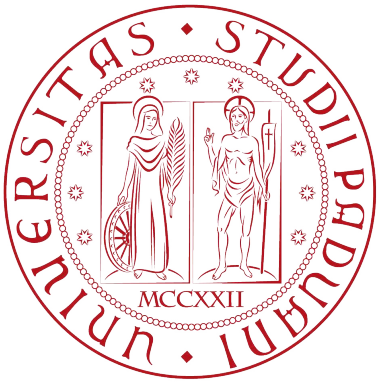


UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Astrophysics and Cosmology

Final Dissertation

Hierarchical Black Hole Mergers in Active Galactic Nuclei

Thesis supervisor:
Prof. Michela Mapelli
Thesis co-supervisor:
Dr. Maria Celeste Artale

Candidate:
Maria Paola Vaccaro

Academic Year 2021/22

Abstract

This work explores the active galactic nucleus (AGN) disk scenario for hierarchical black hole (BH) mergers.

In AGNs, a central super-massive BH (SMBH) is surrounded by an extremely dense gaseous accretion disk. The stellar-sized BHs orbiting the SMBH are thus subject to strong gas torques and experience damping: their orbital eccentricity and inclination with respect to the disk plane are suppressed. Some of the stellar BHs eventually end up orbiting inside the disk, and suffer Type I migration, similar to what happens to planets in protoplanetary disks.
The disk may develop migration traps, where migration stalls and BHs accumulate. This enhances interactions among BHs, with consequent efficient binary formation and, thanks to gas hardening, rapid merger. Because of the deep gravitational potential of the SMBH, the merger remnants are usually retained in the system and they can often go through multiple episodes of pair-up and merger.
Hence, AGNs are promising environments for the formation of BHs in the upper mass gap with $M_{\rm BH} \in [50, 130] M_\odot$ and intermediate-mass black holes (IMBHs) with $M_{\rm BH} \in [10^2, 10^5] M_\odot$.

In this Thesis work, we have developed a semi-analytical code which allows to easily explore BBH production in such environments, without the need of extensive N-body hydrodynamical simulations. Our model includes a formalism for BBH formation in migration traps and BBH evolution, as well as prescriptions for BH spin orientations.

The physical parameters of the system, such as the mass of the SMBH and the density and thickness of the disk, strongly influence BH dynamics. Thanks to the computational efficiency of our code, we are able to explore the properties of BBH mergers in a variety of disk models. We find that, in disks with medium to high gas density or with low thickness, there is a large production of remnant BHs with masses up to a few $\times 10^3 M_\odot$. On the other hand, disks that are thick and diluted are not able to prompt efficient migration, thus they produce no BBH mergers.

Finally, we qualitatively compare our simulation outputs with gravitational-wave data from the LIGO–Virgo–KAGRA collaboration and we deduce that high-mass BBH merger detections could have been produced in AGN disks.

Acknowledgements

Words cannot express my gratitude to Michela Mapelli for introducing me to the fascinating topic of black hole dynamics, for providing the most helpful insights at times when I struggled to sort problems out and, most importantly, for encouraging me every step of the way during this project.

Thanks should also go to the whole DEMOBLACK group and especially to Maria Celeste Artale who significantly helped me out in analyzing Illustris TNG data.

I am grateful to Dario Barone, who worked on improving the user-friendliness of the FASTCLUSTER code and with whom I had many fruitful discussions on the topics of this thesis.

Contents

List of Figures

List of Tables

List of Acronyms

Acronym	Definition
BH	black hole
IMBH	intermediate-mass black hole
SMBH	supermassive black hole
GW	gravitational wave
LVK	LIGO-Virgo-KAGRA
GWTC	GW transient catalog
ISCO	innermost stable circular orbit
BBH	binary black hole
GC	globular cluster
NSC	nuclear star cluster
AGN	active galactic nucleus
IR	infrared
UV	ultraviolet
SSD	Shakura-Sunyaev disk
SG	Sirko-Goodman
TQM	Thompson-Quataert-Murray
IGM	inter-galactic medium
SN	supernova
EM	electromagnetic
DLt	high-density, long-lived, thin
DLT	high-density, long-lived, thick
Dst	high-density, short-lived, thin
DsT	high-density, short-lived, thick
dLt	medium-density, long-lived, thin
dLT	medium-density, long-lived, thick
dst	medium-density, short-lived, thin
dsT	medium-density, short-lived, thick
dLt	low-density, long-lived, thin
dLT	low-density, long-lived, thick
dst	low-density, short-lived, thin
dsT	low-density, short-lived, thick

Chapter 1

Introduction

Recent observations of gravitational waves have provided invaluable information on the population of black holes in the Universe.

In this Chapter we will briefly introduce the nature of black holes, discuss gravitational waves emission and detection, and outline the formation scenarios for binary black holes.

1.1 Black Holes

A black hole (BH) is an incredibly compact object such that all of its mass $M_{\rm BH}$ is contained inside its event horizon. For a non-spinning BH, the radius of the event horizon is defined as the Schwarzshild radius $R_S = 2GM_{\rm BH}/c^2$, where G is the gravity constant, c the speed of light and $M_{\rm BH}$ is the mass of the BH. At the event horizon, the gravity is so strong that nothing, not even photons, can escape from it.

As a comparison, a star like the Sun has a mass $M_\odot \simeq 2 \times 10^{30}\,\mathrm{kg}$, a radius $R_\odot \simeq 7 \times 10^{8}\,\mathrm{m}$ and a density of roughly $\rho_\odot \simeq 1400\,\mathrm{kg/m^3}$. If it were a BH, the same mass would be concentrated in a much smaller sphere of radius $2GM_\odot/c^2 \simeq 3\,\mathrm{km}$ and density of roughly $2 \times 10^{19}\,\mathrm{kg/m^3}$.

Few people know that the idea of the existence of BHs was first postulated in 1783 by Rev. John Michell [64]. He argued that a sufficiently compact star may have a surface escape velocity larger than the speed of light and would thus be invisible. He called these objects 'dark stars'. This idea had very little traction on his contemporaries and was forgotten until his writings re-surfaced in the 1970s [14].
A proper mathematical treatment of BHs had to await 1916, when Albert Einstein formulated the theory of General Relativity and, shortly after, Karl Schwarzschild found a spherically symmetric solution of Einstein's equations in the vacuum. He demonstrated the existence of a characteristic event horizon, the Schwarzschild radius R_S, within which no communication is possible with external observers. Later, Roy Kerr generalized this solution to spinning black holes in 1963 and finally John Wheeler coined the term 'black hole' in 1968. [25]

Nowadays, we have astronomical evidence for the existence of tens of BHs in the local Universe, and we classify them based on their mass:

1. **Stellar BHs** with masses $M_{\rm BH} < 100 M_\odot$,
2. **Intermediate BHs** with masses $100 M_\odot \leq M_{\rm BH} < 10^5\, M_\odot$,
3. **Supermassive BHs** with masses $M_{\rm BH} \geq 10^5\, M_\odot$.

Stellar BHs can form directly from the collapse of massive stars (with initial mass $M_* \gtrsim 20 M_\odot$). The theory of stellar evolution predicts a gap in BH masses between $\sim 60 M_\odot$ and $120 M_\odot$, which is referred to as the upper mass gap or pair-instability mass gap. The reason for this gap is that very massive stars (with helium core masses between $64 M_\odot$ and $133 M_\odot$ [37]) experience an instability in their core due to the creation of electron-positron pairs out of two photons. This leads to a runaway collapse which is predicted to produce a very powerful explosion, called pair-instability supernova, that disrupts the entire star and leaves no compact remnant.
The lower limit of the mass gap has large uncertainty due to our poor understanding of the physics of massive stars and, in particular, of the ${}^{12}{\rm C}\,(\alpha,\gamma)\,{}^{16}{\rm O}$ reaction rate [18, 23].

1.2 Gravitational Waves

According to the theory of General Relativity, mass is the source of space-time curvature. A massive object in an accelerated motion produces ripples in spacetime, called gravitational waves (GWs), which propagate away at the speed of light [35].
In this section we will discuss the emission of gravitational waves by binary systems, their detection and some astrophysical implications.

1.2.1 Emission from compact objects binaries

We consider a binary system made of two compact objects, such as neutron stars or black holes, of masses m_1 and m_2 and we consider them as point-like masses. Assuming circular orbits, their dynamics is equivalent to that of a body of mass $\mu = m_1 m_2 / (m_1 + m_2)$ on a circular orbit of radius R and orbital frequency $\omega_s^2 = G\,(m_1 + m_2)\,/R^3$.
It is useful to define the chirp mass M_c as

$$M_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \tag{1.1}$$

It can be shown [52] that the GW strain[1] at lowest order (quadrupole) generated from such a binary, for the two polarizations + and $\times$, is

$$h_+\,(t) = \frac{4}{r}\left(\frac{G M_c}{c^2}\right)^{5/3}\left(\frac{\omega_s}{c}\right)^{2/3}\frac{1+\cos^2\theta}{2}\cos\left(2\omega_s t_{\rm ret} + 2\phi\right) \tag{1.2}$$

$$h_\times\,(t) = \frac{4}{r}\left(\frac{G M_c}{c^2}\right)^{5/3}\left(\frac{\omega_s}{c}\right)^{2/3}\cos\theta\,\sin\left(2\omega_s t_{\rm ret} + 2\phi\right) \tag{1.3}$$

[1]The strain is $h = \delta L/L$, where L is the distance between two reference points in space and δL is the induced displacement between them. [3]

where r is the distance from the observer, $t_{\rm ret} = t - r/c$ is the retarded time, and θ and ϕ are the zenith and azimuth angle at which the observer is viewing the source. We notice that the GW frequency is equal to twice the orbital frequency of the source: $\omega_{\rm gw} = 2\,\omega_s$.

GWs carry energy away from the source. The total radiated power is approximately [52]

$$P_{\rm gw} = \frac{32}{5}\frac{c^5}{G}\left(\frac{GM_c\omega_{\rm gw}}{2c^3}\right)^{10/3} \tag{1.4}$$

We assume that the point masses have no internal structure, so the only possible source of energy is the orbital energy of the binary $E_{\rm orbit}$. Even in a realistic system of two extended objects, corrections due to their internal structure are of order $(v/c)^{10}$ and can safely be neglected for non-relativistic motion [52].
The orbital energy and its time evolution are

$$E_{\rm orbit} = -\frac{Gm_1m_2}{2R}\,, \qquad\qquad \frac{dE_{\rm orbit}}{dt} = -P_{\rm gw} \tag{1.5}$$

So, if $E_{\rm orbit}$ decreases to compensate for the loss of energy to GWs, R must decrease as well. If R decreases, ω_s increases and, in turn, the GW strain $h_{+,\times}$ increases. This leads to a runaway process where the binary keeps shrinking in time and the amplitude of GWs keeps increasing until, eventually, the binary reaches coalescence.

As the binary loses energy and shrinks, the strain frequency $f_{\rm gw}$ and amplitude A increase. Defining the time to coalescence as $\tau = t_{\rm coales} - t$, their evolution is [52]

$$f_{\rm gw} \equiv \frac{\omega_{\rm gw}}{2\pi} = \frac{1}{\pi}\left(\frac{5}{256}\frac{1}{\tau}\right)^{3/8}\left(\frac{GM_c}{c^3}\right)^{-5/8} \tag{1.6}$$

$$A \equiv \frac{4}{r}\left(\frac{GM_c}{c^2}\right)^{5/3}\left(\frac{\pi f_{\rm gw}}{c}\right)^{2/3} = \frac{1}{r}\left(\frac{GM_c}{c^2}\right)^{5/4}\left(\frac{5}{c\,\tau}\right)^{1/4} \tag{1.7}$$

Both quantities increase continuously as $\tau \to 0$ and diverge at $\tau = 0$. This type of signal is referred to as 'chirping waveform' or simply 'chirp'.
Substituting numerical values in eq. 1.6, we can write

$$f_{\rm gw} \simeq 134\,{\rm Hz}\left(\frac{1.21M_\odot}{M_c}\right)^{5/8}\left(\frac{1\,{\rm s}}{\tau}\right)^{3/8} \tag{1.8}$$

The evolution of R, using the definition of ω_s, the relation $\omega_{\rm gw} = 2\,\omega_s$ and eq. 1.6, is

$$\frac{\dot R}{R} = -\frac{2}{3}\frac{\dot f_{\rm gw}}{f_{\rm gw}} = -\frac{1}{4\tau} \qquad \Longrightarrow \qquad R\,(\tau) = R_0\left(\frac{\tau}{\tau_0}\right)^{1/4} \tag{1.9}$$

where $R_0 = R\,(\tau_0)$ and, consistently with the definition of τ, $R\,(\tau = 0) = 0$.

In truth, the expressions for the GW strain in eq.s 1.2 and 1.3 and for the radiated power in eq. 1.4 are only valid in the approximation of quasi-circular orbits ($\dot\omega_s \ll \omega_s^2$). When a binary shrinks, it eventually reaches a condition where this approximation is no longer valid. In particular there is a minimum value of the radial distance beyond which stable circular

orbits are no longer allowed. This is called the innermost stable circular orbit (ISCO) and, for non rotating BHs and $m_1 \gg m_2$, it is equal to

$$R_{\rm ISCO} = \frac{6\,G\,(m_1+m_2)}{c^2} \tag{1.10}$$

Therefore, for $R \gtrsim R_{\rm ISCO}$, the binary goes through a succession of quasi-circular orbits with progressively decreasing radius. The frequency and amplitude of the emitted GWs follow eq.s 1.6 and 1.7. Then, when the orbital separation approaches $R_{\rm ISCO}$, there is a plunge phase where the radius decreases quickly and a correct estimate of the evolution of $f_{\rm gw}$ and A requires numerical relativity.

The maximum frequency in the quasi-circular orbit approximation is reached when the two objects are separated by a distance $R_{\rm ISCO}$ and it is equal to [52]

$$f_{\rm max} = 2\,(f_s)_{\rm ISCO} = \left(\frac{\omega_s}{\pi}\right)_{\rm ISCO} = \frac{1}{\pi 6\sqrt{6}}\,\frac{c^3}{G\,(m_1+m_2)} \simeq 4.4\,{\rm kHz}\left(\frac{M_\odot}{m_1+m_2}\right) \tag{1.11}$$

In Figure 1.1 we show the overall shape of the waveform of a GW signal, constructed using numerical relativity. There are three well-defined phases: inspiral, merger and ringdown. During the inspiral the binary has a large separation, so the approximation of quasi-circular orbit is valid and the evolution of the signal is given by eq.s 1.6 and 1.7. As the two objects get closer to each other, they eventually come into contact, lose their individual identities and merge to form a final black hole. Both the frequency and amplitude of the signal are maximum at merger.[2] The remnant object is somewhat distorted and, as it settles into its fundamental state, it emits GWs with almost constant frequency and rapidly-decreasing amplitude. This is the ringdown phase.

1.2.2 Detection: the LIGO–Virgo–KAGRA Collaboration

Recent observations of gravitational waves have provided invaluable information on the population of BHs in the Universe. This has been possible thanks to the construction of four GW detectors: two LIGO detectors in the USA, the Virgo detector in Italy and the KAGRA detector in Japan. They are advanced modified Michelson interferometers with arms of length of 3 to 4 km. They are sensitive in the frequency range from $\sim 10\,$Hz up to a few kHz, as visible in Figure 1.2. We will refer to the LIGO–Virgo–KAGRA Collaboration as LVK.

The LVK interferometers are incredibly sensitive to spatial distortions: typical strains from astrophysical sources are on the order of $\delta L/L = 10^{-21}$ or less [3]. Thus, with arms with length of the order of a few kilometers, the interferometers must be sensitive to displacements δL of less than $\sim 10^{-18}\,$m. This is an incredibly small displacement: for comparative purposes, note that the radius of a proton is $\sim 8.4 \times 10^{-16}\,$m [51]. This high sensitivity is achieved by using Fabry-Perot resonant cavities, so that the optical length of the arms is of a few hundred kilometers.

We have seen in the previous section how the frequency of the GW waveform $f_{\rm gw}$ increases, reaches a maximum frequency $f_{\rm max}$ at merger and remains approximately constant during the ringdown.

[2] Although the frequency keeps increasing after the binary reaches the ISCO separation, in the following we will assume that $f_{\rm merger} \simeq f_{\rm max}$ as defined in eq. 1.11.

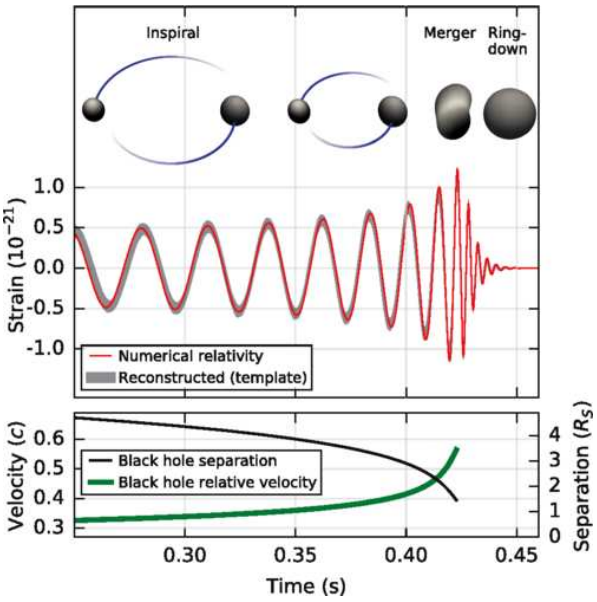


Figure 1.1: Upper panel: Estimated theoretical strain amplitude, given by numerical relativity, for the event GW150914. Lower panel: estimate of separation and relative velocity of the binary members for the event GW150914. Credits to [48].

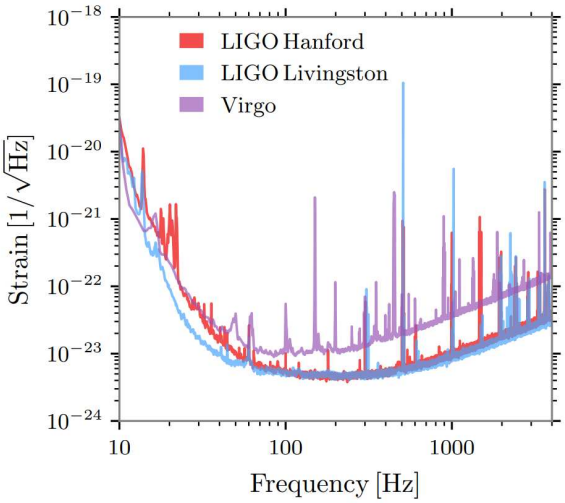


Figure 1.2: The sensitivity of the LIGO–Virgo network for the O3 observing run as characterized by the amplitude spectral density (i.e., the strain per unit square root of frequency). The narrow spectral features are due to mechanical resonances (such as vibrational modes of the suspension fibers), calibration lines and 50 Hz and 60 Hz electric power harmonics. At high frequency, photon shot noise is the main limit to the sensitivity. At low frequency, there are many sources of noise such as seismic and thermal noise, radiation pressure on mirrors, variation in the atmospheric pressure and scattered light [13]. Figure from [49].

Both the behavior of $f_{\rm gw}$ as a function of time (eq. 1.8) and of the maximum frequency $f_{\rm max}$ at the end of the inspiral (eq. 1.11) depend on combinations of the two masses m_1 and m_2. In particular, binaries with larger total masses reach lower maximum frequencies and, since they also have larger chirp masses, their signal are visible for a shorter time interval in the interferometers. This does not generally mean that the inspiral of these binaries is short: it can last for a long time, but it emits at frequencies at which the LVK detectors are not sensitive.
For example, the signal of a binary with masses $m_1 = m_2 = 1.4M_\odot$ enters the detector's bandwidth ($f_{\rm gw} \geq 10\,{\rm Hz}$) at $\tau = 17\,{\rm min}$ prior to coalescence and reaches a maximum frequency $f_{\rm max} = 1.6\,{\rm kHz}$. If we increase the masses up to $m_1 = m_2 = 10M_\odot$, the signal is visible for only $\tau = 38\,{\rm s}$ and the maximum frequency is $f_{\rm max} = 440\,{\rm Hz}$, which is well within the detector's bandwidth. For an even more massive binary with $m_1 = m_2 = 100M_\odot$, the signal is visible for only $\tau = 8\,{\rm s}$ and it reaches a maximum frequency of $f_{\rm max} = 22\,{\rm Hz}$, which is in the higher-noise region of the bandwidth (see Figure 1.2).
Since it is not possible to identify signals that are too short or that have a low signal-to-noise ratio, the LVK interferometers are not sensitive to mergers of high-mass compact binaries.

The most used technique to identify astrophysical GW signals is that of 'matched filtering', which uses a bank of strain templates computed using numerical relativity and searches for similar patterns in the data [49]. The ability to recover the signal of a merger event depends strongly on whether the bank used contains a template with the right parameters.
For example, for a given chirp mass, the banks contain a larger number of templates with mass ratio $q = m_1/m_2 \sim 1$ rather than with low q [2, 49]. Therefore the LVK events identified with matched filtering are biased towards higher q.

The templates are well characterized by combinations of the binary component parameters such as the chirp mass M_c (eq. 1.1), the mass ratio $q = m_2/m_1$ and two combinations of the BH spins called the effective spin $\chi_{\rm eff}$ and the precession spin $\chi_{\rm p}$. These quantities determine the phase evolution during inspiral [47]. Also, the amplitude of the strain is inversely proportional to the distance as in eq. 1.7.

The expression for the effective spin is

$$\chi_{\rm eff} = \frac{m_1\vec{\chi}_1 + m_2\vec{\chi}_2}{m_1 + m_2} \cdot \frac{\vec{L}}{|\vec{L}|} \tag{1.12}$$

where $\vec{L}$ is the orbital angular momentum and $\vec{\chi}_i$ ($i = \{1, 2\}$) are the dimensionless spins, defined as $\vec{\chi}_i = c\vec{S}_i/Gm_i^2$ where $\vec{S}_i$ is the spin angular momentum. For BHs, χ_i can theoretically range from 0 (non-spinning) to 1 (maximally-spinning).

The expression for the precession spin is

$$\chi_{\rm p} = \max\left\{\chi_{1,\perp},\, \frac{q\,(4q+3)}{4+3q}\chi_{2,\perp}\right\} \tag{1.13}$$

where $\chi_{i,\perp}$ is the component of spin perpendicular to the direction of the orbital angular momentum $\vec{L}$ ($i = \{1, 2\}$).

Hence $\chi_{\rm eff}$ is a measure of the BH spin components along the orbital angular momentum vector, while $\chi_{\rm p}$ measures the spin components in the orbital plane. The precession spin

owes its name to the fact that the presence of spin components perpendicular to the orbital angular momentum vector causes a precession of the binary orbital plane.

From the beginning of observations in September 2015 up until the end of the third observation run (O3b) in March 2020, the LVK collaboration has detected 76 transient signals associated with the inspiral and merger of compact binaries with false alarm rate $\leq 1\,\mathrm{yr}^{-1}$ [50]. Of these, 13 merger events have at least one of the progenitors with mass overlapping with the pair-instability mass gap at 90% credible interval. We list the main properties of the upper-mass-gap binaries in table 1.1.

Name	p_{astro}	$m_1/M_\odot$	$m_2/M_\odot$	$m_{\mathrm{rem}}/M_\odot$	χ_{eff}
GW170729	0.98	$50.2^{+16.2}_{-10.2}$	$34.0^{+9.1}_{-10.1}$	$79.5^{+14.7}_{-10.2}$	$0.37^{+0.21}_{-0.25}$
GW190413_134308	0.99	$47.5^{+13.5}_{-10.7}$	$31.8^{+11.7}_{-10.8}$	$78.0^{+16.1}_{-11.5}$	$-0.33^{+0.25}_{-0.29}$
GW190519_153544	> 0.99	$66.0^{+10.7}_{-12.0}$	$40.5^{+11.0}_{-11.1}$	$\mathbf{100.0}^{+13.0}_{-12.9}$	$0.31^{+0.20}_{-0.22}$
GW190521_030229	> 0.99	$95.3^{+28.7}_{-18.9}$	$69.0^{+22.7}_{-23.1}$	$\mathbf{147.4}^{+40.0}_{-16.0}$	$0.03^{+0.32}_{-0.39}$
GW190602_175927	> 0.99	$69.1^{+15.7}_{-13.0}$	$47.8^{+14.3}_{-17.4}$	$\mathbf{110.5}^{+17.9}_{-13.9}$	$0.07^{+0.25}_{-0.24}$
GW190620_030421	0.99	$57.1^{+16.0}_{-12.7}$	$35.5^{+12.2}_{-12.3}$	$88.0^{+17.2}_{-12.4}$	$0.33^{+0.22}_{-0.25}$
GW190701_203306	0.99	$53.9^{+11.8}_{-8.0}$	$40.8^{+8.7}_{-12.0}$	$90.2^{+11.2}_{-8.9}$	$-0.07^{+0.23}_{-0.29}$
GW190929_012149	0.87	$80.8^{+33.0}_{-33.2}$	$24.1^{+19.3}_{-10.6}$	$90.3^{+22.3}_{-14.6}$	$0.01^{+0.34}_{-0.33}$
GW191109_010717	> 0.99	65^{+11}_{-11}	47^{+15}_{-13}	$\mathbf{107}^{+18}_{-15}$	$-0.29^{+0.42}_{-0.31}$
GW191127_050227	0.49	53^{+47}_{-20}	24^{+17}_{-14}	76^{+39}_{-21}	$0.18^{+0.34}_{-0.36}$
GW191230_180458	0.95	$49.4^{+14.0}_{-9.6}$	37^{+11}_{-12}	82^{+17}_{-11}	$-0.05^{+0.26}_{-0.31}$
GW200216_220804	0.77	51^{+22}_{-13}	30^{+14}_{-16}	78^{+19}_{-13}	$0.10^{+0.34}_{-0.36}$

Table 1.1: Selected events from the gravitational wave transient catalog of the LVK collaboration, updated to March 2020. We show events with low FAR ($\leq 1\,\mathrm{yr}^{-1}$) and high primary masses ($\gtrsim 50M_\odot$). The GW transients are named after the date and time (UTC) of observation with the format GWYYMMDD_hhmmss. p_{astro} is the probability of astrophysical origin for the event, m_1 and m_2 are the masses of the binary components, m_{rem} is the mass of the remnant and χ_{eff} is the effective spin. Table from [50].

1.2.3 Astrophysical Implications

With the detection of GWs, it is now easier than it has ever been to identify BHs and study their properties.
The results of the LVK detection runs have not been short of surprises. In particular, the identification of BHs in the pair-instability mass gap (Table 1.1) challenges the models of stellar evolution. There are two possible explanations for the existence of such BHs: either the traditional models of pair-instability supernovae are wrong and it is actually possible to form BHs in the $[\sim 60, 120]\,M_\odot$ mass range from direct stellar collapse [18], or the formation of these BHs requires dynamical formation channels.

In this work, we will explore one of the possible dynamical formation channels for the formation of BHs with masses in the pair-instability mass gap and above.

Also, in a few of the LVK observations, the remnant is an IMBH (highlighted in bold font in Table 1.1). In particular, the most massive and widely-studied IMBH is that produced by the merger event GW190521 (e.g. [1]).
These GW observations provided the first direct detections of IMBHs. Before 2019, several IMBH candidates had been identified as X-ray sources or sources of optical emission lines, as well as via kinematical measurements in massive star clusters, but none of these observations seemed compelling [7]. Ensuring the existence of IMBHs and understanding their formation is of primary importance for astrophysics because they are a link between stellar-mass black holes and SMBHs, so the detection of merger events leading to their formation is an important milestone.

The LVK observations do not rule out the existence of binaries in which one of the members is an IMBH: currently, LVK observations pose an upper limit of $0.056\,\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ for the detection of the merger of equal-mass binary systems with total mass $200\,M_{\odot}$ [2]. The construction of larger terrestrial interferometers such as the Einstein telescope and of space-based interferometers such as LISA will allow for the detection of higher-mass black hole binaries [3].

1.3 Formation scenarios for binary black holes

Binary black holes (BBH) can form in one of two ways: they can either be original binaries or dynamically assembled binaries.

An original BBH comes from the evolution of a binary of two massive stars, where the stars undergo core collapse and create BH remnants with the condition that the binary is not disrupted by the supernova explosions. Instead, a dynamically assembled BBH forms via encounters of BHs in dense stellar environments, such as stellar clusters.
In a dense stellar environment, a BH experiences a large number of close-by interactions with other objects, leading to a number of dynamical processes.

First of all, a BBH can form by direct encounters of three single bodies. Indeed, during the chaotic and resonant evolution of a triple system there can be very close passages between pairs of objects. A close interaction between two objects has two main effects: the system loses energy through gravitational waves production, and tidal interactions become important [75]. Both of these processes can favor the formation of binaries.

Also, a formed binary will have frequent interactions with single bodies. These interactions can have many outcomes depending on the masses and energies of the objects at play:

i) The binary can be widened or destroyed (also said 'ionized') by the three-body interaction. The three objects can therefore leave the interaction as single objects.

ii) The binary can become even more bound (we say that it is 'hardened') by transferring some of its energy to the third body.

iii) One of the members of the binary can be replaced by the intruder. For example, if a binary composed of a BH and a low-mass star undergoes an encounter with a single BH, this can lead to the formation of a BBH [90].

Repeated interactions of a strongly-bound binary with tertiary objects causes multiple occurrences of process (ii), until the binary becomes so tight that its evolution is governed by gravitational waves emission and it is led to merger.

If the environment is dense enough, the merger product can again go through binary formation, hardening and coalescence. This can lead to a chain of events where we dynamically form more and more massive BBHs which produce more and more massive remnants. This process is called hierarchical merging.

The densest stellar environments in the universe are star clusters. There are different kinds of clusters which are classified based on their mass and age.
Globular clusters (GC) are the oldest clusters, with typical ages of 12 Gyr and typical masses of $10^4 - 10^6 M_\odot$. They are quite compact and they typically found in galactic haloes. Young clusters (YSC) have formed recently (they have ages of about 100 Myr) and they are gravitationally loose. They have typical masses of $10^2 - 10^5 M_\odot$. Open clusters (OSC) are similar to YSC but they are smaller and older: they have typical masses of $10^1 - 10^4 M_\odot$ and typical ages of a few Gyr. Nuclear clusters (NSC) are found at the center of most galaxies. They have typical masses $10^4 - 10^8 M_\odot$ and typical ages of a few Gyr. As we will discuss in subsection 2.2.1, they can sometimes coexist with a SMBH.

In this work we will focus on a particular class of NSCs: those that coexist with a SMBH which is in a phase of mass accretion, which is called active galactic nucleus (AGN).

1.4 Aim and Structure of this work

The aim of this work is to study the possibility for the dynamical formation of BBH in active galactic nuclei (AGN). In particular, we want to analyze the properties of the population of compact remnants and study their dependency on the physical features of the AGN.

In chapter 1, we introduce the physical process of BBH inspirals, explain why they produce GWs and describe their detections.

In chapter 2 we define the physical processes at play. We start with a historical introduction on AGN observations and we proceed by explaining the physical properties of AGN components: a SMBH, a gaseous accretion disk and a NSC. Then we describe the interaction of NSC objects with the gaseous accretion disk and how this can lead to the dynamical formation and merger of BBHs.

In chapter 3 we describe our numerical semi-analytical model for BHs evolution in AGNs.

In chapter 4 we show the results of our simulations and discuss the properties of the remnant population.

In chapter 5 we outline our conclusions and discuss some possibilities for future research.

Chapter 2

Dynamical formation of BBHs in AGN Disks

2.1 The zoology of Active Galactic Nuclei

An active galactic nucleus (AGN) is observationally identified as a source in the sky with very bright emission. AGNs cannot be stellar sources because of the peculiar properties of their emission spectra. Normal stars emit black body radiation in a relatively narrow wavelength range (between the near-IR and the near-UV) and their spectra display absorption lines due to the colder outer layers in their atmosphere. Instead, AGNs are powerful emitters of non-thermal radiation covering the entire electromagnetic spectrum from the radio to the γ rays. Also, AGN spectra often have strong emission lines.

AGNs display large variability in their observational properties and in the past they were often categorized as different objects. In this section, we will explain their historic nomenclature and the currently accepted unification scheme that allows us to explain how the same phenomenon can produce different-looking sources.

Seyfert galaxies This class of objects is named after their discoverer Carl Seyfert who noticed that some spiral galaxies had unusually bright nuclei [78]. The spectra of these nuclei show non-thermal continuum emission and contain strong emission lines of high excitation from atoms such as hydrogen, carbon, oxygen, neon and magnesium.
Their are subdivided into two categories based on the width of their emission lines: those displaying broad lines are named Seyfert 1 galaxies, while those that only have narrow lines are called Seyfert 2.
Seyfert 1 galaxies usually have hydrogen emission lines with broadness compatible with gas speeds of $1000 - 5000\,\mathrm{km/s}$, but they often have some much narrower lines coming from forbidden transitions[1], such as [OIII], with width compatible with velocities of a few $100\,\mathrm{km/s}$. Therefore, these two types of emission lines are probably produced in different regions.

[1] Permitted transitions are those allowed by the electric-dipole selection rules, while forbidden transitions are those with zero dipole component but non-zero high-order components. A forbidden transition can happen with much lower probability than a permitted one.

