The Formation and Evolution of Galaxies

Setting the scene

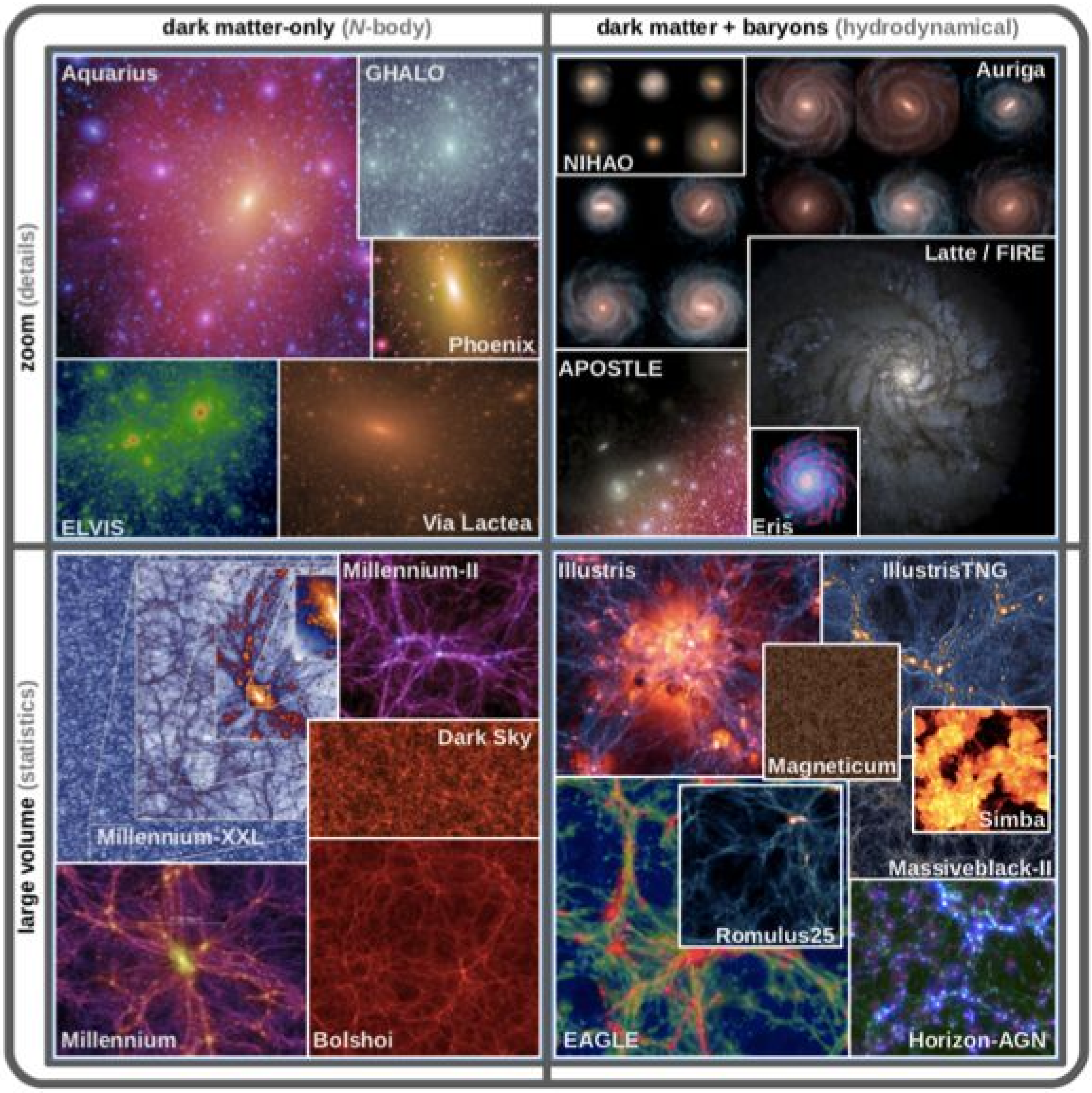
Prior to the recombination epoch, matter and radiation were tightly coupled in the post big-bang plasma. Given the large ratio of photons to baryons, matter felt significant pressure from Thompson scattering, and would have been distributed evenly throughout the cosmos. Only non-interacting particles of dark matter were unsusceptible to this scattering and could form slight overdensities, which amassed to one part in 100,000 by redshift (z) of about 1100 (around 400,000 years after the big bang). The background radiation emitted at recombination remains the only observable signature of these early epochs, after which the universe was plunged into the *dark ages*, an epoch about which we have almost no empirical information.

Yet the post recombination era was the scene in which the skeleton for the large-scale structure began construction. With dark matter able to collapse and produce the first scaffolding of the web-like structure of the cosmos, baryons now freed from Thompson scattering were able to gravitationally follow the structure of the dark matter. With no metals, and all the hydrogen and helium residing in their atomic forms, this era will only be probable via very low frequency radio observations. Yet during this time, the largest density fluctuations grew by around 10 orders of magnitude in 100 Million years, to the time when the first stars are believed to have ignited at z~30.

The formation, evolution, and end-points of the first stars were likely peculiar in many ways: the absence of interstellar metals and under-production of molecular gas will have resulted in differing gas fragmentation, stellar initial mass distributions, and binarity fractions from those expected and observed in the modern universe. Stellar rotation, and the lack of metals will have altered stellar opacities, convective zones, winds, and so on, which will all affect the rates of stellar evolution and emissivities of radiation. The photospheres of these first stars produced the first ionizing photons seen after recombination, producing a radiation field that could inhibit and delay subsequent local star-formation, and slowing the growth of primeval galaxies.

