



# Early galaxy formation and its large-scale effects

Pratika Dayal <sup>a,\*</sup>, Andrea Ferrara <sup>b</sup>

<sup>a</sup> Kapteyn Astronomical Institute, Rijksuniversiteit Groningen, Landleven 12, Groningen, 9717 AD, The Netherlands

<sup>b</sup> Scuola Normale Superiore, Piazza dei Cavalieri 7, Pisa, 56126, Italy



## ARTICLE INFO

### Article history:

Accepted 1 October 2018

Available online 9 October 2018

Editor: M.P. Kamionkowski

### Keywords:

High-Redshift  
Intergalactic medium  
Galaxy formation  
First stars  
Reionization  
Cosmology theory

## ABSTRACT

Galaxy formation is at the heart of our understanding of cosmic evolution. Although there is a consensus that galaxies emerged from the expanding matter background by gravitational instability of primordial fluctuations, a number of additional physical processes must be understood and implemented in theoretical models before these can be reliably used to interpret observations. In parallel, the astonishing recent progresses made in detecting galaxies that formed only a few hundreds of million years after the Big Bang is pushing the quest for more sophisticated and detailed studies of early structures. In this review, we combine the information gleaned from different theoretical models/studies to build a coherent picture of the Universe in its early stages which includes the physics of galaxy formation along with the impact that early structures had on large-scale processes as cosmic reionization and metal enrichment of the intergalactic medium.

© 2018 Elsevier B.V. All rights reserved.

## Contents

1. Galaxy formation: a brief historical perspective .....	3
2. Cosmic evolution in a nutshell .....	6
3. The cosmic scaffolding: dark matter halos .....	8
3.1. The standard cosmological model .....	8
3.2. The start: linear perturbation theory .....	9
3.3. Nonlinear collapse and hierarchical assembly of dark matter halos .....	11
4. Basic physics of galaxy formation .....	14
4.1. Physical ingredients .....	14
4.1.1. Gas cooling .....	15
4.1.2. Star formation .....	15
4.1.3. Radiation fields .....	16
4.1.4. Supernova explosions .....	16
4.1.5. Mass and energy exchange .....	16
4.2. Theoretical tools .....	17
4.2.1. Semi-analytic models (SAMs) .....	17
4.2.2. Hydrodynamic simulations .....	17
4.2.3. Semi-numerical models .....	19
5. The birth of the first galaxies .....	20
5.1. Cold accretion and gas assembly of the first galaxies .....	20
5.2. The formation of the first stars .....	23
6. Heavy elements: production, transport and ejection .....	27

\* Corresponding author.

E-mail addresses: [p.dayal@rug.nl](mailto:p.dayal@rug.nl) (P. Dayal), [andrea.ferrara@sns.it](mailto:andrea.ferrara@sns.it) (A. Ferrara).

6.1.	Metal production and mixing in galaxies .....	27
6.2.	Galactic winds and the enrichment of the IGM .....	29
6.3.	The link between metal enrichment and reionization .....	32
7.	Ionizing radiation from early galaxies: the Epoch of Reionization.....	33
7.1.	The escape of ionizing photons .....	33
7.2.	Growth and properties of ionized regions.....	35
7.3.	Impact of global ultraviolet background on galaxy formation.....	36
7.4.	The sources of reionization: galaxies, quasars or ... ? .....	39
8.	Emerging galaxy properties .....	42
8.1.	High-z galaxies: confronting observations and theory.....	43
8.2.	The high-z UV luminosity function.....	44
8.3.	Stellar mass density .....	47
8.4.	Cosmic star formation rate density .....	47
8.5.	Dust content .....	48
8.6.	Spin and shapes .....	51
8.7.	Sizes.....	53
9.	Open questions and future outlook .....	54
	Acknowledgments .....	55
	References .....	55

---

## List of acronyms

Acronym	Extended name
AGB	Asymptotic Giant Branch
AGN	Active Galactic Nuclei
ALMA	Atacama Large Millimeter Array
AMR	Adaptive Mesh Refinement
BH	Black Holes
BBN	Big Bang Nucleosynthesis
( $\Lambda$ )CDM	(Lambda) Cold Dark Matter
CGM	Circum-galactic medium
CMB	Cosmic Microwave Background
COBE	Cosmic Background Explorer
DLAs	Damped Lyman Alpha systems
DM	Dark Matter
E-ELT	European Extremely Large Telescope
EoR	Epoch of Reionization
GMCs	Giant molecular Clouds
GRBs	Gamma Ray Bursts
HERA	Hydrogen Epoch of Reionization Array
HMF	Halo Mass Function
HST	Hubble Space Telescope
IGM	Inter-galactic medium
IMF	Initial Mass Function
ISM	Inter-stellar medium
JWST	James Webb Space Telescope
LABs	Lyman Alpha Blobs
LAEs	Lyman Alpha Emitters
LBGs	Lyman Break Galaxies
LCGs	Lyman continuum emitting galaxies
LyC	Lyman continuum
LF	Luminosity function
LLS	Lyman limit systems
Lofar	Low Frequency Array
LW	Lyman Werner
MW	Milky Way

---

Acronym	Extended name
PISN	Pair Instability Supernovae
PopIII stars	Population III (metal-free) stars
QSO	Quasi-stellar object
r.m.s.	Root mean square
SAM	Semi-analytic models
SFRD	Star formation rate density
SKA	Square Kilometre Array
SMBH	Super massive black holes
SMD	Stellar mass density
SN	Supernova
SNII	Type II Supernova
SPH	Smoothed-particle hydrodynamics
UV	Ultra-violet
UVB	Ultra-violet background
UV LF	Ultra-violet luminosity function
VLT	Very Large Telescope
WDM	Warm Dark Matter
WMAP	Wilkinson Microwave Anisotropy Probe

## List of symbols

Symbol	Definition
$a(t)$ or $a(z)$	Scale factor of the Universe at time $t$ or redshift $z$
$c$	Speed of light ( $3 \times 10^{10}$ cm s $^{-1}$ )
$\chi_{HI}$	Neutral hydrogen fraction
$f_{esc}$	Escape fraction of H I ionizing photons from the galaxy
$G$	Gravitational constant ( $6.67 \times 10^{-8}$ cm $^3$ gm $^{-1}$ s $^{-2}$ )
$h$	Planck's constant ( $6.67 \times 10^{-27}$ cm $^2$ gm s $^{-1}$ )
$k_B$	Boltzmann's constant ( $1.38 \times 10^{-16}$ cm $^2$ gm s $^{-2}$ K $^{-1}$ )
$H_0$	Hubble constant ( $100h$ km s $^{-1}$ Mpc $^{-1}$ )
$m_p$	Proton mass ( $1.6726 \times 10^{-24}$ gm)
$m_H$	Hydrogen mass ( $1.6737 \times 10^{-24}$ gm)
$\mu$	Mean molecular weight
$\dot{N}_s$	Production rate of LyC photons
$\Omega_b$	Density parameter for baryons
$\Omega_c$	Density parameter cold dark matter
$\Omega_m$	Density parameter for total matter
$\Omega_\Lambda$	Density parameter for dark energy
$\rho_{crit}$	Critical density of the Universe
$R_{vir}$	Virial radius of halo of mass $M_h$
$\sigma_8$	r.m.s. fluctuations on scales of $8h^{-1}$ Mpc
$\tau$	CMB electron scattering optical depth
$T_{vir}$	Virial temperature of halo of mass $M_h$
$T_\gamma$	CMB temperature
$V_{vir}$	Virial temperature of halo of mass $M_h$
$z_{re}$	Reionization redshift

## 1. Galaxy formation: a brief historical perspective

The modern history of galaxy formation theory began immediately after the second World War. In fact, the mere term “galaxy” came into play only after the pivotal discovery, by Hubble in 1925, that the objects until then known as “nebulae”, since their original discovery by Messier towards the end of the 18th century, were indeed of extra-galactic origin. It was then quickly realized that these systems were germane to understanding our own galaxy, the Milky Way (MW). The years immediately after blossomed with key cosmological discoveries: in 1929 Hubble completed his study to prove cosmic expansion and, soon after, in 1933, Zwicky suggested that the dominant fraction of mass constituting clusters of galaxies was unseen, laying the ground for the concept of “Dark Matter” (DM). As early as 1934, one of the first modelling attempts [1] proposed galaxies to have formed out of primordial gas whose condensation process could be followed by applying viscous hydrodynamic equations. Similar pioneering attempts were made by von Weizsäcker [2] who proposed

that galaxies fragmented out of turbulent, expanding primordial gas, although this suggestion was unaware of the later finding that vorticity modes in the early Universe decay rapidly in the linear regime. In addition, Hoyle [3] had already pointed out that the spin of a proto-galaxy could arise from the tidal field of its neighbouring structures. A few years later, Hoyle [4] produced the first results emphasizing the role of gas radiative cooling, fragmentation and the Jeans length (the scale just stable against gravitational collapse) during the collapse of a primordial cloud. These were used to explain the observed masses of galaxies and the presence of galaxy clusters. The limitation of these early models, that attempted to discuss the formation of galaxies independently of a cosmological framework, however, was that a number of ad-hoc assumptions had to be made.

Soon after, Scrimgeour [5] reiterated the point already made by Gamow [6], while working on his proposal of an initially hot and dense initial phase of the Universe ("Hot Big Bang") leading to primordial nucleosynthesis, that the galaxy formation problem must be tied to the cosmological one. As a result of the initially hot state of the Universe, Gamow also predicted the existence of a background of thermal radiation (now known as the Cosmic Microwave Background; CMB) with a temperature of a few degrees Kelvin. As at that time the steady-state model proposed by Bondi, Gold & Hoyle (for an overview see, e.g. Bondi [7]), in which the continual creation of matter enabled the large-scale structure of the Universe to be independent of time despite its expansion, was still under serious consideration, Scrimgeour applied the previous ideas of galaxy formation in that context. His idea was that the gravitational fields of pre-existing galaxies could produce density concentrations in the intergalactic gas resulting in the formation of new galaxies due to gravitational instability. Other ideas, that were sometimes subsequently rejuvenated, were also circulating. The most seminal of these was the "explosive galaxy formation" scenario put forward by Voronitsov-Velyaminov [8] in which galaxies could be torn apart by powerful "repulsive" forces and broken into sub-units. A more modern version of this statement would use the energy deposition by massive stars and, possibly, massive black holes (via processes collectively referred to as "feedback") to trigger a propagating wave of galaxy formation (see e.g. Ostriker and Cowie [9]).

The situation changed dramatically with the experimental discovery of the CMB by Penzias & Wilson in 1965 [10]. Peebles [11] immediately realized that the Jeans scale after the matter–radiation decoupling (i.e. when the hydrogen became neutral), in a Universe filled with black-body radiation, would be close to the mass of present-day globular clusters ( $\approx 10^6 M_\odot$ ). This remarkable result inextricably tied cosmic evolution to galaxy formation. However, Harrison [12] emphasized a problem related to the growth of structure, namely that if galaxies have to emerge out of amorphous initial conditions via a gravitational or thermal instability process, the rate of growth of the perturbations should be short compared to the cosmic age. However, due to Hubble expansion, the growth rate is only a power-law rather than an exponential as in classic theory. Harrison then incorrectly concluded, partly due to the then-sketchy knowledge of the cosmological parameters and the role of dark matter, that this condition was not met. Thus, he favoured the "primordial structure hypothesis" which presupposes that structural differentiation originates with the Universe. In modern language this is equivalent to requiring a very large amplitude of the primordial fluctuation field, later clearly disfavoured by the measurements of CMB anisotropies. A different prediction was made by Silk [13] who instead suggested a much larger ( $\approx 10^{11} M_\odot$ ) critical mass of the first galaxies. This argument was based on his discovery of the "Silk damping" effect that predicts fluctuations below a critical mass to be optically thick and damped out by radiative diffusion on a short time scale as compared to the Hubble expansion time. This proposal, and the contrasting one by Peebles, set the stage for the two scenarios of galaxy formation that survived until the end of the last century. These are the "top-down" scenario, in which small galaxies are formed by fragmentation of larger units, and the "bottom-up" scenario, where large galaxies are hierarchically assembled from smaller systems.

Enormous theoretical efforts were undertaken in the 1970's to discriminate between these two contrasting paradigms of galaxy formation. The idea that galaxies emerged out of fluctuations of the primordial density field was gradually becoming more accepted (notwithstanding the unknown origin and amplitude of the perturbations) after the seminal works by Harrison [14] and Zeldovich [15] who showed that only a specific, scale-invariant form of the fluctuation power spectrum would be compatible with the development of such fluctuations into proto-galaxies. The next urgent step was then to clarify the details of the non-linear stages of the growth of these fluctuations. Zel'dovich [16] obtained an approximate solution, valid for large scale perturbations and pressure-less matter. In this solution, matter would collapse into a disk-like structure (popularly known as a "Zeldovich pancake") before developing a complex shock structure and eventually fragmenting into lower-mass structures. This large-scale approximation seemed appropriate given the Silk damping of small scale perturbations. Moving in a similar direction, Gunn and Gott [17] laid down the basic formalism describing the nonlinear collapse of a perturbation of arbitrary scale embedded in the expanding cosmic flow. They showed that after an initial expansion phase, the radius of the (spherical) perturbation would "turn around", reverting the motion into a contracting one. These results ultimately clarified the key importance of mass accretion onto the initial seed perturbations to form fully-fledged galaxies. This study was nicely complemented by the work of Press and Schechter [18] who developed a simple, yet powerful, method to compute the mass distribution function of collapsed objects from an initial density field made up of random gaussian fluctuations. This paper resulted in strong support for the bottom-up scenario by successfully postulating that once the condensation process has proceeded through several scales, the mass spectrum of condensations becomes "self-similar" and independent of the initially assumed spectrum. Hence, larger mass objects form from the nonlinear interaction of smaller masses. This "Press–Schechter formalism" has become the underlying paradigm of galaxy formation theory and we will devote a later section (Section 3.3) of this review to highlight its main features. Since this result was derived by analysing the statistical behaviour of a system of purely self-gravitating particles, that in modern terms we would classify as dark matter, the theory fell short of the description of baryonic physics. The latter represents a key ingredient in

predicting dissipative processes, star formation and ultimately the actual appearance of observed galaxies. This was pointed out by Larson [19] who worked out a model representing the collapse of an initially gaseous proto-galaxy and the subsequent transformation of gas into stars. In this study, Larson also pioneered the use of numerical simulations that he applied to the solution of the relevant fluid-dynamic equations describing the collapse.

The time was then ripe for the production of the milestone paper by White and Rees [20], built on arguments proposed by a number of other works including Gott and Thuan [21], Silk [22], Binney [23] and Rees and Ostriker [24]. The basic idea was to combine the Press-Schechter theory of gravitational clustering with gas dynamics following the earlier proposals by, e.g., Larson [19]. Dark matter, whose existence in clusters, such as Coma, was becoming accepted, was finally self-consistently included in galaxy formation theories. In practice, the White and Rees [20] model is a two-staged one: dark matter first condenses into halos via pure gravitational collapse after which baryons collapse into the pre-existing potential wells, dissipating their energy via gas-dynamical processes. As a result, the final galaxy properties are partly determined by the assembly of their parent dark matter halo. Once simple, semi-empirical, prescriptions for star formation are implemented one can simultaneously compute statistical properties, like the co-moving number density of galaxies as a function of halo mass/luminosity/stellar mass and redshift, along with their structural/morphological properties (including the gas/stellar mass/star formation rate and sizes). Although admittedly simplified, the model by White and Rees [20] provided a first prediction of the luminosity function (the number density of galaxies in a given luminosity bin) that was roughly consistent with the scarce data available at that time. This class of two-staged galaxy formation models, originating from the above cited original works, has become known as “semi-analytical models” (SAMs). Due to their simplicity and flexibility, although present-day descendants have acquired a remarkable degree of complexity, SAMs are a standard tool in the galaxy formation field. Their more technical features are briefly discussed in Section 4.2.

At approximately the same time, the idea that galaxies are embedded in a dynamically-dominant dark matter halo received decisive support from the observations led by Vera Rubin, starting with the fundamental study presented in Rubin et al. [25]. Measuring the rotation curves of 10 spiral galaxies using the Balmer-alpha (or  $H\alpha$ ) line at 6562 Å in the galaxy rest-frame Rubin and her collaborators reached the surprising conclusion that all rotation curves were approximately flat out to distances as large as  $r = 50$  kpc rather than showing the expected Keplerian decline. They also noted that the maximal velocity was not correlated with the galaxy luminosity. Rather, it was a measure of the total mass and radius, clearly indicating the need for a dark matter halo to explain the observed rotation curves.

However, one crucial element, necessary for completing a coherent framework for galaxy formation, was still missing. This was a precise knowledge of the primordial density fluctuation power spectrum. Indeed, the powerful two-stage approach by White and Rees [20] was still plagued by this limitation. In their study, these authors assumed a simple power-law dependence of the root mean square (r.m.s.) amplitude of the perturbation on mass as  $\sigma \propto M^{-\alpha}$ , with  $\alpha$  essentially being an unknown parameter. Fortunately, soon after Guth [26] and Sato [27] independently proposed the basic ideas of inflation (for a review see, e.g., [28]). They noted that, despite the (assumed) highly homogeneous state of the Universe immediately after the Big Bang, regions separated by more than 1.8 degrees on the sky today should never have been causally connected (the horizon problem). In addition, the initial value of the Hubble constant must be fine-tuned to an extraordinary accuracy to produce a Universe as flat as the one we see today (flatness problem). According to the original proposal (also known as “old inflation”) de-Sitter inflation occurred as a first-order transition to true vacuum. However, a key flaw in this model was that the Universe would become inhomogeneous by bubble collisions soon after the end of inflation. To overcome the problem, Linde [29] and Albrecht and Steinhardt [30], instead, proposed the concept of slow-roll inflation with a second-order transition to true vacuum. Unfortunately, this scenario also suffers from the fine-tuning problem of the time required in a false vacuum to lead to a sufficient amount of inflation. Linde [31] later considered a variant version of the slow-roll inflation called chaotic inflation, in which the initial conditions of scalar fields are chaotic. Chaotic inflation has the advantage of not requiring an initial thermal equilibrium state. Rather, it can start out close to the Planck density, thereby solving the problem of the initial conditions faced by other inflationary models.

Most importantly, in the context of galaxy formation, inflation offered precise predictions for the origin of the primordial fluctuations and their power spectrum. Bardeen et al. [32] showed that fluctuations in the scalar field driving inflation (the “inflaton”) are created on quantum scales and expanded to large scales, eventually giving rise to an almost scale-free spectrum of gaussian adiabatic density perturbations. This is the Harrison-Zeldovich spectrum where the initial density perturbations can be expressed as  $P(k) \propto k^{n_s}$  with the spectral index  $n_s \approx 1$ . Quantum fluctuations are typically frozen by the accelerating expansion when the scales of fluctuations leave the Hubble radius. Long after inflation ends, the scales cross and enter the Hubble radius again. Indeed, the perturbations imprinted on the Hubble patch during inflation are thought to be the origin of the large-scale structure in the Universe. In fact temperature anisotropies, first observed by the Cosmic Background Explorer (COBE) satellite in 1992, exhibit a nearly scale-invariant spectrum as predicted by the inflationary paradigm – these have been confirmed to a high degree of precision by the subsequent Wilkinson Microwave Anisotropy Probe (WMAP) and *Planck* experiments. As  $\sigma \propto M^{-(3+n_s)/6}$  for the inflationary spectrum, the value of  $\alpha$  in the White and Rees [20] theory could now be uniquely determined.

The final step was to compute the linear evolution of the primordial density field up to the recombination epoch. The function describing the modification of the initial spectrum, due to the differential growth of perturbations on different scales and at different epochs, is called the “transfer function” and was computed for dark matter models made by massive (GeV-TeV) particles, collectively known as Cold Dark Matter (CDM) models, by Peebles [33], Bardeen et al. [34] and later improved by the addition of baryons by Sugiyama [35].

Based on the above theoretical advances and exploiting the early availability of computers, it was possible to envision, for the first time, the possibility of simulating, *ab-initio*, the process of structure formation and evolution through cosmic time. This second class of galaxy formation models is referred to as “cosmological simulations” which is another major avenue of research in the field today. More technical details will be given in a following dedicated section (Section 4.2). Here, it suffices to stress how these theoretical and technological advances have revolutionized our view of how galaxies form. In the first attempt, Navarro and Benz [36] simulated the dynamical evolution of both collisionless dark matter particles and a dissipative baryonic component in a flat universe in order to investigate the formation process of the luminous components of galaxies at the centre of galactic dark halos. They assumed gaussian initial density fluctuations with a power spectrum of the form  $P(k) \propto k^{-1}$  meant to reproduce the slope of the power spectrum in the range of scales relevant for galaxies. Although the available computing power for their smoothed-particle hydrodynamic (SPH scheme; for details see Section 4.2) allowed only the outrageously small number of about 7000 particles (compared to the tens of billions used in modern computations) the emergence of the “cosmic web”, made up of filaments, knots and voids, was evident. At the same time, their study represented a change of paradigm marking a transition from the early idealized, loosely informed spherical models to a physically-motivated realization of cosmic structure. A similar effort was also carried out by Katz and Gunn [37], who introduced key technical improvements, such as TREESPH, in which gravitational forces are computed using a hierarchical tree algorithm. The adaptive properties of TREESPH were ideal to solve collapse-type problems.

The amazing progress made in the past century has paved the road for the current research in cosmic structure and galaxy formation whose most modern incarnation we aim at presenting in this review. Although the general scenario has now been established, many fundamental problems remain open, in particular those concerning the formation of the first galaxies. This area represents the current frontier of our knowledge. As such it will be central for the next decade with the advent of many observational facilities that will allow us to peer into these most remote epochs of the Universe.

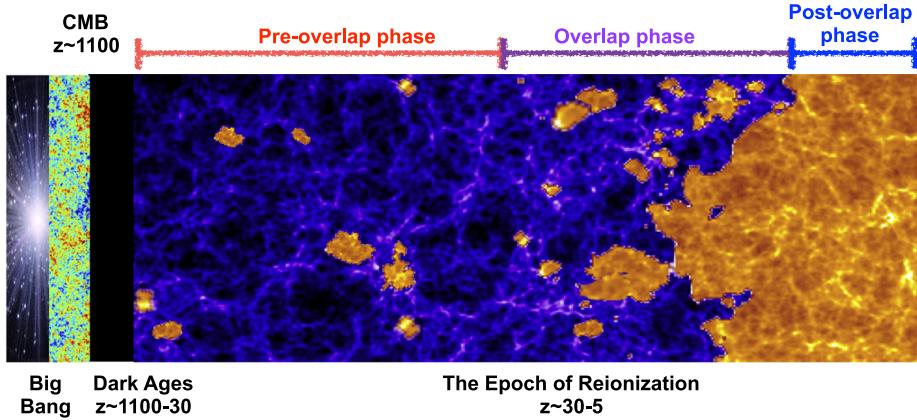
## 2. Cosmic evolution in a nutshell

Our understanding of the Universe has improved remarkably over the last two decades. In addition to strengthening and refining the foundations of the Big Bang standard model, including the Hubble expansion, the CMB and the abundance of light elements, cosmology has given us several genuine, irrefutable surprises. The most prominent among these are that (i) we live in a flat Universe, corroborating predictions of inflationary theories; (ii) roughly 85.3% of cosmic matter is constituted by some kind of, as yet unknown, dark matter particles; (iii) the Hubble expansion is accelerating, possibly due to a non-clustering, negative-pressure fluid called dark energy; (iv) black holes a billion times the mass of the Sun were already in place 1 billion years after the Big Bang; and (v) distant galaxies, and products of their stellar activity, including Gamma-Ray Bursts (GRBs), have recently been detected out to redshifts  $z > 11$ , corresponding to a cosmic age shorter than 0.42 Gyr. Do we have a complete, exact theory to explain these puzzling experimental evidences? Unfortunately, not yet. Gaps remain in the basic cosmological scenario and admittedly some of them are rather large.

It is now widely accepted that the Universe underwent an inflationary phase early on that sourced the nearly scale-invariant primordial density perturbations that have resulted in the large-scale structure we observe today. However, inflation requires non-standard physics and, as of now, there is no consensus on the mechanism that made the Universe inflate; moreover, observations allow only few constraints on the numerous inflationary models available. This inflationary expansion forced matter to cool and pass through a number of symmetry-breaking phase transitions (grand unification theory, electroweak, hadrosynthesis, nucleosynthesis) until recombination when electrons and protons finally found it energetically favourable to combine into hydrogen (or helium) atoms. Due to the large value of the cosmic photon-to-baryon ratio,  $\eta \simeq 10^9$ , this process was delayed until the temperature dropped to 0.29 eV, about 50 times lower than the binding energy of hydrogen. Cosmic recombination proceeded far out of equilibrium because of a “bottleneck” at the  $n = 2$  level of hydrogen since atoms can only reach the ground state via slow processes including the two-photon decay or Lyman-alpha resonance escape. As the recombination rate rapidly became smaller than the expansion rate, the reaction could not reach completion and a small relic abundance of free electrons was left behind at  $z < 1078$ .

Immediately after the recombination epoch, the Universe entered a phase called the *Dark Ages*, where no significant radiation sources existed, as shown in Fig. 1. The hydrogen remained largely neutral at this stage. The small inhomogeneities in the dark matter density field present during the recombination epoch started growing via gravitational instability giving rise to highly nonlinear structures, i.e., collapsed haloes (Section 3). It should, however, be kept in mind that most of the baryons at high redshifts do not reside within these haloes — they are rather found as diffuse gas in the Intergalactic Medium (IGM).

These collapsed haloes formed potential wells, whose depth depends on their mass and, within which baryons could fall. If the mass of the halo was high enough (i.e. the potential well was deep enough), the gas would be able to dissipate its energy, cool via atomic or molecular transitions and fragment within the halo (Section 4). This produced conditions appropriate for the condensation of gas and the formation of stars in galaxies. Once these luminous objects started forming, the Dark Ages were over. The first population of luminous stars and galaxies generated ultraviolet (UV) radiation through nuclear reactions. In addition to galaxies, perhaps an early population of accreting black holes (quasars) and the decay or annihilation of dark matter particles also generated some amount of UV light. The UV photons with energies  $> 13.6$  eV were then able to ionize hydrogen atoms in the surrounding IGM, a process known as “cosmic reionization” (Section 7). Reionization is thus the second major change in the ionization state of hydrogen (and helium) in the Universe (the first being recombination).



**Fig. 1.** A timeline of the first billion years of the Universe. According to our current understanding, immediately after its inception in the Big Bang, the Universe underwent a period of accelerated expansion (“inflation”) after which it cooled adiabatically. At a redshift  $z \sim 1100$ , the temperature dropped to about 0.29 eV at which point matter and radiation decoupled (“decoupling”) giving rise to the CMB and electrons and protons recombined to form hydrogen and helium (“recombination”). This was followed by the cosmic “Dark Ages” when no significant radiation sources existed. These cosmic dark ages ended with the formation of the first stars (at  $z \lesssim 30$ ). These first stars started producing the first photons that could reionize hydrogen into electrons and protons, starting the “Epoch of cosmic Reionization” which had three main stages: the “pre-overlap phase” where each source produced an ionized region around itself, the “overlap phase” when nearby ionized regions started overlapping and the “post-overlap phase” when the IGM was effectively completely ionized. Source: (Reionization simulation credit: Dr. Anne Hutter).

Cosmic reionization, like an Ariadne’s thread, connects many of these gaps. While its evolution is shaped by the matter-energy content and geometry of the Universe, it reflects the way in which galaxies and black holes formed, affects the visibility of distant objects and modifies the properties of the CMB, to mention a few aspects. Disentangling the complicated physical interplay amongst these factors holds the key to filling some of the aforementioned gaps remaining in the standard cosmological model.

According to our current understanding, reionization started around the time when the first structures formed which is currently believed to be around  $z \approx 30$ . In the simplest picture, shown in Fig. 1, each source first produced an ionized region around itself (the “pre-overlap” phase) which overlapped and percolated into the IGM during the “overlap” phase. The process of overlapping seems to have been completed around  $z = 6 - 8$  at which point the neutral hydrogen fraction, expressed as  $\chi_{HI}(z) = n_{HI}(z)/n_H(z)$  where  $n_H$  and  $n_{HI}$  are the densities of hydrogen and neutral hydrogen respectively, fell to values of  $\chi_{HI} < 10^{-4}$ . Following that, a never-ending “post-reionization” (or “post-overlap”) phase started which implies that the Universe is largely ionized at the present epoch. Reionization by UV radiation was also accompanied by heating: electrons released by photo-ionization deposited the photon energy in excess of 13.6 eV into the IGM. This IGM reheating could expel the gas and/or suppress cooling in low mass haloes possibly affecting the cosmic star formation during and after reionization. In addition, the nuclear reactions within the stellar sources potentially altered the chemical composition of the medium if the star exploded as a supernova (Section 6), significantly changing the star formation mode at later stages.

The Epoch of reionization (EoR) is of immense importance in the study of structure formation since, on the one hand, it is a direct consequence of the formation of first structures and luminous sources while, on the other, it affects subsequent structure formation. Observationally, the reionization era represents a phase of the Universe which is only starting to be probed. While the earlier phases are probed by the CMB, the post-reionization phase ( $z < 6$ ) has been mapped out by a number of probes ranging from quasars to galaxies and other sources. In addition to the importance outlined above, the study of Dark Ages and cosmic reionization has acquired increasing significance over the last few years because of the enormous repository of data that is slowly being built-up (Section 8): over the last few years observations have increasingly pushed into the EoR with the number of high-redshift galaxies and quasars having increased dramatically. This has been made possible by a combination of state-of-the-art facilities such as the Hubble Space Telescope (*HST*), Subaru and Very Large telescopes (*VLT*) and refined selection techniques. In the latter category, the Lyman break technique, pioneered by Steidel et al. [38], has been successfully employed to look for Lyman Break Galaxies (LBGs), that are up to three orders of magnitude fainter than the Milky Way, at  $z \simeq 7$  [39,40]. Using the power afforded by lensing, such techniques have now detected viable galaxy candidates at redshifts as high as  $z \simeq 11$ , corresponding to only half a billion years after the Big Bang. Further, the narrow-band Lyman Alpha technique has been successfully used to look for Lyman Alpha Emitters (LAEs) and recently, the broad-band colours are being used to identify galaxies by means of their nebular emission (e.g. [41,42]). Finally, some of the most distant spectroscopically confirmed cosmic objects are GRBs that establish star formation was already well under way at those early epochs, further encouraging deeper galaxy searches. These data sets are already being supplemented by those from facilities including the Atacama Large Millimeter Array (*ALMA*), the Low Frequency Array (*Lofar*) and the Hydrogen Epoch of Reionization Array (*HERA*). In the coming decade, cutting-edge facilities including the James Webb Space Telescope (*JWST*), the Square Kilometre Array (*SKA*) and the European Extremely Large Telescope (*E-ELT*) are expected to further

provide unprecedented glimpses of the early Universe. This observational progress has naturally given rise to a plethora of theoretical models, ranging from analytic calculations to semi-analytic models to numerical simulations, often yielding conflicting results. In the following Sections we will try to summarize these discoveries and progresses in a useful manner. As conventional, we cite all magnitudes in the AB system [43] throughout this review.

### 3. The cosmic scaffolding: dark matter halos

In this Section, we start by discussing the standard Lambda Cold Dark Matter ( $\Lambda$ CDM) cosmological model, which forms the backdrop for galaxy formation, in more detail. We then discuss the collapse of halos in the context of linear perturbation theory before progressing to the non-linear collapse and assembly of the first dark matter halos via merger events and accretion.

#### 3.1. The standard cosmological model

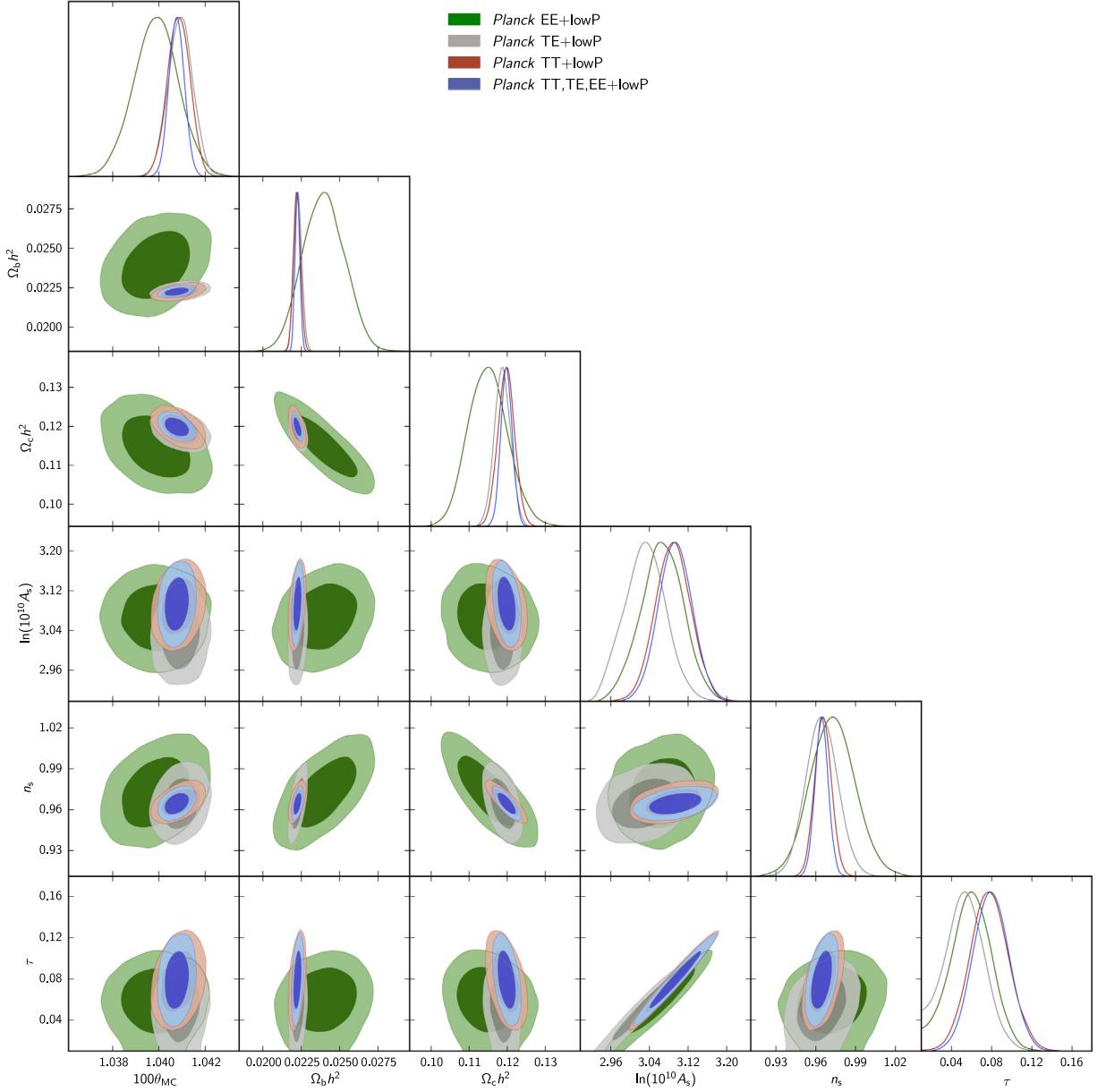
As discussed in Section 2, the Hot Big Bang scenario, the prevailing paradigm describing the Universe from its earliest stages to the present day, envisages it to have started in a hot, dense phase with the energy density largely dominated by radiation down to redshift  $z \simeq 10^4$ . At  $z \simeq 1100$ , for the first time, the temperature of the Universe fell to  $T \simeq 3000$  K, making it energetically favourable for electrons and protons to recombine into neutral hydrogen (H I) atoms. This Epoch of Recombination was imminently followed by the decoupling of the matter and radiation fluids, allowing photons to free-stream to us from the last scattering surface. Such photons are currently observed as the CMB. Indeed, along with the CMB measurements, the homogeneous (e.g. [44]) and isotropic (e.g. [45]) expansion of the Universe and the confirmation of the abundance of light elements predicted by Big Bang Nucleosynthesis (BBN; [46]) form the three pillars of the Hot Big Bang model. This homogeneous, isotropic and expanding Universe is described by the maximally-symmetric Robertson–Walker (RW) metric:

$$ds^2 = c^2 dt^2 - a^2(t) \left\{ \frac{dR^2}{1 - kR^2} + R^2 d\theta^2 + R^2 \sin^2 \theta d\phi^2 \right\}, \quad (1)$$

where  $c$  is the speed of light,  $R$ ,  $r$ ,  $\theta$  and  $\phi$  are co-moving co-ordinates and  $a(t)$  is the cosmic scale factor. Finally,  $k$  represents the curvature of the Universe which is positive, zero and negative for a closed, flat and open Universe, respectively.

Since its accidental discovery, resulting in the award of the 1978 Nobel prize, by Penzias and Wilson in 1965 [10], the CMB has been studied in exquisite detail by a number of observatories including *BOOMERanG* (Balloon Observations of Millimetric Extragalactic Radiation and Geomagnetics), *MAXIMA* (Millimeter Anisotropy eXperiment IMaging Array), *DASI* (Degree Angular Scale Interferometer), *WMAP* and, most recently, the *Planck* satellite. These advances in observational cosmology have led to the establishment of a “concordance”  $\Lambda$ CDM model that specifies the three key components of the Universe: (i) Dark Energy ( $\Lambda$ ); (ii) Cold Dark Matter (CDM) that is now thought to provide the cosmic scaffolding for the entire cosmic web; and (iii) ordinary “baryonic” matter. This cosmological model is fully defined once the 6 basic parameters are derived from CMB data, as shown in Fig. 2. These include:  $\Omega_b h^2$  specifying the baryonic density,  $\Omega_c h^2$  specifying the CDM density,  $100\theta_{MC}$  measuring the sound horizon at last scattering,  $\tau$  measuring the electron scattering optical depth and  $n_s$  and  $\ln(10^{10} A_s)$  specifying the spectral index and amplitude of the initial density perturbations, respectively;  $n_s = 1$  implies the scale-invariant Harrison-Zeldovich spectrum as noted. This model also includes 4 additional derived parameters: the Hubble parameter ( $h$ ), the root mean square mass fluctuations on scales of  $8h^{-1}$  Mpc ( $\sigma_8$ ), the total matter density ( $\Omega_m$ ) and the adimensional density of Dark Energy ( $\Omega_\Lambda$ ). Combining the temperature–temperature (TT) spectra, temperature–E mode polarization (TE) spectra and the E mode–E mode polarization (EE) spectra, the best fit values of these parameters are found to be  $\Omega_b h^2 = 0.02225 \pm 0.00016$ ,  $\Omega_c h^2 = 0.1198 \pm 0.0015$ ,  $100\theta_{MC} = 1.04077 \pm 0.00032$ ,  $\tau = 0.079 \pm 0.017$ ,  $n_s = 0.9645 \pm 0.0049$ ,  $\ln(10^{10} A_s) = 3.094 \pm 0.034$ ,  $h = 0.6727 \pm 0.0066$ ,  $\sigma_8 = 0.831 \pm 0.013$ ,  $\Omega_m = 0.3156 \pm 0.0091$  and  $\Omega_\Lambda = 0.6844 \pm 0.0091$  [47]. Combining temperature data and low multipole polarization data, the latest *Planck* results [48] yield a lower value of  $\tau = 0.055 \pm 0.009$ , which has important implications for reionization and its sources as discussed in Section 7.4. Given its import for reionization, we also show the evolution of  $\tau$ , over 12 years, from *WMAP* data that implied a reionization redshift of  $z_{re} \sim 9$ –14 till the latest *Planck* 2016 results implying a much lower value of  $z_{re} \sim 7.8$ –8.8 in Fig. 3.

The  $\Lambda$ CDM model has been remarkably successful in predicting the large scale structure of the Universe, the temperature anisotropies measured by the CMB and the Lyman Alpha forest statistics (e.g. [50–56]). However, as recently reviewed by Weinberg et al. [57], CDM exhibits a number of small scales problems including: (i) the observed lack of, both, theoretically predicted low-mass (“the missing satellite problem”; [58,59]) and high-mass (“too big to fail problem”; [60,61]) satellites of the Milky Way; (ii) predicting dark matter halos that are too dense (cuspy) as compared to the observationally preferred constant density cores (“the core-cusp problem”; [62–64]), and (iii) facing difficulties in producing typical disks due to ongoing mergers down to  $z \simeq 1$  [65]. The limited success of baryonic feedback in solving these small scale problems (e.g. [61,66]) has prompted questions regarding the nature of dark matter itself. One such alternative candidate is provided by Warm Dark Matter (WDM) with particle masses  $m_\chi \sim \mathcal{O}(\text{keV})$  (e.g. [51,67]). In addition to its particle-physics motivated nature, the WDM model has been lent support by the observations of a 3.5 keV line from the Perseus cluster that might arise from the annihilation of light ( $\sim 7$  keV) sterile neutrinos into photons [68–70]. However, other works [71,72] caution that the

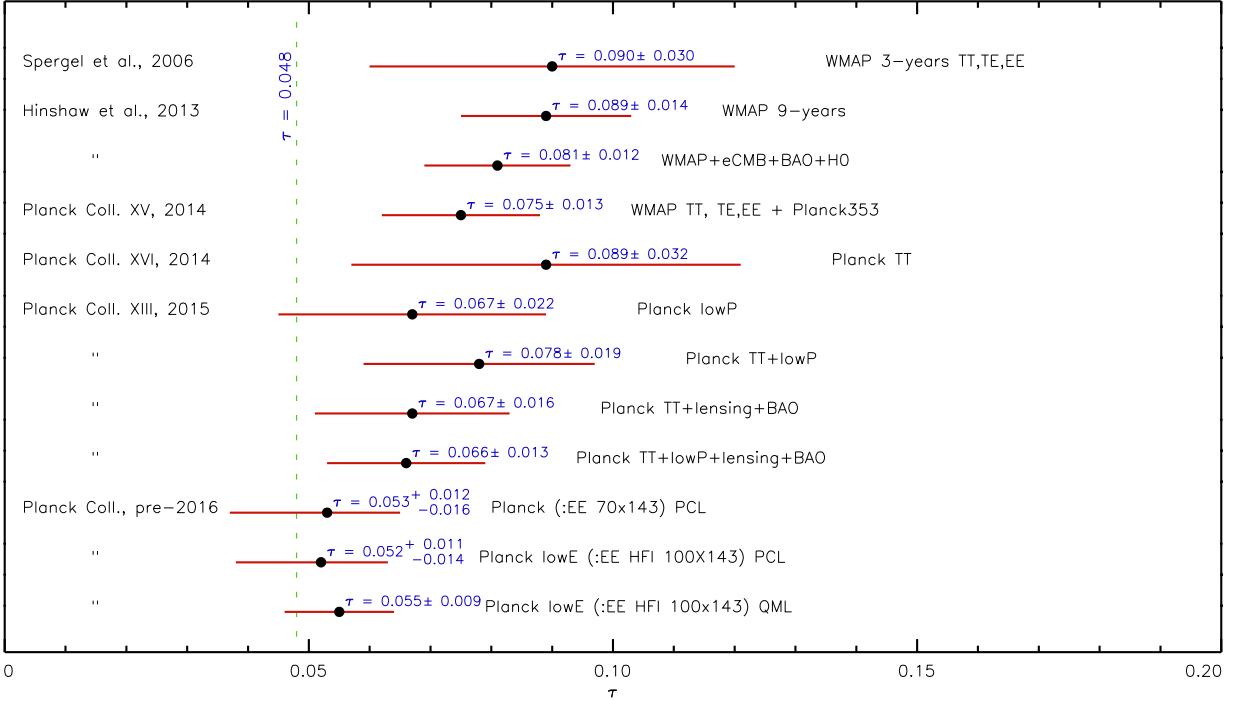


**Fig. 2.** The base  $[\Omega_b h^2, \Omega_c h^2, 100\theta_{MC}, n_s, \tau \text{ and } \ln(10^{10} A_s)]$  and derived  $[h, \sigma_8, \Omega_m \text{ and } \mathcal{Q}_m]$   $\Lambda$ CDM parameter constraints from Planck data [47]. As marked, the green, grey, red and blue contours show parameter values derived using EE and low  $l$  polarization data, TE and low  $l$  polarization data, TT and low  $l$  polarization data and combining TT, TE, EE and low  $l$  polarization data, respectively, with low  $l$  referring to modes with  $l \leq 29$ ; the dark and shaded regions represent the  $1$  and  $2 - \sigma$  contours, respectively. See text in Section 3.1 for details.

power-suppression arising from WDM makes it incompatible with observations, leaving the field open to alternative models including fuzzy CDM consisting of ultra light  $\mathcal{O}(10^{-22} \text{ eV})$  boson or scalar particles [73,74], self-interacting (1 MeV–10 GeV) dark matter [75–77], decaying dark matter [78] and interactive CDM where small-scale perturbations are suppressed due to dark matter and photon/neutrino interactions in the early Universe (e.g. [79,80]). We, however, limit our discussion to the standard  $\Lambda$ CDM model for the purposes of this review.

### 3.2. The start: linear perturbation theory

The growth of small perturbations under the action of gravity was first undertaken by Jeans [81] and Lifshitz [82], with comprehensive derivations in the books by Peebles [83] and Mo et al. [84]. In brief, the aim of linear perturbation theory is to understand the growth of small perturbations of a density contrast  $\delta(\mathbf{x}, t) = \rho(\mathbf{x}, t)/\bar{\rho}(t) \ll 1$  at a given space and time,



**Fig. 3.** The evolution of the electron scattering optical depth,  $\tau$ , over the past 12 years [48], inferred from CMB data from the *WMAP* and *Planck* results as marked. Although not plotted here, the initial *WMAP* 1-year results indicated a very high value of  $\tau = 0.17 \pm 0.04$  resulting in a reionization redshift of  $z_{re} = 20^{+10}_{-9}$  [49]. Since then, the value of  $\tau$  has shown a consistent decrease with the (much lower) final value being  $0.055 \pm 0.009$  yielding a reionization redshift of  $z_{re} \sim 7.8\text{--}8.8$ .

where  $\bar{\rho}(t)$  represents the average background density at time  $t$ . Such growth is described by the linearized form of the hydrodynamical equations (mass and momentum conservation plus the Poisson equation) for a fluid in a gravitational field. These perturbations grow against the background matter distribution represented by an unperturbed medium of constant density and pressure with a null velocity field. Further making reasonable assumptions of homogeneity, isotropy and adiabatic perturbations, these equations can be combined to yield the time-evolution of the density contrast, in comoving co-ordinates, as

$$\frac{d^2\delta(\mathbf{x}, t)}{dt^2} + 2\frac{\dot{a}(t)}{a(t)}\frac{d\delta(\mathbf{x}, t)}{dt} = \frac{c_s^2}{a(t)^2}\nabla^2\delta(\mathbf{x}, t) + 4\pi G\bar{\rho}(t)\delta(\mathbf{x}, t), \quad (2)$$

where  $c_s$  is the fluid sound speed and  $G$  is the Gravitational constant. The physical interpretation of Eq. (2) is straightforward: the gravitationally-driven perturbation growth (second term on the RHS) is opposed by the pressure term (first term on RHS) as well as cosmological expansion (second term on LHS). Writing the density contrast in terms of a Fourier series,  $\delta = \sum \delta_{\mathbf{k}} \exp i\mathbf{k} \cdot \mathbf{x}$ , where  $\mathbf{k}$  is the wave-vector, Eq. (2) can be written as

$$\frac{d^2\delta_{\mathbf{k}}}{dt^2} + 2H(t)\frac{d\delta_{\mathbf{k}}}{dt} = \delta_{\mathbf{k}}\left(4\pi G\bar{\rho}(t) - \frac{k^2 c_s^2}{a(t)^2}\right). \quad (3)$$

The source term (RHS) vanishes at a scale where the pressure gradient balances gravity resulting in the *Jeans wavelength*, the scale that is just stable against collapse, expressed as

$$\lambda_J = \frac{2\pi a(t)}{k_J} = c_s \left( \frac{\pi}{G\bar{\rho}} \right)^{1/2}. \quad (4)$$

The mass enclosed within a radius  $\lambda_J/2$  is referred to as the *Jeans mass* which can be expressed as

$$M_J = \frac{4\pi}{3} \rho \left( \frac{\lambda_J}{2} \right)^3. \quad (5)$$

On the other hand, at  $\lambda \gtrsim \lambda_J$ , the response time for pressure support is longer than the perturbation growth time,

$$t = \frac{1}{(4\pi G\rho)^{1/2}}, \quad (6)$$

and collapse occurs. The power of the Jeans scale lies in the fact that it can be applied to bound structures ranging from galaxy clusters (with physical scale of a few Mpc) to star-forming molecular clouds (on scales of 10–100pc) in the interstellar media (ISM) of galaxies.

These arguments can be extended to a time immediately after matter–radiation decoupling, at  $z \approx 1100$ , in order to obtain the Jeans length and mass. At this time, the sound speed can be approximated assuming a non-relativistic monoatomic gas such that

$$c_s^2 = \left( \frac{5k_B T_\gamma}{3m_p} \right), \quad (7)$$

where  $k_B$  is the Boltzmann constant,  $m_p$  is the proton mass and  $T_\gamma$  is the CMB temperature. This results in a co-moving Jeans length [84]:

$$\lambda_J \approx 0.01(\Omega_b h^2)^{-1/2} \text{ Mpc}, \quad (8)$$

where  $\Omega_b$  is the baryonic density parameter at  $z = 0$ , yielding a constant Jeans mass

$$M_J \approx 1.5 \times 10^5 (\Omega_b h^2)^{-1/2} M_\odot, \quad (9)$$

a value comparable to that of present-day globular clusters. However, we caution that the Jeans mass is not qualitatively accurate in the era of the first galaxies – for these the relevant mass scale is instead provided by the “filtering mass” discussed in Section 7.3.

Inflation predicts that primordial perturbations have a scale invariant power spectrum  $P(k) \propto k^{n_s}$  where the spectral index has been measured to have a value  $n_s = 0.9645 \pm 0.0049$  [47]. The gravitational growth of perturbations modifies the primordial spectrum by depressing its amplitude on scales smaller than the horizon at the matter–radiation equality. The result is that, on small scales,  $P(k) \propto k^{n_s - 4}$ , while the largest scales retain the original quasi-linear spectrum  $\propto k^{n_s}$  (e.g. [34]). Given that most of the power in the standard CDM model is concentrated on small scales, these are the first to go nonlinear, resulting in the formation of bound halos as now explained.