The width of all lines of Seyfert 2 galaxies is compatible with velocities of a few 100 km/s, so both permitted and forbidden lines emanate from the same region. [65]

Whether a region can emit a forbidden line with a significant strength depends mostly on its density. In order to explain this, let us consider an atom with two energetic levels 1 and 2 with $E_1 < E_2$. The equation of balance is [65]

$$n_e n_1 P_{12} = n_e n_2 P_{21} + n_2 A_{21}$$

where n_1 and n_2 are the population densities of states 1 and 2, n_e is the electron density, A_{21} the Einstein coefficient of spontaneous emission, P_{12} the probability for the ion to undergo a transition from state 1 to state 2 in a unit time due to collisions with electrons, and P_{21} the corresponding probability from state 2 to state 1. For forbidden lines, $A_{21} \simeq 0$ ($= 0$ in the dipole approximation).

The emission is mainly produced via spontaneous emission after collisional excitation, so the line luminosity is

$$\mathcal{L} \simeq h\nu_{12} n_2 A_{21} = h\nu_{12} \frac{n_e\, n_1\, P_{12}}{1 + n_e\, P_{21}/A_{21}} = h\,\nu_{12} \begin{cases} n_e\, n_1\, P_{12} & \text{if } A_{21} \gg n_e\, P_{21} \\ n_1\, (P_{12}/P_{21})\, A_{21} & \text{if } A_{21} \ll n_e P_{21}, \end{cases}$$

where h is the Planck constant and ν_{12} is the frequency of a photon absorbed (emitted) by excitation (de-excitation).

We see that in low-density regions ($n_e \ll A_{21}/P_{21}$) the line luminosity $\mathcal{L}$ is independent of A_{21}, so both permitted and forbidden lines can be produced with significant strengths. The opposite is true in high-density regions, where atoms are often subject to de-excitation due to collisions with electrons. Since forbidden lines have a low transition probability, an atom must be in the excited state for a long time before it has a chance to make the transition and this is not possible in a high-density region.

Hence, we can conclude that Seyfert 1 galaxies seem to have high-density regions (no forbidden lines) moving at high speed as well as low-density regions moving at low speed, while Seyfert 2 galaxies only seem to have low-density regions moving at low speed. Of course, this does not rule out the possibility that Seyfert 2 galaxies have high-density regions: all we know it that these regions, if they exist, are blocked from our view.

Radio galaxies Some galaxies have strong emission at radio wavelengths. They were first detected by radio surveys of the sky and, by comparison with optical surveys, almost all of them have been unambiguously identified as elliptical galaxies.

High resolution radio observations of these galaxies usually reveal that the radio emission comes from two lobe structures that extend up to $10^2 - 10^3$ kpc from the central nucleus, where the rest of the galaxy is [65].

Often there are some jet-like structures that stretch from the compact core towards the lobes, suggesting that these jets are responsible for transporting energy from the core out into the radio lobes. The jets are rarely symmetric: often only one of them is observed and, in sources with two jets, one is usually significantly brighter than the other.

Radio emission from jets and lobes of radio galaxies seems to originate from the synchrotron process.

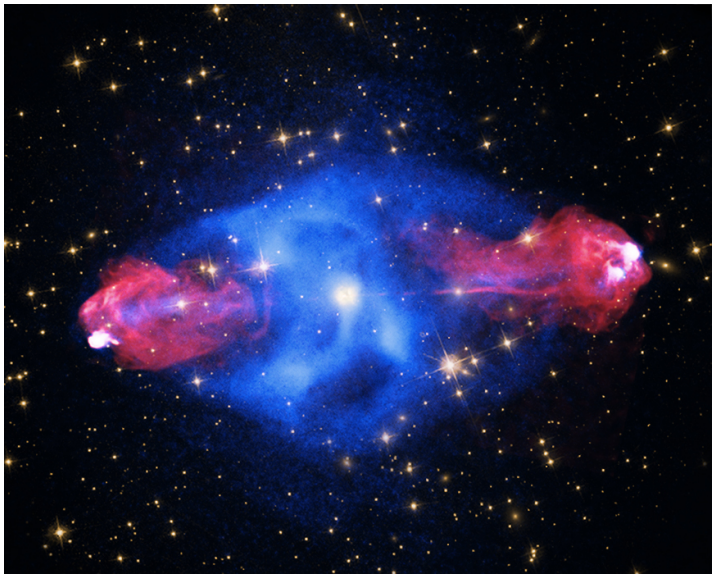


Figure 2.1: Multi-wavelength image of the Cygnus A galaxy. X-ray (in blue) data is from the Chandra Observatory, radio data (in red) is from the Very Large Telescope and optical data (yellow hues) is from the Hubble Space Telescope.

One of the first radio galaxies to be discovered was Cygnus A. In Figure 2.1 there is a multi-wavelength picture of the galaxy, where the radio emission is represented in red and the two-lobe structure is clearly visible.

Seyfert and radio galaxies seem to be drastically different objects. Their morphology is different, as Seyfert galaxies are usually spirals while radio galaxies are usually ellipticals, they emit at different wavelengths and from different spatial regions: most of the emission of Seyfert galaxies comes from the core while most of the radio emission in radio galaxies comes from spatially-extended lobes.
Actually, we will see later that the engine powering up Seyfert galaxies is the same that powers up radio galaxies. The main difference is that radio galaxies display jets that transport energy out of the nucleus of the galaxy. They are presumably made of plasma flowing out at very high speed and they are eventually stopped by collisions with the intergalactic medium, forming wide lobes.

Quasars Quasars ("quasi-stellar radio sources") are very luminous optical sources. They are typically unresolved and their spectra display broad emission lines.
The term quasar is nowadays used interchangeably with the term QSO ("quasi-stellar object") but they have a historical difference: the term "quasar" used to be utilized for sources that emit both in the optical and in the radio, while the term "QSO" was reserved for sources with bright optical emission but invisible in the radio. Because of the similarities in their optical spectra, people have started referring to QSOs as "radio-quiet quasars" and vice versa referring to quasars as "radio-loud QSOs", so the two terms have become synonyms.

Quasars spectra are very similar to those of Seyfert 1 galaxies, the only difference between

these two classes of objects is their luminosity: quasars can be as luminous as 100 ordinary galaxies and they often out-shine their host galaxy [15].
The similarity in spectra suggests that Seyfert galaxies and quasars may be similar kinds of galaxies, where the Seyfert galaxies are the milder form and the quasars are the more extreme and rarer form. Because of their high luminosity, quasars are detected at very large distances where Seyfert galaxies would not be observable.

Observations of quasars with high spectral resolution have identified the presence of fast moving radio-emitting blobs. They sometimes seem to have superluminal motion (i.e. their apparent linear velocity is larger than the light speed c).
This apparently unnatural motion can be explained as the motion of relativistic blobs at a very small angle with the respect to the line of sight [15]. Referring to Figure 2.2, let us consider a source of radiation moving from point A to point B at velocity v in a time δt. The signal emitted at A reached the observer at O at a distance D in a time $t_{AO} = D/c$, while the signal emitted at B, assuming that the angle θ is very small, reaches him at a time $t_{BO} \simeq \delta t + (D - v\delta t \cos\theta)/c$. Therefore, the apparent velocity of the source perpendicular to the line of sight is

$$v_\perp = \frac{v \sin\theta \delta t}{t_{BO} - t_{AO}} \simeq \frac{v \sin\theta}{1 - \frac{v}{c}\cos\theta}$$

If $\theta \ll 1$ and $v \lesssim c$, the apparent velocity $v_\perp$ can be greater than c.

Identifying the morphology of the host galaxy of a quasar is challenging as its nucleus typically overshines the whole galaxy. Nevertheless, the Hubble Space Telescope has observed several hosts of low-redshift quasars. These galaxies seem to have a large diversity in morphology: some are elliptical, some are spiral and some seem to be disturbed or interacting systems [65].

Blazars OVV ("optically violently variables"), also called blazars, are a subclass of quasars. As the name suggests, they are characterized by strong and rapid optical variability, as their optical flux can vary by a significant fraction in less than one day. They are also characterized by a relatively strong polarization of their light, especially in the optical, typically at the level of a few percent (compared to $\lesssim 1\%$ for regular quasars) [65].

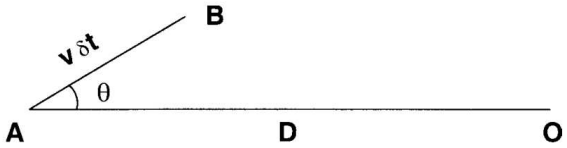


Figure 2.2: A sketch illustrating how superluminal motion arises. Credits to [15].

In Table 2.1 we list a rough estimate of the number densities of the objects we are dealing with. We see that quasars are the rarest of these objects.

The unification scheme It is nowadays believed that all the sources listed above are generated by the same kind of object: a super-massive black hole (SMBH) that is accreting mass via an accretion disk.

Type of object	Number density $\left[\mathrm{Mpc}^{-3}\right]$
Field galaxy	10^{-1}
Seyfert galaxies	10^{-4}
Radio galaxies	10^{-6}
Radio-quiet quasars	10^{-7}
Radio-loud quasars	10^{-9}

Table 2.1: Local number densities of different type of galaxies [65].

We now recognize many similarities between different kinds of active galaxies, which were thought to have nothing in common when they were first discovered.
For example, we know that Seyfert 1 galaxies and quasars are similar, the only difference being that quasars have much more powerful central engines presumably due to a more massive SMBH. Also radio galaxies and radio-loud quasars seem to be the same kind of objects seen at different angles: in quasars the jets are moving directly towards the observer, while in radio galaxies their are moving at an angle and we can see their full extension. The two types of Seyferts are merely the same objects seen from different angles: we can see broad spectral lines (Seyfert 1) if our line of sight is located roughly perpendicular to the accretion disk, so that we can see the central regions where material is moving very fast, whereas we only see narrow lines coming from low-density, slow-moving regions (Seyfert 2) when we are at an angle such that the dusty torus obscures the central part.

The only physical difference seems to be the morphology of the host galaxy. A central engine in a spiral galaxy produces a Seyfert or a radio-quiet quasar depending on whether the central engine is weak or strong, while a central engine in an elliptical galaxy is seen as a radio-loud quasar if the viewing angle is close to the radio jet and is seen as a radio galaxy if the viewing angle is larger [15, 65].

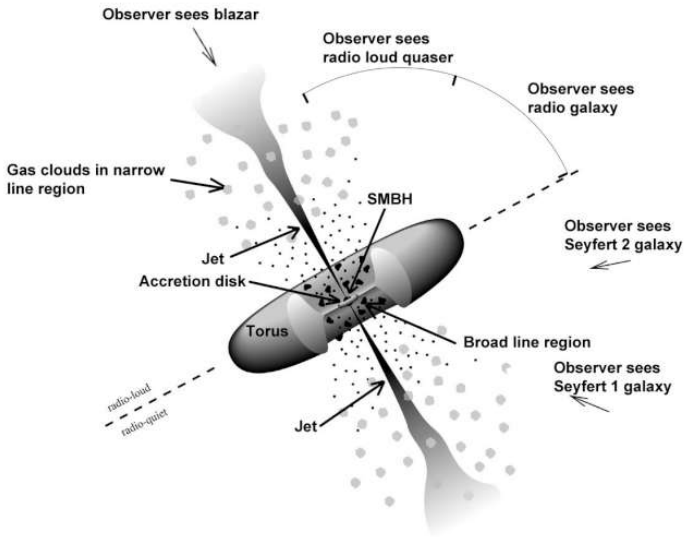


Figure 2.3: Unification scheme for AGNs. Figure credit to Fermi Gamma-ray Space Telescope.

2.2 The AGN Environment

An AGN is an environment where a SMBH is surrounded by a gaseous accretion disk. Sometimes, AGNs also host a nuclear star cluster (NSC). We only consider SMBHs which are compatible with the co-existence with NSCs.
The population of compact objects in the inner part of a NSC interact with the disk's gas due to viscosity effects. We focus in particular on the black hole (BH) component of the NSC population: our goal is to study the pair-up and subsequent merger of binary BHs (BBH) governed by the interaction with gas.

2.2.1 Super Massive Black Holes

SMBHs are, as the name suggests, BHs with very large mass $M_{\rm SMBH} \gtrsim 10^5 M_\odot$, which are expected to be found at the center of massive galaxies.

By definition, SMBHs are not visible because they do not emit light. So, historically, there have been large efforts to confirm their existence and constrain their properties. We will now briefly list of some of the evidence for their existence.

a) Quasars The first astronomical evidence of the existence of SMBHs, starting from the 1960s, was the detection of quasars ("quasi-stellar radio sources") [25].

As anticipated in section 2.1, quasars are objects with incredibly luminous emission up to several 10^{14} times the luminosity of the Sun, which outshine the total emission of their host galaxy. They show large changes in luminosity on timescales as short as 1 hour, suggesting that their size is smaller[2] than one light-hour (equivalent to a few 10^{-5}pc) [65].
Such a bright emission from such a compact area is most likely explained by accretion of matter onto massive BHs.

Indeed, an infalling particle onto a BH can in principle convert between 6% and 42% of its rest-mass energy into radiation, where the lower-limit of 6% is for a non-rotating BH and the upper-limit of 42% is for a maximally-rotating BH [35]. This is much more efficient than the energy conversion that happens in main sequence stars from the nuclear burning of hydrogen into helium, where only about 1% of the hydrogen rest-mass energy is converted into radiation.

Quasars have many other features which make it hard to believe that they could be stellar sources, such as their broad emission lines in the X- and γ-rays and highly collimated radio jets.

b) Kinematics In the early 1990s, when the angular resolution of telescopes was finally good enough thanks to adaptive optics, people started observing the motion of stars in the nucleus of the Milky Way.
Sagittarius A* (SgrA*) was already well known as a compact radio source, so these observations were devoted to measuring its mass via the determination of orbits around it. By the early 2000s, research groups had already identified the orbits of about 10 stars,

[2] If the integrated luminosity of an object changes over a timescale Δt, the size of the emitting region must be smaller than $\sim c\,\Delta t$.

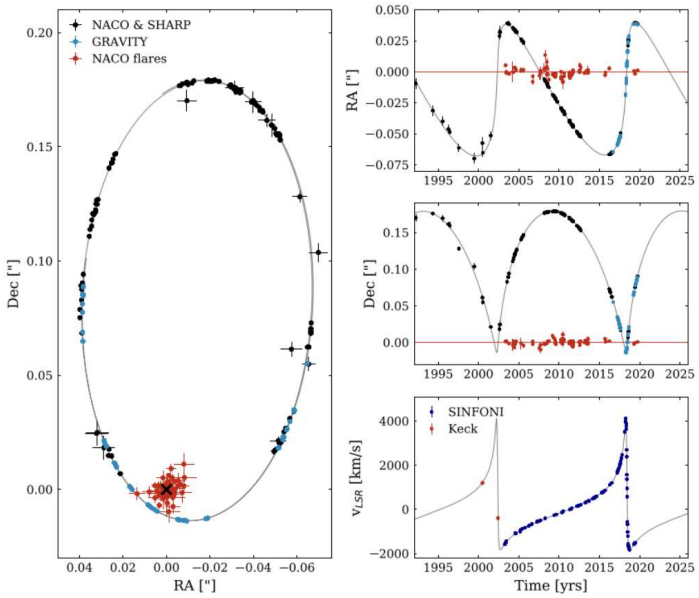


Figure 2.4: Orbit of the star S2/S02 around SgrA*. On the left: astrometric positions. On the right: Measurements of right ascension, declination and line-of-sight velocity. The data (blue and black dots) comes from multiple surveys taken between 1992 and 2019. The gray lines trace the best-fitting general-relativity orbit around a center of mass marked by the black cross at $(0, 0)$ coordinates on the left-hand panel. The red data points mark the positions of the infrared emission from SgrA*. The infrared light from SgrA* is scattered, so a variation in the apparent position of the infrared emission is not indicative of a variation in the position of the source. Figure from [26].

whereas now more that 40 orbits are known in the central light-month ($\sim 3 \times 10^{-2}$ pc) of the galaxy [26].
Each star is observed both with proper motion studies (which determine its motion on the right ascension/declination plane) and Doppler spectroscopy (which determines its velocity in the orthogonal direction), so to establish the three-dimensional structure of its orbit. Most notably, the star S2 (or S02) was observed since 1992 and its motion is well determined. It is on a highly eccentric orbit ($e = 0.88$), with a peri-distance of about 6×10^{-4} pc and an orbital period of 16 years [26]. We show the data and the best-fit orbit in Figure 2.4.

The measured orbits show that the gravitational potential is that of a point mass with mass greater than $10^5 M_\odot$, whose position is that of the radio source SgrA* (within a 4×10^{-5} pc uncertainty) [25]. The size of the radio emission of SgrA* measured at a wavelength of 3.5 mm is about 1 AU ($\sim 5 \times 10^{-6}$ pc) [80].

Such large mass in an object of such a compact size can most reasonably be explained with a massive BH. Nevertheless, the scientific community has not always been unanimous in this interpretation: other options were clusters of smaller objects (neutron stars, stellar BHs and/or brown dwarfs), as well as more exotic objects such balls of heavy fermions held up by degeneracy pressure or boson stars, which are hypothetical stars made of bosons [65]. With the most recent data on the proper motion limit of SgrA* [31, 34], the hypothesis of

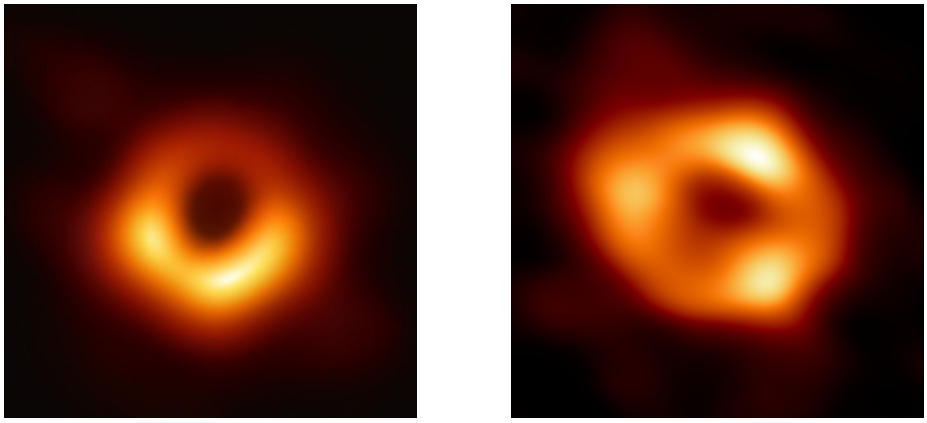


Figure 2.5: Images of M87* (on the left) and SgrA* (on the right) produced by the EHT collaboration.

the dark remnant cluster is rejected beyond any reasonable doubt.

c) Imaging Direct imaging of SMBHs is considered to be a direct proof of the nature of these objects. Imaging a BH is a very compelling task and requires the combination of observations from a worldwide network of radio telescopes.
It was only achieved twice to date: in 2019 for the SMBH M87* at the center of the homonymous galaxy [22] and in 2022 for SgrA* at the center of the Milky Way [17].

We show the images of both M87* and SgrA* in Figure 2.5. They show a bright ring of emission which surrounds a dark inner region called the shadow of the BH. The luminous ring is due to the escape of photons from near the event horizon via an unstable circular orbit. Due to gravitational lensing, the radius of the shadow appears to a far-away observer about 2.5 times larger than the Schwarzschild radius of the BH.

By measuring the size of the BH shadow, we can infer that both M87* and SgrA* are Kerr SMBHs with masses

$$M_{\mathrm{M87}} \simeq 6.5 \times 10^9\,\mathrm{M}_\odot \qquad M_{\mathrm{SgrA}} \simeq 4.0 \times 10^6\,\mathrm{M}_\odot$$

In this work, we are not interested in all SMBHs, but only in those that coexist with NSCs at the center of galaxies.

Coexistence with a NSC Whether or not a galaxy can contain an SMBH seems to depend on its mass: large spheroidal stellar systems, such as elliptical galaxies, invariably host a $10^6 - 10^9 M_\odot$ SMBH, whereas dwarf elliptical galaxies are predominantly dominated by a $10^5 - 10^7 M_\odot$ NSC [33]. One may suppose that galactic evolution can lead to either one or the other scenario, but they are not mutually exclusive. There have been multiple observations of galaxies where an SMBH and an NSC coexist. For example, the Milky Way hosts both an NSC of mass $\sim 3 \times 10^7\,\mathrm{M}_\odot$ [33] and the SMBH SgrA* of mass $\sim 4 \times 10^6\,\mathrm{M}_\odot$. It is interesting to know that the Milky Way's NSC was known since the 1960s, while SgrA* was only discovered in the 1970s [68]. Indeed, identifying the presence of an NSC is

relatively easy because they are luminous sources (see subsection 2.2.3), while spotting an SMBH is much more demanding. Measuring the radiation from accretion of matter onto an SMBH or discerning the kinematics of stars around it are compelling tasks, becoming more and more challenging as the distance increases or the mass of the SMBH decreases.

There is evidence [33] that both SMBHs and NSCs can frequently inhabit galactic spheroids[3] with stellar masses ranging from $10^8 M_\odot$ to $10^{11} M_\odot$. Smaller galaxies typically host NSCs, while larger galaxies typically host SMBHs only.

The relation between the mass of the SMBH and that of a coexisting NSC is [33]

$$\log\left[\frac{M_{\rm SMBH}}{M_{\rm SMBH}+M_{\rm NSC}}\right] = \frac{2}{3}\log\left[\frac{M_{\rm SMBH}}{5\times10^7\,{\rm M}_\odot}\right] \qquad \text{for } M_{\rm SMBH} \leq 5\times10^7\,{\rm M}_\odot \tag{2.1}$$

It is not clear whether the lack of SMBHs in low-mass hosts is an intrinsic property of galaxies or whether it is due to observational limits on low-mass BHs. Therefore, coexistence in low-mass galaxies cannot be ruled out.
Actually, a BH of about $5\times10^4\,{\rm M}_\odot$ was recently identified at the center of the dwarf galaxy RGG 118 [4]. This is technically an IMBH, and its presence is a hint that central massive BHs could be a common feature of lower-mass galaxies as well.

The dearth of NSCs in galaxies with mass greater than $10^{10} M_\odot$, instead, is observationally certain. There are a few possible explanations.
One explanation could be that all standard galaxies start out with NSCs with mass proportional to their spheroidal mass. Then, above some critical mass, dynamical processes become so efficient that they transform a larger and larger fraction of the NSC stars into massive seed central BHs, until there is no NSC star left. It has been shown [83] that this process is possible via runaway tidal capture and tidal disruption events.
Another explanation is that, when the host galaxy is more massive than a few $10^{10} M_\odot$, the accretion of gas onto the central SMBH is so fast that star formation in the NSC might not happen at all [68].

Although the formation and evolution of SMBHs and NSCs must be inevitably linked, they do not show the same scaling relations with their host galaxy. For example, the $M_{\rm NSC}-\sigma$ relation is much shallower than the $M_{\rm SMBH}-\sigma$ one [63, 76]:

$$M_{\rm NSC} \propto \sigma^{\,2.11(\pm0.31)}\,, \qquad M_{\rm SMBH} \propto \sigma^{\,4.73(\pm0.36)} \tag{2.2}$$

where σ is the three-dimensional velocity dispersion of stars in the host spheroid.
Therefore SMBHs and NSCs do not form a single family of central massive objects.

2.2.2 Active Galactic Nuclei Disks

As seen in section 2.1, there is large observational evidence for the existence of AGNs. Nowadays, we also have high-resolution images of some active galaxies that confirm that accretion onto SMBHs happens through accretion disks.

[3] By host galactic spheroids (later called simply host galaxies for shortness) we mean either an elliptical galaxy or the bulge of spiral and lenticular galaxies.

For example, observations of the nucleus of NGC 4258 [66] show a thin, slightly warped disk of radius 0.25 pc (viewed almost edge on) in Keplerian rotation around an unresolved mass of $4 \times 10^7\,\mathrm{M}_\odot$. The disk is clearly visible because it shows water maser emission, probably due to the amplification of a radio-continuum source that shines in the nucleus. It appears to be remarkably thin, with an aspect ratio[4] of roughly 10^{-3} at the location of maser emission. Nevertheless, determining its whole aspect ratio profile or its density profile is unattainable with the current data.

Observations of AGN disks can poorly constrain their physical properties, so we must turn to hydrodynamical models. Unfortunately there can be multiple orders of magnitude variation in density, aspect ratio and lifetimes allowed within common models due to the complexity of modeling magneto-hydrodynamics, turbulence, radiative transport, and plasma physics [57].
In this section, we will illustrate the model we use for the structure of accretion disks. Because of the large uncertainty, we keep the model flexible to allow for significant variations in density, aspect ratio and lifetime.

We assume the AGN disk to be well described by a Shakura-Sunyaev disk (SSD) model [79]. An SSD is geometrically thin,[5] and optically thick, with steady-state accretion onto the central SMBH. Gas turbulence is supposed to be the cause of disk's viscosity; therefore, we can write the the viscosity parameter ν as

$$\nu = \alpha\, c_s\, H, \tag{2.3}$$

where $\alpha \in [0,\, 1]$ is the viscosity coefficient, c_s is the sound speed and H is the disk height. Indeed, in a turbulent medium $\nu = v_{\mathrm{turb}}\, l_{\mathrm{turb}}$, where v_{turb} is the velocity of turbolent cells with respect to the average (Keplerian) gas motion and l_{turb} is the size of the largest turbulent cell. By assuming $v_{\mathrm{turb}} \simeq c_s$ and $l_{\mathrm{turb}} \simeq H$ we find the result of eq. 2.3.

The disk is assumed to be in local thermal equilibrium and to be radiatively efficient, so that it can readily radiate away its viscous heat. Therefore, the disk can stay geometrically thin.

The viscosity coefficient α is a free model parameter. We assume it to be constant over the whole extension of the disk and to have a constant value independently of the other physical properties of the disk. The assumed value is $\alpha = 0.3$ [44].

The problem with SSDs is that they are viscously, thermally, and convectively unstable to perturbations, meaning that they can easily fragment into smaller sub-clouds [45]. Regions of the disk that are stable against fragmentation respect the Toomre's stability criterion which, for Keplerian motion, is:

$$Q = \frac{c_s\, \Omega}{\pi\, G\, \Sigma_g} \gtrsim 1, \tag{2.4}$$

where $\Omega = (G\, M_{\mathrm{SMBH}}/R^3)^{1/2}$ is the Keplerian angular velocity and Σ_g is the surface density of the gas.

Hence, a realistic modeling of an accretion disk requires a few extra assumptions. A possible solution is the model by Sirko & Goodman (hereafter, SG) [81].

[4] The aspect ratio is the ratio between the half-height H of the disk at a certain location and the radial coordinate R.

[5] A geometrically thin disk is such that its aspect ratio $h = H/R$ is always $h \ll 1$

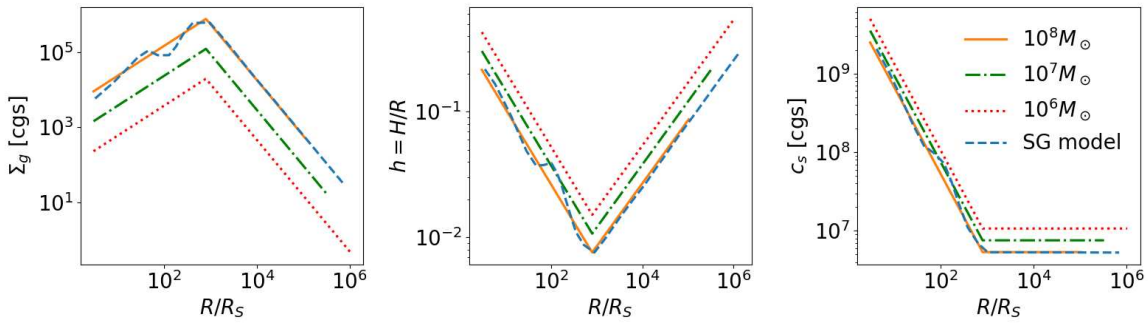


Figure 2.6: Surface density Σ_g, aspect ratio h and sound speed c_s profiles of the SG model for a $10^8 M_\odot$ SMBH (blue dashed) and relative broken power-law best fits (orange solid). The profiles for a $10^7 M_\odot$ and a $10^6 M_\odot$ SMBH are shown in green (dash-dot) and red (dotted).

The SG model fixes the problem of fragmentation by assuming that there is star formation in the outskirts of the disk. Energetic feedback from the newly-formed stars increases the velocity dispersion and sound speed of the gas, maintaining Q close to unity, supporting the disk against global gravitational instability and inhibiting further star formation [7]. As a consequence, the density in the inner part of the disk stays continuous and that region is stable against fragmentation. The model neglects any effects due to magnetic fields and general relativity.
A side consequence of this model is that NSCs can form directly from the outskirts of AGN disks. This can be a possible explanation of the coexistence between SMBHs and NSCs (see subsection 2.2.1).

There are other models for stable accretion disks such as the model by Thompson, Quataert & Murray (hereafter, TQM) [86], who extrapolate a star-forming galaxy disk inward to the SMBH and ensure stability by assuming that the external accretion rate is high enough so that the gas fraction of the disk remains constant, allowing rapid inflow onto the SMBH to continue. This is not an SSD because the source of viscosity are global gravitational instabilities rather than the motion of turbulent cells.
The TQM model produces density and aspect ratio profiles with large discontinuities. Since these discontinuities are placed at radial positions that are important for our subsequent work (see a discussion on migration traps in subsection 3.1.2), they could potentially cause numerical problems. Also, models for the interaction between NSC objects and gas (discussed in sections 2.3 to 2.7) are only reliable for disks with a continuous density and would not be accurate for TQM disks.
Therefore, we prefer the SG model over the TQM one.

We show the surface density, aspect ratio and sound speed profiles of the SG model for a SMBH mass of $10^8 M_\odot$ in Figure 2.6 (blue dashed lines).
We fit the profiles with a broken power law shown in the same figure (orange lines). The best fit power laws are:

$$\Sigma_g\left(R\right) = 7.94\times10^5\,\mathrm{g/cm^2}\begin{cases}\left(\dfrac{R}{10^3\,R_g}\right)^{0.8} & R\leq 10^3\,R_g\\ \left(\dfrac{R}{10^3\,R_g}\right)^{-1.49} & R>10^3\,R_g\end{cases} \tag{2.5}$$

$$h\left(R\right) = 7.59\times10^{-3}\begin{cases}\left(\dfrac{R}{10^3\,R_g}\right)^{-0.6} & R\leq 10^3\,R_g\\ \left(\dfrac{R}{10^3\,R_g}\right)^{0.5} & R>10^3\,R_g\end{cases} \tag{2.6}$$

$$c_s\left(R\right) = 5.37\times10^6\,\mathrm{km/s}\begin{cases}\left(\dfrac{R}{10^3\,R_g}\right)^{-1.1} & R\leq 10^3\,R_g\\ 1 & R>10^3\,R_g\end{cases} \tag{2.7}$$

The SG profiles have only been published for a central SMBH with mass $10^8 M_\odot$. We expect these quantities to have the following power-law dependencies on M_{SMBH} [5]:

$$\Sigma_g \propto \left(M_{\mathrm{SMBH}}\right)^{4/5}\,, \qquad h\propto\left(M_{\mathrm{SMBH}}\right)^{-3/20}\,, \qquad c_s\propto\left(M_{\mathrm{SMBH}}\right)^{3/2} \tag{2.8}$$

We use the SG model for multiple SMBH masses by re-scaling the expressions in eqs. 2.5, 2.6 and 2.7 as determined by the relations of eq. 2.8. We report the full expressions in eqs. 3.11, 3.12 and 3.13.

The lifetime of AGN disks is subject to large uncertainty. Data from quasars observations can poorly constrain the lifetime: different estimates span several order of magnitudes in the range of $10^{-2}-10^3\,\mathrm{Myr}$ [43]. Also, it is not clear whether accretion onto SMBHs happens continuously over a certain time span, or episodically through many cycles of efficient accretion.

One of the most commonly-used probes for quasar lifetimes is the proximity effect: quasars emit a large amount of radiation that effectively ionizes the inter-galactic medium (IGM). The regions around quasars where the ionization level of the IGM is enhanced by quasars radiation, called proximity zones, are observationally identified as regions that emit forests of emission lines. The spatial extent of proximity zones is correlated to the quasar's on-time, i.e. the time during which a quasar has been shining.
The on-time of a quasar is different than its lifetime. Let us suppose that the quasar turns on at a time t_1. If we make an observation at a time t_{obs}, then the on-time is $t_{\mathrm{obs}}-t_1$. The quasar will eventually keep shining for some time after the observation and it will turn off at some time t_2. The lifetime is obviously t_2-t_1 and is generally larger than the on-time.
Therefore, the extent of the proximity zone can only provide direct information on the on-time but not on the lifetime. In order to obtain knowledge about the lifetime we need more complex models.
We use the results from a Bayesian study [43] that performs statistical comparison between the sizes of observed proximity zones and the outcome of the radiative transfer simulations to obtain information about the quasars lifetimes.

They find that the mean quasar lifetime is 1.7 Myr and the 68% credibility interval goes from 0.6 Myr up to 5 Myr.

