University of Groningen

Black Holes during the Epoch of Reionization

Piana, Pietro Olmo

DOI:
[10.33612/diss.173865929](https://doi.org/10.33612/diss.173865929)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Piana, P. O. (2021). *Black Holes during the Epoch of Reionization*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.173865929>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

**rijksuniversiteit
 groningen**

Black Holes during the Epoch of Reionization

Proefschrift

ter verkrijging van de graad van doctor aan de
 Rijksuniversiteit Groningen
 op gezag van de
 rector magnificus prof. dr. C. Wijmenga
 en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

vrijdag 25 juni 2021 om 12:45 uur

door

Pietro Olmo Piana

geboren op 1 september 1991
 te Torino, Italië

Promotor

Prof. dr. Pratika Dayal

Copromotor

Prof. dr. Karina Caputi

Beoordelingscommissie

Prof. dr. Raffaella Morganti

Prof. dr. Raffaella Schneider

Prof. dr. Sune Toft

ISBN: 978-94-6421-397-3

To my family, that always makes things look so much easier.

Alla mia famiglia, che fa sembrare sempre tutto più facile.

Cover designed by the author
Printed by Ipskamp Printing on recycled paper

Contents

	SUMMARY FOR NON-EXPERTS	
	SOMMARIO	
	SAMENVATTING	
1	INTRODUCTION.....	1
1.1	The cosmological context	1
1.1.1	The hot big bang model	2
1.1.2	The formation of the first structures	4
1.2	The first stars, galaxies, and the epoch of reionization	11
1.2.1	The properties of high-redshift galaxies	13
1.2.2	Epoch of reionization	16
1.3	Black holes	20
1.3.1	Black hole seeds	21
1.3.2	Accretion and feedback	23
1.3.3	Observational methods for high-redshift black holes	25
1.3.4	Theoretical methods for black holes in a cosmological context	28
1.4	Thesis outline	30
2	THE HIERARCHICAL ASSEMBLY OF GALAXIES AND BLACK HOLES IN THE FIRST BILLION YEARS: PREDICTIONS FOR THE ERA OF GRAVITATIONAL WAVE ASTRONOMY.....	33
2.1	Introduction	34
2.2	The theoretical model	36
2.2.1	Seeding halos with black holes	36
2.2.2	Star formation and supernova feedback	39
2.2.3	Black hole growth and feedback	43

2.2.4 Smooth-accretion from the intergalactic medium	44
2.2.5 Merging galaxies and black holes	44
2.2.6 The impact of reionization	46
2.3 Comparing theoretical properties to observations	47
2.3.1 The observed UV LF for star formation and black holes	47
2.3.2 The LBG stellar mass density (SMD)	50
2.3.3 The black hole mass function and occupation fraction	52
2.3.4 The black hole-stellar mass relation	54
2.3.5 <i>LISA</i> and GWs from the high- z Universe	56
2.3.6 Conclusions and discussion	65
3 THE MASS ASSEMBLY OF HIGH-REDSHIFT BLACK HOLES . . .	75
3.1 Introduction	75
3.2 Theoretical model	78
3.3 The mass assembly of early black holes and their host halos	80
3.3.1 Black hole accretion rates	82
3.3.2 Joint assembly of black holes and their host halos	86
3.4 Key properties of the black hole population in early galaxies	90
3.4.1 The black hole occupation fraction	90
3.4.2 Time-weighted duty cycle	93
3.4.3 Black hole predictions at $z \gtrsim 4$	94
3.5 Conclusions and discussion	97
4 BLACK HOLE FEEDBACK AND ITS IMPACT ON THE UV LUMINOSITY AND ON THE STELLAR MASS ASSEMBLY OF HIGH-REDSHIFT GALAXIES	101
4.1 Introduction	102
4.2 The Theoretical model	105
4.3 Contribution of AGN to the UV luminosity of early galaxies	110
4.3.1 Contribution of star formation and black holes to the UV luminosity	110
4.3.2 Fractional lifetime for AGN domination of UV luminosity	114
4.3.3 Build-up of the AGN and SF UV LF	115

4.4	Impact of black holes on the assembly of early galaxies	117
4.4.1	Black hole impact on the gas mass assembly of early galaxies	119
4.4.2	Mass loading factors for star formation and black hole feedback	120
4.4.3	Impact of black hole feedback on the M_* – SFR relation	124
4.5	Conclusions and discussion	125
5	CONCLUSIONS AND FUTURE PROSPECTS	129
	BIBLIOGRAPHY	133
	AFTERWORD AND ACKNOWLEDGEMENTS	157

SUMMARY FOR NON-EXPERTS

Black holes are extremely dense and compact objects. It is enough to think that a black hole with the same mass as the Earth would be roughly the same size as a golf ball. Their strong attractive gravitational field bends and curves the geometry of space-time in their vicinity, so that everything that gets close enough and crosses the so-called *event horizon* eventually falls into them, from massive particles to photons. From the physical point of view, black holes are fully defined by their mass and their rotational velocity, assuming they have a null electric charge. Thus, they are pretty easy to describe if studied individually and in an isolated environment. However, the details of their interaction with the surrounding environment are not well understood yet, and are currently studied through both computer simulations and observations.

Researchers agree on the fact that we can find a massive black hole in the centre of almost every galaxy, and if the central black hole is growing then the galaxy is said to host an Active Galactic Nucleus (AGN). During the early phases of galaxy evolution, the collapse of a very massive star or of a clump of pristine gas (formed exclusively by Hydrogen and Helium) can leave behind a black hole seed, which then grows by feeding on the gas and stars that remain trapped in its gravitational field. It is estimated that approximately 90% of the material falling onto the black hole is actually swallowed, while the remaining 10% is turned into energy which is then emitted into the surrounding. This energy covers the whole electromagnetic spectrum, and can appear very luminous to our telescopes. Indeed, the emissions from supermassive black holes (they can reach the total mass of a galaxy like the Milky Way) tell us that these giants were already in place when the Universe was just a few hundred million years old, approximately one fifteenth of its current age. A fraction of the energy emitted by the growing black hole interacts with the gas and dust present in the galaxy (known as interstellar medium, ISM). This interaction is called black hole feedback, and results in an increment of the temperature of the gas particles, which start moving faster and can eventually escape from the galaxy. Since the gas is the raw material out of which stars are

formed, we expect black hole feedback to impact the rate at which new stars are formed within the galaxy. This suggests that understanding the growth of central black holes is important also in the context of galaxy evolution.

From the picture depicted above, we can infer what are the main questions regarding the growth and evolution of massive black holes that still remain unanswered. How can black holes grow so big in a relatively short timescale? How does this growth affect the host galaxies? Given the feedback process, is there a correlation between the physical properties of the black hole with those of the host galaxy? How does black hole growth depend on the size and mass of the host galaxy? These are the main points addressed in this work.

This Thesis

In particular, we study the co-evolution of central black holes and their hosts in the first billion year of the Universe, when galaxies are growing fast but are still very young. In Astronomy we have the possibility of observing objects that are so distant that the light they emit needs to travel for a long time before reaching us, meaning that we see these objects as they were in the past. In fact, the radio emission coming from the farthest observed AGN (or quasar) travelled for more than 13 billion year before reaching Earth. Thanks to this, we now have at our disposal some information about the statistical properties of the very early black hole and galaxy populations. We know about their UV luminosity, but also about their mass (though in this case the measurements still suffer from many uncertainties), their spatial distribution, and their number as a function of time. We have then built *Delphi*, a model to simulate the formation, growth and evolution of galaxies and their central black holes, able to reproduce these observational constraints and at the same time to give us insights of yet unobserved properties.

In Chapter 1 we introduce the questions addressed in the subsequent chapters, explaining what is already known and how our work fits within the context of the current state of the field. In Chapter 2 we effectively build the model, testing its results and making sure that it reproduces the main observations of black hole and galaxy properties at our disposal. We then compute the rate of mergers between black holes. In fact, when galaxies collide, their respective black hole migrate towards the centre of the new resulting galaxy, eventually forming a black hole binary system before merging together. The ever-tightening binary system emits gravitational

waves with very specific waveforms, and the shape of these waves depends on the mass of the black holes. Exploiting this characteristic, the future space mission *LISA*, an interferometer built to detect gravitational waves, will be able to observe black hole mergers at great distances. With our model, we predict the number of mergers between black holes that *LISA* should observe once it becomes operative. This will be most useful to test the validity and some of the assumptions of our model.

Using the same model, in Chapter 3 we try to shed light on how supermassive black holes grow, up to $10^9 - 10^{10}$ solar masses, within the first billion year of the Universe. We closely follow the evolution of both galaxies and their central black holes. These can gain mass either through mergers with other black holes or through accretion of gas and dust from the ISM. While the former mode seems to dominate when black holes are still very young and hosted by small galaxies, the latter becomes gradually more important as time passes. We also find that while small black holes grow faster at later times, bigger black holes show an opposite trend, and seem to grow faster at earlier times.

Finally, in Chapter 4, we study the effect of black hole growth and feedback on the host galaxy. Firstly, this growth causes an increase of the ultraviolet luminosity of the galaxy. With our approach, in particular, we are able to distinguish between the contributions of the growing black hole and that of the stars to the total UV luminosity of the galaxies. This means that we are able to tell which galaxies are dominated, in terms of luminosity, by black hole UV emissions as opposed to those dominated by stellar UV emissions. Indeed, studying the evolution of the UV luminosity coming from the stars of a galaxy gives a lot of information about their history. In addition, we find that the energy released by the growing black hole is actually able to slow down the formation of stars within the host galaxy, by ejecting huge amount of gas from it. Nevertheless, this is not true for galaxies of all masses. Black holes in smaller galaxies are still small and do not produce enough energy to affect the surrounding gas. On the other hand, the strong gravitational pull of the biggest galaxies is able to retain the gas despite the black hole activity. This means that black hole feedback has the biggest impact in galaxies of intermediate mass.

The work presented in this thesis is important especially when looking at the future of Astronomy. Future space telescopes - like the James Webb Space Telescope, *Athena*, *Lynx* - will be able to directly observe the black hole population we have described here, and tell us what of our model works and what still needs to be refined, possibly bringing a better understanding of the exact role played by these supermassive black holes in our Universe.

SOMMARIO

I buchi neri sono corpi estremamente densi e compatti. Basti pensare che un buco nero con la stessa massa della Terra avrebbe più o meno le dimensioni di una pallina da golf. Il loro forte campo gravitazionale curva e deforma la geometria dello spazio-tempo nelle loro immediate vicinanze, di modo che tutto ciò che ci si avvicina troppo e oltrepassa il cosiddetto *orizzonte degli eventi* - si tratti di fotoni senza massa o di particelle pesanti - irrimediabilmente ci cade dentro. Dal punto di vista fisico, e assumendo che abbiano carica elettrica nulla, i buchi neri sono completamente definiti dalla massa e dalla velocità di rotazione. Di conseguenza, se studiati individualmente e isolatamente, sono relativamente facili da descrivere. Ciò che non è ancora propriamente compreso, invece, sono gli effetti della loro interazione con l'ambiente circostante, che gli scienziati cercano di studiare sia con simulazioni numeriche (al computer) sia con osservazioni di buchi neri più o meno lontani.

Gli astronomi sono d'accordo ora nell'affermare che possiamo trovare un buco nero supermassiccio al centro di quasi ogni galassia, e nel caso in cui il buco nero centrale sia in fase di crescita allora diciamo che la sua galassia ospita un nucleo galattico attivo (AGN, dall'inglese *active galactic nucleus*). Durante le prime fasi dell'evoluzione della galassia, la morte di una grande stella o il collasso gravitazionale di una nuvola di gas primordiale (cioè composta esclusivamente da idrogeno ed elio) possono dare luogo ad un buco nero neonato, che può in seguito crescere inglobando il gas e le stelle che rimangono intrappolate nel suo campo gravitazionale. Si stima che circa il 90% della materia che cade sulla superficie del buco nero viene effettivamente incorporato, mentre il rimanente 10% si trasforma in energia elettromagnetica (luce) che viene poi emessa nell'ambiente circostante. La luce rilasciata dal buco nero copre tutto lo spettro elettromagnetico e può apparire molto luminosa ai nostri telescopi; sono proprio i telescopi a dirci che questi buchi neri (che possono anche raggiungere la stessa massa di una galassia come la Via Lattea) già avevano dimensioni enormi solo qualche centinaio di milioni di anni dopo il Big Bang, quindi quando l'Universo aveva circa un quindicesimo dell'età

che gli attribuiamo ora. Una frazione dell'energia emessa dal buco nero interagisce con il gas e la polvere presenti nella galassia (anche chiamati mezzo interstellare). Questa interazione prende il nome di feedback di buco nero, e comporta un incremento della temperatura delle particelle di gas e di polvere, che quindi incominciano a muoversi più velocemente fino a riuscire a liberarsi dall'attrazione gravitazionale della galassia e quindi scappare da essa. Siccome questo gas è anche la materia prima dalla quale nascono le stelle, ci aspettiamo che il feedback di buco nero abbia un impatto sulla velocità con cui la galassia forma nuove stelle. Questo significa che capire come i buchi neri crescono è un tassello importante anche per capire l'evoluzione delle galassie ospiti.

Dato il contesto abbozzato qui sopra, possiamo dedurre quali siano le domande principali riguardo alla crescita e all'evoluzione dei buchi neri che ancora adesso assillano gli studiosi, e a cui proveremo a rispondere in questa tesi. Come fanno i buchi neri a crescere così tanto così in fretta? Qual è l'effetto di questa crescita sulla galassia ospite? Visto il processo di feedback, esiste una correlazione fra le proprietà fisiche dei buchi neri e quelle delle galassie ospiti? Fino a che punto la crescita del buco nero dipende dalla massa e dalla stazza della galassia ospite? Nei prossimi capitoli toccheremo tutti questi punti.

Questa tesi

In questo contesto, studiamo la co-evoluzione dei buchi neri centrali e delle loro galassie durante il primo miliardo di anni di vita dell'Universo, in un momento in cui le galassie sono in forte crescita ma ancora molto giovani e di dimensioni limitate. In astronomia abbiamo la possibilità di osservare oggetti così distanti che la luce che emettono deve viaggiare per lungo tempo prima che ci raggiunga. Questo significa che quando puntiamo un telescopio su uno di questi oggetti stiamo osservando luce antica che ci porta informazioni dal passato. In effetti, le emissioni radio provenienti dagli AGN (anche detti quasar) più lontani che riusciamo ad osservare hanno viaggiato per più di tredici miliardi di anni prima di raggiungere la Terra. Grazie a ciò, adesso abbiamo a disposizione informazioni sulle proprietà statistiche delle primissime generazioni di galassie e buchi neri, e in particolare sulla loro luminosità nella banda ultravioletta (UV) dello spettro, sulla loro massa (anche se in questo caso le misure sono ancora molto incerte), su come sono distribuiti nello spazio, e su come il loro numero varia nel tempo. Abbiamo quindi sviluppato *Delphi*, un modello per simulare la crescita e l'evoluzione della galassie e dei loro buchi neri

centrali che sia in grado di riprodurre le proprietà osservate e nello stesso tempo di darci nuove informazioni riguardo a caratteristiche che ancora sfuggono ai telescopi.

Nel capitolo 1 introduciamo le questioni affrontate nei capitoli successivi, contestualizzando il nostro lavoro rispetto allo stato generale della ricerca sui buchi neri supermassicci dell'Universo giovane. Cercheremo di ricapitolare i progressi, sia osservativi che teorici, avvenuti soprattutto nel corso degli ultimi decenni.

Nel capitolo 2 costruiamo il modello che useremo, ne testiamo i risultati e ci assicuriamo che essi siano compatibili con le osservazioni di galassie e buchi neri a nostra disposizione. Procediamo poi calcolando il numero di collisioni e fusioni fra buchi neri. Infatti, quando due galassie si fondono, i loro rispettivi buchi neri migrano verso il centro della nuova galassia, formando un sistema binario di buchi neri. Man mano che i buchi neri si avvicinano fra loro, orbitando luno attorno all'altro, emettono onde gravitazionali che trasportano un segnale specifico e sempre diverso, a seconda della massa dei buchi neri. Sfruttando questa caratteristica, la futura missione spaziale *LISA*, un interferometro costruito apposta per rilevare onde gravitazionali, sarà in grado di osservare le fusioni fra buchi neri da grande distanza. Con il nostro modello possiamo predire quante di queste fusioni saranno effettivamente osservate da *LISA* una volta che la missione sarà operativa. Il paragone fra i risultati così ottenuti e quelli della missione spaziale sarà estremamente utile nel testare la validità e le assunzioni del nostro modello.

Utilizzando lo stesso modello, nel capitolo 3 tentiamo di far luce su come riescano i buchi neri a crescere, talvolta fino a $10^9 - 10^{10}$ volte la massa del Sole, già durante il primo miliardo di anni di vita dell'Universo (ricordiamo che attualmente l'età dell'Universo è stimata essere di quasi quattordici miliardi di anni). Seguiamo, all'interno del nostro modello, i dettagli dell'evoluzione delle galassie e dei loro buchi neri centrali. Questi ultimi possono crescere o attraverso fusioni con altri buchi neri o con l'accrescimento di gas e polvere dal mezzo interstellare. Se la prima opzione risulta dominante quando i buchi neri sono ancora molto giovani e risiedono in galassie piccole, col tempo la seconda prende gradualmente il sopravvento. In particolare, la crescita dei buchi neri più piccoli si velocizza col passare del tempo, mentre quella dei buchi neri più massicci mostra un andamento opposto, e tende a rallentare durante l'intervallo di tempo considerato.

Infine, nel capitolo 4, studiamo gli effetti della crescita dei buchi neri e del loro feedback sulle galassie ospiti. Come prima cosa notiamo che questo meccanismo porta ad un aumento della radiazione UV emessa

dalle galassie. In particolare, in virtù del nostro approccio, riusciamo a distinguere le due componenti che contribuiscono alla radiazione UV totale che emerge dalle galassie, quella data dall'attività del buco nero e quella proveniente invece dalle stelle. Siamo quindi in grado di seguire l'evoluzione temporale di queste due componenti separatamente, cosa che risulta invece impossibile con le informazioni raccolte dai telescopi. In questo caso la nostra ricerca teorica risulta complementare ai risultati delle osservazioni. Inoltre, dai nostri risultati pare che l'energia emessa dal buco nero in crescita sia effettivamente in grado di rallentare la formazione di nuove stelle, dal momento che causa l'eiezione di importanti quantità di gas dalla galassia. C'è però da notare che questo è vero solo per alcune galassie. Quelle minori, infatti, ospitano buchi neri ancora piccoli che non producono abbastanza energia per espellere il gas. D'altro canto, il forte campo gravitazionale delle galassie più grandi agisce sul gas, prevenendone la fuoriuscita nonostante l'attività del buco nero centrale. Possiamo quindi concludere che il feedback di buco nero produce i massimi effetti nelle galassie di dimensioni intermedie.

IL lavoro presentato in questa tesi è importante specialmente rispetto al futuro della ricerca sui buchi neri antichi. I telescopi spaziali delle prossime generazioni - come il James Webb Space Telescope, *Athena*, *Lynx*, *LISA*, il cui lancio è programmato nell'arco dei prossimi dieci anni - saranno in grado di osservare direttamente la popolazione di buchi neri di cui ci occupiamo in questa tesi, e quindi di dirci quali aspetti del nostro modello funzionano e quali devono essere ancora migliorati. Proprio questo lavoro su due fronti, in cui lo sviluppo dei modelli teorici informa le future missioni osservative e viceversa, ha il potenziale per portarci ad una comprensione più completa e precisa del ruolo giocato dai buchi neri supermassicci nel nostro Universo.

SAMENVATTING

Zwarte gaten zijn extreem dichte en compacte objecten, zo zijn bijvoorbeeld zwarte gaten met dezelfde massa als die van de aarde ongeveer even groot als een golfbal. Dit zorgt ervoor dat er in hun nabije omgeving een sterke zwaartekrachtsveld ontstaat wat de geometrie van de ruimtetijd buigt en kromt. Alles wat hierbij te dichtbij komt, en de zogenaamde event horizon passeert, valt er uiteindelijk in, dit geldt voor zowel massieve deeltjes en fotonen (lichtdeeltjes). Vanuit een fysisch oogpunt worden zwarte gaten gedefinieerd door hun massa en hun rotatiesnelheid, hierbij ervan uitgaande dat ze geen elektrische lading hebben. Zodoende zijn zij gemakkelijk te beschrijven als ze individueel en in een geïsoleerde omgeving worden bestudeerd. De details van hun interactie met de omgeving zijn echter nog niet goed begrepen, wat momenteel bestudeerd wordt door middel van computersimulaties en observaties.

Onderzoekers zijn het erover eens dat er in het centrum van bijna elk sterrenstelsel een (centraal) superzwaar zwart gat gevonden kan worden. Als dit centrale zwarte gat groeit heeft het sterrenstelsel een Actieve Galactische Kern (AGN). Tijdens de vroege fasen van de evolutie van een sterrenstelsel kan de ineenstorting van een zeer massieve ster of een klomp ongerept gas een zaadje voor een zwart gat achterlaten, dat vervolgens groeit door zich te voeden met het gas en de sterren die vastzitten in zijn zwaartekrachtsveld. Geschat wordt dat ongeveer 90% van het materiaal dat op het zwarte gat valt daadwerkelijk wordt opgeslokt, terwijl de resterende 10% wordt omgezet in energie die uitgestoten wordt. Deze energie dekt het hele elektromagnetische spectrum en kan erg lichtgevend zijn. De emissie van superzware zwarte gaten (die soms de massa van ons Melkwegstelsel kunnen bereiken) vertelt ons dat deze reuzen al aanwezig waren toen het heelal slechts een paar honderd miljoen jaar oud was, ongeveer een vijftiende van de huidige leeftijd. Een fractie van de energie die wordt uitgestraald interacteert met het gas en stof dat in het sterrenstelsel aanwezig is (bekend als interstellair medium, ISM). Deze interactie wordt feedback van zwarte gaten genoemd en resulteert in een verhoging van de temperatuur van de gas deeltjes, die als gevolg hiervan sneller gaan

bewegen en uit het sterrenstelsel kunnen ontsnappen. Aangezien het gas de bouwsteen vormt voor sterren, verwachten men deze feedback invloed zal hebben op de snelheid waarmee nieuwe sterren in het sterrenstelsel gevormd worden. Dit suggereert dat de groei van superzwarte zwarte gaten ook belangrijk is voor de evolutie van sterrenstelsels.

De afbeelding hierboven geeft aan wat de belangrijkste onbeantwoord vragen zijn met betrekking tot de groei en evolutie van superzwarte zwarte gaten. Hoe kunnen zwarte gaten in relatief korte tijd zo groot worden? Hoe beïnvloedt de groei van de zwarte gaten de omliggende sterrenstelsels? Hoe zijn de eigenschappen van het centrale zwart gat gecorreleerd met de fysieke kenmerken van het sterrenstelsel? Hoe hangt de groei van zwarte gaten af van de grootte en massa van het omliggende sterrenstelsel? Dit zijn de belangrijkste punten die in dit werk aan de orde komen.

Dit proefschrift

Wij bestuderen de co-evolutie van sterrenstelsel en hun centrale zwarte gaten in de eerste miljard jaar van het heelal, wanneer sterrenstelsels snel groeien maar nog erg jong zijn. In de sterrenkunde beschikt men over de mogelijkheid om objecten te observeren die zo ver weg staan dat het licht dat ze uitstralen een lange tijd moet reizen voordat het ons bereikt: wij zien deze objecten zoals zij waren in het verleden. Zo heeft de radio-emissie van de verste waargenomen AGN er meer dan 13 miljard jaar overgedaan om de aarde te bereiken. Dit licht geeft ons een deel van de statistische informatie die deze zeer jonge zwarte gaten en sterrenstelsel populatie karakteriseert. Wij beschikken voornamelijk over hun lichtsterkte in het ultraviolet (UV), maar ook over hun massa (ondanks dat de meetonzekerheden van deze metingen nog steeds groot zijn), hun ruimtelijke verdeling en hun aantal als functie van de tijd. Wij hebben een computermodel gebouwd, genaamd Delphi, om de vorming, groei en evolutie van sterrenstelsels en hun centrale zwarte gaten te simuleren, om de observaties en hun afgeleide parameters te reproduceren, en tegelijkertijd om inzicht te verkrijgen in nog niet eerder waargenomen eigenschappen.

In Hoofdstuk 2 bouwen we het computermodel, testen we de resultaten en zorgen we ervoor dat het model de belangrijkste huidige observaties van zwarte gaten en eigenschappen van sterrenstelsels reproduceert. Vervolgens berekenen we de snelheid waarmee zwarte gaten fuseren. Wanneer sterrenstelsels botsen, migreren elk van hun zwarte gaten naar het centrum van het nieuwe gevormde sterrenstelsel. Deze zwarte gaten vormen een binair systeem voordat ze uiteindelijk samensmelten. Het

steeds kleiner wordende binaire systeem zendt zwaartekrachtsgolven uit met zeer specifieke golfvormen, waarvan de vorm afhangt van de massa van de zwarte gaten. Gebruikmakend van deze eigenschap zal de toekomstige ruimtemissie *LISA*, een interferometer die is gebouwd om zwaartekrachtsgolven te detecteren, in staat zijn om het samensmelten van zwarte gaten op grote afstanden waar te nemen. Wij voorspellen met ons model het aantal fusies tussen zwarte gaten dat *LISA* zou moeten waarnemen zodra het in werking treedt. Hiermee zullen wij instaat zijn om de validiteit en aannames van ons model te testen.

Met hetzelfde model onderzoeken wij in Hoofdstuk 3 hoe superzware zwarte gaten met een massa tot $10^9 - 10^{10}$ zonnemassa's groeien binnen het eerste miljard jaar van het heelal. Hierbij volgen wij de evolutie van beide sterrenstelsels en hun centrale zwarte gaten op de voet. Deze kunnen toenemen in massa door fusies met andere zwarte gaten of door de aanwas van gas en stof uit het ISM. Terwijl het eerste scenario lijkt te domineren wanneer zwarte gaten nog erg jong zijn en worden omgeven door kleinere sterrenstelsels, wordt het tweede scenario geleidelijk aan belangrijker naarmate de tijd verstrijkt. Wij zien ook dat kleine zwarte gaten op latere tijdstippen sneller groeien, terwijl de grotere zwarte gaten een tegenovergestelde gedrag vertonen en in vroegere tijden sneller lijken te groeien.

Ten slotte bestuderen we in Hoofdstuk 4 hoe de groei en feedback van zwarte gaten het omliggende sterrenstelsel beïnvloed. Ten eerste veroorzaakt de groei een toename in UV licht. Wij zijn met onze aanpak in staat om de UV bijdragen van het groeiende zwarte gat te onderscheiden van die van de sterren in het sterrenstelsel. Dit betekent dat wij kunnen zien welke sterrenstelsels worden gedomineerd door UV emissie van zwarte gaten en welke door stellaire UV emissie, waarbij dat laatste veel informatie geeft over de evolutiegeschiedenis van de sterren. Wij ontdekten dat de energie die vrijkomt tijdens de groei van het zwarte gat een enorme hoeveelheid gas uit het sterrenstelsel duwt, waardoor de vorming van sterren in het omliggende sterrenstelsel vertraagd wordt. Dit geldt echter niet voor alle sterrenstelsels. Zwarte gaten in kleinere (minder zware) sterrenstelsels produceren niet genoeg energie om het omringende gas te beïnvloeden. Aan de andere kant kunnen de grootste (en zwaarste) sterrenstelsels het gas vast te houden door hun sterke gravitatiekracht, ondanks de activiteit van het zwarte gat. Dit betekent dat feedback van zwarte gaten de grootste impact heeft in sterrenstelsels met een gemiddelde massa.

Het werk in dit proefschrift is vooral belangrijk als we kijken naar de toekomst van de astronomie. Toekomstige ruimtetelescopen - zoals de

James Webb Space Telescope, *Athena*, *Lynx* - zullen in staat zijn om de hier beschreven populatie van zwarte gaten direct te observeren. Op deze manier kan ons model getest en verfijnd worden, wat mogelijk een beter begrip oplevert van de rol die deze superzware zwarte gaten spelen in ons heelal.

1. INTRODUCTION

Black holes are often considered among the most exotic and fascinating objects in the Universe. Not only because they exist in an extreme physical state, with a density theoretically comparable to that of atomic nuclei ($\sim 10^{14}$ g/cm³), but also because their elusive nature, a consequence of the fact that photons cannot escape their gravitational field, has triggered the fantasy of generations of scientists, writers, movie directors. Indeed, the detection of a black hole poses several challenges. It is enough to think that they were first theorised in 1916 by Karl Schwarzschild, as a solution of general relativity (Schwarzschild 1916), but we had to wait for more than a century, until 2019, for the first direct image of a black hole and its vicinity (Event Horizon Telescope Collaboration et al. 2019). This picture was obtained only thanks to the unprecedented joint efforts of eight radiotelescopes around the world, 60 different institutes in 20 countries and more than 300 collaborators. Alongside the big place they occupy in the collective imagination, several reasons contribute to making black holes very interesting also from the scientific point of view. In fact, they are thought to be at the centre of most of the galaxies we are able to observe today, in correspondence with the peaks of the density field of the Universe. Their growth and evolution are thus linked to that of the big structures that we can observe. They can grow up to tens of billions of solar masses, and during their growth they convert part of the accreted mass into energy, which is then injected into the surrounding galaxies, affecting their star formation history. The formation, evolution and feedback of the first black holes of the Universe is the object of this thesis.

1.1 The cosmological context

The success of any comprehensive, coherent and self-sustained cosmological model depends on its ability to explain some fundamental observable characteristics of our Universe. Fortunately, this effort is facilitated by the cosmological principle, the underlying basic assumption stating that the physical properties of the Universe are the same when observed from any

location in the Universe and any direction. Indeed, numerical simulations and observations suggest a homogeneity scale of the order of 300 – 400 Mpc (Yadav et al. 2010; Ntelis et al. 2017). Conversely, the distribution of energy and matter is highly perturbed at smaller scales, which are dominated by fluctuations in the cosmic density field taking the form of the walls, filaments and clusters of galaxies forming the cosmic web. The spatial and mass distribution of the structures we can observe in the Universe is not the only key physical property we have to explain: we know that the receding velocity between any two different points of the Universe is proportional to their proper relative distance (the Hubble’s law; Lemaître 1927; Hubble 1929); we measure a cosmic microwave background (CMB) radiation with a temperature of $T_0 = 2.72548 \pm 5.7 \times 10^{-4}$ K (Fixsen 2009) and fluctuations of the order of 10^{-5} ; we observe the presence of ordinary baryonic matter but almost no antimatter, and we measure the baryonic matter to amount to only $\approx 4.7\%$ of the total content of the Universe, which for the remaining part consists of dark matter and dark energy (Planck Collaboration et al. 2016b); finally, we know that Hydrogen accounts for $\approx 75\%$ of the total baryonic matter, Helium for $\approx 25\%$ and other elements are present in traces.

1.1.1 The hot big bang model

The cosmological model that is currently adopted by the scientific community to describe the Universe, successfully reproducing these characteristics, is the so-called Λ CDM model, which is based on the validity of Einstein’s general relativity as a theory of gravity (Einstein 1917) and on the presence of dark matter and dark energy, accounting for $\sim 95\%$ of the content of the Universe (Planck Collaboration et al. 2016b). According to this cosmological picture, the initial phases of the evolution of the Universe are described by the Hot Big Bang Model. This evolution proceeds in steps, each corresponding to a specific physical state defined by the temperature of the cosmic microwave background, which decreases with decreasing redshift according to the relation $T_{\text{CMB}} = T_0(1 + z)$ (Noterdaeme et al. 2011). Initially, about 13.8 billion years ago (Planck Collaboration et al. 2016a), the Universe was in a singular state of extreme density and temperature. Eventually, this singularity broke down and the Universe started to expand. After about 10^{-32} s, the Universe underwent a period of exponential growth called inflation (Guth 1981; Linde 1982; Albrecht & Steinhardt 1982), which increased its size by 60 orders of magnitude within a time span of 10^{-34} seconds. During this period, primordial quantum fluctuations, the seeds of the perturbations that eventually collapsed

into the clusters, walls and filaments of galaxies we see today, arose. The strength of inflation theory is that it explains the origins of these primordial inhomogeneities, and has predictive power with respect to their spectrum. During inflation, the characteristic scale of these first quantum perturbations is stretched from close to the Planckian length to galactic magnitude, while the relative amplitudes remain unchanged. Inflation theory offers natural and simple solutions to other problems as well (for a review about inflation, see Senatore 2017): a period of accelerated expansion of the Universe would explain the high level of homogeneity in the distribution of the energy density across casually disconnected regions (horizon problem), suggesting that these regions were casually connected before inflation took place. Besides, by allowing a period of non-adiabatic exponential expansion, it would also solve the flatness problem, linked to the fine-tuning of the density of the Universe (Guth 1981). At the end of inflation, the Universe consisted of an extremely hot plasma of quarks and gluons, together with other elementary particles; a breakdown of the symmetry between matter and antimatter allowed the first baryons and leptons to form permanently. As expansion proceeds and the temperature drops, quarks can come together to form hadrons, which in turn will eventually be able to form the first nuclei. This process goes under the name of Big Bang nucleosynthesis (Alpher et al. 1948), and the Hot Big Bang Model predicts the respective abundances of the light elements D, ^3He and ^4He , as well as the baryon-to-photon ratio.

At a temperature of around 3000 K and a redshift of 1100, recombination takes place: the Universe is now cool enough to allow the formation of hydrogen and helium atoms, and the lack of free electrons onto which to scatter causes the decoupling of the photons. The photons from the last-scattering surface (CMB) are then free to travel with an infinite mean free path, and we can observe them, redshifted, today (McKellar 1941; Penzias & Wilson 1965).

Despite the numerous important achievements of the Hot Big Bang Model, some unanswered questions and open issues still persist. One of the most recent points raised is that of the primordial abundance of Lithium, since observations do not manage to account for all the amount of Lithium that should have been produced during the Big Bang nucleosynthesis (for a review, see Fields 2011). Nevertheless, this discrepancy might have an Astrophysical source (for instance, the presence of unaccounted systematics in the measurements of Lithium abundance in old metal-poor stars). Finally, we cannot disregard the fact that the Standard Model of Physics is able to accurately describe the interactions between the different types of particles only below the energy threshold represented by the most

powerful particle accelerators, which is of the order of a few TeVs. This means we do not know how to model the very early phase of the Universe, corresponding to much higher energies. In addition, as of now, we do not have an explanation for the observed asymmetry between regular matter and antimatter, nor for the nature of dark matter or dark energy.

1.1.2 The formation of the first structures

With the expanding background, the progressive decrease of the temperature, and the passage to a matter-dominated phase at $z \sim 3600$, eventually gravity becomes the dominant driving force of the evolution of the Universe, and matter starts collapsing onto the local density peaks, forming the first structures.

1.1.2.1 Newtonian linear perturbation theory

Initially, the amplitudes of the fluctuations are small enough for us to be able to approximate their evolution with Newtonian linear perturbation theory. This is valid as long as their contrast density with respect to the background is much less than unity, their size is smaller than the horizon size so that the casual links can be considered instantaneous, and the matter content can be considered as a non-relativistic fluid. We can use this approach since here we are interested in the gravitational collapse of non-relativistic particles like dark and baryonic matter. If the perturbations are relative to the distribution of relativistic particles (like photons and neutrinos) or are characterised by larger amplitudes, we would need the relativistic theory of linear perturbations. For the analytical treatment shown in this and the next subsection, we will mostly refer to the volume *Galaxy and Formation Evolution*, by Mo et al. (2010).

In a gas of baryonic particles the frequent collisions can establish local thermodynamic equilibrium between the particles. This allows us to assume that the mean free path of the particles of the fluid is much smaller than the scale of interest. In addition, this formulation is valid also for a dark matter fluid whose particles have a small velocity dispersion. We can express the density evolution as a function of time t and (comoving) coordinates $\mathbf{x}$ by writing $\rho(\mathbf{x}, t) = \bar{\rho}(t) [1 + \delta(\mathbf{x}, t)]$, where $\bar{\rho}$ is the background average density and $\delta(\mathbf{x}, t)$ is the density perturbation contrast. The evolution of a fluid under the influence of a gravitational potential Φ can then be described by the continuity equation (describing the conservation of matter), the Euler equations (equations of motion) and the Poisson

equation (describing the gravitational field). We can then write the system

$$\frac{\partial \delta}{\partial t} + \frac{1}{a} \nabla \cdot [(1 + \delta) \mathbf{v}] = 0, \quad (1.1)$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{\dot{a}}{a} \mathbf{v} + \frac{1}{a} (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{\nabla \Phi}{a} - \frac{c_s^2}{a} \frac{\nabla \delta}{(1 + \delta)} - \frac{2T}{3a} \nabla S, \quad (1.2)$$

$$\nabla^2 \Phi = 4\pi G \bar{\rho} a^2 \delta, \quad (1.3)$$

where $\mathbf{v}$ is the peculiar velocity of the fluid, $a \equiv (1 + z)^{-1}$ is the scale factor of the Universe and c_s is the characteristic sound speed, T is the temperature and S is the entropy. The $\nabla \equiv \nabla_{\mathbf{x}}$ and $\partial/\partial t$ notations represent respectively the spatial and temporal partial derivative. Notice that if we wanted to study the collapse of dark matter halos without baryonic matter, we could apply the same treatment, neglecting the last two terms on the right-hand side of Eq. 1.2. Linearising the system and solving for δ yields the evolution equation of the density contrast for $\delta \ll 1$

$$\frac{\partial^2 \delta}{\partial t^2} + 2\frac{\dot{a}}{a} \frac{\partial \delta}{\partial t} = 4\pi G \bar{\rho} \delta + \frac{c_s^2}{a^2} \nabla^2 \delta + \frac{2}{3} \frac{\bar{T}}{a^2} \nabla^2 S, \quad (1.4)$$

where S is the entropy and $\bar{T}$ is the background average temperature. We can find the solution for each of the modes of the perturbation by Fourier transforming Eq. 1.4. If we assume the mode function to be a plane wave, we can write

$$\delta(\mathbf{x}, t) = \sum_{\mathbf{k}} \delta_{\mathbf{k}}(t) \exp(i\mathbf{k} \cdot \mathbf{x}), \quad (1.5)$$

$$\delta_{\mathbf{k}}(t) = \frac{1}{V_u} \int \delta(\mathbf{x}, t) \exp(-i\mathbf{k} \cdot \mathbf{x}) d^3 \mathbf{x}, \quad (1.6)$$

where V_u is the volume within which the perturbation is assumed to be periodic. The Fourier-transformed equation becomes then

$$\frac{d^2 \delta_{\mathbf{k}}}{dt^2} + 2\frac{\dot{a}}{a} \frac{d\delta_{\mathbf{k}}}{dt} = \left[4\pi G \bar{\rho} - \frac{k^2 c_s^2}{a^2} \right] \delta_{\mathbf{k}} - \frac{2}{3} \frac{\bar{T}}{a^2} k^2 S_{\mathbf{k}}. \quad (1.7)$$

If we assume isentropic perturbations with adiabatic expansion we can drop the third term on the right-hand side, which represents isocurvature perturbations seeded by spatial variations in pressure. Since we are dealing with small perturbations, we can neglect for now also the Hubble dragging and the second term on the left-hand side, related to the expansion of the Universe. We then obtain two solutions for $\delta_{\mathbf{k}}$, one corresponding to a

decaying mode and the other one to a growing mode. What distinguishes one from the other is the characteristic scale of the perturbation, given a certain mass. By setting $4\pi G\bar{\rho} - k^2 c_s^2/a^2 = 0$, we then obtain the characteristic Jeans length

$$\lambda_J \equiv \frac{2\pi a}{k_J} = c_s \sqrt{\frac{\pi}{G\bar{\rho}}}, \quad (1.8)$$

indicating the distance that a sound wave can travel in a gravitational free-fall time $t_{\text{ff}} \sim (G\rho)^{-1/2}$. Hence, if the characteristic length of the perturbation is $\lambda > \lambda_J$, the pressure can no longer counteract the gravitational contraction and the structure collapses onto itself. On the other hand, if $\lambda < \lambda_J$, the perturbation will propagate through space at sound speed without collapsing. The inclusion of the Hubble drag term, as expected, would slow down the collapse of the perturbation when $\lambda > \lambda_J$ and damp the pressure waves when $\lambda < \lambda_J$. We can also define the Jeans mass as the mass within a sphere of radius $\lambda_J/2$, limit above which the perturbation will collapse

$$M_J = \frac{4\pi}{3} \bar{\rho} \left(\frac{\lambda_J}{2} \right)^3. \quad (1.9)$$

An important solution of Eq. 1.7 is obtained for scales $k \ll k_J$, in the case of an isentropic and pressureless fluid in an expanding Universe. In this case we can write

$$\frac{d^2 \delta_{\mathbf{k}}}{dt^2} + 2 \frac{\dot{a}}{a} \frac{d\delta_{\mathbf{k}}}{dt} = 4\pi G\bar{\rho} \delta_{\mathbf{k}}. \quad (1.10)$$

It is possible to show analytically that the evolution of the decaying mode δ_- is proportional to that of the Hubble constant, with $\delta_- \propto H(t)$. Carroll et al. (1992) instead offered a numerical approximation of the solution corresponding to the growing mode of the perturbation, finding

$$\delta_+ \propto D(z) \propto g(z)/(1+z), \quad (1.11)$$

where $D(z)$ is the linear growth rate, and

$$g(z) \approx \frac{5}{2} \Omega_m(z) \left\{ \Omega_m^{4/7}(z) - \Omega_\Lambda(z) + [1 + \Omega_m(z)/2] [1 + \Omega_\Lambda(z)/70] \right\}^{-1}. \quad (1.12)$$

The advantage of this approximation is that the different modes of the perturbations are not coupled, and we can write

$$\delta_{\mathbf{k}}(z) = D(z) \delta_{\mathbf{k}}^0, \quad (1.13)$$

meaning that each mode of the perturbation can be followed individually, and it grows proportionally to the linear growth rate.

1.1.2.2 Newtonian non-linear perturbation theory

If we expand perturbation theory to higher orders, then coupling terms appear, but they remain negligible in the limit $\delta \ll 1$. When $\delta \gtrsim 1$, these terms need to be accounted for when studying the evolution of the perturbation, and they make it impossible to find an analytical solution to the system formed by Eqs. 1.1, 1.2 and 1.3. Nevertheless, a perturbative solution can be found (for a review see Bernardeau et al. 2002) by expanding the solution, and writing it as

$$\delta(\mathbf{x}, a) = \sum_{n=1}^{\infty} \delta^{(n)}(\mathbf{x}, a) \simeq \sum_{n=1}^{\infty} D_+^n(a) \delta^{(n)}(\mathbf{x}), \quad (1.14)$$

where $\delta^{(n)}(\mathbf{x}) \propto \delta^n(\mathbf{x})$. Such nonlinearities couple the different modes of the perturbation together, violating Gaussianity. In addition, if one substitutes Eq. 1.14 into the non-linearised form of Eq. 1.7, they will get a solution which depends on a coupling term, and this coupling term is proportional to the cosine of the angle between two different perturbations, meaning that perturbations with the same direction of propagation are subject to a stronger coupling. In this sense, we can say that the perturbation modes in the non-linear regime are resonant.

An alternative to the Euler non-linear perturbative approach is the Lagrangian perturbation theory (Zel'Dovich 1970). This elegant approach describes the evolution of the displacement between the initial position of the particles $\mathbf{q}$ and their position $\mathbf{x}$ at a time t by directly linking it to the physical gravitational potential. Nevertheless, this approximation breaks down when the trajectories of the particles start crossing each other, hence its validity is limited to the first stages of structure formation since it cannot account for the oscillations after the first collapse and before the virialisation of the system.

