



Grand unification of AGN activity in the Λ CDM cosmology

N. Fanidakis,¹[★] C. M. Baugh,¹ A. J. Benson,² R. G. Bower,¹ S. Cole,¹ C. Done¹ and C. S. Frenk¹

¹Institute for Computational Cosmology, Department of Physics, University of Durham, Science Laboratories, South Road, Durham DH1 3LE

²Mail Code 130-33, California Institute of Technology, Pasadena, CA 91125, USA

Accepted 2010 July 26. Received 2010 July 18; in original form 2010 January 13

ABSTRACT

We track the coevolution of supermassive black holes (SMBHs) and their host galaxies through cosmic time. The calculation is embedded in the GALFORM semi-analytic model which simulates the formation and evolution of galaxies in a cold dark matter (CDM) universe. The black hole (BH) and galaxy formation models are coupled: during the evolution of the host galaxy, hot and cold gas are added to the SMBH by flows triggered by halo gas cooling, disc instabilities and galaxy mergers. This builds up the mass and spin of the BH, and the resulting accretion power regulates gas cooling and subsequent star formation. The accretion flow is assumed to form a geometrically thin cool disc when the accretion rate exceeds $0.01\dot{M}_{\text{Edd}}$, and a geometrically thick, radiatively inefficient hot flow when the accretion rate falls below this value. The resulting quasar optical luminosity function matches observations well, and the mass of the SMBH correlates with the mass of the galaxy bulge as in the observed $M_{\text{bh}}-M_{\text{bulge}}$ relation. The BH spin distribution depends strongly on whether we assume that the gas in any given accretion episode remains in the same plane or it fragments into multiple, randomly aligned accretion episodes due to its self-gravity. We refer to these cases as the ‘prolonged’ and ‘chaotic’ accretion modes, respectively. In the chaotic accretion model there is a clear correlation of spin with SMBH mass (and hence host galaxy bulge mass). Massive BHs ($M > 5 \times 10^8 M_{\odot}$) are hosted by giant elliptical galaxies and are rapidly spinning, while lower mass BHs are hosted in spiral galaxies and have much lower spin. Using the Blandford–Znajek mechanism for jet production to calculate the jet power, our model reproduces the radio loudness of radio galaxies, low ionization emission regions (LINERS) and Seyferts, suggesting that the jet properties of active galaxy nuclei (AGN) are a natural consequence of both the accretion rate on to and the spin of the central SMBH. This is the first confirmation that a CDM galaxy formation model can reproduce the observed radio phenomenology of AGN.

Key words: methods: numerical – galaxies: active – galaxies: jets – galaxies: nuclei – quasars: general.

1 INTRODUCTION

Active galaxies can be classified according to the importance of the radio emission from their nucleus. Objects showing strong emission at radio wavelengths belong to the ‘radio-loud’ class, whereas those with negligible emission belong to the ‘radio-quiet’ class. Radio-loud active galactic nuclei (AGN) are associated with large-scale radio-emitting jets while radio-quiet AGN show very little or negligible jet activity. Sikora, Stawarz & Lasota (2007) found that AGN

have a bimodal distribution on the radio-optical plane, with radio-loud objects being about 10^3 times brighter in radio than radio-quiet objects (see also Kellermann et al. 1989; Xu, Livio & Baum 1999). The origin of this dichotomy remains unknown. However, it has been proposed in many studies that a first step towards explaining why some AGN launch prominent jets while others do not is to understand the nature of the central engine, the accreting black hole (BH).

AGN jets are believed to extract rotational energy from the BH and accretion disc through magnetic fields. Analytical studies have contributed to an understanding of the nature of the jets (Blandford & Znajek 1977; MacDonald & Thorne 1982; Blandford & Payne

*E-mail: nikolaos.fanidakis@dur.ac.uk

1982; Begelman, Blandford & Rees 1984), but a breakthrough has come from magnetohydrodynamical (MHD) simulations of the accretion flow. These self-consistently calculate the turbulent magnetic field dynamo which is the physical origin of the stresses which transport angular momentum outwards, allowing material to accrete on to the BH (Balbus & Hawley 1998). Embedding such calculations in a fully general relativistic framework produces relativistic jets from the accretion flow, and the collimation and jet power depend on the BH spin (McKinney & Gammie 2004; De Villiers et al. 2005; Hawley & Krolik 2006).

Thus, it now seems clear that the jet power depends on the vertical (poloidal) magnetic field strength close to the BH, but the simulations are still not yet at the level where they can predict this ab initio (e.g. Beckwith, Hawley & Krolik 2008), so analytic models are still required. The most popular of these, the Blandford-Znajek (BZ) mechanism, uses the magnetic field as a means to tap the spin of the BH. This gives a strong spin dependence on the jet power. Indeed, many authors have proposed spin as the physical parameter that determines the radio loudness of an AGN (see Wilson & Colbert 1995; Hughes & Blandford 2003). The spin paradigm, as this idea is called, offers a plausible theoretical explanation for the wide range of jet luminosities in AGN and could be the basis for understanding the observed dichotomy between radio-loud and radio-quiet AGN.

Alternatively, the radio-loud, radio-quiet switch may just be determined by the physical state of the accretion flow. Stellar-mass BH binary systems in our galaxy show a clear spectral transition at about 1 per cent of the Eddington luminosity, $L \sim 0.01L_{\text{Edd}}$, from a hot, optically thin accretion flow at low luminosities (e.g. an advection-dominated accretion flow or ADAF; Narayan & Yi 1994) to a cool, geometrically thin disc (TD; Shakura & Sunyaev 1973) at higher luminosities (Esin, McClintock & Narayan 1997; see e.g. the review by Done, Gierliński & Kubota 2007). The collapse of the radio jet observed across this transition clearly relates to the large drop in pressure (and hence scaleheight, H) between the hot and cool flow (e.g. Fender, Belloni & Gallo 2004). Even without BH spin, this produces a clear dichotomy of radio properties (e.g. Jester 2005).

It seems most plausible that *both* these mechanisms, along with the mass accretion rate, affect jet power. The BZ mechanism has a ‘hidden’ dependence on mass accretion rate and the scaleheight of the flow because the magnetic field strength close to the BH saturates, due to the dynamo, into rough equipartition with the pressure in the flow. This depends on the mass accretion rate for either the disc or the hot flow, but the pressure in the hot flow is much larger than that in the cool disc so the field strength and hence jet power abruptly drop at this transition (Meier 2001, 2002).

Sikora et al. (2007) used these ideas to interpret observations of radio galaxies and concluded that the data could be explained if the jet depends on mass accretion rate, accretion model and spin. However, they also required a mass-spin correlation, such that the most massive BHs, which, according to the $M_{\text{bh}}-M_{\text{bulge}}$ relationship (Magorrian et al. 1998; McLure & Dunlop 2002; Marconi & Hunt 2003; Häring & Rix 2004), reside in massive elliptical galaxies, have higher spin than lower mass BHs, which reside in spirals. They speculated that this could occur through the hierarchical growth of structure, where the major mergers which give rise to giant elliptical galaxies trigger large amounts of gas accretion with a given angular momentum direction on to the BH, spinning it up to the maximal value. By contrast, spiral galaxies have not experienced a recent major merger, and grow mainly through smooth accretion and multiple minor mergers with random angular momentum, resulting in low spin and hence weak jet luminosities.

Some of these ideas have recently been explored using simplified models of galaxy formation by Berti & Volonteri (2008) and Lagos, Padilla & Cora (2009). In this paper, we calculate the growth of BHs by accretion and mergers, their acquisition of spin and their accretion rates within a specific theory of galaxy formation in a cold dark matter (CDM) model which has been extensively tested against a large range of observations of the galaxy population. Not only is our calculation of the joint evolution of BHs and galaxies self-consistent, but we also incorporate several accretion and jet-launching models from the literature which allows us to perform a quantitative comparison with observations of AGN.

The paper is organized as follows. In Section 2, we explain how we calculate the growth of the mass and spin of supermassive black holes (SMBHs) through the accretion of hot gas from the halo (radio mode) and cold gas from the galactic disc and mergers (quasar mode). Most BHs grow primarily by gas accretion except the most massive ones which form late and increase their mass substantially by merging with other SMBHs. A BH is spun up to approximately the maximal value when it accretes its own mass from a flow at constant angular momentum.

In Section 3, we present the resulting distributions of BH mass, gas accretion rate and spin. We consider two possible modes of accretion. In the ‘prolonged accretion’ case, the gas is accreted in a single episode. Even minor mergers trigger gas flows on to the nucleus that often deposit a mass greater than the mass of the recipient BH. Thus, most BHs are typically spun up to the maximal value. This model predicts radio properties that do not reproduce the observed luminosity function. We then consider a ‘chaotic accretion’ case in which the accretion episodes are limited by the self-gravity of the disc. Gas flows on to the nucleus give rise to a series of accretion episodes each typically augmenting the BH mass by only a small factor. Successive accretion events are uncorrelated, resulting in low spins for the relatively low mass BHs which grow primarily by accretion. On the other hand, the most massive holes which build up substantial mass through BH–BH mergers are spun up to high (but not maximal) spin values. In Section 4, we show that adopting this accretion model, the optical luminosity from disc-accreting objects matches the quasar luminosity function reasonably well.

In Section 5, we incorporate an explicit BZ model for the jet power (Meier 2002) and use this to predict the radio luminosity function which, in the chaotic accretion case, agrees well with observations. On the basis of our derived optical and radio luminosity distributions, we propose, in Section 6, a ‘grand unification of AGN activity’, in which the accretion flow and jet luminosity are related to AGN type through their mass, spin and mass accretion rate. Fundamentally, our calculation shows, for the first time, how the coeval growth of BHs and their host galaxies result in optical and radio properties that explain the AGN activity seen in the local Universe.

2 EVOLUTION OF SPIN IN HIERARCHICAL GALAXY FORMATION MODELS

In this section we describe the cosmological processes that initiate and regulate the growth of SMBHs and outline the modelling of the spin evolution due to gas accretion and mergers in hierarchical cosmologies. Our basic modelling tool for our predictions is the GALFORM code (Cole et al. 2000), and we use an update of the Bower et al. (2006) version for modelling the formation and evolution of galaxies in the Λ CDM cosmology. The changes relative to the Bower et al. model are the following. First, the fraction, ϵ_{SMBH} , of the Eddington luminosity of an accreting SMBH that is available for heating the halo during an episode of AGN feedback is set to

0.01 (Bower et al. 2006 use $\epsilon_{\text{SMBH}} = 0.04^1$). Secondly, in starbursts triggered by galaxy mergers or disc instabilities, we assume that the fraction, F_{SMBH} , of the cold gas turned into stars that is accreted on to the BH is 0.01 (Bower et al. 2006 use $F_{\text{SMBH}} = 0.017$). These changes are introduced to improve the modelling in this paper and *do not* affect the fundamental predictions of the Bower et al. model.

BHs are assumed to evolve in mass in accordance with the model developed in Malbon et al. (2007) and Bower et al. (2006). We study the cosmological spin evolution of SMBH seeds of a total of $\sim 4.2 \times 10^6$ galaxies identified in the Millennium N -body simulation (comoving volume of $1.25 \times 10^8 h^{-3} \text{ Mpc}^3$) from redshift 127 to redshift zero (Springel et al. 2005). Here h is defined by $H_0 = h \times 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, where H_0 is the Hubble constant at redshift zero. The cosmology adopted in the simulations is $h = 0.72$, $\Omega_m = 0.25$, $\Omega_b = 0.045$, $\Omega_\Lambda = 0.75$ and $\sigma_8 = 0.9$.²

2.1 The growth of SMBHs

The evolution of BH mass fits naturally in the CDM model of galaxy formation (Kauffmann & Haehnelt 2000; Malbon et al. 2007), where structures grow hierarchically. Small structures form first, then evolve through mergers into large ones. In the event of a galaxy merger, the less massive galaxy (satellite) sinks into the gravitational potential of the massive central galaxy as a result of dynamical friction (Binney & Tremaine 1987). In our model we assume that if the mass of the satellite galaxy is comparable to that of the central galaxy, the merger disrupts the galaxies and result in the formation of an elliptical galaxy (major merger). This event is accompanied by a burst of star formation as the available cold gas from both progenitors is transferred to the centre and transformed into stars. Some of this cold gas reservoir feeds the central SMBH. An additional process of cold gas accretion that contributes to the SMBH mass is the collapse of galaxy discs triggered by dynamical instabilities (Efstathiou, Lake & Negroponte 1982). When the self-gravity of the galactic disc becomes sufficiently large, the disc equilibrium is disrupted resulting in the formation of a bar which enables gas to be transferred towards the centre as the disc. A fraction of that gas is directly fed into the BH through an accretion disc and powers an AGN (Lynden-Bell 1969), while the rest undergoes star formation. For highly efficient accretion activity, the disc/BH system dominates the energetics of the nucleus, and the galaxy becomes visible as a quasar. SMBHs at the centres of galaxies observed in the local Universe are regarded as the remnants of quasar phases at earlier epochs.

The consequence of a minor merger, namely the accretion of satellites of low mass compared to the central galaxy, is usually less dramatic. During a minor merger in our model, cold gas and stellar content of the satellites are added to the central galaxy, and a merger-driven starburst may supply fresh gas to the central BH.

SMBHs acquire part of their mass merging with other SMBHs. The formation of a BH–BH binary and the subsequent coalescence of the BHs is a natural evolutionary stage for a SMBH if the host galaxy experiences a merger. As the galaxies merge, we assume that

¹ Note that due to an error in Bower et al. (2006), cooling luminosities were overestimated by a factor of 4π . Thus, while the paper quotes the efficiency parameter ϵ_{SMBH} as 0.5, this should have been $0.5/4\pi = 0.04$. With this correction, the rest of the parameters and results are unchanged.

² Ω_m , Ω_b and Ω_Λ express the present density of the baryonic, total matter and dark energy components of the Universe relative to the critical density ($\rho_{\text{crit}} = 3H^2/8\pi G$). σ_8 measures the rms mass fluctuations in spheres of radius $8 h^{-1} \text{ Mpc}$ linearly extrapolated to the present day.

the SMBHs sink to the centre by dynamical friction from distant stars or by viscous effects from the surrounding gas. The transition to a bound binary state after the galaxy merger is an open issue (Milosavljević & Merritt 2001). However, it is believed that once the separation of the two BHs becomes small enough, gravitational radiation carries away the remaining angular momentum of the binary. The removal of energy from a SMBH binary due to gravitational wave emission leads to a gradual shrinkage of the relative separation of the two members. There is a point where the eccentricity reaches zero and the orbit circularizes. At that time, the two BHs are very close to each other, and gravitational waves are emitted copiously. The radiated energy is so large that two SMBHs, which are a few au apart, lose all their potential energy within a couple of minutes and inevitably coalesce. After the coalescence is complete, the binary enters the *ringdown* phase, where the merged members settle into a quiescent remnant hole.

A third channel for SMBH growth in our model is provided by diffuse gas in dark matter haloes undergoing quasi-hydrostatic cooling. When a massive halo collapses gas is shock heated out to a radius comparable to the virial radius of the dark matter halo. These haloes have a cooling time that is longer than the free-fall time of the gas and, thus, the gas settles into a quasi-static atmosphere surrounding the galaxy rather than simply falling towards the centre (White & Frenk 1991). This atmosphere – the ‘hot halo’ regime – is in pressure-supported hydrostatic equilibrium and extends beyond the virial radius of the dark matter halo. The galaxy is then supplied with cold gas by cooling flows at the disc centre, which also feed the central SMBH.

The formation of massive hot haloes would lead to the growth of very massive galaxies unless a heating mechanism regulates the cooling flow. Bower et al. (2006) and Croton et al. (2006) invoke energy injection from the central SMBH during the so-called ‘radio mode’ feedback (see also De Lucia et al. 2006; Cattaneo et al. 2007; Lagos, Cora & Padilla 2008). The radio mode is assumed to occur during the quiescent accretion of gas from the hydrostatically supported hot halo on to the SMBH, during which energy from the BH accretion is injected directly into the hot halo suppressing the cooling flow. In the Bower et al. (2006) model, the cooling flow stops when the power from the SMBH is sufficient to offset the rate at which energy is being radiated away.

Bower et al. (2006) adopt a BH growth model in which during a disc instability or galaxy merger, the BH accretes a fixed fraction of the gas that turns into stars in the burst taking into account processes such as feedback and recycling in the galaxy (Malbon et al. 2007). The amount of gas deposited on to the BH is set by an efficiency factor, which determines the fraction of the available gas reservoir for star formation that is accreted by the hole. The value of that parameter is chosen to fit the normalization of the local $M_{\text{bh}} - M_{\text{bulge}}$ relation. Hereafter, we refer to the accretion of cold gas triggered by disc instabilities and galaxy mergers as ‘quasar mode’ accretion and to the accretion from quasi-hydrostatic haloes as ‘radio mode’ accretion, following the terminology introduced by Croton et al. (2006).

2.2 Astrophysical processes affecting BH spin evolution

SMBHs are expected to possess angular momentum $J_{\text{bh}} = |a|GM_{\text{bh}}^2/c$, where a is the spin parameter, $0 \leq |a| \leq 1$. The spin has a significant impact in the close vicinity of the BH. For example, it determines the efficiency for converting matter into radiation in an accretion disc (Novikov & Thorne 1973) and it is believed to influence the formation and direction of the radio jets in AGN

(Blandford & Znajek 1977; MacDonald & Thorne 1982; Begelman et al. 1984). In addition, it is of special interest in the modelling of gravitational waves from BH–BH binaries (Merritt & Milosavljević 2005; Baker et al. 2006b, 2007; Campanelli et al. 2007; Buonanno, Cook & Pretorius 2007).

The evolution of spin is closely related to the channels that contribute to the growth of the BH. Each mechanism for BH growth is associated with different spin evolution. For example, accretion of gas that co-rotates with the BH should spin up the hole (Bardeen 1970), whereas the merger of two equal-mass non-spinning BHs results in a final spin of 0.69 (Baker et al. 2006a; Berti et al. 2007; Hinder et al. 2008). We explain below how these mechanisms that influence the spin of a BH are included in our model.

