SCUOLA
NORMALE
SUPERIORE

Classe di Scienze
Tesi di perfezionamento in Fisica

Radiative transfer simulations of galaxies and AGN at the Epoch of Reionization

Candidato:
Fabio Di Mascia

Relatore:
Prof.ssa Simona Gallerani

Anno accademico 2022/23

Contents

Abstract

In the last two decades, several sky surveys in the optical/near-infrared allowed to the discovery of over ~ 300 quasars when the Universe was less than 1 billion year old. These extremely bright sources are thought to be powered by black holes gaining mass by accretion and merging, which shine as bright Active Galacitc Nuclei (AGN). Spectroscopic follow-ups have estimated black hole masses in the range of $10^8 - 10^{10}$ $\mathrm{M}_\odot$. The existence of these supermassive black holes (SMBHs) challenge current theoretical models of black hole formation and evolution. They also likely represent just the tip of the iceberg of the whole black hole populations at that epoch, making the discovery of the underlying population and of their progenitors essential to better understand their origin and evolution. Moreover, the birth and growth of these objects is also thought to be strictly connected with the evolution of their host galaxies, making their study deeply connected with the theories of structures formation and evolution.

Incredible progress in both the theoretical and observational sides in the latest years has significantly improved our understanding of the these objects. On the observational side, several facilities, such as the radio-interferometers ALMA and NOEMA, the Hubble Space Telescope (HST), the X-ray telescope Chandra, have provided a large amount of high-quality data. On the theoretical side, a lot of effort has been dedicated to perform sophisticated and computationally expensive hydrodynamical simulations in a cosmological framework. However, these progresses also increased the need for a tailored comparison between observations and theoretical predictions.

Light from distant sources is affected by several processes before being collected by our instruments. For example, dust grains absorb UV/optical photons and emit in the infrared, significantly affecting the observational appearance of the first galaxies. Understanding and properly taking into account the processes involved in the propagation of photons through a dusty medium has become more and more important throughout the years as a bridge connecting observations and theory. Radiative transfer calculations have been proved to be a fundamental tool to address this problem, thanks to the possibility to compute the detailed propagation of photons in a wide wavelength range, from the sub-millimeter band up to the X-rays, thus allowing the possibility to make observational predictions of physical observable quantities from cosmological simulations. Their importance is increasing as current facilities continue to provide exquisite data and new science missions, such as the recently-launched James Webb Space Telescope (JWST), and several observatories planned to operate in the next decade (e.g. ATHENA, AXIS, Origins, Roman and Euclid). This development will be fundamental in order to shed light on the first SMBHs in the Universe, as many questions are still unanswered. In particular:

- What are the physical properties of early SMBHs and their hosts?
- What are the spectral signatures of their progenitors?

- How did the AGN activity affect the host galaxy?

In this Thesis, we tackle these questions by combining cosmological hydrodynamic simulations which follow the evolution of early SMBHs with radiative transfer calculations, simulating their emission in dusty galaxies. The Thesis is structured as follows:

- Chapter 1 contains a general overview of the cosmological context of the Thesis, with a description of the standard cosmological model, a brief summary of the history of the Universe and an introduction to the formation of the first structures.
- Chapter 2 introduces the Active Galactic Nuclei (AGN), their structure according to the unified model and their multi-component emission. We also introduce the theoretical and observational background regarding the formation and observation of the first black holes (BHs).
- Chapter 3 is dedicated to dust. We overview the formation and evolution of interstellar dust grains, their main role in shaping the emission of galaxies and main observational properties in AGN.
- In Chapter 4 we provide a broad overview about modern cosmological hydrodynamic simulations, focusing on the different numerical techniques implemented. We then introduce the hydrodynamic simulations we make use in the work presented in the following chapters.
- In Chapter 5 we present the radiative transfer problem and the main numerical strategies adopted to tackle it, focusing in particular on the Monte Carlo approach. We then introduce the numerical setup we adopt with the code SKIRT in order to simulate AGN emission within dust galaxies.
- In Chapter 6 we show our main results regarding the impact of AGN radiation on the dust temperature in the host galaxy. We also investigate the signatures of AGN emission in the infrared band in high-redshift galaxies and make predictions for upcoming mid-infrared (MIR) facilities.
- In Chapter 7 we expand the discussion of the previous chapter, by investigating how the AGN boosts the IR emission of its host galaxy, and affects the estimate of its star formation rate (SFR).
- In Chapter 8 we study the dust properties, i.e. composition, abundance and grain size distribution, in high-redshift quasars by comparing synthetic attenuation curves with observed ones.
- In Chapter 9 we investigate the impact of AGN feedback on the host galaxy gas column density and we make predictions for the X-ray emission.
- Finally, in Chapter 10 we summarise all our results and describe the future prospects.

Part I

Theoretical background

General introduction 1

"The study of cosmology is a delicate dance between what we observe but do not fully understand and what we fully understand but cannot observe" (Abraham Loeb). During the last century, these two sides (what we know and what we observe) have come closer, thanks to enormous progresses both in the theoretical modeling of our Universe and in the observation capabilities, and yet the Universe continues to puzzle us as our journey is far from being over.

1.1 The cosmological model

At the beginning of the XX century, two major events completely changed our perception of the Universe, and of our place within it. With one of its articles of 1905, Einstein revolutionized the classic notion of space and time, which were thought to be universal concepts, by introducing the theory of special relativity (Einstein, 1905). Later on, in 1916, he introduced the theory of general relativity (Einstein, 1916), which describes the gravitational interaction connecting the space-time geometry with the mass-energy distribution. Einstein's theory of gravity is one of the fundamental ingredients of modern cosmology.

In 1924 Hubble discovered that some of the so-called "nebulae", which were thought to reside inside the Milky Way since the original discovery by Messier towards the end of the 18th century, were actually extra-galactic systems (Hubble, 1926a,b, 1929b). These observations revealed that the Universe was much bigger than what previously thought. Not long after, in 1929, he measured the radial velocity of the extra-galactic nebulae, finding a linear correlation with their distance (Hubble, 1929a). This relation, known as the *Hubble's law* (also called *Hubble-Lemaître law*, recognizing the work by Georges Lemaître, who independently found the same result) was the first observational evidence that leads to the theorizing of the cosmic expansion.

Another key discovery of the last century that leaded to the current cosmological model was the discovery of the Cosmic Microwave Background (CMB) in 1965 (Penzias & Wilson, 1965), which favoured the hypothesis of a hot, initial dense state of the Universe ("Hot Big Bang" model), as opposed to a steady-state Universe, which was previously the leading model.

Since then, enormous progresses have been carried in the understanding of our Universe, and in its mathematical description. The current cosmological model is built upon the following three fundamental pillars:

- The homogeneous (Maartens, 2011, e.g.) and isotropic (Saadeh et al., 2016, e.g.) expansion of the Universe;
- The existence of the Cosmic Microwave Background;

- The abundance of elements, with a $\sim$ 75% mass-abundance in hydrogen and $\sim$ 25% in Helium and traces of other light elements (D, ^{7}Be, ^{7}Li), which is remarkably uniform throughout the Universe, suggesting a common origin on cosmological scale, the so-called *primordial nucleosynthesis* (e.g. Alpher et al., 1948).

From a theoretical point of view, the standard cosmological model relies on two simple assumptions:

- The dynamics is regulated by the laws of the General Relativity.
- On sufficiently large scales, the Universe is homogeneous and isotropic, the so-called *cosmological principle* [1].

Originally, the assumption of a homogeneous and isotropic Universe was a simplification made by Einstein to find a solution to its equations. *Homogeneity* means that the physical conditions are the same in any point in space; *isotropy* means that the Universe appears the same for an observer toward each direction, i.e., there is no privileged direction. These assumptions turn out to be a good approximation of the Universe on the largest observable scales, with fluctuations of the order of 10^{-5} around the average, as revealed since the first measurements of the CMB radiation (Smoot et al., 1992).

In the formalism of the General Relativity, the space-time geometry can be expressed as:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{1.1}$$

Here ds represents the space-time interval between two events [2], $dx^\mu = (cdt, dx, dy, dz)$ is a four-dimensions vector, containing the time coordinate (re-scaled by the speed of light c) and the three spatial coordinates, and $g_{\mu\nu}$ is the *metric*, which defines the actual space-time geometry. For a photon, the space-time interval is $ds = 0$. For an homogeneous and isotropic Universe, the space-time interval ds can be easily expressed in spherical coordinates by using the Friedmann-Lemaître-Robertson-Walker (FLRW) metric:

$$ds^2 = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right], \tag{1.2}$$

where $a(t)$ is the cosmic *scale factor*, which describes the expansion of the Universe, and k is the *curvature*, which is found to be essentially flat, i.e. $k \approx 0$, by CMB measurements (de Bernardis et al., 2000; Planck Collaboration et al., 2020).

The metric in eq. 1.2 enters in the Einstein's field equations, which describe the dynamic of the Universe:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu} + \Lambda g_{\mu\nu}, \tag{1.3}$$

where $R_{\mu\nu}$ is the Ricci tensor, R is the curvature scalar, $T_{\mu\nu}$ is the energy-momentum tensor, Λ is the dark energy term or vacuum energy (originally introduced by Einstein as the cosmological constant), and G is the gravitational constant. Under the assumptions of homogeneity and isotropy, the 16 equations in 1.3 reduce to 10 independent equations. They relate the local space-time curvature expressed by the metric (left-hand side), with the matter-energy distribution expressed by the energy-momentum tensor and the dark energy term (right-hand side). The

[1] A more rigorous formulation of the cosmological principle is the following: *at any epoch, the Universe appears the same to all fundamental observers, regardless of their individual locations.* A fundamental observer is in an inertial system with respect to the CMB.

[2] A point in the space-time is called an event, and it is defined by one time coordinate and three spatial coordinates.

beauty of Einstein's equations is that the dynamics of the matter is governed by the geometry, which is in turn determined by the matter itself.

The cosmological constant was introduced by Einstein only at a second time, in order to allow for a static solution to its equations for an homogeneous and isotropic Universe. This constant corresponds to a source with a uniform energy density in space, which provides additional pressure support against gravity contraction. However, when Hubble discovered that distant galaxies were receding from us at a velocity proportional to their distance, thereby suggesting a Universe in expansion, Einstein regretted the cosmological constant as his "biggest blunder". When at the end of the XX century it was discovered that the expansion of the Universe is actually accelerating (Perlmutter et al., 1998, 1999; Riess et al., 1998; Schmidt et al., 1998), cosmologists re-evaluated the cosmological constant. Indeed, a constant, uniform energy density term in an expanding Universe progressively increases its pressure support as long as the expansion proceeds, thereby accelerating the expansion. Today, the constant Λ also gives name to the current standard cosmological model, the Λ-Cold Dark Matter (ΛCDM) model.

It is interesting to express the expansion of the Universe in the formalism of the FLRW metric in eq. 1.2. The distance between two points in the space is given by $r(t) = a(t)R_c$, where R_c is the *comoving* coordinate, i.e. the distance in a reference frame which follows the Hubble expansion. The scale factor simply translated between the physical coordinates and the comoving coordinates. The physical distance between the two points changes over time, and the velocity by which they recede from each other is given by $v = dr/dt = \dot{a}(t)R_c$. We can then relate the distance with the velocity, obtaining the Hubble's law:

$$v = \frac{\dot{a}(t)}{a(t)} r = H(t)r, \tag{1.4}$$

where $H(t) \equiv \dot{a(t)}/a(t)$ is the Hubble constant. It is clear that the Hubble constant is a constant in space, but not in time. Its present value, denoted with H_0 according to the most recent estimates from the CMB is $H_0 = 67.4 \pm 0.5\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Planck Collaboration et al., 2020), although there is a tension with local measurements (e.g. Riess et al., 2016). The Hubble constant has the dimension of the inverse of a time: it corresponds to the expansion rate of the Universe. Its inverse, $1/H(t)$ sets a fundamental time-scale of the Universe.

The expansion of the Universe alters the energy of photons as they travel in space. It can be shown that $cdt/a(t)$ is the same for a photon as it travels from a source since the time of its emission t_1 until its detection by us at t_0:

$$\begin{aligned} \frac{dt_1}{a(t_1)} &= \frac{dt_0}{a(t_0)} \\ dt_0 &= \frac{a(t_1)}{a(t_0)} dt_0, \end{aligned} \tag{1.5}$$

that is: time intervals in the past appear shorter than the present day in an expanding Universe, according to the ratio of the scale factor at the two epochs. Given that the frequency of a photon scales as $\nu \propto dt^{-1}$, the previous equation can also be expressed as:

$$\frac{\nu_1}{\nu_0} = \frac{\lambda_0}{\lambda_1} = \frac{a(t_0)}{a(t_1)}, \tag{1.6}$$

where λ is the wavelength of the photon. We can now introduce the *redshift* z as the relative difference between the observed and emitted wavelengths of a photon:

$$z(t_1) = \frac{\lambda_0 - \lambda_1}{\lambda_0}, \tag{1.7}$$

which can be expressed as:

$$1 + z(t_1) = \frac{a(t_0)}{a(t_1)} = \frac{1}{a(t_1)}, \tag{1.8}$$

using the convention $a(t_0) = 1$. The redshift provides a convenient method to express both the distance of objects and the time in the past at which we are observing them. The term redshift brings its name by the effect just described regarding the photon energy: a photon reaching us from a distance source is detected with a wavelength longer than the one it had at the time of its emission, thereby its energy is shifted toward the red portion of the spectra.

The evolution of the Universe is profoundly regulated by the evolution of the scale factor $a(t)$. For an homogeneous and isotropic Universe, filled by a perfect fluid with density ρ and pressure p, the scale factor evolves according to the Friedmann's equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} \tag{1.9}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right), \tag{1.10}$$

which were first derived by Friedmann from the Einstein's field equations (Friedmann, 1922). Friedmann's equations require to know the energy-density relation of the matter filling the Universe, which is expressed by an equation of state that relates the pressure p with the density ρ. A simple form of the equation of state is $p = w\rho c^2$, where w can assume different values depending on the type of the component:

- *matter*: $P_\mathrm{m} < \frac{1}{3}\rho_\mathrm{m}$, that is $w < 1/3$;
- *radiation*: $P_\mathrm{r} = \frac{1}{3}\rho_\mathrm{r}$, that is $w = 1/3$;
- *vacuum energy*: $P_\Lambda < \frac{1}{3}\rho_\Lambda$, that is $w = -1$.

The vacuum energy term corresponds to a fluid with a negative pressure. From the second Friedmann's equation, we see that the pressure term is related with the second derivative of the scale factor. A Universe in accelerating expansion, as found at the end of the XX century, has to satisfy the condition $\rho + 3p/c^2 < 0$, that is $w < -1/3$, thereby implying the presence of a dark energy term.

From the first Friedmann's equation, we can also derive the *critical density* of the Universe, that is the density (considering all the fluid components) necessary to have a flat Universe, that is the curvature parameter $k = 0$:

$$\rho_\mathrm{c} = \frac{3H^2}{8\pi G} \approx 8.5 \times 10^{-30}\left(\frac{H}{H_0}\right)^2. \tag{1.11}$$

We can then rewrite the first Friedmann's equation by considering the three energy-components contributing to ρ, and rescaling them with the present-day critical density $\rho_{\mathrm{c},0}$:

$$\frac{H(t)}{H_0} = \sqrt{\Omega_\mathrm{m} + \Omega_\mathrm{r} + \Omega_\Lambda}, \tag{1.12}$$

where we have defined $\Omega_i = \rho_i/\rho_{\mathrm{c},0}$ for each component, and $\rho_\Lambda = \Lambda/(8\pi G)$. Finally, we can express each component in terms of its present-day value, considering their scaling with the scale factor. For the matter, the energy density scales with the volume as $\rho_\mathrm{m} \propto a^{-3}$; for the

$\Omega_{\mathrm{m},0}$	$\Omega_{\Lambda,0}$	$\Omega_{\mathrm{b},0}h^2$	h	n_{s}	σ_8
0.321	0.679	0.02212	0.6688	0.8118	0.9626

Table 1.1: The basic set of cosmological parameters in the standard ΛCDM cosmology model. Here $\Omega_{\mathrm{m},0}$ is the matter density, $\Omega_{\Lambda,0}$ is the dark energy density, $\Omega_{\mathrm{b},0}$ is the baryonic density, h is the Hubble constant in units of $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, n_{s} is the scalar spectral index of the primordial fluctuations and σ_8 represents the present-day fluctuation on scales of $R = 8\,h^{-1}$ Mpc (see Section 1.3 for these two parameters). Data from Planck Collaboration et al. (2020).

radiation it scales $\rho_{\mathrm{m}} \propto a^{-4}$ (the additional power comes from the frequency redshift); for the dark energy $\rho_{\Lambda} \propto a^{0}$. With this substitutions, the previous equation becomes:

$$\frac{H(t)}{H_0} = \sqrt{\frac{\Omega_{\mathrm{m},0}}{a^3} + \frac{\Omega_{\mathrm{r},0}}{a^4} + \Omega_{\Lambda,0}}. \tag{1.13}$$

This equation describes the evolution of the Universe in terms of the matter-energy components that fill it. Today we measure $\Omega_{\mathrm{m},0} = 0.315 \pm 0.007$ and $\Omega_{\Lambda,0} = 0.685 \pm 0.007$. These parameters are essential in order to understand the history of our Universe, and they are indeed part of the basic sets of parameters of the standard cosmological model, some of which are summarised in Table 1.1.

It is interesting to note that the matter content of the Universe is consistent with the critical density, i.e. $\Omega_{\mathrm{m},0} + +\Omega_{\mathrm{r},0} + \Omega_{\Lambda,0} \approx 1$, within the errors on the measure. From the Friedmann's equations, this implies $k = 0$, that is: our Universe is essentially flat.

Eq. 1.13 also shows that in different epochs during the evolution of the Universe, a different matter-energy component regulated its evolution, because matter, radiation and the dark energy evolves differently with the scale factor. By equating Ω_{m} and Ω_{r} and considering that $a \propto (1+z)^{-1}$, we find that at $z_{\mathrm{eq}} \approx 3300$, the two component had a comparable energy density. This means that for $z > z_{\mathrm{eq}}$, the Universe was in a *radiation-dominated* era. After that moment, the Universe entered into the *matter-dominated* era, until eventually the dark energy caught up at around $z \approx 0.3$. We are now in the *dark-energy* dominated era. These three phases are illustrated in Figure 1.1. In the next Section, we will provide a brief overview of the main steps of the history of the Universe, finally reaching the Epoch of Reionization (see Section 1.5), where the work of this Thesis is focus on.

1.2 A brief history of the Universe

In order to understand the evolution of the Universe, it is necessary to introduce two time scales. The first time-scale is $t_{\mathrm{H}} = H^{-1}$, which represents the expansion rate of the Universe. If we assume that the Universe is filled with particles with density n, then the interaction rate between them is $\Gamma = n\sigma v$, where σ is the cross section for the considered interaction, and v the relative velocity between two particles interacting. The second natural time-scale is then Γ^{-1}, which depends on the density of the Universe and on its temperature (which sets the velocity of the particles). As long as interactions between particles are frequent enough, i.e. $\Gamma^{-1} \ll H^{-1}$, the Universe evolve as following stages of thermal equilibrium. For this reason it makes sense to define a temperature T for the species in the thermal bath. As the expansion progresses, the temperature decreases and therefore the interaction rate decreases, thereby more and more species go out from the thermal equilibrium, de-couple from the thermal bath.

If we imagine to reverse the time back, an expanding Universe as per the Hubble law should have started with all the matter condensed in a single point, a singularity named *Big Bang*. This

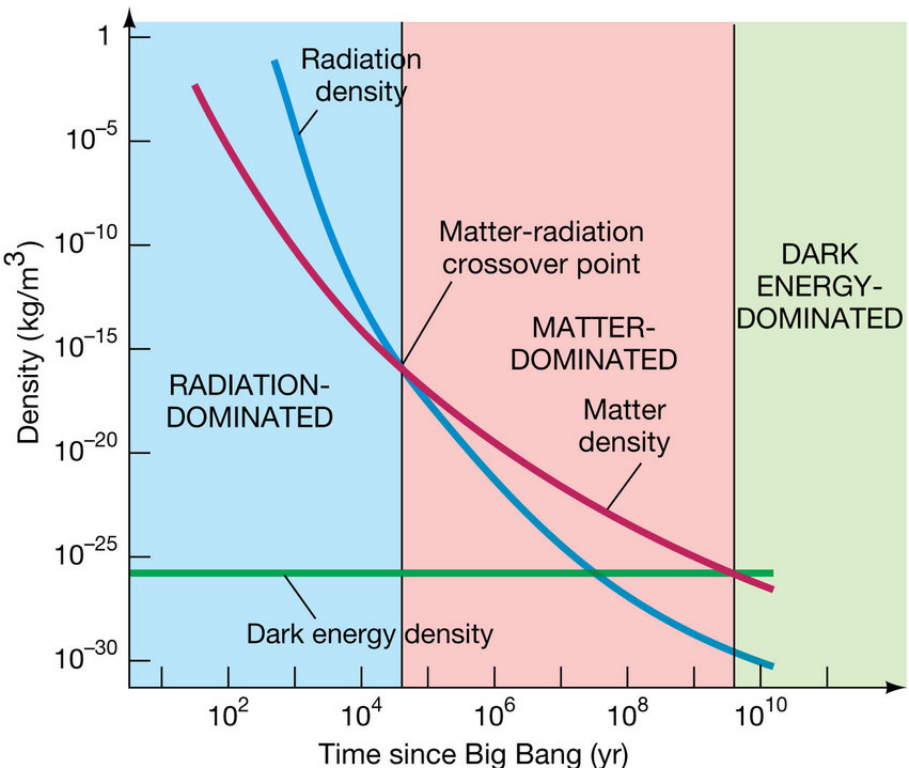

Figure 1.1: A brief history of the Universe with a focus on the domination of its different components. The radiation-dominated era, from the Big Bang until $z \approx 3300$, the matter-dominated era for $0.3 < z < 3300$ and the dark energy era for $z < 0.3$. *Copyright: Pearson Education, Inc. (2011).*

event is estimated to have occurred ≈ 13.7 Gyr ago. The first time it makes sense to talk about is the *Planck scale* ($t \approx 10^{-43}$ s, $T \approx 10^{19}$ GeV). Current theories have no predicting value to describe the status of the matter in these first instants, where the density and the temperature are supposed to be essentially infinite.

It is theorized that shortly after the Big Bang ($t \approx 10^{-36} - 10^{-32}$), a very rapid expansion phase has occurred in an exponential fashion, with the Universe increasing its size by ≈ 60 e-folds, i.e. more than a factor of 10^{25}. This period is named *cosmic inflation*. This phase was proposed in order to explain few problems presented by the original cosmological model (e.g. Linde, 1982). First, as discussed in Section 1.1, the Universe appears to be homogeneous and isotropic on the largest observable scales. This requires casual connection between regions so far apart that light could not have the time to cross from one to another during the history. Secondly, the most recent estimates suggest that the geometry of the Universe is remarkably flat, with a curvature parameter consistent with 0, $k \approx 0$ (see Section 1.1). From the Friedmann's equations, this implies that in the past the Universe had to be even flatter. The homogeneity and flatness of the Universe can either be the consequences of very specific initial conditions that allowed for our existence (the so-called *anthropic principle*), or the result of a physical process. A very rapid expansion, such as the cosmic inflation, would help in solve both these issues. In fact, a tremendous expansion at a rate faster than the speed of light would have displaced far away regions that were previously in equilibrium. Moreover, the flatness can be the consequence of a very fast expansion that stretched the geometry in the early stages of the Universe, removing any dependence on the initial value of the total energy density. Inflation also had another major impact in the history of the Universe, magnifying the primordial quantum fluctuations, leading to the seeds responsible for the formation of the first structures.

After the inflation, an unknown phase known as *baryo-genesis* or *lepto-genesis* generated

an excess of particles over anti-particles. During the following evolution, particles and anti-particles annihilated, leaving a Universe made only by particles as we observe today.

When the temperature dropped to $T \approx 1$ GeV ($t \approx 10^{-9}$ s), quarks could not stay free anymore in the quarks-gluon plasma, thereby condensing and forming protons and neutrons: this phase is known as the *quantum chromo-dynamics (QCD) phase transition*.

At $T \approx 1$ MeV ($t \approx 1$ s), neutrinos decoupled too. Due to the cosmic expansion, the decoupled neutrinos continued to evolve as a black-body with a temperature decaying as $T \propto (1+z)$. Therefore, it is predicted that primordial neutrinos are currently forming a *cosmic neutrino background* (Cνb) with a temperature $T_\nu \approx 1.95$. Given their weakly interacting nature, detecting neutrinos with these low energies ($10^{-6} - 10^{-4}$ eV) is currently unfeasible. However, indirect evidence of the (Cνb) has been found (Betti et al., 2019), and experiments to directly detect it are under development (Follin et al., 2015).

In the next few minutes, the temperature of the thermal bath decreases to the point where protons and neutrons could become bound into atomic nuclei, forming the first elements in the Universe. This process is called *Big Bang nucleosynthesis* (BBN). As mentioned in the Introduction, this process is one of the three pillars of the current cosmological model, as it explains the homogeneous abundance of the light elements (H, D, ^{4}He, ^{7}Li) in the Universe, suggesting a common origin from a process happening on cosmic scales. Furthermore, the physics which regulates the nuclear reactions that occurred during these stages is known, and therefore abundance predictions from BBN can be directly tested with observations, and they constitute one of the major successes of the Standard Model. First of all, the abundances of the elements produced depend on two parameters, which were set in the previous stages of the evolution. Approximately during the same period of neutrinos decoupling, when the temperature was $T \approx 1$ MeV ($t \approx 1$ s), the $beta$-interactions which kept protons and neutrons in equilibrium ceased to be effective and they shift in favor to the less massive protons. After the freeze out, neutrons started to decay, and the neutrons-to-protons ratio had dropped down to $n_n/n_p \approx 1/7$ at the moment when BBN began. The second parameter is the photons-to-baryons ratio. In this Epoch, the Universe was radiation-dominated, and the relative abundance of photons with respect to baryons was very much in favor of the former, with a ratio $\eta = n_\mathrm{b}/n_\gamma \approx 10^{-9}$. Thus, when the energy of the thermal bath dropped below the binding energy of deuterium $E_\mathrm{B} = 2.2$ MeV, the first atoms of deuterium started to form via the reaction $p + n \rightarrow d + \gamma$, but only to get immediately destroyed by the high-energy tail of the photons in the thermal bath (distributed as $e^{-E/kT}$). It was only when the temperature dropped to $\eta^{-1} e^{-E_\mathrm{B}/kT} \approx 1$ (corresponding to $T \approx 0.1$ MeV) that deuterium atoms could survive, and give start to the build-up of the other light elements. Nearly all neutrons end up in ^{4}He atoms during BBN, being the most stable element, whereas the others light elements were produced in much less abundance. It is then straightforward to estimate the helium mass fraction (Bergström & Goobar, 2004):

$$Y = \frac{4n_\mathrm{He}}{n_\mathrm{tot}} = \frac{4(n_n/2)}{n_n + n_p} = \frac{2(n_n/n_p)}{1 + (n_n/n_p)} \approx 0.25.$$

This number is in remarkable agreement with the measurements in metal-poor[3] HII regions (regions of ionized hydrogen around young stars), where the helium abundance is supposed to be close to the primordial value, due to the lack of chemical enrichment (e.g. Fields et al., 2020).

After the BBN, which led to the formation of the first light elements, the Universe was still too hot for the species to be in neutral form, and it continues to grow as a plasma of nuclei

[3]In astrophysics, every element heavier than hyrdogen and helium is referred to as a *metal*, and the metals abundance in mass is called *metallicity*. This is commonly indicated as $Z = 1 - X - Y$, where X and Y are the hydrogen and helium mass abundances respectively.

and electrons, efficiently coupled to photons via the Thomson scattering. During this stage, the Universe was essentially opaque, as the mean free path of the photon before encountering a free electron was very short. At $z \approx 3300$, the Universe entered in the matter-dominated era.

According to the standard Λ-CMD model, at some point during the previous stages, another species, which constitutes the dark matter we see today, de-couple from the thermal bath. The therm *dark matter* indicates a matter component (i.e. its energy-density scales with the scale factor as a^{-3}) which does not seem to interact with the electromagnetic field, and therefore it appears dark. There are several observational evidences suggesting that the amount of matter in the Universe is larger than what can be associated to the luminous one, for example rotation curves of galaxies (e.g. Corbelli & Salucci, 2000), and CMB observations (e.g. Peebles, 1982; Planck Collaboration et al., 2020). In the standard cosmological model, the dark matter is thought to be "cold", i.e. to decouple from the thermal bath when non-relativistic, as opposed to be "hot", and decouple when relativistic. Despite the indirect evidences of the dark matter presence in the Universe, it is still not known what its nature is, and many solutions have been proposed, such as: weakly interacting massive particles (WIMPs; e.g. Bertone 2010), Massive Compact Halo Objects (MACHOs; e.g. Alcock et al. 2000), primordial black holes (PBHs; e.g. Carr et al. 2010) and many more (e.g. Bertone et al., 2005). Currently, the nature of the dark matter and its direct detection represents one of the most important question in modern cosmology, and there is a huge effort in the theoretical and particles physics community to address this problem.

At $z \sim 1100$ ($T \sim 3000$ K), a major phase transition occurred, as it became energetically favourable for protons and electrons to recombine into neutral hydrogen atoms. This event could not occur yet at $T \sim 13.6$ eV, because at that time the photons in the high-energy tail of the Planck distribution were still abundant enough to dissociate most of the atoms, due to the low baryon-to-photon ratio. The transition at $z \sim 1100$ is named *Recombination*. Shortly after, the Universe became transparent to the propagation of light, allowing photons to travel from the last scattering surface, thereby providing a cosmic background containing the imprint of the Recombination. This photon background constitutes the *Cosmic Microwave Background (CMB)*, and it is one of the three fundamental pillars of the concordance cosmological model. The CMB is an invaluable tool to study the primordial Universe, as this signal contains an enormous amount of information regarding, among the many, the matter composition of the Universe and the power spectrum of the primordial fluctuations. After its accidental discovery in 1965 (Penzias & Wilson, 1965), the CMB signal was studied by several experiments, including the soviet mission RELIKT-1 in 1983, the NASA COBE mission in 1989 (Smoot et al., 1992), the BOOMERanG experiments, WMAP (Spergel et al., 2003) and, most recently, the Planck satellite (Planck Collaboration et al., 2020). An image of the sky map of the CMB signal reconstructed from the Planck satellite data is shown in Figure 1.2. After removing the dipole anisotropy due to the motion of the Earth around the Sun, the CMB radiation appears as a uniform black-body radiation, because photons escaping the thermal bath preserved this spectral shape as the Universe expanded. However, due to the cosmic expansion, the temperature measured now is reduced by a factor of $1 + z$, and currently we measure a CMB temperature of $T_0 = 2.7255 \pm 0.0006K$ K (Group et al., 2020). However, fluctuations of 18 μK (order of 10^{-5}) are also present, which reflect very small anisotropies in the primordial Universe, that were the seeds of the formation of large structures, and ultimately of galaxies. The CMB has a richness of information regarding the properties of the early Universe. When decomposing the signal into spherical armonics, the temperature power spectrum reveals several peaks (see Figure 1.3), whose positions and amplitudes depend on the values of the cosmological parameters, such as the aforementioned $\Omega_{\rm b}$, $\Omega_{\rm m}$, Ω_Λ, h, which describes the abundance of matter in the Universe, and $n_{\rm s}$, σ_8, which describe the primordial matter spectrum, and thereby the matter fluctuations. Therefore the CMB represents one of the

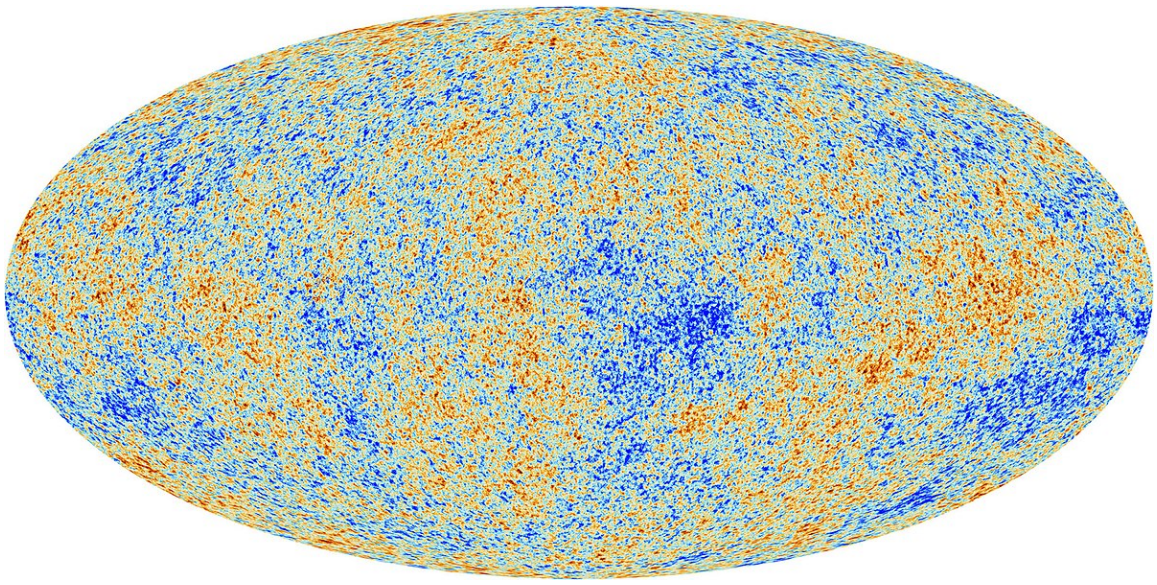

Figure 1.2: An all-sky mollweide view of the CMB signal detected by the Planck satellite, a cosmic relic of the Universe as it was 380000 years after the Big Bang. The signal is remarkably consistent with a black-body with a temperature of $T_0 = 2.7255 \pm 0.0006K$ K, with fluctuations of 18 μK, corresponding to the anisotropies in the primordial Universe, that were the seeds of the first structures formation. *Credit to ESA and the Planck Collaboration*: https://www.esa.int/ESA_Multimedia/Images/2013/03/Planck_CMB.

most important laboratories to test the concordance cosmological model and proving constraints on the fundamental parameters, with which we currently describe our Universe.

After the Recombination, photons could freely travel in the Universe, however there was no source of light yet. This period, called *Dark Ages*, lasted until $z \sim 20 - 30$, when the first bound structures formed, and the first stars appeared. During the Dark Ages, the only emission produced during this epoch was the 21-cm line emission from the spin-flip transition to the ground state of neutral hydrogen. This radiation can be in principle detected in contrast to the CMB, and it entangles a tremendous amount of information regarding the thermal history of the Universe and the formation of the first structures, and it could also provide complementary constraints regarding the matter content of the Universe. For this reason, there is a great undergoing effort from the astrophysical community to detect this signal, with experiments such as the Square Kilometre Array (SKA)[4] and the Hydrogen Epoch of Reionization Array (HERA) (DeBoer et al., 2017).

In a Universe essentially dark, the first matter fluctuations slowly began to grow, eventually leading to the formation of the first dark matter halos, which were the place where the first galaxies formed later on. It is amazing to imagine that the building blocks of the spectacular and diverse Universe that we observe today were formed in the darkness. We will discuss in more details about the birth of the first structures in Section 1.3.

The birth of the first stars marked the end of the Dark Ages and the beginning of a new age, the so-called *Cosmic Dawn*, at approximately $z \sim 20 - 30$. This beginning of this is still beyond the reach of the current facilities, as today the farthest known galaxy resides at $z \sim 11$ (Oesch et al., 2016). As more and more galaxies formed and evolved, their light began to ionize hydrogen atoms in the Inter-Galactic Medium (IGM), causing the last major phase transition of the Universe. This epoch, called *Reionization*, is thought to have lasted between $5 \lesssim z \lesssim 16$, and a lot of ongoing work is focused on constraining the timeline and the spatial evolution of this transition. The work of this Thesis is set on this Epoch, which we discuss in more details in

[4] https://www.skatelescope.org/

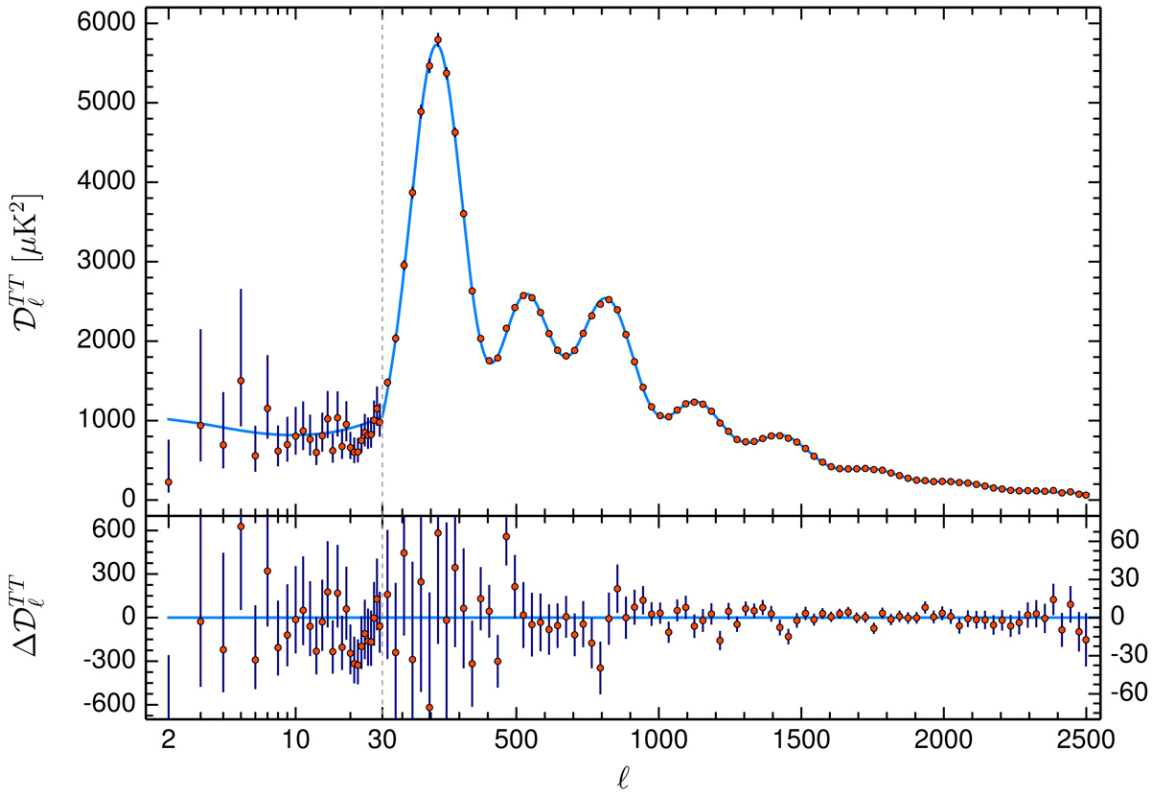

Figure 1.3: Temperature power spectrum of the Cosmic Microwave Background from the Planck 2018 data. The red dots are the Planck data, and the cyan line is the best-fit obtained with the concordance Λ-CDM model. *Credit to the Planck Collaboration* (Planck Collaboration et al., 2020).

Section 1.5.

The Epoch of Reionization ended when the Universe was approximately 1 Gyr old, less than a tenth of the present age (~ 13.7 Gyr). Since then, it continued to evolve, with more and more generations of stars emerging and galaxies slowly assembling into clusters or super-clusters until today. The last major, and relatively recent event, occurred during this period was the transition into the dark energy era at $z \sim 0.3$.

A schematic view of the main Epochs of the history of the Universe is shown in Figure 1.4.

1.3 Structure formation

During the Dark Ages, the small inhomogeneities in the matter distribution in the early Universe began to grow, eventually leading to the formation of gravitationally bound halos. This process is described by the standard model of structure formation (see, for example, the reviews Barkana & Loeb (2001); Dayal & Ferrara (2018), and references therein). The dark matter fluid, which decoupled early from the thermal bath, was the first interested by this process. For this reason the first gravitationally bound structures to form were dark matter halos, which formed the potential wells where the baryons would be accreted, accelerating their collapse, and allowing the formation of the first galaxies.

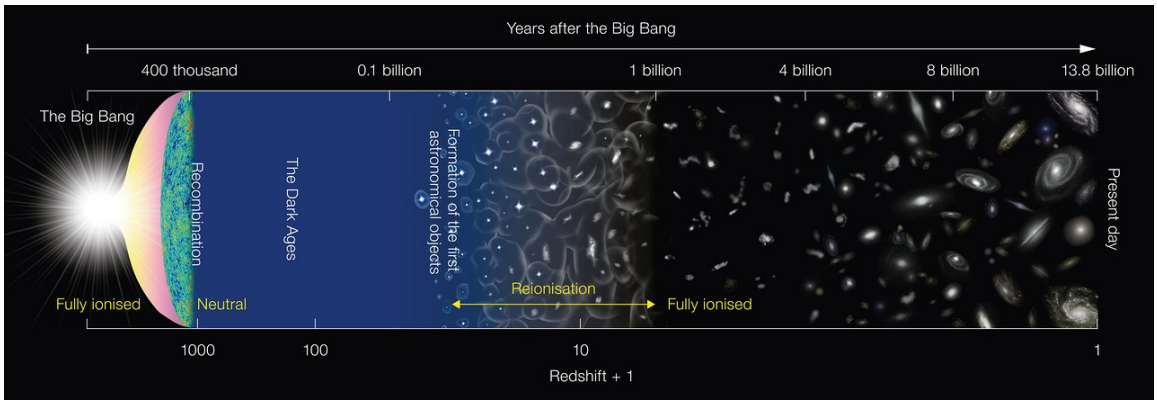

Figure 1.4: A schematic view of the main Epochs of the history of the Universe, from the Big Bang to the present day. *Credit to the National Astronomic Observatory of Japan (NAOJ).*

1.3.1 Linear regime

In this paragraph, we will summarise the basic elements of the structure formation theory. The dark matter interacts only through gravity, so it can be considered as a pressure-less fluid. In the first stage, perturbations are small with respect to the ambient density, with a density contrast $1 + \delta(\mathbf{x}) \equiv \rho(\mathbf{x}, t)/\overline{\rho}(t) \ll 1$, and the growth can be described by the so-called linear regime. The equations of the linear perturbation theory apply to both dark matter and baryons, as long as the densities are low and radiative processes involving baryons can be neglected. In the high density regime, a full Newtonian approach is instead required.

Perturbations grow under the action of the gravity of the matter itself, whereas the expansion of the Universe and the baryonic pressure oppose to the contraction. Such process can be described by the hydrodynamic equations of mass and momentum conservation, which their classical form read:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{1.14}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\frac{1}{\rho}\nabla p - \nabla \phi, \tag{1.15}$$

plus the Poisson equation for the gravitational field:

$$\nabla^2 \phi = 4\pi G \rho. \tag{1.16}$$

In the previous equations, $\rho(\mathbf{x}, t)$ is the total matter density at a given space and time, $\mathbf{v}(\mathbf{x}, t)$ is the fluid velocity, $p(\mathbf{x}, t)$ is the baryonic pressure, and $\phi(\mathbf{x}, t)$ the gravitational potential. In an expanding Universe, it is convenient to write the previous equations in the co-moving reference frame. Moreover, the previous equations can be linearized when considering small perturbations, by expressing them in terms of the density contrast $\delta(\mathbf{x}, t)$. Under these assumptions, the following equation is obtained:

$$\frac{\partial^2 \delta(\mathbf{x}, t)}{\partial t^2} + 2\frac{\dot{a}(t)}{a(t)}\frac{\partial \delta(\mathbf{x}, t)}{\partial t} = \frac{c_s^2}{a^2(t)}\nabla^2 \delta(\mathbf{x}, t) + 4\pi G \overline{\rho}(t)\delta(\mathbf{x}, t). \tag{1.17}$$

Here $a(t)$ is the scale factor and c_s is the fluid sound speed, defined as $c_s^2 = \partial p/\partial \rho$. This equation has the form of the equation for an harmonic oscillator for the density contrast. The source terms on the right hand side are the gravity, which drives the collapse, and the pressure, which opposes

it. The linear term on the left hand side accounts for the expansion of the Universe, which behaves as a damping force opposing to the gravitationally-induced growth of perturbations.

It is convenient to rewrite the previous equation by expressing the density contrast in Fourier modes as $\delta = \sum \delta_k e^{i\mathbf{k}\cdot\mathbf{x}}$, where $\mathbf{k}$ is the wave vector of the k mode:

$$\frac{\partial^2 \delta_k}{\partial t^2} + 2H(t)\frac{\partial \delta_k}{\partial t} = \delta_k \left(4\pi G\overline{\rho}(t) - \frac{c_s^2 k^2}{a^2(t)} \right). \tag{1.18}$$

Written in this form, we can find a scale at which perturbations are just stable against the collapse, by equating the source term to zero:

$$\lambda_J = \frac{2\pi a(t)}{k_J} = c_s \left(\frac{\pi}{G\overline{\rho}} \right)^{1/2}. \tag{1.19}$$

This scale is called the *Jeans wavelength*. The term $t_{\rm ff} = \pi/(G\overline{\rho})^{1/2}$ is the *free-fall time*[5], and it represents the typical time scale over which the collapse happens.

For scales smaller than the Jeans length, $\lambda < \lambda_J$, the oscillator admits a solution in the form of $\delta_k = \delta_0 e^{i(\mathbf{k}\cdot\mathbf{x} \pm \omega t)}$, with $\omega^2 = c_s^2(k^2 - k_J^2) > 0$. This solution represents a density wave that propagates at the speed lower than the sound speed c_s, therefore the pressure is able to prevent the collapse and the perturbations do not grow (the sound crossing time is lower than the free fall time). If instead $\lambda > \lambda_J$, then $\omega^2 = c_s^2(k^2 - k_J^2) < 0$: this solution represents a standing wave whose amplitude increase exponentially with time. In this case, gravity overcomes the pressure and the perturbations can grow.

The mass enclosed in a sphere with radius $\lambda_J/2$ is referred to as the *Jeans mass*, and it can be expressed as:

$$M_J = \frac{4\pi\overline{\rho}}{3}\left(\frac{\lambda_J}{2} \right)^3. \tag{1.20}$$

The Jeans mass represents the minimum mass associated to a density perturbation to collapse under its self-gravity.

1.3.2 The non-linear regime

When the perturbations grow beyond the linear regime, $\delta(\mathbf{x}, t) \sim 1$, the full dynamics must be followed. We will only provide a brief description of the non-linear regime of the perturbations with a simplified model, which nevertheless captures the main concepts. We refer to Barkana & Loeb (2001) for a more detailed description.

We consider spherical perturbations in a uniform density background, in the framework of the so-called *spherical collapse* or *top-hat* model. We consider the dark matter fluid only, thereby ignoring the pressure. We consider a region of initial radius r_i, that is originally expanding following the Hubble flow, with a velocity $\dot{r}_i = H_i r_i$, and an initial density $\rho_i = \rho_0(1 + \delta_i)$. The corresponding mass associated to this region is then:

$$M_i = \frac{4}{3}\overline{\rho} r_i^3. \tag{1.21}$$

The evolution of a shell of radius $r(t)$ under its self-gravity is regulated by the Newton's law:

$$\frac{dr^2(t)}{dt^2} = -\frac{GM}{r^2(t)}. \tag{1.22}$$

[5]The free-fall time is defined as the time required for a sphere of mass M and uniform density ρ to collapse into a point under the only force of gravity. The exact derivation gives $t_{\rm ff} = \pi/(32G\rho)^{1/2}$, however, from a physical standpoint, the different coefficients are not relevant to our discussion.

If the the velocity of the shell is below the escape velocity, the previous equation can be solved by introducing the following parametrization:

$$r(t) = A(1 - cos\theta) \tag{1.23}$$
$$t = B(\theta - sin\theta), \tag{1.24}$$

where θ increases with cosmic time, and A and B are two constants related by $A^3 = GMB^2$. With this substitution, the complete solution for the evolution of the over-density δ, and r and t is:

$$1 + \delta(t) = \frac{9}{2}\frac{(\theta - sin\theta)^2}{(1 - cos\theta)^3} \tag{1.25}$$
$$r \approx \frac{r_i}{2\delta_i}(1 - cos\theta) \tag{1.26}$$
$$t \approx \frac{\delta_i^{-3/2}}{2H_i}(\theta - sin\theta). \tag{1.27}$$

For $\theta = \pi$ we find the turn around radius, at which point the expansion reverses and the shell begins to collapse. We notice that the time of the turn around is proportional to $\delta_i^{-3/2}$, that is: the larger the density contrast, the sooner the collapse begins. The equation for time can also be expressed in terms of the redshift, by using the relation $t_i = 2/3H_i$, which holds during the matter-dominated era, obtaining:

$$(1 + z) = (1 + z_i)\left(\frac{4}{3}\right)^{2/3}\delta_i(\theta - sin\theta)^{-2/3}. \tag{1.28}$$

The moment of the collapse can be associated with $\theta = 2\pi$, at which point:

$$(1 + z_{\rm coll}) = 0.356\ \delta_i\ (1 + z_i). \tag{1.29}$$

This expression is particularly insightful. If we consider fluctuations of $\delta_i \sim 10^{-5}$ at $z_i \sim 1000$, as in the CMB, we find that... structures have not collapsed yet! This implies that without dark matter de-coupling earlier than $z_i \sim 1000$, the galaxies that we see now could not be there!

At the point of the collapse, the shells should converge onto a single point. However, given that the fluid is pressure-less, shells would simply go through the center and start to oscillate, as in a damped harmonic oscillator. During this stage, the collapsed material undergoes a phase of relaxation, in which the gravitational energy is converted into kinetic energy of the particles involved, eventually reaching an equilibrium condition. This final stage is called *virialization*, because the structure satisfies the virial theorem, and the collapsed structure is called a *halo*. An accurate analytical derivation provides the following expressions for the virial radius $R_{\rm vir}$, the virial temperature $T_{\rm vir}$, the circular velocity $v_{\rm vir}$ for an halo of mass M_h, collapsing at redshift z (e.g. Barkana & Loeb, 2001):

$$R_{\rm vir} = 0.784\left(\frac{M_h}{10^8 h^{-1}{\rm M}_\odot}\right)^{1/3}\left[\frac{\Omega_m\Delta_c}{\Omega_m(z)18\pi^2}\right]\left(\frac{1+z}{10}\right)^{-1} h^{-1}{\rm kpc} \tag{1.30}$$
$$v_{\rm vir} = 23.4\left(\frac{M_h}{10^8 h^{-1}{\rm M}_\odot}\right)^{1/3}\left[\frac{\Omega_m\Delta_c}{\Omega_m(z)18\pi^2}\right]^{1/6}\left(\frac{1+z}{10}\right)^{1/2} {\rm km\ s^{-1}} \tag{1.31}$$
$$T_{\rm vir} = 1.98\times 10^4\left(\frac{\mu}{0.6}\right)\left(\frac{M_h}{10^8 h^{-1}{\rm M}_\odot}\right)^{2/3}\left[\frac{\Omega_m\Delta_c}{\Omega_m(z)18\pi^2}\right]^{1/3}\left(\frac{1+z}{10}\right) {\rm K}, \tag{1.32}$$

where μ is the mean molecular weight, and Δ_c and $\Omega_m(z)$ are given by (Bryan & Norman, 1998):

$$\Delta_c = 18\pi^2 + 82(\Omega_m(z) - 1) - 39(\Omega_m(z) - 1)^2, \tag{1.33}$$

$$\Omega_m(z) = \frac{\Omega_m(1+z)^3}{\Omega_m(1+z)^3 + \Omega_\Lambda}. \tag{1.34}$$

The term Δ_c represents the average over-density of the collapsed structure with respect to the background density. It shows that approximately $\rho_{\rm vir} \approx 200\overline{\rho}$.

1.3.3 Statistics of the matter perturbations

The previous discussion provides a good description of the evolution of a single perturbation. However, the Universe is constituted of many evolving perturbations, i.e. structures, which grow and eventually merge into larger ones. Therefore, a statistical description of the halos abundance is required, with the ultimate goal to quantify the number density of halos of given mass per unit volume at each time. First of all, we assume that the density perturbation $\delta(\mathbf{x})$ is gaussian distributed, and we define the correlation between its value at two different coordinates as:

$$\zeta(\mathbf{x}) =< \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{dx}) >, \tag{1.35}$$

where the brackets indicate the expectation value. The two-points correlation function is the Fourier transform of the power spectra:

$$\zeta(\mathbf{x}) = \frac{1}{V}\sum_k P(k)e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{1.36}$$

where $P(k) =< \delta_k^2 > /V$ is the power spectra. This quantity describes the weight of the different modes of the perturbation field. It is usually expressed in the following form (e.g. Eisenstein & Hu, 1999):

$$P(k) = Ak^n T^2(k) D^2(z,k), \tag{1.37}$$

where the parameters are the following: A is a normalization constant related with the present day fluctuations; k^n is the primordial power spectrum set by the inflation (which predicts $n \simeq 1$); $T^2(k)$ is the transfer function, which accounts for modification of the primordial power spectrum due to processes at the matter-radiation equality and Recombination; $D(z,k)$ is the growth factor, which describes the linear growth of the perturbation.

Starting from the power spectra, it is possible to compute the mass variance of a perturbation on a scale R by using a window function, defined as a function $W_R(r)$ that satisfies $\int dr^3 W_R(r) = 1$. The corresponding mass variance is given by:

$$\sigma_M^2(R) \equiv < \delta_M^2 > = \frac{1}{V_W^2}\int_0^\infty \frac{dk}{2\pi^2} k^2 P(k)|W_R(k)|^2, \tag{1.38}$$

where $W_R(k)$ is the Fourier transform of the window function. The mass variance σ_M^2 is a fundamental ingredient to compute the dark matter halos abundance. The most common method used for this computation is the analytical formalism developed by Press & Schechter (1974), which express the number density of collapsed objects of mass larger than M at redshift z as:

$$\frac{dn}{dM} = \frac{\rho_0}{M}\sqrt{\frac{2}{\pi}}\frac{\delta_c(z)}{\sigma_M^2}\frac{d\sigma_M}{dM} exp\left[-\frac{\delta_c^2(z)}{2\sigma_M^2}\right], \tag{1.39}$$

where $\delta_c(z) = 1.686/D(z)$ is the critical collapse over-density threshold.

1.4 Star formation

The formation of the first gravitationally bound structures acted as a catalyst for the birth of the first galaxies, as baryons started to collapse into the potential wells of the dark matter halos. However, the physics of the baryonic collapse is fundamentally different than the one of the dark matter, because of the thermodinamical and radiative processes: the baryonic fluid has a pressure, which opposes to the collapse. As described in Section 1.3.1, the fundamental scale at which perturbations can collapse is the Jeans scal λ_J, which depends on the local density and sound speed. At $z < 30$, the comoving Jeans length is approximately $\lambda_J \approx 0.01(\Omega_{b,0}h^2)^{-1/2}$ Mpc (Mo et al., 2010), which translates into a mass of $M_J \approx 1.5 \times 10^5(\Omega_{b,0}h^2)^{-1/2}\mathrm{M}_\odot$. Therefore, the first baryonic structures to collapse were of the size of modern globular clusters.

As baryons are accreted onto the dark matter halos, they tend to reach the virial temperature of the halo due to the gravitational heating. In order for them to further collapse, they need to lose kinetic energy, and this is possible via radiative losses. Understanding the gas cooling processes in the primordial Universe is paramount in order to understand the birth of the first stars. Cooling processes allow baryons to lose kinetic energy, and then pressure, so they can collapse to much higher densities than the dark matter particles. The main cooling processes in the primordial gas fall into these categories:

- *radiative recombination*: an atom recombines in an excited state, and a non-ionizing photon is emitted in the process;
- *collisional ionization*: thermal energy of the atoms is converted into ionization energy;
- *bound-bound transition*: thermal energy is converted into excitation energy of bound-states transitions;
- *thermal bremsstrahlung*: radiation due to the acceleration of a charged particle in an electric field.

The relative efficiency of these processes depends on the gas physical conditions, in particular on its temperature T, density n and its chemical composition. In general, collisionally-excited lines, in particular of heavy elements (C, N, O, Fe) provide the main cooling mechanism in the gas, however for the pristine gas the situation is different. In Fig. 1.5, it is shown the cooling function $\Lambda(n, T)$ for a primordial gas. At temperature $T \approx 10^5$ K, helium excitation lines dominate the cooling, whereas at higher temperature the free-free processes are more important. At $T \approx 10^4$ K, which is roughly the temperature of the first virialized structures, the bound-bound transitions of atomic hydrogen constitute the most important cooling channel. However, their cooling efficiency drops rapidly below 10^4 K, so another cooling mechanism is needed for the gas to further cool down. This channel is provided by the rotational-vibrational transitions of molecular hydrogen H_2(blue peak in Fig. 1.5).

The formation of molecular hydrogen is then a fundamental step that lead to the birth of the first stars. However, given that the hydrogen atom does not have a dipole moment, it is quite difficult for two H atoms to collide and radiate away the excess thermal energy. In the present-day Universe, dust grains (see Chapter 3) act as catalysts for this process, as hydrogen atoms start to orbit around them and can efficiently radiate away their energy before collide. However, in the primordial gas, there were essentially no metals, and then no dust, so alternative channels were necessary. The most important one at $z < 100$ (for an extensive review see Loeb et al.

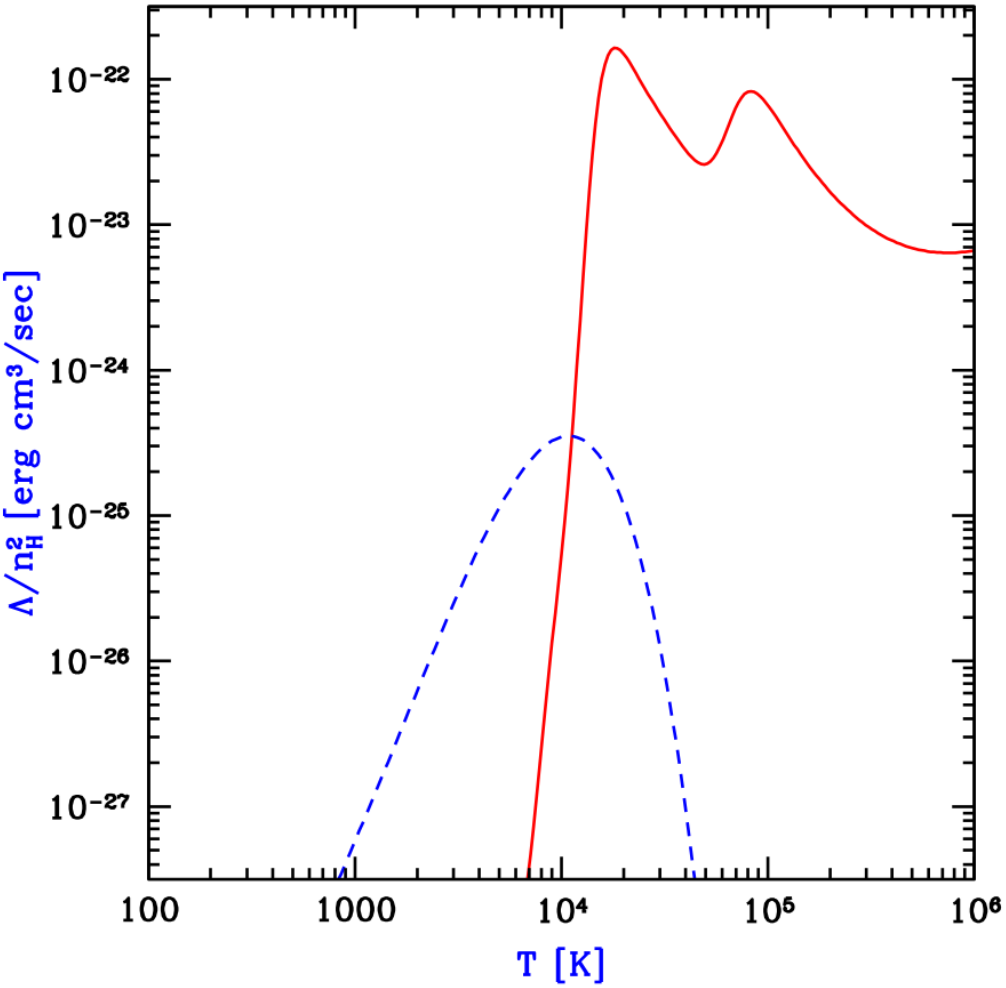

Figure 1.5: Primordial gas cooling function (divided by the square of the hydrogen density) as a function of the gas temperature. The red solid line accounts for the contribution of atomic hydrogen (first peak) and helium (second peak) excitation lines, plus the free-free processes, which dominate at the higher temperatures. The blue-dashed line shows the contribution of the molecular hydrogen, which is the only coolant of the primordial gas at temperatures below 10^4 K. *Figure from Barkana & Loeb (2001).*

2008) is the H^- channel:

$$H + H^- \rightarrow H^- + \gamma \tag{1.40}$$
$$H + H^- \rightarrow H_2 + e^-. \tag{1.41}$$

Molecular hydrogen can also be dissociated via impact with H, H_2, H^+ or electrons. However, the most important destruction mechanism is the photo-dissociation due to photons in the Lyman-Werner band ($\sim 11.2 - 13.6$ eV), which triggers the two-steps Solomon process (Draine & Bertoldi, 1996). When the ideal conditions for the H_2formation are met, gas effectively cools below $T \approx 10^3$ K and the formation of the first stars can proceed. Later on, after the ISM is enriched by heavy elements produced by stars, the cooling rate increases and the gas can further collapse, setting onto a disk on the cooling time scale, which is typically shorter than the free-fall time (e.g. Loeb et al., 2008).

The sites of star formation are Giant Molecular Clouds (GMCs), which forms under the action of self-gravity and whose size and masses are set by the Jeans scale, $\lambda_J \approx c_s t_{\rm ff}$ (here c_s is the sound speed of the disk and $t_{\rm ff}$ the local free-fall time), as discussed in Section 1.3.1. The

structure of the molecular clouds is very complex, with a dense core and filaments, embedded in a more diffuse gas component, with densities spanning a wide range, from 10^2 to 10^7 cm^{-3}. The detailed physical processes that ultimately lead to the birth of stars within these complex structures are still not fully understood, and are currently a field of intense research (see e.g. Bromm (2013); Milosavljević & Safranek-Shrader (2016) for recent reviews on the subject, an references therein).

The first generation of stars, also called Population III stars, is characterized by the essentially metal-free environment in which they form, which ultimately affect their properties and their impact on the thermal and dynamical evolution of the first galaxies. In fact, it is expected that primordial gas is less susceptible to fragmentation during the collapse, because of the low metal content. Therefore, the Initial Mass Function (IMF) of the first stars, i.e. their mass distribution at birth, is thought to be shifted toward larger masses as compared to what is observed for the later generation of stars, the Population II and I (e.g. Bromm, 2013). The first massive stars are likely very hot and very luminous, according to the typical scaling $T \propto M_*^{5/8}$ and $L \propto M_*^{7/2}$ (Dayal & Ferrara, 2018). As a consequence, the spectrum of a massive star is rich in high-energy photons capable to ionize hydrogen ($h\nu > 13.6$ eV) and, in the most-extreme cases, also helium. Moreover, they emit a large amount of radiation capable of photo-dissociating molecular hydrogen ($11.2-13.6$ eV), thereby profoundly affecting the subsequent star formation in the galaxy. Their UV radiation can also ionize, and ultimately disperse, the GMCs, halting further star formation in that region (e.g. Spitzer, 1978). Stars also affect their surrounding not only via radiation, but also via several other processes, which are usually labelled as *mechanical feedback*. Massive stars have short lifetimes ($\lesssim 30$ Myr), and they end their life with powerful explosions named *supernovae*, which cause a shock wave onto the surrounding gas, and create hot ($T \approx 10^6$ K) bubbles. They also inject an enormous amount of energy ($\approx 10^{51}$ erg), which is carried away by neutrinos. These processes overall lead to the formation of bulk motions within the cloud, to an increase of turbulence, and to a reduction of the star formation rate because of the buoyancy of the hot gas, which also tends to leave the galaxy in the form of a *galactic wind*. The picture of star formation is therefore very complex due to the interplay between the newborn stars and their environment, and how they will affect the following episodes of star formation.

The radiation of the first stars had a profound impact on the Universe. In fact, the UV photons began to ionize the neutral hydrogen in the IGM, starting the Reionization of the Universe, which we discuss in more details in Section 1.5.

1.5 The epoch of Reionization

The Reionization of the Universe is a complex multi-scale process, that involves physics from sub-galactic scales up to large cosmological distances (see Mesinger (2016) for a recent review).

The radiation emitted by the first galaxies depends on their individual star formation histories, which are determined by the star formation processes on $\sim$ pc scales. Therefore the star formation rate sets the intrinsic spectra of the galaxy, but the actual amount of UV photons that traverse the IGM depends on their escape fraction f_{esc}, which is still poorly constrained, both from theoretical models and observations (see Dayal & Ferrara (2018) and references therein), with strong variation from galaxy to galaxy.

As the UV photons travel in the IGM, they begin to ionize hydrogen atoms, forming HII regions, where the gas is photo-heated to $\approx 10^4$ K. These bubbles expand until they encounter another one and then merge together. Given that galaxy formation follows the large scale structures distribution, there are regions where the photo-ionizing photons are more abundant than

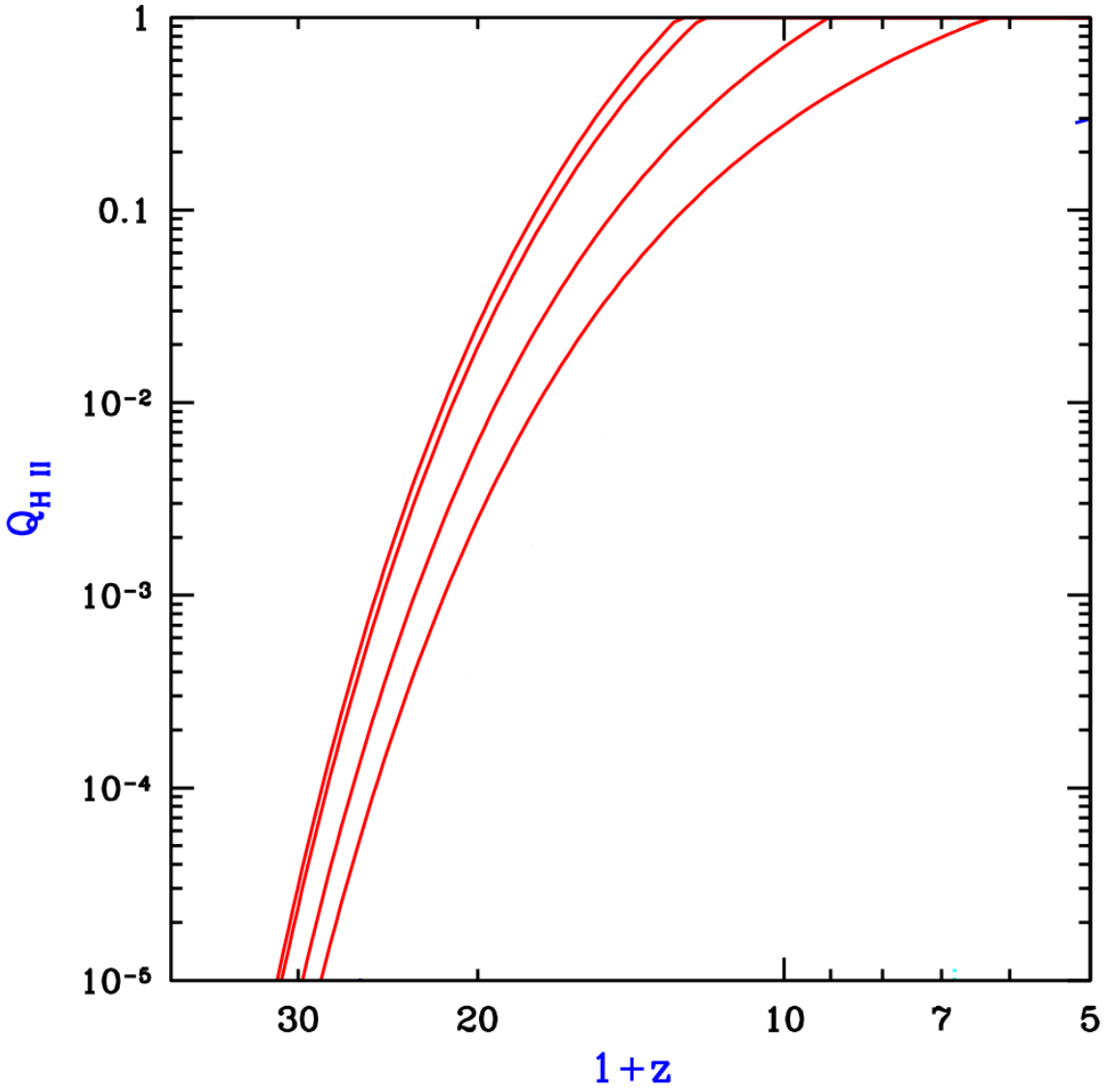

Figure 1.6: Redshift evolution of the volume filling factor $Q_{\rm HII}$ from the semi-analyical calculation by Barkana & Loeb (2001). The red solid curves represent the evolution for different values of the clumping factor $C = 0, 1, 10, 30$. *Adapted from Barkana & Loeb (2001).*

in other. Moreover, the path that the photons are able to cover depends on the gas optical depth: high density regions are able to self-shield themselves and to remain neutral (e.g. Miralda-Escudé et al., 2000; Rahmati et al., 2013). The gas density also determines the recombination rate of the ionized hydrogen, which scales as $\approx n^2$. Therefore, Reionization is a highly non-homogeneous and patchy process.

The main parameter used to describe this process is the filling factor $Q_{\rm HII}$, which represents the volume fraction filled with HII regions. In Figure 1.6, we show the analytical prediction for the evolution of $Q_{\rm HII}$ by Barkana & Loeb (2001), for different values of the clumpiness factor $C =< n_H^2 > / < n_H >^2$, which parametrizes the density fluctuations. This prediction emphasizes the role of the gas distribution on large scales, which does not only affect the local state of the Reionization processes, but also its global timing: in fact, the higher is the clumpiness factor, the more time is needed for Reionization to be completed.

Observational constraints of the EoR come mainly from two types of studies: i) the analysis of the CMB signal anisotropies; ii) the study of the spectra of distant quasars. CMB photons

interact via Thomson scattering with free electrons, whose abundance rises during the Reionization, damping the temperature fluctuations with respect to the ones at the last scattering surface. This damping is easy to detect and it provides an estimate of the integrated optical depth due to Thomson scattering τ_e, with the latest value being $\tau_e = 0.054 \pm 0.007$ (Planck Collaboration et al., 2020). However, this method provides only an integrated constraint, and a functional form for $Q_{\rm HII}$ is needed in order to derive a Reionization history. Distant quasars (see also Section 2.3) provide an additional probe thanks to the Lyman-α[6] absorption features left in their spectra (e.g. Fan et al., 2006; McGreer et al., 2015). However, these sources gives a measure of the ionization state of the IGM along a specific sightlines and therefore can be a biased probe of the IGM state. Overall, current constraints suggest Reionization to be completed at $z \sim 5-6$, but the timing of this process is still not fully understood, and it is currently a field of active research, both from the theoretical and the observational side.

In the near future, experiments such as the Square Kilometre Array (SKA)[7] and the Hydrogen Epoch of Reionization Array (HERA) (DeBoer et al., 2017) are expected to detect the cosmological 21-cm signal, which will be the most powerful probe of the Reionization, providing the opportunity to infer the neutral hydrogen distribution across the cosmic time down to the cosmic dawn, and many other physical properties of the IGM (Park et al., 2019, e.g.).

[6] The Lyman-α line corresponds to the transition of the neutral hydrogen atom from the first excited state to the ground state.

[7] https://www.skatelescope.org/

Active Galactic Nuclei at high-redshift

2

The quest for Active Galactic Nuclei (AGN) in the Universe began with the detection of a surprisingly luminous stellar-like object (Schmidt, 1963), named 3C 273, which was ~ 10 times more luminous than the brightest known galaxies at the time, and it was associated with a radio source. Other similar objects were soon found, and they were called *quasar*, which stands for quasi-stellar objects. Shortly after, it was suggested that these powerful emitting objects could be powered by accretion onto very massive black holes (Lynden-Bell, 1969), a prediction confirmed and discussed in several theoretical and observational studies since then (e.g. Soltan, 1982; Marconi et al., 2004). Today, the number of detected quasars is in the order of $\approx 10^5$ (e.g. Alam et al., 2015), and we know that super-massive black holes (SMBH, $M_{\rm BH} \sim 10^{6-10}$ $\rm M_\odot$) reside in the center of most massive galaxies ($M_\star \sim 10^{9-12}$ $\rm M_\odot$ (e.g. Magorrian et al., 1998a), possibly affecting their evolution (Kormendy & Ho, 2013). In this Chapter we will briefly review the unified model of AGN and we describe its complex emission spectrum. Then, we focus on observations of quasars at high-redshift, which is one of the main topics of this Thesis, and we discuss the main mechanisms of BH formation, and their interaction with the host galaxy.

2.1 The unified model of AGN

Following the definition of Netzer (2015), we will refer to Active Galactic Nuclei (AGN) galaxies containing a massive (10^5 $\rm M_\odot$) black hole, which is accreting at a rate higher than $L_{\rm AGN}/L_{\rm Edd} = 10^{-5}$, where $L_{\rm AGN}$ is the AGN luminosity and $/L_{\rm Edd}$ is the Eddington luminosity[1]. If the luminosity is arger than 10^{13} $\rm L_\odot$, we will call it a quasar.

We will provide here a brief description about the AGN structure and the unification models and we refer to the recent review by Netzer (2015) for more extended discussion. AGN show a large variety of observational properties, but they all present several of the following structures:

- an *accretion disk* on sub-pc scales, which regulates the gas inflow onto the central object;
- high-density clouds at a distance of $0.01 - 1$ pc, moving at large velocities (≈ 1000 km/s), thereby showing broad emission lines, the *broad line region* (BLR);
- an axis-symmetric dusty structure, named *dusty torus*, on scales $0.1 - 10$ pc;

[1]The Eddington luminosity is defined as the luminosity at which the radiation pressure from a radiating source of mass M equals its gravitational attraction exerted on the gas. It represents the maximum luminosity for which gas infall is possible. For pure ionized-hydrogen, it is given by $L_{\rm Edd} = 4\pi G M c m_p/\sigma_T = 1.26 \times 10^{38} (M/\rm M_\odot)$ erg s^{-1}, where G is the gravitational constant, m_p the proton mass, c the speed of light, and σ_T the Thomson scattering cross section for electrons.

- Low-density, low-velocity gas clouds, showing narrow emission lines, the *narrow line region* (NLR), extending from the dusty torus to $100 - 1000$ pc;
- a central radio jet;
- a hot corona surrounding the accretion disk, where electrons are up-scattered via inverse Compton by the high-energy photons emitted by the accretion disk.