To study the formation of a dark matter halo, which eventually will reach a density contrast with respect to the background $\delta \gg 1$, the most accurate approximation with an analytical solution that we can find is the spherical collapse. The equation of motion of a mass shell with spherical symmetry is given by

$$\frac{d^2 r}{dt^2} = -\frac{GM}{r^2} + \frac{\lambda}{3} r, \quad (1.15)$$

where the last term accounts for the expansion of the Universe. Integrating this equation and writing the mass contained within the overdensity as $M = (1 + \delta_i) \Omega_{m,i} \rho_{\text{crit}}(t_i) (4\pi r_i^3/3)$, with t_i indicating the epoch of

initial collapse of the perturbation, we obtain the value for minimal overdensity for spherical collapse $\delta_c(t_{\text{coll}}) \approx 1.686$, which is an extremely good approximation for all realistic cosmologies and epochs considered. As the collapse proceeds, the perturbation tends to virialise and reaches an equilibrium, and at this moment the dark matter halo can be fully characterised by virial radius, temperature and circular velocity. Using the virial theorem, as in Barkana & Loeb (2001), we can derive the value of the overdensity at its final state, $\Delta_f = 18\pi^2 \simeq 178$, and the virial radius, which corresponds to

$$R_{\text{vir}} = 0.784 \left(\frac{M}{10^8 h^{-1} \text{ M}_\odot} \right)^{1/3} \left[\frac{\Omega_m}{\Omega_m^z} \frac{\Delta_f}{18\pi^2} \right]^{-1/3} \left(\frac{1+z}{10} \right)^{-1} h^{-1} \text{ kpc}, \quad (1.16)$$

where $\Omega_m^z = \Omega_m(1+z)^3 / [\Omega_m(1+z)^3 + \Omega_\Lambda + \Omega_k(1+z)^2]$. We can then estimate the virial circular velocity $V_c = (GM/R_{\text{vir}})^{1/2}$ and the virial temperature $T_{\text{vir}} = \mu m_p V_c^2 / (2k_B)$, where μ is the mean molecular weight and m_p the proton mass.

1.1.2.3 Statistical description of the cosmic density field

In order to test the validity of a cosmological perturbation theory, one should first establish a way to relate theory and observations of the cosmic density field. For this purpose, a satisfactory statistical description of the properties of cosmic perturbations is required. This is achieved by studying the moments of the probability distribution function characterising the perturbation field $\delta(\mathbf{x})$. This task is made simpler by the fact that cosmic inflation naturally produces Gaussian or almost Gaussian primordial perturbation fields, that are completely and exclusively determined by the second-order moment of the Gaussian distribution function. In particular, the one-point distribution function of a Gaussian field is defined as

$$P_x(\delta)d\delta = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp -\frac{\delta^2}{2\sigma^2} d\delta, \quad (1.17)$$

which describes the probability that the field δ has a value between δ and $\delta + d\delta$ at the location x , and where σ^2 is the variance of the field. In this case the second moment is the only non-zero moment of the distribution function, it is called the two-point correlation function and is defined as

$$\xi(x) = \langle \delta_1 \delta_2 \rangle, \quad \text{with } x \equiv |\mathbf{x}_1 - \mathbf{x}_2|. \quad (1.18)$$

Note that $\xi(0) = \sigma^2$. In Fourier space the field $\delta(\mathbf{x})$ is written as

$$\delta_{\mathbf{k}} = \frac{1}{V_u} \int \delta_{\mathbf{x}} \exp(-i\mathbf{k} \cdot \mathbf{x}) d\mathbf{x}, \quad (1.19)$$

where $V_u = L_u^3$ is the volume in which the perturbation is periodic, and $\mathbf{k}$ are the modes of the perturbations. The second moment in Fourier space, denominated power spectrum, becomes

$$P(k) \equiv V_u \langle |\delta_{\mathbf{k}}|^2 \rangle \equiv V_u \langle \delta_{\mathbf{k}} \delta_{-\mathbf{k}} \rangle. \quad (1.20)$$

Substituting the definition of $\delta_{\mathbf{k}}$ we obtain

$$P(k) = \int \xi(\mathbf{x}) \exp(-i\delta_{\mathbf{k}} \cdot \delta_{\mathbf{x}}) d^3\delta_{\mathbf{x}} = 4\pi \int_0^\infty \xi(r) \frac{\sin kr}{kr} r^2 dr. \quad (1.21)$$

Conversely, we have that

$$\xi(r) = \frac{1}{(2\pi)^3} \int P(k) \exp(i\mathbf{k} \cdot \mathbf{r}) d^3\mathbf{k} = \frac{1}{(2\pi)^3} \int_0^\infty P(k) \frac{\sin kr}{kr} k^2 dk, \quad (1.22)$$

from which it appears clearly that the power spectrum and the two-point correlation function are one the Fourier transform of the other.

In order to consider the formation of structures at a given mass or size, it is useful to introduce a spherically symmetric window function with characteristic radius R to get the smoothed field

$$\delta_s(\mathbf{x}; R) = \int \delta_0(\mathbf{x}') W(\mathbf{x} + \mathbf{x}'; R) d^3\mathbf{x}', \quad \text{with} \quad \int W(\mathbf{x}; R) d^3\mathbf{x} = 1, \quad (1.23)$$

where $\delta_0(\mathbf{x})$ is the overdensity field linearly extrapolated to the present time. The mass contained within the characteristic radius R can be written as $M = 4/3\pi\bar{\rho}R^3$ for a spherical top-hat filter function. Given this smoothed density field, Press & Schechter (1974) derived an approximation for the mass function of collapsed objects by assuming that the probability that $\delta_s > \delta_c(t)$ corresponds to the fraction of mass elements in halos with mass greater than M at epoch t . Since we are considering a Gaussian field we can write the probability as

$$P[> \delta_c(t)] = \frac{1}{\sqrt{2\pi}\sigma(M)} \int_{\delta_c(t)}^\infty \exp\left[-\frac{\delta_s^2}{2\sigma^2(M)}\right] d\delta_s, \quad (1.24)$$

where

$$\sigma^2(M) = \langle \delta_s^2(\mathbf{x}; R) \rangle = \frac{1}{2\pi^2} \int_0^\infty P(k) \tilde{W}^2(\mathbf{k}R) k^2 dk \quad (1.25)$$

is the mass variance of the smoothed density field and $\tilde{W}(\mathbf{k}R)$ is the Fourier transform of the window function. The problem of this approach is that it does not take into account the fact that underdense regions might be embedded in larger overdense regions, and end up being in collapsed objects anyway. For this reason, a factor of 2 needs to be artificially added to the definition of the fraction of collapsed objects with mass greater than M , and we write $F(> M) = 2P[> \delta_c(t)]$. The number density of collapsed objects with mass within the range $[M, M + dM]$ corresponds then to

$$n(M, t)dM = \frac{\bar{\rho}}{M} \frac{\partial F(> M)}{\partial M} dM \quad (1.26)$$

$$= \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M^2} \frac{\delta_c}{\sigma} \exp\left(-\frac{\delta_c^2}{2\sigma^2}\right) \left| \frac{d \ln \sigma}{d \ln M} \right| dM, \quad (1.27)$$

known as the Press-Schechter halo mass function (HMF). Finally, if we introduce the variable $\nu = \delta_c(t)/\sigma(M)$ and define $f_{\text{ps}}(\nu) = (2/\pi)\nu \exp(-\nu^2/2)$, we can rewrite Eq. 1.27 as

$$n(M, t)dM = \frac{\bar{\rho}}{M^2} f_{\text{ps}}(\nu) \left| \frac{d \ln \nu}{d \ln M} \right| dM. \quad (1.28)$$

Note that Bond et al. (1991) derived the same halo mass function through a more rigorous treatment, called the excursion set formalism, without the need of adding “by hand” the artificial factor of 2.

Numerical simulations, which directly solve the equations of motion of the dark matter particles, have been used as a probe of the accuracy of the halo abundances predicted by the Press-Schechter formalism. As their resolution was gradually improving, these simulations started showing that the Press-Schechter HMF based on spherical collapse tends to overpredict the low-mass end and underpredict the high-mass end of mass function (see for instance Governato et al. 1999). To address this issue, Sheth & Tormen (1999) and Sheth et al. (2001) introduced a modified critical density δ_{ec} that is computed by keeping into account the ellipticity and prolateness of the collapsing structure, which are defined by computing the eigenvalues of the initial tidal tensor of the gravitational potential field. The critical overdensity for ellipsoidal collapse is expressed as

$$\delta_{\text{ec}}(\sigma, t) = \delta_c(t) \left\{ 1 + \beta \left[\frac{\sigma^2}{\sigma_*^2(t)} \right]^\gamma \right\}, \quad (1.29)$$

where $\sigma_*(t) \equiv \delta_c(t)$, $\beta \sim 0.47$ and $\gamma \sim 0.615$. Notice that σ^2 is a monotonically decreasing function of the mass, meaning that while the critical overdensity for ellipsoidal collapse is similar to that for spherical

collapse for high masses, it is higher for smaller objects, which are more influenced by tidal fields. This new definition of the critical density induces changes in the expression of the halo mass function, which can be rewritten by substituting $f_{\text{ps}}(\nu)$ in Eq. 1.28 with

$$f_{\text{ec}}(\nu) = A \left(1 + \frac{1}{\tilde{\nu}^{2q}} \right) f_{\text{ps}}(\tilde{\nu}), \quad (1.30)$$

where $\tilde{\nu} = 0.84\nu$, $q = 0.3$ and $A \approx 0.322$.

1.2 The first stars, galaxies, and the epoch of reionization

The first galaxies start forming once the primordial gas within a limited region of the Universe stops expanding with the background, and is then allowed to collapse and form stars. In addition to the gas already present in the collapsing region, gas is also accreted from the intergalactic medium (IGM). A part of it is shock-heated to the halo virial temperature while falling into the potential well (hot-mode accretion), but the majority of the gas is accreted at temperatures lower than virial (cold-mode accretion). The dynamics of gas accretion during galaxy formation has been studied in depth (for a review, see Dayal & Ferrara 2018), and numerical simulations have shown that quasi-spherical hot-mode accretion dominates for halo masses $M_{\text{h}} \gtrsim 10^{11.5-12} M_{\odot}$, while cold-mode filamentary accretion dominates at lower masses (Kereš et al. 2005; Dekel & Birnboim 2006; Ocvirk et al. 2008; van de Voort et al. 2012). At $z \gtrsim 2.5$, though, gas accretion through cold filaments occurs also in high-mass galaxies. The subsequent evolution of the protogalaxy is driven by the local collapse and cooling timescales of this gas (Rees & Ostriker 1977; White & Rees 1978). As the gas sets in the potential well, it starts cooling with a characteristic timescale $\tau_{\text{cool}} \propto T/[2n\Lambda(n, T)]$ which is proportional to the particle number density n , the gas temperature T , and to a specific cooling function $\Lambda(n, T)$. This proportionality implies that the gas in the centre of the potential well, within the so-called *cooling radius*, where the density is higher, cools faster than in the outskirt. If the free-fall time (i.e. the timescale over which the cloud would collapse under its own gravity), which goes with $\tau_{\text{ff}} \propto (G\rho)^{1/2}$ (where ρ is the mass density), is longer than the cooling time, then gravity cannot counterbalance radiation pressure, and gas fragments into smaller clouds.

As specified in Barkana & Loeb (2001), in metal-free environments cooling is achieved thanks to molecular Hydrogen H_2 , which is formed

thanks to the relic fraction of free electrons left after recombination via the formation channel

Nevertheless, the abundance of free electrons available is too low to produce all the H_2 needed to explain galaxy formation. For this, we need the gravitational action of dark matter. In fact, dark matter is supposedly non-interactive, hence not affected by thermal pressure, and forms the first structures much earlier than baryons. The clustering dark matter exerts a dragging force on the gas, increasing its density, and favouring a secondary H_2 formation channel, whose efficiency is proportional to the density of the gas:

The subsequent increase in the H_2 number density makes the cooling process more efficient, which in turn helps the collapse of the gas and starts a chain reaction. In this situation, the protostellar clouds cool down and can form fragments whose size eventually reaches the Jeans radius, the maximum scale which can ensure a balance between the two forces (Barkana & Loeb 2001). This fragmentation process comes to an end when $\tau_{\text{cool}} \geq \tau_{\text{ff}}$, or in other words when the cooling radiative process becomes inefficient, and gravity prevails. At this point, star formation activity can start. In a primordial metal-free environment, where the only cooling mechanism is provided by inefficient H_2 transitions, the fragmentation process stops much earlier than if the collapsing gas had already been enriched with metals. Numerical simulations tell us that the typical mass of these very first Population-III (Pop-III) stars, with a metallicity of $Z_{\text{crit}} \leq 10^{-5} - 10^{-6} Z_{\odot}$, could vary between $10^2 M_{\odot}$ and $10^4 M_{\odot}$, though it could be even higher in isolated environments (O’Shea & Norman 2007). For this reason, the nuclear reactions occurring in the core of these stars and providing the energy needed for local thermodynamical equilibrium (LTE) follow the well-known pp-chain (proton-proton), which is relatively inefficient. This means that the fuel has to burn at a much higher rate with respect to the stars where the CNO cycle can happen, which dramatically shortens their lifetime (supposed to be of the order of $\sim 10^6$ yr) and makes their spectra much hotter and harder. The emission of these stars peaks in the UV band, with a huge amount of ionizing photons released, approximately one order of magnitudes higher than for Pop-II stars with metallicity $Z > Z_{\text{crit}}$. Doubts on the nature of Pop-III stars still remain, as their presumably short lifetime makes them extremely hard to observe (though see also Windhorst et al. 2018), and their exact formation

mechanism and final fate are still unknown (Bromm & Larson 2004; Ciardi & Ferrara 2005; Wise 2012; Klessen 2019).

1.2.1 The properties of high-redshift galaxies

The study and understanding of the properties of high-redshift galaxies have recently seen impressive progress, thanks to the surveys performed with instruments like the Hubble Space Telescope (HST), the Subaru Telescope, the Very Large Telescope (VLT), Spitzer and Alma. At $z \gtrsim 4$ galaxies are primarily detected through the Lyman-break feature of their spectrum (see for instance Castellano et al. 2010; Bouwens et al. 2011; McLure et al. 2013; Atek et al. 2015a; Oesch et al. 2016; Livermore et al. 2017), emerging from the absorption of Lyman-continuum photons with (rest-frame) wavelength $\lambda < 912\text{\AA}$ by the neutral Hydrogen in the interstellar medium (ISM) along the line of sight (Meier 1976). Wide-field narrow-band techniques are instead used to identify Lyman-alpha emitters up to $z \sim 8$ (Partridge & Peebles 1967; Kashikawa et al. 2011; Matthee et al. 2015; Ouchi et al. 2018), by detecting the Hydrogen recombination line at $\lambda = 1216\text{\AA}$. Its flux can then be used to estimate the star formation rate of the galaxy, through its correlation with the $H\alpha$ emission line (Kennicutt 1998). Nevertheless, given the extremely short mean free path of $\text{Ly-}\alpha$ emission when neutral Hydrogen is present (Madau & Rees 2000), this method becomes ineffective at higher redshift, as the Universe is less and less ionised. The refinement of these observational techniques, and the growth of the high-redshift samples of detected galaxies, provided theorists with an impressive amount of data, which allows for the robust computation of statistical observables against which to tune theoretical cosmological models. We will briefly describe here some of the main diagnostics used at $z > 4$, namely the UV luminosity function, the UV continuum slope, the stellar mass density, the star formation rate density, and the stellar mass function.

The high-redshift UV luminosity function (UV LF) is now the most robust probe of galaxy evolution that we have at $z > 4$. It measures the number of galaxies per unit volume as a function of luminosity, hence conveying precise information about the sheer number of UV photons produced by the galaxies, which is used not only to set constraints on the sources of reionization but also to study their star formation activity. Additionally, the precise shape of the UV LF, when compared with that of the halo mass function, gives us an idea of the star formation efficiency of the different galaxies as a function of total luminosity, and it is determined by the different physical processes able to impact the cooling timescales

of the gas. At faint luminosities, SN feedback in low-mass galaxies are thought to quench star formation by heating the gas of the disk and ejecting it from their shallow potential wells (e.g. Dekel & Silk 1986). The UV LF is closest to the HMF at its knee, indicating that those galaxies have a high star formation efficiency. At the bright end, again, studies indicate that the low- z UV LF is suppressed, possibly due to heating from AGN (Croton et al. 2006; Bower et al. 2006), inefficient cooling in high-mass halos (Binney 1977; Benson et al. 2003) and dust attenuation (Martin et al. 2005). Similarly to the halo mass function, the UV LF can be described by the three (degenerate) parameters of a Schechter function, the normalization ϕ_* , the knee luminosity L_* and the faint-end slope α , and it can be written as

$$\frac{dn}{dL} = \phi(L) = \frac{\phi_*}{L_*} \left(\frac{L}{L_*} \right)^\alpha \exp^{-(L/L_*)}. \quad (1.31)$$

The degeneracy between these parameters means that it is not easy to assess whether the evolution of the galaxy UV LF is driven by number density (ϕ_*) evolution or luminosity (L_*) evolution (Dayal et al. 2013), and studies tend to report contrasting results depending on the luminosity range considered (we will address this issue in more details in Chapter 4). Observationally, the faint-end slope of the UV LF is by now pretty well determined up to $M_{UV} \sim -16$ and $z \sim 10$, thanks to studies of the Hubble Ultra Deep Field (Ellis et al. 2013; Schenker et al. 2013; McLure et al. 2013; Finkelstein et al. 2015; Bouwens et al. 2015b) and the Hubble Frontier Fields (Atek et al. 2015b; Ishigaki et al. 2015; McLeod et al. 2016; Livermore et al. 2017; Ishigaki et al. 2018). The faintest magnitudes have been reached also thanks to the exploitation of gravitational lensing methods (Postman et al. 2012; Atek et al. 2015b; Kawamata et al. 2016; Bouwens et al. 2017; Livermore et al. 2017), and the consistent result is a steepening of the faint-end slope α with increasing redshift (see for instance Stark 2016; Ishigaki et al. 2018, and references therein), revealing an increasingly dominating population of low-luminosity galaxies. After all, this is to be expected, given the steepening of the low-mass end of the theoretical halo mass function with increasing redshift. On the other hand, the bright end is still subject to different uncertainty sources, contaminated as it is by AGN, dust extinction and lensed galaxies above all (Ono et al. 2018). Addressing the AGN contribution to the total UV LF is one of the goals of Chapter 4. For this thesis, the galaxy UV luminosity is modelled using the code STARBURST99 (Leitherer et al. 1999, 2010).

Linked to the galaxy UV luminosity, is the UV spectral slope (Calzetti et al. 1994), generally referred to as β , which is found to correlate pretty

well with dust extinction (Meurer et al. 1997; Reddy et al. 2012). For this reason, β has been used to probe the evolution of the dust content, even though at high redshift continuum measurements are not available and one has to rely on single-colour estimations (Bouwens et al. 2009) or fitting techniques (Finkelstein et al. 2012). It has thus been found that the dust content decreases with decreasing luminosity at a given redshift, which also implies less dusty galaxies at higher redshift, results which have also been confirmed by simulations (Finlator et al. 2011b).

The cosmic star formation rate density (SFRD) is another important volume-averaged observable for high-redshift galaxies (Bouwens et al. 2011; McLure et al. 2013; Oesch et al. 2013). The integration of the UV LF to a limiting magnitude, often chosen to be $M_{UV} = -17.7$ in the Hubble Ultra Deep Fields, yields the UV luminosity density (UVLD), which can then be converted to SFRD through the conversion factor defined in Madau et al. (1998). Observational works have found that the slope of the SFRD is positive for $0 < z < 2$, reaches its peak at $z \sim 2 - 3$ and is negative afterwards, steepening with increasing redshift (Madau & Dickinson 2014). It is important though to point out that inferring the galactic SFR from its UV luminosity needs assumptions for the IMF, the metallicity, and the contribution of binary stars.

Integrating the star formation density over redshift, one obtains the stellar mass density (SMD), defined as the average stellar mass of the Universe per unit volume as a function of redshift. Its measurement, though, is inextricably linked to the uncertainties arising from trying to estimate stellar masses of high-redshift galaxies from broad-band data. SED modelling techniques in fact depend on the assumptions we make about IMF, metallicity of the stars, dust attenuation, nebular emission lines, and age of the stellar population. For this reason, SMD measurements are often accompanied by uncertainties of the order of $0.3 - 0.5$ dex, increasing with redshift, and yield heterogeneous results (González et al. 2010; Labbé et al. 2013; Stark et al. 2013; Duncan et al. 2014; Oesch et al. 2014; Grazian et al. 2015; Song et al. 2016). Nevertheless, it seems reasonable to think that galaxies brighter than $M_{UV} \sim -18$ contain half of the total stellar mass of the Universe at $z = 5$, and this fraction decreases with redshift (Dayal & Ferrara 2018), with low-mass galaxies playing a more important role.

Finally, we consider the stellar mass function (SMF), the distribution of galaxies as a function of mass per unit volume. Clearly, the SMF suffers from the same limitations as the SMD, but it has been shown that the stellar mass correlates, at least at lower redshift, with virtually all galactic properties, such as the star formation rate, metallicity and size (Tremonti

et al. 2004; Noeske et al. 2007; Williams et al. 2010). Testing how such correlations change at higher redshift would offer valuable insights into galaxy evolution. Additionally, abundance matching techniques lining the SMF to the HMF offer a way to compute the baryon-to-dark matter ratio of the galaxies, hence giving us the possibility to study the impact of AGN and SN feedback as a function of stellar mass (Lu et al. 2014). Hybrid observational methods, combining multi-band data, collected for instance with Spitzer and HST, stacking techniques and test against theoretical models, have the power to set tighter constraints especially on the low-mass end of the SMF. In such cases, the slope of the low-mass end has been found to increase with redshift, similarly to what happens with the UV LF, while no significant evolution of the knee mass has been found (Song et al. 2016).

We point out that the high-redshift versions of all of these observables, except for the UV spectral slope, are used in this thesis to tune our model, and are successfully reproduced.

1.2.2 Epoch of reionization

The period of time between the photon decoupling and the formation of the very first collapsed structures is known as the Dark Ages, which ended around $z = 20 - 30$, when cooling processes (especially from molecular hydrogen vibrational and rotational radiative emissions) became efficient enough for the first stars to form in minihalos with $M_{\text{h}} \approx 10^6 M_{\odot}$ (Haiman et al. 1996; Tegmark et al. 1997; Barkana & Loeb 2001). The formation of stars in the first collapsed dark matter halos is accompanied by the production of UV and X-ray photons with energy $E_{\gamma} > 13.6 \text{ eV}$, able to ionise the neutral Hydrogen (HI) present in the intergalactic medium. This process goes under the name of cosmic reionization, is thought to happen between $z \sim 13$ and $z \sim 6$, and it represents the last phase transition in the history of the Universe (for a review, see Barkana & Loeb 2001; Dayal & Ferrara 2018). The details of how reionization happened depend on many different factors: the physical properties of the ionizing sources, like the luminosity, the spectrum, the fraction of ionizing photons which can escape from the galaxy (escape fraction, f_{esc}), the statistical properties of each of the ionizing population (number density, clustering and spatial distribution), and finally also the distribution profile of the gas density field. All these properties affect the initial redshift of reionization, its topology (intended as the size distribution and the evolution of the HII regions) and duration. Beside ionizing the Hydrogen present in the Universe, energetic photons contribute to the overall UV background radiation and directly

affect galaxy evolution, by photo-evaporating gas in low-mass halos with shallow potential wells sitting in ionised regions (Barkana & Loeb 1999; Shapiro et al. 2004). Nevertheless, the picture is complicated by the patchiness and the temporal extent of the reionization process, and one should model the effect of its feedback on every single halo according to its redshift formation, the local strength of the UV background (UVB) and the epoch at which the UVB starts irradiating the halo (Gnedin 2000). This type of feedback impacts the star formation rate and the ionising emissivity of smaller galaxies (Wu et al. 2019), creating a feedback loop that is not trivial to model. For this reason, cosmological hydrodynamic simulations coupled with radiative transfer are necessary to study the effect of an inhomogeneous and evolving UV background on structure formation and evolution (Finlator et al. 2011a; Hasegawa & Semelin 2013; Gnedin & Kaurov 2014; Pawlik et al. 2015; Ocvirk et al. 2016; Katz et al. 2019; Wu et al. 2019; Hutter et al. 2020a,b). In particular, common findings point towards suppression of star formation in halos with masses $M_h \lesssim 10^9 M_\odot$, with galaxies residing in higher-density regions showing sharp reionization-induced quenching at earlier epochs. In any case, stellar feedback seems to play a dominant role in regulating the star formation activity (e.g. Rosdahl et al. 2018), and the impact of reionization feedback on galactic observables such as the faint end of the UV luminosity function (UVLF) and the star formation rate density (SFRD) is minimal (Gnedin & Kaurov 2014; Finlator et al. 2018). Still, the degeneracies intrinsic to all of these models (mainly between the escape fraction and the ionizing emissivity, which in turn depends on the assumed initial mass function and stellar population synthesis models) make it hard to converge towards a single consistent physical model of reionization. In order to break this degeneracy, tighter observational constraints are required, possibly on the escape fraction of ionising photons or on the 21-cm signal power spectrum at high redshift, which would provide us with information about the physical properties of high- z galaxies (Kim et al. 2013; Seiler et al. 2019).

Indeed, a lot of observational effort has gone into estimating the escape fraction of ionizing photons from high-redshift sources as a function of halo mass and redshift. Nevertheless, the attenuation and suppression of the Lyman continuum part of the spectrum (formed by ionizing photons with wavelength $\lambda < 912\text{\AA}$) due to ISM absorption and interlopers along the line of sight pose a serious complication to the observation of Lyman continuum emitters (LCEs) at high redshift. Studies have tried to circumvent this problem by studying low-redshift analogues of high-redshift LCEs (e.g. Leitherer et al. 2016; Izotov et al. 2018), finding that the value of f_{esc} is higher in correspondence of low-density channels dug in the ISM by

SN feedback or turbulence (Kakiichi & Gronke 2019), and hence it is strongly dependent on the line of sight under which the source is observed (Gazagnes et al. 2020). Also for this reason, the many attempts to measure f_{esc} at different redshift (Inoue et al. 2006; Nestor et al. 2011; Boutsia et al. 2011; Mostardi et al. 2015; Shapley et al. 2016; Grazian et al. 2016; Matthee et al. 2017) have not managed to find an agreement yet, with average values ranging from $\sim 1\%$ to $\sim 20\%$ and seemingly increasing with redshift. Still, outliers are found at all redshift, with values up to $\gtrsim 50\%$ (Vanzella et al. 2016, 2018). Besides, these results depend on the assumed stellar population models, the inclusion of photons from binary stars and the underlying dust attenuation model, and hence they come with all the related uncertainties.

Other observational probes can be used to constrain reionization models. Thanks to the high cross-section for the Lyman α absorption of HI atoms, the amount of neutral Hydrogen left in the Universe at each redshift can be tracked by the so-called Lyman α forest. When the light emitted by distant quasars encounters a HI region along its line of sight, absorption features bluewards of the Lyman α wavelength at $1216\text{\AA}$ appear in the quasar's spectrum (Gunn & Peterson 1965; Scheuer 1965). Nevertheless, this technique is very sensitive even to low amounts of HI, and it saturates at $z \gtrsim 6$, probing only the latest stages of reionization. A complementary observational approach consists in studying the 21-cm emission (or absorption, depending on the quantum state of neutral Hydrogen. For a review about the 21-cm physics, see Zaroubi et al. 2012) line, which is produced by the spin-flip transition of the Hydrogen atom and tracks the ionization state of the IGM (Hogan & Rees 1979; Madau et al. 1997; Furlanetto et al. 2006; Zaroubi 2013). In principle, with this technique, one could follow the ionization state of the Universe from the earliest stages of reionization all the way to the end, and a lot of effort has been put into trying to observe the 21-cm emission line as a function of redshift. Radio interferometers such as LOFAR (van Haarlem et al. 2013; Patil et al. 2017), GMRT (Paciga et al. 2011), MWA (Tingay et al. 2013; Bowman et al. 2013) and PAPER (Parsons et al. 2010) have set upper limits on the statistical detection of the 21-cm fluctuations, and next-generation facilities such as HERA (DeBoer et al. 2017) and SKA (Mellema et al. 2013; Koopmans et al. 2015) will help setting even tighter constraints, but no clear detection has been observed so far. Finally, also the optical depth of the Compton scattering of CMB photons on free electrons is an indicator of the ionization state of the Universe (Peebles & Yu 1970; Sunyaev & Zeldovich 1972).

1.2.2.1 The sources of reionization

The current picture maintains that the starting kick to reionization was given by Pop-III stars. Their hard spectra make these stars perfect candidates for reionization, but as soon as the very first generation exploded, the interstellar medium (ISM) is driven above the critical metallicity, preventing a second generation of metal-free stars from forming locally. For this reason, they are supposed to play a role only during the early phases of reionization, but the reality is that we have not obtained a full grip on the actual contributors to reionization yet. Several options are still viable, and it is likely that each of them played a role to some extent. We briefly summarise the state-of-the-art knowledge about what were the suppliers of ionizing photons.

Pop-II stars in low-mass galaxies: low-mass galaxies have been proved to be extremely common at all epochs (Smit et al. 2014; Bouwens et al. 2015b; Smit et al. 2018; Atek et al. 2018; Ouchi et al. 2018; De Barros et al. 2019; Maseda et al. 2020). Depending on their initial mass function and on their escape fraction, they are supposed to provide for the majority of ionizing photons. Pop-II stars, in particular, have higher metallicity and hence softer spectra than Pop-III stars but are much more common (Ciardi & Ferrara 2005). Supernovae explosions, being very energetic events, are also a source of ionizing photons, though their small number makes them unlikely to be an important contributor, even if there are still huge uncertainties on the primordial initial mass function (Johnson & Khochfar 2011). In any case, many theoretical results converge on galaxies in low-mass halos with $M_h \lesssim 10^{9.5} M_\odot$ contributing to the majority of LyC photons at $z > 7$ (Choudhury & Ferrara 2007; Salvaterra et al. 2011; Wise et al. 2014; Dayal et al. 2017a).

Quasars: the contribution of high-redshift quasars, characterised by higher luminosity and escape fraction but also lower number density than regular galaxies, has been discussed extensively in the literature, with some works observing very high AGN number densities (Giallongo et al. 2015), while other surveys found opposite results (Weigel et al. 2015; McGreer et al. 2018). For this reason, the literature has shown contrasting results, with some works suggesting a significant contribution of AGN to reionization (e.g. Mitra et al. 2018; Grazian et al. 2018; Finkelstein et al. 2019), while others found it to be largely subdominant (Onoue et al. 2017; Qin et al. 2017; Hassan et al. 2018; Trebitsch et al. 2018, 2020). Dayal et al. (2020), using the *Delphi* semi-analytic model (SAM) presented in this thesis, found

that AGN account for about 10 – 25% of the total amount of ionising photons at $z \geq 4$. Thanks to their hard spectra, though, they play a much more fundamental role reionizing HeII, which has a binding energy of 54.4 eV and happened at later times (Furlanetto & Oh 2008).

Mini-quasars: powered by intermediate-mass black holes with masses $M_{\text{bh}} = 10^2 - 10^5 M_{\odot}$, they could have provided a significant amount of ionizing photons at the very beginning of reionization at $z \sim 15$ (Madau et al. 2004). The upper limits on their number density can be estimated with the unresolved portion of the cosmic soft X-ray background that we observe today, corresponding to $\approx 8\%$, (Dijkstra et al. 2004; Zaroubi et al. 2007). Some studies have shown they could contribute to up to 20% of the ionizing photons, but better measurements of the reionization state of Helium, as well as fewer uncertainties on the spectral energy distribution and the duty cycle of the mini-quasars, are required to set more precise constraints (Wang et al. 2010; Hao et al. 2015).

Decaying particles and neutrinos: the (supposed) decay of dark matter particles or neutrinos could have played a role, releasing ionizing photons, even though, if present, they should have produced fluctuations in the CMB power spectrum that we do not see. This lack of observational support sets an upper limit on their contribution to reionization of $\sim 10\%$ (Chen & Kamionkowski 2004). This mechanism would help in explaining the high optical depth of Thomson scattering measured by WMAP during the Dark Ages, when star formation should not have started yet (Kogut et al. 2003).

1.3 Black holes

In 1784 John Michell published a paper in the *Philosophical Transactions of the Royal Society of London*, in which he first suggested the existence of “dark stars”. Basing himself purely on Newtonian gravity, and assuming the corpuscular nature of light, he imagined a body whose escape velocity would be higher than the speed of light. Such an object would need to have a radius

$$R \leq \frac{2GM}{c^2}, \quad (1.32)$$

M being its mass and c the speed of light. This early hypothesis was confirmed in 1916 when Karl Schwarzschild found the first exact vacuum solution to Einstein’s field equations of general relativity in the case of

a spherical, electrically neutral and non-rotating mass. The external gravitational field generated by this mass element is described in fact by the line element of the Schwarzschild metric, which in spherical coordinates (t, r, θ, ϕ) reads

$$ds^2 = - \left(1 - \frac{2GM}{c^2 r}\right) c^2 dt^2 + \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2). \quad (1.33)$$

This expression has an isolated singularity with infinite curvature in $R = 0$, while $R = R_{\text{schw}} = 2GM/c^2$ is a singularity only for this coordinate system. The latter is known as the Schwarzschild radius, and it represents the boundary beyond which particles cannot escape the gravitational pull of the black hole, dividing the space-time into two disconnected regions.

Physically speaking, black holes are fully characterised by their mass, charge and angular momentum (Israel 1967). Leaving aside the electric charge, the most complete description for a rotating black hole is given by the Kerr metric, according to which the event horizon becomes dependent on the black hole spin $a \equiv c\mathcal{L}/(GM^2)$ (where $\mathcal{L}$ is the angular momentum), and reads

$$R_{\text{kerr}} = \frac{R_{\text{schw}}}{2} \left(1 + \sqrt{1 - a^2}\right). \quad (1.34)$$

Given the same mass, the fastest a black hole is spinning, the smallest is its Kerr radius.

1.3.1 Black hole seeds

The cornerstone of this thesis is the study of astrophysical high-redshift supermassive black holes inhabiting the centres of massive galaxies (for a recent review, see Inayoshi et al. 2020). Their origin, in fact, remains an open question. The initial seed mass and the properties of the host both affect the subsequent growth and hence the luminosity of the black hole, but degeneracies between the seed mass, the formation redshift and the duty cycle make it challenging to constrain the growth model. So far, three different non-exclusive formation scenarios have been put forward (see also Volonteri 2012; Haiman 2013; Latif & Ferrara 2016a), we will briefly review them here.

Stellar black hole seeds are formed from the collapse of the first generation of high-mass, extremely metal-poor Pop-III stars (Madau & Rees 2001a)

with masses of a few hundreds of solar masses (Bromm & Larson 2004). They represent the most natural progenitors of high-redshift quasars, but the uncertainties on the upper mass limit for Pop-III stars (Bromm 2013) translate into uncertainties on the mass of the remnant black holes. Indeed, the initial seed mass, the time-averaged accretion rate and the contribution of black hole mergers to the growth (hence the initial halo occupation fraction) are still open questions, but even with a simple accretion model it is easy to see that this is a viable black hole formation scenario. A seed with an initial mass $M_{\text{seed}} = 100 M_{\odot}$, growing at the Eddington rate with a radiative efficiency of 10%, would take $\sim 7 \times 10^8$ yr (roughly equivalent to the age of the Universe at $z = 6$) to grow to the size of $3 \times 10^9 M_{\odot}$ (Haiman 2013). Nevertheless, the progressive discovery of supermassive black holes at $z \gtrsim 7$ (see e.g. Bañados et al. 2018) is setting more and more constraints on the minimum initial seed mass or the duty cycle, and studies have started to consider also phases of super-Eddington accretion (Mayer 2019).

Direct-collapse black hole seeds are supposedly formed from the monolithic isothermal collapse of a cloud of pristine gas (originally proposed in Rees 1984). However, for a gas cloud to collapse in that manner, a specific set of conditions must be met. First of all, it requires a high inflow rate of gas, of the order of $0.1 M_{\odot} \text{ yr}^{-1}$, occurring uninterrupted for at least 1 Myr. The gas needs to efficiently shed angular momentum and coagulate in the centre of the potential well, and for this reason, low-mass low-spin high-redshift halos are ideal host candidates. Numerical simulations of the gas dynamics have shown that angular momentum can be dissipated through non-axisymmetric gravitational torques (the so-called *bars-within-bars* mechanism, Choi et al. 2015). In addition, the gas should be metal-free and deprived of molecular Hydrogen, in order to avoid cooling, fragmentation and star formation. An external Lyman-Werner (UV) radiation field, possibly from a nearby star-forming galaxy, is required to photo-dissociate H_2 (Latif et al. 2013b; Habouzit et al. 2016b). Given these conditions, the gas cloud can collapse in a self-similar way, cooling through Lyman-alpha emissions, and numerical simulations have shown that a black hole seed of $\sim 10^5 M_{\odot}$ can form at the centre of the halo, delaying the onset of star formation (Latif et al. 2013c; Dijkstra et al. 2014; Shlosman et al. 2016). In particular, using the *FiBY* (First Billion Year) project, Agarwal et al. (2014) found that DCBH seeds of $M_{\text{bh}} \sim 10^4 - 10^6 M_{\odot}$ can form in small satellite galaxies which have not formed stars yet and are irradiated by a strong LW background from bigger close-by galaxies. Clearly, given their high initial mass, direct-collapse black

holes provide an easier explanation for the formation of supermassive black holes at $z > 7$, but, despite candidates have been proposed (Hartwig et al. 2016), no clear observational counterpart has been detected yet.

Black hole seeds in stellar clusters provide another possible explanation for the formation of high-redshift supermassive black holes (for a review, see Katz 2019). In fact, the first stars of the Universe are expected to form in dense clusters, especially if traces of metal have already contaminated the pristine gas (Klessen 2019). The most massive stars of the cluster sink to the centre through dynamical friction, increasing the cross-section for mergers, and the subsequent repeated stellar collisions have the potential to form - thanks to a runaway process - a very massive star (e.g. Portegies Zwart et al. 1999; Katz et al. 2015), which can then collapse into an intermediate-mass black hole, depending to its mass and metallicity (Heger et al. 2003). Unlike the other two scenarios, the conditions necessary for this seed-forming mechanism to occur (high-density star clusters) have been observed in the local Universe (Espinoza et al. 2009). Nevertheless, the viability of this runaway collapse is very dependable on the initial central mass and density of the cluster (Katz et al. 2015), so it is still unclear whether this process can dominate the formation of supermassive black hole seeds.

1.3.2 Accretion and feedback

Black holes are thought to be surrounded by an accretion disk of material whirling around the central body (Shakura & Sunyaev 1973), which gradually loses gravitational potential energy by emitting gravitational waves. All the material within a radius of $r = 3R_{\text{schw}}$ is eventually accreted by the black hole. A percentage of the accreted mass corresponding to the black hole radiative efficiency ϵ_r is turned into energy which is then re-injected into the surrounding. This process stands at the root of what we call black hole feedback (Fabian 2012). In particular the energy emitted by the accreting black hole is $L = \epsilon_r \dot{M}_{\text{acc}} c^2$, where L is the luminosity emitted by the accreting black hole and $\dot{M}_{\text{acc}}$ is the mass accretion rate. For simplicity, it is commonly assumed $\epsilon_r = 0.1$, even though Novikov & Thorne (1973) first showed the dependence of the radiative efficiency on the black hole spin. Equating the inwards gravitational pull on a test particle and the outwards radiative force, Eddington (1916) derived a theoretical limit on the mass accretion rate and its corresponding luminosity, given by

$$L_{\text{edd}} = \frac{4\pi G M m_p c}{\sigma_T}, \quad (1.35)$$

where m_p is the proton mass and σ_T is the Thomson cross-section. In this thesis we adopt the Eddington mass accretion rate $\dot{M}_{\text{edd}} = L_{\text{edd}}/(\epsilon_r c^2)$ as the maximum accretion rate, but it is important to stress that this holds true when accretion takes place in conditions of spherical symmetry.

Bondi & Hoyle (1944) studied the physics of the accretion in the context of point masses and spherical symmetry but taking into account also the physical properties of the infalling material, finding that all the matter which finds itself gravitationally bound to the black hole is eventually accreted at a rate of

$$\dot{M}_{\text{bh}} = \frac{4\pi G^2 M_{\text{bh}}^2 \rho}{(c_s^2 + v_{\text{rel}}^2)^{3/2}}, \quad (1.36)$$

where c_s is the speed of sound, ρ is the gas density at the Bondi radius $r_B = GM_{\text{bh}}/c_s^2$, and v_{rel} is the gas velocity relative to the black hole. It is common practice in theoretical cosmological models to assume a Bondi-Hoyle accretion rate limited by the Eddington rate. It should be considered, though, that works have shown that super-Eddington accretion can take place given a specific set of conditions (for a review, see Mayer 2019). Such conditions involve either an asymmetric distribution of the medium, like the presence of a disk or supersonic turbulence (Novak et al. 2012; Van Borm et al. 2014), or radiation trapping within the gas flow, which diminishes the radiation pressure onto the infalling material, allowing for higher accretion rates (Begelman 1979; Jiang et al. 2014). High-redshift galaxies, where the dense gas reservoirs of the nuclei are constantly replenished by large-scale gas accretion and the optical depth is high, are more likely to meet these conditions. In particular, it has been argued that super-Eddington accretion is needed to grow small seeds with $M \lesssim 100 M_\odot$ to masses of $\sim 10^9 - 10^{10} M_\odot$ by $z = 6$, but we will address better this issue in Chapter 3.

Historically, black hole feedback has been distinguished between two different types, radiative mode (or quasar mode) for accretion rates $\dot{M}_{\text{bh}} \lesssim 0.02 \dot{M}_{\text{edd}}$ and mechanical mode (or radio mode) for higher accretion rates (Croton et al. 2006; McNamara & Nulsen 2007). Radiative feedback, more effective when the surrounding medium is optically thick, photo-heats and ionises the surrounding gas, decreasing the black hole accretion rate (Park & Ricotti 2012; Jeon et al. 2012) and possibly stopping the inflow of gas with low angular momentum towards the black hole (Pacucci et al. 2015; Inayoshi et al. 2016). The mechanical feedback, instead, takes the form of bipolar relativistic jets (e.g. Blandford & Znajek 1977; Begelman et al. 1984; McNamara et al. 2005; Sadowski et al. 2013),

which can expand through the ISM, shock-heating the gas and driving gas outflows from the host galaxy (for a review, see Wise 2019).

1.3.3 Observational methods for high-redshift black holes

In principle, observations in the X-ray band, with instruments such as *Chandra*, *XMM-Newton* and *NuSTAR*, are best at detecting AGN activity. In fact, inverse Compton scattering of thermal photons onto the accretion disk produces a power-law X-ray emission which is observed practically in all AGN (Sunyaev & Truemper 1979; Fabian 2006; Gilfanov & Merloni 2014; Brandt & Alexander 2015). At the same time, also the expanding jets, characteristic of the AGN radio mode, show an X-ray counterpart resulting from the interaction between the accelerated plasma and the ISM (Worrall 2009). Another advantage of the X-ray band is that its photons are less susceptible to be absorbed, and they represent a less biased tracker of the intrinsic emission than the optical or UV emissions. Finally, the X-ray luminosity, and in particular the X-ray-to-optical luminosity ratio, is a strong diagnostic tool to distinguish AGN from normal galaxies, especially at $1 < z < 5$. Nevertheless, X-ray surveys have been able to probe the quasar population only up to $z = 5$ (e.g. Ueda et al. 2014), and only the *ATHENA* mission will be able look at the X-ray emission of the black hole population at $6 \leq z \leq 10$ (Nandra et al. 2013).