2.2.1 Gas accretion

After the SMBH seed forms at the centre of a galaxy, accretion usually initiates the growth era. We assume that an accretion disc is formed in the equatorial plane of the hole. As proposed by Lynden-Bell (1969), gas parcels in the disc gradually lose angular momentum due to viscous torques exerted by magnetic fields and drift radially inwards until they reach the inner edge of the accretion disc. The inner edge of the disc is usually taken as the location of the last stable orbit (LSO) around the BH. The LSO is a function of the hole's angular momentum and can be written as (Bardeen, Press & Teukolsky 1972)

$$\hat{r}_{\text{iso}} \equiv r_{\text{iso}}/R_g = \{3 + Z_2 \pm [(3 - Z_1)(3 + Z_1 + 2Z_2)]^{1/2}\}, \quad (1)$$

where the gravitational radius, R_g , is defined as half of the Schwarzschild radius of the BH, $R_{\text{Schw}} = 2R_g = 2GM_{\text{bh}}/c^2$, and Z_1, Z_2 are defined in terms of the spin, a , as

$$\begin{aligned} Z_1 &\equiv 1 + (1 - a^2)^{1/3} [(1 + a)^{1/3} + (1 - a)^{1/3}], \\ Z_2 &\equiv (3a^2 + Z_1^2)^{1/2}. \end{aligned} \quad (2)$$

Note that when the spin parameter has a negative sign the BH is counter-rotating with respect to the orbit of a particle around it. Then, for counter-rotating orbits ($-1 \leq a < 0$) $9 \leq \hat{r}_{\text{iso}} < 6$ and, for co-rotating orbits ($0 < a \leq 1$), $6 < \hat{r}_{\text{iso}} \leq 1$.

An important property is the binding energy of the gas at the LSO, defined as the difference between the rest energy of a gas parcel at infinity and its energy at the LSO as measured by an observer at infinity. If $\tilde{\epsilon}$ expresses the energy per unit rest mass, we can define the binding energy as $1 - \tilde{\epsilon}/c^2$. This provides a simple relation for the accretion efficiency,

$$\epsilon \equiv 1 - \tilde{\epsilon}_{\text{iso}}/c^2 = 1 - \sqrt{1 - \frac{2}{3} \frac{1}{\hat{r}_{\text{iso}}}} \quad (3)$$

(Novikov & Thorne 1973) which corresponds to the fraction of the energy released by matter spiralling in towards the BH through a succession of almost circular orbits. Equation (3) shows that accretion of matter on to slowly rotating BHs has moderate efficiency. For non-rotating BHs ($a = 0, \hat{r}_{\text{iso}} = 6$) this is just $1 - \sqrt{8/9}$ or 5.7 per cent. However, for co-rotating matter around rapidly rotating BHs, the efficiency increases significantly reaching $1 - 1/\sqrt{3}$ or 42.3 per cent as $a \rightarrow 1$. This sets an upper limit to the efficiency of the accretion on to BHs, a process obviously much more efficient than thermonuclear burning.

Once the gas reaches the LSO, we assume that it falls directly into the BH. In this way, the accretion carries into the BH the energy per unit mass, $\tilde{\epsilon}$, and angular momentum per unit mass, \tilde{l} , that the gas has at the LSO. Thus, accretion of a rest mass dM_0 leads to a

change in the hole's total mass, M_{bh} , and angular momentum, J_{bh} , equal to

$$dM_{\text{bh}} = (\tilde{\epsilon}_{\text{iso}}/c^2)dM_0, \quad dJ_{\text{bh}} = \tilde{l}_{\text{iso}}dM_0. \quad (4)$$

The change in the hole's spin induced by the accretion of dM_0 is governed by the differential equation

$$\frac{da}{d \ln M_{\text{bh}}} = \frac{1}{M_{\text{bh}}} \frac{c^3}{G} \frac{\tilde{l}_{\text{iso}}}{\tilde{\epsilon}_{\text{iso}}} - 2a. \quad (5)$$

This was integrated by Bardeen (1970) using the explicit expressions for $\tilde{\epsilon}$ and \tilde{l} , resulting in the following solution:

$$a^f = \frac{1}{3} \hat{r}_{\text{iso}}^{1/2} \frac{M_{\text{bh}}}{M_{\text{bh}}^f} \left[1 - \left\{ 3\hat{r}_{\text{iso}} \left(\frac{M_{\text{bh}}}{M_{\text{bh}}^f} \right)^2 - 2 \right\}^{1/2} \right] \quad (6)$$

when $M_{\text{bh}}^f/M_{\text{bh}} \leq \hat{r}_{\text{iso}}^{1/2}$. Here, M_{bh}^f and a^f are the final mass and spin parameter of the BH. If $M_{\text{bh}}^f/M_{\text{bh}} > \hat{r}_{\text{iso}}^{1/2}$, then the final spin is always equal to unity. The expression in equation (6) governs the evolution of a during accretion from an initial state of a co-rotating or counter-rotating disc. According to this, a non-rotating BH will be spun up to a maximum rotation ($a = 1, \hat{r}_{\text{iso}} = 1$) after increasing its mass by $1.44M_{\text{bh}}$ (see Fig. 1). For a maximally rotating BH in a counter-rotating accretion disc ($a = -1, \hat{r}_{\text{iso}} = 9$), an increase in mass of $2M_{\text{bh}}$ is required in order for the BH to be spun down to $a = 0$ and subsequently spun up to maximum rotation. As implied by equation (6), further accretion on to the BH keeps a equal to unity.

Bardeen's calculations are limited to the process of angular momentum transport between the accreted matter and the BH, without taking into account any other processes. In fact, Thorne (1974) argued that the accretion disc radiates and that some of this radiation will be accreted by the hole. Capture of photons with angular momentum opposite to that of the BH will then produce a counteracting torque that prevents spin up beyond the limiting value of $a = 0.998$ ($\hat{r}_{\text{iso}} = 1.23, \epsilon \simeq 0.32$). In addition, axial relativistic outflows of matter in the form of jets may have strong implications for the upper spin limit, since these outflows are accelerated by

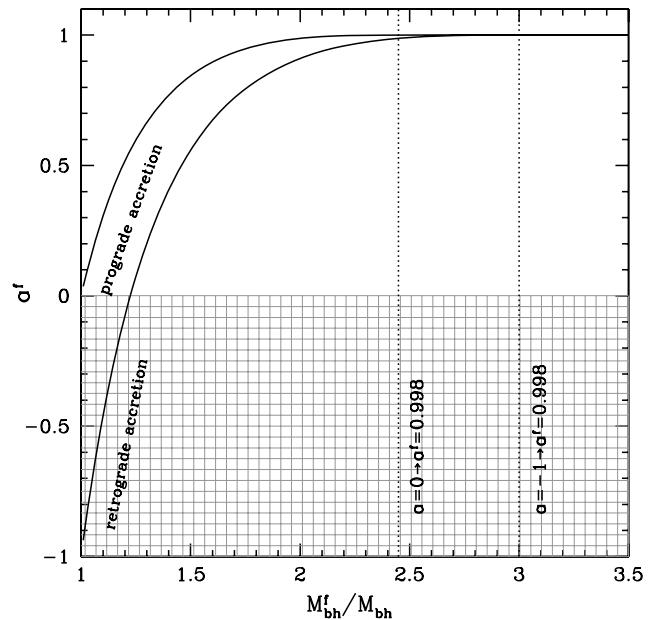


Figure 1. The final BH spin after the accretion of gas in a counter- and co-rotating configuration. The dotted lines indicate the final mass needed to spin up a non-rotating and maximally counter-rotating BH to $a^f = 0.998$.

magnetic fields that are powered by the extraction of the rotational energy of the BH (Hawley & Krolik 2006; Benson & Babul 2009).

2.2.2 Gas accretion through a misaligned disc

In the general case, where the accretion disc does not lie on the equatorial plane of the BH but has a random orientation relative to the angular momentum of the BH, the evolution of the spin is a complicated process. Following the discussion in King et al. (2005), we assume that a TD is inclined at some random angle θ relative to the orientation of the hole's angular momentum vector, \mathbf{J}_{bh} . We denote the angular momentum of the disc as \mathbf{J}_d (see King et al. 2005 and Volonteri, Sikora & Lasota 2007 for a discussion of the nature of \mathbf{J}_d), and define the total angular momentum vector \mathbf{J}_{tot} as

$$\mathbf{J}_{\text{tot}} = \mathbf{J}_{\text{bh}} + \mathbf{J}_d. \quad (7)$$

The angle θ between \mathbf{J}_{bh} and \mathbf{J}_d is defined such that $0 \leq \theta \leq \pi$; the values $\theta = 0$ and $\theta = \pi$ correspond to full alignment and anti-alignment respectively. \mathbf{J}_{tot} is a constant vector i.e. has fixed orientation as well as magnitude for a given accretion event. Its magnitude is given by

$$J_{\text{tot}}^2 = J_{\text{bh}}^2 + J_d^2 + 2J_{\text{bh}}J_d \cos \theta. \quad (8)$$

We further introduce the angle θ_t as the angle between \mathbf{J}_{bh} and \mathbf{J}_{tot} .

When the vectors \mathbf{J}_{bh} and \mathbf{J}_d are misaligned, the tilted orbits of the gas parcels in the accretion disc experience a torque due to the Lense–Thirring effect, which causes the plane of the accretion disc to precess about the rotational axis of the BH (Lense & Thirring 1918; Wilkins 1972; Bardeen & Petterson 1975; Scheuer & Feiler 1996). If viscosity is strong enough this can force the inner parts of the disc to rotate into the equatorial plane of the hole resulting in a ‘warped disc’ (see Fig. 2).

The Lense–Thirring torque can be expressed as

$$\frac{\partial \mathbf{L}}{\partial t} = \Omega_p \times \mathbf{L}, \quad (9)$$

where \mathbf{L} is the angular momentum per unit area of the disc and Ω_p is the precession rate and is given by

$$\Omega_p = \frac{2GJ_{\text{bh}}}{c^2 R^3}, \quad (10)$$

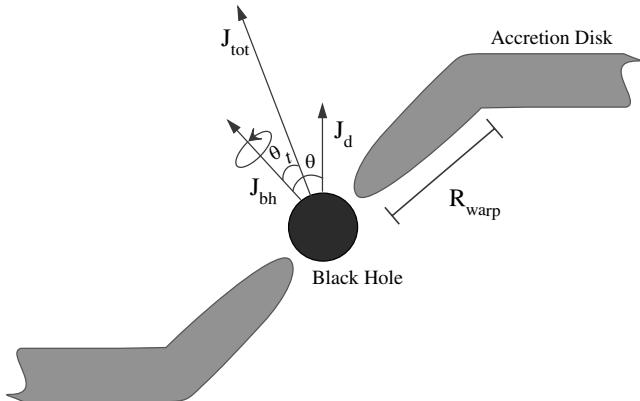


Figure 2. Schematic illustration of a warped accretion disc. \mathbf{J}_{bh} is the angular momentum of the BH ($|J_{\text{bh}}| = |a|GM_{\text{bh}}^2/c$), \mathbf{J}_d is the angular momentum of the disc given by equation (20) and \mathbf{J}_{tot} represents the total angular momentum of the system, $\mathbf{J}_{\text{bh}} + \mathbf{J}_d$.

where R is the distance from the BH (Pringle 1992). The precession time-scale is therefore defined as

$$t_{\text{prec}} \equiv \frac{2\pi}{\Omega_p(R)}, \quad (11)$$

which is proportional to R^3 and thus is much shorter closer to the BH. Other time-scales relevant to this problem are the viscous time-scales of accretion, t_{v_1} , and warp propagation, t_{v_2} ,

$$t_{v_{1,2}} \equiv \frac{R^2}{v_{1,2}(R)}. \quad (12)$$

Here, v_1 and v_2 are the kinematic viscosities acting on velocity gradients parallel and normal to the plane of the disc, respectively. The balance between the time-scales t_{prec} and t_{v_2} determines whether the Lense–Thirring torque is able to align the inner disc with the spin axis. The condition for alignment is $t_{\text{prec}} \lesssim t_{v_2}$ (Natarajan & Armitage 1999). In other words, the disc will be aligned with the spin at radii where the precession time-scale is much shorter than the time-scale of radial diffusion of the warp. The characteristic radius, R_{warp} , of the aligned part of the disc follows from the condition $t_{\text{prec}} = t_{v_2}$.

The evolution of \mathbf{J}_{bh} and \mathbf{J}_d during the precession is determined by the conditions (King et al. 2005)

$$\frac{d}{dt} J_{\text{bh}}^2 = 0, \quad \frac{d}{dt} J_d^2 \leq 0 \quad \text{and} \quad \frac{d}{dt} \cos \theta_t \geq 0. \quad (13)$$

As can be inferred by the first two conditions, the magnitude of \mathbf{J}_{bh} remains constant while the angle θ_t decreases with time, which implies that \mathbf{J}_{bh} always aligns with \mathbf{J}_{tot} . In contrast, \mathbf{J}_d decreases in magnitude as \mathbf{J}_{bh} aligns, which is to be expected, since the total angular momentum has to remain constant.

The end result of the Lense–Thirring precession is a BH which is aligned or anti-aligned with the surrounding accretion disc. The expression for the magnitude of \mathbf{J}_{tot} allows us to examine the final configuration of the system. Obviously, if $J_{\text{bh}}^2 > J_{\text{tot}}^2$ anti-alignment occurs, which requires

$$\cos \theta < -\frac{J_d}{2J_{\text{bh}}}. \quad (14)$$

Hence, a BH with $\theta > \pi/2$ and $2J_{\text{bh}} > J_d$ eventually anti-aligns with the accretion disc. This means that if $J_d > 2J_{\text{bh}}$ then alignment always happens as the disc angular momentum completely overwhelms that of the BH.

Further investigation of the warped discs requires the knowledge of the accretion flow properties. We use the TD solution of Shakura & Sunyaev (1973) for our analysis. In the standard Shakura–Sunyaev disc model, the analytic expression for the warp radius depends on the values of J_{bh} , M_{bh} and the accretion rate \dot{M} of the BH. In terms of the Schwarzschild radius, R_{warp} is written as (Volonteri et al. 2007)

$$\begin{aligned} \frac{R_{\text{warp}}}{R_{\text{Schw}}} &= 3.6 \times 10^3 a^{5/8} \left(\frac{M_{\text{bh}}}{10^8 M_{\odot}} \right)^{1/8} \lambda^{-1/4} \\ &\times \left(\frac{v_2}{v_1} \right)^{-5/8} \alpha^{-1/2}. \end{aligned} \quad (15)$$

Here α is the Shakura–Sunyaev viscosity parameter and $\lambda = L/L_{\text{Edd}}$ is the Eddington ratio. The accretion luminosity, L , and Eddington luminosity, L_{Edd} , are defined as

$$L = \epsilon \dot{M} c^2, \quad (16)$$

with ϵ denoting the accretion efficiency, and

$$L_{\text{Edd}} = \frac{4\pi G M_{\text{bh}} c}{\kappa} = 1.4 \times 10^{46} \left(\frac{M_{\text{bh}}}{10^8 M_{\odot}} \right) \text{ erg s}^{-1}, \quad (17)$$

where $\kappa \sim 0.3 \text{ cm}^2 \text{ g}^{-1}$ is the electron scattering opacity. Associated with the Eddington luminosity is an accretion rate expressed as $\dot{M}_{\text{Edd}} = L_{\text{Edd}}/\epsilon c^2$. This is the accretion rate for which the BH radiates at the Eddington luminosity. It is then convenient to express the physical accretion rate in units of the Eddington accretion rate, $\dot{m} \equiv \dot{M}/\dot{M}_{\text{Edd}}$. We note that the accretion efficiency in \dot{M}_{Edd} is spin dependent; however, we chose to keep $\epsilon = 0.1$ when calculating \dot{M}_{Edd} for simplicity.

Given that the accretion is characterized by a time-scale t_{v_1} , the mass of the disc inside the radius R_{warp} is

$$M_d(R_{\text{warp}}) = \dot{M} t_{v_1}(R_{\text{warp}}), \quad (18)$$

where the accretion time-scale is given by

$$\begin{aligned} t_{v_1} &= \frac{R_{\text{warp}}^2}{v_1} = 3 \times 10^6 a^{7/8} \left(\frac{M_{\text{bh}}}{10^8 M_{\odot}} \right)^{11/8} \\ &\times \lambda^{-3/4} \left(\frac{v_2}{v_1} \right)^{-7/8} \alpha^{-3/2} \text{ yr}. \end{aligned} \quad (19)$$

We can write the total angular momentum, J_d , passing through R_{warp} as

$$J_d(R_{\text{warp}}) \lesssim M_d(R_{\text{warp}})(GM_{\text{bh}}R_{\text{warp}})^{1/2}. \quad (20)$$

In terms of the anti-alignment criterion this gives

$$\begin{aligned} \frac{J_d}{2J_{\text{bh}}} &= \frac{M_d}{aM_{\text{bh}}} \left(\frac{R_{\text{warp}}}{R_{\text{Schw}}} \right)^{1/2} \\ &= 10^{-9} \lambda \left(\frac{t_{v_1}}{1 \text{ yr}} \right) \left(\frac{R_{\text{warp}}}{R_{\text{Schw}}} \right)^{1/2} a^{-1}. \end{aligned} \quad (21)$$

Evaluation of this quantity determines whether or not the anti-alignment condition in a misaligned disc is satisfied.

2.2.3 The case of self-gravity limited discs

The BH growth process in AGN environments involves vast amounts of accreted gas, often comparable to the initial mass of the accreting hole. This amount of mass settles on to the accretion disc around the hole, which often extends to several thousands of gravitational radii. It is usually assumed that the available mass fuel, M_{acc} , is consumed in a single accretion episode, thus providing a supply of constant angular momentum (Volonteri et al. 2007). In this case, the amount of mass consumed is enough to spin up the hole up to $a = 0.998$, even if the BH initially was maximally spinning in a counter-rotating direction. Inevitably, repeated accretion episodes during major galaxy mergers act to spin up the BH to maximum rotation.

