UNIVERSITÀ
DEGLI STUDI
DI TRIESTE

Università degli Studi di Trieste

XXXVII CICLO DEL DOTTORATO DI RICERCA IN

FISICA

The Formation of Supermassive Black Holes From Population III.1 Protostars – Implications for Galaxy Evolution, Black Hole Growth and High-Redshift Observations

---

SETTORE SCIENTIFICO-DISCIPLINARE: ASTRONOMIA E ASTROFISICA

DOTTORANDO:
Vieri Cammelli

COORDINATORE:
Prof. Francesco Longo

SUPERVISORE:
Prof. Pierluigi Monaco

SUPERVISORE:
Prof. Jonathan C. Tan

Anno Accademico 2023-2024

ACKNOWLEDGEMENTS

This thesis was made possible with the effort of many people throughout my PhD adventure.

I would like to start by thanking my supervisors Pierluigi Monaco and Jonathan Tan for their firm support and guidance throughout my scientific and academic experience. I am especially thankful to Jonathan Tan for hosting me as a visiting student at Chalmers University in Gothenburg, Sweden, during a long, cold, bright, and snowy year. I would also like to extend my sincere thanks to Fabio Fontanot and Gabriella De Lucia for their unparalleled assistance in unravelling the astrophysics and coding intricacies of galaxy evolution models. My gratitude goes to Richard Ellis and Matt Hayes for their invaluable help with Hubble data and for the productive time spent together in London and Stockholm working on the project. I warmly thank Jacopo Salvalaggio and Jasbir Singh as my long-term office mates for the unwavering help with my work over the years. Special thanks to Mahsa Sanati and Maya Petkova for engaging discussions related to my work. Finally, I thank my thesis reviewers, Raffaella Schneider and Andy Bunker, for the useful comments they provided.

ABSTRACT

Exploring the origins and evolution of the earliest Supermassive Black Holes (SMBHs) stands as one of the most complex challenges in astrophysics. These immense black holes reside at the cores of the most massive galaxies and wield significant influence over various aspects of galactic development. Their impact extends to feedback mechanisms, transformations in morphology, and potentially contributing to the cosmic reionization process. Despite numerous proposed theoretical explanations regarding the initial formation of SMBH seeds, distinguishing between these theories remains an arduous task due to the lack of precise observational estimates regarding their prevalence in the Universe. According to recent observations, SMBHs are expected to undergo rapid growth during the early stages of their evolution, specifically beyond redshift 6. While there are several proposals for how SMBHs form, including "direct collapse" BHs in metal-free halos, runaway merger of stellar remnant BHs in globular clusters and dark matter annihilation producing supermassive "Population III.1" stars, there is little consensus.

In this thesis work we present an implementation of the Pop III.1 seeding model for SMBHs in a theoretical model of galaxy formation and evolution to assess the growth of the SMBH population and the properties of the host galaxies. The model of Pop III.1 seeding involves SMBH formation at redshifts greater than 20 in dark matter minihalos that are isolated from external radiative and mechanical feedback, parametrized by isolation distance d_{iso}. Within a standard ΛCDM cosmology, we generate dark matter halos using the code `PINOCCHIO` and seed them according to the Pop III.1 scenario, exploring values of d_{iso} from 50 to 100 kpc in proper distance. We consider two alternative cases of SMBH seeding: a Halo Mass Threshold (HMT) model in which all halos $> 7 \times 10^{10}$ $M_\odot$ are seeded with $\sim 10^5$ $M_\odot$ black holes; an All Light Seeds ALS model in which all halos are seeded with low, stellar-mass black holes. We follow the redshift evolution of the halos, populating them with galaxies using the GAlaxy Evolution and Assembly theoretical model of galaxy formation GAEA, including accretion onto the SMBH and related feedback processes.

We first present predictions for the properties of galaxy populations, focusing on stellar masses, star formation rates, and black hole masses. The local, $z \sim 0$ metrics of occupation fraction as a function of the galaxy stellar mass, Galaxy Stellar Mass Function (GSMF), and Black Hole Mass Function (BHMF) all suggest a constraint of $d_{iso} < 75$ kpc. We discuss the implications of this result for the Pop III.1 seeding mechanism.

Secondly, we utilize the semi-analytic approach proposed above to make predictions on physical and observable quantities of our SMBHs at redshifts

higher than 4. In particular we consider their appearance as Active Galactic Nuclei (AGNs) and we predict key quantities such as BH masses, AGN luminosity functions and Eddington ratios, and test whether it is possible to put some constraints on the SMBH seeding scenario. In detail, we compare with data from recent works reporting high redshift observations of AGNs detected with the James Webb Space Telescope (JWST).

In the next part of the thesis, we carried out an observational study of the Hubble Ultra Deep Field (HUDF) seeking the youngest SMBHs. We aim to estimate the SMBH number density across cosmic time by means of AGN variability. In light of their variable radiation emission, we propose that examining AGN activity in these systems through photometric variability represents an unexplored yet reliable method. This work involves recent observations of the Hubble Space Telescope (HST) pointing the HUDF in Cycle 30. This revisit, occurring a decade after the initial observations and matching its depth and wavelength band, offers an optimal time frame to identify AGN variability. The specific wavelength, luminosity range, and redshift of our target make this initiative particularly well-suited for estimating the comoving number density of SMBHs out to $z \sim 9$. The resulting dataset from our program provides valuable insights to possibly differentiate between various scenarios explaining the formation of these enigmatic black holes.

CONTENTS

Acknowledgements iii

List of Figures xi

List of Tables xix

Acronyms xx

1 Introduction 1

1.1 Supermassive black holes . . . 3
1.2 Channels of SMBH seed formation . . . 5
1.2.1 Gas and stars-based mechanisms . . . 8
1.2.2 Extreme physics and early Universe mechanisms . . . 9
1.3 Pop III.1 model . . . 10
1.4 Seeding in simulations . . . 12
1.5 Semi-analytical models of galaxy formation and evolution . . . 14
1.5.1 Seeding in SAMs . . . 15
1.6 High-redshift AGNs and the Hubble Ultra Deep Field . . . 15

2 Host galaxies of Pop III.1 seeded SMBHs 19

2.1 Methods . . . 19
2.1.1 The dark matter skeleton: `PINOCCHIO` . . . 20
2.1.2 GAEA semi-analytical model . . . 21
2.1.3 Building the merger trees . . . 24
2.1.4 Calibrating the trees . . . 27
2.1.5 Seeding SMBHs . . . 31
2.2 Results . . . 33
2.2.1 Occupation fractions . . . 33
2.2.2 Stellar mass function . . . 37
2.2.3 Local black hole mass function . . . 39
2.2.4 The $M_{BH} - M_{star}$ relation . . . 40
2.2.5 Eddington ratios . . . 49
2.3 Summary & Conclusions . . . 52

3 Predicting high redshift SMBHs in the JWST era 55

3.1 Methods . . . 56
3.1.1 Inferring the AGN properties . . . 56
3.2 Results . . . 57
3.2.1 Comparison with JWST-derived properties . . . 61
3.3 Discussion . . . 64
3.4 Conclusions . . . 68
4 A search for SMBHs via variability across cosmic time 71
4.1 Datasets and methods . . . 72
4.1.1 Observations & data reduction . . . 72
4.1.2 Photometry . . . 73
4.1.3 Identifying Variable Sources . . . 74
4.1.4 Difference Imaging Variables . . . 76
4.1.5 Redshift determination and known AGNs . . . 78
4.2 Results . . . 78
4.2.1 Detected Variable Sources . . . 78
4.2.2 Redshift distribution . . . 79
4.2.3 Comparison with previous analysis . . . 82
4.3 Variability as a method to find SMBHs . . . 82
4.4 Discussion & Conclusions . . . 88
5 Final remarks 91
5.1 Summary . . . 91
5.1.1 Host galaxies of Pop III.1 seeded SMBHs . . . 91
5.1.2 High redshift SMBHs in the JWST era . . . 92
5.1.3 A census of SMBHs via AGN variability in the HUDF 92
5.2 Future prospects . . . 93
A Galaxy survival time 97
B Calibrated Galaxy Stellar Mass Function 99
Bibliography 103

LIST OF FIGURES

Figure 1.1 The images of BHs taken by the Event Horizon Telescope Collaboration (2019; 2022). 4

Figure 1.2 Mass accretion history of Eddington limited growth tracks of confirmed high redshift quasars ($z \geqslant 7$) assuming a constant radiative efficiency of 0.1. The shaded horizontal areas indicate the range of seed masses of the BH according to different formation mechanisms. Credit by Wang et al. (2021). 6

Figure 1.3 Original diagram of formation pathways proposed by Rees (1978). 7

Figure 1.4 Evolution of the comoving number density of Supermassive Black Holes (SMBHs), n_{SMBH}, as a function of redshift. Results for the Pop III.1 models with different isolation distances (in proper distance) are shown in coloured lines, see labels. The dotted lines represent the total number of SMBHs that have formed, while the solid lines show the remaining number of seeded halos after accounting for mergers. Additionally, a sample Halo Mass Threshold (Halo Mass Threshold (HMT)) model is depicted by the dashed green line, where each halo with mass greater than $m_{th} = 7.1 \times 10^{10} M_{\odot}$ is seeded. The pink dashed line shows the results, extrapolated to $z \sim 0$, of Chon et al. (2016) assuming Direct Collapse Black Hole (DCBH) forming in a simulation. The black solid square displays an estimate for the number density of SMBHs at $z = 0$ (see text). Adapted form Singh et al. (2023). . 12

Figure 1.5 Deep combined stacked image of the HUDF, captured by the HST across multiple bands between Sept. 24, 2003 and Jan. 16, 2004. This view required over 400 Hubble orbits around the earth collecting ~ 800 exposures. Nearly 10,000 galaxies populate this field. Credits: Beckwith et al. (2006). 17

Figure 2.1 HMF at redshift $z = 0$ from the PINOCCHIO Millennium-like box (solid blue line) and the Millennium simulation (solid red line). Due to the different definition of the halo mass, at $z = 0$ the halo masses in PINOCCHIO are moderately higher compared to the Millennium halos. We calibrate the PINOCCHIO HMF by using a polynomial fit of mass difference at a given number density that is a function of the redshift (dotted blue line). 28

Figure 2.2 *Upper panel:* GSMF at redshift z=0 from the Millennium-like box. Dashed (dotted) lines indicate the contribution of the central (satellite and orphan) galaxies, respectively. While central galaxies show remarkable agreement, PINOCCHIO low mass satellites differ by a factor of ~ 4 in number density with respect to the Millennium ones. *Lower panel:* Total GSMF at $z = 0$ from the Millennium-like box. The symbols show observed data points from Panter et al. (2007), Cole et al. (2001), Bell et al. (2003) and Pérez-González et al. (2008). This plot shows the goodness of the calibration, especially above $10^{10} M_{\odot}$. 30

Figure 2.3 A simplified illustration of the Pop III.1 SMBH seeding scheme (see text) showing the conditions for a star to be isolated enough to be considered as a Pop III.1 star. Adapted from Singh et al. (2023). 31

Figure 2.4 Occupation fraction of seeded galaxies as a function of different galaxy properties for several redshifts. From left to right: halo mass, stellar mass, hot gas mass and SFR. The 6 lines show different seeding mechanisms. 34

Figure 2.5 Occupation fraction of seeded galaxies as a function of the stellar mass at redshift zero. The 6 lines show different seeding mechanisms. In the range of stellar masses between few in 10^8 and 10^{10} $M_{\odot}$, we compare against different observational constraints for the occupation fraction. X-ray sources detected in local surveys pose some lower limit to the occupation fraction as presented by Miller et al. (2015) drawn as the yellow shaded region. In black we show the more stringent constraints as reported by Nguyen et al. (2019) obtained from dynamical findings in a small sample of nearby galaxies. 35

Figure 2.6 Cosmic evolution of the GSMF from redshift 0 up to 3. Solid (dashed) lines indicate the contribution of all (seeded) galaxies, respectively. The shaded areas depict the Poisson uncertainties in each mass bin. Results for Pop III.1 models with 3 values of isolation distance (in proper distance) are shown, as labelled. Observed data are in gray symbols (Panter et al., 2007; Cole et al., 2001; Bell et al., 2003; Pérez-González et al., 2008; Van Den Bosch et al., 2008; Bundy et al., 2005; Drory et al., 2004; Fontana et al., 2006). As a reference, the dotted horizontal line stands for 10 objects per mass bin in the whole volume of the box. The GSMF tells us the isolation distance parameter that better reproduces the observed trend, especially at the massive end. The 50 kpc case agrees with the exponential cut off of the massive population of galaxies from the local relation, where almost every galaxy with a stellar mass higher than $\sim 3 \times 10^{10}$ $M_\odot$ is assumed to host a SMBH, up to $z \sim 1$. At redshift 0 we compare with the HMT and ALS seeding schemes. These two cases mostly overlap in the graph. While ALS well reproduces the local observational trend by construction, HMT naturally seeds every massive galaxy resulting in efficient quenching of star formation in such systems. We use them as reference models to validate the Pop III.1 seeding at redshift z=0. The green dashed line of the HMT scheme tells us that the fraction of seeded galaxies rises sharply up to one above $\sim 10^9$ $M_\odot$ as shown from Fig. 2.5. . . 38

Figure 2.7 BHMF at redshift $z \sim 0$. Solid lines indicate the contribution of different seeding schemes. The shaded areas around the lines depict the Poisson uncertainties in each mass bin. Observed data are taken from a sample of local galaxies (shaded area Mutlu-Pakdil et al., 2016) and from the unbiased corrected relation from Shankar et al. (2020) (hatched area). As a reference, the dotted horizontal line stands for 10 objects per mass bin in the whole volume of the box. 39

Figure 2.8 Scatter plot of the evolution of the $M_{BH} - M_{star}$ relation at several redshifts. Different columns represent various seeding schemes. At redshift $z \sim 0$ we compare our predictions for our seeded galaxies with several best fits from local observations: dynamical measurements from KH+13, including massive bulge-dominated quiescent galaxies, while R+15 use a different mass-to-light ratio, plus an extended sample combining local early- and late-type galaxies by Greene et al. (2020) (hereafter G+20). At $z \sim 2$, we report the R+15 fit for reference. Moving upward, the high-z rows show the comparison with recent results from the JFAINT sample: in Pacucci et al. (2023), they directly fit the data in the redshift range $z \sim 4 - 7$ while Li et al. (2024) estimate an unbiased fit taking into account the uncertainties on the mass measurements and selection effects. Shaded regions illustrate the intrinsic scatter at 1-sigma according to each relation. Coloured symbols show faint AGNs taken from Maiolino et al. (2023), Harikane et al. (2023), Übler et al. (2023) and Kocevski et al. (2023) and reported according to their redshifts. In the top row, results at $z \sim 10$ are shown against the single data point (GN-z11) from Maiolino et al. (2024). 41

Figure 2.9 Power law fits to the $M_{BH} - M_{star}$ relation at $z \sim 0$. Dotted lines depict local fits with associated dispersion (shaded areas). The indices α are shown in the legend. Solid lines indicate the fit results to the median SMBH masses (data points with dispersion) in several stellar mass bins for different seeding models. The stellar mass range covered by the solid lines corresponds to the fitted interval. 42

Figure 2.10 Power law fits to the $M_{BH} - M_{star}$ relation at $z \sim 5$. Dashed lines depict fits to the JFAINT data with associated dispersion (shaded areas). For reference, we also show the dotted line reporting the R+15 fit. The indices α are shown in the legend. *(a) Upper panel:* Solid lines indicate the fit results to the median SMBH masses (data points with dispersion) in several stellar mass bins for different seeding models. The stellar mass range covered by the lines corresponds to the fitted interval in the global range. *(b) Lower panel:* As (a), but the fits extend over the low-mass range. 43

Figure 2.11 Medians (in log scale, solid lines) of the Eddington ratio λ_{edd} as a function of redshift for the different seeding mechanisms. Shaded areas denote the 1-sigma dispersion of the distributions. 50
Figure 2.12 Normalized histograms (in log scale) of the Eddington ratio λ_{edd} for in different redshift bins as a function of the seeding mechanism. The spacing in redshift bins corresponds to a linear spacing in the logarithm of the scale factor a. 51
Figure 3.1 Cosmic evolution of the UV luminosity function at redshift from ~4 up to ~6 for different seeding cases, as labelled. Solid lines indicate model predictions of the composite total contribution from stellar and AGN emission by all galaxies with no dust attenuation. Dashed lines show the same total contribution from active ($\lambda_{Edd} > 0.1$) seeded galaxies only (i.e. the AGN UVLF). The shaded area depicts the uncertainty due to dust attenuation at such early epochs, where lower limit represent the total dust obscuration as assumed by GAEA (see text). Several observations for galaxies, AGNs and composite UVLFs are shown, according to the legend. 59
Figure 3.2 Same as Fig. 3.1 in the redshift range $z \sim 7-9$. 60
Figure 3.3 Bolometric luminosities as a function of the BH masses. Coloured diamonds show estimates from Harikane et al. (2023) and Maiolino et al. (2023) with an average redshift of ~ 5.2. *Top panel:* Black points represent model predictions for the Pop III.1 model with $d_{iso} = 75$ kpc at redshift $z \sim 5$. *Bottom panel:* same as the top panel for the HMT seeding scheme. Dotted gray straight lines denote the bolometric luminosity relative to the Eddington limit for several fractions, from 1 down to 0.01, for a given BH mass. This comparison indicate that our simulated SMBH fail to mimic the activity and the mass of SMBHs as inferred by the Hα line detected in JWST sources. 63
Figure 3.4 *Top panel:* upper limits for bolometric luminosities estimated from M_{UV} against Hα-inferred measurements as from Harikane et al. (2023) and Maiolino et al. (2023). *Bottom panel:* same as in the top panel but for BH mass estimates. 66

Figure 3.5 Bolometric luminosities as a function of the BH masses. Similarly to Fig 3.3, coloured empty diamonds show estimates from Harikane et al. (2023) and Maiolino et al. (2023) now derived from their UV magnitude M_{UV} as explained in the text. Black points display model predictions for the Pop III.1 model with d_{iso} = 75 kpc at redshift $z \sim 5$. This comparison indicates that our simulated SMBH partially reproduce the activity and the mass of SMBHs as inferred by the M_{UV} detected in JWST sources. 67

Figure 4.1 Calibrated photometric uncertainties to identify variable candidates. We plot the photometric variability, measured as the magnitude difference between the two epochs, as a function of the mean magnitude in F105W, F140W and F160W, as labelled. Every grey dot represents a source that has been measured in both observations. Black dots show the intrinsic $1\sigma_m$ scatter of the data measured in bins of 0.5 mag width. Golden circles depict the calibrated photometric 1 σ_m uncertainties in each bin (see text). The inner pair of yellow lines are polynomial fits to these calibrated uncertainties. The next sets of yellow lines show 2 σ_m and 3 σ_m uncertainties, scaled from the 1 σ_m fit. Red squares highlight sources that are estimated to be $\geqslant 3$ σ_m variables. The right-hand panel shows a histogram of the distribution of magnitude differences for all objects detected in the field. 75

Figure 4.2 Histogram of $\Delta m/\sigma_m$ for the total sample of detected sources in F105W, F140W and F160W as labelled, in blue. The red dashed line shows the Gaussian fit to the histogram values. 77

Figure 4.3 Redshift distribution of the variable candidate sources. Different colours refer to the specific significance thresholds for variability detection in units of σ_m. At redshift between $6 < z < 7$ we also include the 2 extra sources detected in H24 which are not selected as variables in this work. 80

Figure 4.4 Sky map of variable candidates in the HUDF. The estimated redshift is shown by the colour of the symbol with reference to the colour-bar. Different symbols refer to different significance levels as indicated in the legend. 80

Figure 4.5 Redshift evolution of SMBH abundance. Comoving number density of SMBHs, n_{SMBH}, is plotted versus redshift, z. Estimated counts of candidate Active Galactic Nucleus (AGN)s in the correspondent volume are represented with the filled diamonds with black edges, for the 2 (turquoise), 2.5 (lime) and 3 σ_m (magenta) significance thresholds, in different redshift bins between $z \sim 0-9$. In each bin, we subtract from the row counts the expected number of false detections according to a Gaussian statistics for the correspondent significance levels. These values are corrected first for variability incompleteness by multiplying for the correction fraction detailed in Table 4.2, and secondly for luminosity function incompleteness (see Table 4.4, second column). For this latter, we collect a compilation of fits to the UVLF at several redshifts (see text, *compilation* case). At $z = 0$ we depict the estimated number density of SMBHs from local galaxies (Banik et al., 2019, black square). The red diamond shows the observational constraint derived by Hayes et al. (2024). Broad emission line sources at $z = 4-7$ (Harikane et al., 2023) are shown with the black diamonds. For the $z \sim 4-7$ sources, the fraction of non-active SMBHs remains highly uncertain, and we leave all points as lower limits. Brown stars indicate the number densities from the known AGN sample of Lyu et al. (2022). Since these Various Pop III.1 formation models (Banik et al., 2019; Singh et al., 2023) with isolation distance parameters, d_{iso}, from 50 to 200 kpc (proper distance) are shown by the coloured solid lines. At low redshifts these decrease compared to the maximum value attained (dotted lines) account for merger events. The green dashed line shows the SMBH seeding assumed in Vogelsberger et al. (2014) based on a Halo Mass Threshold (HMT) above a mass of $7.1 \times 10^{10}\ M_\odot$ (being the shaded region a factor of two variation in this mass scale). Models of SMBH formation via *Direct Collapse* from Chon et al. (2016) are depicted by the pink dotted line. 86

Figure 4.6 Redshift evolution of SMBH abundances as in Fig. 4.5, but with luminosity function corrections from Table 4.4, third column, for the *conservative* case of a Schechter LF (see text). 87

Figure A.1 Estimated total survival time in Gyr for seeded halos for the $d_{iso} = 50$ kpc case as a function of the halo mass ratio and color coded with the mass of the most massive (primary) halo. 98

Figure B.1 GSMF predictions based on the Millennium simulation box (blue lines) compared to the GSMF extracted from our Millennium-like PINOCCHIO box run used for the calibration of the merging times (see Section 2.1.4). *Upper panels*: satellite and central galaxies contributions to the GSMF are reported separately in dotted and dashed lines, respectively. *Lower panels*: total GSMFs. Gray symbols show observational estimates as reported in Fig. 2.2. 100

LIST OF TABLES

Table 2.1 Fitting parameters and derived quantities: the slope α, the intercept $M_{BH,9}$ (at $M_{star} = 10^9 M_\odot$), and dispersion σ are shown at z~0. We also report the average (geometric mean) mass of the 10 most massive BHs. 40

Table 2.2 Fitting parameters and derived quantities: the slope α, the intercept $M_{BH,9}$ (at $M_{star} = 10^9 M_\odot$), and dispersion σ are shown at z~5. We also report the average (geometric mean) mass of the 10 most massive BHs. 44

Table 4.1 Observing epochs and times. 72

Table 4.2 Number of matched AGNs against Lyu et al. (2022) across all the three filters. 78

Table 4.3 General properties of the $z > 6$ variable sources at > 2 σ_m significance. [a]This source correspond to source 1052123 in H24. [b]Spectroscopic redshift from NIRspec (Bunker et al., 2023). [c]Spectroscopic redshift from FRESCO (Oesch et al., 2023). 81

Table 4.4 UV luminosity function correction as function of redshift. The second column indicates the correction factors using different compilations of fits to the UVLF at different redshifts. The values derived from the case adopting a conservative Schechter UVLF are reported in the last column (see text). 84

ACRONYMS

ACS Advanced Camera for Surveys
AGB Asymptotic Giant Branch
AGN Active Galactic Nucleus
ALS All Light Seeds
BH Black Hole
BHMF Black Hole Mass Function
BPT Baldwin, Phillips and Terlevich
DCBH Direct Collapse Black Hole
DM Dark Matter
DMH Dark Matter Halo
DMA Dark Matter Annihilation
EHT Event Horizon Telescope
FoF Friends of Friends
FIR Far Infra-Red
FWHM Full Width Half Maximum
GSMF Galaxy Stellar Mass Function
GW Gravitational Wave
HMT Halo Mass Threshold
HMF Halo Mass Function
HUDF Hubble Ultra Deep Field
HST Hubble Space Telescope
HLSP High-Level Science Products
IMBH Intermediate Mass Black Hole
NIR Near Infra-Red
IR Infra-Red
ISM Inter Stellar Medium
IGM Inter Galactic Medium
JWST James Webb Space Telescope
LIGO Laser Interferometer Gravitational-Wave Observatory
LF Luminosity Function
LPT Lagrangian Perturbation Theory
LRD Little Red Dot
MAST Mikulski Archive for Space Telescopes
MSA Micro-Shutter Array
NANOGrav North American Nanohertz Observatory for Gravitational Waves
NFW Navarro-Frenk-White
PTA Pulsar Timing Array
QSO Quasi Stellar Object
SAM Semi-analytical Model
SED Spectral Energy Distribution

SMBH Supermassive Black Hole
SFR Star Formation Rate
SFRD Star Formation Rate Density
SLSN Super Luminous Supernova
SN Supernova
UV Ultraviolet
UVLF Ultraviolet Luminosity Function
WCS World Coordinate System
WIMP Weakly Interacting Massive Particle
WFC-3 Wide Field Camera 3
ZAMS Zero Age Main Sequence

1

INTRODUCTION

Black Holes (BHs) are known to be a fundamental component of our Universe and among the most intriguing objects out there. They were first theorized as early as 1784 by John Michell and shortly after by Pierre-Simon Laplace (Michell, 1784; Laplace, 1796). Since then a new branch of physics has developed and numerous theories about their formation and evolution have flourished until nowadays. They represent a boundary where our understanding of the laws of physics breaks down, offering a unique window into extreme gravity, quantum mechanics, and the nature of the universe itself.

Since the first pioneering theoretical speculations describing them as stars with enough gravitational force to prevent light from escaping, our understanding of the behaviour of these objects has greatly advanced. Only with the development of Einstein's theory of general relativity over one century ago, the modern understanding of BHs began to take shape. An exact solution to the field equations published by Einstein in the General Relativity (Einstein, 1916) was found by Schwarzschild under the simplistic assumptions of spherically symmetry and non rotating mass distribution (Schwarzschild, 1916). The metric of this solution naturally leads to the definition of BH: when the radial size of a body is smaller than the corresponding Schwarzschild radius ($r_s = 2GM/c^2$, with M being the total mass of the body), we refer to this object as a BH. Note that at Schwarzschild radius the radial component of the metric $(1 - r_s/r)^{-1}$ goes to infinity, causing a so called *singularity*. It can be shown mathematically that, regardless of the choice of coordinates, the only singularity of this solution is found at $r = 0$. This location collects the ending point of all the geodesics inside the sphere defined by $r = r_s$, also known as the *event horizon*. For this reason, any event occurring within this region cannot be observed outside as no signal would be able to escape from eventually collapsing in the singularity at $r = 0$.

In order to recover a more general solution, Kerr (1963) published a metric for a rotating, spherically symmetric BH. While the Schwarzschild metric has a point singularity at the centre, the Kerr metric presents a ring-shaped singularity, known as a *ring singularity*. If the spin is zero, the Kerr metric reduces to the Schwarzschild solution. Shortly after Kerr's work, the solution was extended to account for a charged, rotating BH, resulting in the Kerr-Newman metric (Newman and Janis, 1965; Newman et al., 1965). Regardless the formation process, physically speaking we refer to a BH as a fairly simple object. Via the no/hair theorem, a BH can be fully described by means of three independent properties — mass, angular momentum

and electric charge. This naturally implies that any meaningful insights about its formation and its matter composition is inevitably lost. What we are left with in order to investigate the formation processes is to study the spatial and mass distribution of BHs, as well as how they affect the physical properties of the environment they reside in.

Over the decades, BHs have transitioned from theoretical curiosities to well-established cosmic objects. There are several types of BHs, categorized primarily by their mass. Stellar mass BHs, formed from the gravitational collapse of massive stars, typically range from a few to several tens of solar masses. Today, our understanding of BH formation through stellar collapse has significantly advanced. When a star's internal pressure — maintained by nuclear fusion — runs out, gravity takes over, causing the star to collapse inward. Depending on the star's initial mass, the remnant left behind could become a white dwarf, neutron star, or in fact, a BH. The collapse of massive stars is one of the primary processes that produce stellar mass BHs. These objects are scattered throughout the Universe, residing in binary systems or as isolated remnants of supernovae. On the other end of the mass spectrum, at the centres of galaxies lie SMBHs with masses ranging from millions to billions of times that of the Sun. The formation and growth of these cosmic giants remain an active area of research, with many questions still unresolved.

The origin and existence of BHs have long been the focus of extensive debate. Observational evidence, from the motion of stars around invisible massive objects to the detection of gravitational waves from merging BHs, has consolidated their role in astrophysics. Recently, the Laser Interferometer Gravitational-Wave Observatory (LIGO) Collaboration (LIGO Scientific Collaboration et al., 2015; Abbott et al., 2016) has detected multiple BH mergers, with masses reaching up to $\sim 150 M_\odot$ (Abbott et al., 2020a,b). The Event Horizon Telescope (EHT) Collaboration made history in 2019 by capturing the first-ever image of a BH's shadow in the Messier 87 galaxy, offering visual confirmation of these objects and their behaviour as predicted by relativity (see Fig. 1.1, 2019). It took only few more years to successfully repeat the experiment, in 2022, when the EHT Collaboration captured the Einstein ring cast by the radio emission from the gas surrounding the central object from our own Milky Way (2022). Both these BHs, with masses ranging from $\sim 10^6 - 10^9 M_\odot$, sit at the higher end of the BH mass spectrum, i.e. they are SMBHs. By monitoring a set of 67 pulsars for over 15 years, the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) Consortium together with other Pulsar Timing Arrays (PTAs) collaborations[1] recently found an evidence of a Gravitational Wave (GW) background claimed to be due to SMBH mergers (Agazie et al., 2023a,b).

In this Thesis, we will explore various aspects of SMBHs, from their formation mechanisms and observational signatures to their impact on

1 The European PTA (EPTA Collaboration et al., 2023), the Parkes PTA (PPTA; Reardon et al., 2023) and the Chinese PTA (CPTA; Xu et al., 2023)

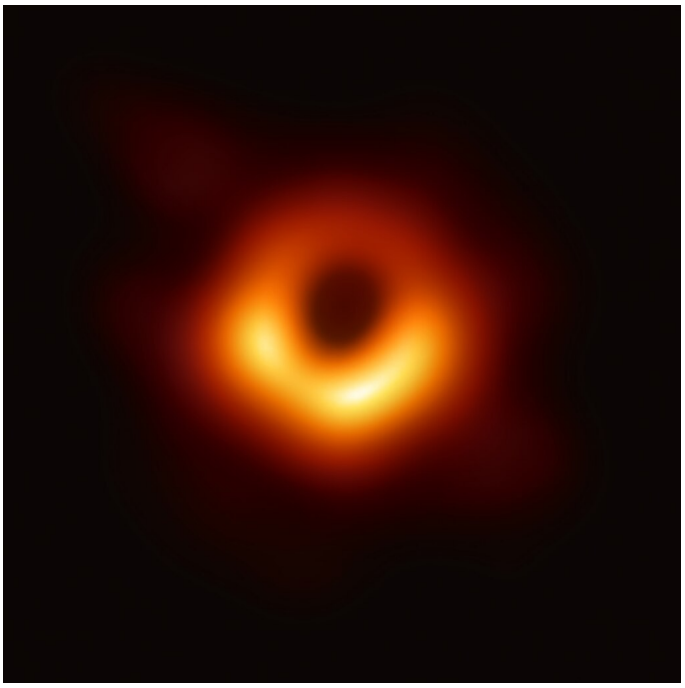
surrounding matter and potential for future discoveries. By delving into both the theoretical foundations and the latest observational breakthroughs, we aim to contribute to the ongoing quest to understand these cosmic powerhouses.

1.1 SUPERMASSIVE BLACK HOLES

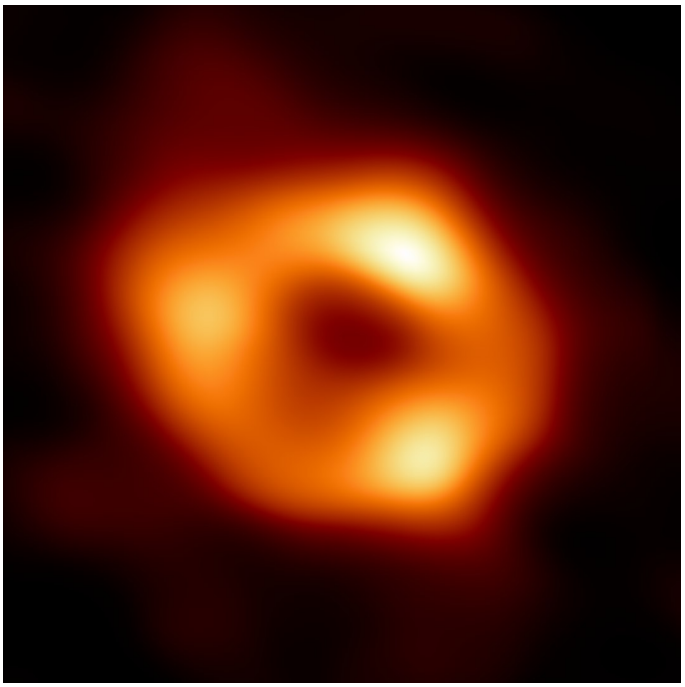
While the origin of stellar mass BHs is relatively well understood, their supermassive counterparts (SMBHs) offer plenty of different interpretations as the formation scenarios are still suffering from poor observational constraints. SMBH are often witnessed at the centre of the almost entire population of massive galaxies, with masses stretching from several tens of thousands up to billions of solar masses (e.g., Ferrarese and Ford, 2005; Volonteri, 2010; Graham, 2016; Inayoshi et al., 2020; Volonteri et al., 2021; Lusso et al., 2023). This mass interval defines the SMBH regime. Their origin is certainly not stellar as they populate a distinct range of the mass spectrum. In fact, as the lightest estimated SMBH mass via dynamical measurements is about $6.800 M_\odot$ in the NGC 205 galaxy (Nguyen et al., 2019), the most massive stellar BHs was detected via gravitational waves in a merger event, carrying a mass of approximately $150 M_\odot$ (GW190521, Abbott et al. (2020a)). Interestingly, observations show an apparent dearth of detected BHs with masses spanning from few hundreds up to several thousands of solar masses. This gap between stellar BHs and SMBHs is usually referred to as the Intermediate Mass Black Hole (IMBH) regime.

The mystery deepens when we examine observations of high-redshift quasars. Yang et al. (2020) and Wang et al. (2021) discovered supermassive black holes (SMBHs) with masses $\gtrsim 10^9 M_\odot$ in extremely bright quasars, J1007+2115 and J0313-1806, at redshifts greater than 7, when the universe was less than 1 billion years old. Additionally, Wang et al. (2019) identified a quasar at $z \sim 6.6$ hosting a SMBH with a mass estimated by Yang et al. (2021) to be around $3.59 \times 10^9 M_\odot$, making it one of the most massive black holes discovered at $z > 6.5$ to date. More recently, the James Webb Space Telescope (JWST)[2] has paved the way for astonishing discoveries regarding the earliest SMBHs. Notably, a growing number of detected galaxies appear to host black hole candidates with masses ranging from $\sim 10^6 - 10^8 M_\odot$ at epochs between $z = 4 - 7$. Among many works, but not limited to, Maiolino et al. (2023) and Harikane et al. (2023) reported a surprisingly large number of faint AGNs in this redshift range, raising questions about their ubiquity in the early universe. Moreover, the recently discovered galaxy GN-z11 has been shown to potentially contain a SMBH with a mass of around $1.6 \times 10^6 M_\odot$ at a redshift of $z \sim 10.6$, as detailed by Bunker et al. (2023) and Maiolino et al. (2023). These recent findings suggest that the cosmic co-evolution between black holes and their host galaxies at these early times likely differs significantly from that of lower-redshift systems. Furthermore, at least a fraction of these early black holes may have formed

2 https://webbtelescope.org/



(a) The SMBH M87* at the centre of Messier 87 galaxy, ~ 16.4 Mpc away from the Sun.



(b) The SMBH Sagittarius A* at the centre of Milky Way, ~ 8 kpc away from the Sun.

Figure 1.1: The images of BHs taken by the Event Horizon Telescope Collaboration (2019; 2022).

and grown rapidly, reaching the SMBH stage when the universe was only a few hundred million years old. On the other hand, the presence of SMBHs in galaxies such as the Milky Way and low-mass SMBHs in some dwarf galaxies today supports a wide mass spectrum of SMBH formation, but also indicate that some SMBHs have not grown much since their birth.

Explaining all these elements constitutes an intricate puzzle for both theoretical models and observational findings. First, SMBHs must have grown in mass over time. In fact, observations have shown that most BHs have accreted mass by several orders of magnitude (e.g., Salucci et al., 2000; Yu, 2002; Marconi and Hunt, 2003). Also, it seems unrealistic to form stellar BHs with masses up to billions of solar masses leading directly to the observed SMBH mass distribution. It is useful to visualise the dispute from a global perspective. There are two main ingredients that are at play in the ecology of SMBH formation: the initial mass of the BH, also known as the *seed* BH, and the *accretion* channel onto the BH. Fig. 1.2 (Wang et al., 2021), illustrates the growth trajectories of some of the most massive and earliest quasars, under the assumption of Eddington-limited accretion throughout their lifetimes. For the most massive black holes, even if they formed very early, at redshifts around $z \sim 30$, the seed black hole mass would need to exceed $10^4 M_\odot$ to reach their observed size. If their formation occurred later, even larger initial seed masses would be required. However, indefinitely pushing the formation epoch earlier to lower the seed mass is not feasible either, as the formation of dark matter halos requires linear density perturbations to reach the critical density, and favourable conditions for star formation and subsequent black hole collapse only arise afterward. Although models of super-Eddington accretion have been suggested (e.g., Kohri et al., 2022), simulations show that the gas supply available to early SMBHs can be substantially limited by feedback from star formation, making sustained Eddington-limited accretion rates unlikely (e.g., O'Shea et al., 2005; Jeon et al., 2023). These challenges highlight the need for models that explain the formation of black holes at the supermassive scale of $\sim 10^5 M_\odot$.

1.2 CHANNELS OF SMBH SEED FORMATION

Most of the proposed formation pathways of SMBHs were first ventured in the pioneering study of Martin Rees (Rees, 1978). The mechanisms we study today are still based on this picture: the idea is that a "gas cloud" in its primordial phase would lead, via different processes, to the formation of a seed BH (Fig. 1.3). In the previous Section we introduce the concept of *seed* BH as the initial stage of a newly born BH, which will then grow in mass and become detectable either via GW emitted in mergers or via electromagnetic radiation powered by gas accretion. In literature, we usually classify the seeds depending on their mass. *Light* seeds are identified by masses spanning from $\sim 10^2 M_\odot$ up to $few \times 10^3 M_\odot$, while we usually refer to masses between 10^4 and $10^6 M_\odot$ as *massive* seeds. Finally, *intermediate* mass seeds lie in the gap between light and massive ones.

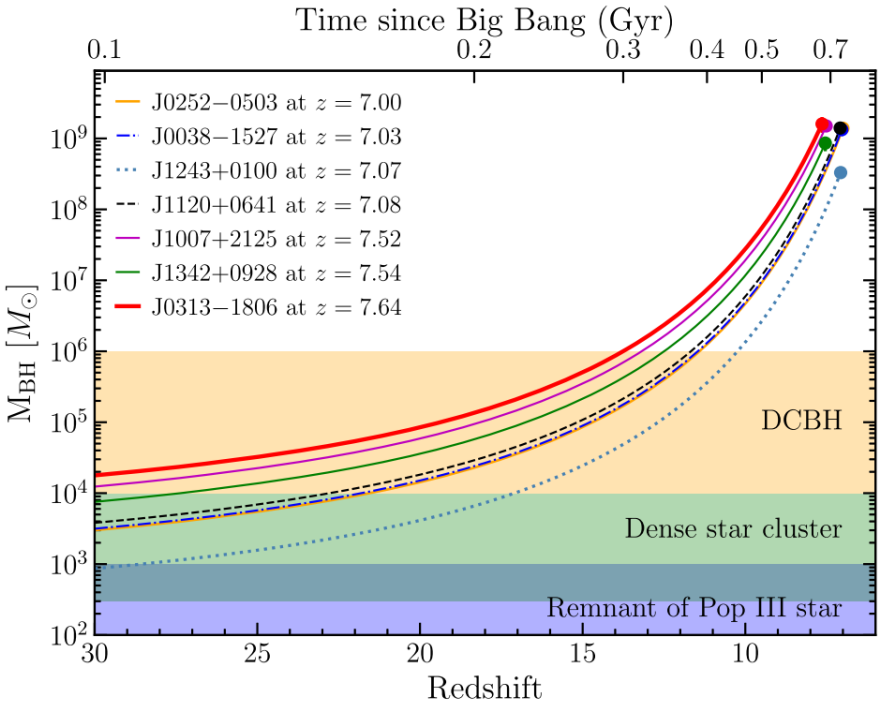


Figure 1.2: Mass accretion history of Eddington limited growth tracks of confirmed high redshift quasars ($z \geqslant 7$) assuming a constant radiative efficiency of 0.1. The shaded horizontal areas indicate the range of seed masses of the BH according to different formation mechanisms. Credit by Wang et al. (2021).

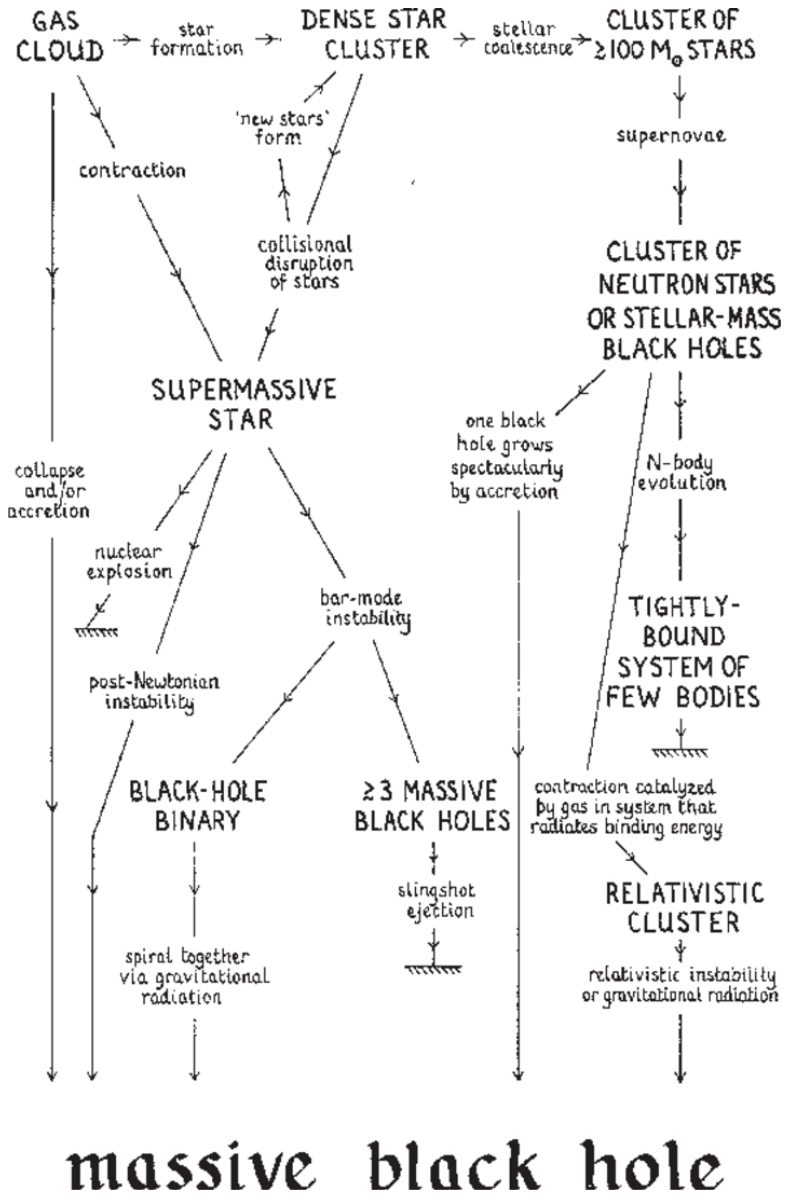


Figure 1.3: Original diagram of formation pathways proposed by Rees (1978).

1.2.1 *Gas and stars-based mechanisms*

The DCBH scenario advances the idea that a primordial gas cloud hosted by a relatively massive, atomically-cooled, UV-irradiated halo of $\sim 10^8 M_\odot$ collapses into a single supermassive star of $\sim 10^{4-6} M_\odot$, subsequently forming a massive seed by $z \sim 10$ (e.g., Bromm and Loeb, 2003; Lodato and Natarajan, 2006; Shang et al., 2010; Montero et al., 2012; Maio et al., 2019; Bhowmick et al., 2022). In some cases, the collapse may go through an intermediate phase where a quasi-star forms. Instead of being powered by nuclear fusion, this quasi-star is fuelled by the accretion of gas onto a small, growing BH at its centre (Begelman et al., 2006; Begelman et al., 2008). For these supermassive black hole seeds to form, a crucial condition is that accretion continues without the gas fragmenting as the protostar grows on a timescale shorter than the thermal, Kelvin-Helmholtz timescale (Haemmerlé et al., 2020). If this can occur, the surface of the protostar stays cool, which limits ultraviolet radiation emission and prevents disruption of the accretion flow. Moreover, uninterrupted accretion without fragmentation requires metal-free halos, which is naturally the case since these halos consist of primordial gas. However, molecular hydrogen formation must be suppressed to prevent fragmentation, as molecular hydrogen acts as a cooling mechanism (Latif et al., 2013) and strong Lyman-Werner radiation is necessary to photodissociate molecular hydrogen (Wise et al., 2019). Habouzit et al. (2016) showed that the rarity of all these conditions combined makes the formation of SMBH via this process relatively uncommon. In addition, cosmological studies have underlined that such conditions are favourable mostly in satellite halos, making the ubiquitous presence of SMBH at the centre of massive galaxies even more challenging by requiring short sinking times scales (Chon and Omukai, 2020).

While the number density of black holes resulting from direct collapse could account for the currently known population of high-redshift quasars, the specific conditions required for this process are considered too rare to explain the entire population of supermassive black holes observed at $z = 0$ (Chon et al., 2016; Wise et al., 2019). Additionally, recent simulations suggest that the supermassive stars forming through this mechanism may not be as large as originally expected: instead of reaching the previously predicted masses, they may only grow to around $\lesssim 10^4 M_\odot$, due to the turbulence in the early stages of galaxy formation that disrupts the accretion flow (Regan et al., 2020).

Another possible pathway to form intermediate or even massive seeds involves runaway stellar mergers in young, dense, metal-poor stellar clusters. These mergers can lead to the formation of stars with masses ranging from $\sim 200 - 10^3 M_\odot$ (e.g., Portegies Zwart et al., 2004). Such large masses can be attained through repeated collisions, especially if massive stars can migrate to the cluster's core, significantly increasing the collision rate before undergoing supernova explosions (Ebisuzaki, 2003). Additionally, scenarios involving gas-driven accretion in dense stellar-mass black hole

clusters have been proposed as a mechanism for forming supermassive black holes (Kroupa et al., 2020), alongside the idea that repeated mergers between minihalos of about $10^5 M_\odot$ could create favourable conditions for SMBH formation (Katz et al., 2015). However, predicting the frequency and likelihood of such dense clusters forming in galaxies is highly challenging. This is primarily because it requires resolving the high resolution processes of individual star formation and evolution, which introduces significant uncertainties and limitation in forecasting the cosmological population of these systems (see, e.g., Boekholt et al., 2018; Chon and Omukai, 2020; Tagawa et al., 2020).