For this reason, the compilation of catalogues of powerful quasars at $z > 6$ has been performed especially thanks to the effort of optical, UV and infrared surveys (for a more extensive review, see Schleicher 2019). In particular, the *Sloan Digital Sky Survey* (SDSS; York et al. 2000) has played a pivotal role in this effort, with a final sample of 52 quasars at $5.7 < z < 6.4$ and with $-29 \leq M_{1450} \leq -24.5$, where M_{1450} is the UV magnitude at $\lambda = 1450\text{\AA}$ (Jiang et al. 2016). Thanks to this sample the bright-end slope of the quasar UV luminosity function has been constrained to be $\beta = -2.8 \pm 0.2$, while the evolution of their number density can be parametrised as $\rho(M_{1450}) = \rho_{(z=6)} \times 10^{\kappa(z-6)}$, with $\kappa = -0.72 \pm 0.11$ for $5 \leq z \leq 6$, hence showing a rapid drop. In addition, the Lyman-alpha forest of the observed spectra have helped in putting constraints on the end of reionization (Fan et al. 2006). Another optical survey, the *Canada-France High- z Quasar Survey* (CFHQS) contributed with a total of 19 quasar detections (Willott et al. 2007, 2009, 2010a), extending the quasar UV luminosity function to fainter magnitudes up to $M_{1450} \sim -22$ and suggesting a double power-law shape with a break at $M_{1450} \sim -25$. To detect quasars at even higher redshift, we need to move from optical to (near-)infrared surveys. Indeed, finding quasars at $z \sim 7$ has been one of

the scientific goals of the *UKIRT (UK infrared telescope) Infrared Deep Sky Survey* (UKIDSS). This effort led to the discovery of some of the highest-redshift supermassive black holes found in the literature, such as ULAS J131911.29+064124.3, with $M_{\text{bh}} \sim 2 \times 10^9 M_{\odot}$ at $z = 0.85$ (Mortlock et al. 2011), and ULAS J134208.10+092838.61, with $M_{\text{bh}} \sim 8 \times 10^8 M_{\odot}$ at $z = 5.4$ (Bañados et al. 2018). In particular, the presence of these giants only 700 million years after the Big Bang, which requires an average accretion rate of $\approx 2.6 M_{\odot} \text{ yr}^{-1}$, raised questions about our understanding of SMBH formation, specifically about how they could have grown so much in such a short time. We will better address this issue in Chapters 3 and 4. Other relevant surveys include the *Pan-STARRS* survey at $z > 5.4$ (Venemans et al. 2015a; Bañados et al. 2016; Mazzucchelli et al. 2017), the *Subaru High- z Exploration of Low-Luminosity Quasars* (SHELLQs), designed to detect quasars up to $z = 7.4$ and to extend the faint end of their UV luminosity function to $M_{1450} \sim -22$ (Matsuoka et al. 2016, 2018), the *Visible and Infrared Survey Telescope for Astronomy* (VISTA) survey (Venemans et al. 2013, 2015b), and the *Dark Energy Survey* (DES; Reed et al. 2015, 2017).

Aside from the direct detection of resolved high-redshift quasars, observational constraints can also be set by studying the unresolved X-ray background (Dijkstra et al. 2004; Salvaterra et al. 2005). In particular, the strong emission of quasars in the hard X-ray band at $z > 6$ would be detected, redshifted, as soft X-ray photons, with energies of 0.5 - 2 KeV. With the help of X-ray surveys such as Chandra, ROSAT and XMM-Newton, it is possible to measure the total soft X-ray background emission, out of which the unresolved portion is obtained by subtracting the contribution from resolved sources and from extended sources, such as galaxy groups and clusters. What is left, estimated to be $\sim 8\%$ (Cappelluti et al. 2017), is attributed to high-redshift quasar activity, and sets constraints on the black hole accreted mass density (Salvaterra et al. 2005).

Finally, since the first observational run of the Advanced LIGO detectors (Abbott et al. 2016), gravitational waves have recently become an additional powerful tool not only to test the validity of general relativity as a theory of gravity and the standard model in general, but also to study (high-redshift) black holes. While LIGO is designed to probe stellar black holes with masses of a few dozens of solar masses, the *Laser Interferometer Space Antenna* (LISA) could in principle detect mergers between black holes as massive as $M_{\text{bh}} = 10^4 - 10^6 M_{\odot}$ up to $z \sim 20$ (Dayal & Ferrara 2018). The advantages of gravitational waves is that they travel unimpeded, providing a direct probe of the intrinsic properties (i.e. mass and redshift) of the black holes (for a review, see Colpi & Sesana 2017). Their study could shed light on the presence of intermediate-mass black holes in dwarf

galaxies, on their mass function as a function of redshift, and even on their role in galaxy formation. Measurements of the merger rates at $z > 4$ can tell us how supermassive black holes formed, and how much mergers contributed to their growth. In addition, if compared to the galaxy merging rates, they could also help us in constraining the halo occupation fraction, as well as in gaining a better understanding of the merging time scales of black hole binaries.

1.3.3.1 Observational properties of the high- z quasar population

Just like for star-forming galaxies, the luminosity function is the most studied observational probe of the high-redshift black holes (Willott et al. 2010b; Jiang et al. 2016; Matsuoka et al. 2018; McGreer et al. 2018), pointing towards a strong decrease in the quasar number density with increasing redshift at $z > 4$. Quasars at $z = 6$ are found to be 100 times less abundant than those at $z = 4$, meaning that the normalization parameter ϕ_* of the AGN UV LF decreases with increasing redshift, while its shape remains similar to itself (Inayoshi et al. 2020). Additionally, contrary to what is observed for the galaxy UV LF, the AGN luminosity function flattens at low luminosities (Matsuoka et al. 2018), hence excluding the possibility of reionization being driven by AGN, especially at high redshift. Nevertheless, obscuration effects that might hide a part of the AGN population from optical and IR (not X-ray) surveys must be taken into account, and some studies have suggested that up to 80% of the quasar population with $\log(L_X) \lesssim 44$ at $3 < z < 6$ might be heavily obscured (Vito et al. 2018a). In addition, results consistently show that the obscured fraction does not evolve with redshift at $z > 2$ (e.g. Iwasawa et al. 2012; Ueda et al. 2014; Liu et al. 2017). The hosts of these high-redshift AGN appear to be similar to low-redshift starburst galaxies, with star formation rates of the order of $10^3 \text{ M}_\odot \text{ yr}^{-1}$, metal-enriched nuclei showing broad emission lines, and important amounts of dust on kpc scales (De Rosa et al. 2014; Mazzucchelli et al. 2017; Valiante et al. 2017; Reed et al. 2019; Shen et al. 2019). This suggests that the hosts of these high- z quasars have been assembled rapidly at early epochs. On the other hand, differences have been reported when the $M_{\text{bh}} - M_*$ relation at low and high redshifts are compared: the dynamical masses of high- z AGN hosts are lower than those at low z (Wang et al. 2016; Decarli et al. 2018; Shimasaku & Izumi 2019), though some observational biases make these results inconclusive (Reines & Volonteri 2015; Lupi et al. 2019). In most cases, estimations of black hole masses at high z rely on correlations between the velocity of Doppler-shifted emission lines such as $\text{H}\beta$ and Mg II and the optical/UV

luminosity (for a review of the method, see Peterson 2010). The problem of the evolution of the $M_{\text{bh}} - M_*$ relation has been studied also with numerical simulations (for a summary, see Habouzit et al. 2020), which generally find similar and tight correlations at high redshift. Nevertheless, they show differences at the low-mass end, depending on the strength of SN feedback and on the details of the sub-grid models implemented for black hole growth, but deeper observations will be needed to differentiate between them.

1.3.4 Theoretical methods for black holes in a cosmological context

It has been shown that in order to reproduce the available observations of the black hole and galaxy mass function, theoretical models must include AGN feedback from accreting black holes at $z \sim 4 - 5$ (Bower et al. 2006; Croton et al. 2006; Lagos et al. 2008). Nevertheless, studying the physics of black hole accretion and feedback and their interplay with the host galaxy in a cosmological context has proven to be very challenging, due to the extremely wide spatial and temporal dynamic ranges involved. While the accretion disk of a supermassive black hole with $M_{\text{bh}} \sim 10^8 M_\odot$ is of the order of 10^{-4} pc (Morgan et al. 2010), the typical box size of a cosmological simulation is around $10^7 - 10^9$ Mpc. Similarly, the time scale of black hole accretion and that of galactic evolution differ by $\sim 2 - 3$ orders of magnitude, making the feedback between the two complex to follow and computationally expensive to model. Usually, the problem is alleviated by introducing sub-grid analytical prescriptions to deal with localised physical processes, and their parameters are tuned against a few galaxy and black hole key observables, such as the mass and luminosity functions (e.g. Vogelsberger et al. 2014; Schaye et al. 2015). Tuned models are then ready to produce predictions for the next-generation surveys, which might validate them or not, providing an indication about what can be improved. Given the focus of this thesis, we will briefly review the most important characteristics related to black hole accretion and feedback of the main models in circulation which explicitly explore the physical properties of the black hole populations at $z \geq 4$.

Semi-analytical models, initially put forward in (White & Rees 1978), are built on analytical or numerical halo merger trees (Parkinson et al. 2008; Guo & White 2014), which follow the accretion and merging history of dark matter halos of different masses. The halos are seeded with baryonic matter, and the redshift evolution of the stellar, gas and black hole components is defined by physically-motivated analytical prescription.

Depending on the model, several quantities can be computed from each halo at every time step, such as star formation and black hole accretion rates, their associated feedback, the chemical enrichment of the gas and the spectral energy distribution (SED) of the galaxies. Despite their computational convenience, the drawback of such methods is that they do not provide the spatial information of the galaxies, precluding the possibility of performing clustering studies. The *GALFORM* model (Cole et al. 2000; Lacey et al. 2016; Griffin et al. 2020; Baugh et al. 2019) is built on the merger tree extracted from the Millennium N-body simulation (Springel et al. 2005). Black hole growth takes place during bursts of star formation triggered by mergers or disk instabilities (quasar mode) or through gas accretion from the halo’s corona (radio mode). The accretion rate is not capped, and the radiative efficiency depends on the black hole spin, which in turn is defined by the transfer of angular momentum from the infalling gas to the accreting black hole. The *L-Galaxies* model (Henriques et al. 2015) is also based on a merger tree from the Millennium simulation. In this case, black hole growth takes place mainly during galaxy mergers, which trigger accretion of cold gas mass proportionally to the galaxy satellite-to-central mass ratio. The radiative efficiency is fixed at 10% and the feedback takes the form of relativistic jets. The *MERAXES* model (Mutch et al. 2016; Qin et al. 2017) employs a Bondi-Hoyle accretion model capped at the Eddington limit, while the *SHARK* model (Lagos et al. 2018) has no accretion rate limit, and black hole growth mainly occurs during starbursts. *SAGE* (Croton et al. 2016) implements a Bondi-Hoyle accretion model and distinguishes between radio and quasar feedback modes, the latter corresponding to the main black hole accretion channel.

Hydrodynamical simulations (Carlberg et al. 1990) are computationally more expensive than SAMs, but they follow from first principles the clustering of dark matter and baryonic particles into halos and galaxies, providing spatial and morphological information. In particular, two different particle-tracking methods have been developed, that approach the problem of the discretization of the fluid in different fashions: smoothed particle hydrodynamics (SPH, Gingold & Monaghan 1977; Lucy 1977) and adaptive mesh refinement (AMR Berger & Colella 1989). The first one is a mesh-free Lagrangian method that directly follows the motion of the fluid elements, effectively discretizing the mass, while the latter is a grid-based Eulerian method that discretizes the volume and allows to adapt the required accuracy of the solving algorithms according to the cell considered. The computing power is the main bottle-neck of these simulations, and there is always a trade-off between the size of the simulated region of the Universe (or box) and its mass resolution. Since we are interested in

the statistics of high-redshift quasars, which are hosted in relatively rare massive galaxies, we focus on simulations with a box size $L \gtrsim 100$ Mpc. The *ILLUSTRIS* simulation (Vogelsberger et al. 2014) is built on the hybrid AREPO code (Springel 2011), based on a moving Voronoi tessellation. The black hole accreted mass is defined by the (uncapped) Bondi-Hoyle-Littleton rate, the rotation of the black holes is not considered and the radiative efficiency is fixed at 5%. The *Horizon-AGN* simulation (Dubois et al. 2014) is built on the AMR RAMSES code (Teyssier 2002). The black hole accretion rate is the same as in *ILLUSTRIS*, but this time capped by the Eddington limit. Despite the fact that the spin parameter of the black hole is tracked, the radiative efficiency is fixed at 10%, and both the quasar and radio feedback modes are implemented. On a sub-volume of this same simulation, Trebitsch et al. (2020) built the *OBELISK* simulation, an attempt to bridge the gap between large-volume cosmological simulations and high-resolution zoom-in studies, managing to study at the same time reionization evolution and halos just above the atomic-cooling limit. The *EAGLE* simulation (Crain et al. 2015; Schaye et al. 2015) is based on the SPH code GADGET (Springel 2005). It implements once again the same Bondi-Hoyle-Littleton accretion rate and a single AGN feedback mode with a radiative efficiency of 10%. On a version of the same code is based *MassiveBlackII* (Khandai et al. 2015), which has an accretion rate capped at twice the Eddington limit and distinguishes between the hot and cold phases of the ISM. Finally, we mention the *BlueTides* simulations (Feng et al. 2015), built on yet another version of GADGET (MP-GADGET) and with a black hole accretion and feedback sub-grid model similar to that of *MassiveBlackII*, but in a box with $L \sim 400$ Mpc, which allows the study of the evolution of the most massive halos at $z > 6$.

1.4 Thesis outline

As we have seen, the study of the high-redshift black hole population is an expanding research topic. The details of black hole accretion and feedback are directly related to galaxy evolution, hence representing a fundamental brick in the construction of theoretical simulations and semi-analytical models. In particular, accretion and feedback from growing black holes are not fully understood yet. In this thesis, we make use of a flexible semi-analytical model to shed light on the growth and evolution of black holes in a cosmological context.

In Chapter 2, we introduce black holes into the Delphi semi-analytic model that is used throughout this thesis. We are now fully tracking the

accretion- and merger-driven hierarchical assembly of the dark matter halo, gas, stellar and black hole masses of high-redshift galaxies. Crucially, we explore several different physical scenarios with two types of black hole seeds (stellar and direct-collapse black holes), the impact of reionization and the impact of instantaneous versus delayed galaxy mergers on the baryonic growth. Using a minimal set of mass- and z -independent free parameters associated with star formation, black hole growth and their associated feedback, and including suppressed black hole growth in lower-mass galaxies, we show that our model successfully reproduces all available data sets for early galaxies and quasars. In particular, we find that the effect of reionization dominates in determining the stellar mass density for observed galaxies as well as the black hole mass function. We then use this model to predict the number of merger events at high- z that the space interferometer *LISA* will predict. As expected, mergers of stellar black holes dominate the merger rates for all scenarios and our model predicts an expected upper limit of about 20 mergers using instantaneous merging and no reionization feedback over the 4-year mission duration. Including the impact of delayed mergers and reionization feedback provides about 12 events over the same observational time-scale. This model has a total of 7 free parameters (the strength of the UVB determining DCBH formation, the threshold star formation efficiency, the fraction of SN and BH energies that can couple to the gas content, the halo mass above which black holes can accrete at the Eddington limit and the fraction of available gas that can be accreted by the black hole). In this paper, my role was to carry out a search in this 7-D parameter space to determine the parameter combinations that could jointly reproduce the observed galaxy (UV LF, SMF) and AGN (UV LF, BHMF) observables at high- z (see appendix A of the paper). This was crucial in determining the final set of parameter values used in this paper.

We have then used this model, tuned to reproduce all available data, in Chapter 3, to study the mass assembly and properties of $z > 4$ black holes across the whole mass range $10^3 < M_{\text{bh}}/M_{\odot} < 10^{10}$. After an initial growth phase dominated by black hole mergers down to $z \sim 7$ (9), supermassive black holes in $z = 4$ halo masses of $M_{\text{h}}|_{z=4} \sim 10^{11.75} (10^{13.4}) M_{\odot}$ mainly grow by gas accretion from the interstellar medium. In particular, we find that most of the accretion occurs in the major branch for $M_{\text{h}}|_{z=4} \sim 10^{11-12} M_{\odot}$ halos, while accretion in secondary branches plays a significant role in assembling the black hole mass in higher-mass halos ($M_{\text{h}}|_{z=4} \gtrsim 10^{12} M_{\odot}$); in addition, while the Eddington ratio increases with decreasing redshift for low-mass ($M_{\text{bh}} < 10^5 M_{\odot}$) black holes, the opposite is true at larger masses. Since the accretion rate depends on the gas mass present in the host halo, the duty cycle of the

Eddington-limited accretion phase – which can last up to ≈ 650 Myr – is crucially linked to the joint assembly history of the black hole and its host halo.

Finally, in Chapter 4, we deepen our analysis to study the impact of black hole growth on high-redshift galaxies, both in terms of the observed UV luminosity and of the star formation rate. To do this, firstly, we assess the contribution of AGN to the total galaxy UV luminosity as a function of stellar mass and redshift. Secondly, we study the evolution of the AGN and stellar luminosity functions, finding that it is driven both by changes in their characteristic luminosities M^* (i.e. evolution of the intrinsic brightness of galaxies) and in their normalizations ϕ^* (i.e. evolution of the number densities of galaxies), depending on the luminosity range considered. Finally, we follow the mass assembly history for three different halo mass bins, finding that the magnitude of AGN-driven outflows depends on the host halo mass. AGN feedback is most effective when the energy emitted by the accreting black hole is approximately 1% of the halo binding energy. In such cases, AGN feedback can drive outflows that are up to 100 times more energetic than SN-driven outflows.

2. THE HIERARCHICAL ASSEMBLY OF GALAXIES AND BLACK HOLES IN THE FIRST BILLION YEARS: PREDICTIONS FOR THE ERA OF GRAVITATIONAL WAVE ASTRONOMY

Pratika Dayal, Elena M. Rossi, Banafsheh Shiralilou, Olmo Piana, Tirthankar Roy Choudhury, Marta Volonteri

– MNRAS, 2019, Volume 486, Issue 2, pp.2336-2350

In this work we include black hole (BH) seeding, growth and feedback into our semi-analytic galaxy formation model, *Delphi*. Our model now fully tracks the, accretion- and merger-driven, hierarchical assembly of the dark matter halo, gas, stellar and BH masses of high-redshift ($z \gtrsim 5$) galaxies. We explore a number of physical scenarios that include: (i) two types of BH seeds (stellar and those from Direct Collapse BH; DCBH); (ii) the impact of reionization; and (iii) the impact of instantaneous versus delayed galaxy mergers on the baryonic growth. Using a minimal set of mass- and z -independent free parameters associated with star formation and BH growth, and their associated feedback, and including suppressed BH growth in lower-mass galaxies, we show that our model successfully reproduces all available data sets for early galaxies and quasars. While both reionization and delayed galaxy mergers have no sensible impact on the evolving ultra-violet luminosity function, the impact of the former dominates in determining the stellar mass density for observed galaxies as well as the BH mass function. We then use this model to predict the *LISA* detectability of merger events at high- z . As expected, mergers of stellar BHs dominate the merger rates for all scenarios and our model predicts an expected upper limit of about 20 mergers using instantaneous merging and no reionization feedback over the 4-year mission duration. Including

the impact of delayed mergers and reionization feedback provides about 12 events over the same observational time-scale.

2.1 Introduction

The detection of Gravitational Waves (GWs), from merging stellar Black Hole (BH) binaries and recently from the mergers of binary neutron stars, have opened a new observable window onto the low- z Universe. Over the next decade, we expect the Laser Interferometer Space Antenna (*LISA*) to detect GW signals from merging massive ($10^4 - 10^6 M_\odot$) black hole binaries at the centre of galaxies from redshifts as high as $z \sim 20$, if such BHs are already in place at those early epochs. These observations perfectly complement galaxy surveys, using the *Hubble* and Subaru telescopes and the forthcoming James Webb Space telescope (*JWST*) and the European Extremely Large telescope (*E-ELT*), that aim at yielding tantalising glimpses of the formation of the earliest galaxies in the era of cosmic dawn (for a recent review see Dayal & Ferrara 2018).

Predictions for such observations naturally require theoretical models that can consistently and simultaneously reproduce existing galaxy and BH observations before predictions are made for higher redshifts. In this work, we introduce BH seeding, growth and feedback into the *Delphi* semi-analytic model of galaxy formation. This model now fully tracks the hierarchical assembly of the dark matter, baryonic and BH components of early galaxies including the impact of feedback associated with star-formation, BHs and reionization. We use the results of our model to forecast the BH merger rate and the associated GW signature for *LISA* between $z \sim 5 - 20$. This redshift window covers the formation epoch of supermassive black hole (SMBH) seeds, whose nature and properties are currently outstanding questions. Three main formation channels are currently discussed (see also Sec. 2.2.1) that involve: (i) the first generation of massive metal-free Population III stars (PopIII; e.g. Madau & Rees 2001b); (ii) the monolithic collapse of gas in assembling protogalaxy (e.g., Loeb & Rasio 1994); and (iii) the core collapse of the first ultra-dense nuclear star clusters (e.g., Devecchi & Volonteri 2009). These channels differ in terms of the expected seed mass and “birth” rate with PopIII remnants being the lightest ($\sim 100 M_\odot$) and most frequent and BHs resulting from gas collapse (Direct Collapse Black Holes; DCBHs) being the most massive and rarest (see Latif & Ferrara 2016b; Hartwig 2019, for recent reviews). Therefore, catching the GW signals from merging BHs between $z \sim 5 - 20$ can shed unique insights on SMBH infancy (Colpi

2019). Electromagnetic searches of SMBH seeds probe the fraction of the SMBH population in an active phase and provide their luminosity, therefore shedding light on the combination of their masses and accretion rates. An alternative but still indirect measure of the SMBH mass, which is one of quantities needed to test seed formation scenarios, can be estimated with spectroscopic observations. On the other hand, *LISA* can detect and directly provide the masses and spins for quiescent SMBHs, as long as they are in a coalescing binary. Therefore these two approaches are truly complementary and both are necessary in order to obtain a full view of the SMBH seed population in the early Universe.

In this work we focus on BH-BH mergers that ensue after the merger of two galaxies, both hosting a central black hole, rather than the merger of SMBHs that are born in a binary (Hartwig et al. 2018). While such calculations have been carried out by previous works (e.g. Sesana et al. 2007, 2011; Barausse 2012), the strength of our model lies, both, in the minimal set of free parameters used for star formation and BHs (and their associated feedback) as well as the number of physical scenarios explored that include: (i) two types of BH seeds (stellar and those from direct-collapse black holes); (ii) the impact of reionization; and (iii) the impact of instantaneous versus delayed galaxy mergers on the baryonic growth of galaxies. We also check the dependence of our results on the DCBH seeds mass used, by showing the *LISA* event rates expected for seed masses higher than for our fiducial model by an order of magnitude, ranging between $10^{4-5} M_{\odot}$. The results of this *heavy DCBH seeds* model are shown in Sec. 2.3.5.3. Our model matches the key observables both for high- z Lyman break galaxies (LBGs)- including the Ultra-violet luminosity function (UVLF), the stellar mass function (SMF), the stellar mass density (SMD), stellar mass-halo mass relations and the mass-to-light (M/L) ratios- and black holes - including their UV LF, the black mass function (BHMF) and the BH mass-stellar mass relations. The GW event rates predicted by this work are therefore benchmarked against all available high- z galaxy and BH data.

The cosmological parameters used in this work are $\Omega_m, \Omega_{\Lambda}, \Omega_b, h, n_s, \sigma_8 = 0.3089, 0.6911, 0.049, 0.67, 0.96, 0.81$, consistent with the latest results from the *Planck* collaboration (Planck Collaboration et al. 2016a). We quote all quantities in comoving units unless stated otherwise and express all magnitudes in the standard AB system (Oke & Gunn 1983a).

2.2 The theoretical model

This work is based on using the code *Delphi* (Dark Matter and the emergence of galaxies in the epoch of reionization), introduced in previous papers including Dayal et al. (2014a, 2015, 2017a,b). In brief, *Delphi* uses a binary merger tree approach to jointly track the build-up of dark matter halos, their baryonic component (both gas and stellar mass) and their star-formation driven spectra through cosmic time. We start by building merger trees for 550 $z = 4$ galaxies, uniformly distributed in the halo mass range of $\log(M_h/M_\odot) = 8 - 13.5$, up to $z = 20$. Each $z = 4$ halo is assigned a co-moving number density by matching to the dn/dM_h value of the $z = 4$ Sheth-Tormen halo mass function (HMF) and every progenitor halo is assigned the number density of its $z = 4$ parent halo; we have confirmed that the resulting HMFs are compatible with the Sheth-Tormen HMF at all z . In terms of feedback, so far, this model has focused on modelling the impact of (TypeII) supernovae (SNII) and reionization feedback on the formation of early galaxies. In this work, we extend our model to include the impact of BH seeding, growth and feedback on early galaxy formation. The very first progenitors of every $z = 4$ halo, that mark the start of its assembly (the so-called “starting leaves”), are assigned an initial gas mass that scales with the halo mass according to the cosmological ratio such that $M_g = (\Omega_b/\Omega_m)M_h$. Depending on their mass and redshift, such starting leaves can also be seeded with a black hole as explained in Sec. 2.2.1. We start by calculating the star formation efficiency of a halo and the gas-mass remaining after SN feedback (Sec. 2.2.2). If a halo hosts a black hole, a part of the gas left after star formation and SN feedback can be accreted onto the black hole and the impact of black hole feedback is included as detailed in Sec. 2.2.3. At each step, we include both the impact of smooth-accretion and mergers in assembling the halo, baryonic (gas and stellar mass) and black hole mass as detailed in Secs. 2.2.4 and 2.2.5. In our endeavour to build a model with minimal free parameters, we limit ourselves to two and four mass- and z -independent free parameters related to star formation and black holes, respectively.

2.2.1 Seeding halos with black holes

First, we decide whether the starting leaves of each halo merger tree are eligible to host a black hole seed. Here we consider the possibility of more than one BH formation mechanism operating in the (early) Universe, as generally expected (e.g., Valiante et al. 2016a). In particular, we explore the two formation channels that yield the least (Pop III) and the most

massive (DCBH) black hole seeds. These black hole seeds are planted *in the starting leaves of any halo* as now detailed.

Figure 1: The cumulative redshift distribution of the number density of light (stellar black hole; solid black line) and heavy (DCBH) seeds in our model. For the latter we show the distribution of DCBH seeds using a value of $\alpha = 30$ (short-dashed blue line) and $\alpha = 300$ (long-dashed light-blue line).

(i) **Heavy seeds:** first postulated as massive ($10^{3-5}M_{\odot}$) black hole seeds to explain the presence of SMBHs at early cosmic epochs (e.g. Loeb & Rasio 1994; Bromm & Loeb 2003), DCBH formation models have been continually refined and developed over the past years (e.g. Begelman et al. 2006, 2008; Regan & Haehnelt 2009; Shang et al. 2010; Johnson et al. 2012; Latif et al. 2013a; Agarwal et al. 2014; Dijkstra et al. 2014; Ferrara et al. 2014; Habouzit et al. 2016a). The current understanding from these works requires the following conditions to be met for a DCBH host: first, the halo should have reached the atomic cooling threshold, with a virial temperature $T_{\text{vir}} \gtrsim 10^4\text{K}$, for the gas to be able to cool isothermally; secondly, the halo should be metal-free to prevent gas fragmentation; and finally, the halo should be exposed to a high enough “critical” Lyman-Werner (LW) background ($J_{\text{crit}} = \alpha J_{21}$). Here $\alpha > 1$ is a free parameter and J_{21} is the LW background expressed in units of $10^{-21}\text{erg s}^{-1}\text{Hz}^{-1}\text{cm}^{-2}\text{sr}^{-1}$ (see e.g. Sugimura et al. 2014). Interested readers are referred to Dayal et al. (2017b) for complete details on how DCBHs are seeded in high- z halos. In

brief, we start by making the reasonable assumption that the starting leaves of any halo are metal-free by virtue of never having accreted metal-enriched gas. Further, we use the stellar population synthesis code Starburst99 (Leitherer et al. 1999) to calculate the LW (11.2 – 13.6 eV) luminosity of each galaxy based on its entire star formation history. This is used to calculate the mean LW emissivity at a given redshift, $\epsilon_{\text{LW}}(z)$, by integrating over all galaxies present at that z . Accounting for fluctuations in the background, most likely around galaxies and from the biased (i.e. clustered) distribution of galaxies, we identify the probability of the starting leaves being irradiated by LW intensity above a critical threshold value as detailed in Dayal et al. (2017b); for the calculations in this work we explore values that range over an order of magnitude such that $\alpha = 30$ and 300. The mass distribution of the seeds is uncertain and depends on the specific physical conditions at birth and on whether the intermediate state of a supermassive star is followed by a brief period of super-Eddington accretion onto the newly born BH (i.e. a quasistar phase). For example, the supermassive star mass depends on the strength of the LW radiation that illuminates the birth site (Latif 2019; Agarwal 2019). On the other hand, the existence of a quasistar phase and its outcome in term of BH seed mass depend on internal rotation and on mass loss in winds (Dotan et al. 2011; Fiacconi & Rossi 2016, 2017). We are therefore left with an uncertain SMBH seed mass that can be bracketed by $10^3 - 10^5 M_\odot$ in halos below $10^9 M_\odot$. Given the halo masses and LW radiation thresholds used in this work (cf. Fig. 5 in Latif 2019), we randomly populate halos in the top half of the calculated probability range with a DCBH seed of mass ranging between $10^3 - 4 M_\odot$ (light DCBH seeds): the number of halos populated with such DCBHs is calculated by matching this DCBH mass function to the probabilistic one (obtained by multiplying the mass function of DCBH hosts with the hosting probability).

(ii) Light seeds: stellar black hole seeds of mass $\sim 10^2 M_\odot$ can be created by the collapse of Pop-III stars in minihalos with $M_h \sim 10^5 M_\odot$. Making the reasonable assumption that halos collapsing from high ($\gtrsim 3.5$)- σ fluctuations in the primordial density field are most likely to host such seeds (e.g. Volonteri et al. 2003; Barausse 2012) results in the host halos being more massive than $M_h \gtrsim 10^{7.2} M_\odot$ at $z \gtrsim 13$. However, given our halo mass range of $10^{8-13.5} M_\odot$, starting leaves have to be assigned these seeds by hand. In this work, we start by populating the starting leaves of any halo, that fulfil the DCBH criterion detailed above, with seed DCBHs. The starting leaves of halos at $z \gtrsim 13$ that fulfil the light seed criterion, but do not contain a DCBH, are then populated with stellar BH seeds mass $M_{\text{bh}} = 150 M_\odot$.

The initial seed distribution obtained with this formalism is shown in Fig. 1. The cumulative number density of stellar BH seeds has a value of about $10^{-3.8} \text{ Mpc}^{-3}$ by $z \sim 4$. This clearly dominates over the cumulative number density of DCBH seeds, which has a value of about $10^{-5.8} (10^{-7.6}) \text{ Mpc}^{-3}$ by $z \sim 4$ for $\alpha = 30$ (300). Note that the models for “light seeds” described above were inspired by early calculations that suggested that only one, very massive, Pop-III star would form in a given halo, right at the centre of the potential well (e.g., Abel et al. 2002). More recent models favour a larger amount of fragmentation, leading to lighter and scattered Pop-III remnants (Jeon et al. 2014; Smith et al. 2018) which are less suitable as SMBH seeds given the difficulty of both accreting material from their surroundings as well as finding the dynamical center of the galaxy/halo via dynamical friction.

Therefore we also run test cases with only DCBHs as a lower limit to the presence of SMBH seeds in primeval galaxies. In terms of general BH population, models including only DCBHs with $\alpha = 300$ fail entirely in reproducing the observed active galactic nuclei (AGN) luminosity function at $z = 5 - 6$, given their extremely low number densities; consider the long-dashed cyan line in Fig. 1 which has a value of about $10^{-7.6} \text{ Mpc}^{-3}$ and compare this to an observed AGN number density of $\sim 10^{-5.5} - 10^{-4.5} \text{ Mpc}^{-3}$ (Vito et al. 2018b; Kulkarni et al. 2019a). Models with $\alpha = 30$ are nearly compatible with the faint end of the observed luminosity function at $z = 5$, but fail to produce enough AGN at $z = 6$ unless the seed mass is above $10^4 M_\odot$. In terms of GW signatures, the results are similar to the reference case but selecting only mergers between heavy seeds. In summary, both “light” and “heavy” seed models have uncertainties and problems associated with their formation and growth. Given this state-of-the-art in the field of BH formation, our model comprehensively explores the allowed parameter space, includes a realistic approach to the dependence of BH growth on the host mass, and presents a thorough comparison with observations in Sec. 2.3 in order to explore the consequences of current uncertainties.

2.2.2 Star formation and supernova feedback

Given our first leaves - with or without a black hole seed - we can A newly-formed stellar population of mass $M_*(z)$ at redshift z can impart the interstellar medium (ISM) with a total SN-II energy E_{SN} given by

$$E_{\text{SN}} = f_*^w E_{51} \nu M_*(z) \equiv f_*^w v_s^2 M_*(z). \quad (2.1)$$

Here, each SN-II is assumed to impart an (instantaneous) explosion energy of $E_{51} = 10^{51} \text{erg}$ to the ISM and $\nu = [134 M_{\odot}]^{-1}$ is the number of SN-II per stellar mass formed for a Salpeter IMF between $0.1 - 100 M_{\odot}$; we maintain this IMF through-out this work. The values of E_{51} and ν yield $v_s = 611 \text{ km s}^{-1}$. Finally, f_*^w is the fraction of the SN explosion energy that couples to gas.

For any given halo, the energy E_{ej} required to unbind and eject the ISM gas not converted into stars can be expressed as

$$E_{\text{ej}}(z) = \frac{1}{2} [M_{\text{gi}}(z) - M_*(z)] v_e^2, \quad (2.2)$$

where $M_{\text{gi}}(z)$ is the *initial* gas mass in the galaxy, prior to any star formation or BH accretion, at epoch z . Further, the escape velocity v_e can be expressed in terms of the halo rotational velocity, v_c , as $v_e = \sqrt{2} v_c$.

We then define the *ejection efficiency*, f_*^{ej} , as the fraction of gas that must be converted into stars to “blow-away” the remaining gas from the galaxy (i.e. $E_{\text{ej}} \leq E_{\text{SN}}$). This can be calculated by imposing $E_{\text{ej}} = E_{\text{SN}}$ leading to

$$f_*^{\text{ej}}(z) = \frac{v_c^2(z)}{v_c^2(z) + f_*^w v_s^2}. \quad (2.3)$$

The *effective star formation efficiency* for any halo is then expressed as

$$f_*^{\text{eff}} = \min[f_*, f_*^{\text{ej}}], \quad (2.4)$$

where f_* is a free parameter representing the maximum instantaneous star formation efficiency - this parameter is fixed by matching to the bright-end of the observed LBG UV LF as explained in Sec. 2.3.1.

In this formalism, the newly formed stellar mass formed at z can be expressed as

$$M_*(z) = M_{\text{gi}}(z) f_*^{\text{eff}}. \quad (2.5)$$

In the spirit of maintaining simplicity, we assume that every stellar population has a fixed metallicity of $0.05 Z_{\odot}$ and each newly-formed stellar population has an age of 2 Myr . Using these parameters with the population synthesis code STARBURST99 (Leitherer et al. 1999), the rest-frame UV luminosity (between 1250 and $1500 \text{ \AA}$) from a newly-formed stellar mass can be expressed as

$$L_*^{\text{UV}} = 10^{33.077} \left(\frac{M_*}{M_{\odot}} \right) \text{ erg s}^{-1} \text{ \AA}^{-1}. \quad (2.6)$$

This star-formation episode must then result in a certain amount of gas, $M_*^{\text{ge}}(z)$, being ejected from the galaxy at the given z -step. The value of $M_*^{\text{ge}}(z)$ depends on whether $f_*^{\text{eff}} = f_*$ or $f_*^{\text{eff}} = f_*^{\text{ej}}$: while the galaxy is an “efficient star-former” in the former case, that can support new stellar mass being formed without losing much of its gas, the latter case is true for a “feedback-limited” system that loses all of its ISM gas after star formation. Mathematically, M_*^{ge} can be calculated as

$$M_*^{\text{ge}}(z) = [M_{\text{gi}}(z) - M_*(z)] \frac{f_*^{\text{eff}}}{f_*^{\text{ej}}}. \quad (2.7)$$

The final gas mass, $M_*^{\text{gf}}(z)$, remaining in the galaxy at that redshift-step, after star formation and SN feedback, can then be expressed as

$$M_*^{\text{gf}}(z) = [M_{\text{gi}}(z) - M_*(z)] \left[1 - \frac{f_*^{\text{eff}}}{f_*^{\text{ej}}} \right]. \quad (2.8)$$

Model	α	UVB	f_*	f_*^w	f_{bh}^w	f_{bh}^{ac}	$f_{Edd}(M_h < M_h^{crit})$	$f_{Edd}(M_h > M_h^{crit})$	ϵ_r
ins1	30	No	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
ins2	300	No	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
ins3	30	Yes	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
ins4	300	Yes	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
tdf1	30	No	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
tdf2	300	No	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
tdf3	30	Yes	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
tdf4	300	Yes	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1

Table 2.1: Values of the free parameters used for the models indicated in column 1. In our notation, α indicates the minimum threshold of LW background radiation, in units of f_{21} , needed for DCBH seed formation, f_* is the star formation efficiency cap, f_*^w (f_{bh}^w) is the fraction of SN (black hole) energy that couples to the ISM and drives outflows, f_{bh}^{ac} is the maximum fraction of gas present in the galaxy which is available for black hole accretion, f_{Edd} is the Eddington ratio at which the black hole can accrete, M_h^{crit} is the threshold halo mass below which black hole growth is stunted and ϵ_r is the black hole radiative efficiency.

2.2.3 Black hole growth and feedback

Once seeded, black holes can grow via accretion and mergers. We discuss black hole growth via accretion in this section and the merger-driven growth is deferred to Sec. 2.2.5 that follows. At any given redshift, the Eddington mass accretion rate, $\dot{M}_{\text{ed}}$, for a black hole of mass M_{bh} can be calculated as

$$\dot{M}_{\text{ed}}(z) = \frac{4\pi G M_{\text{bh}}(z) m_{\text{p}}}{\sigma_{\text{T}} \epsilon_{\text{r}} c}, \quad (2.9)$$

where G is the gravitational constant, m_{p} is the proton mass, σ_{T} is the Thomson scattering optical depth, ϵ_{r} is the black hole radiative efficiency and c is the speed of light. Given our merger tree time-step of $\Delta t = 20 \text{ Myr}$, the total mass that can be accreted at the Eddington rate in one time-step is $M_{\text{Edd}}(z) = x_{\text{r}} \dot{M}_{\text{ed}}(z) \times \Delta t$, where $x_{\text{r}} = (1 - \epsilon_{\text{r}})$. Further, we define the maximum gas mass that can be accreted by the BH in a given time step, $M_{\text{bh}}^{\text{ac}}(z)$, as:

$$M_{\text{bh}}^{\text{ac}}(z) = \min \left[f_{\text{Edd}} M_{\text{Edd}}(z), x_{\text{r}} f_{\text{bh}}^{\text{ac}} M_{*}^{\text{gf}}(z) \right], \quad (2.10)$$

where $f_{\text{bh}}^{\text{ac}}$ is the fraction of the available gas mass left, after star formation and supernova feedback, that can be accreted by the BH, and f_{Edd} , a free-parameter, is the fractional Eddington rate of accretion. Using this formalism, the BH accretes either at a fraction of the Eddington rate or a fraction of the available gas mass, whichever is lower. Matching to the AGN UV LF (Sec. 2.3.1) requires $f_{\text{Edd}} = 7.5 \times 10^{-5}$ below a critical halo mass of $M_{\text{h}}^{\text{crit}} = 10^{11.25} [\Omega_{\text{m}}(1+z)^3 + \Omega_{\Lambda}]^{-0.125}$ (see also Bower et al. 2017) and $f_{\text{Edd}} = 1$ above this mass at any redshift. The resulting accretion will yield a black hole feedback energy of

$$E_{\text{bh}} = f_{\text{bh}}^{\text{w}} \epsilon_{\text{r}} M_{\text{bh}}^{\text{ac}}(z) c^2, \quad (2.11)$$

where f_{bh}^{w} is the efficiency of BH feedback coupling to the gas.

The black hole feedback required to eject the left-over gas (after accretion) can be expressed as

$$E_{\text{bh}}^{\text{ej}}(z) = \frac{1}{2} [M_{*}^{\text{gf}}(z) - M_{\text{bh}}^{\text{ac}}(z)] v_{\text{e}}^2. \quad (2.12)$$

The *effective black hole feedback* is therefore taken to be the minimum between the energy required to eject all the gas up to the maximum value such that $E_{\text{bh}}^{\text{eff}} = \min[E_{\text{bh}}, E_{\text{bh}}^{\text{ej}}]$. The final gas mass left in the halo, that

can be carried over for mergers at the end of the time-step, after black hole accretion and feedback, can then be calculated as

$$M_{\text{bh}}^{\text{gf}}(z) = [M_{*}^{\text{gf}}(z) - M_{\text{bh}}^{\text{ac}}(z)] \left[1 - \frac{E_{\text{bh}}^{\text{eff}}}{E_{\text{bh}}^{\text{ej}}} \right]. \quad (2.13)$$

2.2.4 Smooth-accretion from the intergalactic medium

The merger of halos is accompanied by “smooth accretion” of dark matter from the intergalactic medium (IGM). In the analytic merger tree, this smoothly-accreted dark matter mass is calculated as

$$M_{\text{dm}}^{\text{sa}}(z) = M_{\text{h}}(z) - \sum_{i=1}^N M_{\text{h}}(z + \Delta z), \quad (2.14)$$

where $M_{\text{h}}(z)$ is the halo mass at z and the second term on the RHS denotes the sum of the halo masses of all its progenitors at the previous redshift step.

We make the reasonable assumption that the smooth accretion of dark matter is accompanied by the accretion of a cosmological ratio of gas from the IGM such that the smoothly-accreted gas mass M_{g}^{sa} at z can be written as

$$M_{\text{g}}^{\text{sa}}(z) = \frac{\Omega_{\text{b}}}{\Omega_{\text{m}}} M_{\text{dm}}^{\text{sa}}(z). \quad (2.15)$$

2.2.5 Merging galaxies and black holes

As galaxies merge, in addition to dark matter, they bring in stellar and gas mass, with the latter depending on the star formation and BH accretion efficiencies of the progenitor halos. As detailed in Secs. 2.2.2 and 2.2.3, galaxies forming stars and/or accreting at the limit of BH feedback will only bring stellar mass into their successors resulting in “dry mergers”. On the other hand, halos bringing in both stellar and gas mass result in “wet mergers”. Mergers of galaxies result in a summation of their dark matter,

baryonic and BH masses and total UV luminosities such that

$$M_{\text{dm}}(z) = \sum_{i=1}^N M_{\text{h}}(z + \Delta z) + M_{\text{dm}}^{\text{sa}}(z), \quad (2.16)$$

$$M_{\text{g}}(z) = \sum_{i=1}^N M_{\text{bh}}^{\text{gf}}(z + \Delta z) + M_{\text{g}}^{\text{sa}}(z), \quad (2.17)$$

$$M_{*,\text{tot}}(z) = M_*(z) + \sum_{i=1}^N M_*(z + \Delta z), \quad (2.18)$$

$$M_{\text{bh,tot}}(z) = M_{\text{bh}}(z) + \sum_{i=1}^N M_{\text{bh}}(z + \Delta z), \quad (2.19)$$

$$L_{\text{tot}}^{\text{UV}}(z) = L_*^{\text{UV}}(z) + L_{\text{bh}}^{\text{UV}}(z) + \sum L_*^{\text{UV}}(z + \Delta z). \quad (2.20)$$

Using STARBURST99, we find that the UV luminosity for a burst of stars (normalized to a mass of $1M_{\odot}$ and metallicity $0.05 Z_{\odot}$) decreases with time as

$$\log \left(\frac{L_*^{\text{UV}}(t)}{\text{erg s}^{-1} \text{\AA}^{-1}} \right) = 33.0771 - 1.33 \log(t/t_0) + 0.462, \quad (2.21)$$

where t is the age of the stellar population (in yr) at z and $\log(t_0/\text{yr}) = 6.301$. Finally, we make the limiting assumption that BH luminosity decays away within the time-step of 20 Myrs i.e. BH luminosity is only relevant in the redshift step in which the black hole accretes.

