Why do black holes trace bulges (& central surface densities), instead of galaxies as a whole?

Philip F. Hopkins [ID],[1]★ Sarah Wellons [ID],[2] Daniel Anglés-Alcázar,[3] Claude-André Faucher-Giguère[2] and Michael Y. Grudić[2]

[1]*TAPIR, Mailcode 350-17, California Institute of Technology, Pasadena, CA 91125, USA*
[2]*CIERA and Department of Physics and Astronomy, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA*
[3]*Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, CT 06269-3046, USA*

ABSTRACT

Previous studies of fueling black holes in galactic nuclei have argued (on scales $\sim 0.01-1000$ pc) accretion is dynamical with inflow rates $\dot{M} \sim \eta \, M_{\rm gas}/t_{\rm dyn}$ in terms of gas mass $M_{\rm gas}$, dynamical time $t_{\rm dyn}$, and some η. But these models generally neglected expulsion of gas by stellar feedback, or considered extremely high densities where expulsion is inefficient. Studies of star formation, however, have shown on sub-kpc scales the expulsion efficiency $f_{\rm wind} = M_{\rm ejected}/M_{\rm total}$ scales with the gravitational acceleration as $(1 - f_{\rm wind})/f_{\rm wind} \sim \bar{a}_{\rm grav}/\langle \dot{p}/m_* \rangle \sim \Sigma_{\rm eff}/\Sigma_{\rm crit}$ where $\bar{a}_{\rm grav} \equiv G \, M_{\rm tot}(< r)/r^2$ and $\langle \dot{p}/m_* \rangle$ is the momentum injection rate from young stars. Adopting this as the simplest correction for stellar feedback, $\eta \rightarrow \eta \, (1 - f_{\rm wind})$, we show this provides a more accurate description of simulations with stellar feedback at low densities. This has immediate consequences, predicting the slope and normalization of the $M_{\rm BH} - \sigma$ and $M_{\rm BH} - M_{\rm bulge}$ relation, $L_{\rm AGN} -$SFR relations, and explanations for outliers in compact Es. Most strikingly, because star formation simulations show expulsion is efficient ($f_{\rm wind} \sim 1$) below total-mass surface density $M_{\rm tot}/\pi \, r^2 < \Sigma_{\rm crit} \sim 3 \times 10^9 \, {\rm M_\odot \, kpc^{-2}}$ (where $\Sigma_{\rm crit} = \langle \dot{p}/m_* \rangle/(\pi \, G)$), BH mass is predicted to specifically trace host galaxy properties above a critical surface brightness $\Sigma_{\rm crit}$ (B-band $\mu_{\rm B}^{\rm crit} \sim 19 \, {\rm mag \, arcsec^{-2}}$). This naturally explains why BH masses preferentially reflect bulge properties or central surface densities (e.g. $\Sigma_{\rm 1 \, kpc}$), not 'total' galaxy properties.

Key words: accretion, accretion discs – galaxies: active – galaxies: evolution – galaxies: formation – quasars: general – quasars: supermassive black holes.

1 INTRODUCTION

Understanding the origins, growth and evolution of super-massive black holes (BHs) remains one of the most important unsolved problems in extragalactic astrophysics. It is now well established that most sufficiently-massive galaxies host BHs whose masses correlate with various host galaxy bulge properties (Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Aller & Richstone 2007; Hopkins et al. 2007b; Kormendy, Bender & Cornell 2011; for a review see Kormendy & Ho 2013). The small scatter in these correlations (relative to other galaxy properties; Hopkins, Murray & Thompson 2009b), together with constraints indicating that most BH mass is assembled in an optically bright quasar phase (Soltan 1982; Salucci et al. 1999; Yu & Tremaine 2002; Hopkins, Narayan & Hernquist 2006b), suggests a picture of 'coevolution' between galaxies and accreting BHs visible as active galactic nuclei (AGN) or quasars (Merloni & Heinz 2008). Understanding this 'coevolution' has far-reaching consequences beyond the BHs themselves: for example, it is widely believed that 'feedback' from accreting BHs (in the form of radiation, winds, and jets; Laor et al. 1997; Crenshaw et al. 2000; Dunn et al. 2010; Sturm et al. 2011; Zakamska et al.

2016; Williams et al. 2017) can unbind, expel, or super-heat gas in the vicinity of the BH and throughout the host galaxy (Silk & Rees 1998; King 2003; Di Matteo, Springel & Hernquist 2005; Hopkins et al. 2005a, b; Murray, Quataert & Thompson 2005; Debuhr et al. 2010; Torrey et al. 2020), potentially regulating star formation and galaxy stellar masses (Croton et al. 2006; Hopkins et al. 2006a, 2008) and the structure of the circum-galactic medium around massive galaxies (Ciotti & Ostriker 1997; Cox et al. 2006; Best et al. 2007; Voit et al. 2017).

But modeling the strength of 'feedback' from SMBHs, and their presence in the first place, depends fundamentally on understanding their accretion rates. In understanding how gas is transported from the inter-galactic medium on to BHs, it is especially important to understand, both empirically and theoretically, how gas is transported from scales $\sim 0.1-1000$ pc within the galaxy (where its angular momentum is $\sim 10^7$ times too large to be accreted by the BH directly) into the BH accretion disc (scales $\lesssim 0.01$ pc). These scales include the observational and numerical resolution limits of essentially all resolved galaxy surveys and/or galaxy-scale numerical simulations (Schartmann et al. 2010; Fabian 2012; Hopkins et al. 2014a; Naab & Ostriker 2017; Davé et al. 2019) – so in both empirical and theoretical studies of AGN 'fueling' and its relation to galaxy properties, these are the key scales one wishes to relate to the AGN accretion rate. Moreover, neither well-understood galaxy-scale angular momentum

★ E-mail: phopkins@caltech.edu

transport mechanisms (mergers, galaxy-scale arms/bars), nor well-understood traditional accretion-disc processes (e.g. the MRI and turbulent/viscous stresses), can operate efficiently over most of these scales (especially from ~0.01–10 pc, within the BH radius of influence), leading to one of several 'last parsec problems' (Goodman 2003; Jiang & Goodman 2011). Moreover the assumptions of the classical Bondi–Hoyle (Bondi & Hoyle 1944) or Shakura & Sunyaev (1973) type accretion models are violated by many orders of magnitude on these scales: gas within a galaxy is rapidly cooling ($t_{cool} \ll t_{freefall}$), self gravitating, star forming, turbulent, must lose most of its angular momentum to efficient torques to be accreted, and the potential is dominated by a combination of gas, collisionless stars and dark matter, and the BH itself (Hopkins & Quataert 2010b, 2011b; Angles-Alcazar et al. 2020).

Empirically, it is clear that the best galactic *predictors* of BH mass on these scales are the velocity dispersion and/or stellar mass of the central classical 'bulge,' or nuclear star cluster (NSC) in late-type dwarf galaxies which exhibit no classical bulge, as opposed to e.g. *total* galaxy stellar or disc or halo mass or luminosity or circular velocity (e.g. Mancini & Feoli 2012; Kormendy & Ho 2013; Reines & Volonteri 2015). But this itself presents an important theoretical puzzle, related to the question above of what physics actually drives accretion on these scales. Almost all theoretical models to date of BH mass growth via pure accretion (i.e. 'fueling-limited' models), hierarchical assembly (e.g. BH growth primarily via mergers), and/or self-regulation via feedback (i.e. 'feedback-regulated' models) predict correlations between BH mass and 'gas supply in the galaxy center' or 'depth of the potential' in which the BH sits or 'mass assembled via mergers' (e.g. Silk & Rees 1998; King 2003; Di Matteo et al. 2005; Hopkins et al. 2007a; Peng 2007). These models commonly assume that these properties correlate closely with 'bulge' or NSC mass, but that is not correct in galaxies that are not bulge-dominated.

For example, in almost all galaxies of Sa or later type (including the Milky Way), the bulge *does not* dominate the central potential, relative to either the stellar+gas disc/entire galaxy or the dark matter: this can be seen from simple comparison of $G\,M_{bulge}/R_{bulge}$ versus $G\,M_{disk}/R_{disk}$ and $G\,M_{halo}/R_{halo}$, or more detailed Jeans modeling (Aller & Richstone 2007; McMillan 2017; Taranu et al. 2017). The discrepancy can be orders of magnitude in dwarfs.[1] However, if one considers the potential gradients, i.e. gravitational *acceleration* provided by these components ($\propto G\,M/R^2$), then the bulge often does dominate in the region between the SMBH radius of influence and the outskirts of the bulge – a crucial difference to which we will return below.

The 'gas supply to the galaxy center' is also not particularly well correlated with the bulge mass: nuclear bulge/cluster/disc shapes/densities/masses/radii vary wildly (Ferrarese et al. 1994; Lauer et al. 2007b, 2002; Läsker et al. 2016; Savorgnan & Graham 2016). So there is no reason, in most models for BH growth, why BHs would correlate particularly well with the 'central mass' within an arbitrarily varying annulus $<R$ that happens to correspond to the 'bulge' size. Yet Hopkins et al. (2009b) argued that the total

bulge/NSC mass was, in practice, often a better predictor of M_{BH} compared to e.g. mass within a fixed physical annulus or multiple of the BH 'radius of influence' $R_{ROI} \sim G\,M_{BH}/\sigma^2$. And in dwarfs, the observed SMBHs/AGN and their associated light excess/'bulge' are in fact most often *not* located near the center of mass or center of light of the galaxy (Reines et al. 2020), if such a center can even be defined (it often cannot at $\lesssim$ kpc scales). Regarding 'mass assembled by mergers,' it is increasingly clear that in sub-L_* galaxies galaxy–galaxy mergers play a minor/secondary role in bulge formation (Courteau, de Jong & Broeils 1996; Governato et al. 2010; Hopkins et al. 2010b,c; Puech et al. 2012; Pillepich, Madau & Mayer 2015); even if they do, most of the 'incoming' mass associated with such mergers ends up in an extended halo, rather than a compact bulge, and produces relatively little contribution to BH growth/AGN activity in dwarfs, $\sim L_*$, or Seyfert galaxies (Alexander & Hickox 2012; Kocevski et al. 2012; Heckman & Best 2014; Hopkins, Kocevski & Bundy 2014b). Finally, *none* of the theoretical models described above explain why BHs would correlate more poorly with 'pseudo bulges' and 'nuclear discs' as defined photometrically following Kormendy & Kennicutt (2004), as compared to 'classical' photometric bulges.

