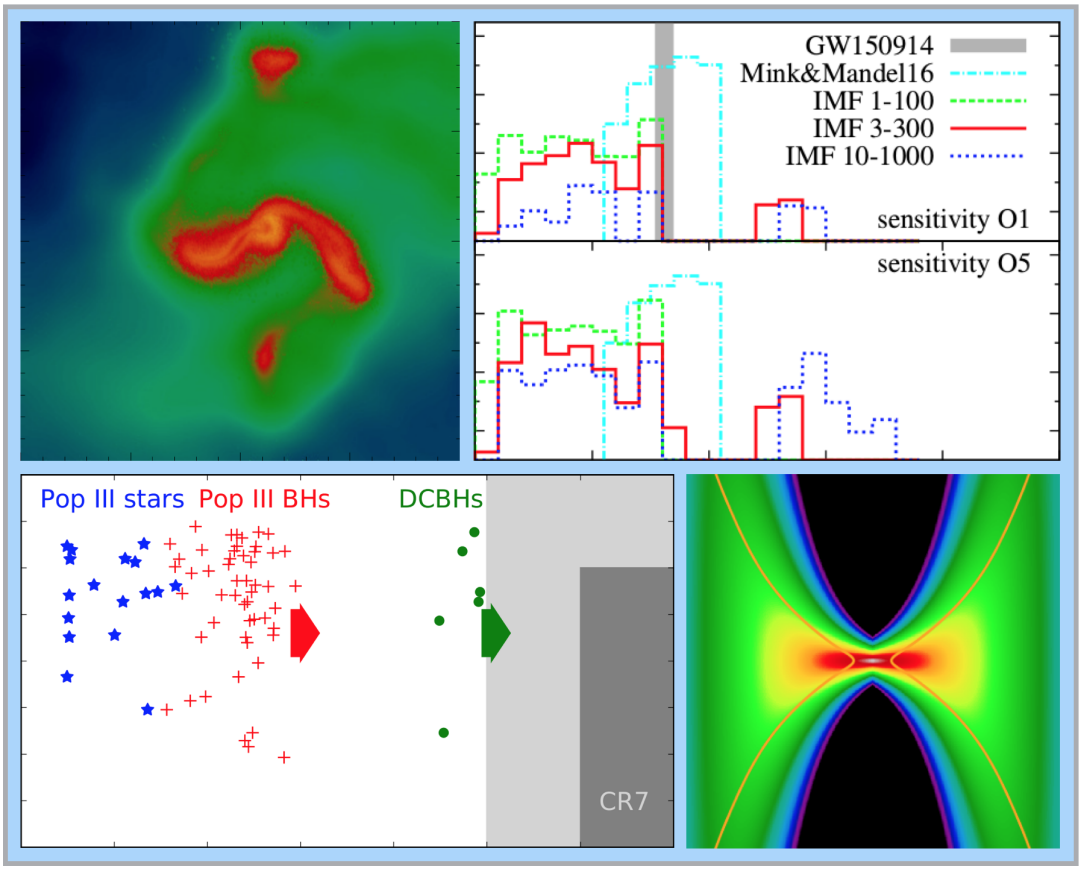
Formation and Growth of the First Supermassive Black Holes



Ph.D. Thesis

by Tilman Hartwig

UNIVERSITÉ PIERRE ET MARIE CURIE

École Doctorale d'Astronomie et Astrophysique d'Île-de-France

PH.D. THESIS

to obtain the title of Doctor in Astrophysics
of the University Pierre et Marie Curie

Presented by

TILMAN HARTWIG

Formation and Growth of the First Supermassive Black Holes

Thesis Advisor: DR. MARTA VOLONTERI

prepared at the Institut d'Astrophysique de Paris, CNRS (UMR 7095),
Université Pierre et Marie Curie (Paris VI)
with financial support from the
European Research Council grant 'BLACK'

Presented and publicly defended on the 22nd of September, 2017 to a jury consisting of

Reviewers:	MICHELA MAPELLI	-	INAF, Padova, Italy
	JARRETT JOHNSON	-	LANL, Los Alamos, USA
Advisor:	MARTA VOLONTERI	-	IAP, Paris, France
President:	FRÉDÉRIC DAIGNE	-	IAP, Paris, France
Examiner:	RALF KLESSEN	-	ZAH, Heidelberg, Germany

Formation and Growth of the First Supermassive Black Holes

Supermassive black holes reside in the centres of most massive galaxies and we observe correlations between their mass and properties of the host galaxies, such as bulge mass or stellar velocity dispersion. Besides this correlation between a galaxy and its central black hole (BH), we see BHs more massive than one billion solar masses already a few hundred million years after the Big Bang. These supermassive BHs at high redshift are just the tip of the iceberg of the entire BH population, but they challenge our understanding of the formation and growth of the first BHs.

With this thesis, I contribute to a better understanding of the formation and growth of the first supermassive BHs with theoretical models and comparisons to recent observations. I improve radiative transfer in 3D hydrodynamical simulations to study the formation of massive BH seeds and analyse the feedback of accreting BHs in their typical host galaxies with a novel 2D analytical approach. Moreover, I model observations of the Lyman-α emitter CR7 and of the first detection of gravitational waves with a semi-analytical model of early star formation.

Our improved method to calculate H_2 self-shielding in simulations of supermassive star and BH seed formation ('direct collapse') yields probabilities to form massive seed BHs that are more than one order of magnitude higher than previously expected. We find that CR7 might be the first candidate to host such a direct collapse BH and we disprove the initially claimed existence of a massive metal-free stellar population in CR7. To better constrain a different seeding scenario for the first BHs, we calculate the merger rate density of binary BHs from the first stars and their detection rates with aLIGO. Our model demonstrates that upcoming detections of gravitational waves in the next decades will allow to put tighter constraints on the properties of the first stars and therefore on formation scenarios of the first BHs. To follow the subsequent accretion of the seed BHs, we develop a 2D analytical model of active galactic nuclei-driven outflows to demonstrate that a more realistic disc profile reduces the amount of gas that is ejected out of the halo, compared to existing 1D models. The outflow prevents gas accretion on to the central BH for only about $\sim 1\,\mathrm{Myr}$, which permits almost continuous gas inflow in the disc plane. Our model yields a correlation of $M_{\mathrm{BH}} \propto \sigma^{4.8}$ between the BH mass and the host's velocity dispersion, close to observations. In the last chapter, we conclude by summarising our results and presenting future research perspectives.

Formation et Croissance des Premiers Trous Noirs Supermassifs

Les trous noirs supermassifs résident dans les centres de la plupart des galaxies massives et on observe des corrélations entre leurs masses et les propriétés de leurs galaxies hôtes, comme la masse du bulbe ou la dispersion des vitesses stellaires. De plus, on observe des trous noirs de plus d'un milliard de masses solaires quelques centaines de millions d'années seulement après le Big Bang. Ces trous noirs supermassifs présents dans l'univers jeune ne sont que le sommet de l'iceberg de l'ensemble de la population de trous noirs, mais ils mettent en question notre compréhension de la formation et de la croissance des premiers trous noirs.

Avec cette thèse, je contribue à une meilleure compréhension de la formation et la croissance des premiers trous noirs supermassifs grâce à des modèles théoriques et des comparaisons avec des observations récentes. J'améliore le transfert radiatif dans les simulations hydrodynamiques pour étudier la formation des graines massives de trous noirs et j'analyse la rétroaction des trous noirs accretant dans leurs galaxies hôtes avec une nouvelle approche analytique en 2D. De plus, je modélise les observations de l'émetteur de Lyman-α CR7 et de la première détection d'ondes gravitationnelles avec un modèle semi-analytique de formation des premières étoiles.

Notre nouvelle méthode améliorant le calcul de la densité de colonne de H_2 dans les simulations de la formation d'étoiles supermassives et de graines de trous noirs («effondrement direct») donne des probabilités pour former des graines massives de trous noirs qui sont plus d'un ordre de grandeur plus élevées que prédit précédemment. Nous trouvons que CR7 pourrait être le premier candidat à héberger un tel trou noir formé par effondrement direct et nous démentons l'existence initialement revendiquée d'une population stellaire massive primordial dans CR7. Nous calculons la densité des taux de fusion des trous noirs binaires des premières étoiles et leurs taux de détection avec aLIGO. Notre modèle démontre que les détections des ondes gravitationnelles à venir au cours des prochaines décennies permettront d'imposer des contraintes plus strictes sur les propriétés des premières étoiles et donc sur les scénarios de formation des premiers trous noirs. Nous développons un modèle analytique en 2D de la rétroaction des noyaux actifs de galaxie pour démontrer qu'un profil de disque plus réaliste réduit la quantité de gaz qui est éjectée du halo par rapport aux modèles 1D existants. La rétroaction empêche l'accretion de gaz sur le trou noir central pendant seulement ~ 1 million d'année environ, ce qui permet une accretion de gaz presque continue dans le plan du disque. Notre modèle donne une corrélation de $M_{\rm BH} \propto \sigma^{4.8}$ entre la masse du trou noir et la dispersion des vitesses stellaires de la galaxie hôte, en concordance avec les observations. Dans le dernier chapitre, nous concluons en résumant nos résultats et en présentant les perspectives de recherche futures.

Contents

1 Introduction

1.1 Motivation

Black holes are one of the most fascinating objects in the Universe. On the one hand, they are extremely simple objects and astrophysical black holes are entirely described by their mass and spin. On the other hand, they cause a variety of different effects on all physical scales and play a crucial role in the formation and evolution of galaxies.

A black hole (BH) is a singularity in space time, created by a mass M that is entirely contained within its Schwarzschild radius

$$R_{\rm SS} = \frac{2GM}{c^2} \approx 3\,{\rm km}\left(\frac{M}{{\rm M}_\odot}\right), \tag{1.1}$$

where G is the gravitational constant and c the speed of light. The curvature of the metric at the Schwarzschild radius is so strong that all trajectories, even those of photons, will inevitably end in the BH. Phrased differently, not even light can escape a BH.

In this thesis we will focus on supermassive BHs in the centres of galaxies, their formation, growth, and feedback. We are especially interested in these processes in the early Universe for two main reasons. First, we observe correlations between the BH mass and different properties of its host galaxy in the local Universe, which points to a possible co-evolution. To understand the population of BHs in their host galaxies and how they evolve together, we want to understand the growth of BHs in their host galaxy from the beginning. Second, already a few hundred million years after the Big Bang, we observe supermassive black holes (SMBHs) of up to $10^{10}\,{\rm M}_\odot$ and do not understand how they gained that much mass within this cosmologically short amount of time. This is because the accretion onto a BH in spherical symmetry is limited via the radiation pressure generated by the accretion disc around a BH to the Eddington rate

$$\dot{M}_{\rm Edd} = \frac{4\pi G M_{\rm BH} m_p}{\epsilon c \sigma_T}, \tag{1.2}$$

where m_p is the proton mass, σ_T the Thompson cross section, and $\epsilon \approx 0.1$ the radiative efficiency of the accretion disc. This yields an e-folding time of $\sim 45\,{\rm Myr}$ for the mass growth rate of BHs. In Figure 1.1 we compare different mass growth rates for a sample of high redshift BHs. To explain the most massive SMBHs at $z > 6$ we need models that produce BH seeds with sufficiently high masses or enable mass accretion close to or even above the Eddington limit.

In this thesis we will discuss the formation and growth of BHs over many orders of magnitude in mass. We will refer to BHs with masses of up to $\sim 100\,{\rm M}_\odot$ as stellar mass BHs and to BHs with masses above $\sim 10^6\,{\rm M}_\odot$ as supermassive black holes (SMBHs). One should also keep in mind the different length scales involved in this problem: the Schwarzschild radius of a stellar mass BH is of the order $10^{-11}\,{\rm pc}$ and even typical SMBHs have sizes below μpc. On the other hand, massive galaxies extend to several hundred kpc and galactic outflows are observed out to Mpc scales. If a galaxy had the size of the earth, the BH would have the size of a black dot on the back of a ladybug.

The thesis is structured as follows: in this first chapter, we introduce the cosmological framework and

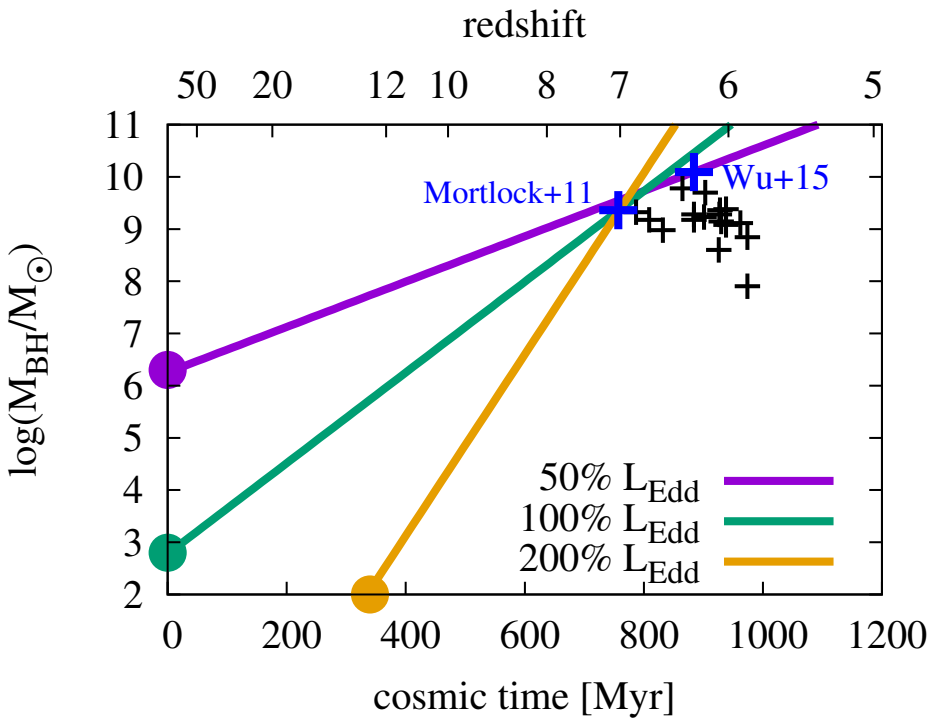


Figure 1.1: Subset of quasars observed at high redshift and their required mass growth rates. The three lines represent accretion at a constant fraction of the Eddington limit and the dots correspond to the required seed masses. It is necessary that high redshift SMBHs start with a high seed mass, accrete constantly above the Eddington limit, or both. The two record holders of this sample in mass (Wu et al. 2015) and redshift (Mortlock et al. 2011) are highlighted in blue.

present the main theories and observational constraints of our current understanding of the formation of the first stars and BHs. In the following chapters we present the results that we obtained during the thesis in different fields. These chapters are mainly based on published papers and we put them into the context of the formation and growth of the first BHs and have updated the references and parts of the discussion with recent results. In the last chapter, we summarise the main findings, conclude and present further perspectives and future research plans.

1.2 Cosmological structure formation

In the early Universe, matter was almost homogeneously distributed. From 2009 to 2013, the Planck satellite analysed temperature fluctuations of the cosmic microwave background (CMB) as a footprint of the tiny but crucial inhomogeneities of the cosmic density field (Figure 1.2). According to the standard model of cosmology, structure has formed hierarchically from those small density fluctuations. The subsequent merging of successively larger structures built up the Universe as we know it today. In order to understand the theoretical framework, the thermal evolution, and non-linear structure formation, we introduce the basic principles in cosmology (mainly based on Bartelmann 2007, 2009; Clark & Glover 2013; Glover 2013a; Hartwig 2014).

1.2.1 Cosmological framework

The standard model of cosmology is based on two assumptions:

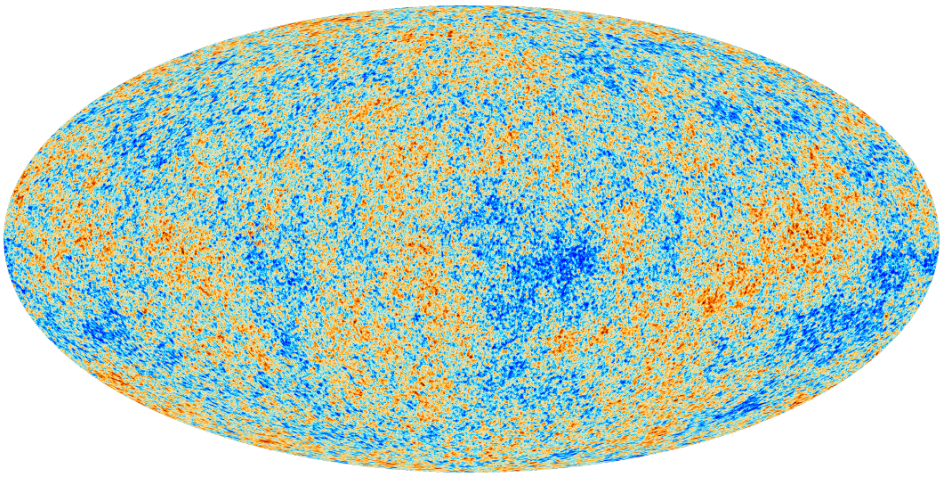


Figure 1.2: Map of temperature fluctuations of the CMB spectrum, showing tiny inhomogeneities about 378,000 years after the Big Bang. The fluctuations are of the order of $\Delta T/T \simeq 10^{-5}$ and correspond to density inhomogeneities (ESA and the Planck Collaboration 2013).

1. On sufficiently large scales, the Universe is isotropic.

2. The position from where we observe the Universe is by no means preferred to any other position (Copernican principle).

These two fundamental assumptions are commonly summarised by claiming that the Universe is "homogeneous and isotropic" (cosmological principle). The only relevant force on cosmological scales is gravity. Due to the cosmological principle, spacetime can be described by a Robertson-Walker metric of the form

$$g = -c^2 \mathrm{d}t^2 + a(t)^2 \left[\mathrm{d}w^2 + f_k(w)^2 \left(\mathrm{d}\theta^2 \sin^2\theta \mathrm{d}\phi^2\right)\right], \tag{1.3}$$

where $a(t)$ is the scale factor and $f_k(w)$ is a function which depends on the radial coordinate w and on the curvature of space-time. The scale factor describes its relative size with $a_{\mathrm{today}} = 1$. The whole dynamics of the Universe is reduced to the dynamics of the scale factor $a(t)$, which can be described by the Friedmann equations

$$\frac{\dot{a}(t)^2}{a(t)^2} = \frac{8\pi G}{3}\rho(t) - \frac{kc^2}{a(t)^2} + \frac{\Lambda}{3}, \tag{1.4}$$

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left(\rho(t) + \frac{3p(t)}{c^2}\right) + \frac{\Lambda}{3}, \tag{1.5}$$

where $\rho(t)$ is the matter density, k parametrises the curvature of space-time, Λ is the cosmological constant, and $p(t)$ the pressure. The Friedmann equations can be combined to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(a(t)^3\rho(t)c^2\right) + p(t)\frac{\mathrm{d}}{\mathrm{d}t}\left(a(t)^3\right) = 0, \tag{1.6}$$

which can be interpreted by means of energy conservation. For non-relativistic matter the adiabatic equation yields the time-dependent matter density $\rho_m(t) = \rho_{m,0}a^{-3}$, whereas for relativistic matter it yields $\rho_r(t) = \rho_{r,0}a^{-4}$ with the current matter density $\rho_{m,0}$ and current energy density of radiation $\rho_{r,0}$. Since the individual energy contributions scale with different powers of a, we can distinguish different epochs in cosmic history and derive their expansion behaviour. The early Universe ($t \lesssim 500\mathrm{yr}$) is

dominated by the radiation energy density and the time evolution of the scale factor in this epoch is given by $a \propto t^{1/2}$, whereas the epoch thereafter is dominated by matter and the scale factor evolves as $a \propto t^{2/3}$ (Einstein-de Sitter limit). The expansion entered the current, dark energy-dominated epoch with $a(t) \propto \exp(t)$, 10^{10} yr after the Big Bang.

Due to the expansion of the Universe, the spectrum of distant galaxies is redshifted with respect to their rest-frame. This redshift z is related to the scale factor of the Universe at emission of the photons by

$$1 + z = \frac{1}{a} \tag{1.7}$$

and represents a commonly used measurement for the age of the Universe and the corresponding distances.

If not stated otherwise, we use the following cosmological parameters in a flat cold dark matter Universe, based on Planck Collaboration et al. (2016):

- Hubble constant: $H_0 = h100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with $h = 0.678$
- Matter density parameter (including baryons): $\Omega_m = 0.309$
- Baryonic matter density parameter: $\Omega_b = 0.048$
- Dark energy density parameter: $\Omega_\Lambda = 0.691$
- Amplitude of the linear power spectrum on the scale of $8h^{-1}\,\mathrm{Mpc}$: $\sigma_8 = 0.823$
- Optical depth to Thomson scattering: $\tau_e = 0.066$
- Spectral power-law index: $n_s = 0.961$
- Helium fraction: $Y_\mathrm{He} = 0.248$

We denote comoving units with a lower-case 'c' in front of the length unit, e.g. "cMpc".

1.2.2 Thermal and chemical evolution

The temperature is a fundamental parameter for the description of the thermal and chemical evolution of the Universe. According to the Stefan-Boltzmann law, the energy density of radiation scales as $\rho_r \propto T^4$, and we have seen above that it also relates to the scale factor by $\rho_r \propto a^{-4}$. Combining these two relations and using Equation 1.7 we derive

$$T_r \propto 1 + z, \tag{1.8}$$

which states that the Universe was hotter in the past. The Universe has to evolve adiabatically on large scales, because any heat flow would violate the cosmological principle.

The Universe starts in a very hot, dense state and cools by expanding. Hence, different particle species freeze out at different times, because the temperature drops and they are no longer in thermal equilibrium. Once the temperature drops below $Tk_B \simeq 800\,\mathrm{keV}$ the equilibrium between protons and neutrons can no longer be maintained. At the moment of "freeze-out", the neutron-to-proton number density is given by

$$\frac{n_n}{n_p} = e^{-\Delta mc^2/k_B T} \simeq \frac{1}{6}, \tag{1.9}$$

where $\Delta mc^2 \simeq 1.4\text{MeV}$ is the mass difference between neutrons and protons. In the first three minutes before fusion can begin producing long-lived heavier elements, neutrons have time to decay and the final neutron-to-proton ratio is

$$\frac{n_n}{n_p} \simeq \frac{1}{7}. \tag{1.10}$$

At this point, nucleosynthesis fuses the existing particles to isotopes and heavier elements, which can be seen in Figure 1.3. Hence, the primordial gas consists of roughly 76% ^{1}H, 24% ^{2}He atoms and traces of

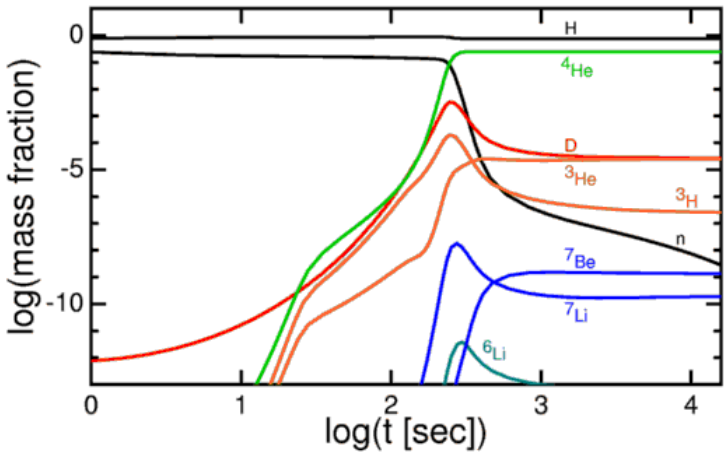


Figure 1.3: Fractional abundance of different primordial species as a function of time for the standard cosmological model. Adapted from Wright (2012).

other isotopes and elements. The gas is still ionised until the reaction

$$\mathrm{H}^+ + \mathrm{e}^- \leftrightarrow \mathrm{H} + \gamma \tag{1.11}$$

freezes out and the gas recombines. This recombination process can be described by Saha's equation and is delayed due to the very large photon-to-baryon ratio. Recombination occurs at $z = 1100 \pm 80$, which corresponds to 378,000 years after the Big Bang. Beyond this point, the gas is mainly neutral, until the first stars and galaxies reionise it again.

1.2.3 Jeans analysis

Obviously (and luckily), the Universe is not completely homogeneous. Even on astronomical scales we observe inhomogeneous, filamentary structures as shown in Figure 1.4. In order to understand structure formation, we first have to review the governing physics. The main set of equations are the equation of state and the equations of hydrodynamics, which are discussed in more detail Section 2.2.1.

In a first, linear approximation we derive the evolution of small perturbations in density

$$\rho(\vec{x}, t) = \rho_0(t) + \Delta\rho(\vec{x}, t) \tag{1.12}$$

and expansion velocity

$$\vec{v}(\vec{x}, t) = \vec{v_0}(t) + \Delta\vec{v}(\vec{x}, t). \tag{1.13}$$

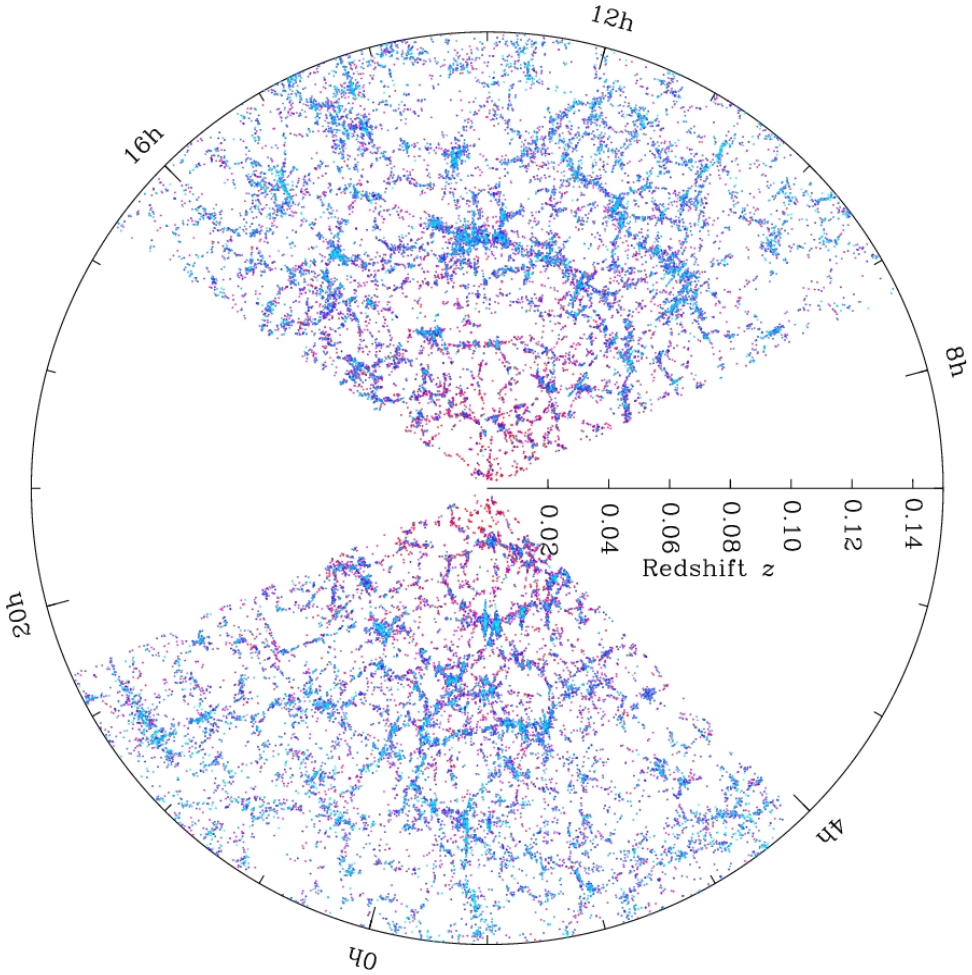


Figure 1.4: Large scale structure in the northern equatorial slice of the SDSS main galaxy redshift sample. The slice is 2.5° thick and galaxies are colour-coded by luminosity. Adapted from SDSS (2008).

By defining the density contrast

$$\delta = \frac{\Delta\rho}{\rho_0}, \tag{1.14}$$

using the isothermal equation of state $p = \rho_0 c_s^2$ with the sound speed c_s and neglecting all quadratic terms in the perturbations, we derive the second-order differential equation

$$\ddot{\delta} + 2H\dot{\delta} = \left(4\pi G\rho_0 + \frac{c_s^2\nabla^2\delta}{a^2}\right), \tag{1.15}$$

where $H = \dot{a}/a$ is the Hubble function. On large scales in an Einstein-de Sitter Universe, $\delta(z) = \delta_0(1+z)^{-1}$ solves this equation, which indicates the growth of density perturbations with time. Decomposing Equation 1.15 into a set of plane waves

$$\delta(\vec{x},t) = \int \frac{\mathrm{d}\vec{k}}{(2\pi)^3}\hat{\delta}(\vec{k},t)e^{-i\vec{k}\vec{x}}, \tag{1.16}$$

yields

$$\ddot{\hat{\delta}} + 2H\dot{\hat{\delta}} = \left(4\pi G\rho_0 - \frac{c_s^2 k^2 \hat{\delta}}{a^2}\right). \tag{1.17}$$

Identifying the right hand side of this equation as the source term of a damped harmonic oscillator equation, the criterion for growing density perturbations is

$$k < k_J = \frac{2\sqrt{\pi G\rho_0}}{c_s}. \tag{1.18}$$

Alternatively, we can define the Jeans length

$$\lambda_J = \frac{2\pi}{k_J} = c_s\sqrt{\frac{\pi}{G\rho_0}}. \tag{1.19}$$

Perturbations smaller than the Jeans length oscillate, whereas perturbations larger than the Jeans length collapse. In a similar fashion we define the Jeans mass

$$M_J = \frac{4\pi}{3}\rho_0\left(\frac{\lambda_J}{2}\right)^3 = \frac{\pi^{5/2}}{6G^{3/2}}c_s^3\rho_0^{-1/2}, \tag{1.20}$$

which describes the minimum mass that a density perturbation must have in order to collapse under its own gravity. At high redshift $z \gg 100$ the Jeans mass is given by (Barkana & Loeb 2001)

$$M_J = 1.35 \times 10^5 \left(\frac{\Omega_m h^2}{0.15}\right)^{-1/2} M_\odot. \tag{1.21}$$

In the low redshift limit, where the coupling between radiation and matter is weak and the gas temperature evolves adiabatically, the Jeans mass is given instead by (Glover 2013a)

$$M_J = 5.18 \times 10^3 \left(\frac{\Omega_m h^2}{0.15}\right)^{-1/2} \left(\frac{\Omega_b h^2}{0.026}\right)^{-3/5} \left(\frac{1+z}{10}\right)^{3/2} M_\odot. \tag{1.22}$$

1.2.4 Non-linear structure formation

In order to derive Equation 1.15 we had to assume $|\delta| \ll 1$. Since δ can become arbitrarily large, we have to find a different description for structure formation in the non-linear regime. One common approach is the spherical collapse model to investigate under which conditions an overdense cloud decouples from the cosmic expansion and collapses to a gravitationally bound object. At the time of decoupling, the cloud has an overdensity of ~ 5.55 with respect to the cosmological background density at the same time. At the point, when we expect the halo to be collapsed in our spherical collapse model, its linear density contrast is

$$\delta_c \simeq 1.69. \tag{1.23}$$

This means that a halo can be considered to be collapsed when its density contrast expected from linear theory reaches a value of δ_c. Assuming the cloud to be in virial equilibrium, the resulting density contrast

with respect to the background density is

$$\Delta = 18\pi^2 \simeq 178. \tag{1.24}$$

The spherical collapse model answers the question whether an overdensity collapses and what we actually mean by "collapsed".

We further quantify the expected number of halos with a certain mass in a given cosmological volume. Press & Schechter (1974) addressed this question by deriving the halo mass function. Their original derivation was wrong by a factor of one half (they argued that it should be there due to normalisation), which was fixed few years later by Bond et al. (1991). The comoving halo number density is then given by

$$N(M,z)\mathrm{d}M = \sqrt{\frac{2}{\pi}}\frac{\rho_0\delta_c}{\sigma_R}\frac{\mathrm{d}\ln\sigma_R}{\mathrm{d}M}\exp\left(-\frac{\delta_c^2}{2\sigma_R^2}\right)\frac{\mathrm{d}M}{M}, \tag{1.25}$$

where σ_R^2 is the smoothed density variance within a sphere of radius R. On small scales the mass function scales as

$$N(M,z)\mathrm{d}M \propto M^{-2}\mathrm{d}M, \tag{1.26}$$

which implies that there are many more low-mass objects than high-mass objects. These findings illustrate the hierarchical structure formation, because massive objects form over time by successively merging low-mass objects.

1.3 Formation of the first stars

In the last section, we discussed the cosmological framework, which defines the initial conditions for star formation in the early Universe. In this section we want to derive characteristics of the first, so called population III (Pop III) stars as possible progenitors of the first SMBHs and as the building blocks of the first galaxies.

The lack of metals in the interstellar medium (ISM) of the early Universe suggests that the characteristic mass of the first stars might be higher than for present-day stars. While early studies found that the first stars have masses of $\sim 100 - 1000\,\mathrm{M}_\odot$, several authors around 2010 advocated the pathway of disc fragmentation around the first protostar to form small multiples of Pop III stars with masses in the range $\sim 10 - 100\,\mathrm{M}_\odot$. A main outcome of a consistent theory of Pop III star formation is the primordial initial mass function (IMF), which defines the number of expected stars per mass. Especially in a primordial environment without any metal line-driven stellar winds, the mass of a Pop III star is the crucial parameter which defines its luminosity, temperature, spectrum, radius, lifetime and its final fate. An overview of the individual mass ranges, proposed by different authors, can be seen in Figure 1.5. The significant variations and uncertainties in the mass ranges illustrate the lack of understanding of the gravoturbulent fragmentation of primordial gas.

Our current understanding of primordial star formation evolved over the last decades (the following brief historical overview is based on Greif et al. 2013). In the 1960's H_2 was considered to be an important coolant in low-metallicity gas, which provides the necessary release of thermal energy for primordial gas clouds to collapse and create protostars (Saslaw & Zipoy 1967; Peebles & Dicke 1968; Hirasawa 1969; Matsuda et al. 1969; Takeda et al. 1969). Subsequent studies modelled Pop III star formation with

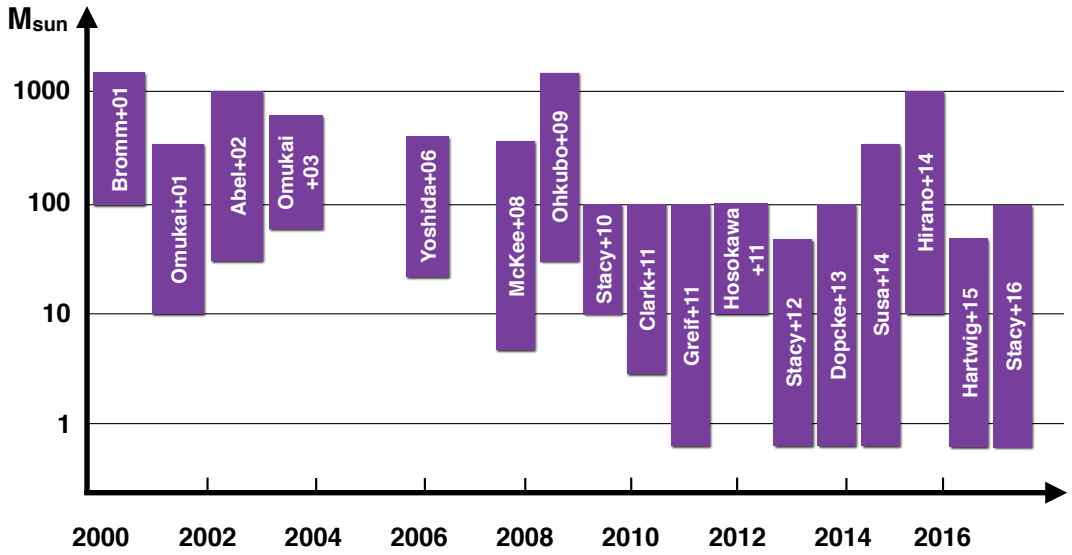


Figure 1.5: Expected mass range of primordial star formation as a function of publication year. The indicated mass ranges are rough estimates and sometimes rely on different assumptions and on various numerical approaches. This plot was inspired by N. Yoshida, complemented with data by Bromm et al. (2001a); Omukai & Palla (2001); Abel et al. (2002); Omukai & Yoshii (2003); Yoshida et al. (2006); McKee & Tan (2008); Ohkubo et al. (2009); Stacy et al. (2010); Clark et al. (2011a); Hosokawa et al. (2011); Stacy et al. (2012); Dopcke et al. (2013); Susa et al. (2014); Hirano et al. (2014); Hartwig et al. (2015b); Stacy et al. (2016). No claim on completeness.

the help of one-zone models, which were able to follow the dynamics with a simplified treatment of radiative cooling (Yoneyama 1972; Hutchins 1976; Silk 1977; Carlberg 1981; Palla et al. 1983; Silk 1983; Kashlinsky & Rees 1983; Carr et al. 1984; Couchman & Rees 1986; Uehara et al. 1996; Tegmark et al. 1997). With increasing computing capacity, three-dimensional simulations of primordial star formation were able to include more physical processes and follow the collapse with a higher spatial resolution (Abel et al. 1998; Bromm et al. 1999; Abel et al. 2000; Bromm et al. 2002; Abel et al. 2002).

Pop III star formation is fundamentally different from present-day star formation: the CMB temperature and mean cosmic density are higher and there are no metals, which provide efficient cooling. According to Bromm (2013), the characteristics of the dark matter minihalos of $\sim 10^5 - 10^6\,\mathrm{M}_\odot$, which hosts the star-forming cloud, is one of the main ingredients of Pop III star formation. The Press-Schechter formalism yields an estimate of the number of dark matter halos as a function of redshift and the Jeans analysis determines whether these halos are gravitationally stable. Even if those halos are gravitationally unstable, the remaining question is whether they can cool in a sufficiently short time (for a detailed discussion about the various cooling mechanisms in primordial gas, see Hartwig et al. 2015b). The virial theorem provides an estimate of the halo mass and temperature:

$$M_{\mathrm{vir}} = 2 \times 10^7 h^{-1} \left(\frac{\mu}{0.6}\right)^{2/3} \Omega_m^{-1/2} \left(\frac{1+z}{20}\right)^{-3/2} M_\odot \tag{1.27}$$

$$T_{\mathrm{vir}} = 3 \times 10^3 \left(\frac{M}{10^6 M_\odot}\right)^{2/3} \left(\frac{1+z}{20}\right) \mathrm{K} \tag{1.28}$$

Since the virial mass is significantly above the estimated Jeans mass, we expect minihalos with masses around M_J and temperatures of roughly $T_{\mathrm{vir}} \simeq 2000\mathrm{K}$ to be the first objects which provide the required conditions for effective H_2 cooling and star formation (Hummel et al. 2012). A further analysis of

the underlying physics yields masses for the minihalo of $10^5 - 10^6 M_\odot$ at virialisation redshifts around $z = 20 - 30$ (Bromm & Larson 2004; Glover 2005; Yoshida et al. 2006; O'Shea & Norman 2007; Greif et al. 2008; McKee & Tan 2008; Yoshida et al. 2008; Glover & Abel 2008; Bromm et al. 2009; Turk et al. 2009; Peters et al. 2010; Turk et al. 2011; Clark et al. 2011a,b; Greif et al. 2012; Stacy et al. 2013; Hirano et al. 2014; Stacy & Bromm 2013; Greif et al. 2013; Bromm 2013). Only a small fraction of the baryons within a virialised object can participate in Pop III star formation, which yields star formation rates of $0.1\% - 1\%$ (Abel et al. 2000; Klessen 2011; Stacy et al. 2013), see also Section 4.4.1.

1.3.1 Primordial chemistry

The second important ingredient in order to understand Pop III star formation is the chemical composition of the primordial gas cloud. Although the number of involved species is much lower than in the present-day Universe, the chemical network is still complex. We only present the most important reactions, whereas Glover & Savin (2009) model a primordial network containing 30 species and 392 reactions. A detailed list of all important chemical reactions and the corresponding rate coefficients can be found in Yoshida et al. (2006); Glover (2015a,b).

The most essential molecule in primordial star formation is molecular hydrogen. Therefore, we focus on its different formation and destruction channels. The chemistry of H_2 in primordial gas has been reviewed in several studies (Abel et al. 1997; Galli & Palla 1998; Stancil et al. 1998; Glover & Abel 2008), while the following overview is mainly based on Clark & Glover (2013) and Glover (2013a). Due to the non-existing dipole moment, H_2 cannot form directly via the radiative association reaction

$$\mathrm{H} + \mathrm{H} \nrightarrow \mathrm{H_2} + \gamma. \tag{1.29}$$

Whereas in the local Universe most H_2 forms via surface reactions on dust grains, this mechanism is not possible in primordial gas. Hence, most H_2 in the early Universe is created via the pathway

$$\mathrm{H} + \mathrm{e} \rightarrow \mathrm{H^-} + \gamma \tag{1.30}$$

$$\mathrm{H^-} + \mathrm{H} \rightarrow \mathrm{H_2} + \mathrm{e^-} \tag{1.31}$$

and about 10% via the pathway

$$\mathrm{H} + \mathrm{H^+} \rightarrow \mathrm{H_2^+} + \gamma \tag{1.32}$$

$$\mathrm{H_2^+} + \mathrm{H} \rightarrow \mathrm{H_2} + \mathrm{H^+}. \tag{1.33}$$

The rate limiting step of these reactions is the rate at which H^- ions form (Turk et al. 2011). However, since the creation and destruction rates of H^- are uncertain by up to an order of magnitude, this also leads to an uncertainty in the creation rate of H_2 . Additionally, H_2 can form via the 3-body processes

$$\mathrm{H} + \mathrm{H} + \mathrm{H} \rightarrow \mathrm{H_2} + \mathrm{H}, \tag{1.34}$$

$$\mathrm{H} + \mathrm{H} + \mathrm{H_2} \rightarrow \mathrm{H_2} + \mathrm{H_2}\,, \tag{1.35}$$

$$\mathrm{H} + \mathrm{H} + \mathrm{He} \rightarrow \mathrm{H_2} + \mathrm{He}. \tag{1.36}$$

Since the possibility of these reactions scale with n^3, formation via the 3-body processes becomes only important at densities above $n \gtrsim 10^8\,\mathrm{cm^{-3}}$ (Omukai et al. 2005; Glover & Abel 2008; Stacy et al. 2013). The onset of 3-body H_2 formation marks an important point in Pop III star formation, because

subsequently most hydrogen is transformed into H_2 and the gas is almost fully molecular at a density of around $n \simeq 10^{12}\,\mathrm{cm}^{-3}$. However, the rate coefficient for this reaction is uncertain by nearly an order of magnitude and the choice of its value influences accretion, long-term stability of the disc, and fragmentation (Turk et al. 2011; Clark et al. 2011b).

In order to destroy H_2, its binding energy of 4.48 eV has to be overcome, which is far above the mean kinetic energy of the CMB photons at the corresponding redshift. Also the collisional dissociation

$$H_2 + H \rightarrow H + H + H, \tag{1.37}$$

$$H_2 + H_2 \rightarrow H + H + H_2 \tag{1.38}$$

or charge transfer with H^+

$$H_2 + H^+ \rightarrow H_2^+ + H \tag{1.39}$$

are ineffective at low temperatures. Hence, primordial gas remains molecular, until a strong ionising radiation field (e.g. by massive stars or black holes) dissociates H_2.

It is mainly dissociated by photons with energies in the range 11.15 to 13.6eV (Lyman-Werner bands), since these photons are not strongly absorbed by neutral hydrogen. Although this process might not be relevant for the initial H_2 abundance in the first minihalos, photodissociation via the so-called Solomon process (Solomon 1965; Stecher & Williams 1967; Abel et al. 1997; Wolcott-Green et al. 2011)

$$H_2 + \gamma \rightarrow H_2^* \rightarrow H + H \tag{1.40}$$

is very important for the accretion and evolution of primordial protostars and may provide a formation channel of massive BH seeds via thermodynamic instability (see Section 1.4.2). However, since Lyman-Werner lines have a spectral width of $\Delta\nu/\nu \simeq 10^{-5}$, only 15% of the absorptions are followed by photodissociation of H_2.

According to Equation 1.10, the mass ratio of helium to hydrogen is about 1/4, which yields a fractional helium abundance of 8.3%. The abundance of molecular hydrogen in the early Universe is between $10^{-6} \lesssim x_{H_2} \lesssim 10^{-3}$ (Galli & Palla 1998; Omukai et al. 2005; Yoshida et al. 2008; Turk et al. 2011; Clark et al. 2011a; Clark & Glover 2013; Bromm 2013). Elemental abundances of the first elements are illustrated in Figure 1.3 as a function of time. Hence, the mean molecular weight of primordial gas is roughly $\mu \simeq 1.23$.

1.3.2 Collapse of a primordial cloud

A characteristic timescale for the collapse of a gas cloud is the free-fall time

$$t_{ff} = \sqrt{\frac{3\pi}{32G\rho}}, \tag{1.41}$$

where ρ is the mean density. A homogeneous, pressureless gas cloud, which is initially at rest collapses within this time. Although these conditions are too idealised, the free-fall time yields a useful characteristic timescale for the collapse of a cloud.

The gas does not simply undergo hierarchical fragmentation, but rather evolves in a highly complex and non-linear fashion. Therefore, we study the different phases and the corresponding physical processes. A state of hydrodynamic and thermodynamic equilibrium is reached at a density of $n \simeq 10^4\,\mathrm{cm}^{-3}$ and

a temperature of $T \simeq 200$K (Glover & Abel 2008; Clark et al. 2011a). Following Bromm (2013), the subsequent collapse can be divided into three phases:

1. Atomic phase: for densities $n \leq 10^8\,\mathrm{cm}^{-3}$ the gas is mostly atomic, except a tiny amount of H_2 .
2. Molecular phase: 3-body H_2 formation sets in and provides an efficient cooling channel until the lines of molecular hydrogen become optically thick at around $n \simeq 10^{14}\,\mathrm{cm}^{-3}$.
3. Approaching protostellar conditions: collision induced emission and H_2 dissociation cooling can provide a release of thermal energy up to densities of $n \simeq 10^{17}\,\mathrm{cm}^{-3}$. Above this value, pressure and temperature decrease steeply and a protostar is formed.

Although the density increases by ten orders of magnitude within the first two phases, the temperature remains almost constant. Only once the H_2 content of the collapsing gas is exhausted, the collapse becomes adiabatic at densities of around $n \simeq 10^{20}\,\mathrm{cm}^{-3}$ (Clark et al. 2011b; Glover & Abel 2008).

Another important ingredient of star formation is turbulence (Klessen 2011). Primordial star formation is mainly dominated by transonic turbulence (Clark et al. 2008; Greif et al. 2013; Stacy et al. 2013; Bromm 2013) and the assembly of the first galaxies marks the onset of supersonic turbulence (Safranek-Shrader et al. 2012a; Bromm 2013).

1.3.3 Fragmentation

Under certain conditions, the cloud does not simply collapse and form one central star, but rather fragment into several clumps. The possibility of fragmentation in primordial clouds is considered by Sabano & Yoshii (1977); Clark et al. (2008); Turk et al. (2009); Stacy et al. (2010); Clark et al. (2011a,b); Greif et al. (2011, 2012); Machida & Doi (2013); Glover (2013a); Bromm (2013); Hartwig et al. (2015b). Generally, one can distinguish between two different types of fragmentation. In the case of the "chemo-thermal" instability a sudden increase in the net cooling rate enables local contraction, while the disc fragmentation relies on instabilities of the accretion disc.

Fragmentation is a chaotic, non-deterministic process and the actual outcome depends sensitively on the initial conditions. Turk et al. (2009) and Greif et al. (2013) simulate different primordial clouds and find them to fragment in one out of five and in two out of nine simulations, respectively. This demonstrates that one can not definitively predict the outcome of fragmentation, but can only access it statistically over a large ensemble of realisations. There are three analytical expressions quantifying the susceptibility of a gas cloud to fragment:

- In order to locally contract instead of globally collapse, a necessary criterion for fragmentation is (Turk et al. 2009)

$$\frac{t_{\mathrm{cool}}}{t_{ff}} < 1, \tag{1.42}$$

  where t_{cool} is the cooling time.

- Toomre (1964) analyses the stability of rotating gas discs and derives the instability criteria

$$Q = \frac{c_s \kappa}{\pi G \Sigma} < 1, \tag{1.43}$$

  where κ is the epicyclic frequency and Σ is the surface density of the disc. For Keplerian discs one can replace κ by the orbital frequency ω. Formally, this criterion is only valid for thin discs and

Goldreich & Lynden-Bell (1965) extended the criterion by requiring $Q < 0.676$ for a finite-thickness isothermal disc to fragment.

- Gammie (2001) investigates the nonlinear outcome of a stability analysis of a Keplerian accretion disc. Based on numerical experiments he derives the instability criterion

$$\frac{3t_{\rm cool}}{\omega} < 1, \tag{1.44}$$

where ω is the orbital frequency. This criterion expresses the possibility that pieces of the disc cool and collapse before they have the opportunity to collide with one another in order to reheat the disc.

In their study on the formation and evolution of primordial protostellar systems, Greif et al. (2012) find the Toomre criterion insufficient for the quantification of gravitational instability and additionally use the Gammie criterion. None of these criteria guarantees fragmentation (Yoshida et al. 2006), but a combination of them yields a reliable quantification of instabilities in the gas cloud (Hartwig et al. 2015b).

Several recent studies report the formation of a rotationally supported disc around Pop III protostars, which might become gravitationally unstable and fragment (Clark et al. 2011b; Hosokawa et al. 2011; Greif et al. 2012, 2013; Glover 2013a; Latif et al. 2013a; Hartwig et al. 2015b; Stacy et al. 2016). This disc extends out to $400 - 1000$AU and has a characteristic temperature of $1500 - 2000$K (Clark et al. 2011b; Hosokawa et al. 2011; Glover 2013a). While the chemo-thermal instability is triggered by the onset of 3-body H_2 formation, disc fragmentation generally occurs at higher densities ($10^{13} - 10^{14}\,{\rm cm}^{-3}$) very close ($r = 1 - 20$AU) to the centre of the cloud (Ripamonti & Abel 2004; Omukai et al. 2005; Clark et al. 2011a,b; Greif et al. 2012; Machida & Doi 2013; Hartwig et al. 2015b). Fragmentation in the central part of the cloud is stopped, when the last cooling mechanism is eliminated by the dissociation of H_2 and the thermal evolution becomes adiabatic.

Magnetic fields might play an important role in the formation of the first stars and they are the focus of several recent studies (Sur et al. 2010; Schober et al. 2012; Machida & Doi 2013; Peters et al. 2014; Latif & Schleicher 2016). Initial seed fields are provided via plasma fluctuations ("Biermann Battery", Biermann 1950). The turbulence generated by the gas infall can amplify magnetic fields very efficiently and on short timescales up to dynamically significant values, leading to a strong, tangled magnetic field (Schober et al. 2012; Latif & Schleicher 2016). The collapsing gas is stabilised by magnetic fields, because angular momentum around the protostar is effectively transferred by magnetic breaking, the gas falls directly on to the protostar without forming a disc, and only a single massive star forms (Machida & Doi 2013; Peters et al. 2014). Magnetic fields significantly affects Population III star formation for

$$B > 10^{-12}\,{\rm G}\left(\frac{n}{1\,{\rm cm}^{-3}}\right)^{-3/2}, \tag{1.45}$$

where n is the number density of the gas, whereas magnetic effects can be ignored and a stellar cluster forms when the natal minihalo field strength is weaker than (Machida & Doi 2013)

$$B < 10^{-13}\,{\rm G}\left(\frac{n}{1\,{\rm cm}^{-3}}\right)^{-3/2}. \tag{1.46}$$

1.3.4 Evolution of protostars

In order to determine the mass accretion rate, feedback and other processes that might influence the protostellar environment, we have to understand the basic principles of primordial protostellar evolution. Except for some quantitative disagreements (see e.g. Turk et al. 2011), there is broad agreement on the qualitative evolution of a Pop III protostar. Although this evolution is a complex interplay of many processes, it can be summarised by a few characteristic steps (Stahler et al. 1986; Omukai & Palla 2001).

1. At densities above $n \gtrsim 10^{19}\,\mathrm{cm}^{-3}$ most of the molecular hydrogen is dissociated, the adiabatic index rises from $\gamma \simeq 1.1$ to $\gamma \simeq 5/3$ and the self-similar solution breaks down (Omukai & Palla 2001; Haardt et al. 2002; Omukai et al. 2005; Yoshida et al. 2006; McKee & Tan 2008; Greif et al. 2012).

2. A quasi-hydrostatic core with a mass of $M_* \simeq 0.01\,\mathrm{M}_\odot$ forms and a shock front develops at the protostellar surface (Omukai & Nishi 1998; Haardt et al. 2002; Greif et al. 2012; Hirano et al. 2014).

3. In the early phase, luminosity and thermal energy are mostly produced by contraction (Omukai & Palla 2001; Krumholz & McKee 2008). Consequently, temperature rises slowly and at a mass of around $M_* \simeq 5\,\mathrm{M}_\odot$ the optically thick outer layers expand to due radiation pressure from the luminosity of the contracting core (Stahler et al. 1986).

4. Hydrogen fusion halts the contraction at higher masses and the star reaches the zero age main sequence (ZAMS) after approximately $10^6\,\mathrm{yr}$ of protostellar evolution (Yoshida et al. 2006; Glover 2013a).

There are some special characteristics for the metal-free stellar evolution, compared to the present-day one. These characteristics of primordial protostellar evolution are summarised in the following list:

- Deuterium burning does not play an important role (Omukai & Palla 2001; Schleicher et al. 2013; Glover 2013a).

- High-mass stars reach the main sequence while they are still accreting (Peters et al. 2010; Klessen 2011).

- Massive star formation cannot be spherically symmetric (Klessen 2011).

- While Pop III protostars are large, fluffy objects, the final stars are smaller than their present day counterparts (Smith et al. 2011; Clark & Glover 2013).

The previously discussed evolution of primordial protostars goes along with an interdependent network of other important processes that are discussed in the following subsections.

Accretion

Accretion is a crucial process in the determination of the final masses of metal-free stars. Due to the chemical composition of the primordial gas and the comparatively high temperature, accretion in the early Universe differs significantly from accretion in present-day star formation. Since there are no dust grains, radiative pressure is less efficient and the accretion rates are consequently higher. A first estimate

for the accretion rate can be derived from the assumption that a Jeans mass is accreted in about one free-fall time:

$$\dot{M} = \frac{M_J}{t_{ff}} \propto c_s^3 \propto T^{3/2} \tag{1.47}$$

The temperature of the star-forming gas in a primordial minihalo is higher than the $\sim 10\,\mathrm{K}$ of present-day star forming clouds (Omukai & Palla 2001; Bergin & Tafalla 2007; Glover 2013a). Consequently, the primordial accretion rates are much higher than local accretion rates, which are of the order of $\dot{M} \simeq 10^{-5}\,\mathrm{M_\odot\,yr^{-1}}$ (Glover & Abel 2008). Typical values for the accretion rate in primordial star formation range from $\dot{M} \simeq 10^{-3}\,\mathrm{M_\odot\,yr^{-1}}$ to $\dot{M} \simeq 10^{-1}\,\mathrm{M_\odot\,yr^{-1}}$ (Ripamonti et al. 2002; Hosokawa et al. 2011; Clark et al. 2011b; Hirano et al. 2014). However, these accretion rates are highly variable in time (Klessen 2011; Smith et al. 2011; Hirano et al. 2014; Stacy et al. 2013; Glover 2013a).

A main question is, when accretion is stopped and by what processes. In order to answer this question, we can distinguish two cases:

- In the case of "smooth accretion", major parts of the infalling envelope are accreted by the central object. Since protostellar radiation cannot halt this inflow, the mass accretion rate decreases rather slowly with increasing stellar mass (Machida & Doi 2013; Hirano et al. 2014).
- In the case of "competitive accretion", the gas is accreted by several protostars. Consequently, each individual protostar has to compete for gas and the reservoir might get exhausted. This scenario (also called "fragmentation induced starvation") is discussed e.g. by Peters et al. (2010).

In reality, there is no clear distinction between these two scenarios and the accretion rate might remain high, although there are several protostars competing for infalling gas. Ultimately, accretion stops due to radiative feedback by the protostars.

Protostellar mass

As already seen in Figure 1.5, the uncertainties in the expected mass range are very large. Since the accretion rate is very high and primordial gas contains almost no metals, Pop III stars can become very massive (Omukai & Palla 2001; Clark et al. 2008). Assuming a constant accretion rate of $\dot{M} \simeq 10^{-3} M_\odot\,\mathrm{yr^{-1}}$ over the typical Kelvin-Helmholtz timescale $t_{KH} \simeq 10^5\mathrm{yr}$ yields a rough upper limit of $100 M_\odot$. Even if we consider continuing accretion during the main sequence lifetime, this increases the upper mass limit only by a factor of a few to $\sim 600 M_\odot$ (Bromm & Loeb 2004). However in reality, protostellar feedback and competitive accretion limit the stellar masses to much smaller values.

Feedback

Generally, radiative feedback from the accreting protostar has several effects on its environment:

- Radiative pressure counteracts the gravitational force and therefore reduces accretion.
- Radiation heats the gas and therefore stabilise it against fragmentation.
- Some photons are energetic enough to photodissociate H_2.

It is still a matter of debate whether these processes actually occur and significantly influence primordial star formation.

Although the number of ionising photons is small in the early protostellar phase, their number steeply rises with mass. Above a protostellar mass of $10-15\,M_\odot$, a star can effectively photodissociate molecular hydrogen (Glover 2000; Glover & Brand 2003; Smith et al. 2011; Klessen 2011; Hosokawa et al. 2011; Hirano et al. 2014). Although accretion feedback heats the gas and even removes the dominant coolant, it is not believed to suppress fragmentation on large scales (Krumholz & McKee 2008; Peters et al. 2010; Clark et al. 2011a,b; Greif et al. 2011; Glover 2013a). However, Smith et al. (2011) and Machida & Doi (2013) find fragmentation to be suppressed in the inner $\sim 20\,\mathrm{AU}$ due to accretion luminosity heating. Even if fragmentation cannot be suppressed by accretion feedback, it is delayed by up to $\sim 1000\,\mathrm{yr}$. The radiation pressure is rather ineffective at early protostellar stages, but it becomes important for protostellar masses above $\sim 50\,M_\odot$ and can even halt further gas accretion (Omukai & Palla 2001; Haardt et al. 2002; Hosokawa et al. 2011, 2012b; Hirano et al. 2014).

Final stages

The final fate of a Pop III star mainly depends on its mass and weakly on its rotation. We summarise these final fates and the corresponding remnants in Table 1.1. Since rotation supports mixing of the

Mass Range $[M_\odot]$	Final Fate	Remnant
$\lesssim 9$	Planetary nebula	White Dwarf
9 – 10	O/Ne/Mg core collapse SN	Neutron Star
10 – 25	Fe core collapse SN	Neutron Star
25 – 40	Weak Fe core collapse SNa	Black hole
40 – 100	Direct collapse (no SN)	Black hole
100 – 140	Pulsational pair-instability SN	Black hole
140 – 260	Pair-instability SN	No remnant
260 – 10^5	Direct collapse (no SN?)	Black hole
$\gtrsim 10^5$	Direct collapse before reaching main sequence	Black hole

Table 1.1: Final stages of non-rotating Pop III stars as a function of their initial mass. Adopted from Karlsson et al. (2013) with data based on Fryer et al. (2001); Heger et al. (2003); Nomoto et al. (2006); Ohkubo et al. (2006)

elements inside the star and therefore leads to a more homogeneous evolution, the minimum mass for a rotating Pop III star to end in a pair-instability supernovae (PISN) is decreased to $\sim 65 M_\odot$ (Chatzopoulos & Wheeler 2012; Yoon et al. 2012; Stacy & Bromm 2013). The same reasoning is valid for the upper mass limit of (pulsational) PISNe: Limongi (2017) shows that rotation and thus chemical mixing yields more massive He cores in metal-poor ($10^{-3} Z_\odot$) stars. The maximum mass for a metal-poor star to explode as a PISNe is therefore $\sim 190\,M_\odot$ (Yoon et al. 2012).

1.3.5 Observational signatures of the first stars

The first stars are too far away in space and time to be directly observable (Zackrisson et al. 2011) and we rely on indirect observations to constrain their properties:

- PISNe are very luminous supernovae of massive, low metallicity stars. If they exist, they can be observed out to high redshifts due to their high explosion energies and allow to probe the Universe prior to reionisation. Their (non-)detection constrains the star formation rate and the IMF of the first stars (Mackey et al. 2003; Scannapieco et al. 2005; Whalen et al. 2013; de Souza et al. 2014; Magg et al. 2016; Hartwig et al. 2017a).

- We can learn about the first stars by observing their direct descendants in our galactic neighbourhood. This approach is called galactic archaeology, the search for stellar fossils in our own Galaxy, which provides precious information about its formation history (Frebel & Norris 2015). In Hartwig et al. (2015a) we show that already the absence of any Pop III remnant sets tight constraints on the lower mass limit of the first stars (see also Ishiyama et al. 2016; Magg et al. 2017; de Bennassuti et al. 2017).

- Second generation stars are those that got enriched by exactly one previous supernova and which consequently carry the chemical fingerprint of this progenitor star. Several groups have already succeeded in determining the most likely masses of such progenitor stars (Aoki et al. 2014; Keller et al. 2014; Placco et al. 2016). In a current study, we predict the most promising host systems of these second generation stars and determine their observational signatures (Hartwig et al. 2017c).

- Gravitational waves (GWs) from the compact remnants of the first stars allow indirect observations of the mass and formation efficiency of Pop III stars (Schneider et al. 2000; Hartwig et al. 2016a; Pacucci et al. 2017a). We will discuss this possibility in more detail in Chapter 3.

- Further constraints on the Pop III IMF come from existing and upcoming observations of gamma ray bursts (Bromm & Loeb 2002; Ma et al. 2016), or the near-infrared background excess (Madau & Silk 2005).

1.4 Formation of the first supermassive black holes

In this section, we introduce existing theories of the formation and growth of the first SMBHs. We first present current observations in the local and distant Universe to motivate our theoretical framework and to highlight open questions. We discuss and compare different formation scenarios of BH seeds and present AGN-driven winds as one potential feedback mechanism that regulates their growth.

1.4.1 Observations

Observations of quasars at high redshifts indicate that SMBHs of several billion solar masses were already assembled in the first billion years after the Big Bang (Fan et al. 2003, 2006a; Willott et al. 2010; Venemans et al. 2013b; De Rosa et al. 2014; Wu et al. 2015). The current record holders are a bright quasar, which hosts a SMBH with a mass of $2 \times 10^9 M_\odot$ at $z = 7.085$ (Mortlock et al. 2011), corresponding to ~ 800 million years after the Big Bang, and a SMBH with $1.2 \times 10^{10} M_\odot$ at $z = 6.30$ (Wu et al. 2015). It is still unclear how these objects were able to acquire so much mass in this short period of time, which in turn raises questions about the formation mechanism and the involved physics.

Besides the observation of these SMBHs at high redshift, there are several observations that indicate a correlation in the local Universe between the mass of the central SMBH and large scale properties of the host galaxy, such as the stellar velocity dispersion σ, luminosity, or the bulge mass (Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Marconi & Hunt 2003; Häring & Rix 2004; Gültekin et al. 2009), see Figure 1.6. These relations imply a possible correlation between the central black hole and its host galaxy, and AGN feedback has been suggested to be responsible for regulating accretion onto the BH and star formation in the galaxy, guiding this co-evolution (see Heckman & Kauffmann 2011, for a review).

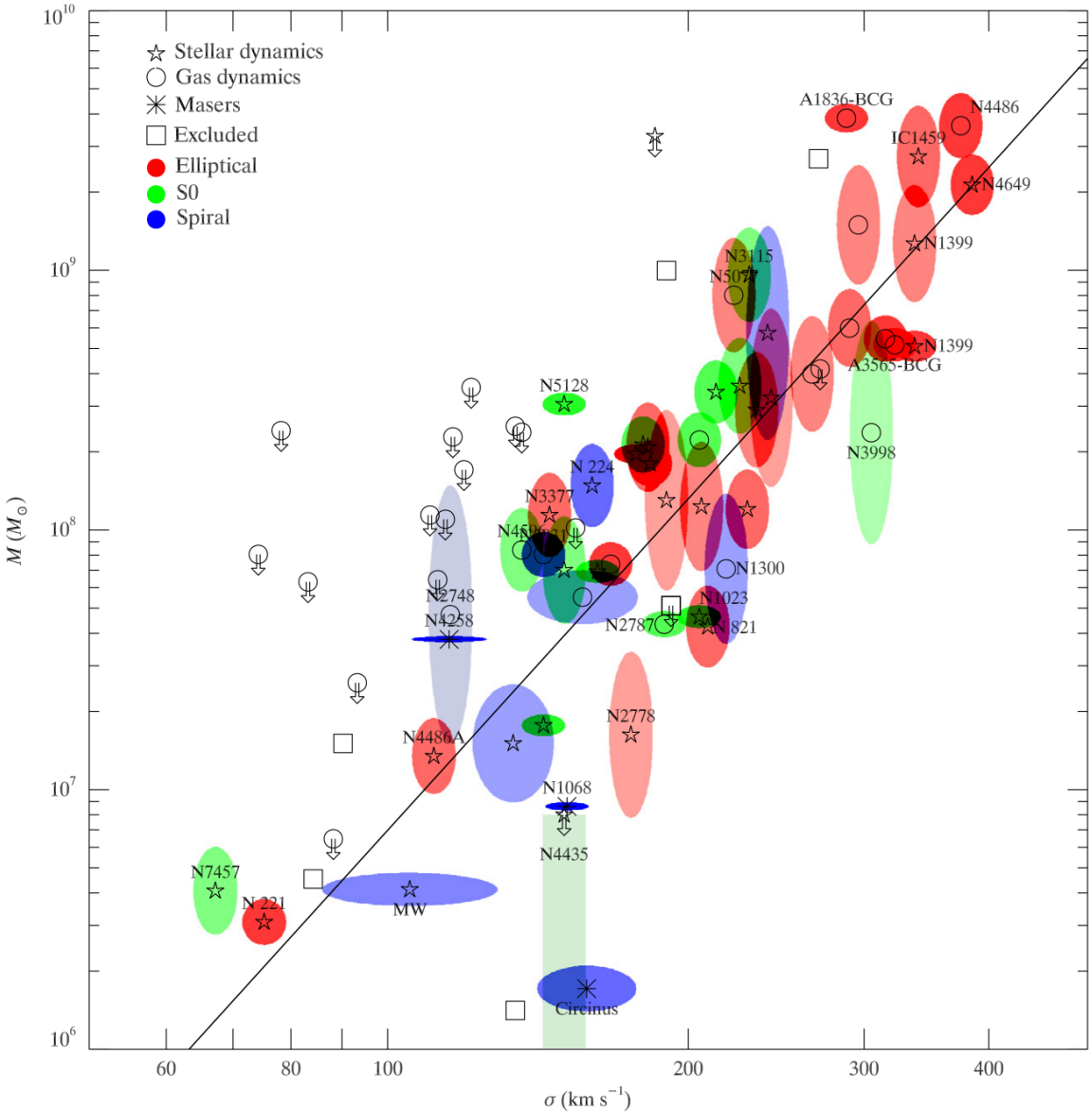


Figure 1.6: M–σ relation based on observations of local galaxies with dynamical measurements. The symbol indicates the method of BH mass measurement: stellar dynamical (pentagrams), gas dynamical (circles), masers (asterisks). The colour of the error ellipse indicates the Hubble type of the host galaxy. Squares are galaxies that are not included in the fit. The line is the best fit relation to the full sample. The correlation indicates a possible co-evolution between a BH and its host galaxy. Adapted from Gültekin et al. (2009).

1.4.2 Seed formation scenarios

We present and briefly review several scenarios to form seed BHs. We focus on the astrophysical standard scenarios to form stellar and more massive BH seeds, but we first present more exotic formation channels. So-called "primordial" BHs have formed during inflation, possibly out of large density inhomogeneities, non-linear metric perturbations, a step in the power spectrum, or cosmic string loops, which result from a scaling solution of strings formed during a phase transition in the very early Universe (Silk & Vilenkin 1984; Bramberger et al. 2015; Khlopov 2010). Overdensities of mass

$$M_{\mathrm{BH}} \approx \frac{c^3 t}{G} \approx 10^{15}\,\mathrm{g}\left(\frac{t}{10^{-23}\,\mathrm{s}}\right), \tag{1.48}$$

could collapse to primordial BHs (Hawking 1971; Carr & Hawking 1974; Carr et al. 2016), where t is the time after the Big Bang. However, the probability to form such a BH is highly improbable in a homogeneously expanding Universe, since it implies metric perturbations of order unity. If they exist, primordial BHs represent a nonrelativistic form of dark matter and their mass density has to be smaller than the dark matter density of the Universe. More stringent constraints come from primordial BHs with masses below 10^{14} g, which should have evaporated until today (Hawking 1975; Khlopov 2010). The (non-)detection of expected evaporated particles and their contribution to cosmic rays put constraints on the formation rate and mass spectrum of potential primordial BHs. Even the absence of observational evidence for primordial BHs is important for cosmology, since it provides unique information about the very early Universe (Khlopov 2010). Upcoming, spaceborne GW detections of BH mergers at high redshift can contribute to further constrain these models.

Pop III remnant BHs

One natural scenario are stellar mass seed black holes with masses up to a few hundred solar masses that are the remnants of Population III (Pop III) stars and then grow by mass accretion or mergers (Madau & Rees 2001; Haiman & Loeb 2001; Volonteri et al. 2003; Yoo & Miralda-Escudé 2004; Haiman 2004; Pelupessy et al. 2007; Tanaka & Haiman 2009; Whalen & Fryer 2012; Madau et al. 2014).