### 3.3. Nonlinear collapse and hierarchical assembly of dark matter halos

Once density perturbations grow beyond the linear regime, i.e.  $\delta(\mathbf{x}, t) \sim 1$ , the full non-linear collapse must be followed. The dynamical collapse of a dark matter halo can be solved exactly in case of specific symmetries, the simplest of which is the collapse of a “top-hat” spherically symmetric density perturbation. The radius of a mass shell in a spherically symmetric perturbation evolves as [84]

$$\frac{d^2r}{dt^2} = -\frac{GM}{r^2} + \frac{\Lambda}{3}r, \quad (10)$$

with the non-zero cosmological constant contributing to the gravitational acceleration. Integrating this equation yields

$$\frac{1}{2} \left( \frac{dr}{dt} \right)^2 - \frac{GM}{r} - \frac{\Lambda c^2}{6} r^2 = \varepsilon, \quad (11)$$

where  $\varepsilon$  is the (constant) specific energy of the mass shell. Let us assume that at its maximum expansion, before the perturbation detaches from the expanding Hubble flow (“turn-around”) and starts to collapse, the shell has a radius  $r_{max}$  at time  $t_{max}$  where

$$t_{max} = \frac{1}{H_0} \left( \frac{\zeta}{\Omega_\Lambda} \right)^{1/2} \int_0^1 dx \left[ \frac{1}{x} - 1 + \zeta(x^2 - 1) \right], \quad (12)$$

where  $\zeta = (\Lambda c^2 r_{max}^3 / 6GM)$ . Further the initial radius  $r_i$  can be linked to  $r_{max}$  as

$$\frac{r_i}{r_{max}} \approx \left( \frac{w_i}{\zeta} \right)^{1/3} \left[ 1 - \frac{1}{5}(1 + \zeta) \left( \frac{w_i}{\zeta} \right)^{1/3} \right], \quad (13)$$

where  $w_i = \Omega_\Lambda(t_i)/\Omega_m(t_i) = (\Omega_\Lambda/\Omega_m)(1 + z_i)^{-3}$ . Combining this with the fact that  $M = (1 + \delta_i)\Omega_{m,i}\bar{\rho}(t_i)(4\pi r_i^3/3)$  yields

$$\delta_i = \frac{3}{5}(1 + \zeta) \left( \frac{w_i}{\zeta} \right)^{1/3}. \quad (14)$$

This can be linearly evolved till the present time to obtain  $\delta_0$  as

$$\delta_0 = \frac{a_0 g_0}{a_i g_i} \delta_i = \frac{3}{5} g_0 (1 + \zeta) \left( \frac{w_i}{\zeta} \right)^{1/3}, \quad (15)$$

where  $w_i = w_0/(1+z_i)^3$  and  $g$  is the linear growth factor. Assuming that the shell collapses at  $t_{col} = 2t_{max}$ , we can derive the linear overdensity at the collapse time as

$$\delta_c(t_{col}) = \frac{3}{5}g(t_{col})(1+\zeta)\left(\frac{w(t_{col})}{\zeta}\right)^{1/3} \approx 1.686[\Omega_m(t_{col})]^{0.0055}, \quad (16)$$

i.e. the linear over-density at collapse time  $t_{col}$  is weakly dependent on the density parameter with  $\delta_{crit} \simeq 1.686$  for all realistic cosmologies [84–86].

Given that the scale factor evolves as  $a \propto t^{2/3}$  in a matter-dominated Universe, the redshift of maximum expansion of the perturbation,  $z_{max}$ , can be related to the collapse redshift,  $z_c$ , as

$$\frac{1+z_{max}}{1+z_c} \propto \left(\frac{t_{col}}{t_{max}}\right)^{2/3} = 2^{2/3}, \quad (17)$$

implying a very swift collapse after turn-around. In reality, however, particles in the mass shell considered will cross the mass shells inside it. Indeed, by  $t = 2t_{max}$  the mass shells enclosed by the collapsed shell will have crossed each other many times resulting in a dark matter halo in virial equilibrium i.e. where the gravitational energy has been converted into the kinetic energy of the particles involved in the collapse. Subsequently, baryons, being the sub-dominant matter component, almost passively start to fall into these dark matter potential wells.

A halo of mass  $M_h$  collapsing at redshift  $z$  has a physical virial radius ( $R_{vir}$ ), circular velocity ( $V_{vir}$ ) and virial temperature ( $T_{vir}$ ) given by [86,87]

$$R_{vir} = 0.784\left(\frac{M_h}{10^8 h^{-1} M_\odot}\right)^{1/3} \left[\frac{\Omega_m \Delta_c}{\Omega_m^z 18\pi^2}\right] \left(\frac{1+z}{10}\right)^{-1} h^{-1} \text{kpc}. \quad (18)$$

$$V_{vir} = 23.4\left(\frac{M_h}{10^8 h^{-1} M_\odot}\right)^{1/3} \left[\frac{\Omega_m \Delta_c}{\Omega_m^z 18\pi^2}\right]^{1/6} \left(\frac{1+z}{10}\right)^{1/2} \text{km s}^{-1}. \quad (19)$$

$$T_{vir} = 1.98 \times 10^4 \left(\frac{\mu}{0.6}\right) \left(\frac{M_h}{10^8 h^{-1} M_\odot}\right)^{2/3} \left[\frac{\Omega_m \Delta_c}{\Omega_m^z 18\pi^2}\right]^{1/3} \left(\frac{1+z}{10}\right) \text{K}, \quad (20)$$

where [88]

$$\Delta_c = 18\pi^2 + 82(\Omega_m^z - 1) - 39(\Omega_m^z - 1)^2, \quad (21)$$

$$\Omega_m^z = \frac{\Omega_m(1+z)^3}{\Omega_m(1+z)^3 + \Omega_\Lambda}, \quad (22)$$

and  $\mu$  is the mean molecular weight.

Once dark matter collapses to form halos, according to the hierarchical CDM structure formation model, these small-scale bound halos merge through time to form successively larger structures. Further, the properties of dark matter halos, such as the density profile, naturally have a critical impact on the properties of the baryons bound in halos. We refer readers to Taylor [89] for a detailed review on dark matter halos, and only focus on two key aspects – relating to their density profiles and the evolution of their number density through time – in what follows.

Press and Schechter [18] were the first to provide an elegant analytic formalism to track the evolution of the number density (per comoving volume) of dark matter halos, the Halo Mass Function (HMF), through time using their “peak-formalism”. Essentially, this formalism identifies peaks above a certain density threshold after smoothing an initial gaussian random density field with a filter. The probability that the density, at a given spatial position and time, exceeds the critical density ( $\delta_{crit} \simeq 1.686$ ) collapsing into a bound object, is given by

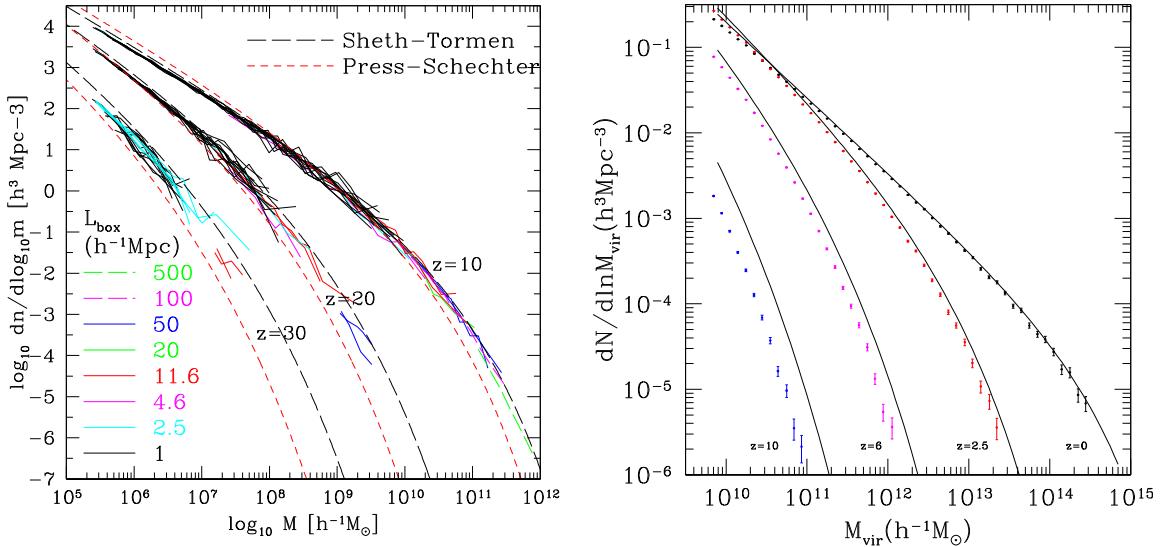
$$p[>\delta_{crit}] = \frac{1}{\sqrt{2\pi}\sigma(M)} \int_{\delta_{crit}}^{\infty} e^{-\frac{\delta_s^2}{2\sigma^2(M)}} d\delta_s = \frac{1}{2} \operatorname{erfc} [\delta_{crit}/\sqrt{2}\sigma(M)], \quad (23)$$

where  $\delta_s$  represents the smoothed density field and  $\sigma(M)$  represents its mass variance, obtained by convolving the initial power spectrum  $P(k)$  with a window function. Press and Schechter soon realized that only half of the (initially over-dense) dark matter would be able to form collapsed objects in this scenario. However, initially under-dense regions can be enclosed inside larger over-dense regions, leading to a finite probability of their resulting in bound objects. They therefore argued that material in initially under-dense regions would be accreted onto collapsed objects, doubling their mass whilst leaving the shape of the mass function unchanged. Including this factor of 2 the Press–Schechter HMF, expressing the number density of halos between mass  $M_h$  and  $M_h + dM_h$  at redshift  $z$ , can be written as

$$n(M_h, z)dM_h = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M_h^2} v e^{-v^2/2} \left| \frac{d \ln v}{d \ln M_h} \right| dM_h, \quad (24)$$

where  $v = \delta_{crit}(z)/\sigma(M)$ . Further,  $\delta_{crit}(z)$  is related to its value at  $z = 0$  such that  $\delta_{crit}(z) = \delta_{crit}(0)D(z)$ , where  $D(z)$  is the linear growth rate normalized to the present day value so that  $D(z=0) = 1$  and  $D(z)$  can be written as [92]:

$$D(z) = \frac{g(z)}{g(0)(1+z)}, \quad (25)$$



**Fig. 4.** A comparison of the Sheth-Tormen HMF with those derived from N-body simulations by Reed et al. [90] (left panel) and Klypin et al. [91] (right panel). As seen, while the Sheth-Tormen HMF is in agreement with simulation results at  $z = 0$  (right panel), it successively over-predicts the number density of halos of all masses with increasing  $z$  in both panels. These results are independent of the halo finder used: while the left panel has used the friends-of-friends halo finder, the right panel has used the Bound-Density-Maxima algorithm.

where

$$g(z) = 2.5\Omega_m[\Omega_m^{4/7} - \Omega_A + (1 + \Omega_m/2)(1 + \Omega_A/70)]^{-1}. \quad (26)$$

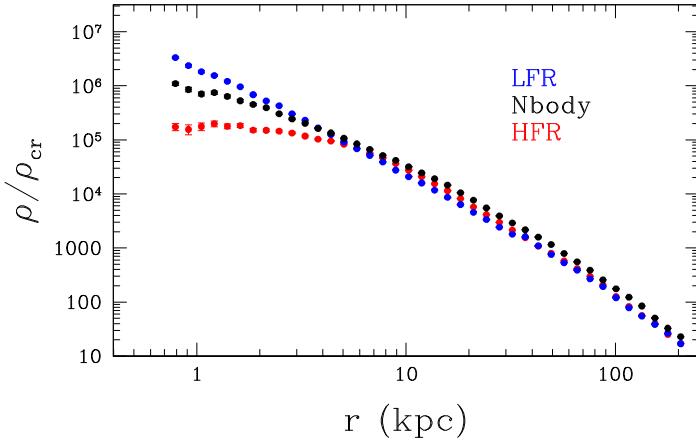
This implies that the Press-Schechter HMF is fully defined once the initial conditions ( $P(k)$ ,  $\sigma(M)$ ) are specified. Eq. (24) shows that the HMF has a characteristic mass,  $M_{h*}(z)$ , below which halos follow a power-law mass-number density relation, with the number density falling off exponentially above this mass. This simple spherical collapse model was extended to the more realistic ellipsoidal collapse scenario by Sheth and Tormen [93,94] resulting in the Sheth-Tormen HMF, whose free parameters are fit by comparing to observations. A number of works have shown that while the Sheth-Tormen HMF faithfully reproduces the number density of low- $z$  systems with mass ranging from dwarfs to clusters, it progressively over-predicts the number density of high-mass halos with increasing  $z$  ([90,91,95] and references therein) as shown in Fig. 4. This inconsistency between the HMFs derived from N-body simulations and the analytical HMF holds irrespective of the (halo finding) technique used to identify particles bound to a given potential well. As shown in Fig. 4, while Reed et al. [90] have used the friends-of-friends halo finder [96], that links all particles within a chosen linking length into a single group, Klypin et al. [91] have used the Bound-Density-Maxima algorithm [97] that locates maxima of density and removes unbound particles to identify bound groups. Motivated by this problem, a number of works (e.g. [91,98]) have suggested correction factors to bring the Sheth-Tormen HMF into better agreement with simulations. For example, Klypin et al. [91] suggest the following correction factor,  $F(z)$ , to ensure better than 10% agreement between analytic and numerical HMFs at all  $z$ :

$$F(z) = \frac{5.5D(z)^4}{1 + [5.5D(z)]^4}. \quad (27)$$

Encoding the abundance of dark matter halos as a function of mass and redshift, the HMF is a powerful probe of cosmology: while the amplitude of the HMF on the scale of galaxy clustering at  $z = 0$  has been used to jointly derive limits on  $\sigma_8$  and the matter density parameter  $\Omega_m$ , the redshift evolution of the amplitude has been used to characterize the dark energy equation of state [99] and even confirm the concordance model of cosmology [100].

As for the dark matter halo density profile, the simplest method of calculating it is to consider a halo to be made up of collapsing concentric shells of different radii and densities. This yields a radial density profile described by a steep power-law with a constant slope (e.g. [101]). The real breakthrough in this field was made when Navarro et al. [63,102] ran high-resolution CDM simulations to show that the internal density profile of dark matter halos is well described by the Navarro, Frenk and White (NFW) profile that is shallower (steeper) than  $r^{-2}$  at small (large) radii such that

$$\rho(r) = \frac{\rho_{\text{crit}}\delta_c}{(r/r_s)(1+r/r_s)^2}, \quad (28)$$



**Fig. 5.** The dark matter density profile (compared to the critical density) of a MW-mass ( $10^{12} M_{\odot}$ ) halo in three different realizations [71]. The black dotted line shows results using an N-body only simulation which predicts an NFW profile. The blue and red dotted lines show results from SPH simulations with low (Kroupa IMF [110] and  $4 \times 10^{50}$  erg of energy per SN) and high (Chabrier IMF [111] and  $10^{51}$  erg of energy per SN) feedback, respectively. As shown, SN feedback in the latter case has a significant impact, resulting in a density profile that is much flatter (and cored in the central 8 kpc) compared to the NFW.

where  $\rho_{crit}$  is the critical density,  $r_s$  represents the halo-dependent scale radius and  $\delta_c$  represents the characteristic over-density defined by

$$\delta_c = \frac{200}{3} \frac{c_h^3}{[\ln(1 + c_h) - c_h/(1 + c_h)]}, \quad (29)$$

where  $c_h = R_{vir}/r_s$  represents the concentration parameter that reflects the critical density of the Universe at the “collapse” redshift.<sup>1</sup> While the NFW profile works extremely well over a wide range of halo masses, ranging from galactic to cluster scales (e.g. [103]), a number of high resolution simulations [104–106] find the profile in the inner-most regions ( $r \lesssim 0.01R_{vir}$ ) to be better fit by the Einasto profile [107]. On the other hand, Burkert [108] have showed that the halo density profiles of dwarfs galaxies are better fit by a Burkert profile of the form

$$\rho(r) = \frac{\rho_0 r_0^3}{(r + r_s)(r^2 + r_s^2)}, \quad (30)$$

where  $\rho_0$  and  $r_s$  are the two free parameters representing the central density and scale radius respectively. Resulting in an isothermal core of fixed density in the inner-most regions, this profile has now been shown to also fit the halo profiles for spirals that are a hundred times more massive [109].

Finally, we note that the concentration parameter depends on the complex assembly history of a given halo: with a longer phase of relatively quiescent growth, massive rare halos are found to be more concentrated with the concentration parameter varying weakly with  $z$ . On the other hand, low mass halos, that have experienced a recent major merger, are found to be less concentrated with  $c_h$  varying rapidly as a function of  $z$  (e.g. [112–114]). A growing body of work has focused on using high-resolution SPH simulations, that simultaneously track the assembly of dark matter halos and the baryonic component, to study the impact of supernova (SN) feedback affecting the dark matter density profile. As in Fig. 5, it has now been shown that the high star formation (and hence SN) efficiency associated with  $M_h \gtrsim 10^{10} M_{\odot}$  halos is powerful enough to perturb the dark matter profile, resulting in a flat, cored, density profile in the central few kpc [71,115–119].

#### 4. Basic physics of galaxy formation

We are now ideally placed to discuss the key physical processes (including gas cooling, star formation, radiation fields and SN feedback) involved in galaxy formation (Section 4.1). We then discuss the key theoretical approaches used to model early galaxies (Section 4.2) before ending by detailing the key physics implemented in a number of (semi-analytic and numerical) models, over a range of physical scales, in Table 1.

##### 4.1. Physical ingredients

Although the physics of gravitational instability, connecting the primordial density fluctuation field to nonlinear dark matter structures, offers a solid starting point for galaxy formation theories, the most difficult challenge is to obtain a sound

<sup>1</sup> While the original NFW paper, and most works following it, denote the concentration parameter by  $c$ , we use the notation  $c_h$  since  $c$  denotes the speed of light throughout this work.

description of the fate of baryons. This is a particularly important point as most of the information obtained from observations is carried by electromagnetic signals from this component. It is therefore critical that any valuable model predicts both the global (such as luminosity functions, star formation rates, stellar and gas masses) and structural (sizes, morphology, accretion, outflows, thermal state of the ISM, turbulence) properties of galaxies. These predictions are far from trivial as they involve a large number of microphysical and hydrodynamical processes interacting on a multi-scale level in a complex feedback network. However, because some of the key processes cannot be described purely from a fundamental physics perspective (a classical example is the way in which star formation is implemented in *any* type of model), one is forced to take a more heuristic approach in which theory must include some *Ansatz* based on empirical evidence. For this reason, the output of any model must be carefully confronted against observations to test whether the assumptions made hold against a wider application range. In the following we try to succinctly describe the key relevant processes for the formation of a galaxy. As it will become clear in the following Sections, though, details of their specific implementation might noticeably change the conclusions reached by different studies. With these caveats, that will be better elucidated later on, we proceed with a first reconnaissance of the key processes involved in galaxy formation.

#### 4.1.1. Gas cooling

Once the baryons have been shock-heated to the virial temperature of their host dark matter halo, further evolution can occur only if the gas is able to radiatively dissipate its internal energy. The associated loss of pressure causes the gas to collapse towards the centre, likely forming a rotationally supported disk-like structure. Cooling processes largely involve collisionally-excited line emission from atoms and continuum radiation (e.g. bremsstrahlung, recombination and collisional ionizations) with heavy elements (including C, N, O and Fe) being the most efficient radiators. However, the first galaxies must have formed in an (almost) primordial environment where the abundances of these elements, as compared to hydrogen (referred to as the “metallicity”), was extremely low. Lacking the key coolants, radiative dissipation had to rely essentially on two species: (a) molecular hydrogen and (b) collisional excitation of the 1s-2p transition of the hydrogen atom i.e. the Lyman-alpha ( $\text{Ly}\alpha$ ) transition. Both processes present some criticalities: while molecular hydrogen, lacking a dipole moment, is an inefficient radiator,  $\text{Ly}\alpha$  requires temperatures  $\gtrsim 10^4$  K for electrons to be sufficiently energetic to excite the transition. Thus, while  $\text{H}_2$  supplies the necessary cooling in relatively small halos with  $T_{vir} \lesssim 10^4$  K,  $\text{Ly}\alpha$  dominates the process for larger halos. Once heavy elements are formed the cooling rates increase, essentially linearly with metallicity ( $Z$ ), and the gas can collapse and form a disk on the cooling timescale,  $t_{cool}$ , that is typically shorter than the free-fall time of the system,  $t_{ff}$ . For a detailed review on the subject see e.g. Ferrara [120].

#### 4.1.2. Star formation

Once gas has cooled into a central, baryon dominated structure, which often resembles a rotating disk, clumps form as a result of the ongoing process of gravitational instability. Large gas agglomerates, known as Giant Molecular Clouds (GMC), form under the action of self-gravity. Their typical masses are, to a first approximation, set by the Jeans length which is equal to the product  $\lambda_J \approx c_s t_{ff}$ , where  $c_s$  is the gas sound speed in the disk, and  $t_{ff}$  the local free-fall time. Turbulent motions and magnetic fields, if present in the early Universe, can contribute extra pressure in addition to the thermal one, playing an important role in supporting the cloud against collapse. These GMCs are highly structured systems, composed of a dense core and filaments, embedded in a more diffuse (interclump) gas component. Arising from the combination of turbulence and gravitational effects, the gas density distribution spans a wide range of values, ranging from  $\approx 10^2 \text{ cm}^{-3}$  in the diffuse, interclump medium to  $> 10^7 \text{ cm}^{-3}$  in the densest cores. As a result, the gas in GMCs, as suggested by the name, is predominantly molecular (including  $\text{H}_2$ , CO and other more complex molecules). How these dense cores turn into stars is a difficult question that has been challenging experts since the 1960's. Although many details are still unclear, a general scenario that emerges is one in which cores, undergoing gravitational collapse, become centrally concentrated, with a density profile that scales approximately as  $\rho(r) \propto r^{-2}$ . In the innermost regions of the collapsing core, the opacity eventually becomes large enough that the gas switches from approximately isothermal to adiabatic behaviour. Once the gas is hot enough to dissociate the molecular hydrogen, a second collapse ensues and a protostar is formed. These complex processes cannot be followed in detail by even the most advanced galaxy formation simulations. Therefore, one usually resorts to empirical prescriptions to determine the fraction of the gas in the core that is turned into stars. This efficiency,  $\epsilon_{core}$ , is relatively high, and observationally inferred to be  $\approx 0.3\text{--}0.5$ . Perhaps more useful in large scale approaches is the analogous overall efficiency per free-fall time for the entire GMC. This value is estimated to be  $\epsilon_{ff} \approx 0.01$  and it is encoded in the empirical Kennicutt–Schmidt relation that links the gas surface density to the star formation rate. Another important quantity to be specified in any galaxy model is the Initial Mass Function (IMF) of the stars, i.e. their mass distribution at birth. Whether the IMF is best described through the Salpeter function as observed in the Galaxy (where the number of stars of a given mass  $N(M_s) \propto M_s^{-2.35}$ ) or follows alternative forms such as the Kroupa [121] or Chabrier [111], as well as its dependence on other physical quantities (e.g. metallicity, environment etc.) remains highly uncertain. Excellent reviews on star formation physics are given in McKee and Ostriker [122] and Krumholz [123].

#### 4.1.3. Radiation fields

Radiation from (massive) stars plays an important role in the thermal and dynamical evolution of galaxies. This is particularly true for the earliest systems, which are characterized by virial temperatures not greatly exceeding those resulting from photo-heating (typically  $\sim 10^4$  K) by UV photons. In addition, radiative emission is fundamental in order to detect and study these systems. The power (i.e. “luminosity”,  $L_s$ ) and the spectrum of the radiation field emitted by a star depends sensitively on its mass ( $M_s$ ). Massive stars tend to be more luminous, approximately following the relation  $L_s \propto M_s^{7/2}$  for most of their lifetime, spent on the so-called main sequence. They are also hotter: assuming that they emit as black bodies, and given the main-sequence mass dependence of the stellar radius,  $R_s \propto \sqrt{M_s}$ , one obtains a temperature that scales as  $T_s \propto M_s^{5/8}$ . Thus the spectrum of massive stars contains a larger fraction of photons with energy  $h\nu > 1$  Ryd = 13.6 eV capable of ionizing hydrogen atoms and, for the extreme case of metal-free (Pop III) stars, also helium. This process creates ionized bubbles through which photons can freely escape from galaxies. The first stars also produce Lyman–Werner (LW) photons (11.2–13.6 eV) establishing a LW background that can impact further star formation (see Section 5.2). Non-ionizing UV photons ( $h\nu < 1$  Ryd) are equally important given that they also heat the gas via photoelectric effect mediated by dust grains; such energy deposition allows the creation of the multiphase ISM structures observed in galaxies. The background created by UV radiation (the UV background; UVB) from stars can also photoevaporate, and ultimately disperse, GMCs hampering further star formation. All these processes are thoroughly discussed in the classical book by Spitzer [124].

#### 4.1.4. Supernova explosions

Massive stars have short lifetimes: as the available energy from nuclear fusion  $\propto M_s$ , the lifetime can be roughly estimated to be  $M_s c^2 / L_s \propto M_s^{-5/2}$ . Thus a massive star with  $M_s \gtrsim 8M_\odot$  only lives for  $\lesssim 30$  Myr, a short timescale compared to the typical dynamical time of a galaxy which is approximately equal to the rotational time  $t_{dyn} = 2\pi R/V$ , where  $R$  and  $V$  are the disk radius and circular velocity, respectively. As an example, the dynamical timescale for the Milky Way is estimated to be  $t_{dyn} \approx 300$  Myr. The core collapse, marking the death of massive stars, ignites a powerful explosion (“supernova”) followed by shock waves propagating into the surrounding gas, and carving hot ( $T \approx 10^6$  K) bubbles. The explosion leaves behind a remnant which could be a neutron star or a black hole (BH).<sup>2</sup> Approximately 99% of the total supernova energy (typically corresponding to 1 Bethe =  $10^{51}$  erg almost independent of  $M_s$ ) is carried away by neutrinos. The ejecta initially expand freely at velocities exceeding 1000 km s<sup>-1</sup>, but once the inertia of the swept up gas becomes important, the expansion enters an adiabatic Sedov–Taylor phase in which the radius of the bubble increases with time only as  $t^{3/5}$  rather than linearly. Finally, radiative energy losses in the post-shock gas become important, causing a drop in the internal energy, resulting in an additional deceleration, until the cold shell simply coasts along conserving momentum. The role of SN energy injection can be hardly overlooked as it might have dramatic effects on the evolution of a galaxy. The production of hot gas quenches star formation (a process often referred to as “mechanical feedback”), shocks inject bulk motions, vorticity and turbulence in the ISM and explosions disperse heavy elements and dust enhancing the cooling ability of the gas. Finally, the hot gas is buoyant and tends to leave the galaxy in the form of a galactic wind (see below). Modelling these processes is extremely difficult and no clear consensus has been reached on how to self-consistently include them in theoretical models and in simulations that cannot properly resolve the relevant physical scales. The interested reader can find a detailed theoretical description of SN explosions in Ostriker and McKee [125].

#### 4.1.5. Mass and energy exchange

Galaxies can hardly be modelled as closed systems. Cosmological simulations show that galaxies form at the intersection of the filaments constituting a complex network, known as the “cosmic web”. Matter flows into galaxies along these “pipes” feeding the system with dark matter and baryons (presumably in a ratio close to the cosmological  $\Omega_c/\Omega_b \approx 5.8$ ). The accreted intergalactic gas represents new fuel for star formation. At the same time, given its more pristine composition, intergalactic gas acts to dilute the heavy element content of the interstellar medium. The cycle is closed by the fact that supernovae associated with newly born stars might eject a sizeable fraction of their gas (and metal) mass into the IGM again (for a review see [126]). Note that, under some conditions, part of the accelerated wind gas might rain back onto the galaxy in a so-called “galactic fountain”. It is also possible that the above dynamical process reaches a sort of steady state in which the rate of gas accretion balances the sum of star formation and outflow rate (this model is informally known as the “bath-tub” model). Accretion is not the only way in which galaxies gain their mass. An alternative mechanism relies on galaxy mergers occurring after close encounters. This is probably the dominant mode of growth at early cosmic times as the merger rate for typical galaxies is a steep function of redshift, scaling as  $\propto (1+z)^\gamma$  with  $\gamma = 2.2\text{--}2.5$ . Mergers in hierarchical formation models are particularly important as, in addition to triggering bursts of star formation driven by the increased density of the compressed gas, they could drive dramatic changes in galaxy morphology and dynamics. For example, elliptical galaxies are thought to result from the merger of two disk galaxies. These issues are reviewed in detail by Conselice [127].

<sup>2</sup> In the particular case of pair instability supernovae, occurring in the narrow range  $140 < M_s/M_\odot < 260$  there is no compact remnant and the star is totally disrupted.

## 4.2. Theoretical tools

The vast range of scales involved in modelling galaxies in a cosmological context, ranging from the sub-pc scales of star formation and its associated feedback to the super-Mpc scales associated with reionization, have spawned a variety of theoretical techniques. Here, we briefly discuss the three main “classes” of theoretical (semi-analytical, semi-numerical and hydrodynamic) models adopted for modelling early galaxies and interested readers are referred to the review by Somerville and Davé [128] for more details on the modelling techniques discussed here. The key aims, parameters (volumes modelled and dark matter particle resolution mass) and physics implemented in a number of theoretical models, ranging from small to large scales, is summarized in Table 1. A fourth, empirical method based on *abundance-matching* attempts to map the properties of the underlying dark matter halos onto observable galaxy properties (e.g. [129,130]). However, given that such models do not explicitly implement physical processes from first principles, we do not discuss them further in this Section.

We start with the assembly of dark matter halos that provide the backbone for galaxy formation. Dark matter merger trees, describing the (merger- and accretion-driven) growth of halos can be constructed in two ways. The first approach, colloquially refereed to as building “binary merger trees” uses the extended Press–Schechter formalism to successively fragment a halo into its (at most two) higher- $z$  progenitors; progenitors below the chosen resolution mass limit are then regarded as “smooth-accretion” of dark matter from the IGM. Alternatively, dark matter only (“N-body simulations”) can be run where the motion and collapse of dark matter particles, initially distributed according to the cosmological initial conditions, are followed by calculating the forces from all other particles using Poisson’s equations. These calculations can either be carried out using particle-based “tree” methods (where forces from families of distant particles are computed via multiple moments), using particle mesh codes (particles move along potential gradients computed via Fourier transform of the density field) or a combination thereof.

Once these codes are run, a key challenge is to identify a bound “halo” and the constituent sub-structures (sub-halos) arising from distinct halos that formed and merged into the central halo at earlier times. Simulations rely on three main categories of halo finding techniques (for a detailed comparison see [131]): (i) *Friends-of-friends (FoFs) method*: that identifies bound groups by linking together all particles separated by a chosen “linking length” which is usually a fraction of the average particle separation in the simulation; (ii) *Spherical over-density (SO) method*: that identifies peaks in the matter density field around which spherical shells are grown out to a point where the density falls below a certain pre-defined value; and (iii) *6D phase-space methods*: an extension of the FoF method, these identify bound halos by including an additional proximity condition in velocity space. We are now in a position to discuss how these dark matter halos can be combined with baryonic calculations to yield galaxy populations.

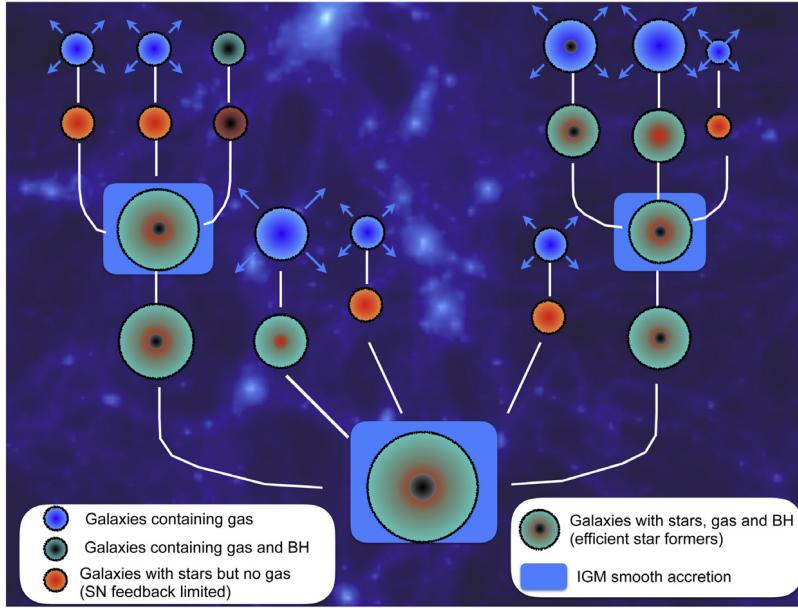
### 4.2.1. Semi-analytic models (SAMs)

SAMs are based on using sets of equations to implement, some or all of, the physical processes discussed in Section 4.1 into dark matter merger trees, jointly tracking the assembly of baryons and their host dark matter halos through cosmic time – a typical SAM is shown in Fig. 6. Depending on the aim, the key processes (gas accretion, star formation, chemical/mechanical/radiative feedback and metal enrichment) can be included with varying degrees of complexity: e.g. gas and metals can be distributed following the dark matter density profile, gas can be split into cold/hot phases, outflows can preferentially carry away metal enriched gas and reionization feedback can be implemented, to name a few. As a result of their very nature of treating baryons through equations, lacking “sub-grid” resolution, these models generally yield average baryonic properties over a given halo. That being said, the appeal of SAM lies in their (relative) simplicity of implementation, the flexibility of incorporating physical processes and the large physical/mass dynamic ranges and physical parameter space that can be explored over reasonably short timescales with relatively low computing requirements.

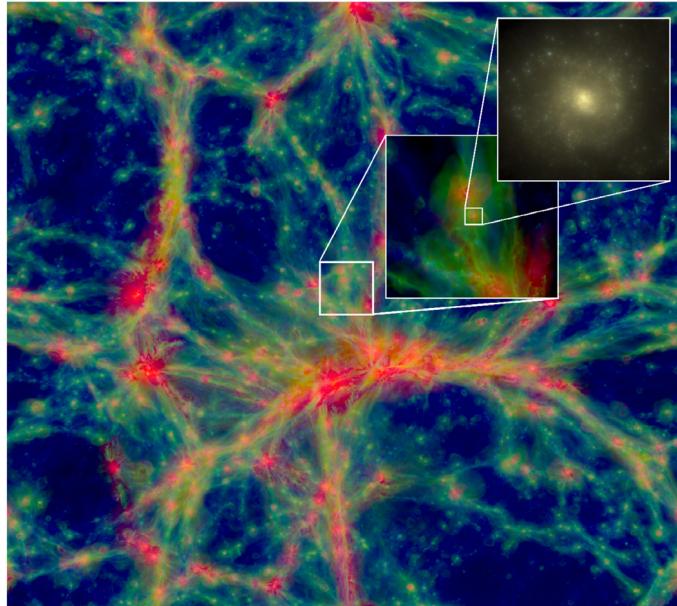
### 4.2.2. Hydrodynamic simulations

The most complex approach to galaxy formation involves hydrodynamic simulations. In this approach, one simultaneously solves the relevant equations of gravity, hydrodynamics and thermodynamics over particles or grid cells that represent dark matter, gas and stars. This is an extremely powerful, albeit computationally expensive, method that can track and predict the interplay between dark matter and baryons and the resulting galaxy/large scale structure properties (depending on the simulation size) once appropriate sub-grid (scales below the resolution limit) recipes for gas cooling, star formation, and feedback are implemented.

Hydrodynamic simulations can be divided into two main categories: (i) *Lagrangian or SPH*: where particles carry the reference frame, and particle properties – including mass, temperature, metallicity – are “smoothed” by weighing over all neighbours closer than a chosen separation (called the linking length). An example of a typical SPH simulation is shown in Fig. 7; and (ii) *Eulerian*: these are based on a cartesian reference frame where fluid particles are discretized into cells and physical properties are calculated across fixed cell boundaries. Given the dynamic range involved in galaxy formation, the latter method has been extended to Adaptive Mesh Refinement (AMR) techniques where cells satisfying a given criterion – for example in density or temperature – are split into sub-cells leading to a higher resolution in such regions. Issues with these techniques (see e.g. [132]), such as the inability of SPH to resolve shocks and the sensitivity of AMR results to bulk velocities, have led to the development of a new class of “unstructured” mesh models, lying between Lagrangian and Eulerian methods, that use Voronoi tessellation to sub-divide the space around particles. Based on polyhedral cells, this



**Fig. 6.** A SAM jointly tracking the (merger and accretion driven) buildup of dark matter halos as well as their baryonic component. SAMs present a powerful tool for capturing the key physics involved in galaxy formation including: the stellar mass growth due to star formation, the merger and accretion driven gas mass growth, the role of SN (and possibly black holes) in ejecting gas from low-mass halos and tracking the resulting impact on the subsequent growth of more massive systems via halo mergers and gas accretion. As shown, galaxies assemble as a result of both wet mergers of “efficient star formers” that do not lose all their gas mass after star formation/black hole accretion and dry mergers of supernova/black hole “feedback limited” systems that lose all their gas resulting in dry mergers. This naturally leads to a variety of galaxy assembly histories and properties for a given halo mass.



**Fig. 7.** A  $100 \times 100 \times 20$  comoving Mpc (cMpc) slice through the EAGLE simulation, illustrating the dynamic range attainable with state-of-the-art SPH simulations. The intensity represents the gas density while the colour indicates the gas temperatures (blue-green-red from cooler to hotter). The inset shows a region 10 cMpc and 60 ckpc on a side. The zoom in to an individual galaxy with stellar mass  $3 \times 10^{10} M_\odot$  shows the optical band stellar light [133].

mesh continually de-forms and re-forms as particles move. Despite its obvious advantage in simultaneously tracking both dark matter and gas particles, the enormous computational effort required for hydrodynamic simulations naturally places a limit on the physical volume that can be simulated and the physical parameter space that can be explored for a given mass resolution.

**Table 1**

An illustrative list of models for the first billion years. For each model we list the key aim (column 2), the simulation technique (column 3), box size (column 4), DM mass resolution (column 5), the key physics implemented with the footnote explaining the letters used (column 6) and the model name and reference (column 7). The explanations of the letters used in columns 1 and 6 follow on the next page.

No.	Main aim	Technique	box size [cMpc]	$M_{DM}$ [ $M_\odot$ ]	Key Physics	Code [reference]
Small-scale models						
1	SF in GMC	Resimulated	$1-10 R_{vir}$	$\sim 10^2-10^6$	AIKO	FIRE [136]
2	SF, GF, EoR	AMR	1	1840	DIKO	[137]
3	GF, EoR	AMR+RT	$4 h^{-1}$	$4 \times 10^6$	EO	EMMA [138]
4	UV fb	SPH+RT	5	$2.5 \times 10^5$	EIO	[139]
5	UV fb, GF	SPH+RT	$3-6 h^{-1}$	$0.18-1.4 \times 10^6$	GIKO	[140]
6	ISM,CGM	AMR	$9.7 h^{-1}$ kpc	$9.5 \times 10^4$	AIKL	[141]
7	UV fb, GF	SPH+RT	$10 h^{-1}$	$4.3 \times 10^7$	GIO	[142]
8	GF, EoR	Eulerian+RT	20	$4.8 \times 10^5$	EIKO	[143]
9	GF, EoR	AMR	40	$3 \times 10^4$	diko	Renaissance [144,145]
10	GF, EoR	AMR+RT	$20-40 h^{-1}$	$7 \times 10^6$	AIO	CROC [146]
Intermediate-scale models						
1	GF, EoR	N-body+semi-numerical RT	100	$3.9 \times 10^6$	DIKO	DRAGONS [147]
2	GF	SAM	–	$M_{min} = 10^8$	CIP	DELPHI [148]
3	EoR (LG)	Eulerian+RT	91	$3.5 \times 10^5$	EIO	CoDA [149]
4	GF, EoR	SPH+ RT	$12.5-100 h^{-1}$	$10^6-8 \times 10^7$	EIKO	Aurora [150]
5	$f_{esc}$	SPH	$10-100 h^{-1}$	$6 \times 10^6-9 \times 10^8$	GIKLM	[151]
6	GF	SPH	$25-100 h^{-1}$	$1.2-9.7 \times 10^6$	FIJKM	EAGLE [133]
7	GF	Unstructured mesh	106	$6.2 \times 10^6$	GIJK	Illustrus [152]
Large-scale models						
1	EoR	N-body+RT	$114-425 h^{-1}$	$0.55-5 \times 10^7$	HO	[134]
2	GF	SPH	$400 h^{-1}$	$1.7 \times 10^7$	GIJKM	BlueTides [153]
3	GF	SAM	$500 h^{-1}$	$1.3 \times 10^9$	BIJKP	GALFORM [154]

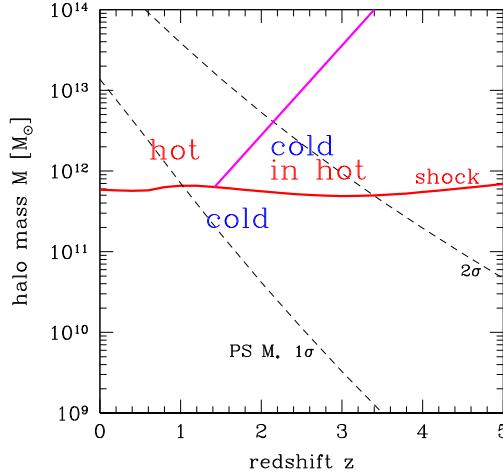
#### 4.2.3. Semi-numerical models

As detailed in Section 7.3, the past years have seen an increasing realization of the necessity of coupling galaxy formation – on kiloparsecs scales – with the impact of the radiative feedback generated during reionization – on tens of Megaparsec scales. Indeed, Iliev et al. [134] have shown that, while  $(100h^{-1}\text{cMpc})^3$  boxes are sufficiently large for deriving convergent reionization histories, the morphology of the ionized bubbles remains poorly described for box sizes smaller than  $(200h^{-1}\text{cMpc})^3$ . The rise of 21 cm cosmology, and associated statistics including the r.m.s, skewness and kurtosis of the differential 21 cm brightness temperature (the temperature of the redshifted 21 cm emission with respect to the CMB), have therefore led to the development of a new class of “semi-numerical” models (for a review see [135]). The key idea is to couple semi-analytic models of galaxy formation, run on large ( $\gtrsim 100\text{cMpc}$ )<sup>3</sup> N-body simulations, with analytic/numerical models of radiative transfer. This is the only computationally tractable approach of consistently describing the complexity of the galaxy formation-reionization interplay as will be discussed in more detail in Section 7.

The explanations of the letters used in column 6 of Table 1.

- A: SFR  $\propto \rho_{\text{H}_2}$
- B: SFR  $\propto \text{mass}_{\text{H}_2}$
- C: SFR  $\propto \text{mass}_{\text{gas}}$
- D: SFR  $\propto \text{mass}_{\text{cold gas}}$
- E: SFR  $\propto \rho_{\text{gas}}$
- F: SFR  $\propto \text{Pressure}_{\text{gas}}$
- G: Multi-phase SF using Springel and Hernquist [155] prescription
- H: SFR  $\propto M_h$
- I: SN feedback
- J: SMBH feedback
- K: Metal enrichment
- L: Dust enrichment
- M: Feedback from homogeneous UVB (e.g. [156,157])
- N: Feedback from self-consistently calculated homogeneous UV fields
- O: Feedback from self-consistently calculated inhomogeneous UVB
- P: SFR = 0 (for  $M_h < M_{cr}$  or  $V_{vir} < V_{cr}$  halos at  $z < z_{re}$ )

Further, in Table 1, GF stands for galaxy formation, UV fb stands for UV feedback and LG stands for Local group.



**Fig. 8.** Cold streams and shock-heated medium as a function of halo mass and redshift [170]. The (nearly) horizontal red curve shows the typical mass above (below) which gas flows into galaxies are predominantly hot (cold). The dashed lines show the halo mass corresponding to 1 and 2- $\sigma$  fluctuations as a function of redshift. The inclined solid curve shows the upper mass limit (roughly corresponding to the rare, massive 2- $\sigma$  fluctuations at  $z \gtrsim 2$ ) where the hot interstellar medium can host cold streams at  $z \approx 1$ -2 allowing disk growth and star formation. On the other hand, similar mass halos at lower  $z$  are all dominated by hot-mode accretion that shuts-off gas supply and star formation.

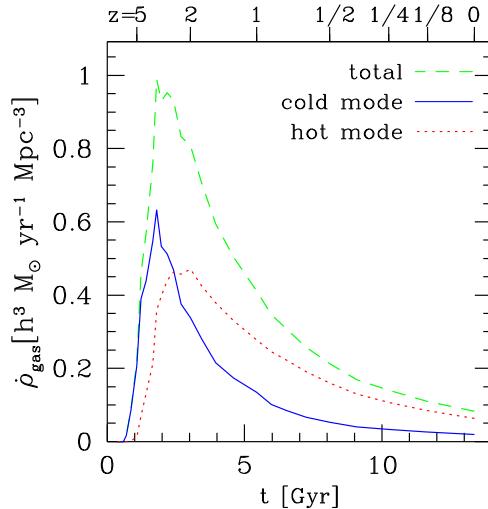
## 5. The birth of the first galaxies

We are now in a position to discuss how early galaxies built-up their gas mass and formed the first stars. As noted in Section 4, the gas in the first galaxies would have been metal-free, being composed of a primordial mixture of ( $\sim 75\%$ ) hydrogen and ( $\sim 25\%$ ) helium. Continuing the classification where solar metallicity stars are termed Population I, the stars forming out of such metal-free gas are referred to as *Population III* stars (PopIII; e.g. [158]): this classification has been refined to include PopIII.1 (PopIII.2; formerly known as PopII.5) stars that formed out of primordial gas (that was photo-ionized) prior to its gravitational collapse into halos (e.g. [159]). While we limit ourselves to an overview of the first stars, we refer interested readers to excellent reviews (e.g. [160–164]) for more details.

### 5.1. Cold accretion and gas assembly of the first galaxies

Once a dark matter halo forms a potential well, it accretes gas from the intergalactic medium to form its interstellar medium and circum-galactic medium (CGM). The standard paradigm of galaxy formation (e.g. [22,24,165]) sketches a picture of “hot-mode accretion” where gas falling into a dark matter potential well, with circular velocity  $V_{vir}$ , is shock heated to the halo virial temperature  $T_{vir}(z) = 10^6(V_{vir}(z)/167 \text{ km s}^{-1})^2 \text{ K}$ . The gas inside a “cooling radius”, within which the cooling time is shorter than the free-fall time, can cool radiatively and settle into a centrifugally supported disk where star formation can take place. Mergers can scatter stars off their orbits, producing spheroidal systems that might re-assemble into a disk after subsequent gas accretion. However, as early as 1977 it was pointed out that, for plausible physical conditions, only a small fraction ( $\approx 10\%$ ) of gas would be heated to  $T_{vir}$ , with most never heated to temperatures above  $2-3 \times 10^4 \text{ K}$  (e.g. [23,37,166]). This has evolved into the current paradigm of “cold-mode accretion” that suggests most gas is accreted at temperatures much lower than  $T_{vir}$ . Indeed, given that cooling is extremely efficient at  $T \approx 1-2.5 \times 10^5 \text{ K}$ , this provides a natural temperature threshold separating hot and cold-mode accretion [167–169].

A number of works (e.g. [170,171]) have shown that pressure can build-up, allowing a stable virial shock to develop, when the compression rate, required to restore pressure, is larger than the cooling rate of the post-shocked gas. However, there are two regimes in which shock heating of gas can be avoided: firstly, low mass galaxies are unable to support stable shocks at any redshift. Secondly, even in the presence of shocks, cold gas can be accreted along filaments that penetrate deep into the halo, feeding the galaxy with gas brought in from much larger distances than in the standard picture of hot-mode accretion [168–174]. This naturally implies a critical *shock-heating* mass,  $M_{shock}$ , below which galaxies pre-dominantly assemble through filamentary cold flows, with quasi-spherical hot-accretion dominating above this scale. As shown in Fig. 8, this critical mass has a value of  $M_{shock} \approx 10^{11.5-12} M_\odot$ , essentially independent of redshift, and represents an equipartition scale where halos accrete their gas equally via hot and cold mode accretion [168,169].  $M_h \lesssim M_{shock}$  halos effectively accrete via the cold-mode at all  $z$ . However, the situation is more complicated for larger mass halos: while cold streams, penetrating through the hot ISM, can contribute to the gas mass of the rare and massive  $M_h \gtrsim M_{shock}$  galaxies at high- $z$  ( $z \gtrsim 2$ ), at lower redshifts such halos are purely dominated by hot-mode accretion [170,171].



**Fig. 9.** Redshift history of total smooth gas accretion (dashed green line) with the deconstructed contributions from cold- (solid blue line) and hot-mode (red dotted line) accretion [168]. These lines show results from all galaxies in a  $22.2 h^{-1}$  cMpc box with  $128^3$  dark matter particles and initially the same number of gas particles. As shown, given the HMF shifting to lower masses with increasing redshift, cold-mode accretion provides the dominant mechanism for building the gas content of galaxies at  $z \gtrsim 2$ , shifting to hot-mode accretion at lower- $z$ .