These first *Population III* stars also produce the first metals, which are delivered back to pollute pristine interstellar media with metals and likely dust. These metals have dramatic effects on all subsequent gas cooling, fragmentation, etc. which will provide advantages to the later generations of Population II stars. The deaths of these first stellar objects are also believed to result in the population of stellar mass black holes, and perhaps the seeds for luminous active galaxies seen at later epochs.

Galaxy formation continues as DM halos continue to grow and merge and potential wells deepen; the accretion of baryonic matter from the cosmic web accelerates driving star formation and early galaxy buildup until the first objects become visible at *z*~10 to deep Hubble Space Telescope imaging. Potentially from the remnants left by the first stellar explosions, black holes grow to attain ‘supermassive’ status, producing the first quasars that are currently observed only shortly afterwards. The mechanisms by which these easly systems grew remain elusive, but soon-to-be commissioned telescopes such as the James Webb Space Telescope and Extremely Large Telescopes on the ground will be charged with determining the properties of these galaxies, and with that the foundations for subsequent galaxy assembly. These galaxies and accreting supermassive black holes around redshifts 6-10 lit up the universe, and produced sufficient ionizing radiation to (re-)ionize the entirety of the remaining gas in between these first generation of galaxies, i.e. the intergalactic medium (IGM).



Vogelsberger et al. (2020): Visual overview of some recent structure and galaxy formation simulations. In left and right panels, DM-only and DM+baryons simulations are shown, respectively. In top panels, zoom simulations are displayed that are able to resolve small scales. In bottom panels, the large volume simulations are visualized.

The first overdensities become visible about a billion years after the big bang, with faint, lower mass galaxies surrounding massive galaxies and luminous AGN. The beginnings of the observable large-scale structure evolve with the buildup of massive galaxies, which proceeds down to redshifts of about 2. Somewhere around this time the Milky Way progenitor experienced its last significant merging event. The star formation rate of the universe increases ten-fold, and individual galaxies become more luminous, massive, and physically larger. Observations of the dust continuum and molecular gas show the assembling galaxies in this epoch to be very rich in gas, likely as a result of accretion from the IGM feeding gas-rich circumgalactic media (CGM), from where the gas settles to form massive disks. On the other hand, winds and outflowing material appear to be ubiquitous in these galaxies, and the complex interplay between gas inflows and outflows, occurring in the circumgalactic environment, regulate the efficiency and ultimately the modes of star formation. Scaling relations between fundamental galaxy properties become clear, with relatively tight sequences emerging between mass, star-formation rates, and abundances of heavy elements and sizes. These relations remain in place, but shift in their loci, down to the present day universe; the processes underpinning our place in the cosmos were established at least 10 billion years ago, if not earlier.

This side of redshift 2, galaxy evolution changes fundamentally. While galaxies continue to grow, the latter 10 Gyr of evolution have seen the total star formation rate of the universe decrease 10-fold. This is likely because of a shut-off in the delivery mechanism, as accretion of gas from the cosmic web slows, combined with a fundamental decrease in the star-formation efficiency. Over the same timeframe, the nodal points of the more diffuse large-scale structure/cosmic web condense to form galaxy clusters with masses 10-100 Trillian (US; 10^{12}) times that of the sun, and high-temperature gaseous intracluster media, and play host to the most massive galaxies known. As mass overdensities grow and cosmic contrast increases, environmental processes contribute more to galaxy formation. The complex interplay of environmental interactions and internal ('secular') processes induce morphological transitions, growing galaxy bulges/spheroids and quenching in a process that begins at the highest galaxy masses. The scaling relations that were tight at $z>2$ bend, and the total cosmic star-formation budget shifts out of dense, dust obscured massive systems into more isolated/group environment disks that in isolation evolve over many billions of years.

The First Galaxies and the Epoch of Reionization

Soon after its inception in the hot big bang, the Universe most likely underwent a period of exponential expansion called inflation, that was responsible for generating the (nearly) scale-invariant primordial density perturbations that form the seeds of all the large-scale structure we see today. After inflation, as the Universe expanded and cooled, roughly 400,000 years after the Big bang, it became cool enough that electrons and protons could combine to form neutral hydrogen atoms - this is known as the "*epoch of recombination*". It is at this point that matter and radiation decoupled from each other, giving rise to the Cosmic Microwave Background (CMB) discussed in the previous section, *Origin and Evolution of the Universe*. Indeed, the CMB has been crucial in inferring the two key cosmic constituents that shape the formation and evolution of galaxies: dark matter (representing about 86% of all matter) that forms the cosmic scaffolding within which galaxies form and baryons (normal matter; comprising 14% of all matter) that interacts with radiation to make up the "visible" Universe. Immediately after the recombination epoch, the Universe entered a phase called the "*Dark Ages*", where no significant radiation sources existed, nor observations can currently probe. At this time, the small inhomogeneities in the dark matter density field present during the recombination epoch grew via gravitational instability giving rise to collapsed haloes. In what is now understood to be the second stage of galaxy formation, diffuse (neutral) hydrogen and helium gas was accreted by these halos from the Intergalactic Medium (IGM), and started cooling via gas-dynamical processes. If the mass of the halo is high enough, the gas would be able to dissipate its energy, cool via atomic or molecular hydrogen transitions and fragment within the halo, giving rise to conditions appropriate for the condensation of gas to form the first stars. The cosmic dark ages end and the era of "*cosmic dawn*" begins with the formation of the first stars a few hundred million years after the Big bang. Forming out of metal-free gas[1] of primordial composition (roughly 75% Hydrogen and 25% Helium, by mass), these stars are classified as Population III (PopIII) stars[2]; this classification has been refined to include PopIII.1 and PopIII.2 stars that formed out of neutral and photoionized primordial gas collapsing into the earliest halos, respectively.