However, already a simple order of magnitude argument illustrates that these seed BHs require extreme conditions to be the progenitors of the most massive BHs at high redshift. Assuming accretion at the Eddington limit, the e-folding time is 45 million years. A seed BH, formed ~ 100 Myr after the Big Bang, accreting at the Eddington limit can therefore grow by a factor of $e^{650/45} \simeq 2 \times 10^6$ until $z = 7.085$. To reach a mass of $2 \times 10^9\,\mathrm{M}_\odot$ within this time, it is therefore necessary to start with a seed mass of about $10^3\,\mathrm{M}_\odot$. This is larger than the mass of a typical Pop III stellar remnant (Clark et al. 2011a,b; Greif et al. 2011; Stacy et al. 2012; Latif et al. 2013a; Hirano et al. 2014; Hartwig et al. 2015b), and also constant accretion at the Eddington limit is very unlikely (Alvarez et al. 2009; Aykutalp et al. 2014; Habouzit et al. 2017). In all seed formation scenarios, BH-BH mergers or a lower radiative efficiency of the accretion disc could help seeds to gain mass more rapidly (Shapiro 2005; Volonteri & Rees 2006). It is still an open question, how these high gas accretion and inflow rates could be sustained during the growth of the SMBH (Johnson & Bromm 2007; Alvarez et al. 2009; Milosavljević et al. 2009b; Johnson et al. 2013b; Jeon et al. 2014) and we further address this question in Chapter 6.

Dense stellar clusters

Very massive stars could form through run-away stellar collisions in dense stellar clusters, which produce remnant BHs of $\gtrsim 1000\,\mathrm{M}_\odot$ (Begelman & Rees 1978; Ebisuzaki et al. 2001; Portegies Zwart et al. 2004; Belczynski et al. 2004b; Omukai et al. 2008; Devecchi & Volonteri 2009; Lupi et al. 2014; Katz et al. 2015; Mapelli 2016; Sakurai et al. 2017).

The typical masses of stellar cluster in atomic cooling halos with $T_{\rm vir} > 10^4\,\mathrm{K}$ are of the order $10^5\,\mathrm{M}_\odot$ and their half mass radii is about 1 pc. A large fraction of these very dense clusters undergoes core collapse before stars are able to complete stellar evolution. Runaway stellar collisions eventually lead to the formation of a very massive star, leaving behind a massive black hole remnant (Devecchi & Volonteri 2009). This scenario favours the metallicity range of about 10^{-5}–$10^{-3}\,Z_\odot$, because above this threshold, stellar winds eject too much gas and prevent the formation of massive stars and below this threshold gas fragmentation is not efficient and we cannot form a dense cluster in the first place. Mapelli (2016) studies the metallicity dependence on the formation of BHs in dense stellar clusters. BHs can form via stellar collisions, even at solar metallicities, but the mass of the final merger product is limited to $\leq 30\,\mathrm{M}_\odot$ due to mass-loss by stellar winds. In the metallicity range $0.1 - 0.01\,Z_\odot$ BHs with masses up to $350\,\mathrm{M}_\odot$ can form (Spera & Mapelli 2017). However, for these BHs with $< 1000\,\mathrm{M}_\odot$, formed above $\sim 0.01\,Z_\odot$, the probability of being ejected out of the cluster is much higher, compared to more massive BHs (Mapelli et al. 2011, 2013).

Alternatively, the central density in the core of a stellar cluster can be high enough for fast mergers of stellar mass BHs. If the binaries are hard enough and gas is funnelled to the centre in a low angular momentum flow to contract the existing stellar cluster, a BH with a final mass of $M_{\rm BH} \gtrsim 10^5\,\mathrm{M}_\odot$ can form, independent of the initial gas metallicity (Davies et al. 2011). The most promising stellar clusters for this scenario have a velocity dispersion in the range $40 - 100\,\mathrm{km\,s^{-1}}$ (Miller & Davies 2012).

Direct Collapse black holes

Another formation scenario is the direct collapse of a protogalactic gas cloud, which yields black hole seed masses of $10^4 - 10^6\mathrm{M}_\odot$ (Rees 1984; Loeb & Rasio 1994; Bromm & Loeb 2003a; Koushiappas et al. 2004; Begelman et al. 2006; Lodato & Natarajan 2006; Volonteri 2010; Shang et al. 2010; Schleicher et al. 2010; Choi et al. 2013; Latif et al. 2013b; Regan et al. 2014a; Latif et al. 2014; Sugimura et al. 2014; Visbal et al. 2014a; Agarwal et al. 2014; Latif et al. 2015; Becerra et al. 2015; Hartwig et al. 2015c; Latif et al. 2016c; Johnson & Dijkstra 2017). To form such a massive seed, a high mass inflow rate of $\gtrsim 0.1\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ is required (Begelman 2010; Hosokawa et al. 2012a, 2013; Schleicher et al. 2013; Ferrara et al. 2014; Latif & Volonteri 2015). Sufficient conditions for such high mass inflow rates are provided in halos with $T_{\rm vir} > 10^4\,\mathrm{K}$ in which gas fragmentation and star formation are suppressed during the collapse (Latif et al. 2013b). To avoid fragmentation, the gas has to be metal-free and a strong radiation background has to photodissociate molecular hydrogen, which otherwise acts as a strong coolant. Strong shocks (Inayoshi & Omukai 2012) or magnetic fields (Sethi et al. 2010; Machida & Doi 2013) can also help to prevent fragmentation in primordial gas. Under these specific conditions, the gas can only cool by atomic hydrogen and collapses monolithically to form a supermassive star (SMS), which later on forms a BH seed (Shapiro 2004; Begelman 2010; Johnson et al. 2012; Hosokawa et al. 2012a, 2013; Inayoshi et al. 2014; Inayoshi & Haiman 2014; Latif et al. 2016c) or a quasi-star, which forms a stellar mass black hole that grows by swallowing its envelope (Begelman et al. 2006, 2008; Volonteri & Begelman 2010; Ball et al. 2011; Schleicher et al. 2013). The characteristic mass of direct collapse BHs is $\sim 10^5\,\mathrm{M}_\odot$ (Latif et al. 2013c; Ferrara et al. 2014), with the exact mass depending on the accretion rate, metallicity, and

LW flux. We will refer to this specific type of direct collapse as 'direct collapse scenario' hereafter.

Based on the strength of the photodissociating radiation, the cloud either monolithically collapses close to isothermality, or is able to efficiently cool and to fragment. The main quantity that discriminates between these two different collapse regimes is the flux in the Lyman–Werner (LW) bands (11.2–13.6 eV). This LW radiation is emitted by the first generations of stars and it is convenient to express the flux in units of $J_{21} = 10^{-21}\,\mathrm{erg\,s^{-1}\,cm^{-2}\,Hz^{-1}\,sr^{-1}}$ (in the following, we use this convention without explicitly writing J_{21}). The so-called critical value J_{crit} sets the threshold above which a halo with $T_{\mathrm{vir}} > 10^4\,\mathrm{K}$ can directly collapse to a SMBH seed. Below this value, the gas is susceptible to fragmentation due to efficient H_2 cooling and the mass infall rates towards the centre are generally lower. However, the values for J_{crit} quoted in the literature span several orders of magnitude from $J_{\mathrm{crit}} = 0.5$ (Agarwal et al. 2016b) to as high as $J_{\mathrm{crit}} \simeq 10^5$ (Omukai 2001; Latif et al. 2015). There are several reasons for this large scatter. First of all, the value of J_{crit} is highly sensitive to the spectral shape of the incident radiation field, with softer radiation fields leading to significant smaller values of J_{crit} (Shang et al. 2010; Sugimura et al. 2014; Agarwal & Khochfar 2015; Latif et al. 2015). Secondly, one-zone calculations (e.g. Omukai 2001) tend to yield lower values of J_{crit} than determinations made using 3D numerical simulations. This is a consequence of the fact that J_{crit} depends to some extent on the details of the dynamical evolution of the gas, which are only approximately captured by one-zone calculations. This dependence on the gas dynamics also leads to J_{crit} varying by a factor of a few from halo to halo (Shang et al. 2010; Latif et al. 2014). Although there seems not to be one universal value of J_{crit} (Agarwal et al. 2016b), it is convenient to use this artificial threshold as a quantification of the direct collapse scenario to test the relevance of different physical processes. Once a process significantly affects the value of J_{crit}, it is very likely that it plays an important role in the formation of SMSs and SMBH seeds.

We note that, besides the presented direct collapse scenario, other gas-dynamical channels to form massive BHs are possible. The high mass infall rate could also be achieved by low angular momentum material, although an additional mechanism of efficient angular momentum transport is still required (Eisenstein & Loeb 1995; Koushiappas et al. 2004). Another possibility are local or global gas instabilities that could transport angular momentum outwards and gas inwards (Shlosman et al. 1989; Begelman et al. 2006). The "bar-within-bars" mechanism is triggered by global tidal torques and amplified by subsequently resulting local hydrodynamical torques. It can redistribute gas on a dynamical time scale and act over a large dynamical range from galactic scales to the centre.

1.4.3 Mass accretion and AGN feedback

To explain the observed SMBHs at $z > 6$ the BH seeds need to grow in mass by five to eight orders of magnitude in a few hundred million years. This poses two problems for the mass growth rate of BHs: first, the gas has to be transported from galactic scales to the central BH and lose its angular momentum to be finally accreted. Second, the mass growth of massive BHs is unavoidably connected to strong radiative feedback from the accretion disc. In Chapter 6, we focus on AGN-driven winds as a likely form of feedback and motivate the main observations and the basic theory in this section.

The accretion discs around SMBHs are nature's most efficient engines to convert gravitational energy of the infalling gas into radiation. Therefore, these extreme sources of energy are expected to suppress further gas accretion and to regulate the galactic gas content and star formation via ionising, thermal, and mechanical feedback. AGN feedback may also self-regulate the growth of the central SMBH and its inclusion in models of galaxy formation improves the match between the simulated and the observed galaxy luminosity function for massive galaxies (Bower et al. 2006; Croton et al. 2006; Somerville et al.

2008; Hirschmann et al. 2014; Croton et al. 2016).

The proposed co-evolution between a galaxy and its central BH is still debated and there are many open questions regarding the nature of AGN feedback (recent reviews on the observational evidence and the theoretical modelling of AGN feedback can be found in Fabian 2012; King & Pounds 2015). Is this relation between the BH mass and galactic properties still valid in galaxies at higher redshift or with lower masses (Greene et al. 2008; Volonteri et al. 2008; Volonteri & Natarajan 2009; Volonteri & Stark 2011; Baldassare et al. 2015; Reines & Volonteri 2015; Habouzit et al. 2017)? How and in which form is the energy from the AGN on μpc scales communicated to the interstellar medium (ISM) on kpc scales? Under which conditions can gas escape the gravitational potential of a galaxy? How could the first BHs grow to masses above $10^9\,\mathrm{M}_\odot$ within the first billion years of the Universe, when strong radiative feedback is expected to suppress further gas inflow on to the BH?

Our current understanding of AGN-driven outflows is based on observations at various wavelengths. The dynamics of the ISM, which is a first indicator of AGN feedback, can be traced by dust, CO, C^+, or CII emission with radio telescopes, such as ALMA. Radio observations reveal large molecular outflows with velocities of $100\,\mathrm{km\,s^{-1}}$ up to a few times $1000\,\mathrm{km\,s^{-1}}$ (Aalto et al. 2012; Cicone et al. 2014). Recently, Maiolino et al. (2017) report star formation inside a galactic outflow, another strong indication for the presence of an outflow of cold gas. In Mrk231, the closest and best studied quasar, we observe velocities of $\sim 1000\,\mathrm{km\,s^{-1}}$ and mass outflow rates of the order $\sim 1000\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ (Feruglio et al. 2010; Rupke & Veilleux 2011; Sturm et al. 2011). These outflows have to be powered by the AGN, because SN-driven outflows can only account for outflow velocities up to $\sim 600\,\mathrm{km\,s^{-1}}$ (Martin 2005; Sharma & Nath 2013). Besides these molecular outflows, the Fe K lines in the X-ray reveal highly ionised outflows with mildly relativistic velocities of $\sim 0.1c$ in several AGNs (Chartas et al. 2002; Pounds et al. 2003; Reeves et al. 2003; Cappi 2006; Gofford et al. 2013). Some ultrafast outflows even have velocities up to $\sim 0.3c$, but the majority has velocities around $\sim 0.1c$ (King & Pounds 2015).

The consensus view of AGN-driven outflows consists of an inner, line-driven disc wind with mildly relativistic velocities (also called "nuclear wind"). This highly ionised wind is shocked once it encounters the denser ISM and drives a forward shock into the ISM, which sweeps up most of the gas (see Figure 1.7). As we discuss in Chapter 6, the details of the driving mechanism depend on the efficiency of cooling in the shocked wind: if the internal energy is radiated away, only the momentum input from the AGN drives the outflow ("momentum-driven"). If cooling is inefficient, the hot shocked wind adiabatically expands and drives the shock into the ISM ("energy-driven"). These different driving mechanisms have fundamental influence on the outflow dynamics and the thermal state of the outflow. Most of the thermal energy is carried by the shocked wind and the shocked ISM accounts for the bulk part of the kinetic energy.

AGN-driven outflows should not be confused with jets. Although both emerge from the accretion disc around massive BHs and affect the ISM and gas accretion on galactic scales, there are fundamental differences: while galactic winds are driven by the disc wind, jets are powered via the Blandford–Znajek mechanism (Blandford & Znajek 1977) or via magnetic fields. Moreover, a jet is well confined with a small opening angle, whereas AGN-driven winds have a wider opening angle or are even assumed to be isotropic.

Several groups have developed 1D analytical models of AGN-driven outflows in these two different regimes (Silk & Rees 1998; Fabian 1999; King 2003; Murray et al. 2005; Silk & Nusser 2010; Ishibashi & Fabian 2012; Zubovas & King 2012; Faucher-Giguère & Quataert 2012). While these models can reproduce several observations, the assumption of a spherically symmetric gas distribution is too simplistic to capture the complexity of a realistic galaxy. Moreover, these models often assume an energy-

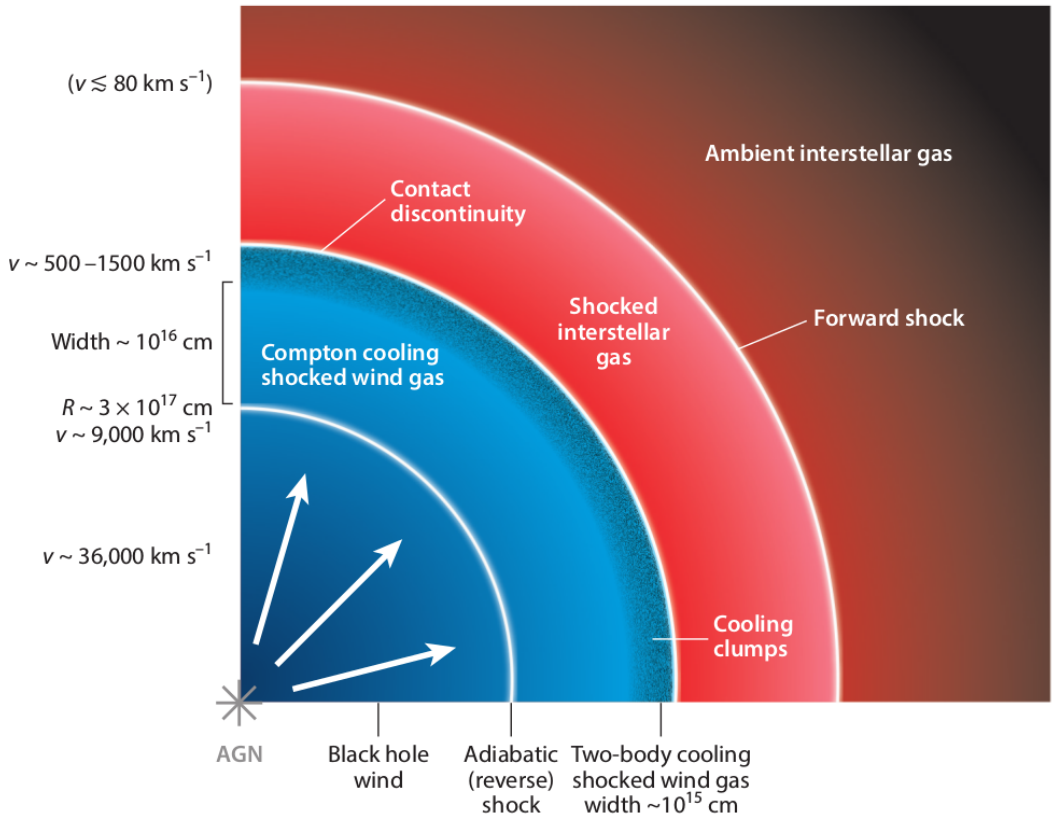


Figure 1.7: Illustration of the shock pattern resulting from the impact of a disc wind (blue) on the ISM (red). The fast inner wind collides with the ambient gas in the host galaxy and is slowed down in a strong shock. The shocked gas exerts a pressure on the ISM and sweeps it up into a dense shell. The shell's motion then drives a milder outward shock into the ambient ISM. Adapted from King & Pounds (2015).

or momentum-driven outflow, without self-consistently solving for the corresponding transition. For a momentum-driven outflow, you find a characteristic BH mass of (King & Pounds 2015)

$$M_\sigma = \frac{f_{\rm gas}\kappa}{\pi G}\sigma^4, \tag{1.49}$$

with the gas mass fraction $f_{\rm gas}$ and gas opacity κ, above which the AGN can eject the gas out of the gravitational potential of its host galaxy and hence prevents further BH growth. The $\propto \sigma^4$ scaling is independent of the underlying density distribution (McQuillin & McLaughlin 2012). If, on the other hand, we assume an energy-driven outflow, this characteristic mass and its scaling with the velocity dispersion is given by

$$M_\sigma = \frac{11 f_{\rm gas}\kappa}{\eta\pi G^2 c}\sigma^5, \tag{1.50}$$

where η is proportional to the mechanical luminosity (Silk & Rees 1998; Haehnelt et al. 1998; Fabian 2012; McQuillin & McLaughlin 2013). Observationally, different values for the exponent α of the M-σ relation, $M_{\rm BH} \propto \sigma^\alpha$, are proposed, ranging from $\alpha = 3.7$ (Gebhardt et al. 2000) to $\alpha = 5.6$ (McConnell & Ma 2013) with other values within this range (see e.g. Ferrarese & Merritt 2000; Tremaine et al. 2002; Gültekin et al. 2009; Kormendy & Ho 2013).

2 How an improved implementation of H_2 self-shielding influences the formation of massive stars and black holes

This chapter is based on the published paper "How an improved implementation of H_2 self-shielding influences the formation of massive stars and black holes" (Hartwig et al. 2015c).

2.1 Motivation

As we have seen in the introduction, one possible pathway to the formation of massive BH seeds is the direct collapse of gas in a pristine halo. One often assumes that a photodissociating LW background keeps the gas atomic at around 10^4 K to obtain the required high mass inflow rates and prevent fragmentation.

An important process in the thermodynamics of this collapse is H_2 self-shielding against LW radiation. It is generally expressed as a suppression factor $f_{\rm sh}$ to the H_2 photodissociation rate (see section 2.2.4). There are several analytic expressions to calculate $f_{\rm sh}$ as a function of the H_2 column density and the gas temperature (Draine & Bertoldi 1996; Wolcott-Green et al. 2011; Richings et al. 2014). Neglecting self-shielding underestimates $J_{\rm crit}$ (Shang et al. 2010) and even among these analytic functions, the value of $J_{\rm crit}$ varies by an order of magnitude (Latif et al. 2014; Sugimura et al. 2014), due to different assumptions of the gas dynamics, which demonstrates the importance of a correct treatment of this effect. Another challenge is the proper determination of the effective H_2 column density for self-shielding, since it is either computationally very expensive or not very accurate. In this study, we test the effect of a more accurate H_2 self-shielding implementation on the formation of massive stars and the direct collapse scenario. In contrast to previous studies, we determine the column densities self-consistently during the simulation and properly account for the Doppler shifts of spectral lines by velocity gradients, which reduce the effective column density. To do so, we use the TREECOL algorithm developed by Clark et al. (2012a) and extended by Hartwig et al. (2015b) to calculate a spherical map of column densities around each cell. This method is based on the hierarchical tree structure used to calculate the gravitational forces between fluid elements in the computational domain and therefore comes with only little additional cost. A more detailed description of this method is provided in Section 2.2.4.

2.2 Methodology

In this section, we present our computational methods. First, we briefly introduce the hydrodynamical code we use and describe the initial conditions and refinement strategy. Then, we present the chemical network and our new approach for the determination of effective H_2 column densities for self-shielding.

2.2.1 Moving-mesh code AREPO

We want to simulate the gravitational collapse of a gas cloud without following the time evolution of every single atom. There are generally two different numerical approaches, one grid-based, the other particle-based. In the first method, the introduction of a spatial grid, discretises space, and the quantities of interest are then assigned to the grid cells (Eulerian approach). In many cases the cell size is refined adaptively to increase resolution when specific conditions are met (Adaptive Mesh Refinement, AMR). Alternatively, we can discretise mass by introducing individual particles which carry information about mass, velocity, and other hydrodynamic quantities (Lagrangian approach). In this case force and hydrodynamical quantities are softened and smoothed over a certain region or a few particles to avoid spurious effects (SPH). For our study of H_2 self-shielding, we use the moving-mesh code AREPO, which is based on a moving unstructured mesh defined by the Voronoi tessellation. It combines the advantages of the two methods: the resolution follows the density automatically and continuously (like in SPH) and the finite volume discretisation yields accurate treatment of instabilities and shocks (like in AMR). The moving-mesh approach therefore avoids the main problems of noise and diffusiveness in SPH simulations and lack of Galilean-invariance in grid-based codes.

To understand the basics of computational hydrodynamics, we briefly review the underlying theory (mainly based on Landau & Lifshitz 1987; Klessen 2002; Hartwig 2014). The time evolution of the gas is given by the four equations of hydrodynamics:

- The conservation of mass is formulated by the continuity equation

$$\frac{\mathrm{d}\rho}{\mathrm{dt}} = \frac{\partial\rho}{\partial t} + \vec{v}(\vec{\nabla}\rho) = -\rho(\vec{\nabla}\vec{v}), \tag{2.1}$$

  where ρ is the density and $\vec{v}$ is the velocity.

- The conservation of momentum is formulated by the Navier-Stokes equation

$$\frac{\mathrm{d}\vec{v}}{\mathrm{dt}} = \frac{\partial\vec{v}}{\partial t} + (\vec{v}\vec{\nabla})\vec{v} = -\frac{\vec{\nabla}p}{\rho} - \vec{\nabla}\phi + \eta\vec{\nabla}^2\vec{v} + \left(\zeta + \frac{\eta}{3}\right)\vec{\nabla}(\vec{\nabla}\vec{v}) \tag{2.2}$$

  with the pressure p, the gravitational potential ϕ and the viscosity coefficients η and ζ.

- The energy equation formulates the conservation of energy

$$\frac{\mathrm{d}\epsilon}{\mathrm{dt}} = \frac{\partial\epsilon}{\partial t} + \vec{v}(\vec{\nabla}\epsilon) = T\frac{\mathrm{ds}}{\mathrm{dt}} - \frac{p}{\rho}(\vec{\nabla}\vec{v}), \tag{2.3}$$

  where ϵ is the energy density, T is the temperature and s is the entropy.

- The Poissons equation

$$\vec{\nabla}^2\phi = 4\pi G\rho \tag{2.4}$$

  relates the gravitational potential to the matter density.

In order to solve the hierarchically nested Maxwell-Boltzmann transport equations, we need a closure equation, where one normally uses the equation of state

$$p = K\rho^\gamma, \tag{2.5}$$

with the adiabatic index γ. These five equations are the theoretical framework of hydrodynamics and further heating and cooling processes can be coupled via the ds-term in the energy equation.

The main challenge of gravitational N-body codes is the calculation of the gravitational force which classically scales as N^2. AREPO solves this problem by clustering the particles into so called "tree nodes" (see Figure 2.1), which reduces the computational effort to a $N \log N$-scaling. The general idea of a tree

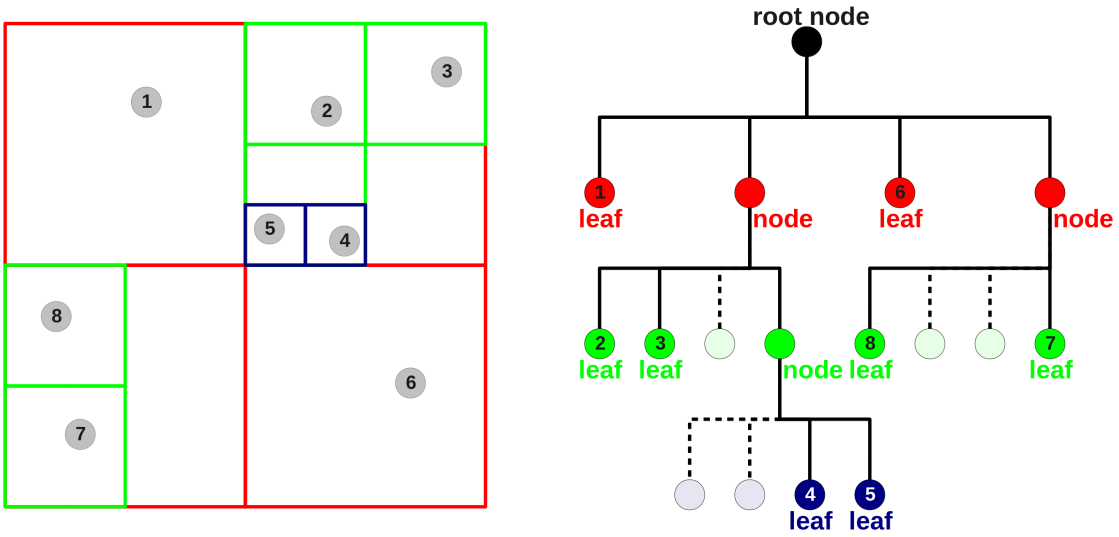


Figure 2.1: Illustration of the tree concept. The simulation volume is decomposed (left) and the resolution elements are clustered into tree nodes (right) to efficiently calculate the gravity. We use a regular grid for better illustration, although the grid in AREPO is an unstructured mesh, generated via Voronoi tessellation. Adapted from Hartwig (2014).

is to cluster all particles that lie within a certain region into one tree node. Hence, we can determine the gravitational potential of the nodes rather than the individual particles. The gravitational force is then determined by "walking" the tree and summing up the contributions of the nodes, which yields a relative error below 1%. In the same way, we will calculate the gas column densities, by summing up matter contributions in different directions, which are clusters in tree nodes.

2.2.2 Initial conditions

We are interested in the collapse of the subset of metal-free halos that is able to cool by atomic hydrogen. The cooling rate of H rises steeply around $T \simeq 10^4\,\mathrm{K}$ and expressed by the virial mass, we are interested in halos with $M_{\mathrm{vir}} \simeq 10^7\,\mathrm{M}_\odot$ at $z \gtrsim 10$. In this study, we focus on the effect of different H_2 self-shielding implementations. We first run a cosmological dark matter only simulation and select the first halos with a mass of $\sim 10^7\,\mathrm{M}_\odot$. Under the assumption that the value of J_{crit} is mainly affected by the gas dynamics within the virial radius, this is a representative candidate for a metal-free halo that directly collapses to a SMBH seed. We create the initial density field at redshift $z = 99$ in a periodic box of $1\,\mathrm{cMpc}\,h^{-1}$ with MUSIC (Hahn & Abel 2011), which generates the displacements and velocities following second-order Lagrangian perturbation theory.

We first run a cosmological simulation with $\sim 2 \times 10^7$ Voronoi cells, which corresponds to a particle mass of $m_{\mathrm{DM}} = 5.1 \times 10^4\,\mathrm{M}_\odot$. We trace the target halo and a region of twice its virial radius to the initial conditions. In a second run, we refine this region of interest and also include gas, which leads to masses

in the highest refined region of $m_{\rm DM} = 100\,{\rm M}_\odot$ and $m_{\rm gas} = 18\,{\rm M}_\odot$. This resolution is set to properly resolve the collapse of the gas up to densities of $n \simeq 10^6\,{\rm cm}^{-3}$. This value is chosen to cover the local thermodynamical equilibrium (LTE) of H_2 at around $n \simeq 10^4\,{\rm cm}^{-3}$, because above this value it is much easier to collisionally dissociate H_2 than at lower densities. Hence, if the gas reaches this value without building up a significant fraction of H_2, it is not going to manage to do so at higher densities either. Regan et al. (2015) study the effect of the dark matter mass resolution that is needed to properly resolve the collapse of halos at high redshift. They find that for typical collapse scenarios with a moderate LW background, $m_{\rm DM} = 100\,{\rm M}_\odot$ is a sufficient resolution, whereas this minimum mass resolution even decreases for the higher LW backgrounds ($J_{\rm LW} = 500$) that we want to study here. Consequently, our dark matter mass resolution is sufficient to properly resolve the collapse.

The finite box size of our simulations might distort the nonlinear effective coupling on the boxscale by not covering all relevant Fourier modes of the power spectrum (see e.g., Seto 1999). However, the choice of our box size is well motivated for a cosmological representative selection of a $10^7\,{\rm M}_\odot$ halo, because we do not want to draw high-precision cosmological probes from these simulations, but rather analyse the collapse behaviour of one specific halo. Another effect of the limited box size is the distortion of the large-scale tidal fields, which might affect the angular momentum of the halos, according to tidal torque theory (see e.g., Fall & Efstathiou 1980). However, the angular momentum budget of halos is dominated by local effects like mergers or the accretion of cold gas streams and only a minor contribution comes from cosmic tidal fields (Danovich et al. 2012; Dubois et al. 2014; Laigle et al. 2015). In any case, the effect of different implementations of H_2 self-shielding on the collapse dynamics should not be affected by the limited box size.

2.2.3 Chemistry

In the following section, we describe the chemical network that we have implemented in AREPO and highlight the most important reactions and rate coefficients. A more extensive discussion of the relevant chemical processes for modelling direct collapse with a strong LW background can be found in Glover (2015a,b). We apply a primordial chemistry network that is originally based on the work by Glover & Jappsen (2007), Glover & Abel (2008), and Clark et al. (2011a). Since the deuterium chemistry does not affect the direct collapse scenario (Glover 2015a), we only follow explicitly the evolution of H, He, H_2, H^+, H^-, H_2^+, He^+, He^{++}, and e^-. Glover (2015a) identifies a minimal subset of reactions that must be included in the chemical model in order to determine $J_{\rm crit}$ accurately. We have made sure to include all of these reactions in our chemical network. Full details regarding our choice of reaction rate coefficients can be found in Clark et al. (2011a) and Glover (2015a,b).

The collapse dynamics depends strongly on the abundance of molecular hydrogen, which is the dominant coolant for temperatures below $\sim 10^4\,{\rm K}$. As we have seen before, molecular hydrogen is mainly formed via the two-step process:

$$\mathrm{H} + \mathrm{e}^- \rightarrow \mathrm{H}^- + \gamma \tag{2.6}$$

$$\mathrm{H} + \mathrm{H}^- \rightarrow \mathrm{H}_2 + \mathrm{e}^-, \tag{2.7}$$

and is primarily destroyed either by collisions with hydrogen atoms

$$\mathrm{H}_2 + \mathrm{H} \rightarrow \mathrm{H} + \mathrm{H} + \mathrm{H}, \tag{2.8}$$

or by the so-called Solomon process (Field et al. 1966; Stecher & Williams 1967):

$$H_2 + \gamma \rightarrow H + H, \tag{2.9}$$

where a LW photon photodissociates the molecule by exciting it from the electric ground state into an excited electronic state. In $\sim 15\%$ of the cases, the electrons do not decay into a bound state, but into the vibrational continuum of the ground state and thereby dissociate the molecule. For the Solomon process, we assume a spectrum of a blackbody with an equivalent temperature of $T_{\rm rad} = 10^5\,{\rm K}$ (T5) that is cut off above 13.6 eV, because photons with higher energies are absorbed by the intergalactic medium. This choice represents the case in which the spectra are dominated by Pop III stars. The second generation of stars is believed to be less massive and can be approximated by a blackbody spectrum with $10^4\,{\rm K} < T_{\rm rad} < 10^5\,{\rm K}$ (Sugimura et al. 2014; Agarwal & Khochfar 2015). It is important to note that the value of $J_{\rm crit}$ depends on the choice of the spectrum, because whereas a T5 spectrum mainly photodissociates the H_2 directly (Equation 2.9), a spectrum with a cooler effective temperature dominantly prohibits H_2 formation by photodetachment of H^- via

$$H^- + \gamma \rightarrow H + e^- \tag{2.10}$$

(Latif et al. 2014; Sugimura et al. 2014; Agarwal & Khochfar 2015; Agarwal et al. 2016b). Since we want to focus on the effects of H_2 self-shielding, we will only consider the T5 spectrum, where this effect plays a dominant role. We also include dissociative tunnelling (Martin et al. 1996), which significantly contributes to the total collisional dissociation rate of H_2 and is therefore necessary for a proper treatment of primordial gas physics (Latif et al. 2014; Glover 2015a).

2.2.4 H_2 self-shielding

Since one photon of the external radiation field can only photodissociate one H_2 molecule, a large column density of molecular hydrogen can protect the inner regions against photodissociation. This process is known as self-shielding. For the photodissociation of H_2, we use the rate coefficient (Glover & Jappsen 2007)

$$k = 1.38 \times 10^{-12} \frac{J_{\rm LW}}{J_{21}} f_{\rm sh}\,{\rm s}^{-1}, \tag{2.11}$$

where the factor $f_{\rm sh} \leq 1$ accounts for the effect of H_2 self-shielding with $f_{\rm sh} = 1$ in the optically thin limit. This rate coefficient corresponds to a normalisation of the radiation field of $J_{\rm LW} = 1$ at 12.87 eV, the middle of the LW bands. The exact treatment of H_2 self-shielding requires a full radiative scheme with line transfer of all the important lines in the LW bands and is therefore prohibitively expensive. However, the shielding factor can be approximately expressed as a function of the H_2 column density and the gas temperature. The latter enters because of the thermal broadening of spectral lines and due to the temperature-dependent excitation of different rotational levels of the H_2 molecule. Draine & Bertoldi (1996) study the structure of stationary photodissociation fronts and propose a self shielding factor of the form

$$f_{\rm sh} = \frac{0.965}{(1+x/b_5)^\alpha} + \frac{0.035}{(1+x)^{0.5}} \exp\left[-\frac{(1+x)^{0.5}}{1180}\right], \tag{2.12}$$

where $x = N_{\rm H_2}/5 \times 10^{14}\,{\rm cm}^{-2}$ is the H_2 column density, $b_5 = b/10^5\,{\rm cm\,s}^{-1}$ is the scaled Doppler parameter of the molecular hydrogen, and $\alpha = 2$. This functional form accounts for the effect of line overlap and has been applied in many studies (Glover & Jappsen 2007; Whalen & Norman 2008; Gnedin

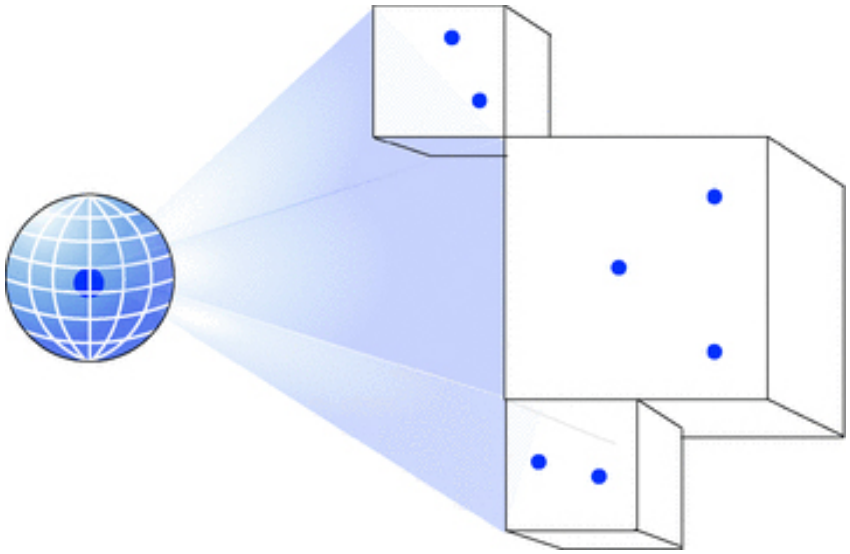


Figure 2.2: Schematic illustration of the TreeCol algorithm. The particles are already clustered into tree nodes (boxes). TreeCol creates a spherical map around each particle by projecting the mass contributions of the tree nodes onto the spherical grid. Adapted from Clark et al. (2012b).

et al. 2009; Shang et al. 2010; Glover et al. 2010; Christensen et al. 2012; Krumholz 2012).

Wolcott-Green et al. (2011) model the photodissociation and self-shielding of H_2 in protogalaxies with three-dimensional simulations, based on post-processing their output data. They found that the formula by Draine & Bertoldi (1996) is only valid for cold or low density gas, in which only the lowest rotational states of H_2 are populated and propose the modification $\alpha = 1.1$ to equation (2.12), which provides a better fit in dense gas for all relevant temperatures. A comparison of both functions (Latif et al. 2014; Sugimura et al. 2014) shows that the application of the newer formula by Wolcott-Green et al. (2011) yields values of $J_{\rm crit}$ that are up to an order of magnitude lower than those derived with the function by Draine & Bertoldi (1996). Another functional form was proposed by Richings et al. (2014) who model shielding against UV radiation in the diffuse interstellar medium. They also include the effect of turbulent gas velocities, but they derive the self-shielding factor for a one-dimensional plane-parallel slab of gas.

In a more realistic scenario, however, we are interested in a collapsing cloud, where the relative velocities between infalling gas particles Doppler-shift the spectral lines. Due to this effect, an H_2 molecule can only shield other H_2 molecules whose relative velocity is smaller than its thermal velocity. Otherwise, the spectral lines are shifted too far and H_2 molecules do not contribute to the effective column density. To account for this effect, Hartwig et al. (2015b) have implemented a new method for the determination of effective column densities in three-dimensional simulations. This method is based on the TREECOL algorithm by Clark et al. (2012a), which directly sums up the individual mass contributions for the column density of each fluid element (see illustration in Figure 2.2). With this information, we create spherical maps of the column density around each Voronoi cell, with 48 equal-area pixels based on the HEALPIX algorithm (Górski et al. 2005). The number of 48 pixels is motivated by HEALPIX, which divides the sphere into 12 equal-area pixels, which can be subdivided into 2^N pixels each. Based on the analysis by Clark et al. (2012a), we chose $N = 2$, since this value provides a sufficient angular resolution for most astrophysical applications, resulting in a solid angle of $\sim 0.26\,\mathrm{sr}$ per pixel.

Two characteristics of the code make TREECOL highly useful for our purpose. First, it uses the tree structure, which is already present in the code to determine the gravitational force. Hence, the deter-

mination of column densities comes with only little additional computational cost. Secondly, we can directly compare the relative velocities of the particles v_r that possibly contribute to the column density to the thermal gas velocity $v_{\rm th}$ of the particle for which we want to calculate the column density. Following Hartwig et al. (2015b), a particle only contributes to the effective column density, if

$$v_r < 1.694 v_{\rm th}, \tag{2.13}$$

where $v_{\rm th}$ is the thermal gas velocity and the numerical factor is a correction of the so-called Sobolev approximation (Sobolev 1960) and takes the true line profile into account. Based on this criterion, we determine the effective column densities and calculate the shielding factors separately for all pixels with equation (2.12) and the exponent $\alpha = 1.1$ by Wolcott-Green et al. (2011). The final shielding factor is the mean of 48 directional-dependent factors. The averaging has to be weighted by the direction-dependent flux if one assumes an anisotropic LW radiation field, such as Regan et al. (2014b, 2016).

With this new approach, we can use formula (2.12) with the Wolcott-Green et al. (2011) exponent $\alpha = 1.1$, which was derived for static gas and extend it to collapsing gas clouds by defining the effective H_2 column densities based on the relative gas velocities. Since our approach automatically accounts for turbulent motions on scales above the spatial resolution, the Doppler parameter b_5 in formula (2.12) does only include the thermal broadening. This approach cannot only be used for H_2 self-shielding, but also for many other radiative transfer processes that rely on the determination of column densities. The only requirement is that the resolution elements are stored in a tree-like structure, which is already the case in most codes that include self-gravity. This method for the determination of effective column densities is tested and explained in more detail in Hartwig et al. (2015b).

Our method of computing effective H_2 column densities is valid as long as the main contribution comes from the core of individual lines that shield themselves. At high column densities, however, the Lorentzian contribution to the line profile becomes important and the corresponding damping wings should be taken into account for the determination of self-shielding. While this effect is negligible at small column densities, Gnedin & Draine (2014) show that it should be taken into account for H_2 column densities of $N_{\rm H_2} > 10^{21}\,{\rm cm}^{-2}$. At these high column densities, the Doppler-shifts induced by relative velocities become less important and eventually the total column density contributes to self-shielding. A more detailed analysis of this effect is given in section 2.3.4, where we show that a correct treatment of the overlap of these damping wings changes the value of the self-shielding factor by less than 5% and has no influence on the determination of $J_{\rm crit}$. In addition, the use of effective column densities is computationally more efficient, since the velocity criterion imposed by equation (2.13) limits the amount of fluid elements that have to be projected. Therefore, we employ this method in our numerical simulations.

We also note that Safranek-Shrader et al. (2012a) use a non-local approach for the determination of H_2 column densities. They study the influence of LW radiation on star formation in the first galaxies and approximate the column density in the following way. For each computational cell, they calculate the column densities in six directions parallel to the coordinate axis. The smallest of these column densities is then used to calculate the self-shielding factor. They show that this is already an improvement, compared to local estimations of the column density. Similar techniques have also been used to study the effects of H_2 self-shielding in giant molecular clouds (see e.g. Nelson & Langer 1997; Glover & Mac Low 2007a,b).

Most previous simulations use an approximation for the H_2 column density, based on a characteristic length scale $L_{\rm char}$, and the assumption that the H_2 density is constant within this length and negligible

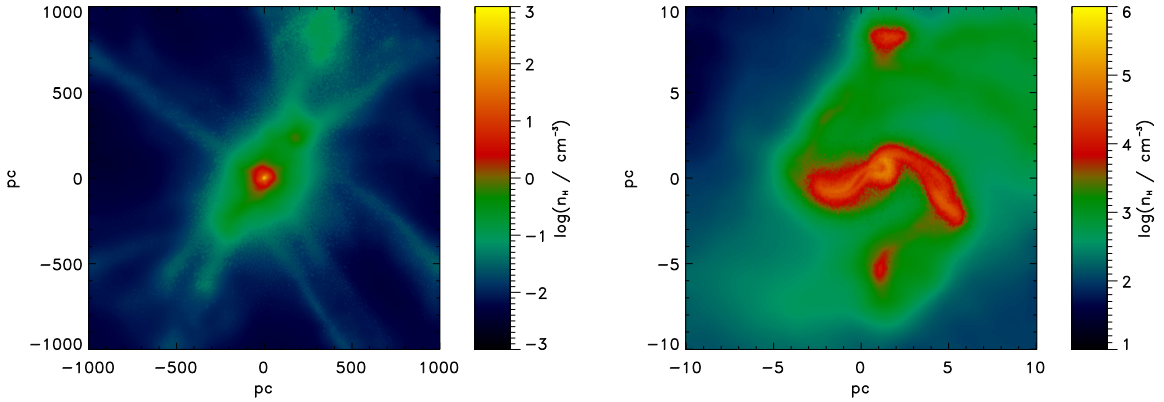


Figure 2.3: Maps of the average number density of hydrogen nuclei along the line of sight for halo C with the TREECOL approach at the moment of collapse ($z \sim 15.1$) at a scale of 2000 pc (left) and 20 pc (right). The background flux is $J_{\rm LW} = 10^3 < J_{\rm crit}$, and we can clearly see the formation of clumps in the central region.

beyond it. The column density is then given by $N_{\rm H_2} = n_{\rm H_2} L_{\rm char}$, where $n_{\rm H_2}$ is the local number density of molecular hydrogen. Assuming that the effect of self-shielding occurs only locally, many simulations (e.g. Shang et al. 2010; Van Borm et al. 2014; Sugimura et al. 2014; Latif et al. 2015; Latif & Volonteri 2015; Glover 2015a,b; Agarwal et al. 2016b) use $L_{\rm char} = L_{\rm J}$ with the Jeans length

$$L_{\rm J} = \sqrt{\frac{15 k_B T}{4\pi G \mu m_p \rho}}. \tag{2.14}$$

Since this is a widely used approximation, we will compare our results obtained with TREECOL to this local approximation.

2.3 Results

In order to increase the statistical significance of our results, we create four independent sets of cosmological initial conditions. The side length of each of these boxes is 1 cMpc h^{-1}, and we select one halo per box. As described in section 2.2.2, the region of interest is refined for each box and we resimulate this set of cosmological zoom-in simulations. At redshift $z = 30$, we switch on the photodissociating background. In reality, the LW radiation increases with cosmic time and the time of its onset also influences the collapse of primordial gas clouds (Visbal et al. 2014b). We use this simplification of an instantaneous onset to be able to focus on the different implementations of the H_2 self-shielding and make it comparable to most previous works in this field. As long as the H_2 abundance has enough time to reach an equilibrium, before the halo of interest starts with the run-away collapse, the results are unaffected by the choice of this redshift. This criterion is fulfilled in all our simulations.

Altogether, we study the collapse of four different halos (A, B, C, and D) with two methods (TREECOL and Jeans approximation) for determining the column density and several different strengths of the LW background. All these runs are independent and the column densities are calculated self-consistently during the simulation. As an example, the structure of halo C is illustrated in Figure 2.3. The halo is embedded in the cosmic web and is fed by several gas streams. The central region shows a lot of substructure and gas clumps, which indicate ongoing fragmentation. The LW radiation is too weak to prevent efficient H_2 cooling and the gas can locally contract before the cloud globally collapses.

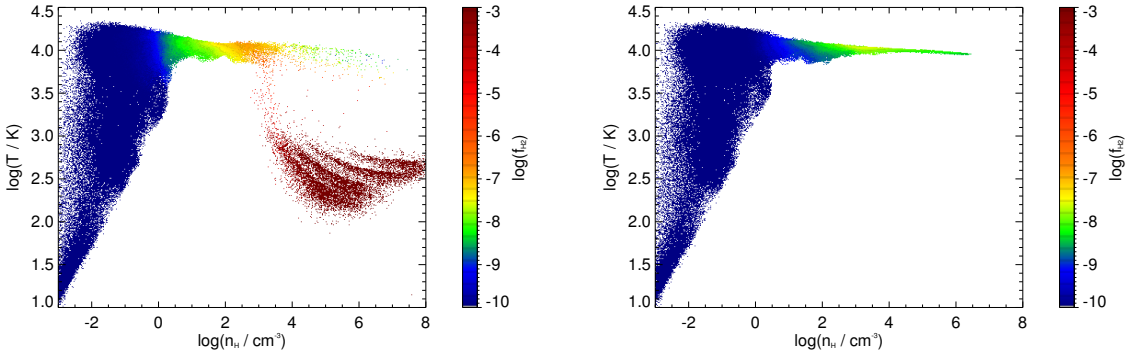


Figure 2.4: Temperature as a function of number density of hydrogen nuclei colour-coded by the H_2 fraction for halo C using TREECOL with $J_{\rm LW} = 10^3$ (left) and $J_{\rm LW} = 10^4$ (right). To better highlight the relevant difference at higher H_2 fractions, we artificially set a lower limit of $f_{H_2} = 10^{-10}$ in these plots. We can clearly see the different collapse behaviours depending on the strength of the LW background. With a high $J_{\rm LW}$, the gas remains hot around 10^4 K with $f_{H_2} \lesssim 10^{-7}$. For a lower strength of the photodissociating background, the fraction of molecular hydrogen rises up to $f_{H_2} \simeq 10^{-3}$ and the gas cools down to several hundred Kelvin. In the latter case, small clumps decouple from the global thermal evolution and we see their imprints as stripe-like structures in the cold high-density gas.

2.3.1 Determination of $J_{\rm crit}$

The value of $J_{\rm crit}$ sets the threshold between the two different collapse regimes. Above this value, the H_2 fraction remains low, the temperature stays around 10^4 K during the collapse, and only one central density peak forms. Below this value, the photo-dissociating background is not strong enough and H_2 line emission cools the collapsing gas to a few hundred K. This collapse typically results in several fragments (Clark et al. 2011a,b; Greif et al. 2011; Stacy et al. 2012; Hirano et al. 2014). In order to discriminate between these two scenarios, we have to analyse the temperature evolution during the collapse. For two typical cases, the phase diagrams are given in Fig. 2.4. During the virialisation of the halo, the gas shock heats to the virial temperature of around 10^4 K. The gas contracts further and remains at this temperature due to cooling by atomic hydrogen. With increasing density, also the column density of H_2 increases and the self-shielding against the external radiation becomes more efficient. As discussed above, H_2 reaches LTE at $n \simeq 10^4\,{\rm cm}^{-3}$, and if the collapse is still isothermal up to these densities, it will proceed isothermally. If the LW background is not strong enough, the H_2 fraction can increase, which in turn increases the column densities and hence leads to a more efficient self-shielding, which consequently increases the H_2 fraction even further. This runaway production of H_2 enables a clear distinction between the two different collapse regimes. An H_2 fraction of $\sim 10^{-3}$ is sufficient to cool the gas to temperatures of a few hundred K and hence to induce gas fragmentation. The individual clumps that form in this latter scenario can also be seen in the phase diagram as stripes in the cold, high-density regime, indicating that their thermal evolution is decoupled from one another.

In order to find the values of $J_{\rm crit}$, we compare the phase diagrams for different values of the LW background. The corresponding plots for the four halos and the two different methods are shown in Figure 2.5. We test and display the background strengths of $J_{\rm LW} = 10^3, 3 \times 10^3, 10^4$ for all halos and then successively bracket the actual value of $J_{\rm crit}$. We stop the simulations after the first snapshot with a peak density of $n \geq 10^6\,{\rm cm}^{-3}$ and compare the temperatures in the density regime above $n \geq 10^4\,{\rm cm}^{-3}$. If the temperature falls below 6000 K, the collapse is regarded as non-isothermal. We note that at 6000 K high gas infall rates are still possible, but once the gas temperature falls below this threshold, it cools

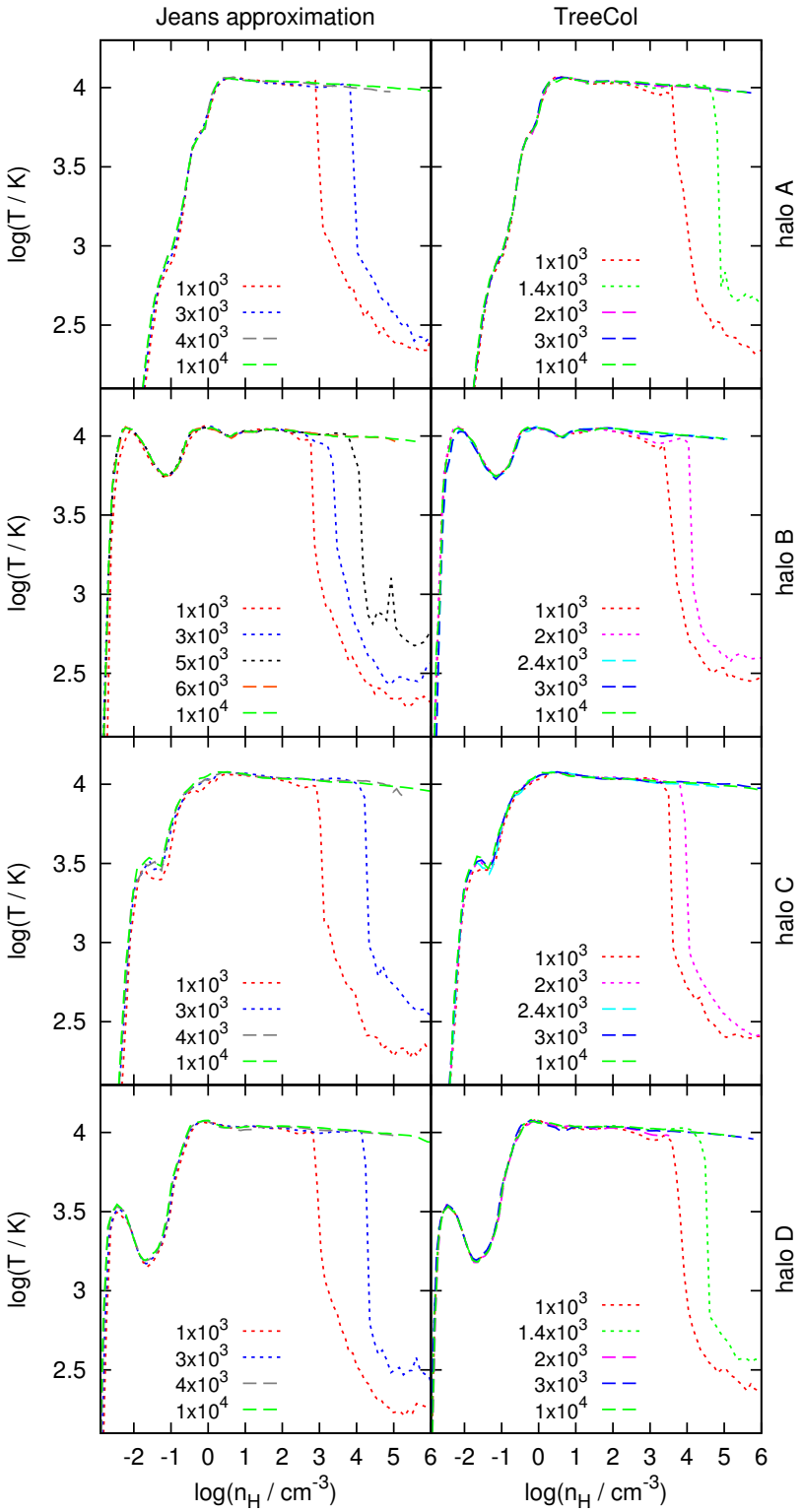


Figure 2.5: Temperature as a function of the number density of hydrogen nuclei for the four halos and the two different methods for determining the H_2 column density. The curves represent several realisations with different LW backgrounds, where the long-dashed lines represent the isothermal collapse and the short-dashed lines the collapse with efficient H_2 cooling. In this and the following figures, the values stated in the legend represent the LW background in units of J_{21}. From these plots, we can read off the critical value for the isothermal collapse, which is systematically lower for the runs based on the TREECOL method.

halo	$J_{\rm crit}$ (Jeans)	$J_{\rm crit}$(TREECOL)	$M_{\rm vir}/10^7\,{\rm M}_\odot$	$z_{\rm coll}$
A	3500	1700	1.8	17.9
B	5500	2200	1.2	14.4
C	3500	2200	1.7	15.1
D	3500	1700	1.1	12.9

Table 2.1: Critical values $J_{\rm crit}$ of the LW background for the four different halos and the two different column density approaches. The $J_{\rm crit}$ determined with the TREECOL method is smaller by about a factor of two in all halos. The halo-to-halo variance of this value is small and the collapse redshifts are distributed in a reasonable range. We also list the virial mass and the collapse redshift, which indicates the time when the halos first reach a density of $n \geq 10^6\,{\rm cm}^{-3}$.

generally to much lower temperatures (Fig. 2.5). With this method, we find the lowest value J_1 for which the collapse is still isothermal and the highest value J_2 for which H_2 can efficiently cool the gas. Due to a limited number of possible realisations, we define the final critical value as the geometrical mean between these two values

$$J_{\rm crit} = \sqrt{J_1 J_2}. \tag{2.15}$$

Due to this finite number of tested $J_{\rm LW}$, the final values have an uncertainty of $\sim 10\%$. The resulting values, the virial masses of the halos, and the collapse redshifts are compared in Table 2.1. First of all, we directly see that $J_{\rm crit}$ is about a factor of two lower in the runs where we use the TREECOL method to calculate the column densities. The reasons for this effect will be discussed in detail below. However, already the results based on the commonly used Jeans approximation are lower than found in previous studies. There are two main reasons for this. First, e.g. Latif et al. (2014) and Sugimura et al. (2014) show that the self-shielding function by Wolcott-Green et al. (2011) yields values of $J_{\rm crit}$ that are up to an order of magnitude lower, compared to those derived with the function by Draine & Bertoldi (1996). Secondly, the ENZO chemical model, which was used by many studies in this field (e.g. Shang et al. 2010; Wolcott-Green et al. 2011; Regan et al. 2014a; Latif et al. 2014), tends to overestimate $J_{\rm crit}$ by about a factor of two (Glover 2015a). Hence, our results based on the Jeans approximation are in agreement with previous studies.

2.3.2 Differences in the H_2 self-shielding

In order to understand the differences induced by the new TREECOL approach, we directly compare the column densities and corresponding self-shielding factors for the two methods. First, we summarise the most important features of this new method to understand these results. With TREECOL, we create a spherical grid with 48 pixels around each Voronoi cell and project the column densities on to this grid. However, we do not calculate the total column densities, because gas can only contribute to the self-shielding if its relative velocity is smaller than ~ 1.7 times the thermal velocity (Equation 2.13). Based on this criterion we determine the spherical maps of effective column densities, which represents the spatial distribution of the self-shielding gas. The self-shielding factors are then calculated based on the column density for each of the pixels separately and then averaged over the 48 directions. This procedure is physically motivated, because the product $J_{\rm LW} f_{\rm sh}$ in equation (2.11) represents an effective photodissociating flux and we simply average these fluxes over 48 different directions.

The importance of this point becomes clear, when we analyse the directional dependence of the column densities in Figure 2.6. We see that the column density distribution around one Voronoi cell is generally highly asymmetric and dominated by the contributions from one direction. In the case of a collapsing halo it is certainly the central high-density peak that yields the strongest contribution to the column

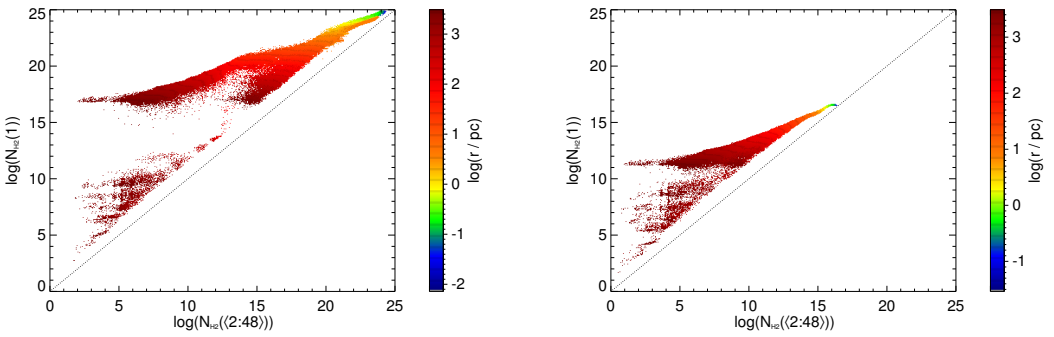


Figure 2.6: Plots of the maximal column density for each Voronoi cell $N_{\mathrm{H2}}(1)$ as a function of the mean H_2 column density averaged over the remaining 47 pixels $N_{\mathrm{H2}}(\langle 2:48 \rangle)$, colour-coded by the radius. On the left, we show the results for halo C and a flux of $J_{\mathrm{LW}} = 10^3 < J_{\mathrm{crit}}$ and on the right for the same halo but with $J_{\mathrm{LW}} = 10^4 > J_{\mathrm{crit}}$. This plot illustrates the large directional dependence of the column densities, because for one Voronoi cell, the column densities in different directions vary by up to ten orders of magnitude. Interestingly, we can also see the more spherically symmetric collapse structure in the isothermal case (right), because the plot converges towards the diagonal for small radii, indicating that the column densities are distributed isotropically close to the centre. Whereas in the left panel, the angular distribution of column densities remains anisotropic.

density. From this direction, obviously, we should not expect any photodissociating radiation. A different way of presenting this important directional dependence of the self-shielding factor is given in Figure 2.7. Averaging the column densities first and then calculating the shielding factor based on the one mean column density is highly biased by the contribution of one dominating direction. Consequently, it is important to properly average the effective photodissociating fluxes, as we do in our simulation. In contrast, a local approximation that only yields one column density might be affected by the central density peak and consequently underestimate the self-shielding factor.

Now, we compare directly the results for the column density and the shielding factor based on the two implementations. In Figure 2.8, we display the column densities of the simulations with the Jeans approximation and the TREECOL approach. These column densities are calculated self-consistently during the run and are therefore not directly comparable. However, we already see that the column densities in the isothermal case remain under $10^{16}\,\mathrm{cm}^{-2}$ and that, for smaller values of J_{LW}, the column density increases already for lower densities. To be able to directly compare the effect of the two different approaches, we use one snapshot of the simulations based on TREECOL and determine the corresponding Jeans column densities by post-processing the data. The comparison of these column densities is given in Figure 2.9. In the isothermal case, the column densities remain below $10^{16}\,\mathrm{cm}^{-2}$ and in this regime, the Jeans approximation overestimates the column densities by approximately one order of magnitude. Consequently, the shielding in this regime is more efficient with the Jeans approximation, and we need a higher LW background to obtain an isothermal collapse. For column densities above $10^{16}\,\mathrm{cm}^{-2}$, the Jeans approximation underestimates the values. We find that these regions, where the Jeans approximation underpredicts the column densities are lower density regions between the clumps (compare e.g. with the right panel in Figure 2.3). In this low-density environment, the Jeans approximation sees only the local gas conditions and predicts a small column density, whereas TREECOL is able to capture the nearby high-density clumps with high H_2 fractions. Consequently, TREECOL yields higher column densities in this regime. However, this underestimation does not affect the value of J_{crit}, because we only get $N_{\mathrm{H_2}} > 10^{16}\,\mathrm{cm}^{-2}$ in halos where $J_{\mathrm{LW}} < J_{\mathrm{crit}}$ and where the gas can undergo runaway cooling. Phrased differently, this underestimation by the Jeans approximation at higher column densities is a consequence

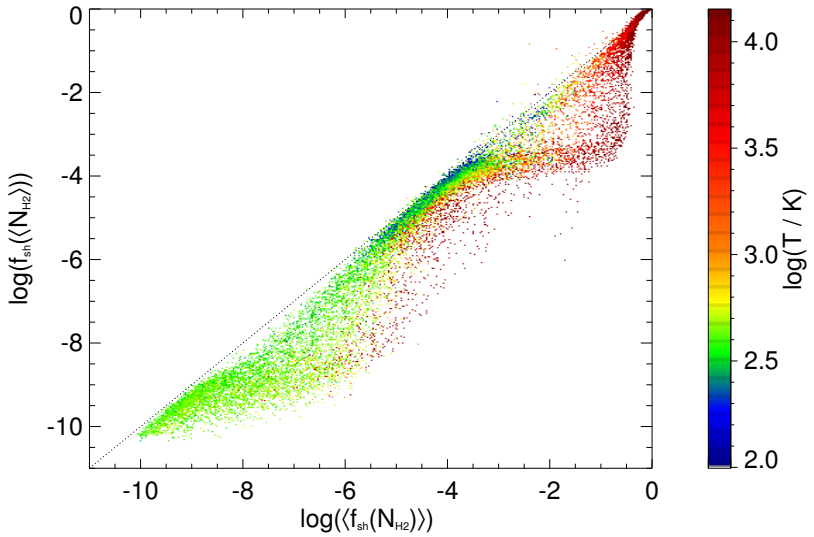


Figure 2.7: Self-shielding factor for Halo C with the TREECOL approximation and $J_{\rm LW} = 10^3 < J_{\rm crit}$, colour-coded by the gas temperature. This plot illustrates the directional dependence of the self-shielding factor. In our simulation, we determine individually the self-shielding factors for the 48 different directions and average them afterwards (horizontal axis of this plot). On the vertical axis we see the self-shielding factor for exactly the same simulation output, but here, we first average the 48 different column densities for each fluid element and calculate the self-shielding factor based on this one averaged column density. The proper direction-dependent treatment of the H_2 self-shielding generally yields a less efficient shielding against LW radiation. This discrepancy is smaller for temperatures of a few hundred Kelvin.

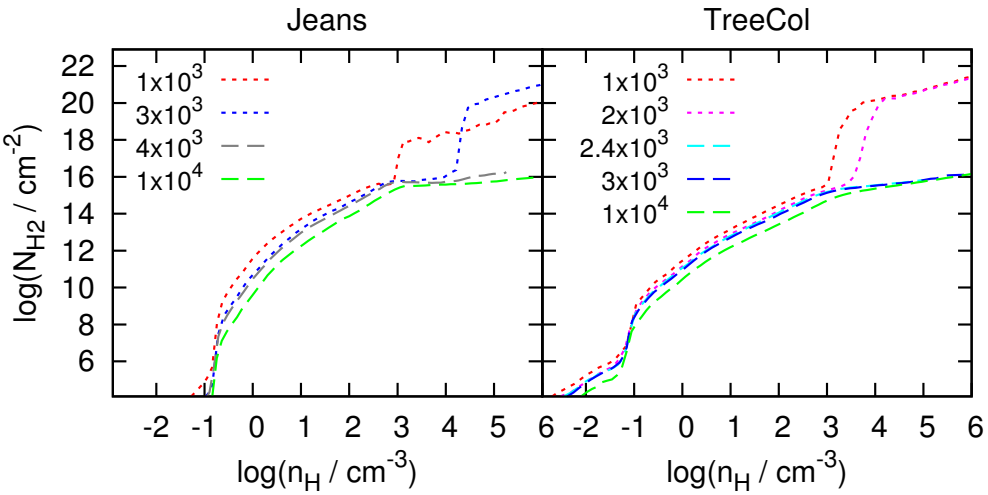


Figure 2.8: Molecular hydrogen column density as a function of the density of hydrogen nuclei for halo C. For the TREECOL runs, we plot the median column density over the 48 pixels in order not to be biased by one dominating direction. The values of the column density are calculated self-consistently during the simulation. Hence, they are not directly comparable for the two different approaches, because the structure of the cloud might be different. However, we can already see important similarities such as a threshold column density of about $10^{16}\,{\rm cm}^{-2}$ above which strong self-shielding enables the formation of sufficient H_2 to cool the cloud efficiently.

of the runaway H_2 cooling and the subsequent fragmentation and not its trigger.

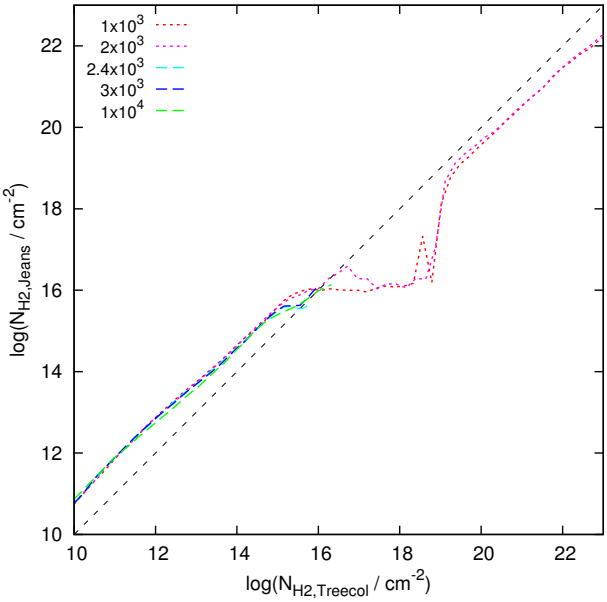


Figure 2.9: Direct comparison of the H_2 column densities for the two different approaches. The data presented here is based on an output of halo C simulated with TREECOL and the corresponding column densities based on the Jeans approximation are calculated based on the output file. Again, we use the median over the 48 pixels for the column densities based on TREECOL. For column densities below $10^{16}\,\mathrm{cm}^{-2}$, the Jeans approximation yields column densities that are higher by about one order of magnitude.

There are two obvious reasons why TREECOL yields lower effective column densities. First of all, it takes into account the three-dimensional distribution of the matter and the declining density radially outwards, whereas the Jeans approximation assumes a constant density within one Jeans length and hence generally overestimates the gas number density. Secondly, only fluid elements within a certain velocity range contribute to the effective column density in TREECOL, which again reduces the value of the column density. Wolcott-Green et al. (2011) have already pointed out that the Jeans approximation generally overestimates the column densities in a static, isothermal slab of gas. Our treatment of the relative velocities further increases this effect.

We can see the same trend for the self-shielding factors in Figure 2.10. In the case of an isothermal collapse, the TREECOL methods yields always higher values for the shielding factor and therefore a less efficient shielding against the LW background. Consequently, a smaller value of J_{crit} provides already enough photodissociating radiation to keep the collapse isothermal.