In the early Universe, no metals were present except traces of lithium. This metal-poor environment supposedly created ideal conditions for matter to accrete without significant fragmentation and form the first generation of stars. These stars, known as Population III (Pop III) stars, are thought to become potential progenitors of SMBH seeds. In this picture, the primary coolant is molecular hydrogen, which is far less efficient than heavier elements. As a result, the temperature of the contracting primordial gas remains higher than that of the metal-rich interstellar medium in the local Universe, leading to reduced fragmentation. However, conventional models of Pop III star formation predict stellar masses of around $\sim 100\,M_\odot$ (e.g., Madau and Rees, 2001; Abel et al., 2002; Bromm et al., 2002; Tan and McKee, 2004; McKee and Tan, 2008a; Hosokawa et al., 2011; Susa et al., 2014). Such stars would only be capable of forming stellar-mass black holes, resulting in relatively low-mass black hole seeds and making their role as SMBH seeds implausible, even though we would expect to find a consistent number of Pop III remnants in the large majority of galaxies.

One challenge facing models that start with light or intermediate mass seeds is the apparent dearth of observed IMBHs in the local universe (Banik et al., 2019; Greene et al., 2020; Volonteri et al., 2021). In addition, it is difficult to model the IMBH seeding in a cosmological context, unless requiring either sub-grid or probabilistic recipes calibrated on high-resolution zoom-in simulations as recently studied by Bhowmick et al. (2024c,b).

1.2.2 *Extreme physics and early Universe mechanisms*

In this Section we briefly mention a class of formation models that have been recently proposed and invoke exotic physics, from different flavours of the standard dark matter paradigm to primordial fluctuation in the early Universe. However, regardless of the appealing properties of these mechanisms, they often rely on ad-hoc assumptions not yet observationally constrained.

If dark matter undergoes self-interaction, then this could provide a mechanism for SMBH seeding via collapse of the halos themselves (e.g., Feng et al., 2021). In fact, the presence of baryons in protogalaxies can significantly speed up this gravothermal evolution, reducing the collapse timescales.

The angular momentum in the central region of the halo can be dissipated through viscosity resulting from the self-interactions. However, for this scenario to produce high-redshift supermassive black holes (SMBHs), the host halos must reside in regions of high-density fluctuations. As a result, SMBHs formed through this mechanism are expected to be rare at high redshifts.

Primordial BHs provide another extreme yet viable formation pathway. They are postulated to take place in the presence of high density contrast fluctuations and the mass of the seed equals the mass of the horizon at the collapse time. Therefore their initial mass can span several *tens* of order of magnitude, from grams up to the ten thousands of solar masses (e.g., García-Bellido, 2019). However, this scenario appears to be disfavoured when requiring not to violate the total matter density (e.g., Carr and Kühnel, 2020) and by the clustering analysis of Shinohara et al. (2023).

1.3 POP III.1 MODEL

Despite the fact that the seeding mechanisms detailed in Section 1.2 are not mutually exclusive, they all tend to predict too few SMBH seeds to be consistent with the observed number density in the local Universe (Banik et al., 2019), hence missing a significant population of SMBHs. In this Thesis, motivated by the above considerations, we focus on a formation mechanism based on Pop III.1 stars as the progenitors of SMBHs. These stars represent a subclass of Population III stars, which were classified by McKee and Tan (2008a) into two categories. Pop III.1 sources are defined as special Pop III stars forming at the centre of dark matter (DM) minihalos in the early universe ($z \gtrsim 20$), which are isolated from any source of stellar or SMBH feedback (McKee and Tan, 2008b). Pop III.2 stars, on the other hand, form in similar dark matter minihalos but are influenced by feedback from external astrophysical sources, which favours gas fragmentation, resulting in lower-mass stars compared to Pop III.1 stars. In semi analytical models, Tan and McKee (2004) and McKee and Tan (2008a) showed that Pop III.1 stars can grow up to $\sim 140 M_\odot$. Greif and Bromm (2006a) also argue similar subdivision of Pop III stars, and present $\geqslant 100\,M_\odot$ masses for Pop III.1 stars, while an order of magnitude less for the Pop III.2 stars. However, simulations by Tan et al. (2010), based on the semi-analytical models of McKee and Tan (2008a) and the simulations of O'Shea and Norman (2007), indicate that the initial mass function (IMF) of Pop III stars peaks around $\sim 100\,M_\odot$, with a tail stretching to $\sim 10^3\,M_\odot$.

Although the upper end of this IMF includes massive stars, it still falls short of explaining the formation of SMBHs with masses on the order of $\sim 10^5\,M_\odot$. As discussed previously, explaining the observed high-redshift quasars would require seed black holes of around $10^3\,M_\odot$ to grow through Eddington-limited accretion for their entire lifetime. However, sustaining such high accretion rates is highly improbable, as radiative and mechanical feedback from the Pop III.1 star would likely disrupt its surroundings and

hinder the continuous growth of the black hole (O'Shea et al., 2005; Johnson and Bromm, 2007; Milosavljević et al., 2009).

One possible physical mechanism that allows Pop III.1 protostars to grow up to high masses, i.e., $\gtrsim 10^4$ $M_\odot$, is the influence of Dark Matter Annihilation (DMA) on the protostellar structure (Spolyar et al., 2008; Natarajan et al., 2009; Freese et al., 2010; Rindler-Daller et al., 2015). The fiducial assumption is that dark matter is composed of Weakly Interacting Massive Particles (WIMPs), which undergo self-annihilation with a weak interaction cross-section. If sufficient WIMPs are captured by the Pop III.1 protostar, then its structure is altered. In particular, the protostar can remain relatively large as it accretes to high masses, thus reducing its ionizing feedback on its own accretion flow. This may enable the efficient accretion of a large fraction of the entire baryonic content of the parent minihalo, i.e., $\sim 10^5$ $M_\odot$, to the Pop III.1 protostar, which subsequently, within a few Myr, collapses to a SMBH. Other minihalos are Pop III.2 sources, i.e., still metal free, but, having been irradiated by UV radiation, have higher free electron abundances leading to greater abundances of H_2 and HD, higher cooling rates, and fragmentation to lower mass, $\sim$ 10 $M_\odot$, stars (Johnson and Bromm, 2006; Greif and Bromm, 2006b).

This formation scenario is a valid alternative to "standard" seeding schemes, and has been applied in cosmological simulations in previous works. Banik et al. (2019) first (hereafter Paper I) and Singh et al. (2023) then (hereafter Paper II) followed the redshift evolution of SMBH seeds and host halos down to the local Universe. In the next paragraph we review the main results and implications of the Pop III.1 model from previous studies.

The Pop III.1 model predicts a characteristic mass of $\sim 10^5 M_\odot$ for SMBH seeds and provides an explanation for the apparent dearth of IMBHs. In addition, the Pop III.1 model foresees that SMBHs form very early in the universe, which suggests a theoretical explanation for the population of high-z SMBHs without the need for sustained Eddington or super-Eddington levels of accretion. Another prediction is that the initial spatial distribution of SMBHs is relatively uniform, i.e., with low levels of clustering, with the seeds separated from each other by the isolation criterion described above. This criterion is parametrized by an isolation distance d_{iso}, expressed in physical units, of the order of $\lesssim$ 100 kpc. The co-moving number density of SMBHs, n_{SMBH}, is sensitive to d_{iso}, with a value of about 100 kpc able to explain the observed local number density of SMBHs of $\sim 5 \times 10^{-3}$ Mpc^{-3}, as shown in Fig. 1.4. The estimate at $z = 0$, presented in Paper I, is based on the assumption that each galaxy with a luminosity greater than $L_{min} = 0.33$ L_* hosts an SMBH. The value of d_{iso} also sets a limit on the formation epoch of SMBHs. For $d_{iso} = 50 - 100$ kpc, most seeds are formed at $z \sim 30$, i.e., with co-moving separations of $\sim$ *few* Mpc, and the process is largely complete by $z \sim 25$, after which n_{SMBH} is nearly constant. This time scale has a weak dependence on d_{iso}, with later seeding happening for smaller d_{iso}. As a result of the initially separated distributions, in the Pop III.1 model BH mergers only begin to occur at relatively late times, i.e.,

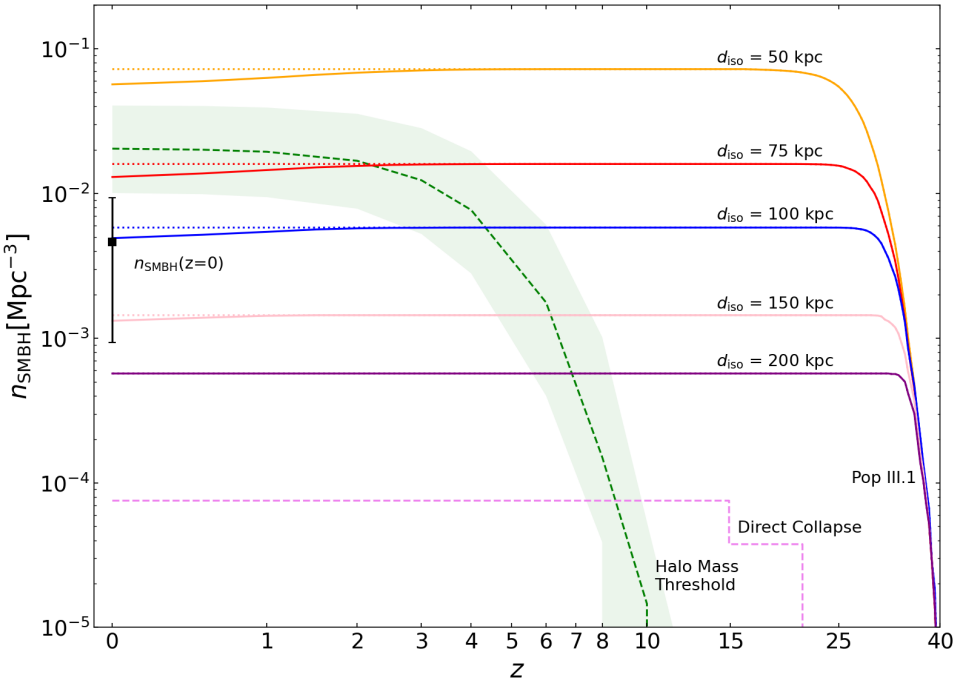


Figure 1.4: Evolution of the comoving number density of SMBHs, n_{SMBH}, as a function of redshift. Results for the Pop III.1 models with different isolation distances (in proper distance) are shown in coloured lines, see labels. The dotted lines represent the total number of SMBHs that have formed, while the solid lines show the remaining number of seeded halos after accounting for mergers. Additionally, a sample Halo Mass Threshold (HMT) model is depicted by the dashed green line, where each halo with mass greater than $m_{th} = 7.1 \times 10^{10} M_{\odot}$ is seeded. The pink dashed line shows the results, extrapolated to $z \sim 0$, of Chon et al. (2016) assuming DCBH forming in a simulation. The black solid square displays an estimate for the number density of SMBHs at $z = 0$ (see text). Adapted form Singh et al. (2023).

$z \lesssim 1$, which causes n_{SMBH} to decrease by modest amounts $\sim 10 - 20\%$ (Singh et al., 2023). Consequently, the SMBH number density remains fairly constant down to the local Universe, with only a small fraction of seeds lost in halo mergers by $z = 0$. Furthermore, by redshift $z \sim 0$ the occupation fraction of seeded halos saturates to unity for the most massive halos as a function of the isolation distance d_{iso}. At redshifts $z \gtrsim 5$, this model makes strong predictions for the number of SMBHs we expect in the early Universe (e.g., their relative contribution to the luminosity functions), underscoring the importance of current and future deep field observations of high redshift AGNs. Such observations are crucial for potentially distinguishing among different seeding mechanisms.

1.4 SEEDING IN SIMULATIONS

Numerical simulations of galaxy formation and evolution in cosmological volumes typically have quite limited spatial resolution and struggle to

capture the processes leading to SMBH seeding. The lack of knowledge about the physical processes and the difficulty to treat them from first principles further limit such numerical approaches. In addition, the real challenge resides in simultaneously simulating a broad range of scales, from mini-halos at $z \sim 20 - 30$ up to the structures we observe in the local Universe. Thus, implemented seeding schemes are generally based on simple threshold models. In particular, many works seed a SMBH once the parent halo reaches a certain threshold in dark matter mass, i.e., a HMT seeding scheme (Sijacki et al., 2007; Matteo et al., 2008). The same idea has been used by Vogelsberger et al. (2014) within the Illustris Project in which a SMBH of mass 1.4×10^5 $M_\odot$ is placed in each halo crossing a mass threshold of $m_{th} = 7.1 \times 10^{10}$ $M_\odot$. In the Evolution and Assembly of GaLaxies and their Environments (EAGLE) simulations, Barber et al. (2016) implemented a similar approach.

These simplistic mechanisms are primarily influenced by the mass resolution of the simulations: because the early stages of galaxy formation cannot be resolved, seed BHs are placed in halos as soon as they are detected with a minimum threshold of particles, and therefore neglecting their prior assembly and accretion history.

Extending to alternative threshold models, in Horizon-AGN simulation (with Dark Matter (DM) mass resolution of 8×10^8 $M_\odot$, Volonteri et al., 2016) implemented lower limits for gas and stellar densities, as well as stellar velocity dispersion, to determine if a galaxy hosts a BH, using a seed mass of 10^5 $M_\odot$. Their formation was restricted to redshifts $z > 1.5$ and all forming BHs had to be separated by at least 50 comoving kiloparsecs to prevent multiple BHs from forming in the same galaxy. Similarly, the `OBELISK` simulation (Trebitsch et al., 2021, based on a sub-volume of the `HORIZON-AGN` simulation with mass resolution of 1.2×10^6 $M_\odot$) used a slightly lower seed mass and applied gas and stellar density thresholds, including gas Jeans instability, requiring an isolation of 50 kpc from other SMBHs to prevent multiple BH formation. In another approach, the `ROMULUS` simulation (Tremmel et al., 2017, particle DM mass $\sim 3 \times 10^5$ $M_\odot$) set criteria based on metallicity limits, gas density thresholds, and a restricted temperature range for SMBH formation, using a seed mass of 10^6 $M_\odot$. Additionally, in the Illustris Project framework, Bhowmick et al. (2022) explored various gas-based SMBH seeding prescriptions and a range of seed masses from around 10^4 to 10^6 $M_\odot$, while keeping a relatively high DM resolution about $\sim 6 \times 10^6$ $M_\odot$. Moreover, in high-resolution zoom-in simulations Bhowmick et al. (2024c,b) and Bhowmick et al. (2024a) explicitly resolved pristine and dense gas clouds forming $\sim 10^3$ $M_\odot$ seeds and built a stochastic seeding scheme that can directly set the initial seeding conditions in lower resolution runs. Using the hydrodynamical cosmological code `RAMSES` , Habouzit et al. (2016) investigated the conditions under which halos can host DCBHs over a large range of resolutions and box sizes as a function of the illuminating Lyman-Werner (LW) background and Supernova (SN) feedback. Under optimistic assumptions, their SMBH number density ranges from 7×10^{-7}

up to 10^{-4} cMpc^{-3}, still a factor of at least $\sim$ 10 lower than the local estimate.

More generally, the latest efforts in simulations have made it possible to cover the whole range of theoretically predicted seed masses, from 10^2 up to 10^6 $M_\odot$. Also, while a constant seed mass is more commonly adopted, recent developments have investigated more sophisticated schemes by considering a distribution of initial seed masses extending from light to heavy seeds (e.g, Ni et al., 2022; Habouzit et al., 2017). In any case, the choice of the seed mass is highly degenerate with the accretion model onto the central engine, particularly in very massive galaxies where the BH can grow by several orders of magnitude and lose memory of the its initial mass by $z \sim 0$. Conversely, lower mass galaxies, far more abundant than their massive counterparts, can hide a significant population of dormant or stunted BHs which would be informative about the initial seed mass. However, despite the improvements in the treatment of sub-grid physics and the implementation of zoom-in approaches, large, cosmological simulations do not include yet a full physical model for SMBH formation (e.g., see Di Matteo et al., 2023).

1.5 SEMI-ANALYTICAL MODELS OF GALAXY FORMATION AND EVOLUTION

A complementary approach with respect to hydrodynamical simulations is provided by Semi-analytical Models (SAMs). SAMs are tools used to simulate the evolution of galaxy populations within Dark Matter Halo (DMH)s by modelling the physical processes that drive the evolution of the baryonic components of dark matter halos by invoking theoretically and/or observationally motivated prescriptions. We stress here that such prescriptions (*sub-grid* physics) are similarly implemented in hydrodynamical simulations. These processes encompass baryonic gas cooling and heating, star formation, gas accretion onto SMBHs, and their related feedback mechanisms. The flexibility of these models enables us to obtain predictions of galaxy properties across cosmological volumes and allows efficient exploration of the associated parameter space in order to study the impact of different physical assumptions (for a review see Somerville and Davé 2015; De Lucia 2019).

On the other hand, the price to pay is the lack of a complete and detailed treatment of the gas hydrodynamics. One limitation of the application of SAMs to the study of the early BH seeding lies in the lack of resolution in dark matter halo trees generated with either N-body simulation or analytical recipes (e.g., from Press and Schechter, 1974; Lacey and Cole, 1993). In fact, dealing with cosmological volumes significantly impacts the modelling of BH seed formation, which depends on the local gas conditions within halos and the DM mass resolution.

1.5.1 *Seeding in SAMs*

Among the BH seeding mechanisms implemented in SAMs, Ricarte and Natarajan (2018) used DCBHs whose mass is calculated proportionally to the DM halo mass. Given their redshift-dependent mass resolution scheme, this results in placing $\gtrsim 10^4$ $M_\odot$ BH seeds at $z \sim 15-20$ in parent halos of minimum mass about 5×10^6 $M_\odot$. In the Cosmic Archaeology Tool (`CAT`) presented in Trinca et al. (2022), both light seeds from Pop III stars and DCBHs are considered to study their contribution to the BH mass function and Luminosity Function (LF) at $4 \lesssim z \lesssim 6$. Similarly, BH seed masses ranging from 10^2 up to 10^5 $M_\odot$ were assigned in the `DELPHI SAM` (Dayal et al., 2019) according to the probability of a halo to host a DCBH or a stellar BH remnant. Sassano et al. (2021) seeded BHs at the centre of galaxies based on the locally derived properties of the halo environment. In particular, several thresholds for the illuminating Lyman-Werner flux, the metallicity and the gas-to-dust ratio were used to determine whether the final BH would be a light, medium or heavy seed. In the `L-GALAXIES` model, Spinoso et al. (2022) studied DCBHs according to spatial variations of the star formation in terms of chemical and radiative feedback. Additionally, since the first mini-halos ($T_{vir} \sim 10^3$ K) lie below their resolution limit ($L_{box} = 100$ Mpc/h), Pop III remnant seeds were placed following the sub-grid approach of Sassano et al. (2021).

1.6 HIGH-REDSHIFT AGNS AND THE HUBBLE ULTRA DEEP FIELD

Observational constraints on SMBH formation theories are clearly needed, as they provide foundational input into galaxy formation models. The most informative observations are likely to come from the largely uncharted epoch prior to $z \sim 6$. Known $z \sim 7.5$ quasars are over 4 orders of magnitude less abundant than galaxies (McLure et al., 2013), but these rare objects trace only the most extreme accretion rates. Recent James Webb Space Telescope (JWST) observations have identified impressive samples of AGN at luminosities more consistent with ordinary ($\simeq L^\star$) galaxies which volume densities appear to be much larger than predicted by extrapolating the luminosity function of bright AGNs down to fainter luminosities (Maiolino et al., 2023; Harikane et al., 2023; Larson et al., 2023). Interestingly, unexpected *Little Red Dots* (LRDs) have been identified at $z > 4$ (Matthee et al., 2024; Wang et al., 2024; Kokorev et al., 2024; Inayoshi and Ichikawa, 2024; Akins et al., 2024; Maiolino et al., 2024), with a smoothly rising red rest-frame optical Spectral Energy Distribution (SED), blue Ultraviolet (UV) continuum and often broad Balmer lines. While it is not clear whether the majority of LRDs are AGNs, a substantial fraction definitely are. Moreover, with volume densities of $\gtrsim 10^{-5}$ cMpc^{-3}, they are expected to significantly contribute to the total SMBH number density and they could help us constrain SMBH seeding mechanisms.

However, despite recent advances, it remains immensely challenging to obtain an accurate census of AGNs at $z \gtrsim 6$ at faint magnitudes. Hot dust tracers (e.g., Stern et al., 2005) are shifted to the Far Infra-Red (FIR), and X-ray and radio facilities are not sufficiently sensitive to identify individual AGN at these distances. High ionization UV emission lines (e.g., He II λ1640, C IV λ1550, N V λ1240) are sometimes observed in luminous targets (e.g., Mainali et al., 2018; Laporte et al., 2017), but they can be inconclusive, as these lines also form in low-metallicity starbursts (Senchyna et al., 2019; Berg et al., 2019; Saxena et al., 2020). NIRSpec has delivered deep IR spectroscopy and several studies (e.g., Cameron et al., 2023; Boyett et al., 2024; Hu et al., 2024) could identify AGN via line flux ratios classification (Baldwin, Phillips and Terlevich (BPT)-like diagrams, Baldwin et al., 1981), although the targeted approach of Micro-Shutter Array (MSA) spectroscopy are always subject to photometric pre-selection biases and will be focused more towards the bright end of the LF. In addition, NIRSpec surveys will face challenges in achieving completeness at $M_{UV} \simeq -17$ to -20, where galaxies are ~ 100 times more abundant, and occupy densities that can distinguish between SMBH seeding mechanisms.

There is one remaining observable that can identify AGNs for all $z \gtrsim 6$ galaxies: their *time* variability. Cohen et al. (2006), Pouliasis et al. (2019) and O'Brien et al. (2024) have all shown the effectiveness of using photometric monitoring with HST to identify variable sources, finding large numbers of AGNs at intermediate luminosities, that evade other selection techniques (X-ray, radio, IR colours).

The deepest field for which there is a long history of exquisite Hubble Space Telescope (HST) observations is the Hubble Ultra Deep Field (HUDF)[3] (Fig. 1.5). In order to probe variability to the highest redshifts, near-infrared imaging with the Wide Field Camera 3 (WFC-3) is essential and the two most relevant historic datasets are the HUDF09 (GO 11563, PI: Illingworth) and HUDF12 (GO 12498, PI: Ellis) imaging campaigns. Crucially, in Hayes et al. (2024) we recently re-imaged the HUDF with HST in the NIR with a campaign taken in 2023 to an equivalent depth as HUDF12 in the F140W filter (GO 17073, PI: Hayes, hereafter HUDF23), leveraging the 10 years that have elapsed since completion of the original datasets. We found three certain AGN and nine further candidates at $6 < z < 7$, providing a new and completely independent lower limit on the SMBH density at high redshift. At a value of $n_{SMBH} = 8 \times 10^{-3} cMpc^{-3}$, this is the largest density so far reported at these redshifts.

This thesis aims to further investigate the Pop III.1 seeding model from two perspectives: (i) by applying it to galaxy formation and evolution using semi-analytic modelling extended to $z = 0$, and (ii) by searching for variable AGN candidates in high-redshift observations within the HUDF. Firstly, starting from a standard ΛCDM cosmological picture, the `PINOCCHIO`

3 https://esahubble.org/images/heic0611b/

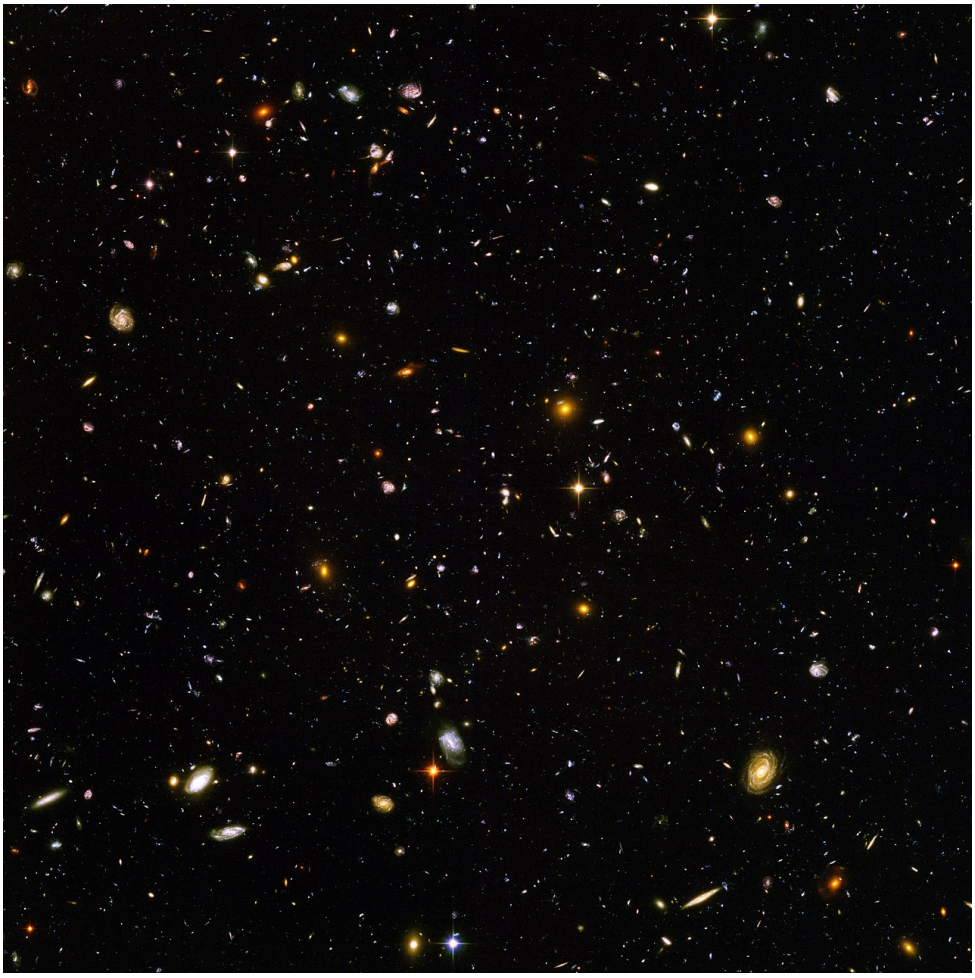


Figure 1.5: Deep combined stacked image of the HUDF, captured by the HST across multiple bands between Sept. 24, 2003 and Jan. 16, 2004. This view required over 400 Hubble orbits around the earth collecting ~ 800 exposures. Nearly 10,000 galaxies populate this field. Credits: Beckwith et al. (2006).

code (Monaco et al., 2002; Munari et al., 2017) traces the evolution of cosmic perturbations and structures based on Lagrangian Perturbation Theory (LPT) and generates catalogues of collapsed objects (i.e. dark matter halos). These halos are eventually seeded with SMBHs according to a specific seeding mechanism. By following the redshift evolution of these structures, we populate them with visible galaxies by exploiting the state-of-the-art SAM GAEA of galaxy formation and evolution (Hirschmann et al., 2016; Fontanot et al., 2020; De Lucia et al., 2024), including accretion onto the central SMBH. This study focuses on three key aspects of SMBHs and their host galaxies: the population properties of seeded and unseeded galaxies, the co-evolution of SMBHs and galaxies as a function of galaxy properties, and the observational implications in terms of AGN activity. Secondly, we conduct a novel variability search for AGNs in the HUDF to provide an initial estimate of the SMBH number density in the reionization era to help us further constrain the origin of these intriguing objects. Given the multiple observed epochs in different filters at matched depths, and the suitable time delays between them for studying AGN variability, the HUDF serves as an optimal benchmark for estimating AGN activity at early epochs. This part of the Thesis builds on the work initiated with Hayes et al. (2024), aiming to locate credible variable sources from low to high redshift as the first step in constraining the volume density of SMBHs via AGN variability across cosmic time. The Thesis is structured as follows.

CHAPTER 2 describes the cosmological simulation used, the coupling with the semi-analytic model and the implementation of the seeding mechanisms. This chapter also details the model prediction of the BH and galaxy population properties with particular emphasis on the BH-galaxy relations.

CHAPTER 3 provides preliminary estimates on the observational properties of SMBHs visible as AGNs from the Pop III.1 model and compares them to high redshift observation from JWST. A description of the assumptions and physical models adopted for BH accretion and luminosities is also discussed.

CHAPTER 4 outlines the methods used to reduce and analyse the HUDF observations. After this the chapter also presents the first results in terms of AGN variable candidates and the estimate of the SMBH number density across cosmic epochs.

CHAPTER 5 condenses the main points of this Thesis work as well as a list of future possibilities of further investigation.

2

GALAXY EVOLUTION AND BLACK HOLE GROWTH FROM SEMI-ANALYTICAL MODELLING

Chapter 1 has introduced the milestone works of Banik et al. (2019) (Paper I) and Singh et al. (2023) (Paper II) where they first applied the Pop III.1 seeding mechanism in a cosmological simulation down to $z = 10$ and $z = 0$, respectively. Among other results, their analysis presents the cosmic evolution of the number density for different isolation distances, the clustering properties of the seeded halos at several redshifts, as well as synthetic spatial distributions of SMBHs on the sky.

In this Chapter we will introduce the simulation framework together with the methods and the assumptions made. In particular, we model SMBH formation in a cosmological box within the standard ΛCDM cosmogony by exploiting halo merger trees generated with `PINOCCHIO` (Monaco et al., 2002; Munari et al., 2017) to simulate the large-scale distribution of DMHs. Subsequently, these halos are seeded according to their isolation state at the formation time (see Paper I and Paper II). By adapting the `PINOCCHIO` DMH merger tree to a suitable format and adding the modelling of subhalos, we follow the redshift evolution of these structures and we populate them with galaxies utilizing the state-of-the-art GAlaxy Evolution and Assembly (GAEA) SAM of galaxy formation and evolution (Hirschmann et al., 2016; Fontanot et al., 2020; De Lucia et al., 2024). Different seeding mechanisms can be explored in this framework, but in this study we have focused mainly on the Pop III.1 model.

In Section 2.1, we describe in detail our new fully semi-analytical approach and the implementation of SMBH seeding and accretion. The main results of this study, together with the possible implications, are presented in Section 2.2. Finally we review our summary and conclusions in Section 2.3. The whole content of this Chapter reports the results from the paper accepted for publication to the Monthly Notices of the Royal Astronomical Society:

- **Vieri Cammelli**, Pierluigi Monaco, Jonathan C. Tan et al., 2024, "*The formation of supermassive black holes from Population III.1 seeds. III. Galaxy evolution and black hole growth from semi-analytical modelling*, MNRAS, 536, 851-870."

2.1 METHODS

In this work, we couple DMH merger trees extracted from cosmological simulations using the `PINOCCHIO` algorithm with the GAEA SAM, providing

predictions of the properties of galaxy populations associated with the DMH distribution at various redshifts. Given the full semi-analytical approach, we can explore wide ranges of physical and/or observational properties of galaxies with flexibility in the choice of the parametrizations adopted for the physical processes at play. In the following we give a brief description of the PINOCCHIO runs and the GAEA model. We then describe how we interface the two codes. We design a specific code which takes the PINOCCHIO halos as inputs, models the physical properties of main and sub halos and returns DMH merger trees structurally equivalent to numerically derived ones. We focus our implementation on different assumptions for SMBH seeding mechanisms within the treatment of accretion modes onto the central black hole already implemented in GAEA (Fontanot et al., 2020).

We note here that the two key tools that we use in our approach have been successfully tested against observational and numerical measurements: (i) the GAEA model is able to reproduce a wide range of observational results at $z = 0$ and up to $z \sim 5$; (ii) the high resolution realizations of halo merger trees constructed with PINOCCHIO result from an approximate method which has been shown to agree well with simulations down to $z = 0$, as shown in the above cited papers and references therein.

In this Thesis we adopt the following nomenclature. Gravitationally bound DM structures that are not hosted by a larger bound structure are called *halos*, the galaxies lying at their centres are called *central galaxies*. Halos may host smaller bound clumps of DM, we call them *subhalos*, and the galaxies at their centre *satellite galaxies*. In a simulation, the disruption of a subhalo may happen before the actual merger should take place, due to limited resolution. Therefore, its associated satellite can exist for some time after the disappearance of its subhalo; in this phase it is named an *orphan galaxy*.

2.1.1 *The dark matter skeleton: PINOCCHIO*

PINOCCHIO (PIN-pointing Orbit Crossing-Collapsed HIerarchical Objects Monaco et al., 2002; Munari et al., 2017) is a semi-analytical code that follows the formation and merger history of dark matter halos in Lagrangian space, that is the space defined by the initial positions of mass elements. One can think of PINOCCHIO as an algorithm applied to the initial conditions of a simulation, where the Lagrangian space is discretized into a grid and each cell is represented by a massive particle. First, an algorithm based on ellipsoidal collapse computes, for each particle, the time at which the particle is deemed to reach a multi-stream region (orbit crossing). The particles are then grouped into massive halos, whose position is estimated using LPT. As such, this code can be seen as a halo finder in Lagrangian space, that provides relatively accurate halo catalogues without running a full numerical simulation. While memory requirements are still high, so that a massive run requires a supercomputer, the computational cost is thousands time lower than an equivalent N-body simulation.

In this work we use the PINOCCHIO run presented in Paper II (Singh et al., 2023) i.e., a cubic box of side 59.7 Mpc (40 h^{-1} Mpc with $h = 0.67$) with standard Planck cosmology (Planck Collaboration, 2020), sampled with 4096^3 particles, for a particle mass of 1.23×10^5 $M_\odot$. The smallest resolved halos were set at 10 particles (an acceptable value for a semi-analytical algorithm), i.e., 1.23×10^6 $M_\odot$. This box was processed in Paper II to compute which mini-halos host SMBH seeds in the Pop III.1 scenario, with fiducial seed mass assumed to be $10^5 M_\odot$. Distributing such a large box on hundreds of nodes makes it difficult to reconstruct massive halos, whose Lagrangian size may be similar or even exceed the size of a computational domain. As explained in Paper II, for evolution from $z = 10$ down to $z = 0$ the box was re-run at a lower resolution (using 1024^3 particles) on a single node, thus avoiding any issue in the domain decomposition. This was achieved by first identifying the particles of the seeded halos in the high-resolution box, and then tagging the corresponding low-resolution particles whose Lagrangian space overlaps with the high-resolution particles' Lagrangian space. These tagged particles correspond to the seeded halos in the low-resolution simulation. In the low-resolution run (used in this work), the dark matter particle mass is $\sim 5 \times 10^7 M_\odot$, and the smallest resolved halo is a factor of 10 more massive. PINOCCHIO is also able to produce the merger history of dark matter halos with continuous time sampling by providing the exact redshift for each individual merger event.

2.1.2 *GAEA semi-analytical model*

GAEA represents an evolution of the original model published by De Lucia and Blaizot (2007). In this study, we utilize the version of the model published in Fontanot et al. (2020) (hereafter F20). The model includes: (a) a comprehensive treatment of chemical enrichment, explicitly addressing differential enrichment linked to Asymptotic Giant Branch (AGB) stars, Type II SNe, and Type Ia SNe (De Lucia et al., 2014); (b) an updated approach to stellar feedback tracing gas ejection via stellar-driven outflows (the model is partially based on results from hydrodynamical simulations, Hirschmann et al., 2016) coupled with a gas re-incorporation timescale dependent on DMH mass (Henriques et al., 2015); (c) an improved model for disc sizes (Xie et al., 2017) tracking angular momentum evolution through mass and energy exchanges within the galaxy; (d) an update modelling for cold gas accretion onto SMBHs (F20). This latter ingredient is relevant for this study and will be described in more detail in the next subsection.

GAEA has been shown to reproduce a wide range of observations. Fontanot et al. (2017a) demonstrated that the evolution of the Galaxy Stellar Mass Function (GSMF) and cosmic Star Formation Rate (SFR) obtained from GAEA are in agreement with measurements available up to $z \sim 7$. Hirschmann et al. (2016) showed that the model reproduces well the observed gas fractions and mass-metallicity relations at $z < 3$, but tends to overpredict the SF activity of low-mass galaxies at low redshift. The model we use in this study

also nicely reproduces the fraction of quiescent galaxies as a function of stellar mass and hierarchy at low-z (De Lucia et al., 2019). Furthermore, the model galaxies exhibit size–mass and angular momentum–mass relations that are in relatively good agreement with observational assessments, both in the local Universe and at higher redshifts (Zoldan et al., 2019). In future work we will extend our analysis to alternate recent versions of the model including distinct treatments for the partitioning of cold gas into atomic and molecular hydrogen (Xie et al., 2017, 2020; De Lucia et al., 2024), as well as a model accounting for a variable stellar initial mass function (Fontanot et al., 2017b, 2018, 2024). However, these two variations will not be considered in the present study.

2.1.2.1 *SMBH accretion and feedback*

The modelling adopted in GAEA for the accretion onto the SMBH is described in detail in F20. We note here that this phenomenon is treated following two main specific prescriptions. A first accretion channel from hot gas, known as *radio-mode*, is modelled according to the implementation of Croton et al. (2006). In this mode, the accretion rate is proportional to the mass of the BH (M_{BH}), to the virial velocity V_{vir} and to the fraction of the hot gas in the DMH (f_{hot}), adjusted by a free parameter k_{radio}:

$$\dot{M}_R = k_{radio} \frac{M_{BH}}{10^8 M_\odot} \frac{f_{hot}}{0.1} \left(\frac{V_{vir}}{200\,km/s} \right)^3 . \quad (1)$$

However, the more luminous AGNs arise from a second accretion mode, traditionally termed the *Quasi Stellar Object* (QSO)-mode: we take advantage of the modelling of cold gas accretion onto SMBHs presented in F20. In particular, we refer to the model implementation defined as F06-GAEA in F20, which is based on prescriptions first described by Fontanot et al. (2006). The occurrence of the AGN phenomenon has been accomplished by using a three phase approach. 1) The first phase requires that a fraction of the cold gas available in the galaxy dissipates a substantial amount of angular momentum and gathers in the central region, turning into a gas reservoir available for accretion onto the BH. 2) The amount of cold gas flowing from the reservoir towards the centre of the galaxy leads to accretion onto the BH. 3) Ultimately, outflows induced by the AGN lead to the expulsion of a portion of the galaxy's gas content.

This model assumes that disc instabilities and galaxy mergers lead to an efficient angular momentum loss and trigger QSO-mode accretion events. Typically this loss of angular momentum results from SFR episodes in the central regions, which inject turbulence and exert radiation drag. We assume that, following merger events, the BH reservoir accretion rate is proportional to the central SFR via the free parameter f_{lowJ} schematized as:

$$\dot{M}^{cs}_{rsv} = f_{lowJ} \psi_{cs}, \quad (2)$$

with ψ_{cs} estimated via the collisional starburst prescriptions from Somerville et al. (2001) and it equals the amount of SFR in the central regions triggered by the merger itself. It is important to note that BHs are assumed to merge instantaneously (no delay is assigned) once the host galaxies have merged.

For disk instabilities, the net result in GAEA involves moving a fraction of stars from the stellar disk to the stellar bulge so as to restore stability (De Lucia et al., 2011). Since there is no star formation associated by construction, we assume the reservoir growth rate to be proportional to the bulge grow rate $\dot{M}_{bulge}$. Hence:

$$\dot{M}^{di}_{rsv} = f_{lowJ}\mu\dot{M}_{bulge}, \tag{3}$$

where the free parameters f_{lowJ} and μ (6×10^{-3} and 10, respectively, as in F20) regulate the fraction of gas accreted due to angular momentum loss.

Once the reservoir gathers gas around the central BH, accretion episodes can be triggered. Following the viscous accretion rate derived by Granato et al. (2004), we define the accretion onto the BH as:

$$\dot{M}_{BH} = f_{BH}\frac{\sigma_B^3}{G}\left(\frac{M_{rsv}}{M_{BH}}\right)^{3/2}\left(1+\frac{M_{BH}}{M_{rsv}}\right)^{1/2}, \tag{4}$$

where σ_B is the velocity dispersion of the bulge component, assumed to scale linearly with V_{vir} as derived by Ferrarese (2002) for a sample of local galaxies.

This prescription, once coupled with the amount of gas accumulated into the reservoir, can induce accretion rates beyond the Eddington limit. We limit the actual accretion rate to:

$$\dot{M}_{max} = 10\frac{M_{BH}}{t_{edd}} = 10\dot{M}_{edd}, \tag{5}$$

where $\dot{M}_{edd}$ is the accretion rate of a BH shining at the Eddington luminosity with a radiative efficiency of 10%, over an Eddington-Salpeter timescale $t_{edd} \simeq 45\,\mathrm{Myr}$. This upper limit is supported by both observational and theoretical findings (see, e.g., Takeo et al., 2019; Jiang et al., 2019; Delvecchio et al., 2020). Some theoretical models indicate that such large accretion rates may occur via intermittent bursts, particularly at higher redshifts (e.g., Inayoshi et al., 2016).

F20 examined the impact of AGN activity on the host galaxy, specifically focusing on its cold gas phase. AGNs are believed to influence the surrounding medium by actively heating it, and eventually leading to the expulsion of cold gas via galactic winds driven by the AGN. SNe explosions combined with the radiation pressure of the AGN are assumed to promote further accretion by compressing part of the Inter Stellar Medium (ISM) ($f_{cen} \sim 10^{-3}$) in the central region. This material is eventually added to the BH reservoir (see Monaco and Fontanot, 2005). Each accretion episode

described by Eq. (4) triggers an AGN-driven outflow, with a rate that is modelled assuming a scaling relation with the BH accretion rate:

$$\dot{M}_{qw} = \epsilon_{qw} \dot{M}_{BH}, \tag{6}$$

where ϵ_{qw} is a free parameter whose value is 320 as reported by F20. Note that this AGN-driven wind scaling shows results consistent with both hydrodynamical cosmological simulations (Brennan et al., 2018) and observational findings (Fiore et al., 2017).

2.1.3 *Building the merger trees*

The standard GAEA runs have been defined on DMH merger trees extracted from the Millennium Simulation suite (Springel et al., 2005). The merger tree format adopted in GAEA is organized as follows. The starting point is a temporal sequence of *snapshots*, where the position and velocity of all particles at a given time is provided. First, DM halos are identified using a classical Friends of Friends (FoF) algorithm. Then, a second algorithm (SUBFIND) identifies subhalos in each FoF group. Finally, the merger trees are built by identifying unique descendants for all subhalos (see Springel et al., 2005, for details). For each identified halo or subhalo, pointers and physical quantities are stored at the same redshifts of the snapshots. These merger trees are then provided to GAEA as a skeleton for the galaxy formation model.

At the resolution of the Millennium simulation, dark matter subhalos *disappear*, i.e. cannot be identified anymore as distinct subhalos, at typically large distances from the halo centre, when the merger is likely incomplete. Since the baryons are more centrally concentrated than dark matter, this time is presumably underestimating the time at which the hosted satellite galaxy is expected to merge with the central galaxy. In GAEA, this numerical effect is mitigated by modelling orphan galaxies: when a subhalo is lost, its satellite galaxy is assigned a residual merger time, estimated using dynamical friction arguments, and its evolution is followed until it merges with the central galaxy of the host halo. The position and velocities of orphan galaxies are obtained by following the most bound particle of the disrupted subhalo.

The formalism and the structure of the `PINOCCHIO` merger trees (see Section 2.1.1) are substantially different compared to the numerical ones. The outputs are provided in the form of merging histories and catalogues of dark matter halos. The merger histories offer a continuous time sampling that uniquely determines the evolution of every single halo along the merger tree, providing the mass, the redshift of first appearance, the merging redshift (if any), the halo ID and other useful pointers. The catalogues supply, according to a time grid, the information about mass, position, velocity and halo ID for all halos. Thus, in order to use `PINOCCHIO` outputs as GAEA inputs, it is necessary to adapt the merger trees to a different format,

and add quantities that are used by GAEA and not available in PINOCCHIO as detailed in the following.

As a first step, for the existing halos, at any given time we linearly interpolate in redshift the position, velocity and mass between the two closest PINOCCHIO catalogues. In fact, both PINOCCHIO and GAEA do not require running on a shared, pre-defined time grid. Depending on the specific scientific application, without re-running PINOCCHIO we can interpolate the physical quantities available in the catalogues according to any given snapshot list. The accuracy of such interpolation is assured by the fact that PINOCCHIO halo positions are predicted with LPT. This allows us to run PINOCCHIO once and then freely choose the snapshot list as input to GAEA afterwards. For the purpose of this study, given the relatively small volume, we set twice as many snapshots as in the Millennium, equally spaced in the logarithm of the scale factor a from redshift 127 to 0. This guarantees the self-consistency with the GAEA predictions based on the Millennium merger trees and to concurrently increase the time sampling for a better time resolution. At redshifts above ~ 4 we also test a time sampling with 4 times the number of snapshots with respect to the Millennium one and we verify that the predictions are stable and robust as one decreases the integration time step in the semi-analytical model.

Secondly, one physical quantity required by GAEA, but unavailable in PINOCCHIO, is the maximum rotational velocity of the halo V_{max}. In dark matter N-body simulations this quantity is estimated as:

$$V_{max} = max\Big(\sqrt{\frac{GM(r)}{r}}\Big). \tag{7}$$

The radius r runs over all the particles bound to the halo. In PINOCCHIO merger trees, such information is not directly available. We utilise instead:

$$V_{max} = f(z)V_{circ}, \tag{8}$$

where a dimensionless factor $f(z)$ (of the order of unity), which is a function of redshift, multiplies the circular velocity of the halo which can be estimated as:

$$V_{circ} = \sqrt[3]{10GH(z)M}, \tag{9}$$

where $H(z)$ is the Hubble parameter and M the mass of the halo defined and calibrated against FoF halo masses (Monaco et al., 2002).

Finally, it is important to note that in PINOCCHIO, when two halos merge (referred to as the *accretion* time), the smaller one gives its mass to the larger one, and is *removed* from the halo list, meaning that it is not further updated. This means we cannot keep track of these halos in the simulation once they become subhalos. Therefore, we model the presence of subhalos by assigning a survival time to each halo that is lost following a merger. This approach integrates the merger tree with information on its subhalo, as detailed by Berner et al. (2022) (see next Section).

2.1.3.1 *Modelling subhalos*

In order to cast the above requirements into the PINOCCHIO-based approach, we have written a *translation* code that models the presence of subhalos and adds physical information to properly mimic and output Millennium-like DMH merger trees. We use results derived from both theoretically and numerically predicted quantities to fill the required physical information and add the treatment of subhalos. In particular, we assign a population of subhalos within each halo by: 1) assuming a spatial distribution for subhalos following a Navarro-Frenk-White (NFW) density profile (Navarro et al., 1997); 2) implementing a statistically derived distribution for the angular momentum of subhalos based on the orbital binding energies observed in numerical simulations (Zentner et al., 2005; Birrer et al., 2014); and 3) assigning a subhalo survival time since accretion (Berner et al., 2022; Boylan-Kolchin et al., 2008). Note that the NFW profile is not accurate for subhalos in the inner regions of the DMHs, but describes well the observed spatial distribution of galaxies (e.g., Gao et al., 2004). We opt to sample the position of subhalos from the NFW distribution at each output time, which means there is no correlation in their spatial position between two consecutive snapshots. In fact, subhalo orbits and positions do not affect the physical treatment of galaxies in the F20 version of the GAEA model.

We define a total merging time of each subhalo since the merging of the two halos (accretion time) as the sum of the survival time of dark matter subhalos, derived from simulations, and the residual merging time of their galaxies at the time subhalos disappear. This is supported by the fact that in our runs of GAEA, satellite and orphan galaxies are equally treated. Our approach takes into account the two phases separately. We initially adopt a halo survival time to shape the first part of the total merging time, practically mimicking the satellite phase t_{sat}. Given the merging subhalo mass m_{sub}, the main halo mass M_{main} and the orbital circularity η at merger, we assume (with reference to Eq. (2.6) from Berner et al., 2022):

$$t_{sat} = A(D)\tau_{dyn}\frac{\left(\frac{M_{main}}{m_{sub}}\right)^{b(D)}}{\log\left(1+\frac{M_{main}}{m_{sub}}\right)}\exp(c\eta), \tag{10}$$

where A and b are functions of the linear growth factor $D(z)$ and c is a free parameter. The timescale τ_{dyn} is the dynamical time at the virial radius, which is assumed to depend on the Hubble constant H only and its value is of the order of $\sim 0.1H^{-1}$.