The formation of the first stars irreversibly changed the Universe and all subsequent galaxy formation in two key ways: Firstly, PopIII stars in the mass range of 140-260 solar masses probably explode as highly energetic Pair Instability Supernovae (PISN), yielding copious amounts of metals and starting the "*metal age*" of the Universe. Secondly, the first population of luminous stars and galaxies generated hydrogen (and helium) ionizing photons, starting the "*epoch of (hydrogen and helium) reionization*". A fraction of the ionizing energy produced by these sources was also responsible for heating the IGM, which resulted in gas being (photo) evaporated from low-mass halos, and the inhibition of their star formation. The Epoch of reionization (EoR) is therefore of immense importance in the study of structure formation because, on the one hand, it is a direct consequence of the formation of first structures and luminous sources while, on the other, it affects all subsequent structure formation. However, due to a number of interlinked complex problems - including the gas masses and star formation rates of early galaxies, the fraction of ionizing photons that could escape out of the galactic environment and contribute to reionization (escape fraction) and the impact of reionization on galaxy formation - the history, sources and progress of reionization remain compelling open questions.

The last few years have seen a golden era for the search of galaxies lying within the EoR. This has been made possible by a combination of state-of-the-art facilities such as the Hubble Space Telescope, Subaru, Keck, and Very Large telescopes, refined selection techniques (such as the Lyman Break technique) and the power of cosmic gravitational lensing (that magnifies the light

---

[1] In Astronomy, all elements heavier than Helium are classified as "metals"

[2] Stars with metal content comparable to that of our Sun are termed Population I (PopI) stars

from distant objects). These have been used to detect sources at redshifts as high as z~11 (only half a billion years after the Big Bang) and galaxies 3 orders of magnitude less luminous than the Milky Way at z~7 (800 million years post-Big Bang). In addition, quasars, powered by accretion onto supermassive black holes are found to be in place well within the first billion years. Dust and metals have been identified in some of these systems with ALMA. Finally, some of the most distant spectroscopically confirmed cosmic objects are Gamma Ray Bursts that confirm star formation was already well under way at those early epochs, and further encouraging deeper galaxy searches. These data sets will soon be supplemented by those from cutting-edge facilities including ALMA, VLT and in the coming decade by galaxy observatories (JWST, Euclid, LSST, ELT), 21cm facilities (SKA and HERA) and gravitational wave observatories (Athena, LISA, Einstein telescope) to build a panchromatic picture of galaxy formation at these early epochs. This observational progress has naturally given rise to a plethora of theoretical models, ranging from analytic calculations to semi-analytic models to numerical simulations. We now discuss the key open questions in the field of early galaxy formation.

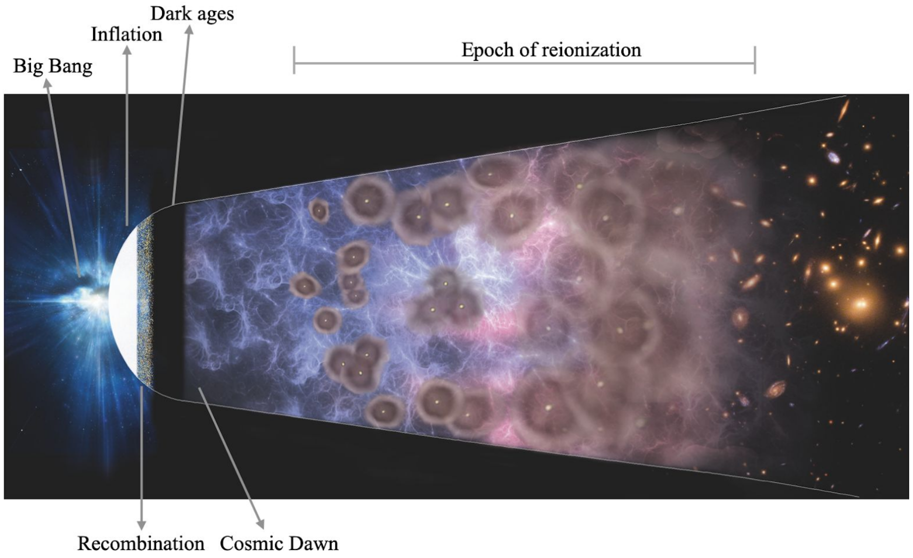


Fig. 1: A timeline of the key cosmic epochs.

Key open questions

1. How did the first galaxies obtain their gas?

According to the current standard theoretical paradigm, early galaxies accrete most of their gas from the IGM through “cold mode accretion” where the accreted gas temperature is much lower than the temperature within the (hot) halo. Providing ready-to-use-gas for star formation, filaments of cold gas that extend deep into the halo are crucial to explain the gas mass assembly and high star formation rates of early galaxies. However, on the theoretical side, the role of supernova and reionization feedback, that can decrease the cold gas mass accreted, remains poorly understood. Gaining a deeper theoretical understanding of how much gas can be gained via cold flows requires enhancing high-resolution simulations (i.e. FIRE, Illustris, Hestia) implementing realistic supernova and turbulent feedback prescriptions in a cosmological setting (so as to capture heating from the reionization process). On the observational front, the small solid angle covered by cold flows has rendered direct observations enormously difficult. In the future, we expect further observational strengthening of this paradigm through indirect evidence of cold flows via Lyman alpha emission (through e.g. sensitive integral field spectrographs) and through absorption-line systems (with e.g. JWST, ELT).