2.3.3 Impossibility of a simple correction factor

Although TREECOL is computationally efficient, it would be useful if there is a simple correction formula to the Jeans approximation that is able to reproduce the self-shielding results obtained from TREECOL at least to some degree. The aim would be to find an updated version of equation (2.12) that uses the column densities $N_{\mathrm{H}_2} = n_{\mathrm{H}_2} L_{\mathrm{J}}$ based on equation (2.14). As a start, we can impose the velocity criterion from equation (2.13) on the matter included in the calculation of the effective H_2 column density. This changes the dependence of the self-shielding factor on the thermal velocity, which is expressed by the exponent α in equation (2.12). Hence, we try to fit the results obtained with TREECOL with the free

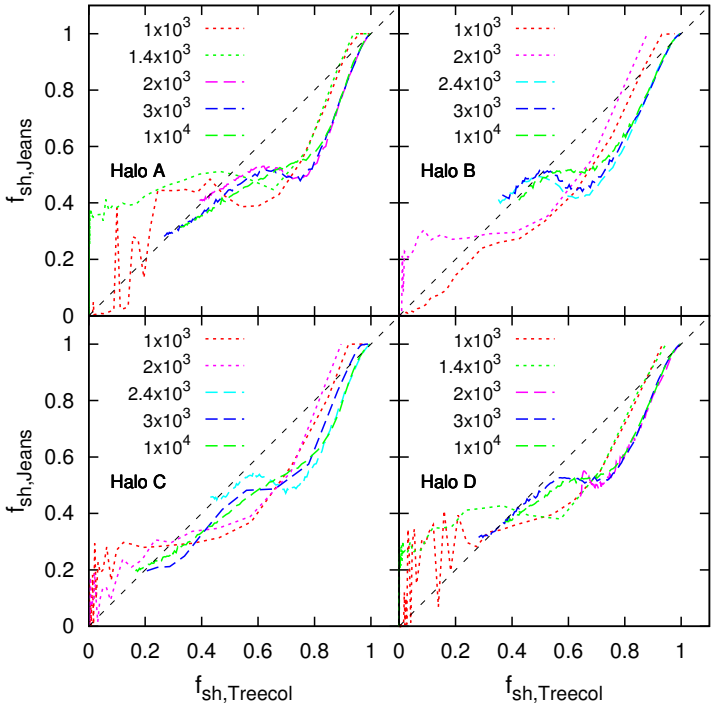


Figure 2.10: Direct comparison of the self-shielding factors for the two different approaches. The data presented here is based on halos simulated with TREECOL and the corresponding Jeans column densities are calculated based on the output files. In the isothermal regime (long-dashed lines), the TREECOL approximation yields higher self-shielding factors and therefore a less efficient shielding against the external photodissociating radiation field.

parameter α_* but with the column densities determined with the Jeans approximation as an input:

$$f_{\rm sh}(N_{\rm H_2,TreeCol})_{\alpha=1.1} = f_{\rm sh}(N_{\rm H_2,Jeans})_{\alpha_*} \tag{2.16}$$

However, this exponent varies with a large scatter between the different halos and with the various strengths of the background radiation. For example, for halo C we find $\alpha_* = 0.6$ for $J_{\rm LW} = 2400$ and $\alpha_* = 1.2$ for $J_{\rm LW} = 4000$, although both collapse isothermally. Most other values are distributed in the range $0.7 < \alpha_* < 1.1$, which illustrates that it is not possible to find a simple correction factor to reproduce the results by TREECOL. This is mainly based on the fact that a proper determination of the self-shielding factor has to take into account the three-dimensional structure of the density and velocity field, as we have seen before. Since the Jeans approximation is only based on local quantities, it cannot capture this structure and consequently fails at reproducing the self-shielding factors.

2.3.4 Effect of damping wings

So far, we assumed that only matter in a certain velocity range contributes to the self-shielding of H_2, because for larger relative velocities the spectral lines are Doppler-shifted too far from the core. As already mentioned in section 2.2.3, this picture changes for higher column densities, where the contribution of the damping wings becomes important and different lines can self-shield each other (Black & Dalgarno 1977; Draine & Bertoldi 1996). This effective broadening of spectral lines makes relative velocities less important and the total H_2 column density then contributes to self-shielding. The original self-shielding formula by Draine & Bertoldi (1996), Equation 2.12 in section 2.2.3, already includes these

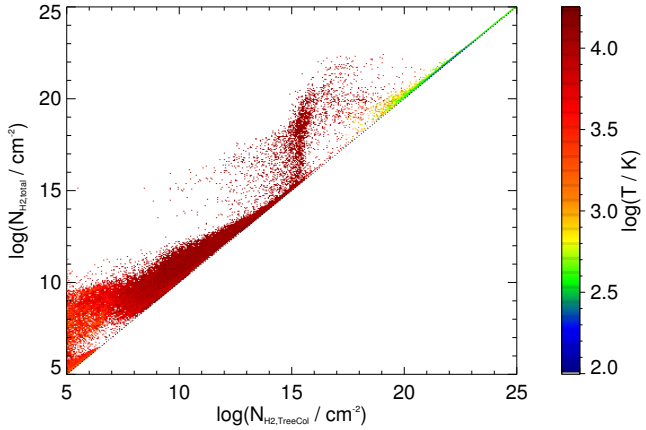


Figure 2.11: Total H_2 column density as a function of the effective H_2 column density colour-coded by the gas temperature for halo C with $J_{\rm LW} = 1000 < J_{\rm crit}$. For the effective column density, we impose the velocity criterion of Equation 2.13 and only matter fulfilling this criterion is included. The excess at $N_{\rm H_2,treecol} \simeq 10^{16}{\rm cm}^{-2}$ is created by the hot gas, which falls on to the cold and dense gas clumps and has therefore a high relative velocity. Hence, the velocity criterion excludes a significant amount of matter, and the total column density is higher in this regime.

two contributions, where the first term corresponds to shielding from the cores of individual lines, and the second term represents the effect of overlapping damping wings, which are not affected by relative velocities. Gnedin & Draine (2014) propose a correction to the self-shielding factor of the form

$$f_{\rm sh} = \frac{0.965}{(1+x_1/b_5)^{\alpha}} + \frac{0.035}{(1+x_2)^{0.5}} \exp\left[-\frac{(1+x_2)^{0.5}}{1180}\right], \tag{2.17}$$

where $x_1 = N_{\rm H_2,treecol}/5\times10^{14}{\rm cm}^{-2}$ is the effective H_2 column density and $x_2 = N_{\rm H_2,total}/5\times10^{14}{\rm cm}^{-2}$ is the total H_2 column density. Gnedin & Draine (2014) use the exponent $\alpha = 2$ based on the original work by Draine & Bertoldi (1996), whereas we apply the more recent value $\alpha = 1.1$ by Wolcott-Green et al. (2011).

We investigate the impact of the damping wings by determining the total H_2 column densities. To do so, we post-process the snapshots of halos C with TREECOL and compare them to the effective column densities (Figure 2.11). The total column density is higher by up to several orders of magnitude. Especially around $N_{\rm H_2,treecol} \simeq 10^{16}{\rm cm}^{-2}$, we see that the total column density is much larger than the effective one. This excess is created by the hot gas that falls towards the centre and has therefore a high relative velocity with respect to these central cold and dense gas clumps. Consequently, the velocity criterion excludes a large contribution, and the total column density is significantly higher than the effective one in this regime. In any case, this comparison strengthens the importance and illustrates the influence of the velocity criterion on the determination of the column densities for self-shielding.

We now use these two column densities and determine the self-shielding factors with equation (2.17) and compare them to the previously used values based on equation (2.12). In the case of $J_{\rm LW} = 10^4 > J_{\rm crit}$, the effective and the total column densities remain below $\sim 10^{16}{\rm cm}^{-2}$, and the contribution of damping wings is negligibly small: the mean deviation between the self-shielding factors is 0.24% with a maximal discrepancy of 0.60%. This is due to the fact that the main contribution to the self-shielding in this regime comes from the cores of individual lines and hence, from the first term in equation (2.17).

For the case with $J_{\rm LW} = 10^3 < J_{\rm crit}$, we have a higher H_2 abundance and consequently higher H_2 column densities. Here, the self-shielding factors differ on average by 2.2% with a maximal deviation of 5.3%. Hence, we conclude that the contribution of the damping wings to the self-shielding is negligibly small in our scenario, compared to other effects and approximations. In particular, it seems not to affect the determination of $J_{\rm crit}$, because this effect only becomes important, once the runaway H_2 production has already set in.

2.3.5 Mass infall rate

The main quantity that determines if a SMS and hence a seed of a SMBH forms is the mass infall rate $M_{\rm in}$, which has to be above $M_{\rm in} \gtrsim 0.1\,{\rm M}_\odot\,{\rm yr}^{-1}$ (Begelman 2010; Hosokawa et al. 2013; Schleicher et al. 2013; Latif et al. 2013b; Ferrara et al. 2014; Latif & Volonteri 2015). An isothermally collapsing cloud at $T \simeq 10^4\,{\rm K}$ is a sufficient criterion to provide the necessary mass infall rates, but is this criterion also necessary? To study this question, and to see if the H_2 self-shielding implementation induces any difference, we analyse the mass infall rate for the different halos and methods. Assuming a spherically symmetric cloud, we determine the mass infall rate by

$$M_{\rm in} = 4\pi r^2 |v_r| \rho, \tag{2.18}$$

where r is the distance from the densest point, v_r is the radial velocity, and ρ is the gas density. We average the radial velocity and the gas density over spherical shells. However, we note that the mass infall rate is generally not spherically symmetric and neither is it constant in time. We therefore take the mean over the last $\sim 10^5\,{\rm yr}$ of evolution to obtain a reasonably smooth and well-defined value (Figure 2.12). The mass infall rates in the isothermal collapse scenarios are in the range $0.1\,{\rm M}_\odot{\rm yr}^{-1} \lesssim M_{\rm in} \lesssim 1\,{\rm M}_\odot{\rm yr}^{-1}$ and tend to be higher than those in the cases with $J_{\rm LW} < J_{\rm crit}$. Especially in the central $\sim 10\,{\rm pc}$, only the isothermal clouds are able to provide mass infall rates above $0.1\,{\rm M}_\odot{\rm yr}^{-1}$, which are necessary for the formation of SMBH seeds. As we have seen before, the self-shielding method influences the value of $J_{\rm crit}$, but if we now compare the regimes below and above $J_{\rm crit}$ separately for both approaches, we do not see any clear systematic difference that is induced by the choice of the self-shielding method.

Assuming that approximately one Jeans mass per free fall time falls towards the central region, the mass accretion rate is given by

$$\dot{M} = \frac{c_s^3}{G} \propto T^{3/2}. \tag{2.19}$$

Consequently, the accretion rate is higher for hotter gas, such as in the isothermally collapsing cloud that retains temperatures of $\sim 10^4\,{\rm K}$ during the collapse. Moreover, the Jeans mass decreases with the temperature and the gas is more susceptible to fragmentation in the case of efficient H_2 cooling (see e.g. Clark et al. 2011a,b). Once cooling becomes efficient and the gas temperature falls, the cooling time becomes shorter than the local freefall time and the gas can locally contract, before the cloud globally collapses. The resulting clumpy structure can be seen in Figure 2.3, whereas the spatial structure of the isothermal collapse is illustrated in Figure 2.13. This smooth, almost spherically symmetric structure without signs of fragmentation enables higher gas infall rates, which leads to the formation of a SMS which then collapses to a SMBH seed. However, one should keep in mind that the different mass infall rates for scenarios below and above $J_{\rm crit}$ are just a trend and also the threshold of $0.1\,{\rm M}_\odot{\rm yr}^{-1}$ is only a rough estimator with other proposed values between 0.01 and $1\,{\rm M}_\odot{\rm yr}^{-1}$ (Begelman 2010; Hosokawa et al. 2012a; Schleicher et al. 2013; Ferrara et al. 2014). We find $M_{\rm in} < 0.1\,{\rm M}_\odot{\rm yr}^{-1}$ also for the isothermal collapse and vice versa. A value of $0.6\,{\rm M}_\odot{\rm yr}^{-1}$ seems to discriminate between the two collapse regimes:

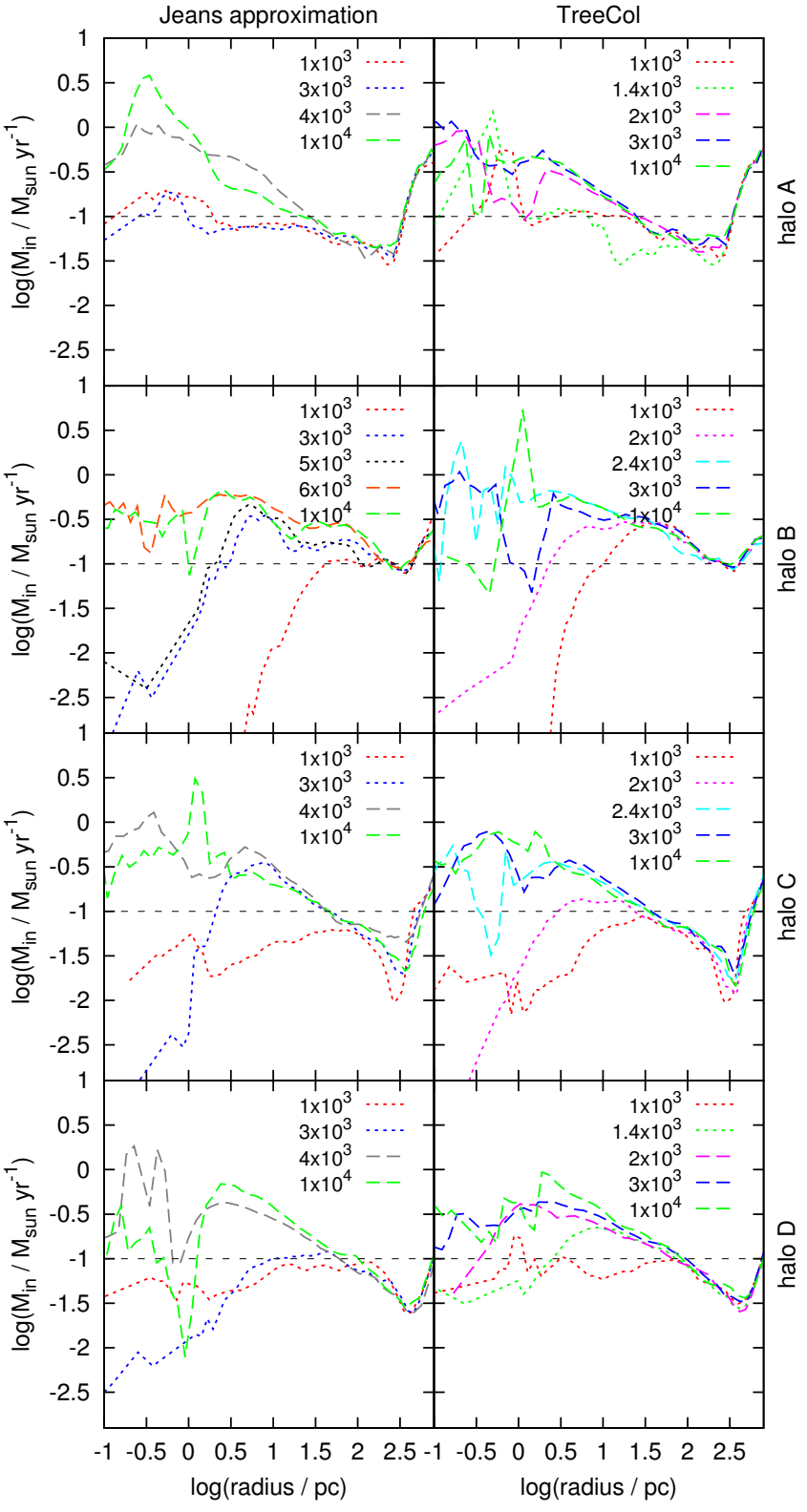


Figure 2.12: Radial profiles of the mass infall rates averaged over the last $\sim 10^5$ yr of the collapse. Simulations with $J_{\rm LW} < J_{\rm crit}$ are shown with short-dashed lines and those with $J_{\rm LW} > J_{\rm crit}$ are shown with long-dashed lines. The black line at $M_{\rm in} = 0.1\,{\rm M}_\odot{\rm yr}^{-1}$ represents the theoretical threshold above which the formation of a SMBH seed is possible. Generally, the mass infall rates are higher in the case of an isothermal collapse and especially in the central regions, the mass infall rates fall significantly below $0.1\,{\rm M}_\odot{\rm yr}^{-1}$ for $J_{\rm LW} < J_{\rm crit}$.

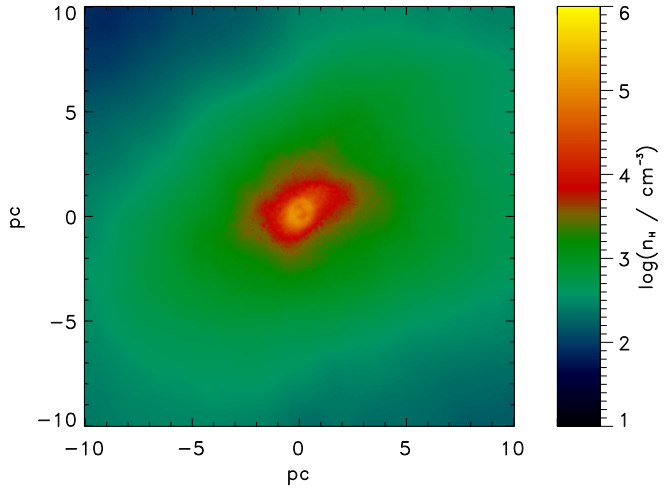


Figure 2.13: Map of the average number density of hydrogen nuclei along the line of sight for halo C with the TREECOL approach at the moment of collapse ($z \sim 15.1$). The background flux is $J_{\rm LW} = 3 \times 10^3 > J_{\rm crit}$ and in contrast to Figure 2.3, we can see no sign of gas fragmentation but a rather smooth, approximately spherically symmetric accretion towards the centre.

all isothermally collapsing clouds yield accretion rates above this value, but only one out of eight clouds with $J_{\rm LW} < J_{\rm crit}$ reaches this accretion rate. The high accretion rates in this one halo (top right panel in Fig. 2.12) are probably due to the radial binning, which artificially increases the mass accretion rate if a dense clump is present in this bin.

Latif & Volonteri (2015) study the infall rates in atomic cooling halos in greater detail. As in our study, they find that it is not always necessary to completely suppress H_2 formation to obtain sufficiently large infall rates to form a SMS. Moreover, they detect a rotationally supported structure in the central parsec, but this rotational support does not halt the collapse and still enables infall rates of $\sim 0.1\,{\rm M}_\odot{\rm yr}^{-1}$. For a more detailed discussion of this topic, we refer the interested reader to Latif & Volonteri (2015).

2.4 Caveats

The attempt to find one universal value of $J_{\rm crit}$ is rather artificial, because the relevant physical processes are too complex to be summarised in one simple number that decides whether we form a SMBH seed or not. Agarwal et al. (2016b) study the value of $J_{\rm crit}$ using one-zone models with a more realistic spectral energy distribution for the external radiation and show that $J_{\rm crit}$ is not one fixed value, but rather a range spanning more than three orders of magnitude. Moreover, we can also have sufficiently high mass infall rates to form a SMS even for $J_{\rm LW} < J_{\rm crit}$. Although the concept of one universal threshold $J_{\rm crit}$ is questionable in the formation scenario of SMBH seeds, it is a convenient quantification to study the influence of different physical processes on the direct collapse scenario.

2.4.1 Stellar spectrum

One should keep in mind that we use a T5 spectrum to study the effect of H_2 self-shielding (section 2.2.3). This is an accurate approximation for stars with a characteristic mass of $\sim 100\,{\rm M}_\odot$, but later stellar populations are believed to have softer stellar spectra and might therefore yield a lower value of $J_{\rm crit}$ (Sugimura et al. 2014; Agarwal & Khochfar 2015; Agarwal et al. 2016b). Latif et al. (2015) show that the

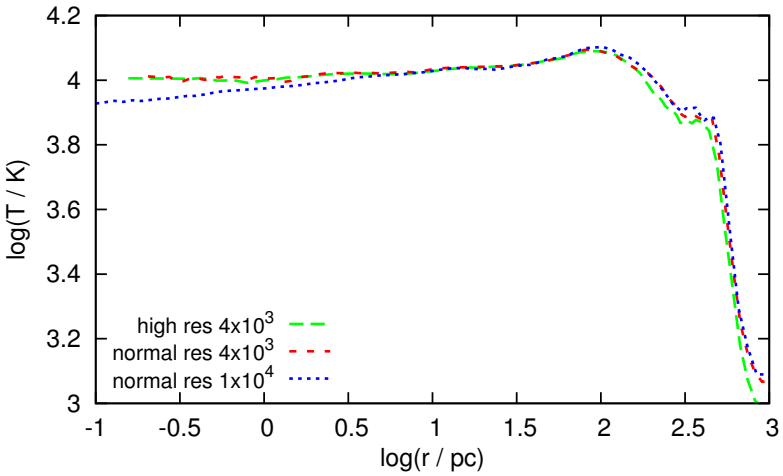


Figure 2.14: Temperature profiles for halo C using the Jeans approximation, where the green long-dashed curve represents an additional run with a two times higher spatial resolution. This should be compared to the red short-dashed curve, which represents a run with the normal resolution and the same background flux. The blue dotted curve shows the profile for the same halo with normal resolution but a higher LW flux. All curves show an isothermal collapse, but the differences induced by a different background radiation are larger than those induced by a higher resolution.

value of $J_{\rm crit}$ only weakly depends on the adopted radiation spectra in the range $2\times10^4\,{\rm K} \leq T_{\rm rad} \leq 10^5\,{\rm K}$ and only stellar populations older than $\sim 1\,{\rm Gyr}$ have lower effective spectral temperatures, which are not relevant for the formation of the first massive BH seeds

2.4.2 Resolution

To test the resolution, we also perform one simulation with a better mass resolution. In our standard approach, we resolve the local Jeans mass by 66 Voronoi cells, which corresponds to a spatial resolution of $\lesssim 0.1$pc. For the high-resolution run, we increase the mass resolution by a factor of eight, which doubles the spatial resolution. The comparison can be seen in Figure 2.14. The run with a higher spatial resolution yields the same temperature profile as the run with the normal resolution. Hence, we can conclude that our normal resolution is high enough to properly resolve the collapse, and our results are not sensitive to the numerical resolution.

2.4.3 Photochemistry

The Wolcott-Green et al. (2011) self-shielding function that we use here is intended for use when the H_2 is rotationally hot and not only the lowest rotational levels are populated. This is a reasonable approximation for densities $n \simeq 10^3$–10^4 cm^{-3}, but at lower densities we would expect most of the H_2 molecules to be in the $J = 0$ or $J = 1$ levels, and in this regime the Wolcott-Green et al. (2011) self-shielding function will underestimate the effectiveness of H_2 self-shielding. In addition, our current treatment of the shielding of H_2 does not account for absorption by the Lyman series lines of atomic hydrogen (Haiman et al. 1997; Schauer et al. 2015). Wolcott-Green & Haiman (2011) show that shielding of H_2 by atomic hydrogen becomes important for column densities of $N_{\rm H} > 10^{23}{\rm cm}^{-2}$. Including this effect may yield larger values for $J_{\rm crit}$ and may hence exacerbate the difference between the Jeans approximation results and the results derived using TREECOL.

We also do not include the photodetachment of H^- by trapped Lymanα photons, which lowers $J_{\rm crit}$ by up to a factor of a few (Johnson & Dijkstra 2017). This process has the strongest impact on the critical flux for the relatively high background radiation temperatures, as we have assumed in our model (T5) and taking it into account would further lower the value of $J_{\rm crit}$.

2.5 Conclusions

We have implemented a new method for the determination of H_2 column densities in the 3D moving mesh code AREPO and used it to study the effect of an improved treatment of H_2 self-shielding on the 'direct collapse' scenario. In a comparison to the previously used Jeans approximation, we find that the effective column densities are generally smaller with our new method and the necessary LW background flux to suppress efficient H_2 cooling is lower by a factor of about 2. More precisely, we find $J_{\rm crit} \simeq 2000$ with our new approach compared to $J_{\rm crit} \simeq 4000$ with the Jeans approximation. The main reason for this difference is the large directional dependence of the self-shielding factor that cannot be captured with one-dimensional methods. Because the detailed morphological and kinematic structure of the cloud matters a lot for the determination of the effective column density, it is also not possible to find a simple correction factor that might reproduce the results based on TREECOL.

Following Inayoshi & Tanaka (2015a), the density of possible direct collapse black hole formation sites scales with $n_{\rm DCBH} \propto J_{\rm crit}^{-5}$ for $J_{\rm LW} > 10^3$. Consequently, the factor of two, by which $J_{\rm crit}$ is lower with our new self-shielding approach, leads to a number density of direct collapse black holes in the early Universe that is about 32 times higher than previously expected. Although the number of expected direct collapse black holes is significantly higher with our new method, the value of $J_{\rm crit} \simeq 2000$ is still too high to explain the number density of SMBH at redshift $z \simeq 6$ of $n_{\rm SMBH} = 10^{-9}\,{\rm Mpc}^{-3}$ (comoving units) only by the isothermal direct collapse scenario. Even under optimistic assumptions, $J_{\rm crit}$ has to be smaller than 1000 to explain the observed number density (see e.g. Dijkstra et al. 2014; Inayoshi & Tanaka 2015a). In Section 5.3.3 we discuss in detail the connection between $J_{\rm crit}$ and the expected number density of direct collapse BHs.

3 Gravitational waves from the remnants of the first stars

Another possible pathway to form the seeds of the first SMBHs are the remnants of the first stars that we could indirectly detect via the mergers of their compact remnants. This chapter is based on the published paper "Gravitational waves from the remnants of the first stars" (Hartwig et al. 2016a). Our catalogues of Pop III binaries are publicly available here: www2.iap.fr/users/volonter/downloads

3.1 Introduction

The first detection of GWs on September 14, 2015 has opened a completely new window to investigate astrophysical processes and phenomena, which are otherwise invisible to observations in the electromagnetic spectrum (but see Loeb 2016; Perna et al. 2016). We had been completely blind to gravitational radiation before and we are now entering a new era of astronomy, which will reveal completely new domains of research and revolutionise astrophysics, comparable to radio astronomy in the first half and X-ray astronomy in the second half of the $20^{\rm th}$ century.

This first event GW150914 was the inspiral and merger of two BHs with masses $M_1 = 36^{+5}_{-4}\,{\rm M}_\odot$ and $M_2 = 29^{+4}_{-4}\,{\rm M}_\odot$ at redshift $z = 0.09^{+0.03}_{-0.04}$ (Abbott et al. 2016b). It was detected by the Advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO) with a false alert probability of $< 2 \times 10^{-7}$ (Abbott et al. 2016b). In Figure 3.1 we show the observed signals in the two gravitational wave detectors. The local merger rate density for GW150914-like systems is $3.4^{+8.8}_{-2.8}\,{\rm yr}^{-1}{\rm Gpc}^{-3}$ (Abbott et al. 2016), while the rate for all possible stellar BHs is $12 - 213\,{\rm yr}^{-1}{\rm Gpc}^{-3}$ (Abbott et al. 2017).

The mass of the merging BHs is much higher than all the previous estimates of stellar-mass BHs based on X-ray binaries. Before GW150914 the existence of BHs with mass above $\sim 10\,{\rm M}_\odot$ (and below the supermassive black hole range, $\gtrsim 10^5\,{\rm M}_\odot$) was never proven conclusively. Low metallicity stars are of special interest, because they loose less metals due to stellar winds during their lifetime, allowing for more massive remnant with respect to high-metallicity stars with the same initial mass (see Figure 3.2). Zero-metallicity stars are therefore of primary interest, because they are believed to be more massive than present-day stars, thus yielding more massive remnants (for reviews see Bromm 2013; Glover 2013b; Greif 2015). A fraction of GW detections can originate from these primordial stars. This was suggested several years ago (Bond & Carr 1984; Schneider et al. 2000, 2001; Belczynski et al. 2004a; Kulczycki et al. 2006; Sesana et al. 2009; Kinugawa et al. 2014; Dominik et al. 2013, 2015) and the detection of GW150914 by LIGO spurred renewed interest (Hartwig et al. 2016a; Kinugawa et al. 2016a; Belczynski et al. 2016c; Kinugawa et al. 2016b; Inayoshi et al. 2017).

In this chapter, we apply a semi-analytical approach to determine the rate density and the detection rate of mergers for aLIGO that originate from the first stars. We specifically investigate the dependence on the primordial initial mass function (IMF), and derive the probability that a certain event is of Pop III origin.

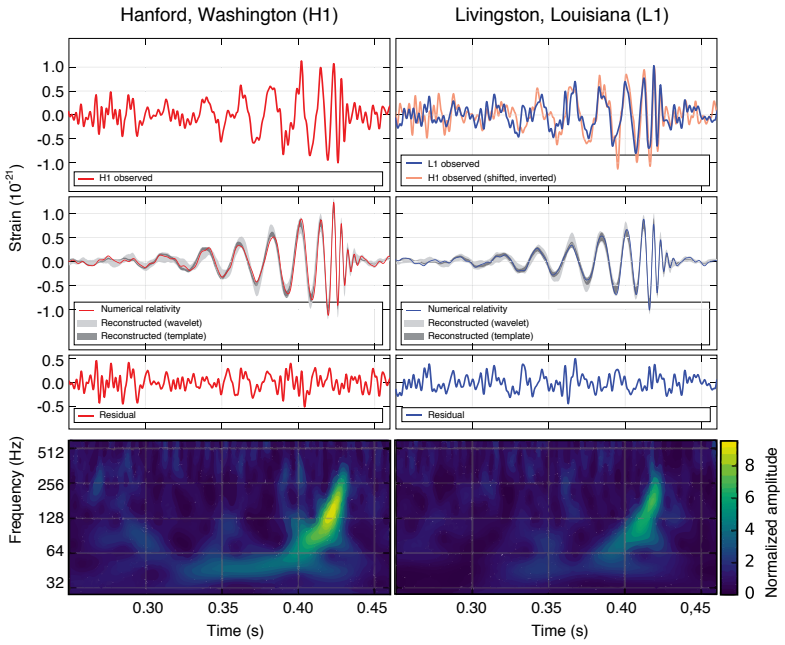


Figure 3.1: The gravitational-wave event GW150914 observed by the LIGO Hanford (left column) and Livingston (right column) detectors. Times are shown relative to September 14, 2015 at 09:50:45 UTC. Top row: strain, shifted for Livingston by the time delay to Hanford. Second row: Gravitational-wave strain projected onto each detector in the 35–350 Hz band. Third row: Residuals after subtracting the filtered numerical relativity waveform from the filtered detector time series. Bottom row: Time-frequency representation of the strain data, showing the signal frequency increasing over time. Adapted from Abbott et al. (2016).

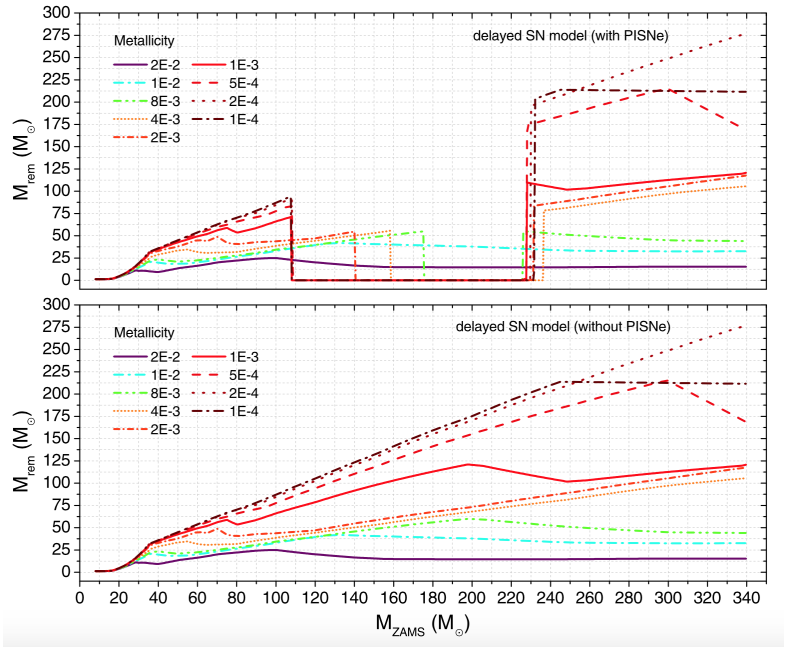


Figure 3.2: Mass of the compact remnant as a function of the initial mass of the star, for different stellar metallicities, with and without the inclusion of PISNe. A metallicity of 0.02 corresponds to solar in this nomenclature. Metallicity increases the mass loss due to stellar winds and hence decreases the remnant mass. This makes metal-poor stars favourable candidates as progenitors of the observed BH-BH mergers. Adapted from Spera et al. (2016), see also Spera & Mapelli (2017).

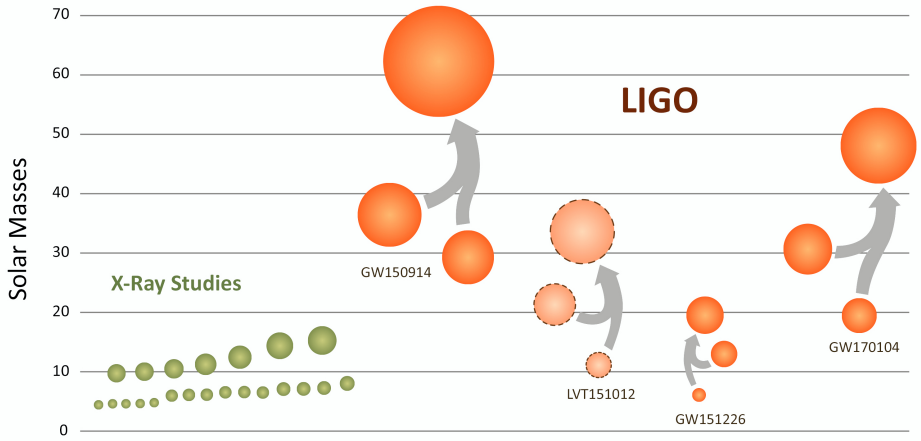


Figure 3.3: LIGO has discovered a new population of black holes with masses that are larger than what had been seen before with X-ray studies alone (green). The three confirmed detections by LIGO (GW150914, GW151226, GW170104), and one lower-confidence detection (LVT151012), point to a population of stellar-mass binary black holes that, once merged, are larger than $10\,\mathrm{M}_\odot$ — larger than what was known before. Image credit: LIGO/Caltech/MIT/Sonoma State (Aurore Simonnet).

This chapter is based on the consequences and interpretation of the first detection GW150914. Since then, there were two and a half more detections, which are illustrated in Figure 3.3. We do not explicitly discuss the other detections in this chapter, but most detected BHs were more massive than what most people previously expected to observe, which triggered new discussion about their possible formation mechanisms.

3.2 Methodology

3.2.1 Self-consistent Pop III star formation

We create a cosmologically representative sample of dark matter merger trees with the GALFORM code based on Parkinson et al. (2008). The merger trees start at $z_{\mathrm{max}} = 50$ and follow Pop III star formation down to $z = 6$, after which we do not expect significant Pop III star formation to occur. The formation of primordial stars is modelled self-consistently, taking into account radiative and chemical feedback. We briefly review the main aspects of the model; see Hartwig et al. (2015a, 2016b); Magg et al. (2016) and Section 4.3 for details.

To form Pop III stars, a halo has to be pristine and its virial temperature has to be high enough to allow for efficient cooling by H_2. We also check the time-scale of dynamical heating due to previous mergers and the photodissociation of H_2 by external Lyman-Werner radiation. Once a halo passes these four criteria, we assign individual Pop III stars to it, based on random sampling of a logarithmically flat initial mass function (IMF) in the mass range between $M_{\mathrm{min}} = 3\,\mathrm{M}_\odot$ and $M_{\mathrm{max}} = 300\,\mathrm{M}_\odot$, motivated by, e.g., Greif et al. (2011), Clark et al. (2011a), Dopcke et al. (2013). The IMF is still uncertain, and therefore, besides the fiducial model ($3 - 300\,\mathrm{M}_\odot$), we also consider a low mass ($1 - 100\,\mathrm{M}_\odot$) and a high mass ($10 - 1000\,\mathrm{M}_\odot$) IMF. A steeper slope of the IMF yields more low-mass and less massive stars and is to first order equivalent to a logarithmically flat IMF with a smaller characteristic mass, as pointed out by Hartwig et al. (2015a). We therefore only vary the range and not the slope of the IMF. The total stellar mass per Pop III-forming halo is set by the star formation efficiency. This parameter is

calibrated to reproduce the optical depth to Thomson scattering of $\tau = 0.066$ (Planck Collaboration et al. 2016), taking into account the contribution by later generations of stars, based on the global cosmic star formation history (Behroozi & Silk 2015). With this approach we populate minihalos with small multiples of primordial stars (see Fig. 9 in Hartwig et al. 2015a).

3.2.2 Binary sampling and evolution

We use the results of the most detailed study of Pop III binary systems to date by Stacy & Bromm (2013). They performed a cosmological simulation initialised at $z = 100$ within a 1.4 cMpc box. The first ten minihalos to form in their computational domain were resimulated with greater refinement to examine the gas collapse and subsequent formation of a Pop III multiple system within each minihalo. Employing sink particles with an accretion radius of 20 AU to represent individual stars, their dynamics and mass accretion was followed for a period of 5000 yr. From this study, we adopt a binary fraction of 36%, which translates into a $\sim$ 50% probability for a single star to have a binary companion (see also Stacy et al. 2016). This fraction might be too low due to the limited spatial resolution and other studies of Pop III star formation might allow larger binary fractions (Clark et al. 2011a; Greif et al. 2011; Smith et al. 2011), but all derived merger and detection rates scale linearly with this binary fraction. Our results can thus readily be rescaled accordingly.

The evolution of the binary system and consequently the nature (BH or NS), the masses, and the time of coalescence of the two compact objects depend mainly on the ZAMS characteristics, respectively the semi-major axis and eccentricity of their orbit and their masses. To illustrate the various effects that could play a role in stellar binary evolution, we illustrate the most important processes in Figure 3.4. This specific channel, evolution, and ZAMS properties was proposed by Belczynski et al. (2016a) as the most likely progenitor of GW150914.

For the pairing of the binaries in our model and the underlying distribution of mass ratios, we apply the "ordered pairing" advocated by Oh et al. (2015), since observations show that massive binaries favour members with similar masses. Hence, we order the primordial stars in one halo by descending mass, check probabilistically if they have a binary companion, and pair the most massive with the second most massive, the third most massive with the fourth most massive and so on. For the ZAMS eccentricity e_0 of each binary system we draw a random value from the thermal distribution $p(e)\,\mathrm{d}e \propto e\,\mathrm{d}e$ with $e_{\mathrm{min}} = 0.1$ and $e_{\mathrm{max}} = 1$ (Kroupa 1995; Dominik et al. 2012; Kinugawa et al. 2014). This distribution agrees qualitatively with that in Stacy & Bromm (2013), see Figure 3.5. The ZAMS semi-major axis a_0 is sampled from the distribution $p(x)\,\mathrm{d}x \propto x^{-1/2}\,\mathrm{d}x$ with $x = \log(a_0/\mathrm{R}_\odot)$, $a_{\mathrm{min}} = 50\,\mathrm{R}_\odot$, and $a_{\mathrm{max}} = 2 \times 10^6\,\mathrm{R}_\odot$. The shape and the lower limit are motivated by Sana et al. (2012) and de Mink & Belczynski (2015), whereas the upper limit is chosen in agreement with Stacy & Bromm (2013). We have verified that the specific choice of these limits does not significantly affect the final results.

Once we have identified the binaries and assigned their ZAMS quantities, we use the tabulated models for stellar binary evolution by de Mink & Belczynski (2015) to calculate the masses of the compact remnants and their delay time t_{del} until coalescence. The delay time is the sum of the time to the formation of the last compact object measured from the ZAMS, and the ensuing inspiral time, t_{insp}, of the two compact objects. The latter is the dominant time-scale for Pop III remnants. We chose their model "N-m2 A.002", which is the best fit to the properties of Pop III stars in terms of IMF, metallicity and evolutionary channels (see also Kinugawa et al. 2014; Belczynski et al. 2016a). It is the model with the lowest available metallicity (10% solar) and with a shallower IMF, which guarantees a better sampling at higher stellar masses. Moreover, model A is a better approximation for metal-free stars

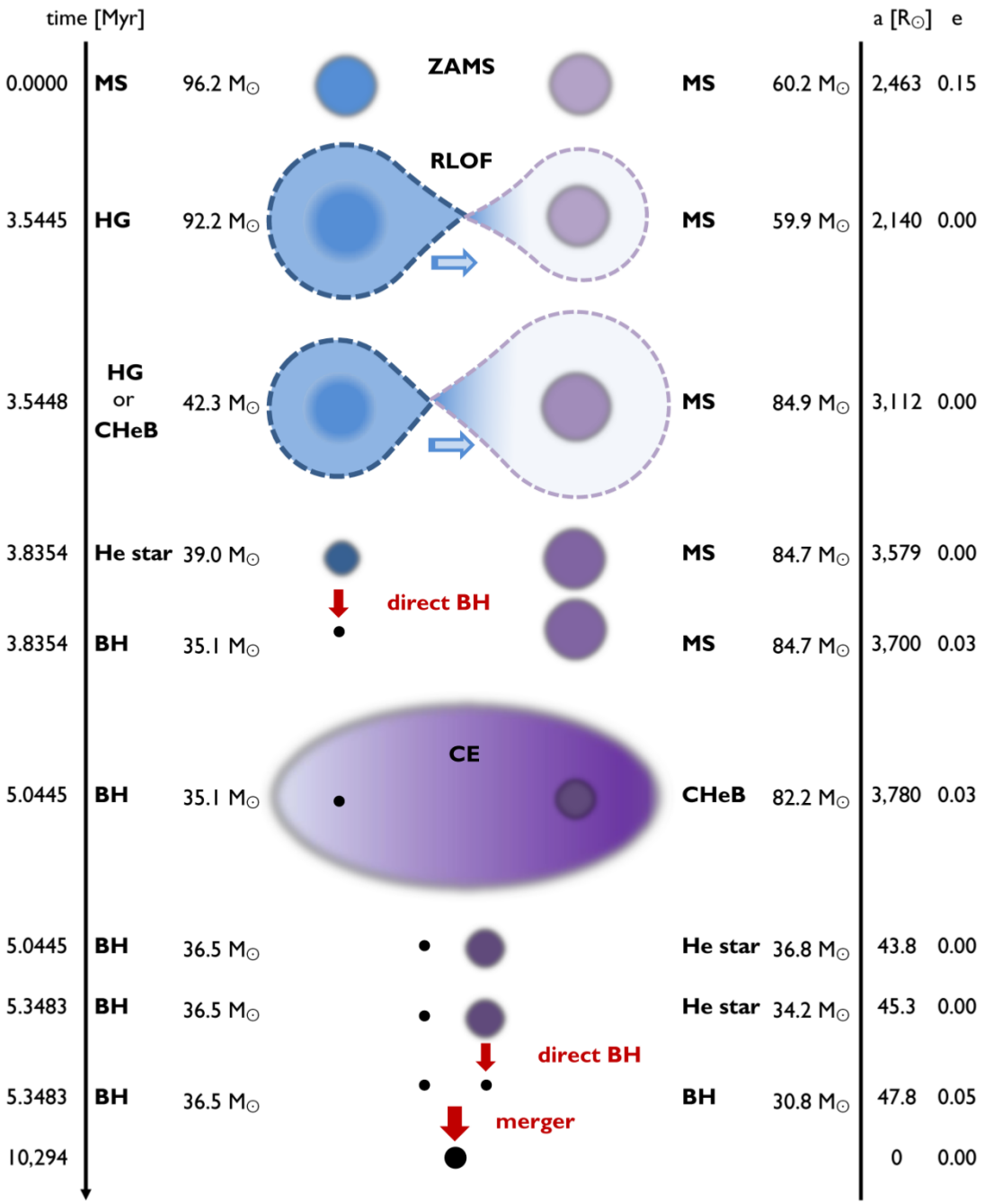


Figure 3.4: Example binary evolution leading to a BH-BH merger similar to GW150914. This example binary formed in a low metallicity environment of 3% solar. Roche lobe overflow (RLOF) can cause mass exchange between the binaries. The common envelope (CE) phase shrinks the separation of the binaries by more than two orders of magnitude, which is crucial for the final merger time, which scales with the separation to the fourth power. The final inspiral of the two BHs until merger is purely caused by the emission of GWs and takes about 10 Gyr in this example (compared to only about 5 Myr for the stellar evolution). Adapted form Belczynski et al. (2016a).

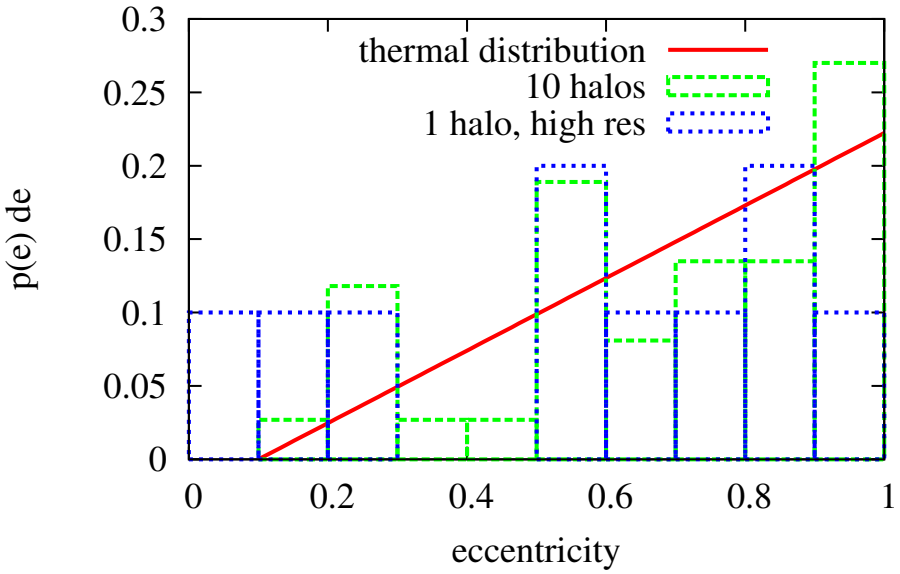


Figure 3.5: Probability distribution of the ZAMS eccentricity. Comparison between Stacy & Bromm (2013) and the thermal distribution, which are in rough agreement.

than the alternative model B, since they explicitly exclude all evolutionary channels where the common envelope is initiated by a Hertzsprung gap donor star in model B. In fact, Pop III stars at the end of their main sequence phase avoid the Hertzsprung gap, because their central temperature during the hydrogen-burning phase is already high enough to ignite helium (Kinugawa et al. 2014). Also Belczynski et al. (2016a) conservatively assume that stellar evolution at even lower metallicities proceeds in the same way as the evolution at $Z = 0.5\% \, Z_{\odot}$.

One should keep in mind the differences between the stellar binary evolution of metal-free and metal-enriched systems, which might lead to systematic errors (Kinugawa et al. 2014, 2016a; Belczynski et al. 2016c). In contrast to metal-enriched stars, Pop III stars lose only a small fraction of their mass due to stellar winds and therefore produce more massive remnants (Mapelli et al. 2013; Spera et al. 2016). Moreover, Pop III stars with a stellar mass of less than $\sim 50\,\mathrm{M}_{\odot}$ evolve as blue giants (not as red giants, like metal-enriched stars) and the resulting stable mass transfer makes the common envelope phase less likely.

The data is tabulated for stellar masses of the individual companions of up to $150\,\mathrm{M}_{\odot}$. For higher masses, we proceed in the following way. We ignore binaries with one star in the mass range $140\,\mathrm{M}_{\odot} \leq M_{*} \leq 260\,\mathrm{M}_{\odot}$, as we do not expect any compact remnants due to pair-instability supernova (PISN) explosions (Heger & Woosley 2002). For stars above $260\,\mathrm{M}_{\odot}$, we consider t_{del} characteristic of stars with $100\,\mathrm{M}_{\odot} \leq M_{*} \leq 140\,\mathrm{M}_{\odot}$, adopt the final black hole masses of primordial stars from Heger & Woosley (2002), and correct the tabulated inspiral time according to (Kinugawa et al. 2014)

$$t_{\mathrm{insp}} \propto {m_1}^{-1} {m_2}^{-1} (m_1 + m_2)^{-1}, \tag{3.1}$$

where m_1 and m_2 are the masses of the binary compact objects. This approach is justified because the tabulated t_{del} show negligible dependence on stellar mass for massive stars.

We have compared our approach for the binary stellar evolution to Kinugawa et al. (2014). After rescaling their results to our self-consistently calculated SFR of Pop III stars (see section 3.3) we find good

qualitative agreement of the intrinsic merger rate density of BHs at high redshift and also quantitative agreement within the error margins at low redshift, which is the more relevant region for possible detections. We hence conclude that despite the approximations and extrapolations, our model of binary stellar evolution yields reasonable results also for Pop III stars.

3.2.3 Detectability

Based on the cosmologically representative, self-consistent sampling of Pop III stars and the corresponding $t_{\rm del}$ of each binary, we determine the intrinsic merger rate density R. This represents the number of compact binary mergers per unit source time and per comoving volume, and is also referred to as the rest-frame merger rate density. To estimate the aLIGO detection rate, we calculate the single-detector signal-to-noise ratio (SNR) ρ for each merger via (Maggiore 2007; Finn & Chernoff 1993; Cutler & Flanagan 1994),

$$\rho^2 = 4 \int_0^\infty \frac{|h(f)|^2}{S_n(f)} {\rm d}f \,, \tag{3.2}$$

where $h(f)$ is the Fourier-domain (sky- and orientation-averaged) GW strain at the detector, and S_n is the noise power spectral density of a single aLIGO detector. We assume that an event is detectable if $\rho > 8$, as conventionally done in the LIGO literature (Abadie et al. 2010; Dominik et al. 2015; Belczynski et al. 2016b; de Mink & Mandel 2016). This translates to SNR larger than 12 for a three-detector network, e.g. the two aLIGOs and advanced Virgo. For the current aLIGO detectors, we use the O1 noise power spectral density (Abbott et al. 2015), whereas to assess detectability when the detectors are in their final design configurations we use the zero-detuning, high-power configuration of Abbott et al. (2009). For $h(f)$, we use either inspiral-only, restricted post-Newtonian waveforms (computing the Fourier transform with the stationary phase approximation, see e.g. Maggiore 2007), or inspiral-merger-ringdown PhenomA (non-spinning) waveforms (Ajith et al. 2008, 2009). We employ the former for BH-NS and NS-NS systems, with a cut-off at the frequency of the innermost stable circular orbit (ISCO; note that the ISCO frequency also corresponds approximately to the merger frequency of NS-NS systems). For BH-BH systems, particularly at high masses, the merger-ringdown contains considerable SNR, hence we use PhenomA waveforms. We then calculate the detection rate as (Haehnelt 1994)

$$\frac{dn}{dt} = 4\pi c \int_{\rho>8} {\rm d}z {\rm d}m_1 {\rm d}m_2 \frac{{\rm d}^2 R}{{\rm d}m_1 {\rm d}m_2} \frac{{\rm d}t}{{\rm d}z} \left(\frac{d_L}{1+z} \right)^2 , \tag{3.3}$$

where the luminosity distance d_L and the derivative of the look-back time with respect to z, ${\rm d}t/{\rm d}z$, are computed with a ΛCDM cosmology, and the integral is restricted to detectable events only ($\rho > 8$).

Finally, we characterise the stochastic GW background of our binary population by the energy density spectrum (see e.g. Phinney 2001; Rosado 2011; Abbott et al. 2016a):

$$\Omega_{\rm GW}(f) = \frac{f}{\rho_c c^2} \int {\rm d}z {\rm d}m_1 {\rm d}m_2 \frac{{\rm d}^2 R}{{\rm d}m_1 {\rm d}m_2} \frac{{\rm d}t}{{\rm d}z} \frac{{\rm d}E_s}{{\rm d}f_s}, \tag{3.4}$$

where ρ_c is the critical density, and ${\rm d}E_s/{\rm d}f_s \propto (f|h(f)|)^2$ is the spectral energy density of a binary, computed at the source-frame frequency $f_s = f(1+z)$ (f being the frequency at the detector). Our model's prediction for $\Omega_{\rm GW}$ should be compared with the 1σ power-law integrated curves (Thrane & Romano 2013) that are given by Abbott et al. (2016a) for the aLIGO/advanced Virgo network in the observing runs O1 (2015-16) and O5 (2020-2022), which represent the network's sensitivity to standard cross-correlation searches (Allen & Romano 1999) of power-law backgrounds.

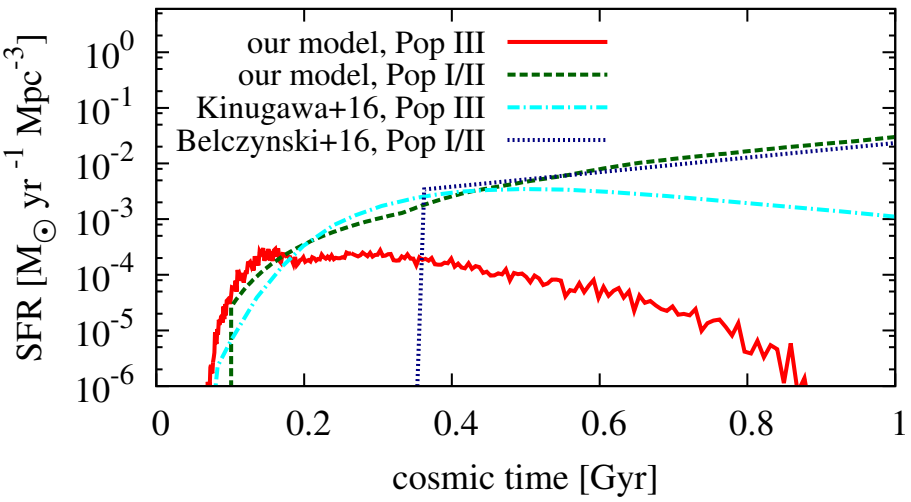


Figure 3.6: Comparison of our SFR for the fiducial IMF with the models used by Kinugawa et al. (2016a) and Belczynski et al. (2016a). Our Pop I/II SFR, which is adopted from Behroozi & Silk (2015), is in good agreement with the corresponding SFR by Belczynski et al. (2016a). For the Pop III stars, our self-consistent modelling yields a peak SFR that is about an order of magnitude lower than the value by Kinugawa et al. (2016a).

3.3 Results

In Figure 3.6, we compare our star formation rate (SFR) to other models. Our self-consistent Pop III SFR, with a peak value of $\mathrm{SFR}_{\mathrm{max}} = 2 \times 10^{-4}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$, is in agreement with Visbal et al. (2015), who show that it cannot exceed a few times $10^{-4}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$, without violating the constraints set by Planck Collaboration et al. (2016). Kinugawa et al. (2014) assume a SFR with a peak value of $\mathrm{SFR}_{\mathrm{max}} = 3 \times 10^{-3}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$, which is about an order of magnitude higher than our result. The SFR is roughly the same for all our Pop III IMFs, because we calibrate each model to match τ. Our SFR decreases more rapidly towards lower redshift than e.g. the SFR in Johnson et al. (2013a), due to the different treatment of reionisation and metal enrichment, which prevents all Pop III star formation in our model for $z < 6$.

The intrinsic merger rate density of compact objects can be seen in Figure 3.7. To compare with Kinugawa et al. (2016a) we rescale their SFR to our peak value, and additionally run a case with their IMF (10-100). Our values are about an order of magnitude lower than the rescaled prediction for Pop III stars by Kinugawa et al. (2016a) in the regime relevant for GW detections, i.e. mergers occurring at late cosmic times. The difference is to be ascribed to the different binary evolution models. Given that Kinugawa et al. (2014) study explicitly Pop III star binary evolution, and we extrapolate models at higher metallicity, we conclude that our estimates are a lower limit to the GW detections of Pop III binaries by up to an order of magnitude. We stress, however, that our Pop III SFR is calculated self-consistently and reproduces the optical depth constraint set by Planck Collaboration et al. (2016), in contrast to Kinugawa et al. (2014, 2016a).

Comparing models with different IMFs, we conclude that the number of expected mergers is dominated by M_{max}. This is because the remnants at low stellar masses are mostly NSs, which constitute only a small contribution to the overall merger rate. The merger rate density generally decreases with higher M_{max}, which is mainly related to the different number of binary systems that can form per minihalo for the different IMFs. At face value, Pop III stars do not yield a major contribution to the total merger rate density, but we recall that our estimates are likely lower limits. Crucially, due to their higher masses,

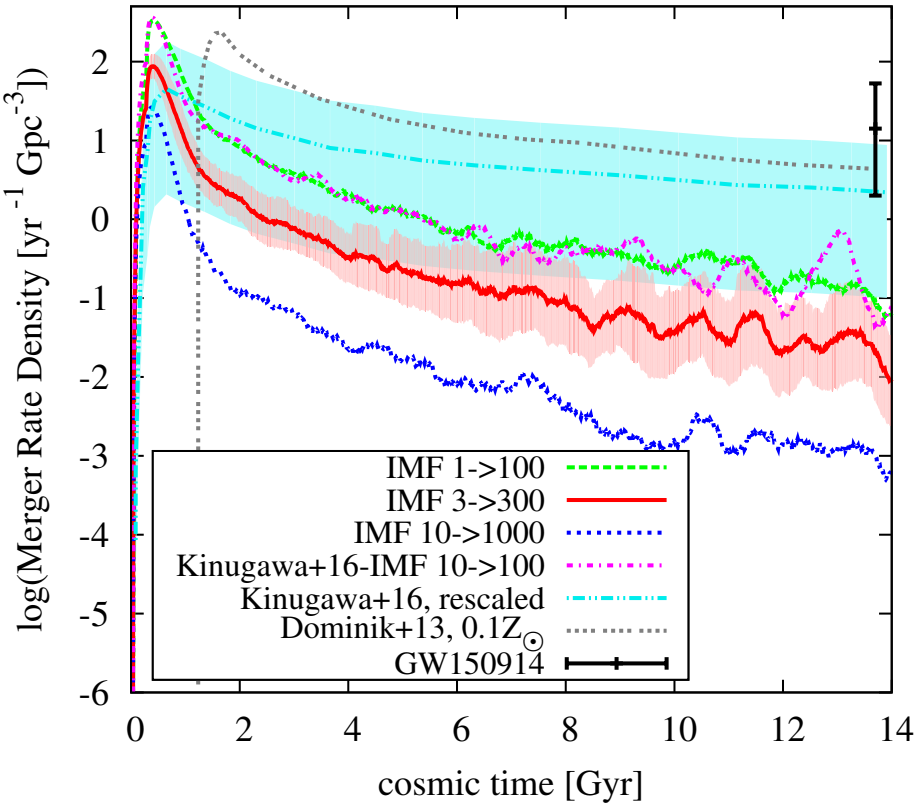


Figure 3.7: Intrinsic merger rate densities (BH-BH, BH-NS and NS-NS) for our models and comparison to the literature. For clarity, we show the statistical variance as shaded region only for our fiducial model. We plot the model by Kinugawa et al. (2016a) for Pop III remnants, (their figure 22, model 'under100, optimistic core merger'), rescaled to our $\mathrm{SFR}_{\mathrm{max}}$ with the corresponding systematic uncertainty. The model by Dominik et al. (2013) determines the merger rate density for stars at 10% solar metallicity (their figure 4). We also show the expected value at $z = 0$ from the GW150914 detection (Abbott et al. 2016).

Pop III BH-BH mergers have strong GW signals, which boost their detection probability over lower-mass BHs formed from later stellar generations.

Another essential question is whether we are able to discriminate mergers of Pop III origin. The most massive remnant BHs for binaries at $0.1\,Z_\odot$ have a mass of $\sim 42\,\mathrm{M}_\odot$ (de Mink & Belczynski 2015) and Mapelli et al. (2017) show that BH binaries merging within the aLIGO horizon have a chirp mass of $\lesssim 35\,\mathrm{M}_\odot$ for a minimal metallicity of 1% solar. All BHs with higher masses must be of Pop III origin (though note that binary BHs with M_{tot} up to $\sim 160 M_\odot$ may form in globular clusters, Belczynski et al. 2014, 2016b). The minimal mass for Pop III remnant BHs above the PISN gap is $M_{\mathrm{PISN}} \approx 200\,\mathrm{M}_\odot$. Since aLIGO will measure the source-frame total mass for $M_{\mathrm{tot}} \gtrsim M_{\mathrm{PISN}}$ within a 20% uncertainty at the 2-σ level (Graff et al. 2015, see also Veitch et al. 2015; Haster et al. 2016), we can use M_{PISN} as the threshold for the *unambiguous detection* of a Pop III remnant BH. For IMFs extending above $300\,\mathrm{M}_\odot$ about one BH-BH merger per decade can be unambiguously attributed to a Pop III binary (Table 3.1). As discussed above, this is plausibly a lower limit because of the too efficient mass loss and hence the underestimated remnant masses in our binary evolution model. However, BH-BH mergers in this mass range could also originate from runaway collapse in young stellar clusters, which are expected to have detection rates with aLIGO of the same order of magnitude as Pop III remnant BHs (Mapelli et al. 2010) or from stellar BHs that have accreted gas during their evolution.

IMF	BH-BH	BH-NS	NS-NS	$m_1 > M_{\rm PISN}$
1-100	5.3	1.4×10^{-2}	7.2×10^{-3}	0
3-300	0.48	2.1×10^{-3}	8.1×10^{-4}	0.011
10-1000	0.12	2.4×10^{-4}	1.1×10^{-5}	0.089

Table 3.1: Detection rates in events per year for aLIGO at final design sensitivity. Assuming a log-normal distribution, we find a statistical scatter (1σ) between different independent realisations of $0.04 - 0.43$ dex, depending on the IMF and type of the merger. For the fiducial IMF (3-300) we expect about one Pop III binary black hole every two years and for the lower mass IMF (1-100) even up to 5 detections per year. The probability to detect a merger that can uniquely be identified as being of Pop III origin ($m_1 > M_{\rm PISN}$) is highest for the high mass IMF (10-1000) with about one detection per decade. Inversely, the strong dependence of the detection rate on the IMF can be used to infer the upper mass limit for Pop III stars.

To distinguish the contributions to the detection rate by different generations of stars (also below $M_{\rm PISN}$) we show the specific detection rates as a function of $M_{\rm tot} = m_1 + m_2$ in Figure 3.8. We compare

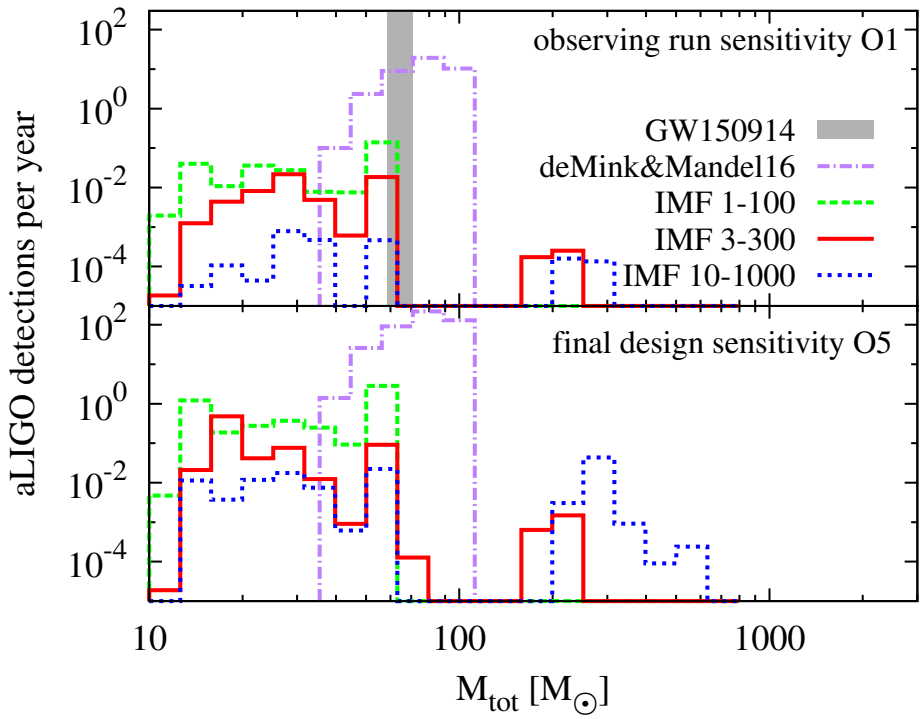


Figure 3.8: Expected number of BH-BH merger detections per year as a function of the total binary mass for the current aLIGO sensitivity (top) and final design sensitivity (bottom). The mass range of GW150914 is indicated by the grey area. With sufficient detections around $M_{\rm tot} \approx 300\,{\rm M}_\odot$ we could discriminate different Pop III IMFs based on their GW fingerprint.

our three different IMFs with a model by de Mink & Mandel (2016). They determine the detection rate of Pop I/II stars for the chemically homogeneous evolutionary channel for binary black hole mergers, which dominates at $30\,{\rm M}_\odot \lesssim M_{\rm tot} \lesssim 100\,{\rm M}_\odot$. For a given $M_{\rm tot}$ the histogram enables to determine the probability that this event originates from the first stars. For GW150914, this probability is $\sim 1\%$ (see Belczynski et al. 2016a; Woosley 2016; Schneider et al. 2017, for different approaches). For detections around $\sim 300\,{\rm M}_\odot$, in addition to unambiguously establishing a Pop III origin, one can even distinguish different Pop III IMFs by their number of detections over the next decades.

We conclude by discussing the contribution from Pop III binaries to the stochastic GW background, shown in Figure 3.9. In our models, the highest stochastic GW background is produced by BH-BH

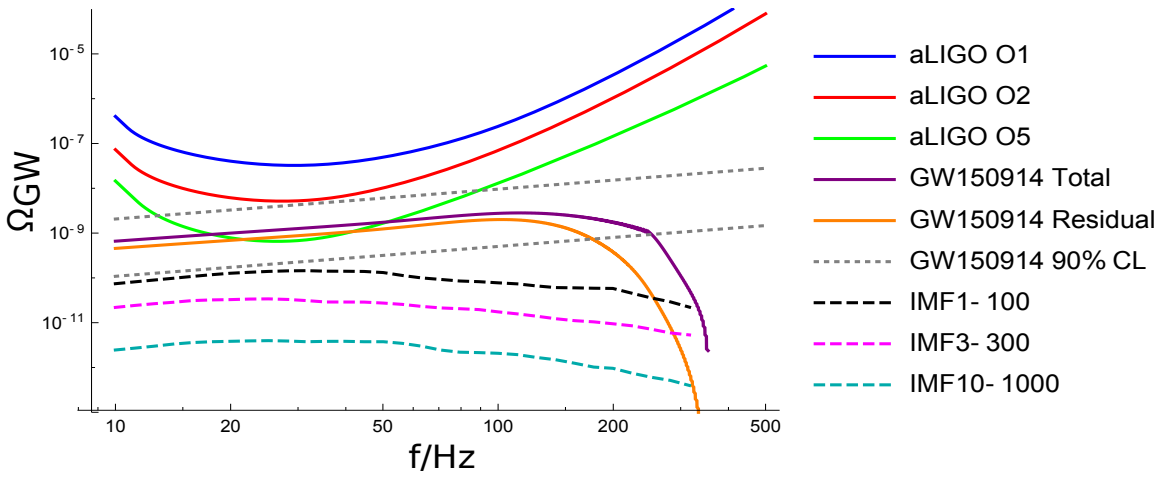


Figure 3.9: Stochastic GW background for our three different IMFs (long-dashed) compared to the inferred background from GW150914 and to the current and future detector sensitivity. The energy density spectra are shown for the total background (purple) and for the residual background (orange), excluding resolved sources, assuming final aLIGO and Virgo sensitivity. The grey dotted lines illustrate the approximate statistical uncertainty. The blue, red, and green curves show the 1σ sensitivity of the network expected for the two first observing runs O1 and O2, and for 2 years at the design sensitivity in O5 (Abbott et al. 2016a). The contribution from Pop III stars to the stochastic GW background will not be detectable with aLIGO.

mergers, and at $f \approx 25$ Hz (where the network is most sensitive), $\Omega_{\rm GW}(25\,{\rm Hz}) \approx 4\times10^{-12}$–$1.4\times10^{-10}$ (the inclusion or removal of resolved sources only makes a negligible difference). For comparison, the stochastic background at 25 Hz inferred from the observation of GW150914 is $\Omega_{\rm GW}(25{\rm Hz}) = 1.1^{+2.7}_{-0.9}\times 10^{-9}$ at 90% confidence level, which is higher than previously expected (Abbott et al. 2016a). The aLIGO/advanced Virgo network 1σ sensitivity (corresponding to SNR = 1) for two years of observation at design sensitivity (O5) is $\Omega_{\rm GW}(25{\rm Hz}) = 6.6\times10^{-10}$. The stochastic background produced by Pop III compact object mergers is therefore negligible at the relevant frequencies (Dvorkin et al. 2016, but see Schneider et al. 2000; Inayoshi et al. 2016).

3.4 Discussion

We have estimated the GW fingerprint of Pop III remnants on the aLIGO data-stream. GWs have the potential to directly detect the remnants of the first stars, and possibly even to constrain the Pop III IMF by observing BH-BH mergers with a total mass around $M_{\rm tot} \approx 300\,{\rm M}_\odot$. The latter is the key to ascertain the impact of the first stars, due to their radiative and supernova feedback, on early cosmic evolution and as progenitors of the first SMBHs. The new GW window ideally complements other probes, such as high-z searches for energetic supernovae with the *James Webb Space Telescope (JWST)*, or stellar archaeological surveys of extremely metal-poor stars (Bromm 2013; Frebel & Norris 2015). We have developed a model which includes Pop III star formation self-consistently, anchored, within the uncertainties, to the *Planck* optical depth to Thomson scattering. The main caveats in this study arise from the still uncertain Pop III binary properties and the corresponding stellar binary evolution.

We find a probability of $\sim$ 1% that GW150914 originates from Pop III stars, although this number may increase with improved future modelling. Crucially, the higher masses of the first stars boost their GW signal, and therefore their detection rate. Up to 5 detections per year with aLIGO at final design sensitivity originate from Pop III BH-BH mergers. Approximately once per decade we should detect

a BH-BH merger that can unambiguously be identified as a Pop III remnant. It is exciting that the imminent launch of the *JWST* nearly coincides with the first direct detection of GWs, thus providing us with two powerful, complementary windows into the early Universe.

GWs from BH binaries originating from Pop III stars may also be detectable by the Einstein Telescope or (in their early inspiral) by LISA (Amaro-Seoane & Santamaría 2010; Sesana 2016), which would allow probing the physics of these systems with unprecedented accuracy.

Recently, Belczynski et al. (2016c) developed a new model of Pop III stellar binary evolution and calculated the merger and detection rates for their compact remnants binaries with aLIGO. Their independent model also predicts a detection rate of $6\,\mathrm{yr}^{-1}$ for aLIGO ($2\,\mathrm{yr}^{-1}$ taking into account its duty cycle), confirming our results.