Each of these components have been subject of dedicated studies across the years, and research work is still ongoing. At the the end of the last century, it was proposed that the large diversity of AGN observations could be described by few parameters, via a so-called AGN unification scheme (e.g. Urry & Padovani, 1995). A schematic representation of this model is shown in Figure 2.1. The idea is that all AGN have the depicted structure, but the luminosity of the central source and the torus inclination angle with respect to our line of sight affect the observational appearance of the AGN. For example, the torus can obscure the radiation along the line of sight the AGN is observed at, effectively making impossible to detect emission from the broad line regions. Therefore, obscured and unobscured AGN are intrisically the same objects, seen at different angles with respect to the torus. This model naturally explains the distinction between Type-I, and Type-II AGN, which are typically defined as follows. Type-I AGN are characterized by broad ($\gtrsim 1000$ km/s) emission lines[2] in the optical/UV (for example, the Balmer lines Hα, Hβ), associated to high-density ($\approx 10^9$ cm^{-3}) gas, and a central source visible in different bands. Low-to-intermediate luminosity Type-I AGN also show narrow emission lines, which are mostly absent at high-luminosity. Type-II AGN show narrow ($\approx 300 - 1000$ km/s) emission lines from the NIR to UV spectral range, typically associated to low-density ($\approx 10^2 - 10^6$ cm^{-3}) gas and an high photo-dissociation rate. The strongest ones are O III 5007 Å, [N II] 6584 Å, O II 3727 Å, O IV 25.9 μm, Ne V 3426 Å, C IV 1549 Å, C III 1909 Å, and the hydrogen Balmer and Lyman lines. Type-II AGN are associated to an X-ray point source.

The torus structure was originally thought to be characterized by a smooth matter distribution (e.g. Antonucci, 1993), with ionized gas in the inner region, but most of the gas in molecular form. The inner radius of the structure depends on the sublimation temperature of dust grains, which is $T_{\rm sub} \approx 1800$ K for graphite grains and $T_{\rm sub} \approx 1800$ K for silicate grains. From these values, it is possible to derive the minimum radius at which the radiation field from the central source allow the grains to not evaporate (e.g. Barvainis, 1987). Today, observations and theoretical works point toward a more complex torus structure, with dense clumps and a low-density interclump material (e.g. Stalevski et al., 2016; Wada et al., 2016; Hönig, 2019; Venanzi et al., 2020).

2.2 AGN emission spectrum

The complex AGN physical structure described in Section 2.1 is also reflected in their Spectral Energy Distribution (SED, which represents the radiation emitted in the different energy bands) due to the variety of physical mechanisms involved. Generally, the AGN SED can be decomposed into the following parts (e.g. Sanders et al., 1989):

[2]We use the standard notation in spectroscopy to indicate emission lines. The roman numeral after the element indicate its ionization state, for example H II is neutral hydrogen, H III is ionized hydrogen, He III is double-ionized helium and so on. A λ before the wavelength means that the line is a doublet. Brackets around a line denote the "forbiddenness" of the line, i.e. if the line violates selection rules for dipolar electric transitions (it is therefore a measure of its likelihood): no brackets are associated to *permitted* lines, a right bracket to *semi-forbidden* transitions (also called intercombination lines: they change the total spin), double brackets indicate *forbidden* ones (they do not have a single electron jump and/or change the total angular momentum).

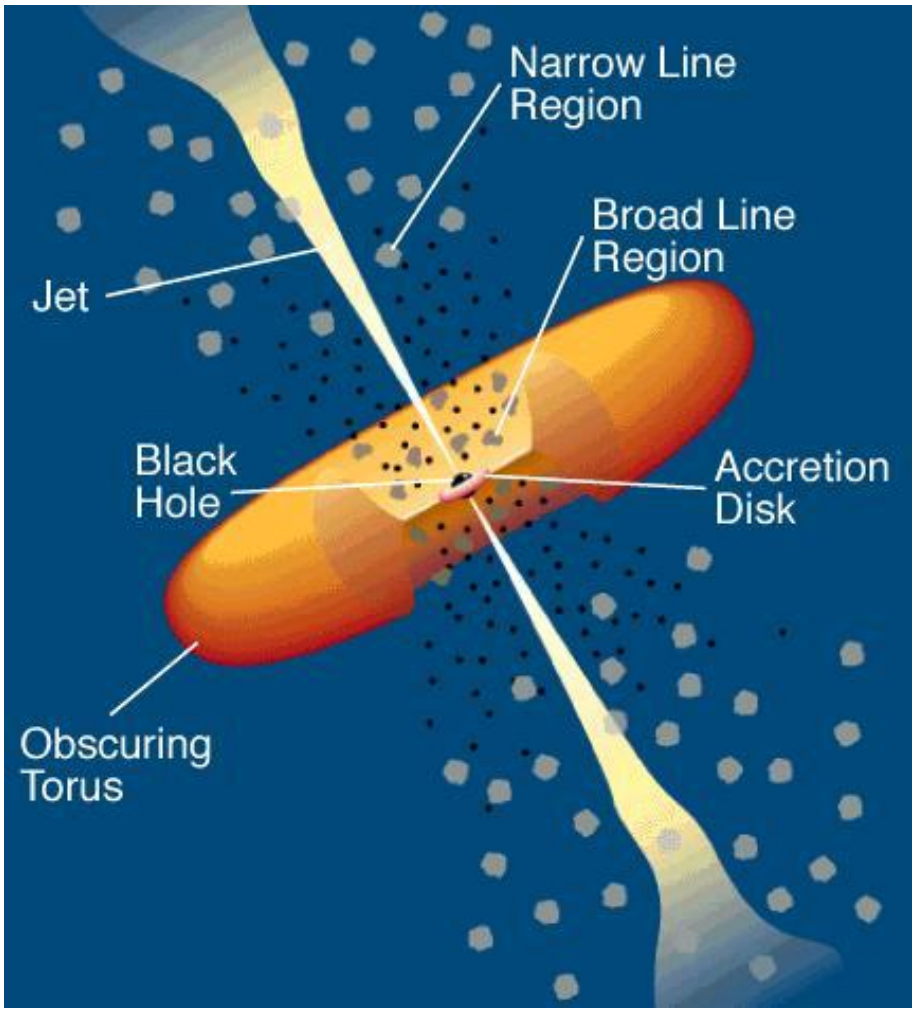

Figure 2.1: The AGN unified model proposed by Urry & Padovani (1995).

- the Big Blue Bump, in the optical/UV range 10 nm-0.3 μm;
- the Infrared Bump, at $2 - 300$ μm;
- the X-ray emission from the hot corona;
- absorption/emission lines from the BLR and NLR;
- radio emission.

We discuss them in more details in the following sections.

2.2.1 X-ray emission

High-energy photons in the X-ray band ($0.5 < E < 10$ keV) are thought to originate from the hot corona via inverse Compton scattering of the optical/UV photons produced in the accretion disk (e.g. Galeev et al., 1979). The hard X-ray ($2 - 10$ keV) part of the spectrum is often modelled as a power-law $N_{\rm ph} \propto E^{-\Gamma}$, where $N_{\rm ph}$ is the number of photons emitted at energy E per unit frequency and time. Observations in the local Universe reveal a remarkably narrow range for the photon index $\langle\Gamma\rangle = 1.8 \pm 0.2$ (see e.g. Merloni et al. 2014 and reference therein). X-ray observations can provide insights about the accretion physics of the source. For instance, the

photon index Γ can be used as a proxy of the accretion rate, with steeper slopes corresponding to higher Eddington ratios (e.g. Lusso & Risaliti, 2017).

X-ray photons can suffer absorption from gas at high column densities, due to Compton scattering and photo-electric absorption. The attenuated flux can be written as $F_\nu^{\rm att} = F_\nu^{\rm int} e^{-\tau}$, where the optical depth τ is given by $\tau = 1.2(\sigma_T + \sigma_{\rm ph})N_{\rm H}$ (e.g. Yaqoob, 1997). Here σ_T is the Thomson cross section, $\sigma_{\rm ph}$ the photo-electic cross section and $N_{\rm H}$ the hydrogen column density. Compton scattering becomes important at densities $N_{\rm H} \geq \sigma_T^{-1} \approx 1.5 \times 10^{24}$ cm^{-2} and for energies above 4 keV and dominates above 10 keV, where it adds to the photo-electric absorption. (Morrison & McCammon, 1983) computed the photo-electric absorption cross section per hydrogen atom as a function of energy in the range $[0.03 - 10]$ keV, assuming a metallicity $Z = 0.0263$. Their results are shown in Figure 2.2. The main features of the cross section are the several jumps corresponding to the K-shell energy of different elements. In fact, an element contributes to the cross section only if the energy of the photons is greater than the energy of the K-shell. We also note that the cross section is a decreasing function of the photon energy. Soft X-ray photons suffer heavier attenuation with respect to the hard ones.

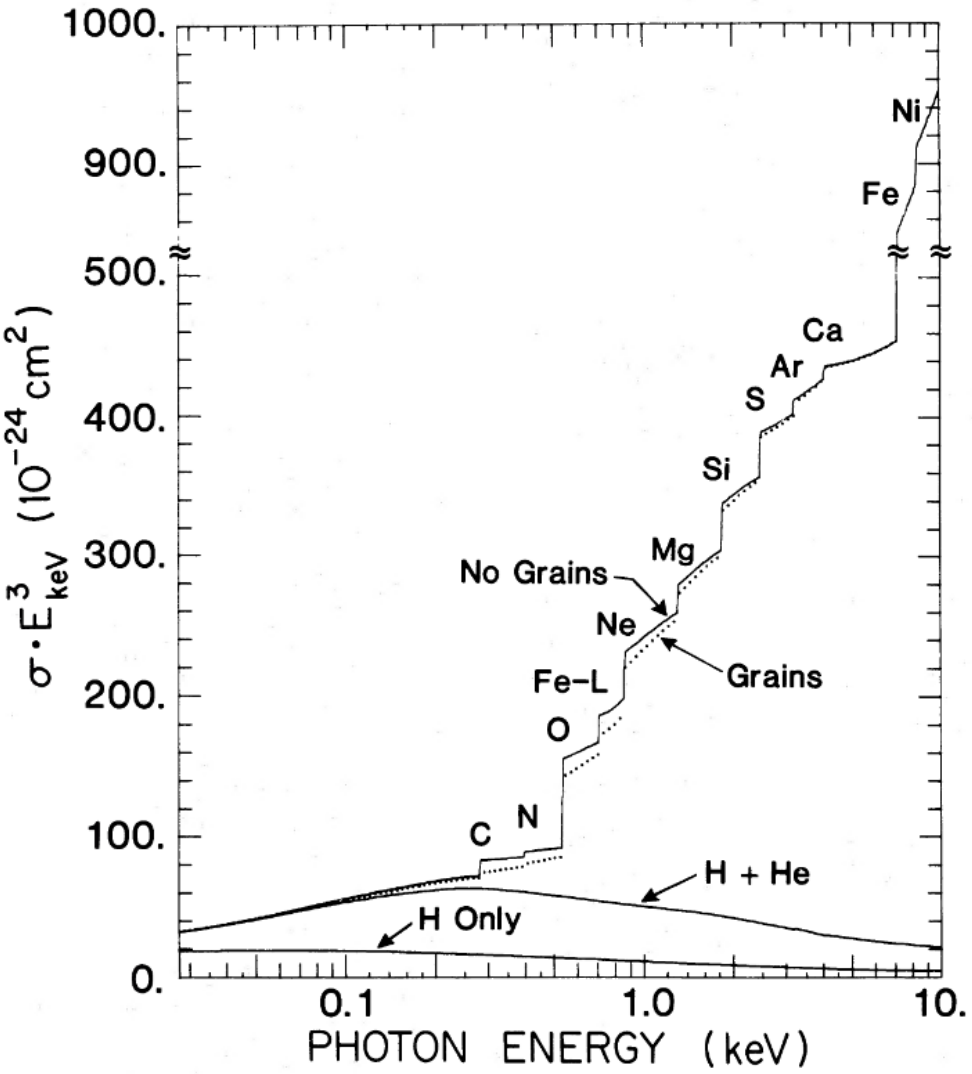

Figure 2.2: Photo-electric absorption cross section per hydrogen atom as a function of energy, scaled by $E/(1{\rm keV})^3$. The solid line indicates the cross section assuming all elements in the gas phase and in neutral atomic form. The dotted line shows the cross section assuming a fraction of the elements condensed into 0.3 μm dust grains. *Figure from Morrison & McCammon (1983).*

X-ray radiation can also be a diagnostic for the column density of a Compton-thin torus ($N_{\rm H} < \sigma_T^{-1} \approx 1.5 \times 10^{24}$ cm^{-2}, σ_T being the Thomson cross-section). The hardness of the spectra, measured as the flux in the soft-to-hard X-ray band $F_{0.5-2\ \rm keV}/F_{2-10\ \rm keV}$ has been shown to be a good proxy (≈ 0.3 dex) for the hydrogen column density in the range $10^{21} - 1.25 \times 10^{24}$ cm^{-2} (e.g. Ueda et al., 2003). In their analysis of a sample of 1310 sources, Merloni et al. (2014) found a column density of $N_{\rm H} \approx 10^{21.5}$ cm^{-2}2 to divide optically obscured and unobscured AGN reasonably well, with 30% of the type-I sources having lower values and a similar fraction of type-II ones larger column densities. The relative importance of the hot corona and the accretion disk to the overall emission is usually expressed via the parameter $\alpha_{\rm OX} = 0.384 \log L_\nu(2{\rm keV})/L_\nu(2500\text{Å})$ (e.g. Brandt & Alexander, 2015). This parameter has been shown to anti-correlate with the quasar UV luminosity (see for example Lusso & Risaliti 2017 and references therein).

Observations at high redshift found no significant evolution of the parameters Γ and $\alpha_{\rm OX}$, suggesting that the accretion physics remains the same across cosmic time (e.g. Vito et al., 2019b).

2.2.2 Radio emission

Most AGN also show radio emission: interestingly enough, many AGN were first identified as radio galaxies before being detected in the optical, with the exception of the historical discovery by Seyfert (1943). They are typically divided into two categories, *radio-loud* AGN and *radio-quiet* AGN, depending on their "loudeness" R, which measures the intensity of the radio flux with respect to the optical one in the B-band[3] (Netzer, 2013):

$$R = \frac{L_\nu(5GHz)}{L_\nu(4400\ \text{Å})} = 1.36 \times 10^5 \frac{L(5GHz)}{L(4400\ \text{Å})}. \tag{2.1}$$

AGN with $R < 10$ are usually considered radio-quiet, whereas above this threshold they are called radio-loud. The radio emission comes from a central core, and it is thought to be produced via self-absorbed synchrotron emission. It is usually well-modelled with a power-law $F_\nu \propto \nu^{-\alpha_r}$, with sources characterized by $\alpha_r < 0.5$ referred to as *flat-spectrum* radio sources and those with $\alpha_r > 0.5$ as *steep-spectrum* radio sources. These two categories also show a distinct morphology in their radio emission, with steep-spectrum sources presenting radio lobes, whereas flat-spectrum sources an high luminosity cores but no lobes.

2.2.3 UV/optical emission: the Big Blue Bump

SMBHs accrete gas via an accretion disk from which thermal emission is generated through dissipative processes (e.g. Shakura & Sunyaev, 1973; Rees, 1984). The accretion disk provides a large fraction ($\sim 10 - 50\%$) of the AGN bolometric luminosity, emitting into optical/ultraviolet (UV) wavelength range (Hopkins et al., 2007; Lusso et al., 2015; Shen et al., 2020a), producing the so called Big Blue Bump in AGN spectra (e.g. Sanders et al., 1989).

The spectrum of an optically-thick accretion disk in which viscous forces are the main source of energy dissipation can be derived with a simple analytical model (e.g. Frank et al., 2002). For such a disk, the temperature distribution follows:

$$T(R) = T_* \left(\frac{R}{R_*}\right)^{-3/4} [1 - (R_*/R)^{1/2}]^{1/4}, \tag{2.2}$$

[3]The B-band is associated to the typical B-filter, which is centered at $\lambda = 445$ nm, with a full-width half maximum 94 nm.

where R_* is the radius of the central object and

$$T_* = \left(\frac{3GM_*\dot{M}_*}{8\pi R_*^3 \sigma_{\rm SB}} \right)^{1/4}, \tag{2.3}$$

with M_* and $\dot{M}_*$ being respectively the mass and the accretion rate of the compact object and $\sigma_{\rm SB}$ is the Stefan-Boltzmann constant. Assuming that each rink of the disk emits as a black body with a temperature distribution $T(R)$ as in eq. 2.2, the spectrum is proportional to:

$$S_\nu \propto \int_{R_{in}}^{R_{out}} \frac{\nu^3}{e^{h\nu/kT(R)} - 1} R dR, \tag{2.4}$$

where R_{in} and R_{out} are the inner and outer radii of the accretion disk. At high frequencies, $h\nu \gg kT_*$, the emission falls down exponentially. At low frequencies, $h\nu \ll kT_*$, $T \propto R^{-3/4}$ and the spectrum can be written as:

$$S_\nu \propto \nu^{1/3} \int_0^{h\nu/kT_{out}} \frac{x^{5/3} dx}{e^x - 1}. \tag{2.5}$$

For intermediate frequencies, $kT_{out} \ll h\nu \ll kT_*$, $S_\nu \propto \nu^{1/3}$, whereas at lower frequencies, $h\nu \ll kT_{out}$, $S_\nu \propto \nu^2$. Summarising, the accretion disk spectrum can be modelled as:

$$S_\nu \propto \begin{cases} \nu^2 & \text{for } h\nu \ll kT_{out} \\ \nu^{1/3} & \text{for } kT_{out} \ll h\nu \ll kT_* \\ e^{-\nu} & \text{for } h\nu \gg kT_* \end{cases} \tag{2.6}$$

From the observational side, the emission in the optical/UV was historically modelled as a single power-law emission in the form $F_\nu \propto \nu^{\alpha_\nu}$. However, a variety of slopes were found depending on the wavelength range considered. Vanden Berk et al. (2001) found $\alpha_\nu = -0.44 \pm 0.1$ for $1300 - 5000$ Å, whereas (Richards et al., 2003) found a distribution in the range $-1.8 < \alpha_\nu < 0.4$, peaking around $\alpha = -0.4$, for a sample of 4576 of $z \lesssim 2.2$ quasars analyzed in the range $1200 \lesssim \lambda \lesssim 6000$ Å. Lusso et al. (2015) found $\alpha_\nu = -0.61 \pm 0.1$ for $912 - 2500$ Å. At higher redshifts, Gallerani et al. (2010) analysed 33 quasars in the redshift range $3.9 \lesssim z \lesssim 6.4$ finding that unreddened quasars are characterised by $\alpha = -1.7 \pm 0.5$, whereas reddened quasars prefer steeper slopes ($\alpha < -2.3$).

2.2.4 Infrared emission

Optical and UV photons from the accretion disk might be absorbed by the dusty structures in the torus around the AGN and then re-emitted in the infrared (e.g. Sanders et al., 1989). This emission can be described as a power-law $F_\nu \propto \nu^\alpha$ in the range $1 - 10$ μm, with the slope ranging $-2 \lesssim \alpha \lesssim -0.5$ (Elvis et al., 1994; Barmby et al., 2006). Given that normal star forming galaxies have a mid-IR slope $\alpha \approx 2$, the mid-IR emission have been used in several works to select AGN in IR surveys (e.g. Lacy et al., 2004; Chung et al., 2014)

2.2.5 Absorption and emission lines

On top of the continua, AGN show a variety of emission and absorption lines, making their spectra an invaluable source of information for the gas composition and physical properties (density, metallicity, ionization state and so on) in their surrounding.

Clouds in the BLR are characterized by high-density ($n_H \approx 10^{9-14}\,\mathrm{cm}^{-3}$) and high-temperature ($T \approx 10^4$ K) gas (e.g. Peterson, 2006). They are partially ionized ($U \approx 10^{-24}$), with the most abundant ions being He II, He III, O IV, O VI , C III, C IV, causing the production of Hα, Lyα, C IV λ1549 Å, O VI λ1035 Å. The density in the BLR is high enough to suppress the forbidden lines, but not all the semi-forbidden ones, like C III] λ1909 Å and O III] λ1663 Å. The neutral regions in the BLR produce strong lines of H I, Mg II, Fe II, with equivalent widths depending on the emissivity of the central source and on covering factor of the clouds.

More distant gas in the NLR is instead characterized by smaller column densities (($n_H \approx 10^{20-21}\ \mathrm{cm}^{-2}$)) and densities ($n_H \approx 10^4\ \mathrm{cm}^{-3}$). The column density in the clouds can be low enough to make them optically thin to helium and hydrogen Lyman continuum, therefore the mean level of ionization can be higher than in the BLR. Moreover, the radiation field is lower, allowing the formation of dust in the molecular gas (see Chapter 3). This low-density gas tends to produce intense forbidden lines, and also coronal lines from highly ionized species. Low-ionization lines, like Fe IIor Mg II, are instead expected to be very weak. Common lines from NLR are [O III] λ 5007 Å, [[N II]] λ 6584 Å, [O II] λ 3727 Å, [O IV] λ 25.9 μm, [Ne V] λ 3426 Å, C IV λ 1549 Åand the hydrogen Balmer and Lyman lines. Examples of high-ionization, narrow emission lines are [Ne VI] 7.66 μm, [Si VI] 1.9 μm in the mid-infrared, and Fe X, Fe XI, Fe XIVin the optical. They are characterized by very large value of U. These lines are referred to as *coronal lines*, and the regions where they are produced is called the *coronal-line region*. Most of the coronal lines are emitted by gas located between the NLR and the BLR. The lines from the highest ionized species originate mostly from regions closer to the central source.

LINERs are AGN characterized by low-ionization, narrow emission lines, such as [[N II]] λ 6584 Å, [S II] λ 6731 Å, [[I]] λ 6300 Åand the Balmer lines. They can be divided into type-I (broad emission lines) and type-II (only narrow emission lines) LINERs.

The relative strength of different ionization lines can also be used to infer the gas physical conditions. Several diagnostic diagrams (also called BPT-diagrams, from the seminal work by Baldwin et al. 1981) have been used to estimate the ionizing flux by comparing the different line ratios, in particular [O III] λ 5007 Å/Hβ, [N II] λ 6584 Å/Hα, [S II] λ 6731 Å/Hα, [I] λ 6300 Å/Hα (e.g. Veilleux & Osterbrock, 1987; Kauffmann et al., 2003; Kewley et al., 2006). These diagrams show well-defined regions that allow to distinguish high-ionization AGN, LINERS, and normal star forming galaxies.

Emission lines can also be used to estimate the black hole mass (see also Section2.5) (e.g. Vestergaard, 2002; Vestergaard & Osmer, 2009) and the gas chemical composition (e.g. Hamann et al., 2002; Matsuoka et al., 2011).

AGN also show absorption lines in their spectra, which are also used to classify them. Broad absorption line (BAL) AGN are type-I AGN, whose spectrum present deep, blue-shifted absorption features. They are divided into two categories: high-ionization BAL AGN, if they show high-ionization strong resonance lines such as C IV λ 1549 Å, Si IV λ 1397 Å, Ne V λ 1240 Å, O VI λ 1035 Å, Lyα, and low-ionization BAL AGN, which show Mg II λ 2798 Åand Fe II. The presence of BAL can be explained by assuming that high-luminosity AGN contain outflowing gas with very large velocities (up to 20% of the speed of light).

AGN can also show narrow absoprtion lines (NALs), typically due to the same resonance

[4]The ionization parameter U is a measure of the ionizing photon density with respect to the gas density. It can be defined as $\int_{E_1}^{E_2} \frac{(L_E/E)dE}{4\pi r^2 c N_\mathrm{H}}$ (e.g. Netzer, 2013, section 5.4), where L_E is the monochromatic luminosity at energy E r is the distance from the ionizing source, N_H is the hydrogen column density. The energies E_1 and E_2 depends on the considered species: for the hydrogen, it is common to set $E_1 = 13.6$ eV and E_2 to infinite. Alternative definitions express the ionization parameter in terms of the flux density, but its physical meaning is essentially the same.

transitions as in BAL AGN, but associated with gas at lower velocities ($\approx 100 - 1000$ km/s).

2.3 The quest for high-redshift quasars

As discussed in Section 2.2, accreting black holes release a significant fraction of their bolometric luminosity in the optical/UV spectra (Hopkins et al., 2007; Lusso et al., 2015; Shen et al., 2020a, e.g.). In the EoR, these photons are redshifted into the near-infrared (NIR) bands, making surveys in these windows the natural strategy for AGN searches.

Over the last two decades, thanks to several optical/NIR surveys, such as the Sloan Digital Sky Survey (SDSS; Fan et al. 2006; Jiang et al. 2009), the UKIDSS Large Area Survey (Venemans et al., 2007), the Canada-France High-z Quasar Survey (CFHQS; Willott et al. 2007), the VISTA Kilo-Degree Infrared Galaxy Survey (VIKING; Venemans et al. 2013, 2015), Pan-STARRS1 (Bañados et al., 2014), the Very Large Telescope Survey Telescope ATLAS survey (Carnall et al., 2015), the Dark Energy Survey (DES; Reed et al. 2015, and the Subaru High-z Exploration of Low-luminosity Quasars (SHELLQs; Kashikawa et al. 2015; Matsuoka et al. 2016), more than 200 quasars have been discovered at the most distant redshifts probed so far ($z \sim 6 - 8$, Mortlock et al. 2011; Bañados et al. 2018; Wang et al. 2018b, 2021a). Follow-up NIR spectroscopical observations of emission lines (e.g. Mg II and C IV) produced by Broad Line Region clouds have confirmed that these sources are powered by $\sim 10^8 - 10^{10}$ $M_\odot$ BHs (Fan et al., 2000; Willott et al., 2003; Kurk et al., 2007; Jiang et al., 2007; Wu et al., 2015; Bañados et al., 2018; Yang et al., 2020b; Wang et al., 2021a). The challenge is to understand how SMBHs have formed in < 1 Gyr, namely the age of the Universe at $z \sim 6$. Theoretical models of black hole accretion are in fact facing serious difficulties in explaining such a rapid growth (e.g. Volonteri et al., 2003; Tanaka & Haiman, 2009; Haiman, 2013; Pacucci et al., 2015a; Lupi et al., 2016), also including the rather uncertain formation mechanism of SMBH seeds (Shang et al., 2010; Schleicher et al., 2013; Latif et al., 2013; Ferrara et al., 2014; Latif & Ferrara, 2016; Inayoshi et al., 2020). We will discuss this point in more details in Section 2.5.

The problem is exacerbated by the unsuccessful search for high-z AGN powered by $\sim 10^{6-7}$ $M_\odot$ BHs (e.g. Xue et al., 2011; Cowie et al., 2020). Whether these sources are too rare (Pezzulli et al., 2017a), and/or too faint to be detected by current optical/NIR survey (Willott et al., 2010; Jiang et al., 2016; Pacucci et al., 2016; McGreer et al., 2018; Matsuoka et al., 2018b; Wang et al., 2019; Kulkarni et al., 2019), and/or their optical/UV emission is obscured by dust (see Chapter 3), remains unclear. This latter hypothesis is supported by at least two observational results: (i) multi-wavelength studies of ~ 1000 local AGN show a decrease in the covering factor of the circumnuclear material with increasing accretion rates due to the increase of the dust sublimation radius of the obscuring material with incident luminosity (e.g. Ricci et al., 2017a); (ii) X-ray observations provide indications that the fraction of obscured AGN increases with redshift (e.g. Vito et al., 2014, 2018), an evidence further supported by studies of Lyα absorption profiles of distant quasars (e.g. Davies et al., 2019). Both these facts resonate with the expectation that early growth of SMBHs, typically characterized by low accretion rates, is buried in a thick cocoon of dust and gas (e.g. Hickox & Alexander, 2018, for a review on this subject).

Signatures of the complex interplay between AGN/stellar radiation and dust grains remain imprinted in the rest-frame UV-to-FIR spectral energy distribution (SED) of galaxies. Therefore, multi-wavelength SED analysis of galaxies and AGN can be used to infer information on their dust properties (mass, temperature, grain size distribution, composition), to shed light on their star formation and nuclear activities, and to quantify the relative contribution of stars and AGN radiation to dust heating (Bongiorno et al., 2012; Pozzi et al., 2012; Berta et al., 2013; Gruppioni et al., 2016). Telescopes sensitive to Mid-Infrared (MIR, $5 \lesssim \lambda_{\rm RF} \lesssim 40$ μm), like

Spitzer (Werner et al., 2004) and Herschel (Pilbratt et al., 2010), and to Far-Infrared (FIR, $45 \lesssim \lambda_{\rm RF} \lesssim 350$ μm) wavelengths (e.g. ALMA, NOEMA) have made possible to study the panchromatic SED of bright ($M_{UV} \lesssim -26$) quasars at $z \sim 6$.

SEDs observations obtained with Herschel and Spitzer in these sources (Leipski et al., 2013, 2014) have been used to disentangle the star formation versus AGN contribution to the total restframe IR emission (TIR, $8 < \lambda_{\rm RF} < 1000$ μm). The result of this study is that star formation may contribute $25 - 60\%$ to the bolometric TIR luminosity, with strong variations from source to source. In particular, Leipski et al. (2014) performed a multi-component SED analysis on a sample of 69 $z > 5$ quasars, finding that a clumpy torus model needs to be complemented by an hot (~ 1300 K) dust component to match the NIR data, and by a cold (~ 50 K) dust component for the FIR emission. This work shows that, in addition to the standard AGN-heated component, a large variety of dust conditions is required to reproduce the observed SED. Yet these kinds of studies are limited to a small sample of bright sources. Future facilities in the rest-frame MIR, such as the proposed Origins Space Telescope (OST; Wiedner et al. 2020) with a sensitivity ~ 1000 higher than its precursors Spitzer and Herschel, will significantly improve our knowledge of dusty galaxies in the Epoch of Reionization.

ALMA and NOEMA observations have provided the opportunity of studying the ISM/CGM properties of bright $z \sim 6$ quasar hosts (e.g. Carilli & Walter, 2013; Gallerani et al., 2017a), by means of rest frame FIR emission lines, as the [CII] line at 158 micron (e.g. Maiolino et al., 2005; Walter et al., 2009; Wang et al., 2013; Venemans et al., 2016; Novak et al., 2019), CO rotational transitions (e.g. Bertoldi et al., 2003b; Walter et al., 2003; Riechers et al., 2009; Gallerani et al., 2014; Venemans et al., 2017a; Carniani et al., 2019a; Li et al., 2020), and the corresponding dust continuum emission (e.g. Bertoldi et al., 2003a; Venemans et al., 2016, 2017b; Novak et al., 2019). These observations have shown that these massive galaxies ($M_{\rm dyn} \sim 10^{10} - 10^{11}$ M$_\odot$) are characterized by high star formation rates ($SFR \sim 100 - 1000$ M$_\odot$ yr^{-1}), and large amount of molecular gas ($\sim 10^{10}$ M$_\odot$) and dust ($\sim 10^{8}$ M$_\odot$), that are typically distributed on galactic scales ($\lesssim 5$ kpc). In some exceptional cases (Maiolino et al., 2012; Cicone et al., 2015), the extension of [CII] emitting gas has been detected up to CGM scales ($\sim 20 - 30$ kpc), possibly driven by fast outflowing gas ($v_{\rm out} \gtrsim 1000$ km s^{-1}) with extreme mass outflow rate of $\dot{M}_{\rm out} \sim 1000$ $M_\odot$ yr^{-1}. Recently, Bischetti et al. (2022) studied high-resolution spectra of 30 quasars at $5.8 \leq z \leq 6.6$ in the context of the Ultimate X-shooter Legacy Survey of Quasars (XQR-30) project (Wang et al., 2021a). They find finding blue-shifted BAL emission lines in 50% of the sample. Their results indicate that powerful outflows (with velocities up to 17% of the speed of light) can be a common phenomenon at high-redshift, with an occurrence a factor of $2 - 3$ higher with respect to quasars at $z \approx 2 - 4.5$.

These FIR data, combined with X-ray and UV observations, have shown the presence of galaxy/AGN companions in the field of $z \sim 6$ quasars. In the X-ray, whereas the fraction of dual AGN can be as high as $\sim 40 - 50\%$ out to $z \sim 4.5$ (e.g. Koss et al., 2012; Vignali et al., 2018; Silverman et al., 2020), at $z \sim 6$, there are only tentative X-ray detections of double systems (e.g. Vito et al. 2019a; Connor et al. 2019, but see also Connor et al. 2020.)

The occurrence of UV detected and sub-mm galaxy (SMG) companions is instead more frequent. Marshall et al. (2020) detected up to nine companions with $-22 \lesssim M_{\rm UV} \lesssim -20$ in the field of view of six quasars at $z \sim 6$ (but see also Mechtley et al., 2012). Decarli et al. (2017) reported the [CII] line and 1 mm continuum ($F_{\rm cont}$) detection of SMGs close to 4 (out of ~ 20) quasars at $z \sim 6$, with $0.2 \lesssim F_{\rm cont} \lesssim 2.0$ mJy, and reported projected distances between ~ 8 and ~ 60 kpc. More recently, Venemans et al. (2020) find that 13 out of their studied 27 quasars have companion galaxies in their field, at projected separations of $3 - 90$ kpc. ALMA data of $z \sim 8$ Lyman Break Galaxies (Laporte et al., 2017; Bakx et al., 2020) have suggested the presence in

these sources of dust hotter than expected ($T \sim 60 - 90\ K$, Behrens et al. 2018; Arata et al. 2019; Sommovigo et al. 2020). The origin of warm dust in early galaxies can be traced back to their (i) large SFR surface densities that favour an efficient heating of dust grains (Behrens et al., 2018) and (ii) more compact structure of molecular clouds (MC) that delays their dispersal by stellar feedback, implying that a large fraction ($\sim 40\%$) of the total UV radiation remains obscured (Sommovigo et al., 2020). Another possibility concerns the presence of obscured, accreting, massive ($\sim 10^8\ M_\odot$) BHs, whose UV luminosity is absorbed by dust located in the ISM of the host ($\lesssim 1$ kpc) and/or into a central obscurer, closer to the active nuclei (~ 1 pc), and heated to temperatures as high as 80-500 K, respectively (Orofino et al., 2021). According to this scenario, buried AGN should be searched for among Lyman break galaxies (LBGs) populating the bright-end of their UV luminosity function ($-24 < M_{\rm UV} < -22$), where indeed a large fraction of objects consists of spectroscopically confirmed AGN (Ono et al., 2018).

Obscured AGN may therefore represent a bridge between LBGs and bright quasars in the galaxy formation process.

2.4 Black hole - galaxy co-evolution

In the last two decades, several observations in the local Universe have revealed that black hole masses correlate with the host-galaxy properties (e.g. Magorrian et al., 1998b; Ferrarese & Merritt, 2000; Gebhardt et al., 2000; Marconi & Hunt, 2003; Gültekin et al., 2009). In particular, the BH masses scale approximately as $M_{\rm BH} \approx 0.001\ M_{\rm bulge}$, where $M_{\rm bulge}$ is the bulge mass of the galaxy. Moreover, it also correlates with the velocity dispersion σ of the stars in the galactic bulge, with a scaling $M_{\rm BH}/(10^8\ {\rm M}_\odot) \approx (\sigma/200\ {\rm km\ s^{-1}})^4$, but recent observations suggest a steeper one $M_{\rm BH} \propto \sigma^5$ (McConnell et al., 2011). These findings are somehow surprising, as the BH radius of influence ($r_{\rm in} = GM_{\rm BH}/\sigma^2 \approx 1 - 100$ pc) is several orders of magnitude smaller than the typical size of galactic bulges. Nevertheless, they suggest a casual link in the formation and growth of the black hole masses and the bulges, which is typically referred to as coevolution (see Kormendy & Ho 2013 for an extensive review on this subject).

These observations motivated the idea that BHs could interact with the host galaxy by injecting energy (*radiative feedback*, e.g. Silk & Rees 1998) or momentum (*mechanical feedback*, e.g King 2003) onto the surrounding gas, effectively removing or heating cold gas, thus preventing further star formation. These models were also able to partially reproduce the observed $M_{\rm BH} - \sigma$ relation. The general picture of AGN feedback can be described as follows. AGN activity is triggered by cold gas accretion and generate feedback, which drives the gas away, thus halting accretion itself, and possibly quenching the star formation in the galaxy. Then, cold gas is able to form again, re-igniting both star formation and further accretion onto the central engine. As a result, an oscillatory behaviour between starburst and AGN activity is expected to occur (Binney & Tabor, 1995; Ciotti & Ostriker, 2007, e.g.), in which star formation and BH accretion affect each other (see also Harrison 2017 for a broader review and reference therein).

Additional observations also support the existence of feedback from AGN, in particular: the relation between the X-ray luminosity and X-ray temperature of the gas in the intra-cluster medium (Markevitch, 1998), the low gas cooling rate in cluster (Fabian, 1994), the presence of high-velocity winds ($\approx 10\%$ of the speed of light) from brightest quasars (Tombesi et al., 2010), the star formation inefficiency in the most massive halos (Behroozi et al. 2013, see also left panel in Fig. 2.3), the similar histories of BH growth and star formation in the Universe (Madau & Dickinson 2014, see also right panel in Fig. 2.3). Moreover, most of galaxy formation models require AGN to inject energy or momentum into the surround gas in the most massive ($M_* \gtrsim 10^{10}\ {\rm M}_\odot$) galaxies (Croton et al., 2006; Somerville et al., 2008a; Ciotti et al., 2010). AGN

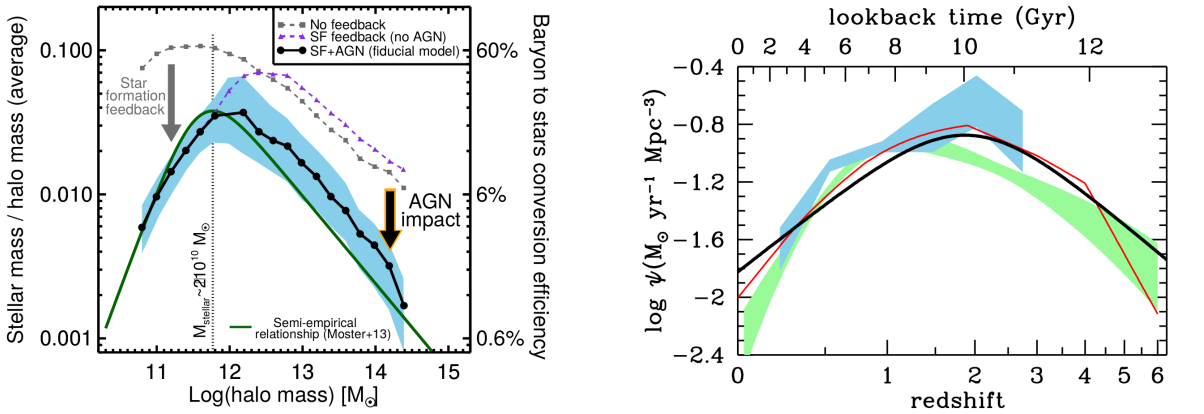

Figure 2.3: *Left panel.* Ratio of the stellar mass to the total halo mass a function of halo mass in semi-empirical relations (green line, Moster et al. 2013) and in semi-analytical models with and without feedback (black and grey lines, Somerville et al. 2008b). The shaded area in cyan indicates the 16th and 84th percentiles of the fiducial model including both AGN feedback and star formation. The impact of stellar feedback is to reduce the conversion of baryons into stars in low mass halos, whereas AGN feedback is invoked to do the same in the most massive halos. *Figure taken from Harrison (2017). Right panel.* Comparison of the star formation rate history (black line) with the black hole accretion history deduced from X-ray data from Shankar et al. (2009) (red line), (Aird et al., 2010) (light green shaded area), (Delvecchio et al., 2014) (light blue shaded area). The comoving black hole accretion rates have been scaled up by a factor of 3300 for a better visual comparison. *Figure taken from Madau & Dickinson (2014).*

produce an incredible amount of energy by releasing $\approx 10\%$ of the accreted rest-mass energy as radiation. For a 10^8 $M_\odot$-BH approximately 10^{54} J are released, ten orders of magnitude more than the typical energy released by a supernova, and three orders of magnitude higher than the typical binding energy of the galaxy. Therefore, only a small fraction of this energy needs to be coupled with the gas in order for the BHs to produce winds or outflows (e.g. Harrison et al., 2018), able to affect the dynamics and the star formation in the host galaxy (e.g. Harrison, 2017). Observations of outflows estimated coupling efficiencies between the kinetic power of the outflow and the AGN luminosity of $\approx 0.001\% - 10\%$ (see for example Fig. 2 in Harrison et al. 2018). However, a large number of observational challenges in inferring the outflows properties still remain, in particular at high redshift (see also Section 2.3). Moreover, outflows can be effectively in quenching the activity in the nuclear regions of galaxies, but how their impact manifests on larger scales is still not understood (e.g. Costa et al., 2014b). We are still missing an observational "smoking gun" that unequivocally proves the AGN impact on the star formation activity in the host galaxy.

The interpretation of the $M_{\rm BH} - M_{\rm bulge}$ relation is also complicated by its evolution with redshift. In particular, at high redshift, BH masses are found to lie above the local relations (e.g. Wang et al., 2010; Neeleman et al., 2021), with a larger scatter with respect to the relation at lower redshifts. Instead, BHs in SMGs at $z \sim 2$ tend to be under-massive. Galaxy mergers have been invoked to explain this correlation (e.g. Peng, 2007; Cattaneo et al., 2009). In fact, this relation is well established in bulges and elliptical galaxies, but not in pseudo-bulges and disks-dominated galaxies, which are not remnant of major mergers. Moreover, multiple galaxy mergers might also average out the properties of the involved galaxies, tightening the relation at lower redshift. Whatever the physical mechanisms responsible for the $M_{\rm BH}$-relations are, they must have been already in place in the early Universe.

2.5 The formation of the first black holes

As mentioned in Section 2.3, several observations reveal that the quasars at $z \gtrsim 6$ are powered by accretion onto BHs with masses of $\sim 10^8 - 10^{10}$ $M_\odot$ BHs (e.g. Mortlock et al., 2011; Wu et al., 2015; Bañados et al., 2018; Yang et al., 2020b; Wang et al., 2021a). These estimates are based on calibrations between the full width at half-maximum (FWHM) of bright BLR emission lines, in particular Mg II λ 2798 Åand C IV λ 1549 Å, with the BHs masses (see e.g. Netzer, 2013). The evidence of SMBHs already in place within the first Gyr of the Universe is one of the biggest puzzles in modern astrophyiscs (e.g. Tanaka & Haiman, 2009; Volonteri, 2010; Schleicher et al., 2013; Latif & Ferrara, 2016; Valiante et al., 2017; Woods et al., 2019; Inayoshi et al., 2020).

Why their existence is surprising can be illustrated as follows. Mass accretion onto a BH is regulated by the mass inflow rate $\dot{M}$ and by the radiative efficiency ϵ_r, which quantifies how much mass is converted into energy during the accretion process. Assuming that the mass inflow rate is limited by the Eddington luminosity $L_{\rm Edd}$ (see footnote [1]), the mass growth rate is regulated by the following equation:

$$\dot{M}_{\rm BH} = \lambda_{\rm Edd} \frac{1-\epsilon_r}{\epsilon_r} \frac{L_{\rm Edd}}{c^2}, \tag{2.7}$$

where $\lambda_{\rm Edd}$ is the ratio between the accretion luminosity and the Eddington luminosity. If the accretion proceeds at the maximum rate for the whole lifetime of the BH (i.e. $\lambda_{\rm Edd} = 1$, the growth proceeds exponentially as $M(t) = M(0)e^{t/t_e}$, with an e-folding timescale, given by:

$$t_e = \frac{\sigma_T c}{4\pi\, G\, M\, m_p} \frac{\epsilon_r}{1-\epsilon_r} = t_S \frac{\epsilon_r}{1-\epsilon_r}, \tag{2.8}$$

where we introduced the *Salpeter time* $t_S \approx 0.45$ Gyr. For a typical value $\epsilon_r \simeq 0.1$, the characteristic growth time scale is $t_S \approx 0.045$ Gyr, which implies that for a BH seed of $M(0) \approx 10^2$ $M_\odot$ (a reasonable mass for a BH originating from a Pop. III star remnant), approximately 20 e-foldings are required to reach 10^9 $M_\odot$, corresponding to 700 Myr. Therefore this light seed requires constant mass inflow at the maximum possible rate, essentially for the entire lifetime of the Universe since its formation, which is a very unlikely scenario. It seems necessary that more massive seeds are required, which would have a jump start in the growth (see Fig. 2.4). We know briefly review the main BHs formation channels proposed in literature.

Since the seminal work by Rees (1984), several mechanisms leading to BH formation have been proposed, which can be divided into three categories:

- BH formation from the collapse of a massive star;
- BHs originated after runaway collapses in stellar clusters;
- Direct Collapse Black Holes (DCBHs): BHs fromed from the monolithic collapse of the gas within a dark matter halo.

Primordial stars in the mass range $40 - 140$ $M_\odot$ M or above 260 $M_\odot$ are expected to collapse into BHs of similar masses (Heger & Woosley, 2002; Heger et al., 2003). If trace amounts of metals are present, all stars with masses $M_* > 40$ $M_\odot$ end their lives by collapsing into a BH, but with mass losses increasing with metallicity. Stars of primordial composition are thus the ones that should retain most of their mass. However, the exact masses of the final BHs and the maximum mass that can be achieved via this channel depend on the Pop III IMF, which is not well constrained (see also Section 1.4). Recent dedicated simulations including UV feedback from the protostar showed that the mass growth is halted at $\approx$ 40 $M_\odot$ (e.g. Hosokawa et al.,

2011), but in principle, if only a single star forms from its mini-halo, it can reach masses up to 10^3 $M_\odot$ (Hirano et al., 2014). Stellar mass BHs seeds are therefore expected to have masses in the range $10 - 10^3$ $M_\odot$, and are thus referred to as *light seeds*. However, as discussed before, it is very unlikely for these seeds to grow up to 10^9 $M_\odot$ via Eddington-limited accretion. For this reason, many works tried to exploit the possibility that, under specific conditions, accretion could be super-Eddington (e.g. Madau et al., 2014; Pacucci et al., 2015b; Lupi et al., 2016; Pezzulli et al., 2017b), i.e. it can proceed at a a rate higher than the Eddington rate. This can happen because at high mass inflow rates the standard radiatively efficient thin disk solution by Shakura & Sunyaev (1973) is no longer valid. In this regime, the viscosity-generated heat does not have time to be radiated away and it is instead advected onto the BH, significantly reducing the radiative efficiency and in turn the radiation from the BH itself that could be coupled to the gas slowing further mass inflow. Therefore, stellar mass BHs are still a possibility as a SMBHs seeds. *Intermediate mass seeds* (10^{3-4} $M_\odot$) can be formed as a result of dynamical processes within collisional stellar clusters. Star collisions and dynamical interactions lead to an increase of the density of the core, which in turn increases interactions and following collisions. In a multi-mass system, a single massive star might dominate the dynamic, significantly growing by collisions with other stars in a runaway process. Dedicated studies (Devecchi & Volonteri, 2009; Katz et al., 2015; Yajima & Khochfar, 2016; Sakurai et al., 2017) have shown that this pathway can be effective in leading to the formation of a massive star of $\approx 10^{3-4}$ $M_\odot$.

The last main mechanism for BH formation is the so-called Direct Collapse Black Hole (DCBH) scenario, which consists in the collapse of most of the gas in a dark matter halo into a single object, leading to the formation of *heavy seeds* of 10^{5-6} $M_\odot$ (e.g. Eisenstein & Loeb, 1995; Bromm & Loeb, 2003; Schleicher et al., 2013; Ferrara et al., 2014). In order for this mechanism to work, gas should rapidly lose angular momentum avoiding fragmentation. These conditions require the gas composition to be primordial, in order to prevent metal lines cooling, and also the suppression of molecular hydrogen formation, which is the other major coolant, which could otherwise lead to the formation of clumps during the collapse. High mass infall to rapidly accumulate material onto the central object requires high gas temperature, given that it scales as $\dot{M} \propto c_s^3/G \propto T^{3/2}$. This makes dark matter halos at $z \sim 15$, with $\gtrsim 10^7$ $M_\odot$ and virial temperature $T \approx 10^4$ K, the ideal birthplaces of DCBHs. Gas in these halos would cool mainly via atomic-hydrogen cooling, and mass infall can reach values of $\gtrsim 0.1$ $M_\odot\,\mathrm{yr}^{-1}$, which are necessary for a single star to form without radiative feedback onto the surrounding gas (e.g. Hosokawa et al., 2013). However, all these conditions are not easily satisfied and research is still ongoing to explore the feasibility of each one in the primordial Universe (see for example the recent reviews by Latif & Ferrara 2016; Inayoshi et al. 2020 and references therein).

To summarise, different mechanisms have been explored to explain the formation and growth of the SMBHs in place at high redshift. However, many aspects are still uncertain and additional work is needed in order to finally understand their origin.

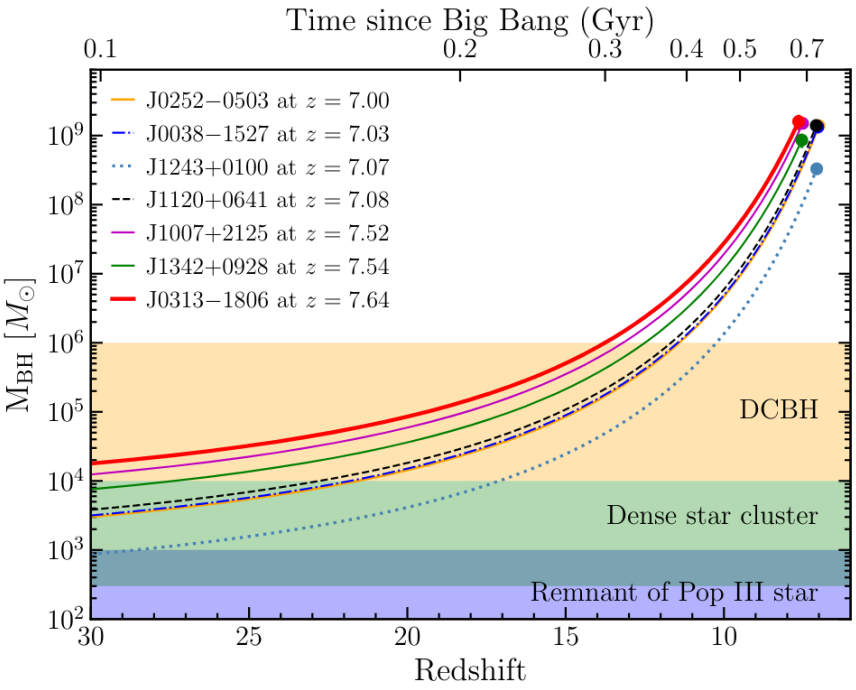

Figure 2.4: The evolution of the black hole masses with redshift according to eq. 2.7, for the observed quasars at $z \geq 7$, assuming $\epsilon_r = 0.1$. The shaded area denotes the typical mass ranges of the seeds produced by the most popular BH formation mechanisms described in this Section. *Figure from Wang et al. (2021a).*

Cosmic dust

3

Cosmic dust is a fundamental component of the ISM of galaxies. It is ultimately a product of stellar evolution and it plays a crucial role both in the evolution of galaxies and in the way we study them. In fact, dust grains enhance the formation of molecular hydrogen, which is the main coolant in primordial gas (see Chapter 1), with an efficiency depending on the grains surfaces (Hirashita & Ferrara, 2002; Cazaux & Tielens, 2004). Dust cooling also becomes relevant in the later stage of star formation, affecting the typical mass of stars and shaping the stellar IMF (Larson, 2005; Omukai et al., 2005; Schneider et al., 2006)). Therefore, both the overall dust content and the dust grain size distribution play an important role in the physical processes regulating star formation.

Dust grains are believed to have sizes ranging from few Åto $\sim \mu$m (e.g. Mathis et al., 1977; Weingartner & Draine, 2001). They also scatter and absorb optical-ultraviolet (UV) radiation from stars and accreting BHs, and thermally re-emit this energy as infrared (IR) photons. Consequently dust affects the spectral energy distribution (SED) of observed galaxies (e.g. Calzetti et al., 2000), the intensity of emission lines and the estimate of the star-formation rate (SFR), which requires corrections for dust extinction (e.g. Kennicutt & Evans, 2012). Therefore, constraining dust properties is essential for interpreting observations and understanding galaxy evolution.

In this Chapter, we provide a brief overview about the main processes regulating dust formation and evolution. We introduce the extinction curves measured in the local Universe and the dust models developed to reproduce them. Finally, we discuss the main observational and theoretical works dedicated to the dust in high-redshift AGN, which is one of the main topic of this Thesis.

3.1 The dust cycle

Dust grains form via the condensation of metals in environments characterized by temperatures of $600 < T < 1800$ K and pressure $10^{-8} < P < 10^{5}$ Pa (e.g. Salpeter, 1974; Frenklach et al., 1989; Frenklach & Feigelson, 1989). These conditions are met in the cold atmosphere of evolved stars in the asymptotic giant branch (AGB) phase (e.g. Ferrarotti & Gail, 2006), which are thought to be the main dust production sources in the local Universe. An alternative pathway for dust formation is represented by expanding supernova (SN) ejecta (SNe; e.g. Kozasa et al., 1989a; Todini & Ferrara, 2001; Nozawa et al., 2003; Bianchi & Schneider, 2007), where the gas conditions for dust condensation are also met. Stars in the mass range $12 - 40$ M$_\odot$, ending their lives as SNe, are expected to produce dust masses of the order of $0.1 - 2$ M$_\odot$ per SN event. However, the reverse shock can destroy a significant fraction ($> 80\%$) of the dust produced, and how

much dust survives is still a matter of debate (e.g. Bianchi & Schneider, 2007; Nozawa et al., 2007; Hirashita, 2015). Both AGB stars and SNe are also able to redistribute the dust produced back into the ISM by means of stellar winds or the SN ejecta, respectively. Additionally, expanding clouds in AGN winds have been proposed as possible sources of dust formation (Elvis et al., 2002).

The metal species able to condense depend on the local metal abundances and on the gas physical conditions, therefore a large variety of grains can be formed. For example, silicate dust is produced mainly by stars with masses $< 1\ \mathrm{M}_\odot$ or $> 4\ \mathrm{M}_\odot$, which are oxygen-rich during the AGB phase. Carbon and SiC dust are instead mainly produced by stars with masses $1\ \mathrm{M}_\odot < M_* < 4\ \mathrm{M}_\odot$ (Zhukovska et al., 2008). Dust grains in SNe ejecta are mainly silicate-based and amorphous carbon (Todini & Ferrara, 2001; Bianchi & Schneider, 2007). After their formation, dust grains undergo several processes which alter the total dust mass and the initial grains size distribution. Once injected into the ISM, they suffer from thermal sputtering if they are swept by SN shocks, or by grain-grain collisions in high temperature gas (shattering, Jones et al., 1996; Yan et al., 2004). The efficiency of this process increases with the radius of the grain (Draine & Salpeter, 1979), therefore large grains suffer from thermal sputtering more than small grains and the initial grain size distribution is effectively skewed toward small grains.

Once the ISM becomes sufficiently enriched with metals, the conditions for accretion of additional metals from the ISM-phase can also be met in dense ISM regions or in molecular clouds (e.g. Spitzer, 1978; Dwek & Scalo, 1980; Draine, 2009). Given that small grains have a larger surface-to-volume ratio as compared to large grains, this process is more effective for the small grains ($a < 0.03\ \mu\mathrm{m}$) with respect to the larger ones. This process does not only boost the abundance of small grains, but it also increases the overall dust mass content. However, its efficiency is still debated (Ferrara et al., 2016). Small grains can also grow via coagulation (Chokshi et al., 1993; Hirashita & Yan, 2009), turning into larger grains. Finally, gas-grains collisions (sputtering) in SN shocks or in hot plasma destroy dust grains and return their metals back to the gas (Draine & Salpeter, 1979; Tielens et al., 1994).

The aforementioned processes are the main drivers of the dust-cycle in galaxies, which is deeply interconnected with their evolution. They have been included in many theoretical works to investigate the evolution of chemical species and dust abundance in galaxies (e.g. Dwek, 1998; Zhukovska et al., 2008; Pipino et al., 2011; Asano et al., 2013b; Hirashita, 2015; Nozawa et al., 2015; Hirashita & Aoyama, 2019). However, despite these efforts, many uncertainties in the theoretical models remain and a complete picture of the dust properties at high-redshift is still missing.

3.2 Dust extinction

The first evidence of interstellar dust came from obscuration of starlight (Trumpler, 1930), also called *extinction*. Measurements of extinction are generally performed by using the so-called pair-method: the spectra of two stars of the same spectral class are compared, assuming that dust extinction becomes negligible at larger wavelength and that one of the sources has little foreground material. The extinction A_λ at a wavelength λ is usually expressed in the form:

$$A_\lambda = -2.5 \log(F_\lambda^{\mathrm{obs}} / F_\lambda^{\mathrm{int}}), \tag{3.1}$$

where F_λ^{obs} is the measured (observed) flux and F_λ^{int} is the intrinsic, i.e. unattenuated flux.

The extinction curves in the MW have been found to vary with the line of sight, depending on the ISM environment they pass through. A useful way to parameterize this variation is the

parameter $R_V = A_V/(A_B - A_V)$, which measures the slope of the extinction curves in the optical band. This has been found to vary between 2.2 and 5.8, with an average $R_V \approx 3.1$ for the diffuse ISM (e.g. Fitzpatrick, 1999). Measurements of the extinction curve have been performed for several lines of sight in the Milky Way (MW; Savage & Mathis, 1979), Small Magellanic Cloud (SMC) and Large Magellanic Cloud (LMC; Fitzpatrick, 1989). The top panel of Figure 3.1 shows the measured average extinction curves of the SMC, LMC and MW in the wavelength range $0.1\ \mu\mathrm{m} < \lambda < 0.5\ \mu\mathrm{m}$. The SMC curve is consistent with a power-law, whereas the MW curve shows a prominent bump around 2175 Å, and a similar feature is visible in the LMC curve. The MW and LMC curves also are less steep than the SMC toward shorter wavelengths. The wavelength dependence of dust extinction entangles information on the grain chemical composition and size distribution, and it is therefore a powerful observational tool to constrain dust models. The 2175 Å bump was shown to be well fitted by a Drude profile (Fitzpatrick & Massa, 1986), whose oscillator strength f per H nucleon requires $n_X f_X/n_H \approx 9.3 \times 10^{-6}$ (Draine, 1989). Given that the oscillator strength per molecule is expected to be $f_x \lesssim 0.5$, this suggested that the responsible molecules had to contain one or more elements between C, O, Mg, Si, Fe, as they are the only ones sufficiently abundant in the ISM (Draine, 2003a). Small, randomly-oriented, spherical grains made by graphite would have an oscillator strength of $f = 0.16$ per C (Draine, 1989), requiring an ISM abundance of $\mathrm{C/H} = 5.8 \times 10^{-5}$, consistent with observations. However, graphite grains do not explain the variation in the full-width-half-maximum (FWHM) of the profile with the line of sight, without a change in the wavelength of the feature. Currently, the most promising explanation for the 2175 Å feature are small polycyclic aromatic hydrocarbons (PAHs) molecules (e.g. Weingartner & Draine, 2001; Li & Draine, 2001a).

Mathis et al. (1977) (hereafter MRN) were the first to develop an interstellar dust model, that was successful in reproducing the observed extinction in the MW. Strong absorption features observed at $9.7\ \mu\mathrm{m}$ and $18\ \mu\mathrm{m}$ are well explained by the stretching and bending modes in silicate grains. The strong extinction feature at 2175 Å is instead reproduced by graphite grains. They proposed a model in which the two populations, silicate and graphite grains, have the same functional form for the grain size distribution:

$$\frac{dn}{da} = C\, n_{\mathrm{H}}\, a^{-3.5}, \tag{3.2}$$

where a is the radius of a grain (assumed to be spherical), n_{H} is the hydrogen number density and C a normalization constant for the grain distribution. The grain size distribution extends from 50 Å to $0.25\ \mu\mathrm{m}$, and the cross sections for each grain was computed by using the Mie theory, which is valid under the assumption that grains are spherical. By extending the MRN model, Draine & Lee (1984) found the values $C = 10^{-25.13}\ \mathrm{cm}^{-2.5}$ and $C = 10^{-25.11}\ \mathrm{cm}^{-2.5}$, for graphite and silicate grains respectively. The MRN model was first challenged by observations of thermal emission from dust in the $3 - 60\ \mu\mathrm{m}$, associated with grains with temperatures $30 - 600$ K, implying the existence of a population of dust grains smaller than 50 Å. Emission features at 3.3, 6.2, 7.7, 8.6, and $11.3\ \mu\mathrm{m}$ were associated to C-H and C-C stretching and bending modes in PAHs (Leger & Puget, 1984), suggesting that carbonaceous grains could be as small as few molecules.

(Weingartner & Draine, 2001) developed a comprehensive model to reproduce the extinction curves of the MW, SMC and LMC, by assuming a mixture of carbonaceous and silicate grains, including a population of small PAHs molecules. These functions are parameterized in order to reproduce the extinction curves along different line of sights, for different values of R_V. Figure 3.2 shows the grain size distributions for the carbonaceous, silicates and PAHs from the

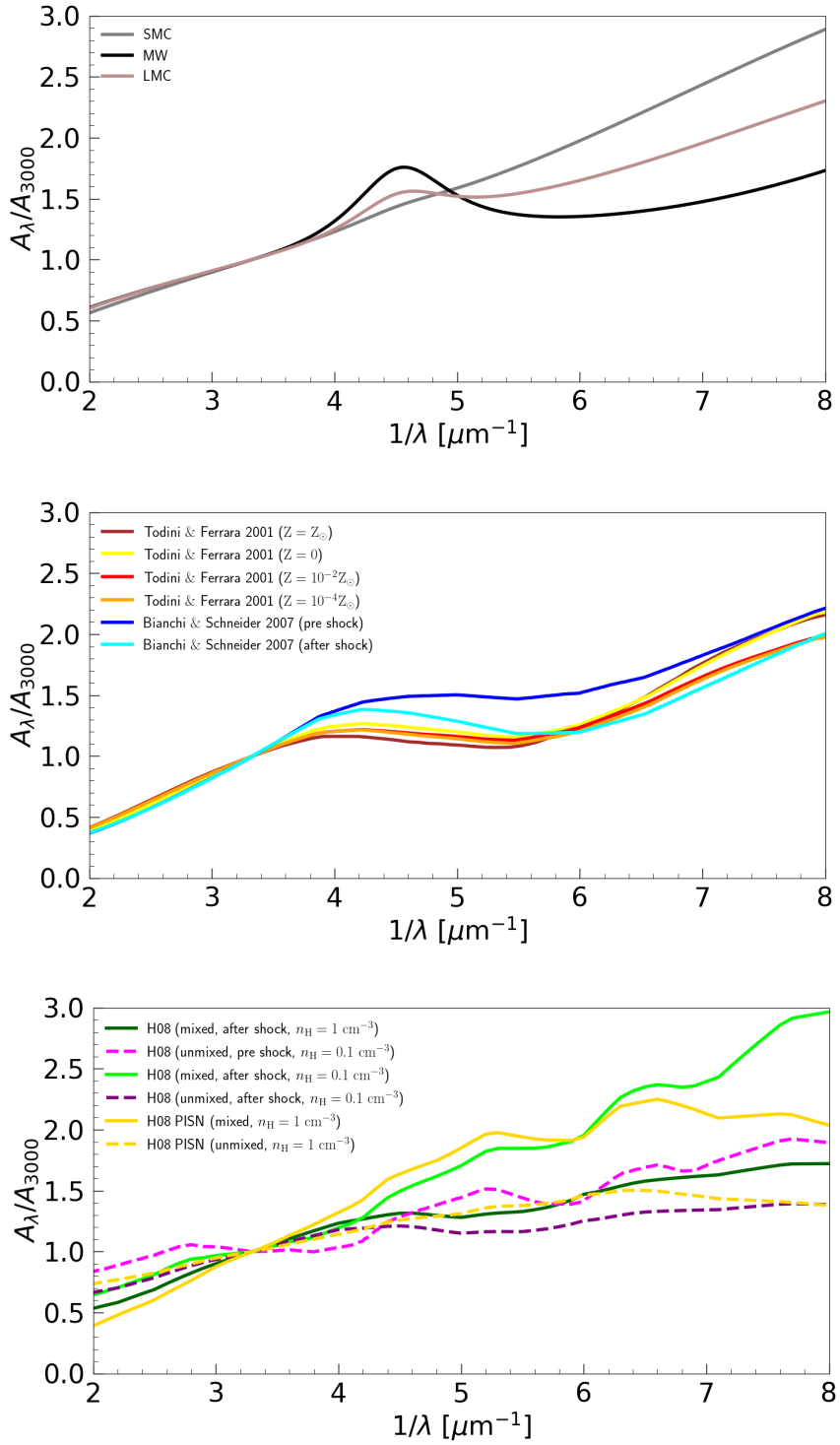

Figure 3.1: *Top panel:* Averaged measured extinction curves of the MW, SMC and LMC by Savage & Mathis (1979); Fitzpatrick (1989). *Middle panel:* Extinction curves predicted by the SN models from Todini & Ferrara (2001) for different values of metallicity of the progenitor star, and from the SN models by Bianchi & Schneider (2007), which also accounts for the reverse shock reprocessing of grains. *Bottom panel:* Extinction curves predicted by the SN models of Hirashita et al. (2008). These models consider two cases for the grain mixing in the SN ejecta: mixed (solid lines) and unmixed (dashed lines); see text for details). They also include dust production from pair-instability SNe (PISN).

(Weingartner & Draine, 2001) that are able to reproduce the average extinction curves in the MW ($R_V = 3.1$ case), SMC and LMC.

Once the grain size distribution is determined, the extinction A_λ can be computed as (Weingartner & Draine, 2001):

$$A_\lambda = 2.5\pi \log e \int d\ln a \frac{dN(a)}{da} a^3 Q_{\rm ext}(\lambda, a), \tag{3.3}$$

where $N(a)$ is the column density of grains with size $\leq a$, and $Q_{\rm ext}(\lambda, a)$ is the extinction factor, which can be computed from Mie theory (Bohren & Huffman, 1984). Many theoretical studies have been dedicated to compute the dielectric functions for silicate and graphite grains, in order to have an accurate determination of the extinction factor, the scattering cross sections and the scattering phase functions (e.g. Draine, 2003b,c).

A similar formalism have been adopted in many works in order to study the extinction curves produced by different source mechanisms (see Section 3.1). Since the first detection of dust in the SN event SN1987A (Moseley et al., 1989; Kozasa et al., 1989b), dust from SN origin has gained particular attention, also because of its appeal in explaining the high dust mass content in early galaxies (see Section 3.3). Todini & Ferrara (2001) developed a model based on classical nucleation theory, exploring the dependence of the extinction curves on the metallicity of the SNe progenitors. Bianchi & Schneider (2007) expanded on this model by exploring a wider range of initial conditions. They also include the destruction and reprocessing of grains by the passage of the reverse shock, thereby studying the evolution of the extinction curve from the time of dust formation ($\sim$ hundreds days after the SN event) to its survival hundreds years later. In particular, they found that $2 - 20\%$ of the initial dust mass survives the reverse shock. Hirashita et al. (2008) computed extinction curves for SN ejecta, building on the study by Nozawa et al. (2007), which focused on the dust destruction by the reverse shock, finding that grains smaller than ~ 0.02 μm are efficiently destroyed in environments characterized by an hydrogen number density higher than 0.1 cm^{-3}. They also explored the effect of mixing of the elements in the ejecta on the formation of dust species, by considering two extreme cases: an unmixed case, where the original onion-like structure of elements is preserved in the helium core, and a mixed case, in which the elements are uniformly mixed. They also included dust production from pair-instability SNe (PISNe), with a progenitor mass of 170 M$_\odot$. As a comparison, the extinction curves produced by these models are shown in the middle and bottom panel of Figure 3.1.

Finally, an additional complication in interpreting the dust properties in a galaxy comes from the geometry. Inferring the dust extinction curve for star-forming galaxies is in general not trivial, because the observed spectrum also depends on the relative distribution of dust and stars. Therefore, observations of star-forming galaxies typically provide an *attenuation* law, which describes the effective reddening of the radiation as compared to the case with no extinction. Calzetti et al. (1994) found an attenuation law for local star-forming galaxies which is flatter than the extinction curve for SMC, LMC and MW. In particular, the resulting curve is characterised by the absence of the 2175 Å dust feature, which is prominent in the MW curve and is usually attributed to polycyclic aromatic hydrocarbon (PAH) nanoparticles (e.g. Li, 2007). This result can be explained with an SMC-like intrinsic curve assuming a clumpy distribution for dust and stars (Gordon et al., 1997; Inoue, 2005) or with a MW-type dust for different geometries (Pierini et al., 2004; Panuzzo et al., 2007). Interpreting observations of star-forming galaxies and inferring their dust properties is therefore quite difficult.

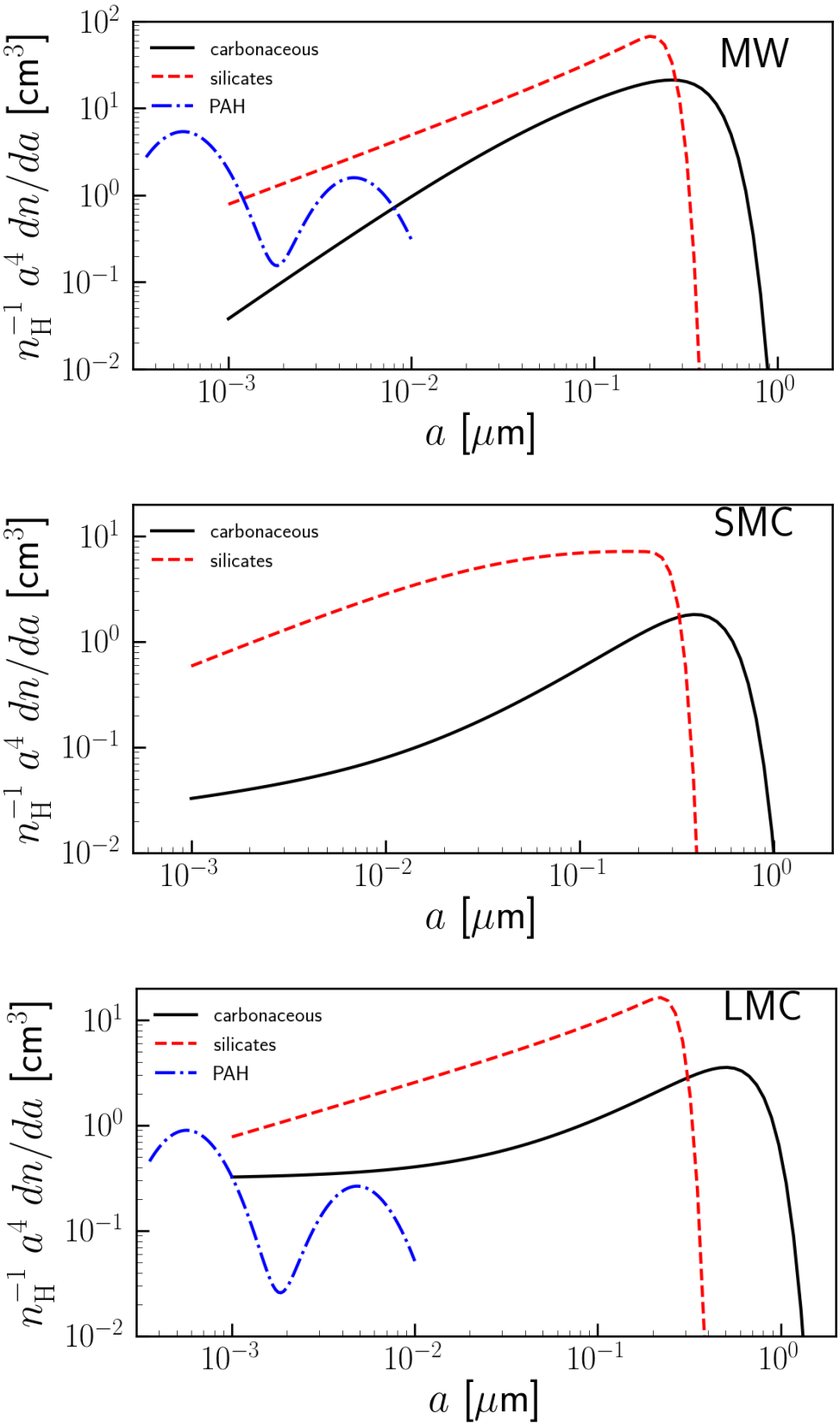

Figure 3.2: Grain size distributions according to the Weingartner & Draine (2001) model that reproduce the MW (for $R_V = 3.1$, *top panel*), SMC (*middle panel*) and LMC (*bottom panel*) extinction curves. The solid black lines show the carbonaceous grains, except the PAHs, which are shown in dash-dotted blue lines; the silicates component is shown in red.

3.3 Dust in high-redshift AGN

Active Galactic Nuclei (AGN), whose UV emission is dominated by the accretion disk powering the central black hole (BH) (e.g. Kormendy & Ho, 2013), represent an alternative place where to study dust extinction properties, possibly avoiding the complications presented by the dust/stars distribution in star-forming galaxies. Observations of mildly reddened quasars at $z < 4.4$ suggested extinction curves compatible with the SMC one (Reichard et al., 2003; Richards et al., 2003; Hopkins et al., 2004). Instead, highly obscured quasars show a prevalence of extinction curves markedly different from any of the SMC/LMC/MW ones, as indicated both by individual sources at $z < 0.7$ (Maiolino et al., 2001a; Gaskell & Benker, 2007), and composite spectra (Gaskell et al. 2004; Czerny et al. 2004, but see also Willott 2005). These works underlined two important features of extinction curves in AGN: (i) a flat (or *grey*) extinction in the far-ultraviolet (FUV), and (ii) the absence of the 2175 Å bump. A grey extinction was interpreted as an indication of the dominance of large grains in the circumnuclear region of AGN, possibly because of the depletion of small grains (Laor & Draine, 1993) or efficient coagulation (Maiolino et al. 2001b; but see also Weingartner & Murray 2002).

The picture is even more complex at high-redshift. Maiolino et al. (2004) analysed the spectrum of the reddened broad absorption line (BAL) quasar SDSSJ1048+4637 at $z = 6.2$, revealing a peculiar extinction curve, being flat at rest-frame $\lambda \gtrsim 1700$ Å and rising at $\lambda \lesssim 1700$ Å, compatible with a supernova origin (Todini & Ferrara, 2001; Hirashita et al., 2005; Bianchi & Schneider, 2007). A similar extinction curve was also inferred for the quasar CFHQS J1509-1749 at $z = 6.12$ (Willott et al., 2007). Gallerani et al. (2010, hereafter G10) studied the optical-near infrared spectra of 33 quasars at $3.9 < z < 6.4$, finding that the extinction curve required to explain dust-reddened quasars (characterised by an attenuation at 3000 Å, A_{3000}, in the range $0.82 < A_{3000} < 2.0$) deviate from the SMC extinction curve, with a tendency to flatten at $\lambda \lesssim 2000$ Å. In the case of BAL quasars, they also computed a mean extinction curve (MECBAL in G10) which displays a significant flattening at $\lambda \lesssim 2000$ Å. These observations suggest that dust properties in AGN at high-redshift might be quite different from those deduced in the local Universe (but see also Hjorth et al., 2013).

A different dust origin at high-redshift is also supported by the timescales involved in dust production mechanisms. The time required for low and intermediate-mass stars ($M < 8$ M$_\odot$) to reach the AGB phase (10^8 yr to few 10^9 yr) is comparable to the age of the Universe at $z \simeq 6$ (Morgan & Edmunds, 2003). The progenitors of core-collapse type II supernovae (SNII) are instead much more short-lived (10^6 yr) and therefore they must be the dominant dust sources at high-z (Gall et al. 2011; Leśniewska & Michałowski 2019, but see also Valiante et al. 2009), provided that a significant fraction of grains survives the reverse shock (Bianchi & Schneider, 2007; Hirashita et al., 2008). This argument makes SNe the most attractive solution to explain the large amounts of dust ($M_{\rm dust} = 10^7 - 10^8$ M$_\odot$) in high-z quasars measured from their rest-frame far-infrared (FIR) emission when the Universe was only 1 Gyr old (e.g. Bertoldi et al., 2003a; Michałowski et al., 2010; Carilli & Walter, 2013; Gallerani et al., 2017b).