2. What are the properties and observational prospects of detecting the first stars?

The standard 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 a PopIII star, as well as their initial mass function (IMF) remain challenging open questions. This is because they depend on a number of poorly understood physical processes such as the initial turbulence of the halo gas, the presence and influence of magnetic fields, relative streaming velocities between baryons and dark matter and possibly even the impact of dark matter annihilation that still need to be self-consistently implemented in high-resolution simulations. On the observational front, a number of works have shown that the JWST and ELT could possibly detect the signatures of PopIII star formation in the high-z Universe: this could either be through spectroscopic and or photometric identification of the HeII (1640 A) emission line or through extremely strong HI Lyman alpha emission that produces peculiar broadband colours. These observations could also be optimized using gravitational lensing to boost the signal of faint galaxies/stars. In addition, the extremely luminous supernovae expected from PopIII stars could be detected at early times: the challenge here lies in building a statistical sample across multiple fields and separated by a long enough time-scale to study the slowly evolving light-curves. The idea would then be to study large fields with e.g. Euclid, WFIRST or LSST whose targets could be followed up by JWST. Finally, the technique of intensity mapping - 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. HeII-CO or HeII-21cm) in order to spatially correlated fluctuations and eliminate contaminants. Finally, closer to home, the field of "stellar archaeology" uses the fact that the metallicity of a star reflects the chemical conditions of its birth clouds to hunt for very old, surviving stars that retain the signatures of metal production in that early-epoch. Abundances derived from high resolution spectrographs on ELTs will provide dramatic impact in this field, targeting halo stars in the Milky Way and other nearby galaxies.

3. What were the key physical properties of early galaxies?

Both theoretical models and observations agree on a picture where the star formation rate density of the Universe progressively increases from the EoR down to the era of cosmic noon at z~3 and a tight correlation exists between the star formation rate and the stellar mass. However, the gas content of early galaxies, and therefore their star formation efficiency, remain open questions. From the theoretical side, zoom-in simulations of galaxies embedded in the cosmic web are being performed to map out the gas masses and temperatures. Observations of the gas mass and its distribution will be crucial in validating such models. Such observations will use molecular and atomic emission lines as well as sub-millimeter continuum emission to infer the gas mass and distribution. To capture the diversity in galaxy populations current instruments need to receive an upgrade (ALMA as foreseen in the ALMA2030 roadmap) or new instruments altogether (e.g. high-frequency at SKA, ngVLA) to be able to probe lower mass galaxies and, more importantly, capture the distribution of gas within galaxies which is critical for understanding the mechanisms controlling the star formation efficiency. In order to find and obtain a proper census of the gas- or dust-rich objects in the early universe, it is vital to push the observations to the lowest possible masses; this is critical for our understanding of the build-up of the most massive galaxies seen today (next section), and the requisite large-area surveys are essential can only be done with new instruments such as AtLAST.

The metal content of early galaxies remains another challenging open question that has been the focus of many theoretical models. How and when metallicity relations - linking key galaxy observables such as the stellar mass, star formation rate, gas-phase metallicity, and cold gas mass - emerged remain open questions although combinations of ALMA, JWST, and ELT data to target molecular, atomic, and ionized ionized phases, will extend fundamental-metallicity relations from intermediate redshifts to these very early epochs.

Finally, explaining the high observed dust-to-stellar mass ratios (~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. Some works now show dust grain growth to be extremely inefficient in the ISM given its high ionization fields, but the key dust sources, masses and their 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.

4. What were the key sources of reionization and how did reionization impact galaxy formation?

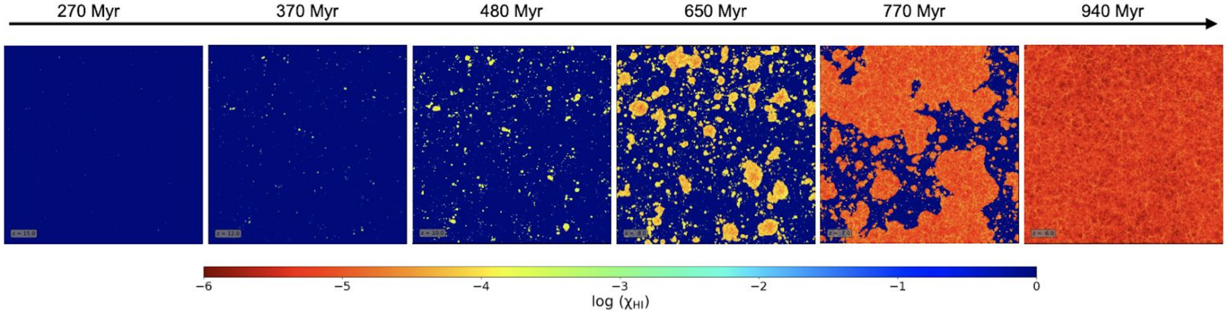
Of the hydrogen ionizing photons produced within a source, only a fraction can escape the galaxy and contribute to reionization. This '*escape fraction*', which is a complex function of the gas/dust content, feedback, and line of sight effects, remains a major unknown for reionization, with both theory and observations finding values ranging widely for individual galaxies. However, forthcoming observations, with JWST and ELTs will identify strong ionizing photon leakers up to z~9 through a variety of indirect methods. This escape fraction, as well as the production efficiency of ionizing photons, will be crucial in determining the relative contribution of galaxies, black holes and other exotic scenarios (such as dark matter annihilations and decays) to reionization. Indeed, while most works focus on galaxies being the key reionization sources, the role of low to intermediate mass black holes, hosted by faint AGN, remains an open question. As a number of theoretical works have shown, the secondary ionizations from black hole activity and the suppression of the most numerous low-mass galaxies due to the rising UV background, could both allow significant reionization contributions from low mass black holes. Large-scale surveys with Euclid and LSST will be invaluable in pinning-down high-redshift AGN number densities to shed more light on this issue.