4 Exploring the nature of the Lyman-α emitter CR7

Sobral et al. (2015) presented observations of a galaxy at $z = 6.6$, which they interpreted as the first detection of the first stars. In this chapter, which is based on the published paper "Exploring the nature of the Lyman-α emitter CR7" (Hartwig et al. 2016b), we explore and compare different interpretations of this peculiar system, which might host an accreting BH.

4.1 Introduction

The first sources of light ushered the Universe out of the cosmic dark ages, thus initiating the long history of star and galaxy formation. In the ΛCDM paradigm of hierarchical structure formation, the first generation of stars were assembled in dark matter minihalos of a few times $10^5\,\mathrm{M}_\odot$. They later merged to form more massive halos of $\gtrsim 10^7\,\mathrm{M}_\odot$, which could be the birthplaces of the first galaxies (Bromm & Yoshida 2011). Understanding the formation of the first galaxies is of paramount importance because they constitute the basic building blocks of present-day galaxies.

The first stars influenced subsequent structure formation by chemical, radiative and mechanical feedback. They enriched the Universe with metals and led to the formation of second-generation Population II (Pop II) stars. As discussed in Section 1.3, various theoretical studies suggest that Pop III stars were on average more massive than Pop II stars because of the absence of efficient coolants in the pristine gas, raising the Jeans mass necessary for fragmentation (see the recent reviews by Bromm 2013; Glover 2013b; Greif 2015). Strong nebular emission lines, such as Lyman-α (Lyα) and those of HeII, are expected to be present in the gas ionised by Pop III stars (Tumlinson & Shull 2000; Bromm et al. 2001b; Oh et al. 2001; Schaerer 2002, 2003). The prime targets to detect these nebular emission lines are the first galaxies, which are expected to host both Pop III and Pop II stars.

Large ground based telescopes and the Hubble Space Telescope (HST) have opened a new window on high redshifts. Lyα has emerged as a powerful probe to detect distant galaxies at $z > 5$. The search for high redshift galaxies has intensified over the past few years and candidate galaxies have now been observed between $z = 8 - 11$ (Bouwens et al. 2011; Ellis et al. 2013; McLeod et al. 2015, 2016; Oesch et al. 2016). Numerous galaxies have been detected at $z > 6$ using both the strong Lyα emission and the Lyman break techniques (Ouchi et al. 2008, 2009, 2010; Vanzella et al. 2011; Ono et al. 2012; Finkelstein et al. 2013; Pentericci et al. 2014; Oesch et al. 2015; Matthee et al. 2015). Recently, Oesch et al. (2016) even claim the observation of a galaxy at $z = 11.1$. Although such galaxies have been detected at these redshifts, the presence of Pop III stars has not been observationally confirmed in any system. On the other hand, detections of high redshift quasars reveal the existence of supermassive black holes (SMBHs) of a few billion solar masses at $z > 6$ (Fan et al. 2006a; Mortlock et al. 2011; Venemans 2015; Wu et al. 2015). Various models have been proposed to explain the formation of SMBHs which include the growth of stellar-mass black holes (BHs) as well as so-called 'direct collapse' black holes (DCBHs) (Loeb &

Rasio 1994; Bromm & Loeb 2003a; Alvarez et al. 2009; Volonteri 2010; Haiman 2013; Latif et al. 2013b, 2015). Direct observational constraints on both seed BH masses and the Pop III initial mass function (IMF) are necessary to understand the formation of the first galaxies and high redshift quasars.

The recent discovery of strong HeII line emission from the Lyα emitter CR7 (COSMOS redshift 7) at $z = 6.6$ by Sobral et al. (2015) may constitute the first detection of either a Pop III star cluster or an accreting seed BH. It is the most luminous Lyα emitter at $z > 6$ and has very strong Lyα and HeII 1640 Å emission lines but no metal lines in the rest-frame UV. Sobral et al. (2015) suggest that CR7 can be explained by the composite spectra of normal metal-enriched stars and primordial stars. Deep HST imaging shows that CR7 is composed of three clumps, A, B, and C. The spectral energy distributions (SEDs) of clumps B and C are best fitted by an evolved stellar population while clump A has strong Lyα and HeII 1640 Å lines and can be explained either by a young primordial stellar population or an accreting BH. Sobral et al. (2015) find that Pop III stars with an age of a few Myr and a total stellar mass of $\sim 10^7\,M_\odot$ with a top-heavy IMF are required to explain such strong emission lines and photoionisation feedback (Xu et al. 2016a; Visbal et al. 2016, 2017) or Lyα radiation drag (Yajima & Khochfar 2017) can promote the formation of massive metal-free galaxies at lower redshift. The required stellar mass in Pop III stars is degenerate with the uncertain IMF and the required stellar mass might even be higher. The possibility that CR7 hosts a BH has also been discussed in Sobral et al. (2015); Dijkstra et al. (2016); Agarwal et al. (2016a, 2017a); Smith et al. (2016); Smidt et al. (2016).

Pallottini et al. (2015) have proposed that the strong Lyα and HeII 1640 Å line emission in CR7 can be explained by either a < 2 Myr old Pop III stellar cluster of $\sim 10^7\,M_\odot$ with a top-heavy IMF or an $\sim 10^5\,M_\odot$ BH formed at $z = 7.3$. Agarwal et al. (2016a), Dijkstra et al. (2016), and Smidt et al. (2016) have also shown that the observations can be explained by an $\sim 10^6 - 10^7\,M_\odot$ BH in the progenitor halos of CR7. But none of the studies has shown how such a large reservoir of metal-free gas can exist at $z = 6.6$ in clump A, which is required both for the formation of a young metal-free Pop III stellar cluster or a DCBH (although see Fumagalli et al. 2011). More observations are required to find how common are sources like CR7 at such redshifts and under what conditions can they form. In addition to these further observations, a self-consistent model, which takes into account both in-situ star formation and metal enrichment, is required to better understand the assembly history and nature of CR7.

In this chapter, we use a semi-analytical model to examine the nature of CR7 and its assembly history. The model includes recipes for Pop III and Pop II star formation as well as a self-consistent treatment of metal enrichment. We consider a range of halo masses, star formation efficiencies and IMFs with a sufficient number of realisations to obtain statistically sound results. We investigate the possibility of a massive BH forming from either a Pop III remnant or a DCBH. The results suggest that our model of Pop III star formation cannot reproduce the observed Lyα and HeII line luminosities and that a DCBH likely powers CR7.

4.2 Observational constraints

CR7, the brightest Lyα emitter at $z > 6$, was first classified as an unreliable high redshift candidate (Bowler et al. 2012) and as a brown dwarf candidate in the Milky Way (Ilbert et al. 2013). It was then found as a Lyα emitter candidate at $z = 6.6$ by Matthee et al. (2015) and spectroscopically confirmed by Sobral et al. (2015). This spectroscopic follow-up revealed strong Lyα and HeII 1640 Å lines with luminosities of $L_{\mathrm{Ly}\alpha} = (8.5 \pm 1.0) \times 10^{43}\,\mathrm{erg\,s^{-1}}$ and $L_{\mathrm{HeII}} = (2.0 \pm 0.8) \times 10^{43}\mathrm{erg\,s^{-1}}$. These observed luminosities imply a line ratio of HeII/Ly$\alpha \approx 0.23$ or a ratio of $Q(\mathrm{He^+})/Q(\mathrm{H}) \approx 0.42$ for ionising photon

$M_*(\,\mathrm{M}_\odot)$	$L_{\mathrm{Ly}\alpha}(\mathrm{erg\,s^{-1}})$	$L_{\mathrm{HeII}}(\mathrm{erg\,s^{-1}})$	$L_{\mathrm{HeII}}/L_{\mathrm{Ly}\alpha}$
1	2.1×10^{28}	4.8×10^{13}	2.4×10^{-15}
10	4.3×10^{36}	4.4×10^{30}	1.0×10^{-6}
100	1.1×10^{39}	1.8×10^{37}	1.6×10^{-2}
1000	1.6×10^{40}	3.7×10^{38}	2.3×10^{-2}

Table 4.1: Line luminosities for single, non-rotating Pop III stars, averaged over their lifetimes with no mass loss (Schaerer 2002). We extrapolate and interpolate linearly between the originally tabulated values. The model assumes an electron temperature, T_e, and density, n_e, of 30000 K and 100 $\mathrm{cm^{-3}}$, which are typical values for gas around the first stars. Decreasing the electron temperature to $T_e = 10000\,\mathrm{K}$ would increase the HeII line luminosity by at most $\sim 10\%$. The luminosities are a very steep function of mass, and the line ratio of HeII/Lyα is $\lesssim 0.02$ for the considered primordial stellar populations.

emission rates. This line ratio corresponds to a very hard spectrum with an effective temperature of $T_\mathrm{eff} > 10^5\,\mathrm{K}$, and even Pop III stars with a mass of $1000\,\mathrm{M}_\odot$ only have $Q(\mathrm{He^+})/Q(\mathrm{H}) \approx 0.1$ (Schaerer 2002).

One possible explanation for this extreme ratio is that a fraction of the Lyα photons was not observed due to trapping in the intergalactic medium (IGM) or dust absorption (Dijkstra et al. 2006, 2007; Dijkstra & Wyithe 2010; Zheng et al. 2010; Latif et al. 2011a,b; Smith et al. 2015a, 2016; Matthee et al. 2016). By post-processing cosmological hydrodynamic simulations with a multiwavelength radiative transfer scheme, Laursen et al. (2009) and Yajima et al. (2014) find that the escape fraction of Lyα photons can be as low as $\sim 10\%$, with the exact value depending on specific properties such as mass, star formation, dust content, or metallicity. They also show that Lyα photons do not escape isotropically, but the flux generally depends on the direction (Dijkstra et al. 2006). Consequently, the observed Lyα luminosity is only a lower limit to the actual luminosity and the line ratio of HeII/Lyα should hence be treated with caution.

In Table 4.1, we list line luminosities and HeII/Lyα ratios for $1-1000\,\mathrm{M}_\odot$ Pop III stars. The line ratios of a primordial stellar population in this mass range are much smaller than the one observed for CR7. The Lyα luminosity can be reduced by trapping or absorption, but we still have to account for the very high HeII luminosity. Put differently, we have to explain a system with $L_{\mathrm{Ly}\alpha} > 8.32\times10^{43}\,\mathrm{erg\,s^{-1}}$ and $L_{\mathrm{HeII}} = 1.95\times10^{43}\mathrm{erg\,s^{-1}}$. We will focus on the HeII luminosity, which is more difficult to explain in the context of early structure formation, because if a primordial stellar population fulfils the HeII constraint, it automatically satisfies the Lyα constraint.

The rate of ionising photons per stellar baryon for metal-free stars in the mass range $300-1000\,\mathrm{M}_\odot$ is almost constant (Bromm et al. 2001b). In this range, models suggest that Pop III stars have an effective temperature of $\sim 100\,\mathrm{kK}$ (Bromm et al. 2001b; Schaerer 2002) and we also see from Table 4.1 that the hardness of the spectrum, quantified by HeII/Lyα, does not change significantly in the mass range $100-1000\,\mathrm{M}_\odot$. Since the binary properties and the initial rotational velocities of Pop III stars are not well known (Stacy et al. 2012, 2016), we use stellar models of single, non-rotating stars. Rotation and binary evolution might however have an influence on the evolution of massive, metal-free stars and can change their stellar lifetime, spectrum, final fate, or the remnant masses (Ekström et al. 2008; de Mink et al. 2013). In Hartwig et al. (2015a) we have investigated the effect of rapidly rotating Pop III stars in our semi-analytical model and did not find a significant influence on the star formation or metal enrichment.

The full width at half maximum (FWHM) line widths of these lines are $(266\pm15)\,\mathrm{km\,s^{-1}}$ for Lyα and $(130\pm30)\,\mathrm{km\,s^{-1}}$ for HeII. Other strong radiative sources, such as Wolf-Rayet stars or AGN, generally produce broader lines with FWHM $\gtrsim 10^3\,\mathrm{km\,s^{-1}}$ (Brinchmann et al. 2008). Recently, Smidt et al.

(2016) demonstrate by post-processing cosmological simulations with a radiative transfer code that a massive BH of $\sim 10^7\,M_\odot$ accreting at 25% of the Eddington limit yields a line width for Lyα that is in good agreement with the observation and a line width of $\sim 210\,\mathrm{km\,s^{-1}}$ for HeII, which is slightly above the observed value. Smith et al. (2016) show with a 1D radiation-hydrodynamics simulation that the observed $160\mathrm{km\,s^{-1}}$ velocity offset between the Lyα and HeII line peaks is more likely to be produced by an accreting BH than by a stellar population with an effective temperature of $10^5\,\mathrm{K}$. Such a stellar population might ionise its environment too efficiently and can account neither for the velocity offset, nor for the spatial extension of $\sim 16\,\mathrm{kpc}$ of the Lyα emitting region. The UV slope of $\beta = -2.3 \pm 0.08$ cannot be used to distinguish between different models, since both a young metal-free stellar population and an accreting BH yield a blue UV slope in this frequency range (Dijkstra et al. 2016).

Another striking feature of CR7 is the absence of metal lines with upper limits of HeII/OIII]1663Å > 3 and HeII/CIII]1908Å > 2.5. This does not indicate the absence of metals, but that the HeII line dominates and hence normal stellar populations are excluded as the only explanation, since they would produce HeII/OIII]1663Å $\lesssim 0.3$, HeII/CIII]1908Å $\lesssim 0.3$ (Gutkin et al. 2016). These upper limits for the metal recombination lines also set the absolute metallicity of the gas (see Section 4.3.5). Due to the above constraints, the main source of ionising photons in CR7 has to be embedded in gas with a low metallicity.

The observed equivalent widths are $\mathrm{EW_{Ly\alpha}} > 230\,\text{Å}$ and $\mathrm{EW_{HeII}} = (80\pm20)\,\text{Å}$. The EW of Ly$\alpha$ is only a lower limit (since no UV continuum is detected) and has consequently no strong constraining power, because both a young stellar population and an accreting BH can yield an Lyα EW of $> 230\,\text{Å}$ (Shields & Ferland 1993; Malhotra & Rhoads 2002; Schaerer 2003). Also Dijkstra et al. (2016) do not use this observational constraint in their spectral fit since it does not yield any additional information. However, the large EW of HeII can be used to confine the age and ambient metallicity of a potential Pop III stellar population. A comparison to the very detailed stellar evolutionary synthesis models of Schaerer (2003) and Raiter et al. (2010) yields a metallicity of $< 10^{-7}$ and a very recent starburst of $< 1\,\mathrm{Myr}$ in order for the models to be consistent with the EW of HeII. Unfortunately, there is not such a sophisticated model with the required parameter dependences for the EWs of an AGN spectrum. Moreover, the determination of the HeII EW for an AGN spectrum would require the modelling of the underlying continuum. This is problematic because the AGN and stellar continuum overlap at these wavelengths and they are degenerate without knowing the respective luminosities (Stark et al. 2015). The detailed modelling of this problem is beyond the scope of the paper.

HST observations show that CR7 is composed of three clumps with a projected separation of $\sim 5\,\mathrm{kpc}$, with one clump hosting a young, blue stellar population and the other clumps hosting older, red populations (see Figure 4.1). This might be evidence for an ongoing merger, which makes it even more important to take the merger history of this system into account. Sobral et al. (2015) find that the best-fitting SED model is a combination of an older $1.6 \times 10^{10}\,M_\odot$, $0.2\,Z_\odot$ population with an age of 360 Myr and a metal-free population with a top-heavy IMF, a total stellar mass of $\sim 10^7\,M_\odot$ and an age of a few Myr (see Figure 4.2). However, a radiation source with a lifetime of at least $10 - 100\,\mathrm{Myr}$ is required to account for the spatial extent of the Lyα emitting region (Smith et al. 2016), which favours an accreting BH over a young, metal-free stellar population.

4.3 Methodology

We use a semi-analytical model that is based on the work of Hartwig et al. (2015a). In this section, we present the model and introduce the techniques we use to explore the nature of CR7.

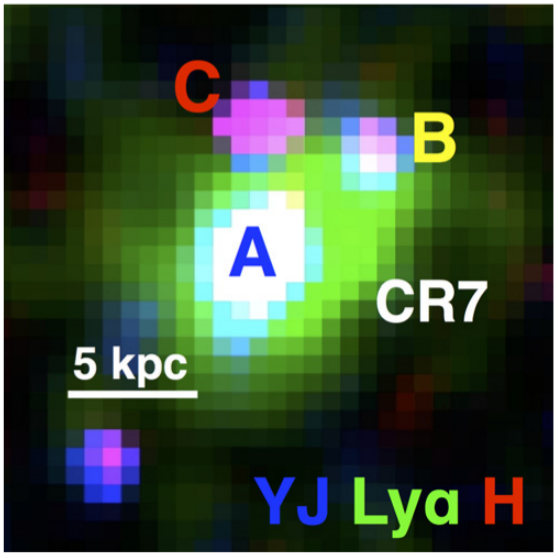


Figure 4.1: False colour composite of CR7. Clump A dominates the Lyα emission and the rest-frame UV light, while the scattered emission, extend all the way to B and part of C, likely indicates a significant amount of gas in the system. Note that the reddest (in rest-frame UV) clump is C, with B having a more intermediate colour and with A being very blue in the rest-frame UV. Adapted from Sobral et al. (2015).

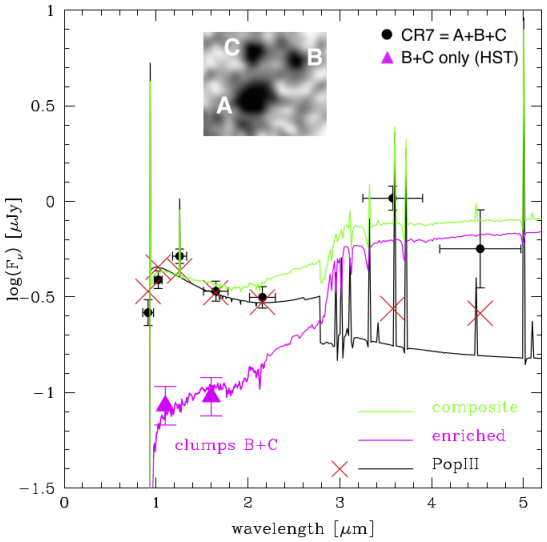


Figure 4.2: Observed SED of CR7 plus HST photometry for clumps B+C (magenta triangles). The black line shows a fit with a pure Pop III SED to the rest-frame UV part; the magenta line the SED of an old simple stellar population with 20% solar metallicity which matches the flux from clumps B+C; the green line shows the predicted SED summing the two populations after rescaling the Pop III SED by a factor of 0.8. The composite SED reproduces well the observed photometry. Adapted from Sobral et al. (2015).

4.3.1 Fiducial model

Cosmological context

According to the hierarchical scenario of structure formation, halos merge over time to form larger structures. The distribution of halo masses as a function of redshift can be described analytically by the model of Press & Schechter (1974). Based on this idea, Bond et al. (1991) and Lacey & Cole (1993) developed methods to construct assembly histories of individual halos that allow the construction of dark matter merger trees. Our code is based on the merger tree algorithm by Parkinson et al. (2008), which generates dark matter merger trees with arbitrary mass resolution. We use a resolution mass of $M_{\rm res} = 2.5 \times 10^5\,{\rm M}_\odot$ for the merger tree, which is sufficient to resolve all halos for the redshifts of interest (Magg et al. 2016). We assume a flat ΛCDM Universe and use the Planck Collaboration et al. (2016) cosmological parameters, most importantly the new optical depth to Thompson scattering $\tau_e = 0.066 \pm 0.016$, which is significantly lower than $\tau_e = 0.0907 \pm 0.0102$ from their previous release (Planck Collaboration 2014).

Halo mass

As a starting point to constructing the merger tree backwards in cosmic time, we need an approximation for the halo mass of CR7 at $z = 6.6$. We use the value of $1.6 \times 10^{10}\,{\rm M}_\odot$ proposed by Sobral et al. (2015) as a fiducial total stellar mass for our model. This mass has been derived with SED fitting under the assumption that CR7 hosts a normal stellar population, although this might not be the case. This

simplifying assumption might lead to errors of a factor of a few in the estimation of the total stellar mass. Moreover, estimates of stellar mass based on SED fitting are subject to uncertainties related to degeneracies between several stellar population parameters (SFR, metallicity, IMF, rotation, binarity).

To determine the halo mass for a given stellar mass we use the model by Behroozi et al. (2013), who constrain SFRs and histories as a function of halo mass up to $z = 8$ from observations of the stellar mass function, the cosmic SFR, and the specific SFR. For a stellar mass of $1.6 \times 10^{10}\,\mathrm{M}_\odot$ it yields a halo mass of

$$M_h = (1.2 \pm 0.2) \times 10^{12}\,\mathrm{M}_\odot, \tag{4.1}$$

where the uncertainty represents the cosmological scatter. This is close to the assumed value of $M_h = 10^{12}\,\mathrm{M}_\odot$ in Agarwal et al. (2016a), but slightly higher than the values derived in Bowler et al. (2017). We use this as a fiducial mass in our model, but also test other halo masses to check the dependence of the results on this very uncertain parameter.

Star formation

For each merger tree, we model the formation of Pop III stars and subsequent stellar populations. For a halo to be able to form Pop III stars, it has to fulfil four criteria. First, the halo has to be metal-free and not be polluted by in-situ star formation from progenitors or enriched externally by supernovae (we test the effect of a critical metallicity threshold in Section 4.4.1). Moreover, the halo has to be above the critical mass

$$M_\mathrm{crit} = 3 \times 10^5 \left(\frac{1+z}{10}\right)^{-3/2} M_\odot \tag{4.2}$$

to enable the primordial gas to cool efficiently and trigger star formation (Hummel et al. 2012). Furthermore, mergers dynamically heat the gas and can delay Pop III star formation. Hence, the mass growth rate of a certain halo has to be below

$$\frac{\mathrm{d}M}{\mathrm{d}z} \lesssim 3.3 \times 10^6\,\mathrm{M}_\odot \left(\frac{M}{10^6 M_\odot}\right)^{3.9} \tag{4.3}$$

to enable primordial star formation (Yoshida et al. 2003), or otherwise star formation is delayed. Finally, Lyman-Werner (LW) photons can also delay or even prevent the collapse of primordial gas by photodissociating H_2. Following Machacek et al. (2001), the fraction of the total gas mass that cools and collapses under the influence of a LW background is

$$f_\mathrm{LW} = 0.06 \ln\left(\frac{M_\mathrm{halo}/M_\odot}{1.25 \times 10^5 + 8.7 \times 10^5 F_\mathrm{LW}^{0.47}}\right), \tag{4.4}$$

where F_LW is the LW flux in units of $10^{-21}\mathrm{erg\ s^{-1}cm^{-2}Hz^{-1}}$. Primordial star formation is suppressed if f_LW falls below zero.

Once we have identified a Pop III star-forming halo, we assign individual stars to it whose masses are randomly sampled from a logarithmically flat IMF between M_min and M_max. The shape of the IMF is motivated by numerical simulations that predict a relatively flat mass distribution that is dominated by high-mass stars (Clark et al. 2011a; Greif et al. 2011; Susa et al. 2014; Hirano et al. 2014; Hosokawa et al. 2016) and the total stellar mass in a halo with virial mass M_vir is

$$M_* = \eta f_\mathrm{LW} \frac{\Omega_b}{\Omega_m} M_\mathrm{vir}, \tag{4.5}$$

where η is the star formation efficiency (SFE) parameter for primordial star formation.

We use spectra tabulated by Schaerer (2002, 2003) to determine the number of ionising and LW photons and the Lyα and HeII line luminosities for each Pop III star individually as a function of its mass. Lifetimes for the stars are also taken from Schaerer (2002, 2003). The associated line luminosities for Pop III stars are listed in Table 4.1. We adopt luminosities averaged over the stellar lifetime, which are smaller than the zero-age main sequence luminosities by a factor of a few. This might be relevant for a recent burst in Pop III stars but does not change our final conclusions.

We model the formation of subsequent generations of stars (Pop II) from the observed cosmic star formation history at high redshift and account for their chemical and radiative feedback (ionising and LW photons). We take cosmic SFRs from Behroozi & Silk (2015), extrapolate them to $z > 15$ and set them to zero for $z > 30$.

Escape fractions

For the escape fraction of ionising radiation we use $f_{\rm esc,III} = 0.5$ for Pop III star-forming halos and $f_{\rm esc} = 0.1$ for later generations of stars in more massive halos. Generally, the escape fraction depends on the halo mass, redshift and the stellar physics (Paardekooper et al. 2013; Trebitsch et al. 2015), but we use these average values for our simplified model, which are in good agreement with previous studies (Johnson et al. 2009; Finkelstein et al. 2012; Wise et al. 2014; Xu et al. 2016b). A higher escape fraction of ionising radiation would lead to a lower SFE, which only weakly affects the final results as we show in Section 4.4.2.

Escape fractions for LW photons are taken from Schauer et al. (2015), who find that they are a function of both halo and Pop III star mass. The escape fraction is dominated by the most massive star in each halo, which first ionises the surrounding gas. Hence, we use the escape fraction values in the far field approximation for the most massive star in each halo.

Metal enrichment and critical metallicity

After its lifetime and depending on its mass, a Pop III star explodes in a supernova (SN) and pollutes its environment with metals. We use metal yields from Heger & Woosley (2010) and assume a Sedov-Taylor expansion of the enriched volume. By summing the volumes of these SN remnants in all halos, we calculate the time-dependent fraction of the total volume that is already polluted. Once a new halo forms, we check statistically whether it is already metal-enriched or still pristine. A key assumption of this model is that the SN remnant is still in the Sedov-Taylor phase when it breaks out of its host halo. This is valid for minihalos, as the HII region expansion has already cleared most of the gas from the halo (Whalen et al. 2004; Kitayama et al. 2004; Alvarez et al. 2006; Abel et al. 2007). For later generations of stars in more massive halos, we assume that they do not contribute to the pollution of the IGM by metals (although see Mac Low & Ferrara 1999).

In principle, a halo can host several Pop III SNe and their metal-enriched remnants can overlap. Since the expansion of the enriched gas is dominated by the most massive (and consequently first) SN to explode (Ritter et al. 2015), we only account for the expansion of the SN of the most massive progenitor per halo. This approximation is valid, as long as there are not too many highly energetic SNe in one halo, which is generally not the case in our models.

We are not only interested in whether a halo is polluted by metals, but also in the metallicity of the polluted gas. This information enables us to allow Pop III star formation not only in pristine gas but also below a certain critical metallicity $Z_{\rm crit}$. Due to the lack of spatial information in the merger tree,

we use the approximation that all metals are deposited in the outer shell of the expanding SN. Hence, we calculate a metal surface density, $\Sigma_m(t)$, for each SN explosion that decreases as the SN expands, and we construct a time dependent probability distribution for these surface densities. When we find that a halo is polluted with metals, we randomly draw a $\Sigma_m(t)$ from the probability distribution and determine the mass of the metals with which the newly formed halo is polluted from $M_m = \Sigma_m(t)\pi R_{\rm vir}^2$, where $R_{\rm vir}$ is its virial radius. Assuming that these metals mix homogeneously with the gas, we calculate its metallicity as follows:

$$Z = \frac{M_m}{0.02\, M_{\rm vir}\, \Omega_b/\Omega_m}, \tag{4.6}$$

which yields $10^{-6} \lesssim Z/Z_\odot \lesssim 10^{-2}$ for external metal enrichment by SNe. The accretion and inflows of pristine gas on to the halos are taken into account self-consistently due to the smooth accretion of gas below the resolution limit of the merger tree. Note that we state all metallicities in units of the solar metallicity, $Z_\odot$.

SFEs based on merger history

As an alternative to a constant SFE, we include another model for star formation based on the merger history of the halos. It has been shown that mergers can enhance star formation because of tidal torques, which allow the efficient transport of gas to the centre of the galaxies where the dense gas can cool and form stars (e.g., Sanders & Mirabel 1996; Cox et al. 2008). This is further supported by observations, which show a negative correlation between star formation indicators and the projected distance of galaxies (Barton et al. 2000; Lambas et al. 2003; Smith et al. 2007). To account for this effect, we use the model of Cox et al. (2008) to determine the SFE. Cox et al. (2008) study the effect of the galaxy mass ratio on merger-driven starbursts with numerical simulations for typical galaxies in the local Universe. For disc galaxies with halo masses in the range $(0.5 - 11.6) \times 10^{11}\,{\rm M}_\odot$, the burst efficiency is best described by a fit of the form

$$\eta_{\rm burst} = \epsilon_{1:1} \left(\frac{M_{\rm sat}}{M_{\rm primary}} \right)^{\alpha}, \tag{4.7}$$

where $\epsilon_{1:1}$ is the burst efficiency for equal mass mergers, $M_{\rm sat}$ is the mass of the satellite and $M_{\rm primary}$ the mass of the primary. The stellar mass per halo is then

$$M_* = \eta_{\rm burst} \frac{\Omega_b}{\Omega_m} M_{\rm vir}. \tag{4.8}$$

For the local Universe and specified mass range, they propose values of $\epsilon_{1:1} = 0.55$ and $\alpha = 0.69$. Applying these values to higher redshifts and less massive halos leads to drastic overestimates of SFRs and premature reionisation ($\tau_e = 0.287$). $\epsilon_{1:1}$ depends more strongly on the absolute mass of the galaxies than α, so we treat it as a free parameter to match the reionisation history of the Universe and keep α constant.

Sampling the halo mass function and calibrating SFEs

To determine the star formation efficiencies η and $\epsilon_{1:1}$ we calibrate them against τ_e, which is a measure of the integrated ionisation history of the Universe. The optical depth is very sensitive to the number of ionising photons in the early Universe and can therefore be used to calibrate the SFE. To do so, we sample the halo mass function at $z = 6.6$ from $10^8 - 10^{13}\,{\rm M}_\odot$ and weight the number of ionising

label	$M_{\rm min}$	$M_{\rm max}$	SFE	$Z_{\rm crit}/Z_\odot$
fiducial	$3\,{\rm M}_\odot$	$300\,{\rm M}_\odot$	Eq. 4.5, $\eta = 0.14$	0
1→100	$1\,{\rm M}_\odot$	$100\,{\rm M}_\odot$	Eq. 4.5, $\eta = 0.20$	0
$Z_{\rm crit}$	$3\,{\rm M}_\odot$	$300\,{\rm M}_\odot$	Eq. 4.5, $\eta = 0.12$	$10^{-3.5}$
merger	$3\,{\rm M}_\odot$	$300\,{\rm M}_\odot$	Eq. 4.8, $\epsilon_{1:1} = 4 \times 10^{-3}$	0

Table 4.2: Overview of the four models we use for primordial star formation. The fiducial model assumes an IMF from $3 - 300\,{\rm M}_\odot$ and only pristine gas can form Pop III stars. In the second model, we change the IMF to slightly lower masses from $1 - 100\,{\rm M}_\odot$. The third model allows Pop III star formation up to a metallicity of $Z_{\rm crit} = 10^{-3.5}\,Z_\odot$. In the last model, we determine the stellar mass per halo based on the merger history. In all the models, we calibrate the SFE to reproduce the optical depth to Thomson scattering.

photons from a given halo by the number density of halos of this mass at that redshift. This yields a cosmologically representative sample, since halos with masses below $10^8\,{\rm M}_\odot$ at $z = 6.6$ hardly produce any ionising photons and halos above $10^{13}\,{\rm M}_\odot$ are very rare at this redshift. For more details and a thorough comparison to analytical models of the halo mass function see Magg et al. (2016).

For 11 equally distributed halo masses in this range we generate 100 merger trees, which yield a statistically representative number of realisations. We then calculate τ_e from the ionisation history of these halos. We chose the SFE accordingly, to match the observed value of $\tau_e = 0.066$. Reionisation is mainly driven by Pop II stars, but we also need the contribution by primordial stars. We do not account for other sources in the total ionising budget in the early Universe, such as quasars (e.g., Volonteri & Gnedin 2009; Madau & Haardt 2015) or high-mass X-ray binaries, and the uncertainty in τ_e might yield different SFEs. However, we demonstrate in Section 4.4.2 that a different SFE has no influence on metal enrichment and our final conclusions.

For this implementation of the halo mass function and 100 randomly generated merger trees we probe a cosmologically representative volume of $\sim 10^6\,{\rm Mpc}^{-3}$. This is statistically sufficient for most of our purposes, but might be too small to probe certain rare scenarios of SMBH seed formation (see Section 4.3.4).

4.3.2 Models of Pop III star formation

We model the star formation history of CR7, focusing on the primordial stellar component, with our semi-analytical approach. To investigate the possibility of having $10^7\,{\rm M}_\odot$ of Pop III stars at $z = 6.6$ and the corresponding high HeII line luminosity and EW we test several models of primordial star formation, which are summarised in Table 4.2. For the 'fiducial' model we assume a logarithmically flat IMF from $M_{\rm min} = 3\,{\rm M}_\odot$ to $M_{\rm max} = 300\,{\rm M}_\odot$ and an SFE $\eta = 0.14$, which yields an optical depth $\tau_e = 0.067$. The mass range of the IMF is consistent with recent simulations (Greif et al. 2011; Susa et al. 2014; Hirano et al. 2014; Hosokawa et al. 2016) and it covers all possible stellar remnants relevant to our model. In a second model, which we label '1→100', we assume a lower-mass IMF from $M_{\rm min} = 1\,{\rm M}_\odot$ to $M_{\rm max} = 100\,{\rm M}_\odot$, motivated by recent simulations of primordial star formation (Clark et al. 2011a; Stacy et al. 2012; Dopcke et al. 2013; Hartwig et al. 2015b; Stacy et al. 2016), which manifest disc fragmentation and hence lower-mass Pop III stars. This model should illustrate the effects of a different mass range of primordial stars, since the actual values are not well constrained. In this second model, we adopt $\eta = 0.20$ which results in $\tau_e = 0.066$.

The main difference between primordial and later generations of star formation is the ability of the gas to cool efficiently. Metal lines can cool the gas to lower temperatures than cooling by molecular hydrogen, which is the most efficient coolant in primordial gas. Hence, the Jeans mass in metal enriched

gas is smaller than in primordial gas and the cloud consequently fragments into more and smaller clumps, which then collapse to form protostars. So far we assumed that Pop III stars form from pristine gas with a top-heavy IMF, but several studies show that even with trace amount of metals or dust a top-heavy IMF is possible (Bromm et al. 2001a; Schneider et al. 2002, 2012; Frebel et al. 2007; Dopcke et al. 2013; Safranek-Shrader et al. 2014).

Consequently, in the third scenario, named '$Z_{\rm crit}$', we assume that Pop III stars form with a top-heavy IMF out of metal enriched gas with a metallicity of $Z < Z_{\rm crit}$ with $Z_{\rm crit} = 10^{-3.5}\,Z_\odot$. Although dust cooling can yield a lower $Z_{\rm crit}$ (Schneider et al. 2012; Dopcke et al. 2013), we adopt this value as a conservative upper limit. We use a flat IMF from $M_{\rm min} = 3\,{\rm M}_\odot$ to $M_{\rm min} = 300\,{\rm M}_\odot$. The effective temperature is lower for a metal enriched star than for a primordial star of the same mass so its spectrum is softer (Bromm et al. 2001b). The LW and Lyα luminosities are only weakly affected, but the HeII line luminosity is generally smaller for $0 < Z \lesssim Z_{\rm crit}$ compared to the metal-free case (Cojazzi et al. 2000; Schaerer 2003). On average, it is smaller by a factor of ~ 10, but the exact value depends on the treatment of stellar winds and metallicity. We will still use the line luminosities for the metal-free case but keep in mind that this yields a strict upper limit for the HeII line. We also verified that the final results are insensitive to the choice of $Z_{\rm crit}$. This is in agreement with Latif et al. (2016a), who show that the fraction of halos that are enriched up to a certain metallicity is only a weak function of the actual metallicity for $10^{-6} < Z/Z_\odot < 10^{-4}$. An SFE $\eta = 0.12$ yields an optical depth $\tau_e = 0.068$.

In the fourth model, which we label 'merger', we couple the star formation to the merger history based on Equation 4.8 with $\epsilon_{1:1} = 4 \times 10^{-3}$. For mergers with a mass ratio below 0.2, we set $\epsilon_{1:1} = 10^{-5}$, which allows about one star to form per halo. The IMF extends from $M_{\rm min} = 3\,{\rm M}_\odot$ to $M_{\rm min} = 300\,{\rm M}_\odot$ and only pristine gas can form Pop III stars. We use radiative feedback in this model only to check if a minihalo can collapse in a given LW background, but we do not use this value to determine the final mass in Pop III stars as we do in the other models. We obtain an optical depth of $\tau_e = 0.067$ with this model.

One should keep in mind that our understanding of primordial star formation is still incomplete and uncertain due to the lack of any direct observations. We try to overcome this uncertainty by implementing different scenarios of Pop III star formation that cover the most likely theories about the formation of the first stars. We have also tested other parameters, such as a primordial IMF extending to masses above $300\,{\rm M}_\odot$ (see Section 4.4.1), but since these results deviate even further from the observational constraints of CR7, we do not explicitly discuss them here.

4.3.3 Pop III remnant black hole

We also investigate the possibility that CR7 hosts a massive BH. Pallottini et al. (2015) show that a BH with an initial mass of $\sim 10^5\,{\rm M}_\odot$ embedded in a halo of total gas mass $10^7\,{\rm M}_\odot$ can account for the observed line luminosities about 100 Myr after formation. This result was derived by coupling a 1D radiation-hydrodynamic code (Pacucci & Ferrara 2015) to the spectral synthesis code CLOUDY (Ferland et al. 2013) as described in detail in Pacucci et al. (2015a). After $\sim 100\,{\rm Myr}$ the gas in the halo is depleted and the HeII luminosity decreases. Generally it is possible to obtain the necessary fluxes at later times if the reservoir of metal-poor gas is large enough. First, we explore the formation scenario in which massive BHs grow from Pop III stellar remnants (Madau & Rees 2001; Haiman & Loeb 2001; Volonteri et al. 2003; Whalen & Fryer 2012).

The final fate of a star depends mainly on its mass (see Table 1.1). Pop III stars with masses of $25\,{\rm M}_\odot \leq M_* \leq 140\,{\rm M}_\odot$ or with $M_* > 260\,{\rm M}_\odot$ collapse to BHs, although the exact ranges depend on

rotation and magnetic field strengths (Karlsson et al. 2013). $140 - 260\,\mathrm{M}_\odot$ Pop III stars are completely disrupted in pair instability SNe that leave no remnant behind. Primordial stars that directly collapse to a BH do not pollute their host halo with metals. This is an important characteristic of primordial stars, which facilitates having an accreting BH in a metal-free environment.

We trace the formation of Pop III remnant BHs in the assembly history of CR7 and merge the BHs if the mass ratio of a merger of two galaxies is > 0.1 (Taffoni et al. 2003; Volonteri et al. 2003; Van Wassenhove et al. 2014). For smaller mass ratios, we only follow the growth of the more massive BH. The rate at which Pop III remnant BHs grow by accretion is a subject of ongoing debate and depends on the gas supply, the depth of the gravitational potential well of the halo, the merger history, and the radiative feedback (Milosavljević et al. 2009b,a; Park & Ricotti 2011a, 2012; Whalen & Fryer 2012; Park & Ricotti 2013; Pacucci et al. 2015a). Existing simulations have indicated that accretion on to stellar-mass Pop III remnants may be substantially suppressed, due to radiation-hydrodynamical feedback (Johnson et al. 2007; Milosavljević et al. 2009b,a; Jeon et al. 2012, 2014; Pacucci et al. 2015b; Park & Ricotti 2011a; Park et al. 2016). Hence, we do not take gas accretion into account and note that the derived values of the BH masses are a strict lower limit. We discuss the issue of gas accretion on to Pop III remnant BHs in more detail in Section 4.5.

4.3.4 Direct collapse black hole

As a second scenario of massive BH formation we study the direct collapse model, in which a $\sim 10^5\,\mathrm{M}_\odot$ seed BH is formed as a consequence of rapid isothermal collapse (Bromm & Loeb 2003a; Begelman et al. 2006; Latif et al. 2013b; Shlosman et al. 2016). We check if a halo is metal-free and if its virial temperature is above $10^4\,\mathrm{K}$, which implies a minimum halo mass of (Glover 2013b)

$$M_{\mathrm{atom}} = 4 \times 10^7\,\mathrm{M}_\odot \left(\frac{z+1}{11}\right)^{-3/2}. \tag{4.9}$$

To compute the LW flux, we adopt the model by Dijkstra et al. (2014) and Habouzit et al. (2016b). We assume that the flux is provided by one nearby star-forming halo, for which the LW luminosity is

$$L_{\mathrm{LW}} = 10^{47} h\bar{\nu}\,\mathrm{s}^{-1} \frac{M_*}{\mathrm{M}_\odot} \left(1 + \frac{t_6}{4}\right)^{-3/2} \exp\left(-\frac{t_6}{300}\right), \tag{4.10}$$

where $h\bar{\nu} = 2 \times 10^{-11}\,\mathrm{erg}$ is the mean energy of a LW photon, M_* is the stellar mass of the halo, and t_6 is the time in Myr after the initial starburst. Here, we assume that 5% of the gas in the halo turns into stars, which is in agreement with the model by Behroozi & Silk (2015) for the redshifts and halo masses of interest. The collapse time of the atomic cooling halo is approximately 10 Myr (Visbal et al. 2014b), which is equal to the minimum time for which a LW flux $> J_{\mathrm{crit}}$ is required. Since the production of LW photons decreases with time after the initial starburst, we use this time as a minimum requirement to produce sufficient LW photons ($t_6 = 10$). Hence, the distance up to which a star-forming halo of mass M_h can provide a flux $\geq J_{\mathrm{LW}}$ is

$$r_{\mathrm{rad}} = 48\,\mathrm{kpc} \left(\frac{M_h}{10^{11}\,\mathrm{M}_\odot}\right)^{1/2} \left(\frac{J_{\mathrm{LW}}}{100}\right)^{-1/2}. \tag{4.11}$$

We use the general convention to express the LW flux in units of $10^{-21}\,\mathrm{erg\,s^{-1}\,cm^{-2}\,Hz^{-1}\,sr^{-1}}$.

For a given star-forming halo, we calculate the pollution radius, which provides a minimum distance

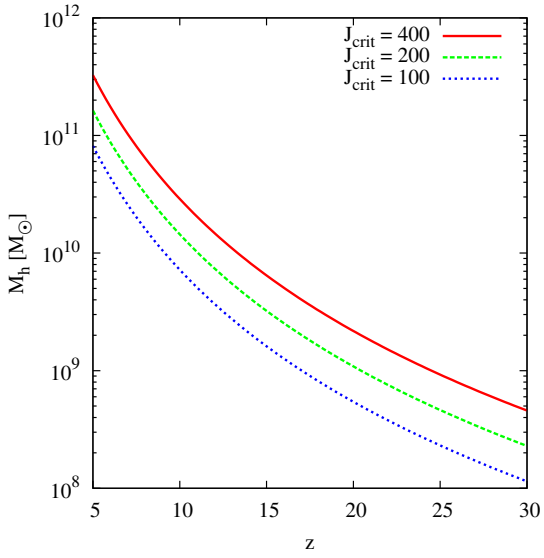


Figure 4.3: Minimum halo mass required to produce $J_{\rm crit}$ without polluting its neighbour halo with metals, plotted as a function of redshift. Depending on the redshift and the required LW flux, a halo of $\sim 10^9 - 10^{11}\,{\rm M}_\odot$ is needed to suppress H_2 formation in its neighbour and produce a DCBH.

between this halo and the atomic cooling halo. Assuming a Sedov-Taylor expansion of the metal enriched galactic winds into the surrounding gas, which has a density of $\Delta = 60$ times the mean density of the IGM (Dijkstra et al. 2014), the radius of metal enrichment can be expressed as

$$r_{\rm metal} = 22\,{\rm kpc}\left(\frac{M_h}{10^{11}\,{\rm M}_\odot}\right)^{1/5}\left(\frac{1+z}{11}\right)^{-6/5}. \tag{4.12}$$

The necessary requirement to form a DCBH is $r_{\rm metal} < r_{\rm rad}$, which translates into a minimum halo mass of

$$M_h > 7.2\times10^9\,{\rm M}_\odot\left(\frac{1+z}{11}\right)^{-4}\left(\frac{J_{\rm LW}}{100}\right)^{5/3} \tag{4.13}$$

for a nearby star-forming halo to provide the LW flux $J_{\rm LW}$. This minimum mass is plotted in Figure 4.3 as a function of redshift for three fluxes. For the considered critical fluxes, halo masses of $\sim 10^9 - 10^{11}\,{\rm M}_\odot$ are required to provide sufficient photodissociating radiation. The halos have to be more massive at lower redshifts because the ambient density of the halos decreases with time, which in turn increases the radius of metal pollution. The critical flux required for isothermal collapse proposed in the literature spans several orders of magnitude and depends on the detailed physics of collapse and the radiation spectrum (Sugimura et al. 2014; Glover 2015a,b; Agarwal et al. 2016b; Latif et al. 2015; Hartwig et al. 2015c). We assume an evolved stellar population with a lower effective temperature here, in contrast to the T5 spectrum in Chapter 2. Assuming a Pop II starburst in a $\sim 10^9\,{\rm M}_\odot$ halo about 10 Myr ago and a distance to this halo of $\gtrsim 10$ kpc, we expect $J_{\rm crit} = 100 - 700$ (Agarwal et al. 2016b). For $J_{\rm crit} \gtrsim 600$ we find no DCBHs in our model, because of the limited cosmological volume that we can simulate. Hence, we vary $J_{\rm crit}$ from 100–400, and discuss our choice of $J_{\rm crit}$ in more detail in Section 4.5.

To identify formation sites of DCBHs we first find a metal-free halo with a mass $\geq M_{\rm atom}$ and its most massive neighbour in the merger tree. These are two halos at the present time step that merge in the next time step. This condition of an incipient merger guarantees a small spatial distance between the two halos. If, for a given $J_{\rm crit}$, the mass of the nearby halo fulfils Equation 4.13, we assume that it

provides a sufficient LW flux without polluting the atomic cooling halo and that a DCBH forms.

Agarwal et al. (2016a) propose that CR7 hosts a DCBH. In $\sim 20\%$ of their merger tree realisations, which represent the mass assembly histories of CR7, a DCBH may form. The formation redshift of the DCBH is $z \sim 20$, which is limited by two factors: at higher redshift, the LW flux, which they calculate from Agarwal et al. (2016b), is not high enough. At lower redshift the formation site of the DCBH is polluted by metals, where they assume that metals are ejected at a constant wind speed of $100\,\mathrm{km\,s^{-1}}$. For a seed mass of $2 \times 10^4\,\mathrm{M_\odot}$, accretion at 40% of the Eddington rate, and an escape fraction of $f_\mathrm{esc} = 0.16$ for Lyα photons, they show that a DCBH is able to reproduce the observed line luminosities of Lyα and HeII. Hence, if the BH is able to accrete metal-free gas for long enough, formation at higher redshifts is also possible. Following Pallottini et al. (2015) and Agarwal et al. (2016a), we assume that a DCBH formed before $z = 7.3$ can account for the observed line luminosities if it accretes low-metallicity gas at $z = 6.6$. Motivated by Latif et al. (2013c) and Ferrara et al. (2014), we assume that DCBHs form with an initial seed mass of $10^4 - 10^5\,\mathrm{M_\odot}$ in our model. These masses, however, are plausibly upper limits, as the strength of the LW flux we assumed is lower than assumed in those papers. Latif & Volonteri (2015), for instance, find that lower LW fluxes result in lower mass concentrations in the precursors of DCBHs when $J_{21} < 1000$.

4.3.5 Determination of the metal tax

A crucial parameter to understand the nature of CR7 is the 'metal tax', respectively the maximum tolerable metallicity that does not violate the observational limits on the metal lines. We use the photoionisation code CLOUDY, version 13.03 (Ferland et al. 2013) to calculate the metal line ratios as a function of the gas metallicity and dust abundance, approximating that the dust abundance and the individual elemental abundances all scale linearly with the total metallicity. We assume an ambient gas density of $10\,\mathrm{cm^{-3}}$ and have checked that the results are only weakly affected by a higher density of $100\,\mathrm{cm^{-3}}$. For the Pop III stellar population, we adopt a 100 kK blackbody spectra, since most of the contribution to the recombination lines of interest might come from such massive, hot stars. To model the emission from the accretion disc of a BH, possibly residing in CR7, we assume an AGN SED based on Richardson et al. (2014). This model is tuned to achieve reasonable agreement along the AGN sequence by matching the HeII/Hβ ratios of observed AGN. The accretion disc of this AGN can be modelled as a multicolour blackbody with a maximum temperature of $T_\mathrm{max} \approx 7 \times 10^5\,\mathrm{K}$. We also tested a different SED to verify that the choice of the AGN model only has minor influence on the line ratio, in which we are interested in. The resulting line ratios as a function of metallicity can be seen in Figure 4.4 for different ionisation parameters (U) and for the two spectral models. The CIII]1908Å/HeII ratio is a strong function of the gas metallicity and the ionisation parameter. For a high ionisation parameter of $\log U = 0$, even solar metallicity gas is in agreement with the observed line ratios. For lower ionisation parameters of $-2 \lesssim \log U \lesssim -1$, as we expect to find for high redshift AGNs (Nagao et al. 2006; Feltre et al. 2016), lower metallicities are required. Given the uncertainties in the model, it is safe to assume

$$Z_\mathrm{limit} = 10^{-2}\,Z_\odot \tag{4.14}$$

as an upper limit of the metallicity in clump A for both scenarios.

A different approach is to further investigate the non-detection of the CIII] doublet at ~ 1908Å. For the spectral resolution of 0.4 Å (Sobral et al. 2015), we expect the EW of CIII] to be below $\lesssim 1$ Å, since it should otherwise be detected as an emission line in the spectrum. This yields an additional

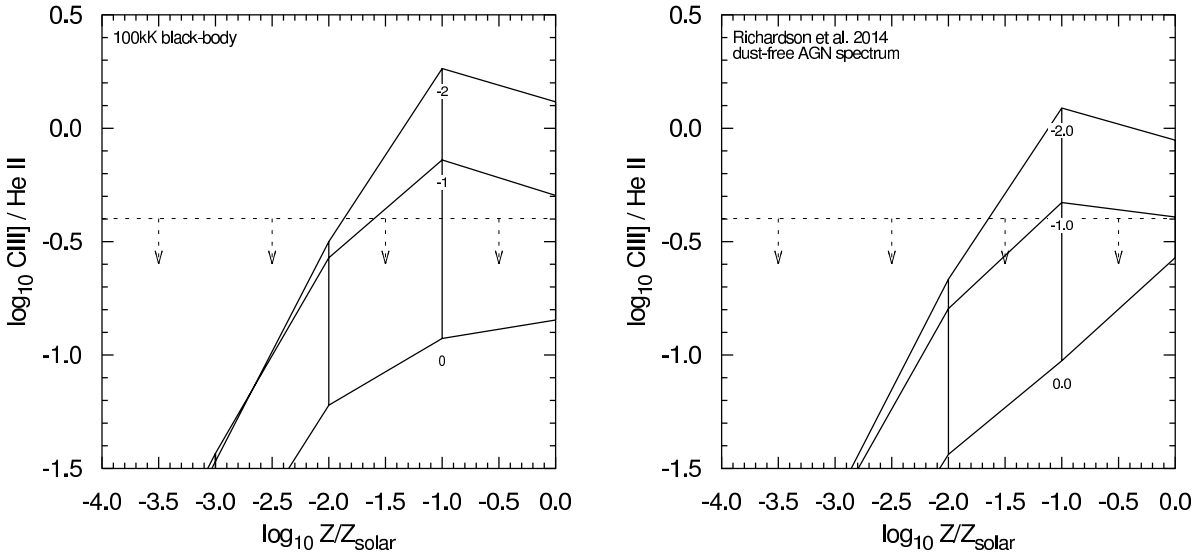


Figure 4.4: Line ratio as a function of the gas metallicity for a 100kK blackbody (left) and an AGN SED (right). The dashed line indicates the upper observational limit of CIII]1908Å/HeII< 0.4 and the three different lines represent different ionisation parameters of the gas. For metallicities below $\sim 10^{-2}\,Z_{\odot}$, all models are in agreement with the observation.

constraint on the gas metallicity as can be seen in Figure 4.5. From the Cloudy model at different ionisation parameters we can constrain that only gas metallicities of $\lesssim 10^{-3}\,Z_{\odot}$ are consistent with the non-detection of the CIII] doublet. Although this is even more constraining than our previous approach based on the line ratios, we use the conservative value of $Z_{\rm limit} = 10^{-2}\,Z_{\odot}$, due to the difficulties in modelling the EW with our simple model. We check the influence of a possibly lower metal tax in the final discussion.

4.4 Results

In this section, we present the main results of our analysis and investigate whether CR7 can be explained with our models of Pop III star formation or by an accreting BH. For each set of parameters, we create 100 independent merger tree realisations and average the derived quantities over the different realisations (if not stated otherwise). This yields a cosmologically and statistically representative sample with statistical scatter of $< 10\%$.

4.4.1 Cosmologically representative models of primordial star formation

We sample the halo mass function from $10^8 - 10^{13}\,{\rm M}_{\odot}$ and weight the contribution to the number of ionising photons by the number density of those halos at $z = 6.6$. This enables us to reproduce the observed value of the optical depth.

The total mass of Pop III stars per halo depends on either the LW background (Equation 4.5) or on the merger history (Equation 4.8). To compare these recipes for star formation, we define the effective SFE as $\eta_{\rm eff} = M_*/M_h$, which is shown in Figure 4.6 as a function of cosmic time. Typical values for $\eta_{\rm eff}$ are $\sim 10^{-4}$ for all the models. Minihalos have masses of the order of $\sim 10^6\,{\rm M}_{\odot}$, and for $\eta_{\rm eff} \sim 10^{-4}$ we form a total stellar mass of $\sim 100\,{\rm M}_{\odot}$ on average per Pop III star-forming minihalo. The random sampling of the IMF leads to statistical variance from halo to halo and there are halos with multiple

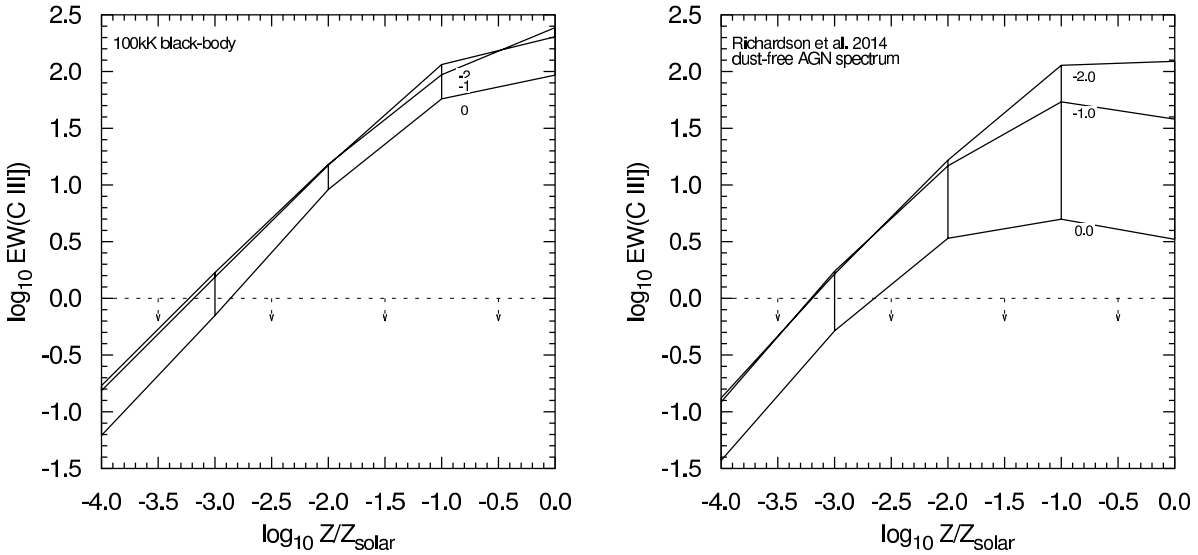


Figure 4.5: Modelled EW of CIII] as a function of the gas metallicity for a stellar population (left) or a AGN spectrum (right). Given the spectral resolution, the EW should be below $\lesssim 1$ Å, which sets an upper limit to the gas metallicity of $\sim 10^{-3}\,Z_\odot$.

Pop III stars and masses $> 100\,\mathrm{M}_\odot$. Indeed, the majority of primordial stars in our model form in multiples of 2-6 stars, as predicted by simulations of early star formation (Clark et al. 2011a; Greif et al. 2011; Stacy et al. 2012; Dopcke et al. 2013; Hartwig et al. 2015b; Stacy et al. 2016).

The cosmic mean Pop III star formation density as a function of time for all four models is shown in Figure 4.7. It is roughly the same for all the models because they are constructed to satisfy the latest constraints on τ_e. The $Z_{\rm crit}$ model's values are slightly higher just shortly before $z = 6.6$ because Pop III stars form also in low-metallicity gas and hence also at later times. Consequently, this model is less affected by metal pollution at lower redshifts. The derived cosmic star formation densities are in agreement with those of Visbal et al. (2015). They show that the recent constraints on τ_e by the Planck Collaboration et al. (2016) limit the mean cosmic star formation density of primordial stars to $\lesssim 10^{-4}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$. The SFR density at $z \sim 6.6$ might not be cosmologically representative due to ongoing Pop III star formation at $z \lesssim 6.6$, which is not captured by our method.

The fact that we have more gas available to form primordial stars in the $Z_{\rm crit}$ model can be seen in the plots of metal-poor gas mass as a function of time in Figure 4.8. At a given redshift, this is the sum of the metal-free gas in all resolved halos in the merger tree. Hence, it is a measure of the maximum available mass to form Pop III stars, assuming an SFE of 100%. For all the models this mass remains below $\sim 10^8\,\mathrm{M}_\odot$. In the model in which we allow primordial stars to also form in metal-enriched gas at $Z < Z_{\rm crit}$, we have more gas to form Pop III stars at lower redshifts and still $10^7\,\mathrm{M}_\odot$ of low-metallicity gas just above $z = 6.6$. These values are derived for $M_h = 1.2 \times 10^{12}\,\mathrm{M}_\odot$.

The stellar mass in Pop III stars and HeII line luminosities are shown in Figure 4.9. For the lower mass IMF there are $10^5\,\mathrm{M}_\odot$ of Pop III stars in CR7, whereas the other models yield values of $10^2 - 10^3\,\mathrm{M}_\odot$. This is in agreement with the results of Xu et al. (2016a), who find in their cosmological simulation $\lesssim 10^3\,\mathrm{M}_\odot$ of Pop III stars in halos at $z = 7.6$. The mass of pristine gas drops steeply before this redshift and only less massive stars with longer lifetimes can survive to be present in CR7. The HeII luminosity is a steep function of the stellar mass, and massive stars are favoured to reproduce the observations. The $Z_{\rm crit}$ model yields larger masses of low-metallicity gas down to lower redshifts and consequently allows primordial star formation at later times. Hence, more massive stars can also survive to contribute

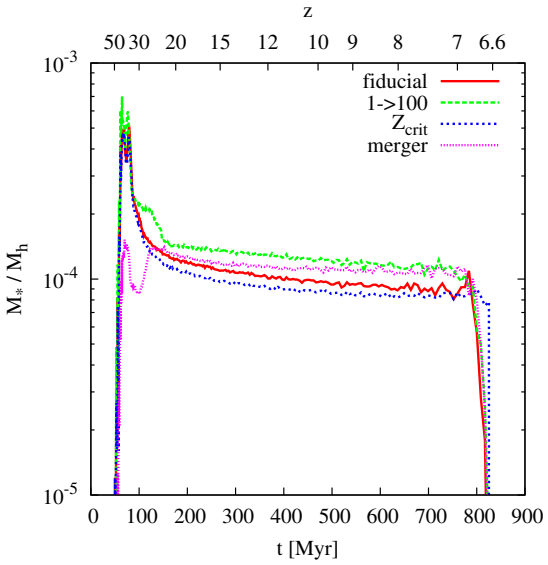


Figure 4.6: Effective SFE for Pop III stars as a function of cosmic time for the four models. After a small early peak, the SFE is approximately constant around 10^{-4}.

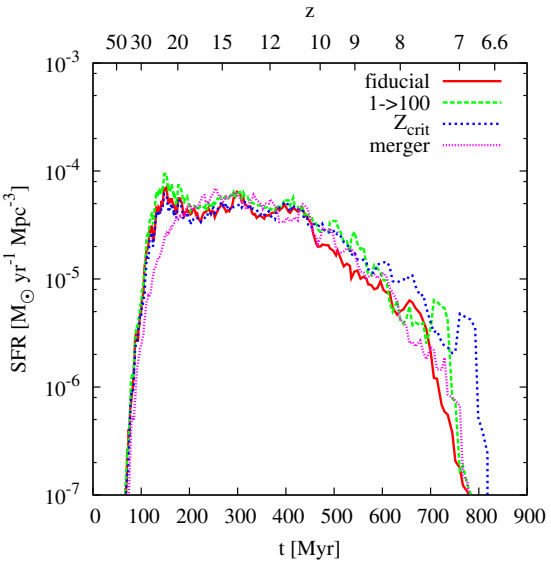


Figure 4.7: Mean cosmic Pop III SFR for all four models per comoving volume. The primordial SFR peaks around redshift 20, and decreases at $z < 12$.

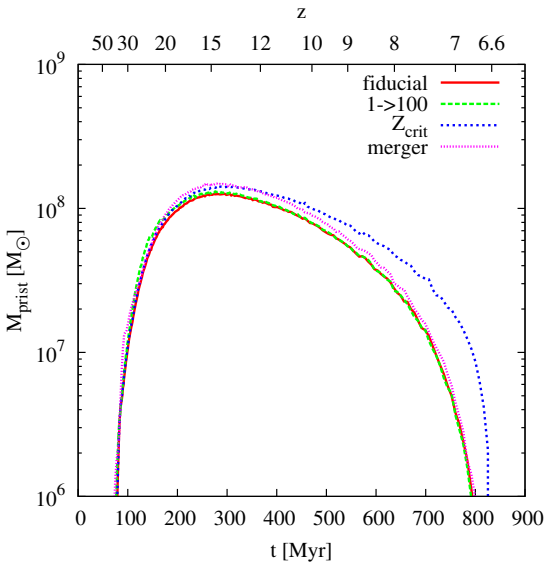


Figure 4.8: Mass of metal-free or metal-poor gas as a function of time for all four models. Here, we assume that the final halo has a mass $M_h = 1.2 \times 10^{12}\,\mathrm{M}_\odot$ at $z = 6.6$ and we only account for gas in its resolved progenitor halos, which explains the rise at early times. This plot illustrates the maximum available amount of gas to form Pop III stars, but even in the most promising Z_{crit} model the mass of gas available for Pop III star formation is limited to $\lesssim 10^8\,\mathrm{M}_\odot$.

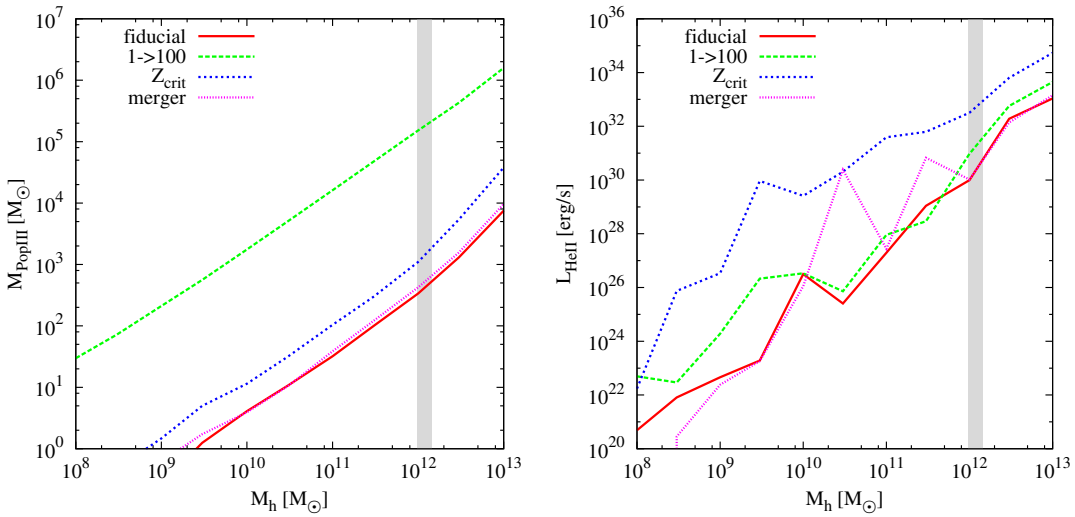


Figure 4.9: Stellar masses in Pop III stars (left) and corresponding HeII luminosities (right) at $z = 6.6$ for a variety of halo masses. Here, we illustrate the entire range of simulated halo masses, and the most likely mass range for CR7 is indicated in grey. The model with a lower mass IMF produces a significantly higher stellar mass in Pop III stars, mainly because of the longer lifetimes of these stars. In the expected mass range of CR7, the stellar primordial mass is limited to $10^2 - 10^5\,\mathrm{M}_\odot$, depending on the model. The Z_{crit} model produces the highest HeII luminosity, because this model is less affected by metal enrichment at lower redshifts and can hence form Pop III stars out to later times. The HeII luminosity is significantly lower than the observed value of $1.95 \times 10^{43}\mathrm{erg\,s^{-1}}$.

to the HeII luminosity at $z = 6.6$. The Z_{crit} model yields the highest HeII luminosities, which are still more than 10 orders of magnitude below the observed value of $1.95 \times 10^{43}\mathrm{erg\,s^{-1}}$. Even for a higher halo mass, which might be possible within the uncertainties of the determination of the stellar and halo mass of CR7 (see Section 4.3.1), the final luminosities are too small. We note that the corresponding Lyα luminosity for Pop III stars in a $\sim 10^{12}\,\mathrm{M}_\odot$ halo is of the order of $10^{36} - 10^{37}\,\mathrm{erg\,s^{-1}}$, which is about 7 orders of magnitude below the observation.

What mostly limits the HeII luminosity is the mass of Pop III stars that survive until $z = 6.6$. To show this effect, we plot the stellar mass distribution of primordial stars in Figure 4.10. The more massive the stars, the shorter their lifetimes and the smaller the probability that they survive long enough to be present down to $z = 6.6$. For $M < 2\,\mathrm{M}_\odot$ this plot represents the IMF, since the lifetimes of these stars are long enough for them to survive until $z = 6.6$. For higher masses we see that the Z_{crit} model is the most likely one to also contain stars that are $\sim 10\,\mathrm{M}_\odot$ since it has the largest amount of gas available for star formation down to lower redshifts. But even these stars are not massive enough to contribute significantly to the HeII luminosity, due to the steep dependence of the HeII luminosity on the stellar mass (see also Table 4.1). We also test more extreme models for the Pop III IMF with a mass range from $M_{\mathrm{min}} = 10\,\mathrm{M}_\odot$ to $M_{\mathrm{max}} = 1000\,\mathrm{M}_\odot$ and find significantly fewer primordial stars at $z = 6.6$ and also a smaller HeII line luminosity than in the fiducial model. For $M_{\mathrm{min}} \gtrsim 50\,\mathrm{M}_\odot$ there are no Pop III stars at all that might contribute to the HeII luminosity at $z = 6.6$, because such massive stars explode within a few Myr as SNe.

Coupling star formation to the merger history of the halos induces a higher scatter in the stellar mass per halo. Hence, the merger model leads to a broader distribution of HeII luminosities at $z = 6.6$. We show the probability distribution function of the luminosities at this redshift in Figure 4.11. The expected HeII luminosities for a halo mass of $M_h = 10^{12}\,\mathrm{M}_\odot$ span more than six orders of magnitude

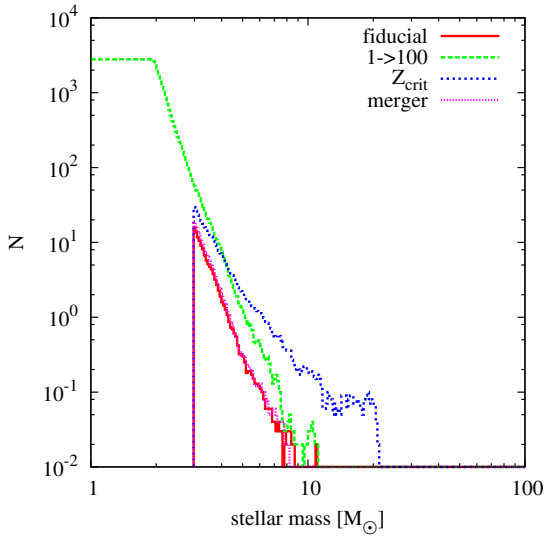


Figure 4.10: Number of stars per logarithmic mass bin that are present in the final $1.2 \times 10^{12}\,\mathrm{M}_\odot$ halo at $z = 6.6$, summed over 100 realisations. The more massive the stars, the shorter their lifetimes and the lower the probability that they survive from their time of formation until the final redshift. The model with the less massive IMF produces many low-mass stars, which survive down to $z = 6.6$ if they are less than $\sim 2\,\mathrm{M}_\odot$. In the Z_{crit} model primordial stars form up to $z \lesssim 6.6$, so more massive survivors are present in the final halo.