Evidence for SN-dust at high-redshift has been also reported in studies of gamma-ray bursts (GRBs) afterglows, as in the case of the GRB071025 afterglow at $z \sim 5$ (Perley et al., 2010; Jang et al., 2011) and GRB050904 at $z = 6.3$ (Stratta et al., 2007). However, both these results are controversial (e.g. Li et al., 2008b; Zafar et al., 2010; Stratta et al., 2011) and studies on larger samples of GRBs up to $z \sim 7$ showed either no significant extinction, or a variety of extinction curves compatible with the SMC, SMC-bar, LMC and MW (Zafar et al., 2011, 2018a,b; Bolmer et al., 2018). Comparing these results with the ones found in AGN is not straightforward, because AGN-host galaxies might have a peculiar dust evolution history with respect to normal

star-forming galaxies (e.g. Nozawa et al., 2015), or quasar spectra might reveal only the dust component in the AGN proximity, whereas GRBs probe the ISM of the host galaxy. Therefore, the question of the dust origin in high-redshift quasars is still open.

Part II

Numerical background

Cosmological simulations

4

Modern astrophysicists can rely on an enormous amount of observational data provided by current facilities. These data constitutes a fundamental tool in order to validate theoretical predictions and models of galaxy formation. However, the problem of galaxy formation and evolution is extremely complex due to its multi-scale and multi-physics nature. Therefore, numerical simulations have been developed in order to tackle this problem, thanks to the possibility to include several physical processes at the same time. These includes, among many others: the physics of gas cooling and heating, star formation, chemical enrichment, feedback from stars and AGN, seeding and growth of massive black holes. Over the last two decades, these simulations have been successful in reproducing a significant number of observations, thereby improving our understanding of the Universe. In this Chapter, we will provide an overview of the main features of modern numerical simulations, their main results, and their current limitations. A detailed review on the subject can be find in Vogelsberger et al. (2020). We then illustrate the hydrodynamic simulations we make use of in the work of this Thesis in Section 4.3.

4.1 Main ingredients

Galaxy formation is an intrinsic multi-scale problem. The physical processes governing their emission and evolution depends on scales of $\approx$ pc, where star formation takes place. Their abundance and clustering properties are instead regulated by evolution on cosmic scales. In order to produce a complete and realistic picture to compare with observations, a variety of physical processes have to be included. The main components constituting a modern cosmological simulation are:

- a cosmological framework;
- Dark Matter and gravitational interactions;
- Baryonic physics (optional).

The large variety of physics included in modern cosmological simulations is schematically shown in Fig. 4.1. We will discuss them in more details in the following sections.

4.1.1 Cosmological framework

As discussed in Section 1.1, observations suggest that our Universe is well described by a cosmological model with a flat geometry, with the matter/energy distribution dominated by a cold-dark matter component and a cosmological constant Λ (e.g. Planck Collaboration et al., 2020).

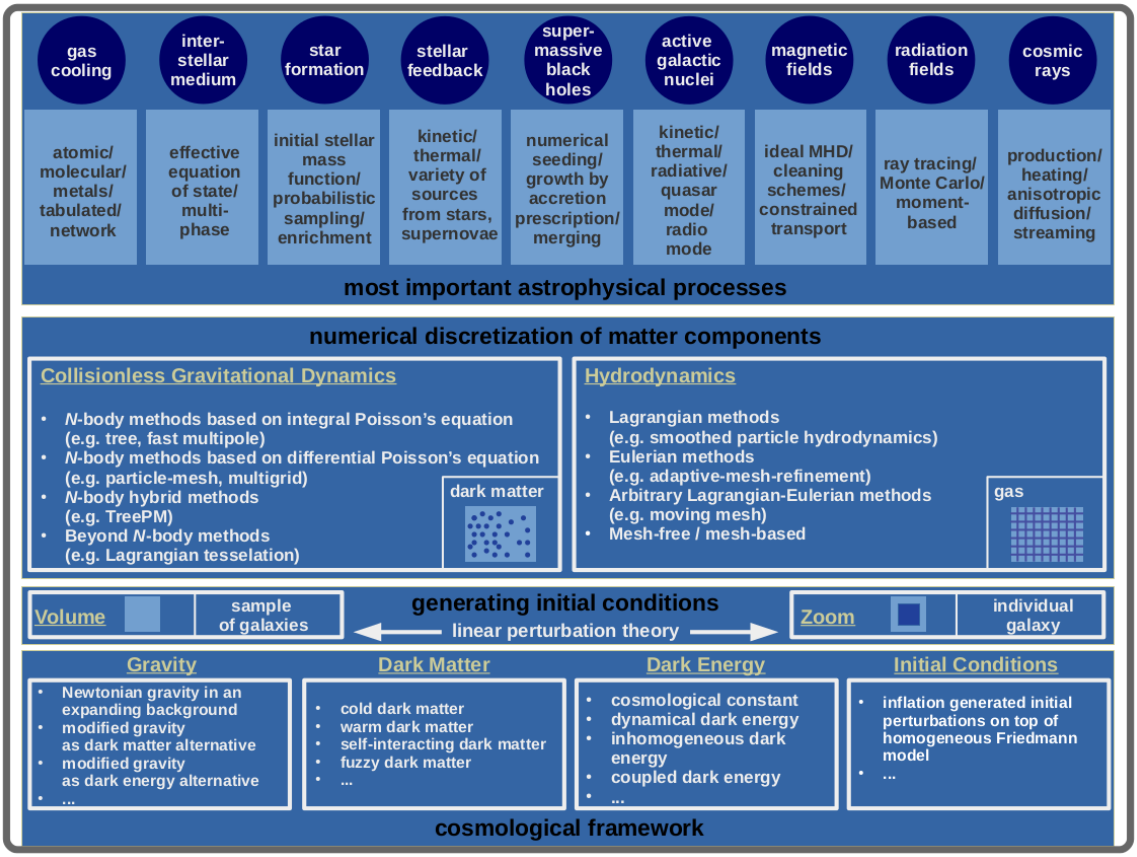

Figure 4.1: Overview of the main components of a cosmological simulation, illustrated from bottom to top. A cosmological framework (see Section 4.1.1) is assumed (bottom box), by specifying the dark matter and dark energy properties, the gravity, and the initial conditions of the simulations. The simulation can be performed either on a large comoving volume to study a large sample of galaxies, or zooming-in on a specific region of the Universe. Several techniques are adopted to solve the dynamics of the dark-matter (see Section 4.1.2) and the fluid equations of the baryonic component (see Section 4.1.3). Finally, a large variety of astrophysical processes is implemented (see Section 4.1.3. *Figure from Vogelsberger et al. (2020).*

Structure formation is expected to proceed hierarchically from small dark matter structures that progressively grows into larger ones, forming the halos which will be the sites of future galaxy formation (see Section 1.3). Thus, the first thing that a cosmological simulation has to include is the flat geometry (described by the FLRW metric), with a specified dark matter, dark energy and baryons compositions. Then, the initial conditions need to be specified, in order to reproduce the perturbations on top of the approximately homogeneous distribution. The matter fluctuations are by their power spectrum $P(k)$, whose functional form is a convolution of the primordial fluctuations set by inflation with the linear growth post-recombination (Eisenstein & Hu 1999, see also Section 1.3.3). Within this framework, the linear density field is typically specified at some time, generally at $z \sim 100$, and then dark matter particle velocities and positions are specified, together with the baryon density, velocity and temperature fields. Cosmological simulations can be divided into two broad categories: *large volume* simulations, and *zoom-in* simulations. In the former, a large volume is simulated and the initial conditions are sampled periodically in the volume. In the zoom-in simulations, a region of interest is selected from a low resolution one, and the number of resolution elements is increased in the small region. Some simulations also employ initial conditions that mimic specific regions of the Universe, such as a matter overdensity (e.g. Hoffman & Ribak, 1991). These are called *constrained* initial conditions. A set of boundary conditions also need to be specified. Typically, periodic boundary conditions are

chosen, as they mimic the homogeneity and isotropy of the matter distribution of the Universe.

4.1.2 Dark Matter

The dark matter plays a crucial role in the structure formation process, by forming the halos that will be the birthplace of galaxies. The evolution of dark matter particles can be described by the collisionless Boltzmann equation, coupled to the Poisson's equation:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \mathbf{v}\frac{\partial f}{\partial r} - \frac{\partial \Phi}{\partial \mathbf{r}}\frac{\partial f}{\partial \mathbf{v}} = 0 \qquad \text{(Boltzmann equation)}$$
$$\nabla^2\Phi = 4\pi G \int f d\mathbf{v} \qquad \text{(Poisson equation)},$$

where $f(\mathbf{r}, \mathbf{v}, t)$ represents the probability density function in the phase space, and Φ is the gravitational potential. This equation is commonly solved via N-body methods, which discretize the equation by sampling the phase-space of dark matter in N points with masses m_i. This technique is more efficient the larger the number of particles adopted, in order to reduce the Poisson noise. The main challenge of N-body methods is to efficiently solve the Poisson equation, i.e. to compute the gravitational forces between the particles. This is generally performed with two different approaches:

- The Poisson equation is solved in its integral form. The forces are computed with a direct summation. In order to reduce the computational effort, the summation is performed in a tree structure, by approximating the contribution from distant particles.

- The Poisson equation is solved in its differential form. This computation is generally achieved by solving the equation in the Fourier space after generating a mesh structure, which is referred to as the particle-mesh (PM) method.

In the first case, gravitational forces are usually softened on small scales, in order to avoid two-body interactions between close particles to produce unphysical results. This is done by employing a softening length ϵ below which particles are treated as a smoothed density field.

Dark-matter-only simulations, despite their simplicity, already provide fundamental insight about the process of structure formation. For example, they show that dark matter begins to collapse in sheets-like structures, then in filaments, and finally forms a web-like structures made of filaments and voids (e.g. Springel et al., 2018). DM-only simulations allow to complement theoretical frameworks to study the halo mass function (e.g. Sheth et al., 2001), the internal halo structure (e.g. Navarro et al., 1996, 1997), and the dark matter distribution (e.g. Springel et al., 2006).

4.1.3 Baryons

Baryons constitute a fundamental component when it comes to compare the predictions of cosmological simulations with the majority of observations, as baryonic physics contains all the processes leading to the formation of stars, which produce the visible light that we can detect. The description of the baryonic component requires first of all the inclusion of thermodynamic state of the gas. In the classical Eulerian formulation, the equations regulating the evolution of

the baryonic fluid can be expressed as:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{4.1}$$

$$\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v} \times \mathbf{v}) + \nabla P = 0 \tag{4.2}$$

$$\frac{\partial \rho e}{\partial t} + \nabla \cdot (\rho e + P)\mathbf{v} = 0. \tag{4.3}$$

Here ρ is the gas density, $\mathbf{v}$ its velocity, P the pressure and $e = u + \frac{1}{2}\mathbf{v}^2$ the energy per unit mass, with u being the internal energy. These equations are then closed with an equation of state for the gas.

In the Eulerian formulation, the hydrodynamic equations are solved on a grid, with a variety of finite-volume, finite-difference techniques to compute the derivatives. Several Godunov's scheme (Godunov, 1959) have been developed over the years to solve the Riemann problem at cells interfaces to correctly resolves shocks and minimize numerical diffusivity (e.g. Colella & Woodward, 1984). Given the large dynamical range of cosmological simulations, most simulations implement an adaptive-mesh refinement (AMR) technique (e.g. Berger & Oliger, 1984), that increases the spatial resolution in specific regions, thereby adapting the grid to the local physical conditions. In the Lagrangian formulation, a mesh-free approach is commonly adopted, by using particles to sample the fluid dynamics and solving the equations for each particle in Lagrangian form. This approach is referred to as Smoothed Particle Hydrodynamics (SPH) formulation, and it has the advantage to ensure conservation of energy, momentum, mass, and entropy, if no artificial viscosity term is present. Other schemes combine the two approaches, for example by using grid deformation techniques. The Voronoi tessellation has gained the most attention due to its mathematical properties and the capability to naturally change its topology, and has been successfully implemented in hydrodynamic simulations (e.g. Springel, 2010).

Simulations also include a large variety of astrophysical processes that complement the fluid equations for the baryonic component, in order to realistic model the galaxy evolution. Most of these processes happen on scales below the spatial resolutions achieved by the simulations, and are therefore implemented via so-called *sub-resolution prescriptions*, which are calibrated with theoretical models and observations. These processes can be broadly summarise as follows:

- **Gas cooling.** Cooling processes (see e.g. Section 1.4) are coupled to the energy equation of the gas. Cooling functions are either tabulated (e.g. Schaye et al., 2015) or implemented explicitly with chemical networks (e.g. Pallottini et al., 2022).

- **Star formation.** Cold and dense gas collapse and eventually form stars. A prescription is assumed to determine the local star formation rate, typically based on the Schmidt-Kennicutt relation (Schmidt, 1959; Kennicutt, 1998). Then, gas is converted into stellar particles using a probabilistic scheme, and the stellar mass is determined by sampling an Initial Mass Function (IMF, e.g. Salpeter 1955; Kroupa 2001; Chabrier 2003). Additional criteria are applied to select the gas that is allowed to be converted into stars, for example considering the gas density (e.g. Schaye et al., 2015) or the Jeans-length (e.g. Hopkins et al., 2014).

- **Stellar Feedback.** Stars inject energy and momentum onto the surrounding gas, mostly in the form of supernovae at the end of their lives, affecting the following star formation within their galaxies (see Section 1.4). Energy can be deposited either thermally or kinetically. In the first case, an ad-hoc prescription is usually implemented in order

to avoid over-cooling (e.g. Stinson et al., 2006). In many cases, wind particles that are hydrodynamically-decoupled from the others, are used to realize local momentum injection, in order to reproduce galactic outflows (e.g. Springel & Hernquist, 2003). Recently, stellar winds, photo-ionization and radiation pressure are starting to be included in simulations (e.g. Agertz et al., 2013). The combination of all the processes implemented reproduces the scenario of a feedback-regulated star formation with an efficiency $\approx 1\%$ in converting gas mass into stars per free fall time, consistent with observations (e.g. Springel & Hernquist, 2003; Hopkins et al., 2014).

- **Supermassive black holes.** Many cosmological simulations include a model to follow the evolution of massive black holes in the early Universe. Given that the detailed process of BH formation cannot be followed in simulations, BHs are seeded in massive halos ($M_{\rm h} \gtrsim 10^9$ M$_\odot$) with a fixed mass depending on the specific BH formation channel considered (see Section 2.5). Gas accretion onto BHs is typically implemented according to the Bondi-Hoyle-Littleton model (Hoyle & Lyttleton, 1939; Bondi & Hoyle, 1944; Bondi, 1952), with some variations like a numerical boosting factor to compensate the limited resolution (e.g. Booth & Schaye, 2009), or a prescription to account for the gas angular momentum (e.g. Anglés-Alcázar et al., 2017). BHs are also allowed to grow via mergers. Similarly to the stars, accreting BHs shining as AGN also provide a form of feedback onto the surrounding gas (see Section 2.4), possibly affecting the whole galaxy and launching large-scale outflows (e.g. Harrison, 2017; Harrison et al., 2018). Many simulations divide the feedback process in two modes: a *quasar* mode, associated with radiatively efficient accretion and a *radio* mode. The first one is associated with radiatively efficient accretion and consists in the deposition of a fraction of the AGN luminosity (proportional to the accretion rate) in the form of energy or momentum onto the gas. The second one is meant to describe AGN jets and it is switched on when the accretion is below a certain threshold.

- **Additional physics.** Many other physical processes can be relevant in order to properly model galaxy evolution, but they are not often included in cosmological simulations, and they require a more tailored implementation. **Magnetic fields** are not important at galaxy-scale but they provide an additional pressure support on small scales, possibly affecting star formation. **Cosmic rays** produced in supernovae explosions and AGN jets enhance the pressure of the ISM and provide an additional heating mechanism for the gas. **Radiation Hydrodynamics** (RHD) affect the thermal and chemical evolution of the gas, and it is particulary relevant for simulations studying the details of the Reionization of the Universe (e.g. Gnedin, 2014). Chapter 5 is dedicated to a more in-depth description of the radiation hydrodynamics and its implementation in numerical simulations.

The sub-grid models described above usually have few free parameters that are calibrated against observations, or specified based on analytical arguments. Simulations usually differ for the detailed implementation of these sub-grid prescriptions, and many results are affected by these choices and can diverge among them. Additional work is still needed in order to refine these models, which currently constitute the main source of uncertainty for the predictions from cosmological simulations.

4.2 Overview of modern simulations

In the last decades, many codes dedicated to cosmological simulations have been developed. We provide here a brief list of the most popular ones, and we refer the interested reader to

Vogelsberger et al. (2020):

- ART (Kravtsov et al., 1997), ART^2 (Li et al., 2008a);
- RAMSES (Teyssier, 2002);
- GADGET-2, GADGET-3 (Springel, 2005);
- AREPO (Springel, 2010);
- ENZO (Bryan et al., 2014);
- GIZMO (Hopkins, 2015).

Fig. 4.2 provide a compendium of simulations developed in the last two decades. On the left side it shows DM-only simulations. They are also divided in large-volume, which provide a large statistical sample of objects, and zoom-in simulations, which follow with higher spatial resolution a specific object in more details. Some notable examples in the first group are: Millennium (Springel et al., 2005), Millennium-2 (Boylan-Kolchin et al., 2009), Bolshoi (Klypin et al., 2011), Millennium-XXL (Angulo et al., 2012), Dark Sky (Skillman et al., 2014). Their volume ranges from $(173\mathrm{Mpc})^3$ of the Millennium-2, which also has the highest mass and spatial resolution (9.4×10^6 $\mathrm{M}_\odot$, and 1.37 kpc respectively) to the $(11628\mathrm{Mpc})^3$ of Dark Sky, which has instead the lowest mass and spatial resolution (5.7×10^{10} $\mathrm{M}_\odot$ and 53.49 kpc respectively).

Some notable examples of the large scale hydrodynamical simulations with baryons are: Illustris (Vogelsberger et al., 2014), Horizon-AGN (Dubois et al., 2014), EAGLE (Schaye et al., 2015), MassiveBlack-2 (Khandai et al., 2015), Bluetides (Feng et al., 2016), Magneticum (Bocquet et al., 2016), MUFASA (Davé et al., 2016), BAHAMAS (McCarthy et al., 2017), Romulus25 (Tremmel et al., 2017), Illustris-TNG (Springel et al., 2018), Simba (Davé et al., 2019). Their volume ranges from the $(25\mathrm{Mpc})^3$ of Romulus25 (which also features the highest mass resolution $m_{\mathrm{DM}}=3.4\times10^5$ $\mathrm{M}_\odot$, $m_{\mathrm{gas}}=2.1\times10^5$ $\mathrm{M}_\odot$), to the $\approx(570\mathrm{Mpc})^3$ of Bluetides and BAHAMAS, with the second one having the lowest mass resolution ($m_{\mathrm{DM}}=5.5\times10^9$ $\mathrm{M}_\odot$, $m_{\mathrm{gas}}=1.1\times10^9$ $\mathrm{M}_\odot$). The spatial resolution (intended as the minimum softening length in SPH codes or the minimum cell size in mesh codes) ranges instead between ≈0.25 kpc of Bluetides, BAHAMAS, Romulus25 to 2.64 kpc of MassiveBlack-2.

4.3 Hydrodynamical simulations used in this work

In the work presented in this Thesis, we make extensive use of the hydrodynamical simulations performed by (Barai et al., 2018) (hereafter B18), and Valentini et al. (2021) (hereafter V21). These simulations were performed in order to study the growth of the SMBHs at high-redshift and the interplay between AGN feedback and the host galaxy properties. We describe them in this section and we refer to the original papers for more details about the numerical setup. We also briefly summarise the main results of these works.

B18 simulations

B18 use a modified version of the Smooth Particle Hydrodynamics (SPH) N-body code GADGET-3 (Springel, 2005) to follow the evolution of a comoving volume of $(500\ \mathrm{cMpc})^3$[1], starting from

[1]Throughout this section ckpc refers to *comoving* kpc and pkpc to *physical* distances in kpc. When not explicitly stated, we are referring to physical distances.

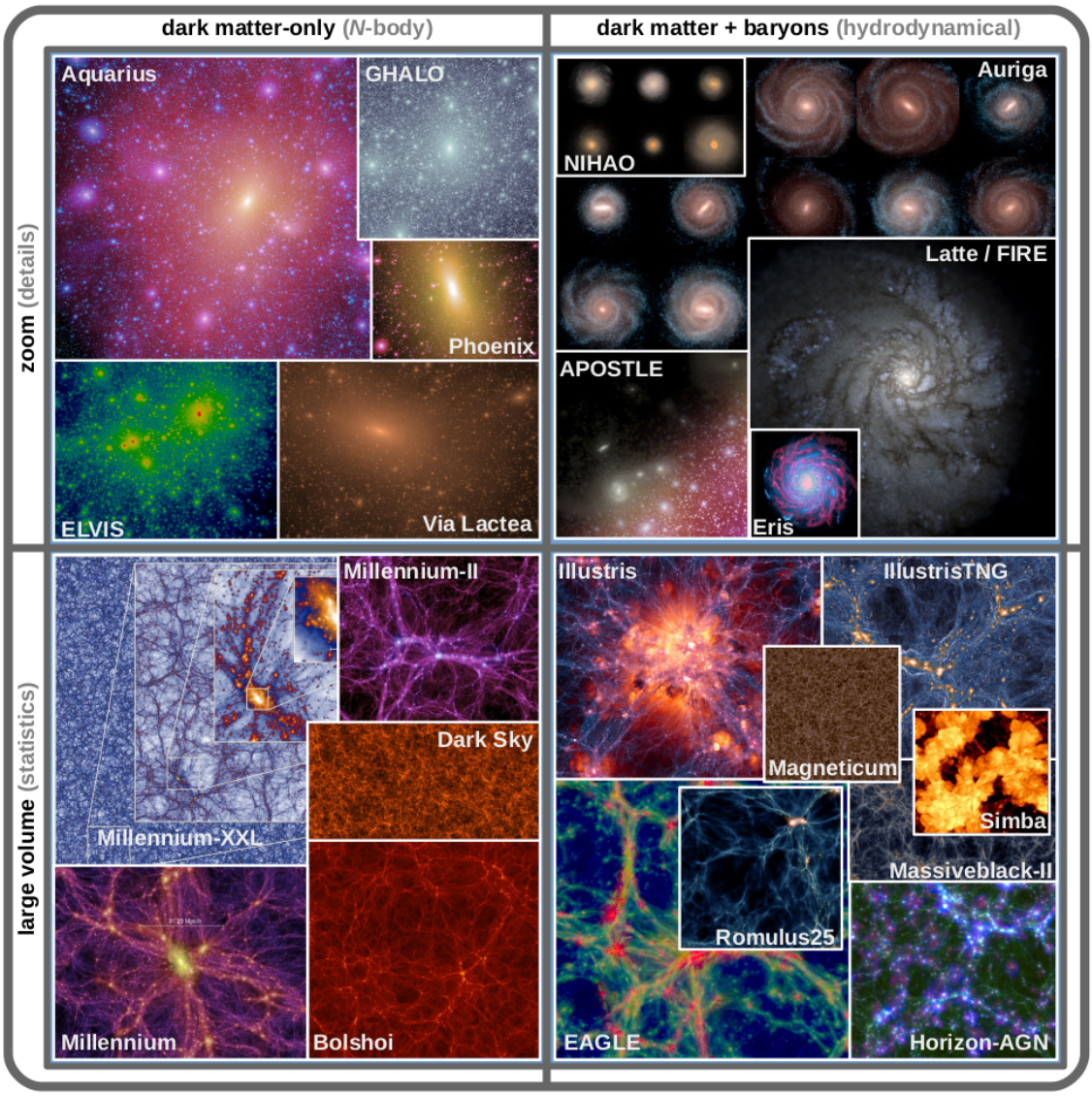

Figure 4.2: Visual representations of some notable recent structure and galaxy formation simulations. They are divided into large volume simulations (*bottom panels*), and zoom-in simulations (*top panels*). *Left panels* show dark matter-only simulations, whereas *right panels* show hydrodynamical simulations with baryons included. *Figure from Vogelsberger et al. 2020.*

cosmological initial condition (IC)[2] generated with MUSIC (Hahn & Abel, 2011) at $z = 100$ and zooming-in on the most massive dark matter (DM) halo inside the box down to $z = 6$[3]. The mass resolution is $m_{\rm DM} = 7.54 \times 10^6$ M$_\odot$ and $m_{\rm gas} = 1.41 \times 10^6$ M$_\odot$ for DM and gas particles, respectively. The softening length for gravitational forces for the high-resolution DM and gas particles is $\epsilon = 1\ h^{-1}$ ckpc, which corresponds to ≈ 210 pc at $z = 6$. For the gas, the smoothing length is determined at each time step according to the local density and typically ranges from 300 pc in the ISM ($n \approx 100$ cm^{-3}) to 6.5 kpc in the CGM ($n \approx 10^{-2}$ cm^{-3}).

The code accounts for radiative heating and cooling according to the tables computed by Wiersma et al. (2009), which also include metal-line cooling. B18 adopts the multiphase model

[2]A flat ΛCDM model is assumed with the following cosmological parameters (Planck Collaboration et al., 2016): $\Omega_{\rm M,0} = 0.3089$, $\Omega_{\Lambda,0} = 0.6911$, $\Omega_{\rm B,0} = 0.0486$, $H_0 = 67.74$ km s^{-1} Mpc^{-1}.

[3]In the low-resolution DM-only simulation, the most massive halo at $z = 6$ has a mass of $M_{\rm halo} = 4.4 \times 10^{12}$ M$_\odot$ (virial radius $R_{200} = 73$ pkpc), massive enough to host luminous AGN, as suggested by clustering studies (e.g. Allevato et al., 2016).

by Springel & Hernquist (2003), which follows the hot and cold phase of the ISM, resulting in an effective equation of state for the gas. Star formation is implemented with a density-based criterion, with a density threshold for star formation of $n_{\rm SF} = 0.13$ cm^{-3} and assuming a Chabrier (2003) initial mass function (IMF) in the mass range $0.1 - 100$ M$_\odot$. Stellar evolution and chemical enrichment are computed for the eleven element species (H, He, C, Ca, O, N, Ne, Mg, S, Si, Fe) tracked in the simulation, following Tornatore et al. (2007). Kinetic feedback from supernovae (SN) is included by relating the wind mass-loss rate ($\dot{M}_{\rm SN}$) with the star formation rate ($\dot{M}_\star$) as $\dot{M}_{\rm SN} = \eta \dot{M}_\star$ and assuming a mass-loading factor $\eta = 2$. The wind kinetic energy is set to a fixed fraction χ of the the SN energy: $\frac{1}{2}\dot{M}_{\rm SN} v_{\rm SN}^2 = \chi \epsilon_{\rm SN} \dot{M}_\star$, where $v_{\rm SN} = 350$ km s^{-1} is the wind velocity and $\epsilon_{\rm SN} = 1.1 \times 10^{49}$ erg M$_\odot^{-1}$ is the average energy released by a SN for each M$_\odot$ of stars formed[4].

In the simulation each BH is treated as a collisionless sink particle and the following seeding prescription is used. When a DM halo – that is not already hosting a BH – reaches a total mass of $M_{\rm h} = 10^9$ M$_\odot$, a $M_{\rm BH} = 10^5$ M$_\odot$ BH is seeded at its gravitational potential minimum location. BHs are allowed to grow by accretion of the surrounding gas or by mergers with other BHs. Gas accretion onto the BH is modelled via the Bondi-Hoyle-Littleton scheme $\dot{M}_{\rm Bondi}$ (Hoyle & Lyttleton, 1939; Bondi & Hoyle, 1944; Bondi, 1952):

$$\dot{M}_{\rm Bondi} = \alpha \frac{4\pi G^2 M_{\rm BH}^2 \rho}{\left(c_{\rm s}^2 + v^2\right)^{3/2}}, \tag{4.4}$$

where G is the gravitational constant, ρ is the gas density, $c_{\rm s}$ is its sound speed, and v is the velocity of the BH relative to the gas, and $\alpha = 100$ is a numerical boost factor adopted to compensate for the low resolution in the regions closest to the BHs. Gas accretion is also capped at the Eddington rate $\dot{M}_{\rm Edd}$, so the final BH accretion rate $\dot{M}_{\rm BH}$ reads as follows:

$$\dot{M}_{\rm BH} = \min(\dot{M}_{\rm Bondi}, \dot{M}_{\rm Edd}). \tag{4.5}$$

To avoid BHs moving from the centre of the halo in which they reside because of numerical spurious effects, we implement BH repositioning or *pinning* (see also e.g. Springel et al., 2005; Sijacki et al., 2007; Booth & Schaye, 2009; Schaye et al., 2015): at each time-step BHs are shifted towards the position of minimum gravitational potential within their softening length. During its growth a BH radiates away a fraction of the accreted rest-mass energy, with a bolometric luminosity

$$L_{\rm bol} = \epsilon_{\rm r} \dot{M}_{\rm BH} c^2, \tag{4.6}$$

where c is the speed of light and $\epsilon_{\rm r}$ is the radiative efficiency. B18 set $\epsilon_{\rm r} = 0.1$, a fiducial value for radiatively efficient, geometrically thin, optically thick accretion disks around a Schwarzschild BH (Shakura & Sunyaev, 1973). A fraction $\epsilon_{\rm f} = 0.05$ of this energy is distributed to the surrounding gas in a kinetic form[5].

In the work described in this Thesis we consider the following three runs performed by B18, starting from the same ICs:

- *noAGN*: control simulation without BHs.
- *AGNsphere*: simulation accounting for BH accretion and AGN feedback. The kinetic feedback is distributed according to a spherical geometry.

[4] In the ISM multiphase model adopted here (Springel & Hernquist, 2003), kicked particles mimicking stellar winds are temporarily hydrodynamically decoupled. This procedure may affect both the properties of the resulting outflows and the structure of the surrounding ISM (e.g. Dalla Vecchia & Schaye, 2008).

[5] We refer to B18 for details about the choice of the value for $\epsilon_{\rm f}$ and the numerical implementation of the kinetic feedback.

- *AGNcone*: same as the *AGNsphere* run, but with kinetic feedback distributed inside a bi-cone with an half-opening angle of $45°$.

We discuss them in more details in Chapter 6.

V21 simulations

In V21 a DM-halo of $M_{\rm h} = 1.12 \times 10^{12}$ $\rm M_\odot$ is chosen for the zoom-in simulation inside a comoving volume of $(148\ \rm cMpc)^3$. A zoom-in region of size $(5.25\ \rm cMpc)^3$ is chosen for re-simulation, with the highest resolution particles of the zoom-in simulation having a mass of $m_{\rm DM} = 1.55 \times 10^6$ $\rm M_\odot$ and $m_{\rm gas} = 2.89 \times 10^5$ $\rm M_\odot$. The gravitational softening lengths employed are $\epsilon_{\rm DM} = 0.72$ ckpc and $\epsilon_{\rm bar} = 0.41$ ckpc for DM and baryon particles respectively, the latter corresponding to ≈ 60 pc at $z = 6$, i.e. a factor of ≈ 3 lower than in B18.

Instead of the multiphase model by Springel & Hernquist (2003), the ISM is described by means of the MUlti Phase Particle Integrator (MUPPI) sub-resolution model (e.g. Murante et al., 2010; Valentini et al., 2020). It features metal cooling, thermal and kinetic stellar feedback, the presence of a UV background, and a model for chemical evolution following Tornatore et al. (2007), and tracking fifteen species (H, He, C, Ca, O, N, Ne, Mg, S, Si, Fe, Na, Al, Ar, Ni). In particular, star formation is implemented with a $\rm H_2$-based prescription instead of a density-based cryterion, as in B18.

BHs are treated as collisionless sink particles, and a $M_{\rm BH} = 1.48 \times 10^5$ $\rm M_\odot$ BH is seeded at the centre of a BH-less DM halo when it reaches a total mass of $M_{\rm h} = 10^9$ $\rm M_\odot$. BHs are allowed to grow by accretion of the surrounding gas or by mergers with other BHs. The accretion rate is estimated separately for the hot and cold phases according to eq. 4.4. For the cold phase, the accretion rate is reduced depending on the angular momentum. The total hot+cold accretion rate is capped at the Eddington rate $\dot{M}_{\rm Edd}$. A fraction of the accreted rest-mass energy is radiated away from the BHs according to eq. 4.6, assuming a radiative efficiency of $\epsilon_{\rm r} = 0.03$. A fraction $\epsilon_{\rm f} = 10^{-4}$ of the radiated luminosity $L_{\rm bol}$ is thermally coupled to the gas surrounding the BHs, and it is isotropically distributed into the gas. In the work of this Thesis, we make use of the fiducial run of the suite by V21, which we will refer to as *AGNthermal*.

In Table 4.1, we report the main numerical features of the runs adopted in the work of this Thesis, and the main physical properties of the zoomed-in halo at $z = 6$ inside a cubic region of 60 kpc size centred on the halo's centre of mass.

	noAGN	*AGNsphere*	*AGNcone*	*AGNthermal*
reference work	Barai et al. (2018)			Valentini et al. (2021)
resolution	$m_{\rm DM} = 7.54 \times 10^6$ M$_\odot$ $m_{\rm gas} = 1.41 \times 10^6$ M$_\odot$ $\epsilon = 1\ h^{-1}$ ckpc			$m_{\rm DM} = 1.55 \times 10^6$ M$_\odot$ $m_{\rm gas} = 2.89 \times 10^5$ M$_\odot$ $\epsilon_{\rm DM} = 0.72$ ckpc, $\epsilon_{\rm bar} = 0.41$ ckpc
subgrid model	Springel & Hernquist (2003)			MUPPI (Valentini et al., 2020)
AGN feedback	no	kinetic, spherical	kinetic, bi-conical	thermal, spherical
$M_{\rm gas}$ [M$_\odot$]	2.4×10^{11}	2.2×10^{11}	1.4×10^{11}	6.8×10^{10}
$M_\star$ [M$_\odot$]	1.5×10^{11}	7.9×10^{10}	7.7×10^{10}	4.0×10^{10}
SFR [M$_\odot$ yr^{-1}]	790	400	160	190
$\dot{M}_{\rm BH}$ [M$_\odot$ yr^{-1}]	0	10	70	36

Table 4.1: Summary of the hydrodynamic runs adopted in this work, presented in Barai et al. (2018) and Valentini et al. (2021). We report: the dark matter (*2nd row*) and baryonic (*3rd row*) particles mass resolution, $m_{\rm DM}$ and $m_{\rm gas}$ respectively; (*4th row*) the gravitational softening length ϵ; (*5th row*) the reference model for the subgrid physics (see text for more details); (*6th row*) the AGN feedback model. We then indicate the main physical properties of the zoomed-in halo at $z = 6$ within a cubic region of 60 kpc size (that corresponds to $\sim 50\%$ of the virial radius: gas mass ($M_{\rm gas}$), stellar mass ($M_\star$), star formation rate (SFR, averaged over the last 10 Myr), and the sum of the accretion rate of all the black holes (BHs) in the selected region ($\dot{M}_{\rm BH}$).

Radiative Transfer

5

Photons from astrophysical sources are what allow us to detect and study them, making the understanding of the mechanisms of emission and transfer to media primary in order to understand our Universe. Studying the radiation from cosmic objects is important not only to analyze the source itself, but also to gain information about the intervening medium, which ultimately alter the radiation traveling toward us. The study of the Radiative Transfer (RT) problem is then one of the fundamental challenges of modern astrophysics. A comprehensive, recent review of the most popular techniques to solve the RT problem, particularly focused on 3D dusty systems can be find in (Steinacker et al., 2013). In this Chapter, we briefly overview the RT problem (Section 5.1), with particular focus on the RT dust problem. We discuss the main computational strategies that have been developed in order to tackle this problem in Section 5.2. In particular, in Section 5.2.4 we focus on the Monte Carlo method, which is at the basis of the RT code SKIRT[1] (Baes & Camps, 2015; Camps et al., 2016), which has been extensively used in this work. Finally, we describe in details the numerical setup for the RT calculations adopted in this Thesis in Section 5.4.

5.1 The equation of radiative transfer

The radiative transfer problem can be described in general by an integro-differential equation as follos. Let $I(\mathbf{x}, \mathbf{n}, \nu, t)$ be the specific intensity of the radiation field at a given point in space $\mathbf{x}$ and at a given time t, with $\mathbf{n}$ indicating the direction of propagation of the radiation and ν its frequency. The specific intensity represents the amount of energy carried by the radiation field per solid angle, unit time and frequency across a surface element perpendicular to $\mathbf{n}$:

$$I(\mathbf{x}, \mathbf{n}, \nu, t) = \frac{dE}{d\mathbf{A} d\Omega dt d\nu}. \tag{5.1}$$

As the radiation travels through a medium, the specific intensity changes as a result of the interactions between the photons and the medium. Excluding line radiative transfer, and scattering processes, the evolution of the radiation field can be described by the following equation (e.g. Chandrasekhar, 1960; Rybicki & Lightman, 1979):

$$\frac{1}{c}\frac{\partial I(\mathbf{x}, \mathbf{n}, \nu, t)}{\partial t} + \mathbf{n} \cdot \nabla I(\mathbf{x}, \mathbf{n}, \nu, t) = -\kappa_{abs}(\mathbf{x}, \nu, t)\rho(\mathbf{x}, t) I(\mathbf{x}, \mathbf{n}, \nu, t) + j(\mathbf{x}, \mathbf{n}, \nu, t). \tag{5.2}$$

Here, the left-hand side of the equation represents the differential variation of the specific intensity over an infinitesimally small intervals in space and time. The right-hand side contains two

[1] Version 8, http://www.skirt.ugent.be.

terms. The first one represents the sink therm of the radiation, which is the loss of energy due to matter-interactions. It is described by the absorption coefficient per unit mass $\kappa_{abs}(\mathbf{x}, \nu, t)$ and it also depends on the local mass density $\rho(\mathbf{x}, t)$ of the medium. In general, a large variety of processes can contribute to the extinction term in eq. 5.2. Photo-excitation or photo-ionization results in a loss of photons from the radiation field, whose energy result in excitation of electrons to higher energy states or ionization of the atoms. Subsequent absorption processes can eventually convert a fraction of the energy of the radiation field into kinetic energy of the atoms, eventually resulting in gas heating. For what concern dust grains, photons can be absorbed and transformed into internal energy, leading to heating of the grain. This process depends on the chemical composition, size and shape of the grain, and it depends on the photon wavelength, as also discussed in Chapter 3. The second term in eq. 5.2 represents the source term, that is the new radiation that is added from the medium at the position $\mathbf{x}$, time t and frequency ν toward the direction $\mathbf{n}$. The emission term accounts for stellar emission, radiation from AGN, emission lines from ionized or excited gas, or Bremsstrahlung. The complexity of the RT problem depends essentially on the the complexity of the RT equation, which is determined by the physical processes responsible for extinction and emission. If scattering is not accounted for, eq. 5.2 can be integrated numerically, as long as the extinction and source terms are known. However, scattering tremendously complicates the problem. Scattering removes radiation from one direction of propagation $\mathbf{n}$ to another direction $\mathbf{n}'$. Therefore, it behaves at the same time as a sink term, by resulting in effective extinction toward $\mathbf{n}$, and as a source term, by adding new radiation toward $\mathbf{n}'$. In case of monochromatic scattering, the frequency of the re-emitted radiation remains the same, but in general the radiation in the new direction can also be re-emitted at different frequency ν'. This process is described by a phase function $\Phi(\mathbf{n}, \mathbf{n}', \mathbf{x}, \nu, \nu')$, which indicates the probability that a photon at frequency ν moving in direction $\mathbf{n}$ is scattered at position $\mathbf{x}$ toward the direction $\mathbf{n}'$ with a new frequency ν'. The phase function satisfies the normalization:

$$\int_{4\pi} \Phi(\mathbf{n}, \mathbf{n}', \mathbf{x}, \nu, \nu') d\Omega = \int_{4\pi} \Phi(\mathbf{n}, \mathbf{n}', \mathbf{x}, \nu, \nu') d\Omega' = 1. \tag{5.3}$$

When adding the scattering term to the RT equation, it reads as:

$$\frac{1}{c}\frac{\partial I(\mathbf{x}, \mathbf{n}, \nu, t)}{\partial t} + \mathbf{n} \cdot \nabla I(\mathbf{x}, \mathbf{n}, \nu, t) = -\kappa_{ext}(\mathbf{x}, \nu, t)\rho(\mathbf{x}, t) I(\mathbf{x}, \mathbf{n}, \nu, t) + j(\mathbf{x}, \mathbf{n}, \nu, t) + \tag{5.4}$$

$$+ \kappa_{sca}(\mathbf{x}, \nu, t)\rho(\mathbf{x}, t) \int_{4\pi} \Phi(\mathbf{n}, \mathbf{n}', \nu, \nu') I(\mathbf{x}, \mathbf{n}', \nu', t) d\Omega', \tag{5.5}$$

where κ_{ext} is the *extinction* coefficient, which contains the loss of energy in the radiation field due to both absorption and scattering, i.e. $\kappa_{ext} = \kappa_{abs} + \kappa_{sca}$.

Eq. 5.4 emphasizes the complexity of the RT problem in its general form. First of all, it is an equation in 6 dimension (3 spatial coordinates, 2 for the propagation of the radiation, 1 for the frequency), therefore the computational cost to solve is high, even with respect to classical hydrodynamics equations, which has 3 less variables. Secondly, it is an integro-differential equation in which the radiation field at different positions and in all direction is coupled. The non-locality of the problem intrinsically makes it difficult to develop efficient techniques to simplify the problem. Moreover, the fact that the extinction and scattering coefficients are spatially varying makes the problem highly non-linear. This problem is also further complicated by the complexity of the 3D spatial distribution characterizing realistic astrophysical dusty systems. Simplifying the geometry in this cases can produce completely different results (e.g. Witt & Gordon, 1996).

The complexity of the problem further increases when we think that dusty systems are composed of grain mixtures, each with different optical properties, due to the grain size and composition. As also discussed in Chapter 3, many theoretical works have been dedicated to compute the optical properties of dusty grains (e.g. Draine, 2003b,c). For what concerns the phase function, the most used parametrization is the one by Henyey & Greenstein (1941), which requires a single-parameter approximation. The Henyey-Greenstein phase function provides a good approximation to the phase functions predicted numerically.

In the case of RT through dusty media, the source term in eq. 5.4 can be decomposed into $j(\mathbf{x}, \mathbf{n}, \nu, t) = j_*(\mathbf{x}, \mathbf{n}, \nu, t) + j_{\rm d}(\mathbf{x}, \mathbf{n}, \nu, t)$, where the first term indicates emission from stellar sources and AGN and the second one emission from dust grains. The dust emissivity term clearly depends on the optical properties of the grains and on the local radiation field, further complicating the RT problem. A common simplifying assumption is that dust grains are in thermal equilibrium with the local radiation field. In this case, the thermal emission of a single grains population i can be approximated as a modified black-body:

$$j_{\rm d}(\mathbf{x}, \mathbf{n}, \nu, t) = \kappa_{abs,i}(\mathbf{x}, \nu)\rho(\mathbf{x})B_\nu(T_i(\mathbf{x}), \tag{5.6}$$

where $B_\nu(T_i(\mathbf{x})$ is the Planck function and $T_i(\mathbf{x})$ is the dust temperature of the population. For a dust mixture, the dust emissivity can be computed by summing over all the populations. The equilibrium dust temperature $T_i(\mathbf{x})$ of each species is computed by equating the energy radiated by the population to the energy absorbed, i.e. assuming energy balance:

$$\int_0^\infty \kappa_{abs,i}(\mathbf{x}, \nu)J(\mathbf{x}, \nu)d\nu = \int_0^\infty \kappa_{abs,i}(\mathbf{x}, \nu)\rho(\mathbf{x})B_\nu(T_i(\mathbf{x})d\nu, \tag{5.7}$$

where $J(\mathbf{x}, \nu, t)$ is the mean intensity of the radiation field, defined as:

$$J(\mathbf{x}, \nu, t) = \frac{1}{4\pi}\int_{4\pi} I(\mathbf{x}, \mathbf{n}, \nu, t)d\nu. \tag{5.8}$$

The assumption of thermal equilibrium works well for large grains. However, small dust grains have small heat capacities, then even a single or a few photons can significantly heat the grain. As a result, small grains will undergo temperature fluctuations, with emission at temperatures higher than the equilibrium temperature.

5.2 Solving the RT problem

Solving a high-dimensional, non-local, non-linear equation requires very high computational resources. Over the last decades, with the amount of observational data to explain increasing and so the computational power, a variety of techniques have been developed in order to solve the RT problem. These can be broadly grouped in the following categories:

- **finite difference;**
- **moment-based methods;**
- **Ray-tracing;**
- **Monte Carlo.**

We will describe them briefly in the following sections.

5.2.1 Finite difference approach

The first strategy to solve the RT equation consists into discretizing it and converting the problem into a set of differential equations (e.g. Stenholm et al., 1991). This approach is somewhat similar to what is done in classical hydrodynamics. However, the solution requires an high computational effort, both in memory and speed, in order to solve the 3D equation, and therefore this technique is mostly used for the 1D and 2D dust RT problem. When applying this technique to the 3D problem, the choice of the grid is very important, because it determines the computational cost and the accuracy of the solution. If the grid is too coarse, the computational cost is low but the accuracy is also low. For this reason, instead of uniform Cartesian grids, typically adaptive refined Cartesian grids are used for this kind of computation, and in the most-advanced versions, the grids is adapted based also on the local radiation field (e.g. Steinacker et al., 2002). This strategy allow to efficiently use the computational domain, as the astrophysical structures to study (filaments, clouds) have very complex 3D geometries that require very high-resolution in a small fraction of the volume of the system, where most of the space can be sampled with a coarser grid.

5.2.2 Moment-based approach

In the moment-based approach, the first moments of the RT equation are considered, in order to expand the solution into spherical harmonics. In this way, the angular dimensions of the problem are removed and the computation becomes easier. In practice, the description of the RT problem is moved from a beam description to a fluid description, in which the bulk motion of the photons is considered. The fluid description allows also to solve the RT equation as the classical hydrodynamics equation, offering the opportunity to combine radiative transfer and hydrodynamcis in the so-called radiation-hydrodynamic (RHD) simulations (e.g. Rosdahl et al., 2013). This approach also has the advantage to be naturally computationally independent by the number of sources considered.

In this framework, the solution of the RT equation is obtained as follows. First, the following quantities are introduced:

$$E(\mathbf{x},\nu,t) \equiv \frac{1}{c}\int I(\mathbf{x},\mathbf{n},\nu,t)\, d\Omega \tag{5.9}$$

$$\mathbf{F}(\mathbf{x},\nu,t) \equiv \int \mathbf{n}\, I(\mathbf{x},\mathbf{n},\nu,t)\, d\Omega \tag{5.10}$$

$$\mathbb{P}(\mathbf{x},\mathbf{n},\nu,t) \equiv \int \mathbf{n}\,\mathbf{n}\, I(\mathbf{x},\mathbf{n},\nu,t)\, d\Omega, \tag{5.11}$$

where $E(\mathbf{x},\nu,t)$ is the monochromatic radiation energy density, $\mathbf{F}(\mathbf{x},\mathbf{n},\nu,t)$ is the monochromatic flux and $\mathbb{P}(\mathbf{x},\mathbf{n},\nu,t)$ is the tensor of the radiation pressure. Then, the 0-th and 1-st moments of the RT equation are considered, integrating over the solid angle, obtaining the following system (e.g. Mihalas & Mihalas, 1984):

$$\frac{\partial E(\mathbf{x},\nu,t)}{\partial t} + \nabla\cdot\mathbf{F}(\mathbf{x},\nu,t) = S(\mathbf{x},\nu,t) - \kappa_{\mathrm{E}}\rho(\mathbf{x},t)cE(\mathbf{x},\nu,t) \tag{5.12}$$

$$\frac{1}{c}\frac{\partial \mathbf{F}(\mathbf{x},\nu,t)}{\partial t} + c\nabla\cdot\mathbb{P}(\mathbf{x},\nu,t) = -\kappa_{\mathrm{F}}\rho(\mathbf{x},t)\mathbf{F}(\mathbf{x},\nu,t), \tag{5.13}$$

where κ_{E} and κ_{F} are the radiation energy and flux weighted mean absorption coefficients respectively, and $S(\mathbf{x},\nu,t)$ is the integral over the solid angle of the total emissivity $j(\mathbf{x},\mathbf{n},\nu,t)$.

This system has three variables, which are the three moments of the specific intensity: $E(\mathbf{x}, \nu, t)$, $\mathbf{F}(\mathbf{x}, \nu, t)$ and $\mathbb{P}(\mathbf{x}, \nu, t)$. In order to solve it, an additional relation is needed, which is often referred to as a *closure relation*. The most-used one is the M1 closure relation (Levermore, 1984), which has been implemented in many numerical codes (e.g. González et al., 2007; Rosdahl & Teyssier, 2015). This formalism has the advantage of depending only on local quantities, but still keeping information of the directionality of the flow. By expressing the radiation pressure as:

$$\mathbb{P}(\mathbf{x}, \nu, t) = \mathbb{D}\ E(\mathbf{x}, \nu, t), \tag{5.14}$$

where $\mathbb{D}$ is the Eddington tensor, which has the following form in the M1 closure:

$$\mathbb{D} = \frac{1-\chi}{2}\mathbb{I} + \frac{3\chi - 1}{2}\mathbf{n} \otimes \mathbf{n}, \tag{5.15}$$

where $\mathbb{I}$ is the identity tensor, $\mathbf{n} = \mathbf{F}/|\mathbf{F}|$ is the direction of the flux and χ is given by:

$$\chi = \frac{3 + 4f^2}{5 + 2\sqrt{4 - 3f^2}}, \tag{5.16}$$

and $f = |\mathbf{F}|/cE$ is the reduced flux. This approximation recovers well the flux in the diffusion regime, when $f \ll 1$ and $\mathbb{D} \simeq 1/3\mathbb{I}$. It also well describes free streaming radiation from a single source, when $f \simeq 1$ and $\mathbb{D} \simeq 1\mathbb{I}$. In presence of multiple sources, in case of shadowing and in all other regimes, the solution is less accurate and un-physical oscillations can be found.

5.2.3 Ray-tracing approach

The Ray-Tracing (RayT) approach focus on solving the RT equation along a specific direction (ray). Along this direction the RT equation can be simplified as follows:

$$\frac{dI(s,\nu}{ds} = -\kappa_{abs}(s,\nu)\rho(s)I(s,\nu) + j(s,\nu), \tag{5.17}$$

where s is the linear coordinate along the ray and we assumed no time dependence for simplicity. The solution has to be computed on each cell of the computational grid, and in each cell the mass density ρ_0, the extinction coefficient per unit mass κ_0 and the emissivity $j_0(\nu)$ are assumed to be constant. In a single cell, the previous equation can then be rewritten in a differential form:

$$I(s + ds, \nu) = I(s,\nu)e^{-\tau_0(\nu)} + \frac{j_0(\nu)ds}{\tau_0(\nu)}\left(1 - e^{-\tau_0(\nu)}\right), \tag{5.18}$$

where $\tau_0(\nu) = \kappa_0(\nu)\rho_0 ds$ is the optical depth in the cell at the frequency ν. In order to get a complete solution, it is necessary to determine which cells in the computational domain are intersected by the ray, find the borders and solve eq. 5.18 within each cell. At the end the variation of the specific intensity along the ray is obtained. The accuracy of the solution depends on the spatial resolution of the grid. In many application, determining the optimal grid to solve the RT problem by using the RayT technique is not trivial. For example, adopting this approach to the commonly used adaptive-mesh refinement grids in hydrodynamic simulations can be too computationally expensive and a strategy to lower the spatial resolution without losing much information is needed. Similarly, high-optical depths require some approximations in order to not decompose the computational domain into many optically-thin cells. The main challenge of the RayT techniques is to find the optimal strategy to cast rays within the computational domain in order to determine the global solution for the specific intensity. An example of the different

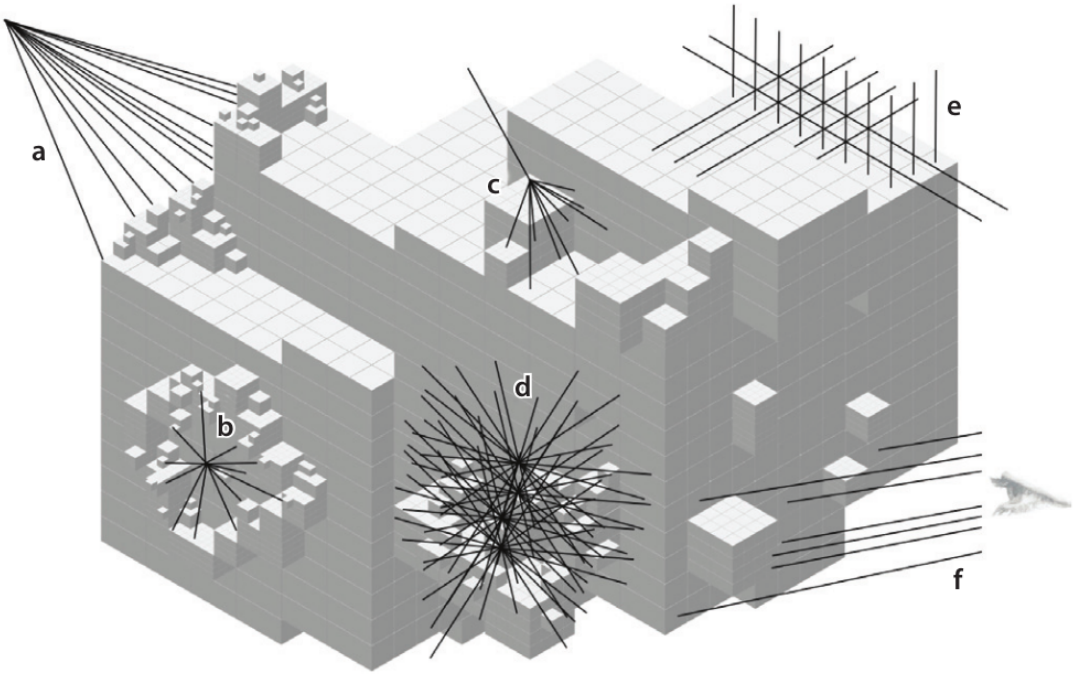

Figure 5.1: Example of the different types of rays used in 3D dust RT calculations, when applied to an AMR grid. For illustrative purposes, only a small fraction of the total number of rays used is shown. Rays can be cast for: (a) a radiation source outside the system; (b) a radiation source inside the system; (c) a scattering event; (d) a region with an optical depth $\tau_{\rm obs} \approx 1$; (e) a coarse portion of the grid; (f) rays to the observer. *Figure from Steinacker et al. (2013).*

types or rays adopted is shown in Fig. 5.1. The main rays are cast from each source toward each grid cells, and for each cell the optical depth from each source to the cell $\tau_{\rm obs}$ is pre-computed. The optical depth from each cell toward the observer is also computed, in order to determine the radiation that is detected from the system. For an in-depth discussion, we refer the reader to the review by Steinacker et al. (2013).

Including scattering radiation in the RayT technique requires a large computational effort, as it requires to store the mean intensity field $J(\mathbf{x}, \nu, t)$ and the specific intensity $I(\mathbf{x}, \mathbf{n}, \nu, t)$ in each cell, for every frequency and direction. The computation can then be performed by casting new rays from the scattering point and use pre-calculated optical depths to solve the RT equation. This treatment is a good approximation for the single-scattering case, when the optical depth is $\tau_{\rm obs} \ll 1$. In systems when multi-scattering has to be taken into account, then other techniques have to be preferred.

5.2.4 Monte Carlo approach

The Monte Carlo (MC) method refers to an algorithm that perform a computation by using a probabilistic approach. This strategy finds application in a large variety of fields, like mathematics, medicine, financial, and physics. All strategies that go below the definition of Monte Carlo have in common the basic idea to solve equations via a stochastic or probabilistic method. The history of MC method go back to the 1733, when Georges-Louis Leclerc, Comte de Buffon (1733), attempted to calculate π from the number of needles falling within an area. The first works in physics using techniques that we would define as MC today date back to the 1930s and 1940s in nuclear physics. Metropolis & Ulam (1949) published the first paper on the MC tech-

nique. In the following decades the MC methods started to be applied also to the RT problem. From then on, a large literature has been dedicated to solve the RT problem in dusty media with the MC approach, with descriptions more and more sophisticated in the last decades thanks to the increasing computational power available (Witt, 1977; Bianchi et al., 1996; Gordon et al., 2001; Ciardi et al., 2001; Juvela, 2005; Jonsson, 2006; Whitney, 2011; Steinacker et al., 2013).

The basic idea of the MC RT is to describe the radiation fields as a large number of photon packets (usually called simply photons). Each packet is followed individually along its travel through the dusty medium. Every property of the packet, such as its birth location, the initial propagation direction, its frequency, the next point of interaction and so on, is determined in a probabilistic way. This step is performed by generating random numbers from a probability distribution function (PDF) that describes the process considered. Essentially, the MC technique does not directly solve the RT equation, but it simulates the RT itself. Due to its stochastic nature, the MC approach does not provide an analytical or an exact solution to the problem, but an approximate one, that is subject to the Poisson noise. This downside can be easily overcome by using a sufficiently large number of photon packets.

The heart of the MC technique is the process of generating random numbers from a the PDF that describes a certain process. A PDF, $p(x)$, is a function that describes the probability density of a certain event, such that $p(x)dx$ is the probability of the event to return a value in the interval $(x, x + dx)$. The probability density is obviously normalized to 1. For this reason, instead of a simple random number generator, MC methods use a *pseudo-random number* generator, that is an algorithm that generates a random number uniformly distributed between 0 and 1. Such number is also called a *univariate*. Computer science have developed many methods to perform this operation in a reasonable accurate way (e.g. Press et al., 1992). In order to generate a random number from another PDF, numerical integration is needed and appropriate methods have been developed, such as the so-called transformation and rejection methods (Devroye, 1986; Press et al., 2002).

5.3 Outline of a Monte Carlo RT solver

The first step in the MC method applied to the RT problem is the generation of the photon packets from a source. For simplicity, let's consider the monochromatic problem at frequency ν, without time dependence, such that the source is described by the emissivity $j_*(\mathbf{x}, \mathbf{n}, \nu)$ and a monochromatic luminosity $L_{\rm tot}(\nu)$. If the emission of the source is sampled with N photon packets, than each of them carries a luminosity $L(\nu) = L_{\rm tot}(\nu)/N$. The starting position of the packet and its direction of propagation needs to be sampled from the PDF $p(\mathbf{x}, \mathbf{n})d\mathbf{x}d\mathbf{n}$ corresponding to the emissivity $j_*(\mathbf{x}, \mathbf{n}, \nu)$:

$$p(\mathbf{x}, \mathbf{n})d\mathbf{x}d\mathbf{n} = \frac{j_*(\mathbf{x}, \mathbf{n}, \nu)}{L_{\rm tot}(\nu)}. \tag{5.19}$$

In case of isotropic emission, the direction can be chosen randomly in the solid angle:

$$p(\mathbf{n})d\mathbf{n} = \frac{d\mathbf{n}}{4\pi} = \frac{sin\theta d\theta d\phi}{4\pi}. \tag{5.20}$$

Generating random values for θ and ϕ can be done by using the transformation method (see below).

The second step of the RT calculation is the determination of the following point of interaction between the photon packet and the intervening medium. This can be determined by consider

the optical depth τ along the path followed by the photon, which is distributed as an exponential function:

$$p(\tau)d\tau = e^{-\tau}d\tau. \tag{5.21}$$

The optical depth that a photon travels before an interaction is drawn from this distribution by using the transformation method. Given that $p(\tau)$ is normalized in the range $[0, +\infty[$, we can simply equate an univariate number $\mathcal{R}$ to the integral between 0 and τ (which is by definition between 0 and 1 as a pseudo-number):

$$\mathcal{R} = \int_0^{\tau} e^{-\tau'}d\tau' = -e^{-\tau} + 1. \tag{5.22}$$

Once $\mathcal{R}$ is generated, the previous expression can be inverted by finding the corresponding optical depth:

$$\tau(\mathcal{R}) = -\ln(1 - \mathcal{R}). \tag{5.23}$$

In a general, it is not always possible to integrate the PDF analytically when applying the transformation method. In that case, the the PDF needs to be integrated numerically. Once the random τ is assigned to the photon, it has to be compared to the optical depth $\tau_{\rm path}$ from the position of the photon to the surface. If $\tau > \tau_{\rm path}$, then the photon escapes the system, otherwise it interacts with the medium. In this case, the position of interaction s along the photon path is determined by converting the optical depth into a physical distance via the following integral equation:

$$\tau = \int_0^{s} \kappa_{\rm ext}(s', \nu)\rho(s')ds'. \tag{5.24}$$

This expression shows a similarity between the MC method and the rayT technique. In fact, the MC method essentially consists in applying the rayT technique to each photon packet. In general, MC methods adopt a dust grid in which the dust density and optical properties are discretized. Solving the previous equation is then done by simply summing along the path travelled by each photon until the optical depth τ is reached.

Once the position of interaction $\mathbf{x} + \mathbf{n}s$ is determined, the photon is moved to this location and it can be either absorbed or scattered. The probability of scattering is simply given by the dust albedo $\mathcal{A} = \kappa_{\rm sca}/\kappa_{\rm ext}$. By extracting a pseudo-random number $\mathcal{R}$, the type of interaction would be a scattering if $\mathcal{R} \leq \mathcal{A}$ and an absorption otherwise. In this way, the actual probability of interaction is accurately reproduced.

If the interaction is an absorption, then the photon energy is stored in the cell and it will later be used to compute the dust emissivity. If instead is a scattering event, a new direction of propagation for the photon has to be computed. The probability $p(\mathbf{n}')d\mathbf{n}'$ of the photon to be re-emitted along the new direction $\mathbf{n}'$ is obtained from the PDF of the corresponding phase function:

$$p(\mathbf{n}')d\mathbf{n}' = \frac{\Phi(\mathbf{n}, \mathbf{n}', \mathbf{x}, \nu)}{4\pi}, \tag{5.25}$$

where $\Phi(\mathbf{n}, \mathbf{n}', \mathbf{x}, \nu)$ is the scattering phase function introduced in Section 5.1 and for simplicity we ignored the possibility that the scattering alters the frequency of the photon.

The steps outline above are repeated until the photon is either absorbed or leave the computational domain. If the dust emission is included in the problem, after the path of each photon has been accounted for, the energy stored in each cell is used to compute the mean intensity of the radiation field $J(\mathbf{x}, \nu, t)$. Then, the dust temperature of the dust grains within each cell is computed, usually assuming energy balance, as described in Section 5.1. This allows the determination of the dust emissivity term $j_{\rm d}(\mathbf{x}, \mathbf{n}, \nu, t)$. It is common that this term is then accounted

for in a new MC computation, such that in the first run only emission from the primary source is considered (i.e. $j_*(\mathbf{x}, \mathbf{n}, \nu, t)$) and in the second one only $j_\mathrm{d}(\mathbf{x}, \mathbf{n}, \nu, t)$. The dust emissivity is then re-computed again and the procedure iterated until a satisfying level of converged is reached.

The MC method has several advantages in solving the RT problem in terms of its flexibility and the possibility to include almost every physical processes involved, with the appropriate stochastic description. However, in its simplest form it is very computational expensive and its application to the 3D dust problem is unfeasible. For this reason, many optimization techniques have been developed in the past years. The central idea of acceleration methods is the concept of *weight*: photon packets can be assigned a weight W, which can be thought as the fraction of the total luminosity carried by that photon. The weight is then used to describe *biased* behaviours. For example, in some cases it is convenient to extract photons from a PDF $q(x)$ that differs from the actual PDF $p(x)$. In these cases, the photon is assigned a weight $W = p(x)/q(x)$ to recover the correct behaviour. A practical example is the case of *biased emission*. In some problems, we are interested in the emission from a source only toward a certain direction. In this case it is convenient to increase the number of photons travelling toward this direction by the appropriate change of the photon weight, obtaining an accurate and fast result without the need to spend a lot of computation time to simulate also a non-interesting region. For an in-depth description of the most common used optimization techniques in 3D dust MC methods, we refer to Steinacker et al. (2002).

A particularly important optimization is the *peel-off* technique, which is very useful when it is necessary to simulate the image of the studied system from a given point of view. In the standard MC RT, only photons that escape the system toward the direction of the observer can contribute to the image, resulting in very poor S/N ratios. This problem can be overcome by assuming that all photons can contribute to the final image, considering every photon that is emitted by a source or scattered toward the direction of the observer during the computation. Each photon is then weighted with the factor:

$$W_\mathrm{obs} = p(\mathbf{n}_\mathrm{obs}) e^{\tau_\mathrm{obs}}, \tag{5.26}$$

where $\mathbf{n}_\mathrm{obs}$ is the direction from the point of emission/scattering toward the observer, τ_obs is the optical depth from that point to the observer, and $p(\mathbf{n}_\mathrm{obs})$ is the probability of the photon to be directed toward the observer.

Another common optimization strategy in MC RT is the use of polychromatic photons. When the problem is not restricted to a single frequency, it is possible to consider photon packets that have photon at all frequencies. In this way the problem is simultaneously solved at all frequencies, and use the biased technique to treat the frequency-dependent processes.

5.4 Radiative transfer setup with SKIRT

Throughout the work described in this Thesis, we use the code SKIRT[2] to post-process snapshots from the hydrodynamic simulations described in Section 4.3 and Section 4.3, in order to make observational predictions of their UV-to-IR spectral energy distribution. SKIRT solves the continuum radiative transfer problem in a dusty medium with a Monte-Carlo approach, by sampling the SED of the sources with a finite number of photon packets (in the following simply referred to as *photons*). Photons are scattered and/or absorbed by dust grains in the simulation volume according to their properties. Dust grains, after being heated up, thermally re-emit the absorbed

[2]Version 8, http://www.skirt.ugent.be.

energy at IR wavelengths. One of the main advantages of the SKIRT code is its flexibility: it allows the user to handle input data from different numerical codes (e.g. Adaptive Mesh Refinement and Smooth Particle Hydrodynamic codes), to account for different dust properties (i.e. grain size distribution and composition), to implement different SEDs for the radiating sources (e.g. stars and accreting BHs), to include many physical mechanisms (e.g. dust stochastic heating and self-absorption).

To relate the energy absorbed by dust with its wavelength-dependent emissivity we adopt the dust models described in Section 5.4. Dust is implemented in SKIRT as described in Section 5.4. We describe the SED adopted in different RT runs for stars and accreting BHs in Section 5.4. We extensively describe the AGN SED we implemented in SKIRT in Section 5.4.

Dust properties

Dust formation, growth and destruction processes are not tracked in the hydrodynamic simulations considered here. Similarly to other RT works (Behrens et al., 2018; Arata et al., 2019; Liang et al., 2019), we derive the dust mass distribution by assuming a linear scaling with the gas metallicity[3] (Draine et al., 2007), parametrizing the mass fraction of metals locked into dust as:

$$f_d = M_{\rm d}/M_Z, \tag{5.27}$$

where $M_{\rm d}$ is the dust mass and M_Z is the total mass of all the metals in each gas particle in the hydrodynamical simulation (see Section 4.3). Gas particles hotter than 10^6 K, are considered dust-free as at these temperatures thermal sputtering is very effective at destroying dust (Draine & Salpeter, 1979; Tielens et al., 1994; Hirashita, 2015). This assumption does not affect the main results of our work, as discussed in Section A.2.

The choice of f_d directly affects the total dust content. The RT calculation is sensitive to the f_d value, which is poorly constrained by high-redshift galaxies observations (see Wiseman et al. 2017 and references therein) and theoretical models (Nozawa et al., 2015). In particular, recent theoretical works (Asano et al., 2013a; Aoyama et al., 2017) suggest that f_d is constant in the early stages of galaxy evolution and then it grows with metallicity up to the Milky-Way (MW) value of $f_d = 0.3$ when/if dust growth becomes important. However, the efficiency of dust growth in the ISM of early galaxies is highly debated (Ferrara et al., 2016). In this work, we consider a constant value of f_d, and focus our attention on how the dust content of galaxies affects their panchromatic SED.

We mainly adopt two different f_d values for the normalization: i) a MW like value ($f_d = 0.3$); ii) a lower value ($f_d = 0.08$) tuned for hydro-simulations (Pallottini et al., 2017; Behrens et al., 2018) to reproduce the observed SED of a $z \sim 8$ galaxy (Laporte et al., 2017).

The properties of dust as chemical composition and grain size distribution are not known in early (AGN-host) galaxies (see Section 3.3 for an extensive discussion). Throughout this Thesis, we consider dust compositions and grain size distributions appropriate for the SMC and MW, by using the results[4] of Weingartner & Draine (2001). In chapter 8, we consider also dust models derived from the Weingartner & Draine (2001) models by modifying the grain size distribution (see Section 8.4). We defer the inclusion of a SN-type extinction curve to a future work. We note that these extinction curves are somewhat intermediate between the SMC and MW curve, with the notable difference that SN-type curves are essentially flat in the wavelength range $1700 - 2500$ Å

[3]Throughout this Thesis the gas metallicity is expressed in solar units, using $Z_\odot = 0.013$ as a reference value (Asplund et al., 2009).

[4]We consider the revised optical properties evaluated in Draine (2003a,b,c).

Dust implementation in SKIRT

Dust is distributed in the computational domain in an octree grid, whose maximum number of levels of refinement for high dust density regions is chosen according to the spatial resolution of the hydrodynamic simulation considered. For the runs from B18, we adopt 8 levels of refinement, achieving a spatial resolution of ≈ 230 pc in the most refined cells, comparable with the softening length in the hydrodynamic simulation (≈ 200 pc at $z = 6$). For the runs from V21, we adopt 10 levels of refinement, with the highest resolved cells having a size of ≈ 59 pc, consistent with the softening length (≈ 87 ppc at $z = 6$). When distributing the dust content derived from the hydrodynamical simulation into an octree grid, a kernel-based interpolation is required in order to convert the dust content from a particles-based distribution into an octree geometry. This procedure leads to a discrepancy between the total amount of dust carried by the SPH gas particles imported from the hydrodynamical simulation and the effective dust content in the computational domain used for the RT calculation. Therefore, it is important to check that the structure of the dust grid adopted achieves sufficient convergence relative to the overall dust content. We checked that the relative difference in the overall dust content is within 1% in all the runs performed. We sample in SKIRT the grain size distribution of graphite, silicates and polycyclic aromatic hydrocarbons (PAHs) using 5 bins for each component. In the SMC model the fraction of dust in PAHs is set to zero.

Grain temperature and emissivity are evaluated by imposing energy balance between the local radiation field and dust re-emission. By default, when dust emission photons propagate, SKIRT accounts for the self-absorption by dust, but it does not take this absorption into account when computing the dust temperature, unless the self-absorption flag is turned on. As this effect may be relevant if dust is IR-optically thick, we have enabled a self-consistent evaluation of the dust temperature, iterating the RT calculation for dust absorption and re-emission until the dust IR luminosity converges within 3%.

We also include non-local thermal equilibrium (NLTE) corrections to dust emission, which include the contribution from small grains that are transiently heated by individual photons. In this case grains of different sizes are no longer at a single equilibrium temperature, but follow a temperature distribution. Behrens et al. (2018) found in their calculations that stochastic heating affects mostly the MIR portion of the SED (rest-frame wavelength $\lesssim 80$ μm) but it has a minor impact on the FIR and (sub)mm emission.

We do not include heating from CMB radiation. As discussed in Section 6.3.2, only a small fraction of dust grains is at a temperature comparable to $T_{\rm CMB}$. We expect this effect to be negligible, as seen a posteriori from the RT results.

We do not include any subgrid model for dust clumpiness. Recent works (e.g. Camps et al. 2016; Trayford et al. 2017; Liang et al. 2021a) that account for subresolution structures of birth clouds harboring young stars (Jonsson et al., 2010), whose typical scales are not resolved by the hydrodynamical simulations, are based on SED templates (Groves et al., 2008) not consistent with our fiducial set up. The stellar emission in the Groves et al. (2008) template is, in fact, calculated from Starbust99 models (Leitherer et al., 1999), by assuming a Kroupa (2002) initial mass function, whereas we model stellar emission using the Bruzual & Charlot (2003) model (see Section 5.4), based on the Chabrier (2003) IMF. Moreover, they include PAH molecules in the dust composition that are instead not considered in our work. We notice that Liang et al. (2021a) find that the Groves et al. (2008) template mainly affects the IR emission from PAHs which is enhanced up to 50% (see Fig. 23 of their paper). Given that in our model we adopt an SMC dust composition (i.e. no PAHs), we do not expect that the inclusion of a subgrid model that accounts for dust clumpiness would significantly affect the main results of our work.

Radiation from stars and AGN

The ultraviolet (UV) radiation field mainly responsible for dust heating is provided by stellar sources and black holes. We describe in the following how the two components are implemented in our model.

Stellar particles in the simulation represent a Single Stellar Population (SSP), i.e. a cluster of stars formed at the same time and with a single metallicity. Given the mass, age and metallicity of the stellar particle imported, SKIRT builds the individual SEDs according to the Bruzual & Charlot (2003) family of stellar synthesis models, placing the sources at the locations of the stellar particles.

Black holes are treated as point source emitters as the typical sizes of the accretion disk and the dusty torus are much smaller ($\lesssim 10$ pc) than the width of the most refined grid cells (≈ 230 pc, see Section 5.4). We implement their emission in SKIRT adopting a SED as described in Section 5.4.

The radiation field is sampled using a grid covering the *rest-frame* wavelength range[5] $[0.1 - 10^3]$ μm. The choice of the lower limit is quite common for RT simulations in dusty galaxies (Schneider et al., 2015; Behrens et al., 2018) and it is motivated by the fact that codes like SKIRT typically do not account for the hydrogen absorption of ionising photons ($\lambda < 912$ Å). The choice of 10^3 μm as the upper limit of the wavelength grid is motivated by the fact that the intrinsic emission from stars and BHs is negligible above this limit. The base wavelength grid is composed of 200 logarithmically spaced bins. If a MW-type dust is used, we add a nested grid with 200 logarithmically spaced bins in the wavelength range $[1-40]$ μm in order to better capture PAH emission (the composite wavelength grid has a total of 320 bins in this case).

A total of 10^6 photon packets per wavelength bin is launched from each source, i.e. stellar particles and BHs[6]. We collect the radiation escaping our computational domain for the 6 lines-of-sight perpendicular to the faces of the cubic computational domain.

AGN Spectral Energy Distribution

The SED of an AGN is shaped by the numerous physical mechanisms involved in the process of gas accretion onto the BH (see discussion in Section 2.2). AGN SED templates are typically based both on theoretical arguments and observations (e.g. Shakura & Sunyaev, 1973; Vanden Berk et al., 2001; Sazonov et al., 2004; Manti et al., 2016; Shen et al., 2020a), possibly including the dusty torus modelling (Schartmann et al., 2005; Nenkova et al., 2008; Stalevski et al., 2012, 2016). For this work, we adopt a composite power-law for the AGN emission written as:

$$L_\lambda = c_i \left(\frac{\lambda}{\mu\mathrm{m}}\right)^{\alpha_i} \left(\frac{L_\mathrm{bol}}{\mathrm{L}_\odot}\right) \mathrm{L}_\odot\ \mu\mathrm{m}^{-1}, \tag{5.28}$$

where i labels the bands in which we decompose the spectra and the coefficients c_i are determined by imposing the continuity of the function based on the slopes α_i. The coefficients c_i and α_i adopted and the relative bands are reported in Table 5.1 and they are chosen as described in the following.

[5]The total AGN bolometric luminosity is distributed from the X-ray to the IR according to the SED adopted (see Section 5.4). The choice of the wavelength range adopted in our simulations affects the fraction of the AGN bolometric luminosity that effectively enters in the calculation (see Fig. 5.2). For the *fiducial* SED introduced in Sec. 5.4, this fraction is $\approx 60\%$, whereas it is $\approx 40\%$ for the *UV-steep* SED.

