A radio view of dust-obscured star formation

Vlugt, D. van der

1 | Introduction

How the present-day Universe formed is a challenging question, and the study of galaxies (including their gas, stars and dust) can be used to help uncover the history of the Universe. In this thesis, we use radio observations to understand how galaxies and star formation evolved from a period when the Universe was only ∼1 billion years old to the present day. Remarkably, the foundations for the formation of galaxies were established shortly after the Big Bang, more than 13 billion years ago. We thus begin with a short introduction to the very early Universe before discussing the specifics of tracing galaxy evolution with deep radio observations. We then give an overview of the scientific chapters of this thesis. We end this chapter with a brief look toward the future.

1.1 The ΛCDM Universe

In 1936, Hubble discovered that the radial velocities of galaxies are related to their distance by

$$v = H_0 d\,, \tag{1.1}$$

where d is the proper distance, an invariant distance measure which is the same for all observers, and H_0 is the Hubble constant, which is measured to be approximately 70km s^{-1} Mpc^{-1} (e.g., Bennett et al. 2013; Brout et al. 2022). The finding of Hubble is best explained by an expanding Universe.[1] Extrapolating back in time would mean that the current observed Universe must have originated from an extremely dense and hot state. This theory is known as the hot Big Bang which can be described with the ΛCDM model. As the currently accepted cosmological model, ΛCDM is the most simple parametrization of the hot Big Bang model and describes the evolution of the Universe with only six parameters. The discovery of the cosmic microwave background (CMB) lends strong support for the hot Big Bang (Penzias & Wilson 1965). The CMB is the afterglow of the early Universe and formed about 380,000 years ago when, after cooling, the protons and electrons in the primordial plasma recombined. Consequently, the photons, previously bound to the matter, could move freely and the Universe became transparent to photons. These photons are observed

[1] Other theories like the steady state model were long considered but have fallen in disfavor in light of the discovery of the CMB; see also Section 1.4.

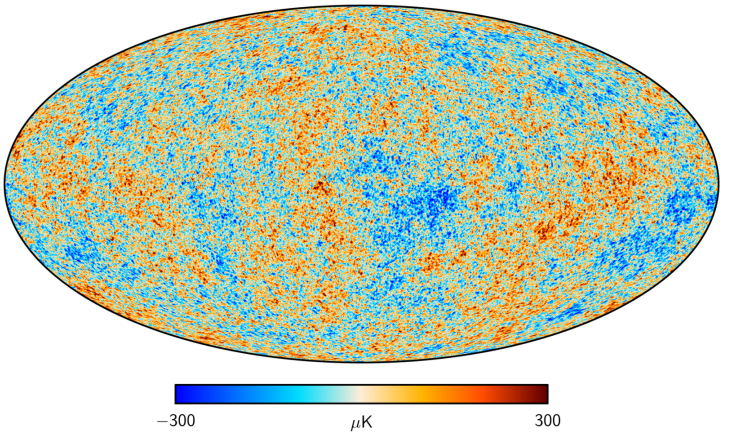

Figure 1.1: The temperature fluctuations in the cosmic microwave background from the four-year Planck satellite observations from the Planck Collaboration et al. (2020). The CMB is the afterglow of the early Universe and constitutes the earliest observable light from the Universe. The CMB is highly uniform and has an average temperature of 2.73 K. The color scale in this figure represents the small deviations in temperature of the order of μK. Slightly over-dense regions in the CMB give rise to future sites of galaxy formation.

as the CMB, redshifted to microwave wavelengths. High-precision measurements of the CMB, shown in Fig. 1.1, show tiny temperature fluctuations at the 10^{-5} level (Baumann 2009) which are associated with the primordial density fluctuations of the Universe.

The primordial fluctuations are thought to have been seeded in the epoch known as 'Inflation' which was introduced in order to explain the flatness, homogeneity and isotropy of the Universe today. The rapid expansion allowed for the 'quantum flutterings' to become classical fluctuations detectable as temperature fluctuations in the CMB (see Fig. 1.1; Clesse 2015). The form of the CMB, an almost perfect black-body spectrum with tiny fluctuations, is in accordance with the Cosmological Principle; the Universe looks statistically the same from all the possible points of view, in all the possible directions in which it is observed. This requires a causal connection in its beginning which is explained by the period of inflation. The entire observable Universe emerged out of the same causal region before the onset of inflation (Clesse 2015).

The ΛCDM model does not explain the nature and origin of dark energy and dark matter. Dark matter was introduced to explain the observed velocity dispersions within galaxy clusters (e.g., Springel et al. 2005) and rotational velocity curves of galaxies (e.g., Rubin & Ford 1970). Both indicated the existence of a large matter component that was not observed. Observations of merging clusters indicate that dark

matter particles are collisionless and only interact through gravity. The most famous example of this is the Bullet cluster, which showed that the dark matter component of the cluster could bypass the gas (Markevitch et al. 2004). The dark matter thus did not show evidence of interactions other than gravitational interactions. However, no experiment to date has successfully detected dark matter particles. The final element of the ΛCDM model is dark energy, which is called 'dark' since its nature remains unclear. The evidence for dark energy, a force acting against the pull of gravity, comes from distance measurements of supernovae showing that the Universe expands and that this expansion is accelerated (for a review see Peebles & Ratra 2003).

1.2 Formation of galaxies

Current theories of structure formation trace structure growth from the gravitational collapse of small perturbations, seen in the CMB, in a Universe dominated by cold dark matter (CDM). In the framework of CDM models, the collapse of structures proceeds bottom-up on larger and larger scales, giving rise to a hierarchy of smaller structures that are incorporated into larger ones at later times (White & Rees 1978).

The evolution of the Universe can be distinguished into different epochs. Dark matter perturbations start to grow fast after the radiation dominated era, when the Universe was dominated by the effects of radiation.[2] Baryonic matter perturbations start to grow after recombination, they then follow the dark matter perturbations and grow in density and size until they turn non-linear. According to the ΛCDM scenario, galaxies form within the dark matter haloes out of the baryons. Unlike dark matter, which is collisionless, gas can dissipate thermal energy through radiative cooling. The hot gas in haloes, heated in viralizaton and mergers, cools radiatively and condenses as it sinks towards the potential well minimum. When a region in such a dense cloud becomes dense enough, it collapses and form stars. Many of these dense clouds together will form groups of stars and eventually galaxies.

In the first phase of structure formation, the Universe remained in a neutral state. When the first stars, galaxies and quasars formed, they produced energetic photons and began to re-ionize the hydrogen in the Universe. This started around $z \approx 15-50$ and is called the epoch of re-ionization (Loeb & Barkana 2001; Barkana & Loeb 2001). Around 1 billion years after the Big Bang (Wise 2019), re-ionization was complete, and structures in the Universe continued forming and growing. After this epoch, the Universe was full of low density ionized plasma, making galaxies observable. The

[2] The rapid expansion of the Universe during the radiation dominated epoch nearly suppresses structure growth during this period.

electrons were sufficiently far apart that the amount of scattering interactions of photons and electrons were much less frequent. At $z \approx 2$ the peak of star formation (e.g., Madau & Dickinson 2014) and quasar activity (e.g., Croton et al. 2006) took place, where most of the black hole and galaxy growth is believed to occur. Matter and dark energy are the dominant components in the present-day Universe ($z = 0$) (e.g., Frieman et al. 2008).

The formation of stars and galaxies is an inefficient process; only $\sim$ 10% of the baryonic mass is locked up in stars and cold gas (e.g., Shull et al. 2012). This means that there is a mechanism suppressing star formation by injecting energy in the interstellar medium (ISM) to prevent the gas in galaxies from cooling. Supernova explosions are such a phenomena as they inject energy in their surrounding medium and can drive galaxy-wide outflows, In addition, it is believed that supermassive black holes (SMBHs) residing in the center of galaxies (active galactic nuclei; AGN) drive outflows. Semi-analytical models have shown that these high-velocity outflows are important for regulating the growth of massive galaxies (e.g., Croton et al. 2006; Sijacki et al. 2007), as their feedback can heat the halo gas surrounding the galaxy which quenches the star formation and growth of the galaxy.

