Galaxy Evolution and Feedback across Different Environments
Proceedings IAU Symposium No. 359, 2020
T. Storchi-Bergmann, W. Forman, R. Overzier & R. Riffel, eds.
doi:10.1017/S1743921320001751

Models for galaxy and massive black hole formation and early evolution

Rainer Weinberger

Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, MS-51, Cambridge, MA 02138, USA
email: rainer.weinberger@cfa.harvard.edu

Abstract. Models for massive black holes are a key ingredient for modern cosmological simulations of galaxy formation. The necessity of efficient AGN feedback in these simulations makes it essential to model the formation, growth and evolution of massive black holes, and parameterize these complex processes in a simplified fashion. While the exact formation mechanism is secondary for most galaxy formation purposes, accretion modeling turns out to be crucial. It can be informed by the properties of the high redshift quasars, accreting close to their Eddington limit, by the quasar luminosity function at peak activity and by low-redshift scaling relations. The need for halo-wide feedback implies a feedback-induced reduction of the accretion rate towards low redshift, amplifying the cosmological trend towards lower accretion rates at low redshift.

Keywords. galaxies: formation, galaxies: nuclei, galaxies: active, methods: numerical, black hole physics, accretion

1. Introduction

Massive black holes (MBHs) are an essential part of cosmological structure formation. Modern simulations of galaxy or galaxy cluster formation rely on models for MBH formation and growth, since feedback effects from these objects have been shown to potentially explain some properties of massive galaxies and galaxy clusters (Somerville & Davé 2015). Two examples are the bimodal distribution of central galaxy colors (Trayford *et al.* 2016; Nelson *et al.* 2018), with more massive galaxies being redder and less star-forming, and the so-called cooling-flow problem, where gas cooling can be commonly observed in galaxy clusters, yet, does not lead to expected levels of star formation (e.g. Fabian 2012).

Cosmological simulations of galaxy formation model the formation of structure of the Universe as an initial value problem, numerically evolving the dark matter and gas distribution over most of cosmic time to redshift zero (for a detailed review, see Vogelsberger *et al.* 2020). While non-radiative, or generally non-dissipative simulations can be readily run and result in virialized halos, the introduction of dissipative terms in the form of radiative cooling of the gas will lead to a runaway collapse which will prohibit a numerical time-integration over cosmic timescales. The introduction of a simple closure at small scales, in which gas exceeding a certain density threshold is transformed to a collisionless 'star-particle' will alleviate this computational problem, however, yield vastly different results when compared to the observed galaxy population.

Significant progress has been made in recent years, showing that a more multi-facetted closure, including the effects of stellar feedback (e.g. Springel & Hernquist 2003; Dalla Vecchia & Schaye 2008) as well as a feedback component from active galactic nuclei (e.g. Sijacki *et al.* 2007; Dubois *et al.* 2012), can produce broad agreement between the simulated and observed galaxy population (e.g. Vogelsberger *et al.* 2013; Crain *et al.* 2015;

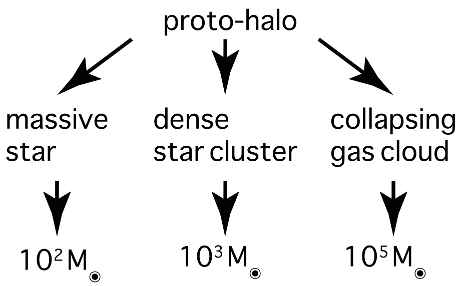

Figure 1. Broad classification of massive black hole seeding channels with expected seed mass.

Pillepich *et al.* 2018). Since these closure or sub-grid models are inspired by a simplified physical understanding, they do not cover the full complexity of the underlying process, therefore limiting the predictive power of the simulation. Yet, these models can be used as a guiding line and a test of plausibility of specific ideas which astrophysical processes are responsible for specific observational signatures.

I will review models for MBHs in cosmological simulations from seeding at high redshift until the onset of quenching around redshifts $z = 1-2$, discuss some of the difficulties related to these models and present open questions about the evolution of MBHs over cosmic time.