We also explore two scenarios for the timescales of galaxy mergers: in the first, halos mergers are accompanied by the mergers of galaxies and their constituent black holes. This instantaneous merging scenario sets the upper limit for the galaxy and black hole merger rate. In the second scenario, we include the fact that galaxies (and their black holes) merge after a “merging” timescale which can be calculated as (Lacey & Cole 1993a)

$$\tau = f_{\text{df}} \Theta_{\text{orbit}} \tau_{\text{dyn}} \frac{M_{\text{host}}}{M_{\text{sat}}} \frac{0.3722}{\ln(M_{\text{host}}/M_{\text{sat}})}, \quad (2.22)$$

where M_{host} is the mass of the host including all the satellites, M_{sat} is the mass of the merging satellite, τ_{dyn} represents the dynamical timescale and f_{df} represents the efficiency of tidal stripping; $f_{\text{df}} > 1$ if tidal stripping is very efficient. For this work, we use $f_{\text{df}} = 1$. Further,

$$\Theta_{\text{orbit}} = \left(\frac{J}{J_{\text{c}}} \right)^{0.78} \left(\frac{r_{\text{c}}}{R_{\text{vir}}} \right)^2, \quad (2.23)$$

where J is the satellite’s specific angular momentum and J_c that of a satellite carrying the same energy and orbiting on a circular orbit. The last term represents the ratio between the circular radius (the radius of a circular orbit with the same energy) and the virial radius of the host. Θ_{orbit} is well modelled by a log-normal distribution such that $\log(\Theta_{\text{orbit}}) = -0.14 \pm 0.26$ (Cole et al. 2000); we randomly sample values from this distribution for each merger. Finally, the dynamical timescale can be calculated as $\tau_{\text{dyn}} = \pi R_{\text{vir}}(z) V_{\text{vir}}(z)^{-1} = 0.1 \pi t_{\text{H}}(z)$ where R_{vir} and V_{vir} are the virial radius and velocity at z . We also make the limiting assumption that satellite galaxies, waiting to merge, neither form stars nor have any BH accretion of gas.

We show results from both scenarios in this work in order to bracket the uncertainty on our understanding of galaxy merger timescales. In reality, however, the second scenario still gives a lower limit to the merger time of BH binaries: additional delays are to be expected once the binary forms after the dynamical friction timescale. The binary has to harden further to reach the separation where gravitational wave emission becomes the dominant source of energy and momentum losses to bring the binary to coalescence and merge within the Hubble time. This additional delay is largely unconstrained, and can range from tens to billions of years (for a review, see Barack et al. 2019). Additionally, for low-mass BHs dynamical friction could be ineffective during the galaxy merger phase since the deceleration of dynamical friction is proportional to the infalling black hole mass resulting in the timescale for orbital decay being inversely proportional to mass.

2.2.6 The impact of reionization

In this work, we also include the effects of the ultra-violet background (UVB) created during reionization which, by heating the ionized IGM to $T \sim 10^4$ K, can have an impact on the baryonic content of low-mass halos (for a recent review see, e.g., Dayal & Ferrara 2018). Given that self-consistently calculating the impact of the UVB on galaxy formation remains an unsolved problem, we consider two scenarios: the first is such that UVB has no impact on the baryonic content of any halo. The second scenario considers maximal UVB feedback in which the gas mass is completely photo-evaporated for all halos below a characteristic virial velocity of $V_{\text{vir}} = 40 \text{ km s}^{-1}$. Critically, whilst leaving the number of mergers unchanged, limiting the gas mass available for accretion onto black holes, the latter scenario results in a lower limit on the mass of the merged galaxies and black holes.

We run the above model for 8 different scenarios (detailed in Table 2.1) that explore 3 key physical effects: (i) the impact of the LW background amplitude in calculating DCBH hosts (Sec. 2.2.1); (ii) the impact of instantaneous versus delayed mergers of galaxies and black holes (Sec. 2.2.5); and (iii) the impact of UV feedback on early galaxies and black holes (Sec. 2.2.6). In what follows, the model *ins1*, that assumes the lowest LW amplitude for DCBHs, instantaneous galaxy mergers and no UV feedback, is denoted as the *fiducial* model and provides the upper limit on our results. On the other hand, with the highest LW amplitude for DCBHs, delayed galaxy (and BH) mergers and maximal UV feedback, the model *tdf4* provides the lower limit to our results.

2.3 Comparing theoretical properties to observations

Now that the theoretical model is in place, we compare model results to a number of observational data sets, for both galaxies and AGN, as detailed in what follows. In terms of galaxies, we use the data sets accumulated for high- z LBGs, detected through a drop in luminosity blue-ward of the Lyman limit at $912\text{\AA}$. The past few years have seen an enormous increase in LBG data due to a combination of state-of-the-art instruments such as (the Wide Field Came 3 onboard) the Hubble Space Telescope (*HST*) as well as refined selection techniques (e.g. Steidel et al. 1999). As for AGN, a number of surveys, including the Sloan Digital Sky Survey (SDSS) and the Canadian-French high- z quasar surveys, and observations with the Subaru telescope, have yielded a statistical sample of AGN/QSO candidates at redshifts as high as $z \simeq 6$. In what follows we compare 4 models that bracket the physically plausible range explored in this work (*ins1*, *ins4*, *tdf1* and *tdf4*), detailed in Table 2.1, with a number of data-sets, including the UV LFs, the stellar mass density, the black hole mass function and the black hole-stellar mass relation.

We note that, given their low number densities, both the “light” and “heavy” DCBH seeding cases yield very similar results for all the observational data-sets discussed. For this reason, we limit our results to the “light DCBH seed” case in this section.

2.3.1 The observed UV LF for star formation and black holes

The observed UV LF (number density of galaxies as a function of the absolute magnitude) and its redshift evolution offer one of the most robust

tests of theoretical models of galaxy formation. We start by calculating the UV magnitudes, separately for star formation and AGN activity, for each theoretical galaxy and computing the associated UV LFs, as shown in Fig. 2. We start by discussing the LBG UV LF: firstly, matching to the bright end of the evolving UV LF requires a maximum star formation efficiency value of $f_* \simeq 2\%$. Secondly, we find that the fiducial model (*ins1*) is in excellent agreement with available LBG observations, ranging between $-22 \lesssim M_{UV} \lesssim -13$, at all $z \sim 5 - 10$ as already shown in our previous works (e.g. Dayal et al. 2014a). The inclusion of a delay in galaxy mergers (*tdf1*) has no sensible impact on the faint-end of the UV LF - this is due to the fact that the progenitors of these low-mass halos are SN feedback limited and hence do not bring in any gas whilst merging (dry mergers) as already pointed out previously (Dayal et al. 2014a). On the other hand, the delay in galaxy mergers leads to an increasing reduction in the gas masses of higher-mass halos whose progenitors are not SN feedback limited and bring in gas in mergers (wet mergers), leading to a slightly steeper bright end. Finally, including the (maximal) impact of reionization feedback (*ins4* and *tdf4*), that photo-evaporates the baryonic content of all galaxies with $V_{vir} \lesssim 40 \text{ km s}^{-1}$, only affects the faint-end of the UV LF and leads to a cut-off at brighter magnitudes ($M_{UV} \sim -14$ to -15) as compared to the continued rise excluding this effect (e.g. in models *ins1* and *tdf1*). In this work, models *ins1* and *tdf4*, therefore, bracket the plausible UV LF range. However, it must be cautioned that the theoretical LBG UV LF has, so far, ignored the impact of dust enrichment which is expected to have a relevant effect in decreasing the luminosities at the bright end.

Focusing on the AGN UV LF, the black hole powered UV LFs for all four models discussed above are found to be in excellent agreement with all available AGN data at $z \sim 5$ and 6 as shown in Fig. 2. We start by noting that given the large masses ($M_h \gtrsim 10^{11.5} M_\odot$) associated with AGN/QSO host halos, the black hole UV LF is only relevant at $M_{UV} \lesssim -21$, corresponding to number densities $\lesssim 10^{-5} [\text{dex}^{-1} \text{Mpc}^{-3}]$ at $z \sim 5$ and 6 . These results are in qualitative agreement with those of Ono et al. (2018) who find 100% of the UV luminosity to come solely from stars for galaxies with $M_{UV} \gtrsim -23$ to -24 . However, given that the AGN number densities are suppressed due to obscuration (see Sec. 2.2.3), calculating the fraction of galaxies dominated by AGN requires a more thorough examination which we defer to a future work. Finally, we note that the contribution of BH-powered luminosity could be one explanation for observed UV LFs that are shallower than the exponentially declining Schechter function at these high- z (e.g. Ono et al. 2018).

Figure 2: The UV LF from $z \simeq 5 - 10$ as marked in the panels. In each panel, the violet points show the available LBG data collected both using space- and ground-based observatories at: (a) $z \simeq 5$ (Bouwens et al. 2007; McLure et al. 2009); (b) $z \simeq 6$ (McLure et al. 2009; Bouwens et al. 2015b; Livermore et al. 2017); (c) $z \simeq 7$ (Bouwens et al. 2010; McLure et al. 2010; Castellano et al. 2010; McLure et al. 2013; Bowler et al. 2014; Livermore et al. 2017; Atek et al. 2015a); (d) $z \simeq 8$ (Bouwens et al. 2010; McLure et al. 2010; Bradley et al. 2012; McLure et al. 2013; Livermore et al. 2017; Atek et al. 2015a); (e) $z \simeq 9$ (McLure et al. 2013; Oesch et al. 2013); and (f) $z \simeq 10$ (Bouwens et al. 2015a; Oesch et al. 2014). In each panel, the yellow points show the AGN data collected at: $z \simeq 5$ (McGreer et al. 2013; Parsa et al. 2018) and $z \simeq 6$ (Willott et al. 2010a; Kashikawa et al. 2015; Parsa et al. 2018; Jiang et al. 2016). In each panel, lines show model UV LFs for galaxies and black holes for the following models summarised in Table 1, that bracket the range of UV LFs allowed in the presence/absence of a UVB and for both instantaneous and delayed (by a merging timescale) merger: *ins1* (galaxies solid black line; BH solid gray line), *ins4* (galaxies short dashed red line; BH short dashed light-red line), *tdf1* (galaxies long dashed green line; BH long-dashed light green line) and *tdf4* (galaxies dot-dashed blue line; BH dot-dashed purple line).

We find that the AGN UV LF is extremely similar for heavy black hole seeds with α varying over an order of magnitude (for 30 to 300) for the four models discussed above. This is probably to be expected given the extremely low number of heavy black hole seeds as compared to the number of light black hole seeds as pointed out in Sec. 2.2.1; the latter therefore clearly dominate the UV LF. As for the merger timescales, including a delay in the mergers of galaxies (and black holes) results in a smaller black hole growth. This is reflected in a lower final black hole mass in a given halo (also see Sec. 2.3.3 that follows). However, this only leads to minor changes in the UV LF which are indistinguishable within the scatter shown by the four models considered here. Further, given the large masses of AGN hosts, reionization feedback has no relevant effect on the AGN UV LF. Finally, looking at the redshift evolution of the AGN UV LF, we find that it shows a sharper redshift evolution compared to the star-formation powered UV LF given the increasing paucity of their high-mass hosts. To quantify this effect, let us focus on a magnitude of $M_{UV} = -20$: while the star formation driven UV LF only evolves by a factor of 3 between $z \sim 5$ and 7, the AGN UV LF (negatively) evolves by roughly three orders of magnitude over the same redshift range.

2.3.2 The LBG stellar mass density (SMD)

We now compare the theoretical SMD to that observationally inferred for LBGs. We start by comparing to observed LBGs with $M_{UV} \lesssim -17.7$ as shown in Fig. 3. As seen, while all four models (*ins1*, *ins4*, *tdf1*, *tdf4*) are in excellent agreement with the data they are offset in normalisation from each-other whilst following very similar slopes such that $SMD \propto (1+z)^{0.42}$. As might be expected, model *ins1* provides the upper limit to the SMD results for observed galaxies. Including the effects of delayed galaxy mergers (*tdf1*) results in a small decrease in the SMD values by about 0.1 dex. However, assuming instantaneous mergers whilst including maximal UVB suppression (*ins4*) only results in a SMD that is different from the fiducial case by a negligible 0.03 dex. These results clearly imply that a delay in the merger timescales is more important than the effect of a UVB for these high mass systems. Finally, the lower limit to the SMD results is provided by model *tdf4* that is about 0.13 dex lower than the fiducial results. These slight changes in the SMD normalisation shows that most of the stellar mass is assembled in massive progenitors (see also Dayal et al. 2013) with low-mass progenitors - that either merge after a dynamical timescale (*tdf1*), are reionization suppressed (*ins4*) or include

Figure 3: The LBG stellar mass density (SMD) as a function of redshift. Points show the observational data collected by: González et al. (2011, red empty circles), Labbé et al. (2013, blue empty triangles), Stark et al. (2013, purple empty squares), Oesch et al. (2014, yellow empty circles), Duncan et al. (2014, red filled squares), Grazian et al. (2015, purple filled circles) and Song et al. (2016, yellow filled triangles). We show results for galaxies with $M_{UV} < -17.7$ which can be directly compared to observational data points for the following models shown in Table 1: *ins1* (solid black line), *ins4* (dot-dashed red line), *tdf1* (long dashed green line) and *tdf4* (dot-dashed blue line). We also show results for the total SMD obtained by summing over all galaxies at a specific z for the same models noted above: *ins1* (solid gray line), *ins4* (dot-dashed light red line), *tdf1* (long dashed light green line) and *tdf4* (dot-dashed purple line).

both these effects (*tdf4*) - contributing only a few percent to the stellar mass for observed galaxies.

On the other hand, the impact of reionization feedback and a delay in the merging timescale are much more dramatic when considering the entire galaxy population (without any limiting magnitudes used). In this case, the fiducial model, *ins1*, shows a slope that evolves with redshift as $SMD \propto (1+z)^{0.24}$. Given that in this case the SMD is dominated by the contribution from low-mass halos, the situation flips as compared to that discussed above: the merger timescale has a negligible effect on the SMD of all galaxies and shows essentially the same amplitude and slope as the fiducial case. However, the UVB suppression of the gas mass of low-mass

halos results in both a decrease in the amplitude (by about 0.23 dex) and a, more dramatic, steepening of the SMD slope such that $\text{SMD} \propto (1+z)^{0.31}$ for models *ins4* and *tdf4*.

2.3.3 The black hole mass function and occupation fraction

Figure 4: The black hole mass function (BHMF) at $z \simeq 6$. We compare observational results (grey line with error bars) from Willott et al. (2010a) to those from our models bracketing the plausible physical range: *ins1* (solid black line), *ins4* (short-dashed red line), *tdf1* (long-dashed green line) and *tdf4* (dot-dashed blue line). As shown, the shape of the BHMF is independent of the inclusion of UV feedback and the merger timescales used. However, the final BH masses are naturally lower when including a delay in the merger timescales as opposed to instantaneous mergers.

We now discuss the black hole mass function (BHMF) which expresses the number density of black holes as a function of their mass, the results of which at $z \simeq 6$ are shown in Fig. 4. As expected, the number density of black holes increase with decreasing BH mass as shown in the Figure. The observed BHMF at $z \sim 6$ extends from $M_{\text{bh}} \sim 10^{7-10} M_{\odot}$. Our theoretical results for all four models discussed above are in good agreement with the data within error bars as seen in the same figure. Naturally, the fiducial model (*ins1*), extending from $M_{\text{bh}} \sim 10^{4.8-8.8} M_{\odot}$, yields the upper

limit to the BHMF. Including a delay in the merger times for black holes (*tdf1*) leads to a decrease in the maximum mass attained by the black holes ($M_{\text{max}} \sim 10^8 M_{\odot}$) showing that gas brought in by merging progenitors halos has a significant contribution to the growth of these high-mass systems. On the other hand, reionization feedback alone (*ins4*) has a negligible effect on the growth of high-mass halos (as discussed in Sec. 2.3.2 above), yielding a BHMF in close agreement with the fiducial one. Finally, the model *tdf4*, including both the impact of delayed mergers and the UVB, yields results quite similar to *tdf1* and, provides the lower limit to the BHMF. We recall that our model is not aimed at (re)producing rare luminous quasars powered by very massive BH (see Valiante et al. 2016a; Pezzulli et al. 2016a, and references therein for models focused on the most massive halos and BHs) but at the bulk of the population of massive BHs. It should therefore not be surprising that the BH mass function does not extend to the highest BH masses observed.

Figure 5: The black hole occupation fraction as a function of halo mass for $z \sim 6$ (black lines), $z \sim 9$ (blue lines) and $z \sim 12$ (red lines) as marked. The solid lines show the occupation fraction for all black holes; the short-dashed and dot-dashed lines show results for Type 1 (stellar black holes) and Type 2+3 (DCBHs), respectively.

We also show the BH occupation fraction in Fig. 5. As shown, galaxies with a halo mass $M_h \gtrsim 10^{10.2}$ have an occupation fraction of 1 by $z \sim 6$. As expected, most of these are stellar black holes except DCBHs that dominate for the most massive halos. The black hole occupation fraction also shifts to progressively lower masses with increasing redshift. This is because of two reasons: first in our model, only starting leaves above $z = 13$ are seeded with black holes; the increasing number of starting leaves forming at lower redshifts are devoid of any black holes. Secondly, low mass halos continually increase in mass with decreasing redshift. We note that our results are qualitatively in good agreement with those obtained from previous works (e.g. Tanaka & Haiman 2009). Finally we stress that the enhancement of the LW seen by any halos only depends on its bias at that redshift and we have ignored the impact of clustered sources that could enhance the LW intensity seen by halos in over-dense environments. Our results must therefore be treated as a lower limit on the DCBH number density and, hence, the Type 2+3 occupation fraction.

2.3.4 The black hole-stellar mass relation

Constraints on the relation between BHs and galaxies at high redshift are scant. In general, since the only confirmed BHs at these redshifts are those powering powerful quasars, the stellar mass of the host cannot be measured (not to mention the stellar velocity dispersion or bulge mass) because the light from the quasar over-shines the host galaxy. The best estimates of the host properties for these powerful quasars are obtained through measures of the cold (molecular) gas properties in sub-mm observations, where a dynamical mass, based on the velocity dispersion of the gas and the radius of the emitting region can be measured (e.g., Venemans et al. 2016; Shao et al. 2017; Decarli et al. 2018, and references therein). For these quasars, the BH to dynamical mass is skewed to values much larger than the ratio of BH to stellar or bulge mass in the local Universe. As discussed in Volonteri & Stark (2011) there are reasons to believe that such high mass ratios should not characterize the whole BH population. Beyond the Malmquist bias causing a more frequent selection of over-massive BHs in low-mass hosts (Lauer et al. 2007; Salviander et al. 2007), only under-massive and low-accretion BHs can explain the lack of widespread AGN detections in LBGs. That BHs in low-mass galaxies are indeed expected to grow slowly and lag behind the host has now been confirmed in many numerical investigations (Dubois et al. 2015; Habouzit et al. 2017; Bower et al. 2017; Anglés-Alcázar et al. 2017). Our implementation of BH growth includes a stunted growth in low-mass galaxies and we obtain a black hole-stellar

mass relation in agreement with numerical investigations, a non-linear scaling where black holes in low-mass galaxies are “stuck” at their initial mass (Habouzit et al. 2017; Bower et al. 2017). BHs in high-mass hosts, on the other hand, can be above the $z = 0$ scaling, as shown in Fig. 6.

Figure 6: The black hole mass-stellar mass relation for $z \simeq 5$ for two models that bracket the expected range: Instantaneous mergers with/without UV feedback (*ins1* using black points and *ins4* using red points; left panel) and mergers after a merging timescale with/without UV feedback (*tdf1* using green points and *tdf4* using blue points; right panel). In both panels we show two relations derived using galaxies in the nearby Universe: $M_{\text{bh}} = 1.4M_* - 6.45$ derived for high stellar mass ellipticals and bulges and $M_{\text{bh}} = 1.05M_* - 4.1$ for moderate luminosity AGN in low-mass halos (Volonteri & Reines 2016), as marked. In each panel we also show the best-fit relation from our model for high stellar mass galaxies: $\log M_{\text{bh}} = 1.25M_* - 4.8$. As seen, our theoretical model yields a non-linear scaling such that black holes in low-mass galaxies are “stuck” at their initial mass; the BH masses of high-mass hosts, on the other hand, are strongly correlated with the stellar mass and are in excellent agreement with the results derived for lower- z high stellar mass galaxies.

Quantitatively, we find that the BH mass-stellar mass relation is strongly correlated for high stellar mass ($M_* \gtrsim 19^{9.5} M_\odot$) galaxies and is best expressed by the relation $M_{\text{bh}} = 1.25M_* - 4.8$ at $z \simeq 5$; the relation flattens below such masses. Including the impact of the UVB (*ins4*) has no impact on this relation at the bright end. However, the suppression of gas mass in low-mass halos naturally results in lower black hole masses by as much as two orders of magnitude for a given stellar mass. As noted above in Sec. 2.3.3, the inclusion of a delay in galaxy merging timescales results in a decrease in the mass of the most massive black holes (by about 0.8 dex) as seen from the right-hand panel of the same figure although it has no impact on the high-mass slope. Further, the results from *ins4* and *tdf4* are quite similar as also expected from the discussion in Sec. 2.3.3

above, yielding the lower-limit to the $M_{\text{bh}} - M_*$ relation. Finally, the best-fit relation derived for high stellar mass galaxies from our model is in excellent agreement with the relation $M_{\text{bh}} = 1.4M_* - 6.45$ derived for high stellar mass ellipticals and bulges in the nearby Universe (Volonteri & Reines 2016).

2.3.5 *LISA* and GWs from the high- z Universe

Once the theoretical model for galaxy and black hole evolution is in place, we can extend our calculations to the gravitational waves expected from the mergers of such high- z black holes. In this work, any merger falls into one of the following 3 categories: (i) *type 1 - stellar black hole mergers*: mergers of two stellar BH seeds; (ii) *type 2 - mixed mergers*: mergers of a stellar BH seed with a DCBH, and (iii) *type 3 - DCBH-DCBH mergers*: extremely rare, these are mergers of two DCBH seeds. In the last category, we also include mergers of a DCBH with a mixed merger in the past. A system with two black holes revolving around each other forms an accelerated mass quadrupole that causes emission of GWs at the expenses of orbital energy with a catastrophic outcome: as the binary emits GWs its semi-major axis shrinks (“inspiral” phase) until the two black holes merge and, after shedding any extra residual energy (“ringdown” phase), a newly born stationary BH forms. The GW signal increases in amplitude and frequency at an accelerated pace with the emission peaking at merger, i.e. roughly at the innermost stable circular orbit (ISCO). The peak frequency at the ISCO for a non-spinning black hole can be expressed as twice

$$f_{\text{ISCO}} = \frac{1}{6\sqrt{6}(2\pi)} \frac{c^3}{GM(1+z)}, \quad (2.24)$$

where M is the total binary mass. Beyond the peak the signal is exponentially damped. Massive black holes ($M_{\text{bh}} > 10^3 M_\odot$) at high redshifts emit at frequencies ($\ll 1$ Hz) much lower than the range of ground based GW detectors. To detect massive BHs through GWs much longer interferometric arms, of a million kilometers are needed, which can be only realised in space. In this section, we forecast the detection performance of the space-based European Space Agency (ESA) mission *LISA* for black hole binaries in the early Universe ($z > 4$), in the evolutionary framework presented above. *LISA* is a space-based GW laser-interferometer, proposed to be launched in 2034, that consists of three spacecrafts in an equilateral triangle constellation. The interferometer’s arms are proposed to be 2.5×10^6 km in length resulting in an optimal frequency range between ≈ 1 mHz to 0.1Hz (Fig. 7).

For each black hole merger, the optimized value of the signal to noise ratio (SNR) associated to the wave model is calculated based on the matched-filtering technique. By assuming the noise to be stationary and Gaussian with zero mean, the SNR is given by

$$\left(\frac{S}{N}\right)^2 = 4 \int_{f_{\min}}^{f_{\max}} \frac{|\tilde{h}(f)|^2}{S_n(f)} df, \quad (2.25)$$

where $|\tilde{h}(f)|$ is the amplitude of the GW signal in frequency domain and $S_n(f)$ is the noise power spectra density (PSD) function. Here $f_{\min}$ is the binary frequency when it is first observed (i.e. at $t = 0$) and $f_{\max}$ is either its frequency at the end of the mission's lifetime or at merger, whichever happens first. In the following sections, we detail the calculation of the signal ($h(f)$; Sec. 2.3.5.1) and the noise (S_n) and integration limits (Sec. 2.3.5.2). We end by presenting our results in Sec. 2.3.5.3.

2.3.5.1 The GW signal

GW detectors generally work with time-dependant scalars, $h(t)$, as their output. The scalar describes the changes in the detector after the passage of waves. In case of laser interferometers $h(t)$ represents the phase shift of the laser beam (or equivalently, the change in the detector's arm length) and can be expressed as

$$h(t) = F_{\times} h_{\times}(t) + F_{+} h_{+}(t), \quad (2.26)$$

where $h_{\times}(t)$ and $h_{+}(t)$ are the GW polarizations. Further, $F_{\times}$ and F_{+} are detector's pattern functions that depend both on the properties of the detector as well as the position of the source in the sky. After averaging the signal over the sky position and transforming it into the frequency domain we obtain

$$|\tilde{h}(f)|^2 = |\tilde{A}(f)|^2 \times |Q^2|, \quad (2.27)$$

where $\tilde{A}(f)$ is the wave amplitude in the frequency domain and Q is a geometrical factor containing information about the pattern functions. For our choice of the detector's configuration we use $Q = \frac{2}{10}\sqrt{3}$, which accounts for averaging over sky position and binary inclination (see Amaro-Seoane et al. 2017).

The calculation of $\tilde{A}(f)$ should in principle be performed with a fully relativistic (NR) numerical code. However, such calculations are numerically expensive and, in fact, only necessary for modelling the

Figure 7: LISA sky-averaged dimensionless sensitivity curve ($\sqrt{f \times S_n}$) as a function of frequency for 4 years of observation time: numerical calculation (Amaro-Seoane et al. 2017, blue line) and analytical approximation (yellow line; model A2N2 from Klein et al. 2016). The numerical curve accounts for the Galactic binary stochastic foreground noise that causes the “bump” around 10^{-3} Hz. We also overplot the dimensionless characteristic strain $h_c = 2f\tilde{h}(f)$ for the averaged (in total mass and redshift) detected BH binaries in the 3 different type of mergers (using the “light” DCBH seeds model) considered in this work. Note that the average detected binary is at an increasing redshift and has a decreasing total mass going from type 3 to type 1 mergers.

highly relativistic end of the inspiral phase and merger. The inspiral phase, where orbital velocities are much lower than the speed of light, can instead be satisfactorily reproduced with an analytical post-Newtonian (PN) formalism. These considerations inspired the so-called “phenomenological models” that give a complete analytical wave model by matching the PN and NR waveforms in the region where the PN approximation breaks down. To calculate $\tilde{\mathcal{A}}(f)$, we model the waveform with the phenomenological model *PhenomC*, which has the advantage of producing the waveform directly in frequency domain, convenient for data-analysis applications. For a detailed description of the code we refer the reader to Santamaría et al. (2010), but we provide below a summary of its functioning mechanism.

Figure 8: The signal-to-noise ratio (SNR) as a function of intrinsic total binary mass and z for 4 year mission duration. The columns from left to right show results for all mergers, type 1 mergers (mergers of two stellar BHs) and type 2 mergers (mergers of a stellar BH and a DCBH); there are no detections of type 3 black holes (mergers of two DCBHs). As marked, the upper and lower rows correspond to results for models *ins1* and *tdf4* with “light” DCBH seeds, respectively. The *LISA* detectability window is such that binaries with $\text{SNR} > 7$ have redshifts between $z = 5 - 13$ and a total mass between $M \simeq 10^{3.5-5.6} M_{\odot}$ with the exact value depending on the model and merger type. The characteristic strain for the average binary is traced in Fig. 7.

2.3.5.2 The instrumental and source noise

We numerically calculate the sky-averaged noise PSD, S_n , for *LISA* using the *LISA*-consortium simulator, that takes into account different instrumental noises as well as the stochastic background from unresolved Galactic binaries. The most notable contribution to the latter comes from Galactic white dwarf binaries that *LISA* is unable to resolve individually (e.g. Amaro-Seoane et al. 2012). The number of these sources is expected to decrease as the mission progresses and a larger number of foreground sources are

detected and removed. *LISA*'s sky-averaged sensitivity curve ($\sqrt{f \times S_n}$) adopted in this paper for the SNR calculation corresponds to a 4-year observing time and is presented (using the blue line) in Fig. 7. For convenience, the frequency limits of the integral (Eqn. 2.25) are instead calculated adopting an analytical fit to *LISA*'s PSD of the form

$$S_n(f) = \frac{20}{3} \frac{4S_{n,\text{acc}}(f) + S_{n,\text{sn}}(f) + S_{n,\text{omn}}(f)}{L^2} \left[1 + \left(\frac{f}{0.41c/(2L)} \right)^2 \right], \quad (2.28)$$

(yellow solid line in Fig. 7) from Klein et al. (2016). In the above equation, L is the detector arm length. Further, $S_{n,\text{acc}}$, $S_{n,\text{sn}}$ and $S_{n,\text{omn}}$ are the noise components due to low-frequency acceleration, shot noise and other measurement noise, respectively. Instead of performing a formal fit to the numerical curve in order to estimate the noise parameters, we adopt the following values from those reported in Klein et al. (2016):

$$\begin{aligned} S_{n,\text{acc}} &= \frac{9 \times 10^{-30}}{(2\pi f)^4} \left(1 + \frac{10^{-4}}{f} \right) [\text{m}^2\text{Hz}^{-1}], \\ S_{n,\text{sn}} &= 2.22 \times 10^{-23} [\text{m}^2\text{Hz}^{-1}], \\ S_{n,\text{omn}} &= 2.65 \times 10^{-23} [\text{m}^2\text{Hz}^{-1}], \end{aligned} \quad (2.29)$$

corresponding to a $L = 2 \times 10^6$ km arm length (model A2N2L6; Klein et al. 2016). Both the numerical and analytical curve in Fig. 7 are calculated for the current *LISA* design, that presents three spacecrafts connected by 6 links.

A visual comparison between our analytical (Eqn. 2.28) and numerical curves shows a sufficiently close match for our purposes for frequencies $> 10^{-4}$ Hz. We note that the analytical curve does not account for a stochastic background (the bump around 10^{-3} Hz). This omission allows low-mass black hole binaries ($\sim 10^3 M_\odot$), that cross the sensitivity curve near its bottom, to stay longer in the observed band. However, we will show that, even with this extra integration time, their SNR is never above the detection threshold. Finally, we over-plot the signals from representative detected merger events. The examples considered in this figure are the average (in total mass and redshift) detected binary for each of the 3 types of mergers considered in this work. Their tracks cross the *LISA* sensitivity curve around $f_{\text{min}} \approx$ a few 10^{-4} Hz where the analytical and numerical sensitivity curves match extremely well.

Model	All ₇	Type 1 ₇	Type 2 ₇	Type 3 ₇	All ₀	Type 1 ₀	Type 2 ₀	Type 3 ₀
ins1	19.8	13	6.8	0.05	300.3	288.3	11.8	0.15
tdf4	12.5	12.1	0.4	0	247.7	247.0	0.62	0.01
ins1 (heavy)	23.3	13	10.3	0.04	300.3	288.3	11.8	0.15
tdf4 (heavy)	12.5	12.1	0.4	0	247.7	247.0	0.62	0.01

Table 2.2: Total number of *LISA* detections expected for a $\text{SNR} > 7$ at $z > 4$ over a 4-year duration of the mission for the two models the bracket the upper and lower limits of the physical parameter space: *ins1* and *tdf4* for both light and heavy DCBH seeds. We show results for the three different types of BH mergers explained in Sec. 2.3.5.2: (i) *type1* - *stellar black hole mergers*: mergers of two stellar black hole seeds; (ii) *type 2* - *mixed mergers*: mergers of a stellar black hole seed with a DCBH, and (iii) *type 3* - *DCBH-DCBH mergers*: extremely rare, these are mergers of two DCBH seeds. While columns 2-5 show results for all, type 1, 2 and 3 mergers, respectively, for a $\text{SNR} > 7$, columns 6-9 show results for the same quantities without imposing any SNR cut.

2.3.5.3 *LISA* detectability of GW from the high- z Universe

To confidently claim detection, the SNR of an event must be above a critical value. Here we adopt the typical *LISA* threshold of $\text{SNR} = 7$. Each row in Fig. 8 represents the calculated SNR values for all simulated binaries in a given model as a function of their total intrinsic mass and the redshift. Starting with the fiducial model (*ins1*; top panels), BH mergers become detectable once they reach masses of $\sim 10^4 M_\odot$ at $z \lesssim 13$. As masses grow with time, these systems can reach SNR values as high as ~ 1000 for a total BH mass around $10^5 M_\odot$ below $z \sim 11$. Binaries with $\text{SNR} > 7$ appear in the redshift range $z \simeq 5 - 13$ and range in total mass between $M \simeq 10^{3.5-5.6} M_\odot$. Allowing a precise estimation of parameters such as distance, sky localisation and chirp mass, these mass and z ranges will therefore be best probed using GWs. Finally, as the black holes grow above $\sim 10^6 M_\odot$, the SNR decreases as the emitted GW signal shifts to lower frequencies, and above $\sim 10^7 M_\odot$ it goes out of the detectability window. While the results remain quite similar for the *tdf4* model, a delay in the merger timescales results in a severe reduction in the number of type 2 mergers as shown from the lower right-most panel of the same figure. Moreover, the detectability window around $10^{4-6} M_\odot$ shifts to slightly lower redshifts. We end by noting that type 2 mergers are rarer than type 1 in both models, with type 3 mergers being the rarest as expected (see Table 2.2) - this is why type 3 mergers are not plotted here.

We now discuss the yearly high- z event detection rate expected from *LISA* using a $\text{SNR} > 7$. Once the black hole merger rate density (per unit comoving volume) of events with $\text{SNR} > 7$ at a given z , $N_{\text{com}}(z)$, is obtained, we convert this into the expected number of mergers per year $d^2N/dzdt$ as (Haehnelt 1994; Arun et al. 2009)

$$\frac{d^2N}{dzdt} = 4\pi c N_{\text{com}}(z) \left(\frac{d_L(z)}{(1+z)} \right)^2 [yr^{-1}], \quad (2.30)$$

where $d_L(z)$ is the luminosity distance at z . The results of this calculation are shown in Fig. 9. As shown, the fraction of *LISA detectable* events rises with decreasing redshift from about 1/400 at $z \simeq 13$ to about 1/10 by $z \simeq 8$ to as high as 1/4 by $z \simeq 5$. As expected, most of these events are type 1 mergers. Quantitatively, by $z \simeq 4$, roughly 66% of detectable mergers are type 1 with about 32% being type 2 mergers with type 3 mergers only contributing 0.3% to the total number. While the qualitative behaviour is quite similar in the *tdf4* case, given the slower black hole mass growth, type 1 mergers significantly increase (contributing about 96% to the cumulative event rate by $z \simeq 4$) while the contribution of type

2 mergers falls to roughly 3%. Crucially, we do not find any type 3 mergers above the detection limit in this case.

Figure 9: The black hole merger event rate (per year) expected as a function of redshift for two models that bracket the physical range probed: *left panel*: *ins1* and *right panel*: *tdf4*. In each panel, the dot-dashed purple line shows the results for all mergers (without any cut in signal to noise ratio) while the solid black line shows the results for all mergers using a value of $\text{SNR} > 7$. The latter is deconstructed into the contribution from ($\text{SNR} > 7$) type1 (green dashed line), type2 “light DCBH” seed (dark blue dashed line) and type3 “light DCBH” seed (red dashed line) mergers. Further, the long-dashed light blue line and dot-dashed pink line show results for mergers with $\text{SNR} > 7$ using a heavier DCBH seed mass of $10^{4-5} M_{\odot}$ for type 2 and type 3 mergers, respectively. These results are in general agreement with those used for *LISA* calculations (e.g. Fig. 3 Klein et al. 2016).

We find that considering the “heavy DCBH seed” model leads to a slight change in these numbers for the *ins1* case: while the cumulative contribution of type 1 mergers drops slightly to 52%, this is compensated by an increase (to 47%) in the cumulative number of detectable type 2 mergers while the number of type 3 mergers remain unchanged. This heavier seed model, however, has no impact on the results from the *tdf4* model

The *total* number of detections per model and merger type for the *LISA* mission (over 4 years) are summarised in Table 2.2. The model *ins1* with “heavy DCBH seeds” yields the highest total detection number of ~ 23 events comprising of ~ 13 type 1 and ~ 10 type 2 mergers. These numbers reduce slightly to about 20 total events comprising of 13 type 1 and 7 type 2 mergers using the “light DCBH seed” model. In contrast, only a dozen events (all of type 1) are expected using model *tdf4*; as expected from the discussion above, the DCBH seed mass has no bearing on these results.

We also calculate the event rate in terms of the redshifted merged mass, $M_z = M(1+z)$, such that

$$\frac{d^2 N}{dM_z dt} = 4\pi c N_{\text{com}}(M_z) \left(\frac{dL(z)}{(1+z)} \right)^2 [yr^{-1}]. \quad (2.31)$$

The results of this calculation, presented in Fig. 10, clearly show the *LISA* detectability preference for BH masses ranging between $10^4 - 10^7 M_\odot$ for type 1 and type 2 mergers for both the *ins1* and *tdf4* models. Type 3 mergers, instead, are detectable in the mass range $10^{5-7} M_\odot$ in the “light” DCBH seed model while being undetectable in the *tdf4* model. Moving on to the “heavy DCBH seed model”, while the mass range remains unchanged for type 1 mergers, the range for both type 2 and type 3 mergers decreases: while the former range between $10^{5-7} M_\odot$ for both the *ins1* and *tdf4* models, the type 3 range lies in the very narrow range of $10^{5.5-6.5} M_\odot$ for the *ins1* case; as expected, the number of mergers of each type in each model are similar to the cumulative numbers quoted above. Practically, however, it would be difficult to distinguish between these different seeding models purely from the detected mass function given all types of merger reside in the same mass range between $10^{4-7} M_\odot$.

Finally, we provide a comparison of our expected event rates with those available in the literature: starting with heavy seeds, all previous studies used DCBH models based on “dynamical” instabilities, of the type advocated by Begelman et al. (2006), Lodato & Natarajan (2006) and Volonteri & Begelman (2010). In this study we have focused on the currently favored (at least by the first star community) “thermodynamical” models, that require a high level of LW background for the formation of seeds. As shown in this paper and in Habouzit et al. (2016a) this model results in much rarer seeds. We find that Klein et al. (2016) predict 3.9 mergers/year at $z > 4$ in the Q3-d model based on Lodato & Natarajan (2006); note that using the same seeding model as Klein et al. (2016), Bonetti et al. (2019) obtain results consistent with previous literature. Further, Sesana et al. (2007) predict 2.2 mergers/year at $z > 4$ in the model based on Begelman et al. (2006), while the LW-based model explored in this paper yields 0.0025-0.035 mergers/year at $z > 4$ as shown in Table 2.2. A comparison with Ricarte & Natarajan (2018), who also use a model based on Lodato & Natarajan (2006) and do not include a LW condition, is more difficult because they show only events with $\text{SNR} > 5$. Using this SNR cut, we find that the peak in the rates for heavy seeds is similarly broad and covers a similar redshift range when comparing our results to theirs although they predict a larger number of events: their peak rate is between 0.5 – 5 events/year while our peak rate goes from

$\sim 0.05 - 0.25$ events/year. To summarise, our lower merger rates for heavy seeds, compared to previous works, is what should be expected for a model that predicts extremely rare seeds. This is the effect of adding the condition on the LW background that previous models had not included.

For light (popIII) seeds Klein et al. (2016) predict 146.3 mergers/year at $z > 4$ (although they extrapolate to 2x this rate in their Table 1 and related text) and Sesana et al. (2007) predict 57.7 mergers/year at $z > 4$. Our model predicts between 62.0 and 75.1 mergers/year at $z > 4$ as shown in Table 2.2. When comparing to Ricarte & Natarajan (2018), again the peak in the rates for light seeds is similarly broad and covers a similar range in redshift but the value of the peak rates are lower in our case. In particular, our type 1 peak rate is 0.75 event/yr in the optimistic (*ins1*) and pessimistic (*tdf4*) models, while in Fig. 9 their peak rates lie between $\sim 5 - 20$ events/year. While our merger rate for light seeds is well within the expectations of the literature, as we made similar assumptions, the results being on the lower side are likely because of the resolution of our merger trees: for instance, Ricarte & Natarajan (2018) have a mass resolution of $\sim 10^6 M_\odot$, while Klein et al. (2016) follows Barausse (2012) who follows Volonteri et al. (2003) (whose trees are used for Sesana et al. 2007) in having a resolution dependent on the halo mass at $z=0$, reaching $10^5 M_\odot$ for halos with mass $< 4 \cdot 10^{12} M_\odot$ at $z=0$ and up to $10^7 M_\odot$ for halos with mass $10^{15} M_\odot$ at $z=0$.

Figure 10: The BH merger event rate (per year) as a function of the redshifted BH mass ($M_z = M_{\text{bh}}(1+z)$). The lines show the same models as noted in Fig. 9.

2.3.6 Conclusions and discussion

In this work, we have included the impact of BH seeding, growth and feedback, into our semi-analytic model, *Delphi*. Our model now jointly

tracks the build-up of the dark matter halo, gas, stellar and black hole masses of high- z ($z \gtrsim 5$) galaxies. We remind the reader that our star formation efficiency is the minimum between the star formation rate that equals the halo binding energy and a saturation efficiency. In the same flavour, the black hole accretion at any time-step is the minimum between the black hole accreting a certain fraction of the gas mass left-over after star formation, up to a fraction of the Eddington limit: while high-mass halos can accrete at the Eddington limit, low-mass halos follow a lower efficiency track. We explore a number of physical scenarios using this model that include: (i) two types of black hole seeds (stellar and those from direct-collapse black holes; DCBH); (ii) the impact of reionization impact; and (iii) the impact of instantaneous versus delayed galaxy mergers on the baryonic growth.

We show that, using a minimal set of mass and z -independent free parameters, our model reproduces all available data-sets for high- z galaxies and BH including the evolving (galaxy and AGN) UV LF, the SMD and the BHMF. Crucially, our model naturally yields a BH mass-stellar mass relation that is tightly coupled for high stellar mass ($M_* \gtrsim 10^{9.5} M_\odot$) halos; lower-mass halos, on the other hand, show a stunted black hole growth. Interestingly, while both reionization feedback and delayed mergers have no impact on the UV LF, the SMD is more affected by reionization feedback as compared to delayed mergers.

We then use this model, bench-marked against all available high- z data, to predict the merger event rate expected for the *LISA* mission. We find that *LISA*-detectable binaries (with $\text{SNR} > 7$) appear in the redshift range $z \simeq 5 - 13$ and range in total mass between $M \simeq 10^{3.5-5} M_\odot$. While type 1 mergers (of two stellar black holes) dominate in all the scenarios studied, type 2 mergers (merger of a stellar black hole and a DCBH) can contribute as much as 32% to the cumulative event rate by $z \sim 4$ in the fiducial (*ins1*) model. However including the impact of reionization feedback and delayed mergers (*tdf4* model) results in a lower black hole growth with type 2 mergers contributing only 3% to the cumulative event rates. Using heavier DCBH seeds results in a larger number of type 2 mergers becoming detectable with *LISA* whilst leaving the results effectively unchanged for the *tdf4* model.

Quantitatively, the model *ins1* with “heavy DCBH seeds” yields the highest total detection number of ~ 23 events comprising of ~ 13 type 1 and ~ 10 type 2 mergers. These numbers reduce slightly to about 20 total events comprising of 13 type 1 and 7 type 2 mergers using the “light DCBH seed” model. In contrast, only a dozen events (all of type 1) are expected using model *tdf4* and the DCBH seed mass has no bearing on these results.