1.3 Tracing galaxy evolution

The evolution of a galaxy is shaped by many factors and complex processes. Observing and characterizing the properties of galaxies across galaxy populations and cosmic time give us one way to gain insight into these processes. Stars are the most visible constituents of galaxies, and thus understanding the evolution of star formation helps us understand how galaxies evolve.

1.3.1 Star formation

Stars form when molecular clouds collapse under their own gravity. To describe the radiation from newly formed stars, we can use a model know as a Simple Stellar Population (SSP). Such a model assumes the stars that it describes to be born at roughly the same time, and also from the same gas, meaning they have the same initial chemical composition (Bruzual 2010). A galaxy can then simply be described as a combination of SSPs. To then model the evolution of stars in the SSP, we can use stellar evolution models, which provide complete sets of stellar evolutionary tracks for stars of different mass and metallicity. The path which a star follows across the Hertzsprung-Russel diagram, a diagram showing the relation between the

luminosity of a star and their stellar classification, is called an evolutionary track. The metallicity of the stars stems from the initial composition as assumed in the SSP and the mass of stars is determined by the assumed initial mass function (IMF), which describes the initial distribution of masses for the population of stars born in molecular clouds. Commonly used forms of the IMF are the Salpeter (1955) power law, the Kroupa (2001) broken power law and the Chabrier (2003) log-normal. To convert the theoretical predictions from the SSPs into observables, theoretical and empirical spectral libraries are needed.

To characterize a galaxy across the electromagnetic spectrum, we make use of spectral energy distributions (SEDs) which describe the total energy emitted as a function of wavelength. An SED is a combination of stellar emission, possible AGN activity, and dust attenuation and emission. An SED can thus be constructed from multiple SSPs taking into account the star formation history (SFH) of the galaxy. The SFH describes the evolution of the amount of star formation over time. Since stars form metals during their evolution, the SFH also includes metallicity evolution. Because the shape of the SFH is complex, it is usually parameterized with a simple analytical function. To now construct the SED, the amount of dust in a galaxy should be taken into account. Dust can attenuate the UV light emitted by stars and this light gets reprocessed by the dust emitting in the infrared. The effect of dust can be seen in Fig. 1.2. The figure shows the unattenuated light in blue and the dust reprocessed emission in red. The comparison of the two lines, especially at optical wavelengths ($\lambda \lesssim 1.0\,\mu$m), shows how the shape of the stellar emission and infrared emission is altered by dust attenuation.

The emission of galaxies observed at different wavelengths can be modelled using an SED fitting code which uses the SSPs, SFHs and dust attenuation models to describe the observations. The fitted SED can be used to estimate fundamental properties such as the stellar mass, dust mass and the star formation rate (SFR; e.g., da Cunha et al. 2015). The SFR is an essential quantity describing the evolution of a galaxy – it describes how quickly a galaxy is converting the available gas into stars over time.

1.3.2 Tracers of star formation

SED models of galaxies have shown us how observables at different wavelengths can be linked to the physical properties of a galaxy, and specifically the star formation. In the following sections, we will discuss how different tracers can be used to study the star formation of a galaxy.

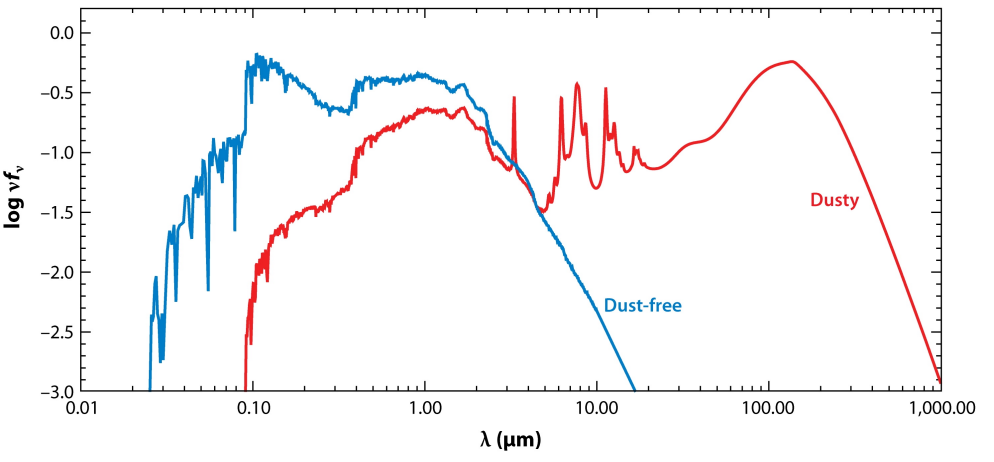

Figure 1.2: Example of a spectral energy distribution (SED) from Conroy (2013). The SED shows the amount of light emitted by a galaxy with the unattenuated light in blue and the attenuated emission in red. Comparing the unattenuated and attenuated emission shows the effect of dust attenuation on a galaxy's SED.

Ultraviolet

In principle, ultraviolet (UV) light is the most direct tracer of SFR in dust-free environments across timescales of $\sim$ 10–200 Myr (Kennicutt & Evans 2012). In the absence of an (unobscured) AGN, the unattenuated light with wavelengths ranging from $\lambda \sim 0.125\,\mu$m to $\lambda \sim 0.25\,\mu$m originates mainly from massive stars with lifetimes of 10–200 Myr and directly traces young stellar populations and hence the UV luminosity and SFR are linearly correlated (Kennicutt 1998). UV observations can thus be used to constrain the unobscured star formation out to very high redshifts ($z \simeq 9$; e.g., McLure et al. 2013; Bouwens et al. 2015; Bowler et al. 2015; Finkelstein et al. 2015; McLeod et al. 2015; Bouwens et al. 2016; Parsa et al. 2016; Mehta et al. 2017; Ono et al. 2018; Oesch et al. 2018; Bouwens et al. 2021). Measuring the SFR becomes uncertain for star-forming galaxies (SFGs) containing dust because dust can attenuate the UV emission. For SFGs that lie on the main sequence (MS, the nearly linear relation between SFR and stellar mass followed by most SFGs) this can result in an attenuation of 0.5–3 magnitudes (e.g., Salim & Narayanan 2020). For strongly dust-obscured galaxies the dust attenuation can be much higher.

UV observations thus need significant and uncertain corrections for dust obscuration. Also UV observations are unable to detect the most extreme SFGs in which star formation is completely enshrouded in dust (e.g., Smail et al. 1997; Lutz et al. 2011; Riechers et al. 2013; Casey et al. 2014a; Dudzevičiūtė et al. 2020). Therefore,

knowledge on how the dust attenuation evolves with redshift is mandatory to study the redshift evolution of SFGs.

Infrared

Dust in the ISM, heated by young massive stars, re-emits absorbed UV light at longer wavelengths and can thus be studied in the far-infrared (FIR) or sub-millimeter (sub-mm) to trace the SFR. The dust itself is partly produced in supernova explosions which occur during the last evolutionary stages of massive stars ($M > 8M_{\odot}$; e.g., Rho et al. 2008). In addition, it has been argued that dust grains grow in the ISM by the accretion of metals on to grains (Inoue 2003). Some dust also originates from the evolutionary stage of low/intermediate mass asymptotic giant branch (AGB) stars (e.g., Gehrz 1989). Because newly formed stars are bright, SFGs have most of their dust heated by young stellar populations and the emitted FIR will thus directly trace the obscured SFR. The SFR and FIR luminosity are linearly correlated (Kennicutt 1998), where the total FIR luminosity is measured between (rest-frame) 8-1000μm. Current FIR observations are able to constrain the dust content and SFR of galaxies up to a redshift $z \sim 6$ (e.g., Rodighiero et al. 2010; Gruppioni et al. 2013; Rowan-Robinson et al. 2016; Koprowski et al. 2017; Dudzevičiūtė et al. 2020; Lim et al. 2020). However, the measurements beyond $z \simeq 2$ are uncertain, as source confusion and blending limit the ability to detect faint objects at low resolution. For example, the resolution of the infrared space telescope *Herschel* ranges from $5''$ to $36''$ whereas a high-redshift galaxy has a typical size of $1''$. Neighboring galaxies can thus blend together in observations of infrared telescopes making it difficult to observe and infer the SFRs of these galaxies.

