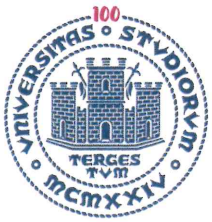


UNIVERSITÀ
DEGLI STUDI
DI TRIESTE

Università degli Studi di Trieste

XXXVI CICLO DEL DOTTORATO DI RICERCA IN

FISICA

The Formation of Supermassive Black Holes From Population III.1 Seeds – Implications for Clustering, Binarity and Gravitational Waves

---

SETTORE SCIENTIFICO-DISCIPLINARE: ASTRONOMIA E ASTROFISICA

DOTTORANDO:
Jasbir Singh

COORDINATORE:
Prof. Francesco Longo

SUPERVISORE:
Prof. Pierluigi Monaco

SUPERVISORE:
Prof. Jonathan C. Tan

Anno Accademico 2022-2023

ABSTRACT

The origin of supermassive black holes (SMBHs) remains an open question in astrophysics. The presence of black holes more massive than a billion solar masses at high redshifts ($z > 6$) challenges many formation mechanisms. Scenarios based on typical stellar-mass seeds struggle to explain the existence of such massive black holes at high redshifts, whereas "direct collapse" based scenarios have difficulties to explain the entire population of these black holes at $z = 0$. One mechanism that can alleviate these problems invokes the collapse of Pop III.1 stars as seeds of SMBHs. These stars are composed of primordial metal free gas forming in dark matter mini halos in the early universe, but with the additional condition of isolation from other stellar or SMBH feedback sources (McKee and Tan, 2008). This isolation condition, parameterized by an isolation distance, d_{iso}, leads to limited fragmentation and thus co-location of the star with the dark matter cusp. If the dark matter particles in the halo are assumed to be Weakly Interacting Massive Particles (WIMPs), then their self-annihilation inside the protostar would release energy which can be captured by the star, allowing it to maintain a relatively cool photosphere (Spolyar et al., 2008; Natarajan et al., 2009; Freese et al., 2010; Rindler-Daller et al., 2015). This results in low levels of ionizing feedback, thus more efficient accretion of almost the entire baryonic content of its natal minihalo, i.e. $\sim 10^5 M_\odot$. In Banik et al. (2019), this seeding mechanism was applied on a cosmological box of $(40h^{-1}\mathrm{Mpc})^3$ using the `PINOCCHIO` code (Monaco et al., 2002; Munari et al., 2017) and the evolution of the seeded halos was followed down to $z = 10$. The number density of SMBHs at that redshift matches the estimated co-moving number density at $z = 0$ for values of $d_{iso} \sim 100$ kpc (proper distance).

In this thesis, we investigate the Pop III.1 seeding mechanism in `PINOCCHIO` simulations of $(40h^{-1}\mathrm{Mpc})^3$ volumes and follow the evolution to the local universe at $z = 0$. We present the methods and tools developed to identify and track the evolution of seeded halos. With these tools, we compute the evolution of the number density of seeded halos and the occupation fraction of halos. We also compute the 3-dimensional 2-point correlation functions and projected correlation functions of the seeded halos, while also accounting for the finite size of the simulation box. The clustering signal shows good agreement with the observations of luminous galaxies in the local universe, and we predict a distinct drop in the clustering at high redshifts at small scales corresponding to the size of the isolation sphere during the formation epoch of the seeds. Then we present a method to distinguish among different seeding mechanisms by creating synthetic Ultra Deep Fields of the same size as the Hubble Ultra Deep Field. We also detail the method used to create a light cone by stitching together multiple pencil beams extracted from the box, and calculating the exact redshifts when each halo enters the light cone.

In the next part of the thesis, we compute the binary statistics of SMBHs within dark matter halos as a function of redshift and for different values of isolation distance. We also investigate the effect on these statistics of different prescriptions for galaxy and SMBH merger timescales. We then compare our results to the latest observations of AGN multiplicity such as the Subaru Strategic Program (Aihara et al., 2018, 2019; Silverman et al., 2020).

In the final part of the thesis, we compute the gravitational wave background (GWB) emanating from SMBH binaries produced via the Pop III.1 seeding mechanism. Since `PINOCCHIO` outputs the merger history of the halos, as done for the multiplicity study, we extract the mergers of seeded halos and assume a time delay between the merger of the halos and the SMBHs occupying them. Then, to calculate the merger rate of the black holes as a function of redshift and chirp mass, we compute the average rate of seeded halo mergers from 10 `PINOCCHIO` boxes. Since `PINOCCHIO` is a dark matter only code, in order to obtain the mass of the SMBHs inhibiting the seeded halos, we use halo mass – black hole mass scaling relations. Then we compute the value of the GWB and explore the impact of different assumptions of time delay and the mass scaling relations. Finally, we do comparisons with the latest results from Pulsar Timing Array experiments, particularly the NANOGrav collaboration (Agazie et al., 2023a).

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AGN Active Galactic Nuclei

BHMF Black Hole Mass Function

CDM Cold Dark Matter

CPTA Chinese Pulsar Timing Array

DM Dark Matter

ESA European Space Agency

GW Gravitational Wave

GWB Gravitational Wave Background

HMT Halo Mass Threshold

HD Hellings and Downs

HSC Hyper Suprime-Cam

HUDF Hubble Ultra Deep Field

IMF Initial Mass Function

JWST James Webb Space Telescope

LIGO Laser Interferometer Gravitational-Wave Observatory

LISA Laser Interferometer Space Antenna

LPT Lagrangian Perturbation Theory

MPI Message Passing Interface

NANOGrav North American Nanohertz Observatory for Gravitational Waves

PTA Pulsar Timing Array

PPTA Parkes Pulsar Timing Array

SDSS Sloan Digital Sky Survey

SMBH Supermassive Black Hole

SSP Subaru Strategic Program

TOA Time of Arrival

UDF Ultra Deep Field

UV Ultraviolet

WIMP Weakly Interacting Massive Particles

1 INTRODUCTION

It has been almost 240 years since the existence of black holes was first postulated by John Michell in 1784 and by Pierre-Simon Laplace independently in 1796 (Michell, 1784; Laplace, 1796). Since then black holes have captured the imagination of scientists and the general public alike. It was only in 1967 that the term *black hole* was coined by John Wheeler. Initially postulated as stars with enough gravitational attraction to pull all the emitted light back towards it, we now have a much better understanding of the behaviour of these objects.

Mere months after Einstein published the General Theory of Relativity (Einstein, 1916), Schwarzschild found an exact solution to the field equations outside a spherically symmetric and non rotating mass distribution (Schwarzschild, 1916). The *Schwarzschild metric*, or the square of the line element describing this spacetime in spherical coordinates is given by:

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

where $r_s = 2GM/c^2$ is the Schwarzschild radius and M being the total mass of the spherical body. If a body becomes compact enough such that its radius becomes smaller than the Schwarzschild radius - it becomes a black hole. It can be shown mathematically that all the geodesics inside this region end at the center, the *singularity*, which means that any object that finds itself inside this region will be trapped and not able to escape outside. At the Schwarzschild radius, the radial component of the metric $(1 - r_s/r)^{-1}$ goes to infinity, although it is only an artefact of the particular choice of coordinates. The metric becomes regular in other coordinates such as the Lemaitre coordinates and Eddington–Finkelstein coordinates. The singularity at $r = 0$ is the only true singularity in this metric. The boundary which separates the universe and the $r < r_s$ region is called the *event horizon*, named after the fact that any event happening inside this region would be unobservable to an outside observer since no signal from the event would emerge outside.

To find a solution which included rotation as well, it took another 47 years when Roy Kerr published an exact solution for a spherically symmetric and rotating black hole (Kerr, 1963). Unlike the Schwarzschild metric, Kerr metric exhibits two event horizons where the purely radial component of the metric goes to infinity. However, this metric also has a region termed *ergosphere*, within which the temporal component of the metric inverts its sign. This implies that any particle inside this region must co-rotate with the inner mass. Also, while in the Schwarzschild metric the singularity at the center is a point, Kerr metric exhibits a circular singularity, known as *ring singularity*. All these additional features of this metric are a result of the spin of the central mass - if the spin equals zero, the metric reduces to a Schwarzschild metric.

Soon after publication of the Kerr metric, it was generalised for a charged and spinning black hole which is now known as the Kerr-Newman metric (Newman and Janis, 1965; Newman et al., 1965).

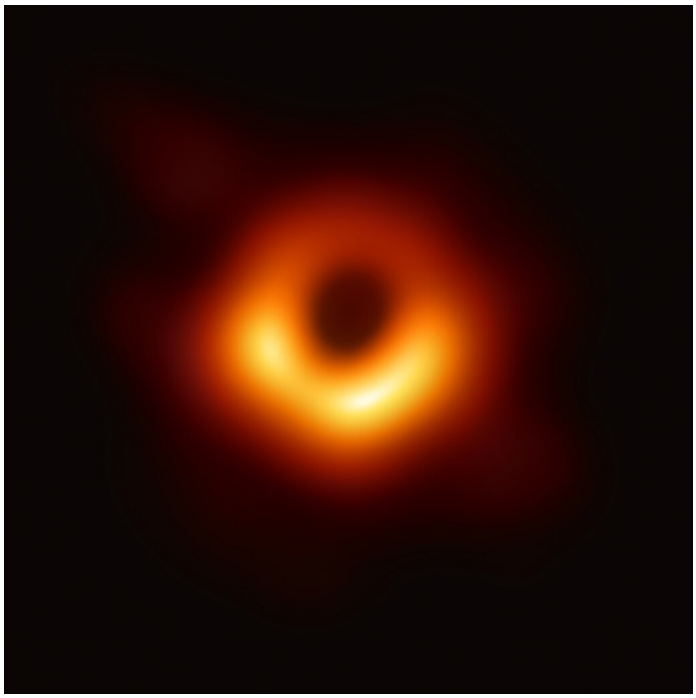
Although a black hole itself is a simple object, the formation process is a lot more complicated. Various mechanisms have been proposed, ranging from the collapse of stars to the fluctuations of spacetime in the early universe. But fundamentally, they all behave the same, irrespective of the formation mechanism. According to the no-hair theorem, a black hole is described completely by three independent properties - mass, electric charge, and angular momentum. This means that any information of the formation mechanism and the properties of the matter (or antimatter) that created the black hole is lost. So in order to investigate their formation mechanisms, we need to rely on their ensemble properties such as their number density and clustering, or the effect of its metric on the surrounding environment or the host galaxy/halo in general.

The formation, and the very existence of black holes have been debated extensively in the past. Regarding the formation, an idealized calculation by Oppenheimer and Snyder (1939) proved that black holes *can* form from ordinary matter, via the collapse of a star made from pressureless "dustball" due to its own self gravity (Hughes, 2005). We now have a much better understanding of the formation of black holes via stars - when the source of internal pressure - nuclear fusion inside a star - dies out, it results in the whole star succumbing to the force of its own gravity and collapsing at the center. Depending on the initial mass of the star, the collapsed remnant can either be a white dwarf, a neutron star, or a black hole. The collapse of stars is one of the most prominent ways to create stellar mass black holes - with masses ranging from a few solar masses to tens of solar masses.

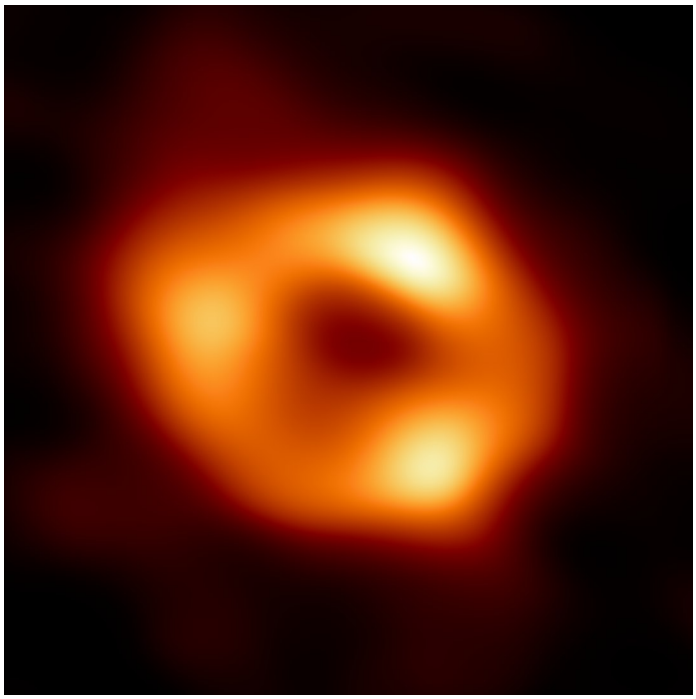
Regarding the existence of black holes, we now have direct evidence of them through the detection of gravitational waves by Laser Interferometer Gravitational-Wave Observatory (LIGO) Collaboration (LIGO Scientific Collaboration et al., 2015; Abbott et al., 2016) emitted by their mergers. At the date of writing this thesis, they have detected multiple stellar mass black holes, with masses reaching as high as $\sim 150 M_\odot$ (Abbott et al., 2020a,b). In addition to this, we even have direct images of black holes now, from the Event Horizon Telescope Collaboration (2019; 2022) who imaged the supermassive black holes at the center of Messier 87 galaxy and our Milky Way Galaxy (figure 1.1). These black holes have masses of the order of $\sim 10^6 - 10^9 M_\odot$, which are on the highest end of the mass spectrum. With these recent images and the GWs from their mergers, there is a definitive proof of their existence. In the next section we discuss in detail the formation of supermassive black holes.

1.1 SUPERMASSIVE BLACK HOLES

As mentioned in the previous section, the origin of stellar mass black holes is well understood, but the same is not completely true for the black holes on the massive end - the supermassive black holes (SMBHs). These black holes



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



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

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

are defined to be more massive than $\sim 10^5 M_\odot$, and are found at the center of most massive galaxies (e.g., Ferrarese and Ford, 2005; Volonteri, 2010; Graham, 2016; Inayoshi et al., 2020; Volonteri et al., 2021; Lusso et al., 2023). The SMBHs imaged by the Event Horizon Telescope mentioned above (figure 1.1), are also at the center of their respective galaxies. The origin of these black holes is one of the most outstanding open questions of contemporary astrophysics.

Discoveries of high redshift quasars, such as J1007+2115 at $z = 7.515$ (Yang et al., 2020) and J0313-1806 at $z = 7.642$ (Wang et al., 2021), which are estimated to host SMBHs with masses $\gtrsim 10^9 M_\odot$, place stringent constraints on SMBH formation and growth scenarios. Wang et al. (2019) found a quasar at $z \sim 6.6$ hosting a SMBH with virial mass around $3.59 \times 10^9 M_\odot$ (Yang et al., 2021), which is one of the most massive black holes discovered at $z > 6.5$. More recently, the continuous influx of astounding results from the James Webb Space Telescope[1] (JWST) has raised more questions than answers regarding the earliest SMBHs. For example, the recently discovered GN-z11 galaxy at redshift $z \sim 10.6$ has hinted at the presence of a SMBH of mass around $1.6 \times 10^6 M_\odot$ (Bunker et al., 2023; Maiolino et al., 2023a). Furthermore, Harikane et al. (2023) detected 10 faint AGNs with masses around $\sim 10^6 - 10^8 M_\odot$ in the redshift range $z = 4 - 6$, while Maiolino et al. (2023b) confirmed 12 more at $4 < z < 7$ with masses $4 \times 10^5 M_\odot$ to $8 \times 10^7 M_\odot$. These recent results, along with others (for e.g., Yue et al., 2023; Stone et al., 2024) show the interesting trend of SMBH masses lying 1 to 2 dex above the local black hole mass - stellar mass relation. Although this may imply a possible evolution of the relation at higher redshifts as summarised in Pacucci et al. (2023), Li et al. (2024) suggests that this may be mainly due to a combination of selection effects and uncertainties in black hole and stellar mass measurements. Nonetheless, the existence of these high redshift SMBHs imply that at least some of them could form and grow efficiently to very high masses by the time the universe was only a few 100 million years old.

The challenge comes in explaining the masses of these behemoths when the universe was still in its infancy. The figure 1.2 taken from Wang et al. (2021) gives a visual representation of this conundrum. The figure shows some of the most massive and earliest quasars and their growth track assuming Eddington-limited accretion for their entire lifetime. For the most massive black holes, even assuming very early formation at $z \sim 30$ with Eddington-limited accretion the SMBH seed mass would need to be $\gtrsim 10^4 M_\odot$ and a later formation epoch would imply even higher seed masses. Going to earlier and earlier times indefinitely to reduce the seed mass is also not an option since the formation of halos with enough baryonic matter and then the formation of stars which will collapse - requires time and favourable conditions. While scenarios of super-Eddington accretion have been proposed (e.g., Kohri et al., 2022), numerical simulations indicate that typical gas supply rates to early-formed SMBHs are impacted by star formation feedback and will be far below the level needed to sustain Eddington-limited accretion rates (e.g.,

1 https://webbtelescope.org/

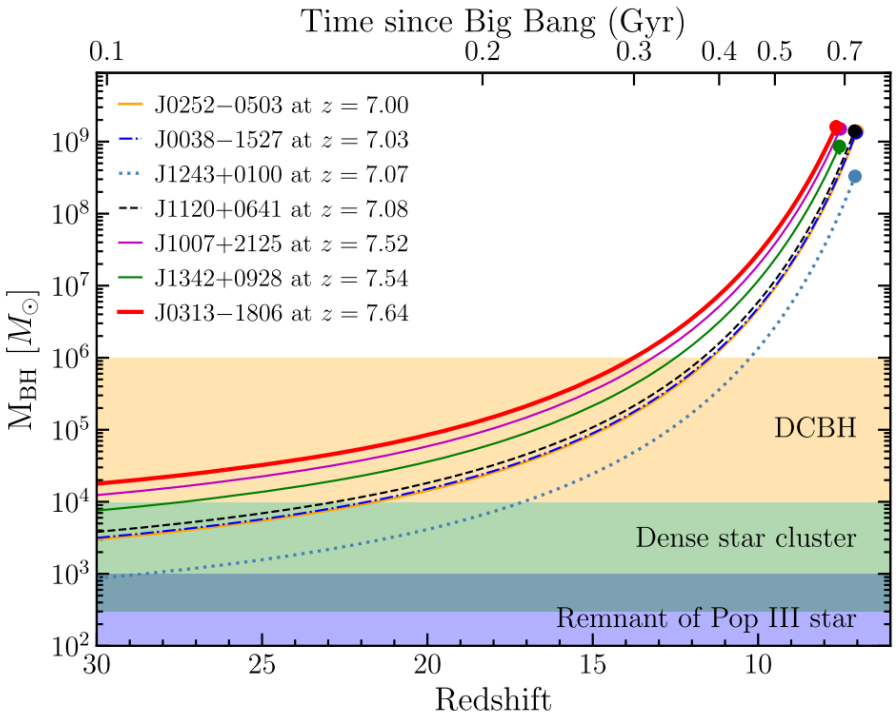


Figure 1.2: The figure from Wang et al. (2021) showing the possible growth tracks of confirmed high redshift quasars ($z \geqslant 7$) assuming an Eddington-limited accretion for the entire lifetime with 0.1 radiative efficiency. The colored horizontal bands refer to the range of seed masses of the black hole form different formation mechanisms.

O'Shea et al., 2005; Jeon et al., 2023). These considerations motivate the need for models of black hole formation at the supermassive, $\gtrsim 10^5 M_\odot$ scale.

1.1.1 *Formation mechanisms*

There are a variety of proposed ideas for the physical mechanism of SMBH formation (e.g., Rees, 1978). One suggested process is "direct collapse", which involves a massive primordial composition gas cloud contained in a relatively massive, atomically-cooled halo of $\sim 10^8 M_\odot$. The cloud collapses into a single, supermassive star of $10^4 - 10^6 M_\odot$ that then forms a SMBH (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., 2022a). It may also happen that the collapse proceeds via an intermediate phase where a quasi-star is formed which is powered by accretion onto a small growing black inside the quasi-star instead of nuclear burning (Begelman et al., 2006, 2008). The main condition required for these supermassive seeds to form is that when the protostar is growing, the accretion can continue without gas fragmentation. This can happen if the surface of the protostar remains cool, which results in less ultraviolet radiation emerging from the protostar and hence no disruption to the accretion flow. And the possibility of cool outer layers increases if the protostar grows in timescales shorter than the Kelvin-Helmholtz timescale

(Haemmerlé et al., 2020). Furthermore, for the accretion to continue without fragmentation, the halos need to be metal free - which is already satisfied since they only have primordial gas composition. However, the formation of molecular hydrogen must be suppressed, which provides a cooling channel and hence fragmentation (Latif et al., 2013). This requires strong Lyman-Werner radiation to photodisassociate the molecular hydrogen (Wise et al., 2019). The alignment of all these conditions naturally to create SMBHs makes this scenario relatively rare (Habouzit et al., 2016).

Although the number density of black holes emerging from direct collapse would be enough to explain the currently known population of high redshift quasars, the somewhat contrived conditions required for this scenario are not thought to be common enough to explain the total observed population of SMBHs at $z = 0$ (Chon et al., 2016; Wise et al., 2019). Furthermore, recent simulations have shown that the supermassive stars forming via this mechanism might not be as massive as initially predicted, but only reaching $\lesssim 10^4 M_\odot$, due to the turbulent environment present in the initial stages of galaxy formation which disrupts the accretion flow (Regan et al., 2020).

Another mechanism to form intermediate, or even supermassive black holes is through runaway stellar mergers in young and dense metal-poor clusters to create stars with masses of the order $\sim 200 - 10^3 M_\odot$ (e.g., Portegies Zwart et al., 2004). This mass can be reached through repeated collisions if the massive stars can reach the cluster core to increase the collision rate drastically before exploding as supernovae (Ebisuzaki, 2003). Gas accretion driven compression of a dense cluster of stellar mass black holes to form a SMBH has also been proposed (Kroupa et al., 2020), along with the repeated collisions among minihalos of $10^5 M_\odot$ to create favourable conditions (Katz et al., 2015). However, in general predicting whether the conditions needed for such dense clusters arise in galaxies and at what rate is very challenging given the need to resolve the formation and evolution of individual stars, so predictions for the cosmological population of such systems are highly uncertain (see, e.g., Boekholt et al., 2018; Chon and Omukai, 2020; Tagawa et al., 2020).

Another class of SMBH seeding model considers the very first, so-called Population (Pop) III stars as potential progenitors. These were the first stars to form in the early universe, in metal-free and metal-poor environments. These environmental conditions provide the perfect opportunity to accrete matter onto the protostar without the issue of fragmentation. The only cooling agent in this case in molecular hydrogen. Since it is an inefficient coolant compared to heavier metals, the temperature of the gravitationally contracting primordial gas is higher than the interstellar medium - resulting in reduced fragmentation. Conventional models of Pop III star formation predict stellar masses that are "only" $\sim 100\, M_\odot$ (e.g., Madau and Rees, 2001; Abel et al., 2002; Bromm et al., 2002; Tan and McKee, 2004; McKee and Tan, 2008; Hosokawa et al., 2011; Susa et al., 2014), which would only have the ability to produce stellar-mass black holes, i.e., relatively low-mass seeds.

More exotic models involving modification of the standard cold dark matter paradigm have also been proposed. For example, 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). An even more extreme scenario is one in which SMBHs are primordial black holes, although this appears to be disfavoured by the clustering analysis of Shinohara et al. (2023).

1.1.2 *Seeding in simulations*

Given the uncertainty of SMBH formation models and the difficulty of resolving the small-scale physics, cosmological simulations have typically made very simplified assumptions for the SMBH seeding process based on the properties of the parent halo or galaxy. One of the simplest and most widely used model is the halo mass threshold (HMT) seeding scheme based on the methods developed by Sijacki et al. (2007) and Di Matteo et al. (2008), in which a seed black hole is assumed to form in a halo crossing a certain mass threshold. The Illustris project (Vogelsberger et al., 2014) used this mechanism to add SMBHs of mass $1.4 \times 10^5 M_\odot$ in each halo which crosses a mass threshold of $m_{th} = 7.1 \times 10^{10} M_\odot$. A similar approach was used in the Evolution and Assembly of GaLaxies and their Environments (EAGLE) simulation (Barber et al., 2016). These kind of mechanisms are largely driven by the mass resolution of the simulations: since the early phases of galaxy formation are not resolved, seed black holes are placed in halos as soon as they are resolved with a minimum number of particles.

More recent simulations have taken into consideration additional properties of the host galaxy for SMBH seeding. For example, the Horizon-AGN simulation (Volonteri et al., 2016) required gas and stellar densities and stellar velocity dispersion to exceed certain thresholds for a galaxy to form a black hole, with a seed mass of 10^5 $M_\odot$ adopted. In addition, all the forming black holes needed to be separated by at least 50 comoving kpc, and their formation was only allowed down to $z = 1.5$. Adopting similar criteria, the OBELISK simulation (Trebitsch et al., 2021) also applied conditions of gas and stellar density needing to exceed certain thresholds, including Jeans instability of the gas, as well as a required isolation of 50 kpc from other SMBHs to avoid multiple black holes forming in the same galaxy. If all these conditions were satisfied, then a black hole of $3 \times 10^4 M_\odot$ was assigned to the galaxy. In another approach, the ROMULUS simulation (Tremmel et al., 2017) employed criteria of a limit on metallicity, a threshold on gas density, and a restricted temperature range for SMBH formation, with a seed mass of $10^6 M_\odot$ adopted. In yet another example, Bhowmick et al., 2022b have considered a variety of gas-based SMBH seeding prescriptions and a range of seed masses from $\sim 10^4$ to 10^6 $M_\odot$. While the investigation of certain thresholds of physical quantities for SMBH formation is an advance on a simple HMT models, the above studies are still far from being a complete physical description of SMBH formation.

1.2 POP III.1 MODEL

In this thesis, we focus on a formation mechanism based on Pop III.1 stars as the progenitors of SMBHs. These stars are a subcategory of Pop III stars, which were distinguished by McKee and Tan (2008) into Pop III.1 and Pop III.2 stars. The former are the ones which are formed first in dark matter minihalos in a given region of the universe and so are isolated from other stellar or SMBH feedback sources. The latter also form with primordial composition of gas in dark matter minihalos, but are influenced by external astrophysical sources of feedback. This results in a greater possibility of fragmentation and hence lower masses compared to Pop III.1 stars. Semi analytical models presented by Tan and McKee (2004) and McKee and Tan (2008) showed that Pop III.1 stars can grow up to $\sim 140 M_\odot$. Greif and Bromm (2006) also argues similar subdivision of Pop III stars, and present $\geqslant 100 M_\odot$ masses for Pop III.1 stars, while an order less for the Pop III.2 stars. However, Tan et al. (2010) applied the semi analytical models of McKee and Tan (2008) to simulations of O'Shea and Norman (2007) and found an Initial Mass Function (IMF) of the stars peaking at $\sim 100 M_\odot$, but with a tail extending to $\sim 10^3 M_\odot$.

The higher end of the obtained IMF is massive, but not enough to reach the scales of SMBHs of $\sim 10^5 M_\odot$. As mentioned before, to explain the masses of the observed high redshift quasars, seed masses around of $10^3 M_\odot$ would need Eddington-limited accretion through its entire lifetime. Such rate of prolonged accretion is unlikely to be sustained, due to the radiative and mechanical feedback of the Pop III.1 star on its surroundings (O'Shea et al., 2005; Johnson and Bromm, 2007; Milosavljević et al., 2009).

One interesting way to increase the mass of these stars further is by considering the physics of dark matter particles inside their natal minihalo. If the dark matter particles are constituted by Weakly Interacting Massive Particles (WIMP), then the energy released by their self-annihilation inside the protostar can be captured by it (Spolyar et al., 2008; Natarajan et al., 2009; Freese et al., 2010; Rindler-Daller et al., 2015). This effect will be the strongest at the center of the halo, where the density of dark matter particles is the highest. Thus, this condition requires the co-location of the star with the cusp of the dark matter halo. The energy injection from the self-annihilation results in the protostar to be in a swollen, relatively cool state, minimizing the ionizing feedback in its surroundings. This allows efficient accretion of the baryonic content of the halo; essentially the entire $\sim 10^5 M_\odot$ baryonic mass of the halo can be accreted before the star collapses into a SMBH after a few Myr.

This Pop III.1 seeding mechanism, which is based on locating isolated minihalos, was applied in a cosmological simulation by Banik et al. (2019) (hereafter Paper I). The evolution was followed from high redshifts down to $z = 10$. The main free parameter in the model is the *isolation distance* (d_{iso}), i.e., how far a newly forming minihalo needs to be from previously formed halos in order to be a Pop III.1 source. For a fiducial value of $d_{iso} = 100$ kpc (proper distance), the model yields co-moving number densities of SMBHs that match the estimated level of the known $z = 0$ SMBH population. Note,

that in this case (and all other reasonable cases) most minihalos do not form Pop III.1 sources. Rather, most are Pop III.2 sources, which are metal free, but having been disturbed by radiative feedback are expected to undergo significant fragmentation to form only lower-mass (e.g., $\sim 10\,M_\odot$) stars (Greif and Bromm, 2006).

This thesis aims to carry out further investigation of the Pop III.1 seeding mechanism, by applying it to cosmological simulations extending to $z = 0$. We particularly focus on three aspects of the SMBHs formed from this mechanism: the ensemble properties of the seeded halos and the SMBHs hosted in them, the dual AGN statistics, and finally the gravitational wave background generated by the mergers of these black holes. The thesis is structured as follows.

CHAPTER 2 presents a brief introduction of the theoretical formulation of the gravitational wave background, and the latest results from Pulsar Timing Array experiments.

CHAPTER 3 first describes the cosmological simulation used, and the numerical methods developed to implement the seeding mechanism using the simulation outputs. This chapter also details the ensemble results of the seeded halos, namely the number density, the occupation fraction, and the clustering.

CHAPTER 4 outlines the methods used to create a synthetic Ultra Deep Field from the simulation boxes. After this the chapter also explores the dependence of the seeding parameters on the number of seeds in the field.

CHAPTER 5 provides preliminary estimates on the fraction of dual AGNs from the Pop III.1 model, after describing simple assumptions and physical models adopted for merging timescales.

CHAPTER 6 describes in detail the methods used to calculate the gravitational wave background from the Pop III.1 seeding mechanism, and its comparisons with the latest observational results.

CHAPTER 7 summarizes the main conclusions of the thesis, and the future scope of the presented work.

2 GRAVITATIONAL WAVE BACKGROUND

It took almost 100 years since the prediction of gravitational waves (GWs) by the General Theory of Relativity (Einstein, 1916) to directly detect them through the merger of stellar mass black holes. This signal was detected by the Laser Interferometer Gravitational-Wave (LIGO) Observatory in September 2015 which matched the waveform of a stellar binary black hole merger with more than 5.1σ significance (Abbott et al., 2016). Since then more mergers have been detected, including a neutron star - neutron star merger (Abbott et al., 2017). Most of the compact objects detected through these mergers, either the merging black holes or the remnants, have masses of around a few tens of solar masses, with the maximum around $\sim 150 M_\odot$ (Abbott et al., 2020a,b). This is due to the design sensitivity of the detector, which has a sensitivity peak at $\sim$ 100 Hz (LIGO Scientific Collaboration et al., 2015), most suitable for detecting the inspiral frequency just before and during the merger for stellar mass black holes.

For black holes at the extreme end of the mass spectrum - around $\sim 10^8 - 10^{10} M_\odot$, which are found at the center of massive galaxies, their binaries emit GWs with slowly evolving frequencies, superimposing to form a broadband signal in the nHz range - the Gravitational Wave Background (GWB; e.g., Rajagopal and Romani 1995; Wyithe and Loeb 2003; Sesana et al. 2004; Burke-Spolaor et al. 2019). This background is of particular interest to us since the Pop III.1 model predicts supermassive black hole seeds formation in the early universe which will later form binaries and create a GWB. In this chapter we present a brief introduction to the theoretical formulation of the background, and then the advancements in the detection of this background using arrays of pulsars. Finally, we end with the latest results from the detection of the background.

2.1 THEORETICAL FORMULATION

In this section, we present an overview of the key concepts and relevant formulae to calculate the GWB based on a population of merging SMBH binaries. Let f_r be the frequency of the gravitational waves emitted by a source in its frame, which is at redshift z. Then the redshifted frequency f observed at Earth or in the observer's frame is related to the source frame frequency by $f_r = f(1+z)$. Let the energy emitted by the source in gravitational waves between frequencies f_r and $f_r + df_r$ be $(dE_{gw}/df_r)df_r$.