2. Different phases

The evolution of MBHs can be divided into 4 phases. First, MBH seed formation; second, early growth and the highest redshift quasars; third, the peak of the cosmic accretion rate density, and finally the epoch in which active galactic nucleus (AGN) feedback impacts the evolution of the entire host galaxy.

2.1. *Seeding*

One of the most uncertain aspects of MBHs is their formation. Since structure formation of non-dissipative components such as dark matter stops at virialization, further gravitational collapse is only possible via dissipative processes, i.e. radiative cooling of gas. It is therefore unsurprising that the formation of MBHs depends critically on the physics of radiative cooling, which in itself depends on chemical composition as well as external radiation fields. Theories for different channels of high redshift MBH formation have been around for some time (see Rees 1984 and Volonteri 2010 for reviews on this topic), yet the precise mechanism and possible observational evidence for it are subject to active research. Figure 1 shows a broad categorization into three different channels.

The first is the stellar-remnant channel, in which black holes of mass of order 10^2 $M_\odot$ are produced as remnants of massive, so-called population III stars (Carr *et al.* 1984). These short-lived stars are the first stars in the Universe, and form only in the absence of chemically enriched gas in so-called mini-halos.

The second channel is operating when cooling is slightly more efficient. Cool gas can form in the halo center, fragment and collapse into individual stars, thus forming a dense star cluster. In this star cluster, through collisional n-body dynamics, core-collapse can occur (Begelman & Rees 1978), which leads to the formation of a MBH in the center, with masses of order 10^3 $M_\odot$ (Devecchi & Volonteri 2009).

A third channel is possible if the formation of molecular hydrogen and consequently cooling to low temperatures is inhibited, e.g. by a sufficiently strong UV radiation flux. This leads to a larger Jeans length, i.e. prevents fragmentation, and a direct collapse of

https://doi.org/10.1017/S1743921320001751 Published online by Cambridge University Press

an entire massive gas cloud into a MBH seed (Bromm & Loeb 2003). The expected mass of a black hole forming via this channel is of the order of $10^5\,\mathrm{M}_\odot$.

While none of these channels is fully understood, there are only a comparably small number of properties of crucial importance for cosmological simulations of galaxy formation: when these seed MBHs form, which mass they have and how frequent they are.

Given the typical mass resolution of a cosmological volume simulation (targeted towards a $z = 0$ galaxy population) is around 10^5 to $10^6\,\mathrm{M}_\odot$, it is often omitted to distinguish between different seed scenarios. Instead, a commonly used way to seed MBHs in simulations is to simply assume that they are present in every halo exceeding a specific mass, typically around 10^{10}–$10^{11}\,\mathrm{M}_\odot$. While the exact numbers are somewhat arbitrary, this is a numerically very robust way to introduce MBHs in the simulation. However, implicitly assumes that low-mass MBHs grow in the same way as halos, which leads to a relatively flat distribution of seed times. A metallicity and gas density based seeding, which is an alternative and used in some simulations, leads to a peak of seeding at high redshift, with practically no seeding events at lower redshift, reflecting more the theoretical expectation that MBH seeds require a low metallicity environment to form (Tremmel *et al.* 2017). While these vastly different ways to introduce MBHs in the simulation likely lead to very different predictions about the early and the low-mass MBH population, it is important to keep in mind that the properties of the high-mass population is strongly influenced by gas accretion and hierarchical merging of halos (Weinberger *et al.* 2018), which leads to similar properties at low redshift independent of the details of the seeding.

2.2. *Early growth*

Once formed, MBHs grow in two different ways: via mergers with other black holes, and via gas accretion. While mergers will contribute, the initial growth is dominated by rapid accretion of gas. Evidence for this is provided by the existence of high redshift quasars with associated MBH masses of order $10^9\,\mathrm{M}_\odot$ at redshifts $\gtrsim 7$ (Bañados *et al.* 2018). These high mass MBHs at these redshifts place strict constraints on the combination of seed redshift, seed mass and maximum accretion rate at which a MBH can accrete. Assuming this maximum accretion rate is the Eddington limit, it becomes very hard to explain these black holes from population III remnants. Viable solutions are high-mass seeds from direct collapse or accretion rates that exceed the Eddington limit (see Smith *et al.* 2017 for a more detailed discussion).