Ground-based sub-mm/mm continuum observations of dusty galaxies can help to overcome some of the problems in FIR observations (e.g., Chapman et al. 2005; Hodge et al. 2013; Swinbank et al. 2014; Dunlop et al. 2017; Dudzevičiūtė et al. 2020; Zavala et al. 2021). In particular, ground-based interferometer arrays (e.g., the Atacama Large Millimeter/sub-millimeter Array (ALMA) which will be discussed in more detail in Section 1.4.1) offer high-resolution observations and hence do not suffer from source blending, though their fields of view are typically small. In addition, singly ionized carbon ([CII]) has been found to correlate with SFR and could therefore be used as a dust-unbiased SFR indicator at low and possibly at high-z (e.g., De Looze et al. 2014; Matthee et al. 2019). [CII] was recently used in several studies conducted with ALMA to study the SFR and its evolution with redshift (Gruppioni et al. 2020; Khusanova et al. 2021; Loiacono et al. 2021).

Radio

Long-wavelength ($\lesssim$ 100 GHz) radio emission is another tracer of recent star formation (Condon 1992). Radio emission in galaxies below rest-frame frequencies $\lesssim$ 30 GHz is dominated by synchrotron radiation arising from cosmic ray electrons gyrating in the galaxy's magnetic fields (e.g., Sadler et al. 1989; Condon 1992; Clemens et al. 2008; Tabatabaei et al. 2017). These charged cosmic-ray particles are accelerated in shocks launched by supernovae of stars with M $> 8\,M_{\odot}$ in SFGs. These massive stars have lifetimes of $\lesssim 3 \times 10^8$ years, so their supernova rates are proportional to the recent SFR. However, SMBHs in AGN can also accelerate the electrons that produce synchrotron emission. The radio emission originating from star formation is a tracer of the SFR which is, unlike UV, not attenuated by dust. In contrast to FIR observations, radio observations often have a high spatial resolution and can cover larger areas of the sky than interferometric sub-mm observations with high angular resolution. Therefore, radio observations offer an alternative way of studying the star formation in galaxies.

1.4 Determining star formation with radio observations

Radio observations are used in this thesis to trace the star formation in galaxies. Radio waves of astronomical origin were first discovered in 1931 by Karl Jansky using a self-constructed radio telescope. Since then, the capabilities of radio telescopes have been immensely improved. Not only have the antennas themselves been upgraded, but also new techniques such as interferometry, developed in the 1950s, have been introduced.

Aperture synthesis

Before 1950 it was impossible to connect observed radio emission with known objects observed at optical wavelengths. The angular resolution of the single-dish radio telescopes that existed back then was too bad. Fortunately, interferometry offered a solution, and Martin Ryle developed the aperture synthesis technique needed to conduct radio observations using multiple antennas. The angular resolution of a telescope is given as $\theta \sim \lambda/D$ where λ is the wavelength and D is the diameter of the telescope. This means that to reach a resolution of $\sim 1''$ at $\lambda \sim 1$ mm, a 2km diameter dish would be needed. Instead, an array of smaller dishes can be used to achieve

high angular resolutions at radio frequencies. When observing a galaxy with an array, the signal from the observed object arrives at each antenna at a different time (due to different travel lengths) depending on the location of the antenna in the array. In aperture synthesis, the signals are then combined in a correlator, where the time delay is measured and compensated for. The time delay gives positional information about the emitting object. The distance between the antennas is called the baseline length and in an interferometer the angular resolution is now, instead of one number, a range of angular sizes: $\lambda/B_{\rm max} < \theta < \lambda/B_{\rm min}$ with $B_{\rm max}$ the maximum baseline in the array and $B_{\rm min}$ the minimum baseline in the array. The full potential of aperture synthesis is only realized when large numbers of antennas are used simultaneously.

One pair of antennas will only sample one baseline in one direction whereas three antennas will give three baselines, which means sampling the source six times. Thus, for N antennas the source will get sampled N(N-1) times. In addition, more antennas will sample more of the source over time using the Earth's rotation. The rotation will constantly change the position of the antennas with regard to the observed source, causing the direction of the baselines to change with respect to the observed section of the sky and thus sampling different parts of the source. With the development of aperture synthesis, high-resolution radio observations became possible, enabling easier cross-correlation with optical sources. The new technique helped pinpoint strong radio emission observed in, for example, the optically faint galaxy Cygnus A (e.g., Jennison & Das Gupta 1953) and the young supernova remnant Crab Nebula (e.g., Bolton et al. 1949). Observations using aperture synthesis also showed compact strongly emitting radio sources named quasi-stellar objects without an obvious optical counterpart (e.g., Matthews & Sandage 1963). We now know these quasi-stellar objects as quasars which are extremely luminous AGN.

Since the 1950s, not only have radio telescopes been upgraded and developed but also sophisticated calibration techniques have been introduced to produce robust radio maps. To produce a radio image using the observations of an interferometer, many steps such as phase and flux calibration, removing corrupted data, and imaging and performing a deconvolution of the image are needed. This process is called data reduction, and it is designed to mitigate known issues in radio observations. The steps to produce radio continuum observations are covered in **Chapter Two** of this thesis.

The VLA

The radio observations presented in this thesis were carried out with The Very Large Array (VLA). This array was constructed in the 1970s and uses the aperture synthesis

technique. The VLA is an interferometer situated on the Plains of San Agustin in USA at an altitude of approximately 2000m. The array consists of 27 antennas, of 25m diameter, distributed on-site in a Y-shaped configuration. The resulting total collecting area of all antennas is equivalent to a 130m dish. The distances between the antennas can be changed, giving the array the ability to increase or decrease the angular resolution. The four different configurations that can be set are named, from highest to lowest angular resolution, A, B, C and D. The original VLA underwent a major upgrade in 2012 that was emphasized by renaming it the 'Karl G. Jansky Very Large Array'. The radio observations presented in this thesis profited from this upgrade as the significantly increased bandwidths of the upgraded array now allow for observations of the radio sky down to several hundred nJy beam^{-1} sensitivities. The upgraded VLA provides continuous frequency coverage from 1 to 50 GHz, with two additional narrower bands centered around 90cm and 4m.

Number counts and luminosity functions

The depth achievable in modern radio surveys now allows for observations of faint (μJy) radio-emitting objects, thus enabling constraints on the nature of these populations (e.g., Rujopakarn et al. 2016; Murphy et al. 2017; Smolčić et al. 2017a; Owen 2018; Bondi et al. 2018; Mauch et al. 2020). The simplest statistical analysis that can be performed with a flux-limited radio survey is counting the number of galaxies. Despite the simplicity, this can provide very useful information, as the shape of the number of sources as a function of flux is tightly related to the evolutionary properties of the sources and the geometry of the Universe. If the Universe were Euclidean, the number of sources expected to be found above some flux density S would be $N \propto S^{-3/2}$, or described as the differential number counts: $n = dN/dS \propto S^{-5/2}$. The 2C radio survey (Shakeshaft et al. 1955) showed that the number counts differed from the counts expected from an Euclidean Universe, which was evidence for an evolving Universe (Ryle & Scheuer 1955). Due to some controversy, known as the 'Sidney-Cambridge controversy' (Sullivan 1984), about the understanding of source confusion in the 2C radio survey, it took the discovery of the CMB (Penzias & Wilson 1965) to confirm the prediction of the expanding Universe and the hot Big Bang model, but this also showed the power of early radio surveys.