Next, we define $\Omega_{gw}(f)$ as the present-day GW energy density per logarithmic (observer frame) frequency interval, normalized to the critical energy

density ρ_c. Then the total present day energy density of gravitational radiation $\mathcal{E}_{gw}$ can be calculated as

$$\mathcal{E}_{gw} \equiv \int_0^\infty \rho_c c^2 \Omega_{gw}(f) \frac{df}{f}, \tag{2}$$

where $\rho_c = 3H_0^2/(8\pi G)$. Now let us consider a population of coalescence events per unit comoving volume occurring between redshift z and $z + dz$, and define it by $(dn/dz)dz$. In other words the number density of events per unit redshift is defined as dn/dz. Then in a homogeneous and isotropic universe, the total present day energy density $\mathcal{E}_{gw}$ is equal to the sum of the energy densities radiated at each redshift from all the sources (Phinney, 2001):

$$\mathcal{E}_{gw} = \int_0^\infty \int_0^\infty dz \frac{dn(z)}{dz} \frac{1}{1+z} \frac{dE_{gw}}{df_r} f_r \frac{df_r}{f_r} \tag{3}$$

$$= \int_0^\infty \int_0^\infty \left(\frac{dn(z)}{dz} \frac{1}{1+z} f_r \frac{dE_{gw}}{df_r} \right) dz \frac{df}{f}. \tag{4}$$

Here the factor $1/(1+z)$ accounts for the redshift of the gravitons since emission. Now equating equation 2 with equation 4, we get:

$$\rho_c c^2 \Omega_{gw}(f) = \int_0^\infty \frac{dn(z)}{dz} \frac{1}{1+z} f_r \frac{dE_{gw}}{df_r}\bigg|_{f_r = f(1+z)} dz. \tag{5}$$

Finally, using equation 2 from Phinney 2001 to relate the energy density with the GW amplitude of the spectrum, we can rewrite the last equation as:

$$\rho_c c^2 \Omega_{gw}(f) = \frac{\pi}{4} \frac{c^2}{G} f^2 h_c^2(f) = \int_0^\infty \frac{dn(z)}{dz} \frac{1}{1+z} f_r \frac{dE_{gw}}{df_r}\bigg|_{f_r = f(1+z)} dz, \tag{6}$$

where $h_c(f)$ is the characteristic amplitude of the GW spectrum over logarithmic frequency interval $d\ln f = df/f$. We can rearrange this equation to finally obtain the expression for the spectrum:

$$h_c^2(f) = \frac{4G}{\pi c^2} \frac{1}{f^2} \int_0^\infty \frac{dn(z)}{dz} \frac{1}{1+z} \frac{dE_{gw}}{d\ln f_r}\bigg|_{f_r = f(1+z)} dz. \tag{7}$$

This equation gives the important result that the spectrum is independent of cosmology, and only depends on the rate of merger events and the energy spectrum of the gravitational wave emission. For a population of SMBH binaries, which are far from their last stable circular orbit frequencies, the energy spectrum, or the energy emitted per logarithmic frequency interval is given by (Thorne, 1987)

$$\frac{dE_{gw}}{d\ln f_r} = \frac{1}{3G} (G\mathcal{M})^{5/3} (\pi f_r)^{2/3}, \tag{8}$$

where $\mathcal{M}$ is the source/rest frame chirp mass, defined for two objects of masses m_1 and m_2 orbiting around each other as $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$. The frequency of the GW f_r is twice the orbital frequency f_{orb}. This equation 8 is the Newtonian approximation which is sufficient for our

purposes, and it is valid for cases when the orbit is shrinking only due to the emission of GWs without any eccentricity.

Considering again the number density of the mergers, we can update it to include the dependence on the chirp masses of the binaries:

$$\frac{dn}{dt} = \int_0^\infty \frac{d^2n}{dz d\mathcal{M}} d\mathcal{M}, \tag{9}$$

where $d^2n/dz d\mathcal{M}$ is the comoving number density of mergers per unit redshift and chirp mass. Using equations 7, 8 and 9, we can now write the expression for characteristic amplitude of the GWB for a population of inspiraling massive binaries as

$$h_c^2(f) = \frac{4G}{\pi c^2}\frac{1}{f^2}\int_0^\infty dz \int_0^\infty d\mathcal{M} \frac{d^2n}{dz d\mathcal{M}} \frac{1}{1+z}\frac{dE_{gw}}{d\ln f_r} \tag{10}$$

$$= \frac{4G^{5/3}}{3\pi^{1/3}c^2}\frac{1}{f^{4/3}}\int_0^\infty dz \int_0^\infty d\mathcal{M}(1+z)^{-1/3}\mathcal{M}^{5/3}\frac{d^2n}{dz d\mathcal{M}}. \tag{11}$$

From this simplified version of the characteristic amplitude (equation 11), we can conclude that to calculate the background we just need the number density of the merger events per unit redshift and chirp mass, and that the frequency dependence of the spectrum is independent of the number density. This expression also motivates the common expression for the background as a power law of the form

$$h_c(f) = A_{yr}\left(\frac{f}{yr^{-1}}\right)^{-\alpha}, \tag{12}$$

where A_{yr} is the amplitude at a reference frequency of $1yr^{-1}$, and $\alpha = 2/3$ for the ideal case. The coefficient A_{yr} contains all the information of the source population, which turns out to be independent of the frequency of the spectrum in the ideal case.

Another way to interpret the equation 11 is by expressing the comoving number density of mergers in terms of $d^3N/dz d\mathcal{M} d\ln f_r$, which depicts the number of comoving number of binaries emitting in a given logarithmic interval with chirp mass in the range $[\mathcal{M}, \mathcal{M}+d\mathcal{M}]$, and redshifts belonging to the interval $[z, z+dz]$ (Rajagopal and Romani, 1995; Jaffe and Backer, 2003; Sesana et al., 2008):

$$\frac{d^2n}{dz d\mathcal{M}} = \frac{d^3N}{dz d\mathcal{M} d\ln f_r}\frac{d\ln f_r}{dt_r}\frac{dt_r}{dz}\frac{dz}{dV_c}, \tag{13}$$

where

$$\frac{d\ln f_r}{dt_r} = \frac{96}{5}\frac{(G\mathcal{M})^{5/3}}{c^5}(\pi f_r)^{8/3} \tag{14}$$

$$\frac{dt_r}{dz} = -\frac{1}{H(z)(1+z)} \tag{15}$$

$$\frac{dz}{dV_c} = \frac{H(z)}{4\pi d_c^2(z)c}. \tag{16}$$

In the expressions above, $d \ln f_r / dt_r$ is the rate of change of frequency of the emission due to gravitational radiation, and t_r is the time in rest frame. dt_r/dz is the lightcone term, and dV_c is the differential comoving volume of the shell at redshift z. Using all these expressions, we can calculate $d^3N/dz d\mathcal{M} d \ln f_r$ as:

$$\frac{d^3N}{dz d\mathcal{M} d \ln f_r} = \frac{5}{24} \frac{c^6}{\pi^{5/3}} (G\mathcal{M})^{-5/3} (1+z) d_c^2(z) f_r^{-8/3} \frac{d^2n}{dz d\mathcal{M}}. \tag{17}$$

This expression allows us to populate the universe with GW emitting binaries, using as input the number density of coalescences $d^2n/dz d\mathcal{M}$, which can be from an analytical expression, or extracted from a simulation.

Let us also consider the strain of GWs emitted by a binary in circular orbit, whose orbit is decaying via the emission of the waves. The leading order contribution is given by (Thorne, 1987)

$$h_s(f_r) = \frac{8\pi^{2/3}}{10^{1/2}} \frac{(G\mathcal{M})^{5/3}}{c^4 d_c(z)} f_r^{2/3}, \tag{18}$$

where $d_c(z)$ is the comoving distance to redshift z. Equipped with this equation, and equations 11 and 17, we apply a little algebraic manipulation to arrive at the following expression for the strain of the GWB (Sesana et al., 2008):

$$h_c^2(f) = \int_0^\infty dz \int_0^\infty d\mathcal{M} \frac{d^3N}{dz d\mathcal{M} d \ln f_r} h_s^2(f_r). \tag{19}$$

This equation provides a simple interpretation of the background: the square of the background is given by integral of all the sources emitting in a logarithmic frequency bin $d \ln f_r$ multiplied by the squared strain of the individual source. We will use this expression later in the thesis to calculate the background for the Pop III.1 model (Chapter 6). This equation also highlights an important aspect of calculating the background, especially at higher frequencies where the contribution mostly comes from the most massive binaries. It may happen that there is on average less than one source with a particular chirp mass, at a particular redshift and emitting in a particular frequency bin: this means that the term $d^3N/dz d\mathcal{M} d \ln f_r$ could become less than one. But since this term gives the number of sources in that particular 3-dimensional $(z, \mathcal{M}, \ln f_r)$ bin, it should be set to zero as there are no "fractional" sources in a real universe. This was not being considered in equation 11, even though it is equivalent with equation 19. This *discreteness* of the sources at high frequencies causes a drop in the amplitude of the spectrum, which would otherwise follow a simple power law with $h_c(f) \propto f^{-2/3}$. For an extensive discussion on this, we refer the reader to Sesana et al. (2008). In our work, we consider the discreteness of the sources.

2.2 DETECTING THE BACKGROUND

To detect the GWB, the same basic principle which is used to detect the GWs terrestrially is used - measure the changes in the distance between

two freely falling references masses, which could be due to the passing of GWs between the two objects (Pirani, 1956, 2009). The physical limitations of the terrestrial observatories (such as LIGO), make them sensitive to the ~ 100 Hz GW frequency band, as mentioned before. To detect the nHz waves which are the characteristic of SMBHs, Sazhin (1978) and Detweiler (1979) proposed monitoring the changes in regularity of pulsar emission as the evidence for a GW passing between the solar system and the pulsars. Their extremely precise Time of Arrival (TOA)s of the pulses allows them to be used as accurate clocks, and any small change in their timings can be detected. In a seminal work by Hellings and Downs (1983), they suggested that the correlations between the perturbations of Time of Arrivals (TOAs) of multiple pulsars could reveal a GW signal. To this regard, Romani (1989) and Foster and Backer (1990) proposed a way to detect the GWB buried in the pulsar TOA noise by monitoring a pulsar timing array of highly stable millisecond pulsars.

In the present day we now have multiple groups and consortia monitoring pulsars for multiple years to detect tiny changes in the TOAs which are attributed to the GWB. The North American Nanohertz Observatory for Gravitational Waves[1] (NANOGrav) is one such consortium who recently found an evidence of a background after monitoring a set of 67 pulsar for a baseline exceeding 15 years (Agazie et al., 2023a; Agazie et al., 2023b). The results were released together with other Pulsar Timing Arrays (PTAs) collaborations, namely the European PTA[2] (EPTA Collaboration et al., 2023), the Parkes PTA[3] (PPTA; Reardon et al., 2023) and the Chinese PTA (CPTA; Xu et al., 2023), which all show an evidence of the GWB.

Focusing on the results from the NANOGrav Collaboration (the NANOGrav 15 year data), figure 2.1 shows the relative position of the monitored pulsars with respect to our galaxy. All the pulsars were monitored for at least 3 years, with some of them monitored over 16 years. After analysing the TOAs from all these pulsars, the interpulsar phase-coherent correlation was found to be in good agreement with the correlation expected from a signal of GW origin (HD correlation; Hellings and Downs, 1983). Figure 2.2 shows these correlations in interpulsar angular separation bins. These correlations are calculated assuming the signal is due to an ideal GWB or equivalently $\alpha = 2/3$ (in the figure γ is the spectral index defined as $\gamma = 3 + 2\alpha = 13/3$). As we can see, the measured signal follows the HD correlation, providing evidence for a GWB.

Figure 2.3 shows the posteriors of the amplitude of the spectrum and the spectral index at 1yr^{-1} and 0.1yr^{-1} reference frequencies obtained through the Bayesian analysis of the signal (Agazie et al., 2023a). The posterior medians and the $5\% - 95\%$ quantiles for the amplitude are $A_{\text{yr}} = 6.4^{+4.2}_{-2.7} \times 10^{-15}$ and for the spectral index $\gamma = 3.2^{+0.6}_{-0.6}$. It is worth mentioning that although in literature the models usually present A_{yr}, i.e., the spectrum at 1yr^{-1} reference frequency, the constraints from NANOGrav are usually derived from

1 https://nanograv.org/
2 https://www.epta.eu.org/
3 https://www.atnf.csiro.au/research/pulsar/ppta/

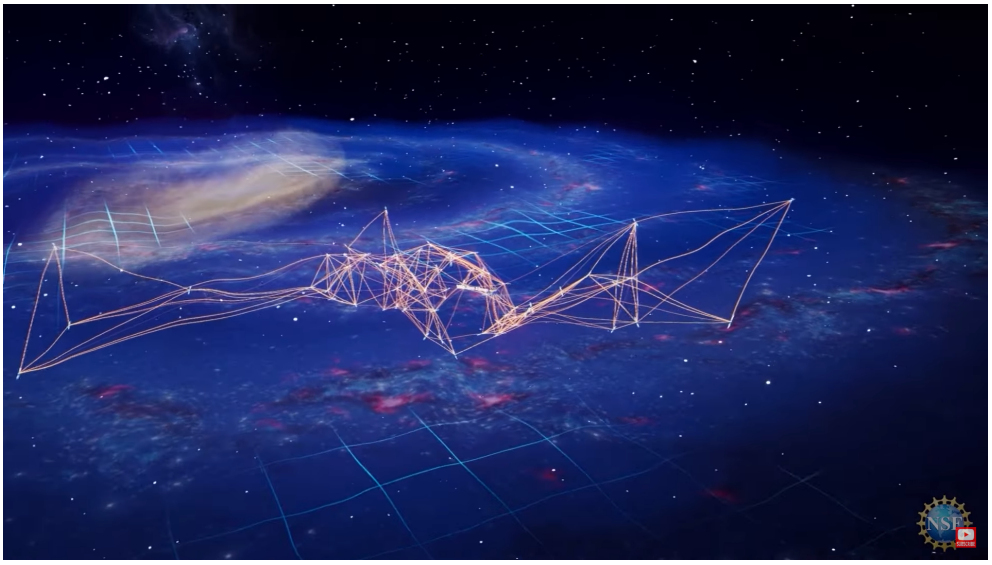


Figure 2.1: The relative position of all the pulsars monitored by NANOGrav for the 15 year data set. All the vertices represent the position of the pulsars. The position of our Solar System is inside the dense group of pulsars in the center of the image, but slightly offset to the left. Credits: NANOGrav Collaboration (https://www.youtube.com/watch?v=c2EKbvnee3o)

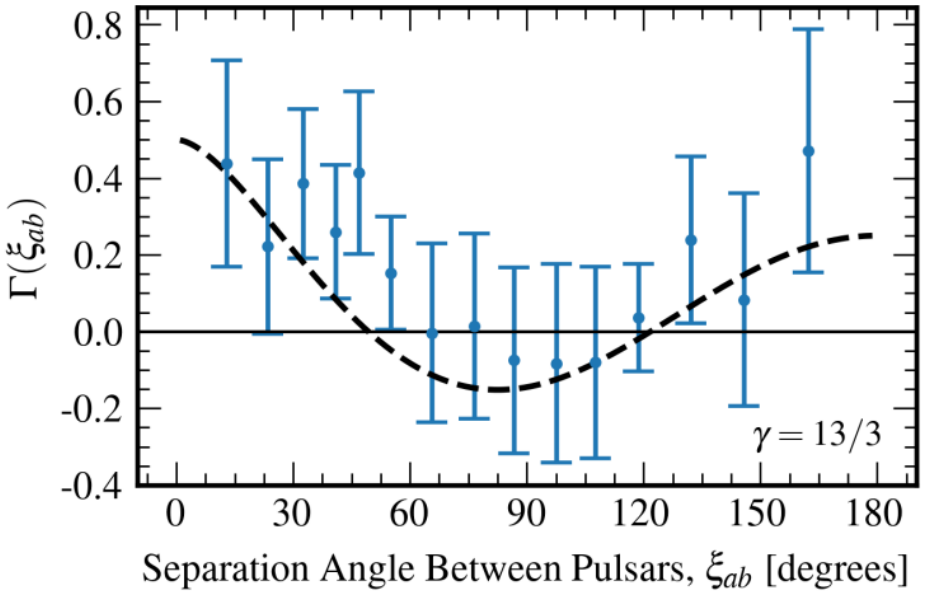


Figure 2.2: The interpulsar angular-separation-binned correlation from 2211 distinct pairs from 67 pulsars, assuming an ideal GW background with $\alpha = 2/3$, or $\gamma = 13/3$. The theoretical HD correlation is shown in black dashed line. The horizontal black line represents no correlation. Credits: Agazie et al. (2023a).

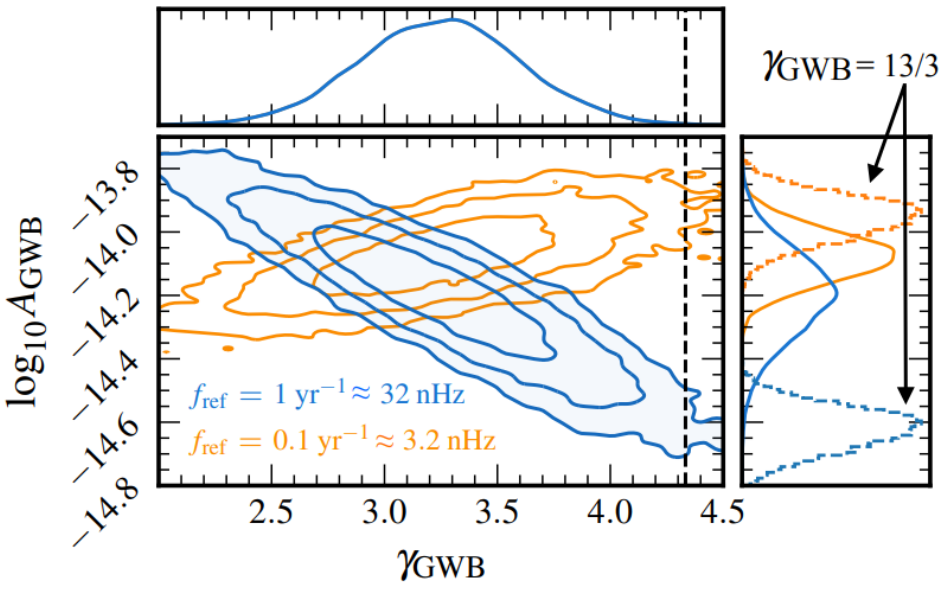


Figure 2.3: The $1\sigma, 2\sigma$ and 3σ credible region posteriors for the value of the GWB and the spectral index γ at reference frequencies of $1{\rm yr}^{-1}$ (*blue* contours) and $0.1{\rm yr}^{-1}$ (*orange* contours). The dashed colored distributions in the plot on the right show the posteriors after assuming $\gamma = 13/3$. Credits: Agazie et al. (2023a).

lower frequencies, since they have higher signal-to-noise ratio and the most sensitive frequency band is around $\sim 0.1{\rm yr}^{-1}$. The main takeaways from this figure are:

- The amplitude of the background is slightly on the higher end compared to previous predictions from various models for a background from SMBH binaries (e.g., Sesana et al., 2008; McWilliams et al., 2014; Ravi et al., 2015; Kulier et al., 2015; Zhu et al., 2019; Simon, 2023). This could imply that the signal is coming from a more massive population of black holes than previously assumed.
- The posterior for the spectral index γ is in moderate tension with the theoretical $\gamma = 13/3$ for GW-only induced inspiral of the binaries. The smaller value implies that the orbital shrinkage of the binaries is significantly affected by the binaries' interaction with the surrounding gas and stars.

The figure 2.4 shows the strain of the GWB, calculated using the NANOGrav 15 yr dataset (Agazie et al., 2023). As mentioned before, the lower frequency bins have lower signal-to-noise ratio which is apparent from the size of the distributions in the violin plot. The environmentally driven evolution of binaries model, and only GW-induced model - both fit well to the data, although the former provides a slightly better fit, especially at the lowest frequency bin where the environmental effects are the most prominent (e.g., Sesana, 2010; Kelley et al., 2017; Biava et al., 2019; Bortolas et al., 2021). From the figure we can also observe that the GW only evolution drops below the ideal power law

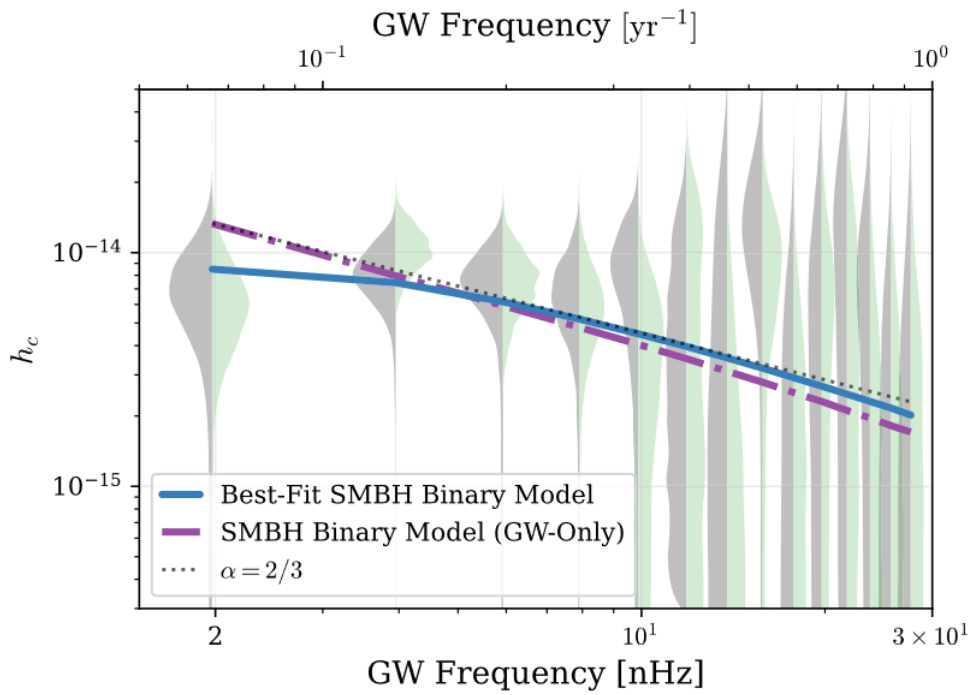


Figure 2.4: The NANOGrav 15 yr GWB spectrum posteriors translated into the characteristic strain. The violin plot depicts the posteriors, while the two colors represent the output from two different estimators. The solid blue line shows the best fit environmentally driven binary evolution, and the purple line shows the strain for only GW-induced evolution. The dotted black line shows the ideal power law model with $\alpha = 2/3$. Credits: Agazie et al. (2023)

model at higher frequencies due to the discreteness of the sources mentioned in §2.1.

With the current signal-to-noise ratio, the signal cannot be definitely attributed to the SMBH binaries, even though all the results derived are consistent with binaries. Nonetheless, the possibility of the signal originating from non-astrophysical sources and new physics is still present. Afzal et al. (2023) investigated the possibility of the signal originating from cosmic inflation, scalar-induced GWs, cosmic strings, and more. Although some of the models provide a better fit to the data compared to the inspiraling SMBH binaries, it is believed that the improvement in noise modelling and the understanding of the SMBH binary signals will weaken the evidence against new physics. In the near future, the results of NANOGrav 15 year data set will be combined with results from other PTA experiments to increase the number of monitored pulsars and the signal-to-noise ratio. This will provide an improved understanding of the signal and most probably confirm the SMBH binary origin of the background.

In Chapter 6, we calculate the GWB from SMBH binaries in the Pop III.1 model and compare it with the NANOGrav 15 year results.

3

EVOLUTION OF SEEDED HALOS DOWN TO THE LOCAL UNIVERSE

As mentioned in Chapter 1, Banik et al. (2019) (Paper I) applied the Pop III.1 seeding mechanism in a cosmological simulation down to $z = 10$. They presented the evolution of the number density for different isolation distances up to $z = 10$, the clustering of the seeded halos at that redshift, and also the synthetic sky maps, among other results. In this chapter, we take the same seeding mechanism and push all the above mentioned results plus the occupation fraction down to redshift zero using the latest version of PINOCCHIO code. We start by giving an overview of the PINOCCHIO code, then present the seeding scheme and its methodological application to the simulation outputs. After this we present all the results obtained from the application of this seeding scheme, such as the number density, the occupation fraction and the clustering. We finish by presenting a simple comparison of the projected correlation function of the seeded halos with observations. The entire content of this chapter presents the results from the paper published in the Monthly Notices of the Royal Astronomical Society:

- **Jasbir Singh**, Pierluigi Monaco, Jonathan C. Tan, 2023, "*The formation of supermassive black holes from Population III.1 seeds. II. Evolution to the local universe*", Monthly Notices of the Royal Astronomical Society, 525, 969-982.

3.1 PINOCCHIO SIMULATIONS

As in Paper I, to test our Pop III.1 seeding mechanism, we used the PINOCCHIO code (Monaco et al., 2002; Munari et al., 2017) to generate a cosmological box of 59.7 Mpc (40 h^{-1} Mpc for $h = 0.67$) with standard Planck cosmology (Planck Collaboration 2020) and study the formation of Dark Matter (DM) (mini-)halos in that box. PINOCCHIO uses Lagrangian Perturbation Theory (Lagrangian Perturbation Theory (LPT), e.g., Moutarde et al., 1991) to approximate the evolution of cosmological perturbations in a ΛCold Dark Matter (CDM) universe. This allows the code to run much faster compared to N-body codes, while also reproducing, with a good accuracy, the mass and accretion history of halos (Monaco et al., 2013; Monaco, 2016). For a given set of initial conditions, the code generates outputs in the form of catalogs at different redshifts, which contain mass, position and velocity of the DM halos, and a complete information of the merger histories of all the halos, with continuous time sampling. All these outputs are essential for efficiently implementing the seeding scheme (described in the next section), and for obtaining the exact redshift of mergers among seeded halos (discussed in detail in Chapter 5).

This code was written for applications in cosmology, where huge volumes with moderate mass resolution are requested, and its performance heavily depends on the mass resolution adopted. To resolve minihalos of $\sim 10^6 M_\odot$ it is necessary to sample a 59.7 Mpc box with 4096^3 particles; this results in a particle mass of $1.23 \times 10^5 M_\odot$, and we adopted a minimum mass of 10 particles (that would be unacceptable for an N-body simulation, but it is acceptable for a semi-analytic code like PINOCCHIO), resulting in a minihalo mass of $1.23 \times 10^6 M_\odot$.

Such a large simulation can only be run on a supercomputer, distributing the computation on a large number of Message Passing Interface (MPI) tasks. The construction of halos from collapsed particles is performed in Lagrangian space: the box is divided in sub-boxes, and the grouping algorithm is run on the the particles belonging to its domain. Halos near or across the sub-box borders would not be constructed correctly, so the sub-box is augmented with a "boundary layer" (a ghost region) whose size should scale with the Lagrangian radius R_{max} of the largest halo one expects to find in the simulation volume (that can be of the order of several Mpc). This implies an overhead in memory that can be significant. When dividing a small box into many tasks, the size of the sub-boxes can be of the same order of (if not larger than) R_{max}, making the memory overhead unsustainable. The constraint is weakened by stopping the simulation at higher redshift, when R_{max} is still small. As a result, with V4 of PINOCCHIO (Munari et al., 2017) used in Paper I, we were only able to push the simulation down to $z = 10$.

We use here the novel V5 of the code, that implements a number of numerical techniques to improve memory efficiency. This code will be presented elsewhere, the strategy to perform halo construction at high resolution is the following. The sub-box is augmented with a boundary layer as large as needed, but instead of storing the properties of all particles in the augmented sub-box we start by storing only the particles that lie in the sub-box (excluding the boundary layer) and are predicted to collapse by $z = 0$. Then the halo construction code is run once, collecting a tentative list of halos; after, all the particles that are in the boundary layer and lie within N_{Lag} times the Lagrangian size of any formed halos are added to the list of particles. After collecting the extra information, the halo construction code is run again, generating the final list of halos. Memory occupation thus depends on the parameter N_{Lag}; our tests show that $N_{Lag} = 3$ guarantees a convergent result, but an extreme run such as the one we present here was possible only by using $N_{Lag} = 2$. The 59.7 Mpc box with full 4096^3 resolution was thus run to $z = 0$ on 800 MPI tasks over 100 computing nodes (each with 256 GB of RAM), so the domain was divided into $6 \times 6 \times 7.5$ Mpc sub-volumes for halo construction. The resulting halo mass function showed two problems that are presented in greater detail in Appendix A. We discuss here their nature and their implications.

As a consequence of the difficulty of calibrating the formation of halos with a very steep power spectrum, the mass of the first halos is underestimated by a factor of ~ 2 at $z \sim 30$, decreasing to a negligible value at $z \sim 10$. This is a known trend in PINOCCHIO visible, e.g., in Figure 1 of Munari et al., 2017 where

the $z = 3$ halo MF is slightly underestimated in those tests. We are working to improve this prediction, but we do not consider this as a showstopper for several reasons: our seed BHs are already predicted to form very early, so this underestimation only causes us to be slightly conservative in their formation redshift, i.e., in fact they would already have formed at slightly higher z. In our simple modeling we are assuming here immediate formation of the protostar and then the SMBH, whereas in reality this might take several Myr or even tens of Myr. The time span that separating $z = 32$ from $z = 29$ is only ~ 14 Myr, so neglecting astrophysical timescales leads to an overestimation of formation redshift, which compensates against the underestimation problem. Finally, the minihalo threshold mass can be consider to be a second free parameter of the modeling (although one that has physical motivation to be close to $10^6 M_\odot$), so one can simply consider our predictions to be valid for minihalo masses of $2.5 \times 10^6\ M_\odot$. We add to these arguments the fact that inaccuracies in halo masses do not propagate as inaccuracies in halo positions, that are crucial outcomes of our seeding scheme.

A more serious problem is connected to the inaccurate reconstruction of halos more massive than $10^{12} M_\odot$. Indeed, the small size of the sub-box domain for constructing halos results in a poor reconstruction of massive halos. This problems makes predictions at $z = 0$ unreliable. We thus produced the same box at a lower resolution, sampled with 1024^3 particles, on a single MPI task on a 256 GB node. Again, this was possible thanks to V5 of the code. In this case halo construction is as good as it can be. However, the identification of halos that contain seed SMBHs has been performed in the high resolution box, and though the simulations share the same large-scale structure, matching massive halos in the two boxes is not a clean procedure. We then resorted to this algorithm: starting from the fact that one low-resolution particle contains 64 high-resolution ones, we calculated which particle in the lower resolution box includes the seeded mini-halo, and assigned the seed to the halo that contains that specific low-resolution particle. We checked that results at $z = 0$ produced with the low- and high-resolution simulations were consistent, with a significant difference in halo clustering of halos more massive than a certain threshold that is an expected consequence of the inaccurate mass reconstruction and the known relation of halo bias with halo mass. In the following we will present results at $z = 0$ based on the low resolution box, unless mentioned otherwise.

3.2 SEEDING SCHEME

To determine which halos are seeded with a Pop III.1 star and thence SMBH, consider the scenario depicted in Fig. 3.1, unfolding in the early universe. The figure shows three stars A, B and C in different halos where only A and C become Pop III.1 stars whereas B is a Pop III.2 star, depending on the separation and formation order. Star A formed first, which then influenced its environment within a sphere of radius equal to d_{feedback}, expected to be primarily radiative feedback. Since this star is in a pristine primordial gas without the influence of any feedback from nearby stars, it is defined to be

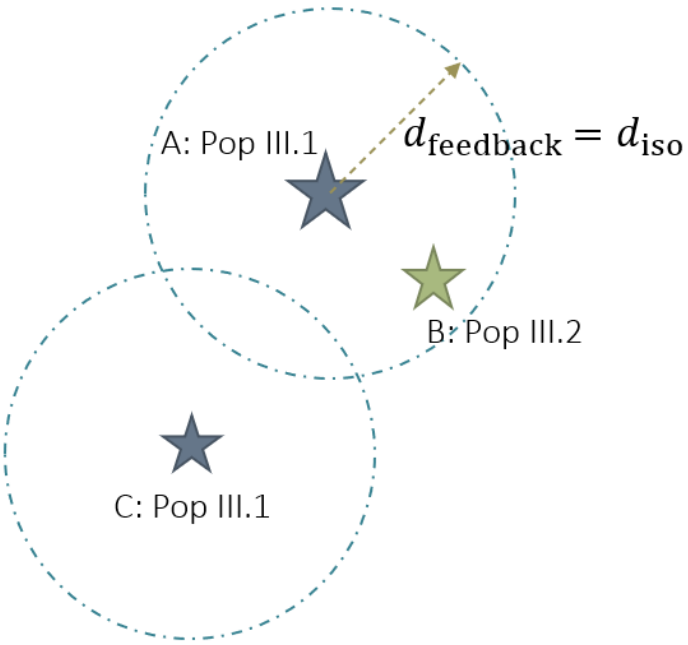


Figure 3.1: A schematic illustration of the Pop III.1 SMBH seeding scenario depicting the conditions for a star to be isolated enough to be considered as a Pop III.1 star (see text).

a Pop III.1 star. Star B, which subsequently forms at a distance less than d_{feedback} from star A, is affected by the feedback and hence is a Pop III.2 star (or even a Pop II star if it has been chemically polluted). Finally, star C forms beyond the regions affected by feedback from sources A and B, and is thus also assigned to be a Pop III.1 star and thus an SMBH. For the model considered here, the feedback distance is set equal to the isolation distance d_{iso}. So effectively, the condition for a star to be regarded as a Pop III.1 star is that when it is forming, there should be no previously formed halos present in the sphere of radius d_{iso}. We consider d_{iso} as a free parameter in our theory and vary it to match the observed number density of the SMBHs in the local Universe.