Next, we add a second time calibrated by estimating the residual merging time of orphan galaxies assigned by GAEA as a function of the halo masses at the time the subhalos disappear in the Millennium simulation. This was originally derived from the classical dynamical friction timescale (Chandrasekhar, 1943; Binney and Tremaine, 1987), adapted to be applied in SAMs (e.g., De Lucia and Blaizot, 2007; Boylan-Kolchin et al., 2008; De Lucia et al., 2010), which requires inputs of the position of the subhalo

at the time it is lost in the simulation. Since this information is not available in PINOCCHIO, we assume a dependence only on the mass ratio between the merging halo (m_{sub}) and the main halo (M_{main}) at the accretion time. We also attempt to factor out the dependence on the redshift by directly fitting the ratio between the extra orphan time t_{orph} and the age of the Universe, t_{age}, at the accretion time. Based on the above considerations, we derive a formula for the orphan time, given by:

$$t_{orph} = t_{age} k \frac{\left(\frac{m_{sub}}{M_{main}}\right)^{\alpha}}{\log\left(1 + \frac{m_{sub}}{M_{main}}\right)^{\beta}}, \tag{11}$$

where k, α and β are free parameters (see the next Section for the calibrated values).

Finally, the total merging time of the satellite galaxy is the sum of the estimated halo survival time as in Eq. (10) and the modelled extra time for orphan galaxies given by Eq. (11)(see Appendix A). Note that while the individual contributions from these time scales vary depending on the mass resolution of the DMH trees, their combined total remains resolution-independent. This is because the total merging time is determined by the halo mass ratio and the accretion redshift. Hence, directly adding the total merging time to newly identified subhalos ensures that the construction of the PINOCCHIO merger trees does not depend on the DM mass resolution.

2.1.4 *Calibrating the trees*

As discussed in Section 2.1.3, our reference GAEA realization is run on merger trees extracted from the Millennium Simulation. This is a numerical realization of a cosmological volume of side 500 h^{-1}Mpc assuming the WMAP1 lambda cold dark matter cosmology ($\Omega_\Lambda = 0.75$, $\Omega_m = 0.25$, $\Omega_b = 0.045$, $n = 1$, $\sigma_8 = 0.9$, and $H_0 = 73\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, Springel et al., 2005). We run a PINOCCHIO box that matches the Millennium one in terms of particle mass resolution, volume and cosmological parameters. This run has been used to calibrate the time scales discussed in the previous section.

The different definition of halo mass between the two simulations results in DM halo masses which are on average a factor of ~2-3 more massive in PINOCCHIO with respect to the Millennium (see Fig. 2.1). This difference requires us to calibrate, at each snapshot, the PINOCCHIO Halo Mass Function (HMF) to match the Millennium one. To do so, we bin the range of cumulative number density n between the maximum halo mass for which we count at least ten objects and the minimum halo mass corresponding to the resolution of the simulation ($\sim 5 \times 10^8 M_\odot$). In such an interval, we fit both the PINOCCHIO and Millennium HMFs and we retrieve their ratio in halo mass for any given number density. Then, we multiply the PINOCCHIO halo mass by the correspondent correction factor. This operation provides halos with the same mass distribution of the Millennium simulation, thus avoiding introducing systematic bias in subsequent steps. With a similar

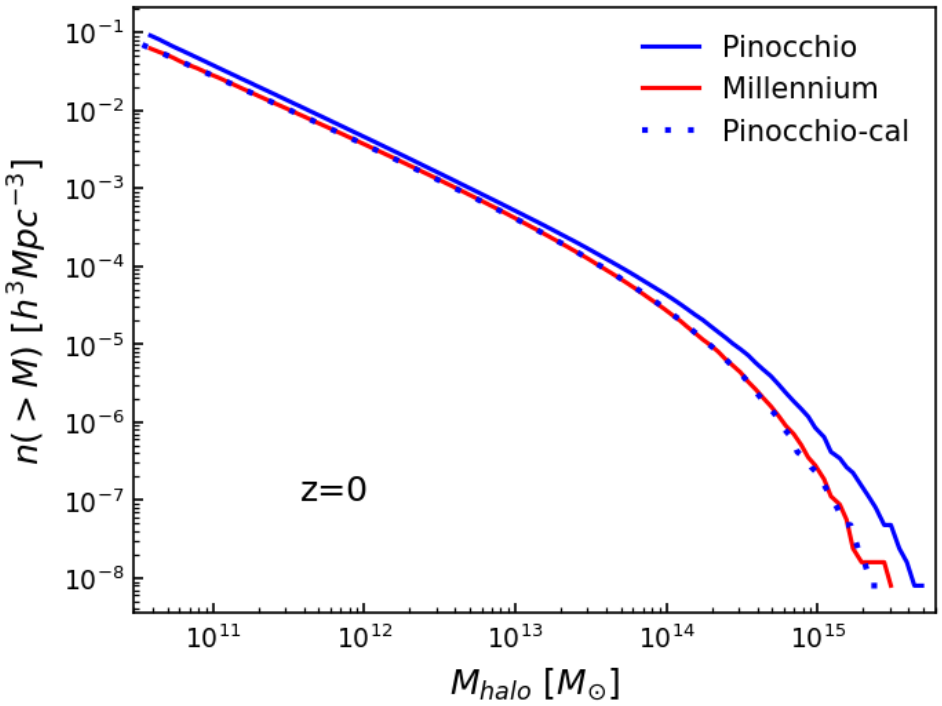


Figure 2.1: HMF at redshift $z = 0$ from the PINOCCHIO Millennium-like box (solid blue line) and the Millennium simulation (solid red line). Due to the different definition of the halo mass, at $z = 0$ the halo masses in PINOCCHIO are moderately higher compared to the Millennium halos. We calibrate the PINOCCHIO HMF by using a polynomial fit of mass difference at a given number density that is a function of the redshift (dotted blue line).

strategy, we calibrate the distribution of the maximum rotational velocity of the halo, V_{max}, introducing a correction factor in order to match the distribution extracted from the Millennium main halos as a function of redshift (see Eq. (8)). Additionally, the process of the assembly of the merger trees involves parameters we need to adjust, specifically the ones associated with the estimate of the merging time. For this scope, we run the GAEA semi-analytical model on the *translated* merger trees generated from the PINOCCHIO box as described above. Then we aim at calibrating the parameters involved in the construction of the merging times t_{sat} and t_{orph} by reproducing the GSMF at redshift $z = 0$. In fact, the merging history of galaxies plays a crucial role in the build-up of their stellar and gas mass content, as well as shaping the cosmic Star Formation Rate Density (SFRD) and AGN activity across the age of the Universe.

We explore different combinations of the satellite t_{sat} and orphan t_{orph} time scales in order to reproduce the normalization and the position of the knee of the GSMF at $z \sim 0$. In particular the total merging time of the satellite galaxies impacts on the total number density in terms of their relative contribution to the GSMF. Therefore, we calibrate the parameters in Eqs. (10) and (11) by matching at the same time the GSMF of central and satellite galaxies obtaining the following values:

$$
\begin{aligned}
A &= \begin{cases} 0.39 & \text{for } D(z) > 0.8 \\ 0.195 + \frac{0.195}{0.2}(D(z) - 0.6) & \text{for } 0.8 \geqslant D(z) \geqslant 0.6 \\ 0.195 \left(\frac{D(z)}{0.6}\right)^2 & \text{for } D(z) < 0.6 \end{cases} \\
b &= 1.015 \cdot D(z) \\
c &= 1.3 \\
k &= 0.3 \\
\alpha &= 0.94 \\
\beta &= 1.7.
\end{aligned}
\tag{12}
$$

Fig. 2.2 shows the result of our best estimate of the GSMF from both populations of galaxies (upper panel). The GSMF of centrals using the PINOCCHIO merger trees agrees well with the Millennium-based predictions for stellar masses above few times $10^{10} M_\odot$, throughout the knee until the exponential cut-off at the high-mass end. On the other hand, satellites show a substantial difference (a factor of up to ~ 4) at the low mass end of the GSMF, i.e., at masses below $\sim 10^{10}$ $M_\odot$, even though the total gap in the bottom panel is dominated by the central galaxies. We argue that this difference is due to the intrinsic differences in the features of the PINOCCHIO and Millennium simulations in the definition and construction of merger events (orbit crossing vs FoF). As a consequence, this effect causes a dearth of small mass halos as one approaches $\sim$ few 10^9 $M_\odot$, explaining the gap in the number density of galaxies. However, if one compares the total

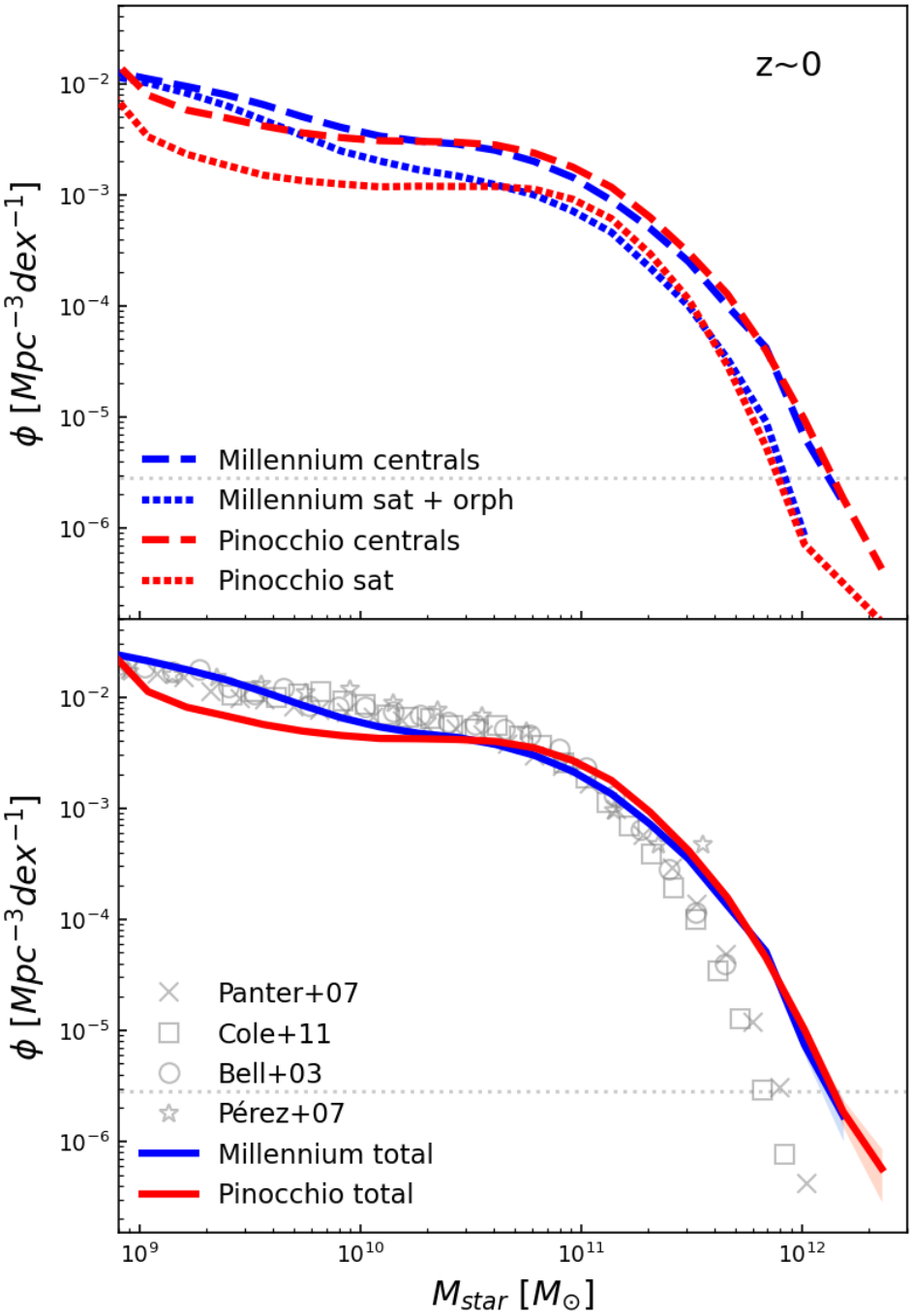


Figure 2.2: *Upper panel:* GSMF at redshift z=0 from the Millennium-like box. Dashed (dotted) lines indicate the contribution of the central (satellite and orphan) galaxies, respectively. While central galaxies show remarkable agreement, PINOCCHIO low mass satellites differ by a factor of ~ 4 in number density with respect to the Millennium ones. *Lower panel:* Total GSMF at $z = 0$ from the Millennium-like box. The symbols show observed data points from Panter et al. (2007), Cole et al. (2001), Bell et al. (2003) and Pérez-González et al. (2008). This plot shows the goodness of the calibration, especially above $10^{10} M_\odot$.

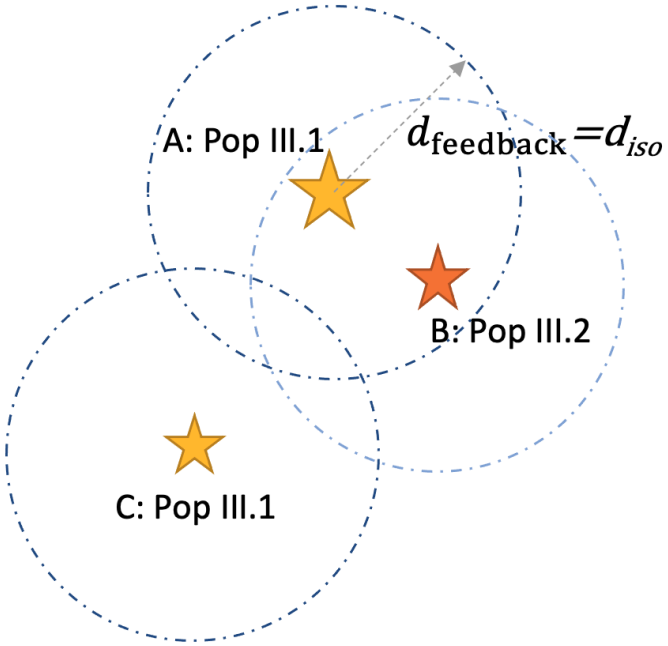


Figure 2.3: A simplified illustration of the Pop III.1 SMBH seeding scheme (see text) showing the conditions for a star to be isolated enough to be considered as a Pop III.1 star. Adapted from Singh et al. (2023).

GSMF combining all types of galaxies (bottom panel), the difference in the low-mass end reduces to a factor of ~ 2 or less. We also verify that at earlier epochs, when the fraction of satellite galaxies is lower, the impact of the merging time is marginal. Our GSMF reproduces the trend obtained from the Millennium-based predictions with a better agreement compared to $z \sim 0$.

Bearing in mind these caveats, we consider the calibration resulting in the GSMF depicted in Fig. 2.2 as being sufficiently accurate for the scope of this work and we opt not to calibrate the GAEA physical parameters keeping the values proposed by F20 (see also Appendix B).

2.1.5 *Seeding SMBHs*

The primary goal of our study is to investigate the impact of the Pop III.1 SMBH seeding scheme on galaxy and SMBH properties. This seeding scheme is illustrated in Fig. 2.3 (for a more detailed introduction of the seeding we refer to Banik et al., 2019; Singh et al., 2023). In the figure, three stars, denoted as A, B, and C, reside in different halos. Among them, only A and C evolve into Pop III.1 supermassive protostars, while B is designated as a lower mass Pop III.2 star. This classification depends on their physical separation from other sources at the time of their formation.

Star A, forming first, exerts its influence within a sphere of radius equal to $d_{feedback}$, primarily driven by radiative feedback. Since this star originates in a pristine primordial gas environment, devoid of any influence from

neighbouring stars, it falls under the category of a Pop III.1 star. Star B, on the other hand, forms at a distance less than $d_{feedback}$ from star A, subjecting it to the effects of feedback. Consequently, it is classified as a Pop III.2 star (or even a Pop II star if the gas cloud it forms from is chemically enriched). Finally, star C forms beyond the regions influenced by feedback from stars A and B. As a result, it is are designated as another Pop III.1 source.

The physical mechanism that may allow SMBH formation is the impact of dark matter annihilation heating on the structure of the Pop III.1 protostar (Spolyar et al., 2008; Natarajan et al., 2009; Rindler-Daller et al., 2015). The expected effect is to keep the protostar relatively large and cool, thus emitting a weak flux of ionising radiation. Efficient collapse of the baryonic content of the minihalo to the Pop III.1 protostar, i.e., yielding a mass of $\sim 10^5\,M_\odot$, is assumed to occur. After a few Myr of evolution, this source is then expected to form a SMBH of similar mass. Thus the initial mass of the SMBH seeds that we implement in the Pop III.1 scheme is $M_{BH} = 10^5\,M_\odot$.

The main parameter of the Pop III.1 model is the isolation distance, d_{iso}, which is required for a minihalo to form a Pop III.1 source. Banik et al. (2019) and Singh et al. (2023) have shown that values of $d_{iso} \simeq 50$ to 100 kpc (proper distance) are sufficient to yield overall numbers of SMBHs consistent with estimates of local number densities at $z = 0$. We will thus consider values of $d_{iso} = 50, 75, 100$ kpc in our modelling.

We also consider two other seeding models. The first of these is the BH seeding scheme that is implemented within GAEA and described in detail in Xie et al. (2017). Recall that this scheme has been tested at the Millennium simulation resolution and is intended to be a sub-grid model that is not attempting to mimic any seeding scenario. Whenever a new DMH is resolved in the PINOCCHIO simulation, we assign it a BH mass (M_{BH}) scaled with the parent DMH mass (M_{DM}):

$$M_{BH} = \left(\frac{M_{DM}}{10^{10} M_\odot h^{-1}}\right)^{1.33} \frac{10^{10} M_\odot h^{-1}}{3 \times 10^6}, \tag{13}$$

where the index of the relation is derived from Volonteri et al. (2011). We refer to this scheme in which all halos are seeded as All Light Seeds (ALS). Note that this is an *ad-hoc* model that mimics the formation of seeds from Pop III star remnants and estimates their evolution by accretion before the halo is resolved. Given the relatively high mass resolution of our PINOCCHIO box (see Section 2.1.1), applying Eq. (13) to the M_{DM} mass distribution results in BH seed masses of the order of $\sim 10^{1-2} M_\odot$. This mechanism effectively depicts a scenario where every single halo is seeded with a Pop III remnant BH and can be considered an implementation of the standard light-seed Pop III model described in Chapter 1. Note that this seeding scheme has been tested previously only at the resolution of the order of the Millennium Simulation (Xie et al., 2017). The GAEA version presented in F20 and used in the model calibration described in Section 2.1.4 is based on this latter seeding model.

Finally, for continuity with Paper II we also consider the HMT seeding scheme, where all halos with mass $> M_{thr} = 7.1 \times 10^{10}$ $M_\odot$ are seeded with a 1.4×10^5 $M_\odot$ BH. This is the typical seeding scheme used in hydro simulations (e.g., Vogelsberger et al., 2014), however the value of the threshold mass is determined by the mass resolution of simulations able to resolve galaxies in a cosmological box. The box used to obtain our results has a much higher resolution, so this scheme can be considered as a toy model for a seeding scheme where BH seeds appear later in the evolution of the Universe.

2.2 RESULTS

Here we present the predicted properties of the galaxy populations based on DMH merger trees extracted from the 40 Mpc h^{-1} `PINOCCHIO` box described in Section 2.1.1 and *translated* as explained in Section 2.1.3. We evaluate the Pop III.1 seeding model for three cases of $d_{iso} = 50$, 75, and 100 kpc (see Section 2.1.5), as well as for the ALS and HMT seeding schemes. Unless otherwise specified, we apply the accretion scheme onto the SMBH described in Section 2.1.2.1. Throughout this work, the three cases of isolation distance $d_{iso} = 50$, 75 and 100 kpc are depicted in orange, red and blue, respectively, while the ALS and HMT models are shown in magenta and green, respectively.

2.2.1 *Occupation fractions*

We examine the cosmic evolution of the fraction of galaxies that are seeded with SMBHs, i.e., the SMBH *occupation* fraction, as a function of various physical properties. In Fig. 2.4 we show the evolution of the occupation fraction of seeded galaxies as a function of various galaxy properties from redshift $z \sim 10$ down to $z \sim 0$. Different columns indicate different galaxy quantities, from left to right: halo mass, stellar mass, hot gas mass, and SFR. The ALS seeding scheme is by construction always identical to unity. For a more meaningful comparison in terms of occupation fractions of SMBHs, here we also consider the ALS case applying a cut in BH mass above 10^5 $M_\odot$ (hereafter depicted with the label ALS-cut).

At all epochs, the SMBH occupation fractions as a function of halo mass, stellar mass and hot gas mass have similar behaviours across Pop III.1 seeding models. As the isolation distance decreases, the occupation fraction rises. SMBHs tend to reside in the more massive systems, as the Pop III.1 model places them in the first locally formed mini-halos, which are more likely to be the progenitors of the largest structures later on. While the high-mass end occupation fraction as a function of halo mass reaches unity for all models at redshift zero, this is not the case for the occupation fraction as a function of stellar mass, i.e., only $\sim 60\%$ of galaxies with highest level of stellar mass ($\sim 10^{12}$ $M_\odot$) are seeded in the $d_{iso} = 100$ kpc case, but rising to about 100% in the $d_{iso} = 50$ kpc case. The various Pop III.1 cases show

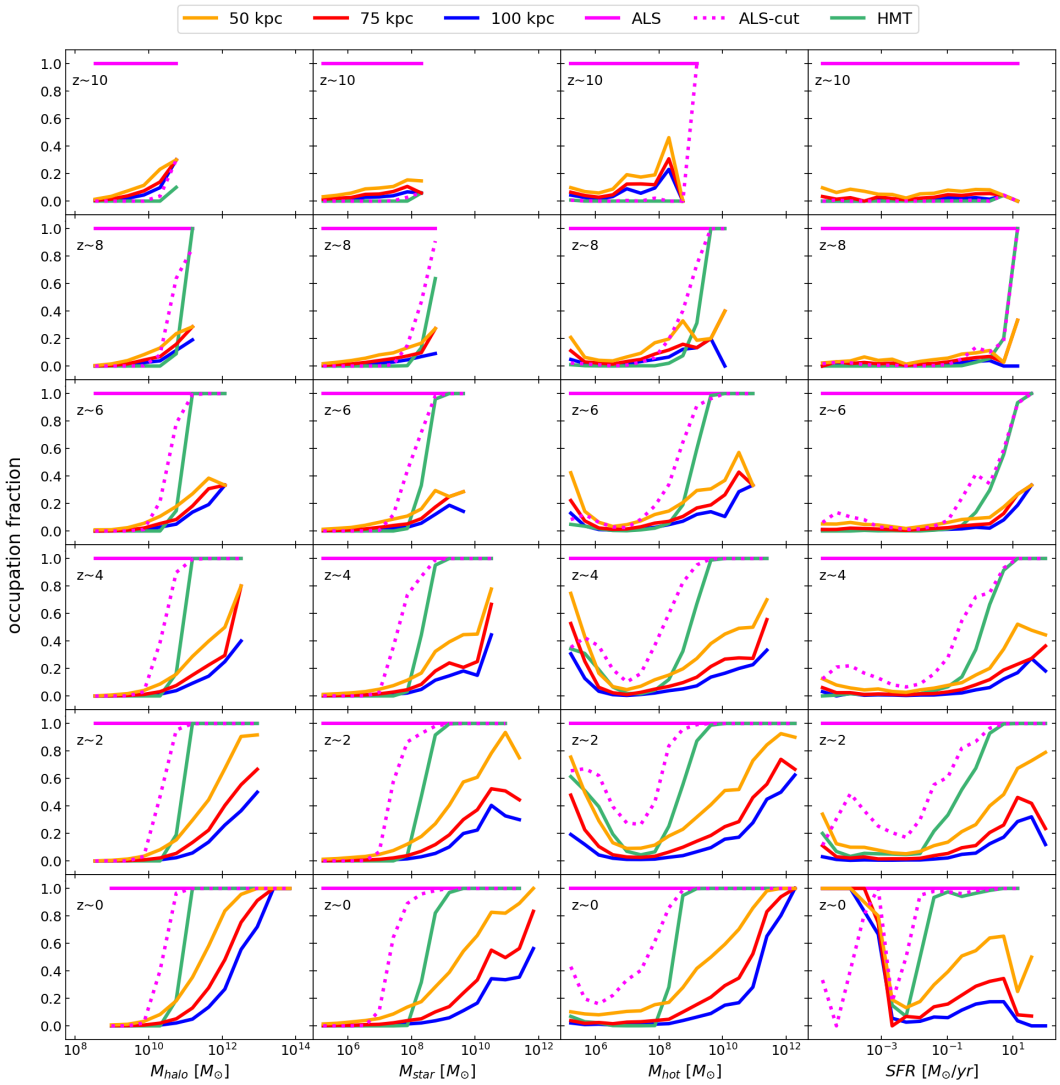


Figure 2.4: Occupation fraction of seeded galaxies as a function of different galaxy properties for several redshifts. From left to right: halo mass, stellar mass, hot gas mass and SFR. The 6 lines show different seeding mechanisms.

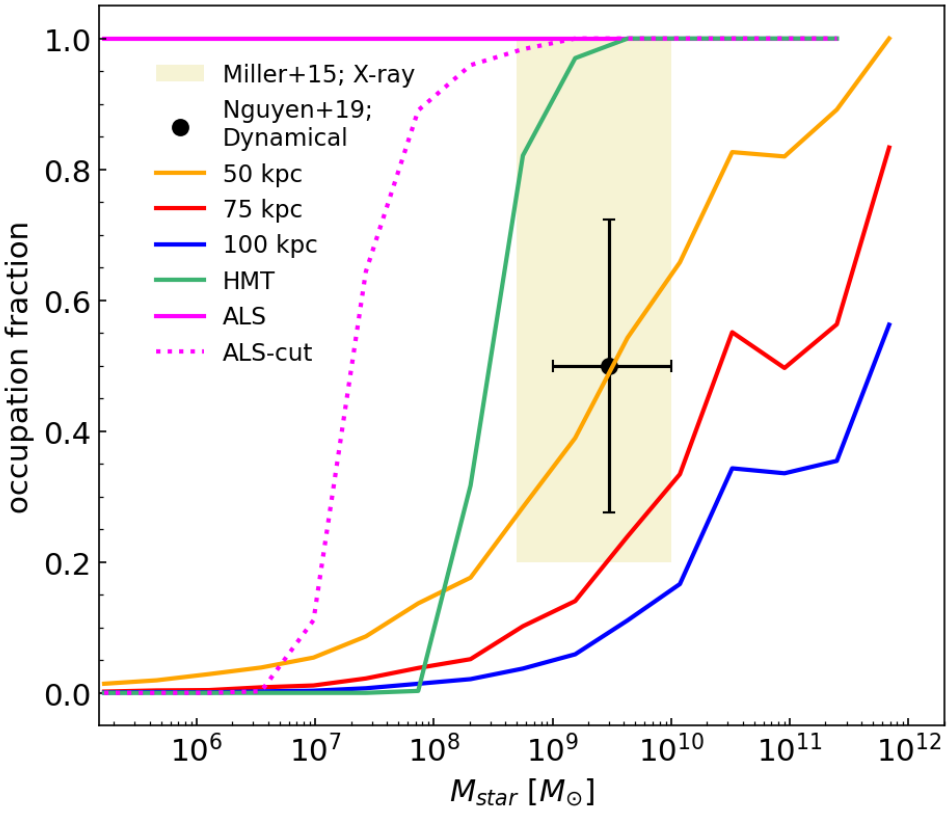


Figure 2.5: Occupation fraction of seeded galaxies as a function of the stellar mass at redshift zero. The 6 lines show different seeding mechanisms. In the range of stellar masses between few in 10^8 and 10^{10} $M_\odot$, we compare against different observational constraints for the occupation fraction. X-ray sources detected in local surveys pose some lower limit to the occupation fraction as presented by Miller et al. (2015) drawn as the yellow shaded region. In black we show the more stringent constraints as reported by Nguyen et al. (2019) obtained from dynamical findings in a small sample of nearby galaxies.

occupation fractions that rise gradually for $M_{star} \gtrsim 10^7\ M_\odot$ to $10^9\ M_\odot$ as d_{iso} increases from 50 to 100 kpc. On the other hand, the HMT model shows an occupation fraction that rises very steeply to unity for stellar masses $\gtrsim few \times 10^9\ M_\odot$ across all epochs. The ALS-cut case presents a similar shape, but shifted downward by one order of magnitude to $10^8\ M_\odot$ by z~0. This reflects the fact that, when a low mass seeding scenario is applied, the viscous accretion mode soon establishes a BH growth rate proportional to the SFR reproducing the local scaling at the latest epochs ($M_{star} \sim 10^3\ M_{BH}$). Therefore, applying a cut in M_{BH} corresponds to imposing a sharp threshold for M_{star}.

Thus a key difference between the Pop III.1 models and the HMT and ALS seeding schemes is that the Pop III.1 models have much smaller SMBH occupation fractions at relatively high values of M_{star}. This implies a population of unseeded galaxies reaching large values of M_{star}. Indeed, in the absence of AGN feedback, these systems keep forming stars with relatively high rates at all redshifts, which influences the SMBH occupation fraction versus SFR, discussed below. A related consequence is that the Pop III.1 models have a larger scatter in the $M_{star} - M_{halo}$ relation, discussed in Section 2.2.4.

When considering SMBH occupation fraction as a function of hot gas mass, to the extent that more massive halos and/or stellar components correlate with M_{hot}, we see similar trends. However, at intermediate redshifts we see that there is a population of seeded galaxies with very low values of M_{hot}, which are attributed to systems in which AGN feedback prevents the build up of hot gas mass.

Fig. 2.4 also shows the evolution of the occupation fractions as a function of SFR. At later epochs, from redshift ~ 4 down to intermediate redshifts ($z \sim 2$), one sees how seeded galaxies, mostly residing in more massive galaxies, tend to have larger SFR. At the same time AGN feedback begins to influence the gas in these galaxies, leading to an efficient suppression of star formation in massive systems. This demonstrates the effectiveness of radio-mode AGN feedback in keeping galaxies passive. By redshift $z \sim 0$ the fraction of passive galaxies, characterized by SFRs lower than $10^{-3}\ M_\odot/yr$, reaches SMBH occupation fractions of unity across all models. However, we see that the ALS-cut scheme goes up to one for SFR around $10^{-3}\ M_\odot/yr$, while it drops down to small fractions for slightly larger values of SFR. One should keep in mind that ALS-cut does not track the whole population of seeded galaxies, but only the ones having $M_{BH} > 10^5\ M_\odot$. As opposed to the other models, this latter sub sample of galaxies is not exclusive in terms of radio-mode AGN feedback since this also affects galaxies with BHs just below the BH mass cut, contaminating the measurements of the occupation fraction for the SFR as well as for the hot gas.

As a point of comparison with observational data, in Fig. 2.5 we illustrate constraints for the fraction of observed local galaxies hosting a SMBH in the range of stellar masses between few in 10^8 and $10^{10}\ M_\odot$. Observationally,

dynamical and X-ray measurements based on local galaxies seem to agree that at least $\gtrsim 50\%$ of the population in this mass range host a SMBH with mass $\sim 10^{4-6}$ $M_\odot$ (see discussion in Greene et al., 2020). In particular, Nguyen et al. (2019) show that out of 10 galaxies within a distance of 4 Mpc for which dynamical measurements are available, only 5 SMBHs are detected, inferring a lower limit for the occupation fraction. Additionally, Miller et al. (2015) advocate a lower limit of 20% with a most probable estimate of around 70% based on the detection of X-ray sources in local galaxies within a comparable stellar mass.

Fig. 2.5 illustrates that current observational measurements are not of sufficient precision to discriminate between the different models. If the quality and statistics of observational samples can be increased to reduce the uncertainties, then the presented models could be distinguished. However, it is interesting to note that our simulations present a better agreement in terms of the occupation fraction when d_{iso} approaches the value of 50 kpc, as we allow more halos to be seeded. Conversely, the 75 and 100 kpc cases struggle to produce large enough occupation fractions while the ALS-cut and HMT model tends to create more numerous SMBHs although still in agreement with Miller et al. (2015). On the other hand, the ALS scenario would be difficult to test against current observations as the totality of galaxies in this mass range would host a BH, whose mass would span from stellar seeds up to the SMBH regime.

2.2.2 *Stellar mass function*

In Fig. 2.6 we illustrate the cosmic evolution of the GSMFs evaluated for the Pop III.1 seeding models from $z \sim 0$ up to 3. For the redshift zero case, we see that below stellar masses of approximately $3 \times 10^{10} M_\odot$ the predicted functions match reasonably well down to low masses of $M_{star} \sim 10^8$ $M_\odot$. In this regime, the majority of galaxies are non-seeded, and so there is little impact, i.e., due to AGN feedback, of varying d_{iso}. On the other hand, reducing d_{iso} from 100 to 50 kpc has a dramatic impact at the high-mass end of the GSMF. Specifically, while the 75 and 100 kpc models struggle to reproduce the exponential cut-off of the GSMF above a few $\times 10^{10}$ $M_\odot$, the 50 kpc case is able to maintain good agreement with the observational data. This is due to the action of AGN feedback in suppressing the build-up of high stellar mass galaxies. We see that the ALS and HMT models, having similar, saturated occupation fractions at high stellar masses, are also able to reproduce the local high-mass end of the GSMF. It is worth stressing that the agreement with observational constraints cannot be improved by a different parameter calibration, as the lack of a central SMBH in unseeded galaxies prevents the onset of Radio-mode feedback in the first place, and stellar feedback alone is insufficient to reduce the number of massive galaxies while keeping a good agreement at lower masses and using realistic loading factors (White and Frenk, 1991).

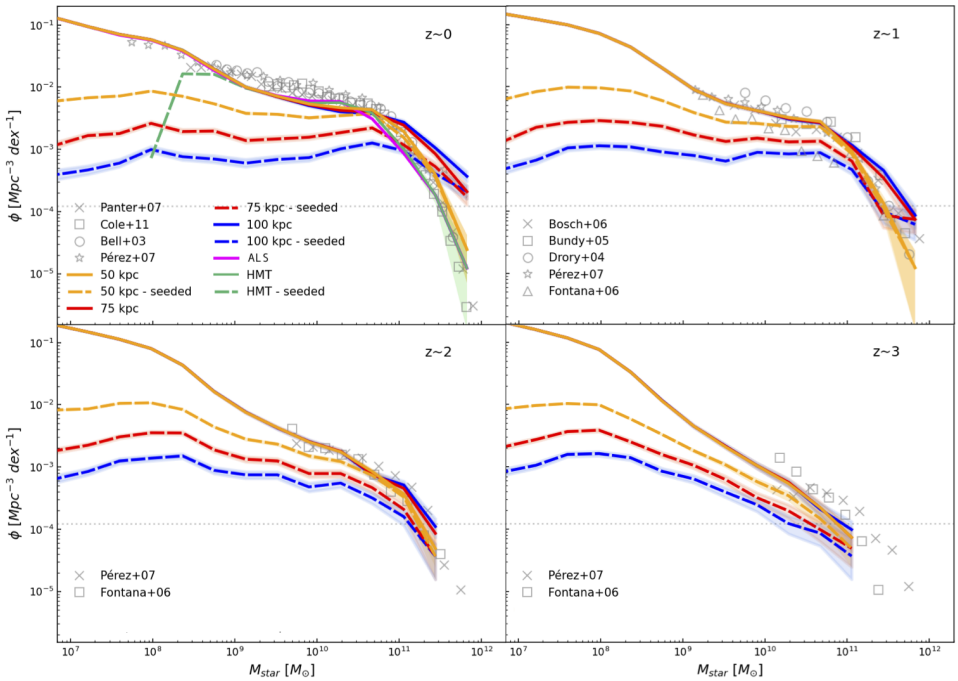


Figure 2.6: Cosmic evolution of the GSMF from redshift 0 up to 3. Solid (dashed) lines indicate the contribution of all (seeded) galaxies, respectively. The shaded areas depict the Poisson uncertainties in each mass bin. Results for Pop III.1 models with 3 values of isolation distance (in proper distance) are shown, as labelled. Observed data are in gray symbols (Panter et al., 2007; Cole et al., 2001; Bell et al., 2003; Pérez-González et al., 2008; Van Den Bosch et al., 2008; Bundy et al., 2005; Drory et al., 2004; Fontana et al., 2006). As a reference, the dotted horizontal line stands for 10 objects per mass bin in the whole volume of the box. The GSMF tells us the isolation distance parameter that better reproduces the observed trend, especially at the massive end. The 50 kpc case agrees with the exponential cut off of the massive population of galaxies from the local relation, where almost every galaxy with a stellar mass higher than $\sim 3 \times 10^{10}$ $M_\odot$ is assumed to host a SMBH, up to $z \sim 1$. At redshift 0 we compare with the HMT and ALS seeding schemes. These two cases mostly overlap in the graph. While ALS well reproduces the local observational trend by construction, HMT naturally seeds every massive galaxy resulting in efficient quenching of star formation in such systems. We use them as reference models to validate the Pop III.1 seeding at redshift z=0. The green dashed line of the HMT scheme tells us that the fraction of seeded galaxies rises sharply up to one above $\sim 10^9$ $M_\odot$ as shown from Fig. 2.5.

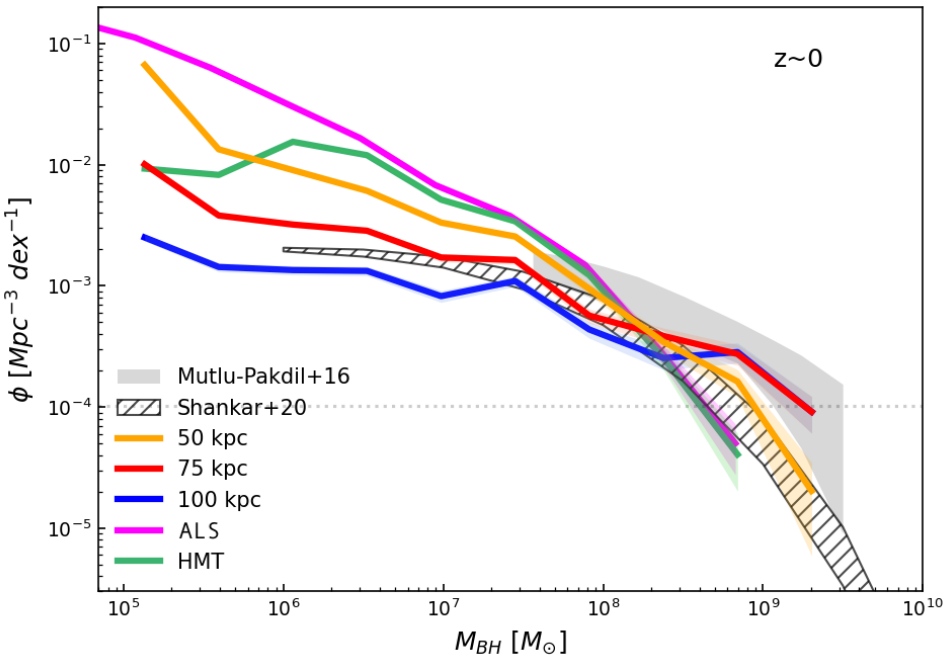


Figure 2.7: BHMF at redshift $z \sim 0$. Solid lines indicate the contribution of different seeding schemes. The shaded areas around the lines depict the Poisson uncertainties in each mass bin. Observed data are taken from a sample of local galaxies (shaded area Mutlu-Pakdil et al., 2016) and from the unbiased corrected relation from Shankar et al. (2020) (hatched area). As a reference, the dotted horizontal line stands for 10 objects per mass bin in the whole volume of the box.

Moving to higher redshifts, the differences between seeded and non-seeded galaxies become less visible in terms of average stellar mass distribution, as the SMBH are not big enough to affect the evolution of the host galaxy by means of radio-mode feedback. This also reflects in the GSMF: if we look at the evolution of the GSMF up to redshift $z \gtrsim 2$ right at the peak of the cosmic SFR, the different isolation distance scenarios become more similar and all of them tend to agree with the observed data (see higher redshift panels). This makes it more difficult to use such an observable to discern among the seeding criteria, i.e., d_{iso}, at $z \gtrsim 1$.

2.2.3 *Local black hole mass function*

Here we present basic predictions for the mass distribution of the SMBH population at $z \sim 0$. As shown in Fig. 2.7, the three d_{iso} cases predict different trends for the Black Hole Mass Function (BHMF). As a reference, observational constraints are compared to our predictions, derived from a sample of local galaxies (shaded area Mutlu-Pakdil et al., 2016) and from the unbiased corrected relation from Shankar et al. (2020) (hatched area). Overall, we see that the $d_{iso} = 75$ kpc case predicts number densities of SMBHs in approximate agreement with the observational data. The

Table 2.1: Fitting parameters and derived quantities: the slope α, the intercept $M_{BH,9}$ (at $M_{star} = 10^9 M_\odot$), and dispersion σ are shown at z~0. We also report the average (geometric mean) mass of the 10 most massive BHs.

	z~0				
	50 kpc	**75 kpc**	**100 kpc**	**HMT**	**ALS**
α_{vh}	1.38	1.71	1.69	1.37	1.29
$\log(M_{BH,9,vh}[M_\odot])$	4.99	4.35	4.31	5.16	5.37
σ_{vh}[dex]	0.74	0.98	0.79	0.32	0.29
α_{high}	0.73	0.76	0.89	0.83	0.82
$\log(M_{BH,9,high}[M_\odot])$	6.08	6.05	5.97	6.06	6.19
σ_{high}[dex]	0.61	0.68	0.63	0.43	0.40
α_{low}	0.63	0.61	0.56	1.29	0.80
$\log(M_{BH,9,low}[M_\odot])$	6.19	6.20	6.19	6.18	6.32
σ_{low}[dex]	0.50	0.43	0.40	0.50	0.51
$\log(\overline{M}_{BH,max10}[M_\odot])$	8.99	9.26	9.23	8.62	8.65

$d_{iso} = 100$ kpc case falls slightly below the observed values, while the $d_{iso} = 50$ kpc case is slightly higher.

In the high mass end, we see that both the HMT and ALS models do not produce as many very high mass SMBHs as the Pop III.1 models. We expect this is due to a combination of factors: 1) competition between SMBHs for available gas mass; 2) a greater degree of AGN radio-mode feedback, due to larger occupation fractions, which reduces the amount of gas available for SMBH accretion. Conversely, the HMT and ALS models predict the presence of much larger populations of lower-mass SMBHs. In particular, the ALS model, which invokes seeds down to stellar mass scales, predicts a large population of IMBHs.

2.2.4 *The $M_{BH} - M_{star}$ relation*

Here we consider the relation between central SMBH mass and galactic stellar mass for the different models of SMBH seeding. Recall that, as described above, these results are based on the fiducial SMBH growth model adopted by F20. The scatter plots in Fig. 2.8 show the populations of seeded galaxies for the various seeding schemes (columns) at redshift zero (bottom row), with high redshift results shown in the other rows.

2.2.4.1 *Redshift zero results*

Focusing first on the local results, we see a general trend of more massive black holes being found in galaxies with larger stellar masses. Since the

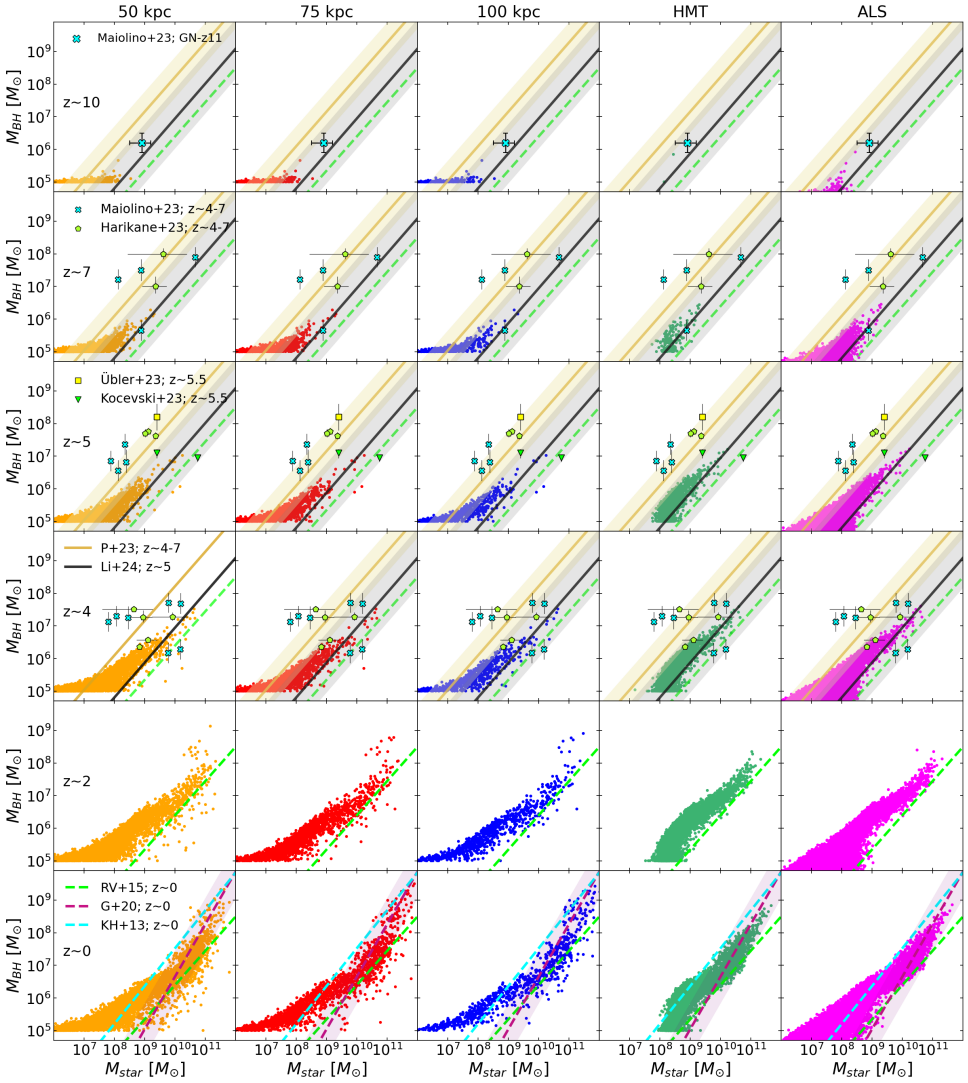


Figure 2.8: Scatter plot of the evolution of the $M_{BH} - M_{star}$ relation at several redshifts. Different columns represent various seeding schemes. At redshift $z \sim 0$ we compare our predictions for our seeded galaxies with several best fits from local observations: dynamical measurements from KH+13, including massive bulge-dominated quiescent galaxies, while R+15 use a different mass-to-light ratio, plus an extended sample combining local early- and late-type galaxies by Greene et al. (2020) (hereafter G+20). At $z \sim 2$, we report the R+15 fit for reference. Moving upward, the high-z rows show the comparison with recent results from the JFAINT sample: in Pacucci et al. (2023), they directly fit the data in the redshift range $z \sim 4 - 7$ while Li et al. (2024) estimate an unbiased fit taking into account the uncertainties on the mass measurements and selection effects. Shaded regions illustrate the intrinsic scatter at 1-sigma according to each relation. Coloured symbols show faint AGNs taken from Maiolino et al. (2023), Harikane et al. (2023), Übler et al. (2023) and Kocevski et al. (2023) and reported according to their redshifts. In the top row, results at $z \sim 10$ are shown against the single data point (GN-z11) from Maiolino et al. (2024).