In this letter, we combine qualitative scalings common to many of the accretion models described above with a simple correction, generally neglected in simulation prescriptions, for the mass fraction expelled by *stellar* feedback from star formation on sub-kpc scales ('between' the simulation-resolved scales and accretion disc), and show that this provides an immediate and natural resolution to the questions above.

2 THEORY

The problem of accretion from sub-kpc scales described in Section 1 has been studied in detail in many papers, for example the series by Hopkins & Quataert (2010a, b, 2011a, b), subsequently explored further in other work (e.g. Hopkins et al. 2012a; Anglés-Alcázar, Özel & Davé 2013; Anglés-Alcázar et al. 2017a, b; Davé et al. 2019; Thomas et al. 2019; Angles-Alcazar et al. 2020, and others discussed below). These studies generically showed that on these scales, accretion is regulated by 'gravitational torques' from a combination of asymmetries in the potential, interactions between the collisionless (stars + dark matter) and gas components, and shocks/dissipation in the gas, giving rise to an accretion rate of the form:

$$\dot{M}_{acc} = \eta\,M_{gas}\,\Omega \sim \eta\,\frac{f_{gas}\,V_c^3}{G} \sim \eta\,\frac{4\pi\,G^2 M_{tot}^2 \rho}{V_c^3}, \qquad (1)$$

where $M_{gas} \approx \pi\,\Sigma_{gas}\,R^2$ is the gas mass within some annulus R, $\Omega = V_c/R$ is the dynamical frequency, and η is some relatively-weakly-varying function which describes the magnitude of whatever torques actually remove angular momentum and allow for accretion. For example, in the model from Hopkins & Quataert (2011b) $\eta \approx 0.01\,(M_S/M_d)^{1/6}\,[1 + 3\,M_{d,9}^{1/3}\,(M_{gas}/M_d)]^{-1} \sim 0.001$ where $M_S = M_{BH} + M_{\alpha\,disc}$ is the total 'sink' (BH + accretion disc) system mass, and $M_{d,9} \equiv M_d/10^9\,M_\odot$ with $M_d(< R)$ the total mass in a 'discy' (rotation-dominated) component. A number of subsequent, independent idealized theoretical studies (Kim, Seo & Kim 2012; Alig et al. 2013; Li, Shen & Emsellem et al. 2015; Kim 2015; Inayoshi et al. 2019) have validated the qualitative scaling above for similar assumptions, and detailed observations of galactic nuclei have appeared to confirm both the dominance of gravitational torques, and the approximate scaling of inflow rates with dynamical nuclear properties as predicted by these models (Combes et al. 2013; Esquej et al. 2014; García-Burillo et al. 2014; Querejeta et al. 2016). Broadly

---

[1] Using standard abundance-matching relations from Behroozi et al. (2019) and assuming Navarro, Frenk & White (1996) haloes, the central potential from the DM *alone* in sub-L_* (dwarf) galaxies scales as $\Phi(r \to 0) \sim (250\,\mathrm{km\,s^{-1}}\,[M_*/10^{10}\,M_\odot]^{1/6})^2$ – much larger than the potential from the bulge or NSC (or stellar disc), and very weakly dependent on stellar mass, while e.g. the BH and bulge/NSC mass scales super-linearly with stellar mass as $M_{BH} \propto M_{bulge/NSC} \propto M_*^{2-4}$ (Reines & Volonteri 2015; Graham & Scott 2015).

speaking, even quite different accretion models have arrived at scalings which qualitatively follow equation (1) on similar scales.[2]

However, all of these studies essentially neglected the possibility that gas would be efficiently expelled from the galactic nucleus by *stellar* feedback (e.g. radiation pressure, stellar mass loss, and SNe explosions), before it could accrete into the BH accretion disc. This includes models which treat stellar feedback as a 'sub-grid' process influencing the ISM but either not driving strong outflows or simply driving outflows with a by-hand fixed 'efficiency' $\dot{M}_{\rm out} \sim \dot{M}_*$, as well as those which neglect it entirely. A couple of subsequent studies (e.g. Wada, Papadopoulos & Spaans 2009; Hopkins et al. 2016; Angles-Alcazar et al. 2020; Kawakatu, Wada & Ichikawa 2020) have revisited this problem with simulations that explicitly include the relevant stellar feedback processes. However, these were simulations of nuclear discs intended to model extremely bright QSOs with enormous surface mass densities (or accelerations), $\Sigma_{\rm eff} \equiv M_{\rm tot}(< R)/\pi R^2 \gtrsim 10^5 \, {\rm M}_\odot \, {\rm pc}^{-2}$, where stellar feedback (even from vigorous SNe explosions) is unable to unbind large quantities of gas, and served primarily to 'thicken' the nuclear disc (potentially explaining features of the obscuring 'torus'; Wada & Norman 2002; Thompson et al. 2005).

Under less-extreme conditions, many theoretical (Wutschik, Schleicher & Palmer 2013; Torrey et al. 2017; Grudić et al. 2018, 2019a) and observational (Vollmer, Beckert & Davies 2008; Izumi, Kawakatu & Kohno 2016) studies have pointed out that stellar feedback can in principle easily expel most of the gas from galactic nuclei, dramatically suppressing accretion rates on to the BH. This can occur 'indirectly' or 'directly.' In the 'indirect' sense, efficient stellar feedback can, in a cosmological sense, lead to a given dark matter halo producing a much-less-massive, lower-density galaxy, which in turn produces sub-kpc conditions less conducive to BH growth (see discussion in e.g. Bower et al. 2017; Habouzit, Volonteri & Dubois 2017). These effects would therefore be implicit in the accretion models discussed above. But stellar feedback can also 'directly' restrict accretion through a given annulus in a galaxy given fixed larger-scale conditions, by ejecting some of that material in the annulus which would otherwise have lost its angular momentum (e.g. Dubois et al. 2015; Grudić et al. 2019a). The latter is the case of interest here. While such behaviour has been qualitatively observed in simulations, a simple quantitative parametrization of its effects is still lacking. Therefore consider: the simplest parameterzation of this effect is to take

$$\eta \to (1 - f_{\rm wind}) \, \eta, \qquad (2)$$

where $f_{\rm wind} \equiv M_{\rm ejected}/M_{\rm gas,\,total}$ represents the fraction of gas expelled by stellar feedback from within the annulus.

As shown in detail in Torrey et al. (2017), in galactic nuclei, the scalings for star formation and $f_{\rm wind}$ are essentially the same as in massive GMC complexes, as opposed to 'galactic' outflow/star formation models. This is fundamentally because on spatial scales $\sim 0.1-1000$ pc, the dynamical times $t_{\rm dyn} \sim 0.5 \, {\rm Myr} \, (R/100 \, {\rm pc}) (200 \, {\rm km \, s^{-1}}/V_c)$ are much shorter than the $t_{\rm fb} \sim 30-100 \, {\rm Myr}$ time-scales over which most stellar feedback is deposited. So gas flows in, converts to stars on some number of free-fall times (as in a 'single burst'), but the stars formed then rapidly expel gas from the central regions as they age and SNe begin to explode (akin to GMC destruction): no 'steady state' is possible when $t_{\rm dyn} \ll t_{\rm fb}$.