We end with a few caveats. Firstly, given that we do not consider (the realistic case of) recoil and black hole ejection from the host halos, all black holes remain bound to halos. Secondly, the enhancement of the LW seen by any halos only depends on its bias at that redshift. This effectively means that we ignore the impact of the local environment on the LW intensity seen by any halo, and this may lead to an increase in seed formation and mergers in more biased regions. Thirdly, we have not included black hole seeds from stellar dynamical channels which have a milder metallicity dependence and should have a number density intermediate between stellar and direct-collapse black holes (e.g., Devecchi et al. 2012; Lupi et al. 2014); DCBH models that are metallicity-independent can also provide an additional channel increasing the black hole merger rate over cosmic time (Volonteri & Begelman 2010; Bonoli et al. 2014). Finally, we have used a very crude mode for reionization feedback that ignores the patchiness of reionization - in our model, halos either remain unaffected by the UVB or halos below a certain chosen virial velocity have all of their gas mass completely photo-evaporated. We aim to address each of these intricacies in detail in future works.

Acknowledgments

PD acknowledges support from the European Research Council’s starting grant ERC StG-717001 (“DELPHI”). PD and OP acknowledge support from the European Commission’s and University of Groningen’s CO-FUND Rosalind Franklin program. MV acknowledges funding from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 614199, project “BLACK”). Finally, PD thanks A. Mazumdar for his scientific inputs which have greatly added to the paper.

APPENDIX A: THE ROLE OF PARAMETERS

We now briefly discuss the interplay between the baryonic processes implemented in our model and black hole growth. Gas accretion onto the host galaxy (either from the IGM or through galaxy mergers), star formation, SN feedback and AGN feedback all contribute in regulating the amount of gas mass in the galaxies, thereby having a direct effect on black hole growth. The extent of their impact changes across the different growth phases of the black hole and its host. Increasing the star formation rate threshold f_* or the coupling factor f_*^w means that SN feedback are more effective in ejecting gas, especially in low-mass galaxies. These parameters are tuned to match the galaxy UV luminosity function, the stellar mass function, the stellar mass density and the star formation rate density. As the host halo grows and the potential well deepens, though, SN feedback becomes less effective. The value of $f_{\text{Edd}}(M_h < M_h^{\text{crit}})$ (the fraction of the Eddington rate at which the black hole is allowed to accrete), is tuned to match the faint end of the AGN UV luminosity function. Once $M_h > M_h^{\text{crit}}$ and the Eddington-limited accretion phase has started, the impact of AGN feedback grows, along with the black hole mass, with an exponential trend.

We tested the values of the parameters strictly related to black hole growth (the critical halo mass threshold for black hole accretion M_h^{crit} , the maximum fraction of gas that the black hole is allowed to accrete within a single time step $f_{\text{bh}}^{\text{ac}}$, and the coupling factor between the energy emitted by the black hole and the gas of the host galaxy f_{bh}^w) by looking at how changing them would affect the AGN (and stellar) luminosity function and the black hole mass function at $z = 6$. In Fig. 11 we vary M_h^{crit} while keeping $f_{\text{bh}}^{\text{ac}}$ and f_{bh}^w fixed, while in Figs. 12 and 13 we vary $f_{\text{bh}}^{\text{ac}}$ and f_{bh}^w respectively. These three parameters are tuned to reproduce simultaneously both the AGN UVLF and the black hole mass function. A lower value of M_h^{crit} (mod1 and mod2) means that black holes start growing at earlier times and in halos with lower final mass, leading to an increased number density of black holes with $M_{\text{bh}} > 10^4 M_\odot$. This translates into a higher normalisation of the AGN UVLF. In addition, since longer black hole activity means more gas ejected by AGN feedback, the star formation activity decreases at masses $M_h > M_h^{\text{crit}}$, affecting the stellar UVLF. Conversely, a higher M_h^{crit} (mod4 and mod5) results in lower number densities of black holes with $M_{\text{bh}} > 10^4 M_\odot$, reduced AGN feedback, and a negligible effect on the stellar UVLF. The parameter $f_{\text{bh}}^{\text{ac}}$ directly intervenes in Eq. 2.10: a higher $f_{\text{bh}}^{\text{ac}}$ - or a lower f_{bh}^w - translates into a later switch of the growth mode to sub-Eddington accretion rates, yielding a longer duty cycle. For low values of $f_{\text{bh}}^{\text{ac}}$ (mod13 and mod14) the growth of massive black holes

is reduced, with repercussions on both the AGN UVLF and on the BHMF, which show lower normalisations. At higher values (mod16 and mod17) the growth becomes self-regulated: an increase in $f_{\text{bh}}^{\text{ac}}$ means less gas mass available during the subsequent time step, compensating the original change in $f_{\text{bh}}^{\text{ac}}$. In any case, the stellar UVLF remains effectively unchanged. The parameter f_{bh}^{w} , instead, regulates the amount of gas mass ejected by AGN feedback, with a direct effect on the star-forming capabilities of the galaxy in the subsequent time step and hence visible effects on the stellar UVLF: higher coupling factors (mod9 and especially mod10) bring its normalisation (and that of the AGN UVLF) lower. The black hole mass function, on the other hand, is affected only at the high-mass end and only for very high values of f_{bh}^{w} (mod10).

Model	f_{bh}^{w}	$f_{\text{bh}}^{\text{ac}}$	$M_{\text{h},0}^{\text{crit}}$
fid	0.003	0.00055	11.25
mod1	0.003	0.00055	10
mod2	0.003	0.00055	10.7
mod3	0.003	0.00055	11.6
mod4	0.003	0.00055	12
mod5	0.0001	0.00055	11.25
mod6	0.0005	0.00055	11.25
mod7	0.001	0.00055	11.25
mod8	0.006	0.00055	11.25
mod9	0.01	0.00055	11.25
mod10	0.1	0.00055	11.25
mod11	$\mathbf{0.003} \times (\mathbf{M_{\text{h}}}/\mathbf{M_{\text{h}}^{\text{crit}}})^{1/2}$	0.00055	11.25
mod12	$\mathbf{0.003} \times (\mathbf{M_{\text{h}}}/\mathbf{M_{\text{h}}^{\text{crit}}})^{-1/2}$	0.00055	11.25
mod13	0.003	0.00005	11.25
mod14	0.003	0.0001	11.25
mod15	0.003	0.00035	11.25
mod16	0.003	0.0015	11.25
mod17	0.003	0.0055	11.25
mod18	0.003	$\mathbf{0.00055} \times (\mathbf{M_{\text{h}}}/\mathbf{M_{\text{h}}^{\text{crit}}})^{1/3}$	11.25
mod19	0.003	$\mathbf{0.00055} \times (\mathbf{M_{\text{h}}}/\mathbf{M_{\text{h}}^{\text{crit}}})^{-1/3}$	11.25

Table 2.3: We show here the values of the parameters used to produce the luminosity functions and the black hole mass functions in Figs. 11, 13, 12. Parameter values are in bold when different from the fiducial model. $M_{\text{h},0}^{\text{crit}} = \log(M_{\text{h}}^{\text{crit}}/[\Omega_m(1+z)^3 + \Omega_\Lambda]^{-0.125})$.

Figure 11: *upper panel:* the intrinsic AGN (solid lines) and stellar (dashed lines) UV luminosity functions for different values of M_h^{crit} are shown. The thick green line represents our fiducial model. *Lower panel:* same as in the upper panel for the black hole mass function. For a comparison of our results with observational UVLFs see Fig. 2.

Figure 12: *upper panel:* the intrinsic AGN (solid lines) and stellar (dashed lines) UV luminosity functions for different values of $f_{\text{bh}}^{\text{ac}}$ are shown. The thick green line represents our fiducial model. *Lower panel:* same as in the upper panel for the black hole mass function. For a comparison of our results with observational UVLFs see Fig. 2.

Figure 13: *upper panel:* the intrinsic AGN (solid lines) and stellar (dashed lines) UV luminosity functions for different values of f_{bh}^w are shown. The thick green line represents our fiducial model. *Lower panel:* same as in the upper panel for the black hole mass function. For a comparison of our results with observational UVLFs see Fig. 2.

APPENDIX B: PHENOM C

Usually, PN-NR matching techniques operate in the time domain between the 15th and the 5th orbit (the matching window) before the merger of the binary system. Within this interval, both techniques are in fact well behaved. The matching can be done by interpolating between the NP parameters, calculated at an earlier phase of the inspiral, and the NR parameters, calculated for the inspiral phase after the matching window, and subsequently transposing the hybrid waveform from the time domain to the frequency domain (see for instance Ajith et al. 2007). Conversely, *PhenomC* operates the matching directly in the frequency domain. It makes use of different NR codes to model the later phases of the inspiral (effectively covering the black hole mass ratio range $q = 1 - 4$; check Table 1 of Santamaría et al. (2010) for a recap of the characteristics of the NR codes used). On the other hand, the long, early inspiral phase is studied with a PN analytical treatment, under the assumption of a weak (non-relativistic) gravitational field. When this is the case, the orbital phase and amplitude can be calculated analytically via a Taylor expansion of the evolution equation of orbital frequency, which directly links the energy loss due to the emission of gravitational waves to the variation of the orbital frequency (eq. 3.5 of Santamaría et al. 2010).

Once an expression $\tilde{h}(f)$ for the frequency-domain waveform in both the PN and NR case is found, we can calculate the “distance” between the two signals over a frequency interval (notice that this is possible thanks to the fact that the pre-ringdown frequency of a gravitational wave is a monotonic function of time). In order to obtain the measured signal, we need to couple the theoretical signal with the noise of the detector. Since the noise is considered to be a zero-mean stochastic process, it is independent on frequency, and for this reason it is particularly convenient to perform calculations in the frequency domain. Once a way to distinguish the two wave functions - in the presence of noise - is given, it is possible to define a 1-parameter family of waveforms which interpolates between them.

3. THE MASS ASSEMBLY OF HIGH-REDSHIFT BLACK HOLES

*Olmo Piana, Pratika Dayal, Marta Volonteri, Tirthankar Roy
Choudhury*

– MNRAS, 2021, Volume 500, Issue 2, pp.2146-2158

We use the *Delphi* semi-analytic model to study the mass assembly and properties of high-redshift ($z > 4$) black holes over a wide mass range, $10^3 < M_{\text{bh}}/M_{\odot} < 10^{10}$. Our black hole growth implementation includes a critical halo mass ($M_{\text{h}}^{\text{crit}}$) below which the black hole is starved and above which it is allowed to grow either at the Eddington limit or proportionally to the gas content of the galaxy. As a consequence, after an initial growth phase dominated by black hole mergers down to $z \sim 7$ (9), supermassive black holes in $z = 4$ halo masses of $M_{\text{h}}|_{z=4} \sim 10^{11.75} (10^{13.4}) M_{\odot}$ mainly grow by gas accretion from the interstellar medium. In particular, we find that: (i) while most of the accretion occurs in the major branch for $M_{\text{h}}|_{z=4} \sim 10^{11-12} M_{\odot}$ halos, accretion in secondary branches plays a significant role in assembling the black hole mass in higher-mass halos ($M_{\text{h}}|_{z=4} \gtrsim 10^{12} M_{\odot}$); (ii) while the Eddington ratio increases with decreasing redshift for low-mass ($M_{\text{bh}} < 10^5 M_{\odot}$) black holes, it shows the opposite trend for larger masses. In addition, since the accretion rate depends on the gas mass present in the host halo, the duty cycle of the Eddington-limited accretion phase – which can last up to ≈ 650 Myr – is crucially linked to the joint assembly history of the black hole and its host halo.

3.1 Introduction

High-redshift observations have discovered that about 200 quasars, powered by accretion onto supermassive black holes (SMBH), were already in place in the first billion years of the Universe (e.g. Fan et al. 2001;

Mortlock et al. 2011; Mazzucchelli et al. 2017; Bañados et al. 2018; Matsuoka et al. 2018). The masses of these black holes can be as high as $M_{bh} \sim 1.2 \times 10^{10} M_{\odot}$ at a redshift of $z = 6.3$ (Wu et al. 2015) and $M_{bh} \sim 7.8 \times 10^8 M_{\odot}$ at $z = 7.54$ (Bañados et al. 2018). The discoveries of such early SMBHs have re-ignited the - still largely unanswered - question of how they could have grown so large so fast (for a recent review see Inayoshi et al. 2020). Growing a black hole of $10^9 M_{\odot}$ from a seed of about $100 M_{\odot}$ (at a reasonable radiative efficiency of 10%) would require near-Eddington accretion for almost 800 Myr - this corresponds to the age of the universe at $z \sim 6.8$ (Haiman & Loeb 2001; Tanaka 2014). The problem is exacerbated by the fact that high-velocity recoil kicks (of up to $500\text{--}1000 \text{ km s}^{-1}$) can significantly reduce the efficiency of black hole growth through mergers (Fitchett 1983; Favata et al. 2004; Haiman 2004; Volonteri & Rees 2006) by expelling them from the shallow potential wells of low-mass halos. While some studies have tried to alleviate this problem by allowing black holes to undergo short episodes of super-Eddington accretion within the context of a slim accretion disk model (Haiman 2004; Yoo & Miralda-Escudé 2004; Volonteri & Rees 2005; Lupi et al. 2016; Pezzulli et al. 2016b; Mayer 2019), others have suggested that pristine gas clouds could collapse isothermally and without fragmentation to yield direct-collapse black hole (DCBH) seeds with mass $M_{bh} \sim 10^4 - 10^5 M_{\odot}$ (Haehnelt & Rees 1993; Loeb & Rasio 1994; Eisenstein & Loeb 1995; Bromm & Loeb 2003; Begelman et al. 2006; Habouzit et al. 2016b; Luo et al. 2018). Alternatively, Boco et al. (2020) recently proposed that heavy black hole seeds of masses $10^4 - 10^6 M_{\odot}$ could form via merging of compact stellar remnants in young galaxies with dense gas environments.

A huge amount of theoretical effort has been dedicated to modelling black hole evolution at high-redshifts using both semi-analytic models (SAMs; e.g. Volonteri et al. 2008; Croton et al. 2006; Bower et al. 2006; Hirschmann et al. 2012; Barausse 2012; Sesana et al. 2014; Mutch et al. 2016; Lacey et al. 2016; Valiante et al. 2016b; Marshall et al. 2019) and hydrodynamic simulations (e.g. Schaye et al. 2015; Khandai et al. 2015; Di Matteo et al. 2017; Dubois et al. 2014; Volonteri et al. 2016; Vogelsberger et al. 2014; Sijacki et al. 2015; Latif et al. 2018; Habouzit et al. 2019; Huang et al. 2020). While SAMs offer speed and flexibility in exploring a larger parameter space and uncovering the properties of the black hole population as a whole, simulations offer the advantage of tracking the dynamics of the different galactic components. Nevertheless, zoom-in simulations are required both to reach the parsec-scale resolution needed to accurately model black hole accretion during the early growth phase (Lupi et al. 2019), when the internal structure of the host gas cloud plays

a major role (Beckmann et al. 2019), and to follow the (sub-) kiloparsec-scale gas inflows from the cosmic web which drive black hole growth at later evolutionary stages (Costa et al. 2014; Richardson et al. 2016); these techniques, naturally, do not allow a statistical study of the black hole population as a whole.

In general these works are tuned to reproduce the low-redshift black hole–host galaxy data. Episodes of fast black hole growth are triggered by galaxy mergers, disk instabilities or bursts of star formation, with the black hole accretion rate often assumed to be equal to the Bondi-Hoyle rate and capped at one or two times the Eddington limit. Since the efficiency of black hole growth through mergers, which are much more common at high redshift, is hampered by the black hole merging timescale (Tremmel et al. 2017) and by recoil kicks (Haiman 2004), most works agree that SMBHs at $z > 6$ mainly grow by gas accretion (see Wise 2019, for a review). Despite this progress in understanding the evolution of the high-redshift black hole population, many questions still remain open.

In this paper we want to study how the black hole physical properties and growth histories vary as a function of host halo mass across the entire population. To this end, we use the semi-analytic model *Delphi*, that has been shown to reproduce all key observables for galaxies and active galactic nuclei (AGN) at $z \gtrsim 5$ (including the AGN and stellar ultra-violet luminosity function, the stellar mass function, the stellar mass density and the black hole mass function, Dayal & Ferrara 2018) in addition to reionization observables (the electron scattering optical depth and emissivity constraints, Dayal et al. 2020). The key strengths of this model lie in that: (i) it is seeded with two types of black hole seeds (stellar and direct collapse); (ii) the black hole accretion rate is linked to the host halo mass; and (iii) it uses a minimal set of free parameters to model star formation, black holes and their associated feedback. In particular, we make use of the results from numerical simulations (Dubois et al. 2015; Bower et al. 2017; Habouzit et al. 2017) finding that black hole growth is suppressed in low-mass halos. This is because SN feedback in low-mass galaxies hinders the accumulation of gas mass around the central regions. With this picture in mind, we assume that black holes can not accrete below a critical halo mass threshold $M_{\text{h}}^{\text{crit}}$. So far, *Delphi* is the only SAM explicitly implementing such a mechanism to regulate BH growth.

The cosmological parameters used in this work correspond to Ω_{m} , Ω_{Λ} , Ω_{b} , h , n_{s} , $\sigma_8 = 0.3089, 0.6911, 0.049, 0.67, 0.96, 0.81$ (Planck Collaboration et al. 2016a). We quote all quantities in comoving units, unless stated otherwise, and express all magnitudes in the standard AB system (Oke & Gunn 1983b).

3.2 Theoretical model

In this work, we use the semi-analytic code *Delphi* (Dark matter and the emergence of galaxies in the epoch of reionization; Dayal et al. 2014b; Dayal & Ferrara 2018), which jointly tracks the assembly of the dark matter, baryonic and black hole components of high-redshift ($z \gtrsim 4$) galaxies. In brief, starting at $z = 4$, we build analytic merger trees for 550 halos, equally separated in log space in the mass range $M_h = 10^8 - 10^{13.5} M_\odot$, up to $z = 20$ in time steps of 20 Myr. Each $z = 4$ halo is assigned a number density according to the Sheth-Tormen halo mass function (HMF, ?) at $z = 4$. We then assign the same number density also to all the progenitors of the halo; we have confirmed that the resulting HMFs are in accord with the Sheth-Tormen HMF at all redshifts for $z \sim 5 - 20$.

The very first progenitors of each tree are seeded with an initial gas mass proportional to the halo mass such that $M_{\text{gi}} = (\Omega_b/\Omega_m) M_h$. Such first halos irradiated by a Lyman-Werner (LW) background of intensity (J_{LW}) larger than a critical value $J_{\text{crit}} = \alpha J_{21}$ (where $J_{21} = 10^{-21} \text{ergs}^{-1} \text{Hz}^{-1} \text{cm}^{-2} \text{sr}^{-1}$ and α is a free parameter) are assigned a DCBH seed with a mass between $10^3 - 10^4 M_\odot$. First halos (at $z > 13$) not fulfilling this criteria are instead assigned a stellar black hole (SBH) seed with a mass of $150 M_\odot$. Starting from these first progenitors and going forward in time, at each redshift step the halos and their baryonic components can grow both through the smooth accretion of dark matter and gas from the intergalactic medium (IGM) and through mergers; similarly, black holes grow through accretion from the interstellar medium (ISM) and mergers.

At a given time, the star formation efficiency is defined as the minimum between the type-II Supernovae (SNII) energy required to unbind the rest of the gas and an upper threshold f_* . In the interest of simplicity, each newly formed stellar population follows a Salpeter initial mass function (Salpeter 1955) over $0.1 - 100 M_\odot$, has a metallicity $Z = 0.05 Z_\odot$ and an age of 2 Myr. A fraction of the gas mass $M_*^{\text{gf}}(z)$ left in the halo after the star formation and SNII feedback can be accreted onto the black hole. The black hole accretion rate depends on the mass and the redshift of the host halo through a critical halo mass $M_h^{\text{crit}}(z) = 10^{11.25} [\Omega_m(1+z)^3 + \Omega_\lambda]^{-0.125}$, below which the black hole is stuck in a stunted accretion mode. Once the halo outgrows M_h^{crit} , at each time step its black hole is allowed to accrete either a fixed fraction $f_{\text{bh}}^{\text{ac}}$ of the gas mass present in the galaxy or a mass corresponding to what it would accrete if it were growing at the Eddington rate, whichever is lower. The physical reasoning behind M_h^{crit} is the following: while gas inflows towards the central black hole are disrupted by SN feedback in low-mass galaxies (see for instance Dubois

et al. 2015; Habouzit et al. 2017; Bower et al. 2017; Anglés-Alcázar et al. 2017), Lupi et al. (2019) have shown that the black hole mass growth speeds up at later times, once the host galaxy reaches a stellar mass of $M_* \approx 10^{10} M_\odot$, and SN feedback is not effective in keeping the gas away from the central regions anymore. In our work, the transitional host mass $M_h^{\text{crit}}(z)$ has been tuned to simultaneously reproduce the number densities of low-luminosity AGN at $z \geq 5$ and the black hole mass function at $z = 6$ (see Dayal et al. 2020).

The mass accreted by the black hole at any redshift step can then be written as

$$M_{\text{bh}}^{\text{ac}}(z) = \min \left[f_{\text{Edd}} M_{\text{Edd}}(z), (1 - \epsilon_r) f_{\text{bh}}^{\text{ac}} M_*^{\text{gf}}(z) \right], \quad (3.1)$$

where $M_{\text{Edd}}(z) = \dot{M}_{\text{Edd}} \times \Delta t$ is the mass accreted by the black hole in a time step assuming the Eddington growth rate, $f_{\text{bh}}^{\text{ac}} = 5.5 \times 10^{-4}$ represents the maximum fraction of the total gas mass left in the host galaxy that can be accreted onto the black hole, $\epsilon_r = 0.1$ is the radiative efficiency of the black hole and f_{Edd} is defined as

$$f_{\text{Edd}} = \begin{cases} 7.5 \times 10^{-5} & M_h(z) < M_h^{\text{crit}}(z) \\ 1 & M_h(z) \geq M_h^{\text{crit}}(z) \end{cases} \quad (3.2)$$

A fixed fraction $f_{\text{bh}}^{\text{w}} = 3 \times 10^{-3}$ of the total energy emitted by the accreting black hole is allowed to couple to the gas content of the host halo.

While we assume the mergers between dark matter halos to occur instantaneously, we explore two prescriptions for the mergers of their baryonic components: instantaneous, and merging after a delay induced by dynamical friction (Lacey & Cole 1993b). Finally, we include the impact of reionization feedback. Creating an ultra-violet (UV) heating background, reionization can photo-evaporate the gas mass from low-mass halos (Dayal & Ferrara 2018). We implement this feedback in its maximal form by suppressing the gas content of all halos with a virial velocity $V_{\text{vir}}(z) < 40 \text{ km s}^{-1}$.

To test the robustness of our approximations, we present results for two cases: (i) the *instantaneous* model (*ins1*) corresponds to galaxies and black holes merging instantaneously with their host halos, no UV feedback and DCBH seeds using a value of $\alpha = 30$; (ii) the *delayed* model (*tdf4*) includes the prescription for the dynamical friction delay, a higher value of $\alpha = 300$ for DCBH seed formation (which effectively reduces the number of DCBH seeds by a factor of ≈ 50) and the impact of the reionization feedback. These models and their free parameter values used are summarised in Table 2.1.

We now briefly discuss the interplay between the baryonic processes implemented in our model and black hole growth. Gas accretion onto the host galaxy (either from the IGM or through galaxy mergers), star formation, SN feedback and AGN feedback all contribute in regulating the amount of gas mass in the galaxies, thereby having a direct effect on black hole growth. The extent of their impact changes across the different growth phases of the black hole and its host. Increasing the star formation rate threshold f_* or the coupling factor f_*^w means that SN feedback are more effective in ejecting gas, especially in low-mass galaxies. These parameters are tuned to match the galaxy UV luminosity function, the stellar mass function, the stellar mass density and the star formation rate density. As the host halo grows and the potential well deepens, though, SN feedback becomes less effective. The value of $f_{\text{Edd}}(M_h < M_h^{\text{crit}})$ (the fraction of the Eddington rate at which the black hole is allowed to accrete), is tuned to match the faint end of the AGN UV luminosity function. Once $M_h > M_h^{\text{crit}}$ and the Eddington-limited accretion phase has started, the impact of AGN feedback grows, along with the black hole mass, with an exponential trend. We tested the values of the parameters strictly related to black hole growth (M_h^{crit} , $f_{\text{bh}}^{\text{ac}}$ and f_{bh}^w) by looking at how changing them would affect the AGN luminosity function and the black hole mass function. An increase in M_h^{crit} leads to a decrease of the amplitude of the black hole mass function at $M_{\text{bh}} \gtrsim 10^4 M_\odot$, to a lower normalisation of the AGN luminosity function and to a slightly higher number density of very luminous AGN. In fact a higher M_h^{crit} means a higher gas mass at the onset of the Eddington-limited accretion phase, which can then be sustained for a longer period of time. The parameter $f_{\text{bh}}^{\text{ac}}$ directly intervenes in Eq. 3.1: a higher $f_{\text{bh}}^{\text{ac}}$ - or a lower coupling factor f_{bh}^w - translates into a later switch of the growth mode to sub-Eddington accretion rates, yielding a longer duty cycle. Once the black hole has exited the Eddington-limited phase, an increase in $f_{\text{bh}}^{\text{ac}}$ or in the coupling factor f_{bh}^w would mean less gas mass available for the subsequent time step, compensating the change in $f_{\text{bh}}^{\text{ac}}$ and f_{bh}^w . In this regime black hole growth is self-regulated. These parameters are in turn tuned to reproduce simultaneously both the AGN UV LF and the black hole mass function.

3.3 The mass assembly of early black holes and their host halos

In this section, we start by discussing the evolution of the black hole accretion rates (expressed in terms of the Eddington fraction) as a function

Figure 1: The halo mass–black hole mass relation at $z = 5, 7$ and 9 , as marked. The upper and lower panels show results for the *ins1* and *tdf4* cases, respectively. The different symbols (green filled circles, blue filled triangles and orange empty squares) in each panels show the black hole accretion channels marked; the cyan filled diamonds represent DCBHs. In the $z = 5$ panel, the solid, dashed and dotted lines show the (average) observational results inferred for the local universe (Ferrarese 2002), for $0.1 \lesssim z \lesssim 0.3$ (Bandara et al. 2009), and theoretical predictions at $z = 5$ (Croton 2009), respectively.

of the halo mass at different redshifts in Sec. 3.3.1. We then discuss the merger- and accretion-driven assembly of the black hole and dark matter components of early galaxies in Section 3.3.2.

3.3.1 Black hole accretion rates

The relation between black holes and their host galaxies remains poorly constrained at high redshifts. Given that the only confirmed black holes at high redshifts are those powering luminous quasars, where the light from the accreting black hole over-shines that from the host galaxy, the stellar mass/bulge mass and the stellar velocity of the host cannot be measured. The best estimates of the dynamical mass for quasars are then obtained through measurements of the cold molecular gas properties in sub-millimetre observations (e.g. Venemans et al. 2016; Shao et al. 2017; Decarli et al. 2018, and references therein). Black hole masses instead are generally inferred using rest-frame ultra-violet indicators (e.g. Vestergaard 2002; McLure & Jarvis 2002).

In Fig. 1 we show the black hole mass-halo mass relation at $z = 5 - 9$ for both the *ins1* and the *tdf4* scenarios, color-coded by the black hole accretion channel. We start by discussing the results of the *ins1* model, which yields the upper limit to the black hole mass. Independently of redshift, black holes residing in low-mass halos ($M_h \lesssim 10^9 M_\odot$) do not accrete at all, since SNII feedback ejects all of the gas mass from the host galaxies. The double tail at $M_h \sim 10^{9-11} M_\odot$ is due to the presence of two types of black hole seeds: the upper part is comprised of (mixed) black holes resulting from the mergers of DCBHs and SBHs, while the lower part is comprised of SBHs. The black holes in this halo mass range show *stunted growth*, given that they accrete at $7.5 \times 10^{-5} M_{\text{Edd}}$. Interestingly, since DCBH seeds are assigned solely on the basis of halo bias (see discussion in Dayal & Ferrara 2018), they cover almost the entire halo mass range at all the redshifts studied. Once the critical halo mass threshold is reached, black holes can start accreting at the Eddington rate (*Eddington-limited* phase); this phase ends when the *gas-limited* condition $M_*^{\text{gf}}(z) f_{\text{bh}}^{\text{ac}} (1 - \epsilon_r) < M_{\text{Edd}}(z)$ is reached. This happens for the highest-mass halos with $M_h \gtrsim 10^{12} M_\odot$ at $z = 5$. Our black hole mass-halo mass relation is therefore described by a triple power-law at $z \sim 5$: flat at $M_h \lesssim M_h^{\text{crit}}$, almost vertical at $M_h^{\text{crit}} \lesssim M_h \lesssim 10^{11.6} M_\odot$ and $\log(M_{\text{bh}}/M_\odot) \propto \log(M_h/M_\odot)^{1.4}$ at $M_h \gtrsim 10^{11.6} M_\odot$. On the other hand, notice that black holes at $z \gtrsim 7$ never enter the last accretion phase. Finally, the slight increase of M_h^{crit} with decreasing redshift naturally results in more massive halos (and hence black holes) accreting at the Eddington rate with decreasing redshift.

We then look at results from the *tdf4* scenario that includes the impact of the UV background and a dynamical friction-induced time delay in mergers. As seen from the same figure, the $M_{\text{bh}}-M_h$ relation in this

scenario follows a similar slope as the *ins1* model for $M_h \gtrsim 10^{11} M_\odot$ at all the redshifts considered. However, as a result of the higher LW threshold used, the total number of DCBH black hole seeds in the *tdf4* model is lower by a factor of ≈ 50 . In addition, this model has a larger scatter (by about a factor of two) in the black hole masses at a given halo mass compared to the *ins1* model. This is in accord with previous works (e.g. Peng 2007) that find the number of merging events to anti-correlate with the width of the scatter in the black hole mass-host mass relation.

We find that for halos with $M_h \gtrsim 10^{11} M_\odot$, for both models, while the slope of our $M_{bh}-M_h$ relation is consistent with a number of other (observational and theoretical) works, its normalisation is bracketed by the observational/theoretical ranges. This is quite encouraging given the different proxies used to estimate the halo masses in low- z observational works, ranging from bulge properties (Ferrarese 2002) to gravitational lens modelling (Bandara et al. 2009), as well as theoretical model results extrapolating local relations to $z = 5$ (Croton 2009).

From an observational perspective the stellar mass is easier to measure than the halo mass, especially since telescopes like JWST will be able to estimate stellar masses also at high redshifts. Therefore, as a complement to Fig. 1, we show in Fig. 2 the $M_* - M_{bh}$ relationship as predicted from our model. A debate has been going on in the literature regarding the evolution of this relation (see the introduction of Suh et al. 2020, for a review on the current state of the discussion). In brief, some works inferred a positive evolution with redshift of the $M_{bh}-M_*$ ratio, suggesting that at earlier times black hole growth was more efficient than that of the host galaxies (see for instance Peng et al. 2006; Merloni et al. 2010; Decarli et al. 2010; Caplar et al. 2018), while other authors found no significant difference in the mass ratios of high- and low-redshift AGN, both with observational (Shields et al. 2003; Shankar et al. 2009; Salviander et al. 2015; Shen et al. 2015) and theoretical (Thomas et al. 2019) studies. To complicate things further, selection biases at high redshift might affect these measurements by favouring the detection of overmassive active black holes (Lauer et al. 2007). From Fig. 2 it appears that at $z = 5$, for galaxies with masses $M_* \gtrsim 10^{9.5} M_\odot$, our relation is ~ 1 dex higher with respect to the low-redshift measurements from Suh et al. (2020) and the relation for moderate-luminosity local AGN shown in Reines & Volonteri (2015), pointing towards a positive evolution of the relation. For the same mass range, our relation also seems to be in very good agreement with the relation inferred for AGN in old high-mass elliptical galaxies (Reines & Volonteri 2015). In conclusion, our results point to black holes hosted by massive galaxies above $10^{9.5} M_\odot$ to grow efficiently and be more massive

Figure 2: The stellar mass–black hole mass relation at $z = 5, 7$ and 9 , as marked. The upper and lower panels show results for the *ins1* and *tdf4* cases, respectively. The different symbols (green filled circles, blue filled triangles and orange empty squares) in each panels show the black hole accretion channels marked; the cyan filled diamonds represent DCBHs. In the $z = 5$ panel, we plot the observational results for local moderate-luminosity AGN (solid line) and high-mass ellipticals (long-dashed line) inferred in Reines & Volonteri (2015), the results of the *SIMBA* simulation at $z = 5$ from Thomas et al. (2019) (short-dashed line), and those observed at $z = 0 - 2$ from Suh et al. (2020) (grey shaded areas).

than black holes in galaxies with the same stellar mass at lower redshift, while black holes hosted in low-mass galaxies below $10^{9.5} M_{\odot}$ are less massive and follow a steeper relation (Volonteri & Stark 2011). The different behaviour at low and high stellar mass is simply a reflection of the behaviour at low and high halo mass.

Figure 3: The redshift evolution of the (logarithm of the) average Eddington ratio $\bar{\lambda}_{\text{Edd}}$ for the different black hole mass bins marked. Solid and dot-dashed lines show results for the *ins1* and *tdf4* scenarios, respectively.

With this picture in mind, we can now look at the evolution of the Eddington ratio, defined as $\lambda_{\text{Edd}} = M_{\text{bh}}^{\text{ac}}/M_{\text{Edd}}$, which is indicative of the black hole growth efficiency. By construction, we have that $\lambda_{\text{Edd}} = 7.5 \times 10^{-5}$ for halo masses $M_{\text{h}} \lesssim 10^{11.5} M_{\odot}$, while λ_{Edd} is generally between 0.1 and 1 for black holes in higher-mass halos. The results for the average Eddington ratios ($\bar{\lambda}_{\text{Edd}}$) as a function of redshift and for different black hole mass bins are shown in Fig. 3. Low-mass black holes ($M \sim 10^{3-4} M_{\odot}$) reside in progressively more massive halos with decreasing z , eventually starting to accrete at the Eddington rate (also see Fig. 1), leading to an increase of $\bar{\lambda}_{\text{Edd}}$ from 10^{-3} at $z = 9$ to 0.1 by $z = 4$. Moving on, approximately 90% of the black holes in the mass range $10^4 - 10^5 M_{\odot}$ are in the Eddington-limited regime at $z = 4$, with $\bar{\lambda}_{\text{Edd}} \approx 0.9$, while the rest of the black holes sits in low-mass halos and accretes at $\bar{\lambda}_{\text{Edd}} = 7.5 \times 10^{-5}$. Black holes with masses $10^{5-6} M_{\odot}$ generally accrete at $\bar{\lambda}_{\text{Edd}} = 1$ at all redshifts, as also seen from

Fig. 1. As black holes grow above $10^6 M_\odot$ and switch to the gas-limited accretion phase, the average Eddington ratio drops from $\bar{\lambda}_{\text{Edd}} = 1$ to progressively lower values (~ 0.1). This is because as M_{bh} increases, so does the Eddington mass, with the consequence that more massive black holes switch to the gas-limited phase at higher redshifts.

Compared to the *ins1* model, in the *tdf4* scenario, the largest difference is seen for the lowest mass bins ($M_{\text{bh}} = 10^{3-5} M_\odot$). We also see a slight increase in $\bar{\lambda}$ for $M_{\text{bh}} \gtrsim 10^6 M_\odot$. This is because a delay in the baryonic mass assembly in the *tdf4* scenario causes (i) the black holes to enter the Eddington-limited phase with a lower mass, and (ii) stay in this phase longer since $M_{\text{Edd}}(z) < M_*^{\text{gf}}(z) f_{\text{bh}}^{\text{ac}} (1 - \epsilon_r)$ for a longer amount of time.

3.3.2 Joint assembly of black holes and their host halos

We now study the relative contribution of mergers and accretion to the joint assembly of black holes and their host halos at $z > 4$. For both black holes and halos, the merged mass is that brought in by all the progenitors of the previous z -step; the accreted mass is the ISM and IGM mass that is accreted onto the black hole and halo, respectively. Here we explore results for 21 galaxies over three halo mass bins (at $z = 4$) each: low-mass halos ($M_{\text{h}} \sim 10^{9.9-10.1} M_\odot$), intermediate-mass halos ($M_{\text{h}} \sim 10^{11.65-11.85} M_\odot$) and high-mass halos ($M_{\text{h}} \sim 10^{13.3-13.5} M_\odot$). We start by discussing results from the *ins1* scenario. Naturally, both the *ins1* and *tdf4* scenarios show the same results for halo assembly; the dynamical friction delay only affects the baryonic component.

Low-mass halos (panel a of Fig. 4) start assembling their dark matter mass at $z \sim 14$. Most of the halo growth is driven by accretion down to $z \sim 4$, at which point mergers start contributing to the halo mass assembly in equal measure. On the other hand, their black hole assembly is dominated by stochastic mergers across the whole redshift range $z \sim 13-4$. The accretion mode is sub-dominant given the negligible accretion rates ($\sim 7.5 \times 10^{-5} M_{\text{edd}}$), and contributes less than 10% of the total black hole mass by $z \sim 4$.

Intermediate-mass halos (panel b of the same figure) naturally start assembling earlier, at $z \sim 17.5$. Accretion dominates the mass assembly until $z \sim 7$, below which point mergers start contributing equally. While mergers dominate the black hole mass build-up down to $z \sim 7$, thereafter Eddington-limited accretion takes over, dominating the black hole mass budget by ~ 2 orders of magnitude by $z = 4$. By then, more than 95% of this black hole accretion has taken place in the major branch (see Fig. 5), wherein black holes accrete in the Eddington-limited phase. The high

Figure 4: The cumulative mass assembly of the black hole and the host dark matter halo as a result of mergers (M_{bh}^{mer} and M_h^{mer}) and accretion (M_{bh}^{acc} and M_h^{acc}) as a function of redshift. We show results for $z = 4$ halo masses in the range $M_h \sim 10^{9.9-10.1} M_\odot$ (upper panel), $M_h \sim 10^{11.65-11.85} M_\odot$ (middle panel) and $M_h \sim 10^{13.3-13.5} M_\odot$ (lower panel), as marked, averaging over 21 halos in each panel. The solid and dashed lines in each panel show results for the *ins1* and *tdf4* cases, respectively; shaded regions show the $1-\sigma$ variance for the *ins1* scenario.

Figure 5: The redshift evolution of the ratio of the cumulative mass accreted by black holes in the major branch with respect to that accreted in the whole merger tree of a galaxy (averaged over 21 halos) for $z = 4$ halo masses of $M_h \sim 10^{11.65-11.85} M_\odot$ and $M_h \sim 10^{13.3-13.5} M_\odot$, as marked. The solid and dashed lines in each panel show results for the *ins1* and *tdf4* cases, respectively; shaded regions show the $1 - \sigma$ variance for the *ins1* scenario.

variance here is due to the fact that different halos, with their different assembly histories, overcome the critical halo mass for Eddington-limited accretion at different redshifts.

The dark matter mass assembly of high-mass halos is dominated by accretion until $z \sim 6$, below which mergers take over as the main growth mechanism (see also Cattaneo 2002). As for black holes, mergers dominate (and increase the mass by about three orders of magnitude) between $z \sim 17.5 - 9$; accretion dominates instead the later phases, contributing about 90% to the final mass at $z \sim 4$. It is interesting also to notice that for high-mass halos the variance in the merged components is generally lower, and significant only at $z \leq 7$, when stochastic major mergers take place. Further, more than one progenitor is able to overcome M_h^{crit} , and on average about 20 (75%) of the mass is assembled in the major branch at $z \sim 10$ (4) as shown in Fig. 5.

Figure 6: As a function of z , we show the ratio of the cumulative black hole mass to the cumulative halo mass (averaged over 21 halos) for $z = 4$ halo masses of $M_h \sim 10^{11.65-11.85} M_\odot$ and $M_h \sim 10^{13.3-13.5} M_\odot$, as marked. The upper and lower panels show results for the *ins1* and *tdf4* cases, respectively; shaded regions show the associated $1 - \sigma$ variance.

In the *tdf4* case the black hole mass assembly proceeds slower, due to fewer mergers (because of dynamical friction) that bring in lower gas masses (because of the photo-evaporation of gas from low-mass halos).

To summarise, while black holes residing in low-mass halos can grow only through (rare) mergers, those in intermediate- and high-mass halos predominantly grow through mergers only in the early phases, before switching to an accretion-dominated growth phase. As expected, mergers are less important in the *tdf4* scenario, contributing less than 10% to the total mass.

Finally, in Fig. 6 we plot the redshift evolution of the ratio of the cumulative black hole mass to the cumulative halo mass for the same halo mass bins as in Fig. 4. As shown, black holes in low-mass halos (when present) are starved throughout the redshift range $z = 14 - 4$, driving the mass ratio to a constant decrease. Intermediate-mass halos show a flat evolution at $z > 13$, when the black hole seed formation efficiency is still 100% (i.e. every newly-born halo has a black hole). At $z \sim 13 - 6$, the ratio drops as new halos form without black hole seeds. At $z \lesssim 6$, the ratio shows a steep rise as the black holes residing in the biggest progenitors enter the Eddington-limited accretion phase. This results in a rise from $M_{\text{bh}}/M_{\text{h}} \sim 10^{-7.4}$ at $z = 6$ to $M_{\text{bh}}/M_{\text{h}} \sim 10^{-4.5}$ at $z = 4$. While the evolution of the highest-mass halos is qualitatively similar to intermediate-mass halos, their black holes already enter the Eddington-limited growth channel at $z \approx 8$. Sustaining this growth for a longer time, they show a mass ratio that increases from $M_{\text{bh}}/M_{\text{h}} \sim 10^{-6.8}$ at $z = 8$ to $M_{\text{bh}}/M_{\text{h}} \sim 10^{-3.6}$ at $z = 4$. Again, intermediate-mass halos show a larger variance with respect to high-mass halos as a result of the diverse merging and accretion histories of their black holes.

3.4 Key properties of the black hole population in early galaxies

We now explore the key properties of BHs in early galaxies, specifically focusing on their occupation fraction and duty cycles in this section.

3.4.1 The black hole occupation fraction

We define the black hole occupation fraction (f_{bh}) as the fraction of halos hosting a *central* black hole. In the case of merging systems, this refers to the fraction of central galaxies hosting a black hole. We note that this specification is important only in the *tdf4* scenario; in the *ins1* case galaxies

Figure 7: The redshift evolution of the minimum halo mass for which the central black hole occupation fraction $f_{bh} = 1$ (thick lines) and $f_{bh} = 0.1$ (thin lines) for both the *ins1* (solid lines) and the *tdf4* (dashed lines) scenarios, as marked.

merge at the same time as their halos. In our model, halos are populated with SBH seeds down to $z = 13$. Lower- z halos contain a black hole if they are either seeded with a DCBH or if they gain one through mergers. The latter, and dominant, mechanism naturally implies that the minimum halo mass for which $f_{bh} = 1$ increases with decreasing redshift. Indeed, as shown in Fig. 7, we find that $f_{bh} = 0.1$ (i.e. 10% of halos contain a BH) for a minimum halo mass of $M_h \sim 10^{8.4} M_\odot$ at $z \sim 12$, which increases to $M_h \sim 10^{9.5} M_\odot$ by $z = 5$. Reaching $f_{bh} = 1$ naturally requires a higher halo mass of $M_h \sim 10^{8.9}$ ($10^{10.5}$) $M_\odot$ at $z \sim 12$ (5). As seen from the same plot, there is no sensible difference in the *ins1* and *tdf4* results for $f_{bh} = 0.1$. However, the minimum halo mass for $f_{bh} = 1$ is about 0.3 – 0.4 dex higher in the *tdf4* model as compared to *ins1* model. This is because a merging galaxy and its black hole reach the central galaxy at a later time step as compared to the instantaneous (*ins1*) scenario.