[6]We verified that the number of packets used is sufficient to achieve numerical convergence by comparing the results with control simulations with 5×10^5 photon packets per wavelength bin.

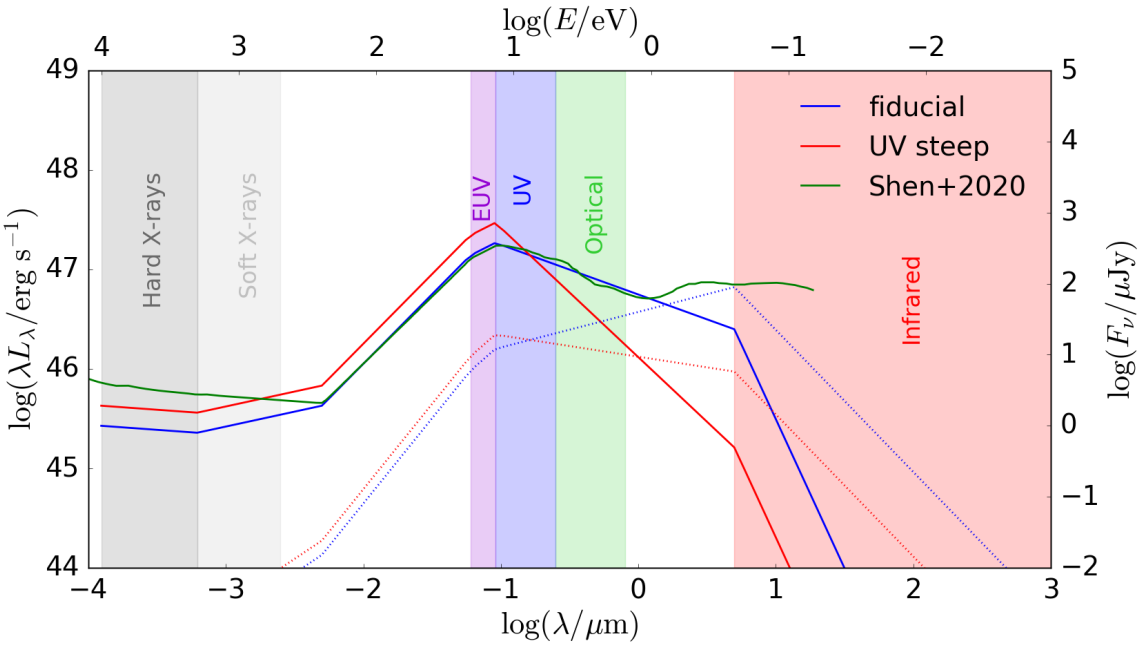

Figure 5.2: AGN SED for a bolometric luminosity $L_{\rm bol} = 10^{13}$ $L_\odot$: the *fiducial* SED ($\alpha_{\rm UV} = -1.5$) is shown with a blue thick line; the *UV-steep* SED ($\alpha_{\rm UV} = -2.3$) is shown with a red thick line. We plot the SED template derived in Shen et al. (2020a) for comparison with a thick green line, re-scaling it in order to have the same $L_{2500\text{Å}}$ of the fiducial SED. The SEDs differ mainly at wavelength longer than the UV band, with the UV-steep SED dropping faster than the other two. The fiducial SED is in very good agreement with the Shen et al. (2020a) SED up to ≈ 2 μm, from where the contribution by dust in the torus and in the galaxy included in their IR template begins to dominate the emission in their SED. As a reference for the SED plots in the following, we also plot our two SEDs as the F_ν (in μJy) vs λ with dotted lines, keeping the same colour legend.

For the X-ray band, based on the results by Piconcelli et al. (2005) and Fiore et al. (1994) in the hard ($2 - 10$ keV, $\alpha_{\rm X,hard} = -1.1 \pm 0.1$) and soft ($0.5 - 2$ keV, $-0.7 < \alpha_{\rm X,soft} < 0.3$) band, respectively, we consider $\alpha_{\rm X,hard} = -1.1$ and $\alpha_{\rm X,soft} = -0.7$. Consistently with Shen et al. (2020a), in the wavelength range $50 - 600$ Å we use $\alpha = 0.4$ (the slope chosen for the soft X-ray band is then adopted up to 50 Å for continuity). For the Extreme UV band (EUV, $600 < \lambda < 912$ Å) we use $\alpha_{\rm EUV} = -0.3$ as in Lusso et al. (2015). We also note that this value is consistent with the constraints by Wyithe & Bolton (2011, $-1.0 < \alpha_{\rm EUV} < -0.3$) based on the analysis of near-zones observed around high redshift quasars.

The analysis of a large sample (4576) of $z \lesssim 2.2$ quasars (Richards et al., 2003) spectra in the range $1200 \lesssim \lambda \lesssim 6000$ Å has shown that the spectral slopes are distributed in the range ($-2.6 < \alpha < -0.2$) and peak around $\alpha = -1.6$. In the $912 < \lambda < 2500$ Å band, Lusso et al. (2015) have constructed a stacked spectrum of 53 quasars at $z \sim 2.4$ finding $\alpha = -1.39 \pm 0.01$. Moreover, Gallerani et al. (2010) have analysed 33 quasars in the redshift range $3.9 \lesssim z \lesssim 6.4$ finding that unreddened quasars are characterised by $\alpha = -1.7 \pm 0.5$, whereas reddened quasars prefer steeper slopes ($\alpha < -2.3$). Finally, from a theoretical point of view, the classical black-body composition for a Shakura & Sunyaev (1973)-disk predicts that $F_\nu \propto \nu^{1/3}$ (see also Section 2.2.3), which translates into $\alpha = -2.3$. Given the uncertain value of the slope for wavelengths longer than 912 Å, we consider two possible values for the slope in the range from the UV to NIR: $\alpha_{\rm UV} = -1.5$, which is representative of unreddened quasars, and $\alpha_{\rm UV} = -2.3$. We will refer to these two models as the *fiducial* and *UV-steep* model, respectively.

At longer wavelengths, the intrinsic AGN emission is expected to follow the Rayleigh-Jeans tail regime $F_\nu \propto \nu^2$, which corresponds to $\alpha_{\rm IR} = -4$. The transition between the UV slope and

the IR one increases with the black hole mass (Shakura & Sunyaev, 1973; Pringle, 1981; Sazonov et al., 2004). In this work, we adopt a transition wavelength $\lambda_{\rm trans} = 5\ \mu$m. This component represents the IR emission from the accretion disk only. We did not include the emission from the hot dust component from the torus because we cannot resolve the scales ($1 - 10$ pc) of the torus itself. We discuss how this affects our results in Section 6.5.1.

	hard X	soft X	X to EUV	EUV	UV to NIR	NIR to FIR
	$[2-10]$ keV	$[6.2-60]$ Å	$[50-600]$ Å	$[600-912]$ Å	$[0.0912-5]\ \mu$m	$[5-10^3]\ \mu$m
c (fiducial)	2	0.042	14.133	1.972	0.111	6.225
α (fiducial)	-1.1	-0.7	0.4	-0.3	-1.5	-4.0
c (UV-steep)	0.003	0.066	22.499	3.140	0.026	0.402
α (UV-steep)	-1.1	-0.7	0.4	-0.3	-2.3	-4.0

Table 5.1: Coefficients of our AGN SEDs models as expressed in eq. 5.28. The slopes α_i and the ranges of the piece-wise decomposition were chosen as explained in Section 5.4. Imposing the continuity of the function determines the coefficients c_i. The SED built in this way is by construction normalised to the bolometric luminosity of the source expressed in $L_\odot$ according to eq. 5.28.

The fiducial and UV-steep SEDs adopted in this work are shown in Fig. 5.2 with blue and red lines, respectively. We also report with a green line the SED derived by Shen et al. (2020a). Our bolometric corrections[7] (reported in Table 5.2) are consistent with the ones by Shen et al. (2020a) (reported in the top panel of their Fig. 2), for $L_{\rm bol} \approx 10^{47}$ erg s^{-1}. We further calculate the $\alpha_{\rm OX} = 0.384 \log L_\nu(2{\rm keV})/L_\nu(2500\text{Å})$ index for our SEDs and find that it is in agreement with observations of $z \sim 6$ quasars (e.g. Nanni et al., 2017; Gallerani et al., 2017b; Vito et al., 2019b).

SED model	$\frac{L_{\rm bol}}{L_{\rm X,hard}}$	$\frac{L_{\rm bol}}{L_{\rm X,soft}}$	$\frac{L_{\rm bol}}{L_{\rm UV}}$	$\frac{L_{\rm bol}}{L_{\rm B}}$	$\alpha_{\rm OX}$
fiducial	130	130	3.4	6.0	-1.65
UV-steep	80	81	3.1	13.6	-1.51

Table 5.2: Bolometric corrections ($L_{\rm bol}/L_{\rm band}$) and $\alpha_{\rm OX}$ for the *fiducial* ($\alpha_{\rm UV} = 1.5$) and *UV-steep* ($\alpha_{\rm UV} = 2.3$) AGN SED models adopted in this work. The bands used to compute the bolometric corrections are defined as: hard X-ray $[2-10]$ keV, soft X-ray $[0.5-2]$ keV, UV $[0.1-0.3]\ \mu$m. $L_{\rm B}$ is defined as λL_λ at $\lambda = 4400$ Å. The two models mostly differ for the luminosity in the B band. For a bolometric luminosity $L_{\rm bol} = 10^{47}$ erg s^{-1}, our bolometric corrections are consistent with the observational constraints reported in the top panel of Fig. 2 by Shen et al. (2020a).

[7] Consistently with Shen et al. (2020a), we express the UV band luminosity as $\nu_{1450\text{Å}} L_{\nu_{1450\text{Å}}}$, the B band as $\nu_{4400\text{Å}} L_{\nu_{4400\text{Å}}}$, whereas the soft [hard] X-ray luminosity is the integrated luminosity in the 0.5-2 [2-10] keV band.

Part III

Analysis and results

AGN imprints on the Infrared emission of $z \sim 6$ galaxies

6

6.1 Introduction

In Section 2.5, we discussed how the existence of supermassive black holes (SMBHs, 10^{8-10} $M_\odot$) within the first Gyr of the Universe (e.g. Mortlock et al., 2011; Wu et al., 2015; Bañados et al., 2018; Yang et al., 2020b; Wang et al., 2021a) is one of the biggest puzzles in modern astrophyiscs (e.g. Tanaka & Haiman, 2009; Volonteri, 2010; Schleicher et al., 2013; Latif & Ferrara, 2016; Valiante et al., 2017; Woods et al., 2019; Inayoshi et al., 2020). Whether they form from light seeds (10^{1-2} $M_\odot$) as remnants of Pop-III stars (e.g. Madau et al., 2014; Pacucci et al., 2015a; Lupi et al., 2016), intermediate seeds (10^{3-4} $M_\odot$) from runaway collisions in stellar clusters (Devecchi & Volonteri, 2009; Katz et al., 2015; Yajima & Khochfar, 2016; Sakurai et al., 2017), or massive seeds (10^{5-6} $M_\odot$) from direct collapse of pristine halos (e.g. Schleicher et al., 2013; Ferrara et al., 2014; Latif & Ferrara, 2016; Inayoshi et al., 2020) remain open questions. Moreover, how they were able to grow up to the observed masses in such a short time period, possibly experiencing episodes of super-Eddington accretion (e.g. Pacucci et al., 2015a; Pezzulli et al., 2016), is not yet fully understood.

This problem is further complicated by the unsuccessful search for high-z AGN powered by $\sim 10^{6-7}$ $M_\odot$ BHs (e.g. Xue et al., 2011; Cowie et al., 2020). Whether these sources are too rare (Pezzulli et al., 2017a), and/or too faint to be detected by current optical/NIR survey (see Section 2.3), and/or their optical/UV emission is obscured by dust (see Chapter 3), remains unclear. This latter hypothesis is also supported by X-ray observations suggesting that the fraction of obscured AGN increases with redshift (e.g. Vito et al., 2014, 2018). Moreover, the early growth of SMBHs, typically characterized by low accretion rates, is expected to be buried in a thick cocoon of dust and gas (e.g. Hickox & Alexander, 2018, for a review on this subject). In this scenario, a certain fraction of UV photons are absorbed and/or scattered by dust grains in gas clouds in the host galaxy. By transferring energy and momentum to the surrounding dusty environment, AGN radiation can substantially affect the conditions of the interstellar (ISM) and circumgalactic (CGM) medium of the host galaxy in several ways (see Section 2.4). UV radiation heats the dust, leading grains to re-emit in the far-infrared.

In the latest years, ALMA and NOEMA observations have provided the opportunity of studying the ISM/CGM properties of bright $z \sim 6$ quasar hosts via rest frame FIR emission lines, such as the [CII] line at 158 micron and CO rotational transitions, and the corresponding dust continuum emission (see Section 2.3 and reference therein). ALMA data of $z \sim 8$ Lyman Break Galaxies (Laporte et al., 2017; Bakx et al., 2020) have suggested the presence in these sources of dust hotter than expected ($T \sim 60 - 90$ K, Behrens et al. 2018; Arata et al. 2019; Sommovigo et al. 2020). The origin of warm dust in early galaxies can be traced back to their (i) large SFR surface

densities that favour an efficient heating of dust grains (Behrens et al., 2018) and (ii) more compact structure of molecular clouds (MC) that delays their dispersal by stellar feedback, implying that a large fraction ($\sim 40\%$) of the total UV radiation remains obscured (Sommovigo et al., 2020). Another possibility concerns the presence of obscured, accreting, massive ($\sim 10^8\ M_\odot$) BHs, whose UV luminosity is absorbed by dust located in the ISM of the host ($\lesssim 1$ kpc) and/or into a central obscurer, closer to the active nuclei (~ 1 pc), and heated to temperatures as high as 80-500 K, respectively (Orofino et al., 2021). According to this scenario, buried AGN should be searched for among Lyman break galaxies (LBGs) populating the bright-end of their UV luminosity function ($-24 < M_{\rm UV} < -22$), where indeed a large fraction of objects consists of spectroscopically confirmed AGN (Ono et al., 2018).

Obscured AGN may therefore represent a bridge between LBGs and bright quasars in the galaxy formation process. In this appealing scenario, the following questions arise: *(i) If high-z galaxies contain an obscured AGN, does this imply warmer dust temperatures? (ii) Is there a relation between the dust temperature and the BH accretion rate?(iii) What are the most promising spectral ranges and observational strategies to detect obscured AGN?* To answer these questions it is necessary to build up a model that follows the co-evolution of BHs with their host galaxy from their birth up to the formation of SMBHs powering $z \sim 6$ quasars, while accounting for AGN and stellar feedback. The final aim is to produce synthetic multi-wavelength SEDs that can be directly compared with the aforementioned observations of $z \sim 6$ quasars to validate the underlying galaxy-BH formation model. This can be done by post-processing cosmological hydro-dynamical simulations with dust radiative transfer calculations.

Several works in the past years made use of radiative transfer simulations to understand the AGN contribution to the total IR emission of a galaxy, mainly focusing on Ultra Luminous Infrared Galaxies (ULIRGs), and late-stage mergers (e.g. Chakrabarti et al., 2007; Chakrabarti & Whitney, 2009; Younger et al., 2009; Snyder et al., 2013; Roebuck et al., 2016; Blecha et al., 2018). However, these studies are limited up to $z \sim 3$ and they rely on hydrodynamical simulations in which the initial conditions of both the dark matter and gas components were set with analytical prescriptions. Recently, Schneider et al. (2015) have studied the origin of the infrared emission in SDSS J1148+5251, a $z \sim 6$ quasar, by applying dust RT calculations to the output of a semi-analytical merger tree code finding that the dust heating by the AGN radiation may contribute up to 70% of the total IR luminosity. This is consistent with the results found by Li et al. (2008a) that computed RT calculations on hydrodynamical simulations of luminous quasars to reproduce the SED of SDSS J1148+5251. They also found that the AGN contribution to the IR emission is significant, because dust heating is dominated by the central source during the quasar-phase.

In this Chapter, we investigate the imprints of AGN in the IR emission of $z \sim 6$ galaxies by post-processing cosmological hydrodynamic simulations of SMBHs formation (Barai et al., 2018, hereafter B18) with dust RT calculations performed by using the code SKIRT (Baes & Camps, 2015; Camps et al., 2016). The B18 simulations studied the growth of SMBHs ($10^8 - 10^9$ M$_\odot$ at $z = 6$) and the impact of different AGN feedback prescriptions on their host galaxies, residing in a $\sim 10^{12}$ M$_\odot$ dark matter halo.

This Chapter is organised as follows: in Section 6.2 we summarise the main properties of the hydrodynamical simulations adopted (Section 6.2.1) and the model adopted for the radiative transfer calculations (Section 6.2.2). We present our results in Section 6.3 and we compare them with observations in Section 6.4. We then make predictions for the proposed mission ORIGINS in Section 6.5. Finally we summarise our results in Section 6.6 along with our conclusions.

6.2 Numerical setup

6.2.1 Hydrodynamic Simulations

In this work we make use of the cosmological hydrodynamic simulations of B18, which we briefly summarise here (see Section 4.3 and the original paper for more details). These simulations are performed with a modified version of the SPH N-body code GADGET-3 (Springel, 2005) and follow the evolution of a comoving volume of $(500\ \mathrm{cMpc})^3$, starting from cosmological initial conditions (IC) generated with MUSIC (Hahn & Abel, 2011) at $z = 100$, and zooming-in on the most massive dark matter (DM) halo inside the box down to $z = 6$[1].

Radiative heating and cooling according to the tables computed by Wiersma et al. (2009), which also include metal-line cooling. The multiphase model by Springel & Hernquist (2003) is implemented to describe ISM processes on subgrid resolution. Star formation is implemented with a density-based criterion, with a density threshold for star formation of $n_{\mathrm{SF}} = 0.13\ \mathrm{cm}^{-3}$ and assuming a Chabrier (2003) initial mass function (IMF) in the mass range $0.1 - 100\ \mathrm{M}_{\odot}$. Stellar evolution and chemical enrichment are computed following Tornatore et al. (2007), and kinetic feedback from SNe is also included.

BHs are included in the simulations by seeding a $M_{\mathrm{BH}} = 10^5\ \mathrm{M}_{\odot}$ BH in a DM halo when it reaches a total mass of $M_{\mathrm{h}} = 10^9\ \mathrm{M}_{\odot}$, and they are allowed to grow via gas accretion or mergers with other BHs. The former is modelled via the Bondi-Hoyle-Littleton scheme $\dot{M}_{\mathrm{Bondi}}$ (Hoyle & Lyttleton, 1939; Bondi & Hoyle, 1944; Bondi, 1952):

$$\dot{M}_{\mathrm{Bondi}} = \alpha \frac{4\pi G^2 M_{\mathrm{BH}}^2 \rho}{\left(c_{\mathrm{s}}^2 + v^2\right)^{3/2}}, \tag{6.1}$$

where G is the gravitational constant, ρ is the gas density, c_{s} is its sound speed, and v is the velocity of the BH relative to the gas, and $\alpha = 100$ is a numerical boost factor adopted to compensate for the low resolution in the regions closest to the BHs. Gas accretion is also capped at the Eddington rate $\dot{M}_{\mathrm{Edd}}$, so the final BH accretion rate $\dot{M}_{\mathrm{BH}}$ reads as follows:

$$\dot{M}_{\mathrm{BH}} = \min(\dot{M}_{\mathrm{Bondi}}, \dot{M}_{\mathrm{Edd}}). \tag{6.2}$$

During its growth a BH radiates away a fraction of the accreted rest-mass energy, with a bolometric luminosity

$$L_{\mathrm{bol}} = \epsilon_{\mathrm{r}} \dot{M}_{\mathrm{BH}} c^2, \tag{6.3}$$

where c is the speed of light and $\epsilon_{\mathrm{r}} = 0.1$ is the radiative efficiency (e.g. Shakura & Sunyaev, 1973). A fraction $\epsilon_{\mathrm{f}} = 0.05$ of this energy is distributed to the surrounding gas in a kinetic form.

In this work we consider the following three runs performed by B18, starting from the same ICs:

- *noAGN*: control simulation without BHs.
- *AGNsphere*: simulation accounting for BH accretion and AGN feedback. The kinetic feedback is distributed according to a spherical geometry.
- *AGNcone*: same as the *AGNsphere* run, but with kinetic feedback distributed inside a bi-cone with an half-opening angle of $45°$.

[1] In the low-resolution DM-only simulation, the most massive halo at $z = 6$ has a mass of $M_{\mathrm{halo}} = 4.4 \times 10^{12}\ \mathrm{M}_{\odot}$ (virial radius $R_{200} = 73$ pkpc), massive enough to host luminous AGN, as suggested by clustering studies (e.g. Allevato et al., 2016).

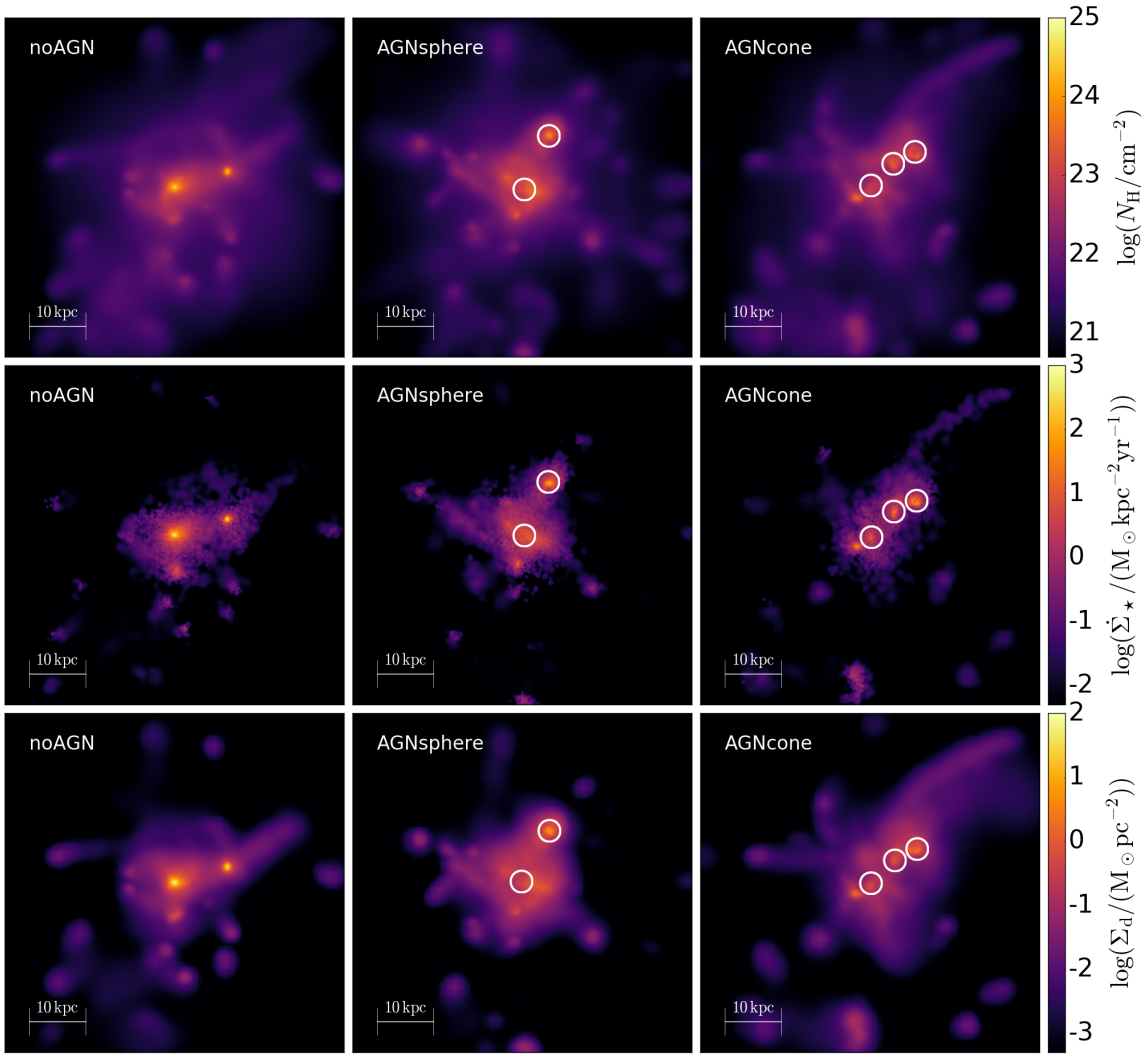

Figure 6.1: Morphology of the most massive halo at $z = 6.3$ inside a cubic box of 60 kpc size for the three cosmological simulations of B18: *noAGN* (left column), *AGNsphere* (middle column) and *AGNcone* (right column). The top, middle and bottom panel show the hydrogen column density, the star formation rate and the dust surface density (assuming a dust-to-metal ratio $f_d = 0.08$, see Section 5.4), respectively. White empty circles show the location of BHs accreting at $\dot{M}_{\rm BH} > 1\ {\rm M}_\odot\,{\rm yr}^{-1}$.

In Fig. 6.1 we show the hydrogen column density (top row) and the star formation rate (middle row) for the zoomed-in halo in the three simulations for a line of sight aligned with the angular momentum of the particles inside the selected region. In the following, this is our reference line of sight. From the top row, it can be seen that the central region, corresponding to the main galaxy, is characterised by the highest column density in all the runs. It reaches values of $N_H \sim 6 \times 10^{24}\ {\rm cm}^{-2}$ in the *noAGN* run, whereas it is an order of magnitude lower when AGN feedback is included. This is because kinetic feedback kicks gas away from the accreting BHs. In turn, the decreased gas density quenches the overall SFR density. In fact, star formation rate densities $\Sigma_{\rm SFR}$ as high as $\Sigma_{\rm SFR} \approx 600\ {\rm M}_\odot\,{\rm yr}^{-1}\,{\rm kpc}^{-2}$ are found in the *noAGN* run, in sharp contrast with those in the *AGNsphere* ($\approx 130\ {\rm M}_\odot\,{\rm yr}^{-1}\,{\rm kpc}^{-2}$, characterized by a total BH accretion rate $\dot{M}_{\rm BH} = 3.1\ {\rm M}_\odot\,{\rm yr}^{-1}$), and *AGNcone* ($\approx 50\ {\rm M}_\odot\,{\rm yr}^{-1}\,{\rm kpc}^{-2}$, $\dot{M}_{\rm BH} = 89\ {\rm M}_\odot\,{\rm yr}^{-1}$) cases. The same trend is observed also for the total SFR, as reported in Table 4.1.

run	AGN feedback	$M_{\rm gas}$ [$M_\odot$]	$M_\star$ [$M_\odot$]	SFR [$M_\odot\,{\rm yr}^{-1}$]	$\dot{M}_{\rm BH}$ [$M_\odot\,{\rm yr}^{-1}$]	$M_{\rm UV}$ [mag]
noAGN	no	2.9×10^{11}	1.2×10^{11}	600	-	
AGNsphere	spherical	2.1×10^{11}	6.5×10^{10}	312	3.1	-24.32
AGNcone	bi-conical	1.4×10^{11}	7.0×10^{10}	189	89	-27.97

Table 6.1: Summary of the hydrodynamic runs of B18 used in this work. For each run, we indicate the feedback model used in the simulation and the main physical properties of the zoomed-in halo at $z = 6.3$ within a cubic region of 60 kpc size (that corresponds to $\sim 50\%$ of the virial radius): gas mass ($M_{\rm gas}$), stellar mass ($M_\star$), star formation rate (SFR, averaged over the last 10 Myr), and the sum of the accretion rate of all the black holes (BHs) in the selected region ($\dot{M}_{\rm BH}$). We further associate to $\dot{M}_{\rm BH}$ an intrinsic UV magnitude $M_{\rm UV}$ (see Section A.3).

6.2.2 Radiative Transfer setup

We post-process the snapshots at $z = 6.3$ of the three selected hydrodynamic simulations in B18 by using the publicly available Monte-Carlo, radiative transfer code SKIRT[2](Baes & Camps 2015; Camps et al. 2016. see Chapter 5 for more details).

Dust formation, growth and destruction processes are not tracked in the simulations by B18. Similarly to other RT works (Behrens et al., 2018; Arata et al., 2019; Liang et al., 2019), we derive the dust mass distribution by assuming a linear scaling with the gas metallicity[3] (Draine et al., 2007), parametrizing the mass fraction of metals locked into dust as:

$$f_d = M_{\rm d}/M_Z, \tag{6.4}$$

where $M_{\rm d}$ is the dust mass and M_Z is the total mass of all the metals in each gas particle in the hydrodynamical simulation. The choice of f_d directly affects the total dust content. The RT calculation is sensitive to the f_d value, which is poorly constrained by high-redshift galaxies observations (see Wiseman et al. 2017 and references therein) and theoretical models (Nozawa et al., 2015). In this work, we consider a constant value of f_d, and focus our attention on how the dust content of galaxies affects their panchromatic SED.

We adopt two different f_d values for the normalization: i) a MW like value ($f_d = 0.3$); ii) a lower value ($f_d = 0.08$) tuned for hydro-simulations (Pallottini et al., 2017; Behrens et al., 2018) to reproduce the observed SED of a $z \sim 8$ galaxy (Laporte et al., 2017). The dust surface density distribution derived in the $f_d = 0.08$ case is shown in the bottom row of Fig. 6.1. High dust surface density regions correspond to active star forming regions where gas metal enrichment is more pronounced. Therefore, gas and dust density, and SFR are generally correlated in our simulations, as can be seen in Fig.6.1.

The dust properties, such as chemical composition and grain size distribution, are not known in early (AGN-host) galaxies, as also discussed in Section 3.3. For the time being, we assume a dust composition and grain size distribution appropriate for the SMC by using the results[4] of Weingartner & Draine (2001).

Dust is distributed in the computational domain in an octree grid with a maximum of 8 levels of refinement for high dust density regions, achieving a spatial resolution of ≈ 230 pc in the most refined cells, comparable with the softening length in the hydrodynamic simulation (≈ 200 physical pc at $z = 6.3$). We verify in Section A.1 that the number of refinement levels adopted in our fiducial setup is sufficient to achieve converge of the results. Adopting an SMC-like dust, the

[2]Version 8, `http://www.skirt.ugent.be`.

[3]Throughout this chapter the gas metallicity is expressed in solar units, using $Z_\odot = 0.013$ as a reference value (Asplund et al., 2009).

[4]We consider the revised optical properties evaluated in Draine (2003a,b,c).

RT run name	Hydro run name	Radiation field	AGN SED	f_d
*noAGN*008	noAGN	stars		0.08
*noAGN*03	noAGN	stars		0.3
*AGNsphere*008	AGNsphere	stars + BHs	fiducial	0.08
*AGNsphere*03	AGNsphere	stars + BHs	fiducial	0.3
*AGNcone*008	AGNcone	stars + BHs	fiducial	0.08
*AGNcone*03	AGNcone	stars + BHs	fiducial	0.3
*AGNcone*008*UVsteep*	AGNcone	stars + BHs	UV-steep	0.08
*AGNcone*03*UVsteep*	AGNcone	stars + BHs	UV-steep	0.3

Table 6.2: SKIRT post-processing runs performed. The first column labels the RT simulation, the second column indicates the corresponding hydrodynamical run, the third column specifies the radiation field included (e.g. stars with or without black holes), the fourth column specifies the AGN SED used (if black holes are present), and the fifth column contains the dust to metal ratio f_d adopted.

grain size distribution of graphite and silicates is sampled with 5 bins for each component. Gas particles hotter than 10^6 K, are considered dust-free as at these temperatures thermal sputtering is very effective at destroying dust (Draine & Salpeter, 1979; Tielens et al., 1994; Hirashita, 2015). This assumption does not affect the main results of our work, as discussed in Section A.2.

Dust self-absorption of thermally re-emitted photons in IR-optically thick regions is accounted for. We also include non-local thermal equilibrium (NLTE) corrections to dust emission. We do not include heating from CMB radiation. As discussed in Section 6.3.2, only a small fraction of dust grains is at a temperature comparable to $T_{\rm CMB}$. We expect this effect to be negligible, as seen a posteriori from the RT results.

The ultraviolet (UV) radiation field mainly responsible for dust heating is provided by stellar sources and black holes. Stellar radiation is implemented with the stellar synthesis models by Bruzual & Charlot (2003). Black holes are treated as point source emitters as the typical sizes of the accretion disk and the dusty torus are much smaller ($\lesssim 10$ pc) than the width of the most refined grid cells (≈ 230 pc, see above). We model their emission by adopting a composite power-law SED as described in Section 5.4, making use in this work of both the fiducial and UV-steep SEDs. The radiation field is sampled using a grid covering the *rest-frame* wavelength range $[0.1 - 10^3]$ μm, composed of 200 logarithmically spaced bins. A total of 10^6 photon packets per wavelength bin is launched from each source, i.e. stellar particles and BHs[5]. We collect the radiation escaping our computational domain for the 6 lines-of-sight perpendicular to the faces of the cubic computational domain.

We perform RT calculations on the three hydrodynamic simulations presented in section 6.2.1. For each hydro-simulation we vary the dust to metal ratio from $f_d = 0.08$ to $f_d = 0.3$; for the *AGNcone* run we consider both the AGN SEDs described in Section 5.4. We end up with a total of 8 post-processed runs, as reported in Table 6.2.

[5]We verified that the number of packets used is sufficient to achieve numerical convergence by comparing the results with control simulations with 5×10^5 photon packets per wavelength bin.

6.3 Results

In this section we present the results obtained through our RT calculations. We first present in Section 6.3.1 the morphology of the ultraviolet (UV, $1000 - 3000$ Å) and total infrared (TIR, $8 - 1000$ μm) emission and discuss how it is affected by the presence of the AGN and total dust content. Then, in Section 6.3.2 we derive the dust temperature in the different runs. Finally, we discuss in Section 6.3.3 the synthetic SEDs resulting from our calculations.

6.3.1 Overview

In Fig. 6.2, we show the UV (top row) and TIR (middle row) emission maps derived for the runs *noAGN* (left column), *AGNsphere* (middle column), *AGNcone* (right column) for $f_d = 0.08$. In Fig. 6.3 we show the same maps but for $f_d = 0.3$. We use the same line of sight as in Fig. 6.1.

By comparing the TIR maps with the dust surface density (Fig. 6.1, bottom row) we see that the morphology of the TIR emission matches the dust distribution, as expected. Moreover, the brightest TIR spots in the *noAGN* (AGN runs) correspond to the locations of the most highly star forming regions (accreting BHs), responsible for the dust grains heating. We discuss in more details the dust temperature in Section 6.3.2.

For what concerns UV emission, in the *noAGN* case, its distribution correlates with the star formation surface density (see middle row in Fig.6.1); in the AGN runs, the brightest spots are located in correspondence of the AGN positions, identified by white circles. Noticeably, whereas in the *AGNcone* simulation with $f_d = 0.08$ three peaks appear in the UV emission map (labelled as A, B, and C in Fig. 6.2), corresponding to the AGN positions[6], in the case $f_d = 0.3$ only one of them survives to the strong dust obscuration. In Section 6.4.1 we investigate in further details the contribution to the total SED of the different components traced by the UV and TIR maps.

Table 6.3 reports the UV and TIR luminosities before and after the dust-reprocessing of the radiation. We find that, in the *noAGN* run $84 - 94\%$ of the total UV emission is extincted by dust if $f_d = 0.08 - 0.3$. For comparison, in the *AGNcone* and *AGNsphere* runs, the same fraction is $77 - 99\%$ and $54 - 95\%$, respectively. Overall we find that in our simulated dusty galaxies ($M_{\rm d} \gtrsim 3 \times 10^7 M_\odot$) a large fraction ($\gtrsim 50\%$) of UV emission is obscured by dust, with some lines of sight characterised by 1% of UV transmission.

The range reported for the UV reprocessed luminosity in Tab. 6.3 refers to the variation occurring along different lines of sights: the minimum and maximum values differ by a factor that can be as high as ~ 6 in the AGN runs. We expect UV luminosity variations along different lines of sight even larger than the ones we find, if UV radiation would intersect dense, compact, dusty, molecular clouds, whose sizes ($\lesssim 100$ pc) and complex internal structure ($\sim 1 - 10$ pc, Padoan & Nordlund, 2011; Padoan et al., 2014; Vallini et al., 2017) are not resolved by of our simulations.

According to the Unified Model (Urry & Padovani, 1995), the classification between TypeI (unobscured) and TypeII (obscured) AGN is based on the presence of a dusty, donut-like shaped structure that is responsible for anisotropic obscuration in the circum-nuclear region (< 10 pc). Our results show that large UV luminosity variations with viewing angle, in addition to the ones due to the torus, arise from the inhomogeneous distribution of dusty gas surrounding the accreting BH, on ISM scales ($\gtrsim 200$ pc; see also Gilli et al., 2014).

[6]The accretion rate quoted in Table 4.3 for the case *AGNcone* is in fact the sum of the accretion rates of the most active black holes in the simulations ($\dot{M}_{\rm BH} \approx 32$, 7 and 50 $\rm M_\odot\,yr^{-1}$, for the sources A, B, and C, respectively).

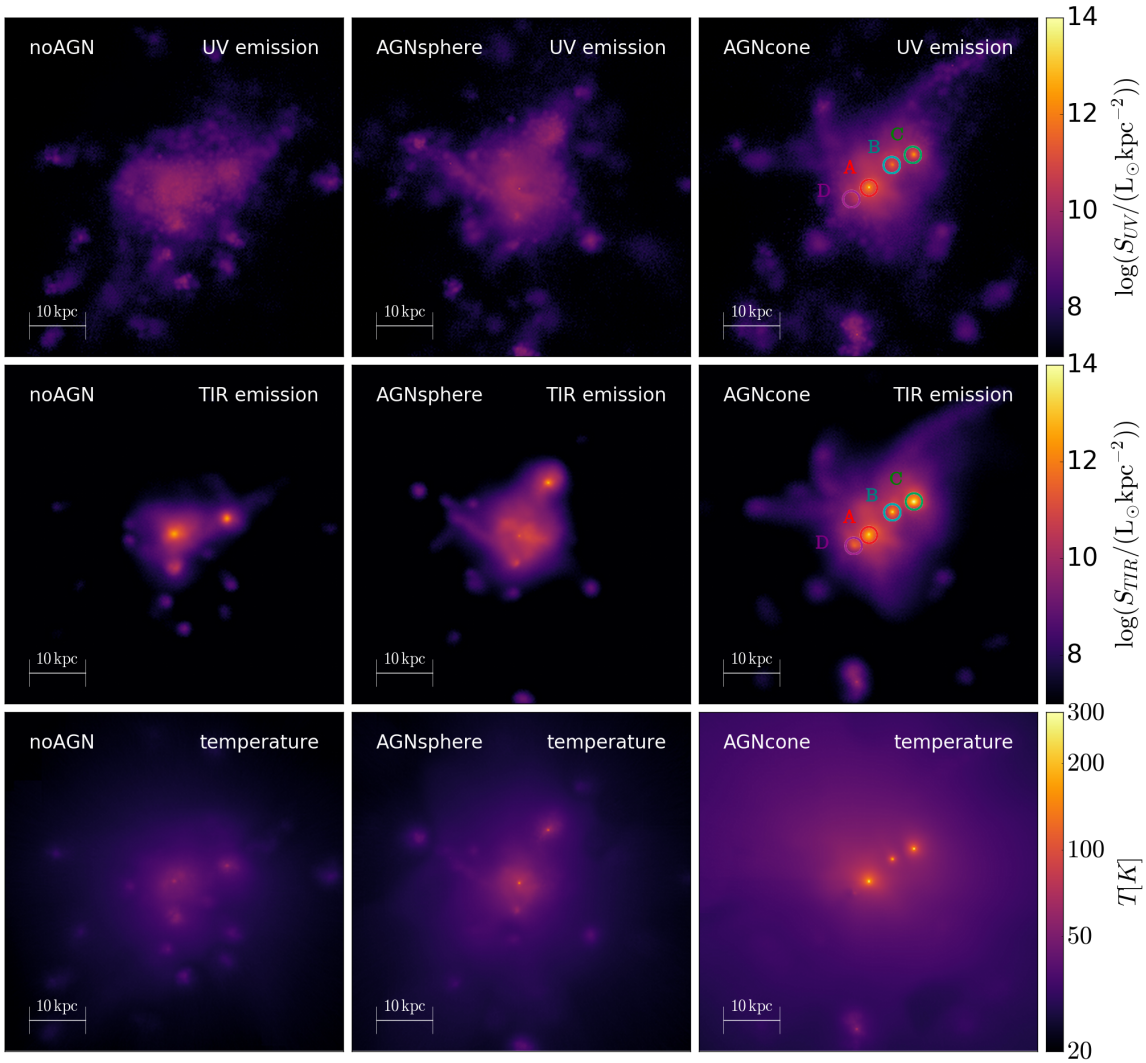

Figure 6.2: UV (top row), TIR (middle row), luminosity-weighted dust grain temperature (bottom row) maps for the runs with $f_d = 0.08$. The maps shown are produced with the same line of sight used in Fig. 6.1. We mark the four most luminous sources in TIR for the *AGNcone* runs, which will be discussed in more details in Section 6.4.1.

6.3.2 Dust temperature

One of the key physical quantities derived from RT calculations is the mass-weighted dust temperature ($\langle T_\mathrm{d} \rangle_M$). In what follows, we first describe how we compute the luminosity-weighted dust temperature ($\langle T_\mathrm{d} \rangle_L$, see Behrens et al., 2018; Sommovigo et al., 2020) and compare this value with $\langle T_\mathrm{d} \rangle_M$; then we discuss how the dust temperature is affected by the total amount of dust, and different types of UV sources (stars vs AGN).

Luminosity- vs. mass-weighted T_d

To compute $\langle T_\mathrm{d} \rangle_L$, we assume that each dust cell emits as a grey body[7] $L_{TIR} \propto M_\mathrm{d} T_\mathrm{d}^{4+\beta_\mathrm{d}}$, where β_d is the dust emissivity index[8]. $\langle T_\mathrm{d} \rangle_L$, finally depends on the total amount of dust M_d in the

[7]This approximation holds only for dust cells that are optically thin to IR radiation, although we caveat that a small number of cells in the simulation is actually optically thick.

[8]The actual value of β_d depends on the RT calculation. For example, Behrens et al. (2018) found a value of $\beta_\mathrm{d} = 1.7$. For computing the luminosity-weighted temperature, we assume $\beta_\mathrm{d} = 2$. This choice does not

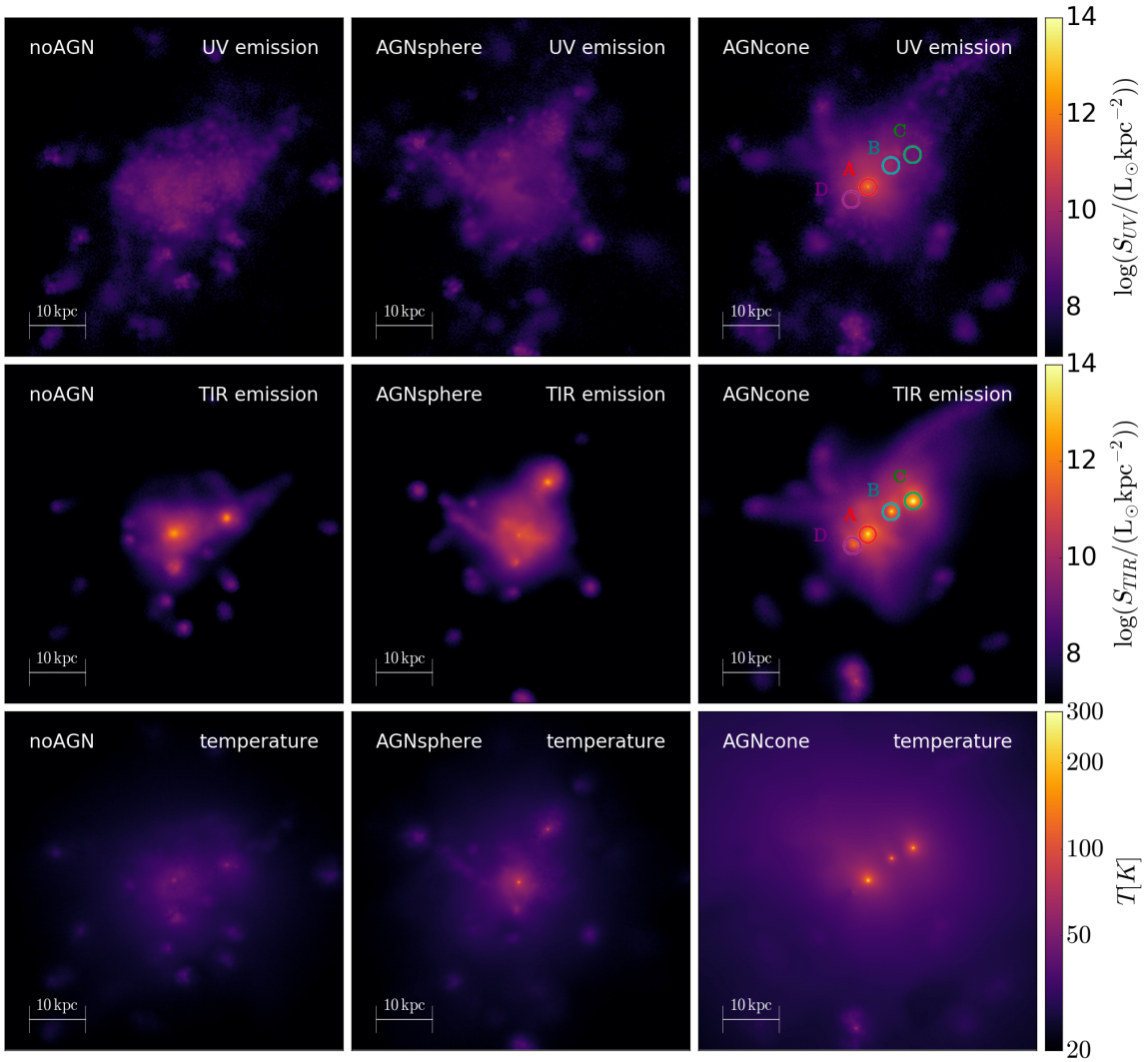

Figure 6.3: Same as in figure 6.2 but for for $f_d = 0.3$.

simulation, determined by our choice of f_d. Runs with $f_d = 0.3$ are characterised by average dust temperatures $\sim 10\%$ lower with respect to the corresponding runs with $f_d = 0.08$. This is because the same UV energy is distributed over a larger amount of dust mass.

In Fig. 6.4 we show the $\langle T_\mathrm{d} \rangle_L$ PDF (blue histograms), compared with the mass-weighted $\langle T_\mathrm{d} \rangle_M$ one (red histograms) for the *noAGN*, *AGNsphere* and *AGNcone* simulations with $f_d = 0.08$, as a reference case. In each run, the PDF of $\langle T_\mathrm{d} \rangle_M$ peaks at lower dust temperatures with respect to $\langle T_\mathrm{d} \rangle_L$. The difference between the mass-weighted and luminosity-weighted temperatures is particularly evident in the runs in which AGN radiation is included. In particular, the spikes of the luminosity-weighted histograms correspond to dust cells in the immediate proximity of accreting BHs. This dust component constitutes only a small fraction of the total mass, but it provides a significant contribution to the overall luminosity, as further discussed in the next section.

significantly affect the final results: the estimate of $\langle T_\mathrm{d} \rangle_L$ varying $1.5 < \beta_\mathrm{d} < 2.5$ is within 10% of the value reported in Table 5.

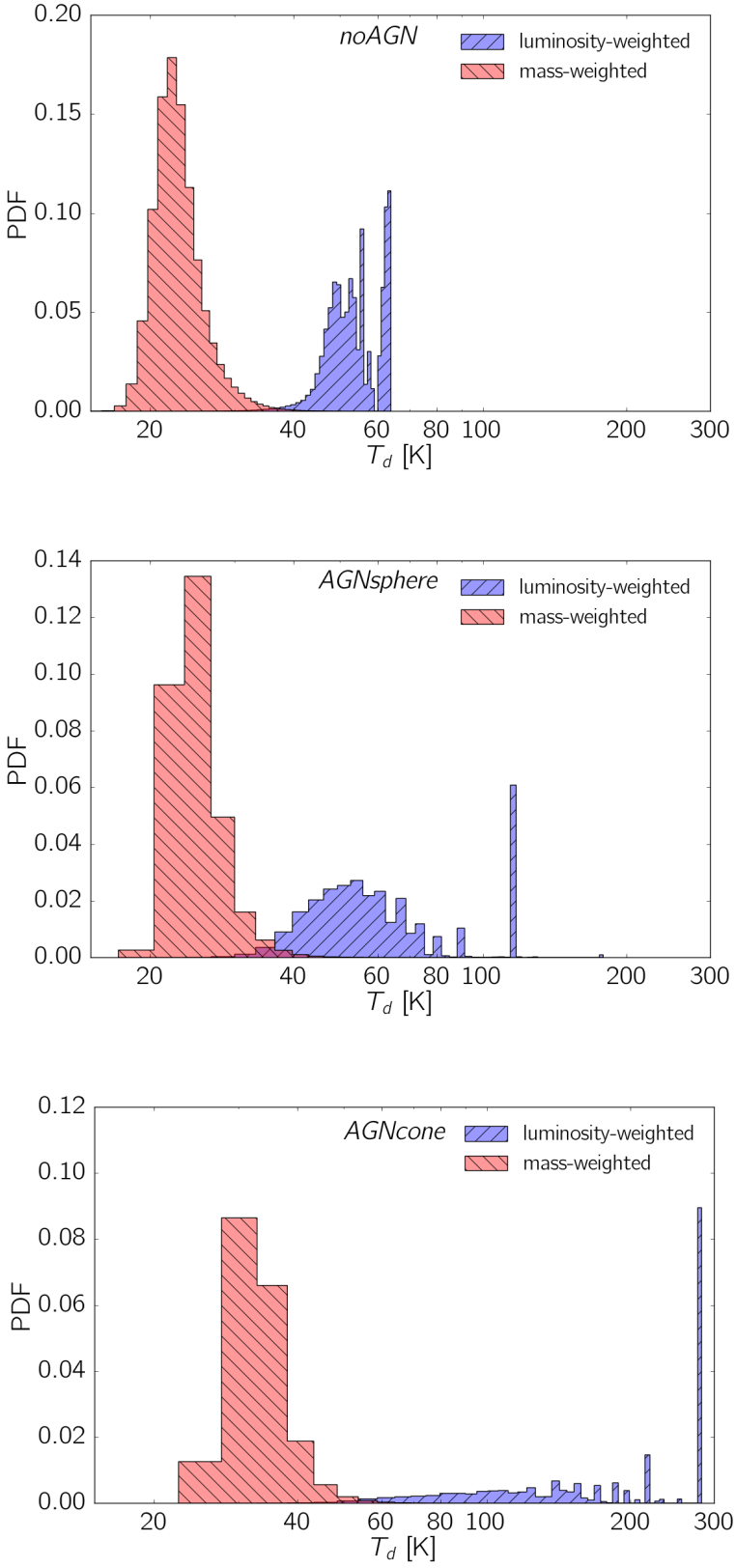

Figure 6.4: Mass-weighted (red histograms) and luminosity-weighted (blue histograms) dust grains temperature PDF. The panels refer to: (top) *noAGN*, (middle) *AGNsphere* and (bottom) *AGNcone*. As a reference case, we show the results for $f_d = 0.08$.

RT run	$L_{\rm UV}$ [10^{11} L$_\odot$]	$L_{\rm TIR}$ [10^{12} L$_\odot$]	M$_{\rm d}$ [10^7 M$_\odot$]	$\langle T_{\rm d}\rangle_L$ [K]	$T_{\rm d}^{\rm min/max}$ [K]	$L_{\rm UV}^{intr}$ [10^{12} L$_\odot$]	$\tau_{\rm UV}$
*noAGN*008	$5.7-6.7$	$4.4-4.6$	9.2	54 ± 6	$15-64$	4.1	$1.81-1.97$
*noAGN*03	$3.2-3.9$	$4.7-4.9$	34	48 ± 6	$13-57$	4.1	$2.35-2.56$
*AGNsphere*008	$7.0-17$	$3.3-3.4$	5.1	70 ± 27	$17-179$	3.7	$0.78-1.65$
*AGNsphere*03	$2.8-7.9$	$4.4-4.6$	19	62 ± 25	$15-178$	3.7	$1.54-2.57$
*AGNcone*008	$27-90$	$43-50$	3.3	208 ± 78	$22-282$	39	$1.47-2.67$
*AGNcone*03	$5.8-32$	$54-71$	13	182 ± 69	$20-272$	39	$2.50-4.20$

Table 6.3: Overview of the main physical properties of the galaxies for the RT runs performed (see Table 6.1). The table contains: (first column) the name of the run, (second column) the processed UV (integrated in the band $1000-3000$ Å) luminosity $L_{\rm UV}$, (third column) the processed total infrared (integrated in the band $8-1000$ μm) luminosity $L_{\rm TIR}$, (fourth column) the total dust mass contained in the simulated region $M_{\rm d}$, (fifth column) the *luminosity-weighted* temperature of the dust grains $\langle T_{\rm d}\rangle_L$, reported as the mean of the PDF within one standard deviation, (sixth column) the minimum and maximum value the dust grains temperature, (seventh column) the intrinsic (i.e, not dust-processed) UV luminosity $L_{\rm UV}^{\rm intr}$, (eighth column) the effective UV optical depth $\tau_{\rm UV}$, estimated as as $e^{-\tau_{\rm UV}} = L_{\rm UV}/L_{\rm UV}^{\rm intr}$. For the dust-processed UV, TIR luminosities and UV optical depth we report the range bracketed by the six line of sights considered for each simulation.

Stars and AGN contribution to dust heating

The brightest TIR spots in Fig. 6.2 and 6.3 in the *noAGN* (AGN runs) correspond to the locations of the most highly star forming regions (accreting BHs). Whereas in the *noAGN* run the maximum $T_{\rm d}$ value is about 60 K, in the AGN runs, dust grains reach luminosity-weighted temperatures $T_{\rm d}\gtrsim 200$ K close to BHs, and $T_{\rm d}\approx 60$ K in the diffuse gas.

We underline that in the *noAGN* run $\langle T_{\rm d}\rangle_L$ is up to 4 times lower with respect to the AGN runs despite having a star formation rate 3 times higher. These results indicate the dominant role played by AGN radiation in the dust heating. This is particularly evident if we compare in more details the run *noAGN* and *AGNsphere*. In the *noAGN* case, $L_{\rm UV}^{\rm intr} = L_{\rm UV,stars} = 4.1\times10^{12}$ L$_\odot$; in the *AGNsphere* case, $L_{\rm UV}^{\rm intr} = L_{\rm UV,stars} + L_{\rm UV,BH} = (2.3+1.4)\times10^{12}L_\odot = 3.7\times10^{12}$ L$_\odot$.

Thus, although the UV budget in the *AGNsphere* run is mostly provided by stars, and the total UV intrinsic luminosity is comparable to the *noAGN* case, $T_{\rm d}$ peaks at higher temperature values if BH accretion is present. In Fig. 6.5, we compare the fraction of mass (left panel) and TIR luminosity[9] (right panel) from dust with a temperature above a certain threshold for the three runs. In the *noAGN* run the TIR luminosity is arising from dust with $T_{\rm d}\lesssim 50$ K. In the *AGNsphere* (*AGNcone*) run $>50\%$ of the TIR luminosity is arising from dust with $T_{\rm d}\gtrsim 70$ K ($T_{\rm d}\gtrsim 150$ K); this warm dust only constitutes 0.1% of the total dust mass. This confirms that a small mass fraction of warm dust dominates the IR emission, as expected from the scaling $L_{\rm d}\propto M_{\rm d}T_{\rm d}^{4+\beta_{\rm d}}$.

[9]The luminosity is computed assuming $\beta_{\rm d}=2$ for consistency with the temperature PDF. The resulting luminosity varying $1.5<\beta_{\rm d}<2.5$ differs by $\approx 10\%$ from the quoted values.

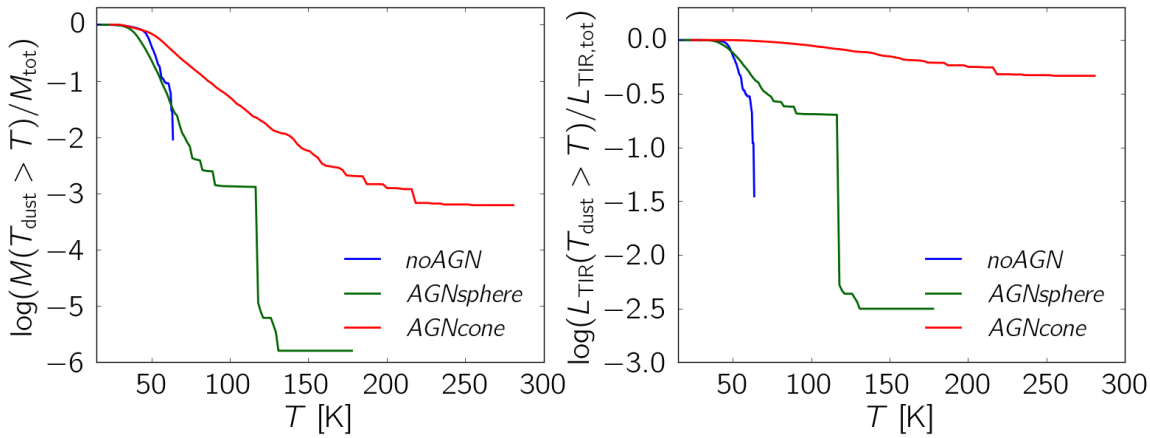

Figure 6.5: Mass fraction (left) and TIR luminosity fraction (right) of dust with a temperature $T_{\rm d} > T$ as a function of the temperature T for the runs *noAGN* (blue line), *AGNsphere* (green line), *AGNcone* (red line). Results for $f_d = 0.08$ are shown.

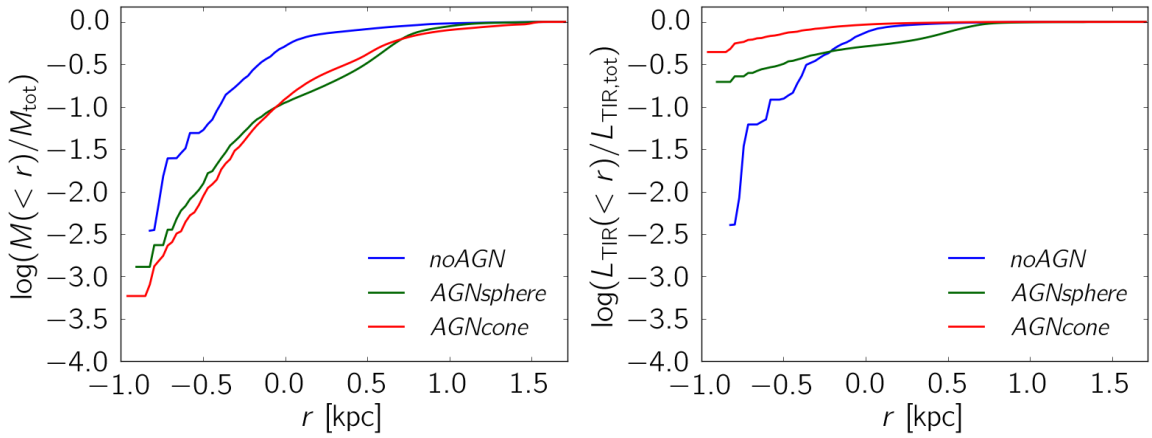

Figure 6.6: Cumulative mass fraction (left) and TIR luminosity fraction (right) of the dust at a distance r from AGN location or most star forming regions. The lines show the results for *noAGN* (blue line), *AGNsphere* (green line) and *AGNcone* (red line) with $f_d = 0.08$.

Spatial extent of FIR emitting regions

Fig. 6.6 shows the fraction of dust mass and infrared luminosity as a function of the distance[10] from the regions with the highest star formation for the *noAGN* case and from the BHs with the highest accretion rate for the run *AGNsphere* and *AGNcone*.

In the *noAGN* case, the dust mass within $r \lesssim 300$ pc represents $\sim 0.3\%$ of the total dust, and it provides $\sim 3\%$ of the total IR luminosity. In the *AGNsphere* (*AGNcone*) case, only $\sim 0.1\%$ (~ 0.06 %) of the total dust mass is found at $r \lesssim 300$ pc from an accreting BH but it contributes 20% ($\sim 40\%$) of the total IR luminosity.

[10]Given that there are multiple accreting BHs, we selected the 2 (3) most active ones in the *AGNsphere AGNcone* run and 2 most accreting star forming regions (the main galaxy and its largest satellite) in the *noAGN* run. For each cell containing dust in the octree grid we evaluate the distance from each reference source and then we consider the minimum one for this calculation.

6.3.3 Synthetic Spectral Energy Distributions

Fig. 6.7 shows the intrinsic flux density from stars (dashed line) and AGN (dotted line) for the first six runs reported in Table 6.2. The higher value of the flux density from stars in the *noAGN* run with respect to both AGN runs is due to the negative AGN feedback that in the AGN simulations quenches the star formation rate in the host galaxy (see Section 3.7 of B18 for an extensive discussion on this topic). This effect is more pronounced in the *AGNcone* run since it is characterised by a black hole accretion rate that is a factor of ~ 30 higher than in *AGNsphere* (see Table 6.1). The total intrinsic flux (dotted-dashed line) is comparable between *noAGN* and *AGNsphere* (see also Table 6.3).

We now analyse the differences between the reprocessed flux density (observed, solid line) resulting from our calculations, focusing on the rest-frame NIR ($1 \lesssim \lambda_{\rm RF} \lesssim 5$ μm), MIR ($5 \lesssim \lambda_{\rm RF} \lesssim 40$ μm), and FIR ($40 \lesssim \lambda_{\rm RF} \lesssim 350$ μm) wavelength ranges.

The intrinsic NIR flux is suppressed by ≈ 10 times in all runs; the highest rest-frame UV attenuation is seen in the *AGNcone* run, with some (all) lines of sight showing a flux reduced by ≈ 100 times for $f_d = 0.08$ ($f_d = 0.3$). However, for a fixed dust content, the *AGNcone* run still provides the highest rest-frame UV flux. In this wavelength range, the SED is nearly constant in the *noAGN* run whereas it increases toward larger wavelengths in the runs with AGN, as a consequence of the contribution from accretion. The observed optical-NIR flux depends both on the radiation field and dust content.

For what concerns the MIR, at short wavelengths, ($\lambda_{\rm RF} \sim 4 - 6$ μm), the SED is dominated by the almost unattenuated emission from stars and/or AGN; the *AGNcone* SED is ~ 30 times brighter than the *AGNsphere* one as a consequence of its higher BH activity. At longer wavelengths ($\lambda_{\rm RF} > 6$ μm), the observed flux arises from heated dust IR emission. The flux density in this wavelength range is the result of the sum of multiple greybodies, each emitting at different temperatures, according to the luminosity-weighted dust temperature PDF discussed in Section 6.3.2. The warm dust in AGN runs produces a MIR excess with respect to the *noAGN* run, and shifts the peak of the emission toward shorter wavelengths: $\lambda^{\rm peak}_{noAGN} = 59.4$ μm, $\lambda^{\rm peak}_{AGNsphere} = 54.1$ μm and $\lambda^{\rm peak}_{AGNcone} = 27.0$ μm.

Finally, the Rayleigh-Jeans tail of the FIR emission is mostly sensitive to the total dust content. In fact, by comparing the $f_d = 0.08$ and $f_d = 0.3$ cases, we find that the flux at 1 mm scales almost linearly with the dust mass, without a strong dependence on the radiation source.

To summarise, the SED in the NIR wavelength range depends both on the dust mass (for fixed dust properties) and the type of source (stars and/or AGN); the MIR retains information almost solely on the type of source: the presence of an AGN enhances the flux and shifts the peak of the emission at shorter wavelengths; the flux in the Rayleigh-Jeans tail of the FIR emission mostly depends on the total dust content.

6.4 Comparison with $z \sim 6$ quasar data

To test the results of our model (SPH simulation post-processed with RT calculations), we compare in Fig. 6.8 our predictions from the *AGNcone* run ($M_{\rm UV} = -27.97$) with multi-wavelength (NIR to FIR) observations of $z \sim 6$ bright ($-29 \lesssim M_{\rm UV} \lesssim -26$) quasars (see Table A.1 in Appendix A.5).

In the NIR, our predicted SEDs are underluminous with respect to the flux of TNG/GEMINI spectra (grey lines in Fig. 6.8). This mismatch cannot be solved by decreasing the dust content, since by assuming $f_d < 0.08$ the synthetic SEDs would become underluminous in the FIR with respect to ALMA data. We instead suggest that a better agreement with observations can be

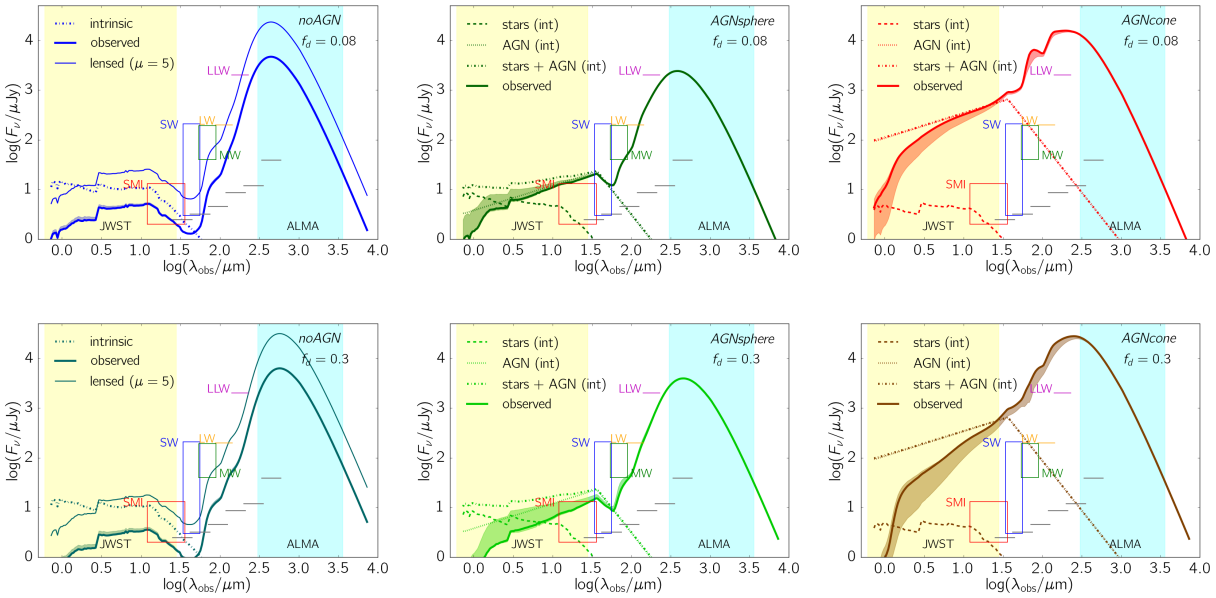

Figure 6.7: Intrinsic and processed (observed) SED for the first six runs in Table 6.2. The first column refers to *noAGN*, the second to *AGNsphere* and the third to *AGNcone*, whereas the first row to $f_d = 0.08$, and the second to $f_d = 0.3$. The solid line shows the observed flux for the reference line of sight, whereas the shaded area brackets the scatter in the observed SED between the six lines of sights used for the computation. The intrinsic flux is also shown with a a dot-dashed line. In the runs with AGN, the individual components are also shown: radiation from stars is denoted with dashed lines, radiation from AGN with dotted lines. In *noAGN* runs, thin solid lines indicate the flux that would be observed if the galaxy is magnified by a factor $\mu = 5$. Sensitivity bands of JWST and ALMA are shown as yellow and cyan shaded regions respectively. The grey lines indicate the sensitivity reached by the ORIGINS telescope at 5σ in 1 hr of observing time. The colored rectangles and horizontal lines indicate the sensitivity of the two instruments of the SPICA telescope: SMI ($\lambda_{\rm obs} = 27\ \mu$m, red rectangle), and SAFARI, in photometric mapping mode at short (SW, $\lambda_{\rm obs} = 45\ \mu$m, blue rectangle), mid (MW, $\lambda_{\rm obs} = 72\ \mu$m, green rectangle), long (LW, $\lambda_{\rm obs} = 115\ \mu$m, orange line) and very long wavelengths (LLW, $\lambda_{\rm obs} = 185\ \mu$m, violet line). The upper sides of rectangles represent the sensitivity that will be reached by SPICA at 5σ in 1 hr of observing time. The bottom side of rectangles represents the maximum sensitivity reachable with SPICA, and it is obtained by considering the confusion limit flux at 3σ (such a high sensitivity can be reached in the case of follow-up observations). If the confusion limit is reached in less than 1 hr, it is shown as a single line.

obtained by assuming an extinction curve flatter than the SMC (Gallerani et al. 2010; Di Mascia et al. 2021b, see also Chapter 8).

For what concerns the comparison in the MIR, models with $\alpha_{\rm UV}^{\rm fid}$ are in good agreement both with Spitzer/Herschel photometric data and with the slope/shape of templates by Hernán-Caballero & Hatziminaoglou (2011) resembling Spitzer/IRS spectra. Vice-versa, models with $\alpha_{\rm UV}^{\rm steep}$ are both under-luminous with respect to Spitzer/IRAC observations at $\lambda_{\rm obs} = 24\ \mu$m (namely $\lambda_{\rm RF} \sim 3\ \mu$m at $z \sim 6$) and show a slope in the MIR that does not agree with observed spectra.

We underline that the model with $\alpha_{\rm UV}^{\rm steep}$ can be possibly reconciled with observations if the torus is included. In fact, a dust component with temperature close to sublimation (~ 1500 K) would enhance the MIR emission exactly at the Spitzer/IRAC wavelengths[11]. We give a first estimate of the impact of the torus emission on our predicted SEDs in Section 6.5.1 (Fig. 6.10) and we defer the inclusion of the torus into our model to a future study.

By comparing our predicted SEDs with FIR observations, we note that both models ($f_d = 0.08 - 0.3$) provide a reasonable match with FIR data, independently on the assumed UV slope (fiducial vs UV-steep). We find that the models with a larger dust-to-metal ratio $f_d = 0.3$

[11] The emission of a greybody at temperature $T_{\rm d}$ and with $\beta_{\rm d} = 2$ peaks at $\lambda_{peak} = (2.9 \times 10^3)/T_{\rm d}\ \mu$m

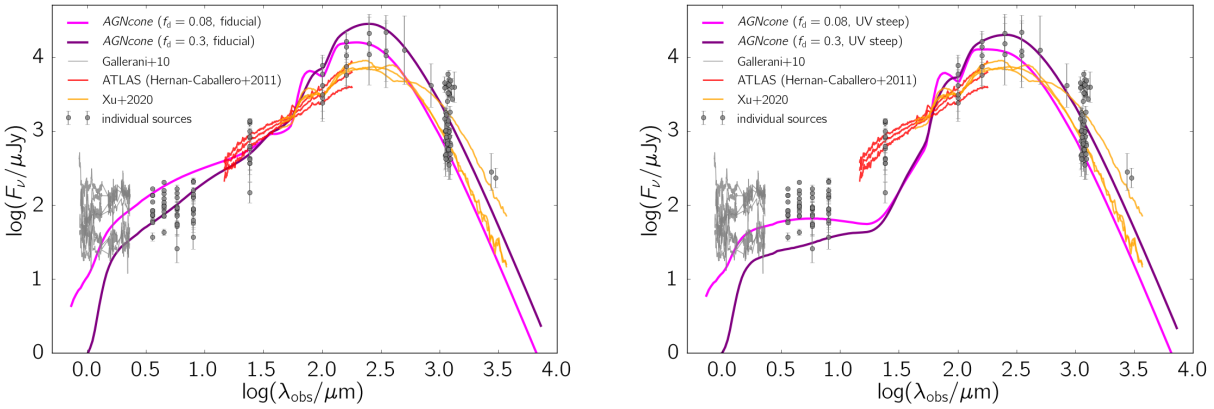

Figure 6.8: *Left panel*: Comparison between synthetic fiducial SEDs and observations of $z \sim 6$ quasars (grey circles, see Table A.1 in Appendix A.5). The magenta and violet solid curves denote the models for $f_d = 0.08$ and $f_d = 0.3$, respectively. Grey lines represent dust-reddened $z \gtrsim 6$ quasar spectra taken with the TNG/GEMINI ($A_{3000} > 0.8$, see Table 1 in Gallerani et al., 2010). The spectra are calibrated by using the measures of λL_λ at 1450 Å provided in Table 1 by Juarez et al. (2009). We further show with orange lines the template obtained from the analysis of 125 quasar spectra at $0.020 < z < 3.355$ taken with Spitzer/IRS (the tree lines correspond to the median SED, the 25^{th} and 75^{th} percentiles, in the case of the luminous, $\log(\lambda L_{5100}) > 44.55$, sub-sample by Hernán-Caballero & Hatziminaoglou (2011)), and with red lines the IR AGN SED derived by Xu et al. (2020) from 42 quasars at $z < 0.5$ (see their Table 3). The spectra of these low redshift sources are reported to get a hint of the spectral slope of AGN in the MIR; no spectroscopy information is available so far in the case of $z \sim 6$ quasars. *Right panel*: same as left panel, but for the UV-steep AGN SED models.

are slightly preferred, since in the $f_d = 0.08$ case we can only explain the less luminous FIR sources.

Hereafter, we consider as fiducial the model with $\alpha_{\mathrm{UV}}^{\mathrm{fid}}$ and $f_d = 0.3$.

6.4.1 Multiple merging system

The most massive halo in the *AGNcone* run at $z = 6.3$ hosts a merging system of multiple sources, three of which are AGN (A, B, C) and one is a normal star forming galaxy (D). We show in Fig. 6.9 the SEDs extracted from individual sources. In our simulated system, source A is the most luminous UV source, providing $\sim 70\%$ of the total UV flux. However, it does not correspond to the most accreting BH, which is instead powering source C, distant ~ 10 kpc from A. Despite having the highest intrinsic UV budget, this source is fainter than A in the UV because it is enshrouded by dust: source C is in fact the most luminous IR source of the system and provides $\sim 70\% - 80\%$ ($\sim 40\%$) of the MIR (FIR) flux. The second brightest UV source in our system is source D.

By comparing our synthetic SEDs with HST and ALMA data[12] (Marshall et al., 2020; Decarli et al., 2017), we found that sources A, B and D would be detectable and resolved with HST; for what concerns the FIR, given the angular resolution of current ALMA data (i.e. 1" that corresponds to ~ 6 kpc at $z = 6.3$), it is not possible to disentangle source B and D from A, whereas source C would show up as an SMG companion, even brighter than source A (as in the case of CFHQ J2100-1715 by Decarli et al., 2017). To summarise, our study shows that, consistently with HST and ALMA observations, bright ($M_{\mathrm{UV}} \leq -26$) $z \sim 6$ quasars (e.g. source A in our simulations) are part of complex, dust-rich merging systems, possibly containing highly ac-

[12] We do not consider constraints from MIR observations since individual sources cannot be resolved at these wavelengths as a consequence of the poor angular resolution. We further refer to Vito et al. in preparation for a detailed comparison with X-ray observations.