Further complicating the issue, this heating background could slow the progress of reionization by photo-evaporating gas from low-mass halos. However, this impact critically depends on the emissivity of reionization sources, the strength of the heating background and the relative redshifts of the source, and that at which the background finally affects the gas mass of an object, to name a few. As a result, the impact of the heating background on star formation remains an open question. Self-consistently coupling galaxy formation and reionization on large-scales with high-resolution is the key goal of many ongoing theoretical models (Bluetides, Dragons, Astraeus). From the observational end, observations of the faintest galaxies with the JWST will elucidate the faint end of the ultra-violet luminosity function to yield insights on the impact of reionization feedback.

5. What was the topology and history of reionization?

The past decade has witnessed the emergence of a concordance picture in which reionization ended within the first billion years. This has been complemented by observations (HST, Subaru and VISTA telescopes) that provide tantalising glimpses of galaxy formation at the same epochs. Despite this enormous progress near the end of the EoR, details of both the history and topology (starting close to galaxies and progressing outwards -- '*inside out*' -- or the reverse) remain poorly understood. This is because they depend on a number of inter-linked processes including the abundance and ionizing photon production efficiency of galaxies, feedback from supernovae and photo-ionization, ionizing photon escape, and the space- and time-dependent density of the IGM. A huge amount of theoretical effort has been dedicated to answering these questions using high-resolution, large-scale models with some now close to the resolution and volume limits required (e.g. Bluetides, Astraeus).

In the next decades, forthcoming radio observatories (e.g. SKA and Hera) aim to map out the reionization topology using the 21cm emission from the spin-flip transition of neutral hydrogen. However, confirming the high-redshift nature of this emission, understanding the reionization sources and inferring whether the reionization topology DOES BLA OR BLA will require cross-correlating 21cm data with an independent galaxy data-set. Galaxies identified by means of their Lyman Alpha emission (1216 A) will provide the ideal galaxy data-set given their precise redshifts, as well as the influence of HI on the Lyman Alpha emission profile. A high-priority challenge will be to identify surveys that – combined with SKA – will maximize the extraction of information otherwise not attainable with SKA data alone and determine the survey strategies (areas/depths) that will maximise the synergy between the SKA and state-of-the-art (LAE and LBG) observatories including JWST, ELT, EUCLID, and existing 8-10m class telescopes.



Hutter et al. (2020): An example of how reionization proceeds through cosmic time using a numerical model (Astraeus) that couples reionization and galaxy formation in the first billion years. The numbers above the arrow show the age of the Universe; blue colours show neutral hydrogen while red/yellow colours show ionized regions. As shown, starting off filled with neutral hydrogen, the Universe is slowly reionized within the first billion years by (mostly) star formation. Of course, the exact shapes and distribution of the ionized hydrogen depend on the sources and their distribution as discussed above. The color bar shows the fraction of hydrogen that is neutral.

Galaxy Evolution

Background

The evolution of galaxies cannot be separated from the evolution of structure in the Universe in the form of virialized structures known as dark matter (DM) halos. The multifaceted challenge of understanding the evolution of galaxies comes down to understanding how matter cools and assembles near the centers of halos, how star-formation is started, proceeds and is influenced by feedbacks and radiation, giving rise to a remarkably diverse observed galaxy population that nonetheless shows great regularity in its properties as evidenced by mass scaling relations and a multitude of correlations between a diverse array of parameters. State-of-the-art hydrodynamical simulations and semi-analytical models capture this diversity quite accurately, but the physical models that describe all processes (with the exception of gravity) are prescriptions that work ‘in practice’ and not *ab initio* physical calculations. At the same time, empirical models, using the DM halo mass distribution as a starting point, seem to give a reasonably good description of how the occupation of halos by galaxies evolves over time, in particular in terms of the stellar-to-halo mass ratio. This success is enabled by the now accurate and precise measurements of the evolution of the star-formation activity and the mass function of galaxies across cosmic time (Fig XXX).