We then deconstruct f_{bh} as a function of halo mass for different black hole mass bins, the results of which are shown in Fig. 8. Firstly, as expected, independent of the redshift and model used, there is a clear trend of black holes of increasing mass being hosted in increasingly massive halos. In

Figure 8: The central black hole occupation fraction (f_{bh}) as a function of halo mass for the different black hole mass bins noted. The occupation fractions are shown using lines for the *ins1* case and as shaded areas for the *tdf4* case. The solid thin cyan line shows the cumulative black hole fraction for the *ins1* case.

terms of the most massive black holes, while by $z \simeq 9$ the most massive halos ($M_h \sim 10^{11.5} M_\odot$) host black holes as massive as $\sim 10^6 M_\odot$, this increases to $M_{bh} \sim 10^{9.5} M_\odot$ by $z = 5$ for halos with $M_h \sim 10^{13.4} M_\odot$. At each redshift the overlap between the different occupation fraction curves is indicative of the intrinsic scatter of the M_{bh} – M_h relation shown in Fig. 1. Indeed, while at $z = 9$ $M_h \sim 10^{11-12} M_\odot$ halos host black holes with $M_{bh} \sim 10^4$ – $10^6 M_\odot$, at $z = 5$ the same halo masses host black holes spanning four orders of magnitude in mass, between $10^{3.5}$ – $10^{7.5} M_\odot$. This is mainly due to the black holes residing in these halos entering the Eddington-limited accretion regime. Further, deviations from the average M_{bh} – M_h relation, which depend on the black hole mass at the time the halo reaches M_h^{crit} , are amplified by the exponential growth of black holes during the Eddington-limited growth phase.

In the *tdf4* case, halos of a given mass range typically host black holes of slightly lower masses compared to *ins1*. This is due to a combination of delayed mergers and (to a lower extent) lower gas masses, given that mergers are the dominant mode of black hole growth at $z \gtrsim 9$ in intermediate- and high-mass halos. As accretion overtakes the mass build-up at lower z , results from the *ins1* and *tdf4* models come into closer agreement.

3.4.2 Time-weighted duty cycle

We now explore the black hole duty cycle defined as the fractional lifetime a black hole spent in the stunted, Eddington-limited and gas-limited accretion modes. This is done for all the $z = 4$ halos that host a black hole, accounting for all of the progenitors along the merger tree. The time-weighted duty cycle for each accretion mode is then defined as $\tau_i = t_i/t_{\text{tot}}$, where t_i indicates the time spent in each of the three accretion channels (i) and t_{tot} is the total lifetime of the black hole.

As expected in a hierarchical growth scenario, most of the lifetime of the progenitors of the final black hole is spent in small halos, where they are subject to stunted accretion. In particular, black hole progenitors of halos with $M_h|_{z=4} < M_h^{\text{crit}}$ spend all of their lifetime in a stunted accretion mode or – if they are in SN-feedback dominated halos (which are more common among the progenitors of high-mass halos) – not accreting at all. For $M_h|_{z=4} \gtrsim 10^{11.5} M_\odot$ halos, the time spent in both the Eddington and gas-limited regimes decreases with increasing halo mass, again due to an increasing fraction of low-mass halos in their merger trees. The black holes in intermediate-mass halos spend at most 5% and 1% of their lifetime

Figure 9: Logarithm of the fraction (τ_i) of the cumulative lifetime of the black hole spent in the stunted (blue line), Eddington-limited (cyan line) and gas-limited (red line) accretion modes as a function of the $z = 4$ halo mass ($M_h|_{z=4}$). The orange line shows black holes in SNII-feedback dominated galaxies that do not accrete at all.

accreting in the Eddington- and gas-limited modes, respectively; these drop to 0.1% and 1% for the highest-mass halos.

The most important difference upon implementing dynamical friction delay and UV feedback is that the black hole lifetime spent in non-accreting halos increases by at least a factor of 3 for the highest-mass halos (and up to an order of magnitude for the lowest-mass halos) due to UV suppression of their gas mass.

Finally, the variations (up to one order of magnitude) between the duty cycles of halos of very similar final masses is driven by the assembly histories of both black holes and their host halos, which determine both the halo and black hole masses, as well as the amount of gas mass available for accretion onto the black hole.

3.4.3 Black hole predictions at $z \gtrsim 4$

We conclude our analysis by proposing a couple of testable predictions of our model. In Fig. 10 we show the resulting X-ray luminosity function (XLF) at $z = 5, 7$ and 8. Since our model provides us with the bolometric

luminosity information of each AGN, we build the luminosity function just by binning our AGN sample. We adopt the Duras et al. (2020) X-ray bolometric correction, while the obscured fraction as a function of luminosity is taken from Ueda et al. (2014). We compare our results to the $z = 5$ observational XLF from Ueda et al. (2014) (left panel), and we also compare the $7 < z < 8$ results of other SAMs and simulations taken from Fig. 3 of Amarantidis et al. (2019) to our XLF at $z = 7$ (central panel) and $z = 8$ (right panel). We use our intrinsic and obscured X-ray luminosity functions as upper and lower limit on the XLF, since for the other results the upper and lower limits correspond to the estimations for the maximum and minimum obscuration effects. We point out that our lower limit represents the median obscured XLF obtained from 1000 realisations of the obscuration model: for each realisation, in each luminosity bin, we randomly draw a subsample of the AGN populating that bin corresponding to the obscured fraction and manually set their AGN luminosity to zero. The red error bars represent the 16th and 84th percentiles of the resulting set of obscured XLFs. We can easily notice in the $z=7$ and $z=8$ panels that introducing the threshold halo mass M_h^{crit} for black hole growth delays the build-up of our XLF with respect to that of the other models, but this delay disappears at $z=5$ when comparing it to the observational results.

Together with the luminosity function, the mass function is the other main predictable statistical observable, though as of now it is still very loosely constrained at high redshift (Willott et al. 2010a). In Fig. 11 we show our $z = 5 - 7$ black hole mass function. Here we focus on the results from the *ins1* model and deconstruct it to show the contributions from all three black hole accretion channels. The general behaviour follows that expected from the hierarchical structure formation model: black holes with $M_{\text{bh}} \lesssim 10^7 M_\odot$ follow a power law with their number densities being virtually constant at $z \sim 5 - 7$. On the other hand, for $M_{\text{bh}} \gtrsim 10^7 M_\odot$, the number density of high-mass black holes show an exponential cut-off, specially at $z \sim 5 - 6$, in addition to showing a rapid evolution between $z \sim 6 - 7$. In particular, the mass bin $10^6 M_\odot \lesssim M_{\text{bh}} \lesssim 10^7 M_\odot$ experiences the greatest variation of number density in the shortest timescale, increasing by almost three orders of magnitude between $z = 7$ and $z = 6$ (i.e. in less than 300 Myr). As a consequence of the different evolution in number densities of the low- and high-mass ends of the BHMF, at $z = 5$ black holes with $M_{\text{bh}} \sim 10^{4.5} M_\odot$ and $M_{\text{bh}} \sim 10^7 M_\odot$ have very similar number densities. The exponential growth of black holes entering the Eddington-limited accretion regime (i.e. all black holes with $M_{\text{bh}} \gtrsim 10^4 M_\odot$) makes them move from one black hole mass bin to the next one in a very short time span, while the halo mass grows much slower. Therefore, the shallow slope

Figure 10: Evolution of the AGN X-ray luminosity function at $z = 5, 7, 8$. The black shaded areas represent our results, for which the upper and lower limit are set respectively by the intrinsic and the obscured XLF, where the obscured fraction as a function of luminosity is taken from Ueda et al. (2014). The red error bars represent the $1 - \sigma$ uncertainties on our lower limit. The other coloured shaded areas represent results from various models taken from Fig. 3 of Amarañtidis et al. (2019), their upper and lower limits corresponding to the minimum and maximum obscuration effect. The orange solid line shows the observational results from Ueda et al. (2014).

of the BHMF in the intermediate mass range $10^4 M_\odot < M_{\text{bh}} < 10^7 M_\odot$ at $z = 5$ is a direct effect of the critical halo mass $M_{\text{h}}^{\text{crit}}$ imposed in our model.

As noted before, we reiterate that the black hole accretion channel strongly depends on its mass: while low-mass black holes with $M_{\text{bh}} \sim 10^4 M_\odot$ accrete at very low (7.5×10^{-5}) Eddington rates, intermediate-mass black holes can accrete at the Eddington rate while the most massive black holes show a variable Eddington ratio.

We also compare our BHMF to other theoretical and observational works in the same figure. At $z = 6$, our results are in qualitative agreement with the observational BHMF built (using spectroscopically-selected quasars) by Willott et al. (2010a), as well as the theoretical results of Volonteri & Reines (2016) (based on extrapolating the observed $z = 0$ black hole–host galaxy correlations to higher redshifts). At $z = 5$, our results are in reasonable agreement with those from the *EAGLE* simulation (Rosas-Guevara et al. 2016). We are also in qualitative agreement with the results from the *Horizon-AGN* project (Volonteri et al. 2016) and the AGN synthesis model by Merloni & Heinz (2008), although these models seem to hint to a characteristic knee mass lower by ≈ 0.5 dex. Finally at $z = 5-6$, our results are within the range bracketed by the *GALFORM* (Griffin et al. 2020) semi-analytic model (higher than our BHMF by more than 1 dex at $z = 6$ for $10^6 M_\odot \lesssim M_{\text{bh}} \lesssim 10^8 M_\odot$) and the *Dragons* semi-analytic model (Qin et al. 2017), which predicts a lower number density (by about 0.5 dex for $10^6 M_\odot \lesssim M_{\text{bh}} \lesssim 10^8 M_\odot$ at $z = 5$). This mismatch is indicative of the different black hole seeding and growth mechanisms implementations. Griffin et al. (2020) assume that black hole growth proceeds proportionally to the star formation rate in the host galaxy, and that accretion episodes can occur also during galaxy mergers able to drive gas towards the central regions. Given the absence of any halo mass threshold for the growth of black holes, they start assembling at earlier times with respect to our model. We then expect their BHMF to evolve faster than ours at higher redshift. In Qin et al. (2017), instead, black holes are allowed to accrete a (variable) fraction of the hot gas present in the galaxy, while cold gas can only be accreted only during mergers; there is no such distinction between cold and hot gas accretion in our model, which might explain why we produce more massive black holes.

3.5 Conclusions and discussion

In this work, we have used the *Delphi* semi-analytic model to study the joint assembly of high-redshift ($z > 4$) black holes and their host halos.

Figure 11: Redshift evolution of the black hole mass function for the *ins1* scenario (solid cyan line). This is deconstructed to show the mass function for black holes accreting in the stunted mode (dashed blue line), Eddington-limited mode (long-dashed red line) and the gas-limited modes (green dotted line). We also show results from other theoretical and observational works at: $z=5$ from the Eagle simulations (Rosas-Guevara et al. 2016, black solid line), the DRAGONS model (Qin et al. 2017), the Horizon-AGN simulation (Volonteri et al. 2016, green shaded area) and (Merloni & Heinz 2008, purple shaded area), and at $z=6$ we show theoretical results from the GALFORM model (Griffn et al. 2020, green dashed line) and from Volonteri & Reines (2016, black dashed line) and observational results (Willott et al. 2010a, orange shaded area).

We implement a physically-motivated black hole growth mechanism that includes stunted growth in (gas-poor) low-mass halos while high-mass halos above the critical halo mass $M_h^{\text{crit}}(z)$ can accrete at the Eddington rate. After tuning the parameters of the model for the fiducial *ins1* scenario (see Table 2.1), we also consider a more realistic second case (*tdf4*) in which the mergers of the baryonic components are delayed by a dynamical friction timescale, a maximal UV feedback scenario is implemented and the LW BG threshold for DCBH seed formation is higher. We summarise here the key results.

1. First of all, as a consequence of our seeding mechanism and accretion model, in halos with $M_h > M_h^{\text{crit}}$ the black hole mass is assembled mostly via mergers until $z \sim 8$. Notably, this remains true even when the delay in mergers due to dynamical friction is added to the model. In fact, at such high redshift the mass ratios between halos are on average closer to unity, and the merging timescales are shorter. At $z < 8$ accretion takes over.
2. By $z = 4$ we find that in halos with $M_h \approx 10^{11.75} (10^{13.4}) M_\odot$ up to 95% (75%) of the final black hole mass is accreted within the major branch of the merger tree, while the rest comes from mergers with black holes hosted in secondary branches. While bigger halos have more progenitors growing above M_h^{crit} , in intermediate-mass halos only the major branch of the tree is able to fuel Eddington-limited accretion.
3. Finally, the average Eddington ratio $\bar{\lambda}_{\text{Edd}}$ seems to have a different evolution for black holes of different masses: for $M_{\text{bh}} \gtrsim 10^5 M_\odot$ we have $\bar{\lambda}_{\text{Edd}} \approx 1$ at $z > 6$, while it decreases down to ≈ 0.1 at lower redshift. Low-mass black holes with $M_{\text{bh}} < 10^5 M_\odot$, instead, show an opposite trend, increasing from $\bar{\lambda}_{\text{Edd}} \ll 1$ at high redshift to $\bar{\lambda}_{\text{Edd}} \approx 0.1 - 1$ at $z = 4$.

We tested our black hole seeding mechanism, finding that varying the initial DCBH seed mass our results remain effectively unchanged. This is in agreement with previous works (e.g. Ricarte & Natarajan 2018) that conclude that a hybrid scenario with a bimodal black hole seed mass distribution (such as ours) yields basically the same results as a light-seed one, given the low number density of DCBHs compared to SBHs.

To conclude, the reader should keep in mind a few caveats. First, we are not accounting for gravitational recoil kicks in mergers between massive black holes, which could either offset the central (merged) black hole or eject it altogether, hampering its subsequent growth (Blecha et al. 2016). Nevertheless, spin alignment between the merging black holes might reduce the effectiveness of this mechanism, especially if the merging

system is embedded in a circum-binary gaseous disk (Dotti et al. 2010). We also point out that the black hole merging timescales are still probably severely underestimated: not only we do not model the black hole binary inspiralling phase after the galaxies have merged, but we are also not accounting for the final parsec problem, which requires stellar scattering processes for allowing the black holes to coalesce together. In addition, we do not consider that in this low-mass regime the stellar and gaseous components can actually drive the dynamics of the black holes, introducing stochasticity in their orbits and possibly affecting the black hole merger rate, as shown by Pfister et al. (2019). This last mechanism might widely stretch or even prevent black holes from merging at all. We should also mention that the intensity of the Lyman-Werner background necessary to form DCBH seeds is still very poorly constrained in the literature, and values of α higher (lower) than the ones we have chosen in Table 2.1 would result in a lower (higher) number density of DCBH seeds forming at $8 < z < 13$. Finally, we manage to form black holes as massive as $10^{10} M_{\odot}$ at $z = 5$ by allowing them to accrete a fraction of the total gas mass present in the host galaxy, rather than only the gas confined in the central regions. For these reasons the black hole masses that our model *ins1* yields should be considered as an upper limit.

We will dedicate a next chapter to study how the black hole growth portrayed here affects the evolution of the stellar mass of the host galaxy.

Acknowledgements

We thank the anonymous referee for helping us improve the clarity of the paper. OP and PD acknowledge support from the European Commission’s and University of Groningen’s CO-FUND Rosalind Franklin program. PD also acknowledges support from the European Research Council’s starting grant ERC StG-717001 (“DELPHI”) and from the NWO grant 016.VIDI.189.162 (“ODIN”). TRC acknowledges the support of the Department of Atomic Energy, Government of India, under project no. 12-R&D-TFR-5.02-0700.

Data Availability

The data underlying this paper, which have been produced by our model, are available on reasonable request to the corresponding author.

4. BLACK HOLE FEEDBACK AND ITS IMPACT ON THE UV LUMINOSITY AND ON THE STELLAR MASS ASSEMBLY OF HIGH-REDSHIFT GALAXIES

*Olmo Piana, Pratika Dayal, Marta Volonteri, Tirthankar Roy
Choudhury*

– Submitted to MNRAS

We employ the Delphi semi-analytical model to study the impact of black hole growth on high-redshift galaxies, both in terms of the observed UV luminosity and of the star formation rate. To do this, firstly, we assess the contribution of AGN to the total galaxy UV luminosity as a function of stellar mass and redshift. We estimate, together with their duty cycle, the observed fraction of AGN-dominated galaxies and find that they outnumber stellar-dominated galaxies at $M_{\text{UV}} \lesssim -24$ and $z = 6$. Secondly, we study the evolution of the AGN and stellar luminosity functions (LFs), finding that it is driven both by changes in their characteristic luminosities M^* (i.e. evolution of the intrinsic brightness of galaxies) and in their normalizations ϕ^* (i.e. evolution of the number densities of galaxies), depending on the luminosity range considered. Finally, we follow the mass assembly history for three different halo mass bins, finding that the magnitude of AGN-driven outflows depends on the host halo mass. Galaxies in halos at $z = 4$ with $M_{\text{h}} \sim 10^{11.75} M_{\odot}$ are most affected and show a corresponding decrease of the star formation rate by a factor of three. Indeed, AGN feedback is most effective when the energy emitted by the accreting black hole is approximately 1% of the halo binding energy. In such cases, AGN feedback can drive outflows that are up to 100 times more energetic than SN-driven outflows.

4.1 Introduction

The UV luminosity function (UVLF) is the main observational probe to study the evolution of galaxies at high redshift, thanks to important ground-based and space-based facilities such as the HST (Hubble Space Telescope), the VLT (Very Large Telescope), Alma and Subaru telescopes. The stellar UV luminosity can be used as a proxy for the underlying star formation activity (Ouchi et al. 2005; Lee et al. 2009; McLure et al. 2009; Stark et al. 2009; Labbé et al. 2010; González et al. 2010; McLure et al. 2013; Finkelstein et al. 2015; Bouwens et al. 2015b; Livermore et al. 2017; Ishigaki et al. 2018; Ono et al. 2018) and it also seems to correlate with other physical properties such as the size of the host galaxy (Kawamata et al. 2018). Several works have put a lot of effort into studying the galaxy UVLF at different epochs, characterising its shape in the redshift range $z = 4 - 8$ (Castellano et al. 2010; McLure et al. 2010; Oesch et al. 2014; Atek et al. 2015a; Bouwens et al. 2015b; Bowler et al. 2015; McLeod et al. 2016; Livermore et al. 2017; Bouwens et al. 2017; Ishigaki et al. 2018). Nevertheless, as pointed out by Ono et al. (2018), its bright end still suffers from contamination coming from misidentified AGN and strongly-lensed sources, as well as from inaccuracies given by photometric redshift determination and low statistics. Alongside the galaxy UVLF, the high-redshift AGN UVLF has been studied to gain an understanding of the evolution of supermassive black holes and about the contribution of quasars to reionisation with contrasting results (Madau et al. 2004; Giallongo et al. 2015; Grazian et al. 2018; Dayal et al. 2019; Trebitsch et al. 2018, 2020), given that the AGN emission tracks the accretion rate of the central black holes. The study of the AGN UVLF, though, is complicated by the obscuration effects of the torus surrounding the central black hole, depending on its orientation with respect to the line of sight (Antonucci & Miller 1985; Urry & Padovani 1995), and this might result in incomplete samples. For this reason, and because of low statistics, the AGN UVLF is less precisely constrained, and it is still unclear whether the collected data are better fitted by a Schechter function or by a double power-law (Willott et al. 2010a; McGreer et al. 2013; Kashikawa et al. 2015; Jiang et al. 2016; Yang et al. 2016; Manti et al. 2017; Parsa et al. 2018; Kulkarni et al. 2019b). Being able to disentangle the UV contributions from the AGN and stellar components to the total UV luminosity would help in setting more precise constraints on the bright end of the galaxy and AGN UVLF (Ono et al. 2018).

Observations of luminous accreting supermassive black holes at $z > 7$ (Mortlock et al. 2011; Bañados et al. 2018) showed that these objects are

able to emit enormous amounts of energy, up to 10^{46} erg s $^{-1}$, supposedly exceeding the energy emitted by Supernovae (SN) and leaving an imprint on their assembly histories. It has been suggested that this activity might be the cause for the correlations between the physical properties of the black holes and those of the host galaxies observed at low redshift (see Kormendy & Ho 2013, for a review). The absolute magnitude of the bulges is found to be proportional to the logarithm of the mass of the black holes already in Kormendy & Richstone (1995). Later works confirmed this finding and unveiled even tighter correlations between black hole masses and the mass, luminosity and velocity dispersion of the bulges (Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Häring & Rix 2004; Gültekin et al. 2009; Kormendy & Ho 2013; McConnell & Ma 2013; Savorgnan & Graham 2015), finding mass ratios $M_{\text{bh}}/M_{\text{bulge}} \approx 0.002 - 0.006$ with a scatter of $\approx 0.25 - 0.3$ dex. Nevertheless, the situation at $z > 4$ is still unclear, and the validity of these relations remains unknown. It has instead been shown that numerical simulations need to incorporate AGN feedback - able to quench star formation - in order to explain the statistical properties of massive elliptical galaxies and the offset between the high-mass end of the halo mass function and that of the galaxy mass function (Silk & Rees 1998; Di Matteo et al. 2005; Croton et al. 2006; Bower et al. 2006; Sijacki et al. 2007; Vogelsberger et al. 2014; Schaye et al. 2015). The implementation of negative AGN feedback also appears to yield galaxy scaling relations which are in better agreement with observations (Dubois et al. 2013). Nevertheless, we do know that AGN feedback is not the only viable quenching mechanism, and simulations that fail to account for this might overestimate the impact of AGN feedback. For instance, other simulations (e.g. Dekel & Birnboim 2006; Bower et al. 2017) show that the transition of galaxies at $z \lesssim 3$ from cold to hot accretion mode might also play an important role in galaxy quenching, by hindering the accretion of filaments of cold gas that would provide fuel for star formation for stellar masses of $M_* \gtrsim 3 \times 10^{10} M_{\odot}$. In addition, numerical simulations and semi-analytic models often struggle in including environmental effects such as strangulation, ram-pressure stripping, tidal stripping, and harassment, which have been shown to be responsible for emptying galaxies of their gas and quenching the star formation activity at lower stellar masses and redshift (Hirschmann et al. 2014). For this reason, it is still unclear what role these mechanisms play at high redshift.

Observationally, widespread detection of AGN-driven outflows in active galaxies at $z \lesssim 3$ (Fischer et al. 2010; Feruglio et al. 2010; Greene et al. 2011; Aalto et al. 2012; Harrison et al. 2012) seemed to support the

hypothesis that AGN feedback might remove molecular gas from the disk of the host galaxies and quench the star formation activity. On the other hand, outflows at higher redshift have hardly been detected (Maiolino et al. 2012; Cicone et al. 2015), though stacking spectral analysis performed on AGN samples at $z \sim 6$ show signatures of extended outflows (Gallerani et al. 2018; Stanley et al. 2019). Some studies have tried to establish correlations between the presence of these outflows and the star formation rates of the galaxies, but the results remain inconclusive so far: thanks to Integral Field Units techniques, local spatial anti-correlation between AGN-driven outflows and the star formation rate within the galaxy has been found (e.g. Cano-Díaz et al. 2012), but whether these outflows are able to quench star formation on a galactic scale is still unclear. At the same time, signs of an enhancement of star formation activity due to increased gas pressure in outflows have also been observed (Maiolino et al. 2017), and in some cases the simultaneous presence of negative and positive feedback within the same galaxy was revealed (Cresci et al. 2015), confirming the prediction of Silk (2013). Unfortunately, these observations often come with several caveats. The determination of reliable star formation rates depends on the ability to accurately model dust extinction, or on the availability of high-quality far-infrared observations. In addition, as hinted earlier, building statistically significant samples of AGN and star-forming galaxies of similar stellar masses to compare with each other has proven to be difficult, given that precise stellar mass measurements require accurate dust and stellar population modelling, besides being able to accurately distinguish between the AGN and the stellar emissions.

Within this picture, the goal of this paper is first to assess the importance of the AGN contribution to the total UV luminosity of galaxies as a function of their total luminosity (Section 4.3), in order to provide a tool to disentangle the stellar and the AGN UV luminosity functions, and secondly to study at what galaxy mass range AGN feedback are most efficient in driving outflows, strongly affecting the star formation history of the host galaxy (Section 4.4).

The cosmological parameters used in this work are Ω_m , Ω_Λ , Ω_b , h , n_s , $\sigma_8 = 0.3089, 0.6911, 0.049, 0.67, 0.96, 0.81$ (Planck Collaboration et al. 2016a). We quote all quantities in comoving units, unless stated otherwise, and express all magnitudes in the standard AB system (Oke & Gunn 1983b).

4.2 The Theoretical model

In this work we use the semi-analytic code *Delphi* (Dark matter and the emergence of galaxies in the epoch of reionization) to study the role of star formation and black holes and their associated feedback in determining both the UV luminosities and physical properties of early galaxies (Dayal et al. 2014b, 2019). We briefly describe the model here and interested readers are referred to our earlier papers (Dayal et al. 2019; Piana et al. 2021) for full details. Our model is based on semi-analytic dark matter halo merger trees which track the accretion and merging history of 550 $z = 4$ dark matter halos with (equally separated) masses between $\log(M_h/M_\odot) = 8 - 13.5$. We use a binary merger tree algorithm (Parkinson et al. 2008) to reconstruct merger trees up to $z = 20$ in timesteps of 20 Myr. Each halo at $z = 4$ is assigned a number density according to the Sheth-Tormen halo mass function (HMF) at $z = 4$, and this number density is propagated throughout all of its merger tree so that the Sheth-Tormen HMF is matched at each redshift. At each time step, dark matter halos can grow both via smooth accretion from the intergalactic medium (IGM) and mergers with other halos. Similarly, the gas mass contained in the galaxy is replenished both through smooth accretion of gas from the IGM and through mergers; for the former, we make the reasonable assumption that smooth accretion of dark matter is accompanied by accretion of a cosmological fraction of gas mass. The starting leaves of the merger trees are seeded with black holes: in the *fiducial* model, these seeds are $10^{3-4} M_\odot$ direct-collapse black holes (DCBHs) if the virial temperature of the halo is $\gtrsim 10^4$ K and the Lyman-Werner (LW) background impinging on the halo is $30J_{21}$, where J_{21} is the LW background expressed in units of $10^{-21} \text{ erg s}^{-1} \text{ Hz}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$ (Sugimura et al. 2014). Starting leaves at $z > 13$ not meeting these criterion are instead seeded with a $150 M_\odot$ stellar black hole (SBH).

The initial gas mass can form stars with an effective efficiency f_*^{eff} . This is defined as the minimum value between the star formation efficiency whose corresponding SN-II feedback is enough to expel the rest of the gas from the host galaxy and an upper threshold value ($f_* = 2\%$). A fraction of the gas mass left after star formation and SNII gas ejection can be accreted onto the black hole. This accreted mass is given by

$$M_{\text{bh}}^{\text{ac}}(z) = \min [f_{\text{Edd}} M_{\text{Edd}}(z), f_{\text{bh}}^{\text{ac}} M_{\text{g}}^{\text{sn}}(z)], \quad (4.1)$$

where $M_{\text{Edd}} = \dot{M}_{\text{Edd}} \Delta t$ ($\Delta t = 20$ Myr) is the Eddington mass, f_{Edd} is the Eddington fraction, $f_{\text{bh}}^{\text{ac}}$ represents the maximum fraction of the total gas mass left after star formation and SNII feedback (M_*^{gf}) that can be

accreted onto the black hole, and $\epsilon_r = 0.1$ is the radiative efficiency of the black hole. The value of f_{Edd} depends on the halo potential, given that black hole growth is stunted in halos below a characteristic mass $M_h^{\text{crit}}(z) = 10^{11.25} [\Omega_m(1+z)^3 + \Omega_\lambda]^{0.125}$ (see Bower et al. 2017; Habouzit et al. 2017). In our model, f_{Edd} is a step function such that

$$f_{\text{Edd}} = \begin{cases} 7.5 \times 10^{-5} & M_h < M_h^{\text{crit}} \\ 1 & M_h \geq M_h^{\text{crit}} \end{cases} \quad (4.2)$$

The black hole mass accretion rate at any time-step is then $\dot{M}_{\text{bh}} = M_{\text{bh}}^{\text{ac}}(z)/\Delta t$. A fraction f_{bh}^{w} of the total energy emitted by the accreting black hole couples to the gas and drives outflows. The remaining gas mass, $M_{\text{bh}}^{\text{gf}}$, represents the final gas mass at the end of the time step. Our model parameters (f_* , f_{bh}^{w} , $f_{\text{bh}}^{\text{ac}}$ and f_{Edd}) are tuned to simultaneously reproduce the key observables for both star forming galaxies (such as the evolving ultra-violet LF, stellar mass function, star formation rate density) and AGN (the AGN UVLF and the black hole mass function). While we explicitly include visibility corrections for black holes as described in what follows, the star formation efficiency has been tuned to reproduce the *observed* LBG UV LF and should therefore be thought of as the dust-attenuated (observed) value. For this reason, we do not include any dust correction for star formation in this work. We summarise the values of the key model parameters in Table 4.1.

Parameter	Description	$f_{\text{ueda}}^{\text{unabs}}$	$f_{\text{merloni}}^{\text{unabs}}$
ϵ_{r}	radiative efficiency of black hole accretion	0.1	0.1
f_*	star formation efficiency threshold	0.02	0.02
f_*^{w}	fraction of SN energy that couples to the gas	0.1	0.1
f_{bh}^{w}	fraction of AGN energy that couples to the gas	0.003	0.0035
$f_{\text{bh}}^{\text{ac}}$	fraction of available gas mass that black holes can accrete	5.5×10^{-4}	10^{-4}
$f_{\text{Edd}}(M_{\text{h}} < M_{\text{h}}^{\text{crit}})$	black hole accretion rate in fraction of Eddington	7.5×10^{-5}	7.5×10^{-5}
$f_{\text{Edd}}(M_{\text{h}} > M_{\text{h}}^{\text{crit}})$	black hole accretion rate in fraction of Eddington	1	1
α	LW background threshold for DCBH formation (in units of J_{21})	30	30

Table 4.1: Here we list the free parameters of the model in column 1, we describe them in column 2, and we show their values in the cases in which we apply AGN visibility corrections from Ueda et al. (2014) and Merloni et al. (2014) in columns 3 and 4, respectively.

Each new stellar population is assumed to form following a Salpeter initial mass function (Salpeter 1955) in the mass range $0.1 - 100 M_{\odot}$, with a fixed metallicity $Z = 0.05 Z_{\odot}$ and an age of 2 Myr. Its time-dependent UV luminosity L_{*}^{UV} is computed using *STARBURST99* (Leitherer et al. 1999), and decays with increasing age of the stellar population. Hence, the total stellar luminosity of a galaxy at a specific time step is the sum of the contribution coming from the new star forming event plus that from the star forming events occurred within each of the progenitors during the previous time step, such that $L_{*,\text{tot}}^{\text{UV}}(z) = L_{*}^{\text{UV}}(z) + \sum L_{*,\text{tot}}^{\text{UV}}(z + \Delta z)$, with the sum running over all the progenitors of the galaxy. With respect to the black holes, instead, we can write their bolometric luminosity as

$$L_{\text{bh}} = \frac{\epsilon_r (1 - f_{\text{bh}}^{\text{w}}) \dot{M}_{\text{bh}} c^2}{\Delta t}. \quad (4.3)$$

We first convert this bolometric luminosity into the B-band luminosity as

$$\log \left(\frac{L_{\text{bh}}}{\nu_{\text{B}} L_{\text{bh}}^{\nu_{\text{B}}}} \right) = 0.80 - 0.067\Lambda + 0.017\Lambda^2 - 0.0023\Lambda^3, \quad (4.4)$$

where $\Lambda = \log (L_{\text{bh}}/L_{\odot}) - 12$ and $L_{\text{bh}}^{\nu_{\text{B}}}$ is expressed in solar luminosity (Marconi et al. 2004). Assuming an AGN spectral slope of $L_{\text{bh}}^{\nu} \propto \nu^{-0.44}$, we can convert the B-band luminosity to UV luminosity at 1375 Å ($L_{\text{bh}}^{\text{UV}}$). The total UV luminosity of the galaxy is then defined as

$$L_{\text{tot}}^{\text{UV}}(z) = L_{*,\text{tot}}^{\text{UV}}(z) + L_{\text{bh}}^{\text{UV}}(z). \quad (4.5)$$

Note that in this formulation we are assuming that the black hole luminosity decays instantaneously.

A certain fraction of the total AGN population is obscured by the torus and dust surrounding the central engine. While observations point to the fraction of obscured AGN evolving with the AGN luminosity, these are limited to $z \lesssim 2$ (e.g. Ueda et al. 2003; La Franca et al. 2005; Hasinger 2008; Iwasawa et al. 2012; Ueda et al. 2014). In this work, we use the obscured fractions computed in the works by Ueda et al. (2014) and Merloni et al. (2014), which are defined as functions of the X-ray luminosity of the black holes L_X computed across the energy range 2–10 keV. Defining unobscured AGN to be those with a column density $\log(N_H) < 22$, (Ueda et al. 2014) find an unobscured fraction

$$f_{\text{unabs}} = \frac{1 - \psi}{1 + \psi}, \quad (4.6)$$

Figure 1: The redshift-independent unobscured AGN fractions as a function of the X-ray luminosity inferred by Ueda et al. (2014, solid line) and Merloni et al. (2014, dashed line).

where $\psi = 0.43 [1 + \min(z, 2)]^{0.48} - 0.24 (\log L_X - 43.75)$. This function saturates at 0.008 at the low-luminosity end and at 0.73 at the high-luminosity end. On the other hand, Merloni et al. (2014) find

$$f_{\text{unobs}} = 1 - 0.56 + \frac{1}{\pi} \arctan \left(\frac{43.89 - \log L_X}{0.46} \right). \quad (4.7)$$

These unobscured fractions are shown in Fig. 1. As seen, the results from Ueda et al. (2014) yield a higher (lower) fraction of unobscured AGN as compared Merloni et al. (2014) corrections below (above) $L_X \sim 10^{43} \text{ erg s}^{-1}$. This necessitates a slight change in the AGN-associated free parameters between these models, as shown in Table 4.1, to match to the observed AGN UV LF at $z \sim 5 - 7$.

The number density of unobscured AGN in a given luminosity bin can be written as $\tilde{n}_{\text{AGN}} = f_{\text{unobs}}(L_X) \times n_{\text{AGN}}$, where n_{AGN} is the intrinsic number density of AGN populating the same luminosity bin. We randomly select a fraction $f_{\text{unobs}}(L_X)$ of AGN populating any given bin - for these we assume that all the luminosity reaches the observer; the rest of the AGN are assumed to be completely obscured in AGN UV Luminosity. This

implies that the total UV Luminosity is $L_{\text{tot}}^{\text{UV}} = L_{*}^{\text{UV}} + L_{\text{bh}}^{\text{UV}}$ for unobscured AGN and $L_{\text{tot}}^{\text{UV}} = L_{*}^{\text{UV}}$ for obscured AGN (see also Ricci et al. 2017; Dayal et al. 2019). This process introduces some fluctuations in the estimation of the luminosity function since black holes falling in the same luminosity bin do not necessarily share the same number density (especially at low masses). To solve this issue, we perform 1000 realisations of the sampling procedure and compute the median number density of unobscured AGN in each luminosity bin.

Finally, while in the *fiducial* model black holes are assumed to merge at the same time as their parent halos, we also explore a scenario (*tdf4*) where we implement (i) delayed black hole mergers due to dynamical friction according to Lacey & Cole (1993b); (ii) a maximal reionization feedback that photoevaporates the gas mass in halos with rotational velocities $v_c < 40 \text{ km s}^{-1}$; (iii) a higher threshold of Lyman-Werner background for DCBH seed formation, with $\alpha = 300J_{21}$. The results of this latter scenario are detailed in Sec. 4.4.

4.3 Contribution of AGN to the UV luminosity of early galaxies

We now study the relative contribution of star formation and black holes to the UV luminosities of early galaxies (Sec. 4.3.1) before discussing the fractional lifetime of galaxies dominated by UV luminosity from black holes (Sec. 4.3.2).

4.3.1 Contribution of star formation and black holes to the UV luminosity

We start by discussing the importance of black hole luminosity in star-forming galaxies. Distinguishing AGN from SF galaxies with the sole use of photometry has proven to be a challenge. This is true especially for galaxies with $M_{\text{UV}} \gtrsim -23.5$, where LBGs might become dominant (Ono et al. 2018; Adams et al. 2020), given that the AGN UV LF is still very uncertain, with faint-end slopes ranging from $\alpha \sim -1.3$ (Matsuoka et al. 2018; Akiyama et al. 2018) to $\alpha \sim -2$ (McGreer et al. 2018; Shin et al. 2020). Some studies have tried to impose a condition on the compactness of the source, for it to be considered an AGN (Kashikawa et al. 2015; Matsuoka et al. 2018), but this does not ensure that all and only LBGs are filtered out. For this reason, Bowler et al. (2021) opted for a distinction based on the size and morphology of the galaxies, in addition to archival spectral information,

4.3.1. Contribution of star formation and black holes to the UV luminosity

to estimate the fraction of AGN per luminosity bin for $M_{\text{UV}} > -24.5$ at $z = 4$. In addition, knowing the AGN- or SF-dominated galaxy fraction as a function of luminosity would also help in obtaining better constraints on the bright end of the high-redshift galaxy UV luminosity function by correcting for this possible contaminant citep[though other sources of uncertainty - e.g. cosmic variance - might still persist, e.g.][ono2018. We now employ our model to predict the fraction of galaxies dominated by stellar activity at $z = 5 - 7$. It is important to note that our modelled AGN UVLF extends only to $M_{\text{UV}}^{\text{bh}} \sim -26$ at $z = 5$; we are therefore unable to probe the high-luminosity end of this UV LF.

Figure 2: The $M_{\text{UV}} - M_*$ relation at $z = 5, 6$ and 7 . For each galaxy, we show the *intrinsic* UV magnitudes summing over star formation and black hole accretion (blue filled circles), star formation only (orange empty diamonds) and black hole accretion only (cyan filled triangles).

We show the *intrinsic* UV magnitudes for the contribution from star formation (M_{UV}^*), black hole accretion ($M_{\text{UV}}^{\text{bh}}$), and their sum ($M_{\text{UV}}^{\text{tot}}$) as a function of stellar mass at $z \sim 5 - 7$ in Fig. 2. Starting at $z = 5$, we see that the black hole magnitude falls off very quickly at $M_{\text{UV}}^{\text{tot}} \gtrsim -19$ where black holes accrete at $7.5 \times 10^{-5} \dot{M}_{\text{Edd}}$. Once black holes overcome this threshold and enter the Eddington-limited regime at $M_* \gtrsim 10^9 M_\odot$, their contribution to the total UV luminosity rapidly grows, spanning a wide UV magnitude range with $M_{\text{UV}}^{\text{bh}} \sim -10(-22)$ mag for $M_* \sim 10^9 (10^{10}) M_\odot$, indicative of their exponential growth. As a consequence, we see a steepening of the $M_{\text{UV}}^{\text{tot}} - M_*$ relation at $M_* \gtrsim 10^9 M_\odot$: the more massive the host galaxy, the more important the contribution of black hole activity to its intrinsic UV emission. The huge spread in $M_{\text{UV}}^{\text{bh}}$ for $10^9 M_\odot \lesssim M_* \lesssim 10^{10} M_\odot$

corresponds to an increase by up to 2 magnitudes in the scatter of the $M_{\text{UV}}^* - M_*$ relation. This scatter decreases again at the high-mass end of the relation, where AGN activity dominates the UV emission. The qualitative situation is similar at $z \sim 6$ for $M_* \sim 10^{9-10} M_\odot$ galaxies; the black hole emission outshines the stellar emission by ≈ 1.5 magnitudes at $M_* \gtrsim 10^{11} M_\odot$. Finally, at $z = 7$, the black hole emission is sub-dominant with respect to the SF component of the UV magnitude at all masses, as supermassive black holes with $M_{\text{bh}} \gtrsim 10^6 M_\odot$ have just started to form at this epoch (Piana et al. 2021).

We then study, in Fig. 3, the ratio between the observed UV luminosities of the black hole and stellar components ($L_{\text{UV}}^{\text{bh}}/L_{\text{UV}}^*$) as a function of total UV magnitude and Eddington ratio for the AGN obscuration models (Ueda et al. 2014; Merloni et al. 2014) discussed in Section 2.2. In particular, the orange line shows the average of the median UV ratio computed for 1000 realisations of the obscuration model, while the dots represent, as an example, one single realisation. At all redshifts, we find the median ratio to increase with decreasing total UV magnitudes. Starting with the obscuration fractions from the Ueda et al. (2014) model, at $z = 5$, Eddington accretion is confined to galaxies with $-23 \lesssim M_{\text{UV}}^{\text{tot}} \lesssim -19$, where halos have just grown above $M_{\text{h}}^{\text{crit}}$. At lower magnitudes of $M_{\text{UV}}^{\text{tot}} \lesssim -23$, galaxies start accreting a fraction ($\sim 10^{-4}$) of the available gas mass instead, as per our formalism. The UV ratio distribution of such galaxies is quite steady with magnitude, since in this regime black holes grow proportional to the gas mass present in the galaxy, like the stellar mass. At $z \sim 6$, most galaxies accrete at the Eddington rate at $M_{\text{UV}}^{\text{tot}} \lesssim -20$ (except a few of the brightest ones) with the black hole luminosity becoming increasingly important for brighter galaxies. At $z \sim 7$, while most galaxies continue accreting at the Eddington rate, the black hole luminosity is at most 10% of the stellar luminosity. Finally, we point out that in the case in which we apply the visibility corrections from Merloni et al. (2014) the UV ratio stabilises at a slightly lower value, due to the fact that in order to compensate for a higher fraction of unobscured AGN (see Fig. 1) we lower the maximum fraction of gas in the host galaxy that the black hole can accrete. In addition, the higher unobscured fraction (see Fig. 1) leads to more galaxies populating the high-luminosity bins, though the qualitative trends are the same as for the previous case.

In Fig. 4 we show the fraction of galaxies whose UV luminosity is dominated by the stellar emission, as a function of total UV magnitude. This allows us to perform a comparison with the results obtained by Ono et al. (2018), who measured the fraction of spectroscopically-identified galaxies (intended as the number of spectroscopically-confirmed galaxies

4.3.1. Contribution of star formation and black holes to the UV luminosity

Figure 3: The ratio between the UV luminosity from black hole accretion and star formation as a function of total UV magnitude for all of the galaxies present at $z = 5 - 7$ for the black hole correction models noted in each row. We remind the reader that while our star formation efficiency is tuned to reproduce observations (in particular the galaxy UV LF and SMF), we have corrected for AGN obscuration effects by randomly sampling the galaxies in each bin (see Sec. 2.2 for more details). The orange solid line represents the median UV ratio averaged over 1000 realisations of the obscuration effects (as explained at the bottom of Sec. 2.2) and the error bars show the average 16th and 84th percentile of the distribution. The data points show instead one random realisation, with the colour coding indicating the black hole accretion Eddington ratio.

divided by that of spectroscopically-confirmed galaxies plus AGN) in high-redshift catalogues, though with wide error bars. Our model reproduces these results at $z = 5$ and 6, if we define SF-dominated galaxies those for which $L_{\text{UV}}^{\text{sf}} > 100 L_{\text{UV}}^{\text{bh}}$. In particular, galaxies are fully dominated by SF UV emission for $M_{\text{UV}}^{\text{tot}} \gtrsim -19.5$, while AGN activity dominates the luminosity of all galaxies with $-26 \lesssim M_{\text{UV}}^{\text{tot}} \lesssim -25$ at $z = 5$ and 6. Also, the

Figure 4: The fraction of galaxies with a UV ratio $L_{\text{UV}}^{\text{sf}} > L_{\text{UV}}^{\text{bh}}$ (orange shaded region) and $L_{\text{UV}}^{\text{sf}} > 100 L_{\text{UV}}^{\text{bh}}$ (cyan shaded region) as a function of the total UV magnitude at $z = 5, 6$ and 7 ; the extent of the shaded areas represent the difference between the 84th and 16th percentile (i.e. the 1-sigma error) of the distribution in each bin for 1000 realisations of the two black hole obscuration models considered in the upper and lower rows. At $z \sim 5, 6$, the solid points show the points collected by Ono et al. (2018) with error bars showing the $1 - \sigma$ error.

magnitude bin in which the numbers of AGN-dominated and SF-dominated galaxies are roughly equal falls around $M_{\text{UV}} \sim -24$ (-22) both at $z = 6$ and $z = 5$ when the Ueda (Merloni) AGN visibility corrections are applied, indicating that there is only minimal redshift evolution of the fraction of SF-dominated galaxies between $z = 6$ and 5 . As noted, this result can be useful to estimate the expected fraction of AGN in large high-redshift photometric surveys and to correct the bright end of the galaxy UV LF for this contribution.