A simple analytic model for $f_{\rm wind}$ in this limit is given by Fall, Krumholz & Matzner (2010), as updated in Grudić et al. (2020): upon forming, a mass $M_{*,\,\rm young}$ of young stars ($\ll 100$ Myr old) within the nucleus in an area $A \sim \pi R^2$ will inject momentum into the surrounding gas (via feedback) at a rate:

$$\frac{{\rm d}\dot{P}_{\rm fb}}{{\rm d}A} \sim \langle \dot{p}/m_* \rangle \, \frac{M_{*,\,\rm young}}{A}, \qquad (3)$$

where

$$\langle \dot{p}/m_* \rangle \sim ({\rm a \ few}) \, \frac{L_*/c}{m_*} \sim 1000 \, \frac{L_\odot}{{\rm M}_\odot \, c} \sim 10^{-7} \, \frac{\rm cm}{\rm s^2}, \qquad (4)$$

is the momentum injection rate per stellar mass, for a well-sampled IMF.[3] This will expel the remaining gas when ${\rm d}\dot{P}_{\rm fb}/{\rm d}A$ exceeds the force per unit area on the gas from gravity:

$$\frac{{\rm dForce}_{\rm grav}}{{\rm d}A} \sim \bar{a}_{\rm grav} \, \frac{M_{\rm gas}}{A} \sim \frac{G \, M_{\rm tot}}{R^2} \, \frac{M_{\rm gas}}{R^2} \sim G \, \Sigma_{\rm eff} \, \Sigma_{\rm gas} \,, \qquad (5)$$

with

$$\bar{a}_{\rm grav} \equiv \frac{G \, M_{\rm tot}(< r)}{r^2}, \qquad (6)$$

$$\Sigma_{\rm eff} \equiv \frac{M_{\rm tot}(< r)}{\pi \, r^2}, \qquad (7)$$

defined inside a spherical annulus of radius r. Equating ${\rm dForce}_{\rm grav}/{\rm d}A$ (equation 5) and ${\rm d}\dot{P}_{\rm fb}/{\rm d}A$ (equation 3) and solving for $M_{\rm gas}$ to obtain the gas mass which can be expelled gives:

$$\frac{M_{\rm gas,\,expelled}}{M_{*,\,\rm young}} \sim \frac{M_{\rm ejected}}{M_{\rm retained}} = \frac{f_{\rm wind}}{1 - f_{\rm wind}} \sim \frac{\langle \dot{p}/m_* \rangle}{\bar{a}_{\rm grav}}, \qquad (8)$$

i.e.

$$1 - f_{\rm wind} \approx \frac{\bar{a}_{\rm grav}}{\langle \dot{p}/m_* \rangle + \bar{a}_{\rm grav}} = \frac{\Sigma_{\rm eff}}{\Sigma_{\rm crit} + \Sigma_{\rm eff}}, \qquad (9)$$

with

$$\Sigma_{\rm crit} = \frac{\langle \dot{p}/m_* \rangle}{\pi \, G} \sim 3000 \, \frac{{\rm M}_\odot}{\rm pc^2} = 3 \times 10^9 \, \frac{{\rm M}_\odot}{\rm kpc^2} \sim 0.6 \, \frac{\rm g}{\rm cm^2} \,. \qquad (10)$$

Because, in essentially all reasonable models on scales $\sim 1-1000$ pc, *most* of the retained mass goes into star formation rather than inflow to the BH, we can safely neglect the correction for inflow itself in this derivation of $f_{\rm wind}$.[4]

---

[2] For example, (1) assuming a constant accretion rate per free-fall time simply gives $\eta = $ constant, by definition. (2) The 'gravito-turbulent'-type models motivated by Gammie (2001), applied to star-forming discs with Toomre $Q \sim 1$ as in Thompson, Quataert & Murray (2005), Kawakatu & Wada (2008), Hopkins & Christiansen (2013) give $\eta \approx 0.1 \, (M_{\rm d}/M_{\rm tot})^2 \sim$ constant. (3) 'Ballistic accretion' (Hobbs et al. 2011) gives $\eta \approx (h/R)^{-1} \exp{(-0.6 \, R^2/h^2)}$ which is constant if the discs are thick ($h \sim R$) or we assume $h/R \sim$ constant, or scales similarly to 'gravitoturbulent' cases if we take $Q \sim$ constant. (4) A generalized version of the Shu (1977) self-similar scaling for a collapsing isothermal sphere, allowing for non-gas contributions to the potential and turbulence, gives $\eta \approx (1 + \Delta_v^2)^{-3/2}$ with $\Delta_v^2 \equiv (c_s^2 + \sigma_{\rm turb}^2/3 + |\langle \delta {\bf v} \rangle|^2)/V_c^2$ (with sound speed c_s, 3D gas velocity dispersion $\sigma_{\rm turb}$, and bulk BH-gas relative velocity $\delta {\bf v}$; see Hopkins et al. 2006a; Di Matteo et al. 2008). (5) The estimator in e.g. Hobbs et al. (2012) for 'Bondi-like' accretion in a halo (ignoring turbulence and relative motion) is simply this with $\eta = (1 + c_s^2/V_c^2)^{-3/2} \approx 1$.

[3] Crucially, the quantity $\langle \dot{p}/m_* \rangle$ for a 'young' (ZAMS or age $\lesssim 30$ Myr) is approximately independent of whether the dominant stellar feedback comes from radiation pressure, expanding HII regions, O/B winds, or SNe; see Leitherer et al. (1999), Bruzual & Charlot (2003), Hopkins, Quataert & Murray (2012b), Agertz et al. (2013), Kim & Ostriker (2015).

[4] More formally, there is some subtle ambiguity in equation (8) in how precisely to relate $M_{\rm retained}$, $M_{*,\,\rm young}$, and some (generally much smaller) mass accreted through an annulus in the same time, which can be addressed

Alternatively, adopting a continuum limit within each annulus as gas moves to the BH, we can revisit the derivation of equation (1) in Hopkins & Quataert (2011b). There, we solved a steady-state model calculating the strength of gravitational torques driving a *total* inflow rate within each annulus, coupled to the continuity equation, with $\dot{M}_{in}(R) = \dot{M}_{in}(R + dR) - \dot{M}_*(R < R' < R + dR) = \dot{M}_{in}(R + dR) - 2\pi\, R\, dR\, \dot{\Sigma}_*(R)$, i.e. accounting for gas lost to star formation within each annulus. If we modify this to also include gas lost in winds, then $\dot{\Sigma}_*(R) \rightarrow \dot{\Sigma}_* + \dot{\Sigma}_{wind} = [1 + \eta_{wind}(R)]\,\dot{\Sigma}_*$ where $\eta_{wind} \equiv \dot{\Sigma}_{wind}/\dot{\Sigma}_* \sim \langle \dot{p}/m_* \rangle / \bar{a}_{grav}$ within each annulus. While the exact solutions to this are, in general, numerical, simply taking $\eta \rightarrow \eta\,(1 - f_{wind})$ with f_{wind} from equation (9) provides a remarkably good approximation to the full solution, and is exact in small and large $\Sigma_{eff}(R)$ limits. Since Σ_{eff} in the analytic model increases monotonically as $R \rightarrow 0$, the 'loss' term f_{wind} is dominated by the largest radii, e.g. R where it is evaluated: for $\Sigma_{eff} \gg \Sigma_{crit}$, the exact solution is unmodified from Hopkins & Quataert (2011b), for $\Sigma_{eff} \ll \Sigma_{crit}$, it is multiplied by one power of $\Sigma_{eff}/\Sigma_{crit}$, as expected.

3 COMPARISON TO NUMERICAL SIMULATIONS

Equation (9) is actually remarkably well supported by both explicit numerical MHD simulations of GMC/star cluster/nuclear disc formation with explicit, resolved stellar feedback physics (Colín et al. 2013; Gavagnin et al. 2017; Geen et al. 2017; Grudić et al. 2018; Kim et al. 2018; Grudić & Hopkins 2019) as well as observations (Vollmer et al. 2008; Murray & Rahman 2010; Grudić et al. 2019b; Kruijssen et al. 2019), as shown in Fig. 1.[5] A wide range of different numerical codes, methods, and treatments of stellar feedback, including simulations of both GMCs as well as nuclear stellar discs support such a scaling. We compare e.g. an arbitrary variant dependence on $\Sigma_{eff}/\Sigma_{crit}$ fitting different simulations that give somewhat different detailed behaviour,[6] in order to illustrate that even with systematic differences in physics and numerical methods, different simulations predict a relation qualitatively similar to our simple order of magnitude estimate. These scalings and the results in Fig. 1 also immediately explain why the previous simulations of 'QSO-scale nuclear discs' discussed above, with $\Sigma_{eff} \sim 10^5\,M_\odot\,pc^{-2} \gg \Sigma_{crit}$ (so $1 - f_{wind} \approx 1$) saw essentially negligible effects on the accretion rate scaling (compared to equation 1) including explicit stellar feedback, while lower-resolution cosmological simulations of high-redshift, lower-mass galaxies (primarily dwarfs in low-luminosity AGN phases), with $\Sigma_{eff} \ll \Sigma_{crit}$ at their resolution limits, found that stellar feedback tended to 'blow out' most of the gas ($1 - f_{wind} \ll 1$) before it could accrete, dramatically suppressing $\dot{M}_{BH}$ (Dubois et al. 2015; Habouzit et al. 2017).

Fig. 2 tests this explicitly in high-resolution numerical simulations of inflow rates into the central $\ll 1$ pc around a BH from $\sim 1{-}1000$ pc radii discs, including detailed stellar feedback models identical to the

more accurately with the continuum limit models discussed below. But since this is simply an order-of-magnitude argument and the behaviour is identical in the relevant limits we study below, we can neglect this ambiguity for now.
[5]Since these are idealized simulations, f_{wind} can be easily measured as the fraction of the initial gas mass which is entirely expelled.
[6]Note the difference between the somewhat-larger efficiency predicted by Kim et al. (2018) in Fig. 1 and other plotted cases owes in part to the fact that Kim et al. (2018) included only UV radiation as a stellar feedback mechanism, but also to more detailed numerical and methodological differences discussed in detail therein and in Hopkins & Grudić (2019), Grudić & Hopkins (2019).
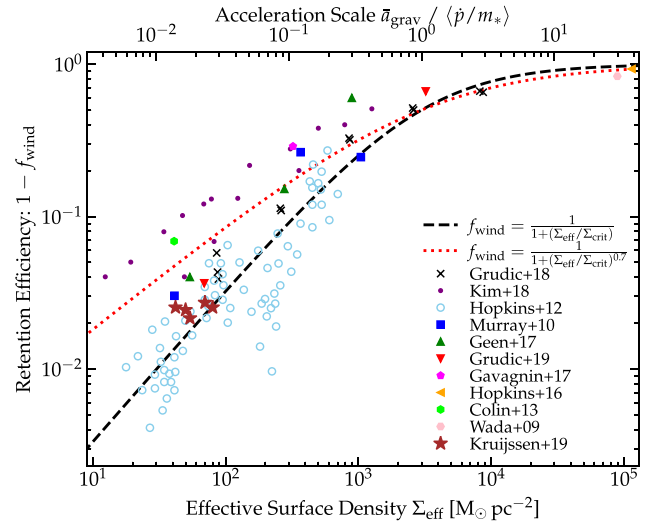