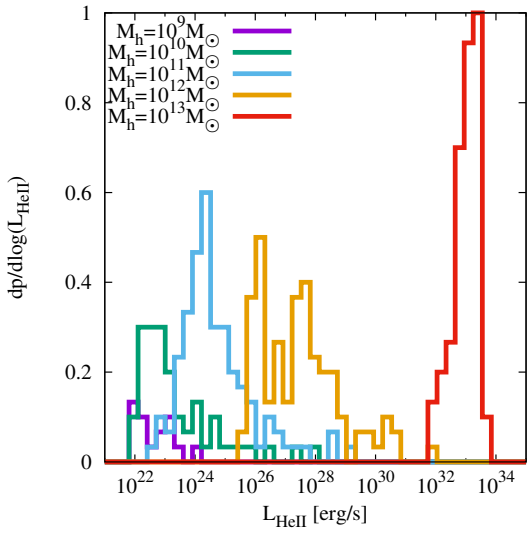


Figure 4.11: Probability distribution function for the expected HeII line luminosities in the merger model for a variety of final halo masses. Due to the random sampling of merger histories, the scatter in the final HeII luminosity in this model is broader compared to the other models. However, even the scatter and the associated probabilities for large values of L_{HeII} cannot account for the missing 10 orders of magnitude between model and observation.

with a maximum value of $\sim 10^{32}\,\mathrm{erg\,s^{-1}}$. The recipe for star formation in mergers (Equation 4.8) is only an extrapolation to lower-mass halos and higher redshifts. But even in the most optimistic case, in which all the pristine gas turns into Pop III stars during a major merger of two $\sim 10^6\,\mathrm{M}_\odot$ minihalos, ~ 100 such mergers are required just above $z = 6.6$ to account for $10^7\,\mathrm{M}_\odot$ of Pop III stars at that redshift. It is therefore difficult to explain the HeII emission in CR7 with our models for primordial star formation.

4.4.2 Alternative scenarios of primordial star formation

The host halo of CR7 corresponds to a rare $\sim 3\sigma$ peak in the cosmological density field at $z = 6.6$. Such rare halos have a comoving number density of only $\sim 10^{-7}\,\mathrm{Mpc^{-3}}$ at this redshift, and it is possible that star formation within these rare objects might proceed differently than in the average galaxy at this redshift. So far, we have assumed that CR7 is cosmologically "representative" and our star formation recipes reproduce the reionisation history of the Universe. Since CR7 cannot be explained in this way, we now drop the constraint of the optical depth, which enables us to vary the SFE of primordial star formation. In other words, we no longer require the mean Pop III SFE in the progenitors of CR7 to be the same as the global mean required to produce the right Thomson scattering optical depth, but instead treat it as a free parameter. We show the effect of changing the SFE in Figure 4.12. Even if we change the SFE by 4 orders of magnitude, the mass of zero-metallicity gas is only weakly affected. In none of the models this mass is sufficient to explain the observational signatures of CR7, and the HeII line luminosity is limited to $\lesssim 3 \times 10^{32}$ erg s^{-1}. To understand this rather weak dependence of pristine gas mass and final luminosity on the SFE, we investigate the two limiting cases for the metal enrichment

model. As described in Section 4.3.1, we only account for the SN expansion of the most massive star per halo, which could either be a core-collapse SN with an explosion energy of $E_0 = 1.2 \times 10^{51}$ erg, or an pair-instability supernova (PI SN) with $E_0 = 10^{52} - 10^{53}$ erg, depending on the mass of the star. For a very low value of η, we form only one star per halo and the probability that this star explodes is $\sim 45\%$ for a logarithmically flat IMF from $3 - 300\,\mathrm{M}_\odot$. For a very high value of η, we form many Pop III stars per halo and have a correspondingly high probability to obtain a star that explodes as a highly-energetic PI SN. The radius of the metal enriched volume is approximately proportional to $E_0^{1/5}$ so it is only weakly affected by SNe with different explosion energies. This explains the small variations in the mass of pristine gas between the models with different SFEs. Having several SNe going off in one halo might break the conservative assumption of our metal enrichment model that the pollution of metals is dominated by the most massive SN. Considering multiple SNe per minihalo would lead to even less pristine gas at lower redshifts.

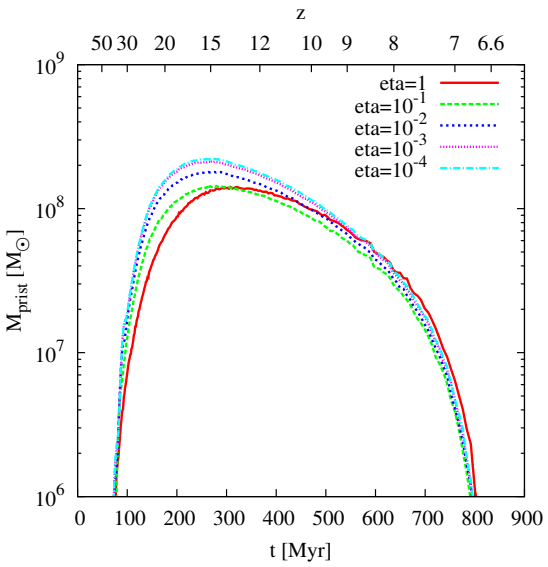


Figure 4.12: Mass of pristine gas in a $1.2 \times 10^{12}\,\mathrm{M}_\odot$ halo as a function of time for a range of SFE parameters. The probability that a halo pollutes its environment with metals is not a strong function of the SFE, but only depends on the probability that there is at least one highly-energetic SN in the halo.

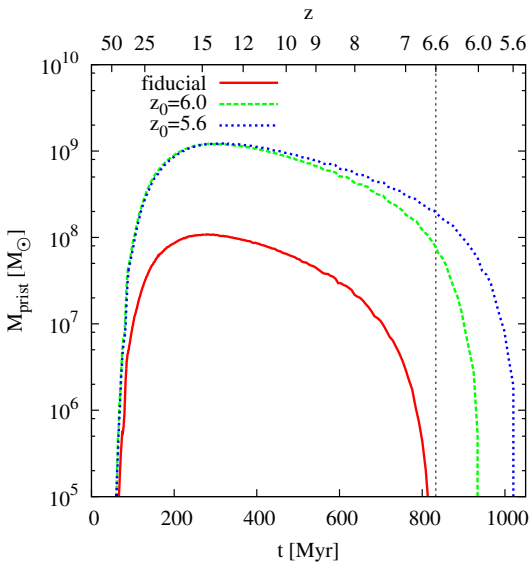


Figure 4.13: Mass of pristine gas as a function of cosmic time for models that assume a final mass of $1.2 \times 10^{12}\,\mathrm{M}_\odot$, but several final redshifts, for which the merger trees are created backwards in time. The gas mass available at this redshift is much higher for those models, but we would still need an SFE of 10% and to overcome other problems (see text).

We have assumed that CR7 is one halo at $z = 6.6$, although we clearly see three distinct clumps. By construction, the final halo is polluted with metals because its progenitors were enriched by SNe or external metal enrichment. Consequently, there is no Pop III star formation at $z = 6.6$. In an alternative scenario, we now assume that CR7 is an ongoing merger and that the three clumps will merge to one halo in 100 Myr ($z_0 = 6.0$) or 200 Myr ($z_0 = 5.6$). Based on Behroozi et al. (2013), we estimate the halo masses of halo A and C to be $M_A = (3 \pm 0.6) \times 10^{10}\,\mathrm{M}_\odot$ and $M_C = (6 \pm 1.2) \times 10^{11}\,\mathrm{M}_\odot$. The merger time of these halos with the more massive halo B are $\gtrsim 50$ Myr at this redshift (Boylan-Kolchin et al. 2008). For $z_0 \lesssim 5.6$, the masses of the third and fourth most massive halo tend to be approximately equal, which does not match the constellation of CR7, where we only observe three clumps (see also Section 4.4.6). This limits the possible range to $5.6 \lesssim z_0 \lesssim 6.0$.

These redshifts are the starting points for constructing merger trees backwards in time so that we can

determine the primordial stellar mass and corresponding HeII luminosity at $z = 6.6$. The primordial gas mass is shown as a function of time in Figure 4.13. The primordial gas mass at $z = 6.6$ is much higher for the halos whose merger tree extends down to redshifts $z_0 < 6.6$. The two additional models yield $\sim 10^8\,\mathrm{M}_\odot$ of pristine gas at $z = 6.6$. The corresponding total masses of Pop III stars are $2 \times 10^4\,\mathrm{M}_\odot$ and $1 \times 10^5\,\mathrm{M}_\odot$, which are 2-3 orders of magnitude higher than in the fiducial model but still not sufficient to account for the observational constraints. The HeII luminosities are $7 \times 10^{39}\,\mathrm{erg\,s^{-1}}$ and $3 \times 10^{40}\,\mathrm{erg\,s^{-1}}$ for the $z_0 = 6.0$ and 5.6 models, respectively. If Pop III stars form instantaneously out of the pristine gas with 10% efficiency just above $z = 6.6$, the mass of pristine gas in these models would be sufficient to explain observations of CR7. However, there is no plausible mechanism that could trigger an instantaneous starburst of this intensity, which is required to explain both the luminosity and the EW of the HeII emission.

Although the models with $z_0 < 6.6$ could in principle explain CR7, there are two shortcomings. First, the primordial stars are distributed over all the CR7 progenitors. In our model, we just add up all the Pop III stars at $z = 6.6$ although they should be confined to clump A. This additional constraint would limit the Pop III stars (and their total mass) to just this one clump (see Section 4.4.6). Another effect becomes important at lower redshifts that we have not considered. After the reionisation of the Universe, photoionisation heating counteracts the cooling in minihalos and might prevent their collapse (Abel & Haehnelt 1999; Pawlik et al. 2009), which further limits the number of Pop III stars that can form at lower redshifts (Magg et al. 2017).

4.4.3 Pop III remnant black holes

While an accreting BH can also explain the line luminosities from CR7, these observations cannot yet differentiate between seed BH formation mechanisms. We first study BH formation in the progenitor halos of CR7 by mergers of Pop III stellar remnants. We follow the mass assembly history of these BHs in merger tree realisations and illustrate 30 randomly selected histories for three host halo masses in Figure 4.14. The Pop III remnant BHs grow to $10^4 - 10^5\,\mathrm{M}_\odot$ by $z = 6.6$, depending on the mass of the halo. Since we only account for mass growth due to mergers of BHs, these values should be treated as a lower limit. We discuss the effect of additional mass accretion in Section 4.5. For the $10^{12}\,\mathrm{M}_\odot$ halo, we expect a BH with a mass of $\sim 10^5\,\mathrm{M}_\odot$ at $z \approx 7$, which could explain CR7 (Pallottini et al. 2015; Pacucci et al. 2016). Note that this quantifies the most massive BH in all progenitors of CR7. In Section 4.4.6 we investigate explicitly those BHs that reside in clump A. Previous studies assume that the BH can accrete low metallicity gas down to $z = 6.6$. This is a strong assumption, and we show in Section 4.4.5 that the host halos of stellar remnant BHs of this mass are generally polluted before $z = 6.6$.

4.4.4 Direct collapse black hole

We also consider under which conditions a DCBH can form in the progenitor halos of CR7. The DCBH formation rate density is shown in Figure 4.15. DCBHs form in the five models at redshifts $z \geq 7.3$. Generally, a higher halo mass and a lower value of J_{crit} facilitates the formation of DCBHs. This is because a higher halo mass yields more progenitors and hence more possible formation sites for a DCBH, whereas for a lower value of J_{crit} halos with lower masses can provide sufficient flux to enable isothermal collapse. The peak formation rate of DCBHs is around $z = 15$ and the rate decreases steeply at $z < 10$. In our models no DCBH can form at $J_{\mathrm{crit}} > 600$ in the progenitors of CR7, whereas this is also limited by the finite number of merger tree realisations. Our implementation yields results similar to those by Dijkstra et al. (2014) and Habouzit et al. (2016b). The difference emerges from our implementation

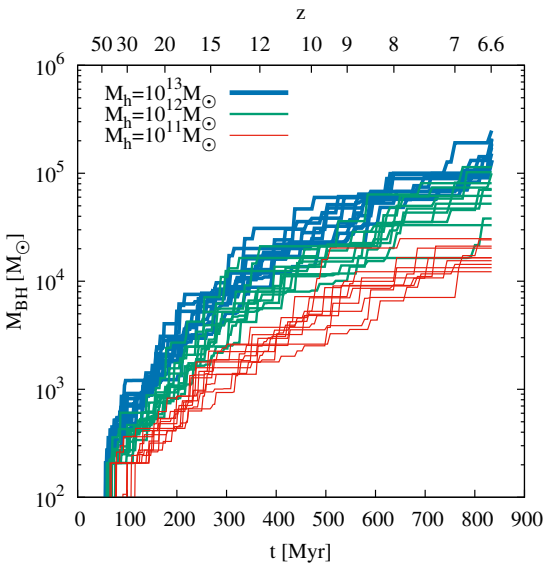


Figure 4.14: Growth history for Pop III stellar remnant BHs that gain mass via galaxy mergers. Here we show 10 random realisations for each halo mass of $10^{11}\,M_\odot$ (red), $10^{12}\,M_\odot$ (green), and $10^{13}\,M_\odot$ (blue). The BHs can reach final masses of $10^4 - 10^5\,M_\odot$ by $z = 6.6$. These are the most massive BHs in all CR7 progenitors and restricting it to the less massive clump A will lower the BH masses as we will see below.

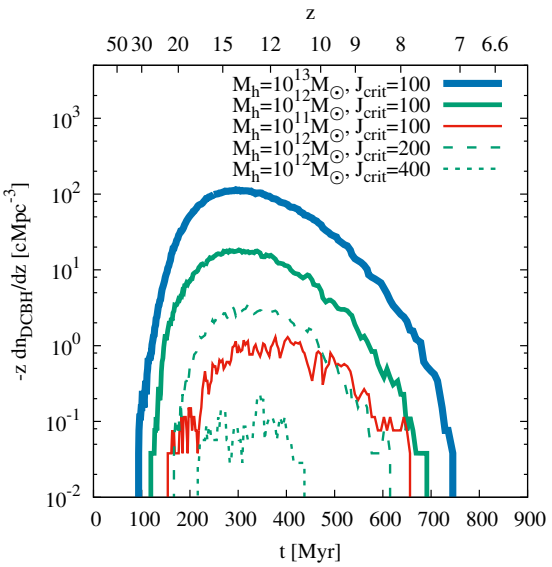


Figure 4.15: DCBH formation rate as a function of redshift for several halo masses and values of $J_{\rm crit}$. The formation rate is normalised to the comoving volume of the main halo, where the red, green, and blue lines are for halo masses of $10^{11}\,M_\odot$, $10^{12}\,M_\odot$, and $10^{13}\,M_\odot$, respectively. The solid, dashed, and dotted green lines show the numbers of DCBHs per redshift for $J_{\rm crit} = 100, 200, 400$, respectively. The comoving volume is calculated as the mass of the main halo M_h divided by the average cosmic density, which yields an upper limit. In all these cases, we can form DCBHs during the assembly of the halo. A larger halo mass and a lower value of $J_{\rm crit}$ facilitates the formation of a DCBH down to lower redshifts. In none of the cases can we form a DCBH at redshifts $z < 7.3$.

of Pop III star formation and the fixed time of 10 Myr that we require for an atomic cooling halo to collapse, whereas they assume a redshift-dependent collapse time, which is generally longer. Therefore, our assumptions are more optimistic, and similar to Agarwal et al. (2012, 2014).

The difference between our results and those of Agarwal et al. (2016a) arise mainly from the treatment of metal enrichment. They assume a constant velocity for the enriched winds that yields a window of 50 Myr in which DCBH formation is possible before the atomic cooling halo is polluted with metals. In our model, we follow the pollution of individual halos self-consistently. However, the important question is not only if DCBHs form in the progenitor halos of CR7, but also if those halos can remain below $Z_{\rm limit}$ for long enough. Otherwise, we should see the imprint of those metals in the spectrum.

4.4.5 Mass of metal-poor gas

We have shown in the previous sections that DCBHs and Pop III remnant BHs can reach the masses needed to explain the Lyα and the HeII luminosities of CR7. However, another important constraint are the upper limits on the OIII]1663Å and CIII]1908Å lines. The absence of these recombination lines requires the photon source to be embedded in metal-poor gas at $z = 6.6$. In this section, we quantify the amount of low-metallicity gas that is present in the progenitors of CR7. We use the fiducial model and

account for gas in halos with a metallicity $Z < Z_{\rm limit}$. In previous sections we analysed the total mass of low-metallicity gas at a given redshift in all progenitors, but now we examine the maximum mass of metal-poor gas in a single halo at a given time. Its evolution over time is shown in Figure 4.16. This plot shows that metal-poor halos with a gas mass of $\sim 10^8\,\mathrm{M}_\odot$ are present down to redshift $z = 6.6$. The more important question is if these metal-poor halos can also host a sufficiently massive BH to explain the observed line luminosities. To answer this question we show the mass of low-metallicity gas in halos that host either a DCBH or a Pop III BH remnant in Figure 4.17. Here, we plot the mass

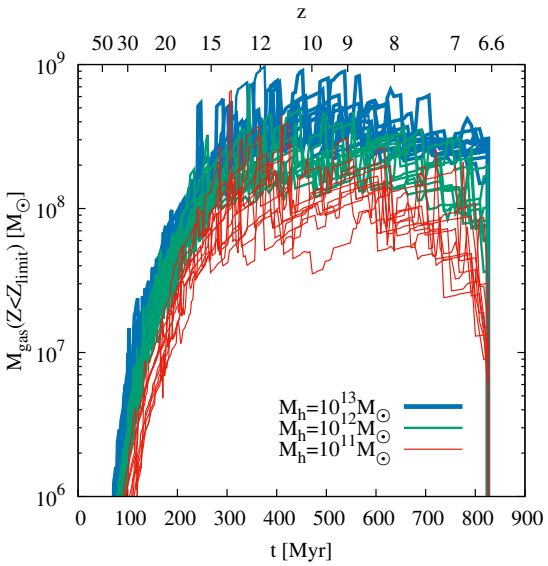


Figure 4.16: Maximum gas mass in the given halo with $Z < Z_{\rm limit}$. We plot 10 random realisations for final halo masses of $10^{11}\,\mathrm{M}_\odot$ (red), $10^{12}\,\mathrm{M}_\odot$ (green), and $10^{13}\,\mathrm{M}_\odot$ (blue). Even at low redshifts, it is possible for halos to have a low metallicity and a gas mass of $\sim 10^8\,\mathrm{M}_\odot$.

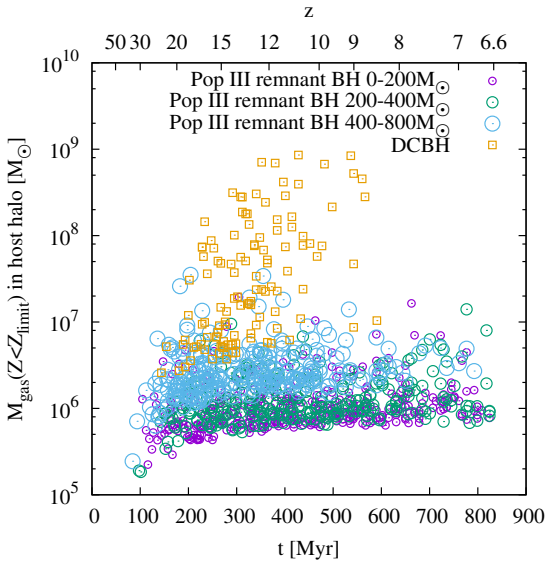


Figure 4.17: Mass of metal-poor gas ($Z < Z_{\rm limit}$) surrounding BHs of different origin. The squares denote DCBHs, the circles are Pop III stellar remnant BHs and the colours and sizes of the circles indicates the mass of the BH. The data is taken one time step before the halos are polluted with metals to $Z > Z_{\rm limit}$, so the symbols mark until when a given type of BH can be surrounded by the indicated mass of metal-poor gas. By construction, the host halo of a DCBH is metal-free at the moment of formation, but shortly afterwards it merges with the enriched halo that previously provided the required LW flux.

of low-metallicity gas surrounding BHs at the moment before the halo is polluted to $Z > Z_{\rm limit}$. This enables us to quantify the lowest redshift at which a BH can reside in a pocket of low-metallicity gas. For the Pop III remnant scenario, we expect BHs in halos with a gas mass of $\lesssim 10^7\,\mathrm{M}_\odot$ down to $z = 6.6$ but these BHs have masses of $\lesssim 800\,\mathrm{M}_\odot$, much smaller than the required mass of $\geq 10^5\,\mathrm{M}_\odot$. This is mainly due to the fact that either the BH progenitor itself pollutes its host halo with metals at the end of its lifetime or that another Pop III star in the same halo enriches it with metals. The only possibility to remain below $Z_{\rm limit}$ is to have Pop III stars that collapse directly to a BH. In our model, Pop III BH remnants only grow by mergers with other BHs and might be polluted during the merger. Only a few BHs merge and remain in a low-metallicity environment. We find few Pop III remnant BHs that can grow to $\sim 800\,\mathrm{M}_\odot$ before their host halo is polluted to $Z > Z_{\rm limit}$.

In contrast, the DCBHs reside by construction in low-metallicity halos with a much higher gas mass of up to $\sim 10^9\,\mathrm{M}_\odot$. These higher masses can be reached because we quench Pop III star formation in the host halo with a strong LW flux from a nearby star-forming halo. We find that DCBHs reside in a metal-

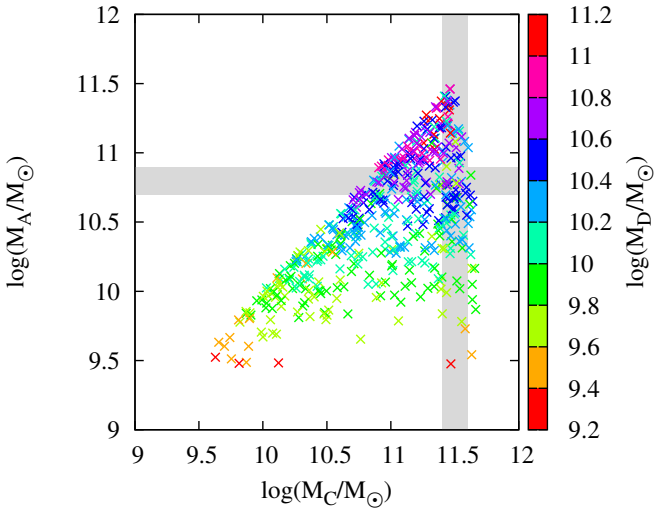


Figure 4.18: Masses of the second, third, and fourth most massive halos at $z = 6.6$, assuming that they merge to one halo within 100 Myr. The grey region indicates the required mass range for clump A and C of CR7 and we show 512 different merger tree realisations. Since we do not see a fourth clump within a projected distance of ~ 10 kpc to CR7, we require that the fourth most massive clump in the merger tree is significantly less massive than clump A. In this sample, we find constellations in the right mass range and with $M_D/M_A \lesssim 0.1$. The most massive clump B has always a mass of $\sim 10^{12}\,\mathrm{M}_\odot$.

free environment for only several tens of Myr before they are polluted by merging with the neighbour halo. DCBHs therefore reside in low-metallicity halos only down to $z \approx 8.5$, which is ~ 240 Myr before $z = 6.6$. When interpreting this result, one should keep in mind that the rapidly-changing metallicity of the host halo due to the merger is an artefact of our recipe for DCBH formation. For this analysis, we use an optimistic value of $J_{\mathrm{crit}} = 100$. The actual value can be up to an order of magnitude higher (Section 4.5) so the derived values should be treated with caution, since a higher J_{crit} might inhibit the formation of DCBHs.

4.4.6 Comparison of scenarios

In this section, we compare the three models and estimate, which has the highest probability to reproduce the observations of CR7. To do so, we assume that CR7 is an ongoing merger of the three clumps A, B, and C, which merge at $z_0 = 6.0$. This is the most plausible scenario to obtain the right constellation of masses in CR7 as shown in Figure 4.18. The most massive halo B always fulfils the requirement of $M_B \approx 10^{12}\,\mathrm{M}_\odot$, whereas only $< 10\%$ of the merger tree realisations yield masses in the right range for halo A and C. As an additional constraint, we require a significant gap between the masses of the third and the fourth most massive halo at $z = 6.6$, because we only see three clumps in CR7 and there is no evidence of a fourth equally massive clump. Hence, we assume that the fourth most massive halo D in the merger tree should be at least an order of magnitude less massive than halo A ($M_D/M_A \lesssim 0.1$). This mass distribution is the anticipated constellation of CR7. We analyse the metallicity, and estimate the masses of both Pop III stars (with the fiducial model) and BHs resulting from the two different seeding scenarios in halo A at $z = 6.6$. The results are shown in Figure 4.19. For a given scenario, this plot illustrates the possible mass-metallicity combinations of the three scenarios. For Pop III stars, the stellar mass in clump A at $z = 6.6$ remains always below $10^3\,\mathrm{M}_\odot$ and the halo has a metallicity of $\sim 10^{-3}\,Z_\odot$. Pop III remnant BHs might reside in metal-poor halos at $z = 6.6$, but are not massive

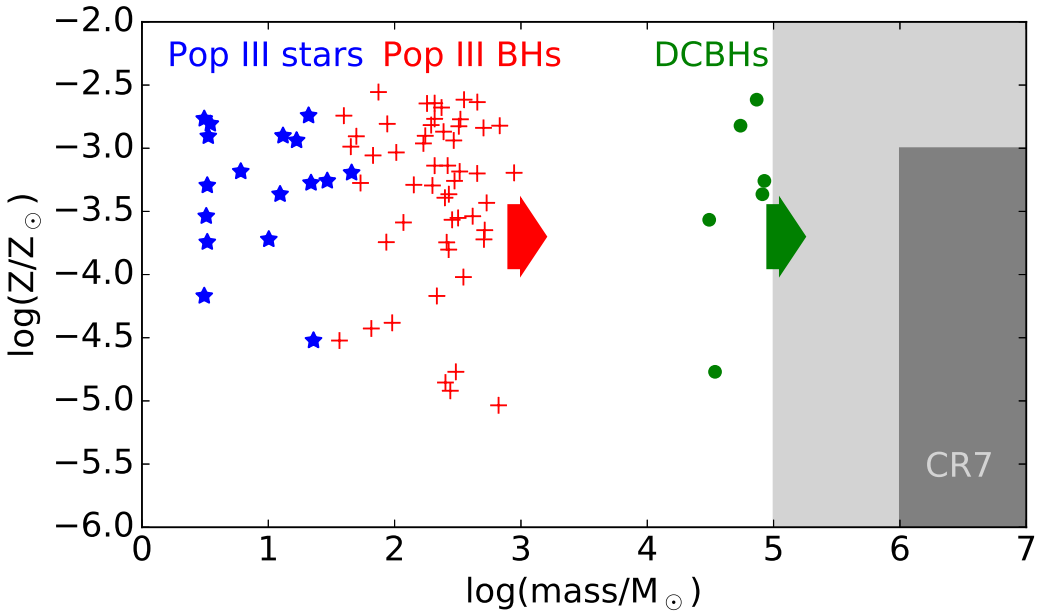


Figure 4.19: Comparison of the three different models, regarding the metallicity of clump A at $z = 6.6$ and the mass of the stellar population or of the BH, respectively. The grey shaded area indicates the region in which we expect the system to have the same observational signature as CR7. A Pop III stellar population has to have a mass of $\sim 10^7\,\mathrm{M}_\odot$ and a BH should be in the mass range $10^5 - 10^7\,\mathrm{M}_\odot$ to yield the observed line luminosities. The upper limit of the metallicity is set by Z_{limit}, in order not to violate the observed metal tax, but lower values are favourable. BHs of different origin can reside in sufficiently metal-poor halos, but, without including growth by accretion neither Pop III remnants nor DCBHs are sufficiently massive to account for the observations. The arrows for the BH populations symbolise possible mass growth by gas accretion; DCBHs require less mass growth by accretion to account for the spectral properties of CR7.

enough ($\lesssim 10^3\,\mathrm{M}_\odot$) to explain CR7. Only the DCBH scenario can explain the observed line luminosities. For the given mass constellation of the halos A, B, C, and D, the probability that clump A hosts Pop III stars is 37%, that it hosts a Pop III remnant BH is 98% and that it hosts a DCBH is 10%. These probabilities add up to over 100%, because due to merging of the progenitor halos, clump A can contain Pop III stars and BHs at the same time. Comparing these results to, e.g., Figure 4.14 shows that it is crucial to treat CR7 as three individual clumps, since the BH mass in clump A is generally lower than the most massive BH in all CR7 progenitor halos. The third most massive halo A is less affected by metal enrichment, but the Pop III remnant BHs do not merge to masses above $10^3\,\mathrm{M}_\odot$. Under these conditions and with the optimistic assumption of $J_{\mathrm{crit}} = 100$ we find that $\sim 90\%$ of the DCBHs reside in sufficiently metal-poor gas at $z = 6.6$.

Only $\sim 0.5\%$ of all the merger tree realisations yield the right mass constellation and out of these, only 9% produce results that are consistent with the observations of CR7. According to the Press-Schechter formalism, the number density of the host halo ($M_h = 1.2 \times 10^{12}\,\mathrm{M}_\odot$ at $z = 6.0$) is $n \lesssim 10^{-5}\,\mathrm{Mpc}^{-3}$. This yields an expected abundance of $n \lesssim 5 \times 10^{-9}\,\mathrm{Mpc}^{-3}$ for objects with the same constellation and observational signature as CR7. These estimates might even be lower with a smaller value of Z_{limit}, suggested by the non-detection of CIII]. The survey of Matthee et al. (2015) covers $5\,\mathrm{deg}^2$ in the redshift

range $z = 6.5 - 6.6$, which corresponds to an observed volume of $\sim 4.3 \times 10^6\,\mathrm{Mpc}^3$. The volume in the COSMOS fields, where CR7 was found, is $\sim 1.5 \times 10^5\,\mathrm{Mpc}^3$ and the expected number density of CR7-like sources is of the order $10^{-6} - 10^{-7}\,\mathrm{Mpc}^{-3}$ (Pallottini et al. 2015; Visbal et al. 2016). CR7, therefore, appears to have been a fortunate discovery, for the design of that survey. Larger surveys can confirm whether the mechanisms suggested in this paper occur with the expected probability or if we might have to adopt the model and assumptions to account for such rare sources.

There are other galaxies near the epoch of reionisation that have comparable observational signatures. Matthee et al. (2015) find more than ten Lyα emitter candidates potentially similar to CR7, which are worth being further investigated. Also the 'Himiko' galaxy is very bright and extended in Lyα, consists of three clumps of which one is very blue, and shows no sign of metals emission lines in the rest-frame UV (Ouchi et al. 2009, 2013; Zabl et al. 2015). However, no HeII emission has be confirmed for this galaxy, which disfavours Himiko as a potential host of Pop III stars.

4.5 Caveats

The main shortcoming of our semi-analytical approach is the lack of spatial information in the merger tree and hence the simplified treatment of metal enrichment. We check statistically if a halo is polluted, which yields reasonable results on average over many merger tree realisations. Once a halo is polluted we assume that the metals mix homogeneously with the gas and assign a single metallicity to the halo. However, metal-enriched winds may not mix effectively with the dense gas in minihalos and a large fraction of this gas might remain at low metallicity (Cen & Riquelme 2008; Pallottini et al. 2014; Ritter et al. 2015; Smith et al. 2015b). Three-dimensional hydrodynamical simulations are required to obtain a more reliable metal enrichment history and identify pockets of metal-free gas.

We do not explicitly account for accretion on to BHs. In our model, the DCBHs have a seed mass of $10^4 - 10^5\,\mathrm{M}_\odot$ and the Pop III remnant BHs only gain mass by mergers. The accretion of large masses of gas is necessary to account for the observed BH mass density at $z = 6$ (Tanaka & Haiman 2009). Pop III remnant BHs form in minihalos with shallow potential wells, which are unable to retain the photoionised heated gas (Johnson et al. 2007; Alvarez et al. 2009). Also the energy of a single supernova from a massive star is sufficient to clear the halo of accretable gas. The average time that Pop III remnant BHs can accrete mass in our model is $\sim 400\,\mathrm{Myr}$. To grow from a seed mass of $\sim 200\,\mathrm{M}_\odot$ to a final mass of $10^6\,\mathrm{M}_\odot$ at $z = 6.6$ they would have to accrete constantly at 80% the Eddington rate. Using cosmological zoom-in simulations, Jeon et al. (2012) study the assembly of the first galaxies under feedback from a central BH. They find that the accretion rate on to Pop III remnant BHs remains always below $10^{-6}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ for a radiative efficiency of $\epsilon = 0.1$. This is below the required accretion rate to account for the BH in CR7 at $z = 6.6$. Hence, the mass of the Pop III remnant BHs in our model might generally be higher, but they cannot grow sufficiently massive to account for the observations (but see Lupi et al. 2016).

In our analysis, we investigate $J_{\mathrm{crit}} = 100 - 400$ as the LW flux required to enable isothermal direct collapse of an atomic cooling halo. The probability of forming a DCBH is a steep function of this value (see section 5.3.3) and we do not find any DCBHs in our model at $J_{\mathrm{crit}} > 600$. The required flux varies from halo to halo (Shang et al. 2010; Latif et al. 2014) and also depends on the stellar spectrum. Pop III stars with effective temperatures of $\sim 10^5\,\mathrm{K}$ require $J_{\mathrm{crit}} \approx 10^5$, whereas subsequent stellar populations have lower effective temperatures and require a smaller J_{crit} (Bromm & Loeb 2003a; Shang et al. 2010; Sugimura et al. 2014; Agarwal & Khochfar 2015; Latif et al. 2015). In our model, a young population of

Pop II stars provides the required flux for isothermal direct collapse. We therefore expect $J_{\rm crit} < 1000$ (Shang et al. 2010; Agarwal et al. 2016b), but the exact value remains an open question.

A third way to form SMBHs in the early Universe is the fragmentation of gas in a low-metallicity halo into a dense nuclear cluster of low-mass stars and the subsequent build-up of a $\sim 10^3\,{\rm M}_\odot$ star via runaway mergers in the cluster (Portegies Zwart et al. 2004; Devecchi & Volonteri 2009; Lupi et al. 2014; Mapelli 2016). This scenario results in the creation of a $10^3\,{\rm M}_\odot$ BH and requires enough low-metallicity gas to form the cluster. Although this channel can generally produce as many BHs as Pop III BH remnants, it faces the same problem as the other scenarios: how can such a BH grow to sufficient masses without the host halo being polluted by metals? Given the formation criteria for a dense cluster we expect that it might yield results comparable to those of the Pop III remnant scenario and hence not be able to explain the properties of CR7.

4.6 Conclusion

We have explored the nature of CR7, a Lyα emitter at $z = 6.6$ (Matthee et al. 2015; Sobral et al. 2015), which has very strong Lyα and HeII signatures without any detection of metal lines in the spectrum. Using a semi-analytical merger tree model, we have investigated a variety of formation histories for CR7 and tested different scenarios for its origin.

Sobral et al. (2015) originally proposed that a recent Pop III starburst with a total stellar mass of $\sim 10^7\,{\rm M}_\odot$ or an accreting BH can account for the observational constraints. Based on our current understanding of Pop III star formation, we show that such a starburst seems not to be possible for several reasons. The mass of metal-free gas decreases with time since more and more halos are polluted by SNe. Hence, after peaking around $z \sim 15$, the cosmic SFR for Pop III stars declines sharply at lower redshifts. Moreover, very hot massive stars are required for the HeII line emission and the short lifetimes of these stars consequently requires a very recent burst, $\lesssim 1\,{\rm Myr}$ to be in agreement with the observed EW of HeII. Our model fails to reproduce the observed HeII line luminosity by about 10 orders of magnitude. Besides this fiducial model of primordial star formation, we also tested various models in which Pop III stars form in gas with a metallicity $Z < Z_{\rm crit}$, in which star formation is based on the merger history of the halo, and in which we adopt a different IMF. None of these models can explain CR7 as a primordial star cluster.

If CR7 hosts Pop III stars, the metal pollution by these first stars must be significantly less efficient than previously assumed and we need a mechanism to form $10^7\,{\rm M}_\odot$ of metal-free stars synchronised in one halo within $\sim 1\,{\rm Myr}$ (Visbal et al. 2016; Yajima & Khochfar 2017)

We also investigate the possibility that CR7 hosts an accreting BH. This scenario seems more appealing than Pop III stars because DCBHs form down to $z \sim 7.3$ and Pop III BH remnants can form down to $z = 6.6$. Several groups have shown that such BHs can reproduce the observational constraints (Pallottini et al. 2015; Agarwal et al. 2016a; Smith et al. 2016; Dijkstra et al. 2016; Smidt et al. 2016; Pacucci et al. 2016). However, they assume that the BH is embedded in metal-poor gas at $z = 6.6$, which is difficult to obtain with our model because when a Pop III star BH forms, it is very likely that a SN also enriches the halo with metals. Furthermore, we assume that Pop III star BHs mainly grow via mergers and consequently the already low possibility of finding a BH in a metal-poor halo shrinks with every merger with another halo, which could be enriched with metals. Finally, we only find Pop III BH remnants with masses $< 10^3\,{\rm M}_\odot$ in low-metallicity halos with gas masses of $\sim 10^6\,{\rm M}_\odot$ at $z = 6.6$.

The most promising explanation for CR7 is an accreting DCBH. By construction, they form in metal-

free halos and may remain in metal-poor gas until $z = 6.6$. Under the optimistic condition of $J_{\rm crit} = 100$, we find that a small fraction of systems can host DCBHs that are able to reproduce the line luminosities of CR7 without violating the upper limit of the metal line luminosities. Our findings are supported by other observational constraints. Only an accreting BH can account for the spatial extension of the Lyα emitting region and for the velocity offset between the Lyα and HeII line peaks, because this velocity offset requires a source lifetime of $> 10\,$Myr (Smith et al. 2016). A stellar source in CR7 would however require a recent burst of $\lesssim 1\,$Myr, and a very low metallicity of $< 10^{-7}$ to account for the large EW of HeII (Schaerer 2003; Raiter et al. 2010). We note that we are not able to confirm, if an accreting BH is able to produce the EW of HeII, which is an important question to address in the future.

There are other observations that might help to better understand the nature of CR7 and distinguish between the different formation scenarios. The dwarf galaxy I ZW 18 at a distance of 18 Mpc also shows strong HeII emission and has a very low metallicity (Kehrig et al. 2015, 2016). It might host either metal-free or Wolf-Rayet stars, which could account for these observational features. Smaller halos, such as I ZW 18, might remain unpolluted by their progenitors down to lower redshift and e.g. photoionisation heating might prevent star formation for a long time (Visbal et al. 2016, 2017). The study of the assembly history of such systems might reveal interesting new insights that can help to understand more massive low-metallicity counterparts at higher redshift.

Another observational signature of whether CR7 hosts a BH or a Pop III stellar population is the X-ray flux. The expected luminosity from X-ray binaries is $L_{\rm x} \approx 10^{40}\,{\rm erg\,s^{-1}}$ for a SFR of $1\,{\rm M_\odot\,yr^{-1}}$ (Glover & Brand 2003; Grimm et al. 2003; Mineo et al. 2012), which is at least two orders of magnitude lower than the emission from a BH of $\sim 10^6\,{\rm M_\odot}$. CR7 was not detected as a point source in the Chandra COSMOS Survey. Therefore its X-ray luminosity in the energy range $0.2 - 10\,$keV is less than $10^{44}\,{\rm erg\,s^{-1}}$ (Elvis et al. 2009). For a BH with a mass of $10^6\,{\rm M_\odot}$, accreting at 40% of the Eddington rate, and a bolometric correction based on Marconi et al. (2004) we find $L_{\rm x} \lesssim 2 \times 10^{42}\,{\rm erg\,s^{-1}}$ in the observer rest frame $0.2 - 10\,$keV band. This translates into a flux of $F_{\rm x} \lesssim 8.2 \times 10^{-18}\,{\rm erg\,cm^{-2}\,s^{-1}}$ or $\sim 1.4 \times 10^{-7}$ counts per second with Chandra in the $0.2 - 10\,$keV band. Hence, integration times of several years might be required, which seems infeasible until the next generation of X-ray telescopes, such as ATHENA (Nandra et al. 2013), could approach this problem with a significantly higher effective collecting area. Alternatively, if many more DCBH candidates similar to CR7 are found, one could stack deep X-ray observations of these objects and attempt to detect their X-ray emission using the stacked image. In this case, a successful detection would imply that some significant fraction of the candidates are indeed DCBHs, although it would not allow us to say with certainty that any particular candidate is a DCBH. With this method, however, there is no X-ray detected in the stack of Lyα emitters in the Chandra COSMOS Legacy survey (Civano et al. 2016), down to a total depth of 1.38 Ms corresponding to a luminosity of $10^{43}\,{\rm erg\,s^{-1}}$ (at the redshift of the sources) in the $0.5 - 2\,$keV band (Civano, private communication). Another method to identify DCBHs was recently proposed by Pacucci et al. (2016). They identify two objects with a robust X-ray detection found in the CANDELS/GOODS-S survey with a steep spectral slope in the infrared and an extremely red spectra. They argue, based on their assumptions, that this can be explained by either a DCBH or an atypically high SFR of $> 1000\,{\rm M_\odot\,yr^{-1}}$. Another explanation of this spectra could be a dusty AGN without the constraint of being a DCBH.

More crucially, validation of the general scenario of a low metallicity in CR7 requires further deep spectroscopy to estimate the level of metallicity and the source of the hard spectrum in CR7. The focus should hence be to obtain deeper spectra from UV to near-IR to improve the limits on the metal line emission and to further constrain the different formation histories. Beyond current instruments, such as MOSFIRE and X-SHOOTER, NIRSpec on *JWST* will be the ideal instrument to probe sources such as

CR7 (Pacucci et al. 2016, 2017b), which provide a tantalising preview into the initial episodes of star and BH formation, which will fully be elucidated in the near future with an array of next-generation observational facilities.

4.6.1 New observations of CR7

After the publication of our paper, Bowler et al. (2017) presented new observations of CR7 in the optical, near-infrared, and mid-infrared, which affect the possible interpretation of CR7. In contrast to the earlier observation by Sobral et al. (2015), they find an oxygen emission line, a bluer colour, and a lower EW of HeII by using Spitzer/IRAC data, which extends up to one order of magnitude deeper than previous observations. They also find no evidence that components B and C are particularly red as claimed by Sobral et al. (2015), and derive considerably lower stellar masses ($< 10^{10}\,M_\odot$) for those clumps than derived by e.g. Agarwal et al. (2016a). They assign this difference to high uncertainties on the photometric SED fitting and the degeneracy between the stellar age and dust attenuation (Curtis-Lake et al. 2013). Bowler et al. (2017) claim that the [3.6]-[4.5] colour of CR7 cannot be explained by a primordial stellar population or a DCBH, but they propose a low-mass, narrow-line AGN or a young low-metallicity star-burst as alternative explanations.

However, (Bowler et al. 2017) base their interpretation on only one proposed DCBH spectra by Agarwal et al. (2016a) and did not consider different models. Later studies show that also a DCBH in the centre of CR7 could be in agreement with the recent observation and reproduce the [3.6]-[4.5] colour (Agarwal et al. 2017a; Pacucci et al. 2017b). Moreover, the interpretation of CR7 as a low-mass, narrow-line AGN or a DCBH are not exclusive. The majority of studies agrees that CR7 does not host a massive population of metal-free stars and the new detection of an oxygen emission line seems to confirm this interpretation.

Our study was mainly based on the original observation by Sobral et al. (2015), but our model is also in agreement with the recent observations presented by Bowler et al. (2017):

- their derived gas metallicity of $0.005\,Z_\odot$ is within the limits of our derived metal tax of $< 0.01\,Z_\odot$ for clump A.
- we conclude that CR7 cannot host the initially proposed amount of massive Pop III stars.
- we demonstrate that a DCBH could form in the progenitors of CR7 and Agarwal et al. (2017a); Pacucci et al. (2017b) show that such a BH is in agreement with the recent observations.

Upcoming HST grism spectroscopy of CR7 will provide a more reliable estimate of the line flux and therefore help to further constraint the nature of CR7.

Recently, Shibuya et al. (2017) present a reanalysis of the VLT/X-SHOOTER spectrum of CR7 with three different, independent methods. They confirm a strong Lyα emission line, but they do not detect a HeII emission line at 1640 Å. They obtain an upper limit of the HeII luminosity of CR7 of $< 10^{43}\,\mathrm{erg\,s^{-1}}$, which is still in agreement with the assumptions in our model.

5 Statistical predictions for the first black holes

This chapter is based on the invited book chapter "Statistical predictions for the first black holes" that will be published in "Formation of the First Black Holes" (Hartwig 2017).

5.1 Observational constraints

The optically selected quasar sample in the SDSS survey yields a lower limit for the abundance of SMBHs with $> 10^9\,\mathrm{M}_\odot$ at $z = 6$ of about $1/\mathrm{cGpc}^3$ (Fan et al. 2006b; Venemans et al. 2013a; De Rosa et al. 2014). As we will see below, most current models predict sufficiently many BH seeds to explain the number of BHs at high redshift, but not all models can reproduce the high masses of some SMBHs at $z > 6$. The direct collapse (DC) scenario, which yields massive BH seeds (see Section 5.3), cannot account for the number of less massive BHs at lower redshift under conservative parameter assumptions (Greene 2012; Reines et al. 2013). Moreover, not all halos are able to fuel gas to their centre at high rates over several hundred million years to grow the BH seeds to their final masses (see Chapter 6 for the feedback limited growth rates).

The discovery of the Lyman-α emitter CR7 at $z = 6.6$ (Sobral et al. 2015; Bowler et al. 2017) might yield another constraint on the number density of direct collapse BHs (DCBHs). The strong Lyman-α and HeII emission with large equivalent widths, the steep UV slope, and the absence of metal lines in the original observation by Sobral et al. (2015) suggests a massive population of metal-free stars with a total stellar mass of $\sim 10^7\,\mathrm{M}_\odot$. However, several studies demonstrate that the observational signature is more likely to emerge from an accreting DCBH in the centre of CR7 (Pallottini et al. 2015; Agarwal et al. 2016a; Hartwig et al. 2016b; Smidt et al. 2016; Smith et al. 2016; Pacucci et al. 2017b, Chapter 4) than by a population of metal-free stars (but see Visbal et al. 2016; Bowler et al. 2017). The expected number density of CR7-like sources is of the order $10^{-6} - 10^{-7}\,\mathrm{Mpc}^{-3}$ (Pallottini et al. 2015; Hartwig et al. 2016b; Visbal et al. 2016), which yields an additional observational constraint on the abundance of the first BHs.

In the next section we will see that stellar mass BHs, as the remnants of the first stars (see Section 2.1.), are sufficiently abundant and it is not their number density that poses a problem, but rather the necessary mass accretion over more than six orders of magnitude and the associated feedback that pose the biggest challenge (see Section 2.5). We will hence mostly focus on DCBHs, which are predicted to be born with a higher seed mass, but which form only under very peculiar conditions and it is worth to investigate the probability of these environmental conditions of DC to assess the likelihood and the halo occupation fraction for DCBHs.

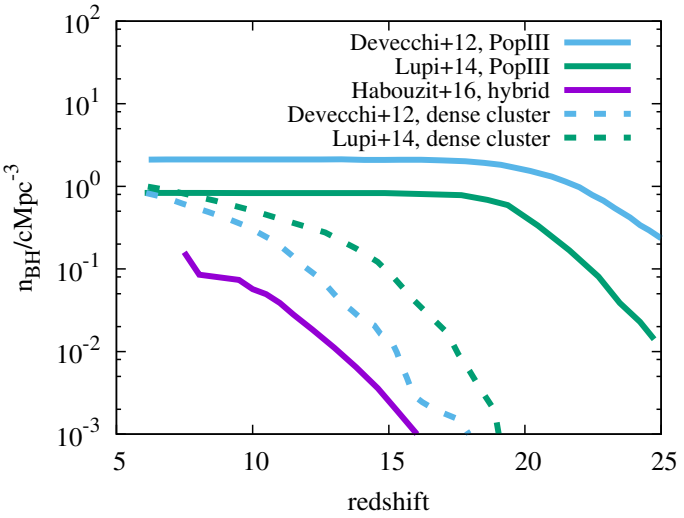


Figure 5.1: Comparison of the comoving number density of stellar mass BHs as a function of redshift based on models by Devecchi et al. (2012), Lupi et al. (2014), and Habouzit et al. (2016c). To convert from the given mass density to number density we assume an average mass of $100\,M_\odot$ for Pop III remnant BHs and of $700\,M_\odot$ (Devecchi et al. 2012) and $10^3\,M_\odot$ (Lupi et al. 2014) for BHs formed via runaway collisions in dense stellar clusters. The number of black holes in Habouzit et al. (2016c) is a lower limit, because their simulation does not resolve minihalos as potential formation sites. The values by Lupi et al. (2014) are an upper limit with $\xi = 3$, quantifying the contraction of the cluster and the redistribution of gas. The number density of DCBHs is always below $10^{-5}\,\mathrm{cMpc}^{-3}$ (Habouzit et al. 2016c), but see Section 5.3 for a more detailed discussion.

5.2 Probability for stellar mass seed black holes

Stellar mass seed BHs of $100 - 1000\,M_\odot$ can either form as remnants of the first metal-free stars (see Section 2.1) or as a consequence of runaway stellar collisions in metal-poor nuclear clusters (see Section 2.2). Devecchi et al. (2012) estimate the seed density for these two scenarios with a semi-analytical model of structure formation. Their model confirms recent observations that all halos with masses above $10^{11}\,M_\odot$ host a BH and predicts a BH occupation fraction of 10% for halos less massive than $10^9\,M_\odot$. Habouzit et al. (2016c) demonstrate that the number density of stellar mass seed BHs is $\gtrsim 10^{-2}\,\mathrm{cMpc}^{-3}$ at $z < 10$, significantly higher than the number density of DCBH seeds in their numerical simulations. Their model is a hybrid between Pop III remnant BHs and those formed in dense stellar clusters, because to distinguish both models one would need to resolve much higher gas densities. Moreover, they allow seed BHs to form in slightly metal enriched halos, which includes potential dense stellar clusters. We compare the predicted number densities of stellar mass BHs in Figure 5.1. The number density of BH seeds formed by runaway stellar collisions in nuclear clusters is slightly below the number density of Pop III remnant BHs, but both are significantly above the expected number density of DCBHs (see section 5.3). However, since BH formation via runaway collisions in dense stellar clusters can also occur in slightly metal-enriched gas, this formation channel dominates the mass density of BH seeds at later times (Devecchi et al. 2012). In their model, BHs form via runaway stellar collision only in the narrow metallicity range $10^{-3.5} < Z/Z_\odot < 10^{-3}$, but other studies and recent simulations show a wider range of conditions for this BH formation channel (Mapelli & Bressan 2013; Mapelli 2016; Sakurai et al. 2017).

Hence, the bottleneck to form SMBHs at high redshift from stellar mass BHs is not the abundance of their seeds but rather the required high accretion rate over a sufficiently long time and the associated feedback. In the following section we will hence focus on DCBHs and quantify the probability to form such a massive BH seed.

5.3 Probability for the direct collapse scenario

The idea of forming a massive BH as the result of monolithic collapse of a proto-galactic gas cloud was first proposed by Rees (1984). Since then, various flavours of this formation scenario have emerged with different processes triggering the collapse of the cloud: galaxy mergers (Mayer et al. 2010), gas instabilities (Begelman et al. 2006; Choi et al. 2013), or contraction due to radiative cooling by atomic hydrogen (Bromm & Loeb 2003b).

The main requirement for these scenarios is a high mass infall rate so that the accretion timescale of the infalling gas is shorter than the Kelvin-Helmholtz timescale of the protostar (Hosokawa et al. 2013; Sakurai et al. 2016). Under this condition, the stellar radius monotonically increases with mass and enough gas can be accreted before the star reaches the main sequence of hydrogen burning. In this section we mostly focus on the DC scenario as a massive BH seed forming as the result of the thermodynamically triggered collapse of a pristine gas cloud with high mass infall rates and suppressed fragmentation. We will also comment on alternatives and weaken some of the assumptions, such as the formation of massive seed BHs in metal-poor gas or under the influence of dust cooling.

Pristine halos that cool mainly by atomic hydrogen are expected to have high mass infall rates, because they isothermally collapse at $T \approx 10^4\,\mathrm{K}$ and the inefficient atomic cooling prevents fragmentation of the gas and allows for high mass infall rates ($\dot{M}_{\rm in} \propto T^{3/2}$). A photodissociating flux from a nearby star-forming galaxy is required to keep the gas atomic. Phrased differently, a pair of halos has to be well synchronised in space and time to trigger DC (Visbal et al. 2014a; Regan et al. 2017): they have to be close enough to generate a sufficiently high photodissociating flux, but they have to be sufficiently far separated to prevent metal enrichment and tidal stripping (Chon et al. 2016; Regan et al. 2017). Regarding the temporal synchronisation, there is only a short window after the initial starburst in which the LW flux is high enough, but the metal-enriched, SN-driven winds have not yet reached the neighbouring halo (Dijkstra et al. 2014; Habouzit et al. 2016c; Hartwig et al. 2016b; Agarwal et al. 2016a).

5.3.1 Atomic cooling halos

The cooling function of hydrogen rises steeply around $\sim 8000\,\mathrm{K}$ and we follow the convention to define halos with $T_{\rm vir} \gtrsim 10^4\,\mathrm{K}$ as atomic cooling halos. The gas in an atomic cooling halo collapses isothermally at around $10^4\,\mathrm{K}$, which results in a high mass infall rate towards the centre (Latif & Volonteri 2015). The virial temperature of a halo of mass M at redshift z is given by

$$T_{\rm vir} = 2 \times 10^4 \left(\frac{M}{10^8\,\mathrm{M}_\odot}\right)^{2/3} \left(\frac{1+z}{10}\right)\,\mathrm{K} \tag{5.1}$$

for atomic gas of primordial composition (Loeb 2010) and the critical mass for an atomic cooling halo as a function of redshift is hence given by

$$M_{\rm ac} = 4 \times 10^7\,\mathrm{M}_\odot \left(\frac{1+z}{10}\right)^{-3/2}. \tag{5.2}$$

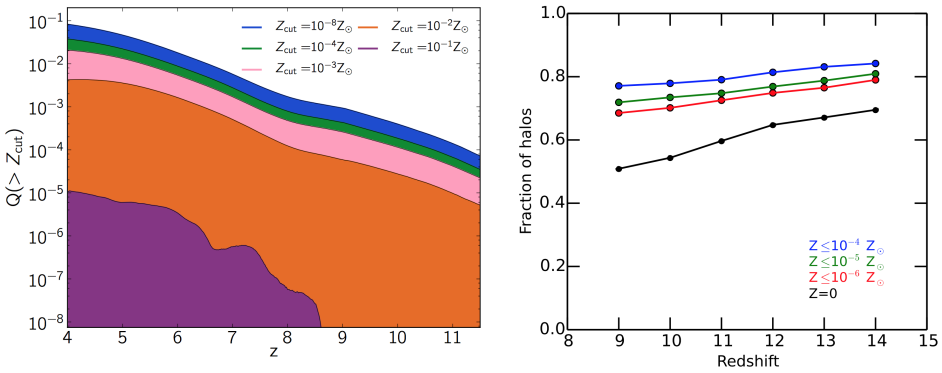


Figure 5.2: Left: Redshift evolution of the metal volume filling factor for different values of the metallicity cut, $Z_{\rm cut}$, adapted from Pallottini et al. (2014). Right: Fraction of halos in the mass range $2 \times 10^7 - 10^8\,{\rm M}_\odot$ with metallicities below a given value as a function of redshift. The fraction of halos with $Z \leq 10^{-5} Z_\odot$ is about 1.5 times higher than the metal-free halos, adapted from Latif et al. (2016b). At the redshifts of interest, more than half of the atomic cooling halos are still pristine.

The number density of atomic cooling halos can be calculated via Press-Schechter theory (Press & Schechter 1974), including the corrections by Sheth et al. (2001):

$$n_{\rm ac} = \int_{M_{\rm ac}}^{\infty} \sqrt{\frac{2}{\pi}} \frac{\rho_m}{M} \frac{-{\rm d}(\ln \sigma)}{{\rm d}M} \nu_c \exp(-\nu_c^2/2)\,{\rm d}M, \tag{5.3}$$

with $\nu_c = \delta_{\rm crit}(z)/\sigma(M)$, where $\delta_{\rm crit}(z)$ is the critical overdensity for collapse, $\sigma(M)$ is the variance of the density power spectrum at mass M, and ρ_m is the mean matter density of the Universe. However, this constrain is not very restrictive because any host galaxy of a SMBH once passed this mass threshold of atomic cooling and was hence a potential candidate for DC. In the next section we will discuss more constraining criteria such as the photodissociating radiation and the absence of metals.

5.3.2 Pristine gas

In order to prevent fragmentation and maintain a high temperature and hence mass infall rate of the collapsing gas cloud, the halo has to be pristine (but see Omukai et al. 2008; Latif et al. 2016b; Agarwal et al. 2017b). Metals are a much more efficient coolant than hydrogen and the higher cooling rate makes gas more susceptible to fragmentation and prevents DC via ordinary star formation. Halos can either be enriched internally by Pop III stars or externally by SN-driven metal-enriched winds from nearby star-forming halos and external enrichment seems to be more efficient in suppressing the formation of DCBHs (Dijkstra et al. 2014). Figure 5.2 illustrates how the fraction of pristine gas and halos evolves with redshift. At $z = 4$ only $\lesssim 10\%$ of the cosmological volume is enriched with metals (see also Johnson et al. 2013a). However, this volume filling factor also accounts for voids without star formation. When we consider atomic cooling halos as the halos of interest for DC, Latif et al. (2016b) find that more than half of them are still pristine until $z \approx 9$, but this is an upper limit since they do not include Pop III enrichment nor took clustering into account. These candidate halos have not been externally enriched and also star formation in these halos was suppressed by e.g. a photodissociating background (Machacek et al. 2001). Considering this criterion for DC independently does not seem very constraining, because every second target halo is still pristine at the redshift of interest. However, we discuss in the next section that DC generally requires a nearby star-forming galaxy, which drastically reduces the probability of not being polluted by metals.

5.3.3 Photodissociating radiation

A photodissociating LW flux is another crucial ingredient for DC. The concept of a critical flux threshold, $J_{\rm crit}$, above which a DCBH can form is well established in the literature, but there is no consensus on its value and most likely it is a distribution of critical values depending on e.g. the spectral shape of the stellar population (Shang et al. 2010; Sugimura et al. 2014), the anisotropy of the illuminating flux (Regan et al. 2014b, 2016), the treatment of H_2 self-shielding (Wolcott-Green et al. 2011; Hartwig et al. 2015c), the details of the chemistry model (Glover 2015a,b), and the star formation history (Agarwal et al. 2016b). Although there is no universal value for $J_{\rm crit}$, it is convenient to quantify the importance of different effects by this value. Once we have chosen a threshold value for DC, we can quantify the probability to find a pristine atomic cooling halo that is exposed to a LW flux above this threshold. In this section we will review and summarise different models to address this question and to derive the number density of DCBH formation sites. We use the standard notation for the LW flux and express its intensity in units of $J_{21} = 10^{-21} \mathrm{erg\,s^{-1}\,cm^{-2}\,sr^{-1}\,Hz^{-1}}$.

The build-up of the cosmological LW background is self-regulated (Johnson et al. 2008), because a high flux suppresses star formation and hence the production of further LW photons. The mean value in the redshift range $z = 10 - 20$ is of the order $0.1\,J_{21}$ (Johnson et al. 2013a), but possible formation sites for DC are hidden in the high end tail of the $J_{\rm LW}$ distribution.

Ahn et al. (2009) present the first self-consistent radiative transfer calculations of the inhomogeneous LW background. By using high-resolution N-body simulations and an equivalent grey opacity for the radiative transfer they show that the LW background is inhomogeneous and correlates with the clustering of large-scale structure. In contrast to ionising photons in the high redshift Universe, LW photons can travel cosmological distances due to the small prereionisation H_2 abundance of $\sim 10^{-6}$ and hence an optical depth of $\tau_{\rm LW} < 1$ (Shapiro et al. 1994). However, due to the expansion of the Universe they get redshifted out of the LW bands and can no longer contribute to H_2 photodissociation once their energy is below $\sim 11.5\,\mathrm{eV}$ (see also Haiman et al. (2000) for a discussion of radiative transfer of LW at high redshift on cosmological scales).

Ahn et al. (2009) define the "LW horizon" as the maximum comoving distance from a source that an H_2 dissociating photon can reach and determine this radius to be $\sim 100\alpha\,\mathrm{cMpc}$, with

$$\alpha = \left(\frac{h}{0.7}\right)^{-1} \left(\frac{\Omega_m}{0.27}\right)^{-1/2} \left(\frac{1+z_s}{21}\right)^{-1/2}, \tag{5.4}$$

where z_s is the redshift of the source. Within this LW horizon the cosmic expansion reduces the effective LW flux by

$$f_{\rm mod} = \begin{cases} 1.7\exp[-(r_{\rm cMpc}/116.29\alpha)^{0.68}] - 0.7 & \text{for} \quad r_{\rm cMpc}/\alpha \leq 100 \\ 0 \quad \text{for} \quad r_{\rm cMpc}/\alpha > 100, \end{cases} \tag{5.5}$$

where $r_{\rm cMpc}$ is the distance from the source in comoving Mpc.

Due to the long mean free path of LW photons, many sources contribute to the local flux and Dijkstra et al. (2008) find that $> 99\%$ of all halos are exposed to a LW flux within a factor of two of the mean value. They calculate the LW background based on the clustering of halos and account for the non-linear Eulerian bias of the two-point correlation function (Iliev et al. 2003), which fits the correlation function derived in N-body simulations. Accounting for this non-linear bias is crucial since we are especially interested in close pairs of halos, which are the most promising candidates for DC (Dijkstra et al. 2008; Visbal et al. 2014a; Regan et al. 2017). The differential probability to find a halo of mass M at redshift

z in the distance r to an atomic cooling halo is given by

$$\frac{\mathrm{d}^2 P_1(M,r,z)}{\mathrm{d}M\,\mathrm{d}r} = 4\pi r^2 (1+z)^3 [1+\xi(M_{\rm ac},M,r,z)] \frac{\mathrm{d}n}{\mathrm{d}M}, \tag{5.6}$$

where $\xi(M_{\rm ac},M,r,z)$ is the two-point correlation function, which yields the excess probability to find two halos of masses $M_{\rm ac}$ and M at distance r, and $\mathrm{d}n/\mathrm{d}M$ denotes the Press-Schechter mass function (Press & Schechter 1974; Sheth et al. 2001). Dijkstra et al. (2008) model the LW luminosity of each halo by randomly sampling from a log-normal distribution of UV luminosities per unit star formation rate, $P_2/\mathrm{d}\log(L_{\rm LW})$. The probability distribution function (PDF) of a given atomic cooling halo to be exposed to a LW flux $J_{\rm LW}$ is hence given by

$$\frac{\mathrm{d}P}{\mathrm{d}\log(J_{\rm LW})} = \int_{M_{\rm min}}^{\infty} \mathrm{d}M \int_{r_{\rm min}}^{\infty} \mathrm{d}r \frac{\mathrm{d}^2 P_1}{\mathrm{d}M\,\mathrm{d}r} \frac{P_2}{\mathrm{d}\log(L_{\rm LW})}. \tag{5.7}$$

Dijkstra et al. (2008, 2014) and Inayoshi & Tanaka (2015b) assume $M_{\rm min} = M_{\rm ac}$, but the choice of the minimal radius $r_{\rm min}$ is debated, because its value is crucial: small values will populate the high-end tail of the PDF because the flux scales with the inverse square of the distance, but a too small value is physically prohibited to prevent external enrichment of the atomic cooling by the star forming halo and to prevent tidal stripping due to dynamical effects (Chon et al. 2016; Regan et al. 2017). See Dijkstra et al. (2008) for a discussion of how the choice of $r_{\rm min}$ (and other model parameters) affects the PDF of $J_{\rm LW}$. The high-end tail of the PDF is dominated by close pairs of halos and the illuminating flux, as seen by the atomic cooling halo, should hence be highly anisotropic. Moreover, the relative velocity of this close pair of halos is relevant for the escape fraction of LW photons out of the star-forming galaxy, because photons might be shifted into the LW bands in the far-field (Schauer et al. 2015, 2017).

In Dijkstra et al. (2014) they improve their previous model by explicitly accounting for pollution by metal-enriched winds from the star forming galaxy. They approximate the radius of the metal enriched region around a star-forming halo by

$$r_{\rm metals} = 0.3\,\mathrm{kpc} \left(\frac{M_*}{10^5\,\mathrm{M}_\odot}\right)^{1/5} \left(\frac{n}{1\,\mathrm{cm}^{-3}}\right)^{-1/5} \left(\frac{t}{\mathrm{Myr}}\right)^{2/5}, \tag{5.8}$$

where M_* is the stellar mass and n the density of the gas into which the SN-driven metal bubble expands. Based on this metal polluted radius, Dijkstra et al. (2014) correct the PDF of $J_{\rm LW}$ (Eq. 5.7) by multiplying it with $\Theta(r - r_{\rm metals})$ to exclude atomic cooling halos as potential candidates for DC if they are too close and already polluted. This approach was also used in Habouzit et al. (2016a); Hartwig et al. (2016b) to estimate the likelihood of finding DCBH candidates.

A comparison of the literature on the PDF of $J_{\rm LW}$ can be seen in Figure 5.3. All studies broadly agree on the shape and mean value, which increases with decreasing redshift and all studies predict potential DCBH hosts, which are exposed to a flux above $J_{\rm crit} \approx 10^2 - 10^3$. The final density of DCBHs depends crucially on the high-end slope, which differs amongst the authors. While the recent simulation-based results of Agarwal et al. (2012); Chon et al. (2016); Johnson et al. (2013a) agree on the high-end slope, the models by Dijkstra et al. (2008); Ahn et al. (2009) yield a steeper slope. Inayoshi & Tanaka (2015b), who follow the approach by Dijkstra et al. (2008, 2014), find an even steeper slope of $\mathrm{d}p/\mathrm{d}\log(J_{\rm LW}) \propto J_{\rm LW}^{-5}$ in the range $10^3 \lesssim J_{\rm LW} \lesssim 10^4$ and they assign the discrepancy to a different choice in $r_{\rm min}$. Besides quantitative differences, this comparison illustrates the importance of constraining $J_{\rm crit}$ and its dependence on the physics and environment, because a small change in $J_{\rm crit}$ can result in an

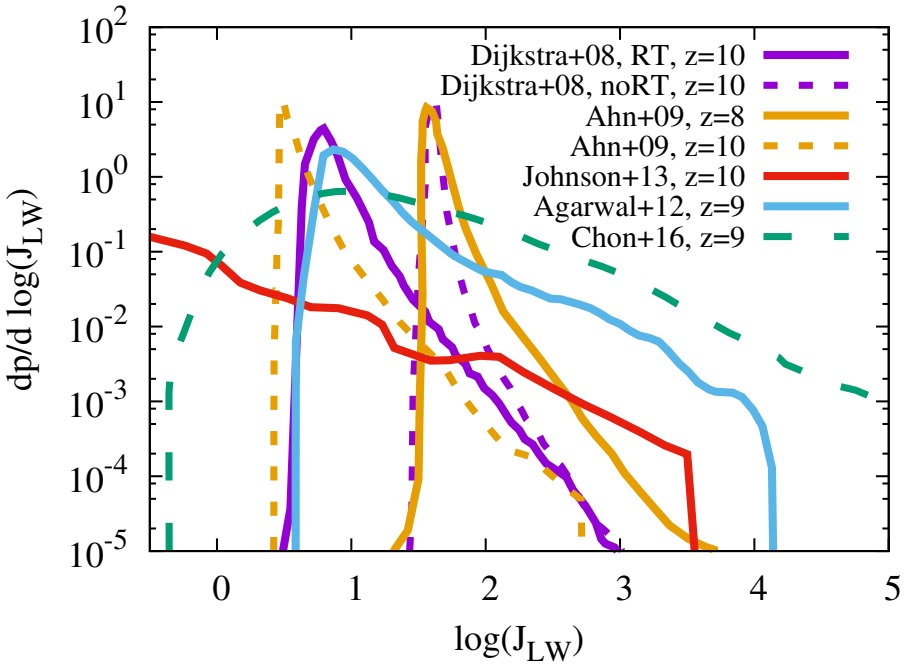


Figure 5.3: Comparison of the PDF of the LW flux. The data is based on Dijkstra et al. (2008); Ahn et al. (2009); Johnson et al. (2013a); Agarwal et al. (2012); Chon et al. (2016) (purple, orange, red, blue, and green, respectively). Note the slightly different redshifts and how the peak of this distribution shift towards higher fluxes with decreasing redshift due to the increasing cosmic star formation rate. All studies determine the LW flux based on the clustering of halos. Chon et al. (2016); Agarwal et al. (2012) derive the distribution for pristine halos, Ahn et al. (2009) show the PDF for all grid cells inside their computational domain, Johnson et al. (2013a) for primordial gas with $n \geq 1\,\mathrm{cm}^{-3}$, and Dijkstra et al. (2008) provide the flux as seen by atomic cooling halos of mass $M = 4 \times 10^7\,\mathrm{M}_\odot$. In their fiducial model (noRT) they assume the IGM to be transparent to LW photons and in their second model with radiative transfer they account for intergalactic H_2, which reduces the contribution from far away sources and hence the mean and minimum value of the PDF. However, the distribution at high J_{LW} is not affected by intergalactic absorption, because it is dominated by close pairs of halos.

immense change of the probability to form DCBHs.

The sharp cut-off at low J_{LW} is due to the very long mean free path of LW photons and hence the mean, median, and minimum value of this distribution are very close. The less sharp cut-off at low J_{LW} in Johnson et al. (2013a); Chon et al. (2016) is caused by the smaller box size and the hence smaller number of sources contributing to the local flux. In the other studies, the box sizes are larger or they implicitly account for the LW background radiation by the cosmic star formation history. The LW background depends also strongly on the environment. The "local" background is much higher in the biased region of a $z = 6$ quasar, compared to a random halo or even a synchronised halo pair (Valiante et al. 2016). Such high σ-peaks in the cosmic density field therefore naturally ease the formation of massive BH seeds via the DC scenario.

Another crucial question is for how long an atomic cooling halo has to be exposed to a photodissociating flux above J_{crit} for it to collapse isothermally. The collapse time of the gas core of a typical atomic cooling halo is $t_{\mathrm{coll}} \approx 10\,\mathrm{Myr}$ at $z = 10$ (Visbal et al. 2014a) with $t_{\mathrm{coll}} \propto (1+z)^{-3/2}$. Other authors require the conditions for DC to be fulfilled over at least the freefall time, which is of the order 100 Myr at the redshifts of interest and yields more conservative estimates (Dijkstra et al. 2014). Habouzit et al. (2016c) analyse the impact of the choice of the collapse time on the probability to find DCBH candidates

and find a significant difference: at $z = 7.3$ they find 17 (3) DC formation sites in their computational volume for an assumed collapse time of 10 Myr (150 Myr).