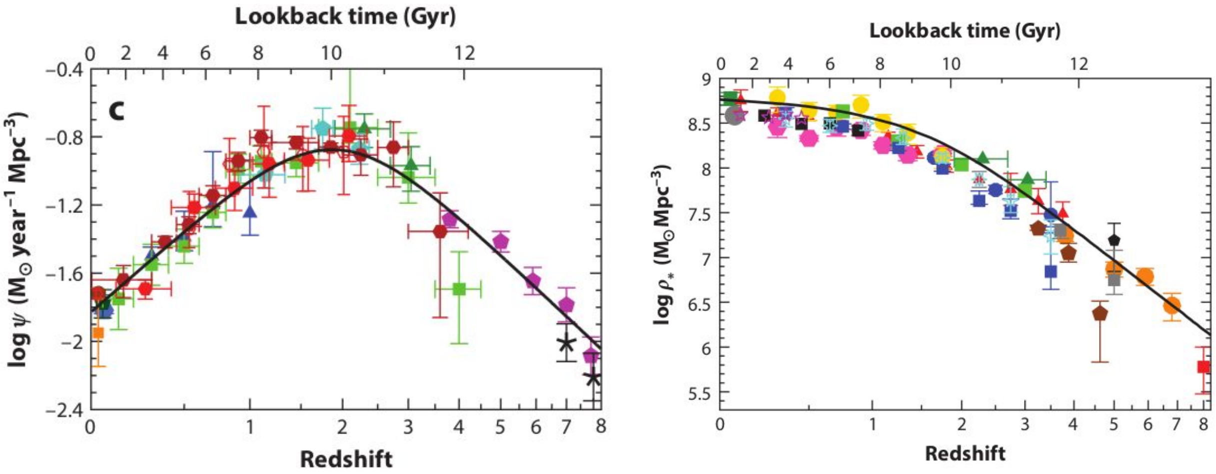


Fig XXX. Madau & Dickinson 2014: Evolution of cosmic (volume averaged) star formation rate density, and its integral, the sellar mass density. The star formation rate density begins low at early times, increases to a peak after around 4 billion years, before steadily declining into the modern, evolved, usinverse.

A central tenet is that most of the ordinary, non-dark matter (usually referred to as baryonic matter) is *not* in the form of stars and does not even reside inside galaxies (even though it is believed that perhaps the majority of matter at one point resided in a galaxy for at least some time). Therefore, understanding galaxy evolution comes down to understanding the entire ‘*baryon cycle*’ (Fig XXX), i.e., in what form the material outside galaxies -- in the inter- and circumgalactic medium (CGM) -- exists, and how material cycles through the different phases of low-density warm/hot gas in the halo, and cool/cold gas and dust within the galaxy. The processes that govern this cycle determine how much fuel for star formation is available at any point in time, and act as a backdrop for understanding the star (and galaxy) formation process itself. Both in hydrodynamical simulations and in observations, the rate of star formation is to first order predicted, at all cosmic times, by the amount of cold, dense gas available. But how efficiently and where exactly in a galaxy stars form out of the dense gas is only understood in a crude sense, limiting the fidelity of even the most recent hydrodynamical simulations in terms of galaxy structure.

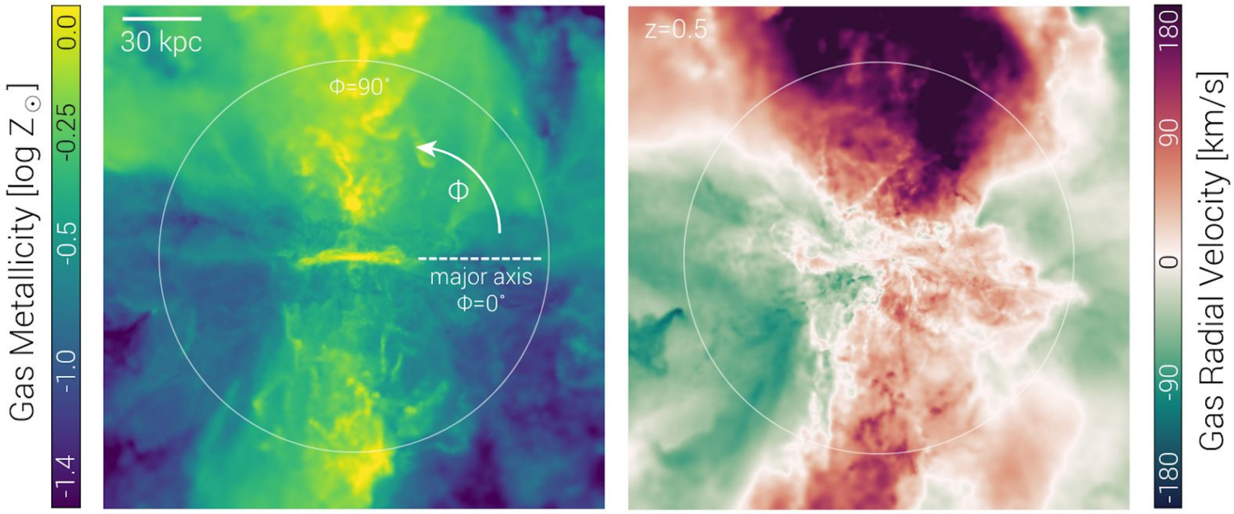


Fig XXX: the baryon cycle is complex. Inflows and outflows interact. It all takes place in the CGM. The CGM is thus an interface regime/reservoir, which encodes both cosmic gas accretion (from the outside) and feedback (from the inside). It has a multitude of observable signatures, in both gas absorption and emission, particularly in the UV and x-ray.

Another key challenge for galaxy formation models has been to explain the inferred low star formation efficiency (the conversion of gas mass into stellar mass) in galaxies and dark matter halos. Indeed, a key prediction of LCDM cosmology is the halo mass function, which is known to diverge from the observed stellar mass function of galaxies at both high and low masses. This suggests that the star formation efficiency must drop in these two different regimes. To reconcile theory with observations, state-of-the-art models and simulations invoke **`negative feedback’** from various astrophysical processes. Essentially all models need to invoke `supernovae feedback’ at low masses, and, at high masses, `AGN feedback’ from the central supermassive black holes that are known to populate the centers of all massive galaxies: energy injection into the halo in the form of jets and outflows is believed to offset cooling and thereby prevents the replenishment of the galaxies’ cold gas reservoir, inhibiting the star-formation process. This mechanism must act efficiently, as many present-day massive galaxies have been largely devoid of star formation for half the age of the Universe. The only way for these galaxies to grow is through merging with other galaxies, adding to their stellar mass budget and altering their dynamical structure.