Figure 1. Scaling of the 'retention factor' or wind loss factor $f_{wind} \equiv M_{ejected}/M_{gas,\,total}$ measured in simulations and observations of $\lesssim 100$ pc-scale structures: simulations of molecular clouds (Colín, Vázquez-Semadeni & Gómez 2013; Gavagnin et al. 2017; Geen, Soler & Hennebelle 2017; Kim, Kim & Ostriker 2018; Grudić & Hopkins 2019), simulations including galactic nuclei & discs (Hopkins et al. 2012b; Grudić et al. 2018), circum-BH disc simulations (Wada et al. 2009; Hopkins et al. 2016), and observed GMCs (Murray & Rahman 2010) and galactic nuclei (Kruijssen et al. 2019). We compare the simple predicted theoretical scaling from Fall et al. (2010), Grudić et al. (2020) (equation 9; $f_{wind}^{-1} = 1 + (\Sigma_{eff}/\Sigma_{crit}) = 1 + \bar{a}_{grav}/\langle \dot{p}/m_* \rangle$), and a slight (arbitrary) variant fit ($f_{wind}^{-1} = 1 + (\Sigma_{eff}/\Sigma_{crit})^{0.7}$) which illustrates the theoretical uncertainties.
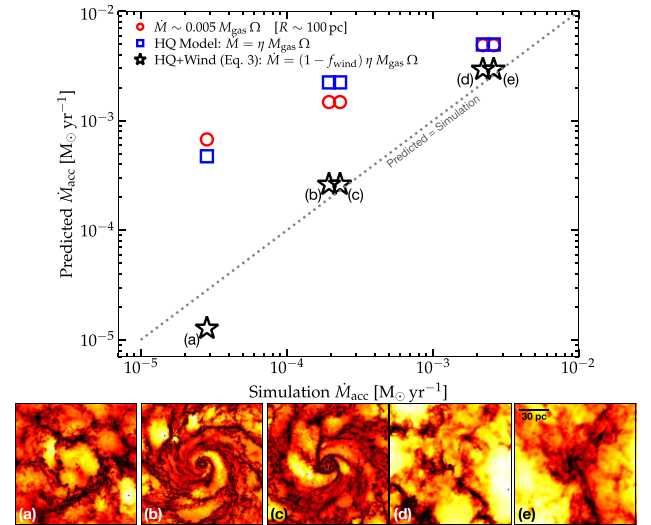

Figure 2. Direct test of different BHAR estimators in 5 simulations: we simulate a BH surrounded by an exponential gas + stellar disc (scale-length 100 pc) with star formation and stellar feedback as in Hopkins et al. (2016), with initial masses $(M_{BH}, M_{gas}, M_{*,\,disc})/M_\odot$ of (1e5, 5e6, 5e6) (a), (1e7, 5e6, 5e6) (b+c), (1e7, 5e8, 5e8) (d+e), each run for ~ 10 dynamical times. Models b, c and d, e differ in the initial value of $Q = 2.5,\,0.5$, respectively, given to the disc. We measure $\dot{M}_{acc}$ as resolved gravitational capture of bound gas within < 0.1 pc (averaged over the simulation duration), and compare to the predicted $\dot{M}_{acc}$ from the reference models, evaluating M_{gas}, Ω, Σ_{eff}, etc. at $R = 100$ pc and $t = 0$. We compare: $\eta = 0.005 = $ constant; 'HQ,' the Hopkins & Quataert 2011b model for η in equation (1); and 'HQ + Wind,' our proposed correction to this taking f_{wind} from equation (9). Images show a gas density projection for each run [scale bar labeled in (e)]; at the lowest Σ_{eff} (a) we directly see stellar feedback evacuating the nuclear region.

no-AGN-feedback ('No_BAL') simulations in Hopkins et al. (2016). We repeat their previous simulations with the same code and physics: the only difference is that we rescale the initial nuclear disc and BH masses such that $\Sigma_{\rm eff}$ ranges from $\sim 10^2 - 10^4\,{\rm M}_\odot\,{\rm pc}^{-2}$, while they considered only a case with $\Sigma_{\rm eff} \sim 10^5\,{\rm M}_\odot\,{\rm pc}^{-2}$. We compare the true accretion rates predicted by the high resolution sims, to the accretion rate that would be inferred by an analytic estimator based on the global simulation initial conditions or as a sub-grid model in a simulation at a lower, more typical resolution. As expected, assuming $f_{\rm wind} \approx 0$ (i.e. taking equation 1 without modification) works increasingly well at the highest $\Sigma_{\rm eff}$, corresponding to the highest-$\dot{M}$ cases here. But at lower $\Sigma_{\rm eff}$ and $\dot{M}$, ignoring this term leads to order of magnitude or more over estimation of $\dot{M}_{\rm acc}$, while incorporating the simple $1 - f_{\rm wind}$ scaling predicted by equation (9) provides a remarkably good fit to the full simulation results (despite very different inflow structures in the different regimes; see Hopkins & Quataert 2010b).

Briefly, we note in applications of equation (9) in simulations which do explicitly include stellar feedback, that since BH accretion rates are generally evaluated in some resolution-scale kernel around the BH, one should evaluate and apply the $f_{\rm wind}$ correction within the approximately the same kernel, since that is precisely the scale where (by definition) explicit stellar feedback will cease to be resolved. But some care is needed and (like with any sub-grid model) the range of applicable scales is finite. If, for example, the unresolved region is so small that the radial infall time-scale for the gas is much shorter than the time-scale for stars to form and begin producing feedback there (e.g. $\lesssim 10^5$ yr), then stellar feedback should not have a noticeable effect on SMBH accretion on these scales.

4 CONSEQUENCES

This simple analytic expression has a number of interesting scaling properties and consequences. In a time-averaged sense, ignoring variations in accretion efficiency through the BH accretion disc, $\dot{M}_{\rm BH} \propto \eta\,(1 - f_{\rm wind})\,M_{\rm gas}(< R)\,\Omega \sim \eta\,(\Sigma_{\rm eff}/(\Sigma_{\rm crit} + \Sigma_{\rm eff}))\,M_{\rm gas}(< R)\,\Omega$. BH growth is dominated by episodes at high accretion rates, which for this estimator are dominated by periods with high gas fractions and $\Sigma_{\rm eff} \gtrsim \Sigma_{\rm crit}$ in the central $\sim$ kpc; these have a characteristic integrated duration $\Delta t \sim$ a few $t_{\rm dyn} = \tau/\Omega$ (where $\tau \sim$ a few, before star formation, outflows, or accretion itself deplete the gas),[7] so $M_{\rm BH} \sim \dot{M}_{\rm BH}\,\Delta t$. With this toy model in mind, consider:

(i) **The Connection Between BHs and Bulges:** because $\dot{M}_{\rm BH}$ decreases rapidly when $\Sigma_{\rm eff} \ll \Sigma_{\rm crit}$, the final BH mass is essentially proportional to the mass of gas at $\Sigma_{\rm eff} \gtrsim \Sigma_{\rm crit}$ (most of which forms stars, as $f_{\rm wind} < 1$ at these densities) in the galaxy center. In other words, the BH growth is specifically sensitive primarily *to the mass at high surface densities* in the galaxy center. But in nearly all studies of BH-host galaxy scalings, the 'bulge' is *defined photometrically as excess light above the central surface brightness of the disc* (Kormendy & Richstone 1995; Magorrian et al. 1998; Kormendy 1999; Ferrarese & Merritt 2000; Gebhardt et al. 2000). This is in fact how such bulges (or NSCs) are usually observationally defined (and measured via e.g. B/D decomposition; Ferrarese et al. 2006). The

critical surface mass density $\Sigma_{\rm crit} \sim 3 \times 10^9\,{\rm M}_\odot\,{\rm kpc}^{-2}$ corresponds, for an old stellar population, to a B-band surface brightness $\mu_{\rm B}^{\rm crit} \sim 18 - 20$ mag arcsec^{-2} (ignoring surface-brightness dimming at high redshifts). This corresponds very neatly with typical $\mu_{\rm B}$ above which bulges or NSCs appear (references above and e.g. Allen et al. 2006; Fisher & Drory 2008)! That is not an accident, as the *same* $\langle \dot{p}/m_* \rangle$ or $\Sigma_{\rm crit}$ appears (via $f_{\rm wind}$) in the self regulation of star formation that regulates galaxy mass profiles/surface densities (see Grudić et al. 2019a, 2020). But the physical interpretation is quite different, as here it is not AGN but stellar feedback doing the 'regulation.'

What is striking here is that, unlike many BH accretion rate models, this depends *explicitly* on surface brightness/density (the same quantity that defines bulges/NSCs), in a *non-linear* manner. This provides an obvious, natural explanation for the fact that BHs appear to better correlate with the properties of these 'central light' excesses, instead of just the galaxy properties as a whole, or the central potential (which, especially in discs with small bulges/NSCs, can easily be dominated by the more extended DM halo and disc), or properties of the disc, or circular velocity/halo mass (Tremaine et al. 2002; Gultekin et al. 2009b; Kormendy et al. 2011; Kormendy & Bender 2011; Reines & Volonteri 2015). It also naturally explains secondary correlations with Sersic index (Graham & Driver 2007; Graham & Scott 2015), as higher n_s is a direct reflection of the central high-Σ light component, and why 'pseudobulges' as defined in e.g. Kormendy & Kennicutt (2004), Fisher & Drory (2008), Kormendy & Bender (2012), which feature disc-like low n_s (flat/low central surface brightness profiles) correlate more poorly with BH mass (Greene, Ho & Barth 2008; Hu 2008; Fisher et al. 2012; Kormendy & Ho 2013).