Source counts from radio surveys are currently recognized as essential data in describing the different radio source populations. At high flux densities, the source counts are well-constrained and found to be dominated by AGNs that follow a smooth power-law distribution down to $S_{1.4\mathrm{GHz}} \sim 1\,\mathrm{mJy}$ (e.g., Condon & Mitchell 1984; Windhorst et al. 1990). Below 1 mJy, the Euclidean-normalized source counts flatten

(e.g., Richards 2000; Huynh et al. 2005; Biggs & Ivison 2006; Owen & Morrison 2008; Bondi et al. 2008; Padovani et al. 2015). It is now widely accepted that this observed flattening is due to the emergence of SFGs and radio-quiet AGN, which begin to contribute significantly only at these faint flux densities (e.g., Rowan-Robinson et al. 1993; Seymour et al. 2004; Padovani et al. 2009). Deeper radio observations and $P(D)$ analyses on confusion-limited surveys show evidence of a further steepening of the number counts below $S_{1.4\mathrm{GHz}} \sim 50\,\mu$Jy (e.g., Condon et al. 2012; Vernstrom et al. 2014, 2016; Smolčić et al. 2017a; Prandoni et al. 2018; Mauch et al. 2020). The composition of this ultra-faint radio population was previously uncertain, but it was expected from simulations and observations that the fraction of SFGs will become significant (> 60 %) below $S_{1.4\mathrm{GHz}} \sim 100\,\mu$Jy (Smolčić et al. 2017b). Constraints on the ultra-faint radio populations at high resolution are useful for predictions for future radio surveys with new and upcoming facilities such the ASKAP, MeerKAT, ngVLA, and SKA, where source confusion noise may be an issue. The number counts presented in **Chapter Two** are constrained with observations achieving sub-μJy r.m.s noise levels, and are thus probing the ultra-faint radio population.

A more advanced analysis of radio surveys involves measuring the cosmic density of sources in bins of luminosity, thus constraining the luminosity function (LF). To construct the LF, the redshift need to be established for the observed sources. Determining the LF for different galaxy populations (e.g., SFG or AGN) can help constrain the evolutionary paths of these populations. The local LF is well-established (e.g., Condon et al. 2002; Best et al. 2005; Mauch & Sadler 2007; Condon et al. 2019) and often used to constrain the LF out to higher redshifts. To do this, the local LF is evolved to fit the high-redshift data. We can distinguish two extreme cases for this evolution: 1) pure luminosity evolution (PLE), meaning the density of sources is constant but luminosities vary with cosmic epoch, and 2) pure density evolution (PDE), which implies the co-moving density of sources of any luminosity varies. Radio studies to-date have observed radio LFs but struggled to reach the knee of the LF ($L_\star$), or the point at which there is a break from the power law describing the LF, at $z > 1$. Because these studies are most sensitive to the SFG population above the knee, the density and luminosity evolution parameters may become degenerate, preventing a precise estimate of the knee location. The VLA-COSMOS 1.4 GHz Large Project and the VLA-COSMOS 3 GHz Large Project from, respectively, Smolčić et al. (2009b) and Novak et al. (2017) thus assumed pure luminosity evolution rather than luminosity and density evolution (Condon & Mitchell 1984) in order to fit the radio LF out to $z \sim 5$. Recently, Malefahlo et al. (2022) used a Bayesian approach to reach below the 5σ detection limit of Novak et al. (2017) but still only constrained pure

luminosity evolution. Enia et al. (2022) used a 1.4 GHz-selected sample to constrain the evolution of the radio LF up to $z \sim 3.5$ by fitting a modified Schechter function (equivalent to fitting both luminosity and density evolution). By constraining the LF and its evolution as is described in **Chapter Three**, it is possible to estimate the cosmic star formation rate density as a function of cosmic time.

1.4.1 The radio spectrum

Radio emission in SFGs has two components related to star formation: free-free emission and synchrotron radiation. Free-free emission is produced during an interaction between charged particles (electrons and protons). Young massive stars ($M \geq 10 - 15 M_{\odot}$) create short-lived ionized regions know as HII-region from which the free-free emission originates (Kennicutt & Evans 2012). Free-free emission provides thus a first-hand view on star formation. In addition, unlike UV emission, dust obscuration plays only a minor role, and therefore free-free emission is potentially the most accurate tracer of star formation (e.g., Mezger & Henderson 1967; Turner & Ho 1983, 1985; Klein & Graeve 1986; Kobulnicky & Johnson 1999; Murphy et al. 2012, 2015; Nikolic & Bolton 2012).

Free-free emission is described by a flat power-law $S \propto \nu^{-0.1}$ (e.g., Condon 1992); see Fig. 1.3. At rest-frame frequencies of $\nu \geq 30$ GHz the total radio emission is dominated by free-free radiation (e.g., Condon 1992; Murphy et al. 2011; Klein et al. 2018). This constitutes the faintest part of the radio SED, as can be seen in Fig. 1.3. At $\nu \geq 200$ GHz, free-free emission is no longer the dominant emission term, as thermal emission from dust grains becomes more prevalent. Free-free emission provides information on the star formation process independent of FIR and synchrotron emission. However, isolating the free-free component and measuring its flux density is difficult. Recently, studies have now done this successfully in samples of galaxies beyond the local universe (Algera et al. 2021, 2022). The work in these companion studies was made possible by the ultra-deep 10 GHz radio observations presented in **Chapter Two**.

In comparison to free-free emission, synchrotron emission can be more easily observed. It arises from fairly old ($\geq 10^7$ yr) electrons that traveled some distance (≥ 1 kpc) from the supernova remnants associated with massive stars ($M \geq 8 M_{\odot}$) in which they were accelerated (Condon 1992). The emission is described by a power-law $S \propto \nu^{-0.7}$ (e.g., Condon 1992) as can be seen in Fig. 1.3. Such massive stars are short lived, and as such their supernova rate is related to a galaxy's recent star formation rate, albeit with a short delay (≤ 30 Myr; Bressan et al. 2002). Problematically for

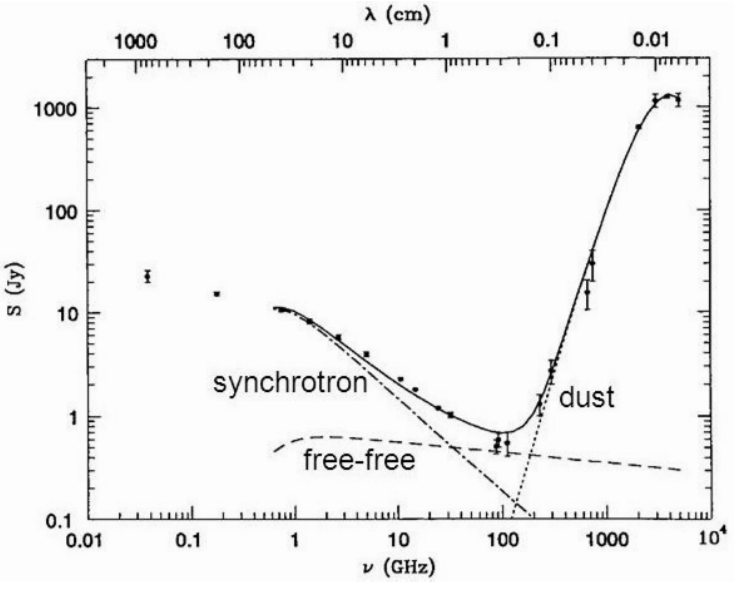

Figure 1.3: Example of a spectral energy distribution (SED) of the starburst galaxy M82 adapted from Klein et al. (1988). The SED shows the flux density as a function of frequency, in this case, from the radio to the far infrared. The solid line shows the total emission from the galaxy. The dot-dashed line shows the synchrotron emission and the dashed line shows the free-free thermal emission. The dotted line $\nu \geq 100$ GHz shows the dust thermal emission. At low frequencies, the synchrotron component is dominant, while at $\nu \geq 30$ GHz, free-free emission overtakes the synchrotron contribution. At higher frequencies, $\nu \geq 200$ GHz, dust overtakes free-free emission.

studying star formation with radio observations, not only supernovae but also AGN can produce powerful radio emission.