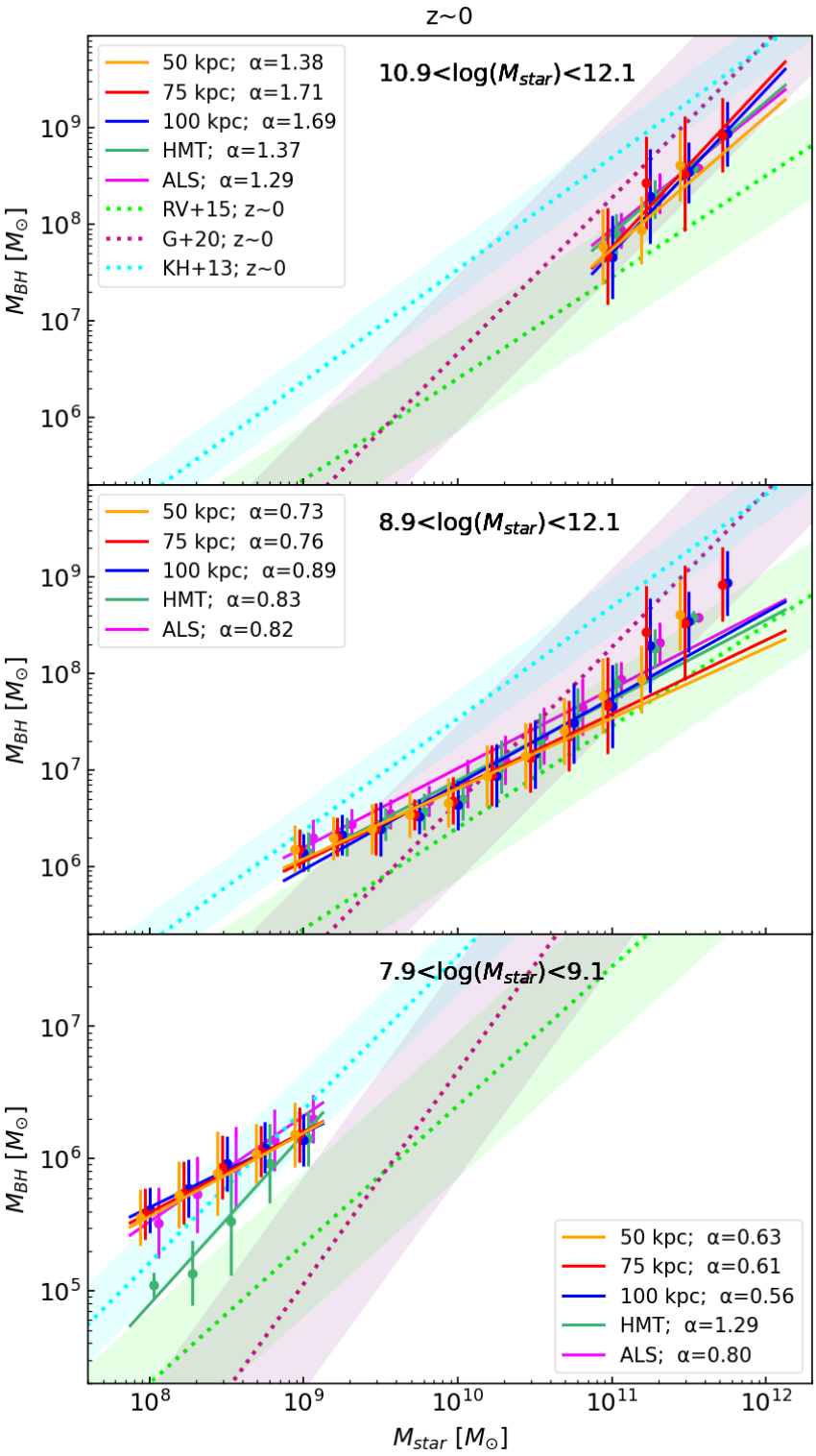


Figure 2.9: Power law fits to the $M_{BH} - M_{star}$ relation at $z \sim 0$. Dotted lines depict local fits with associated dispersion (shaded areas). The indices α are shown in the legend. Solid lines indicate the fit results to the median SMBH masses (data points with dispersion) in several stellar mass bins for different seeding models. The stellar mass range covered by the solid lines corresponds to the fitted interval.

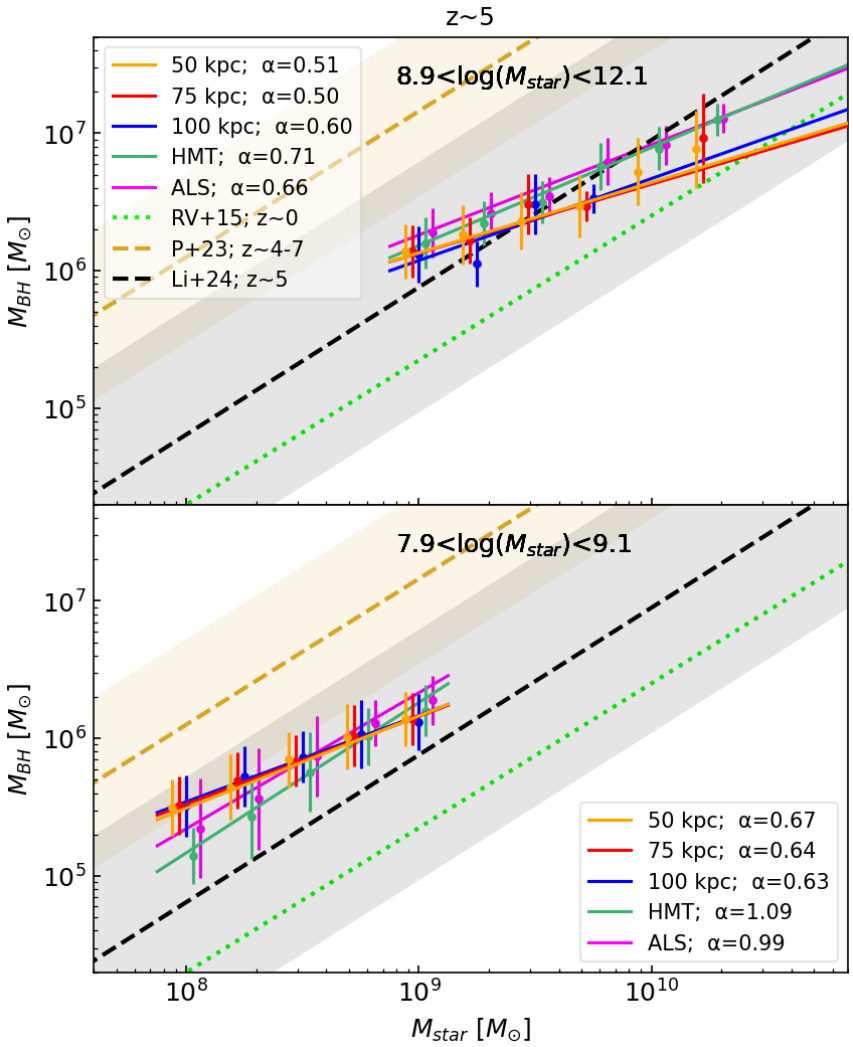


Figure 2.10: Power law fits to the $M_{BH} - M_{star}$ relation at $z \sim 5$. Dashed lines depict fits to the JFAINT data with associated dispersion (shaded areas). For reference, we also show the dotted line reporting the R+15 fit. The indices α are shown in the legend. *(a) Upper panel:* Solid lines indicate the fit results to the median SMBH masses (data points with dispersion) in several stellar mass bins for different seeding models. The stellar mass range covered by the lines corresponds to the fitted interval in the global range. *(b) Lower panel:* As (a), but the fits extend over the low-mass range.

Table 2.2: Fitting parameters and derived quantities: the slope α, the intercept $M_{BH,9}$ (at $M_{star} = 10^9 M_\odot$), and dispersion σ are shown at z~5. We also report the average (geometric mean) mass of the 10 most massive BHs.

	z~5				
	50 kpc	**75 kpc**	**100 kpc**	**HMT**	**ALS**
α_{high}	0.75	-	-	0.83	0.78
$\log(M_{BH,9,high}[M_\odot])$	6.08	-	-	6.13	6.19
σ_{high}[dex]	0.46	-	-	0.39	0.36
α_{low}	0.65	0.60	0.58	0.96	1.04
$\log(M_{BH,9,low}[M_\odot])$	6.07	6.02	6.01	6.13	6.23
σ_{low}[dex]	0.39	0.38	0.35	0.49	0.58
$\log(\overline{M}_{BH,max10}[M_\odot])$	6.75	6.65	6.56	6.91	6.93

Pop III.1 model assumes seed masses of $10^5\ M_\odot$, we see a flattening of the $M_{BH} - M_{star}$ relation to this level as one goes to lower mass galaxies. This feature is not seen in the HMT model since requiring a threshold halo mass of $7 \times 10^{10}\ M_\odot$ naturally imposes an effective threshold on M_{star} for seeded galaxies. This feature is also not seen in the ALS seeding scheme, since the seed masses in this model are much smaller than $10^5 M_\odot$.

In general, the $M_{BH} - M_{star}$ relation can be approximated as a power law, especially if one restricts to limited ranges in stellar mass. We thus fit the following function to the distributions:

$$M_{BH} = M_{BH,9} \left(\frac{M_{star}}{10^9\ M_\odot} \right)^\alpha \tag{14}$$

and consider a stellar mass of $10^9\ M_\odot$ as fiducial scale at which to divide *low-mass* and *high-mass* galactic systems. However, we set a lower limit of $M_{star} = 10^8\ M_\odot$, which is designed to make the metrics of the low-mass case easier to compare to observed systems. We also consider a *very high-mass* range at $M_{star} > 10^{11}\ M_\odot$. We carry out power law fits to the binned median values of M_{BH}, i.e., giving equal weight to the different bins of M_{star}, which are distributed evenly in logarithm and spaced every ~0.33 dex. We require at least 5 sources in a given bin to include it in the fit. Every median value is then weighted with its corresponding standard error, which in the least populated high-mass bins is about 0.3 dex. We report the maximum range of the black hole distribution via a metric $\overline{M}_{BH,max10}$, which is the geometric mean mass of the ten most massive black holes in the simulation domain of $\sim (60\,\mathrm{Mpc})^3$. We also measure the dispersion, σ about each best-fit power law $M_{BH} - M_{star}$ relation, averaging the dispersions in each mass bin equally. The redshift zero results for α, $M_{BH,9}$ and σ for the low-, high- and very high-mass cases, as well as $\overline{M}_{BH,max10}$, are reported in Table 2.1 for the various seeding schemes. These power law fits are also plotted in Fig. 2.9.

For galaxies with $M_{star} < 10^9\,M_\odot$, the properties of the best fit power law show that the Pop III.1 models have the most shallow power law indices, i.e., $\alpha_{low} \simeq 0.6$, with the $d_{iso} = 100$ kpc case having the shallowest index of 0.56. In comparison, the HMT model has $\alpha_{low} \simeq 1.3$, while the ALS model has $\alpha_{low} \simeq 0.8$. On the other hand the amplitude of the power law fits, as measured by $M_{BH,9}$, show very little variation ($\lesssim 0.1$ dex) between the models. The dispersion about the power law is similar among the models at 0.5 dex, except for the two larger d_{iso} cases where it drops at 0.4 dex for the 100 kpc model. This trend in the Pop III.1 models results from the fact that for smaller values of d_{iso}, a broader variety of halos is seeded, leading to a wider range of growth and star formation histories that then lead to greater scatter in the $M_{BH} - M_{star}$ relation.

In the high-mass regime, with $M_{star} > 10^9\,M_\odot$, the processes leading to black hole growth and star formation result in an $M_{BH} - M_{star}$ relation that is relatively similar between the different seeding schemes, so it becomes harder to distinguish the models from the statistics of their $M_{BH} - M_{star}$ relations. Nevertheless, from the results shown in Table 2.1 we notice that, in contrast to the low-mass regime, the Pop III.1 models have systematically higher values of dispersion about their best-fit power laws, i.e., $\sigma_{high} \simeq 0.7$, while the HMT model has $\sigma_{high} \simeq 0.5$ and the ALS model 0.4. Similar conclusions can be drawn for the very high-mass regime. We expect that this is caused by the Pop III.1 models having quite large volumes that are prevented from forming seeds, i.e., via the isolation distance criteria, and these volumes can contain relatively massive halos that form galaxies with relatively high M_{star} that either never host a SMBH (see the comparatively low occupation fractions in Fig. 2.5 of the Pop III.1 models) or gain a relatively low-mass SMBH from later merger with a smaller halo/galaxy.

We also notice that at the highest masses there is evidence for steepening of the power law behaviour of the $M_{BH} - M_{star}$ relation. This is most noticeable for the Pop III.1 models, which tend to have a greater number of more massive SMBHs (see Fig. 2.7). After inspecting the accretion histories of some example cases, we attribute this as being due to reduced impact of AGN feedback in the Pop III.1 models given their overall smaller numbers of SMBHs. Indeed, AGN feedback processes act as regulators of the cold gas content in galaxies, by heating their cold gas component and displacing it to the hot phase. This implies that unseeded galaxies in Pop III.1 models tend to accumulate larger cold gas reservoirs with respect to seeded ones. Therefore, when these galaxies became satellites in massive haloes and merge with a central seeded galaxies, they provide more cold gas available for accretion onto the central SMBH with respect to seeded galaxies (see §2.1.2.1), thus enhancing the final BH mass achieved compared to the HMT and ALS cases.

To quantify these differences, we refer to the power law fits for the very high-mass regime (see Table 2.1). In the top panel of Fig. 2.9 we see that for larger values of d_{iso} Pop III.1 models show a steeper slope up to ~1.7 for the 100 and 75 kpc cases, while other seeding criteria present

comparable shallower trends with power law indexes of about 1.3-1.4. This is a direct consequence of the already mentioned competition effect on the BH accretion (e.g., see also the BHMF in Fig. 2.7). In terms of scatter, we observe larger values in this stellar mass regime for the Pop III.1 model.

Related to the very-high mass regime of the $M_{BH} - M_{star}$ relation, we see that $\overline{M}_{BH,max10}$ is higher in the Pop III.1 models, i.e., $1 - 2 \times 10^9$ $M_\odot$, than in the HMT or ALS models, where it is $\lesssim 5 \times 10^8$ $M_\odot$. We again attribute this to reduced competition for gas in the Pop III.1 models due to smaller overall numbers of SMBHs. Also, the increased number of BH mergers in HMT and especially in ALS is not significant for the overall growth of the central SMBH, as the majority of mergers can contribute only up to a few percent of the total mass. Recall that the extra seeds in this two schemes are likely to happen in relatively low-mass halos since the most massive ones are seeded in all the scenarios.

We compare our $z = 0$ results to various observational constraints (see Fig. 2.9). In particular, we consider the empirical $M_{BH} - M_{star}$ relations obtained from: 1) the $M_{BH} - M_{bulge}$ relation from Kormendy and Ho (2013) (hereafter KH+13), who use dynamical measurements of massive bulge-dominated quiescent galaxies and who find an intrinsic scatter about the relation of about 0.3 dex; 2) the Reines and Volonteri (2015) (hereafter R+15) relation, which fits the same $M_{BH} - M_{star}$ data set of KH+13, but with an updated mass-to-light ratio, resulting in a lower normalization; 3) the Greene et al. (2020) (hereafter G+20) relation, which considers low-mass ($\sim 10^5$ $M_\odot$) SMBHs and their host galaxies. We note that the inclusion of low-mass objects in the G+20 fit down to stellar masses of $\sim 10^9 M_\odot$ causes the sample to be sparse and potentially biased for stellar masses below $10^{10} M_\odot$, where the number of objects measured via dynamical methods is small. We note also that low-mass and faint SMBHs are more likely to be missed in observational surveys, potentially biasing the derived relations (e.g., Shankar et al., 2020).

In the range of measured stellar masses from $\sim 10^9$ $M_\odot$ up to 10^{12} $M_\odot$, all theoretical predictions show agreement within the 1-sigma bands of the observational fits provided by R+15 and by the sample used in G+20. As discussed above, the largest differences between the theoretical models in this regime are in the very high-mass regime, so this may be a promising area for future, more detailed observational tests.

In the low-mass regime, i.e., $\lesssim 10^9$ $M_\odot$, there are larger differences between the models. As discussed the Pop III.1 models have shallower indices and higher amplitudes than the HMT and ALS models, and these appear to be less in agreement with the R+15 and G+20 observational results. However, as mentioned the data are relatively sparse in this regime and potentially subject significant systematic uncertainties and observational biases. From the theoretical point of view the precise amplitude and power law index is also sensitive to choice of a single seed mass of 10^5 $M_\odot$. Still, since the models show greatest differences in this regime, we consider that

improving the observational constraints at these lower values of M_{star} is a high priority.

2.2.4.2 *High-redshift results*

Moving to higher redshifts, in the first five rows of Fig. 2.8 we show the evolution of the $M_{BH} - M_{star}$ relation of the simulated seeded galaxies from $z \sim 10$ down to $z \sim 2$. To guide the eye, the light green dashed line shows the local relation from R+15. Table 2.2 reports power law fit results for the $z \sim 5$ sample.

As we proceed to higher redshift, the most obvious feature is that there are relatively few massive galaxies and associated highest-mass SMBHs. To help quantify this trend, Table 2.2 reports the results for $\log\overline{M}_{BH,max10}$ for the different seeding schemes at $z \sim 5$. We see that $\log(\bar{M}_{BH,max10} \simeq 7$ for most of the models, dropping to 6.56 for the Pop III.1 model with $d_{iso} = 100$ kpc. This is likely to reflect the fact that this case has the fewest SMBHs and so a reduced sampling of the relatively rare conditions that lead to the strongest growth. In the high-mass regime, the $d_{iso} = 75$ and 100 kpc cases do not form sufficient SMBHs for us to measure the $M_{BH} - M_{star}$ relation. For the other seeding schemes, we see quite similar power law fits, but with the Pop III.1 $d_{iso} = 50$ kpc case having moderately higher dispersion. In the low-mass regime, as at $z = 0$, we see significantly shallower indices in the Pop III.1 models ($\alpha_{low} \simeq 0.6$) compared to the HMT and ALS models ($\alpha_{low} \simeq 1.0$). Furthermore, the Pop III.1 models have smaller dispersions than the HMT and ALS cases. Among the different Pop III.1 models, the effect of varying the isolation distance d_{iso} is relatively hard to distinguish from the $M_{BH} - M_{star}$ relation fits.

Considering the evolution with redshift, among the various seeding schemes, the HMT is the one showing the largest evolution from $z \sim 10$ down to 4 as the process of seeding halos with SMBHs starts relatively late (i.e., $z \sim 10$, see Paper II). In the ALS seeding model, the BH seeds grow relatively fast and by $z \gtrsim 7$ they have already caught up with the SMBH populations formed from heavy seeds. This is due to the combination of the viscous accretion mode onto the BH and the small seed mass. In fact, seeding with stellar mass BHs allows these objects to accrete several times their own mass as the BH accretion rate goes up to 10 times the Eddington limit and they keep accumulating mass until they become massive enough to self regulate their own growth primarily via AGN radio-mode feedback (see also Fig. 2.11).

In Fig. 2.8 we also compare the theoretical models with recent results obtained from deep JWST galaxy surveys where it has been possible to find very faint AGNs otherwise undetected with other facilities (hereafter JFAINT sample). Due to the relatively small volume of our `PINOCCHIO` box (~ 60 : cMpc per side; see Section 2.1.1), caution is required when comparing these theoretical results with the full JFAINT sample, which is derived from

surveys covering larger volumes. In particular, larger simulation volumes are required to make comparison with the rarest, very luminous quasars.

Focusing on redshifts 4 to 7, we plot data from Maiolino et al. (2023) (twelve low-luminosity AGNs), Harikane et al. (2023) (ten objects), Kocevski et al. (2023) (2 objects) and Übler et al. (2023) (1 object). This JFAINT sample includes sources available from different surveys, specifically from the CEERS (Finkelstein et al., 2023), JADES (Eisenstein et al., 2023) and ERO (Pontoppidan et al., 2022) programs. In the $z \sim 10$ row we report the single data point (GN-z11) from Maiolino et al. (2024). Except for GN-z11 which displays narrow Balmer lines, it is worth noting that the JFAINT sample has been selected by looking for the broad component of the Hα and Hβ lines by means of NIRspec requiring a specific threshold for the Full Width Half Maximum (FWHM). Additionally, BH masses are measured using a relation connecting the BH mass to the FWHM of the broad line, mainly Hα, calibrated locally using the decomposition of the broad and narrow components of the emission lines (see Maiolino et al., 2024; Harikane et al., 2023). Stellar masses are instead typically estimated via SED fitting, which carries significant uncertainties to corrections related to dust attenuation, assumptions about metallicity, and AGN contribution.

This sample of JWST-detected AGNs has been used to estimate the $M_{BH} - M_{star}$ at earlier epochs. Pacucci et al. (2023) (P+23) performed a direct fit of the JFAINT data in order to derive the intrinsic $M_{BH} - M_{star}$ relation at a mean redshift of ~ 5 (golden solid line), concluding that high-z SMBHs tend to be overmassive by a factor of 10-100 with respect to the local relation. This interpretation implies that, even with very massive seeds, super-Eddington accretion episodes are required to frequently occur at high redshift. More recently, Li et al. (2024) (Li+24) have presented a study of the same sample of objects as in P+23 including a detailed analysis of the possible biases. In fact, uncertainties due the measurement of both BH and stellar mass and selection effects caused by flux-limited detection may lead to biased conclusions towards an overmassive population of BHs and their host galaxies. According to these authors, the observed data can be explained by assuming that the intrinsic $M_{BH} - M_{star}$ relation is more similar to the local one (e.g., of KH+13). Their result, shown by the black line, suggests that local and high-z relations behave in a similar way, where the JFAINT sample is an extremely biased selection towards the most luminous objects.

Our predicted BH masses are, on average, lower by one order of magnitude in BH mass with respect to the most massive BHs in the JFAINT sample (with mean redshift $\simeq 5.2$), while we can reach reasonably high stellar masses by redshift ~ 5. Our $M_{BH} - M_{star}$ relations do not match the P+23 fit, while showing a much better agreement with the estimate of Li+24. If the P+23 results are confirmed as a robust estimate for the true high-z $M_{BH} - M_{star}$ relation, there are two possible arguments that can explain the lack of very massive BHs in our model predictions. On the one hand the accretion scheme onto the central BH used in this study was calibrated by F20 in the GAEA framework to match AGN bolometric luminosity functions

up to redshift ~ 4. The fact that our realizations do not recover such high BH masses may be connected with the physics implemented in the growth model driven by gas momentum loss and viscous accretion, which sets the timescale of the accretion rate being inversely proportional to the BH mass as in Eq. (4). However, at higher redshift cold gas accretion may be expected to be a more continuous process (e.g., Inayoshi et al., 2016) that does not need trigger events such as galaxy mergers and/or disc instabilities to happen. If true, this could lead SMBHs to grow on shorter timescales and/or at higher efficiencies than we have so far implemented (see the discussion in Section 2.1.2.1).

Additionally, the limited size of our PINOCCHIO box also limits our ability to sample extreme, rare objects that undergo the fastest, most efficient accretion. This volume effect is particularly evident by looking at the top row at $z \sim 10$, where only single objects can occasionally match the GN-z11 data point within its error bars. We note that our simulated volume is at least a factor of 10 smaller than those probed by the JADES and CEERS programs.

Furthermore, the measurements of BH masses using local calibrations may be significantly biased upward (by up to one order of magnitude), as suggested recently by Abuter et al. (2024). Other considerations are that for the Pop III.1 and HMT models we seed galaxies with a delta function distribution of BH mass at 10^5 $M_\odot$. More realistically, we would include a range of masses with dispersion, and this would enhance the chance to obtain a few, larger SMBH masses.

2.2.5 *Eddington ratios*

Here we examine the distribution of Eddington ratios, λ_{edd}, that are present in our modelled SMBH populations. Figs. 2.11 and 2.12 show in different formats the distribution of λ_{edd} across different seeding models as a function of redshift. The solid lines in Fig. 2.11 show the median λ_{edd} of the different models, with shaded areas being the 1σ dispersion as a function of redshift. Fig. 2.12 shows the normalized histograms of λ_{edd} in ten redshift bins. For Pop III.1 models, the distributions, including median values, of λ_{edd} are relatively similar. We do notice that, especially at higher redshifts, the cases with larger d_{iso} have a higher proportion of higher accretion rate SMBHs. Comparing high and low redshifts, at higher redshifts the SMBHs tend to have higher accretion rates, with a median λ_{edd} of $\sim 10^{-3}$ to 10^{-2} down to $z \sim 1$ for $d_{iso} = 75$ and 100 kpc cases and $z \sim 2$ for $d_{iso} = 50$ kpc. After this, SMBHs start to self-regulate their growth via AGN feedback. This decreases the available gas which in turn lowers the average accretion to $\lambda_{edd} \sim 10^{-4}$ by $z \sim 0$. However, the histograms in Fig. 2.12 show that the SMBHs have a tail towards accretions close to the Eddington limit (or even super-Eddington), which would manifest as luminous AGN, including quasars.

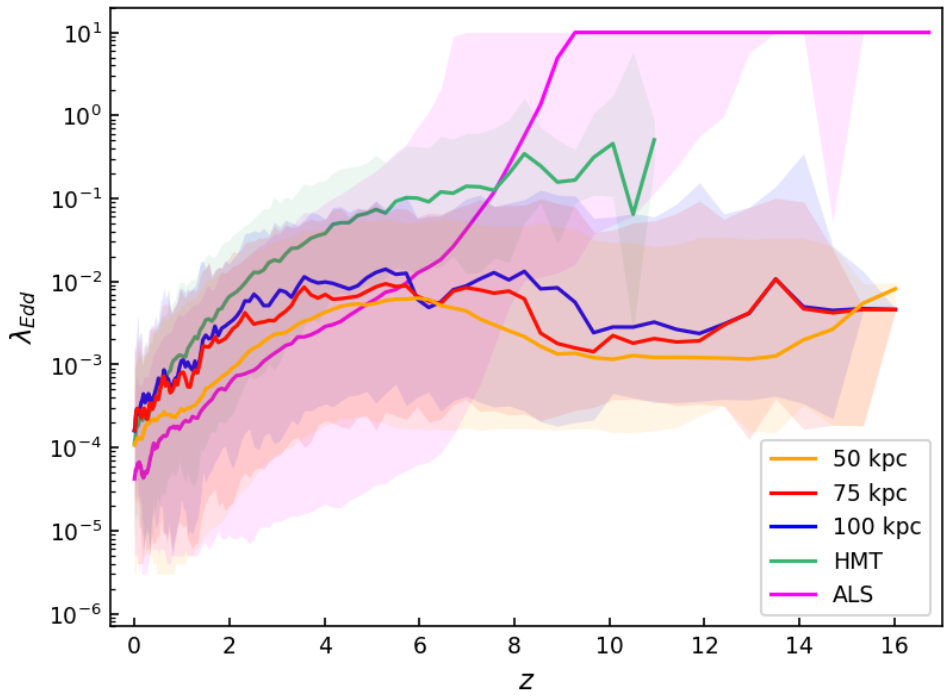


Figure 2.11: Medians (in log scale, solid lines) of the Eddington ratio λ_{edd} as a function of redshift for the different seeding mechanisms. Shaded areas denote the 1-sigma dispersion of the distributions.

On the other hand, the HMT and ALS models exhibit different distributions of λ_{edd}. In the case of HMT, the BH seeds appear at lower redshift ($z \sim 10$) in already massive halos where a substantial amount of gas available for accretion has gathered during its previous history. At this point, the central BHs are able to accrete efficiently resulting in λ_{edd} of about an order of magnitude larger than in the Pop III.1 models. However, below $z \sim 3$ the HMT BHs tend to align with the behaviour of the Pop III.1 populations.

For the ALS model, the BH seeds are significantly lower in mass than the other seeding models considered in this study. This explains the fact that down to $z \sim 8$ the large majority of BHs are accreting at super-Eddington rates (limited at 10 times as a basic model assumption). In fact, in the viscous accretion prescription the accretion rate is proportional to the ratio between the BH gas reservoir and the BH mass. Until the ALS BHs reach the supermassive regime, this ratio takes values much larger with respect to the other models, explaining the trends observed in Figs. 2.12 and 2.11. At lower redshifts, these systems tend to run out of gas available for BH accretion, as all halos are seeded (e.g., see the BHMF in Fig. 2.7). We note that by $z \sim 2$ and down to the local Universe, all the models have generally similar distributions, although the ALS model retains an excess of very high, super-Eddington accretors, mostly being low- and intermediate-mass BHs.

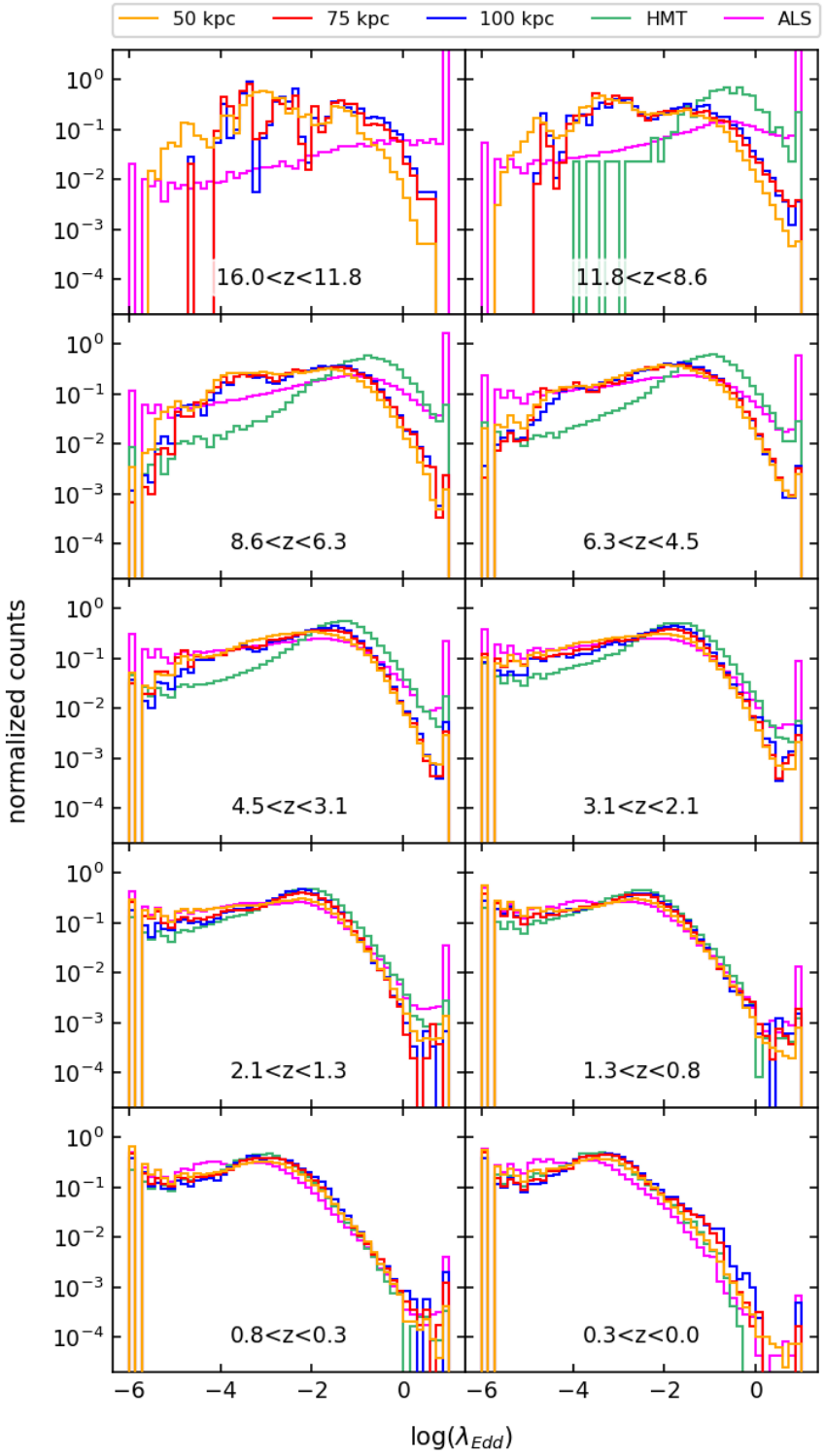


Figure 2.12: Normalized histograms (in log scale) of the Eddington ratio λ_{edd} for in different redshift bins as a function of the seeding mechanism. The spacing in redshift bins corresponds to a linear spacing in the logarithm of the scale factor a.

2.3 SUMMARY & CONCLUSIONS

In this Chapter we have introduced a novel semi-analytical approach that accounts for different SMBH seeding scenarios within theoretical models of galaxy formation and evolution. It utilizes merger trees generated by the PINOCCHIO code for dark matter halos, extends them by incorporating subhalos, and then applies the GAEA semi-analytical model to populate these halos with observable galaxies. This approach allows us to investigate a wide range of galaxy properties by adjusting various parameters governing galaxy formation without expensive N-body and/or hydrodynamical simulations. The evolution of subhalos and their merging with the main halos is implemented via physically motivated models that have been calibrated on simulations.

We first adapted the structure of PINOCCHIO merger trees to the Millennium simulation format by adding information about the subhalos. We assume a spatial distribution for subhalos following a Navarro-Frenk-White density profile (NFW, Navarro et al., 1997), statistical prescriptions for the angular momentum (Zentner et al., 2005; Birrer et al., 2014) and a subhalo survival time since accretion (Berner et al., 2022; Boylan-Kolchin et al., 2008). This ensures that the GAEA model will be able to run on PINOCCHIO-generated halo merger trees. We calibrated this method on a Millennium-like PINOCCHIO box making sure that we fed the semi-analytical model with a consistent halo mass function. By anchoring the calibration to the local observed GSMF, we estimated a total survival time for satellite galaxies, which is tuned to reproduce the exponential cut-off of the high-mass end of the predicted GSMF at $z \sim 0$. This approach makes it possible to apply our fully semi-analytical pipeline to a variety of scientific cases beyond the scope of the presented implementation.

We have primarily focused on the impact of implementing different mechanisms for seeding SMBHs, especially focusing on the Pop III.1 model, which postulates a new mechanism for the formation of all SMBHs. We have investigated three values of the isolation distance that is needed for a given minihalo to be a Pop III.1 source. For comparison, we have examined the HMT used by the Illustris-TNG simulations in which every halo exceeding a mass of 7.1×10^{10} $M_\odot$ is seeded with a BH of mass 1.4×10^5 $M_\odot$. As another example case, we also considered predictions for the standard seeding scheme implemented in GAEA based on ALS. Here, the initial mass of the BH seed scales with the initial halo mass, resulting in light seeds of the order of stellar mass BHs.

Within our framework, we have explored the implications of this set of seeding models when applied to cosmological volumes in a galaxy formation and evolution framework down to low redshifts. Unlike other astrophysical models, the Pop III.1 scenario presents the earliest and least clustered distribution of seeds, affording relatively longer periods of time for black hole growth via accretion, reducing the need for sustained modes of super-Eddington accretion.

A list of our main findings follows:

- **SMBH occupation fractions**: for the Pop III.1 models, SMBH seeds are predominantly abundant in massive galaxies, with the occupation fraction increasing as d_{iso} decreases. By $z \sim 0$, the occupation fraction reaches unity for halo masses above 10^{13} $M_\odot$ across all models. AGN feedback significantly influences the thermal state of the gas and the SFR, quenching seeded galaxies by $z \sim 0$. The observational measurements of the occupation fraction in the local Universe as a function of stellar mass suggest that $d_{iso} < 75$ kpc. In contrast, the HMT and ALS schemes produce too many BHs in systems with $M_{star} \sim 10^{9-10}$ $M_\odot$.

- **Local GSMF**: the AGN radio-mode feedback affects the shape of the GSMF by decreasing the SFR in massive galaxies, if hosting a SMBH, and causing the exponential cut-off at the high-mass end. At $z \sim 0$, the low-mass end of the GSMFs obtained from Pop III.1 models is generally similar to the predictions of the HMT and ALS models and with the observations, as in this regime the number densities are regulated by SN feedback. Above few $\times 10^{13}$ $M_\odot$, smaller isolation distances are favoured to reproduce the quenching of the majority of massive galaxies, consistent with $d_{iso} \lesssim 75$ kpc scenarios.

- **Local BHMF**: the slope of the BHMF at low SMBH masses is found to be a crucial way to distinguish different seeding models. Our favoured cases with $d_{iso} < 75$ kpc match the high-mass end of the BHMF and is in reasonable agreement with estimates at the low-mass end. However, it should be noted that the observational constraints are relatively uncertain in this regime and many lower-mass SMBHs may be missed in current surveys.

- **Local $M_{BH} - M_{star}$ relation**: the predicted $M_{BH} - M_{star}$ relations at redshift zero show some differences between the models. The main differences are a trend for the Pop III.1 models to have modestly steeper indices in the very high-mass regime, i.e., with $M_{star} > 10^{11}$ $M_\odot$. Larger differences are present in the low-mass regime, i.e., with $M_{star} < 10^9$ $M_\odot$, which reflects the imprint of the assumed mass scale of the seeds, i.e., $M_{BH} = 10^5$ $M_\odot$. We note that so far we have made very simple assumptions for this seed mass. We also note that the observational data in the regime are subject to significant uncertainties due to the difficulty of obtaining a complete census of SMBHs in this regime.

- **High-z $M_{BH} - M_{star}$ relation**: at high redshifts, comparison of the model $M_{BH} - M_{star}$ relations with observational constraints derived from JWST-detected AGN candidates remains open to debate as the probed ranges of BH and stellar masses are likely biased upward by the most luminous objects and not well sampled by the limited volume of the simulation considered in this study.

- **SMBH accretion rates**: the distribution of λ_{edd} suggests that, within the viscous accretion model for the BHs adopted in GAEA, massive seeds do not grow very efficiently in their early phases, while light seeds tend to accrete at the maximum rate allowed (i.e., 10 times Eddington).
- **Constraining** d_{iso}: all three local, $z \sim 0$ metrics of occupation fraction as a function of the galaxy stellar mass, galaxy GSMF, and BHMF suggest a constraint of $d_{iso} < 75$ kpc. Such a value places a constraint on physical models for the isolation distance, e.g., due to photoionization from the Pop III.1 source.

Expanding on this final point, a reference scale for radiative feedback from the Pop III.1 sources themselves is the radius of the Strömgren sphere of a supermassive, $\sim 10^5\,M_\odot$ protostar, which may have a final phase of evolution that involves being on or close to the Zero Age Main Sequence (ZAMS) for several Myr. Such a star is expected to have a H-ionizing photon luminosity of $S \sim 10^{53}$ H-ionizing photons per second and to heat its HII region to temperatures of $T \sim 30,000$ K. Then, the radius of the HII region adopting a mean intergalactic medium density is

$$r_{HII} = 59.6 S_{53}^{1/3} T_{3e4}^{0.27} \left(\frac{n_H}{n_{H,z=30}} \right)^{-2/3} \text{kpc}, \tag{15}$$

where $S_{53} \equiv S/(10^{53}\,s^{-1})$, $T_{3e4} \equiv T/(3 \times 10^4\,K)$, and $n_{H,z=30}$ is the mean number density of H nuclei in the Inter Galactic Medium (IGM) at $z = 30$. We note that this mean density scales as $(1+z)^3$, so by $z = 20$, the mean density drops by a factor of 0.310, which would increase r_{HII} by a factor of 2.18. We also note that the actual size of the HII region may be limited by R-type expansion, with the timescale to establish ionization equilibrium being longer than 10 Myr. This would tend to make the size of the HII region somewhat smaller than the estimate given in Eq. (15). In spite of these uncertainties, we consider the close correspondence of this ionization feedback scale with the constraint on $d_{iso} < 75$ kpc derived from our semi-analytical modelling of galaxy evolution and SMBH growth to indicate that this feedback process may well play an important role in setting the conditions for Pop III.1 supermassive star and SMBH formation, with the regions affected by the HII regions forming lower-mass Pop III.2 stars (Johnson and Bromm, 2006; Greif and Bromm, 2006b) (see Chapter 1).

3

PREDICTING HIGH REDSHIFT SMBHS IN THE JWST ERA

The JWST, launched at the end of 2021, has paved the way toward uncharted territories in the field of early AGN activity. Its scientific discoveries are already challenging our understanding of the high redshift Universe ($\gtrsim 4$) and pushing the limit beyond any expectation. The telescope's unparalleled Infra-Red (IR) spectral sensitivity and imaging resolution have led the very first detection of a population of fainter ($M_{UV} \sim -17$) AGNs and their host galaxies, featuring broad Balmer lines with FWHM $\gtrsim 1000 km/s$ at $4 < z < 7$ (Harikane et al., 2023; Maiolino et al., 2023; Kocevski et al., 2023; Matthee et al., 2024; Kocevski et al., 2024; Greene et al., 2024). Although the broad emission lines are generally believed to originate from the broad-line regions of the AGN, supported by the detection of a broad component of Balmer lines, the nature of these lines remains a subject of debate. Nonetheless, in this context, we will refer to these objects as AGNs. Among them, even more recent JWST observation have unveiled a significant population of sources, typically exhibiting a blue UV excess, red optical slope and relatively compact morphology. These objects are known as LRDs (e.g., Matthee et al., 2024; Kokorev et al., 2024; Greene et al., 2024; Kocevski et al., 2024; Akins et al., 2024).

The presence of AGNs in these objects, if further confirmed, would have a dramatic impact on early galaxy formation and evolution, the study of SMBH seeding schemes as well as on their role in the reionization of the Universe. Recent claims indicate that the estimated space density of JWST-detected AGNs could be as large as $10^2 - 10^3$ times the extrapolated faint end of the quasar luminosity function (Maiolino et al., 2023; Harikane et al., 2023; Kocevski et al., 2024). Assuming an escape fraction consistent with bright quasars, they could well become key actors in the overall ionizing budget and contribute to a significant fraction of the reionization process (e.g., Giallongo et al., 2015; Finkelstein et al., 2019; Giallongo et al., 2019; Boutsia et al., 2021; Grazian et al., 2022; Madau et al., 2024). Hence, the standard picture in which star forming galaxies are the primary sources responsible for the reionization (e.g., Robertson et al., 2015), may have to be revised.

However, this population of AGNs reveals rather peculiar and unique hallmarks. Most notably, they often go undetected in the X-ray surveys, contrary to expectations for typical AGNs (for instance, in terms of the ratio between Hα and X-ray luminosity). Even in stacked X-ray images of 71 JWST-detected AGNs between $2 < z < 11$, Maiolino et al. (2024) found no detection in Chandra observation, despite the BHs being significantly overmassive with respect to the host galaxy stellar masses when compared

to local scaling relations (e.g., R+15). In parallel, other studies have ventured the idea that the observed broad emission could arise from different processes. For instance, the intense emission and extreme compactness of LRDs could be explained via SED fitting of starburst galaxies (Pérez-González et al., 2024). Moreover, multi-epoch monitoring of relatively massive AGNs ($\sim 10^8 M_\odot$) has shown no variability over time scales longer than typically associated with an accreting SMBH of this mass (Kokubo and Harikane, 2024). Viable alternative options offered by this study span from fast outflows to Raman scattering of the UV stellar continuum in order to explain the broad emission. These various interpretations underscore the urgent need for further investigation into these peculiar yet widespread objects.

In this Chapter we delve into the observational properties of our modelled SMBHs and we focus our predictions on the latest results in the observational landscape, particularly within the latest JWST findings. The aim of this study is to test the Pop III.1 seeding model detailed in Chapter 1 and applied in galaxy formation and evolution models as presented in Chapter 2. Using this scheme, we utilize the model proposed by Cammelli et al. (2024) to make predictions on physical and observable quantities of AGNs, such as BH masses, AGN luminosity functions and Eddington ratios, at $z > 4$ and test whether it is possible to put some constraints on the possible SMBH seeding scenario. In particular we compare with data from recent works reporting a sample of faint AGNs detected with JWST.

In Section 3.1, we describe the methods and assumptions to estimate the properties of our SMBHs visible as AGNs. The main results and the comparison with the observed data are presented in Section 3.2, while Sections 3.3 and 3.4 summarize the results and discuss possible tensions and viable interpretations of such comparisons. The preliminary results of this Chapter will be presented in the future paper:

- **Vieri Cammelli**, Jonathan C. Tan, Pierluigi Monaco et al., in prep., "*The formation of supermassive black holes from Population III.1 seeds. IV. Observational implications for early SMBH populations in the JWST era.*"

3.1 METHODS

Here we extend the approach introduced in the previous Chapter and published in Cammelli et al. (2024) to predictions of observational properties of SMBH populations at high redshift. We will consider the seeding schemes outlined in Section 2.1.5.

3.1.1 *Inferring the AGN properties*

Basic AGN properties are inferred or derived directly from the outputs of the GAEA model coupled with the `PINOCCHIO` merger trees. In fact, BH masses and accretion rates onto the central BH are calculated on-the-fly within the algorithm and provided in the output. Note that the BH accretion

rate is derived according to Eq. 4. Instead, bolometric luminosities L_{bol} are calculated from the accretion rate $\dot{M}_{BH}$ following the approach of F20. For radiative efficient AGNs, meaning Eddington ratios λ_{edd} larger than 0.1, we have:

$$L_{bol} = \frac{\epsilon_{rad}}{1-\epsilon_{rad}} \dot{M}_{BH} c^2, \tag{16}$$

where c is the speed of light and ϵ_{rad} is the radiative efficiency, here assumed equal to 15%. For less efficient AGNs, defining L_{edd} as the Eddington luminosity, we consider (see Churazov et al., 2005; Hirschmann et al., 2014):

$$L_{bol} = 10 L_{edd} \lambda_{edd}^2. \tag{17}$$

Moreover, numerical simulations suggest that very low, Radio-mode, accreting BHs result in highly inefficient systems, with ϵ_{rad} as low as 2% (e.g., Sądowski and Gaspari, 2017). In the following, we do not attempt to model any dust obscuration correction to the intrinsic luminosity of the AGNs. Finally, in order to compare our predictions to observational constraints in the intrinsic UV band, we adopt bolometric corrections from Shen et al. (2020):

$$\frac{L_{bol}}{L_{UV}} = c_1 \left(\frac{L_{bol}}{10^{10} L_\odot}\right)^{k_1} + c_2 \left(\frac{L_{bol}}{10^{10} L_\odot}\right)^{k_2}, \tag{18}$$

where $c_1 = 1.862$, $k_1 = -0.361$, $c_2 = 4.87$ and $k_2 = -0.063$. Conversely, for the stellar emission coming from the host galaxies, GAEA already outputs a list of magnitudes in several pre-selected bands, including the UV. Attenuated magnitudes are also available and calculated within the model as detailed in De Lucia and Blaizot (2007) by means of a combined approach which accounts for both the homogeneous ISM component and the molecular clouds around newly formed stars (Devriendt et al., 1999; Charlot and Fall, 2000).

3.2 RESULTS

In the previous Chapter, we presented the $M_{BH} - M_{star}$ relation at various redshifts, comparing it against JWST observations in the range $4 < z < 7$, particularly from Maiolino et al. (2023), Harikane et al. (2023), Übler et al. (2023), and Kocevski et al. (2023). Despite the significant uncertainties in these observational measurements, it is important to note that our estimated BH masses are, on average, about one order of magnitude lower than those of JWST-detected faint AGNs at a given stellar mass, and they remain consistent with local scaling relations. We now turn to a comparison of the model Ultraviolet Luminosity Function (UVLF)s with several observational constraints. The UVLF is often employed as a key metric for evaluating our understanding of the physical processes occurring within galaxies and AGNs from a simulation perspective. Both star formation and AGNs exhibit

significant emission in the UV, and specific characteristics of the SED in this band can serve to disentangle between of the star formation rate and BH accretion rate. As a result, the UVLF reflects the overall activity within AGNs and their host galaxies.