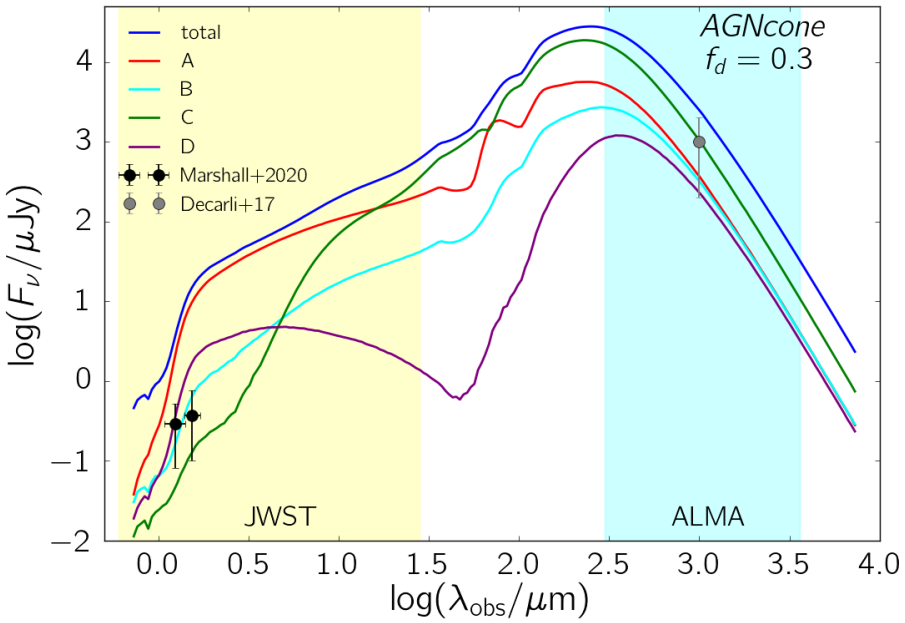

Figure 6.9: Comparison between the SEDs extracted from sources A, B, C, D in the field of view of the *AGNcone* run with $f_d = 0.3$, keeping the same color legend as in Fig. 6.3. The total SED is instead plotted with a blue solid line. Black points indicate rest-frame UV limits from deep HST observations by Marshall et al. (2020), whereas the grey point FIR fluxes for star-forming companion galaxies around quasars from Decarli et al. (2017).

creting BHs (e.g. source A, B and C with $\gtrsim 5\ \mathrm{M}_\odot\ \mathrm{yr}^{-1}$) and star forming galaxies (e.g. source D). Deeper and higher resolution ALMA data and JWST observations are required to better characterize the properties of galaxy companions in the field of view of $z \sim 6$ quasars.

6.5 Guiding future MID-IR facilities

Given the good agreement between our results and currently available $z \sim 6$ quasars observations, we can use our simulations to make predictions for the proposed Origins Space Telescope (OST[13]; Wiedner et al. 2020). OST covers the wavelength range $2.8 - 588\ \mu\mathrm{m}$, and is designed to make broad-band imaging (Far-IR Imager Polarimeter, FIP), low resolution ($R \sim 300$) wide-area/deep spectroscopic surveys, and high resolution ($R \sim 40000 - 300000$) pointed observations (with the Origins Survey Spectrometer, OSS). We further consider the capability of detecting IR emission from $z \sim 6$ quasars through a 2.5 m diameter infrared telescope, cryocooled to 8 K that covers the wavelength range $12 - 230\ \mu\mathrm{m}$, and is designed to make high-resolution ($R \sim 28000$) in the near-infrared ($12 - 18\ \mu\mathrm{m}$) and mid-infrared ($30 - 37\ \mu\mathrm{m}$) broad band mapping, and small field spectroscopic and polarimetric imaging at 100, 200 and $350\ \mu\mathrm{m}$. These are the characteristics of the Space Infrared Telescope for Cosmology and Astrophysics (SPICA; e.g. Spinoglio et al. 2017; Gruppioni et al. 2017; Egami et al. 2018; Roelfsema et al. 2018), an infrared space mission, initially considered as a candidate for the M5 mission, but cancelled in October 2020 (Clements et al., 2020).

The *noAGN* case is detectable by ORIGINS in five bands, corresponding to $\approx 6 - 80\ \mu\mathrm{m}$ rest-frame. ORIGINS would be able to probe the SED of highly star forming galaxies ($\mathrm{SFR} \sim 600\ \mathrm{M}_\odot\ \mathrm{yr}^{-1}$) at wavelengths shorter than the peak wavelength, which is crucial in order to have a solid determination of the dust temperature (Behrens et al., 2018; Sommovigo et al.,

[13]OST is a concept study for a 5.9 m diameter infrared telescope, cryocooled to 4.5 K, that has been presented to the United States Decadal Survey in 2019 for a possible selection to NASA's large strategic science missions.

2020). The *noAGN* case falls just below the SPICA sensitivity threshold. We thus consider the possibility of observing lensed galaxies with SPICA; the thin solid SED in the *noAGN* panels in Fig. 6.7 accounts for a magnification factor $\mu \sim 5$. Our results show that highly star forming galaxies (SFR ~ 600 M$_\odot$ yr^{-1}) without an active AGN will be at the SPICA reach if lensed by a factor $\mu \gtrsim 5$.

For what concerns the *AGNsphere* case, the simulated run corresponds to a faint AGN ($M_{\rm UV} = -23.4$; X-ray luminosity $L_{\rm X} \sim 10^{44}$ erg s^{-1}). This kind of sources is not easily detectable through UV and X-ray observations: (i) less than 20 $z \sim 6$ quasars fainter than $M_{\rm UV} = -23.75$ have been discovered so far (Matsuoka et al., 2018b); (ii) none $z \sim 6$ quasar with $L_{\rm X} < 4 \times 10^{44}$ erg s^{-1} has been detected so far with Chandra (Vito et al., 2019b). Our predictions show that the SED of a faint AGN is instead well above ORIGINS' sensitivities at all wavelengths and also above the sensitivities of two SPICA's bands for all the simulations we performed. This result emphasises the important role that future MIR facilities would have in studying the faint-end of the UV and X-ray luminosity function in $z \sim 6$ AGN.

The *AGNcone* runs show that quasars with $M_{UV} < -25$ are very easily detectable both by ORIGINS and SPICA at a signal-to-noise ratio high enough to get good quality spectra even in these very distant sources. We notice that only ~ 20 quasars have been detected so far with the Spitzer / Herschel telescopes at $z \gtrsim 6$ (Leipski et al., 2014; Lyu et al., 2016), and most of them ($> 80\%$) are bright ($M_{\rm UV} < -26$). Quasars fainter than $M_{UV} = -26$ have been detected so far at mm wavelengths at $> 5\sigma$ only in two $z \geq 6$ quasars (J1048-0109 and P167-13 by Venemans et al., 2018).

Our results show that the ORIGINS telescope will be an extremely powerful instrument for studying the properties of the most distant galaxies and quasars known so far.

6.5.1 Unveiling faint/obscured AGN

Combined ALMA data with follow-up JWST and /or[14] ORIGINS observations will be crucial to discover faint/obscured AGN and to distinguish them from galaxies without an active nuclei. In fact, by comparing the predicted fluxes in ORIGINS band 1 and/or MIRI band at 29 μm ($F_{29\mu\rm m}$) with the ones in ALMA band 7 $F_{\rm band7}$, we find:

$$\frac{F_{29\mu\rm m}(\textit{AGNsphere})/F_{\rm band7}(\textit{AGNsphere})}{F_{29\mu\rm m}(\textit{noAGN})/F_{\rm band7}(\textit{noAGN})} \approx 8 - 10,$$

meaning that we expect a a MIR-to-FIR excess of one order of magnitude in the case of a faint AGN host galaxy (*AGNsphere*) with respect to a star forming galaxy without AGN (*noAGN*). This result shows that by following up with JWST and/or ORIGINS sources already detected with ALMA it will be possible to discriminate between star forming galaxies and faint/obscured AGN.

We note that, given the limited resolution ($\sim$ 200 pc) of the hydrodynamical simulations adopted in this work, we cannot resolve the torus ($\sim$ 0.1 $-$ 10 pc) that is, therefore, not included in our modelling. The presence of a dusty torus surrounding accreting BHs provides an additional source of MIR emission boosting the MIR excess expected in AGN. This can increase both the detectability of faint quasars with a SPICA-like telescope and the possibility of exploiting the synergy between ALMA and MIR facilities to unveil dust-obscured AGN. For example, we qualitatively show in Fig. 6.10 how our predicted SEDs would change with the inclusion of the emission from the dusty torus. For this comparison we consider the *AGNsphere*

[14]The James Webb Space Telescope is planned to fly on October 31, 2021, with 10 years of operation goal. The proposed ORIGINS mission is planned for launch in the early 2030s, so it will ideally continue the work of JWST.

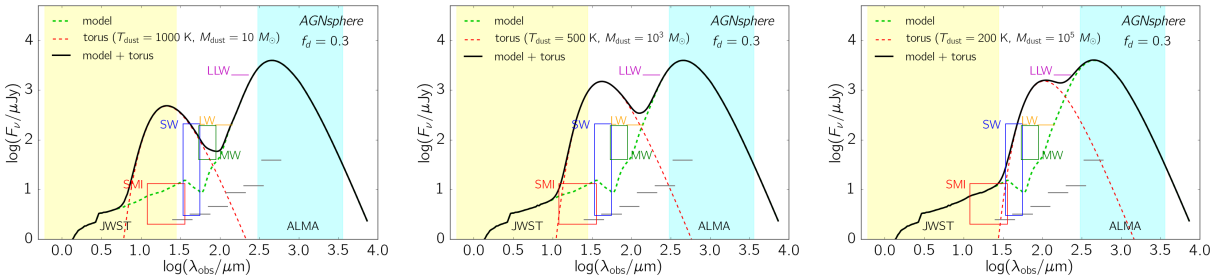

Figure 6.10: Predicted SEDs with the inclusion of the dusty torus emission to our models for the *AGNsphere* case with $f_d = 0.3$. The green dashed line refers to the original model, the red dashed line to the torus and the black solid line to the sum of the two. The torus emission is modelled as a greybody with $M_{\rm dust}$ and $T_{\rm dust}$ as specified in the panel, and $\beta_{\rm dust} = 2$.

case ($M_{\rm UV} = -23.4$), since we aim to investigate the ability of MIR telescopes to unveil faint AGN. As a proof of concept, we simply model the torus emission as a single-temperature $T_{\rm dust}$ greybody, with dust mass $M_{\rm dust}$, and $\beta_{\rm dust} = 2$. We consider different models to cover the range in masses ($10^1 - 10^5$ M$_\odot$) and temperatures ($200 - 1200$ K) constrained by theoretical models (Schartmann et al., 2005; Stalevski et al., 2016) and observations (García-Burillo et al., 2016, 2019).

The MIR emission from the torus brings the SEDs of the faint AGN of our model easily within the reach of a SPICA-like telescope, for a wide range of the torus parameters considered. This further expands the potential of future MIR telescopes in the discovery and the study of faint AGN at high-redshift.

We stress that this is a rough estimate that – among other things – neglects the torus geometry, i.e. the fact that UV emission is extinguished along the equatorial plane. Therefore UV photons would escape only towards the polar regions, reducing the amount of dust on $\lesssim 200$ pc scales directly irradiated by the AGN and possibly the IR emission coming from high temperature ($T_{\rm d} \sim 200 - 300$ K) regions. We plan to include the torus emission in a consistent way in our model in a future work and to further examine its impact on our results and on the potential of future facilities.

6.6 Summary and conclusions

In this Chapter, we have considered a suite of zoom-in cosmological hydrodynamic simulations of a massive halo ($\sim 10^{12}$ M$_\odot$) at $z \sim 6$ (Barai et al., 2018). The set of simulations include a control simulation of a highly star forming galaxy (SFR ~ 600 M$_\odot$ yr^{-1}) without BHs (called *noAGN* run), and two simulations with accreting BHs that account for AGN kinetic feedback distributed according to a spherical (*AGNsphere* run) and bi-conical (*AGNcone* run) geometry. These two different feedback prescriptions result in different SFRs of the host galaxy (~ 300 and ~ 200 M$_\odot$ yr^{-1} in *AGNsphere* and *AGNcone*, respectively), and AGN activity (~ 3 and ~ 90 M$_\odot$ yr^{-1} in *AGNsphere* and *AGNcone*, respectively).

We performed dusty radiative transfer calculations of the three runs in post-process by exploiting the code SKIRT (Baes et al., 2003; Baes & Camps, 2015; Camps & Baes, 2015; Camps et al., 2016) with the aim of understanding the impact of radiative feedback on the observed spectral energy distributions (SEDs) of $z \sim 6$ galaxies. We have considered (i) intrinsic AGN SEDs defined by a composite power-law $F_\lambda \propto \lambda^\alpha$ constrained through observational and theoretical arguments; (ii) SMC dust properties (grain size distribution and composition); (iii) different total dust mass content (parametrized in terms of the dust-to-metal ratios $f_{\rm d} = 0.08$ and $f_d = 0.3$),

and we have explored how different assumptions affect the observational properties of galaxies in the Epoch of Reionization (EoR). By analyzing the synthetic emission maps and SEDs resulting from our calculations we have found the following results:

- In dusty galaxies ($M_{\rm d} \gtrsim 3 \times 10^7$ M$_\odot$) a large fraction ($\gtrsim 50\%$) of UV emission is obscured by dust.
- Large UV luminosity variations with viewing angle, can be at least partially due to the inhomogeneous distribution of dusty gas on scales $\gtrsim 100$ pc.
- Simulations including AGN radiation show the presence of a clumpy, warm ($\approx 200 - 300$ K) dust component, in addition to a colder ($\approx 50 - 70$ K) and more diffuse dusty medium, heated by stars; warm dust provides up to 50% of the total infrared luminosity, though constituting only a small fraction ($\lesssim 0.1\%$) of the overall mass content.

We have tested our model by comparing the simulated SEDs with observations of $z\sim 6$ bright ($M_{UV} \lesssim -26$) quasars, the only class of sources for which multi-wavelength observations, ranging from the optical-NIR to the mm, are available so far. For what concerns the intrinsic SEDs, we have considered two variations for the rest-frame UV band: a fiducial value $\alpha_{\rm UV}^{\rm fid} = -1.7$ suggested by observations of unreddened quasars (Richards et al., 2003), and a steeper slope $\alpha_{\rm UV}^{\rm steep} = -2.3$ supported by observations of reddened quasars (Gallerani et al., 2010) and theoretical arguments (Shakura & Sunyaev, 1973). The main findings of this comparison are the following:

- We find a good agreement between simulations and both MIR (Spitzer/Herschel) and millimetric (ALMA) data, in the case of $\alpha_{\rm UV}^{\rm fid} = -1.7$. In the rest-frame UV, our predicted SEDs are underluminous with respect to data, suggesting peculiar extinction properties (Gallerani et al. 2010; Di Mascia et al. 2021b, see also Chapter 8).
- The case $\alpha_{\rm UV}^{\rm steep} = -2.3$ cannot explain the Spitzer/IRAC flux at $\lambda_{\rm obs} = 24$ μm and show a slope in the MIR that does not agree with Spitzer/IRS spectra. This discrepancy can be possibly alleviated by adding to our model the emission arising from a dusty torus with $T_{\rm d} \sim 1500$ K, close to sublimation temperature of graphite and silicate grains (Netzer, 2015).
- Quasars powered by SMBHs are part of complex, dust-rich merging systems, containing both multiple accreting BHs and star forming galaxies that, because of strong dust absorption, are below the detection limit of current deep optical-NIR observations (Mechtley et al., 2012), but appear as SMG companions, consistently with recent shallow ALMA data Decarli et al. (2017). Deeper ALMA and future JWST observations are required to study the environment in which $z \sim 6$ quasars form and evolve.

Given the good agreement between our results and rest-frame MIR observations, we exploit our simulations to make predictions for the proposed Origins Space Telescope (OST; Wiedner et al. 2020), a possible selection to NASA's large strategic science missions, and for a MIR telescope with the same technical specifications of the Space Infrared Telescope for Cosmology and Astrophysics (SPICA; e.g. Spinoglio et al. 2017; Gruppioni et al. 2017; Egami et al. 2018; Roelfsema et al. 2018), an infrared space mission, initially considered as a candidate for the M5 mission, but cancelled in October 2020 (Clements et al., 2020). We end up with the following conclusions:

- Highly star forming galaxies ($\mathrm{SFR} \sim 600\ \mathrm{M_\odot\,yr^{-1}}$) without an active AGN will be easily detected by ORIGINS. It will also be able to probe the peak of the dust emission, allowing a solid estimate of the dust temperature in star forming galaxies at high redshift. These galaxies would also be detected by a SPICA-like telescope, if lensed by a factor $\mu \gtrsim 5$.
- Bright high-z quasars ($M_{UV} < -26$) are detectable with ORIGINS/SPICA at a signal-to-noise ratio high enough to get high quality spectra even in these very distant sources.
- The FIR/MIR flux ratio in star forming galaxies is one order of magnitude higher with respect to AGN hosts, even in the case of low accretion rates ($\dot{M}_{\mathrm{BH}} \sim 3\ \mathrm{M_\odot\,yr^{-1}}$). By following up with ORIGINS/SPICA galaxies already detected with ALMA it will be possible to unveil faint and/or dust-obscured AGN, whose fraction is expected to be large ($> 85\%$) at high redshift (e.g. Vito et al. 2014, 2018; see also Davies et al. 2019). Our FIR/MIR estimate is quite conservative, because our model does not include the emission from the dusty torus, which is expected to boost the MIR flux by up to two order of magnitudes in ORIGINS/SPICA bands.

These results highlight the importance of a new generation of MIR telescopes to understand the properties of dusty galaxies and AGN at the EoR.

Impact of AGN on SFR estimates

7

7.1 Introduction

In the last two decades hundreds of quasars have been discovered within the first Gyr of the Universe (e.g. Yang et al., 2020b; Wang et al., 2021a), as discussed in Section 2.5. These objects are powered by accreting supermassive black holes (SMBHs, 10^{8-10} $M_\odot$, e.g Wu et al. 2015), shining as Active Galactic Nuclei (AGN). Observations in the local Universe reveal that black hole masses correlate with the host-galaxy properties (e.g. Magorrian et al., 1998b; Ferrarese & Merritt, 2000; Gebhardt et al., 2000; Marconi & Hunt, 2003; Gültekin et al., 2009), suggesting a co-evolution between the the galaxy and the central SMBH (e.g. Kormendy & Ho, 2013; Harrison, 2017), possibly mediated by a form of feedback from the AGN via energy/momentum injections onto the surrounding gas (e.g. Silk & Rees, 1998; King, 2003). In this picture, AGN activity might also regulate the star formation in the host galaxy (e.g. Carniani et al., 2016; Harrison et al., 2018). A connection between the two quantities is also suggested by the similar shape of the black hole accretion density and star formation rate density across cosmic time (e.g. Aird et al., 2015).

In the latest years, ALMA and NOEMA observations have provided the opportunity to study the properties of the interstellar medium (ISM) in several $z \sim 6$ bright quasar-host galaxies via their [C II]158 μm and CO emission, as well as the underlying dust continuum (e.g. Gallerani et al., 2017a; Decarli et al., 2018; Venemans et al., 2018; Carniani et al., 2019b; Venemans et al., 2020; Neeleman et al., 2021). These studies revealed the presence of large molecular gas reservoirs ($M_{\rm gas} \gtrsim 10^{10}$ $M_\odot$), dust masses ($M_{\rm dust} \approx 10^{7-8}$ $M_\odot$), and ongoing star formation rates (SFRs) as high as $\approx 10^{2-3}$ $M_\odot$ yr^{-1} (e.g. Decarli et al. 2018; Venemans et al. 2020) in the host galaxies.

These estimates rely on the assumption that the far-infrared (FIR) emission is mainly due to stellar light reprocessed by dust. The FIR photometry is generally fitted with a grey-body function (e.g. Carniani et al., 2019b), from which the dust mass $M_{\rm dust}$ is determined, whereas the dust temperature $T_{\rm dust}$ is assumed to be in the range $40 - 60$ K. The total infrared (TIR) luminosity, $L_{\rm TIR}$, is then computed by integrating the grey-body function in the wavelength range $8 - 1000$ μm, and a SFR $- L_{\rm TIR}$ calibration (e.g. Kennicutt & Evans, 2012) is used to infer the star formation rate in the host galaxy. This computation is very sensitive to the value of $T_{\rm dust}$, which is often set a priori because in most cases only one or a few FIR photometric data points are available (e.g. Venemans et al., 2020).

Recently, Walter et al. (2022) obtained high angular resolution measurements of the [C II] emission and underlying dust continuum in the $z = 6.9$ quasar J2348-3054. From the fit of the Spectral Energy Distribution (SED) they derived $T_{\rm dust} = 84.9^{+8.9}_{-10.5}$ K, and then inferred a SFR of 4700 $M_\odot$ yr^{-1}. This estimate is a factor of ≈ 10 higher than the one suggested by the [C II]-SFR

relation in Herrera-Camus et al. (2018). They also found $T_{\rm dust} \geq 134$ K in the innermost region (≈ 110 pc), implying an extreme star formation rate density of 10^4 M$_\odot$ yr^{-1} pc^{-2}, which is only marginally consistent with the Eddington limit for star formation (Thompson et al., 2005). The authors argued that this result might also be explained by a contribution from the central AGN to the dust heating in the innermost ≈ 100 pc.

Numerical simulations constitute a complementary tool to investigate the dust properties in the ISM of galaxies, thanks to the possibility to support them with radiative transfer calculations (e.g. Behrens et al., 2018). Recently, McKinney et al. (2021) simulated a post-merger dust-enshrouded AGN, representative of submillimiter galaxies (SMGs) at $z \sim 2-3$. They find that AGN can dominate the dust heating on kpc-scales, boosting the IR emission at $\lambda \gtrsim 100\ \mu$m (typically associated with dust-reprocessed stellar light) by up to a factor of 4. As a consequence, standard FIR-based calibrations would over-estimate the star formation rates by a comparable factor. This finding is also supported by empirical works suggesting that the AGN contribution to the IR emission increases as a function of AGN power in $z \lesssim 2.5$ sources, questioning the extreme star formation rates ($\gtrsim 1000$ M$_\odot$ yr^{-1}) inferred for the IR-brightest sources (Symeonidis & Page, 2021; Symeonidis et al., 2022).

In this Chapter, we test the goodness of the assumption that dust heating is dominated by stellar radiation in high-z quasar-hosts exploiting numerical simulations. In Chapter 6 we studied the AGN contribution to the dust heating in $z \sim 6$ galaxies by post-processing the cosmological hydrodynamical simulations by B18 with the radiative transfer code SKIRT (Baes & Camps, 2015; Camps et al., 2016). We found that in normal star-forming galaxies the dust temperature distribution is consistent with typical values adopted ($40-50$ K) in the SED fitting of $z > 5$ sources. However, in runs including AGN radiation the dust temperature increases to $T_{\rm dust} > 100$ K in the regions closest (≈ 200 pc) to the AGN, consistent with the finding by Walter et al. (2022). We expand the work presented in the previous Chapter by analyzing the synthetic SEDs predicted by our simulations to investigate whether the SFR inferred from the IR emission is overestimated when the AGN contribution is not accounted for.

7.2 Numerical Methods

7.2.1 Hydrodynamical simulations

In this work we adopt the suites of cosmological hydrodynamic simulations studied in B18 and in V21 (already described in details in Section 4.3 and 4.3 respectively). These simulations make use of a modified version of the Smooth Particle Hydrodynamics (SPH) N-body code GADGET-3 (Springel, 2005) to follow the evolution of a $\sim 10^{12}$ M$_\odot$ dark matter (DM) halo in a zoom-in fashion from $z = 100$ to $z = 6$. A flat ΛCDM cosmology is assumed, with: $\Omega_{\rm M,0} = 0.3089$, $\Omega_{\Lambda,0} = 0.6911$, $\Omega_{\rm B,0} = 0.0486$, $H_0 = 67.74$ km s^{-1} Mpc^{-1} (Planck Collaboration et al., 2016). The two suites differ for the particle resolution and subgrid physics implemented, as follows.

B18 simulations

In B18 a DM-halo of $M_{\rm h} = 4.4 \times 10^{12}$ M$_\odot$ at $z = 6$ (virial radius $R_{200} = 73$ pkpc[1]) is chosen inside a comoving volume of $(500\ {\rm cMpc})^3$ for re-simulation in a zoom-in region of $(5.21\ {\rm cMpc})^3$. The highest resolution DM and gas particles in the zoom-in region have a mass of

[1]Throughout this paper ckpc refers to *comoving* kpc and pkpc to *physical* distances in kpc. When not explicitly stated, we are referring to physical distances.

$m_{\rm DM} = 7.54 \times 10^6$ M$_\odot$ and $m_{\rm gas} = 1.41 \times 10^6$ M$_\odot$, respectively. Gravitational forces are softened on a scale of $\epsilon = 1\ h^{-1}$ ckpc, which corresponds to ≈ 210 pc at $z = 6$. The physics of the ISM is described by the multiphase model of Springel & Hernquist (2003), assuming a density threshold of $n_{\rm SF} = 0.13$ cm^{-3} for star formation and a Chabrier (2003) initial mass function (IMF) in the mass range $0.1 - 100$ M$_\odot$. The code accounts for radiative heating and cooling, for stellar winds, supernova feedback and metal enrichment. Stellar evolution and chemical enrichment are computed following Tornatore et al. (2007).

BHs are included in the simulation by placing a $M_{\rm BH} = 10^5\ h^{-1}$ M$_\odot$ BH seed at the centre of a $M_{\rm h} = 10^9\ h^{-1}$ M$_\odot$ DM halo, if it does not host a BH already. BHs can grow either by gas accretion or via mergers with other BHs. The former process is modelled via the Bondi-Hoyle-Littleton scheme (Hoyle & Lyttleton, 1939; Bondi & Hoyle, 1944; Bondi, 1952), and the accretion rate is capped at the Eddington rate $\dot{M}_{\rm Edd}$. A fraction of the accreted rest-mass energy is radiated away with a bolometric luminosity:

$$L_{\rm bol} = \epsilon_{\rm r} \dot{M}_{\rm BH} c^2, \tag{7.1}$$

where c is the speed of light and $\epsilon_{\rm r} = 0.1$ is the radiative efficiency. AGN feedback from the accreting BHs is modelled by distributing a fraction $\epsilon_{\rm f} = 0.05$ of the energy irradiated by the BHs onto the surrounding gas in kinetic form. In this work we consider the runs *AGNsphere* and *AGNcone* from B18, in which AGN feedback is distributed in a spherical symmetry and in a bi-cone with an half-opening angle of $45°$, respectively.

V21 simulations

In V21 a DM-halo of $M_{\rm h} = 1.12 \times 10^{12}$ M$_\odot$ is chosen for the zoom-in simulation inside a comoving volume of $(148\ {\rm cMpc})^3$. A zoom-in region of size $(5.25\ {\rm cMpc})^3$ is chosen for re-simulation, with the highest resolution particles of the zoom-in simulation having a mass of $m_{\rm DM} = 1.55 \times 10^6$ M$_\odot$ and $m_{\rm gas} = 2.89 \times 10^5$ M$_\odot$. The gravitational softening lengths employed are $\epsilon_{\rm DM} = 0.72$ ckpc and $\epsilon_{\rm bar} = 0.41$ ckpc for DM and baryon particles respectively, the latter corresponding to ≈ 60 pc at $z = 6$, i.e. a factor of ≈ 3 lower than in B18. Instead of the multiphase model by Springel & Hernquist (2003), the ISM is described by means of the MUlti Phase Particle Integrator (MUPPI) sub-resolution model (e.g. Murante et al., 2010; Valentini et al., 2020). It features metal cooling, thermal and kinetic stellar feedback, the presence of a UV background, and a model for chemical evolution, following Tornatore et al. (2007). In particular, star formation is implemented with a H_2-based prescription instead of a density-based cryterion, as in B18.

A fraction of the accreted rest-mass energy is radiated away from the BHs according to eq. 7.1, assuming a radiative efficiency of $\epsilon_{\rm r} = 0.03$. A fraction $\epsilon_{\rm f} = 10^{-4}$ of the radiated luminosity $L_{\rm bol}$ is thermally coupled to the gas surrounding the BHs, and it is isotropically distributed into the gas. BHs with $M_{\rm BH} = 10^5 h^{-1}$ M$_\odot$ are seeded in a $10^9 h^{-1}$ M$_\odot$ BH-less DM halo, and they grow by accretion and mergers. In this work, we make use of the fiducial run of the suite by V21, which we will refer to as *AGNthermal*.

7.2.2 Radiative transfer

For each snapshot of the hydrodynamical simulations we identify all the AGN with $L_{\rm bol} > 10^{10}$ L$_\odot$ as per eq. 7.1. We choose for the RT post-processing a number of snapshots sufficient to well sample the AGN luminosity range between 10^{10-14} L$_\odot$. We select for each RT simulation a cubic region of 60 kpc size ($\sim 50\%$ of the virial radius), centered on the center of mass of the

most-massive halo. We choose snapshots from different simulations in order to make our results not biased by the subgrid physics and AGN feedback prescription implemented[2]. By choosing multiple snapshots for each run, we aim to make our results more general and not depending on the specific gas/stellar morphology and AGN/star formation activity of a single snapshot. The selected snapshots are post-processed with SKIRT[3] (Baes & Camps, 2015; Camps et al., 2016). For the RT setup, a dust component and the radiation sources need to be specified (see Chapter 5 for more details about the RT setup).

Given that the processes related to the dust production and destruction are not explicitly followed in the hydrodynamic simulations adopted in this work, the dust distribution is assumed to track the one of the metals. We assume a linear scaling (e.g. Draine et al., 2007) parameterized by the dust-to-metal ratio $f_{\rm d} = M_{\rm d}/M_Z$, where $M_{\rm d}$ is the dust mass and M_Z is the total mass of all the metals in each gas particle. This parameter acts as a normalization factor for the overall dust content. In order to make our results less dependent on the specific choice of this parameter, we adopt two values for $f_{\rm d}$: a MW-like value, $f_{\rm d} = 0.3$, and a lower value, $f_d = 0.08$, found to reproduce the observed SED of a $z \sim 8$ galaxy (Behrens et al., 2018). We adopt dust optical properties of the Small Magellanic Cloud (SMC, Weingartner & Draine 2001). We assume gas particles hotter than 10^6 K to be dust-free because of thermal sputtering (e.g. Draine & Salpeter, 1979).

Dust is distributed in the computational domain in an octree grid, whose maximum number of levels of refinement for high dust density regions is chosen according to the spatial resolution of the hydrodynamic simulations. For the runs from B18, we adopt 8 levels of refinement, achieving a spatial resolution of ≈ 230 pc in the most refined cells, comparable with the softening length in the hydrodynamic simulation (≈ 210 pc at $z = 6$); for the runs from V21, we adopt 10 levels of refinement, with the highest resolved cells having a size of ≈ 59 pc, consistent with the softening length (≈ 87 pc at $z = 6$). For illustration purposes, in Fig. 7.1 we show the dust surface density distribution for the snapshot at $z = 6.1$ of the run *AGNthermal*, which we consider as a representative case. Within the simulated computational box, three quasar-hosts galaxies with $L_{\rm bol} > 10^{10}$ L$_\odot$ are present, and are labelled in the figure.

We make use of the stellar synthesis models by Bruzual & Charlot (2003) to implement the stellar radiation, according to the mass, age and metallicity of each stellar particle. For the black holes, we use the composite power-law AGN SED we introduced in DM21a, which is derived on the basis of several observational and theoretical works (Shakura & Sunyaev, 1973; Fiore et al., 1994; Richards et al., 2003; Sazonov et al., 2004; Piconcelli et al., 2005; Gallerani et al., 2010; Lusso et al., 2015; Shen et al., 2020a). The AGN SED reads:

$$L_\lambda = c_i \left(\frac{\lambda}{\mu{\rm m}}\right)^{\alpha_i} \left(\frac{L_{\rm bol}}{{\rm L}_\odot}\right) {\rm L}_\odot\ \mu{\rm m}^{-1}, \tag{7.2}$$

where i labels the bands in which we decompose the spectra and the coefficients c_i are determined by imposing the continuity of the function based on the slopes α_i, as detailed in Table 2 in DM21a. In particular, we make use of the *fiducial* AGN SED, characterized by an UV spectral slope $\alpha_{\rm UV} = -1.5$. The SED is then normalized according to the bolometric luminosity of the AGN (see eq. 7.1). The radiation field is sampled by using a grid composed of 200 logarithmically spaced bins, covering the *rest-frame* wavelength range $[0.1 - 10^3]$ μm, with 10^6 photon packets launched from each source per wavelength bin.

In order to isolate the contribution of the AGN to dust heating, we perform each RT run with and without AGN radiation: we refer to the first group of simulations as *AGN on*, and to

[2]Different feedback prescriptions can significantly affect the gas/metals distribution and the star formation, as already shown in the original work by B18 and V21, and also in DM21a (see their Figure 1) and Vito et al. (2022).

[3]Version 8, `http://www.skirt.ugent.be`.

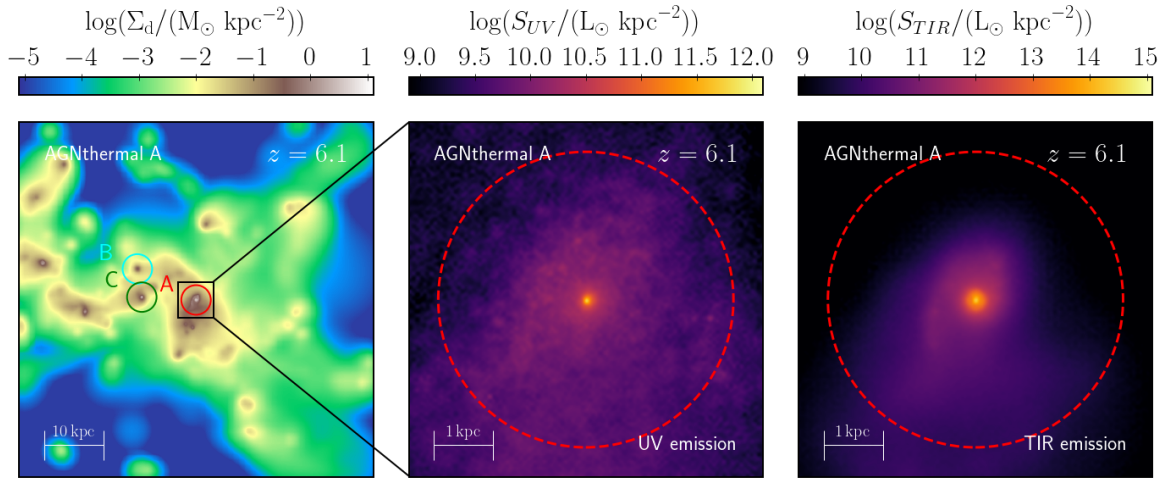

Figure 7.1: Maps of relevant properties relative to a representative snapshot, the $z = 6.1$ one of the run *AGNthermal*. *Left panel:* Surface density distribution of the dust component within the computational box of 60 kpc size. The coloured circles indicate the 2.5 kpc region around each quasar-host with $L_{\rm bol} > 10^{10}$ L$_\odot$, which we use to compute the dust temperature PDF in Section 7.3.1 and the synthetic SEDs in Section 7.3.2. *Middle panel:* UV emission predicted by our RT calculations around source A in a smaller region of 6 kpc size. The radius of 2.5 kpc is shown with a dashed red line. *Right panel:* Same as the middle panel, but for the TIR emission.

the second as *AGN off*. The runs *AGN off* and *AGN on* share the same dust/stellar content and distribution: they differ from each other solely in the AGN contribution to the radiation field.

As an illustrative case, we show in the middle and right panel of Fig. 7.1 the UV ($0.1-0.3\,\mu$m) and total-infared (TIR, $8-1000\,\mu$m) emission in a zoomed region around source A.

7.3 AGN contribution to dust heating

We now investigate the contribution of the AGN radiation to the dust heating, by comparing the dust temperature distribution of the *AGN on* and *AGN off* runs. We first analyze the actual dust temperature in the host galaxies in Section 7.3.1; then we discuss in Section 7.3.2 the dust temperature that would be inferred for our quasar-hosts by treating our synthetic SEDs as mock observations.

7.3.1 Physical dust temperature

For each RT run, we select the region within 2.5 kpc around each AGN, which encloses the AGN host. This size is comparable to the galaxy sizes in the simulations and to the dust continuum size of the majority of the quasar-hosts observed at $z \sim 6$ (e.g. Venemans et al., 2020). The dust masses in the AGN hosts are in the range $3.1\times10^6-2.0\times10^8$ M$_\odot$, whereas the AGN bolometric luminosities are $L_{\rm bol} = 10^{10-14}$ L$_\odot$. We compute the probability distribution function (PDF) of the dust temperatures in this region for the *AGN off* and *AGN on* runs as done in DM21a. For each AGN, we consider the cells in the dust grid within 2.5 kpc from the BH, and we compute the mass-weighted dust temperature PDF $\langle T_{\rm d}\rangle_M$ by weighting the dust temperature in each cell $T_{\rm d,i}$ with its dust mass $M_{\rm d,i}$. We also compute the luminosity-weighted dust temperature PDF $\langle T_{\rm d}\rangle_L$, assuming that each cell emits as a grey-body $L_{\rm TIR,i} \propto M_{\rm d,i} T_{\rm d,i}^{4+\beta_{\rm d}}$, where $\beta_{\rm d}$ is the dust emissivity index (see Section 7.3.2), which for this calculation is set to $\beta_{\rm d} = 2$[4]. Finally, we compute the

[4]We verified that varying $1.5 < \beta_{\rm d} < 2.5$ does not change significantly our results.

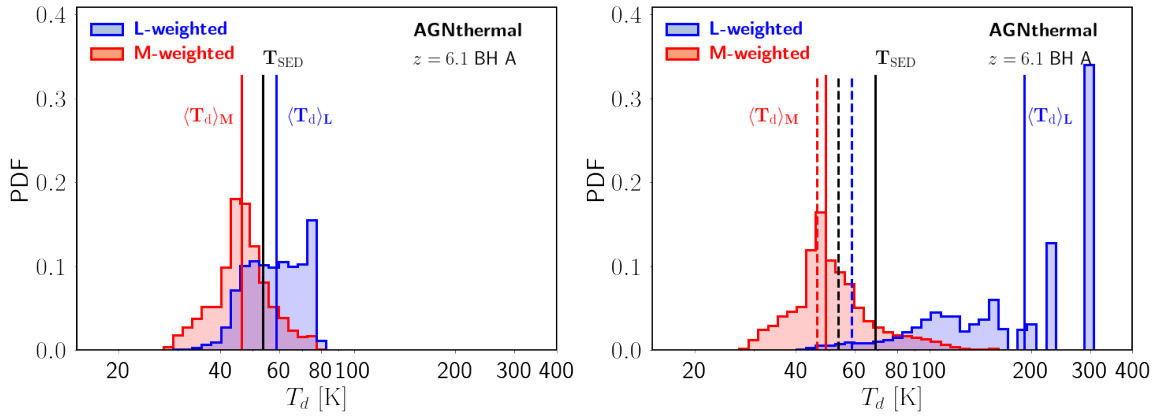

Figure 7.2: *Left panel: probability Distribution Function (PDF) of the mass-weighted (red) and luminosity-weighted (blue) dust temperature (see Section 7.3.1) of the computational cells within* 2.5 *kpc from a representative AGN (i.e. source A in Fig. 7.1) in the* AGN off *case. The red and blue solid lines mark the median values of the mass-weighted and luminosity weighted PDFs,* $T_{\rm d,M}$ *and* $T_{\rm d,L}$*, respectively. The black solid line indicates the value of* $T_{\rm SED}$ *derived from the SED fitting (see Section 7.3.2).* Right panel: *same as the left panel, but for the* AGN on *case. For comparison, we also show in this panel the values of* $T_{\rm d,M}$*,* $T_{\rm d,L}$ *and* $T_{\rm SED}$ *for the* AGN off *case with dashed lines, adopting the same color legend as in the left panel.*

median values of the PDFs within the 2.5 kpc regions, $\langle T_{\rm d} \rangle_M$ and $\langle T_{\rm d} \rangle_L$, and we indicate these medians with $T_{\rm d,M}$ and $T_{\rm d,L}$, respectively. The latter is expected to be dominated by the hottest regions in the AGN proximity, whereas the former should be more representative of the bulk of the dust in the ISM. As an example, in Fig. 7.2, we compare the PDFs in the *AGN off* and *AGN on* cases derived for the representative quasar-host also used in Fig. 7.1. The presence of AGN radiation shifts the dust temperature distribution toward higher values. This effect is significant for the luminosity-weighted distribution, but it is also noticeable in the mass-weighted one.

In order to analyze in a more systematic way the effect of the AGN radiation on the dust temperature distributions, we compare the median values of the weighted PDFs for all of the considered *AGN off* and *AGN on* cases in Fig. 7.3. We find that although $T_{\rm d,L} < 70$ K in the *AGN off* cases, in agreement with the values usually assumed in observations to derive $M_{\rm d}$ and SFR, it increases up to $T_{\rm d,L} \approx 250$ K in the *AGN on* runs, when AGN radiation is accounted for, showing a strong trend with AGN luminosity. This result is consistent with what found in Di Mascia et al. (2021a), that AGN radiation can effectively heat dust in the ISM on ≈ 200 pc scales from the AGN.

A similar trend is also found for the mass-weighted dust temperature, even if less pronounced. In fact, while $T_{\rm d,M} \lesssim 50$ K in the *AGN off* runs, it increases up to $T_{\rm d,M} \approx 80$ K in the *AGN on* runs. This result has important consequences: it implies that AGN can heat the bulk of the ISM dust in their host galaxies, not only in their proximity. The behaviour of $T_{\rm d,M}$ is also found to correlate with the AGN luminosity. In both the mass-weighted and luminosity-weighted distributions the median dust temperatures in *AGN on* runs start to deviate (see Fig. 7.3) from the *AGN off* values already at $L_{\rm bol} \approx 10^{12}$ L$_\odot$ for most of the AGN hosts. For $L_{\rm bol} \gtrsim 10^{13}$ L$_\odot$, $T_{\rm d,M}$ increases by $20-50\%$. This is a significant effect: since the TIR luminosity scales approximately as $L_{\rm TIR} \propto T_{\rm dust}^{4+\beta}$ (where $\beta \approx 2$ is the dust emissivity index, see Section 7.3.2), a 50% boost from the AGN radiation results in a 10-fold increase of the total IR luminosity. Given that most of the quasars detected at $z > 6$ have bolometric luminosities larger than 10^{13} L$_\odot$ (e.g. Yang et al., 2021a), our calculations suggest that AGN dominate the dust heating in the majority of these sources.

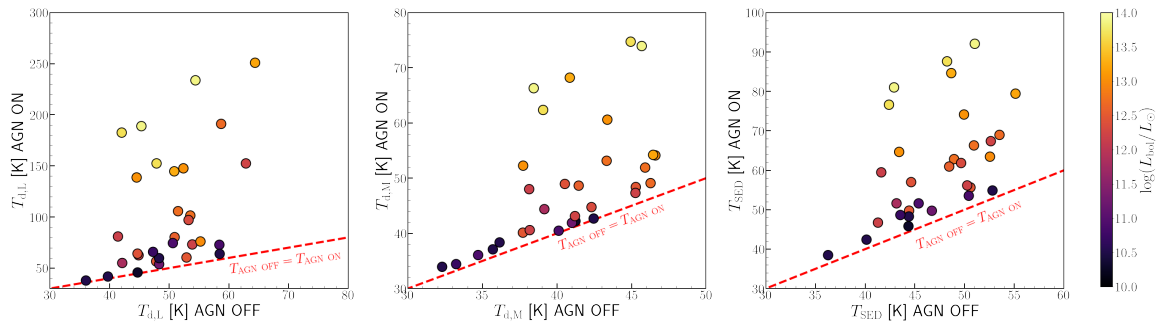

Figure 7.3: Comparison of different proxies of dust temperature distribution between the *AGN off* and *AGN on* cases for all the considered AGN and snapshots. *Left panel*: median of the luminosity-weighted dust temperature $T_{\rm d,L}$; *Middle*: median of the mass-weighted dust temperature $T_{\rm d,M}$; *Right*: SED-derived dust temperature $T_{\rm SED}$ (see Section 7.3.2). Each run is colour-coded according to the AGN bolometric luminosity. The dashed red lines marks the locus where temperatures in the *AGN on* and *AGN off* cases are equal.

7.3.2 Dust temperature inferred from the SED

We derive the synthetic SED of each AGN-host by considering the radiation emitted within 2.5 kpc from the position of the BHs, consistently with the analysis in Section 7.3. The SED of a representative quasar-host is shown in Fig. 7.4. We then treat the SEDs as mock observations and we perform an SED fitting of the FIR emission.

The dust emission in the optically-thin limit can be expressed as (e.g. Carniani et al., 2019b):

$$S^{\rm obs}_{\nu_{\rm obs}} = \frac{1+z}{d_L^2}\, M_{\rm dust}\, \kappa_\nu\, B_\nu(T_{\rm SED}), \tag{7.3}$$

where $S^{\rm obs}_{\nu_{\rm obs}}$ represents the observed flux at the observed frequency $\nu_{\rm obs}$ and $B_\nu(T_{\rm SED})$ is the black-body emission at the dust temperature $T_{\rm SED}$. We underline that $T_{\rm SED}$ should not be interpreted as a "true" dust temperature of the dust component in the galaxy, which has instead a complex multi-temperature distribution (see Appendix A in Sommovigo et al. 2021a for an in-depth discussion). The dust opacity κ_ν is usually expressed as a power-law $\kappa_0 \left(\frac{\lambda_0}{\lambda}\right)^\beta$, with the parameters κ_0, λ_0 and the dust emissivity index β derived from theoretical models (e.g. Weingartner & Draine, 2001; Bianchi & Schneider, 2007) or from observations (e.g. Beelen et al., 2006; Valiante et al., 2011). We adopt the dust opacity κ_ν appropriate for the SMC dust model from Weingartner & Draine (2001), the same used in the RT simulations. This leaves the dust mass $M_{\rm dust}$ and the dust temperature $T_{\rm dust}$ as the only free parameters. In most observations of $z \sim 6$ quasar-hosts only one or a few data points in the rest-frame FIR are available (e.g. Venemans et al., 2020), without probing the SED peak, therefore the estimate of $T_{\rm SED}$ is very uncertain, with few notable exceptions (e.g. Pensabene et al., 2021). In order to avoid this problem, we consider the FIR portion of our synthetic SEDs at $\lambda_{\rm obs} > 200\ \mu$m (corresponding to more than 70 bins of the wavelength grid), which is the wavelength range typically probed by $z \sim 6$ quasars observations, and we assign a 10% error to the flux at each wavelength, in order to mimic the typical experimental uncertainties (e.g. Venemans et al., 2020).

In Fig. 7.2 we mark the value of $T_{\rm SED}$ for the representative quasar-host shown in Fig. 7.1. $T_{\rm SED}$ has an intermediate value between $\langle T_{\rm d}\rangle_M$ and $\langle T_{\rm d}\rangle_L$. It also tends to be more similar to the median of the mass-weighted distribution rather than to the luminosity-weighted one in the *AGN on* cases. In the right-most panel of Fig. 7.3 we show how the measured values of $T_{\rm SED}$ change when including AGN radiation for all the simulated quasar-hosts. We find a significant increase of the estimated $T_{\rm SED}$ when AGN radiation is included, from $T_{\rm SED} \lesssim 60$ K in the *AGN off* to $T_{\rm SED} \approx 90$ K in the *AGN on* runs. In particular, we notice that all the simulated galaxies

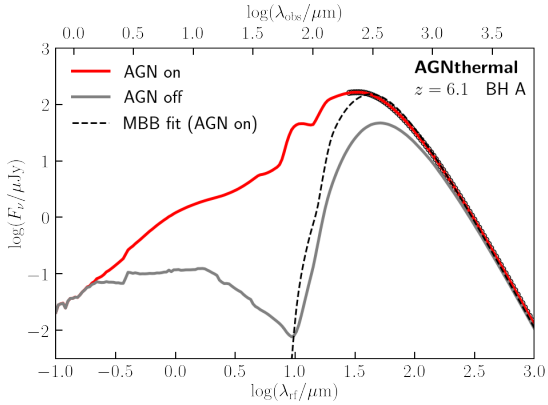

Figure 7.4: Spectral Energy Distribution (SED) emitted from the 2.5 kpc size area selected around a representative quasar-host (source A in Fig. 7.1). The solid red line shows the flux for the *AGN on* RT run and the grey solid line the *AGN off* case. The black dashed line indicate the SED fit of the rest-frame FIR part of the SED performed according to the modified black-body function (eq. 7.3). The black circles indicate the points of the SED used for the fit.

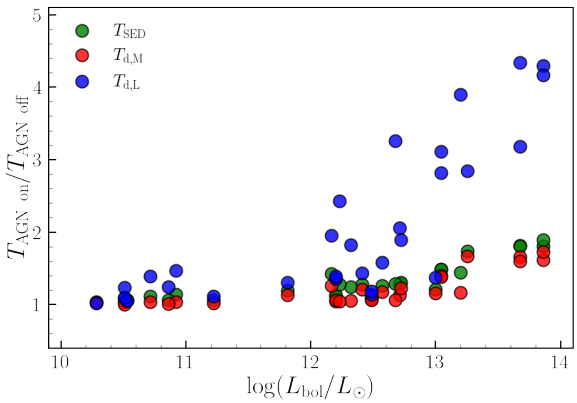

Figure 7.5: Ratio between the *AGN on* and *AGN off* runs for the median of the luminosity-weighted temperature $T_{\rm d,L}$ (blue circles), the median of the mass-weighted temperature $T_{\rm d,M}$ (red circles), and $T_{\rm SED}$ (green circles) as a function of the bolometric luminosity of the AGN.

with $T_{\rm SED} \gtrsim 60$ K are among *AGN on* runs, suggesting that galaxies with $T_{\rm SED} \gtrsim 60$ K are likely AGN-powered. A trend with the AGN bolometric luminosity is also present, with the most luminous AGN showing the largest increase of $T_{\rm SED}$. This result is consistent with the behaviour of the mass-weighted and luminosity-weighted PDFs shown in Fig. 7.3, as $T_{\rm SED}$ partly reflects the actual dust temperature distribution in the AGN host. Therefore, the correlation found in Fig. 7.3 is a direct consequence of the AGN heating of the ISM dust in the simulated galaxies, as seen in the behaviour of $T_{\rm d,L}$ and $T_{\rm d,M}$ (left and middle panel of Fig. 7.3).

In Fig. 7.5 we show the ratio of the values obtained for these three indicators of the dust temperature distribution. As expected, all the ratios increase with $L_{\rm bol}$, with the shallowest (steepest) rise for the mass-weighted (luminosity-weighted) median; $T_{\rm SED}$ is intermediate between the two. In particular, we note that for the brightest AGN, $L_{\rm bol} \gtrsim 10^{13}$ L$_\odot$, the SED-derived dust temperature $T_{\rm SED}$ increases by a factor of almost 2, similarly to the median of the mass-weighted $T_{\rm d,M}$.

7.4 Star formation rate estimate from the FIR

The analysis in Section 7.3 shows that in the simulated $z \sim 6-7$ galaxies AGN radiation can contribute significantly to dust heating, affecting the temperature of the bulk of the dust in the ISM. Given that the star formation rate estimates in high-z quasar-hosts are based on the assumption that stellar radiation is the only source of dust heating in the host galaxy, we investigate the impact of the AGN contribution on the inferred SFRs values. We now take advantage of our RT simulations by using our synthetic SEDs as mock observations as in Section 7.3.2. Then, we estimate the star formation rate in the host galaxy from the FIR luminosity, assuming that dust is heated only by stars. Finally, we compare the inferred star formation rate with the actual star formation rate in the quasar hosts, which is known from the hydrodynamic simulations.

For each simulated AGN host, we estimate the star formation rate from our synthetic SEDs as follows. We compute the total IR luminosity $L_{\rm TIR}$ by integrating equation 7.3 over the wavelength range $8-1000\ \mu$m. Then, we make use of the $\rm SFR_{FIR}-L_{TIR}$ calibration in Kennicutt & Evans (2012), based on Murphy et al. (2011). This relation provides an estimate of the *obscured* star formation rate, i.e. the stellar radiation that is absorbed by dust, *assuming that dust heating is purely due to stars*. We perform this computation both for the *AGN off* and the *AGN on* runs, in order to quantify the AGN contribution to the SFR estimate.

For *AGN off* runs we also include the *unobscured* star formation rate, by adopting the $\rm SFR_{UV}-L_{FUV}$ calibration from Kennicutt & Evans (2012)[5]. The contribution of $\rm SFR_{UV}$ is not accounted for in the *AGN on* runs, consistently with observational works, since the UV emission is expected to be dominated by the quasar. We note that accounting also for $\rm SFR_{UV}$ would increase the discrepancies between the inferred SFRs in the *AGN on* and *AGN off* cases, thus reinforcing our conclusions.

We compare the SFR inferred from the SED as described above with the actual SFR in the AGN hosts, which is known from the hydrodynamical simulations. The SFR is computed in the region within 2.5 kpc from the position of each BH in order to be consistent with the synthetic SED. We consider the SFR averaged over the past 100 Myr, which we found to provide the best agreement with the UV calibration in the *AGN off* runs. Fig. 7.6 presents the ratio between the SED-based SFR and the true SFR as a function of the AGN bolometric luminosity. For the runs without AGN radiation the SFR estimated from the SED fitting is overall in agreement with the actual star formation rate in the AGN hosts (see black histogram in the right panel of Fig. 7.6). This consistency check demonstrates that our procedure correctly recovers the SFR in the host galaxy if dust heating due to AGN radiation is negligible.

In the *AGN on* runs, the SFR inferred from the SED tends to over-estimate the actual star formation rate, with a discrepancy that becomes significant (a factor ≈ 3) at bolometric luminosities $L_{\rm bol} = 10^{12}\ \rm L_\odot$, and approximately an order of magnitude of discrepancy for most of the runs at $L_{\rm bol} \gtrsim 10^{13}\ \rm L_\odot$. This behaviour is consistent with our findings in Section 7.3, in particular with the trend of $T_{\rm SED}$ between *AGN on* and *AGN off* runs as a function of AGN luminosity. As discussed in Section 7.3, an increase by a factor of ≈ 2 in the dust temperature results in a $\approx 10\times$ increase of the total infrared luminosity. Given that the IR luminosity is assumed to come from dust-reprocessed stellar light only, also the inferred SFR increases by the same factors.

We fit the relation $\log(\rm SFR_{SED}/SFR_{true})$ vs $\log(L_{\rm bol}/\rm L_\odot)$ with a power-law expression, find-

[5]We use the results from Madau & Dickinson (2014) – see their Fig. 4 – to correct the conversion factor, in order to take into account the different Initial Mass Function (IMF) adopted in the simulations (Chabrier, 2003) with respect to the one used in the Kennicutt & Evans (2012) calibration (Kroupa & Weidner, 2003).

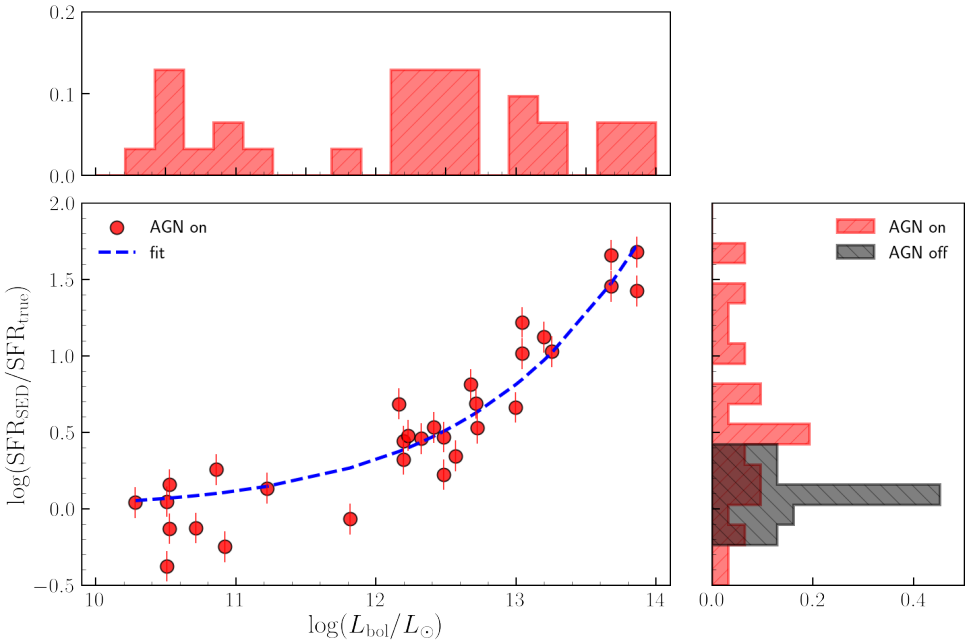

Figure 7.6: Ratio of the total SFR (estimated via FIR SED fitting, as explained in Section 7.4) to the true SFR in the AGN hosts as a function of the luminosity of the AGN. The blue dashed line indicates the exponential fit as in eq. 7.4. The upper inset shows the distribution of AGN luminosities. The inset on the right shows the distribution of the $\mathrm{SFR_{SED}}$ to $\mathrm{SFR_{true}}$ ratio for the *AGN on* and *AGN off* runs in red and black histograms respectively. We note that the distribution of the SFR estimates of the *AGN off* runs is consistent with the actual SFR.

ing:

$$\log\left(\frac{\mathrm{SFR_{SED}}}{\mathrm{SFR_{true}}}\right) = 7.68\times10^{-14}\left[\log\left(\frac{L_{\mathrm{bol}}}{\mathrm{L_\odot}}\right)\right]^{11.7}. \tag{7.4}$$

This relation quantifies the over-estimate of the star formation rate in AGN-hosts at $z \gtrsim 6$, as a function of their bolometric luminosity. It can be applied to correct the SFR in quasar-hosts with observed luminosity in the range $10^{10} \lesssim L_{\mathrm{bol}}/\mathrm{L_\odot} \lesssim 10^{14}$, and a well-constrained T_{SED} from the SED fitting. We assign a 0.15 dex uncertainty to this expression based on the scatter between the same run performed with different values of the dust-to-metal ratio.

As a practical case, we apply this correction to the quasar J2348-3054, recently studied in Walter et al. (2022). Instead of $\mathrm{SFR} \approx 4700\ \mathrm{M_\odot\,yr^{-1}}$, we find a $\mathrm{SFR} \approx 370\ \mathrm{M_\odot\,yr^{-1}}$, in broad agreement with the [C II]-inferred $\mathrm{SFR} \approx 530\ \mathrm{M_\odot\,yr^{-1}}$ from the relation by Herrera-Camus et al. (2018).

7.5 Summary and conclusions

In this Chapter, we investigated whether stellar radiation can be considered the main source of dust heating in $z \sim 6$ quasar-hosts, and the reliability of the star formation rate (SFR) inferred from the total infrared (TIR, $8-1000\ \mu\mathrm{m}$) emission in these galaxies. We combined cosmological hydrodynamic simulations with radiative transfer (RT) calculations, and simulated each galaxy with and without AGN radiation, in order to isolate the AGN contribution to dust heating.

We find that AGN with $L_{\mathrm{bol}} \gtrsim 10^{13}\,\mathrm{L_\odot}$ effectively heat the bulk of the dust in the interstellar medium (ISM) on galaxy scales, and not only the dust on ≈ 100 pc from their surrounding. As a result, the actual IR emission that comes from dust-reprocessed stellar light can be significantly over-estimated.

We quantify this effect by applying the $\mathrm{SFR} - L_{\mathrm{TIR}}$ relation by Kennicutt & Evans (2012) to the synthetic Spectral Energy Distribution (SEDs) of the simulated quasar-hosts, and compare the results with the "true" SFR in the hydrodynamic simulations. We find that the SFR tends to be overestimated by a factor of ≈ 3 (≈ 30) for $L_{\mathrm{bol}} \approx 10^{12}\ \mathrm{L}_{\odot}$ ($L_{\mathrm{bol}} \approx 10^{13}\ \mathrm{L}_{\odot}$, see Fig. 7.6). We also provide a simple relation (eq. 7.4) that quantifies the overestimate of the SFR in terms of the AGN luminosity.

We note that our results might be sensitive to the spatial resolution of the hydrodynamic simulations adopted in this work, the assumed models for the ISM physics, and to the numerical setup of the RT calculations. It would be valuable to repeat this analysis using simulations achieving higher spatial resolutions and a different implementation of the radiative transfer calculations.

The dust attenuation law in $z \sim 6$ quasars

8

8.1 Introduction

As discussed in Section 3.3, the high dust mass content ($M_{\rm dust} = 10^7 - 10^8$ M$_\odot$) in $z \sim 6$ quasar-hosts (e.g. Bertoldi et al., 2003a; Michałowski et al., 2010; Carilli & Walter, 2013; Gallerani et al., 2017b) suggests that the dust production mechanisms in place in the early Universe might be different than in the local one (Gall et al. 2011; Leśniewska & Michałowski 2019, but see also Valiante et al. 2009).

Individual measurements of extinction curves in $z > 6$ quasars also point in this direction, suggesting dust properties compatible with SN origin (Maiolino et al., 2004; Willott et al., 2007). Extinction curves in dust-reddened ($0.82 < A_{3000} < 2.0$) quasars at $3.9 < z < 6.4$ were found to deviate from the SMC extinction curve, with a tendency to flatten at $\lambda \lesssim 2000$ Å (Gallerani et al., 2010, hereafter G10).

Studies of gamma-ray bursts (GRBs) afterglows have been instead more controversial (e.g. Li et al., 2008b; Zafar et al., 2010; Stratta et al., 2011), showing a variety of extinction curves, compatible with the ones found in the local Universe (Zafar et al., 2011, 2018a,b; Bolmer et al., 2018). However, comparing these results with the ones found in AGN is not straightforward, as AGN-host galaxies might have a peculiar dust evolution history with respect to normal star-forming galaxies (e.g. Nozawa et al., 2015), or quasar spectra might reveal only the dust component in the AGN proximity, whereas GRBs probe the ISM of the host galaxy.

In the last decades, many theoretical works investigated the evolution of chemical species and dust abundance in galaxies (e.g. Dwek, 1998; Zhukovska et al., 2008; Pipino et al., 2011; Asano et al., 2013b; Hirashita, 2015; Nozawa et al., 2015; Hirashita & Aoyama, 2019). They showed that the grain size distribution evolves significantly throughout the galaxy evolution. In particular, Hirashita (2015); Nozawa et al. (2015); Hirashita & Aoyama (2019), and they have been used to make predictions for extinction curves of high-redshift galaxies. However, despite these efforts, many uncertainties in the theoretical models remain and a complete picture of the dust properties at high redshift is still missing.

A complementary approach to the problem is to study the dust properties at high-redshift by post-processing hydrodynamical simulations with radiative transfer calculations (e.g. Hou et al., 2017; Behrens et al., 2018; Narayanan et al., 2018; Aoyama et al., 2020; Shen et al., 2020b). This strategy has the advantage to combine the knowledge of the dust composition and grain size distribution of a theoretical model with the detailed gas/stars distributions predicted by the hydrodynamical simulations. This allows to reliably compute the SED and the attenuation curve, which can then be compared with data, gaining insight about the dust content and composition of the observed source. However, most of these studies focused on normal star-forming galaxies,

run	AGN feedback	$M_{\rm gas}$ [$\rm M_\odot$]	$M_\star$ [$\rm M_\odot$]	SFR [$\rm M_\odot\,yr^{-1}$]	$\dot{M}_{\rm BH}$ [$\rm M_\odot\,yr^{-1}$]
noAGN	no	2.9×10^{11}	1.2×10^{11}	600	-
AGNcone	bi-conical	1.4×10^{11}	7.0×10^{10}	189	89

Table 8.1: Summary of the main physical properties of the zoomed-in halo at $z = 6.3$ within a cubic region of 60 kpc size for the two hydrodynamic runs from B18 used in this work. For each run, we indicate: the feedback model used in the simulation, the gas mass ($M_{\rm gas}$), the stellar mass ($M_\star$), the star formation rate (SFR, averaged over the last 10 Myr), and the sum of the accretion rate of all the BHs in the selected region ($\dot{M}_{\rm BH}$).

and the AGN contribution has often been neglected in this context. An exception is the work by Li et al. (2008a), who applied radiative transfer calculations to cosmological hydrodynamical simulation, taking into account also AGN radiation and different dust models. They were able to reproduce the SED of the quasar SDSS J1148+5251, and inferring its dust properties, finding that dust from supernova origin best explain the observations. However, this work was limited to the analysis of a single quasar.

In chapter 6, we studied the AGN contribution to the IR emission in AGN-host galaxies at high-redshift, by post-processing cosmological hydrodynamical simulations of $z \simeq 6$ quasars (B18) with SKIRT (Baes & Camps, 2015; Camps et al., 2016). In this chapter, we make use of the same simulations to investigate dust attenuation properties at high-redshift. We explore different dust models, in which we vary the dust chemical composition, mass content and grain size distribution. Our goal is to constrain these quantities by comparing our synthetic SEDs with current observations of a large sample of bright $z \simeq 6$ quasars.

This chapter is organised as follows: in Section 8.2 we briefly summarise the hydrodynamical simulations and the numerical setup for the RT calculations. In Section 8.3 we present the attenuation curves obtained with our calculations, and investigate how their shape is affected by the AGN activity. In Section 8.4 we compare our results with observations of $z \simeq 6$ quasars; in Section 8.4.1 we identify the best-fit dust model. In Section 8.5 we discuss the implications of our results and some caveats of the model adopted. Summary and conclusions are given in Section 8.6.

8.2 Numerical setup

For this work, we make use of the zoom-in cosmological hydrodynamic simulations by B18, already described in details in Section 4.3. We make use of the *AGNcone* run, in which kinetic feedback distributed inside a bi-cone with an half-opening angle of $45°$, and of the control simulation with no BHs (*noAGN*). We focus on the $z = 6.3$ snapshots of these runs, already analysed in chapter 6. The different physical properties among these runs in terms of gas morphology, star formation rate and black hole activity allow us to study the relative impact of dust and radiation sources distribution on the observed attenuation curves in AGN-host galaxies. For our analysis we select a cubic region of 60 physical kpc size centred on the halo's centre of mass (the virial radius of the most massive halo is ≈ 60 kpc at this redshift). We report the main physical properties of the zoomed-in halo inside this region in Table 8.1.

The radiative transfer setup is also similar to the one used in the work described in chapter 6. In particular, the dust mass distribution is again derived by assuming a linear scaling with the gas metallicity (Draine et al., 2007), parametrizing the mass fraction of metals locked into dust as:

$$f_{\rm d} = M_{\rm d}/M_Z, \tag{8.1}$$

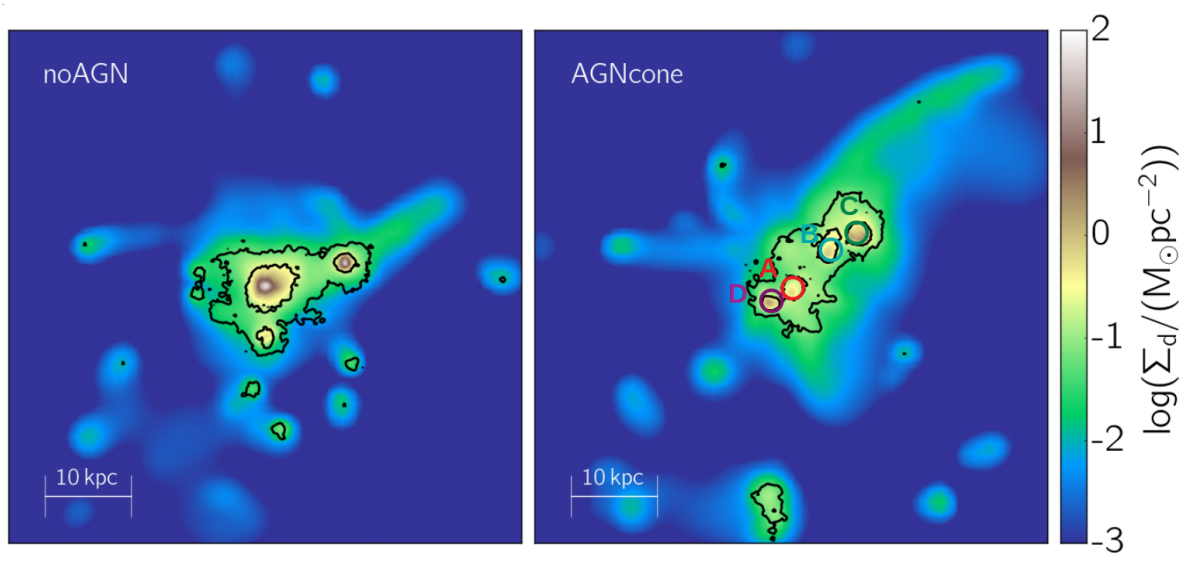

Figure 8.1: Dust surface density distribution of the most-massive halo at $z = 6.3$ in *noAGN* (left) and *AGNcone* (right) for a cubic region with size 60 kpc. A dust to metal ratio of $f_{\rm d} = 0.08$ is assumed. Coloured circles indicates the position of three active BHs (source A, B, C) and a star-forming galaxy (source D), part of a merging system (see Section 6.4.1). The black contours enclose regions with intrinsic UV emission higher than $S_{\rm UV} = 10^9$ L$_\odot$ kpc^{-2} (outer contours) and $S_{\rm UV} = 10^{10}$ L$_\odot$ kpc^{-2} (inner contours).

where $M_{\rm d}$ is the dust mass and M_Z is the total mass of all the metals in each gas particle in the hydrodynamical simulation. Gas particles hotter than 10^6 K are assumed to be dust-free because of thermal sputtering (Draine & Salpeter, 1979; Tielens et al., 1994). The parameter $f_{\rm d}$ acts as a normalization factor for the total dust content. In this work, we adopt a value ($f_d = 0.08$) tuned for hydro-simulations (Pallottini et al., 2017; Behrens et al., 2018) to reproduce the observed SED of a $z \sim 8$ galaxy (Laporte et al., 2017) and a MW-like value ($f_{\rm d} = 0.3$).

Dust is distributed in the computational domain in an octree grid with a maximum of 8 levels of refinement for high dust density regions, achieving a spatial resolution of ≈ 230 pc in the most refined cells, comparable with the softening length in the hydrodynamic simulation (≈ 200 physical pc at $z = 6.3$). In Figure 8.1, we show the dust distribution derived under these assumptions for our reference line of sight (los), namely the one aligned with the angular momentum of the particles inside the selected region. Regions with higher dust densities correspond to active star-forming regions, where gas metal enrichment is more effective. The dust distribution is affected by the AGN feedback, as also discussed in chapter 6. In *noAGN*, the higher star formation rate and the absence of AGN feedback lead to a more compact dust distribution, with a surface density $\Sigma_{\rm d}$ that peaks at $\Sigma_{\rm d} \approx 50$ M$_\odot$ pc^{-2}, whereas in *AGNcone* it reaches $\Sigma_{\rm d} \approx 2$ M$_\odot$ pc^{-2}.

The dust chemical composition and its detailed grain size distribution in early (AGN-host) galaxies are still a matter of debate, as discussed in Section 3.3 and in Section 8.1. We consider dust compositions and grain size distributions appropriate for the SMC and Milky Way (MW), by using the results of Weingartner & Draine (2001). However, we consider also dust models derived from the Weingartner & Draine (2001) models by modifying the grain size distribution (see Section 8.4). We defer the inclusion of a SN-type extinction curve to a future work. We note that these extinction curves are somewhat intermediate between the SMC and MW curve, with the notable difference that SN-type curves are essentially flat in the wavelength range $1700 - 2500$ Å We sample in SKIRT the grain size distribution of graphite, silicates and polycyclic aromatic hydrocarbons (PAHs) using 5 bins for each component. In the SMC model the fraction of dust in PAHs is set to zero. Note that the assumed dust properties do not necessarily imply that the resulting *attenuation* curves in our simulations match the extinction curve of the dust

model used, because attenuation curves depend not only on the dust properties but also on the geometry of the distributions of dust and radiative sources and on radiative transfer effects (e.g. Witt & Gordon, 1996; Bianchi et al., 2000; Behrens et al., 2018; Narayanan et al., 2018). We discuss this point in detail in Section 8.3.

8.3 Synthetic attenuation curves

The extinction curve entangles information on the dust composition and grain size distribution, which determine how dust grains absorb and scatter photons at different wavelengths. The absolute value of the attenuation at a given wavelength depends also on the overall dust content. Moreover, the spatial distribution of radiative sources and absorbers play a crucial role in determining the observed flux via radiative transfer effects. The same extinction curve, i.e. the same dust properties, can produce different attenuation curves depending on the dust-sources geometry (see e.g. Witt & Gordon, 1996; Bianchi et al., 2000; Seon & Draine, 2016; Behrens et al., 2018; Narayanan et al., 2018; Salim & Narayanan, 2020; Liang et al., 2021b). Dust attenuation is therefore a complex process that depends on dust content, dust composition and on the geometry of the system.

Fig. 8.2 shows the attenuation curves resulting from the RT simulations performed by DM21a for the runs *noAGN*008, *noAGN*03, *AGNcone*008, *AGNcone*03. The *noAGN* run (left panels) is representative of a high star-forming ($\mathrm{SFR} \approx 600\ \mathrm{M_\odot\,yr^{-1}}$) $z \sim 6$ galaxy, whereas the *AGNcone* run (right panels) is representative of a bright, $M_{UV} \sim -26$, $z \sim 6$ quasar. This comparison allows us to study the role of AGN (if any) in shaping the attenuation curve of its host galaxy. The attenuation curves shown are obtained by normalizing A_λ to its value at 3000 Å, where $A_\lambda = 1.086\tau_\lambda$, and $\tau_\lambda = -\ln(F_\lambda^{\rm obs}/F_\lambda^{\rm int})$, with $F_\lambda^{\rm obs}$ and $F_\lambda^{\rm int}$ being the observed and the intrinsic flux, respectively. In these simulations an intrinsic SMC-like extinction curve is adopted. In the next two subsections, we investigate the origin of the different attenuation curves, by focusing on the slope and the los-to-los variations.

8.3.1 Slope of the attenuation curve

The attenuation curves in the *noAGN* run are flatter than in *AGNcone*, for a fixed dust-to-metal ratio. In particular, the ratio of the attenuation at 0.1 ($A_{0.1}$) and 1 μm (A_1) rest-frame is $A_{0.1}/A_1 \approx 3-4$ in *noAGN*, whereas it can be as high as ≈ 10 in *AGNcone* runs. The flattening of the attenuation curve in the *noAGN* case is due to the high dust optical depth characterising this run ($\Sigma_{\rm d} \approx 50\ \mathrm{M_\odot\,pc^{-2}}$ for $f_{\rm d} = 0.08$, see Fig. 8.1). Regions characterized by these large surface densities become optically thick even to rest-frame NIR photons. This is evident by comparing the intrinsic and observed flux in the SED (see Fig. 8 in DM21a). Instead, the dust surface density is at most $\approx 2\ \mathrm{M_\odot\,pc^{-2}}$ in *AGNcone*, such that the attenuation is significant at the shortest wavelengths (where the extinction is higher because of the dust optical properties), but the dust is optically thin in the NIR.

The importance of the dust surface density in shaping the attenuation curve in our simulations can also be appreciated by comparing the results of the same run for different dust contents. In fact, for the reference los, the ratio between the attenuation at 0.1 μm, $A_{0.1}$, and the attenuation at 1 μm, A_1, decreases in *AGNcone* from the case $f_{\rm d} = 0.08$ to $f_{\rm d} = 0.3$. As a result, the attenuation curves in *AGNcone* tend to become flatter when increasing the dust content.