Simulations show that the rate of cold-gas accretion, which is both a function of halo mass and redshift, can be approximated as [175]

$$\dot{M}_b \simeq 7.36 \left( \frac{M_h}{10^{12} M_\odot} \right)^{1.15} (1+z)^{2.25} f_{0.184} \quad [M_\odot \text{yr}^{-1}], \quad (31)$$

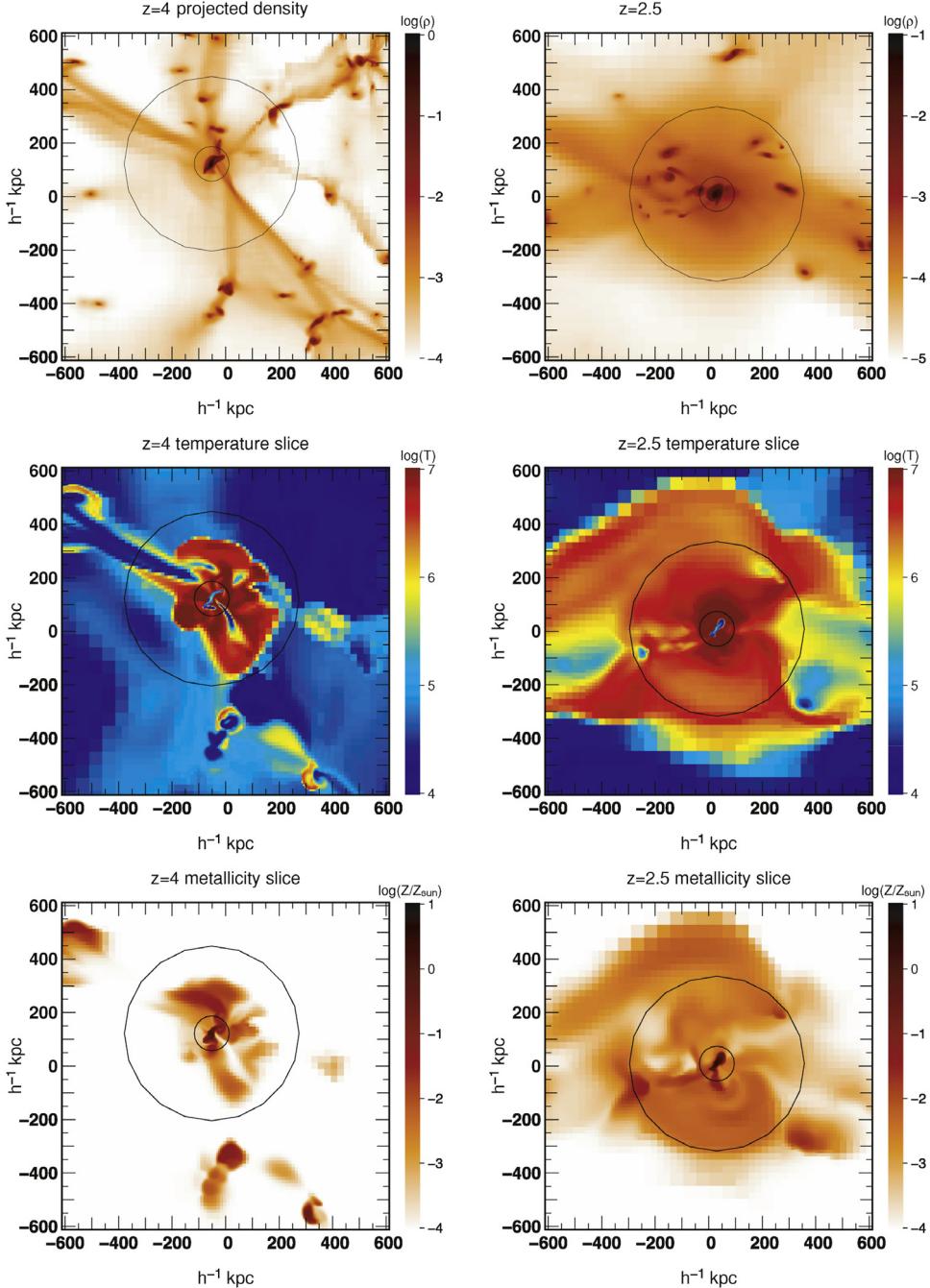
where  $f_{0.184} = 0.184$  is the baryonic fraction using the latest cosmological values [48]. As shown in Fig. 9, this increase in the accretion rate, combined with the HMF shifting to lower masses at earlier times, naturally results in cold-mode accretion becoming increasingly important at high- $z$  [168]; we however caution the reader that such studies are limited to  $z \lesssim 5$  and can therefore not be extrapolated to higher redshifts. Indeed, most ( $\simeq 60\%$ ) gas is accreted cold, with  $T \simeq 10^4$  K at  $z \simeq 5$ . As the cold-flow rate drops with decreasing redshift a bimodal temperature distribution emerges, at  $z \simeq 2.5$ , with roughly half the gas being accreted cold (at  $10^{4.5}$  K) and the other half hot (at  $10^{6.5}$  K) culminating with most gas being accreted via hot-mode accretion, with  $T \simeq 10^{5-7}$  K, by  $z \simeq 0$ .

Running high-resolution AMR simulations Ocvirk et al. [169] have confirmed such results for a MW-mass ( $M_h = 2 \times 10^{12} M_\odot$ ) galaxy as shown in Fig. 10. As can be seen, such a galaxy, at  $z = 4$  is predominantly fed by well-defined cold streams ( $T \lesssim 10^{5.5}$  K) at  $R_{vir}$  with the gas reaching  $T_{vir}$  only after penetrating as deep as  $0.5R_{vir}$  (see also [168]). On the other hand, a similar-mass galaxy at  $z = 2.5$  accretes from a hot quasi-spherical medium with  $T \gtrsim 10^{6.5}$  K at all  $r \lesssim R_{vir}$ . Critically, this model highlighted a crucial missing ingredient: the metallicity of the accreted gas that, together with its density, determines the cooling rate. As seen from Fig. 10, gas accreted via the cold mode shows a wide range of metallicities ranging from metal-poor ( $Z \sim 10^{-4} Z_\odot$ ) filamentary gas tunnelling in from  $\sim 2R_{vir}$  to the central galaxy to metal-rich ( $Z \sim 0.1 Z_\odot$ ) satellites. In comparison, gas accreted via the hot-mode has a much higher metallicity,  $Z \gtrsim 0.01 Z_\odot$ , at all radii.

However, the importance of cold-accretion sensitively depends on the feedback strength implemented in simulations [176]. Firstly, SN feedback can heat a significant fraction of the ISM gas, diluting incoming cold flows, reducing the importance of cold-flow accretion. Secondly, photo-evaporation or heating of gas from low mass halos due to the ionizing UVB (see Section 7.3) has a bearing on the importance of cold-flows in two ways: (i) the lack or decrease of gas in low-mass halos results in a lower amount of “merged gas”, with the rest being reclassified as “cold accretion”; and (ii) the UVB establishes an IGM temperature floor such that only halos with  $T_{vir} > T_{IGM}$  can accrete in the cold mode, thereby affecting the minimum value of  $M_{\text{shock}}$  [177].

By providing ready-to-use gas, cold flows are critical to the process of galaxy formation: firstly, they are responsible for assembling most of the fuel for star formation in high- $z$  galaxies, with mergers only playing an important role in the assembly of groups and clusters [148,173,177]. Secondly, they fuel the extremely high star formation rates ( $\simeq 100-200 M_\odot \text{yr}^{-1}$ ) observed in galaxies at  $z \simeq 2-3$  whose extended clumpy disks are incompatible with the compact, perturbed kinematics expected from mergers (see e.g. [171]). Thirdly, they are directly responsible for most of the disk growth at high- $z$  where filamentary accretion can keep disks intact [171,177]. Finally, they are probably responsible for the observed bimodality in galaxy populations separating massive, red ellipticals from low mass, blue star forming disks (e.g. [178]).

Despite their importance in driving galaxy formation, observations of cold accretion have taken a long time to build up. A key reason is the small solid angle covered by cold-flow accretion (e.g. [171,179,180]). Forming at the intersections



**Fig. 10.** Simulation maps [169] showing cold accretion onto a typical galaxy with  $M_h = 2 \times 10^{12} M_\odot$  at  $z = 4$  (left column) and  $z = 2.5$  (right column). The rows show the projected density (top panels), temperatures (middle panels) and metallicity (bottom panels). In each panel, the large and small circles show the location of  $R_{\text{vir}}$  and  $0.2R_{\text{vir}}$ , respectively. While this galaxy would predominantly accrete via cold, high-density filaments showing a wide range of metallicities at  $z = 4$ , it would accrete from a hot quasi-spherical, higher metallicity medium at a lower redshift of  $z = 2.5$ .

of filaments, rare, massive halos are typically fed by a few filaments with an opening angle of  $20^\circ$ – $30^\circ$  or a few percent of the sphere ( $\sim 0.4 \text{ rad}^2$ ). On the other hand, forming deep inside filaments, lower-mass halos accrete from a wider opening angle [171]. Using Lyman Limit Systems (LLS) and Damped Lyman Alpha systems (DLAs) as probes of cold-mode accretion, Faucher-Giguère and Kereš [179] find a covering fraction of  $< 15\%$  for a  $3 \times 10^{11} M_\odot$  halo at  $z = 2$  that increases to  $\sim 30\%$  for a lower mass ( $7 \times 10^{10} M_\odot$ ) halo at  $z = 4$ . This low-covering fraction, compared to the order unity covering fraction for galactic winds, leads to signatures of cold streams being swamped by those of outflows in stacked spectra [179].

Additionally, signatures of inflows can easily be masked by interstellar absorption at the systemic redshift [181] and when selecting absorbers via metal-lines, ambiguity remains in separating a galactic scale outflow falling back onto the galaxy (“galactic fountain”) from material being accreted from the IGM for the first time. As a result, observers have mostly had to resort to hunting for indirect signatures of cold flows although there are a few hints of direct detections, as summarized below.

(i) *Direct evidences of cold flow accretion:* Using a sample of about 680 UV-selected galaxies at  $z \simeq 2.4$ , Rakic et al. [182] find the optical depth of Ly $\alpha$  absorbers to be stronger in the transverse direction than along the line-of-sight, which has been interpreted as resulting from cold accretion on scales of 1.4–2 pMpc. Further, Rauch et al. [183,184] have detected extended, filamentary Ly $\alpha$  emission near faint LAEs, possibly powered by cold-gas accretion, which might be lower-mass analogues of the Ly $\alpha$  blobs (LABs) discussed next. Finally, Martin et al. [185,186] claim the Ly $\alpha$  emitting filament detected near a quasi-stellar object (QSO) at  $z = 2.8$  and the associated proto-disk (of mass  $4 \times 10^{12} M_\odot$ ) could only have formed as a result of cold-accretion from the cosmic web.

(ii) *Observing cold flows via Ly $\alpha$  blobs:* First detected at  $z \simeq 3$  by Steidel et al. [187], LABs are extremely luminous ( $L_{Ly\alpha} \approx 10^{44} \text{ erg s}^{-1}$ ) and extended ( $\sim 100 \text{ kpc}$ ) regions of Ly $\alpha$  emission. A number of theoretical studies have shown that LABs might result from the dissipation of gravitational energy through collisional excitations as cold gas flows into the potential well of a massive ( $M_h \gtrsim 10^{12} M_\odot$ ) dark matter halo (e.g. [188–191]). Indeed, ruling out alternative sources, namely Ly $\alpha$  emission originating from a dust-enshrouded highly star-bursting galaxy with large-scale superwinds or being powered by QSO activity, a number of works have claimed the detection of LABs as unambiguous evidence of cold-accretion at  $z \simeq 2$ –3 [192–195] although see [196] for a summary to the contrary.

(iii) *Observing cold flows via Lyman Limit (LLS) and Damped Ly $\alpha$  (DLA) systems:* Simulations show that cold-accretion is a dominant contributor to the cross-section of optically thick LLS ( $10^{17} < N_{\text{HI}} < 10^{20.3} \text{ cm}^{-2}$ ) and low-column density DLAs ( $10^{20.3} < N_{\text{HI}} < 10^{21} \text{ cm}^{-2}$ ), with most of this gas never having experienced a virial shock [174,179,197]. Indeed it has been argued that, despite comprising  $\lesssim 20\%$  of the LLS population at  $z \simeq 2.5$ –3.5 [198,199], low-metallicity LLS in the proximity of metal-enriched galaxies present a compelling case for metal-poor cold accretion (e.g. [200–202]).

(iv) *Observing cold flows via absorption line systems:* While individual low-metallicity metal line absorbers, e.g. MgII [181, 203,204], have been suggested as evidences of cold accretion, it is difficult to make strong statements from such claims given H I near galaxies does not necessarily need to be inflowing despite its low-metallicity.

## 5.2. The formation of the first stars

Once the first galaxies gain and cool their gas, the first stars can finally form a few hundred million years after the Big Bang, ending the “cosmic dark ages” of the Universe. Forming out of gas of primordial composition these PopIII stars bridge the gap between the primordial ( $Z \sim 5 \times 10^{-9} Z_\odot$ ) gas predicted by BBN and the lowest-metallicity Population II stars ( $Z \sim 0.005$ – $0.05 Z_\odot$ ) observed. In addition, these stars could help explain the constant Lithium abundance of metal-poor halo stars (which is lower than BBN predictions) in addition to providing the seeds for super-massive black holes (see e.g. [87]).

The first stars are predicted to form in low mass ( $M_h \simeq 10^6 M_\odot$ ) halos (the so-called minihalos) at  $z \simeq 20$ –30 that, corresponding to  $3$ – $4\sigma$  peaks in the gaussian random field of primordial density perturbations (e.g. [205]), can host dense gas (with hydrogen number density  $n_H \gtrsim 500/\text{cm}^3$ ) which can cool on timescales shorter than the free-fall time. Assuming a top-hat collapse and idealized virialized halos, these structures can attain a molecular hydrogen fraction that scales with the virial temperature as  $f_{H_2} \propto T_{\text{vir}}^{1.5}$  [205]. Given that the virial temperature of these halos is  $\sim 10^3$  K, i.e. well below the atomic cooling threshold temperature of  $10^4$  K, gas cooling in these halos predominantly relies on H<sub>2</sub> [206]. As early as in 1961, it was realized that H<sub>2</sub> can form in the gas phase through the following channel [207]:



where the free electrons are either remnants from the epoch of recombination [208] or are produced by collisional ionizations in accretion shocks during galaxy assembly [209]. H<sub>2</sub> can also form through the alternative channel:



However, as a result of H<sup>−</sup> (Eq. (32)) forming more efficiently as compared to H<sub>2</sub><sup>+</sup> (Eq. (34)), the former channel dominates in most circumstances (e.g. [205]).

One of the most critical quantities pertaining to PopIII stars is their “characteristic” mass indicating the scale, below which the IMF flattens or declines and, that contains most of the stellar mass depending on the chemical, dynamical and thermal properties of the primordial gas. In the early 2000’s, two main numerical approaches were adopted to model PopIII(1) stars. The first, using AMR, allowed simulating the formation of PopIII stars ranging from cosmological ( $\sim 10$  Kpc) to protostellar ( $\sim 10^{-3}$  pc) scales [210,211]. On the other hand, although possessing a smaller dynamical range, the SPH approach allowed modelling sink particles and a wider range of initial conditions [212,213]. Irrespective of the initial conditions and

methodologies, however, both approaches found PopIII stars to be very massive, with masses ranging between  $M_s = 30\text{--}300M_\odot$ . They also show a “preferred” state of the primordial gas, solely arising as a result of the micro-physics of H<sub>2</sub> cooling, corresponding to a critical temperature  $T_c \sim 200$  K and a critical density  $n_c \sim 10^4 \text{ cm}^{-3}$ . This critical temperature arises as a result of the fact that collisional excitations and radiative decays of rotational transitions are responsible for cooling gas at  $T \lesssim 10^3$  K, with the two lowest H<sub>2</sub> rotational energy levels being separated by about 512 K. Although collisions with H atoms populating the high-energy tail of the Maxwell–Boltzmann distribution can lead to slightly lower temperatures, H<sub>2</sub> cooling cannot proceed below  $T \lesssim 100$  K. On the other hand, the critical density arises as a result of the competition between radiative decays (that cool the gas) and collision de-excitations (that do not). Indeed, at  $n < n_c$ , the cooling rate is  $\propto n^2$ , decreasing to  $\propto n$  for higher densities. Once these characteristic properties are reached, H<sub>2</sub> cooling is rendered inefficient and the gas undergoes a period of quasi-hydrostatic contraction, cooling and collapsing once the Jeans (or Bonnor–Ebert; [214,215]) mass is reached at

$$M_J \simeq 500 \left( \frac{T}{200 \text{ K}} \right)^{3/2} \left( \frac{n}{10^4 \text{ cm}^{-3}} \right)^{-1/2} M_\odot, \quad (36)$$

where  $T$  and  $n$  represent the gas temperature and particle number density, respectively. This gas eventually reaches a density of  $n \sim 10^{8\text{--}10} \text{ cm}^{-3}$  at which point three-body H<sub>2</sub> formation ( $3\text{H} \rightarrow \text{H}_2 + \text{H}$ ) becomes effective [216]. Continued collapse leads to the gas becoming optically thick to H<sub>2</sub> at  $n \sim 10^{10} \text{ cm}^{-3}$  [217] before it finally becomes optically thick also in the continuum at  $n \sim 10^{16} \text{ cm}^{-3}$  [218]. The cooling then transitions from H<sub>2</sub> dissociation cooling to being fully adiabatic once all the H<sub>2</sub> is destroyed. We briefly note that PopIII.2 star formation is intrinsically different in the initial stages because a higher ionization fraction in the recombining gas enables enhanced H<sub>2</sub> formation. Allowing the gas to cool to a lower temperature, this increases the effectiveness of cooling via HD (deuterated hydrogen molecule) that can cool the gas to the CMB temperature [219, e.g.]. This lower temperature and higher critical density of HD produce lower mass ( $\lesssim 100M_\odot$ ) fragments (see Eq. (36)) after which the collapse proceeds as described for PopIII.1 stars.

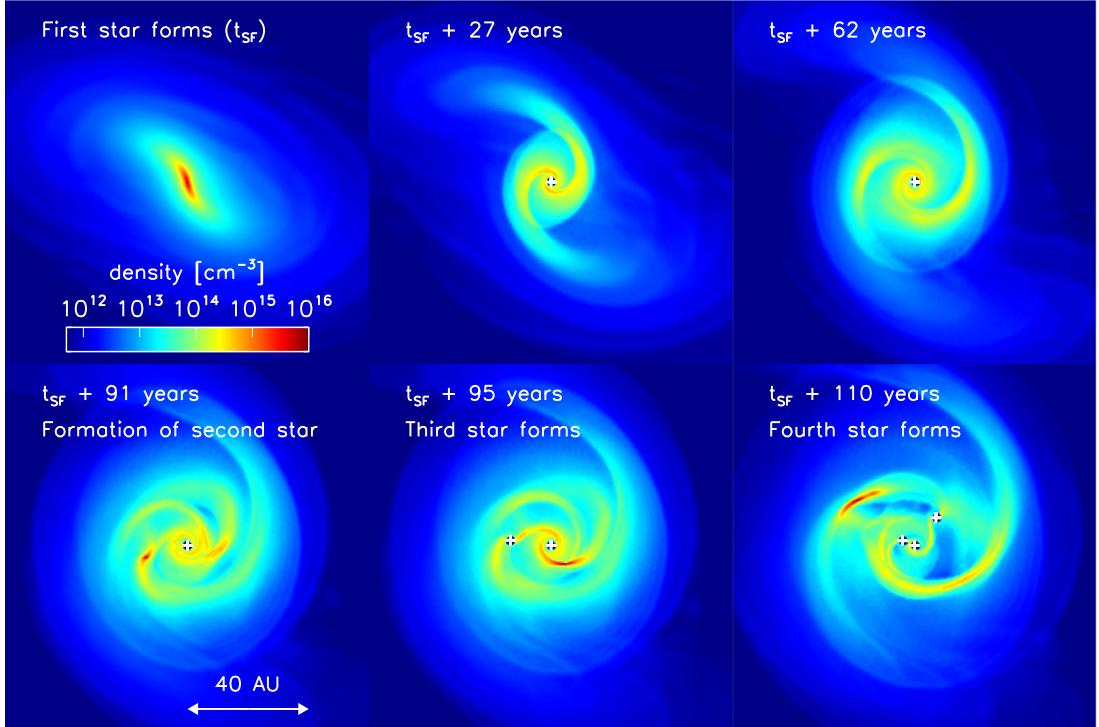
Initially, it was thought that each of these primordial protostellar clouds would form a single star, since no gas fragmentation was seen in simulations (e.g. [211,218–220]), yielding a final PopIII stellar mass of the order of  $30\text{--}300M_\odot$ . However, a paradigm shift occurred in year 2010 due to the growth in brute-force computational power that allowed simulations to use the concept of “sink particles” [221] and follow analyses beyond the formation of a protostar. The main idea was to replace gravitationally-bound gas that had collapsed below the resolution scale by a sink particle of the same mass that would only accrete gas gravitationally. Incorporating sink particles, both SPH [222–224] and AMR [225–229] simulations started showing gas settling into a protostellar disk around a hydrostatic core. This accretion disk is highly susceptible to gravitational fragmentation, as shown in Fig. 11, even for velocity dispersions as low as 20% of the sound speed, with a more active and widespread fragmentation in PopIII.1 halos [223]. The multiplicity (fragmentation into binary or multiple-stellar systems) and normalization of the resulting mass function depends on details including the sink implementation [227], gas metallicity – with lower metallicity gas being more prone to fragmentation [226] – and the simulation resolution [230]. Such multiplicity, however, could possibly be reduced by the late-time viscous merging of fragments in the disk as shown by recent works (e.g. [231]).

Interestingly the mass of the initial PopIII hydrostatic core corresponding to  $M_{\text{core}} \sim 5 \times 10^{-3}M_\odot$ , forming in the centre of the dense ( $n_H \sim 10^{22} \text{ cm}^{-3}$ ) collapsing cloud, is almost the same as that inferred for PopII star formation [232] since it essentially corresponds to the Jeans mass within the accreting protostellar disk. The final mass, however, depends on the accretion rate,  $\dot{M}_{\text{acc}} \propto c_s^3/G \propto T^{3/2}$ , a result that has now been confirmed using three-dimensional numerical simulations [233,234]. A comparison of typical gas temperatures in the birth clouds of PopIII ( $T \sim 200$  K) and PopII ( $T \sim 10$  K) stars clearly shows that the former accrete at a rate that is about two orders of magnitude higher than PopII star formation. The accretion rate is nevertheless limited by the Eddington luminosity,  $L_{\text{edd}}$ , such that

$$G \frac{M_p}{R_p} \dot{M}_{\text{acc}} \simeq L_{\text{edd}} \rightarrow \dot{M}_{\text{acc}} \simeq \frac{L_{\text{edd}}}{G} \frac{R_p}{M_p}, \quad (37)$$

where  $M_p$  and  $R_p$  are the mass and radius of the central protostar. Using a value of  $R_p = 5R_\odot$ , typical for PopIII stars, yields  $\dot{M}_{\text{acc}} \simeq 5 \times 10^{-3}M_\odot \text{ yr}^{-1}$ , a result that has also been confirmed numerically [211,217,233,235]. With accretion, the mass of the protostellar core grows as  $M_p \sim \int \dot{M}_{\text{acc}} dt \propto t^{0.45}$  which naturally implies that objects forming at lower redshifts have larger accretion rates as also shown by simulations [233]. However, it must be noted that disk fragmentation produces strong gravitational torques that cause the accretion rate to become extremely variable [224]. If the accretion rate becomes sub-critical early on, the dramatic expansion of  $R_p$ , that can stop accretion (see Eq. (37)), can be avoided resulting in the star reaching a final mass as high as  $700M_\odot$  at an age of about 3 Myr [160]. Interestingly, two-dimensional simulations, that do not follow disk fragmentation or non-asymmetric three-dimensional effects, obtain similar results with a final stellar mass of the order of  $1\text{--}1000M_\odot$  [236].

Whether or not such a high mass can be achieved, however, depends on a number of heavily debated feedback effects, such as H<sub>2</sub> photodissociation and radiation pressure from Ly $\alpha$  photons, that can terminate accretion. On the one hand, McKee and Tan [219] have shown that feedback effects do not significantly reduce the accretion rate until the ionized hydrogen (HII) region extends beyond the gravitational escape radius which generally occurs once the star has already reached  $50\text{--}100M_\odot$ , with a similar range of PopIII masses ( $\sim 1\text{--}100M_\odot$ ) being found by more recent hydrodynamic simulations [237]. However, hydrodynamic simulations also show scenarios where (i) most of the PopIII ionizing radiation escapes along the poles [238],



**Fig. 11.** The density evolution in a 120 AU region around the first protostar, showing the build-up of the protostellar disk and its eventual fragmentation on the order of a few tens of Myrs [224]; wakes in the low-density regions are produced by the previous passage of the spiral arms. These results have been obtained by re-simulating a collapsing over-dense region of  $1000M_{\odot}$  taken from a 200 kpc SPH simulation implemented with the idea of sink particles, allowing the authors to follow the fragmentation of the proto-stellar disk.

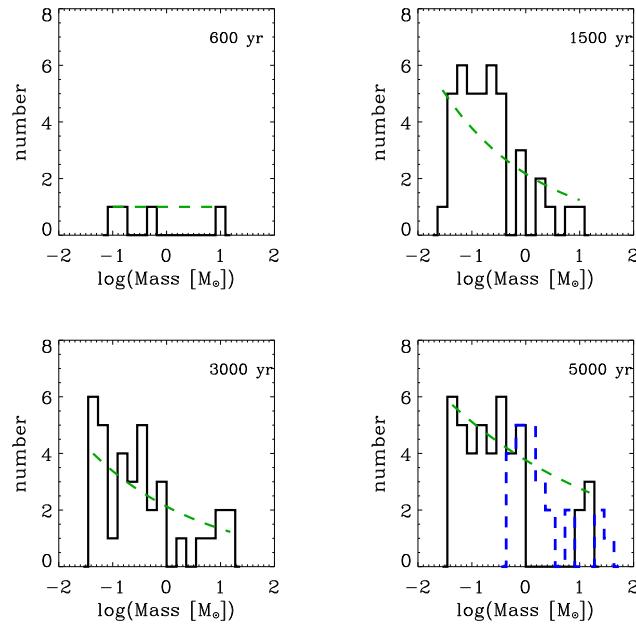
photo-evaporating the disk and halting accretion at about 0.7 Myr, resulting in a final PopIII stellar mass of  $43M_{\odot}$  [218]; or (ii) accretion rates drop, although gas can still fragment forming binaries, once the region of gas photo-heated by LW reaches the disk size, limiting PopIII stars to having masses  $\sim 0.05\text{--}20M_{\odot}$  [239,240]. These authors find a wide range of PopIII masses resulting in a broad, logarithmically-flat(ish) top-heavy IMF (see also [222,223,225]) as shown in Fig. 12.

Given the hierarchical assembly of galaxies in the  $\Lambda$ CDM model the feedback from these PopIII stars, significantly impacting the formation and evolution of subsequent galaxy generations (e.g. [241,242]), can be broadly divided into 3 main categories [87]:

(i) *Chemical feedback*: PopIII stars in the mass range of  $140\text{--}260M_{\odot}$  explode as Pair Instability Supernovae (PISN), with the range being extended down to  $85M_{\odot}$  for a rapidly rotating progenitor [243], completely disrupting the progenitor [244]; PopIII stars outside this range end up as black holes. These PISN have extremely large metal yields, on average converting roughly half their mass into metals, with iron making up the largest portion of the metal mass for the most massive PISN [245]. It has been numerically shown that, with PISN explosion energies of the order of  $10^{53}$  ergs, roughly 90% of these metals can escape their host minihalos and enrich the neighbouring  $\sim 1\text{kpc}$  to a metallicity  $Z \gtrsim 10^{-2}Z_{\odot}$  about 4 Myrs after the explosion [246,247]. Simulations of SN blastwaves [234,248] show that the early enrichment from PopIII stars is very inhomogeneous with low-density voids being enriched earlier than higher density filaments and virialised halos [249], extending the era of PopIII star formation. It is therefore clear that, enriching the Universe, albeit patchily, with its first metals, these PopIII stars mark the start of the end of metal-free star formation. The fine structure line cooling from singly ionized C (CII) and neutral oxygen (OI) enables gas to cool, to lower temperatures [246], and fragment leading to a shift from a top-heavy PopIII to a normal bottom-heavy metal-enriched PopII/I star formation once the critical metallicity threshold of  $Z_{\text{crit}} \simeq 10^{-3.5}\text{--}10^{-5\pm 1}$  is exceeded [250–254]. Further fragmentation can be achieved via dust continuum cooling if the dust-to-gas ratio exceeds a value of  $2.6\text{--}6.3 \times 10^{-9}$  [255,256]. The exact value of  $Z_{\text{crit}}$  is rendered unimportant by the fact that even a few SN explosions can enrich a  $10^8M_{\odot}$  halo above  $Z_{\text{crit}}$  [241,242,257–259]. However, even in the presence of metals, the minimum temperature floor is set by the CMB,  $T_{\gamma} = 2.732(1+z)$ , resulting in a  $z \simeq 15$  characteristic stellar mass of

$$M_s \sim 20 \frac{n_f}{10^4 \text{ cm}^{-3}} M_{\odot}, \quad (38)$$

where  $n_f$  is the density at which gas opacity prevents further fragmentation, giving rise to the idea of PopIII.2 star formation modulating the gap between PopIII and PopII stars (e.g. [260]).



**Fig. 12.** Simulating PopIII stellar systems starting from cosmological initial conditions and resolving scales as small as 1AU, [240] show the time evolution – from 600–5000 yr – of the sink mass function (black histograms) that clearly show the fragmentation of metal-free gas. The dashed green line in each panel shows the least-square fit to the function  $dN/dm \propto M_p^{-x}$  where  $x = (0, 1.24, 1.2, 1.13)$  at  $t = (600, 1500, 3000, 5000)$  yrs as shown. The dashed blue line in the last panel shows results from a lower resolution simulation without feedback.

(ii) *Mechanical feedback*: PISN have explosion energies  $E_{\text{PISN}} \sim 10^{51-53}$  erg [245], which is at least an order of magnitude larger than that produced by ordinary metal-enriched Type II SN (SNII; [261]). Such high explosion energies are larger than the binding energy of low-mass halos at early cosmic epochs (e.g.  $\sim 3 \times 10^{50}$  erg for a  $10^6 M_\odot$  halo at  $z \sim 20$ ), leading to a complete disruption of gas mass in the host halo [234,248], suppressing further star formation for a Hubble time  $\sim 300$  Myrs at  $z \sim 10$  [248,259]. However, Muratov et al. [262] show that the effects of such outflows are temporary, with cosmological inflows of gas restoring the baryon fraction to its universal value. Further, these authors show that, both, metal production and stellar feedback from PopII stars overtake those from PopIII stars in a few tens to hundreds of Myrs, depending on the galaxy mass. However, it must be noted that SN effects, although being debilitating for the host halo, could have exerted a positive mechanical feedback on neighbouring minihalos by shock-compressing the gas in their cores [248].

(iii) *Radiative feedback*: PopIII stars have a very hard spectrum, measured as the ratio of singly ionized He and hydrogen ionizing photons. The ionizing rate ratio is  $Q(\text{He}^+)/Q(H) \sim -1.4$  for PopIII stars as compared to  $\sim -2.6$  for stars with  $Z = 0.25 Z_\odot$ . PopIII stars also produce about 3 times more hydrogen ionizing photons compared to stars with  $Z \simeq Z_\odot$  for the same underlying IMF [263]. As a result, PopIII stars initiate both the epoch of hydrogen and helium reionization. Following the growth of HII regions from PopIII stars using radiation hydrodynamic simulations, a number of works have shown that a  $100 M_\odot$  star can create an HII region of about 3 kpc [264–266]. The extremely high IGM temperatures of about  $T \sim 20,000$  K in such ionized regions can result in the photo-evaporation of the gas from low-mass halos as discussed in Section 7.3 that follows. Indeed, while the threshold virial mass for PopIII star formation is effectively independent of redshift in the absence of a heating UVB, varying between  $1.5 - 7 \times 10^5 M_\odot$  at  $z \sim 19-33$  [233], it increases with decreasing redshift as more of the IGM is ionized. The HII ionization front can also drive shock waves (with a speed approximately equal to  $\sim 30$  km/s) that can expel about 90% of the gas from minihalos (escape velocity  $\sim 3$  km/s), leaving behind a warm ( $T \sim 3 \times 10^4$  K), diffuse ( $0.1 \text{ cm}^{-3}$ ) medium, delaying subsequent star formation by as much as 100 Myrs [267].

The UV radiation from the first stars would have also established a cosmological LW background radiation field (e.g. [268]).  $\text{H}_2$  can be easily dissociated by LW photons through the Solomon process ( $\text{H}_2 + \gamma \rightarrow \text{H}_2^* \rightarrow 2\text{H}$ ), where  $\text{H}_2^*$  is an excited electronic state. Although galaxies more massive than the earliest  $\sim 10^6 M_\odot$  minihalos could self-shield, the LW background could photo-dissociate their  $\text{H}_2$  at a rate  $\propto J_{21} \min[1, (N_{\text{H}_2}/10^{14} \text{ cm}^{-2})^{-0.75}]$  (with the first and second terms representing the LW background intensity and the shielding factor) in a static medium (e.g. [160,269]). In the absence of  $\text{H}_2$ , cooling can proceed via atomic transitions in  $T_{\text{vir}} > 10^4$  K halos where gas at the virial temperature can collapse into massive ( $\sim 10^6 M_\odot$ ) compact clouds in the centre of the halo, providing the seeds for the Super Massive Black Holes (SMBH;  $M_{\text{BH}} \sim 10^9 M_\odot$ ) [270,271] observed by the Sloan Digital Sky Survey (SDSS) at  $z \sim 6$  [272]. However, as mentioned above, since  $\text{H}_2$  formation is catalyzed by the presence of free electrons, processes such as X-ray production from SN remnants [273], accretion onto black holes [274], the presence of HII regions [275] and collisional ionizations driven by structure formation [209] can also have a positive effect on  $\text{H}_2$  formation. We end by noting that if PopIII stars had only

reached masses up to a few tens of  $M_{\odot}$  (as per the calculations of [237–239]), their disruptive effects would be less important than those discussed here.

We now briefly discuss observations of metal-free stars. We start with the field of “stellar archaeology” which focuses on hunting for metal-poor stars in the local Universe and we refer readers to a recent review [276] for more details. In brief, the idea is to use the fact that the metallicity of a star reflects the chemical conditions of its birth clouds – the chemical abundance of a star can therefore be used as a proxy for its age. Combining this with the fact that the lifetime of a star (of mass  $M_s$ ) scales roughly as  $10(M_{\odot}/M_s)$  Gyr, observations aim to hunt for  $M \lesssim 0.8M_{\odot}$  stars, with a lifetime  $\sim 19$  Gyr, in the Galactic halo and in satellite dwarfs, both, to find metal-free stars and to get hints on the (metal) enrichment of the Universe in its first few hundred million years. Moving farther to high- $z$ , as recognized by Partridge and Peebles [277], the earliest galaxies can be confirmed through nebular recombination lines (see Section 8.1) including Ly $\alpha$  (1216 Å) and Hell (1640 Å). Starting with Ly $\alpha$ , Kashikawa et al. [278] have detected a LAE at  $z = 6.538$  with a Ly $\alpha$  equivalent width of  $EW = 872^{+844}_{-298}$  Å that either implies an age  $\lesssim 4$  Myr and a metallicity of  $Z \lesssim 10^{-5}Z_{\odot}$  or a clumpy dust distribution. Additionally, as a result of their extremely hot atmospheres (e.g. [263]), PopIII stars produce copious amounts of Hell ionizing photons that can result in a Hell recombination line with EW of  $\gtrsim 100$  Å (e.g. [279]). A number of works have now shown that the JWST will be able to detect the signatures of PopIII star formation in the high- $z$  Universe. This could be achieved by detecting the Hell emission line from  $z = 10$  galaxies containing a minimum PopIII stellar mass of  $10^5 M_{\odot}$ . Alternatively, using an integration time of 100 h [280] the JWST could detect both the H $\alpha$  and Hell lines in  $10^9 M_{\odot}$  halos if the PopIII IMF is top-heavy [281]. Finally, it is possible to identify Pop III galaxies at  $z = 8–10$  from their peculiar colours resulting from the presence of strong (Ly $\alpha$  and H $\alpha$ ) nebular emission lines and the absence of any metal emission lines (e.g. [OIII] $\lambda 5007$  Å). This method could be sensitive to Pop III mass fractions as low as  $\lesssim 1\%$  of the total stellar mass [282]. These observations could also be rendered more sensitive using gravitational lensing to boost the luminosity of faint background sources (e.g. [283]). In addition to line emission, the extremely luminous PISN expected from PopIII stars ( $\sim 130–250 M_{\odot}$ ) could be detected by the hundreds by the JWST [284]. The challenge in this technique lies in building a statistical sample across multiple fields and separated by a long enough cadence to study the slowly evolving light-curves. Finally, the technique of intensity mapping – that focuses on observing the cumulative large scale fluctuations in intensity from faint, undetectable sources – could be employed by a future 2m space-based ultraviolet observatory to hunt for PopIII star formation at  $z = 10–20$  using multiple lines (e.g. Hell-CO or Hell-21 cm) in order to spatially correlate fluctuations and eliminate contaminants [285].

## 6. Heavy elements: production, transport and ejection

BBN is a key component of the Hot Big Bang model that explains the synthesis of light elements in the early Universe (for a detailed explanation see e.g. [286]). Taking place at an epoch when temperatures ranged between  $\sim 10–0.1$  MeV, BBN predicts that a substantial mass fraction of  ${}^4\text{He} = 0.2485 \pm 0.0002$  and trace amounts (of the order of  $10^{-5}$ ) of Deuterium/H and  ${}^3\text{He}/\text{H}$  formed about 3 min after the Big Bang, using the latest *Planck* cosmological parameters [55]. As the expansion of the Universe cooled matter, thermonuclear fusion could not proceed efficiently beyond  ${}^7\text{Li}$  resulting in an extremely low abundance of  ${}^7\text{Li}/\text{H} \sim 10^{-9}$ ; BBN is inefficient in generating  ${}^6\text{Li}$ ,  ${}^9\text{Be}$  and  ${}^{10,11}\text{B}$ . Moving further along the atomic table, Carbon is produced along inefficient paths involving  ${}^{11}\text{B}$  with radiative capture of  ${}^{12}\text{C}$  leading to production of trace amounts of Nitrogen and Oxygen with abundances of  $({}^{12}\text{C}, {}^{14}\text{N}, {}^{16}\text{O})/\text{H} = (4.4 \times 10^{-16}, 2.6 \times 10^{-17}, 1.8 \times 10^{-20})$  (e.g. [287]). Therefore metals heavier than  ${}^4\text{He}$  were only formed in substantial quantities after the first stars started the “metal-age” of the Universe at a (median) redshift of  $z = 65.8$ , roughly 30 Myrs after the Big Bang, at which there is a 50% probability of finding the first star (e.g. [288]). Producing the first heavy metals after BBN, these stars paved the way for the transition from PopIII to metal-rich PopII star formation as explained in what follows.

### 6.1. Metal production and mixing in galaxies

Before discussing the metal production in (PopIII and PopII) stars, we introduce a few key concepts. Firstly, the metallicity of gas is defined as the ratio of the total metal mass (of elements heavier than  ${}^4\text{He}$ ) to the total gas mass such that  $Z = \sum_{i>He} M_i/M_{\text{gas}}$ . This can be normalized to the solar metallicity that has a value  $Z_{\odot} = 0.0122$  [289]. The abundance of a specific element (say, Iron) relative to the Sun can then be expressed as

$$[\text{Fe}/\text{H}] = \log\left(\frac{n_{He}}{n_H}\right) - \log\left(\frac{n_{Fe}}{n_H}\right)_{\odot}.$$

The amount of metals produced by a given stellar population depends on a number of key parameters including the stellar mass, the initial stellar metallicity, the SN explosion energy and the IMF.

As noted in Section 5.2 the first stars formed out of gas of primordial composition with  $Z < Z_{\text{crit}} \sim 10^{-3.5} – 10^{-5.5} Z_{\odot}$  (e.g. [252,254]), had masses of up to  $300 M_{\odot}$  and a top-heavy IMF. PopIII stars in the range of  $140 – 260 M_{\odot}$  explode as PISN, releasing roughly half their mass into metals [245]. We can carry out a simple order-of-magnitude estimate of the volume that can be metal-polluted by such PISN following the calculations in Ferrara [290]: let us assume that the total amount of metals produced by a PISN is  $M_z \approx 2E_{51}M_{\odot}$  where the SN explosion energy is expressed in terms of  $E_{51} = E/10^{51}$  erg. Assuming a fraction  $f_w$  of the explosion energy to be converted into kinetic energy, these metals will be dispersed over a volume  $V = f_w E / \rho c_s^2$  where  $\rho$  and  $c_s$  are the density and sound speed in the surrounding medium, respectively. This

**Table 2**

SN metal yields [Oxygen (column 5), Silicon (column 6), Sulphur (column 7), Iron (column 8) and total (column 9)] for a range of SN models (column 1) using a Salpeter IMF with the lower and upper SN mass limits shown in columns 2 and 3, respectively, and the initial metallicity values shown in column 4. The IMF limits for the entire stellar population range between  $0.1\text{--}100M_{\odot}$  for SN [297] and between  $0.07\text{--}100M_{\odot}$  for the SNII+HN (hypernova) case [298], yielding a SNII rate of  $\nu = [280M_{\odot}]^{-1}$  SNII. In this specific table, solar metallicity corresponds to  $Z_{\odot} = 0.02$ .

SN	$M_l$ [ $M_{\odot}$ ]	$M_u$ [ $M_{\odot}$ ]	$Z$ [ $Z_{\odot}$ ]	$y_O$ [ $M_{\odot}$ ]	$y_{Si}$ [ $M_{\odot}$ ]	$y_S$ [ $M_{\odot}$ ]	$y_{Fe}$ [ $M_{\odot}$ ]	$y_Z$ [ $M_{\odot}$ ]	Reference
SNII	12	40	$10^{-2}$	1.430	0.133	0.064	0.136	2.03	[297]
SNII	12	40	1	1.560	0.165	0.078	0.115	2.23	[297]
SNII+HN	13	40	1	1.806	0.144	0.061	0.108	3.11	[298]
SNIa				0.148	0.158	0.086	0.744	1.23	[299]

can be used to obtain an estimate of the corresponding shell mass enriched by metals as  $M_m = \rho V = f_w E/c_s^2$ . Using typical values, of  $c_s = 10$  km/s and  $f_w = 0.1$  appropriate for high- $z$  galaxies (see e.g. [148]), yields a shell metallicity of  $\langle Z \rangle = M_Z/M_m = 3.2 \times 10^{-3}Z_{\odot} \gg Z_{crit}$ . Although the shell mass used here is an upper limit, given that radiative losses can decelerate the progress of the shock wave and we have assumed perfect mixing, a single PopIII star with an explosion energy of  $E = 10E_{51}$  can enrich the surrounding  $5 \times 10^5 M_{\odot}$  of gas above the critical metallicity, quenching all further PopIII star formation.

Continuing along these lines, we can also calculate the volume density of PISN required to enrich all of the gas in the Universe at  $z = 15$  (at critical density) to  $Z_{crit} = 10^{-5}Z_{\odot}$ , assuming each PISN to have a mass of  $100M_{\odot}$ , 45% of which is converted into metals, as:

$$\frac{1}{V} = \frac{\rho(z=15)Z_{crit}}{M_Z} \sim 10 \text{ Mpc}^{-3}. \quad (39)$$

This implies that 1 PISN in every  $\approx 10^8 M_{\odot}$  halo would be sufficient to enrich the entire IGM to  $Z_{crit}$ , modulo the fact that we have neglected metal enrichment from PopII stars and assumed all metals from PopIII to be homogeneously distributed throughout the IGM. In reality, however, the clustering of  $3\text{--}4\sigma$  density fluctuations, that host the first stars, results in the densest regions being patchily metal-enriched first as shown by simulations and discussed in Section 6.2 that follows. However, a number of simulations have shown that, given that PopIII stars contribute less than 1–10% to the total star formation at redshifts as early as  $z \simeq 15$  [291], they have a negligible contribution to the total metal budget of the Universe [229,292–294]. We now discuss metal production from PopII stars which are regarded as the key cosmic metal factories.

The IMF-averaged PopII SN yield of a given heavy element,  $y_i$ , can be calculated as

$$y_i = \frac{\int_{M_l}^{M_u} M_i \phi(M) dM}{\int_{M_l}^{M_u} \phi(M) dM}, \quad (40)$$

where  $\phi(M)$  expresses the IMF which has a slope of  $-2.35$  for the commonly used Salpeter IMF [295]. Further,  $M_l$  and  $M_u$  express the lower and upper mass ranges of stars that can explode as SNII and contribute to the metal content. However given that stars with  $M_s = 50 \pm 10M_{\odot}$  form black holes without contributing to the metal content [296] and the uncertain pre-SN evolution of  $8\text{--}11M_{\odot}$  stars, SN metal yields have only been tabulated between  $12\text{--}40M_{\odot}$  [297]. For this specific case, the number of SNII per unit stellar mass is found to be  $\nu = [280M_{\odot}]^{-1}$ . We tabulate the yields of some common elements in Table 2 - we note that while metal yields are almost independent of metallicity for  $10^{-4} \lesssim Z \lesssim 10^{-2}$ , they vary by about 10% above this value [297] as shown in the table.

Using the same calculations as carried out for PopIII star formation above, and using the yield and  $\nu$  values from Table 2, we can also calculate the average metallicity of a galaxy after an epoch of PopII star formation resulting in a total stellar mass  $M_*$  as

$$\langle Z \rangle = y_Z \nu M_* \times \frac{c_s^2}{f_w E_{51}}. \quad (41)$$

Plugging in a reasonable estimate for the star formation efficiency corresponding to 0.3% [148] (which yields  $M_* = 1.4 \times 10^6$ ), again using  $f_w = 0.1$ ,  $c_s = 10$  km/s, and an initial stellar metallicity of  $10^{-2}Z_{\odot}$ , we find  $\langle Z \rangle \sim 0.16Z_{\odot}$  for SNII metals only. This simple estimate assumes that most of the metals mix in the galaxy and do not escape from it via outflows (i.e. a closed-box solution), so it is perhaps an upper limit to  $Z$  for a given efficiency/yield as discussed in Section 6.2 that follows.

However, the details of how metals produced in stars are distributed and mixed into the surrounding gas and on increasingly large scales remains an outstanding question. Historically, this process has represented a major factor of uncertainty in building galaxy evolution models, and particularly so when numerical simulations are involved. This is because the mixing process involves a number of complex phenomena, such as conductive interfaces, dynamical instabilities, diffusion and turbulence cascades, acting on an extremely wide range of scales in a multi-phase ISM. As a side note, the term mixing is often used synonymously with dispersal in the literature; the two terms, however, involve physically distinct processes affecting different spatial scales. While dispersal strictly refers to the transport of bulk enriched material on

(super-)galactic scales independently of its degree of homogeneity, mixing is produced by the above mentioned microscopic processes enforcing a homogeneous chemical composition.

Metals initially contained in dense supernova ejecta are separated from the ambient medium by a contact discontinuity. The first step is then to break such separating surfaces. When the cooling time in the cavity becomes shorter than its age, a cooled shell will form that rapidly becomes unstable [300,301] to Rayleigh–Taylor (RT) and Kelvin–Helmholtz (KH) instabilities. At this stage metals are distributed in fragments whose size depends on the energetics of the SN explosion and the thermodynamical properties of the ambient medium. The RT fingers produced by RT instabilities penetrate deep into the high metallicity gas and may be eroded by KH instabilities due to the passage of rapidly moving hot gas, i.e., a mixing layer [302]. The shear flows produced in the mixing layer efficiently mix the freshly produced metals with more pristine gas. These fragments are then processed by thermal conduction (as they are still immersed into the hot cavity gas) and shocks. Both processes heat up the metal enriched gas until a sort of equilibrium temperature, close to the geometric mean of the fragments and cavity gas, is attained. This gas will then cool by emitting characteristic optical, infrared and ultraviolet lines. Cooling often involves non-equilibrium ionization and self-photoionization.

The transfer of heat eventually balances the temperature of the processed ejecta and ambient gas. However mixing requires also transfer of matter. This process results from diffusion [303] eventually resulting in a perfect mixing of the new heavy elements with those already present in the gas. While a complete theory of mixing is lacking, a simple dimensional analysis can provide some insight. The characteristic diffusion time over a length  $L$  is  $t_D = L^2/D$ , where  $D$  is the diffusion coefficient which is usually expressed as the product of the particle mean free path,  $\ell$ , and the r.m.s. velocity,  $\sigma = (3kT/\mu m_H)^{1/2}$ . Using standard values for the density and temperature of the hot halo and the H I disk, one finds that the diffusion time scale is much shorter in the halo ( $t_D \approx 4 \times 10^5$  yr) than in the disk ( $t_D \approx 6 \times 10^{11}$  yr). The fact that the mixing timescale in the disk is even larger than the present Hubble time raises a thorny problem. Thus, it seems that efficient mixing can be guaranteed only if enriched matter cycles from the star forming regions into the halo and then rains back in a sort of galactic fountain [304].

Alternatively one can envisage the presence of strong turbulent motions, initially suggested by Roy and Kunth [305] whose natural mixing timescale would be the eddy turnover time,  $L/\sigma_T$ , where  $\sigma_T$  is the r.m.s. velocity of the turbulent fluid. In practice, this process is akin to that at work in the mixing layers discussed above, and can uniformly mix metals on scales down to a pc. Along these lines, de Avillez and Mac Low [306] used simulations to compute the timescale for the variance decay of concentration fields with no continuing sources in SN-driven interstellar turbulence. However, they did not give a detailed study or discussion of the central physical issues in turbulent mixing, such as the production of small-scale scalar fluctuations and their effect on the mixing timescales. Similarly, while Klessen and Lin [307] considered the dispersal of tracer particles in supersonic turbulence, they did not study mixing in the sense of homogenization of concentration fluctuations.

Detailed modelling of metal mixing is not particularly abundant in the literature. Recent hydrodynamics simulations have employed parametric models to account for unresolved sub-grid metal mixing. One of the most successful examples is given by the approach developed by Pan and Scannapieco [308], who have carried out the first systematic numerical study of passive scalar mixing in isothermal supersonic turbulence. Other studies have, instead, concentrated on the effects of mixing on the global chemical properties of galaxies (e.g. [309,310], and in the enrichment process and chemical signatures of the first stars, e.g. [311–314]). More sophisticated hydrodynamics simulations incorporating a realistic multi-phase ISM, and improved microphysics and turbulence treatment are required to understand the details of metal mixing in galaxies.

## 6.2. Galactic winds and the enrichment of the IGM

Radiative cooling is very efficient in dense low-mass halos at high- $z$ . Left unchecked, this leads to an over-production of stars and too many baryons being locked up in condensed halos (as compared to observations), a problem canonically termed “overcooling” [315,316]. This problem can be alleviated by introducing SN feedback that can reduce the star-formation efficiency of small halos by ejecting their gas and quenching further star formation (e.g. [155,248,317]). We start by discussing the maximum amount of star formation that can be supported by a halo, purely based on equating the halo binding energy to the SN ejection energy [148]. The formation of an amount  $M_*(z)$  of stars at redshift  $z$  can impart the ISM with a total SNII energy ( $E_{SN}$ ) given by

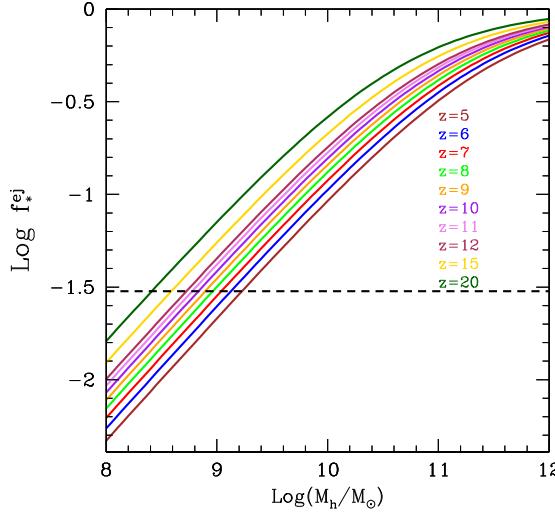
$$E_{SN} = f_w E_{51} v M_*(z) \equiv f_w V_s^2 M_*(z), \quad (42)$$

where each SNII is assumed to impart an (instantaneous) explosion energy of  $E_{51}$  to the ISM. Further, assuming all stars with mass larger than  $8M_\odot$  to explode as SNII yields  $v = [134 M_\odot]^{-1}$  for a Salpeter IMF between  $0.1 - 100 M_\odot$ . The values of  $E_{51}$  and  $v$  yield  $V_s = 611 \text{ km s}^{-1}$ . Finally,  $f_w$  is the fraction of the SN explosion energy that is converted into kinetic energy and drives winds.