Chon et al. (2016) demonstrate that the requirement of having a flux of at least $J_{\rm crit}$ at the time of virialisation of the atomic cooling halos is too strict for DC. They reduce the LW intensity artificially to 2% of its original value at virialisation and still find the halo to collapse isothermally. They argue that the local flux increases over time due to the approaching halos by an order of magnitude and that lower intensities than previously thought are sufficient to induce DC (see also Fernandez et al. 2014).

5.3.4 Tidal and ram pressure stripping

The conditions for DC are generally met in close halo pairs, where one halo provides the photodissociating radiation and the other one is still pristine and about to pass the mass threshold of atomic hydrogen cooling. However, dynamical effects such as tidal forces and ram pressure stripping can no longer be neglected for the thermal evolution of the gas collapse (Chon et al. 2016). Tidal stripping occurs in close pairs of galaxies, where the tidal force exerted by the more massive galaxy results in an effective transfer of gas and stars from the lower mass to the more massive galaxy. Ram pressure stripping can expel gas out of the gravitational potential of a galaxy that moves with a sufficiently high relative velocity with respect to the ambient medium.

To study these dynamical effect of the environment, Chon et al. (2016) select DC candidate halos in a semi-analytical model and perform 3D hydrodynamical follow-up simulations of these pristine atomic cooling halos that are illuminated by a sufficiently high LW flux. Out of 42 selected halos they find only two successful candidates and show that tidal forces and ram pressure stripping prevent the gas collapse in the other cases.

They also point out that while minor mergers tend to disrupt the gas distribution and prevent collapse, major mergers rather help in triggering the collapse due to the induced gravitational instability at the halo centre, caused by the asymmetric mass distribution on large scales (see also Mayer et al. 2010, 2015). In summary, they find that only 5% of the atomic cooling halos that are pristine and illuminated with $J_{\rm LW} \geq J_{\rm crit}$ eventually collapse to form a DCBH.

5.3.5 Density of direct collapse seed black holes

In this section, we compare different estimates for the probability to form DCBHs, their expected number density, and the resulting occupation fractions. As we have seen before, the PDF to find pristine atomic cooling halos exposed to a certain LW flux (Eq. 5.7) can be calculated analytically (Dijkstra et al. 2008, 2014; Inayoshi & Tanaka 2015b), by post-processing cosmological N-body simulations (Ahn et al. 2009; Agarwal et al. 2012), or by performing cosmological simulations (Johnson et al. 2013a; Agarwal et al. 2014; Habouzit et al. 2016a; Chon et al. 2016). The number density of DCBHs can then be approximated by

$$n_{\rm DCBH}(z) = \int_{M_{\rm ac}}^{\infty} {\rm d}M \frac{{\rm d}n}{{\rm d}M} P_{\rm DCBH}(\geq J_{\rm crit}, z). \tag{5.9}$$

However, this approach assumes that all pristine atomic cooling halos that are exposed to $> J_{\rm crit}$ collapse to a DCBH, but the analytical and semi-analytical models do not account for hydrodynamical effects such as tidal or ram pressure stripping. Chon et al. (2016) show that these effects can reduce the number number density of DCBHs by more than one order of magnitude. A comparison of the number density of DCBH as a function of redshift by different authors can be seen in Figure 5.4. The number density increases with decreasing redshift and with decreasing $J_{\rm crit}$, but even for the same value of $J_{\rm crit}$ the

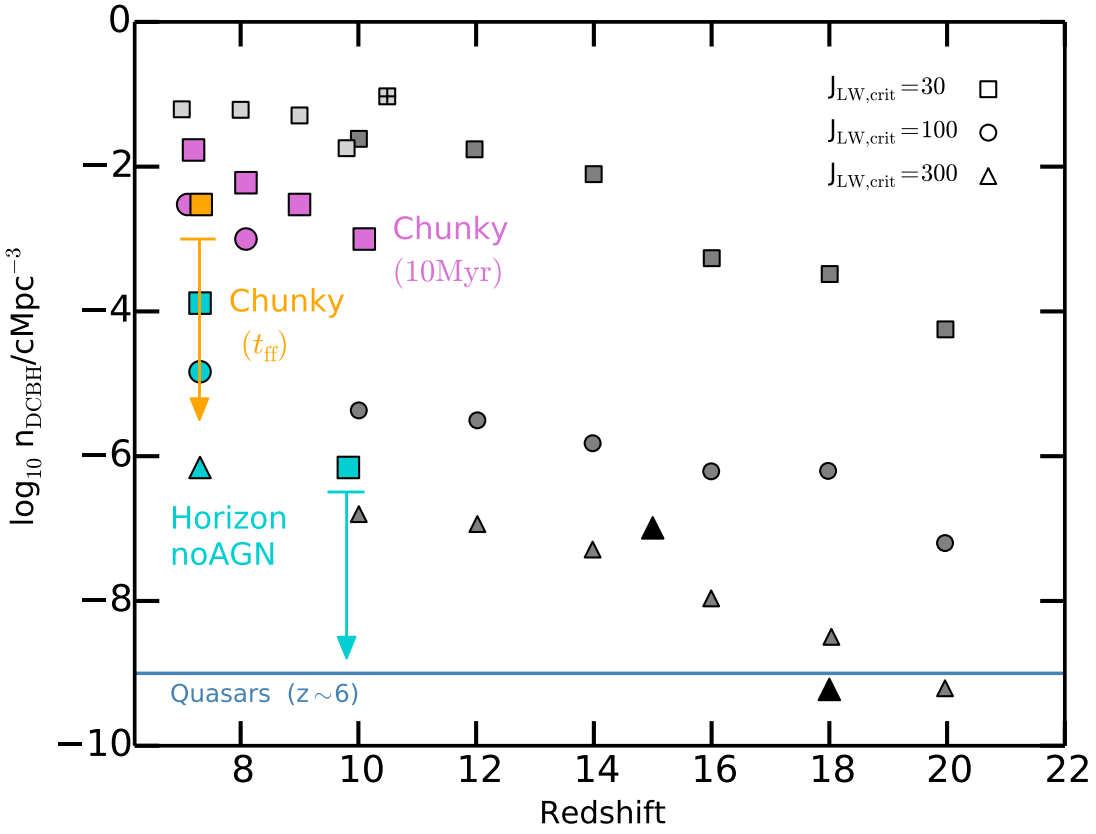


Figure 5.4: Comoving number density of DCBH formation sites, as a function of redshift. Symbol shapes represent different radiation intensity thresholds. Squares: $J_{\rm crit} = 30$, circles: $J_{\rm crit} = 100$, triangles: $J_{\rm crit} = 300$. The horizontal solid blue line shows the comoving number density of quasars at $z \sim 6$. The light grey crossed square at $z = 10.5$ is from the hydrodynamical simulation by Agarwal et al. (2014), the light grey squares in the range $z = 7 - 10$ are from Agarwal et al. (2012), dark grey squares and black triangles are the results of Dijkstra et al. (2014) and Valiante et al. (2016), respectively. The orange square and purple symbols show the number density for Habouzit et al. (2016a) for different assumptions on the collapse time. The cyan squares, circle and triangle represent the large-scale cosmological simulation Horizon-noAGN (Dubois et al. 2014b). Adapted from Valiante et al. (2017).

absolute values of the number density differ by several orders of magnitude. All studies agree that chemical enrichment from previous star formation in progenitor halos or by metal-enriched winds from nearby star-forming halos play a crucial role in determining the final number density of DCBHs. Also the implementation of SN feedback strongly affects the possibility to form a DCBH (Habouzit et al. 2016a). Hydrodynamical simulations suffer from the limited volume which allows in some cases only for upper limits on the DCBH density (Agarwal et al. 2014; Habouzit et al. 2016a). Other simulations find values of 2.5×10^{-4} cMpc^{-3} for $J_{\rm crit} = 100$ (Chon et al. 2016) and $(0.1 - 5) \times 10^{-6}$ cMpc^{-3} for $J_{\rm crit} = 300$ (Habouzit et al. 2016c). Using an analytical estimate for the clustering of halos, (Yue et al. 2014) find that the DCBH mass density rises from $\sim 5\,{\rm M}_\odot\,{\rm Mpc}^{-3}$ at $z \sim 30$ to the peak value $\sim 5 \times 10^5\,{\rm M}_\odot\,{\rm Mpc}^{-3}$ at $z \sim 14$ in their fiducial model. This value is larger than the total SMBH mass density at $z = 0$ (Soltan 1982), which has two reasons: first, they assume low values of the critical LW flux with the highest values tested in their models of $J_{\rm crit} = 150$ and second, they include the positive feedback of DCBHs on their own formation, which fosters a rapid multiplication of the number of DCBHs. We refer the interested

reader to Habouzit et al. (2016a) for a quantitative comparison of the different semi-analytical models.

Bellovary et al. (2011) use cosmological simulations to derive the BH occupation fraction and they show that halos with $M > 3 \times 10^9\,\mathrm{M}_\odot$ host massive BHs regardless of the efficiency of seed formation. Habouzit et al. (2016a) find, based on post-processing state-of-the-art cosmological simulations (Dubois et al. 2014), that $\sim 30\%$ of halos more massive than $10^{11}\,\mathrm{M}_\odot$ at $z = 6$ have at least one progenitor that fulfils the conditions for DC, assuming a critical value of $J_{\mathrm{crit}} = 30$. However, this occupation fraction is a strong function of the critical flux and falls below $1 - 10\%$ for $J_{\mathrm{crit}} > 100$. In Habouzit et al. (2016c) they confirm an occupation of $\sim 10\%$ for Pop III remnant BHs and an occupation fraction of the order $10^{-6} - 10^{-5}$ for DCBHs (see also Tanaka & Haiman 2009).

Although there are still large theoretical uncertainties in the number density of DCBH formation sites, caused by e.g. the model of metal enrichment, the value of J_{crit}, or the necessary duration of the photodissociating feedback, all models are able to yield more than 1 DCBH per cGpc^3 by $z = 6$, which is the lower limit set by current observations. The most optimistic scenarios are even able to explain the majority of BHs in local galaxies via the DC scenario. The remaining bottleneck is to grow these seed BHs to SMBHs by circumventing the self-regulating accreting feedback as we will see in the next chapter.

In this chapter, we have presented the rather consensual picture on the statistical predictions on BH formation in the early Universe. However, we note that most presented results are subject of ongoing debate and reality might lie beyond what we currently assume to be realistic. We have presented DC as a promising pathway to form massive BH seeds, but other formation channels than the isothermal collapse may also be successful, or different scenarios, such as runaway collisions in the first stellar clusters could be more efficient than we have assumed (Katz et al. 2015; Sakurai et al. 2017). Moreover, the conservative conditions that the gas for the BH formation via gas-dynamical processes has to be pristine might be too strict and BH seed formation is also possible in gas that is slightly enriched with metals or dust (Omukai et al. 2008; Mayer et al. 2015; Latif et al. 2016b; Agarwal et al. 2017b). All these caveats affect the predicted formation efficiency and number density of BH seeds in the early Universe.

6 2D analytical model for AGN-driven outflows in galaxy discs

This chapter is based on the paper "Active Galactic Nuclei outflows in galaxy discs", submitted to MNRAS (Hartwig et al. 2017b).

6.1 Motivation

We discuss the growth of the first BHs and the feedback associated with this accretion. We do not model the accretion history self-consistently, but focus on AGN-driven winds. Note that AGN-driven winds are one possible negative feedback process to prevent the infall and accretion of gas on to the central BH. We do not account for feedback via radiation pressure, SNe, jets, or an interplay of these processes.

Numerical 3D simulations of AGN-driven outflows demonstrate that the outflow takes the path of least resistance and propagates preferentially along the poles of the host galaxy (Gabor & Bournaud 2014; Costa et al. 2015; Bieri et al. 2017). This results in highly anisotropic outflows as found by Costa et al. (2014) and observationally confirmed by Cicone et al. (2015). However, 3D simulations remain expensive, have explored limited parameter space, and have not yet focused on low-mass or high redshift galaxies.

Our aim is to develop a flexible, but general model to study AGN feedback in galaxies where rotational support in the gas component is non-negligible. We derive a 2D analytical model for AGN-driven outflows: two dimensions are sufficient to capture the geometry of a disc-like galaxy, which allows simultaneously an AGN-feeding inflow and an ISM-ejecting outflow, but it is still simple enough to be treated analytically and flexible enough to explore a large parameter space with this model. We include a self-consistent transition from momentum- to energy-driving and derive observables, such as velocities, mass outflow rates, and momentum loading.

6.2 Methodology

In this section we describe the basic equations that govern our 2D analytical approach, justify necessary approximations, and introduce our physically motivated galaxy model.

6.2.1 Galaxy model

We present and motivate the individual components and characteristic radii of our analytical approach.

Gas density profile of the disc

Our main focus is to model AGN-driven outflows in low-mass galaxies, which represent the first galaxies at high redshift or dwarf galaxies at lower redshift. We assume that the gas distribution in these

galaxies can be approximated by a 2D axisymmetric density profile. The surface brightness of many spiral galaxies can be described by an exponential profile (Sérsic 1963; Freeman 1970; Courteau et al. 1996). Smit et al. (2017) recently report the observation of two $z \approx 6.8$ galaxies with a clear velocity gradient in CII, which might suggest the presence of a turbulent, rotation-dominated disc. Based on imaging spectroscopy of the Hα emission line, Genzel et al. (2017) confirm that the observed velocity profiles of high-z disc galaxies favour a thick exponential disc profile. We therefore assume an exponential profile for the surface density

$$\Sigma(R) = \Sigma_0 \exp(-R/R_0), \tag{6.1}$$

where R_0 is the scaling radius, given by (Mo et al. 1998):

$$R_0 = \frac{\lambda}{\sqrt{2}} \left(\frac{j_d}{m_d} \right) R_{\rm vir}, \tag{6.2}$$

with j_d being the fraction of the specific angular momentum of the halo that goes into the disc, m_d the fraction of the halo mass that settles into the disc, and λ the spin parameter of the halo. For typical halos we assume $j_d \approx m_d \approx \lambda \approx 0.05$ and hence

$$R_0 \approx 0.035 R_{\rm vir}. \tag{6.3}$$

The scale height of an isothermal, self-gravitating disc is given by

$$H(R) = \frac{c_s^2}{\pi G \Sigma(R)} = \frac{c_s^2}{\pi G \Sigma_0} \exp(R/R_0), \tag{6.4}$$

where c_s is the sound speed of the gas. The derivation of this relation assumes a locally constant surface density and the real scale height might be smaller in regions where the disc surface density changes significantly with radius. For the total density profile of the disc we require

$$\Sigma(R) = \int_{-\infty}^{\infty} \rho(R,h) {\rm d}h. \tag{6.5}$$

By assuming

$$\rho(R,h) \propto \cosh^{-2}(h/H(R)) \tag{6.6}$$

we find

$$\rho(R,h) = \frac{\Sigma_0}{2H(R)} \exp(-R/R_0) \cosh^{-2}(h/H(R)). \tag{6.7}$$

The gas mass inside a radius R for the exponential disc is given by

$$M(<R) = 2\pi\Sigma_0 R_0 \left[R_0 - \exp(-R/R_0)(R + R_0) \right]. \tag{6.8}$$

For reasonable values of the spin parameter λ, the second term is negligible. Assuming that $m_d M_h$ is the mass of the disc within $R_{\rm vir}$ we can calculate the normalisation of the surface density as

$$\Sigma_0 = \frac{m_d M_h}{\pi (\lambda R_{\rm vir})^2}, \tag{6.9}$$

where M_h is the mass of the halo.

Dark matter

Following Costa et al. (2014), we assume a spherically symmetric Hernquist profile (Hernquist 1990) of the form

$$\rho_{DM} = (1 - f_{\rm gas}) \frac{M_h}{2\pi} \frac{a}{r(r+a)^3} \tag{6.10}$$

and

$$M_{DM}(< r) = (1 - f_{\rm gas}) M_h \frac{r^2}{(r+a)^2}, \tag{6.11}$$

with $f_{\rm gas} = \Omega_b/\Omega_m = 0.17$ and the scaling radius $a = 0.1 R_{\rm vir}$.

1D reference model

To highlight the differences of a 2D model, we also compare our new model to a spherical 1D gas distribution for which we assume that the radial gas profiles follows a Hernquist profile with

$$\rho_{\rm gas} = \frac{f_{\rm gas} M_h}{2\pi} \frac{a}{r(r+a)^3}, \tag{6.12}$$

where $a = 0.1 R_{\rm vir}$.

Fiducial galaxy model

For our fiducial model we assume typical parameters of a $z = 6$ galaxy with a halo mass of $M_{\rm h} = 10^8\,{\rm M}_\odot$. The BH has a mass of $M_{\rm BH} = 10^5\,{\rm M}_\odot$ and shines at an Eddington ratio of $f_{\rm Edd} = 0.3$. We assume that the AGN shines for a time of $t_{\rm on} = 10\,{\rm Myr}$ with a constant luminosity that is given by the black hole mass and the Eddington ratio. The resulting surface density normalisation and disc scale height at the scale radius are $\Sigma_0 = 0.28\,{\rm g\,cm^{-2}}$ and $H(R_0) = 2.6\,{\rm pc}$, respectively. In section 6.3.2 we vary all these model parameters independently to analyse their influence on the outflow dynamics. We do not include stars in our model of AGN-driven outflows.

A smaller gas mass fraction, e.g. at lower redshift, would decrease the surface density normalisation. Also, gas in high-z galaxies is turbulent and has significant pressure support (Genzel et al. 2008; Förster Schreiber et al. 2009; Safranek-Shrader et al. 2012b), which would additionally increase the disc scale height. These effects would lower the ISM density in the disc plane and permit the AGN-driven wind to propagate to larger radii.

Initial and final radii

We start the integration of the outward driven shock at an initial radius $R_{\rm min}$. The physical interpretation of this radius is the innermost distance where the disc wind starts to sweep up the interstellar medium (ISM). We use the self-gravity radius of the accretion disc as a proxy for $R_{\rm min}$ and we express it in units of the Schwarzschild radius,

$$R_{\rm SS} = \frac{2GM_{\rm BH}}{c^2} \approx 3 \times 10^5 \frac{M_{\rm BH}}{{\rm M}_\odot}\,{\rm cm}. \tag{6.13}$$

For standard values of the disc viscosity and Eddington ratio the outer radius of the disc (i.e. self-gravity radius) and hence the inner radius for our integration is given by (Shakura & Sunyaev 1973; Goodman & Tan 2004; King et al. 2008; Dotti et al. 2010):

$$R_{\rm min} \approx 10^5 R_{\rm SS}, \tag{6.14}$$

which yields consistent results for all the tested parameter combinations and for black hole masses, spanning many orders of magnitude.

We integrate the equations out to arbitrary radii, but the physically interesting regime is within the virial radius. Hence, we follow the outflow in each direction until it either becomes subsonic, or until it reaches the virial radius of the halo. Since the escape velocity at the virial radius is roughly equal to the sound speed, we assume that the wind escapes the gravitational potential of the halo if it reaches the virial radius and is still supersonic.

6.2.2 Quantifying the outflow

We define a set of parameters which quantify the efficiency, dynamics, and nature of the outflow.

We want to introduce a radius, $R_{\rm max}$, out to which the outflow has to push the gas at least to stop further gas accretion on to the central BH. Due to the higher gas density in the disc plane, it is a sufficient criterion for the outflow to reach $R_{\rm max}$ in the disc plane. A sufficient criterion to prevent further gas accretion on to the central BH is $R_{\rm max} = R_{\rm vir}$ and a necessary criterion is $R_{\rm max} > R_{\rm soi}$, where $R_{\rm soi}$ is the sphere of influence, within which the dynamics is dominated by the black hole.

Once pushed out to a certain radius, the question is if this gas can reach the BH in a sufficiently short amount of time to start refuelling the AGN. We hence relate $R_{\rm max}$ to the recovery time $t_{\rm ff}$, which we define as the minimum time the gas needs to fall back towards the central BH, after being pushed out to the radius $R_{\rm max}$. Setting the free-fall time to the sound crossing time yields for $R_{\rm max}$ in the disc plane

$$t_{\rm ff} = 5\,{\rm Myr}\left(\frac{R_{\rm max}}{R_0}\right)\left(\frac{1+z}{7}\right)^{-3/2}. \tag{6.15}$$

There is no unique criterion on the absolute value of $R_{\rm max}$ or $t_{\rm ff}$ to shut off further gas accretion. We therefore present and discuss both relative to the scale radius of the disc and to the lifetime of the AGN. If $t_{\rm ff}$ is of the order of the AGN lifetime, we can assume that the outflow is efficient enough to significantly suppress further mass accretion.

The ejected mass $M_{\rm eject}$ is the gas mass that is ejected from the gravitational potential of the halo, i.e. that passes the virial radius with a velocity larger than the escape velocity. The ejection angle $\theta_{\rm eject}$ is the last angle (measured from the disc normal) for which gas can escape the halo. It defines the opening angle of a cone, which is cleared off gas by the AGN-driven wind. Gas outside this cone is pushed to $R_{\rm max}(\theta)$, but falls back towards the galactic centre over a time $\gtrsim t_{\rm ff}$. The free-fall time in the disc plane is defined as the time which gas that has been pushed to a certain radius needs at least to fall back towards the central BH, once the shock becomes subsonic (see Eq. 6.15). This is an approximate quantification for how long an AGN can shut off its own gas supply.

We also compare the mass accretion rate of the central BH:

$$\dot{M}_{\rm acc} = \frac{L}{\eta c^2}, \tag{6.16}$$

to the mass outflow rate:

$$\dot{M}_{\rm out} = \sum_i M_{\rm shell,i}\frac{v_{\rm shell,i}}{R_{\rm shell,i}}, \tag{6.17}$$

where the mass and velocity of the shell resolution elements are summed over all angles i. The mass outflow rate generally varies with time, but to make it more accessible for a direct comparison to the mass accretion rate, we also define the mass outflow rate at the sphere of influence, which is one unique

number. In theory, the shell crosses the sphere of influence at different times for different angles and we need to sum over shell elements at different "crossing-times". However, in practice, the shell crosses the sphere of influence for all angles in the same time step, because the density profile of the exponential disc is close to spherical in the centre and there is no preferred direction. Comparing these two mass rates quantifies the efficiency of the AGN to convert infalling, accreted mass into a mass outflow.

The ratio $R_{\rm perp}/R_{\rm disc}$ quantifies the asymmetry of the outflow. At any time, $R_{\rm disc}$ is the position of the shock front in the disc plane and $R_{\rm perp}$ perpendicular to it. If this ratio is equal to one, the wind propagates almost spherically symmetric. For values above unity the outflow develops peanut-shaped lobes. This ratio illustrates the importance and improvement of a 2D treatment with respect to 1D wind solutions.

We define the momentum boost, which is often referred to as "mechanical advantage", as the ratio of the total momentum of the AGN-driven wind divided by the total momentum input by photons:

$$\frac{\sum_i |\mathbf{p}_i|}{tL/c}. \tag{6.18}$$

The vectorial sum would be equal to zero, but summing the absolute values of the momentum allows to quantify the efficiency of the AGN to accelerate the gas. In the momentum-driven regime we expect

$$\sum_i |\mathbf{p}_i| \leq tL/c, \tag{6.19}$$

but the adiabatic expansion in the energy-driven case can yield mechanical advantages above unity. Observations find momentum boost in the range $\sim$ 2–30 (Rupke & Veilleux 2011; Sturm et al. 2011) and 3D simulations yield momentum boosts of $1-30$ (Cicone et al. 2014; Bieri et al. 2017).

6.2.3 Shock acceleration

There are two systematically different types of outflows: if the shocked wind cools efficiently, the outflow is accelerated only by the momentum transfer and the outflow is called *momentum-driven*. A momentum-driven outflow develops a thin shock layer. For less efficient cooling, the shock-heated gas remains at temperatures of $> 10^7$ K and the adiabatic expansion of the hot shocked wind accelerates the shock into the ISM. In this case, a thick shock layer develops and the outflow is labelled *energy-driven*, because the internal energy of the adiabatically expanding shock-heated wind drives the outflow.

In a disc-like geometry we expect the shock to propagate faster into the direction of lower density. This implies that a self-consistent 2D treatment of the wind propagation is important to capture all the relevant physics of the shock dynamics.

We can assume that the shock front is locally plane-parallel and in the rest frame of the shock front, the Rankine-Hugoniot jump conditions for a plane parallel shock are:

$$\rho_1 u_1 = \rho_2 u_2 \tag{6.20}$$

$$\rho_1 u_1^2 + P_1 = \rho_2 u_2^2 + P_2 \tag{6.21}$$

$$\frac{1}{2}\rho_1 u_1^2 + \epsilon_1 + \frac{P_1}{\rho_1} = \frac{1}{2}\rho_2 u_2^2 + \epsilon_2 + \frac{P_2}{\rho_2}, \tag{6.22}$$

where the index 1 refers to the pre-shock and the index 2 to the post-shock conditions of the density ρ,

velocity u, pressure P, and the specific internal energy ϵ. For a strong shock with Mach number

$$\mathcal{M}_1 = \left(\frac{\rho_1 u_1^2}{\gamma P_1}\right)^{1/2} \gg 1, \tag{6.23}$$

we find

$$\frac{\rho_2}{\rho_1} \approx \frac{\gamma+1}{\gamma-1} = 4 \tag{6.24}$$

$$P_2 \approx \frac{2}{\gamma+1}\rho_1 u_1^2 = \frac{4}{3}\rho_1 u_1^2 \tag{6.25}$$

$$T_2 \approx \frac{2(\gamma-1)}{(\gamma+1)^2}\frac{m}{k_\mathrm{B}}u_1^2 = \frac{3}{16}\frac{m}{k_\mathrm{B}}u_1^2 \tag{6.26}$$

$$\mathcal{M}_2 \approx \left(\frac{\gamma-1}{2\gamma}\right)^{1/2} \approx 0.45, \tag{6.27}$$

where the last equalities are valid for an adiabatic coefficient of $\gamma = 5/3$ (Shull & Draine 1987). Hence, a shock converts supersonic gas into denser, slower moving, higher pressure, subsonic gas.

The density of the shocked wind depends on the available cooling channels. For a momentum-driven strong shock, the post-shock density is given by $\rho_2 = 4\rho_1$ (Eq. 6.24). In the energy-driven case the density of the shocked wind depends on the thickness of the shock. We generally refer to the shock position as the contact discontinuity between the shocked wind and the shocked ISM, and the thickness of the shocked wind is given by

$$\Delta R = R - R_\mathrm{sw}, \tag{6.28}$$

where R and R_sw are the position of the contact discontinuity and of the wind shock, respectively (see illustrations of outflow dynamics in Zubovas & King 2012; Faucher-Giguère & Quataert 2012; King & Pounds 2015). The radius where the wind is shocked for an energy-driven outflow is given by Faucher-Giguère & Quataert (2012):

$$R_\mathrm{sw} \approx R\sqrt{\frac{v}{v_\mathrm{in}}}, \tag{6.29}$$

and the thickness of the shocked wind is $\Delta R = R(1-(v/v_\mathrm{in})^{1/2})$. We do not use this thickness of the shock explicitly for the integration of the equations of motion, but we use it implicitly to calculate the efficiency of radiative cooling (Section 6.2.4).

Equation of motion: momentum-driven

The shock is accelerated by the momentum input of the AGN, counteracted by gravity (King 2010; Ishibashi & Fabian 2014, 2015; King & Pounds 2015):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[M_\mathrm{shell}(r)\dot{r}\right] = \frac{f_\mathrm{edd}L_\mathrm{edd}}{c} - G\frac{M_\mathrm{shell}(r)(M_\mathrm{DM}(<r)+M_\mathrm{BH})}{r^2} - F_\mathrm{selfgrav}, \tag{6.30}$$

where M_shell is the mass of the swept up ISM and

$$L_\mathrm{edd} = \frac{4\pi c G M_\mathrm{BH} m_p}{\sigma_T} \tag{6.31}$$

is the Eddington luminosity. The term $F_{\rm selfgrav}$ is not explicitly mentioned in most studies (King 2010; Ishibashi & Fabian 2014, 2015; King & Pounds 2015), but it accounts for the swept up gas in the shell and the gravity contribution from the gas outside the shell. The gravitational force of the ambient gas is negligible (in spherical symmetry it cancels out), but the self gravity of the shell has to be included.

The term L/c describes the total momentum per time that can be transferred to the gas via single scatter events. For Thomson scattering we assume that the directions of the photons after one scattering are random and that the momentum transfer from secondary scattering events cancels out. However, there is a probability $0 < p \leq 1$ for the photons to escape freely without any interaction with the ISM:

$$\ln(p) = -\frac{\sum_i \sigma_{t,i}}{d\Omega r^2} = -\frac{M_{\rm shell}}{m_p}\frac{\sigma_t}{d\Omega r^2} = -\kappa\Sigma_{\rm shell} = -\tau_{\rm shell}, \tag{6.32}$$

with the opacity $\kappa = \sigma_t/m_p = 0.4{\rm cm}^2\,{\rm g}^{-1}$. The fraction of the total force input L/c that actually couples to the gas is hence

$$\frac{L}{c}\left(1 - e^{-\tau_{\rm shell}}\right). \tag{6.33}$$

We do not explicitly account for this effect and show in Section 6.3.1 that it is negligibly small and irrelevant in most of the scenarios considered. However, see Ishibashi & Fabian (2015) for a more detailed discussion of this effect and for the influence of multiscattering of IR photons on dust grains.

Equation of motion: energy-driven

In the energy conserving case, the shocked ISM is not accelerated by the direct momentum input, but by the adiabatic expansion of the shocked wind. The momentum equation then reads

$$\frac{\rm d}{{\rm d}t}[M_{\rm shell}(R)\dot{R}] + F_{\rm grav} = 4\pi R^2 P, \tag{6.34}$$

where P is the pressure in the shocked wind, which can be calculated via energy conservation

$$(\gamma - 1)\frac{\rm d}{{\rm d}t}(PV) = \frac{\eta}{2} f_{\rm Edd} L_{\rm Edd} - P\frac{\rm d}{{\rm d}t}V - F_{\rm grav}\dot{R}, \tag{6.35}$$

where V is the volume and η sets the mechanical luminosity, which can be expressed as $\eta = v_{\rm in}/c$ (see section 6.2.5). This energy conservation assumes that the thermal energy of the shocked wind is much higher than its kinetic energy, which is generally the case (Faucher-Giguère & Quataert 2012).

6.2.4 Cooling and heating

In this section we discuss the most important heating and cooling mechanisms of the gas. We demonstrate that inverse Compton cooling is the only relevant mechanism on the scales of interest (Ciotti & Ostriker 1997) and discuss why radiative cooling of the wind shock can be neglected.

Compton cooling

The Compton cooling time of gas is given by

$$t_{\rm compton} = \frac{3 m_e c}{8\pi\sigma_T U_{\rm rad}}\frac{m_e c^2}{E}, \tag{6.36}$$

with the radiation energy density

$$U_{\rm rad} = \frac{f_{\rm Edd} L_{\rm Edd}}{4\pi R^2 c} \tag{6.37}$$

and the internal energy in the shocked gas

$$E = \frac{9 m_p v_{\rm in}^2}{16}, \tag{6.38}$$

where $v_{\rm in}$ is the initial velocity of the disc wind, before it shocks with the ISM (see Section 6.2.5) and m_p and m_e are the masses of the proton and electron, respectively. The Compton cooling time is hence given by (see also Faucher-Giguère & Quataert 2012)

$$t_{\rm compton} = \frac{2}{3\pi f_{\rm edd}} \frac{c}{G M_{\rm BH}} \left(\frac{m_e}{m_p}\right)^2 \left(\frac{v_{\rm in}}{c}\right)^{-2} R^2. \tag{6.39}$$

Two temperature medium

Inverse Compton cooling acts only on the electrons. Faucher-Giguère & Quataert (2012) analyse the cooling properties of a medium with different temperatures of the electrons and ions. They determine the characteristic time-scale to achieve thermal equilibrium ($T_e = T_p$) to be

$$t_{ei} = \frac{3 m_e m_p}{8(2\pi)^{1/2} n_p e^4 \ln\Lambda} \left(\frac{k_B T_e}{m_e} \frac{k_B T_p}{m_p}\right)^{3/2}, \tag{6.40}$$

where e is the elementary charge and $\Lambda \approx 40$. Assuming $T_p = 10 T_e = 10^{10}\,\mathrm{K}$ and a proton density of $n_p = 1\mathrm{cm}^{-3}$ as typical post-shock conditions, this time-scale is of the order $t_{ei} \approx 1\,\mathrm{Myr}$. Faucher-Giguère & Quataert (2012) show that this time-scale is longer than the Compton cooling time and might delay the cooling by up to two orders of magnitude. We do not include this effect of a two-temperature medium in our calculations and discuss its effect in Section 6.4.4.

Radiative cooling

We only consider inverse Compton cooling in our calculation of the cooling time. To quantify the possible contribution of radiative cooling, we use the cooling function by Sutherland & Dopita (1993) in the functional form by Tozzi & Norman (2001) with the units $10^{-22}\,\mathrm{erg\,cm^3\,s^{-1}}$:

$$\Lambda = \begin{cases} 1.544 \left(\frac{T}{2\times10^4\,\mathrm{K}}\right)^6 & \text{for } T \leq 2\times10^4\,\mathrm{K} \\ 6.72 \left(\frac{T}{2.32\times10^5\,\mathrm{K}}\right)^{0.6} & \text{for } 2\times10^4\,\mathrm{K} < T \leq 2.32\times10^5\,\mathrm{K} \\ 8.8\times10^9 \left(\frac{T}{\mathrm{K}}\right)^{-1.7} + 1.7\times10^{-5}\left(\frac{T}{\mathrm{K}}\right)^{0.5} + 0.063 & \text{for } T > 2.32\times10^5\,\mathrm{K} \end{cases} \tag{6.41}$$

The post-shock temperature is given by Eq. 6.26 and we calculate the gas density based on Eq. 6.28 or Eq. 6.24, depending on whether the shock is energy- or momentum-driven, respectively. The radiative cooling time becomes:

$$t_{\rm rad} = \frac{9 v_{\rm in}^2 \mu m_{\rm p}^2}{32 \Lambda \rho_{\rm shock}}, \tag{6.42}$$

where $\rho_{\rm shock}$ is the corresponding post-shock density and Λ the cooling function. The radial profile of the radiative cooling time and a comparison to other characteristic time-scales can be seen in Figure 6.1. Within $\sim 0.01\,\mathrm{pc}$ the velocity of the swept up material is close to the initial wind velocity and therefore the shock would be infinitely thin (Eq. 6.29). This implies a very high density and an artificially short

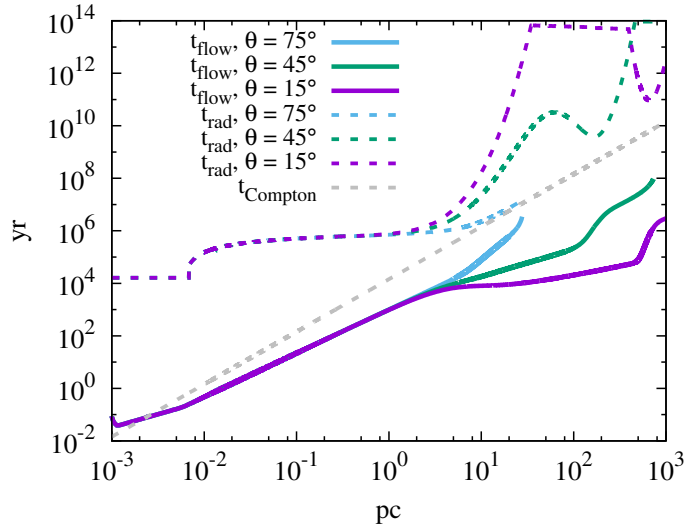


Figure 6.1: Radial profiles of characteristic time-scales for our fiducial model. The different colours indicate different angles with respect to the disc normal and the Compton cooling time in this radial range is independent of the angle (grey).

cooling time. In this case we set the post-shock density to $\rho_{\rm shock} = 4\rho$ to avoid singularities. On all scales of interest, the radiative cooling time is always longer than the Compton cooling time and can be neglected.

6.2.5 Mechanical luminosity and initial wind velocity

The post-shock temperature is set by the pre-shock wind velocity $v_{\rm in}$ (Eq. 6.26). By assuming that the initial wind mass rate $\dot{M}_{\rm in}$ is equal to the accretion rate on to the black hole $\dot{M}_{\rm BH}$, with a radiative efficiency of $\eta \approx 0.1$ for the AGN, we obtain

$$v_{\rm in} = \eta c \approx 30000\,{\rm km\,s^{-1}}. \tag{6.43}$$

The assumption $\dot{M}_{\rm in} = \dot{M}_{\rm BH}$ might not be accurate, but we use it as a working hypothesis, following e.g. Faucher-Giguère & Quataert (2012); Costa et al. (2014); King & Pounds (2015) and we explore the dependence on $v_{\rm in}$ in Section 6.3.2. The assumption of $v_{\rm in} = 0.1c$ is supported by observations of the Doppler-shifted Fe K band absorption lines, which are observed in many local AGNs (Pounds et al. 2003; Tombesi et al. 2010; Tombesi 2016). Ostriker et al. (2010) provide a more detailed and complete discussion of the momentum driving and mechanical luminosity.

This assumption enters in the calculation of the cooling time and for the mechanical input of the AGN. If we assume that the initial mass outflow rate is only a fraction of the accretion rate:

$$\dot{M}_{\rm in} = f_{\rm acc}\dot{M}_{\rm BH}, \tag{6.44}$$

we would find $v_{\rm in} = \eta f_{\rm acc} c$ for the initial velocity and the energy equation would change to

$$\frac{\eta}{2}L_{\rm AGN} \rightarrow \frac{\eta}{2}f_{\rm acc}L_{\rm AGN} = \frac{\eta}{2}f_{\rm acc}f_{\rm Edd}L_{\rm Edd}. \tag{6.45}$$

Hence, in the parameter study of the outflow dynamics, $f_{\rm acc}$ and $f_{\rm Edd}$ enter the equation of motion as

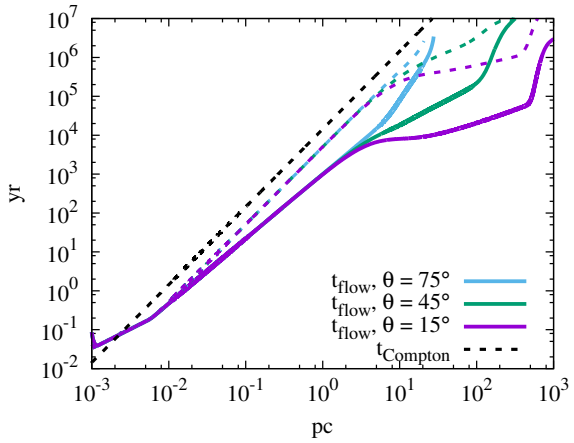


Figure 6.2: Characteristics times as a function of radius for different angles with respect to the disc normal for our fiducial model. The black dashed line shows the Compton cooling time (Eq. 6.39) and the solid lines show the flow time for the self-consistent driving, i.e. energy-driven when $t_{\rm flow} < t_{\rm Compton}$. The dashed lines illustrate the dynamics of an outflow that is always momentum-driven. The transition from momentum- to energy-driven occurs at $\sim 2 \times 10^{-3}$ pc and the momentum-driven solution is less efficient in driving an outflow.

a product, which allows us to keep one parameter fixed ($f_{\rm acc} = 1$) and vary only the Eddington ratio.

We assume that this initial line-driven disc wind is isotropic. This is supported by observations of ultrafast outflows in a large fraction of observed AGNs. Unless we see all AGNs from a very particular angle, these AGN-driven winds have large opening angles with an almost spherical driving mechanism (King & Pounds 2015).

6.2.6 Momentum- to energy-driven transition

The main problem arises since the nature of the outflow depends on the dynamics of the shock and the shock dynamics depends on the nature of the outflow. We self-consistently check at any time if the outflow is energy- or momentum-driven (see Figure 6.2). We expect the transition to happen at different radii for different angles with respect to the disc plane. This implies that there are phases where momentum- and energy-driving occur at the same time in a given galaxy. For the energy input by the AGN in driving the adiabatic expansion, we correct it by the solid angle $\mathrm{d}\Omega$ of the energy-driven outflow, because for $4\pi - \mathrm{d}\Omega$ we expect the energy input to be radiated away in the momentum-driven shock front. However, the remaining question is, which volume represents the correct basis for the energy equation of the adiabatic expansion: the entire volume enclosed by the shock front or only the volume enclosed by the energy-driven shock front? A Gedankenexperiment can help to settle this question: if we only had to consider the volume of the cone enclosed by the energy-driven shock (perpendicular to the disc plane, because we expect higher velocities in this direction), then what happens at the boundary of this cone? The adiabatically expanding volume does not care about whether it is enclosed by an energy- or momentum-driven shock front, but rather wants to expand in all directions. For the same reason, we have to adopt our treatment of the adiabatically expanding volume once the shock becomes subsonic: the gas still wants to expand and can do so into the ISM with $\sim c_s$, even if the shock front is dissolved by turbulence. For directions into which the shock becomes subsonic, we hence follow the propagation of the swept-up mass and the adiabatically expanding volume separately. We have tested the influence of these methods and find that the difference is small.

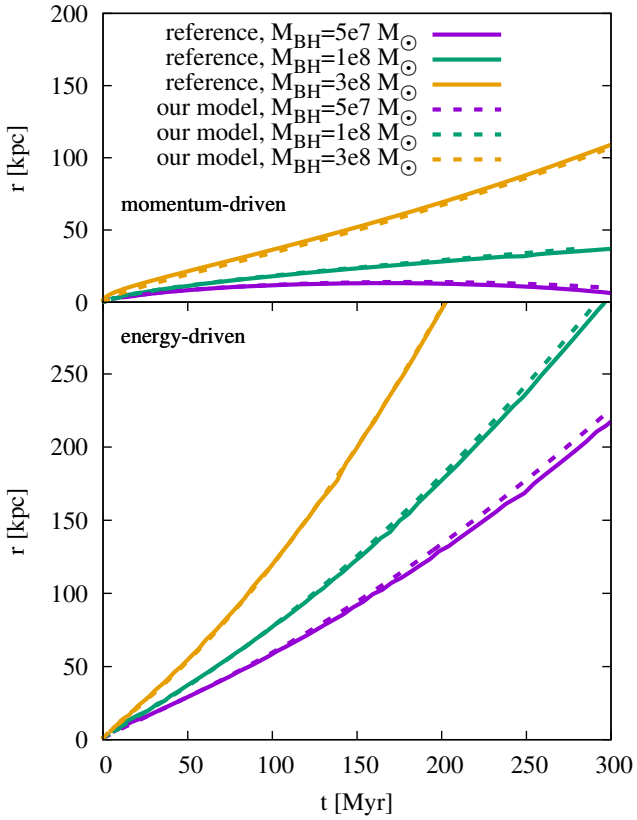


Figure 6.3: Comparison of our model with the results by King (2005) and Costa et al. (2014) for different black hole masses. Here, we assume the same spherical density distribution as Costa et al. (2014), but integrate the equations of motion with our 2D model. Our results agree well with the analytical solution by King (2005) and with the 3D simulations by Costa et al. (2014).

6.2.7 Validation of the model: comparison to 1D solution

To test the robustness and convergence of the method, we first compare the results of our 2D analytical model to the results of Costa et al. (2014) and King (2005). We assume $M_h = 10^{12}\,\mathrm{M}_\odot$ for the halo mass, a Hernquist profile for the dark matter and the gas, where the gas mass is $M_{\mathrm{gas}} = f_{\mathrm{gas}} M_h$ with $f_{\mathrm{gas}} = 0.17$ and the scaling radius $a = 28\,\mathrm{kpc}$. The virial radius is $R_{\mathrm{vir}} = 163\,\mathrm{kpc}$, the AGN shines at the Eddington luminosity and for the black hole we assume three different masses of $M_{\mathrm{BH}} = 5 \times 10^7, 10^8, 3 \times 10^8\,\mathrm{M}_\odot$. We assume that there is no gas within the sphere of influence of the BH, which is $\sim 500\,\mathrm{pc}$ for these configurations. The results can be seen in Figure 6.3.

Our 2D analytical model is able to reproduce the outflow dynamics of the 1D model by King (2005) and of the 3D simulation by Costa et al. (2014) in a spherical gas profile. We then adopt the same functional forms for the gas distribution (Hernquist profile) and apply it to our fiducial model with $M_h = 3 \times 10^7\,\mathrm{M}_\odot$ and $M_{\mathrm{BH}} = 10^5\,\mathrm{M}_\odot$. The direct comparison of the outflow dynamics in 1D and 2D is given in Figure 6.4. For the same gas mass and AGN luminosity, the outflow in the 2D model can only eject mass within a cone with an opening angle of $\sim 45°$, whereas the same outflow in the 1D model ejects all the gas out of the virial radius. The main reason for this lower efficiency in the 2D scenario is that the pressure in the 1D model cannot escape, but builds up over time. In the 2D scenario, however, the pressure decreases with time due to the fast expansion of the energy-driven wind perpendicular to the disc.

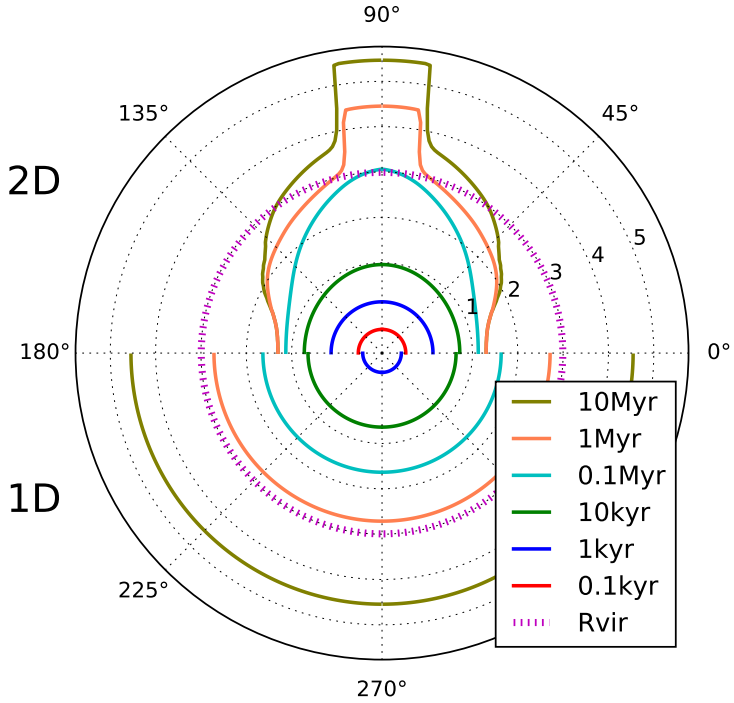


Figure 6.4: Time evolution of the position of the shock front in polar coordinates. Top: 2D solution with the galactic disc plane at 0°. Bottom: 1D spherical solution. The concentric grey rings indicate the radius in $\log(r/\mathrm{pc})$ from 10 pc to 100 kpc and the magenta circle represents the virial radius at 980 pc. In the disc plane, the shock cannot propagate beyond a maximal radius of ~ 20 pc and only gas at $\gtrsim 45°$ escapes the virial radius. The 1D wind is slower at $\lesssim 30$ kyr, but ejects more gas out of the virial radius at later times.

6.3 Results

6.3.1 Standard case (fiducial parameters)

We show the time evolution of the main dynamical quantities for our fiducial model in Figure 6.5. Within ~ 0.01 pc the swept up mass and hence the gas inertia is negligibly small and the dynamics is dominated by the impinging disc wind, which has an initial velocity of $v_{\mathrm{in}} = 0.1c$ in our fiducial model. Our model is insensitive to the choice of the initial radius of integration R_{min}, as long as it is within this disc wind-dominated region. When the cooling time becomes longer than the flow time, the pressure rises by about ten orders of magnitude and the outflow becomes energy-driven. The outflow slows down due to the inertia of the swept up ISM, which was at rest first. Within ~ 10 pc the density profile is almost spherically symmetric, but we can see a clear difference for the outflow beyond this radius for different angles with respect to the disc plane. While the shock front in the disc plane cannot escape the virial radius, the gas perpendicular to the disc plane gets ejected out of the galaxy.

We further illustrate the outflow dynamics in Figure 6.6. Depending on the angle, the outflow has typical velocities of 100–10000 $\mathrm{km\,s^{-1}}$ in our fiducial model. The bottom panel nicely shows how the outflow gets reaccelerated in the perpendicular direction beyond ~ 5 pc due to the exponentially decreasing ISM density. We show the time evolution of the shock structure in Figure 6.7. The outflow starts spherically symmetric, but develops an elongated shape, perpendicular to the disc plane after > 10 kyr. After ~ 1 Myr, the outflow has stopped in the disc plane in our fiducial model and keeps propagating along the poles.

To analyse the influence of the individual driving and restoring forces, we plot the different contribu-

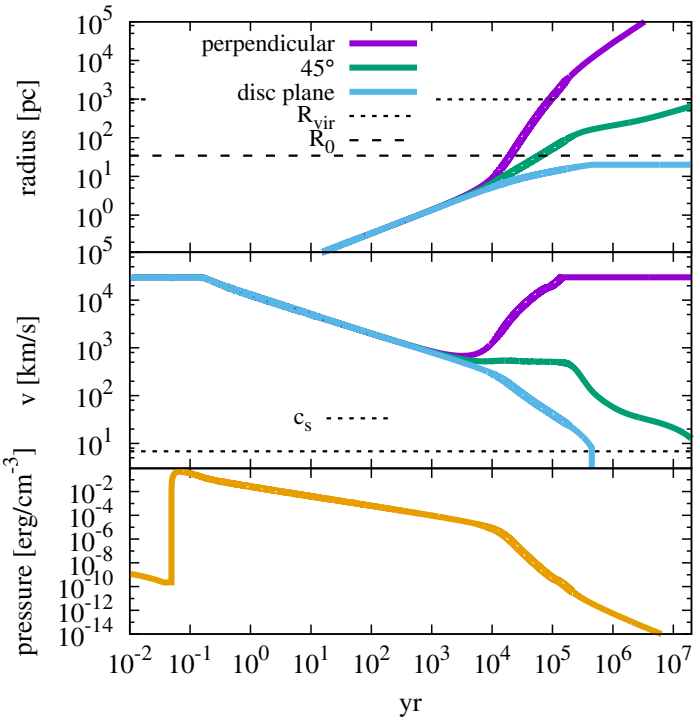


Figure 6.5: Time evolution of different quantities for our fiducial galaxy model. Top panel: position of the shock for different angles with respect to the disc. The virial and disc scaling radius are indicated by the black dashed/dotted lines. Middle panel: velocity for different angles and the wind becomes subsonic in the disc plane at around 10^6 yr. Lower panel: the pressure of the shocked wind rises rapidly once the gas can no longer cool efficiently.

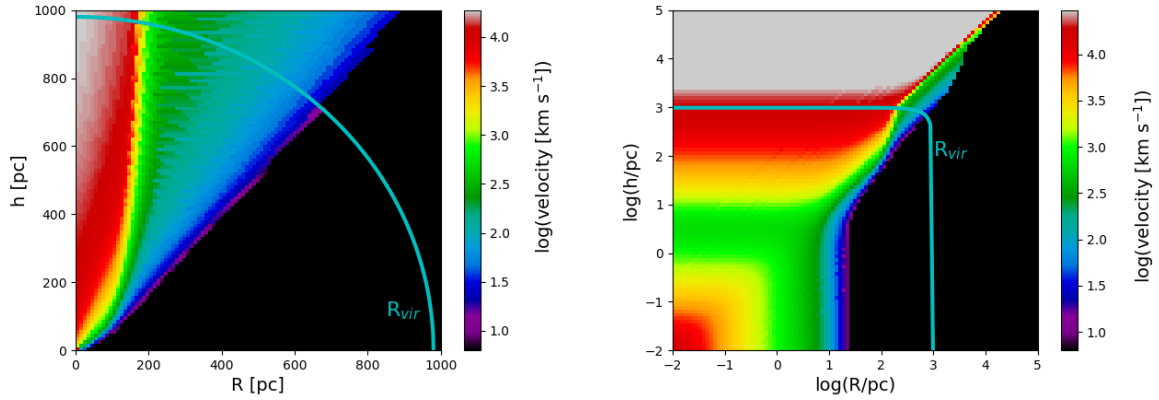


Figure 6.6: Velocity maps for our fiducial galaxy model with the galactic disc along the x-axis and the virial radius highlighted in cyan. The linear scaling in the left panel illustrates nicely the overall geometry and the logarithmic scaling in the right panel highlights different zones of interest: at small radii, the wind starts with a high velocity, but is decelerated by the increasing mass of the swept up gas. Whereas the velocity is a decreasing function with radius in the disc plane, the shock gets accelerated in the direction perpendicular to the disc beyond ~ 5 pc due to the strongly decreasing gas density in this direction.

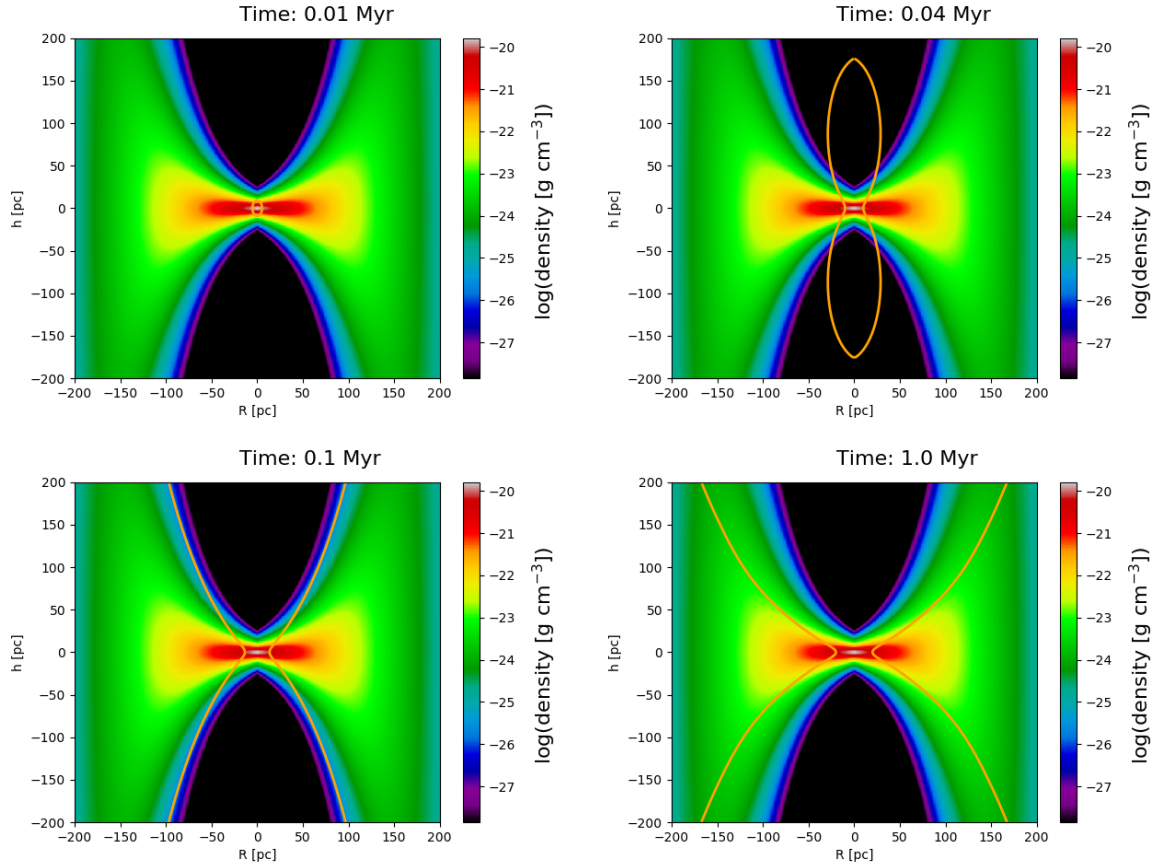


Figure 6.7: Time evolution of the shock front (orange) with the initial density distribution in the background. The shock front starts spherically symmetric but it quickly develops an elongated shape in the direction of least resistance, i.e. perpendicular to the disc plane. The virial radius for the galaxy is $\sim 1\,\mathrm{kpc}$.

tions to the effective acceleration on a shell element in Figure 6.8. The main forces are the acceleration by the AGN (either L/c or the pressure of the adiabatically expanding shell) and the inertia of the swept-up gas, which has to be accelerated ($\dot{M}_{\mathrm{shell}}(r)\dot{r}$). The gravity from the BH and DM, as well as the self-gravity of the swept-up gas are negligibly small on all scales of interest. The total acceleration changes sign several times: within $\lesssim 0.01\,\mathrm{pc}$ the wind accelerates the shell until the mass of the newly swept up material becomes significant. Within $\lesssim 5\,\mathrm{pc}$, the wind decelerates due to the inertia of the swept up gas. Perpendicular to the disc, the wind accelerates again at $\gtrsim 5\,\mathrm{pc}$ due to the decreasing gas density and hence smaller gas inertia. In the disc plane, the velocity is too small and the shock becomes subsonic. The deceleration and subsequent acceleration perpendicular to the disc plane around $\sim 5\,\mathrm{pc}$ can also be seen in Figure 6.6.

The contribution of IR and UV account for absorption in the UV and re-emission in the IR, following Ishibashi & Fabian (2015) with the IR and UV opacities $\kappa_{\mathrm{IR}} = 5\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ and $\kappa_{\mathrm{UV}} = 10^3\mathrm{cm}^2\,\mathrm{g}^{-1}$. Even for solar metallicity, these effects are subdominant compared to other contributions, but most importantly they are only relevant in the momentum-driven regime, where the momentum of the photons couples directly to the gas. See also Bieri et al. (2017) for a more detailed discussion of the contribution of different photon groups to AGN feedback.

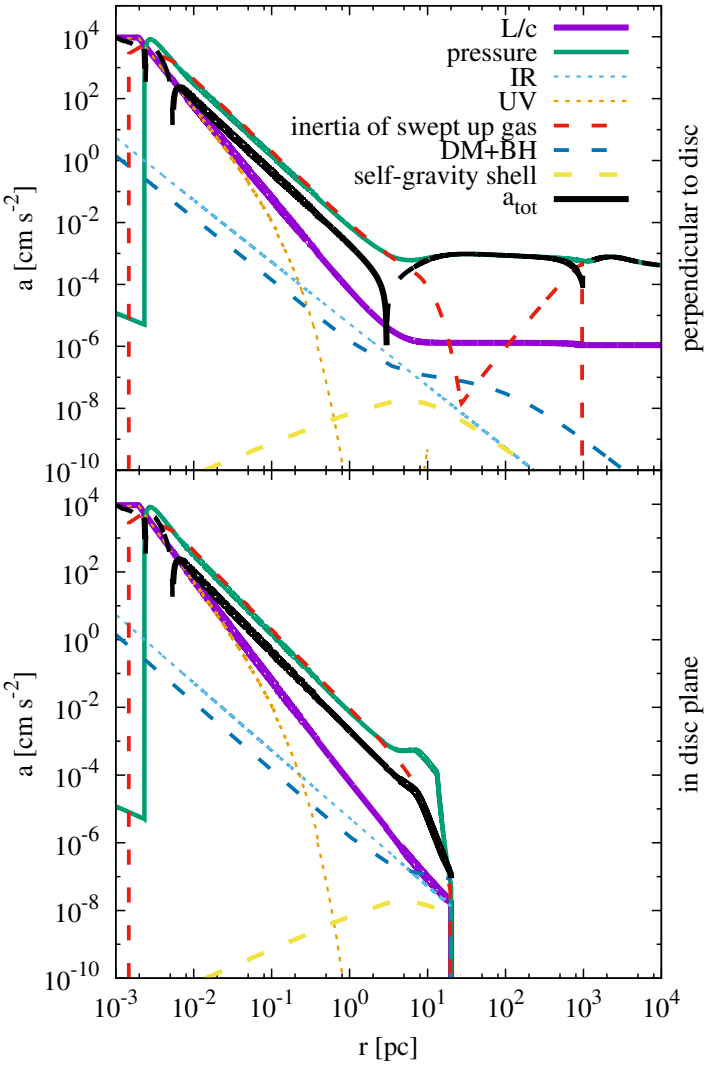


Figure 6.8: Accelerations as a function of radius for the different contributions perpendicular to the disc (top) and in the disc plane (bottom) for our fiducial galaxy model. Solid lines are outward and dashed lines are inward contributions. Dotted lines represent accelerations that were calculated a posteriori and are not included self-consistently in the dynamics. The black line represents the total acceleration. The dynamics is dominated by the equilibrium between acceleration due to the AGN and the inertia of the swept up ISM.

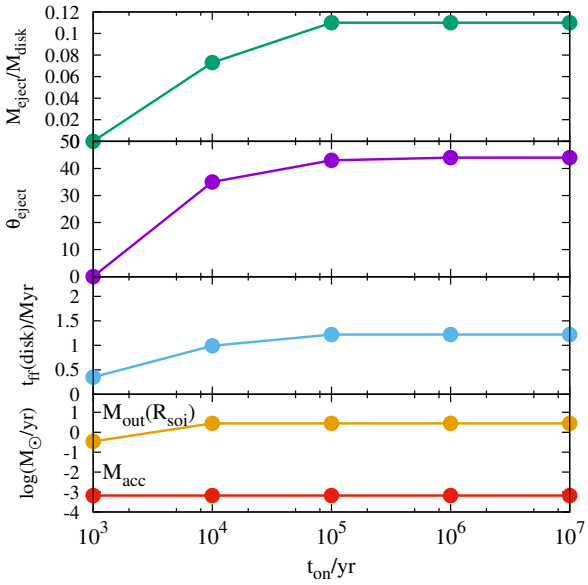


Figure 6.9: Dependence of different quantities on the lifetime of the AGN. The ability of the outflow to remove gas from the galaxy increases with $t_{\rm on}$ up to maximum values above $t_{\rm on} \geq 0.1\,{\rm Myr}$.

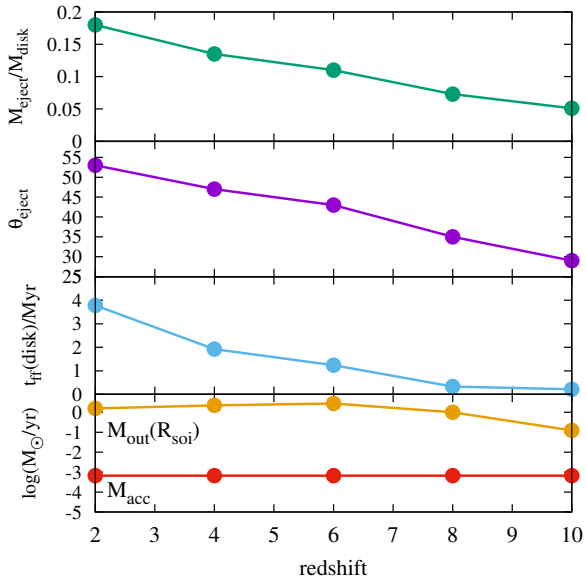


Figure 6.10: Dependence of different quantities on redshift. The decreasing efficiency with increasing redshift is due to the deeper gravitational potential at high redshift, because of the smaller virial radius.

6.3.2 Parameter study

The advantage of our analytical model is that we can easily explore large sets of parameters to study the dependence of the AGN-driven wind on e.g. the BH mass, redshift, or the AGN lifetime. In this section, we present different parameter studies and analyse their effect on the dynamics of the outflow. Some of the tested parameter combinations represent extreme scenarios and might be very rare in nature or not appear at all. We will also comment on the likelihood of the presented parameter choices.

AGN lifetime

We vary the time for which the AGN shines with a constant luminosity, $t_{\rm on}$, before we set its luminosity to zero, see Figure 6.9. For $t_{\rm on} > 0.1\,{\rm Myr}$ the dynamics does not change, or phrased differently, for our fiducial galaxy model it does not make a difference whether the AGN is active for 0.1 or 10 Myr. In addition, this illustrates that the most significant part of the momentum transfer happens in the first $\sim 10^5\,{\rm yr}$. If the AGN shines only for $t_{\rm on} \lesssim 10^3\,{\rm yr}$ it cannot eject any gas out of the galaxy. The mass outflow rate increases with time as the AGN-driven wind sweeps up more ISM and reaches its peak value $\sim 30\,{\rm kyr}$ after the AGN starts to shine. At this time, the asymmetry of the outflow starts to develop, the pressure-driving in the disc plane is less efficient because the internal energy can escape more easily perpendicularly to the disc plane, and the shock in the disc plane slows down. For $t_{\rm on} > 30\,{\rm kyr}$ the mass outflow rate is independent of the lifetime. Typical AGN lifetimes are of the order 10 Myr with variations of the accretion rate on time-scales as short as 1000 yr (Park & Ricotti 2011b; Sugimura et al. 2017; Negri & Volonteri 2017).

Redshift

In our fiducial model we focus on low-mass galaxies at $z = 6$, but our 2D approach can also be extrapolated to galaxies at other redshift and we show the redshift dependence of the AGN-driven outflow

in Figure 6.10. The AGN-driven outflow is generally less efficient at higher redshift due to the deeper gravitational potential; i.e., a halo at higher redshift has a smaller virial radius for the same M_h, due to the higher mean cosmic density. This trend would be amplified by a smaller gas mass fraction at lower redshift, which we do not include in our model: with cosmic time, gas is converted into stars and the remaining ISM exerts a weaker resistance against the outflow, which therefore can propagate further.

Initial wind velocity

The initial wind velocity is crucial, because it is the least well constrained parameter and directly influences the inner wind velocity, hence the Compton cooling time, and consequently the transition from momentum- to energy-driving (Figure 6.11). For $v_{\rm in} \leq 0.1c$ the choice of the initial wind velocity does not affect the long-term evolution and the dynamics on galactic scales. However, for values slightly larger than our fiducial value $v_{\rm in} = 0.1c$, the transition to energy-driving in the disc occurs at significantly later times (see Section 6.3.3 for a more detailed discussion).

We can observe another interesting feature: the later the transition from momentum- to energy-driving occurs, the more asymmetric the outflow will be in the late phase ($R_{\rm perp}/R_{\rm disc}$ at $t \gtrsim 10^6\,{\rm yr}$). Except for the tiny contribution of the shell self-gravity, the momentum-driven expansion in one direction does not depend on the outflow dynamics in another direction. This is in contrast to the energy-driven case, where the accelerating pressure depends on the volume enclosed by the shock front, which directly couples the outflow dynamics in different directions. Consequently, the energy-driven outflow develops an asymmetry once it encounters a path of least resistance, which makes the pressure fall and the acceleration in the disc plane smaller. In contrast, the momentum-driven outflow starts more spherical, even at larger radii, where the asymmetric density profile allows the outflow to propagate into a low-density directions. However, when the outflow now becomes energy-driven, with a higher momentum boost, the pressure does not continue to accelerate the outflow in the disc plane, because there is already a very well defined path of least resistance in the perpendicular direction. In addition, the volume enclosed by the shock front in the moment of the late transition is smaller than the volume of an outflow with an early transition to energy-driving at the same time. Hence, the pressure rises above the value of the pressure of the outflow with an early transition, compare $P(t \approx 10^5\,{\rm yr})$.

At the same time we can see from the time evolution of the momentum boost that this higher pressure (at $\sim 10^5\,{\rm yr}$) does not drive the acceleration in all directions to the same degree. It rather boosts the expansion perpendicular to the disc plane, for which the outflow has already paved a path of least resistance. We confirm that the momentum boost depends on the confinement time, i.e. the period over which the outflow is enclosed by gas of roughly the same density.

AGN luminosity

We have verified that the outflow dynamics on all relevant scales is determined solely by the AGN luminosity, i.e. the product of Eddington ratio and BH mass and not their individual values, see Figure 6.12. A small BH with a high Eddington ratio is therefore equivalent to a more massive BH with a smaller Eddington ratio, keeping in mind that high Eddington ratios are much less abundant then low or intermediate values (Habouzit et al. 2017). The indicated freefall times yield the time that gas needs at least to fall back towards the centre once the AGN is off. It scales linearly with the radius out to which the gas is pushed in the disc plane with $t_{\rm ff} = 5\,{\rm Myr}(R/R_0)$. In Section 6.3.3 we discuss the origin of the transition to momentum-driving at higher AGN luminosities.