(ii) **The $M_{\rm BH} - \sigma$ Relation:** the central velocity dispersion of a galaxy scales as $\sigma^2 \sim G\,M_{\rm tot}(< R_e)/R_e$. For galaxies (including most disc + bulge systems) where at the effective radius R_e, the effective surface density $\Sigma_{\rm eff}$ is below $\Sigma_{\rm crit}$ (i.e. $\mu_{\rm eff} \gtrsim \mu_{\rm crit}$, where $\mu_{\rm crit}$ and $\mu_{\rm crit}$ are the approximate surface brightness values in some band corresponding to stellar surface densities of $\sim \Sigma_{\rm eff}$ and $\sim \Sigma_{\rm crit}$, respectively), this implies $M_{\rm BH} \sim \tau\,\eta\,(\Sigma_{\rm eff}/\Sigma_{\rm crit})\,M_{\rm gas} \sim (\tau\,f_{\rm gas}/G^2\Sigma_{\rm crit})\,\eta\,(G\,M_{\rm tot}/R)^2 \sim (\tau\,f_{\rm gas}\,\eta/G^2\Sigma_{\rm crit})\,\sigma_{\rm 3D}^4 \sim 10^{8.5}\,{\rm M}_\odot\,(\tau\,f_{\rm gas}\,\eta/0.001)\,(\sigma_{\rm 1D}/200\,{\rm km\,s}^{-1})^4 \propto \sigma^4$, in excellent agreement with the relation observed (Gültekin et al. 2009a; Kormendy & Ho 2013; Läsker et al. 2016), *especially* for low-mass BHs in small/dwarf/late-type host galaxies with effective surface densities $\ll \Sigma_{\rm crit}$ (Barth et al. 2004; Peterson et al. 2005; Baldassare et al. 2015). This is demonstrated explicitly in preliminary cosmological simulation tests in Fig. 3. Note that this is similar to the derivation in King (2003), Murray et al. (2005), McLaughlin, King & Nayakshin (2006) of $M_{\rm BH} \propto \sigma^4$ for self-regulation via single-scattering radiation pressure (momentum flux $\dot{p} = L/c$) for an Eddington-limited BH, not by accident, because $\langle \dot{p}/m_* \rangle$ is order of magnitude similar to $\sim L/c$ for the stars (whether it comes in actual radiation, stellar winds, or SNe) and the L of young stars is dominated by approximately Eddington-limited massive stars (see Grudić et al. 2020).

(iii) **The $M_{\rm BH} - M_{\rm bulge}$ relation:** on the other hand, if most of the galaxy stellar mass lies above $\Sigma_{\rm crit}$ [$\mu_{\rm eff} \lesssim \mu_{\rm crit}$, i.e. 'pure (classical/dense) bulge' systems], then $1 - f_{\rm wind} \sim 1$, and the SFE is order unity, so we simply have $M_{\rm BH} \sim \eta\,M_* \sim 0.001\,(\tau\,\eta/0.001)\,M_{\rm bulge}$. In other words, going from $\mu_{\rm eff} \gg \mu_{\rm crit}$ to $\mu_{\rm eff} \ll \mu_{\rm crit}$, this predicts a transition from $M_{\rm BH} - \sigma$ to $M_{\rm BH} - M_{\rm bulge}$ being the more 'causal' or 'intrinsic' relation. This is somewhat similar to suggestions of a 'break' in $M_{\rm BH} - \sigma$ owing to the well-observed break in the Faber–Jackson relation, where dry merging would lead to a dominant $M_{\rm BH} -$

---

[7] In our order-of-magnitude arguments here, it makes no difference whether the BH grows most of its mass in a single 'event' with duration $\sim \Delta t$, or several events with similar conditions and total (sum) duration $\sim \Delta t$. If BHs grow primarily via many independent events which each contributes very little mass, or via BH–BH mergers, a different treatment would be needed.
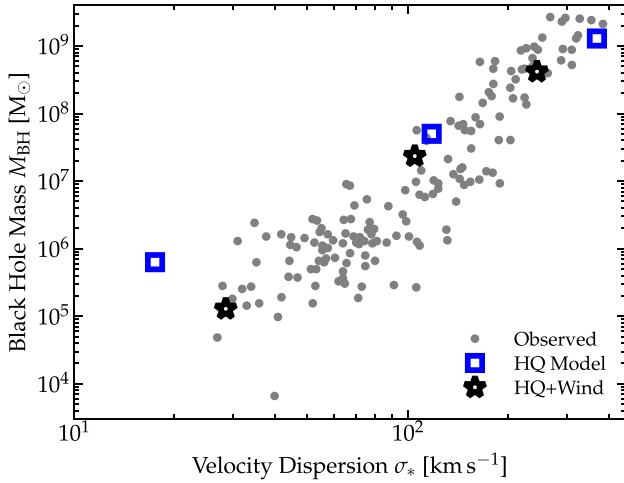
Figure 3. Preliminary comparison of cosmological simulations of three galaxies (to be studied in Wellons et al., in prep), with star formation and stellar feedback following the FIRE project Hopkins et al. (2018), including accreting BHs as in e.g. Anglés-Alcázar et al. (2017b), using two different sub-grid accretion models: (1) the HQ model as in Fig. 2 and Anglés-Alcázar et al. (2017a), and (2) the identical prescription adding the same $(1 - f_{wind})$ correction factor from Fig. 2 and equation (9). We compare to the compilation of observed BHs in Baldassare et al. (2020). For massive systems the proposed correction is a minor effect on total cosmic BH growth, but for dwarfs without massive bulges, ignoring the proposed correction could significantly over-predict BH growth.

M_{bulge} relation at larger masses (Aller & Richstone 2007; Lauer et al. 2007a; McConnell & Ma 2013; Graham & Scott 2015; Sahu, Graham & Davis 2019b; Posti & Fall 2021), but in this case the discriminating criterion is surface brightness based.[8] However it is important to note the caveat that there is no obvious difference in the scatter between $M_{BH}(\sigma)$ and $M_{BH}(M_{bulge})$ observed at present in massive ellipticals (McConnell & Ma 2013; Sahu, Graham & Davis 2019a). But again, because of the natural connection to surface density/acceleration, this argument would explain why the lowest-mass BHs in small hosts (with photometric 'bulges' with relatively low central surface brightness) appear to be 'low' relative to an extrapolated $M_{BH} - M_*$ relation while agreeing better with $M_{BH} - \sigma$ (see Fig. 3 and Barth et al. 2004; Peterson et al. 2005; Greene & Ho 2007; Kormendy & Ho 2013; Baldassare et al. 2015).

(iv) **Mild Redshift Evolution:** it is well established that the progenitors of giant elliptical galaxies today had their central, high-surface brightness 'cores' in place at high redshifts $z \gtrsim 2$ (Bezanson et al. 2009; Hopkins et al. 2009c), and grew primarily in both size and mass via dry merging of smaller systems which accrete the extended 'envelope' of low-surface brightness material and ICL (van Dokkum et al. 2010; Wellons et al. 2016). These 'cores' (whether cuspy or 'cored' in their nuclear profile) easily exceed Σ_{crit}; so if the BH is sensitive to the mass above Σ_{crit} it would reflect essentially the entire galaxy mass in the progenitor. The subsequent merging would contribute negligible material at $> \Sigma_{crit}$, and even the merging BHs are unlikely to sink via dynamical friction (Hopkins et al. 2008,

2009a, d), so M_{bulge} will increase but M_{BH} will not, leading to redshift evolution in M_{BH}/M_{bulge} (as proposed in Croton 2006; Hopkins et al. 2010a). But the effect would be mild, because these galaxies have probably only grown by a factor of ~ 2 in stellar mass (making this the upper limit to redshift evolution in M_{BH}/M_{bulge} to $z \sim 2 - 4$), consistent with observational limits (Suh et al. 2020).

(v) **'Outliers' in Compact Es:** for the same reasons, at similar total M_{bulge}, high-surface brightness cEs will have most of their stellar mass at densities $> \Sigma_{crit}$, while giant Es might have a significant mass fraction below Σ_{crit}, implying the cE would have a larger M_{BH} from these scalings. This is consistent with some claims for observed 'outliers' (McConnell & Ma 2013; Seth et al. 2014; Trakhtenbrot et al. 2015; Walsh et al. 2016; Liepold et al. 2020); however, we stress that the effect saturates, as once most of the mass is at $\Sigma_{eff} \gtrsim \Sigma_{crit}$, there is no 'additional' dependence on compactness, also consistent with the relatively modest limits on such dependence in e.g. Ni et al. (2019).