For any given halo, the energy  $E_{ej}$  required to unbind and eject all the ISM gas can be expressed as

$$E_{ej} = \frac{1}{2} [M_{g,i}(z) - M_*(z)] V_e^2, \quad (43)$$

where  $M_{g,i}(z)$  is the gas mass in the galaxy at epoch  $z$ ; the term  $M_{g,i}(z) - M_*(z)$  accounts for the fact that SN explosions have to eject the part of the initial gas mass not converted into stars. Further, the escape velocity  $V_e$  can be expressed in terms of the halo rotational velocity,  $V_{vir}$ , as  $V_e = \sqrt{2}V_{vir}$ .



**Fig. 13.** The ejection efficiency ( $f_*^{ej}$ ) as a function of halo mass for  $z \simeq 5 - 20$ ; this is the star-formation efficiency required to eject all the gas from the galaxy and quench further star formation. The horizontal line shows the upper limit of  $f_* = 0.03$  required to match to high- $z$  observations (Section 8.2); we assume galaxies with  $f_*^{ej} > 0.03$  saturate at an effective efficiency of  $f_*^{eff} = f_* = 0.03$ . We have assumed each SN imparts an explosion energy of  $E_{51} = 10^{51}$  erg, of which a fraction  $f_w = 0.1$  drives winds and the SN rate ( $\nu$ ) is calculated assuming a Salpeter IMF between  $0.1 - 100 M_\odot$ . These values of  $E_{51}$  and  $\nu$  lead to  $V_s = 611$  km s $^{-1}$  as detailed in the text. Further, we note that as a result of the deepening potential wells, halos of a given mass can sustain higher  $f_*^{ej}$  values with increasing redshift.

We then define the *ejection efficiency*,  $f_*^{ej}$ , as the fraction of gas that must be converted into stars to “blow-away” the remaining gas from the galaxy (i.e.  $E_{ej} \leq E_{SN}$ ). This can be calculated as

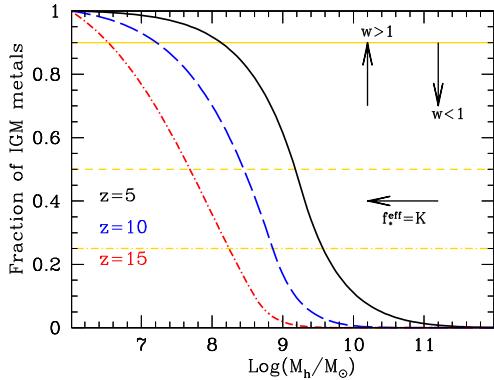
$$f_*^{ej}(z) = \frac{V_{vir}^2(z)}{V_{vir}^2(z) + f_w V_s^2}. \quad (44)$$

This represents the maximum fraction of gas that can be converted into stars *at a given instant of time*. The relation between  $f_*^{ej}$  and the halo mass, as a function of redshift, is shown in Fig. 13. The *effective star formation efficiency* can then be expressed as

$$f_*^{eff} = \min[f_*, f_*^{ej}], \quad (45)$$

where  $f_*$  is the saturation star formation efficiency. This represents the maximum fraction of gas that can be converted into stars in a galaxy, at a given time, without expelling all the remaining gas. Since  $V_{vir}$  scales with the halo mass, “efficient star-formers” (hosted by large dark matter halos) can continuously convert a fraction ( $f_*$ ) of their gas into stars, while feedback-limited systems can form stars with a maximum efficiency dictated by  $f_*^{ej}$  that decreases with decreasing halo mass. Matching the bright and faint ends of the evolving UV luminosity function (UV LF) at  $z \simeq 5 - 10$  requires  $f_* = 0.03$  and  $f_w = 0.1$  as will be shown in the later Section 8.2.

In addition to gas, these SN explosions also transport metals into the IGM thereby enriching it continually through time. We can carry out simple, order of magnitude, estimate to study the key IGM metal polluters as follows (e.g. [148]): assuming each halo to contain initially a cosmological gas mass fraction (i.e.  $M_{g,i} = (\Omega_b/\Omega_c)M_h$ ) and using the effective star formation efficiency calculated above, we can obtain the stellar mass as  $M_* = f_*^{eff}M_{g,i}$  and an outflowing gas mass  $M_{out} = (M_{g,i} - M_*)[f_*^{eff}/f_*^{ej}]$ . This produces a metal mass equal to  $M_z = y_Z \nu M_*$  resulting in a metallicity  $Z = M_z[M_{g,i} - M_*]^{-1}$ . Assuming perfect metal mixing in the ISM, this yields an outflowing metal mass  $M_{outz} = M_{out}Z$ . Multiplying this by the number density of the halo as obtained from the HMF, we obtain the cumulative (fractional) IGM metal contribution as a function of halo mass and redshift; note that, in the case of perfect metal-mixing, outflows do not impact the metallicity of the ISM ( $Z = M_z/M_g$ ) since the metal and gas mass losses compensate for each other. The results of this calculation are shown in Fig. 14. As seen, if galaxies residing in halos with masses as low as  $10^6 M_\odot$  form stars, 90% (50%) of IGM metals are provided by  $M_h \lesssim 10^{8.1} (10^{9.5}) M_\odot$  halos at  $z = 5$  with these numbers decreasing to  $M_h \lesssim 10^{6.5} (10^{7.8}) M_\odot$  by  $z = 15$ . If star formation is suppressed below the atomic cooling threshold, larger halos will naturally provide a larger fraction of IGM metals. This clearly shows the IGM metal pollution to be primarily driven by low-mass halos that lose their gas and metals in outflows. On the other hand, as a result of their larger potential wells, larger-mass halos keep most of their gas and metals bound within them at all  $z$ . As also shown, assuming a constant star formation efficiency, as opposed to a halo mass dependent one, would shift the metal contribution towards lower-mass halos. Relaxing assumptions of perfect mixing of metals and gas would result in a larger (smaller) metal enrichment if the outflowing gas were to be preferentially metal-rich ( $w > 1$ ) or metal-poor ( $w < 1$ ) where  $w$  is the ratio between the metallicity of the outflowing gas and the ISM metallicity.



**Fig. 14.** The cumulative fractional contribution to the IGM metal enrichment from galaxies below the halo mass value on the x-axis using the effective star formation efficiency as explained in the text in Section 6.2. As marked, the solid, dashed and dot-dashed lines show the results at  $z = 5, 10$  and  $15$ , respectively. The horizontal solid, dashed and dot-dashed lines show cumulative fractional contributions of  $90\%$ ,  $50\%$  and  $25\%$ , respectively. As shown, while galaxies below  $M_h \sim 10^{7.8} M_\odot$  contribute half of the total IGM metal budget at  $z = 15$ , this value shifts to higher values of  $M_h \sim 10^{9.5} M_\odot$  by  $z = 5$ . Changing the model assumptions, such as assuming outflows to be preferentially metal-rich ( $w > 1$ ) or metal-poor ( $w < 1$ ) can increase and decrease the fractional contributions. Further, assuming a constant star formation efficiency for all halos, as opposed to a mass-dependent value, will result in a larger fractional contribution from lower masses as marked.

The metal-enrichment of the IGM, driven by a central source (SNII in our case), can be calculated in more detail using the physical scenario detailed in [290,301,318,319]: the SNII explosion exerts pressure on the surrounding hot isothermal plasma with pressure  $p_b$  and temperature  $T_b$  resulting in the formation of a hot bubble, surrounded by a cool, dense spherical shell of mass  $M_m$ , radius  $R_m$ , volume  $V_m$  and thickness  $\delta R_m$ , expanding into an ambient medium of baryonic density  $\rho_b$  (the IGM in this case). The shell is driven by the thermal pressure of the internal gas which has to accelerate new material onto the expanding shell in addition to overcoming gravitational collapse. The equations of mass, energy and momentum conservation for this shell in an expanding Universe can be written as:

$$\frac{\dot{M}_m}{M_m} = \frac{1}{R_m^3 \rho_b} \frac{d}{dt} (R_m^3 \rho_b) \quad (46)$$

$$\frac{d}{dt} \dot{R}_m = \frac{8\pi G p_b}{\Omega_b H^2 R_m} - \frac{3}{R_m} (\dot{R}_m - H R_m)^2 - (\Omega - 0.5\Omega_m) \frac{H^2 R_m}{2} \quad (47)$$

$$\dot{E} = L - p_b dV_m/dt = L - 4\pi p_b R_m^2 \dot{R}_m. \quad (48)$$

where  $\Omega$  is the total density parameter and  $L$  is the luminosity that incorporates all sources of heating and cooling of the plasma.

The main features of this expanding bubble can be summarized as follows: for bubble ages  $t \ll t_H$ , both the impact of gravity and the Hubble flow are negligible, resulting in  $R \propto t^{3/5}$ . As  $t \rightarrow t_H$ , the evolution becomes much more complicated given that SN explosion energy input has ceased which, combined with the impact of gravity, results in a slower, decelerating expansion with  $R \propto t^{1/4}$ . Finally, at  $t \gg t_H$ , the shell gets frozen into the Hubble flow with  $R \propto t^{2/3}$  for  $\Omega = 1$ . However, at timescales much longer than the source lifetime, the metal-bubble volume,  $V_m$ , can be approximated by the Sedov-von Neumann-Taylor solution – the asymptotically adiabatic non cosmological limit to the equations above – resulting in

$$V_m = \frac{4\pi}{3} R_m^3 = \left( \frac{E_{51} v M_* t^2}{\bar{\rho}} \right)^{3/5}, \quad (49)$$

where  $E_{51} v M_*$  represents the total explosion energy input by SN during the formation of a stellar mass  $M_*$ ,  $t$  is the time for which the bubble expands and  $\bar{\rho}$  is the mean IGM density. Using numerical simulations, [294] find essentially all simulated bubbles expand for  $t \sim 250$  Myrs. Again, as above, using a typical  $10^{9.5} M_\odot$  halo at high- $z$  ( $z \simeq 7$ ) with  $M_* = 1.4 \times 10^6$ ,  $v = [134 M_\odot]^{-1}$ ,  $E_{51} = 10^{51}$  ergs and  $\bar{\rho}(z = 7)$ , we find  $R_m = 4.18$  kpc which is about 4 times larger than the virial radius of this halo ( $R_{vir} = 1.04$  kpc).

We can take one step further and calculate the metal filling fraction of the Universe,  $Q_m$ . As noted in Section 3.3, structures grow hierarchically in the  $\Lambda$ CDM framework resulting in a large number of low mass halos below the knee of the HMF which can be described by a power-law such that  $dN/d \log M_h \propto M_h^{(n_{eff}-3)/6}$  with  $n_{eff} = -2.2$  (e.g. [290]); we ignore the exponentially decreasing number density of high-mass halos that keep most of their gas and metals bound within their potential. As explained above, hot, metal-enriched gas driven by SN winds can escape its host halo, shock the IGM and form a blast-wave sweeping up a region of space which scales as  $R_m \propto E^{3/5}$  in the Sedov-Taylor phase. The filling factor of these

bubbles can then be calculated as [290,301]

$$Q_m \propto \frac{E^{3/5} dN/d\log M_h}{EdN/d\log M_h} \propto (dN/d\log M_h)^{2/5} \propto M_h^{5.2/15}. \quad (50)$$

### 6.3. The link between metal enrichment and reionization

The analytic calculations above, that find star-forming halos with the lowest masses to be the key metal-polluters, have now also been confirmed by semi-analytic models (e.g. [320]) - these show most of the metals to come from high- $z$  galaxies fainter than an absolute UV magnitude of  $M_{UV} = -15$ , corresponding to the faintest galaxies seen so far using the Hubble frontier Field data [40,321]. We caution the reader that this result has explicitly excluded galaxy clustering and re-accretion of metal-rich ejecta onto galaxies (the galactic fountain scenario) given numerical simulations are required for such analysis. As will be seen in Section 7.4, these low-mass halos are also considered to be the key reionization sources, establishing a critical link between the joint processes of reionization and the metal enrichment of the IGM in the first billion years.

The proper volume,  $V_{II}$ , of an ionized region can be calculated as [322] and see also Section 7.2:

$$\frac{dV_{II}}{dt} = \frac{dN_s}{dt} \frac{M_* f_{esc}}{n_H} - \frac{V_{II}}{t_{rec}} + 3H V_{II}, \quad (51)$$

where  $dN_s/dt$  is the production rate of ionizing photons and has an initial value (at  $t_0 = 2$  Myrs) of  $\dot{N}_s(0) = 10^{46.6} \text{ s}^{-1} \text{ M}_\odot^{-1}$  for a  $0.1 - 100 M_\odot$  Salpeter IMF and  $Z = 0.05 Z_\odot$  using *Starburst99* [323,324] and undergoes a time evolution as  $\dot{N}_s(t) = \dot{N}_s(0) - 3.92 \log(t/t_0) + 0.7$  [325]. Further,  $f_{esc}$  is the escape fraction of ionizing photons from the ISM as explained in Section 7.1. Given that the source lifetime is generally less than the recombination timescale and the Hubble time, we can write

$$V_{II} \simeq \frac{dN_s}{dt} \frac{M_* f_{esc}}{n_H} = f_{esc} \frac{N_\gamma M_*}{\bar{\rho}}, \quad (52)$$

where  $N_\gamma = 5 \times 10^3$  is the number of H I ionizing photons produced per baryon bound in stars [290].

We can now compare the IGM volume ionized by star formation in a galaxy to that metal-enriched by it as [290]

$$\frac{V_{II}}{V_z} = \left( \frac{E_{51} v M_* t^2}{\bar{\rho}} \right)^{3/5} \frac{\bar{\rho}}{f_{esc} N_\gamma M_*} \propto \left( \frac{\bar{\rho}}{M_*} \right)^{2/5} \propto n_*^{2/5}, \quad (53)$$

where  $n_*$  is the mean comoving number density of galaxies with stellar mass  $M_*$ . This implies that the metal enrichment tracks reionization more efficiently if the key sources are more numerous, i.e. have low masses.

Finally, we look at the expected mean metallicity of the IGM once reionization is complete following the calculations in [290]: the mass of metals produced by a stellar population of mass  $M_*$  can be expressed as

$$M_z = v y_z M_* = v y_z \mu m_p N_{b*} = v y_z \mu m_p \frac{N_s}{N_\gamma}, \quad (54)$$

where  $\mu = 1.22$  is the mean molecular weight and  $N_{b*}$  is the number of baryons in stars and we have used  $N_s = N_\gamma N_{b*}$ . Then using the fact that  $\kappa(z_{re}) n_H / (1 - Y)$  photons are required by the end of reionization, we obtain

$$M_z = v y_z \mu m_p \frac{\kappa(z_{re}) n_H}{(1 - Y) N_\gamma}. \quad (55)$$

Here,  $\kappa(z_{re})$  accounts for the fact that ionized hydrogen can recombine in over-dense regions, requiring multiple ionizing photons to maintain reionization which can be approximated as

$$\kappa(z_{re}) = \frac{t_H(z_{re})}{t_{rec}} \approx \frac{H_0 \Omega_m (1 + z_{re})^{3/2}}{[\chi_{HI} n_H \alpha_B(T) C]^{-1}} = 2.9 \left( \frac{1 + z_{re}}{7} \right)^{3/2}, \quad (56)$$

where  $\alpha_B(T)$  is the hydrogen case B recombination co-efficient and we have used a clumping factor  $C = \langle n_{\text{HII}}^2 \rangle / \langle n_{\text{HII}} \rangle^2 \simeq 3$  where the brackets denote an average over a cosmologically representative volume; these terms are detailed in Section 7.2 that follows. Combining Eqs. (55), (56) and the fact that  $\langle Z \rangle = M_z / M_H = M_z / (m_p n_H)$ , we obtain [290]

$$\langle Z \rangle = 5.5 \times 10^{-4} \left( \frac{1 + z_{re}}{7} \right)^{3/2} Z_\odot, \quad (57)$$

implying that the IGM, on average, is polluted well above  $Z_{crit}$  by the end of reionization. However, this does not preclude pockets of pristine gas in which a small fraction of PopIII stars can still form. Interestingly, an earlier reionization results in a higher average IGM metallicity value – this is because the higher the redshift, the denser is the Universe, requiring a larger number of ionizing photons to maintain reionization that results in a correspondingly higher value of the metal enrichment.

## 7. Ionizing radiation from early galaxies: the Epoch of Reionization

As noted in the last section, star formation (both PopIII and II) results in Lyman continuum (LyC;  $\hbar\nu > 1$  Ryd) photons that can ionize hydrogen, starting the Epoch of (hydrogen) Reionization; although quasars, powered by accretion of gas onto a black hole, also produce ionizing photons, we mostly focus our attention on star-forming galaxies in this review. Given the complex density fields of gas and dust in the ISM, however, only a small fraction of LyC photons actually emerge into the IGM. We start by discussing the escape fraction of hydrogen ionizing photons in Section 7.1 before moving on to the growth of ionized regions (Section 7.2) and the impact of reionization on galaxy formation (Section 7.3) and end by examining the key reionization sources (Section 7.4).

### 7.1. The escape of ionizing photons

The escape fraction of LyC photons,  $f_{\text{esc}}$ , denotes the ratio between the ionizing photons ( $\lambda < 912 \text{ \AA}$  in the rest-frame) that emerge out of the galactic environment, and ionize the IGM, and the total amount produced within a galaxy. The escape fraction is a crucial quantity for reionization: determining the history and topology of reionization, it directly impacts galaxy formation by establishing the UVB that can photo-evaporate gas from low-mass halos in ionized regions. On the other hand, the ionizing photons absorbed in the ISM ( $\propto [1 - f_{\text{esc}}]$ ) lead to the production of nebular emission lines. Significantly affecting the broadband fluxes [326], these can cause an appreciable variation in the inferred properties (e.g. the stellar mass and ages) of high-z galaxies [327,328]. However, a key bottleneck in reionization studies is that the value of  $f_{\text{esc}}$ , and even its dependence on galaxy mass and redshift, remain poorly understood at best as now discussed.

Pinning down  $f_{\text{esc}}$  observationally has presented a massive challenge. This is because it is nearly impossible to directly detect LyC photons at  $z \gtrsim 4$  given the increasing abundance of LLS (column density  $\simeq 10^{17} \text{ cm}^{-2}$ ) that lead to a 5-fold decrease in the LyC transmission from  $z \simeq 0$  to  $z > 3.5$  [329,330]. While LyC photons can be detected at slightly lower redshifts, inferring  $f_{\text{esc}}$  still requires knowing their intrinsic production rate which is hard to undertake since Balmer lines (specially H $\alpha$ ; used to infer this value) shift into the near-IR at  $z > 3$ . Therefore, what is generally measured is the flux ratio at 1500 Å versus the LyC,  $(L_{1500}/L_{\text{LyC}})_{\text{obs}}$ , which can be linked to the intrinsic flux ratio of the underlying stellar population,  $(L_{1500}/L_{\text{LyC}})_{\text{int}}$ , as (e.g. [331])

$$(L_{1500}/L_{\text{LyC}})_{\text{obs}} = (L_{1500}/L_{\text{LyC}})_{\text{int}} \times 10^{-0.4(A_{1500} - A_{\text{LyC}})} e^{\tau_{\text{HI,IGM}}} e^{\tau_{\text{HI,ISM}}}. \quad (58)$$

Here,  $A_{1500} - A_{\text{LyC}}$  is the differential dust extinction (between 1500 and 912 Å in the rest-frame) in magnitudes and  $e^{\tau_{\text{HI,IGM}}}$  and  $e^{\tau_{\text{HI,ISM}}}$  are the IGM and ISM optical depths to LyC photons, respectively. A relative escape fraction ( $f_{\text{esc}}^{\text{rel}}$ ) can then be defined as

$$f_{\text{esc}}^{\text{rel}} = \frac{(L_{1500}/L_{\text{LyC}})_{\text{int}}}{(L_{1500}/L_{\text{LyC}})_{\text{obs}}} e^{\tau_{\text{HI,IGM}}}, \quad (59)$$

which can be turned into an absolute  $f_{\text{esc}}$  using the conversion

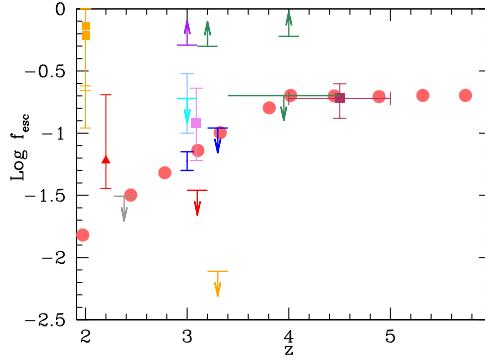
$$f_{\text{esc}} = 10^{-0.4A(1500 \text{ \AA})} f_{\text{esc}}^{\text{rel}}, \quad (60)$$

where the attenuation at 1500 Å,  $A(1500 \text{ \AA})$  can be related to the colour excess as  $A(1500 \text{ \AA}) = 10.33E(B - V)$  (e.g. [331]) using the Calzetti [332] dust reddening law inferred from local starbursts.

We show a compilation of the available observational estimates of  $f_{\text{esc}}$  at  $z \simeq 0 - 6$  in Fig. 15, which are now detailed; this comparison must, however, be treated with caution given the lack of uniformity between the different samples shown. Combining estimates of  $f_{\text{esc}}$  both from direct observations of LyC photons and those inferred indirectly from the observed ionizing background intensity, Inoue et al. [333] find the median  $f_{\text{esc}}$  value to rise by a factor of about 10 from  $f_{\text{esc}} \lesssim 1\%$  at  $z \lesssim 1$  to  $f_{\text{esc}} \simeq 10\%$  at  $z \gtrsim 4$ . An enormous amount of observational effort has been invested to supplement these numbers: using a ground-based custom-built narrow band filter to study the  $z \simeq 3.09$  cluster SSA22, Nestor et al. [334] find  $\langle f_{\text{esc}} \rangle = 12\%^{+11}_{-6}$  for LBGs, with LAEs showing a much higher value of  $\langle f_{\text{esc}} \rangle \gtrsim 17\%$ . Including spectroscopic data, these numbers were later refined to  $\langle f_{\text{esc}}^{\text{LBG}} \rangle \sim 5 - 7\%$  and  $\langle f_{\text{esc}}^{\text{LAE}} \rangle \sim 10 - 30\%$  [335]. These results were interpreted as implying high (low)  $f_{\text{esc}}$  values along lines-of-sight where SN feedback had (not) cleared the ISM of gas and dust. These are in good agreement with the value of  $\langle f_{\text{esc}} \rangle = 5.9\%^{+14.5}_{-4.2}$  inferred for  $z \sim 2.2$  galaxies [336] using ground-based data and  $\langle f_{\text{esc}} \rangle \lesssim 3.1\%$  at  $z \sim 2.38$  [337] and  $f_{\text{esc}} \lesssim 2.7 - 3.4\%$  using HST data (assuming  $E(B - V) = 0.1$ ) at  $z \sim 3.1$  [338]. These values are in accord with the low values inferred for bright galaxies ( $L \gtrsim 0.5L_{z=3}^*$ <sup>3</sup>) with  $f_{\text{esc}} < 11.1\% (0.77\%)$  at  $z \sim 3.3$  [340] ([341] where we have assumed  $E(B - V) = 0.1$ ) and  $f_{\text{esc}} \lesssim 5 - 20\%$  at  $z \sim 3.4 - 4.5$  [342] depending on how the data are selected (e.g. by magnitude or redshift). Taken together, these results imply an escape fraction that effectively decreases with increasing mass or luminosity while broadly showing an increase with redshift.

Complicating this picture, however, observations have revealed outliers at all redshifts ( $z \sim 2 - 5$ ) that show  $f_{\text{esc}}$  values larger than 20% and even as high as 100%. For example, Naidu et al. [343] find  $f_{\text{esc}} = 60\%^{+40}_{-38}, 72\%^{+28}_{-48}, 62\%^{+38}_{-51}$  for 3 galaxies at  $z \sim 2$ ; they detect 3 other active galactic nuclei (AGN) hosts that show values between 13 – 100%. At  $z \sim 3$ , Mostardi

<sup>3</sup>  $L_{z=3}^*$  is the characteristic luminosity at  $z = 3$  corresponding to  $M_{\text{UV}} = -21.0$  [339] using the AB magnitude system [43].



**Fig. 15.** A summary of observationally inferred  $f_{\text{esc}}$  values as a function of  $z$ . The plotted points show observational data collected by Inoue et al. [333, filled circles], Nestor et al. [334, violet filled squares], Nestor et al. [335, dark and light blue bars showing limits inferred for LBGs and LAEs, respectively], Matthee et al. [336, red filled triangle], Naidu et al. [343, orange filled squares] and Leethochawalit et al. [348, maroon filled square]. The downward facing arrows showing upper limits are from the work by Vasei et al. [337, grey], Siana et al. [338, red], Boutsia et al. [340, blue], Grazian et al. [341, orange], Vanzella et al. [342, green] and Mostardi et al. [344, cyan]. Finally, the upward facing arrows showing lower limits are from the work by Shapley et al. [345, purple] and Vanzella et al. 2016, 2017 [green; 327, 328]. Although showing a large scatter, these observations might be broadly taken to indicate a positive correlation of  $f_{\text{esc}}$  with redshift.

et al. [344] have inferred  $f_{\text{esc}} \sim 14 - 19\%$  for an extremely young galaxy with age less than 50 Myr and Shapley et al. [345] have inferred  $f_{\text{esc}} > 51\%$  even for a typical  $L_*$  galaxy. At  $z \sim 3.2$ , Vanzella et al. [346] have found a unique, compact Green Pea galaxy (Ion2) that shows an [OIII]4959,5007+H $\beta$  rest-frame equivalent width of 1600 Å and  $f_{\text{esc}} \gtrsim 50\%$  depending on the exact IGM attenuation used; in a future work [347] they have supplemented this with the serendipitously discovered “Ion3” at  $z \sim 4$  that also shows  $f_{\text{esc}} > 60\%$ . Finally, using gravitational lensing, Leethochawalit et al. [348] have inferred  $f_{\text{esc}} \sim 19 \pm 6\%$  at  $z \sim 4 - 5$ .

These results involve a few caveats: first, the escape fractions critically depend on the assumed star formation history and age of the underlying stellar population (that determines the intrinsic 1500 Å -LyC ratio), the inclusion of photons from binary stars, the E(B – V) value along the line of sight, and the precise model used to account for the IGM attenuation of LyC photons. Further, a number of studies [338,349,350] have found the LyC emission to be offset from the UV by as much as 1–1.5” – results presented above that assume co-spatial fluxes to constrain  $f_{\text{esc}}$  must therefore be used with caution. Moreover, other studies [338,341,350] find  $\gtrsim 50\%$  of bright LBGs to be contaminated by low- $z$  interlopers, requiring high resolution imaging and deep spectroscopy to distinguish LyC emitters from foreground galaxies. Finally, some works (e.g. [351]) caution that the standard LBG colour selection technique, requiring no flux blue-ward of LyC for a successful detection, is intrinsically biased against selecting LyC emitters. Indeed, this work finds  $f_{\text{esc}} \simeq 33 \pm 7\%$  for LyC emitting galaxies (LCGs) that drops by half, to  $f_{\text{esc}} \simeq 16 \pm 4\%$ , when considering the entire LBG+LCG population although crucially, these authors do not account for any IGM attenuation.

On the theoretical front, constraining  $f_{\text{esc}}$  requires coupling realistic renditions of ISM properties with a full radiative transfer code. The complexity of the problem has necessitated a number of theoretical approaches: we start with semi-empirical techniques that focus on inferring  $f_{\text{esc}}$  values that yield the correct CMB electron scattering optical depth for a galaxy population matched to observations. For example, it has been shown that simultaneously reproducing the high- $z$  UV LF, the electron scattering optical depth  $\tau \sim 0.08$  from the WMAP7- and WMAP9-year data [54,352] and Ly $\alpha$  forest statistics requires one of the following conditions be met [353–355]: (i) the UV LF should either be extrapolated to magnitudes as faint as  $M_{\text{UV}} \sim -10$  or  $-13$  compared to the current detection limits of  $M_{\text{UV}} \simeq -17$ ; or (ii)  $f_{\text{esc}}$  should increase from about 4% at  $z \simeq 4$  to about 100% at  $z > 10$ ; or (iii) a hybrid solution can be found wherein undetected galaxies contribute significantly and  $f_{\text{esc}}$  evolves modestly. However, the lower value of  $\tau = 0.05 - 0.06$  measured by Planck [47,48] is fully consistent with lower, and more reasonable,  $f_{\text{esc}}$  values ranging between 10% [356,357] and 20% [325,358]. Making the assumption of a minimum star formation rate required to blow outflows and clear channels for the escape of ionizing radiation [359], on the other hand, naturally find an (inverted) mass-dependence such that  $f_{\text{esc}}$  increases with mass. However, Ferrara and Loeb [360] caution that  $f_{\text{esc}}$  values can be large only if channels are opened within the first few Myrs after star formation, prior to SN explosion, since SN occur only after ionizing photon production has already started declining. Modelling low-mass  $T_{\text{vir}} \lesssim 10^4$  K galaxies, these authors show that  $f_{\text{esc}} = 0$  before the ionizing front generated by star formation reaches the virial radius, and effectively increases to 1, at a star formation efficiency of  $\sim 10^{-3}$ , once the ionization front breaks out to  $r \gtrsim R_{\text{vir}}$ .

Using SPH or AMR techniques, several authors have reported a wide range of  $f_{\text{esc}}$  values for low mass halos such that  $f_{\text{esc}} \simeq 0.15 - 0.6$  for  $M_h \lesssim 10^{6-7} M_\odot$  [145,361,362] and decreases to  $f_{\text{esc}} \simeq 0.05 - 0.4$  for  $M_h \lesssim 10^{8-9} M_\odot$  [145,151,363]. On the other hand, a number of works [364–366] find much lower values of  $f_{\text{esc}} \simeq 10^{-5} - 0.1$ , with only 30% of  $> 10^8 M_\odot$  halos showing  $f_{\text{esc}} \gtrsim 0.01$ . The variance amongst these results clearly demonstrates the complexity of the dependence of  $f_{\text{esc}}$  on the ISM density, clumping and turbulence structure as well as the availability of low-density lines-of-sight. The

results for high-mass ( $M_h \gtrsim 10^{11} M_\odot$ ) halos are in better agreement, yielding  $f_{esc}$  values ranging between 0.01 – 0.07 at  $z \simeq 3 - 9$  [151,367–369]. Interestingly, all of these works find  $f_{esc}$  to decrease with an increase in the halo mass because of a more clumped or higher density ISM. The only exception is the work by Gnedin et al. [369] who find  $f_{esc} \simeq 0.02$  for  $M_h \gtrsim 10^{11} M_\odot$  galaxies as a result of young stars lying outside the H I disk;  $f_{esc}$  decreases for low mass galaxies because of stars being deeper embedded inside thicker and/or more extended H I disks. Finally, in terms of its redshift evolution, most works converge on  $f_{esc} \simeq 0.01 - 0.02$  at  $z \simeq 2 - 4$  increasing with  $z$  to  $\simeq 0.06 - 0.1$  at  $z \gtrsim 3.6$  to as high as 0.8 at  $z \gtrsim 10.4$  [151,367,370].

These  $f_{esc}$  values are critically linked to a number of physical parameters: for example,  $f_{esc}$  values are found to be positively linked to the star formation rate and are the highest during the most intense starbursts, varying by as much as 1–6 orders of magnitude as a result of SN creating channels for ionizing photons [363,364,369,371]. The IMF is another critical quantity impacting  $f_{esc}$ . A Salpeter IMF results in lesser SN feedback and hence higher star formation rates implying a higher  $f_{esc}$  [363,368]. Indeed, Wise and Cen [363] find  $f_{esc}$  to increase from 0.25 for  $M_h > 10^8 M_\odot$  halos and a top-heavy IMF to 0.4 using a Salpeter IMF. Although anisotropic line of sight and orientation (edge-on or face-on viewing angle) effects can vary  $f_{esc}$  by about an order of magnitude [145,151,364,365,367,369], all these calculations agree that the effects of dust are not critical at high redshifts — since H I and gas densities are correlated, in regions of high H I density where dust masses can be high,  $f_{esc} \sim 0$  in any case, while regions of low H I density (where  $f_{esc}$  is high) do not contain much dust. Finally, the exact radius ( $0.5\text{--}2 R_{vir}$ ) at which  $f_{esc}$  is measured in not critical given most absorption takes place deep in the ISM, with clumps and filaments at large  $R_{vir}$  having a marginal effect due to their small solid angle for absorption [363].

We end with a few caveats. The first is that most massive stars exist in binary systems, a fact that is only starting to be included in population synthesis models. Enhancing the population of massive stars (at  $\gtrsim 3$  Myrs) after the burst, this binary fraction can lead to a larger  $f_{esc}$ , by as much as a factor of 3–10, by virtue of leakage channels produced after massive star formation [372] and/or a slower decline in the ionizing luminosity [373]. Further, a number of works [374,375] have shown that run-away stars, that travel at speeds of more than  $30 \text{ km s}^{-1}$  and can comprise as much as 30% of all massive stars in the MW, can travel 0.1–1 kpc from their dense birth clouds into low-density regions, increasing  $f_{esc}$  by a factor of 2–8 depending on the halo mass, galaxy geometry and the mechanism producing the runaway. A number of works also caution that  $f_{esc}$  values can be over-estimated by calculations that do not explicitly resolve high-density molecular clouds [362] or model the early phases of galaxy formation [361]. On the other hand, ignoring the high turbulence expected in high- $z$  galaxies [234,376,377] can under-estimate  $f_{esc}$  by factors of 2–4 even in media where the optical depth is  $\simeq 1$  [378].

We end on the positive note that forthcoming observations might present novel ways of shedding light on  $f_{esc}$ : spectroscopy with the *JWST* might potentially be used to identify strong LyC leakers ( $f_{esc} \gtrsim 50\%$ ) even up to  $z \simeq 9$  combining estimates of the UV spectral slope ( $\beta$ ; calculated between 1300–3000 Å in the rest frame) with the H $\beta$  emission line equivalent width, modulo caveats regarding the age, metallicity and dust attenuation [379]. Further, the [OIII] $\lambda 5007$ /[OII] $\lambda 3727$  ratio seems positively correlated with  $f_{esc}$  in density bound HII regions as shown by studies of local LyC leakers [380] that can be extrapolated to high- $z$  (see e.g. [381] and references therein). Finally, the H $\alpha$  and visible continuum surface brightness profiles, either around bright ( $L \gtrsim 5L_*$ ) galaxies or stacked faint *JWST* sources, could shed light on  $f_{esc}$  [382].

## 7.2. Growth and properties of ionized regions

As LyC photons stream out of sources and escape into the IGM they start interacting with the gas composed mostly of hydrogen and helium atoms. The main effect is that an HII region, bounded by an ionization front, will form around the source within which the gas is photo-heated to temperatures of the order of a few  $10^4 \text{ K}$  (e.g. [383]). For an ideal situation in which the IGM density is uniform the shape of the HII region will be spherical; in more realistic cases the shape will be largely determined by the distribution of density inhomogeneities producing a fluctuating opacity to the propagating photons.

Let us assume that a source emits LyC photons at a rate,  $\dot{N}_s$ . Then the physical expansion velocity (in excess of the Hubble flow) is regulated by the difference between the ionization and recombination rates occurring within the HII regions. Mathematically this yields [322,384]:

$$4\pi r_I^2 n_H \left( \frac{dr_I}{dt} - H r_I \right) = \dot{N}_s - \frac{4}{3}\pi r_I^3 n_H^2 C(z) \alpha_B(T), \quad (61)$$

where  $r_I$  is the proper radius of the ionization front,  $n_H$  is the mean hydrogen density of the IGM and  $C(z)$  is the clumping factor. Finally,  $\alpha_B(T)$  is the Case B radiative recombination rate to the  $n = 2$  level of hydrogen, allowing for ionizations from the ground state by the diffuse radiation emitted following recombinations to the ground state; if diffuse radiation is treated separately, then the Case A recombination rate should be used. The previous equation makes three assumptions: (i) the thickness of the I-front is negligible with respect to  $r_I$ ; (ii) the propagation velocity of the ionization front is supersonic with respect to the ambient gas and therefore the dynamical response of the gas to the sudden heating can be neglected (“R-type” front); and (iii) the velocity of the front is sub-luminal: this assumption might break down in the initial phases of the evolution (small  $r_I$ ). While Eq. (61) is very similar to the classical one for the Str’omgren radius first derived by Shapiro and Giroux [322], in an expanding universe the mean density decreases with time and ionized regions hardly have time to fill their Str’omgren radius before they overlap with another bubble.

As already mentioned, the presence of inhomogeneities in the gas distribution has essentially two different effects. First, it produces variations, which might be substantial, in the optical depth to ionizing radiation along different directions [385–387] as a result of which some of the line of sights from the source may be totally blocked. The gas behind these obscuring over-densities cannot be reached by ionizing photons and therefore remains neutral. As a result the I-front surface loses its spherical symmetry and becomes jagged, expanding faster into low-density voids. Second, the recombination rate, which is  $\propto n_{\text{H}}^2$ , is strongly enhanced in over-dense regions resulting in a considerably decreased final ionized volume [388]. These IGM inhomogeneities are expected [389,390] as a result of the typically lognormal distribution of the density field deduced from observations of post-reionization IGM (the so-called Ly $\alpha$  forest). However these are relatively mild and typically correspond to weakly non-linear fluctuations with over-densities  $\lesssim 30$ . However, due to gravitational clustering, the environment of a given galaxy might contain a number of lower mass nonlinear structures, i.e. mini-halos. These structures, which might be  $\approx 10^5$  times denser than the mean IGM [391], act as powerful sinks of ionizing photons. Because the clumping factor may be dominated by rare dense structures, establishing its value from numerical simulations is difficult and depends strongly on the numerical resolution adopted. A simple fitting formula has been provided in Iliev et al. [392],

$$C(z) \simeq 26.2917 \exp[-0.1822z + 0.003505z^2], \quad (62)$$

which is applicable in the redshift range  $6 < z < 30$ . However, this clumping factor can be significantly reduced once the impact of the reionization UV background is taken into account [393], as shown in Section 7.3 that follows.

As individual bubbles around sources grow with time they are likely to overlap with other bubbles produced by nearby sources. This process, by which the entire cosmic volume becomes ionized, can be described in statistical terms. Analogous to the metal enrichment of the IGM, we introduce the porosity parameter,  $Q_{\text{HII}}$ , i.e. the product of the number density of bubbles times the average volume of each of them.<sup>4</sup> Hence we define the time (or, equivalently, redshift) when  $Q_{\text{HII}} = 1$  as the reionization epoch.

The evolution of  $Q_{\text{HII}}(z)$  is simply given by [394,395]:

$$-H(z)(1+z) \frac{dQ_{\text{HII}}}{dz} = \frac{\dot{n}_S(z)}{n_{\text{H}}(0)} - \frac{C(z)Q_{\text{HII}}}{t_{\text{rec}}(z)}, \quad (63)$$

where  $n_{\text{H}}(0)$  is the comoving number density of hydrogen at  $z = 0$  and  $\dot{n}_S(z)$  is the production rate of ionizing photons per comoving volume. The recombination timescale,  $t_{\text{rec}}$ , at redshift  $z$  can be expressed as

$$t_{\text{rec}}(z) = [\chi_{\text{HI}} n_{\text{H}}(z) C(z) \alpha_B(T)]^{-1}. \quad (64)$$

Once  $Q_{\text{HII}} > 1$ , all the cosmic gas has been ionized. Towards the final phases of the reionization process the mean free path of the photons becomes comparable or larger than the Hubble radius  $c/H(z)$ . This implies that each patch of gas can see radiation coming from all the sources (the emissivity) and a roughly uniform UVB is established. In practice, though, the mean free path is limited to  $\lambda_{\text{mfp}} = c/H(z)(dN_{\text{LLS}}/dz)^{-1}$ , where  $N_{\text{LLS}}$  is the number of optically thick over-dense LLS. At  $z = 4$ , for example, Bolton and Haehnelt [396] estimate that  $dN_{\text{LLS}}/dz = 3.3$ . The intensity of the UVB is then simply written as  $J(z) \simeq \dot{n}_S(z)\lambda_{\text{mfp}}$ . Although the previous approach provides a broad-brush description of the reionization progress, and has been widely used given its simple analytical solution, the complexity of the problem can only be fully described by cosmological radiative transfer simulations. For more details, specially on the spatial distribution of ionized and neutral gas, we defer the reader to the extensive review by Meiksin [395].

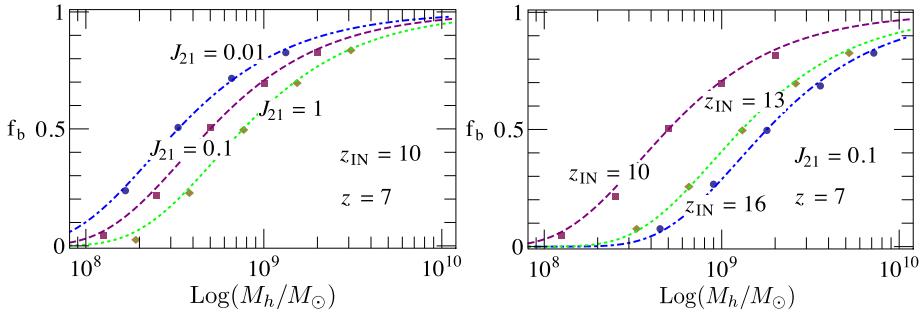
Finally, the above analytical treatment might not catch some deeper physical aspects related to the reionization topology. Furlanetto and Oh [397] pointed out that the naive view of discrete ionized bubbles growing independently of each other throughout reionization might be qualitatively wrong. Reionization, in fact, is inherently a problem in percolation and phase transitions physics. According to this scenario, reionization exhibits two distinct percolation phase transitions. In the first, occurring when the neutral hydrogen fraction is  $\chi_{\text{HI}} \approx 0.9$ , the bulk of the ionized gas is quickly incorporated into a unique infinitely large ionized region. In the second (when  $\chi_{\text{HI}} \approx 0.1$ ), a unique neutral region breaks into discrete neutral regions. The model predicts that the discrete ionized regions follow a power law size distribution, with nearly equal filling factors per logarithmic interval in bubble volume up to some characteristic size. This percolation theory could provide quite a precise description of reionization topology and evolution. It represents a promising research field over the next few years, particularly in view of the 21 cm experiments that aim to detect high- $z$  H I through its spin-flip transition emission.

### 7.3. Impact of global ultraviolet background on galaxy formation

We now discuss the feedback-effects of the UVB created by galaxies during reionization on subsequent galaxy formation. This is critical since feedback could, in principle, inhibit the star-forming capabilities of the very galaxies driving reionization, delaying its progress and changing its morphology.

During reionization, photoionization heating from the continually rising UVB can raise the gas temperature of about  $2 \times 10^4$  K in ionized regions as noted above. This temperature is equal to the virial temperature (see Eq. (20)), thereby being equal to the gravitational potential, of a bound halo of mass  $\simeq 10^{8.25} M_{\odot}$  at  $z \sim 15$ . As expected, the equivalence

<sup>4</sup> Sometimes the filling factor of the ionized regions,  $f_{\text{HII}} = (1 - e^{-Q_{\text{HII}}})$ , is also used; for  $Q_{\text{HII}} \ll 1$  they are almost equivalent.



**Fig. 16.** The baryon fraction,  $f_b$ , as a function of the halo mass at  $z = 7$  for different reionization histories [402]. The left panel shows results for a UVB intensity ( $J_{21}$ ) that varies over two orders of magnitude fixing the redshift ( $z_{in} = 10$ ) at which the halo is exposed to the UVB. As expected,  $f_b$  decreases with an increasing  $J_{21}$  value. The right panel shows results for a fixed  $J_{21} = 0.1$  but where  $z_{in}$  varies from 10 to 16. As shown, the earlier a halo of a given mass is exposed to the UVB, the lower is its resulting  $f_b$  value.

between the external radiation field and the internal binding energy can have two effects: firstly, it can photo-evaporate gas from the small potential wells of the lowest mass galaxies (in the minihalo/low mass end of atomic cooling halo regime). Secondly, it can increase the gas pressure, associated with the increase in IGM temperature, increasing the Jeans scale for gas collapse (e.g. [142,398,399]) and inhibiting gas from accreting onto galaxies. Both these effects can combine to suppress further star formation. However, since the baryonic content of a halo exposed to a UVB depends on multiple parameters, including the redshift, the thermal history and the intensity of the UVB, both the minimum halo mass that can host gas and the halo baryon fraction during reionization remain a matter of debate [137,139,400–403]. The gas mass in a halo of total mass  $M_t$  exposed to a UVB can be parameterized as [398]

$$M_g = \frac{f_b M_t}{[1 + (2^{\alpha/3} - 1)(M_{cr}/M_t)^\alpha]^{3/\alpha}}, \quad (65)$$

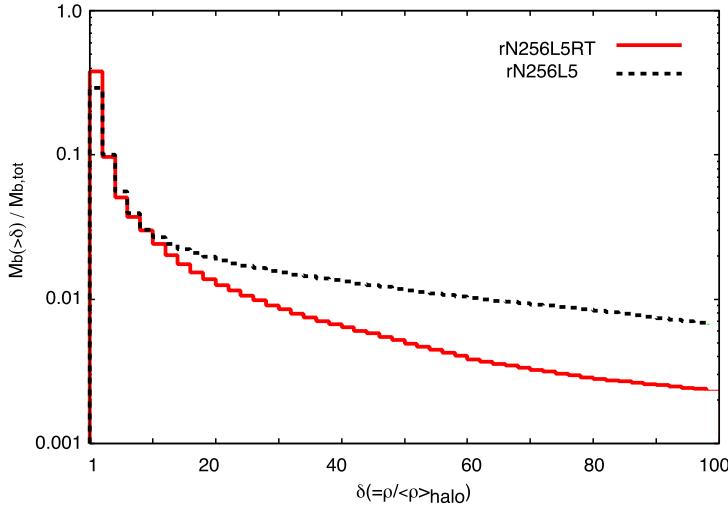
where  $f_b = \Omega_b/\Omega_c$  is the cosmological baryon fraction,  $\alpha$  is a free parameter and  $M_{cr}$  is the “characteristic” halo mass that can retain half its gas mass. The value of  $M_{cr}$  remains a matter of contention: Gnedin [398] find ( $\alpha = 1$ )  $M_c = M_F \simeq 10^{9.2} M_\odot$  at  $z = 6$  (see also [139,140,404] who find comparable mass suppression scales) where  $M_F$  is the filtering mass over which baryonic perturbations are smoothed in linear theory; this result is consistent with that found by Naoz et al. [405] and Okamoto et al. [401] at  $z \gtrsim 6$ . On the other hand, in the low- $z$  post-reionization regime, Okamoto et al. [401] find that, rather than the linear filtering scale, the characteristic mass that can accrete gas depends on the gas temperature at  $R_{vir}$ : if the gas temperature exceeds  $T_{vir}(R_{vir})$ , at a density that is a third of the halo virial over-density, no accretion takes place resulting in a much lower characteristic mass as compared to  $M_F$  (see also [404]). This result is supported by the work of Hoefft et al. [400] who also find the mass that can retain half its baryons as  $M_{cr} \ll M_F$  – this result is driven by the equilibrium temperature between photo-heating and radiative cooling at a characteristic over-density of  $\delta \simeq 1000$ ; only halos capable of compressing gas to this density despite a UVB can accrete gas. These authors also note that halos are never completely baryon-free since cold-phase gas and baryons bound into stars will resist evaporation. The discrepancy in these results possibly arises because mass scales have either been evaluated at the mean density [398] or at the halo virial densities [400,401].

In reality however, a spherically collapsing gas cloud never encounters densities within an order of magnitude of the cosmic mean [399] – while a gas cloud encountering gas densities comparable to the virial density of a halo will be able to radiate away its energy and keep collapsing, the bottleneck occurs at densities an order of magnitude less (more) than the virial density (cosmic density) where gas cannot cool efficiently. Using 1D collapse simulations, the above formalism has been extended to account for the delay in hydrodynamic response of baryons exposed to a UVB by calculating the baryon fraction as  $f_b(z) = 2^{-M_{cr}(z)/M_h}$  [402] where

$$M_{cr}(z) = 2.8 \times 10^9 M_\odot J_{21}^{0.17} \left( \frac{1+z}{10} \right)^{-2.1} \left[ 1 - \left( \frac{1+z}{1+z_{in}} \right)^2 \right]^{2.5}. \quad (66)$$

Here  $J_{21}/(10^{-21} \text{ erg s}^{-1} \text{ Hz}^{-1} \text{ cm}^{-2} \text{ sr}^{-1})$  expresses the intensity of the UVB and  $z_{in}$  is the redshift at which the halo of mass  $M_h$  is exposed to the UVB. As shown in Fig. 16, this formalism results in  $f_b$  decreasing with an increasing intensity of the UVB for a fixed value of  $z_{in}$  (left panel). Further, as expected, the earlier a halo is exposed to a UVB of a given intensity, the lower is its baryon fraction (right panel of the same figure).

In addition to the IGM, the UVB can also impact the ISM density: radiation hydrodynamic simulations show that accounting for UV feedback inside the ISM can decrease the baryon mass residing in the most overdense regions. As shown in Fig. 17 this severely suppresses the star formation capabilities of low-mass halos [139]. As UVB feedback affected halos merge, their lower baryon fraction values propagate through to the next generations of halos resulting in an observable impact on the SFR– $M_h$  [140] and  $M_* - M_h$  relations for higher mass halos [325].



**Fig. 17.** The impact of the UVB on ISM clumping [139]. This plot shows the fraction of baryonic mass residing in regions with a baryon density exceeding the value on the x-axis for a typical  $6.7 \times 10^9 M_\odot$  halo. As clearly seen, including UV feedback (red solid line) smoothes out the baryon mass in the most over-dense (inner-most) regions as compared to a case without UV feedback (dashed black line) by as much as a factor of 3, severely suppressing the star-formation rate.