Simulations of these high redshift quasars are very challenging, since the low number density of these objects requires to simulate a significant fraction of the visible universe to obtain a meaningful sample. An illustration of the scales involved is shown in Figure 2. The mean inter-object separation of high redshift quasars is of the order 10^9 pc, beyond the reach of most cosmological simulations targeting galaxy formation. Simply increasing the simulated volume is not possible, since there are resolution requirements to consider for modeling MBH seeds, even the direct collapse ones, as well as for black hole accretion. Fully satisfying these two opposing requirements is not possible at present day, yet some studies exist trying to address this problem in cosmological volume simulations that stop at high redshift (Di Matteo *et al.* 2017) as well as dedicated zoom simulations focusing on single halos (Smidt *et al.* 2018). In the latter case it has recently become possible to include the effects of radiation self-consistently in the simulation, which, by definition is crucial for objects accreting at the Eddington limit.

Cosmological volume simulations to $z = 0$ focusing on galaxy formation to-date do neither include radiation-hydrodynamics nor have the volume to produce these rare, high redshift quasars, which implies on the one hand that dedicated simulations are required to study them, on the other hand that their presence and abundance cannot be used as a

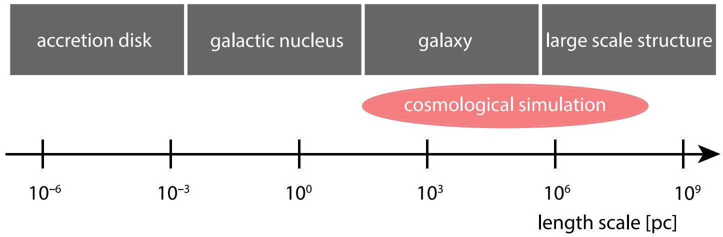

Figure 2. Spatial scales relevant for massive black holes. Cosmological simulations typically cover spatial scales from 300 Mpc to 30 pc. Therefore the covered volume is too small to contain rare objects such as the most luminous quasars, while the resolution limits require to marginalize over 6 orders of magnitude in spatial scales and related small-scale processes around massive black holes.

direct constraint for simulations. Yet, the general notion that MBH at high redshift seem to be able to accrete at or close to their Eddington limit is reassuring that the commonly employed assumption in simulations that accretion is limited to the Eddington rate is not unreasonable (note that due to the lack of radiation in these simulations, the simulated MBHs could have super-Eddington accretion rates at high redshift if not limited by the accretion model).

2.3. *Peak of activity*

Towards lower redshift, at the peak of the MBH accretion rate and star formation rate density at $z = 2{-}3$, the quasar luminosity function can be determined since a significant fraction of active galactic nuclei are observable. This quasar luminosity function is an important constraint on gas accretion onto MBHs. From a modeling perspective, the estimate of the accretion rate for the bulk of the of MBH population is very uncertain. While for very high redshifts and for the most luminous objects the assumption of accretion at the Eddington limit, i.e. a radiation pressure limited accretion, is a reasonable one, this ceases to be the case for the less extreme cases (Weinberger *et al.* 2018). For these less extreme cases it is hard to determine the limiting factor for accretion (possible factors are angular momentum, cooling, interactions with small-scale outflows, ...), let alone to estimate the accretion rate accurately from properties at galactic scales. Many models used in simulations are based on the Bondi accretion, frequently with some modifications. More recently, simulations using other prescriptions such as relations based on a torque, i.e. angular momentum limited accretion, have been performed and have shown to yield orders of magnitude different results for the same large scale conditions (Anglés-Alcázar *et al.* 2013). Considering the large range of unresolved scales (see Figure 2), and the fact that these modes assume different limiting factors for accretion, this discrepancy is not entirely surprising. But considering this uncertainty, it is rather surprising that these models are at all able to produce reasonable agreement with both low redshift MBH scaling relations, as well as high redshift quasar luminosity function constraints.