Figures 3.1 and 3.2 show the evolution of the predicted UVLFs as a function of redshift ($4 < z < 9$) for different seeding schemes. As in Chapter 2, we maintain the same nomenclature and colour coding: orange, red, and blue represent the Pop III.1 model at different isolation distances d_{iso} of 50, 75, and 100 kpc, respectively; green indicates the HMT scheme; and magenta denotes the ALS scenario. The grey dotted horizontal lines indicate a value corresponding to 10 objects per magnitude bin across the entire volume of the box, serving as a conservative sampling limit for our simulations. The orange solid line represents the composite UVLF from the Pop III.1 model at 50 kpc in a dust-free scenario, meaning no attenuation has been applied to the stellar emission. The 75 and 100 kpc cases are not shown, as their composite UVLFs coincide with that of the 50 kpc model, making them indistinguishable at this scale. This overlap results from the limited volume of our simulation box, which extends the model predictions primarily within the galaxy dominated regime of the UVLF (e.g., Finkelstein and Bagley, 2022). Consequently, variations in the d_{iso} seeding distance, which influence SMBH abundances, have minimal impact on the composite UVLF. The orange shaded area indicates the possible range due to dust attenuation in the intrinsic stellar emission in the UV band, with the lower limit corresponding to the fiducial attenuation law described in Section 3.1. We also report the total (composite) UVLF for the HMT and ALS schemes as a green and a magenta solid line, respectively. Note that the total luminosity of the single objects is calculated by adding up both the AGN and host galaxy luminosity. Notably, all models exhibit only mild evolution of total UVLFs with redshift and generally reproduce the observational estimates reasonably well. It is important to recall that the model parameters have been calibrated to match local and lower redshift ($z \lesssim 4$) scaling relations and galaxy property distributions, including UVLFs (see Section 2.1). Thus, our estimated UVLFs constitute genuine predictions of the model.

The UVLFs of AGNs are instead shown in dotted lines. In order to count active SMBHs, in the AGN UVLF we select seeded galaxies with an Eddington ratio $\lambda_{Edd} > 0.1$. The relative contribution of AGNs to the total UVLF — i.e., the AGN UVLF — varies significantly depending on the seeding scheme. Focusing on the Pop III.1 models, they shows the expected trend observed in the previous Chapter regarding the SMBH occupation fractions as a function of various galaxy properties. As d_{iso} decreases, a greater fraction of halos meets the isolation criterion, leading to an increasing number of halos — and consequently, galaxies — being seeded with SMBHs. At fainter magnitudes (around $M_{UV} \sim -17$), the AGN UVLF more than doubles on average between the 100 kpc and 50 kpc cases across all redshifts. Similarly to the trend observed in the GSMF, we find that as we examine larger luminosities (and thus larger masses), the AGN UVLFs

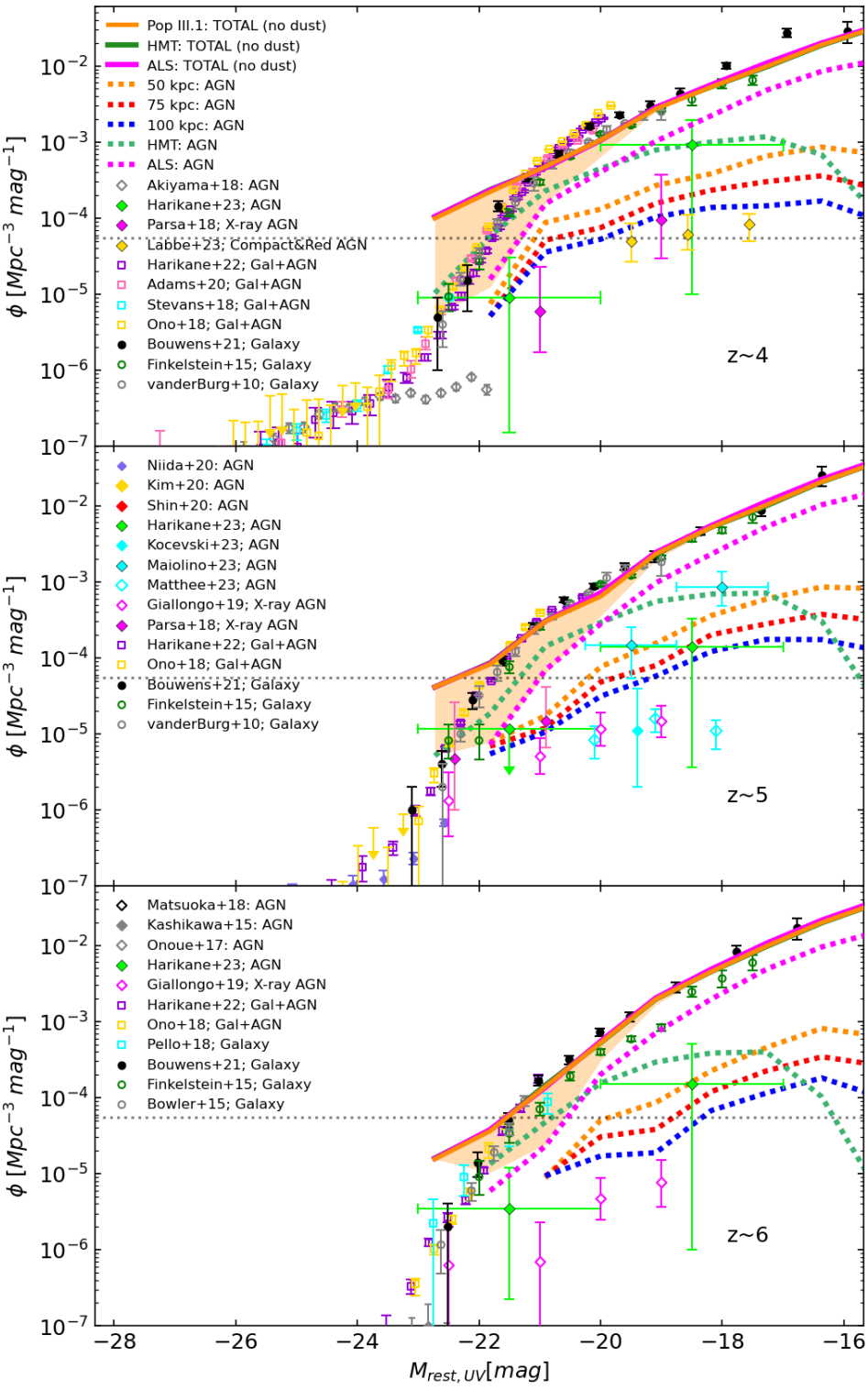


Figure 3.1: Cosmic evolution of the UV luminosity function at redshift from ~4 up to ~6 for different seeding cases, as labelled. Solid lines indicate model predictions of the composite total contribution from stellar and AGN emission by all galaxies with no dust attenuation. Dashed lines show the same total contribution from active ($\lambda_{\rm Edd} > 0.1$) seeded galaxies only (i.e. the AGN UVLF). The shaded area depicts the uncertainty due to dust attenuation at such early epochs, where lower limit represent the total dust obscuration as assumed by GAEA (see text). Several observations for galaxies, AGNs and composite UVLFs are shown, according to the legend.

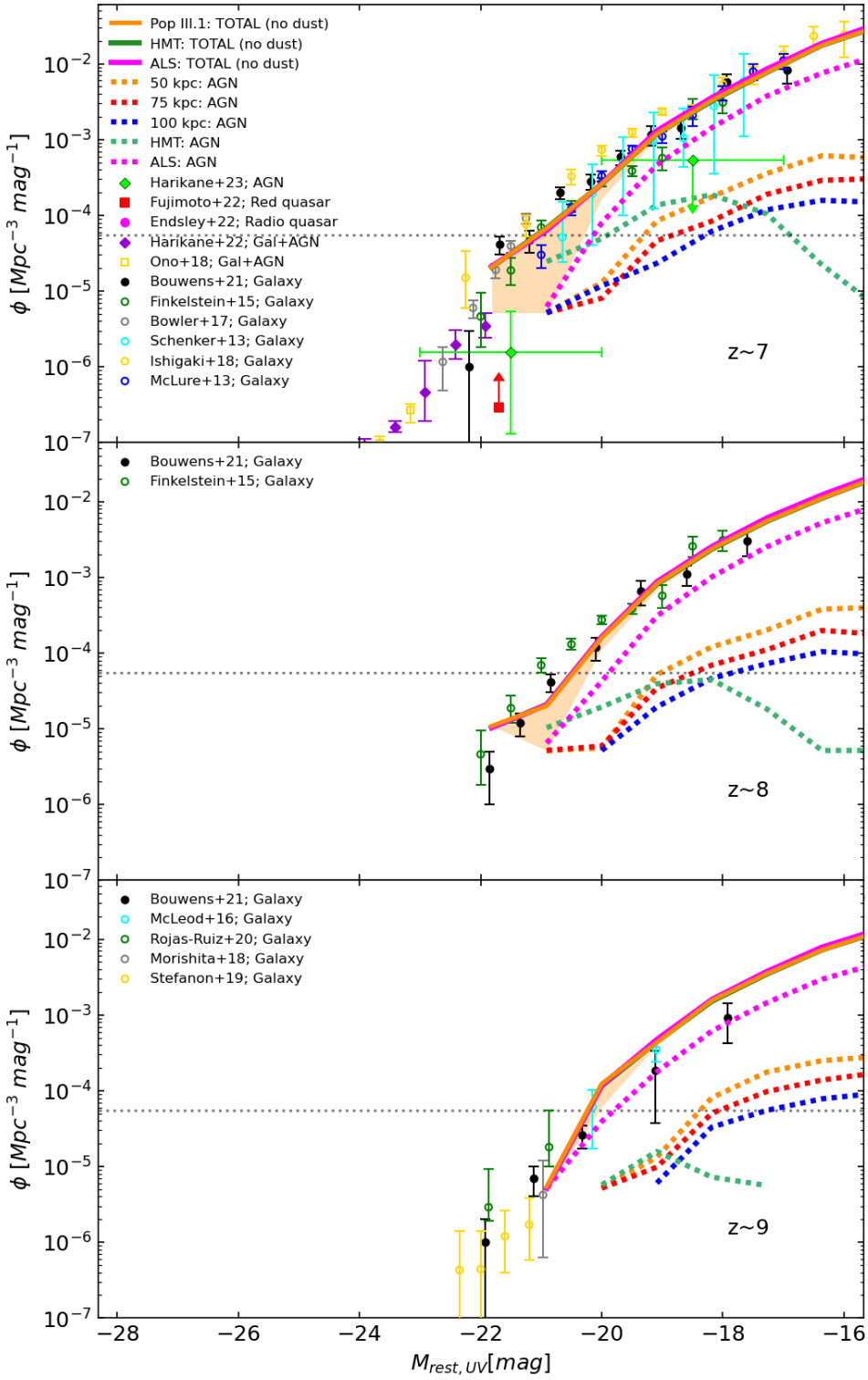


Figure 3.2: Same as Fig. 3.1 in the redshift range $z \sim 7 - 9$.

tend to converge towards a similar value. The most luminous galaxies, residing in the most massive halos, are seeded consistently, regardless of d_{iso}. The predictions for the ALS also show a mild evolution over the considered redshift range, with an average fraction of AGNs across different magnitudes of ~ 0.5. In contrast, the HMT model exhibits a different evolution for the AGN UVLF. This scheme seeds halos based on a threshold criterion, resulting in a rapid increase in the number density of SMBHs below a certain redshift, which is determined by the mass threshold. For the values considered here for the HMT scheme, the majority of seeds appear between redshifts 10 and 3. Consequently, the trend in AGN UVLF reflects this behaviour: at $z \sim 9$, only a few luminous objects are seeded, contributing marginally to the overall UVLF. As we move to lower redshifts, the fraction of seeded AGNs quickly approaches unity, already reaching this level at $z \sim 7$ for the most luminous sources. By $z \sim 4$, almost the entirety of the UVLF brighter than $M_{UV} \sim -20$ is dominated by AGNs.

For a comparison with observational data, in Fig. 3.1 we compile observational constraints for galaxy, AGN and composite UVLFs from several studies [1]. The latest JWST derived AGN UVLFs are depicted as diamonds [2]. It is noteworthy that the Pop III.1 predictions align relatively well with the UVLFs derived from Harikane et al. (2023) and Maiolino et al. (2023) within the redshift range of $4 \lesssim z \lesssim 7$, both in terms of slope and normalization. This model would naturally predict the large abundance of high redshift faint AGNs witnessed by JWST. In contrast, at the high-luminosity end, the HMT scheme generates a larger number of AGNs, exhibiting a much flatter slope compared to the Pop III.1 model and the observed UVLFs, while showing a decline at fainter magnitudes. The ALS seeding scenario predicts on average systematically larger values with respect to the observation, even though they marginally lie within the uncertainties. However, it would be hard to explain an AGN fraction of about 0.5 extending all the way to the faintest objects, e.g. the significant impact that AGNs would have on the reionization of the Universe, even though latest observations claim a non negligible contribution from AGNs (see the introduction to this Chapter).

3.2.1 *Comparison with JWST-derived properties*

In the previous Section, we showed that the comparison with the UVLFs based on UV magnitudes detected by JWST shows a remarkably strong agreement for the Pop III.1 model in terms of both shape and normalization.

1 Pre-JWST data are from Akiyama et al. (2018), Parsa et al. (2018), Labbe et al. (2023), Harikane et al. (2022), Adams et al. (2020), Stevans et al. (2018), Ono et al. (2018), Finkelstein et al. (2015), Bouwens et al. (2021), van der Burg et al. (2010), Giallongo et al. (2019), Niida et al. (2020), Kim et al. (2020), Shin et al. (2020), Matsuoka et al. (2018), Kashikawa et al. (2015), Onoue et al. (2017), Pelló et al. (2018), Bowler et al. (2015), Fujimoto et al. (2022), Bowler et al. (2017), Schenker et al. (2013), Ishigaki et al. (2018), McLure et al. (2013), McLeod et al. (2016), Rojas-Ruiz et al. (2020), Morishita et al. (2018) and Stefanon et al. (2019)

2 JWST observational data are from Harikane et al. (2023), Maiolino et al. (2023), Kocevski et al. (2023) and Matthee et al. (2023)

This finding encourages further exploration of this comparison. Specifically, Harikane et al. (2023) and Maiolino et al. (2023) provide estimates for bolometric luminosities and BH masses for these proposed faint AGNs. As discussed at the beginning of this Chapter, these measurements are based on the broad emission of Balmer lines, particularly Hα. The measured line fluxes are used to derive line luminosities, which are then converted into bolometric luminosities through a bolometric correction. Additionally, the mass of the BH can be inferred by assuming a relationship which depends on the luminosity and the FWHM of the Hα line (see next Section). It is important to note that these relationships, calibrated at $z \sim 0$, are assumed to remain valid within the redshift range examined in the surveys. These estimates can be directly compared to the corresponding properties derived from GAEA, as detailed in Section 3.1.

In Fig. 3.3 we focus on the comparison between the estimated bolometric luminosities L_{bol} and BH masses M_{BH} as from Harikane et al. (2023) and Maiolino et al. (2023) against 2 specific realizations of our seeding schemes: the Pop III.1 model with $d_{iso} = 75$ kpc (top panel) and the HMT scheme (bottom panel). Given the range of luminosities, BH masses and UVLFs, we consider that these two realizations are representative and informative about the comparison. As the average redshift of the JWST sources is about ~ 5.2, we compare with model predictions at a similar redshift ~ 5. In both panels, we see that our model predictions fail to reproduce the bulk of the JWST data points. The Pop III.1 model produces on average bolometric luminosities and BH masses that are at most $\gtrsim 1$ order of magnitude lower than the observational constraints. In the bottom panel, the HMT scheme shows similar results in terms of the most extreme objects, although its overall distribution indicates a higher level of activity compared to the Pop III.1 model. This is consistent with the distribution of Eddington ratios described in Figures 2.11 and 2.12. However, as discussed in the previous Section, the HMT model tends to slightly overpredict the AGN UVLFs at $z \sim 5$ (see Fig 3.1) while the Pop III.1 scenario aligns well with the observational data.

This denotes a complex picture: although the Pop III.1 model reproduces the AGN UVLFs for the faint JWST-detected AGNs, and the HMT case even overpredicts them, both scenarios fail to match the quantities estimated from JWST Hα broad emission measurements. It is important to note that, while this may partly arise from the limited volume of our simulation box — preventing access to larger luminosities and BH masses — the apparent inconsistency with AGN UVLFs would persist. In fact, observationally, the estimated BH masses and bolometric luminosities as well as the AGN UVLFs are derived from the same set of objects. Conversely, if using similar assumptions as above, running our simulations on a larger box would indeed recover more massive and luminous BHs; however, they would appear at brighter magnitudes in AGN UVLFs compared to the observed data. In the following section, we explore this apparent tension further and propose potential issues.

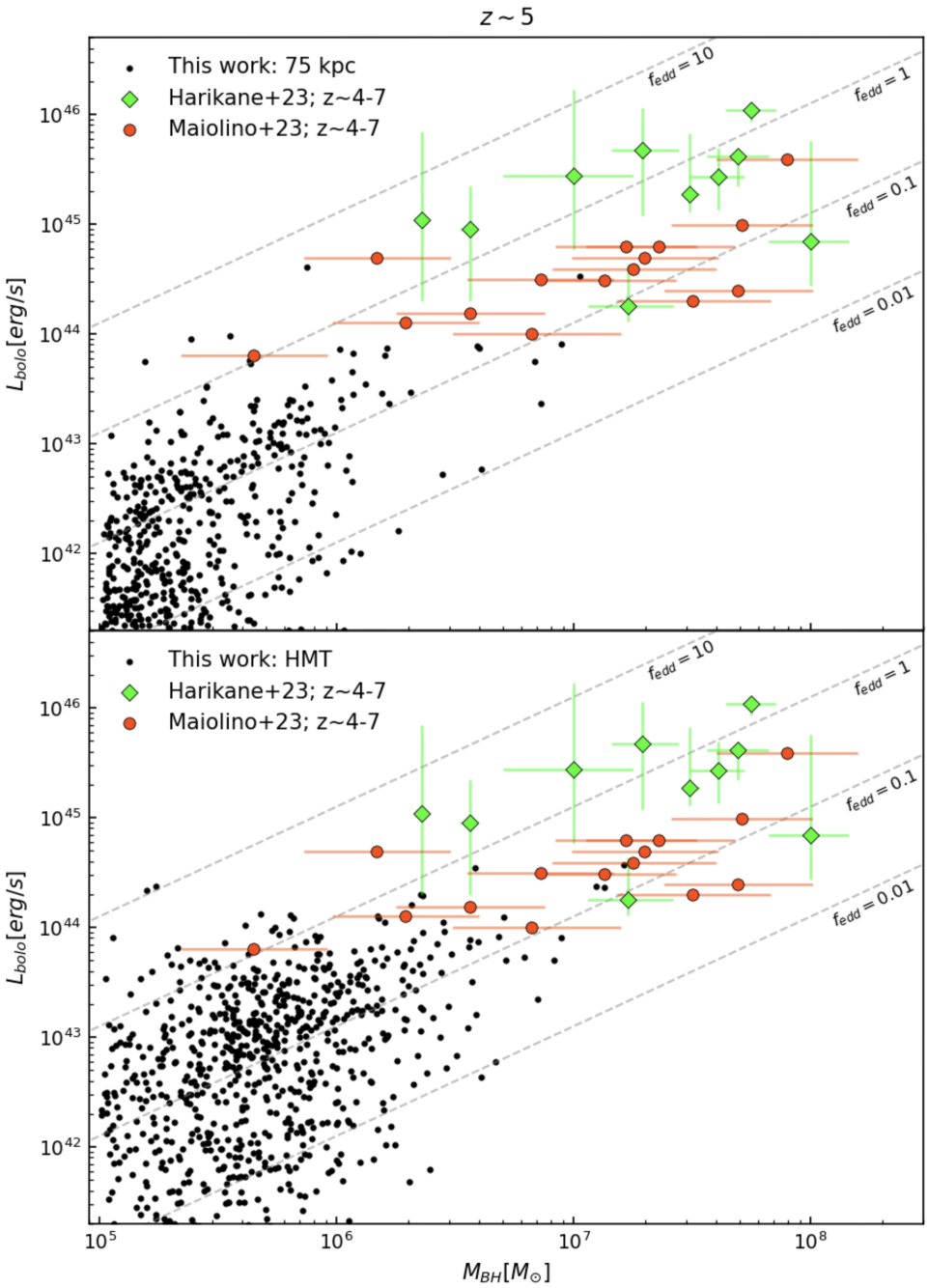


Figure 3.3: Bolometric luminosities as a function of the BH masses. Coloured diamonds show estimates from Harikane et al. (2023) and Maiolino et al. (2023) with an average redshift of ~ 5.2. *Top panel:* Black points represent model predictions for the Pop III.1 model with $d_{iso} = 75$ kpc at redshift $z \sim 5$. *Bottom panel:* same as the top panel for the HMT seeding scheme. Dotted gray straight lines denote the bolometric luminosity relative to the Eddington limit for several fractions, from 1 down to 0.01, for a given BH mass. This comparison indicate that our simulated SMBH fail to mimic the activity and the mass of SMBHs as inferred by the Hα line detected in JWST sources.

3.3 DISCUSSION

Before discussing the inconsistencies between our predictions and JWST-derived data, it is important to first review the potential sources of uncertainty that may affect the observational estimates. The AGN UVLFs, as described by Harikane et al. (2023) and Maiolino et al. (2023), are calculated by applying a multiplication factor to the galaxy UVLF, which accounts for the fraction of AGNs identified in the survey based on redshift and magnitude bins. While it seems reasonable to assume that the AGN UVLF behaves similarly to the galaxy UVLF in this magnitude range, this assumption may introduce systematic biases. Furthermore, the number of detected AGNs in these surveys is highly sensitive to the selection criteria used (e.g., Hα FWHM, signal-to-noise ratios, and the FWHM of forbidden lines). Even small variations in these criteria can lead to significant differences in the number of AGNs identified.

On the other hand, estimates of BH masses and bolometric luminosities rely on locally calibrated relations, which may not be applicable at higher redshifts. Recent work by Abuter et al. (2024) suggests that using relations calibrated at $z \sim 0$ may overestimate BH masses by a factor of 3-4, or even up to 10. Similarly, bolometric corrections at high redshifts may behave differently, especially given the unique properties of these AGNs, as discussed in the introduction of this chapter.

The arguments presented suggest that significant uncertainties may be influencing the results, and caution should be exercised when interpreting them. Nevertheless, we treat these estimates as direct values and propose an alternative perspective on the conclusions drawn by Harikane et al. (2023) and Maiolino et al. (2023). As noted earlier, we are unable to simultaneously reconcile the AGN UVLFs with the inferred BH masses and/or bolometric luminosities. Specifically, while we can reproduce the AGN UVLF, the resulting BH masses are significantly lower by up to $\gtrsim 1$ order of magnitude. Moreover, our BH accretion rates are a fraction of the Eddington rate (see Fig. 2.12), and we envisage no peculiarities there. To address these discrepancies, provided that BH masses and bolometric luminosities are estimated accurately, Harikane et al. (2023) proposed that a high extinction of $A_v \sim 4$ magnitudes in the UV band could be assumed to explain the reported trend in the UV luminosity function. However, this would result in a bolometric AGN LF that might conflict with reasonable extrapolations from lower redshifts, potentially exceeding current constraints on the quasar bolometric LF at $z \sim 3$ (e.g., Fiore et al., 2012; Ueda et al., 2014).

This sample of AGNs detected by JWST was initially selected using NIRCam imaging, which enabled measurements of M_{UV}. The targets were then followed up with NIRSpec to obtain spectroscopic redshifts and detailed spectra. While the NIRCam observations rely on straightforward photometry of the sources, the NIRSpec data require spectral line fitting and decomposition to separate the AGN emission from that of the host galaxy. Although the spectral information provides crucial insights into the nature

of the emission — whether originating from stars or the AGN — it may be subject to systematic errors, particularly when the spectrum is poorly resolved, and the derived quantities depend heavily on strong assumptions, as previously discussed.

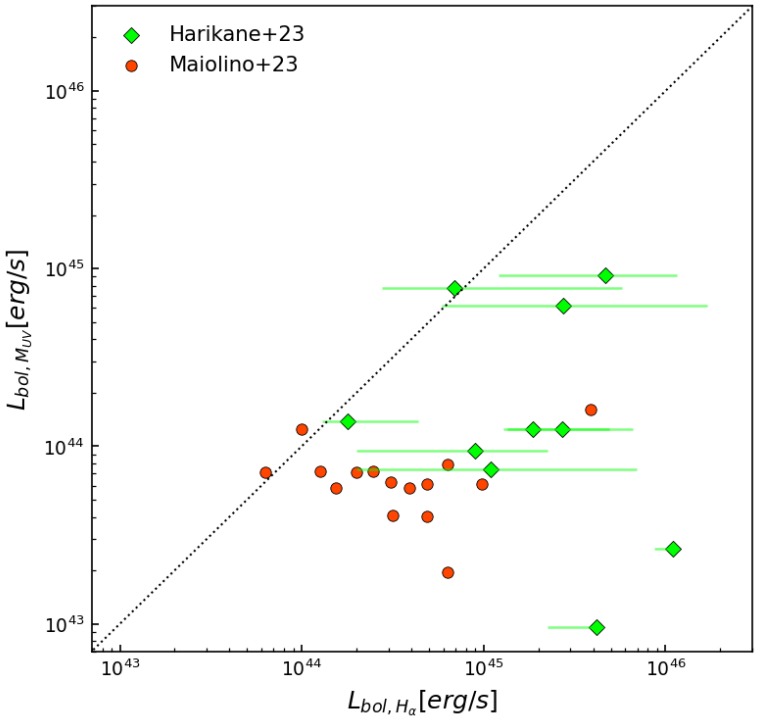
As a key test to address potential caveats, we calculate black hole masses and bolometric luminosities directly from the UV magnitude (M_{UV}), deriving first the bolometric luminosity and then the Hα luminosity. For bolometric corrections, we adopt again those from Shen et al. (2020) in Eq. (18) and the methods described by Harikane et al. (2023) and Maiolino et al. (2023). The black hole masses are estimated in terms of the width $FWHM_{H_\alpha}$ and the luminosity L_{H_α} of the broad component of the Hα:

$$\log\left(\frac{M_{BH}}{M_\odot}\right) = C + \alpha \log\left(\frac{L_{H_\alpha}}{10^{42}\ erg/s}\right) + \beta \log\left(\frac{FWHM_{H_\alpha}}{10^3\ km/s}\right), \tag{19}$$

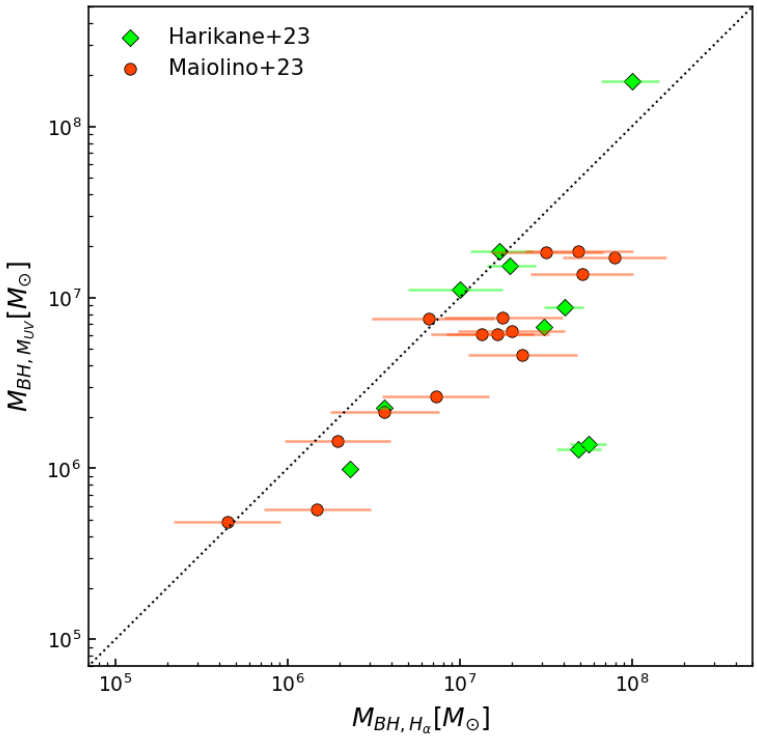
where the values of the normalization factor C and the exponents α and β vary according to Harikane et al. (2023) ($C = 6.30$, $\alpha = 0.55$ and $\beta = 2.06$) and Maiolino et al. (2023) ($C = 6.60$, $\alpha = 0.47$ and $\beta = 2.06$) and reference therein. For the $FWHM_{H_\alpha}$, we use the measurements reported in these works. In our estimates, we do not attempt to fully deblend the AGN component from the host galaxy's light. Instead, as a first approximation, we assume that the entire UV radiation observed in M_{UV} comes from the AGN, meaning that our estimated values should be considered upper limits for both the bolometric luminosities and the black hole masses.

In Fig. 3.4 we show the results of our measurements of bolometric luminosities (top panel) and BH masses (bottom panel) for the sample of AGNs detailed above. For bolometric luminosities, our M_{UV}-inferred values, despite being treated as upper limits, are typically $\gtrsim 1-2$ dex smaller than those derived from the Hα luminosity, with only a few objects displaying consistent L_{bol}. The bottom panel shows the estimated BH masses via the two different methods. Similar to the bolometric luminosities, the M_{UV}-inferred masses tend to be smaller, by up to one order of magnitude, though the discrepancy with the Hα-derived estimates is less pronounced compared to the bolometric luminosities. This difference is partly due to the fact that BH mass estimates from Eq. (19) depend on bolometric luminosity, via Hα luminosity, raised to the power of β, which is typically calibrated to be around 0.5 at $z \sim 0$.

As a consistency check, we also attempt to compare our M_{UV}-inferred masses and luminosities from the sample of Harikane et al. (2023) and Maiolino et al. (2023) against model predictions as in Fig. 3.3. Fig. 3.5 shows the comparison with the synthetic data based on the Pop III.1 model with d_{iso} of 75 kpc. For the M_{UV}-inferred estimates, we find that the tension with the predicted AGN UVLFs, previously highlighted in Fig. 3.1, is partially alleviated. Our model predictions now successfully capture the majority of the JWST-observed sample, particularly for the most luminous and massive objects.



(a)



(b)

Figure 3.4: *Top panel:* upper limits for bolometric luminosities estimated from M_{UV} against Hα-inferred measurements as from Harikane et al. (2023) and Maiolino et al. (2023). *Bottom panel:* same as in the top panel but for BH mass estimates.

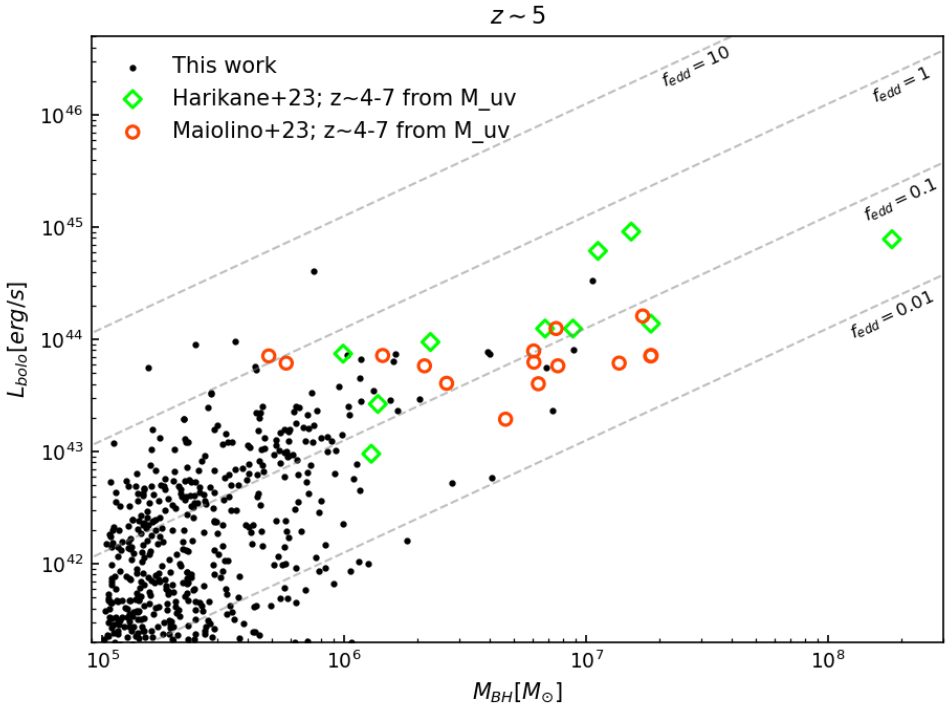


Figure 3.5: Bolometric luminosities as a function of the BH masses. Similarly to Fig 3.3, coloured empty diamonds show estimates from Harikane et al. (2023) and Maiolino et al. (2023) now derived from their UV magnitude M_{UV} as explained in the text. Black points display model predictions for the Pop III.1 model with $d_{iso} = 75$ kpc at redshift $z \sim 5$. This comparison indicates that our simulated SMBH partially reproduce the activity and the mass of SMBHs as inferred by the M_{UV} detected in JWST sources.

Given these results, we consider that it will be crucial to further investigate the apparent tension between observed data and predicted quantities. Our model predictions can only partially align with the observational results, as shown in Section 3.2. We recall here that model predictions are based on basic and widely used assumptions (see Section 3.1) and we do not envisage systematic issues dramatically affecting the model estimates. A potential caveat lies in the limited observational sample, likely under-representative of the broader SMBHs population at these epochs (e.g., Li et al., 2024) and potentially representing the rarest and most luminous objects. Aside the discrepancies between our predictions of bolometric luminosities and BH masses and the AGN UVLFs, we notice a potential inconsistency within the JWST-derived data themselves. Our revised observational estimates contrast with the values reported in literature. Even under the assumption that AGN activity dominates the UV emission in these objects, i.e. effectively providing upper limits for the derived AGN properties, our M_{UV}-inferred estimates lie typically well below those inferred from Hα estimates (see Fig. 3.4).

3.4 CONCLUSIONS

In this Chapter we utilise the semi-analytical approach introduced in Chapter 2 to make predictions of observable quantities of SMBHs in terms of AGN properties. We focus our effort in comparing model predictions of AGN physical properties against observational constraints, such as BH masses, bolometric luminosities and UVLFs . We collect forecasts from the different SMBH seeding scenarios as detailed in Section 2.1.5 in the redshift range from $4 < z < 9$, with particular emphasis on $z \lesssim 7$ where recent results from JWST observational campaigns are available. Our aim is to test whether the comparison with high redshift observational data can help in disentangling different SMBH formation mechanisms, including the Pop III.1 model. Given the large uncertainties and apparent inconsistencies in the JWST estimated properties, we also attempt to re-derive upper limit measurements and discuss possible issues.

Our main results include:

- **UVLFs**: the total combined UVLFs show strong agreement with observed data, particularly at the brightest magnitudes, within the uncertainties of dust attenuation. Across all SMBH seeding models, the UVLFs exhibit mild cosmic evolution up to $z \sim 9$, with no significant dependence on the seeding scheme. The Pop III.1 model accurately reproduces AGN UVLFs for faint AGNs detected by JWST, while the HMT model tends to overestimate them, especially for $M_{UV} \lesssim -20$.
- **BH Masses and Bolometric Luminosities**: while the Pop III.1 model aligns with the AGN UVLFs observed by JWST, it fails to match the data for BH masses and bolometric luminosities, with differences

up to $\sim 1-2$ orders of magnitude. The HMT model shows similar discrepancies, despite its overprediction of the bright-end AGN UVLF.

- **Revised BH Mass and Luminosity Estimates**: to address this inconsistency, we revisited the methods used to derive bolometric luminosities and BH masses in the literature. Instead of using Balmer line fluxes and FWHMs, we calculated upper limits based on the assumption that the UV emission is entirely AGN-driven. These UV-based estimates differ significantly from the Hα-derived values, with the latter exceeding our upper-limit estimates for both bolometric luminosities and BH masses.
- **Partial Reconciliation with Model Predictions**: comparing model predictions with M_{UV}-derived estimates partially resolves the observed tension, but highlights the need for future observations to address potential sources of systematic errors. These may include issues with bolometric corrections, scaling relations for BH mass estimates, and line flux/FWHM measurements.
- **Further Investigation Required**: current observational constraints are insufficient to distinguish between different SMBH formation models. Additional analysis is needed to reduce uncertainties and resolve key inconsistencies between model predictions and observational data.

4

A SEARCH FOR SMBHS VIA VARIABILITY ACROSS COSMIC TIME

The previous Chapters have highlighted the pressing need for observational constraints on SMBH formation theories. The search for AGNs in the early Universe is driven by several key factors. Each AGN detected guarantees the presence of at least one SMBH, providing a lower limit on the abundance of SMBHs. Accurately measuring their abundance is essential for testing competing SMBH formation theories (e.g., Rees, 1978; Volonteri, 2010; Banik et al., 2019; Inayoshi et al., 2020; Singh et al., 2023; Hayes et al., 2024). The observed number density of SMBHs, along with their inferred seeding mechanisms, serves as a crucial empirical input for galaxy formation simulations, where AGN play a fundamental role in regulating galaxy evolution, luminosity functions, star formation quenching, and other critical processes. The most valuable insights are expected to emerge from the largely unexplored epoch before $z \sim 6$. Recently, thanks to the groundbreaking capabilities of the JWST launched in 2021, numerous studies have reported the detection of AGNs during the reionization period (e.g., Harikane et al., 2023; Matthee et al., 2023; Greene et al., 2023; Maiolino et al., 2023; Bogdán et al., 2024; Kokorev et al., 2024). However, to meaningfully differentiate between competing formation theories, studies must be able to constrain the comoving density of SMBHs down to levels of 1 per 10^3 Mpc^3, corresponding to host galaxies with luminosities significantly below the characteristic luminosity L^* (Banik et al., 2019).

In this Chapter we introduce what we consider may be a crucial diagnostic that can independently identify AGN across cosmic time and different galaxy populations, namely *time variability*. Specifically we address this task by re-imaging the historic deep HST fields, the HUDF, thereby studying variability in the faintest galaxies, at the volume densities needed to discriminate between SMBH seeding scenarios. In Hayes et al. (2024) (hereafter H24) we report first results from the detailed study of variability between different epochs. The main result of H24 is the report of three high-z, i.e., between $z = 6 - 7$, AGN candidates identified via variability and implication of these results for $n_{SMBH} \gtrsim 7 \times 10^{-3}$ cMpc^{-3} and thus possibly constraining SMBH seeding models. Here we present the full variability analysis of the HUDF09 (GO 11563, PI: Illingworth), HUDF12 (GO 12498, PI: Ellis) and HUDF23 (GO 17073, PI: Hayes) data sets, including detection of AGN at various thresholds of significance and over the full redshift range of the source population.

In Section 4.1, we describe the observational campaigns and the methodology applied in the data reduction as well as the analysis to select

Table 4.1: Observing epochs and times.

Year	Filters	Orbits	GO# / PI
2008-9	F105W	24	11563 / Illingworth
	F160W	53	
2012	F105W	72	12498 / Ellis
	F140W	30	
	F160W	26	
2023	F140W	30	17073 / Hayes

This table lists the observing epochs, filters used, orbits, and the corresponding GO numbers and PIs.

variable candidates for AGNs. The main results are presented in Section 4.2, while Section 4.3 discusses variability as a method to seek SMBHs. The implications in the landscape of SMBH seeding formation mechanisms are summarized in Section 4.4. The results presented in this Chapter will be part of the future paper:

- **Vieri Cammelli**, Jonathan C. Tan, Matthew J. Hayes, Richard S. Ellis, Pierluigi Monaco et al., in prep., "*Glimmers in the Cosmic Dawn. II. A census of supermassive black holes via photometric variability across cosmic time.*"

4.1 DATASETS AND METHODS

In this Section we present the observational datasets and the methodology used in this study, following a similar approach to H24, with particular emphasis on the search for variable objects.

4.1.1 *Observations & data reduction*

To enhance our sensitivity to high-redshift variable AGN, we targeted the HUDF due to its provision of samples with the highest comoving volume density. This field was initially observed in the optical with Advanced Camera for Surveys (ACS) (Beckwith et al., 2006), followed by key Near Infra-Red (NIR) mode observations with WFC-3/IR during 2008-2009 under HUDF09, capturing images in three filters (F105W, F125W, and F160W) over 192 orbits. In 2012, the field was re-imaged under HUDF12, significantly deepening the F105W and F160W exposures and adding a fourth filter, F140W, to search for Lyman break galaxies at $z \sim 8$.

To search for photometric variability across all sources within the HUDF IR footprint, we re-imaged the field in September 2023 using the F140W filter, replicating the centre, field orientation, and depth (30 orbits) of the HUDF12 observation. We processed the F140W image with the `calwfc3`

pipeline and `astrodrizzle` (STSCI Development Team, 2012) software, using High-Level Science Products (HLSP) from the Mikulski Archive for Space Telescopes (MAST) as reference images. Additionally, we re-processed the F140W image from the UDF12 campaign to confirm our methods align with the depth of the HLSP image. Concurrently, we independently re-processed the F105W and F160W images from the 2009 and 2012 epochs to search for variable sources over the shorter, earlier time baseline. This setup enables variability searches across three epochs: the period from 2009 to 2012 is covered by the F105W and F160W filters, while the span from 2012 to 2023 is sampled by the F140W filter alone (see Table 4.1 for details).

4.1.2 *Photometry*

For a given filter and time baseline, we run `Source Extractor` (Bertin and Arnouts, 1996) on both epochs. We use an rms map based on the weight map produced by `astrodrizzle` for each filter, which is itself an inverse variance image based on the input exposures that contributed to each pixel. For each detection, `Source Extractor` outputs the geometric properties of the detection ellipse, specifically World Coordinate System (WCS) and pixel coordinates, and photometric properties of the detected object, namely the flux, the magnitude and their respective errors estimated accordingly.

Rather than using `Source Extractor` in the standard *single mode*, we opt for the *dual mode* approach. In single mode, both source extraction and photometry are performed on the same input image. However, dual mode allows for source detection and aperture determination using one image (the *detection image*), while photometric measurements are made on a second image (the *photometry image*). In contrast to H24, where we implement the dual mode approach using the single epochs themselves as detection images in both directions, we employ a common detection image for any given filter and epoch. Specifically, we use a combined stacked frame from multiple epochs for the unconvolved F105W, F140W, and F160W images, along with the original HLSP in F125W, all matched to the same WCS. This method provides a 'global' list of source coordinates and apertures that remain consistent across different runs of `Source Extractor`, ensuring one-to-one correspondence for all extracted photometric catalogues in any given filter.

Instead of conducting 'extended' galaxy photometry, such as using Kron-like or moment-centred apertures as done by `Source Extractor`, we aim to obtain photometry centred on the brightest, unresolved sources within each galaxy determined from the ultra-deep stack. To achieve this, we rely on the coordinates of the brightest pixels reported by `Source Extractor` and the background images it generates during its background subtraction process. At these coordinates, we perform aperture photometry within 4-pixel diameter ($0\farcs26$) apertures, applying local background subtraction to exclude local (non-compact) galaxy light.

We then assemble a photometric catalogue for each individual filter (epoch) by running `Source Extractor` in the dual mode. As stated above, the dual mode assures us that a possible photometric variability, if any, comes from the same region of the sky, once every frame have been co-aligned with respect to the HLSP. Hence, for each image, we obtain photometry of each galaxy nucleus, where AGN variability is expected. We consider that off centre variability would be likely due to SN and/or stellar cluster activity in the outskirts of the galaxy. However, such an effect could contaminate the high redshift ($z > 6$) population of variable candidates, being likely unresolved objects. Given the infinitesimal probability to observe a SN/stellar cluster in the footprint of the HUDF at high redshift ($\sim 10^{-4}$), we regard this effect as a negligible contamination. Moreover, detecting a SN at these epochs would itself be a remarkable discovery.

Finally, we compare these local aperture magnitudes in each image relative to the photometric uncertainty of each source. Additionally, the drizzling process causes artificially underestimated uncertainties due to the correlation of signal between adjacent pixels. We account for this effect by multiplying the uncertainties by a correction factor, assumed to be constant over the processed area (e.g., Casertano et al., 2000; Fruchter and Hook, 2002, see Fig. 4.1). It is worth underlining that our photometric measurements do not aim for accurate estimation of the absolute flux, rather to a comparison between two different epochs in terms of relative variation.

4.1.3 *Identifying Variable Sources*

Following H24, we identify variables in each matched pair of filters using two techniques: comparing the nuclear/central photometry of galaxies at different epochs and detecting residual sources in pair-subtracted images in any given filter. Regions near the image edges are excluded due to the dithering pattern causing excess noise and we focus solely on a central region of the HUDF covering $123'' \times 139''$.

4.1.3.1 *Photometric comparison*

The photometric comparison method follows a similar approach to that of O'Brien et al. (2024). We first correct for a systematic offset of approximately 0.01 magnitudes that is observed at all magnitudes, likely due to slightly imperfect zeropoints in images taken many years apart. Under the assumption that the vast majority of sources in the field will not vary, we calculate the standard deviation of the Δm distribution in 0.5 mag bins, where $\Delta m = m_1 - m_2$ is the magnitude difference. The magnitude labels always refer to the time ordering of the different epochs, being 1 the first visit and 2 the second one. We then propagate the uncertainty on the magnitude difference derived from our photometric catalogues in the following way:

$$\sigma_{\Delta_m} = \left(\frac{2.5}{\ln 10}\right)\sqrt{\left(\frac{\sigma_{F_1}}{F_1}\right)^2 + \left(\frac{\sigma_{F_2}}{F_2}\right)^2}, \tag{20}$$

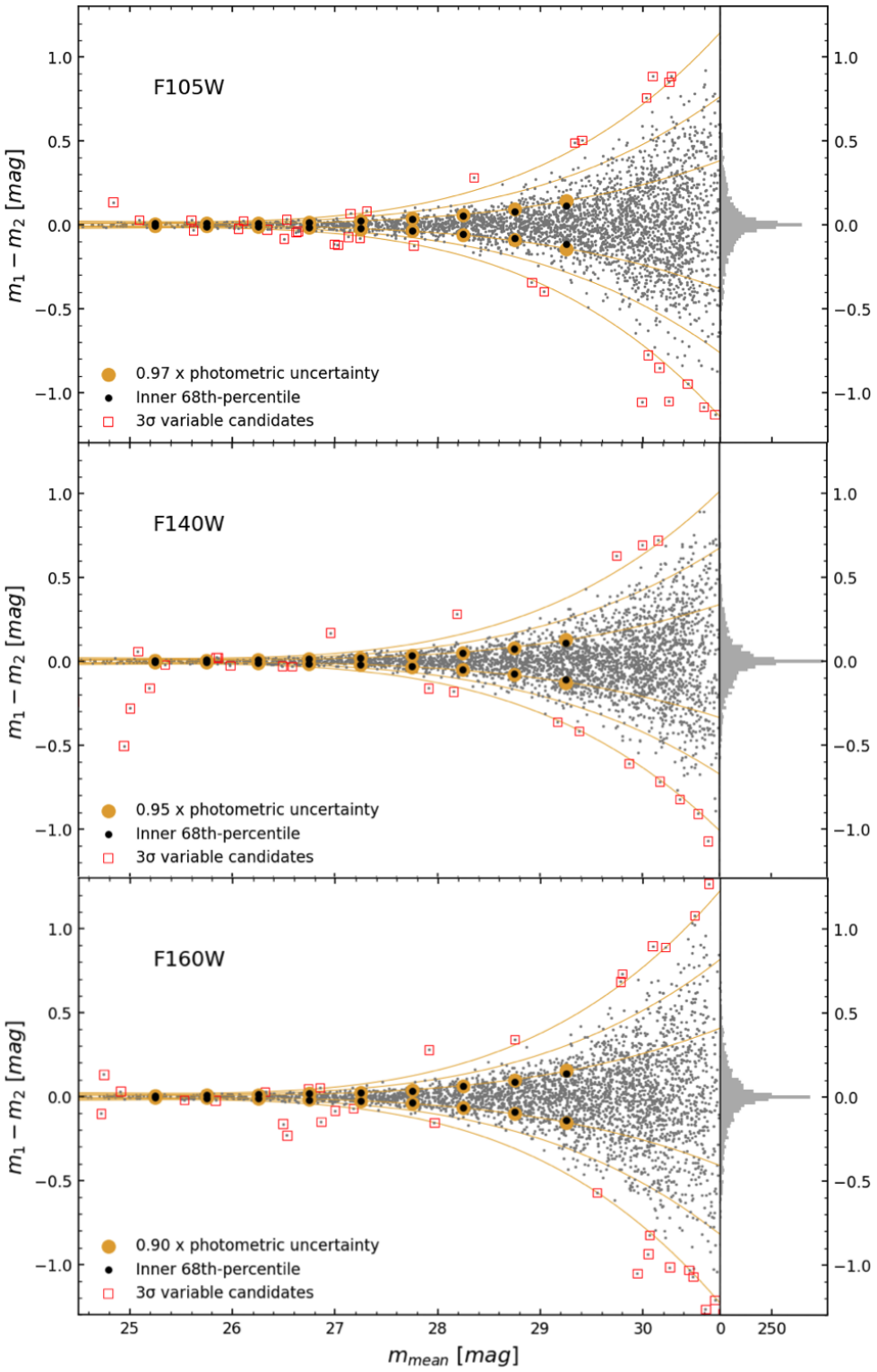


Figure 4.1: Calibrated photometric uncertainties to identify variable candidates. We plot the photometric variability, measured as the magnitude difference between the two epochs, as a function of the mean magnitude in F105W, F140W and F160W, as labelled. Every grey dot represents a source that has been measured in both observations. Black dots show the intrinsic $1\sigma_m$ scatter of the data measured in bins of 0.5 mag width. Golden circles depict the calibrated photometric 1 σ_m uncertainties in each bin (see text). The inner pair of yellow lines are polynomial fits to these calibrated uncertainties. The next sets of yellow lines show 2 σ_m and 3 σ_m uncertainties, scaled from the 1 σ_m fit. Red squares highlight sources that are estimated to be $\geqslant 3$ σ_m variables. The right-hand panel shows a histogram of the distribution of magnitude differences for all objects detected in the field.

where σ_F and F are the flux error and the flux as estimated by `Source Extractor`, respectively.

