases.

In the bottom right panels of Fig. 5, we plot the number of stars and sBHs per radial annulus. The number of stars is calculated assuming a constant $\dot{\Sigma}_\star$ over an AGN lifetime of $t_{\rm AGN} = 10^8$ yr. We give the total mass and number of stars and sBHs in Table 1. Note that $M_{\rm tot}$ in the sBH-only and sBH-only-chimney models never exceeds the mass of the SMBH, justifying the continued use of a Keplerian potential. This is not true for the higher $\dot{M}_{\rm out}$, starburst-only cases.

There are also slightly more sBHs in the $\dot{M}_{\rm out} = 0.3$ and 0.6 $\dot{M}_{\rm Edd}$ sBH-only models than in their sBH-only-chimney counterparts. This result is somewhat counter-intuitive given that in the chimney scenario more sBHs are required to produce the same amount of heating. Exterior to the opacity gap ($r \gtrsim 0.1$ pc) the number of sBHs in the chimney models is greater than in the sBH-only models, as expected. The result is a slight increase in mass flux through the opacity gap in the sBH-only models relative to the chimney models. This slight increase in mass flux in the disc interior requires a larger number of sBHs in the sBH-only case to stabilize the disc, thereby explaining the slight enhancement in number in the sBH-only case.

6.2 Heat mixing in a multiphase disc

The homogeneity of the disc fluid depends on the number of sBHs embedded in it relative to the volume of gas the sBHs can be expected to ionize. This latter value can be approximated by the volume within the Strömgren radius,

$$R_{\rm Str} = \left(\frac{\int_{h_{\rm p}\nu_1}^{\infty} \frac{L_{{\rm sBH}_\nu}}{h_{\rm p}\nu} {\rm d}\nu}{\frac{4\pi}{3} n_{\rm H}^2 \alpha_{\rm B}} \right)^{1/3} \simeq 1.9 \times 10^{-4}\ {\rm pc} \left(\frac{\rho}{10^{-16}\ {\rm g\,cm^{-3}}} \right)^{-\frac{2}{3}} . \quad (34)$$

Here, $\alpha_{\rm B} = 2.59 \times 10^{-13}$ is the Case B recombination cross-section for H I (Osterbrock & Ferland 2006), $n_{\rm H} = \rho/(X m_{\rm p})$ is the number density of Hydrogen, and $X = 0.75$ is the mass fraction of Hydrogen for primordial gas. We parametrize $R_{\rm Str}$ in terms of ρ by assuming Eddington-capped accretion.

The Strömgren radius approximates the size of the ionized bubble around a single sBH, assuming the ionization profile is a step function (i.e. the gas is completely ionized interior to $R_{\rm Str}$ and completely neutral beyond). If the mean free path of the average ionizing photon is long, the ionization profile flattens and the boundary of the ionized bubble is blurred (Kramer & Haiman 2008). The ionizing mean free path $\Lambda_{\rm ion}$, is calculated in much the same way as $\Lambda_{\rm X}$ in Section 4.1, but does not include absorption by dust. Dust absorption reduces the ionizing emission but does not contribute to the ionization fraction. That is,

$$\Lambda_{\rm ion} = (n_{\rm H}\sigma_{\rm H\,I} + n_{\rm He}\sigma_{\rm He\,I})^{-1} \times \left[\sqrt{1 + \frac{n_{\rm H}(1.5\sigma_{\rm C} + \sigma_{\rm DS})}{(n_{\rm H}\sigma_{\rm H\,I} + n_{\rm He}\sigma_{\rm He\,I})}} \right]^{-1} . \quad (35)$$

The ionizing photon-weighted, average mean free path is then calculated according to

$$\bar{\Lambda}_{\rm ion} \equiv \frac{\int_{h_{\rm p}\nu_1}^{\infty} \frac{L_{{\rm sBH}_\nu}}{h_{\rm p}\nu} \exp(-\tau_{\rm DA}) \Lambda_{\rm ion} {\rm d}\nu}{\int_{h_{\rm p}\nu_1}^{\infty} \frac{L_{{\rm sBH}_\nu}}{h_{\rm p}\nu} \exp(-\tau_{\rm DA}) {\rm d}\nu} \simeq 9.9 \times 10^{-4}\ {\rm pc} \left(\frac{\rho}{10^{-16}\ {\rm g\,cm^{-3}}} \right)^{-1} , \quad (36)$$

where we have assumed the ionizing flux is attenuated due to dust absorption according to $\tau_{\rm DA} = n_{\rm H}\sigma_{\rm DA} h$, and sBH accretion is Eddington capped. We can assess the importance of this blurring by comparing $\bar{\Lambda}_{\rm ion}$ with the Strömgren radius. Setting $R_{\rm Str} = \bar{\Lambda}_{\rm ion}$ we can solve for $\rho \simeq 1.4 \times 10^{-14}$ g cm^{-3}. For a marginally stable disc, this density translates to a distance of $\simeq 0.17$ pc. We plot equations (34) and (36) as a function of distance in Fig. 6. For $r < 0.17$ pc, $R_{\rm Str} > \bar{\Lambda}_{\rm photo}$ suggesting that the sBHs carve out distinct, ionized bubbles in the neutral disc fluid. Beyond 0.17 pc, we expect the gas around embedded sBHs to have a broadened ionization profile.

Fig. 6 also shows the average distance between sBHs, $\bar{d}_{\rm sBH} = S_{\rm sBH}^{-1/2}$, and the scale height h in all four sBH-only models. We can conclude that ionization bubbles do not typically overlap, because $\bar{\Lambda}_{\rm photo} < \bar{d}_{\rm sBH}$, although these lengths become comparable at large radii. This finding is consistent with our discussion in Section 4.1, in which we found larger escape fractions in the outer disc, where $\bar{\Lambda}_{\rm photo}$ approaches h. Most importantly, $R_{\rm Str} \ll \bar{d}_{\rm sBH}$ across all sBH-only models and radii, indicating that the disc may be treated as neutral

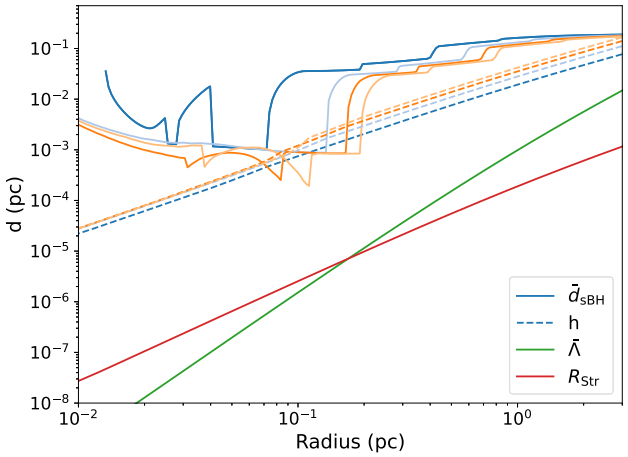

Figure 6. The scale height h (dashed) and average distance between sBHs $\bar{d}_{\rm sBH}$ for sBH-only models. As in Fig. 5, boundary conditions are differentiated by line colours: dark blue, light blue, dark orange, and light orange. The Strömgren radius $R_{\rm Str}$ is plotted in red, and the average mean free path $\bar{\Lambda}_{\rm photo}$ is plotted in green. $R_{\rm Str}$ and $\bar{\Lambda}_{\rm photo}$ never exceed $\bar{d}_{\rm sBH}$, justifying our assumption of a neutral disc.

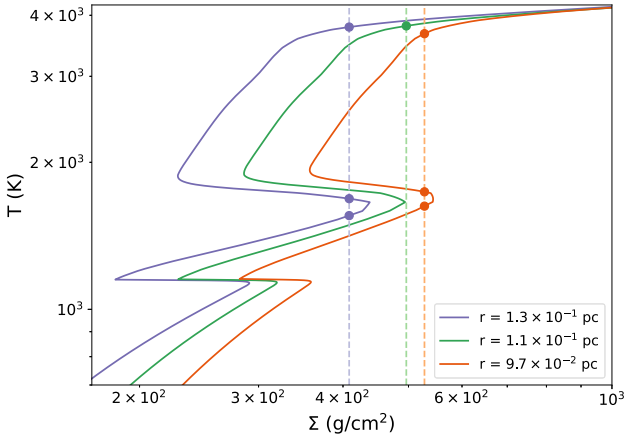

Figure 7. Thermal equilibrium curves in the (Σ, T) plane for the sBH-only models, where $\dot{M}_{\rm out} = 1.0\dot{\rm M}_{\rm Edd}$. Solid curves show equilibria at nearby radii: 1.3×10^{-1} pc (purple),1.1×10^{-1} pc (green), and 9.7×10^{-2} pc (orange). Vertical dashed lines indicate the $\Sigma = \Sigma(r, \dot{M})$ derived from the steady-state model. Finally, dots mark the solution or set of solutions at each radius. Steady-state solutions exhibit discontinuous jumps across varying r, with low temperature solutions ($\sim 1.8 \times 10^3$ K) available at $r = 1.3 \times 10^{-1}$ pc and $r = 9.7 \times 10^{-2}$ pc but not 1.1×10^{-1} pc.

and the effects of electron scattering can be neglected. Moreover, we expect photons produced by recombination will fall primarily in the IR and optical bands, where their mean free path is longer, and they can be assumed to uniformly heat the disc, although the details of these processes need to be addressed in future work.

6.3 Degenerate solutions

In the starburst-only, sBH-only, and sBH-only-chimney disc models, there are ranges of $\dot{M}$ and r across which multiple steady-state solutions exist. Typically, these solutions can be categorized into three branches, a low-temperature (~ 1000 K) 'cold' solution, a high-temperature 'hot' solution ($\sim 2000 - 3000$ K) and an intermediate, thermally unstable solution. This multibranch behaviour is classically associated with outbursts in dwarf novae, manifesting at around 7000 K due to a steep rise in opacity linked to H^- scattering (e.g. Lasota 2001). In contrast, we find degenerate solutions in the cooler, outer regions of AGN – straddling the opacity 'bump' from molecular absorption by water vapour at ~ 2000 K.

The existence of multiple solutions in the starburst-only disc models has been noted across previous works. TQM05 approached the degeneracy by prioritizing the lowest temperature solution in all cases. They argue that both the intermediate- and high-temperature solutions are likely to be unstable and can therefore be discarded, using formal stability arguments in the intermediate solution case and qualitative physical arguments in the high temperature case. In the starburst-only disc models discussed here we follow the TQM05 approach, choosing the 'cold' solutions when degeneracies arise.

Following the same approach in the sBH-only and sBH-only-chimney cases, however, does not always yield physical results. At $\lesssim 0.1$ pc in the two highest $\dot{M}_{\rm out}$ cases, the high-mass flux forces the disc from the degenerate solution regime on to the 'hot,' optically thin branch. Because $T_{\rm eff} \simeq T\tau^{-1/4}$, $T_{\rm eff}$ increases dramatically along with accretion on to sBHs ($\dot{\Sigma}_{\rm sBH}$) that scales with $T_{\rm eff}^4$. The steep decrease in $\dot{M}$ that results leads to reappearance of the low-temperature solution branch at interior radii.

This situation is illustrated in Fig. 7, where three thermal equilibrium curves at $r = 1.3 \times 10^{-1}$ pc (purple), $r = 1.1 \times 10^{-1}$ pc (green), and $r = 9.7 \times 10^{-2}$ pc (orange) are shown in the (Σ, T) plane. The non-trivial shapes of the curves in Fig. 7 follow from the temperature-dependence of the opacity. Points mark the steady-state solution(s) along these curves. Note that as we move from the outermost to the innermost radius, we begin with three possible solutions (1.3×10^{-1} pc), drop to one solution ($r = 1.1 \times 10^{-1}$ pc), and then jump back to three solutions again (9.7×10^{-2} pc). If we were to choose the lowest temperature solution in all cases, our modelled disc solutions would oscillate between the cold and hot solution branches. Physically, we do not expect such abrupt jumps in temperature and opacity. Although not taken into account formally here, discs experience some radial heat transport through radiation and turbulence, very likely smoothing such extreme variations in opacity and temperature. In an effort to avoid this highly oscillatory and unphysical radial profile, we choose the solution whose opacity is closest to that found in the adjacent exterior radius. Thus, in the high $\dot{M}_{\rm out}$ cases, once the jump to the hot branch of the solutions is made, a high temperature and low opacity is maintained in the disc interior.

In the lowest $\dot{M}_{\rm out}$ case the same process leads to jumps between 'cold' and 'hot' solutions, although shifted slightly towards the SMBH – at a distance of $r \lesssim 5 \times 10^{-2}$ pc. When $\dot{M}_{\rm out} = 0.3\ \dot{M}_{\rm Edd}$, the disc profiles are the same regardless of which solution-picking algorithm we use.

6.4 Emerging spectra

AGNs are observed to produce significant flux, spanning nearly nine decades in frequency, from radio to X-ray (Elvis et al. 1994). Emission across such a broad wave-band must arise from distinct regions and physical mechanisms in the disc. Several continuum features are seen consistently across observed AGN spectra, and have been used to constrain sources of emission, including an increase in emission in the optical-ultraviolet (UV) known as the 'UV-bump', flux exceeding classic α-disc models in the radio, IR and X-ray bands, and an inflection point at $\sim$1 eV or ~ 1 μm.