3.3 SEED IDENTIFICATION IN THE DARK MATTER CATALOGS

To perform the seed identification analysis from the dark matter catalogs generated by `PINOCCHIO`, we first divided the entire redshift range (from $z = 0$ to the redshift when the first minihalo forms, $z \approx 40$) into small bins of widths ranging from $\Delta z = 1, 2$ or 3, depending on the output catalogs available, which in turn depends on the relative change in positions of (mini)halos. The bins are wider at high redshifts, but smaller at lower redshifts. Then for each redshift interval $[z_l, z_h]$ where $(z_h > z_l)$, we utilised k-d tree data structure to create a three dimensional map in position space of all the halos existing between z_h and z_l. The positions used to create the tree are taken from the output catalog of `PINOCCHIO` at the lower redshift of the interval

(z_l). Since the positions are not updated once the tree is constructed, we account for the change in the positions within this redshift interval by finding the maximum change (δ) of position among all the halos existing for the entire redshift range. Then for each minihalo crossing the mass threshold of $10^6 M_\odot$ (or as in the nomenclature of `PINOCCHIO`: "appearing") at a redshift $z_{app} \in (z_l, z_h]$, we perform a ball search using the *k*-d tree to find all the halos around the appearing minihalo within a sphere of radius $d_{iso} - 2\delta$[1]. If there exists even a single halo at the redshift z_{app} within this sphere, then this minihalo is flagged as a halo containing a non-Pop III.1 star at its center. If there are no halos existing at this redshift, then the ball search is performed again with the same minihalo at the center, but this time within a sphere of radius $d_{iso} + 2\delta$. Then for all the halos existing at redshift z_{app} within the shell of radius $d_{iso} \pm 2\delta$, we find the exact distance between the minihalo at the center and all these halos using the exact positions at z_{app}. If this distance is greater than d_{iso} for all the halos within the shell, then the minihalo at the center is flagged as a Pop III.1 source, i.e., an SMBH-seeded halo. This process is repeated for each minihalo crossing the threshold mass within the two redshifts, and then this whole procedure is performed again for all the redshift intervals, until the whole redshift range is covered. In this way we are able check the isolation condition for each minihalo appearing in the cosmological box and find all the seeded minihalos.

At smaller redshifts, the change in positions of the halos (δ) within the redshift intervals becomes comparable to the isolation distance. This implies that the quantity $d_{iso} - 2\delta$ can become negative (in our simulation box, this happens at around $z \approx 15$ for $d_{iso} = 50$ kpc). In this case, the ball search is directly performed in a sphere of radius $d_{iso} + 2\delta$, and then the exact distances between the minihalo at the center and all the other halos existing at z_{app} is calculated.

This division of the entire redshift interval and creating the *k*-d only at specific redshifts is performed to avoid reconstructing the tree with the up-to-date position at every instance a new minihalo appears. Since the number of minihalos is very large, it becomes highly expensive computationally to reconstruct the tree with updated positions each time a new minihalo appears.

3.4 RESULTS

3.4.1 *Number density evolution*

As explained in the last section and in Paper I, we identify SMBH-seeded halos by the condition that the isolation sphere of radius d_{iso} around a newly forming minihalo is not populated by any other existing halo (of mass greater than our minihalo threshold mass). The obtained results for the evolution of number density for different values of d_{iso} (in proper distance units) are shown in Fig. 3.2. The colored dotted lines show the number density

1 A factor of 2 is multiplied with δ to account for the change in position of both the minihalo at the center of the sphere and all the other halos within the sphere.

evolution of total number of SMBHs, whereas the colored solid lines show the number density for seeded halos (which are slightly smaller, especially at lower redshifts, due to mergers). Compared to the number densities in Figure 1 of Paper I, the values obtained here are moderately lower (by a factor of 1.45 for 100 kpc and 1.65 for 50 kpc) because we have considered periodic boundary conditions when identifying the seeds, which was not done in Paper I.

Fig. 3.2 also shows some observational estimates of n_{SMBH}. An estimate at $z = 0$, presented in Paper I, is calculated by assuming that each galaxy with luminosity greater than $L_{min} = 0.33L_*$ hosts an SMBH, with the error bar around this point assuming a range of L_{min} from 0.1 to 1.0 L_*. Note, L_* is the characteristic luminosity corresponding to $M_B = -19.7 + 5\log h = -20.55$ (e.g., Norberg et al., 2002). Recent observations of high redshift AGNs from JWST surveys have started providing lower limits on the number density of SMBHs in the early universe, with one such estimate presented by Harikane et al., 2023, from the sample of Nakajima et al., 2023 (black diamonds in Fig. 3.2). Their estimate provides lower bounds on the observed number density of Type I AGN at redshifts $z = 4$ to 7.

From Fig. 3.2 illustrates the expected behaviour that as the isolation distance is reduced, the number of formed SMBHs increases, i.e., it is easier to satisfy the isolation distance criterion. We can also conclude that for a certain range of d_{iso} ($\approx$ 90 kpc to 170 kpc), the number density obtained is in reasonable agreement with the $z = 0$ estimate. Thus, the case with $d_{iso} = 200$ kpc is disfavoured simply by its inability to produce enough SMBHs. A key feature of the fiducial model, i.e., with $d_{iso} = 100$ kpc, is that *all* SMBHs have formed very early in the Universe: the process is essentially complete by $z \simeq 25$.

Fig. 3.2 also shows results for an example halo mass threshold (HMT) model (shown by green dashed line) in which each halo more massive than $m_{th} = 7.1 \times 10^{10} M_\odot$ is seeded (e.g., the Illustris project: Vogelsberger et al., 2014; Sijacki et al., 2015, etc.); note, this seeding scheme is driven by the mass resolution of the simulation, i.e., halos are seeded as soon as they are resolved with a sufficient number of particles). The main difference compared to the fiducial Pop III.1 model is in the overall number density of SMBHs at $z \gtrsim 5$.

We also show the results of a simulation by Chon et al. (2016) modeling the formation of SMBHs via the direct collapse mechanism. Here they simulated a $20h^{-1}$Mpc box and found only two SMBHs formed (at $z \simeq 15$ and 21). Even though this simulation was only run down to $z = 9$, the number density is not expected to increase much at lower redshifts, given the conditions assumed to be needed for direct collapse, i.e., massive, irradiated, tidally-stable, metal-free halos. While this model allows some SMBHs to form relatively early, as discussed in Chapter 1, the overall number densities achieved by this mechanism are much smaller than are needed to explain the entire observed SMBH population.

We quantify the number of mergers that occur between seeded halos in the Pop III.1 models. Table 3.1 shows the total number of SMBHs that formed ($N_{SMBH,form}$) and the number of halos containing them at $z = 0$ ($N_{SMBH}(z = 0)$). Assuming efficient merging of SMBHs that are in the same

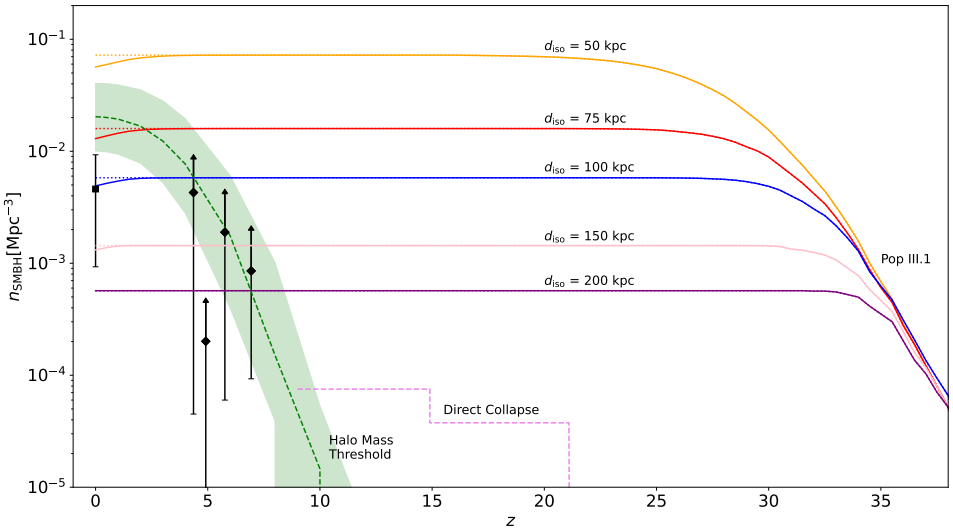


Figure 3.2: Evolution of the comoving number density of SMBHs, n_{SMBH}, for different theoretical models. Results for Pop III.1 models with several values of isolation distance (in proper distance) are shown, as labelled. The dotted lines show the total number of SMBHs that ever formed, while the solid lines show remaining number of seeded halos after accounting for mergers. An example Halo Mass Threshold (HMT) model is shown by the dashed green line in which each halo with mass higher than $m_{th} = 7.1 \times 10^{10} M_{\odot}$ is seeded (see text). The green shaded region shows the effect of lowering and raising m_{th} by a factor of 2. The violet dashed line shows the results of a simulation modeling SMBH formation via direct collapse (Chon et al., 2016). The black solid square indicates an estimate for the number density of SMBHs at $z = 0$ assuming each galaxy with luminosity higher than $L_{min} = 0.33 L_*$ contains one SMBH, with the range shown by the error bar obtained by varying L_{min} from $0.1 L_*$ to L_*. The black diamonds are estimated lower limits of n_{SMBH} from JWST observations of Type I AGN (Harikane et al., 2023).

Table 3.1: Total number of formed SMBHs ($N_{SMBH,form}$), total number of SMBHs remaining at $z = 0$ assuming efficient mergers ($N_{SMBH}(z = 0)$), the difference between these ($\Delta N_{SMBH} = N_{SMBH,form} - N_{SMBH}(z = 0)$), which is equivalent to the number of mergers, and the fraction of original SMBHs that are destroyed by mergers ($f_{merger} = \Delta N_{SMBH}/N_{SMBH,form}$).

d_{iso} [kpc]	$N_{SMBH,form}$	$N_{SMBH}(z = 0)$	ΔN_{SMBH}	$f_{merger}(\%)$
50	15356	12051	3305	21.52
75	3394	2760	634	18.68
100	1234	1043	191	15.48
150	306	280	26	8.50
200	121	116	5	4.13

halo, then the number of mergers is $\Delta N_{SMBH} = N_{SMBH,form} - N_{SMBH}(z = 0)$. A feature of the Pop III.1 seeding mechanism is that SMBHs are initially spread out from each other, so that there are relatively few binary SMBHs and few mergers. A detailed analysis of the mergers including the binary AGN number densities, and the gravitational wave background emanating from these mergers is discussed in later chapters of this thesis.

A caveat of our seeding model is that at small redshifts, around $\lesssim 6$, the isolation distance in comoving units becomes so small that many minihalos that appear after this redshift start satisfying the isolation criteria. This effect would result in an increase in number density by around 2 orders of magnitude by $z = 0$ from the converged values around $z \approx 20$, for all cases of d_{iso}. However, since reionization has completed by $z \approx 8$ (Planck Collaboration 2020), we assume that the formation of Pop III.1 sources is also not possible below this redshift. Hence, in our analysis, we set a limit of seed formation to be only possible until $z = 8$. For most cases of the isolation distances we considered ($\geqslant 75$ kpc), the number density is already converged at redshifts greater than $z = 20$. However, for the case of 50 kpc, new seeds still keep on appearing until $z = 8$ (although below $z = 15$ the total number only increases by about 1%).

In Figure 3.3, we show a visual representation of the seeded halos in the box at different redshifts, for all the isolation distances considered in Fig. 3.2. As discussed, the 50 kpc case is the most crowded with the highest number of seeded halos at every epoch shown. Initially all the seeds emerge in a relatively unclustered manner, but eventually the clustering increases as lower-mass seeded halos migrate towards more massive halos and merge with them in overdense regions. We perform a more detailed analysis of clustering in §3.4.3.

3.4.2 *Occupation fraction of seeded halos*

From observations of local galaxies, it appears that almost all massive galaxies contain a nuclear SMBH. This implies that the SMBH occupation fraction

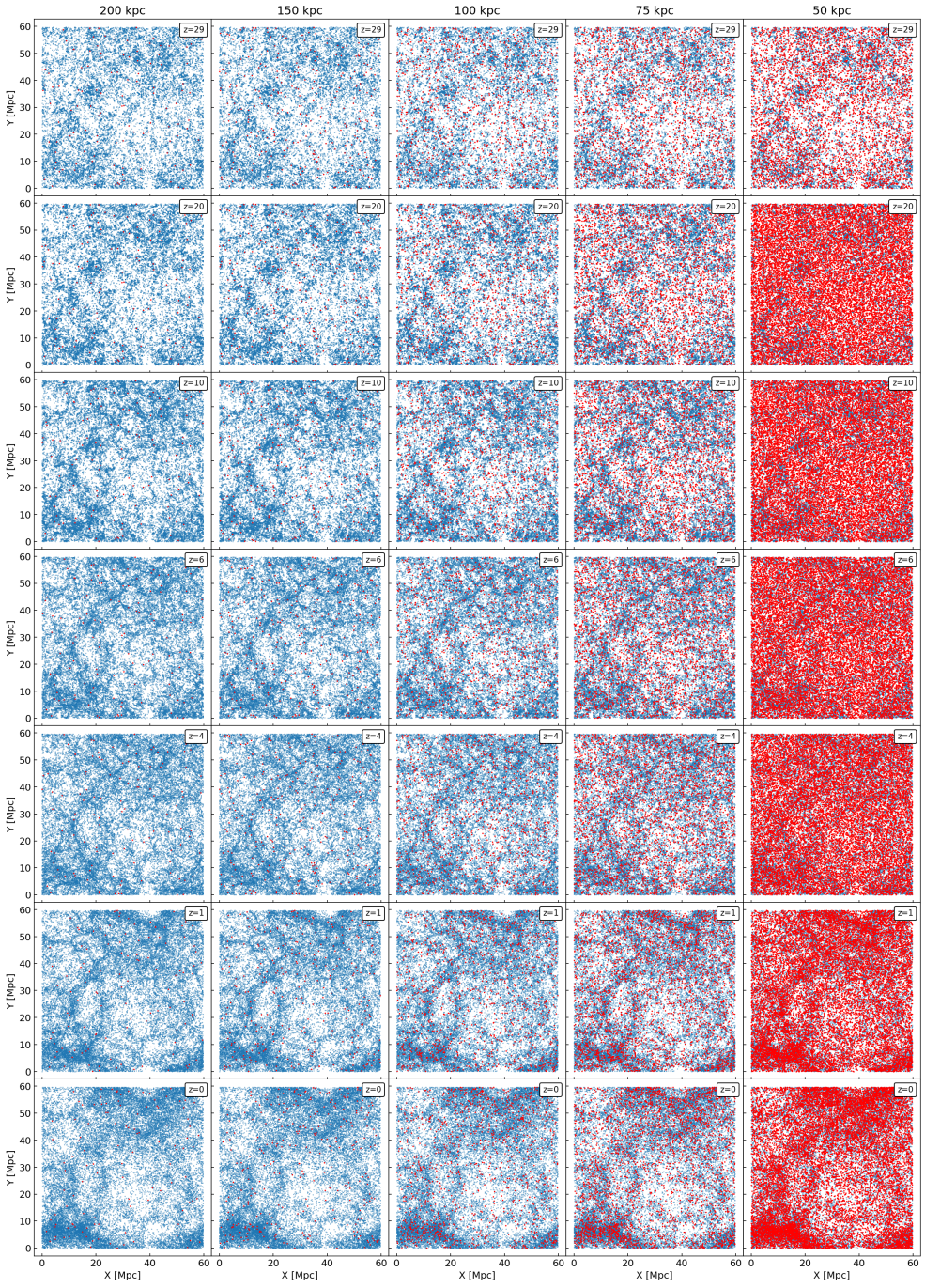


Figure 3.3: Projection of the positions of seeded halos (*red*) and non-seeded halos (*blue*) in the XY plane of the box for different isolation distances. The redshift is shown in the top right corner of each panel (same for each row). Only the 30,000 most massive non-seeded halos within each panel are shown for ease of visualisation.

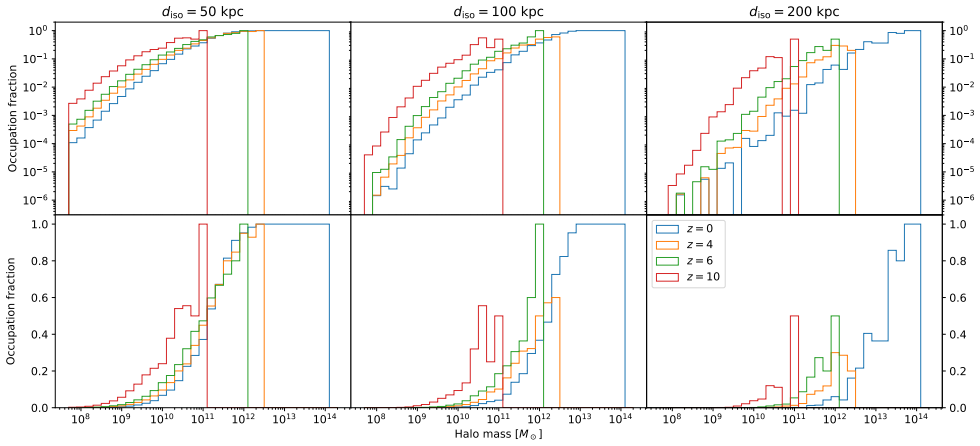


Figure 3.4: Evolution of SMBH occupation fraction of halos for different cases of d_{iso}. Top row depicts the fraction in log scale, while the bottom row shows the same data in linear scale. The mass bins are divided into equal bins of width 0.2 dex.

of halos should approach unity as halo mass rises. Figure 3.4 shows the evolution of occupation fraction from one realization of our 59.7 Mpc box, through 4 different redshifts for halos ranging from $[5 \times 10^7, 2 \times 10^{14}]M_\odot$ (the upper limit of the mass range is chosen to include the most massive halo at $z = 0$ in our 1024^3 resolution simulation box, measuring $1.2 \times 10^{14}M_\odot$). As expected, with the decrease in the isolation distance, more and more halos are seeded and hence the occupation fraction is higher compared to the same mass range for larger d_{iso}. All the fractions at $z = 0$ approach unity for the most massive halos, independent of the isolation distance. Interestingly, the most massive halo is not always occupied by an SMBH throughout the redshift evolution in our simulations. For example, at $z = 4$ there can be significant fractions of the most massive halos, i.e., $\sim 10^{12}M_\odot$, that are not seeded, as in the case of $d_{iso} = 100$ kpc. Figure 3.4 also shows that for $d_{iso} = 200$ kpc the occupation fraction for halos with masses $\sim 10^{12}M_\odot$ at $z = 0$ is quite small, $\lesssim 0.1$, which is a further indication that it produces too few SMBHs.

Figure 3.5 shows the evolution of the cumulative occupation fraction, i.e., for all halos more massive than $\{10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}\}M_\odot$, for three different cases of isolation distance. If we consider only the most massive halos ($> 10^{13}M_\odot$), the fraction is close to one (as also evident from Fig. 3.4). At a given redshift, as we consider less massive halos, the occupation fraction decreases. At a given mass threshold, as we move out to higher redshift the occupation generally rises, since these halos become relatively more extreme members of the global halo population. Interestingly, the occupation fraction for all halos more massive than 10^8 and $10^9M_\odot$ ($10^{10}M_\odot$ as well, although to a lower degree) at $z = 0$ differ by factors of approximately 10 among the three cases of isolation distances considered, reflecting the same differences in the global number densities at $z = 0$ (see Fig. 3.2).

To obtain a better understanding of the mass function of the seeded halos, in Figure 3.6 we present the distribution functions of these halos for the

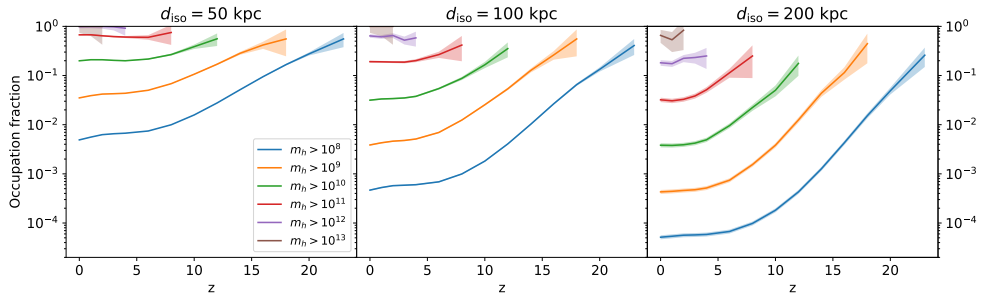


Figure 3.5: Cumulative occupation fractions of halos having masses greater than a given value (see legend). The shaded region represents $\pm 1\sigma$ error due to counting statistics.

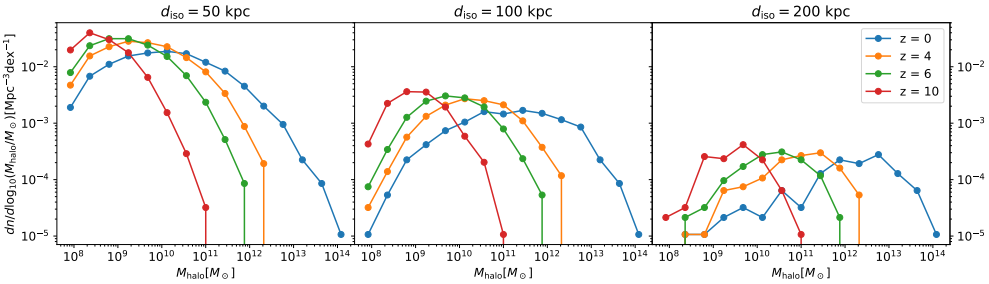


Figure 3.6: Mass function of seeded halos at different redshifts for $d_{iso} = 50$, 100 and 200 kpc cases (left to right).

$d_{iso} = 50$, 100 and 200 kpc cases, including their evolution with redshift. We see that, as expected, these mass functions evolve to higher masses as the universe evolves from $z = 10$ down to $z = 0$. The peak of the seeded halo mass function is lower for smaller values of d_{iso}. However, the distributions are quite broad, indicating significant fractions of SMBHs in relatively low-mass halos, even at $z = 0$. In a future paper in this series (Cammelli et al., in prep), these seeded halo mass functions and the properties of their host galaxies will be compared to SMBH census data, especially focusing on properties derived in the local universe.

3.4.3 *Clustering*

We perform a clustering analysis using the `CORRFUNC` library (Sinha and Garrison, 2020) for `PYTHON`, and the results are shown in Fig. 3.7. By sampling r in 20 logarithmic bins of $r_{min} = 0.5$ Mpc/h to $r_{max} = 13.3$ Mpc/h, we evaluate the 3D 2-point correlation function[2] (2pcf) $\xi_{hh}(r)$ for all halos more massive than $10^{10} M_\odot$ at $z = 0$. Since `PINOCCHIO` only evolves dark matter halos, the information of substructures such as subhalos within halos is not stored or tracked. This implies that only radial scales greater than the size of a typical dark matter halo (3 to 4 Mpc at $z = 0$), are relevant for

2 All the correlation functions presented in this section have been corrected by analytically adding large scale clustering modes corresponding to scales larger than the box size. Refer to appendix B for more details.

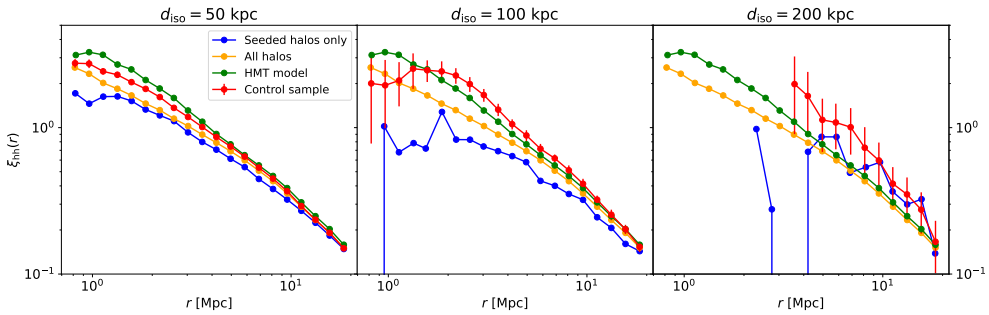


Figure 3.7: The 3D 2-point correlation function for the seeded halos more massive than $10^{10}M_\odot$, at $z = 0$ for different isolation distances. The blue points show the correlation function for only the halos containing SMBHs, while the orange points show the correlation for all the halos, with or without an SMBH. For the red points, we randomly select halos from the pool of all the halos, but with the same number and mass distribution as the seeded halos. The error bars indicate 1σ deviations from the mean value from randomly sampling 50 times. The green points show the correlation for halos seeded according to the halo mass threshold (HMT) scheme, in which all the halos greater than $m_{th} = 7.1 \times 10^{10}M_\odot$ are seeded.

consideration. In other words, the correlation function presented here does not include the one-halo term. From the figure, we observe that the clustering of the SMBH-seeded halos (blue points) is always lower compared to other cases. This is expected because of the nature of our model, which results in larger distances between SMBHs and hence smaller clustering amplitude. The plots for $d_{iso} = 50$ and 100 kpc clearly depict this, while the case of 200 kpc suffers from low number statistics. The red points, which represent the clustering of random halos with the same number and mass distribution as of the seeded halos, are generally more than 1σ higher than the blue points, except at the largest scales. This can be clearly seen for the fiducial case of 100 kpc. We also show the clustering for the fiducial case of HMT schemes with $m_{th} = 7.1 \times 10^{10}M_\odot$ (Sijacki et al., 2015), depicted by green points. This model also generally shows higher clustering than our Pop III.1 seeding model. Thus a clustering analysis of census of a local Universe ($z = 0$) survey of all (or a significant fraction) of SMBHs has the potential to distinguish between these SMBH seeding mechanisms.

In Figure 3.8, we show the evolution of the projected correlation function for the $d_{iso} = 50$ and 100 kpc cases (blue lines), compared to halos with the same mass and number distribution as the respective seeded halos (red lines). As seen in the 3D 2pcf, the clustering of the seeded halos is always lower than the randomly selected halos and this trend is observed even at higher redshifts. Furthermore, there is a significant drop of the clustering amplitude of the seeded halos for scales lower than $d_{iso}(\bar{z}_{form})$ (vertical grey band), a signature of feedback cleared bubbles, first discussed in Paper I for $z \geqslant 10$. Here we see that this signature of suppressed clustering persists to lower redshift, although is gradually diminished as the Universe evolves to a more clustered state.

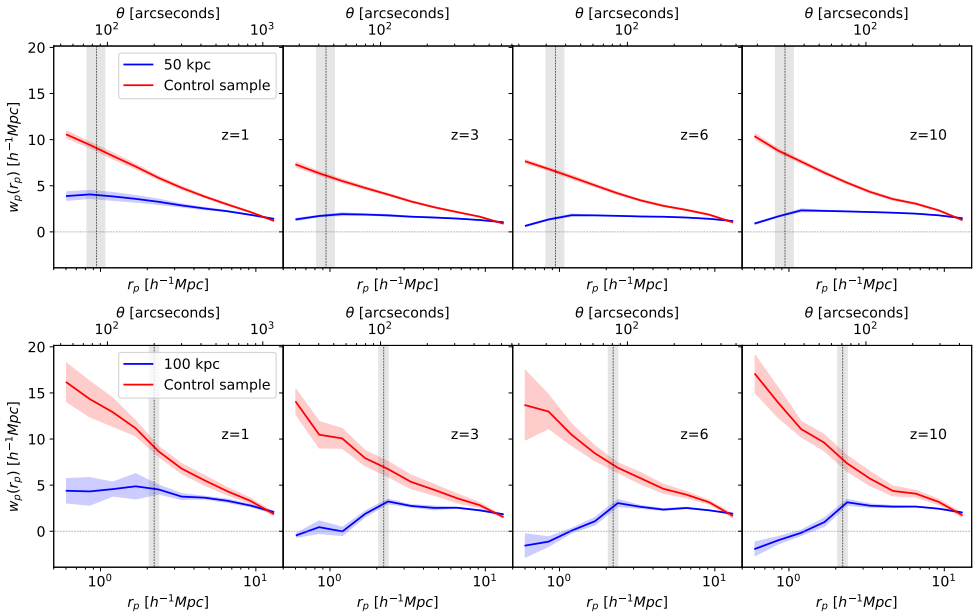


Figure 3.8: Evolution of projected correlation function for $d_{iso} = 50$ kpc (top row) and 100 kpc (bottom row) cases. The blue line is the average after computing the correlation of the seeds from 3 orthogonal sides of the box and the shaded region represents the 1σ spread. The control sample is the correlation of halos selected randomly but with the same mass and number distribution as the seeded halos at that redshift. The red line refers to the average after randomly sampling 10 times and the shaded region refers to 1σ deviations from the mean. The vertical grey line refers to the size of the isolation radius at the mean formation redshift ($d_{iso}(\bar{z}_{form})$) of the seeded halos, and the grey region represents 1σ deviation from the mean. For 100 kpc, $\bar{z}_{form} = 32.08$, and for 50 kpc, $\bar{z}_{form} = 27.14$. The angular axis on top of each panel corresponds to the angular scale of r_p projected on the sky at the respective redshift.

We emphasise that comparing our clustering predictions at redshifts greater than 1 or 2 is not feasible with currently available observational data. The measurements from a range of luminosity of AGNs at these redshifts imply minimum halo masses of $\sim 5 \times 10^{11} h^{-1} M_\odot$ at $z \sim 3$ (Allevato et al., 2014) to more than $10^{12} h^{-1} M_\odot$ at $z \sim 4$ (He et al., 2018). For our 59.7 Mpc box, the number of seeded halos above these thresholds are quite low. For instance, for the $d_{iso} = 100$ kpc case, only around 6% of sources are above this threshold at $z = 3$ and only 0.7% sources are more massive than $10^{12} h^{-1} M_\odot$ at $z = 4$. If we apply these halo mass cuts on our seeded halos, then the clustering signal is too noisy to make any decent comparison with the observational data. Moreover, at high halo masses the occupation fraction approaches unity, so for the measured clustering of bright AGNs, hosted in relatively massive halos, we expect that they may cluster as their host halos, with no appreciable difference with respect to currently used models. More data on AGN, especially those that are present in lower-mass halos/galaxies is needed to test the models.

As a crude comparison, in Figure 3.9 we include the clustering measurements from Zehavi et al. (2011), who performed the projected clustering analysis of volume-limited sample of 570,000 galaxies from the Seventh Data Release (Abazajian et al., 2009) of the Sloan Digital Sky Survey (Sloan Digital Sky Survey (SDSS), York et al., 2000). The galaxies used in their data extend out to $z = 0.25$, with a median redshift of $z \sim 0.1$. We compare our results at $z = 0$ for $d_{iso} = 50$ and 100 kpc, along with the HMT scheme, with their galaxy luminosity threshold cut result for $M_r < -19.0$. We computed the relation between DM halo mass and r-band absolute magnitude by comparing the clustering amplitude of `PINOCCHIO` DM halos with Zehavi et al.'s measurements, minimising the χ^2 of the clustering amplitude only for $r_p > 3 h^{-1}$ Mpc (to avoid the one-halo clustering scales); for $M_r < -19.0$ we find a clustering-matched halo mass of $M_{PIN}^{-19.0} = 1.91 \times 10^{12} h^{-1} M_\odot$, higher than the value suggested in that paper ($M_{zehavi}^{-19.0} = 2.55 \times 10^{11} h^{-1} M_\odot$); this is not surprising, given the different cosmology assumed in 2011. We then applied this halo mass cut on our $d_{iso} = 50$ and 100 kpc sources, as well as the HMT scheme, and compared the projected correlation function for the $M_r < -19.0$ threshold galaxies in Figure 3.9. For the region of interest, the clustering of the seeded halos shows good agreement, within the errors, with the observations. The $d_{iso} = 50$ kpc correlation completely overlaps the HMT one because all the sources more massive than $M_{PIN}^{-19.0}$ are seeded in this model. Also, at this high-mass cut, most of the $d_{iso} = 50$ kpc sources are also seeded in the $d_{iso} = 100$ kpc model, and hence their clustering follows similar trends. This is due to the fact that the occupation fraction approaches unity for the most massive halos (see §3.4.2) for all the isolation distances, and since the mass cut is high, this means that most, if not all, the halos are seeded, regardless of the isolation distance.