4.3.2 Fractional lifetime for AGN domination of UV luminosity

We now discuss the fractional lifetime that galaxies spend in the black hole UV luminosity dominated phase (defined as $L_{\text{UV}}^{\text{sf}} \lesssim 100 L_{\text{UV}}^{\text{bh}}$), the results

for which are shown in Fig. 5. For each halo at $z = 4$ with final mass $M_h > M_h^{\text{crit}}$ we walk through the merger tree, computing the fraction of time spent in the AGN-dominated phase by all of its progenitors (solid lines) and just the major branch (dashed lines). Considering the full merger tree and no dust attenuation of the black hole luminosity, galaxies with $M_h \sim 10^{13.5} M_\odot$ spend only about 1% of their cumulative lifetime in black hole UV luminosity dominated mode, since their progenitors mass function is mostly dominated by low-mass halos. On the other hand, lower-mass halos with $M_h \sim 10^{11.5} M_\odot$ spend a factor of 3 longer time in this phase, due to their lower number of progenitors and hence the fact that their total average lifetime is shorter. When visibility corrections are implemented, the difference between the high-mass and low-mass ends is smaller due to the fact that the obscured fraction increases at lower-masses. In general, dust corrections bring down the black hole UV luminosity-dominated fractional lifetime by approximately a factor of 10 at all masses, since dust obscuration effects apply only to black hole luminosity and not to the stellar luminosity. Looking at the major branches of the halos, $\log(\tau_{\text{agn}}) \sim 0.64$ (0.18) for the progenitors of halos with $M_h \approx 10^{13.5}$ ($10^{11.5}$) $M_\odot$ in the intrinsic case, since the major branch of higher-mass halos enters the Eddington-limited accretion phase at earlier times (Piana et al. 2021). When obscuration effects are applied, these numbers go down by a factor of 4 (12) for halos with $M_h \approx 10^{13.5}$ ($10^{11.5}$) $M_\odot$, as expected by the mass dependence of the obscured fraction of AGN. Finally, the variance of τ_{agn} within each mass bin is directly linked to the specific variations in the mass assembly history (especially for halos with final masses $M_h \lesssim 10^{12.3} M_\odot$).

4.3.3 Build-up of the AGN and SF UV LF

We now study the evolving AGN and LBG UV LF from $z \sim 5 - 10$ as shown in Fig. 6. The shape of the high-redshift LBG UV LF has now been extremely well-constrained observationally within the magnitude interval $-22.6 \lesssim M_{UV} \lesssim -14.5$ (Oesch et al. 2010; McLure et al. 2013; Bradley et al. 2014; Atek et al. 2015a; Bowler et al. 2015; McLeod et al. 2016; Bouwens et al. 2017; Livermore et al. 2017). However, it remains debated as to whether its evolution is primarily driven by an evolution in the underlying number density (Ouchi et al. 2009; McLure et al. 2009) or the characteristic knee luminosity (Castellano et al. 2010; Bouwens et al. 2011; Grazian et al. 2011; Bouwens et al. 2012). On the other hand, the shape of the AGN UV LF remains debated at high- z , ranging from a Schechter function (Manti et al. 2017) to a double-power law (Jiang et al. 2016; Parsa et al. 2018; Kulkarni et al. 2019b).

Figure 5: As a function of the final halo mass at $z = 4$, we show the fractional lifetime for which $L_{\text{UV}}^{\text{sf}} < 100 L_{\text{UV}}^{\text{bh}}$. The lines represent the median and the shaded areas represent the 16th and 84th percentiles of the distribution within each bin. The dashed and solid lines show results for the major branch and the full tree for the intrinsic and corrected black hole luminosities, as marked.

In Fig. 6 we show the LBG and AGN UV LF at $z \sim 5 - 10$, deconstructed into the contributions from the progenitors of the three halo mass bins discussed in the previous section. Starting with the LBG UV LF, we find that while the progenitors of $M_{\text{h}}|_{z=4} = 10^{9.9-10.1} M_{\odot}$ halos have UV magnitudes $-14.5 \lesssim M_{\text{UV}}^* \lesssim -11.5$ at $z \sim 10$, these widen to $-15.5 \lesssim M_{\text{UV}}^* \lesssim -9$ by $z \sim 5$. At all redshifts, they dominate the number density at the faintest end of the UV LF. As expected, the progenitors of $M_{\text{h}}|_{z=4} = 10^{11.65-11.85} M_{\odot}$ halos extend to brighter luminosities, with $M_{\text{UV}}^* \sim -19(-21)$ at $z \sim 10(5)$, dominating the UV LF number density in the intermediate UV range. Finally, the progenitors of halos as massive as $M_{\text{h}}|_{z=4} = 10^{13.3-13.5} M_{\odot}$ are responsible for building-up the brightest end of the UV LF. This confirms a picture wherein the evolution of the UV LF depends on the luminosity range probed (see also Dayal et al. 2013): the evolution of the bright end

of the LF is due to a fairly steady increase in the UV luminosity in the most massive LBGs. The evolution of low luminosity end, on the other hand, involves a mix of positive and negative luminosity evolution (as low-mass galaxies temporarily brighten then fade) coupled with a positive and negative density evolution as new low-mass galaxies form at each redshift while others are consumed by merging.

As for the AGN UV LF, we start by noting that its amplitude exceeds that of the LBG UV LF at $z \sim 5 - 6$ for $M_{\text{UV}} \lesssim -23$. However, as a result of the black hole-growth conditions imposed by our model (where black holes in halos below a characteristic mass can only grow at 7.5×10^{-5} of the Eddington rate), the LBG UV LF dominates at all magnitudes at $z \gtrsim 7$. We also find that the progenitors of $M_{\text{h}}|_{z=4} = 10^{9.9-10.1} M_{\odot}$ halos, which spend their lives below the halo mass threshold, do not contribute to the considered AGN UV LF, since their black holes are stuck in the low Eddington regimes and their magnitudes is $M_{\text{UV}}^{\text{bh}} \gtrsim -10$. For halos in the intermediate mass bin, as we showed in Piana et al. (2021), only the major branches of these halos grow above $M_{\text{h}}^{\text{crit}}$, and they do so at different epochs according to the details of their assembly history. This means that the progenitors of these halos enter the AGN UV LF at different times, and sit in different magnitude bins across the range $-21 \lesssim M_{\text{UV}} \lesssim -12$ ($-18.5 \lesssim M_{\text{UV}} \lesssim -11$) at $z = 5$ (6). Interestingly, we find that the AGN UVLF of the progenitors of the highest-mass bin is relatively flat at all epochs, especially for $z \lesssim 7$ and $M_{\text{UV}} \gtrsim -22$. This is because these AGN, accreting at Eddington rate, reside in halos of relatively similar masses (hence number densities: see Fig. 2) with $M_{\text{h}} \gtrsim M_{\text{h}}^{\text{crit}}$. On the other hand, the highest-luminosity end of the AGN UVLF is populated by black holes accreting proportional to the gas mass present in the host galaxy, and hence tracks the exponential fall-off of the halo mass function.

4.4 Impact of black holes on the assembly of early galaxies

We can now discuss the impact of black hole growth on the gas content (and hence star formation rates) of their hosts, the mass loading factors for SNII and black hole feedback and the impact of black hole feedback on the star formation rate-stellar mass relation.

Figure 6: The evolving UV LF from $z \sim 5 - 10$, as marked, for black hole accretion and star formation. We show how the progenitors of galaxies with $z = 4$ halo masses $M_h = 10^{9.9-10.1} M_\odot$ (dotted lines), $M_h = 10^{11.65-11.85} M_\odot$ (thin solid lines) and $M_h = 10^{13.3-13.5} M_\odot$ (dashed lines) populate the intrinsic star formation (thick solid blue line) and black hole (thick solid cyan) UV LF (blue empty histogram). We also show observational results for the AGN UVLF at $z = 5$ (McGreer et al. 2013; Parsa et al. 2018) and $z = 6$ (Willott et al. 2010a; Kashikawa et al. 2015; Jiang et al. 2016; Parsa et al. 2018; Matsuoka et al. 2018) and for the galaxy UVLF at $z = 5$ (Bouwens et al. 2015b; Ono et al. 2018) and $z = 6$ (Bouwens et al. 2007; McLure et al. 2009; Willott et al. 2013; Bouwens et al. 2015b; Finkelstein et al. 2015).

4.4.1 Black hole impact on the gas mass assembly of early galaxies

We start by tracking the gas and stellar mass assembly of early galaxies (in the same halo mass bins discussed above) to understand the impact AGN feedback, the results for which are shown in Fig. 7. We show results for both the *fiducial* model (left column) and the *tdf4* model (right column). Starting with the *fiducial* model, we see that the initial gas mass (M_{gi}) grows with time as galaxies assemble their halos and gas mass through accretion and mergers. The trend of the newly-formed stellar mass (M_*) closely follows that of the initial gas mass at all halo masses, with the star formation efficiency being capped at 2%. In turn, the trend of the ejected gas mass due to SN feedback (M_*^{ge}) depends on M_* . At early times and low masses, M_*^{ge} is of the same order of magnitude as the initial gas mass, and towards $z = 4$ it steadily decreases with increasing mass, down to $\sim 30\%$ (upper panels), 3% (middle panels) and 1.5% (bottom panels) of M_{gi} respectively for each mass bin. This is easily explicable by the fact that, as the halo potentials deepen, galaxies are able to retain more of their gas mass after SNII feedback.

With respect to black holes, as also discussed above, their growth is fully suppressed for the lowest mass bin where halo masses are less than $M_{\text{h}}^{\text{crit}}$. As a result, black holes can only accrete at 7.5×10^{-5} of the Eddington rate. Black hole accretion (in the plot shown as $M_{\text{bh}}^{\text{ac}}$) becomes relevant for $M_{\text{h}}|_{z=4} \gtrsim 10^{11.75} (10^{13.5}) M_{\odot}$ halos at $z \sim 8$ (12) when the host halos grow above $M_{\text{h}}^{\text{crit}}$ and black holes can start accreting either at the Eddington rate or a fraction of the available gas mass. The ejected gas mass due to black hole feedback ($M_{\text{bh}}^{\text{ge}}$) is proportional to the gas mass accreted by the black hole $M_{\text{bh}}^{\text{ac}}$, and it tends to decrease with decreasing redshift and the growth of the potential well. In particular, for the intermediate (highest) mass bin, black hole feedback starts dominating the instantaneous budget of the expelled mass at $z = 5$ (6), with the final $M_{\text{bh}}^{\text{ge}}/M_*^{\text{ge}}$ ratio reaching ≈ 10 (1.5) at $z = 4$. In fact, the highest mass bin is characterised by a higher fraction of low-mass progenitors which do not host any black hole activity, thus showing a lower $M_{\text{bh}}^{\text{ge}}/M_*^{\text{ge}}$ ratio at $z = 4$ with respect to the intermediate-mass bin.

We briefly touch upon the impact of black hole feedback on the star formation rates - this is discussed in more detail in Sec. 4.4.3 that follows. For intermediate-mass halos, we also see that the increasing trend of the gas mass ejected by black hole feedback affects the initial gas mass - and hence the star formation rate - at $z \lesssim 5$, with $M_{\text{bh}}^{\text{ge}} \approx 0.1$ (0.3) $\times M_{\text{gi}}$ at $z = 5$ (4). In the high-mass bin, instead, we see that $M_{\text{bh}}^{\text{ge}} \approx 0.1$ (0.05) $\times M_{\text{gi}}$

at $z = 6$ (4). The huge variance in the black hole accretion rate (and in the corresponding feedback) represents the fact that black holes residing in halos of similar final mass enter the Eddington-limited accretion phase at different times, depending on the build-up of the host halo mass. We can then define the final gas mass recovered at the end of the time step as $M_{\text{bh}}^{\text{gf}} = M_{\text{gi}} - M_{*} - M_{*}^{\text{ge}} - M_{\text{bh}}^{\text{ac}} - M_{\text{bh}}^{\text{ge}}$, which yields $M_{\text{bh}}^{\text{gf}} \sim 0$ at $z \gtrsim 11$ (16) for halos with $M_{\text{h}} \sim 10^{10}$ ($10^{11.75}$) $M_{\odot}$ and $M_{\text{bh}}^{\text{gf}} \sim M_{\text{gi}}$ at $z \lesssim 10$ (8) for halos with $M_{\text{h}} \sim 10^{11.75}$ ($10^{13.4}$).

Finally, galaxies show a delay in the mass assembly of halos in the low (intermediate) mass bin for the *tdf4* model. As shown in the same figure, low- (intermediate-) mass galaxies start assembling at $z \sim 9.5$ (17.5) in the *tdf4* model as compared to $z \sim 14.5$ (18.5) in the *fiducial* case. Nevertheless, this difference becomes negligible at lower redshift, when the mass budget is dominated by massive galaxies which are unaffected by reionization feedback and delayed mergers. In fact, the delay induced by dynamical friction increases with the host-to-satellite halo mass ratio, which means only minor mergers with the smallest halos are affected. For this reason, we expect no significant change in the statistical properties of galaxies or AGN at $z \lesssim 10$.

4.4.2 Mass loading factors for star formation and black hole feedback

We now look at the efficiency with which star formation and black hole activity drive gas outflows in Fig. 8. We define the mass loading factor for star formation $\eta_{*}(z) = \dot{M}_{*}^{\text{ge}}(z) / \sum \text{SFR}(z)$ as the ratio between the gas outflow rate due to SNII feedback and the star formation rate, while that for black holes $\eta_{\text{bh}}(z) = \dot{M}_{\text{bh}}^{\text{ge}}(z) / \sum \dot{M}_{\text{bh}}(z)$ corresponds to the ratio between the ejected and accreted mass by the black hole. The sums run over all the progenitors of a halo at a specific redshift z .

When we compute the mass loading factors across the full merger trees, we see that black hole accretion is much more efficient than star formation in driving outflows per unit of mass, and given the same halo mass bin we have that $\eta_{\text{bh}} \approx 10^{2-3} \times \eta_{*}$ at all redshifts, despite the higher coupling factor for SN feedback (we remind the reader that we set $f_{*}^{\text{w}} = 0.1$ and $f_{\text{bh}}^{\text{w}} = 0.003$). In both cases, the mass loading factor is defined by the balance between the energy emitted by the growing black hole (stellar) mass on one hand, and the continuously increasing binding energy of the host halo on the other. As a consequence, both the SN and AGN mass loading factors decrease towards lower redshift and for higher mass bins.

Figure 7: The mass accretion histories for halos of three different halo masses at $z = 4$. From top to bottom: $M_h|_{z=4} = 10^{10}, 10^{11.75}, 10^{13.4} M_\odot$, each averaged over 21 different halos of similar mass. On the left panels we follow, in the *ins1* case, the differential evolution of the gas mass present at the beginning of the time step (M_{gi}), the newly-formed stellar mass (M_*), the instantaneous gas mass ejected because of SN feedback (M_*^{ge}), the mass accreted by the black hole ($M_{\text{bh}}^{\text{ac}}$), the gas mass ejected from the galaxy because of black hole feedback ($M_{\text{bh}}^{\text{ge}}$) and the gas mass left in the host at the end of the time step ($M_{\text{bh}}^{\text{gf}}$). In the right panels we show the same quantities for the *tdf4* scenario. The shaded areas represent the standard deviation of the averaged sample within each time step.

Figure 8: Time evolution of the mass loading factors both for black holes (upper panel) and star formation (lower panel) averaged over the 21 halos for the three different halo mass ranges considered in Fig. 7. The solid lines represent the total mass loading factors obtained summing up - at each redshift - the mass components of all the progenitors of a halo; dashed lines show mass loading factors considering the major branch of the merger tree only.

4.4.2. Mass loading factors for star formation and black hole feedback

In more details, the lowest-mass bin shows a delayed formation of black holes, as we have seen in Fig. 7. At $z > 11$ these black holes do not have gas to accrete, since they reside in halos whose gas mass has been fully ejected by SN feedback. At $z \sim 11$, halos can retain some of the gas after star formation but they are still below M_h^{crit} , and $\eta_{\text{bh}} \sim 10^4$. From this point onward the trend steadily decreases with the increasing host mass, down to $10^{3.7}$ at $z = 4$. In the intermediate- (high-) mass bin, η_{bh} also shows a decreasing trend. At $z \sim 8$ (12) for intermediate- (high-) mass halos we have an exponential increase in $\dot{M}_{\text{bh}}$ due the start of the Eddington-limited accretion phase, and the steepening becomes ever more pronounced, dropping below 10^4 . The break visible at $z = 5$ (9) at values of $\eta_{\text{bh}} \sim 10^3$ ($10^{2.8}$) for the intermediate- (high-) mass bin is due to the major branches of the trees entering the self-regulated regime and growing less efficiently (i.e. with lower Eddington ratios) proportionally to the gas mass. As the potential well of the host halo continues to deepen, especially for the high-mass bin, the decreasing trend steepens again. At $z = 4$ the difference between the intermediate- and high-mass bin is of more than one order of magnitude, with $\eta_{\text{bh}} \sim 10^{2.75}$ and $\eta_{\text{bh}} \sim 10^{1.5}$ respectively.

With respect to the SN mass loading factor, instead, the difference between the intermediate- and high-mass bins is only ≈ 0.2 orders of magnitude, constant across redshift, and at $z = 4$ we have that $\eta_* \sim 1.8$ and $\eta_* \sim 1$ respectively. In fact, the star formation activity, unlike black hole activity, is dominated by the contribution from low-mass progenitors. For the same reason, the major branch plays a lesser role in determining the trend of the SN mass loading factor less than for the black hole mass loading factor. The lowest-mass bin consistently shows a higher value, with $\eta_* \sim 13$ at $z = 4$, since they have shallower potential wells.

If we take into account only the major branch of the different merger trees, the evolution of both the AGN and SN mass loading factor is straightforward, and we can write, at each redshift,

$$\eta_*(z) = \frac{\dot{M}_*^{\text{ge}}(z)}{SFR(z)} = \frac{1 - f_*^{\text{eff}}}{f_*^{\text{ej}}(z)} \sim \frac{v_c^2(z) + \epsilon_{SN}}{v_c^2(z)}, \quad (4.8)$$

where $\epsilon_{SN} = f_*^w v_s^2$ is the SN energy density (i.e. per unit of formed stellar mass) that couples to the gas and v_s is the velocity of the SN-driven winds, taken to be 611 km s^{-1} . Similarly, for the AGN mass loading factor limited to the evolution of the major branch we have

$$\eta_{\text{bh}}(z) = \frac{\dot{M}_{\text{bh}}^{\text{ge}}(z)}{\dot{M}_{\text{bh}}(z)} = \frac{v_c^2(z) + \epsilon_{bh}}{v_c^2(z)}, \quad (4.9)$$

where in this case the black hole energy density coupling to the gas (per unit of accreted mass) is defined as $\epsilon_{bh} = \epsilon_r f_{bh}^w c^2$. We can then see how in both panels the mass loading factor relative to the major branches decreases with increasing rotational velocity of the halo.

4.4.3 Impact of black hole feedback on the M_* – SFR relation

To better visualise the impact of black hole feedback on star formation, we show in Fig. 9 the distribution of our galaxies in the stellar mass-star formation rate plane in the redshift range $z = 5 - 10$; the colour coding represents the logarithm of the number density of galaxies in each bin. As expected, low-mass galaxies dominate the distribution in terms of number density. In addition, the low-mass end shows a double tail at $z \sim 8 - 10$, due to the fact that low-mass galaxies, after a first star formation burst, might eject most of their gas and quench their SF activity. At $z \sim 5$, galaxies with $M_* \sim 10^{5-9} M_\odot$ follow the relation $\log(\text{SFR}) = \log(M_*/M_\odot) - 8.75$. The y -intercept of this relation increases with increasing redshift, and at $z = 10$, the star formation rates are a factor of 3 higher than at $z = 5$ for a given stellar mass. Galaxies with $M_* \gtrsim 10^9 M_\odot$ at $z \sim 5$ exhibit a decrease in the star formation rate compared to this relation. This is because these galaxies are hosts to active black hole which eject considerable amounts of gas mass. We show this by plotting, underneath each panel, the logarithm of the average ratio between the black hole feedback energy coupling to the gas and the host halo binding energy, as a function of stellar mass. Very low-mass galaxies with $M_* \lesssim 10^6 M_\odot$ host black holes which are not accreting at all and whose growth is fully suppressed, due to the fact that the gas in these galaxies has been completely ejected by SN feedback. When the energy released by the accreting black hole reaches $\approx 10^{-2.5}$ times the binding energy of the gas the host halo, at $z \sim 6$, the galaxy starts straying from the sequence. At $M_* \gtrsim 10^{10.5} M_\odot$, on the other hand, the binding energy of the halos is high enough to retain the gas within their potential well, the E_{bh}/E_h ratio decreases down to $\sim 10^{-3}$, and the star formation rates can grow up to $\sim 10^3 M_\odot \text{ yr}^{-1}$. Our results suggest that AGN feedback, at least at $z > 6$, is more effective for the intermediate-mass range, rather than for the high-mass range, implying that the quenching of the highest-mass galaxies occurs either at lower redshift or through different mechanisms. Finally, we point out that our results at $z = 6$ are in good agreement with the observed M_* -SFR relation inferred in Khusanova et al. (2020).

Figure 9: Density plot of the M_* -SFR relation at $z = 5 - 10$, as marked. Underneath each panel we plot the logarithm of the average black hole accretion energy-to-halo binding energy ratio as a function of stellar mass. The solid black line represents our relation at $z = 5$ and follows the relation $\log(\text{SFR}) = \log(M_*/M_\odot) - 8.75$, while the blue dots are observational data taken at $5.6 < z < 6.6$ (Khusanova et al. 2020).

4.5 Conclusions and discussion

In this work we use the *Delphi* semi-analytic model of galaxy formation and evolution to address the impact of black hole activity on the host galaxies

at $z > 4$, following the mass and luminosity build-up of the AGN and stellar components across cosmic time. In order to allow a comparison with observation, we implement the AGN visibility corrections taken from Ueda et al. (2014) and Merloni et al. (2014). In the first half of the paper, we discuss the contribution of the AGN emission to the total UV luminosity of the galaxy, as well as the build-up of the galaxy and AGN UV luminosity functions, while in the second part we show the impact that AGN feedback has on the mass assembly history of the galaxy.

- We find that the observed fraction of galaxies for which the AGN UV luminosity outshines the stellar component by a factor of 100 is dominant for magnitudes $M_{\text{UV}}^{\text{tot}} \lesssim -24$ (-23) at $z = 6$ when we employ the Ueda's (Merloni's) correction model, and that this magnitude becomes fainter with time. Our findings are in good agreement with the galaxy fraction observationally inferred in Ono et al. (2018), though the uncertainties on the obscured fraction of AGN, which is sensitive on the sample of high-redshift AGN that has been used to estimate it, prevent us from reaching unambiguous results.
- The fractional lifetime that the major branches of our merger trees spend in the AGN mode is $\approx 63\%$ for $M_{\text{h}}|_{z=4} \sim 10^{13.5} M_{\odot}$, decreasing with decreasing halo mass down to $\approx 18\%$ for halos with $M_{\text{h}}|_{z=4} \sim 10^{11.75} M_{\odot}$, though the observed duty cycle, after obscuration effects, is respectively of $\approx 16\%$ and 1% .
- We study the build-up of the AGN and stellar UV luminosity functions, finding that the stellar UV LF of the progenitors of low-mass galaxies (driving the evolution of the faint end of the total UVLF) evolves towards brighter luminosities only down to $z \sim 8$, while that of high-mass galaxies (populating the bright end of the total UVLF) does not stop evolving across the redshift range considered. This is an indication of how the galaxy UVLF evolution is driven by actual brightness evolution at the high-luminosity end. An evolution towards fainter luminosities, due to the action of SN feedback in low-mass galaxies, is also found.
- According to our model, SN-driven outflows dominate the instantaneous gas outflow rate down to $z \sim 5$ (6) for galaxies in halos with $M_{\text{h}}|_{z=4} \sim 10^{11.4}$ ($10^{13.5}$) $M_{\odot}$. After then, AGN-driven mass outflows take over.
- We compute the mass loading factor for SN- and AGN- driven outflows, finding higher values for lower-mass galaxies in both cases, with $\log(\eta_{\text{bh}}) \sim 10^{3.7}$, $10^{2.8}$ and $10^{1.5}$ respectively for the low,

intermediate and high mass bin at $z = 5$. Similarly, for SN feedback, we find respectively $\log(\eta_{\text{sn}}) \sim 13, 1.8$ and 1 .

- We find that the impact of AGN feedback is greater for galaxies in intermediate-mass halos with $M_{\text{h}}|_{z=4} \sim 10^{11.75} M_{\odot}$ than at higher masses: a deeper potential well and hence a higher halo binding energy can retain part of the gas in outflows. In particular, when the energy emitted by the accreting black hole is approximately 1% of the halo binding energy, we see a corresponding decrease in the star formation rate by a factor of 3. This occurs in galaxies of mass $\sim 10^{9.5} M_{\odot}$ at $z = 5$.

Acknowledgments

OP and PD acknowledge support from the European Commission's and University of Groningen's CO-FUND Rosalind Franklin program. PD also acknowledges support from the European Research Council's starting grant ERC StG-717001 ("DELPHI") and from the NWO grant 016.VIDI.189.162 ("ODIN").

Data Availability

The data underlying this paper, which have been produced by our model, are available on reasonable request to the corresponding author.

5. CONCLUSIONS AND FUTURE PROSPECTS

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

– XXX

In the work presented in this thesis we have applied our semi-analytical model *Delphi* for galaxy and black hole formation and evolution to the study of the physical properties of the high-redshift black hole population and their hosts. Tuned to reproduce the main statistical observables of the high-redshift Universe (the galaxy and black hole UV luminosity functions, the stellar and black hole mass functions, the stellar mass density, the UV luminosity density), the model allows us to follow the mass assembly history of the dark matter halos, their galaxies, and the central supermassive black holes from $z = 20$ to $z = 4$, by building the merger trees of 550 halos with masses of $M_h = 10^8 - 10^{13.5} M_\odot$ at $z = 4$. Our model implements a double black hole seeding mechanism, including both stellar and direct-collapse black holes. We use a physically-motivated growth recipe that assumes that in halos below a critical mass $M_h^{\text{crit}}(z)$ gas accretion onto the black hole is stunted, hindered by the effect of Supernovae feedback which eject gas from the host galaxy (Bower et al. 2017). Above this critical halo mass, the black hole grows according to a minimal condition, accreting at each time step the minimum between the mass it would accrete if it grew at the Eddington rate and a fixed fraction of the total gas present in the host galaxy. In order to assess the solidity of the model we implement different scenarios and compare the corresponding results. When reionization feedback is taken into account, the mass assembly history of low-mass galaxies is slightly delayed, but this does not leave an imprint in later evolutionary phases. Implementing delay in galaxy mergers due to dynamical friction has a similar effect, since the delay is higher for lower-mass halos. Finally, increasing the necessary flux of Lyman-Werner background radiation needed for direct-collapse black

hole seed formation does not affect the lower- z black hole population, since stellar black holes vastly dominate the density of the seeds.

We will recap here the main conclusions and results obtained from our model in each of the the chapters included in this thesis.

Chapter 2

In the first chapter (Dayal et al. 2019) we study the gravitational wave signals that black hole mergers would generate, in order to predict the detection rate of the forthcoming space-based interferometer *LISA*. In the first part of the paper, we show that our model *Delphi* is able to successfully reproduce the observations available at $z > 4$ of the galaxy and AGN UV luminosity functions, the LBG stellar mass density and the black hole mass function. In the second part, we compute the black hole merger rate as expected to be observed by *LISA*, assuming a minimum signal-to-noise ratio $\text{SNR} = 7$. The detected binary systems appear in the redshift range $z = 5 - 13$ and total mass interval $M_{\text{tot}} = 10^{3.5} - 10^5 M_{\odot}$, and are dominated by mergers between stellar mass black holes. The expected total number of detections during the whole duration of the mission is ~ 23 (12) in the *ins1* (*tdf4*) scenario.

Chapter 3

In the second chapter (Piana et al. 2021) we follow the mass assembly history of the high- z black holes and their host dark matter halos, to understand what mechanisms drive the evolution of the black hole population. We find that mergers dominate black hole growth down to $z \sim 8$ in both the *ins1* and *tdf4* cases (with this limiting redshift being higher for higher-mass halos), while gas accretion takes over at lower redshift. By $z = 4$, $\sim 95\%$ (75%) of the final black hole mass is accreted within the major branch of the merger tree for halos with final masses $M_{\text{h}} \approx 10^{11.75}$ (13.4) $M_{\odot}$. Finally, we study the behaviour of the average Eddington ratio for black holes in different mass bins. We find an increasing trend for black holes with final masses $M_{\text{bh}} < 10^5 M_{\odot}$, with $\bar{\lambda}_{\text{Edd}} \ll 1$ at high redshift and growing up to $\sim 0.1 - 1$ by $z = 4$. On the other hand, high-mass black holes with $M_{\text{bh}} > 10^5 M_{\odot}$ show a decreasing trend from $\bar{\lambda}_{\text{Edd}} = 1$ at $z \gtrsim 6$ and decreasing to ~ 0.1 by $z = 4$.

Chapter 4

In the third and final chapter, at present under review, we discuss the impact of black hole growth on the host galaxies. From the observational point of view, we study the contribution to the total UV luminosity given by the AGN and stellar components of high-redshift galaxies, finding that AGN-dominated galaxies represent the majority for $M_{\text{UV}} \lesssim -24$ at $z = 6$, with their fraction increasing towards brighter luminosities and lower redshift. Their intrinsic (corrected for obscuration) duty cycles range from $\sim 18\%$ (1%) to $\sim 63\%$ (18%) depending on the host halo mass. We show that AGN-driven gas outflows dominate over SN-driven outflows for $z \lesssim 6$ (5) in halos with final mass $M_{\text{h}} \sim 10^{13.4}$ ($10^{11.75}$) $M_{\odot}$. Despite the AGN duty cycle being highest for the most massive halos, AGN feedback is more efficient in driving outflows in intermediate-mass halos, with $M_{\text{h}} \sim 10^{11.5} M_{\odot}$, in which the ratio between the energy emitted by the black hole and the binding energy of the halo reaches ~ 0.01 . Noticeably, the impact of AGN feedback on the mass assembly history of the most massive galaxies seems almost negligible.

Future Prospects

We are lucky enough to live in a critical moment for the study of the high-redshift black hole population. In 2016, the first direct detections of gravitational waves from the merger of black hole and neutron star binary systems (Abbott et al. 2016) paved the way for the development of multi-messenger astronomy, and it is only a matter of time before we will be able to collect both gravitational and spectral information about such systems. Upcoming space-based facilities will help in understanding and shedding light on many of the topics touched in this thesis. We have seen how the space interferometer *LISA* will be able to detect black hole mergers at $z = 5 - 13$ and with final masses $10^{3.5} - 10^5 M_{\odot}$. At the same time, space X-ray telescopes like *Athena* and *Lynx* will probe the population of accreting black holes at $z > 6$, covering the mass range $10^4 - 10^8 M_{\odot}$, while the James Webb Space Telescope will be invaluable to understand the physical properties of the host galaxies of these black holes, and will also be able to distinguish between different black hole seeding mechanisms (Natarajan et al. 2017). The development and increasing accuracy of theoretical models like the one presented in this thesis will serve the purpose of guiding these observational efforts, which in turn will provide new information to inform the models and increase their accuracy.

Indeed, the model we have built for this project leaves of course space for improvement. We do not include the effects of recoil kicks in black hole mergers, which could lower the mass accreted through mergers. We do not account for the black hole binary inspiraling phase and for the final parsec problem, which means that we are likely underestimating the merging timescales. Adding spatial information to the halos of our merger trees would help in better constraining the flux of the Lyman-Werner background in which each galaxy is immersed, in order to gain a more precise determination of the hosts of direct-collapse black holes. In addition, it would allow us to perform differential studies of the evolution of central and satellite galaxies. Including a model for the evolution of the dust content of galaxies, instead, would provide a direct way to account for obscuration effects, without having to rely on observational surveys, which at these redshift are limited and yield heterogeneous results. Finally, the reionization model we explore here is very basic and does not properly account for the patchiness and topology nor for the temporal evolution of the process. For that, again, including spatial information about the merger trees would be of vital importance.

Clearly, a lot of work still needs to be done, but thanks to the powerful eyes of the next generation telescopes we can look at it with curious optimism.

BIBLIOGRAPHY

- Aalto, S., Garcia-Burillo, S., Muller, S., et al. 2012, *A&A*, 537, A44
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *Physical Review X*, 6, 041015
- Abel, T., Bryan, G. L., & Norman, M. L. 2002, *Science*, 295, 93
- Adams, N. J., Bowler, R. A. A., Jarvis, M. J., et al. 2020, *MNRAS*, 494, 1771
- Agarwal, B. 2019, Primordial gas collapse in the presence of radiation: direct collapse black hole or Population III star?, ed. M. Latif & D. Schleicher, 115–124
- Agarwal, B., Dalla Vecchia, C., Johnson, J. L., Khochfar, S., & Paardekooper, J.-P. 2014, *MNRAS*, 443, 648
- Ajith, P., Babak, S., Chen, Y., et al. 2007, *Classical and Quantum Gravity*, 24, S689
- Akiyama, M., He, W., Ikeda, H., et al. 2018, *PASJ*, 70, S34
- Albrecht, A. & Steinhardt, P. J. 1982, *Phys. Rev. Lett.*, 48, 1220
- Alpher, R. A., Bethe, H., & Gamow, G. 1948, *Physical Review*, 73, 803
- Amarantidis, S., Afonso, J., Messias, H., et al. 2019, *MNRAS*, 485, 2694
- Amaro-Seoane, P., Aoudia, S., Babak, S., et al. 2012, *Classical and Quantum Gravity*, 29, 124016
- Amaro-Seoane, P., Audley, H., Babak, S., et al. 2017, *ArXiv e-prints*
- Anglés-Alcázar, D., Faucher-Giguère, C.-A., Quataert, E., et al. 2017, *MNRAS*, 472, L109

BIBLIOGRAPHY

- Antonucci, R. R. J. & Miller, J. S. 1985, *ApJ*, 297, 621
- Arun, K. G., Babak, S., Berti, E., et al. 2009, *Classical and Quantum Gravity*, 26, 094027
- Atek, H., Richard, J., Jauzac, M., et al. 2015a, *ApJ*, 814, 69
- Atek, H., Richard, J., Kneib, J.-P., et al. 2015b, *ApJ*, 800, 18
- Atek, H., Richard, J., Kneib, J.-P., & Schaerer, D. 2018, *MNRAS*, 479, 5184
- Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, *ApJS*, 227, 11
- Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, *Nature*, 553, 473
- Bandara, K., Crampton, D., & Simard, L. 2009, *ApJ*, 704, 1135
- Barack, L., Cardoso, V., Nissanke, S., et al. 2019, *Classical and Quantum Gravity*, 36, 143001
- Barausse, E. 2012, *MNRAS*, 423, 2533
- Barkana, R. & Loeb, A. 1999, *ApJ*, 523, 54
- Barkana, R. & Loeb, A. 2001, *Phys. Rep.*, 349, 125
- Baugh, C. M., Gonzalez-Perez, V., Lagos, C. d. P., et al. 2019, *MNRAS*, 483, 4922
- Beckmann, R. S., Devriendt, J., & Slyz, A. 2019, *MNRAS*, 483, 3488
- Begelman, M. C. 1979, *MNRAS*, 187, 237
- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1984, *Reviews of Modern Physics*, 56, 255
- Begelman, M. C., Rossi, E. M., & Armitage, P. J. 2008, *MNRAS*, 387, 1649
- Begelman, M. C., Volonteri, M., & Rees, M. J. 2006, *MNRAS*, 370, 289
- Benson, A. J., Bower, R. G., Frenk, C. S., et al. 2003, *ApJ*, 599, 38
- Berger, M. J. & Colella, P. 1989, *Journal of Computational Physics*, 82, 64
- Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, *Phys. Rep.*, 367, 1

-
- Binney, J. 1977, *ApJ*, 215, 483
- Blandford, R. D. & Znajek, R. L. 1977, *MNRAS*, 179, 433
- Blecha, L., Sijacki, D., Kelley, L. Z., et al. 2016, *MNRAS*, 456, 961
- Boco, L., Lapi, A., & Danese, L. 2020, *ApJ*, 891, 94
- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, *ApJ*, 379, 440
- Bondi, H. & Hoyle, F. 1944, *MNRAS*, 104, 273
- Bonetti, M., Sesana, A., Haardt, F., Barausse, E., & Colpi, M. 2019, *MNRAS*, 486, 4044
- Bonoli, S., Mayer, L., & Callegari, S. 2014, *MNRAS*, 437, 1576
- Boutsia, K., Grazian, A., Giallongo, E., et al. 2011, *ApJ*, 736, 41
- Bouwens, R. J., Illingworth, G. D., Franx, M., et al. 2009, *ApJ*, 705, 936
- Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, *ApJ*, 670, 928
- Bouwens, R. J., Illingworth, G. D., González, V., et al. 2010, *ApJ*, 725, 1587
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011, *ApJ*, 737, 90
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015a, *ApJ*, 803, 34
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015b, *ApJ*, 803, 34
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, *ApJ*, 752, L5
- Bouwens, R. J., Oesch, P. A., Illingworth, G. D., Ellis, R. S., & Stefanon, M. 2017, *ApJ*, 843, 129
- 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
- Bowler, R. A. A., Adams, N. J., Jarvis, M. J., & Häußler, B. 2021, *MNRAS*, 502, 662
- Bowler, R. A. A., Dunlop, J. S., McLure, R. J., et al. 2015, *MNRAS*, 452, 1817
- Bowler, R. A. A., Dunlop, J. S., McLure, R. J., et al. 2014, *MNRAS*, 440, 2810

BIBLIOGRAPHY

- Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, *PASA*, 30, e031
- Bradley, L. D., Trenti, M., Oesch, P. A., et al. 2012, *ApJ*, 760, 108
- Bradley, L. D., Zitrin, A., Coe, D., et al. 2014, *ApJ*, 792, 76
- Brandt, W. N. & Alexander, D. M. 2015, *A&A Rev.*, 23, 1
- Bromm, V. 2013, *Reports on Progress in Physics*, 76, 112901
- Bromm, V. & Larson, R. B. 2004, *ARA&A*, 42, 79
- Bromm, V. & Loeb, A. 2003, *ApJ*, 596, 34
- Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, *ApJ*, 429, 582
- Cano-Díaz, M., Maiolino, R., Marconi, A., et al. 2012, *A&A*, 537, L8
- Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2018, *ApJ*, 867, 148
- Cappelluti, N., Li, Y., Ricarte, A., et al. 2017, *ApJ*, 837, 19
- Carlberg, R. G., Couchman, H. M. P., & Thomas, P. A. 1990, *ApJ*, 352, L29
- Carroll, S. M., Press, W. H., & Turner, E. L. 1992, *ARA&A*, 30, 499
- Castellano, M., Fontana, A., Paris, D., et al. 2010, *A&A*, 524, A28
- Cattaneo, A. 2002, *MNRAS*, 333, 353
- Chen, X. & Kamionkowski, M. 2004, *Phys. Rev. D*, 70, 043502
- Choi, J.-H., Shlosman, I., & Begelman, M. C. 2015, *MNRAS*, 450, 4411
- Choudhury, T. R. & Ferrara, A. 2007, *MNRAS*, 380, L6
- Ciardi, B. & Ferrara, A. 2005, *Space Sci. Rev.*, 116, 625
- Cicone, C., Maiolino, R., Gallerani, S., et al. 2015, *A&A*, 574, A14
- Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, *MNRAS*, 319, 168
- Colpi, M. 2019, *Probing the formation of the seeds of supermassive black holes with gravitational waves*, ed. M. Latif & D. Schleicher, 241–268
- Colpi, M. & Sesana, A. 2017, *Gravitational Wave Sources in the Era of Multi-Band Gravitational Wave Astronomy*, 43–140

-
- Costa, T., Sijacki, D., Trenti, M., & Haehnelt, M. G. 2014, MNRAS, 439, 2146
- Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937
- Cresci, G., Mainieri, V., Brusa, M., et al. 2015, ApJ, 799, 82
- Croton, D. J. 2009, MNRAS, 394, 1109
- Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
- Croton, D. J., Stevens, A. R. H., Tonini, C., et al. 2016, ApJS, 222, 22
- Dayal, P., Choudhury, T. R., Bromm, V., & Pacucci, F. 2017a, ApJ, 836, 16
- Dayal, P., Choudhury, T. R., Pacucci, F., & Bromm, V. 2017b, MNRAS, 472, 4414
- Dayal, P., Dunlop, J. S., Maio, U., & Ciardi, B. 2013, MNRAS, 434, 1486
- Dayal, P. & Ferrara, A. 2018, Phys. Rep., 780, 1
- Dayal, P., Ferrara, A., Dunlop, J. S., & Pacucci, F. 2014a, MNRAS, 445, 2545
- Dayal, P., Ferrara, A., Dunlop, J. S., & Pacucci, F. 2014b, MNRAS, 445, 2545
- Dayal, P., Mesinger, A., & Pacucci, F. 2015, ApJ, 806, 67
- Dayal, P., Rossi, E. M., Shiralilou, B., et al. 2019, MNRAS, 486, 2336
- Dayal, P., Volonteri, M., Choudhury, T. R., et al. 2020, MNRAS, 495, 3065
- De Barros, S., Oesch, P. A., Labbé, I., et al. 2019, MNRAS, 489, 2355
- De Rosa, G., Venemans, B. P., Decarli, R., et al. 2014, ApJ, 790, 145
- DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, PASP, 129, 045001
- Decarli, R., Falomo, R., Treves, A., et al. 2010, MNRAS, 402, 2453
- Decarli, R., Walter, F., Venemans, B. P., et al. 2018, ApJ, 854, 97
- Dekel, A. & Birnboim, Y. 2006, MNRAS, 368, 2
- Dekel, A. & Silk, J. 1986, ApJ, 303, 39
- Devecchi, B. & Volonteri, M. 2009, ApJ, 694, 302

BIBLIOGRAPHY

- Devecchi, B., Volonteri, M., Rossi, E. M., Colpi, M., & Portegies Zwart, S. 2012, *MNRAS*, 421, 1465
- Di Matteo, T., Croft, R. A. C., Feng, Y., Waters, D., & Wilkins, S. 2017, *MNRAS*, 467, 4243
- Di Matteo, T., Springel, V., & Hernquist, L. 2005, *Nature*, 433, 604
- Dijkstra, M., Ferrara, A., & Mesinger, A. 2014, *MNRAS*, 442, 2036
- Dijkstra, M., Haiman, Z., & Loeb, A. 2004, *ApJ*, 613, 646
- Dotan, C., Rossi, E. M., & Shaviv, N. J. 2011, *MNRAS*, 417, 3035
- Dotti, M., Volonteri, M., Perego, A., et al. 2010, *MNRAS*, 402, 682
- Dubois, Y., Gavazzi, R., Peirani, S., & Silk, J. 2013, *MNRAS*, 433, 3297
- Dubois, Y., Pichon, C., Welker, C., et al. 2014, *MNRAS*, 444, 1453
- Dubois, Y., Volonteri, M., Silk, J., et al. 2015, *MNRAS*, 452, 1502
- Duncan, K., Conselice, C. J., Mortlock, A., et al. 2014, *MNRAS*, 444, 2960
- Duras, F., Bongiorno, A., Ricci, F., et al. 2020, *A&A*, 636, A73
- Eddington, A. S. 1916, *MNRAS*, 77, 16
- Einstein, A. 1917, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften* (Berlin), 142
- Eisenstein, D. J. & Loeb, A. 1995, *ApJ*, 443, 11
- Ellis, R. S., McLure, R. J., Dunlop, J. S., et al. 2013, *ApJ*, 763, L7
- Espinoza, P., Selman, F. J., & Melnick, J. 2009, *A&A*, 501, 563
- Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019, *ApJ*, 875, L1
- Fabian, A. C. 2006, *Astronomische Nachrichten*, 327, 943
- Fabian, A. C. 2012, *ARA&A*, 50, 455
- Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001, *AJ*, 122, 2833
- Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, *AJ*, 132, 117
- Favata, M., Hughes, S. A., & Holz, D. E. 2004, *ApJ*, 607, L5