Recently, however, King, Pringle & Hofmann (2008) argued that the end result of the accretion growth channel might be completely different if we take into account the fact that an accretion disc becomes self-gravitating at some radius, R_{sg} , where its mass exceeds $M_{\text{sg}} \sim (H/R)M_{\text{bh}}$. As a result, the mass of the disc is limited by its self-gravity to $\Delta M_{\text{episode}} \ll M_{\text{acc}}$, which gives rise to a series of $N \simeq M_{\text{acc}}/\Delta M_{\text{episode}}$ well-separated accretion episodes. King et al. suggest that these accretion episodes are randomly oriented around the BH, an assumption supported by observations indicating that there is no apparent relation between the accretion disc (or radio jet) orientation and the host galaxy disc (Nagar & Wilson 1999; Kinney et al. 2000). They further argue that this could be the result of intense star formation outside R_{sg} randomizing the input gas direction, which could provide a qualitative explanation for the ring of stars seen in the near vicinity of the central BH in the Milky Way (Genzel et al. 2003).

The much smaller angular momentum associated with each accretion episode means that, in general, $J_d < 2J_{\text{bh}}$, so in the chaotic accretion model counter-alignment occurs in a fraction (King et al. 2005)

$$f = \frac{1}{2} \left(1 - \frac{J_d}{2J_{\text{bh}}} \right) \simeq 1/2 \quad (22)$$

of the accretion episodes. Therefore, counter- and co-alignment are equally likely outcomes of the Lense–Thirring effect. However, accretion of gas in a counter-rotating disc is more efficient in spinning up the BH since the gas is being dumped on to the BH from a larger distance and, thus, carries more angular momentum into the BH. Hence, a succession of counter-aligned and co-aligned accretion episodes with equal frequency should systematically spin down the BH, resulting in a global spin distribution oscillating around zero.

The physics of self-gravitating accretion discs is described by Pringle (1981). Briefly, the criterion that the self-gravity of a disc be negligible is the requirement that the gravitational force along the \hat{z} direction be dominated by the central BH. This can be expressed as the surface density of the disc, Σ , being negligible compared to the quantity $M_{\text{bh}}H/R^3$, or in terms of the disc mass

$$M_d(< R) \ll \frac{H}{R} M_{\text{bh}}. \quad (23)$$

The self-gravity becomes marginally important when $M_d \simeq (H/R)M_{\text{bh}}$, which defines the self-gravity radius, R_{sg} , given by

$$\frac{R_{\text{sg}}}{R_{\text{Schw}}} = 1.5 \times 10^3 \epsilon^{8/27} \left(\frac{M_{\text{bh}}}{10^8 M_{\odot}} \right)^{-26/27} \lambda^{-8/27} \alpha^{14/27}. \quad (24)$$

The semi-thickness of the disc in the Shakura–Sunyaev model is given by

$$\begin{aligned} \frac{H}{R} &= 1.36 \times 10^{-3} \epsilon^{-1/5} \left(\frac{M_{\text{bh}}}{10^8 M_{\odot}} \right)^{-1/10} \lambda^{1/5} \\ &\times \left(\frac{R}{R_{\text{Schw}}} \right)^{1/20} \alpha^{-1/10}. \end{aligned} \quad (25)$$

Hence, by replacing R with R_{sg} in expression (25), we obtain an analytic expression for the disc mass within R_{sg} ,

$$\begin{aligned} M_{\text{sg}} &= \frac{H}{R} M_{\text{bh}} \Big|_{R=R_{\text{sg}}} = 2.13 \times 10^5 \epsilon^{-5/27} \\ &\times \left(\frac{M_{\text{bh}}}{10^8 M_{\odot}} \right)^{23/27} \lambda^{5/27} \alpha^{-2/17} M_{\odot}. \end{aligned} \quad (26)$$

As we described in Section 2.2.2, once the disc forms with a non-zero misalignment angle about the spin axis of the hole, it will be subject to the Lense–Thirring precession. If the radius, R_{sg} , is greater than R_{warp} , then

$$J_d = J_d(< R_{\text{warp}}). \quad (27)$$

If, however, $R_{\text{sg}} < R_{\text{warp}}$, then the entire disc will be subject to Lense–Thirring precession, thus aligning itself in the equatorial plane of the BH.

2.2.4 BH–BH binary coalescence

We now examine the evolution of spin during BH binary coalescence. During the merger of the two binary members, the smaller BH adds its spin and orbital angular momentum at the LSO to the spin of the larger one. Even if the progenitor holes do not possess any orbital angular momentum, the final remnant will preserve the

residual orbital momentum of the binary (i.e. the angular momentum that has not been radiated away). Thus, the remnant will always be a Kerr BH.

Recent breakthroughs in numerical relativity have provided robust simulations of BH mergers by directly solving the Einstein equations in a fully relativistic frame (Campanelli et al. 2007; Herrmann et al. 2007; Marronetti et al. 2007). These simulations have explored the parameter space for different configurations of initial masses and spins, allowing accurate measurements of the final spin. For example, a merger of two equal mass BHs both with $a = 0$ in a circular orbit results in a Kerr BH with $a^f \simeq 0.69$ (Baker et al. 2006a). This is a robust prediction and is also valid for eccentric orbits with eccentricities smaller than 0.4 (Hinder et al. 2008).

Analytic fits extend the predictions for a^f to the entire space of parameters and reproduce closely all the available numerical data. In the case where the masses are unequal but the spins are zero, the final spin can be estimated by the analytic expression

$$a^f \simeq 2\sqrt{3} \frac{q}{(1+q)^2} - 2.029 \frac{q^2}{(1+q)^4}, \quad (28)$$

where $q = M_2/M_1 \leq 1$ is the mass ratio of the progenitor BHs (Berti et al. 2007).

A more general analytic fitting formula for predicting the final spin of any given binary configuration has been provided by Rezzolla et al. (2008a,b). In their analysis, they assume that the final spin parameter vector can be expressed as

$$\mathbf{a}^f = \frac{1}{(1+q)^2} (\mathbf{a}_1 + \mathbf{a}_2 q^2 + \boldsymbol{\ell} q), \quad (29)$$

where $\mathbf{a}_{1,2} = c J_{bh}^{1,2}/(GM_{1,2}^2)$ and $\boldsymbol{\ell} = \boldsymbol{\ell}'/(M_1 M_2)$, with $\boldsymbol{\ell}'$ defining the difference between the orbital angular momentum, \mathbf{l} , when the binary is widely separated, and the angular momentum radiated away in gravitational waves before the merger, \mathbf{j}_{rad} , namely

$$\boldsymbol{\ell}' = \mathbf{l} - \mathbf{j}_{rad}. \quad (30)$$

The vector $\boldsymbol{\ell}'$ is taken to be parallel to the orbital momentum vector throughout the evolution of the binary, an assumption which is not strictly valid since the system could radiate away angular momentum in a non-symmetric way. However, the error introduced by this assumption is relatively small for the binary configurations studied in the simulations of Rezzolla et al. A further assumption is that the mass radiated away in gravitational waves is negligible, as it accounts for only a small fraction (5–7 per cent) of the total mass energy of the binary configurations analysed. Under these assumptions, Rezzolla et al. (2008a,b) proposed an analytic expression for the magnitude of the final spin, given by

$$|a^f| = \frac{1}{(1+q)^2} \left[|\mathbf{a}_1|^2 + |\mathbf{a}_2|^2 q^4 + 2\mathbf{a}_2 \mathbf{a}_1 q^2 \cos \varphi + 2(\mathbf{a}_1 \cos \vartheta + \mathbf{a}_2 q^2 \cos \xi) \boldsymbol{\ell} q + \boldsymbol{\ell}^2 q^2 \right]^{1/2}, \quad (31)$$

with the cosine angles φ , ϑ and ξ defined as

$$\cos \varphi \equiv \hat{\mathbf{a}}_1 \cdot \hat{\mathbf{a}}_2, \quad \cos \vartheta \equiv \hat{\mathbf{a}}_1 \cdot \hat{\boldsymbol{\ell}}, \quad \cos \xi \equiv \hat{\mathbf{a}}_2 \cdot \hat{\boldsymbol{\ell}}. \quad (32)$$

The norm of $\boldsymbol{\ell}$ is given by

$$|\boldsymbol{\ell}| = \frac{s_4}{(1+q^2)^2} (|\mathbf{a}_1|^2 + |\mathbf{a}_2|^2 q^4 + 2|\mathbf{a}_1||\mathbf{a}_2|q^2 \cos \varphi) + \left(\frac{s_5 \mu + t_0 + 2}{1+q^2} \right) (|\mathbf{a}_1| \cos \vartheta + |\mathbf{a}_2| q^2 \cos \xi) + 2\sqrt{3} + t_2 \mu + t_3 \mu^2. \quad (33)$$

Here μ expresses the symmetric mass ratio, $\mu = q/(1+q)^2 = M_1 M_2 / (M_1 + M_2)^2$, and the coefficients take the values $s_4 = -0.129$, $s_5 = -0.384$, $t_0 = -2.686$, $t_2 = -3.454$, $t_3 = 2.353$. The final spin as given by equation (31) is in good agreement with numerical data, with residuals of less than 3 per cent.

2.3 Simulating the spin evolution

We finally describe our method for calculating the evolution of BH spin. During the formation history of the BHs we take into account spin changes due to the accretion of gas and mergers, as described in Section 2.2. The initial population of seeds is assumed to be non-rotating. These seeds grow through accretion of gas during disc instabilities, galaxy mergers, and radio mode accretion and through BH–BH mergers.

2.3.1 Thin discs

To model the physics of the accreted gas, we assume that when $\dot{m} \geq 0.01$ the gas forms an accretion disc around the BH whose physics is described by the standard Shakura–Sunyaev disc model. We assume that the gas available to be fed into the BH after a galaxy merger or the collapse of a dynamically unstable disc is accreted over a time-scale proportional to the dynamical time-scale of the galactic bulge that hosts the BH. A fixed proportionality factor of 4 is used in this analysis resulting in typical accretion time-scales of the order of 10^7 – 10^8 yr. Note that this factor is ~ 10 longer than the one chosen by Malbon et al. (2007). The new value is introduced in order to match the local quasar luminosity function and does not affect the galaxy formation model. The time-scale for accretion during the radio mode is computed directly by the galaxy formation model and depends on the cooling time-scale of the gas in the hot halo.

The accretion disc is assumed to be randomly oriented relative to the spin axis of the hole. If $\theta \neq 0^\circ, 180^\circ$, we assume that the disc precesses around the BH as described in Section 2.2.2. In brief, for a given BH mass, M_{bh} , we determine the numerical values of the warp parameters, R_{warp} , $M_d(R_{warp})$, and the angular momenta $J_d(R_{warp})$ and J_{bh} . The ratio $J_d/2J_{bh}$ allows us, through the criterion in equation (14), to check if anti-alignment or alignment occurs. We then allow the BH to accrete an amount of gas equal to $M_d(R_{warp})$ and evolve the spin using equation (6). Finally, we update the mass of the BH and estimate the hole’s new spin and the disc precession by calculating the new angles θ and θ_r . This process is repeated until: (a) the disc aligns or counter-aligns ($\theta = 0^\circ$ or 180°) with the hole’s spin, in which case we just consume the rest of the available gas and evolve the spin according to equation (6), or (b) the accretion disc is entirely consumed, without being able to align or counter-align itself with the spin. In almost all cases we find that the accretion disc ultimately aligns (or anti-aligns) itself on the equatorial plane of the BH.

In addition to the case where the entire gas reservoir is consumed within one accretion episode with constant angular momentum orientation (prolonged accretion), we test the model proposed by King et al. (2008) in which the mass of the accretion disc is limited by its self-gravity (chaotic accretion). We repeat the same steps as before. However, in this case $M_d = M_{sg}$. Once M_{sg} is consumed, we update the mass of the hole and determine the mass of the new disc, M'_{sg} , computed from the updated BH mass. At this point we make the assumption that the disc retains no memory of the angular momentum of the initial flow and that it settles into a random orientation

around the BH with $M_d = M'_{sg}$. We then add the next M_{sg} until the disc is consumed.

It is important to clarify at this point that the two accretion models considered here serve as a means for studying the effect of accretion on BH spin. The forthcoming analysis that follows and results do not rely on any specific model for the physical processes that drive the gas flows towards the vicinity of the BH. Several mechanisms for feeding the SMBH with gas could result in similar spin distributions to the chaotic accretion case considered here (see e.g. Heller, Shlosman & Englmaier 2001 for the properties of non-self-gravitating gaseous bars in barred disc galaxies and Hobbs et al. 2010 for an investigation of the effect of supernova-driven turbulence on the fuelling of SMBHs in galactic bulges). Here, our main goal is to explore the effect of different SMBH spin distributions on properties of the host galaxy.

2.3.2 ADAFs

Many of the accretion events in our simulations are characterized by low, sub-Eddington, accretion rates, with $\dot{m} \leq 0.01$. For such low accretion rates, the gas flow has low density and is unable to cool efficiently since radiative cooling does not balance the energy generated by viscosity. Thus, the viscous energy is trapped in the gas as entropy and ultimately advected into the hole. This type of accretion is known as an ADAF (Rees 1982; Narayan & Yi 1994; Abramowicz et al. 1995).

ADAFs have a number of distinct properties that some of them will be essential for the analysis in later sections (see Sections 5 and 6). For example, for an ADAF around a BH, only a fraction of the standard accretion luminosity, $L = \epsilon \dot{M} c^2$, is emitted as radiation. The remainder of the viscously dissipated energy is stored in the gas as entropy, resulting in hot flows with almost virial temperatures. We note that, as shown by Ichimaru (1977), the ions and the electrons in an ADAF are not thermally coupled and, thus, reach different temperatures. This two-temperature virialized plasma flow is optically thin and, for high-viscosity parameters ($\alpha \sim 0.2\text{--}0.3$), it acquires a quasi-spherical geometry around the BH ($H \sim R$), which resembles spherical Bondi accretion. However, despite the geometrical similarity, the dynamics of the ADAF are fundamentally different, since the accretion is entirely due to dissipation via viscous forces rather than gravity.

In general, the accretion flow can have a rather complicated structure. For example, it is possible that a TD dominates the geometry of the outer parts, then switches to an ADAF as the flow approaches the BH (Esin et al. 1997). The transition to the thick flow depends on the accretion rate and usually occurs closer to the BH for higher accretion rates. Such a complex disc model for describing the accretion flows during the radio mode is, however, beyond the scope of this paper. Here, we assume a simple configuration of a quasi-spherical corona around the hole with any transition to a TD occurring far from the BH. Precession effects due to misalignment between the disc and the BH are modelled as in the TD case.

2.3.3 Binary mergers

Finally, we consider the spin changes due to binary coalescence. Following the merger of two galaxies in GALFORM, each harbouring a SMBH, the one hosted by the satellite sinks towards the SMBH of the central galaxy and eventually forms a binary. In our model, BH mergers tend to occur in gas-poor environments, and thus we assume that torques from accreting flows that might be present

during the formation of the binary do not influence the orientation of the BH spins prior to the merger. We therefore take the BH spin vectors in equation (31) to be randomly oriented relative to each other and the orbital angular momentum. In addition, we assume that all binaries will ultimately merge on a very short time-scale, effectively right after the host galaxies merger and induced accretion have been completed.

During the merger, the smaller BH plunges into the larger one carrying along its angular momentum at the LSO. At this point we neglect any mass loss due to gravitational wave emission, as it corresponds to a small fraction of the total mass energy of the system. In addition, we do not take into account the effects of gravitational recoil due to asymmetric emission of gravitational waves in binary mergers. In our model, mergers with recoil velocities as high as $v_{\text{recoil}} \sim 500 \text{ km s}^{-1}$ (Libeskind et al. 2006) could displace BH remnants. However, we assume that dynamical friction relocates them in the galactic centre. Therefore, we do not expect gravitational recoil to have a significant impact on the spin distribution of the remnant BHs. The spin of the final remnant is finally estimated using the fitting formulae in equations (31) and (33), as described in Section 2.2.4.

3 EVOLUTION OF BH MASS AND SPIN

3.1 Evolution of BH mass

We now briefly present our main predictions for BH growth, using our updates (see Section 2) to the model developed by Bower et al. (2006) and Malbon et al. (2007). In Fig. 3 we show the fraction of all BHs that accrete at a given rate in the quasar and radio modes,

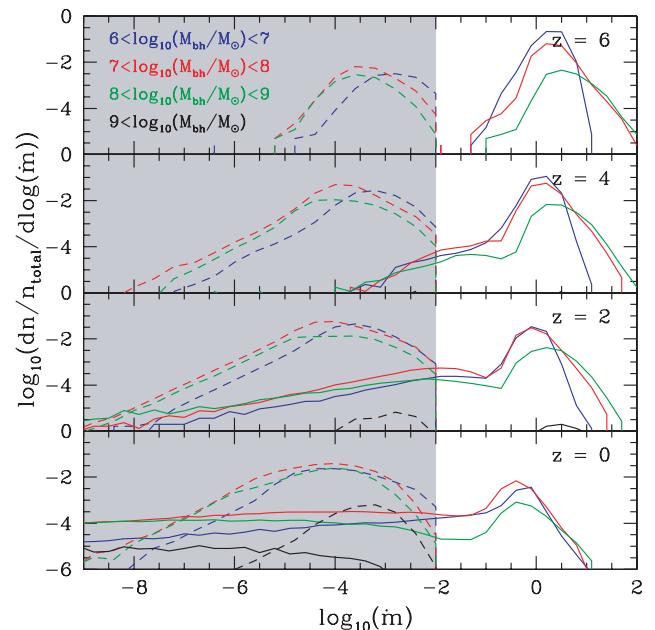


Figure 3. The distribution of accretion rates for selected BH mass ranges at various redshifts. Line styles represent two different accretion modes: solid lines for the quasar mode and dashed lines for the radio mode. The shaded area represents the regime where the accretion flow is described by an ADAF. The accretion rates in the radio mode are truncated at $\dot{m} = 0.01$, in accordance with our model of AGN feedback. In the present Universe, the quasar mode accretion peaks at $\dot{m} \simeq 0.5$. However, the peak shifts to higher values at higher redshifts, indicating that accretion was more efficient in the past.

at different redshifts. All BHs that accrete in the radio mode during the redshift bin are included, but the quasar mode is episodic so we only include objects which experience a starburst within a given redshift bin. Our updated parameters give a longer duration for the accretion episode in the quasar mode (four times rather than 0.5 times the dynamical time-scale of the bulge) than assumed in Malbon et al. (2007). This results in the majority of BHs with mass $10^6 - 10^8 M_\odot$ which accrete in the quasar mode at $z = 0$ having $\dot{m} \simeq 0.5$, as observed (Heckman et al. 2004). At higher redshift the distributions shift to higher mass accretion rates, with a mean $\dot{m} \simeq 3$ at $z = 6$. This implies that in a hierarchical universe, accretion of gas was more efficient at early epochs, so the rate of growth of BHs is faster in the past. By contrast, accretion activity during the radio mode peaks at very low accretion rates ($\dot{m} \approx 10^{-4}$) at $z = 0$, with a long tail to lower values together with some spread to the maximum allowed value of 0.01 (imposed by the AGN feedback model). This upper limit to the mass accretion rate limits the shift to higher \dot{m} at higher z , so the main evolutionary trend is that there are fewer very low \dot{m} radio mode objects at high z .