We also compare our predicted attenuation curves with the Calzetti et al. (1994) curve, and we find that *noAGN* is in very good agreement, whereas the curves in *AGNcone* are much steeper. We also consider the results obtained by Seon & Draine (2016), which performed radiative

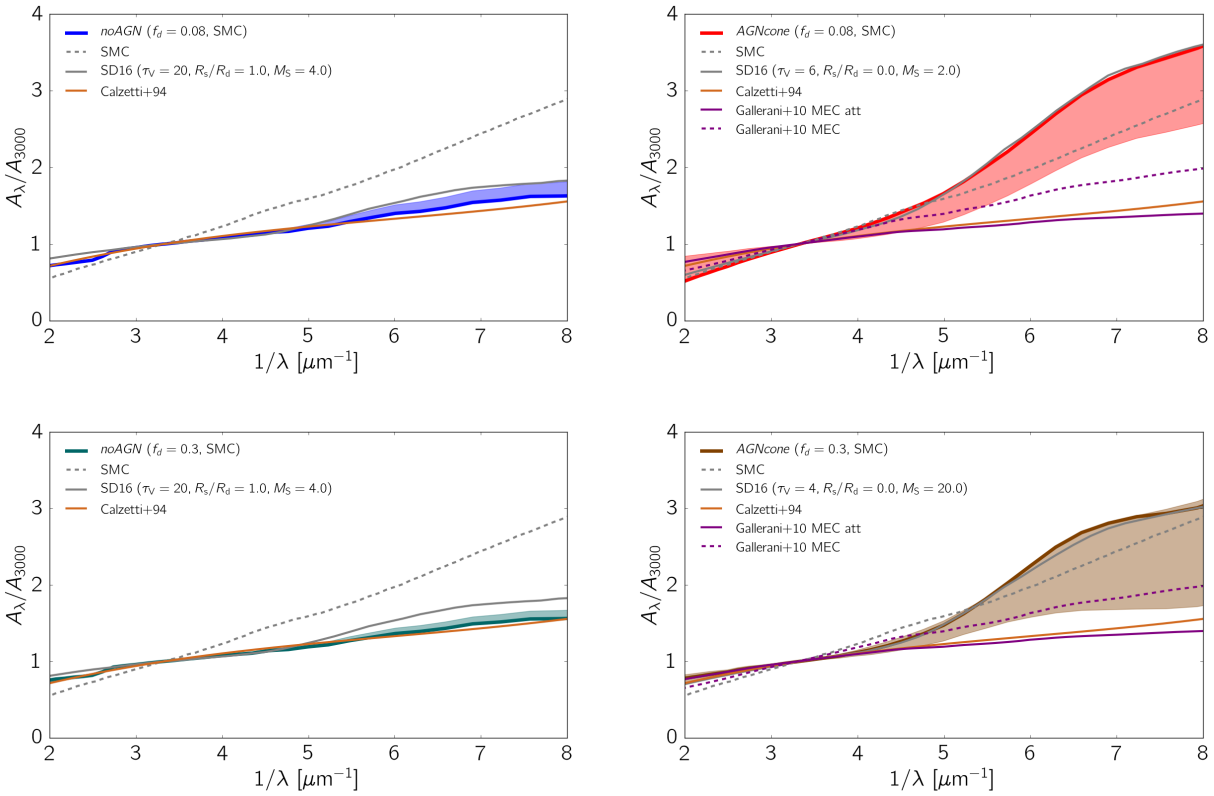

Figure 8.2: Attenuation curves (normalised to their value at 3000Å) corresponding to four of the runs performed in DM21a (see their Figure 8). The first column refer to *noAGN*, the second to *AGNcone*. The first row shows the case $f_{\rm d} = 0.08$, the second row $f_{\rm d} = 0.3$. The shaded area indicate the scatter between different lines of sight, used for the RT calculation. The solid line refers to our reference los. We also plot the SMC extinction curve (grey dashed line) and the Calzetti attenuation curve (light brown solid line, Calzetti et al., 1994). We also show with a grey solid line the attenuation curves derived in Seon & Draine (2016) (assuming an SMC extinction curve) that best match our attenuation curves for the reference los, and we quote the corresponding parameters $(\tau_{\rm V}, R_{\rm s}/R_{\rm d}, M_{\rm S})$ describing the geometry adopted for that model (see text for more details). For *AGNcone* we also show in purple the extinction curve (MEC, dashed line) and attenuation curve (MEC attenuated, solid line) obtained in G10 for high-redshift quasars.

transfer simulations of a single turbulent cloud, in order to study how the geometry of the cloud shapes the attenuation curves at fixed extinction curve (i.e. SMC, MW, LMC). A comparison between our results and simulations of a single cloud is quite difficult, but it can give useful insights. In the Seon & Draine (2016) models, dust and stars are assumed to be distributed in spherical symmetry inside a turbulent cloud (characterised by a Mach number $M_{\rm S}$), within a radius $R_{\rm d}$ and $R_{\rm s}$ respectively. $R_{\rm s} = 0$ represents a cloud with a bright point source at its center, whereas $R_{\rm s}/R_{\rm d} = 1$ corresponds to a case in which dust and stars are uniformly distributed. A model is uniquely determined by the triplet ($\tau_{\rm V}$, $R_{\rm s}/R_{\rm d}$, $M_{\rm S}$), where $\tau_{\rm V}$ is the V-band optical depth.

In each panel of Fig. 8.2 we show the model in Seon & Draine (2016) that best fits the attenuation curve for the reference los, based on a minimum squares criterion. For the *noAGN* run, we find that this model is characterised by a high optical depth ($\tau_{\rm V} = 20$, which is the maximum value adopted in their work) and by $R_{\rm s}/R_{\rm d} = 1$. This further confirms our explanation that the flatness of the attenuation curves in *noAGN* is determined by a very high dust surface density. The ratio $R_{\rm s}/R_{\rm d} = 1$, corresponding to the case of stars and dust well mixed, is consistent with

the compactness of the galaxies in *noAGN* (see also the left panel in Fig. 8.1). For the *AGNcone* run, the best-fit model by Seon & Draine (2016) has a lower optical depth $\tau_V = 4$ with respect to *noAGN*, which is again consistent with the dust surface density, and a ratio $R_s/R_d = 0$. This case corresponds to a geometry with a bright point source at the center of the cloud, which is a reasonable approximation for the geometry in the *AGNcone*.

The relation between the dust optical depth and the steepness of the attenuation curves, and how they change among the different simulations, can be understood with the following analytical argument. For this calculation, we make use of two simplified geometries, but it is useful to get a physical insight about how the dust optical depth affects the shape of the attenuation curves.

The attenuation at a specific wavelength λ is defined as: $A_\lambda = 1.086\tau_\lambda = -1.086\ln(T(\tau_\lambda, \zeta_\lambda, \xi_\lambda))$, where T is the *transmissivity*, i.e. the fraction of emergent light. ζ_λ and ξ_λ are functions of the albedo $\omega(\lambda)$ and the asymmetry parameter $g(\lambda)$ of the Henyey-Greenstein scattering phase function at the wavelength considered. The albedo is defined as $\omega(\lambda) = \sigma_{scat}(\lambda)/\sigma_{ext}(\lambda)$, where $\sigma_{scat}(\lambda)$ and $\sigma_{ext}(\lambda)$ are the scattering and extinction cross sections at the wavelength λ, respectively. We consider two different geometries: i) a point source surrounded by dust in a spherically symmetric distribution (*point source model*); ii) a sphere with dust and emitters uniformly distributed (*intermixed model*). For the first case, we adopt the classical solution for the transmissivity derived by Code (1973):

$$T_{ps}(\tau_\lambda, \zeta_\lambda, \xi_\lambda) = \frac{2}{(1+\zeta_\lambda)e^{\xi_\lambda\tau_\lambda} + (1-\zeta_\lambda)e^{-\xi_\lambda\tau_\lambda}}, \tag{8.2}$$

where the functions ζ_λ and ξ_λ have the following functional form:

$$\zeta_\lambda(\omega_\lambda, g_\lambda) = \sqrt{(1-\omega_\lambda)/(1-\omega_\lambda g_\lambda)} \tag{8.3}$$

$$\xi_\lambda(\omega_\lambda, g_\lambda) = \sqrt{(1-\omega_\lambda)(1-\omega_\lambda g_\lambda)}. \tag{8.4}$$

For the homogenous sphere, we make use of the solution obtained by Osterbrock (1989) (see also Appendix C in Városi & Dwek 1999):

$$T_{mix}(\tau_\lambda) = \frac{3}{4\tau_\lambda}\left\{1 - \frac{1}{2\tau_\lambda^2} + \left(\frac{1}{\tau_\lambda} + \frac{1}{2\tau_\lambda^2}\right)e^{-2\tau_\lambda}\right\}, \tag{8.5}$$

which expresses the escaping probability for photons in a sphere where sources and absorbers are homogeneously distributed and the opacity from the centre to the surface is τ_λ. We then rewrite eq. 8.2 and 8.5 by expressing the optical depth τ_λ in terms of the optical depth in the V-band, $\tau_\lambda = (\sigma_{ext}, \lambda/\sigma_{ext,V})\tau_V$, and making use of the formulas above, we can write the attenuation at a given wavelength A_λ in terms of the V-band optical depth.

In order to explore how the attenuation changes with optical depth, both in the optical/UV regime and in the NIR, we compute the attenuation at $0.1\ \mu m$ and $1\ \mu m$. For this calculation, we adopt the appropriate values for ω_λ, g_λ and σ_{ext} according to the results by Draine (2003b) for the SMC model (see their Fig. 4). In the left panel of Fig. 8.3 we plot the ratio $A_{0.1}/A_1$ as a function of τ_V with thick solid lines, for the point source (purple) and intermixed (orange) model; dashed lines represent $A_{0.1}$ and A_1, separately.

At low optical depths, the ratio $A_{0.1}/A_1$ converges to a constant in both cases, which is a function only of the optical properties of the dust mixture assumed. In particular,

$$\frac{A_{0.1}}{A_1} \xrightarrow{\tau_V \ll 1} \frac{\zeta_{0.1}\xi_{0.1}\sigma_{0.1}}{\zeta_1\xi_1\sigma_1} \simeq 46.87 \qquad \text{(point source)}$$

$$\frac{A_{0.1}}{A_1} \xrightarrow{\tau_V \ll 1} \frac{\sigma_{0.1}}{\sigma_1} \simeq 21.81 \qquad \text{(intermixed)},$$

where ζ_λ, ξ_λ, σ_λ have been evaluated at the wavelength indicated in the subscript. At high optical depths, the ratio converges to:

$$\frac{A_{0.1}}{A_1} \xrightarrow{\tau_V \gg 1} \frac{\sigma_{0.1}}{\sigma_1} \simeq 21.81 \qquad \text{(point source)}$$

$$\frac{A_{0.1}}{A_1} \xrightarrow{\tau_V \gg 1} 1 \qquad \text{(intermixed).}$$

In both cases, the ratio $A_{0.1}/A_1$ decreases with increasing optical depth. This effect is limited in the point source model, whereas it is much more significant in the homogeneous sphere model, for which this ratio tends to 1, corresponding to a flat attenuation curve.

We compare this analytical prediction with our simulations. We convert the dust surface density into an optical depth by using the extinction coefficient $\kappa_V = 1.72 \times 10^4\ \mathrm{cm^2\,g^{-1}}$. The maximum value of the V-band optical depth estimated in this way over the field of view for the chosen line of sight is $\approx$ 200 and $\approx$ 7 for *noAGN* and *AGNcone* respectively, assuming $f_\mathrm{d} = 0.08$.

We emphasize that the quoted opacity is the *slab* opacity, i.e. it represents the attenuation expected for a light beam passing through a slab with the assumed dust density and properties. The corresponding attenuation would read simply as $e^{-\tau}$, with τ being the opacity considered. However the *effective* opacity depends also on two essential ingredients: i) radiative transfer effects; ii) the relative distribution of dust grains and radiation sources. Both these factors can contribute in making the effective opacity much lower than the slab one.

Therefore, in order to make a more fair comparison of the results from the RT simulations (which include all the physical processes in play) with our analytical calculation (which include RT effects, but it is limited to a point-source geometry), we need to examine in more details how the slab opacity is distributed in the hydrodynamical runs. In fact, for a given line of sight, the optical depth varies from region to region, while the attenuation curves shown in Fig. 8.2 refer to the whole field of view.

We compute the Probability Distribution Function (PDF) of the slab opacity from the dust distribution maps in Fig. 8.1. We perform the calculation for the reference line of sight, in order to directly compare the results with the estimates provided above. We restrict the computation to the subregion of the field of view for which the slab V-band optical depth is $\tau_\mathrm{V} \gtrsim 10^{-3}$, and we compute the PDF by weighting τ_V with the *intrinsic* UV emissivity, in order to emphasize the contribution from the regions responsible for the emission.

In the right panel of Fig. 8.3 we show the UV-weighted PDF of the slab τ_V for *noAGN* (blue) and *AGNcone* (red), for the selected subregion of the field of view. The low density regions ($\tau_\mathrm{V} \ll 1$) contribute only to a small fraction of the weighted distribution, i.e. $\lesssim 15\%$ in *noAGN* and $\lesssim 1\%$ in *AGNcone*. The weighted fraction of dust in the range $1 < \tau_\mathrm{V} < 10$ is $\approx 16\%$ in *noAGN*, whereas it constitutes almost the totality ($\approx 98\%$) in *AGNcone*. This is because the dusty regions around active BHs in *AGNcone*, which are the most UV-emitting regions (see black contours in Fig. 8.1), are characterised by optical depths in this range. In *noAGN*, the most star-forming regions have $\tau_\mathrm{V} > 10$, which account for $\approx 69\%$ of the weighted distribution. Overall, the fraction of the field of view observed that has very high slab opacity is much higher in *noAGN* than in *AGNcone*. The low $A_{0.1}/A_1$ ratio in *noAGN* suggests that the sphere model is a reasonable approximation of the simulated object, characterised by a compact distribution. This is also consistent with the comparison with the attenuation curves derived by Seon & Draine (2016). *AGNcone* is instead better described by a combination of the point source and intermixed models. In fact, the point source model alone predicts a $A_{0.1}/A_1$ ratio a factor of ≈ 4 higher than what we find from the simulations in the relevant optical depth range,

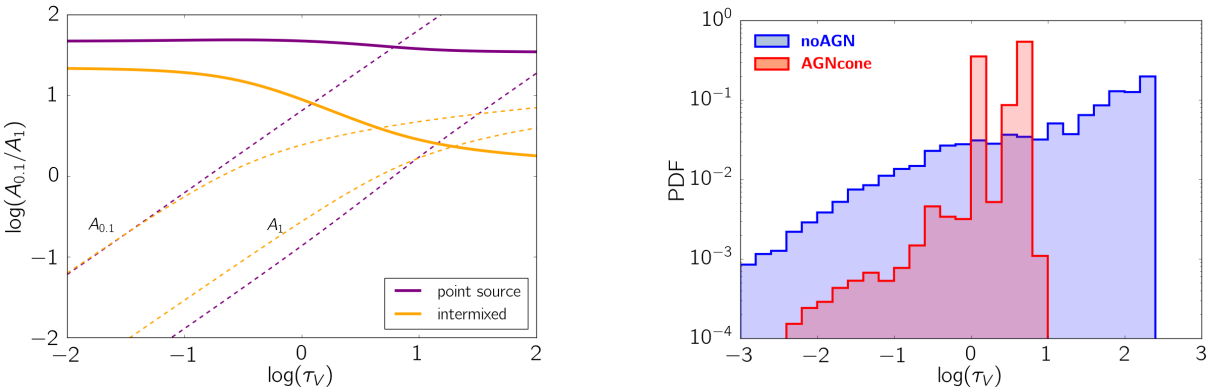

Figure 8.3: *Left panel:* Ratio (solid lines) of the attenuations $A_{0.1}$ and A_1 (dashed lines), at 0.1 μm and 1 μm respectively, as a function of the V-band optical depth $\tau_{\rm V}$. The attenuation is computed from equations 8.2 and 8.5, for the point source model (purple line) and the intermixed model (orange line) respectively, adopting parameter values appropriate for SMC dust. *Right panel*: PDF of the *slab* dust optical depth derived from the dust density distribution in the V-band, $\tau_{\rm V}$, for the reference los, assuming $f_{\rm d} = 0.08$. The PDF is weighted by the intrinsic UV emissivity in the field of view. The histograms show the case *noAGN* (blue) and *AGNcone* (red).

$1 < \tau_{\rm V} < 10$. Despite the simplicity of these two analytical models, they are still able to explain why the resulting attenuation curve in *noAGN* is almost flat, whereas the ratio $A_{0.1}/A_1$ reaches up to ≈ 10 in AGN runs.

The different optical depth distribution between the *noAGN* and *AGNcone* runs is the result of AGN feedback effects. The kinetic feedback prescription adopted in Barai et al. (2018) (see Section 4.3) generates powerful outflows that affect the gas distribution (and therefore the dust distribution) in the host galaxy, removing gas from the central regions and distributing it over several kpc, causing a more diffuse gas distribution as compared with the *noAGN* case. Therefore, the steep attenuation curves we predict for AGN-host galaxies are a direct consequence of a gas distribution affected by AGN activity. We further discuss this point in Section 8.5.3.

8.3.2 Line of sight attenuation variations

The los-to-los variation of the slope of the attenuation curves is much smaller in *noAGN* than in AGN runs, as can be seen from the width of the shaded regions in Fig. 8.2. In fact, $A_{0.1}/A_1 \approx 3-4$ in *noAGN*, while it is $\approx 6-9$ in *AGNcone*, for the $f_{\rm d} = 0.08$ case. This scatter is also due to the role of AGN in shaping the gas distribution in its surrounding[1]. In the *noAGN* case, the gas distribution is more compact and uniform around star-forming regions with respect to *AGNcone*. As a result, not only the attenuation curve for each los is flat, but also the scatter between each los is small. In *AGNcone*, feedback from accreting BHs alters the surrounding matter distribution by cleaning gas along some directions: some los are characterised by low dust column densities, resulting into steep attenuation curves, as the extreme ones in *AGNcone*, some other by high dust column densities, providing flat attenuation curve, similar to the ones found in *noAGN*.

[1]We underline that we are focusing here on how the attenuation curve changes with the line of sight, whereas in Section 8.3.1 we focus on the slope of the attenuation curve for a single los.

8.4 Comparison with quasar data

In Section 8.3, we studied the attenuation curves predicted by our simulations from a general point of view, in order to get insight about how the AGN activity affects the shape of the attenuation curves. In this Section we compare our results with observations of high-redshift quasars.

In Fig. 8.4, we show the synthetic SEDs obtained from our RT calculations for the *AGN-cone* run ($M_{\rm UV} = -27.97$ before accounting for dust attenuation) and we compare them with observations of $z \geq 6$ bright ($-29 \lesssim M_{\rm UV} \lesssim -26$) quasars (see Table A.1 in Appendix A.5 for all the sources of the sample) in the rest-frame UV-to-FIR (see also Fig. 9 in DM21a). We find qualitatively an overall agreement between the results of our model with $f_{\rm d} = 0.3$ and the fiducial AGN SED[2] with the photometric data points, with a notable exception. In the optical/UV band, the predicted SEDs luminosities fall short with respect to the data by G10; also, they show a steep decline in the far-UV, which is not observed. Different dust models might help in solving these mismatches. In fact, G10 fitted the spectra with extinction curves flatter than the SMC, hinting at different dust composition and/or grain size distribution at high-redshift.

In order to verify this working hypothesis, we considered several models, which we summarised in the following. We assume the dust-to-metal ratio to vary between $f_{\rm d} = 0.08$ and $f_{\rm d} = 0.3$, by considering also an intermediate value $f_{\rm d} = 0.15$, in addition to the ones considered in the previous analysis. Given that a steep slope ($-2.9 < \alpha_{UV} < -2.3$) is suggested in the analysis of reddened high-z quasars by G10 (see their Table 2), we include both the fiducial and UV-steep AGN SEDs in this calculation (see Section 5.4). We adopt a grain size distribution and composition that match the extinction properties either of the SMC or the MW (the latter was not included in the previous work). We also consider alternative dust models, which have the same dust composition of the fiducial SMC/MW cases, but with a cut in the dust grain distribution at $a_{\rm min}$. We use $a_{\rm min} = 0.01$ μm and $a_{\rm min} = 0.1$ μm. We note that the grain size distribution of the SMC/MW dust models by Weingartner & Draine (2001) implemented in SKIRT extends down to $a_{\rm min} = 0.001$ μm for silicates and carbonaceous grains, and down to $a_{\rm min} = 0.0003548$ μm for PAH molecules. Grain size distributions deficient in small dust grains are suggested by flat attenuation curves measured in AGN (e.g. Gaskell et al., 2004; Czerny et al., 2004; Gallerani et al., 2010) and justified by theoretical models (e.g. Hirashita & Aoyama, 2019; Nozawa et al., 2015).

We end up with a total of 20 models, which are summarised in Table 8.2. We show as a reference the *total* SED of these models in the left-hand panel in Fig. 8.4.

8.4.1 Best-fit model

In order to find the model that best matches the observational data, we perform a χ^2 analysis by comparing our synthetic SEDs with the observational data of the quasars selected. In the following, we will refer to rest-frame wavelengths $1 \lesssim \lambda \lesssim 5$ μm as Near-Infrared (NIR), $5 \lesssim \lambda \lesssim 40$ μm as Mid-Infrared (MIR) and $40 \lesssim \lambda \lesssim 350$ μm as Far-Infrared (FIR).

We first group the data in different bins: two bins for the rest-frame UV/optical wavelengths (0.145 μm and 0.3 μm rest-frame), ten bins for the NIR/MIR/FIR corresponding to the photometric bands of the Spitzer/Herschel telescopes, and a single bin for the rest-frame FIR, at 1.18 mm[3]. However, our models do not include the emission from hot dust in the torus around

[2] In DM21a we considered two values for the dust-to-metal ratio, $f_{\rm d} = 0.08$ and $f_{\rm d} = 0.3$, an SMC-like dust composition and both the fiducial and the UV-steep AGN SED.

[3] The rest-frame optical/UV data for the quasars in the sample are the spectra studied in G10 taken with the TNG/GEMINI telescopes. We take the flux at rest-frame 0.145 μm and 0.3 μm from these spectra. Regarding

AGN, which might be significant in the NIR, as we also pointed out in DM21a (see Figure 11). To overcome this limitation, we exclude from our analysis those bins characterised by a torus contamination above 5%[4]. This procedure excludes the bins at $8\ \mu$m and $24\ \mu$m. However, we argue that our results are not much affected by this choice, because the strongest constraints to the χ^2 fitting come from the rest-frame UV/optical data, as discussed below. We end up with a total of eleven bins at the following wavelengths: $0.145\ \mu$m (rest-frame), $0.3\ \mu$m (rest-frame), $3.6\ \mu$m, $4.5\ \mu$m, $5.8\ \mu$m, $100\ \mu$m, $160\ \mu$m, $250\ \mu$m, $350\ \mu$m, $500\ \mu$m, 1.18 mm. We assign to each bin an *average flux estimate*, given by the mean value of all the fluxes in the bin, and an error, given by the standard deviation. Given that the fluxes span few order of magnitudes, we consider for each bin the logarithm of the flux. For what concerns the optical/UV bins, since errors in the observed spectra are not provided at these wavelengths, we assume a 10% error on the average value of the logarithm of the flux. Furthermore, given that if a bin has few sources, the standard deviation might produce very small errors, we impose a minimum relative error for the average measure equal to 10% of the mean value for that bin.

For what concerns the comparison between our predicted SEDs and optical/NIR/FIR observations, we first remind that our simulated field of view contains a merging system, composed by three AGN, which we labelled as A, B, C (see Section 6.4.1 and Fig. 8.1) and one star-forming galaxy (source D). We use the SED of source (A) at $0.145\ \mu$m and $0.3\ \mu$m rest-frame, and the SED of source (C) at 1.18 mm[5], since they are the brightest in the UV and IR, respectively (for a detailed discussion, see Section 6.4.1). For what concerns the NIR/MIR/FIR, we have to take into account that the resolution of Spitzer and Herschel data does not allow to resolve the two dominant emitting regions[6]. Thus, for the bins from $3.6\ \mu$m to $500\ \mu$m, we compare the data with the total synthetic SED.

Finally, we compute the reduced χ^2 between the binned data and the predictions by our models. In particular, we use the following formula:

$$\chi^2 = \frac{1}{N} \sum \left(\frac{\log F_{\rm model} - F^{\log}_{\rm obs}}{\sigma^{\log}_{\rm obs}} \right)^2 , \tag{8.6}$$

where the sum is performed over the total N bins, $F_{\rm model}$ is the flux of our synthetic SEDs, $F^{\log}_{\rm obs}$ is the average flux of the logarithm of the data in a given bin as

$$F^{\log}_{\rm obs} = \langle \log F_{\rm obs} \rangle_{\rm bin}, \tag{8.7}$$

and the error considered in each bin reads as

$$\sigma^{\log}_{\rm obs} = \max(0.1 \log F_{\rm obs}, \sigma(\log F_{\rm obs})), \tag{8.8}$$

the rest-frame FIR measures by ALMA, we collect into a single bin the points with observed wavelengths between 1.1 mm and 1.3 mm. Then, we simply assume that they are all taken at the average wavelength of the sample, namely 1.18 mm. In the other cases the wavelength of the bin simply corresponds to the observed wavelength.

[4]We model the torus emission as a single-temperature greybody, with dust mass $M_{\rm torus} = 10\ {\rm M}_\odot$, $T_{\rm torus} = 1000$ K and emissivity $\beta_{\rm torus} = 2$. The contamination from the torus is then computed as the ratio between the flux from the torus component to the total (torus plus model) flux. We acknowledge that this is a simplification of the torus physical properties, whose temperature is expected to follow a power-law profile $\propto T_{\rm torus}^{-3/4}$, with a maximum corresponding to the sublimation temperature of the grains at $\approx 1500 - 1800$ K (e.g. Netzer, 2015). However, we find that the wavelength range mostly affected by the torus emission is consistent with what found in SED fitting of AGN including a torus component (e.g. Hernán-Caballero et al., 2016).

[5]The SED of an individual source is extracted from a circle of 2.5 kpc centred at the position of the source in the field of view.

[6]Source (A) is in fact only ~ 10 kpc away from source (C).

where $\sigma(\log F_{\rm obs})$ represents one standard deviation of the logarithm of the fluxes in a bin. For each model, we considered all the six lines of sight for which the RT calculation is performed and we take the lowest χ^2.

The results of the χ^2 analysis applied to our 20 models are reported in Table 8.2. We find six models with $\chi^2 \leq 1$, which are favoured by this analysis. *Among these six models, five of them have a grain size distribution cut at* $a_{\rm min} = 0.1$ μm. This is mainly a consequence of the constraints imposed by the optical/UV data, which are best explained by those models that have a grey extinction at these wavelengths. The only marginally favored ($\chi^2 = 1.00$) model without the cut $a_{\rm min} = 0.1$ μm has a MW-type intrinsic extinction curve, which produces strong PAHs features, whose presence in high-redshift quasars is still questioned[7].

Five of the six models have either a small or intermediate dust-to-metal ratio ($f_{\rm d} = 0.08$ or $f_{\rm d} = 0.15$), suggesting dust masses of $(3.3 - 6.2) \times 10^7$ M$_\odot$. In particular, the models with $f_{\rm d} = 0.15$ have the lowest χ^2, indicating that an intermediate dust content provides the best compromise between the attenuation at the shortest wavelengths and re-emission in the FIR. The best model with $f_{\rm d} = 0.3$ has $\chi^2 = 0.89$, so it is only marginally favoured, emphasizing the fact that models with high f_d lead to a stronger attenuation in the optical/UV than suggested by the data, even if a grey extinction curve is assumed.

Overall, we find no strong preference in favor of any specific dust model composition, with three models having an intrinsic SMC-type dust and three models MW-type dust. However, this is mostly due to the fact that the grain size distribution of silicates and carbonaceous grains are quite similar for the SMC and MW dust models if considering only grains with sizes larger than 0.1 μm. In particular, we underline that, in the MW models with $a_{\rm min} = 0.1$ μm, PAH molecules are effectively removed. All of the six models have the fiducial AGN SED, therefore we conclude that the UV-steep SED is disfavoured by this analysis.

In the right-hand panel of Fig. 8.4 we show the best fit model, given by $f_{\rm d} = 0.15$, the fiducial AGN SED and the cut in the grain size distribution at $a_{\rm min} = 0.1$ μm from an original SMC-type dust composition. The cut in the grain size distribution results in a grey extinction at short wavelengths, and provides the best agreement with the optical/UV data, whereas the intermediate dust content given by $f_{\rm d} = 0.15$ (i.e. dust masses of $\approx 6 \times 10^7$ M$_\odot$) gives the best compromise between UV attenuation and IR emission. We further comment on the attenuation curves in Section 8.4.2.

8.4.2 Attenuation curves in high-redshift quasars

The χ^2 analysis in Section 8.4.1 provides us with important insights concerning the dust abundance, composition and grain size distribution in $z \sim 6$ quasars. We can therefore infer an attenuation curve that correctly reproduces the current observations, according to our framework.

In the left panel of Fig. 8.5 we show the attenuation curve for our best-fit model, which has the fiducial AGN SED, $f_{\rm d} = 0.15$ and a SMC-like dust composition, with a grain size distribution cut at $a_{\rm min} = 0.1$ μm. It displays an almost flat attenuation curve, with small los-to-los variations. Compared to the SMC extinction curve, it shows a large deviation at short wavelengths, where the SMC curve steepens. The second best-fit model has the same setup

[7]The presence of PAHs in AGN hosting galaxies is unlikely both for observational and theoretical reasons. From the observational point of view, Spitzer/IRS spectra of low redshift quasars do not show PAHs features as prominent as in normal star-forming galaxies (e.g. Hernán-Caballero et al., 2016). From the theoretical side, these complex, small, fragile molecules may not survive to the intense radiation of AGN, because of Columb explosion (Tazaki et al., 2020) and/or drift-induced sputtering (Tazaki & Ichikawa, 2020).

AGN SED	$f_{\rm d}$	Dust model	$a_{\rm min}$ [μm]	χ^2
fiducial	0.08	SMC	-	1.10
fiducial	0.3	SMC	-	3.23
fiducial	0.08	MW	-	1.00
fiducial	0.3	MW	-	3.99
UV-steep	0.08	SMC	-	1.48
UV-steep	0.3	SMC	-	4.82
UV-steep	0.08	MW	-	2.34
UV-steep	0.3	MW	-	7.00
fiducial	0.08	SMC	0.01	1.06
fiducial	0.3	SMC	0.01	3.10
fiducial	0.08	SMC	0.1	0.68
fiducial	0.3	SMC	0.1	0.97
fiducial	0.08	MW	0.1	0.60
fiducial	0.3	MW	0.1	1.31
UV-steep	0.08	SMC	0.1	1.31
UV-steep	0.3	SMC	0.1	3.17
UV-steep	0.08	MW	0.1	1.71
UV-steep	0.3	MW	0.1	4.14
fiducial	0.15	SMC	0.1	0.52
fiducial	0.15	MW	0.1	0.57

Table 8.2: The set of 20 models considered in our χ^2 analysis, in order to find the best match with the observed sample. The columns indicate: (first) the AGN intrinsic SED adopted, (second) the dust-to-metal ratio, (third) the dust model as in Weingartner & Draine (2001), (fourth) the minimum grain size in the modified grain size distribution (if it is not specified, no cut is applied, and the standard grain size distribution is considered), (fifth) the χ^2 for the model.

of the best one, except with a MW-like dust composition. However, the results are almost the same, because the modified grain size distribution in these two models result in a very similar dust extinction curve. The corresponding attenuation curve is shown in the right panel of Fig. 8.5. We note that in the MW-derived model the characteristic bump at 2175 Åis absent because both the PAH molecules and small graphite grains, which contribute to the feature (e.g. Li & Draine, 2001b), are removed by the cut at $a_{\rm min} = 0.1$ μm. The attenuation curves of these two models predicted by our calculation are slightly flatter than the Calzetti et al. (1994) one, and they are in excellent agreement with the average attenuation[8] curve deduced in G10, with a small

[8]We note that G10 refer to their results in terms of *extinction* curve. This notation is commonly used in the case of quasars, typically represented as a point source plus a foreground screen geometry. However, in this work we find that the attenuation curve substantially differs from the extinction curve even in the case of AGN (see for example Fig. 8.2 and Fig. 8.5). This implies that a simplified geometry cannot fairly represent quasars for at least two reasons: 1) emission from stars cannot be neglected in these systems since quasar host galaxies are typically highly star-forming ($SFR \sim 10^2 - 10^3$ M$_\odot$ yr^{-1}); 2) dust distribution in quasar host galaxies is shaped by quasar feedback and presents a complex, disturbed morphology. In light of these results, the curves found by G10 from their fitting procedure should be interpreted as *attenuation* curves, and not as extinction curves.

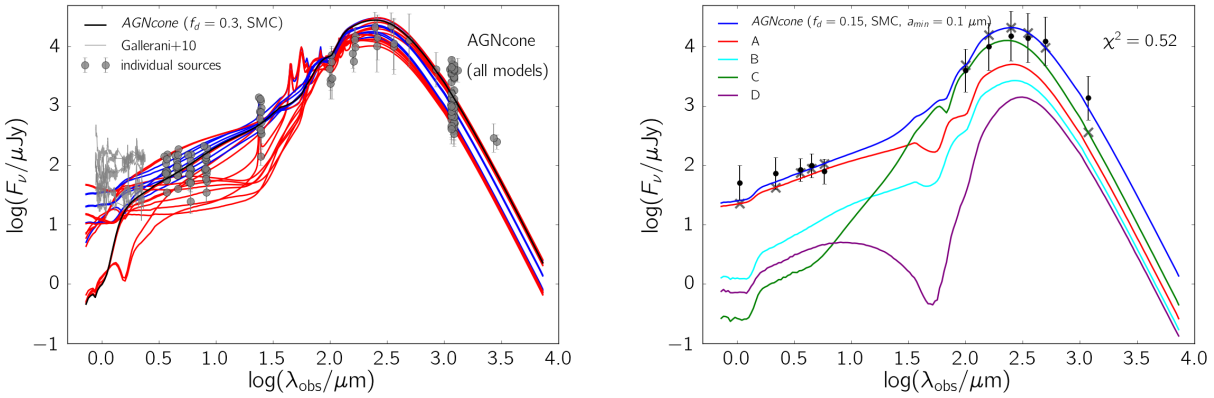

Figure 8.4: *Left:* Comparison of the *total* SEDs of the 20 models considered in the χ^2-analysis (coloured lines) with UV-to-FIR data. The model $f_{\rm d} = 0.3$, SMC-type dust and the fiducial AGN SED, considered as fiducial in DM21a, is plotted with a black solid line. We also color-code each model according to the value of χ^2: models with $\chi^2 > 1$ are shown in red, models with $\chi^2 \leq 1$ in blue. The photometric data of the sources considered are shown with grey points, redshifted at the redshift of the simulated snapshot. The optical/UV spectra for the $z > 6$ quasars analysed in G10 are shown with grey lines. *Right:* Synthetic SEDs of the *individual* sources for the best fit model, which has $f_{\rm d} = 0.15$, fiducial AGN SED, and SMC dust composition modified with $a_{\rm min} = 0.1$ μm. The color legend for each source is the same as in Fig. 8.1, and the SED of the whole field of view is shown in blue. In particular, the SED of the UV-bright source (A) is plotted in red, while the one of the IR-bright source (C) in green. We mark with grey crosses the fluxes of our synthetic SEDs used for the χ^2 computation. We also show the observed data grouped into eleven bins (see text for more details) with black circles. The χ^2 of the best fit model is reported in the top right corner.

deviation only at the shortest wavelengths, where our curves are flatter.

We underline that the attenuation curves of our two best models shown in Fig. 8.5 are derived by collecting the radiation of the whole field of view. However, for a fair comparison with the observations of optical/UV-selected quasars, it might be more appropriate to investigate the attenuation curve relative to the brightest UV source in the field of view. We checked that the attenuation curve inferred by considering only the radiation from source A differ by less than 10 % from the attenuation curve derived by considering the whole field of view, shown in in Fig. 8.5.

8.5 Discussion

8.5.1 Implications

Our calculations support a scenario in which dust in AGN-hosting galaxies is dominated by large grains (i.e. $a_{\rm min} = 0.1$ μm). This can either be explained by very efficient coagulation of small grains into larger ones in the dense molecular clouds of quasar-hosts (see e.g Nozawa et al., 2015) or by physical mechanisms that preferentially destroy small grains in the extreme environments around AGN. Recently, two processes have been studied in this context. First, dust destruction by charging can wipe out small grains due to the intense radiation field in the AGN proximity (Tazaki et al., 2020). Second, grains might be destroyed by drift-induced sputtering while they are moved from the equatorial plane toward the polar region by the radiation pressure from the AGN. This mechanism is more effective at destroying small grains than large grains, effectively selecting the dust component that is directly irradiated by the central source and that is responsible of the extinction (Tazaki & Ichikawa, 2020). The two processes might also act at the same time.

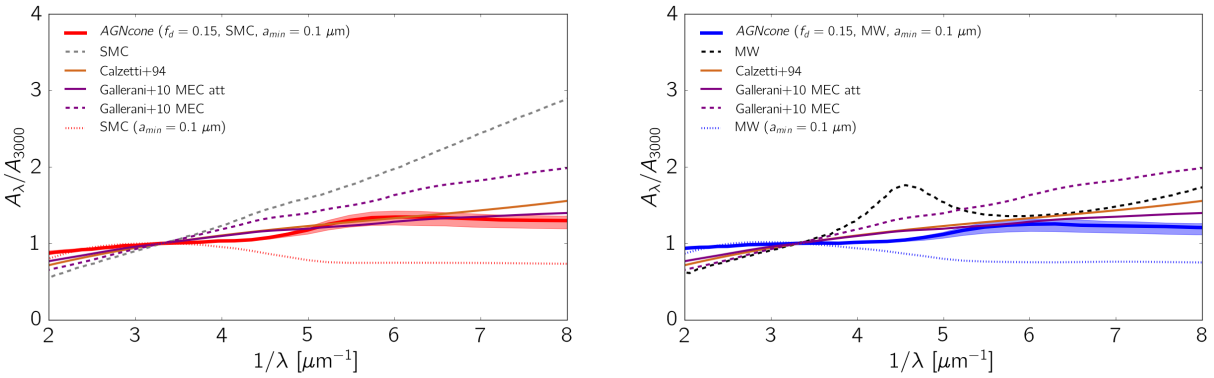

Figure 8.5: Attenuation curves for the best-fit models, characterised by $f_{\rm d} = 0.15$, fiducial AGN SED, $a_{\rm min} = 0.1$ μm and SMC (left panel) or MW (right panel) original dust composition. As a reference, we also plot the SMC extinction curve (grey dashed line), in the left panel, whereas we show the MW extinction curve (black dashed line) in the right panel. We also plot the Calzetti attenuation curve (light brown solid line, Calzetti et al., 1994), and the extinction curve (MEC, purple dashed line) and attenuation curve (MEC attenuated, purple solid line) obtained in G10 for high-redshift quasars. Finally, we show the intrinsic extinction curve for the modified grain size distribution used in the SMC (red dotted line) and MW (blue dotted line) case.

The framework used in this work assumes a linear scaling relation between dust and metals; furthermore, the same grain size distribution is assumed for all the dust in the field of view. Therefore, we cannot discriminate between a scenario in which small grains are absent in most of the ISM of AGN-host galaxies, or if the grain size distribution supported by our analysis reflects only the properties of a dust component located in the proximity of the AGN. An accurate description of the dust content and the grain size distribution would require to follow in detail the processes involved in the grains formation and destruction, whose efficiency depends on the local gas physical conditions, such as density, temperature, metallicity (see e.g. Hirashita & Aoyama, 2019; Aoyama et al., 2020). This detailed treatment of the dust physics is beyond the scope of our work.

Moreover, if large grains are dominant because of dust production mechanisms in place at high redshift, this evidence further supports the role of SNe as the main dust factories in the early Universe (Todini & Ferrara, 2001). A SN-origin for dust at high redshift was also suggested by the extinction curve of J1048+4637, measured by Maiolino et al. (2004), which flattens at $\lambda > 1700$ Å and then rises again at $\lambda < 1700$ Å. This evidence suggests that, even though we find attenuation curves in AGN to be flatter than the SMC on average (G10, as also found in), we cannot exclude the presence of a minor component of small grains, which can be unveiled by studying individually each source. We plan to improve our work in the future in this direction, and to include also models of dust grains from SN origin in our RT calculations. Furthermore, in order to reproduce the extinction curve of J1048+4637 measured by Maiolino et al. (2004), Nozawa et al. (2015) had to assume in their dust models that carbonaceous grains were in the form of amorphous carbon and not graphite, which allowed them to remove the 2175 Å bump. We also plan to consider models with amorphous carbon in a future work.

8.5.2 Effective dust-to-metal ratio in presence of efficient dust destruction

The models favored by our χ^2-analysis in Section 8.4.1 all lack small grains. In practice, their grain size distribution has the same functional form of the original model used (i.e. SMC or MW) with a cut at $a_{\rm min} = 0.1$ μm. Instead, the dust mass considered in the simulations for these models, which is inferred from the dust-to-metal ratio (equation 8.1), is the same as for the models with the standard grain size distribution, i.e. it is simply re-distributed among larger

grains.

In this Section, we assume that the dominance of large grains is the result of some destruction process in AGN environments (see e.g. Tazaki et al., 2020; Tazaki & Ichikawa, 2020). Under this hypothesis, the effective dust production in the AGN-hosting galaxy is higher than if destruction is not accounted for. Therefore, the *effective* dust-to-metal ratio (i.e. before accounting for the removal of small grains) is actually higher, and this discrepancy depends on the fraction of grains destroyed. In the following, we quantify the factor that relates the assumed dust-to-metal ratio with the effective one for the models with $a_{\rm min} = 0.1\ \mu$m, under the assumption that small grains are destroyed.

Given a grain size distribution dn/da, which extends from a minimum grain size $a_{\rm min}$ to a maximum grain size $a_{\rm max}$, the total dust mass can be expressed as:

$$M_{\rm d}^{\rm total} = \frac{M_{\rm gas}}{\overline{\mu} m_{\rm p}} \int_{a_{\rm min}}^{a_{\rm max}} \frac{4}{3} \pi \delta_{\rm g} a^3 \left(\frac{dn}{da} \right) da, \tag{8.9}$$

where $\overline{\mu}$ is the average molecular weight of the gas, $m_{\rm p}$ is the proton mass, and $\delta_{\rm g}$ represents the bulk density of the grain, which we assume to be $\delta_{\rm carb} = 2.24\ {\rm g\,cm^{-3}}$ for carbonaceous grains and $\delta_{\rm sil} = 3.5\ {\rm g\,cm^{-3}}$ for silicate grains as in Weingartner & Draine (2001). The total mass fraction of grains smaller than $0.1\ \mu$m for a fixed species is then given by the ratio:

$$\begin{aligned} \frac{M_{\rm d}^{\rm small}}{M_{\rm d}^{\rm total}} &= \frac{\frac{M_{\rm gas}}{\overline{\mu} m_{\rm p}} \int_{a_{\rm min}}^{0.1\mu{\rm m}} \frac{4}{3} \pi \delta_{\rm g} a^3 \left(\frac{dn}{da} \right) da}{\frac{M_{\rm gas}}{\overline{\mu} m_{\rm p}} \int_{a_{\rm min}}^{a_{\rm max}} \frac{4}{3} \pi \delta_{\rm g} a^3 \left(\frac{dn}{da} \right) da} = \frac{\int_{a_{\rm min}}^{0.1\mu{\rm m}} a^3 \left(\frac{dn}{da} \right) da}{\int_{a_{\rm min}}^{a_{\rm max}} a^3 \left(\frac{dn}{da} \right) da} = \\ &= \frac{\int_{a_{\rm min}}^{0.1\mu{\rm m}} a^3 \left(\frac{dn}{da} \right) da}{\int_{a_{\rm min}}^{0.1\mu{\rm m}} a^3 \left(\frac{dn}{da} \right) da + \int_{0.1\mu{\rm m}}^{a_{\rm max}} a^3 \left(\frac{dn}{da} \right) da} = \\ &= \frac{\mathcal{R}_{\rm small}}{\mathcal{R}_{\rm small} + \mathcal{R}_{\rm large}} = \frac{\mathcal{R}_{\rm small}}{\mathcal{R}_{\rm total}}, \end{aligned}$$

where we introduced the *reduced mass* of small and large grains $\mathcal{R}_{\rm small}$ and $\mathcal{R}_{\rm large}$ respectively, which do not depend on the bulk density. We also indicate $\mathcal{R}_{\rm total} = \mathcal{R}_{\rm small} + \mathcal{R}_{\rm large}$ for a single species. In SKIRT the SMC and MW dust models are implemented according to the model by Weingartner & Draine (2001)[9]. By using the appropriate values for the SMC model, we get:

$$\frac{\mathcal{R}_{\rm small}^{\rm SMC,sil}}{\mathcal{R}_{\rm total}^{\rm SMC,sil}} = 0.62, \tag{8.10}$$

$$\frac{\mathcal{R}_{\rm small}^{\rm SMC,carb}}{\mathcal{R}_{\rm total}^{\rm SMC,carb}} = 0.20. \tag{8.11}$$

Thus, in the SMC models without grains $\lesssim 0.1\ \mu$m, 62% of the dust mass in silicates and 20% of the dust mass in carbonaceous grains of the original grain size distribution is removed. These results are summarised in Fig. 8.6. Overall, the amount of dust mass in small grains (i.e. removed) is:

$$\frac{M_{\rm d}^{\rm SMC,small}}{M_{\rm d}^{\rm SMC,total}} = \frac{\delta_{\rm sil}\, \mathcal{R}_{\rm small}^{\rm SMC,sil} + \delta_{\rm carb}\, \mathcal{R}_{\rm small}^{\rm SMC,carb}}{\delta_{\rm sil}\, \mathcal{R}_{\rm total}^{\rm SMC,sil} + \delta_{\rm carb}\, \mathcal{R}_{\rm total}^{\rm SMC,carb}} = 0.59. \tag{8.12}$$

Therefore, the overall dust mass formed, i.e. without considering destruction of small grains, is a factor $1/(1-0.59) \approx 2.5$ higher. This implies that, for a given value of the dust-to-metal

[9]Grain size distribution as in equations 4, 5 and 6, coefficients from their Table 1 and Table 3.

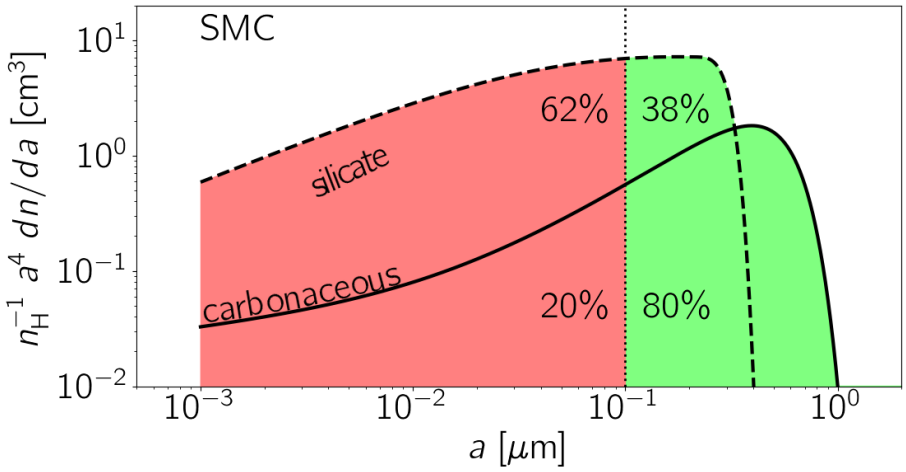

Figure 8.6: Grain size distributions for the SMC, according to the Weingartner & Draine (2001) model. The solid line refers to carbonaceous grains and the dashed line to silicates. The red regions indicate the grains removed in our models without small grains, i.e. $a_{\rm min} = 0.1\ \mu$m, while the green regions the grains kept. The percentage quoted refer to the mass fraction of the corresponding grains removed (when the number is in the red region) or kept (green region), for silicate (upper values, close to the dashed line) or carbonaceous (lower values, close to the solid line).

ratio $f_{\rm d}$, an SMC model without small grains has an effective dust-to-metal ratio of $f_{\rm d}^{\rm eff} \approx 2.5 f_{\rm d}$. For the best-fitting model found in Section 8.4.1, which has $f_{\rm d} = 0.15$, we derive $f_{\rm d}^{\rm eff} = 0.38$, which is very similar to the Milky-Way value. This might indicate that dust production at high-redshift was already quite efficient in AGN-hosting galaxies. The calculation illustrated above also implies that there is a maximum dust-to-metal ratio for which we can assume the existence of a physical process responsible for the selective removal of grains smaller than $0.1\ \mu$m, which is $f_{\rm d}^{\rm max} = 0.4$.

8.5.3 AGN feedback and dust spatial distribution

In Section 8.3, we studied how the shape of the attenuation curves of a simulated galaxy depends, among many other factors, on the dust-sources geometry. In particular, we find that the attenuation curves in AGN-host galaxies tend to be steeper than in normal star-forming galaxies because of lower dust densities. We argue that this behaviour is a consequence of the AGN feedback, which drives powerful outflows, pushing away the gas from the BH surrounding and outside the galaxies, thus reducing the overall gas density. This finding motivates the need for assuming a peculiar grain size distribution in order to reproduce the grey attenuation curves suggested by optical/UV data. We are able to match the observations if grains smaller than $a_{\rm min} = 0.1\ \mu$m are removed (see Section 8.4.1).

However, another possibility to ease the tension with the data is that the dust distribution in observed quasars is actually much more compact than what predicted by the hydro-simulations considered in this work. For example, it is possible that the kinetic feedback implemented is too effective in powering winds, and that other prescriptions, such as thermal feedback, might instead be better suited to reproduce the observations. If that is the case, then the higher density would also result in higher dust attenuation. Therefore, in order to investigate this point, we consider the absolute dust attenuation at 3000 Å, A_{3000}, of our best-model and we compare its value with the observations. The six quasars with optical/UV measurements used in Section

8.4.1 have $A_{3000} = 0.82 - 2.00$ (G10). Our best-fit model has $A_{3000} = 1.32$, which is nicely consistent with these results. Dust surface densities an order of magnitude higher (required to mimic a grey extinction curve in *AGNcone*, see Section 8.3) would result in an attenuation of a factor of ≈ 10 higher than what predicted, which is too high with respect to the observations. Furthermore, with a higher dust attenuation, a higher intrinsic magnitude would also be required in order to match the far-UV flux of observations. Given that in our framework the bolometric luminosity of an AGN is directly linked with its accretion rate (see Section 4.3), this would likely imply super-Eddington accretion. Therefore, we argue that the gas distribution predicted by the hydrodynamic simulations adopted is consistent with observations. Nevertheless, we plan to investigate attenuation curves in simulated galaxies with different feedback prescriptions in a future work.

8.5.4 Caveats

In this work, we consider the snapshot at $z = 6.3$ of the run *AGNcone* performed in B18, and we post-process it with RT calculations in order to compare it with the observations of $z \sim 6$ quasars. It is natural to question how representative the chosen DM halo and the specific snapshot we selected are of the quasars sample we compare with. In the low resolution, DM-only simulation performed in B18 the selected DM halo has a mass of $M_{\rm halo} = 4.4{\times}10^{12}\,{\rm M}_{\odot}$ at $z = 6$, which is compatible with results from clustering studies (e.g. Allevato et al., 2016). However, this DM halo experiences a specific merging and accretion history and cannot be representative of the whole population of high-redshift quasars. Moreover, the evolution of the halo and its host galaxies is also affected by the numerical model implemented, in particular on the AGN feedback, the subgrid physics adopted and the resolution limits. As such, our results might be sensitive to the structure of the simulated halo, and to the properties down to ISM scales, which are ultimately affected by the details of the numerical models implemented. It would be valuable to repeat this analysis by exploring several realizations of the Universe via cosmological hydrodynamical simulations, by considering a wide range of DM halo masses, histories, and the parameters describing the AGN feedback and the subgrid physics. This is beyond the scope of this work.

Moreover, in the analysis performed in Section 8.4.1, we compare our simulated AGN with the average properties of the $z \sim 6$ quasars. It would be valuable to also perform this analysis by considering the observed quasars individually, and finding for each of them a simulated object that best matches their SEDs. However, there is only handful of objects with a SED sampled well enough for such calculation. Furthermore, a very high number of simulations would be required for this kind of study.

The snapshot at $z = 6.3$ we focus on was already analysed in DM21a, where we show it matches reasonably well the average properties of high-redshift quasars, underestimating only the UV emission. In particular, this system is consistent with recent observations of quasars-SMGs companions (e.g. Decarli et al., 2017). However, three AGN are active (see Section 6.4.1) at this snapshot, thus it might represent a peculiar moment in the DM halo evolution. As a check, we post-process additional snapshots (see Section A.4), when a different number of AGN is active. We find that in all cases the dust-attenuation in the UV make our simulated objects under-luminous with respect to the observations, confirming that the need of less extinction at short wavelength is not special of the snapshot selected for the detailed analysis in this work.

As discussed in Section 4.3, the simulations considered in this work achieve a maximum spatial resolution of ≈ 200 pc, and a maximum density of $\approx 200\ {\rm cm}^{-3}$. Therefore, we are not able to resolve dense, clumpy molecular structures on $\sim$ pc scales, which might dominate the

extinction with respect to the diffuse gas in the ISM. Even the highest resolution cosmological simulations do not reach sub-pc scale-resolution, i.e. cannot follow in detail all the physical processes inside such clumpy molecular clouds. Resolving such structures might decrease the effective opacity of the medium, because of their low volume filling factor, as shown in radiative transfer simulations in two-phases clumpy media (e.g. Witt & Gordon, 1996; Bianchi et al., 2000). This might in turn imply a higher dust mass than suggested by our simulations in order to match the observational data. However, this effect is highly dependent on the ISM properties, for example Decataldo et al. (2020) find that only 10% of the UV flux can escape such clumpy structures. While dedicated works will be needed to fully explore the subject, the predicted attenuation curve steepening should only be marginally affected by such differential variation. Thus we expect our main conclusion that small grains are removed in AGN environment to still hold.

8.6 Summary and conclusions

We have studied dust attenuation in $z \sim 6$ quasar-host galaxies by post-processing with radiative transfer (RT) calculations a suite of cosmological zoom-in hydrodynamical simulations. Our main goal has been to clarify whether extinction curves such as those found in the the Small Magellanic Cloud (SMC) or in the Milky Way (MW) can explain multi-wavelength observations of high-redshift quasars spectra. Using the cosmological simulations presented in B18, in which the evolution of a $\sim 10^{12}$ $\mathrm{M}_{\odot}$ dark matter halo at $z \simeq 6$ is followed in a zoom-in fashion, we focus on the snapshot at $z = 6.3$ of the run *AGNcone*, previously analyzed in detail in DM21a. At this evolutionary stage, the simulated halo hosts three galaxies with active black holes at their centres, corresponding to an (unattenuated) UV magnitude of $M_{\mathrm{UV}} = -28$. We have post-processed the snapshot with the code SKIRT (Baes et al., 2003; Baes & Camps, 2015; Camps & Baes, 2015; Camps et al., 2016), which solves the radiative transfer problem in dusty systems. We considered different models for the overall dust mass content, dust composition, grain size distribution, and intrinsic AGN spectrum. We have first examined the effects of relative dust-sources geometry on the synthetic attenuation curves by comparing the results of *AGNcone* with a control run without BHs, namely *noAGN*. We find that, at fixed dust composition, *noAGN* shows flatter attenuation curves and smaller los-to-los variations than *AGNcone* (Fig. 8.2).

We quantify the steepness of the attenuation curves by the ratio $A_{0.1}/A_1$, which compares the total attenuation at 0.1 μm and 1 μm respectively. We provide an analytical model that relates this ratio with the V-band optical depth, for two simplified geometries: i) a point source surrounded by a spherical distribution of dust (*point source model*) and ii) a sphere with dust and emitters uniformly distributed (*intermixed model*). In both cases, the ratio $A_{0.1}/A_1$ decreases for $\tau_{\mathrm{V}} \gg 1$, and in the intermixed model it goes to 1, thereby resulting in a flat attenuation curve. The sphere model well describes the compact dust/stars distribution in *noAGN* run, and the low $A_{0.1}/A_1$ ratio predicted at high optical depths is consistent with the most (intrinsic) UV-emitting regions having $\tau_{\mathrm{V}} > 10$ in this run (Fig. 8.3, right panel). In the *AGNcone* run, they are characterised by $1 < \tau_{\mathrm{V}} < 10$, and the value of the $A_{0.1}/A_1$ ratio lies in the middle of the results of the point source and intermixed models. Overall, the fact that the attenuation curve in *noAGN* is flatter than in *AGNcone* is well explained by the higher dust optical depths. The different optical depths, i.e. dust distribution, among the two runs is a consequence of the AGN feedback, which drives powerful outflows that remove gas from the central regions, and distribute it over several kpc. We conclude that the steep synthetic attenuation curves predicted for the simulated AGN-host galaxies in *AGNcone* are a direct consequence of a gas distribution affected by AGN activity.

We then compare the synthetic Spectral Energy Distributions (SEDs) obtained by RT post-

processing *AGNcone* with multi-wavelength observations of a sample of bright $z \sim 6$ quasars with comparable UV magnitudes ($-29 \lesssim M_{\rm UV} \lesssim -26$). Among the twenty models considered, only six are favoured by a χ^2-analysis (Table 8.2 and Figure 8.4). Five of these six models have a modified grain size distribution, obtained from the one underlying the original dust model (SMC-like or MW-like) by removing grains with sizes $a < a_{\rm min} = 0.1$ μm. The attenuation curves inferred for these models are close to flat (Fig. 8.5). These findings are consistent with the results by G10, who fit the spectra of a sample of $3.9 \leq z \leq 6.4$ reddened quasars, suggesting extinction curves flatter than the SMC.

The standard dust models tend to produce attenuation curves that are too steep to match the data, because of the low optical depths in AGN-host galaxies caused by AGN activity. A modified grain size distribution is needed in order to reconcile the synthetic SEDs with the optical/UV data. Therefore we caution toward the applicability of the SMC extinction law in high-z quasar-hosts when interpreting observations.

Our calculations finally suggest a dust-to-metal ratio $f_{\rm d} \lesssim 0.15$, which implies dust masses $M_{\rm d} \lesssim 6 \times 10^7$ M$_\odot$ for the AGN-hosts in the sample. However, if we attribute the modified grain size distribution to efficient destruction processes in place in AGN-environments (e.g. Tazaki & Ichikawa, 2020; Tazaki et al., 2020), the original dust mass produced might instead be a factor ≈ 2.5 higher, thus implying an effective dust-to-metal ratio of ≈ 0.4 before accounting for dust removal (Fig. 8.6). This would suggest very efficient dust production in high-redshift AGN-hosting galaxies. Alternatively, the dominance of large grains in AGN-hosts might suggest a supernova origin for dust at high redshift (e.g. Todini & Ferrara, 2001) or efficient coagulation of small grains into larger ones (see e.g Nozawa et al., 2015). Which of the two scenarios applies, i.e. efficient destruction of small grains vs. preferential production of large grains, cannot be assessed from our calculations, because of limited spatial resolution and the lack of detailed physics in the hydrodynamical simulations. A framework describing the dust content and grain size evolution according to the local gas physical conditions (e.g. density, temperature, metallicity) in the galaxy, paired with a higher spatial resolution, is required to disentangle the two scenarios. These improvements are left for a future work.

The upcoming James Webb Space Telescope (JWST) will significantly improve current rest-frame optical/NIR data with deeper observations and it will provide us with high-resolution spectra, effectively probing the spectral region more sensitive to dust attenuation with an unprecedented sensitivity. This will drive a significant contribution in our understanding of the dust origin in the early Universe.

Feedback effect on the observable properties of $z > 6$ AGN

9

9.1 Introduction

In Section 2.5 we discussed how the existence of super-massive black holes (SMBHs) with masses larger than billion solar masses at $z \gtrsim 6$ (e.g., Bañados et al., 2018; Matsuoka et al., 2018b; Wang et al., 2021a), when the Universe was < 1 Gyr old, challenges our current understanding of SMBH and galaxy formation and evolution, and is thus one of the most pressing open issues in modern astrophysics (e.g., Woods et al., 2019). Their distance and faintness make observations of these objects difficult and strongly biased towards the most luminous and massive accreting SMBHs. A complementary approach is to use numerical simulations as tools to study the largely unknown phases of SMBH growth in the early Universe (e.g. Tanaka & Haiman, 2009; Sijacki et al., 2009; Habouzit et al., 2016, 2019). However, observed properties of high-redshift accreting SMBHs, or active galactic nuclei (AGN), and predictions of numerical simulations have been compared only seldom (e.g. Ni et al., 2020; Habouzit et al., 2021; Di Mascia et al., 2021a; Zana et al., 2022).

An important ingredient entering in numerical simulations focused on the early growth of SMBHs is the effect of AGN feedback (e.g. Costa et al., 2014a, 2020; Barai et al., 2018; Habouzit et al., 2019; Valentini et al., 2021), as it is often considered to have a major role in shaping the evolution of AGN and galaxies along the whole cosmic history (e.g., Fiore et al., 2017). In particular, optically-selected luminous quasi-stellar objects (QSOs) in the early Universe often present evidence for the launching of fast and massive multi-phase outflows (e.g., Maiolino et al. 2012; Cicone et al. 2015; Bischetti et al. 2019; Carniani et al. 2019a; Schindler et al. 2020; Izumi et al. 2021; but see also, e.g., Decarli et al. 2018; Novak et al. 2020; Meyer et al. 2022), which are expected to affect the observable properties of the QSOs themselves and their host galaxies, such as X-ray oscuration, UV extinction, and gas content (e.g., Brusa et al., 2015; Ni et al., 2020).

Outflows observed in QSOs are though to originate from fast nuclear winds, which, in turn, may be accelerated by several physical mechanisms, including radiation pressure, due to UV photons produced in the accretion disc, on dust grains or on partially ionized gas mediated by UV transitions, and magnetic effects (e.g. Proga et al., 2000; Murray et al., 2005; Fabian et al., 2008; Yuan et al., 2015; Ricci et al., 2017b). The physical scales involved in these processes are those of the accretion disk (e.g. Giustini & Proga, 2019). Since such scales cannot be resolved by large-scale cosmological simulations, different authors have modeled AGN feedback using several different recipes (e.g., Barai et al. 2018; Costa et al. 2020; Ni et al. 2020). Moreover, the effect of the outflow on the surrounding material can potentially depend on its geometry (e.g. Zubovas et al., 2016). Since the exact acceleration physics, and thus launching direction, of

nuclear winds is not well understood, numerical simulations typically assume either spherical (e.g., Feng et al., 2016) or bi-conical (e.g., Sala et al., 2021) outflow geometry as study cases.

Beside the properties of the individual galaxies hosting accreting SMBHs, numerical simulations provide also information on the environment of high-redshift luminous AGN. While these objects are expected to reside in the peaks of the dark matter halo distribution, which are generally characterized by large overdensities of galaxies (e.g., Costa et al. 2014a; Wise et al. 2019), although with some scatter (e.g., Habouzit et al. 2019), observations struggle to provide us with a clear view of typical high-redshift QSO environment. In fact, $z > 6$ QSOs have been reported to reside in a variety of environments, including underdense, normal, and overdense regions (e.g. Ota et al. 2018; Mazzucchelli et al. 2019; Overzier 2021). The first spectroscopically confirmed galaxy overdensity around a $z > 6$ QSOs was presented recently by Mignoli et al. (2020), followed by a tentative confirmation of another structure by Overzier (2021).

A significant fraction ($\approx 40\%$) of $z \gtrsim 6$ QSOs has ALMA-detected dusty companion galaxies at distances of a few kpc (e.g. Willott et al., 2017; Decarli et al., 2018; Neeleman et al., 2019; Venemans et al., 2020). These satellite galaxies might host heavily reddened and buried AGN (see for example Section 6.4.1), although currently there is no strong observational evidence for the presence of accreting SMBHs in their centres (e.g., Connor et al., 2019, 2020; Vito et al., 2019a, 2021). Such objects would be typically brighter than inactive galaxies, expecially in the X-ray band. Therefore, their predicted number in numerical simulations can be tested against observational results to infer how well simulations approximate reality.

In this Chapter, we present a study of the effect of AGN kinetic feedback on the observable properties of $z > 6$ AGN in cosmological simulations. In particular, we analyse a set of numerical simulations presented by Barai et al. (2018) with different kinetic feedback prescriptions, focusing on the most massive SMBH at $z = 6$ and its surrounding environment. We extract multiwavelength observables such as column density and radial extent of the gas distributed in the host galaxies, UV and X-ray AGN fluxes, and number of satellite AGN detectable over small (i.e., a few kpc) distances from the central SMBH. We compare these properties with results from multiwavelength observations.

The Chapter is structured as follows. In Section 9.2 we describe the numerical setup of the simulations, the AGN selection, and the method used to measure the gas column density and distribution. In Section 9.3 we discuss the redshift evolution of the column densities for the considered AGN. In Section 9.4 we present the observable properties of the simulated AGN and their host galaxies, and we compare them with empirical findings. In Section 9.5 we investigate the presence of multiple AGN systems over scales of a few kpc, and we compare their detectability rates in the X-ray band with results from observations of high-redshift AGN. Finally, in Section 9.6 we discuss and interpret the results, and in Section 9.7 we provide a summary.

9.2 Method

9.2.1 Numerical model

We consider the simulation runs *AGNcone* and *AGNsphere* by B18, which include kinetic feedback. We provide here a summary of the numerical setup and refer to the original works for an in-depth discussion.

B18 used a modified version of the Smooth Particle Hydrodynamics (SPH) N-body code GADGET-3 (Springel, 2005) to follow the evolution of a comoving volume of $(500\ \mathrm{Mpc})^3$, starting from cosmological initial condition generated with MUSIC (Hahn & Abel, 2011) at $z = 100$, and zooming-in on the most massive (i.e., $4 \times 10^{12}\ \mathrm{M}_\odot$) dark matter (DM) halo, corresponding to a

$\approx 3\sigma$ overdensity (e.g., Barkana & Loeb, 2001), inside the box down to $z = 6$. Therefore, the final zoomed-in simulations focus by construction on a highly biased cubic region, with a volume of (5.21 Mpc)^3. The highest level of the simulation has a mass resolution of $m_{\rm DM} = 7.54 \times 10^6$ $\rm M_\odot$ and $m_{\rm gas} = 1.41 \times 10^6$ $\rm M_\odot$ for DM and gas particles, respectively. The softening length for gravitational forces for these high-resolution DM and gas particles is $R_{\rm soft} = 1h^{-1}$ ckpc.

The code accounts for gas heating and cooling (including metal-line cooling) depending on the gas metal content, based on eleven element species (H, He, C, Ca, O, N, Ne, Mg, S, Si, Fe) that are tracked in the simulation (Tornatore et al., 2007). Star formation in the inter-stellar medium (ISM) is implemented following the multiphase effective subresolution model by Springel & Hernquist (2003), adopting a density threshold for star formation of $n_{SF} = 0.13$ cm^{-3}. The simulations include stellar winds, supernovae feedback, and metal enrichment, and assume a Chabrier (2003) initial mass function in the mass range $0.1 - 100$ $\rm M_\odot$ (Tornatore et al., 2007; Barai et al., 2013; Biffi et al., 2016).

When a DM halo that is not already hosting a black hole (BH) reaches a total mass of $M_{\rm h} \geq 10^9$ $\rm M_\odot$, a $M_{\rm BH} = 10^5$ $\rm M_\odot$ BH is seeded at its centre. BHs are treated as collisionless sink particles and are allowed to grow by accretion of the surrounding gas or by mergers with other BHs. Gas accretion onto BHs is modelled via the classical Bondi-Hoyle-Littleton accretion rate $\dot{M}_{\rm Bondi}$ (Hoyle & Lyttleton, 1939; Bondi & Hoyle, 1944; Bondi, 1952), capped at the Eddington rate $\dot{M}_{\rm Edd}$:

$$\dot{M}_{BH} = \min(\dot{M}_{\rm Bondi}, \dot{M}_{\rm Edd}). \tag{9.1}$$

Accreting BH radiate away a fraction $\epsilon_{\rm r}$ of the accreted rest-mass energy, with a bolometric luminosity

$$L_{\rm bol} = \epsilon_{\rm r} \dot{M}_{\rm BH} c^2, \tag{9.2}$$

where c is the speed of light. B18 fixed the radiative efficiency to $\epsilon_{\rm r} = 0.1$, a fiducial value for radiatively efficient, geometrically thin, optically thick accretion disks around a Schwarzschild BH (Shakura & Sunyaev, 1973).

A fraction $\epsilon_{\rm f} = 0.05$ of the total output energy is distributed to the surrounding gas in a kinetic form[1]. In *AGNcone* the kinetic energy is distributed along two cones with a half-opening angle of $45°$. The direction of the cone axis is chosen randomly for each BH at the seeding time, and is kept fixed throughout the simulation (Barai et al., 2018), similarly to what is done in Zubovas et al. (2016). Instead, the AGN feedback in *AGNsphere* pushes away the gas particles along random directions, thus mimicking a spherical geometry.

9.2.2 AGN selection

We analyse the simulation snapshots in steps of $\Delta z = 0.2$ from $z = 10$ to $z = 8$ and $\Delta z = 0.1$ from $z = 8$ to $z = 6$. In particular, we follow the most massive SMBH at $z = 6$ in each simulation set, and consider a box with side size of 60 kpc centred on it. We refer to all of the SMBHs in the box accreting at $\dot{M}_{BH} > 0.02\,\rm M_\odot\,yr^{-1}$ (i.e., $L_{bol} \approx 10^{44}$ erg s^{-1}) as AGN. Fig. 9.1 presents the BH mass evolution of AGN in the two simulations. Each AGN is labelled with the initial letter of the run (C for *AGNcone*, S for *AGNsphere*).

AGNcone forms two very massive ($> 10^9\,M_\odot$) BHs at $z < 7$, while only less massive BHs are formed in the *AGNsphere* run. This behaviour is linked to the implementation of the feedback: *AGNcone* allows the gas to accrete continuously along the equatorial directions, while the lack of a preferential direction along which the outflow is launched in *AGNsphere* does not allow

[1] We refer to B18 for details about the choice of the value for $\epsilon_{\rm f}$ and the numerical implementation of the kinetic feedback.

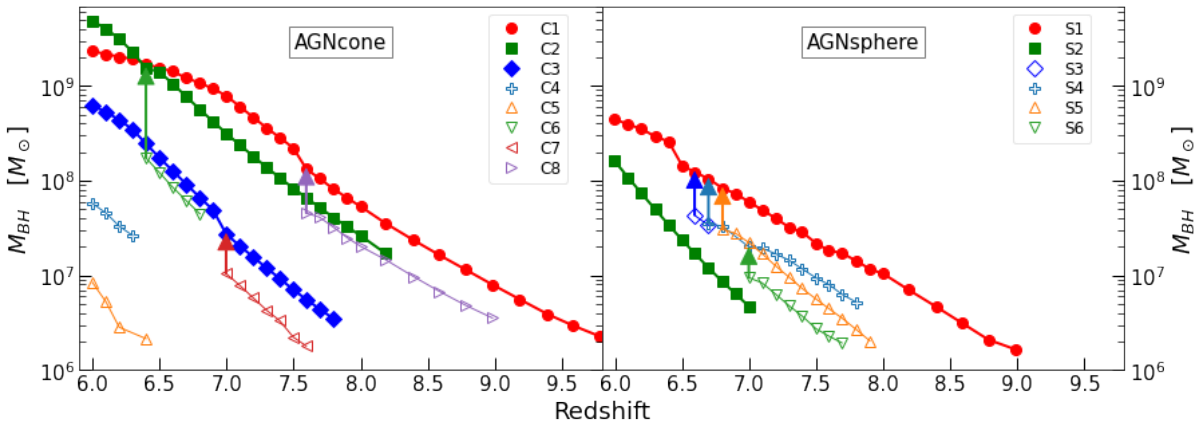

Figure 9.1: BH masses as a function of redshift for the *AGNcone* (left) and *AGNsphere* (right) runs. Only SMBHs accreting at $\dot{M} > 0.02\,\mathrm{M}_\odot$ are considered. The arrows mark the mergers between BHs. AGN considered in Section 9.3 and Section 9.4 (i.e., those that reach $z = 6$ with $M_{BH} > 10^8\,\mathrm{M}_\odot$) are plotted as filled symbols.

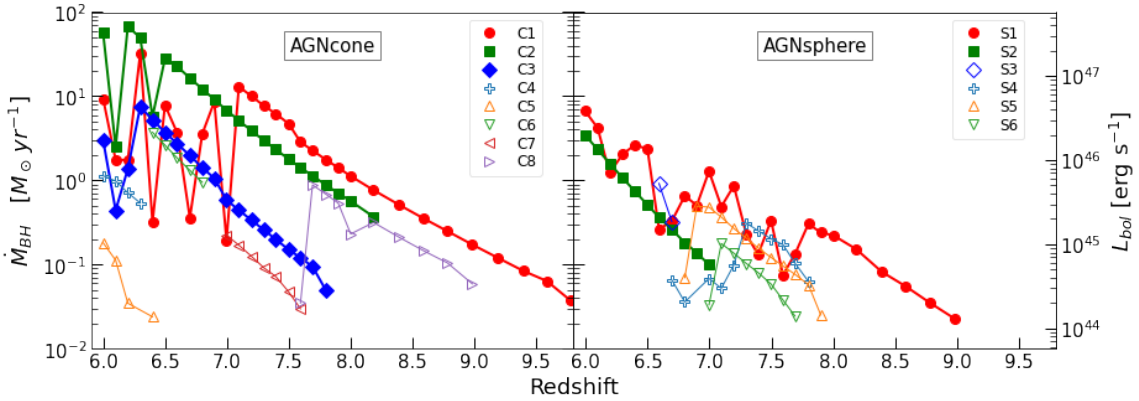

Figure 9.2: Mass accretion rate as a function of redshift for the *AGNcone* (left) and *AGNsphere* (right). The corresponding bolometric luminosity (Eq. 4.6) is reported in the right axis.

for a steady and efficient accretion onto the SMBH. This effect can be appreciated in Fig. 9.2: the accretion rate of *AGNcone* is generally higher than that of *AGNsphere*, at least up to $\dot{M} \approx 1 - 30\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$. At higher accretion rates, which are reached by the most accreting BHs at $z < 7$, AGN feedback prevents further increase of the accretion rate.

Hereafter, we focus our analysis on the AGN that reach $z = 6$ with $M_{BH} > 10^8\,\mathrm{M}_\odot$ and $L_{bol} > 10^{46}\,\mathrm{erg\,s}^{-1}$; (see filled symbols in Fig. 9.1 and Fig. 9.2), which we refer to as "bright AGN" (i.e., C1, C2, and C3 in *AGNcone*; S1 and S2 in *AGNsphere*). These BH mass and luminosity values are typical of known $z > 6$ QSOs (e.g., Yang et al., 2021b), allowing us to compare the physical properties of simulated and observed AGN in a consistent way. We note that, since the simulations focus on a single cosmic region at high redshift, the derived expectations on the AGN observable properties might be affected by cosmic variance.

9.2.3 Gas column density and radial distribution

Here we describe the method that we use to derive the distribution of hydrogen, helium, and metal column densities in the ISM for galaxies hosting AGN in the considered simulations. We make use of the hydrogen column density to derive the observational properties predicted by the two considered simulations.

We estimate the distribution of the column densities for the bright AGN in the simulations by launching 1000 randomly selected lines of sight (LOSs) toward each AGN from a distance $d = 30$ kpc. Each LOS is considered as the axis of a cylinder with basis radius of $R_{\rm soft}$. We note that the resolutions of the simulations do not allow us to probe structures on smaller scales, as, for instance, the existence of a dusty torus on pc scales. Then, each cylinder is divided along its length into bins of $l_{\rm bin} = 0.25$ kpc width, for a total of $\frac{d}{l_{\rm bin}} = 120$ radial bins. We compute the density of each chemical element in a bin of the cylinder from the mass carried by each particle included in that bin. With this approach, we also obtain the radial distribution of the gas density. Finally, we integrate along the cylinder to compute the total column density of hydrogen (N_H) and of the other elements. The resulting total N_H is not sensitive to reasonably different values of $l_{\rm bin}$ (i.e., from 0.25 kpc to 1 kpc). Therefore, we used $l_{\rm bin} = 0.25$ as this value allows us to sample well the radial distribution of the gas (see Section 9.4.2). Fig. 9.3 (upper panel) presents an example of the derived column-density map centred on the QSO C1 in *AGNcone*. Each circle represents one of the 1000 random LOSs, which sample homogeneously the entire solid angle as seen from C1.