We compare the models for different BH masses, at the same halo mass, fixing the Eddington ratio

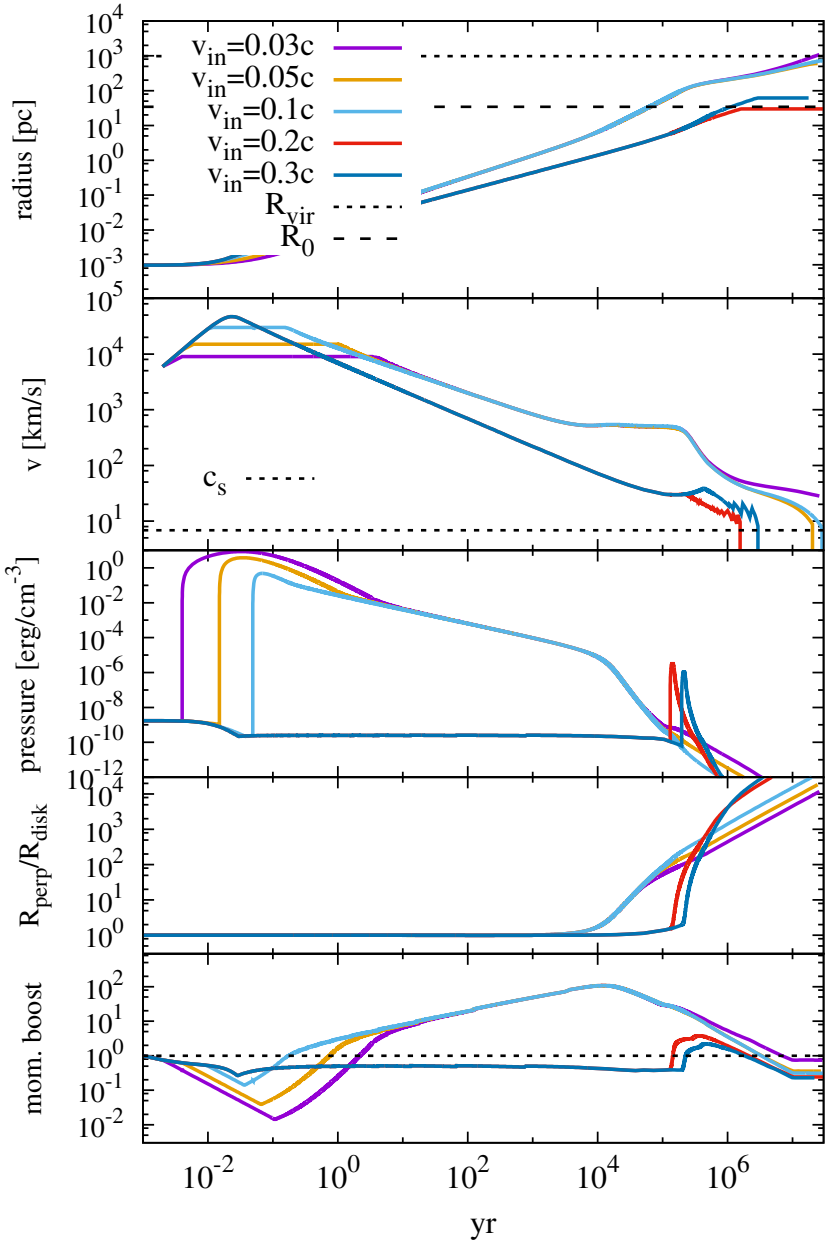


Figure 6.11: Time evolution of the AGN-driven outflow for different values of the initial wind velocity for the direction $\theta = 45°$. For values above our fiducial parameter of $v_{\rm in} = 0.1c$ the transition to energy-driving occurs at much larger radii (red, dark blue). Also, the evolution of the asymmetry of the outflow is remarkable: while the shock is more spherical in the momentum-driven phase, it becomes more asymmetric after transition to energy-driving than the outflow that is energy-driven from early on (see text).

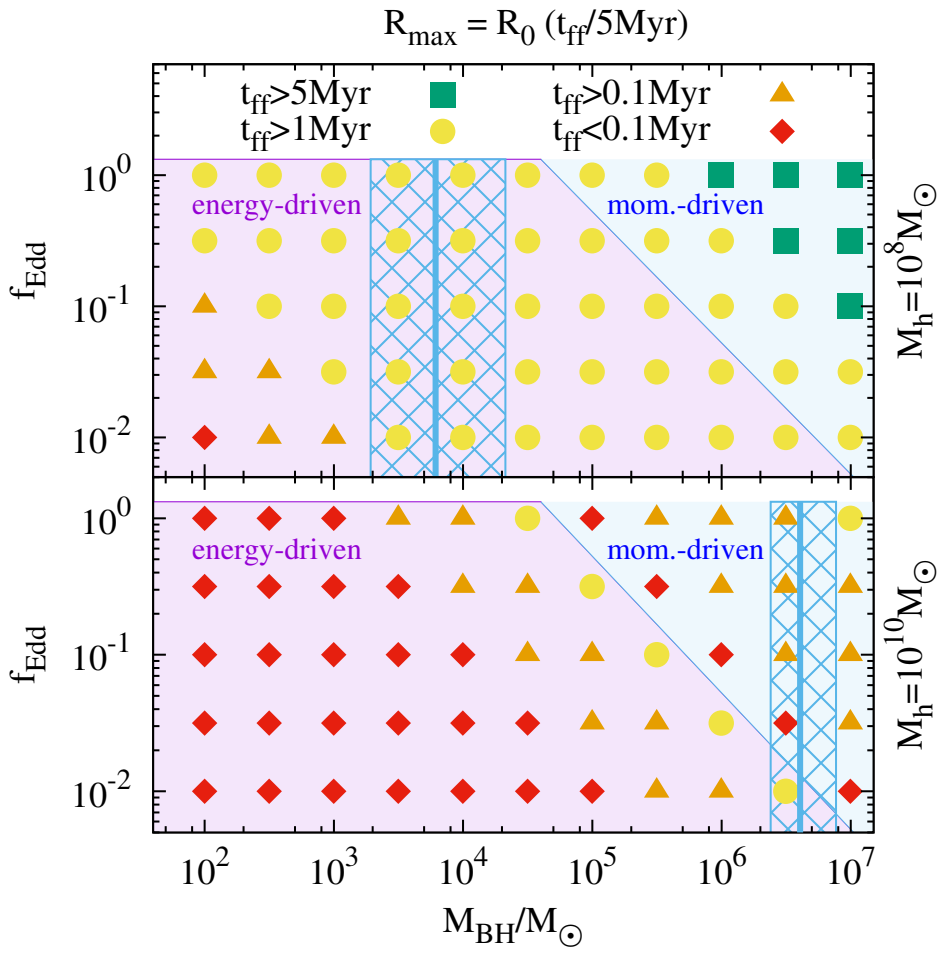


Figure 6.12: We compare the outflow, quantified by $t_{\rm ff}$ in the disc plane, for different combinations of the Eddington ratio and the BH mass for two different halo masses: $10^8\,{\rm M}_\odot$ (top panel) and $10^{10}\,{\rm M}_\odot$ (bottom panel). The blue vertical lines and hashed regions indicate the typical $M_{\rm BH}$ for these halo masses, based on the M–σ relation (Gültekin et al. 2009). The outflow dynamics depends only on the product $f_{\rm Edd}M_{\rm BH}$ and we clearly see two ranges of energy- and momentum-driven outflows, separated by a luminosity of $\sim 10^{43}\,{\rm erg\,s^{-1}}$, as we will discuss in Section 6.3.3. The same AGN in a two orders of magnitude more massive halo generates fallback times of the gas that are about one order of magnitude shorter.

at $f_{\rm Edd} = 0.3$, in Figure 6.13 and 6.14. We also observe an interesting trend for the momentum boost, where we observe peak values of $\gtrsim 100$, with the lowest mass black BH generating the highest momentum boosts. Phrased differently, a lower mass BH is more efficient in converting its momentum and energy input into momentum of the outflow. For all considered BH masses, the AGN can drive outflows perpendicular to the disc beyond $R_{\rm vir}$, but only for $M_{\rm BH} \gtrsim 10^5\,{\rm M}_\odot$ the AGN can eject gas at $\theta \geq 45°$, whereas the wind powered by less massive BHs is not strong enough and it becomes subsonic before reaching $R_{\rm vir}$ in the diagonal direction.

The dependence of the outflow on the Eddington ratio for different BH masses, at fixed halo mass, is shown in Figure 6.12. For small BHs ($M_{\rm BH} \leqslant 10^3\,{\rm M}_\odot$) outflows cannot push the gas beyond disc radius even for $f_{\rm Edd} = 1$. A large range of the parameter space allows for outflows to affect gas in the whole disc (yellow and green symbols in the upper panel of Figure 6.12), possibly regulating the time over which an AGN has to remain off before new gas flows in to feed a new accretion episode as well as star formation in the disc.

To push gas beyond R_0, which is still only $\sim 3\%$ of the virial radius, a BH should be as massive as 1% of the halo mass and have $f_{\rm Edd} = 1$, or 10% of the halo mass and have $f_{\rm Edd} > 0.1$. Note that these ratios between BH and halo mass are very large. The virial velocity of a halo, i.e. the circular velocity at the virial radius, is given by $V_{\rm vir} = 200\,{\rm km\,s^{-1}}(M_h/10^{11}\,{\rm M}_\odot)^{1/3}$ at $z = 6$ (Ferrarese 2002; Wyithe & Loeb 2002; Volonteri et al. 2011). Assuming that $\sigma = V_{\rm vir}/\sqrt{3}$ and the $z = 0$ scaling between BH mass and σ (Gültekin et al. 2009) yields

$$M_{\rm BH} = 6000\,{\rm M}_\odot \left(\frac{M_h}{10^8\,{\rm M}_\odot}\right)^{1.41} \tag{6.46}$$

for the typical BH mass in a halo of mass M_h.

Simulations suggest that to correctly reproduce the luminosity function of AGN only a small fraction ($\lesssim 25\%$) of high-z BHs in galaxies with stellar mass $< 10^{10}\,{\rm M}_\odot$ is expected to have Eddington ratios above 0.01, with the majority of BHs accreting at $f_{\rm Edd} < 10^{-4}$ (Habouzit et al. 2017). For these more realistic parameter combinations, with BHs on the M-σ relation and small Eddington ratios, we expect only short fallback times of the order $\lesssim 1\,{\rm Myr}$. Therefore, BHs can grow almost continuously and the accretion phases are interrupted by comparably short episodes of AGN feedback: a $\sim 10^4\,{\rm M}_\odot$ BH in a $10^8\,{\rm M}_\odot$ halo has a recovery time of $\sim 1\,{\rm Myr}$, indicating that accretion is not interrupted for long. A $\sim 10^6\,{\rm M}_\odot$ BH in a $10^{10}\,{\rm M}_\odot$ halo creates even shorter recovery times, making BH growth easier in higher mass systems.

The role of AGN outflows in driving large amounts of gas out of halos appears limited, if they are the only source of feedback. A viable possibility is that AGN feedback needs SN feedback as a precursor. Costa et al. (2014) and Prieto et al. (2017) show that AGN feedback is inefficient without the aid of SN feedback: SN winds heat the cold gas in the galaxy, creating a rarefied environment where energy injection from AGN feedback can easily accelerate the gas. In the next section we discuss how an outflow can still regulate the duty cycle of BHs.

In the following, we do not vary the BH mass and the Eddington ratio independently, but keep the Eddington ratio at $f_{\rm Edd} = 0.3$ and vary the BH mass and consequently the AGN luminosity. The results for a different Eddington ratio can be rescaled accordingly.

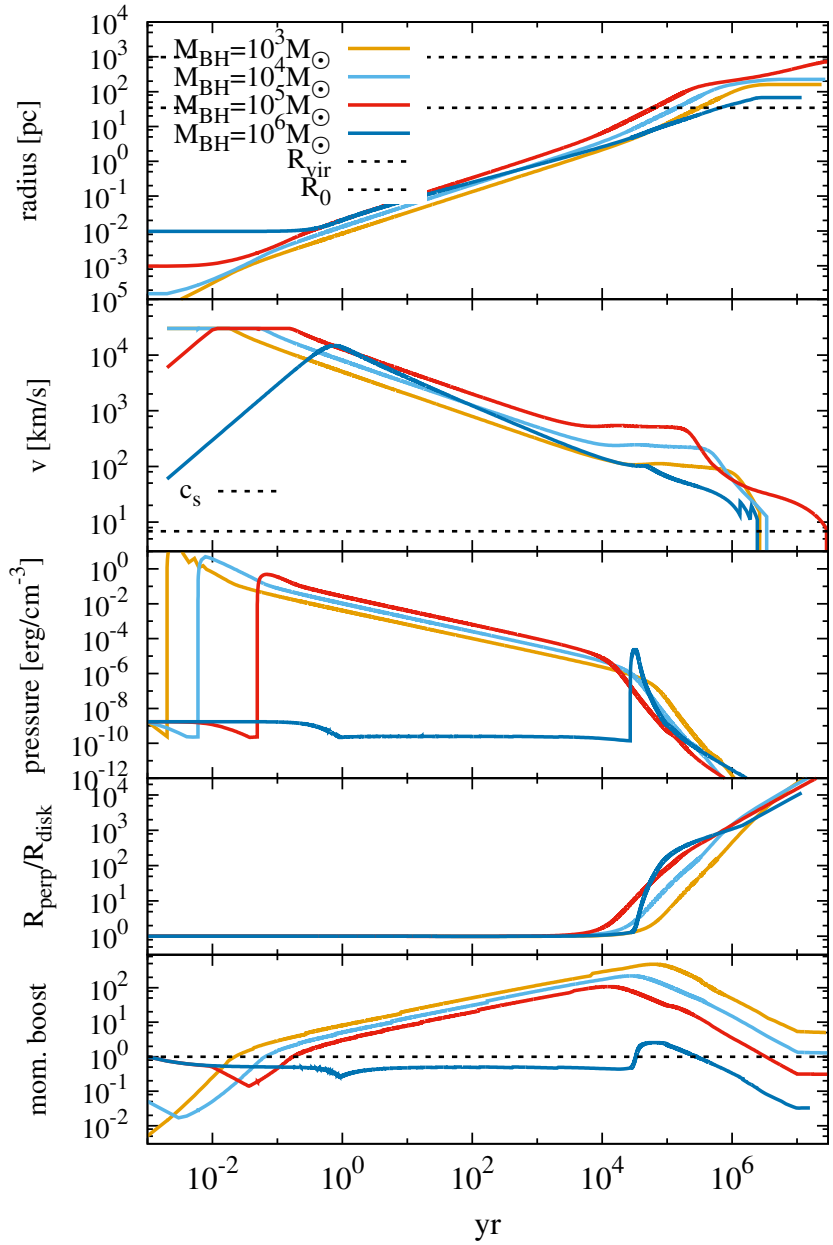


Figure 6.13: Time evolution of the AGN-driven outflow for different black hole masses, i.e. different AGN luminosities for a fixed $f_{\rm Edd} = 0.3$ with $M_h = 10^8\,{\rm M}_\odot$. The models with $M_{\rm BH} \leq 10^5\,{\rm M}_\odot$ experience an early transition from momentum- to energy-driving (see sudden jump in the pressure) and the model with $M_{\rm BH} = 10^6\,{\rm M}_\odot$ experiences a later transition (see also section 6.3.3 for more details). The radius and velocity represent the evolution at 45°.

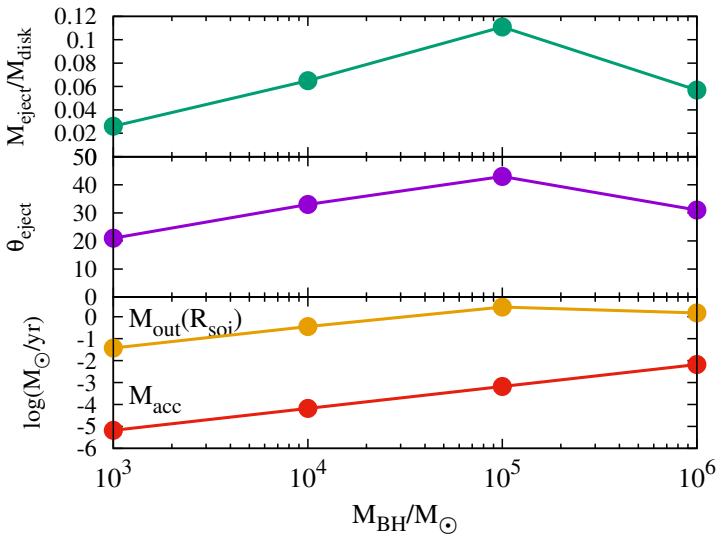


Figure 6.14: Dependence of the gas ejection efficiency and the mass outflow rate on the BH mass for $M_h = 10^8\,\mathrm{M}_\odot$ and $f_\mathrm{Edd} = 0.3$. The ability to eject gas from the galaxy has a local peak at $M_\mathrm{BH} \approx 10^5\,\mathrm{M}_\odot$, because for higher BH masses, i.e. AGN luminosities, the outflow is momentum-driven in the disc (see Section 6.3.3).

Halo mass

We compare the outflow efficiency for different combinations of the halo and BH mass in Fig 6.15. The general trend is that the outflow becomes more powerful with a higher AGN luminosity and lower halo mass. However, there is a discontinuity of the trend at $M_\mathrm{BH} \approx 10^5\,\mathrm{M}_\odot$, which we will discuss in more detail in the next section. We further note that contours of equal efficiency are steeper than linear in the $\log(M_h)$-$\log(M_\mathrm{BH})$ plane. Even for the most extreme scenario of a $\sim 10^4\,\mathrm{M}_\odot$ BH accreting at 30% Eddington in a $10^6\,\mathrm{M}_\odot$ halo, the outflow hardly pushes the gas beyond R_0 in the disc midplane. This indicates that the gas recovery times after AGN-driven winds are very short, of the order $\sim 1\,\mathrm{Myr}$.

6.3.3 Momentum-driven outflows for AGN luminosities above 10^{43} erg/s

We see in Figure 6.16 that the transition from an initially momentum-driven to an energy-driven outflow occurs at different radii or not at all depending on the halo and black hole mass for a fixed Eddington ratio. The nature of this transition can be better seen in Figure 6.17, where we show the time evolution of the pressure in a halo with $M_h = 2 \times 10^9\,\mathrm{M}_\odot$ and an outflow powered by BHs of different mass. For $M_\mathrm{BH} \leq 10^5\,\mathrm{M}_\odot$ the transition to energy-driving occurs early (note that this is not our fiducial galaxy model, for which the transition occurs already at $t < 0.1\,\mathrm{yr}$). For BH masses of the order $10^5 - 10^6\,\mathrm{M}_\odot$ the outflow is always momentum-driven and for even higher BH masses of $\gtrsim 10^6\,\mathrm{M}_\odot$ we see a late transition to energy-driving (small peak in the red curve around $t \approx 10^6\,\mathrm{yr}$). This trend can also be seen from a comparison of the relevant time-scales in Figure 6.18. The three different regimes can be understood as follows: for a low BH mass (i.e. AGN luminosity) we see an early transition to energy-driving, because of the initial smaller acceleration by the AGN (green, light-blue). For significantly higher AGN luminosities (dark-blue, purple) the outflow remains momentum-driven and reaccelerates beyond $\sim 5\,\mathrm{pc}$ (compare Fig. 6.6), which decreases the flow-time and makes the outflow energy-driven. The intermediate cases (orange, red) also remain momentum-driven, but become subsonic before they reach $\sim 5\,\mathrm{pc}$.

We can understand these transitions at systematically different radii based on simple analytical arguments. The Compton cooling time as a function of radius r is given by Eq. 6.39 or for typical

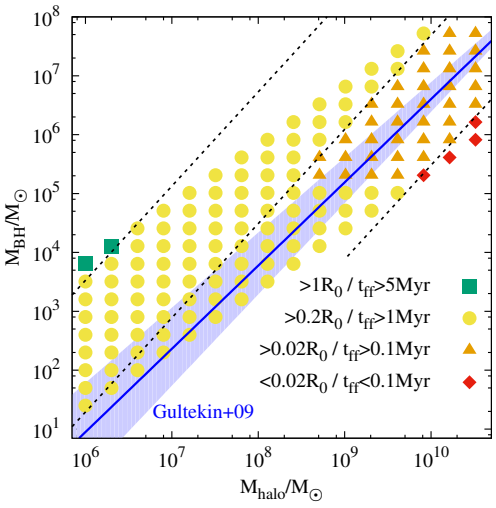


Figure 6.15: Efficiency of ejecting gas and preventing further accretion for different combinations of the BH and halo mass, quantified by the maximum radius the outflow reaches in the disc plane. The black dashed lines represent $M_{\rm BH} \propto M_h^{1.6}$, illustrating thresholds of constant maximal radii. The blue line and shaded area show the expected BH masses and their error margins (Gültekin et al. 2009). Note that this relation is valid for local and more massive galaxies and should rather guide the eye and indicate typical host environments for a given BH mass than be used for a quantitative comparison.

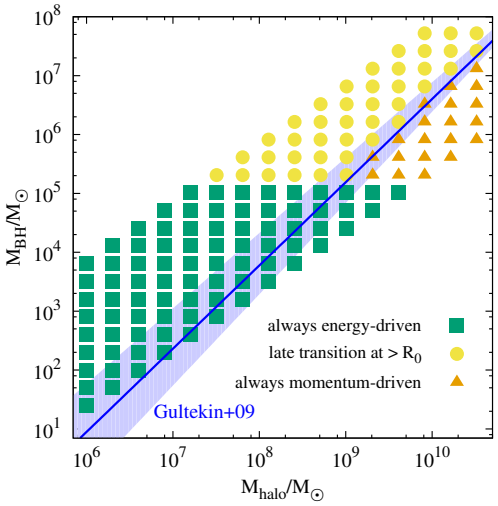


Figure 6.16: Parameter study of the transition radius in the disc from momentum- to energy-driven outflows at 30% Eddington. We can identify three different regimes: the transition occurs at very small radii close to $R_{\rm min}$ (green), the transition occurs at larger radii $\gtrsim R_0$ (yellow), or there is no transition at all and the outflow is always momentum-driven (orange). The last case also includes scenarios where the transition occurs at $> R_{\rm vir}$, which is irrelevant for our study.

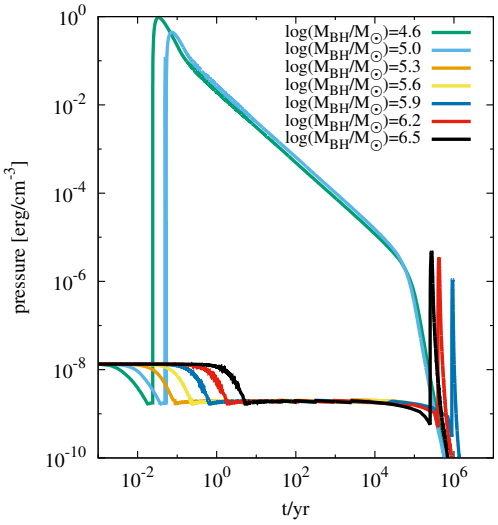


Figure 6.17: Time evolution of the pressure in a halo with a mass of $2 \times 10^9\,{\rm M}_\odot$ for BH masses spanning 2 orders of magnitude. This represents a vertical slice in Figure 6.16. The transition to an energy-driven outflow occurs at different times or not at all, depending on the BH mass, i.e. on the AGN luminosity.

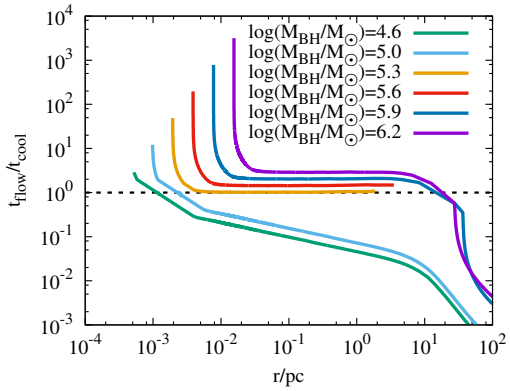


Figure 6.18: For a fixed halo mass of $M_h = 2 \times 10^9\,{\rm M}_\odot$ we plot the ratio of the flow and the cooling time of the AGN-driven outflow as a function of radius (position of the shell) for different BH masses in the disc plane. All outflows start momentum driven, but if the flow time becomes shorter than the cooling time, Compton cooling can no longer cool the shock efficiently and the shock becomes energy-driven. However, above $M_{\rm BH} \gtrsim 10^5\,{\rm M}_\odot$ the cooling time is always shorter than the flow time in the disc and both have the same scaling with radius of $t \propto r^2$. For even higher AGN luminosities, a late transition to energy-driving occurs on galactic scales.

parameters:

$$t_{\mathrm{compton}} = 1.4\,\mathrm{kyr}\left(\frac{f_{\mathrm{Edd}}}{0.3}\right)^{-1}\left(\frac{M_{\mathrm{BH}}}{10^6\,\mathrm{M}_\odot}\right)^{-1}\left(\frac{v_{\mathrm{in}}}{0.1c}\right)^{-2}\left(\frac{r}{1\,\mathrm{pc}}\right)^{2}. \tag{6.47}$$

To derive the relevant flow time, we have to make several simplifying assumptions. If the transition to energy-driving occurs, it occurs at $< 1\,\mathrm{pc}$ (compare Figure 6.18), which is within the sphere of influence and the density in this inner part can be assumed to be constant. Furthermore we neglect the gravity of the BH. This assumption might seem to be in contradiction to being within the BH's sphere of influence, but the gravitational force is proportional to the mass of the swept up ISM, which is negligibly small at the radii of interest (compare Figure 6.8). The equation of motion is then given by

$$\frac{\mathrm{d}}{\mathrm{d}t}(M_{\mathrm{shell}}v) = \frac{L}{c}, \tag{6.48}$$

with the shell's velocity $v = \dot{r}$ and the shell mass $M_{\mathrm{shell}} = 4/3\pi r^3\rho_0$. The central gas density ρ_0 can be derived from the exponential disc profile as

$$\rho_0 = \frac{\Sigma_0}{2H_0} = \frac{G}{2\pi c_s^2}\frac{(m_d M_{\mathrm{halo}})^2}{(\lambda R_{\mathrm{vir}})^4}, \tag{6.49}$$

where $H_0 = H(R = 0)$ is the disc scale height in the centre. The differential equation can be solved by a function of the form

$$r(t) = r_0\left(\frac{t}{t_0}\right)^{0.5}$$
$$v(t) = \frac{r_0}{2}(t_0 t)^{-0.5},$$

which yields directly the same scaling with radius for the flow time as for the cooling time:

$$t_{\mathrm{flow}} = \frac{r}{v} = 2t \propto r^2. \tag{6.50}$$

We note again that this result is only valid for the inner part of the disc, where the density can be assumed to be constant and the mass of the swept up gas is still negligible for the gravity. For the flow time as a function of radius we obtain

$$t_{\mathrm{flow}} = \sqrt{\frac{8\pi c\rho_0}{3L}}r^2 \propto m_d\lambda^{-2}(1+z)^{3/2}f_{\mathrm{Edd}}^{-1/2}M_{\mathrm{BH}}^{-1/2}R^2. \tag{6.51}$$

Or expressed with fiducial values and fixed $m_d = \lambda = 0.05$ at $z = 6$:

$$t_{\mathrm{flow}} = 2.2\,\mathrm{kyr}\left(\frac{f_{\mathrm{Edd}}}{0.3}\right)^{-1/2}\left(\frac{M_{\mathrm{BH}}}{10^6\,\mathrm{M}_\odot}\right)^{-1/2}\left(\frac{r}{1\,\mathrm{pc}}\right)^{2}. \tag{6.52}$$

Setting it equal to the cooling time we can conclude that the outflow is always momentum-driven if the black hole and Eddington ratio fulfil

$$\frac{f_{\mathrm{Edd}}}{0.3}\frac{M_{\mathrm{BH,crit}}}{4\times10^5\,\mathrm{M}_\odot} > \left(\frac{v_{\mathrm{in}}}{0.1c}\right)^{-4}\left(\frac{1+z}{7}\right)^{-3}, \tag{6.53}$$

or expressed via the AGN luminosity

$$L_{\rm AGN} > 10^{43}\,{\rm erg\,s^{-1}} \left(\frac{v_{\rm in}}{0.1c}\right)^{-4} \left(\frac{1+z}{7}\right)^{-3}. \tag{6.54}$$

Note the independence on the halo mass and the strong dependence on the not well-constrained initial velocity of the inner wind $v_{\rm in}$. Uncertainties in this parameter can change the resulting transition luminosity by up to three orders of magnitude. The analytically derived BH transition mass, $4 \times 10^5\,{\rm M_\odot}$, is close to the one obtained in the 2D simulation of $\sim 2 \times 10^5\,{\rm M_\odot}$. The small difference is related to the necessary approximations in order to analytically solve the equation of motion. The strong redshift dependence indicates that the transition mass above which a BH can no longer drive an efficient energy-driven wind in the disc is higher at low redshift, caused by the redshift dependence of the central density.

The parameter dependences can be explained as follows: a higher Eddington ratio $f_{\rm Edd}$ yields more photons by the AGN, which boost the cooling via inverse Compton scattering. At lower redshift, the gravitational potential is shallower for the same halo mass, which makes the shell velocity higher, the flow time smaller, and hence requires more efficient cooling to sustain a momentum-driven outflow. The initial wind velocity is directly proportional to the post-shock temperature (Eq. 6.26) and hence defines the internal energy of the shocked wind.

Momentum-driven outflows around the threshold $L_{\rm AGN} \approx 10^{43}\,{\rm erg\,s^{-1}}$ or slightly above generate a smaller momentum boost (Figure 6.13), are less powerful in ejecting gas out of the galaxy (Figure 6.14), and drive the gas in the disc plane to smaller maximal radii (Fig 6.12, 6.15). However, more luminous AGNs with $M_{\rm BH} \approx 10^7\,{\rm M_\odot}$ and $f_{\rm Edd} = 1.0$ in $10^8\,{\rm M_\odot}$ halos are the only systems that create recovery times of the gas in excess of 5 Myr (Figure 6.12) via momentum-driving. However, such a combination of Eddington ratio, BH and halo mass is very unlikely.

6.3.4 Gas ejection perpendicular to the disc

So far, we have focussed on the gas dynamics in the disc plane and have seen that AGN-driven winds in 2D push the gas in the disc plane at most to about the scale radius R_0. If we instead focus on the gas dynamics and outflow velocities at higher inclination, we can study the wind on galactic scales and relate it to observations of gas dynamics in high redshift galaxies. In Figure 6.19, we compare the outflow velocities at the virial radius for different angles with respect to the disc. For better comparison, we normalise the velocities to the virial velocity of the corresponding halos,

$$V_{\rm vir} = 20\,{\rm km\,s^{-1}} \left(\frac{M_h}{10^8\,{\rm M_\odot}}\right)^{1/3}. \tag{6.55}$$

A low halo mass and a high AGN luminosity generate outflows with higher velocities. As before, this general trend is interrupted by the transition between energy- and momentum-driving around a BH mass of $10^5\,{\rm M_\odot}$. Above this threshold, the outflow is momentum-driven in the disc and a part of the post-shock thermal energy is radiated away. This reduces the pressure and therefore the adiabatic acceleration in all directions.

There is a large parameter range, for which gas at $\theta \geq 45°$ cannot escape the halo (red diamonds). In contrast, close to the disc normal, gas is ejected with high velocities of $> 10^4\,{\rm km\,s^{-1}}$. AGN-driven outflows eject gas out of the galaxy and the direction-dependent velocities at $R_{\rm vir}$ range from $\lesssim 100\,{\rm km\,s^{-1}}$ to mildly relativistic velocities close to the disc normal, in agreement with observations (Chartas et al. 2002; Pounds et al. 2003; Cappi 2006; Feruglio et al. 2010; Rupke & Veilleux 2011; Aalto et al. 2012;

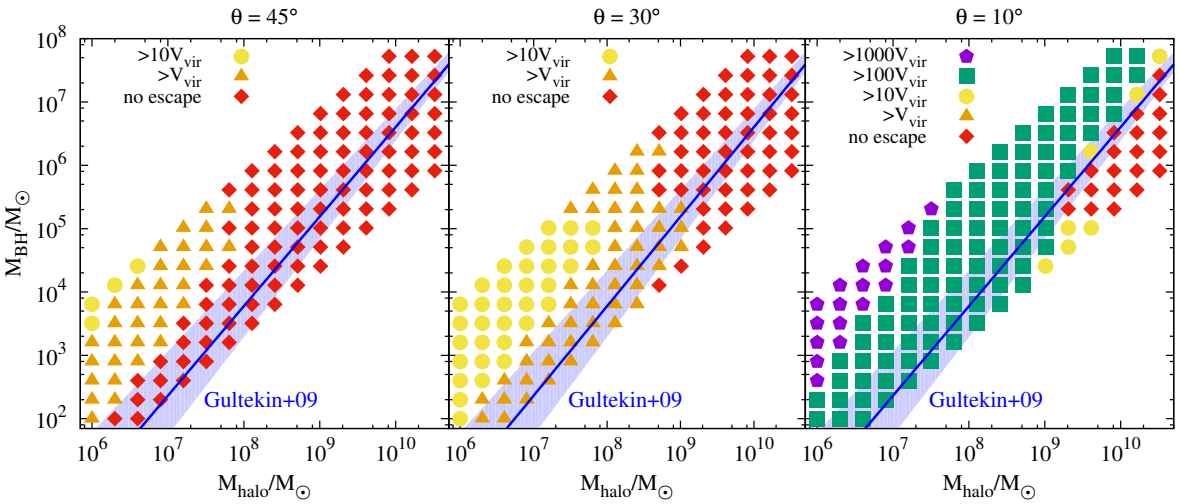


Figure 6.19: Outflow velocities at the virial radius, normalised to the virial velocity for different angles with respect to the disc normal. All models are with $f_{\rm Edd} = 0.3$ and the general trend is that more massive BHs in less massive galaxies create higher outflow velocities, close to the disc normal even in excess of $10^4\,{\rm km\,s^{-1}}$. However, there is also a wide range of models that can not eject gas at $\theta > 45°$ beyond the virial radius.

Gofford et al. 2013; Cicone et al. 2014). The subset of models that can not even eject gas at $\theta = 10°$ (red diamonds in the right panel) are the models where the outflow is always momentum-driven (compare Figure 6.16).

6.3.5 Comparison to spherical case

In this section we highlight the importance of a 2D approach by comparing it directly to the solutions obtained in a 1D model. We assume that the gas in the 1D model follows the distribution of the DM, which we describe by a Hernquist profile (Eq. 6.12). We first compare the general dynamics in our fiducial model in Figure 6.20. The 1D outflow is momentum-driven for a longer time and therefore starts with smaller velocities. However, once the 1D outflow becomes energy-driven, it ejects all the gas out of the virial radius, whereas the 2D model can only eject a smaller fraction of the total gas. We further quantify this effect in Fig 6.21. In a spherical model the ejected mass is either 0% for $t_{\rm on} \leq 10^5\,{\rm yr}$ or 100% for $t_{\rm on} > 10^5\,{\rm yr}$, whereas in our 2D model the ejected mass increases gradually with $t_{\rm on}$, and reaches saturation at 11% for $t_{\rm on} \gtrsim 10^5\,{\rm yr}$. In this regime, the 1D model overestimates the efficiency of an AGN-driven outflow by one order of magnitude in our fiducial galaxy model.

In Figure 6.22 we plot the efficiency of the 1D model in ejecting gas out to a certain radius. To quantify the efficiency of the outflow in the 1D Hernquist profile we apply the same scaling radius, R_0, as for 2D exponential disc (Eq. 6.3) to check if the outflow can reach at least this radius. The enclosed gas mass within R_0 in both scenarios is roughly the same with less than a factor two difference. There is a large parameter range for which 1D models can eject all the gas out of a galaxy. This range is of special interest, because it encloses halo masses of $\sim 10^7 - 10^9\,{\rm M_\odot}$, which are typical masses of dwarf galaxies at low redshift and of high redshift galaxies, which might host the first massive BH seeds. Above a certain halo mass of $\sim 10^9\,{\rm M_\odot}$, which corresponds to a typical BH mass of $10^5\,{\rm M_\odot}$, the AGN (at 30% Eddington) is not powerful enough to eject gas out of the galaxy.

We compare these results to Park et al. (2016) and Pacucci et al. (2015c), who use 1D radiation-hydrodynamics simulations to study the effect of radiative feedback on the growth of BHs in small halos.

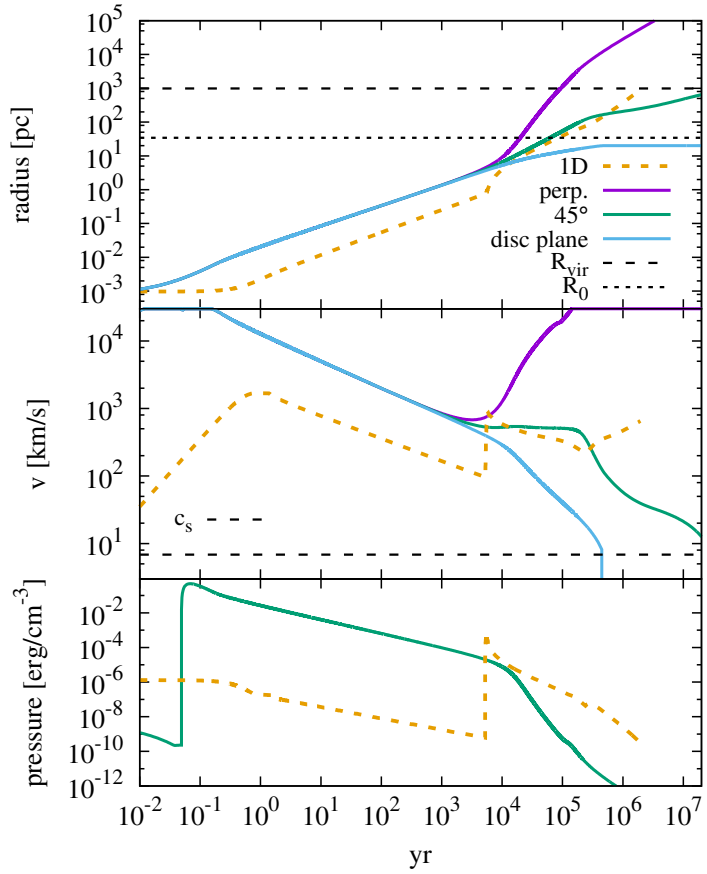


Figure 6.20: Same as Figure 6.5, but with the additional solution of a 1D model (orange, dotted), where we distribute the disc gas mass following a Hernquist profile. The transition to momentum-driving occurs at later time and larger radius, which is partially caused by the different radial density profiles for the Hernquist profile in 1D and an exponential disc in 2D. In the spherical model, all gas is ejected beyond $R_{\rm vir}$.

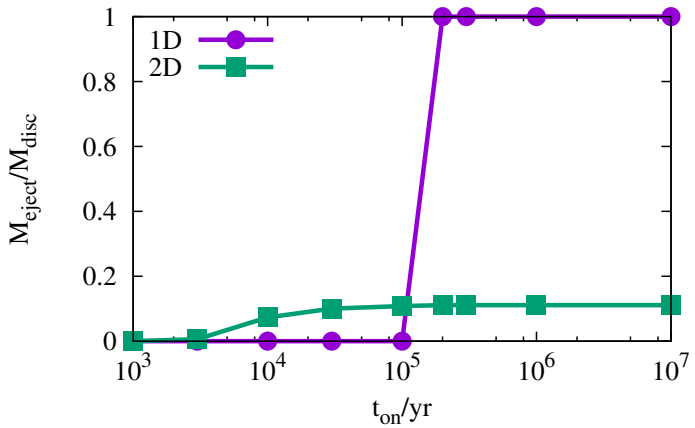


Figure 6.21: Fraction of the disc mass that is ejected by the AGN-driven wind as a function of the time over which the AGN shines with a constant luminosity (fiducial model with $f_{\rm Edd} = 0.3$). In the 1D model either no or all the gas is ejected out of the virial radius and this transition occurs at $t_{\rm on} \approx 10^5$ yr. In our 2D model we see a gradual rise of the ejected mass until it reaches the final value of $\sim 10\%$ at $t_{\rm on} \gtrsim 10^5$ yr. This demonstrates the strength and importance of a 2D treatment of AGN-driven feedback, compared to a simplistic 1D model.

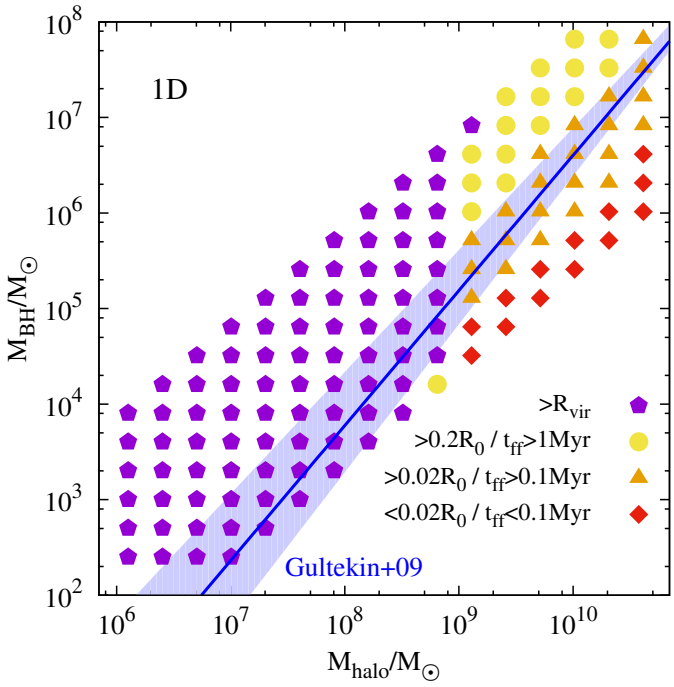


Figure 6.22: Maximum radius the AGN-driven wind reaches in spherical symmetry for different combinations of the BH and halo mass (compare to the 2D results in Figure 6.15). The purple pentagons in this plot indicate complete gas ejection out of the galaxy.

They show that light BHs of $\sim 100\,M_\odot$ cannot grow efficiently via radiation-regulated accretion. Pacucci et al. (2015c) further derive an analytical expression for this critical BH mass, which is in the range 10–$10^6\,M_\odot$, depending on the accretion scenario and the host properties. In our 1D comparison model, we find a comparable mass range of $M_{\rm BH} \approx 10^2 - 10^7\,M_\odot$, where the BH can completely eject the gas out of a host halo ($M_h \lesssim 10^9\,M_\odot$) and hence prevent rapid mass growth of the BH. For a fixed halo mass and Eddington ratio in two dimensions, a higher mass BH generates stronger outflows and ejects the gas to larger radii. This makes lower mass BHs less efficient in stopping their own gas supply and hence more susceptible for mass growth at high duty cycles. This is the inverse trend of Pacucci et al. (2015c) and Park et al. (2016), who find low mass BHs to be more efficient in stopping their own gas supply. These differences are likely caused by different assumptions on what drives the outflow: Pacucci et al. (2015c) and Park et al. (2016) study the direct influence of radiation pressure from the BH on the accretion process, while we model AGN-driven winds. For instance, the critical mass for radiation pressure to significantly change the mass accretion rate (Pacucci et al. 2015c) and reduce growth is related to the significance of the Eddington limit. The inflow rate is determined by the halo properties. If a given halo provides the same amount of gas inflow to a small black hole or a large black hole, it will be the smaller black hole that will reach the Eddington limit first, having its growth stunted.

We note that the differences between the 1D and 2D model presented in this section might partially depend on the different scaling of the density with radius and hence on the different radii where the outflow becomes energy-driven. However, the qualitative differences between the 1D and 2D model remain still valid beyond these differences in the density profile.

6.4 Discussion

We have developed a new 2D analytical model of AGN-driven outflows and present the results for various galaxies and a range of typical conditions. The additional dimension with respect to previously existing 1D models allows to account for the direction-dependent density profiles. This has significant consequences on the maximally ejected gas mass fraction, on the momentum boost, and on the condition to stop further accretion on the central BH.

6.4.1 Advantage of 2D

With a more realistic disc-profile we find the outflow to propagate preferentially towards low density regions. In the momentum-driven regime this is simply due to the lower column densities perpendicular to the disc and hence the lower gas masses that have to be accelerated. In the energy-driven regime, however, the evolution of the entire confined volume is relevant for the acceleration of the shock. Once the energy-driven shock has created a chimney perpendicular to the disc, the pressure and internal energy can escape and the driving in the disc plane is significantly reduced.

Naturally, we find that an AGN drives the gas out to different radii for different angles, before the shock front becomes subsonic or crosses the virial radius. In 1D models there is only one threshold AGN luminosity above which all the ISM is ejected and below which the AGN wind is not efficient enough. For none of our models the AGN can eject all the gas out of the galaxy.

If we define the criteria to stop further accretion on to the SMBH via the minimum time that the swept up gas needs at least to fall back to the centre, or equivalently that gas is swept out at least to a given radius in the disc plane, we find that this condition results in a proportionality of $M_{\rm BH} f_{\rm Edd} \propto M_h^{1.6}$ (Figure 6.15). The power of this scaling is independent of the exact maximal radius or minimal time chosen. However, the absolute value of this threshold has to be quite small ($R_{\rm max} \lesssim 0.01 R_{\rm vir}$), if we want to reproduce the observed M–σ relation with our 2D analytical model of AGN feedback (blue line in Figure 6.15).

For $\sigma \approx V_{\rm vir} \propto M_{\rm h}^{1/3}$ (Ferrarese 2002; Loeb 2010; Volonteri et al. 2011) we find a scaling between the BH mass and the halo velocity dispersion of

$$M_{\rm BH} \propto \sigma^{4.8}, \tag{6.56}$$

which is close to the proposed slope of $M_{\rm BH} \propto \sigma^5$ for an energy driven outflow in 1D models (Silk & Rees 1998; Haehnelt et al. 1998; Fabian 2012; McQuillin & McLaughlin 2013).

We do not include SN feedback in our model, which is also expected to change this relation between the galaxy mass and the efficiency to eject gas. We compare SN-driven outflows to AGN-driven winds in more detail in (Dashyan et al. 2017). Moreover, our properties of the galactic disc are intrinsically linked to the DM halo by assuming that a constant fraction of the gas with a certain fraction of the angular momentum settles into the disc (Mo et al. 1998). Therefore, we cannot disentangle the effect of our 2D density distribution and the underlying DM halo on the M-σ scaling with our analytical model.

Our new 2D model further allows for new interpretations of existing observations. Cicone et al. (2015) present CII and FIR continuum observations with the Bure Interferometer of a quasar and its host galaxy at $z = 6.4189$. The CII reveals velocities up to $1400\,{\rm km\,s^{-1}}$ and outflow with a complex morphology out to $\sim 30\,{\rm kpc}$. They find no evidence for a regular rotation pattern and identify 48 individual clumps with dynamical times in the range $10^{6.6-8.0}\,{\rm yr}$. They interpret the spread in outflow times as a non-constant AGN luminosity causing various outbursts. However, it could also be related to projection effects or

simply different flow times in different directions with respect to the disc plane (compare Figure 6.2). This supports the importance of modelling AGN outflows in 2D to allows for an outflow with different velocities in different directions.

6.4.2 Outflow efficiency as a function of the AGN luminosity

The efficiency of an outflow can be defined in different ways: either by its ability to convert input energy into momentum of the outflow (momentum boost), or by its ability to evacuate the ISM within a certain radius and preventing further gas accretion for a certain time (freefall time). In this section, we use these two quantifications to compare AGN-driven outflows in the momentum- and in the energy-driven regime.

BHs of lower mass or accreting at lower Eddington ratio have a higher momentum boost, i.e., they have a stronger push relative to the input momentum. The momentum loading of the outflow increases with decreasing AGN luminosity, because if the luminosity is too high, the outflow rapidly propagates towards lower density regions and paves a path of least resistance perpendicular to the disc plane. Consequently, the shocked wind can adiabatically expand in this direction without transferring its momentum to the denser gas in the disc plane. An outflow driven by a lower luminosity AGN remains confined for longer and therefore has more time to transfer its momentum to the gas.

However, the strength of the outflow, defined as its ability to push gas out to a given radius, e.g., the disc radius R_0, is higher the more massive and highly accreting BHs are (Figure 6.12, 6.15). High luminosity AGN are more efficient in driving an outflow out to large galactic radii and hence suppress their own gas supply for a significant amount of time. For instance Figure 6.12 shows that in a halo of $10^8\,\mathrm{M}_\odot$ an AGN with luminosity below $< 10^{40}\,\mathrm{erg\,s^{-1}}$ shuts off accretion for only $< 1\,\mathrm{Myr}$, an AGN with luminosity $10^{40}\,\mathrm{erg\,s^{-1}} < L_\mathrm{AGN} \lesssim 10^{43}\,\mathrm{erg\,s^{-1}}$ shuts off accretion for $\sim 1\,\mathrm{Myr}$, and an AGN with luminosity above $> 10^{44}\,\mathrm{erg\,s^{-1}}$ shuts off accretion for $> 10\,\mathrm{Myr}$. These AGN luminosities of $\log(L_\mathrm{AGN}/\mathrm{erg\,s^{-1}}) = 40,\ 43,\ 44$ correspond to BH masses of roughly $10^2,\ 10^5,\ 10^6\,\mathrm{M}_\odot$ at $f_\mathrm{Edd} = 1.0$. In a more massive halo these time-scales are even shorter, since they depend on the virial radius.

For AGN luminosities $\gtrsim 10^{43}\,\mathrm{erg\,s^{-1}}$, the outflow remains momentum-driven in the disc out to galactic radii. For AGN luminosities around this threshold, the momentum-driven outflow is less powerful in ejecting gas out of the galactic potential and in preventing further mass accretion in the disc plane, compared to an energy-driven outflow in the same galaxy. However, at even higher AGN luminosities, the AGN is more efficient and can prevent further gas infall for $> 10\,\mathrm{Myr}$, whereas the energy-driven outflow can only push gas in the disc plane out to distances, which correspond to a recovery time of $\sim 1\,\mathrm{Myr}$. The main reason for this difference is that the direct momentum input acts more isotropically and works still in the higher-density disc plane, even if there is already a path of least resistance in the perpendicular direction. In contrast to the energy-driven outflow, where all the pressure can escape perpendicular and does not push the gas in the disc plane to larger radii.

6.4.3 Driving mechanism

For low AGN luminosities, the transition from momentum- to energy-driving in the disc plane occurs on very small, i.e. sub-pc, scales in our models. For higher AGN luminosities beyond $L_\mathrm{AGN} \gtrsim 10^{43}\,\mathrm{erg\,s^{-1}}$, we observe that the outflow remains momentum-driven in the disc. In all models we see a mildly relativistic, energy-driven outflow close to the disc normal. This outflow develops due to the low column density in this direction, which causes very high accelerations. The ejected mass in this cone is negligible compared to the gas mass in the disc. Previous 1D models find that the transition from momentum-

to energy-driving occurs on scales of $\sim 1\,\mathrm{kpc}$ (Ciotti & Ostriker 1997; King 2003; King et al. 2011; Zubovas & King 2012), several orders of magnitude larger than what we find in our 2D model. In our 1D comparison model this transition occurs around 1 pc (Figure 6.20), but our halo mass is lower than in other 1D studies.

Bourne & Nayakshin (2013) model the expected X-ray signature of inverse Compton cooling in a one temperature medium and with thermally decoupled electrons in the post-shock wind. They argue that current observations of AGN do not show evidence of Compton cooling from a one temperature medium and weak constraints on a possible cooling from a two-temperature medium. This observation supports the theory of energy-driven winds on galactic scales that do not radiate away their thermal internal energy. This suggests that most AGN-driven outflows are energy-driven.

We derive a characteristic AGN luminosity above which the wind remains momentum driven in the galactic disc (Eq. 6.53). For our fiducial parameters this characteristic luminosity corresponds to a BH mass of the order $\sim 10^5\,\mathrm{M}_\odot$, which is independent of the halo mass, but it strongly depends on the redshift and the initial wind velocity.

In this paper, we focused on galactic outflows driven by the inner disc wind of the AGN. However, the ISM could also be accelerated directly by the radiation pressure from the AGN (Ishibashi & Fabian 2014, 2015; Thompson et al. 2015; Costa et al. 2017). Although this mechanism can be efficient close to the SMBH, radiation pressure can accelerate gas only up to an effective optical depth of order unity. For electron scattering this transparency radius, where $\tau \approx 1$, is (King & Pounds 2014, 2015)

$$R_{\mathrm{tr}} \approx 50\,\mathrm{pc}\left(\frac{f_{\mathrm{gas}}}{0.17}\right)\left(\frac{\sigma}{200\,\mathrm{km\,s^{-1}}}\right)^2, \tag{6.57}$$

beyond which direct driving by radiation pressure from the AGN cannot accelerate galactic outflows. The inclusion of dust increases the efficiency of direct radiation pressure on the ISM, due to the higher cross section of dust and the consequently larger transparency radius of several kpc (Ishibashi & Fabian 2012, 2015; King & Pounds 2015). The momentum input of such a radiation pressure-driven wind can be comparable to that of a momentum-driven wind, but the main difference is the frequency-dependent cross section of dust (with a peak in the UV). Consequently, not all gas might experience the radiation pressure due to its locally low opacity or due to self-shielding by higher opacity regions. Even hybrid models might be realistic, where the inner wind does not directly emerge from the disc, but is accelerated by radiation pressure. In any case, if radiation pressure on electrons and dust drives galactic gas outflows, the accelerated ISM prefers the path of least resistance, which requires a 2D treatment.

6.4.4 Caveats

We treat the ISM as homogeneous and assume that the AGN-driven shock sweeps up all the enclosed gas. Real galaxies have a multiphase ISM with HII regions and dense molecular clumps, which alter this simplistic treatment. Theoretical models of the impact of AGN winds (Bieri et al. 2017) and jets (Wagner et al. 2012, 2013) on a fractal ISM show that dense clumps can be dissolved by IR photons that can penetrate into these clumps. The efficiency of the AGN-driven wind depends on the size and filling factor of these dense clumps. A galaxy with many small isolated clouds experiences efficient cloud dispersion compared to a galaxy with fewer but bigger cloud complexes. By neglecting the realistic ISM structure, we overestimate the outflow efficiency, because a fractal ISM predefines already paths of lest resistance and the momentum transfer to dense clumps is smaller compared to a homogeneous ISM. On the other hand, the low density gas between the clumps is accelerated and ejected more easily.

The shock can also form dense cold clumps via dynamical instabilities (Costa et al. 2015), which we do not include in our model. The contact discontinuity is Rayleigh-Taylor unstable and might give rise to the formation of clumps that form behind the shock front and decouple from the outflow. These denser entities could also be the constituents of the high-speed molecular outflows. Recently, Ferrara & Scannapieco (2016) find in their 3D simulation that clumps form at the transition from momentum- to energy-driving, but they get rapidly dissolved by the hot shock gas flowing past them (see also Bieri et al. 2017; Richings & Faucher-Giguere 2017).

We assume the AGN luminosity to be constant during the lifetime of the AGN. This is a necessary assumption to clearly disentangle the influence of different input parameters. Different groups study the radiation regulated accretion on to the BH and demonstrate the importance of multidimensional hydrodynamical simulation to self-consistently follow the accretion and resulting luminosity (Park & Ricotti 2011b; Sugimura et al. 2017; Negri & Volonteri 2017). The accretion rate and hence the AGN luminosity oscillate by up to two orders of magnitude on time-scales of several thousand years, shorter than the lifetime of the AGN in our model. Gilli et al. (2017) investigate galactic outflows driven by an AGN with exponentially increasing luminosity, i.e. they self-consistently account for the mass growth of the central BH at constant Eddington ratio. They find that the late time expansion of the radius in their 1D model is also exponential, irrespective of the details of the driving mechanism.

We only distinguish between the momentum- and energy-driven phase, whereas Faucher-Giguère & Quataert (2012) also account for the "intermediate partially radiative bubble stage". In this phase, the cooling time is shorter than the flow time but longer than the crossing time of the shock and the shock cools only partially. Although we agree that a more realistic transition from one driving mechanism to another is necessary, this does not affect our fiducial model since the transition occurs already very early and is very sharp.

The weak collisional coupling between protons and electrons in the shocked wind can increase the Compton cooling time by two orders of magnitude (Section 6.2.4). We do not implicitly include this effect, but as pointed out by Faucher-Giguère & Quataert (2012), the more realistic treatment of the shock as a two-temperature medium significantly decreases the efficiency of inverse Compton cooling. This does not change anything in our fiducial model, because the transition to energy-driven occurs already at very small radii. However, the AGN transition luminosity, above which the outflow is always momentum-driven in the disc (Eq. 6.53) will be higher as it scales $L_{\rm AGN,crit} \propto t_{\rm compton}^2$.

An additional increase of the Compton cooling time can be achieved by the inclusion of Compton heating. Sazonov et al. (2004) determine the equilibrium Compton temperature of gas exposed to a characteristic AGN spectrum to be $T_C = 2 \times 10^7\,\mathrm{K}$, which is insensitive to obscuration effects. The equilibrium Compton temperature is equal to the mean photon energy averaged over the AGN spectra. They conclude that the characteristic Compton heating and cooling rates per particle should be the same within a factor of ~ 2. Most of the Compton cooling will be provided by the IR component, whereas the Compton heating is dominated by the high-energy component. The maximum radius out to which an AGN can heat low density, highly ionised gas by Compton heating is

$$r_C = 0.4\,\mathrm{kpc} \left(\frac{f_{\rm Edd}}{0.1}\right)^{1/2} \left(\frac{M_{\rm BH}}{10^8\,\mathrm{M}_\odot}\right)^{1/2}. \tag{6.58}$$

For ionization parameters below $\xi \lesssim 10^5$ the gas cannot be heated to T_C, due to other more efficient cooling mechanisms. A more detailed discussion of the effect of Compton heating is given in Sazonov et al. (2004, 2005).

Hydrodynamical simulations of a geometrically thin and optically thick accretion disc show that the disc wind might not be isotropic, as assumed in our model, but has a covering factor of $\sim 20\%$ (Proga et al. 2000). Moreover, the initial wind velocity itself might depend on the angle to the disc with higher velocities in the perpendicular direction and a possible self-shadowing effect by the outer part of the disc amplifies this anisotropy (Sugimura et al. 2017).

6.5 Summary and conclusion

AGN-driven winds can transport energy and momentum from sub-pc scales around the BH to galactic scales and thereby regulate the co-evolution of a galaxy and its central BH. We present a new 2D analytical model for AGN-driven outflows and demonstrate the importance of a more realistic gaseous disc profile for the outflow dynamics. In contrast to simplistic 1D models, we predict smaller gas ejection fractions and shorter fallback times, both facilitating an efficient mass growth of BHs via feeding from galactic scales. This is related to the energy-driven nature of the outflows, which pave a path of least resistance perpendicular to the disc and hence prevent efficient driving in the disc plane.

Our main results are as follows:

- For typical high redshift galaxies, the AGN can eject at most $\sim 10\%$ of the ISM out of the halo, whereas 1D models predict complete gas ejection under similar conditions. At high redshift, the ejected gas mass fractions are lower due to the deeper gravitational potential, compared to similar galaxies at low redshift.
- The characteristic time for which the AGN can suppress further gas supply is remarkably short, of the order a few million years. We find AGNs with a low luminosity to be more efficient in converting their input energy into outflow momentum, because the swept-up medium is confined for longer by the shock front and has consequently more time to store up internal energy.
- We also find a systematic transition in the outflow nature: for AGN luminosities below $10^{43}\,\mathrm{erg\,s^{-1}}$ the outflow is energy-driven, independent of the halo properties. At higher luminosities, the outflow remains momentum-driven in the disc plane.
- Independently of the exact criterion to suppress further gas accretion, we find a slope of $M_{\mathrm{BH}} \propto \sigma^{4.8}$ for the M–σ relation.

Our new model highlights the importance of a realistic 2D density profile to correctly predict the ejected gas mass fraction and fallback time. These results can be used as an improved subgrid model in cosmological simulations or semi-analytical models of galaxy formation to predict the efficiency of AGN feedback. To improve the model with a more realistic feeding prescription, we want to implement a self-consistent connection between the inflows and outflows, see the future projects below.

7 Conclusion

In this chapter we summarise the main results of the PhD thesis, highlight how they contribute to open questions in the field, and present future research plans and perspectives, based on this work.

7.1 Summary

With this thesis I have contributed to several of the open questions about the formation and growth of the first SMBHs. I have used numerical simulations to study the formation of the first seed BHs. I have modelled and interpreted recent observations of the Lyα emitter CR7 and the first detection of gravitational waves and made predictions for upcoming detections of BH-BH mergers from the first stars. I have also developed a new 2D analytical model of AGN-driven outflows to better understand the feedback caused by the growth of BHs. The main results of the individual chapters are:

- H_2 self-shielding is an important process for the direct collapse scenario to form massive BH seeds. We have shown that an improved, self-consistent treatment of self-shielding in 3D cosmological simulations reduces the critical LW flux by a factor of two and therefore boosts the number of formation sites of DCBHs by more than one order of magnitude.

- We have calculated the merger rate density of binary BHs from the first stars and their detection rates with aLIGO. There is a $\sim 1\%$ probability that the first detection GW150914 was of primordial origin. Upcoming detections in the next decades will allow to put tighter constraints on the primordial IMF and therefore on SMBH formation scenarios.

- The strong Lyα and HeII emission in CR7 cannot be caused by a primordial stellar population. An accreting BH in the centre of CR7, which might be formed via direct collapse, is more likely to account for the observational features. This would be the first observational evidence of a DCBH.

- We have reviewed the statistics of the formation of the first SMBHs. We discuss the main processes that influence the probability to form BH seeds via different scenarios and compare the probability of the different criteria. We demonstrate that metal pollution and the LW flux are the most constraining ingredient and highlight the importance of tidal stripping, which has been advocated by several recent publications.

- We have developed a 2D analytical model of AGN-driven outflows and demonstrate that a more realistic disc profile reduces the amount of gas that is ejected out of the halo, compared to existing 1D models. AGN-driven winds stop gas accretion on to the central BH for only about $\sim 1\,$Myr, which still allows for almost continuous gas inflow in the disc plane.

We want to highlight the new and original aspect of this thesis: we were the first to demonstrate that CR7 cannot host the originally proposed mass of metal-free stars and after the detection of GW150914, we were the first to consider its possible primordial origin. In both cases, other groups have independently

confirmed our main conclusions. Our study of AGN-driven winds is the first analytical solution of this problem in two dimensions and constitutes the starting point for various follow-up projects (see below).

Our work was driven by two main observations that we want to understand: the existence of SMBHs already a few hundred million years after the Big Bang and low redshift correlations between galaxies and their central BHs. The extreme quasars at high redshift are only the tip of the iceberg of the entire population of BHs. Consequently, their formation and growth might be different compared to the main population of BHs in the Universe. Their growth is a complex interplay of accretion, feedback, mergers, and ejections. Different channels for the formation and various mechanisms to enable efficient mass growth might work in parallel to account for the different observations in the low and high redshift Universe. With our study of H_2 self-shielding we demonstrate that there should be more DCBHs than previously assumed. Consequently, the direct collapse scenario or dynamical instabilities might provide massive BH seeds to account for the most massive quasars at high redshift. However, the vast majority of BHs might form as Pop III remnant BHs or BHs formed in dense stellar clusters. Their self-regulated growth in galaxies might explain the correlation, as our 2D model of AGN-driven winds suggests.

Upcoming detections of next-generation telescopes will help to improve our understanding of the formation and growth of the first SMBHs. Future X-ray missions and near infrared facilities, such as *JWST*, will be able to observe accreting BHs at $z \gtrsim 6$, constraining the accretion properties of massive BHs at early times (Volonteri 2010; Volonteri et al. 2017). Especially the GW detector LISA will be of paramount importance to detect massive BHs in the high redshift Universe (Sesana et al. 2007, 2009). The detection rates and mass distribution of BH-BH mergers enable a discrimination between different formation scenarios (Figure 7.1). The upcoming detections with LISA will be able to distinguish two

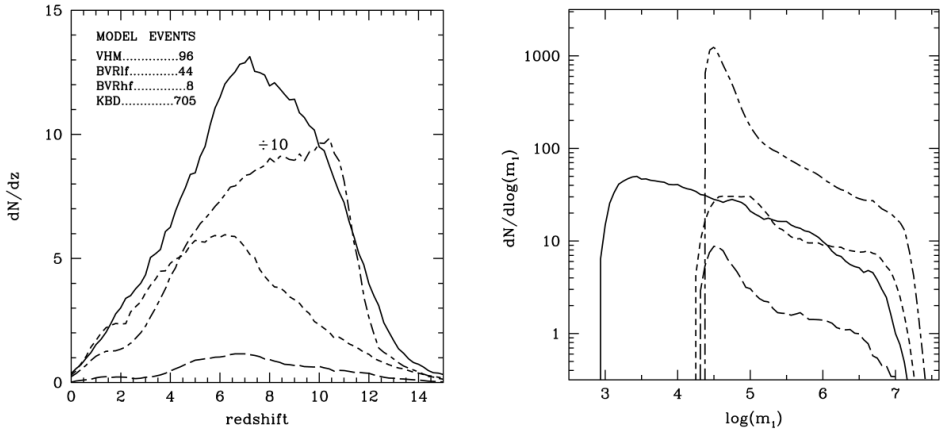


Figure 7.1: Detection rates (SNR > 5) with LISA in a 3-year mission for different BH formation scenarios: Pop III seed BHs based on Volonteri et al. (2003) (VHM, solid lines); massive BHs formed via dynamical instabilities based on Koushiappas et al. (2004) (KBD, short–long-dashed lines) and based on Begelman et al. (2006) with efficient (BVRhf, long-dashed lines) and mild (BVRlf, short-dashed lines) chemical feedback. Left: Redshift distribution of massive binary BHs. The number of events predicted by the KBD model is divided by a factor of 10. The top left-hand corner label lists the total number of expected detections. Right: Mass function of the more massive member of massive binary BHs. Adapted from Sesana et al. (2007).

populations of BHs: stellar mass BHs forming at $z = 10 - 20$ or more massive BHs, forming at later redshift. Constraining the formation scenarios with GWs will also put tighter limits on the possible growth histories.

7.2 Perspectives and future projects

We briefly present future projects that we want to realise based on the results of this thesis. Besides these explicitly presented projects, there remains a lot of work to be done in the different fields covered. The individual caveat sections point e.g. to necessary improvements for the models.

7.2.1 Semi-analytical model of self-regulated BH growth

The main advantage of our 2D analytical model of AGN-driven outflows is the more realistic gas distribution, compared to simplistic 1D solutions and the computational efficiency, compared to 3D hydrodynamical simulations. This makes our new approach the perfect subgrid model for upcoming studies: based on the galaxy properties, such as halo mass, angular momentum, gas temperature, and redshift, we can calculate the expected effect of AGN-driven winds. These tabulated outflow efficiencies (e.g. ejected gas mass fraction, recovery time) serve as input for more sophisticated models of BH growth.

For an isolated galaxy, we can study periodic episodes of gas ejection by the AGN and the redistribution of the ISM while the AGN is off. In a more realistic scenario, we also take into account the merger history of the host galaxy, star formation, and gas supply from outside the halo. Moreover, we aim at implementing a self-consistent model for the feeding of the BH from galactic scales, based on e.g. Lodato & Natarajan (2006); Devecchi et al. (2010). We want to couple our 2D model to the semi-analytical merger tree that we have already used to constrain the nature of CR7 (Hartwig et al. 2016b) and the expected GW signal from the first stars (Hartwig et al. 2016a). Starting with a self-consistent seeding scenario of BHs in the high redshift Universe, we follow their mass accretion and the self-regulated feedback through cosmic time. The aim is to correctly reproduce the observed population of BHs in their host galaxies at low and high redshift.

7.2.2 Constraining the nature of the first stars with Galactic archaeology

As we have seen, the first stars are one possible path to the formation of the first SMBHs and their feedback sets the scene for most other proposed scenarios. To better understand the nature of the first stars, I will work on the following project in collaboration with Prof. Naoki Yoshida at the University of Tokyo: we will create mock observations of second generation stars and constrain the IMF of Pop III stars by combining state-of-the-art simulations and observations of stellar fossils in the Milky Way.

An interesting class of stars are those that got enriched by exactly one previous supernova and that consequently carry the chemical fingerprint of this progenitor star. We will combine high-resolution simulations of chemical enrichment in the young Milky Way with observations of galactic archaeology to answer the question: if there were Pop III stars of mass M, what is the probability to find a second generation star in a galactic archaeology survey that carries this characteristic chemical fingerprint? This probability distribution, deconvolved with the distribution of currently observed second generations stars, directly yields an estimate of the Pop III IMF.

In a second step, we will create mock observations of these second generation stars by accounting for surface pollution, observational biases, selection effects, survey strategies, and the sensitivity of the telescopes. We will apply the same pipeline as for observed second generation stars to determine the most likely progenitor masses of our simulated stars. We will do this by comparing the metal abundances in the stellar atmosphere with tabulated models and catalogues of the elemental yields of SN explosions for different progenitor masses. This approach will allow us to derive the likelihood of indirectly detecting

a Pop III star of a specific mass by its characteristic metal abundances imprinted in a second generation star.

Our results will have two direct impacts. First, the community can use our estimate of the Pop III IMF as an input for their models, which will improve the predictions and analysis of upcoming observational campaigns of the first galaxies and black holes with spaceborne telescopes and GW detectors. Second, we can optimise future galactic archaeology surveys to find more second generation stars based on our mock observations, because with current selection criteria we might miss some of these rare and precious fossils.

7.2.3 Bayesian meta–analysis of the Pop III IMF

There is a variety of different approaches to constrain the Pop III IMF but so far there has not been a dedicated effort to consolidate these theoretical and observational predictions. In another planned project, I am going to combine these constraints in a Bayesian meta-analysis to obtain our current understanding of the primordial IMF. During my PhD, I have acquired the necessary skills and knowledge for this meta-analysis: stellar archaeology (Hartwig et al. 2015a; Magg et al. 2017; Hartwig et al. 2017c), numerical simulations of Pop III star formation (Hartwig et al. 2015b), PISNe to probe the first stars (Magg et al. 2016; Hartwig et al. 2017a), gravitational waves from the first stars (Hartwig et al. 2016a), and the formation of massive metal-free stars under Lyman-Werner feedback (Hartwig et al. 2015c). An illustration of the Bayesian meta-analysis can be seen in Figure 7.2.

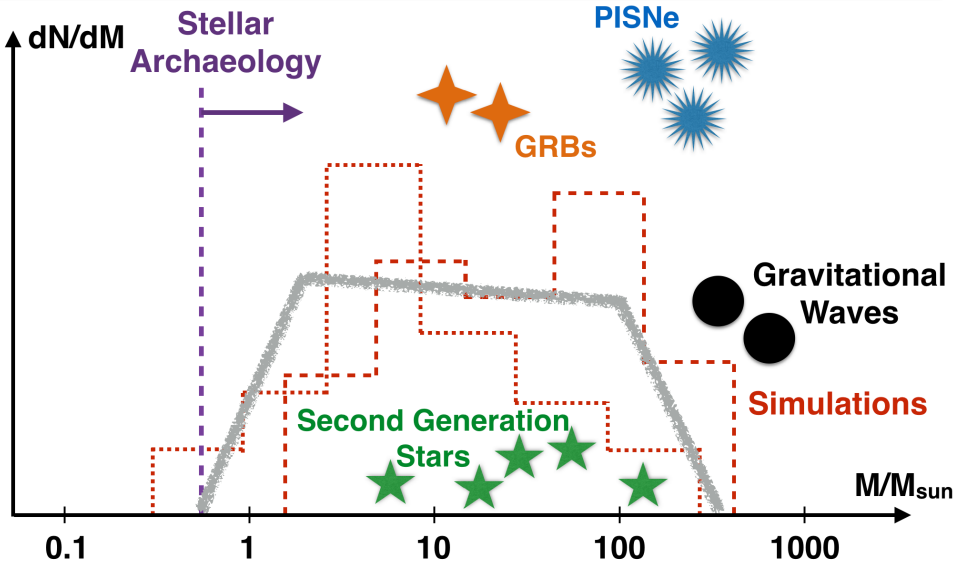


Figure 7.2: I will combine different constraints on the Pop III IMF in a Bayesian meta analysis: constraints from stellar archaeology, second generation stars, numerical simulations, SNe in the high redshift Universe, gamma-ray bursts (GRBs), and gravitational waves will be quantitatively compared and consolidated.