In our model, we will consider two possible values of lifetimes. In some of our simulations we will assume a lifetime of 1 Myr, consistently with the mean quasar lifetime found from [43]. We also consider the option that quasars may be longer-lived and assume a lifetime of 10 Myr in other simulations. In all simulations, we always consider a single accretion episode.

To allow for the wide variety of parameters suitable for AGN disks and explore their relation to the dynamical assembly of BH binaries, we will examine eight possible scenarios. Given a SMBH mass, we will consider:

- Three possible values of the surface density at the migration trap. We will call the relative models high, medium and low density disk and we will indicate them with a 'D' , a lowercase 'd' and an italic lowercase '*d*' respectively.

- Two possible values of the lifetime. We will call the relative models long-lived and short-lived disk and we will indicate them with an 'L' or an 's' respectively.

- Two possible values of the aspect ratio at the migration trap. We will call the relative models thin and thick[6] disk) and we will indicate thick with a capital 'T' and thin with a lowercase 't'.

Every disk type will be indicated by an acronym. For example, the high-density, long-lived and thin disk will be referred to as 'DLt' whereas the medium-density, short-lived and thick disk as 'dsT'.
See section 4.1 for more details on the chosen parameters.

We assume no evolution for the AGN disk: in our model its physical properties remain unaltered for its whole lifetime.

2.2.3 Nuclear Star Clusters

NSCs are dense and massive clusters that, as their name suggests, are found in the innermost region of most galaxies. They are quite an ubiquitous feature of galaxies: their hosts can exhibit different morphology and a wide range of masses between $10^8 M_\odot$ and $10^{10} M_\odot$ [68]. Observationally, they are easy to identify because they over-shine any other stellar cluster and because of their unique location at the dynamical center of the host galaxy. However, studying their properties can be hard due to patchy dust extinction which affects galactic nuclei.
Even though there are hundreds of studies of NSCs, their formation mechanism is not certain yet. The two main theories are that they formed from the inspiral of star clusters into the center of the galaxy or from in-situ star formation triggered by high gas densities in the galactic nucleus [68].

The parameter governing the size of a NSC is the effective radius (or half-light radius) $r_{\rm eff}$, defined as the radius within which half the cluster light is contained. The typical size of a NSC is $r_{\rm eff} \simeq 3.3\,{\rm pc}$, comparable to the size of most globular clusters.

[6] This nomenclature is just for comparison, as both disk models will be geometrically thin.

NSC masses, on the other hand, are much higher than globular clusters'. The typical mass of a NSC is around $10^6 - 10^7 M_\odot$ and can be as high as $10^9 M_\odot$, while the typical mass of a Milky Way globular cluster is around $10^5 - 10^6 M_\odot$.
Hence, NSCs are the densest known stellar environments, which can reach surface densities of $10^6 \mathrm{M}_\odot/\mathrm{pc}^2$ or more. [68]

We defer a detailed discussion of mass relations to subsection 3.1.3. In particular, the relation between the NSC and the host galaxy mass is in eq. 3.14, while the relation between the NSC mass and its radial dimension is in 3.15.

The shape of a NSC is that of a spheroidal with typical ellipticity lower than 0.2. However, NSC with high mass can sometimes have a very flattened shape with ellipticity up to 0.6 [68]. In the following work, we will neglect the existence of such elongated NSCs and we will assume them to be spherical.

We assume the NSC to have an age such that its stellar population is composed of light-mass stars ($M \lesssim 2\, M_\odot$) that are still in the main sequence phase, burning hydrogen in their cores, and by the remnants of higher-mass stars. The radial distribution of these objects will not be uniform. Because of dynamical mass segregation [62], heavier objects will sink rapidly towards the core.
We trust that the mass fraction of BHs in the central region of a mass-segregated NSC is [5]

$$f_{\mathrm{BH}} \simeq 0.04 \tag{2.9}$$

We do not consider stellar evolution in our setting: the small and mid-sized stars ($M \sim 1\, \mathrm{M}_\odot$) have a typical main sequence phase of 10 Gyr, which is much longer than a typical AGN episode which lasts, at most, a few tens of Myr. So the time span of the simulation can be considered as an instant during the lifetime of these stars. More massive stars have shorter lifetimes, but we consider them to be already extinct and to have already formed their final remnants.
For the same reason, we neglect star formation during AGN episodes.

Finally, we neglect the existence of original binaries in the NSC, assuming all components of the cluster to be single objects.

2.3 Gas Capture

When NSC objects orbit around the central SMBH, their orbits can cross the disk and will gather some of the disk gas, causing them to be subject to strong gas drag. This can have the effect of damping both the inclination i and the eccentricity e of their orbit:

$$e \to 0\,, \qquad i \to 0$$

Therefore, after a sufficient number of laps, these objects will have circular orbits embedded in the disk [59]. This process is called gas capture or orbital damping.

The gas accretion and subsequent gas drag are significant only for prograde orbiters (i.e. objects which orbit in the same direction as the disk). We neglect any variation in the orbits of retrograde orbiters.

If gas drag is sufficiently strong, the eccentricity variation is given by [9, 19]

$$\overline{\frac{de}{dt}} = \begin{cases} -\kappa e & \text{for } e \leq 2\,h \\ -\kappa e^{-2} & \text{for } e > 2\,h \end{cases} \tag{2.10}$$

where the average is over one orbit and h is the disk's aspect ratio.
Similarly, for the inclination [19]: (i is in radians)

$$\overline{\frac{di}{dt}} = \begin{cases} -\kappa i & \text{for } i \leq h \\ -\kappa i^{-2} & \text{for } i > h \end{cases} \tag{2.11}$$

Eccentricity and inclination damping is therefore faster for objects with an initially small e and i (exponential decay) while for larger e and i it has initially a power-law behavior followed by an exponential decay when $e \leq 2\,h$ or $i \leq h$. Overall, it happens over a typical timescale $1/\kappa$ for all initial eccentricities and inclinations.

We define the timescale as $t_{\rm damp} = (\kappa)^{-1}$. For an object of mass m on an initial orbit of semi major axis A around a SMBH of mass $M_{\rm SMBH}$ it is [38, 59]

$$t_{\rm damp} = \frac{M_{\rm SMBH}^2\, h^4}{m\, \Sigma_g\, A^2\, \Omega} \tag{2.12}$$

where h is the aspect ratio of the disk, Σ_g is the surface density of the gas and Ω is the Keplerian angular velocity around the SMBH.
The damping timescale $t_{\rm damp}$ is inversely proportional to the mass of the object m, so more massive objects are damped more efficiently.

In this model we neglect the change in mass due to gas accretion. Therefore, during the orbital evolution, m is a constant quantity.
We also neglect any change in the disk structure due to the presence of embedded objects.

Notation We will often use coordinates such as semi-major axes and radii. We use capital letters A and R to refer to orbits around the central SMBH while we use lower-case letters a and r for orbits inside a binary system.

2.4 Migration

Once a BH (or a star) is embedded in the disk, it exchanges angular momentum with the surrounding gas and is subject to gas torques. Torques can be both positive or negative, leading to outward or inwards migration, respectively.

Similarly to what happens to planet seeds in protoplanetary disks, migration can happen in two different ways called Type I and Type II.

Type I Small to medium-mass objects are subject to Type I migration, meaning that they change their radial position in the disk without significantly perturbing the density distribution of the disk itself.

For an object of mass m orbiting on a circumference at radius R this happens on a timescale [6, 59]

$$t_{\rm migr} \simeq \frac{M_{\rm SMBH}^2\, h^2}{m\, \Sigma_g\, R^2\, \Omega}. \tag{2.13}$$

Differently from eq. 2.12, here we are considering a radius R rather than a semi major axis A because gas capture happens necessarily before migration,[7] so the orbits have already been circularized when migration sets-in.

Similarly to the damping time (eq. 2.12), the migration time is also proportional to the inverse of the mass, so more massive objects migrate faster.

In our disk model (see subsection 2.2.2), torques are positive in the inner region of the disk (where the slope of the surface density is positive) and they are negative in the outer region (where the slope of the surface density is negative) [7]. See Figure 2.6 as a reference.
Therefore, Type I migration is directed outwards in the inner disk and inwards in the outer disk. At the location where the torques change sign, called a migration trap, migration will stall leading to a large accumulation of objects.
Hence, after a timescale $t_{\rm migr}$ (eq. 2.13), the migrating object will be in the migration trap.

As stars migrate towards regions of the disk with higher stellar density, they increase their probability of encounters with other objects and the nature of these encounters can hinder their migration. The fate of a star is different based on the object it interacts with.
For example, it can have a collision with another star and create a more massive one, which has a shorter lifetime and can explode as a supernova (SN). The remnant of an SN explosion receives a strong kick at birth [27] which can push it out of the AGN disk. Alternatively, if it comes close to a compact object, such as a BH, the star can be stripped of its outer layers, losing a significant fraction of its mass, or it can even be swallowed by the BH if it crosses its innermost stable circular orbit (ISCO) radius. If the star survives the encounter, its migration would not be hampered but it would be slowed down due to a decrease in mass.
Finally, if a star interacts with a hard binary it can increase its kinetic energy according to the Heggie's law [36] (see eq. 3.30 for more details), receiving what is called a recoil kick. Therefore, it can be scattered to a different location of the disk.
Because of their higher mass and because their migration is less impeded than what happens to stars, we only consider the migration of BHs in the disk.

Type II Larger objects, on the other hand, can open gaps in the disk. This happens because the motion of a massive object exerts an intense tidal perturbation on the disk, which effectively pushes material away from the orbit's trail [12]. This is called Type II migration.
An object of mass m can open a gap in the disk if [58]

$$q > \sqrt{\frac{\alpha}{0.09}\, h^5} \tag{2.14}$$

where $q = m/M_{\rm SMBH}$ is the mass ratio with respect to the central SMBH, α is the viscosity parameter and $h = h\,(R)$ is the aspect ratio of the disk at radius R.

[7] Migration can only set-in when $i = 0$ and the orbit is embedded in the disk. Because of gas drag, $i \to 0$ and $e \to 0$ on similar timescales. See section 2.3.

Taking as fiducial values $\alpha = 0.3$ and $h = 0.02$, the minimum mass to open a gap in the disk is $m \simeq 10^{-4}\, M_{\rm SMBH}$.
If an object opens a gap in the disk, assuming that no gas is able to cross the gap, its migration follows the viscous evolution of the disk's gas. In the inner disk the gas is accreting onto the SMBH, so the object would follow the inward motion of the gas. In the outer disk, instead, the disk can be expanding and Type II migration would be directed outwards. The timescale for Type II migration is the timescale for the viscous evolution of the disk [59].

$$t_{\rm migr,\,II} = t_{\rm visc} = \frac{1}{\alpha\, h^2\, \Omega} \tag{2.15}$$

However, pressure forces in the disk push to close the gap. So, even if an object is massive enough to open a gap, the latter can stay open against pressure forces only if [12, 58]

$$q \gtrsim \alpha\, (40\, h)^2 \tag{2.16}$$

For the same fiducial values ($\alpha = 0.3$ and $h = 0.02$), this conditions entails $m \gtrsim 0.2\, M_{\rm SMBH}$. An object with such a great mass would generally be a SMBH itself. We will verify a posteriori that with our model we never produce such massive BHs via hierarchical merging, so Type II migration is not possible.

2.5 Pair-up of Binary Black Holes

Depending on the physical characteristics of the AGN (such as viscosity, gas density, aspect ratio and SMBH mass), gas capture and migration can happen on short timescales.
When these processes are efficient, they can lead to a large abundance of BHs in the narrow region of the migration trap. Not only is there a large BH density but also all BHs in the migration traps are on similar orbits (prograde and quasi-Keplerian), so their relative velocities of encounter are small.
Under these conditions, it is easy for two BHs to become gravitationally bound in a binary. Therefore, efficient damping and migration lead to efficient binary pair-up.

In this work, we assume that the pair-up of a primary and a secondary BH is immediate as soon as the primary reaches the migration trap. We show in Appendix A that the densities that can be reached in the migration trap are high enough to allow for this approximation.

The pairing timescale of a BBH is therefore

$$t_{\rm pair} = t_{\rm damp} + t_{\rm migr} + t_{\rm in} \tag{2.17}$$

where $t_{\rm damp}$ is the damping timescale (eq. 2.12), $t_{\rm migr}$ is the migration one (eq. 2.13) and $t_{\rm in}$ is the formation time of the primary BH since the time of formation of the AGN disk. We are neglecting stellar evolution, so for first-generation BHs $t_{\rm in} = 0$.
For BHs of a later generation N, the formation timescale obviously keeps track of the evolution of the previous generations n as

$$t_{\rm in}^{(N)} = \sum_{n=1}^{N-1} \left(t_{\rm damp}^{(n)} + t_{\rm migr}^{(n)} \right) \tag{2.18}$$

The pair-up process can perturb the orbit around the SMBH, but if damping is efficient the orbit circularizes quickly. We always assume the binary to be bound on a circular orbit in the migration trap.

The motion of the BBH around the SMBH is quasi-Keplerian [85].
We assume that the disk's particles have a Keplerian velocity (in its computation we only account for the mass of the SMBH and of the inner NSC objects, while we neglect the mass of the disk)

$$v_{\rm Kepl}\left(R\right) = \sqrt{\frac{G\left(M_{\rm SMBH} + M_{\rm NSC}(R)\right)}{R}} \tag{2.19}$$

Objects (both BHs and BBHs) have a velocity deviation $\delta v_{\rm BH}$ with respect to the local disk which is distributed as a Gaussian with zero mean and standard deviation [85].

$$\beta = 0.2\,\frac{v_{\rm Kepl}\left(R\right)}{\sqrt{3}} \tag{2.20}$$

Therefore, objects in an orbit at radius R have a velocity

$$v\left(R\right) = v_{\rm Kepl}\left(R\right) + \delta v_{\rm BH}\left(R\right) \tag{2.21}$$

For BBHs in the migration trap, this quantity must obviously be evaluated at the migration trap radius $R_{\rm trap}$.

2.6 Evolution of a BBH in the AGN disk

When two BHs pair-up in the migration trap, they form a BBH which is embedded in the disk. During its internal orbit, the binary exchanges angular momentum with the gas and is efficiently hardened.

The semi major axis a and eccentricity e of a binary of masses m_1 and m_2 evolve due to gas hardening according to the following equations [39]:

$$\dot{a}_{\rm gas} = -\frac{24\pi\alpha c_s^2\Sigma_g(1+e)^2 a}{\mu\Omega_b} \tag{2.22}$$

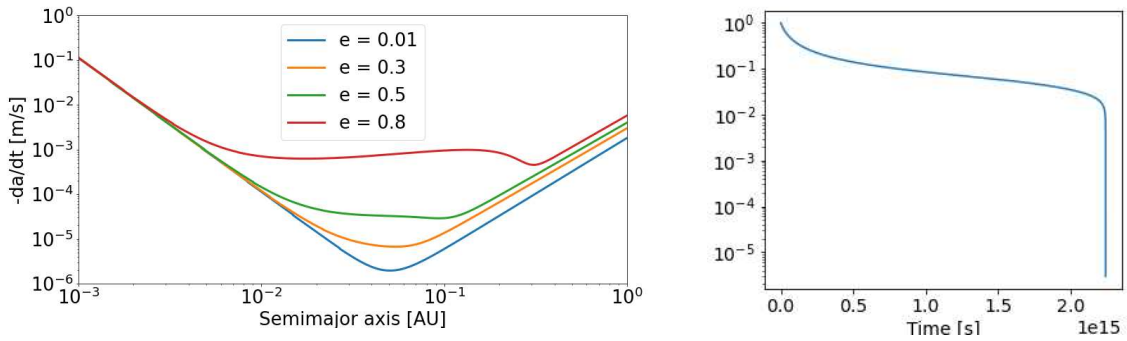
$$\dot{e}_{\rm gas} = \frac{12\pi\alpha c_s^2\Sigma_g}{\mu\Omega_b}\frac{(1-e^2)^{1/2}(1+e)^2}{e}\left[1-(1-e^2)^{1/2}\right], \tag{2.23}$$

where $\mu = m_1 m_2/(m_1+m_2)$ is the reduced mass and $\Omega_b = \sqrt{G(m_1+m_2)/a^3}$ is the Keplerian orbital frequency.
The previous equations are valid in the assumption that gravitational torques from the binary clear a cavity in the surrounding gas distribution, which remains circular and is surrounded by a circumbinary disk.

The binary also hardens due to the effect of gravitational-wave (GW) emission, which will govern the evolution at small semi major axes. The evolution of the semi-major axis a and eccentricity e due to GW hardening as [70]

$$\dot{a}_{\rm GW} = -\frac{64}{5}\,\frac{G^3 m_1 m_2(m_1+m_2)}{c^5 a^3(1-e^2)^{7/2}}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right) \tag{2.24}$$

$$\dot{e}_{\rm GW} = -\frac{304}{15}\,\frac{G^3 m_1 m_2(m_1+m_2)e}{c^5 a^4(1-e^2)^{5/2}}\left(1+\frac{121}{304}e^2\right) \tag{2.25}$$



(a) (b)

Figure 2.7: (a) Orbital decay rate $\dot{a}$ displayed as a function of the orbital separation a for different initial eccentricities ($e = 0.01 - 0.8$). Reproduction of a plot from [39]. (b) Evolution of the semimajor axis as a function of the delay time. For both of these plots we used the same fiducial values as [39]: $\alpha = 0.1$, $M_{\rm SMBH} = 10^7 M_\odot$, $h = 0.01$, $m_1 = m_2 = 25 M_\odot$.

The interaction with other objects also contributes to the hardening of hard binaries according to the Heggie's law [36] (see eq. 3.30 for more details). Here we completely neglect the hardening effect due to three-body interactions because they occur on a timescale larger than gas hardening.

The overall evolution of the binary is described by:

$$\dot{a} = \dot{a}_{\rm gas} + \dot{a}_{\rm GW}\,, \qquad \dot{e} = \dot{e}_{\rm gas} + \dot{e}_{\rm GW} \tag{2.26}$$

The results from the integration of these equations are shown in Figure 2.7. We observe two distinct trends corresponding to the disc-driven regime at large separations and the GW-driven regime at small separations, respectively. The transition between the two regimes happens when the $\dot{a}$ is minimum in modulus. The rate of the orbital decay in the gas-hardening regime decreases with decreasing separation because the BBH exchanges angular momentum with the gas less efficiently, while the orbital decay rate in the GW-driven regime increases with decreasing separation. Therefore, once GW emission dominates over gas hardening, the BBH is rapidly led to merger.

The semimajor axis continuously decreases until coalescence. The elapsed time from the pair-up is called delay time and is referred to as $t_{\rm del}$. Thus, the merging time is computed as

$$t_{\rm merg} = t_{\rm pair} + t_{\rm del} \tag{2.27}$$

2.6.1 Potential electromagnetic signatures of BBH mergers in AGNs

AGNs are the only known environments where stellar mass BHs are driven to merger in the presence of substantial gas densities. This can theoretically produce luminous electro-magnetic (EM) radiation.
The inspiral and merger of a BBH in a gaseous medium will produce shocks that propagate in the surrounding gas, heating the gas and giving rise to EM emission, and additionally a BBH may potentially accrete matter at super-Eddington rates, which can produce fast and

bright EM transients [61]. Also, accretion onto intermediate-mass BHs (IMBHs) produces broad-band EM radiation of a nature similar to what happens in quasars due to accretion onto SMBHs.

If the binary or the IMBH is buried within the accretion disk, emission can be reprocessed as it passes through the optically thick disk. This implies that high-energy radiation can be converted into optical or infrared radiation, and that the emission can be spread-out in time. On the other hand, if the disk is thin enough[8], the binary is able to clear a cavity in the disk during its inspiral, so its emission will leave the AGN without needing to pass through the disk and will reach the observer directly. This high-energy emission should be detectable in the X-rays by the Chandra observatory and in the γ-rays by the Fermi Space Telescope [5].

Detecting this EM radiation would be useful because it would allow us to identify BBHs (and also other kinds of binaries such as BH - neutron stars and binary neutron stars) before they merge and it would provide valuable information on the binary population in AGNs. Most AGNs are too luminous to be able to identify short electromagnetic transients, so observational campaigns focus on low-luminosity AGNs.
To date, no correlation has been observed between short high-energy bursts in the local universe and AGNs [61].

2.7 Spin signature

Embedded objects can weakly perturb the surface-density profile of the AGN gaseous disk, resulting in gas torques which tend to align both the BH spin vectors $\vec{\chi}$ and the binaries' orbital angular momentum vectors $\vec{L}$ with the angular momentum $\vec{J}$ of the disk itself. The efficiency of this effect depends on the density and age of the disk [10].
On the other hand, three-body encounters of BBHs with other objects tend to misalign $\vec{\chi}_{1,2}$ relative to $\vec{L}$ [84]. The competing effects of the gaseous disk and dynamical encounters on BBHs determine the distribution of BBH spin orientations.

We describe the distribution of BH spins in AGNs by using a phenomenological model [87] which is able to capture the salient features predicted from theoretical models.
We need to consider the interplay of three different phenomena:

i) The effect of gaseous torques on BH spins $\vec{\chi}$:

   BHs are expected to have randomly oriented spins at formation [56, 85]. As their orbits become embedded in the disk and they start migrating, they accrete a certain amount of mass Δm from the gaseous disk, resulting in a torque that tends to align their spin with the disk's angular momentum $\vec{J}$. The magnitude of the torque depends on Δm.

   If the disk is dense and long-lived, fully embedded BHs accrete a mass $\Delta m \geq 0.01\, m_i$, where m_i is the initial mass of the BH. Therefore, the torque from gas accretion efficiently reorients the BH spin vector $\vec{\chi}$ to align with the angular momentum vector $\vec{J}$ of the AGN disk.

   Otherwise, if the disk is dilute, $\Delta m \ll 0.01\, m_i$ and the BH spins are only torqued in alignment with $\vec{J}$ on long timescales $\tau \gtrsim 5\,\mathrm{Gyr}$. Similarly, even if AGN disks are dense, but typically short-lived ($\tau \lesssim 1\,\mathrm{Myr}$), this effect is weaker.

[8] The disk needs to have height H smaller than the Hill radius (eq. 3.29): $H < r_{\mathrm{Hill}}$ [39]

ii) The effect of gaseous torques on BBH orbital angular momentum $\vec{L}$:

When BBHs are embedded in the gaseous disk, they are subject to gaseous torques due to the exchange of angular momentum. This not only has the effect of efficiently hardening the binary, as we have seen in section 2.6, but it also damps the binary's eccentricity and drives $\vec{L}$ into alignment with $\vec{J}$. This is especially effective for long-lived AGNs ($\tau \gtrsim 5\,\mathrm{Gyr}$).

iii) The effect of dynamical encounters on BBH orbital angular momentum $\vec{L}$:

Encounters of BBHs with compact objects in the disk or in the spherical NSC component can randomize the orbital angular momentum of the BBH [84]. The efficiency of this process is inversely related with the efficiency of gas capture, which in turn increases for higher gas density and longer disk lifetime.

In Table 2.2, we display a short overview of the alignment of $\vec{\chi}_{1,2}$ and $\vec{L}$ for different properties of the disk. We consider the short-lived ($\tau \lesssim 5\,\mathrm{Gyr}$) vs long-lived ($\tau \gtrsim 5\,\mathrm{Gyr}$) cases as well as the high-density ($\overline{\rho_g} \gtrsim 10^{-11}\mathrm{g/cm^3}$) versus low-density ($\overline{\rho_g} \lesssim 10^{-11}\mathrm{g/cm^3}$) cases.
In this work we are often concerned with the surface gas density Σ_g, but the spin alignment is influenced by the volumetric gas density $\rho_g = \Sigma_g/2H$, where H is the disk's height. In order to avoid any confusion, we will use the terms 'dense' and 'dilute' to refer respectively to the high ρ_g or low ρ_g scenarios, and we will state high, medium or low Σ_g explicitly.

	Long-lived	Short-lived
Dense	$\vec{L} \parallel \vec{J}$	$\vec{L} \parallel \vec{J}$
	$\vec{\chi}_1 \parallel \vec{L}$	$\vec{\chi}_1 \parallel \vec{L}$
	$\vec{\chi}_2 \parallel \vec{\chi}_1$	$\vec{\chi}_2$ isotropic
Dilute	$\vec{L}$ isotropic	$\vec{L}$ isotropic
	$\vec{\chi}_1 \parallel \vec{J}$	$\vec{\chi}_1$ isotropic
	$\vec{\chi}_2 \parallel \vec{\chi}_1$	$\vec{\chi}_2$ isotropic

Table 2.2: Alignment of spin and angular momentum for different physical characteristics of the AGN disk. In the table, $\vec{J}$, $\vec{L}$, $\vec{\chi}_1$ and $\vec{\chi}_2$ are the AGN disk angular momentum, the BBH orbital angular momentum and the spins of the first and second BH.

2.8 Multiple BH generations: hierarchical mergers

We introduce here the concept of BH generation.
A BH which is the direct result of stellar evolution is called a first-generation (1g) BH. When two first-generation BHs form a binary and coalesce, the remnant is called a second-generation (2g) BH. In general, an Nth-generation (Ng) BH is the result of the repeated merger of N 1g BHs. For instance, an Ng BH can either be the result of an Mg-1g merger (where M + 1 = N) or of an Mg-Lg merger (where M + L = N, $L > 1$).

At birth, merger remnants receive a relativistic kick $v_{\rm kick}$ [53] which pushes the merger out of the migration trap and possibly out of the inner core of the NSC.
Since the gravitational potential well of the SMBH is very deep and the magnitude of the kick velocity is usually much smaller than the magnitude of the Keplerian velocity, the probability of ejection from the system is small. We will see this in more detail in chapter 4. Therefore, merger remnants' orbits are usually slightly perturbed: they typically find themselves out of the migration trap but not too far from it so that they can still interact with the disk. From there, they migrate again into the trap and they eventually find a new companion there. This gives rise to hierarchical mergers.

A single BH can only go through a finite number of hierarchical mergers before it comes across one of the following scenarios:

i) The disk has evaporated, therefore damping, migration and any effect due to gas torques stop.

ii) The relativistic kick received at merger is so strong that the remnant is ejected from the AGN.

iii) The number of BHs in the NSC is finite, therefore the maximum mass that can be accreted is limited and the BH may not find any companions to pair-up with.

iv) The BH is so massive that it opens a gap in the disk and can only move from its radial location due to Type II migration.

The scenario (i) is straightforward to evaluate by checking if, at generation g, it holds:

$$t_{\rm merg}^{(g)} \leq \tau \tag{2.28}$$

where τ is the disk lifetime and $t_{\rm merg}$ is computed as in eq. 2.27.

Assessing the scenario (ii) requires to compute the final velocity after kick and compare it with the escape velocity:

$$\|\vec{v}_{\rm fin}\| = \|\vec{v}_{\rm in} + \vec{v}_{\rm kick}\| \leq \|\vec{v}_{\rm esc}\| \tag{2.29}$$

$$v_{\rm esc}\,(R) = \sqrt{\frac{2\,G\,(M_{\rm SMBH} + M_{\rm NSC}(R)\,)}{R}} \tag{2.30}$$

We only consider the gravitational potential of the SMBH and of the inner NSC, neglecting the mass of the gaseous disk.

Scenario (iii) is taken into account ensuring that the BH does not accrete more mass than what is available in the inner NSC in form of other BHs:

$$M_{\rm acc} \leq M_{\rm BH}^{\rm max} \tag{2.31}$$

The mass accreted by the BH is the sum of its initial mass m_1 and of the masses of all the secondaries $m_2^{(g)}$ it pairs-up and merges with:

$$M_{\rm acc} = m_1 + \sum_g m_2^{(g)} \tag{2.32}$$

The index g represents the generation number.

The calculation of $M_{\rm BH}^{\rm max}$ is deferred to the following chapter (see eq.s 3.16 to 3.18).

The constraint on the accretion mass in eq. 2.31 works well as an upper bound for hierarchical mergers since reaching this condition means that just one seed BH was able to accrete all the mass available in the form of BHs in the migration trap.

Scenario (iv) is evaluated by checking whether both of the conditions in eqs. 2.14 and 2.16 are respected. For typical values of viscosity and aspect ratio, these conditions entail

$$m \gtrsim 10^{-1} M_{\rm SMBH}$$

When an object does Type II migration, it is bound to its radial location in the disk and can only move *with* the disk on its viscous timescale. See section 2.4 for more details.

First-generation BHs are never massive enough to satisfy the required conditions, hence only merger remnants have the potential of opening gaps in the disk. A merger remnant is the result of the coalescence of a BBH in the migration trap, but it receives a kick at birth which pushes it out of the trap. So, if a Nth-generation BH is able to open a gap, it surely does it out of the migration trap.
This has two consequences. The first is that the Type II migrator will never reach the migration trap. The second is that the gap it creates will prevent some Type I migrators from reaching the trap: it will intercept inward-moving migrators if it is located at a radius greater that the trap's, or it will intercept outward-moving migrators if the opposite is true. These intercepted BHs can potentially pair-up and merge with the Type II migrator, although their merger would not be assisted by gas hardening.

There is another element to consider to fully characterize the outcome of Type II migration: it can only happen for very high BH masses. Even if such a massive remnant could be generated in the AGN disk, it would surely be the result of a high-generation merger. This would surely happen in the later stages of the disk's lifetime. At that point, we can safely assume that most of the BHs have already migrated in the migration trap. Therefore, we can neglect any further pair-up event on the Type II migrator.

This is why we consider the onset of Type II migration to be one of the processes that can halt hierarchical mergers.

Even if Type II migrators cannot be subject to any further merger event, the presence of gaps has noteworthy consequences on disk density profile. In the instance of Type II migration, we bluntly neglect any changes in the disk's density.
Luckily, in our model the production of BHs that can do Type II migration turns out to be an impossible event (see chapter 4), so neglecting the change in the disk's density has no catastrophic consequences.

2.9 Other physical processes

There are some other physical processes, besides those described from section 2.3 to section 2.8, that can become relevant in an AGN environment.
Here we briefly list some of these processes which we neglect in our model.

First of all, in our model we assume that the pair-up (and eventually the merger) of a BBH can only happen in the migration trap. Of course this is not entirely realistic because we expect to have some hard original binaries in the NSC which can merger before reaching the migration trap and, furthermore, two BHs could lie close enough to each other to pair-up even before reaching the migration trap.
Locations in the disk outside of the migration trap are called 'the bulk'. In [60] they find that, although more than 50% of mergers happen in the bulk, hierarchical merger is only efficient in the migration trap. Hence, as an editorial choice, we neglect the possibility of pair-up in the bulk since we are interested in the formation of IMBHs via hierarchical mergers.

Also, we are neglecting three-body interactions and their effects on BH scattering and BBH hardening. This is due to the dearth of models for these interactions: in the literature there are some works (e.g. [16, 24] used in Appendix A) which assume an isotropic distribution of velocities and are appropriate for spherical star clusters. In a Keplerian disk geometry, the distribution of velocities is much different and these models are not appropriate [57].

Moreover, we neglect dynamical interactions with the SMBH and gravitational perturbations caused by the presence of IMBHs. For example, in [20] they show that the presence of IMBHs in the disk of mass may enhance the ionization of BBHs, therefore decreasing the merger rate.

Finally, we are entirely neglecting the evolution of the disk. The efficiency of all processes described in sections 2.3–2.8 strongly depends on the disk's density and aspect ratio. If these quantities evolve over time, all dynamical processes will be affected as well.

Chapter 3

Methods: Semi-analytical model

We have seen that the interaction with the gaseous disk can drive the dynamical encounter of black hole binaries and potentially lead to the formation of IMBHs.
In this Chapter we describe the semi-analytical model used in order to produce a statistical sample of BBHs formed in AGN disks, with the goal of exploring all the possible outcomes.
We expanded the semi-analytic code FASTCLUSTER [54, 55], which already included prescriptions for hierarchical mergers in globular, young and nuclear star clusters. We incorporated a parallel model for AGNs.

3.1 Setup of system parameters

3.1.1 Super-massive black hole (SMBH)

SMBH mass In order to determine the mass of the SMBH, we use the catalog of active SMBHs of IllustrisTNG. This is a cosmological magnetohydrodynamical simulation of galaxy formation in a cubic box of size $(L_{\rm TNG})^3$.
There are multiple versions of IllustrisTNG which have a different simulation volume and a different resolution. We consider the ones with the smallest and the biggest simulation volume respectively: TNG50 [67, 72] and TNG300 [71, 82].

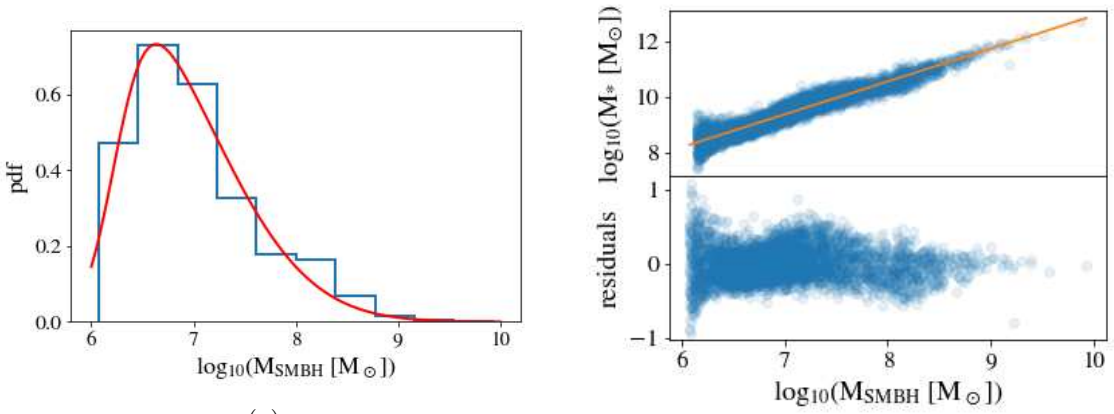
The TNG50 has $L_{\rm TNG} \simeq 50\,\mathrm{Mpc}$ and a resolution of about $(24\,\mathrm{kpc})^3$. Instead, the TNG300 has a larger simulation volume with $L_{\rm TNG} \simeq 300\,\mathrm{Mpc}$ but a coarser resolution of about $(121\,\mathrm{kpc})^3$. They produce a different population of SMBHs that correlate differently with the host galaxy mass.