The regularity of the formation process of massive galaxies, and the imprint of the various evolutionary stages (fast initial growth, steady star formation, quiescence), are evidenced by the tight, empirical correlations that have proved to be a powerful tool to understand galaxy evolution. Dynamical scaling relations such as the Fundamental Plane and the Tully-Fisher relation connect the total mass (set by the halo) with the baryonic mass (set by the integral of cooling/star formation and outflows), and the close connection between galaxy structure and age serves as a strong constraint on any galaxy formation model.

In summary, a very solid framework -- based on excellent observations of galaxies across cosmic time -- exists to describe the evolution of galaxies, but many missing, elementary parts of the physical model are needed to explain the regularity and rich, detailed properties of galaxies and their evolution through cosmic time.

Key questions

In order to make progress the following key questions need to be addressed:

1. What DM halo properties (besides mass) determine the properties of a galaxy?
Current empirical models of galaxy evolution have mostly focussed on reconstructing the stellar-to-halo mass relation and its evolution. But halo mass is clearly not the only parameter that determines the fate of the galaxy it hosts. The age of the halo, its concentration, its angular momentum, its merger history and its large-scale environment must play crucial roles, which must be understood if we aim to explain the diversity of properties, even of galaxies with the same halo mass. Furthermore, it has been argued that also the baryon fraction, which is usually set to the cosmological value, might not actually be constant with halo mass and cosmic time, complicating the interpretation of the stellar-to-halo mass relation and galaxy evolution in general.

2. What is the thermodynamic state of the gaseous halos that surround galaxies?
Observations and recent, high-resolution simulations have hinted that the physical properties of this circumgalactic medium are much more complex than a quasi-static, homogeneous, virialized hot atmosphere of classical theory. The recent idea that a large amount of relatively cool material exists in the form of small clouds in an environment of much hotter and less dense gas is a major observational puzzle, with no firm theoretical understanding of the physical origin of such cloudlets. The presence of cold gas in the CGM will also modulate the halo cooling process, as well as the fueling of star formation and the coupling of `feedback’ to the halo. What observational signatures does the CGM present? What do they tell us about the cosmic baryon cycle, i.e. the presence, properties, and interaction between inflows and outflows? Furthermore,

it is crucial to know the angular momentum of infalling gas to understand its role in quenching star formation activities.

3. How much material escapes from and cycles back to galaxies, and what drives these flows?

The enriched gas surrounding galaxies serves as direct evidence for the importance of outflows. Outflows, driven by stellar winds and supernovae and/or Active Galactic Nuclei (AGN) are routinely observed, but it has so far been impossible to quantify the amount of mass in these flows, how they vary over time, and how enriched material cycles back onto the galaxy at later times. Yet understanding this `feedback' and the associated cycle determines the stellar mass of the galaxy and its chemical composition. It will be crucial to understand the flows of gas within galaxies as well, as this determines the degree of mixing, the composition of newly formed stars and the level of chemical enrichment of outflows.

4. How do stars form out of dense, molecular gas?

There is a tight correlation between the amount of dense gas and the amount of star formation. That is, the amount of fuel in a galaxy determines the amount of star formation: the average efficiency is assumed to be constant. Combined with the outflows of gas driven by processes associated with the star formation itself, this gives rise to the idea of self regulation: more cold gas leads to more star formation, driving stronger outflows, which reduces the gas content and the level of star formation. However large variations in the star formation efficiency are observed within galaxies, which over time leads to large departures in morphology compared to what is expected for a constant efficiency (as is assumed in hydrodynamical simulations). This is closely related to understanding the conditions under which molecular gas is formed, becomes self-gravitating, and star formation is triggered. This includes knowing the role played by mechanisms like cloud-cloud collisions, dynamical/orbital features, or turbulence among others which have all been proposed to promote gas self-gravity. It is also important to understand how the metal enrichment of the ISM affects the efficiency of star formation across cosmic time as metals act as coolants and are fundamental in dust production, and hence shielding molecules from dissociating radiation.

The interpretation of this variation in efficiency is currently under debate, and understanding this phenomenon is crucial for understanding the diversity in galaxy morphologies.

5. What is the star-formation history of individual galaxies?

So far, we have considered the star-formation rate and the total stellar mass (which is the integral of the star formation history). The observed tight relation between the star-formation rate and the total stellar mass of galaxies, the so-called main sequence of star-forming galaxies, holding for almost 90% of the cosmic time, indicates that stars form mostly through steady-state processes rather than in violent episodes. However, this provides only the crudest picture of the formation history of the stellar population. The observed scatter of the main-sequence relation reflects the complexity of the star-formation history of individual galaxies: some simulations predict that galaxies undergo fluctuations of star formation activity on short time scales (0.2-2 Gyr) that could be the results of morphological compaction, variations of accretion, or self-regulation from cooling, star formation, and outflows. However, claims have been made that the long term fluctuations (~10 Gyr) are also driven by differences in initial halo masses.

Much progress has been made in the past two decades by comparing studies of the galaxy population with cosmological simulations. However, the next step is to compare the evolutionary history of individual galaxies with evolutionary paths predicted by simulations. This requires high-quality spectroscopy to reconstruct the full star formation history of galaxies to inter the properties of their progenitors at earlier cosmic times. Without this key piece of information we would rely only on population studies at different cosmic epochs, which has been the main mode of operation of much of the field in the past two decades, but it has proven extremely challenging to reconstruct the evolutionary paths of individual galaxies. In order to compare the evolutionary histories of individual galaxies in cosmological simulations with those of real galaxies across cosmic time, their star-formation histories need to be known.