In Fig. 8, we show the spectra of the starburst, sBH-only-chimney, and sBH-only disc models (top, middle, and bottom panels,

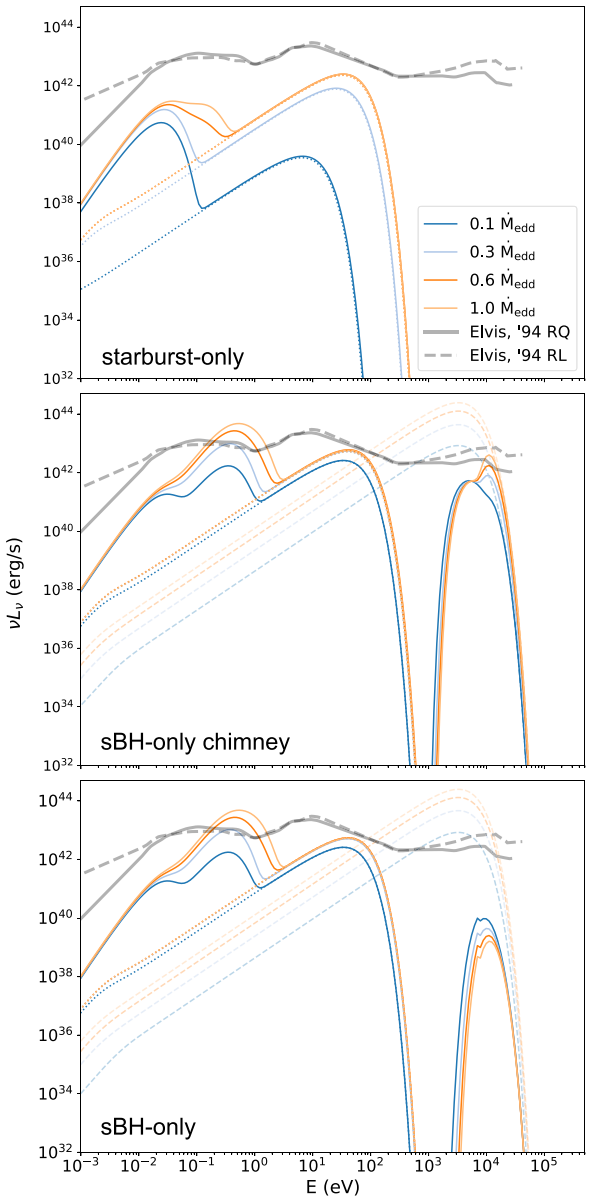

Figure 8. SEDs emerging in the starburst-only (top panel), sBH-only-chimney (middle panel), and sBH-only (bottom panel) disc solutions presented in Fig. 5. Solid curves show the composite thermal spectrum emitted by the AGN disc. In the sBH-only and sBH-only-chimney models, the solid spectrum also includes the escaping SED of the embedded sBH population – visible as a high-energy bump peaking at ~ 10 keV. In all three panels, the dotted curves show the SED of a standard Shakura–Sunyaev α-disc, with an $\dot{M}$ matching the inner mass flux of each model and $\alpha = 0.3$. The dashed lines in the bottom two panels illustrate the total contributed sBH flux. We discuss these modelled spectra in terms of several component features, including a mid-IR bump ($E \lesssim 0.1$ eV), a near-IR bump (0.1 eV $\lesssim E \lesssim 1$ eV), a UV-bump (1 eV $\lesssim E \lesssim 20$ eV), and an X-ray bump that peaks at ~ 10 keV. The grey lines show the averaged observed emission from radio quiet (solid) and radio loud (dashed) AGN taken from Elvis et al. (1994), normalized by a factor of 5×10^{-3}, such that the integrated strength of UV emission is $L_{UV} = L_{Edd} \simeq 5.8 \times 10^{44}$ erg s^{-1}.

respectively), computed assuming that every annulus in the disc radiates as a black body. In the preceding subsections, we examine the UV, IR, and X-ray emission predicted by our models and compare our results with the mean SED for radio quiet (RQ) and radio loud (RL) quasars from Elvis et al. (1994). These reference spectra are plotted in grey, and scaled by a factor of 5×10^{-3}. This scaling corresponds approximately to the ratio of the integrated UV luminosity of the Elvis et al. (1994) data to the Eddington luminosity of a 4×10^6 M$_\odot$ BH. This provides a useful benchmark for comparison against our relatively low-mass SMBH models.

6.4.1 *The UV-bump*

The UV-bump is conventionally ascribed to thermal emission from the viscously heated disc interior (Malkan 1983; Czerny & Elvis 1987) with energies $\gtrsim 1$ eV. The relationship between the classic thin disc spectrum and the UV-bump present in our model spectra is illustrated explicitly by plotting the Shakura–Sunyaev disc spectra (dotted lines) with $\alpha = 0.3$ and mass flux equal to $\dot{M}(R_{in})$ of our corresponding model spectra. Critically, the Shakura–Sunyaev spectra are indistinguishable from their sBH/starburst supported counterparts in the UV-bump domain, suggesting that this feature is parametrized entirely by the mass flux ($\dot{M}$) close to the SMBH and the mass of the SMBH.

Note that the UV-bump predicted by our models is bluer than the bump present in the Elvis et al. (1994) composite spectra. This discrepancy may be explained by recalling that Shakura–Sunyaev discs have a maximum temperature of $k_B T \sim 10(\dot{m}/10^8 M_\odot)^{1/4}$ eV where $\dot{m} = \dot{M}/\dot{M}_{Edd}$ (Shakura & Sunyaev 1973). The relatively low mass of our adopted SMBH, which is a factor of ~ 100 times lower than the typically 10^8–10^9 M$_\odot$ SMBHs powering the bright quasars in the Elvis et al. (1994) sample, can explain the apparent shift of our UV-bump to a factor of 2–3 higher energies.

The spectra shown in Fig. 8 are also faint relative to most observed AGN. With bolometric luminosities of $\sim 10^{-2}$ L$_{Edd}$ or 5×10^{44} erg s^{-1}, these models are significantly less luminous than typical bright AGNs, which emit closer to their Eddington limit with luminosities of $\sim 10^{46}$ erg s^{-1}. The relatively low luminosities we see in Fig. 8 are not uncommon for discs surrounding low-mass SMBHs (Greene & Ho 2007), but suggest that these results cannot be extrapolated to higher mass systems whose accretion rates commonly reach or exceed the Eddington limit. Interestingly, AGN emitting at a lower fraction of their Eddington rate have distinct spectra characterized by the absence of a 'big blue bump' and instead exhibit a broad excess in the mid-IR (e.g. Ho 2008 and references therein); these are spectral features that are somewhat better aligned with our modelled spectra, but also indicative of their limited scope. The dependence of these models on SMBH mass is discussed further in Section 8.

6.4.2 *IR emission*

The mass supply rate and the spectrum have a much less direct relationship in the IR regime, where starburst/sBH heating exceeds viscous heating and sets T_{eff}. The result is IR emission well in excess of a standard steady Shakura-Sunyaev α-disc. Our starburst-only models align with results from TQM05. Low-mass supply rates ($\dot{M}_{out} < \dot{M}_{crit}$) result in depleted inner disc mass flux and a correspondingly small, red-shifted UV-bump. In the $\dot{M}_{out} = 0.1\ \dot{M}_{Edd}$ case, flux from the outer, star-forming region actually dominates thermal emission – contrary to observations. Higher $\dot{M}_{out}$ models

have a broader and subdominant IR-bump. The broadening of the IR peak is the result of additional star formation and increased $T_{\rm eff}$ in the opacity gap at ~ 0.1 pc.

At energies between 10^{-3} and 1 eV, observed emission also diverge from the classic α-disc, with an apparent emission excess in the IR. TQM05 concluded that this emission was broadly consistent with their modelled results. This excess is, however, commonly attributed to the reprocessing of inner disc flux by dust in a circumnuclear torus (Antonucci 1993). Still, as of yet, there is no universally accepted model template for torus emission. Additionally, factors such as contamination from host galaxy IR light and complex torus models with numerous degrees of freedom have led to a wide range of results (Netzer 2015). Several models have been proposed in which the source of IR emission is a warped or flared disc (Sanders et al. 1989; Baskin & Laor 2018; Landt et al. 2023) and imaging by the GRAVITY collaboration suggests that, in the case of NGC 1068 the near-infrared (NIR) emission is local to a thin disc (GRAVITY Collaboration 2020).

Additional constraints on the NIR emission come from reverberation mapping studies in AGN. These studies show that the NIR bump typically lags the optical/UV continuum by tens to hundreds of days (Suganuma et al. 2006; Koshida et al. 2014). These results are suggestive of dust heated and reprocessed by a circumnuclear torus as opposed to local emission by stars/sBHs in the disc. However, reverberation mapping campaigns have limitations in accurately separating the accretion disc contribution from toroidal dust emission. This limitation has been recently emphasized in the spectroscopic reverberation studies by Landt et al. (2019, 2023), in which a 'constant red component' is observed in the NIR spectrum of NGC 5548 and Mrk 876. This component is separate from the variable emission attributed to hot dust and represents an enhanced emission in the NIR over the standard α-disc spectrum. This behaviour is consistent with auxiliary heating from a disc-embedded population.

In our sBH-only and sBH-only-chimney models, emission in the IR is further enhanced, because the sBHs allow a higher mass inflow rate into the inner regions of the disc. The result is a larger surface density interior to 0.1 pc, requiring a higher heating rate to stabilize this region. The resulting IR bump is shifted to higher energies – dominated by the near, not the mid, IR. Higher $\dot{M}_{\rm out}$ leads to further enhancement of the NIR relative to the rest of the spectrum, making it the dominant spectral feature in the $\dot{M}_{\rm out} = 0.3$, 0.6, and 1.0 $\dot{\rm M}_{\rm Edd}$ sBH models. As illustrated in the Elvis et al. (1994) composite spectra, such strong NIR emission is not commonly observed in AGN – indicating that sBH-only supported discs are rare. This constraint is discussed in more detail in Section 7.3.

6.4.3 *X-ray emission*

Hard X-ray emission (> 2 keV) in AGN has been widely attributed to inverse Compton scattering of disc photons in a hot corona sandwiching the AGN (Shapiro, Lightman & Eardley 1976; Sunyaev & Titarchuk 1980; Haardt & Maraschi 1991). The emission is characterized by a power-law shape between 2 and 10 keV. These X-ray photons may also reflect off the nearby disc, creating the so-called Compton hump, an excess of hard X-rays over the power-law continuum, peaking between ∼20 and 30 keV (Netzer 2015). The observed spectrum is therefore the sum of primary continuum and reprocessed X-rays.

In the bottom two panels of Fig. 8, there is a spectral bump at $\sim$ 10 keV due to emission from embedded sBHs escaping the disc.

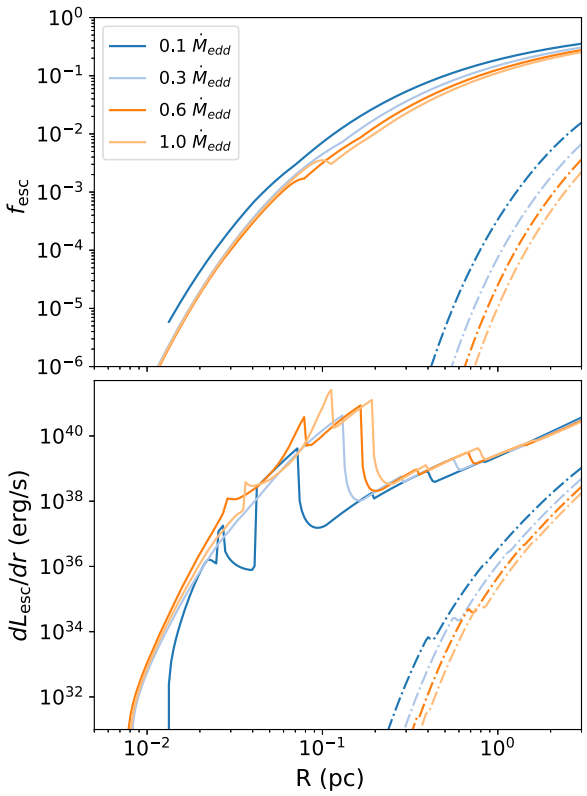

Figure 9. Escape fraction $f_{\rm esc}$ (top panel) and escape luminosity per log-radial bin $\mathrm{d}L_{\rm esc}/dr$ (bottom panel) as a function of distance from the central SMBH for the sBH-only-chimney models (solid lines) and sBH-only models (dot–dashed lines). Across all disc models, $f_{\rm esc}$ is largest in the outer disc. However, the large sBH population in the opacity gap region combined with a slightly higher $f_{\rm esc}$ means that a $\mathrm{d}L_{\rm esc}/\mathrm{d}r$ is still significant in the chimney models at distances between $2 \times 10^{-2} - 2 \times 10^{-1}$ pc.

From lowest to highest $\dot{M}_{\rm out}$, the ratio of the escaping sBH flux to the bolometric luminosity in the sBH-only-chimney models is 6.9×10^{-2}, 3.4×10^{-2}, 2.6×10^{-2}, and 2.8×10^{-2}. In the sBH-only models, the ratios are 9.3×10^{-4}, 1.2×10^{-4}, 3.1×10^{-5}, and 1.1×10^{-5}. The significant contribution of X-ray flux to bolometric luminosity in the chimney case is consistent with the lower optical depth and enhanced $f_{\rm esc}$ in these models. Aside from this additional X-ray flux, the intrinsic disc flux of the sBH-only and sBH-only-chimney models is approximately the same.