To calibrate our estimates of the uncertainties, we compare the observed standard deviations of the Δm distributions in magnitude bins of 0.5 mag width over the range from 25 to 29.5 mag with the estimated mean photometric uncertainties in each bin from Eq. (20). We then compute the total average scale factor that globally adjusts the uncertainty reported in Eq. (20) to match the observed Δm standard deviation. For the F105W, F140W, and F160W filters the resulting calibration scale factors are 0.97, 0.95, and 0.90, respectively. Next we fit a polynomial function to the calibrated estimates of the 1σ uncertainties in each magnitude bin. It is by a comparison of this estimated 1σ uncertainty as a function of magnitude that a given source's significance of variability is ultimately assessed in a two-epoch observation in a given filter. Fig. 4.1 reports the Δm of all the detected sources in the corresponding filter (grey points). The golden curves represent the inner $1\sigma_m$ range of the distribution at each magnitude and are polynomial fits to the golden points (and scaled by a factor of 2 and 3 σ_m). This lines determine the variability level of every object which we assign a value calculated as the ratio between Δ_m and the associated 1 σ_m variation at that specific magnitude.

We consider three thresholds of significance for variability: 2 σ_m; 2.5 σ_m; and 3 σ_m. For a source to be classified as a candidate variable, it must have a $\Delta m/\sigma_m \geqslant 2$ in at least one filter. For a visual interpretation, we refer to the red squares in Fig. 4.1, where we use $\Delta m/\sigma_m \geqslant 3$ in each filter for which we apply the same photometric error calibration and selection procedure of the variable candidate sources.

Our verification process, which includes examining the number of sources in the images and the distribution of $\Delta m/\sigma_m$, shows that this criterion results in a contamination level from non-varying sources down to magnitudes as faint as ~ 30 that we need to account for. In order to do that, we calculated the number of false detections expected at a given significance level assuming a Gaussian statistics. We discuss this in detail in Section 4.3. Fig. 4.2 reports the histograms of $\Delta m/\sigma_m$ for each filter/epoch. The red dashed lines are Gaussian fits to the blue histogram.

4.1.4 *Difference Imaging Variables*

As a complement to variability identified via aperture photometry, we also generate pair-subtracted images for each filter. From these images we may identify variable sources that were not detected via aperture photometry. We are also able to assess the morphology of photometric variables. In particular, single AGNs are expected to appear as compact, unresolved variable sources. For this process, each image is first smoothed with a small Gaussian kernel of 0.5 pixels. Then images taken in the same filters were directly subtracted, resulting in six "difference images". We then applied

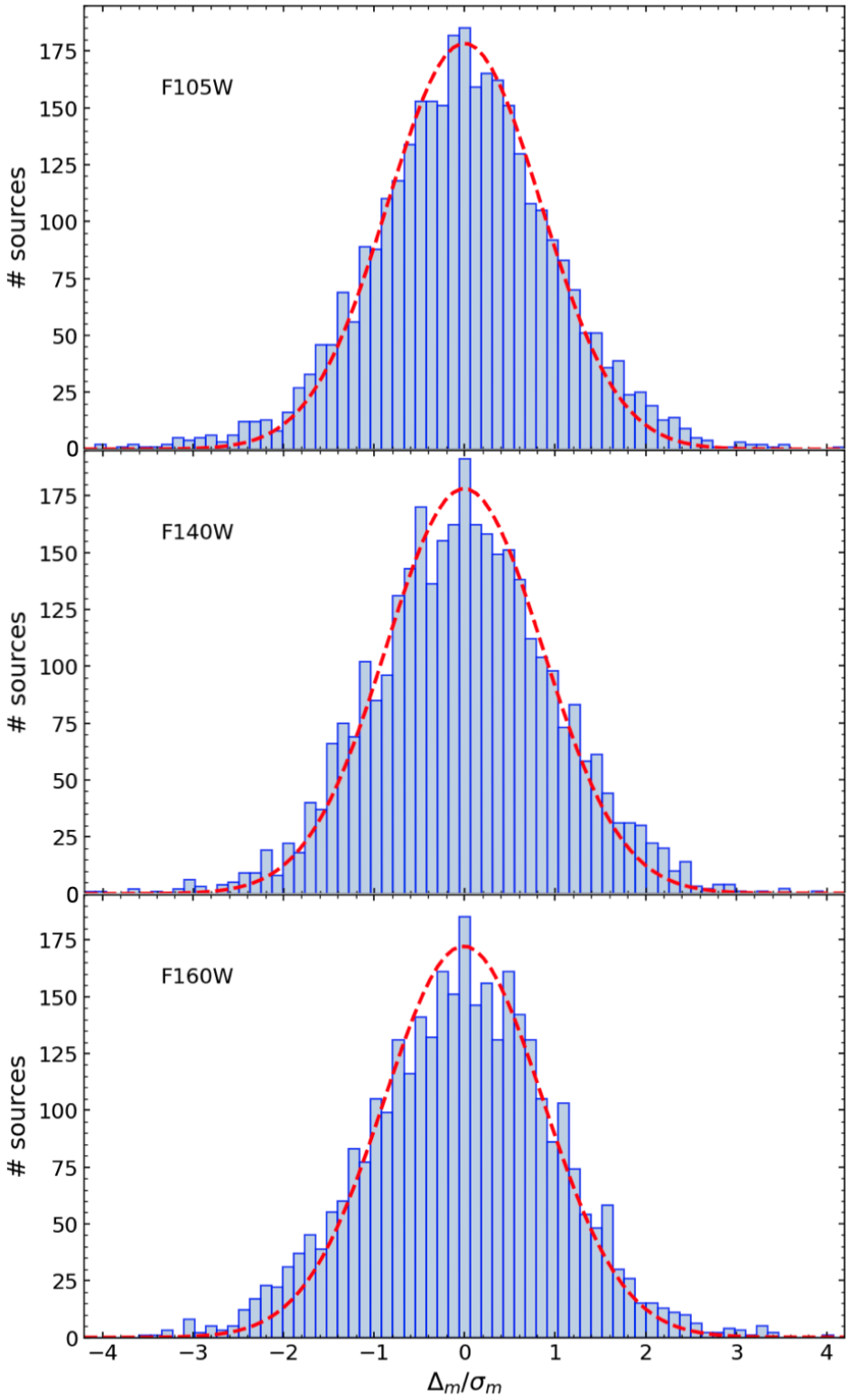


Figure 4.2: Histogram of $\Delta m/\sigma_m$ for the total sample of detected sources in F105W, F140W and F160W as labelled, in blue. The red dashed line shows the Gaussian fit to the histogram values.

Table 4.2: Number of matched AGNs against Lyu et al. (2022) across all the three filters.

Significance Level	Fraction of AGNs	Corr. factor
$> 2\ \sigma$	13/31	2.38
$> 2.5\ \sigma$	9/31	3.44
$> 3\ \sigma$	6/31	5.17

`Source Extractor` to these difference images, adjusting the effective gains to account for increased sky noise due to the subtraction process, which mimics a shallower integration time. Each detected source is manually inspected, and spurious artefacts are removed. Here we retain sources with a variability significance of 5 σ or higher.

4.1.5 *Redshift determination and known AGNs*

We cross-correlate each of these sources with positions from known redshift catalogues, including spectroscopic data from VLT/MUSE (Bacon et al., 2023), the JADES GTO program for spectroscopic redshifts using NIRSpec (Bunker et al., 2023), photometric redshifts from NIRCam+HST (Rieke et al., 2023), and the Ultraviolet UDF photometric catalogue (Rafelski et al., 2015). Fig. 4.3 show the redshift distribution derived from the cross-match with the cited catalogues. Note that the depicted redshift is the spectroscopic one when available, and the photometric one otherwise.

Additionally, we cross-correlate with known AGNs in the GOODS-S field, as compiled by Lyu et al. (2022). For this latter we apply the same mask to exclude the image edges, as described at the beginning of this Section. In Table 4.2 we report the number of matched AGNs in the field against the more complete sample of Lyu et al. (2022) as a function of the significance threshold for variability detection (in units of σ_m). We consider that such fractions can be used to account for variability incompleteness, as detailed in the next Section.

4.2 RESULTS

4.2.1 *Detected Variable Sources*

Photometric variability ($m_1 - m_2$) of sources in the HUDF in the three filters as a function of mean magnitude is shown in Fig. 4.1, along with estimates of the 1 σ, 2 σ and 3 σ uncertainties. We can identify a total of 76 sources with photometric variability significance larger than 3 σ_m in at least one or more filters. We also consider a significance level of 2.5 σ_m, for which we find 149 variable candidates. Lowering the significance threshold at 2 σ_m, we retain a catalogue of 443 objects for which either photometric or spectroscopic redshift are available from the cross-match against published

data as in Section 4.1.5. H24 confirm the photometric variability as a method in finding convincing AGN. In fact, among the selected sources we detect two confirmed AGNs at $z = 2 - 3.2$, which illustrate the robustness of this technique at redshifts where they can be independently confirmed. Among the best studied regions of the extragalactic sky over the past 2 decades, the HUDF has been targeted in all different kinds of survey across the electromagnetic spectrum: independent AGN confirmation methods span from deep optical and NIR spectroscopy to VLT/MUSE and HST/grism, the deepest X-ray imaging from Chandra, radio continuum imaging, and deep Spitzer imaging to measure MIR colours (see the discussion in Section 4.4). H24 also present results for transient events, most likely SN, and shows how they can act as interlopers at lower redshift ($z \sim 2 - 3$). Nonetheless, Super Luminous Supernovas (SLSNs) are more of a concern as they can reach above our limits in UV brightness (Yan et al., 2018). However, using rate calculations from Moriya et al. (2019), we regard the SLSN scenario very rare as the rate per year of such an event in the HUDF footprint would be $\sim 7 \times 10^{-4}$. We present here the revised results, with particular focus upon the statistical and spatial properties of candidate variable objects up to redshift $z \lesssim 9$. We first include the redshift distribution of our sources. We then present the estimated number density of our confident variable sources. The study of the morphology and photometric inferences of the sources considered here is beyond the scope of this work and will be further investigated in a future work.

4.2.2 *Redshift distribution*

Fig. 4.3 shows the redshift histograms of our sample of selected variable sources. Different colours depict different samples according to the significance level (> 2, 2.5 and 3 σ_m) as detected in at least one filter. We see that the vast majority of candidates resides at $z \lesssim 4$ for all cases of significance threshold, even thought the fraction of sources that we retain at $z \gtrsim 4$ is an increasing function of the required σ_m level. This suggests that we can asses the *true* variability of the high redshift objects with more robustness. We also note that with respect to H24 we are able to recover (at $> 2\,\sigma_m$), on top of the 3 candidates at $z > 6$, a few extra tens of sources extending up to a maximum (photometric) redshift of 8.74, with a total of 4 objects at $z > 8$, 7 within $7 < z < 8$, and 16 in the range $6 < z < 7$. These numbers add up to a total of 27 sources at $z > 6$. Among them, we can retrieve a spectroscopic redshift for 3 candidates. Table 4.3 reports the properties of this sample of high-z variable candidates.

For a visual representation of the field, Fig. 4.4 shows the spatial distribution of the variable sources superimposed on the HUDF footprint in the sky. Different symbols indicate the different ranges of the maximum, across the 3 filters, significance level in units of σ_m as detailed in the legend, colour coded according to the estimated redshift from Fig. 4.3.

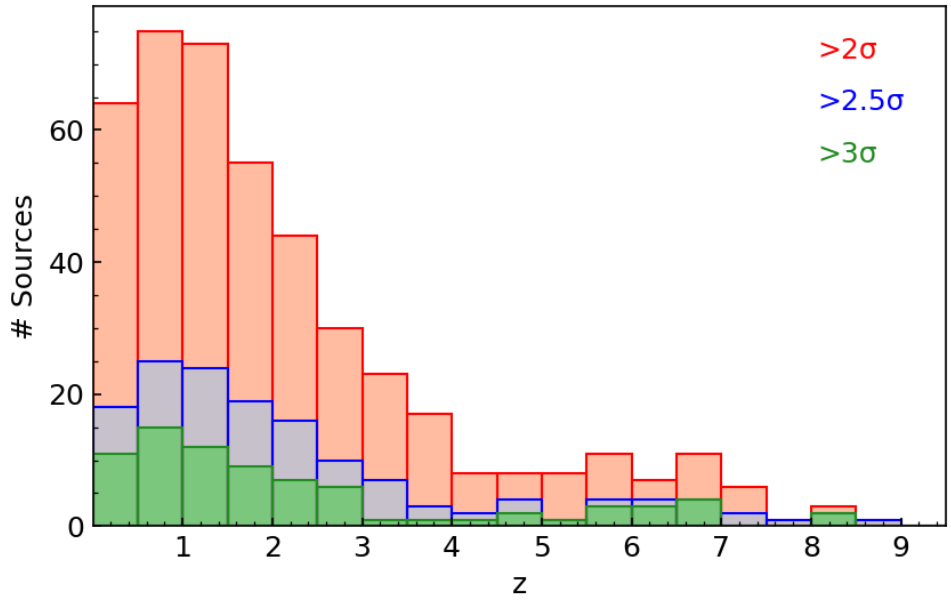


Figure 4.3: Redshift distribution of the variable candidate sources. Different colours refer to the specific significance thresholds for variability detection in units of σ_m. At redshift between $6 < z < 7$ we also include the 2 extra sources detected in H24 which are not selected as variables in this work.

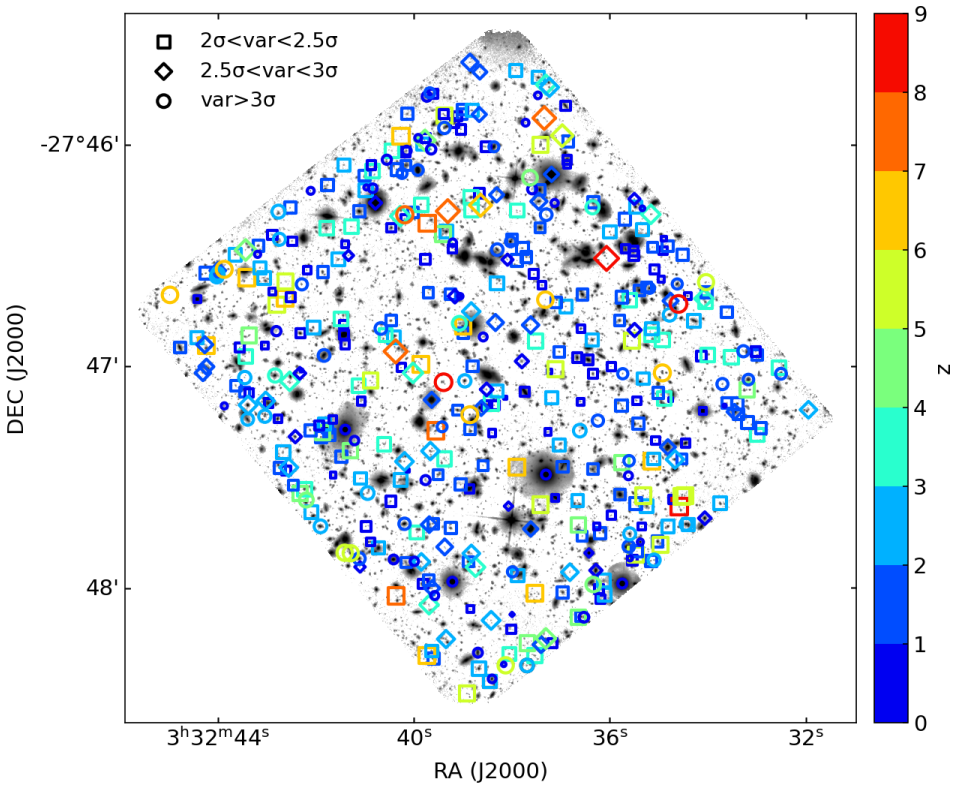


Figure 4.4: Sky map of variable candidates in the HUDF. The estimated redshift is shown by the colour of the symbol with reference to the colour-bar. Different symbols refer to different significance levels as indicated in the legend.

ID	RA	Dec	$m_{1,F105W}$	$m_{2,F105W}$	$m_{1,F160W}$	$m_{2,F160W}$	$m_{1,F140W}$	$m_{2,F140W}$	Redshift	σ_{F105W}	σ_{F160W}	σ_{F140W}
	J2000/Gaia	J2000/Gaia	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)				
3026	53.15032	-27.77523	29.74 ± 0.40	30.15 ± 0.35	29.94 ± 0.36	30.01 ± 0.56	29.69 ± 0.26	30.23 ± 0.44	$8.74^{+0.62}_{-0.58}$	-1.60	-0.30	**-2.63**
785	53.14414	-27.79393	30.18 ± 0.55	30.38 ± 0.41	30.16 ± 0.47	29.93 ± 0.51	30.17 ± 0.35	30.79 ± 0.49	$8.16^{+0.03}_{-0.10}$	-0.61	0.83	**-2.21**
2617	53.14410	-27.77863	29.55 ± 0.32	29.93 ± 0.28	30.23 ± 0.45	29.85 ± 0.55	30.35 ± 0.56	29.65 ± 0.26	$8.09^{+6.93}_{-0.14}$	-1.70	1.42	**3.28**
1829	53.16414	-27.78452	29.57 ± 0.50	29.10 ± 0.13	29.09 ± 0.15	28.97 ± 0.20	28.95 ± 0.13	29.02 ± 0.15	$8.02^{0.23}_{-0.11}$	**3.25**	0.95	-0.77
2165	53.16824	-27.78230	30.00 ± 0.50	30.44 ± 0.43	30.53 ± 0.59	30.51 ± 0.86	30.74 ± 0.66	30.05 ± 0.39	$7.53^{+0.46}_{-0.29}$	-0.53	-0.01	**2.52**
3351	53.16382	-27.77174	30.37 ± 0.62	30.49 ± 0.39	30.19 ± 0.39	31.19 ± 1.15	30.20 ± 0.39	30.20 ± 0.41	$7.27^{+0.11}_{-0.75}$	-0.35	**-2.60**	-0.02
3326	53.16745	-27.77203	30.08 ± 0.49	30.81 ± 0.58	30.86 ± 0.68	31.18 ± 1.65	30.88 ± 0.72	30.16 ± 0.42	7.27[b]	**-2.27**	-0.68	**2.41**
3270	53.16555	-27.77267	29.57 ± 0.34	30.02 ± 0.29	29.60 ± 0.26	30.14 ± 0.70	29.67 ± 0.28	29.76 ± 0.30	$7.17^{+0.09}_{-0.07}$	-1.98	**-2.31**	-0.58
294	53.16821	-27.80069	30.18 ± 0.62	30.63 ± 0.53	31.05 ± 0.99	30.25 ± 0.74	30.64 ± 0.75	30.14 ± 0.46	$7.05^{+0.03}_{-0.11}$	-1.39	**2.04**	1.83
1297	53.16482	-27.78826	29.18 ± 0.21	29.19 ± 0.12	28.71 ± 0.09	28.52 ± 0.12	28.70 ± 0.09	28.61 ± 0.09	7.228[c]	0.13	**2.16**	1.25
3790	53.15560	-27.76475	30.74 ± 0.81	29.99 ± 0.28	29.87 ± 0.32	29.77 ± 0.44	30.02 ± 0.31	29.72 ± 0.27	$7.00^{+5.83}_{-2.97}$	**2.53**	0.42	1.52
2875	53.18091	-27.77676	30.34 ± 0.67	29.84 ± 0.31	29.63 ± 0.25	29.37 ± 0.36	29.72 ± 0.30	29.75 ± 0.30	$6.91^{+0.10}_{-0.03}$	**2.11**	1.41	-0.18
1906	53.14556	-27.78379	28.17 ± 0.08	28.22 ± 0.05	28.14 ± 0.05	28.20 ± 0.09	28.06 ± 0.05	28.25 ± 0.07	6.87[c]	-0.56	-0.81	**-3.26**
1511[a]	53.16193	-27.78706	28.48 ± 0.16	28.22 ± 0.06	28.05 ± 0.07	27.77 ± 0.08	27.83 ± 0.05	27.99 ± 0.06	$6.74^{+0.04}_{-0.07}$	**4.18**	**5.76**	**-4.07**
2919	53.18284	-27.77612	28.56 ± 0.13	28.60 ± 0.09	28.57 ± 0.10	28.64 ± 0.16	28.33 ± 0.08	28.05 ± 0.06	$6.74^{+0.06}_{-0.04}$	-0.23	-0.64	**5.48**
2709	53.17763	-27.77828	30.98 ± 1.26	30.74 ± 0.62	31.17 ± 1.19	30.30 ± 0.77	30.53 ± 0.60	31.7 ± 1.8	$6.72^{+0.25}_{-0.27}$	0.60	**2.09**	**-2.85**
3724	53.16780	-27.76613	30.49 ± 1.21	30.71 ± 0.62	30.26 ± 0.50	30.51 ± 0.89	30.95 ± 0.85	30.27 ± 0.47	$6.72^{+5.58}_{-0.08}$	-0.59	-0.83	**2.20**
2047	53.16616	-27.78321	30.65 ± 0.89	30.59 ± 0.48	31.02 ± 0.94	30.16 ± 0.64	30.47 ± 0.60	30.74 ± 0.69	$6.67^{+0.30}_{-0.14}$	0.23	**2.24**	-0.90
1123	53.14649	-27.79053	30.22 ± 0.48	30.51 ± 0.51	30.98 ± 0.77	32.4 ± 2.0	31.05 ± 0.84	30.33 ± 0.47	$6.58^{+0.63}_{-1.31}$	-0.89	**-2.08**	**2.17**
1058	53.15796	-27.79101	30.03 ± 0.65	30.73 ± 0.40	29.96 ± 0.31	30.28 ± 0.71	30.06 ± 0.28	30.14 ± 0.39	$6.56^{+6.37}_{-0.001}$	**-2.27**	-1.18	0.38
93	53.16560	-27.80507	30.60 ± 1.08	30.38 ± 0.62	30.75 ± 0.94	30.68 ± 1.54	30.71 ± 1.07	30.09 ± 0.64	$6.53^{+4.83}_{-0.43}$	0.72	0.12	**2.26**
305	53.15643	-27.80051	30.02 ± 0.49	29.96 ± 0.30	30.51 ± 0.65	29.87 ± 0.48	30.23 ± 0.45	30.17 ± 0.42	$6.26^{+0.02}_{-0.19}$	0.43	**2.13**	0.27
2215	53.18433	-27.78184	30.18 ± 0.53	29.99 ± 0.28	30.76 ± 0.65	30.18 ± 0.75	30.67 ± 0.60	30.11 ± 0.46	$6.21^{+0.16}_{-0.38}$	0.89	1.62	**2.06**
3391	53.16105	-27.77130	29.92 ± 0.41	29.76 ± 0.21	29.78 ± 0.31	30.50 ± 0.67	29.89 ± 0.28	30.00 ± 0.32	$6.21^{+0.05}_{-0.31}$	0.87	**-2.55**	-0.52
2286	53.16257	-27.78037	27.10 ± 0.03	27.16 ± 0.02	27.51 ± 0.03	27.48 ± 0.04	27.51 ± 0.03	27.48 ± 0.03	$6.11^{+0.02}_{-0.01}$	**-2.03**	1.01	1.54
2736	53.18744	-27.77804	29.44 ± 0.47	30.53 ± 0.65	31.04 ± 1.25	31.72 ± 5.86	30.93 ± 0.87	30.37 ± 0.59	$6.08^{+0.03}_{-0.03}$	**-4.45**	-1.17	1.74
2680	53.15550	-27.77841	29.97 ± 0.46	30.33 ± 0.39	31.43 ± 1.45	30.33 ± 0.74	30.09 ± 0.38	30.99 ± 0.89	$6.01^{+0.06}_{-0.05}$	-1.23	**2.50**	**-3.06**

Table 4.3: General properties of the $z > 6$ variable sources at $> 2\ \sigma_m$ significance. [a]This source correspond to source 1052123 in H24. [b]Spectroscopic redshift from NIRspec (Bunker et al., 2023). [c]Spectroscopic redshift from FRESCO (Oesch et al., 2023).

4.2.3 *Comparison with previous analysis*

Here we discuss the differences between the photometric analysis conducted in H24 and the motivations that led us to present the updated version in this work. Briefly, while using the same `Source Extractor` configuration, in H24 we run the dual mode, using one of the two epochs as the detection image and performing photometric measurements in both epochs. This process was carried out in both time directions, depending on which epoch was used as the detection frame. Despite the straightforward approach, the final outcome has several disadvantages. i) Since the detection image is not the deepest available, `Source Extractor` places the apertures based on the peak pixel in the specific epoch. This introduces susceptibility to the peculiar variability effects of the single epoch, which generally may not be reflected in the other epoch due to off-centre variable emission. ii) Slight uncorrected misalignment between the two epochs, taken several years apart, can shift the coordinate systems, thereby distorting the *detected* variability. iii) The combination of these factors prevents us from considering sources that become fainter over time between the detection and measurement images (see H24 for details). As a result, when calculating the σ_m to evaluate the significance of the sources, one has to rely on a one-sided distribution, which may under-represent the total sample and complicate the calculation of calibrated uncertainties (see Section 4.1.3).

These considerations led us to revise the procedure and adopt a *global* approach, using an ultra-deep stack as the detection frame across all epochs/filters (see previous Section 4.1.2). It is worth noting that the different methodologies yield some variation in the reported list of variable sources. In particular, H24 identifies three remarkable variable sources at $6 < z < 7$. Of these three objects, two are recovered in this study, including one selected via difference imaging, while the third source is missed in the current approach due to differences in aperture placement affecting the variability assessment in H24 (see Section 4.1.2).

4.3 VARIABILITY AS A METHOD TO FIND SMBHS

Through photometric monitoring conducted in three epochs — 2008/2009, 2012, and 2023 — we discover several variable sources in the Hubble Ultra Deep Field. We advocate that variability searches are a highly effective and comprehensive tool for identifying AGN in deep imaging surveys. While only a portion of AGN can be detected through their variability, the significant advantage of this approach lies in its ability to survey the entire field with no prior selection, offering a fully multiplexed view. Unlike other methods such as radio or X-ray diagnostics, which are limited by detection thresholds, variability searches ensure that any object captured in imaging data can be examined for activity.

Historically, high-redshift luminous quasars have been known for some time (Fan et al., 2006; Mortlock et al., 2011; Bañados et al., 2018), though they

are typically much brighter, by four magnitudes or more, than the sources we detect. More recently, there have been reports of AGN at even higher redshifts with luminosities closer to $L^\star$ (Maiolino et al., 2023; Harikane et al., 2023; Larson et al., 2023). However, our survey, which is restricted to the small volume of the HUDF, has already identified several likely AGN in the $6 < z < 7$ redshift range, as also detailed in H24. This suggests a higher co-moving number density of AGN at these epochs than what has been previously reported. In the following we first revise the results in terms of estimated SMBH number density from our previous analysis done in H24. Secondly we described how we extend the approach to the full sample of variable candidate sources.

In H24, we estimate for the first time the comoving number density of SMBHs via a variability study. Considering the survey probes a volume of approximately 10,500 comoving Mpc^3 within the redshift range of 6 to 7, we calculate a number density estimate of $n_{SMBH} \sim 2.9 \times 10^{-4}$ cMpc^{-3}. It is important to note that only a subset of AGN within a given luminosity range is expected to be detectable via variability searches. In fact, as mentioned in Section 4.1, Lyu et al. (2022) published a catalogue of confirmed AGN in the GOODS-S field, which fully encompasses the HUDF region targeted in our study. Of the 31 AGN from their catalogue that fall within our HUDF footprint, we detect 8 objects through photometric variability, with redshifts ranging from 0.6 to 3.2. Thus, assuming the ratio 8/31 can be used to roughly estimate a variability completeness correction factor, this boosts the estimated number of variables by the correspondent factor of 3.85. We stress here that Lyu et al. (2022) identified these AGN through various methods, including mid-IR colours (4 AGN), X-ray luminosities (7 AGN), radio loudness (7 AGN), optical spectroscopy (1 AGN), and, in some cases, variability. Several AGNs were identified by more than one diagnostic technique.

Crucially, none of the selected candidates have been previously reported through photometric variability, including those found in earlier works using HST imaging by Lyu et al. (2022) and the 45 sources reported by Cohen et al. (2006). It is reasonable to expect discrepancies between NIR long-baseline studies considered here (and in H24) and earlier surveys. Notably, the imaging data from these previous studies are shallower, taken at shorter wavelengths, and covered shorter time baselines compared to our re-imaging of the HUDF in the near-infrared. Additionally, none of the AGN in our sample with redshifts greater than 6 are present in the Lyu et al. (2022) catalogue, likely because they are too faint to be identified using traditional techniques given the current observational limits.

It is important to note that our survey is sensitive only to AGN with a minimum luminosity of $M_{UV} \simeq -18.6$, meaning this estimate still represents a lower bound on the true value of n_{SMBH}. In contrast, Harikane et al. (2023) detected AGN via broad emission lines and estimated n_{SMBH} from $z \sim 4$ to 7 by extrapolating their assumed luminosity function down to $M_{UV} \simeq -17$. To ensure a fair comparison, in H24 we apply the same method to our data,

Table 4.4: UV luminosity function correction as function of redshift. The second column indicates the correction factors using different compilations of fits to the UVLF at different redshifts. The values derived from the case adopting a conservative Schechter UVLF are reported in the last column (see text).

Redshift	Corr. factor UVLF compilation	Corr. factor UVLF conservative
$z \sim 0.3$	0.05	0.09
$z \sim 0.8$	0.74	0.15
$z \sim 1.5$	0.15	0.29
$z \sim 2.5$	0.41	0.50
$z \sim 3.5$	0.72	0.78
$z \sim 4.5$	1.37	1.19
$z \sim 5.5$	2.8	1.81
$z \sim 6.5$	5.8	2.8
$z \sim 7.5$	12.9	4.5
$z \sim 8.5$	29.3	7.4

estimating a luminosity function correction factor (assuming $M^\star = -20$ and $\alpha = -2$) of approximately 7. This correction derived in H24 increases the estimate of n_{SMBH} to $\sim 7.7 \times 10^{-3}$ cMpc^{-3}.

In this work we adopt a similar strategy and extend this approach to the whole sample of variable candidate sources. Additionally, from the total sample of variable candidates, as shown in the histogram in Fig.4.3, in each redshift bin we first correct by subtracting the expected number of false detections (for every significance level, 2, 2.5 or 3 σ_m) by assuming a Gaussian statistics on the entire number of detected sources in the field. This enables us to explore the trend of n_{SMBH} as a function of both redshift and the variability significance threshold. Concurrently, we are able to assess the robustness of our methodology in relation to the criteria used to classify source variability. As noted in Section 4.1.5, Table 4.2 presents the estimates for the variability completeness correction factor at three different significance levels. As discussed earlier, this serves as the initial correction to account for AGNs which do not show variability in the considered time window between the two epochs.

To obtain a more accurate measurement of the n_{SMBH}, it is essential to estimate a correction for the luminosity function. Given the substantial differences in redshift for the sources in our sample, such a correction cannot be considered constant. The UVLF evolves over time, peaking at cosmic noon ($z \sim 2$) and steadily decreasing in both normalization and $L^\star$ as redshift increases. Additionally, the detection limit of the survey, expressed in absolute magnitude M_{lim}, needs to be updated as a function of redshift. As a result, the relative contribution to the total volume density of galaxies

within the magnitude range between M_{lim} and -17 changes over time. We address this by calculating the luminosity function correction based on a compilation of UVLF estimates. Specifically, we integrate the double power-law fits detailed in Finkelstein and Bagley (2022) for the redshift range $z \sim 3$ to 9. At lower redshift, we integrate UVLF fits from GALEX and UVCANDELS survey programs (Arnouts et al., 2005; Sun et al., 2024). At $z \lesssim 4$ the correction to apply turns around as the detection limit overcomes the UVLF lower bound ($M_{UV} \sim -17$). We refer to this case as the *compilation* UVLF corrections. For continuity with previous works, we also consider a simple case in which we model the UVLF with a Schechter function (Schechter, 1976) similar to H24 by assuming $M_* = -20$ and $\alpha = -1.2$. These parameter are estimated by requiring what it would be the UVLF in order to obtain a roughly constant n_{SMBH} as a function of the redshift. We will refer to this latter as the *conservative* UVLF correction case. In Table 4.4 we report the UVLF correction factors.

Fig. 4.5 presents our results for the estimated comoving number density of SMBHs, n_{SMBH}, as a function of redshift calculated in bins (delimited by the gray dotted vertical lines) from $z \gtrsim 0$ to 9 for the *compilation* case. It also shows previous estimates from Harikane et al. (2023) and from H24 in black and red diamonds, respectively. Error bars are calculated as the sum in quadrature of the uncertainties due to Poisson errors on both the raw number counts (corrected for the false detections) and on the variability completeness correction fraction of matched AGNs (Table 4.2) plus the uncertainty on the luminosity function correction (Table 4.4). For this latter, we adopt a conservative relative error of 10% given the uncertainties in the integration carried mainly by the detection limit magnitude and the lower bound of -17 as reported by Harikane et al. (2023). Our estimates are consistent with our previous analysis in H24 even though slightly higher in the $z \sim 6-7$ bin and systematically higher than Harikane et al. (2023) at redshift larger than 5, while similar values are obtained in the interval at $z \sim 4-5$. At a given redshift, our estimates from different significance levels show a fairly stable trend, suggesting that the methodology is not a strong function of this choice. In the $z \sim 3-4$ redshift bin we notice that our number densities indicate a large scatter as a function of the significance threshold, due to the dearth of 3 σ_m variable candidates in this redshift range with respect to the 2 σ_m counts (Fig. 4.3). Overall n_{SMBH} shows a slightly decreasing trend up to redshift about 4-5, where $n_{SMBH} \sim 10^{-3}$ cMpc^{-3}, and then it raises by $\gtrsim 1$ order of magnitude at the highest redshift bins up to $\sim 3 \times 10^{-2}$ cMpc^{-3}. Solid lines show predictions from the Pop III.1 model as a function of different isolation distances (Singh et al., 2023). Dashed lines illustrate predictions from the HMT model in green and from the Direct Collapse model in pink. The black square at $z = 0$ is calculated by Banik et al. (2019) by assuming that every galaxy at luminosities greater than 0.33 L* is hosting an SMBH, with uncertainties ranging from 0.1 L* to L*. Since the occupation fraction of galaxies has been shown not to be a step function (e.g., Cammelli et al., 2024), these values should be regarded as a

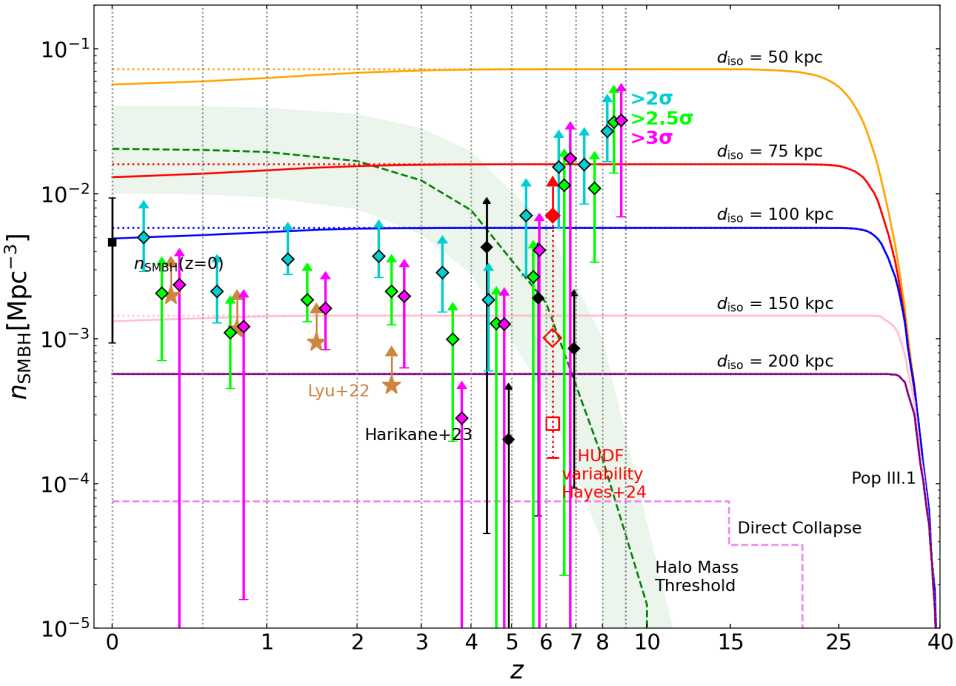


Figure 4.5: Redshift evolution of SMBH abundance. Comoving number density of SMBHs, n_{SMBH}, is plotted versus redshift, z. Estimated counts of candidate AGNs in the correspondent volume are represented with the filled diamonds with black edges, for the 2 (turquoise), 2.5 (lime) and 3 σ_m (magenta) significance thresholds, in different redshift bins between $z \sim 0-9$. In each bin, we subtract from the row counts the expected number of false detections according to a Gaussian statistics for the correspondent significance levels. These values are corrected first for variability incompleteness by multiplying for the correction fraction detailed in Table 4.2, and secondly for luminosity function incompleteness (see Table 4.4, second column). For this latter, we collect a compilation of fits to the UVLF at several redshifts (see text, *compilation* case). At $z = 0$ we depict the estimated number density of SMBHs from local galaxies (Banik et al., 2019, black square). The red diamond shows the observational constraint derived by Hayes et al. (2024). Broad emission line sources at $z = 4-7$ (Harikane et al., 2023) are shown with the black diamonds. For the $z \sim 4-7$ sources, the fraction of non-active SMBHs remains highly uncertain, and we leave all points as lower limits. Brown stars indicate the number densities from the known AGN sample of Lyu et al. (2022). Since these Various Pop III.1 formation models (Banik et al., 2019; Singh et al., 2023) with isolation distance parameters, d_{iso}, from 50 to 200 kpc (proper distance) are shown by the coloured solid lines. At low redshifts these decrease compared to the maximum value attained (dotted lines) account for merger events. The green dashed line shows the SMBH seeding assumed in Vogelsberger et al. (2014) based on a Halo Mass Threshold (HMT) above a mass of 7.1×10^{10} $M_\odot$ (being the shaded region a factor of two variation in this mass scale). Models of SMBH formation via *Direct Collapse* from Chon et al. (2016) are depicted by the pink dotted line.

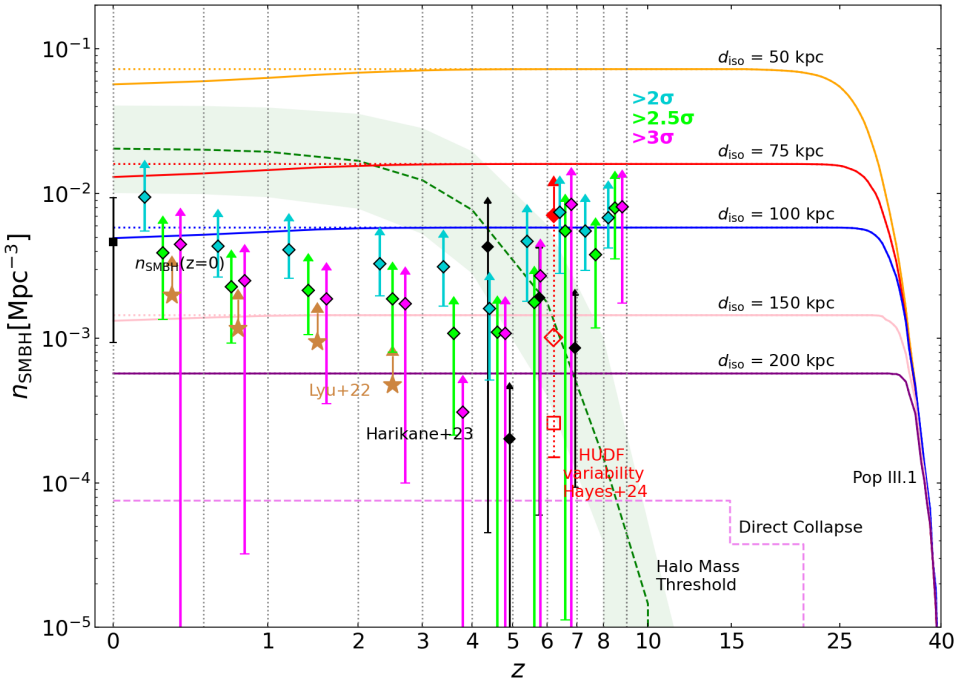


Figure 4.6: Redshift evolution of SMBH abundances as in Fig. 4.5, but with luminosity function corrections from Table 4.4, third column, for the *conservative* case of a Schechter LF (see text).

lower bound to constrain the population of SMBHs in the local Universe. Seeding models that do not reproduce this minimum values at $z \sim 0$ will not be able to explain the origin of the whole population of SMBHs, although a combination of different seeding schemes still could.

In Fig. 4.6 we show the n_{SMBH} estimates as in Fig. 4.5, but for the simple case of UVLF corrections discussed above. In this scenario, our estimates denote a similar trend with respect to the compilation case, but with smaller deviations from a constant behaviour. The value of n_{SMBH} spans from $\sim 10^{-3}$ cMpc^{-3} at $z \sim 4-5$ up to almost 10^{-2} cMpc^{-3} at both $z \gtrsim 0$ and $z \sim 6-9$. We stress that in both cases, UVLF corrections are highly uncertain as they are poorly measured at magnitudes as faint as $M_{UV} \sim -17$, especially at high redshift. Their impact can dramatically affect our final estimates of n_{SMBH}.

We also emphasize that since not all AGN are detected down to $M_{UV} \simeq -17$, this estimate should still be treated as a lower limit. At these redshifts, n_{SMBH} can only be empirically estimated by counting AGN, which includes only those SMBHs that are in an active, rapidly accreting state. To determine the *true* comoving density, an additional correction would be required for the fraction of SMBHs that are currently passive. However, we present all density values as lower limits due to the lack of observational data to estimate this correction at these epochs.

4.4 DISCUSSION & CONCLUSIONS

In this work we provide estimates of the n_{SMBH} as a function of redshift from the local Universe out to $z \sim 9$. We achieve this by comparing the photometric variability in different epochs taken from HST observations of the HUDF. Thanks to archival data, such a methodology has been applied to different WFC-3 filters: in F105W and F160W we exploit previous survey programs from Illingworth (2009) and Ellis et al. (2013), while in F140W we utilize the more recent epoch from Hayes et al. (2024). We argue that photometric variability can be used as a relatively robust tool to seek AGNs, especially at high redshift, as it does not require any specific conditions but being visible in imaging data. Other, more common, diagnostics for AGN are instead limited by several observational challenges, for instance in terms of spectral resolution, signal-to-noise ratio or detection limits.

Our estimated measurements can impact and constraint the current knowledge of SMBH formation mechanisms and their implementation in simulations. For instance, the resulting abundance of the HMT scheme in terms of SMBH number density as a function of redshift is depicted by the green line in Fig. 4.5. By redshift $z \simeq 6.5$, this model yields an SMBH number density of approximately $n_{SMBH} \sim 10^{-3}$ cMpc^{-3}, with the number increasing rapidly as the universe evolves. At redshift $z = 0$, the SMBH number density reaches around $n_{SMBH} \simeq 2 \times 10^{-2}$ cMpc^{-3}, which is close to observational estimates of $n_{SMBH}(z = 0) \sim 10^{-3} - 10^{-2}$ cMpc^{-3} (Vika et al., 2009; Banik et al., 2019). However, if a substantial population of SMBHs have been ejected from their host galaxies or have extremely low luminosities, these estimates may underestimate the true $n_{SMBH}(z = 0)$.

An alternative formation scenario is the DCBH model, where SMBHs originate from metal-free, UV-irradiated, and atomically cooled halos, which are relatively massive (see Chapter 1). This mechanism has been proposed to explain the formation of large SMBH seeds in the early universe (e.g., Bromm and Loeb, 2003). Nevertheless, simulations of this process have been struggling to get a sufficient number of SMBHs in place to match the observed population at $z = 0$. Among others, Chon et al. (2016) found a density of $n_{SMBH} \sim 10^{-4}$ cMpc^{-3}, while a total SMBH density ranging from $n_{SMBH} \sim 10^{-7} - 10^{-6}$ cMpc^{-3} was derived by Wise et al. (2019).

Predictions for the evolution of the comoving number density of SMBHs originating from Pop III.1 seeds have been made by Banik et al. (2019) and Singh et al. (2023). A key parameter in the Pop III.1 model is the *isolation distance* d_{iso}, which refers to the minimum distance a Pop III.1 source must be from already-formed sources to remain undisturbed. This distance is largely determined by the impact of radiative feedback. To account for the observed population of low-redshift SMBHs, with a density of $n_{SMBH}(z = 0) \sim 5 \times 10^{-3}$ cMpc^{-3}, the isolation distance must be approximately $d_{iso} \simeq 100$ kpc (proper distance), or about 3 cMpc at the typical formation redshift of $z \sim 30$. The model predicts that the number density of SMBHs, n_{SMBH},

remains roughly constant from redshift $z \sim 25$ to $z = 0$, with only minor reductions over time due to mergers (see Fig. 4.5).

More recent efforts include the implementation of the Pop III.1 model in galaxy formation and evolution models, following the co-evolution of SMBH and the host galaxy (Cammelli et al., 2024). In the previous Chapters we have shown, based on galaxy population properties and scaling relation, that d_{iso} must be smaller than 75 kpc in order to reproduce the local GSMF and SMBH occupation fraction. The HUDF variability estimates of $n_{SMBH}(z)$ presented here, for the *compilation* case, span from $few \times 10^{-3}$ cMpc^{-3} at redshifts $z \lesssim 5$ rising by almost one order of magnitude up to $\sim 3 \times 10^{-3}$ cMpc^{-3} by $z \sim 8-9$. We also note that in the lower redshift range ($z \lesssim 5$), we are less sensitive to AGN variability as the intrinsic UV-optical emission from the central engine, responsible for the variability, progressively drops out of the considered IR filters, moving to the visible spectrum. Additionally, time dilation decreases as we observe at lower redshift. Since IR variability is expected to occur on longer time scales with respect to the central variation in the UV-optical (for a review, e.g. Lyu and Rieke, 2022), this might wash out the variable signal at later epochs. This suggests that our estimate of the $n_{SMBH}(z)$ may be underestimated at these redshifts and overall they indicate a more plausible scenario where d_{iso} is $\lesssim 75$ kpc. Similar conclusions can be drawn for the estimates of $n_{SMBH}(z)$ at $z \gtrsim 6$ in the simple UVLF correction case, with a slightly less constraining value for d_{iso} smaller than 100 kpc. For HMT models reproducing such number densities at $z \gtrsim 6$ would translate in setting the halo mass threshold for SMBH formation to be $\lesssim 3 \times 10^{10}$ M$_\odot$. However, these models increasingly face tension as they predict a significantly higher abundance of SMBHs than the estimated n_{SMBH} at $z = 0$. Moreover, in the high redshift regime, the HUDF variability-based estimate of n_{SMBH} is about 100 times larger than the direct collapse model prediction by Chon et al. (2016), and at least 10,000 times greater than the estimate provided by Wise et al. (2019).