This model results in a *bimodal* mass accretion rate distribution for all BH masses and redshifts. Activity triggered by accretion of cold gas from the galaxy disc, replenished by galaxy mergers, has a mean mass accretion rate which is ~ 1000 times higher than that of activity fed by radio mode accretion. However, the relative importance of these two modes varies with BH mass. This is shown in more detail in Fig. 4 which shows the 20–80 and 5–95 percentile mass accretion rates as a function of BH mass at $z = 0$. More than 5 per cent of the lower mass AGN have accretion via a TD ($\dot{m} > 0.01$). However, this fraction drops with mass at BH masses of $\sim 10^8 M_\odot$. More massive galaxies are hosted by large ellipticals which are gas poor and are dominated by radio mode accretion.

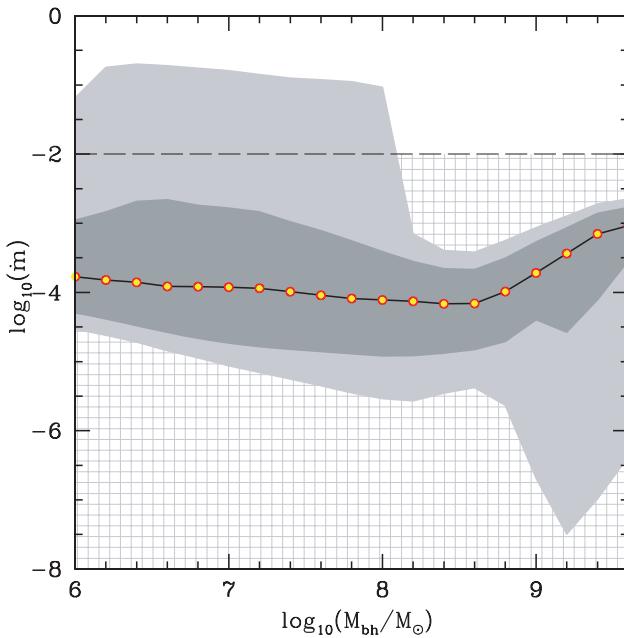


Figure 4. The correlation of accretion rate with BH mass at redshift zero. The median (red–yellow points), 20–80 percentiles (dark grey) and 5–95 percentiles (light grey) are shown. The shaded background indicates the region where the accretion disc is modelled as an ADAF. Only lower mass BHs ($< 10^8 M_\odot$), i.e. those hosted by lower mass galaxies (typically spirals), have enough gas left to trigger accretion via a TD ($\dot{m} > 0.01$). More massive BHs have little cold gas and undergo radio mode accretion which (by construction) has $\dot{m} < 0.01$.

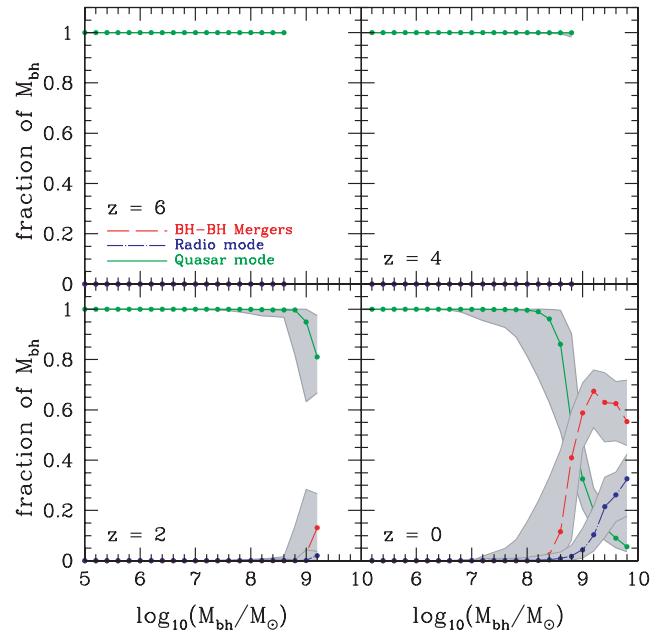


Figure 5. Contribution to the final BH mass at different redshifts from gas accretion in quasar mode (solid green line), radio mode (dashed blue line) and BH–BH mergers (dash-dotted red line) in the updated Bower et al. (2006) model. The lines represent the medians and the shaded region the 20–80 percentiles of the distributions. The dominant channel of BH growth for BHs with $M_{bh} \lesssim 10^8 M_\odot$ is the quasar mode. Above that mass, mergers and radio mode become the dominant growth channels.

Fig. 5 shows the contribution from the different growth channels to the final BH mass at different redshifts. At high redshifts, SMBHs with masses up to $\sim 10^8 M_\odot$ build their mass almost exclusively through the accretion of cold gas (quasar mode). This forms either a TD or an ADAF depending on \dot{m} (see also Fig. 3). Accretion of gas from the hot halo of massive galaxies during the radio mode always occurs with $\dot{m} \leq 0.01$ by construction, thus giving rise to an ADAF. This low upper limit to the mass accretion rate in this mode means that it makes comparatively little impact on BH growth, except for masses greater than $10^8 M_\odot$. Note that in the original Bower et al. (2006) model, the radio mode is responsible for building most of the BH mass above $10^8 M_\odot$, while in the updated version presented here, both BH mergers and radio mode contribute significantly for high-mass BHs.

In Fig. 6, we compare the relationship between the BH mass and host bulge mass predicted by our model to the observational data. The model predicts an almost linear relation and reproduces the observations reasonably well. The scatter in BH mass is ~ 1 dex for bulge masses below $10^{11} M_\odot$ and becomes gradually smaller for higher bulge masses. The resulting BH mass functions are shown in Fig. 7. These show that the present-day Universe is dominated by BHs with masses of $\sim 10^7 - 10^8 M_\odot$ which is in good agreement with the recent results from the Sloan Digital Sky Survey (SDSS) of Heckman et al. (2004). The shape of the mass functions is fairly similar at all redshifts. The main difference occurs at the high-mass end where the BH mass function extends to higher masses at low redshifts. For $M_{bh} \lesssim 10^8 M_\odot$ the mass function is almost constant and flat up to $z = 2$. This indicates that these BH are already in place at $z = 2$. Above that mass, there is a sharp decrease in the space density of BH at $z = 2$ and 4. At $z = 0$, the mass function has a shallower slope. The estimated BH mass density at $z = 0$ is $\rho_{bh} =$

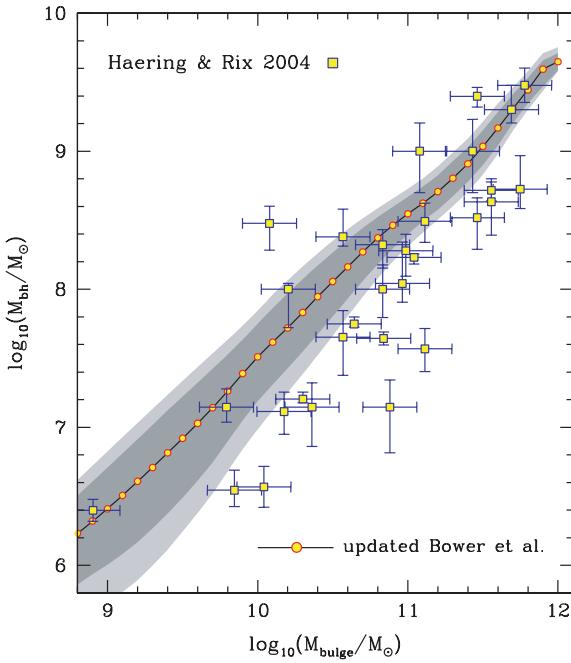


Figure 6. The $M_{\text{bh}} - M_{\text{bulge}}$ relation predicted by the updated Bower et al. (2006) model (solid line). The shaded areas indicate the 10–90 (light) and 20–80 (dark) percentile spread of the theoretical predictions. The observational data are taken from Häring & Rix (2004).

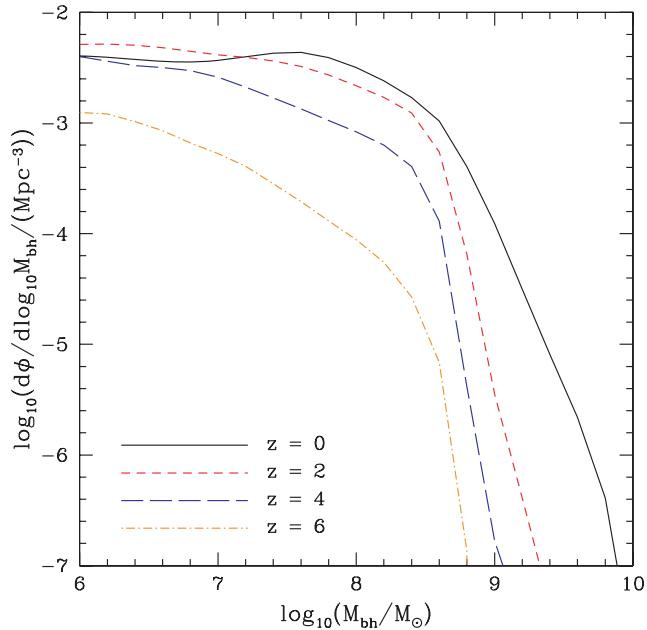


Figure 7. The predicted mass functions at $z = 0, 2, 4$ and 6 for the BH population in our simulation.

$3.84 \times 10^5 M_{\odot} \text{ Mpc}^{-3}$, which is in good agreement with the value implied by the X-ray background (Fabian & Iwasawa 1999).

3.2 Global BH spin distributions

The histograms in Fig. 8 show the distributions of BH spins at different redshifts (left-hand panel) and for different mass ranges (right-hand panel) for all galaxies that host BHs with $M_{\text{bh}} \geq 10^6 M_{\odot}$. The top and bottom panels correspond to the chaotic and prolonged

accretion models, respectively. The BH spin distributions predicted by the two accretion models are already established at $z = 6$ and do not change appreciably with time. The major channel for BH growth at $z = 6$ is the quasar mode and this generally takes place via a TD. The accreted mass in each episode is generally larger than the mass of the BH, so this spins it up to maximal in the prolonged accretion model. The break up of the accretion into much smaller individual events in the chaotic model leads to a random walk spin distribution around a low spin value. None the less, the plots in Fig. 5 show that BH–BH mergers contribute to the most massive BHs at late redshifts, as does accretion with $\dot{m} < 0.01$ in both the radio and quasar mode. In order to gain further insight into the effect of these growth channels we show the different spin distributions as a function of BH mass at $z = 0$ in the right-hand panels.

As already mentioned, low-mass BHs are built via thin-disc accretion in the quasar mode (Fig. 5) with mean $\dot{m} \sim 0.5$. For a BH of mass $10^6 M_{\odot}$ a typical accretion event in the quasar mode contains $10^6 M_{\odot}$ of gas, so in the prolonged accretion model this will spin the BH up to maximal. However, in the chaotic accretion model in which the size of the disc is limited by its self-gravity, the mass of the gas that settles on the disc cannot be more than $\sim (H/R) \times 10^6 M_{\odot} \sim 10^4 M_{\odot}$. Thus, a $10^6 M_{\odot}$ accretion event consists of ~ 100 accretion episodes with $\Delta M_{\text{episode}} \ll M_{\text{bh}}$, all randomly oriented around the BH. As a consequence, 10^6 – $10^8 M_{\odot}$ BHs experience a net spin-down to modest values centred around $a^f \sim 0.15$.

The growth channel for more massive BHs ($M_{\text{bh}} > 10^8 M_{\odot}$) includes a large fraction of objects which accrete in the radio mode or quasar mode with $\dot{m} < 0.01$, i.e. via a geometrically thick ADAF. However, we do not include full modelling of this process as our spin evolution calculations always assume a Shakura–Sunyaev TD rather than the appropriate ADAF equations. We do not expect this to have a large effect on the spin distributions as most of the BH mass and spin are built up from high-mass accretion rates.

The mergers between massive BHs have a significant impact on the final spin distributions. According to the analytic fitting formula, equation (31) from Rezzolla (2008, 2009), the post-merger spin of the final remnant depends strongly on the masses of the progenitors. For example, BHs acquiring a final spin greater than 0.69 after a merger are the end product of binaries of comparable mass ($q \simeq 1$) in which the members already had significant spins (recall that $a^f = 0.69$ is the final spin of a binary of equal-mass non-spinning BHs). Mergers between BHs of comparable mass are common only for the most massive BHs ($M_{\text{bh}} \gtrsim 10^8 M_{\odot}$) in our simulation.³ This is because these BHs are hosted by massive elliptical galaxies that have experienced at least one major merger (merger between galaxies of equal mass) in their past history (Parry et al. 2009). Hence, their BHs, which are of similar mass, will acquire a final spin close to 0.69 or higher (depending on the value of the initial spins) after they coalesce. In contrast to the most massive BHs, lower mass BHs ($M_{\text{bh}} \lesssim 10^8 M_{\odot}$) which experience a merger are involved mostly in minor mergers ($q \ll 1$). These do not have a significant

³ At redshift zero nearly 19 per cent of the BHs with $M_{\text{bh}} \geq 10^6 M_{\odot}$ have experienced at least one merger. The low merger rate of BHs merely reflects the relative minor effect that mergers have in the formation of galactic spheroids (except in very massive ellipticals) in the Bower et al. (2006) model, as found by Parry, Eke & Frenk (2009). For example, at redshift zero about 6 per cent of the BHs with $M_{\text{bh}} \geq 10^6 M_{\odot}$ have experienced more than one merger and only ~ 8 per cent have had a recent major merger.

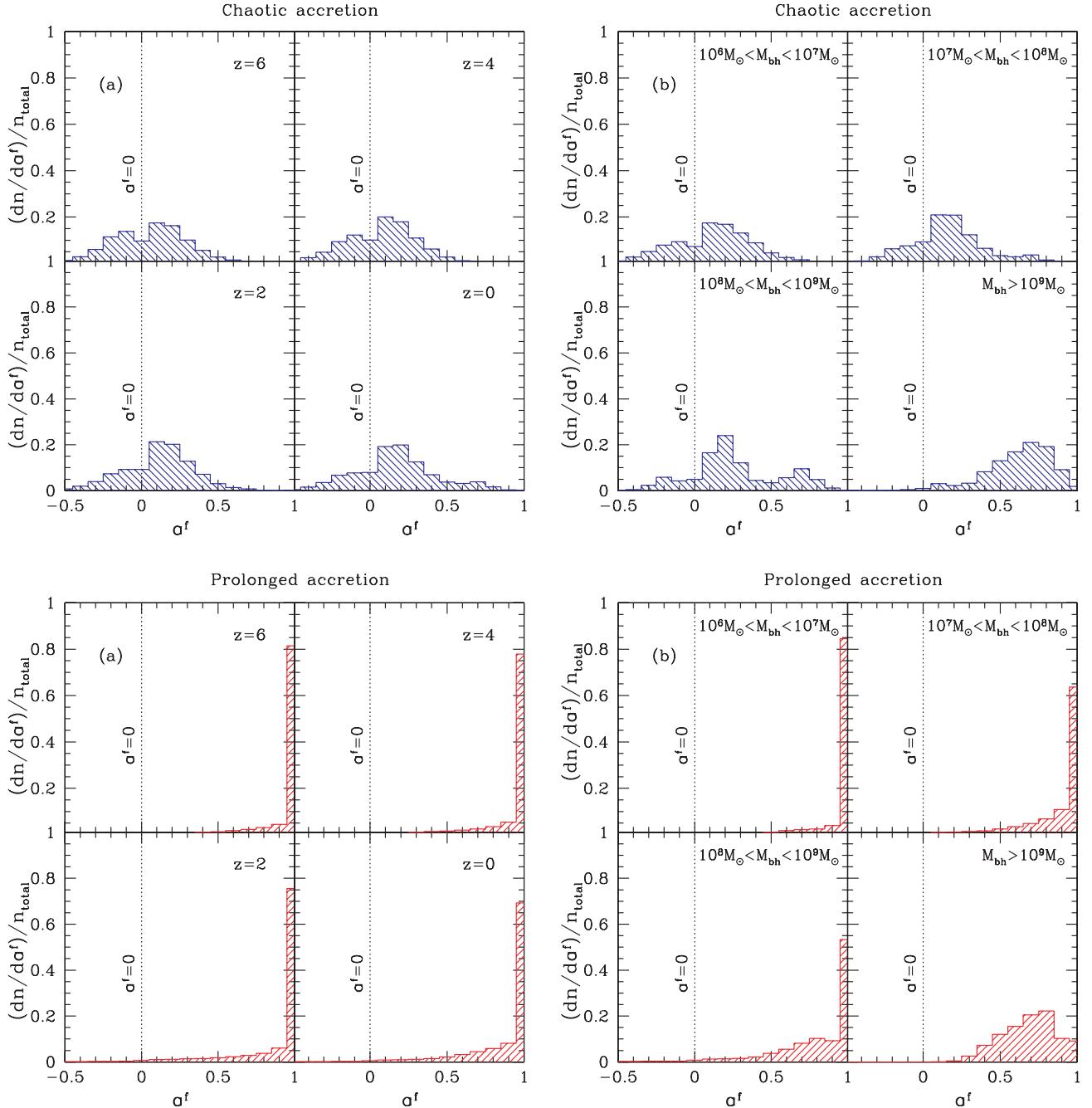


Figure 8. Model predictions for the final BH spin distributions. (a) The spin evolution due to BH mergers and accretion and (b) the distribution of BH spins at $z = 0$ in different mass bins in the chaotic (upper panels) and prolonged (lower panels) accretion models. The main difference between the two accretion modes is that the chaotic case results in low-mass BHs having spins distributed around zero, whereas the prolonged case results in a distribution peaked at around $a^f = 0.998$.

impact on the final spin of the remnant which is dominated by the more massive member of the binary.