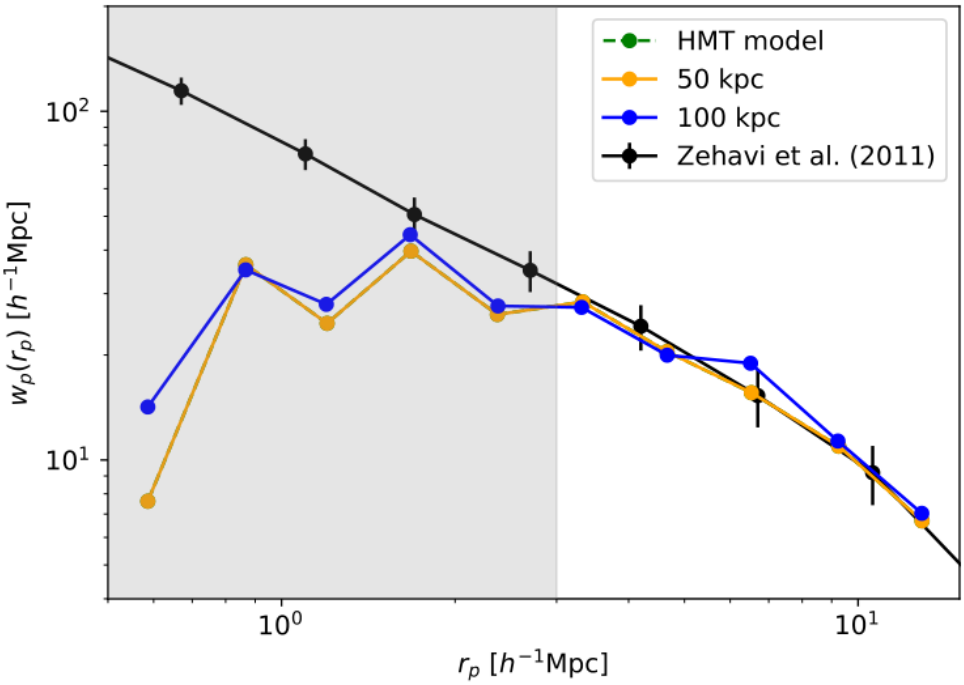


Figure 3.9: Comparison of the results for the projected correlation function $w_p(r_p)$ obtained from our simulations for $d_{iso} = 50$ kpc, 100 kpc and the HMT scheme at $z = 0$ with the observational data from Zehavi et al. (2011) for a $M_r < -19.0$ magnitude cut. The shaded region shows scales smaller than the size of a typical halo at $z = 0$, i.e., $r_p < 3h^{-1}$ Mpc, which are not of interest for our comparison due to limitations of our model (lack of sub-halos). The HMT scheme and 50 kpc models overlap, as all halos above the threshold are seeded for that value of d_{iso}.

4

ULTRA DEEP FIELD

As discussed in the previous chapter, the Pop III.1 model predicts a high number of SMBHs at high redshifts compared to other models. Most of the seeds are already in place by redshift $z \sim 25$ for the fiducial isolation distance. This implies that one potential way to investigate different seeding mechanisms is to calculate the number density of SMBHs at high redshifts. Recent results from JWST surveys have already started to shed light on the high-z SMBH population (Harikane et al., 2023; Maiolino et al., 2023b), as discussed in 3.4.1.

Another interesting avenue to target the high redshift population is provided by the Hubble Ultra Deep Field (HUDF, Beckwith et al., 2006; Ellis et al., 2013). The original HUDF, shown in figure 4.1, was released in 2004 using the then newly installed Advanced Camera for Surveys (ACS) in 2002, and the Near Infrared Camera and Multi-object Spectrometer (NICMOS) (Beckwith et al., 2006). Located in the constellation Fornax, this image shows more than 10,000 objects with most of them galaxies. The million second exposure allowed the instrument to detect light from galaxies in the early universe, corresponding to redshifts even higher than 7. The deep exposure of the field presents the possibility to detect faint AGNs at high redshifts, which would be powered by SMBHs. The redshift space distribution of the AGNs can thence provide valuable constraints on different seeding mechanisms.

In this chapter, we explain the methods used to create a synthetic ultra deep field with different seeding schemes, using the simulation volume described in the previous chapter. We then present the predictions for the number of sources which could be potentially observed in the UDF. Finally, we briefly mention our successful observational proposals to re-image the HUDF to search for high redshift AGNs. The content of this chapter presents the results from the paper published in the Monthly Notices of the Royal Astronomical Society:

- **Jasbir Singh**, Pierluigi Monaco, Jonathan C. Tan, 2023, "*The formation of supermassive black holes from Population III.1 seeds. II. Evolution to the local universe*", Monthly Notices of the Royal Astronomical Society, 525, 969-982.

4.1 SYNTHETIC FIELD CONSTRUCTION

To construct the synthetic field, we first need to calculate the total volume of the field, since this can give us an idea of the number of `PINOCCHIO` boxes of 40 Mpc/h side length we need. So in order to calculate the volume of the field, let us assume that the field is placed symmetrically along the x-axis, as shown schematically in figures 4.2 and 4.3. The field extends from

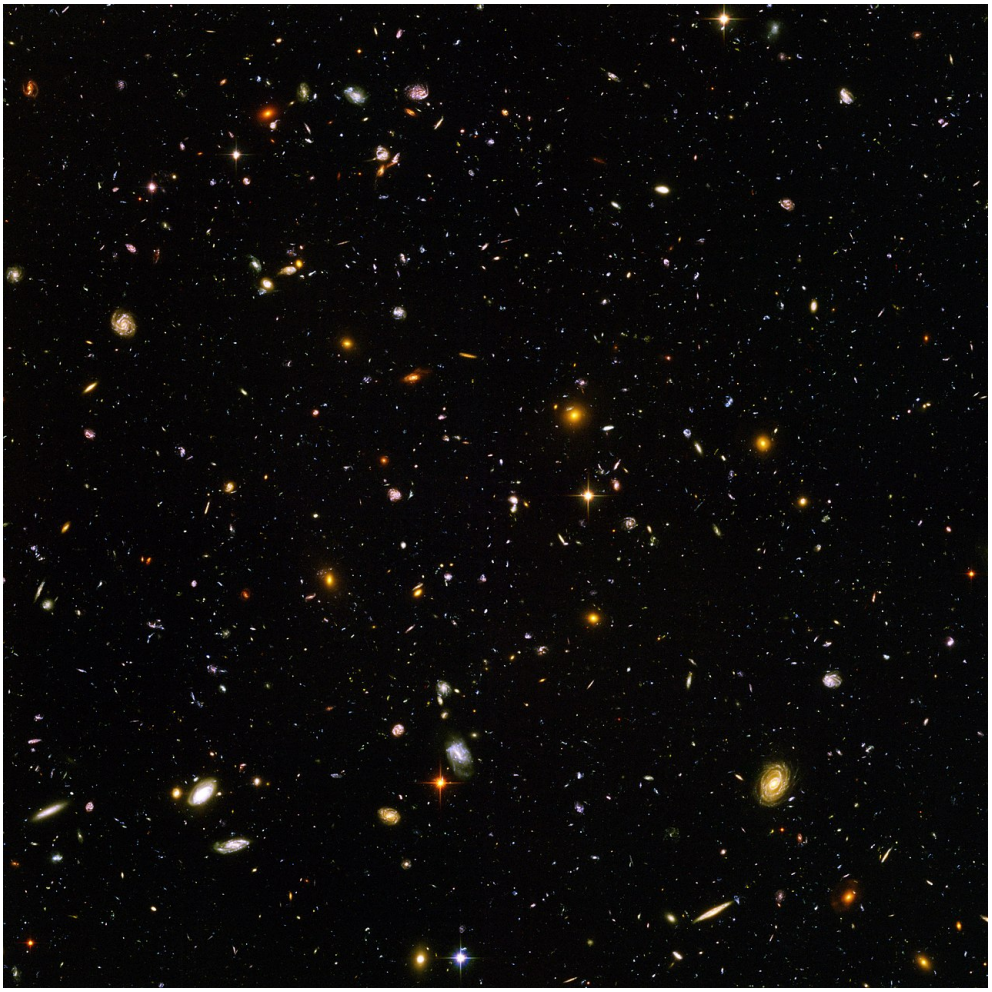


Figure 4.1: Hubble Ultra Deep Field. Credits: Beckwith et al., 2006

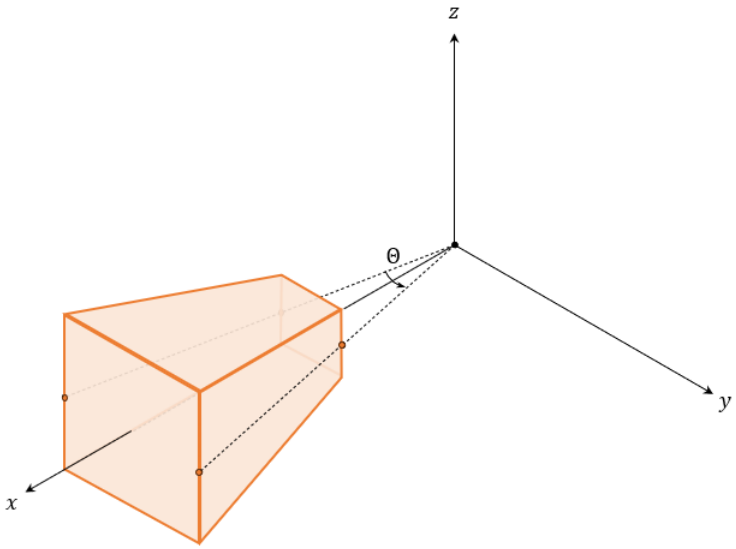


Figure 4.2: Isometric view of the UDF volume, placed on the x-axis.

the observer to the particle horizon, but we are interested in the comoving volume from r_i to r_f, which denote the comoving radial distance to the initial and final redshifts of the field respectively. The opening angle of the field is denoted by Θ, and is assumed to be the same for both the polar and azimuthal angles. Assuming spherical coordinates, the volume of the field can be calculated by integrating the differential volume element. From the schematic figures of the field mentioned above, it is straightforward to obtain the limits of integration for all the axes: the radial coordinate $r : r_i \to r_f$, the azimuthal angle $\phi : -\Theta/2 \to \Theta/2$, and the polar angle (with respect to the z-axis) $\theta : \pi/2 - \Theta/2 \to \pi/2 + \Theta/2$. After integrating, we obtain the general expression for the volume:

$$V = \frac{2}{3}\Theta \sin\frac{\Theta}{2}\left(r_f^3 - r_i^3\right). \tag{20}$$

If we are aiming for a field ranging from $z_i = 4$ to $z_f = 16$, then $r_i = d_c(z = 4) = 4,917.34$ Mpc/h and $r_f = d_c(z = 16) = 7,081.14$ Mpc/h. For the opening angle, considering $\Theta = 2.4$ arcminutes, or 6.98×10^{-4} radians. Using these dimensions, the volume of the field using equation 20 turns out to be 38,367.78 $(\text{Mpc/h})^3$. This volume is smaller than the volume of our box, which is 64,000 $(\text{Mpc/h})^3$, implying that the field can be constructed by using one single box.

Next, we can start constructing the field using the snapshots of the box captured between redshifts z_i and z_f, slicing them into *pencil beams* and stacking them together end to end. To determine the size of the beams, we need to consider the (comoving) side length of the field at the initial and final redshifts. At $z = 4$, the side length is equal to $d_c(z = 4) \times \Theta = 3.43$ Mpc/h, and 4.94 Mpc/h at the final redshift $z = 16$. Since we are using a 40 Mpc/h side box, we can take the snapshot of the box and divide it into 25 pencil beams of $8 \times 8 \times 40$ Mpc/h dimensions, as shown in the figure 4.4. This size of the beams will ensure that the entire synthetic field will be encompassed

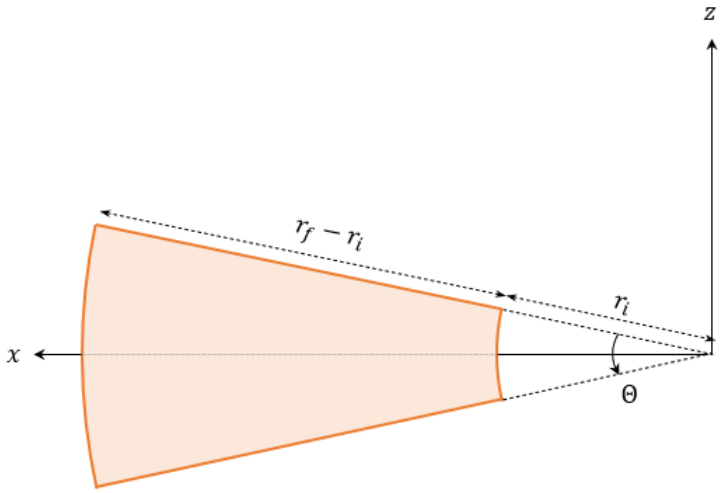


Figure 4.3: The projection of the UDF volume on the $x - z$ plane.

by the pencil beams up to $z = 16$, and even beyond. Then selecting the beams in random, we stack them next to each other to cover the entire range from z_i to z_f, as illustrated in figure 4.5. To include more randomness and reduce repetitions of halos, the pencil beams are created from all three orthogonal sides of the box, and then rotated and combined together to make one long beam. The figure 4.5 shows the configuration of some of the pencil beams through which the synthetic field is carved, assuming the observer is at $x = 0$. But before this is done, the position and redshift of all the halos inside the pencil beams when they enter the light cone are calculated by solving for z in the equation

$$|\vec{x}(z)| = d_c(z) \tag{21}$$

for each halo. Here $\vec{x}(z)$ represents the position of the halo from the snapshots after placing them in the pencil beam. Once the position of all the halos in the light cone is determined, then all the halos lying inside the field with opening angle Θ are selected to construct the synthetic ultra deep field. The next section presents the synthetic field obtained for different isolation distances and their results.

4.2 RESULTS AND DISCUSSION

Using the method described in the previous section, we create the synthetic fields of 2.4 arcminutes side for different SMBH seeding criteria. In figure 4.6, we show the synthetic UDFs of only the seeded halos for the Pop III.1 model with $d_{iso} = 50$ and 100 kpc for the redshift range $z \in [4.0, 16.0]$. These fields are from the highest resolution run of `PINOCCHIO` with 4096^3 particles. As expected, the field for $d_{iso} = 50$ kpc is much more densely populated with seeded halos compared to 100 kpc. For these fields, the number of halos present for the 50 kpc case is more than 12 times the halos in the 100 kpc case. This difference in the number of halos is also reflected through the number density evolution, which is discussed in §3.4.1.

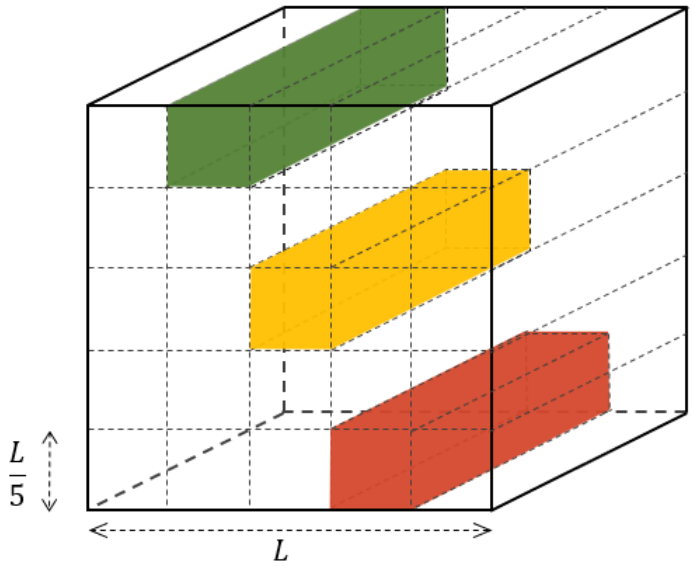


Figure 4.4: The snapshot of the box split into 25 pencil beams, with 3 highlighted in different colors. Here L is the size length of the box, equal to 40 Mpc/h.

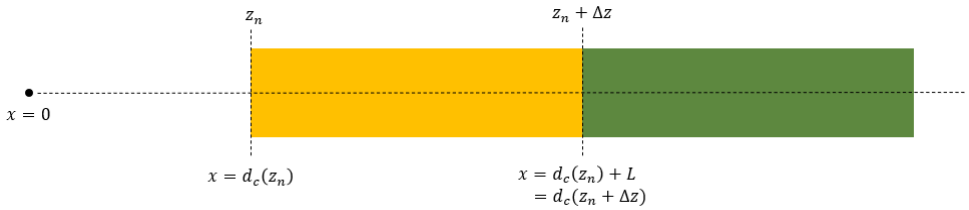
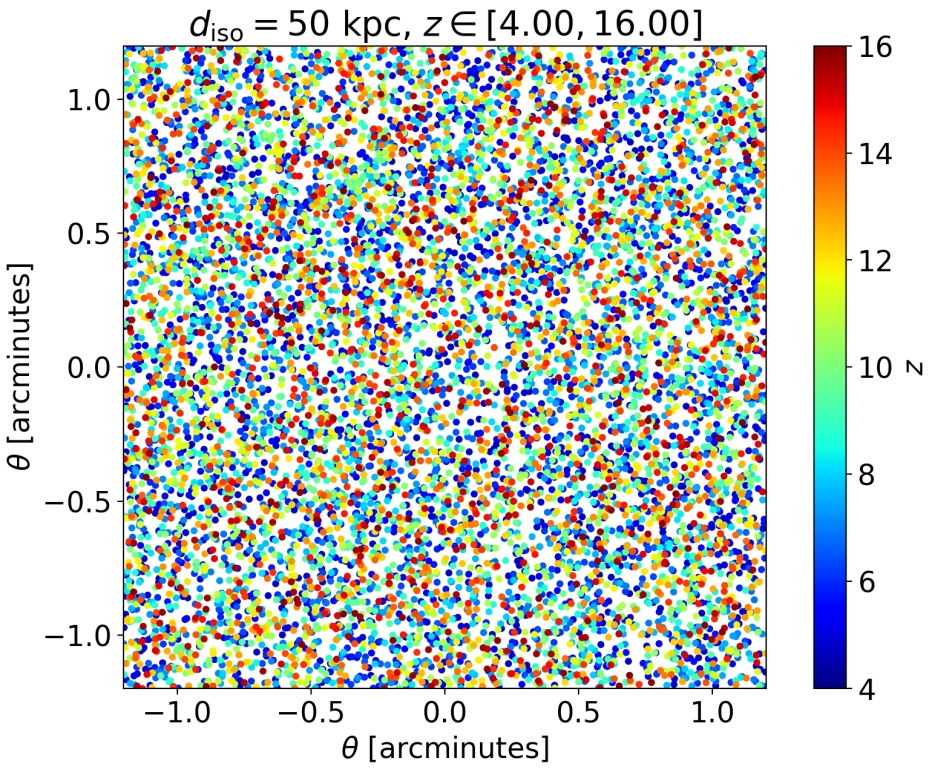
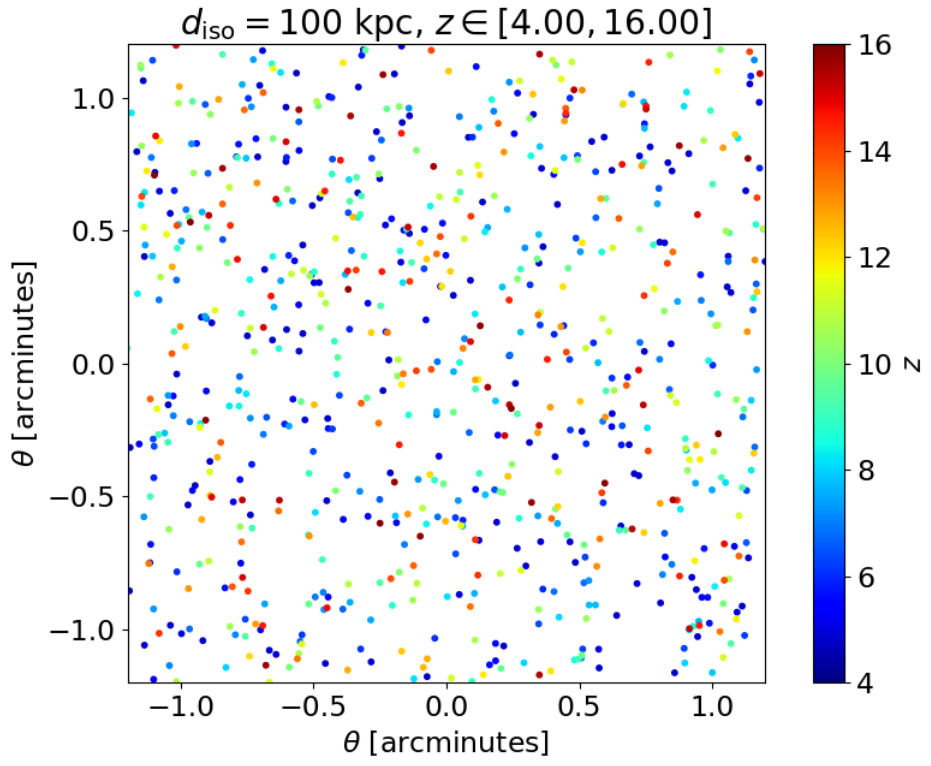


Figure 4.5: The side view of two pencil beams stacked, which are randomly selected. The observer is at $x = 0$ and z_n corresponds to the redshift of the snapshot. Δz corresponds to the size to the length of the box L in redshift space. Since they are symmetrically placed on the x-axis, this figure also represents the top view of the beams.



(a) 50 kpc



(b) 100 kpc

Figure 4.6: Synthetic HUDF consisting of only the seeded halos for d_{iso} = 50 kpc and 100 kpc cases over a redshift range from 4 to 16.

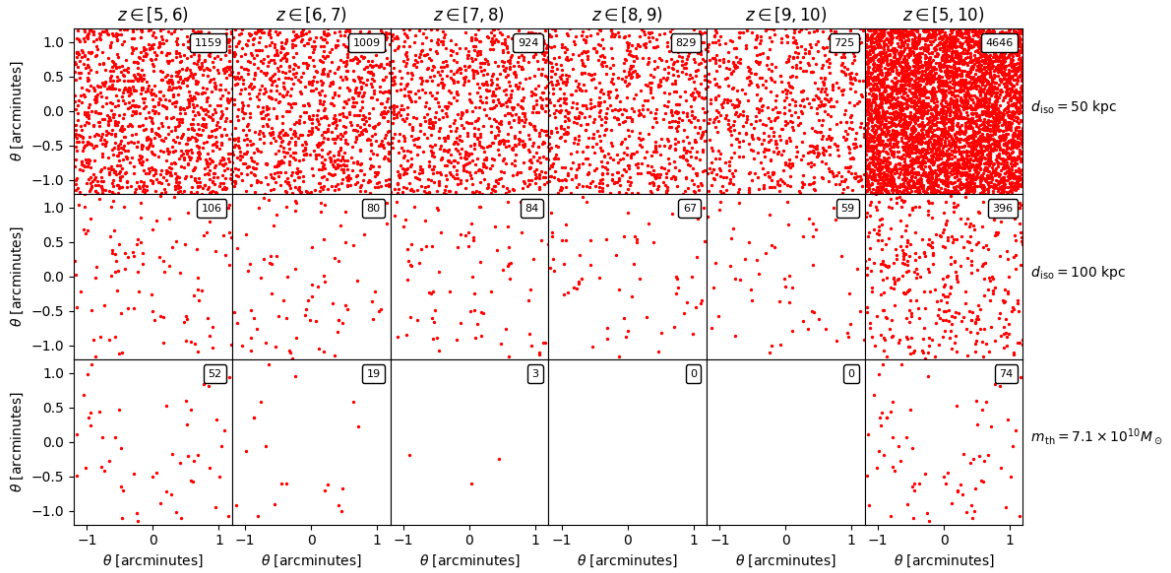


Figure 4.7: The distribution of SMBHs in redshift intervals in the range $z = 5 - 10$ in a synthetic HUDF, where the last column shows all the sources. The first row shows the case for d_{iso} =50 kpc. The second row shows the case for d_{iso} =100 kpc. The third row shows the distribution from the fiducial HMT scheme with $m_{th} = 7.1 \times 10^{10} M_\odot$. The total number of SMBHs in each panel are indicated in the top right corners of each.

To further investigate the distribution of the seeded halos in the redshift space in the synthetic field, we present figure 4.7 which shows the number of seeded halos within the redshift range $z = 5 - 10$ in bins of $\Delta z = 1$. The top two rows show the seeded halos from the Pop III.1 model, whereas the last row shows the seeded halos from the HMT numerical scheme with $m_{th} = 7.1 \times 10^{10} M_\odot$. The total number of sources in the field (*last column*) for the fiducial d_{iso} =100 kpc model is five times higher than the fiducial HMT scheme. From the observational perspective, these seeded halos would be hosting luminous galaxies at their center with AGNs which may or may not be passive. Even accounting for the obscuration factor and the duty cycle, the large difference in the number of SMBHs observed would still be significant enough for different isolation distances or seeding models, especially at higher redshifts. As shown in this figure and the figure 3.2, large differences in the number density at higher redshifts ($\gtrsim 6$ or 7) implies that even detecting a handful of sources at these redshifts would help constrain different seeding mechanisms. In the following section we describe some observational campaigns which will help in detecting AGNs in the HUDF.

Due to the inherent randomness in the division of the simulation box and the stacking order of the pencil beams, we created multiple realisations of the field from the same box to improve the statistics on the number of seeded halos in different redshift bins. Table 4.1 shows the number of seeded halos in the field for an extended redshift range by averaging over 100 random realisations of the light cone, and also by integrating the global number density obtained from the box over the field volume. Almost all the values

Table 4.1: Number of seeded halos in our synthetic HUDF, calculated by averaging over 100 random realizations of the light cone (*From light cone* column) and by integrating the global number density (*From number density* column) over the redshift ranges, for $d_{iso} = 100$ kpc and the fiducial HMT scheme with $m_{th} = 7.1 \times 10^{10} M_{\odot}$. The errors on the averaged values correspond to 1σ deviations. Note that all the numbers are rounded to the nearest integer.

z range	100 kpc		HMT	
	From light cone	From number density	From light cone	From number density
4-5	110 ± 8	101	86 ± 19	105
5-6	92 ± 6	90	36 ± 10	49
6-7	85 ± 5	81	13 ± 6	18
7-8	74 ± 6	73	3 ± 2	7
8-9	69 ± 5	66	1 ± 1	1
9-10	60 ± 5	60	0	0
10-11	57 ± 5	54	0	0
11-12	50 ± 5	50	0	0
12-13	47 ± 4	46	0	0
13-14	42 ± 4	43	0	0
14-15	40 ± 5	40	0	0
15-16	40 ± 5	37	0	0

calculated from the number density in the box lies within the 1σ dispersion of the values from averaging over the 100 light cones, implying a stable and realistic variation of the number of seeded halos as expected from cosmic variance. Furthermore, the numbers from the integration also reflect the drastic difference in the number of sources in the different seeding schemes at high redshifts.

4.3 OBSERVATIONAL CAMPAIGNS

As mentioned before in this chapter, the HUDF provides the perfect window to the high redshift universe with plenty of luminous and faint sources which could be potential AGNs. To this regard, we submitted two proposal to re-observe the HUDF again to search for AGNs in the field and obtain estimates on the number density of them at high redshifts ($z \gtrsim 5$) which got successfully accepted:

- **Hubble Space Telescope** Cycle 30 proposal "*AGN variability at cosmic dawn: a census of the youngest supermassive black holes*"
  PI: Prof. Matthew James Hayes, Stockholm University, ID: 17073, 30 allocated orbits[1]
  The goal of this proposal is to re-image the HUDF to search for AGNs via variability at high redshifts. We propose to obtain a 30-orbit image of the field in WFC3/F140W, to compare with a similar image taken in 2012

1 https://archive.stsci.edu/proposal_search.php?id=17073&mission=hst

(Ellis et al., 2013). Since the HST completed the requested observations in 2023, the latest results provide a decade long baseline which, due to cosmological time delay, corresponds to an optimal restframe delay of $1-1.4$ yr in the redshifts $z = 10-6$. This timescale is on par with the variability timescales of these AGNs of $0.5-1.5$ years (Caplar et al., 2017; Li et al., 2018; Sánchez-Sáez et al., 2019). This will allow us to maximize the possibility of detecting the varying AGNs.

From the obtained data, we were able to identify 71 sources with the variability analysis, assuming a variation of 3σ or higher in the brightness in all the filters and epochs considered. A subset of the identified sources are presented in Hayes et al. (2024), which include 2 confirmed AGN at $z = 2-3.2$, 3 supernovae (two at $z \sim 1$, while the third has no distance measurement), and 3 more possible AGN at $z > 6$ (7 total in the complete dataset). Using these high redshift AGN sources, we estimated the comoving number density of SMBHs after correcting for variability and luminosity incompleteness. This gives a value of 8×10^{-3} Mpc^{-3} at $z = 6-7$, matching closely with the number density at these redshifts for SMBHs from halos seeded with $d_{iso} = 100$ kpc (see figure 3.2).

- **James Webb Space Telescope** Cycle 2 proposal "*A complete spectroscopic census of the faintest galaxies and AGN at the dawn of galaxy formation*"
  PI: Prof. Matthew James Hayes, Stockholm University, ID: 3290, 48.5 hours[2]
  The goal of this proposal is to use the NIRSpec instrument of JWST to perform an extremely deep and blind exposure of the HUDF to detect the faintest emission line sources at redshifts $7-11$. This will allow us to detect the faintest AGNs in early galaxies to help constrain SMBH formation models. To achieve this, we devised an innovative new NIRSpec mode called "*Multi-Object Coronography for Slitless Spectroscopy*". According to this technique, we propose to perform a slitless survey of the HUDF by taking advantage of existing Hubble imaging - by closing all the micro-shutter of the Micro-Shutter Assembly to mask all the bright sources pre-identifed from Hubble imaging. This will filter out more than 99.99% of contaminating photons and enable us to obtain high resolution rest frame Ultraviolet (UV) spectra of $z > 7$ galaxies.

2 https://www.stsci.edu/cgi-bin/get-proposal-info?observatory=JWST&id=3290

5

DUAL AGNs

An inevitable consequence of the hierarchical growth of structure is that multiple SMBHs formed in separate halos can end up in the same halo and eventually the same galaxy. There have been a few observations of dual AGNs in the same galaxies which are separated by a few kiloparsecs, although they remain rare (Colpi, 2014; Volonteri et al., 2021). The difficulty in identifying these systems have already been well documented (for e.g., Koss et al., 2018; De Rosa et al., 2019; Silverman et al., 2020; Severgnini et al., 2021; Mannucci et al., 2022) with challenges present in every observation frequency band. For example, in the optical spectra of *Sloan Digital Sky Survey* (SDSS, York et al., 2000; Abazajian et al., 2009), they are identified through double-peaked O[III] line. But this signature is not unique, and may be indicating other phenomena occurring close to a single AGN, such as matter outflows. In other frequency bands, the signal may masquerade as star formation processes, star forming galaxies, weakly accreting black holes, etc. To confirm the presence of dual AGNs, the potential candidates detected in surveys need to be followed up with dedicated multi-wavelength follow-up programs (De Rosa et al., 2019).

From the numerical point of view, cosmological hydrodynamical simulations (e.g., Magneticum: Hirschmann et al. 2014; Horizon-AGN: Dubois et al. 2014; Illustris: Vogelsberger et al. 2014; MUFASA: Davé et al. 2016; Romulus: Tremmel et al., 2017; IllustrisTNG: Springel et al. 2018) provide samples of simulated galaxies to predict and compare results with large observational surveys. They can model the evolution and growth of SMBHs with increasing complexity of sub-grid physics to account for the interaction of the black hole with the in situ galactic matter, and in particular the AGN feedback (e.g., Springel et al., 2005a; Fabjan et al., 2010; Steinborn et al., 2015; Valentini et al., 2020). These simulations are also able to track the mergers of galaxies hosting SMBHs which results in multiple SMBHs present in the same galaxy; accreting and glowing as dual or multiple AGNs. The challenge often faced by these simulations is simulating a large enough volume to make a fair comparison with the surveys, while retaining enough resolution to be able to resolve AGN pairs with separations down to kpc scales. Nonetheless, they can still provide good estimates on the fraction of dual AGNs spanning separation ranges allowed by the resolution. On the other hand, isolated simulations of mergers allow us to study the mergers with extremely high resolution, but fail to provide any constrain on the dual AGN fraction which requires a whole population of AGN in a cosmological volume (De Rosa et al., 2019).

In approximate dark matter only codes like `PINOCCHIO` which do not resolve substructures within a halo, one needs to adopt prescriptions to model various properties of the subhalo and the galaxies, and their SMBHs contained in them to estimate the number of dual AGN systems and their fraction. In

particular, the trajectories of the subhalo, the dynamical friction timescale, galaxy merging timescales, etc, need to be modeled. In this chapter, we present the simple prescriptions for these timescales and calculate the dual AGN fraction from the Pop III.1 model for various isolation distances. We first present the merger rate of halos from 10 PINOCCHIO boxes and then the simple models used to estimate the time taken from the halo merger to the hosted galaxies' merger. Finally, we present preliminary results on the fraction of dual AGNs in the model and test its dependence on the various assumed timescales.