In addition to such “negative feedback”, the UVB can also have a positive (accelerating) effect on reionization by decreasing the IGM clumping. Governing the recombination rate of the gas, the clumping factor is one of the key parameters determining the progress of reionization and the persistence of ionization in over-dense regions. Simulations have now confirmed that photo-ionization heating can act as an effective pressure term by increasing the Jeans mass and smoothing out density perturbations on small scales [406–408], especially if X-ray sources had already pre-heated the gas (e.g. [409,410]). This leads to a clumping factor that reduces with  $z$  as [406,411]:

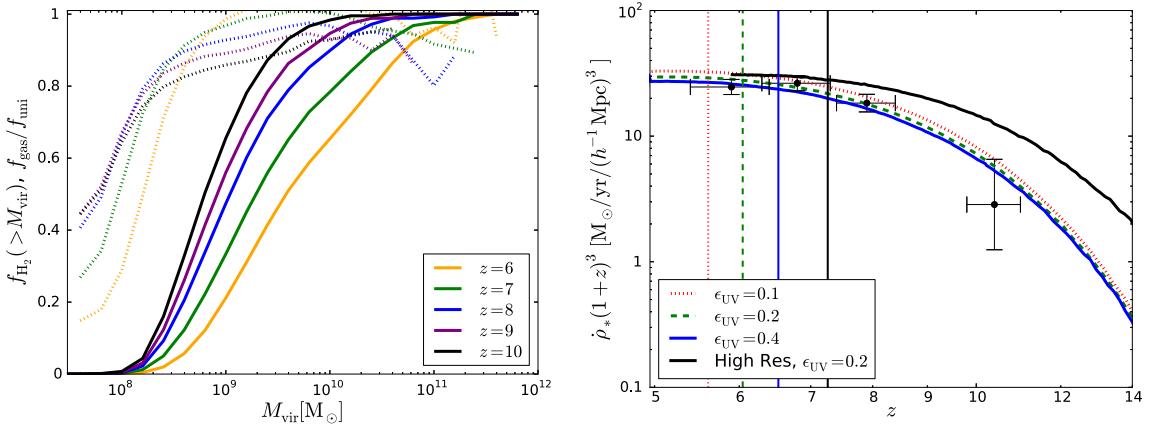
$$C(z) = 1 + 43 z^{-1.71}. \quad (67)$$

As expected, a decrease in the recombination rate has a positive impact on the number of photons needed to complete and maintain reionization.

In terms of the UVB impact on the reionization topology, McQuinn et al. [412] have coupled N-body simulations with a RT code to show that, for a given neutral fraction, the UVB suppression of star formation in ionized regions results in HII regions becoming larger (by as much a factor of 4). These HII regions are also more spherical since reionization shifts to being driven by massive rare sources. Indeed, simulations show a more uniform distribution of HII regions, peaked on smaller scales, and the large scale power reduced by tens of percents in the presence of a UVB [403]. However, both these authors point out that the reionization morphology is primarily dependent on the value of the average neutral hydrogen fraction,  $\langle \chi_{HI} \rangle$ , and only secondly on the nature and properties of the sources.

Finally, as noted in Section 3.1, CDM exhibits a number of problems in matching theoretical predictions with observations on small scales. One of the key manifestations of this is the observed lack of theoretically predicted satellites of the MW (“the missing satellite problem”; [58,59]). One possible explanation is the UVB feedback from reionization: using N-body simulations and abundance matching, Busha et al. [413] have shown that theoretical results match a number of observables (including the luminosity function, radial distribution and circular velocity distribution) for MW dwarfs if only halos with  $T_{vir} \gtrsim 10^5$  K can form stars assuming a reionization redshift of  $z_{re} = 8^{+3}_{-2}$ . Further, changing  $z_{re}$  can change the dwarf satellite number density by as much as two orders of magnitude. Similar results have been obtained Lunnan et al. [414] using N-body simulations of the MW coupled with a semi-analytic galaxy formation model where suppressing star formation in halos with  $V_{vir} \lesssim 50 \text{ km s}^{-1}$  can change the satellite numbers by a factor of 3–4 as  $z_{re}$  changes between 7.4 – 11.9. Further, Keck and *HST* photometry of (6) MW dwarfs indicate 100% of their stars having formed by  $z \sim 3$  – the similarity of their ancient population suggests a synchronized truncation of star formation driven by an external influence such as reionization [415]. In terms of galaxy populations, the increasing deficit of baryonic mass with decreasing halo mass inferred, using *SDSS* and *ALFALFA* (Arecibo Legacy Fast ALFA survey) observations, might have a contribution from reionization feedback in addition to SN-driven outflows [416].

A number of calculations, summarized below, have been used to couple UV feedback and galaxy formation. Details concerning the box sizes and resolutions can be found in Table 1 in Section 4.2. Essentially, the problem is one of the physical scales that need to be modelled: while galaxies have physical scales of a few kpc, reionization and its impact needs to be modelled on scales of  $\sim 100$  Mpc (as required for a convergent topology [134]), especially in the end stages when almost all galaxies are immersed in the UVB. A number of different techniques have therefore been adopted to circumvent the problem. These range from semi-analytic models [402,403] to N-body simulations/semi-analytic models coupled with



**Fig. 18.** *Left panel:* The average gas fraction (dotted lines) and the molecular gas fraction (solid lines) as a function of  $z$  using high-resolution ( $20h^{-1}$  Mpc;  $1024^3$ ) simulation runs from Gnedin and Kaurov [146]. As shown, the low-mass halos most affected by UV feedback contain little molecular gas and therefore, do not form stars in any case. *Right panel:* the impact of UV feedback on the global star formation rate density for different simulation runs [146]. The black line corresponds to the high resolution simulation shown in the left panel, with the other lines showing results for the fiducial ( $40h^{-1}$  Mpc;  $1024^3$ ) model for three different values of  $\epsilon_{\text{UV}}$  – the escape fraction of ionizing radiation up to the resolution limit; black points show observational data from [417]. As shown, reionization does not impact the global star formation rate in these simulations.

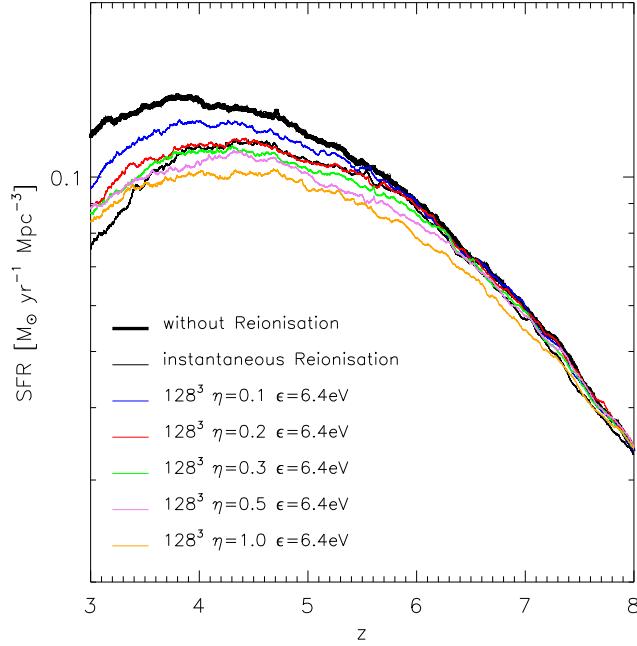
RT [147,418] to high-resolution, small volume radiation-hydrodynamic simulations [138–140,142,146,419,420] to high-resolution, intermediate volume radiation-hydrodynamic simulations [149]. These have led to often conflicting scenarios in which the UVB either has a negligible effect or fully suppresses star-formation in low mass galaxies as now summarized:

(i) *Negligible effects.* A number of works (e.g. [403,421]) note that accounting for the delay in the hydrodynamic response of baryons to the impinging UVB is crucial to prevent (a spurious) overestimation of the impact of the UVB. Using a variety of approaches mentioned above, the star formation rate density (and by extension reionization) is found to be impervious to the UVB unless the key reionization sources are either molecular-cooling driven [403] rapidly losing their gas after SN explosions [147,420] or low-mass galaxies that contain little to no molecular gas in the first place [146,147] as also shown in Fig. 18. The impact of the UVB would also be mitigated if the gas fuelling star-formation had already cooled into halos by the time it is exposed to the UVB [421]. Further, although cumulatively decreasing the gas content of early galaxies through time, the UVB can only suppress the number of ionizing photons-to-baryon ratio by about a factor  $\lesssim 2$  towards the end stages of reionization, when most galaxies are immersed in ionized regions. This corresponds to changing the end redshift of reionization by  $\Delta z \lesssim 0.5$  [403]. Further, Gnedin and Kaurov [146] show that, given that the galaxies affected by UV feedback contain little to no molecular gas in the first place, UV feedback has no observable impact (slope changes by 0.1 for a unit shift in  $z_{\text{re}}$ ) on the UV LF.

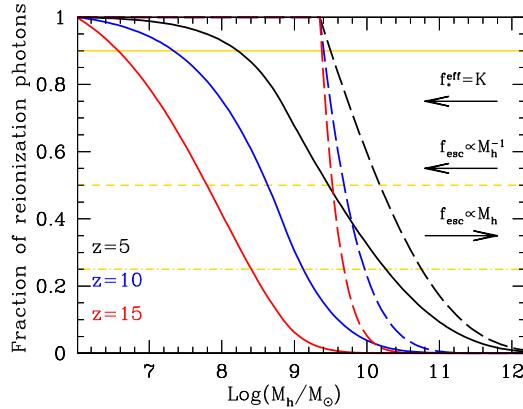
(ii) *Strong effects.* On the contrary, several other works find UV feedback to have a large effect both on reionization and the properties of high-redshift galaxies. Combining SPH simulations with an optically thin Eddington tensor approximation, Petkova and Springel [142] find that the increase in gas temperature, and the corresponding decrease in gas densities due to pressure effects, leads to gas cooling and collapsing more slowly. This results in a decrease in the SFR density at a given  $z$  as shown in Fig. 19. This effect is found to be the most pronounced for the lowest mass halos ( $T_{\text{vir}} \simeq T_{\text{gas}}$ ). Further, the SFR efficiency decreases with an increase in  $f_{\text{esc}}$  ( $\eta$  in Fig. 19) since a larger number of ionizing photons are available to heat the gas; although [423] find similar results, they show that SN feedback is also responsible for dispersing the cold, dense gas in the halo core. This is in agreement with results from radiation hydrodynamic simulations that show the UVB can suppress star formation in halos with  $T_{\text{vir}} < 10^5$  K, corresponding to  $M_h \lesssim 10^{9.2} \text{M}_{\odot}$  at  $z = 6$  [140]. These authors caution, however, that the effects of SN feedback and the UVB couple positively at small scales – outflows decrease the star formation rate, resulting in a lower UVB, permitting low-mass halos to retain their gas. Further, they also find that, even, low-mass halos recently exposed to a UVB are resilient to its effects since most of the gas has cooled to high densities and will form stars for many dynamical timescales. Although [139] find similar trends of UV feedback suppressing the star formation rate density by a factor of 5–6 at  $z = 6$ , their work results in much stronger feedback as compared to other works [140,142,363]. This arises from their inclusion of two ISM effects: (a) the internal UV field generated by star formation inside low mass ( $M_h \lesssim 10^9 \text{M}_{\odot}$ ) halos causes the dispersal of high density regions, and (b) the merger of low-mass systems into larger ones destroys dense gas preventing  $\text{H}_2$  formation and cooling.

#### 7.4. The sources of reionization: galaxies, quasars or ...?

We are now in a position to discuss the key reionization sources. Reionization should naturally begin with the formation of the first metal-free stars at  $z \simeq 15$ , with the key sources shifting to larger metal-enriched halos, powered by PopIII/II stars and even black holes, at later times. However, this naive picture is complicated by the fact that the sources of reionization



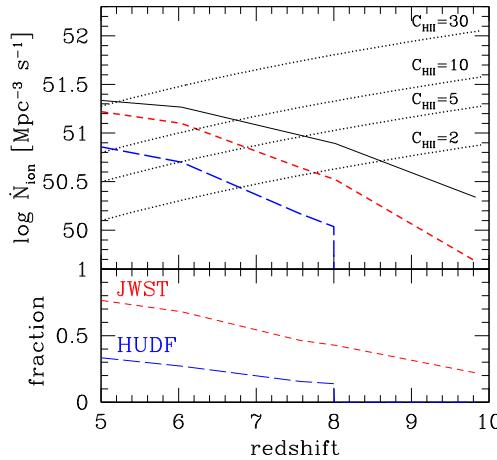
**Fig. 19.** The impact of the UVB on the star formation rate density (SFRD; [142]). As marked, the thin black line shows the SFRD, at its maximum value, in a model without any UV feedback. The think black line shows the SFRD in the case of instantaneous reionization at  $z = 6$  using the Haardt & Madau background [422]. The other lines show the values coupling GADGET with a moments method of radiative transfer for ISM escape fraction ( $\eta$ ) values ranging between 10–100%. As expected, the impact of UV feedback increases with increasing  $\eta$ .



**Fig. 20.** The cumulative fractional contribution to LyC photons from galaxies below the halo mass value on the  $x$ -axis. Assuming each halo to contain a cosmological baryon fraction, a constant  $f_{esc}$  and a constant clumping factor  $C = 1$ , we show results (for  $z = 5, 10$  and  $15$  as marked) using the maximal star formation efficiency–halo mass relation derived in Section 6.2. The solid lines show results without any UV feedback. Dashed lines show results for a maximal UV feedback scenario where all galaxies with  $V_{vir} < 50 \text{ km s}^{-1}$  are assumed to be devoid of baryons due to photo-evaporation (i.e.  $f_b = 0$ ). The horizontal lines show the 25, 50 and 90% fractions. As seen, while  $M_h \lesssim 10^{8.7} M_\odot$  galaxies can provide half of the LyC photons using a halo mass dependent star formation efficiency, this value increases to  $M_h \lesssim 10^{9.8} M_\odot$  if only galaxies above  $V_{vir} > 50 \text{ km s}^{-1}$  can support star formation. Changing the model assumptions, such as assuming  $f_{esc}$  to increase (decrease) with the halo mass will result in a larger fractional contribution from higher (lower)  $M_h$  values. Further, assuming a constant star formation efficiency for all halos, as opposed to a mass-dependent value, will result in a larger fractional contribution from lower masses as marked.

depend on a number of poorly understood parameters including the minimum halo mass that can host enough cool gas to allow star formation, the star formation/black hole accretion rates and intrinsic stellar populations at high-redshifts, the escape fraction of LyC photons and the impact of the UVB on the gas mass (and hence star formation rates) of low-mass halos. It is precisely because of these reasons that the nature and properties of the reionization sources, ranging from star forming galaxies to QSOs to exotic scenarios (e.g. decaying dark matter) remain debated.

Assuming galaxies to be the key reionization sources one can try and estimate the contribution of different halo masses to the ionizing photon budget at that redshift, the results of which are shown in Fig. 20.

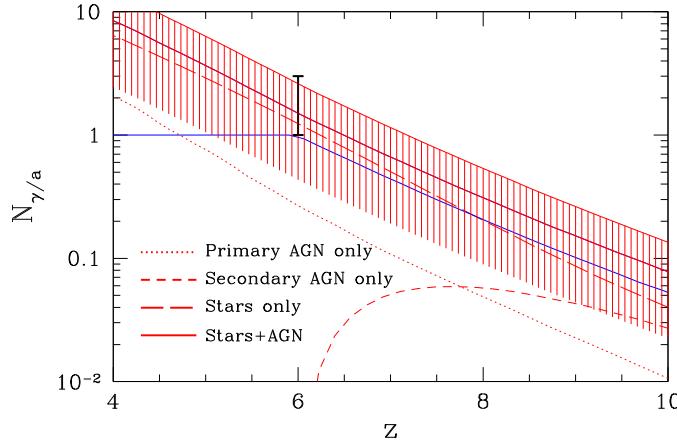


**Fig. 21.** Results from hydrodynamic simulations showing the contributions of galaxies of different magnitudes to reionization [426]. *Upper panel:* Redshift evolution of the total specific ionization rate (solid line). The short (long)-dashed line corresponds to galaxies detectable by *JWST* (*HST/WFC3*); dotted lines show the specific IGM recombination rate for different values of the clumping factor. *Bottom panel:* Fraction of ionizing photons coming from galaxies identified by *JWST* and in the *HUDF*. These calculations have assumed a fixed value of  $f_{\text{esc}} = 0.2$ . As shown, while the *HST* (using a magnitude limit of  $M_{\text{UV}} \sim -18$ ) can only detect (massive) galaxies providing about 40% of reionization photons, the *JWST* is critically required to peer deep enough ( $M_{\text{UV}} \sim -15$ ) to be able to detect the sources providing 80% of the total ionizing photon budget.

We start with the HMF and assign each galaxy a cosmological baryon fraction and a star formation efficiency that provides enough SN energy to unbind the rest of the gas, up to a maximum threshold which we take to be 0.03, as in Section 6.2. Further, we assume constant values for both  $f_{\text{esc}}$  and  $C$ . Allowing galaxies with halo masses as low as  $10^6 M_{\odot}$  to form stars, half of the LyC at  $z = 15$  (5) come from halos with  $M_h \lesssim 10^{7.8} (10^{9.6}) M_{\odot}$ ; larger halos will naturally provide a larger fraction of LyC photons if star formation is suppressed below the atomic cooling threshold. Complicating the calculation, we assume a *maximal* UV feedback scenario that completely suppresses the baryon fraction (i.e.  $f_b = 0$ ) for all galaxies with  $V_{\text{vir}} \lesssim 50 \text{ km s}^{-1}$ . This results an increase in the halo masses providing half of the LyC photons to  $M_h \lesssim 10^{9.6} (10^{10.4}) M_{\odot}$  at  $z = 15$  (5). As marked in the same plot, changing the model assumptions will naturally change these results: for example, assuming  $f_{\text{esc}}$  to increase (decrease) with the halo mass will result in a larger fractional contribution from higher (lower) mass halos. Further, assuming a constant star formation efficiency for all halos will result in a larger fractional contribution from lower masses. These simple calculations demonstrate that, for almost all reasonable physical scenarios, most LyC photons are provided by low-mass halos at high- $z$ . Higher mass halos, however, might become more important at later times due to an increase in the number density of the largest sources and UV suppression of star formation in low-mass halos [145,366,424,425].

This broad picture agrees with the majority of theoretical and observational works. Observationally, different groups have combined arguments involving the UV luminosity density contributed by galaxies of different masses [427], the steep faint-end slope of the LBG luminosity function (e.g. [428]),  $f_{\text{esc}}$  being higher for galaxies with bluer UV slopes [429], and the abundance and luminosity distribution of galaxies from the Hubble Ultra Deep Fields [355,358] to conclude that galaxies with  $M_{\text{UV}} \simeq -10$  to  $-15$  alone can reionize the IGM. This is accord with theoretical results that claim galaxies alone could have reionized the Universe [140,142,146,151,406,424,430–432]. Using roughly constant values of  $f_{\text{esc}}$  across the entire galaxy population, semi-analytic models [421,424,431], semi-numerical models [433,434], SPH simulations [368,426], combinations of SPH simulations and radiative transfer [140,151,364–366] and AMR simulations [361] all converge on low-mass ( $M_h \lesssim 10^{9.5} M_{\odot}$  and  $M_{\text{UV}} \gtrsim -17$ ) halos providing the bulk ( $\gtrsim 50\%$ ) of LyC photons at  $z \gtrsim 7$ . We show one such result, from SPH simulations, in Fig. 21 [426]. This figure clearly shows that the photon output required to maintain reionization increases with the IGM clumping factor. Furthermore, it shows that, with its magnitude limit of  $M_{\text{UV}} \sim -15$ , the *JWST* will be crucial in shedding light on the sources that provide roughly 80% of LyC photons.

However, the contribution from galaxies of different masses (or luminosities) sensitively depends on the details of the feedback prescriptions implemented and values used for  $f_{\text{esc}}$  and  $C$ . For example, while reionization finishes as early as  $z_{\text{re}} \sim 8$  for  $f_{\text{esc}} = 1$ , it gets delayed to  $z_{\text{re}} \sim 5.5$  using a lower  $f_{\text{esc}} = 0.2$  [422]. Using AMR simulations, where  $f_{\text{esc}}$  decreases from  $\sim 50\%$  for  $M_h \sim 10^{7-7.5} M_{\odot}$  halos to  $\sim 5\%$  for larger halos, [145,425] find that while low-mass halos dominate the reionization budget at early epochs, larger  $M_h \gtrsim 10^9 M_{\odot}$  halos can provide up to 50% of the total ionization budget at later epochs ( $z \lesssim 10$ ; see also [424]). On the other hand, the few models that either assume  $f_{\text{esc}}$  to scale with the star formation rate density (and hence with the halo mass) or quench star formation in any cell with  $T \gtrsim 2 \times 10^4 \text{ K}$  (resulting in a severe star formation suppression in  $M_h \lesssim 3 \times 10^9 M_{\odot}$  halos) naturally find much larger halos, with  $M_h \sim 10^{10-11} M_{\odot}$  ( $M_{\text{UV}} \gtrsim -18$ ), dominating the total reionization budget. Interestingly, coupling N-body simulations with RT, Iliev et al. [435] have shown that H I 21cm line statistics (e.g. the power spectra) can be used to distinguish between reionization driven solely by high-mass ( $\gtrsim 10^9 M_{\odot}$ ) halos as opposed to one where both high- and low-mass ( $10^{8-9} M_{\odot}$ ) halos contribute. These calculations have been extended



**Fig. 22.** Total ionizations-to-atom ratio as a function of redshift assuming quasars with large seeds with high spins from Volonteri and Gnedin [436]; the hatched region shows the estimated observational uncertainty associated with the model. Dotted, short-dashed, and long-dashed red lines show the contributions from primary and secondary ionization from quasars and from stars respectively. As shown, secondary ionizations from quasars can have a significant impact at  $z \gtrsim 8$ , with galaxies dominating the budget at lower redshifts.

by Dixon et al. [418] who show that, while the overall shape and skewness of the 21 cm signal are quite robust to the exact nature of source suppression, box size and resolution, higher order statistics (e.g. the kurtosis of the 21 cm power spectrum) could be used to distinguish between reionization purely driven by low-mass versus high-mass halos.

Quasars might also significantly contribute to reionization [356,436–438], specially at  $z \gtrsim 8$  if ionizations by secondary electrons are accounted for, with galaxies taking over as the dominant reionization sources at  $z \lesssim 6$  as shown in Fig. 22. The contribution of AGN to reionization has witnessed a resurgence after the recent claim of extremely high faint AGN number densities measured by Giallongo et al. [439] at  $z \gtrsim 4$ . It has been shown that if such a claim is confirmed, and this high comoving emissivity persists up to  $z \simeq 10$ , then AGN alone could drive reionization with little/no contribution from galaxies [437]. However, in addition to the fact that later works have not been able to corroborate these high AGN number densities, an AGN-dominated reionization scenario is in tension with current constraints on the thermal history of the IGM [440]. Further, AGN-driven reionization results in a slower evolution of the H $\alpha$  Ly $\alpha$  forest as compared to observations and is therefore ruled out at the  $2 - \sigma$  level; the data therefore favours a model where galaxies with  $f_{esc} \sim 10\%$  dominate in the initial reionization stages with AGN being relevant only at  $z \lesssim 6$  [438].

In terms of the imprint on reionization, Hassan et al. [432] have used a semi-numerical framework to show that AGN-dominated models produce larger ionized regions, reflected in the 21 cm power being twice at large at all scales. Further, using SPH simulations processed with RT Kakiuchi et al. [441] show that, although UV obscured quasars can lead to a broader ionization front as compared to galaxies and unobscured quasars, they can imprint a distinctive morphology on the HII regions only if their ionizing photon contribution is consistently larger than galaxies since the onset of reionization. Finally, [442] raise the point that the quasar impact depends on its duty cycle, leaving potentially observable imprints on the 21 cm signal in their direct environment.

Finally, we list a few sources going beyond metal-rich star formation and quasars:

(a) *PopIII stars*: Post-processing SPH simulations with RT, Paardekooper et al. [365] show that, although they could have started the process of reionization, given their short life-times and low numbers, PopIII stars contribute negligibly to reionization; similar results have been found by semi-analytic models [356] and AMR simulations [404]. These results are borne out by the findings of Maio et al. [291] who find that PopIII stars contribute  $\lesssim 10\%$  to the total star formation rate density at  $z \lesssim 15$  even for an “extreme” PopIII Salpeter IMF between  $0.1 - 100 M_\odot$ .

(b) *Runaway stars*: Contributing 50–90% of the total ionizing radiation escaping galaxies and enhancing  $f_{esc}$  values by as much as a factor of 7 in low-mass galaxies ( $M_{UV} = -12$  to  $-16$  at  $z = 10$ ), runaway stars could have had a significant contribution to reionization [374,375] although Ma et al. [362] find a much lower enhancement.

(b) *Decaying/annihilating DM*: While decaying/annihilating light dark matter particles (of mass 1–10 Mev) and sterile neutrinos (2–8 keV) can contribute to partial early reionization at  $z \lesssim 100$ , their contribution to the electron scattering optical depth is minimal ( $\tau \lesssim 0.01$ ; [443]) if the unresolved X-ray background emissivity constraints are to be respected (e.g. [444,445]). The main contribution of dark matter annihilation/decays lies in an earlier and more homogeneous heating of the IGM. This can leave detectable imprints, specially on the height and positions of the 21 cm peaks at large scales ( $k \sim 0.1 \text{ Mpc}^{-1}$ ; [446]), which can be used to constrain the dark matter–energy injection rates and lifetimes [447].

## 8. Emerging galaxy properties

Having described the underlying theory for early galaxy formation, we devote this section to a detailed comparison between theoretical results and observational data. This (iterative) process has been enormously successful, both, in interpreting the available measurements and improving theoretical models as now described.

### 8.1. High- $z$ galaxies: confronting observations and theory

The past decade has seen the “golden age” in the hunt for high- $z$  galaxies, made possible by observatories including the *HST*, the 8.2 m *Subaru*, the 8.2 m *VLT*, the 85 cm *Spitzer* and the 10 m *Keck*. Providing coverage from the optical through to mid infra-red wavelengths, these facilities have allowed unprecedented constraints on the physical properties of early galaxy populations. We start by briefly summarizing the two main approaches used to detect high- $z$  galaxies:

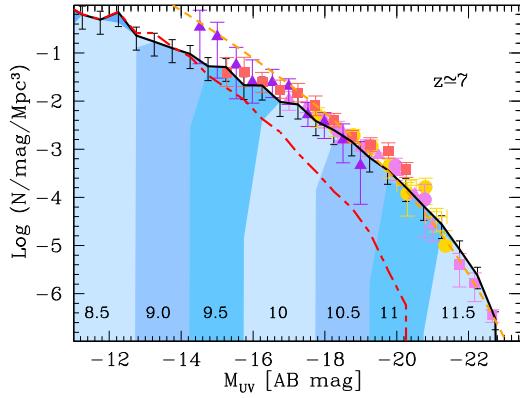
- The Lyman break technique to identify Lyman Break Galaxies (LBGs): First proposed by Meier [448], this approach relies on successfully identifying the 912 Å Lyman break caused by the absorption of LyC photons by neutral ISM gas. At  $z \gtrsim 5$ , the Lyman Break effectively shifts to 1216 Å, with photons of lower wavelengths being absorbed by neutral hydrogen in both the ISM and the intervening IGM. This “Gunn–Peterson” absorption trough [449] in high- $z$  quasar spectra has been a crucial input in inferring the end redshift of reionization (e.g. [272,450]). As shown in the following sections, LBG selections have yielded thousands of candidates with  $z$  as high as  $z \simeq 11$ , thanks to two key instruments – the Advanced Camera for Surveys (ACS) and the Wide Field Camera3 (WFC3) onboard the *HST* [39,40,321,417,451–465], shedding light on the hitherto hidden early galaxies that had formed in the first billion years.
- The Ly $\alpha$  line to identify Lyman Alpha Emitters (LAEs): As detailed in Section 7.1, most hydrogen ionizing photons are absorbed by H I in the ISM of galaxies, with only a small fraction ( $f_{\text{esc}}$ ) escaping into the IGM. Given the large densities associated with the ISM, this ionized gas recombines on timescales of a few Myr (Eq. (64)) giving rise to the Ly $\alpha$  recombination line at 1216 Å in the galaxy rest-frame. Using the case-B recombination approximation, where the recombined H I is optically thick and there is a 2/3 probability of a recombination leading to a Ly $\alpha$  photon [466], the intrinsic Ly $\alpha$  luminosity can be calculated as  $L_{\alpha} = [2/3](1 - f_{\text{esc}})\dot{N}_s h v_{\alpha}$  [erg s $^{-1}$ ]. As noted in Section 6.3,  $\dot{N}_s$  is the number of H I ionizing photons produced by the stellar population and  $v_{\alpha}$  is the Ly $\alpha$  frequency. Despite the elegant proposal of using the Ly $\alpha$  line as a beacon for high- $z$  galaxies as early as 1967 [277], the first LAEs were only successfully identified three decades after the publication of this seminal paper [467–469]. Since these early detections, narrow band techniques have identified hundreds of LAEs up to  $z \simeq 8$  [470–482]. However, a key drawback of the Ly $\alpha$  line is its extremely high optical depth against neutral hydrogen [483]:

$$\tau_{\alpha} = \frac{\pi e^2 f \lambda_{\alpha}}{m_e c h(z)} \chi_{\text{HI}}(z) n_H(z) \simeq 1.5 \times 10^5 \chi_{\text{HI}}(z) h^{-1} \Omega_M^{-1/2} \left( \frac{\Omega_b h^2}{0.019} \right) \left( \frac{1+z}{8} \right)^{3/2}, \quad (68)$$

where  $f$  is the oscillator strength,  $\lambda_{\alpha}$  is the Ly $\alpha$  wavelength and  $e$  and  $m_e$  represent the electron charge and mass, respectively. As seen from this equation,  $\tau_{\alpha} \gg 1$  for a galaxy embedded in an IGM with  $\chi_{\text{HI}} = 0.1$  resulting in a severe attenuation of the Ly $\alpha$  luminosity. Such a galaxy would probably *not* be visible as a LAE. Statistically, this implies a decreasing fraction of star forming galaxies being visible in the Ly $\alpha$  with increasing  $\chi_{\text{HI}}$ . An indication of this effect is already available from the apparent Ly $\alpha$  LF (number density of galaxies as a function of the Ly $\alpha$  luminosity) which, showing effectively no evolution between  $z \simeq 3.1–5.7$  [472], shows a steep drop at higher  $z$  with the knee luminosity ( $L_*$ ) at  $z = 6.6$  being about 50% of the value at  $z = 5.7$  [475,479]. This decline has also been confirmed by other groups [473,484–486]. The increasing dearth of LAEs at high- $z$  has prompted searches for other emission lines which include CIII] $\lambda$  1909 Å [487], [OIII]+H $\beta$  [41,488,489] and CIV] $\lambda$  1548 Å [490,491]. Interestingly, such strong line emitters have been sign-posts for Ly $\alpha$  emission: for example Roberts-Borsani et al. [489] have observed strong line emitters at  $z \simeq 7–9$ , the follow up spectroscopy of which shows all (four) sources to have a Ly $\alpha$  line. Suggesting up to 50% of the galaxy population being in an extreme emission phase as a result of their young stellar populations and intense ionizing fields, these methods provide an excellent opportunity, both for direct detections using the Near-Infrared Spectrograph (NIRSpec) onboard the *JWST* as well as targets for Ly $\alpha$  emission.

In terms of the relation between LAEs and LBGs, there is consensus that these are both drawn from the same underlying population (e.g. [492,493]); the fraction of time a galaxy is visible as an LAE or LBG depends on its underlying stellar mass [494]. While the decrease in the fraction of LBGs showing Ly $\alpha$  emission (with increasing redshift) has been attributed to an increasingly neutral IGM (e.g. [495,496]), this interpretation is complicated by a number of caveats. These include the EW criterion used to select LAEs (e.g. [472]), the degeneracy between the (redshift-dependent) effects of ISM dust and IGM attenuation on Ly $\alpha$  photons [493], a three dimensional degeneracy between  $f_{\text{esc}}$ , ISM and IGM attenuation as well as the ionization structure encountered along a given line-of-sight [497].

We also briefly mention other techniques used to hunt for high- $z$  galaxies that include (for a full discussion see e.g. [498,499]): (a) *The Balmer break technique*: which relies on identifying the Balmer break at 3646 Å in the rest-frame of the galaxy. However, given the fact that high- $z$  stellar populations are generally younger than a Gyr and that the Balmer break only causes a factor of 2 drop in the flux, this technique has not yielded any galaxies not already identified by the LAE and LBG techniques [500]; (b) *Super Massive Black Holes*: the presence of SMBHs has helped identify QSOs at  $z$  as high as  $z \simeq 7.5$  [501,502]. Observations at  $z \sim 6–7$  have firmly established that the rarest, most massive ( $10^{12–13} M_{\odot}$ ) and brightest galaxies ( $M_{\text{UV}} \lesssim -23.5$ ) host SMBHs with mass  $\sim 10^{7–8} M_{\odot}$  solar masses [503–506]. While cross-correlating CANDELS and *Chandra* X-ray data, Giallongo et al. [439] have recently detected AGN activity in galaxies that are about 4 magnitudes fainter, with  $M_{\text{UV}} \sim -19$ , this result has not been confirmed by later observations with the Hyper Suprime Camera on *Subaru* that, instead, roughly find a 100% galaxy fraction down to  $M_{\text{UV}} \simeq -22$  at  $z \simeq 4–6$  (e.g. [506]); (c) *Gamma Ray Bursts (GRBs)*: arising from the death throes of massive stars [507], their extreme luminosity has resulted in GRBs being observed up to



**Fig. 23.** The LBG UV LF at  $z \simeq 7$  [148]. The solid black line shows the results using the fiducial model, i.e. using the maximal star formation efficiency-halo mass relation derived in Section 6.2. This model also includes the impact of mergers, smooth accretion from the IGM and SN-powered gas ejection on both the gas and stellar mass build-up of high- $z$  galaxies; the black error bars show Poissonian errors arising from the luminosity dispersion in each bin. The dashed red line shows that the UV LF would have been severely under-estimated (with increasing brightness) had we not considered the gas brought in by mergers. The orange line shows the results obtained by multiplying the halo mass function (where each galaxy is assigned a cosmological gas-to-DM ratio) with a constant star-formation efficiency  $f_* = 0.9\%$ . Finally, the numbers within the shaded contours indicate the host halo mass. Points show observational results from: Oesch et al. [511, filled violet circles], Bouwens et al. [456, filled yellow circles], Castellano et al. [464, empty yellow squares], McLure et al. [461, filled violet triangles], Bowler et al. [457, filled violet squares], Atek et al. [40, filled red squares] and Livermore et al. [321, filled purple triangles].

$z \simeq 8.4$  with *Swift* [508,509]. The low surface brightness of their host galaxies, however, has limited the study of their physical properties. In what follows, we will solely focus on the LAE and LBG techniques that have, so far, been the most successful at collecting statistically significant samples of  $z \gtrsim 5$  galaxies.

## 8.2. The high- $z$ UV luminosity function

Describing the number density of galaxies as a function of their observed UV luminosity,<sup>5</sup> the UV LF and its  $z$ -evolution is one of the most robust observables available for high- $z$  galaxies. In the simplest form, the UV LF can be constructed by assigning each galaxy a cosmological ratio of baryons to DM ( $\Omega_b/\Omega_c$ ) and multiplying this with a (constant) star-formation efficiency value, say  $f_*$ . This (no feedback) model naturally results in the UV LF tracing the same shape as the HMF as shown (by the dashed line) in Fig. 23. This UV LF can therefore be parameterized using the Schechter function [510] to obtain the number of galaxies as a function of luminosity ( $dn/dL$ ) as:

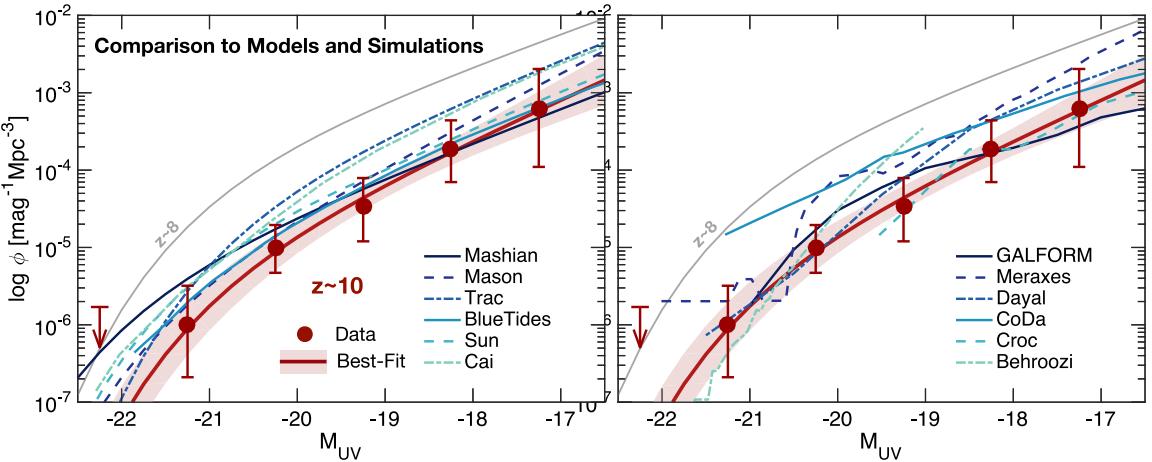
$$\frac{dn}{dL} = \phi(L) = \frac{\phi_*}{L_*} \left( \frac{L}{L_*} \right)^\alpha \exp^{-\left( \frac{L}{L_*} \right)}. \quad (69)$$

This UV LF is fully defined once the three free parameters –  $\phi_*$ ,  $L_*$  and  $\alpha$ , signifying the characteristic number density and luminosity, and the faint-end slope, respectively – are specified. As expected, this function is a power-law at the low-luminosity end (at  $L < L_*$ ) that drops off exponentially above the characteristic (or “knee”) luminosity. Given the general notation of expressing most high- $z$  observables in terms of the UV magnitude, substituting  $\phi(M_{\text{UV}}) dM_{\text{UV}} = \phi(L) d(-L)$  and  $M_{\text{UV}} - M_{\text{UV}*} = -2.5 \log(L/L_*)$ , where  $M_{\text{UV}*}$  is the characteristic (or knee) magnitude, the Schechter function can be written as

$$\frac{dn}{dM_{\text{UV}}} = \phi(M_{\text{UV}}) = \frac{\ln 10}{2.5} \phi_* \left( 10^{0.4(M_{\text{UV}*} - M_{\text{UV}})} \right)^{\alpha+1} \exp^{-10^{0.4(M_{\text{UV}*} - M_{\text{UV}})}}. \quad (70)$$

Starting with observations, a combination of space-based surveys (e.g. the Hubble Ultra Deep Field and the eXtreme Deep Field), carried out with the WFC3 onboard the HST (e.g. [456,459,461,464,511,512]), the Hubble CLASH lensing surveys (e.g. [40,321,463,513]) and ground-based observations (e.g. [458,462,514]) have allowed mapping out the high- $z$  LBG UV LF over 8 magnitudes from  $M_{\text{UV}} \sim -14.5$  to  $M_{\text{UV}} \sim -22.6$  at  $z \simeq 7$  as shown in Fig. 23. Observations indicate the bright-end at  $z \simeq 7$  to be better fit by a double power-law that steepens to the Schechter function by  $z \simeq 5$  which could possibly arise as a result of increasing dust extinction, the onset of AGN feedback or even lensing magnification biases [458,506]. Since the three UV LF free parameters are degenerate when inferred by fitting to observations, it is generally agreed that the faint-end steepens with  $z$  ([512] and references therein) which has critical implications for reionization (see Section 7.4). Disentangling whether the  $z$ -evolution of the UV LF is driven by an evolution in  $\phi_*$  (density evolution; [459]) or  $L_*$  (luminosity evolution; [456]) has remained more challenging. However, theoretical works [515] show that this evolution depends on the luminosity

<sup>5</sup> The intrinsic UV luminosity has an initial value (at  $t_0 = 2$  Myrs) of  $L_{\text{UV}} = 10^{33.07} \text{ erg s}^{-1} \text{Å}^{-1} M_\odot^{-1}$  for a  $0.1 - 100 M_\odot$  Salpeter IMF and  $Z = 0.05 Z_\odot$  using *Starburst99* [323,324]. It undergoes a time evolution as  $L_{\text{UV}}(t) = L_{\text{UV}}(0) - 1.33 \log(t/t_0) + 0.462$  [325].



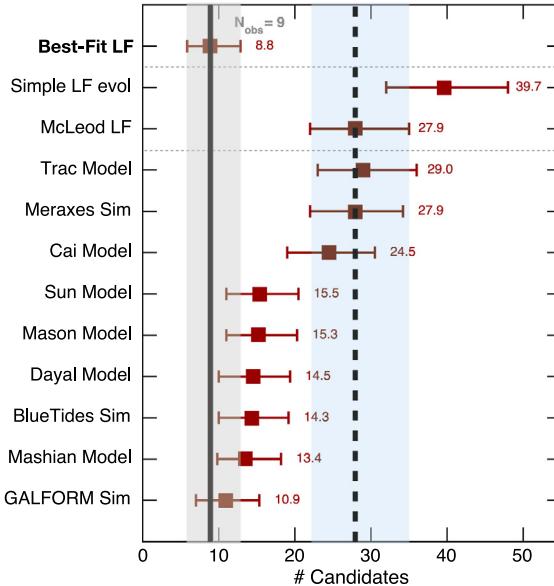
**Fig. 24.** A comparison of the UV LF at  $z \sim 10$  from observations (filled circles) compared with a range of theoretical models (for details see [518]). Despite the diversity of physical approaches and processes implemented, all models predict the UV LFs that, do not differ by more than a factor of 3–4 over the luminosity range and, evolve significantly from  $z \sim 8$  (light grey line) to  $z \sim 10$ .

range probed: the steady brightening of the bright end of the LF is driven primarily by genuine physical luminosity evolution and arises due to a fairly steady increase in the UV luminosity (and hence star-formation rates) in the most massive LBGs; at the faint luminosity end, the evolution of the UV LF involves a mix of positive and negative luminosity evolution (as low-mass galaxies temporarily brighten then fade) coupled with both positive and negative density evolution (as new low-mass galaxies form, and other low-mass galaxies are consumed by merging). We note that, in addition to incompleteness corrections and cosmic variance, a number of observational biases, including magnification uncertainties (e.g. [321,512]) and the estimated size-luminosity relation ([346,516,517] and see also Section 8.7), can have a significant impact on the observed number counts at the faint-end of the UV LF.

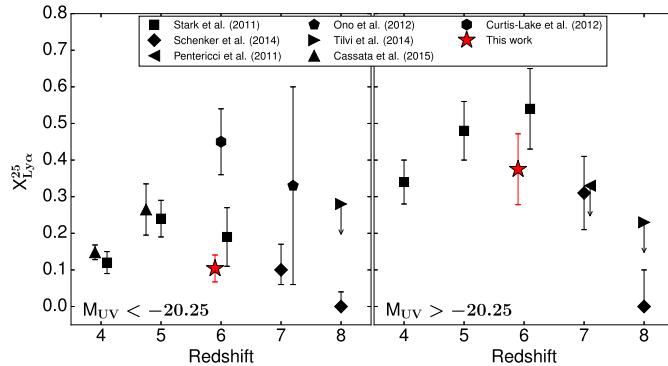
Reproducing the bright-end of the UV LF at  $z \sim 5 – 10$  requires an increasing star formation efficiency (assuming a cosmological baryon fraction). This results in galaxies of the same UV luminosity residing in halos that are twice as massive at  $z \simeq 5$  as at  $z \simeq 8$  leading to a steepening in the predicted faint-end slope, from  $\alpha_{\text{HMF}} = -2.11 \rightarrow -2.32$  from  $z = 5 \rightarrow 8$  (see e.g. [148,499,515]). Further, as noted in Sections 4.1.4 and 6.2, SN feedback is critical for reproducing the faint-end of the UV LF. Using the feedback model introduced in Section 6.2, where  $f_*^{\text{eff}} = \min[f_*, f_*^{\text{ej}}]$ , we find that the observed evolving UV LF from  $z \sim 5 – 10$  is well fit by a model (for details see [148]) where galaxies have a maximum star formation efficiency value of  $f_* \sim 3\%$  (at any given time) and where roughly 10% of the total SN energy coupling to the gas (shown by the solid line in Fig. 23). We also revisit the question of whether galaxies on the UV LF gain their gas through mergers or gas accreted along with dark matter (smooth accretion) from the IGM by switching off gas being brought in by mergers in any part of the merger tree (short-long dashed line in Fig. 23). As clearly shown, gas brought in by mergers is not very important at the faint end since the tiny progenitors of these galaxies are feedback limited in the fiducial model, resulting in largely dry mergers. However, the progenitors of increasingly luminous galaxies are critical to building up the large gas reservoirs that power star formation. Indeed, the LF at the bright end drops to 10% of the fiducial model value ( $M_{\text{UV}}$  decreases by  $\approx 2$  magnitudes) if the gas brought in by mergers is not taken into account. However, we caution the reader that smoothly accreted gas in a progenitor will be considered as “accreted” gas for the successor halo — these numbers therefore represent an upper limit on the importance of accretion.

On the other hand, calculations of the stellar mass build-up show most of it is assembled in the largest progenitors of any galaxy [515]: this results in a global picture of the progenitors of the most massive ( $M_h \gtrsim 10^{11.5} M_\odot$ ) high- $z$  galaxies being “conveyor belts” of gas that is turned into stars in the deepest potential wells of their most massive progenitors.

Given its import, the UV LF has been modelled in a number of ways, ranging from abundance matching to semi-analytic to hydrodynamic simulations, an illustrative example of which, at  $z = 10$ , is shown in Fig. 24. Interestingly, despite the wide variety of approaches followed and parameters used, most of these models are in reasonable agreement with the data. This is probably driven by a number of physical effects that can have a degenerate effect on the UV LF including, but not limited to: the impact of the SN and UVB feedback in decreasing the gas content of low mass dark matter halos, the impact of dust in decreasing the observed luminosity from the most massive halos and the propagation of gas mass through the galaxy assembly history. As noted, a number of observational biases (i.e. the size correction, lensing magnification correction, luminosity limits probed) can impact the faint-end slope and need to be constrained in more detail. These complications are clearly reflected in the fact that both observations and theoretical results expect anywhere between 10–40  $z \simeq 10$  candidate LBGs based, mostly, on the value and luminosity extent of the faint-end slope, as shown in Fig. 25. However, precisely as a result of these degeneracies, both the faint end slope and the cut-off mass/luminosity to which galaxies can keep forming stars remain a matter of debate [148,149,433,519].

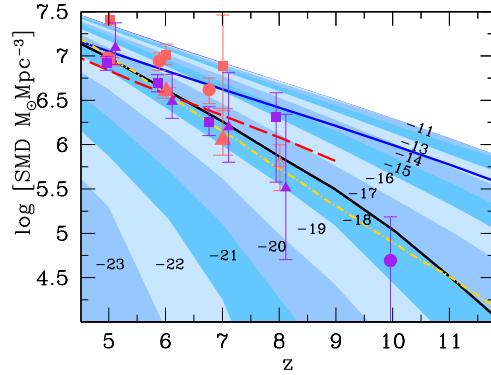


**Fig. 25.** A comparison between the expected number of galaxy candidates at  $z \sim 10$  from theory and observations [518]. Vertical lines show observational results from Oesch et al. [518, solid line] and McLeod et al. [513, dashed line], with the shaded regions showing the  $1-\sigma$  errors. Extrapolating the  $z \sim 4 - 8$  LF evolutionary trends to  $z \sim 10$  yields the “simple LF evol” model which is expected to be an upper limit on the number of  $z \sim 10$  galaxies [518]. All the other points show theoretical results. Despite using a wide range of modelling approaches and physical ingredients, essentially all models are in agreement with the range of 9 – 35 galaxies expected from observations.



**Fig. 26.** A compilation of the fraction of LBGs with  $\text{EW}(\text{Ly}\alpha) > 25 \text{ \AA}$  ( $X_{\text{Ly}\alpha}^{25}$ ) at  $4 \leq z \leq 8$  [528]. The left and right panels show results for the brightest ( $M_{\text{UV}} < -20.25$ ) and faintest ( $M_{\text{UV}} > -20.25$ ) galaxies. This includes results from Stark et al. [495], Schenker et al. [529], Pentericci et al. [530], Ono et al. [531], Tilvi et al. [532], Cassata et al. [533] and Curtis-Lake et al. [534]; the red stars show results from De Barros et al. [528]. As shown, all works converge on a steep drop in the fraction of faint LBGs showing Ly $\alpha$  emission (with an EW larger than 25  $\text{\AA}$ ) at  $z \gtrsim 6.5$  with a larger scatter seen for rare, bright sources.

Moving on, the UV LF can also be constructed for LAEs (see e.g. Fig. 14; [498]) that, as already noted, shows a steep decline at  $z \gtrsim 6$  when compared to the LBG UV LF. This is also reflected in a steep decrease in the fraction of (faint;  $M_{\text{UV}} > -20.25$ ) LBGs detected through their Ly $\alpha$  emission as shown in Fig. 26; the observational trends, showing a large scatter, are more unclear for the brightest galaxies. While it might be tempting to attribute this solely to an increase in the average value of  $\chi_{\text{HI}}$  with redshift, the interpretation is actually quite complex and could arise from a combination of one or more physical effects including: an evolution of the underlying halo mass function [520,521], evolution of galaxy physical properties (star formation rates and dust content) with redshift [493,520–522], the number density evolution of the Lyman Limit absorption systems [523] and even the different CGM of massive and low mass halos [524]. Indeed, it has also been cautioned that simple colour-cuts, in EW and UV magnitude, that do not mimic the actual observational strategies used to select LAEs and LBGs, might also bias theoretical results [525] and that integrating down to extremely faint magnitudes might make most LBGs appear as LAEs [493,526,527]. The advent of the JWST will be crucial in shedding more light on the absence (or presence) of Ly $\alpha$  emission in LBGs.



**Fig. 27.** Redshift evolution of the cosmic stellar mass density. Points show the observed estimates, using a multitude of mass and magnitude cuts and both considering and ignoring the impact of nebular emission lines, from: González et al. [536], red filled circles; Labb   et al. [535], red empty triangles; Stark et al. [537], purple filled squares; Oesch et al. [538], purple filled circles; Duncan et al. [539], red filled squares; Grazian et al. [540], red filled triangles; and Song et al. [541], purple filled triangles; the error bars show the  $1 - \sigma$  error. The solid black line shows DELPHI results integrated down to a UV magnitude limit of  $M_{\text{UV}} = -18$  [148] with the shaded regions showing the contribution of the different magnitude bins marked to SMD through cosmic time; the solid blue line shows the magnitude limits at which galaxies provide 50% to the total SMD at that redshift. The dot-dashed yellow line and the dashed red line show results from the DRAGONS project [147] and the EAGLE simulations [133], respectively.