Finally, provided that the estimated values presented here for the SMBH number density will be further consolidated at high redshift, especially in the *compilation* case for UVLF corrections, this result would be in tension with the local estimate of n_{SMBH} at $z = 0$ (e.g., Banik et al., 2019). Assuming negligible effect of mergers in reducing the overall number of SMBHs, this would suggest that we are ignoring a potentially significant population of SMBH in the local Universe and $z \sim 0$ estimates would need to be revised.

4.4.0.1 *Close SMBH pairs at the cosmic Dawn*

Similarly to the analysis carried out in H24, a valuable test for SMBH seeding schemes consists in measuring the actual distance between candidate AGN pairs. It is helpful to remind here that the Pop III.1 model with $d_{iso} \simeq 100$ kpc in proper distance would correspond to a comoving separation of about 3 cMpc (if the relative motion by $z \sim 7$ is negligible). Also, the HUDF footprint at this epoch would be ~ 5.5 cMpc on a side. In our sample, at

$z > 6$ we are left with the 27 sources listed in Table 4.3 and 3 of them have spectroscopic redshift available. Since we cannot asses the true distance with even one photometric redshift in a single pair (already a $\Delta z = 0.01$ at $z \sim 7$ would be ~ 2.5 cMpc, comparable to d_{iso}), we have searched among the pairs having both spectroscopic redshift available. In this high-z regime, we find none and no constraint can be used to rule out specific scenarios. Among the overall sample, we identified 24 pairs separated by less than 3 Mpc in the redshift range $0.6 < z < 1.5$. However, caution must be exercised at later epochs as SMBH host galaxies start merging and SMBHs can be likely found orbiting around each other in close pairs. In order to use this test as constraint for SMBH formation scenarios, spectroscopic confirmation of high redshift variable candidates are crucially needed. On a similar note, confirming the multiplicity nature of these source would assess the prediction of the Pop III.1 model as only single SMBHs powering these AGNs would be a viable option.

5

FINAL REMARKS

In this Thesis work, we have first studied the implication of the Pop III.1 seeding mechanism when embedded in a global approach within a galaxy formation model. In particular, we have investigated the ensemble properties of galaxy populations and how the presence of such massive SMBH seeds affects their evolution. We have looked at the occupation fractions as a function of several galaxy properties and their evolution through cosmic time as well as at the scaling relations between the host galaxy and the SMBH. The predicted growth of the SMBHs has been used to predict their observational contribution to the galaxy LF, fitting the estimated masses and luminosities into the recent landscape of JWST observations. In the second part of the Thesis we have illustrated results from a recent observational campaign with HST using a novel technique to identify, via photometric variability, potential candidates for high redshift SMBHs. In the following we present the main conclusions of each Chapter.

5.1 SUMMARY

5.1.1 *Host galaxies of Pop III.1 seeded SMBHs*

In Chapter 2, we present a new semi-analytical approach to modelling SMBH seeding scenarios within galaxy formation models, incorporating the PINOCCHIO code for dark matter halos and the GAEA model for galaxy evolution. By leveraging merger trees and refined models for subhalos, this method allows the study of a wide range of galaxy and SMBH properties with a computational cost of a fully semi-analytical approach. Our calibration against the Millennium Simulation highlights the capability to produce consistent halo and galaxy properties across cosmic time, as well as merging times for satellites galaxies.

Key findings include the impact of different seeding models — particularly the Pop III.1 seeding scenario — on the occupation fractions, GSMF, the role of AGN feedback, and mass functions of SMBHs. For the Pop III.1 models, SMBH seeds tend to form at early epochs ($z \sim 25 - 30$), exhibiting a high occupation fraction in massive galaxies, especially with an isolation distance $d_{iso} < 75$ kpc, which our models suggest as fiducial values to reproduce local constraints. This distance aligns with the radius of ionizing feedback from massive Pop III.1 stars, hinting that radiative feedback from these sources may be a decisive factor in regulating SMBH seeding.

The predictions of the M_{BH}-M_{star} relations for both low and high redshifts demonstrate some distinct trends among seeding scenarios, with the Pop III.1 model yielding a modestly steeper relation at the high-mass end. At lower masses, different initial seed masses are shown to affect the slope of the BHMF, although uncertainties in observational constraints make definitive conclusions challenging. Moreover, AGN feedback significantly influences the shape of the GSMF by reducing the SFR and quenching massive galaxies, consistent with observed trends in local galaxies.

5.1.2 *High redshift SMBHs in the JWST era*

We utilized the semi-analytical approach described in Chapter 2 to make predictions on the observable properties of SMBHs, focusing on AGN characteristics. Specifically, we compared model predictions for AGN properties — such as black hole masses, bolometric luminosities, and UVLFs — with observational constraints. We focused on the redshift range $4 < z < 9$, where recent JWST observations are available. Our objective was to assess whether high-redshift data could help distinguish between different SMBH formation models, including the Pop III.1 scenario. Given the uncertainties and apparent inconsistencies in JWST estimates, we also re-evaluated upper limit measurements and considered potential sources of error.

The combined UVLFs across seeding models align closely with observed data at the brightest magnitudes, accounting for dust attenuation uncertainties. The total UVLFs show limited cosmic evolution up to $z \sim 9$, with minimal sensitivity to the choice of seeding model. Notably, the Pop III.1 model matches the faint AGN UVLFs observed by JWST, while the HMT model tends to overestimate the UVLFs for $M_{UV} \lesssim -20$. However, our predictions diverge from JWST estimates of BH masses and bolometric luminosities by up to 1–2 orders of magnitude. To address these inconsistencies, we re-examined the methods for deriving bolometric luminosities and BH masses. By assuming UV emission is entirely AGN-driven, we calculated upper limits for these quantities. These UV-based estimates were significantly lower than those derived from Hα measurements, with the latter exceeding our UV-based upper limits. Using M_{UV}-derived estimates alleviates some tension between model predictions and observed data.

5.1.3 *A census of SMBHs via AGN variability in the HUDF*

In the final Chapter, we presented redshift-dependent estimates of SMBH number densities, n_{SMBH}, extending from the local universe up to $z \sim 9$. Utilizing photometric variability data from HST HUDF observations across multiple epochs, we estimated SMBH densities by identifying candidate AGNs through variability analysis, bypassing the limitations of traditional AGN diagnostics.

Our findings suggest that AGN variability can reliably identify SMBHs, especially at high redshift, where standard AGN indicators often fall short. The resulting SMBH number densities from our analysis allow for a comparison with model predictions. For the Pop III.1 model, at $z \gtrsim 6$ this suggests an isolation distance $d_{iso} \lesssim 75$ kpc, which aligns with our results from the previous Chapters. Meanwhile, models based on DCBHs show significant underprediction for n_{SMBH}.

We also attempt to quantify the geometrical separation in close pairs of SMBHs. Our analysis is limited due to insufficient spectroscopic redshifts among observed AGN pairs. However, at intermediate redshifts $0.6 < z < 1.5$, we detect several candidate pairs within a 3 cMpc separation, which may offer a future test for seeding mechanisms, pending spectroscopic confirmation.

5.2 FUTURE PROSPECTS

The Pop III.1 model examined in this Thesis provides valuable insights into supermassive black hole formation and incorporates beyond-standard-model physics by considering Weakly Interacting Massive Particles as dark matter candidates. The numerous predictions and results presented here have significant potential for further exploration of this model, as well as other SMBH seeding models.

- The current Pop III.1 model implementation assumes feedback sources are present in every halo. Moving forward, we aim to refine this by limiting feedback sources to other Pop III.1 seeded halos or possibly including Pop III.2 seeded halos as well. Additionally, we plan to introduce a time-dependent isolation radius to more accurately represent the expansion of radiation bubbles over time.
- A next major step that is needed to enable further comparison of model results with observational data in the JWST era requires a better understanding of possible observational systematics. Further observations are needed to identify and address systematic sources of error, including potential issues with bolometric corrections, scaling relations, and line flux/FWHM measurements. Current observational constraints are inadequate to conclusively distinguish between SMBH formation models. Additional research is necessary to reduce uncertainties and address discrepancies between model predictions and observational data.
- Another promising avenue is to develop different SMBH growth models to make predictions for their observability and to possibly alleviate challenging tensions. Especially in the high redshift regime, our assumed model for SMBH growth might not be able to capture the initial accretion phase of these objects. We plan to investigate several recipes, from simple Eddington limited accretion scenarios to SFR driven accretion schemes.

- Our calibrated galaxy survival time for satellite galaxies can serve as a first step to calculate the merging rates of SMBH binaries. Under simplistic assumptions, it will be possible to predict the gravitational wave emission from these different seeding scenarios for the current and next generation of GW detectors.
- The variability analysis to search for SMBH underscores the need for further investigation and refinement. Future work on variability data and spectroscopic follow-up at high redshift will be crucial for assessing SMBH multiplicity and seeding scenarios, especially to discriminate different SMBH formation mechanisms.

In conclusion, the Pop III.1 model offers an attractive mechanism to explain the origin of supermassive black holes across cosmic time. The future prospects listed above draw a promising and interesting picture for the development of new methods and techniques that will allow for a better understanding of our Universe.

A

GALAXY SURVIVAL TIME

As discussed in Section 2.1.4, we adopt the sum of two different time scales when dealing with galaxy mergers. First, we estimate the time the subhalo will survive within the main halo group, which has been evaluated via comparison against N-body simulations Berner et al. (2022). Secondly we keep track of the orphan galaxies by adding a second time which mimics the phase during which a galaxy would orbit around the central one and eventually merge with it. In Fig. A.1 we present the estimated total survival time of satellite galaxies as a function of the halo mass ratio calculated at the halo merging time taken from PINOCCHIO using the values reported in Eq. (12). In particular, we report data from the seeded halos for the $d_{iso} = 50\,\mathrm{kpc}$ case. This distribution of points has been calibrated to reproduce the Millennium one as obtained from the standard implementation of GAEA. The impact of the total survival time proposed in Eq. (10) and (11) introduces a dual dependence. Mergers with comparable halo masses tend to last on average few Gyr and typically this happens when the two merging halos belong to the low mass end of the halo mass function. As one increases the mass of the primary halo, it will eventually merge with even smaller halos. For these cases the total survival time will overshoot the age of the Universe and the two structures will never merge. This is in agreement with findings in simulations (Somerville and Davé, 2015). The more comparable are the encounters, the more efficient we expect the interaction to be (i.e., via dynamical friction, ram pressure stripping, tidal disruption forces), resulting in a shorter time scale. Conversely, on average small galaxies joining a massive central galaxy group are thought to undergo very minor physical effects, ending up orbiting around the central galaxy undisturbed for tens or thousands of Hubble times.

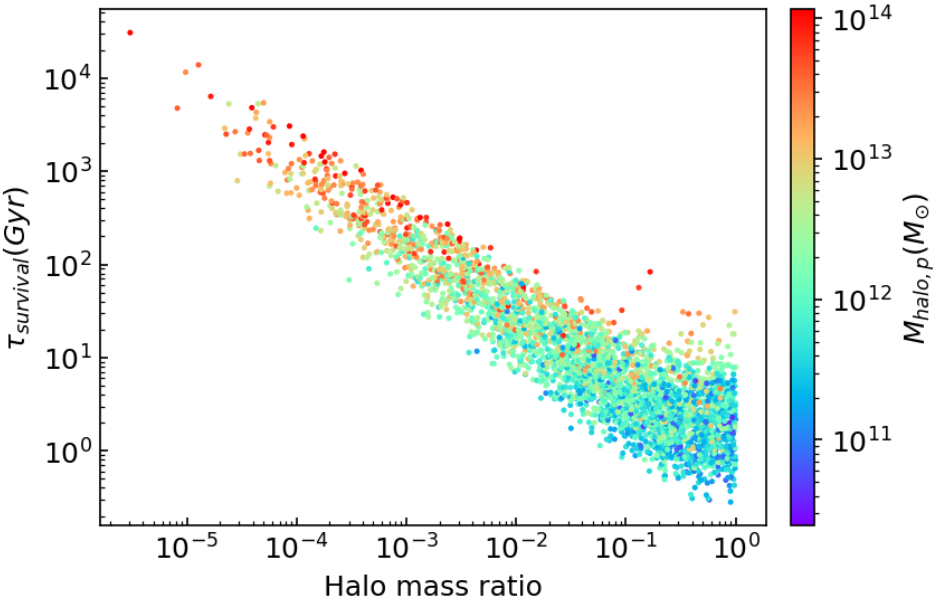


Figure A.1: Estimated total survival time in Gyr for seeded halos for the $d_{iso} = 50$ kpc case as a function of the halo mass ratio and color coded with the mass of the most massive (primary) halo.

B

CALIBRATED GALAXY STELLAR MASS FUNCTION

The galaxy survival time introduced in Section 2.1.3 and showed in Appendix A as a function of the halo mass ratio plays a crucial role in shaping the GSMF. Once again we note that our calibration process has been performed aiming at reproducing the Millennium-based result at redshift 0, and not against observational data. In this appendix we show the results from the calibrated GSMFs at redshifts $z \sim 1$ to 3. In particular we emphasize that our PINOCCHIO-based GSMF well reproduces the Millennium-based predictions up to $z \sim 3$ and above. The agreement between the 2 different trends is depicted in Fig. B.1, where we indicate Millennium- and PINOCCHIO-based GSMFs in blue and red, respectively. This supports the goodness of our calibration of the galaxy survival time in order to obtain reliable model galaxy catalogues.

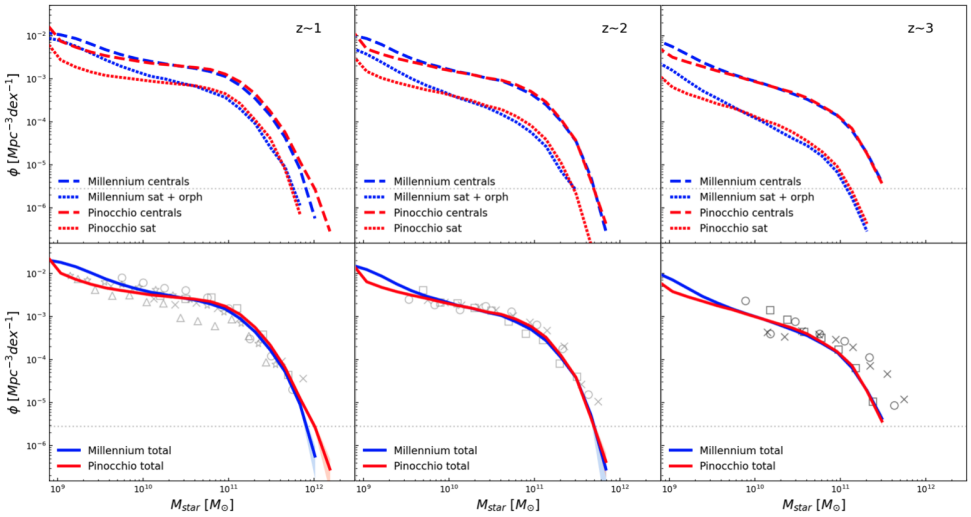


Figure B.1: GSMF predictions based on the Millennium simulation box (blue lines) compared to the GSMF extracted from our Millennium-like PINOCCHIO box run used for the calibration of the merging times (see Section 2.1.4). *Upper panels*: satellite and central galaxies contributions to the GSMF are reported separately in dotted and dashed lines, respectively. *Lower panels*: total GSMFs. Gray symbols show observational estimates as reported in Fig. 2.2.

BIBLIOGRAPHY

Abbott, B. P. et al. (Feb. 2016). 'Observation of Gravitational Waves from a Binary Black Hole Merger'. In: *Physical Review Letters* 116.6, 061102, p. 061102. DOI: 10.1103/PhysRevLett.116.061102. arXiv: 1602.03837 [gr-qc].

Abbott, R. et al. (2020a). 'GW190521: A Binary Black Hole Merger with a Total Mass of 150$M_\odot$'. In: *Phys. Rev. Lett.* 125.10, p. 101102. DOI: 10.1103/PhysRevLett.125.101102. arXiv: 2009.01075 [gr-qc].

— (2020b). 'Properties and Astrophysical Implications of the 150 $M_\odot$ Binary Black Hole Merger GW190521'. In: *Astrophys. J. Lett.* 900.1, p. L13. DOI: 10.3847/2041-8213/aba493. arXiv: 2009.01190 [astro-ph.HE].

Abel, Tom, Greg L. Bryan and Michael L. Norman (Jan. 2002). 'The Formation of the First Star in the Universe'. In: *Science* 295.5552, pp. 93–98. DOI: 10.1126/science.295.5552.93. arXiv: astro-ph/0112088 [astro-ph].

Abuter, R. et al. (Mar. 2024). 'A dynamical measure of the black hole mass in a quasar 11 billion years ago'. en. In: *Nature* 627.8003, 281–285. ISSN: 1476-4687. DOI: 10.1038/s41586-024-07053-4.

Adams, N. J. et al. (May 2020). 'The rest-frame UV luminosity function at z ~ 4: a significant contribution of AGNs to the bright end of the galaxy population'. In: *MNRAS* 494.2, pp. 1771–1783. DOI: 10.1093/mnras/staa687. arXiv: 1912.01626 [astro-ph.GA].

Agazie, Gabriella et al. (July 2023a). 'The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background'. In: *The Astrophysical Journal Letters* 951.1, L8, p. L8. DOI: 10.3847/2041-8213/acdac6. arXiv: 2306.16213 [astro-ph.HE].

Agazie, Gabriella et al. (July 2023b). 'The NANOGrav 15 yr Data Set: Observations and Timing of 68 Millisecond Pulsars'. In: *The Astrophysical Journal Letters* 951.1, L9, p. L9. DOI: 10.3847/2041-8213/acda9a. arXiv: 2306.16217 [astro-ph.HE].

Akins, Hollis B. et al. (June 2024). 'COSMOS-Web: The over-abundance and physical nature of "little red dots"–Implications for early galaxy and SMBH assembly'. In: *arXiv e-prints*, arXiv:2406.10341, arXiv:2406.10341. DOI: 10.48550/arXiv.2406.10341. arXiv: 2406.10341 [astro-ph.GA].

Akiyama, Masayuki et al. (Jan. 2018). 'The quasar luminosity function at redshift 4 with the Hyper Suprime-Cam Wide Survey'. In: *PASJ* 70, S34, S34. DOI: 10.1093/pasj/psx091. arXiv: 1704.05996 [astro-ph.GA].

Arnouts, S. et al. (Jan. 2005). 'The GALEX VIMOS-VLT Deep Survey Measurement of the Evolution of the 1500 Å Luminosity Function'. In: *ApJ* 619.1, pp. L43–L46. DOI: 10.1086/426733. arXiv: astro-ph/0411391 [astro-ph].

Bañados, Eduardo et al. (Jan. 2018). 'An 800-million-solar-mass black hole in a significantly neutral Universe at a redshift of 7.5'. In: *Nature*

553.7689, pp. 473–476. DOI: 10.1038/nature25180. arXiv: 1712.01860 [astro-ph.GA].

Bacon, Roland et al. (Feb. 2023). 'The MUSE Hubble Ultra Deep Field surveys: Data release II'. In: *A&A* 670, A4, A4. DOI: 10.1051/0004-6361/202244187. arXiv: 2211.08493 [astro-ph.GA].

Baldwin, J. A., M. M. Phillips and R. Terlevich (Feb. 1981). 'Classification parameters for the emission-line spectra of extragalactic objects.' In: *Publications of the Astronomical Society of the Pacific* 93, pp. 5–19. DOI: 10.1086/130766.

Banik, Nilanjan, Jonathan C. Tan and Pierluigi Monaco (2019). 'The Formation of Supermassive Black Holes from Population III.1 Seeds. I. Cosmic Formation Histories and Clustering Properties'. en. In: *MNRAS* 483.3. arXiv: 1608.04421, 3592–3606. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/sty3298.

Banik, Nilanjan, Jonathan C. Tan and Pierluigi Monaco (Mar. 2019). 'The formation of supermassive black holes from Population III.1 seeds. I. Cosmic formation histories and clustering properties'. In: *Monthly Notices of the Royal Astronomical Society* 483.3, pp. 3592–3606. DOI: 10.1093/mnras/sty3298. arXiv: 1608.04421 [astro-ph.GA].

Barber, Christopher et al. (July 2016). 'The origin of compact galaxies with anomalously high black hole masses'. In: *MNRAS* 460.1, 1147–1161. ISSN: 0035-8711. DOI: 10.1093/mnras/stw1018.

Beckwith, Steven V. W. et al. (Nov. 2006). 'The Hubble Ultra Deep Field'. In: *The Astronomical Journal* 132.5, pp. 1729–1755. DOI: 10.1086/507302. arXiv: astro-ph/0607632 [astro-ph].

Begelman, Mitchell C., Elena M. Rossi and Philip J. Armitage (July 2008). 'Quasi-stars: accreting black holes inside massive envelopes'. In: *Monthly Notices of the Royal Astronomical Society* 387.4, pp. 1649–1659. DOI: 10.1111/j.1365-2966.2008.13344.x. arXiv: 0711.4078 [astro-ph].

Begelman, Mitchell C., Marta Volonteri and Martin J. Rees (July 2006). 'Formation of supermassive black holes by direct collapse in pre-galactic haloes'. In: *MNRAS* 370.1, 289–298. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2006.10467.x.

Bell, Eric F. et al. (2003). 'The Optical and Near-Infrared Properties of Galaxies. I. Luminosity and Stellar Mass Functions'. In: *ApJ* 149, 289–312. ISSN: 0067-0049. DOI: 10.1086/378847.

Berg, Danielle A. et al. (Mar. 2019). 'The Chemical Evolution of Carbon, Nitrogen, and Oxygen in Metal-poor Dwarf Galaxies'. In: *ApJ* 874. ADS Bibcode: 2019ApJ...874...93B, p. 93. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab020a.

Berner, Pascale et al. (Nov. 2022). 'Rapid simulations of halo and subhalo clustering'. In: *Journal of Cosmology and Astroparticle Physics* 2022. ADS Bibcode: 2022JCAP...11..002B, p. 002. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2022/11/002.

Bertin, E. and S. Arnouts (June 1996). 'SExtractor: Software for source extraction.' In: *A&A, Supplement* 117, pp. 393–404. DOI: 10.1051/aas:1996164.

Bhowmick, Aklant K. et al. (Feb. 2022). 'Impact of gas spin and Lyman-Werner flux on black hole seed formation in cosmological simulations: implications for direct collapse'. In: *Monthly Notices of the Royal Astronomical Society* 510.1, pp. 177–196. DOI: 10.1093/mnras/stab3439. arXiv: 2107.06899 [astro-ph.GA].

Bhowmick, Aklant K et al. (Oct. 2022). 'Probing the z $\gtrsim$ 6 quasars in a universe with IllustrisTNG physics: impact of gas-based black hole seeding models'. In: *MNRAS* 516.1, 138–157. ISSN: 0035-8711. DOI: 10.1093/mnras/stac2238.

Bhowmick, Aklant K. et al. (July 2024a). 'Growth of high redshift supermassive black holes from heavy seeds in the BRAHMA cosmological simulations: Implications of overmassive black holes'. In: *MNRAS*. ADS Bibcode: 2024MNRAS.tmp.1789B. ISSN: 0035-8711. DOI: 10.1093/mnras/stae1819. URL: https://ui.adsabs.harvard.edu/abs/2024MNRAS.tmp.1789B.

Bhowmick, Aklant K. et al. (July 2024b). 'Introducing the BRAHMA simulation suite: signatures of low-mass black hole seeding models in cosmological simulations'. In: *MNRAS* 531. ADS Bibcode: 2024MNRAS.531.4311B, 4311–4335. ISSN: 0035-8711. DOI: 10.1093/mnras/stae1386.

Bhowmick, Aklant K. et al. (Apr. 2024c). 'Representing low-mass black hole seeds in cosmological simulations: A new sub-grid stochastic seed model'. In: *MNRAS* 529. ADS Bibcode: 2024MNRAS.529.3768B, 3768–3792. ISSN: 0035-8711. DOI: 10.1093/mnras/stae780.

Binney, James and Scott Tremaine (1987). *Galactic dynamics*. Princeton Univ. Press Princeton, NJ, p. 747.

Birrer, Simon et al. (Sept. 2014). 'A Simple Model Linking Galaxy and Dark Matter Evolution'. In: *ApJ* 793. ADS Bibcode: 2014ApJ...793...12B, p. 12. ISSN: 0004-637X. DOI: 10.1088/0004-637X/793/1/12.

Boekholt, T. C. N. et al. (May 2018). 'Formation of massive seed black holes via collisions and accretion'. In: *Monthly Notices of the Royal Astronomical Society* 476.1, pp. 366–380. DOI: 10.1093/mnras/sty208. arXiv: 1801.05841 [astro-ph.GA].

Bogdán, Ákos et al. (Jan. 2024). 'Evidence for heavy-seed origin of early supermassive black holes from a z $\approx$ 10 X-ray quasar'. In: *Nature Astronomy* 8, pp. 126–133. DOI: 10.1038/s41550-023-02111-9. arXiv: 2305.15458 [astro-ph.GA].

Boutsia, Konstantina et al. (May 2021). 'The Luminosity Function of Bright QSOs at z $\sim$ 4 and Implications for the Cosmic Ionizing Background'. In: *ApJ* 912.2, 111, p. 111. DOI: 10.3847/1538-4357/abedb5. arXiv: 2103.10446 [astro-ph.GA].

Bouwens, R. J. et al. (2021). 'New Determinations of the UV Luminosity Functions from z 9 to z 2 show a remarkable consistency with halo growth and a constant star formation efficiency'. In: *The Astronomical Journal* 162.2. arXiv:2102.07775, p. 47. ISSN: 0004-6256, 1538-3881. DOI: 10.3847/1538-3881/abf83e.

Bowler, R. A. A. et al. (Sept. 2015). 'The galaxy luminosity function at z $\sim$ 6 and evidence for rapid evolution in the bright end from z $\sim$ 7 to 5'.

In: *MNRAS* 452.2, pp. 1817–1840. DOI: 10.1093/mnras/stv1403. arXiv: 1411.2976 [astro-ph.GA].

Bowler, R. A. A. et al. (Apr. 2017). 'Unveiling the nature of bright z ~ 7 galaxies with the Hubble Space Telescope'. In: *MNRAS* 466.3, pp. 3612–3635. DOI: 10.1093/mnras/stw3296. arXiv: 1605.05325 [astro-ph.GA].

Boyett, Kit et al. (Jan. 2024). 'Extreme emission line galaxies detected in JADES JWST/NIRSpec I: inferred galaxy properties'. In: *arXiv e-prints*. ADS Bibcode: 2024arXiv240116934B. DOI: 10.48550/arXiv.2401.16934. URL: https://ui.adsabs.harvard.edu/abs/2024arXiv240116934B.

Boylan-Kolchin, Michael, Chung-Pei Ma and Eliot Quataert (Jan. 2008). 'Dynamical friction and galaxy merging time-scales'. In: *MNRAS* 383.1, 93–101. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2007.12530.x.

Brennan, Ryan et al. (June 2018). 'Momentum-driven Winds from Radiatively Efficient Black Hole Accretion and Their Impact on Galaxies'. en. In: *ApJ* 860.1, p. 14. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aac2c4.

Bromm, Volker, Paolo S. Coppi and Richard B. Larson (Jan. 2002). 'The Formation of the First Stars. I. The Primordial Star-forming Cloud'. In: *The Astrophysical Journal* 564.1, pp. 23–51. DOI: 10.1086/323947. arXiv: astro-ph/0102503 [astro-ph].

Bromm, Volker and Abraham Loeb (Oct. 2003). 'Formation of the First Supermassive Black Holes'. en. In: *ApJ* 596.1, p. 34. ISSN: 0004-637X. DOI: 10.1086/377529.

Bundy, Kevin, Richard S. Ellis and Christopher J. Conselice (June 2005). 'The Mass Assembly Histories of Galaxies of Various Morphologies in the GOODS Fields'. In: *ApJ* 625. ADS Bibcode: 2005ApJ...625..621B, 621–632. ISSN: 0004-637X. DOI: 10.1086/429549.

Bunker, Andrew J. et al. (Sept. 2023). 'JADES NIRSpec Spectroscopy of GN-z11: Lyman-α emission and possible enhanced nitrogen abundance in a z = 10.60 luminous galaxy'. In: *Astronomy & Astrophysics* 677, A88, A88. DOI: 10.1051/0004-6361/202346159. arXiv: 2302.07256 [astro-ph.GA].

Cameron, Alex J. et al. (Sept. 2023). 'JADES: Probing interstellar medium conditions at z ~ 5.5–9.5 with ultra-deep JWST/NIRSpec spectroscopy'. en. In: *A&A* 677, A115. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202346107.

Cammelli, Vieri et al. (July 2024). 'The formation of supermassive black holes from Population III.1 seeds. III. Galaxy evolution and black hole growth from semi-analytic modelling'. In: *arXiv e-prints*, arXiv:2407.09949, arXiv:2407.09949. DOI: 10.48550/arXiv.2407.09949. arXiv: 2407.09949 [astro-ph.GA].

Cammelli, Vieri et al. (Nov. 2024). 'The formation of supermassive black holes from Population III.1 seeds. III. Galaxy evolution and black hole growth from semi-analytic modelling'. In: *MNRAS* 536.1, 851–870. ISSN: 0035-8711. DOI: 10.1093/mnras/stae2663.

Carr, Bernard and Florian Kühnel (Oct. 2020). 'Primordial Black Holes as Dark Matter: Recent Developments'. In: *Annual Review of Nuclear and Particle Science* 70, pp. 355–394. DOI: 10.1146/annurev-nucl-050520-125911. arXiv: 2006.02838 [astro-ph.CO].

Casertano, S. et al. (Dec. 2000). 'WFPC2 Observations of the Hubble Deep Field South'. In: *The Astron. J.* 120, pp. 2747–2824. DOI: 10.1086/316851. eprint: astro-ph/0010245.

Chandrasekhar, S. (Mar. 1943). 'Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction.' In: *ApJ* 97. ADS Bibcode: 1943ApJ....97..255C, p. 255. ISSN: 0004-637X. DOI: 10.1086/144517.

Charlot, Stéphane and S. Michael Fall (Aug. 2000). 'A Simple Model for the Absorption of Starlight by Dust in Galaxies'. In: *ApJ* 539.2, pp. 718–731. DOI: 10.1086/309250. arXiv: astro-ph/0003128 [astro-ph].

Chon, Sunmyon and Kazuyuki Omukai (May 2020). 'Supermassive star formation via super competitive accretion in slightly metal-enriched clouds'. In: *Monthly Notices of the Royal Astronomical Society* 494.2, pp. 2851–2860. DOI: 10.1093/mnras/staa863. arXiv: 2001.06491 [astro-ph.GA].

Chon, Sunmyon et al. (Nov. 2016). 'COSMOLOGICAL SIMULATIONS OF EARLY BLACK HOLE FORMATION: HALO MERGERS, TIDAL DISRUPTION, AND THE CONDITIONS FOR DIRECT COLLAPSE'. en. In: *Apj* 832.2, p. 134. ISSN: 0004-637X. DOI: 10.3847/0004-637X/832/2/134.

Churazov, E. et al. (Oct. 2005). 'Supermassive black holes in elliptical galaxies: switching from very bright to very dim'. In: *MNRAS* 363.1, pp. L91–L95. DOI: 10.1111/j.1745-3933.2005.00093.x. arXiv: astro-ph/0507073 [astro-ph].

Cohen, S. H. et al. (Mar. 2006). 'Clues to Active Galactic Nucleus Growth from Optically Variable Objects in the Hubble Ultra Deep Field'. en. In: *ApJ* 639.2, p. 731. ISSN: 0004-637X. DOI: 10.1086/499577.

Cole, Shaun et al. (Sept. 2001). 'The 2dF galaxy redshift survey: near-infrared galaxy luminosity functions'. In: *MNRAS* 326. ADS Bibcode: 2001MNRAS.326..255C, 255–273. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2001.04591.x.

Croton, Darren J. et al. (2006). 'The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies'. In: *MNRAS* 365.1. arXiv: astro-ph/0508046, 11–28. ISSN: 00358711, 13652966. DOI: 10.1111/j.1365-2966.2005.09675.x.

Dayal, Pratika et al. (June 2019). 'The hierarchical assembly of galaxies and black holes in the first billion years: predictions for the era of gravitational wave astronomy'. In: *MNRAS* 486.2, 2336–2350. ISSN: 0035-8711. DOI: 10.1093/mnras/stz897.

De Lucia, Gabriella (May 2019). 'Lighting Up Dark Matter Haloes'. In: *Galaxies* 7. ADS Bibcode: 2019Galax...7...56D, p. 56. DOI: 10.3390/galaxies7020056.

De Lucia, Gabriella and Jérémy Blaizot (Feb. 2007). 'The hierarchical formation of the brightest cluster galaxies'. In: *MNRAS* 375.1, 2–14. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2006.11287.x.

De Lucia, Gabriella, Michaela Hirschmann and Fabio Fontanot (Feb. 2019). 'Nature versus nurture: what regulates star formation in satellite galaxies?' In: *MNRAS* 482.4, 5041–5051. ISSN: 0035-8711. DOI: 10.1093/mnras/sty3059.

De Lucia, Gabriella et al. (2010). 'A semi-analytic model comparison - gas cooling and galaxy mergers'. In: *MNRAS*. arXiv:1003.3021, no–no. ISSN: 00358711, 13652966. DOI: 10.1111/j.1365-2966.2010.16806.x.

De Lucia, Gabriella et al. (June 2011). 'Times, environments and channels of bulge formation in a Lambda cold dark matter cosmology'. In: *MNRAS* 414.2, 1439–1454. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.18475.x.

De Lucia, Gabriella et al. (Nov. 2014). 'Elemental abundances in Milky Way-like galaxies from a hierarchical galaxy formation model'. In: *MNRAS* 445.1, 970–987. ISSN: 0035-8711. DOI: 10.1093/mnras/stu1752.

De Lucia, Gabriella et al. (July 2024). 'Tracing the quenching journey across cosmic time'. In: *A&A* 687. ADS Bibcode: 2024A&A...687A..68D, A68. ISSN: 0004-6361. DOI: 10.1051/0004-6361/202349045.

Delvecchio, I. et al. (Mar. 2020). 'The Evolving AGN Duty Cycle in Galaxies Since z ∼ 3 as Encoded in the X-Ray Luminosity Function'. en. In: *ApJ* 892.1, p. 17. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab789c.

Devriendt, J. E. G., B. Guiderdoni and R. Sadat (Oct. 1999). 'Galaxy modelling. I. Spectral energy distributions from far-UV to sub-mm wavelengths'. In: *A&A* 350, pp. 381–398. DOI: 10.48550/arXiv.astro-ph/9906332. arXiv: astro-ph/9906332 [astro-ph].

Di Matteo, Tiziana, Daniel Angles-Alcazar and Francesco Shankar (Apr. 2023). 'Massive black holes in galactic nuclei: Theory and Simulations'. In: *arXiv e-prints*. ADS Bibcode: 2023arXiv230411541D. DOI: 10.48550/arXiv.2304.11541. URL: https://ui.adsabs.harvard.edu/abs/2023arXiv230411541D.

Drory, N. et al. (June 2004). 'The Munich Near-Infrared Cluster Survey (MUNICS). VI. The Stellar Masses of K-Band-selected Field Galaxies to z 1.2'. In: *ApJ* 608. ADS Bibcode: 2004ApJ...608..742D, 742–751. ISSN: 0004-637X. DOI: 10.1086/420781.

EPTA Collaboration et al. (Oct. 2023). 'The second data release from the European Pulsar Timing Array. I. The dataset and timing analysis'. In: *Astronomy & Astrophysics* 678, A48, A48. DOI: 10.1051/0004-6361/202346841. arXiv: 2306.16224 [astro-ph.HE].

Ebisuzaki, T. (2003). 'Missing Link Found? - The "runaway" path to supermassive black holes'. In: *Astrophysical Supercomputing using Particle Simulations*. Ed. by Junichiro Makino and Piet Hut. Vol. 208. IAU Symposium, p. 157.

Einstein, A. (Jan. 1916). 'Die Grundlage der allgemeinen Relativitätstheorie'. In: *Annalen der Physik* 354.7, pp. 769–822. DOI: 10.1002/andp.19163540702.

Eisenstein, Daniel J. et al. (June 2023). 'Overview of the JWST Advanced Deep Extragalactic Survey (JADES)'. In: *arXiv e-prints* arXiv:2306.02465. arXiv:2306.02465 [astro-ph]. DOI: 10.48550/arXiv.2306.02465. URL: http://arxiv.org/abs/2306.02465.

Ellis, Richard S. et al. (Jan. 2013). 'The Abundance of Star-forming Galaxies in the Redshift Range 8.5-12: New Results from the 2012 Hubble Ultra

Deep Field Campaign'. In: *The Astrophysical Journal Letters* 763.1, L7, p. L7. DOI: 10.1088/2041-8205/763/1/L7. arXiv: 1211.6804 [astro-ph.CO].

Event Horizon Telescope Collaboration et al. (Apr. 2019). 'First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole'. In: *The Astrophysical Journal Letters* 875.1, L1, p. L1. DOI: 10.3847/2041-8213/ab0ec7. arXiv: 1906.11238 [astro-ph.GA].

Event Horizon Telescope Collaboration et al. (May 2022). 'First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way'. In: *The Astrophysical Journal Letters* 930.2, L12, p. L12. DOI: 10.3847/2041-8213/ac6674.

Fan, Xiaohui et al. (July 2006). 'Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z~6 Quasars. II. A Sample of 19 Quasars'. In: *The Astron. J.* 132.1, pp. 117–136. DOI: 10.1086/504836. arXiv: astro-ph/0512082 [astro-ph].

Feng, Wei-Xiang, Hai-Bo Yu and Yi-Ming Zhong (June 2021). 'Seeding Supermassive Black Holes with Self-interacting Dark Matter: A Unified Scenario with Baryons'. In: *The Astrophysical Journal Letters* 914.2, L26, p. L26. DOI: 10.3847/2041-8213/ac04b0. arXiv: 2010.15132 [astro-ph.CO].

Ferrarese, Laura (Oct. 2002). 'Beyond the Bulge: A Fundamental Relation between Supermassive Black Holes and Dark Matter Halos'. In: *The Astrophysical Journal* 578. ADS Bibcode: 2002ApJ...578...90F, 90–97. ISSN: 0004-637X. DOI: 10.1086/342308.

Ferrarese, Laura and Holland Ford (Feb. 2005). 'Supermassive Black Holes in Galactic Nuclei: Past, Present and Future Research'. In: *Space Science Reviews* 116.3-4, pp. 523–624. DOI: 10.1007/s11214-005-3947-6. arXiv: astro-ph/0411247 [astro-ph].

Finkelstein, Steven L. and Micaela B. Bagley (2022). 'On the Coevolution of the AGN and Star-forming Galaxy Ultraviolet Luminosity Functions at 3 < z < 9'. In: *ApJ* 938. ADS Bibcode: 2022ApJ...938...25F, p. 25. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac89eb.

Finkelstein, Steven L. et al. (Sept. 2015). 'The Evolution of the Galaxy Rest-frame Ultraviolet Luminosity Function over the First Two Billion Years'. In: *ApJ* 810.1, 71, p. 71. DOI: 10.1088/0004-637X/810/1/71. arXiv: 1410.5439 [astro-ph.GA].

Finkelstein, Steven L. et al. (July 2019). 'Conditions for Reionizing the Universe with a Low Galaxy Ionizing Photon Escape Fraction'. In: *ApJ* 879.1, 36, p. 36. DOI: 10.3847/1538-4357/ab1ea8. arXiv: 1902.02792 [astro-ph.CO].

Finkelstein, Steven L. et al. (Mar. 2023). 'CEERS Key Paper. I. An Early Look into the First 500 Myr of Galaxy Formation with JWST'. en. In: *ApJ* 946.1, p. L13. ISSN: 2041-8205. DOI: 10.3847/2041-8213/acade4.

Fiore, F. et al. (Jan. 2012). 'Faint high-redshift AGN in the Chandra deep field south: the evolution of the AGN luminosity function and black hole demography'. In: *A&A* 537, A16, A16. DOI: 10.1051/0004-6361/201117581. arXiv: 1109.2888 [astro-ph.CO].

Fiore, F. et al. (2017). 'AGN wind scaling relations and the co-evolution of black holes and galaxies'. In: *Astronomy and Astrophysics* 601. ADS

Bibcode: 2017A&A...601A.143F, A143. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201629478.

Fontana, A. et al. (Dec. 2006). 'The Galaxy mass function up to z in the GOODS-MUSIC sample: into the epoch of formation of massive galaxies'. en. In: *A&A* 459.33, 745–757. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361:20065475.

Fontanot, Fabio, Michaela Hirschmann and Gabriella De Lucia (June 2017a). 'Strong Stellar-driven Outflows Shape the Evolution of Galaxies at Cosmic Dawn'. en. In: *ApJ* 842.2, p. L14. ISSN: 2041-8205. DOI: 10.3847/2041-8213/aa74bd.

Fontanot, Fabio et al. (2006). 'The effect of stellar feedback and quasar winds on the active galactic nucleus population'. en. In: *MNRAS* 373.3, 1173–1187. ISSN: 0035-8711, 1365-2966. DOI: 10.1111/j.1365-2966.2006.11094.x.

Fontanot, Fabio et al. (Feb. 2017b). 'Variations of the stellar initial mass function in semi-analytical models: implications for the mass assembly and the chemical enrichment of galaxies in the gaea model'. In: *MNRAS* 464.4, 3812–3824. ISSN: 0035-8711. DOI: 10.1093/mnras/stw2612.

Fontanot, Fabio et al. (Apr. 2018). 'Variations of the stellar initial mass function in semi-analytical models – II. The impact of cosmic ray regulation'. In: *MNRAS* 475.2, 2467–2479. ISSN: 0035-8711. DOI: 10.1093/mnras/stx3323.

Fontanot, Fabio et al. (2020). 'The rise of active galactic nuclei in the galaxy evolution and assembly semi-analytic model'. In: *MNRAS* 496. ADS Bibcode: 2020MNRAS.496.3943F, 3943–3960. ISSN: 0035-8711. DOI: 10.1093/mnras/staa1716.

Fontanot, Fabio et al. (Apr. 2024). 'Variation of the stellar initial mass function in semi-analytical models. III. Testing the cosmic-ray regulated integrated galaxy-wide initial mass function'. en. In: *Astronomy & Astrophysics*. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202348688. URL: https://www.aanda.org/articles/aa/abs/forth/aa48688-23/aa48688-23.html.

Freese, Katherine et al. (June 2010). 'SUPERMASSIVE DARK STARS: DETECTABLE IN JWST'. en. In: *ApJ* 716.2, p. 1397. ISSN: 0004-637X. DOI: 10.1088/0004-637X/716/2/1397.

Fruchter, A. S. and R. N. Hook (Feb. 2002). 'Drizzle: A Method for the Linear Reconstruction of Undersampled Images'. In: *Publications of the ASP* 114, pp. 144–152. DOI: 10.1086/338393. eprint: astro-ph/9808087.

Fujimoto, S. et al. (Apr. 2022). 'A dusty compact object bridging galaxies and quasars at cosmic dawn'. In: *Nature* 604.7905, pp. 261–265. DOI: 10.1038/s41586-022-04454-1. arXiv: 2204.06393 [astro-ph.GA].

Gao, L. et al. (Aug. 2004). 'Galaxies and subhaloes in ΛCDM galaxy clusters'. In: *MNRAS* 352. ADS Bibcode: 2004MNRAS.352L...1G, L1–L5. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2004.08098.x.

García-Bellido, Juan (Dec. 2019). 'Primordial black holes and the origin of the matter-antimatter asymmetry'. In: *Philosophical Transactions of the*

Royal Society of London Series A 377.2161, p. 20190091. DOI: 10.1098/rsta.2019.0091.

Giallongo, E. et al. (June 2015). 'Faint AGNs at z > 4 in the CANDELS GOODS-S field: looking for contributors to the reionization of the Universe'. In: *A&A* 578, A83, A83. DOI: 10.1051/0004-6361/201425334. arXiv: 1502.02562 [astro-ph.CO].

Giallongo, E. et al. (Oct. 2019). 'Space Densities and Emissivities of Active Galactic Nuclei at z > 4'. In: *ApJ* 884.1, 19, p. 19. DOI: 10.3847/1538-4357/ab39e1.

Graham, Alister W. (Jan. 2016). 'Galaxy Bulges and Their Massive Black Holes: A Review'. In: *Galactic Bulges*. Ed. by Eija Laurikainen, Reynier Peletier and Dimitri Gadotti. Vol. 418. Astrophysics and Space Science Library, p. 263. DOI: 10.1007/978-3-319-19378-6_11. arXiv: 1501.02937 [astro-ph.GA].

Granato, Gian Luigi et al. (Jan. 2004). 'A Physical Model for the Coevolution of QSOs and Their Spheroidal Hosts'. en. In: *ApJ* 600.2, p. 580. ISSN: 0004-637X. DOI: 10.1086/379875.

Grazian, Andrea et al. (Jan. 2022). 'The Space Density of Ultra-luminous QSOs at the End of Reionization Epoch by the QUBRICS Survey and the AGN Contribution to the Hydrogen Ionizing Background'. In: *ApJ* 924.2, 62, p. 62. DOI: 10.3847/1538-4357/ac33a4. arXiv: 2110.13736 [astro-ph.GA].

Greene, Jenny E., Jay Strader and Luis C. Ho (2020). 'Intermediate-Mass Black Holes'. In: *Annual Review of Astronomy and Astrophysics* 58.1, 257–312. DOI: 10.1146/annurev-astro-032620-021835.

Greene, Jenny E. et al. (Sept. 2023). 'UNCOVER spectroscopy confirms a surprising ubiquity of AGN in red galaxies at $z > 5$'. In: *arXiv e-prints*. ADS Bibcode: 2023arXiv230905714G. DOI: 10.48550/arXiv.2309.05714. URL: https://ui.adsabs.harvard.edu/abs/2023arXiv230905714G.

Greene, Jenny E. et al. (Mar. 2024). 'UNCOVER Spectroscopy Confirms the Surprising Ubiquity of Active Galactic Nuclei in Red Sources at z > 5'. In: *ApJ* 964.1, 39, p. 39. DOI: 10.3847/1538-4357/ad1e5f. arXiv: 2309.05714 [astro-ph.GA].

Greif, Thomas H. and Volker Bromm (Nov. 2006a). 'Two populations of metal-free stars in the early Universe'. In: *Monthly Notices of the Royal Astronomical Society* 373.1, pp. 128–138. DOI: 10.1111/j.1365-2966.2006.11017.x. arXiv: astro-ph/0604367 [astro-ph].

— (Nov. 2006b). 'Two populations of metal-free stars in the early Universe'. In: *MNRAS* 373.1, pp. 128–138. DOI: 10.1111/j.1365-2966.2006.11017.x. arXiv: astro-ph/0604367 [astro-ph].

Habouzit, Mélanie, Marta Volonteri and Yohan Dubois (July 2017). 'Blossoms from black hole seeds: properties and early growth regulated by supernova feedback'. In: *MNRAS* 468.4, pp. 3935–3948. DOI: 10.1093/mnras/stx666. arXiv: 1605.09394 [astro-ph.GA].

Habouzit, Mélanie et al. (Nov. 2016). 'On the number density of 'direct collapse' black hole seeds'. In: *Monthly Notices of the Royal Astronomical*

Society 463. ADS Bibcode: 2016MNRAS.463..529H, 529–540. ISSN: 0035-8711. DOI: 10.1093/mnras/stw1924.

Haemmerlé, L. et al. (Apr. 2020). 'Formation of the First Stars and Black Holes'. In: *Space Science Reviews* 216.4, 48, p. 48. DOI: 10.1007/s11214-020-00673-y. arXiv: 2003.10533 [astro-ph.GA].