In the IllustrisTNG public data every SMBH is marked out by a flag describing whether it is accreting mass via an AGN disk or not. We focus on the sub-population of SMBHs that are active at the considered time, i.e. the AGNs.
Also, in this work we limit ourselves to the data at redshift $z = 0$.

`(TNG50)` The active SMBH population of TNG50 is well described by a skewed Gaussian (eq. 3.1, where erf is the error function) with best fit parameters $x_0 = 6.23 M_\odot$, $\alpha = 3.79$ and $\omega = 0.94$ (see Figure 3.1a).

$$f(x) = \frac{1}{\sqrt{2\pi}\omega}\, \exp^{-(x-x_0)^2/2\omega^2} \left[1 + \mathrm{erf}\left(\alpha \frac{x - x_0}{\sqrt{2}\omega}\right)\right] \tag{3.1}$$

We sample x from the distribution of eq. 3.1 using rejection sampling. The resulting mass of the SMBH is $M_{\mathrm{SMBH}} = 10^x$.



(a) (b)

Figure 3.1: (a) Catalog of active SMBH masses from Illustris TNG50 (blue histogram) and best-fit skewed gaussian (red plot). (b) Relation between the mass of the SMBH and the mass of the host galaxy in the catalog of active SMBH of Illustris TNG50. The residuals are $\log\left(M_*^{\mathrm{true}}/M_\odot\right) - \log\left(M_*^{\mathrm{fit}}/M_\odot\right)$, i.e. the difference between the real value and the best-fit value.

`(TNG300)` The active SMBH population of TNG300 is roughly described by a power law (eq. 3.2) with best-fit parameters $A = 0.1$, $x_0 = 6.05$ and $n = -1.3$ (see Figure 3.2a).

$$g(x) = A\,(x - x_0)^n \tag{3.2}$$

The resulting SMBH population is strongly restricted to $\log\left(M_{\mathrm{SMBH}}/M_\odot\right) \lesssim 7$.
We sample x from the distribution of eq. 3.2 using inverse random sampling. With this sampling procedure, we ignore the low-density bump at $\log\left(M_{\mathrm{SMBH}}/M_\odot\right) \sim 8$ visible in Figure 3.2a. The mass of the SMBH is $M_{\mathrm{SMBH}} = 10^x$.

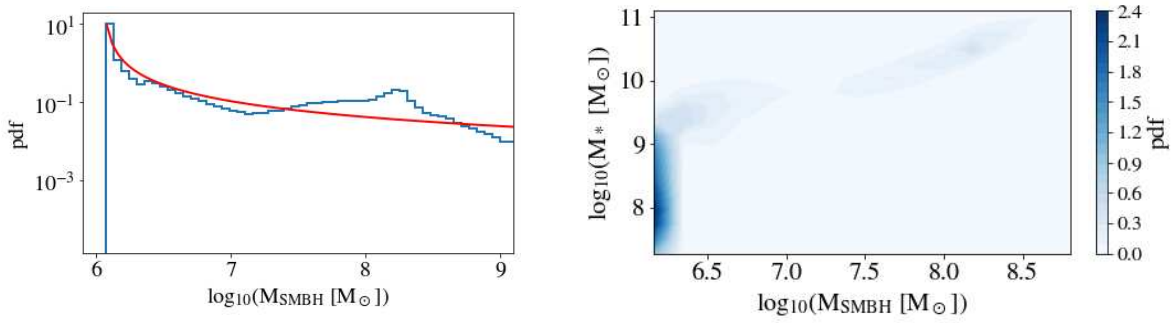
Gravitational radius It is useful to define the gravitational radius R_g as

$$R_g = \frac{GM_{\mathrm{SMBH}}}{c^2} \tag{3.3}$$

We stress that, contrary to what is often done in the literature, we define it as half of a Schwarzschild radius rather than as a whole Schwarzschild radius.

Host galaxy mass In general, the SMBH mass M_{SMBH} correlates with the mass of the host galaxy M_*. Once we know M_{SMBH}, we can determine M_*.

`(TNG50)` The Illustris TNG50 data show a strong correlation between the mass of the SMBH and the mass of the host galaxy, as shown in Figure 3.1b where the data are in blue and a power-law best fit in orange.



(a) (b)

Figure 3.2: (a) Catalog of active SMBH masses from Illustris TNG300 (blue histogram) and best-fit power law (red plot). (b) Relation between the mass of the SMBH and the mass of the host galaxy in the catalog of active SMBH of Illustris TNG300. The residuals are $\log\left(M_*^{\text{true}}/M_\odot\right) - \log\left(M_*^{\text{fit}}/M_\odot\right)$, i.e. the difference between the real value and the best-fit value.

The best-fit equation (shown in Figure 3.3a) is

$$\log\left(M_*/M_\odot\right) = 0.798\,\log\left(M_{\text{SMBH}}/M_\odot\right) - 0.499 \tag{3.4}$$

After selecting the SMBH mass, we sample $\log\left(M_*/M_\odot\right)$ from a gaussian distribution with mean centered on the corresponding value determined by the relation 3.4 and width $\sigma = 0.25$, so that most of the data falls under the $\pm 2\sigma$ dispersion.

`(TNG300)` In the Illustris TNG300 simulation the SMBH mass correlates weakly with the host galaxy mass. As seen in Figure 3.2b, in the range of interest for the SMBH mass ($\log\left(M_{\text{SMBH}}/M_\odot\right) \lesssim 7$) the host galaxy mass seems uncorrelated with the SMBH mass, except for a low-density blob at $\log\left(M_*/M_\odot\right) \gtrsim 9$ which we ignore.
We sample the host galaxy mass uniformly in the range populated in Figure 3.2b.

$$\log\left(M_*/M_\odot\right) \sim \mathcal{U}\left(7.5\,, 9.0\right) \tag{3.5}$$

Coexistence with a NSC We are only interested in SMBHs which can coexist with a NSC. According to [33], this is only possible for galaxy masses such that

$$\log\left(M_*/M_\odot\right) \lesssim 10.5 \tag{3.6}$$

Galaxies with masses higher than this threshold cannot host a NSC.

`(TNG50)` Using eq. 3.4, this translates to an upper limit on the SMBH mass. We only consider active SMBH with masses such that

$$\log\left(M_{\text{SMBH}}/M_\odot\right) \lesssim 7.88 \tag{3.7}$$

`(TNG300)` Sampling the host galaxy mass as in eq. 3.5, the consition of eq. 3.6 is automatically respected.

3.1.2 AGN disk

Radial extension of the disk The disk is assumed to have radial extension between the ISCO radius for a non-rotating BH, which we call $R_{\rm min}$ in this context, and an outer radius $R_{\rm max}$ beyond which the disk's self-gravity becomes important [32, 88]. For $R > R_{\rm max}$ the disk becomes fragmented and experiences star formation, so the viscous interaction with the BHs has a different behavior than described in section 2.3 and section 2.4 and we neglect it.

$$R_{\rm min} = 6\, R_g = \frac{6\, G M_{\rm SMBH}}{c^2} \tag{3.8}$$

$$R_{\rm max} = 0.1\,{\rm pc} \left(\frac{M_{\rm SMBH}}{10^6 M_\odot} \right)^{1/2} \tag{3.9}$$

Migration trap As anticipated in section 2.4, disks can have migration traps. These are locations where torques change sign from positive (leading to outward migration) to negative (leading to inward migration), so that migration will stall and BHs will pile up.

The position of migration traps in the disk can be determined by computing the sign of torques, as done in [7]. They find that migration traps are found at locations where the slope of the gas surface density profile changes sign from positive to negative, i.e. at local maxima. In this work, they consider both the SG [81] and the TQM [86] disk models (see subsection 2.2.2 for more details).
They find that in a SG disk there two local maxima in the gas surface density and therefore two migration traps: there is an inner trap at $49\, R_g$ and an outer one at $662\, R_g$. For a TQM disk there is only one trap at $490\, R_g$, where the slope of density profile goes from positive to negative.
At the migration trap radius of the TQM model, there is a discontinuity of almost two orders of magnitudes in the surface density profile. Since the BBHs of our model pair-up and merge in the migration trap, we often need to use the surface density *at* the migration trap in our code. Therefore, using the TQM density profile can be numerically dangerous because the density is undefined at the migration trap.

For this reason, we choose to use a SG disk model. We notice from Figure 2.6 (original model is the blue line) that the inner migration trap coincides with a local maximum in the density profile, while the outer migration trap corresponds to a global maximum.
For simplicity, we ignore the local overdensity in the disk at $49\, R_g$ and we only assume the existence of the outer migration trap. We define the trap radius as

$$R_{\rm trap} = 662\, R_g = 662\, \frac{G M_{\rm SMBH}}{c^2} \tag{3.10}$$

Physical parameters of the disk We assume a viscosity parameter $\alpha = 0.3$ and an accretion rate onto the SMBH of $0.1\, \dot{M}_{\rm Edd}$. The physical parameters of the disk are functions of the mass of the SMBH and the distance R from the SMBH. The expression for

the gas surface density Σ_g, the disk aspect ratio h and the sound speed c_s are the following:

$$\Sigma_g\left(R, M_{\rm SMBH}\right) = 7.94\times10^5\,{\rm g/cm^2}\left(\frac{M_{\rm SMBH}}{10^8 M_\odot}\right)^{4/5}\begin{cases}\left(\dfrac{R}{10^3\,R_g}\right)^{0.8} & R\le 10^3\,R_g\\ \left(\dfrac{R}{10^3\,R_g}\right)^{-1.49} & R>10^3\,R_g\end{cases} \tag{3.11}$$

$$h\left(R, M_{\rm SMBH}\right) = 7.59\times10^{-3}\left(\frac{M_{\rm SMBH}}{10^8 M_\odot}\right)^{-3/20}\begin{cases}\left(\dfrac{R}{10^3\,R_g}\right)^{-0.6} & R\le 10^3\,R_g\\ \left(\dfrac{R}{10^3\,R_g}\right)^{0.5} & R>10^3\,R_g\end{cases} \tag{3.12}$$

$$c_s\left(R, M_{\rm SMBH}\right) = 5.37\times10^6\,{\rm km/s}\left(\frac{M_{\rm SMBH}}{10^8 M_\odot}\right)^{3/2}\begin{cases}\left(\dfrac{R}{10^3\,R_g}\right)^{-1.1} & R\le 10^3\,R_g\\ 1 & R>10^3\,R_g\end{cases} \tag{3.13}$$

The radial dependence is a simplified version of the Sirko-Goodman model [81] (eq.s 2.5–2.7), which is re-scaled to allow for different $M_{\rm SMBH}$ (eq. 2.8).

3.1.3 Nuclear star cluster (NSC)

NSC mass Once we know the galaxy mass, we need to determine the mass of its NSC. We have already ensured that we are only considering galaxies that can host a NSC in subsection 3.1.1.

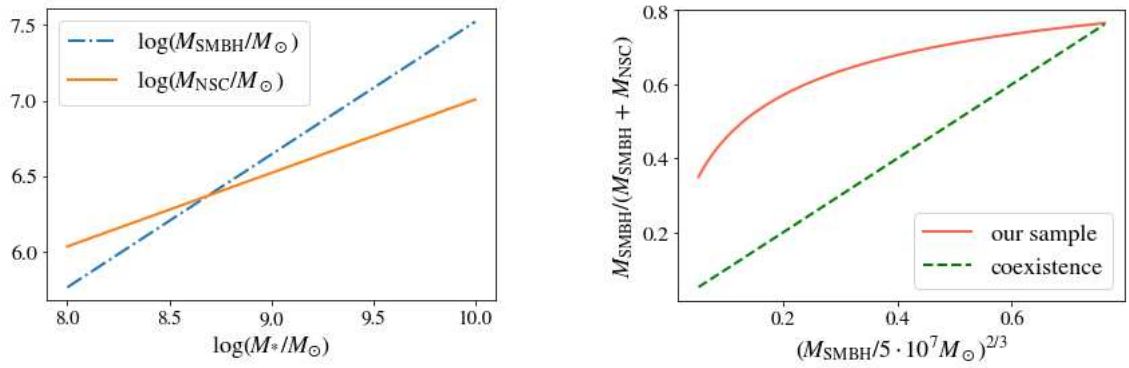
According to a large review on NSC [68] (see Fig. 12 of the reference), considering only early-type galaxies, we see that the galaxy mass and NSC mass are related according to:

$$\log\left(M_{\rm NSC}/M_\odot\right) = 0.485\,\log\left(M_*/M_\odot\right) + 2.16 \tag{3.14}$$

This relation is shown in Figure 3.3a.
To account for the spread in the data, we sample $\log\left(M_{\rm NSC}/M_\odot\right)$ from a gaussian distribution with mean centered on the corresponding value determined by relation 3.14 and width $\sigma = 0.4$, so that most of the data falls under the $\pm2\sigma$ dispersion.

With this sampling method, we generate a NSC mass that is consistent with the prescription for the coexistence with a SMBH discussed in subsection 2.2.1.
We verify this consistency in Figure 3.3b: the green line is the prescription for coexistence of eq. 2.1 and the red line is the relation we obtain with our procedure. It is clear from the figure that, although the two relations are compatible with each other, with our prescription we slightly under-estimate the NSC mass.

NSC effective radius According to a review of the current NSC knowledge [68], the effective radius of a NSC mildly correlates to its mass. In Figure 3.4 we show the data points in pink and a power-law best fit in blue.



(a) (b)

Figure 3.3: (a) The SMBH and the NSC mass as functions of the host galaxy mass. This does not account for statistical fluctuations. (b) Relation between the SMBH mass and the NSC mass: the red line is the relation followed by our sample obtained with the procedure outlined in sections 3.1.1–3.1.3 for Illustris TNG50, the green dashed line is the relation they should follow for coexistence (eq. 2.1).

The best-fit relation between the mass on the NSC and the effective radius is

$$\log\left(r_h/\mathrm{pc}\right) = 0.228\,\log\left(M_{\mathrm{NSC}}/M_\odot\right) - 0.797 \tag{3.15}$$

To account for the spread in the data, we sample $\log\left(r_h/\mathrm{pc}\right)$ from a Gaussian distribution with mean centered on the corresponding value determined by relation 3.15 and width $\sigma = 0.2$, so that most of the data falls under the $\pm 2\sigma$ dispersion.

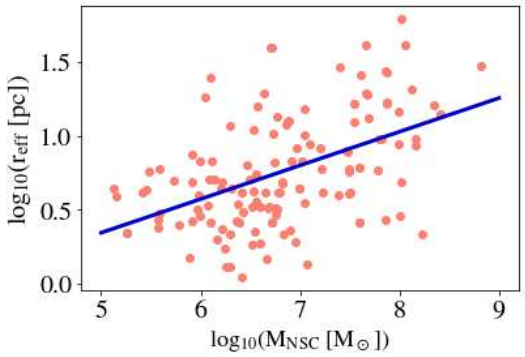


Figure 3.4: Relation between the effective radius (or half-light radius) of a NSC and its mass. Data from [68] is in pink and the best-fit power-law relation is in blue.

Number of BHs in the NSC The spatial distribution of the stars is approximated with a Plummer model, which is a simple model for a spherical distribution of mass [73]. The Plummer mass density profile is

$$\rho\left(R\right) = \frac{3M_0}{4\pi a_{\mathrm{PL}}^3}\left(1 + \frac{R^2}{a_{\mathrm{PL}}^2}\right)^{-\frac{5}{2}} \tag{3.16}$$

The mass enclosed within a radius R is given by

$$M(R) = M_0 \frac{R^3}{\left(R^2 + a_{\rm PL}^2\right)^{3/2}} \tag{3.17}$$

where R is the distance from the SMBH, M_0 is the total mass of the NSC and $a_{\rm PL} = r_h/1.3$ is the scale parameter for the Plummer model. The parameters M_0 and r_h are given by eq.s 3.14 and 3.15.

The mass fraction of BHs in the AGN disk is sampled from a Gaussian with mean 0.04 [5] and standard deviation 0.01 (arbitrary): $f_{\rm BH} \sim \mathcal{N}(0.04\,, 0.01\,)$

The number of BH in the AGN and their cumulative mass are given by the following equations:

$$N_{\rm BH} = \frac{f_{\rm BH}\, f_{\rm trap}}{2} \frac{M(R_{\rm max})}{\langle m \rangle}\,, \qquad M_{\rm BH}^{\rm max} = N_{\rm BH}\, \langle m_{\rm BH} \rangle \tag{3.18}$$

Where $M(R)$ is defined in eq. 3.17, the mean black hole mass $\langle m_{\rm BH} \rangle$ is computed from data obtained from population synthesis simulations, the distance $R_{\rm max}$ is defined in eq. 3.9, the factor 1/2 is to account for prograde orbiters only (statistically half of the total BH population in the NSC), and the mean stellar mass $\langle m \rangle$ is computed using the Kroupa initial mass function [46].
The parameter $f_{\rm trap}$ is the fraction between the number of BHs that are able to reach the migration trap on a timescale shorter than the disk's lifetime and the total number of BHs that interact with the disk: this study focuses on BBH pair-ups in the migration trap, therefore we are not interested in any BHs that live outside of that location. For an operational definition of $f_{\rm trap}$, see section 3.3.

Velocity dispersion The velocity dispersion of stars scales with the SMBH mass roughly as [63]

$$\sigma = 200\,{\rm km/s} \left(\frac{M_{\rm SMBH}}{10^8 M_\odot}\right)^{1/5} \tag{3.19}$$

The value of the exponent is appropriate for unbarred galaxies, whereas it should be raised to 1/6 when including barred galaxies in the sample [76]. We choose to neglect barred galaxies because the motion of stars is far from isotropic and their spatial distribution is far from spherical, so the Plummer distribution of eq. 3.16 is not a valid approximation.

3.2 Setup of initial black hole population

Once we have determined the parameters characterizing the system, we want to simulate the evolution of BHs in it. In order to have statistical significance we sample the masses, spins and initial positions of N BHs and follow their evolution in the system individually.

Mass We randomly draw the 1st-generation masses m_1 from a catalog of masses obtained with the population synthesis code MOBSE [27, 28, 29, 30]. See Figure 3.5.
We assume metallicity $Z = 0.02$, i.e. approximately solar, matching the typical metallicity at the center of massive galaxies in the local Universe.

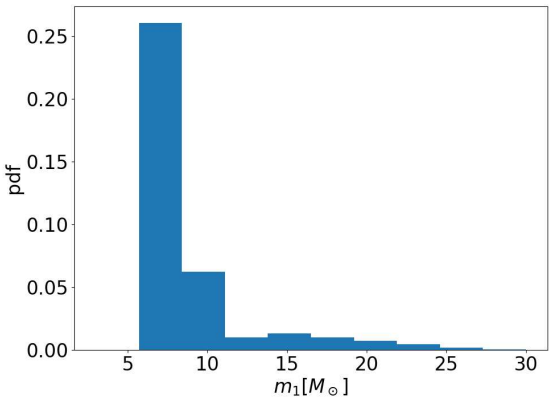


Figure 3.5: Initial mass function for the first generation of BHs.

Spin magnitude We draw the dimensionless spin magnitude χ_1 from a Maxwellian distribution with one-dimensional root-mean square $\sigma_\chi = 0.1$, truncated at $\chi = 1$. We choose $\sigma_\chi = 0.1$ because it is quite reminiscent of the spins inferred from the third GW transient catalog (GWTC-3, [50]).

Spin tilt The way we set the spin tilt depends on the density and lifetime of the disk, as discussed in section 2.7.
We only care about the alignment of $\vec{\chi}_{1,2}$ with $\vec{L}$ (e.g. in the dilute and long-lived case it is not relevant that $\vec{\chi}_1 \parallel \vec{J}$), so the whole model is just characterized by two parameters: σ_1 and σ_2.

$$\vec{\chi}_1 = \chi_1 \begin{pmatrix} \sin\theta_1 \, \cos\phi_1 \\ \sin\theta_1 \, \sin\phi_1 \\ \cos\theta_1 \end{pmatrix} \qquad \vec{\chi}_2 = \chi_2 \begin{pmatrix} \sin\theta_2 \, \cos\phi_2 \\ \sin\theta_2 \, \sin\phi_2 \\ \cos\theta_2 \end{pmatrix} \tag{3.20}$$

We sample the cosine of the angle θ_1 between $\vec{\chi}_1$ and $\vec{L}$ from a truncated Gaussian centered in 1 with standard deviation σ_1 [87]

$$\cos\theta_1 \sim \mathcal{N}_t\,(1\,,\sigma_1\,) \tag{3.21}$$

where we set $\sigma_1 = 0.1$ or $\sigma_1 = 10$ based on the physical characteristics of the disk.

Setting $\sigma_1 = 0.1$ means that the spin of the primary is aligned with the orbital angular momentum of the binary, namely $\vec{\chi}_1 \parallel \vec{L}$. As explained in section 2.7 (see Table 2.2), this happens if the disk is dense.
Instead, setting $\sigma_1 = 10$ is equivalent to sampling $\cos\theta_1$ from a uniform distribution in the interval $[0\,,1]$ and coincides with the isotropic case. The direction $\vec{\chi}_1$ is isotropic only if the disk is dilute and short-lived. If the disk is dilute and long-lived, $\vec{\chi}_1$ is aligned with the disk's angular momentum $\vec{J}$ but the BBH orbital angular momentum $\vec{L}$ is isotropic (see again Table 2.2). Therefore, $\vec{\chi}_1$ and $\vec{L}$ will not be aligned and we can set $\sigma_1 = 10$ also in this case.

We sample the cosine of the angle θ_2 between $\vec{\chi}_2$ and $\vec{L}$ from a truncated Gaussian centered in $\cos\theta_1$ with standard deviation σ_2

$$\cos\theta_2 \sim \mathcal{N}_t\,(\cos\theta_1\,,\sigma_2\,) \tag{3.22}$$

where we set $\sigma_2 = 0.1$ or $\sigma_2 = 10$ based on the physical characteristics of the disk.

Setting $\sigma_2 = 0.1$ means that the spin of the secondary is aligned with that of the primary, namely $\vec{\chi}_2 \parallel \vec{\chi}_1$. As seen in section 2.7, this happens is the disk is long-lived. Setting $\sigma_2 = 10$ is equivalent to sampling $\cos\theta_2$ from a uniform distribution in the interval $[0\,,1]$ and coincides with the isotropic case. This happens if the disk is short-lived (see Table 2.2).

Summarizing, we set $\sigma_i = 0.1$ in case of strong alignment or $\sigma_i = 10$ in the isotropic case (i.e. we sample from an uniform distribution). In Table 3.1 we illustrate the choice of $\sigma_{1,2}$ in different scenarios.
For the azimuthal direction, $\cos\phi_{1,2}$ are drawn from uniform distributions.

	Long-lived	Short-lived
Dense	$\sigma_1 = 0.1$ $\sigma_2 = 0.1$	$\sigma_1 = 0.1$ $\sigma_2 = 10$
Dilute	$\sigma_1 = 10$ $\sigma_2 = 0.1$	$\sigma_1 = 10$ $\sigma_2 = 10$

Table 3.1: Standard deviations of the gaussian distributions from which $\cos\theta_{1,2}$ are sampled, as in eq.s 3.21 and 3.22.

The model for the spin tilts $\theta_{1,2}$ has a direct effect on the effective spin of the resulting population. Using the definition of the effective spin in eq. 1.12 and $\vec{\chi}_i \cdot \hat{L} = \chi_i \cos\theta_i$, we can write the effective spin as

$$\chi_{\rm eff} = \frac{m_1\chi_1\cos\theta_1 + m_2\chi_2\cos\theta_2}{m_1 + m_2} \tag{3.23}$$

We show the effective spin distribution in the four scenarios in Figure 3.6.
We remind that setting $\sigma_1 = 0.1$ implies that $\theta_1 \sim 0$, while setting $\sigma_2 = 0.1$ implies that $\theta_2 \sim \theta_1$. It is easy to see from eq. 3.23 that $\sigma_1 = 0.1$ is a necessary condition for the effective spin distribution to have a peak at $|\chi_{\rm eff}| \neq 0$.
The effective spin distribution has a peak on $|\chi_{\rm eff}| \neq 0$ if and only if the disk's volumetric density is high.

Initial positions We randomly draw an initial radial position for each BH. The radial extension of the accretion disk is very small compared to the typical dimension of a NSC, so we consider the numerical density of objects at radii $< R_{\rm max}$ to be uniform in radius.

$$R \sim \mathcal{U}\left(R_{\rm min}\,, R_{\rm max}\right) \tag{3.24}$$

This is the radius R used in eq. 2.13. We also use it as the semi major axis A of eq. 2.12, implicitly assuming orbits with negligible eccentricity. This is acceptable because $t_{\rm damp} \ll t_{\rm migr}$, so circularization is very fast.

3.3 First generation

After setting up the characteristics of the first generation of BHs, we follow their evolution in the disk.

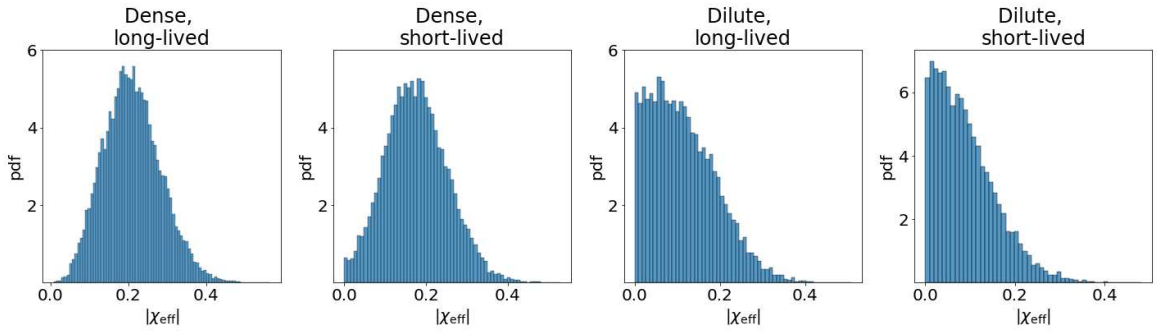


Figure 3.6: Effective spin distributions resulting from the alignment of the spins of the primary and secondary BHs $\vec{\chi}_{1,2}$ with the orbital angular momentum of the binary $\vec{L}$.

Pair-up Prograde BHs that cross the disk during their evolution will change their orbit due to the effects of damping and migration described respectively in section 2.3 and section 2.4.
Using eq. 2.12 and 2.13, we evaluate the damping and migration timescale for each BH. So after a time

$$t_{\rm pair}^{(1)} = t_{\rm damp}^{(1)} + t_{\rm migr}^{(1)} \tag{3.25}$$

we consider each BH to be in a circular orbit in the migration trap at a radius given by eq. 3.10. We assume that the BHs already formed before the formation of the AGN disk, so we neglect their formation time.
Of course, respecting the constraint in eq. 2.28, if the pairing time is too long, the system is removed from the simulation and is not evolved.

We compute the fraction $f_{\rm trap}$ of BHs that reach the migration trap by counting the number of first-generation BHs for which $t_{\rm pair}^{(1)} < T$, where T is the disk's lifetime, and dividing it by N (i.e. the total number of first-generation BHs in the simulation). This parameter is used for the computation of the maximum mass that can be accreted by a single BH, as in eq. 3.18.

Since the density of BHs in the migration trap is high, we assume that the pair-up with a secondary BH is instantaneous upon entrance in the migration trap. See Appendix A for more details.

Secondary mass We need to determine the properties of the newly-formed BBH, let us start with the secondary mass.
We allow for Ng-Mg mergers, but we always assume $M \leq N$. So a first generation primary BH ($N = 1$) will necessarily pair up with a first generation secondary ($M = 1$).
We defer the discussion of the model for Ng-Mg pairing to section 3.4. Here we describe the model for 1g-1g mergers, which is of course a subcase of the more general model.
For the first generation, the secondary mass m_2 is randomly drawn with probability distribution [69]

$$p\left(m_2|m_1\right) \propto \left(m_1 + m_2\right)^4 \tag{3.26}$$

between $m_2^{\rm min} = 3M_\odot$ and $m_2^{\rm max} = m_1$.

The previous equation comes from a model that includes mass segregation and other effects happening in star clusters, producing a bias towards higher m_2.

Secondary spin magnitude The secondary spin modulus is set in the same way as for the primary, as described in section 3.2.

Secondary spin tilt The secondary spin tilt is set by sampling the cosine of the angle θ_2 as in eq. 3.22. The value of the standard deviation σ_2 depends on the physical characteristics of the disk as in table 3.1.

BBH semi-major axis We assign the initial semi major axis a of the binary sampling from a distribution (following [85])

$$p(a) \propto a^{9/2} \qquad \text{for} \quad a \in [a_{\min}, a_{\max}] \tag{3.27}$$

$$a_{\min} = 1 R_{\odot}\,, \quad a_{\max} = r_{\text{Hill}} \tag{3.28}$$

The dependency $a^{9/2}$ comes from a calculation of the formation rate of binaries in dense environments [8], the lower limit $a_{\min}$ is arbitrary while the upper limit $a_{\max}$ is chosen to ensure that binaries can form only when the two bodies undergo a strong encounter.
The upper limit $a_{\max} = r_{\text{Hill}}$ requires an encounter between the primary and the secondary with impact parameter less than the BBH Hill radius, which is computed accounting for the fact that in our model the BBH forms in the migration trap

$$r_{\text{Hill}} = R_{\text{trap}} \left(\frac{m_1 + m_2}{3\, M_{\text{SMBH}}} \right)^{1/3} \tag{3.29}$$

This is necessary because, if you have two objects at distances greater than r_{Hill}, the gravity of the SMBH dominates and prevents the formation of the BBH.

BBH eccentricity We set the initial eccentricity e following a thermal distribution $p(e) \propto 2e$ for e between 0 and 1.

This law was initially obtained by [40] for a thermalized population of binaries, i.e. a population for which the distribution of energies follows a Boltzmann distribution. It is a standard procedure for N-body and semi-analytical codes to use this eccentricity distribution as an initial condition [55, 85]

Soft binaries According to the Heggie's law [36] a binary can survive in a star cluster only if it is hard, i.e. if its binding energy E_b is larger than the average kinetic energy of a field star $\langle E_k \rangle$.

$$E_b = \frac{G m_1 m_2}{2a} \geq \langle E_k \rangle = \frac{1}{2} \langle m \rangle \sigma^2 \tag{3.30}$$

where a is the semi-major axis of the binary, $\langle m \rangle$ is the average mass of a star in the NSC and σ is the three-dimensional velocity dispersion.
This is because the interaction between a binary and a field star has a different effect based on their relative energy. If the binary is soft (i.e. $E_b < \langle E_k \rangle$) the binary will gain energy and

become softer ($E_b \to 0$), whereas if the binary is hard (i.e. $E_b > \langle E_k \rangle$), the binary will lose some of its energy and become harder ($|E_b|$ increases).
Therefore, in a dense environment where the interaction between binaries and single stars is frequent, soft binaries will expand and eventually be ionized ($a \to \infty$) whereas hard binaries will shrink ($a \to 0$).

Obviously, when a binary hardens its extra energy is transferred to the field star which increases its kinetic energy. For the conservation of momentum, both the binary and the star will receive a kick. Since typically $m_1 + m_2 > \langle m \rangle$, the star has a higher kick velocity than the binary.

Hence, out of all the BBHs, we dynamically evolve only those for which eq. 3.30 is respected. If a BBH does not satisfy this condition, we do not consider it any further. Otherwise, we evolve it by gas hardening and by GW emission, as described below.

BBH hardening The binary is now in the migration trap, so we set $R = R_{\rm trap}$ and update the quantities in eq. 3.11-3.13 accordingly.
We perform the integration of the hardening eq. 2.26 using the Euler method and an adaptive time-step. The time-step adaptation works such that we reduce (or increase) the time-step Δt by a factor of 10 (or 2) whenever the percentage change of a between two time-steps is $> 1\%$ (or $\leq 0.1\%$).
We refer to the integration time (i.e. the delay time between pair-up and, eventually, merger) as $t_{\rm del}$.
If the time from the beginning of the interaction with the disk (i.e. $t_{\rm pair} + t_{\rm del}$) is larger than the lifetime of the disk, that means that the disk has evaporated before the binary could merge. The BBH keeps hardening due to gravitational-wave emission only. From this moment onward, we integrate eq.s 2.24 and 2.25.

We consider the merging to happen when the binary members cross the ISCO radius of a non-spinning black hole with mass equal to the total mass of the binary system

$$a \leq r_{\rm ISCO} = 6\,\frac{G\,(m_1 + m_2)}{c^2} \tag{3.31}$$

with a tolerance of $0.1\,r_{\rm ISCO}$. This happens at a time $t_{\rm merg} = t_{\rm pair} + t_{\rm del}$.