### 8.3. Stellar mass density

The stellar mass,  $M_*$ , is one of the most fundamental galaxy properties, encapsulating information about its entire star-formation history. However, achieving accurate stellar mass estimates of early galaxies from broad-band data has been difficult because it depends on the assumptions made regarding the IMF, the strength of nebular emission, the star formation history, stellar metallicity and dust; the latter three parameters are degenerate, adding to the complexity of the problem. Indeed, accounting for the presence of nebular emission lines can result in a downward revision of the stellar ages and masses by as much as a factor of 3 [327,535]. Although properly constraining  $M_*$  ideally requires rest-frame near infra-red data, which will be provided by future instruments such as the Mid-Infrared instrument (MIRI) onboard the *JWST*, broad-band *HST* and *Spitzer* data are already being used to infer the stellar mass density (SMD; stellar mass per unit volume) of galaxies at redshifts as high as  $z \simeq 10$ , as shown in Fig. 27.

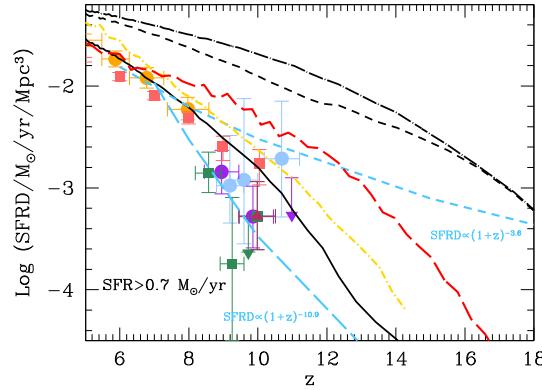
In brief, the observed SMD values inferred from LBGs at  $z \simeq 5 - 10$  have been collected using a multitude of mass and luminosity limits and making different physical assumptions regarding the presence of nebular emission lines: e.g. while a number of groups have computed the SMD down to a limiting magnitude of  $M_{\text{UV}} = -18$  [535–538], some account for the impact of nebular emission [535,537] while the others do not. On the other hand, a number of works compute the SMD by integrating the Schechter (stellar mass) function between  $M_* = 10^8 - 10^{13} M_\odot$  [539–541]. Interestingly, all of these results are in accord, within error bars, perhaps demonstrating the lack of leverage in the SMD inferred using these different means. As per these measurements, the SMD-z relation has a slope that scales as  $(1 + z)^{-0.36}$  for  $z \gtrsim 4$ , although the error bars and paucity of data preclude any strong conclusions at  $z \gtrsim 8$ .

We compare these observations to the results from three theoretical models, one each demonstrating the semi-analytic (DELPHI; [148]), semi-numerical (DRAGONS; [147]) and hydrodynamic (EAGLE; [133]) approaches, as shown in Fig. 27. As shown, using a magnitude cut of  $M_{\text{UV}} \lesssim -17.7$ , within error bars, the results from all three models are in reasonable agreement with each other as well the data despite the different physical ingredients implemented, the baselines used to calibrate model parameters and even the numerical methods used.

Conducting a census of the total stellar mass we find that, as a result of their enormous numbers, small, faint galaxies contain most of the stellar mass in the Universe at  $5 \leq z \leq 12$  as shown in Fig. 27. Indeed, galaxies brighter than current observational limits ( $M_{\text{UV}} \leq -18$ ) contain about 50% of the total stellar mass at  $z \simeq 5$ . This value then steadily decreases with redshift such that observed galaxies contain a quarter of the total stellar mass at  $z \simeq 6.5$  and only 10% of the total stellar mass at  $z \simeq 9$  (i.e. at redshifts  $z \geq 6.5$ , most of the stellar mass of the Universe is locked up in galaxies too small to have been detected so far). The next generation of space instruments such as the *JWST*, along with future developments in the use of gravitational lensing (e.g. in the Frontier Fields), which might extend stellar mass estimates down to  $M_{\text{UV}} \sim -14$ , will play an important role in revealing about half (a quarter) of the stellar mass in the Universe up to  $z \simeq 8$  ( $z \simeq 9.5$ ).

### 8.4. Cosmic star formation rate density

The cosmic star formation rate density (SFRD), inferred from the UV luminosity, is an important indicator of active star formation over the past few Myrs, given that UV luminosity decays as roughly  $(t/2 \text{ Myr})^{-1.3}$ . A number of observational works have focused on estimating the SFRD using sources with  $M_{\text{UV}} \lesssim -17.7$  from the *HST* Ultra deep (HUDF) surveys [451, 456,461,511,542] which have now been pushed to even fainter magnitudes using the CLASH lensing surveys [513,543–545];



**Fig. 28.** The evolution of the cosmic star formation rate density (SFRD) above a star-formation limit of  $0.7 M_{\odot} \text{ yr}^{-1}$  corresponding to  $M_{\text{UV}} \sim -17.7$ . The points show data collected by Bouwens et al. [453,547], orange filled circles, the CLASH lensing surveys (light blue filled circles; [543–545]), Ellis et al. [542], purple filled circles], McLeod et al. [513, red filled squares], Oesch et al. [451, green filled squares] and Oesch et al. [538, maroon filled triangle]. The downward facing arrows showing upper limits are from Ellis et al. [542, purple] and Oesch et al. [451, green]. The black lines shows DELPHI [148] results integrated down to a UV magnitude limit of  $M_{\text{UV}} = -17.7$  (solid line),  $M_{\text{UV}} = -15$  (dashed line) and for all galaxies (dot-dashed line). The dot-dashed gold line and the dashed red line show results from the DRAGONS project [147] and the EAGLE simulations [133], respectively. We also show the SFRD- $z$  trend inferred from low- $z$  galaxies which evolves as  $\propto (1+z)^{-3.6}$  (short-dashed blue line) with  $z > 8$  LBGs showing a much steeper fall-off  $\propto (1+z)^{-10.9}$  (long-dashed blue line) [538]. Although showing different slopes for the  $z$ -evolution of the SFRD, as of now, all three models (DELPHI, DRAGONS and EAGLE) are in accord with the observations within error bars.

these results are shown in Fig. 28. Essentially, observations integrate the UV LF to the chosen magnitude limit (e.g. a limit of  $M_{\text{UV}} = -17.7$  roughly corresponding to a SFR rate of  $0.7 M_{\odot} \text{ yr}^{-1}$ ). This integrated UV luminosity density is converted into a SFRD, generally, using the conversion factor from Madau et al. [546].<sup>6</sup>

It must be noted that inferring the UV luminosity density has been a herculean task at these high- $z$  due to a number of complexities including: dust corrections, the contamination from emission line galaxies, from local (L, M, T or Y) dwarfs and/or AGN, the impact of cosmic variance given the small fields and the uncertainties associated with the adopted lensing magnifications. Despite these complications, the data inferred from all fields is in excellent agreement as shown in the same figure. Given the paucity of statistics, however, the  $z$ -evolution of the SFRD has remained a matter of debate, with observational works finding a much steeper slope of  $\text{SFRD} \propto (1+z)^{-10.9 \pm 2.5}$  at  $z \gtrsim 8$  (e.g. [538]) as compared to the shallower  $\text{SFRD} \propto (1+z)^{-3.6}$  inferred for lower redshift data (e.g. [428,453]). Of course, it must be noted that the underlying metallicity, IMF, lack of binary stars and constancy of the SFR for about a 100 Myrs prior to observations are all implicit assumptions in this conversion. For example, this conversion would face an upward revision by a factor of about 2 if these galaxies had particularly young ages (e.g. [456]).

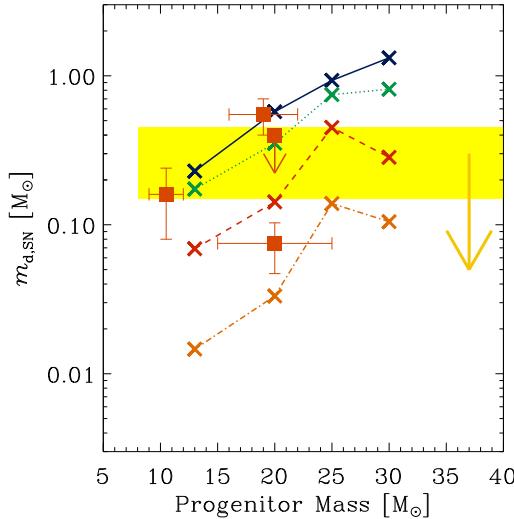
We then compare these observations to theoretical models (DELPHI, DRAGONS and EAGLE) finding that these results, whilst agreeing with observations at  $z \lesssim 8$ , do not show the steep drop seen at higher- $z$ . Interestingly, while the results from the semi-analytic (DELPHI; [148]) and semi-numerical (DRAGONS; [147]) models are in excellent agreement with each other, possibly as a result of the similar physics implemented, the hydrodynamical simulation (EAGLE; [133]) results show a much shallower slope. Forthcoming data, with e.g. the JWST and E-ELT, will be instrumental in pinning down the observations and disentangling the impacts of the observational biases from the physical effects (e.g. gas accretion, mergers, the impact of SN and UV feedback) driving the star formation in these early systems.

### 8.5. Dust content

Dust plays a critical role, both, in shaping the physics as well as visibility of early galaxies. For example, dust surfaces offer the main production sites for H<sub>2</sub> formation (e.g. [548]), inducing the formation of molecular clouds and enhancing star formation activity [549,550]. Dust cooling also induces fragmentation of molecular clouds (e.g. [551]) thereby shaping the stellar IMF [252]. Finally, dust grains absorb UV light, which is re-emitted in the Infra-red bands (see e.g. [552]), affecting both the observed luminosity as well as the spectral slopes of early galaxies. Dust can be produced by three key sources in the high- $z$  universe:

(i) *Core Collapse Supernovae (SNII)*: exploding on timescales of less than 28 Myr [553], stars with mass  $M_s > 8M_{\odot}$  end their lives as “Core Collapse Supernovae”. The classical nucleation theory [554] provides the backbone for the dust produced by such sources: solid materials can condense out of vapour in the supersaturated state once the partial pressure of the given species exceeds the vapour pressure in the condensed phase. This well-defined condensation barrier naturally results in clusters of a “critical” size which then grow into macroscopic dust grains by the accretion of material. A number of theoretical

<sup>6</sup> Where a SFR of  $1 M_{\odot} \text{ yr}^{-1} = 1.4 \times 10^{-28} L_{\text{UV}} [\text{erg s}^{-1} \text{ Hz}^{-1}]$  for a Salpeter IMF between  $0.1 - 100 M_{\odot}$  using STARBURST99.



**Fig. 29.** The dust mass formed in a SN as a function of the progenitor mass, in the range of  $8 - 40M_{\odot}$ , at the zero age main sequence assuming all the dust mass in the ALMA-detected  $z \simeq 6.6$  LAE – “Himiko” – is produced by SN [568]. As shown, this assumption results in an upper limit of  $m_d = 0.15 - 0.45M_{\odot}$  of dust produced per SN, for different dust species, shown by the shaded area. Crosses connected by the solid, dotted, dashed and dot-dashed lines show theoretical predictions of [570] for hydrogen number densities corresponding to  $n_H = 0, 0.1, 10 \text{ cm}^{-3}$  in the ambient medium. The filled squares with error bars show observed dust masses detected by *Herschel* for nearby SN remnants [562,563,571–573].

models have used this theory to predict that silicate and carbon grains form in the SN ejecta, a few hundred days after the explosion, with condensation efficiencies in the range 0.1–3, obtaining  $0.1 - 1M_{\odot}$  of dust per supernova for progenitors in the range of  $12 - 40M_{\odot}$  (e.g. [555]). While there is now clear observational evidence for dust produced by supernovae [556–559], the dust yields remain a matter of debate. This is because supernovae observations, although limited to only a few objects, show dust masses of the order of  $10^{-5} - 10^{-3}M_{\odot}$  hinting at condensation efficiencies that are about two orders of magnitude lower than theoretical predictions. On the other hand, observations of young galactic SN remnants with *Spitzer* show freshly formed dust masses  $\sim 0.015 - 0.05M_{\odot}$  [560,561] with *Herschel* far-infrared observations yielding estimates as high as  $0.4 - 0.7M_{\odot}$  for SN1987A [562] with follow-up ALMA observations confirming this to be the newly formed dust component [563]. One possible means of reconciling these apparently divergent results is to account for partial destruction of newly formed dust in the SN reverse shock, with dust destruction being more efficient for small grains [564] and with increasing ISM densities [550,564].

(ii) *Asymptotic Giant (AGB) Branch stars:* All stars with initial masses in the range of  $0.8 - 8 M_{\odot}$  on the main sequence evolve through the thermally pulsating AGB phase, losing mass at a very high rate, before becoming white dwarfs. Dust can form, by condensation of solid particles in the outflowing gas, in the cool, dense atmospheres during this stage, leading to optically thick circum-stellar dust shells. The dust mixtures formed depend on the relative abundances of carbon and oxygen (C/O) of the ejected matter. Refining the models for dust production from AGBs as a function of their initial stellar mass and metallicity, Zhukovska et al. [565] predict that while silicate dust is mostly produced by oxygen-rich stars (M type with C/O < 1 and S type with C/O ≈ 1) with an initial mass  $M_s < 1.5M_{\odot}$  and  $M_s > 4M_{\odot}$ , carbon and SiC dust production is, instead, dominated by stars with an initial mass in the range of  $1.5 - 4 M_{\odot}$  during the carbon-rich phase (called C type with C/O > 1); iron dust is produced in all AGB stars. Interestingly, carbon dust of mass  $10^{-5} - 10^{-2}M_{\odot}$  can be produced at all metallicities. However, comparable amounts of silicate dust can only be produced at  $Z \sim Z_{\odot}$  requiring the presence of sufficient Si, O, Mg and Fe, which have to be gathered from the pre-enriched ISM, given they cannot be produced by the star itself [566]. Given their evolutionary timescales of about a Gyr before the onset of dust production, AGB stars were thought to play a negligible role in the dust enrichment of early galaxies [567]. However, it has now been shown that while both SN and AGBs can contribute to the dust enrichment in the early Universe after several hundred Myrs of star formation [566], high mass AGB stars ( $M_s > 3M_{\odot}$ ) can be the dominant dust producers only if each SN produces  $\lesssim 3 \times 10^{-3}M_{\odot}$  of dust [559]. These arguments have been applied to LAEs given their predominantly young populations – using the ALMA (1.2 mm) upper limit of  $52.1 \mu\text{Jy}$  for Himiko, an exceptionally bright  $z \simeq 6.6$  LAE, Hirashita et al. [568] (and see also Hirashita et al. [569]) infer an each SN to produce a dust mass  $m_d < 0.15 - 0.45M_{\odot}$  as shown in Fig. 29.

(iii) *Grain growth in the ISM:* Finally, dust mass can also grow by the accretion of gas-phase metals onto pre-existing dust grains in cold, dense phase of the ISM [574,575]. One of the key parameters determining such dust growth is the metallicity. To quantify this, Asano et al. [576] have introduced the concept of a “critical metallicity” at which the contribution of AGB stars and SN equal that of the dust mass growth in the ISM.

In the simplest approximation, the time evolution of the dust mass,  $dM_d/dt$ , can be calculated as (e.g. [566,576–579])

$$\frac{dM_d}{dt} = \dot{Y}_d + \dot{G}_{ac} + \dot{I}_d - \mathcal{D}_{ISM}\psi(t) - \dot{D}_{SN} - \dot{O}_d. \quad (71)$$

Here, the rate of increase of dust mass depends on the rate at which dust produced by SN and AGB stars is injected into the ISM ( $\dot{Y}_d$ ), the rate at which the dust mass grows in the ISM ( $\dot{G}_{ac}$ ) and the rate of dust infall from the IGM ( $\dot{I}_d$ ). These are countered by a number of effects that, instead, decrease the dust mass:  $D_{ISM}\psi(t)$  expresses the rate of dust abstraction (assimilation) into further star formation where  $D_{ISM}$  is the ISM dust-to-gas ratio and  $\psi(t)$  is the star formation rate. As noted, dust can also be destroyed by SN shocks in the diffuse ISM ( $\dot{D}_{SN}$ ) and assuming a homogeneous mixture of ISM gas and dust, dust can be lost in outflows of gas into the IGM ( $\dot{O}_d$ ). We now explain the key terms (1,2 and 6 on the RHS) in more detail. Firstly, the dust yield can be calculated as [566]:

$$\dot{Y}_d = \int_{m_*(t)}^{m_{up}} m_d(M_s, Z)\phi(M_s)\psi(t - \tau_m)dM_s, \quad (72)$$

where the lower integration limit refers to the mass of a star with lifetime  $\tau_m = t$ ,  $m_{up}$  is the upper mass limit of the stellar IMF and  $m_d$  is the dust mass produced by a star of mass  $M_s$  and metallicity  $Z$ . Further, the dust accretion rate can be expressed as (e.g. [580])

$$\dot{G}_{ac} = X_c \frac{M_d(t)}{\tau_{acc}}, \quad (73)$$

where  $X_c$  is the mass fraction of cold gas where dust can grow on a timescale  $\tau_{acc}$  which critically depends on the density ( $n_{mol}$ ), temperature ( $T_{mol}$ ) and metallicity ( $Z$ ) of the cold phase such that [576,579]

$$\tau_{acc} \approx 20 \text{ Myr} \times \left( \frac{n_{mol}}{100 \text{ cm}^{-3}} \right)^{-1} \left( \frac{T_{mol}}{50 \text{ K}} \right)^{-1/2} \left( \frac{Z}{Z_\odot} \right)^{-1}. \quad (74)$$

Finally, the rate of dust destruction can be written out as  $\dot{D}_{SN} = M_d/\tau_{dest}$  where the destruction timescale  $\tau_{dest}$  is [581,582]

$$\tau_{dest}(t) = \frac{M_g(t)}{v \dot{M}_* \epsilon M_{sn} (100 \text{ km s}^{-1})}, \quad (75)$$

where  $M_g$  is the gas mass,  $\dot{M}_*$  is the star formation rate and  $\epsilon$  is the efficiency of dust destruction in a SN-shocked ISM which has typical values between 0.1 – 0.5 for varying densities of the ISM and magnetic field strengths [581,583]. Further,  $M_{sn}(100 \text{ km s}^{-1})$  is the mass accelerated to  $100 \text{ km s}^{-1}$  by the SN blast wave and has a value of  $6.8 \times 10^3 M_\odot$  [582]. However, the relative importance of the above mentioned processes of dust formation and destruction in early galaxies remain under debate. For example, Ferrara et al. [584] have shown that ISM dust growth remains problematic, as in the diffuse ISM (where grains reside most of the time), accretion rates are very slow, dust temperatures are high and growth is impeded by the Coulomb barrier. They therefore conclude that the impact of dust destruction by shocks has probably been severely overestimated.

In the absence of direct detection of far infra-red dust emission, the key quantity used to characterize the impact of dust has been the relation between the spectral slope ( $\beta$ ; measured between 1500–3000 Å) in the galaxy rest frame and the observed UV magnitude – the “Meurer relation” [585]. Although calibrated using low- $z$  starburst galaxies and assuming  $\beta = -2.23$ , where  $\beta$  is naturally a function of the age, metallicity and IMF of the underlying stellar population and the dust extinction curve (expressing the extinction as a function of wavelength), this relation has now been widely applied to  $z \gtrsim 5$  galaxies to infer the  $\beta - M_{UV}$  – or the so-called “color magnitude relation” [547,586–589]. A number of theoretical works have now confirmed the existence of this relation [493,590,591], highlighting the fact that the  $\beta$  slope becomes increasing bluer (shallow) with decreasing luminosities and increasing redshifts and that the scatter [589] arises because of a mix of intrinsically blue (dust-free) and dust enriched older objects, specially at  $-18 \lesssim M_{UV} \lesssim -19$  [591]. Indeed, this decrease in the impact of dust on the UV luminosity reflects in the fact that while, theoretically,  $z \simeq 8$  galaxies require almost no dust to match to the observed UV LF, matching to the  $z \simeq 5$  UV LF requires dust extinction at all luminosities brighter than  $M_{UV} \simeq -18$  [426,497,515,522,592–594].

The presence of far infra-red dust emission, e.g. from ALMA for  $z \simeq 5.25$  galaxies [595], can also be used to get a hint on the dust attenuation through the infra-red excess-UV spectral slope (IRX- $\beta$ ) relation [585]. Here,  $IRX = \log(F_{IR}/F_{1600})$  is the ratio between the far infra-red dust emission and the UV flux at 1600 Å in the galaxy rest-frame. Theoretical works now show that, given their lower stellar masses and lower chemical maturity, high- $z$  galaxies are characterized by lower IRX values and bluer  $\beta$  slopes compared to lower- $z$  counterparts for the same colours. Capak et al. [595] point out that the IRX values can also be used an indicator of the dust formation mechanisms – for example, significant dust contribution from grain growth can induce a considerable scatter in the IRX at a given value of the  $\beta$  slope. Given their blue slopes, most high- $z$  galaxies only had upper limits on the dust-mass until the discovery of dust in the  $z = 7.5$  galaxy A1689-zD1 and A2744<sub>y</sub>D4 at  $z = 8.38$  as summarized in Table 3. Theoretical works have shown that even using the upper limit of the stellar dust yields (from both SN and AGBs) results in dust masses that are about two orders of magnitude lower than those observed for both A1689-zD1 and A2744<sub>y</sub>D4 and that the only way of reconciling the observed dust masses with theory is to include grain growth on timescales as short as 0.2 Myr which imply a density of about  $10^4 \text{ cm}^{-3}$  in the cold phase of the ISM [580] (and see also [596]). Mancini et al. [580] also make the point that although grain growth becomes the dominant process as early as  $z \sim 10 - 12$ , the average contribution from AGBs can be large as 40% and can therefore not be ignored. However, Behrens et al. [597] point out that given that the far infra-red luminosity scales as  $L_{IR} \propto M_{dust} T_{dust}^6$ , these observations could be explained by reasonable dust masses if the temperature is ramped up by a factor 2–3. Further observations with ALMA will be imperative in shedding more light on the (currently debated) dust masses of early galaxies.

**Table 3**

Summary of the high- $z$  galaxies named in column 1 with the redshifts (column 2), UV magnitudes (column 3), stellar masses (column 4) and dust masses (column 5). All quantities are computed using the homogeneous stellar mass-magnitude relations and dust temperatures and emissivity values as in [580].

Name	$z$	$M_{\text{UV}}$	$\text{Log}(M_*/M_\odot)$	$\text{Log}(M_d/M_\odot)$	Reference
A1703-zD1	6.800	-20.3	$9.2 \pm 0.3$	$<7.36$	[598]
z8-GND-5296	7.508	-21.4	$9.7 \pm 0.3$	$<8.28$	[598]
HCM6A	6.560	-20.8	$9.5 \pm 0.3$	$<7.71$	[599]
IOK-1	6.960	-21.3	$9.7 \pm 0.3$	$<7.43$	[600]
Himiko	6.595	-21.7	$9.9 \pm 0.3$	$<7.30$	[601]
BDF-3299	7.109	-20.44	$9.3 \pm 0.3$	$<7.02$	[602]
BDF-512	7.008	-20.49	$9.3 \pm 0.3$	$<7.36$	[602]
BDF-46975	6.844	-21.49	$9.8 \pm 0.3$	$<7.38$	[602]
A1689-zD1	7.500	-19.7	$9.0 \pm 0.3$	$7.51 \pm 0.2$	[603]
A2744yD4	8.38	-	$9.29^{+0.24}_{-0.11}$	$6.74^{+0.65}_{-0.17}$	[604]

### 8.6. Spin and shapes

We start by discussing the spin of high- $z$  galaxies before discussing their shapes. Galaxy formation theory (e.g. [165,605]) suggests that gas within a cooling radius can settle into a disk inside a dark matter halo. Assuming the gas and dark matter were initially well-mixed, such that the distribution of angular momentum of each mass element in the disk was the same as that in the halo, Fall and Efstathiou [165] deduced the angular momentum profile for an exponential disk and that the disk radius  $R_d \propto \lambda$ . Here,  $\lambda$  is the pre-collapse angular momentum of the halo which is generally parameterized as [606]

$$\lambda \equiv \frac{J|E|^{1/2}}{GM^{5/2}}, \quad (76)$$

where  $J$ ,  $E$  and  $M$  represent the total angular momentum, energy and mass of the system where the angular momentum is acquired through tidal interactions with neighbouring objects [606]. As early as 1987, Barnes and Efstathiou [607] used CDM N-body simulations to show that  $\lambda$  is insensitive to the initial primordial density perturbation spectrum and follows a log-normal distribution with a median value of  $\lambda \sim 0.05$ . They also confirmed that the angular momentum of pre-collapse halos grows linearly with time, as predicted by linear tidal-torque theory [608]. Bullock et al. [609] re-defined  $\lambda$  so as to be easily calculated from N-body simulations as

$$\lambda' = \frac{J_{vir}}{\sqrt{2}M_{vir}V_{vir}R_{vir}}, \quad (77)$$

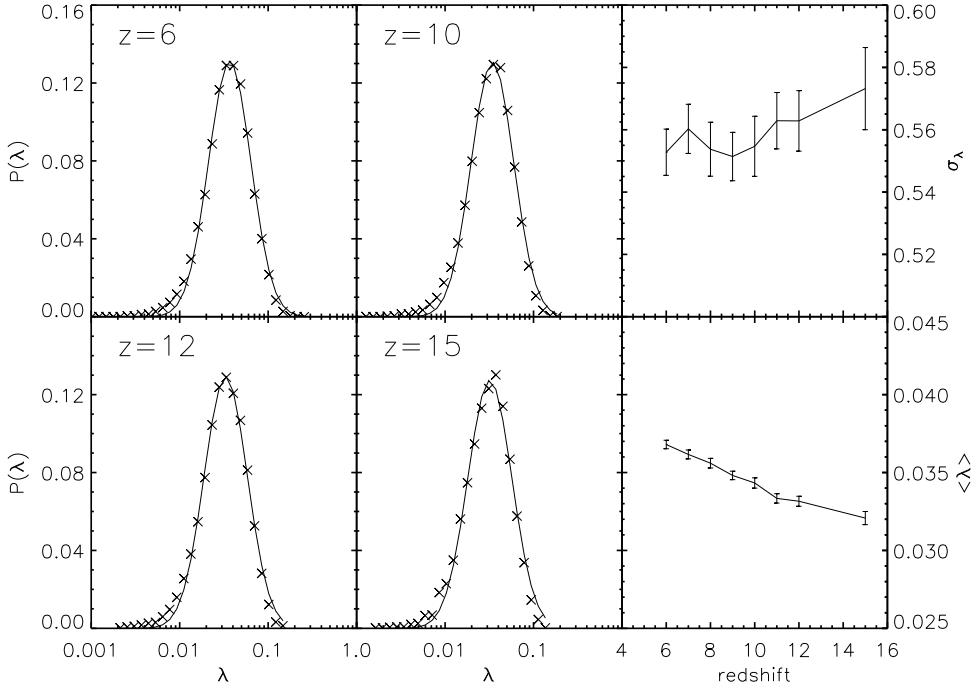
where  $J_{vir}$  represents the angular momentum enclosed within the virial mass  $M_{vir}$ . Further,  $\lambda' = \lambda \times f(c_h)$  where  $c_h$  is the concentration parameter such that  $f(c_h) \simeq [2/3 + (c_h/21.5)^{0.7}]$  [605]. Robust against the exact choice of the outer radius,  $\lambda'$  converges to  $\lambda$  at the virial radius of a singular isothermal halo. N-body simulations too show  $\lambda$  to follow a log-normal probability distribution function with a peak at  $\lambda \sim 0.05$  (e.g. [607]),  $\lambda \sim 0.04 \pm 0.5$  [609,610] and  $\lambda \sim 0.04 \pm 0.55$  [611], almost irrespective of mass or  $z$  [610,611] as also shown in Fig. 30. This implies a relation between the disk and halo radius through  $R_d = \lambda R_{vir}$  such that the size distribution should reflect the underlying dark matter halo size distribution because gas acquires most of its angular momentum with dark matter halos before collapse. Interestingly, the spin correlates with the environment, with galaxies with high spin parameters being more clustered in accord with the tidal-torque theory (e.g. [607,611]).

Further, the cumulative mass distribution of the specific angular momentum has been found to have a *universal* shape that can be fit by a two-parameter ( $\mu$  and  $j_0$  below) function of the form [609]

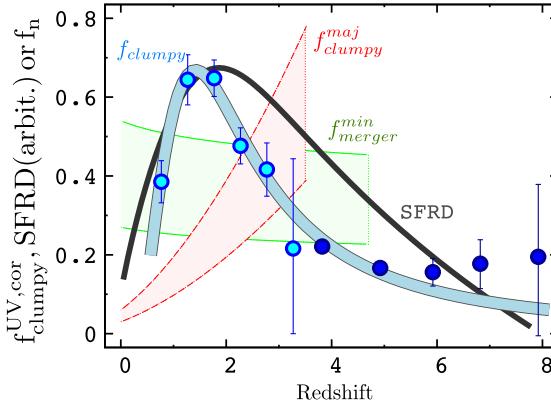
$$M(< j) = M_h \frac{\mu j}{j_0 + j}; \quad \mu > 1, \quad (78)$$

where  $j$  represents the specific angular momentum of a halo of mass  $M_h$ . Naturally, the profile has a maximum specific angular momentum  $j_{max} = j_0/(\mu - 1)$ , follows a power-law at  $j \lesssim j_0$  and flattens at  $j \gtrsim j_0$ . Here,  $\mu$  acts as a shape parameter, leading to a pure power-law (pronounced bend) for  $\mu \gg 1$  ( $\mu \rightarrow 1$ ). Given that galaxies assemble their mass through a mix of relatively-quiescent accretion and mergers, this relation probably arises as a combination of the tidal-torque theory and the non-linear process of angular momentum transfer from satellite orbits during mergers (see e.g. [609]); [610] also show that  $\lambda$  peaks sharply during major mergers and shows a steady decline during the gradual accretion of small satellites.

As for the shapes of high- $z$  galaxies, as early as 2004, Ferguson et al. [615] found a semi-major to semi-minor axis ratio of  $b/a = 0.65$  for  $386 z \sim 4$  galaxies selected via HST, with luminosities between  $0.7 - 5L_*$ , suggesting LBGs are not spheroids. This result was borne out by the larger sample (4700 LBGs between  $z \simeq 2.5 - 5$ ) collected by Ravindranath et al. [616] who found 40% of LBGs to have exponential light-profiles, similar to disks, with only about 30% showing the high concentrations seen for spheroids. A significant fraction (about 30%) showed light profiles shallower than an exponential, exhibiting multiple cores or disturbed morphologies, indicative of close pairs or on-going mergers. As shown in Fig. 31, Shibuya et al. [612] show that while their observed LBGs are consistent with disk-like surface brightness profiles, the fraction of “clumpy” galaxies in



**Fig. 30.** N-body simulations, for a box size of  $2.46h^{-1}$  Mpc with a DM resolution mass of  $7.3 \times 10^3 M_\odot$ , run from  $z \sim 100 - 6$  that show the spin parameter for all halos for  $z = 15 - 6$  as marked in the left two columns [611]. As shown by the solid line, the distribution is well-fit by a log-normal function. The right-most column shows the change in the mean and standard deviation of the log-normal fit for all the outputs studied.



**Fig. 31.** The redshift evolution of the “clumpy” galaxy fraction at  $z \simeq 0 - 8$  for star forming galaxies (filled cyan circles) and LBGs (filled blue circles) with  $L_{\text{UV}} = 0.3 - 1L_{g=3}$  [612] with Poissonian error bars. For comparison, the solid black line shows the cosmic SFRD from [613] which has been arbitrarily shifted and scaled along the y-axis. The shaded red and green regions show the galaxy major and minor merger fractions, respectively, assuming that the merger observability timescale ranges from 1 to 2 Gyr [614].

the UV, corrected for detection incompleteness,  $f_{\text{clumpy}}^{\text{uv,cor}}$ , whilst rising from  $z \simeq 8$  to 1 shows a subsequent decreases to  $z \simeq 0$ , tracking the cosmic star formation rate density, well described by the following function:

$$f_{\text{clumpy}}^{\text{uv,cor}} = a \times \frac{(1+z)^b}{1 + [(1+z)/c]^d}, \quad (79)$$

where the best fit parameters are found to be  $a = 0.035$ ,  $b = 4.6$ ,  $c = 2.2$  and  $d = 6.7$ . These results are in good agreement with those found by other groups (see Fig. 4; [612]) who find  $f_{\text{clumpy}}^{\text{UV}} \simeq 20\%$  at  $z \simeq 3 - 5$  [617],  $f_{\text{clumpy}}^{\text{UV}} \simeq 40\%$  at  $z \simeq 7$  [618] and  $f_{\text{clumpy}}^{\text{UV}}$  rising from  $\simeq 10\%$  at  $z \simeq 6 - 7$  to as much as  $50\%$  by  $z \simeq 8$  [619]. LAE observations too show the same clumpy structure, both at  $z \sim 3.1$  [620] and, to a smaller extent, at  $z \simeq 4 - 6$  [478]. As shown in Fig. 31, the observed redshift trend of  $f_{\text{clumpy}}^{\text{uv,cor}}$  over  $z \simeq 0 - 8$  requires minor mergers at  $z \lesssim 1$  and  $z \gtrsim 2.5$  whilst requiring most mergers to be major at  $z \simeq 1 - 2$ . A more elegant explanation is therefore offered by the violent disk instability model (e.g. [175,621]) where clumps form

in unstable regions with a Toomre parameter value below the critical value of unity in thick, gas-rich disks. Depending on the inflows of cold gas, this model predicts the clumping fraction to evolve with the cosmic star formation rate density, as observed. Finally, in terms of galaxy properties, Shibuya et al. [612] find that  $f_{\text{clumpy}}^{\text{UV}}$  increases with the UV luminosity, star formation rate density and the blue-ness of the spectral slope, although the error bars are quite significant.

On the theoretical front, simulations show high- $z$  galaxies to be dominated by disks although the halo mass at which disks start dominating remains a matter of debate: while Romano-Díaz et al. [622] find a 100% disk fraction in  $z \simeq 10$  galaxies for a total mass  $\gtrsim 10^9 M_\odot$ , Feng et al. [623] find disks to dominate (disk fraction of roughly 70%) in much larger  $z = 8 - 10$  galaxies with stellar masses  $\gtrsim 10^{10} M_\odot$ . On the other hand, other works (e.g. [281]) find disks in galaxies with masses as low as  $8 \times 10^7 - 1.2 \times 10^8 M_\odot$ . Forthcoming observations with the *JWST* and *E-ELT* will be crucial in shedding more light on the shapes of early galaxies.

### 8.7. Sizes

Indicating the dynamical state resulting from the processes (of feedback, merging, inflows, outflows) associated with galaxy assembly the size evolution of galaxies, measured through the half-light radius ( $R_{\text{hl}}$ ) and the size-luminosity relation, are extremely useful parameters for constraining theoretical models. Mo et al. [605] have shown that, for an infinitesimally thin disk with an exponential surface density profile, the disk size  $R_d$ , which can be used as a proxy of  $R_{\text{hl}}$ , is related to  $R_{\text{vir}}$  as

$$R_d = \frac{\lambda}{\sqrt{2}} \frac{j_d}{\mu_d} R_{\text{vir}}, \quad (80)$$

where  $\mu_d$  and  $j_d$  are the fraction of mass and angular momentum in the disk relative to the halo. This relation can be used to derive the expected  $z$ -evolution of  $R_d$  through the redshift evolution of  $R_{\text{vir}}$  using the virial theorem (Section 3.3) such that:

$$R_d(z) \propto R_{\text{vir}}(z) \propto M_h^{1/3} H^{-2/3}(z), \quad (81)$$

where we use  $M_h = (4\pi/3)R_{\text{vir}}^3 \bar{\rho}$ , assuming a collapse density of 200 times the critical density. Using  $H(z) = H(0)(\Omega_m(1+z)^3 + \Omega_\Lambda)^{1/2}$  and ignoring  $\Omega_\Lambda$ , which is reasonable at  $z > 2$ , at a fixed halo mass we obtain

$$R_d \propto (1+z)^{-1}. \quad (82)$$

The virial theorem can also be used to link  $R_d$  to the galaxy velocity and to the luminosity through the Tully–Fisher [624] and Faber–Jackson [625] relations that find  $L \propto V_{\text{vir}}$ . For a given  $L$ , this yields

$$R_d \propto \frac{V_{\text{vir}}}{H(z)} \propto (1+z)^{-1.5}. \quad (83)$$

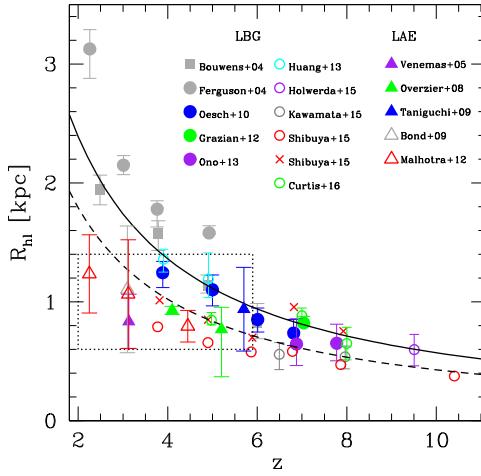
These two behaviours bracket the range of  $R_d = f_n(z)$  allowed using these simple arguments.

Liu et al. [626] have used a 100 Mpc N-body simulation with a DM particle resolution mass of  $2.6 \times 10^6 h^{-1} M_\odot$  to calculate  $R_d$  assuming  $j_d/\mu_d = 1$  and calculating  $\lambda$  as in Eq. (77). Estimating  $R_{\text{hl}} = 1.678 R_d$ , these authors find  $R_{\text{hl}} \propto (1+z)^{-m}$  with  $m = 2.0 \pm 0.07$  for  $z \sim 5 - 10$  LBGs with luminosities in the range  $0.3 - 1L_{z=3}^*$ . Further, they find  $R_{\text{hl}} \propto L^\beta$  where  $\beta$  steepens from  $0.25 \pm 0.02$  at  $z \sim 5$  to  $0.36 \pm 0.03$  at  $z \sim 10$  implying galaxies of a given luminosity become more compact with increasing  $z$ , as expected. These authors also point out that SN feedback is required in order to reproduce observed galaxy sizes at  $z \sim 5$ ; the lack of SN feedback results in a younger stellar population that produces more luminosity, leading to a shallower  $R_d - L$  relation. These theoretical results are slightly steeper than those inferred from observational campaigns, shown in Fig. 32, which have been used to infer  $m \sim -1.5$  [615,627],  $m = -1.12 \pm 0.17$  [511],  $m = -1.2 \pm 0.1$  [619,628,629] and  $m = -1$  [612,630–632] in the range  $(0.3 - 1)L_{z=3}^*$ . However, Curtis-Lake et al. [633] point out that using the modal value of the size-distribution results in a very shallow relation at  $z \sim 4 - 8$  fit by  $m = 0.2 \pm 0.26$  which is consistent with the null hypothesis of no size evolution.

Observations have detected a number of other correlations including those linking the size and stellar mass where galaxies with increasing stellar masses have larger  $R_{\text{hl}}$  values over  $z \sim 0$  to 7 [628,639,640]. However, at high- $z$ , the UV luminosity is a much more robust quantity so that we can express  $R_{\text{hl}} \propto L_{\text{UV}}^\alpha$ . Indeed, the theoretical value of  $\alpha = 0.25$  is in good agreement with the values inferred observationally:  $\alpha \sim 0.3 - 0.5$  at  $z \sim 7$  [516],  $\alpha = 0.24$  at  $z \sim 7$  [631],  $\alpha = 0.22 - 0.25$  at  $z \sim 4, 5$  [634] and  $\alpha = 0.27$  at  $z \sim 8$  [612]. Exploiting the power of the Hubble frontier Fields, *MUSE* and *VLT*, the  $R_{\text{hl}} - L_{\text{UV}}$  tend has now been confirmed to hold down to magnitudes as faint as  $M_{\text{UV}} = -15$  to  $-17$ , yielding point source profiles with  $R_{\text{hl}} \sim 160 - 240 \text{ pc}$  at  $z \sim 2 - 8$  [641] and  $R_{\text{hl}} \sim 16 - 140 \text{ pc}$  at  $z \sim 6$  [346] with  $\alpha \sim 0.5$ .

The size distribution of LBGs can be well fit by a log-normal distribution with a minor evolution in the peak value from  $z \sim 2.3 - 7$  [511,634] with a tail of larger galaxies building up towards lower- $z$  driven either by hierarchical build up or an enhanced accretion of cold gas [511,615]. The values of luminosity and the half light radius have also been used to infer the average star formation surface densities whose values are found to be  $\sum_{\text{SFR}} \sim 1.9 - 2 M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$  at  $z \sim 4 - 8$  for  $(0.3 - 1)L_{z=3}^*$  LBGs, hinting at very similar star formation efficiencies for these LBGs [511,629]. However, these values are slightly lower than  $\sum_{\text{SFR}} \sim 4.1 (5.8) M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$  at  $z \sim 6 - 7 (8)$  inferred by Kawamata et al. [619].

On the other hand, galaxies detected as LAEs are found to be more compact than LBGs [642] and show a  $z$ -independent, constant small size  $\sim 1.6 \text{ kpc}$  at  $z \sim 2.5 - 6.5$  [478,620,636–638]. At the same redshifts, LBGs are bigger by a factor of 1.2–2.5, with the sizes for these two populations converging at  $z \gtrsim 5$  [638].



**Fig. 32.** The observed half light radius for LAEs and LBGs as marked in the legend. The LBG data sets refer to galaxies with  $0.3 - 1L_{*,z=3}$  from Bouwens et al. [630], Oesch et al. [511], Grazian et al. [516], Ono et al. [629], Huang et al. [634], Holwerda et al. [631], Shibuya et al. [635], with circles and crosses showing median and average results respectively and Curtis-Lake et al. [633], as marked. The Ferguson et al. [615] points are for bright LBGs with  $L = 0.7 - 5L_{*,z=3}$  which might be responsible, at least in part, for their larger estimates of  $R_{hl}$  compared to the other studies. The LAE data sets have been collected by Venemans et al. [636], Overzier et al. [637], Taniguchi et al. [478], Bond et al. [620] and Malhotra et al. [638], as marked. As shown by the solid and dashed black lines, the LBG size evolution is well-fit by a relation  $R_{hl} \propto (1+z)^{-1.1}$  [612,630] with different normalizations obtained by different groups. On the other hand, LAEs (demarcated by the dotted box) show a much shallower size evolution with  $z$ , with LAEs and LBGs effectively having the same sizes at  $z \gtrsim 5$  [638].

Finally, we note that the size-luminosity relation has enormous implications for the faint end of the UV LF. This is because the effects of completeness correction become increasingly severe if faint high- $z$  galaxies are extended sources. While the faint end of the UV LF could be as steep as  $\alpha = -2$  if faint galaxies are extended ( $\sim 0.2 - 0.3'$  at apparent J-band magnitudes of  $= 28 - 29$ ), it would be much shallower with  $\alpha \sim -1.7$  if faint galaxies can be approximated as point objects [516]. Indeed, the latest observations of faint ( $M_{UV} = -15$  to  $-17$ ) LBGs at  $z \sim 6 - 8$  suggest these to fit the latter scenario, resulting in smaller completeness corrections and shallower faint end slopes by a factor of 2–3 [346,517]. The  $R_{hl} - L_{UV}$  relation also has connotations for the  $z$ -evolution of the UV LF: although the surface brightness is constant for both populations, while the evolving sizes of LBGs (coupled with the constant surface brightness) result in an evolving UV LF, the constant small sizes of LAEs can explain their non-evolving UV LF at  $z \sim 3 - 6$  [638]. Finally, Zick et al. [643] caution that the faint, compact sources seen in the Hubble frontier fields are most probably globular clusters hosted by faint galaxies rather than being faint galaxies themselves.

## 9. Open questions and future outlook

Despite the enormous progress in our understanding of the inter-twined processes of galaxy formation and reionization in the first billion years, a number of issues remain unsolved as it should now appear clear from this review. While a number of these issues might remain unsolved or debated in the short term, where possible we provide an outlook on how others might be solved.

### 1. What is the nature of dark matter?

As shown in Section 3, while the cold dark matter paradigm has been exceptionally successful at explaining the large scale structure of the Universe, its small scale problems have left the field open to proposals of *alternative* dark matter models including warm dark matter, interactive cold dark matter, self-interacting dark matter and annihilating dark matter, to name a few. While detections of a 3.5 kev line from the Perseus cluster could possibly arise as a result of dark matter annihilations, given the lack of any experimental evidence of super-symmetry, the nature of dark matter remains one of the greatest puzzles in modern cosmology.

### 2. IMF of the first stars

As shown in Section 5.2, the classic paradigm, of metal-free gas collapsing to form one massive star, underwent a change in 2010 when simulations started finding metal-free gas to be extremely prone to fragmentation. However, the final mass of the PopIII star, as well as the IMF remain open questions. Observations of metal-poor stars in the Milky Way and detections of PopIII stars in the early Universe, using for example the *JWST* would be invaluable on shedding more light on these issues.

### 3. Inhomogeneous metal enrichment in the first billion years

The metal enrichment of the ISM remains poorly understood in the early Universe. How and when metallicity relations – linking key galaxy observables such as the stellar mass, star formation rate, gas-phase metallicity and possibly even the cold gas mass – emerged remains an open question. In the near future, a combination of observations with ALMA (of the molecular component) and the *JWST* (of nebular emission lines to infer the gas-phase metallicity) could help extend lower- $z$  mass-metallicity relations to these early cosmic epochs.

#### 4. LyC escape fraction and its dependence on galaxy properties and redshift

As shown in Section 7.1, the escape fraction of LyC photons, being a complex function of the ISM gas and dust content, supernova feedback, turbulence and line of sight, remains a major open question for reionization, with both theory and observations finding values ranging anywhere between  $< 1\%$  to  $\gtrsim 50\%$ . However, forthcoming observations, with the JWST might potentially yield strong LyC leakers up to  $z \simeq 9$ . In addition, indirect methods such as those combining estimates of the UV spectral slope ( $\beta$ ) with the  $H\beta$  emission line, using the  $[OIII]\lambda 5007/[OII]\lambda 3727$  ratio as a proxy for  $f_{esc}$ , or exploiting  $H\alpha$  and visible continuum surface brightness profiles with the JWST could help shed more light on the mass- and redshift-dependence of  $f_{esc}$ .

#### 5. Topology and history of reionization

As shown in Section 7, both the reionization history and topology (inside-out versus outside-in) depend on a number of inter-linked processes including the abundance- and redshift evolution of star forming galaxies, the supernova- and reionization-feedback dependent intrinsic star formation (and hence LyC photon production) rates, the escape fraction of LyC photons from the galaxies and the IGM clumping factor to name a few. As of now, the  $z$ -evolution of LAEs and the fraction of LBGs showing  $Ly\alpha$  emission have been used to get hints on the, average and piecewise, reionization history between  $z \sim 5.7 - 7$ . However, over the next decade, 21 cm experiments such as the SKA and Hera will be used, not just to obtain tomographical maps of the reionization process, but also shed light on the topology of reionization by cross-correlating 21 cm and galaxy data (e.g. [644–648]).

#### 6. Impact of the UVB background on galaxy formation

As shown in Section 7.3, while the UV background, could, on the one hand, slow down the progress of reionization by photo-evaporating gas from low-mass halos, this effect could, to some extent, be countered by the UVB smoothing out the IGM clumping. However, the UV impact on both these processes critically depends on the emissivity from reionization sources, the strength of the UVB and the relative redshifts of the source and that at which the UVB impinges on a source, to name a few. As a result, the impact of the UVB on both star formation and the IGM clumping factor remain open questions. Although observations of the faintest galaxies with the JWST might shed light on the faint end of the UV LF, decoupling the impact of supernova versus UV feedback will, possibly, not be tractable.

#### 7. Role of AGN, intermediate mass black holes and micro quasars in reionization

As shown in Section 7.4, while most works focus on galaxies to be the key reionization sources, the role low to intermediate mass black holes, hosted by faint AGN or micro quasars, remains an open question. Indeed, as a number of works have shown, the secondary ionizations from black hole activity and the suppression of the most numerous low-mass galaxies due to the rising UVB, could both allow significant, or at least non-negligible, reionization contributions from low mass black holes.

#### 8. How far could the faint-end of the UV LF extend?

Star formation in galaxies requires the dark matter halos to be able to, hold on to, and cool, a sufficient amount of, gas. However, given that both supernova and UV feedback can, to some extent, heat or blow-out the gas from low-mass halos, the lower limit of halos that can form stars, and hence contribute to reionization, remains an open question. Naturally this implies that both the faint-end slope and the cut-off magnitude (beyond which star formation is suppressed) remain open questions as of now. In the future, using JWST to observe lensed galaxies might possibly shed light on the faint-end turn-over point of the UV LF and its redshift evolution, if not the cut-off magnitude.

#### 9. Dust formation and obscuration

As shown in Section 8.5, explaining the extremely high observed dust-to-stellar mass ratios ( $\sim 0.2\%$ ) of some high- $z$  galaxies requires most of the dust be grown by accretion in the ISM, on timescales that are 10 times shorter than those inferred for the Milky Way, with only a small fraction provided by supernovae and AGB stars. Given that works show dust grain growth to be extremely inefficient in the ISM given its high ionization fields, the key dust sources, masses and dust impact on the observability of high- $z$  galaxies remain open questions. Pinning down the dust masses of high- $z$  galaxies, ALMA will be invaluable in shedding more light on this issue over the next years.

### Acknowledgments

PD acknowledges support from the ERC Starting Grant DELPHI H2020/717001, and the European Commission and University of Groningen CO-FUND Rosalind Franklin program. AF acknowledges support from the ERC Advanced Grant INTERSTELLAR H2020/740120. This research was also partly supported by the Munich Institute for Astro- and Particle Physics (MIAPP), Italy of the DFG cluster of excellence “Origin and Structure of the Universe”. PD acknowledges the generous hospitality of the Scuola Normale Superiore, where part of this work was carried out. The authors are grateful to the anonymous referee, Rychard Bouwens, Volker Bromm, Emma Curtis-Lake, Benedetta Ciardi, Avishai Dekel, Nick Gnedin, Anne Hutter, Anupam Mazumdar, Pascal Oesch, Andrea Pallottini, Olmo Piana, Renske Smit, and Eros Vanzella for their invaluable scientific inputs. Finally PD dedicates this review to her grandmother – the first mathematician in the family.

### References

- [1] G. Strömgren, *Astrophys. J.* 80 (1934) 327.
- [2] C.F. von Weizsäcker, *Astrophys. J.* 114 (1951) 165.
- [3] F. Hoyle, *Problems of Cosmical Aerodynamics*, 1951, p. 195.