5.1 HALO MERGER RATE

In PINOCCHIO, when a smaller halo merges with a larger halo, the former disappears from both the merger tree and the snapshots. But due to the continuous time sampling of merger histories, the exact redshift of the merger is stored, along with the mass of both the halos at that redshift. After identifying all the seeded halos using the seed identification algorithm (§3.3), we follow their merger trees to identify mergers among the seeded halos. Then using the tree to obtain the exact merging redshift, we calculate the rate of mergers per unit redshift and volume. Figure 5.1 shows this rate, after averaging over 10 boxes of 40 Mpc/h side length. It is calculated by counting the number of mergers in a redshift bin, and then dividing by the width of the bin and the box volume. From the figure, we can observe that the rate for all the isolation distances show similar increasing trend up to $z \sim 1$, and then exhibit a small drop. It is interesting to note that the difference in the rate among different isolation distances at lower redshift is of the same order as the difference in the number density (figure 3.2), whereas it is higher at higher redshifts.

To get a better idea of the population of halos merging with each other, we present the halo mass function of only the seeded halos in figure 5.2. The mass functions of the secondary halos (less massive, *dotted lines*) in the merger, all converge at higher masses, irrespective of the isolation distance. The primary halos (more massive, *solid lines*) also show convergence at higher masses, although to a slightly lesser degree. This convergence points back to the occupation fraction of seeded halos approaching unity at the higher mass end, regardless of the isolation distance (§3.4.2). This observation highlights that the most massive seeded halos are prone to hosting multiple SMBHs, a consequence of undergoing a greater number of mergers — a distinctive signature of the hierarchical growth of structures. We also present the mass ratio distribution in figure 5.3 of the merging halos. We see a rising trend for mass ratio of all the isolation distance cases, with most mergers happening among halos of similar sizes.

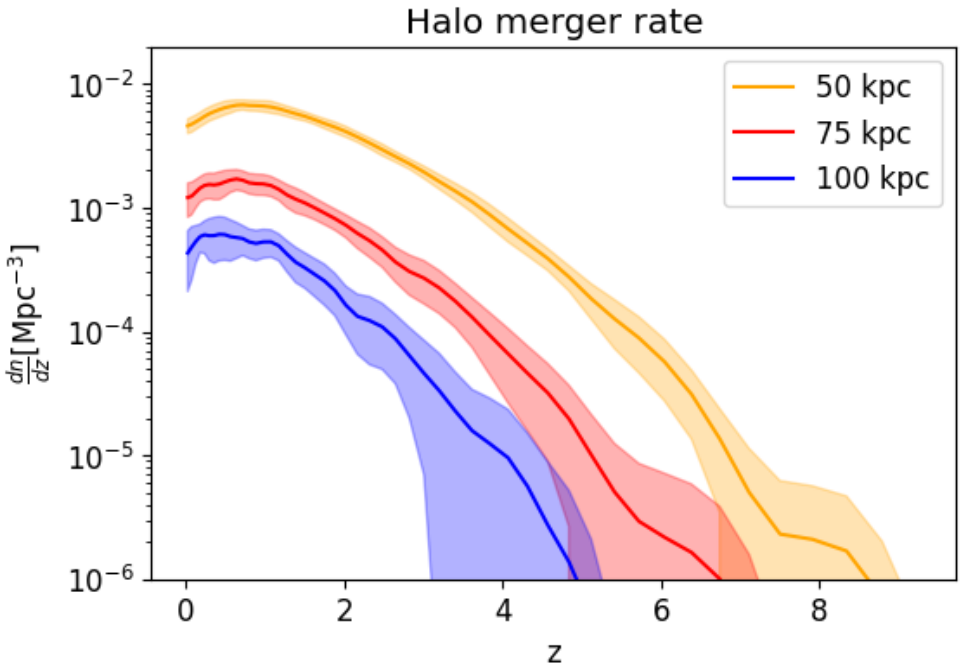


Figure 5.1: Average halo merger rate per unit volume of seeded halos for different isolation distances from ten 40 Mpc/h boxes. The shaded area depicts 1σ deviation from the average.

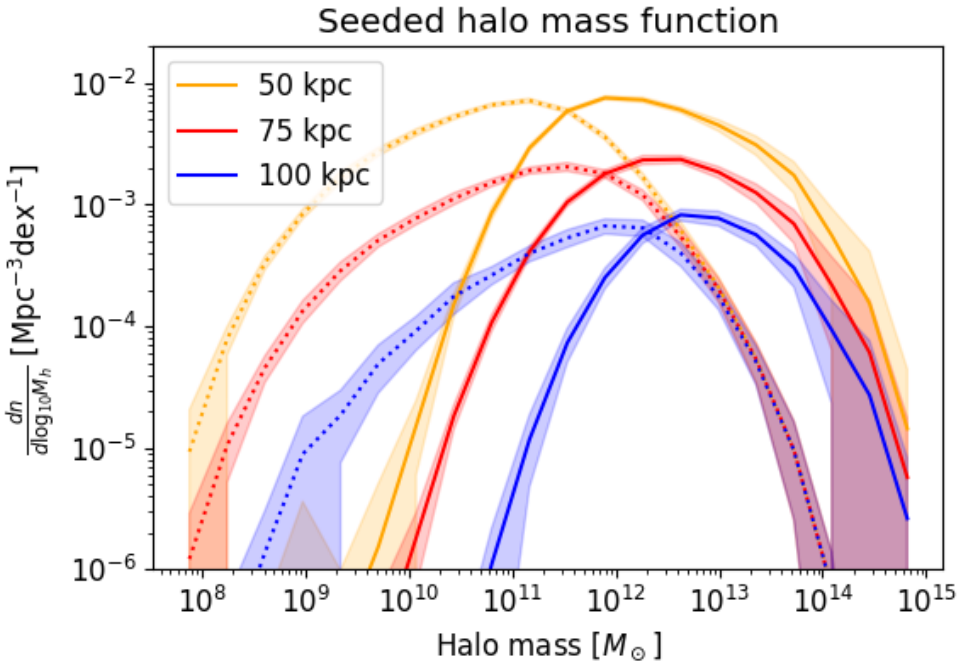


Figure 5.2: The halo mass function of the seeded halos after averaging over 10 simulation boxes. The solid lines represent the mass function for the primary (more massive) halos and the dotted represent the same for the secondary (less massive) halos in the merger. The shaded area depicts 1σ deviation from the average.

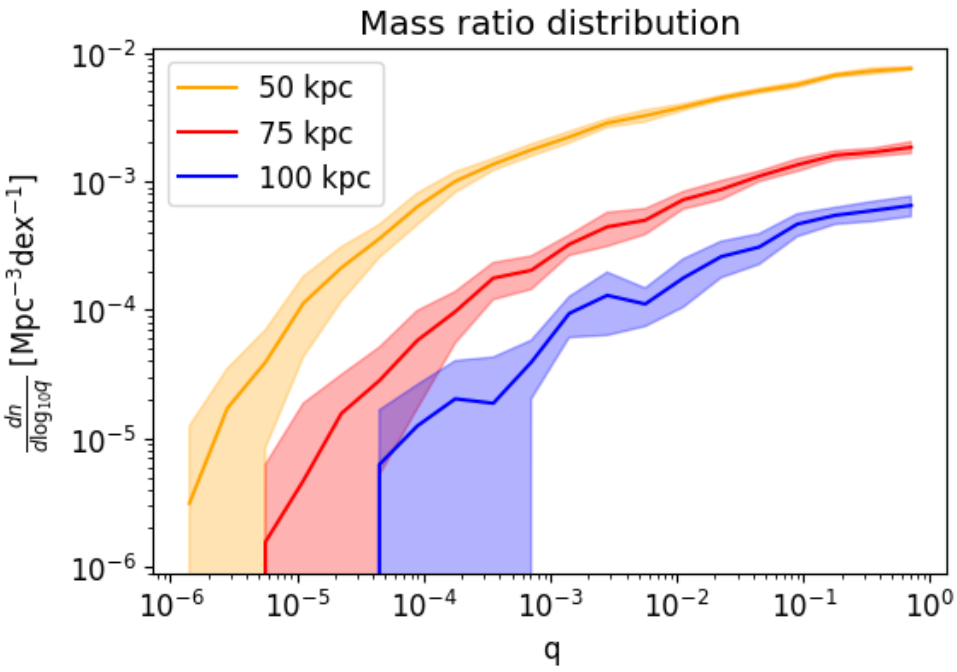


Figure 5.3: The halo mass ratio distribution of the merging halos, averaged from 10 simulation volumes. The y-axis shows the number density of halos per logarithmic ratio bins. The mass ratios are divided in equal bins of 0.3 dex width.

5.2 MERGING TIMESCALES

5.2.1 *Fixed delay model*

In PINOCCHIO, halo mergers are treated as an instantaneous process - when the smaller halo finds itself within a certain fraction of the Lagrangian radius of the larger halo, the merger of the two halos is established (Monaco et al., 2002). The further evolution of the smaller halo - now a subhalo - inside the larger halo is not resolved. This implies that we need to adopt some treatment of timescales for the tidal stripping of the halo, the sinking timescale of the subhalo towards the center due to dynamical friction, the merging time of the galaxies hosted by these halos, and finally the time taken to form a SMBH binary separated by a few kiloparsecs. In the first approximation, we assume a fixed time delay encompassing all the above processes and denote it by $\tau_{H\to G}$, which in summary represents the time delay between the beginning of the halo merger and the conclusion of the galaxy merger. Furthermore, we also assume that the dual AGN pair has formed right after the merging of the galaxies. In the simplest case, $\tau_{H\to G}$ can be assumed to be zero, but we consider cases where it is equal to a few 100 million years, up to a few billion years as well, representing a more realistic timescale (more details in the next section).

5.2.2 *Physical delay model*

To model the time scales for the halo and galaxy merger, we take the PinGAEA model from Cammelli et al. (in prep) who populated the dark matter halos of PINOCCHIO with galaxies using the state-of-the-art semi-analytic model GAEA of galaxy formation and evolution (Hirschmann et al., 2016; Fontanot et al., 2020). The model parameters of GAEA are originally calibrated to the results from the Millennium Simulation suite (Springel et al., 2005b).

Using this model, we set $\tau_{H\to G}$ as the sum of the halo survival time t_{sat}, which is based on equation 2.6 from Berner et al., 2022, and the merging time of the orphan galaxy t_{orph} (from the stripped subhalo) based on Chandrasekhar dynamical friction timescale (Chandrasekhar, 1943) adapted and applied to semi analytical models (e.g., Boylan-Kolchin et al., 2008). Denoting the mass of the main halo by M_{main}, for the subhalo by m_{sub}, and the orbital circularity at the merger by η, the halo survival time is given by:

$$t_{sat} = A(D)\tau_{dyn}\frac{\left(\frac{M_{main}}{m_{sub}}\right)^{b(D)}}{\log\left(1+\left(\frac{M_{main}}{m_{sub}}\right)^{1.2}\right)}\exp(c\eta)\exp\left(-0.2\frac{m_{sub}}{M_{host}}\right), \quad (22)$$

where $D(z)$ is the growth factor, and all the remaining parameters are defined as:

$$A = \begin{cases} 0.39 & D(z) > 0.8 \\ 0.195 + \frac{0.195}{0.2}(D-0.6) & 0.8 \geqslant D(z) \geqslant 0.6 \\ 0.195\left(\frac{D(z)}{0.6}\right)^2 & 0.6 > D(z) \end{cases} \quad (23)$$

$$b = 1.015D(z) \quad (24)$$

$$c = 1.3. \quad (25)$$

The dynamical timescale τ_{dyn} is assumed to be given by $\sim 0.1H^{-1}$, where H is the Hubble constant. Since we don't have information on the orbital circularity η, we sample it from a probability distribution obtained using numerical simulations (Zentner et al., 2005; Birrer et al., 2014). For calculating the merging time of the orphan galaxy, we use:

$$t_{orph} = 0.12t_{age}\frac{\left(\frac{m_{sub}}{M_{main}}\right)^{0.94}}{\log\left(1+\frac{m_{sub}}{M_{main}}\right)^{1.7}}, \quad (26)$$

where t_{age} is the age of the universe at the time of halo merger. The keen reader will notice a small deviation of some of the exponents and parameters, and the form itself of the formulae for t_{sat} and t_{orph} from standard expressions. This is done to consistently reproduce both the orphan merging time distribution and the stellar mass function at redshift 0. For more details on how this calibration is done, we refer the reader to Cammelli et al. (in prep).

To see this model in effect, we present figure 5.4 which shows $\tau_{H\to G}$ for merging halos seeded from $d_{iso} = 50$ kpc from one simulation volume.

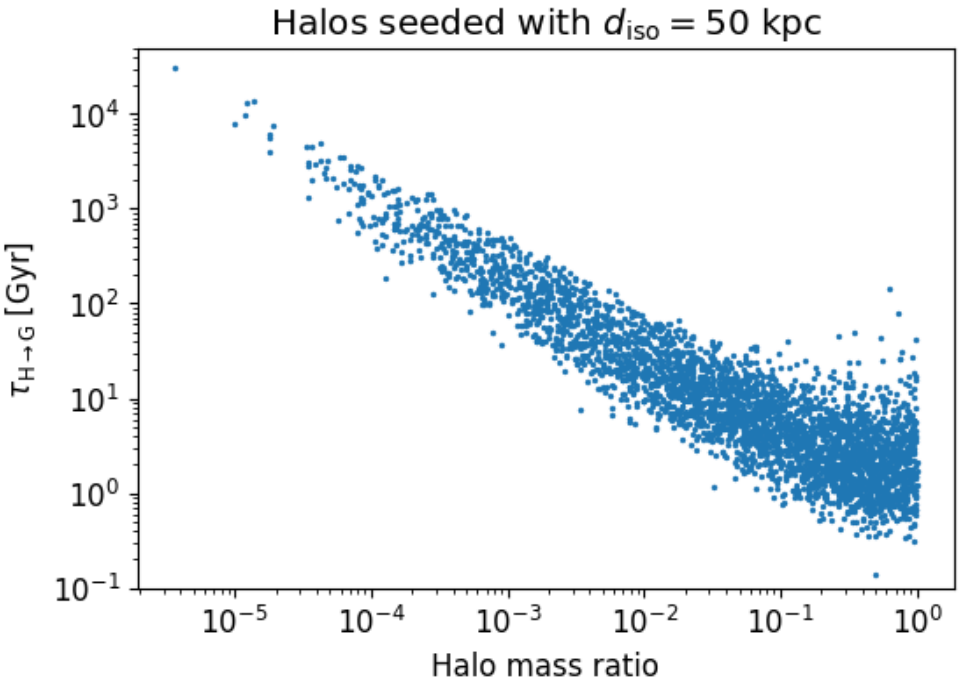


Figure 5.4: The time delay $\tau_{H\to G}$ for merging halos seeded from the $d_{iso} = 50$ kpc case from one simulation.

According to this delay model, some of the mergers will take more than the age of the universe to complete, but there are still more than 60% of halo mergers which will take less time than that. In the next section, we apply this model, along with the fixed time delay version to calculate the dual fraction with simplistic assumptions on the visibility phase of the AGNs.

5.3 DUAL AGN FRACTIONS

With the tools mentioned in the previous section, we present here the preliminary results for the fraction of dual AGNs. When performing observational surveys, a dual AGN system can only be identified if both the SMBHs which are present in the same galaxy are accreting. Fortunately, the very process of merger of galaxies which would bring multiple SMBHs together into a single galaxy, has been predicted to initiate AGN triggering activities by various studies (e.g., Di Matteo et al., 2005; Ellison et al., 2011; Silverman et al., 2011; Kocevski et al., 2015; Koss et al., 2018). On the other hand, actually observing the AGNs accreting which are not obscured and individually resolving the close pairs have been a challenging endeavor (e.g., De Rosa et al., 2019; Silverman et al., 2020).

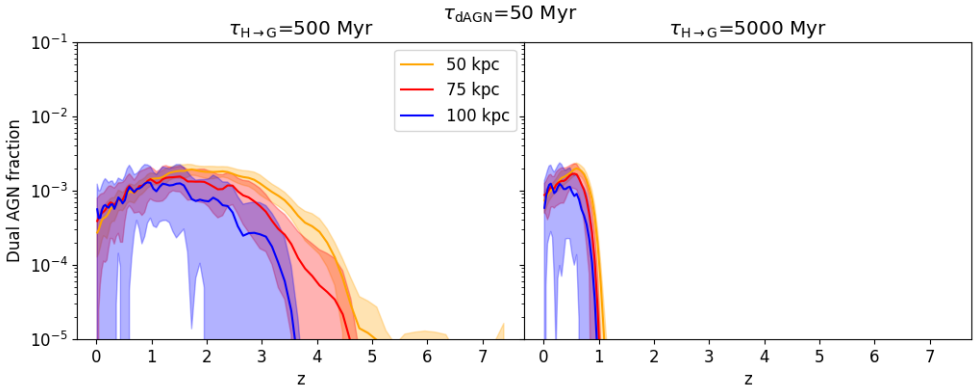
In this preliminary work, we assume that all dual SMBHs within a halo or galaxy are unobscured, and each black hole in a binary is simultaneously accreting and glowing as an AGN for τ_{dAGN} time. Since estimating this timescale based on the galaxy properties is out of scope of this thesis, we assume constant values based on Salpeter timescale (Salpeter, 1964). This

time, roughly equal to 50 Myr, refers to the time taken by a black hole to grow by a factor of e with Eddington limited accretion. Here we will consider $\tau_{dAGN} = 50$ and 500 Myr.

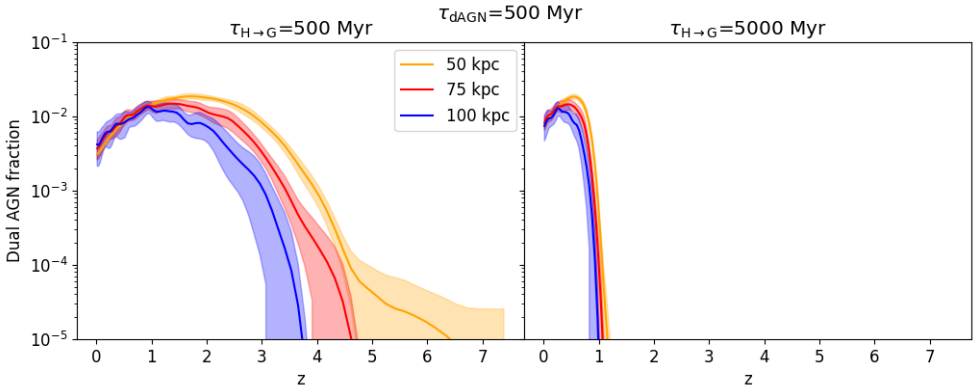
We first consider the fixed delay model (§5.2.1), with two values for $\tau_{H\to G}$: 500 Myr and 5000 Myr. The former is closer to the lower edge of the distribution of time delays in figure 5.4, while the later is closer to the median which is around 7.23 Gyr. In figure 5.5, we present the effect of changing the two time delay parameters. It is trivial to observe that $\tau_{H\to G}$ affects the spread of the fraction across cosmic time, whereas τ_{dAGN} mostly affects the amplitude of the fraction. The higher value of $\tau_{H\to G}$ predicts that the fraction decreases drastically beyond redshift 1, which is not consistent with the observations and simulations alike (Silverman et al., 2020). This implies that assuming all the halo and galaxy mergers take 5 Gyr to complete is rather an extreme assumption. Another interesting thing to notice is that the fraction for all the isolation distances tend to converge at lower redshifts. This is due to the similar scaling of the merger rate (figure 5.1) and the total number density (figure 3.2) for all isolation distances around $z \sim 0$, which results in similar ratio of these two quantities.

To get a more realistic fraction, we use the `PinGAEA` model mentioned in section §5.2.2 to obtain the value of $\tau_{H\to G}$ for each merger. Figure 5.6 shows the fraction using this model, and it predicts a more even spread in cosmic time. The amplitude again depends on τ_{dAGN} as before, and the fractions again converge at lower redshifts.

Finally, we use this physical model to make a simple comparison with the observation result of the fraction from Silverman et al. (2020), who measured the fraction up to $z \sim 3.5$. Exploiting the capabilities of the Subaru Strategic Program (SSP, Aihara et al., 2018, 2019) with Hyper Suprime-Cam (HSC, Miyazaki et al., 2018), they searched for SDSS quasars falling with the Hyper Suprime-Cam (HSC) survey footprint for dual quasar candidates. For calculating the dual AGN fraction, they selected 96 candidates with projected physical separation between 5 and 30 kpc and a minimum brightness ratio of 0.1 of the quasar pairs. Furthermore, they placed a lower limit on the mass of black holes while making comparison with different simulations - $m_{SMBH} > 10^8 M_\odot$, which we also consider. Since we do not directly have mass of the black holes inside the DM halos from `PINOCCHIO`, we again use the `PinGAEA` model to calculate the mass using its dark matter - black hole mass scaling relation which is calibrated to the local universe (more details in Chapter 6). Also, we do not have the distances between the black holes, so we do not consider the separation criterion. But we do take into account the brightness ratio by only considering those pairs which have a black hole mass ratio greater than 0.1. However, after applying all these cuts, we find that we have only $\sim$ 10 pairs for $d_{iso} = 50$ kpc, and only $\sim$ 6 for the 100 kpc case. To improve the statistics, we consider different black hole mass cutoffs and present the results in figure 5.7. Using $\tau_{dAGN} = 500$ Myr for all the cases, we can see a stark difference in the number of dual AGNs mentioned in the upper right corner of each panel for different mass cuts. The trend of converging fractions is still present for black holes with mass



(a) $\tau_{dAGN} = 50$ Myr



(b) $\tau_{dAGN} = 500$ Myr

Figure 5.5: Dual AGN fraction for different cases of τ_{dAGN} and $\tau_{H\to G}$ for three different isolation distances. The solid lines represent the average over 10 simulation volumes, and the shaded region gives 1σ spread around the mean.

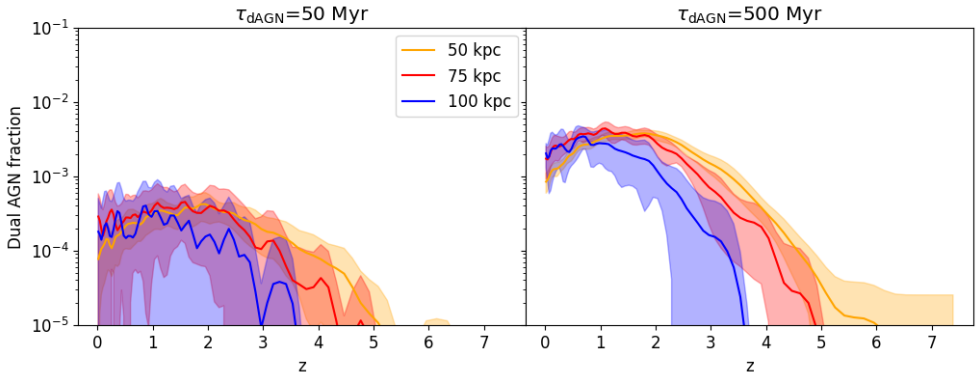


Figure 5.6: Similar to figure 5.5, but with $\tau_{H\to G}$ calculated from `PinGAEA` (Cammelli et al, in prep).

Table 5.1: The average value of the dual AGN fraction up to $z < 2$ for different mass cuts from figure 5.7.

d_{iso} [kpc]	Dual AGN fractions (in units of 10^{-4})		
	$m_{SMBH} > 10^6 M_\odot$	$m_{SMBH} > 10^7 M_\odot$	$m_{SMBH} > 10^8 M_\odot$
50	25.2 ± 9.8	5.65 ± 1.25	0.27 ± 0.14
75	31.8 ± 7.5	14.2 ± 2.1	0.99 ± 0.70
100	25.5 ± 4.9	16.8 ± 4.3	2.11 ± 1.20

greater than $10^6 M_\odot$, but there's a slight dip for halos seeded with 50 kpc for higher mass cuts. This is due to the fact that the higher mass black holes reside in higher mass halos, and the occupation fraction of those halos tends to one (§3.4.2). This implies that the set of most massive black holes for each isolation distance is the same, but they are divided by the total number of black holes - which is different. Since 50 kpc has the most number of black holes, this results in its average fraction being lower than the larger isolation distance cases.

We also see from the figure that there is not much evolution of the fraction from redshifts 0 to $\sim$ 2. In table 5.1, we present the average fraction up to redshift 2 for all the considered mass cuts. Comparing to the fraction observed by Silverman et al. (2020) up to $z \sim 3.5$, which is $(2.6 \pm 1.8) \times 10^{-3}$, we get similar fraction for the smallest mass cut ($> 10^6 M_\odot$). Although a fairer comparison should be with the mass cut used by Silverman et al. (2020) as mentioned before, which gives almost 12 times smaller fraction for the 100 kpc case and $\sim$ 100 times smaller for the 50 kpc case.

Comparing our results to hydrodynamical simulations at higher redshifts, we find that these simulations predict an increasing trend of the fraction up to $z \sim 5$, and then a decrease at redshifts beyond (De Rosa et al., 2019). The exact value of the fraction depends on various factors, including cuts on SMBH or galaxy mass, luminosity, physical separation, and more. Hydrodynamical simulations, which account for baryons, often include all the aforementioned properties, facilitating comparisons with observations. For example, the fraction calculated by Yue et al. (2021) from a single quasar pair at $z \sim 5-6$ is around an order of magnitude lower than the simulations, although the observation is a lower limit and one of the SMBH is obscured, while the simulation results include both obscured and unobscured AGNs. On the other hand, more recent observations using JWST (Maiolino et al., 2023b; Perna et al., 2023) are revealing that the fraction is higher at higher redshifts compared to simulations, pointing to an even greater abundance of dual AGNs.

To conclude this work, we would like to point out that this preliminary comparison with observations should be taken with a grain of salt, since observations are often mired in known *and unknown* selection effects, which makes faithful comparison with simulations challenging. Nonetheless, simulations can still provide a valuable insight into the fraction, especially hydrodynam-

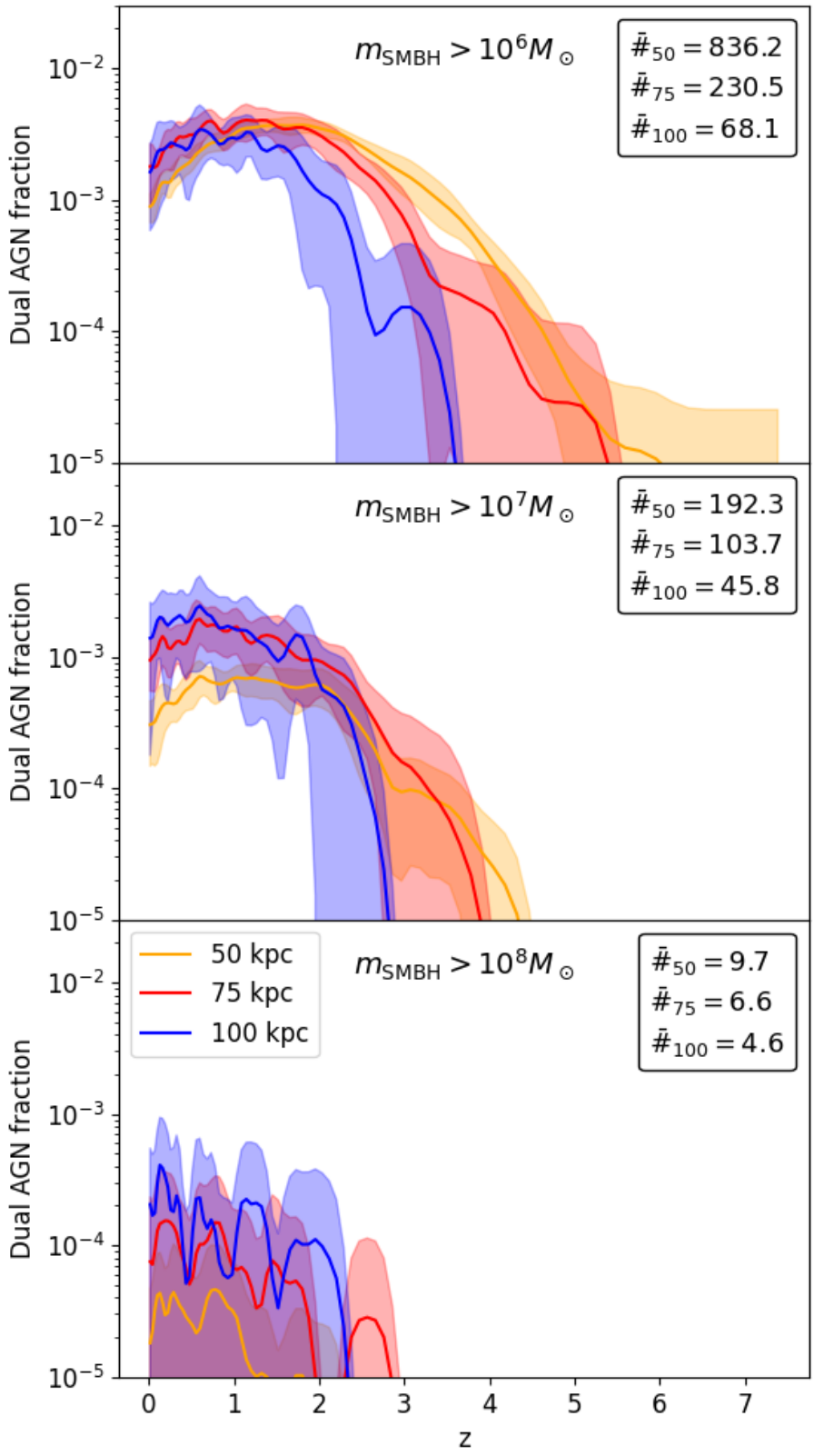


Figure 5.7: The dual fraction for different mass cuts with physical delay model and $\tau_{dAGN} = 500$ Myr. The numbers in the upper right corner of each panel depicts the average number of dual AGNs for respective isolation distances.

ical simulations which are able to treat AGN physics to small scales. In this chapter we presented a simple model which allowed us to calculate the fraction for the Pop III.1 model with relatively simple assumptions. In the future we plan to improve on these predictions by using a more sophisticated prescription for the interaction of the SMBH with the galactic environment and using larger simulation boxes to obtain more massive black holes and improve the statistics.

6

GWB FROM THE POP III.1 MODEL

After discussing the dual AGN fraction rate in the previous chapter, the natural question to ask next is what is the gravitational wave background generated by these SMBHs orbiting around each other. In Chapter 2, we studied how a population of the most massive SMBH binaries create a background of superimposed and unresolved GWs which can be detected by PTA experiments. Since the Pop III.1 model also predicts black holes of these masses which can end up in the same dark matter halo and galaxy (Banik et al., 2019; Singh et al., 2023), in this chapter we aim to calculate the GWB generated by them after making some simple assumptions. We begin by calculating the merger rate of SMBHs in the model, and then discuss the mass scaling relations needed to assign masses to the said black holes. We then present the estimated background, and the impact of various assumptions we make. Finally we end with the comparison of our results with the latest NANOGrav results (Agazie et al., 2023a).

6.1 SMBH MERGER RATE

Let's recall from §2.1 that in order to calculate the background, we need the number density of mergers of the black holes per unit redshift and chirp mass: $d^2n/dzd\mathcal{M}$. As mentioned multiple times in this thesis, from `PINOCCHIO` we only have the information of the dark matter halos hosting the seed SMBHs. But we also have the merger histories of these halos through which we can investigate the mergers among seeded halos and their exact redshifts of occurrence. In order to obtain the redshifts of black hole mergers and their masses, we need to make some assumptions. We start by focusing on calculating the BH merger redshifts, or the number density of mergers per unit redshift: dn/dz.

6.1.1 *SMBH merger rate per unit redshift*

To estimate the merging redshift of the black holes present inside seeded halos, we need to model the time taken from the halo merger to the black hole merger. The first part of this timescale - the time elapsed from the beginning of the halo merger to the conclusion of the galaxy merger - is already treated in Chapter 5. Denoted previously by $\tau_{H\to G}$ (§5.2), we again consider two cases: constant values for this timescale, and the physical model from Cammelli et al. (in prep).