Merger remnant When a binary merges, we model the mass and spin of the merger remnant using fitting formulas from numerical relativity, as described by [41].

Relativistic kick At birth, merger remnants receive a relativistic kick $v_{\rm kick}$ because of the transfer of linear momentum caused by asymmetries in GW emission. We use the model in eq. (14.202) of [53] for the magnitude of the kick.
We draw a random kick direction by randomly sampling two angles (spherical coordinates) θ and ϕ from the distributions $f(\phi)$ and $f(\theta)$.

$$f(\phi) = \frac{\pi}{2} \qquad f(\theta) = \frac{\sin(\theta)}{2} \tag{3.32}$$

The relativistic kick $v_{\rm kick}$ pushes the merger remnant out of the migration trap. We compute the new velocity as

$$v_{\rm fin} = \begin{pmatrix} v_{\rm Kepl}(R_{\rm trap}) + \cos\phi\, \sin\theta\, v_{\rm kick} \\ \sin\phi\, \sin\theta\, v_{\rm kick} \\ \cos\theta\, v_{\rm kick} \end{pmatrix} \tag{3.33}$$

where $v_{\rm Kepl}\,(R_{\rm trap})$ is the Keplerian velocity in the migration trap computed as in eq. 2.19. We assume Keplerian motion with negligible eccentricity after the kick.
The position of the remnant is updated to

$$R_{\rm fin} = \frac{G\,(M_{\rm SMBH} + M_{\rm NSC}(r)\,)}{v_{\rm fin}^2} \tag{3.34}$$

Also, the quantities in eq. 3.11-3.13 are update accordingly.

As a safety check, we ensure that the new radial position $R_{\rm fin}$ is lower than the maximum radius of the disk $R_{\rm max}$ (eq. 3.9), meaning that the remnant can experience damping and be embedded in the disk. Otherwise, we discard the remnant and do not consider it for future generations.

3.4 Nth generations

Merger products are generally retained by the deep potential well, allowing for hierarchical BBH mergers.
After each merger events, we check that the requirements (i)-(iv) listed in section 2.8 are met. We briefly remind them here.

i) The merging time must be smaller than the disk lifetime, as in eq. 2.28. If this requirement is not met, it means that the disk has evaporated and there can no longer be neither damping nor migration into the migration trap.

ii) The velocity of the merger remnant $v_{\rm fin}$ (eq. 3.33) must be smaller than the escape velocity, as in eq. 2.29. This ensures that the remnant is retained in the inner NSC and can keep interacting with the disk.

iii) The mass accreted by the BH over its lifetime must be lower than the total mass available in the inner NSC, as in eq. 2.31. Otherwise, the remnant can't find any companion to pair-up with.

iv) The mass of the merger remnant must be too small to open a gap in the disk. If it does (eq.s 2.14 and 2.16), it is a Type II migrator and it can no longer pair-up with other BHs in the migration trap.

If any of the conditions above is not met, we discard the merger remnant and don't consider it for future generations.

We follow the evolution of Nth-generation BHs with a procedure similar to the one outlined for first-generation BHs in section 3.2 and section 3.3. We report here any difference from the steps outlined in the previous sections.

Primary mass and spin In hierarchical merging, the remnant of an (N-1)th-generation merger acts as the primary BH for the Nth-generation. So the primary mass and spin are simply set as the remnant mass and spin of the previous generation, computed according to [41].

Primary initial position Similarly, the primary initial position is set as the position of the merger remnant of the previous generation, set in eq. 3.34. This value is used to compute the pairing time in eq. 2.17, where for the Nth generation

$$t_{\rm in}^{(N)} = t_{\rm merg}^{(N-1)} \tag{3.35}$$

Secondary mass The way we set the secondary mass for later generations is quite different than what it's been done for the first generation.
In our model, we allow for Ng-Mg mergers, but we always assume M ≤ N. So, first of all, we need to establish the generation M of the secondary.

Following [89], we assume that the probability that a given merger generation M is chosen for the companion is proportional to the number of 1g BHs required to synthesize it:

$$p\,(M) \propto 2^{-(M-1)} \tag{3.36}$$

We sample M from the probability distribution in eq. 3.36 using inverse random sampling.

Once we know the generation of the secondary, we can determine its mass.
First of all, we generate a 1g seed determining its mass and spin as in section 3.2. Then we let it go through a certain number of ng-1g mergers[1] until it creates an Mg remnant.
At each step, the primary will be the remnant of a previous merger event on the seed and as usual its mass m_1 and spin will be computed according to [41]. The secondary mass instead will be sampled from

$$p(m_2|m_1) \propto (m_1 + m_2^{\rm max})^4 \tag{3.37}$$

between $m_2^{\rm min} = 3M_\odot$ and $m_2^{\rm max}$.
Here $m_2^{\rm max}$ is determined based on the value of m_1. In particular, we care about whether m_1 is larger that the maximum mass $m_1^{\rm max}$ of a 1g BH in the input sample coming from the population synthesis code MOBSE [28, 27, 29, 30].

$$\text{if} \quad m_1 \leq m_1^{\rm max} \quad \Rightarrow \quad m_2^{\rm max} = m_1 \tag{3.38}$$

$$\text{else} \quad \Rightarrow \quad m_2^{\rm max} = m_1^{\rm max} \tag{3.39}$$

This is a modification of eq. 3.26.
With this choice, if the primary has mass compatible with a 1g BH ($m_1 \leq m_1^{\rm max}$) we sample the secondary as seen previously in eq. 3.26 (model from [69]).
Otherwise, if the primary is a higher-generation BH ($m_1 > m_1^{\rm max}$), we keep the same analytical form as in eq. 3.26 but we force the secondary to have mass compatible with a 1g BH ($m_2 \leq m_1^{\rm max}$).

This calculation is quite fast because we only compute the remnant mass and spin at each step, neglecting the merger times.

[1] We choose to consider only ng-1g mergers rather than ng-mg for simplicity.

Remnant generation As per the definition of generation number stated in section 2.8, the generation of the remnant is

$$N' = N + M \tag{3.40}$$

since N' stellar BHs were required for its assembly.

We keep iterating the procedure of section 3.4 until all BHs have failed to meet at least one of the requirements (i)-(iii) above.

Chapter 4

Results

In this Chapter we describe the results of our set of simulations.
We discuss the timescale of the evolution of the BH population, the mass and spin of the remnants, and the number of generations of hierarchical merging that BHs are able to achieve.

4.1 Description of runs

Each of our simulations is characterized by four properties:

1. The disk's lifetime

   We consider a disk to be either short-lived, with a lifetime of 5 Myr, or long-lived, with a lifetime of 10 Myr. We label them with the letter 's' or 'L' respectively.

2. The disk's density

   We have seen in subsection 2.2.2 and subsection 3.1.2 that the surface density of the disk varies as

   $$\Sigma_g\left(R, M_{\rm SMBH}\right) = \Sigma_0 \left(\frac{M_{\rm SMBH}}{10^8 M_\odot}\right)^{4/5} \begin{cases} \left(\dfrac{R}{10^3\, R_g}\right)^{0.8} & R \leq 10^3\, R_g \\ \left(\dfrac{R}{10^3\, R_g}\right)^{-1.49} & R > 10^3\, R_g \end{cases} \tag{4.1}$$

   We consider three scenarios: medium surface density with $\Sigma_0 = 7.91 \times 10^5\,\mathrm{g/cm^2}$, high surface density with $\Sigma_0 = 7.91 \times 10^6\,\mathrm{g/cm^2}$ (we multiply by a factor 10), and low surface density with $\Sigma_0 = 7.91 \times 10^4\,\mathrm{g/cm^2}$ (we divide by a factor 10).

   We label high density models with a capital letter 'D', medium density with a lowercase 'd' and low density with an italic lowercase '*d*'.

3. The disk's thickness

Similarly, the aspect ratio varies as

$$h\left(R, M_{\rm SMBH}\right) = h_0 \left(\frac{M_{\rm SMBH}}{10^8 M_\odot}\right)^{-3/20} \begin{cases} \left(\dfrac{R}{10^3\, R_g}\right)^{-0.6} & R \leq 10^3\, R_g \\ \left(\dfrac{R}{10^3\, R_g}\right)^{0.5} & R > 10^3\, R_g \end{cases} \tag{4.2}$$

We consider the thick scenario with $h_0 = 7.59 \times 10^{-3}$ and the thin scenario with $h_0 = 7.59 \times 10^{-4}$ (we divide by a factor 10).

We label thick disks with a capital 'T' and thin disks with an lowercase 't'.

4. The SMBH population

   In the runs labeled by '50' we extract the SMBH mass by the TNG50 population (Figure 3.1a), whereas in those labeled by '300' we extract it by the TNG300 population (Figure 3.2a).

Each run will be referred to with an acronym. For example, the high-density, long-lived and thick disk for the TNG50 population of SMBHs is called DLT_50, while the low-density, short-lived and thin one for the TNG300 population is *d*st_300.
See Table 4.1 for a summary of the details of each run.

We note that, with these definitions, the SG model described in subsection 2.2.2 is thick and medium-density. These terms are to be considered as comparative, since all disk models considered in this work are geometrically thin and have a relatively large gas density.

In each run, we set up the SMBH mass and corresponding AGN environment parameters as described in section 3.1. Then we fix the properties of 10^4 BHs as outlined in section 3.2. Each of these BHs evolves independently in a separate system. We repeat the procedure four times with four different SMBH masses, so the outputs are the combined results from four AGNs with different properties. The properties of each NSC are in Table 4.3 at the end of this Chapter.
Considering multiple SMBH masses as well as multiple initial BH masses and positions increases the statistical significance of each run.

4.2 Timescale analysis

We remind here that the most relevant physical timescales in our model are the damping and migration timescales (eq.s 2.12 and 2.13) after which a prograde BH reaches the migration trap, as well as the delay timescale $t_{\rm del}$ between BBH pair-up and merger (i.e. the integration time of $\dot{a}_{\rm gas}$ in eq. 2.26).
For convenience, we write here the expressions for the damping and migration timescales:[1]

$$t_{\rm damp} \simeq \frac{M_{\rm SMBH}^2\, h^4}{m\, \Sigma_g\, R^2\, \Omega} \qquad\qquad t_{\rm migr} = \frac{M_{\rm SMBH}^2\, h^2}{m\, \Sigma_g\, R^2\, \Omega}$$

[1]Here we write $t_{\rm damp}$ in terms of the radius R rather than the semimajor axis A because, as stated in section 3.2, they are equivalent in our model.

Run	Model name	Illustris version	Surface density	Lifetime	Thickness
1	**DLt_50**	TNG50	High	Long	Thin
2	**DLT_50**	TNG50	High	Long	Thick
3	**Dst_50**	TNG50	High	Short	Thin
4	**DsT_50**	TNG50	High	Short	Thick
5	**dLt_50**	TNG50	Medium	Long	Thin
6	**dLT_50**	TNG50	Medium	Long	Thick
7	**dst_50**	TNG50	Medium	Short	Thin
8	**dsT_50**	TNG50	Medium	Short	Thick
9	***d*Lt_50**	TNG50	Low	Long	Thin
10	***d*LT_50**	TNG50	Low	Long	Thick
11	***d*st_50**	TNG50	Low	Short	Thin
12	***d*sT_50**	TNG50	Low	Short	Thick
13	**DLt_300**	TNG300	High	Long	Thin
14	**DLT_300**	TNG300	High	Long	Thick
15	**Dst_300**	TNG300	High	Short	Thin
16	**DsT_300**	TNG300	High	Short	Thick
17	**dLt_300**	TNG300	Medium	Long	Thin
18	**dLT_300**	TNG300	Medium	Long	Thick
19	**dst_300**	TNG300	Medium	Short	Thin
20	**dsT_300**	TNG300	Medium	Short	Thick
21	***d*Lt_300**	TNG300	Low	Long	Thin
22	***d*LT_300**	TNG300	Low	Long	Thick
23	***d*st_300**	TNG300	Low	Short	Thin
24	***d*sT_300**	TNG300	Low	Short	Thick

Table 4.1: Description of runs. The horizontal line separates the TNG50 and the TNG300 runs.

As we have discussed in section 2.6, gravitational wave emission only becomes relevant once the semi-major axis is small, and it quickly drives the binary to merge. Therefore, the delay timescale is mostly governed by gas hardening and it can be roughly approximated as

$$t_{\rm del} \simeq \frac{a}{\dot{a}_{\rm gas}} = -\frac{\mu\,\Omega_b}{24\pi\,\alpha\,c_s^2\,\Sigma_g\,(1+e)^2} \tag{4.3}$$

In Figure 4.1 we show the timescales for the first generation of TNG50 runs. The timescale ranges for TNG300 runs are analogous, so they are not shown. We see that there is a large variability both in $t_{\rm damp}$ and in $t_{\rm migr}$, which is representative of the large variability in the SMBH mass (randomly sampled from either eq. 3.1 or eq. 3.2) and the initial position R (randomly sampled between $R_{\rm min}$ and $R_{\rm max}$, eq.s 3.8 and 3.9).
In each run, only a fraction of first generation BHs has both damping and migration timescales shorter than the disk's lifetime. We call this fraction $f_{\rm trap}$ and we show its value

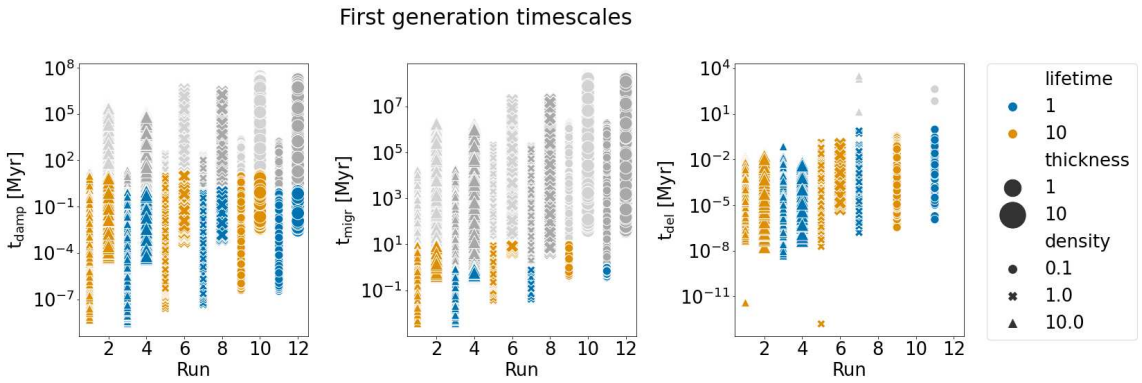


Figure 4.1: Damping, migration and delay timescales for the first generation of all runs. As indicated in the legend, circles are for low density, crosses are for medium densities and triangles for high densities; small sizes are for thin disks and large sizes for thick disks; blue is for short-lived disks and orange for long-lived disks. Gray color marks that the timescale is longer that the disk's lifetime.

for all runs in Table 4.2. BHs that do not respect this condition are shown in gray in Figure 4.1.
In particular, $f_{\rm trap}$ is zero for runs models dsT, dLT and dsT for both TNG50 and TNG300. Therefore, these types of disks are not able to produce any BBH pair-up.
For models DsT, dLT and dst, $f_{\rm trap}$ is small but not null. In the following section we will see that these types of disks are able to form some BBH mergers, but they do not produce many hierarchical mergers generations.

	DLt	DLT	Dst	DsT	dLt	dLT	dst	dsT	dLt	dLT	dst	dsT
TNG50	0.49	0.17	0.34	6e-3	0.33	4e-3	0.17	0.0	0.17	0.0	6e-3	0.0
TNG300	0.48	0.14	0.32	3e-3	0.31	4e-3	0.16	0.0	0.15	0.0	3e-3	0.0

Table 4.2: Fraction $f_{\rm trap}$ of BHs that reach the trap within the disk's lifetime for each of the runs in Table 4.1.

In Figure 4.2 we show the average timescale for first generation BHs as a function of the SMBH mass. From the explicit expression of $t_{\rm damp}$ and $t_{\rm migr}$ we would expect both timescales to increase with the square of $M_{\rm SMBH}$. Instead from the figure we notice that, for a given disk density and thickness, all timescales decrease as a function of $M_{\rm SMBH}$.
This trend is easy to explain if we consider the non-explicit dependencies on $M_{\rm SMBH}$: the surface density, aspect ratio and sound speed of the disk depend on $M_{\rm SMBH}$ as in eq. 2.8, while the maximum radial extension of the disk scales with the square root of $M_{\rm SMBH}$ as in eq. 3.9. Using $R \sim R_{\rm max}$ as a rough approximation, the expected trends are

$$t_{\rm damp} \propto M_{\rm SMBH}^{-2/5}\,, \qquad t_{\rm migr} \propto M_{\rm SMBH}^{-1/10}\,, \qquad t_{\rm del} \propto M_{\rm SMBH}^{-19/5} \tag{4.4}$$

By a fit of the data for DLt_50 and Dst_50 models (shown in Figure 4.4a), we see that the damping and migration timescales decrease even more steeply than expected, while the

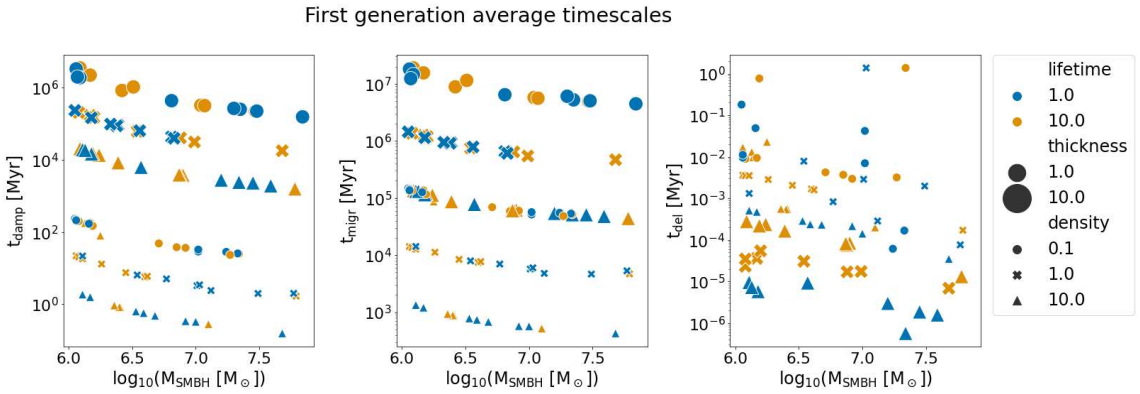


Figure 4.2: Average damping, migration and delay timescales for the first generation of all runs. The legend is analogous to that of Figure 4.1. DLt (Dst) is marked as orange (blue) small triangles, DLT (DsT) as orange (blue) large triangles, dLt (dst) as orange (blue) small crosses, dLT (dsT) as orange (blue) large crosses, dLt (dst) as orange (blue) small dots, dLT (dsT) as orange (blue) large dots.

delay timescale decrease less steeply than expected:

$$t_{\rm damp} \propto M_{\rm SMBH}^{-0.7}\,, \qquad t_{\rm migr} \propto M_{\rm SMBH}^{-0.2}\,, \qquad t_{\rm del} \propto M_{\rm SMBH}^{-1.2} \tag{4.5}$$

In particular, the behavior of $t_{\rm damp}$ and $t_{\rm migr}$ is irrespective of the disk's lifetime, while $t_{\rm del}$ is shorter for short-lived models that for their long-lived counterpart.
For the dense and thick case, the delay timescale decreases for longer lifetimes

From Figure 4.2 we clearly see that, when increasing the thickness by a factor of 10, damping timescales increase by 4 orders of magnitudes and migration timescales increase by 2 orders of magnitude, as expected from the explicit expressions. We would not expect any change in the delay timescale when changing the thickness, but instead it seems to decrease by up to 2 orders of magnitudes (it is roughly 2 for the dLt versus dLT case). This is merely an effect of the fact than many BHs in the thick models never reach the migration trap and eqs. 2.26 are not integrated for them: the lower the fraction of BHs that reach the migration trap for a given run, the more we are biased towards higher primary masses (an higher m gives lower $t_{\rm damp}$ and $t_{\rm migr}$) and therefore lower delay times (an higher $\Omega_b \propto (m_1+m_2)^{1/2}$ gives lower $t_{\rm del}$).
Similarly, the damping and migration timescales decrease by 1 order of magnitude when increasing the surface density by a factor of 10, as expected from the explicit expressions. The delay timescale is expected to decrease by a factor of 10, but this effect is polluted by the selection effects of BHs reaching (or not) the migration trap.

In Figure 4.3 we show the average timescales for all generations. When performing the average, we include only the BHs for which the pairing timescale $t_{\rm pair} = t_{\rm damp} + t_{\rm migr}$ is lower than the disk's lifetime: $t_{\rm pair} < \tau$.
Any deviation for the dependency on $M_{\rm SMBH}$ from what noted in Figure 4.2 is due to selection effects. We also notice that there is no substantial difference between the TNG50 and the TNG300 runs.

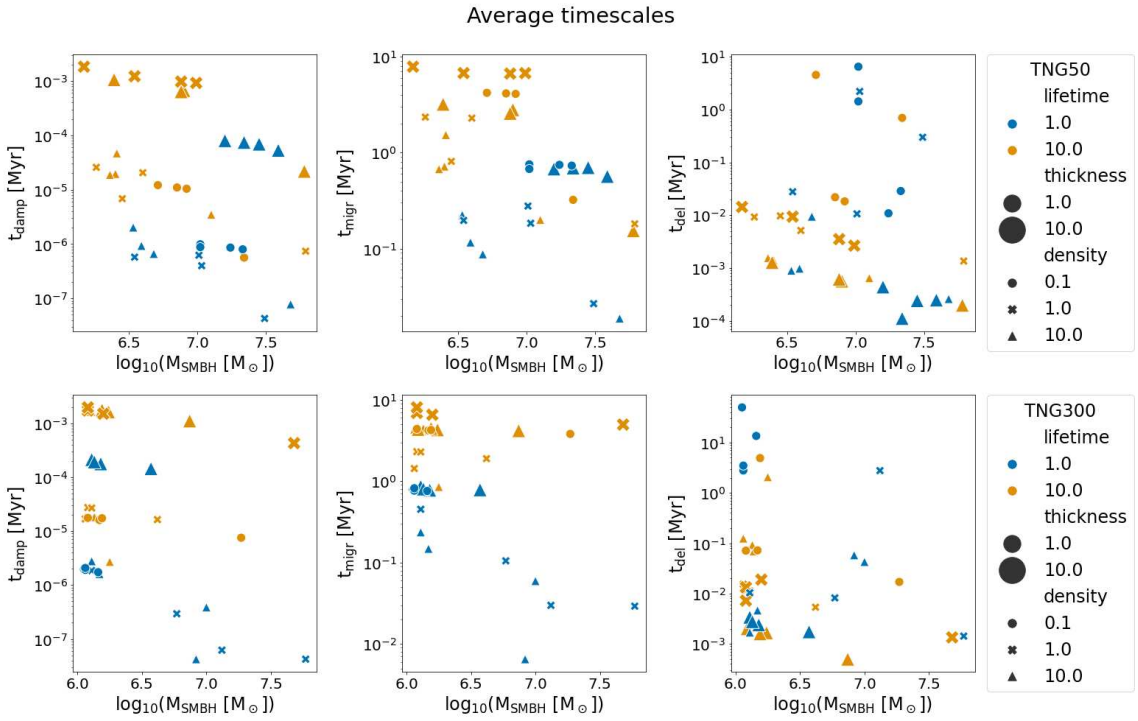
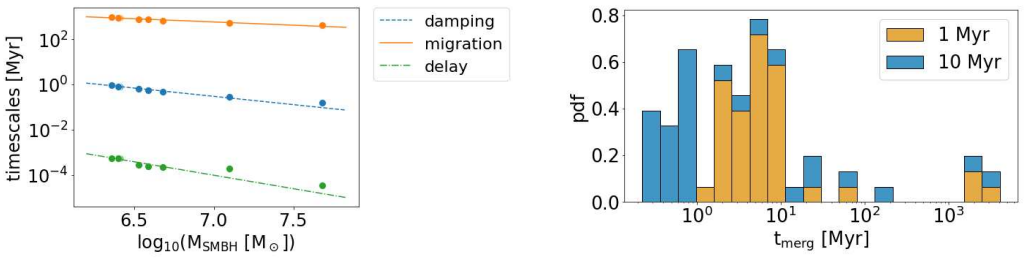


Figure 4.3: Damping, migration and delay timescales averaged over all generation for BHs that reach the migration trap within the disk's lifetime. The legend is analogous to that of Figure 4.2. The top row is for TNG50 models and the bottom row for TNG300.



(a) (b)

Figure 4.4: (a) Damping, migration and delay timescales for the first generation of models DLt_50, and Dst_50. Each timescale to SMBH mass relation is fitted with a power-law function; the best fit laws are shown as lines. (b) Merging time histogram for all simulations. The color is representative of the disk's lifetime: blue identifies models with lifetime of 1 Myr, while orange identifies those with lifetime of 10 Myr.

In Figure 4.4b we show a histogram of the merging timescale for the population of BHs that can experience a pair-up in the migration trap (i.e. that respect the condition $t_{\rm pair} < \tau$). Most mergers happen within the disk's lifetime but, even after the evaporation of the disk, BBHs can keep hardening due to GW emission only (eq.s 2.24 and 2.25) so that some of the mergers happen tens of millions of years after the end of the AGN phase.

In conclusion, for the BHs that are able to reach the migration trap and form BBHs, the typical damping timescale is of the order of $10^{-7} - 10^{-3}$ Myr, the typical migration timescale is of the order of $10^{-2} - 10^{1}$ Myr and the typical damping timescale is of the order of $10^{-4} - 10^{-1}$ Myr. Overall, the evolution is governed by the migration timescale.

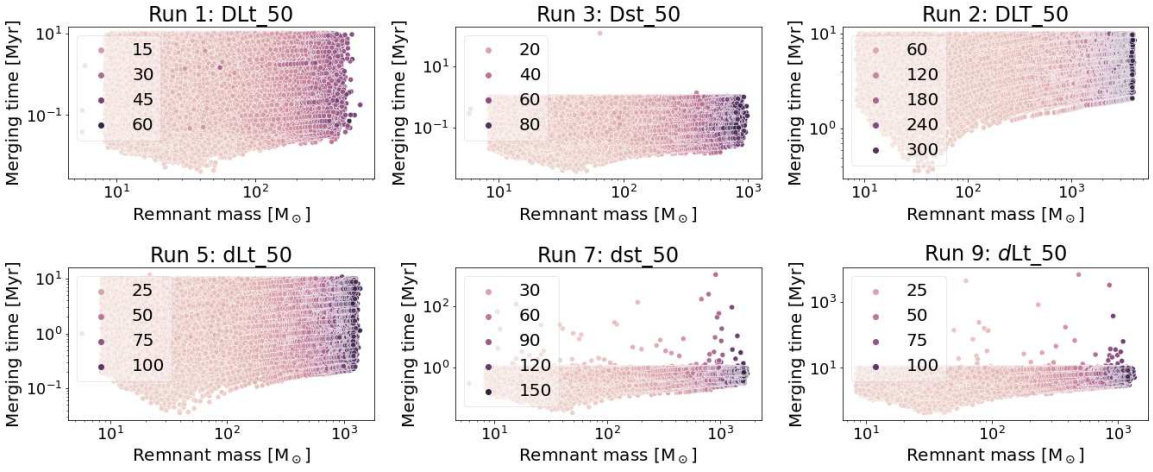
4.3 Mass function

As pointed out in the previous section, dsT, *d*LT and *d*sT models do not give rise to any merger events in the migration trap, so we will not include them in our discussion.

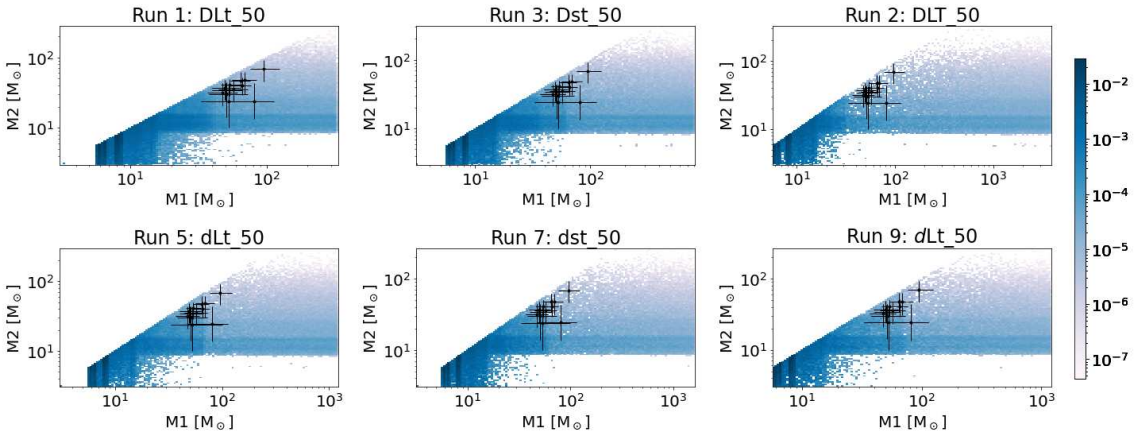
BHs in disks with high $f_{\rm trap}$ produce many hierarchical mergers events, while those in disks with $f_{\rm trap} \lesssim 10^{-2}$ can only produce few (if any) hierarchical merger generations before the condition on the maximum accretable mass (eq. 2.31) is met.
We are mostly interested on high-generation hierarchical mergers for the production of remnant BHs with masses in the upper gap and above, so we will only discuss the results from models with high $f_{\rm trap}$ (referring to Table 4.2, these are the DLt, DLT, Dst, dLt, dst and *d*Lt models). Instead, the results for DsT, dLT and *d*st disk models (i.e. those with low $f_{\rm trap}$) are shown in Appendix B.
Also, using data from Illustris TNG50 or TNG300 does not significantly affect the results. Therefore we only show plots for TNG50 runs in this section, while the analogous plots for TNG300 runs are in Appendix B.

In models where the fraction of BHs that reach the trap is $f_{\rm trap} \gtrsim 10^{-2}$, the dynamical evolution of the BH population gives rise to a large number of hierarchical mergers. As shown in Figure 4.5a, they are able to produce BH remnants with masses that span several orders of magnitude (roughly from tens of solar masses up to a few thousands of solar masses) via merger events that happen throughout the whole disk lifetime: the onset of the merger events happens sooner in models where the damping and migration timescales are shorter (such as DLt) and, in all cases, the production of BH remnants only stops when the disk evaporates. Remnants with formation timescales longer than the disk's lifetime, particularly numerous for dst and *d*Lt, come from BBHs that formed before the evaporation of the disk ($t_{\rm pair} \leq \tau$) and kept hardening in the absence of gas due to GW emission ($t_{\rm merg} > \tau$).
We show the remnant mass distributions and binary chirp mass distributions for the high-$f_{\rm trap}$ models in Figure 4.12a and 4.12b, respectively.

An important parameter for the characterization of a BBH is the ratio between the primary and secondary masses, called the mass ratio q.
The mass ratio of our simulated events is compatible with that of mass-gap BBHs observed by the LVK collaboration (listed in Table 1.1) in all models shown in Figure 4.5b, although in the DLT model they all fall in a lower-probability area.



(a)



(b)

Figure 4.5: (a) Remnant mass and merging timescale for all merger events of each disk type. The shade of the points is representative of their hierarchical merger generation: remnants with higher mass are produced via merger events of higher generation. (b) In blue: Density distribution of primary and secondary masses for all merger events of each disk type. In black: LVK merger events with primary in the upper mass gap and low FAR (in Table 1.1)

Therefore the observed merger events may have been produced in AGNs with the following properties: either medium to high density and thin aspect ratio (with either short or long lifetime), or low density, thin aspect ratio and long lifetime.
Geometrical thinness ($h_0 \lesssim 10^{-3}$) seems to be a necessary condition to reproduce the LVK merger events in Table 1.1. This is evident when comparing the DLt and DLT models: although they are both able to produce some hierarchical mergers, the fraction $f_{\rm trap}$ of BHs that are able to reach the migration trap is too low in the thick scenario, so that the maximum mass condition of eq. 2.31 is typically met for remnants with lower masses.