6. How does the structure of galaxies evolve?

In general terms, we know that the morphological patterns we see today (disks, bars, bulges,...) already emerged at least 10 billion years ago. At those earlier cosmic times, spheroidal galaxies, disks and bars were all less common, even though large disks have been found to exist even as early as z=4. Instead, irregular galaxies were more common, their morphologies perhaps reflecting rapid gas accretion and fast growth. As these galaxies gain in mass, they are thought to `settle’ into disks that then remain stable for many orbital periods, and for massive galaxies a gradual loss of angular momentum possibly due to merging leads to a more spherical structure. This brief evolutionary story leaves many blank spaces. In general, we do not know how, along its star formation history, a galaxy moves through the various morphological stages. Kinematic signatures across cosmic time must be combined with knowledge of star formation histories to shed light on this.

Upcoming Facilities and Future Developments

Observational

1. **The James Webb Space Telescope.** For photometric structure. Improving stellar mass estimates at mid and high-redshifts because of access to the rest NIR. Determination of fundamental galaxy properties at mid redshifts, using rest-optical emission lines.
2. **Extremely Large Telescopes.** Kinematic structure of galaxies. Diagnosing ionized gas properties at high redshifts with rest ultraviolet emission lines.
3. **Ultraviolet Telescopes.** Many of the questions relating to the thermodynamic conditions in the CGM will be answered by future generations of ultraviolet and X-ray telescopes. Because the temperatures typically lie in this hard-to-probe temperature regimes above 10^5 K, the strongest observational tracers are resonance lines that form in the UV and soft X-ray (e.g. C IV, O VI, N V, and O VII), and also the Lyman-alpha line of atomic hydrogen. With the demise of HST and its UV coverage, and, a new

UV-sensitive telescope is a must. Large apertures will be a requirement for the high-resolution, high signal-to-noise ratio measurements required to study this gas phase using absorption line measurements in background quasars. Similarly, individual galaxies will be studied by mapping the same gas in emission from requiring high sensitivity to diffuse, low surface brightness emission.

4. **X-rays.** Future telescopes (e.g. ATHENA) will provide enhanced sensitivity and spectral resolution. This will allow us to (a.) constrain the great unknown of current hydro sims: gas fractions/impact of feedback at Milky Way and group-mass scales, where different feedback models currently produce a large diversity of results. (b.) Models currently make interesting predictions for the X-ray emission around galaxies of different types, e.g. the galaxy-CGM connection as probed by high-energy observables.
5. Large HI disks (SKA), expand to higher z with HI.
6. Radio galaxy surveys (low SB to see jets, compare with X ray morphology/cavities)
7. High-S/N continuum spectroscopic of distant galaxies (like, LEGA-C). MOONS, E-ELT multi-obj spec.
8. **Far-infrared and submillimeter Facilities.** The tremendous progress in measuring star formation and towards understanding the physics underlying the star formation progress was only possible due to IR space missions (ISO, *Spitzer*, *Herschel*) and the advent of ALMA. As the most important cooling lines for the ISM are in the FIR and the properties of dust and its role in the star formation process are far from understood, new observing facilities in the FIR to submm domaine with survey capabilities that deliver higher angular resolution and higher sensitivity are required to make progress on uncovering the physics and providing a full census of the star formation activity (cold gas reservoir) across cosmic time. However the lack of new FIR space instruments that could allow for high resolution access to FIR key lines affects our understanding of the critical phase of star formation that is hidden from optical light.

Theoretical

1. Generalization of Halotools / Universe Machine type methods; larger-box, high-res simulation (simply put: TNG300 volume or bigger at FIRE resolution). Needed to control for effect on baryons on small-scale clustering. Relates to cosmological parameter estimation, i.e. we cannot achieve the stated goals of “precision cosmology” with upcoming missions of enormous scope (e.g. EUCLID) with gravity-only simulations: need to develop effective tools to account for, or marginalize over, the impact of baryons.. Mention machine learning approaches?
2. Simulations: Two current frontiers to advance the realism and physical fidelity of cosmological and galaxy-scale simulations: **(i)** missing/new physics. While this used to include magnetic fields, these are becoming increasingly common in hydrodynamical simulations. Still not typically included, and of undoubted importance: cosmic rays (production, transport, and coupling), and full radiation hydrodynamics -- RHD; i.e. self-consistent coupling, on-the-fly, of the radiation emitted by stars, SMBHs, hot gas, and other sources. **(ii)** resolving the cold phase of interstellar gas. This is the reservoir

from which stars are actually born, which in current cosmological (large-volume) simulations is modeled in rather crude approaches, e.g. as an unresolved, two-phase medium. But star formation occurs at scales much smaller than the galaxy as a whole. Further, new and complex physics (such as non-ideal MHD effects, or non-equilibrium, multi-species chemistry/cooling networks) come into play. Essential to make predictions for molecules (H2, CO) which is where much of our knowledge of the gaseous properties of high-redshift galaxies now comes from (e.g. ALMA). Handling both (i) and (ii) implies a need for very high numerical resolution, as well as efficient numerical techniques to deal with the multi-scale, multi-physics complexity of the problem.

9. Simulations: Other key questions include: (a) better method to estimate the 'stability' of cold, star-forming gas, i.e. the physical dependencies of star formation efficiency are unknown, and we need to move beyond empirical, Kennicutt-Schmidt based relations in the models for them to become predictive at these scales.