Then the remaining part of the timescale is the time taken by the binary to merge - which is usually the most uncertain. This involves modelling the internal dynamics of a galaxy and the interaction of stellar objects with the SMBHs. Initially during a galaxy merger, at separation of $\sim$ kpc, the dynamical

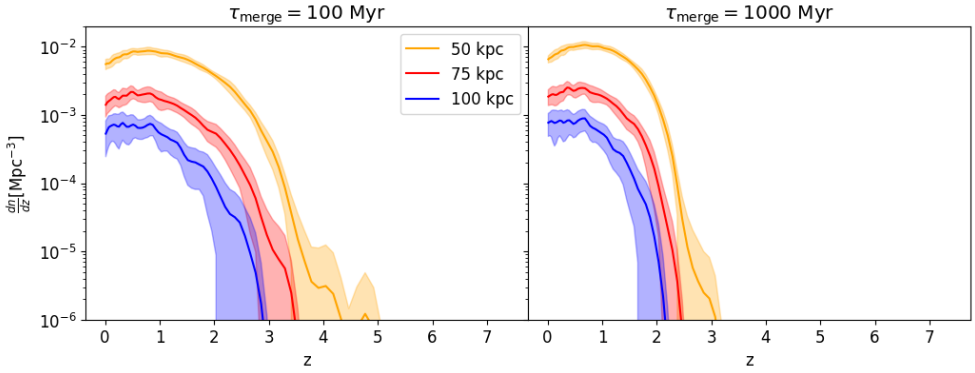
friction is enough to bring the SMBHs toward the galactic center. Afterwards, the shrinking takes place mostly due to three-body interactions with stars which helps transfer the angular momentum away from the binaries into the stars (e.g., Quinlan, 1996; Sesana et al., 2006; Kelley et al., 2017; Barber et al., 2024). But it may happen that the binaries do not encounter enough interactions to shrink further within Hubble timescales to reach separations where the GWs dominate the orbit decay (~ 0.01 pc for total binary mass of $\sim 10^8 M_\odot$) - the so-called *final parsec problem* (e.g., Milosavljević and Merritt, 2001; Merritt and Milosavljevic, 2005). While various mechanisms have been proposed to address this problem (e.g., Yu, 2002; Khan et al., 2011; Holley-Bockelmann and Khan, 2015; Sesana and Khan, 2015; Bonetti et al., 2016), for the purposes of this work, we assume that the problem has been solved.

We introduce the parameter $\tau_{\rm merge}$ which denotes binary lifetime of all the binaries (same as τ_f in Agazie et al. 2023) and consider two values of this parameter for all the binaries: 100 Myr and 1000 Myr. In figure 6.1 we show the SMBH merger rate density for both the cases of fixed $\tau_{H\to G}$ and the one obtained from the physical model (§5.2.2). The merger rate from constant $\tau_{H\to G}$ exhibits a peak around $z \sim 1$, and then decreases a bit towards $z = 0$, essentially mirroring the halo merger rate (figure 5.1). On the other hand, the merger rate after considering the physical model for $\tau_{H\to G}$ shows an increasing rate down to $z = 0$. The effect of varying $\tau_{\rm merge}$ is not too pronounced for both the cases, and the final rate at $z = 0$ is similar. It does however affects the spread of the rate throughout the redshifts, which is expected. Nonetheless, for calculating the GWB we are mostly concerned with the lower redshifts since the majority of the signal comes from this regime.

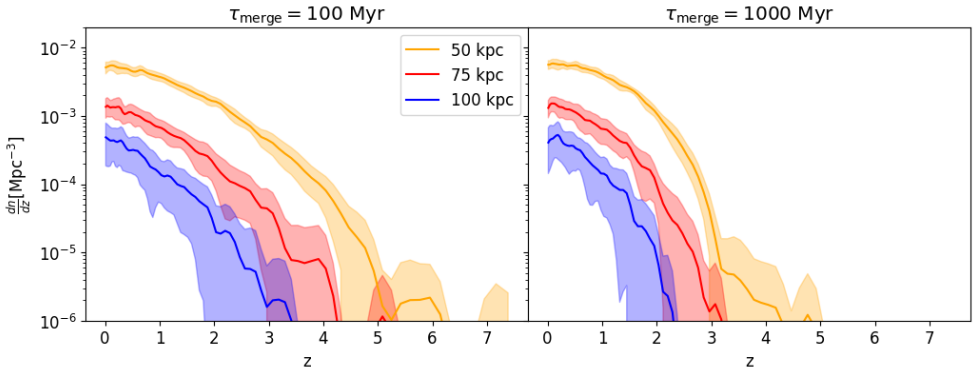
With the merger rate per unit redshift set, we can focus our attention on calculating the total merger rate per unit chirp mass. Calculating the chirp mass requires the information of the masses of the black holes. From `PINOCCHIO` we already have the masses of the merging DM halos, so we can turn to halo mass - black hole mass scaling relations, which we explore next.

6.1.2 *Mass scaling relations*

Over the past couple of decades, there has been a mounting evidence of the symbiotic relationship between the growth of a SMBH and its host galaxy - from the observational and simulation point of view alike (e.g., Ferrarese, 2002; Di Matteo et al., 2003; Jiang et al., 2011; Beifiori et al., 2012; Sabra et al., 2015; Mutlu-Pakdil et al., 2018; Davis et al., 2019; Tripodi et al., 2024). We consider only the relations at $z = 0$ since the majority of the signal comes from lower redshifts and there is not much evolution of the relations expected at these redshifts. We show the scaling relations considered in this work in figure 6.2, along with the scaling relation of stellar mass of a galaxy with its host halo dark matter mass (the y-axis of the plot corresponds to the stellar mass for this relation) from Behroozi et al. (2013). We include this stellar mass scaling relation as a reference for an upper limit on the masses of the black holes which the halos can host - it is highly unlikely that the



(a) $\tau_{H\to G} = 1000$ Myr



(b) $\tau_{H\to G}$ calcuated from `PinGAEA` model (Cammelli et al., in prep)

Figure 6.1: The comoving number density of the mergers of SMBHs from Pop III.1 model with different values of $\tau_{H\to G}$ and t_{merg}. The solid lines represent the average over 10 simulation volumes, and the shaded region gives 1σ spread around the mean.

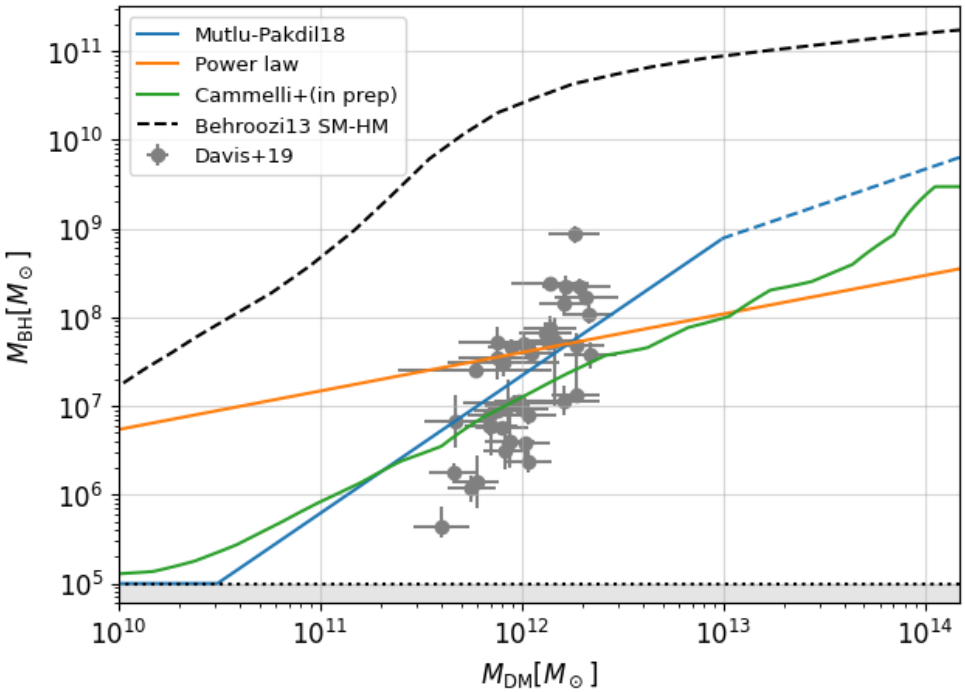


Figure 6.2: The dark matter halo mass M_{DM} - black hole mass M_{BH} scaling relations used in this work, and the stellar mass - dark matter halo mass from Behroozi et al. (2013). The blue dashed line represents our modification to the *Mutlu-Pakdil18* relation to better mimic the black holes in higher halo masses. The grey dots show the observational data from Davis et al. (2019). The black dotted line shows our lower limit on the SMBH masses.

mass of a SMBH is equal to or greater than the stellar mass of its galaxy. The *Mutlu-Pakdil18* relation is taken from Mutlu-Pakdil et al. (2018), in which they used the Illustris simulation to predict the SMBH - dark matter mass relation. However, at dark matter masses above $10^{13} M_\odot$, their relation overshoots compared to the masses of the SMBHs in their simulation, so we halved the slope above those halo masses. The *Cammelli+(in prep)* relation comes from the `PinGAEA` model (Cammelli et al., in prep) introduced in Chapter 5, which takes into consideration the $10^5 M_\odot$ mass of the SMBH seeds. Finally, *Power law* refers to a simple power law relation taking into consideration that the seeding mechanism we employ in our simulation seeds $10^6 M_\odot$ halos with black holes of $10^5 M_\odot$. We then assume that the black hole mass follows the power law, passing through the point slightly above the median black hole mass from the data set of Davis et al. (2019). This relation (and also the *Cammelli+(in prep)* relation) in turn predicts a SMBH of $\sim 10^8 M_\odot$ in a dark matter halo of $10^{13} M_\odot$ - equivalent to having a $\sim$ 25 times more massive black hole than Sgr A* in a halo $\sim$ 10 times more massive than the Milky Way halo. Finally, since all our Pop III.1 seeds are $10^5 M_\odot$, we use this as a floor value for the mass of SMBHs in all the relations, depicted by the horizontal black dotted line in the figure.

While using the mass relations, we also add some dispersion σ_{BH} to simulate the spread in black hole masses. Although this dispersion could

be a free parameter, the freedom in varying it is limited by the black hole mass function (BHMF). To test the impact of our dispersion, we calculate the BHMF using a dispersion of 0.3 dex for all the seeded halos at $z = 0$, and compare it with observational estimates. Figure 6.3 shows the BHMF calculated using the scaling relations discussed above. The BHMF from Mutlu-Pakdil et al. (2016) is estimated from observations of SMBHs from Sérsic index distribution of 208 early and late type galaxies in the local universe, covering a comoving volume of $3.37 \times 10^4 h_{67.77}^{-3} \mathrm{Mpc}^3$. On the other hand, the BHMF from Shankar et al. (2020) is estimated by using AGN clustering to constrain black hole - stellar mass relation to avoid observational biases. This results in a unbiased/intrinsic mass function which is more strictly constrained than other mass functions estimated through observations. Comparing these BHMFs with the ones obtained from the scaling relations for the seeded halos, we can conclude the following for each relation:

- The *Mutlu-Pakdil18* relation is in good agreement with the intrinsic BHMF at lower masses and then with the observational one, but overshoots it at the highest masses. At the lowest mass end, the mass function rises due to the rising abundance of low mass seeded halos for smaller isolation distances (see the $z = 0$ case in figure 5.2). This is due to assuming a floor of $10^5 M_\odot$ for all SMBH masses, which results in all seeded halos below $\sim 3 \times 10^{10} M_\odot$ being assigned SMBHs of $10^5 M_\odot$ mass.

- The *Power law* relation produces a peak at around $\sim 10^7 M_\odot$ which is expected since it predicts the black holes almost two orders of magnitude higher at around $10^{10} M_\odot$ halos compared to other relations (figure 6.2). At the higher mass end, the mass functions follows the intrinsic BHMF, although it fails to produce the most massive black holes.

- The *Cammelli+ (in prep)* relation produces the best fit of the intrinsic BHMF compared to all three scaling relations considered. This is not completely unexpected since the `PinGAEA` model is calibrated to reproduce the $z = 0$ BHMF for the Pop III.1 model (Cammelli et al., in prep). At the lowest mass end, the behaviour is quite similar to the *Mutlu-Pakdil18* relation since the condition of the seed masses of SMBHs being $10^5 M_\odot$ is already considered in the `PinGAEA` model. We are currently investigating this scaling relation at low halo masses to gain a better understanding of the growth of SMBHs in these halos.

One final observation from the BHMFs: all the mass functions for different isolation distances tend to converge at the highest end because the occupation fraction of the most massive halos tend to one irrespective of the isolation distance (§3.4.2).

Now with the masses of all the SMBHs in binaries estimated using these scaling relations, we can calculate the complete number density of mergers per unit redshift and the chirp mass. This allows us to finally calculate the background for the Pop III.1 model, which we present in the next section.

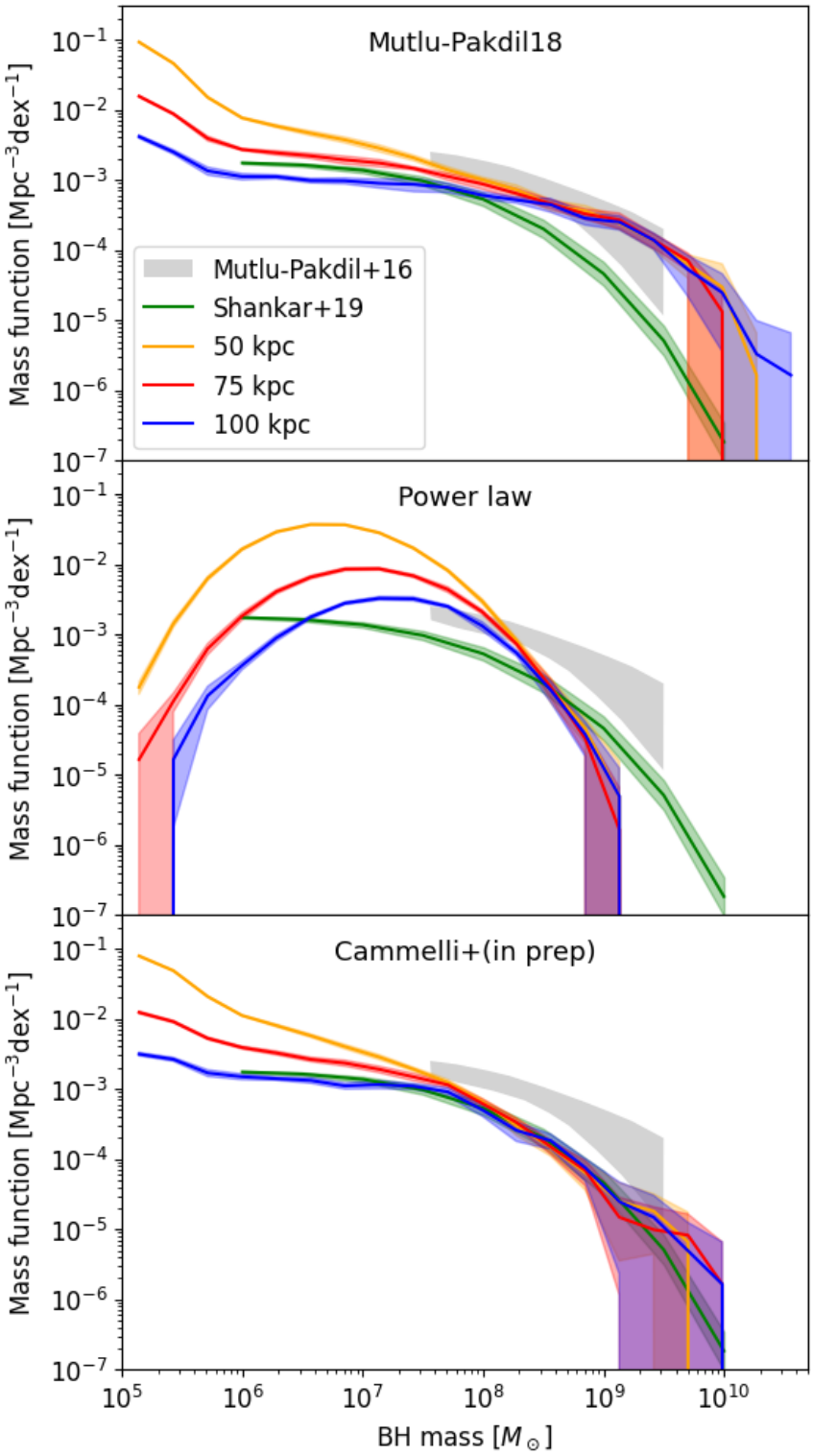


Figure 6.3: The black hole mass function for the mass relations considered in this work, with mass dispersion $\sigma_{\rm BH} = 0.3$ dex. The grey shaded area represents the estimate from Mutlu-Pakdil et al. (2016) and the green region shows the intrinsic BHMF from Shankar et al. (2020). The BHMF from the Pop III.1 model shows the average over 10 boxes, with 1σ spread around the mean.

6.2 GWB RESULTS

Equipped with the information on how to calculate the merger rate per unit redshift and chirp mass from the previous section, we can now calculate the GWB with the methods elaborated in Chapter 2. We explore a total of four cases of timescale combinations: $\tau_{merg} = \{100, 1000\}$ Myr cases for both $\tau_{H\to G} = 1000$ Myr and $\tau_{H\to G}$ calculated from `PinGAEA`.

The procedure to calculate the spectrum can be briefly summarized as follows:

- From the 10 $40h^{-1}$ Mpc side length `PINOCCHIO` boxes, we compute the rate of merger $d^2n/dzd\mathcal{M}$ 100 times per box for each isolation distance and mass scaling relation, while also considering the dispersion in mass $\sigma_{BH} = 0.3$ dex. Note that for the 100 realizations from 1 box, each will have the same redshift distribution of mergers, but slightly varied mass distribution due to the dispersion.
- Once we have the 100×10 realizations of the merger rate from all the boxes, we take its average to compute the average rate of mergers.
- Using this average rate, we calculate 100 realizations of the GWB using equation 19 since we consider the discreteness of sources at high frequencies. This is done by calculating the number of sources in each 3-dimensional $(z, \mathcal{M}, \ln f_r)$ bin from the equation 17 and then performing Poisson sampling of each bin 100 times for as many realizations.

With this procedure, we calculate the average GWB spectrum and plot it in figures 6.4, 6.5, 6.6, and 6.7 to explore the effects of different assumptions of the timescales and the mass scaling relations. The following conclusions can be drawn from these spectra:

- All the spectra show "spikes" at the higher frequency end, due to the discreteness of the sources at those frequencies (Sesana et al., 2008).
- Some individual realizations (*light blue lines*) may rise above the measured spectrum for both $\gamma = 3.2$ and $13/3$, but in general the average tends to stay lower, especially at higher frequencies.
- The constant $\tau_{H\to G}$ predicts a higher value of the background compared to the background from the physical model. This is due to the fact that in the physical model, the merger times self-consistently depend on the properties of the halos (mainly the mass ratios) and it may happen that some of the halo mergers take more than the age of the universe to complete. But assuming a constant value of this timescale forces all the mergers to happen in a relatively shorter time, inflating the number of mergers happening within the past lightcone.
- The effect of varying τ_{merge} seems minimal, which is not surprising given the merger rate at lower redshifts have small dependence on it (see figure 6.1).

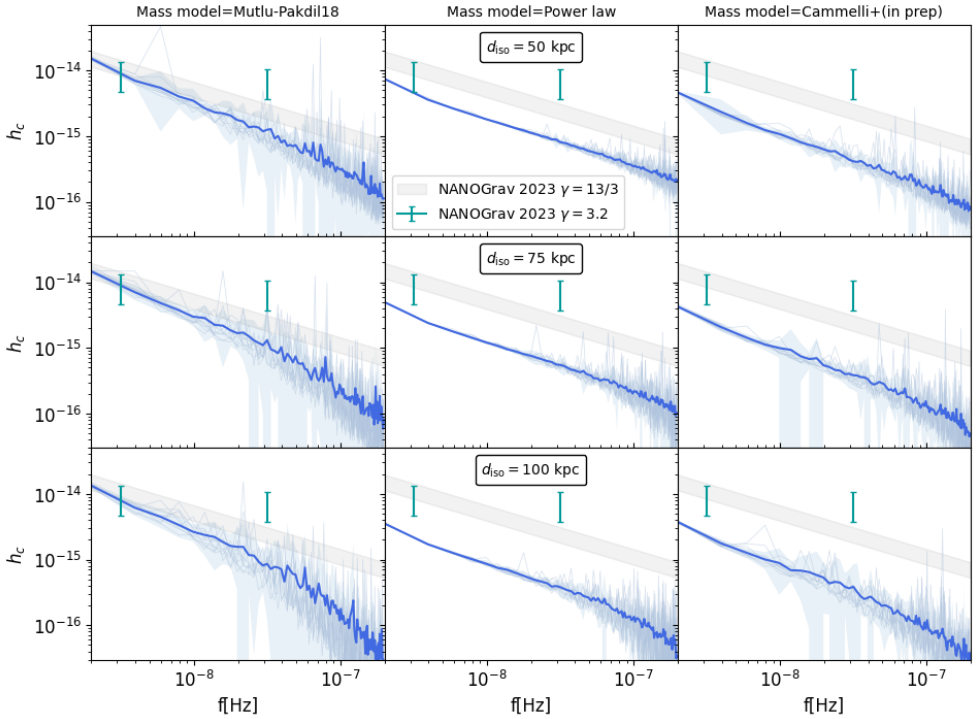


Figure 6.4: The GWB from Pop III.1 model for three isolation distances, each occupying a row with the respective label in the middle column. The columns show the spectrum from all the mass scaling relations used. These results are for $\tau_{H\to G} = 1000$ Myr and $\tau_{merge} = 100$ Myr. The solid blue line shows the average from 100 realizations of the universe, and the blue shaded region shows the 1σ spread around the mean. The light blue lines show 10 random realizations of the spectrum. The two turquoise errorbars show the posteriors of the spectrum measured from NANOGrav 15 year data set (Agazie et al., 2023a) for $f = \{0.1, 1\}\,\mathrm{yr}^{-1}$ frequencies, without assuming an ideal GW-only orbit decay (which gives the posterior for the spectral index as $\gamma = 3.2 \pm 0.6$, §2.2). The grey band also depicts the posteriors for the observed spectrum, but with the prior assumption of GW-only orbit decay ($\alpha = 2/3$).

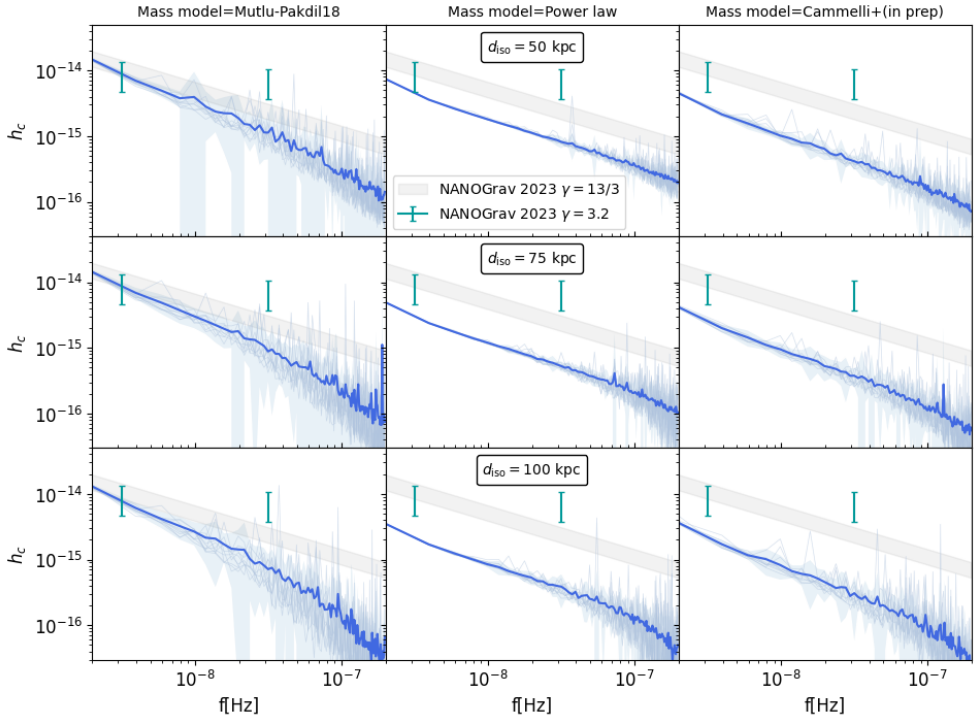


Figure 6.5: Same as figure 6.4, but for $\tau_{H\to G} = 1000$ Myr and $\tau_{merge} = 1000$ Myr.

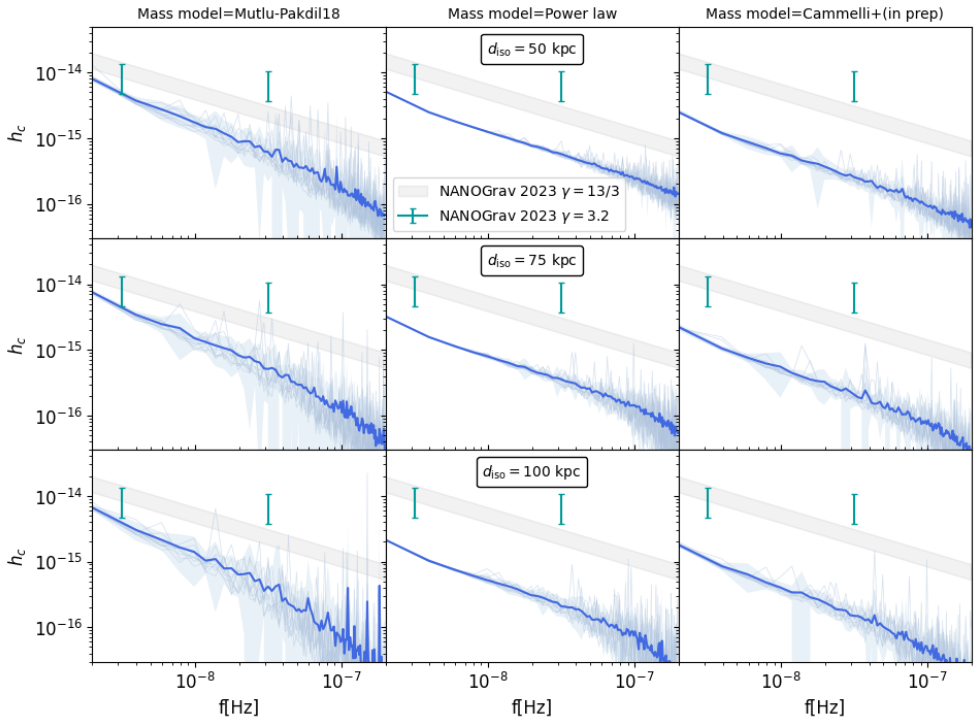


Figure 6.6: Same as figure 6.5, but for $\tau_{H\to G}$ calculated from `PinGAEA` and $\tau_{merge} =$ 100 Myr.

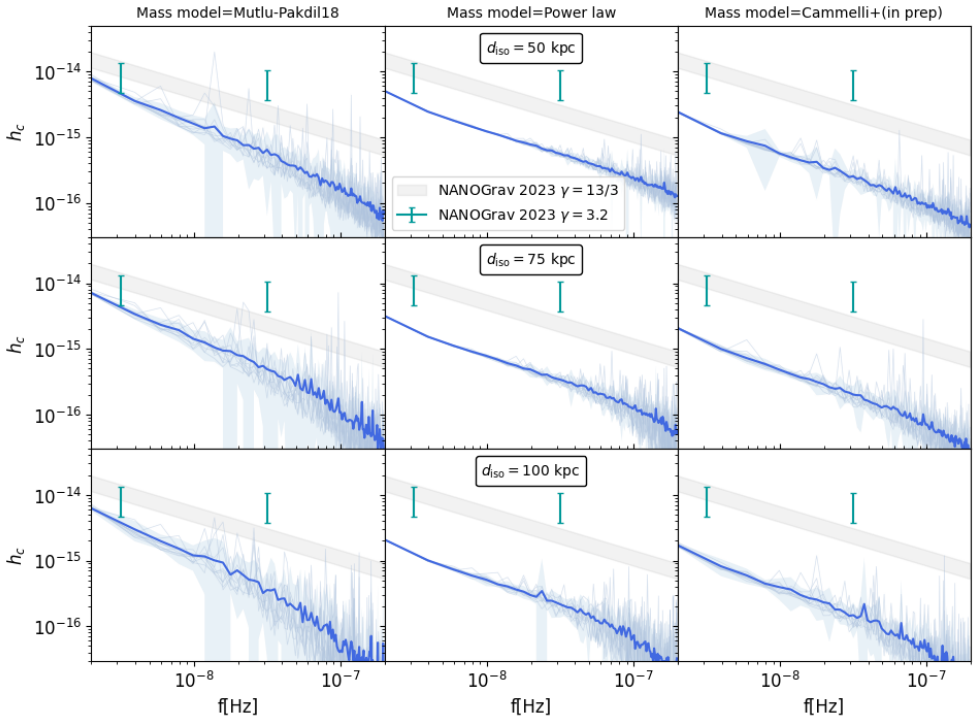
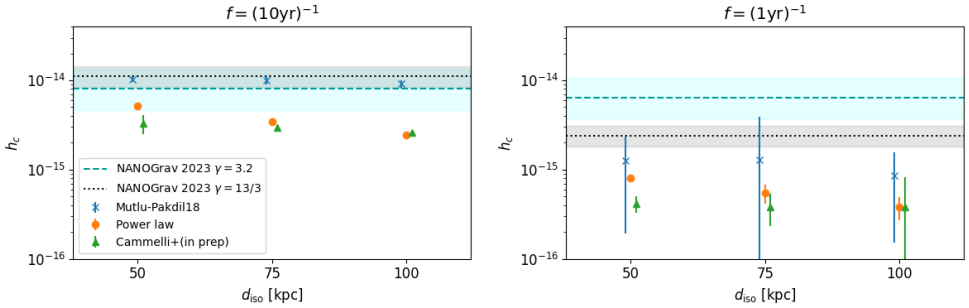


Figure 6.7: Same as figure 6.6, but for $\tau_{H\to G}$ calculated from PinGAEA and $\tau_{merge} = 1000$ Myr.

To better compare the value of the background at the reference frequencies of $f = \{0.1, 1\}\,yr^{-1}$ with the NANOGrav results (Agazie et al., 2023a), we present figures 6.8, 6.9, 6.10, and 6.11 which show the spectrum for all the cases considered at only these frequencies. From these figures it is trivial to conclude that all the models considered are closer to the observed spectrum at $0.1 yr^{-1}$ frequency than at $1 yr^{-1}$. Also, there is more dispersion at the higher reference frequency due to the effect of discreteness mentioned before. At lower frequencies there are much more sources so it results in a smoother spectrum.

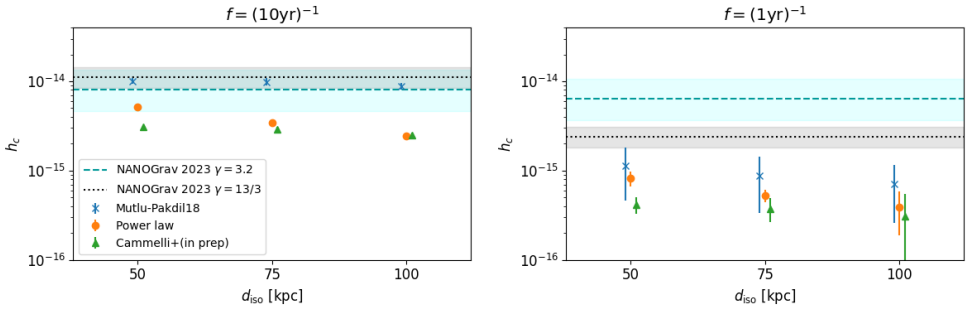
We also observe that there is not a strong dependence on the isolation distance parameter, which can be explained by considering that most of the background comes from the most massive binaries, which tend to be the same for all the isolation distances (see the high mass end of the BHMFs in figure 6.3 which is for all the SMBHs at $z = 0$, but the conclusion holds for the SMBHs in binaries as well). On the other hand, the dependence on the modelling of $\tau_{H\to G}$ and the mass scaling relation is stronger.

Considering further the dependence on the scaling relations, the *Mutlu-Pakdil18* always predicts the highest background, which is expected due to the bottom-heavy nature of the BHMF with our assumed σ_{BH}. It predicts the most heaviest SMBHs compared to the other relations and thus the highest amplitude of the background. The relation from PinGAEA predicts the lowest amplitude of the background, although this relation provides the best agreement with the local estimated BHMF. This scaling relation, combined with the halo to the galaxy merging time delay from PinGAEA model



(a) $\mathrm{f} = 0.1\,\mathrm{yr}^{-1}$ (b) $\mathrm{f} = 1\,\mathrm{yr}^{-1}$

Figure 6.8: The value of the background at the references frequencies of $\mathrm{f} = \{0.1, 1\}\,\mathrm{yr}^{-1}$ for the case with $\tau_{\mathrm{H}\to\mathrm{G}} = 1000$ Myr and $\tau_{\mathrm{merge}} = 100$ Myr.