First, I will combine the different observations and determine the likelihoods for the IMF-generating parameters such as $M_{\rm min}$, $M_{\rm max}$, or the slope(s). The challenge is to quantify the observational biases and selection effects. Then, I compare this posterior IMF to the IMFs predicted by simulations and can identify crucial input parameters, such as resolution, magnetic fields, or protostellar evolution and verify to which degree a given simulation is in agreement with the compilation of observations. In a Bayesian framework, the variance multipliers that consistently come out of this analysis yield the appropriate weighting for the combination of IMFs based on different simulations. This quantitative framework of

how to combine different simulated Pop III IMFs allows to derive one standardised initial mass function of the first stars.

The main result will be the distribution of most likely IMFs for the first stars that can be used as input in many other models. Another precious outcome is the information which is the most valuable approach in constraining the Pop III IMF with the highest density of information and which range of the IMF has to be better probed by either simulations or observations in the future. This can be quantified with the Kullback-Leibler divergence (Kullback & Leibler 1951) to measure the information gain of upcoming observations or simulations and to optimise the required experimental design accordingly.

Acknowledgements

My research and the successful completion of this PhD Thesis would not have been possible without the support of many people, whom I would like to thank for their contributions.

- My PhD supervisor Marta Volonteri for her guidance, help, and support. From her I've learned invaluable lessons about astrophysics, good scientific practice, and about life in general. Marta, your critical mind and wisdom prevented me several times from doing nonsense and guided me towards the realisation of this PhD thesis. Many thanks for your support and the freedom to realise my own projects and ideas.

- Ralf Klessen and Volker Bromm as my collaborators for guidance and help, also beyond astrophysics.

- Mattis Magg for his contributions to my semi-analytical model, improvements, corrections, and new ideas. Mattis, I owe you several cakes by now!

- My coauthors Simon Glover, Muhammad Latif, Enrico Barausse, Athena Stacy, Dan Whalen, Eric Pellegrini, and Gohar Dashyan for indispensable contributions, inspiring discussions, and critical comments on the manuscripts.

- The BH/AGN group at IAP for weekly inspiring, critical discussions and help.

- Naoki Yoshida and his research group for their hospitality during my internship at the University of Tokyo. ありがとうございます

- The current and former PhD students at IAP, especially bureau 10 for the main motivation to come to IAP daily, Caterina, Siwei, and Alba for support in difficult times, and Melanie and Rebekka for help and support at the beginning.

- The members of my PhD defence jury Frederic Daigne, Michela Mapelli, Jarrett Johnson, and Ralf Klessen for their time and effort, especially to the two reviewers Michela and Jarrett for carefully reading this manuscript and providing their critical and helpful feedback.

- The anonymous referees of our papers, who also contributed to the improvement of the presented results.

- All the lovely colleagues, who I met at various conferences and with whom I shared crazy and more realistic ideas about possible future projects.

- Chantal Le Vaillant, Agnes Le Guerer, and the Ecole Doctorale for support and help with the bureaucracy.

- Marta, Judith, Melanie, Hugo, Jesse, Caterina, Helen, and Sunmyon for cross-reading parts of my PhD Thesis and providing me with helpful corrections and ideas of improvement.

- Besides all the scientific support, I want to thank my friends and family for all the help, affirmation, and support in the last years and especially during the past few weeks. Moreover, I want to thank my Liberian friend Wilfred for not only sharing the common utopia to improve the world by education, but also for collectively realising the teacher workshops in West Africa.

- Finally, all the resources that simplify daily work: the open dictionary leo.org, stackoverflow for saving hundreds of hours of debugging, arXiv and ADS, and further public software: we are very grateful for the free availability of the code MUSIC to generate cosmological initial conditions. We thank the GALFORM team to make their code publicly available. Calculations in this thesis were performed with version 13.04 of CLOUDY, last described by Ferland et al. (2013).

- I acknowledge funding under the European Community's Seventh Framework Programme (FP7/ 2007-2013) via the European Research Council Grants 'BLACK' under the project number 614199.

List of Figures

List of Tables

Bibliography

Aalto, S., Muller, S., Sakamoto, K., et al. 2012, Astron. and Astrophys., 546, A68

Abadie, J., et al. 2010, Classical and Quantum Gravity, 27, 173001

Abbott, B. P., et al. 2009, Advanced LIGO anticipated sensitivity curves, https://dcc.ligo.org/LIGO-T0900288/public

Abbott, B. P., et al. 2015, LIGO Document G1501223-v3, H1 Calibrated Sensitivity Spectra Oct 1 2015 (Representative for Start of O1), https://dcc.ligo.org/LIGO-G1501223/public

Abbott, B. P., et al. 2016, Physical Review X, 6, doi:10.1103/PhysRevX.6.041015

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016a, Physical Review Letters, 116, 131102

Abbott, B. P., et al. 2016b, Phys. Rev. Lett., 116, 061102

Abbott, B. P., et al. 2016, Physical Review Letters, 116, doi:10.1103/PhysRevLett.116.061102

Abbott, B. P., et al. 2016, ArXiv e-prints, arXiv:1602.03842

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, Physical Review Letters, 118, 221101

Abel, T., Anninos, P., Norman, M. L., & Zhang, Y. 1998, Astrophys. J., 508, 518

Abel, T., Anninos, P., Zhang, Y., & Norman, M. L. 1997, Nature, 2, 181

Abel, T., Bryan, G. L., & Norman, M. L. 2000, Astrophys. J., 540, 39

—. 2002, Science, 295, 93

Abel, T., & Haehnelt, M. G. 1999, Astrophys. J. Lett., 520, L13

Abel, T., Wise, J. H., & Bryan, G. L. 2007, Astrophys. J. Lett., 659, L87

Agarwal, B., Dalla Vecchia, C., Johnson, J. L., Khochfar, S., & Paardekooper, J.-P. 2014, Monthly Notices Roy. Astron. Soc., 443, 648

Agarwal, B., Johnson, J. L., Khochfar, S., et al. 2017a, Monthly Notices Roy. Astron. Soc., 469, 231

Agarwal, B., Johnson, J. L., Zackrisson, E., et al. 2016a, Monthly Notices Roy. Astron. Soc., 460, 4003

Agarwal, B., & Khochfar, S. 2015, Monthly Notices Roy. Astron. Soc., 446, 160

Agarwal, B., Khochfar, S., Johnson, J. L., et al. 2012, Monthly Notices Roy. Astron. Soc., 425, 2854

Agarwal, B., Regan, J., Klessen, R. S., Downes, T. P., & Zackrisson, E. 2017b, ArXiv e-prints, arXiv:1703.08181

Agarwal, B., Smith, B., Glover, S., Natarajan, P., & Khochfar, S. 2016b, Monthly Notices Roy. Astron. Soc., 459, 4209

Ahn, K., Shapiro, P. R., Iliev, I. T., Mellema, G., & Pen, U.-L. 2009, Astrophys. J., 695, 1430

Ajith, P., et al. 2008, Phys. Rev., D77, 104017, [Erratum: Phys. Rev.D79,129901(2009)]

—. 2009, Phys. Rev. D, 79, 129901

Allen, B., & Romano, J. D. 1999, Phys. Rev. D, 59, 102001

Alvarez, M. A., Bromm, V., & Shapiro, P. R. 2006, Astrophys. J., 639, 621

Alvarez, M. A., Wise, J. H., & Abel, T. 2009, Astrophys. J. Lett., 701, L133

Amaro-Seoane, P., & Santamaría, L. 2010, Astrophys. J., 722, 1197

Aoki, W., Tominaga, N., Beers, T. C., Honda, S., & Lee, Y. S. 2014, Science, 345, 912

Aykutalp, A., Wise, J. H., Spaans, M., & Meijerink, R. 2014, Astrophys. J., 797, 139

Baldassare, V. F., Reines, A. E., Gallo, E., & Greene, J. E. 2015, Astrophys. J. Lett., 809, L14

Ball, W. H., Tout, C. A., Żytkow, A. N., & Eldridge, J. J. 2011, Monthly Notices Roy. Astron. Soc., 414, 2751

Barkana, R., & Loeb, A. 2001, Physics Reports, 349, 125

Bartelmann, M. 2007, Lecture Notes "Cosmology"

—. 2009, Lecture Notes "General Relativity"

Barton, E. J., Geller, M. J., & Kenyon, S. J. 2000, Astrophys. J., 530, 660

Becerra, F., Greif, T. H., Springel, V., & Hernquist, L. E. 2015, Monthly Notices Roy. Astron. Soc., 446, 2380

Begelman, M. C. 2010, Monthly Notices Roy. Astron. Soc., 402, 673

Begelman, M. C., & Rees, M. J. 1978, Monthly Notices Roy. Astron. Soc., 185, 847

Begelman, M. C., Rossi, E. M., & Armitage, P. J. 2008, Monthly Notices Roy. Astron. Soc., 387, 1649

Begelman, M. C., Volonteri, M., & Rees, M. J. 2006, Monthly Notices Roy. Astron. Soc., 370, 289

Behroozi, P. S., & Silk, J. 2015, Astrophys. J., 799, 32

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, Astrophys. J., 770, 57

Belczynski, K., Bulik, T., & Rudak, B. 2004a, Astrophys. J. Lett., 608, L45

Belczynski, K., Buonanno, A., Cantiello, M., et al. 2014, Astrophys. J., 789, 120

Belczynski, K., Holz, D. E., Bulik, T., & O'Shaughnessy, R. 2016a, Nature, 534, 512

Belczynski, K., Repetto, S., Holz, D. E., et al. 2016b, Astrophys. J., 819, 108

Belczynski, K., Ryu, T., Perna, R., et al. 2016c, ArXiv e-prints, arXiv:1612.01524

Belczynski, K., Sadowski, A., & Rasio, F. A. 2004b, Astrophys. J., 611, 1068

Bellovary, J., Volonteri, M., Governato, F., et al. 2011, Astrophys. J., 742, 13

Bergin, E. A., & Tafalla, M. 2007, Ann. Rev. Astron. Astrophys., 45, 339

Bieri, R., Dubois, Y., Rosdahl, J., et al. 2017, Monthly Notices Roy. Astron. Soc., 464, 1854

Biermann, L. 1950, Zeitschrift Naturforschung Teil A, 5, 65

Black, J. H., & Dalgarno, A. 1977, ApJS, 34, 405

Blandford, R. D., & Znajek, R. L. 1977, Monthly Notices Roy. Astron. Soc., 179, 433

Bond, J. R., & Carr, B. J. 1984, Monthly Notices Roy. Astron. Soc., 207, 585

Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, Astrophys. J., 379, 440

Bourne, M. A., & Nayakshin, S. 2013, Monthly Notices Roy. Astron. Soc., 436, 2346

Bouwens, R. J., Illingworth, G. D., Labbe, I., et al. 2011, Nature, 469, 504

Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, Monthly Notices Roy. Astron. Soc., 370, 645

Bowler, R. A. A., McLure, R. J., Dunlop, J. S., et al. 2017, Monthly Notices Roy. Astron. Soc., 469, 448

Bowler, R. A. A., Dunlop, J. S., McLure, R. J., et al. 2012, Monthly Notices Roy. Astron. Soc., 426, 2772

Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2008, Monthly Notices Roy. Astron. Soc., 383, 93

Bramberger, S. F., Brandenberger, R. H., Jreidini, P., & Quintin, J. 2015, Journal of Cosmology and Astroparticle Physics, 6, 007

Brinchmann, J., Pettini, M., & Charlot, S. 2008, Monthly Notices Roy. Astron. Soc., 385, 769

Bromm, V. 2013, Reports on Progress in Physics, 76, 112901

Bromm, V., Coppi, P. S., & Larson, R. B. 1999, Astrophys. J. Lett., 527, L5

—. 2002, Astrophys. J., 564, 23

Bromm, V., Ferrara, A., Coppi, P. S., & Larson, R. B. 2001a, Monthly Notices Roy. Astron. Soc., 328, 969

Bromm, V., Kudritzki, R. P., & Loeb, A. 2001b, Astrophys. J., 552, 464

Bromm, V., & Larson, R. B. 2004, Ann. Rev. Astron. Astrophys., 42, 79
Bromm, V., & Loeb, A. 2002, Astrophys. J., 575, 111
—. 2003a, Astrophys. J., 596, 34
—. 2003b, Astrophys. J., 596, 34
—. 2004, Nature, 9, 353
Bromm, V., & Yoshida, N. 2011, Ann. Rev. Astron. Astrophys., 49, 373
Bromm, V., Yoshida, N., Hernquist, L., & McKee, C. F. 2009, Nature, 459, 49
Cappi, M. 2006, Astronomische Nachrichten, 327, 1012
Carlberg, R. G. 1981, Monthly Notices Roy. Astron. Soc., 197, 1021
Carr, B., Kühnel, F., & Sandstad, M. 2016, Phys. Rev. D, 94, 083504
Carr, B. J., Bond, J. R., & Arnett, W. D. 1984, Astrophys. J., 277, 445
Carr, B. J., & Hawking, S. W. 1974, Monthly Notices Roy. Astron. Soc., 168, 399
Cen, R., & Riquelme, M. A. 2008, Astrophys. J., 674, 644
Chartas, G., Brandt, W. N., Gallagher, S. C., & Garmire, G. P. 2002, Astrophys. J., 579, 169
Chatzopoulos, E., & Wheeler, J. C. 2012, Astrophys. J., 748, 42
Choi, J.-H., Shlosman, I., & Begelman, M. C. 2013, Astrophys. J., 774, 149
Chon, S., Hirano, S., Hosokawa, T., & Yoshida, N. 2016, Astrophys. J., 832, 134
Christensen, C., Quinn, T., Governato, F., et al. 2012, Monthly Notices Roy. Astron. Soc., 425, 3058
Cicone, C., et al. 2014, Astron. and Astrophys., 562, A21
Cicone, C., Maiolino, R., Gallerani, S., et al. 2015, Astron. and Astrophys., 574, A14
Ciotti, L., & Ostriker, J. P. 1997, Astrophys. J. Lett., 487, L105
Civano, F., Marchesi, S., Comastri, A., et al. 2016, Astrophys. J., 819, 62
Clark, P. C., & Glover, S. C. O. 2013, Lecture Notes "The Dark Ages of the Universe"
Clark, P. C., Glover, S. C. O., & Klessen, R. S. 2008, Astrophys. J., 672, 757
—. 2012a, Monthly Notices Roy. Astron. Soc., 420, 745
—. 2012b, Monthly Notices Roy. Astron. Soc., 420, 745
Clark, P. C., Glover, S. C. O., Klessen, R. S., & Bromm, V. 2011a, Astrophys. J., 727, 110
Clark, P. C., Glover, S. C. O., Smith, R. J., et al. 2011b, Science, 331, 1040
Cojazzi, P., Bressan, A., Lucchin, F., Pantano, O., & Chavez, M. 2000, Monthly Notices Roy. Astron. Soc., 315, L51
Costa, T., Rosdahl, J., Sijacki, D., & Haehnelt, M. 2017, ArXiv e-prints, arXiv:1703.05766
Costa, T., Sijacki, D., & Haehnelt, M. G. 2014, Monthly Notices Roy. Astron. Soc., 444, 2355
—. 2015, Monthly Notices Roy. Astron. Soc., 448, L30
Couchman, H. M. P., & Rees, M. J. 1986, Monthly Notices Roy. Astron. Soc., 221, 53
Courteau, S., de Jong, R. S., & Broeils, A. H. 1996, Astrophys. J. Lett., 457, L73
Cox, T. J., Jonsson, P., Somerville, R. S., Primack, J. R., & Dekel, A. 2008, Monthly Notices Roy. Astron. Soc., 384, 386
Croton, D. J., Springel, V., White, S. D. M., et al. 2006, Monthly Notices Roy. Astron. Soc., 365, 11
Croton, D. J., Stevens, A. R. H., Tonini, C., et al. 2016, ApJS, 222, 22
Curtis-Lake, E., McLure, R. J., Dunlop, J. S., et al. 2013, Monthly Notices Roy. Astron. Soc., 429, 302
Cutler, C., & Flanagan, É. E. 1994, Phys. Rev. D, 49, 2658
Danovich, M., Dekel, A., Hahn, O., & Teyssier, R. 2012, Monthly Notices Roy. Astron. Soc., 422, 1732
Dashyan, G., Silk, J., Mamon, G. A., Dubois, Y., & Hartwig, T. 2017, MNRAS in prep.
Davies, M. B., Miller, M. C., & Bellovary, J. M. 2011, Astrophys. J. Lett., 740, L42

de Bennassuti, M., Salvadori, S., Schneider, R., Valiante, R., & Omukai, K. 2017, Monthly Notices Roy. Astron. Soc., 465, 926

de Mink, S. E., & Belczynski, K. 2015, Astrophys. J., 814, 58

de Mink, S. E., Langer, N., Izzard, R. G., Sana, H., & de Koter, A. 2013, Astrophys. J., 764, 166

de Mink, S. E., & Mandel, I. 2016, Monthly Notices Roy. Astron. Soc., 460, 3545

De Rosa, G., Venemans, B. P., Decarli, R., et al. 2014, Astrophys. J., 790, 145

de Souza, R. S., Ishida, E. E. O., Whalen, D. J., Johnson, J. L., & Ferrara, A. 2014, Monthly Notices Roy. Astron. Soc., 442, 1640

Devecchi, B., & Volonteri, M. 2009, Astrophys. J., 694, 302

Devecchi, B., Volonteri, M., Colpi, M., & Haardt, F. 2010, Monthly Notices Roy. Astron. Soc., 409, 1057

Devecchi, B., Volonteri, M., Rossi, E. M., Colpi, M., & Portegies Zwart, S. 2012, Monthly Notices Roy. Astron. Soc., 421, 1465

Dijkstra, M., Ferrara, A., & Mesinger, A. 2014, Monthly Notices Roy. Astron. Soc., 442, 2036

Dijkstra, M., Gronke, M., & Sobral, D. 2016, Astrophys. J., 823, 74

Dijkstra, M., Haiman, Z., Mesinger, A., & Wyithe, J. S. B. 2008, Monthly Notices Roy. Astron. Soc., 391, 1961

Dijkstra, M., Haiman, Z., & Spaans, M. 2006, Astrophys. J., 649, 14

Dijkstra, M., Lidz, A., & Wyithe, J. S. B. 2007, Monthly Notices Roy. Astron. Soc., 377, 1175

Dijkstra, M., & Wyithe, J. S. B. 2010, Monthly Notices Roy. Astron. Soc., 408, 352

Dominik, M., Belczynski, K., Fryer, C., et al. 2012, Astrophys. J., 759, 52

—. 2013, Astrophys. J., 779, 72

Dominik, M., Berti, E., O'Shaughnessy, R., et al. 2015, Astrophys. J., 806, 263

Dopcke, G., Glover, S. C. O., Clark, P. C., & Klessen, R. S. 2013, Astrophys. J., 766, 103

Dotti, M., Colpi, M., Maraschi, L., Perego, A., & Volonteri, M. 2010, in Astronomical Society of the Pacific Conference Series, Vol. 427, Accretion and Ejection in AGN: a Global View, ed. L. Maraschi, G. Ghisellini, R. Della Ceca, & F. Tavecchio, 19

Draine, B. T., & Bertoldi, F. 1996, Astrophys. J., 468, 269

Dubois, Y., Pichon, C., Welker, C., et al. 2014, Monthly Notices Roy. Astron. Soc., 444, 1453

Dvorkin, I., Vangioni, E., Silk, J., Uzan, J.-P., & Olive, K. A. 2016, Monthly Notices Roy. Astron. Soc., 461, 3877

Ebisuzaki, T., Makino, J., Tsuru, T. G., et al. 2001, Astrophys. J. Lett., 562, L19

Eisenstein, D. J., & Loeb, A. 1995, Astrophys. J., 443, 11

Ekström, S., Meynet, G., Chiappini, C., Hirschi, R., & Maeder, A. 2008, Astron. and Astrophys., 489, 685

Ellis, R. S., McLure, R. J., Dunlop, J. S., et al. 2013, Astrophys. J. Lett., 763, L7

Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJS, 184, 158

ESA and the Planck Collaboration. 2013, Best Map Ever of the Universe

Fabian, A. C. 1999, Monthly Notices Roy. Astron. Soc., 308, L39

—. 2012, Ann. Rev. Astron. Astrophys., 50, 455

Fall, S. M., & Efstathiou, G. 1980, Monthly Notices Roy. Astron. Soc., 193, 189

Fan, X., Strauss, M. A., Richards, G. T., Hennawi, J. F., & Becker, e. 2006a, Astron. J., 131, 1203

Fan, X., Strauss, M. A., Schneider, D. P., Becker, R. H., & White, e. 2003, Astron. J., 125, 1649

Fan, X., Strauss, M. A., Becker, R. H., et al. 2006b, Astron. J., 132, 117

Faucher-Giguère, C.-A., & Quataert, E. 2012, Monthly Notices Roy. Astron. Soc., 425, 605

Feltre, A., Charlot, S., & Gutkin, J. 2016, Monthly Notices Roy. Astron. Soc., 456, 3354

Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, Rev. Mex. Astron. Astrofis., 49, 137
Fernandez, R., Bryan, G. L., Haiman, Z., & Li, M. 2014, Monthly Notices Roy. Astron. Soc., 439, 3798
Ferrara, A., Salvadori, S., Yue, B., & Schleicher, D. 2014, Monthly Notices Roy. Astron. Soc., 443, 2410
Ferrara, A., & Scannapieco, E. 2016, Astrophys. J., 833, 46
Ferrarese, L. 2002, Astrophys. J., 578, 90
Ferrarese, L., & Merritt, D. 2000, Astrophys. J. Lett., 539, L9
Feruglio, C., Maiolino, R., Piconcelli, E., et al. 2010, Astron. and Astrophys., 518, L155
Field, G. B., Somerville, W. B., & Dressler, K. 1966, Ann. Rev. Astron. Astrophys., 4, 207
Finkelstein, S. L., Papovich, C., Dickinson, M., et al. 2013, Nature, 502, 524
Finkelstein, S. L., Papovich, C., Ryan, R. E., et al. 2012, Astrophys. J., 758, 93
Finn, L. S., & Chernoff, D. F. 1993, Phys. Rev. D, 47, 2198
Förster Schreiber, N. M., et al. 2009, Astrophys. J., 706, 1364
Frebel, A., Johnson, J. L., & Bromm, V. 2007, Monthly Notices Roy. Astron. Soc., 380, L40
Frebel, A., & Norris, J. E. 2015, Ann. Rev. Astron. Astrophys., 53, 631
Freeman, K. C. 1970, Astrophys. J., 160, 811
Fryer, C. L., Woosley, S. E., & Heger, A. 2001, Astrophys. J., 550, 372
Fumagalli, M., O'Meara, J. M., & Prochaska, J. X. 2011, Science, 334, 1245
Gabor, J. M., & Bournaud, F. 2014, Monthly Notices Roy. Astron. Soc., 441, 1615
Galli, D., & Palla, F. 1998, Astron. and Astrophys., 335, 403
Gammie, C. F. 2001, Astrophys. J., 553, 174
Gebhardt, K., Bender, R., Bower, G., et al. 2000, Astrophys. J. Lett., 539, L13
Genzel, R., et al. 2008, Astrophys. J., 687, 59
—. 2017, Nature, 543, 397
Gilli, R., Calura, F., D'Ercole, A., & Norman, C. 2017, ArXiv e-prints, arXiv:1703.05961
Glover, S. 2000, in The First Stars, ed. A. Weiss, T. G. Abel, & V. Hill, 261
Glover, S. 2005, Space Science Reviews, 117, 445
Glover, S. 2013a, in Astrophysics and Space Science Library, Vol. 396, Astrophysics and Space Science Library, ed. T. Wiklind, B. Mobasher, & V. Bromm, 103
Glover, S. 2013b, in Astrophysics and Space Science Library, Vol. 396, The First Galaxies, ed. T. Wiklind, B. Mobasher, & V. Bromm, 103
Glover, S. C. O. 2015a, Monthly Notices Roy. Astron. Soc., 451, 2082
—. 2015b, Monthly Notices Roy. Astron. Soc., 453, 2901
Glover, S. C. O., & Abel, T. 2008, Monthly Notices Roy. Astron. Soc., 388, 1627
Glover, S. C. O., & Brand, P. W. J. L. 2003, Monthly Notices Roy. Astron. Soc., 340, 210
Glover, S. C. O., Federrath, C., Mac Low, M.-M., & Klessen, R. S. 2010, Monthly Notices Roy. Astron. Soc., 404, 2
Glover, S. C. O., & Jappsen, A.-K. 2007, Astrophys. J., 666, 1
Glover, S. C. O., & Mac Low, M.-M. 2007a, ApJS, 169, 239
—. 2007b, Astrophys. J., 659, 1317
Glover, S. C. O., & Savin, D. W. 2009, Monthly Notices Roy. Astron. Soc., 393, 911
Gnedin, N. Y., & Draine, B. T. 2014, Astrophys. J., 795, 37
Gnedin, N. Y., Tassis, K., & Kravtsov, A. V. 2009, Astrophys. J., 697, 55
Gofford, J., Reeves, J. N., Tombesi, F., et al. 2013, Monthly Notices Roy. Astron. Soc., 430, 60
Goldreich, P., & Lynden-Bell, D. 1965, Monthly Notices Roy. Astron. Soc., 130, 125
Goodman, J., & Tan, J. C. 2004, Astrophys. J., 608, 108

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, Astrophys. J., 622, 759
Graff, P. B., Buonanno, A., & Sathyaprakash, B. S. 2015, Phys. Rev. D, 92, 022002
Greene, J. E. 2012, Nature Communications, 3, 1304
Greene, J. E., Ho, L. C., & Barth, A. J. 2008, Astrophys. J., 688, 159
Greif, T. H. 2015, Comput. Astrophys. Cosmol., 2, 3
Greif, T. H., Bromm, V., Clark, P. C., et al. 2012, Monthly Notices Roy. Astron. Soc., 424, 399
Greif, T. H., Johnson, J. L., Klessen, R. S., & Bromm, V. 2008, Monthly Notices Roy. Astron. Soc., 387, 1021
Greif, T. H., Springel, V., & Bromm, V. 2013, Monthly Notices Roy. Astron. Soc., 434, 3408
Greif, T. H., Springel, V., White, S. D. M., et al. 2011, Astrophys. J., 737, 75
Grimm, H.-J., Gilfanov, M., & Sunyaev, R. 2003, Monthly Notices Roy. Astron. Soc., 339, 793
Gültekin, K., Richstone, D. O., Gebhardt, K., et al. 2009, Astrophys. J., 698, 198
Gutkin, J., Charlot, S., & Bruzual, G. 2016, Monthly Notices Roy. Astron. Soc., 462, 1757
Haardt, F., Ripamonti, E., Colpi, M., & Ferrara, A. 2002, Astrophys. and Space Science, 281, 479
Habouzit, M., Volonteri, M., & Dubois, Y. 2017, Monthly Notices Roy. Astron. Soc., 468, 3935
Habouzit, M., Volonteri, M., Latif, M., Dubois, Y., & Peirani, S. 2016a, Monthly Notices Roy. Astron. Soc., 463, 529
Habouzit, M., Volonteri, M., Latif, M., et al. 2016b, Monthly Notices Roy. Astron. Soc., 456, 1901
—. 2016c, Monthly Notices Roy. Astron. Soc., 456, 1901
Haehnelt, M. G. 1994, Monthly Notices Roy. Astron. Soc., 269, 199
Haehnelt, M. G., Natarajan, P., & Rees, M. J. 1998, Monthly Notices Roy. Astron. Soc., 300, 817
Hahn, O., & Abel, T. 2011, Monthly Notices Roy. Astron. Soc., 415, 2101
Haiman, Z. 2004, Astrophys. J., 613, 36
Haiman, Z. 2013, in Astrophysics and Space Science Library, Vol. 396, Astrophysics and Space Science Library, ed. T. Wiklind, B. Mobasher, & V. Bromm, 293
Haiman, Z., Abel, T., & Rees, M. J. 2000, Astrophys. J., 534, 11
Haiman, Z., & Loeb, A. 2001, Astrophys. J., 552, 459
Haiman, Z., Rees, M. J., & Loeb, A. 1997, Astrophys. J., 476, 458
Häring, N., & Rix, H.-W. 2004, Astrophys. J. Lett., 604, L89
Hartwig, T. 2014, Master's thesis, University of Heidelberg, Germany
Hartwig, T. 2017, Formation of the First Black Holes, ed. D. R. G. Schleicher & M. A. Latif, Vol. in prep. (World Scientific)
Hartwig, T., Bromm, V., Klessen, R. S., & Glover, S. C. O. 2015a, Monthly Notices Roy. Astron. Soc., 447, 3892
Hartwig, T., Bromm, V., & Loeb, A. 2017a, MNRAS in prep.
Hartwig, T., Clark, P. C., Glover, S. C. O., Klessen, R. S., & Sasaki, M. 2015b, Astrophys. J., 799, 114
Hartwig, T., Glover, S. C. O., Klessen, R. S., Latif, M. A., & Volonteri, M. 2015c, Monthly Notices Roy. Astron. Soc., 452, 1233
Hartwig, T., Volonteri, M., Bromm, V., et al. 2016a, Monthly Notices Roy. Astron. Soc., 460, L74
Hartwig, T., Volonteri, M., & Dashyan, G. 2017b, ArXiv e-prints, arXiv:1707.03826
Hartwig, T., Latif, M. A., Magg, M., et al. 2016b, Monthly Notices Roy. Astron. Soc., 462, 2184
Hartwig, T., Yoshida, N., Klessen, R. S., et al. 2017c, MNRAS in prep.
Haster, C.-J., Wang, Z., Berry, C. P. L., et al. 2016, Monthly Notices Roy. Astron. Soc., 457, 4499
Hawking, S. 1971, Monthly Notices Roy. Astron. Soc., 152, 75
Hawking, S. W. 1975, Communications in Mathematical Physics, 43, 199

Heckman, T. M., & Kauffmann, G. 2011, Science, 333, 182
Heger, A., Fryer, C. L., Woosley, S. E., Langer, N., & Hartmann, D. H. 2003, Astrophys. J., 591, 288
Heger, A., & Woosley, S. E. 2002, Astrophys. J., 567, 532
—. 2010, Astrophys. J., 724, 341
Hernquist, L. 1990, Astrophys. J., 356, 359
Hirano, S., Hosokawa, T., Yoshida, N., et al. 2014, Astrophys. J., 781, 60
Hirasawa, T. 1969, Progress of Theoretical Physics, 42, 523
Hirschmann, M., Dolag, K., Saro, A., et al. 2014, Monthly Notices Roy. Astron. Soc., 442, 2304
Hosokawa, T., Hirano, S., Kuiper, R., et al. 2016, Astrophys. J., 824, 119
Hosokawa, T., Omukai, K., & Yorke, H. W. 2012a, Astrophys. J., 756, 93
Hosokawa, T., Omukai, K., Yoshida, N., & Yorke, H. W. 2011, Science, 334, 1250
Hosokawa, T., Omukai, K., Yoshida, N., & Yorke, H. W. 2012b, in American Institute of Physics Conference Series, Vol. 1480, American Institute of Physics Conference Series, ed. M. Umemura & K. Omukai, 91–96
Hosokawa, T., Yorke, H. W., Inayoshi, K., Omukai, K., & Yoshida, N. 2013, Astrophys. J., 778, 178
Hummel, J. A., Pawlik, A. H., Milosavljević, M., & Bromm, V. 2012, Astrophys. J., 755, 72
Hutchins, J. B. 1976, Astrophys. J., 205, 103
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, Astron. and Astrophys., 556, A55
Iliev, I. T., Scannapieco, E., Martel, H., & Shapiro, P. R. 2003, Monthly Notices Roy. Astron. Soc., 341, 81
Inayoshi, K., & Haiman, Z. 2014, Monthly Notices Roy. Astron. Soc., 445, 1549
Inayoshi, K., Hirai, R., Kinugawa, T., & Hotokezaka, K. 2017, Monthly Notices Roy. Astron. Soc., 468, 5020
Inayoshi, K., Kashiyama, K., Visbal, E., & Haiman, Z. 2016, Monthly Notices Roy. Astron. Soc., 461, 2722
Inayoshi, K., & Omukai, K. 2012, Monthly Notices Roy. Astron. Soc., 422, 2539
Inayoshi, K., Omukai, K., & Tasker, E. 2014, Monthly Notices Roy. Astron. Soc., 445, L109
Inayoshi, K., & Tanaka, T. L. 2015a, Monthly Notices Roy. Astron. Soc., 450, 4350
—. 2015b, Monthly Notices Roy. Astron. Soc., 450, 4350
Ishibashi, W., & Fabian, A. C. 2012, Monthly Notices Roy. Astron. Soc., 427, 2998
—. 2014, Monthly Notices Roy. Astron. Soc., 441, 1474
—. 2015, Monthly Notices Roy. Astron. Soc., 451, 93
Ishiyama, T., Sudo, K., Yokoi, S., et al. 2016, Astrophys. J., 826, 9
Jeon, M., Pawlik, A. H., Bromm, V., & Milosavljević, M. 2014, Monthly Notices Roy. Astron. Soc., 440, 3778
Jeon, M., Pawlik, A. H., Greif, T. H., et al. 2012, Astrophys. J., 754, 34
Johnson, J. L., & Bromm, V. 2007, Monthly Notices Roy. Astron. Soc., 374, 1557
Johnson, J. L., Dalla Vecchia, C., & Khochfar, S. 2013a, Monthly Notices Roy. Astron. Soc., 428, 1857
Johnson, J. L., & Dijkstra, M. 2017, Astron. and Astrophys., 601, A138
Johnson, J. L., Greif, T. H., & Bromm, V. 2007, Astrophys. J., 665, 85
—. 2008, Monthly Notices Roy. Astron. Soc., 388, 26
Johnson, J. L., Greif, T. H., Bromm, V., Klessen, R. S., & Ippolito, J. 2009, Monthly Notices Roy. Astron. Soc., 399, 37
Johnson, J. L., Whalen, D. J., Fryer, C. L., & Li, H. 2012, Astrophys. J., 750, 66
Johnson, J. L., Whalen, D. J., Li, H., & Holz, D. E. 2013b, Astrophys. J., 771, 116

Karlsson, T., Bromm, V., & Bland-Hawthorn, J. 2013, Reviews of Modern Physics, 85, 809
Kashlinsky, A., & Rees, M. J. 1983, Monthly Notices Roy. Astron. Soc., 205, 955
Katz, H., Sijacki, D., & Haehnelt, M. G. 2015, Monthly Notices Roy. Astron. Soc., 451, 2352
Kehrig, C., Vílchez, J. M., Pérez-Montero, E., et al. 2015, Astrophys. J. Lett., 801, L28
—. 2016, Monthly Notices Roy. Astron. Soc., doi:10.1093/mnras/stw806
Keller, S. C., Bessell, M. S., Frebel, A., et al. 2014, Nature, 506, 463
Khlopov, M. Y. 2010, Research in Astronomy and Astrophysics, 10, 495
King, A. 2003, Astrophys. J. Lett., 596, L27
—. 2005, Astrophys. J. Lett., 635, L121
King, A., & Pounds, K. 2015, Ann. Rev. Astron. Astrophys., 53, 115
King, A. R. 2010, Monthly Notices Roy. Astron. Soc., 402, 1516
King, A. R., & Pounds, K. A. 2014, Monthly Notices Roy. Astron. Soc., 437, L81
King, A. R., Pringle, J. E., & Hofmann, J. A. 2008, Monthly Notices Roy. Astron. Soc., 385, 1621
King, A. R., Zubovas, K., & Power, C. 2011, Monthly Notices Roy. Astron. Soc., 415, L6
Kinugawa, T., Inayoshi, K., Hotokezaka, K., Nakauchi, D., & Nakamura, T. 2014, Monthly Notices Roy. Astron. Soc., 442, 2963
Kinugawa, T., Miyamoto, A., Kanda, N., & Nakamura, T. 2016a, Monthly Notices Roy. Astron. Soc., 456, 1093
Kinugawa, T., Nakamura, T., & Nakano, H. 2016b, ArXiv e-prints, arXiv:1610.00305
Kitayama, T., Yoshida, N., Susa, H., & Umemura, M. 2004, Astrophys. J., 613, 631
Klessen, R. S. 2002, Presentation "Lecture on SPH Basics" held at the EU Network Conference "Numerical Modelling of Young Star Clusters" in Cardiff, Wales
Klessen, R. S. 2011, in EAS Publications Series, Vol. 51, EAS Publications Series, ed. C. Charbonnel & T. Montmerle, 133–167
Kormendy, J., & Ho, L. C. 2013, Ann. Rev. Astron. Astrophys., 51, 511
Koushiappas, S. M., Bullock, J. S., & Dekel, A. 2004, Monthly Notices Roy. Astron. Soc., 354, 292
Kroupa, P. 1995, Monthly Notices Roy. Astron. Soc., 277, astro-ph/9508084
Krumholz, M. R. 2012, Astrophys. J., 759, 9
Krumholz, M. R., & McKee, C. F. 2008, Nature, 451, 1082
Kulczycki, K., Bulik, T., Belczyński, K., & Rudak, B. 2006, Astron. and Astrophys., 459, 1001
Kullback, S., & Leibler, R. A. 1951, The annals of mathematical statistics, 22, 79
Lacey, C., & Cole, S. 1993, Monthly Notices Roy. Astron. Soc., 262, 627
Laigle, C., Pichon, C., Codis, S., et al. 2015, Monthly Notices Roy. Astron. Soc., 446, 2744
Lambas, D. G., Tissera, P. B., Alonso, M. S., & Coldwell, G. 2003, Monthly Notices Roy. Astron. Soc., 346, 1189
Landau, L. D., & Lifshitz, E. M. 1987, Fluid mechanics (Oxford, England; New York: Pergamon Press)
Latif, M. A., Bovino, S., Grassi, T., Schleicher, D. R. G., & Spaans, M. 2015, Monthly Notices Roy. Astron. Soc., 446, 3163
Latif, M. A., Bovino, S., Van Borm, C., et al. 2014, Monthly Notices Roy. Astron. Soc., 443, 1979
Latif, M. A., Omukai, K., Habouzit, M., Schleicher, D. R. G., & Volonteri, M. 2016a, Astrophys. J., 823, 40
—. 2016b, Astrophys. J., 823, 40
Latif, M. A., & Schleicher, D. R. G. 2016, Astron. and Astrophys., 585, A151
Latif, M. A., Schleicher, D. R. G., & Hartwig, T. 2016c, Monthly Notices Roy. Astron. Soc., 458, 233
Latif, M. A., Schleicher, D. R. G., Schmidt, W., & Niemeyer, J. 2013a, Astrophys. J. Lett., 772, L3

Latif, M. A., Schleicher, D. R. G., Schmidt, W., & Niemeyer, J. C. 2013b, Monthly Notices Roy. Astron. Soc., 433, 1607

—. 2013c, Monthly Notices Roy. Astron. Soc., 436, 2989

Latif, M. A., Schleicher, D. R. G., Spaans, M., & Zaroubi, S. 2011a, Monthly Notices Roy. Astron. Soc., 413, L33

Latif, M. A., & Volonteri, M. 2015, Monthly Notices Roy. Astron. Soc., 452, 1026

Latif, M. A., Zaroubi, S., & Spaans, M. 2011b, Monthly Notices Roy. Astron. Soc., 411, 1659

Laursen, P., Sommer-Larsen, J., & Andersen, A. C. 2009, Astrophys. J., 704, 1640

Limongi, M. 2017, ArXiv e-prints, arXiv:1706.01913

Lodato, G., & Natarajan, P. 2006, Monthly Notices Roy. Astron. Soc., 371, 1813

Loeb, A. 2010, How Did the First Stars and Galaxies Form? (Princeton University Press)

—. 2016, Astrophys. J. Lett., 819, L21

Loeb, A., & Rasio, F. A. 1994, Astrophys. J., 432, 52

Lupi, A., Colpi, M., Devecchi, B., Galanti, G., & Volonteri, M. 2014, Monthly Notices Roy. Astron. Soc., 442, 3616

Lupi, A., Haardt, F., Dotti, M., et al. 2016, Monthly Notices Roy. Astron. Soc., 456, 2993

Ma, Q., Maio, U., Ciardi, B., & Salvaterra, R. 2016, ArXiv e-prints, arXiv:1610.03594

Mac Low, M.-M., & Ferrara, A. 1999, Astrophys. J., 513, 142

Machacek, M. E., Bryan, G. L., & Abel, T. 2001, Astrophys. J., 548, 509

Machida, M. N., & Doi, K. 2013, Monthly Notices Roy. Astron. Soc., 435, 3283

Mackey, J., Bromm, V., & Hernquist, L. 2003, Astrophys. J., 586, 1

Madau, P., & Haardt, F. 2015, Astrophys. J. Lett., 813, L8

Madau, P., Haardt, F., & Dotti, M. 2014, Astrophys. J. Lett., 784, L38

Madau, P., & Rees, M. J. 2001, Astrophys. J. Lett., 551, L27

Madau, P., & Silk, J. 2005, Monthly Notices Roy. Astron. Soc., 359, L37

Magg, M., Hartwig, T., Agarwal, B., et al. 2017, ArXiv e-prints, arXiv:1706.07054

Magg, M., Hartwig, T., Glover, S. C. O., Klessen, R. S., & Whalen, D. J. 2016, Monthly Notices Roy. Astron. Soc., 462, 3591

Maggiore, M. 2007, Gravitational Waves. Vol. 1: Theory and Experiments, Oxford Master Series in Physics (Oxford University Press)

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, Astron. J., 115, 2285

Maiolino, R., Russell, H. R., Fabian, A. C., et al. 2017, Nature, 544, 202

Malhotra, S., & Rhoads, J. E. 2002, Astrophys. J. Lett., 565, L71

Mapelli, M. 2016, Monthly Notices Roy. Astron. Soc., 459, 3432

Mapelli, M., & Bressan, A. 2013, Monthly Notices Roy. Astron. Soc., 430, 3120

Mapelli, M., Giacobbo, N., Ripamonti, E., & Spera, M. 2017, ArXiv e-prints, arXiv:1708.05722

Mapelli, M., Huwyler, C., Mayer, L., Jetzer, P., & Vecchio, A. 2010, Astrophys. J., 719, 987

Mapelli, M., Ripamonti, E., Zampieri, L., & Colpi, M. 2011, Monthly Notices Roy. Astron. Soc., 416, 1756

Mapelli, M., Zampieri, L., Ripamonti, E., & Bressan, A. 2013, Monthly Notices Roy. Astron. Soc., 429, 2298

Marconi, A., & Hunt, L. K. 2003, Astrophys. J. Lett., 589, L21

Marconi, A., Risaliti, G., Gilli, R., et al. 2004, Monthly Notices Roy. Astron. Soc., 351, 169

Martin, C. L. 2005, Astrophys. J., 621, 227

Martin, P. G., Schwarz, D. H., & Mandy, M. E. 1996, Astrophys. J., 461, 265

Matsuda, T., Satō, H., & Takeda, H. 1969, Progress of Theoretical Physics, 42, 219
Matthee, J., Sobral, D., Oteo, I., et al. 2016, Monthly Notices Roy. Astron. Soc., 458, 449
Matthee, J., Sobral, D., Santos, S., et al. 2015, Monthly Notices Roy. Astron. Soc., 451, 4919
Mayer, L., Fiacconi, D., Bonoli, S., et al. 2015, Astrophys. J., 810, 51
Mayer, L., Kazantzidis, S., Escala, A., & Callegari, S. 2010, Nature, 466, 1082
McConnell, N. J., & Ma, C.-P. 2013, Astrophys. J., 764, 184
McKee, C. F., & Tan, J. C. 2008, Astrophys. J., 681, 771
McLeod, D. J., McLure, R. J., & Dunlop, J. S. 2016, Monthly Notices Roy. Astron. Soc., 459, 3812
McLeod, D. J., McLure, R. J., Dunlop, J. S., et al. 2015, Monthly Notices Roy. Astron. Soc., 450, 3032
McQuillin, R. C., & McLaughlin, D. E. 2012, Monthly Notices Roy. Astron. Soc., 423, 2162
—. 2013, Monthly Notices Roy. Astron. Soc., 434, 1332
Miller, M. C., & Davies, M. B. 2012, Astrophys. J., 755, 81
Milosavljević, M., Bromm, V., Couch, S. M., & Oh, S. P. 2009a, Astrophys. J., 698, 766
Milosavljević, M., Couch, S. M., & Bromm, V. 2009b, Astrophys. J. Lett., 696, L146
Mineo, S., Gilfanov, M., & Sunyaev, R. 2012, Monthly Notices Roy. Astron. Soc., 419, 2095
Mo, H. J., Mao, S., & White, S. D. M. 1998, Monthly Notices Roy. Astron. Soc., 295, 319
Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, Nature, 474, 616
Murray, N., Quataert, E., & Thompson, T. A. 2005, Astrophys. J., 618, 569
Nagao, T., Maiolino, R., & Marconi, A. 2006, Astron. and Astrophys., 447, 863
Nandra, K., Barret, D., Barcons, X., et al. 2013, ArXiv e-prints, arXiv:1306.2307
Negri, A., & Volonteri, M. 2017, Monthly Notices Roy. Astron. Soc., 467, 3475
Nelson, R. P., & Langer, W. D. 1997, Astrophys. J., 482, 796
Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., & Maeda, K. 2006, Nuclear Physics A, 777, 424
Oesch, P. A., van Dokkum, P. G., Illingworth, G. D., et al. 2015, Astrophys. J. Lett., 804, L30
Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, Astrophys. J., 819, 129
Oh, S., Kroupa, P., & Pflamm-Altenburg, J. 2015, Astrophys. J., 805, 92
Oh, S. P., Haiman, Z., & Rees, M. J. 2001, Astrophys. J., 553, 73
Ohkubo, T., Nomoto, K., Umeda, H., Yoshida, N., & Tsuruta, S. 2009, Astrophys. J., 706, 1184
Ohkubo, T., Umeda, H., Maeda, K., et al. 2006, Astrophys. J., 645, 1352
Omukai, K. 2001, Astrophys. J., 546, 635
Omukai, K., & Nishi, R. 1998, Astrophys. J., 508, 141
Omukai, K., & Palla, F. 2001, Astrophys. J., 561, L55
Omukai, K., Schneider, R., & Haiman, Z. 2008, Astrophys. J., 686, 801
Omukai, K., Tsuribe, T., Schneider, R., & Ferrara, A. 2005, Astrophys. J., 626, 627
Omukai, K., & Yoshii, Y. 2003, Astrophys. J., 599, 746
Ono, Y., Ouchi, M., Mobasher, B., et al. 2012, Astrophys. J., 744, 83
O'Shea, B. W., & Norman, M. L. 2007, Astrophys. J., 654, 66
Ostriker, J. P., Choi, E., Ciotti, L., Novak, G. S., & Proga, D. 2010, Astrophys. J., 722, 642
Ouchi, M., Ellis, R., Ono, Y., et al. 2013, Astrophys. J., 778, 102
Ouchi, M., Ono, Y., Egami, E., et al. 2009, Astrophys. J., 696, 1164
Ouchi, M., Shimasaku, K., Akiyama, M., et al. 2008, ApJS, 176, 301
Ouchi, M., Shimasaku, K., Furusawa, H., et al. 2010, Astrophys. J., 723, 869
Paardekooper, J.-P., Khochfar, S., & Dalla Vecchia, C. 2013, Monthly Notices Roy. Astron. Soc., 429, L94
Pacucci, F., & Ferrara, A. 2015, Monthly Notices Roy. Astron. Soc., 448, 104

Pacucci, F., Ferrara, A., Grazian, A., et al. 2016, Monthly Notices Roy. Astron. Soc., 459, 1432
Pacucci, F., Ferrara, A., Volonteri, M., & Dubus, G. 2015a, Monthly Notices Roy. Astron. Soc., 454, 3771
Pacucci, F., Loeb, A., & Salvadori, S. 2017a, ArXiv e-prints, arXiv:1706.09892
Pacucci, F., Pallottini, A., Ferrara, A., & Gallerani, S. 2017b, Monthly Notices Roy. Astron. Soc., 468, L77
Pacucci, F., Volonteri, M., & Ferrara, A. 2015b, Monthly Notices Roy. Astron. Soc., 452, 1922
—. 2015c, Monthly Notices Roy. Astron. Soc., 452, 1922
Palla, F., Salpeter, E. E., & Stahler, S. W. 1983, Astrophys. J., 271, 632
Pallottini, A., Ferrara, A., Gallerani, S., Salvadori, S., & D'Odorico, V. 2014, Monthly Notices Roy. Astron. Soc., 440, 2498
Pallottini, A., Ferrara, A., Pacucci, F., et al. 2015, Monthly Notices Roy. Astron. Soc., 453, 2465
Park, K., & Ricotti, M. 2011a, Astrophys. J., 739, 2
—. 2011b, Astrophys. J., 739, 2
—. 2012, Astrophys. J., 747, 9
—. 2013, Astrophys. J., 767, 163
Park, K., Ricotti, M., Natarajan, P., Bogdanović, T., & Wise, J. H. 2016, Astrophys. J., 818, 184
Parkinson, H., Cole, S., & Helly, J. 2008, Monthly Notices Roy. Astron. Soc., 383, 557
Pawlik, A. H., Schaye, J., & van Scherpenzeel, E. 2009, Monthly Notices Roy. Astron. Soc., 394, 1812
Peebles, P. J. E., & Dicke, R. H. 1968, Astrophys. J., 154, 891
Pelupessy, F. I., Di Matteo, T., & Ciardi, B. 2007, Astrophys. J., 665, 107
Pentericci, L., Vanzella, E., Fontana, A., et al. 2014, Astrophys. J., 793, 113
Perna, R., Lazzati, D., & Giacomazzo, B. 2016, Astrophys. J. Lett., 821, L18
Peters, T., Klessen, R. S., Mac Low, M.-M., & Banerjee, R. 2010, Astrophys. J., 725, 134
Peters, T., Schleicher, D. R. G., Smith, R. J., Schmidt, W., & Klessen, R. S. 2014, Monthly Notices Roy. Astron. Soc., 442, 3112
Phinney, E. S. 2001, ArXiv e-prints, astro-ph/0108028
Placco, V. M., Frebel, A., Beers, T. C., et al. 2016, ArXiv e-prints, arXiv:1609.02134
Planck Collaboration. 2014, Astron. and Astrophys., 571, A16
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, Astron. and Astrophys., 594, A13
Portegies Zwart, S. F., Baumgardt, H., Hut, P., Makino, J., & McMillan, S. L. W. 2004, Nature, 428, 724
Pounds, K. A., Reeves, J. N., King, A. R., et al. 2003, Monthly Notices Roy. Astron. Soc., 345, 705
Press, W. H., & Schechter, P. 1974, Astrophys. J., 187, 425
Prieto, J., Escala, A., Volonteri, M., & Dubois, Y. 2017, Astrophys. J., 836, 216
Proga, D., Stone, J. M., & Kallman, T. R. 2000, Astrophys. J., 543, 686
Raiter, A., Schaerer, D., & Fosbury, R. A. E. 2010, Astron. and Astrophys., 523, A64
Rees, M. J. 1984, Ann. Rev. Astron. Astrophys., 22, 471
Reeves, J. N., O'Brien, P. T., & Ward, M. J. 2003, Astrophys. J. Lett., 593, L65
Regan, J. A., Johansson, P. H., & Wise, J. H. 2014a, Astrophys. J., 795, 137
—. 2014b, Astrophys. J., 795, 137
—. 2015, Monthly Notices Roy. Astron. Soc., 449, 3766
—. 2016, Monthly Notices Roy. Astron. Soc., 459, 3377
Regan, J. A., Visbal, E., Wise, J. H., et al. 2017, Nature Astronomy, 1, 0075
Reines, A. E., Greene, J. E., & Geha, M. 2013, Astrophys. J., 775, 116

Reines, A. E., & Volonteri, M. 2015, Astrophys. J., 813, 82

Richardson, C. T., Allen, J. T., Baldwin, J. A., Hewett, P. C., & Ferland, G. J. 2014, Monthly Notices Roy. Astron. Soc., 437, 2376

Richings, A. J., & Faucher-Giguere, C.-A. 2017, ArXiv e-prints, arXiv:1706.03784

Richings, A. J., Schaye, J., & Oppenheimer, B. D. 2014, Monthly Notices Roy. Astron. Soc., 442, 2780

Ripamonti, E., & Abel, T. 2004, Monthly Notices Roy. Astron. Soc., 348, 1019

Ripamonti, E., Haardt, F., Ferrara, A., & Colpi, M. 2002, Monthly Notices Roy. Astron. Soc., 334, 401

Ritter, J. S., Sluder, A., Safranek-Shrader, C., Milosavljević, M., & Bromm, V. 2015, Monthly Notices Roy. Astron. Soc., 451, 1190

Rosado, P. A. 2011, Phys. Rev. D, 84, 084004

Rupke, D. S. N., & Veilleux, S. 2011, Astrophys. J. Lett., 729, L27

Sabano, Y., & Yoshii, Y. 1977, Publ. of the Astron. Soc. of Japan, 29, 207

Safranek-Shrader, C., Agarwal, M., Federrath, C., et al. 2012a, Monthly Notices Roy. Astron. Soc., 426, 1159

—. 2012b, Monthly Notices Roy. Astron. Soc., 426, 1159

Safranek-Shrader, C., Milosavljević, M., & Bromm, V. 2014, Monthly Notices Roy. Astron. Soc., 438, 1669

Sakurai, Y., Vorobyov, E. I., Hosokawa, T., et al. 2016, Monthly Notices Roy. Astron. Soc., 459, 1137

Sakurai, Y., Yoshida, N., Fujii, M. S., & Hirano, S. 2017, ArXiv e-prints, arXiv:1704.06130

Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Science, 337, 444

Sanders, D. B., & Mirabel, I. F. 1996, Ann. Rev. Astron. Astrophys., 34, 749

Saslaw, W. C., & Zipoy, D. 1967, Nature, 216, 976

Sazonov, S. Y., Ostriker, J. P., Ciotti, L., & Sunyaev, R. A. 2005, Monthly Notices Roy. Astron. Soc., 358, 168

Sazonov, S. Y., Ostriker, J. P., & Sunyaev, R. A. 2004, Monthly Notices Roy. Astron. Soc., 347, 144

Scannapieco, E., Madau, P., Woosley, S., Heger, A., & Ferrara, A. 2005, Astrophys. J., 633, 1031

Schaerer, D. 2002, Astron. and Astrophys., 382, 28

—. 2003, Astron. and Astrophys., 397, 527

Schauer, A. T. P., Whalen, D. J., Glover, S. C. O., & Klessen, R. S. 2015, Monthly Notices Roy. Astron. Soc., 454, 2441

Schauer, A. T. P., Agarwal, B., Glover, S. C. O., et al. 2017, Monthly Notices Roy. Astron. Soc., 467, 2288

Schleicher, D. R. G., Palla, F., Ferrara, A., Galli, D., & Latif, M. 2013, Astron. and Astrophys., 558, A59

Schleicher, D. R. G., Spaans, M., & Glover, S. C. O. 2010, Astrophys. J. Lett., 712, L69

Schneider, R., Ferrara, A., Ciardi, B., Ferrari, V., & Matarrese, S. 2000, Monthly Notices Roy. Astron. Soc., 317, 385

Schneider, R., Ferrara, A., Natarajan, P., & Omukai, K. 2002, Astrophys. J., 571, 30

Schneider, R., Ferrari, V., Matarrese, S., & Portegies Zwart, S. F. 2001, Monthly Notices Roy. Astron. Soc., 324, 797

Schneider, R., Graziani, L., Marassi, S., et al. 2017, ArXiv e-prints, arXiv:1705.06781

Schneider, R., Omukai, K., Bianchi, S., & Valiante, R. 2012, Monthly Notices Roy. Astron. Soc., 419, 1566

Schober, J., Schleicher, D., Federrath, C., et al. 2012, Astrophys. J., 754, 99

SDSS. 2008, Online Material: SDSS Legacy Survey (http://www.sdss.org/legacy/)

Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41
Sesana, A. 2016, Physical Review Letters, 116, 231102
Sesana, A., Gair, J., Mandel, I., & Vecchio, A. 2009, Astrophys. J. Lett., 698, L129
Sesana, A., Volonteri, M., & Haardt, F. 2007, Monthly Notices Roy. Astron. Soc., 377, 1711
Sethi, S., Haiman, Z., & Pandey, K. 2010, Astrophys. J., 721, 615
Seto, N. 1999, Astrophys. J., 523, 24
Shakura, N. I., & Sunyaev, R. A. 1973, Astron. and Astrophys., 24, 337
Shang, C., Bryan, G. L., & Haiman, Z. 2010, Monthly Notices Roy. Astron. Soc., 402, 1249
Shapiro, P. R., Giroux, M. L., & Babul, A. 1994, Astrophys. J., 427, 25
Shapiro, S. L. 2004, Coevolution of Black Holes and Galaxies, 103
—. 2005, Astrophys. J., 620, 59
Sharma, M., & Nath, B. B. 2013, Astrophys. J., 763, 17
Sheth, R. K., Mo, H. J., & Tormen, G. 2001, Monthly Notices Roy. Astron. Soc., 323, 1
Shibuya, T., Ouchi, M., Harikane, Y., et al. 2017, ArXiv e-prints, arXiv:1705.00733
Shields, J. C., & Ferland, G. J. 1993, Astrophys. J., 402, 425
Shlosman, I., Choi, J.-H., Begelman, M. C., & Nagamine, K. 2016, Monthly Notices Roy. Astron. Soc., 456, 500
Shlosman, I., Frank, J., & Begelman, M. C. 1989, Nature, 338, 45
Shull, J. M., & Draine, B. T. 1987, in Astrophysics and Space Science Library, Vol. 134, Interstellar Processes, ed. D. J. Hollenbach & H. A. Thronson, Jr., 283–319
Silk, J. 1977, Astrophys. J., 211, 638
—. 1983, Monthly Notices Roy. Astron. Soc., 205, 705
Silk, J., & Nusser, A. 2010, Astrophys. J., 725, 556
Silk, J., & Rees, M. J. 1998, Astron. and Astrophys., 331, L1
Silk, J., & Vilenkin, A. 1984, Physical Review Letters, 53, 1700
Smidt, J., Wiggins, B. K., & Johnson, J. L. 2016, Astrophys. J. Lett., 829, L6
Smit, R., Bouwens, R. J., Carniani, S., et al. 2017, ArXiv e-prints, arXiv:1706.04614
Smith, A., Bromm, V., & Loeb, A. 2016, Monthly Notices Roy. Astron. Soc., 460, 3143
Smith, A., Safranek-Shrader, C., Bromm, V., & Milosavljević, M. 2015a, Monthly Notices Roy. Astron. Soc., 449, 4336
Smith, B. D., Wise, J. H., O'Shea, B. W., Norman, M. L., & Khochfar, S. 2015b, Monthly Notices Roy. Astron. Soc., 452, 2822
Smith, B. J., Struck, C., Hancock, M., et al. 2007, Astron. J., 133, 791
Smith, R. J., Glover, S. C. O., Clark, P. C., Greif, T., & Klessen, R. S. 2011, Monthly Notices Roy. Astron. Soc., 414, 3633
Sobolev, V. V. 1960, Moving envelopes of stars
Sobral, D., Matthee, J., Darvish, B., et al. 2015, Astrophys. J., 808, 139
Solomon, P. M. 1965, PhD thesis, The University of Wisconsin - Madison.
Soltan, A. 1982, Monthly Notices Roy. Astron. Soc., 200, 115
Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, Monthly Notices Roy. Astron. Soc., 391, 481
Spera, M., Giacobbo, N., & Mapelli, M. 2016, Memorie della Societa Astronomica Italiana, 87, 575
Spera, M., & Mapelli, M. 2017, ArXiv e-prints, arXiv:1706.06109
Stacy, A., & Bromm, V. 2013, Monthly Notices Roy. Astron. Soc., 433, 1094
Stacy, A., Bromm, V., & Lee, A. T. 2016, Monthly Notices Roy. Astron. Soc., 462, 1307

Stacy, A., Greif, T. H., & Bromm, V. 2010, Monthly Notices Roy. Astron. Soc., 403, 45

—. 2012, Monthly Notices Roy. Astron. Soc., 422, 290

Stacy, A., Greif, T. H., Klessen, R. S., Bromm, V., & Loeb, A. 2013, Monthly Notices Roy. Astron. Soc., 431, 1470

Stahler, S. W., Palla, F., & Salpeter, E. E. 1986, Astrophys. J., 302, 590

Stancil, P. C., Lepp, S., & Dalgarno, A. 1998, Astrophys. J., 509, 1

Stark, D. P., Walth, G., Charlot, S., et al. 2015, Monthly Notices Roy. Astron. Soc., 454, 1393

Stecher, T. P., & Williams, D. A. 1967, Astrophys. J. Lett., 149, L29

Sturm, E., et al. 2011, Astrophys. J. Lett., 733, L16

Sugimura, K., Hosokawa, T., Yajima, H., & Omukai, K. 2017, Monthly Notices Roy. Astron. Soc., 469, 62

Sugimura, K., Omukai, K., & Inoue, A. K. 2014, Monthly Notices Roy. Astron. Soc., 445, 544

Sur, S., Schleicher, D. R. G., Banerjee, R., Federrath, C., & Klessen, R. S. 2010, Astrophys. J. Lett., 721, L134

Susa, H., Hasegawa, K., & Tominaga, N. 2014, Astrophys. J., 792, 32

Sutherland, R. S., & Dopita, M. A. 1993, ApJS, 88, 253

Taffoni, G., Mayer, L., Colpi, M., & Governato, F. 2003, Monthly Notices Roy. Astron. Soc., 341, 434

Takeda, H., Satō, H., & Matsuda, T. 1969, Progress of Theoretical Physics, 41, 840

Tanaka, T., & Haiman, Z. 2009, Astrophys. J., 696, 1798

Tegmark, M., Silk, J., Rees, M. J., et al. 1997, Astrophys. J., 474, 1

Thompson, T. A., Fabian, A. C., Quataert, E., & Murray, N. 2015, Monthly Notices Roy. Astron. Soc., 449, 147

Thrane, E., & Romano, J. D. 2013, Phys. Rev. D, 88, 124032

Tombesi, F. 2016, Astronomische Nachrichten, 337, 410

Tombesi, F., Cappi, M., Reeves, J. N., et al. 2010, Astron. and Astrophys., 521, A57

Toomre, A. 1964, Astrophys. J., 139, 1217

Tozzi, P., & Norman, C. 2001, Astrophys. J., 546, 63

Trebitsch, M., Blaizot, J., & Rosdahl, J. 2015, ArXiv e-prints, arXiv:1510.06949

Tremaine, S., Gebhardt, K., Bender, R., et al. 2002, Astrophys. J., 574, 740

Tumlinson, J., & Shull, J. M. 2000, Astrophys. J. Lett., 528, L65

Turk, M. J., Abel, T., & O'Shea, B. 2009, Science, 325, 601

Turk, M. J., Clark, P., Glover, S. C. O., et al. 2011, Astrophys. J., 726, 55

Uehara, H., Susa, H., Nishi, R., Yamada, M., & Nakamura, T. 1996, Astrophys. J. Lett., 473, L95

Valiante, R., Agarwal, B., Habouzit, M., & Pezzulli, E. 2017, ArXiv e-prints, arXiv:1703.03808

Valiante, R., Schneider, R., Volonteri, M., & Omukai, K. 2016, Monthly Notices Roy. Astron. Soc., 457, 3356

Van Borm, C., Bovino, S., Latif, M. A., et al. 2014, Astron. and Astrophys., 572, A22

Van Wassenhove, S., Capelo, P. R., Volonteri, M., et al. 2014, Monthly Notices Roy. Astron. Soc., 439, 474

Vanzella, E., Pentericci, L., Fontana, A., et al. 2011, Astrophys. J. Lett., 730, L35

Veitch, J., Pürrer, M., & Mandel, I. 2015, Physical Review Letters, 115, 141101

Venemans, B. 2015, Nature, 518, 490

Venemans, B. P., Findlay, J. R., Sutherland, W. J., et al. 2013a, Astrophys. J., 779, 24

—. 2013b, Astrophys. J., 779, 24

Visbal, E., Bryan, G. L., & Haiman, Z. 2017, Monthly Notices Roy. Astron. Soc., 469, 1456

Visbal, E., Haiman, Z., & Bryan, G. L. 2014a, Monthly Notices Roy. Astron. Soc., 445, 1056

—. 2015, Monthly Notices Roy. Astron. Soc., 453, 4456

—. 2016, Monthly Notices Roy. Astron. Soc., 460, L59

Visbal, E., Haiman, Z., Terrazas, B., Bryan, G. L., & Barkana, R. 2014b, Monthly Notices Roy. Astron. Soc., 445, 107

Volonteri, M. 2010, A&ARv, 18, 279

Volonteri, M., & Begelman, M. C. 2010, Monthly Notices Roy. Astron. Soc., 409, 1022

Volonteri, M., & Gnedin, N. Y. 2009, Astrophys. J., 703, 2113

Volonteri, M., Haardt, F., Ghisellini, G., & Della Ceca, R. 2011, Monthly Notices Roy. Astron. Soc., 416, 216

Volonteri, M., Haardt, F., & Madau, P. 2003, Astrophys. J., 582, 559

Volonteri, M., Lodato, G., & Natarajan, P. 2008, Monthly Notices Roy. Astron. Soc., 383, 1079

Volonteri, M., & Natarajan, P. 2009, Monthly Notices Roy. Astron. Soc., 400, 1911

Volonteri, M., & Rees, M. J. 2006, Astrophys. J., 650, 669

Volonteri, M., Reines, A., Atek, H., Stark, D. P., & Trebitsch, M. 2017, ArXiv e-prints, arXiv:1704.00753

Volonteri, M., & Stark, D. P. 2011, Monthly Notices Roy. Astron. Soc., 417, 2085

Wagner, A. Y., Bicknell, G. V., & Umemura, M. 2012, Astrophys. J., 757, 136

Wagner, A. Y., Umemura, M., & Bicknell, G. V. 2013, Astrophys. J. Lett., 763, L18

Whalen, D., Abel, T., & Norman, M. L. 2004, Astrophys. J., 610, 14

Whalen, D., & Norman, M. L. 2008, Astrophys. J., 673, 664

Whalen, D. J., & Fryer, C. L. 2012, Astrophys. J. Lett., 756, L19

Whalen, D. J., Fryer, C. L., Holz, D. E., et al. 2013, Astrophys. J. Lett., 762, L6

Willott, C. J., Delorme, P., Reylé, C., et al. 2010, Astron. J., 139, 906

Wise, J. H., Demchenko, V. G., Halicek, M. T., et al. 2014, Monthly Notices Roy. Astron. Soc., 442, 2560

Wolcott-Green, J., & Haiman, Z. 2011, Monthly Notices Roy. Astron. Soc., 412, 2603

Wolcott-Green, J., Haiman, Z., & Bryan, G. L. 2011, Monthly Notices Roy. Astron. Soc., 418, 838

Woosley, S. E. 2016, Astrophys. J. Lett., 824, L10

Wright, E. L. 2012, Online Tutorial by Edward Wright on Big Bang Nucleosynthesis: http://www.astro.ucla.edu/ wright/BBNS.html

Wu, X.-B., Wang, F., Fan, X., et al. 2015, Nature, 518, 512

Wyithe, J. S. B., & Loeb, A. 2002, Astrophys. J., 581, 886

Xu, H., Norman, M. L., O'Shea, B. W., & Wise, J. H. 2016a, Astrophys. J., 823, 140

Xu, H., Wise, J. H., Norman, M. L., Ahn, K., & O'Shea, B. W. 2016b, Astrophys. J., 833, 84

Yajima, H., & Khochfar, S. 2017, Monthly Notices Roy. Astron. Soc., 467, L51

Yajima, H., Li, Y., Zhu, Q., et al. 2014, Monthly Notices Roy. Astron. Soc., 440, 776

Yoneyama, T. 1972, Publ. of the Astron. Soc. of Japan, 24, 87

Yoo, J., & Miralda-Escudé, J. 2004, Astrophys. J. Lett., 614, L25

Yoon, S.-C., Dierks, A., & Langer, N. 2012, Astron. and Astrophys., 542, A113

Yoshida, N., Abel, T., Hernquist, L., & Sugiyama, N. 2003, Astrophys. J., 592, 645

Yoshida, N., Omukai, K., & Hernquist, L. 2008, Science, 321, 669

Yoshida, N., Omukai, K., Hernquist, L., & Abel, T. 2006, Astrophys. J., 652, 6

Yue, B., Ferrara, A., Salvaterra, R., Xu, Y., & Chen, X. 2014, Monthly Notices Roy. Astron. Soc., 440, 1263

Zabl, J., Norgaard-Nielsen, H. U., Fynbo, J. P. U., et al. 2015, Monthly Notices Roy. Astron. Soc., 451, 2050

Zackrisson, E., Rydberg, C.-E., Schaerer, D., Östlin, G., & Tuli, M. 2011, Astrophys. J., 740, 13

Zheng, Z., Cen, R., Trac, H., & Miralda-Escudé, J. 2010, Astrophys. J., 716, 574

Zubovas, K., & King, A. 2012, Astrophys. J. Lett., 745, L34