When the effects on BH spin of mergers and accretion are combined, we find the following result. BHs of mass $\gtrsim 5 \times 10^8 M_\odot$ that have experienced a major merger will acquire spins greater than ~ 0.69 ; since these mergers often occur in gas-poor environments, the spin distribution for these masses is not significantly altered by the accretion. For lower mass BHs, accretion quickly erases the characteristic post-merger spins, because the growth is dominated by intense accretion during the quasar mode. For chaotic accretion,

this results in a bimodal distribution of spins for BHs with mass $10^8 - 10^{10} M_\odot$. The first peak is located at $a^f \sim 0.15$ and corresponds to BHs that have had a major merger accompanied by an accretion episode in which the post-merger spin is quickly erased and kept low. The second peak is located at $a^f \sim 0.7 - 0.8$ and corresponds to BHs that experienced a merger in the absence of accreting flows and acquired a final spin characteristic of post-merger BH remnants.

We note that BHs in our cosmological model have low merger rates at high redshift ($z \sim 6$). As redshift decreases, the merger rate increases and reaches a maximum at $z \sim 1.5$. Eventually, at $z = 0$

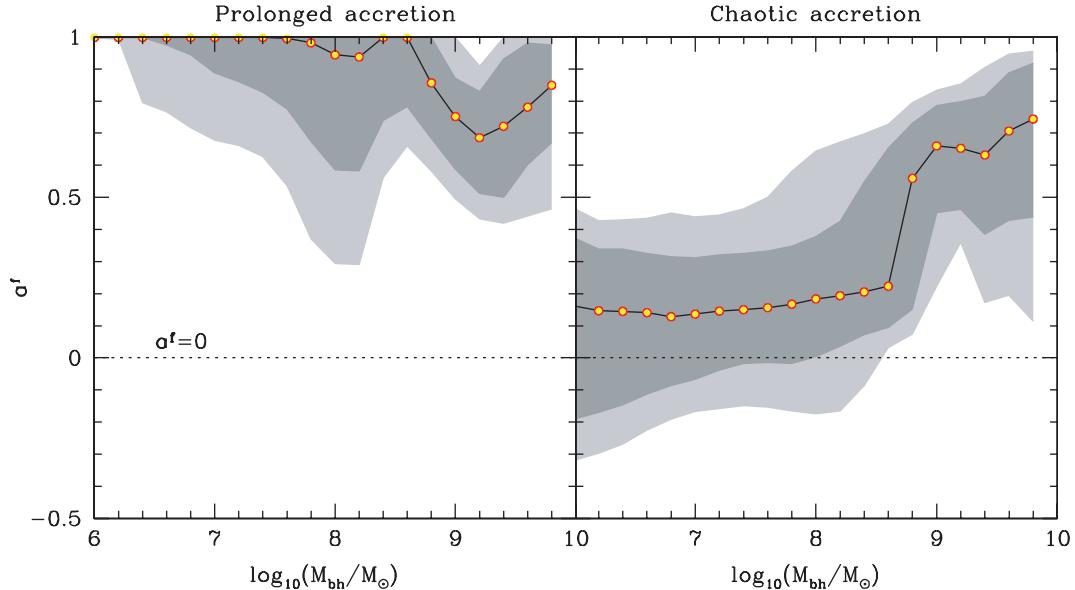


Figure 9. The correlation of spin with BH mass (a) for the prolonged and (b) the chaotic accretion models. The solid lines show the medians and the shaded areas the 10–90 (light) and 20–80 (dark) percentile spreads of the distributions.

nearly 19 per cent of the BHs with $M_{\text{bh}} \geq 10^6 M_{\odot}$ have experienced at least one merger. Therefore, the significance of mergers increases with decreasing redshift. This accounts for the decrease with redshift of the rapidly rotating BHs in the prolonged model.

We finally show in Fig. 9 the correlation between the BH mass and spin. This figure confirms the conclusion of Fig. 8: BHs that grow their mass through chaotic accretion display a strong correlation between their mass and spin. In this case we find that the BH mass correlates with the host-galaxy morphology such that small BHs are usually found in spiral galaxies, whereas massive BHs are found in massive ellipticals. Thus, the apparent correlation between BH mass and spin indicates a strong correlation between BH spin and host-galaxy morphology. As a consequence, we expect to find rapidly rotating BHs in principle at the centres of very massive elliptical galaxies. However, these are also the objects which have low mass accretion rates (see Fig. 4), and so predominantly accrete via an ADAF whereas some fraction of the lower mass, low spin BH can accrete via a TD. By contrast, in the case of prolonged accretion there is no apparent correlation between the BH mass and spin since most of the BHs exhibit very high spin independently of mass. Thus, rapidly rotating BHs are found in most types of galaxy. The correlation of mass and mass accretion rate (Fig. 4) remains the same, so the most massive ellipticals harbour the most massive BHs and accrete via an ADAF while a small fraction of the lower mass BHs in spirals can accrete via a TD. Note that the correlation between M_{bh} and spin for BH masses greater than $10^9 M_{\odot}$ is similar to that predicted by the chaotic model since these BHs include more growth through BH merger activity, which is independent of the accretion model.

4 OPTICAL LUMINOSITIES OF AGN AND THE QUASAR LUMINOSITY FUNCTION

We now explore whether our models are capable of explaining the observed properties of AGN. In particular, we present our predictions for the quasar luminosity function assuming that quasar activity is driven by accretion on to a SMBH. In our model, SMBHs with $\dot{m} > 0.01$ are assumed to accrete through a TD, whereas those

with accretion rates below this are assumed to accrete through an ADAF. In the thin-disc regime the bolometric luminosity is given by the standard expression

$$L = \epsilon \dot{M} c^2. \quad (34)$$

The accretion efficiency, ϵ , is taken to vary with the hole's spin according to equation (3). When the accretion rate becomes super-Eddington ($\dot{m} > 1$), the bolometric luminosity is limited to $(1 + \ln \dot{m}) L_{\text{Edd}}$ (Shakura & Sunyaev 1973). However, we do not restrict the accretion rate if the flow becomes super-Eddington. For the ADAFs, we assume that the plasma acquires a two-temperature configuration. The bolometric luminosity of the disc in this case is given by Mahadevan (1997)

$$L_{\text{bol,ADAF}} = 1.3 \times 10^{38} \left(\frac{M_{\text{bh}}}{M_{\odot}} \right) \left(\frac{\dot{m}^2}{\alpha^2} \right) \left(\frac{\beta}{0.5} \right) \text{ erg s}^{-1}, \quad (35)$$

where β is related to the Shakura–Sunyaev viscosity parameter α through the relation $\alpha \approx 0.55(1 - \beta)$ (Hawley, Gammie & Balbus 1995). The value of α is taken to be 0.1 for all objects in our samples.

Observationally, the quasar bolometric luminosity function can be estimated by combining measurements in many different wavelengths (Hopkins, Richards & Hernquist 2007; Shankar, Weinberg & Miralda-Escudé 2009). We can compare the predictions of our model to the estimated bolometric luminosity function simply by counting the number of quasars radiating at a given luminosity bin. In this analysis, however, we will attempt to calculate the optical b_J -band luminosity function since we are ultimately interested in comparing the strength of the radio jet emission of an AGN relative to its disc optical emission (see Section 6).

The conversion factor between bolometric accretion luminosity and b_J -band luminosity, f_{b_J} , is assumed to be constant and its value is adjusted to match the observed b_J -band luminosity function. We take $f_{b_J} = 0.08$ for the prolonged accretion and $f_{b_J} = 0.2$ for the chaotic accretion model. The substantial difference between the conversion factors for the two accretion models reflects the strong dependence of the bolometric luminosity on the spin through the accretion efficiency. An increase in spin means that the disc can extend closer in towards the BH, so the same mass accretion rate

produces more UV emission for the same b_J -band luminosity. Hence the b_J band carries a smaller fraction of the bolometric luminosity. We check our conversion factors derived from fitting the quasar luminosity function to those derived from a Novikov–Thorne (fully relativistic) version of the Shakura–Sunyaev disc equations, assuming that the b_J band is centred at 4400 Å, with a full width at half-maximum (FWHM) of 980 Å. We find that our assumption is in good agreement with the theoretical calculations as long as we assume that the disc luminosity is Eddington limited and similarly they are in good agreement with the values of 0.07–0.15 suggested by Elvis et al. (1994) from analysis of the SEDs of a large sample of quasars. Thus, we adopt these values for the rest of our calculations.

In deriving the luminosity function, we assume that only thin-disc-accreting objects have the intense ionizing flux which leads to their identification as quasars. We further assume that a fraction of 40 per cent of the total number of quasars is obscured in the b_J band by the torus (Polletta et al. 2008). Finally, absolute b_J magnitudes in the Vega system are obtained using the relation

$$M_{b_J} = -10.44 - 2.5 \log(L_{b_J}/10^{40} \text{ erg s}^{-1}). \quad (36)$$

Fig. 10 shows the resulting quasar luminosity function at redshift zero for both the prolonged-accretion model and the chaotic-accretion model. The models are compared to the observed luminosity functions in the redshift range $0.4 < z < 0.68$ estimated from the 2dF (Croom et al. 2004) and 2dF-SDSS quasi-stellar object (QSO) survey (Croom et al. 2009). The two models agree with each other, although the prolonged-accretion model predicts a somewhat higher volume density of quasars in the $-21 \leq M_{b_J} \leq -23$ and $-24 \leq M_{b_J} \leq -25$ range. The inflection at $M_{b_J} \simeq -24$ that sep-

rates the two regions is a reflection of the $10^8 M_\odot$ dip in the $M_{\text{bh}}-a$ correlation seen in Fig. 9. Both models predict steeper faint-end slope ($-20 \leq M_{b_J} \leq -22$) than the observed luminosity functions. A similarly steep faint end is also seen in the soft X-ray luminosity function at $z \sim 0.4$ (Miyaji, Hasinger & Schmidt 2001; Hasinger, Miyaji & Schmidt 2005). The differences between the two models in this regime arise because the BHs in the prolonged model that contribute to the faint end have much higher spins, and thus higher bolometric luminosities than in the chaotic model.

In conclusion, we can reproduce the optical luminosity function reasonably well with either the chaotic or prolonged accretion models. A consistency check is provided by calculating the average accretion efficiency. Assuming only the accreting sources contributing to the quasar luminosity function we find $\langle \epsilon \rangle = 14.2$ per cent ($\langle a \rangle = 0.86$) for the prolonged model and $\langle \epsilon \rangle = 6.1$ per cent ($\langle a \rangle = 0.12$) for the chaotic model. The latter is in good agreement with $\langle \epsilon \rangle = 6.7$ per cent inferred by optically selected AGN (Martínez-Sansigre & Taylor 2009) and consistent with the low efficiency that characterizes accretion at $z = 0$ suggested by Wang et al. (2009).

5 AGN RADIO LOUDNESS AND THE SPIN PARADIGM

In this section we investigate how the total radio output of an AGN is related to the accretion process and the central BH properties, and we present a model where we combine standard accretion disc theory (including ADAFs), BH evolution and radio jet production to understand the observed AGN radio loudness.

5.1 Modelling the jet emission in AGN

We construct a model for studying the radio loudness of AGN based on our current understanding of the formation and acceleration of jets. The starting point is the galaxy formation model described in Section 2. Galaxies become active every time the central SMBH experiences an accretion episode, allowing us to track the evolution of BH mass, spin and mass accretion rate distributions as described in the previous section. We combine this with the prescriptions for jet luminosity in the BZ model (Meier 2002). For consistency, we also investigate the more general case where the jet taps energy from both the BH and the accretion disc (Blandford & Payne 1982), which is described in the hybrid model presented by Meier (2001).

5.1.1 The BZ jet model

AGN jets in the BZ jet model are exclusively powered by extraction of the rotational energy of the BH and, thus, they are formed only in AGN hosting rotating SMBHs. The mechanical energy of a jet is proportional to the square of the poloidal magnetic field, B_{pol} , at the horizon of the BH (Blandford & Znajek 1977)

$$L_{\text{jet}} \propto B_{\text{pol}}^2 M_{\text{bh}}^2 a^2. \quad (37)$$

This is a second-order perturbative solution for the spin parameter a and is used to approximate the solution for slowly rotating BHs, $a \ll 0.998$. By including higher-order corrections it can be shown that the dependence of L_{jet} on a can be much steeper for rapidly rotating BHs ($L_{\text{jet}} \sim a^4$; Tanabe & Nagataki 2008; see also Tchekhovskoy, Narayan & McKinney 2010 for a recent study of the dynamical range of L_{jet} predicted by these solutions). For this analysis we will adopt the original approximation of Blandford & Znajek (1977) for simplicity.

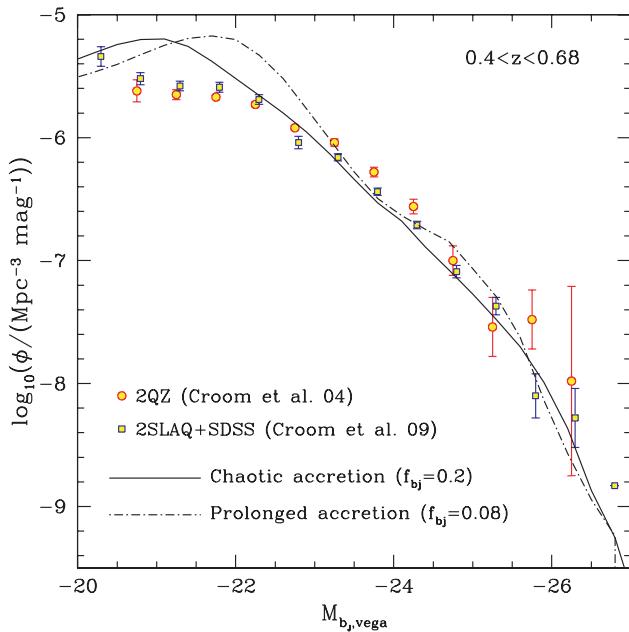


Figure 10. The model quasar luminosity function at $z = 0.5$ for the prolonged (dash-dotted line) and chaotic (solid line) accretion models. The luminosity function includes only AGN assumed to be powered by a TD ($m > 0.01$) that accrete in the quasar mode. The best fits to the observational data require that 8 per cent (prolonged model) and 20 per cent (chaotic model) of the bolometric luminosity be emitted in the b_J band. It further requires that 40 per cent of the total number of AGN be obscured in both accretion models. The quasar luminosity functions from the 2dF (Croom et al. 2004) and 2dF-SDSS QSO survey (Croom et al. 2009) over the range $0.4 < z < 0.68$ are shown with circles and boxes, respectively.

There have been several models that give the strength of B_{pol} around a rotating BH. In most cases, B_{pol} is expressed in terms of the azimuthal component, B_ϕ , as $B_{\text{pol}} \approx (H/R) B_\phi$. Therefore, the strength of the jet depends critically on whether the disc is geometrically thin or thick. For ADAFs ($H \sim R$) and TDs ($H \ll R$), this yields (see Meier 2002 and references therein)

$$L_{\text{jet,ADAF}} = 2 \times 10^{45} \left(\frac{M_{\text{bh}}}{10^9 M_\odot} \right) \left(\frac{\dot{m}}{0.01} \right) a^2 \text{ erg s}^{-1}, \quad (38)$$

$$L_{\text{jet,TD}} = 2.5 \times 10^{43} \left(\frac{M_{\text{bh}}}{10^9 M_\odot} \right)^{1.1} \left(\frac{\dot{m}}{0.01} \right)^{1.2} a^2 \text{ erg s}^{-1}. \quad (39)$$

In the super-Eddington regime ($\dot{m} \geq 1$), we assume that the flow remains in a TD state as there are as yet no models to describe the behaviour of the radio jet in this regime. The value of the viscosity parameter, α , in the ADAF and thin-disc regimes, is set to 0.1. Note that, according to equation (38) the mechanical jet power at the top of the ADAF branch is $L_{\text{jet}} = a^2 L_{\text{bol}} \sim 0.01 a^2 L_{\text{Edd}} \text{ erg s}^{-1}$. This upper limit to the jet luminosity in the ADAF regime is the origin of our revision to the fraction of the Eddington luminosity available for jet feedback into the halo (see Section 2). For low spin BHs, the BZ jet power drops substantially below this.

5.1.2 The hybrid jet model

In the BZ mechanism, the jet is assumed to be powered directly by the extraction of the rotational energy of the BH, neglecting any energy that could be extracted from the disc. However, the field lines frozen to the accreting matter on the disc may generate collimated outflows as a response to the differential rotation of the plasma, even in the case when the BH is not rotating (Blandford & Payne 1982). The outflow in this case is powered by the extraction of the rotational energy of the disc, rather than the BH. Thus, in the general case, both the disc and the hole could contribute to the production of jets.