- [4] F. Hoyle, *Astrophys. J.* 118 (1953) 513.
- [5] D.W. Sciama, *Mon. Not. R. Astron. Soc.* 115 (1955) 3–14.
- [6] G. Gamow, *Astron. J.* 58 (1953) 39.
- [7] H. Bondi, *Cosmology*, Cambridge [Eng.] University Press, 1952.
- [8] B. Vorotilov-Velyaminov, *Astron. J.* 66 (1961) 551.
- [9] J.P. Ostriker, L.L. Cowie, *Astrophys. J. Lett.* 243 (1981) L127–L131.
- [10] A.A. Penzias, R.W. Wilson, *Agron. J.* 142 (1965) 419–421.
- [11] P.J.E. Peebles, *Astrophys. J.* 142 (1965) 1317.
- [12] E.R. Harrison, *Astron. J.* 72 (1967) 303.
- [13] J. Silk, *Nature* 218 (1968) 453–454.
- [14] E.R. Harrison, *Phys. Rev. D* 1 (1970) 2726–2730.
- [15] Y.B. Zeldovich, *Mon. Not. R. Astron. Soc.* 160 (1972) 1P.
- [16] Y.B. Zel'dovich, *Astron. Astrophys. S.* 5 (1970) 84, arXiv:1507.02679.
- [17] J.E. Gunn, J.R. Gott III, *Astrophys. J.* 176 (1972) 1.
- [18] W.H. Press, P. Schechter, *Agron. J.* 187 (1974) 425–438.
- [19] R.B. Larson, *Mon. Not. R. Astron. Soc.* 169 (1974) 229–246.
- [20] S.D.M. White, M.J. Rees, *Mon. Not. R. Astron. Soc.* 183 (1978) 341–358.
- [21] J.R. Gott III, T.X. Thuan, *Astrophys. J.* 204 (1976) 649–667.
- [22] J. Silk, *Astrophys. J.* 211 (1977) 638–648.
- [23] J. Binney, *Astrophys. J.* 215 (1977) 483–491.
- [24] M.J. Rees, J.P. Ostriker, *Mon. Not. R. Astron. Soc.* 179 (1977) 541–559.
- [25] V.C. Rubin, N. Thonnard, W.K. Ford Jr., *Astrophys. J. Lett.* 225 (1978) L107–L111.
- [26] A.H. Guth, *Phys. Rev. D* 23 (1981) 347–356.
- [27] K. Sato, *Mon. Not. R. Astron. Soc.* 195 (1981) 467–479.
- [28] A. Mazumdar, J. Rocher, *Phys. Rep.* 497 (2011) 85–215, arXiv:1001.0993.
- [29] A.D. Linde, *Phys. Lett. B* 116 (1982) 335–339.
- [30] A. Albrecht, P.J. Steinhardt, *Phys. Rev. Lett.* 48 (1982) 1220–1223.
- [31] A.D. Linde, *Phys. Lett. B* 129 (1983) 177–181.
- [32] J.M. Bardeen, P.J. Steinhardt, M.S. Turner, *Phys. Rev. D* 28 (1983) 679–693.
- [33] P.J.E. Peebles, *Astrophys. J. Lett.* 263 (1982) L1–L5.
- [34] J.M. Bardeen, J.R. Bond, N. Kaiser, A.S. Szalay, *Astrophys. J.* 304 (1986) 15–61.
- [35] N. Sugiyama, *Astrophys. J. Suppl.* 100 (1995) 281, arXiv:astro-ph/9412025.
- [36] J.F. Navarro, W. Benz, *Astrophys. J.* 380 (1991) 320–329.
- [37] N. Katz, J.E. Gunn, *Astrophys. J.* 377 (1991) 365–381.
- [38] C.C. Steidel, M. Giavalisco, M. Pettini, M. Dickinson, K.L. Adelberger, *Agron. J.* 462 (1996) L17.
- [39] R.S. Ellis, et al., *Astrophys. J.* 763 (2013) L7.
- [40] H. Atek, et al., *Astrophys. J.* 814 (1) (2015) 69.
- [41] R. Smit, R.J. Bouwens, I.e.a. Labb  , *Astrophys. J.* 784 (2014) 58.
- [42] A. Zitrin, I. Labbe, S.e.a. Belli, *Agron. J.* 810 (1) (2015) L12.
- [43] J.B. Oke, J.E. Gunn, *Agron. J.* 266 (1983) 713–717.
- [44] R. Maartens, *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 369 (2011) 5115–5137.
- [45] D. Saadeh, S.M. Feeney, A. Pontzen, H.V. Peiris, J.D. McEwen, *Phys. Rev. Lett.* 117 (13) (2016) 131302.
- [46] R.A. Alpher, H. Bethe, G. Gamow, *Phys. Rev.* 73 (1948) 803–804.
- [47] Planck Collaboration, P.A.R. Ade, N. Aghanim, M. Arnaud, et al., *Astron. Astrophys.* 594 (2016) A13, arXiv:1502.01589.
- [48] Planck Collaboration, N. Aghanim, M. Ashdown, et al., *Astron. Astrophys.* 596 (2016) A107, arXiv:1605.02985.
- [49] C.L. Bennett, M. Halpern, C. Hinshaw, et al., *Astrophys. J. Suppl.* 148 (2003) 1–27, arXiv:astro-ph/0302207.
- [50] P.J.E. Peebles, *Physical Cosmology*, in: *Princeton Series in Physics*, Princeton University Press, Princeton, NJ, 1971.
- [51] G.R. Blumenthal, S.M. Faber, J.R. Primack, M.J. Rees, *Nature* 311 (1984) 517–525.
- [52] J.R. Bond, A.S. Szalay, *Astrophys. J.* 274 (1983) 443–468.
- [53] S. Cole, W.J. Percival, J.A. Peacock, et al., *Mon. Not. R. Astron. Soc.* 362 (2005) 505–534.
- [54] G. Hinshaw, D. Larson, E. Komatsu, et al., *Astrophys. J. Suppl.* 208 (2013) 19.
- [55] Planck Collaboration, P.A.R. Ade, N. Aghanim, C. Armitage-Caplan, et al., *Astron. Astrophys.* 571 (2014) A16.
- [56] A. Slosar, V. Ir  i  , D. Kirkby, et al., *J. Cosmol. Astropart. Phys.* 4 (2013) 26.
- [57] D.H. Weinberg, J.S. Bullock, F. Governato, R. Kuzio de Naray, A.H.G. Peter, *Proc. Natl. Acad. Sci.* 112 (2015) 12249–12255, arXiv:1306.0913.
- [58] B. Moore, T. Quinn, F. Governato, J. Stadel, G. Lake, *Mon. Not. R. Astron. Soc.* 310 (1999) 1147–1152.
- [59] A. Klypin, A.V. Kravtsov, O. Valenzuela, F. Prada, *Astrophys. J.* 522 (1999) 82–92.
- [60] M. Boylan-Kolchin, J.S. Bullock, M. Kaplinghat, *Mon. Not. R. Astron. Soc.* 415 (2011) L40–L44, arXiv:1103.0007.
- [61] M. Boylan-Kolchin, J.S. Bullock, M. Kaplinghat, *Mon. Not. R. Astron. Soc.* 422 (2012) 1203–1218.
- [62] B. Moore, S. Ghigna, F. Governato, G. Lake, T. Quinn, J. Stadel, P. Tozzi, *Astrophys. J. Lett.* 524 (1999) L19–L22, arXiv:astro-ph/9907411.
- [63] J.F. Navarro, C.S. Frenk, S.D.M. White, *Astrophys. J.* 490 (1997) 493–508.
- [64] K. Subramanian, R. Cen, J.P. Ostriker, *Astrophys. J.* 538 (2000) 528–542.
- [65] R.F.G. Wyse, in: J.G. Funes, E.M. Corsini (Eds.), *Galaxy Disks and Disk Galaxies*, in: *Astronomical Society of the Pacific Conference Series*, 230, 2001, pp. 71–80.
- [66] R. Teyssier, A. Pontzen, Y. Dubois, J.I. Read, *Mon. Not. R. Astron. Soc.* 429 (2013) 3068–3078.
- [67] P. Bode, J.P. Ostriker, N. Turok, *Astrophys. J.* 556 (2001) 93–107.
- [68] E. Bulbul, M. Markevitch, A. Foster, R.K. Smith, M. Loewenstein, et al., *Astrophys. J.* 789 (2014) 13, arXiv:1402.2301.
- [69] A. Boyarsky, O. Ruchayskiy, D. Iakubovskyi, J. Franse, *Phys. Rev. Lett.* 113 (2014) 251301, arXiv:1402.4119.
- [70] N. Cappelluti, E. Bulbul, A. Foster, et al., *Astrophys. J.* 854 (2018) 179, arXiv:1701.07932.
- [71] A.V. Macci  , G. Stinson, C.B. Brook, et al., *Astrophys. J. Lett.* 744 (2012) L9.
- [72] A. Schneider, D. Anderhalden, A.V. Macci  , J. Diemand, *Mon. Not. R. Astron. Soc.* 441 (2014) L6–L10, arXiv:1309.5960.
- [73] L. Hui, J.P. Ostriker, S. Tremaine, E. Witten, *Phys. Rev. D* 95 (4) (2017) 043541, arXiv:1610.08297.
- [74] X. Du, C. Behrens, J.C. Niemeyer, *Mon. Not. R. Astron. Soc.* 465 (2017) 941–951, arXiv:1608.02575.
- [75] D.N. Spergel, P.J. Steinhardt, *Phys. Rev. Lett.* 84 (2000) 3760–3763.
- [76] M. Rocha, A.H.G. Peter, J.S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, J. O  norbe, L.A. Moustakas, *Mon. Not. R. Astron. Soc.* 430 (2013) 81–104, arXiv:1208.3025.

- [77] M. Vogelsberger, J. Zavala, C. Simpson, A. Jenkins, Mon. Not. R. Astron. Soc. 444 (2014) 3684–3698, [arXiv:1405.5216](#).
- [78] M.-Y. Wang, A.H.G. Peter, L.E. Strigari, A.R. Zentner, B. Arant, S. Garrison-Kimmel, M. Rocha, Mon. Not. R. Astron. Soc. 445 (2014) 614–629, [arXiv:1406.0527](#).
- [79] C. Boehm, P. Fayet, R. Schaeffer, Phys. Lett. B 518 (2001) 8–14.
- [80] C. Dvorkin, K. Blum, M. Kamionkowski, Phys. Rev. D 89 (2) (2014) 023519.
- [81] J.H. Jeans, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 199 (1902) 1–53.
- [82] E.M. Lifshitz, Zh. Eksp. Teor. Fiz. 16 (1946) 587–602.
- [83] P.J.E. Peebles, Principles of Physical Cosmology, Princeton University Press, 1993.
- [84] H. Mo, F.C. van den Bosch, S. White, Galaxy Formation and Evolution, Cambridge University Press, Cambridge, UK, 2010.
- [85] P.J.E. Peebles, The Large-Scale Structure of the Universe, Princeton University Press, NJ, 1980, p. 435.
- [86] R. Barkana, A. Loeb, Phys. Rep. 349 (2001) 125–238.
- [87] B. Ciardi, A. Ferrara, Space Sci. Rev. 116 (2005) 625–705.
- [88] G.L. Bryan, M.L. Norman, Astrophys. J. 495 (1998) 80–99.
- [89] J.E. Taylor, Adv. Astron. 2011 (2011) 604898, [arXiv:1008.4103](#).
- [90] D.S. Reed, R. Bower, C.S. Frenk, A. Jenkins, T. Theuns, Mon. Not. R. Astron. Soc. 374 (2007) 2–15.
- [91] A.A. Klypin, S. Trujillo-Gomez, J. Primack, Astrophys. J. 740 (2011) 102.
- [92] S.M. Carroll, W.H. Press, E.L. Turner, Annu. Rev. Astron. Astrophys. 30 (1992) 499–542.
- [93] R.K. Sheth, G. Tormen, Mon. Not. R. Astron. Soc. 308 (1999) 119–126.
- [94] R.K. Sheth, G. Tormen, Mon. Not. R. Astron. Soc. 329 (2002) 61–75.
- [95] W.A. Watson, I.T. Iliev, A. D’Alorio, et al., Mon. Not. R. Astron. Soc. 433 (2013) 1230–1245.
- [96] M. Davis, G. Efstathiou, C.S. Frenk, S.D.M. White, Astrophys. J. 292 (1985) 371–394.
- [97] A. Klypin, J. Holtzman, Particle-Mesh code for cosmological simulations, ArXiv Astrophysics e-prints, 1997, [arXiv:astro-ph/9712217](#).
- [98] H. Trac, R. Cen, P. Mansfield, Astrophys. J. 813 (2015) 54.
- [99] S.W. Allen, A.E. Evrard, A.B. Mantz, Annu. Rev. Astron. Astrophys. 49 (2011) 409–470.
- [100] I. Harrison, P. Coles, Mon. Not. R. Astron. Soc. 421 (2012) L19–L23.
- [101] E. Bertschinger, Astrophys. J. Suppl. 58 (1985) 39–65.
- [102] J.F. Navarro, C.S. Frenk, S.D.M. White, Astrophys. J. 462 (1996) 563.
- [103] A.D. Ludlow, R.E. Angulo, Mon. Not. R. Astron. Soc. 465 (2017) L84–L88, [arXiv:1610.04620](#).
- [104] V. Springel, J. Wang, M. Vogelsberger, et al., Mon. Not. R. Astron. Soc. 391 (2008) 1685–1711, [arXiv:0809.0898](#).
- [105] J. Stadel, D. Potter, B. Moore, et al., Mon. Not. R. Astron. Soc. 398 (2009) L21–L25, [arXiv:0808.2981](#).
- [106] A.A. Dutton, A.V. Macciò, Mon. Not. R. Astron. Soc. 441 (2014) 3359–3374, [arXiv:1402.7073](#).
- [107] J. Einasto, Tr. Astrofiz. Inst. Alma-Ata 5 (1965) 87–100.
- [108] A. Burkert, Astrophys. J. Lett. 447 (1995) L25.
- [109] P. Salucci, A. Burkert, Astrophys. J. Lett. 537 (2000) L9–L12, [arXiv:astro-ph/0004397](#).
- [110] P. Kroupa, C.A. Tout, G. Gilmore, Mon. Not. R. Astron. Soc. 262 (1993) 545–587.
- [111] G. Chabrier, Astrophys. J. Lett. 586 (2003) L133–L136.
- [112] L. Gao, J.F. Navarro, S. Cole, C.S. Frenk, S.D.M. White, V. Springel, A. Jenkins, A.F. Neto, Mon. Not. R. Astron. Soc. 387 (2008) 536–544.
- [113] D.H. Zhao, Y.P. Jing, H.J. Mo, G. Börner, Astrophys. J. 707 (2009) 354–369.
- [114] A.D. Ludlow, J.F. Navarro, R.E. Angulo, M. Boylan-Kolchin, V. Springel, C. Frenk, S.D.M. White, Mon. Not. R. Astron. Soc. 441 (2014) 378–388.
- [115] T.K. Chan, D. Kereš, J. Oñorbe, P.F. Hopkins, A.L. Muratov, C.-A. Faucher-Giguère, E. Quataert, Mon. Not. R. Astron. Soc. 454 (2015) 2981–3001.
- [116] F. Governato, C. Brook, L. Mayer, et al., Nature 463 (2010) 203–206.
- [117] F. Governato, A. Zolotov, A. Pontzen, et al., Mon. Not. R. Astron. Soc. 422 (2012) 1231–1240.
- [118] A. Di Cintio, C.B. Brook, A.A. Dutton, A.V. Macciò, G.S. Stinson, A. Knebe, Mon. Not. R. Astron. Soc. 441 (2014) 2986–2995.
- [119] A. Pontzen, F. Governato, Nature 506 (2014) 171–178.
- [120] A. Ferrara, in: A. Loeb, A. Ferrara, R.S. Ellis (Eds.), Saas-Fee Advanced Course 36: Cold Aqueous Planetary Geochemistry with FREZCHEM, 2008, pp. 161–258.
- [121] P. Kroupa, Mon. Not. R. Astron. Soc. 322 (2001) 231–246, [arXiv:astro-ph/0009005](#).
- [122] C.F. McKee, E.C. Ostriker, Annu. Rev. Astron. Astrophys. 45 (2007) 565–687, [arXiv:0707.3514](#).
- [123] M.R. Krumholz, Notes on Star Formation, 2015, ArXiv e-prints, [arXiv:1511.03457](#).
- [124] L. Spitzer, Physical Processes in the Interstellar Medium, 1978.
- [125] J.P. Ostriker, C.F. McKee, Rev. Modern Phys. 60 (1988) 1–68.
- [126] S. Veilleux, G. Cecil, J. Bland-Hawthorn, Annu. Rev. Astron. Astrophys. 43 (2005) 769–826, [arXiv:astro-ph/0504435](#).
- [127] C.J. Conseilice, Annu. Rev. Astron. Astrophys. 52 (2014) 291–337, [arXiv:1403.2783](#).
- [128] R.S. Somerville, R. Davé, Annu. Rev. Astron. Astrophys. 53 (2015) 51–113.
- [129] P.S. Behroozi, J. Silk, Astrophys. J. 799 (2015) 32.
- [130] N. Mashian, P.A. Oesch, A. Loeb, Mon. Not. R. Astron. Soc. 455 (2016) 2101–2109.
- [131] A. Knebe, S.R. Knollmann, S.I. Muldrew, et al., Mon. Not. R. Astron. Soc. 415 (2011) 2293–2318, [arXiv:1104.0949](#).
- [132] V. Springel, Mon. Not. R. Astron. Soc. 401 (2010) 791–851.
- [133] J. Schaye, R.A. Crain, R.G. Bower, et al., Mon. Not. R. Astron. Soc. 446 (2015) 521–554.
- [134] I.T. Iliev, G. Mellema, K. Ahn, P.R. Shapiro, Y. Mao, U.-L. Pen, Mon. Not. R. Astron. Soc. 439 (2014) 725–743.
- [135] H.Y. Trac, N.Y. Gnedin, Adv. Sci. Lett. 4 (2011) 228–243, [arXiv:0906.4348](#).
- [136] P.F. Hopkins, D. Kereš, J. Oñorbe, C.-A. Faucher-Giguère, E. Quataert, N. Murray, J.S. Bullock, Mon. Not. R. Astron. Soc. 445 (2014) 581–603, [arXiv:1311.2073](#).
- [137] J.H. Wise, T. Abel, M.J. Turk, M.L. Norman, B.D. Smith, Mon. Not. R. Astron. Soc. 427 (2012) 311–326, [arXiv:1206.1043](#).
- [138] D. Aubert, N. Depratis, P. Ocvirik, Mon. Not. R. Astron. Soc. 454 (2015) 1012–1037.
- [139] K. Hasegawa, B. Semelin, Mon. Not. R. Astron. Soc. 428 (2013) 154–166, [arXiv:1209.4143](#).
- [140] K. Finlator, R. Davé, F. Özel, Astrophys. J. 743 (2011) 169, [arXiv:1106.4321](#).
- [141] A. Pallottini, A. Ferrara, S. Gallerani, L. Vallini, R. Maiolino, S. Salvadori, Mon. Not. R. Astron. Soc. 465 (2017) 2540–2558, [arXiv:1609.01719](#).
- [142] M. Petkova, V. Springel, Mon. Not. R. Astron. Soc. 412 (2011) 935–946, [arXiv:1008.4459](#).
- [143] M.L. Norman, D.R. Reynolds, G.C. So, R.P. Harkness, J.H. Wise, Astrophys. J. Suppl. 216 (2015) 16, [arXiv:1306.0645](#).
- [144] B.W. O’Shea, J.H. Wise, H. Xu, M.L. Norman, Astrophys. J. Lett. 807 (2015) L12, [arXiv:1503.01110](#).
- [145] H. Xu, J.H. Wise, M.L. Norman, K. Ahn, B.W. O’Shea, Astrophys. J. 833 (2016) 84, [arXiv:1604.07842](#).
- [146] N.Y. Gnedin, A.A. Kaurov, Astrophys. J. 793 (2014) 30.
- [147] S.J. Mutch, P.M. Geil, G.B. Poole, P.W. Angel, A.R. Duffy, A. Mesinger, J.S.B. Wyithe, Mon. Not. R. Astron. Soc. 462 (2016) 250–276.
- [148] P. Dayal, A. Ferrara, J.S. Dunlop, F. Pacucci, Mon. Not. R. Astron. Soc. 445 (2014) 2545–2557, [arXiv:1405.4862](#).

- [149] P. Ocvirk, N. Gillet, P.R. Shapiro, et al., Mon. Not. R. Astron. Soc. 463 (2016) 1462–1485.
- [150] A.H. Pawlik, A. Rahmati, J. Schaye, M. Jeon, C. Dalla Vecchia, Mon. Not. R. Astron. Soc. 466 (2017) 960–973, arXiv:1603.00034.
- [151] H. Yajima, J.-H. Choi, K. Nagamine, Mon. Not. R. Astron. Soc. 412 (2011) 411–422.
- [152] M. Vogelsberger, S. Genel, V. Springel, et al., Mon. Not. R. Astron. Soc. 444 (2014) 1518–1547.
- [153] Y. Feng, T. Di-Matteo, R.A. Croft, S. Bird, N. Battaglia, S. Wilkins, Mon. Not. R. Astron. Soc. 455 (2016) 2778–2791.
- [154] C.G. Lacey, C.M. Baugh, C.S. Frenk, et al., Mon. Not. R. Astron. Soc. 462 (2016) 3854–3911, arXiv:1509.08473.
- [155] V. Springel, L. Hernquist, Mon. Not. R. Astron. Soc. 339 (2003) 289–311.
- [156] F. Haardt, P. Madau, in: D.M. Neumann, J.T.V. Tran (Eds.), Clusters of Galaxies and the High Redshift Universe Observed in X-rays, 2001, arXiv:astro-ph/0106018.
- [157] C.-A. Faucher-Giguère, A. Lidz, M. Zaldarriaga, L. Hernquist, Astrophys. J. 703 (2009) 1416–1443, arXiv:0901.4554.
- [158] H.E. Bond, Astrophys. J. 248 (1981) 606–611.
- [159] N. Yoshida, K. Omukai, L. Hernquist, Astrophys. J. Lett. 667 (2007) L117–L120, arXiv:0706.3597.
- [160] V. Bromm, R.B. Larson, Annu. Rev. Astron. Astrophys. 42 (2004) 79–118.
- [161] S. Glover, Space Sci. Rev. 117 (2005) 445–508.
- [162] V. Bromm, N. Yoshida, L. Hernquist, C.F. McKee, Nature 459 (2009) 49–54.
- [163] V. Bromm, N. Yoshida, Annu. Rev. Astron. Astrophys. 49 (2011) 373–407.
- [164] J.H. Wise, First Light: A Brief Review, 2012, ArXiv e-prints, arXiv:1201.4820.
- [165] S.M. Fall, G. Efstathiou, Mon. Not. R. Astron. Soc. 193 (1980) 189–206.
- [166] M.A. Fardal, N. Katz, J.P. Gardner, L. Hernquist, D.H. Weinberg, R. Dave, Astrophys. J. 562 (2001) 605–617, arXiv:astro-ph/0007205.
- [167] Y. Birnboim, A. Dekel, Mon. Not. R. Astron. Soc. 345 (2003) 349–364.
- [168] D. Kereš, N. Katz, D.H. Weinberg, R. Davé, Mon. Not. R. Astron. Soc. 363 (2005) 2–28.
- [169] P. Ocvirk, C. Pichon, R. Teyssier, Mon. Not. R. Astron. Soc. 390 (2008) 1326–1338.
- [170] A. Dekel, Y. Birnboim, Mon. Not. R. Astron. Soc. 368 (2006) 2–20.
- [171] A. Dekel, Y. Birnboim, G. Engel, J. Freundlich, T. Goerdt, M. Mumcuoglu, E. Neistein, C. Pichon, R. Teyssier, E. Zinger, Nature 457 (2009) 451–454, arXiv:0808.0553.
- [172] D. Kereš, N. Katz, M. Fardal, R. Davé, D.H. Weinberg, Mon. Not. R. Astron. Soc. 395 (2009) 160–179.
- [173] F. van de Voort, J. Schaye, C.M. Booth, M.R. Haas, C. Dalla Vecchia, Mon. Not. R. Astron. Soc. 414 (2011) 2458–2478.
- [174] F. van de Voort, J. Schaye, G. Altay, T. Theuns, Mon. Not. R. Astron. Soc. 421 (2012) 2809–2819.
- [175] A. Dekel, A. Zolotov, D. Tweed, M. Cacciato, D. Ceverino, J.R. Primack, Mon. Not. R. Astron. Soc. 435 (2013) 999–1019.
- [176] A.J. Benson, R. Bower, Mon. Not. R. Astron. Soc. 410 (2011) 2653–2661.
- [177] A.M. Brooks, F. Governato, T. Quinn, C.B. Brook, J. Wadsley, Astrophys. J. 694 (2009) 396–410.
- [178] G. Kauffmann, T.M. Heckman, S.D.M. White, et al., Mon. Not. R. Astron. Soc. 341 (2003) 54–69.
- [179] C.-A. Faucher-Giguère, D. Kereš, Mon. Not. R. Astron. Soc. 412 (2011) L118–L122.
- [180] T. Goerdt, A. Dekel, A. Sternberg, O. Gnat, D. Ceverino, Mon. Not. R. Astron. Soc. 424 (2012) 2292–2315.
- [181] K.H.R. Rubin, J.X. Prochaska, D.C. Koo, A.C. Phillips, Astrophys. J. Lett. 747 (2012) L26.
- [182] O. Rakic, J. Schaye, C.C. Steidel, G.C. Rudie, Astrophys. J. 751 (2012) 94.
- [183] M. Rauch, G.D. Becker, M.G. Haehnelt, J.-R. Gauthier, S. Ravindranath, W.L.W. Sargent, Mon. Not. R. Astron. Soc. 418 (2011) 1115–1126.
- [184] M. Rauch, G.D. Becker, M.G. Haehnelt, J.-R. Gauthier, W.L.W. Sargent, Mon. Not. R. Astron. Soc. 429 (2013) 429–443.
- [185] D.C. Martin, M. Matuszewski, P. Morrissey, J.D. Neill, A. Moore, S. Cantalupo, J.X. Prochaska, D. Chang, Nature 524 (2015) 192–195.
- [186] D.C. Martin, M. Matuszewski, P. Morrissey, J.D. Neill, A. Moore, C.C. Steidel, R. Trainor, Astrophys. J. Lett. 824 (2016) L5.
- [187] C.C. Steidel, K.L. Adelberger, A.E. Shapley, M. Pettini, M. Dickinson, M. Giavalisco, Astrophys. J. 532 (2000) 170–182.
- [188] M. Dijkstra, Z. Haiman, M. Spaans, Astrophys. J. 649 (2006) 37–47.
- [189] M. Dijkstra, A. Loeb, Mon. Not. R. Astron. Soc. 400 (2009) 1109–1120.
- [190] T. Goerdt, A. Dekel, A. Sternberg, D. Ceverino, R. Teyssier, J.R. Primack, Mon. Not. R. Astron. Soc. 407 (2010) 613–631.
- [191] J. Rosdahl, J. Blaizot, Mon. Not. R. Astron. Soc. 423 (2012) 344–366.
- [192] K.K. Nilsson, J.P.U. Fynbo, P. Møller, J. Sommer-Larsen, C. Ledoux, Astron. Astrophys. 452 (2006) L23–L26.
- [193] D.J.B. Smith, M.J. Jarvis, Mon. Not. R. Astron. Soc. 378 (2007) L49–L53.
- [194] D.J.B. Smith, M.J. Jarvis, M. Lacy, A. Martínez-Sansigre, Mon. Not. R. Astron. Soc. 389 (2008) 799–805.
- [195] D.K. Erb, M. Bogosavljević, C.C. Steidel, Astrophys. J. Lett. 740 (2011) L31.
- [196] S. Cantalupo, in: A. Fox, R. Davé (Eds.), Gas Accretion onto Galaxies, in: Astrophysics and Space Science Library, vol. 430, 2017, p. 195, arXiv: 1612.00491.
- [197] M. Fumagalli, J.X. Prochaska, D. Kasen, A. Dekel, D. Ceverino, J.R. Primack, Mon. Not. R. Astron. Soc. 418 (2011) 1796–1821.
- [198] M. Fumagalli, J.M. O'Meara, J.X. Prochaska, Mon. Not. R. Astron. Soc. 455 (2016) 4100–4121.
- [199] N. Lehner, J.M. O'Meara, J.C. Howk, J.X. Prochaska, M. Fumagalli, The Cosmic Evolution of the Metallicity Distribution of Ionized Gas Traced by Lyman Limit Systems, 2016, rXiv e-prints, arXiv:1608.02588.
- [200] J. Ribaldo, N. Lehner, J.C. Howk, J.K. Werk, T.M. Tripp, J.X. Prochaska, J.D. Meiring, J. Tumlinson, Astrophys. J. 743 (2011) 207.
- [201] N.H.M. Crighton, J.F. Hennawi, J.X. Prochaska, Astrophys. J. Lett. 776 (2013) L18.
- [202] M. Fumagalli, S. Cantalupo, A. Dekel, S.L. Morris, J.M. O'Meara, J.X. Prochaska, T. Theuns, Mon. Not. R. Astron. Soc. 462 (2016) 1978–1988.
- [203] M. Giavalisco, E. Vanziella, S. Salimbeni, et al., Astrophys. J. 743 (2011) 95.
- [204] C.L. Martin, A.E. Shapley, A.L. Coil, K.A. Kornei, K. Bundy, B.J. Weiner, K.G. Noeske, D. Schiminovich, Astrophys. J. 760 (2012) 127.
- [205] M. Tegmark, J. Silk, M.J. Rees, A. Blanchard, T. Abel, F. Palla, Astrophys. J. 474 (1997) 1, arXiv:astro-ph/9603007.
- [206] W.C. Saslaw, D. Zipoy, Nature 216 (1967) 976–978.
- [207] M.R.C. McDowell, The Observatory 81 (1961) 240–243.
- [208] S. Seager, D.D. Sasselov, D. Scott, Astrophys. J. Suppl. 128 (2000) 407–430, arXiv:astro-ph/9912182.
- [209] M.-M. Mac Low, J.M. Shull, Astrophys. J. 302 (1986) 585–589.
- [210] T. Abel, G.L. Bryan, M.L. Norman, Astrophys. J. 540 (2000) 39–44, arXiv:astro-ph/0002135.
- [211] T. Abel, G.L. Bryan, M.L. Norman, Science 295 (2002) 93–98, arXiv:astro-ph/0112088.
- [212] V. Bromm, P.S. Coppi, R.B. Larson, Astrophys. J. Lett. 527 (1999) L5–L8, arXiv:astro-ph/9910224.
- [213] V. Bromm, P.S. Coppi, R.B. Larson, Astrophys. J. 564 (2002) 23–51, arXiv:astro-ph/0102503.
- [214] R. Ebert, Z. Phys. 37 (1955) 217.
- [215] W.B. Bonnor, Mon. Not. R. Astron. Soc. 116 (1956) 351.
- [216] F. Palla, E.E. Salpeter, S.W. Stahler, Astrophys. J. 271 (1983) 632–641.
- [217] N. Yoshida, K. Omukai, L. Hernquist, T. Abel, Astrophys. J. 652 (2006) 6–25, arXiv:astro-ph/0606106.
- [218] N. Yoshida, K. Omukai, L. Hernquist, Science 321 (2008) 669, arXiv:0807.4928.
- [219] C.F. McKee, J.C. Tan, Astrophys. J. 681 (2008) 771–797, arXiv:0711.1377.

- [220] V. Bromm, A. Loeb, *New Astron.* 9 (2004) 353–364, [arXiv:astro-ph/0312456](#).
- [221] M.R. Bate, I.A. Bonnell, N.M. Price, *Mon. Not. R. Astron. Soc.* 277 (1995) 362–376, [arXiv:astro-ph/9510149](#).
- [222] A. Stacy, T.H. Greif, V. Bromm, *Mon. Not. R. Astron. Soc.* 403 (2010) 45–60, [arXiv:0908.0712](#).
- [223] P.C. Clark, S.C.O. Glover, R.S. Klessen, V. Bromm, *Astrophys. J.* 727 (2011) 110, [arXiv:1006.1508](#).
- [224] P.C. Clark, S.C.O. Glover, R.J. Smith, T.H. Greif, R.S. Klessen, V. Bromm, *Science* 331 (2011) 1040, [arXiv:1101.5284](#).
- [225] M.J. Turk, T. Abel, B. O’Shea, *Science* 325 (2009) 601, [arXiv:0907.2919](#).
- [226] M.N. Machida, K. Omukai, T. Matsumoto, S.-I. Inutsuka, *Mon. Not. R. Astron. Soc.* 399 (2009) 1255–1263, [arXiv:0907.3257](#).
- [227] T.H. Greif, V. Springel, S.D.M. White, S.C.O. Glover, P.C. Clark, R.J. Smith, R.S. Klessen, V. Bromm, *Astrophys. J.* 737 (2011) 75, [arXiv:1101.5491](#).
- [228] J. Prieto, P. Padoan, R. Jimenez, L. Infante, *Astrophys. J. Lett.* 731 (2011) L38, [arXiv:1101.5163](#).
- [229] H. Xu, J.H. Wise, M.L. Norman, *Astrophys. J.* 773 (2013) 83, [arXiv:1305.1325](#).
- [230] M.J. Turk, J.S. Oishi, T. Abel, G.L. Bryan, *Astrophys. J.* 745 (2012) 154, [arXiv:1112.4479](#).
- [231] S. Hirano, V. Bromm, *Mon. Not. R. Astron. Soc.* 470 (2017) 898–914, [arXiv:1612.06387](#).
- [232] K. Omukai, R. Nishi, *Astrophys. J.* 508 (1998) 141–150, [arXiv:astro-ph/9811308](#).
- [233] B.W. O’Shea, M.L. Norman, *Astrophys. J.* 654 (2007) 66–92, [arXiv:astro-ph/0607013](#).
- [234] **J.H. Wise, M.J. Turk, T. Abel, *Astrophys. J.* 682 (2008) 745–757.**
- [235] K. Omukai, F. Palla, *Astrophys. J.* 589 (2003) 677–687, [arXiv:astro-ph/0302345](#).
- [236] S. Hirano, T. Hosokawa, N. Yoshida, H. Umeda, K. Omukai, G. Chiaki, H.W. Yorke, *Astrophys. J.* 781 (2014) 60, [arXiv:1308.4456](#).
- [237] H. Susa, K. Hasegawa, N. Tominaga, *Astrophys. J.* 792 (2014) 32, [arXiv:1407.1374](#).
- [238] T. Hosokawa, K. Omukai, N. Yoshida, H.W. Yorke, *Science* 334 (2011) 1250, [arXiv:1111.3649](#).
- [239] A. Stacy, T.H. Greif, V. Bromm, *Mon. Not. R. Astron. Soc.* 422 (2012) 290–309, [arXiv:1109.3147](#).
- [240] A. Stacy, V. Bromm, A.T. Lee, *Mon. Not. R. Astron. Soc.* 462 (2016) 1307–1328, [arXiv:1603.09475](#).
- [241] **J.H. Wise, T. Abel, *Astrophys. J.* 685 (2008) 40–56.**
- [242] **T.H. Greif, S.C.O. Glover, V. Bromm, R.S. Klessen, *Astrophys. J.* 716 (2010) 510–520.**
- [243] E. Chatzopoulos, J.C. Wheeler, *Astrophys. J.* 760 (2012) 154, [arXiv:1210.1617](#).
- [244] C.L. Fryer, S.E. Woosley, A. Heger, *Astrophys. J.* 550 (2001) 372–382, [arXiv:astro-ph/0007176](#).
- [245] A. Heger, S.E. Woosley, *Astrophys. J.* 567 (2002) 532–543, [arXiv:astro-ph/0107037](#).
- [246] V. Bromm, A. Loeb, *Nature* 425 (2003) 812–814, [arXiv:astro-ph/0310622](#).
- [247] G. Chiaki, H. Susa, S. Hirano, *Mon. Not. R. Astron. Soc.* 475 (2018) 4378–4395, [arXiv:1801.01583](#).
- [248] T.H. Greif, J.L. Johnson, V. Bromm, R.S. Klessen, *Astrophys. J.* 670 (2007) 1–14, [arXiv:0705.3048](#).
- [249] R. Cen, M.A. Riquelme, *Astrophys. J.* 674 (2008) 644–652, [arXiv:0704.3545](#).
- [250] **V. Bromm, A. Ferrara, P.S. Coppi, R.B. Larson, *Mon. Not. R. Astron. Soc.* 328 (2001) 969–976.**
- [251] **R. Schneider, A. Ferrara, P. Natarajan, K. Omukai, *Astrophys. J.* 571 (2002) 30–39.**
- [252] R. Schneider, K. Omukai, A.K. Inoue, A. Ferrara, *Mon. Not. R. Astron. Soc.* 369 (2006) 1437–1444, [arXiv:astro-ph/0603766](#).
- [253] F. Santoro, J.M. Shull, *Astrophys. J.* 643 (2006) 26–37.
- [254] B.D. Smith, S. Sigurdsson, *Astrophys. J. Lett.* 661 (2007) L5–L8.
- [255] R. Schneider, K. Omukai, S. Bianchi, R. Valiante, *Mon. Not. R. Astron. Soc.* 419 (2012) 1566–1575.
- [256] G. Chiaki, R. Schneider, T. Nozawa, K. Omukai, M. Limongi, N. Yoshida, A. Chieffi, *Mon. Not. R. Astron. Soc.* 439 (2014) 3121–3127, [arXiv:1401.5057](#).
- [257] T. Karlsson, J.L. Johnson, V. Bromm, *Astrophys. J.* 679 (2008) 6–16.
- [258] U. Maio, S. Khochfar, J.L. Johnson, B. Ciardi, *Mon. Not. R. Astron. Soc.* 414 (2011) 1145–1157.
- [259] M. Jeon, V. Bromm, A.H. Pawlik, M. Milosavljević, *Mon. Not. R. Astron. Soc.* 452 (2015) 1152–1170, [arXiv:1501.01002](#).
- [260] J. Mackey, V. Bromm, L. Hernquist, *Astrophys. J.* 586 (2003) 1–11, [arXiv:astro-ph/0208447](#).
- [261] S.E. Woosley, T.A. Weaver, *Annu. Rev. Astron. Astrophys.* 24 (1986) 205–253.
- [262] A.L. Muratov, O.Y. Gnedin, N.Y. Gnedin, M. Zemp, *Astrophys. J.* 773 (2013) 19, [arXiv:1212.1157](#).
- [263] D. Schaerer, *Astron. Astrophys.* 397 (2003) 527–538.
- [264] T. Kitayama, N. Yoshida, H. Susa, M. Umemura, *Astrophys. J.* 613 (2004) 631–645, [arXiv:astro-ph/0406280](#).
- [265] D. Whalen, T. Abel, M.L. Norman, *Astrophys. J.* 610 (2004) 14–22, [arXiv:astro-ph/0310283](#).
- [266] M.A. Alvarez, V. Bromm, P.R. Shapiro, *Astrophys. J.* 639 (2006) 621–632.
- [267] J.L. Johnson, T.H. Greif, V. Bromm, *Astrophys. J.* 665 (2007) 85–95, [arXiv:astro-ph/0612254](#).
- [268] B. Ciardi, A. Ferrara, T. Abel, *Astrophys. J.* 533 (2000) 594–600.
- [269] B.T. Draine, F. Bertoldi, *Astrophys. J.* 468 (1996) 269, [arXiv:astro-ph/9603032](#).
- [270] J.A. Regan, M.G. Haehnelt, *Mon. Not. R. Astron. Soc.* 396 (2009) 343–353.
- [271] M. Volonteri, *Science* 337 (2012) 544.
- [272] X. Fan, C.L. Carilli, B. Keating, *Annu. Rev. Astron. Astrophys.* 44 (2006) 415–462.
- [273] S.P. Oh, K.M. Nollett, P. Madau, G.J. Wasserburg, *Astrophys. J. Lett.* 562 (2001) L1–L4, [arXiv:astro-ph/0109400](#).
- [274] M.E. Machacek, G.L. Bryan, T. Abel, *Mon. Not. R. Astron. Soc.* 338 (2003) 273–286, [arXiv:astro-ph/0209326](#).
- [275] S.P. Oh, Z. Haiman, *Mon. Not. R. Astron. Soc.* 346 (2003) 456–472, [arXiv:astro-ph/0307135](#).
- [276] A. Frebel, J.E. Norris, *Annu. Rev. Astron. Astrophys.* 53 (2015) 631–688, [arXiv:1501.06921](#).
- [277] R.B. Partridge, P.J.E. Peebles, *Agron. J.* 147 (1967) 868.
- [278] N. Kashikawa, T. Nagao, J. Toshikawa, et al., *Astrophys. J.* 761 (2012) 85, [arXiv:1210.4933](#).
- [279] J.L. Johnson, T.H. Greif, V. Bromm, R.S. Klessen, J. Ippolito, *Mon. Not. R. Astron. Soc.* 399 (2009) 37–47.
- [280] E. Zackrisson, C.-E. Rydberg, D. Schaerer, G. Östlin, M. Tuli, *Astrophys. J.* 740 (2011) 13.
- [281] A.H. Pawlik, M. Milosavljević, V. Bromm, *Astrophys. J.* 731 (2011) 54.
- [282] E. Zackrisson, A.K. Inoue, C.-E. Rydberg, F. Duval, *Mon. Not. R. Astron. Soc.* 418 (2011) L104–L108.
- [283] R.A. Windhorst, F.X. Timmes, J.S.B. Wyithe, et al., *Astrophys. J. Suppl.* 234 (2018) 41, [arXiv:1801.03584](#).
- [284] A. Mesinger, B.D. Johnson, Z. Haiman, *Astrophys. J.* 637 (2006) 80–90, [arXiv:astro-ph/0505110](#).
- [285] E. Visbal, Z. Haiman, G.L. Bryan, *Mon. Not. R. Astron. Soc.* 450 (2015) 2506–2513, [arXiv:1501.03177](#).
- [286] E.W. Kolb, M.S. Turner, *Front. Phys.*, Vol. 69., 1990.
- [287] F. Iocco, G. Mangano, G. Miele, O. Pisanti, P.D. Serpico, *Phys. Rev. D* 75 (8) (2007) 087304, [arXiv:astro-ph/0702090](#).
- [288] S. Naoz, S. Noter, R. Barkana, *Mon. Not. R. Astron. Soc.* 373 (2006) L98–L102, [arXiv:astro-ph/0604050](#).
- [289] M. Asplund, N. Grevesse, A.J. Sauval, in: T.G. Barnes, F.N. Bash (Eds.), *Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis*, in: *Astronomical Society of the Pacific Conference Series*, 336, 2005, p. 25.
- [290] A. Ferrara, Understanding the Epoch of Cosmic Reionization: Challenges and Progress 423 (2016) 163, [arXiv:1511.01120](#).
- [291] U. Maio, B. Ciardi, K. Dolag, L. Tornatore, S. Khochfar, *Mon. Not. R. Astron. Soc.* 407 (2010) 1003–1015, [arXiv:1003.4992](#).
- [292] L. Tornatore, A. Ferrara, R. Schneider, *Mon. Not. R. Astron. Soc.* 382 (2007) 945–950.
- [293] B.D. Oppenheimer, R. Davé, K. Finlator, *Mon. Not. R. Astron. Soc.* 396 (2009) 729–758, [arXiv:0901.0286](#).

- [294] A. Pallottini, A. Ferrara, S. Gallerani, S. Salvadori, V. D'Odorico, Mon. Not. R. Astron. Soc. 440 (2014) 2498–2518, [arXiv:1403.1261](#).
- [295] E.E. Salpeter, *Astrophys. J.* 121 (1955) 161.
- [296] T. Tsujimoto, K. Nomoto, Y. Yoshii, M. Hashimoto, S. Yanagida, F.-K. Thielemann, Mon. Not. R. Astron. Soc. 277 (1995) 945–958.
- [297] S.E. Woosley, T.A. Weaver, *Astrophys. J. Suppl.* 101 (1995) 181.
- [298] K. Nomoto, N. Tominaga, H. Umeda, C. Kobayashi, K. Maeda, Nuclear Phys. A 777 (2006) 424–458, [arXiv:astro-ph/0605725](#).
- [299] B.K. Gibson, M. Loewenstein, R.F. Mushotzky, Mon. Not. R. Astron. Soc. 290 (1997) 623–628, [arXiv:astro-ph/9706138](#).
- [300] S.F. Gull, Mon. Not. R. Astron. Soc. 161 (1973) 47–69.
- [301] P. Madau, A. Ferrara, M.J. Rees, *Astrophys. J.* 555 (2001) 92–105, [arXiv:astro-ph/0010158](#).
- [302] J.D. Slavin, J.M. Shull, M.C. Begelman, *Astrophys. J.* 407 (1993) 83–99.
- [303] G. Tenorio-Tagle, *Astron. J.* 111 (1996) 1641.
- [304] G. Pezzulli, F. Fraternali, Mon. Not. R. Astron. Soc. 455 (2016) 2308–2322, [arXiv:1510.04289](#).
- [305] J.-R. Roy, D. Kunth, *Astron. Astrophys.* 294 (1995) 432–442, [arXiv:astro-ph/9410023](#).
- [306] M.A. de Avillez, M.-M. Mac Low, *Astrophys. J.* 581 (2002) 1047–1060, [arXiv:astro-ph/0208441](#).
- [307] R.S. Klessen, D.N. Lin, *Phys. Rev. E* 67 (4) (2003) 046311, [arXiv:astro-ph/0302527](#).
- [308] L. Pan, E. Scannapieco, *Astrophys. J.* 721 (2010) 1765–1782, [arXiv:1008.0665](#).
- [309] Y. Feng, M.R. Krumholz, *Nature* 513 (2014) 523–525, [arXiv:1408.6543](#).
- [310] L. Armillotta, M.R. Krumholz, Y. Fujimoto, Mixing of metals during star cluster formation: statistics and implications for chemical tagging, 2018, ArXiv e-prints, [arXiv:1807.01712](#).
- [311] R. Cen, M.A. Riquelme, *Astrophys. J.* 674 (2008) 644–652, [arXiv:0704.3545](#).
- [312] M. Jeon, V. Bromm, A.H. Pawlik, M. Milosavljević, Mon. Not. R. Astron. Soc. 452 (2015) 1152–1170, [arXiv:1501.01002](#).
- [313] B.D. Smith, J.H. Wise, B.W. O'Shea, M.L. Norman, S. Khochfar, Mon. Not. R. Astron. Soc. 452 (2015) 2822–2836, [arXiv:1504.07639](#).
- [314] R. Sarmento, E. Scannapieco, S. Cohen, *Astrophys. J.* 854 (2018) 75, [arXiv:1710.09878](#).
- [315] A.J. Benson, R.G. Bower, C.S. Frenk, C.G. Lacey, C.M. Baugh, S. Cole, *Astrophys. J.* 599 (2003) 38–49, [arXiv:astro-ph/0302450](#).
- [316] V. Springel, L. Hernquist, Mon. Not. R. Astron. Soc. 339 (2003) 312–334, [arXiv:astro-ph/0206395](#).
- [317] M.-M. Mac Low, A. Ferrara, *Astrophys. J.* 513 (1999) 142–155, [arXiv:astro-ph/9801237](#).
- [318] M. Tegmark, J. Silk, A. Evrard, *Astrophys. J.* 417 (1993) 54, [arXiv:astro-ph/9306021](#).
- [319] M.V. Medvedev, A. Loeb, *Astrophys. J.* 768 (2013) 113, [arXiv:1212.0330](#).
- [320] J. Bremer, P. Dayal, E.V. Ryan-Weber, Mon. Not. R. Astron. Soc. (2018) [arXiv:1803.05450](#).
- [321] R.C. Livermore, S.L. Finkelstein, J.M. Lotz, *Astrophys. J.* 835 (2017) 113, [arXiv:1604.06799](#).
- [322] P.R. Shapiro, M.L. Giroux, *Astrophys. J. Lett.* 321 (1987) L107–L112.
- [323] C. Leitherer, D. Schaerer, J.D. Goldader, et al., *Astrophys. J. Suppl.* 123 (1999) 3–40, [arXiv:astro-ph/9902334](#).
- [324] C. Leitherer, P.A. Ortiz Otálvaro, F. Bresolin, R.-P. Kudritzki, B. Lo Faro, A.W.A. Pauldrach, M. Pettini, S.A. Rix, *Astrophys. J. Suppl.* 189 (2010) 309–335, [arXiv:1006.5624](#).
- [325] P. Dayal, T.R. Choudhury, V. Bromm, F. Pacucci, *Astrophys. J.* 836 (2017) 16, [arXiv:1501.02823](#).
- [326] E. Zackrisson, N. Bergvall, E. Leitet, *Astrophys. J. Lett.* 676 (2008) L9, [arXiv:0802.3696](#).
- [327] D. Schaerer, S. de Barros, *Astron. Astrophys.* 502 (2009) 423–426.
- [328] S. de Barros, D. Schaerer, D.P. Stark, *Astron. Astrophys.* 563 (2014) A81, [arXiv:1207.3663](#).
- [329] A.K. Inoue, I. Iwata, Mon. Not. R. Astron. Soc. 387 (2008) 1681–1692, [arXiv:0804.2951](#).
- [330] J.X. Prochaska, J.M. O'Meara, G. Worseck, *Astrophys. J.* 718 (2010) 392–416, [arXiv:0912.0292](#).
- [331] B. Siana, H.I. Teplitz, J. Colbert, et al., *Astrophys. J.* 668 (2007) 62–73.
- [332] D. Calzetti, *Astron. J.* 113 (1997) 162–184, [arXiv:astro-ph/9610184](#).
- [333] A.K. Inoue, I. Iwata, J.-M. Deharveng, Mon. Not. R. Astron. Soc. 371 (2006) L1–L5.
- [334] D.B. Nestor, A.E. Shapley, C.C. Steidel, B. Siana, *Astrophys. J.* 736 (2011) 18.
- [335] D.B. Nestor, A.E. Shapley, K.A. Kornei, C.C. Steidel, B. Siana, *Astrophys. J.* 765 (2013) 47.
- [336] J. Matthee, D. Sobral, P. Best, A.A. Khostovan, I. Oteo, R. Bouwens, H. Röttgering, Mon. Not. R. Astron. Soc. 465 (2017) 3637–3655, [arXiv:1605.08782](#).
- [337] K. Vasei, B. Siana, A.E. Shapley, A.M. Quider, A. Alavi, M. Rafelski, C.C. Steidel, M. Pettini, G.F. Lewis, *Astrophys. J.* 831 (2016) 38, [arXiv:1603.02309](#).
- [338] B. Siana, A.E. Shapley, K.R. Kulas, et al., *Astrophys. J.* 804 (2015) 17.
- [339] C.C. Steidel, K.L. Adelberger, M. Giavalisco, M. Dickinson, M. Pettini, *Astrophys. J.* 519 (1999) 1–17.
- [340] K. Boutsia, A. Grazian, E. Giallongo, et al., *Astrophys. J.* 736 (2011) 41.
- [341] A. Grazian, E. Giallongo, R. Gerbasi, et al., *Astron. Astrophys.* 585 (2016) A48.
- [342] E. Vanzella, M. Giavalisco, A.K.e.a. Inoue, *Astrophys. J.* 725 (2010) 1011–1031.
- [343] R.P. Naidu, P.A. Oesch, N. Reddy, et al., *Astrophys. J.* 847 (2017) 12, [arXiv:1611.07038](#).
- [344] R.E. Mostardi, A.E. Shapley, C.C. Steidel, R.F. Trainor, N.A. Reddy, B. Siana, *Astrophys. J.* 810 (2015) 107, [arXiv:1506.08201](#).
- [345] A.E. Shapley, C.C. Steidel, A.L. Strom, M. Bogosavljević, N.A. Reddy, B. Siana, R.E. Mostardi, G.C. Rudie, *Astrophys. J. Lett.* 826 (2016) L24, [arXiv:1606.00443](#).
- [346] E. Vanzella, S. de Barros, K. Vasei, et al., *Astrophys. J.* 825 (2016) 41.
- [347] E. Vanzella, M. Nonino, G. Cupani, et al., Mon. Not. R. Astron. Soc. 476 (2018) L15–L19, [arXiv:1712.07661](#).
- [348] N. Leethochawalit, T.A. Jones, R.S. Ellis, D.P. Stark, A. Zitrin, *Astrophys. J.* 831 (2016) 152, [arXiv:1606.05309](#).
- [349] I. Iwata, A.K. Inoue, Y. Matsuda, et al., *Astrophys. J.* 692 (2009) 1287–1293.
- [350] E. Vanzella, Y. Guo, M. Giavalisco, et al., *Astrophys. J.* 751 (2012) 70.
- [351] J. Cooke, E.V. Ryan-Weber, T. Garel, C.G. Díaz, Mon. Not. R. Astron. Soc. 441 (2014) 837–851.
- [352] E. Komatsu, K.M. Smith, J. Dunkley, et al., *Astrophys. J. Suppl.* 192 (2011) 18.
- [353] M. Kuhlen, C.-A. Faucher-Giguère, Mon. Not. R. Astron. Soc. 423 (2012) 862–876.
- [354] S. Mitra, A. Ferrara, T.R. Choudhury, Mon. Not. R. Astron. Soc. 428 (2013) L1–L5.
- [355] B.E. Robertson, S.R. Furlanetto, E. Schneider, et al., *Astrophys. J.* 768 (2013) 71.
- [356] S. Mitra, T.R. Choudhury, A. Ferrara, Mon. Not. R. Astron. Soc. 454 (2015) L76–L80.
- [357] G. Sun, S.R. Furlanetto, Mon. Not. R. Astron. Soc. 460 (2016) 417–433, [arXiv:1512.06219](#).
- [358] B.E. Robertson, R.S. Ellis, S.R. Furlanetto, J.S. Dunlop, *Astrophys. J. Lett.* 802 (2015) L19.
- [359] M. Sharma, T. Theuns, C. Frenk, R. Bower, R. Crain, M. Schaller, J. Schaye, Mon. Not. R. Astron. Soc. 458 (2016) L94–L98.
- [360] A. Ferrara, A. Loeb, Mon. Not. R. Astron. Soc. 431 (2013) 2826–2833.
- [361] J.H. Wise, V.G. Demchenko, M.T. Halicek, M.L. Norman, M.J. Turk, T. Abel, B.D. Smith, Mon. Not. R. Astron. Soc. 442 (2014) 2560–2579.
- [362] X. Ma, D. Kasen, P.F. Hopkins, C.-A. Faucher-Giguère, E. Quataert, D. Kereš, N. Murray, Mon. Not. R. Astron. Soc. 453 (2015) 960–975.
- [363] J.H. Wise, R. Cen, *Astrophys. J.* 693 (2009) 984–999.
- [364] J.-P. Paardekooper, F.I. Pelupessy, G. Altay, C.J.H. Kruip, *Astron. Astrophys.* 530 (2011) A87.
- [365] J.-P. Paardekooper, S. Khochfar, C. Dalla Vecchia, Mon. Not. R. Astron. Soc. 429 (2013) L94–L98.