AGN

When studying the star formation in galaxies using radio observations, it is important to verify that the observed radio emission is indeed produced by stars. AGN and SFGs are difficult to disentangle in the radio regime because an accreting SMBH in an AGN can also accelerate the electrons that produce synchrotron emission. Synchrotron emission can thus not only be used to study star formation but also to study the black hole accretion activity in the Universe (e.g., Jarvis & Rawlings 2000; Smolčić et al. 2009a; Rigby et al. 2011; McAlpine et al. 2013; Best et al. 2014; Delvecchio et al. 2014, 2017; Sabater et al. 2019). When radio luminosity is used as a proxy for tracing star formation, AGN are considered as contamination of the star-forming sample.

AGN are traditionally classified in two main categories: steep ($\alpha <$ -0.5) and flat-

spectrum ($\alpha >$ -0.5), where the radio spectrum is described by a power-law $S \propto \nu^{\alpha}$. Not only SFGs but also AGN can thus have a steep radio spectrum ($\alpha \sim -0.7$). Flatter spectra are typical of compact core emission and are thus only associated with AGN. Constraining the radio spectrum can thus be one method to separate AGN from SFGs. However, using only this method would still leave many AGN in the sample. In addition, the diagnostic might not be available for the vast majority of the sample since multiple radio continuum frequency observations are needed. Therefore, several other methods have been developed for identifying different types of AGN and separating them from SFGs (e.g Hickox et al. 2009; Mendez et al. 2013; Smolčić et al. 2017b; Delvecchio et al. 2017) using mid-IR data, far-IR data, X-ray information and multi-band optical/IR photometry. A companion study to this thesis by Algera et al. (2020a) used the ultra-deep radio survey presented in **Chapter Two** and the multi-wavelength data available over the field to investigate the composition of the radio population. The sources identified as SFGs are used in **Chapter Three** to determine the evolution of SFR with redshift.

Radio source redshifts

To perform more advanced analysis with a sample of radio sources, such as measuring the LF, redshifts are needed. Fitting a SED model to the stellar and dust emission is a commonly used tool to derive the redshift. However, since the radio spectrum is a featureless power law (Fig 1.3), it can not, alone, be used to constrain the redshift. Multi-wavelength data, either from photometric or spectroscopic surveys, are thus needed to find redshifts for radio sources. Photometric surveys give the average brightness of a galaxy over a fixed wavelength band, with a single measurement point per band, whereas spectroscopic surveys deliver a high-resolution spectrum of a galaxy. Probing photometric redshifts with photometric surveys is cheaper and can be done over larger survey areas, but it delivers more uncertain redshifts. Obtaining a full spectrum with a spectroscopic survey is time consuming but gives more information with details on atomic and molecular lines used to infer the spectroscopic redshifts. These redshifts are measured to high precisions and are therefore more reliable than photometric redshifts. However, it is not cheap to obtain spectroscopic redshifts for large numbers of galaxies, particularly when they are dusty. Targeting CO transitions at millimeter wavelengths with, e.g., ALMA, is a relatively efficient method of obtaining accurate redshifts for radio sources with photometric redshifts.

Redshifts via CO line scans with ALMA

ALMA is an interferometer that started operating in 2011. It is currently the largest millimeter telescope in the world (Wootten & Thompson 2009), and is located on the Chajnantor plateau, in the Atacama Desert, Chile, at an altitude of over 5000m. The array is comprised of 66 antennae: the main array of 50 12m antennae, an array of twelve 7m antennae (Atacama Compact Array; ACA), and a further four 12m antennae (Total Power Array; TPA). The main array offers angular resolutions from a few arcseconds to milli-arcseconds and is a factor of 10-20× more sensitive than its precursors for spectral line detections. Additionally, ALMA is 10-100× more sensitive for continuum detections. ALMA has enabled many advances in our understanding of star formation at $z > 1$ (Hodge & da Cunha 2020). Of particular interest for this thesis is ALMA's frequency coverage (35-950 GHz) and sensitivity which enables it to target CO transitions and thus obtain spectroscopic redshifts for gas/dust-rich sources at high ($z > 1$) redshift. This technique is further discussed and used in **Chapter Five**.

1.4.2 Dust-unbiased star formation

Radio observations probe recent star formation ($\lesssim 30$ Myr; Bressan et al. 2002) and are especially valuable for three reasons: 1) Radio emission is not overwhelmed by stellar populations older than 10^8 yr; 2) radio maps can be made with sub-arcsecond resolution; and most importantly 3) radio emission is not attenuated by dust, so that even the most extreme SFGs enshrouded in dust can be observed (e.g., Condon 1992). The discovery of sub-millimeter galaxies (SMGs) (e.g., Downes et al. 1999; Smail et al. 1999; Frayer et al. 2004) showed the importance of a more complete census of star formation history over cosmic time. In particular, since these galaxies are heavily attenuated in the rest-frame UV (e.g., Simpson et al. 2014; Franco et al. 2018; Dudzevičiūtė et al. 2020), they are often missed in UV observations. Via high angular resolution ALMA observations, we can uncover faint SMGs missed by UV-surveys: also classified as 'optically dark' sources (e.g., Gruppioni et al. 2020; Gómez-Guijarro et al. 2022; Shu et al. 2022; Xiao et al. 2023). Radio observations can also be used to identify 'optically dark' galaxies as has been recently shown (Talia et al. 2021; Enia et al. 2022). To determine the SFR of these galaxies observed in radio surveys, the radio luminosity needs to be linked to the SFR. However, the steps to translate observed synchrotron emission to a SFR, such as supernova explosion, acceleration of electrons by the supernova, propagation of cosmic rays and energy loss, are non-trivial and poorly understood. Fortunately, observations show a tight and ubiquitous

correlation between FIR and radio luminosity, offering a way to derive SFRs from radio observations.

The far-infrared-radio correlation

Given the different timescales of the various processes involved, it is quite remarkable that radio emission triggered by star formation is empirically found to correlate well with the FIR emission of SFGs: the FIR-radio correlation (e.g., Helou et al. 1985). The correlation is parameterized using the dimensionless parameter q_{TIR} as introduced by Helou et al. (1985):

$$q_{\mathrm{TIR}} = \log_{10}\left(\frac{L_{\mathrm{FIR}}}{3.75 \times 10^{12}\ \mathrm{W}}\right) - \log_{10}\left(\frac{L_{1.4\,\mathrm{GHz}}}{\mathrm{W\ Hz^{-1}}}\right). \tag{1.2}$$

where FIR-luminosity L_{FIR} is measured between 8 and 1000 μm, and $L_{1.4\,\mathrm{GHz}}$ is the radio luminosity at 1.4 GHz. The radio SFR can now be derived using the correlation as $\mathrm{SFR} \propto 10^{q_{\mathrm{TIR}}}\, L_{1.4\,\mathrm{GHz}}$ (e.g., Delhaize et al. 2017). Besides being used to calibrate radio luminosity as a tracer of SFR, the FIR-radio correlation is also often used for the classification of galaxies. For example, a sample of galaxies observed in a radio survey used for constraining the star formation at a given redshift should only consist of sources with radio emission originating from star formation. Therefore one would ideally quantify the emission coming from star formation and AGN in all sources. It is, however, easier to simply remove sources that show an excess in radio emission compared to what is expected from the FIR-radio correlation (radio-excess AGN, e.g., Del Moro et al. 2013; Delvecchio et al. 2017; Algera et al. 2020a). Radio-loud AGN are easily removed by this method, as these sources show a large offset from the FIR-radio correlation. A major uncertainty is the ability to distinguish composite sources, which emit low-level AGN emission, from SFGs (e.g., Padovani et al. 2009; Bonzini et al. 2013).