The main reason for this, in most simulations, is the self-termination of rapid accretion due to AGN feedback towards late times (however, see Anglés-Alcázar *et al.* 2013). Feedback at late times, unlike accretion, acts on resolved, galactic scales, which makes modeling easier.

2.4. *Downsizing and quenching*

Towards low redshift, both the star formation rate density as well as the MBH accretion rate density decrease towards redshift zero. Cosmological simulations reveal that this

https://doi.org/10.1017/S1743921320001751 Published online by Cambridge University Press

trend with redshift is not solely caused by feedback, but also visible in the global gas accretion rate density onto halos (van de Voort *et al.* 2011). The drastic decrease in number density of very luminous AGN can consequently, at least in part, be caused by a decrease in luminosity of individual AGNs due to decreased fueling, in combination with a steep negative slope of the quasar luminosity function at high luminosities. Therefore, the existence of these observed global trends cannot be interpreted as evidence for AGN feedback (however the trends might be enhanced due to AGN feedback).

The need for AGN feedback in cosmological simulations is more evident when trying to reproduce a population of massive, central galaxies with sustained low star formation rates, and consequently red intrinsic colors, as well as the X-ray properties of galaxy clusters. Matching observations in this respect has so far only been possible by invoking efficient AGN feedback in galaxies more massive than the Milky Way (e.g. Weinberger *et al.* 2017), a feature all cosmological simulations aiming to reproduce these high mass objects have in common (e.g. Khandai *et al.* 2015; Beckmann *et al.* 2017; Davé *et al.* 2019).

One of the key remaining questions in which simulations differ is how AGN feedback comes to be efficient in massive galaxies, while not being that relevant for the less massive galaxy population that remains star forming to the present day. Different simulations overcome this problem in different ways, some pointing towards the properties of stellar feedback (e.g. Bower *et al.* 2017), others achieve a similar effect by a change in mode of AGN feedback (Weinberger *et al.* 2018), possibly induced by small-scale accretion disk physics or a change in black hole spin (Bustamante & Springel 2019). Future observations, for example of the hot, soft X-ray emitting halo gas might be able to rule out certain models (Oppenheimer *et al.* 2020; Truong *et al.* 2020).

Another major aspect that requires more detailed study is the coupling of the energy released by the AGN with the host galaxy. In cosmological simulations, this is implicitly assumed when constructing a model on kpc scales. Studying this in more detail on smaller scales (e.g. Cielo *et al.* 2018), will be necessary to make a convincing case that whatever, for now, is assumed and required in cosmological simulations, is actually realistic.

3. Summary

Cosmological simulations require AGN feedback to reproduce the properties of massive galaxies. This need for AGN feedback makes it necessary to parameterize the rich and complex physics of MBH formation and evolution in a simplified fashion, but at the same time also allows to investigate the evolution of MBHs over cosmic time in a realistic environment. Some important takeaways are:

- There might be different seeding channels, however they are not modeled in most cosmological simulations, and the high-mass MBHs have likely lost all information about seeding. However, information about seeding can be obtained from low-mass MBHs and gravitational wave events.
- High redshift quasars indicate that early growth in the most extreme environments is close to Eddington-limited. This is very informative for accretion rate estimates at high redshift.
- Accretion rates of less extreme MBHs are significantly more difficult to estimate due to poorly understood physics at unresolved scales.
- Towards low redshift, AGN feedback is required to produce massive, quiescent central galaxies. This trend likely amplifies the general reduction of star formation rate density and downsizing, however is not the sole cause for it.
- The physical cause of the transition from a growth dominated to a feedback dominated regime is still debated, with upcoming observations having the potential to rule out some scenarios.