The escaping sBH flux peaks just beyond the low-energy edge of the Compton hump (∼10 keV). The apparent absence of an observed emission excess at this energy in typical AGN spectra, beyond what can be attributed to coronal emission or Compton upscattering, suggests a constraint on the number of sBHs expected in AGN. Moreover, the hard X-ray emission from sBHs should differ from intrinsic AGN X-ray emission in several important ways. Notably, Compton hump emission is expected to be polarized (Podgorný et al. 2023), experience reverberation lags coincident with coronal variability (Uttley et al. 2014; Kara et al. 2015; Zoghbi, Miller & Cackett 2021), and may be constrained to the inner disc (Fabian et al. 1989; Miniutti & Fabian 2004) – none of which apply to the flux emitted from the sBHs modelled here.

In Fig. 9, we plot $f_{\rm esc}$ (top panel) and $L_{\rm esc}$ (bottom panel) as a function of disc radius for the sBH-only-chimney (solid lines)

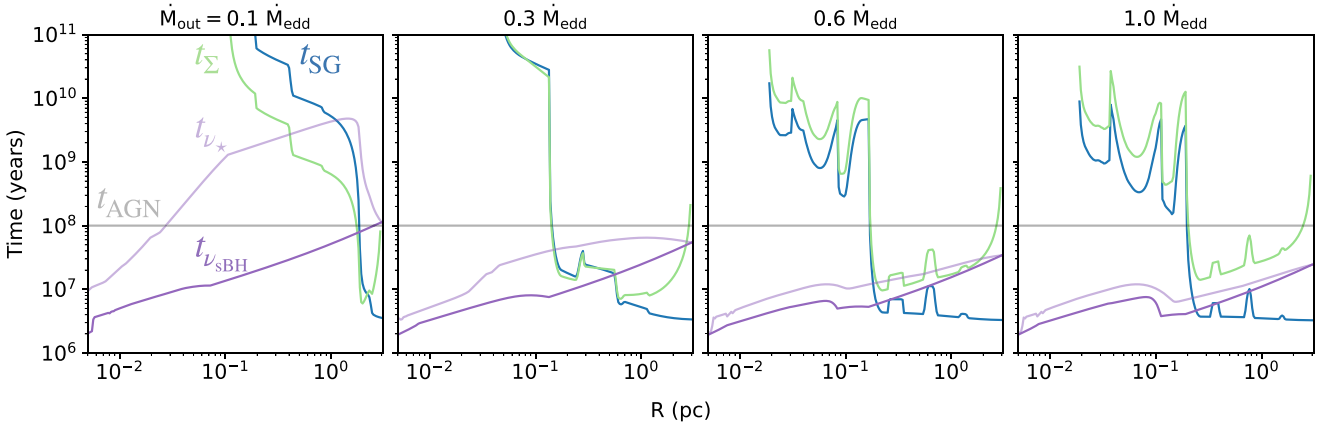

Figure 10. Anticipated time-scales for sBH population growth t_{SG} (blue) and surface density change t_Σ, assuming that the disc is initially in a starburst-only state and evolves toward the sBH-only state. The grey line marks the expected lifetime of the disc $t_{AGN} = 10^8$ yr. Light and dark purple lines indicate the viscous time-scales for the starburst-only and sBH-only disc, respectively. From left to right, panels show these time-scales for the $\dot{M}_{out} = 0.1\ \dot{M}_{Edd}$, $0.3\ \dot{M}_{Edd}$, $0.6\ \dot{M}_{Edd}$, and $1.0\ \dot{M}_{Edd}$. These time-scales provide a useful point of reference for the evolving models described in Section 7.

and sBH-only (dot–dashed lines) models. Note that in the sBH-only models the escaping X-ray emission is limited to the outer disc, beyond 0.1 pc and the opacity gap, while in the sBH-chimney models, L_{esc} persists out to 0.01 pc. Note that the energy of escaped emission is primarily a function of Σ, which ranges from 10^{-1} to 10^1 g cm^{-2} at these radial distances. X-rays escape when $\tau_X = 1$ or $\sigma_X \sim 10^{-23}$ to 10^{-25} cm^2. From Fig. 2, we see that these cross-sections correspond to energies straddling 10 keV – confirmed by the location of the X-ray peak in Fig. 8.

Both because it is constrained to the outer disc and is sourced from upwards of 100 individual sBHs, the escaped X-ray flux should, on average, exhibit minimal variability over long time-scales. In contrast, hard X-ray emission observed in AGN typically varies on very short time-scales – years (Soldi et al. 2014) or even hours (Ponti et al. 2012) – further suggesting the X-ray emitting region is compact (e.g. De Marco et al. 2013; Ursini et al. 2020). Additionally, the fractional root mean square variability amplitude can be as a high as 50 per cent (Caballero-Garcia et al. 2012). To ensure sBH emission does not dominate AGN emission in the 5–10 keV band or impact variability significantly at those energies, we estimate that sBHs can contribute only ∼1–10 per cent of observed X-ray emission. Assuming intrinsic AGN X-ray luminosity is ∼10 per cent of the bolometric luminosity, we conservatively estimate a normalized sBH contribution limit of $\sim 10^{-3}\ L_{bol}$.

This limit is violated by our chimney models, which predict an excess X-ray emission exceeding $10^{-2}\ L_{bol}$. The absence of such detectable excess in observations strongly suggests that the chimney models cannot accurately describe the behaviour of embedded sBHs in AGN discs. Note too that the relative strength of the X-ray peak in the low $\dot{M}_{out}$ sBH-only cases indicated that constraining our models may be more feasible by examining the spectra of low-luminosity AGN.

The prediction of an X-ray emission excess in AGN spectra from sBHs suggests a new observational constraint on the number of sBHs embedded in the disc and further study of X-ray flux from AGN is likely to place tight upper limits on the number of sBHs present in the outer radii of AGN discs. A more thorough numerical analysis, incorporating additional physics such as coronal emission and other deviations from thermal thin-disc spectra in the emission spectra of sBHs, is left to future work.

6.5 sBH population growth time-scale

The steady-state models developed in this section treat stars and sBHs independently, but our goal is to understand how the rate of star formation responds to the growth of the sBH remnant population. From this perspective, the starburst-only model can be viewed as the 'initial state' of the AGN disc while the sBH-only disc approximates a hypothetical 'final state.' The time required to evolve from the initial to the final state is calculated by assuming a constant star formation rate, according to

$$t_{SG} = \frac{S_{sBH}}{\dot{\Sigma}_\star} \frac{\int_{m_{min}}^{m_{max}} m_\star^{1-\delta} dm_\star}{\int_{m_{trans}}^{m_{max}} m_\star^{-\delta} dm_\star} \simeq 2.2 \times 10^7\ \text{yr} \times \left(\frac{S_{sBH}}{41\ \text{pc}^{-2}}\right) \left(\frac{\dot{\Sigma}_\star}{1.8 \times 10^{-3} \dot{M}_{Edd}\ \text{pc}^{-2}}\right)^{-1}. \quad (37)$$

In Fig. 10, we have plotted t_{SG} as a function of radius (blue) for each of our $\dot{M}_{out}$ cases. For comparison, we have also emphasized the expected AGN lifetime $t_{AGN} = 10^8$ yr with a horizontal grey line.

In the three highest $\dot{M}_{out}$ models, t_{SG} is between 10^6 and 10^7 yr when $r > 0.1$ pc, exterior to the opacity gap. We therefore expect that this outer region will be fully supported by sBHs well within the disc lifetime. Interior to the opacity gap, t_{SG} increases sharply such that $t_{SG} > t_{AGN}$, suggesting that the inner disc never reaches an sBH-only supported state. We conclude that the outer disc will quickly reach a single steady-state disc solution – not unlike the 'pile-up' solution described in Gilbaum & Stone (2022). The inner disc, in contrast, will continue to change over the disc lifetime as the sBH population grows.

In the $\dot{M}_{out} = 0.1\,\dot{M}_{Edd}$ case, $t_{SG} < t_{AGN}$ exterior to $r \gtrsim 1$ pc, suggesting that a relatively small fraction of the disc will reach a steady-state in the disc's lifetime. Note, however, that our calculation of t_{SG} does not account for changes in $\dot{M}$ between the initial and final disc states. Beyond $r \gtrsim 1$ pc in the starburst-only model, $\dot{\Sigma}_\star$ is comparable to the higher $\dot{M}_{out}$ cases – the high star formation rate depletes the disc and both $\dot{M}$ and $\dot{\Sigma}_\star$ drop precipitously interior to $r \simeq 1$ pc. The sBH-only model, on the other hand, maintains a nearly constant $\dot{M}$ across the disc and requires auxiliary pressure support from sBHs down to $r \sim 0.1$ pc. As the outer disc evolves from the starburst-only to the sBH-only state, the mass flux through

the disc will increase and is expected to ignite star formation in the disc interior. Enhanced star formation will decrease the time required for the disc to reach the sBH-only, final state. While the change in mass flux is most pronounced in the lowest $\dot{M}_{\rm out}$ case, we expect it to affect star formation over time regardless of boundary conditions.

While $t_{\rm SG}$ represents the time required for the disc to reach an sBH-only state, the local viscous time yields an estimate over which the gas surface density can change to adjust to the evolving number of sBHs. We plot the viscous time in the starburst-only ($t_{\nu_\star}$) and sBH-only disc models ($t_{\nu_{\rm sBH}}$) in Fig. 10. The change in surface density between the starburst-only and sBH-only disc models is parametrized by the time-scale t_Σ, calculated as

$$t_\Sigma = \frac{\Sigma_\star}{|\Sigma_\star - \Sigma_{\rm sBH}|} t_{\rm SG}, \tag{38}$$

that is, the surface density in the starburst-only case divided by the approximate rate at which the surface density changes as the sBH population grows. If the viscous time-scale exceeds t_Σ, then the transition time required to reach the steady-state solution is longer than the rate at which we assume Σ is changing. Note that, in the two highest $\dot{M}_{\rm out}$ cases, $t_\Sigma \lesssim t_{\nu_\star}$ and $t_{\nu_{\rm sBH}}$, suggesting Σ can efficiently reach equilibrium according to the changing mass flux through the disc. But for lower $\dot{M}_{\rm out}$, $t_\Sigma > t_{\nu_\star}$, and $t_{\nu_{\rm sBH}}$, where $t_{\rm SG} < t_{\rm AGN}$, thus limiting the applicability of our model in these cases. We expect that a fully time-dependent, hydrodynamic model will be necessary to evaluate these cases, although this falls outside the scope of the current paper.

It is useful at this point to revisit our discussion of the effect of sBH growth on our disc models. As mentioned in Section 5, we do not take into account the growth of sBHs, setting the sBH mass to $10\,{\rm M}_\odot$ in the models described here and in the following section. We assume, in accordance with Fig. 10, that the sBHs in the outer disc will be seeded within the first ten million years of the disc lifetime. If they then undergo Eddington capped accretion for the remaining $\sim 9 \times 10^7$ yr, both their accretion rate and mass will have increased by approximately an order of magnitude. By comparing the rates of mass flux through the disc and gas consumption by sBHs presented in Fig. 5, we can see that an order of magnitude increase in accretion by sBHs is still well below $\dot{M}$. We therefore do not expect sBH growth to significantly change the evolution of the interior disc and can therefore safely neglect it.

Finally, we note that over the lifetime of the AGN, the SMBH will grow at approximately the same rate as the mass flux through the disc interior. We do not incorporate this growth into our models but can estimate a change in mass of approximately $10^6\ {\rm M}_\odot$ assuming a growth rate of $0.1\ \dot{\rm M}_{\rm Edd}$ over 10^8 yr. This amounts to an error of 25 per cent in our models, which should not significantly affect our results.

7 DISC EVOLUTION AS A SEQUENCE OF STEADY STATES

In this section, we model changes in AGN over time assuming that their evolution can be approximated as a series of steady states. This choice is justified so long as $t_{\rm th} \ll t_{\rm dyn}$ and $t_\nu < t_\Sigma$. The former assures that the disc can self-regulate and, as discussed in Section 3.1, is satisfied for reasonable disc parameters. The latter is not satisfied for all boundary conditions, but is expected when $\dot{M}_{\rm out} \gtrsim 0.6\ \dot{M}_{\rm Edd}$. Nevertheless, for completeness, we analyse discs with $\dot{M}_{\rm out} = 1.0$, 0.6, 0.3, and $0.1\ \dot{M}_{\rm Edd}$, but highlight the caveat that the discs with lower mass accretion rates will require full, time dependent, hydrodynamic equations to confirm the results found here.

We assume, for all models in this section, that the X-ray emission from sBHs interacts with the disc via photoionization, dust absorption and scattering, and Compton scattering. Our X-ray opacity is therefore consistent with the sBH-only models as opposed to the chimney models described in Section 6. Although, as discussed there, we do not expect the choice of X-ray opacity to have a significant effect on the physical disc structure, it does affect the emerging hard X-ray spectrum.

7.1 Star formation and sBH accretion in the evolving disc

The dashed lines in Fig. 11 show the mass fluxes through the disc for our four boundary conditions. Solid and dotted lines indicate the corresponding local star formation rate and sBH accretion rate, respectively. From top to bottom the four panels in Fig. 11 provide snapshots of the disc at 10, 25, 50, and 100 Myr.