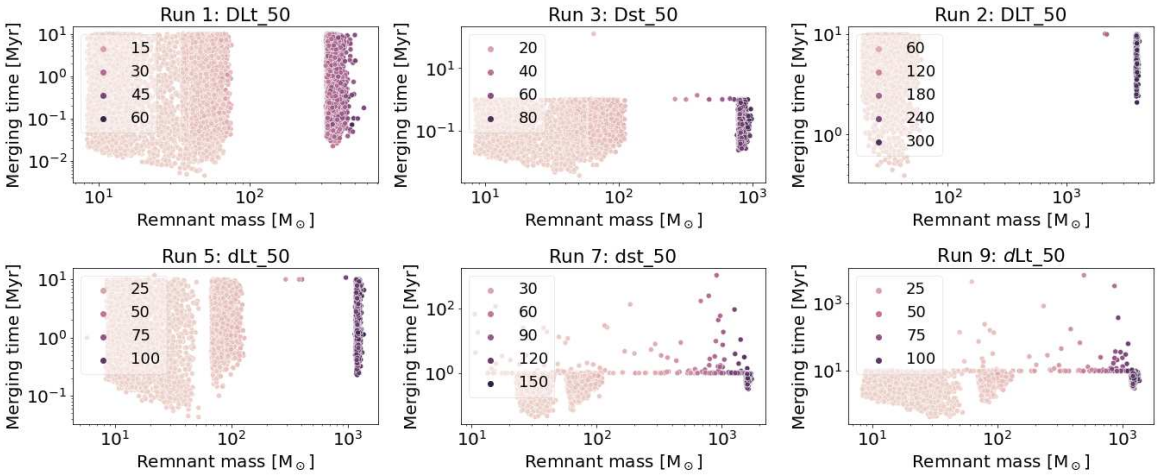


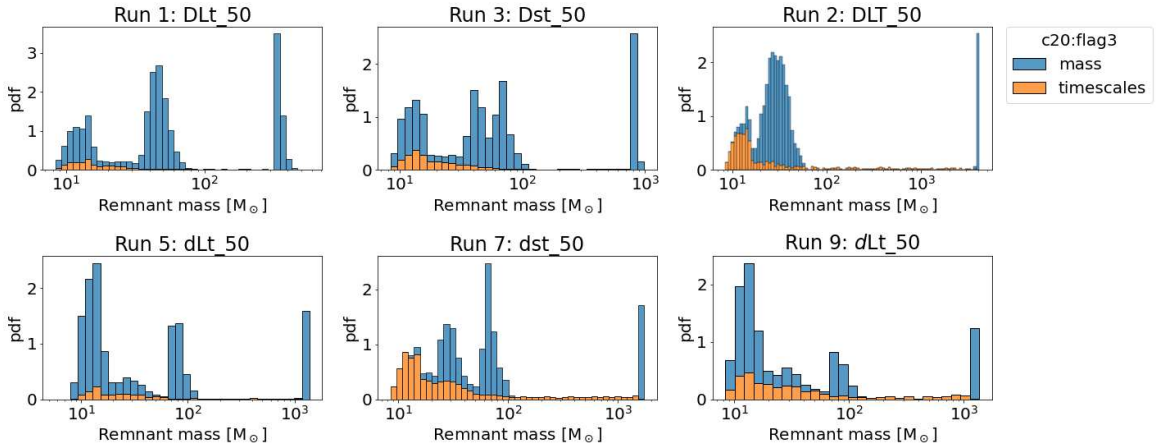
Figure 4.6: Mass and formation timescale for the final remnant population of each disk type. The shade of the points is representative of their hierarchical merger generation.

Simulated models, although they can predict merger events with similar parameters to those observed by the LVK instruments, also predict many other merger events with higher masses or lower mass ratio than what the observations suggest.
As discussed in subsection 1.2.2, events with higher total mass $m_{\rm tot} = m_1 + m_2$ are difficult to see in the LVK interferometers because the maximum frequency $f_{\rm max}$ of the GW signal, emitted at the end of the inspiral phase, is too low. Also, because of the non-completeness of template banks used for matched filtering, the LVK catalog is biased towards high mass ratio $q = m_2/m_1$.
Hence, it is not clear whether high-$m_{\rm tot}$ and low-q events are an artificial product of the simulations or whether they actually happen in AGNs, because these merger events are currently difficult to detect due to observational limits or template banks incompleteness. In one of the future developments of this thesis work, we will transform our models into mock observations, by evaluating their detection efficiency (as in e.g. [11, 54]). This will allow us to compare apples with apples and obtain more meaningful information.

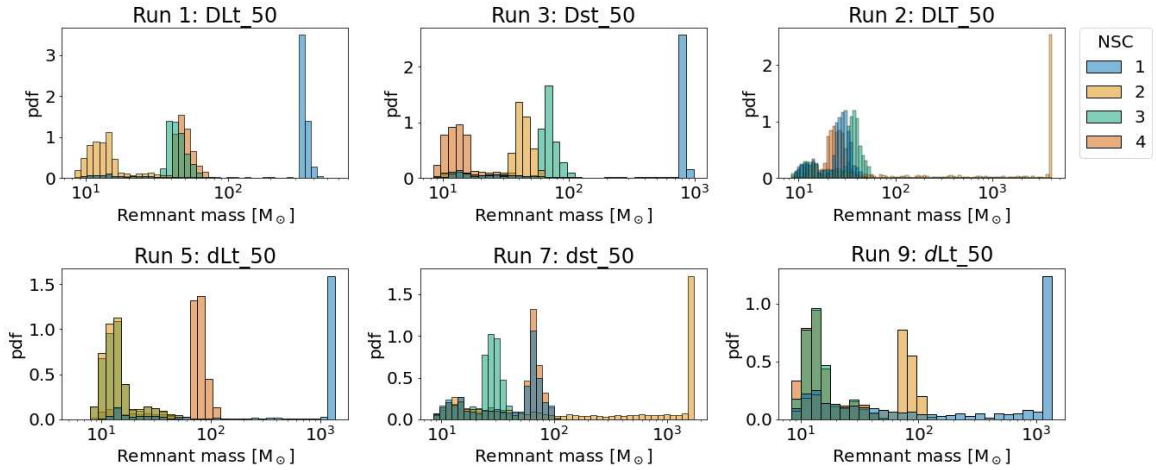
In Figure 4.6, we show the final remnant masses and formation timescales for the high $f_{\rm trap}$ models. By final remnants we mean the remnant of the last hierarchical merger generation of each seed BH, in other words the final remnants are those that are left around the SMBH after the evaporation of the disk.
Analyzing the final remnants is not useful to understand which LVK events are produced in AGNs, but it can help us understand how the inner NSC population changes after an accretion event onto the SMBH. This is relevant for two main reasons:

i) If a single SMBH can go through multiple AGN flare-ups, the final remnant population of one AGN event is the starting population of the next.

ii) When BHs orbit around a SMBH, they lose energy via GW emission (as in eq. 1.5) until, eventually, they merge with the SMBH. These binaries emit low GW frequencies (e.g. the maximum GW frequency for a $10^6 M_\odot$ SMBH and a $10^2 M_\odot$ IMBH, according



(a)



(b)

Figure 4.7: (a) Final remnants mass function for each disk type. The color indicates the cause of hierarchical mergers' cessation, as indicated in the legend. (b) Final remnant mass function for each NSC. The NSCs are indicated in the legend by a label between 1 and 4.

to eq. 1.11, is of a few 10^{-3}Hz), which will be observable with space-based detectors such as LISA [3].

The final remnant mass profiles are shown in Figure 4.7a, where we also indicate which of the reasons listed in section 2.8 caused hierarchical mergers to end.
Interestingly, there is no evidence of Type II migration nor of ejection from the cluster. We will discuss this in detail in section 4.5.

Each mass function is the superposition a smooth distribution (orange in Figure 4.7a) and some quite strong peaks (blue). The smooth distribution is formed by BHs that could not proceed in hierarchical mergers because pairing or merger timescales were too long with respect to the disk's lifetime (eq. 2.28), while the sharp peaks are all formed by BHs that

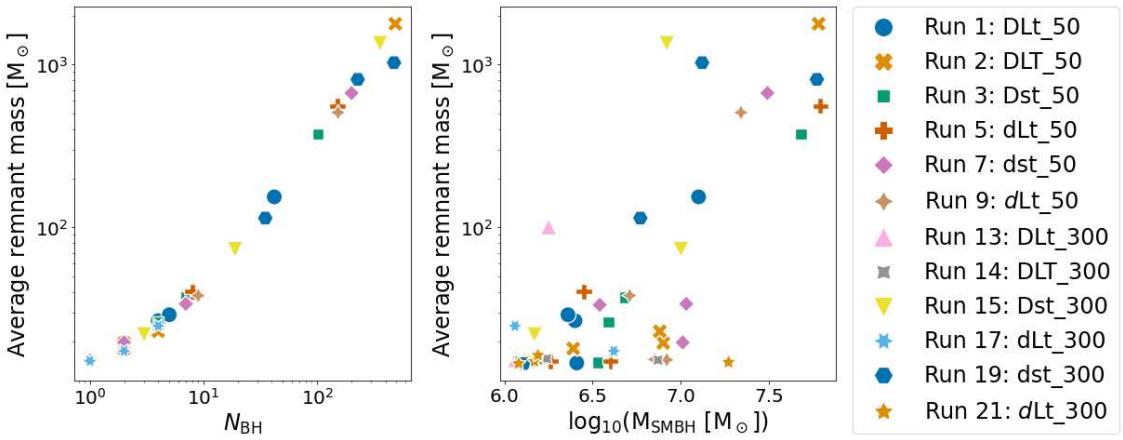


Figure 4.8: Average remnant masses for each run, shown as a function of the BH number (left-hand panel) and of the SMBH mass (right-hand panel). We only display models that produce significant hierarchical mergers for both TNG50 and TNG300. Runs that only produce a few merger generations display no correlation with neither of these quantities.

cleaned up the migration trap of all other BHs that migrated inside it (eq. 2.31). Indeed, although the condition on the maximum accretable mass is a requirement on the progenitor masses throughout the history of accretion onto the seed BH, it effectively constraints the final remnant mass in a narrow interval thus creating narrow peaks in the final mass functions.

The position of the mass peaks is sensitive on the maximum accretable mass $M_{\rm BH}^{\rm max}$. Referring to eqs. 3.17 and 3.18, there are a few relevant parameters: the BH fraction $f_{\rm BH}$, the NSC mass M_0 and the NSC effective radius r_h which collectively define the NSC properties, as well as the parameter $f_{\rm trap}$ which characterizes the dynamics of the BH population. Indeed, as shown in Figure 4.7b, the location of the peak is a unique feature of each NSC.

We compute the average remnant mass for each NSC and plot it in Figure 4.8 against the number of BHs in the migration trap and against the SMBH mass. We notice strong correlation with $N_{\rm BH}$ and mild correlation with the SMBH mass.
The strong correlation with $N_{\rm BH}$ is predictable from the observations above: the maximum accretable mass condition is crucial for the determination of the remnant mass distribution since the remnants can have masses up to $M_{\rm BH}^{\rm max}$ or slightly above. The correlation with the SMBH mass, instead, is mostly due to the shorter average timescales, as discussed in the previous section and seen in Figure 4.2.

As visible from the figures in Appendix B, the simulations performed assuming a TNG300 SMBH mass distribution behave qualitatively in the same way. The main difference is that, since the SMBH distribution is biased towards lower $M_{\rm SMBH}$, they tend to produce lower remnant masses.

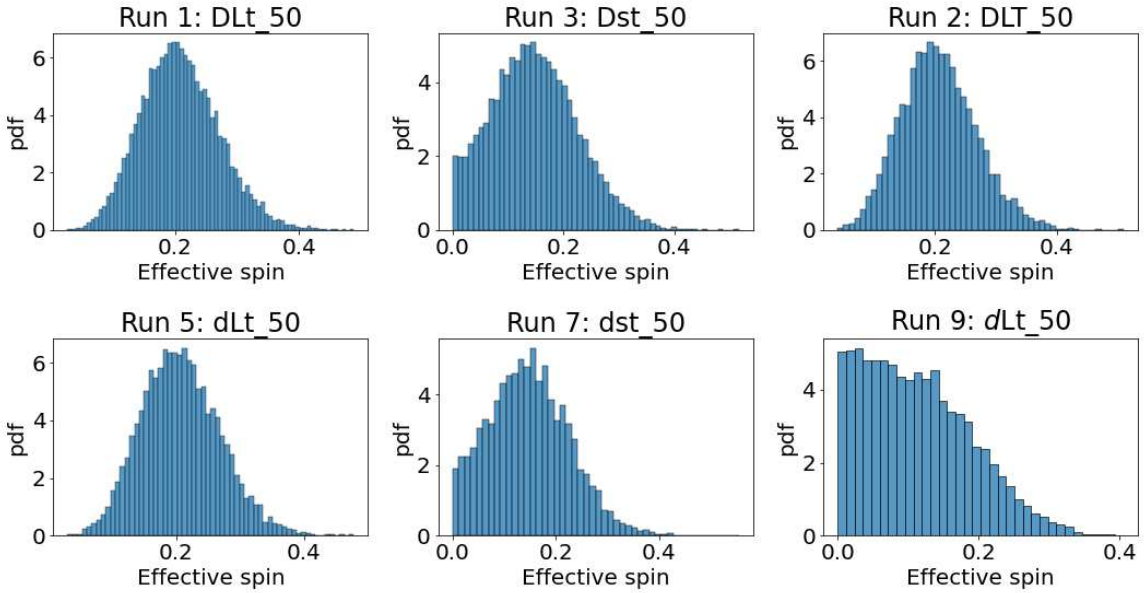


Figure 4.9: Spin distributions for models that produce significant hierarchical mergers.

4.4 Spin function

Gravitational waves' detection is sensitive to the effective spin, which is a particular combination of the primary and secondary spin as in eqs. 1.12 and 3.23, which physically conveys the BH spin components along the orbital angular momentum vector.

The spin alignment mechanism in AGN disks (see section 2.7) is sensitive to the volumetric gas density, which can be computed from the surface density Σ_g and the half-height of the disk H as

$$\rho_g = \frac{\Sigma_g}{2\,H} \tag{4.6}$$

For the choice of the surface density Σ_g and aspect ratio $h = H/R$ in our models (see section 4.1), the average[2] volumetric density values are reported in the following table:

	Thin	Thick
High density	$\overline{\rho_g} \simeq 10^{-9}\,\mathrm{g/cm^3}$	$\overline{\rho_g} \simeq 10^{-10}\,\mathrm{g/cm^3}$
Medium density	$\overline{\rho_g} \simeq 10^{-10}\,\mathrm{g/cm^3}$	$\overline{\rho_g} \simeq 10^{-11}\,\mathrm{g/cm^3}$
Low density	$\overline{\rho_g} \simeq 10^{-11}\,\mathrm{g/cm^3}$	$\overline{\rho_g} \simeq 10^{-12}\,\mathrm{g/cm^3}$

For the purposes of spin alignment, a disk is considered to have a low volumetric density when $\overline{\rho_g} \lesssim 10^{-11}\mathrm{g/cm^3}$ [87]. Hence, lower Σ_g does not automatically imply low ρ_g. Indeed, the medium-density model is considered dense in the thin case and dilute in the thick one.

[2] Averaged over the radial component.

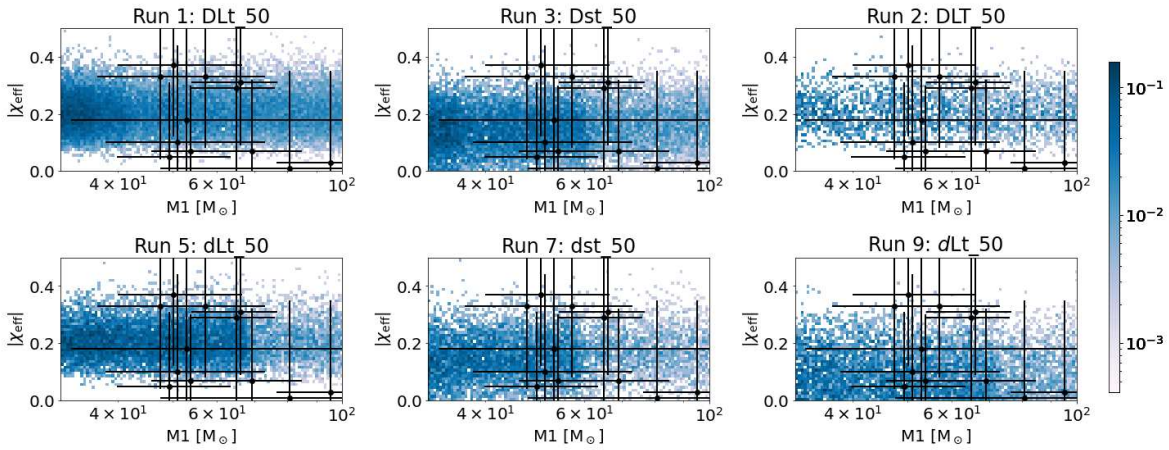


Figure 4.10: In blue: Density distribution of primary mass and effective spin for all merger event of each disk type. In black: LVK merger events with primary in the upper mass gap and low FAR (in Table 1.1).

We remind here that, in this work, the terminology 'dense' or 'dilute' refers respectively to models with high or low ρ_g. Instead, we use the terms 'high-', 'medium-' or 'low- density' to indicate models with different values of Σ_g.

We show the effective spin distribution for the merger remnants of the models producing hierarchical mergers in Figure 4.9. The DLt, DLT and dLt models are dense and long-lived, so they produce an effective spin distribution that in strongly peaked on $|\chi_{\rm eff}| \sim 0.2$; the Dst and dst are dense and short-lived, so their effective spin distribution still has a peak at $|\chi_{\rm eff}| \sim 0.2$ but it is somewhat more skewed towards $|\chi_{\rm eff}| = 0$. The only disk model that produces high-mass binaries with average effective spin $\chi_{\rm eff} = 0$ is the dLt, which is dilute and long-lived.

The only disk archetype which does not produce hierarchical mergers is the short-lived and dilute one (since the dsT, dst and dsT models only produce remnants up to the second hierarchical merger generation as in Figure B.4).

In conclusion, there is no univocal prediction for the spin distribution of high-mass BBHs in AGNs: depending on the disk's physical properties, the effective spin can have a peak on $|\chi_{\rm eff}| = 0.2$ or it can be peaked on $|\chi_{\rm eff}| = 0$.

As shown in Figure 4.10, the effective spins of upper-mass gap BBH mergers detected by LVK are compatible with the effective spin distributions produced by the simulated mergers. Nonetheless, it is not possible to draw definitive conclusions by this qualitative analysis because of the large errorbars on the detected effective spins.

4.5 Causes of hierarchical merger cessation

As we have explained in section 2.8 and section 3.4, hierarchical mergers stop for one of these four conditions:

i) Timescales: the pairing of merging timescale is longer than the disk's lifetime, so by the time the final remnant is formed the disk does not exist anymore (eq. 2.28).

ii) Ejection: the remnant velocity is larger than the escape velocity (eq. 2.29).

iii) Mass: the sum of the seed and secondary masses throughout the whole merger history is larger than the maximum mass which can be accreted in the migration trap (eq. 2.31).

iv) Type II migration: the remnant creates a gap in the disk, so it does not migrate in the migration trap (eq.s 2.14 and 2.16).

As seen in Figure 4.7a, the only two conditions that are relevant in the evolution of the BH population are (i) and (iii), i.e. the requirements on timescales and mass.

There is no merger remnant massive enough to open a gap in the disk in any of our simulations. Therefore, for the physical properties we assumed, Type II migration is not feasible. Although, if we were to consider disk models with lower viscosity or lower thickness, this result could change.
The impossibility of Type II migration is also a direct consequence of the way we imposed the condition on the maximum accretable mass (eq.s 2.31 and 3.18): if we were to relax this condition, for instance considering NSCs of larger masses or lower effective radii or assuming a larger BH fraction, we would allow for the formation of merger remnants of higher mass which might be able to undertake Type II migration.

Similarly, there are no ejections: the relativistic kick velocity is much smaller than the Keplerian velocity, so the kick only slightly perturbs the original orbit. For example, considering a SMBH with mass $10^7 M_\odot$ in a NSC with mass $10^7 M_\odot$ and effective radius 2 pc, the Keplerian velocity of an object in the migration trap is roughly $v_\mathrm{K} \simeq 10^4$ km/s. Instead, the kick velocity received by the merger remnant of an equal-mass BBH with $m_1 = m_2 = 50\, M_\odot$ is roughly 100 km/s.
The gravitational potential well of the SMBH is so deep that no remnant BH is able to escape from it.

Checking for condition (ii) may therefore seem unnecessary, and indeed in many previous works (e.g. [61]) the possibility of ejection is discarded entirely. Nonetheless, if we were to change some assumptions it might become relevant.
Changing the assumptions on the SMBH mass would not have significant effects: referring to eq.s 2.30 and 3.10 and considering that $M_\mathrm{NSC}(R_\mathrm{trap}) \ll M_\mathrm{NSC} \sim M_\mathrm{SMBH}$ since $R_\mathrm{trap} \ll r_h$, it is easy to see that $v_\mathrm{esc}(R_\mathrm{trap})$ is approximately constant for different SMBH masses. Instead, if we were to relax the assumption that BBH mergers can only happen in the migration trap, the escape velocity could be lower and condition (ii) might be influential. As an extreme example, for a merger event at the maximum radius R_max (eq. 3.9) of a disk with $M_\mathrm{SMBH} = 10^7 M_\odot$, the escape velocity would be of merely 500 km/s which is comparable to the typical kick velocity, so ejections would be possible.

Therefore, when considering only mergers in the migration trap and with our assumptions on the physical properties of the disk, the only relevant constraints are conditions (i) and (iii). Both of these conditions have hidden dependencies on the SMBH mass: as seen in section 4.2, higher M_SMBH yield lower pairing timescales and make it slightly harder to meet

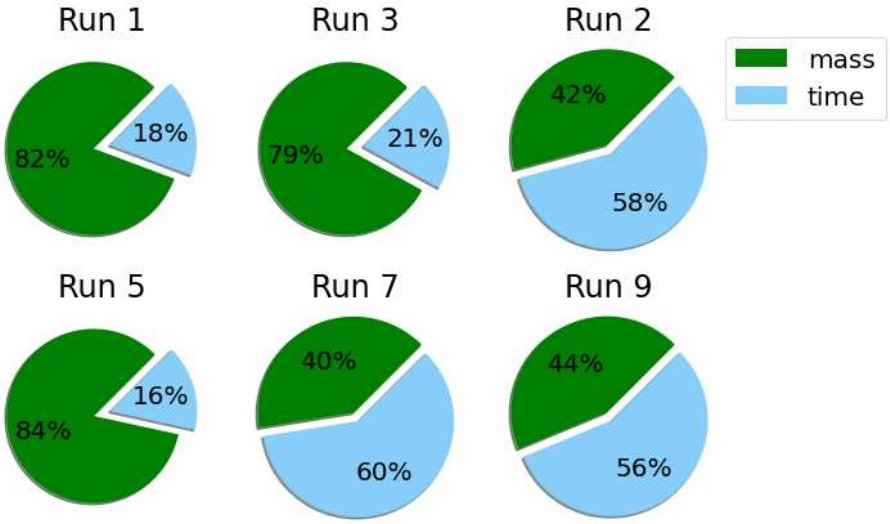


Figure 4.11: Reasons why BHs stop doing hierarchical mergers. Green is rejection for the maximum-mass condition (eq. 2.31), light blue is rejection for the timescale condition (eq. 2.28). We only show runs that produce significant hierarchical mergers. In all other runs, the timescale condition of the predominant reason for mergers cessation.

the maximum timescale condition (eq. 2.28); similarly, higher $M_{\rm SMBH}$ correlate with higher $M_{\rm NSC}$ (Figure 3.3a) which in turn induces a higher maximum mass (eq.s 3.17 and 3.18).

In order to compare the results of different disk models, we perform some runs with the properties shown in Table 4.1 and a fixed SMBH mass of $10^7 M_\odot$. The fraction of BHs that stop hierarchical mergers for conditions (i) and (iii) are illustrated in Figure 4.11.
There are two possible behaviors: either the proportion between reaching the timescale limit or the mass limit is roughly 20:80 (for models DLt, Dst and dLt), or it is 60:40 (for models DLT, dst and dLt). This directly correlated with the fraction $f_{\rm trap}$ of BHs that is successfully able to complete migration (see Table 4.2): for models for which $f_{\rm trap} \gtrsim 30\%$, the average timescales are short and hierarchical mergers continue until the seed BH cleans up the whole migration trap (eq. 2.31); instead, in models with lower $f_{\rm trap}$, the timescales are shorter and condition (i) is typically met earlier than condition (iii) in the accretion history of a seed BH.

The condition on the maximum accretable mass, which was set as an upper limit, is relevant in all runs where BHs perform a significant number of hierarchical mergers and produces prominents peaks in the final remnant mass functions. One may wonder whether reaching this upper limit is realistic, or whether we are missing too much physics.
There are two main ingredients that are missing from our model which might influence this result: three-body scattering and disk structure evolution.

The most frequent kind of multi-body interaction in dense environment is that between a binary and a single object [75]. After the interaction, if the binary is hardened to a smaller semi-major axis, both the binary and the third object receive a recoil kick [36] which is easy to compute with a simple energy balance calculation. For example, for a reference SMBH of

mass $10^7 M_\odot$ (for which, according to eq. 3.19, the velocity dispersion is $\sigma \sim 120\,\mathrm{km/s}$), assuming that all three objects have a mass of $10\,M_\odot$ and that the semi-major axis of the binary (initially equal to 1 AU) is halved, the binary receives a kick of $\sim 50\,\mathrm{km/s}$ while the single objects receives a kick of $\sim 100\,\mathrm{km/s}$.
The magnitude of this kick is comparable with the relativistic kick received by merger remnants. As we have discussed previously, kicks of this magnitude can cause ejections when the interactions happen at large radii (where the escape velocity, in eq. 2.30, is smaller) and they become more and more negligible as the radius decreases. Hence three body scattering can be effective at hindering migration but, once the objects reach the migration trap, three body scattering is completely negligible. So including this effect in the model would cause a decrease in f_{trap} and, consequently, a shift in the position of the peaks towards lower masses.

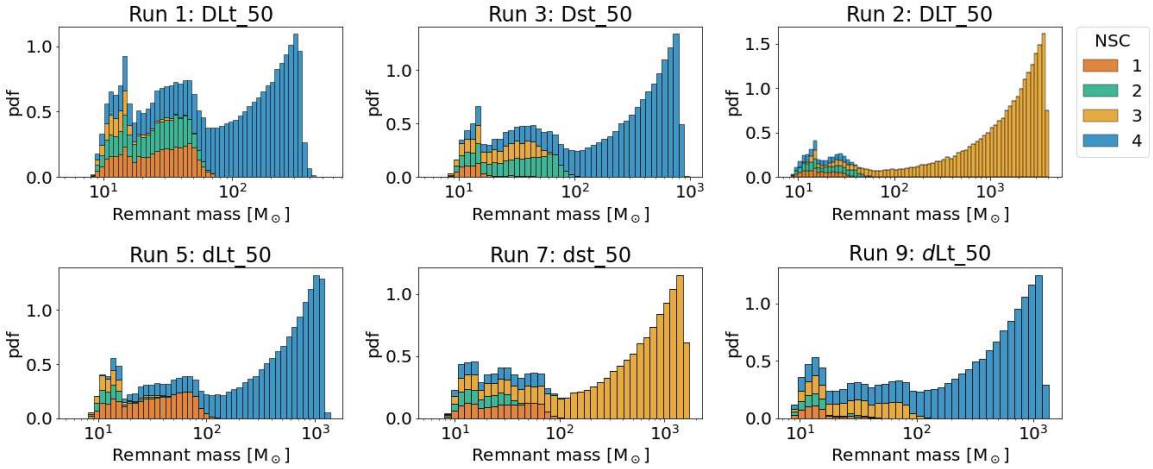
Furthermore, in our model we assumed the structure of the disk to be constant throughout its whole lifetime. Instead, in a more realistic model it would gradually change. For example, one would expect that the surface density gradually decreases due to SMBH feedback [21] and star formation in the outer skirts of the disk.
Decreasing the surface density Σ_g causes an increase in the damping and migration timescales (eq.s 2.12 and 2.13) as well as in the gas-hardening timescale (eq. 4.3), which all scale as $1/\Sigma_g$. So these processes become slower towards the end of the AGN phase, which means that we are currently underestimating the fraction of BHs that stop hierarchical mergers due to the conditions on timescales.
On the other hand, as visible in Figure 4.6, final merger remnants in the maximum-mass peaks are formed with timescales that can be significantly lower than the disk's lifetime. Therefore, although including the disk's evolution in our model can decrease the fraction of final merger remnants in these peaks, it should not erase them completely.

In conclusion, the effect of three-body scattering is to decrease the maximum accretable mass, while disk structure evolution makes it harder (but not impossible) for a seed BH to accrete the whole mass available.

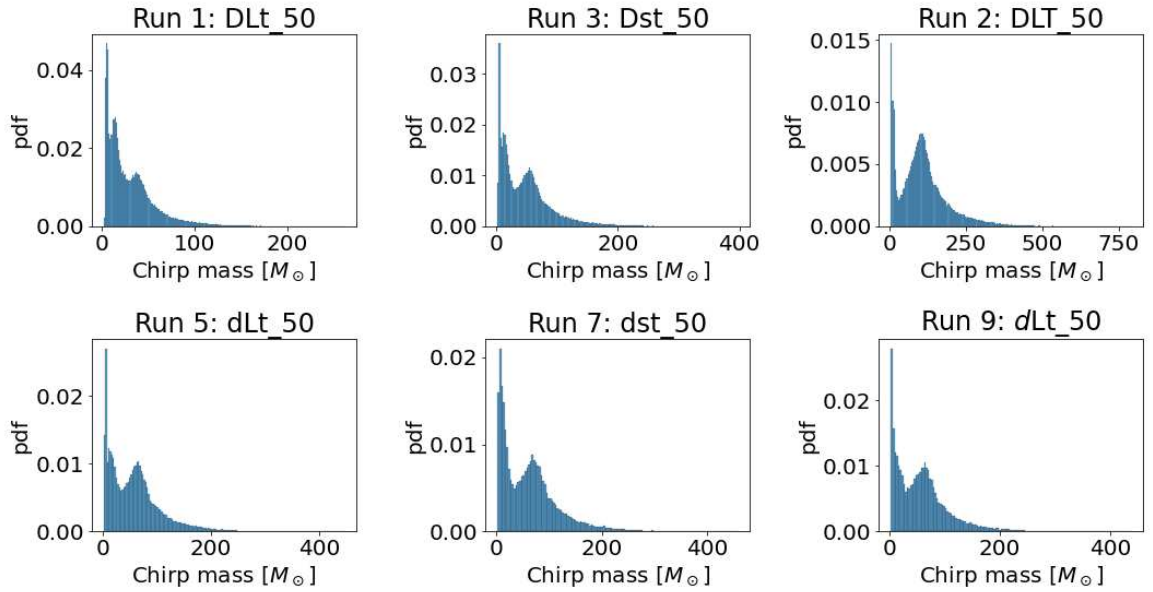
NSC index	Run	$\log M_{\mathrm{SMBH}}$ $[M_\odot]$	$\log M_{\mathrm{NSC}}$ $[M_\odot]$	r_h [pc]	Run	$\log M_{\mathrm{SMBH}}$ $[M_\odot]$	$\log M_{\mathrm{NSC}}$ $[M_\odot]$	r_h [pc]
1	1	7.1	7.0	7.08	13	6.14	6.19	5.8
2		6.41	6.05	10.46		6.25	6.42	10.09
3		6.4	6.35	3.5		6.13	5.26	3.09
4		6.36	6.66	3.81		6.06	6.8	7.87
1	2	6.9	6.32	4.25	14	6.24	7.2	15.07
2		7.78	7.4	5.15		6.09	6.1	3.43
3		6.88	7.1	8.71		6.87	5.73	5.42
4		6.39	6.51	3.04		6.19	5.6	6.56
1	3	7.68	7.18	9.01	15	7.0	6.07	3.01
2		6.59	7.17	6.53		6.92	6.85	2.02
3		6.68	6.49	3.96		6.11	5.47	4.64
4		6.53	6.96	9.92		6.17	6.61	2.84

1	4	7.59	7.33	3.86	16	6.18	6.31	5.01
2		7.45	7.21	6.36		6.11	5.49	3.99
3		7.34	6.85	4.16		6.13	6.07	8.73
4		7.2	7.1	3.54		6.57	5.85	7.34
1	5	7.79	7.21	10.37	17	6.06	6.92	3.22
2		6.6	6.67	8.17		6.08	6.81	4.72
3		6.26	5.84	6.93		6.11	6.05	4.58
4		6.45	7.02	4.26		6.62	6.43	5.04
1	6	6.99	6.82	4.95	18	6.08	6.65	2.85
2		6.88	6.78	6.88		6.08	7.11	6.52
3		6.54	6.48	3.12		6.2	6.45	1.79
4		6.17	6.04	4.33		7.68	6.46	15.28
1	7	6.54	6.17	2.18	19	6.11	5.63	7.82
2		7.49	7.04	4.16		7.12	7.02	3.68
3		7.01	7.35	14.36		6.77	7.19	12.76
4		7.03	7.24	7.87		7.77	6.67	12.68
1	8	6.81	6.74	5.79	20	6.18	6.15	4.82
2		6.83	6.24	2.31		6.38	5.87	1.97
3		6.56	6.12	3.93		6.33	6.67	6.77
4		6.39	6.58	6.85		6.05	6.15	2.54
1	9	7.34	6.54	2.58	21	6.17	5.99	3.11
2		6.71	7.01	4.23		7.27	5.24	7.38
3		6.92	6.49	7.37		6.19	5.47	3.04
4		6.85	6.33	6.37		6.08	6.06	5.07
1	10	6.42	6.27	14.11	22	6.08	5.52	2.25
2		7.04	6.76	3.24		6.51	6.08	3.65
3		7.47	7.01	10.64		6.09	6.13	1.9
4		7.07	6.21	3.0		6.17	5.7	2.69
1	11	7.33	6.52	3.54	23	6.05	5.73	3.51
2		7.24	7.19	8.72		6.16	7.55	5.38
3		7.02	6.82	2.56		6.06	6.08	5.19
4		7.02	6.8	2.25		6.06	5.55	5.25
1	12	7.48	6.66	3.68	24	7.3	5.48	9.67
2		7.35	6.71	6.41		6.06	6.02	4.76
3		6.81	5.98	3.91		6.09	6.17	3.44
4		7.84	6.58	6.98		6.07	7.32	10.28

Table 4.3: Characterizing parameters for each simulated NSC. For each run, the NSCs are identified by an index between 1 and 4.