Harikane, Yuichi et al. (Mar. 2022). 'GOLDRUSH. IV. Luminosity Functions and Clustering Revealed with 4,000,000 Galaxies at z 2-7: Galaxy-AGN Transition, Star Formation Efficiency, and Implication for Evolution at z > 10'. In: *ApJ* 259.1, 20, p. 20. DOI: 10.3847/1538-4365/ac3dfc. arXiv: 2108.01090 [astro-ph.GA].

Harikane, Yuichi et al. (Dec. 2023). 'A JWST/NIRSpec First Census of Broad-line AGNs at z = 4-7: Detection of 10 Faint AGNs with M BH 106-108 $M_{\odot}$ and Their Host Galaxy Properties'. In: *ApJ* 959. ADS Bibcode: 2023ApJ...959...39H, p. 39. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ad029e.

Hayes, Matthew J. et al. (Mar. 2024). 'Glimmers in the Cosmic Dawn: A Census of the Youngest Supermassive Black Holes by Photometric Variability'. In: *arXiv e-prints*, arXiv:2403.16138, arXiv:2403.16138. DOI: 10.48550/arXiv.2403.16138. arXiv: 2403.16138 [astro-ph.GA].

Henriques, Bruno M. B. et al. (2015). 'Galaxy formation in the Planck cosmology – I. Matching the observed evolution of star formation rates, colours and stellar masses'. In: *MNRAS* 451.3, 2663–2680. ISSN: 0035-8711. DOI: 10.1093/mnras/stv705.

Hirschmann, Michaela, Gabriella De Lucia and Fabio Fontanot (2016). 'Galaxy assembly, stellar feedback and metal enrichment: the view from the GAEA model'. In: *MNRAS* 461. ADS Bibcode: 2016MNRAS.461.1760H, 1760–1785. ISSN: 0035-8711. DOI: 10.1093/mnras/stw1318.

Hirschmann, Michaela et al. (Aug. 2014). 'Cosmological simulations of black hole growth: AGN luminosities and downsizing'. In: *Monthly Notices of the Royal Astronomical Society* 442.3, pp. 2304–2324. DOI: 10.1093/mnras/stu1023. arXiv: 1308.0333 [astro-ph.CO].

Hosokawa, Takashi et al. (Dec. 2011). 'Protostellar Feedback Halts the Growth of the First Stars in the Universe'. In: *Science* 334.6060, p. 1250. DOI: 10.1126/science.1207433. arXiv: 1111.3649 [astro-ph.CO].

Hu, Weida et al. (Aug. 2024). 'Characterizing the Average Interstellar Medium Conditions of Galaxies at z ~ 5.6–9 with Ultraviolet and Optical Nebular Lines'. In: *ApJ* 971. ADS Bibcode: 2024ApJ...971...21H, p. 21. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ad5015.

Illingworth, Garth (July 2009). *Galaxies at z~7-10 in the Reionization Epoch: Luminosity Functions to <0.2L* from Deep IR Imaging of the HUDF and HUDF05 Fields*. HST Proposal ID 11563. Cycle 17.

Inayoshi, Kohei, Zoltán Haiman and Jeremiah P. Ostriker (July 2016). 'Hyper-Eddington accretion flows on to massive black holes'. In: *MNRAS* 459.4, 3738–3755. ISSN: 0035-8711. DOI: 10.1093/mnras/stw836.

Inayoshi, Kohei and Kohei Ichikawa (Feb. 2024). 'Birth of Rapidly Spinning, Overmassive Black Holes in the Early Universe'. In: *arXiv e-prints*,

arXiv:2402.14706, arXiv:2402.14706. DOI: 10.48550/arXiv.2402.14706. arXiv: 2402.14706 [astro-ph.GA].

Inayoshi, Kohei, Eli Visbal and Zoltán Haiman (Aug. 2020). 'The Assembly of the First Massive Black Holes'. In: *Annual Review of Astronomy and Astrophysics* 58. ADS Bibcode: 2020ARA&A..58...27I, 27–97. ISSN: 0066-4146. DOI: 10.1146/annurev-astro-120419-014455.

Ishigaki, Masafumi et al. (Feb. 2018). 'Full-data Results of Hubble Frontier Fields: UV Luminosity Functions at z ~ 6-10 and a Consistent Picture of Cosmic Reionization'. In: *ApJ* 854.1, 73, p. 73. DOI: 10.3847/1538-4357/aaa544. arXiv: 1702.04867 [astro-ph.GA].

Jeon, Junehyoung et al. (Apr. 2023). 'Observability of Low-Luminosity AGN in the Early Universe with JWST'. In: *arXiv e-prints*, arXiv:2304.07369, arXiv:2304.07369. DOI: 10.48550/arXiv.2304.07369. arXiv: 2304.07369 [astro-ph.GA].

Jiang, Yan-Fei, James M. Stone and Shane W. Davis (July 2019). 'Super-Eddington Accretion Disks around Supermassive Black Holes'. en. In: *ApJ* 880.2, p. 67. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab29ff.

Johnson, Jarrett L. and Volker Bromm (Feb. 2006). 'The cooling of shock-compressed primordial gas'. In: *MNRAS* 366.1, pp. 247–256. DOI: 10.1111/j.1365-2966.2005.09846.x. arXiv: astro-ph/0505304 [astro-ph].

— (Feb. 2007). 'The aftermath of the first stars: massive black holes'. In: *Monthly Notices of the Royal Astronomical Society* 374.4, pp. 1557–1568. DOI: 10.1111/j.1365-2966.2006.11275.x. arXiv: astro-ph/0605691 [astro-ph].

Kashikawa, Nobunari et al. (Jan. 2015). 'The Subaru High-z Quasar Survey: Discovery of Faint z ~6 Quasars'. In: *ApJ* 798.1, 28, p. 28. DOI: 10.1088/0004-637X/798/1/28. arXiv: 1410.7401 [astro-ph.GA].

Katz, Harley, Debora Sijacki and Martin G. Haehnelt (Aug. 2015). 'Seeding high-redshift QSOs by collisional runaway in primordial star clusters'. In: *Monthly Notices of the Royal Astronomical Society* 451.3, pp. 2352–2369. DOI: 10.1093/mnras/stv1048. arXiv: 1502.03448 [astro-ph.GA].

Kerr, Roy P. (Sept. 1963). 'Gravitational Field of a Spinning Mass as an Example of Algebraically Special Metrics'. In: *Physical Review Letters* 11.5, pp. 237–238. DOI: 10.1103/PhysRevLett.11.237.

Kim, Yongjung et al. (Dec. 2020). 'The Infrared Medium-deep Survey. VIII. Quasar Luminosity Function at z ~ 5'. In: *ApJ* 904.2, 111, p. 111. DOI: 10.3847/1538-4357/abc0ea. arXiv: 2010.05859 [astro-ph.GA].

Kocevski, Dale D. et al. (Sept. 2023). 'Hidden Little Monsters: Spectroscopic Identification of Low-mass, Broad-line AGNs at z > 5 with CEERS'. In: *ApJ* 954. ADS Bibcode: 2023ApJ...954L...4K, p. L4. ISSN: 0004-637X. DOI: 10.3847/2041-8213/ace5a0.

Kocevski, Dale D. et al. (Apr. 2024). 'The Rise of Faint, Red AGN at z > 4: A Sample of Little Red Dots in the JWST Extragalactic Legacy Fields'. In: *arXiv e-prints*, arXiv:2404.03576, arXiv:2404.03576. DOI: 10.48550/arXiv.2404.03576. arXiv: 2404.03576 [astro-ph.GA].

Kohri, Kazunori, Toyokazu Sekiguchi and Sai Wang (Aug. 2022). 'Cosmological 21-cm line observations to test scenarios of super-Eddington

accretion on to black holes being seeds of high-redshifted supermassive black holes'. In: *Physical Review D* 106.4, 043539, p. 043539. DOI: 10.1103/PhysRevD.106.043539. arXiv: 2201.05300 [astro-ph.CO].

Kokorev, Vasily et al. (June 2024). 'A Census of Photometrically Selected Little Red Dots at 4 < z < 9 in JWST Blank Fields'. In: *ApJ* 968. ADS Bibcode: 2024ApJ...968...38K, p. 38. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ad4265.

Kokubo, Mitsuru and Yuichi Harikane (July 2024). 'Challenging the AGN scenario for JWST/NIRSpec broad Hα emitters/Little Red Dots in light of non-detection of NIRCam photometric variability and X-ray'. In: *arXiv e-prints*, arXiv:2407.04777, arXiv:2407.04777. DOI: 10.48550/arXiv.2407.04777. arXiv: 2407.04777 [astro-ph.GA].

Kormendy, John and Luis C. Ho (Aug. 2013). 'Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies'. en. In: *Annual Review of Astronomy and Astrophysics* 51.Volume 51, 2013, 511–653. ISSN: 0066-4146, 1545-4282. DOI: 10.1146/annurev-astro-082708-101811.

Kroupa, Pavel et al. (2020). 'Very high redshift quasars and the rapid emergence of super-massive black holes'. In: *MNRAS*. arXiv:2007.14402, staa2276. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/staa2276.

LIGO Scientific Collaboration et al. (Apr. 2015). 'Advanced LIGO'. In: *Classical and Quantum Gravity* 32.7, 074001, p. 074001. DOI: 10.1088/0264-9381/32/7/074001. arXiv: 1411.4547 [gr-qc].

Labbe, Ivo et al. (June 2023). 'UNCOVER: Candidate Red Active Galactic Nuclei at 3<z<7 with JWST and ALMA'. In: *arXiv e-prints*, arXiv:2306.07320, arXiv:2306.07320. DOI: 10.48550/arXiv.2306.07320. arXiv: 2306.07320 [astro-ph.GA].

Lacey, Cedric and Shaun Cole (June 1993). 'Merger rates in hierarchical models of galaxy formation'. In: *Monthly Notices of the Royal Astronomical Society* 262. ADS Bibcode: 1993MNRAS.262..627L, 627–649. ISSN: 0035-8711. DOI: 10.1093/mnras/262.3.627.

Laplace, Pierre-Simon (1796). *Exposition Du Systéme Du Monde. Tome 2*. A Paris, de l'imprimerie du Cercle-Social, rue du Théâtre-Français, n° 4. L'An IV de la République française.

Laporte, Nicolas et al. (Dec. 2017). 'A Spectroscopic Search for AGN Activity in the Reionization Era'. en. In: *ApJ* 851, p. 40. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aa96a8.

Larson, Rebecca L. et al. (Aug. 2023). 'A CEERS Discovery of an Accreting Supermassive Black Hole 570 Myr after the Big Bang: Identifying a Progenitor of Massive z > 6 Quasars'. en. In: *ApJ* 953, p. L29. ISSN: 2041-8205. DOI: 10.3847/2041-8213/ace619.

Latif, M. A. et al. (Aug. 2013). 'Black hole formation in the early Universe'. In: *Monthly Notices of the Royal Astronomical Society* 433.2, pp. 1607–1618. DOI: 10.1093/mnras/stt834. arXiv: 1304.0962 [astro-ph.CO].

Li, Junyao et al. (Feb. 2024). 'Tip of the iceberg: overmassive black holes at 4'. In: *arXiv e-prints*. ADS Bibcode: 2024arXiv240300074L. DOI: 10.48550/arXiv.2403.00074. URL: https://ui.adsabs.harvard.edu/abs/2024arXiv240300074L.

Lodato, Giuseppe and Priyamvada Natarajan (Oct. 2006). 'Supermassive black hole formation during the assembly of pre-galactic discs'. In: *Monthly Notices of the Royal Astronomical Society* 371.4, pp. 1813–1823. DOI: 10.1111/j.1365-2966.2006.10801.x. arXiv: astro-ph/0606159 [astro-ph].

Lusso, Elisabeta, Rosa Valiante and Fabio Vito (2023). 'The Dawn of Black Holes'. In: *Handbook of X-ray and Gamma-ray Astrophysics. Edited by Cosimo Bambi and Andrea Santangelo*, 122, p. 122. DOI: 10.1007/978-981-16-4544-0_116-1.

Lyu, Jianwei and George Rieke (2022). 'Infrared Spectral Energy Distribution and Variability of Active Galactic Nuclei: Clues to the Structure of Circumnuclear Material'. In: *Universe* 8.6. ISSN: 2218-1997. DOI: 10.3390/universe8060304. URL: https://www.mdpi.com/2218-1997/8/6/304.

Lyu, Jianwei et al. (2022). 'AGN Selection and Demographics in GOODS-S/HUDF from X-Ray to Radio'. en. In: *ApJ* 941.2, p. 191. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac9e5d.

Madau, Piero and Martin J. Rees (Apr. 2001). 'Massive Black Holes as Population III Remnants'. In: *ApJ* 551. ADS Bibcode: 2001ApJ...551L..27M, L27–L30. ISSN: 0004-637X. DOI: 10.1086/319848.

Madau, Piero et al. (Aug. 2024). 'Cosmic Reionization in the JWST Era: Back to AGNs?' In: *ApJ* 971.1, 75, p. 75. DOI: 10.3847/1538-4357/ad5ce8. arXiv: 2406.18697 [astro-ph.CO].

Mainali, Ramesh et al. (Sept. 2018). 'Spectroscopic constraints on UV metal line emission at z $\simeq$ 6–9: the nature of Lyα emitting galaxies in the reionization era'. In: *MNRAS* 479, 1180–1193. ISSN: 0035-8711. DOI: 10.1093/mnras/sty1640.

Maio, Umberto et al. (May 2019). 'The seeds of supermassive black holes and the role of local radiation and metal spreading'. In: *Publications of the Astronomical Society of Australia* 36, e020, e020. DOI: 10.1017/pasa.2019.10. arXiv: 1811.01964 [astro-ph.GA].

Maiolino, Roberto et al. (Aug. 2023). 'JADES. The diverse population of infant Black Holes at 4<z<11: merging, tiny, poor, but mighty'. In: *arXiv e-prints*, arXiv:2308.01230, arXiv:2308.01230. DOI: 10.48550/arXiv.2308.01230. arXiv: 2308.01230 [astro-ph.GA].

Maiolino, Roberto et al. (2023). 'JADES. The diverse population of infant Black Holes at 4<z<11: merging, tiny, poor, but mighty'. In: *arXiv e-prints* arXiv:2308.01230. arXiv:2308.01230. DOI: 10.48550/arXiv.2308.01230. URL: http://arxiv.org/abs/2308.01230.

Maiolino, Roberto et al. (Mar. 2024). 'A small and vigorous black hole in the early Universe'. en. In: *Nature* 627.8002, 59–63. ISSN: 1476-4687. DOI: 10.1038/s41586-024-07052-5.

Maiolino, Roberto et al. (May 2024). 'JWST meets Chandra: a large population of Compton thick, feedback-free, and X-ray weak AGN, with a sprinkle of SNe'. In: *arXiv e-prints*, arXiv:2405.00504, arXiv:2405.00504. DOI: 10.48550/arXiv.2405.00504. arXiv: 2405.00504 [astro-ph.GA].

Marconi, Alessandro and Leslie K. Hunt (May 2003). 'The Relation between Black Hole Mass, Bulge Mass, and Near-Infrared Luminosity'. In: *ApJ*

Letter 589.1, pp. L21–L24. DOI: 10.1086/375804. arXiv: astro-ph/0304274 [astro-ph].

Matsuoka, Yoshiki et al. (Dec. 2018). 'Subaru High-z Exploration of Low-luminosity Quasars (SHELLQs). V. Quasar Luminosity Function and Contribution to Cosmic Reionization at z = 6'. In: *ApJ* 869.2, 150, p. 150. DOI: 10.3847/1538-4357/aaee7a. arXiv: 1811.01963 [astro-ph.GA].

Matteo, Tiziana Di et al. (Mar. 2008). 'Direct Cosmological Simulations of the Growth of Black Holes and Galaxies'. en. In: *ApJ* 676.1, p. 33. ISSN: 0004-637X. DOI: 10.1086/524921.

Matthee, Jorryt et al. (June 2023). 'EIGER. II. First Spectroscopic Characterization of the Young Stars and Ionized Gas Associated with Strong Hβ and [O III] Line Emission in Galaxies at z = 5-7 with JWST'. In: *ApJ* 950.1, 67, p. 67. DOI: 10.3847/1538-4357/acc846. arXiv: 2211.08255 [astro-ph.GA].

Matthee, Jorryt et al. (Mar. 2024). 'Little Red Dots: An Abundant Population of Faint Active Galactic Nuclei at z ~ 5 Revealed by the EIGER and FRESCO JWST Surveys'. In: *ApJ* 963. ADS Bibcode: 2024ApJ...963..129M, p. 129. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ad2345.

McKee, Christopher F. and Jonathan C. Tan (2008a). 'The Formation of the First Stars. II. Radiative Feedback Processes and Implications for the Initial Mass Function'. In: *The Astrophysical Journal* 681.2, pp. 771–797. DOI: 10.1086/587434. URL: https://doi.org/10.1086/587434.

— (July 2008b). 'The Formation of the First Stars. II. Radiative Feedback Processes and Implications for the Initial Mass Function'. en. In: *ApJ* 681.2, p. 771. ISSN: 0004-637X. DOI: 10.1086/587434.

McLeod, D. J., R. J. McLure and J. S. Dunlop (July 2016). 'The z = 9-10 galaxy population in the Hubble Frontier Fields and CLASH surveys: the z = 9 luminosity function and further evidence for a smooth decline in ultraviolet luminosity density at z⩾ 8'. In: *MNRAS* 459.4, pp. 3812–3824. DOI: 10.1093/mnras/stw904. arXiv: 1602.05199 [astro-ph.GA].

McLure, R. J. et al. (July 2013). 'A new multifield determination of the galaxy luminosity function at z = 7-9 incorporating the 2012 Hubble Ultra-Deep Field imaging'. In: *MNRAS* 432.4, pp. 2696–2716. DOI: 10.1093/mnras/stt627. arXiv: 1212.5222 [astro-ph.CO].

Michell, John (1784). 'On the means of discovering the distance, magnitude, etc. of the fixed stars, in consequence of the diminution of the velocity of their light, in case such a diminution should be found to take place in any of them, and such other data should be procured from observations, as would be farther necessary for that purpose.' In: *Philosophical Transactions of the Royal Society* 74, pp. 35–57. DOI: 10.1098/rstl.1784.0008.

Miller, Brendan P. et al. (Jan. 2015). 'X-Ray Constraints on the Local Supermassive Black Hole Occupation Fraction'. In: *ApJ* 799. ADS Bibcode: 2015ApJ...799...98M, p. 98. ISSN: 0004-637X. DOI: 10.1088/0004-637X/799/1/98.

Milosavljević, Milos et al. (June 2009). 'Accretion onto "Seed" Black Holes in the First Galaxies'. In: *The Astrophysical Journal* 698.1, pp. 766–780. DOI: 10.1088/0004-637X/698/1/766. arXiv: 0809.2404 [astro-ph].

Monaco, Pierluigi and Fabio Fontanot (May 2005). 'Feedback from quasars in star-forming galaxies and the triggering of massive galactic winds'. In: *MNRAS* 359.1, 283–294. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2005.08884.x.

Monaco, Pierluigi, Tom Theuns and Giuliano Taffoni (2002). 'The pinocchio algorithm: pinpointing orbit-crossing collapsed hierarchical objects in a linear density field'. en. In: *MNRAS* 331.3, 587–608. ISSN: 00358711, 13652966. DOI: 10.1046/j.1365-8711.2002.05162.x.

Montero, Pedro J., Hans-Thomas Janka and Ewald Müller (Mar. 2012). 'RELATIVISTIC COLLAPSE AND EXPLOSION OF ROTATING SUPERMASSIVE STARS WITH THERMONUCLEAR EFFECTS'. en. In: *ApJ* 749.1, p. 37. ISSN: 0004-637X. DOI: 10.1088/0004-637X/749/1/37.

Morishita, T. et al. (Nov. 2018). 'The Bright-end Galaxy Candidates at z ~ 9 from 79 Independent HST Fields'. In: *ApJ* 867.2, 150, p. 150. DOI: 10.3847/1538-4357/aae68c. arXiv: 1809.07604 [astro-ph.GA].

Moriya, Takashi J. et al. (Apr. 2019). 'First Release of High-Redshift Superluminous Supernovae from the Subaru HIgh-Z SUpernova CAmpaign (SHIZUCA). I. Photometric Properties'. In: *ApJ, Supplement* 241.2, 16, p. 16. DOI: 10.3847/1538-4365/ab07c5. arXiv: 1801.08240 [astro-ph.HE].

Mortlock, Daniel J. et al. (June 2011). 'A luminous quasar at a redshift of z = 7.085'. In: *Nature* 474.7353, pp. 616–619. DOI: 10.1038/nature10159. arXiv: 1106.6088 [astro-ph.CO].

Munari, Emiliano et al. (2017). 'Improving fast generation of halo catalogues with higher order Lagrangian perturbation theory'. In: *MNRAS* 465. ADS Bibcode: 2017MNRAS.465.4658M, 4658–4677. ISSN: 0035-8711. DOI: 10.1093/mnras/stw3085.

Mutlu-Pakdil, Burçin, Marc S. Seigar and Benjamin L. Davis (Oct. 2016). 'THE LOCAL BLACK HOLE MASS FUNCTION DERIVED FROM THE MBH–P AND THE MBH–n RELATIONS'. en. In: *ApJ* 830.2, p. 117. ISSN: 0004-637X. DOI: 10.3847/0004-637X/830/2/117.

Natarajan, Aravind, Jonathan C. Tan and Brian W. O'Shea (Feb. 2009). 'DARK MATTER ANNIHILATION AND PRIMORDIAL STAR FORMATION'. en. In: *ApJ* 692.1, p. 574. ISSN: 0004-637X. DOI: 10.1088/0004-637X/692/1/574.

Navarro, Julio F., Carlos S. Frenk and Simon D. M. White (Dec. 1997). 'A Universal Density Profile from Hierarchical Clustering'. en. In: *ApJ* 490.2, p. 493. ISSN: 0004-637X. DOI: 10.1086/304888.

Newman, E. T. and A. I. Janis (June 1965). 'Note on the Kerr Spinning-Particle Metric'. In: *Journal of Mathematical Physics* 6.6, pp. 915–917. DOI: 10.1063/1.1704350.

Newman, E. T. et al. (June 1965). 'Metric of a Rotating, Charged Mass'. In: *Journal of Mathematical Physics* 6.6, pp. 918–919. DOI: 10.1063/1.1704351.

Nguyen, Dieu D. et al. (Feb. 2019). 'Improved Dynamical Constraints on the Masses of the Central Black Holes in Nearby Low-mass Early-type Galactic Nuclei and the First Black Hole Determination for NGC 205'. In: *ApJ* 872. ADS Bibcode: 2019ApJ...872..104N, p. 104. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aafe7a.

Ni, Yueying et al. (June 2022). 'The ASTRID simulation: the evolution of supermassive black holes'. In: *MNRAS* 513.1, pp. 670–692. DOI: 10.1093/mnras/stac351. arXiv: 2110.14154 [astro-ph.GA].

Niida, Mana et al. (Dec. 2020). 'The Faint End of the Quasar Luminosity Function at z ~ 5 from the Subaru Hyper Suprime-Cam Survey'. In: *ApJ* 904.2, 89, p. 89. DOI: 10.3847/1538-4357/abbe11. arXiv: 2010.00481 [astro-ph.GA].

O'Shea, Brian W. and Michael L. Norman (Jan. 2007). 'Population III Star Formation in a ΛCDM Universe. I. The Effect of Formation Redshift and Environment on Protostellar Accretion Rate'. In: *The Astrophysical Journal* 654.1, pp. 66–92. DOI: 10.1086/509250. arXiv: astro-ph/0607013 [astro-ph].

O'Shea, Brian W. et al. (July 2005). 'Forming a Primordial Star in a Relic H II Region'. In: *The Astrophysical Journal Letters* 628.1, pp. L5–L8. DOI: 10.1086/432683. arXiv: astro-ph/0503330 [astro-ph].

Oesch, P. A. et al. (Oct. 2023). 'The JWST FRESCO survey: legacy NIRCam/-grism spectroscopy and imaging in the two GOODS fields'. In: *MNRAS* 525.2, pp. 2864–2874. DOI: 10.1093/mnras/stad2411. arXiv: 2304.02026 [astro-ph.GA].

Ono, Yoshiaki et al. (Jan. 2018). 'Great Optically Luminous Dropout Research Using Subaru HSC (GOLDRUSH). I. UV luminosity functions at z ~ 4-7 derived with the half-million dropouts on the 100 deg^2 sky'. In: *PASJ* 70, S10, S10. DOI: 10.1093/pasj/psx103. arXiv: 1704.06004 [astro-ph.GA].

Onoue, Masafusa et al. (Oct. 2017). 'Minor Contribution of Quasars to Ionizing Photon Budget at z ~ 6: Update on Quasar Luminosity Function at the Faint End with Subaru/Suprime-Cam'. In: *ApJ Letter* 847.2, L15, p. L15. DOI: 10.3847/2041-8213/aa8cc6. arXiv: 1709.04413 [astro-ph.GA].

O'Brien, Rosalia et al. (Jan. 2024). 'TREASUREHUNT: Transients and Variability Discovered with HST in the JWST North Ecliptic Pole Time Domain Field'. In: *arXiv e-prints*. ADS Bibcode: 2024arXiv240104944O. DOI: 10.48550/arXiv.2401.04944. URL: https://ui.adsabs.harvard.edu/abs/2024arXiv240104944O.

Pacucci, Fabio et al. (Nov. 2023). 'JWST CEERS and JADES Active Galaxies at z = 4-7 Violate the Local M_{BH}-$M_{\star}$ Relation at >3σ: Implications for Low-mass Black Holes and Seeding Models'. In: *ApJ* 957.1, L3, p. L3. DOI: 10.3847/2041-8213/ad0158. arXiv: 2308.12331 [astro-ph.GA].

Panter, Benjamin et al. (2007). 'The star formation histories of galaxies in the Sloan Digital Sky Survey'. In: *MNRAS* 378. ADS Bibcode: 2007MNRAS.378.1550P, 1550–1564. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2007.11909.x.

Parsa, Shaghayegh, James S. Dunlop and Ross J. McLure (Mar. 2018). 'No evidence for a significant AGN contribution to cosmic hydrogen reionization'. In: *MNRAS* 474.3, pp. 2904–2923. DOI: 10.1093/mnras/stx2887. arXiv: 1704.07750 [astro-ph.GA].

Pelló, R. et al. (Dec. 2018). 'The WIRCam Ultra Deep Survey (WUDS). I. Survey overview and UV luminosity functions at z ~ 5 and z ~ 6'.

In: *A&A* 620, A51, A51. DOI: 10.1051/0004-6361/201732080. arXiv: 1809.03373 [astro-ph.GA].

Pérez-González, Pablo G. et al. (June 2024). 'What Is the Nature of Little Red Dots and what Is Not, MIRI SMILES Edition'. In: *ApJ* 968.1, 4, p. 4. DOI: 10.3847/1538-4357/ad38bb. arXiv: 2401.08782 [astro-ph.GA].

Planck Collaboration et al. (Sept. 2020). 'Planck 2018 results. VI. Cosmological parameters'. In: *Astronomy & Astrophysics* 641, A6, A6. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].

Pontoppidan, Klaus M. et al. (Sept. 2022). 'The JWST Early Release Observations'. en. In: *The Astrophysical Journal Letters* 936.1, p. L14. ISSN: 2041-8205. DOI: 10.3847/2041-8213/ac8a4e.

Portegies Zwart, Simon F. et al. (Apr. 2004). 'Formation of massive black holes through runaway collisions in dense young star clusters'. In: *Nature* 428.6984, pp. 724–726. DOI: 10.1038/nature02448. arXiv: astro-ph/0402622 [astro-ph].

Pouliasis, E et al. (Aug. 2019). 'Robust identification of active galactic nuclei through HST optical variability in GOODS-S: comparison with the X-ray and mid-IR-selected samples'. In: *MNRAS* 487.3, 4285–4304. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1483.

Press, William H. and Paul Schechter (Feb. 1974). 'Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation'. In: *ApJ* 187. ADS Bibcode: 1974ApJ...187..425P, 425–438. ISSN: 0004-637X. DOI: 10.1086/152650.

Pérez-González, Pablo G. et al. (Mar. 2008). 'The Stellar Mass Assembly of Galaxies from z = 0 to z = 4: Analysis of a Sample Selected in the Rest-Frame Near-Infrared with Spitzer'. In: *ApJ* 675, 234–261. ISSN: 0004-637X. DOI: 10.1086/523690.

Rafelski, Marc et al. (July 2015). 'UVUDF: Ultraviolet Through Near-infrared Catalog and Photometric Redshifts of Galaxies in the Hubble Ultra Deep Field'. In: *The Astron. J.* 150.1, 31, p. 31. DOI: 10.1088/0004-6256/150/1/31. arXiv: 1505.01160 [astro-ph.GA].

Reardon, Daniel J. et al. (July 2023). 'Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array'. In: *The Astrophysical Journal Letters* 951.1, L6, p. L6. DOI: 10.3847/2041-8213/acdd02. arXiv: 2306.16215 [astro-ph.HE].

Rees, M. J. (Oct. 1978). 'Quasars'. In: *The Observatory* 98, pp. 210–223.

Regan, John A. et al. (Dec. 2020). 'The Formation of Very Massive Stars in Early Galaxies and Implications for Intermediate Mass Black Holes'. In: *The Open Journal of Astrophysics* 3.1, 15, p. 15. DOI: 10.21105/astro.2008.08090. arXiv: 2008.08090 [astro-ph.GA].

Reines, Amy E. and Marta Volonteri (Oct. 2015). 'RELATIONS BETWEEN CENTRAL BLACK HOLE MASS AND TOTAL GALAXY STELLAR MASS IN THE LOCAL UNIVERSE'. en. In: *The Astrophysical Journal* 813.2, p. 82. ISSN: 0004-637X. DOI: 10.1088/0004-637X/813/2/82.

Ricarte, Angelo and Priyamvada Natarajan (Feb. 2018). 'Exploring SMBH assembly with semi-analytic modelling'. In: *MNRAS* 474. ADS Bib-

code: 2018MNRAS.474.1995R, 1995–2011. ISSN: 0035-8711. DOI: 10.1093/mnras/stx2851.

Rieke, Marcia J. et al. (Nov. 2023). 'JADES Initial Data Release for the Hubble Ultra Deep Field: Revealing the Faint Infrared Sky with Deep JWST NIRCam Imaging'. In: *ApJ, Supplement* 269.1, 16, p. 16. DOI: 10.3847/1538-4365/acf44d. arXiv: 2306.02466 [astro-ph.GA].

Rindler-Daller, T. et al. (Jan. 2015). 'DARK STARS: IMPROVED MODELS AND FIRST PULSATION RESULTS'. en. In: *ApJ* 799.2, p. 210. ISSN: 0004-637X. DOI: 10.1088/0004-637X/799/2/210.

Robertson, Brant E. et al. (Apr. 2015). 'Cosmic Reionization and Early Star-forming Galaxies: A Joint Analysis of New Constraints from Planck and the Hubble Space Telescope'. In: *ApJ* 802.2, L19, p. L19. DOI: 10.1088/2041-8205/802/2/L19. arXiv: 1502.02024 [astro-ph.CO].

Rojas-Ruiz, Sofía et al. (Mar. 2020). 'Probing the Bright End of the Rest-frame Ultraviolet Luminosity Function at z = 8-10 with Hubble Pure-parallel Imaging'. In: *ApJ* 891.2, 146, p. 146. DOI: 10.3847/1538-4357/ab7659. arXiv: 2002.06209 [astro-ph.GA].

STSCI Development Team (Dec. 2012). *DrizzlePac: HST image software*. Astrophysics Source Code Library, record ascl:1212.011.

Salucci, Paolo et al. (Sept. 2000). 'The masses of black holes in the nuclei of spirals'. In: *MNRAS* 317.2, pp. 488–496. DOI: 10.1046/j.1365-8711.2000.03622.x. arXiv: astro-ph/9812485 [astro-ph].

Sassano, Federica et al. (Sept. 2021). 'Light, medium-weight, or heavy? The nature of the first supermassive black hole seeds'. In: *MNRAS* 506.1, 613–632. ISSN: 0035-8711. DOI: 10.1093/mnras/stab1737.

Saxena, A. et al. (Apr. 2020). 'The properties of He IIλ1640 emitters at z ∼ 2.5–5 from the VANDELS survey'. en. In: *A&A* 636, A47. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201937170.

Schechter, P. (Jan. 1976). 'An analytic expression for the luminosity function for galaxies.' In: *ApJ* 203. ADS Bibcode: 1976ApJ...203..297S, 297–306. ISSN: 0004-637X. DOI: 10.1086/154079.

Schenker, Matthew A. et al. (May 2013). 'The UV Luminosity Function of Star-forming Galaxies via Dropout Selection at Redshifts z ∼7 and 8 from the 2012 Ultra Deep Field Campaign'. In: *ApJ* 768.2, 196, p. 196. DOI: 10.1088/0004-637X/768/2/196. arXiv: 1212.4819 [astro-ph.CO].

Schwarzschild, Karl (Jan. 1916). 'Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie'. In: *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pp. 189–196.

Senchyna, Peter et al. (Sept. 2019). 'Extremely metal-poor galaxies with HST/COS: laboratories for models of low-metallicity massive stars and high-redshift galaxies'. In: *MNRAS* 488, 3492–3506. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1907.

Shang, Cien, Greg L. Bryan and Z. Haiman (Feb. 2010). 'Supermassive black hole formation by direct collapse: keeping protogalactic gas H_2 free in dark matter haloes with virial temperatures T_{vir} > rsim 10^4 K'. In: *Monthly Notices of the Royal Astronomical Society* 402.2, pp. 1249–

1262. DOI: 10.1111/j.1365-2966.2009.15960.x. arXiv: 0906.4773 [astro-ph.CO].

Shankar, Francesco et al. (Jan. 2020). 'Constraining black hole-galaxy scaling relations and radiative efficiency from galaxy clustering'. In: *Nature Astronomy* 4, pp. 282–291. DOI: 10.1038/s41550-019-0949-y. arXiv: 1910.10175 [astro-ph.GA].

Shen, Xuejian et al. (Jan. 2020). 'The bolometric quasar luminosity function at z = 0-7'. In: *MNRAS* 495.3, pp. 3252–3275. DOI: 10.1093/mnras/staa1381. arXiv: 2001.02696 [astro-ph.GA].

Shin, Suhyun et al. (Apr. 2020). 'The Infrared Medium-deep Survey. VII. Faint Quasars at z ~ 5 in the ELAIS-N1 Field'. In: *ApJ* 893.1, 45, p. 45. DOI: 10.3847/1538-4357/ab7bde. arXiv: 2003.00709 [astro-ph.GA].

Shinohara, Takumi et al. (Apr. 2023). 'Supermassive primordial black holes: a view from clustering of quasars at z ~ 6'. In: *arXiv e-prints*, arXiv:2304.08153, arXiv:2304.08153. DOI: 10.48550/arXiv.2304.08153. arXiv: 2304.08153 [astro-ph.CO].

Sijacki, Debora et al. (Sept. 2007). 'A unified model for AGN feedback in cosmological simulations of structure formation'. In: *MNRAS* 380.3, 877–900. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2007.12153.x.

Singh, Jasbir, Pierluigi Monaco and Jonathan C Tan (2023). 'The formation of supermassive black holes from Population III.1 seeds. II. Evolution to the local universe'. In: *MNRAS* 525.1, 969–982. ISSN: 0035-8711. DOI: 10.1093/mnras/stad2346.

Sądowski, Aleksander and Massimo Gaspari (June 2017). 'Kinetic and radiative power from optically thin accretion flows'. In: *MNRAS* 468.2, pp. 1398–1404. DOI: 10.1093/mnras/stx543. arXiv: 1701.07033 [astro-ph.HE].

Somerville, Rachel S. and Romeel Davé (2015). 'Physical Models of Galaxy Formation in a Cosmological Framework'. In: *Annual Review of Astronomy and Astrophysics* 53.1. arXiv: 1412.2712, 51–113. ISSN: 0066-4146, 1545-4282. DOI: 10.1146/annurev-astro-082812-140951.

Somerville, Rachel S., Joel R. Primack and S. M. Faber (Feb. 2001). 'The nature of high-redshift galaxies'. In: *MNRAS* 320.4, 504–528. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2001.03975.x.

Spinoso, Daniele et al. (2022). 'Multi-flavour SMBH seeding and evolution in cosmological environments'. In: *arXiv:2203.13846*. arXiv: 2203.13846. URL: http://arxiv.org/abs/2203.13846.

Spolyar, Douglas, Katherine Freese and Paolo Gondolo (Feb. 2008). 'Dark Matter and the First Stars: A New Phase of Stellar Evolution'. In: *Physical Review Letters* 100.5, p. 051101. DOI: 10.1103/PhysRevLett.100.051101.

Springel, Volker, Tiziana Di Matteo and Lars Hernquist (Aug. 2005). 'Modelling feedback from stars and black holes in galaxy mergers'. In: *Monthly Notices of the Royal Astronomical Society* 361.3, pp. 776–794. DOI: 10.1111/j.1365-2966.2005.09238.x. arXiv: astro-ph/0411108 [astro-ph].

Stefanon, Mauro et al. (Sept. 2019). 'The Brightest z $\gtrsim$ 8 Galaxies over the COSMOS UltraVISTA Field'. In: *ApJ* 883.1, 99, p. 99. DOI: 10.3847/1538-4357/ab3792. arXiv: 1902.10713 [astro-ph.GA].

Stern, Daniel et al. (Sept. 2005). 'Mid-Infrared Selection of Active Galaxies'. In: *ApJ* 631. ADS Bibcode: 2005ApJ...631..163S, 163–168. ISSN: 0004-637X. DOI: 10.1086/432523.

Stevans, Matthew L. et al. (Aug. 2018). 'Bridging Star-forming Galaxy and AGN Ultraviolet Luminosity Functions at z = 4 with the SHELA Wide-field Survey'. In: *ApJ* 863.1, 63, p. 63. DOI: 10.3847/1538-4357/aacbd7. arXiv: 1806.05187 [astro-ph.GA].

Sun, Lei et al. (Sept. 2024). 'The Ultraviolet Luminosity Function at 0.6 < z < 1 from UVCANDELS'. In: *ApJ* 972.1, 8, p. 8. DOI: 10.3847/1538-4357/ad5540. arXiv: 2311.15664 [astro-ph.GA].

Susa, Hajime, Kenji Hasegawa and Nozomu Tominaga (Sept. 2014). 'The Mass Spectrum of the First Stars'. In: *The Astrophysical Journal* 792.1, 32, p. 32. DOI: 10.1088/0004-637X/792/1/32. arXiv: 1407.1374 [astro-ph.GA].

Tagawa, Hiromichi, Zoltan Haiman and Bence Kocsis (Mar. 2020). 'Making a Supermassive Star by Stellar Bombardment'. In: *The Astrophysical Journal* 892.1, 36, p. 36. DOI: 10.3847/1538-4357/ab7922. arXiv: 1909.10517 [astro-ph.GA].

Takeo, Eishun et al. (Sept. 2019). 'Super-Eddington growth of black holes in the early universe: effects of disc radiation spectra'. In: *MNRAS* 488.2, 2689–2700. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1899.

Tan, Jonathan C. and Christopher F. McKee (Mar. 2004). 'The Formation of the First Stars. I. Mass Infall Rates, Accretion Disk Structure, and Protostellar Evolution'. In: *The Astrophysical Journal* 603.2, pp. 383–400. DOI: 10.1086/381490. arXiv: astro-ph/0307414 [astro-ph].

Tan, Jonathan C., Britton D. Smith and Brian W. O'Shea (2010). 'Protostellar Feedback Processes and the Mass of the First Stars'. In: *AIP Conference Proceedings* 1294.1, pp. 34–40. DOI: 10.1063/1.3518887. eprint: https://aip.scitation.org/doi/pdf/10.1063/1.3518887. URL: https://aip.scitation.org/doi/abs/10.1063/1.3518887.

Trebitsch, Maxime et al. (Sept. 2021). 'The OBELISK simulation: Galaxies contribute more than AGN to H I reionization of protoclusters'. en. In: *A&A* 653, A154. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202037698.

Tremmel, M. et al. (Sept. 2017). 'The Romulus cosmological simulations: a physical approach to the formation, dynamics and accretion models of SMBHs'. In: *MNRAS* 470.1, 1121–1139. ISSN: 0035-8711. DOI: 10.1093/mnras/stx1160.

Trinca, Alessandro et al. (Mar. 2022). 'The low-end of the black hole mass function at cosmic dawn'. In: *MNRAS* 511. ADS Bibcode: 2022MNRAS.511..616T, 616–640. ISSN: 0035-8711. DOI: 10.1093/mnras/stac062.

Ueda, Yoshihiro et al. (May 2014). 'Toward the Standard Population Synthesis Model of the X-Ray Background: Evolution of X-Ray Luminosity and Absorption Functions of Active Galactic Nuclei Including Compton-thick Populations'. In: *ApJ* 786.2, 104, p. 104. DOI: 10.1088/0004-637X/786/2/104. arXiv: 1402.1836 [astro-ph.CO].

Van Den Bosch, Frank C. et al. (June 2008). 'The importance of satellite quenching for the build-up of the red sequence of present-day galaxies'. In: *MNRAS* 387.1, 79–91. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2008.13230.x.

Vika, Marina et al. (Dec. 2009). 'The Millennium Galaxy Catalogue: the M_{bh}-$L_{spheroid}$ derived supermassive black hole mass function'. In: *MNRAS* 400.3, pp. 1451–1460. DOI: 10.1111/j.1365-2966.2009.15544.x. arXiv: 0908.2102 [astro-ph.CO].

Vogelsberger, Mark et al. (Oct. 2014). 'Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe'. In: *MNRAS* 444.2, 1518–1547. ISSN: 0035-8711. DOI: 10.1093/mnras/stu1536.

Volonteri, M. et al. (Aug. 2016). 'The cosmic evolution of massive black holes in the Horizon-AGN simulation'. In: *MNRAS* 460.3, 2979–2996. ISSN: 0035-8711. DOI: 10.1093/mnras/stw1123.

Volonteri, Marta (July 2010). 'Formation of supermassive black holes'. In: *The Astronomy and Astrophysics Review* 18.3, pp. 279–315. DOI: 10.1007/s00159-010-0029-x. arXiv: 1003.4404 [astro-ph.CO].

Volonteri, Marta, Mélanie Habouzit and Monica Colpi (2021). 'The origins of massive black holes'. en. In: *Nature Reviews Physics* 3.1111, 732–743. ISSN: 2522-5820. DOI: 10.1038/s42254-021-00364-9.

Volonteri, Marta, Priyamvada Natarajan and Kayhan Gültekin (2011). 'HOW IMPORTANT IS THE DARK MATTER HALO FOR BLACK HOLE GROWTH?' en. In: *ApJ* 737.2, p. 50. ISSN: 0004-637X. DOI: 10.1088/0004-637X/737/2/50.

Wang, Feige et al. (Oct. 2019). 'Exploring Reionization-era Quasars. III. Discovery of 16 Quasars at $6.4 \lesssim z \lesssim 6.9$ with DESI Legacy Imaging Surveys and the UKIRT Hemisphere Survey and Quasar Luminosity Function at $z \sim 6.7$'. In: *The Astrophysical Journal* 884.1, 30, p. 30. DOI: 10.3847/1538-4357/ab2be5. arXiv: 1810.11926 [astro-ph.GA].

Wang, Feige et al. (Jan. 2021). 'A Luminous Quasar at Redshift 7.642'. In: *ApJ* 907.1, L1, p. L1. DOI: 10.3847/2041-8213/abd8c6. arXiv: 2101.03179 [astro-ph.GA].

Wang, Feige et al. (Feb. 2024). 'A Massive Protocluster Anchored by a Luminous Quasar at $z = 6.63$'. In: *The Astrophysical Journal Letters* 962.1, L11, p. L11. DOI: 10.3847/2041-8213/ad20ef. arXiv: 2402.01844 [astro-ph.GA].

White, Simon D. M. and Carlos S. Frenk (Sept. 1991). 'Galaxy Formation through Hierarchical Clustering'. In: *The Astrophysical Journal* 379. ADS Bibcode: 1991ApJ...379...52W, p. 52. ISSN: 0004-637X. DOI: 10.1086/170483.

Wise, John H. et al. (Jan. 2019). 'Formation of massive black holes in rapidly growing pre-galactic gas clouds'. In: *Nature* 566.7742, pp. 85–88. DOI: 10.1038/s41586-019-0873-4. arXiv: 1901.07563 [astro-ph.GA].

Wise, John H. et al. (Feb. 2019). 'Formation of massive black holes in rapidly growing pre-galactic gas clouds'. en. In: *Nature* 566.77427742, 85–88. ISSN: 1476-4687. DOI: 10.1038/s41586-019-0873-4.

Xie, Lizhi et al. (July 2017). 'H2-based star formation laws in hierarchical models of galaxy formation'. In: *MNRAS* 469.1, 968–993. ISSN: 0035-8711. DOI: 10.1093/mnras/stx889.

Xie, Lizhi et al. (Nov. 2020). 'The influence of environment on satellite galaxies in the GAEA semi-analytic model'. In: *MNRAS* 498. ADS Bibcode: 2020MNRAS.498.4327X, 4327–4344. ISSN: 0035-8711. DOI: 10.1093/mnras/staa2370.

Xu, Heng et al. (July 2023). 'Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I'. In: *Research in Astronomy and Astrophysics* 23.7, 075024, p. 075024. DOI: 10.1088/1674-4527/acdfa5. arXiv: 2306.16216 [astro-ph.HE].

Yan, Lin et al. (May 2018). 'Far-UV HST Spectroscopy of an Unusual Hydrogen-poor Superluminous Supernova: SN2017egm'. In: *ApJ* 858.2, 91, p. 91. DOI: 10.3847/1538-4357/aabad5. arXiv: 1711.01534 [astro-ph.HE].

Yang, Jinyi et al. (July 2020). 'Pōniuā'ena: A Luminous z = 7.5 Quasar Hosting a 1.5 Billion Solar Mass Black Hole'. In: *The Astrophysical Journal Letters* 897.1, L14, p. L14. DOI: 10.3847/2041-8213/ab9c26. arXiv: 2006.13452 [astro-ph.GA].

Yang, Jinyi et al. (Dec. 2021). 'Probing Early Supermassive Black Hole Growth and Quasar Evolution with Near-infrared Spectroscopy of 37 Reionization-era Quasars at $6.3 < z \leqslant 7.64$'. In: *The Astrophysical Journal* 923.2, 262, p. 262. DOI: 10.3847/1538-4357/ac2b32. arXiv: 2109.13942 [astro-ph.GA].

Yu, Qingjuan (Apr. 2002). 'Evolution of massive binary black holes'. In: *Monthly Notices of the Royal Astronomical Society* 331.4, pp. 935–958. DOI: 10.1046/j.1365-8711.2002.05242.x. arXiv: astro-ph/0109530 [astro-ph].

Zentner, Andrew R. et al. (May 2005). 'The Physics of Galaxy Clustering. I. A Model for Subhalo Populations'. en. In: *ApJ* 624.2, p. 505. ISSN: 0004-637X. DOI: 10.1086/428898.

Zoldan, Anna et al. (Aug. 2019). 'The evolution of sizes and specific angular momenta in hierarchical models of galaxy formation and evolution'. In: *MMRAS* 487.4, 5649–5665. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1670.

van der Burg, R. F. J., H. Hildebrandt and T. Erben (Nov. 2010). 'The UV galaxy luminosity function at z = 3-5 from the CFHT Legacy Survey Deep fields'. In: *A&A* 523, A74, A74. DOI: 10.1051/0004-6361/200913812. arXiv: 1009.0758 [astro-ph.CO].

Übler, Hannah et al. (Sept. 2023). 'GA-NIFS: A massive black hole in a low-metallicity AGN at z ~ 5.55 revealed by JWST/NIRSpec IFS'. In: *A&A* 677. ADS Bibcode: 2023A&A...677A.145U, A145. ISSN: 0004-6361. DOI: 10.1051/0004-6361/202346137.