Looking at the top panel of Fig. 11, we see that within the first 10 Myr, the $\dot{M}_{\rm out} = 1.0$ and $0.6\ \dot{M}_{\rm Edd}$ cases are largely supported by heating from sBHs at distances beyond 0.2 pc, although rings of star formation persist at $r \simeq 0.2$, 0.7, and 1 pc. This is consistent with the sBH growth time ($t_{\rm SG}$) calculated for the outer disc in Fig. 10, where peaks in $t_{\rm SG}$ correspond to a slight shifting of the opacity away from the SMBH in the sBH-only models relative to the starburst-only models.

Star formation spans a broader portion of the disc in the $\dot{M}_{\rm out} = 0.3$ and $0.1\ \dot{\rm M}_{\rm Edd}$ cases after 10 Myr. These disc models have not built up a large enough population of sBHs to support the outer disc. This is particularly clear in the lowest $\dot{M}_{\rm out}$ case where sBHs have yet to be seeded at distances less than 1 pc. Recall that in both of these cases, $\dot{M}_{\rm out} < \dot{M}_c$, and their initial starburst-only state is characterized by a high star formation rate in the outer disc and mass depletion in the inner disc. As the sBH population in the outer disc grows, star formation decreases, allowing increased mass flow and igniting star formation in the inner disc. This behaviour can also be seen in the top panel of Fig. 13 where we plot the mass flux into the disc interior over time. Note that initially the $\dot{M}_{\rm out} = 0.3$ and $0.1\ \dot{\rm M}_{\rm Edd}$ cases have $\dot{M}_{\rm in} = 2 \times 10^{-2}$ and $1 \times 10^{-4}\ \dot{M}_{\rm Edd}$, respectively. But the mass flux increases over time, reaching a maximum rate of $0.07\ \dot{M}_{\rm Edd}$ within 6 Myr in the $0.3\ \dot{M}_{\rm Edd}$ case and 20 Myr in the $0.1\ \dot{M}_{\rm Edd}$ case.

At 25 Myr, star formation is shut down beyond ~ 0.2 pc across all modelled discs. There are regions of the outer disc where sBH heating exceeds what necessary to maintain marginal stability. This is particularly clear in the two highest $\dot{M}_{\rm out}$ cases where $h/r > 0.1$, as shown in Fig. 12. Moreover, because we omit sBH growth and capture from the surrounding nuclear cluster, we likely underestimate the degree of excess heating. These inflated regions could experience irradiation from the disc interior, contributing to NIR variability.

25 Myr also marks a distinct transition in the behaviour of the highest $\dot{M}_{\rm out}$ disc. As shown in the second panel from the top of Fig. 11, the star formation rate jumps by nearly an order of magnitude at $r \simeq 0.2$ pc. The peak in star formation causes the mass flux through the disc to drop and, as a result, star formation ceases in the disc interior. This jump in star formation may be understood as a sudden switch between degenerate solutions: from a 'cool', high optical depth, radiation pressure dominated solution to a 'hot', low optical depth, gas pressure dominated solution. The jump is forced by the added heating from accreting sBHs, which increases $T_{\rm eff}$ and $p_{\rm rad} \propto \tau T_{\rm eff}^4$. To maintain hydrostatic equilibrium, τ and $p_{\rm rad}$ drop such that $p_{\rm gas} > p_{\rm rad}$.

The gulf between the hot and cold solution parameters is more concretely illustrated in Fig. 12, where we have plotted relevant disc parameters at 10, 25, 50, and 100 Myr, including τ, $p_{\rm rad}$, and $p_{\rm gas}$.

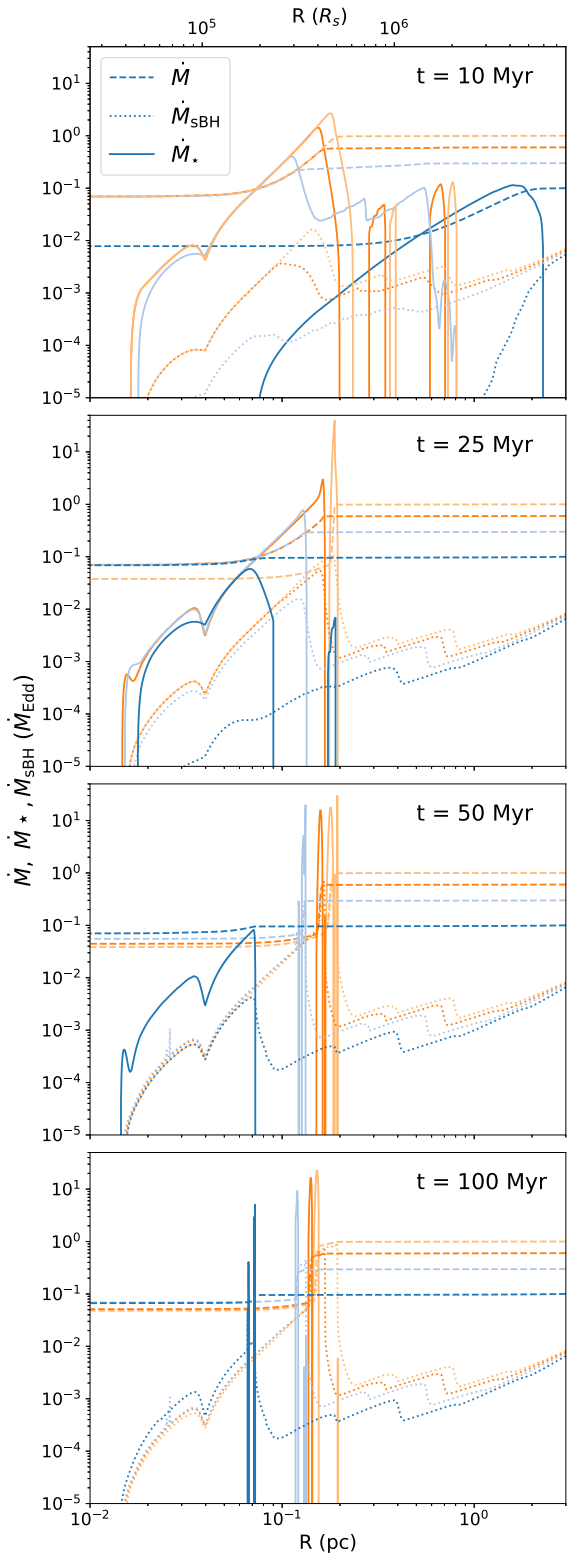

Figure 11. The total star formation rate ($\dot{M}_\star = 2\pi r^2 \dot{\Sigma}_\star$; solid line), the mass flux in the background disc ($\dot{M}$; dashed line), and the total accretion rate on to sBHs ($\dot{M}_{\rm sBH} = 2\pi r^2 \dot{\Sigma}_{\rm sBH}$; dash–dotted line) as a function of disc radius for boundary conditions $\dot{M}_{\rm out} = 1.0$ (light orange), 0.6 (dark orange), 0.3 (light blue), and 0.1 (dark blue) $\dot{M}_{\rm Edd}$. Moving downwards, panels illustrate $\dot{M}$ profiles at 10, 25, 50, and 100 Myr. In all models, the discs evolve towards a state in which star formation is amplified but limited to a narrow annulus. This star formation peak persists across the rest of the disc lifetime.

Note that between 10 and 25 Myr, there is a sharp drop in τ and $p_{\rm rad}$ at 0.1 pc in the $\dot{M}_{\rm out} = \dot{M}_{\rm Edd}$ case, corresponding to the jump in the star formation rate.

By 50 Myr, star formation in the $\dot{M}_{\rm out} = 0.6$ and 0.3 $\dot{M}_{\rm Edd}$ cases is limited to narrow rings, peaking at $r \simeq 0.16$ and 0.09 pc respectively. And by 100 Myr all of our disc models exhibit this behaviour. Although limited to a narrow ring, the cumulative mass consumption rate by stars actually increases. This is also illustrated in the top panel of Fig. 13, where $\dot{M}_{\rm in}$ drops steeply from 0.07 $\dot{M}_{\rm Edd}$ to 0.038, 0.044, 0.055, and 0.067 from highest to lowest $\dot{M}_{\rm out}$. The drops are slightly shifted in time, occurring at 25, 29, 40, and 83 Myr, with lower $\dot{M}_{\rm out}$ models taking longer to shift star formation into the hot state. In the middle panel of Fig. 13, we have plotted the rates of mass consumption by stars and sBHs over time. Note that the drops in $\dot{M}_{\rm in}$ shown in the top panel correspond to small increases in $\dot{M}_\star$ seen in the middle panel. Mass consumption by sBHs, on the other hand, increases over time at a steady rate – remaining subdominant to star formation over the disc lifetime.

After 100 Myr, the star formation peak in the $\dot{M}_{\rm out} = \dot{M}_{\rm Edd}$ case has shifted slightly towards the SMBH, moving from $r \simeq 0.19$ pc to 0.15 pc. As predicted in Section 6.5, the time-scale required for this region to become fully supported by sBHs is $\sim 10^{10}$ yr. However, the enhanced rate of star formation reduces this time-scale to approximately the lifetime of the disc. Here, we choose to stop evaluating our models after 100 Myr, although we expect that the star-forming rings would continue to move inwards as they populated their respective regions with sBHs. However, after 100 Myr, we expect the growth of sBHs in the outer disc and in the increasingly populated inner disc to have a non-negligible effect on the disc structure.

Our models are suggestive of a nuclear disc of stars, not unlike that observed in the Galactic Centre (Ghez et al. 2005). Measurements of stellar orbits by GRAVITY Collaboration (2022) have allowed for high-precision determination of the gravitational potential around Sgr A∗. They find that the extended mass enclosed in 2.4×10^4 $R_{\rm s}$ cannot exceed 3×10^4 $M_\odot$. In the bottom panel of Fig. 14, we plot the combined mass of stars and sBHs within a distance R after 100 Myr. We find that across all models, the mass distribution is within the bounds measured by GRAVITY Collaboration (2022). Moreover, observations of X-ray binaries indicate that there are $\simeq 10^4$ sBHs within 1 pc of Sgr A∗ (Hailey et al. 2018) and there is evidence that a population of 10^4 sBHs within 0.1 pc is necessary to produce the quasi-thermal eccentricities of a population of stars in the Galactic Centre known as 'S-stars' (Antonini 2014). However, the dynamical impact of disc loss and subsequent interactions between embedded objects has not been thoroughly investigated. Given these complicating factors, direct comparison between observed and predicted mass distributions will require extensive and careful modelling.

7.2 Numerical resolution and convergence

Our code requires us to verify convergence of our results with respect to both the temporal and spatial resolution used. The spatial resolution is particularly critical given that, for a large part of the discs's lifetime, star formation and the resulting mass flux change occurs within a very limited radial range. We perform radial resolution tests for $\dot{M}_{\rm out} = \dot{M}_{\rm Edd}$, assuming resolution convergence in these tests to apply for other reasonable boundary conditions. Our tests include a base resolution of 500 logarithmically spaced points in the Toomre unstable region – between $R_{\rm out}$ and 0.01 pc. The rest of the disc is less highly resolved with 304 points between 0.01 pc and $R_{\rm in}$. We then further enhance the resolution interior