- [366] J.-P. Paardekooper, S. Khochfar, C. Dalla Vecchia, Mon. Not. R. Astron. Soc. 451 (2015) 2544–2563, [arXiv:1501.01967](#).
- [367] A.O. Razoumov, J. Sommer-Larsen, *Astrophys. J.* 668 (2007) 674–681.
- [368] A.O. Razoumov, J. Sommer-Larsen, *Astrophys. J.* 710 (2010) 1239–1246.
- [369] N.Y. Gnedin, A.V. Kravtsov, H.-W. Chen, *Astrophys. J.* 672 (2008) 765–775.
- [370] A.O. Razoumov, J. Sommer-Larsen, *Astrophys. J. Lett.* 651 (2006) L89–L92.
- [371] M. Trebitsch, J. Blaizot, J. Rosdahl, J. Devriendt, A. Slyz, Mon. Not. R. Astron. Soc. 470 (2017) 224–239, [arXiv:1705.00941](#).
- [372] X. Ma, P.F. Hopkins, D. Kasen, E. Quataert, C.-A. Faucher-Giguère, D. Kereš, N. Murray, A. Strom, Mon. Not. R. Astron. Soc. 459 (2016) 3614–3619.
- [373] J. Rosdahl, H. Katz, J. Blaizot, et al., Mon. Not. R. Astron. Soc. 479 (2018) 994–1016, [arXiv:1801.07259](#).
- [374] C. Conroy, K.M. Kratter, *Astrophys. J.* 755 (2012) 123.
- [375] T. Kimm, R. Cen, *Astrophys. J.* 788 (2014) 121.
- [376] T.H. Greif, J.L. Johnson, R.S. Klessen, V. Bromm, Mon. Not. R. Astron. Soc. 387 (2008) 1021–1036.
- [377] S. Sur, E. Scannapieco, E.C. Ostriker, *Astrophys. J.* 818 (2016) 28.
- [378] M. Safarzadeh, E. Scannapieco, *Astrophys. J. Lett.* 832 (2016) L9, [arXiv:1611.00372](#).
- [379] E. Zackrisson, A.K. Inoue, H. Jensen, *Astrophys. J.* 777 (2013) 39.
- [380] K. Nakajima, M. Ouchi, Mon. Not. R. Astron. Soc. 442 (2014) 900–916, [arXiv:1309.0207](#).
- [381] A.L. Faisst, *Astrophys. J.* 829 (2016) 99, [arXiv:1605.06507](#).
- [382] L. Mas-Ribas, J.F. Hennawi, M. Dijkstra, F.B. Davies, J. Stern, H.-W. Rix, *Astrophys. J.* 846 (2017) 11, [arXiv:1707.07713](#).
- [383] J. Miralda-Escudé, M.J. Rees, Mon. Not. R. Astron. Soc. 266 (1994) 343–352.
- [384] A. Meiksin, P. Madau, *Astrophys. J.* 412 (1993) 34–55.
- [385] A. Maselli, S. Gallerani, A. Ferrara, T.R. Choudhury, Mon. Not. R. Astron. Soc. 376 (2007) L34–L38, [arXiv:astro-ph/0608209](#).
- [386] J.S. Bolton, M.G. Haehnelt, Mon. Not. R. Astron. Soc. 381 (2007) L35–L39, [arXiv:0705.3558](#).
- [387] A.-C. Eilers, F.B. Davies, J.F. Hennawi, J.X. Prochaska, Z. Lukić, C. Mazzucchelli, *Astrophys. J.* 840 (2017) 24, [arXiv:1703.02539](#).
- [388] I.T. Iliev, E. Scannapieco, P.R. Shapiro, *Astrophys. J.* 624 (2005) 491–504, [arXiv:astro-ph/0411035](#).
- [389] J. Miralda-Escudé, M. Haehnelt, M.J. Rees, *Astrophys. J.* 530 (2000) 1–16, [arXiv:astro-ph/9812306](#).
- [390] J. Chardin, E. Puchwein, M.G. Haehnelt, Mon. Not. R. Astron. Soc. 465 (2017) 3429–3445, [arXiv:1606.08231](#).
- [391] P.R. Shapiro, I.T. Iliev, A.C. Raga, Mon. Not. R. Astron. Soc. 348 (2004) 753–782, [arXiv:astro-ph/0307266](#).
- [392] I.T. Iliev, G. Mellema, P.R. Shapiro, U.-L. Pen, Mon. Not. R. Astron. Soc. 376 (2007) 534–548, [arXiv:astro-ph/0607517](#).
- [393] I.T. Iliev, P.R. Shapiro, A.C. Raga, Mon. Not. R. Astron. Soc. 361 (2005) 405–414, [arXiv:astro-ph/0408408](#).
- [394] P. Madau, F. Haardt, M.J. Rees, *Astrophys. J.* 514 (1999) 648–659, [arXiv:astro-ph/9809058](#).
- [395] A.A. Meiksin, *Rev. Modern Phys.* 81 (2009) 1405–1469, [arXiv:0711.3358](#).
- [396] J.S. Bolton, M.G. Haehnelt, Mon. Not. R. Astron. Soc. 382 (2007) 325–341, [arXiv:astro-ph/0703306](#).
- [397] S.R. Furlanetto, S.P. Oh, Mon. Not. R. Astron. Soc. 457 (2016) 1813–1827, [arXiv:1511.01521](#).
- [398] N.Y. Gnedin, *Astrophys. J.* 542 (2000) 535–541, [arXiv:astro-ph/0002151](#).
- [399] Y. Noh, M. McQuinn, Mon. Not. R. Astron. Soc. 444 (2014) 503–514, [arXiv:1401.0737](#).
- [400] M. Hoeft, G. Yepes, S. Gottlöber, V. Springel, Mon. Not. R. Astron. Soc. 371 (2006) 401–414, [arXiv:astro-ph/0501304](#).
- [401] T. Okamoto, L. Gao, T. Theuns, Mon. Not. R. Astron. Soc. 390 (2008) 920–928, [arXiv:0806.0378](#).
- [402] E. Sobacchi, A. Mesinger, Mon. Not. R. Astron. Soc. 432 (2013) L51–L55, [arXiv:1301.6776](#).
- [403] E. Sobacchi, A. Mesinger, Mon. Not. R. Astron. Soc. 432 (2013) 3340–3348, [arXiv:1301.6781](#).
- [404] J.H. Wise, T. Abel, *Astrophys. J.* 684 (2008) 1–17.
- [405] S. Naoz, R. Barkana, A. Mesinger, Mon. Not. R. Astron. Soc. 399 (2009) 369–376, [arXiv:0906.0349](#).
- [406] A.H. Pawlik, J. Schaye, E. van Scherpenzeel, Mon. Not. R. Astron. Soc. 394 (2009) 1812–1824, [arXiv:0807.3963](#).
- [407] K. Finlator, S.P. Oh, F. Ozel, R. Davé, Mon. Not. R. Astron. Soc. 427 (2012) 2464–2479, [arXiv:1209.2489](#).
- [408] J.M. Shull, A. Harness, M. Trenti, B.D. Smith, *Astrophys. J.* 747 (2012) 100.
- [409] Z. Haiman, *Nature* 472 (2011) 47–48, [arXiv:1104.1189](#).
- [410] G. Knevitt, G.A. Wynn, C. Power, J.S. Bolton, Mon. Not. R. Astron. Soc. 445 (2014) 2034–2048, [arXiv:1409.1267](#).
- [411] F. Haardt, P. Madau, *Astrophys. J.* 746 (2012) 125, [arXiv:1105.2039](#).
- [412] M. McQuinn, A. Lidz, O. Zahn, S. Dutta, L. Hernquist, M. Zaldarriaga, Mon. Not. R. Astron. Soc. 377 (2007) 1043–1063, [arXiv:astro-ph/0610094](#).
- [413] M.T. Busha, M.A. Alvarez, R.H. Wechsler, T. Abel, L.E. Strigari, *Astrophys. J.* 710 (2010) 408–420, [arXiv:0901.3553](#).
- [414] R. Lunnan, M. Vogelsberger, A. Frebel, L. Hernquist, A. Lidz, M. Boylan-Kolchin, *Astrophys. J.* 746 (2012) 109, [arXiv:1105.2293](#).
- [415] T.M. Brown, J. Tumlinson, M. Geha, et al., *Astrophys. J.* 796 (2014) 91, [arXiv:1410.0681](#).
- [416] E. Papastergis, A. Cattaneo, S. Huang, R. Giovanelli, M.P. Haynes, *Astrophys. J.* 759 (2012) 138, [arXiv:1208.5229](#).
- [417] R.J. Bouwens, G.D. Illingworth, P.A. Oesch, et al., *Agron. J.* 803 (2015) 34.
- [418] K.L. Dixon, I.T. Iliev, G. Mellema, K. Ahn, P.R. Shapiro, Mon. Not. R. Astron. Soc. 456 (2016) 3011–3029.
- [419] A.H. Pawlik, M. Milosavljević, V. Bromm, *Astrophys. J.* 767 (2013) 59, [arXiv:1208.3698](#).
- [420] A.H. Pawlik, J. Schaye, C. Dalla Vecchia, Mon. Not. R. Astron. Soc. 451 (2015) 1586–1605, [arXiv:1501.01980](#).
- [421] M. Raičević, T. Theuns, C. Lacey, Mon. Not. R. Astron. Soc. 410 (2011) 775–787.
- [422] F. Haardt, P. Madau, *Astrophys. J.* 461 (1996) 20, [arXiv:astro-ph/9509093](#).
- [423] C.M. Simpson, G.L. Bryan, K.V. Johnston, B.D. Smith, M.-M. Mac Low, S. Sharma, J. Tumlinson, Mon. Not. R. Astron. Soc. 432 (2013) 1989–2011, [arXiv:1211.1071](#).
- [424] T.R. Choudhury, A. Ferrara, Mon. Not. R. Astron. Soc. 380 (2007) L6–L10.
- [425] P. Chen, M.L. Norman, H. Xu, J.H. Wise, Fully Coupled Simulation of Cosmic Reionization. III. Stochastic Early Reionization by the Smallest Galaxies, 2017, ArXiv e-prints, [arXiv:1705.00026](#).
- [426] R. Salvaterra, A. Ferrara, P. Dayal, Mon. Not. R. Astron. Soc. 414 (2011) 847–859.
- [427] S.L. Finkelstein, C. Papovich, R.E. Ryan, et al., *Astrophys. J.* 758 (2012) 93.
- [428] R.J. Bouwens, G.D. Illingworth, P.A. Oesch, et al., *Astrophys. J. Lett.* 752 (2012) L5, [arXiv:1105.2038](#).
- [429] K. Duncan, C.J. Conselice, Mon. Not. R. Astron. Soc. 451 (2015) 2030–2049, [arXiv:1505.01846](#).
- [430] G.D. Becker, J.S. Bolton, Mon. Not. R. Astron. Soc. 436 (2013) 1023–1039, [arXiv:1307.2259](#).
- [431] P. Dayal, T.R. Choudhury, F. Pacucci, V. Bromm, Mon. Not. R. Astron. Soc. 472 (2017) 4414–4421, [arXiv:1705.00632](#).
- [432] S. Hassan, R. Davé, S. Mitra, K. Finlator, B. Ciardi, M.G. Santos, Mon. Not. R. Astron. Soc. 473 (2018) 227–240.
- [433] C. Liu, S.J. Mutch, P.W. Angel, A.R. Duffy, P.M. Geil, G.B. Poole, A. Mesinger, J.S.B. Wyithe, Mon. Not. R. Astron. Soc. 462 (2016) 235–249, [arXiv:1512.00563](#).
- [434] Y. Qin, S.J. Mutch, G.B. Poole, C. Liu, P.W. Angel, A.R. Duffy, P.M. Geil, A. Mesinger, J.S.B. Wyithe, Mon. Not. R. Astron. Soc. 472 (2017) 2009–2027, [arXiv:1703.04895](#).
- [435] I.T. Iliev, G. Mellema, P.R. Shapiro, U.-L. Pen, Y. Mao, J. Koda, K. Ahn, Mon. Not. R. Astron. Soc. 423 (2012) 2222–2253, [arXiv:1107.4772](#).
- [436] M. Volonteri, N.Y. Gnedin, *Astrophys. J.* 703 (2009) 2113–2117, [arXiv:0905.0144](#).

- [437] P. Madau, F. Haardt, *Astrophys. J. Lett.* 813 (2015) L8, [arXiv:1507.07678](#).
- [438] S. Mitra, T.R. Choudhury, A. Ferrara, *Mon. Not. R. Astron. Soc.* 473 (2018) 1416–1425, [arXiv:1606.02719](#).
- [439] E. Giallongo, A. Grazian, F. Fiore, et al., *Astron. Astrophys.* 578 (2015) A83.
- [440] A. D'Alloisio, P.R. Upton Sanderbeck, M. McQuinn, H. Trac, P.R. Shapiro, *Mon. Not. R. Astron. Soc.* 468 (2017) 4691–4701, [arXiv:1607.06467](#).
- [441] K. Kakichi, L. Graziani, B. Ciardi, A. Meiksin, M. Compostella, M.B. Eide, S. Zaroubi, *Mon. Not. R. Astron. Soc.* 468 (2017) 3718–3736, [arXiv:1607.07744](#).
- [442] F. Bolgar, E. Eames, C. Hottier, B. Semelin, Imprints of quasar duty cycle on the 21cm signal from the Epoch of Reionization, 2018, ArXiv e-prints, [arXiv:1802.05880](#).
- [443] M. Mapelli, A. Ferrara, E. Pierpaoli, *Mon. Not. R. Astron. Soc.* 369 (2006) 1719–1724, [arXiv:astro-ph/0603237](#).
- [444] R. Salvaterra, F. Haardt, A. Ferrara, *Mon. Not. R. Astron. Soc.* 362 (2005) L50–L54, [arXiv:astro-ph/0507208](#).
- [445] M. McQuinn, *Mon. Not. R. Astron. Soc.* 426 (2012) 1349–1360, [arXiv:1206.1335](#).
- [446] C. Evoli, A. Mesinger, A. Ferrara, *J. Cosmol. Astropart. Phys.* 11 (2014) 024, [arXiv:1408.1109](#).
- [447] S.R. Furlanetto, S.P. Oh, E. Pierpaoli, *Phys. Rev. D* 74 (10) (2006) 103502, [arXiv:astro-ph/0608385](#).
- [448] D.L. Meier, *Agron. J.* 207 (1976) 343–350.
- [449] J.E. Gunn, B.A. Peterson, *Agron. J.* 142 (1965) 1633–1641.
- [450] R.H. Becker, X. Fan, R.L. White, et al., *Astron. J.* 122 (2001) 2850–2857, [arXiv:astro-ph/0108097](#).
- [451] P.A. Oesch, R.J. Bouwens, G.D. Illingworth, et al., *Agron. J.* 773 (2013) 75.
- [452] P.A. Oesch, G. Brammer, P.G. van Dokkum, et al., *Agron. J.* 819 (2016) 129.
- [453] R.J. Bouwens, G.D. Illingworth, M. Franx, H. Ford, *Agron. J.* 670 (2007) 928–958.
- [454] R.J. Bouwens, G.D. Illingworth, V. González, et al., *Agron. J.* 725 (2010) 1587–1599.
- [455] R.J. Bouwens, G.D. Illingworth, I. Labbe, et al., *Nature* 469 (2011) 504–507.
- [456] R.J. Bouwens, G.D. Illingworth, P.A. Oesch, et al., *Agron. J.* 737 (2011) 90.
- [457] R.A.A. Bowler, J.S. Dunlop, R.J. McLure, et al., *Mon. Not. R. Astron. Soc.* 440 (2014) 2810–2842.
- [458] R.A.A. Bowler, J.S. Dunlop, R.J. McLure, et al., *Mon. Not. R. Astron. Soc.* 452 (2015) 1817–1840.
- [459] R.J. McLure, J.S. Dunlop, M.e.a. Cirasuolo, *Mon. Not. R. Astron. Soc.* 403 (2010) 960–983.
- [460] R.J. McLure, J.S. Dunlop, L. de Ravel, et al., *Mon. Not. R. Astron. Soc.* 418 (2011) 2074–2105.
- [461] R.J. McLure, J.S. Dunlop, R.A.A. Bowler, et al., *Mon. Not. R. Astron. Soc.* 432 (2013) 2696–2716.
- [462] L.D. Bradley, M. Trenti, P.A. Oesch, et al., *Agron. J.* 760 (2012) 108.
- [463] L.D. Bradley, A. Zitrin, D. Coe, et al., *Agron. J.* 792 (2014) 76.
- [464] M. Castellano, A. Fontana, D. Paris, et al., *Astron. Astrophys.* 524 (2010) A28.
- [465] M. Castellano, A. Fontana, K. Boutsia, et al., *Astron. Astrophys.* 511 (2010) A20.
- [466] D.E. Osterbrock, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, 1989.
- [467] A. Dey, H. Spinrad, D. Stern, J.R. Graham, F.H. Chaffee, *Astrophys. J. Lett.* 498 (1998) L93–L97.
- [468] H. Spinrad, D. Stern, A. Bunker, A. Dey, K. Lanzetta, A. Yahil, S. Pascalelle, A. Fernández-Soto, *Agron. J.* 116 (1998) 2617–2623.
- [469] E.M. Hu, L.L. Cowie, R.G. McMahon, *Astrophys. J. Lett.* 502 (1998) L99–L103, [arXiv:astro-ph/9803011](#).
- [470] M. Ouchi, K. Shimasaku, M. Akiyama, et al., *Astrophys. J. Lett.* 620 (2005) L1–L4.
- [471] Y. Taniguchi, M. Ajiki, T. Nagao, et al., *Publ. Astron. Soc. Japan* 57 (2005) 165–182.
- [472] M. Ouchi, K. Shimasaku, M. Akiyama, et al., *Astrophys. J. Suppl. Ser.* 176 (2008) 301–330.
- [473] M. Ouchi, K. Shimasaku, H. Furusawa, et al., *Astrophys. J.* 723 (2010) 869–894.
- [474] M. Iye, K. Ota, N. Kashikawa, H. Furusawa, T. Hashimoto, T. Hattori, Y. Matsuda, T. Morokuma, M. Ouchi, K. Shimasaku, *Nature* 443 (2006) 186–188.
- [475] N. Kashikawa, K. Shimasaku, M.A. Malkan, et al., *Agron. J.* 648 (2006) 7–22.
- [476] K. Ota, M. Iye, N. Kashikawa, et al., *Agron. J.* 677 (2008) 12–26.
- [477] K. Ota, M. Iye, N. Kashikawa, et al., *Astrophys. J.* 722 (2010) 803–811.
- [478] Y. Taniguchi, T. Murayama, N.Z. Scoville, et al., *Agron. J.* 701 (2009) 915–944.
- [479] N. Kashikawa, K. Shimasaku, Y. Matsuda, et al., *Agron. J.* 734 (2011) 119.
- [480] J.J.A. Matthee, D. Sobral, A.M. Swinbank, et al., *Mon. Not. R. Astron. Soc.* 440 (2014) 2375–2387.
- [481] J. Matthee, D. Sobral, S. Santos, H. Röttgering, B. Darvish, B. Mobasher, *Mon. Not. R. Astron. Soc.* 451 (2015) 400–417.
- [482] M. Ouchi, Y. Harikane, T. Shibuya, et al., *Pub. AJ* 70 (2018) S13, [arXiv:1704.07455](#).
- [483] P. Madau, M.J. Rees, *Astrophys. J. Lett.* 542 (2000) L69–L73, [arXiv:astro-ph/0006271](#).
- [484] E.M. Hu, L.L. Cowie, A.J. Barger, P. Capak, Y. Kakazu, L. Trouille, *Astrophys. J.* 725 (2010) 394–423, [arXiv:1009.1144](#).
- [485] S. Santos, D. Sobral, J. Matthee, *Mon. Not. R. Astron. Soc.* 463 (2016) 1678–1691, [arXiv:1606.07435](#).
- [486] K. Ota, M. Iye, N. Kashikawa, et al., *Astrophys. J.* 844 (2017) 85, [arXiv:1703.02501](#).
- [487] D.P. Stark, J. Richard, S. Charlot, et al., *Mon. Not. R. Astron. Soc.* 450 (2015) 1846–1855.
- [488] R. Smit, R.J. Bouwens, M. Franx, et al., *Astrophys. J.* 801 (2015) 122.
- [489] G.W. Roberts-Borsani, R.J. Bouwens, P.A. Oesch, et al., *Astrophys. J.* 823 (2016) 143.
- [490] D.P. Stark, G. Walth, S. Charlot, et al., *Mon. Not. R. Astron. Soc.* 454 (2015) 1393–1403.
- [491] K.B. Schmidt, K.-H. Huang, T. Treu, et al., *Astrophys. J.* 839 (2017) 17, [arXiv:1702.04731](#).
- [492] A. Verhamme, D. Schaefer, H. Atek, C. Tapken, *Astron. Astrophys.* 491 (2008) 89–111, [arXiv:0805.3601](#).
- [493] P. Dayal, A. Ferrara, *Mon. Not. R. Astron. Soc.* 421 (2012) 2568–2579, [arXiv:1109.0297](#).
- [494] A. Hutter, P. Dayal, V. M'uller, *Mon. Not. R. Astron. Soc.* 450 (2015) 4025–4034, [arXiv:1503.05201](#).
- [495] D.P. Stark, R.S. Ellis, M. Ouchi, *Astrophys. J. Lett.* 728 (2011) L2.
- [496] L. Pentericci, E. Vanzella, A. Fontana, et al., *Astrophys. J.* 793 (2014) 113, [arXiv:1403.5466](#).
- [497] A. Hutter, P. Dayal, A.M. Partl, V. M'uller, *Mon. Not. R. Astron. Soc.* 441 (2014) 2861–2877, [arXiv:1401.2830](#).
- [498] J.S. Dunlop, in: T. Wiklind, B. Mobasher, V. Bromm (Eds.), *The First Galaxies*, vol. 396, 2013, p. 223.
- [499] D.P. Stark, *Annu. Rev. Astron. Astrophys.* 54 (2016) 761–803.
- [500] T. Wiklind, M. Dickinson, H.C. Ferguson, M. Giavalisco, B. Mobasher, N.A. Grogin, N. Panagia, *Agron. J.* 676 (2008) 781–806.
- [501] C.J. Willott, P. Delorme, C. Reylé, et al., *Agron. J.* 139 (2010) 906–918.
- [502] D.J. Mortlock, S.J. Warren, B.P. Venemans, et al., *Astrophys. J.* 725 (2010) 1011–1031.
- [503] C.J. Willott, J. Bergeron, A. Omont, *Agron. J.* 801 (2015) 123.
- [504] N. Kashikawa, Y. Ishizaki, C.J. Willott, et al., *Agron. J.* 798 (2015) 28.
- [505] L. Jiang, I.D. McGreer, X. Fan, et al., *Astrophys. J.* 833 (2016) 222, [arXiv:1610.05369](#).
- [506] Y. Ono, M. Ouchi, Y.e.a. Harikane, *Pub. AJ* (2017) arXiv:1704.06004.
- [507] S.E. Woosley, J.S. Bloom, *Ann. Rev. Astron. Astrophys.* 44 (2006) 507–556.
- [508] R. Salvaterra, M. Della Valle, S. Campana, et al., *Nature* 461 (2009) 1258–1260.
- [509] N.R. Tanvir, D.B. Fox, A.J. Levan, et al., *Nature* 461 (2009) 1254–1257.
- [510] P. Schechter, *Astrophys. J.* 203 (1976) 297–306.

- [511] P.A. Oesch, R.J. Bouwens, C.M. Carollo, et al., *Astrophys. J. Lett.* 709 (2010) L21–L25.
- [512] R.J. Bouwens, P.A. Oesch, G.D. Illingworth, R.S. Ellis, M. Stefanon, *Astrophys. J.* 843 (2017) 129, [arXiv:1610.00283](#).
- [513] D.J. McLeod, R.J. McLure, J.S. Dunlop, *Mon. Not. R. Astron. Soc.* 459 (2016) 3812–3824, [arXiv:1602.05199](#).
- [514] R.J. Bouwens, P.A. Oesch, I. Labb  , et al., *Astrophys. J.* 830 (2016) 67, [arXiv:1506.01035](#).
- [515] P. Dayal, J.S. Dunlop, U. Maio, B. Ciardi, *Mon. Not. R. Astron. Soc.* 434 (2013) 1486–1504, [arXiv:1211.1034](#).
- [516] A. Grazian, M. Castellano, A. Fontana, et al., *Astron. Astrophys.* 547 (2012) A51.
- [517] R.J. Bouwens, G.D. Illingworth, P.A. Oesch, H. Atek, D. Lam, M. Stefanon, *Astrophys. J.* 843 (2017) 41, [arXiv:1608.00966](#).
- [518] P.A. Oesch, R.J. Bouwens, G.D. Illingworth, I. Labb  , M. Stefanon, *Astrophys. J.* 855 (2018) 105, [arXiv:1710.11131](#).
- [519] N.Y. Gnedin, *Astrophys. J. Lett.* 825 (2016) L17, [arXiv:1603.07729](#).
- [520] M. Dijkstra, J.S.B. Wyithe, Z. Haiman, *Mon. Not. R. Astron. Soc.* 379 (2007) 253–259, [arXiv:astro-ph/0611195](#).
- [521] P. Dayal, A. Ferrara, S. Gallerani, *Mon. Not. R. Astron. Soc.* 389 (2008) 1683–1696, [arXiv:0807.2975](#).
- [522] P. Dayal, A. Ferrara, A. Saro, *Mon. Not. R. Astron. Soc.* 402 (2010) 1449–1457, [arXiv:0907.4989](#).
- [523] J.S. Bolton, M.G. Haehnelt, *Mon. Not. R. Astron. Soc.* 429 (2013) 1695–1704, [arXiv:1208.4417](#).
- [524] L.H. Weinberger, G. Kulkarni, M.G. Haehnelt, T.R. Choudhury, E. Puchwein, *Mon. Not. R. Astron. Soc.* 479 (2018) 2564–2587, [arXiv:1803.03789](#).
- [525] M. Dijkstra, J.S.B. Wyithe, *Mon. Not. R. Astron. Soc.* 419 (2012) 3181–3193, [arXiv:1108.3840](#).
- [526] C.C. Steidel, M. Bogosavljevi  , A.E. Shapley, J.A. Kollmeier, N.A. Reddy, D.K. Erb, M. Pettini, *Astrophys. J.* 736 (2011) 160, [arXiv:1101.2204](#).
- [527] F. Leclercq, R. Bacon, L. Wisotzki, et al., *Astron. Astrophys.* 608 (2017) A8, [arXiv:1710.10271](#).
- [528] S. De Barros, L. Pentericci, E. Vanzella, et al., *Astron. Astrophys.* 608 (2017) A123, [arXiv:1710.01784](#).
- [529] M.A. Schenker, R.S. Ellis, N.P. Konidaris, D.P. Stark, *Astrophys. J.* 795 (2014) 20, [arXiv:1404.4632](#).
- [530] L. Pentericci, A. Fontana, E. Vanzella, et al., *Astrophys. J.* 743 (2011) 132, [arXiv:1107.1376](#).
- [531] Y. Ono, M. Ouchi, B. Mobasher, et al., *Astrophys. J.* 744 (2012) 83.
- [532] V. Tilvi, C. Papovich, S.L. Finkelstein, et al., *Astrophys. J.* 794 (2014) 5, [arXiv:1405.4869](#).
- [533] P. Cassata, L.A.M. Tasca, O. Le F  vre, et al., *Astron. Astrophys.* 573 (2015) A24, [arXiv:1403.3693](#).
- [534] E. Curti  -Lake, R.J. McLure, H.J. Pearce, et al., *Mon. Not. R. Astron. Soc.* 422 (2012) 1425–1435, [arXiv:1110.1722](#).
- [535] I. Labb  , P.A. Oesch, R.J. Bouwens, et al., *Astrophys. J. Lett.* 777 (2013) L19.
- [536] V. Gonz  lez, I. Labb  , R.J. Bouwens, G. Illingworth, M. Franx, M. Kriek, *Astrophys. J. Lett.* 735 (2011) L34, [arXiv:1008.3901](#).
- [537] D.P. Stark, M.A. Schenker, R. Ellis, B. Robertson, R. McLure, J. Dunlop, *Astrophys. J.* 763 (2013) 129, [arXiv:1208.3529](#).
- [538] P.A. Oesch, R.J. Bouwens, G.D. Illingworth, et al., *Astrophys. J.* 786 (2014) 108, [arXiv:1309.2280](#).
- [539] K. Duncan, C.J. Conselice, A. Mortlock, et al., *Mon. Not. R. Astron. Soc.* 444 (2014) 2960–2984, [arXiv:1408.2527](#).
- [540] A. Grazian, A. Fontana, P. Santini, et al., *Astron. Astrophys.* 575 (2015) A96, [arXiv:1412.0532](#).
- [541] M. Song, S.L. Finkelstein, M.L.N. Ashby, et al., *Astrophys. J.* 825 (2016) 5, [arXiv:1507.05636](#).
- [542] R.S. Ellis, R.J. McLure, J.S. Dunlop, et al., *Astrophys. J. Lett.* 763 (2013) L7, [arXiv:1211.6804](#).
- [543] W. Zheng, M. Postman, A. Zitrin, et al., *Nature* 489 (2012) 406–408, [arXiv:1204.2305](#).
- [544] D. Coe, A. Zitrin, M. Carrasco, et al., *Astrophys. J.* 762 (2013) 32, [arXiv:1211.3663](#).
- [545] R.J. Bouwens, L. Bradley, A. Zitrin, et al., *Astrophys. J.* 795 (2014) 126, [arXiv:1211.2230](#).
- [546] P. Madau, L. Pozzetti, M. Dickinson, *Astrophys. J.* 498 (1998) 106–116, [arXiv:astro-ph/9708220](#).
- [547] R.J. Bouwens, G.D. Illingworth, P.A. Oesch, et al., *Astrophys. J.* 754 (2012) 83, [arXiv:1109.0994](#).
- [548] S. Cazaux, A.G.G.M. Tielens, *Astrophys. J.* 604 (2004) 222–237.
- [549] H. Hirashita, A. Ferrara, *Mon. Not. R. Astron. Soc.* 337 (2002) 921–937, [arXiv:astro-ph/0209034](#).
- [550] D. Yamasawa, A. Habe, T. Kozasa, T. Nozawa, H. Hirashita, H. Umeda, K. Nomoto, *Astrophys. J.* 735 (2011) 44, [arXiv:1104.0728](#).
- [551] K. Omukai, T. Tsuribe, R. Schneider, A. Ferrara, *Astrophys. J.* 626 (2005) 627–643, [arXiv:astro-ph/0503010](#).
- [552] P. Dayal, H. Hirashita, A. Ferrara, *Mon. Not. R. Astron. Soc.* 403 (2010) 620–624, [arXiv:0908.1571](#).
- [553] F. Matteucci, L. Greggio, *Astron. Astrophys.* 154 (1986) 279–287.
- [554] J. Feder, K. Russell, J. Lotte, G. Pound, *Adv. Phys.* 15 (1966) 111–178.
- [555] P. Todini, A. Ferrara, *Mon. Not. R. Astron. Soc.* 325 (2001) 726–736, [arXiv:astro-ph/0009176](#).
- [556] D.H. Wooden, D.M. Rank, J.D. Bregman, F.C. Witteborn, A.G.G.M. Tielens, M. Cohen, P.O. Pinto, T.S. Axelrod, *Astrophys. J. Suppl.* 88 (1993) 477–507.
- [557] A. Elmhamdi, I.J. Danziger, N. Chugai, et al., *Mon. Not. R. Astron. Soc.* 338 (2003) 939–956, [arXiv:astro-ph/0209623](#).
- [558] T. Kozasa, T. Nozawa, N. Tominaga, H. Umeda, K. Nomoto, in: T. Henning, E. Gr  un, J. Steinacker (Eds.), *Cosmic Dust – Near and Far*, in: *Astronomical Society of the Pacific Conference Series*, vol. 414, 2009, p. 43, [arXiv:0903.0217](#).
- [559] C. Gall, J. Hjorth, A.C. Andersen, *Astron. Astrophys. Rev.* 19 (2011) 43, [arXiv:1108.0403](#).
- [560] J. Rho, T. Kozasa, W.T. Reach, J.D. Smith, L. Rudnick, T. DeLaney, J.A. Ennis, H. Gomez, A. Tappe, *Astrophys. J.* 673 (2008) 271–282, [arXiv:0709.2880](#).
- [561] J. Rho, W.T. Reach, A. Tappe, U. Hwang, J.D. Slavin, T. Kozasa, L. Dunne, *Astrophys. J.* 700 (2009) 579–596.
- [562] M. Matsuuwa, E. Dwek, M. Meixner, et al., *Science* 333 (2011) 1258, [arXiv:1107.1477](#).
- [563] R. Indebetouw, M. Matsuuwa, E. Dwek, et al., *Astrophys. J. Lett.* 782 (2014) L2, [arXiv:1312.4086](#).
- [564] S. Bianchi, R. Schneider, *Mon. Not. R. Astron. Soc.* 378 (2007) 973–982, [arXiv:0704.0586](#).
- [565] S. Zhukovska, H.-P. Gail, M. Trieloff, *Astron. Astrophys.* 479 (2008) 453–480, [arXiv:0706.1155](#).
- [566] R. Valiante, R. Schneider, S. Bianchi, A.C. Andersen, *Mon. Not. R. Astron. Soc.* 397 (2009) 1661–1671, [arXiv:0905.1691](#).
- [567] E. Dwek, F. Galliano, A.P. Jones, *Astrophys. J.* 662 (2007) 927–939, [arXiv:0705.3799](#).
- [568] H. Hirashita, A. Ferrara, P. Dayal, M. Ouchi, *Mon. Not. R. Astron. Soc.* 443 (2014) 1704–1712, [arXiv:1406.6762](#).
- [569] H. Hirashita, D. Burgarella, R.J. Bouwens, *Mon. Not. R. Astron. Soc.* 472 (2017) 4587–4597, [arXiv:1709.02526](#).
- [570] T. Nozawa, T. Kozasa, A. Habe, E. Dwek, H. Umeda, N. Tominaga, K. Maeda, K. Nomoto, *Astrophys. J.* 666 (2007) 955–966, [arXiv:0706.0383](#).
- [571] M.J. Barlow, O. Krause, B.M. Swinyard, et al., *Astron. Astrophys.* 518 (2010) L138, [arXiv:1005.2688](#).
- [572] H.L. Gomez, O. Krause, M.J. Barlow, et al., *Astrophys. J.* 760 (2012) 96, [arXiv:1209.5677](#).
- [573] M. Otsuka, J.T. van Loon, K.S. Long, et al., *Astron. Astrophys.* 518 (2010) L139, [arXiv:1005.2787](#).
- [574] B.T. Draine, in: T. Henning, E. Gr  un, J. Steinacker (Eds.), *Cosmic Dust – Near and Far*, in: *Astronomical Society of the Pacific Conference Series*, vol. 414, 2009, p. 453, [arXiv:0903.1658](#).
- [575] A.P. Jones, J.A. Nuth, *Astron. Astrophys.* 530 (2011) A44.
- [576] R.S. Asano, T.T. Takeuchi, H. Hirashita, A.K. Inoue, *Earth, Planets, and Space* 65 (2013) 213–222, [arXiv:1206.0817](#).
- [577] E. Dwek, *Astrophys. J.* 501 (1998) 643, [arXiv:astro-ph/9707024](#).
- [578] A.K. Inoue, *Earth, Planets, and Space* 63 (2011) 1027–1039.
- [579] M. de Bennassuti, R. Schneider, R. Valiante, S. Salvadori, *Mon. Not. R. Astron. Soc.* 445 (2014) 3039–3054, [arXiv:1409.5798](#).
- [580] M. Mancini, R. Schneider, L. Graziani, R. Valiante, P. Dayal, U. Maio, B. Ciardi, L.K. Hunt, *Mon. Not. R. Astron. Soc.* 451 (2015) L70–L74, [arXiv:1505.01841](#).
- [581] C.F. McKee, *Astrophys. J.* 345 (1989) 782–801.
- [582] U. Lisenfeld, A. Ferrara, *Astrophys. J.* 496 (1998) 145–154, [arXiv:astro-ph/9705037](#).

- [583] C.G. Seab, J.M. Shull, *Astrophys. J.* 275 (1983) 652–660.
- [584] A. Ferrara, S. Viti, C. Ceccarelli, *Mon. Not. R. Astron. Soc.* 463 (2016) L112–L116, arXiv:1606.07214.
- [585] G.R. Meurer, T.M. Heckman, D. Calzetti, *Astrophys. J.* 521 (1999) 64–80, arXiv:astro-ph/9903054.
- [586] S.L. Finkelstein, C. Papovich, B. Salmon, et al., *Astrophys. J.* 756 (2012) 164, arXiv:1110.3785.
- [587] J.S. Dunlop, A.B. Rogers, R.J. McLure, et al., *Mon. Not. R. Astron. Soc.* 432 (2013) 3520–3533, arXiv:1212.0860.
- [588] R.J. Bouwens, G.D. Illingworth, P.A. Oesch, et al., *Astrophys. J.* 793 (2014) 115, arXiv:1306.2950.
- [589] A.B. Rogers, R.J. McLure, J.S. Dunlop, et al., *Mon. Not. R. Astron. Soc.* 440 (2014) 3714–3725, arXiv:1312.4975.
- [590] S.M. Wilkins, A. Bunker, W. Coulton, R. Croft, T. di Matteo, N. Khandai, Y. Feng, *Mon. Not. R. Astron. Soc.* 430 (2013) 2885–2890, arXiv:1302.1387.
- [591] M. Mancini, R. Schneider, L. Graziani, R. Valiante, P. Dayal, U. Maiò, B. Ciardi, *Mon. Not. R. Astron. Soc.* 462 (2016) 3130–3145, arXiv:1607.04285.
- [592] S.L. Finkelstein, R.E. Ryan Jr., C. Papovich, et al., *Astrophys. J.* 810 (2015) 71, arXiv:1410.5439.
- [593] Z. Khakheleva-Li, N.Y. Gnedin, *Astrophys. J.* 820 (2016) 133, arXiv:1601.00641.
- [594] S.M. Wilkins, Y. Feng, T. Di Matteo, R. Croft, C.C. Lovell, P. Thomas, *Mon. Not. R. Astron. Soc.* 473 (2018) 5363–5369, arXiv:1710.01976.
- [595] P.L. Capak, C. Carilli, G. Jones, et al., *Nature* 522 (2015) 455–458, arXiv:1503.07596.
- [596] M.J. Michałowski, *Astron. Astrophys.* 577 (2015) A80, arXiv:1503.08210.
- [597] C. Behrens, A. Pallottini, A. Ferrara, S. Gallerani, L. Vallini, *Mon. Not. R. Astron. Soc.* 477 (2018) 552–565, arXiv:1802.07772.
- [598] D. Schaefer, F. Boone, M. Zamojski, J. Staguhn, M. Dessauges-Zavadsky, S. Finkelstein, F. Combes, *Astron. Astrophys.* 574 (2015) A19, arXiv:1407.5793.
- [599] N. Kanekar, J. Wagg, R.R. Chary, C.L. Carilli, *Astrophys. J. Lett.* 771 (2013) L20, arXiv:1305.6469.
- [600] K. Ota, F. Walter, K. Ohta, et al., *Astrophys. J.* 792 (2014) 34, arXiv:1405.5387.
- [601] M. Ouchi, R. Ellis, Y. Ono, et al., *Astrophys. J.* 778 (2013) 102, arXiv:1306.3572.
- [602] R. Maiolino, S. Carniani, A. Fontana, et al., *Mon. Not. R. Astron. Soc.* 452 (2015) 54–68, arXiv:1502.06634.
- [603] D. Watson, L. Christensen, K.K. Knudsen, J. Richard, A. Gallazzi, M.J. Michałowski, *Nature* 519 (2015) 327–330.
- [604] N. Laporte, R.S. Ellis, F. Boone, F.E. Bauer, D. Quénard, G.W. Roberts-Borsani, R. Pelló, I. Pérez-Fournon, A. Streblyanska, *Astrophys. J. Lett.* 837 (2017) L21, arXiv:1703.02039.
- [605] H.J. Mo, S. Mao, S.D.M. White, 295 (1998) 319–336.
- [606] P.J.E. Peebles, *Astrophys. J.* 155 (1969) 393.
- [607] J. Barnes, G. Efstathiou, *Astrophys. J.* 319 (1987) 575–600.
- [608] A.G. Doroshkevich, *Astrophysics* 6 (1970) 320–330.
- [609] J.S. Bullock, A. Dekel, T.S. Kolatt, A.V. Kravtsov, A.A. Klypin, C. Porciani, J.R. Primack, *Astrophys. J.* 555 (2001) 240–257.
- [610] M. Vitvitska, A.A. Klypin, A.V. Kravtsov, R.H. Wechsler, J.R. Primack, J.S. Bullock, *Astrophys. J.* 581 (2002) 799–809.
- [611] A.J. Davis, P. Natarajan, *Mon. Not. R. Astron. Soc.* 393 (2009) 1498–1502.
- [612] T. Shibuya, M. Ouchi, M. Kubo, Y. Harikane, *Astrophys. J.* 821 (2016) 72.
- [613] P. Madau, M. Dickinson, *Annu. Rev. Astron. Astrophys.* 52 (2014) 415–486, arXiv:1403.0007.
- [614] J.M. Lotz, P. Jonsson, T.J. Cox, D. Croton, J.R. Primack, R.S. Somerville, K. Stewart, *Astrophys. J.* 742 (2011) 103, arXiv:1108.2508.
- [615] H.C. Ferguson, M. Dickinson, M. Giavalisco, et al., *Astrophys. J. Lett.* 600 (2004) L107–L110.
- [616] S. Ravindranath, M. Giavalisco, H.C. Ferguson, et al., *Astrophys. J.* 652 (2006) 963–980, arXiv:astro-ph/0606696.
- [617] C.J. Conselice, J. Arnold, *Mon. Not. R. Astron. Soc.* 397 (2009) 208–231, arXiv:0904.4250.
- [618] L. Jiang, E. Egami, X. Fan, et al., *Astrophys. J.* 773 (2013) 153, arXiv:1303.0027.
- [619] R. Kawamata, M. Ishigaki, K. Shimasaku, M. Oguri, M. Ouchi, *Astrophys. J.* 804 (2015) 103.
- [620] N.A. Bond, E. Gawiser, C. Gronwall, R. Ciardullo, M. Altmann, K. Schawinski, *Astrophys. J.* 705 (2009) 639–649.
- [621] A. Dekel, R. Sari, D. Ceverino, *Astrophys. J.* 703 (2009) 785–801, arXiv:0901.2458.
- [622] E. Romano-Díaz, J.-H. Choi, I. Shlosman, M. Trenti, *Astrophys. J. Lett.* 738 (2011) L19, arXiv:1106.1214.
- [623] Y. Feng, T. Di Matteo, R. Croft, A. Tenneti, S. Bird, N. Battaglia, S. Wilkins, *Astrophys. J. Lett.* 808 (2015) L17, arXiv:1504.06618.
- [624] R.B. Tully, J.R. Fisher, *Astron. Astrophys.* 54 (1977) 661–673.
- [625] S.M. Faber, R.E. Jackson, *Astrophys. J.* 204 (1976) 668–683.
- [626] C. Liu, S.J. Mutch, G.B. Poole, P.W. Angel, A.R. Duffy, P.M. Geil, A. Mesinger, J.S.B. Wyithe, *Mon. Not. R. Astron. Soc.* 465 (2017) 3134–3142.
- [627] N.P. Hathi, R.A. Jansen, R.A. Windhorst, S.H. Cohen, W.C. Keel, M.R. Corbin, R.E. Ryan Jr., *Astron. J.* 135 (2008) 156–166.
- [628] M. Mosleh, R.J. Williams, M. Franx, V. Gonzalez, R.J. Bouwens, P. Oesch, I. Labbe, G.D. Illingworth, M. Trenti, *Astrophys. J. Lett.* 756 (2012) L12.
- [629] Y. Ono, M. Ouchi, E. Curtis-Lake, et al., *Astrophys. J.* 777 (2013) 155.
- [630] R.J. Bouwens, G.D. Illingworth, J.P. Blakeslee, T.J. Broadhurst, M. Franx, *Astrophys. J. Lett.* 611 (2004) L1–L4.
- [631] B.W. Holwerda, R. Bouwens, P. Oesch, R. Smit, G. Illingworth, I. Labbe, *Astrophys. J.* 808 (2015) 6.
- [632] R.J. Allen, G.G. Kacprzak, K. Glazebrook, et al., ArXiv e-prints (2016) arXiv:1612.05262.
- [633] E. Curtis-Lake, R.J. McLure, J.S. Dunlop, et al., *Mon. Not. R. Astron. Soc.* 457 (2016) 440–464.
- [634] K.-H. Huang, H.C. Ferguson, S. Ravindranath, J. Su, *Astrophys. J.* 765 (2013) 68.
- [635] T. Shibuya, M. Ouchi, Y. Harikane, *Astrophys. J. Suppl.* 219 (2015) 15, arXiv:1503.07481.
- [636] B.P. Venemans, H.J.A. Röttgering, G.K. Miley, et al., *Astron. Astrophys.* 431 (2005) 793–812.
- [637] R.A. Overzier, R.J. Bouwens, N.J.G. Cross, et al., *Astrophys. J.* 673 (2008) 143–162.
- [638] S. Malhotra, J.E. Rhoads, S.L. Finkelstein, N. Hathi, K. Nilsson, E. McLinden, N. Pirzkal, *Astrophys. J. Lett.* 750 (2012) L36.
- [639] M. Barden, H.-W. Rix, R.S. Somerville, et al., *Astrophys. J.* 635 (2005) 959–981.
- [640] A.W. Graham, C.C. Worley, *Mon. Not. R. Astron. Soc.* 388 (2008) 1708–1728.
- [641] R. Bouwens, M. Aravena, R. Decarli, et al., ArXiv e-prints (2016) arXiv:1606.05280.
- [642] C.C. Dow-Hygelund, B.P. Holden, R.J. Bouwens, et al., *Astrophys. J.* 660 (2007) 47–61.
- [643] T.O. Zick, D.R. Weisz, M. Boylan-Kolchin, *Mon. Not. R. Astron. Soc.* (2018).
- [644] D. Vrbanec, B. Ciardi, V. Jelić, et al., *Mon. Not. R. Astron. Soc.* 457 (2016) 666–675, arXiv:1509.03464.
- [645] E. Sobacchi, A. Mesinger, B. Greig, *Mon. Not. R. Astron. Soc.* 459 (2016) 2741–2750, arXiv:1602.04837.
- [646] A. Hutter, P. Dayal, V. Müller, C.M. Trott, *Astrophys. J.* 836 (2017) 176, arXiv:1605.01734.
- [647] A. Hutter, C.M. Trott, P. Dayal, *Mon. Not. R. Astron. Soc.* 479 (2018) L129–L133, arXiv:1806.07902.
- [648] K. Kubota, S. Yoshiura, K. Takahashi, K. Hasegawa, H. Yajima, M. Ouchi, B. Pindor, R.L. Webster, *Mon. Not. R. Astron. Soc.* 479 (2018) 2754–2766, arXiv:1708.06291.