While cosmological simulations have been remarkably successful over the past years, presenting plausible scenarios of MBH evolution (and rule out a number of alternative ones), future studies that connect modeling of individual processes from first principle with the cosmological evolution are needed to gain further understanding about MBH formation and evolution.

References

Anglés-Alcázar, D., Özel, F., & Davé, R. 2013, *ApJ*, 770, 5
Bañados, E., Venemans, B. P., Mazzucchelli, C., *et al.* 2018, *Nature*, 553, 473
Beckmann, R. S., Devriendt, J., Slyz, A., *et al.* 2017, *MNRAS*, 472, 949
Begelman, M. C. & Rees, M. J. 1978, *MNRAS*, 185, 847
Bower, R. G., Benson, A. J., Malbon, R., *et al.* 2006, *MNRAS*, 370, 645
Bower, R. G., Schaye, J., Frenk, C. S., *et al.* 2017, *MNRAS*, 465, 32
Bromm, V. & Loeb, A. 2003, *ApJ*, 596, 34
Bustamante, S. & Springel, V. 2019, *MNRAS*, 490, 4133
Carr, B. J., Bond, J. R., & Arnett, W. D. 1984, *ApJ*, 277, 445
Cielo, S., Bieri, R., Volonteri, M., *et al.* 2018, *MNRAS*, 477, 1336
Crain, R. A., Schaye, J., Bower, R. G., *et al.* 2015, *MNRAS*, 450, 1937
Dalla Vecchia, C. & Schaye, J. 2008, *MNRAS*, 387, 1431
Davé, R., Anglés-Alcázar, D., Narayanan, D., *et al.* 2019, *MNRAS*, 486, 2827
Devecchi, B. & Volonteri, M. 2009, *ApJ*, 694, 302
Di Matteo, T., Croft, R. A. C., Feng, Y., *et al.* 2017, *MNRAS*, 467, 4243
Dubois, Y., Devriendt, J., Slyz, A., *et al.* 2012, *MNRAS*, 420, 2662
Fabian, A. C. 2012, *ARAA*, 50, 455
Khandai, N., Di Matteo, T., Croft, R., *et al.* 2015, *MNRAS*, 450, 1349
Nelson, D., Pillepich, A., Springel, V., *et al.* 2018, *MNRAS*, 475, 624
Oppenheimer, B. D., Bogdán, Á., Crain, R. A., *et al.* 2020, *ApJL*, 893, L24
Pillepich, A., Springel, V., Nelson, D., *et al.* 2018, *MNRAS*, 473, 4077
Rees, M. J. 1984, *ARAA*, 22, 471
Sijacki, D., Springel, V., Di Matteo, T., *et al.* 2007, *MNRAS*, 380, 877
Smidt, J., Whalen, D. J., Johnson, J. L., *et al.* 2018, *ApJ*, 865, 126
Smith, A., Bromm, V., & Loeb, A. 2017, *Astronomy and Geophysics*, 58, 3.22
Somerville, R. S., & Davé, R. 2015, *ARAA*, 53, 51
Springel, V. & Hernquist, L. 2003, *MNRAS*, 339, 289
Trayford, J. W., Theuns, T., Bower, R. G., *et al.* 2016, *MNRAS*, 460, 3925
Tremmel, M., Karcher, M., Governato, F., *et al.* 2017, *MNRAS*, 470, 1121
Truong, N., Pillepich, A., Werner, N., *et al.* 2020, *MNRAS*, 494, 549
van de Voort, F., Schaye, J., Booth, C. M., *et al.* 2011, *MNRAS*, 415, 2782
Vogelsberger, M., Genel, S., Sijacki, D., *et al.* 2013, *MNRAS*, 436, 3031
Vogelsberger, M., Marinacci, F., Torrey, P., *et al.* 2020, *Nature Reviews Physics*, 2, 42
Volonteri, M. 2010, *A&ARv*, 18, 279
Weinberger, R., Springel, V., Hernquist, L., *et al.* 2017, *MNRAS*, 465, 3291
Weinberger, R., Springel, V., Pakmor, R., *et al.* 2018, *MNRAS*, 479, 4056