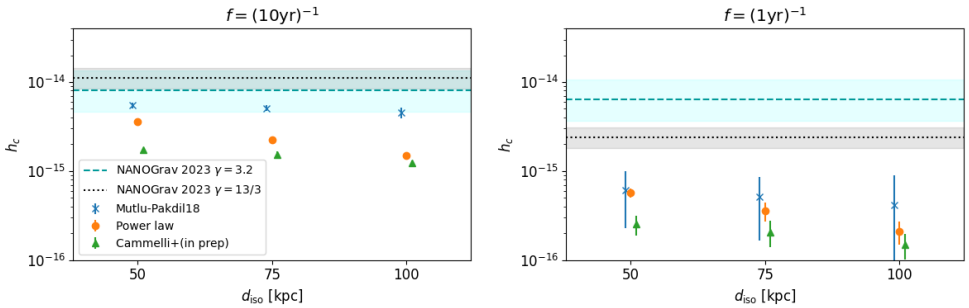


(a) $\mathrm{f} = 0.1\,\mathrm{yr}^{-1}$ (b) $\mathrm{f} = 1\,\mathrm{yr}^{-1}$

Figure 6.9: Same as figure 6.8, but for $\tau_{\mathrm{H}\to\mathrm{G}} = 1000$ Myr and $\tau_{\mathrm{merge}} = 1000$ Myr.

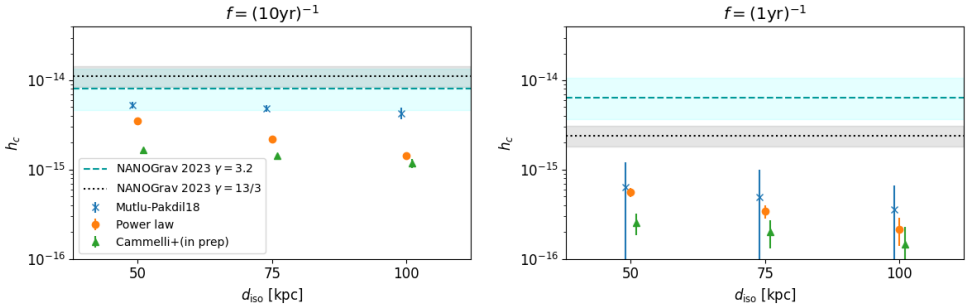
($\tau_{\mathrm{H}\to\mathrm{G}}$) gives the most physically motivated estimation of the background. The amplitude is a bit lower than the measured signal, but this is a known trend from most semi-analytical models (§2.2). Furthermore, the isolation criteria tends to result in lower number of mergers of SMBHs due to the initial isolation of the seeds, and hence lower the background (although this isolation reduces as structure evolves, as shown in the clustering analysis in §3.4.3). The amplitude of the background from the *Power Law* model lies in the middle of the other two scaling relations. Even though it produces less massive SMBHs at the highest mass end compared to the relation from `PinGAEA`, it makes up in the signal by having many more slightly less massive SMBHs, around $\sim 10^7 \mathrm{M}_\odot$ (figure 6.3).

Even though the GWB does not give the most stringent constraints on the isolation distance, it does give a reasonable estimate of the predicted background, which is slightly lower than the measured. To search for another way to have more constraints on this model, we can start by considering again the BHMF from 6.3. This model predicts quite a lot of SMBHs at $< 10^7 \mathrm{M}_\odot$ where the mass function is also not well constrained (Shankar et al., 2020). The mergers of SMBHs of these masses ($10^5 \sim 10^7 \mathrm{M}_\odot$) is in the sweet spot for detection by the upcoming Laser Interferometer Space Antenna (LISA)



(a) $\mathrm{f} = 0.1\ \mathrm{yr}^{-1}$ (b) $\mathrm{f} = 1\ \mathrm{yr}^{-1}$

Figure 6.10: Same as figure 6.9, but for $\tau_{\mathrm{H}\to\mathrm{G}}$ calculated from `PinGAEA` and $\tau_{\mathrm{merge}} = 100$ Myr.



(a) $\mathrm{f} = 0.1\ \mathrm{yr}^{-1}$ (b) $\mathrm{f} = 1\ \mathrm{yr}^{-1}$

Figure 6.11: Same as figure 6.10, but for $\tau_{\mathrm{H}\to\mathrm{G}}$ calculated from `PinGAEA` and $\tau_{\mathrm{merge}} = 1000$ Myr.

mission[1] (Amaro-Seoane et al., 2017), led by the European Space Agency[2] (ESA) which is expected to launch in 2035. This mission will be able to detect mergers of SMBHs with chirp mass around those masses at extremely high redshifts ($z \gtrsim 10$). Combining the detections from Laser Interferometer Space Antenna (LISA) and the PTA experiments will provide a more complete understanding of the merger rate of black holes of different masses in a wider redshift range. Hence, they will be invaluable in constraining the isolation distance since we predict differences in the merger rates which increase with the redshift (figure 6.1).

1 https://www.lisamission.org/
2 https://www.esa.int/

7

FINAL REMARKS

In this thesis, we investigated the Pop III.1 seeding mechanism and the properties of the dark matter halos seeded with the SMBHs produced by it. In particular, we investigated the ensemble properties such as the number density, occupation fraction and the clustering of seeded halos. We then computed the dual AGN fraction rate as well, after making simple assumptions about the timescales involved. Finally, we presented the GWB generated by the SMBH binaries and compared with the latest observations. In the following sections, we provide a summary of the three main projects defining this thesis, and the future perspectives as well.

7.1 SUMMARY AND CONCLUSIONS

7.1.1 *Pop III.1 seeds*

In Chapter 1, we introduced the Pop III.1 stars which take advantage of the isolation and WIMP dark matter self-annihilation to grow up to $\sim 10^5 M_\odot$ before collapsing to form SMBHs (McKee and Tan, 2008; Tan et al., 2010; Freese et al., 2010; Rindler-Daller et al., 2015). Since this collapse happens in the early universe around $z \sim 25 - 35$, the SMBHs produced have enough time to grow and explain the detection of high redshift quasars, without invoking Eddington, or super-Eddington limited accretion for their entire lifetime. The main parameter of the model is the isolation distance, which defines the initial isolation of the seeded (mini-)halos in the early universe. Banik et al. (2019) used this seeding mechanism to investigate number of SMBHs which can be produced by it, the occupation fraction of the seeded halos, and the clustering properties, all up to $z = 10$.

The first project of this thesis was to extend the results presented in Banik et al. (2019) down to $z = 0$ (Chapter 3). In order to achieve this, we used the latest version of `PINOCCHIO` code and developed an algorithm to identify all seeded halos based on the isolation criteria. We then followed the evolution of all the seeded halos up to $z = 0$ and found that for certain values of the isolation distance ($d_{iso} = 100$ kpc), the number of seeds produced by this model matches well with the observed local number density. The evolution of the number density also depicted that most of the seeds formed quite early in the universe, around $z \sim 25$, and the numbers stay almost constant throughout, only dipping slightly at redshifts closer to zero due to mergers. We also compared the evolution of the number density with other models, and found that the number density of SMBHs from the Pop III.1 model at redshifts greater than ~ 6 or 7 is higher compared to other models. This provides avenues to test the seeding mechanism, by estimating the number density at higher redshifts.

The occupation fraction of the seeded halos uncovered the observed correlation of the most massive halos always being occupied by SMBHs at lower redshifts, irrespective of the isolation distance. The clustering analysis revealed that at $z = 0$ the 3d 2-point correlation function of the seeded halos exhibit slightly lower levels of clustering compared to a sample of all the halos, and to the HMT numerical scheme. This signal could potentially be observable in a comprehensive survey of SMBHs in a large volume of the local universe. We also studied the evolution of clustering through the projected correlation function, and found evidence at high redshifts ($z \gtrsim 3$) of feedback cleared bubbles at scales comparable to the comoving size of the isolation radius at mean formation redshifts. In other words, we found a significant dip of the clustering amplitudes at scales smaller than the isolation distance at the formation epoch. At lower redshifts, this signal fades away due to the evolution of the universe into a more clustered state. As a final clustering result, we compared the projected correlation function with data from Zehavi et al. (2011) who used SDSS (Abazajian et al., 2009) data to compute the function for bright galaxies. After making appropriate mass cuts for seeded halos, we found excellent agreement for $d_{iso} = 50$ and 100 kpc with their observations at $z = 0$.

Building on the result of high number density of seeds at high redshifts, in Chapter 4 we presented methods to create a synthetic Ultra Deep Field with seeded halos from the Pop III.1 model, in order to compare with the Hubble Ultra Deep Field (Beckwith et al., 2006; Ellis et al., 2013). This comparison aimed at predicting the number of SMBHs which could be visible in the field at high redshifts. We presented synthetic fields for $d_{iso} = 50$ and 100 kpc and the number of seeds present within the redshift bins from $z = 5$ to 10. We also presented similar results for the HMT seeding scheme and concluded that estimating the number of AGN sources present at high redshifts ($z \gtrsim 7$) in the field could provide strict constraints on different formation mechanisms, and help distinguish between isolation distances as well.

Finally, the results presented in this project helped motivate successful Hubble Space Telescope and James Webb Space Telescope proposals. Both the proposals aimed at reobserving the HUDF to search for high redshift AGNs to get estimates on the number density of SMBHs. This will help in comparing different SMBH formation mechanisms and also help constrain the isolation distance parameter for the Pop III.1 model.

7.1.2 *Dual AGNs*

Due to the hierarchical growth of structure in the universe, we find multiple galaxies with more than one SMBHs, which may or may not be accreting and glowing as AGNs at the same time. The aim of this project was to get a preliminary estimate on the fraction of binary AGNs and its evolution with SMBHs from the Pop III.1 seeding mechanism (Chapter 5). To estimate this fraction, we needed to estimate the time delay between the mergers of seeded halos, and then the mergers between the hosted galaxies ($\tau_{H\rightarrow G}$). From `PINOCCHIO` we already have the redshifts of halo mergers, so we adopted

simple models to calculate the time taken for the host galaxies to merge. First, we considered a constant value for all the halos as the most simplistic case, and then we used the PinGAEA model (Cammelli et al., in prep) to get a physical estimate of this timescale based on the properties of the halos. Right after this time delay, we assumed that a binary of SMBHs is formed and both the black holes are accreting simultaneously for some time (τ_{dAGN}). Since calculating the time while both the black holes are active is out of scope of this thesis, we again assumed simple values based on Salpeter timescale (Salpeter, 1964) (the value itself, and also ten times the value).

With these timescale assumptions, we presented the fraction of dual AGNs from the Pop III.1 model, for $d_{iso} = 50, 75$, and 100 kpc cases. Regardless of the value for $\tau_{H\to G}$, the fraction for all three isolation distances tend to converge at $z = 0$, although at higher redshifts the fraction tends to be higher for smaller isolation distances. The physical model for $\tau_{H\to G}$ predicts a flatter fraction, or in other words the fraction remains more or less constant from $z = 0$ to 2. Whereas assuming a constant timescale for all the halos tends to produce stronger evolution of the fraction. The variation of τ_{dAGN} effected the overall spread in redshift space, especially for the constant $\tau_{H\to G}$ case, while higher values of τ_{dAGN} resulted in higher amplitude of the fraction from the physical model for $\tau_{H\to G}$.

Finally, we presented a simple estimate of the fraction to compare with the results from Silverman et al. (2020) who used the imaging of Hyper Suprime-Cam from Subaru Strategic Program (Miyazaki et al., 2018; Aihara et al., 2018, 2019) to confirm some dual AGN candidates from SDSS quasar sample. They measured the fraction up to $z \sim 3.5$ and found a minimal evolution. Applying the same mass cuts as mentioned in their paper but without considering the AGNs' projected separation cuts, we estimated the fraction from the physical time delay model with $\tau_{dAGN} = 500$ Myr. We found that the number of sources in our analysis were too few to make a robust comparison, but lowering the mass cuts gave us fractions similar to the measured one for redshift up to $z \sim 2$.

7.1.3 *SMBH merger rate and the GWB*

The final project of this thesis involved calculating the gravitational wave background generated by the SMBH binaries in the Pop III.1 model (Chapter 6). This work builds on the previous project in which we modelled the timescales for the halo merger up to the binary formation ($\tau_{H\to G}$). From the theoretical framework described in Chapter 2, we learned that in order to estimate the background we need the number density of SMBH mergers per unit redshift and chirp mass. The first part of this rate required us to calculate the redshifts of the SMBH binary mergers. To estimate the merger redshifts, we adopted the same timescale $\tau_{H\to G}$ as used in the dual AGN project. We further assumed that the final-parsec problem is solved, and the binaries are already in their GW-dominated orbit decay phase after the $\tau_{H\to G}$ time duration had elapsed. For the value of this timescale, we again assumed a constant value, and the value from the PinGAEA model. Then to reach the end of the evolution,

we assumed a constant lifetime of all the binaries τ_{merge}. We considered two values for this lifetime: 100 Myr and 1000 Myr. With all these timescales, we calculated the final merger redshift of all the binaries and hence the SMBH merger rate per unit redshift.

For evaluating the chirp mass for each merger, we used dark matter - black hole mass scaling relations to estimate the mass of the SMBHs based on halo masses which we obtained from `PINOCCHIO`. Using these relations we also calculated the black hole mass function at redshift $z = 0$ and compared with the observed mass functions to gain a better understanding of the scaling relations and their effect on black hole masses for the Pop III.1 model. With this information, we finally calculated the average number density of SMBH mergers per unit redshift and chirp mass from 10 simulation volumes.

Using the merger rate, we calculated the GWB for all the mass relations and timescales assumed. Comparing the results with the latest results from NANOGrav (Agazie et al., 2023a), we found that the background from the Pop III.1 tends to be slightly lower than the observed levels. We also found a strong dependence of the mass scaling relation - the relation which predicts more massive black holes had the highest amplitude of the background. The background calculated using the scaling relation from `PinGAEA`, in combination with using the $\tau_{H\rightarrow G}$ timescale from this model gives the most physically motivated results for the Pop III.1 model. The amplitude of the background is slightly low but unsurprising, since this is an expected outcome from various semi-analytical models (e.g., Sesana et al., 2008; McWilliams et al., 2014; Zhu et al., 2019; Simon, 2023).

We also found a weak dependence of the background on the isolation distance, which can be attributed to the fact that the value of the background is dominated by the population of most massive binaries, which in the Pop III.1 model tends to be the same for each isolation distance. The main difference in the number density of mergers for all the isolation distances comes from black holes around $10^5 - 10^7 M_\odot$, which do not contribute heavily to the background. But mergers of black holes in this mass range would be the prime target of the upcoming LISA, and the detections by this experiment will eventually help in constraining the isolation distances through the number density of mergers.

7.2 FUTURE PROSPECTS

The Pop III.1 model explored in this thesis offers an interesting insight into the formation of supermassive black holes, while also including physics beyond the standard model in the form of Weakly Interacting Massive Particles as dark matter candidates. Although we already made a lot of predictions for the SMBHs created with this seeding mechanism, there is still plenty of scope for further investigation of this model, which we highlight below.

- In the current implementation of the Pop III.1 model, we search for minihalos which are isolated from *every other* halo (Chapter 3). This was based on the assumption that there are feedback sources in every halo.

But we plan on including further conditions such as only considering other Pop III.1 seeded halos as feedback sources, or Pop III.2 seeded halos as well. Furthermore, we plan on introducing time dependence on the isolation radius as well, which reflects the growth of radiation bubbles more realistically.

- In the clustering analysis performed in Chapter 3, we did not consider radial scales smaller than the typical sizes of halos at lower redshifts due to the lack of substructure information. By introducing sub-halo physics via semi-analytical models, we would be able to better model the clustering at lower radial scales and make more robust predictions to compare with surveys such as the SDSS (Zehavi et al., 2011).

- In the calculation of dual AGN fractions in Chapter 5, we assumed a simplistic timescale for the visibility phase of the AGNs in the binaries. This was done because predicting the duty cycle of the AGN requires intricate information of the surrounding matter in the galaxy and a proper treatment of AGN feedback. In this regard, an immediate improvement can be made by calculating the duty cycle and the accretion rate using semi-analytical models of galaxy growth and mergers. This will provide the opportunity to have a more realistic treatment of AGN interactions and a more refined fraction. We can then also begin to physically model the separation between the AGN pairs which was neglected in our results before.

- The improvements in the modelling of dual AGNs and the galaxy interactions will directly translate to an improved treatment of the merger rate of black holes in this model (Chapter 6). This will particularly benefit the modelling of the time taken for the SMBHs to sink at the center of the host galaxy before their orbit starts decaying primarily due to GW emission. Furthermore, the latest NANOGrav results (Agazie et al., 2023) point to a potentially significant interaction of the galaxy environment with the binaries, which signifies a proper modelling of those interactions.

- Finally, as seen from the BHMFs (Chapter 6), there is a significant population of binary black holes with masses around $10^5 - 10^7 M_\odot$. In the future we plan to explore the mergers of these black hole to make predictions for LISA, which will primarily target these black hole binaries. These predictions will also help in constraining the isolation distance, since the difference in the merger rate for black holes around these masses show a strong dependence on the isolation distance.

In conclusion, the Pop III.1 model offers an interesting mechanism to form supermassive black holes in the early universe, naturally explaining the detections of high redshift quasars. The future prospects listed here will help in developing the model further and to make more robust predictions and comparisons with the ongoing and upcoming influx of high redshift data.

A

MATCHING FULL- AND LOW-RESOLUTION PINOCCHIO RUNS

We run the 59.7 Mpc box at the full resolution of 4096^3 particles and at a lower resolution of 1024^3 particles. These resolutions correspond to particle masses of $1.23 \times 10^5 M_\odot$ and $7.87 \times 10^6 M_\odot$. The minimum mass for halos has been set to 10 particles in both cases. Figure A.1 shows the mass function of the full-resolution box at high redshift, where it is evident that the early growth of massive halos is slower than in a universal model (in this case the fit to the friends-of-friends halo mass function of Crocce et al., 2010). We stress that there is no reason to believe that this analytic fit is accurate at such low masses, but we conservatively assume that the disagreement is due to an inaccuracy of PINOCCHIO.

Figure A.2 shows the halo mass function for the low-resolution box (thin lines) and the full-resolution run (thick lines). At high masses the agreement of the high-resolution box with the analytic prediction is poor, while this is not the case for the low-resolution run where the box has not been divided into different domains. Figure A.3 shows the consistency of the seeding fraction among the high-resolution box and a set of lower and lower resolution runs, where seeding of halos is decided by checking which particle in Lagrangian space contains the halos that is seeded in the full resolution box.

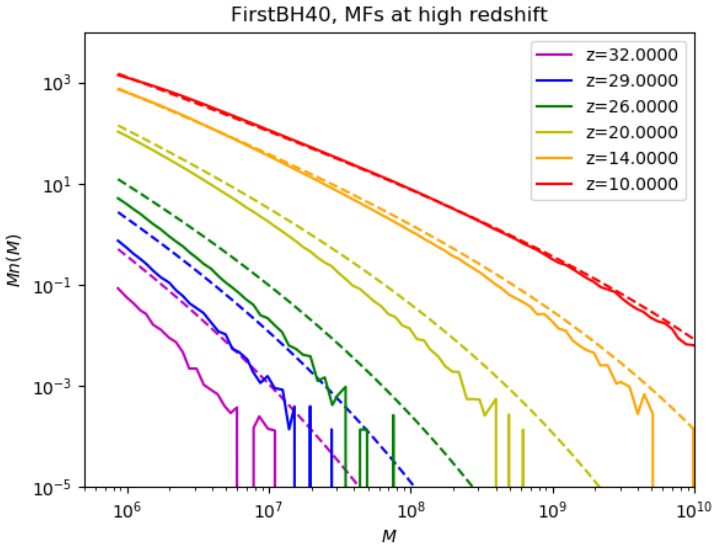


Figure A.1: Halo mass function of the full-resolution box at high redshift. Lines are color-coded in redshift (see legend). Solid lines refer to PINOCCHIO catalogs; dashed lines to the Crocce et al. (2010) analytic fit.

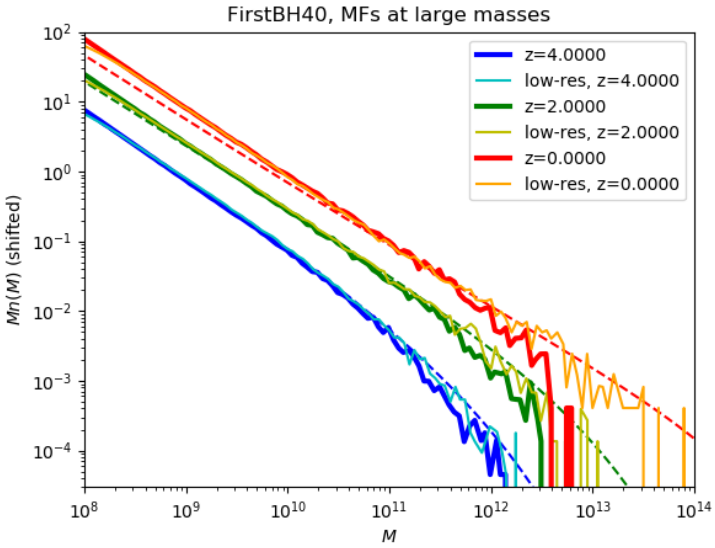


Figure A.2: Halo mass function of the full- (thick solid lines) and low-resolution (thin solid lines) boxes at low redshift. Lines are color-coded in redshift (see legend). Dashed lines are the Crocce et al. (2010) analytic fit.

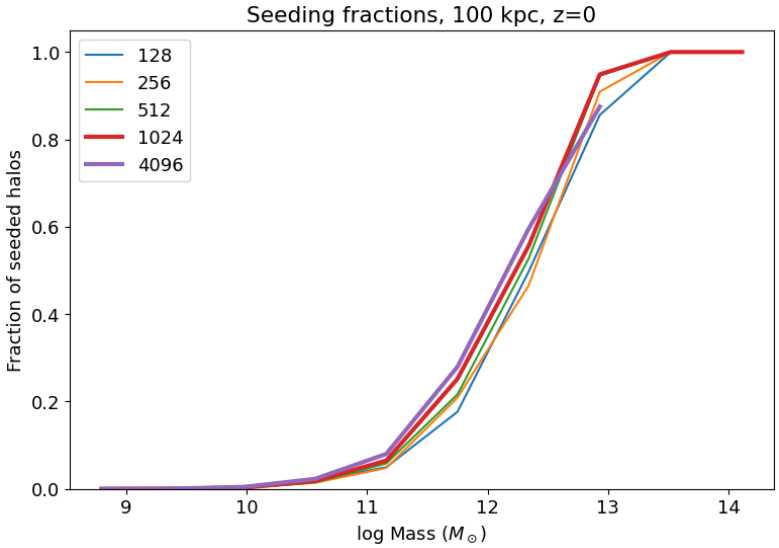


Figure A.3: Fraction of halos of a given mass that contain a seed SMBH, for $d_{iso} = 100$ kpc. Resolution is color-coded (see legend). Thicker lines emphasize the full-resolution (4096) and low-resolution (1024) runs.

B

LARGE-SCALE CLUSTERING MODES

When we use the estimators such as the CORRFUNC library to find the auto correlation of halos in our 59.7 Mpc box, the correlation function only contains the clustering modes smaller than the box size. If we want to make a simplistic comparison of our results with a large survey which sampled a much larger volume, we can do so by analytically adding the larger scale clustering modes. To understand how we achieve this, we examine the analytic expression for calculating the 3D 2pcf for halos for the entire volume of the Universe:

$$\xi_{hh}(r) = \frac{1}{2\pi^2}\int_0^\infty dk k^2 b_h^2 P(k)\frac{\sin kr}{kr}, \tag{27}$$

where ξ_{hh} is the correlation function of halos, b_h is the halo bias, and $P(k)$ is the matter power spectrum. This integral can be split in two parts:

$$\begin{aligned}\xi_{hh}(r) = &\underbrace{\frac{1}{2\pi^2}\int_0^{k_{box}} dk k^2 b_h^2 P(k)\frac{\sin kr}{kr}}_{\text{Large scale contribution } \xi_{LS}(r)} \\ &+ \underbrace{\frac{1}{2\pi^2}\int_{k_{box}}^\infty dk k^2 b_h^2 P(k)\frac{\sin kr}{kr}}_{\text{PINOCCHIO contribution } \xi_{PIN}(r)} \\ &= \xi_{LS}(r) + \xi_{PIN}(r),\end{aligned} \tag{28}$$

where $k_{box} = 2\pi/L_{box}$, with $L_{box} = 59.7$ Mpc in our box. The large scale contribution refers to the clustering modes of radial scale going from L_{box} to infinity, and the PINOCCHIO contribution refers to all the modes of radial scale from 0 to L_{box}. Since the correlation estimator returns ξ_{PIN}, we calculated the large scale contribution by using the linear matter power spectrum from CAMB python library and halo bias from COLOSSUS python library (Diemer, 2018), using the bias model of Comparat et al., 2017, and then numerically integrated the power spectrum to obtain ξ_{LS}.

To make a direct continuation of the angular clustering as shown in Banik et al., 2019 (their Figure 10), we present the angular clustering evolution of seeded halos in Figure B.1 without the large scale corrections added. This figure and Figure 3.8 essentially show the same information, with the only difference that the figure presented here is in angular scale, and without the large scale modes.

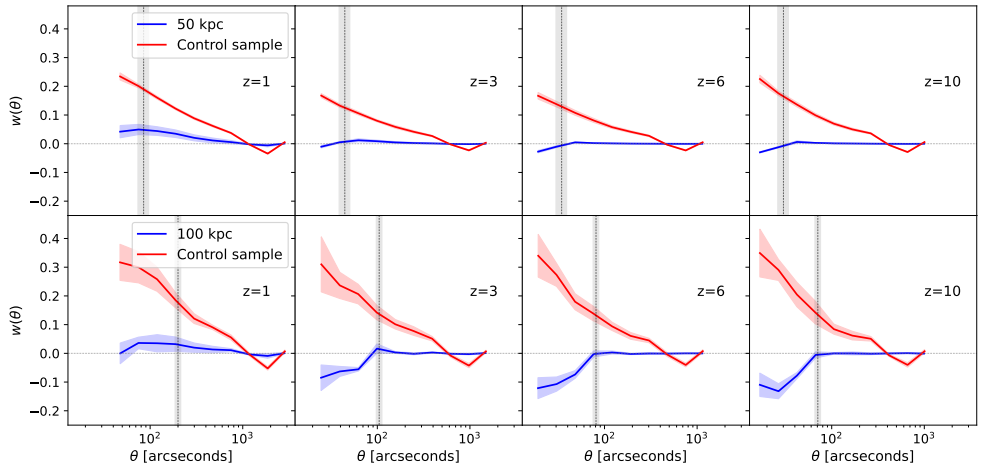


Figure B.1: Evolution of angular correlation function for 50 and 100 kpc isolation distances. The large scale modes are not added in the evaluation of this function. The labels are the same as in Figure 3.8.

BIBLIOGRAPHY

Abazajian, Kevork N. et al. (June 2009). 'The Seventh Data Release of the Sloan Digital Sky Survey'. In: *Astrophysical Journal Supplement* 182.2, pp. 543–558. DOI: 10.1088/0067-0049/182/2/543. arXiv: 0812.0649 [astro-ph].

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, B. P. et al. (Oct. 2017). 'GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral'. In: *Physical Review Letters* 119.16, 161101, p. 161101. DOI: 10.1103/PhysRevLett.119.161101. arXiv: 1710.05832 [gr-qc].

Abbott, R. et al. (2020a). 'GW190521: A Binary Black Hole Merger with a Total Mass of $150M_{\odot}$'. 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].

Afzal, Adeela et al. (2023). 'The NANOGrav 15 yr Data Set: Search for Signals from New Physics'. In: *Astrophys. J. Lett.* 951.1, p. L11. DOI: 10.3847/2041-8213/acdc91. arXiv: 2306.16219 [astro-ph.HE].

Agazie, Gabriella et al. (2023). 'The NANOGrav 15 yr Data Set: Constraints on Supermassive Black Hole Binaries from the Gravitational-wave Background'. In: *Astrophys. J. Lett.* 952.2, p. L37. DOI: 10.3847/2041-8213/ace18b. arXiv: 2306.16220 [astro-ph.HE].

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].

Aihara, Hiroaki et al. (Jan. 2018). 'The Hyper Suprime-Cam SSP Survey: Overview and survey design'. In: *Publications of the Astronomical Society of Japan* 70, S4, S4. DOI: 10.1093/pasj/psx066. arXiv: 1704.05858 [astro-ph.IM].

Aihara, Hiroaki et al. (Dec. 2019). 'Second data release of the Hyper Suprime-Cam Subaru Strategic Program'. In: *Publications of the Astronomical Society of Japan* 71.6, 114, p. 114. DOI: 10.1093/pasj/psz103. arXiv: 1905.12221 [astro-ph.IM].

Allevato, V. et al. (Nov. 2014). 'Clustering of Moderate Luminosity X-Ray-selected Type 1 and Type 2 AGNS at Z ~3'. In: *The Astrophysical Journal* 796.1, 4, p. 4. DOI: 10.1088/0004-637X/796/1/4. arXiv: 1409.7693 [astro-ph.GA].

Amaro-Seoane, Pau et al. (Feb. 2017). 'Laser Interferometer Space Antenna'. In: *arXiv e-prints*, arXiv:1702.00786, arXiv:1702.00786. DOI: 10.48550/arXiv.1702.00786. arXiv: 1702.00786 [astro-ph.IM].

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: *Monthly Notices of the Royal Astronomical Society* 460.1, pp. 1147–1161. DOI: 10.1093/mnras/stw1018. arXiv: 1602.04819 [astro-ph.GA].

Barber, Jordan, Debatri Chattopadhyay and Fabio Antonini (Jan. 2024). 'Black hole binary mergers in dense star clusters: the importance of primordial binaries'. In: *Monthly Notices of the Royal Astronomical Society* 527.3, pp. 7363–7381. DOI: 10.1093/mnras/stad3600. arXiv: 2310.02376 [astro-ph.GA].

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: *Monthly Notices of the Royal Astronomical Society* 370.1, pp. 289–298. DOI: 10.1111/j.1365-2966.2006.10467.x. arXiv: astro-ph/0602363 [astro-ph].

Behroozi, Peter S., Risa H. Wechsler and Charlie Conroy (June 2013). 'The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8'. In: *The Astrophysical Journal* 770.1, 57, p. 57. DOI: 10.1088/0004-637X/770/1/57. arXiv: 1207.6105 [astro-ph.CO].

Beifiori, A. et al. (Jan. 2012). 'On the correlations between galaxy properties and supermassive black hole mass'. In: *Monthly Notices of the Royal Astronomical Society* 419.3, pp. 2497–2528. DOI: 10.1111/j.1365-2966.2011.19903.x. arXiv: 1109.6265 [astro-ph.CO].

Berner, Pascale et al. (Nov. 2022). 'Rapid simulations of halo and subhalo clustering'. In: *Journal of Cosmology and Astroparticle Physics* 2022.11, 002, p. 002. DOI: 10.1088/1475-7516/2022/11/002. arXiv: 2112.08389 [astro-ph.CO].

Bhowmick, Aklant K. et al. (Feb. 2022a). '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. 2022b). 'Probing the z $\gtrsim$ 6 quasars in a universe with IllustrisTNG physics: impact of gas-based black hole seeding models'. In: *Monthly Notices of the Royal Astronomical Society* 516.1, pp. 138–157. DOI: 10.1093/mnras/stac2238. arXiv: 2205.05717 [astro-ph.GA].

Biava, Nadia et al. (Aug. 2019). 'The lifetime of binary black holes in Sérsic galaxy models'. In: *Monthly Notices of the Royal Astronomical Society* 487.4, pp. 4985–4994. DOI: 10.1093/mnras/stz1614. arXiv: 1903.05682 [astro-ph.GA].

Birrer, Simon et al. (Sept. 2014). 'A Simple Model Linking Galaxy and Dark Matter Evolution'. In: *The Astrophysical Journal* 793.1, 12, p. 12. DOI: 10.1088/0004-637X/793/1/12. arXiv: 1401.3162 [astro-ph.CO].

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].

Bonetti, Matteo et al. (Oct. 2016). 'Post-Newtonian evolution of massive black hole triplets in galactic nuclei - I. Numerical implementation and tests'. In: *Monthly Notices of the Royal Astronomical Society* 461.4, pp. 4419–4434. DOI: 10.1093/mnras/stw1590. arXiv: 1604.08770 [astro-ph.GA].

Bortolas, Elisa et al. (Sept. 2021). 'The Competing Effect of Gas and Stars in the Evolution of Massive Black Hole Binaries'. In: *The Astrophysical Journal Letters* 918.1, L15, p. L15. DOI: 10.3847/2041-8213/ac1c0c. arXiv: 2108.13436 [astro-ph.HE].

Boylan-Kolchin, Michael, Chung-Pei Ma and Eliot Quataert (Jan. 2008). 'Dynamical friction and galaxy merging time-scales'. In: *Monthly Notices of the Royal Astronomical Society* 383.1, pp. 93–101. DOI: 10.1111/j.1365-2966.2007.12530.x. arXiv: 0707.2960 [astro-ph].