To assess the effect of feedback on the column density (Section 9.3), we also consider an additional simulation run presented in Barai et al. (2018), that is identical in terms of initial conditions and physical prescriptions to the *AGNcone* and *AGNsphere*, except that BHs are not seeded. The only type of feedback in this run, which we refer to as *noAGN*, is due to supernovae explosions (see Barai et al. 2018 for detailed discussion).

We associate each AGN in a simulation to the corresponding galaxy in the *noAGN* runs following a method similar to that described in Zana et al. (2022): first, we identify the DM halo hosting the AGN as the one having its centre of mass closest to the position of the SMBH. Then, we identify the corresponding halo in the *noAGN* run by cross-matching the DM particle IDs in the two runs, and selecting the halo in *noAGN* which shares the largest fraction of particles with the initial AGN halo, further imposing that the mass difference must be within a factor of 10-50.%[2] Finally, we repeat the procedure described above on the selected halo in *noAGN*, and derive the column density distribution in absence of AGN feedback. At $z > 8$, the redshifts at which the *noAGN* snapshots are taken are significantly different from those of the runs including AGN, making the DM-halo match procedure highly uncertain. Thus, we limit the identification of the AGN-hosting galaxies counterparts in the *noAGN* run to $z < 8$.

9.3 Column density evolution

Fig. 9.4 presents the evolution of the column density for bright AGN in the *AGNcone* and *AGNsphere* simulations. Considering the *AGNcone* simulation, the AGN column densities are similar to, or slightly lower than, those derived for the corresponding galaxies in the *noAGN* run until the AGN accretion rate reaches $\dot{M} \approx 10 - 30\,{\rm M}_\odot\,{\rm yr}^{-1}$. This happens at $z \approx 7$ for C1 and C2, and $z \approx 6.3$ for C3 (see Fig. 9.2). At later times, the AGN column density drops significantly by up to ≈ 1 dex and the accretion rate starts to oscillate. The 10% and 90% percentiles span up to one order of magnitude, especially at $z = 6 - 7$, when the accretion rates reach the maximum values, producing the most powerful conical outflows.

Instead, the column densities of the corresponding galaxies in *noAGN* (grey stripes in Fig. 9.4) keep on increasing relatively smoothly. This finding confirms the AGN N_H drop and the presence of unobscured LOSs to the effect of the conical kinetic feedback. At low accretion rates the produced outflow cannot stop the infall of material, but once the accretion rate reaches

[2]The exact threshold is adjusted at each time step in order to find at least one halo counterpart.

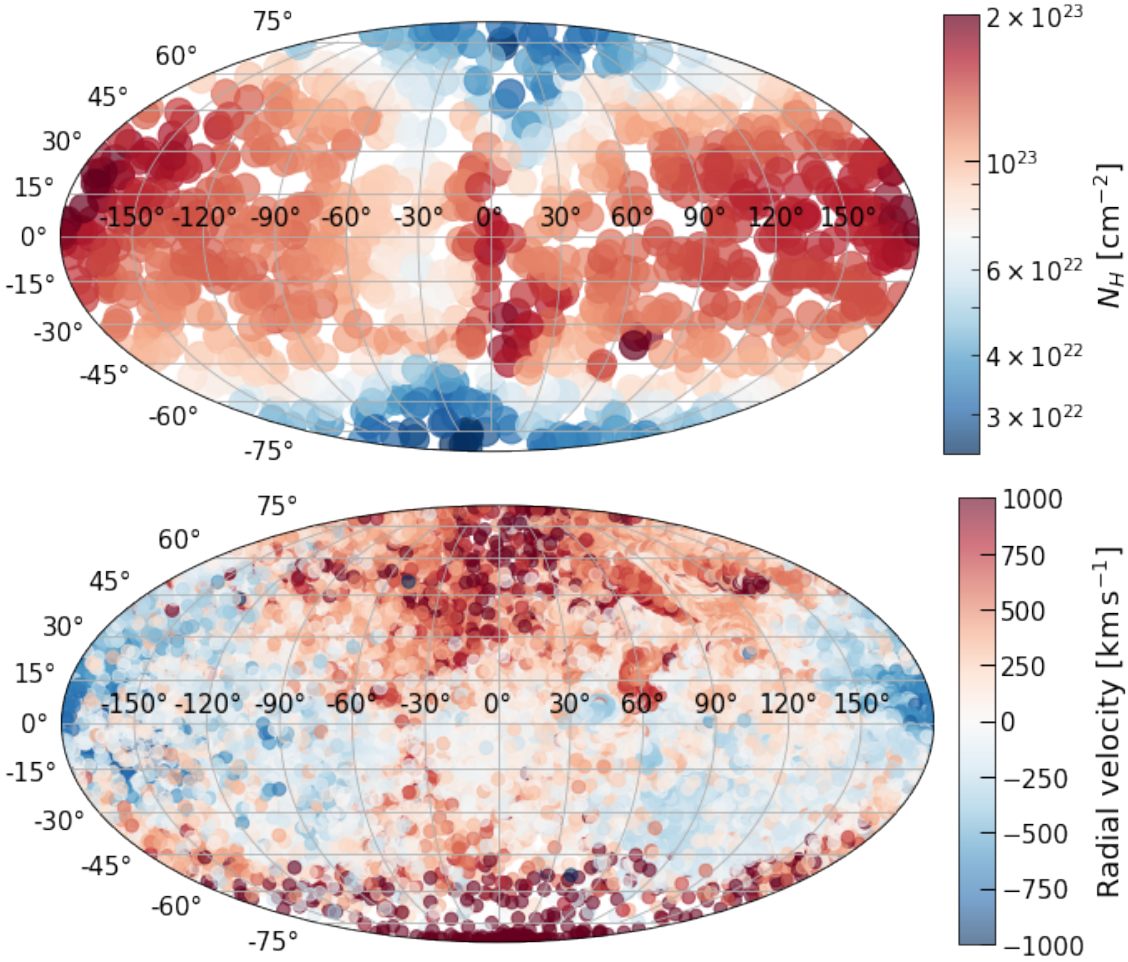

Figure 9.3: *Upper panel*: Mollweide projection of the column density along 1000 random lines of sights centred on the QSO C1 at $z = 7.1$. *Lower panel*: Mollweide projection of the radial velocities of all particles within 10 kpc from C1 at $z = 7.1$. The different sampling of the maps is intended to show the homogeneity of the 1000 LOSs used to compute N_H in the upper panel, and the velocity of the individual gas particles in the lower panel. The map is aligned with the outflow cone direction. Regions where the particles have high positive velocities correspond to the two cones along which the kinetic energy is distributed by the AGN feedback in the *AGNcone* simulation. Such cones are characterised by the lowest values of column densities.

high enough values, the energy carried by the outflow impacts a significant part of the gas in the halo, hindering further infalling, especially along the conical outflow directions. As a result, the N_H decreases, as well as the AGN accretion rate, until more material is allowed to accrete, producing a new burst of powerful feedback. Such a cyclic activity explains the decreasing median N_H, the wider N_H distribution, and the oscillating $\dot{M}$ behaviour at later cosmic times. This result is in qualitative agreement with the self-regulation scenario discussed by, e.g., Sijacki et al. (2009); Dubois et al. (2013); Costa et al. (2014a); Feng et al. (2014); Richardson et al. (2016); Trebitsch et al. (2019), according to which the AGN feedback controls the growth of the black hole and limits the duration of high accretion episodes by emptying the host galaxy gas reservoir, provided that the accretion rate is sufficiently high.

However, we note that the physical interpretation of our results is complicated by the effect that one AGN may have on other AGN-hosting galaxies passing through its feedback cone. In fact, C1, C2, and C3 in *AGNcone* at $z < 7$ are always closer than 30 kpc, and reach minimum distances as small as 4 kpc. At these distances, powerful outflows launched from one AGN may affect nearby galaxies (e.g., Zana et al., 2022).

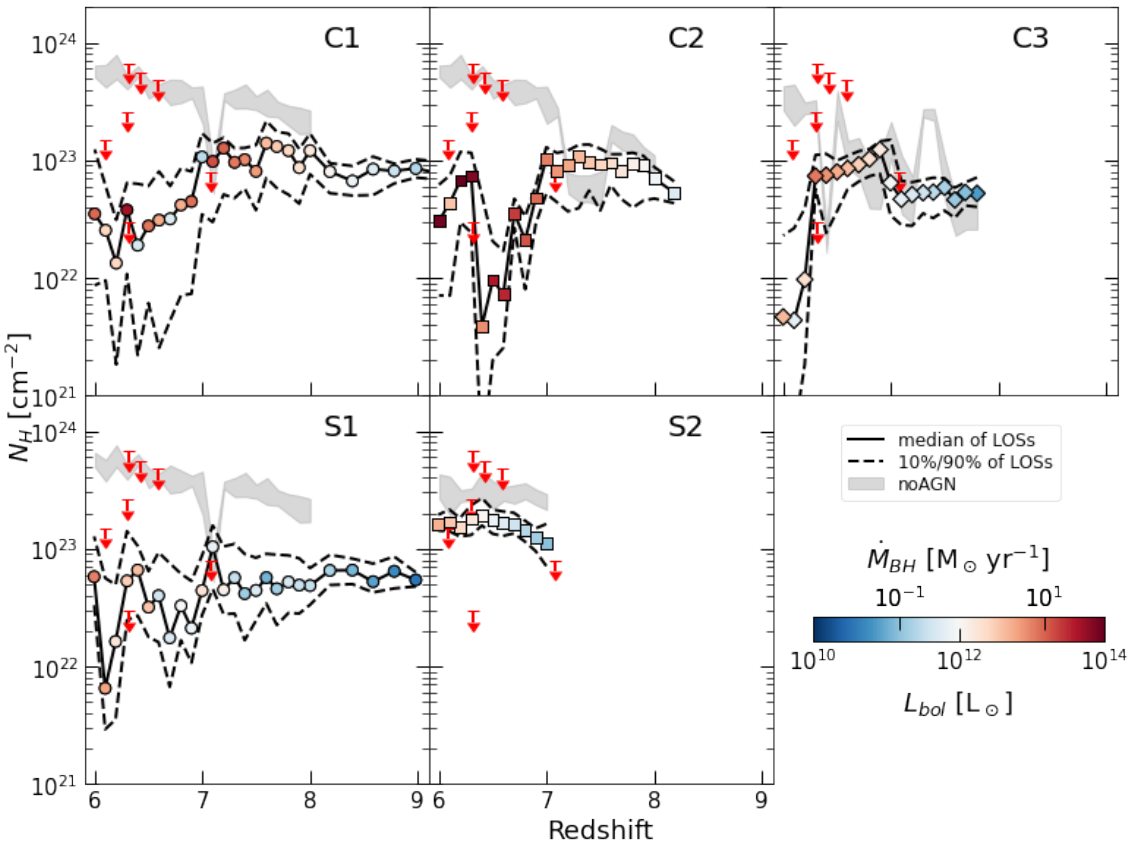

Figure 9.4: Evolution of column density for bright AGN in the *AGNcone* (C1, C2, C3) and *AGNsphere* (S1, S2) simulations. We show the median value (solid line, color coded according to the AGN bolometric luminosity and accretion rate), and the 10% and 90% percentiles (dashed lines) computed by launching 1000 lines of sight. The gray stripes enclose the 10% to 90% percentiles of the column densities of matched galaxies in the same simulation sets where, however, BHs have not been seeded (i.e., the *noAGN* case). To compare with observational results (Section 9.4.1), the red arrows mark the 3σ upper limits derived for X-ray detected QSOs with > 10 counts from Nanni et al. (2018) and Connor et al. (2019).

As an example of the feedback effect on the column density, in Fig. 9.3, we compare the N_H map centred on C1 with the radial velocity map of all particle within 10 kpc from C1. The maps correspond to $z = 7.1$, when C1 reaches a local maximum in accretion rate before the strong AGN feedback starts to impact significantly the N_H (Fig. 9.4) and $\dot{M}$ starts to oscillate. Comparing the column density map (upper panel) with the map of the radial velocity of individual particles (lower panel), we notice that the two conical outflows, identified as regions with positive radial velocities, correspond to LOSs with low column densities. Such LOSs are those along which high-redshift AGN are more easily to be detected in the rest-frame UV band, as we investigate in details in Section 9.4.3. Fig. 9.5 presents the fraction of LOSs along which $N_H < 10^{22}$ cm^{-2} (solid lines) and $N_H < 10^{23}$ cm^{-2} (dashed lines) for each bright AGN. Hereafter, we use the widely used threshold $N_H = 10^{22}$ cm^{-2} to separate obscured and unobscured AGN.[3] For instance, Merloni et al. (2014) found that such a value returns the best agreement between samples of obscured AGN as defined in optical (e.g., narrow emission-line AGN) and X-ray bands. From Fig. 9.5 we infer that only at $z \lesssim 7$ a fraction of the LOSs would appear as unobscured. In particular, at $z \lesssim 7$ C1 presents unobscured LOSs over $10 - 40\%$ of the solid angle, while this fraction is much more variable with redshift (i.e., $0 - 80\%$) for C2 and C3.

[3]However, we note that we consider the dust extinction as a more relevant quantity when we study the AGN rest-frame UV emission in Section 9.4.3.

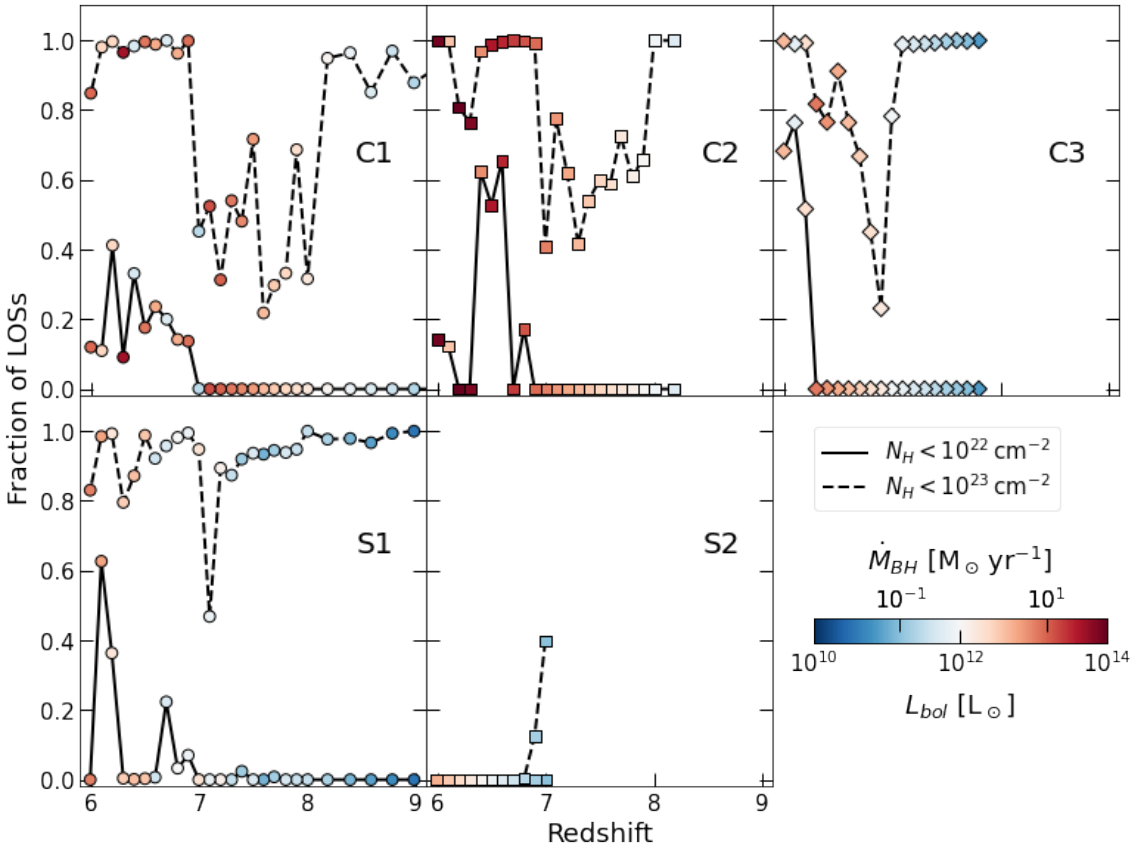

Figure 9.5: Fraction of lines of sight obscured by column densities $< 10^{22}\,\mathrm{cm}^{-1}$ (solid lines) and $< 10^{23}\,\mathrm{cm}^{-1}$ (dashed lines) as a function of redshift for the bright AGN in the *AGNcone* (C1, C2, C3) and *AGNsphere* (S1, S2) simulations. The symbols are color coded according to the AGN bolometric luminosity and accretion rate.

The most massive BH in the *AGNsphere* simulation, S1, follows a somewhat similar N_H evolution to that of the AGN in *AGNcone*: a roughly constant median N_H value up to $z \approx 7$ followed by a slightly decreasing and wider N_H distribution (Fig. 9.4), and the appearance of unobscured LOSs (Fig. 9.5) at later cosmic times. However, some differences exist: first, the AGN N_H is always significant lower than that of the corresponding galaxy in the *noAGN* run (grey stripe), even at $z > 7$. Secondly, the column density drop at $z < 7$ is not as strong as in the *AGNcone* case. Finally, at $z > 7$ the accretion rate of S1 is not as smooth as in the *AGNcone* case, and keeps on increasing even at $z < 7$. These differences may be due to the prescripted geometry of the kinetic feedback in the *AGNsphere* case, in which gas particles are accelerated in a random direction during every accretion event. Therefore, in contrast with the *AGNcone* case, there is no preferential direction (i.e., the equatorial plane of the conical outflow) along which material can keep on accreting undisturbed for long periods of time at $z > 7$. In particular, the accretion rate of S1 never exceeds $\approx 10\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, which is the approximate threshold after which the AGN kinetic feedback affects more evidently the N_H distribution and the accretion rate of AGN in the *AGNcone* run.

The column density evolution of S2, instead, does not appear to be strongly influenced by the AGN feedback. Although the median N_H is slightly lower than the values found in the *noAGN* case, it remains constant with time, and does not drop even at $z < 6.5$, when S2 reaches similar accretion rate to S1. As a result, S2 would never appear as an unobscured AGN. We note that the typical column density of S2 is a factor ≈ 3 higher than that of S1 at any redshift, and its accretion rate rises smoothly from $z = 7$ to $z = 6$. These properties suggest that higher accretion

rates than the values reached by S2 are required in order to launch outflows powerful enough to sweep away the gas in the case of large column densities (e.g., Trebitsch et al. 2019), even when kinetic energy is distributed along random directions by the AGN feedback.

The median values of N_H we derive from the Barai et al. (2018) simulations are consistent with typical values found by Lupi et al. (2022). However, the resolution of that work is $\times 85$ higher than our simulations, and allows the authors to sample compact regions of dense gas with $N_H \gtrsim 10^{24}\,\mathrm{cm}^{-2}$, especially at $z > 8$, when AGN feedback has not yet affected significantly the ISM distribution and density in the host galaxies. One of the main methodological differences with that work is that we compare the ISM densities in the same galaxies in which SMBHs are actively accreting or are not seeded at all. Thus, we probe directly the effect of AGN feedback on the ISM in the host galaxy.

9.4 Comparison with observations

In this section, we compare the observable properties derived from the N_H distributions of the AGN predicted by the simulations (Section 9.3) with observational results. In particular we focus on the comparison with constraints from X-ray observations (Section 9.4.1), the radial distribution of the gas reservoirs (Section 9.4.2), and the observed UV magnitudes (Section 9.4.3).

9.4.1 X-ray obscuration

X-ray observations are routinely used to constrain the column density of obscuring material along the LOSs of AGN. Low and moderate values of column densities ($N_H \lesssim 10^{22}$ cm^{-2}) can absorb soft X-ray photons (rest-frame energies $\lesssim 2$ keV), whereas larger column densities are required to absorb a high fraction of more energetic photons. However, X-ray observations of high-redshift QSOs (e.g. Vito et al., 2019b; Wang et al., 2021b) sample rest-frame energies $E > 3$ keV, and are thus sensitive only to high column densities ($N_H \gtrsim 3\times10^{23}$ cm^{-2}), at least at the sentivities of currently available facilities. Moreover, all of the known $z > 6$ QSOs have been selected based on their unobscured rest-frame UV emission (i.e., they are optically classified as type 1 QSOs), and thus are not expected to be heavily obscured in the X-ray band. For these reasons, existing X-ray observations of bright $z > 6$ QSOs provide us with only loose upper limits of N_H. The downward-pointing red arrows in Fig. 9.4 are the observed upper limits on N_H derived for a sample of $z > 6$ QSOs by Nanni et al. 2017 and Connor et al. 2019, with typical luminosities $L_{bol} = 10^{46} - 10^{47}\,\mathrm{erg\,s^{-1}}$. The column densities derived for bright AGN in all of the considered simulations are lower than, or consistent with, such loose upper limits. Although the N_H values found for the *noAGN* case are typically higher, they are still consistent with some measured upper limits. Therefore, the constraints on N_H obtained from X-ray observations of high-redshift QSOs only marginally favour the presence of kinetic feedback.

We note that constraining AGN obscuration using X-ray observations requires an assumption on gas metallicity, as X-ray photons are mainly absorbed by metal atoms. Typically, solar metallicity is assumed, whereas the ISM metallicity of the host galaxies of the AGN in the B18 simulations is sub-solar (e.g., by factors of $\approx 2-3$ at $z = 6$; e.g., Zana et al. in prep.). This consideration reinforces the overall consistency between the N_H values constrained from X-ray observations and found in the simulations, as significantly larger column densities would be required in the case of sub-solar metallicities to produce X-ray obscuration in excess to that observed in real QSOs. In Section 9.5 we discuss the X-ray detectability of the QSOs in the simulations.

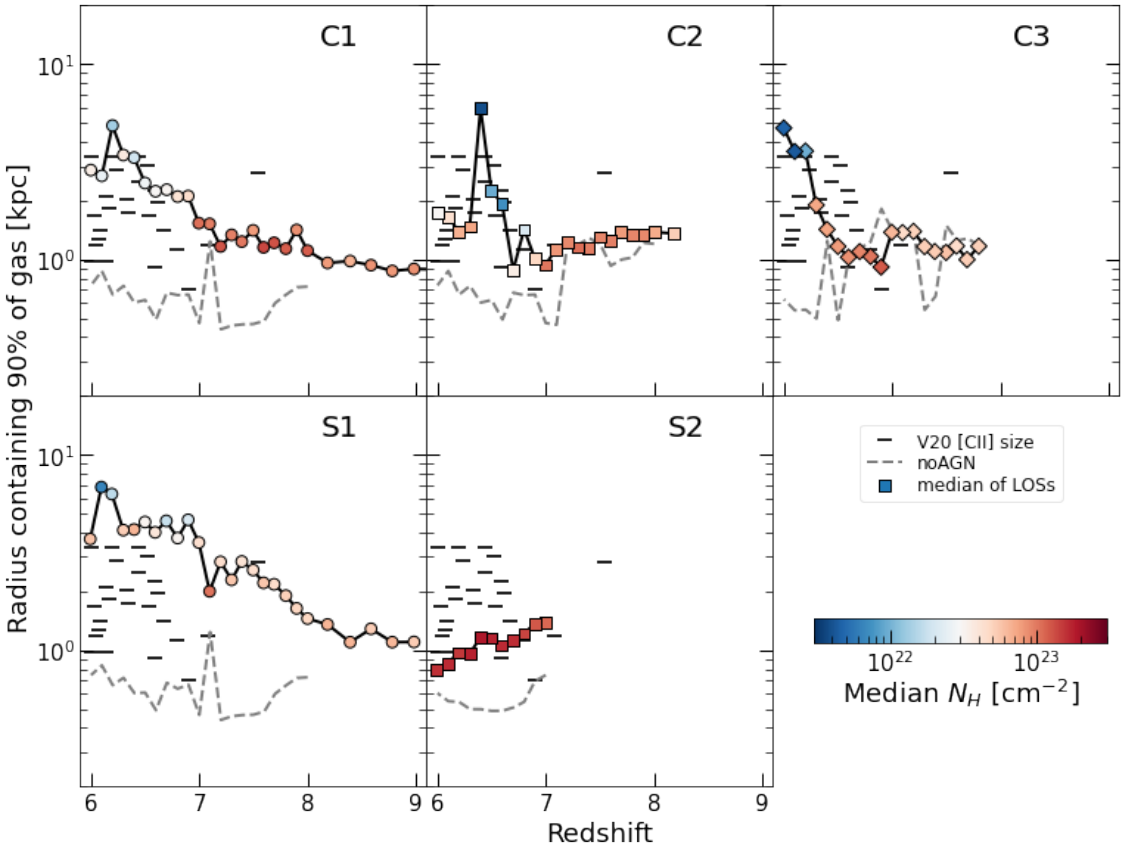

Figure 9.6: Median radius (R_{90}, computed over all of the lines of sight) containing the 90% of the total gas as a function of redshift. The color code indicates the median total N_H, averaged over all of the lines of sight. The dashed grey lines mark the same quantity computed for matched galaxies in the same simulation sets where, however, BHs have not been seeded (i.e., the *noAGN* case). The black ticks mark R_{90} for 25 $z > 6$ QSOs, as estimated from the [C II] emission beam-deconvolved sizes presented by Venemans et al. (2020).

9.4.2 Gas radial distribution

We investigate the effect of kinetic feedback on the observable sizes of the gas reservoirs in high-redshift QSOs. From the radial distribution of N_H derived for each LOS in Section 9.2, we computed the radius from the centre of the galaxy which includes 90% of the gas contributing to the total N_H. Then, for each galaxy, we computed the median value considering all of the 1000 LOSs, and define it as R_{90}. We use such a quantity to quantify the size of the gas reservoir in a galaxy.

Fig. 9.6 presents R_{90} as a function of redshift for every bright AGN in the *AGNcone* and *AGNsphere* simulations, as well as for the matched galaxies in the *noAGN* runs. All of the bright AGN in the *AGNcone* simulation (C1, C2, C3) have a similar evolution of R_{90}: their gas reservoir sizes are constant ($\approx$ 1 kpc) at $z \gtrsim 7$. At lower redshift, where N_H decreases due to strong effect of the kinetic feedback, which is proportional to $\dot{M}$ and L_{bol} (see the color-code of the circles in Fig. 9.4 and Fig. 9.6), R_{90} increases up to several kpc. This behaviour is expected considering that the AGN feedback applies a mechanical push to the surrounding gas particles. In fact, the size of the gas reservoir in the *noAGN* run, where the AGN feedback lacks (grey dashed lines in Fig. 9.6), remains constant or tends to even decrease at later cosmic times.

The evolution of R_{90} for S1 in the *AGNsphere* simulation is similar to that of the AGN in the *AGNcone* simulation. However, the increase of R_{90} is stronger and begins at earlier cosmic times. We recall that the accretion rate of S1 is typically lower than that of the AGN in *AGNcone* (see

Fig. 9.2), and therefore the stronger evolution of R_{90} is not due to intrinsically stronger outflows launched by the AGN, but, as discussed in Section 9.3, to the different geometry of the outflow: being launched along random directions at every accretion event, the outflow is more likely to transmit the kinetic energy to the gas particles in the galaxy even at low or moderate accretion rates. Instead, S2 does not follow the same evolution as S1. On the contrary, R_{90} decreases to sub kpc values approaching $z = 6$. As discussed in Section 9.3 we ascribe this behaviour to the relatively low accretion rate, which does not produce feedback strong enough to efficiently affect the gas distribution in the host galaxy.

We compare our findings with the observed extent of the [C II] emission of 25 $z > 6$ QSOs presented by Venemans et al. (2020), assuming that the [C II] emission line is a good tracer of the spatial extent of the total gas reservoir (e.g., Zanella et al., 2018; Sommovigo et al., 2021b). We used the major axis of the deconvolved [C II] emission size (Tab. 3 of Venemans et al. 2020), which represents the FWHM of the emitting source, and converted it into the radius that includes 90% of the [C II] light, assuming a Gaussian distribution.[4] The resulting values are reported in Fig. 9.6) as black ticks at the redshift of each QSO.

The AGN in the *AGNcone* simulation have R_{90} consistent with the observed values, while the median gas size of S1 is larger at nearly every redshift. S2 have a size consistent with the most compact QSOs in the Venemans et al. (2020) sample. However, this comparison is not fair: the ISM in S2 produces very large column densities at all redshifts and all LOSs (Fig. 9.4), and thus large expected values of dust extinction. All of the QSOs studied in Venemans et al. (2020) are instead rest-frame UV selected objects: we lack observational information about the extent of the gas reservoirs of buried high-redshift QSOs, as is S2. In all cases, the median gas size of the *noAGN* control galaxies are smaller than the observed values for QSOs, suggesting that kinetic feedback is required to produce the gas extents observed in real QSOs.

9.4.3 UV magnitudes

In Section 9.4.1 we discussed how the available X-ray observations of $z > 6$ QSOs are not sensitive to the column density values that we derived for bright AGN in the simulations. Instead, the rest-frame UV emission of high-redshift AGN is expected to be severely affected by dust extinction even for low values of N_H. In this section, we compare the expected rest-frame UV magnitudes of bright AGN in the simulations with the observed values of known $z > 6$ QSOs.

We assumed that the intrinsic (i.e., unextincted) rest-frame UV spectra of the AGN-hosting galaxies in the simulations are dominated by the AGN (i.e., we do not include stellar emission) and are well represented by the Vanden Berk et al. (2001) composite spectrum of type 1 QSOs, rescaled to their bolometric luminosity via the bolometric correction of Venemans et al. (2016) and Decarli et al. (2018)

$$\log\left(\frac{L_{bol}}{\mathrm{erg\,s^{-1}}}\right) = 4.553 + 0.911 \times \log\left(\frac{\lambda L_\lambda(1450\text{Å})}{\mathrm{erg\,s^{-1}}}\right). \tag{9.3}$$

We assumed a simple uniform slab of dust located in front of each AGN and an SMC extinction curve, and computed the measured rest-frame UV flux as

$$F_\lambda^{\mathrm{obs}} = F_\lambda^{\mathrm{intr}} e^{-\tau_\lambda}, \tag{9.4}$$

[4] We note that the conclusions hold if we use an exponential profile (e.g., Fujimoto et al. 2020) and convert the FWHM values reported by Venemans et al. (2020) into exponential scale lengths. In this case, we obtain larger radii than in the Gaussian case by a factor of ≈ 1.75.

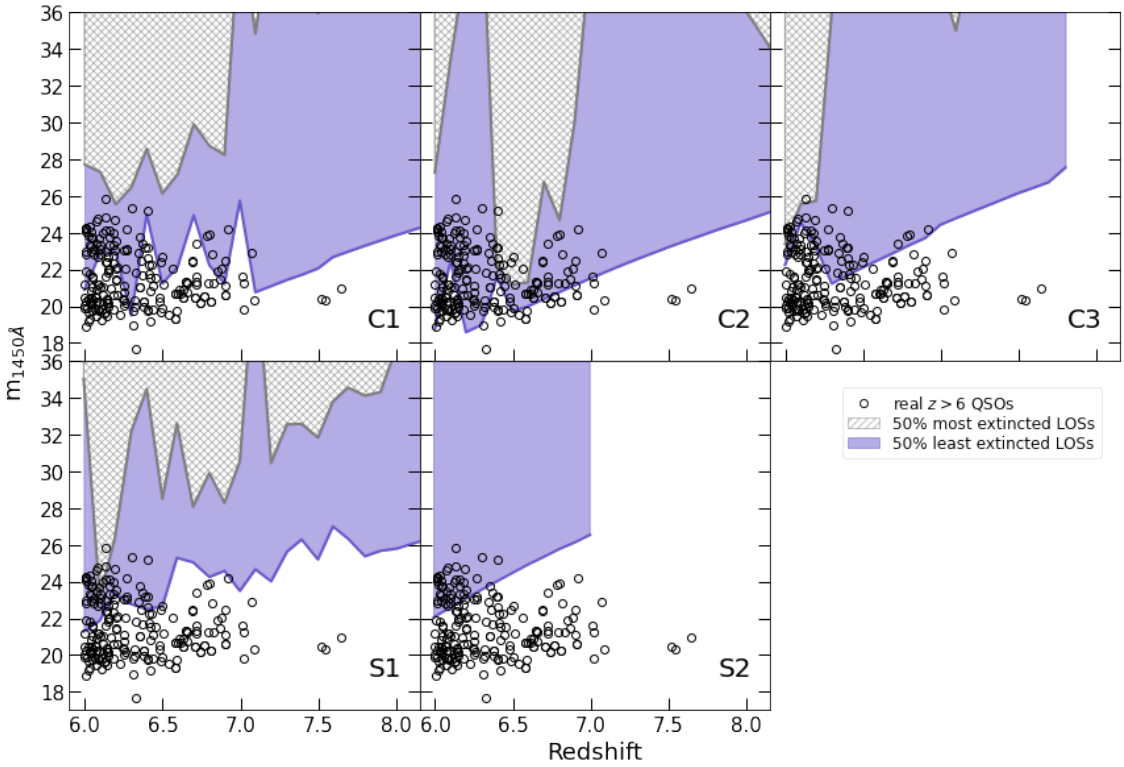

Figure 9.7: Apparent magnitude at rest-frame $\lambda = 1450$ Å as a function of redshift for bright AGN in the *AGNcone* (C1, C2, C3) and *AGNsphere* (S1, S2) simulations. The purple regions encompass the 50% least extincted LOSs, while the grey hatched regions represent the 50% most extincted LOSs. The grey circles are $z > 6$ QSOs collected from Bañados et al. (2016, 2018), Chehade et al. (2018), Matsuoka et al. (2018b,a, 2019a,b), Mazzucchelli et al. (2017a), Reed et al. (2017), Tang et al. (2017), Wang et al. (2017, 2018a,b, 2019, 2021a), and Yang et al. (2020a,b).

where $\tau_\lambda = k_\lambda \Sigma_m f_{dust}$, k_λ is the extinction cross section at wavelength λ, Σ_m is the mass column density of metals, which we computed in Section 9.2.3, and the fraction of metal mass locked into dust is assumed to be $f_{dust} = 0.15$ as in Di Mascia et al. (2021b). Finally, we computed the apparent magnitude at the wavelength corresponding to rest-frame 1450 Å, that is m_{1450}.

For all considered AGN, the metal mass is computed from the column densities of metals derived in Section 9.2.3 for 1000 LOSs at every simulation snapshot. Thus, we obtain a distribution of 1000 values of m_{1450} at every redshift. In Fig. 9.7 we show the magnitudes obtained for the 50% least (purple regions) and most (grey hatched regions) extincted LOSs. To allow for a comparison with observations, we add the magnitudes of a sample of $z > 6$ QSOs collected from Bañados et al. (2016, 2018), Chehade et al. (2018), Matsuoka et al. (2018b,a, 2019a,b), Mazzucchelli et al. (2017a), Reed et al. (2017), Tang et al. (2017), Wang et al. (2017, 2018a,b, 2019, 2021a), Yang et al. (2020a,b), with typical magnitudes of $19 \lesssim m_{1450} \lesssim 24$.

Among the considered simulations, *AGNcone* produces the UV brightest AGN, which are consistent with the magnitudes of known QSOs at $z \lesssim 7$. As discussed in Section 9.3, such redshift range corresponds to the period when the AGN strong kinetic feedback strongly affects the gas column density in the host galaxy, strongly suggesting that known, optically selected $z > 6$ AGN are indeed observed preferentially along directions where AGN feedback has cleared the LOS of most of the gas and dust. This prediction is hard to be tested observationally. Not only estimating the outflow direction is a difficult task, but the incidence of outflow in high-redshift AGN itself is still a matter of debate (e.g., Maiolino et al. 2012; Cicone et al. 2015; Bischetti et al. 2019; Novak et al. 2020; Izumi et al. 2021; Meyer et al. 2022). Moreover, $z > 6$ QSOs might have been detected along LOSs which have been previously cleared of most of the gas and dust

by past outflows. In this respect, a caveat arises from the numerical implementation of the ISM properties in the Barai et al. (2018) simulations, which, as described in Section 9.2.1, follow the prescription of Springel & Hernquist (2003). This model does not capture the ISM porosity and therefore is not able to resolve clumpy structures on pc scales. Resolving such structures might decrease the effective opacity of the medium and possibly produce more unobscured lines of sight, even in the absence of AGN feedback.

In Fig. 9.7, only $< 50\%$ of the LOSs of an individual AGN have extinction values small enough to reproduce the observed magnitudes. We computed the probability that multiple AGN appear as UV bright (i.e., $m_{1450} \lesssim 24$) sources along the same LOS, and found that it is negligible. This result is consistent with observations, according to which, to date, no such a system of multiple UV-bright AGN has been discovered at high redshift.

The most luminous AGN in the *AGNsphere* run, S1, reaches magnitudes as bright as the observed values only at $z \approx 6.5$, while it fails at reproducing the magnitudes of $z > 6.5$ QSOs. This is due to the lower accretion rate, and thus lower intrinsic luminosity, of S1 than the accretion rates of bright AGN in *AGNcone*. The large column density of S2 results in dramatic extinction levels along all of the LOSs, such that only along a small fraction of the LOSs S2 has apparent magnitude consistent with those of observed high-redshift QSOs, despite its intrinsic luminosity being similar to that of S1 at $z < 6.5$ (Fig. 9.2).

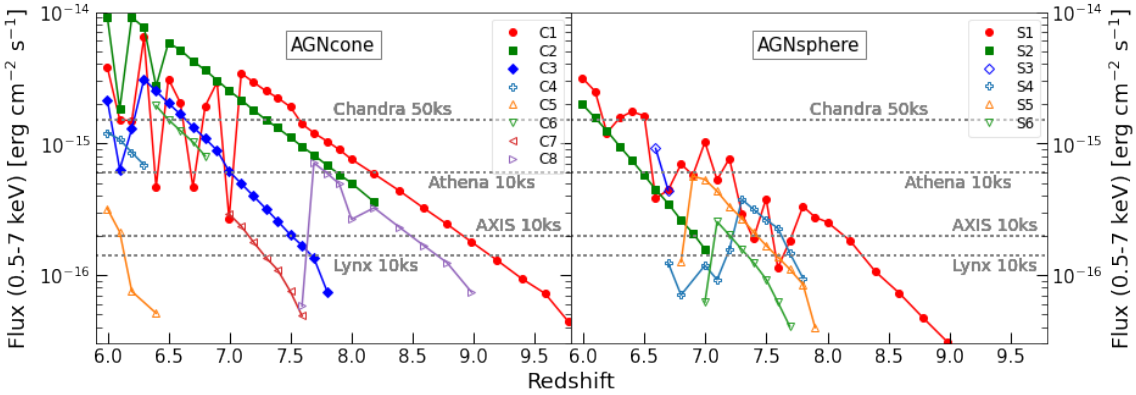

Figure 9.8: Expected X-ray flux in the 0.5–7 keV band as a function of redshift for the *AGNcone* (*left panel*), and *AGNsphere* (*right panel*) runs. For each AGN at each redshift, we assumed the median N_H computed for 1000 lines of sight. The horizontal dotted lines mark the flux limit computed for *Chandra* (50 ks observation), *Athena* (10 ks), *AXIS* (10 ks), and *Lynx* (10 ks).

9.5 Multiple high-redshift AGN on 1-10 arcsec scales

Typical separations between AGN in the Barai et al. (2018) simulations are $\approx 5 - 50$ kpc, corresponding to only a few arcseconds in projection. To date, no multiple AGN system has been discovered observationally at $z > 6$ (e.g., Greiner et al. 2021), with the highest redshift AGN pair being recently discovered at $z = 5.7$ (Yue et al., 2021). This result could be due to dust extinction preventing the detection of other possible accreting SMBHs close to high-redshift QSOs, as we found in our simulations (Section 9.4.3). Alternatively, QSOs observed at $z \gtrsim 6$ intrinsically have no AGN satellite. The latter hypothesis implies that the simulations overpredict the number of bright AGN, due to, e.g., the specific numerical setup and seeding prescription. In addition, as discussed in Section 9.2.1, the simulations focus on an overdense region, which maximizes the probability of forming multiple SMBHs, and thus bright AGN, in a small volume.

To investigate better the relation between the predicted and observed number of systems of multiple AGN at high redshift, in Section 9.5.1 we produce mock X-ray observations with the *Chandra* X-ray observatory[5] based on the *AGNcone* and *AGNsphere* simulations. Then, in Section 9.5.2 we compute the probability of detecting multiple AGN on small angular separations, and compare the findings with observational results. Finally, in Section 9.5.3 we investigate the potential of future X-ray facilities in detecting possible multiple faint AGN over small scales around bright high-redshift QSOs.

9.5.1 Mock X-ray observations

As discussed in Section 9.4.1, the column densities that we derived in Section 9.3 for simulated $z > 6$ AGN have a negligible effect on the X-ray emission at the observed-frame energies probed by X-ray telescopes, allowing us to factor out the effect of varying N_H along different LOSs. However, we have to take into account another effect related to the specific choice of the LOS: the emission of different AGN might be blended along some LOSs due to projection effects, and appear as a single X-ray source. This effect might be important as the projected angular separations of the AGN in the considered simulations are comparable with the angular resolution of *Chandra* (i.e., $\approx 0.5''$), which is the existing X-ray observatory with the sharpest view.

We produce mock observations using the SOXS v. 3.0 software,[6] using *Chandra* response matrices and ancillary files suitable for Cycle 20. SOXS accounts for three background components: a uniform Galactic component, a cosmic background due to point-like sources, and an instrumental component. For each simulation, we produce two sets of mock images, assuming an exposure time of 30 ks or 50 ks, which are typical lengths of real *Chandra* observations of $z > 6$ QSOs (e.g., Vito et al., 2019a; Wang et al., 2021b). For each set, we considered 100 random LOSs, along which all AGN have been projected on the sky plane according to their tridimensional positions in the simulations. This allows us to statistically take into account 1) the possible blending of multiple sources due to projection effects, and 2) the Poisson fluctuations of the number of detected X-ray photons at a given intrinsic flux.

We convert the bolometric luminosities of AGN in the simulations into X-ray luminosities in the rest-frame $2-10$ keV energy band using the Duras et al. (2020) relation. Then, we compute the fluxes in the 0.5-7 keV band (i.e., one of the standard energy bands used to analyse *Chandra* observations) for every AGN, and use them as input values to simulate the images. We adopt intrinsic powerlaw emission with photon index $\Gamma = 2$. This is a typical value for AGN up to $z \approx 6.5$ (e.g. Nanni et al., 2017; Vito et al., 2019b), although Vito et al. (2019b) and Wang et al. (2021b) find hints for a steepening at higher redshifts. We also include absorption due to the measured value of column density along the considered LOS, although, as discussed above, the produced obscuration is negligible for our high-redshift objects, and a Galactic absorption component with $N_H = 5 \times 10^{20}$ cm^{-2}. These computations have been performed with XSPEC v.12.11 (Arnaud 1996; model $phabs \times zvphabs \times powerlaw$)[7], which takes into account all the main absorption processes affecting X-ray photons (see also Section 2.2.1). Fig. 9.8 presents the expected X-ray flux of every AGN in the simulations as a function of redshift.

[5] https://cxc.harvard.edu/
[6] https://hea-www.cfa.harvard.edu/soxs/
[7] https://heasarc.gsfc.nasa.gov/xanadu/xspec/

9.5.2 X-ray detection of multiple AGN

We ran a blind source detection procedure on the *Chandra* mock observations in the 0.5-7 keV band using the *wavdetect* tool in CIAO v.4.12[8] (Fruscione et al., 2006), with a significance threshold of 10^{-5}, over an area corresponding to < 30 kpc from the central QSO, to be consistent with the volume considered throughout this work (see Section 9.2). We repeated this procedure for all snapshots in the $z = 6 - 7$ range, which includes most of the $z > 6$ QSOs observed with *Chandra*, thus allowing for a fair comparison with real observations.

Fig. 9.9 presents the number of AGN detected in the mock *Chandra* observations with 30 ks and 50 ks exposures, averaged over the 100 LOSs, for each simulation. *AGNcone* predicts an average of ≈ 1 detectable AGN already with relatively short exposures (30 ks) and multiple detected X-ray sources using slightly longer observations (50 ks) over all of the considered redshift range. Instead, according to the *AGNsphere* run, 30 ks (50 ks) *Chandra* observations of $z \gtrsim 6.2$ ($z \gtrsim 6.5$) should typically return no detected source, but the probability to detect one or more AGN increases quickly approaching $z = 6$.

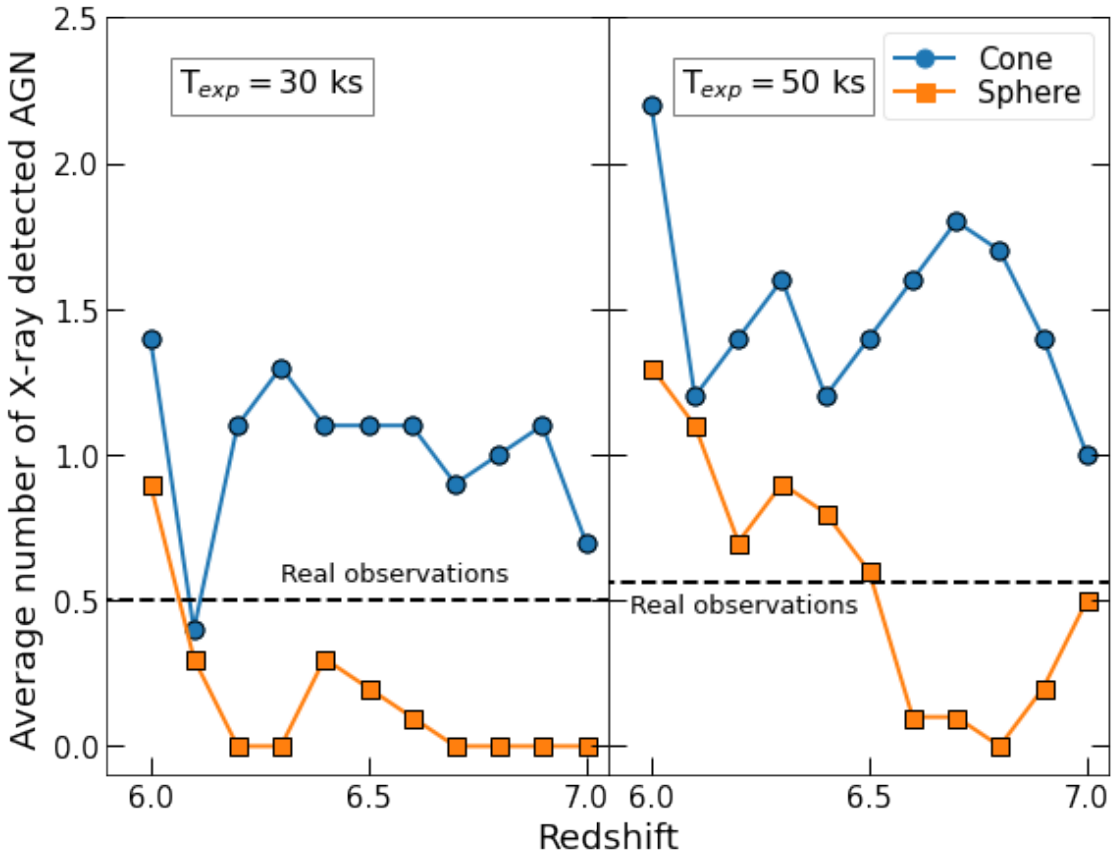

Figure 9.9: Average number of detected X-ray sources, averaged over 100 LOSs, detectable in the two simulations within < 30 kpc from the central AGN with 30 ks (left) and 50 ks (right) *Chandra* observations. The black dashed line mark the average number of detected sources in real observations of $z > 6$ QSOs.

In order to compare these results with real data, we collected all of the available *Chandra* observations of $z = 6 - 7$ QSOs with exposure times of 20-40 ks and 40-80 ks (Tab. 9.1). The median exposure time of the 20-40 ks (40-80 ks) observations is 38 ks (54 ks) and the median redshift of the targeted QSOs is $z = 6.4$ ($z = 6.5$). These values are well matched to our sets of 30 ks and 50 ks mock images, respectively. We repeated the detection procedure described

[8] https://cxc.harvard.edu/ciao4.12/

Table 9.1: Comparison sample of *Chandra* observations of $z = 6 - 7$ QSOs (see Section 9.5.2).

ID	z	Ref	ObsID	t_{exp} [ks]	N_{det}
(1)	(2)	(3)	(4)	(5)	(6)
20-40 ks sample					
J002429.77+391319.0	6.621	W21	20416	20	0
J005006.67+344521.6	6.253	V19	20393	34	1
J022601.87+030259.4	6.541	V19	20390	26	1
J084229.43+121850.5	6.076	V19	20392	29	0
J104819.09-010940.2	6.676	W21	20415	35	0
J150941.78-174926.8	6.122	V19	20391	27	1
J152637.84-205000.7*	6.586	C20	22165	33	0
J163033.90+401209.7	6.065	V19	5618	27	1
40-80 ks sample					
J010953.13-304726.3	6.791	V19	20398,22214	66	0
J030516.92-315055.9	6.614	V19	20394	50	0
J103027.11+052455.1*	6.308	N17	19926	50	1
J111033.98-132945.6*	6.515	V19	20397	54	0
J114816.65+525150.4	6.419	G17	17127	78	1
J164121.73+375520.2	6.047	V19	20396,21961	54	1
J203210.0-211402.3*	6.24	C19	20470	45	1
J223255.14+293032.3	6.666	V19	20395	54	1
J234833.34-305410.0	6.902	W21	20414	42	0

(1) ID of targeted QSO; (2) redshift of targeted QSO; (3) reference for published X-ray data. C19: Connor et al. (2019). C20: Connor et al. (2020). G17: Gallerani et al. (2017b). N17: Nanni et al. (2017). V19: Vito et al. (2019b). W21: Wang et al. (2021b). (4) *Chandra* observation ID considered in this work; (5) Exposure time; (6) number of detected X-ray sources according to the procedure described in Section 9.5.2. * These QSOs have been observed with multiple ObsIDs, resulting in longer total exposure times than those reported here. We only consider the reported ObsIDs to allow for a fair comparison with our 30 ks and 50 ks mock observations.

above on the real *Chandra* observations, considering only an area of $R < 30$ kpc from the targeted QSO, to allow for a fair comparison with the mock image results. We stress that the blind detection procedure prevents any bias related to rest-frame UV pre-selection of possible X-ray sources.

The last column of Tab. 9.1 reports the number of detected sources in the real observations,[9] which are almost equally split between no detected source and one detected source (i.e., the targeted QSO): the average numbers of detected X-ray sources in one observation are 0.50 and 0.56 for the 20-30 ks and 40-80 ks samples, respectively. Similar values are obtained by splitting each sample according to its median redshift. Comparing these results with the expected numbers of detected sources in simulations (Fig. 9.9), we find that *AGNcone* overestimates the number of detectable AGN at all redshifts, assuming both 30 ks and 50 ks exposure times. Instead, *AGNsphere* underestimates such number assuming 30 ks observations, while shows a strong dependence on redshift for longer exposures: at $z > 6.5$ and $z < 6.5$ it underestimates and overestimates, respectively, the average number of detected X-ray sources.

Due to the small sample sizes of real QSO observations and the narrow range covered by the number of detectable X-ray sources, it is difficult to provide a quantitatively robust comparison with the predictions from simulations. Nonetheless, we attempt to do it by comparing the normalized histograms of detected sources in the mock and real observations over the entire $z = 6 - 7$ range (Fig. 9.10). This is justified by the relatively flat redshift distribution of the QSOs targeted by real observations (Tab. 9.1). For each set of mock images, we computed the two-sample Anderson-Darling test.[10] The null hypothesis is that the mock and real observations are drawn from the same parent population, for what the number of detected X-ray sources is concern. We found that the null hypothesis can be rejected with high significance (i.e., Anderson-Darling test sigificance level $\lesssim 0.001$) for almost all combinations of simulations and exposure times: Fig. 9.10 confirms that *AGNcone* and *AGNsphere* overestimate and underestimate, respectively, the number of detectable X-ray sources. Mock simulations of *AGNsphere* with $t_{exp} = 50$ ks is the only set for which the null hypothesis cannot be rejected, although this simulation is not consistent with real observations for $t_{exp} = 30$ ks. It is worth noting that few $z > 6$ QSOs have been pointed with long *Chandra* exposures (100–500 ks; e.g. Nanni et al. 2018, Connor et al. 2020, Vito et al. 2021). Some of these observations were performed to check the presence of faint and possibly obscured AGN around $z > 6$ QSOs, for which companion galaxies have been detected with ALMA and HST. However, to date, no solid detection of such satellite AGN has been obtained (Vito et al. 2019a, 2021; Connor et al. 2019, 2020).

9.5.3 Predictions for future X-ray facilities

The high sensitivities of future X-ray facilities will allow us to push the search for AGN satellites of luminous optically selected QSOs at $z > 6$ down to intrinsic luminosities significantly lower than those probed with *Chandra*. In Fig. 9.8 we report as dotted grey lines the approximate expected sensitivity limits of future missions such as *Athena*/WFI (Nandra et al., 2013), *AXIS* (Mushotzky et al., 2019; Marchesi et al., 2020), and *Lynx*/HDXI (Gaskin et al., 2019), each one computed assuming 10 ks exposure time, and compare them with the sensitivity of a 50 ks *Chandra* observation. We computed these values by simulating X-ray observations of an X-ray

[9] We note that for almost all of the QSOs considered here, the results of the blind detection procedure agree with what reported in the literature, but for J084229.43+121850.5. Vito et al. (2019b) reported a detection of X-ray emission from this QSO, while here we report it as undetected. This apparent discrepancy is due to the different detection procedure (i.e., blind detection vs. rest-frame UV pre-selection of the target position) and significance threshold.

[10] We used the *anderson_ksamp* method of the SciPy package (Virtanen et al., 2020).

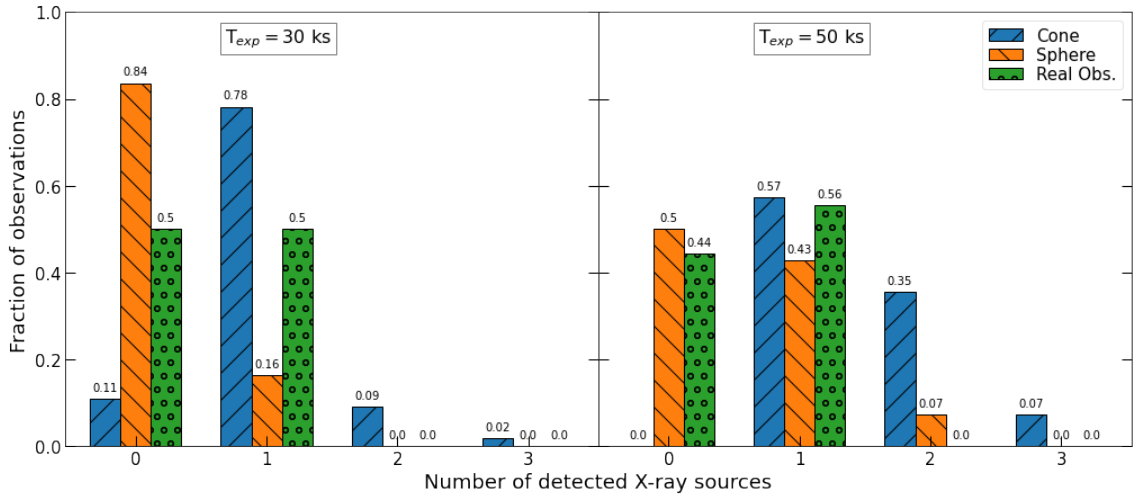

Figure 9.10: Normalized histograms of the number of detected X-ray sources in the mock and real *Chandra* observations of $z = 6 - 7$ AGN, for $t_{exp} = 30$ ks (left) and 50 ks (right).

source, assuming a simple power-law spectrum with photon index $\Gamma = 2$ and varying flux. In particular, for each instrument, we loaded response matrices and background files[11] in XSPEC, and computed the expected source and background count rates in a region including $\approx 90\%$ of the expected point spread function (PSF); i.e., $R = 1''$ for *Chandra*, *AXIS*, and *Lynx*, and $R = 5''$ for *Athena*. Then, we computed the flux that returns a binomial no-source detection probability (i.e., P_B; Weisskopf et al., 2007) such that $(1 - P_B) = 0.997$, corresponding to 3σ in the Gaussian approximation.

Fig. 9.8 shows that all of the considered next-generation X-ray mission will provide us with a huge improvement in the capability of detecting faint AGN at $z > 6$, including satellite AGN around bright QSOs at $z > 6$, in a fraction of the time of a typical *Chandra* observation. Fig. 9.11 presents simulated X-ray observations with *Chandra* (50 ks), *Lynx* (10 ks), *AXIS* (10 ks), and *Athena* (10 ks) of a representative snapshot (i.e., $z = 6.5$) and LOS of the two simulation runs. The satellite AGN will appear as multiple X-ray sources on a few arcsec scales. This implies that, in addition to high sensitivity, excellent angular resolution, such as that provided by *AXIS* and *Lynx*, is required to detect them individually. To probe this issue, we performed a blind detection run with *wavdetect* on these images, and compared the detected sources (black stars in Fig. 9.11) with the input AGN (colored circles): the identification of close objects like C1 and C2 is difficult even with missions with ≈ 0.5 arcsec angular resolution. The problem is clearly more evident with *Athena*, due to its PSF of a few arcsec.

9.6 Discussion

As discussed in Section 9.2.1, the outflow directions in the considered simulations are assumed not to be physically related to the host-galaxy properties and to be time-independent. In particular, the *AGNcone* simulation does not assume the outflow to be perpendicular to the plane of the host galaxy, as suggested by several observations of kpc-scale outflows or radio jets in the local universe (e.g., García-Burillo et al., 2014; Cresci et al., 2015; Morganti et al., 2015; Venturi et al., 2021), where the outflow geometry can be studied in details, and by some numerical simulations (e.g., Hopkins et al., 2012).

[11] We use real response matrices and background files for *Chandra*, and the preliminary files included in SOXS for *Lynx*, *AXIS*, and *Athena*.

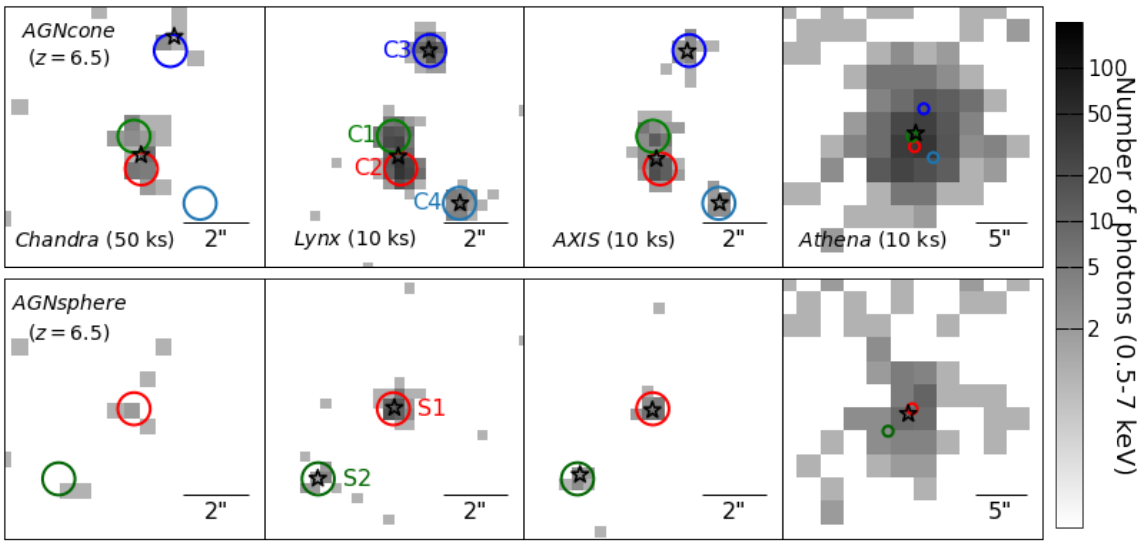

Figure 9.11: Simulated X-ray observations in the 0.5–7 keV band of the most-massive AGN at $z = 6.5$ and the surrounding satellite AGN in the *AGNcone* (upper row) and *AGNsphere* (lower row) simulations. From the leftmost to the rightmost columns, we simulated observations with *Chandra*/ACIS-S (50 ks), *Lynx*/HDXI (10 ks), *AXIS* (10 ks), and *Athena*/WFI (10 ks). For presentation purpose, the angular scale of the *Athena* image is different from the other cases, due to the larger PSF. The circles mark the location of the simulated AGN for a representative line of sight, and are color coded as in Fig. 9.1. The black stars mark the position of X-ray detected sources obtained with a blind detection procedure.

Several physical mechanisms can concur in the acceleration of winds at sub-pc scales that eventually produce large-scale outflows, including magneto-hydrodynamic effect (e.g., Sadowski et al. 2013), thermal driving (e.g., Proga 2007), radiation pressure acceleration, either applied on dust (e.g., Ishibashi & Fabian 2015) or mediated by UV transitions (e.g. Proga & Kallman, 2004; Mizumoto et al., 2021), which might produce outflows with different geometries. Moreover, the outflow geometry might be affected by interactions with the surrounding environment as the outflow expands (e.g. Nelson et al., 2019; Talbot et al., 2021), and might change with time. Cosmological simulations cannot describe in detail such a complex, and largely unknown, physics and evolution of outflows with relatively simple numerical recipes.

The goal of this work is to investigate the effect of two particular large-scale outflow geometries (i.e., a spherical outflow and a bi-conical outflow parametrized as described in Section 9.2.1) on the observable properties of high-redshift AGN, regardless of the sub-grid physical mechanisms responsible for their acceleration. Extensive numerical simulations with identical initial conditions and physics except for the outflow parameters would be required to check whether and how the results are sensitive to different choices of the outflow parameters.

Kinetic feedback produced during the phases of fast accretion of SMBHs in the Barai et al. (2018) simulations has a significant impact on the surrounding material and is required to match the predicted observable properties of bright AGN with observational results. One of the strongest piece of evidence is represented by the study of the gas extent in the AGN host galaxies (Fig. 9.6): the gas reservoirs in the *noAGN* case (i.e., in absence of AGN feedback) are always more compact than those derived from ALMA observations of $z > 6$ QSOs (see also, e.g., van der Vlugt & Costa 2019). The effect of AGN feedback pushes the gas in the host galaxies to larger distances (i.e., up to a few kpc) from the centres, in agreement with observations (e.g., Cicone et al. 2015; Bischetti et al. 2019; Venemans et al. 2020; Izumi et al. 2021). Although other mechanisms related to AGN feedback may produce such an observable, by, for instance, preventing gas infall from large scales (e.g., Trussler et al. 2020) or causing fluctuations

in the gravitational potential, which may lead to a radial migration of the material (e.g., van der Vlugt & Costa 2019), Barai et al. (2018) found that the mechanical removal of gas from the inner region of the host galaxies is the main process that affects their gas content in their simulations. We underline that also some $5 < z < 7$ star-forming ($1 - 70\,\mathrm{M_\odot\,yr^{-1}}$) galaxies have been found to show both an extended [C II] halo (e.g., Fujimoto et al., 2020) and broad wings in the [C II] emission-line profile (e.g., Gallerani et al., 2018; Ginolfi et al., 2020), suggestive of outflows possibly powered by a yet undetected accreting MBH (e.g., Orofino et al., 2021).

At $z < 7$ the feedback produces a general decrease of the N_H (Fig. 9.4), allowing for the appearance of unobscured (i.e., $N_H < 10^{22}\ \mathrm{cm^{-2}}$) LOSs (Fig. 9.5). Such directions are most probably those along which known $z > 6$ QSOs are preferentially observed, as the rest-frame UV selection of these objects requires low dust extinction. In fact, at $z \lesssim 6.5$, when the feedback effect is the strongest, bright AGN in the *AGNcone* simulation are able to reach the UV magnitudes observed for known $z > 6$ QSOs (Fig. 9.7). However, such LOSs represent only a fraction of the total LOSs of an AGN (see also, e.g., Ni et al. 2020; Trebitsch et al. 2019; Lupi et al. 2022): more than half of the LOSs would appear too faint to be selected as high-redshift objects in current optical/near-IR surveys, suggesting that a large fraction of the high-redshift, intrinsically luminous QSO population is observationally missed due to strong UV extinction produced by the ISM only. The presence of a dusty torus on pc scales, which is not included in the simulations we have analysed, would further increase such a fraction.

The outflow geometry likely plays an important role: in the case of a conic outflow, SMBH accretion proceeds at maximum efficiency through equatorial infalling of gas until $\dot{M} \approx 10 - 30\,\mathrm{M_\odot\,yr^{-1}}$ (Fig. 9.2), producing BHs with masses of $> 10^9\,\mathrm{M_\odot}$ at $z = 6 - 7$ (Fig. 9.1). At these accretion rates, the feedback regulates further accretion and reduces the typical obscuring column density, in particular along the cone direction (Fig. 9.3). In the case of outflows launched along random directions, the feedback can affect the growth of the SMBH and the N_H distribution even at lower accretion rates, resulting in $< 10^9\,\mathrm{M_\odot}$ BHs at $z = 6$, provided that the gas in the host galaxy is not too dense, as in the case of S2. Thus, the ISM properties (i.e., N_H and radial size of the gas) of the brightest AGN in the *AGNsphere* run is in agreement with observations. However, hindering the formation of $> 10^9\,\mathrm{M_\odot}$ BHs, the spherical geometry of the feedback in *AGNsphere* prevents AGN from reaching intrinsic luminosities comparable to known $z > 6$ QSOs at most redshifts (Fig. 9.7).

Interestingly, even the most luminous AGN in *AGNcone* cannot explain the detection of UV-bright QSOs at $z \approx 7.5$ (Fig. 9.7), due to the combination of the relatively small BH masses, and hence low accretion rates, which, by construction, are capped at the Eddington rate, and typically high N_H at that early cosmic time in this simulation. The existence of bright QSOs at $z \approx 7.5$ (e.g., Bañados et al., 2018; Wang et al., 2021b) requires different physical conditions for the SMBH formation and mass growth from those adopted in the considered simulations.[12] Future numerical simulations may explore such conditions as viable ways to reconcile the expected and observed properties of $z > 7$ AGN. Non-mutually exclusive possibilities are:
(a) different BH seeding mechanisms, that is, bright and massive QSOs discovered at $z \approx 7.5$ may be grown from more massive BH seeds or have been seeded at earlier redshift than the SMBHs in the simulations.
(b) Sustained periods of super-Eddington accretion at $z > 7.5$, whereas in the simulations the SMBH accretion rate is capped at the Eddington limit.
(c) Mass accretion characterized by a lower radiative efficiency than the value used in the simulations (i.e., $\epsilon_r = 0.1$). In this case, the mass that is not converted into radiation contributes to

[12] As mentioned in Section 9.2.2, cosmic variance may affect our conclusions, as the simulations focus on a single cosmic region at high redshift.

the growth of SMBH, which can reach higher masses than those found in simulations at a given time. For instance, Davies et al. (2019) report observational evidence for possible low radiation efficiency ($\epsilon_r \approx 0.001$) in high-redshift QSOs.

(d) High-redshift AGN typically reside in regions which are even more overdense than that investigated in the Barai et al. (2018) simulations, thus favouring the formation of SMBHs at earlier epochs. However, this possibility would arguably make the discrepancy between the observed and expected number of multiple X-ray detected AGN on small scales even worse. In addition, observational studies return contradictory results on the typical large-scale environment of high-redshift AGN (e.g., Ota et al., 2018; Mazzucchelli et al., 2019; Mignoli et al., 2020; Overzier, 2021).

The analysis that we have performed demonstrates that the comparison between several observable properties of AGN predicted by the Barai et al. (2018) simulations and the observational results, including both the properties of the individual galaxies and the environment, can help us to validate the recipes and assumptions adopted in numerical simulations. In particular, we found that AGN in the considered simulations match the gas radial distributions and apparent UV magnitudes of high-redshift QSOs. In addition, the same set of simulations has been demonstrated to reproduce well a number of physical properties of $z > 6$ QSOs, such as dust properties (Di Mascia et al., 2021b), multi-wavelength spectral energy distribution (Di Mascia et al., 2021a), and the number of UV-detected and [C II]-detected satellite galaxies (Zana et al., 2022).

However, we also found that the predicted number of X-ray detectable satellite AGN located over small scales around luminous high-redshift QSOs both in the *AGNcone* and *AGNsphere* simulations does not agree with the observational results. This observable is relatively easy to estimate from simulations as it depends primarily on the BH accretion rate only, once a suitable conversion to X-ray luminosity is assumed. Moreover, gas and dust absorption does not affect significantly the observed X-ray emission from high-redshift AGN, as opposed to UV emission, up to high column densities ($\log \frac{N_{\rm H}}{\rm cm^{-2}} \approx 23.5 - 24.0$; see Section 9.4.1 and Section 9.4.3).

The mismatch between the number of multiple X-ray detected AGN on small scales between simulations and observations may be related to numerical issues and physical prescriptions. In particular, the simplistic BH seeding recipe implemented in the considered simulations (i.e., a $10^5\,{\rm M}_\odot$ BH is placed in the centre of a galaxy when this reaches a given mass threshold) naturally leads to the formation of a large number of SMBHs, that would appear as bright AGN at later cosmic times. Similar seeding recipies have been commonly adopted by most cosmological simulations (e.g., Costa et al. 2014a, Di Matteo et al. 2017, Barai et al. 2018, Smidt et al. 2018, Lupi et al. 2019, Valentini et al. 2021), and typically mimic the "heavy seed" formation channel for SMBHs (e.g., Lodato & Natarajan, 2006; Ferrara et al., 2014). However, theoretical models of "heavy seed" formation require stringent physical conditions on, e.g., metallicity, physical state of the gas, ad radiation fields (e.g., Ferrara et al., 2014). Accounting for such conditions in cosmological simulations is particularly difficult, but would reduce the number of formed SMBHs, and thus the discrepancy with observational results.

Another possibility is that observed QSOs at high redshift do not reside in regions as dense as those probed in the analysed simulations (but see, e.g., Zana et al. 2022). In this case, the formation of multiple SMBHs is expected to be hindered, helping us reconcile the expected number of X-ray sources with observational results. In addition, we would also expect to form less massive BHs, with direct consequences on the observational expectations discussed in this Chapter, as the BH mass is tightly linked with the maximum accretion rate, and thus AGN luminosity and feedback strength. Qualitatively, we would expect to derive fainter rest-frame UV and X-ray fluxes, weaker feedback, and, as a consequence (see Fig. 9.6), more compact gas reservoirs (i.e., similar to the *noAGN* case) than the values discussed in Section 9.4.2, Section 9.4.3,

and Section 9.5. Future X-ray facilities will provide us with the required sensitivity and angular resolution to investigate the presence of multiple faint AGN around bright high-redshift QSOs down to unprecedented flux limits (see Section 9.5.3).

9.7 Summary and conclusions

We studied the observable properties of $z = 6 - 10$ bright AGN in a suite of zoom-in cosmological simulations by Barai et al. (2018) characterized by the inclusion of AGN kinetic feedback with either bi-conical (namely, *AGNcone*) and spherical (*AGNsphere*) outflow geometry. We focused our investigation on the gas column density and size in the host galaxies, the AGN rest-frame UV magnitude and X-ray fluxes, and the detectability of systems of multiple AGN over a few kpc scale in the X-ray band. We compared these quantities with a control simulation in which SMBHs are not seeded (i.e., *noAGN*), and observational results of $z > 6$ AGN. We summarize our findings as follows.

- *AGNcone* produces three bright AGN that grow up to $5 \times 10^8 < M_{\rm BH} < 5 \times 10^9\,{\rm M}_\odot$ at $z = 6$. These objects are characterized by a steady increase of their accretion rate up to $\approx 10 - 30\,{\rm M}_\odot\,{\rm yr}^{-1}$. Once such high values are reached (at $z \approx 6.5 - 7$), the strong AGN feedback prevents further increase of the accretion rate. This behaviour is linked with the bi-conical geometry of the outflow, that allows steady infalling of material along the equatorial directions, at least until the feedback grows strong enough to affect most of the gas in the galaxy halo. In *AGNsphere*, the spherical geometry of the outflow affects gas accretion already at low and moderate SMBH growth rate. For this reason, the two bright AGN produced in *AGNsphere* reach lower values of BH masses (i.e., $2 \times 10^8 < M_{\rm BH} < 5 \times 10^8\,{\rm M}_\odot$) and accretion rates ($\dot{M} < 10\,{\rm M}_\odot\,{\rm yr}^{-1}$) than objects in *AGNcone*.

- AGN host galaxies in *AGNsphere* have gas column densities of $N_H \approx 10^{23}\,{\rm cm}^{-2}$ from their formation up to $z = 6.5 - 7$, when N_H presents a remarkable drop due to the strong AGN feedback. In fact, the N_H in matched galaxies in *noAGN* continues to increase during the entire considered redshift range. The brightest AGN in *AGNsphere* presents a similar behaviour as those in *AGNcone*, although the N_H is typically slightly lower. We interpret this difference again as due to the assumed spherical symmetry of the outflow. Instead, the second bright AGN in *AGNsphere* do not reach accretion rate sufficiently high to significantly affect the gas in the host galaxy. Our findings are consistent with the upper limits on N_H recently reported for a set of $z > 6$ AGN observed in the X-rays.

- Kinetic feedback is required to match the gas extent reported for high-redshift QSOs (i.e., up to a few kpc). In fact, galaxies in *noAGN* present typical gas sizes of < 1 kpc, while the extents of the gas reservoirs of AGN in *AGNcone* and *AGNsphere* increase up to the observed values of a few kpc at $z \lesssim 7$. The exception is the second bright AGN in *AGNsphere*, due to its relatively low values of accretion rate.

- All AGN in the simulations would appear as obscured (i.e., $N_H > 10^{22}\,{\rm cm}^{-2}$) along all lines of sight (LOSs) at $z > 7$. These objects would be missed by currently employed UV-based selection methods, which are heavily affected by dust extiction, and would require observations in different bands (e.g., X-ray or infrared) to be unveiled. At later cosmic times, a fraction of LOSs (up to $\approx 80\%$, depending on the specific AGN and redshift) have $N_H < 10^{22}\,{\rm cm}^{-2}$. These are the preferential directions along which known, UV-selected $z > 6$ QSOs are observed.

- Under simple, but reasonable, assumptions on the gas-to-dust mass scaling and dust distribution, we estimate the apparent UV magnitudes (m_{1450}) of the AGN in the simulations along different LOSs. We found that AGN in *AGNcone* have m_{1450} consistent with those observed for real high-redshift QSOs (i.e., $m_{1450} < 25$) along $\lesssim 50\%$ of the LOSs at $z < 7$. AGN in *AGNsphere*, instead, have fainter magnitudes, due to the lower instrinsic luminosities, and, for the second AGN, the high extinction levels along most of the LOSs. No AGN in the simulations can reproduce the observed UV magnitudes of the few $z \approx 7.5$ QSOs known to date, whose formation and accretion history are likely not well captured by the prescriptions assumed in the simulations.