Based on these considerations, Meier (2001) proposed a hybrid model for the radio loudness of AGN in which the jet luminosity depends on the rotation of both the BH and the accretion disc. The model includes two distinct accretion states, the ADAF and TD, in which the BH can be rotating or non-rotating. The hybrid model has a weaker dependence on the spin of the BH than the standard BZ mechanism. When $a \neq 0$ the jet luminosity scales with spin as

$$L_{\text{jet,ADAF}}^{\text{Kerr}} \propto (0.55 f^2 + 1.5 f a + a^2), \quad (40)$$

$$L_{\text{jet,TD}}^{\text{Kerr}} \propto (1 + 1.1 a + 0.29 a^2), \quad (41)$$

where f and g are parameters related to the angular velocity of the disc and the azimuthal magnetic field, respectively [see Meier 2001 for more details; also Nemmen et al. (2007) for a fully relativistic account of the hybrid model]. The model appears to be a more realistic interpretation of the jets seen from MHD simulations, since these are still able to power jets even from Schwarzschild BH accretion. We have tested this model by coupling it to our predicted spin distributions. However, we find that the weak dependence on spin in this model results in small dynamic range in radio luminosity, and the results are very similar to the predictions of the BZ jet model for the prolonged accretion case (see Section 6). We therefore prefer to contrast the prolonged accretion model with the chaotic accretion model coupled to the spin dependence of the BZ jet model.

5.2 From jet power to radio luminosity

The mechanical BZ jet luminosities presented in the previous section are converted into radio luminosities according to the theoretical calculations of Heinz & Sunyaev (2003) (see also Falcke & Biermann 1995). Heinz & Sunyaev derived the non-linear dependence between the jet flux and the physical parameters M_{bh} and \dot{m} for radiatively efficient and inefficient flows. These authors found that the core flux at frequency, F_v , of a jet scales as $M_{\text{bh}}^{\xi_1} \dot{m}^{\xi_2}$, with $\xi_1 = \xi_2 = 17/12$ for ADAF systems and $\xi_1 = 17/12$, $\xi_2 = 0$ when the disc is radiation-pressure supported.⁴ A similar dependence on the accretion rate and BH mass seems to be implied by the fundamental plane of BH activity in the case of radiatively inefficient discs (Merloni, Heinz & di Matteo 2003; Falcke, Körding & Markoff 2004). Hence, when jets are launched in ADAFs we assume that

$$L_{\text{R,ADAF}} \propto (M_{\text{bh}} \dot{m})^{1.42}. \quad (42)$$

When the ADAF collapses to a TD we assume that for the range of accretion rates we are interested in, $0.01 \leq \dot{m} \lesssim 100$, the flow is radiation-pressure supported and thus

$$L_{\text{R,TD}} \propto M_{\text{bh}}^{1.42} \quad (43)$$

Since we know how L_{jet} depends on M_{bh} and \dot{m} we can work out the relation between L_{R} and L_{jet} for both accretion regimes. For example, in the BZ model, equations (38) and (39) in combination with equations (42) and (43) give

$$L_{\text{R,ADAF}} = A_1 (M_{\text{bh}} \dot{m})^{0.42} L_{\text{jet,ADAF}}, \quad (44)$$

$$L_{\text{R,TD}} = A_2 M_{\text{bh}}^{0.32} \dot{m}^{-1.2} L_{\text{jet,TD}}, \quad (45)$$

where A_1 and A_2 are normalization constants. In order to constrain the values of these constants, we assume that $A_1/A_2 \simeq 100$ which arises from the fact that the mechanical power of a jet in an ADAF is approximately 100 times higher than that of a jet in a TD (equations 38 and 39). A_1 is then constrained by fitting the predictions of each accretion model for the bright end of the radio luminosity function to the observations (see Section 5.4). When we consider the prolonged accretion model this gives $A_1 = 0.05$ and $A_2 = A_1/100 = 0.0005$. For the chaotic accretion model, $A_1 = 0.07$ and $A_2 = A_1/100 = 0.0007$. The uncertainty in the parameters A_1 and A_2 introduces an arbitrariness into the model, which, however, is unavoidable since the normalization of L_{R} is not calculable from first principles.

5.3 The $\log L_{\text{B}} - \log L_{\text{R}}$ plane for a $10^8 M_\odot$ BH

The theoretical predictions of the BZ model for the evolution of AGN jets on the $\log L_{\text{B}} - \log L_{\text{R}}$ plane are illustrated in Fig. 11. The optical luminosity, L_{B} , is obtained as discussed in Section 4. The jet strength is explicitly determined by the three fundamental parameters of our model: the spin, the accretion rate and the BH mass (and also by the values of the normalization constants A_1 and A_2). The predictions of the models are shown for a BH mass of $10^8 M_\odot$ spinning at $a = 0.01, 0.1, 0.5$ and 0.998 . The different accretion regimes are represented by different background shadings. As shown in the diagram, a BH that accretes through an ADAF develops jets whose radio luminosity increases with the accretion

⁴ The values for the scaling indices ξ_1 and ξ_2 are calculated for a flat spectrum.

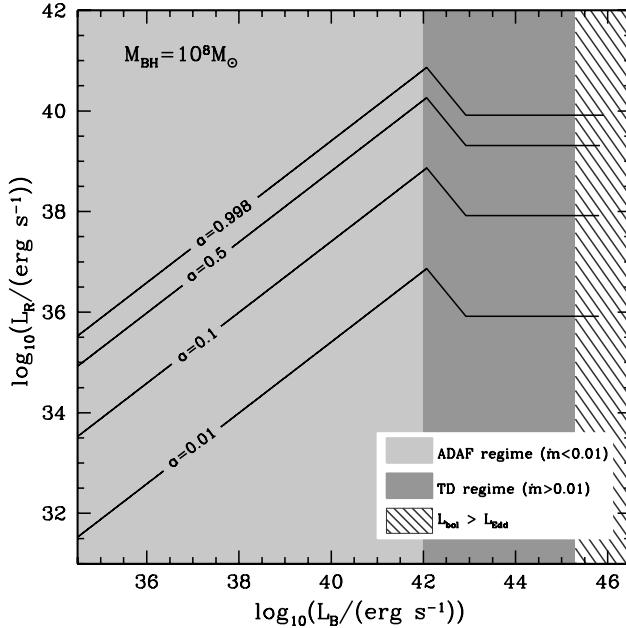


Figure 11. The predictions of the BZ model for the optical and radio luminosities of an AGN. Results are shown for an AGN hosting a SMBH with $M_{\text{bh}} = 10^8 M_{\odot}$ and rotating with spin $a = 0.01, 0.1, 0.5$ and 0.998 . The viscosity parameter α is fixed to 0.1 in both the ADAF and thin-disc regimes. We represent the different accretion regimes with different background shadings. The sharp drop in radio luminosity seen in all cases is due to the transition from the ADAF to the TD regimes when the accretion rate exceeds $0.01L_{\text{Edd}}$. The rightmost shaded area depicts the region where accretion becomes super-Eddington. In this region, the optical luminosity is set to $(1 + \ln \dot{m})L_{\text{Edd}}$. The discontinuity above $10^{42} \text{ erg s}^{-1}$ in accretion luminosity is due to the fact that TDs are radiatively efficient and thus, for the same accretion rate, are more luminous than an ADAF.

rate on the $\log L_B - \log L_R$ plane. When the accretion rate becomes higher than 0.01 , the accretion flow enters the thin-disc regime and the jet collapses by a factor roughly equal to the difference in H/R between the two very different accretion flows.

In the TD regime, the mechanical jet power increases with \dot{m} but this does not translate into an increase in observed radio flux. Instead, the radio flux remains constant, while the optical flux increases with mass accretion rate. Thus, the much lower radio emission is sustained throughout the TD regime and we do not expect objects powered by TDs to be radio loud. Our model extends this behaviour to super-Eddington mass accretion rates, although we note that this is almost certainly an underestimate of the jet and radio luminosities as H/R for such flows increases again to ~ 1 . Thus, there is a clear accretion mode switch in the models. The brightest radio sources are the highest spin BHs accreting at the maximum rate for an ADAF. Since this is $L/L_{\text{Edd}} = 0.01$, this maximum radio luminosity scales also with the mass of the BH.

5.4 The radio luminosity function

A first test of the jet model is the radio luminosity function of AGN. In Fig. 12, we show the predictions of the BZ jet model for the radio luminosity function in both the prolonged and chaotic accretion cases and compare with the estimates from the Two-degree-Field Galaxy Redshift Survey (2dFGRS; Sadler et al. 2002) and from

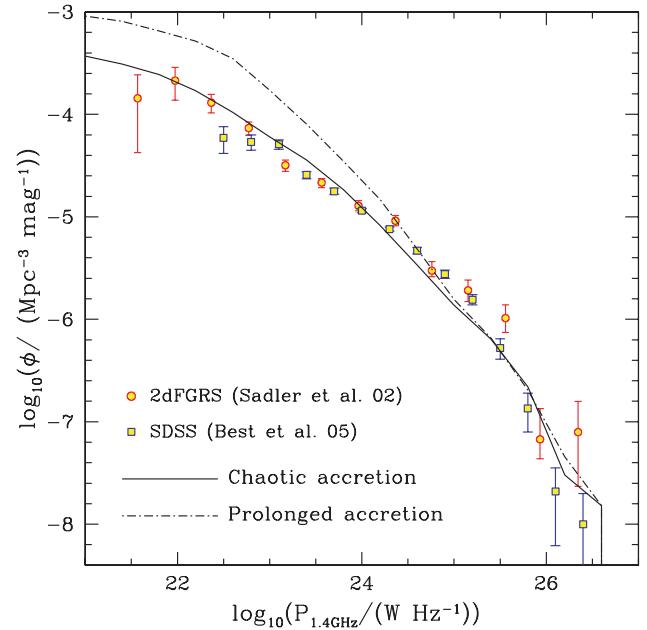


Figure 12. The radio luminosity function of AGN at 1.4 GHz . The lines show our model predictions for the prolonged (dash-dotted) and chaotic (solid) accretion models. The symbols show observational measurements for the local radio luminosity function from Sadler et al. (2002) for the 2dFGRS (circles) and Best et al. (2005) for the SDSS (boxes). The predictions of the BZ model, in combination with the chaotic accretion model, reproduce the observed radio luminosity function reasonably well.

the SDSS (Best et al. 2005).⁵ The model radio luminosity function is derived for systems powered by ADAFs and TDs, including super-Eddington objects, in a simulation volume of $V = 1.25 \times 10^8 \text{ Mpc}^3$.

The coupling between the different spin distributions and the BZ mechanism gives interestingly different predictions for the two accretion models. The chaotic model shows a fairly good agreement with the observations throughout the entire energy spectrum. The faint end is dominated mainly by galaxies that host slowly rotating BHs whereas the bright end is dominated by galaxies that host rapidly rotating massive BHs. The majority of these BHs accrete in the radio mode. A characteristic break in the slope at $\sim 10^{25} \text{ W Hz}^{-1}$ separates the two regimes.

In contrast, in the prolonged accretion case the same sample of galaxies shows a different distribution of luminosities. The model overproduces the number density of faint radio sources by a factor of ~ 3 compared to the chaotic model, due to the fact that these luminosity bins are populated by galaxies with rapidly rotating BHs whose radio luminosity is too large. This trend extends down to luminosities of $10^{25} \text{ W Hz}^{-1}$. Above that luminosity, the model is in reasonably good agreement with the observations. Since the mechanism that spins those BHs up, namely the BH mergers, is present in both models, the two predicted luminosity functions agree with each other very well at $10^{25}-10^{26} \text{ W Hz}^{-1}$. Note that the parameters A_1 and A_2 introduced in Section 5.2 affect only the normalization of the radio luminosity function. The shape is a prediction of the model and depends only on the distribution of BH masses, spins and accretion rates.

⁵ The results from 2dFGRS have been adjusted to the cosmology adopted by Best et al. (2005).

6 PREDICTIONS FOR THE $L_B - L_R$ AGN ACTIVITY

6.1 The distribution of galaxies on the optical–radio plane

We now explore the distribution of AGN on the log L_B –log L_R plane. We consider a sample of galaxies brighter than $L_*(M_V \simeq -20.5)$. This distribution is displayed in Fig. 13, where we plot the B band versus the radio luminosity of AGN for both the prolonged and chaotic accretion models (left- and right-hand plots, respectively). All galaxies are volume weighted and those accreting through a TD are weighted according to the fraction of their lifetime during which they are active. The colour coding corresponds to different types of accretion: blue shows AGN powered by an ADAF, red shows those powered by a TD and green indicates the super-Eddington objects.

Along with the model predictions, we plot the radio and optical luminosities of a sample of AGN-powered radio sources studied by Sikora et al. (2007). We include only local objects, i.e. those at $z < 0.14$, so this excludes the very brightest radio and optical sources. The evolution of the sources across cosmic time will be explored in a future paper.

The data form two distinct sequences on the log L_B –log L_R plane. The upper sequence represents radio-selected galaxies and the lower sequence optically selected objects. The objects in the upper sequence are on average ~ 3 orders of magnitudes more radio loud than those in the lower sequence. These radio bright objects are Fanaroff and Riley class I (FRI) galaxies, broad line region galaxies (BLRGs) and radio-loud quasars (RLQs). These are thought to be powered by very massive BHs (McLure & Dunlop 2002), hosted by elliptical galaxies. Objects in the lower, radio-quiet sequence

include Seyferts, low ionization emission regions (LINERs) and radio-quiet quasars (RQQs).

The observational data suffer from multiple selection effects, which may be responsible for the apparent dichotomy seen in the distribution (see discussion by Sikora et al. 2007). The fact that some FRI galaxies have much higher radio power than predicted by our model could be due to the fact that these data include *lobe* power, whereas we only model the *core* radio emission. However, the locus of the data points sets upper limits for the radio luminosity of radio-quiet and radio-loud objects, and provides a rough visual description of the general bulk properties of the different populations.

In both accretion models, AGN span a wide range of optical and radio luminosities. AGN powered by ADAFs generically produce more radio and less optical emission than those powered by a TD. Thus, the accretion mode switch at $\dot{m} = 0.01$ from ADAF to TD produces a marked transition in radio loudness. All our models show that the radio-loud sequence lies at the upper end of the radio luminosity envelope for SMBH accreting via a hot flow. Conversely, the radio-quiet locus lies predominantly on the TD points, although the lowest luminosity radio-quiet AGN (LINERS) lie on the lower end of the ADAF regime. Thus, the *major* switch between radio-loud and radio-quiet is the change in jet properties as the accretion flow collapses from a hot, geometrically thick configuration, to a cool, geometrically thin TD (Jester 2005).

None the less, there can also be an additional spin dependence which enhances the difference in radio properties, although this depends strongly on the details of the jet modelling and the spin distribution. The BZ jet models are strongly dependent on spin, so their radio luminosity amplifies differences in the spin distribution. In the prolonged accretion model, almost all SMBHs have maximal

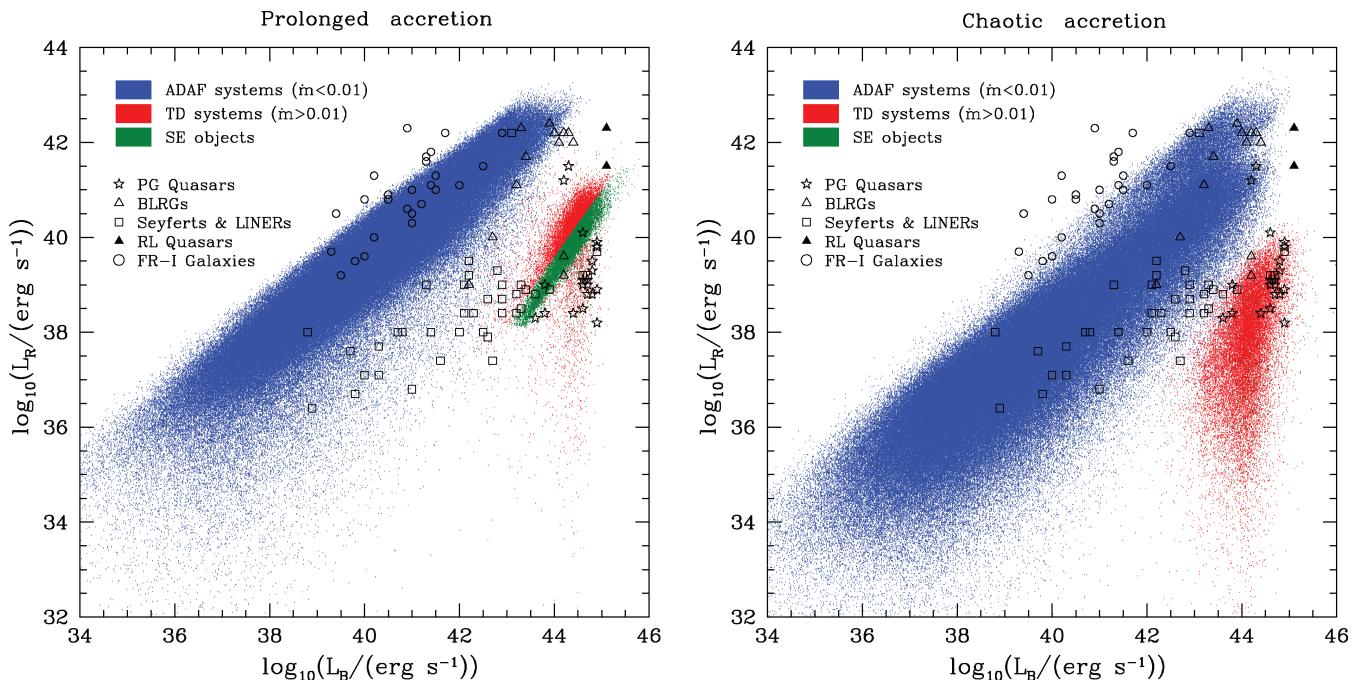


Figure 13. Radio luminosity versus B -band nuclear luminosity for the BZ model in (a) the prolonged and (b) the chaotic accretion models. Only galaxies with $z < 0.14$, $M_V < -20.5$ and $M_{\text{BH}} > 10^6 M_{\odot}$ are shown. The blue points represent galaxies whose SMBHs accrete via an ADAF. These include galaxies experiencing *both* the radio and quasar modes. The red points represent galaxies powered by TDs around SMBHs. The green points denote super-Eddington objects. The observational data are taken from Sikora et al. (2007): BLRGs are shown by open triangles; RLQs by filled triangles; Seyfert galaxies and LINERs by open squares; FRI radio galaxies by open circles and PG Quasars by open stars.

spin, so there is little dispersion in radio flux. Conversely, for the chaotic model, the correlation between mass and spin means that the low-mass and hence lower optical luminosity objects have dramatically lower radio power than the higher mass BH in both the ADAF and TD regime.