(vi) **Quenching and Central Surface Densities ('Σ_1'):** In the last few years studies have shown that a number of galaxy and BH properties, particularly related to 'quenching,' are closely correlated with the central surface density of the galaxy, often parametrized as "$\Sigma_1 \equiv M_*(< 1\,\text{kpc})/\pi\,(1\,\text{kpc})^2$ (Franx et al. 2008; Cheung et al. 2012; van der Wel et al. 2012; Whitaker et al. 2012; Huertas-Company et al. 2016; Ellison et al. 2018; Lee et al. 2018). It is immediately obvious that the model here predicts such a correlation with BH growth: for example, this would automatically explain recent studies showing that BH growth rates and AGN activity increases with Σ_1 at otherwise fixed galaxy properties (Ni et al. 2019). But more strikingly, the most robust observation of interest is that Σ_1 correlates strongly with whether or not a galaxy is 'quenched,' with the quenched fraction increasing rapidly around a critical $\Sigma_1 \sim 3 \times 10^9 \, M_\odot \, \text{kpc}^{-2}$ (Cheung et al. 2012; Barro et al. 2017) – remarkably similar to the predicted Σ_{crit}! Phenomenologically, many have argued this could be a signature of quenching driven by AGN feedback, if BHs were somehow sensitive to Σ_1 (Pandya et al. 2017; Rodríguez-Puebla et al. 2017; Chen et al. 2020). The models here predict a natural explanation for precisely such a dependence of BH growth, and therefore AGN feedback, on Σ_1 around Σ_{crit}: for example, in a model where e.g. the integrated BH feedback energy deposition scales $E_{fb} \sim \int \epsilon_{fb} L_{AGN} \, dt \sim \epsilon_{fb} \, 0.1 \, M_{BH} c^2 \sim \epsilon_{fb} \, 0.1 \, \eta \, (\Sigma_1/\Sigma_{crit}) \, M_{gas}$ (using the scalings above), comparing this to the binding energy of the halo gas ($E_{halo} \sim f_{bar} V_c^2 M_{halo}$) assuming a universal baryon fraction $M_{gas} \sim f_{bar} M_{halo}$, we have $E_{fb} \gtrsim E_{halo}$ for $\Sigma_1 \gtrsim \Sigma_{crit} (0.01/\epsilon_{fb}) (0.001/\eta) (V_c/200\,\text{km s}^{-1})^2$, remarkably similar to the observed quenching 'ridgeline' (Chen et al. 2020).

(vii) **The L_{AGN}-SFR Relation:** it is observationally well-established that galactic star formation scales with surface density (Kennicutt 1998). In fact, standard theoretical models of the Kennicutt–Schmidt (KS) relation generically predict $\dot{\Sigma}_* \sim t_*^{-1} (\Sigma_{eff}/\Sigma_{crit}) \Sigma_{gas}$,[9] with the same Σ_{crit} appearing because stellar feedback self regulates the local SFR (Hopkins, Quataert & Murray 2011; Ostriker & Shetty 2011; Faucher-Giguère, Quataert & Hopkins 2013). Combining this with our expression for $\Sigma_{eff} \lesssim \Sigma_{crit}$ (where these derivations of the KS relation are valid), we immediately obtain $\langle \dot{M}_{acc} \rangle \sim \eta \, (t_* \Omega) \, \dot{M}_*$ where $t_* \Omega \sim 1$ is only weakly dependent

---

[8]Of course, such a break owing to the role of dry-merging around $\sim 10^{10} \, M_\odot$ would owe in part to galactic star formation being quenched at higher masses, which may relate to Σ_{crit} as discussed herein (Posti & Fall 2021). This means these predictions may all be coupled in a non-trivial manner.

[9]Here $t_* \equiv \langle p/m_* \rangle / \langle \dot{p}/m_* \rangle = (\int \langle \dot{p}[t]/\dot{m}_* \rangle \, dt) / \langle \dot{p}[t=0]/\dot{m}_* \rangle \sim 100 \, \text{Myr}$ where $\langle p/m_* \rangle$ is the time-integrated momentum injected by a single stellar population (SSP) while $\langle \dot{p}/m_* \rangle$ is the instantaneous rate for a zero-age SSP (both averaged over the stellar IMF).

on galaxy properties. Using standard bolometric conversions, this can be written as $\langle L_{\mathrm{AGN}} \rangle \sim 0.1\,(\eta\,t_* \,\Omega/0.001)\,\langle L_{\mathrm{SF}} \rangle$ (or e.g. X-ray AGN luminosity versus IR luminosity from star formation: $L_{\mathrm{X,\,AGN}} \sim 0.004\,(\eta\,t_*\,\Omega/0.001)\,L_{\mathrm{IR}}$, in excellent agreement with the observed relation when AGN variability and selection effects are properly included (Hickox et al. 2014; Grimmett et al. 2020). Indeed, observations may specifically indicate a closer correlation between BH accretion and bulge/compact star formation (as compared to galaxy-wide SFRs), which would naturally follow from this (Yang et al. 2019; Ni et al. 2021).

(viii) **Off-Nuclear Fueling/AGN:** although accretion models of the form in equation (1) do not require (unlike e.g. Bondi–Hoyle accretion) that the BH dominates the potential on all scales, they *do* assume that the BH resides near the local center/minimum of the potential, so that gas which loses angular momentum or energy tends (on average) to move 'inwards' or 'towards' the BH. If a BH is ejected or free-moving through the galaxy (as seen in many dwarfs; Reines et al. 2020), this is no longer valid and equation (1) will tend to over estimate $\dot{M}_{\mathrm{BH}}$. While our simple $(1 - f_{\mathrm{wind}})$ correction is not specifically designed to address this situation, it does have the effect of reducing $\dot{M}_{\mathrm{BH}}$ when BHs are off center, as Σ_{eff} is lower, providing at least a partial improvement in accuracy.

5 CONCLUSIONS

We consider the simplest-possible extension to standard models of AGN/SMBH accretion (parametrized as $\dot{M}_{\mathrm{acc}} \sim \eta(...)\,M_{\mathrm{gas}}(< r)\,\Omega(r)$) from galactic nuclei scales ($\sim 0.1 - 1000\,\mathrm{pc}$), to account for the role of *stellar* feedback ejecting gas from smaller scales before it reaches the AGN accretion disc. As shown in Torrey et al. (2017), when the dynamical time $t_{\mathrm{dyn}} \sim \Omega^{-1} \sim r/V_c$ is less than the stellar evolution time-scale for most SNe ($t_* \sim 100\,\mathrm{Myr}$), the presence of gas in galactic nuclei (and hence its ability to accrete further inwards) is regulated by stellar feedback, with efficient feedback able to eject most gas from the nucleus (not necessarily the galaxy) at low densities. Simple analytic models, detailed simulations of molecular clouds and nuclear gas discs, and direct observations all argue that the efficiency of this ejection scales in a simple manner with the gravitational acceleration $\bar{a}_{\mathrm{grav}} \equiv G\,M_{\mathrm{enc}}(< r)/r^2$ or 'effective surface density' $\Sigma_{\mathrm{eff}} \equiv M_{\mathrm{enc}}(< r)/\pi\,r^2$, as $M_{\mathrm{ejected}}/M_{\mathrm{retained}} \sim \langle \dot{p}/m_* \rangle/\bar{a}_{\mathrm{grav}} \sim \Sigma_{\mathrm{crit}}/\Sigma_{\mathrm{eff}}$, where $\langle \dot{p}/m_* \rangle \sim 10^{-7}\,\mathrm{cm\,s^{-2}}$ ($\Sigma_{\mathrm{crit}} \equiv \langle \dot{p}/m_* \rangle/\pi\,G \sim 3 \times 10^9\,\mathrm{M_\odot\,kpc^{-2}}$) is the momentum flux per unit mass in feedback (radiation + stellar mass loss + SNe) from a zero-age main sequence IMF-integrated stellar population. This leads to a 'correction factor' to accretion models which ignored such stellar feedback-driven ejection, of the form $\eta \to \eta\,(1 - f_{\mathrm{ejected}}) \sim \eta\,\Sigma_{\mathrm{eff}}/(\Sigma_{\mathrm{eff}} + \Sigma_{\mathrm{crit}})$.

We show that this immediately resolves some discrepancies between various high-resolution simulation studies of accretion and inflows in galactic nuclei. Simulations which included explicit resolved stellar feedback, but focused on quasar-level, extremely dense gaseous torii or nuclear discs with $\Sigma_{\mathrm{eff}} \gtrsim 10^5\,\mathrm{M_\odot\,cm^{-2}} \gg \Sigma_{\mathrm{crit}}$ (Wada et al. 2009; Hopkins et al. 2016) have found accretion rates $\dot{M}_{\mathrm{acc}}$ in good agreement with older simulations that did not include explicit stellar feedback-driven outflows at all (e.g. Hopkins & Quataert 2010b, 2011a, b), while simulations with lower central densities (representing discs or dwarf galaxies, with little nuclear gas) found much lower inflow rates (Dubois et al. 2015; Anglés-Alcázar et al. 2017b; Torrey et al. 2017).

We go on to show that with this correction factor, the resulting approximate expression for BH accretion rates has a number of interesting properties. Most importantly, because $1 - f_{\mathrm{ejected}} \sim 1$ when $\Sigma_{\mathrm{eff}} \gtrsim$

Σ_{crit} and $1 - f_{\mathrm{ejected}} \sim \Sigma_{\mathrm{eff}}/\Sigma_{\mathrm{crit}} \ll 1$ when $\Sigma_{\mathrm{eff}} \ll \Sigma_{\mathrm{crit}}$, this predicts that BH mass should be correlated most directly with the mass in the galaxy center above a critical effective (total-mass) surface density $\sim \Sigma_{\mathrm{crit}}$ (which, from the same feedback-regulation model, should mostly turn into stars at these high densities). This corresponds to an intrinsic stellar surface brightness $\mu_{\mathrm{B}} \sim (18 - 20)\,\mathrm{mag\,arcsec^{-2}}$ for an old stellar population (depending on age, metallicity, etc.). This corresponds remarkably well to the characteristic surface brightness above which 'bulges' dominate the light. In fact, in essentially all studies of BH-host galaxy scalings, 'bulge' properties are defined *photometrically*, as excess surface brightness above the disc around the BH. This is true even when the bulge does not contain enough total mass to dominate the central potential or escape velocity from the galaxy center – where models which predict BH mass traces binding/kinetic/potential energy or escape velocity would predict a better correlation between BH properties and discs, instead of bulges (which is not observed). This also immediately explains why BHs do not simply correlate with 'central mass' within some fixed physical aperture, as many models also predict, but with specific photometric *features* of galaxies. In short, this simple stellar-feedback-regulated scaling therefore immediately explains why, in fact, *bulge* properties appear to predict BH masses.