- The presence of multiple bright AGN over scales of a few kpc led us to investigate their detectability in X-ray observations with *Chandra*, and to compare the results with real observations of $z > 6$ QSOs. We found that the *AGNcone* run significantly overpredicts the number of X-ray detected multiple AGN at high redshift. Instead, *AGNsphere* produces AGN with lower rate of X-ray detection than typical values derived in relatively shallow (i.e., 30 ks) observations, while it is consistent with the results obtained with longer (i.e., 50 ks) observations.

These results demonstrate that the AGN in the considered simulations have physical properties consistent with those of real QSOs for what the column density and extent of the gas in the host galaxies and the UV magnitudes are concerned. A bi-conical geometry for the outflow is favored over a spherical geometry, as it reproduces AGN with the high luminosities and SMBH masses observed for $z = 6 - 7$ QSOs. However, both simulations cannot explain the recent discovery of luminous QSOs at $z \approx 7.5$, which may have been formed at higher redshift than the assumed seeding time in our simulations, or may have undergone extensive periods of super-Eddington accretion.

Moreover, we showed that the number of multiple AGN detectable in X-ray band over few kpc scales is the observable property that the considered simulations struggled the most to reproduce. We propose that this issue can be due to the simplistic BH seeding methods generally implemented in cosmological simulations, that do not account for the complex physics related with the formation and rapid growth of massive BHs in the early Universe. Future X-ray observatories will provide us with the sensitivity required to investigate the possible presence of multiple faint AGN satellites around luminous QSOs at high redshift.

Part IV

Conclusions and future prospects

Summary

10

The existence of supermassive black holes (SMBHs) powering bright quasars already during the first Gyr of life of our Universe is one of the most important and fascinating puzzles in modern Cosmology. The processes regulating their formation and growth are also profoundly interwoven with the evolution of their host galaxies. Therefore, the study of early black holes is fundamental in order to understand the evolution of galaxies across cosmic history. However, our knowledge is limited by the data provided by current observational facilities, that are likely able to probe just the tip of the iceberg of the whole black hole population at high redshift. Among many, the following questions are still open:

- What are the physical properties of early SMBHs and their hosts?
- What are the spectral signatures of their progenitors?
- How did the AGN activity affect the host galaxy?

In this PhD Thesis, we tackled the aforementioned questions by exploiting two suites of zoom-in cosmological hydrodynamic simulations, which follow the evolution of SMBHs and their host galaxies under different prescriptions for the AGN feedback. In particular we analyse a run with kinetic spherical feedback (*AGNsphere*), a run with kinetic bi-conical feedback (*AGNcone*) and a run with thermal spherical feedback (*AGNthermal*). We then combined these simulations with advanced radiative transfer calculations, performed with the code SKIRT, making observational predictions for early SMBHs and exploring the contribution and the signatures of AGN emission to the observed spectrum. A multi-wavelength study of these objects, from the Infrared (IR) band to the optical/UV and up to the X-ray wavelengths, allows us to explore different processes related with the intrinsic properties of this primeval objects, the gas distribution around them, and the dust properties in early galaxies.

First, we parameterized the intrinsic AGN SED via a composite power-law $F_\lambda \propto \lambda^\alpha$ constrained through observational and theoretical arguments, and we implement this model in a customized version of the SKIRT code. We then investigated how different assumptions on dust mass content (expressed in terms of the dust-to-metal ratios $f_\mathrm{d} = 0.08$ and $f_d = 0.3$) affect the observed spectral energy distributions (SEDs) of $z \sim 6$ galaxies. We found that in the most massive galaxies ($M_\mathrm{d} \gtrsim 3 \times 10^7\ \mathrm{M}_\odot$) a large fraction ($\gtrsim 50\%$) of UV emission is obscured by dust, with a significant scatter (up to a factor of ≈ 5) between different lines of sight (los), due to the inhomogeneous distribution of dusty gas on scales $\gtrsim 100$ pc. The simulated quasars reside in complex, dust-rich merging systems, containing both multiple accreting SMBHs and star forming galaxies. However, due to the strong dust absorption, many of the AGN can be undetected in the optical/NIR and appear as SMG companions, consistently with recent ALMA data.

We have also investigated the dust temperature in AGN hosts, finding that a clumpy, warm ($\approx 200 - 300$ K) dust component, coexists with a colder ($\approx 50 - 70$ K) and more diffuse dusty medium, heated by stars. In galaxies hosting the most-active AGN, this warm dust component provides up to 50% of the total infrared luminosity, though constituting only a small fraction ($\lesssim 0.1\%$) of the overall mass content. The presence of warm dust is also imprinted in the galaxy SEDs, with an increase of the flux in the MIR band. The enhanced MIR flux can be exploited by future MIR facilities, like the proposed Origins Space Telescope (OST). We also compared our synthetic SEDs with the sensitivity limits of OST, finding that both highly star forming galaxies and bright quasars ($M_{UV} < -26$) at high-z can be detected with high signal-to-noise ratios even at this large distances. Finally, we found that the FIR/MIR flux ratio in star forming galaxies is one order of magnitude higher with respect to AGN hosts, even in the case of low accretion rates ($\dot{M}_{\rm BH} \sim 3\ {\rm M}_{\odot}\ {\rm yr}^{-1}$). We predict that, by following up with OST galaxies already detected with ALMA it will be possible to unveil faint and/or dust-obscured AGN, whose fraction is expected to be large ($> 85\%$) at high redshift.

Next, we used our numerical framework to investigate the reliability of the star formation rates inferred for high redshift quasar-hosts, which are based on the assumption that the IR emission is dominated by dust-reprocessed stellar light. We performed multiple radiative transfer computations to carefully isolate the contribution of AGN radiation, and we found that AGN with $L_{\rm bol} \gtrsim 10^{13}\,{\rm L}_{\odot}$ effectively heat the bulk of the dust in the interstellar medium (ISM) on galaxy scales, and not only the dust on ≈ 100 pc from their surrounding. By applying a calibration of the ${\rm SFR} - L_{\rm TIR}$ relation we find that the ${\rm SFR}$ tends to be overestimated by a factor of ≈ 3 (≈ 30) for $L_{\rm bol} \approx 10^{12}$ ${\rm L}_{\odot}$ ($L_{\rm bol} \approx 10^{13}$ ${\rm L}_{\odot}$). Our results suggest that the SFR in all the high-z quasar-hosts might be over-estimated by at least an order of magnitude.

Next, we focused on the optical/UV portion of the spectra, in order to study the dust properties in $z \sim 6$ quasars. Our main goal have been to investigate whether extinction curves such as those found in the the Small Magellanic Cloud (SMC) or in the Milky Way (MW) can explain multi-wavelength observations of high-redshift quasars spectra. We first examined the effects of relative dust-sources geometry on the synthetic attenuation curves, finding that AGN feedback lead to steeper attenuation curves with higher los-to-los variations with respect to galaxies not experiencing AGN feedback. We provide an analytical model that relates the steepness of the attenuation curve with the V-band optical depth, finding that the attenuation curves tend to be flat with high dust optical depths. The steepness of attenuation curves in presence of AGN feedback is ultimately determined by gas distribution affected by AGN activity, as powerful outflows remove gas from the central regions, reducing the overall column densities.

We then compared the synthetic SEDs produced from the *AGNcone* run, which has the intrinsically brightest AGN ($M_{\rm UV} \lesssim -26$), with multi-wavelength data of a sample of bright $z \sim 6$ quasars with comparable UV magnitudes ($-29 \lesssim M_{\rm UV} \lesssim -26$). Among the twenty different RT models implemented, the ones favoured by a χ^2-analysis have a grain size distribution lacking grains with sizes $a < a_{\rm min} = 0.1\ \mu{\rm m}$. This grain size distribution is needed in order to reconcile the synthetic SEDs with the optical/UV data. The attenuation curves inferred for these models are close to flat and consistent with recent analysis of $3.9 \leq z \leq 6.4$ reddened quasars. Therefore we caution toward the applicability of local extinction laws (e.g SMC or MW) in high-z quasar-hosts when interpreting observations. The dominance of large grains in AGN-hosts might either indicate efficient selective dust destruction processes in place in AGN-environments, or alternative dust-formation and evolution mechanisms, such as supernova origin or efficient coagulation of small grains into larger ones.

Finally, we focused on the high-energy spectrum, by investigating the gas column density and size in the host galaxies, the AGN rest-frame UV magnitude and X-ray fluxes, and the de-

tectability of systems of multiple AGN over a few kpc scale in the X-ray band. First, we found that bi-conical feedback allows steady infalling of material along the equatorial directions, whereas a spherical geometry of the outflow affects gas accretion already at low and moderate SMBH growth rate. Overall, kinetic feedback is required to match the gas extent reported for high-redshift QSOs (i.e., up to a few kpc). The gas column density is found to evolve with redshift: all AGN in the simulations would appear as obscured (i.e., $N_H > 10^{22}\,\mathrm{cm}^{-2}$) along all lines of sight at $z > 7$. At later cosmic times, a fraction of los (up to $\approx 80\%$, depending on the specific AGN and redshift) have $N_H < 10^{22}\,\mathrm{cm}^{-2}$. These are the preferential directions along which known, UV-selected $z > 6$ QSOs are observed.

We made use of the computed column densities to estimate the apparent UV magnitudes (m_{1450}) of the AGN in the simulations along different los. We find that AGN in *AGNcone* have m_{1450} consistent with those observed for real high-redshift QSOs (i.e., $m_{1450} < 25$) along $\lesssim 50\%$ of the LOSs at $z < 7$. Instead, in *AGNsphere*, they have fainter magnitudes, either due to the lower instrinsic luminosities, or to high extinction levels along most of the LOSs. These results demonstrate that the AGN in the considered simulations have physical properties consistent with those of real QSOs for what the column density and extent of the gas in the host galaxies and the UV magnitudes are concerned. A bi-conical geometry for the outflow is favored over a spherical geometry, as it reproduces AGN with the high luminosities and SMBH masses observed for $z = 6 - 7$ QSOs. However, both simulations cannot explain the recent discovery of luminous QSOs at $z \approx 7.5$, which may have been formed at higher redshift than the assumed seeding time in our simulations, or may have undergone extensive periods of super-Eddington accretion.

Finally, we investigated their detectability in X-ray observations with *Chandra*, and compare the results with real observations of $z > 6$ QSOs. We found that the *AGNcone* run significantly overpredicts the number of X-ray detected multiple AGN at high redshift. Instead, *AGNsphere* produces AGN with lower rate of X-ray detection than typical values derived in relatively shallow (i.e., 30 ks) observations, while it is consistent with the results obtained with longer (i.e., 50 ks) observations. We propose that this issue can be due to the simplistic BH seeding methods generally implemented in cosmological simulations, that do not account for the complex physics related with the formation and rapid growth of massive BHs in the early Universe.

Overall, the work presented in this Thesis represents few steps towards a better understanding of SMBHs at high-redshift and in answering the open questions we highlighted above:

- *What are the physical properties of early SMBHs and their hosts?* The simulations adopted in our work suggest that SMBHs at $z > 6$ are characterized by very high accretion rates ($\gtrsim 10\ \mathrm{M}_\odot\,\mathrm{yr}^{-1}$) and bolometric luminosities ($\gtrsim 10^{13}\ \mathrm{L}_\odot$). However, they are obscured by high column densities ($N_H > 10^{22}\,\mathrm{cm}^{-2}$) for the majority of their lifetimes, and only at $z < 7$ they become unobscured along most of the lines of sight. Moreover, we find the ISM dust in their hosts to be be dominated by large grains with sizes $a \gtrsim 0.1\ \mu\mathrm{m}$. This results in their attenuation curves to be grey, consistently with recent data.

- *What are the spectral signatures of their progenitors?* The complex gas distribution in early galaxies results in large variations of the UV radiation with the line of sight, with some systems that can appear completely UV-obscured by dust absorption. However, we predict them to be very bright at MIR wavelengths, due to the presence of a warm dust component in their surroundings.

- *How did the AGN activity affect the host galaxy?* We find AGN to suppress the star formation rate in their host galaxy via kinetic feedback. Feedback is also responsible of reducing the gas column density at later times, clearing lines of sight from dust obscuration. On

the other hand, AGN radiation heats the dust in the host galaxy, increasing the overall flux in the IR. As a result, we caution that the SFR of the host galaxy measured from its IR emission can be over-estimated by over an order of magnitude.

New and planned science missions, such as the recently-launched James Webb Space Telescope (JWST), Athena, AXIS, the Origin Space Telescope (OST), Roman and Euclid, planned to operate in the next decades, will provide an invaluable amount of data related with the formation and growth of the first BHs in the Universe. The work presented in this Thesis constitutes a first step in order to make a bridge between the data of next generation telescopes and the predictions from new and more sophisticated cosmological simulations. These objects would be missed by currently employed UV-based selection methods, which are heavily affected by dust extinction, and would require observations in different bands (e.g., X-ray or infrared) to be unveiled.

In particular, we highlight the importance of a new generation of MIR and X-ray telescopes, in order to unveil dust-obscured and faint AGN. These objects would be missed by currently employed UV-based selection methods, which are heavily affected by dust extinction, and would require observations in different bands (e.g., X-ray or infrared) to be unveiled. This will significantly increase the population of known $z > 6$ SMBHs and help to understand the processes regulating their growth. Future X-ray observatories will provide us with the sensitivity required to investigate the possible presence of multiple faint AGN satellites around luminous QSOs at high redshift. The upcoming James Webb Space Telescope (JWST) will significantly improve current rest-frame optical/NIR data with deeper observations and it will provide us with high-resolution spectra, effectively probing the spectral region more sensitive to dust attenuation with an unprecedented sensitivity. This will drive a significant contribution in our understanding of the dust origin in the early Universe.

Supplementary material

A

A.1 Convergence of the dust grid

The dust content derived from the hydrodynamical simulations is distributed in an octree grid with a maximum of 8 level of refinement, achieving a maximum resolution of $\sim$ 234 pc, as described in Section 5.4. This spatial resolution is comparable with the resolution of the hydrodynamical simulations, i.e. $\sim$ 200 pc at $z = 6$. In this Section, we check if the number of refinement levels adopted in our fiducial setup is sufficient to achieve converge of the results. We perform three control simulations, in which the maximum refinement levels are 6, 7 and 9, corresponding to a spatial resolution of 937 pc, 469 pc and 117 pc respectively. In Fig. A.1, we show the SED plot for the *AGNcone* run, adopting $f_d = 0.08$ and the fiducial AGN SED, for the aforementioned values of the maximum refinement levels. The four SEDs mainly differ in the MIR range ($6 - 15$ μm rest-frame). The MIR emission increases when increasing the number of refinement levels, because dust around AGN, which is heated to the highest temperatures, is better resolved. However, the variation between our fiducial model and the model at the highest resolution is less than 30% in the MIR band, thus we conclude that the spatial resolution of the dust grid adopted in our calculations is sufficient to achieve reasonable numerical convergence.

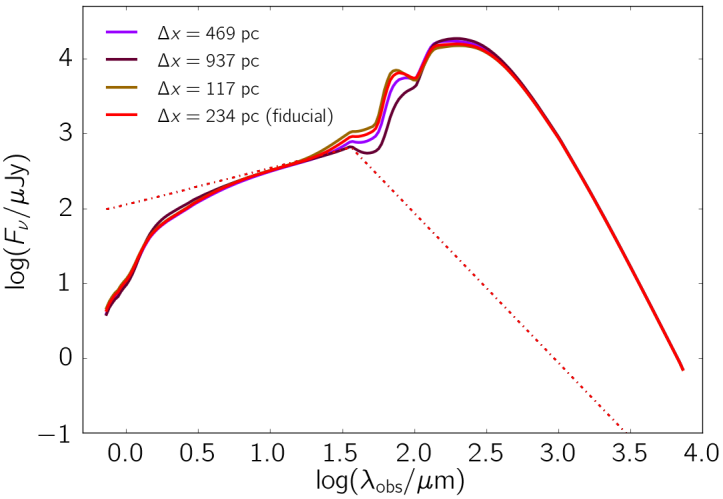

Figure A.1: Spectral Energy Distribution of the *AGNcone* run ($f_d = 0.08$, fiducial AGN SED) for different numbers of the maximum refinement levels: 6, 7, 8 (fiducial) and 9, corresponding to 937 pc, 469 pc, 234 pc (fiducial) and 117 pc, respectively.

A.2 Dust thermal sputtering

We have assumed that dust grains with temperature above a given threshold ($T > 10^6$ K) are destroyed by thermal sputtering (Draine & Salpeter, 1979; Tielens et al., 1994; Hirashita, 2015), as commonly done in simulations (Liang et al., 2019; Ma et al., 2019). However, this dust destruction process may be inefficient in the proximity of AGN, because of grain charging (Tazaki & Ichikawa, 2020; Tazaki et al., 2020). To quantify how this assumption affects our results we re-run the *AGNcone* model with the lower dust content, i.e. $f_d = 0.08$ (fifth row in Table 6.2), after removing the threshold on the dust temperature. In this case, the mass of emitting dust is a factor ~ 2 higher with respect to the fiducial run ($M_{\rm d} = 6 \times 10^7$ M$_\odot$). In Fig. A.2 we compare the SEDs obtained with $f_d = 0.08$ and $f_d = 0.3$ (red and brown lines, respectively) with the model in which dust sputtering is ignored (grey line). The higher dust mass in the model without dust sputtering increases both the attenuation in the UV and the re-emission in the FIR. The resulting SED lies between the $f_d = 0.08$ and $f_d = 0.3$ model results, underlining that the temperature threshold adopted does not affect significantly the main results of our work.

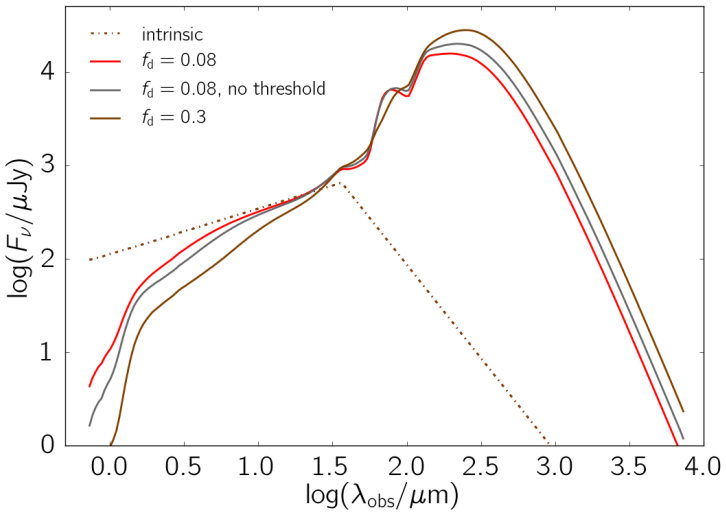

Figure A.2: Comparison of the SEDs of the *AGNcone* run, assuming $f_d = 0.08$ (red), $f_d = 0.3$ (brown) and $f_d = 0.08$ without dust sputtering (grey).

A.3 $\dot{M}_{\rm BH} - M_{\rm UV}$ relation

For a radiation efficiency $\epsilon_r = 0.1$, the bolometric luminosity $L_{\rm bol}$ can be related to the BH accretion rate as follows:

$$L_{\rm bol} \approx 1.5 \times 10^{12} \left(\frac{\dot{M}_{\rm BH}}{\rm M_\odot\ yr^{-1}} \right) \rm L_\odot. \tag{A.1}$$

Using the bolometric corrections reported in Table 5.2, we can convert the accretion rate into an UV luminosity by multiplying the bolometric luminosity by a factor[1] $f_{\rm UV} = L_{\rm UV}/L_{\rm bol}$. Then, we adopt the definition of the AB magnitude

$$m_{\rm AB} = -2.5 \log F_\nu - 48.6,$$

[1] In the case of the fiducial SED, $f_{\rm UV} \approx 0.29$. The results are however very similar in the case of the UV-steep SED.

where F_ν is in cgs units, and we express $M_{\rm UV}$ in terms of the product λL_λ:

$$M_{\rm UV} = 89.9 - 2.5 \log \left(\frac{\lambda L_\lambda}{\rm erg\ s^{-1}} \right), \tag{A.2}$$

where[2] ν and λL_λ are evaluated at $\lambda = 1450$ Å. By combining the previous equations we obtain:

$$M_{\rm UV} = -23.1 - 2.5 \log_{10} \left(\frac{\dot{M}_{\rm BH}}{\rm M_\odot\ yr^{-1}} \right). \tag{A.3}$$

A.4 Dust properties at different redshifts

The analysis shown in Chapter 8 is focused on the snapshot at $z = 6.3$. In order to check that the conclusion we derive are not sensitive to this particular choice, we also post-process the snapshots at $z = 6.2, 6.7, 7$, exploring moments in the halo history, characterised by a different AGN activity. In particular: at $z = 7$ a single AGN with $\dot{M}_{\rm BH} = 7\ {\rm M_\odot\ yr^{-1}}$ is present; at $z = 6.7$, an AGN with $\dot{M}_{\rm BH} = 16\ {\rm M_\odot\ yr^{-1}}$ is in place, with other two much less active ones with $\dot{M}_{\rm BH} \lesssim 2\ {\rm M_\odot\ yr^{-1}}$; at $z = 6.2$, a very powerful AGN with $\dot{M}_{\rm BH} = 68\ {\rm M_\odot\ yr^{-1}}$ dominate the emission, with other two AGN accreting at $\dot{M}_{\rm BH} \lesssim 2\ {\rm M_\odot\ yr^{-1}}$. In each of these snapshots the intrinsic UV emission is provided mainly by a single source.

In Fig. A.3 we show the SEDs at these snapshots for a dust model with $f_{\rm d} = 0.08$ and an SMC extinction curve (left panel) and for the best-fit model found in Section 8.4.1 (right panel). We find that the SEDs produced with the former dust model fail to reproduce the UV data, as for the snapshot at $z = 6.3$. Instead, the discrepancy with the observations is significantly reduced for the best-fit model.

The models that do not agree well with the data have a lower intrinsic bolometric luminosity with respect to the observations, suggesting that the mismatch is due to the fact that the observed quasars are probably in a more active phase with respect to the AGN in the simulation at these snapshots.

This analysis confirms that the need of less extinction at short wavelength is not special of the snapshot at $z = 6.3$.

A.5 Collection of high-redshift quasars

In the work presented in this Thesis, we compare the results of our model with multi-wavelength (NIR to FIR) observations of $z \sim 6$ bright ($-29 \lesssim M_{\rm UV} \lesssim -26$) quasars. The objects used for this comparison are shown in Table A.1 below.

[2] In this expression we do not include k-corrections for the distance modulus (μ) calculations. At $z = 6.3$, the difference between μ and the k-corrected one μ_k is $\mu = \mu_k + 2.1$.

source	z	$M_{\rm UV}$	Reference
J1030+0524	6.31	-27.12	[1-2,8]
J1048+4637	6.23	-27.60	[1-2,8]
J1148+5251	6.43	-27.85	[1-8]
J1306+0356	6.03	-26.76	[1-2,8-9]
J1602+4228	6.07	-26.85	[1-2,8]
J1623+3112	6.25	-26.71	[1-2,8]
J1630+4012	6.07	-26.16	[1-2,8]
J0353+0104	6.07	-26.56	[8]
J0818+1722	6.00	-27.44	[8]
J0842+1218	6.08	-26.85	[8,9]
J1137+3549	6.01	-27.15	[8]
J1250+3130	6.13	-27.18	[8]
J1427+3312	6.12	-26.48	[8]
J2054-0005	6.04	-26.15	[8]
P007+04	6.00	-26.58	[9]
P009-10	6.00	-26.50	[9]
J0142-3327	6.34	-27.76	[9]
P065-26	6.19	-27.21	[9]
P065-19	6.12	-26.57	[9]
J0454-4448	6.06	-26.41	[9]
P159-02	6.38	-26.74	[9]
J1048-0109	6.68	-25.96	[9]
J1148+0702	6.34	-26.43	[9]
J1207+0630	6.04	-26.57	[9]
P183+05	6.44	-26.99	[9]
P217-16	6.15	-26.89	[9]
J1509-1749	6.12	-27.09	[9]
P231-20	6.59	-27.14	[9]
P308-21	6.23	-26.30	[9]
J2211-3206	6.34	-26.65	[9]
J2318-3113	6.44	-26.06	[9]
J2318-3029	6.15	-26.16	[9]
P359-06	6.17	-26.74	[9]
J0100+2802	6.33	-29.30	[10]
P338+29	6.66	-26.01	[14]
J0305-3150	6.61	-26.13	[15]

Table A.1: Quasars used for the comparison with the prediction by our model. Columns indicate: (first) source name, (second) redshift, (third) $M_{\rm UV}$ and (fourth) references for the photometric data used in the comparison, according to the legend. [1] Gallerani et al. (2010); [2] Juarez et al. (2009); [3] Walter et al. (2003); [4] Bertoldi et al. (2003b); [5] Riechers et al. (2009); [6] Gallerani et al. (2014); [7] Stefan et al. (2015); [8] Leipski et al. (2014); [9] Venemans et al. (2018); [10] (Wang et al., 2016); [11] (Venemans et al., 2012); [12] (Venemans et al., 2017b); [13] (Willott et al., 2017); [14] (Mazzucchelli et al., 2017b); [15] (Venemans et al., 2016).

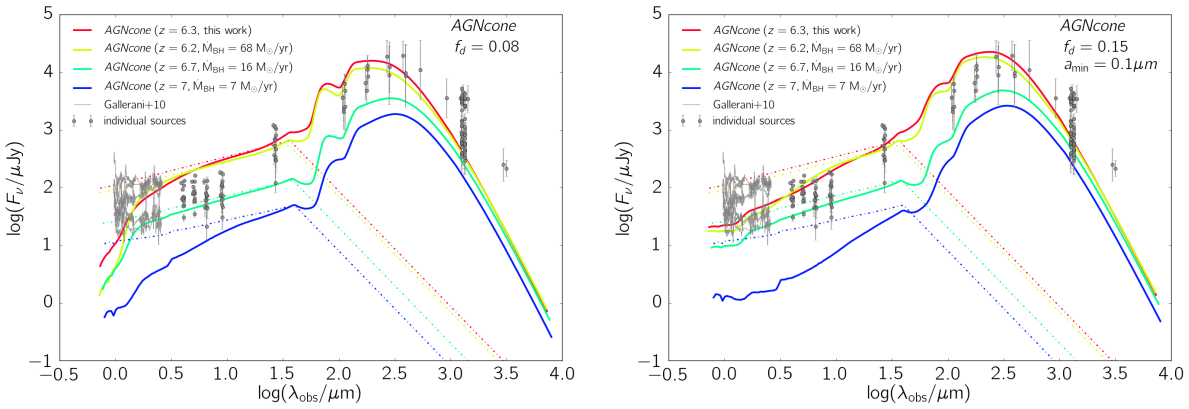

Figure A.3: Comparison of the predicted SEDs with quasar observations for the snapshots at $z = 6.2, 6.3, 6.7, 7$ (with $z = 6.3$ being the snapshot analyzed in this work). The left panel shows the result obtained with a dust model with $f_{\rm d} = 0.08$ and SMC extinction curve. The right panel refers to the best-fit model, characterised by $f_{\rm d} = 0.15$, SMC extinction curve and minimum grain size $a_{\rm min} = 0.1\ \mu$m. The quasar data shown refer to the sources in Table A.1.

Ringraziamenti

Bibliography

Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, ApJ, 770, 25

Aird J., et al., 2010, MNRAS, 401, 2531

Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, MNRAS, 451, 1892

Alam S., et al., 2015, ApJS, 219, 12

Alcock C., et al., 2000, ApJ, 542, 281

Allevato V., et al., 2016, ApJ, 832, 70

Alpher R. A., Bethe H., Gamow G., 1948, Physical Review, 73, 803

Anglés-Alcázar D., Faucher-Giguère C.-A., Quataert E., Hopkins P. F., Feldmann R., Torrey P., Wetzel A., Keres D., 2017, MNRAS, 472, L109

Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046

Antonucci R., 1993, ARA&A, 31, 473

Aoyama S., Hou K.-C., Shimizu I., Hirashita H., Todoroki K., Choi J.-H., Nagamine K., 2017, MNRAS, 466, 105

Aoyama S., Hirashita H., Nagamine K., 2020, MNRAS, 491, 3844

Arata S., Yajima H., Nagamine K., Li Y., Khochfar S., 2019, MNRAS, 488, 2629

Arnaud K. A., 1996, in Jacoby G. H., Barnes J., eds, Astronomical Society of the Pacific Conference Series Vol. 101, Astronomical Data Analysis Software and Systems V. p. 17

Asano R. S., Takeuchi T. T., Hirashita H., Inoue A. K., 2013a, Earth, Planets, and Space, 65, 213

Asano R. S., Takeuchi T. T., Hirashita H., Nozawa T., 2013b, MNRAS, 432, 637

Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481

Bañados E., et al., 2014, AJ, 148, 14

Bañados E., et al., 2016, ApJS, 227, 11

Bañados E., et al., 2018, Nature, 553, 473

Baes M., Camps P., 2015, Astronomy and Computing, 12, 33

Baes M., et al., 2003, MNRAS, 343, 1081

Bakx T. J. L. C., et al., 2020, MNRAS, 493, 4294

Baldwin J. A., Phillips M. M., Terlevich R., 1981, Publ. Astr. Soc. Pac., 93, 5

Barai P., et al., 2013, MNRAS, 430, 3213

Barai P., Gallerani S., Pallottini A., Ferrara A., Marconi A., Cicone C., Maiolino R., Carniani S., 2018, MNRAS, 473, 4003

Barkana R., Loeb A., 2001, Phys. Rep., 349, 125

Barmby P., et al., 2006, ApJ, 642, 126

Barvainis R., 1987, ApJ, 320, 537

Beelen A., Cox P., Benford D. J., Dowell C. D., Kovács A., Bertoldi F., Omont A., Carilli C. L., 2006, ApJ, 642, 694

Behrens C., Pallottini A., Ferrara A., Gallerani S., Vallini L., 2018, MNRAS, 477, 552

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57

Berger M. J., Oliger J., 1984, Journal of Computational Physics, 53, 484

Bergström L., Goobar A., 2004, Cosmology and Particle Astrophysics. Springer, doi:10.1007/3-540-37719-0

Berta S., et al., 2013, A&A, 551, A100

Bertoldi F., Carilli C. L., Cox P., Fan X., Strauss M. A., Beelen A., Omont A., Zylka R., 2003a, A&A, 406, L55

Bertoldi F., et al., 2003b, A&A, 409, L47

Bertone G., 2010, Nature, 468, 389

Bertone G., Hooper D., Silk J., 2005, Phys. Rep., 405, 279

Betti M. G., et al., 2019, JCAP, 2019, 047

Bianchi S., Schneider R., 2007, MNRAS, 378, 973

Bianchi S., Ferrara A., Giovanardi C., 1996, ApJ, 465, 127

Bianchi S., Ferrara A., Davies J. I., Alton P. B., 2000, MNRAS, 311, 601

Biffi V., et al., 2016, ApJ, 827, 112

Binney J., Tabor G., 1995, MNRAS, 276, 663

Bischetti M., Maiolino R., Carniani S., Fiore F., Piconcelli E., Fluetsch A., 2019, A&A, 630, A59

Bischetti M., et al., 2022, arXiv e-prints, p. arXiv:2205.00021

Blecha L., Snyder G. F., Satyapal S., Ellison S. L., 2018, MNRAS, 478, 3056

Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, MNRAS, 456, 2361

Bohren C. F., Huffman D. R., 1984, Nature, 307, 575

Bolmer J., Greiner J., Krühler T., Schady P., Ledoux C., Tanvir N. R., Levan A. J., 2018, A&A, 609, A62

Bondi H., 1952, MNRAS, 112, 195

Bondi H., Hoyle F., 1944, MNRAS, 104, 273

Bongiorno A., et al., 2012, MNRAS, 427, 3103

Booth C. M., Schaye J., 2009, MNRAS, 398, 53

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150

Brandt W. N., Alexander D. M., 2015, A&A Rev., 23, 1

Bromm V., 2013, Reports on Progress in Physics, 76, 112901

Bromm V., Loeb A., 2003, ApJ, 596, 34

Brusa M., et al., 2015, A&A, 578, A11

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

Bryan G. L., et al., 2014, ApJS, 211, 19

Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, ApJ, 429, 582

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Camps P., Baes M., 2015, Astronomy and Computing, 9, 20

Camps P., Trayford J. W., Baes M., Theuns T., Schaller M., Schaye J., 2016, MNRAS, 462, 1057

Carilli C. L., Walter F., 2013, ARA&A, 51, 105

Carnall A. C., et al., 2015, MNRAS, 451, L16

Carniani S., et al., 2016, A&A, 591, A28

Carniani S., et al., 2019a, MNRAS, 489, 3939

Carniani S., et al., 2019b, MNRAS, 489, 3939

Carr B. J., Kohri K., Sendouda Y., Yokoyama J., 2010, Phys. Rev. D, 81, 104019

Cattaneo A., et al., 2009, Nature, 460, 213

Cazaux S., Tielens A. G. G. M., 2004, ApJ, 604, 222

Chabrier G., 2003, Publ. Astr. Soc. Pac., 115, 763

Chakrabarti S., Whitney B. A., 2009, ApJ, 690, 1432

Chakrabarti S., Cox T. J., Hernquist L., Hopkins P. F., Robertson B., Di Matteo T., 2007, ApJ, 658, 840

Chandrasekhar S., 1960, Radiative transfer. Dover Publ. Inc.

Chehade B., et al., 2018, MNRAS, 478, 1649

Chokshi A., Tielens A. G. G. M., Hollenbach D., 1993, ApJ, 407, 806

Chung S. M., et al., 2014, ApJ, 790, 54

Ciardi B., Ferrara A., Marri S., Raimondo G., 2001, MNRAS, 324, 381

Cicone C., et al., 2015, A&A, 574, A14

Ciotti L., Ostriker J. P., 2007, ApJ, 665, 1038

Ciotti L., Ostriker J. P., Proga D., 2010, ApJ, 717, 708

Clements D. L., Serjeant S., Jin S., 2020, Nature, 587, 548

Code A. D., 1973, in Greenberg J. M., van de Hulst H. C., eds, IAU Symposium Vol. 52, Interstellar Dust and Related Topics. p. 505

Colella P., Woodward P. R., 1984, Journal of Computational Physics, 54, 174

Connor T., et al., 2019, ApJ, 887, 171

Connor T., et al., 2020, arXiv e-prints, p. arXiv:2007.14571

Corbelli E., Salucci P., 2000, MNRAS, 311, 441

Costa T., Sijacki D., Trenti M., Haehnelt M. G., 2014a, MNRAS, 439, 2146

Costa T., Sijacki D., Haehnelt M. G., 2014b, MNRAS, 444, 2355

Costa T., Pakmor R., Springel V., 2020, arXiv e-prints, p. arXiv:2006.05997

Cowie L. L., Barger A. J., Bauer F. E., González-López J., 2020, ApJ, 891, 69

Cresci G., et al., 2015, A&A, 582, A63

Croton D. J., et al., 2006, MNRAS, 365, 11

Czerny B., Li J., Loska Z., Szczerba R., 2004, MNRAS, 348, L54

Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431

Davé R., Thompson R., Hopkins P. F., 2016, MNRAS, 462, 3265

Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827

Davies F. B., Hennawi J. F., Eilers A.-C., 2019, ApJL, 884, L19

Dayal P., Ferrara A., 2018, Phys. Rep., 780, 1

DeBoer D. R., et al., 2017, Publ. Astr. Soc. Pac., 129, 045001

Decarli R., et al., 2017, Nature, 545, 457

Decarli R., et al., 2018, ApJ, 854, 97

Decataldo D., Lupi A., Ferrara A., Pallottini A., Fumagalli M., 2020, MNRAS, 497, 4718

Delvecchio I., et al., 2014, MNRAS, 439, 2736

Devecchi B., Volonteri M., 2009, ApJ, 694, 302

Devroye L., 1986, Non-Uniform Random Variate Generation. Springer-Verlag

Di Mascia F., et al., 2021a, MNRAS, 503, 2349

Di Mascia F., Gallerani S., Ferrara A., Pallottini A., Maiolino R., Carniani S., D'Odorico V., 2021b, MNRAS, 506, 3946

Di Matteo T., Croft R. A. C., Feng Y., Waters D., Wilkins S., 2017, MNRAS, 467, 4243

Draine B., 1989, in Allamandola L. J., Tielens A. G. G. M., eds, Vol. 135, Interstellar Dust. p. 313

Draine B. T., 2003a, ARA&A, 41, 241

Draine B. T., 2003b, ApJ, 598, 1017

Draine B. T., 2003c, ApJ, 598, 1026

Draine B. T., 2009, in Henning T., Grün E., Steinacker J., eds, Astronomical Society of the Pacific Conference Series Vol. 414, Cosmic Dust - Near and Far. p. 453 (arXiv:0903.1658)

Draine B. T., Bertoldi F., 1996, ApJ, 468, 269

Draine B. T., Lee H. M., 1984, ApJ, 285, 89

Draine B. T., Salpeter E. E., 1979, ApJ, 231, 77

Draine B. T., et al., 2007, ApJ, 663, 866

Dubois Y., Gavazzi R., Peirani S., Silk J., 2013, MNRAS, 433, 3297

Dubois Y., et al., 2014, MNRAS, 444, 1453

Duras F., et al., 2020, A&A, 636, A73

Dwek E., 1998, ApJ, 501, 643

Dwek E., Scalo J. M., 1980, ApJ, 239, 193

Egami E., et al., 2018, Publ. Astr. Soc. Australia, 35, 48

Einstein A., 1905, Annalen der Physik, 322, 891

Einstein A., 1916, Annalen der Physik, 354, 769

Eisenstein D. J., Hu W., 1999, ApJ, 511, 5

Eisenstein D. J., Loeb A., 1995, ApJ, 443, 11

Elvis M., et al., 1994, ApJS, 95, 1

Elvis M., Marengo M., Karovska M., 2002, ApJL, 567, L107

Fabian A. C., 1994, ARA&A, 32, 277

Fabian A. C., Vasudevan R. V., Gandhi P., 2008, MNRAS, 385, L43

Fan X., et al., 2000, AJ, 120, 1167

Fan X., et al., 2006, AJ, 131, 1203

Feng Y., Di Matteo T., Croft R., Khandai N., 2014, MNRAS, 440, 1865

Feng Y., Di-Matteo T., Croft R. A., Bird S., Battaglia N., Wilkins S., 2016, MNRAS, 455, 2778

Ferrara A., Salvadori S., Yue B., Schleicher D., 2014, MNRAS, 443, 2410

Ferrara A., Viti S., Ceccarelli C., 2016, MNRAS, 463, L112

Ferrarese L., Merritt D., 2000, ApJL, 539, L9

Ferrarotti A. S., Gail H. P., 2006, A&A, 447, 553

Fields B. D., Olive K. A., Yeh T.-H., Young C., 2020, JCAP, 2020, 010

Fiore F., Elvis M., McDowell J. C., Siemiginowska A., Wilkes B. J., 1994, ApJ, 431, 515

Fiore F., et al., 2017, A&A, 601, A143

Fitzpatrick E., 1989, in Allamandola L. J., Tielens A. G. G. M., eds, IAU Symposium Vol. 135, Interstellar Dust. p. 37

Fitzpatrick E. L., 1999, Publ. Astr. Soc. Pac., 111, 63

Fitzpatrick E. L., Massa D., 1986, ApJ, 307, 286

Follin B., Knox L., Millea M., Pan Z., 2015, Phys. Rev. Lett., 115, 091301

Frank J., King A., Raine D. J., 2002, Accretion Power in Astrophysics: Third Edition. Cambridge University Press

Frenklach M., Feigelson E. D., 1989, ApJ, 341, 372

Frenklach M., Carmer C. S., Feigelson E. D., 1989, Nature, 339, 196

Friedmann A., 1922, Zeitschrift fur Physik, 10, 377

Fruscione A., et al., 2006, CIAO: Chandra's data analysis system. D.R. Silvia & R.E. Doxsey eds, p. 62701V, doi:10.1117/12.671760

Fujimoto S., et al., 2020, ApJ, 900, 1

Galeev A. A., Rosner R., Vaiana G. S., 1979, ApJ, 229, 318

Gall C., Hjorth J., Andersen A. C., 2011, A&A Rev., 19, 43

Gallerani S., et al., 2010, A&A, 523, A85

Gallerani S., Ferrara A., Neri R., Maiolino R., 2014, MNRAS, 445, 2848

Gallerani S., Fan X., Maiolino R., Pacucci F., 2017a, Publ. Astr. Soc. Australia, 34, e022

Gallerani S., et al., 2017b, MNRAS, 467, 3590

Gallerani S., Pallottini A., Feruglio C., Ferrara A., Maiolino R., Vallini L., Riechers D. A., Pavesi R., 2018, MNRAS, 473, 1909

García-Burillo S., et al., 2014, A&A, 567, A125

García-Burillo S., et al., 2016, ApJL, 823, L12

García-Burillo S., et al., 2019, A&A, 632, A61

Gaskell C. M., Benker A. J., 2007, arXiv e-prints, p. arXiv:0711.1013

Gaskell C. M., Goosmann R. W., Antonucci R. R. J., Whysong D. H., 2004, ApJ, 616, 147

Gaskin J. A., et al., 2019, Journal of Astronomical Telescopes, Instruments, and Systems, 5, 021001

Gebhardt K., et al., 2000, ApJL, 539, L13

Gilli R., et al., 2014, A&A, 562, A67

Ginolfi M., et al., 2020, A&A, 633, A90

Giustini M., Proga D., 2019, A&A, 630, A94

Gnedin N. Y., 2014, ApJ, 793, 29

Godunov S. K., 1959, Mat. Sb., Nov. Ser., 47, 271

González M., Audit E., Huynh P., 2007, A&A, 464, 429

Gordon K. D., Calzetti D., Witt A. N., 1997, ApJ, 487, 625

Gordon K. D., Misselt K. A., Witt A. N., Clayton G. C., 2001, ApJ, 551, 269

Greiner J., Bolmer J., Yates R. M., Habouzit M., Bañados E., Afonso P. M. J., Schady P., 2021, A&A, 654, A79

Group P. D., et al., 2020, Progress of Theoretical and Experimental Physics, 2020

Groves B., Dopita M. A., Sutherland R. S., Kewley L. J., Fischera J., Leitherer C., Brandl B., van Breugel W., 2008, ApJS, 176, 438

Gruppioni C., et al., 2016, MNRAS, 458, 4297

Gruppioni C., et al., 2017, Publ. Astr. Soc. Australia, 34, e055

Gültekin K., et al., 2009, ApJ, 698, 198

Habouzit M., Volonteri M., Latif M., Dubois Y., Peirani S., 2016, MNRAS, 463, 529

Habouzit M., Volonteri M., Somerville R. S., Dubois Y., Peirani S., Pichon C., Devriendt J., 2019, MNRAS, 489, 1206

Habouzit M., et al., 2021, MNRAS,

Hahn O., Abel T., 2011, MNRAS, 415, 2101

Haiman Z., 2013, in Wiklind T., Mobasher B., Bromm V., eds, , Vol. 396, The First Galaxies. Astrophysics and Space Science Library, p. 293 (arXiv:1203.6075)

Hamann F., Korista K. T., Ferland G. J., Warner C., Baldwin J., 2002, ApJ, 564, 592

Harrison C. M., 2017, Nature Astronomy, 1, 0165

Harrison C. M., Costa T., Tadhunter C. N., Flütsch A., Kakkad D., Perna M., Vietri G., 2018, Nature Astronomy, 2, 198

Heger A., Woosley S. E., 2002, ApJ, 567, 532

Heger A., Fryer C. L., Woosley S. E., Langer N., Hartmann D. H., 2003, ApJ, 591, 288

Henyey L. G., Greenstein J. L., 1941, ApJ, 93, 70

Hernán-Caballero A., Hatziminaoglou E., 2011, MNRAS, 414, 500

Hernán-Caballero A., Hatziminaoglou E., Alonso-Herrero A., Mateos S., 2016, MNRAS, 463, 2064

Herrera-Camus R., et al., 2018, ApJ, 861, 95

Hickox R. C., Alexander D. M., 2018, ARA&A, 56, 625

Hirano S., Hosokawa T., Yoshida N., Umeda H., Omukai K., Chiaki G., Yorke H. W., 2014, ApJ, 781, 60

Hirashita H., 2015, MNRAS, 447, 2937

Hirashita H., Aoyama S., 2019, MNRAS, 482, 2555

Hirashita H., Ferrara A., 2002, MNRAS, 337, 921

Hirashita H., Yan H., 2009, MNRAS, 394, 1061

Hirashita H., Nozawa T., Kozasa T., Ishii T. T., Takeuchi T. T., 2005, MNRAS, 357, 1077

Hirashita H., Nozawa T., Takeuchi T. T., Kozasa T., 2008, MNRAS, 384, 1725

Hjorth J., Vreeswijk P. M., Gall C., Watson D., 2013, ApJ, 768, 173

Hoffman Y., Ribak E., 1991, ApJL, 380, L5

Hönig S. F., 2019, ApJ, 884, 171

Hopkins P. F., 2015, MNRAS, 450, 53

Hopkins P. F., et al., 2004, AJ, 128, 1112

Hopkins P. F., Richards G. T., Hernquist L., 2007, ApJ, 654, 731

Hopkins P. F., Hernquist L., Hayward C. C., Narayanan D., 2012, MNRAS, 425, 1121

Hopkins P. F., Keres D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Hosokawa T., Omukai K., Yoshida N., Yorke H. W., 2011, Science, 334, 1250

Hosokawa T., Yorke H. W., Inayoshi K., Omukai K., Yoshida N., 2013, ApJ, 778, 178

Hou K.-C., Hirashita H., Nagamine K., Aoyama S., Shimizu I., 2017, MNRAS, 469, 870

Hoyle F., Lyttleton R. A., 1939, Proceedings of the Cambridge Philosophical Society, 35, 405

Hubble E. P., 1926a, ApJ, 63, 236

Hubble E. P., 1926b, ApJ, 64, 321

Hubble E., 1929a, Proceedings of the National Academy of Science, 15, 168

Hubble E. P., 1929b, ApJ, 69, 103

Inayoshi K., Visbal E., Haiman Z., 2020, ARA&A, 58, 27

Inoue A. K., 2005, MNRAS, 359, 171

Ishibashi W., Fabian A. C., 2015, MNRAS, 451, 93

Izumi T., et al., 2021, ApJ, 914, 36

Jang M., Im M., Lee I., Urata Y., Huang K., Hirashita H., Fan X., Jiang L., 2011, ApJL, 741, L20

Jiang L., Fan X., Vestergaard M., Kurk J. D., Walter F., Kelly B. C., Strauss M. A., 2007, AJ, 134, 1150

Jiang L., et al., 2009, AJ, 138, 305

Jiang L., et al., 2016, ApJ, 833, 222

Jones A. P., Tielens A. G. G. M., Hollenbach D. J., 1996, ApJ, 469, 740

Jonsson P., 2006, MNRAS, 372, 2

Jonsson P., Groves B. A., Cox T. J., 2010, MNRAS, 403, 17

Juarez Y., Maiolino R., Mujica R., Pedani M., Marinoni S., Nagao T., Marconi A., Oliva E., 2009, A&A, 494, L25

Juvela M., 2005, A&A, 440, 531

Kashikawa N., et al., 2015, ApJ, 798, 28

Katz H., Sijacki D., Haehnelt M. G., 2015, MNRAS, 451, 2352

Kauffmann G., et al., 2003, MNRAS, 346, 1055

Kennicutt Robert C. J., 1998, ApJ, 498, 541

Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349

King A., 2003, ApJL, 596, L27

Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102

Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Koss M., Mushotzky R., Treister E., Veilleux S., Vasudevan R., Trippe M., 2012, ApJL, 746, L22

Kozasa T., Hasegawa H., Nomoto K., 1989a, ApJ, 344, 325

Kozasa T., Hasegawa H., Nomoto K., 1989b, ApJL, 346, L81

Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73

Kroupa P., 2001, MNRAS, 322, 231

Kroupa P., 2002, Science, 295, 82

Kroupa P., Weidner C., 2003, ApJ, 598, 1076

Kulkarni G., Worseck G., Hennawi J. F., 2019, MNRAS, 488, 1035

Kurk J. D., et al., 2007, ApJ, 669, 32

Lacy M., et al., 2004, ApJS, 154, 166

Laor A., Draine B. T., 1993, ApJ, 402, 441

Laporte N., et al., 2017, ApJL, 837, L21

Larson R. B., 2005, MNRAS, 359, 211

Latif M. A., Ferrara A., 2016, Publ. Astr. Soc. Australia, 33, e051

Latif M. A., Schleicher D. R. G., Schmidt W., Niemeyer J., 2013, MNRAS, 433, 1607

Leger A., Puget J. L., 1984, A&A, 137, L5

Leipski C., et al., 2013, ApJ, 772, 103

Leipski C., et al., 2014, ApJ, 785, 154

Leitherer C., et al., 1999, ApJS, 123, 3

Leśniewska A., Michałowski M. J., 2019, A&A, 624, L13

Levermore C. D., 1984, JQSRT, 31, 149

Li A., 2007, in Ho L. C., Wang J. W., eds, Astronomical Society of the Pacific Conference Series Vol. 373, The Central Engine of Active Galactic Nuclei. p. 561

Li A., Draine B. T., 2001a, ApJ, 554, 778

Li A., Draine B. T., 2001b, ApJ, 554, 778

Li Y., et al., 2008a, ApJ, 678, 41

Li A., Liang S. L., Kann D. A., Wei D. M., Klose S., Wang Y. J., 2008b, ApJ, 685, 1046

Li J., et al., 2020, ApJ, 889, 162

Liang L., et al., 2019, MNRAS, 489, 1397

Liang L., Feldmann R., Hayward C. C., Narayanan D., Çatmabacak O., Keres D., Faucher-Giguére C.-A., Hopkins P. F., 2021a, MNRAS,

Liang L., Feldmann R., Hayward C. C., Narayanan D., Çatmabacak O., Keres D., Faucher-Giguére C.-A., Hopkins P. F., 2021b, MNRAS,

Linde A. D., 1982, Physics Letters B, 108, 389

Lodato G., Natarajan P., 2006, MNRAS, 371, 1813

Loeb A., Ferrara A., Ellis R. S., 2008, First Light in the Universe. Springer

Lupi A., Haardt F., Dotti M., Fiacconi D., Mayer L., Madau P., 2016, MNRAS, 456, 2993

Lupi A., Volonteri M., Decarli R., Bovino S., Silk J., Bergeron J., 2019, MNRAS, 488, 4004

Lupi A., Volonteri M., Decarli R., Bovino S., Silk J., 2022, MNRAS, 510, 5760

Lusso E., Risaliti G., 2017, A&A, 602, A79

Lusso E., Worseck G., Hennawi J. F., Prochaska J. X., Vignali C., Stern J., O'Meara J. M., 2015, MNRAS, 449, 4204

Lynden-Bell D., 1969, Nature, 223, 690

Lyu J., Rieke G. H., Alberts S., 2016, ApJ, 816, 85

Ma X., et al., 2019, MNRAS, 487, 1844

Maartens R., 2011, Philosophical Transactions of the Royal Society of London Series A, 369, 5115

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Madau P., Haardt F., Dotti M., 2014, ApJL, 784, L38

Magorrian J., et al., 1998a, AJ, 115, 2285

Magorrian J., et al., 1998b, AJ, 115, 2285

Maiolino R., Marconi A., Salvati M., Risaliti G., Severgnini P., Oliva E., La Franca F., Vanzi L., 2001a, A&A, 365, 28

Maiolino R., Marconi A., Oliva E., 2001b, A&A, 365, 37

Maiolino R., Schneider R., Oliva E., Bianchi S., Ferrara A., Mannucci F., Pedani M., Roca Sogorb M., 2004, Nature, 431, 533

Maiolino R., et al., 2005, A&A, 440, L51

Maiolino R., et al., 2012, MNRAS, 425, L66

Manti S., Gallerani S., Ferrara A., Feruglio C., Graziani L., Bernardi G., 2016, MNRAS, 456, 98

Marchesi S., et al., 2020, A&A, 642, A184

Marconi A., Hunt L. K., 2003, ApJL, 589, L21

Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, MNRAS, 351, 169

Markevitch M., 1998, ApJ, 504, 27

Marshall M. A., et al., 2020, ApJ, 900, 21

Mathis J. S., Rumpl W., Nordsieck K. H., 1977, ApJ, 217, 425

Matsuoka K., Nagao T., Marconi A., Maiolino R., Taniguchi Y., 2011, A&A, 527, A100

Matsuoka Y., et al., 2016, ApJ, 828, 26

Matsuoka Y., et al., 2018a, The Astrophysical Journal Supplement Series, 237, 5

Matsuoka Y., et al., 2018b, ApJ, 869, 150

Matsuoka Y., et al., 2019a, ApJ, 872, L2

Matsuoka Y., et al., 2019b, ApJ, 883, 183

Mazzucchelli C., et al., 2017a, ApJ, 849, 91

Mazzucchelli C., et al., 2017b, ApJ, 849, 91

Mazzucchelli C., et al., 2019, ApJ, 881, 163

McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936

McConnell N. J., Ma C.-P., Gebhardt K., Wright S. A., Murphy J. D., Lauer T. R., Graham J. R., Richstone D. O., 2011, Nature, 480, 215

McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499

McGreer I. D., Fan X., Jiang L., Cai Z., 2018, AJ, 155, 131

McKinney J., Hayward C. C., Rosenthal L. J., Martínez-Galarza J. R., Pope A., Sajina A., Smith H. A., 2021, ApJ, 921, 55

Mechtley M., et al., 2012, ApJL, 756, L38

Merloni A., et al., 2014, MNRAS, 437, 3550

Mesinger A., ed. 2016, Understanding the Epoch of Cosmic Reionization Astrophysics and Space Science Library Vol. 423, doi:10.1007/978-3-319-21957-8.

Metropolis N., Ulam S., 1949, Journal of the American Statistical Association, 44, 335

Meyer R. A., et al., 2022, ApJ, 927, 152

Michałowski M. J., Watson D., Hjorth J., 2010, ApJ, 712, 942

Mignoli M., et al., 2020, A&A, 642, L1

Mihalas D., Mihalas B. W., 1984, Foundations of radiation hydrodynamics. New York: Oxford University Press

Milosavljević M., Safranek-Shrader C., 2016, in Mesinger A., ed., Astrophysics and Space Science Library Vol. 423, Understanding the Epoch of Cosmic Reionization: Challenges and Progress. p. 65 (arXiv:1511.01110), doi:10.1007/978-3-319-21957-8_3

Miralda-Escudé J., Haehnelt M., Rees M. J., 2000, ApJ, 530, 1

Mizumoto M., Nomura M., Done C., Ohsuga K., Odaka H., 2021, MNRAS, 503, 1442

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution. Cambridge University Press

Morgan H. L., Edmunds M. G., 2003, MNRAS, 343, 427

Morganti R., Oosterloo T., Oonk J. B. R., Frieswijk W., Tadhunter C., 2015, A&A, 580, A1

Morrison R., McCammon D., 1983, ApJ, 270, 119

Mortlock D. J., et al., 2011, Nature, 474, 616

Moseley S. H., Dwek E., Glaccum W., Graham J. R., Loewenstein R. F., 1989, Nature, 340, 697

Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121

Murante G., Monaco P., Giovalli M., Borgani S., Diaferio A., 2010, MNRAS, 405, 1491

Murphy E. J., et al., 2011, ApJ, 737, 67

Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569

Mushotzky R., et al., 2019, in Bulletin of the American Astronomical Society. p. 107 (arXiv:1903.04083)

Nandra K., et al., 2013, preprint, (arXiv:1306.2307)

Nanni R., Vignali C., Gilli R., Moretti A., Brand t W. N., 2017, A&A, 603, A128

Nanni R., et al., 2018, A&A, 614, A121

Narayanan D., Conroy C., Davé R., Johnson B. D., Popping G., 2018, ApJ, 869, 70

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Neeleman M., et al., 2019, ApJ, 882, 10

Neeleman M., et al., 2021, ApJ, 911, 141

Nelson D., et al., 2019, MNRAS, 490, 3234

Nenkova M., Sirocky M. M., Ivezić Z., Elitzur M., 2008, ApJ, 685, 147

Netzer H., 2013, The Physics and Evolution of Active Galactic Nuclei. Cambridge University Press

Netzer H., 2015, Annual Review of Astronomy and Astrophysics, 53, 365

Ni Y., Di Matteo T., Gilli R., Croft R. A. C., Feng Y., Norman C., 2020, MNRAS, 495, 2135

Novak M., et al., 2019, ApJ, 881, 63

Novak M., et al., 2020, ApJ, 904, 131

Nozawa T., Kozasa T., Umeda H., Maeda K., Nomoto K., 2003, ApJ, 598, 785

Nozawa T., Kozasa T., Habe A., Dwek E., Umeda H., Tominaga N., Maeda K., Nomoto K., 2007, ApJ, 666, 955

Nozawa T., Asano R. S., Hirashita H., Takeuchi T. T., 2015, MNRAS, 447, L16

Oesch P. A., et al., 2016, ApJ, 819, 129

Omukai K., Tsuribe T., Schneider R., Ferrara A., 2005, ApJ, 626, 627

Ono Y., et al., 2018, Pub. Astron. Soc. Japan, 70, S10

Orofino M. C., Ferrara A., Gallerani S., 2021, MNRAS, 502, 2757

Osterbrock D. E., 1989, Astrophysics of gaseous nebulae and active galactic nuclei. University Science Books

Ota K., et al., 2018, ApJ, 856, 109

Overzier R. A., 2021, arXiv e-prints, p. arXiv:2111.05446

Pacucci F., Volonteri M., Ferrara A., 2015a, MNRAS, 452, 1922

Pacucci F., Volonteri M., Ferrara A., 2015b, MNRAS, 452, 1922

Pacucci F., Ferrara A., Grazian A., Fiore F., Giallongo E., Puccetti S., 2016, MNRAS, 459, 1432

Padoan P., Nordlund a., 2011, ApJ, 730, 40

Padoan P., Federrath C., Chabrier G., Evans N. J. I., Johnstone D., Jørgensen J. K., McKee C. F., Nordlund a., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 77 (arXiv:1312.5365), doi:10.2458/azu_uapress_9780816531240-ch004

Pallottini A., Ferrara A., Bovino S., Vallini L., Gallerani S., Maiolino R., Salvadori S., 2017, MNRAS, 471, 4128

Pallottini A., et al., 2022, MNRAS, 513, 5621

Panuzzo P., Granato G. L., Buat V., Inoue A. K., Silva L., Iglesias-Páramo J., Bressan A., 2007, MNRAS, 375, 640

Park J., Mesinger A., Greig B., Gillet N., 2019, MNRAS, 484, 933

Peebles P. J. E., 1982, ApJL, 263, L1

Peng C. Y., 2007, ApJ, 671, 1098

Pensabene A., et al., 2021, A&A, 652, A66

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Perley D. A., et al., 2010, MNRAS, 406, 2473

Perlmutter S., et al., 1998, Nature, 391, 51

Perlmutter S., et al., 1999, ApJ, 517, 565

Peterson B. M., 2006, in Alloin D., ed., , Vol. 693, Physics of Active Galactic Nuclei at all Scales. Springer Berlin Heidelberg, p. 77, doi:10.1007/3-540-34621-X_3

Pezzulli E., Valiante R., Schneider R., 2016, MNRAS, 458, 3047

Pezzulli E., Valiante R., Orofino M. C., Schneider R., Gallerani S., Sbarrato T., 2017a, MNRAS, 466, 2131

Pezzulli E., Valiante R., Orofino M. C., Schneider R., Gallerani S., Sbarrato T., 2017b, MNRAS, 466, 2131

Piconcelli E., Jimenez-Bailón E., Guainazzi M., Schartel N., Rodríguez-Pascual P. M., Santos-Lleó M., 2005, A&A, 432, 15

Pierini D., Gordon K. D., Witt A. N., Madsen G. J., 2004, ApJ, 617, 1022

Pilbratt G. L., et al., 2010, A&A, 518, L1

Pipino A., Fan X. L., Matteucci F., Calura F., Silva L., Granato G., Maiolino R., 2011, A&A, 525, A61

Planck Collaboration et al., 2016, A&A, 594, A13

Planck Collaboration et al., 2020, A&A, 641, A6

Pozzi F., et al., 2012, MNRAS, 423, 1909

Press W. H., Schechter P., 1974, ApJ, 187, 425

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in FORTRAN. The art of scientific computing. Cambridge University Press

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2002, Numerical recipes in C++ : the art of scientific computing. Cambridge University Press

Pringle J. E., 1981, ARA&A, 19, 137

Proga D., 2007, in Ho L. C., Wang J. W., eds, Astronomical Society of the Pacific Conference Series Vol. 373, The Central Engine of Active Galactic Nuclei. p. 267 (arXiv:astro-ph/0701100)

Proga D., Kallman T. R., 2004, ApJ, 616, 688

Proga D., Stone J. M., Kallman T. R., 2000, ApJ, 543, 686

Rahmati A., Pawlik A. H., Raicević M., Schaye J., 2013, MNRAS, 430, 2427

Reed S. L., et al., 2015, MNRAS, 454, 3952

Reed S. L., et al., 2017, MNRAS, 468, 4702

Rees M. J., 1984, ARA&A, 22, 471

Reichard T. A., et al., 2003, AJ, 126, 2594

Ricci C., et al., 2017a, Nature, 549, 488

Ricci C., et al., 2017b, ApJ, 835, 105

Richards G. T., et al., 2003, AJ, 126, 1131

Richardson M. L. A., Scannapieco E., Devriendt J., Slyz A., Thacker R. J., Dubois Y., Wurster J., Silk J., 2016, ApJ, 825, 83

Riechers D. A., et al., 2009, ApJ, 703, 1338

Riess A. G., et al., 1998, AJ, 116, 1009

Riess A. G., et al., 2016, ApJ, 826, 56

Roebuck E., Sajina A., Hayward C. C., Pope A., Kirkpatrick A., Hernquist L., Yan L., 2016, ApJ, 833, 60

Roelfsema P. R., et al., 2018, Publ. Astr. Soc. Australia, 35, e030

Rosdahl J., Teyssier R., 2015, MNRAS, 449, 4380

Rosdahl J., Blaizot J., Aubert D., Stranex T., Teyssier R., 2013, MNRAS, 436, 2188

Rybicki G. B., Lightman A. P., 1979, Radiative processes in astrophysics. Wiley

Saadeh D., Feeney S. M., Pontzen A., Peiris H. V., McEwen J. D., 2016, Phys. Rev. Lett., 117, 131302

Sadowski A., Narayan R., Penna R., Zhu Y., 2013, MNRAS, 436, 3856

Sakurai Y., Yoshida N., Fujii M. S., Hirano S., 2017, MNRAS, 472, 1677

Sala L., Cenci E., Capelo P. R., Lupi A., Dotti M., 2021, MNRAS, 500, 4788

Salim S., Narayanan D., 2020, ARA&A, 58, 529

Salpeter E. E., 1955, ApJ, 121, 161

Salpeter E. E., 1974, ApJ, 193, 579

Sanders D. B., Phinney E. S., Neugebauer G., Soifer B. T., Matthews K., 1989, ApJ, 347, 29

Savage B. D., Mathis J. S., 1979, ARA&A, 17, 73

Sazonov S. Y., Ostriker J. P., Sunyaev R. A., 2004, MNRAS, 347, 144

Schartmann M., Meisenheimer K., Camenzind M., Wolf S., Henning T., 2005, A&A, 437, 861

Schaye J., et al., 2015, MNRAS, 446, 521

Schindler J.-T., et al., 2020, ApJ, 905, 51

Schleicher D. R. G., Palla F., Ferrara A., Galli D., Latif M., 2013, A&A, 558, A59

Schmidt M., 1959, ApJ, 129, 243

Schmidt M., 1963, Nature, 197, 1040

Schmidt B. P., et al., 1998, ApJ, 507, 46

Schneider R., Omukai K., Inoue A. K., Ferrara A., 2006, MNRAS, 369, 1437

Schneider R., Bianchi S., Valiante R., Risaliti G., Salvadori S., 2015, Astronomy and Astrophysics, 579, A60

Seon K.-I., Draine B. T., 2016, ApJ, 833, 201

Seyfert C. K., 1943, ApJ, 97, 28

Shakura N. I., Sunyaev R. A., 1973, A&A, 500, 33

Shang C., Bryan G. L., Haiman Z., 2010, MNRAS, 402, 1249

Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, ApJ, 690, 20

Shen X., Hopkins P. F., Faucher-Giguère C.-A., Alexander D. M., Richards G. T., Ross N. P., Hickox R. C., 2020a, MNRAS, 495, 3252

Shen X., et al., 2020b, MNRAS, 495, 4747

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877

Sijacki D., Springel V., Haehnelt M. G., 2009, MNRAS, 400, 100

Silk J., Rees M. J., 1998, A&A, 331, L1

Silverman J. D., et al., 2020, arXiv e-prints, p. arXiv:2007.05581

Skillman S. W., Warren M. S., Turk M. J., Wechsler R. H., Holz D. E., Sutter P. M., 2014, arXiv e-prints, p. arXiv:1407.2600

Smidt J., Whalen D. J., Johnson J. L., Surace M., Li H., 2018, ApJ, 865, 126

Smoot G. F., et al., 1992, ApJL, 396, L1

Snyder G. F., Hayward C. C., Sajina A., Jonsson P., Cox T. J., Hernquist L., Hopkins P. F., Yan L., 2013, ApJ, 768, 168

Soltan A., 1982, MNRAS, 200, 115

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008a, MNRAS, 391, 481

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008b, MNRAS, 391, 481

Sommovigo L., Ferrara A., Pallottini A., Carniani S., Gallerani S., Decataldo D., 2020, MNRAS, 497, 956

Sommovigo L., Ferrara A., Carniani S., Zanella A., Pallottini A., Gallerani S., Vallini L., 2021a, MNRAS, 503, 4878

Sommovigo L., Ferrara A., Carniani S., Zanella A., Pallottini A., Gallerani S., Vallini L., 2021b, MNRAS, 503, 4878

Spergel D. N., et al., 2003, ApJS, 148, 175

Spinoglio L., et al., 2017, Publ. Astr. Soc. Australia, 34, e057

Spitzer L. J., 1978, Physical Processes in the Interstellar Medium. Wiley

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2010, MNRAS, 401, 791

Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776

Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137

Springel V., et al., 2018, MNRAS, 475, 676

Stalevski M., Fritz J., Baes M., Nakos T., Popović L. C., 2012, MNRAS, 420, 2756

Stalevski M., Ricci C., Ueda Y., Lira P., Fritz J., Baes M., 2016, MNRAS, 458, 2288

Stefan I. I., et al., 2015, MNRAS, 451, 1713

Steinacker J., Bacmann A., Henning T., 2002, JQSRT, 75, 765

Steinacker J., Baes M., Gordon K. D., 2013, ARA&A, 51, 63

Stenholm L. G., Stoerzer H., Wehrse R., 1991, JQSRT, 45, 47

Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074

Stratta G., Maiolino R., Fiore F., D'Elia V., 2007, ApJL, 661, L9

Stratta G., Gallerani S., Maiolino R., 2011, A&A, 532, A45

Symeonidis M., Page M. J., 2021, MNRAS, 503, 3992

Symeonidis M., et al., 2022, arXiv e-prints, p. arXiv:2205.11645

Talbot R. Y., Sijacki D., Bourne M. A., 2021, arXiv e-prints, p. arXiv:2111.01801

Tanaka T., Haiman Z., 2009, ApJ, 696, 1798

Tang J.-J., et al., 2017, MNRAS, 466, 4568

Tazaki R., Ichikawa K., 2020, ApJ, 892, 149

Tazaki R., Ichikawa K., Kokubo M., 2020, ApJ, 892, 84

Teyssier R., 2002, A&A, 385, 337

Thompson T. A., Quataert E., Murray N., 2005, ApJ, 630, 167

Tielens A. G. G. M., McKee C. F., Seab C. G., Hollenbach D. J., 1994, ApJ, 431, 321

Todini P., Ferrara A., 2001, MNRAS, 325, 726

Tombesi F., Cappi M., Reeves J. N., Palumbo G. G. C., Yaqoob T., Braito V., Dadina M., 2010, A&A, 521, A57

Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, MNRAS, 382, 1050

Trayford J. W., et al., 2017, MNRAS, 470, 771

Trebitsch M., Volonteri M., Dubois Y., 2019, MNRAS, 487, 819

Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, MNRAS, 470, 1121

Trumpler R. J., 1930, Publ. Astr. Soc. Pac., 42, 214

Trussler J., Maiolino R., Maraston C., Peng Y., Thomas D., Goddard D., Lian J., 2020, MNRAS, 491, 5406

Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886

Urry C. M., Padovani P., 1995, Publ. Astr. Soc. Pac., 107, 803

Valentini M., et al., 2020, MNRAS, 491, 2779

Valentini M., Gallerani S., Ferrara A., 2021, MNRAS, 507, 1

Valiante R., Schneider R., Bianchi S., Andersen A. C., 2009, MNRAS, 397, 1661

Valiante R., Schneider R., Salvadori S., Bianchi S., 2011, MNRAS, 416, 1916

Valiante R., Agarwal B., Habouzit M., Pezzulli E., 2017, Publ. Astr. Soc. Australia, 34, e031

Vallini L., Ferrara A., Pallottini A., Gallerani S., 2017, MNRAS, 467, 1300

Vanden Berk D. E., et al., 2001, The Astronomical Journal, 122, 549–564

Városi F., Dwek E., 1999, ApJ, 523, 265

Veilleux S., Osterbrock D. E., 1987, ApJS, 63, 295

Venanzi M., Hönig S., Williamson D., 2020, arXiv e-prints, p. arXiv:2007.13554

Venemans B. P., McMahon R. G., Warren S. J., Gonzalez-Solares E. A., Hewett P. C., Mortlock D. J., Dye S., Sharp R. G., 2007, MNRAS, 376, L76

Venemans B. P., et al., 2012, ApJL, 751, L25

Venemans B. P., et al., 2013, ApJ, 779, 24

Venemans B. P., et al., 2015, MNRAS, 453, 2259

Venemans B. P., Walter F., Zschaechner L., Decarli R., De Rosa G., Findlay J. R., McMahon R. G., Sutherland W. J., 2016, ApJ, 816, 37

Venemans B. P., et al., 2017a, ApJ, 845, 154

Venemans B. P., et al., 2017b, ApJL, 851, L8

Venemans B. P., et al., 2018, ApJ, 866, 159

Venemans B. P., et al., 2020, ApJ, 904, 130

Venturi G., et al., 2021, A&A, 648, A17

Vestergaard M., 2002, ApJ, 571, 733

Vestergaard M., Osmer P. S., 2009, ApJ, 699, 800

Vignali C., et al., 2018, MNRAS, 477, 780

Virtanen P., et al., 2020, Nature Methods, 17, 261

Vito F., Gilli R., Vignali C., Comastri A., Brusa M., Cappelluti N., Iwasawa K., 2014, MNRAS, 445, 3557

Vito F., et al., 2018, MNRAS, 473, 2378

Vito F., et al., 2019a, A&A, 628, L6

Vito F., et al., 2019b, A&A, 630, A118

Vito F., et al., 2021, A&A, 649, A133

Vito F., Di Mascia F., Gallerani S., Zana T., Ferrara A., Carniani S., Gilli R., 2022, MNRAS, 514, 1672

Vogelsberger M., et al., 2014, MNRAS, 444, 1518

Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nature Reviews Physics, 2, 42

Volonteri M., 2010, A&A Rev., 18, 279

Volonteri M., Haardt F., Madau P., 2003, ApJ, 582, 559

Wada K., Schartmann M., Meijerink R., 2016, ApJL, 828, L19

Walter F., et al., 2003, Nature, 424, 406

Walter F., Riechers D., Cox P., Neri R., Carilli C., Bertoldi F., Weiss A., Maiolino R., 2009, Nature, 457, 699

Walter F., et al., 2022, ApJ, 927, 21

Wang R., et al., 2010, ApJ, 714, 699

Wang R., et al., 2013, ApJ, 773, 44

Wang R., et al., 2016, ApJ, 830, 53

Wang F., et al., 2017, ApJ, 839, 27

Wang F., et al., 2018a, preprint, p. arXiv:1810.11926 (arXiv:1810.11926)

Wang F., et al., 2018b, ApJL, 869, L9

Wang F., et al., 2019, ApJ, 884, 30

Wang F., et al., 2021a, ApJL, 907, L1

Wang F., et al., 2021b, ApJ, 908, 53

Weingartner J. C., Draine B. T., 2001, ApJ, 548, 296

Weingartner J. C., Murray N., 2002, ApJ, 580, 88

Weisskopf M. C., Wu K., Trimble V., O'Dell S. L., Elsner R. F., Zavlin V. E., Kouveliotou C., 2007, ApJ, 657, 1026

Werner M. W., et al., 2004, ApJS, 154, 1

Whitney B. A., 2011, Bulletin of the Astronomical Society of India, 39, 101

Wiedner M. C., et al., 2020, arXiv e-prints, p. arXiv:2012.02731

Wiersma R. P. C., Schaye J., Smith B. D., 2009, MNRAS, 393, 99

Willott C. J., 2005, ApJL, 627, L101

Willott C. J., McLure R. J., Jarvis M. J., 2003, ApJL, 587, L15

Willott C. J., et al., 2007, AJ, 134, 2435

Willott C. J., et al., 2010, AJ, 139, 906

Willott C. J., Bergeron J., Omont A., 2017, ApJ, 850, 108

Wise J. H., Regan J. A., O'Shea B. W., Norman M. L., Downes T. P., Xu H., 2019, Nature, 566, 85

Wiseman P., Schady P., Bolmer J., Krühler T., Yates R. M., Greiner J., Fynbo J. P. U., 2017, A&A, 599, A24

Witt A. N., 1977, ApJS, 35, 1

Witt A. N., Gordon K. D., 1996, ApJ, 463, 681

Woods T. E., et al., 2019, Publ. Astr. Soc. Australia, 36, e027

Wu X.-B., et al., 2015, Nature, 518, 512

Wyithe J. S. B., Bolton J. S., 2011, MNRAS, 412, 1926

Xu J., Sun M., Xue Y., 2020, ApJ, 894, 21

Xue Y. Q., et al., 2011, ApJS, 195, 10

Yajima H., Khochfar S., 2016, MNRAS, 457, 2423

Yan H., Lazarian A., Draine B. T., 2004, ApJ, 616, 895

Yang G., et al., 2020a, MNRAS, 491, 740

Yang J., et al., 2020b, ApJL, 897, L14

Yang J., et al., 2021a, ApJ, 923, 262

Yang J., et al., 2021b, ApJ, 923, 262

Yaqoob T., 1997, ApJ, 479, 184

Younger J. D., Hayward C. C., Narayanan D., Cox T. J., Hernquist L., Jonsson P., 2009, MNRAS, 396, L66

Yuan F., Gan Z., Narayan R., Sadowski A., Bu D., Bai X.-N., 2015, ApJ, 804, 101

Yue M., Fan X., Yang J., Wang F., 2021, ApJL, 921, L27

Zafar T., Watson D. J., Malesani D., Vreeswijk P. M., Fynbo J. P. U., Hjorth J., Levan A. J., Michałowski M. J., 2010, A&A, 515, A94

Zafar T., Watson D., Fynbo J. P. U., Malesani D., Jakobsson P., de Ugarte Postigo A., 2011, A&A, 532, A143

Zafar T., et al., 2018a, MNRAS, 479, 1542

Zafar T., et al., 2018b, MNRAS, 480, 108

Zana T., Gallerani S., Carniani S., Vito F., Ferrara A., Lupi A., Di Mascia F., Barai P., 2022, MNRAS, 513, 2118

Zanella A., et al., 2018, MNRAS, 481, 1976

Zhukovska S., Gail H. P., Trieloff M., 2008, A&A, 479, 453

Zubovas K., Bourne M. A., Nayakshin S., 2016, MNRAS, 457, 496

de Bernardis P., et al., 2000, Nature, 404, 955

van der Vlugt D., Costa T., 2019, MNRAS, 490, 4918