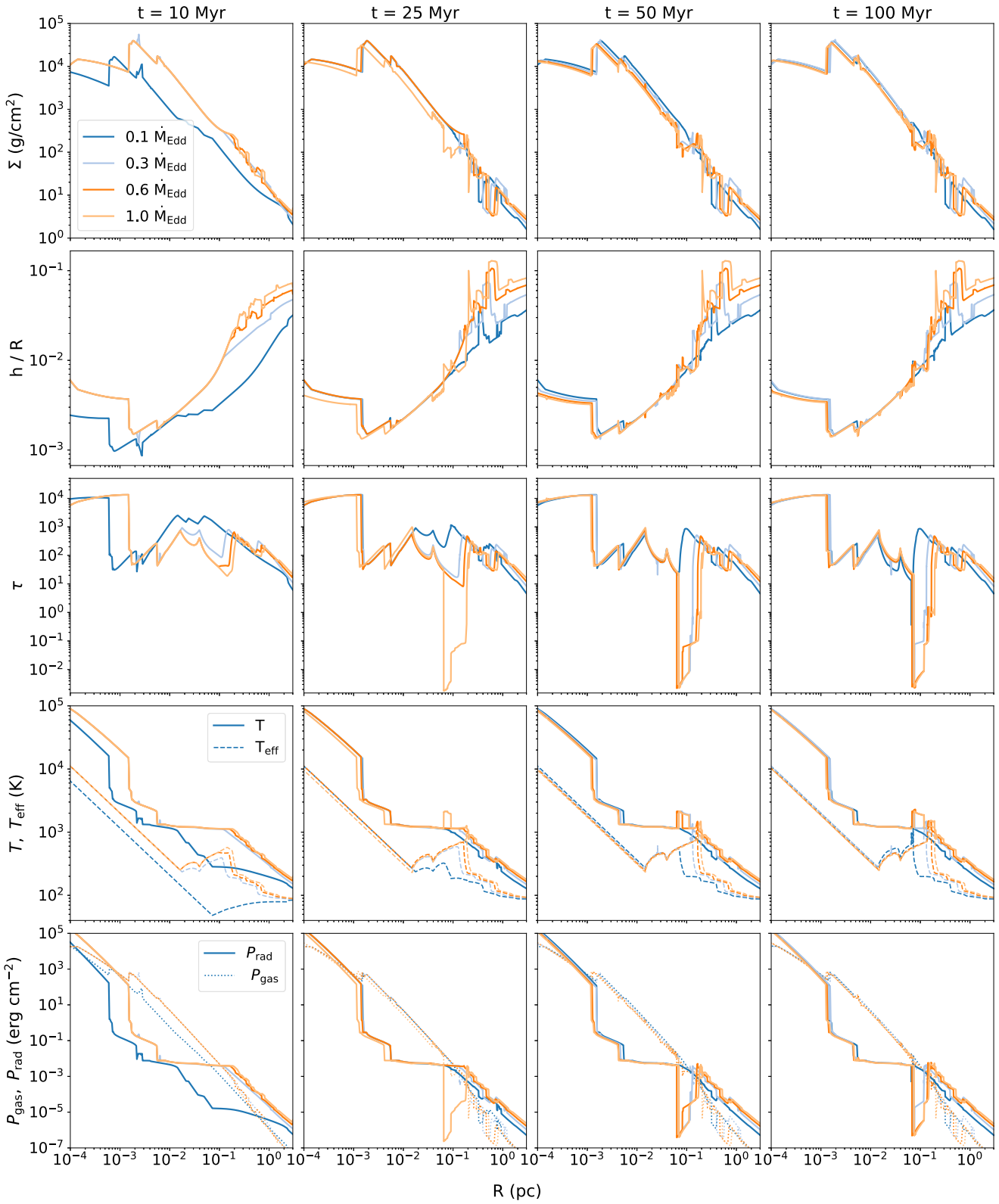

Figure 12. From top to bottom, different rows show the profiles of different physical quantities: surface density (Σ); ratio of scale height to radius h/R; optical depth (τ); temperature (T, solid line) and effective temperature ($T_{\rm eff}$, dashed line); and radiation pressure ($p_{\rm rad}$, solid line) and gas pressure ($p_{\rm gas}$, dotted line). Columns from left to right show snapshots at 10, 25, 50, and 100 Myr. Colours are used to differentiate boundary conditions as in Fig. 11.

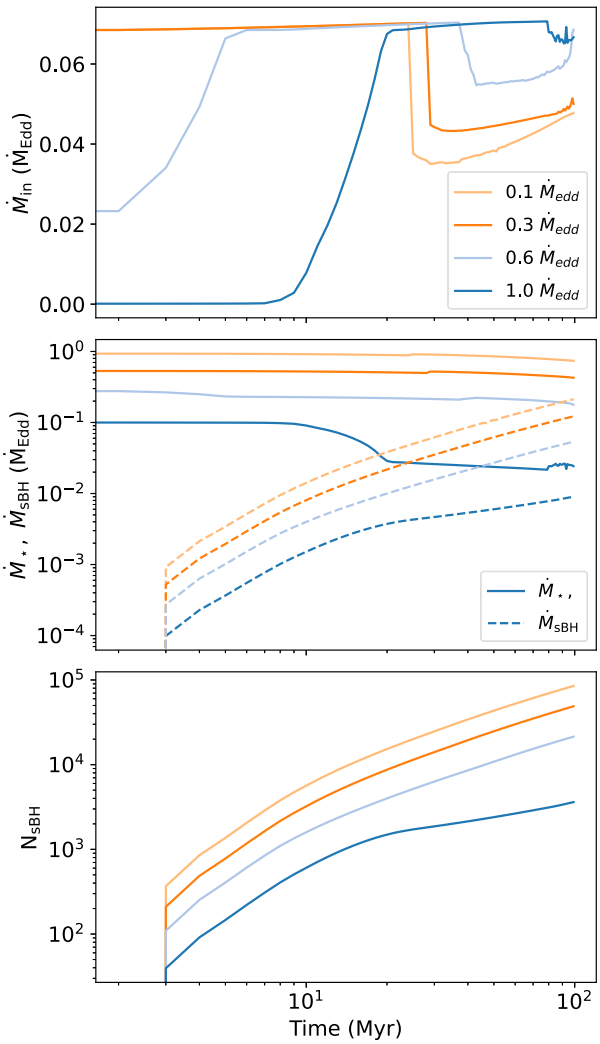

Figure 13. Top panel: Mass flux through the inner disc $\dot{M}_{in}$ as a function of time for the four different boundary conditions: $\dot{M}_{out} = 1\,\dot{M}_{Edd}$ (light orange), $0.6\,\dot{M}_{Edd}$ (dark orange), $0.3\,\dot{M}_{Edd}$ (light blue), and $0.1\,\dot{M}_{Edd}$ (dark blue). Middle panel: Total mass consumption by stars, i.e. $\sum_{i=0}^{n} 2\pi r_i \Delta r_i \dot{\Sigma}_\star(r_i)$ (solid line) and sBHs (dashed line) over time. Bottom panel: Total number of sBHs in the disc over time.

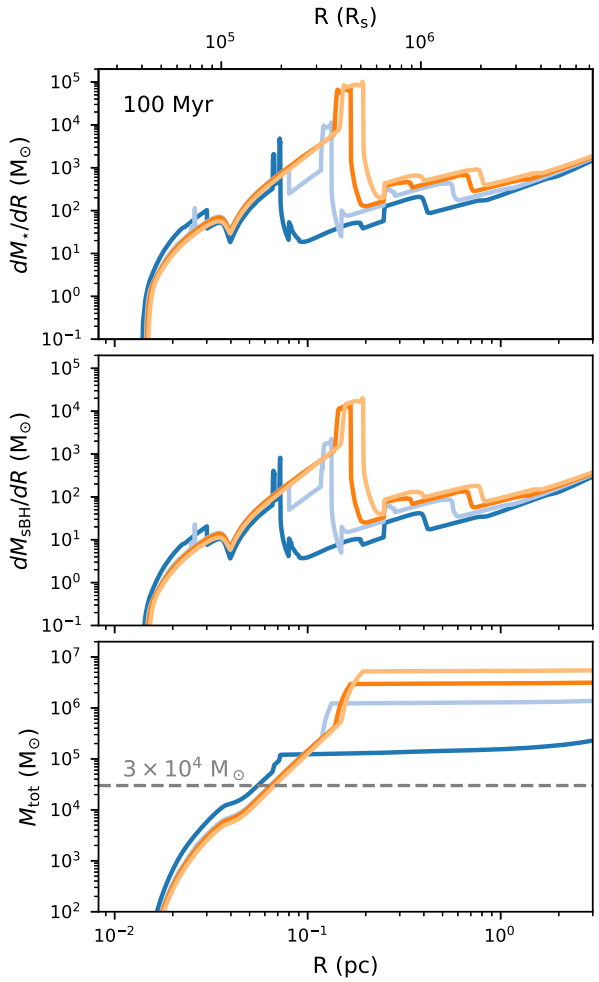

Figure 14. Top panel: Distribution of stellar mass as a function of distance from the SMBH after 100 Myr. As in previous figures, the different colours indicate different mass supply rates. Middle panel: Distribution of sBH mass. Bottom panel: Combined mass of stars and sBHs within a distance R of the SMBH. The grey dashed line indicates the current maximum extended mass within 2.4×10^4 R_s of the Galactic Centre, as determined by GRAVITY Collaboration (2022).

to the opacity gap – the region where star formation persists in narrow rings –increasing the number of logarithmically spaced radial points between 2.5×10^{-5} and 10^{-2} by 400, 500, and 600 points, representing an enhancement in resolution over the base model of 80, 180, and 280 points, respectively.

In the top panel of Fig. 15, we plot $\dot{M}_{in}$ over 100 Myr for each of our radial resolution tests. We note that this parameter is sensitive to small changes in the star formation rate, making it a sensitive probe of resolution convergence. While all of our tests exhibit the same general behaviour, $\dot{M}_{in}$ drops slightly earlier in the two lowest resolution cases. The two lowest resolution cases also see increased fluctuation in $\dot{M}_{in}$ after the drop. The difference between a resolution enhancement of 180 and 280 points is, however, minor, so we conclude that at these resolutions our model has converged. In the discs examined here we use the highest radial resolution from our tests, although decreasing the resolution in the opacity gap by 100 points yields comparable results.

The bottom panel of Fig. 15 shows the results of our time resolution tests, comparing $\dot{M}_{in}$ for models with $\Delta t = 0.5$, 1, 2.5, and 5 Myr. Our plots indicate that, for resolutions at or below $\Delta t = 2.5$ Myr, our models converge. Even $\Delta t = 5$ Myr looks fairly smooth until $t \simeq 70$ Myr, at which point $\dot{M}_{in}$ increases significantly, diverging from values found at higher resolution solutions. These results suggest that our models are much less sensitive to time resolution than radial resolution, and converge for $\Delta t \leq 2.5$ Myr. This is reasonable given that 2.5 Myr is the approximate time delay between star formation and evolution of the most massive stars off the MS. In our models, we use $\Delta t = 1$ Myr.

7.3 Emerging spectral energy distributions

In Fig. 16, we show the spectra emerging from our evolving disc models. From the top-most to bottom-most panel, our spectra are taken at 10, 25, 50, and 100 Myr – as in the disc solutions presented in Figs 11 and 12. These spectra include both the intrinsic disc flux (extending to 2×10^2 eV) and the escaped X-ray flux from

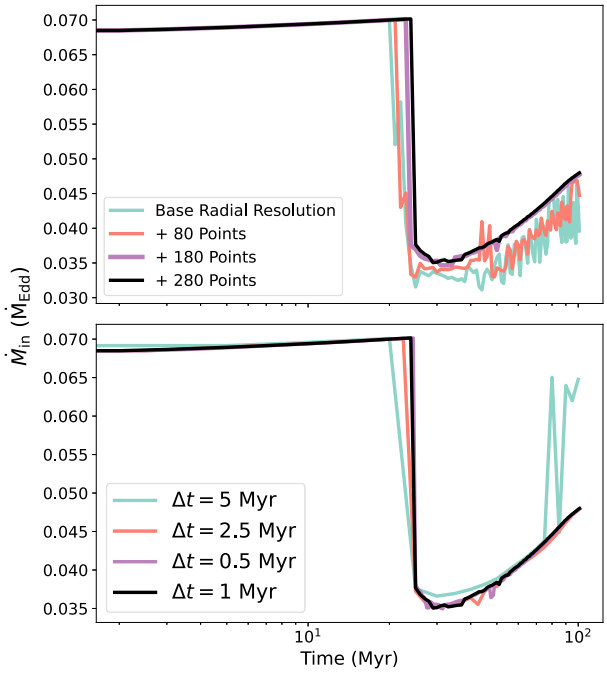

Figure 15. Evolution of $\dot{M}_{\rm in}$ over 100 Myr for different numerical resolutions in space and time. Black lines indicate the resolution used in the evolving models described here. All resolution tests share the boundary condition $\dot{M}_{\rm out} = \dot{M}_{\rm Edd}$. Top panel: Spatial resolution tests including a base radial resolution of 500 points between $R_{\rm out} = 3$ and 0.01 pc (blue) and enhanced resolutions between 0.25 and 0.01 pc by an additional 80 (red), 180 (purple), and 280 (black) points. Models converge for resolutions exceeding 180 additional points. Bottom panel: Temporal resolution tests where Δt is set to 5 Myr (blue), 2.5 Myr (red), 0.5 Myr (purple), and 1 Myr (black). Models converge for $\Delta t \leq 2.5$ Myr.

embedded sBHs (peaking at ∼10 keV). As in Fig. 8, the composite quasar spectra from Elvis et al. (1994), normalized by a factor of 5×10^{-3}, are plotted in all panels for reference.

As discussed in Section 6.4.3, the escaped flux from sBHs comes almost exclusively from the outer disc. This population of sBHs is seeded within the first ∼25 Myr of the disc lifetime, after which the escaped sBH X-ray emission is constant. From smallest to largest $\dot{M}_{\rm out}$, the ratios of escaped sBH emission to bolometric luminosity at 25 Myr are 8.0×10^{-3}, 5.0×10^{-3}, 3.5×10^{-3}, and 3.3×10^{-3}. In general, these values represent a maximum over the AGN lifetime. At earlier times, the sBH population in the outer disc is still growing, and its emission contributes a smaller fraction to the total. At later times, the sBH emission has stabilized but thermal disc emission increases. An exception occurs in the lowest $\dot{M}_{\rm out}$ case, where the fractional contribution of escaped sBH emission is an order of magnitude larger at 10 Myr than at 25 Myr, as the SMBH feeding rate – and thus the bolometric luminosity – is significantly reduced compared to the other disc models.