We also show that this scaling leads immediately to the observed BH–σ relation, directly, especially in lower-mass host galaxies, and explains a wide variety of secondary correlations or lack thereof (e.g. why BHs appear to correlate more poorly with photometrically defined 'pseudobulges'; secondary correlations with galaxy compactness, Sersic index, redshift, and position on the Faber–Jackson relation). And we show that, during active accretion phases, if we invoke the same stellar-feedback regulated arguments commonly used to explain the galactic Schmidt–Kennicutt star formation scalings, we immediately predict a correlation between *mean* AGN luminosity (albeit with large variability expected) and galactic SFR, in agreement with that observed. As a result, if AGN feedback plays a critical role in galaxy quenching, the argument here may also play a critical role explaining the 'critical' value of central surface density $\Sigma_1 \sim 3 \times 10^9\,\mathrm{M_\odot\,kpc^{-2}}$ above which galaxies tend to be quenched – which is observed to be remarkably similar to the predicted Σ_{crit} where AGN accretion is efficient.

Of course, our study here is a simple analytic investigation of dimensional scalings. More refined models will require further, high-resolution numerical simulations and observations of gas in galactic nuclei to test these scalings and calibrate exact coefficients as well as detailed dependence on e.g. gas properties, stellar mass distributions, dynamical state of galaxies, etc. We also stress that we neglect AGN feedback here, as an additional regulator of BH accretion. Of course, AGN can eject mass directly from accretion disc/jet scales (this would appear as some sub-grid 'efficiency' in models here); they can also regulate inflow on these scales by driving large-scale outflows, changing the properties (e.g. M_{gas}, Σ_{eff}) which determined $\dot{M}_{\mathrm{acc}}$, but this does not necessarily change our scaling (equation 1) for $\dot{M}_{\mathrm{acc}}$. Determining whether there is a more complex non-linear interplay again requires self-consistent simulations.

DATA AVAILABILITY STATEMENT

The data supporting the plots within this article are available on reasonable request to the corresponding author. A public version of the GIZMO code is available at http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html.

REFERENCES

Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, ApJ, 770, 25
Alexander D. M., Hickox R. C., 2012, New A Rev., 56, 93
Alig C., Schartmann M., Burkert A., Dolag K., 2013, ApJ, 771, 119
Allen P. D., Driver S. P., Graham A. W., Cameron E., Liske J., de Propris R., 2006, MNRAS, 371, 2
Aller M. C., Richstone D. O., 2007, ApJ, 665, 120
Anglés-Alcázar D., Özel F., Davé R., 2013, ApJ, 770, 5
Anglés-Alcázar D., Davé R., Faucher-Giguère C.-A., Özel F., Hopkins P. F., 2017a, MNRAS, 464, 2840
Anglés-Alcázar D., Faucher-Giguère C.-A., Quataert E., Hopkins P. F., Feldmann R., Torrey P., Wetzel A., Kereš D., 2017b, MNRAS, 472, L109
Angles-Alcazar D. et al., 2020, ApJ, 917, 53
Baldassare V. F., Reines A. E., Gallo E., Greene J. E., 2015, ApJ, 809, L14
Baldassare V. F., Dickey C., Geha M., Reines A. E., 2020, ApJ, 898, L3
Barro G. et al., 2017, ApJ, 840, 47
Barth A. J., Ho L. C., Rutledge R. E., Sargent W. L. W., 2004, ApJ, 607, 90
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Best P. N., von der Linden A., Kauffmann G., Heckman T. M., Kaiser C. R., 2007, MNRAS, 379, 894
Bezanson R., van Dokkum P. G., Tal T., Marchesini D., Kriek M., Franx M., Coppi P., 2009, ApJ, 697, 1290
Bondi H., Hoyle F., 1944, MNRAS, 104, 273
Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, MNRAS, 465, 32
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Chen Z. et al., 2020, ApJ, 897, 102
Cheung E. et al., 2012, ApJ, 760, 131
Ciotti L., Ostriker J. P., 1997, ApJ, 487, L105+
Colín P., Vázquez-Semadeni E., Gómez G. C., 2013, MNRAS, 435, 1701
Combes F. et al., 2013, A&A, 558, A124
Courteau S., de Jong R. S., Broeils A. H., 1996, ApJ, 457, L73
Cox T. J., Di Matteo T., Hernquist L., Hopkins P. F., Robertson B., Springel V., 2006, ApJ, 643, 692
Crenshaw D. M. et al., 2000, AJ, 120, 1731
Croton D. J., 2006, MNRAS, 369, 1808
Croton D. J. et al., 2006, MNRAS, 365, 11
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Debuhr J., Quataert E., Ma C., Hopkins P., 2010, MNRAS, 406, L55
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, ApJ, 676, 33
Dubois Y., Volonteri M., Silk J., Devriendt J., Slyz A., Teyssier R., 2015, MNRAS, 452, 1502
Dunn J. P. et al., 2010, ApJ, 709, 611
Ellison S. L., Sánchez S. F., Ibarra-Medel H., Antonio B., Mendel J. T., Barrera-Ballesteros J., 2018, MNRAS, 474, 2039
Emsellem E., Renaud F., Bournaud F., Elmegreen B., Combes F., Gabor J. M., 2015, MNRAS, 446, 2468
Esquej P. et al., 2014, ApJ, 780, 86
Fabian A. C., 2012, ARA&A, 50, 455
Fall S. M., Krumholz M. R., Matzner C. D., 2010, ApJ, 710, L142
Faucher-Giguère C.-A., Quataert E., Hopkins P. F., 2013, MNRAS, 433, 1970
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Ferrarese L., van den Bosch F. C., Ford H. C., Jaffe W., O'Connell R. W., 1994, AJ, 108, 1598
Ferrarese L. et al., 2006, ApJ, 644, L21
Fisher D. B., Drory N., 2008, AJ, 136, 773
Fisher D. B., Bolatto A., Drory N., Combes F., Blitz L., Wong T., 2012, ApJ, 764, 174
Franx M., van Dokkum P. G., Förster Schreiber N. M., Wuyts S., Labbé I., Toft S., 2008, ApJ, 688, 770
Gammie C. F., 2001, ApJ, 553, 174
García-Burillo S. et al., 2014, A&A, 567, A125
Gavagnin E., Bleuler A., Rosdahl J., Teyssier R., 2017, MNRAS, 472, 4155
Gebhardt K. et al., 2000, ApJ, 539, L13
Geen S., Soler J. D., Hennebelle P., 2017, MNRAS, 471, 4844
Goodman J., 2003, MNRAS, 339, 937
Governato F. et al., 2010, Nature, 463, 203
Graham A. W., Driver S. P., 2007, ApJ, 655, 77
Graham A. W., Scott N., 2015, ApJ, 798, 54
Greene J. E., Ho L. C., 2007, ApJ, 670, 92
Greene J. E., Ho L. C., Barth A. J., 2008, ApJ, 688, 159
Grimmett L. P., Mullaney J. R., Bernhard E. P., Harrison C. M., Alexander D. M., Stanley F., Masoura V. A., Walters K., 2020, MNRAS, 495, 1392
Grudić M. Y., Hopkins P. F., 2019, MNRAS, 488, 2970
Grudić M. Y., Hopkins P. F., Faucher-Giguère C.-A., Quataert E., Murray N., Kereš D., 2018, MNRAS, 475, 3511
Grudić M. Y., Hopkins P. F., Quataert E., Murray N., 2019a, MNRAS, 483, 5548
Grudić M. Y., Hopkins P. F., Lee E. J., Murray N., Faucher-Giguère C.-A., Johnson L. C., 2019b, MNRAS, 488, 1501
Grudić M. Y., Boylan-Kolchin M., Faucher-Giguère C.-A., Hopkins P. F., 2020, MNRAS, 496, L127
Gültekin K. et al., 2009a, ApJ, 698, 198
Gultekin K. et al., 2009b, ApJ, 698, 198
Habouzit M., Volonteri M., Dubois Y., 2017, MNRAS, 468, 3935
Heckman T. M., Best P. N., 2014, ARA&A, 52, 589
Hickox R. C., Mullaney J. R., Alexander D. M., Chen C.-T. J., Civano F. M., Goulding A. D., Hainline K. N., 2014, ApJ, 782, 9
Hobbs A., Nayakshin S., Power C., King A., 2011, MNRAS, 413, 2633
Hobbs A., Power C., Nayakshin S., King A. R., 2012, MNRAS, 421, 3443
Hopkins P. F., Christiansen J. L., 2013, ApJ, 776, 48
Hopkins P. F., Grudić M. Y., 2019, MNRAS, 483, 4187
Hopkins P. F., Quataert E., 2010a, MNRAS, 405, L41
Hopkins P. F., Quataert E., 2010b, MNRAS, 407, 1529
Hopkins P. F., Quataert E., 2011a, MNRAS, 411, L61
Hopkins P. F., Quataert E., 2011b, MNRAS, 415, 1027
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Martini P., Robertson B., Springel V., 2005a, ApJ, 630, 705
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2005b, ApJ, 632, 81
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006a, ApJS, 163, 1
Hopkins P. F., Narayan R., Hernquist L., 2006b, ApJ, 643, 641
Hopkins P. F., Hernquist L., Cox T. J., Robertson B., Krause E., 2007a, ApJ, 669, 45
Hopkins P. F., Hernquist L., Cox T. J., Robertson B., Krause E., 2007b, ApJ, 669, 67
Hopkins P. F., Cox T. J., Kereš D., Hernquist L., 2008, ApJS, 175, 390