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'. In: *The Astrophysical Journal* 596.1, pp. 34–46. DOI: 10.1086/377529. arXiv: astro-ph/0212400 [astro-ph].

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].

Burke-Spolaor, Sarah et al. (June 2019). 'The astrophysics of nanohertz gravitational waves'. In: *The Astronomy and Astrophysics Review* 27.1, 5, p. 5. DOI: 10.1007/s00159-019-0115-7. arXiv: 1811.08826 [astro-ph.HE].

Caplar, Neven, Simon J. Lilly and Benny Trakhtenbrot (Jan. 2017). 'Optical Variability of AGNs in the PTF/iPTF Survey'. In: *The Astrophysical Journal*

834.2, 111, p. 111. DOI: 10.3847/1538-4357/834/2/111. arXiv: 1611.03082 [astro-ph.GA].

Chandrasekhar, S. (Mar. 1943). 'Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction.' In: *The Astrophysical Journal* 97, p. 255. DOI: 10.1086/144517.

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. (Dec. 2016). 'Cosmological Simulations of Early Black Hole Formation: Halo Mergers, Tidal Disruption, and the Conditions for Direct Collapse'. In: *The Astrophysical Journal* 832.2, 134, p. 134. DOI: 10.3847/0004-637X/832/2/134. arXiv: 1603.08923 [astro-ph.GA].

Colpi, Monica (Sept. 2014). 'Massive Binary Black Holes in Galactic Nuclei and Their Path to Coalescence'. In: *Space Science Reviews* 183.1-4, pp. 189–221. DOI: 10.1007/s11214-014-0067-1. arXiv: 1407.3102 [astro-ph.GA].

Comparat, Johan et al. (Aug. 2017). 'Accurate mass and velocity functions of dark matter haloes'. In: *Monthly Notices of the Royal Astronomical Society* 469.4, pp. 4157–4174. DOI: 10.1093/mnras/stx1183. arXiv: 1702.01628 [astro-ph.CO].

Crocce, Martín et al. (Apr. 2010). 'Simulating the Universe with MICE: the abundance of massive clusters'. In: *Monthly Notices of the Royal Astronomical Society* 403.3, pp. 1353–1367. DOI: 10.1111/j.1365-2966.2009.16194.x. arXiv: 0907.0019 [astro-ph.CO].

Davé, Romeel, Robert Thompson and Philip F. Hopkins (Nov. 2016). 'MUFASA: galaxy formation simulations with meshless hydrodynamics'. In: *Monthly Notices of the Royal Astronomical Society* 462.3, pp. 3265–3284. DOI: 10.1093/mnras/stw1862. arXiv: 1604.01418 [astro-ph.GA].

Davis, Benjamin L., Alister W. Graham and Françoise Combes (May 2019). 'A Consistent Set of Empirical Scaling Relations for Spiral Galaxies: The (v_{max}, M_{oM})-(σ_0, M_{BH}, ϕ) Relations'. In: *The Astrophysical Journal* 877.1, 64, p. 64. DOI: 10.3847/1538-4357/ab1aa4. arXiv: 1901.06509 [astro-ph.GA].

De Rosa, Alessandra et al. (Dec. 2019). 'The quest for dual and binary supermassive black holes: A multi-messenger view'. In: 86, 101525, p. 101525. DOI: 10.1016/j.newar.2020.101525. arXiv: 2001.06293 [astro-ph.GA].

Detweiler, S. (Dec. 1979). 'Pulsar timing measurements and the search for gravitational waves'. In: *The Astrophysical Journal* 234, pp. 1100–1104. DOI: 10.1086/157593.

Di Matteo, Tiziana, Volker Springel and Lars Hernquist (Feb. 2005). 'Energy input from quasars regulates the growth and activity of black holes and their host galaxies'. In: *Nature* 433.7026, pp. 604–607. DOI: 10.1038/nature03335. arXiv: astro-ph/0502199 [astro-ph].

Di Matteo, Tiziana et al. (Aug. 2003). 'Black Hole Growth and Activity in a Λ Cold Dark Matter Universe'. In: *The Astrophysical Journal* 593.1, pp. 56–68. DOI: 10.1086/376501. arXiv: astro-ph/0301586 [astro-ph].

Di Matteo, Tiziana et al. (Mar. 2008). 'Direct Cosmological Simulations of the Growth of Black Holes and Galaxies'. In: *The Astrophysical Journal* 676.1, pp. 33–53. DOI: 10.1086/524921. arXiv: 0705.2269 [astro-ph].

Diemer, Benedikt (Dec. 2018). 'COLOSSUS: A Python Toolkit for Cosmology, Large-scale Structure, and Dark Matter Halos'. In: *Astrophysical Journal Supplement* 239.2, 35, p. 35. DOI: 10.3847/1538-4365/aaee8c. arXiv: 1712.04512 [astro-ph.CO].

Dubois, Y. et al. (Oct. 2014). 'Dancing in the dark: galactic properties trace spin swings along the cosmic web'. In: *Monthly Notices of the Royal Astronomical Society* 444.2, pp. 1453–1468. DOI: 10.1093/mnras/stu1227. arXiv: 1402.1165 [astro-ph.CO].

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.

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].

Ellison, Sara L. et al. (Dec. 2011). 'Galaxy pairs in the Sloan Digital Sky Survey - IV. Interactions trigger active galactic nuclei'. In: *Monthly Notices of the Royal Astronomical Society* 418.3, pp. 2043–2053. DOI: 10.1111/j.1365-2966.2011.19624.x. arXiv: 1108.2711 [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.

Fabjan, D. et al. (Jan. 2010). 'Simulating the effect of active galactic nuclei feedback on the metal enrichment of galaxy clusters'. In: *Monthly Notices of the Royal Astronomical Society* 401.3, pp. 1670–1690. DOI: 10.1111/j.1365-2966.2009.15794.x. arXiv: 0909.0664 [astro-ph.CO].

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* 578.1, pp. 90–97. DOI: 10.1086/342308. arXiv: astro-ph/0203469 [astro-ph].

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].

Fontanot, Fabio et al. (Aug. 2020). 'The rise of active galactic nuclei in the galaxy evolution and assembly semi-analytic model'. In: *Monthly Notices of the Royal Astronomical Society* 496.3, pp. 3943–3960. DOI: 10.1093/mnras/staa1716. arXiv: 2002.10576 [astro-ph.CO].

Foster, R. S. and D. C. Backer (Sept. 1990). 'Constructing a Pulsar Timing Array'. In: *The Astrophysical Journal* 361, p. 300. DOI: 10.1086/169195.

Freese, Katherine et al. (2010). 'SUPERMASSIVE DARK STARS: DETECTABLE INJWST'. In: *The Astrophysical Journal* 716.2, pp. 1397–1407. DOI: 10.1088/0004-637x/716/2/1397. URL: https://doi.org/10.1088/0004-637x/716/2/1397.

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].

Greif, Thomas H. and Volker Bromm (Nov. 2006). '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].

Habouzit, Mélanie et al. (Nov. 2016). 'On the number density of 'direct collapse' black hole seeds'. In: *Monthly Notices of the Royal Astronomical Society* 463.1, pp. 529–540. DOI: 10.1093/mnras/stw1924. arXiv: 1601.00557 [astro-ph.GA].

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. 2023). 'JWST/NIRSpec First Census of Broad-Line AGNs at z=4-7: Detection of 10 Faint AGNs with M_BH$\sim 10^{6}$-10^{7} M_sun and Their Host Galaxy Properties'. In: *arXiv e-prints*, arXiv:2303.11946, arXiv:2303.11946. DOI: 10.48550/arXiv.2303.11946. arXiv: 2303.11946 [astro-ph.GA].

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].

He, Wanqiu et al. (Jan. 2018). 'Clustering of quasars in a wide luminosity range at redshift 4 with Subaru Hyper Suprime-Cam Wide-field imaging'. In: *Publications of the Astronomical Society of Japan* 70, S33, S33. DOI: 10.1093/pasj/psx129. arXiv: 1704.08461 [astro-ph.GA].

Hellings, R. W. and G. S. Downs (Feb. 1983). 'Upper limits on the isotropic gravitational radiation background from pulsar timing analysis.' In: *The Astrophysical Journal Letters* 265, pp. L39–L42. DOI: 10.1086/183954.

Hirschmann, Michaela, Gabriella De Lucia and Fabio Fontanot (Sept. 2016). 'Galaxy assembly, stellar feedback and metal enrichment: the view from the GAEA model'. In: *Monthly Notices of the Royal Astronomical Society* 461.2, pp. 1760–1785. DOI: 10.1093/mnras/stw1318. arXiv: 1512.04531 [astro-ph.GA].

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].

Holley-Bockelmann, Kelly and Fazeel Mahmood Khan (Sept. 2015). 'Galaxy Rotation and Rapid Supermassive Binary Coalescence'. In: *The Astrophysical Journal* 810.2, 139, p. 139. DOI: 10.1088/0004-637X/810/2/139. arXiv: 1505.06203 [astro-ph.GA].

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].

Hughes, Scott A. (Nov. 2005). 'Trust but verify: The case for astrophysical black holes'. In: *arXiv e-prints*, hep-ph/0511217, hep–ph/0511217. DOI: 10.48550/arXiv.hep-ph/0511217. arXiv: hep-ph/0511217 [hep-ph].

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, pp. 27–97. DOI: 10.1146/annurev-astro-120419-014455. arXiv: 1911.05791 [astro-ph.GA].

Jaffe, A. H. and D. C. Backer (Feb. 2003). 'Gravitational Waves Probe the Coalescence Rate of Massive Black Hole Binaries'. In: *The Astrophysical Journal* 583.2, pp. 616–631. DOI: 10.1086/345443. arXiv: astro-ph/0210148 [astro-ph].

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, Jenny E. Greene and Luis C. Ho (Aug. 2011). 'Black Hole Mass and Bulge Luminosity for Low-mass Black Holes'. In: *The Astrophysical Journal Letters* 737.2, L45, p. L45. DOI: 10.1088/2041-8205/737/2/L45. arXiv: 1107.4103 [astro-ph.CO].

Johnson, Jarrett L. and Volker Bromm (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].

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].

Kelley, Luke Zoltan, Laura Blecha and Lars Hernquist (Jan. 2017). 'Massive black hole binary mergers in dynamical galactic environments'. In: *Monthly Notices of the Royal Astronomical Society* 464.3, pp. 3131–3157. DOI: 10.1093/mnras/stw2452. arXiv: 1606.01900 [astro-ph.HE].

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.

Khan, Fazeel Mahmood, Andreas Just and David Merritt (May 2011). 'Efficient Merger of Binary Supermassive Black Holes in Merging Galaxies'. In: *The Astrophysical Journal* 732.2, 89, p. 89. DOI: 10.1088/0004-637X/732/2/89. arXiv: 1103.0272 [astro-ph.CO].

Kocevski, Dale D. et al. (Dec. 2015). 'Are Compton-thick AGNs the Missing Link between Mergers and Black Hole Growth?' In: *The Astrophysical Journal* 814.2, 104, p. 104. DOI: 10.1088/0004-637X/814/2/104. arXiv: 1509.03629 [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].

Koss, Michael J. et al. (Nov. 2018). 'A population of luminous accreting black holes with hidden mergers'. In: *Nature* 563.7730, pp. 214–216. DOI: 10.1038/s41586-018-0652-7. arXiv: 1811.03641 [astro-ph.GA].

Kroupa, Pavel et al. (Nov. 2020). 'Very high redshift quasars and the rapid emergence of supermassive black holes'. In: *Monthly Notices of the Royal Astronomical Society* 498.4, pp. 5652–5683. DOI: 10.1093/mnras/staa2276. arXiv: 2007.14402 [astro-ph.GA].

Kulier, Andrea et al. (Feb. 2015). 'Understanding Black Hole Mass Assembly via Accretion and Mergers at Late Times in Cosmological Simulations'. In: *The Astrophysical Journal* 799.2, 178, p. 178. DOI: 10.1088/0004-637X/799/2/178. arXiv: 1307.3684 [astro-ph.CO].

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].

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.

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<z<7 found by JWST are not inconsistent with the local $\mathcal{M}_{\rm BH}$-$\mathcal{M}_{\star}$ relation'. In: *arXiv e-prints*, arXiv:2403.00074, arXiv:2403.00074. DOI: 10.48550/arXiv.2403.00074. arXiv: 2403.00074 [astro-ph.GA].

Li, Zefeng et al. (July 2018). 'The Ensemble Photometric Variability of Over 10^5 Quasars in the Dark Energy Camera Legacy Survey and the Sloan Digital

Sky Survey'. In: *The Astrophysical Journal* 861.1, 6, p. 6. DOI: 10.3847/1538-4357/aac6ce. arXiv: 1805.07747 [astro-ph.GA].

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.

Madau, Piero and Martin J. Rees (Apr. 2001). 'Massive Black Holes as Population III Remnants'. In: *The Astrophysical Journal Letters* 551.1, pp. L27–L30. DOI: 10.1086/319848. arXiv: astro-ph/0101223 [astro-ph].

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. (May 2023a). 'A small and vigorous black hole in the early Universe'. In: *arXiv e-prints*, arXiv:2305.12492, arXiv:2305.12492. DOI: 10.48550/arXiv.2305.12492. arXiv: 2305.12492 [astro-ph.GA].

Maiolino, Roberto et al. (Aug. 2023b). '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].

Mannucci, F. et al. (Aug. 2022). 'Unveiling the population of dual and lensed active galactic nuclei at sub-arcsec separations'. In: *Nature Astronomy* 6, pp. 1185–1192. DOI: 10.1038/s41550-022-01761-5. arXiv: 2203.11234 [astro-ph.GA].

McKee, Christopher F. and Jonathan C. Tan (2008). '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.

McWilliams, Sean T., Jeremiah P. Ostriker and Frans Pretorius (July 2014). 'Gravitational Waves and Stalled Satellites from Massive Galaxy Mergers at z <= 1'. In: *The Astrophysical Journal* 789.2, 156, p. 156. DOI: 10.1088/0004-637X/789/2/156. arXiv: 1211.5377 [astro-ph.CO].

Merritt, David and Milos Milosavljevic (2005). 'Massive black hole binary evolution'. In: *Living Rev. Rel.* 8, p. 8. DOI: 10.12942/lrr-2005-8. arXiv: astro-ph/0410364.

Michell, John (1784). 'On the means of discovering the distance, magnitude, c. 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.

Milosavljević, Milos and David Merritt (Dec. 2001). 'Formation of Galactic Nuclei'. In: *The Astrophysical Journal* 563.1, pp. 34–62. DOI: 10.1086/323830. arXiv: astro-ph/0103350 [astro-ph].

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].

Miyazaki, Satoshi et al. (Jan. 2018). 'Hyper Suprime-Cam: System design and verification of image quality'. In: *Publications of the Astronomical Society of Japan* 70, S1, S1. DOI: 10.1093/pasj/psx063.

Monaco, P. et al. (Aug. 2013). 'An accurate tool for the fast generation of dark matter halo catalogues'. In: *Monthly Notices of the Royal Astronomical Society* 433.3, pp. 2389–2402. DOI: 10.1093/mnras/stt907. arXiv: 1305.1505 [astro-ph.CO].

Monaco, Pierluigi (Oct. 2016). 'Approximate Methods for the Generation of Dark Matter Halo Catalogs in the Age of Precision Cosmology'. In: *Galaxies* 4.4, 53, p. 53. DOI: 10.3390/galaxies4040053. arXiv: 1605.07752 [astro-ph.CO].

Monaco, Pierluigi, Tom Theuns and Giuliano Taffoni (Apr. 2002). 'The pinocchio algorithm: pinpointing orbit-crossing collapsed hierarchical objects in a linear density field'. In: *Monthly Notices of the Royal Astronomical Society* 331.3, pp. 587–608. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2002.05162.x. eprint: https://academic.oup.com/mnras/article-pdf/331/3/587/3360787/331-3-587.pdf. URL: https://doi.org/10.1046/j.1365-8711.2002.05162.x.

Montero, Pedro J., Hans-Thomas Janka and Ewald Müller (Apr. 2012). 'Relativistic Collapse and Explosion of Rotating Supermassive Stars with Thermonuclear Effects'. In: *The Astrophysical Journal* 749.1, 37, p. 37. DOI: 10.1088/0004-637X/749/1/37. arXiv: 1108.3090 [astro-ph.CO].

Moutarde, F. et al. (Dec. 1991). 'Precollapse Scale Invariance in Gravitational Instability'. In: *The Astrophysical Journal* 382, p. 377. DOI: 10.1086/170728.

Munari, Emiliano et al. (Mar. 2017). 'Improving fast generation of halo catalogues with higher order Lagrangian perturbation theory'. In: *Monthly Notices of the Royal Astronomical Society* 465.4, pp. 4658–4677. DOI: 10.1093/mnras/stw3085. arXiv: 1605.04788 [astro-ph.CO].

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'. In: *The Astrophysical Journal* 830.2, 117, p. 117. DOI: 10.3847/0004-637X/830/2/117. arXiv: 1607.07325 [astro-ph.GA].

Mutlu-Pakdil, Burçin et al. (Feb. 2018). 'The Illustris simulation: supermassive black hole-galaxy connection beyond the bulge'. In: *Monthly Notices of the Royal Astronomical Society* 474.2, pp. 2594–2606. DOI: 10.1093/mnras/stx2935. arXiv: 1711.03958 [astro-ph.GA].

Nakajima, Kimihiko et al. (Jan. 2023). 'JWST Census for the Mass-Metallicity Star-Formation Relations at z=4-10 with the Self-Consistent Flux Calibration and the Proper Metallicity Calibrators'. In: *arXiv e-prints*, arXiv:2301.12825, arXiv:2301.12825. DOI: 10.48550/arXiv.2301.12825. arXiv: 2301.12825 [astro-ph.GA].

Natarajan, Aravind, Jonathan C. Tan and Brian W. O'Shea (2009). 'DARK MATTER ANNIHILATION AND PRIMORDIAL STAR FORMATION'. In: *The Astrophysical Journal* 692.1, pp. 574–583. DOI: 10.1088/0004-637x/692/1/574. URL: https://doi.org/10.1088/0004-637x/692/1/574.

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.

Norberg, Peder et al. (Nov. 2002). 'The 2dF Galaxy Redshift Survey: the bJ-band galaxy luminosity function and survey selection function'. In: *Monthly Notices of the Royal Astronomical Society* 336.3, pp. 907–931. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2002.05831.x. eprint: https://academic.oup.com/mnras/article-pdf/336/3/907/2959521/336-3-907.pdf. URL: https://doi.org/10.1046/j.1365-8711.2002.05831.x.

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].

Oppenheimer, J. R. and H. Snyder (Sept. 1939). 'On Continued Gravitational Contraction'. In: *Physical Review* 56.5, pp. 455–459. DOI: 10.1103/PhysRev.56.455.

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: *The Astrophysical Journal Letters* 957.1, L3, p. L3. DOI: 10.3847/2041-8213/ad0158. arXiv: 2308.12331 [astro-ph.GA].

Perna, Michele et al. (Oct. 2023). 'A surprisingly high number of dual active galactic nuclei in the early Universe'. In: *arXiv e-prints*, arXiv:2310.03067, arXiv:2310.03067. DOI: 10.48550/arXiv.2310.03067. arXiv: 2310.03067 [astro-ph.GA].

Phinney, E. S. (July 2001). 'A Practical theorem on gravitational wave backgrounds'. In: arXiv: astro-ph/0108028.

Pirani, F. A. E. (1956). 'On the Physical significance of the Riemann tensor'. In: *Acta Phys. Polon.* 15, pp. 389–405. DOI: 10.1007/s10714-009-0787-9.

— (May 2009). 'Republication of: On the physical significance of the Riemann tensor'. In: *General Relativity and Gravitation* 41.5, pp. 1215–1232. DOI: 10.1007/s10714-009-0787-9.

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].

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].

Quinlan, Gerald D. (July 1996). 'The dynamical evolution of massive black hole binaries I. Hardening in a fixed stellar background'. In: 1.1, pp. 35–56. DOI: 10.1016/S1384-1076(96)00003-6. arXiv: astro-ph/9601092 [astro-ph].

Rajagopal, Mohan and Roger W. Romani (June 1995). 'Ultra–Low-Frequency Gravitational Radiation from Massive Black Hole Binaries'. In: *The Astrophysical Journal* 446, p. 543. DOI: 10.1086/175813. arXiv: astro-ph/9412038 [astro-ph].

Ravi, V. et al. (Mar. 2015). 'Prospects for gravitational-wave detection and supermassive black hole astrophysics with pulsar timing arrays'. In: *Monthly Notices of the Royal Astronomical Society* 447.3, pp. 2772–2783. DOI: 10.1093/mnras/stu2659. arXiv: 1406.5297 [astro-ph.CO].

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].

Rindler-Daller, T. et al. (2015). 'DARK STARS: IMPROVED MODELS AND FIRST PULSATION RESULTS'. In: *The Astrophysical Journal* 799.2, p. 210. DOI: 10.1088/0004-637x/799/2/210. URL: https://doi.org/10.1088/0004-637x/799/2/210.

Romani, R.W. (1989). 'Timing a millisecond pulsar timing array'. In: *Timing Neutron Stars*. Springer Netherlands.

Sabra, Bassem M. et al. (Apr. 2015). 'The Black Hole-Dark Matter Halo Connection'. In: *The Astrophysical Journal* 803.1, 5, p. 5. DOI: 10.1088/0004-637X/803/1/5. arXiv: 1502.00775 [astro-ph.GA].

Salpeter, E. E. (Aug. 1964). 'Accretion of Interstellar Matter by Massive Objects.' In: *The Astrophysical Journal* 140, pp. 796–800. DOI: 10.1086/147973.

Sánchez-Sáez, P. et al. (May 2019). 'The QUEST-La Silla AGN Variability Survey: Selection of AGN Candidates through Optical Variability'. In: *Astrophysical Journal Supplement* 242.1, 10, p. 10. DOI: 10.3847/1538-4365/ab174f. arXiv: 1904.04844 [astro-ph.GA].

Sazhin, M. V. (Feb. 1978). 'Opportunities for detecting ultralong gravitational waves'. In: *Soviet Astronomy* 22, pp. 36–38.

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

Sesana, A., A. Vecchio and C. N. Colacino (Oct. 2008). 'The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays'. In: *Monthly*

Notices of the Royal Astronomical Society 390.1, pp. 192–209. DOI: 10.1111/j.1365-2966.2008.13682.x. arXiv: 0804.4476 [astro-ph].

Sesana, Alberto (Aug. 2010). 'Self Consistent Model for the Evolution of Eccentric Massive Black Hole Binaries in Stellar Environments: Implications for Gravitational Wave Observations'. In: *The Astrophysical Journal* 719.1, pp. 851–864. DOI: 10.1088/0004-637X/719/1/851. arXiv: 1006.0730 [astro-ph.CO].

Sesana, Alberto, Francesco Haardt and Piero Madau (Nov. 2006). 'Interaction of Massive Black Hole Binaries with Their Stellar Environment. I. Ejection of Hypervelocity Stars'. In: *The Astrophysical Journal* 651.1, pp. 392–400. DOI: 10.1086/507596. arXiv: astro-ph/0604299 [astro-ph].

Sesana, Alberto and Fazeel Mahmood Khan (Nov. 2015). 'Scattering experiments meet N-body - I. A practical recipe for the evolution of massive black hole binaries in stellar environments'. In: *Monthly Notices of the Royal Astronomical Society* 454.1, pp. L66–L70. DOI: 10.1093/mnrasl/slv131. arXiv: 1505.02062 [astro-ph.GA].

Sesana, Alberto et al. (Aug. 2004). 'Low-Frequency Gravitational Radiation from Coalescing Massive Black Hole Binaries in Hierarchical Cosmologies'. In: *The Astrophysical Journal* 611.2, pp. 623–632. DOI: 10.1086/422185. arXiv: astro-ph/0401543 [astro-ph].

Severgnini, P. et al. (Feb. 2021). 'A possible sub-kiloparsec dual AGN buried behind the galaxy curtain'. In: *Astronomy Astrophysics* 646, A153, A153. DOI: 10.1051/0004-6361/202039576. arXiv: 2012.09184 [astro-ph.GA].

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].

Shinohara, Takumi et al. (Apr. 2023). 'Supermassive primordial black holes: a view from clustering of quasars at $z \sim 6$'. 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: *Monthly Notices of the Royal Astronomical Society* 380.3, pp. 877–900. DOI: 10.1111/j.1365-2966.2007.12153.x. arXiv: 0705.2238 [astro-ph].

Sijacki, Debora et al. (July 2015). 'The Illustris simulation: the evolving population of black holes across cosmic time'. In: *Monthly Notices of the Royal Astronomical Society* 452.1, pp. 575–596. ISSN: 0035-8711. DOI: 10.1093/mnras/stv1340. eprint: https://academic.oup.com/mnras/article-pdf/452/1/575/4931846/stv1340.pdf. URL: https://doi.org/10.1093/mnras/stv1340.

Silverman, J. D. et al. (Dec. 2011). 'The Impact of Galaxy Interactions on Active Galactic Nucleus Activity in zCOSMOS'. In: *The Astrophysical Journal* 743.1, 2, p. 2. DOI: 10.1088/0004-637X/743/1/2. arXiv: 1109.1292 [astro-ph.CO].

Silverman, John D. et al. (Aug. 2020). 'Dual Supermassive Black Holes at Close Separation Revealed by the Hyper Suprime-Cam Subaru Strategic Program'. In: *The Astrophysical Journal* 899.2, 154, p. 154. DOI: 10.3847/1538-4357/aba4a3. arXiv: 2007.05581 [astro-ph.GA].

Simon, Joseph (June 2023). 'Exploring Proxies for the Supermassive Black Hole Mass Function: Implications for Pulsar Timing Arrays'. In: *The Astrophysical Journal Letters* 949.2, L24, p. L24. DOI: 10.3847/2041-8213/acd18e. arXiv: 2306.01832 [astro-ph.GA].

Singh, Jasbir, Pierluigi Monaco and Jonathan C. Tan (Oct. 2023). 'The formation of supermassive black holes from Population III.1 seeds. II. Evolution to the local universe'. In: *Monthly Notices of the Royal Astronomical Society* 525.1, pp. 969–982. DOI: 10.1093/mnras/stad2346. arXiv: 2301.11464 [astro-ph.GA].

Sinha, Manodeep and Lehman H. Garrison (2020). 'CORRFUNC - a suite of blazing fast correlation functions on the CPU'. In: *Monthly Notices of the Royal Astronomical Society* 491.2, pp. 3022–3041. DOI: 10.1093/mnras/stz3157.

Spolyar, Douglas, Katherine Freese and Paolo Gondolo (2008). 'Dark Matter and the First Stars: A New Phase of Stellar Evolution'. In: *Phys. Rev. Lett.* 100 (5), p. 051101. DOI: 10.1103/PhysRevLett.100.051101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.051101.

Springel, Volker, Tiziana Di Matteo and Lars Hernquist (Aug. 2005a). '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].

Springel, Volker et al. (June 2005b). 'Simulations of the formation, evolution and clustering of galaxies and quasars'. In: *Nature* 435.7042, pp. 629–636. DOI: 10.1038/nature03597. arXiv: astro-ph/0504097 [astro-ph].

Springel, Volker et al. (Mar. 2018). 'First results from the IllustrisTNG simulations: matter and galaxy clustering'. In: *Monthly Notices of the Royal Astronomical Society* 475.1, pp. 676–698. DOI: 10.1093/mnras/stx3304. arXiv: 1707.03397 [astro-ph.GA].

Steinborn, Lisa K. et al. (Apr. 2015). 'A refined sub-grid model for black hole accretion and AGN feedback in large cosmological simulations'. In: *Monthly Notices of the Royal Astronomical Society* 448.2, pp. 1504–1525. DOI: 10.1093/mnras/stv072. arXiv: 1409.3221 [astro-ph.GA].

Stone, Meredith A. et al. (Mar. 2024). 'Undermassive Host Galaxies of Five z ~ 6 Luminous Quasars Detected with JWST'. In: *The Astrophysical Journal* 964.1, 90, p. 90. DOI: 10.3847/1538-4357/ad2a57. arXiv: 2310.18395 [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].

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.

Thorne, K.S. (1987). 'Gravitational radiation'. In: *Three Hundred Years of Gravitation*. Cambridge University Press.

Trebitsch, Maxime et al. (Sept. 2021). 'The OBELISK simulation: Galaxies contribute more than AGN to H I reionization of protoclusters'. In: *Astronomy Astrophysics* 653, A154, A154. DOI: 10.1051/0004-6361/202037698. arXiv: 2002.04045 [astro-ph.GA].

Tremmel, M. et al. (Sept. 2017). 'The Romulus cosmological simulations: a physical approach to the formation, dynamics and accretion models of SMBHs'. In: *Monthly Notices of the Royal Astronomical Society* 470.1, pp. 1121–1139. DOI: 10.1093/mnras/stx1160. arXiv: 1607.02151 [astro-ph.GA].

Tripodi, R. et al. (Jan. 2024). 'HYPERION. The SMBH-galaxy co-evolution at $z > 6$ and the build-up of massive galaxies'. In: *arXiv e-prints*, arXiv:2401.04211, arXiv:2401.04211. DOI: 10.48550/arXiv.2401.04211. arXiv: 2401.04211 [astro-ph.GA].

Valentini, Milena et al. (Jan. 2020). 'Impact of AGN feedback on galaxies and their multiphase ISM across cosmic time'. In: *Monthly Notices of the Royal Astronomical Society* 491.2, pp. 2779–2807. DOI: 10.1093/mnras/stz3131. arXiv: 1911.02572 [astro-ph.GA].

Vogelsberger, Mark et al. (Oct. 2014). 'Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe'. In: *Monthly Notices of the Royal Astronomical Society* 444.2, pp. 1518–1547. DOI: 10.1093/mnras/stu1536. arXiv: 1405.2921 [astro-ph.CO].

Volonteri, M. et al. (Aug. 2016). 'The cosmic evolution of massive black holes in the Horizon-AGN simulation'. In: *Monthly Notices of the Royal Astronomical Society* 460.3, pp. 2979–2996. DOI: 10.1093/mnras/stw1123. arXiv: 1602.01941 [astro-ph.GA].

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 (Sept. 2021). 'The origins of massive black holes'. In: *Nature Reviews Physics* 3.11, pp. 732–743. DOI: 10.1038/s42254-021-00364-9. arXiv: 2110.10175 [astro-ph.GA].

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: *The Astrophysical Journal Letters* 907.1, L1, p. L1. DOI: 10.3847/2041-8213/abd8c6. arXiv: 2101.03179 [astro-ph.GA].

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].

Wyithe, J. Stuart B. and Abraham Loeb (June 2003). 'Low-Frequency Gravitational Waves from Massive Black Hole Binaries: Predictions for LISA and Pulsar Timing Arrays'. In: *The Astrophysical Journal* 590.2, pp. 691–706. DOI: 10.1086/375187. arXiv: astro-ph/0211556 [astro-ph].

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].

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].

York, Donald G. et al. (Sept. 2000). 'The Sloan Digital Sky Survey: Technical Summary'. In: *The Astronomical Journal* 120.3, pp. 1579–1587. DOI: 10.1086/301513. arXiv: astro-ph/0006396 [astro-ph].

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].

Yue, Minghao et al. (Nov. 2021). 'A Candidate Kiloparsec-scale Quasar Pair at $z = 5.66$'. In: *The Astrophysical Journal Letters* 921.2, L27, p. L27. DOI: 10.3847/2041-8213/ac31a9. arXiv: 2110.12315 [astro-ph.GA].

Yue, Minghao et al. (Sept. 2023). 'EIGER V. Characterizing the Host Galaxies of Luminous Quasars at $z\mathrm{rsim}6$'. In: *arXiv e-prints*, arXiv:2309.04614, arXiv:2309.04614. DOI: 10.48550/arXiv.2309.04614. arXiv: 2309.04614 [astro-ph.GA].

Zehavi, Idit et al. (July 2011). 'Galaxy Clustering in the Completed SDSS Redshift Survey: The Dependence on Color and Luminosity'. In: *The Astrophysical Journal* 736.1, 59, p. 59. DOI: 10.1088/0004-637X/736/1/59. arXiv: 1005.2413 [astro-ph.CO].

Zentner, Andrew R. et al. (May 2005). 'The Physics of Galaxy Clustering. I. A Model for Subhalo Populations'. In: *The Astrophysical Journal* 624.2, pp. 505–525. DOI: 10.1086/428898. arXiv: astro-ph/0411586 [astro-ph].

Zhu, Xing-Jiang, Weiguang Cui and Eric Thrane (Jan. 2019). 'The minimum and maximum gravitational-wave background from supermassive binary black holes'. In: *Monthly Notices of the Royal Astronomical Society* 482.2, pp. 2588–2596. DOI: 10.1093/mnras/sty2849. arXiv: 1806.02346 [astro-ph.GA].

COLOPHON

The formation of supermassive black holes from Population III.1 seeds – implications for clustering, binarity and gravitational waves,

Università degli Studi di Trieste Ciclo XXXVI

Final Version as of 8th April 2024 (.