For all boundary conditions, the ratio of escaped sBH emission to total disc luminosity predicted by our evolving disc models exceeds the sBH-only cases described in Section 6 and the observational detection limit of $\sim 10^{-3}$ $L_{\rm bol}$ estimated in Section 6.4.3. In all cases except for the 10 Myr snapshot in the lowest $\dot{M}_{\rm out}$ case, the discrepancy is relatively small – within an order of magnitude of the imposed limit – but underscores the clear-cut constraint provided by X-ray observations. As noted earlier, a detailed numerical investigation of this constraint is beyond the scope of this work and

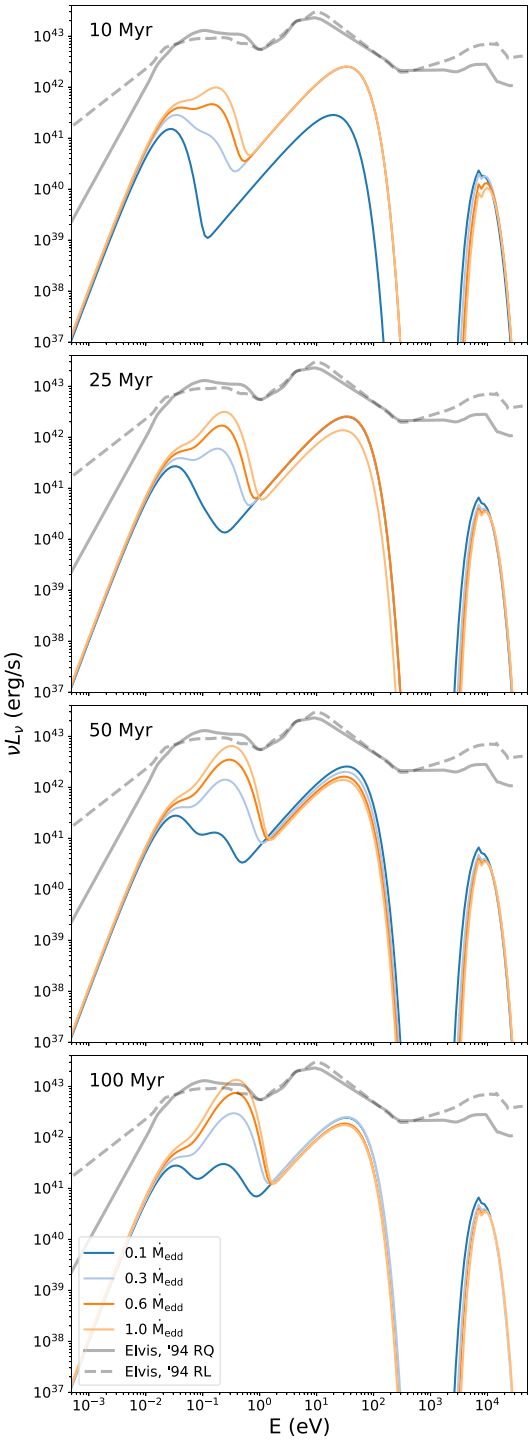

Figure 16. SEDs emerging from our evolving disc models shown at 10, 25, 50, and 100 Myr (from top to bottom panel) as in the disc solutions presented in Fig. 12. The spectra shown include both the intrinsic disc emission and the escaping SED of the embedded sBH population – visible as a high-energy bump peaking at $\sim 10^4$ eV. Grey lines indicate the averaged observed emission from radio quiet (solid) and radio loud (dashed) AGN taken from Elvis et al. (1994), normalized by a factor of 5×10^{-3} as in Fig. 8. Spectral features apparent in the steady-state models (Fig. 8) can also be recognized here, including a mid-IR bump ($E \lesssim 0.1$ eV), a NIR bump (0.1 eV $\lesssim E \lesssim 1$ eV), a UV-bump (1 eV $\lesssim E \lesssim 20$ eV), and an X-ray bump that peaks at ~ 10 eV.

would require a more sophisticated analysis of sBH emission beyond a multicolour black-body spectrum. Still, the expected sBH flux, spanning 2–20 keV, lies well within the range detectable by *NuSTAR* (3–79 keV; Harrison et al. 2013). Observational techniques, such as reverberation mapping, polarimetric analysis, and the measurement of microlensing-induced time delays (Tie & Kochanek 2018), are essential for disentangling sBH emission from the thermal emission intrinsic to the AGN disc. These methods leverage distinguishing characteristics of escaped sBH X-ray emission predicted by these models: it originates in the outer disc, is unpolarized, and varies on long time-scales.

Unlike the X-ray bump, which primarily originates from sBHs in the outer disc, NIR emission reflects the number of sBHs and their contribution to disc heating interior to the opacity gap. As a result, prior to 25 Myr, the intrinsic disc spectra closely resemble the starburst-only spectra shown in Fig. 8. At 10 Myr, the three highest $\dot{M}_{\rm out}$ models display broadened IR peaks, in which the mid-IR (MIR) emission ($E \lesssim 0.1$ eV) is dominant or comparable to the NIR emission (0.1 eV $\lesssim E \lesssim$ 1 eV), and both are subdominant to the UV peak (1 eV $\lesssim E \lesssim$ 20 eV). The lowest $\dot{M}_{\rm out}$ case does not exhibit a NIR peak at 10 Myr, but has significantly enhanced UV emission over the starburst-only model, driven by a two-order-of-magnitude increase in the mass flux reaching the disc interior.

Over time, increased mass flow to the disc interior drives enhanced star formation rates and sBH seeding, strengthening emission in the NIR. By 100 Myr the NIR bump in the three highest $\dot{M}_{\rm out}$ cases is the dominant feature in their respective spectra – with peak emission nearly two orders of magnitude larger than the adjacent $\sim$ 1 μm dip. In contrast, the Elvis et al. (1994) data suggest that the depth of the 1 μm dip is generally a factor of $\sim$2 smaller than the NIR peak. While significant dispersion exists in the composite spectra, with variations up to an order of magnitude cited by Elvis et al. (1994), and dominant IR peaks consistent with our data have been seen in a handful of AGN (Brown et al. 2019), the degree of divergence between our results and composite spectra suggest that our fiducial model needs to be modified to accommodate a broader variety of AGN.

There are several ways the IR and amplitude can be reduced. First, as Fig. 16 shows, if star formation ceased after $\sim$25 Myr, due to feedback or lack of fuel, then the IR bump would not grow to values well above the UV emission. Alternatively, if the outer disc size was truncated, eliminating much of the star- and BH-forming outer discs, it would similarly diminish the IR bump (as also noted by Sirko & Goodman 2003). Finally, if the angular momentum transport was more rapid than we adopted in the α-prescription, for example due to global gravitational torques, then, for fixed boundary conditions, the surface density and the IR amplitude would be lower. Moreover, reducing the number of sBHs necessary to support the disc also suppresses the excess X-ray emission from sBHs. We examine this final dependency as well as spectral dependence on SMBH mass in the next section.

8 VISCOSITY AND SMBH MASS DEPENDENCE

The spectra predicted by both the single source (Section 6) and evolving (Section 7) disc models diverge from typically observed AGN in three significant ways. First, the maximum mass flux reaching the central SMBH is $\lesssim$ 0.1 $\dot{M}_{\rm Edd}$ with a corresponding maximum UV luminosity of $\lesssim 10^{-2}\, L_{\rm Edd}$. As noted in Section 6.4.1, such low accretion rates and luminosities are characteristic of the lower end of observed values and highlight the limitations of these models in reproducing the full range of AGN accretion rates and luminosities. Secondly, the heating required to stabilize the outer disc generally tends to raise $T_{\rm eff}$ and produce IR emission well in excess of a standard α-disc. In a star-only supported disc, the resulting IR bump is roughly consistent with typical quasar spectra, but when sBH accretion feedback is included, the strength of the predicted IR bump exceeds even the most generous observational limits. Thirdly, we predict an emission excess at $\sim$10 keV due to embedded sBHs in the outer disc, inconsistent with commonly observed AGN spectra. This detection limit puts a tight constraint on the number of sBHs expected in the disc as discussed in Sections 6.4.3 and 7.3.

These discrepancies can be mitigated by increasing the disc viscosity. As noted in the previous section, higher viscosity reduces the surface density for a given $\dot{M}$, which in turn lowers the auxiliary heating necessary to stabilize the disc, damping the escaped X-ray flux from sBHs and the thermal IR emission. Additionally, lowering heating requirements lowers the corresponding rate of mass consumption by the embedded population and allows a greater mass flux to reach the disc interior.

Viscosity prescriptions beyond the standard α-disc have been invoked by previous works to explain high rates of mass transport in luminous AGN (Shlosman & Begelman 1989; Shlosman et al. 1990; Goodman 2003) and by TQM05 in ULIRGs. These works assume the presence of large-scale asymmetries (e.g. bars, spiral arms, or magnetic stresses), which induce rapid gas inflow. In Goodman (2003) and TQM05, these global torques were incorporated into steady state disc equations under the assumption that enhanced radial velocity should approach a constant fraction m of the sound speed, i.e. $V_{\rm r} = mc_{\rm s}$, and equation (3) is replaced by (Goodman 2003, TQM05):

$$\dot{M} = 4\pi m \Omega h^2 \rho r\,. \tag{39}$$

This simple prescription is derived from the well-known parametrization of spiral density waves (Binney & Tremaine 2008), typically applied to galactic discs. However, non-axisymmetric structures are expected to form in discs with masses exceeding $\gtrsim 0.1\ M_{\rm bh}$ (Hopkins & Quataert 2011). These structures may also be initiated and amplified by density perturbations due to ongoing star formation as well larger-scale density perturbations in the surrounding galaxy (Shlosman, Frank & Begelman 1989).

To illustrate the effect of enhanced viscosity by global torques on our disc models, we incorporate equation (39) into our single source, star-only and sBH-only models. As in Section 6, we assume Milky Way-like parameters, $M = 4 \times 10^6$ M$_\odot$, $\sigma = 75$ km s^{-1}, and $R_{\rm out} \simeq 3$ pc. Fig. 17 presents the resulting gas mass flux and mass consumption by stars (left, top panel) and sBHs (right, top panel) for $\dot{M}_{\rm out}$ ranging from 0.1 to 64 $\dot{M}_{\rm Edd}$. The corresponding emergent SEDs are shown in the bottom panels.

In the starburst-only models, the mass flux reaching the inner disc achieves a maximum value of $\simeq 6\,\dot{M}_{\rm Edd}$ when $\dot{M}_{\rm out} = 16$ and $64\,\dot{M}_{\rm Edd}$, nearly two orders of magnitude larger than for the discs shown in Section 6. The emerging spectra are bright and UV-dominant, with UV emission exceeding 10^{44} erg s^{-1}.

For mass supply rates $\lesssim 6\ \dot{M}_{\rm Edd}$, star formation is suppressed, resulting in a nearly constant mass flux across the disc. The resulting spectra exhibit weak IR emission relative to the UV-bump. For $\dot{M}_{\rm out} = 0.1\ \dot{M}_{\rm Edd}$, the low star formation rate and constant mass flux starkly contrast with the α-viscosity disc under identical boundary conditions, where the mass flux is reduced by three orders of magnitude. In the disc presented here, a significant fraction of the disc mass is not consumed by stars because the temperature at the outer edge of the disc is lowered to $T \simeq 50$ K. In this regime, opacity becomes less sensitive to temperature, scaling as $\kappa \propto T^{\beta}$, where $\beta < 2$. In the radiation pressure dominated outer disc, the scaling

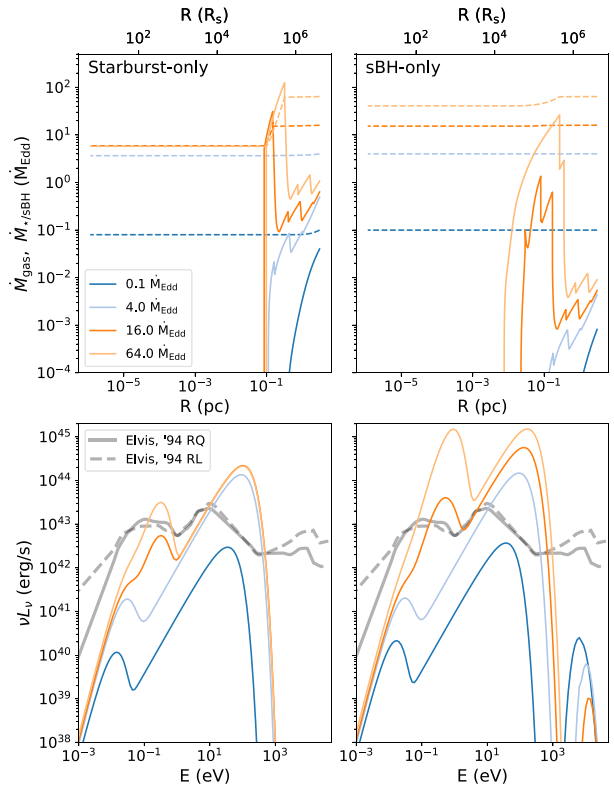

Figure 17. Top panels: Gas disc mass flux (solid lines) and mass consumption by stars (dashed lines, left panel) and sBHs (dashed lines, right panel) for single source AGN disc models with central SMBH mass of 4×10^6 M$_\odot$ and assuming disc viscosity enhanced by global torques where $m = 0.2$. Mass supply rates are distinguished by colour: 1 $\dot{M}_{\rm Edd}$ (dark blue), 8 $\dot{M}_{\rm Edd}$ (light blue), 16 $\dot{M}_{\rm Edd}$ (dark orange), and 32 $\dot{M}_{\rm Edd}$ (light orange). Bottom panels: SEDs emerging from single source, star-only (left) and sBH-only (right), discs plotted in corresponding top panels. The Elvis et al. (1994) data are plotted in grey and normalized such that the strength of UV band emission (1 eV $\leq E \leq 10^2$ eV) is $L_{\rm UV} = L_{\rm Edd} \simeq 5.8 \times 10^{44}$ erg s^{-1}.