There are also subtle differences in optical luminosity between the prolonged and chaotic models, as the higher efficiency of high spin accretion means that the same mass accretion rate gives rise to a higher luminosity. Thus, there are more super-Eddington sources in the prolonged accretion model, in which low-mass objects have high spin, than in the chaotic model.

6.2 The distribution of galaxies on the $\mathcal{R} - \lambda$ plane

In addition to the AGN optical and radio output, the model allows us to study how the *radio loudness* \mathcal{R} of AGN depends on various physical parameters such as the BH mass and the Eddington ratio $\lambda = L_{\text{bol}}/L_{\text{Edd}}$. The radio loudness measures the radio to optical flux ratio, $\mathcal{R} \equiv L_{v_R}/L_{v_{\text{opt}}}$ (Richards et al. 2006). Following the definition of \mathcal{R} in Sikora et al. (2007), we consider the total radio flux at 5 GHz, $L_R = v_{5\text{GHz}} L_{v_{5\text{GHz}}}$, and therefore we express the radio loudness as $\mathcal{R} = (v_B/v_{5\text{GHz}}) \times L_R/L_B = 1.36 \times 10^5 L_R/L_B$, given that the B band is centred at a wavelength of 4400 Å.

We plot the theoretical predictions for the radio loudness of the AGN in our sample along with the observational data set of Sikora et al. (2007) in Fig. 14. Our predictions suggest a clear inverse correlation between \mathcal{R} and λ with a substantial scatter in both accretion models (upper panels in Fig. 14). In the ADAF regime, the distribution of objects shows a correlation of the form $\mathcal{R} \propto \lambda^{-0.4}$ [one can derive the correlation between \mathcal{R} and λ using equations (44), (38)

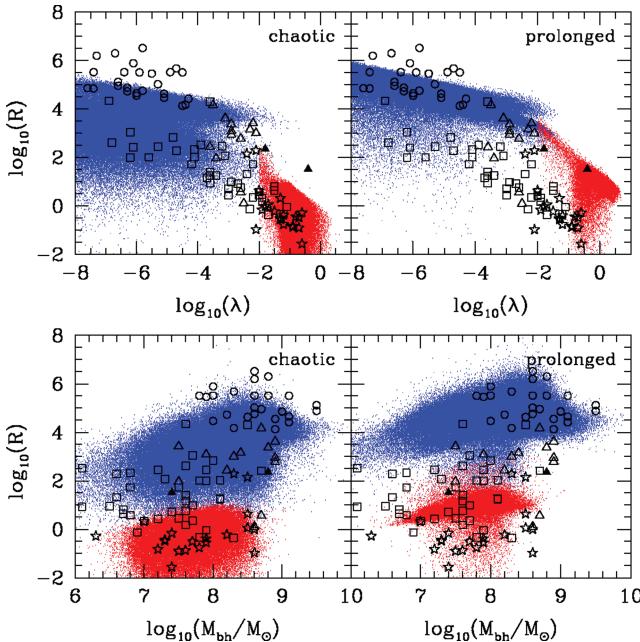


Figure 14. Scatter plots of radio loudness [$\mathcal{R} = (v_B/v_{5\text{GHz}}) \times L_R/L_B$] versus BH mass (top panels) and λ parameter (bottom panels) in the chaotic and prolonged accretion models for the sample in Fig. 13 (galaxies with $z < 0.14$, $M_V < -20.5$ and $M_{\text{bh}} > 10^6 M_\odot$). Different colours represent different accretion regimes: blue for ADAF systems and red for thin-disc systems. The observational data are taken from Sikora et al. (2007): BLRGs are shown by open triangles, RLQs by filled triangles, Seyfert galaxies and LINERs by open squares, FRI radio galaxies by open circles and PG Quasars by open stars.

and (35)] which is driven mainly by the strong dependence of the disc luminosity on the accretion rate, $L_{\text{bol}} \propto \dot{m}^2$. The correlation between \mathcal{R} and λ becomes steeper in the thin-disc regime, where the data approximately follow $\mathcal{R} \propto \lambda^{-1}$. The dependence of \mathcal{R} on λ in the thin-disc regime arises from $\mathcal{R} \propto L_{\text{bol}}^{-1} \propto \dot{m}^{-1} \propto \lambda^{-1}$.

In both models, the radio-quiet AGN ($\log \mathcal{R} < 1$; following Ho 2002) are preferentially found in the $\log \lambda \gtrsim -2$ regime, while the radio-loud sources ($\log \mathcal{R} > 1$) populate exclusively the $\log \lambda \lesssim -2$ regime. When compared to the data from Sikora et al. (2007), the predictions of the chaotic model give an adequate representation of the AGN on the $\mathcal{R} - \lambda$ plane. The model reproduces well the loci of the different AGN populations and the slope of the overall distribution. Objects such as quasars and Seyferts are radiating at $\sim(0.01 - 1) L_{\text{Edd}}$, while the radio galaxies are characterized by very low sub-Eddington luminosities ($< 0.01 L_{\text{Edd}}$).

By contrast, in the prolonged model there is an absence of objects with $\log \mathcal{R} < 3$ in the $-4 \lesssim \lambda \lesssim 3$ regime. This is mainly because most objects pile up in the upper envelope of the distribution due to their high radio luminosities (and BH spin values). We note again that FR I galaxies (open circles) have higher values of radio loudness than the models. This could be due to the fact that the measured radio luminosities for these objects include power from the lobes.

The objects in our sample populating the radio-loud part of the $\mathcal{R} - \lambda$ plane are associated mainly with giant ellipticals that host very massive BHs (see below). Accretion on to these BHs in the ADAF regime produces significant radio power because the spin is very high. In combination with the low optical luminosities of these objects, this gives rise to very high \mathcal{R} values. Therefore, these specific properties of the host of the radio-loud AGN may suggest a correlation between \mathcal{R} and M_{bh} . However, when we plot \mathcal{R} versus M_{bh} (lower panels in Fig. 14) we find that in both models there is no apparent correlation between these two quantities. The data display significant scatter along the M_{bh} axis which is due to the complex dependence of \mathcal{R} on M_{bh} through the optical and jet luminosity expressions in equations (44), (45) and (35). Both models reproduce the loci of the different AGN population in the Sikora et al. data set remarkably well. ADAF systems, mainly identified with radio-loud sources, span a wide range of BH masses ($10^6 - 10^{10} M_\odot$). In contrast, radio-quiet objects have BH masses in the range $\sim 10^7 - 10^{8.5} M_\odot$. Finally, in both models all the sources hosting BH with $M_{\text{bh}} \gtrsim 10^{8.5} M_\odot$ are exclusively radio-loud.

6.3 A physical view of the ‘chaotic accretion’ population

While both accretion models give a fair representation of the locus of the heterogeneous radio-optical data points from Sikora et al. (2007), the chaotic accretion model provides a better fit to the much more homogeneous data from the radio luminosity function described in Section 5.4. Hence, in this section, we focus on the chaotic accretion model and explore the results in greater detail.

In order to gain further insight into the predictions of the model, we show in Fig. 15 how the main parameters that determine the radio loudness of an AGN, a , M_{bh} and \dot{m} , are distributed across the $\log L_B - \log L_R$ plane for the chaotic accretion model.

Comparing these plots to the data in Fig. 13, the brightest radio-loud galaxies i.e. FR I and BLRG, are identified with very massive ($10^8 - 10^9 M_\odot$), rapidly rotating BHs (spins $\gtrsim 0.8$) accreting at the top of the ADAF branch at $\dot{m} \sim 0.01$. In our galaxy formation model, these objects are hosted exclusively by giant ellipticals and their central BHs grow through mergers and through the radio mode. The RLQs (only two in our volume) are the extreme end of this population, and can be matched by the highest mass objects

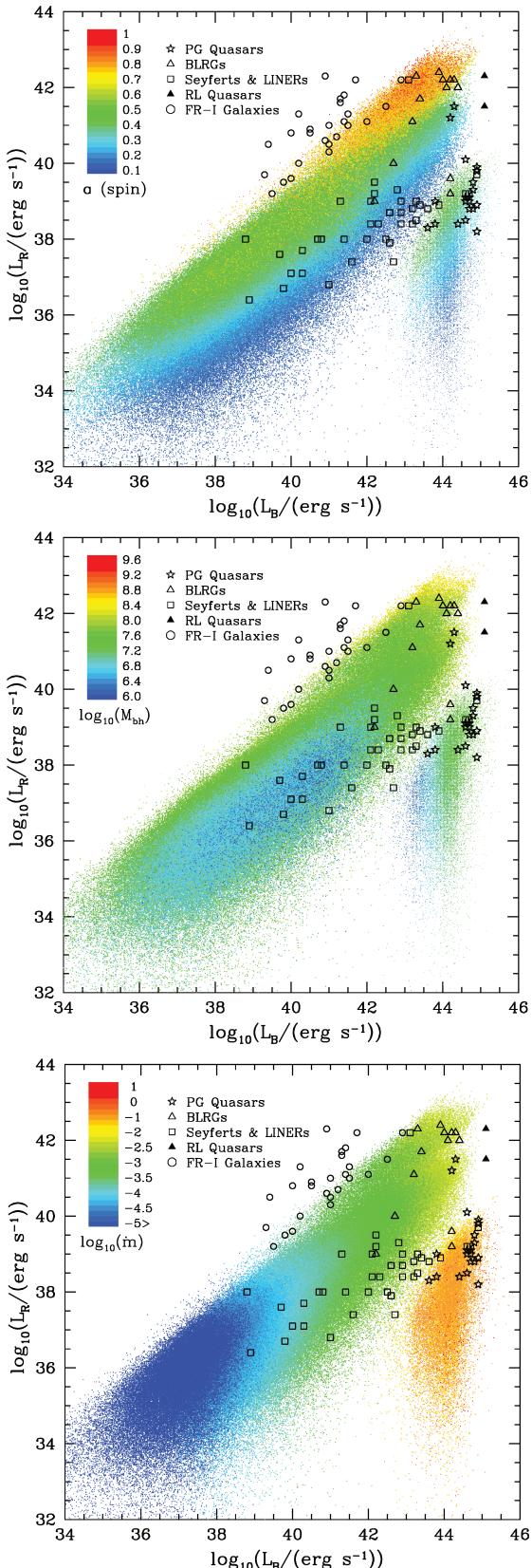


Figure 15. Scatter plot of radio luminosity versus B -band nuclear luminosity for the chaotic accretion model (as in Fig. 13), with model points coloured-coded according to BH spin (top), mass (middle) and accretion rate (bottom). The colours represent different ranges of values as indicated in each key.

in our sample. One issue with identifying these objects with the ADAF branch is that they typically have high excitation spectra, showing that there is a bright UV disc (see e.g. Marchesini, Celotti & Ferrarese 2004). However, the top of the ADAF branch is where the transition to a TD takes place. Observations of stellar-mass BH binary systems show that this transition is complex, probably taking on a composite structure with the TD replacing the hot flow at progressively smaller radii (see e.g. the review by Done et al. 2007). Thus, better modelling of the details of the transition may well be able to reproduce the required UV ionising spectrum.

The Palomar–Green (PG) Quasars are mostly concentrated at the top end of the narrow TD sequence, extending from $\sim 10^{43}$ to $\sim 10^{43}$ erg s^{-1} in optical luminosity and from $\sim 10^{38}$ to $\sim 10^{40}$ erg s^{-1} in radio luminosity. The higher mass accretion rates required to accrete via a TD are only generally possible in lower mass BHs, $10^7\text{--}10^8 M_\odot$ (see Fig. 4), so these have similar optical luminosities to the top of the ADAF branch as their higher mass accretion rate is partially cancelled by the lower mass BHs. These lower mass BHs are hosted in lower mass galaxies, predominantly spirals, and have lower spin, but the majority of the suppression in radio power comes from the ~ 2 orders of magnitude stronger jet emission in the ADAF regime than in the TD regime at $\dot{m} = 0.01$. However, a few of the optically brightest radio-quiet objects are powered by Eddington (or even super-Eddington) accretion on to $10^8 M_\odot$ BHs, and these could be hosted by ellipticals.

Seyfert and LINER galaxies constitute an interesting sample in our model because of their heterogeneity. Those exhibiting high nuclear luminosities ($L_B \gtrsim 10^{43} \text{ erg s}^{-1}$) (mainly Seyferts) appear to occupy both the ADAF and thin-disc regimes, whereas those with lower luminosities (LINERs) are generally powered by an ADAF. These are the most numerous type of AGN seen in the local Universe. The majority of these objects have BHs with masses of $10^7\text{--}10^8 M_\odot$. Thus, they have low-to-moderate spin and are relatively quiet at radio luminosities. These are the same population of objects as the PG Quasars, but at lower mass accretion rates.

7 DISCUSSION

The process of BH growth in our model is dominated by the accretion of large amounts of gas that allow a BH to double its mass several times during a Hubble time. The evolution of the spin of the BH is strongly influenced by the nature of the accretion episodes. It is often assumed that BHs accrete gas via a disc with constant angular momentum. In this case (the prolonged accretion model), BHs are systematically spun up during accretion, with most of them ending up with maximal spin. Even minor mergers or disc instabilities, typical of the growth of the bulge of spiral galaxies, can trigger a gas flow on to the BH of mass comparable to that of the BH itself and thus lead to $a = 1$. By contrast, if the accreting material fragments at its self-gravity radius and the associated star formation randomizes the angular momentum direction of each of the fragments (the chaotic accretion model), then accretion proceeds in multiple *randomly oriented* episodes. This results in a low spin BH. However, for the most massive BHs, the major growth channel is not accretion but BH–BH mergers, and these generate fairly rapid spins of 0.7–0.8. Thus, in the prolonged accretion model, the most massive BHs have slightly lower spin than the bulk of the BH population, whereas in the chaotic model they have larger spin.

Our results for the spin distributions are in good agreement with those of Berti & Volonteri (2008). However, our predictions for the prolonged-accretion model deviate significantly from the results of

Lagos et al. (2009). These authors obtain a strong spin bimodality when they consider accretion of gas via an accretion disc of constant angular momentum. The difference between the two approaches seems from the different amounts of cold gas available for accretion in the two models (see Kim et al. 2010 for a recent study of the cold-gas abundance in the Bower et al. 2006 model). In the original Bower et al. (2006) model and in the updated version that we have adopted in this study, the typical amount of cold gas accreted by a 10^6 – $10^8 M_{\odot}$ BH is enough to spin it up to the maximum value. This is not true, however, for the Lagos et al. model where cold gas accretion includes several episodes with $M_{\text{gas}} < M_{\text{bh}}$ that do not spin up the BH efficiently. The difference in the final spin distributions illustrates the impact of the underlying galaxy formation model on the inferred properties of the BHs at the centre of galaxies.

We have used the resulting distributions of BH mass, spin and mass accretion rate to calculate the properties of AGN, assuming that any accretion at a rate $\dot{m} < 0.01$ proceeds via an ADAF, while accretion at higher rates forms a standard TD. The optical luminosity is then fairly straightforward to calculate, but the radio luminosity depends on the jet model. This is poorly understood, but both observations and theoretical models agree that the collapse of the thick ADAF into a TD leads to a similar collapse of the jet emission by several orders of magnitude. Thus, there is already a clear accretion mode switch in radio-loudness predicted by these models (Maccarone, Gallo & Fender 2003; Jester 2005).

The main issue then is how jet power couples to spin. In the prolonged accretion model, there is a very small range in spin since all BHs are spinning rapidly. Thus, *independently of the detailed jet model*, the slope of the radio luminosity function is directly given by the relative numbers of low- and high-mass BHs accreting in the ADAF regime. This is steeper than observed since lower mass BH are much more numerous than higher mass ones. Instead, in the chaotic accretion model, the observed radio luminosity function can be reproduced *if the jet power depends strongly on spin*. The a^2 dependence on spin power in the BZ jet models is sufficient to make a large difference in radio luminosity between the numerous low-mass, low-spin BHs and the much rarer high-mass, high-spin BHs. This model reproduces the shape of the observed radio luminosity function very well.

A mass–spin correlation was also suggested by Sikora et al. (2007) to explain the range in the radio-to-accretion disc luminosity ratio seen in their (very heterogeneous) sample of AGN. However, the model they propose to establish this correlation is rather different from ours which is based on an actual calculation of galaxy formation from CDM initial conditions. Instead, Sikora et al. speculated that the high spin resulted from major mergers which triggered large gas flows with constant angular momentum direction on to the nucleus, spinning up the BH and producing an elliptical galaxy. Conversely, they argued that low spin resulted from minor mergers of randomly aligned satellite galaxies (a random walk with spin up, spin down). These leave the galactic gas disc intact, producing a spiral galaxy. However, as discussed above, since most of the mass comes from the disc of the host galaxy, even minor mergers result in enough gas flowing to the centre to spin the BHs up to maximal. Fragmentation and chaotic accretion are necessary in order to produce low-spin BHs. Major mergers are indeed the key to the high spin of the most massive BHs, but this is the result of BH–BH mergers, not of gas accretion.

One of the distinguishing features of our work is that the properties of the BHs and their associated AGN are calculated ab initio within the context of a full model of galaxy formation in a Λ CDM

universe. This model has been shown to agree with a large variety of observational data such as galaxy luminosity functions in various passbands and at different epochs, galaxy colours, the cosmic star formation history and structural scaling properties such as the Tully–Fisher relation and the $M_{\text{BH}}-\sigma$ relations (Bower et al. 2006). In this work, we have augmented the galaxy formation model with calculations of BH spin and of the optical and radio output during accretion on to the BH. The two different accretion models we have explored (prolonged and chaotic) span the range of likely spin distributions and are good templates to investigate how the optical and radio luminosities of accreting systems depend on BH spin. As shown in Sections 5 and 6 and discussed in the previous paragraphs, the chaotic model gives a reasonable overall match to the observations, suggesting that a model in which BH spins develop a bimodal distribution is the most plausible.