Radio-SFR calibrations and source classification rely thus on this empirical FIR-radio correlation, which appears to hold across more than five magnitudes in luminosity and persists out to high redshifts (e.g., Helou et al. 1985; Yun et al. 2001; Bell 2003), albeit with ill-constrained redshift evolution (e.g., Sargent et al. 2010; Magnelli et al. 2015; Calistro Rivera et al. 2017a; Delhaize et al. 2017). Though, there is some discussion about whether the redshift evolution can be ascribed to selection biases (Sargent et al. 2010; Algera et al. 2020b; Smith et al. 2021; Molnár et al. 2021) or whether the FIR-radio correlation should have an additional dependency on mass (Delvecchio et al. 2021; McCheyne et al. 2022). From a theoretical point of view,

the observed evolution is difficult to explain. The evolution contradicted the evolution predicted by theoretical studies: high redshift galaxies were expected to become radio-dim due to inverse Compton losses on the CMB but observations showed that these galaxies appeared increasingly bright. The exact physical processes driving the relation remain unclear, and predictions for the redshift evolution of the FIR-radio correlation are also often conflicting (e.g., Murphy 2009; Lacki & Thompson 2010). Dealing with selection biases is non-trivial and not accounting for such biases would lead to artificial observed evolution in the FIR-radio correlation (e.g., Sargent et al. 2010).

The selection biases are most easily explained by introducing K-corrections which convert the flux of a source from observed-frame to rest-frame. The strength of the K-correction depends on the shape of the SED. For example, FIR/sub-mm observations profit from a negative K-correction for high-redshift sources as these observations probe the Rayleigh-Jeans tail of the dust SED. Hence, a dusty galaxy with a fixed far-infrared luminosity and temperature will have an almost constant apparent flux density in the sub-millimeter waveband as a function of redshift. Radio observations lack this negative k-correction, and sources at higher redshift thus become increasingly faint. As a result, radio sources at higher redshifts ($z \geq 2.5$) that would be detectable with FIR/sub-mm observations become too faint for most radio surveys to be detectable. Differences in the sensitivity of radio- and FIR-surveys to galaxies at high redshift if not properly taken into account, will thus result in a biased result.

Measurements of the FIR-radio correlation also suffer from contamination of AGN. The FIR-radio correlation only applies for SFGs, since the correlation arises because the same population of massive stars that heats up dust-producing FIR emission also produces supernovae that generate synchrotron radiation. An AGN in galaxies can also be a source of radio emission. These AGN-hosting galaxies will therefore be offset from the FIR-radio correlation. Such sources should thus be removed from the sample to measure the FIR-radio correlation without introducing an artificial evolution (Molnár et al. 2018). Since the FIR-radio correlation is an integral part of the estimation of SFRs from radio observations, studying the correlation and understanding its origin remains of great importance. The impact of the assumed FIR-radio correlation on the measurements of the evolution of the SFR with redshift will be further discussed in **Chapter Three**.

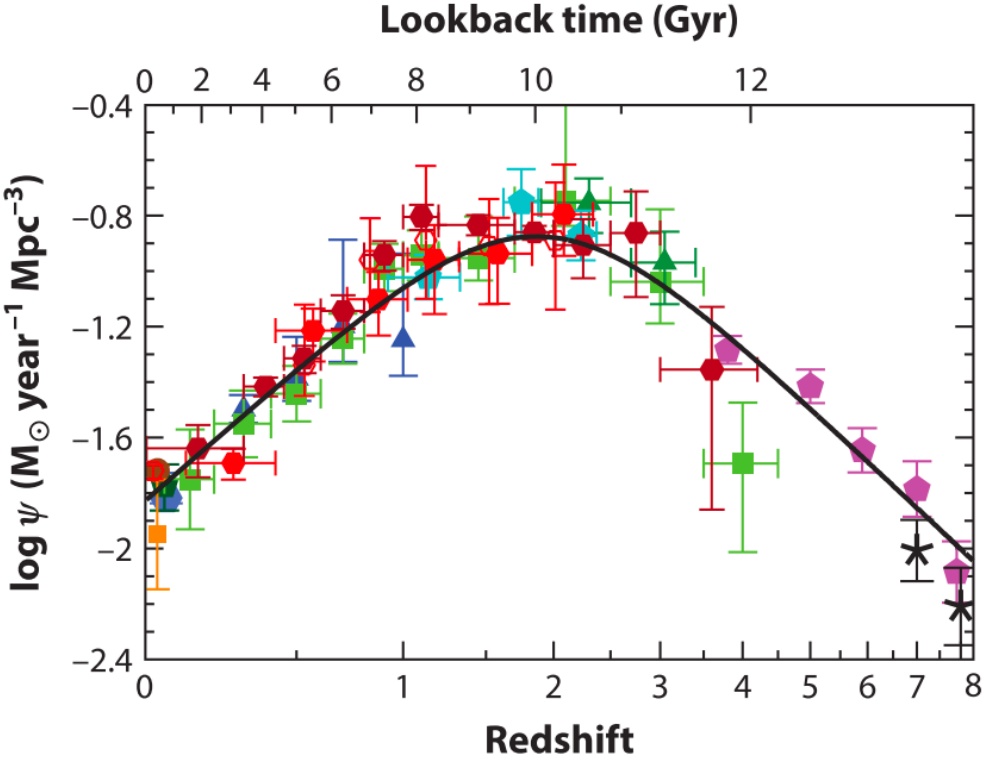

Figure 1.4: A compilation of measurements of the cosmic star formation rate density (SFRD) history by Madau & Dickinson (2014). The function shows the amount (mass) of stars created from gas per year and cubic mega-parsec as a function of cosmic time. From the compilation, we can see that the SFRD rapidly rises up to $z \sim 2$ and declines afterwards. Red and orange points represent measurements from IR observations, while green, blue and purple points are from UV observations. Most measurements at high redshift ($z > 2$) depend on UV observations.

1.5 The cosmic star formation rate density

The evolution of the star formation rate per unit volume, or the star formation rate density (SFRD; Madau & Dickinson 2014), as a function of redshift is one of the most fundamental tools to understand how star formation proceeds globally across cosmic time. It gives the SFR across all galaxies, at a given moment in cosmic history. Several tracers can be used to trace the SFR – and thus the SFRD – as discussed in Section 1.3.2. A compilation showing various UV (e.g., McLure et al. 2013; Bouwens et al. 2015; Bowler et al. 2015; Finkelstein et al. 2015; McLeod et al. 2015; Bouwens et al. 2016; Parsa et al. 2016; Mehta et al. 2017; Ono et al. 2018; Oesch et al. 2018; Bouwens et al. 2021) and IR-based (e.g., Rodighiero et al. 2010; Gruppioni et al. 2013; Rowan-Robinson et al. 2016; Koprowski et al. 2017; Dudzevičiūtė et al. 2020; Lim et al. 2020) measurements of the SFRD from Madau & Dickinson (2014) is shown in Fig. 1.4. As can be seen from this compilation, the SFRD increases from the formation of the first galaxies at $z > 10$ up until $z \sim 2$, after which the SFRD declines. The peak at $z \sim 1-3$ is known as the 'epoch of galaxy assembly'. During this period half the stellar mass observed in the present-day universe was formed. Studying galaxies at this epoch is thus crucial to understand how galaxies evolved from this period to the present day. In addition, it is well established that the majority of the star formation happens on galaxies that lie on the MS.