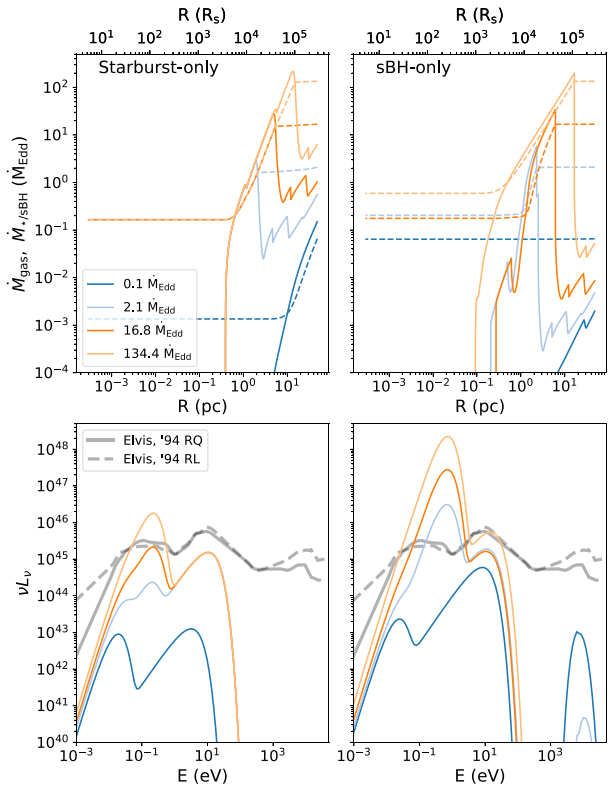

Figure 18. Assuming an SMBH mass of $M = 10^9$ M$_\odot$, we plot mass flux and emergent SEDs for varying $\dot{M}_{\rm out}$, as in Fig. 17. We take $\sigma = 300$ km s^{-1}, $R_{\rm out} = 48$ pc. In the bottom panels, we again scale the Elvis et al. (1994) data so that $L_{\rm UV} = L_{\rm Edd} \simeq 1.4 \times 10^{47}$ erg s^{-1}.

relations $\dot{\Sigma}_\star \propto \Sigma/\kappa$ and $T \propto \Sigma^{1/2}$ imply $\dot{\Sigma}_\star \propto \Sigma^{1-\beta/2}$, that is $\dot{\Sigma}_\star$ scales with Σ. The temperature of the α-viscosity disc with the same boundary conditions is $\simeq 100$ K at $R_{\rm out}$ and $\beta = 2$, making $\dot{\Sigma}_\star$ independent of Σ. In these cases, the critical mass supply rate $\dot{M}_{\rm c} \simeq 2\pi R^2 \dot{\Sigma}_\star \simeq 0.26\ \dot{M}_{\rm Edd}$ sets the threshold between discs that experience significant depletion and those that do not.

In the sBH-only cases, significant mass consumption occurs only for $\dot{M}_{\rm out} = 64\ \dot{M}_{\rm Edd}$, where the inner disc mass flux is 40 $\dot{M}_{\rm Edd}$, reduced by a factor of approximately $\sim 2/3$. In this regime, accretion heating from sBHs produces strong IR heating, comparable to the UV bump and about an order of magnitude larger than emission at the 1 μm dip. For all other cases, the mass flux through the disc remains nearly constant, so that the UV luminosities scale with $\dot{M}_{\rm out}$, and IR emission is a subdominant component of the spectra. The sBH contribution to the X-ray band is also significantly reduced. From lowest to highest $\dot{M}_{\rm out}$, the ratios of escaped sBH X-ray emission to bolometric luminosity are 2.7×10^{-3}, 1.4×10^{-5}, 4.9×10^{-7}, and 1.0×10^{-8}. As in the α-viscosity models, this ratio maximized by lowering the mass supply rate. Only in the lowest $\dot{M}_{\rm out}$ case does the sBH luminosity fraction exceed 10^{-3} $L_{\rm bol}$ – the rough detection limit discussed in Section 6.4.3.

In Fig. 18, we examine the effects of global torques for more commonly observed, high-mass AGN. We take $M = 10^9$ M$_\odot$, $R_{\rm out} = 48$ pc, and $\sigma = 300$ km s^{-1}. We use mass supply rates of 0.1, 2.1, 16.8, and 134.4 $\dot{M}_{\rm Edd}$, or 0.25, 8.0, 64, and 512 $\dot{M}_{\rm c}$, where $\dot{M}_{\rm c} = 6.7$ M$_\odot$ yr^{-1}. The resulting behaviour closely resembles that of α-disc models, where a 'depleted state' emerges for subcritical mass supply rates. Beyond this threshold, the mass flux stabilizes at just above 0.1 $\dot{M}_{\rm Edd}$ (or ~ 4 M$_\odot$ yr^{-1}). Constraints on mass feeding are also discussed in TQM05, where disc accretion rates are found to be similarly limited. This results in a maximum disc luminosity that is nearly an order of magnitude too faint to match the most luminous observed AGN. In the sBH-dominated models, UV emission is slightly enhanced at the highest mass supply rates, but this enhancement comes at the cost of anomalously high IR emission, creating an inconsistency with observed SEDs.

These findings suggest that an additional mechanism is required to explain the spectra of typical AGN discs, one that avoids the apparent trade-off between UV and IR emission inherent in these models. Potential candidates include enhanced viscosity in luminous AGN, driven by magnetic effects, or kinetic pressure from super-Eddington winds. Further exploration of these mechanisms could provide a more comprehensive framework for understanding the observed properties of AGN discs, particularly in the high-luminosity regime.

9 SUMMARY AND CONCLUSIONS

In this work, we used a semi-analytical approach to model the dual feedback mechanisms of sBH accretion and star formation in AGN discs. We develop a set of 1D, steady-state equations that incorporate a population of embedded sBHs seeded by preceding generations of star formation. In this way, we formulate an implicitly

time-dependent system, through which we can approximate the evolving structure of the AGN disc and its spectral signatures.

We assume that radiation from embedded stars and sBHs heat the disc uniformly. As in TQM05, we expect that the UV radiation from massive stars is absorbed and scattered by dust grains which reprocess the UV emission in the IR. For this reason, the disc is always considered optically thick to UV emission. We therefore weigh the stellar mass to luminosity conversion efficiency by the UV flux of the stellar population.

Remnant sBHs, on the other hand, are expected to emit largely in the X-ray band. These high-energy photons interact with the neutral disc via ionization, dust absorption and scattering, and Compton scattering. The resulting effective X-ray optical depth can be below unity, particularly in the outer disc where surface density is low. The escaping radiation does not contribute to the disc heating and reduces the effective mass to luminosity conversion efficiency of sBHs.

Under these assumptions and using parameters appropriate to the Galactic Centre, we construct steady-state AGN models supported by sBHs only and compared them to star-formation supported discs. These models are analogous to the 'pile-up' solutions described in Gilbaum & Stone (2022), but we account for the radial depletion of gas by sBH accretion and assume efficient heat mixing in the disc. We use these models to justify our assumption of a neutral gas disc, showing that the Strömgren sphere for an embedded sBH does not exceed the average distance between sBHs in the Toomre-Q unstable region of the disc. Moreover, we find that ionization bubbles around the sBHs can be well approximated as a step function in the disc interior – suggesting that AGN may maintain a two phase ionization structure at distances ≤ 0.17 pc from the SMBH.

Beginning with a starburst-supported disc, and assuming star formation seeds the sBH population, we compute a sequence of steady state models whose structure is determined by the growing population of accreting remnants. From our starburst-only and sBH-only models we can predict a time-scale over which we expect the disc to evolve. Prior to 25 Myr, our steady-state sequence models evolve consistently with these predicted time-scales. But as the sBH population grows, star formation in the disc interior is enhanced. This jump in star formation depletes the mass flux in the disc interior, resulting in a narrow ring of star formation that persists over the remaining lifetime of the disc.

We use our models to predict the effect of sBH heating on the emerging disc spectra, and identify tentative observational signatures, including enhanced emission in the NIR and the addition of a hard X-ray component. The latter comes from escaped emission from accreting sBHs and peaks at $\sim$10 keV – overlapping an X-ray bump observed in AGN spectra commonly referred to as the 'Compton hump' and attributed to reflected emission from the X-ray corona at the centre of the disc (Zoghbi et al. 2013, 2021; Kara et al. 2015). The apparent absence of an emission excess below the Compton hump suggests an upper limit on the size of the embedded sBH population. Moreover, we expect emission from sBHs to differ from the Compton reflection hump in several key ways. Specifically, the escaped X-ray flux comes primarily from sBHs embedded at distances greater than 0.2 pc from the SMBH, local to the outer disc. We also expect escaped X-ray flux to vary on long time-scales and be unpolarized, distinguishing it from the reflected coronal emission. Our models indicate that the relative strength and observability of this X-ray component is maximized at low-mass feeding rates and during the early phases of disc evolution, when the outer disc has been populated by sBHs, while the inner disc remains primarily supported by stars.

As demonstrated by both Sirko & Goodman (2003) and TQM05, the inclusion of heating from star formation in AGN disc models tends to produce IR emission well in excess of a standard steady Shakura–Sunyaev α-disc. Additionally, mass consumed in star formation reduces the gas mass flux reaching the disc interior, limiting the SMBH feeding rate and UV luminosity to sub-Eddington rates. sBH formation in the outer disc exacerbates this behaviour, because the sBHs allow a higher mass inflow rate into the inner regions of the disc, and the resulting larger surface density requires additional heating to stabilize these regions. As a result, over time, the IR bump becomes the dominant feature in our modelled spectra. Although precise comparison with observed SEDs is beyond the scope of this simplified model, the broad results are not consistent with the UV-dominated spectra typically seen in AGN.

There are several ways the IR amplitude can be reduced. Younger or radially truncated discs would see limited sBH formation, resulting in suppressed IR emission. Alternatively, more rapid angular momentum transport via global torques reduces the auxiliary heating necessary to support the disc, thereby reducing the IR bump. We examine the latter effect by incorporating enhanced viscosity into our fiducial, Milky Way-like model as well as for a more typically observed, high-mass AGN with an SMBH mass of 10^9 $M_\odot$. We find that using a viscosity prescription that incorporates global torques alleviates the IR strength, X-ray excess, and UV faintness for our fiducial, low-mass AGN models. However, for higher mass AGN, the anomalously high IR and faint UV emission persist, suggesting that our models are insufficient to explain the spectral shapes of luminous AGN. Resolving this will likely require the inclusion of additional physics, such as further enhanced viscosity driven by magnetic effects or non-thermal sources of pressure support including super-Eddington winds from embedded stars and sBHs.

While beyond the scope of the present paper, future work will aim to quantify these constraints. Any such analysis must also account for the contribution from circumnuclear tori to AGN IR emission, as inferred from time-lags between IR and UV variability. The upcoming GRAVITY + upgrade offers a promising avenue to disentangle the fraction of NIR emission intrinsic to AGN discs. GRAVITY has already enabled precise localization of NIR emission, as demonstrated by observations of NGC 1068, which revealed a disc-like structure rather than the traditionally assumed torus-like morphology (GRAVITY Collaboration 2020). The GRAVITY + upgrade will improve resolution by a factor of $\sim$2, expanding the number of AGN whose NIR and broad-line region (BLR) emission can be imaged and improving imaging of these regions in previously observed AGN. These advancements will enable a more detailed characterization of the spatial distribution of NIR emission, placing additional constraints on models involving heating by embedded sBHs.

As our paper was being finalized for submission, we became aware of a related preprint by Zhou et al. (2024). While their modelling approach is quite different, they discuss the modified emission from AGN discs due to the presence of embedded sBHs. Focusing on the optical spectrum near 5000 Å, they find that embedded sBHs increase $T_{\rm eff}$ in the outer disc and that the spatially extended heating from sBHs may be resolved in microlensing observations. These results are consistent with what we found, although our model includes many additional physical ingredients, and addresses the time-evolution of the system.

In this work, we focus on the contributions of star formation and consequent, *in situ* sBH accretion to stabilize AGN discs. To enhance the clarity of our results, we have chosen to simplify our model in a number of ways. First, we have assumed the disc

evolution can be treated as a sequence of steady states, and we neglect to account for the full hydrodynamic behaviour of the disc. We expect this simplification to have the greatest impact on models with $\dot{M}_{\rm out} < \dot{M}_{\rm crit} \simeq 0.13\ \dot{M}_{\rm Edd}$, in which the mass flux through the inner disc changes by several orders of magnitude. Secondly, we have neglected to incorporate the dynamical processes influencing the size and location of sBHs and stars in the disc. These processes include (1) angular momentum exchange between embedded objects and the gas disc, resulting in migration (Bellovary et al. 2016; Secunda et al. 2019, 2020; Tagawa et al. 2020). (2) Capture of stars and sBHs from the surrounding nuclear cluster (e.g. Bartos et al. 2017). (3) Dynamical interaction between embedded objects including scatterings and mergers. We expect to address both the hydrodynamical evolution of these discs and dynamical evolution of the stellar and sBH population embedded in them in future work.