-
- Feng, Y., Di Matteo, T., Croft, R., et al. 2015, *ApJ*, 808, L17
- Ferrara, A., Salvadori, S., Yue, B., & Schleicher, D. 2014, *MNRAS*, 443, 2410
- Ferrarese, L. 2002, *ApJ*, 578, 90
- Ferrarese, L. & Merritt, D. 2000, *ApJ*, 539, L9
- Feruglio, C., Maiolino, R., Piconcelli, E., et al. 2010, *A&A*, 518, L155
- Fiacconi, D. & Rossi, E. M. 2016, *MNRAS*, 455, 2
- Fiacconi, D. & Rossi, E. M. 2017, *MNRAS*, 464, 2259
- Fields, B. D. 2011, *Annual Review of Nuclear and Particle Science*, 61, 47
- Finkelstein, S. L., D'Aloisio, A., Paardekooper, J.-P., et al. 2019, *ApJ*, 879, 36
- Finkelstein, S. L., Papovich, C., Salmon, B., et al. 2012, *ApJ*, 756, 164
- Finkelstein, S. L., Ryan, Russell E., J., Papovich, C., et al. 2015, *ApJ*, 810, 71
- Finlator, K., Davé, R., & Özel, F. 2011a, *ApJ*, 743, 169
- Finlator, K., Keating, L., Oppenheimer, B. D., Davé, R., & Zackrisson, E. 2018, *MNRAS*, 480, 2628
- Finlator, K., Oppenheimer, B. D., & Davé, R. 2011b, *MNRAS*, 410, 1703
- Fischer, J., Sturm, E., González-Alfonso, E., et al. 2010, *A&A*, 518, L41
- Fitchett, M. J. 1983, *MNRAS*, 203, 1049
- Fixsen, D. J. 2009, *ApJ*, 707, 916
- Furlanetto, S. R. & Oh, S. P. 2008, *ApJ*, 681, 1
- Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2006, *MNRAS*, 365, 1012
- Gallerani, S., Pallottini, A., Feruglio, C., et al. 2018, *MNRAS*, 473, 1909
- Gazagnes, S., Chisholm, J., Schaerer, D., Verhamme, A., & Izotov, Y. 2020, *A&A*, 639, A85

BIBLIOGRAPHY

- Gebhardt, K., Bender, R., Bower, G., et al. 2000, *ApJ*, 539, L13
- Giallongo, E., Grazian, A., Fiore, F., et al. 2015, *A&A*, 578, A83
- Gilfanov, M. & Merloni, A. 2014, *Space Sci. Rev.*, 183, 121
- Gingold, R. A. & Monaghan, J. J. 1977, *MNRAS*, 181, 375
- Gnedin, N. Y. 2000, *ApJ*, 542, 535
- Gnedin, N. Y. & Kaurov, A. A. 2014, *ApJ*, 793, 30
- González, V., Labbé, I., Bouwens, R. J., et al. 2011, *ApJ*, 735, L34
- González, V., Labbé, I., Bouwens, R. J., et al. 2010, *ApJ*, 713, 115
- Governato, F., Babul, A., Quinn, T., et al. 1999, *MNRAS*, 307, 949
- Grazian, A., Castellano, M., Koekemoer, A. M., et al. 2011, *A&A*, 532, A33
- Grazian, A., Fontana, A., Santini, P., et al. 2015, *A&A*, 575, A96
- Grazian, A., Giallongo, E., Boutsia, K., et al. 2018, *A&A*, 613, A44
- Grazian, A., Giallongo, E., Gerbasi, R., et al. 2016, *A&A*, 585, A48
- Greene, J. E., Zakamska, N. L., Ho, L. C., & Barth, A. J. 2011, *ApJ*, 732, 9
- Griffin, A. J., Lacey, C. G., Gonzalez-Perez, V., et al. 2020, *MNRAS*, 492, 2535
- Gültekin, K., Richstone, D. O., Gebhardt, K., et al. 2009, *ApJ*, 698, 198
- Gunn, J. E. & Peterson, B. A. 1965, *ApJ*, 142, 1633
- Guo, Q. & White, S. 2014, *MNRAS*, 437, 3228
- Guth, A. H. 1981, *Phys. Rev. D*, 23, 347
- Habouzit, M., Li, Y., Somerville, R. S., et al. 2020, *arXiv e-prints*, arXiv:2006.10094
- Habouzit, M., Volonteri, M., & Dubois, Y. 2017, *MNRAS*, 468, 3935
- Habouzit, M., Volonteri, M., Latif, M., Dubois, Y., & Peirani, S. 2016a, *MNRAS*, 463, 529
- Habouzit, M., Volonteri, M., Latif, M., et al. 2016b, *MNRAS*, 456, 1901

-
- Habouzit, M., Volonteri, M., Somerville, R. S., et al. 2019, MNRAS, 489, 1206
- Haehnelt, M. G. 1994, MNRAS, 269, 199
- Haehnelt, M. G. & Rees, M. J. 1993, MNRAS, 263, 168
- Haiman, Z. 2004, ApJ, 613, 36
- Haiman, Z. 2013, The Formation of the First Massive Black Holes, ed. T. Wiklund, B. Mobasher, & V. Bromm, Vol. 396, 293
- Haiman, Z. & Loeb, A. 2001, ApJ, 552, 459
- Haiman, Z., Rees, M. J., & Loeb, A. 1996, ApJ, 467, 522
- Hao, J.-M., Yuan, Y.-F., & Wang, L. 2015, MNRAS, 451, 1875
- Häring, N. & Rix, H.-W. 2004, ApJ, 604, L89
- Harrison, C. M., Alexander, D. M., Swinbank, A. M., et al. 2012, MNRAS, 426, 1073
- Hartwig, T. 2019, Statistical predictions for the first black holes, ed. M. Latif & D. Schleicher, 161–175
- Hartwig, T., Agarwal, B., & Regan, J. A. 2018, MNRAS, 479, L23
- Hartwig, T., Latif, M. A., Magg, M., et al. 2016, MNRAS, 462, 2184
- Hasegawa, K. & Semelin, B. 2013, MNRAS, 428, 154
- Hasinger, G. 2008, A&A, 490, 905
- Hassan, S., Davé, R., Mitra, S., et al. 2018, MNRAS, 473, 227
- Heger, A., Fryer, C. L., Woosley, S. E., Langer, N., & Hartmann, D. H. 2003, ApJ, 591, 288
- Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, MNRAS, 451, 2663
- Hirschmann, M., De Lucia, G., Wilman, D., et al. 2014, MNRAS, 444, 2938
- Hirschmann, M., Somerville, R. S., Naab, T., & Burkert, A. 2012, MNRAS, 426, 237
- Hogan, C. J. & Rees, M. J. 1979, MNRAS, 188, 791

BIBLIOGRAPHY

- Huang, K.-W., Ni, Y., Feng, Y., & Di Matteo, T. 2020, *MNRAS*, 496, 1
- Hubble, E. 1929, *Proceedings of the National Academy of Science*, 15, 168
- Hutter, A., Dayal, P., Legrand, L., Gottlöber, S., & Yepes, G. 2020a, *arXiv e-prints*, arXiv:2008.13215
- Hutter, A., Dayal, P., Yepes, G., et al. 2020b, *arXiv e-prints*, arXiv:2004.08401
- Inayoshi, K., Haiman, Z., & Ostriker, J. P. 2016, *MNRAS*, 459, 3738
- Inayoshi, K., Visbal, E., & Haiman, Z. 2020, *ARA&A*, 58, 27
- Inoue, A. K., Iwata, I., & Deharveng, J.-M. 2006, *MNRAS*, 371, L1
- Ishigaki, M., Kawamata, R., Ouchi, M., et al. 2015, *ApJ*, 799, 12
- Ishigaki, M., Kawamata, R., Ouchi, M., et al. 2018, *ApJ*, 854, 73
- Israel, W. 1967, *Physical Review*, 164, 1776
- Iwasawa, K., Gilli, R., Vignali, C., et al. 2012, *A&A*, 546, A84
- Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018, *MNRAS*, 478, 4851
- Jeon, M., Pawlik, A. H., Bromm, V., & Milosavljević, M. 2014, *MNRAS*, 440, 3778
- Jeon, M., Pawlik, A. H., Greif, T. H., et al. 2012, *ApJ*, 754, 34
- Jiang, L., McGreer, I. D., Fan, X., et al. 2016, *ApJ*, 833, 222
- Jiang, Y.-F., Stone, J. M., & Davis, S. W. 2014, *ApJ*, 796, 106
- Johnson, J. L. & Khochfar, S. 2011, *ApJ*, 743, 126
- Johnson, J. L., Whalen, D. J., Fryer, C. L., & Li, H. 2012, *ApJ*, 750, 66
- Kakiichi, K. & Gronke, M. 2019, *arXiv e-prints*, arXiv:1905.02480
- Kashikawa, N., Ishizaki, Y., Willott, C. J., et al. 2015, *ApJ*, 798, 28
- Kashikawa, N., Shimasaku, K., Matsuda, Y., et al. 2011, *ApJ*, 734, 119
- Katz, H. 2019, *Black hole formation in the first stellar clusters*, ed. M. Latif & D. Schleicher, 125–143
- Katz, H., Kimm, T., Haehnelt, M. G., et al. 2019, *MNRAS*, 483, 1029

-
- Katz, H., Sijacki, D., & Haehnelt, M. G. 2015, *MNRAS*, 451, 2352
- Kawamata, R., Ishigaki, M., Shimasaku, K., et al. 2018, *ApJ*, 855, 4
- Kawamata, R., Oguri, M., Ishigaki, M., Shimasaku, K., & Ouchi, M. 2016, *ApJ*, 819, 114
- Kennicutt, Robert C., J. 1998, *ARA&A*, 36, 189
- Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, *MNRAS*, 363, 2
- Khandai, N., Di Matteo, T., Croft, R., et al. 2015, *MNRAS*, 450, 1349
- Khusanova, Y., Le Fèvre, O., Cassata, P., et al. 2020, *A&A*, 634, A97
- Kim, H.-S., Wyithe, J. S. B., Park, J., & Lacey, C. G. 2013, *MNRAS*, 433, 2476
- Klein, A., Barausse, E., Sesana, A., et al. 2016, *Phys. Rev. D*, 93, 024003
- Klessen, R. 2019, *Formation of the first stars*, ed. M. Latif & D. Schleicher, 67–97
- Kogut, A., Spergel, D. N., Barnes, C., et al. 2003, *ApJS*, 148, 161
- Koopmans, L., Pritchard, J., Mellema, G., et al. 2015, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 1
- Kormendy, J. & Ho, L. C. 2013, *ARA&A*, 51, 511
- Kormendy, J. & Richstone, D. 1995, *ARA&A*, 33, 581
- Kulkarni, G., Worseck, G., & Hennawi, J. F. 2019a, *MNRAS*, 488, 1035
- Kulkarni, G., Worseck, G., & Hennawi, J. F. 2019b, *MNRAS*, 488, 1035
- La Franca, F., Fiore, F., Comastri, A., et al. 2005, *ApJ*, 635, 864
- Labbé, I., González, V., Bouwens, R. J., et al. 2010, *ApJ*, 716, L103
- Labbé, I., Oesch, P. A., Bouwens, R. J., et al. 2013, *ApJ*, 777, L19
- Lacey, C. & Cole, S. 1993a, *MNRAS*, 262, 627
- Lacey, C. & Cole, S. 1993b, *MNRAS*, 262, 627
- Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2016, *MNRAS*, 462, 3854
- Lagos, C. D. P., Cora, S. A., & Padilla, N. D. 2008, *MNRAS*, 388, 587

BIBLIOGRAPHY

- Lagos, C. d. P., Tobar, R. J., Robotham, A. S. G., et al. 2018, *MNRAS*, 481, 3573
- Latif, M. A. 2019, Black hole formation via gas-dynamical processes, ed. M. Latif & D. Schleicher, 99–113
- Latif, M. A. & Ferrara, A. 2016a, *PASA*, 33, e051
- Latif, M. A. & Ferrara, A. 2016b, *PASA*, 33, e051
- Latif, M. A., Schleicher, D. R. G., Schmidt, W., & Niemeyer, J. 2013a, *MNRAS*, 433, 1607
- Latif, M. A., Schleicher, D. R. G., Schmidt, W., & Niemeyer, J. 2013b, *MNRAS*, 430, 588
- Latif, M. A., Schleicher, D. R. G., Schmidt, W., & Niemeyer, J. C. 2013c, *MNRAS*, 436, 2989
- Latif, M. A., Volonteri, M., & Wise, J. H. 2018, *MNRAS*, 476, 5016
- Lauer, T. R., Tremaine, S., Richstone, D., & Faber, S. M. 2007, *ApJ*, 670, 249
- Lee, K.-S., Giavalisco, M., Conroy, C., et al. 2009, *ApJ*, 695, 368
- Leitherer, C., Hernandez, S., Lee, J. C., & Oey, M. S. 2016, *ApJ*, 823, 64
- Leitherer, C., Ortiz Otálvaro, P. A., Bresolin, F., et al. 2010, *ApJS*, 189, 309
- Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, *ApJS*, 123, 3
- Lemaître, G. 1927, *Annales de la Société Scientifique de Bruxelles*, 47, 49
- Linde, A. D. 1982, *Physics Letters B*, 108, 389
- Liu, T., Tozzi, P., Wang, J.-X., et al. 2017, *ApJS*, 232, 8
- Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, *ApJ*, 835, 113
- Lodato, G. & Natarajan, P. 2006, *MNRAS*, 371, 1813
- Loeb, A. & Rasio, F. A. 1994, *ApJ*, 432, 52
- Lu, Y., Wechsler, R. H., Somerville, R. S., et al. 2014, *ApJ*, 795, 123
- Lucy, L. B. 1977, *AJ*, 82, 1013

-
- Luo, Y., Ardaneh, K., Shlosman, I., et al. 2018, MNRAS, 476, 3523
- Lupi, A., Colpi, M., Devecchi, B., Galanti, G., & Volonteri, M. 2014, MNRAS, 442, 3616
- Lupi, A., Haardt, F., Dotti, M., et al. 2016, MNRAS, 456, 2993
- Lupi, A., Volonteri, M., Decarli, R., et al. 2019, MNRAS, 488, 4004
- Madau, P. & Dickinson, M. 2014, ARA&A, 52, 415
- Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429
- Madau, P., Pozzetti, L., & Dickinson, M. 1998, ApJ, 498, 106
- Madau, P. & Rees, M. J. 2000, ApJ, 542, L69
- Madau, P. & Rees, M. J. 2001a, ApJ, 551, L27
- Madau, P. & Rees, M. J. 2001b, ApJ, 551, L27
- Madau, P., Rees, M. J., Volonteri, M., Haardt, F., & Oh, S. P. 2004, ApJ, 604, 484
- Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285
- Maiolino, R., Gallerani, S., Neri, R., et al. 2012, MNRAS, 425, L66
- Maiolino, R., Russell, H. R., Fabian, A. C., et al. 2017, Nature, 544, 202
- Manti, S., Gallerani, S., Ferrara, A., Greig, B., & Feruglio, C. 2017, MNRAS, 466, 1160
- Marconi, A., Risaliti, G., Gilli, R., et al. 2004, MNRAS, 351, 169
- Marshall, M. A., Mutch, S. J., Qin, Y., Poole, G. B., & Wyithe, J. S. B. 2019, MNRAS, 488, 1941
- Martin, D. C., Seibert, M., Buat, V., et al. 2005, ApJ, 619, L59
- Maseda, M. V., Bacon, R., Lam, D., et al. 2020, MNRAS, 493, 5120
- Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2018, PASJ, 70, S35
- Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2016, ApJ, 828, 26
- Matthee, J., Sobral, D., Best, P., et al. 2017, MNRAS, 465, 3637
- Matthee, J., Sobral, D., Santos, S., et al. 2015, MNRAS, 451, 400

BIBLIOGRAPHY

- Mayer, L. 2019, Super-Eddington accretion; flow regimes and conditions in high-*z* galaxies, ed. M. Latif & D. Schleicher, 195–222
- Mazzucchelli, C., Bañados, E., Venemans, B. P., et al. 2017, *ApJ*, 849, 91
- McConnell, N. J. & Ma, C.-P. 2013, *ApJ*, 764, 184
- McGreer, I. D., Fan, X., Jiang, L., & Cai, Z. 2018, *AJ*, 155, 131
- McGreer, I. D., Jiang, L., Fan, X., et al. 2013, *ApJ*, 768, 105
- McKellar, A. 1941, Publications of the Dominion Astrophysical Observatory Victoria, 7, 251
- McLeod, D. J., McLure, R. J., & Dunlop, J. S. 2016, *MNRAS*, 459, 3812
- McLure, R. J., Cirasuolo, M., Dunlop, J. S., Foucaud, S., & Almaini, O. 2009, *MNRAS*, 395, 2196
- McLure, R. J., Dunlop, J. S., Bowler, R. A. A., et al. 2013, *MNRAS*, 432, 2696
- McLure, R. J., Dunlop, J. S., Cirasuolo, M., et al. 2010, *MNRAS*, 403, 960
- McLure, R. J. & Jarvis, M. J. 2002, *MNRAS*, 337, 109
- McNamara, B. R. & Nulsen, P. E. J. 2007, *ARA&A*, 45, 117
- McNamara, B. R., Nulsen, P. E. J., Wise, M. W., et al. 2005, *Nature*, 433, 45
- Meier, D. L. 1976, *ApJ*, 207, 343
- Mellema, G., Koopmans, L. V. E., Abdalla, F. A., et al. 2013, *Experimental Astronomy*, 36, 235
- Merloni, A., Bongiorno, A., Bolzonella, M., et al. 2010, *ApJ*, 708, 137
- Merloni, A., Bongiorno, A., Brusa, M., et al. 2014, *MNRAS*, 437, 3550
- Merloni, A. & Heinz, S. 2008, *MNRAS*, 388, 1011
- Meurer, G. R., Heckman, T. M., Lehnert, M. D., Leitherer, C., & Lowenthal, J. 1997, *AJ*, 114, 54
- Mitra, S., Choudhury, T. R., & Ferrara, A. 2018, *MNRAS*, 473, 1416
- Mo, H., van den Bosch, F. C., & White, S. 2010, *Galaxy Formation and Evolution*

-
- Morgan, C. W., Kochanek, C. S., Morgan, N. D., & Falco, E. E. 2010, *ApJ*, 712, 1129
- Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, *Nature*, 474, 616
- Mostardi, R. E., Shapley, A. E., Steidel, C. C., et al. 2015, *ApJ*, 810, 107
- Mutch, S. J., Geil, P. M., Poole, G. B., et al. 2016, *MNRAS*, 462, 250
- Nandra, K., Barret, D., Barcons, X., et al. 2013, arXiv e-prints, arXiv:1306.2307
- Natarajan, P., Pacucci, F., Ferrara, A., et al. 2017, *ApJ*, 838, 117
- Nestor, D. B., Shapley, A. E., Steidel, C. C., & Siana, B. 2011, *ApJ*, 736, 18
- Noeske, K. G., Faber, S. M., Weiner, B. J., et al. 2007, *ApJ*, 660, L47
- Noterdaeme, P., Petitjean, P., Srianand, R., Ledoux, C., & López, S. 2011, *A&A*, 526, L7
- Novak, G. S., Ostriker, J. P., & Ciotti, L. 2012, *MNRAS*, 427, 2734
- Novikov, I. D. & Thorne, K. S. 1973, in *Black Holes (Les Astres Occlus)*, 343–450
- Ntelis, P., Hamilton, J.-C., Le Goff, J.-M., et al. 2017, *J. Cosmology Astropart. Phys.*, 2017, 019
- Ocvirk, P., Gillet, N., Shapiro, P. R., et al. 2016, *MNRAS*, 463, 1462
- Ocvirk, P., Pichon, C., & Teyssier, R. 2008, *MNRAS*, 390, 1326
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2010, *ApJ*, 709, L16
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2013, *ApJ*, 773, 75
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2014, *ApJ*, 786, 108
- Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, *ApJ*, 819, 129
- Oke, J. B. & Gunn, J. E. 1983a, *Astrophysical Journal*, 266, 713
- Oke, J. B. & Gunn, J. E. 1983b, *ApJ*, 266, 713
- Ono, Y., Ouchi, M., Harikane, Y., et al. 2018, *PASJ*, 70, S10

BIBLIOGRAPHY

- Onoue, M., Kashikawa, N., Willott, C. J., et al. 2017, *ApJ*, 847, L15
- O'Shea, B. W. & Norman, M. L. 2007, *ApJ*, 654, 66
- Ouchi, M., Hamana, T., Shimasaku, K., et al. 2005, *ApJ*, 635, L117
- Ouchi, M., Harikane, Y., Shibuya, T., et al. 2018, *PASJ*, 70, S13
- Ouchi, M., Mobasher, B., Shimasaku, K., et al. 2009, *ApJ*, 706, 1136
- Paciga, G., Chang, T.-C., Gupta, Y., et al. 2011, *MNRAS*, 413, 1174
- Pacucci, F., Volonteri, M., & Ferrara, A. 2015, *MNRAS*, 452, 1922
- Park, K. & Ricotti, M. 2012, *ApJ*, 747, 9
- Parkinson, H., Cole, S., & Helly, J. 2008, *MNRAS*, 383, 557
- Parsa, S., Dunlop, J. S., & McLure, R. J. 2018, *MNRAS*, 474, 2904
- Parsons, A. R., Backer, D. C., Foster, G. S., et al. 2010, *AJ*, 139, 1468
- Partridge, R. B. & Peebles, P. J. E. 1967, *ApJ*, 147, 868
- Patil, A. H., Yatawatta, S., Koopmans, L. V. E., et al. 2017, *ApJ*, 838, 65
- Pawlik, A. H., Schaye, J., & Dalla Vecchia, C. 2015, *MNRAS*, 451, 1586
- Peebles, P. J. E. & Yu, J. T. 1970, *ApJ*, 162, 815
- Peng, C. Y. 2007, *ApJ*, 671, 1098
- Peng, C. Y., Impey, C. D., Rix, H.-W., et al. 2006, *ApJ*, 649, 616
- Penzias, A. A. & Wilson, R. W. 1965, *ApJ*, 142, 419
- Peterson, B. M. 2010, in *American Astronomical Society Meeting Abstracts*, Vol. 215, *American Astronomical Society Meeting Abstracts #215*, 373.01
- Pezzulli, E., Valiante, R., & Schneider, R. 2016a, *MNRAS*, 458, 3047
- Pezzulli, E., Valiante, R., & Schneider, R. 2016b, *MNRAS*, 458, 3047
- Pfister, H., Volonteri, M., Dubois, Y., Dotti, M., & Colpi, M. 2019, *MNRAS*, 486, 101
- Piana, O., Dayal, P., Volonteri, M., & Choudhury, T. R. 2021, *MNRAS*, 500, 2146

-
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016a, A&A, 594, A13
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, A&A, 594, A13
- Portegies Zwart, S. F., Makino, J., McMillan, S. L. W., & Hut, P. 1999, A&A, 348, 117
- Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
- Press, W. H. & Schechter, P. 1974, ApJ, 187, 425
- Qin, Y., Mutch, S. J., Poole, G. B., et al. 2017, MNRAS, 472, 2009
- Reddy, N., Dickinson, M., Elbaz, D., et al. 2012, ApJ, 744, 154
- Reed, S. L., Banerji, M., Becker, G. D., et al. 2019, MNRAS, 487, 1874
- Reed, S. L., McMahon, R. G., Banerji, M., et al. 2015, MNRAS, 454, 3952
- Reed, S. L., McMahon, R. G., Martini, P., et al. 2017, MNRAS, 468, 4702
- Rees, M. J. 1984, ARA&A, 22, 471
- Rees, M. J. & Ostriker, J. P. 1977, MNRAS, 179, 541
- Regan, J. A. & Haehnelt, M. G. 2009, MNRAS, 396, 343
- Reines, A. E. & Volonteri, M. 2015, ApJ, 813, 82
- Ricarte, A. & Natarajan, P. 2018, MNRAS, 481, 3278
- Ricci, F., Marchesi, S., Shankar, F., La Franca, F., & Civano, F. 2017, MNRAS, 465, 1915
- Richardson, M. L. A., Scannapieco, E., Devriendt, J., et al. 2016, ApJ, 825, 83
- Rosas-Guevara, Y., Bower, R. G., Schaye, J., et al. 2016, MNRAS, 462, 190
- Rosdahl, J., Katz, H., Blaizot, J., et al. 2018, MNRAS, 479, 994
- Sadowski, A., Narayan, R., Penna, R., & Zhu, Y. 2013, MNRAS, 436, 3856
- Salpeter, E. E. 1955, ApJ, 121, 161
- Salvaterra, R., Ferrara, A., & Dayal, P. 2011, MNRAS, 414, 847

BIBLIOGRAPHY

- Salvaterra, R., Haardt, F., & Ferrara, A. 2005, *MNRAS*, 362, L50
- Salviander, S., Shields, G. A., & Bonning, E. W. 2015, *ApJ*, 799, 173
- Salviander, S., Shields, G. A., Gebhardt, K., & Bonning, E. W. 2007, *ApJ*, 662, 131
- Santamaría, L., Ohme, F., Ajith, P., et al. 2010, *Phys. Rev. D*, 82, 064016
- Savorgnan, G. A. D. & Graham, A. W. 2015, *MNRAS*, 446, 2330
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, 446, 521
- Schenker, M. A., Robertson, B. E., Ellis, R. S., et al. 2013, *ApJ*, 768, 196
- Scheuer, P. A. G. 1965, *Nature*, 207, 963
- Schleicher, D. R. G. 2019, Current observational status, ed. M. Latif & D. Schleicher, 223–239
- Schwarzschild, K. 1916, *Abh. Konigl. Preuss. Akad. Wissenschaften Jahre 1906,92, Berlin,1907, 1916*, 189
- Seiler, J., Hutter, A., Sinha, M., & Croton, D. 2019, *MNRAS*, 487, 5739
- Senatore, L. 2017, in *New Frontiers in Fields and Strings (TASI 2015)*, ed. J. Polchinski & et al., 447–543
- Sesana, A., Barausse, E., Dotti, M., & Rossi, E. M. 2014, *ApJ*, 794, 104
- Sesana, A., Gair, J., Berti, E., & Volonteri, M. 2011, *Phys. Rev. D*, 83, 044036
- Sesana, A., Volonteri, M., & Haardt, F. 2007, *MNRAS*, 377, 1711
- Shakura, N. I. & Sunyaev, R. A. 1973, *A&A*, 24, 337
- Shang, C., Bryan, G. L., & Haiman, Z. 2010, *MNRAS*, 402, 1249
- Shankar, F., Bernardi, M., & Haiman, Z. 2009, *ApJ*, 694, 867
- Shao, Y., Wang, R., Jones, G. C., et al. 2017, *ApJ*, 845, 138
- Shapiro, P. R., Iliev, I. T., & Raga, A. C. 2004, *MNRAS*, 348, 753
- Shapley, A. E., Steidel, C. C., Strom, A. L., et al. 2016, *ApJ*, 826, L24
- Shen, Y., Greene, J. E., Ho, L. C., et al. 2015, *ApJ*, 805, 96

-
- Shen, Y., Wu, J., Jiang, L., et al. 2019, *ApJ*, 873, 35
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, 323, 1
- Sheth, R. K. & Tormen, G. 1999, *MNRAS*, 308, 119
- Shields, G. A., Gebhardt, K., Salviander, S., et al. 2003, *ApJ*, 583, 124
- Shimasaku, K. & Izumi, T. 2019, *ApJ*, 872, L29
- Shin, S., Im, M., Kim, Y., et al. 2020, *ApJ*, 893, 45
- Shlosman, I., Choi, J.-H., Begelman, M. C., & Nagamine, K. 2016, *MNRAS*, 456, 500
- Sijacki, D., Springel, V., Di Matteo, T., & Hernquist, L. 2007, *MNRAS*, 380, 877
- Sijacki, D., Vogelsberger, M., Genel, S., et al. 2015, *MNRAS*, 452, 575
- Silk, J. 2013, *ApJ*, 772, 112
- Silk, J. & Rees, M. J. 1998, *A&A*, 331, L1
- Smit, R., Bouwens, R. J., Carniani, S., et al. 2018, *Nature*, 553, 178
- Smit, R., Bouwens, R. J., Labbé, I., et al. 2014, *ApJ*, 784, 58
- Smith, B. D., Regan, J. A., Downes, T. P., et al. 2018, *MNRAS*, 480, 3762
- Song, M., Finkelstein, S. L., Ashby, M. L. N., et al. 2016, *ApJ*, 825, 5
- Springel, V. 2005, *MNRAS*, 364, 1105
- Springel, V. 2011, arXiv e-prints, arXiv:1109.2218
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, 435, 629
- Stanley, F., Jolly, J. B., König, S., & Knudsen, K. K. 2019, *A&A*, 631, A78
- Stark, D. P. 2016, *ARA&A*, 54, 761
- Stark, D. P., Ellis, R. S., Bunker, A., et al. 2009, *ApJ*, 697, 1493
- Stark, D. P., Schenker, M. A., Ellis, R., et al. 2013, *ApJ*, 763, 129
- Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, *ApJ*, 519, 1

BIBLIOGRAPHY

- Sugimura, K., Omukai, K., & Inoue, A. K. 2014, *MNRAS*, 445, 544
- Suh, H., Civano, F., Trakhtenbrot, B., et al. 2020, *ApJ*, 889, 32
- Sunyaev, R. A. & Truemper, J. 1979, *Nature*, 279, 506
- Sunyaev, R. A. & Zeldovich, Y. B. 1972, *Comments on Astrophysics and Space Physics*, 4, 173
- Tanaka, T. & Haiman, Z. 2009, *ApJ*, 696, 1798
- Tanaka, T. L. 2014, *Classical and Quantum Gravity*, 31, 244005
- Tegmark, M., Silk, J., Rees, M. J., et al. 1997, *ApJ*, 474, 1
- Teyssier, R. 2002, *A&A*, 385, 337
- Thomas, N., Davé, R., Anglés-Alcázar, D., & Jarvis, M. 2019, *MNRAS*, 487, 5764
- Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, *PASA*, 30, e007
- Trebitsch, M., Volonteri, M., & Dubois, Y. 2020, *MNRAS*, 494, 3453
- Trebitsch, M., Volonteri, M., Dubois, Y., & Madau, P. 2018, *MNRAS*, 478, 5607
- Tremaine, S., Gebhardt, K., Bender, R., et al. 2002, *ApJ*, 574, 740
- Tremmel, M., Karcher, M., Governato, F., et al. 2017, *MNRAS*, 470, 1121
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, 613, 898
- Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, *ApJ*, 786, 104
- Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, *ApJ*, 598, 886
- Urry, C. M. & Padovani, P. 1995, *PASP*, 107, 803
- Valiante, R., Agarwal, B., Habouzit, M., & Pezzulli, E. 2017, *PASA*, 34, e031
- Valiante, R., Schneider, R., Volonteri, M., & Omukai, K. 2016a, *MNRAS*, 457, 3356
- Valiante, R., Schneider, R., Volonteri, M., & Omukai, K. 2016b, *MNRAS*, 457, 3356

-
- Van Borm, C., Bovino, S., Latif, M. A., et al. 2014, *A&A*, 572, A22
- van de Voort, F., Schaye, J., Altay, G., & Theuns, T. 2012, *MNRAS*, 421, 2809
- van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, *A&A*, 556, A2
- Vanzella, E., de Barros, S., Vasei, K., et al. 2016, *ApJ*, 825, 41
- Vanzella, E., Nonino, M., Cupani, G., et al. 2018, *MNRAS*, 476, L15
- Venemans, B. P., Bañados, E., Decarli, R., et al. 2015a, *ApJ*, 801, L11
- Venemans, B. P., Findlay, J. R., Sutherland, W. J., et al. 2013, *ApJ*, 779, 24
- Venemans, B. P., Verdoes Kleijn, G. A., Mwebaze, J., et al. 2015b, *MNRAS*, 453, 2259
- Venemans, B. P., Walter, F., Zschaechner, L., et al. 2016, *ApJ*, 816, 37
- Vestergaard, M. 2002, *The Astrophysical Journal*, 571, 733
- Vito, F., Brandt, W. N., Yang, G., et al. 2018a, *MNRAS*, 473, 2378
- Vito, F., Brandt, W. N., Yang, G., et al. 2018b, *MNRAS*, 473, 2378
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *MNRAS*, 444, 1518
- Volonteri, M. 2012, *Science*, 337, 544
- Volonteri, M. & Begelman, M. C. 2010, *MNRAS*, 409, 1022
- Volonteri, M., Dubois, Y., Pichon, C., & Devriendt, J. 2016, *MNRAS*, 460, 2979
- Volonteri, M., Haardt, F., & Madau, P. 2003, *ApJ*, 582, 559
- Volonteri, M., Lodato, G., & Natarajan, P. 2008, *MNRAS*, 383, 1079
- Volonteri, M. & Rees, M. J. 2005, *ApJ*, 633, 624
- Volonteri, M. & Rees, M. J. 2006, *ApJ*, 650, 669
- Volonteri, M. & Reines, A. E. 2016, *ApJ*, 820, L6
- Volonteri, M. & Stark, D. P. 2011, *MNRAS*, 417, 2085
- Wang, R., Wu, X.-B., Neri, R., et al. 2016, *ApJ*, 830, 53

BIBLIOGRAPHY

- Wang, Y.-Y., Wang, L., Xiang, S.-P., et al. 2010, *Research in Astronomy and Astrophysics*, 10, 199
- Weigel, A. K., Schawinski, K., Treister, E., et al. 2015, *VizieR Online Data Catalog*, J/MNRAS/448/3167
- White, S. D. M. & Rees, M. J. 1978, *MNRAS*, 183, 341
- Williams, R. J., Quadri, R. F., Franx, M., et al. 2010, *ApJ*, 713, 738
- Willott, C. J., Albert, L., Arzoumanian, D., et al. 2010a, *AJ*, 140, 546
- Willott, C. J., Delorme, P., Omont, A., et al. 2007, *AJ*, 134, 2435
- Willott, C. J., Delorme, P., Reyl, C., et al. 2009, *AJ*, 137, 3541
- Willott, C. J., Delorme, P., Reyl, C., et al. 2010b, *AJ*, 139, 906
- Willott, C. J., McLure, R. J., Hibon, P., et al. 2013, *AJ*, 145, 4
- Windhorst, R. A., Timmes, F. X., Wyithe, J. S. B., et al. 2018, *ApJS*, 234, 41
- Wise, J. H. 2012, *First Light: A Brief Review*
- Wise, J. H. 2019, *Formation of the First Black Holes*, 177–194
- Wise, J. H., Demchenko, V. G., Halicek, M. T., et al. 2014, *MNRAS*, 442, 2560
- Worrall, D. M. 2009, *A&A Rev.*, 17, 1
- Wu, X., Kannan, R., Marinacci, F., Vogelsberger, M., & Hernquist, L. 2019, *MNRAS*, 488, 419
- Wu, X.-B., Wang, F., Fan, X., et al. 2015, *Nature*, 518, 512
- Yadav, J. K., Bagla, J. S., & Khandai, N. 2010, *MNRAS*, 405, 2009
- Yang, J., Wang, F., Wu, X.-B., et al. 2016, *ApJ*, 829, 33
- Yoo, J. & Miralda-Escud, J. 2004, *ApJ*, 614, L25
- York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, *AJ*, 120, 1579
- Zaroubi, S. 2013, *The Epoch of Reionization*, ed. T. Wiklind, B. Mobasher, & V. Bromm, Vol. 396, 45
- Zaroubi, S., de Bruyn, A. G., Harker, G., et al. 2012, *MNRAS*, 425, 2964

Zaroubi, S., Thomas, R. M., Sugiyama, N., & Silk, J. 2007, MNRAS, 375, 1269

Zel'Dovich, Y. B. 1970, A&A, 500, 13

AFTERWORD AND ACKNOWLEDGEMENTS

This has been a long journey. Nevertheless, I started my PhD with the feeling that I did not know anything about Astrophysics, and little has changed with respect to that. What helps, is that now I am a little more aware of the fact that I am in good company. Despite being one of the oldest sciences, Astronomy is not a typical one. It cannot be considered a proper experimental science like Chemistry or Biology, at least not until we will be able to reproduce a star, a galaxy or a black hole in a controlled environment, possibly long enough to allow for a complete study. This makes it extremely hard to properly test a theory. In addition, despite being considered an exact science like Mathematics, it does not leave space for a deterministic study of the cause-effect dynamics driving the evolution of the Universe (or of a single star, for that matter). There are just too many variables and unknowns that require assumptions, approximations and parametrizations if we are to deal with them. Still, in recent times, the technological development of telescopes has allowed astronomers to collect vast amounts of new information from the deepest corners of the Universe. Indeed, some of these instruments represent the finest accomplishments of the human engineering mind: I am in awe every time I think about what we are able to see through their eyes. At the same time, progressively more refined computer simulations are starting to compensate for the lack of experimental labs. For these reasons, I cannot help being optimistic and feeling lucky for having the chance to approach Astrophysics in this specific historical period.

But enough with the science. This section is a celebration of the people. Even given all the interest and fascination for Astrophysics that I have been cultivating since my teenage years, I do believe that what made this experience invaluable, and the last four years and a half truly enjoyable, is the group of people I managed to find along the way. Indeed, part of the merit goes to the Kapteyn Institute, able somehow to collect from all over the world some of the most kind, warmhearted, and fun people I have met, who made me feel at home since day one. Even if it is true that writing the acknowledgements gives a certain sense of closure, I can confidently state that many of the people I will mention here will not get rid of me so easily.

Pratika, I want to thank you for giving me the opportunity of being here in the first place. You have been my supervisor and guide through the meanders of academic science for a considerable amount of time, and I really appreciate the

effort you put in it. Our interactions have not always been the easiest: being quite different people, the professional mask we are supposed to wear at work maybe is not always the best possible interface to find a common ground where we can relate to each other more easily. The important thing, though, is that we got here, and I am truly grateful to you for this. I will take the chance also to thank Karina, for always providing, when needed, adequate support. I want to thank my reading committee, professors Raffaella Morganti, Raffaella Schneider and Sune Toft, who took it upon themselves to read and comment my thesis, which with its 32 pages of introduction and no figures is probably not the lightest read. I am grateful also to professors Marta Volonteri and Tirthankar Roy Choudhury for the profitable collaborations and for always giving useful comments on all the work shown in this thesis.

A big thank is reserved for my paranymphs; Enrico, the ease and naturalness with which we became intimate friends is something special. Your endless curiosity for the “other”, different and new, and your way of looking at things with both eyes wide open, receptive, and at the same time still lucid are inspiring to me. Simon, it might not seem so, since we have chosen our next destinations to be as far from each other as we can, but I already miss you. I consider myself very lucky for having shared a home with you, and I thank every day your knee accident that made you leave football for Astronomy, sending you here. For being somebody who does not like to talk much, you still manage to leave an impressive number of unforgettable memories.

A sincere thank to the whole *Delphi* research group, Jonas, Laurent, Ruslan, Nelly, Anne, Graziano, with whom I shared not only interminable discussions about codes and papers, always driven by the implicit curiosity and enthusiasm for the subject, but also pizzas and gingerbread houses. Jonas, your natural kindness and open predisposition towards people were immediately apparent to me when you, the first person to welcome me here, opened the door of the Kapteyn Institute and saved me from a storm on the morning of my first day of work. Laurent, with your frankness, transparency, lightness of spirit and grumpy moments, your friendship is always fun and refreshing. I thank my office mates, Niels, Cristiana, Laura, Avanti, Sara, Bram, which would provide me with much needed distractions and human contact, intermittent reliefs from the long days spent in front of a computer screen. Bram, we shared the office only virtually but we compensated with live football nights, and thank you for helping me with the Dutch translation of the summary. Thanks to all the people who make Kapteyn feel like a big family house, Martine, Christa, Ramona, Maria. Thank you Lucia for all your help in the past four years, including reading my future on tarot cards. Thank you Martin, Eite and Leon, for providing punctual computer assistance when needed. Thanks to Marc, for being my PBC mentor, and to Reynier, Scott and Leon for running the Institute with careful dedication while I was there.

I want to thank my Groningen friends, whose company I profusely enjoyed during the last years. Jonas, there is no coming back from sharing a home together, and it is nice to know that wherever I will be I can always count on a second

home. Vangelis, thank you for chilling with me for four years, for making holidays and trips and Dutch courses always more enjoyable with your caring, calm and propositional spirit. Pier, our serendipitous reunion here in Groningen brought us closer than ever, and I loved the time spent writing, discussing, playing games and watching movies with you and Martyna. Katya, I remember you were smiling to me in the corridors even before we properly met, I cannot tell you how much you have always made me feel welcomed. We have spent together so many dinners, concerts, parties and trips, and I have a feeling that many more will come. Nika, thank you for your contagiously positive and non-conventional eyes you always share with those around you. Smaran, thank you for your ironic and insightful mind. We have discovered quite a few interests in common, even if we share them only from a distance sometimes. Ale, thank you for all the nights spent chatting about more or less absurd objects, scenarios, timelines, random pieces of information; that is one interesting way to get to know each other. Danielito, thank you for your never-failing uplifting personality. Thank you Ioana, for sharing with me a few of the last but definitely most intense months. Kristiina, thanks for being a dedicated bouldering partner, and thanks to you and Karlis for mastering the fine art of the dinner party. Valentina and Peter, your cheerful selves always put me in a great mood, thank you! Thank you Fra, for sharing your passion for cinema with me and for always keeping an attentive eye on the general state of things. Pranav, I have had a lot of fun with you, and thanks for taking over Boterdiep, I hope you will enjoy it as much as I have. Thanks to Andrea, to his eternal smile, and to Pavel for the fun time spent together, usually playing a game or eating, but mostly laughing. Marina, thank you for showing me that there is somebody with more troubles in taking small (or big) decisions than me. Julia, thanks for the wine shared together and for taking care of Pastis and Anubis, I appreciate that a lot, really. Alex, thank you for your funny imitations. Thank you Grega; your visit here was short, but talking to you, even if not nearly as often as we used to, is always grounding and lightening. Thanks to all the big Kapteyn family, past and present: Anastasia, Teymoor, Filippo, Marisa, Giulio, Davide Punzo, Davide Massari, Crescenzo, Antonino, Suma, Pooja, Daniele, Hyoyin, Georg, Sampath, Tadeja, Samira, Helmer, Saikat, Eduardo, Elahe, Umit, Tirna, Fernanda, Kostas, Nick, Andres, Fangyou, Emma, Kirill, Will, Jack, Caterina, and also to those that I have not had the pleasure to meet because of the COVID19 pandemic.

Thank you to the whole Astromundus crew, with whom this journey through Astrophysics started. Grega, Mija, Efsan, Miles, Lawrence, Dina, Eli, Deepak, Abhi, Dan, Arpine, Kristhell, Alexander, Simon, Noemi, Froylan, I avidly preserve the memories of our time spent around Europe.

Un grande grazie ai miei amici di Torino, Fil, Jac, FraD, FraG, BettaM, BettaU, Lauri, Opi, Pietro, Anna, Sera, Liz, Pippo. Il fatto che a prescindere dagli anni e dalle distanze ancora decidiamo di passare intere serate a giocare a Uno online (e che ci divertiamo pure!) è semplicemente rassicurante. Avrei tante altre smancerie da distribuire ma le lascio alla vostra immaginazione.

Infine, un grazie alla mia famiglia, mamma, papà, Ge e Co. In tutte queste mie peregrinazioni, di cui ancora non si vede la fine, voi siete sempre il centro del mondo. Ogni volta, immancabilmente, il vostro supporto incondizionato si fa sentire forte e saldo, e rende tutto più facile. Se ho finito questo dottorato con tranquillità è soprattutto grazie al vostro caldo abbraccio che di parole bisogno non ha. E lo stesso vale per tutto il resto della Famiglia! Nonna, nonno, zie e zii, cugine e cugino, ogni volta che ci ritroviamo tutti insieme in qualche modo si finisce sempre in grasse risate, che sarebbero da incorniciare.