Owing to the sensitivity of telescopes such as the Hubble Space Telescope (*HST*), most of the studies measuring the SFRD at high-redshift ($z > 2$) use UV observations (Fig. 1.4), a wavelength range prone to dust attenuation. Therefore, understanding how much of the star formation is obscured by dust is mandatory to study the redshift evolution of the SFRD, particularly as dust obscuration might play a non–negligible role even out to $z > 4$ (Casey et al. 2018; Bouwens et al. 2020). This motivates the work using radio observations as a dust-unbiased tracer of the SFRD presented in **Chapter Three**.

1.5.1 Dust-obscured star formation

The importance of dust emission, and thus dust-obscured star formation, was first seen from the comparison of the FIR background emission with the UV and optical background. The FIR background emission was first measured with the Cosmic Background Explorer (COBE) satellite. This instrument set out to assess the energy spectrum of the Universe at far-infrared wavelengths and additionally measured an emission that could not be explained by galactic sources: the cosmic infrared

background (e.g., Puget et al. 1996). The infrared background is approximately comparable to that of the UV, which suggests that total luminosity UV surveys only probe about half of all the star formation that has ever occurred. The other half of the star formation is thus obscured by dust.

Since then, many observational studies have been conducted and reached varying conclusions about the relative contribution of the dust-obscured population to the cosmic SFRD. To study star formation at high redshifts, FIR observations are key, since they trace the dust in galaxies, as explained in Section 1.3.2. However, the constraints on the SFRD beyond $z \simeq 2$ are uncertain because the measurements of the FIR LF used to derive the SFRD become more challenging. Source confusion and blending limit the ability to detect faint objects in low resolution *Herschel*/SPIRE observations at $z \simeq 3-4$. Such observations are thus biased towards a less representative population of bright sources. In addition, these observations can be significantly contaminated by AGN, as these sources are more numerous at high redshift (Gruppioni et al. 2013; Symeonidis & Page 2021).

ALMA observations have helped to overcome some of these problems by offering high-resolution observations and being less susceptible to AGN contamination, as they are predominantly sensitive to the cool dust in the star-forming population at high redshift (Hodge & da Cunha 2020). But even with these advantages, sub-mm observations are still impractical to carry out over the large areas that are needed to overcome cosmic variance, which can have a strong impact on any counting statistic (e.g., Moster et al. 2011; Simpson et al. 2019; Gruppioni et al. 2020; Loiacono et al. 2021), because of the small field of view. Cosmic variance in the sub-mm can be overcome by combining a wide-field single dish observation with interferometric follow-up observations (Simpson et al. 2020). However, such observations are expensive. Therefore, we still have poor constraints on the high-redshift dusty SFG population, which was illustrated by Casey et al. (2018) using two extreme models as shown in Fig. 1.5. The dust-poor line in the figure assumes the dusty SFG population peaks at $z \sim 2$, and that UV-luminous sources dominate the SFRD at $z \sim 4$; sources such as the 'optically dark' sources mentioned in Section 1.4.2 are rare in this model. The dust-rich line assumes dusty SFGs are as numerous at $z \sim 2$ as at $z \sim 4$. Neither of the two extreme models could be ruled out by the data sets available.

Recently, Zavala et al. (2021) presented a FIR-based estimate of the dust-obscured star formation out to $z = 6-7$ using the Mapping Obscuration to Reionization with ALMA (MORA) survey. Using only 13 sources detected sources at 2mm, number counts at 1.2 and 3mm, and the semi-empirical modelling from Casey et al. (2018), they estimated the dust-obscured star formation to be $\sim$35% of the total SFRD at

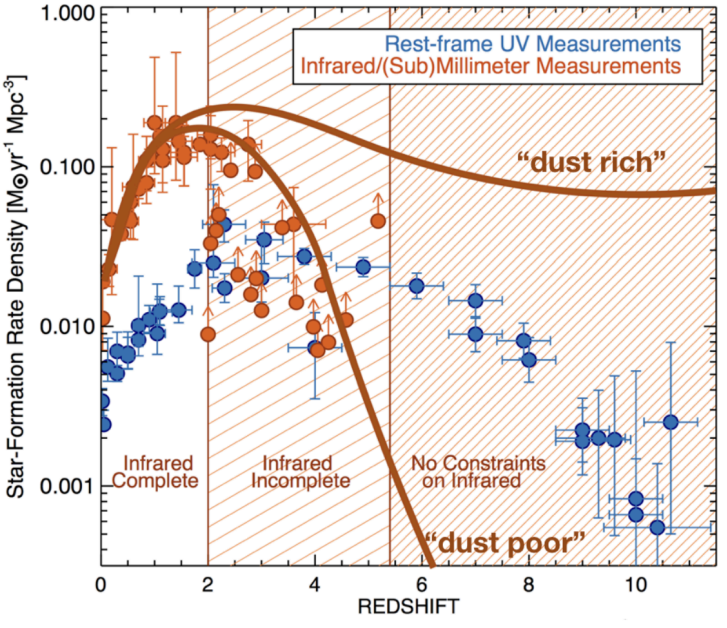

Figure 1.5: A compilation of cosmic star formation rate density measurements at UV wavelengths (blue points) and millimeter wavelengths (orange points) by Casey et al. (2018). The UV measurements have not been corrected for dust attenuation. Using deep HST imaging, the UV measurements reach beyond $z > 10$. Observations of obscured emission in galaxies is incomplete at $z > 2$ and completely unconstrained at $z > 5$. Casey et al. (2018) tests two dramatically different possibilities for the obscured fraction of the SFRD, shown by the dark orange lines.

$z \sim 5$, albeit with significant uncertainties due to low number statistics. This study suggests that dusty star formation already plays an important role at high redshift.

In addition to FIR observations, radio observations can be a powerful way to constrain the importance of dust-obscured star formation out to high redshift (e.g., Novak et al. 2017; Ocran et al. 2020; Cochrane et al. 2023). Radio observations also have the advantage of observing a larger field of view, thus having lower costs to obtain a statistical sample with which to constrain the SFRD. Previously, Novak et al. (2017) used the VLA-COSMOS 3 GHz Large Project to derive the SFRD. By observing the 2 deg^2 COSMOS field down to an r.m.s. of 2.5 μJy, they were able to observe main sequence galaxies out to $z \sim 1.5$. To trace main sequence galaxies out to higher redshifts, however, radio observations with sub-μJy r.m.s. noise levels are necessary. **Chapter Three** describes how we use a new ultra-deep radio survey to do just this.

Thanks to their sensitivity to dust-obscured star formation, radio observations

can also be used to uncover the 'optically dark' galaxy population introduced in Section 1.4.2: extremely dust-obscured sources that are invisible even in deep optical/UV imaging. Different observational works have studied these sources and show that the contribution of 'optically dark' sources appears to increase up to $z \sim 3.5$ and decrease thereafter. However, the constraints still span a wide range of over an order of magnitude at $3 \lesssim z \lesssim 6$. For example, Franco et al. (2018) found that 20% of the 1.1mm sources are not detected with HST down to a depth of $H \sim 28$ mag, while Wang et al. (2019) and Gruppioni et al. (2020) determined that the contribution of their dark sample ranged from, respectively, $\sim 10\,\%$ of the SFRD from LBGs at similar redshifts to equal to the total extinction-corrected contribution from all the known UV-selected galaxies at $z \sim 5$. Recently, observations of 'optically dark' sources have also been conducted with James Webb Space Telescope (*JWST*; Barrufet et al. 2023; Pérez-González et al. 2023). They determined that these sources are mostly dusty SFGs and found a similar contribution as determined by Wang et al. (2019). Radio observations have been shown to identify 'optically dark' galaxies that contribute most significantly to the cosmic SFRD (Talia et al. 2021; Enia et al. 2022). **Chapters Four & Five** discuss the discovery and verification of 'optically dark' sources in new deep radio observations, as well as their contribution to the SFRD.