Finally, we do not account for the growth of the sBHs or stars in these models. We show that the former will not change the qualitative conclusions drawn here, although we expect sBH growth to affect disc structure for AGN lifetimes exceeding 100 Myr. The latter, on the other hand, remains an interesting and open question. Recent work has suggested that stars embedded in AGN discs may evolve to be very massive, or perhaps even 'immortal', replenishing hydrogen gas via accretion at the same rate it is burned in the core (Cantiello et al. 2021; Jermyn et al. 2021, 2022). As discussed in Chen & Lin (2024), such a population would diminish the effects of remnants on disc structure and evolution. On the other hand, super solar abundances observed in AGN offer strong evidence in support of ongoing stellar evolution (Huang, Lin & Shields 2023). To account for this apparent discrepancy between theory and observation, Ali-Dib & Lin (2023) suggest that stellar lifetimes in AGN may be limited by gap formation in the disc, He accumulation in stars, and suppression of mixing by the radiative layer in stellar interiors. According to this 'Stellar Evolution and Pollution in AGN Discs' (SEPAD) model, star formation and evolution off the MS would lead to sufficient levels of heavy element disc pollution to account for observed abundances. Still, these models suggest a stellar IMF that may diverge significantly from field stars and vary based on disc location. This alternative stellar evolution model, and the remnant population it would create, offers another interesting avenue for future work.

To summarize, we find that incorporating sBH remnants into AGN disc models has significant impact on the disc structure, evolution, and corresponding spectra. We emphasize that star formation in AGN discs persists across the disc lifetime, although its radial extent becomes increasingly limited over time, confined to a narrow band at the edge of the opacity gap or $\simeq 0.2$ pc for parameters appropriate to the Galactic Centre, considered here. These processes predict a unique spectral AGN signature, including enhanced IR emission and an additional contribution to the X-ray band attributed to escaped flux from accreting sBHs. Neither of these features are commonly observed in AGN, but highlight observational constraints on AGN models that include auxiliary heating in order to stabilize the outer disc. This feature is likely reconcilable with observations once additional physics is incorporated into our models, which we defer to dedicated follow-up study.

ACKNOWLEDGEMENTS

We thank Semih Tuna, Erin Kara, Yuri Levin, Brian Metzger, Andrei Beloborodov, as well as an anonymous referee, whose thoughtful insights significantly improved this work. MEM has been supported by *GFSD*. ZH acknowledges financial support from NASA grants 80NSSC22K0822 and 80NSSC24K0440 and NSF grant AST-2006176. RP gratefully acknowledges support by NSF award AST-2006839. HT was supported by the National Key R&D Program of China (grant no. 2021YFC2203002).

Software: PANDAS (McKinney 2010), IPYTHON (Pérez & Granger 2007), MATPLOTLIB (Hunter 2007), SCIPY (Virtanen et al. 2020), NUMPY (Oliphant 2006), and JUPYTER (Kluyver et al. 2016).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Ali-Dib M., Lin D. N. C., 2023, MNRAS, 526, 5824
Antonini F., 2014, ApJ, 794, 106
Antonucci R., 1993, ARA&A, 31, 473
Bartos I., Kocsis B., Haiman Z., Márka S., 2017, ApJ, 835, 165
Baruteau C., Cuadra J., Lin D. N. C., 2011, ApJ, 726, 28
Baskin A., Laor A., 2018, MNRAS, 474, 1970
Bell K. R., Lin D. N. C., 1994, ApJ, 427, 987
Bellovary J. M., Mac Low M.-M., McKernan B., Ford K. E. S., 2016, ApJ, 819, L17
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton, NJ, p. 478
Blandford R. D., Begelman M. C., 1999, MNRAS, 303, L1
Boekholt T. C. N., Rowan C., Kocsis B., 2023, MNRAS, 518, 5653
Bondi H., 1952, MNRAS, 112, 195
Brown M. J. I., Duncan K. J., Landt H., Kirk M., Ricci C., Kamraj N., Salvato M., Ananna T., 2019, MNRAS, 489, 3351
Buzzoni A., 2002, AJ, 123, 1188
Caballero-Garcia M. D., Papadakis I. E., Nicastro F., Ajello M., 2012, A&A, 537, A87
Cantiello M., Jermyn A. S., Lin D. N. C., 2021, ApJ, 910, 94
Chen Y.-X., Lin D. N. C., 2024, ApJ, 967, 88
Collin S., Zahn J.-P., 1999, A&A, 344, 433
Czerny B., Elvis M., 1987, ApJ, 321, 305
De Marco B., Ponti G., Cappi M., Dadina M., Uttley P., Cackett E. M., Fabian A. C., Miniutti G., 2013, MNRAS, 431, 2441
DeLaurentiis S., Epstein-Martin M., Haiman Z., 2023, MNRAS, 523, 1126
Dittmann A. J., Cantiello M., Jermyn A. S., 2021, ApJ, 916, 48
Draine B. T., 2003, ApJ, 598, 1026
Elvis M. et al., 1994, ApJS, 95, 1
Fabian A. C., Rees M. J., Stella L., White N. E., 1989, MNRAS, 238, 729
Fabj G., Nasim S. S., Caban F., Ford K. E. S., McKernan B., Bellovary J. M., 2020, MNRAS, 499, 2608
GRAVITY Collaboration, 2020, A&A, 634, A1
GRAVITY Collaboration, 2022, A&A, 657, L12
Gangardt D., Trani A. A., Bonnerot C., Gerosa D., 2024, MNRAS, 530, 3689
Ghez A. M., Salim S., Hornstein S. D., Tanner A., Lu J. R., Morris M., Becklin E. E., Duchêne G., 2005, ApJ, 620, 744
Gilbaum S., Stone N. C., 2022, ApJ, 928, 191
Goodman J., 2003, MNRAS, 339, 937
Goodman J., Tan J. C., 2004, ApJ, 608, 108
Greene J. E., Ho L. C., 2007, ApJ, 670, 92
Grishin E., Gilbaum S., Stone N. C., 2024, MNRAS, 530, 2114
Haardt F., Madau P., 1996, ApJ, 461, 20
Haardt F., Maraschi L., 1991, ApJ, 380, L51
Hailey C. J., Mori K., Bauer F. E., Berkowitz M. E., Hong J., Hord B. J., 2018, Nature, 556, 70
Harrison F. A. et al., 2013, ApJ, 770, 103
Ho L. C., 2008, ARA&A, 46, 475
Hopkins P. F., Quataert E., 2011, MNRAS, 415, 1027
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Huang J., Lin D. N. C., Shields G., 2023, MNRAS, 525, 5702

Inayoshi K., Haiman Z., 2016, ApJ, 828, 110
Inayoshi K., Haiman Z., Ostriker J. P., 2016, MNRAS, 459, 3738
Jermyn A. S., Dittmann A. J., Cantiello M., Perna R., 2021, ApJ, 914, 105
Jermyn A. S., Dittmann A. J., McKernan B., Ford K. E. S., Cantiello M., 2022, ApJ, 929, 133
Kara E. et al., 2015, MNRAS, 449, 234
Kluyver T. et al., 2016, in Loizides F., Schmidt B., eds, Positioning and Power in Academic Publishing: Players, Agents and Agendas. IOS Press, p. 87, available at https://ebooks.iospress.nl/publication/42900
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Koshida S. et al., 2014, ApJ, 788, 159
Kramer R. H., Haiman Z., 2008, MNRAS, 385, 1561
Landt H. et al., 2019, MNRAS, 489, 1572
Landt H. et al., 2023, ApJ, 945, 62
Lasota J.-P., 2001, New A Rev., 45, 449
Leigh N. W. C. et al., 2018, MNRAS, 474, 5672
Levin Y., 2003, preprint (arXiv:astro-ph/0307084)
Levin Y., 2007, MNRAS, 374, 515
Li R., Lai D., 2024, MNRAS, 529, 348
Li Y.-P., Dempsey A. M., Li S., Li H., Li J., 2021, ApJ, 911, 124
Li J., Rodet L., Lai D., 2024, MNRAS, 528, 1198
Lu J. R., Do T., Ghez A. M., Morris M. R., Yelda S., Matthews K., 2013, ApJ, 764, 155
MacLeod M., Lin D. N. C., 2020, ApJ, 889, 94
Malkan M. A., 1983, ApJ, 268, 582
Malkan M. A., Sargent W. L. W., 1982, ApJ, 254, 22
Martini P., 2004, in Ho L. C., ed., Carnegie Observatories Astrophysics Series, Vol. 1: Coevolution of Black Holes and Galaxies. Cambridge Univ. Press, Cambridge, p.169
McKee C. F., Ostriker J. P., 1977, ApJ, 218, 148
McKernan B., Ford K. E. S., Lyra W., Perets H. B., 2012, MNRAS, 425, 460
McKinney W., 2010, in van der Walt S., Millman J., eds, Proc. 9th Python in Sci. Conf. Austin, Texas, p. 51, available at https://proceedings.scipy.org/articles/proceedings-2010
Miniutti G., Fabian A. C., 2004, MNRAS, 349, 1435
Murray C. D., Dermott S. F., 1999, Solar System Dynamics,Cambridge Univ. Press, Cambridge
Netzer H., 2015, ARA&A, 53, 365
Oliphant T. E., 2006, Guide to NumPy. Provo, UT, available at: http://www.tramy.us/
Osterbrock D. E., Ferland G. J., 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei. University Science Books, Mill Valley, p. 19
Panamarev T., Shukirgaliyev B., Meiron Y., Berczik P., Just A., Spurzem R., Omarov C., Vilkoviskij E., 2018, MNRAS, 476, 4224
Pérez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21
Podgorný J., Dovčiak M., Goosmann R., Marin F., Matt G., Różańska A., Karas V., 2023, MNRAS, 524, 3853
Ponti G., Papadakis I., Bianchi S., Guainazzi M., Matt G., Uttley P., Bonilla N. F., 2012, A&A, 542, A83
Raithel C. A., Sukhbold T., Özel F., 2018, ApJ, 856, 35
Rosenthal M. M., Chiang E. I., Ginzburg S., Murray-Clay R. A., 2020, MNRAS, 498, 2054
Rybicki G. B., Lightman A. P., 1986, Radiative Processes in Astrophysics. Wiley, Weinheim, p. 36
Salaris M., Cassisi S., 2005, Evolution of Stars and Stellar Populations. Wiley, West Sussex, England
Salpeter E. E., 1955, ApJ, 121, 161
Salpeter E. E., 1964, ApJ, 140, 796
Sanders D. B., Phinney E. S., Neugebauer G., Soifer B. T., Matthews K., 1989, ApJ, 347, 29
Secunda A., Bellovary J., Mac Low M.-M., Ford K. E. S., McKernan B., Leigh N. W. C., Lyra W., Sándor Z., 2019, ApJ, 878, 85
Secunda A. et al., 2020, ApJ, 903, 133
Semenov D., Henning T., Helling C., Ilgner M., Sedlmayr E., 2003, A&A, 410, 611
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shapiro S. L., Lightman A. P., Eardley D. M., 1976, ApJ, 204, 187
Shields G. A., 1978, Nature, 272, 706
Shlosman I., Begelman M. C., 1989, ApJ, 341, 685
Shlosman I., Frank J., Begelman M. C., 1989, Nature, 338, 45
Shlosman I., Begelman M. C., Frank J., 1990, Nature, 345, 679
Sirko E., Goodman J., 2003, MNRAS, 341, 501
Soldi S. et al., 2014, A&A, 563, A57
Spera M., Mapelli M., 2017, MNRAS, 470, 4739
Spera M., Mapelli M., Bressan A., 2015, MNRAS, 451, 4086
Stone N. C., Metzger B. D., Haiman Z., 2017, MNRAS, 464, 946
Suganuma M. et al., 2006, ApJ, 639, 46
Sunyaev R. A., Churazov E. M., 1996, Astron. Lett., 22, 648
Sunyaev R. A., Titarchuk L. G., 1980, A&A, 86, 121
Syer D., Clarke C. J., Rees M. J., 1991, MNRAS, 250, 505
Tagawa H., Haiman Z., Kocsis B., 2020, ApJ, 898, 25
Tagawa H., Kocsis B., Haiman Z., Bartos I., Omukai K., Samsing J., 2021a, ApJ, 907, L20
Tagawa H., Kocsis B., Haiman Z., Bartos I., Omukai K., Samsing J., 2021b, ApJ, 908, 194
Tagawa H., Kimura S. S., Haiman Z., Perna R., Tanaka H., Bartos I., 2022, ApJ, 927, 41
Thompson T. A., Quataert E., Murray N., 2005, ApJ, 630, 167(TQM05)
Tie S. S., Kochanek C. S., 2018, MNRAS, 473, 80
Toomre A., 1964, ApJ, 139, 1217
Toyouchi D., Hosokawa T., Sugimura K., Nakatani R., Kuiper R., 2019, MNRAS, 483, 2031
Ursini F. et al., 2020, A&A, 634, A92
Uttley P., Cackett E. M., Fabian A. C., Kara E., Wilkins D. R., 2014, A&AR, 22, 72
Virtanen P. et al., 2020, Nature Methods, 17, 261
Yang Y. et al., 2019, Phys. Rev. Lett., 123, 181101
Zhou S., Sun M., Liu T., Wang J.-M., Wang J.-X., Xue Y., 2024, ApJ, 966, L9
Zoghbi A., Reynolds C., Cackett E. M., Miniutti G., Kara E., Fabian A. C., 2013, ApJ, 767, 121
Zoghbi A., Miller J. M., Cackett E., 2021, ApJ, 912, 42

This paper has been typeset from a TEX/LATEX file prepared by the author.