8 A UNIFICATION SCHEME FOR THE AGN ACTIVITY

Finally, we present an interpretation of the optical and radio signatures of an active nucleus at different evolutionary stages of the host galaxy within a chaotic accretion framework. This is illustrated schematically in Fig. 16, where we show the relative position of the different types of active galaxies on the fundamental parameter (a , \dot{m} and M_{bh}) plane (left-hand panel; the a and M_{bh} axes have been merged into one since the two quantities correlate as shown in the right-hand panel of Fig. 9) and the jet/accretion-luminosity plane (right-hand panel). We also speculate on the outcome of super-Eddington accretion. Such objects are rare in the local Universe, but are increasingly important at higher redshift (see Fig. 3).

We start by considering the sources with the lowest mass BHs. During a minor galaxy merger, cold gas from the disc is transferred to the bulge, triggering star formation and BH growth. This process represents a natural mechanism for the gradual growth of spheroids in spiral galaxies (Parry et al. 2009). Accretion of gas on to the BH usually occurs at sub-Eddington rates through a TD which turns the host spheroid into the bright nucleus of a Seyfert galaxy. The bright UV flux from this disc produces a strongly ionising spectrum, resulting in strong emission lines. However, gas flows close to the disc may also contribute to the emission (or absorption) spectrum of the nucleus. The orientation of the central engine relative to a distant observer may have an important impact on the observed spectral features of the source and could account for the Type 1 and 2 sub-classification of Seyferts.

If the host galaxy experiences a major merger or a galactic disc instability, then the entire galactic gas disc is assumed to lose most of its angular momentum, participate in a starburst and is added to the stellar spheroid mass. This supplies the central region of the galaxy with large amounts of cold gas that feed the SMBH with several solar masses of material per year at a near- or super-Eddington rate. Thus, the nucleus becomes exceptionally luminous, resulting in a quasar. The collapse of the cold gas reservoir also initiates an intense starburst which could be contemporaneous with the quasar phase. However, during the quasar phase, the galaxy is normally too dim to be seen against the vast amounts of radiation produced by the central engine. The radio output of the central engine when the accretion rate becomes super-Eddington is not clear since our models do not extend to that regime. Perhaps if the magnetic field strength close to the BH is enhanced due to a transition to a thick disc when the flow exceeds the Eddington limit, strong jets may be launched establishing the galaxy as a RLQ.

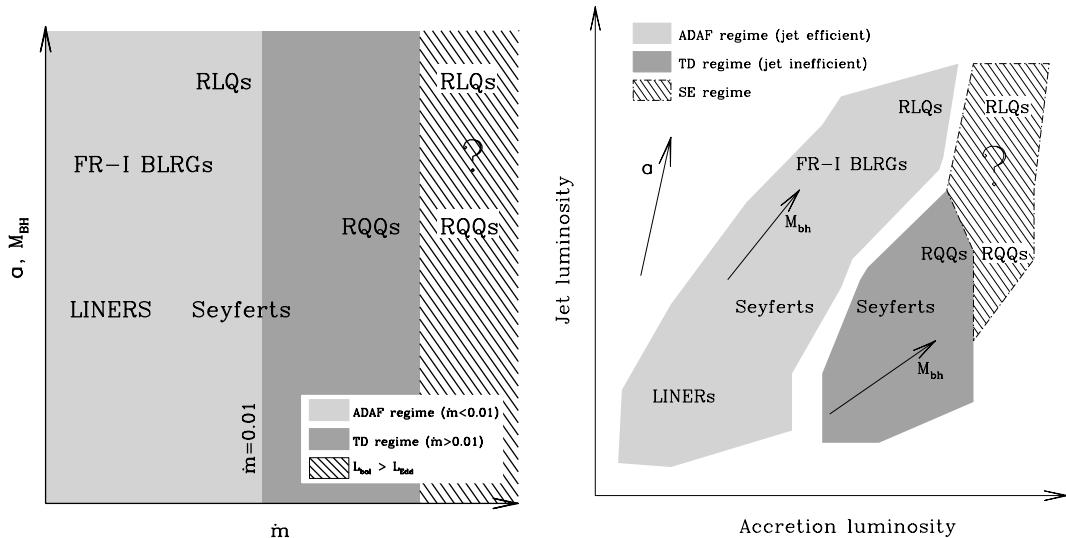


Figure 16. Left: the location of the different AGN types on the fundamental parameter (a, \dot{m}, M_{bh}) plane. The a and M_{bh} axes have been merged into one since the two parameters are correlated as shown in the right-hand plot of Fig. 9 in the chaotic accretion model. The shaded areas represent the different accretion regimes indicated by the keys. Right: the location of the different AGN types on the optical–radio plane. The shaded areas represent different accretion regimes. The arrows show how the physical parameters vary on the plane. Radio galaxies and LINERs lie in the ADAF regime. Seyferts are preferentially found in the thin-disc regime, even though a substantial population of Seyferts may be powered by an ADAF. RQQs accrete at relatively higher accretion rates than Seyferts and RLQs lie at the top of the ADAF branch. We speculate that some of the latter may be found also in the super-Eddington regime.

The duration of quasar activity is usually a few hundred million years. At the end of it (and the associated starburst), the system will have consumed or ejected its cold gas leaving a stellar bulge. In most haloes, the ejected gas falls back into the galaxy and a new galactic disc is formed on a dynamical time-scale. Further gas accretion, or galaxy mergers, may subsequently trigger another period of quasar activity and initiate a new growth era for the SMBH. Eventually, the halo may grow massive enough that gas is no longer able to cool rapidly and the halo enters the hydrostatic cooling regime. In this case, the end-product may be an elliptical galaxy in which further gas cooling is restricted by the radio-mode feedback. By construction, this accretion on to the BH takes place in the ADAF regime and the host elliptical appears as a radio galaxy. In the extreme case where the accretion occurs at the top of the ADAF branch and the central SMBH is rapidly rotating the galaxy is identified as a RLQ. The orientation of the central engine and jets may explain the different subcategories of RLQs. For example, if the jet axis lies close to the line-of-sight of a distant observer, the source may be visible as a blazar.

9 CONCLUSIONS

We have presented a model of AGN activity in which the evolution of the galaxy is calculated in its full cosmological context and the luminosity of AGN at radio and optical wavelengths is calculated using a model for gas accretion and the generation of jets. We first considered the evolution of SMBH spin. We have found that the different astrophysical processes that influence the growth of SMBHs have a significant effect on the global spin distribution. For example, if accretion of gas of constant angular momentum dominates the growth of SMBHs, the associated holes will be rapidly rotating. However, if the gas that is fed into the SMBH has random angular momentum, a bimodal spin distribution results. In this case, high spin values occur only for the most massive BHs ($\gtrsim 10^8 M_\odot$), and these are mainly due to gas-poor major mergers, where the BH growth is dominated by the BH–BH merger.

We have coupled this mass, spin and mass accretion rate evolution to a model for the accretion flow and jet. The accretion flow is assumed to form a geometrically thick, hot, radiatively inefficient flow (ADAF) for $\dot{m} < 0.01$, and to collapse to a TD at higher mass accretion rates. The jet power couples strongly to the accretion mode since it depends on the vertical (poloidal) magnetic field component close to the BH horizon. The collapse by two orders of magnitude in the scaleheight of the flow results in a similar drop in radio power. This already produces a dichotomy in radio properties which explains the distinction between radio-loud and radio-quiet objects.

However, we also find that for our model to match the slope of the observed radio luminosity function, it is necessary for low-mass BHs to generate relatively less jet power than high-mass BHs. This is readily achieved in models where the jet couples *strongly* to spin, such as in the classic BZ jet mechanism, *and* where lower mass BHs have lower spin than the most massive BHs, as in our chaotic accretion model.

Coupling the chaotic accretion model to the BZ jet mechanism results in an AGN population which reproduces the diversity of nuclear activity seen in the local Universe. In particular, the model accounts for the radio and optical luminosities of the FRI, BLRG, Seyfert and LINER galaxy populations. This is the first consistent demonstration that a great part of the phenomenology of AGN can be naturally explained by the coeval evolution of galaxies and BHs, coupled by AGN feedback, in a CDM universe. In future work, we will address the evolution of AGN across cosmic time, as a crucial test of galaxy formation models.

ACKNOWLEDGMENTS

We thank an anonymous referee for useful suggestions that have improved our paper. NF acknowledges receipt of a fellowship funded by the European Commission's Framework Programme 6, through the Marie Curie Early Stage Training project MEST-CT-2005-021074. AJB acknowledges the support of the Gordon & Betty

Moore Foundation. CSF acknowledges a Royal Society Wolfson Research Merit Award. This work was supported in part by an STFC Rolling Grant to the Institute for Computational Cosmology.

REFERENCES

- Abramowicz M. A., Chen X., Kato S., Lasota J.-P., Regev O., 1995, ApJ, 438, L37
- Baker J. G., Centrella J., Choi D.-I., Koppitz M., van Meter J., 2006a, Phys. Rev. D, 73, 104002
- Baker J. G., Centrella J., Choi D.-I., Koppitz M., van Meter J., 2006b, Phys. Rev. D, 96, 111102
- Baker J. G., McWilliams S. T., van Meter J. R., Centrella J., Choi D.-I., Kelly B. J., Koppitz M., 2007, Phys. Rev. D, 75, 124024
- Balbus S. A., Hawley J. F., 1998, Rev. Modern Phys., 70, 1
- Bardeen J. M., 1970, Nat, 226, 64
- Bardeen J. M., Petterson J. A., 1975, ApJ, 195, L65
- Bardeen J. M., Press W. H., Teukolsky S. A., 1972, ApJ, 178, 347
- Beckwith K., Hawley J. F., Krolik J. H., 2008, ApJ, 678, 1180
- Begelman M. C., Blandford R. D., Rees M. J., 1984, Rev. Modern Phys., 56, 255
- Benson A. J., Babul A., 2009, MNRAS, 397, 1302
- Berti E., Volonteri M., 2008, ApJ, 684, 822
- Berti E., Cardoso V., Gonzalez J. A., Sperhake U., Hannam M., Husa S., Brügmann B., 2007, Phys. Rev. D, 76, 064034
- Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005, MNRAS, 362, 9
- Binney J., Tremaine S., 1987, Galactic Dynamics. Princeton University Press, Princeton, NJ
- Blandford R. D., Payne D. G., 1982, MNRAS, 199, 883
- Blandford R. D., Znajek R. L., 1977, MNRAS, 179, 433
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
- Buonanno A., Cook G. B., Pretorius F., 2007, Phys. Rev. D, 75, 124018
- Campanelli M., Lousto C. O., Zlochower Y., Krishnan B., Merritt D., 2007a, Phys. Rev. D, 75, 064030
- Campanelli M., Lousto C., Zlochower Y., Merritt D., 2007b, ApJ, 659, L5
- Cattaneo A. et al., 2007, MNRAS, 377, 63
- Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
- Croom S. M., Smith R. J., Boyle B. J., Shanks T., Miller L., Outram P. J., Loaring N. S., 2004, MNRAS, 349, 1397
- Croom S. M. et al., 2009, MNRAS, 399, 1755
- Croton D. J. et al., 2006, MNRAS, 365, 11
- De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499
- De Villiers J.-P., Hawley J. F., Krolik J. H., Hirose S., 2005, ApJ, 620, 878
- Done C., Gierliński M., Kubota A., 2007, A&AR, 15, 1
- Efstathiou G., Lake G., Negroponte J., 1982, MNRAS, 199, 1069
- Elvis M. et al., 1994, ApJS, 95, 1
- Esin A. A., McClintock J. E., Narayan R., 1997, ApJ, 489, 865
- Fabian A. C., Iwasawa K., 1999, MNRAS, 303, L34
- Falcke H., Biermann P. L., 1995, A&A, 293, 665
- Falcke H., Körding E., Markoff S., 2004, A&A, 414, 895
- Fender R. P., Belloni T. M., Gallo E., 2004, MNRAS, 355, 1105
- Genzel R. et al., 2003, ApJ, 594, 812
- Häring N., Rix H.-W., 2004, ApJ, 604, L89
- Hasinger G., Miyaji T., Schmidt M., 2005, A&A, 441, 417
- Hawley J. F., Krolik J. H., 2006, ApJ, 641, 103
- Hawley J. F., Gammie C. F., Balbus S. A., 1995, ApJ, 440, 742
- Heckman T. M., Kauffmann G., Brinchmann J., Charlton S., Tremonti C., White S. D. M., 2004, ApJ, 613, 109
- Heinz S., Sunyaev R. A., 2003, MNRAS, 343, L59
- Heller C., Shlosman I., Englmaier P., 2001, ApJ, 553, 661
- Herrmann F., Hinder I., Shoemaker D. M., Laguna P., Matzner R. A., 2007, Phys. Rev. D, 76, 084032
- Hinder I., Vaishnav B., Herrmann F., Shoemaker D. M., Laguna P., 2008, Phys. Rev. D, 77, 081502
- Ho L. C., 2002, ApJ, 564, 120
- Hobbs A., Nayakshin S., Power C., King A., 2010, preprint (arXiv:1001.3883)
- Hopkins P. F., Richards G. T., Hernquist L., 2007, ApJ, 654, 731
- Hughes S. A., Blandford R. D., 2003, ApJ, 585, L101
- Ichimaru S., 1977, ApJ, 214, 840
- Jester S., 2005, ApJ, 625, 667
- Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576
- Kellermann K. I., Sramek R., Schmidt M., Shaffer D. B., Green R., 1989, AJ, 98, 1195
- Kim H.-S., Baugh C. M., Benson A. J., Cole S., Frenk C. S., Lacey C. G., Power C., Schneider M., 2010, preprint (arXiv:1003.0008)
- King A. R., Lubow S. H., Ogilvie G. I., Pringle J. E., 2005, MNRAS, 363, 49
- King A. R., Pringle J. E., Hofmann J. A., 2008, MNRAS, 385, 1621
- Kinney A. L., Schmitt H. R., Clarke C. J., Pringle J. E., Ulvestad J. S., Antonucci R. R. J., 2000, ApJ, 537, 152
- Lagos C. D. P., Cora S. A., Padilla N. D., 2008, MNRAS, 388, 587
- Lagos C. D. P., Padilla N. D., Cora S. A., 2009, MNRAS, 395, 625
- Lense J., Thirring H., 1918, Phys. Z., 19, 156
- Libeskind N. I., Cole S., Frenk C. S., Helly J. C., 2006, MNRAS, 368, 1381
- Lynden-Bell D., 1969, Nat, 223, 690
- Maccarone T. J., Gallo E., Fender R., 2003, MNRAS, 345, L19
- MacDonald D., Thorne K. S., 1982, MNRAS, 198, 345
- McKinney J. C., Gammie C. F., 2004, ApJ, 611, 977
- McLure R. J., Dunlop J. S., 2002, MNRAS, 331, 795
- Magorrian J. et al., 1998, AJ, 115, 2285
- Mahadevan R., 1997, ApJ, 477, 585
- Malbon R. K., Baugh C. M., Frenk C. S., Lacey C. G., 2007, MNRAS, 382, 1394
- Marchesini D., Celotti A., Ferrarese L., 2004, MNRAS, 351, 733
- Marconi A., Hunt L. K., 2003, ApJ, 589, L21
- Marronetti P., Tichy W., Brügmann B., González J., Hannam M., Husa S., Sperhake U., 2007, Classical Quantum Gravity, 24, 43
- Martínez-Sansigre A., Taylor A. M., 2009, ApJ, 692, 964
- Meier D. L., 2001, ApJ, 548, L9
- Meier D. L., 2002, New Astron. Rev., 46, 247
- Merloni A., Heinz S., di Matteo T., 2003, MNRAS, 345, 1057
- Merritt D., Milosavljević M., 2005, Living Rev. Relativ., 8, 8
- Milosavljević M., Merritt D., 2001, ApJ, 563, 34
- Miyaji T., Hasinger G., Schmidt M., 2001, A&A, 369, 49
- Nagar N. M., Wilson A. S., 1999, ApJ, 516, 97
- Narayan R., Yi I., 1994, ApJ, 428, L13
- Natarajan P., Armitage P. J., 1999, MNRAS, 309, 961
- Nemmen R. S., Bower R. G., Babul A., Storchi-Bergmann T., 2007, MNRAS, 377, 1652
- Novikov I. D., Thorne K. S., 1973, Black Holes. Les Astres Occlus, p. 343
- Parry O. H., Eke V. R., & Frenk C. S., 2009, MNRAS, 396, 1972
- Polletta M., Weedman D., Höning S., Lonsdale C. J., Smith H. E., Houck J., 2008, ApJ, 675, 960
- Pringle J. E., 1981, ARA&A, 19, 137
- Pringle J. E., 1992, MNRAS, 258, 811
- Rees M. J., 1982, in Riegler G. R., Blandford R. D., eds, AIP Conf. Proc. Vol. 83, The Galactic Center. Am. Inst. Phys., New York, p. 166
- Rezzolla L., Diener P., Dorband E. N., Pollney D., Reisswig C., Schnetter E., Seiler J., 2008a, ApJ, 674, L29
- Rezzolla L., Barausse E., Dorband E. N., Pollney D., Reisswig C., Seiler J., Husa S., 2008b, Phys. Rev. D, 78, 044002
- Richards G. T. et al., 2006, ApJS, 166, 470
- Sadler E. M. et al., 2002, MNRAS, 329, 227
- Scheuer P. A. G., Feiler R., 1996, MNRAS, 282, 291
- Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
- Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, ApJ, 690, 20
- Sikora M., Stawarz Ł., Lasota J.-P., 2007, ApJ, 658, 815
- Springel V. et al., 2005, Nat, 435, 629
- Tanabe K., Nagataki S., 2008, Phys. Rev. D, 78, 024004
- Tchekhovskoy A., Narayan R., McKinney J. C., 2010, ApJ, 711, 50
- Thorne K. S., 1974, ApJ, 191, 507

Volonteri M., Sikora M., Lasota J.-P., 2007, ApJ, 667, 704
Wang J.-M. et al., 2009, ApJ, 697, L141
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Wilkins D. C., 1972, Phys. Rev. D, 5, 814

Wilson A. S., Colbert E. J. M., 1995, ApJ, 438, 62
Xu C., Livio M., Baum S., 1999, AJ, 118, 1169

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.