(a)



(b)

Figure 4.12: (a) Remnant mass distribution for all merger events of each disk type. The color of the bars indicates in which NSC the merger event happened. The NSC are indicated with a label between 1 and 4 as illustrated in Table 4.3. (b) Chirp mass function for all merger events of each disk type.

Chapter 5

Conclusions

Thanks to gravitational wave (GW) observations, we are now more knowledgeable than ever before in the field of black hole (BH) astrophysics. Nonetheless, some questions still remain unanswered. In particular, there is no unequivocal interpretation for binary black hole (BBH) mergers in which one, or both, of the components have mass in the pair-instability mass gap. The detection of such events has stirred up the scientific community: there have been substantial efforts to explain their formation either by reviewing the current models of stellar evolution or by exploring the possibility of dynamical formation of massive BHs via hierarchical mergers.
This thesis fits in the field of BH dynamical formation. Indeed, we investigated the AGN scenario for hierarchical BBH mergers. The AGN environment is peculiar because of the presence of a dense gaseous disk which strongly influences black hole evolution. Therefore, we focused on viscous effects caused by the interaction with gas.

We built a semi-analytical model, preferring it over more accurate N-body or hydro-dynamical simulations so that it would allow us to quickly explore the parameter space. We used a state-of-the-art model for the structure of an AGN disk, which features the presence of a migration trap. We borrowed models from planetary astrophysics to predict how efficiently a BH becomes embedded in the disk and migrates into the migration trap. We implemented a prescription for the pair-up of a BBH in the migration trap and we modeled its evolution accounting for gas hardening and GW emission. Then we evaluated whether the remnant from the BBH merger can return to the migration trap and pair-up again, giving rise to hierarchical mergers.

We performed 10^4 simulations of the evolution of a seed stellar BH in 96 different AGN systems of 12 different kinds. We analyzed the effects on the BBH population of the central SMBH mass, the AGN lifetime, the gas density and disk's thickness.
We find that larger SMBH masses, larger gas densities, lower thicknesses and longer lifetimes correlate with the production of BBHs with higher mass.
For 6 of our 12 disk models, the aftermath of dynamical evolution gives rise to a remnant population with masses up to a few $10^3 M_\odot$ in less than 10 Myr. Hence, AGN disks can be factories for the efficient production of upper mass-gap BHs and of intermediate-mass BHs.

The 6 disk models that produce high-mass BBHs are: disks with medium to high density [i.e., $\Sigma_0 = \Sigma_g \left(R = 10^3 R_g, M_{\rm SMBH} = 10^8 M_\odot\right) \sim 10^6 - 10^7 {\rm g/cm^2}$] and low thickness [i.e.,

$h_0 = h\left(R = 10^3 R_g, M_{\rm SMBH} = 10^8 M_\odot\right) \sim 10^{-3}]$ with any lifetime, disks with high density ($\Sigma_0 \sim 10^7 {\rm g/cm^2}$), high thickness ($h_0 \sim 10^{-2}$) and long lifetime ($\tau = 10\,{\rm Myr}$), and disks with low density ($\Sigma_0 \sim 10^5 {\rm g/cm^2}$), low thickness ($h_0 \sim 10^{-3}$) and long lifetime ($\tau = 10\,{\rm Myr}$).
The remnant mass functions are qualitatively similar in these 6 scenarios, as they are mostly influenced by the maximum accretable mass which accounts for the total mass of BHs that are able to reach the migration trap during the evolution of the system. In turn, the maximum accretable mass strongly depends on the SMBH mass and the physical characteristics of the nuclear star cluster such as its mass and effective radius. Indeed, every single nuclear star cluster has a unique remnant mass function.

Of these 6 disk models, those with medium to high density produce a population of BBHs with average effective spin $|\chi_{\rm eff}| \sim 0.2$ due to the efficient mechanism of spin alignment. On the other hand, disks with lower densities produce a population of BBHs with $|\chi_{\rm eff}| \sim 0$.

We added our AGN model to a pre-existing code called FASTCLUSTER [54, 55] which already included prescriptions for BH dynamics in globular, young and nuclear star clusters. Thanks to this code, we are now able compare catalogs of BBH mergers in different dynamical channels using the same numerical code and the same underlying initial conditions (e.g., the same initial BH mass function). This gets rid of any bias that might arise by using different numerical codes for different environments: any differences are not due to the numerical approach adopted, but to the intrinsic physical differences among channels.
The enhanced version of FASTCLUSTER, which now also includes AGNs among the possible dynamical environments, will be used in the near future to construct mock samples of GW for each environment. These samples will account for the LIGO–Virgo–KAGRA observational biases.
Therefore, by comparing our mock data with the observational data from the detectors, we will be able to infer the mixing fractions of each environment as in [11, 54]. In other words, we will be able to infer how many of the currently observed BH mergers are produced in AGNs.
Not only will this be crucial for the interpretation of the origin of GW transients, but it will also provide valuable information on the structure of AGN disks, which is currently largely uncertain. For instance, identifying a large fraction of high-mass BBHs produced in AGNs would confirm the existence of migration traps in the disk.
As a future perspective, we plan on making this model more realistic by accounting for the evolution of the disk's structure in time and for three-body interactions.

Bibliography

[1] Abbott, R., Abbott, T. D., Abraham, S., and Acernese, F. et al. Properties and astrophysical implications of the 150 $M_\odot$ binary black hole merger GW190521. *The Astrophysical Journal*, 900(1):L13, sep 2020.

[2] Abbott, R., Abbott, T. D., Acernese, F., and Ackley, K. et al. Search for intermediate-mass black hole binaries in the third observing run of Advanced LIGO and Advanced Virgo. *A&A*, 695(A84):25, 2022.

[3] Bailes, M., Berger, B. K., and Brady, P. R. et al. Gravitational-wave physics and astronomy in the 2020s and 2030s. *Nature Reviews Physics*, 3:344–366, 2021.

[4] Baldassare, V. F., Reines, A. E., Gallo, E., and Greene, J. E. A $\sim$50,000$M_\odot$ solar mass black hole in the nucleus of RGG 118. *The Astrophysical Journal*, 809(1):L14, aug 2015.

[5] Bartos, I., Kocsis, B., Haiman, Z., and Márka, S. Rapid and bright stellar-mass binary black hole mergers in active galactic nuclei. *The Astrophysical Journal*, 835(2):165, jan 2017.

[6] Baruteau, C., Crida, A., Paardekooper, S. J., Masset, F., Guilet, J., Bitsch, B., Nelson, R., Kley, W., and Papaloizou, J. Planet-Disk Interactions and Early Evolution of Planetary Systems. In *Protostars and Planets VI*, page 667, January 2014.

[7] Bellovary, J. M., Mac Low, M., McKernan, B., and Saavik Ford, K. E. Migration traps in disks around supermassive black holes. *The Astrophysical Journal*, 819(2):L17, mar 2016.

[8] Binney J. and Tremaine S. *Galactic Dynamics: Second Edition*. Priceton University Press, 2008.

[9] Bitsch, B. and Kley, W. Orbital evolution of eccentric planets in radiative discs. *A&A*, 523:A30, 2010.

[10] Bogdanović, T., Reynolds, C. S., and Coleman Miller, M. Alignment of the spins of supermassive black holes prior to coalescence. *The Astrophysical Journal*, 661(2):L147–L150, may 2007.

[11] Bouffanais, Y., Mapelli, M., Santoliquido, F., Giacobbo, N., Di Carlo, U. N., Rastello, S., Artale, M. C., and Iorio, G. New insights on binary black hole formation channels after GWTC-2: young star clusters versus isolated binaries. *Monthly Notices of the Royal Astronomical Society*, 507(4):5224–5235, 09 2021.

[12] Bryden, G., Chen, X., Lin, D. N. C., Nelson, R. P., and Papaloizou, J. C. B. Tidally induced gap formation in protostellar disks: Gap clearing and suppression of protoplanetary growth. *The Astrophysical Journal*, 514(1):344–367, mar 1999.

[13] Buikema, A. et al. Sensitivity and performance of the advanced ligo detectors in the third observing run. *American Physical Society*, 102:062003, Sep 2020.

[14] Chodos, A. November 27, 1783: John Michell anticipates black holes. *American Physical Society News*, 18(10), 2009.

[15] Choudhuri, A. R. *Astrophysics for Physicists*. Cambridge University Press, 2010.

[16] Coleman Miller, M. and Lauburg, V. M. Mergers of stellar-mass black holes in nuclear star clusters. *The Astrophysical Journal*, 692(1):917–923, feb 2009.

[17] Event Horizon Telescope Collaboration. First Sagittarius A* Event Horizon Telescope Results. IV. Variability, Morphology, and Black Hole Mass. *The Astrophysical Journal Letters*, 930(L15), 2022.

[18] Costa, G., Bressan, A., Mapelli, M., Marigo, P., Iorio, G., and Spera, M. Formation of GW190521 from stellar evolution: the impact of the hydrogen-rich envelope, dredge-up, and $^{12}C(\alpha, \gamma)^{16}O$ rate on the pair-instability black hole mass gap. *Monthly Notices of the Royal Astronomical Society*, 501(3):4514–4533, 12 2020.

[19] Cresswell, P., Dirksen, G., Kley, W., and Nelson, R. P. On the evolution of eccentric and inclined protoplanets embedded in protoplanetary disks. *A&A*, 473(1):329–342, 2007.

[20] Deme, B., Meiron, Y., and Kocsis, B. Intermediate-mass black holes' effects on compact object binaries. *The Astrophysical Journal*, 892(2):130, apr 2020.

[21] Ruszkowski, M. et al. Supermassive black hole feedback. *BAAS*, 2019.

[22] Event Horizon Telescope Collaboration. First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole. *The Astrophysical Journal Letters*, 875(1):L4, April 2019.

[23] Farmer, R., Renzo, M., de Mink, S. E., Marchant, P., and Justham, S. Mind the gap: The location of the lower edge of the pair-instability supernova black hole mass gap. *The Astrophysical Journal*, 887(1):53, dec 2019.

[24] Fragione, G. and Silk, J. Repeated mergers and ejection of black holes within nuclear star clusters. *Monthly Notices of the Royal Astronomical Society*, 498(4):4591–4604, 08 2020.

[25] Genzel, R. *Massive Black Holes: Evidence, Demographics and Cosmic Evolution*, pages 93–119. Springer International Publishing, Cham, 2021.

[26] Genzel, R. Nobel lecture: A forty-year journey. *Rev. Mod. Phys.*, 94:020501, Jun 2022.

[27] Giacobbo, N. and Mapelli, M. The progenitors of compact-object binaries: impact of metallicity, common envelope and natal kicks. *Monthly Notices of the Royal Astronomical Society*, 480(2):2011–2030, October 2018.

[28] Giacobbo, N. and Mapelli, M. The impact of electron-capture supernovae on merging double neutron stars. *Monthly Notices of the Royal Astronomical Society*, 482(2):2234–2243, January 2019.

[29] Giacobbo, N. and Mapelli, M. Revising Natal Kick Prescriptions in Population Synthesis Simulations. *The Astrophysical Journal*, 891(2):141, March 2020.

[30] Giacobbo, N., Mapelli, M., and Spera, M. Merging black hole binaries: the effects of progenitor's metallicity, mass-loss rate and Eddington factor. *Monthly Notices of the Royal Astronomical Society*, 474(3):2959–2974, March 2018.

[31] Gillessen, S., Plewa, P. M., Eisenhauer, F., Sari, R., Waisberg, I., Habibi, M., Pfuhl, O., George, E., Dexter, J., von Fellenberg, S., Ott, T., and Genzel, R. An Update on Monitoring Stellar Orbits in the Galactic Center. *The Astrophysical Journal*, 837(1):30, March 2017.

[32] Goodman, J. Self-gravity and quasi-stellar object discs. *Monthly Notices of the Royal Astronomical Society*, 339(4):937–948, 03 2003.

[33] Graham, A. W. and Spitler, L. R. Quantifying the coexistence of massive black holes and dense nuclear star clusters. *Monthly Notices of the Royal Astronomical Society*, 397(4):2148–2162, 08 2009.

[34] Gravity Collaboration. A geometric distance measurement to the Galactic center black hole with 0.3% uncertainty. *A&A*, 625:L10, May 2019.

[35] Hartle, J. B. *Gravity: An Introduction to Einstein's General Relativity: Pearson New International Edition*. Pearson Education Limited, 2013.

[36] Heggie, D. C. Binary Evolution in Stellar Dynamics. *Monthly Notices of the Royal Astronomical Society*, 173(3):729–787, 12 1975.

[37] Hirschi, R. Very Massive and Supermassive Stars: Evolution and Fate. In *Handbook of Supernovae*, page 567. Springer International Publishing, 2017.

[38] Horn, B., Lyra, W., Mac Low, M. M., and Sándor, Z. Orbital migration of interacting low-mass planets in evolutionary radiative turbolant models. *The Astrophysical Journal*, 750(1):34, apr 2012.

[39] Ishibashi, W. and Gröbner, M. Evolution of binary black holes in agn accretion discs: Disc-binary interaction and gravitational wave emission. *A&A*, 639:A108, 2020.

[40] Jeans, J. H. The origin of binary systems. *Monthly Notices of the Royal Astronomical Society*, 79:408, April 1919.

[41] Jiménez-Forteza, X., Keitel, D., Husa, S., Hannam, M., Khan, S., and Pürrer, M. Hierarchical data-driven approach to fitting numerical relativity data for nonprecessing binary black holes with an application to final spin and radiated energy. *Phys. Rev. D*, 95:064024, Mar 2017.

[42] Kepler, S. O., Kleinman, S. J., Nitta, A., Koester, D., Castanheira, B. G., Giovannini, O., Costa, A. F. M., and Althaus, L. White dwarf mass distribution in the SDSS. *Monthly Notices of the Royal Astronomical Society*, 375(4):1315–1324, 02 2007.

[43] Khrykin, I. S., Hennawi, J. F., Worseck, G., and Davies, F. B. The first measurement of the quasar lifetime distribution. *Monthly Notices of the Royal Astronomical Society*, 505(1):649–662, 05 2021.

[44] King, A. R., Pringle, J. E., and Livio, M. Accretion disc viscosity: how big is alpha? *Monthly Notices of the Royal Astronomical Society*, 376(4):1740–1746, 03 2007.

[45] Kocsis, B., Yunes, N., and Loeb, A. Observable signatures of extreme mass-ratio inspiral black hole binaries embedded in thin accretion disks. *Physical Review D*, 84:024032, Jul 2011.

[46] Kroupa, P. On the variation of the initial mass function. *Monthly Notices of the Royal Astronomical Society*, 322(2):231–246, Apr 2001.

[47] LIGO Scientific Collaboration and Virgo Collaboration. Binary Black Hole Mergers in the First Advanced LIGO Observing Run. *Physical Review X*, 6(4):041015, October 2016.

[48] LIGO Scientific Collaboration and Virgo Collaboration. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016.

[49] LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration. Gwtc-3: Compact binary coalescences observed by ligo and virgo during the second part of the third observing run, 2021.

[50] LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration. The population of merging compact binaries inferred using gravitational waves through GWTC-3. *arXiv e-prints*, page arXiv:2111.03634, November 2021.

[51] Lin, Y.-H., Hammer, H.-W., and Meißner, U.-G.S. New insights into the nucleon's electromagnetic structure. *Phys. Rev. Lett.*, 128:052002, Feb 2022.

[52] Maggiore, M. *Gravitational Waves: Volume 1: Theory and Experiments*. Gravitational Waves. Oxford University Press, 2008.

[53] Maggiore, M. *Gravitational Waves: Volume 2: Astrophysics and Cosmology*. Gravitational Waves. Oxford University Press, 2018.

[54] Mapelli, M., Bouffanais, Y., Santoliquido, F., Arca Sedda, M., and Artale, M. C. The cosmic evolution of binary black holes in young, globular, and nuclear star clusters: rates, masses, spins, and mixing fractions. *Monthly Notices of the Royal Astronomical Society*, 511(4):5797–5816, 02 2022.

[55] Mapelli, M., Dall'Amico, M., Bouffanais, Y., Giacobbo, N., Arca Sedda, M., Artale, M. C., Ballone, A., Di Carlo, U. N., Iorio, G., Santoliquido, F., and Torniamenti, S. Hierarchical black hole mergers in young, globular and nuclear star clusters: the effect of metallicity, spin and cluster properties. *Monthly Notices of the Royal Astronomical Society*, 505(1):339–358, May 2021.

[56] Mapelli, M. and Gualandris, A. Star formation and dynamics in the galactic centre. In *Astrophysical Black Holes*, pages 205–272. Springer International Publishing, 2016.

[57] McKernan, B., Ford, K. E. S., Callister, T., Farr, W. M., O'Shaughnessy, R., Smith, R., Thrane, E., and Vajpeyi, A. LIGO–Virgo correlations between mass ratio and effective inspiral spin: testing the active galactic nuclei channel. *Monthly Notices of the Royal Astronomical Society*, 514(3):3886–3893, 06 2022.

[58] McKernan, B., Ford, K. E. S., Kocsis, B., Lyra, W., and Winter, L. M. Intermediate-mass black holes in AGN discs – II. Model predictions and observational constraints. *Monthly Notices of the Royal Astronomical Society*, 441(1):900–909, 05 2014.

[59] McKernan, B., Ford, K. E. S., Lyra, W., and Perets, H. B. Intermediate mass black holes in AGN discs – I. Production and growth. *Monthly Notices of the Royal Astronomical Society*, 425(1):460–469, 09 2012.

[60] McKernan, B., Ford, K. E. S., O'Shaugnessy, R., and Wysocki, D. Monte Carlo simulations of black hole mergers in AGN discs: Low χ eff mergers and predictions for LIGO. *Monthly Notices of the Royal Astronomical Society*, 494(1):1203–1216, 04 2020.

[61] McKernan, B., Saavik Ford, K. E., Bellovary, J., Leigh, N. W. C., Haiman, Z., Kocsis, B., Lyra, W., Mac Low, M.-M., Metzger, B., O'Dowd, M., Endlich, S., and Rosen, D. J. Constraining stellar-mass black hole mergers in AGN disks detectable with LIGO. *The Astrophysical Journal*, 866(1):66, oct 2018.

[62] Merritt, D. *Dynamics and Evolution of Galactic Nuclei*. Princeton series in astrophysics. Princeton University Press, 2013.

[63] Merritt, D. and Ferrarese, L. The M-σ Relation for Supermassive Black Holes. *The Astrophysical Journal*, 547(1):140–145, jan 2001.

[64] Michell, J. On the means of discovering the distance, magnitude, &c. of the fixed stars. *Philosophical Transactions of the Royal Society*, 74:35–54, 1784. In a letter to Henry Cavendish.

[65] Mo, H., van den Bosch, F., and White, S. *Galaxy Formation and Evolution*. Cambridge University Press, 2010.

[66] Moran, J. M. The Black-Hole Accretion Disk in NGC 4258: One of Nature's Most Beautiful Dynamical Systems. In *Frontiers of Astrophysics: A Celebration of NRAO's 50th Anniversary*, volume 395 of *Astronomical Society of the Pacific Conference Series*, page 87, 2008.

[67] Nelson, D., Pillepich, A., Springel, V., Pakmor, R., Weinberger, R., Genel, S., Torrey, P., Vogelsberger, M., Marinacci, F., and Hernquist, L. First results from the TNG50 simulation: galactic outflows driven by supernovae and black hole feedback. *Monthly Notices of the Royal Astronomical Society*, 490(3):3234–3261, aug 2019.

[68] Neumayer, N., Seth, A., and Böker, T. Nuclear star clusters. *The Astronomy and Astrophysics Review*, 28(1), jul 2020.

[69] O'Leary, R. M., Meiron, Y., and Kocsis, B. Dynamical formation of black hole binaries in the first detected mergers by ligo. *The Astrophysical Journal*, 824(1):L12, jun 2016.

[70] Peters, P. C. Gravitational radiation and the motion of two point masses. *Phys. Rev.*, 136:B1224–B1232, Nov 1964.

[71] Pillepich, A., Nelson, D., Hernquist, L., Springel, V., Pakmor, R., Torrey, P., Weinberger, R., Genel, S., Naiman, J. P., Marinacci, F., and Vogelsberger, M. First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies. *Monthly Notices of the Royal Astronomical Society*, 475(1):648–675, dec 2017.

[72] Pillepich, A., Nelson, D., Springel, V., Pakmor, R., Torrey, P., Weinberger, R., Vogelsberger, M., Marinacci, F., Genel, S., van der Wel, A., and Hernquist, L. First results from the TNG50 simulation: the evolution of stellar and gaseous discs across cosmic time. *Monthly Notices of the Royal Astronomical Society*, 490(3):3196–3233, sep 2019.

[73] Plummer, H. C. On the problem of distribution in globular star clusters. *Monthly notices of the royal astronomical society*, 71:460–470, 1911.

[74] Salpeter, E. E. The Luminosity Function and Stellar Evolution. *The Astrophysical Journal*, 121:161, January 1955.

[75] Samsing, J., MacLeod, M., and Ramirez-Ruiz, E. Formation of tidal captures and gravitational wave inspirals in binary-single interactions. *The Astrophysical Journal*, 846(1):36, aug 2017.

[76] Scott, N. and Graham, A. W. Updated mass scaling relations for nuclear star clusters and a comparison to supermassive black holes. *The Astrophysical Journal*, 763(2):76, jan 2013.

[77] Seeds, M. and Backman, D. *Astronomy: The Solar System and Beyond*. Cengage Learning, 2009.

[78] Seyfert, C. K. Nuclear emission in spiral nebulae. *The Astrophysical Journal*, 97:28, 1943.

[79] Shakura, N. I. and Sunyaev, R. A. Black holes in binary systems. Observational appearance. *AAP*, 24:337–355, January 1973.

[80] Shen, Z. Q., Lo, K. Y., Liang, M.-C., Ho, P. T. P., and Zhao, J.-H. A size of 1 AU for the radio source Sgr A* at the centre of the Milky Way. *Nature*, 438:62–64, 2005.

[81] Sirko, E. and Goodman, J. Spectral energy distributions of marginally self-gravitating quasi-stellar object discs. *Monthly Notices of the Royal Astronomical Society*, 341(2):501–508, May 2003.

[82] Springel, V., Pakmor, R., Pillepich, A., Weinberger, R., Nelson, D., Hernquist, L., Vogelsberger, M., Genel, S., Torrey, P., Marinacci, F., and Naiman, J. First results from the IllustrisTNG simulations: matter and galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, 475(1):676–698, dec 2017.

[83] Stone, N. C., Küpper, A. H. W., and Ostriker, J. P. Formation of massive black holes in galactic nuclei: runaway tidal encounters. *Monthly Notices of the Royal Astronomical Society*, 467(4):4180–4199, 03 2017.

[84] Tagawa, H., Haiman, Z., Bartos, I., and Kocsis, B. Spin evolution of stellar-mass black hole binaries in active galactic nuclei. *The Astrophysical Journal*, 899(1):26, aug 2020.

[85] Tagawa, H., Haiman, Z., and Kocsis, B. Formation and evolution of compact-object binaries in AGN disks. *The Astrophysical Journal*, 898(1):25, jul 2020.

[86] Thompson, T. A., Quataert, E., and Murray, N. Radiation pressure–supported starburst disks and active galactic nucleus fueling. *The Astrophysical Journal*, 630(1):167–185, sep 2005.

[87] Vajpeyi, A., Thrane, E., Smith, R., McKernan, B., and Saavik Ford, K. E. Measuring the properties of active galactic nuclei disks with gravitational waves. *The Astrophysical Journal*, 931(2):82, may 2022.

[88] Yang, Y., Bartos, I., Haiman, Z., Kocsis, B., Márka, Z., Stone, N. C., and Márka, S. AGN disks harden the mass distribution of stellar-mass binary black hole mergers. *The Astrophysical Journal*, 876(2):122, may 2019.

[89] Zevin, M. and Holz, D. E. Avoiding a cluster catastrophe: Retention efficiency and the binary black hole mass spectrum, 2022.

[90] Ziosi, B. M., Mapelli, M., Branchesi, M., and Tormen, G. Dynamics of stellar black holes in young star clusters with different metallicities – II. Black hole–black hole binaries. *Monthly Notices of the Royal Astronomical Society*, 441(4):3703–3717, 06 2014.

Appendix A

Instantaneous pair-up in the migration trap: is it a realistic assumption?

The idea of an instantaneous pair-up (section 3.3) may turn up their noses to most physicists because, as we know, no process in physics is instantaneous. In this appendix, we will explain why we believe that this is a reasonable approximation.

As stated in section 2.3 and section 2.4, all prograde objects that cross the disk during their orbits will migrate and eventually reach migration traps. The timescale for this phenomenon scales as the inverse of the object's mass (eq. 2.12 and 2.13). Therefore, if there is a large number of objects interacting with the disk, after a certain time there will be a large mass density in the migration trap.

We will verify that the density in the migration trap $\rho_{\rm trap}$ is indeed quite high. So we expect that, when a BH enters into the migration trap, it will form a binary with dynamical processes in a negligible time.
We do not know the timescale for dynamical formation in a disk geometry, but we can guess its order of magnitude by using the timescale for dynamical formation of a binary via three-body encounters in a spherical geometry.
For an object of mass m_1 the three-body binary formation timescale is [24]

$$t_{\rm 3bb} = 125\,{\rm Myr}\left(\frac{10^6 {\rm M}_\odot/{\rm pc}^3}{\rho_{\rm trap}}\right)^2\left(\zeta^{-1}\frac{\sigma_{1D}}{30\,{\rm km/s}}\right)^9\left(\frac{20\,{\rm M}_\odot}{m_1}\right)^5 \tag{A.1}$$

where $\zeta \leq 1$ accounts for deviations from equipartition of the BH subsystem (we assume perfect equipartition, $\zeta = 1$) and $\sigma_{1D} = \sigma/\sqrt{3}$ is the one-dimensional velocity dispersion (assuming an isotropic distribution of stellar velocities).

The problem is: how do we know that the companion is another BH and that the newly-born binary is indeed a BBH? Well, we do not.

As seen in subsection 2.2.2, the BH fraction is expected to be only about 4%, so with a probability of 96% the companion will not be a BH.

Although forming a BBH directly is quite rare, according to [90] BBHs form efficiently in dense environments because of dynamical exchanges: a BH can replace a lighter star into a BH-star binary.

Having initially a binary system formed by a BH of mass m_1 and an object of mass m_*, the object is replaced by a BH of mass m_2 on a timescale [16]

$$t_{\rm exch} = 3\,{\rm Gyr}\left(\frac{0.01}{f_{\rm bin}}\right)\left(\frac{10^6 {\rm M}_\odot/{\rm pc}^3}{\rho_{\rm trap}}\right)\left(\frac{\sigma}{50\,{\rm km/s}}\right)\left(\frac{12\,{\rm M}_\odot}{m_1+m_2+m_*}\right)\left(\frac{1\,{\rm AU}}{a_{\rm hard}}\right) \tag{A.2}$$

Where $f_{\rm bin}$ is the binary fraction (we assume $f_{\rm bin} = 0.1$, arbitrarily), σ is the three-dimensional velocity dispersion (eq. 3.19) and $a_{\rm hard}$ is the minimum semi-major axis of a hard binary system.

$$a_{\rm hard} = \frac{G\,\langle m\rangle}{\sigma^2} \tag{A.3}$$

We take $m_* = \langle m\rangle$, where $\langle m\rangle = 1\,{\rm M}_\odot$ is the average mass of an object in the system.

The expression in eq. A.2 is valid in the assumption of an isotropic distribution of stars (i.e. in globular clusters or young star clusters). This is of course not true in an AGNs where the stars are distributed in a disk geometry. Therefore, we cannot use this expression in our model directly but we only employ it to estimate the order of magnitude of the timescale.

With the following derivation, we will show that the exchange timescale $t_{\rm exch}$ (eq. A.2) is typically small and can safely be neglected.

First of all, we want to estimate the density of objects in the migration trap $\rho_{\rm trap}$ and see how it changes as a function of time. This will be quite a rough estimate because we only care about the order of magnitude of the exchange timescale.
We generate the mass of 10^4 objects from a Salpeter[1] initial mass function $\xi\,(m)$ [74] between $0.08\,M_\odot$ and $150\,M_\odot$ using inverse random sampling.

$$\xi\,(m)\,dm = \xi_0\left(\frac{m}{M_\odot}\right)^{-2.35}\left(\frac{dm}{M_\odot}\right) \tag{A.4}$$

Due to stellar evolution, the mass function of the objects interacting with the disk (i.e. the present-day mass function or PDMF) will be quite different from the IMF.
From stellar evolution models, we know that stars will produce different remnants according to their initial mass $m_{\rm in}$ [15]. We assume stars with masses lower than $2\,M_\odot$ to not have evolved yet.
Since we are only interested in the average behavior of these objects, we take the mass $m_{\rm fin}$ of each component of the final population to have mass equal to the average mass of its object type. We make direct use of the IMF for the estimate of MS stars average mass, we use [42] for an estimate of WD masses, [77] for NS's and MOBSE data [28, 27, 29, 30] for BH's.

[1]The Salpeter IMF predicts an overdensity of lighter stars compared to more modern IMF prescriptions, but we use it in this context for simplicity.

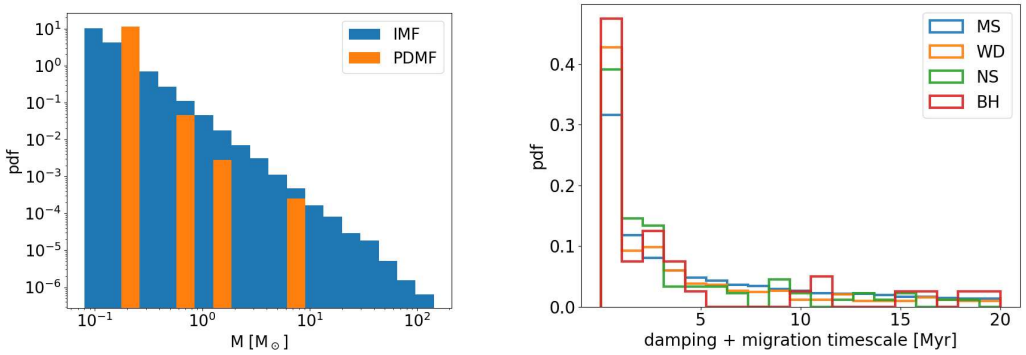


Figure A.1: On the left: Salpeter IMF in blue and toy PDMF in orange. On the right: Damping and migration timescales for the components of the NSC population.

The relations between initial and final masses are:

$$\begin{aligned}
m_{\rm in} \in [\,0.08\,,2\,]\,M_\odot \quad &\Rightarrow \quad \text{Main Sequence (MS)} \quad \Rightarrow \quad m_{\rm fin,MS} = 0.2\,M_\odot \\
m_{\rm in} \in (\,2\,,8\,]\,M_\odot \quad &\Rightarrow \quad \text{White Dwarf (WD)} \quad \Rightarrow \quad m_{\rm fin,WD} = 0.6\,M_\odot \\
m_{\rm in} \in (\,8\,,20\,]\,M_\odot \quad &\Rightarrow \quad \text{Neutron Star (NS)} \quad \Rightarrow \quad m_{\rm fin,NS} = 1.4\,M_\odot \\
m_{\rm in} \in (\,20\,,150\,]\,M_\odot \quad &\Rightarrow \quad \text{Black Hole (BH)} \quad \Rightarrow \quad m_{\rm fin,BH} = 8.7\,M_\odot
\end{aligned}$$

The IMF and toy PDMF are shown in Figure A.1.

Now we want to consider the interaction between the PDMF objects and the AGN gaseous disk. We take as a fiducial value for the SMBH mass the average from the Illustris TNG50 SMBH population (eq. 3.1).

$$M_{\rm SMBH} = 10^{6.84}\,M_\odot$$

We assume the disk to be dense, thin and long-lived.
We extract the mass of the host galaxy and of the NSC according to eq.s 3.4 and 3.14 respectively. We establish the effective radius of the NSC according to eq. 3.15. Since we want to reproduce the average behavior, we neglect any statistical scattering around these values.

We compute the total mass $M\,(R_{\rm max})$ of the objects interacting with the disk using eq.s 3.17 and 3.9. We re-scale the number of objects $N^{\rm sample}$ in the toy PDMF sample to N so that their total mass is comparable with $M\,(R_{\rm max})$:

$$M_{\rm sample} = \sum_{\rm XX} N^{\rm sample}_{\rm XX} m_{\rm fin,XX}\,, \quad N_{\rm XX} = N^{\rm sample}_{\rm XX} \left(\frac{M(R_{\rm max})}{M_{\rm sample}} \right) \tag{A.5}$$

where XX = MS, WD, NS, BH.