Hopkins P. F., Lauer T. R., Cox T. J., Hernquist L., Kormendy J., 2009a, ApJS, 181, 486
Hopkins P. F., Murray N., Thompson T. A., 2009b, MNRAS, 398, 303
Hopkins P. F., Bundy K., Murray N., Quataert E., Lauer T. R., Ma C.-P., 2009c, MNRAS, 398, 898
Hopkins P. F., Hernquist L., Cox T. J., Kereš D., Wuyts S., 2009d, ApJ, 691, 1424
Hopkins P. F., Bundy K., Hernquist L., Wuyts S., Cox T. J., 2010a, MNRAS, 401, 1099
Hopkins P. F. et al., 2010b, ApJ, 715, 202
Hopkins P. F. et al., 2010c, ApJ, 724, 915
Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950
Hopkins P. F., Hayward C. C., Narayanan D., Hernquist L., 2012a, MNRAS, 420, 320
Hopkins P. F., Quataert E., Murray N., 2012b, MNRAS, 421, 3488
Hopkins P. F., Keres D., Onorbe J., Faucher-Giguere C.-A., Quataert E., Murray N., Bullock J. S., 2014a, MNRAS, 445, 581
Hopkins P. F., Kocevski D. D., Bundy K., 2014b, MNRAS, 445, 823
Hopkins P. F., Torrey P., Faucher-Giguère C.-A., Quataert E., Murray N., 2016, MNRAS, 458, 816
Hopkins P. F. et al., 2018, MNRAS, 480, 800
Hu J., 2008, MNRAS, 386, 2242
Huertas-Company M. et al., 2016, MNRAS, 462, 4495
Inayoshi K., Ichikawa K., Ostriker J. P., Kuiper R., 2019, MNRAS, 486, 5377
Izumi T., Kawakatu N., Kohno K., 2016, ApJ, 827, 81
Jiang Y.-F., Goodman J., 2011, ApJ, 730, 45
Kawakatu N., Wada K., 2008, ApJ, 681, 73
Kawakatu N., Wada K., Ichikawa K., 2020, ApJ, 889, 84
Kennicutt R. C. Jr, 1998, ApJ, 498, 541
Kim C.-G., Ostriker E. C., 2015, ApJ, 815, 67
Kim W.-T., Seo W.-Y., Kim Y., 2012, ApJ, 758, 14
Kim J.-G., Kim W.-T., Ostriker E. C., 2018, ApJ, 859, 68
King A., 2003, ApJ, 596, L27
Kocevski D. D. et al., 2012, ApJ, 744, 148
Kormendy J., 1999, in Merritt D. R., Valluri M., Sellwood J. A., eds, ASP Conf. Ser. Vol. 182, Galaxy Dynamics - A Rutgers Symposium. Astron. Soc. Pac., San Francisco, p. 124
Kormendy J., Bender R., 2011, Nature, 469, 377
Kormendy J., Bender R., 2012, ApJS, 198, 2
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kormendy J., Kennicutt R. C. Jr, 2004, ARA&A, 42, 603
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Kormendy J., Bender R., Cornell M. E., 2011, Nature, 469, 374
Kruijssen J. M. D. et al., 2019, Nature, 569, 519
Laor A., Fiore F., Elvis M., Wilkes B. J., McDowell J. C., 1997, ApJ, 477, 93
Läsker R., Greene J. E., Seth A., van de Ven G., Braatz J. A., Henkel C., Lo K. Y., 2016, ApJ, 825, 3
Lauer T. R. et al., 2002, AJ, 124, 1975
Lauer T. R. et al., 2007a, ApJ, 662, 808
Lauer T. R. et al., 2007b, ApJ, 664, 226
Lee B. et al., 2018, ApJ, 853, 131
Leitherer C. et al., 1999, ApJS, 123, 3
Li Z., Shen J., Kim W.-T., 2015, ApJ, 806, 150
Liepold C. M., Quenneville M. E., Ma C.-P., Walsh J. L., McConnell N. J., Greene J. E., Blakeslee J. P., 2020, ApJ, 891, 4
Magorrian J. et al., 1998, AJ, 115, 2285
Mancini L., Feoli A., 2012, A&A, 537, A48
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
McLaughlin D. E., King A. R., Nayakshin S., 2006, ApJ, 650, L37
McMillan P. J., 2017, MNRAS, 465, 76
Merloni A., Heinz S., 2008, MNRAS, 388, 1011
Murray N., Rahman M., 2010, ApJ, 709, 424
Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569

Naab T., Ostriker J. P., 2017, ARA&A, 55, 59
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Ni Q., Yang G., Brandt W. N., Alexander D. M., Chen C. T. J., Luo B., Vito F., Xue Y. Q., 2019, MNRAS, 490, 1135
Ni Q. et al., 2021, MNRAS, 500, 4989
Ostriker E. C., Shetty R., 2011, ApJ, 731, 41
Pandya V. et al., 2017, MNRAS, 472, 2054
Peng C. Y., 2007, ApJ, 671, 1098
Peterson B. M. et al., 2005, ApJ, 632, 799
Pillepich A., Madau P., Mayer L., 2015, ApJ, 799, 184
Posti L., Fall S. M., 2021, A&A, 649, A119
Puech M., Hammer F., Hopkins P. F., Athanassoula E., Flores H., Rodrigues M., Wang J. L., Yang Y. B., 2012, ApJ, 753, 128
Querejeta M. et al., 2016, A&A, 588, A33
Reines A. E., Volonteri M., 2015, ApJ, 813, 82
Reines A. E., Condon J. J., Darling J., Greene J. E., 2020, ApJ, 888, 36
Rodríguez-Puebla A., Primack J. R., Avila-Reese V., Faber S. M., 2017, MNRAS, 470, 651
Sahu N., Graham A. W., Davis B. L., 2019a, ApJ, 876, 155
Sahu N., Graham A. W., Davis B. L., 2019b, ApJ, 887, 10
Salucci P., Szuszkiewicz E., Monaco P., Danese L., 1999, MNRAS, 307, 637

Savorgnan G. A. D., Graham A. W., 2016, ApJS, 222, 10
Schartmann M., Burkert A., Krause M., Camenzind M., Meisenheimer K., Davies R. I., 2010, MNRAS, 403, 1801
Seth A. C. et al., 2014, Nature, 513, 398
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shu F. H., 1977, ApJ, 214, 488
Silk J., Rees M. J., 1998, A&A, 331, L1
Soltan A., 1982, MNRAS, 200, 115
Sturm E. et al., 2011, ApJ, 733, L16+
Suh H., Civano F., Trakhtenbrot B., Shankar F., Hasinger G., Sand ers D. B., Allevato V., 2020, ApJ, 889, 32
Taranu D. S. et al., 2017, ApJ, 850, 70
Thomas N., Davé R., Anglés-Alcázar D., Jarvis M., 2019, MNRAS, 487, 5764
Thompson T. A., Quataert E., Murray N., 2005, ApJ, 630, 167
Torrey P., Hopkins P. F., Faucher-Giguère C.-A., Vogelsberger M., Quataert E., Kereš D., Murray N., 2017, MNRAS, 467, 2301
Torrey P. et al., 2020, MNRAS, 497, 5292
Trakhtenbrot B. et al., 2015, Science, 349, 168
Tremaine S. et al., 2002, ApJ, 574, 740
van der Wel A. et al., 2012, ApJS, 203, 24
van Dokkum P. G. et al., 2010, ApJ, 709, 1018
Voit G. M., Meece G., Li Y., O'Shea B. W., Bryan G. L., Donahue M., 2017, ApJ, 845, 80
Vollmer B., Beckert T., Davies R. I., 2008, A&A, 491, 441
Wada K., Norman C. A., 2002, ApJ, 566, L21
Wada K., Papadopoulos P. P., Spaans M., 2009, ApJ, 702, 63
Walsh J. L., van den Bosch R. C. E., Gebhardt K., Yıldırım A., Richstone D. O., Gültekin K., Husemann B., 2016, ApJ, 817, 2
Wellons S. et al., 2016, MNRAS, 456, 1030
Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, ApJ, 754, L29
Williams R. J., Maiolino R., Krongold Y., Carniani S., Cresci G., Mannucci F., Marconi A., 2017, MNRAS, 467, 3399
Wutschik S., Schleicher D. R. G., Palmer T. S., 2013, A&A, 560, A34
Yang G., Brandt W. N., Alexander D. M., Chen C. T. J., Ni Q., Vito F., Zhu F. F., 2019, MNRAS, 485, 3721
Yu Q., Tremaine S., 2002, MNRAS, 335, 965
Zakamska N. L. et al., 2016, MNRAS, 459, 3144

This paper has been typeset from a TeX/LaTeX file prepared by the author.