We set the initial position of these objects as in eq. 3.24 and we compute the damping and migration timescale using eq.s 2.12 and 2.13. The resulting timescales for the different components of the NSC population are shown in Figure A.1.

We compute the total mass of objects in the migration trap $M_{\rm trap}$ by summing the masses of objects that have already migrated at a certain time. We convert the mass into density by

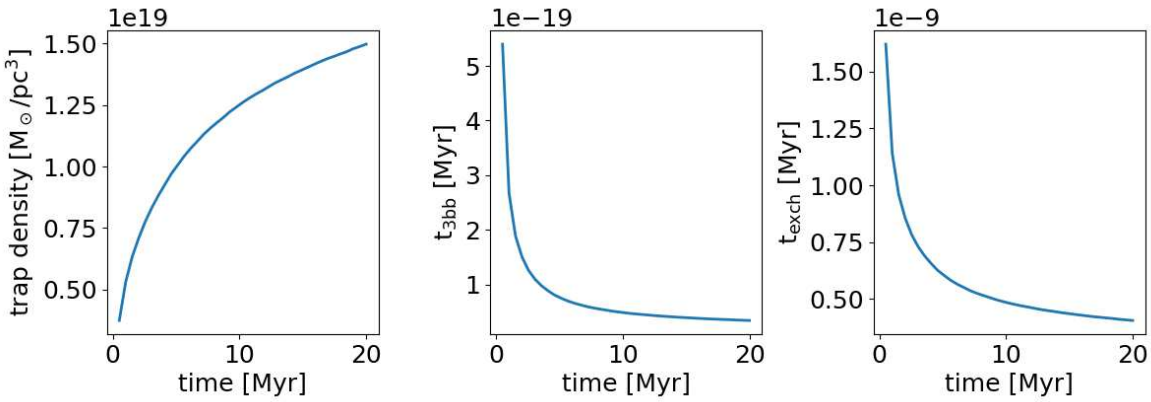


Figure A.2: On the left: Mass density in the migration trap as a function of time. On the right: Exchange timescale as a function of time.

dividing it by the migration trap volume. We assume the trap to be an annulus with a square section (equal height and width).

$$\rho_{\rm trap} = \frac{1}{2}\,\frac{M_{\rm trap}}{2\pi(h_{\rm trap})^2(R_{\rm trap})^3} \tag{A.6}$$

where the factor 1/2 is to account for prograde rotators only. $R_{\rm trap}$ is the migration trap radius (eq. 3.10) and $h_{\rm trap}$ is the aspect ratio at the migration trap (eq. 3.12).
See Figure A.2 for the evolution of mass and density in the migration trap.
We see that the density in the migration trap during the AGN disk lifetime (time $\leq 10\,$Myr) is very high: $\rho_{\rm trap} \sim 10^{19}\,{\rm M}_\odot/{\rm pc}^3$. Hence the pairing time with an object in the migration trap is expected to be negligible.

This value of the density in the migration trap is not accurate both because the model used for the PDMF is naive and, most importantly, because we are oversimplifying two- and three-bodies interactions. As mentioned in section 2.4, the interaction of light-mass objects (such as stars and white dwarfs) with compact objects (such as neutron stars and black holes) can potentially strip or destroy the former, hindering their migration, while the interaction of light-mass objects with hard binaries can easily scatter them out of the migration trap, effectively decreasing the trap density.
Besides, the assumption that the migration trap has equal height and width is arbitrary: there's no reason why the trap could not be have a larger width and therefore a larger volume than assumed.
We are also neglecting the effects of BBH mergers that can kick the remnants out of the migration trap.

Nevertheless, knowing the value of the density in the migration trap, we compute the dynamical binary formation timescale as in eq. A.1.
As shown in Figure A.2, the typical pairing timescale in a spherical geometry is of the order of 10^{-13} years, i.e. 10^{-5} seconds. Therefore we can safely assume that the pairing timescale in a disk geometry will be negligible as well.

We compute the exchange timescale $t_{\rm exch}$ as in eq. A.2. As shown in Figure A.2, the typical exchange timescale is of the order of 10^{-3} years, i.e. of a few hours, which is very small compared to the timescales at play (see chapter 4).

Although this is a toy model and this estimation is not reliable, we can reasonably assume that the pairing and exchange timescales are negligible with respect to the other timescales at play.

In conclusion, the pair-up timescale after entrance in the migration trap can safely be neglected in a typical AGN system.

Appendix B

Additional figures

This appendix contains all the alternative figures that we did not include in the main text for brevity. In particular, here we report the results for the Illustris TNG300 run.

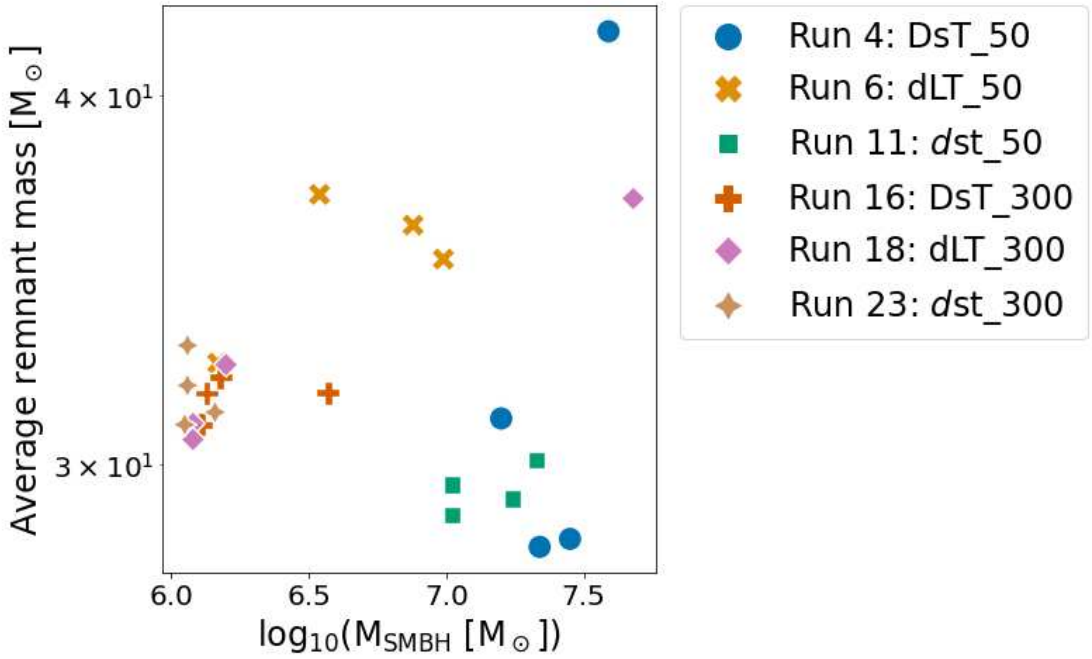


Figure B.1: Same as Figure 4.8 for low-merger runs.

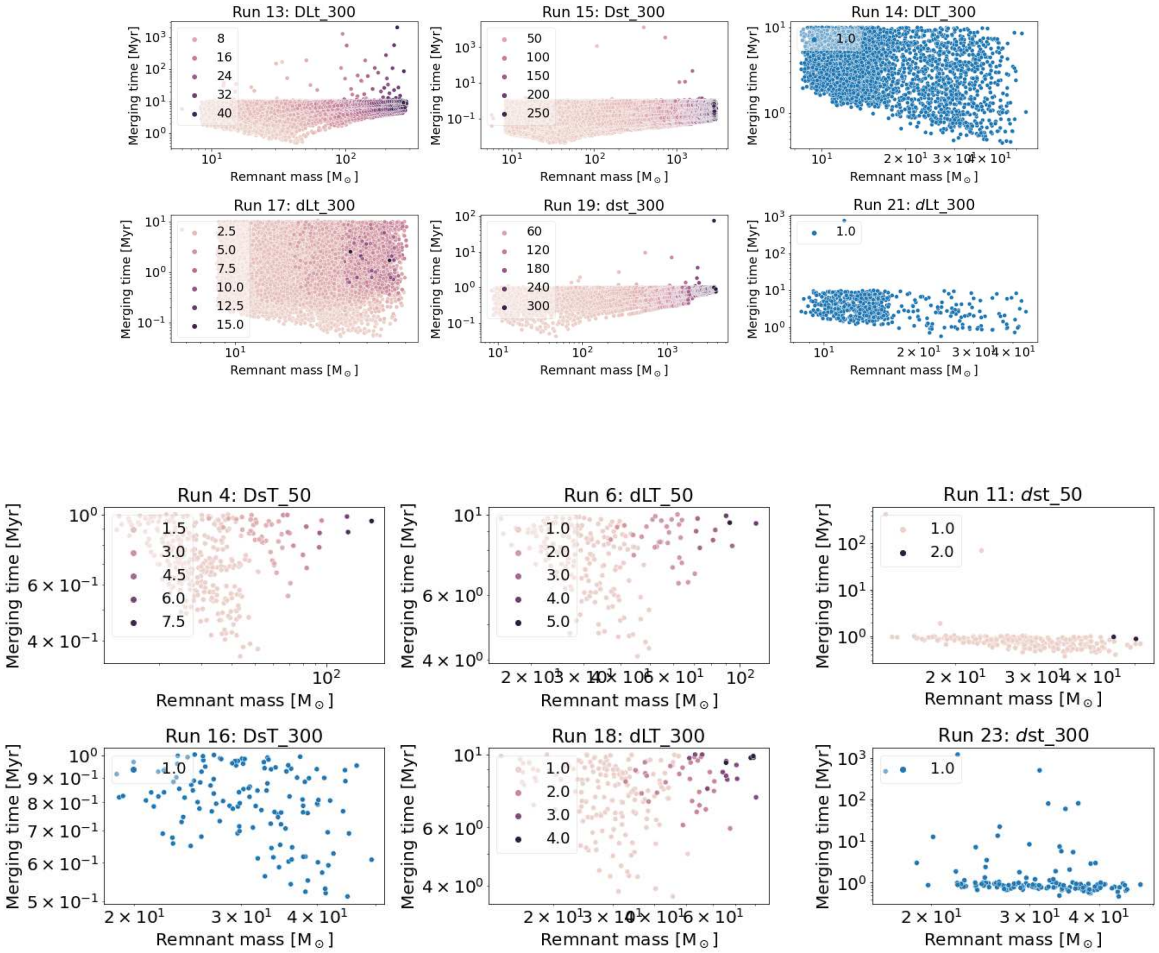


Figure B.2: Same as Figure 4.5a for TNG300 and low-merger runs.

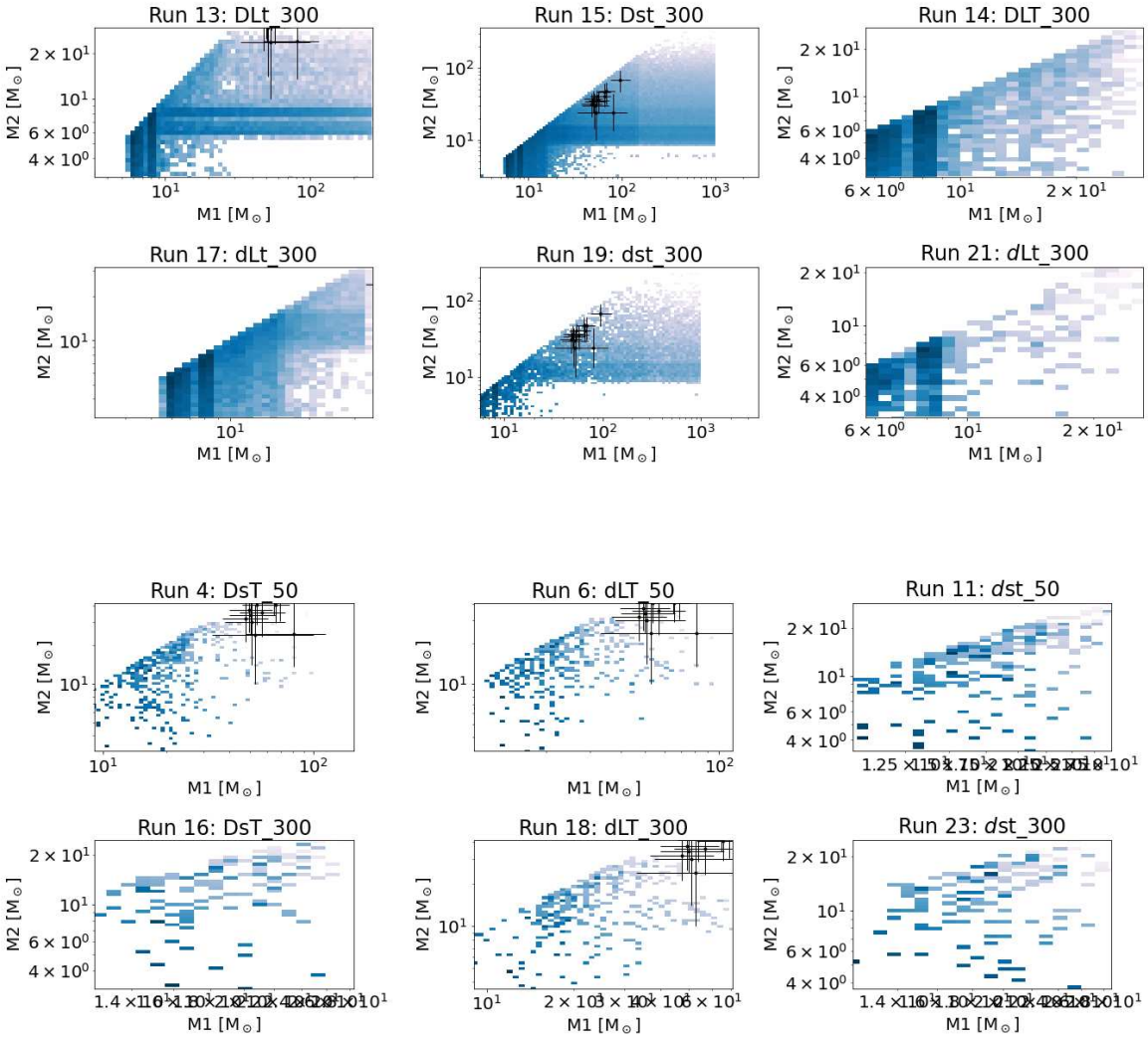


Figure B.3: Same as Figure 4.5b for TNG300 and low-merger runs.

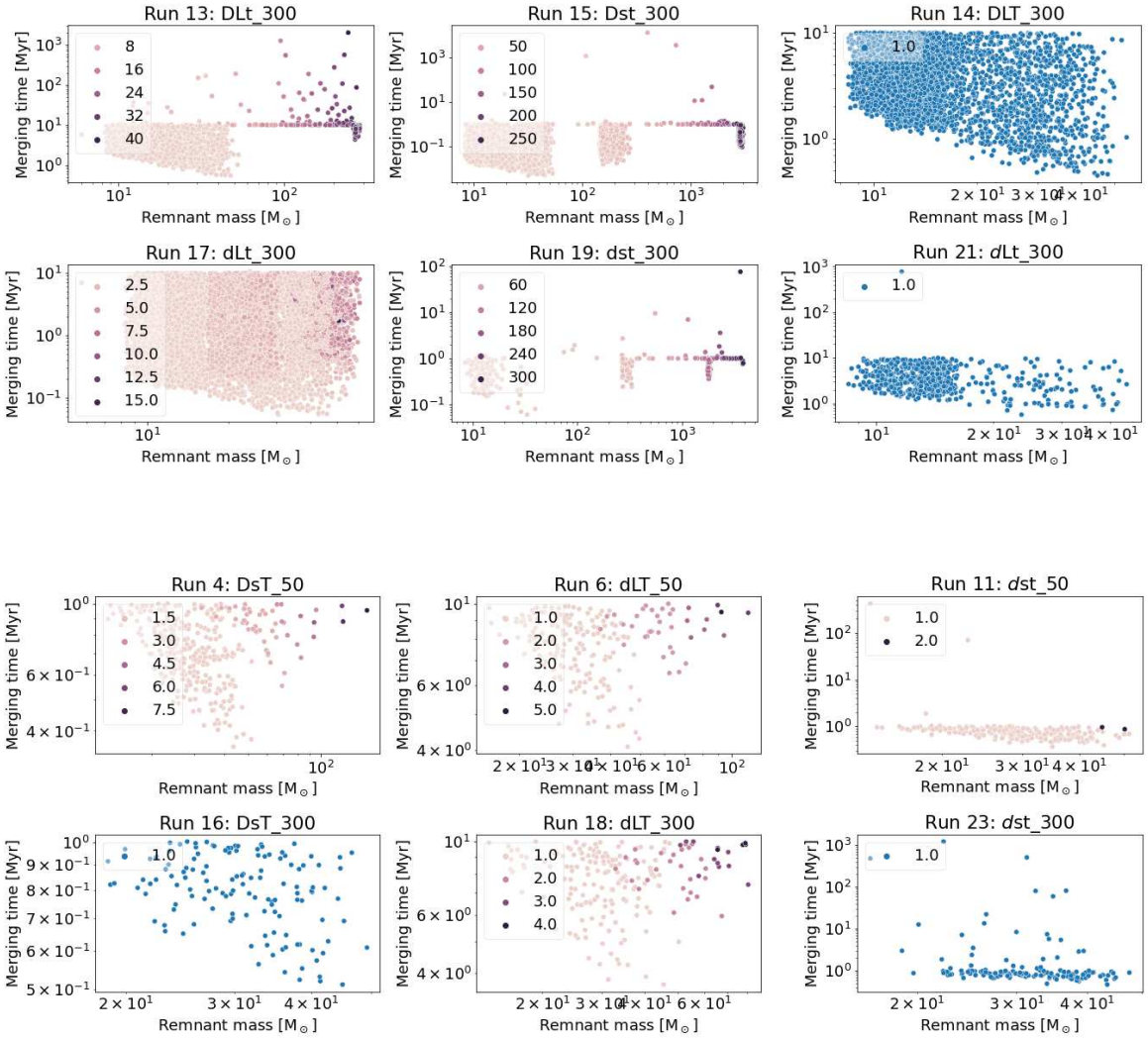


Figure B.4: Same as Figure 4.6 for TNG300 and low-merger runs.

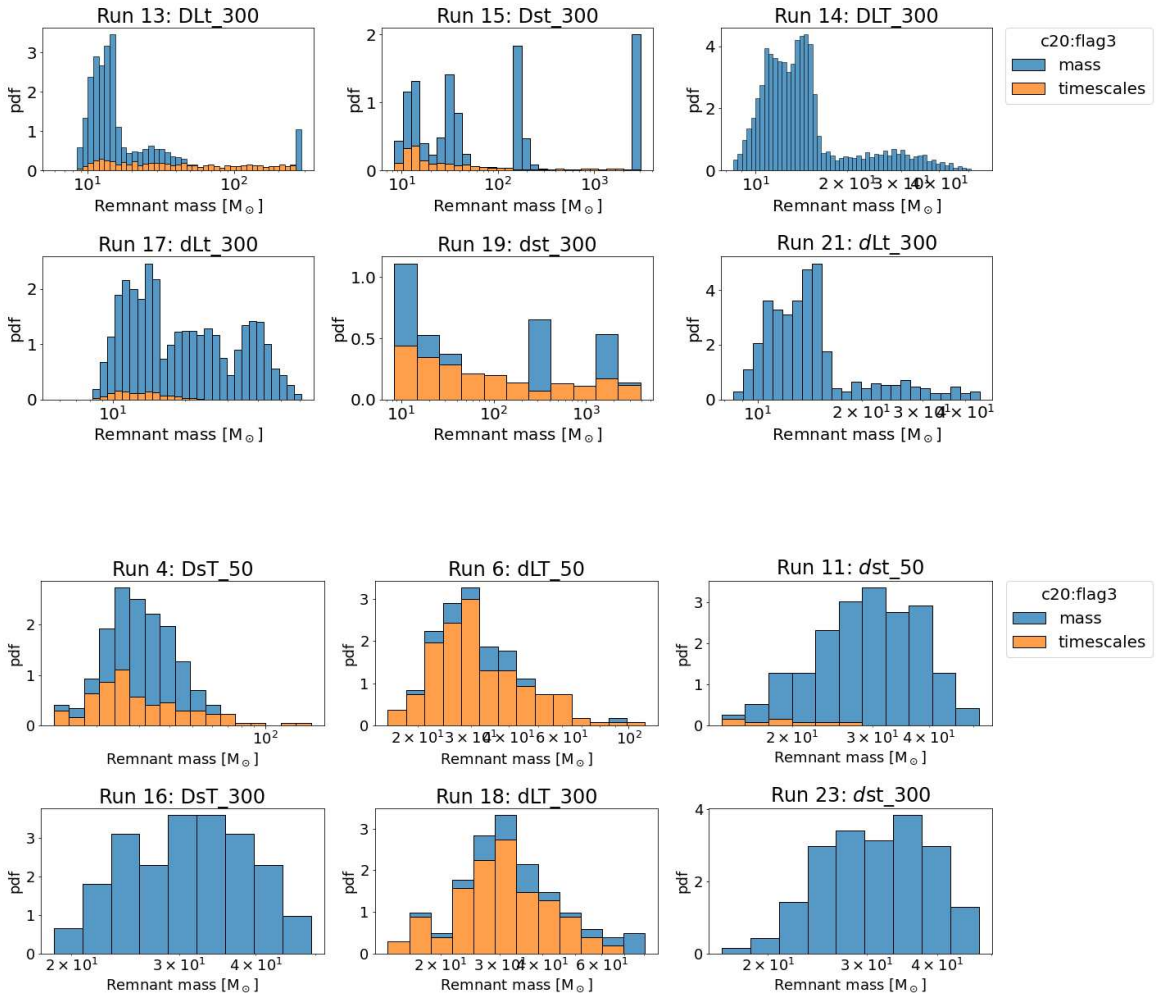


Figure B.5: Same as Figure 4.7a for TNG300 and low-merger runs.

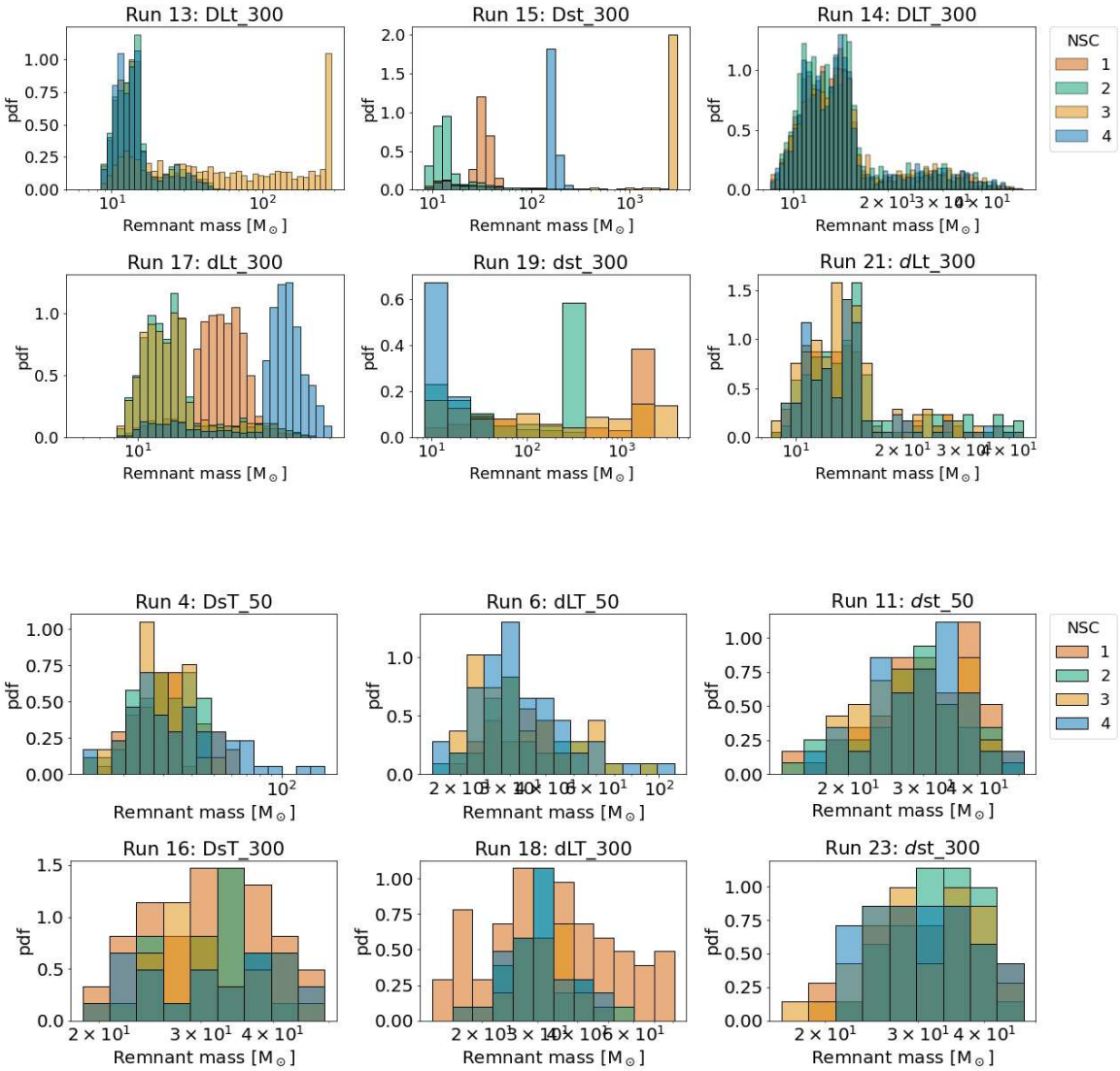


Figure B.6: Same as Figure 4.7b for TNG300 and low-merger runs.

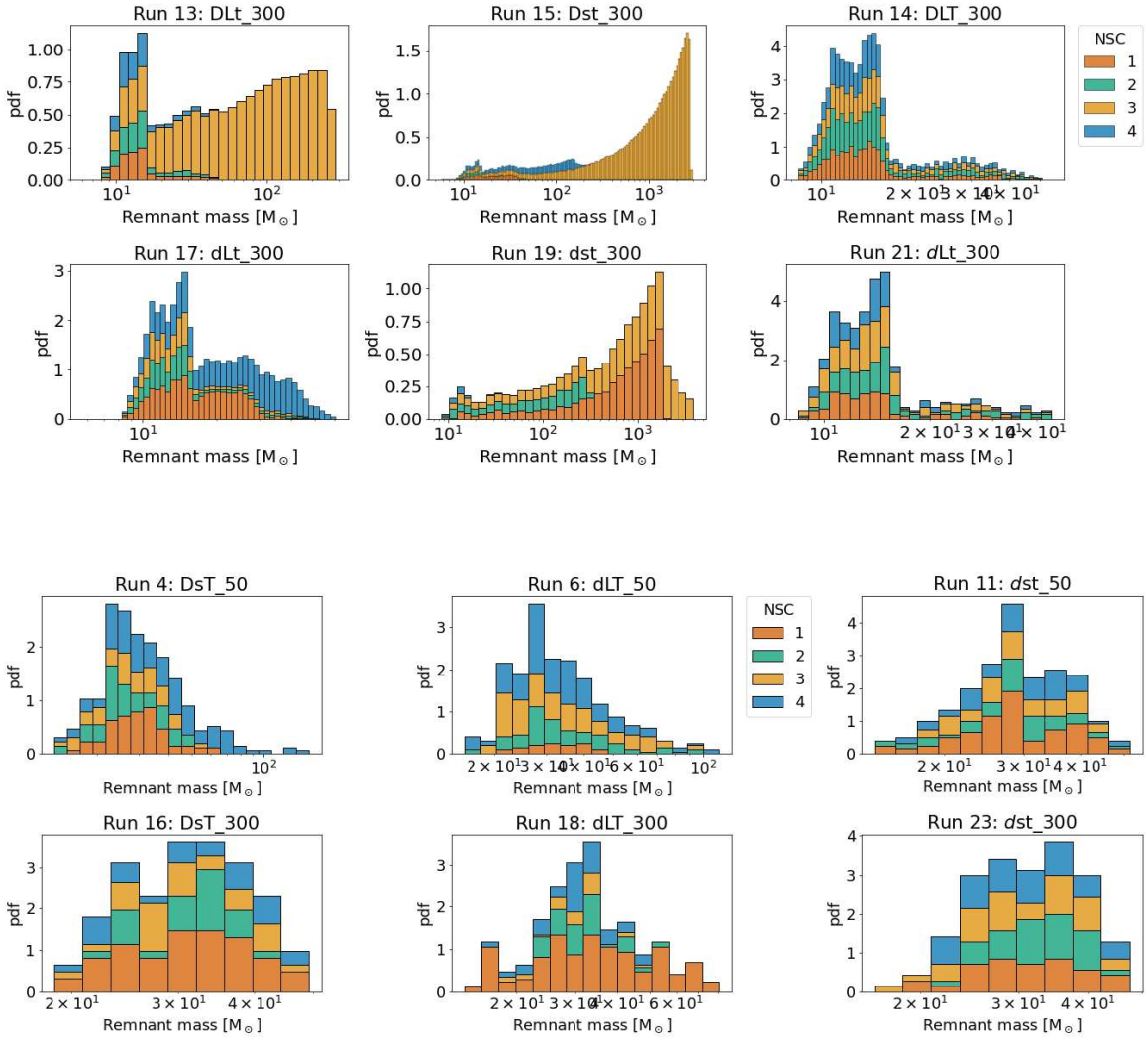


Figure B.7: Same as Figure 4.12a for TNG300 and low-merger runs.

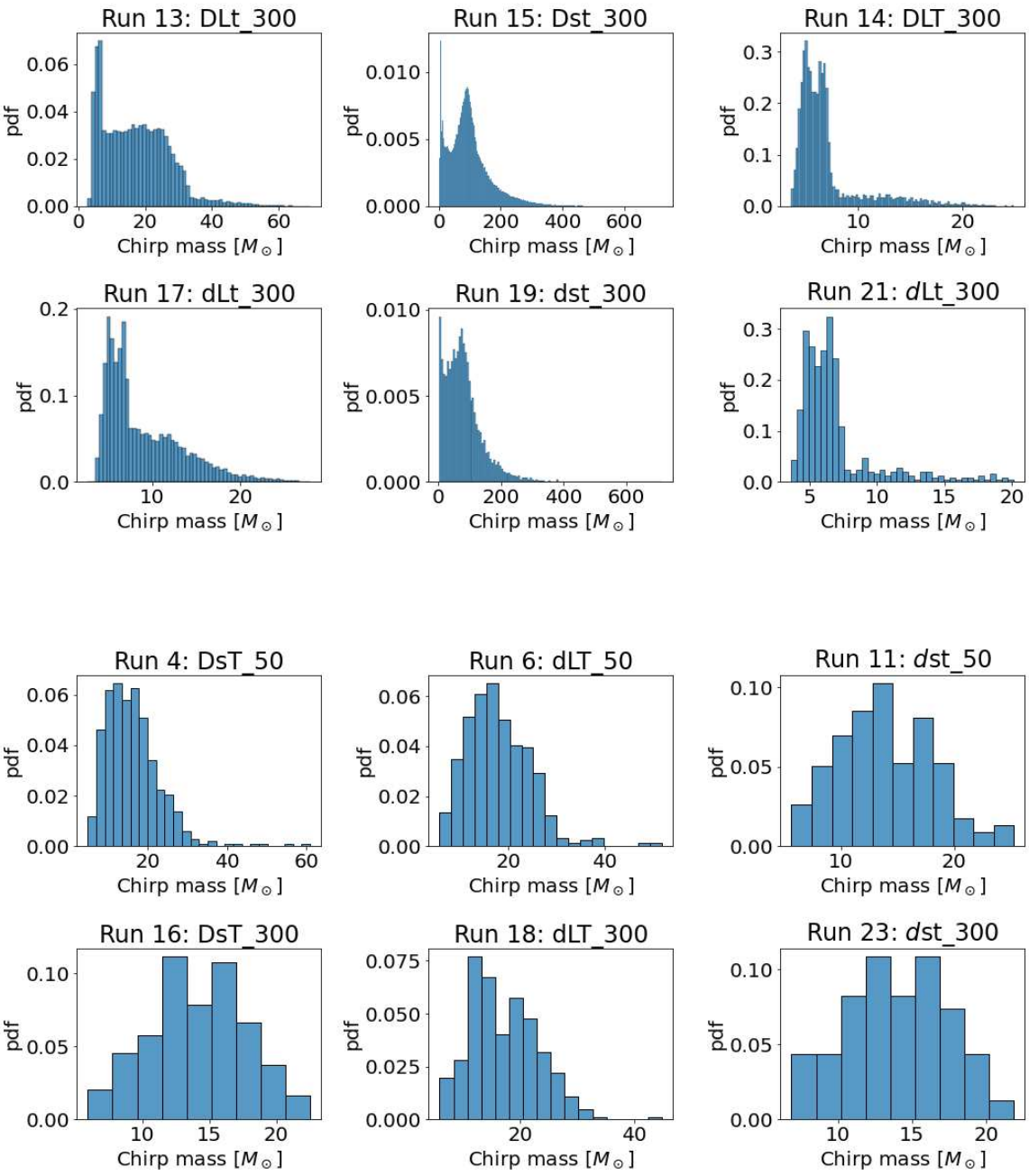


Figure B.8: Same as Figure 4.12b for TNG300 and low-merger runs.