The COSMOS-XS survey

In order to carry out the analyses described above, this thesis presents the VLA COSMOS-XS survey, which provides new deep radio observations at two frequencies: 10 and 3 GHz (i.e., the VLA's X- and S-bands). The survey was conducted in the COSMOS field (Scoville 2007), a 2 deg^2 field which has been observed with all leading ground-based and satellite facilities, yielding a rich multi-wavelength data set (e.g., Laigle et al. 2016; Weaver et al. 2022). The COSMOS field was previously observed with the VLA at 1.4 GHz ($\sigma \sim 10-15\,\mu$Jy beam^{-1}, Schinnerer et al. 2010), and more recently at 3 GHz ($\sigma \sim 2.3\,\mu$Jy beam^{-1}), yielding about four times more radio sources compared to the 1.4 GHz data (Smolčić et al. 2017a). Although these observations provide valuable data over the entire 2 deg^2 COSMOS field, they still lack the depth to detect main sequence galaxies above $z \sim 1.5$. The COSMOS-XS survey presented in this thesis is one of the deepest radio surveys to-date, reaching sub-μJy sensitivities, and giving it the ability to trace main sequence galaxies over the full 'epoch of galaxy assembly' ($z \sim 1-3$). Importantly, it is ~ 5 times deeper than the previous 3 GHz observations conducted in the COSMOS field (Smolčić et al. 2017a). When combined with the rich multi-wavelength data, this survey yields a unique data set with which to study the faintest radio source populations that can currently be probed.

1.6 This thesis

In this thesis, we present the radio observations from the brand-new VLA COSMOS-XS survey and use them to trace dust-unbiased star formation over cosmic time. The thesis is organized as follows:

In **Chapter Two**, we discuss the details of the COSMOS-XS survey and present the radio catalogs. The deep 10 and 3 GHz observations enable us to investigate the Euclidean-normalized source counts down to the μJy level. We show that our observations are consistent within the uncertainties with other results at 3 and 1.4 GHz, but extend to fainter flux densities than previous direct detections.

In **Chapter Three**, we use the 3 GHz-selected sample of SFGs identified in the COSMOS-XS survey to study the evolution of the radio LF. We present evidence for significant density evolution over the observed redshift range, and we use the radio LFs to derive the dust-unbiased SFRD out to $z \sim 4.6$. Using this dust-unbiased survey, we present evidence for a significant underestimation of the SFRD based on the UV LFs at high redshift.

In **Chapter Four**, we use the COSMOS-XS survey to identify a robust sample of 'optically dark' galaxies. Using a new 'Super-deblended' FIR catalog based on the COSMOS-XS priors, we derive FIR-based photometric redshifts for the sources between $2 < z < 5$. We then quantify their contribution to the total SFRD, suggesting that they play a non–negligible role at high ($z > 3$) redshift.

In **Chapter Five**, we present new ALMA spectral scan observations with the aim of determining spectroscopic redshifts for a sub-sample of 10 radio-selected 'optically dark' galaxies. We find CO-based redshift solutions that confirm that the 'optically dark' sources targeted lie at $z \gtrsim 3$, with a median redshift of $z = 3.95 \pm 0.62$. By integrating the luminosity functions found in Chapter Four to the flux limit of the targeted sub-sample, we determine that the SFRD is in agreement with the SFRD found for the 10 ALMA detected 'optically dark' sources. This confirms the cosmic importance of these sources at high ($z > 3$) redshift.

1.7 The future

At the time of writing, the first results from the James Webb Space Telescope (*JWST*) have been published and started providing ground-breaking views of galaxy formation (e.g., Bouwens et al. 2023). *JWST* is equipped with four instruments – the Near Infrared Camera (NIRCam) and Near Infrared Spectrograph (NIRSpec), an imager and spectrograph respectively operating at $\lambda = 0.6 - 5\mu$m, the Mid-Infrared Instru-

ment (MIRI), a camera and spectrograph operating at $\lambda = 5 - 27\mu$m and the Near Infrared Imager and Slitless Spectrograph (NIRISS), a spectrograph operating at the same wavelength as NIRSpec but with lower resolution and a wider field of view. *JWST* enables sub-arcsecond resolution near-infrared sensitive observations of dust-obscured galaxies at rest-frame optical-NIR wavelengths from which the structure of dust emission can be measured out to high redshifts. Early studies already showed the power of *JWST* to study 'optically dark' galaxies, sources discussed in this thesis, and to provide a complete census of these galaxies at $z = 2 - 8$ based on rest-frame optical imaging (e.g., Barrufet et al. 2023; Pérez-González et al. 2023). The NIRcam and MIRI also allow to identify possible AGN contamination, and will thus improve the selection of SFGs needed to constrain the evolution of star formation out to high redshift ($z > 3$).

The 6-m single-dish telescope Cerro Chajnantor Atacama Telescope-prime telescope (CCAT-p) is set for first light in 2024. It operates at mm/sub-mm wavelengths (1.4mm-350μm). It will have a very wide field-of-view (a diameter of 2° at 350μm) but a limited angular resolution. One of the primary goals of CCAT-p is to trace the evolution of the dust obscured star formation in galaxies. The planned 350μm surveys covering $\sim$100 deg^2 will detect $\geq$ 600,000 galaxies of which $\geq$ 1000 are expected to be in the redshift range between $z \sim 5$ and $z \sim 8$ (Chapman et al. 2022). With these IR emission measurements, this CCAT-p survey will provide extensive measurements of the dust-obscured star formation activity at an epoch in which the amount of dust-obscuration in SFGs is still debated as discussed in this thesis.

Radio interferometry will be revolutionized thanks to the Square-Kilometre Array (SKA). This interferometer will observe between 50 MHz and 24 GHz and will extend to flux density limits more than three orders of magnitude deeper than currently possible (De Zotti et al. 2019). The construction of the SKA started in December 2022 and the deployment will be phased. The first phase (SKA-mid) will consist of 197 dishes operating at 350 MHz-15.3 GHz (with the aim of being expanded to 20 GHz) in South Africa with maximum baselines of 156km. These baselines provide sub-arcsecond resolutions with fields of view with diameters of $\sim$10–100 arcmin. The lower frequency companion (SKA-low) will operate in Australia at 50–350 MHz and will consist of $\sim$131,000 antennae with maximum baselines of 65km. The basic increase in depths obtainable with this array will be an opening up of direct detections of MS galaxies at higher redshift than currently possible with the JVLA as discussed in this thesis (e.g., De Zotti et al. 2019). A survey like the COSMOS-XS survey could be undertaken with the SKA within a fraction of the observing time used for the COSMOS-XS survey.

The Low Frequency Array (LOFAR; van Haarlem et al. 2013) is a SKA pathfinder interferometer working at low (< 1 GHz) and ultra-low (< 100 MHz) radio frequencies. It is centered in the Netherlands, with antennas around Europe, and took its first observations of the radio sky in 2010. LOFAR consists of two sets of antennas in each station, the Low Band Antennas (LBA, 10-90 MHz) and the High Band Antennas (HBA, 110-240 MHz). It is able to map 10 deg^2 regions of the sky in a single pointing with high resolutions. Prospects for future studies of radio-selected SFGs with LOFAR are bright. For example, the LOFAR Deep Fields survey, reaching sensitivities of ~ 13 μJy beam^{-1}, will build up a large sample sample of SFGs enabling the study of dust-obscured star formation as described in this thesis. In addition, LOFAR is currently updated to LOFAR 2.0. This is a staged expansion of the technical capabilities of LOFAR and will ensure that LOFAR remains scientifically relevant throughout the decade. The upgrade includes improved electronics, better designed LBA dipoles, ~ 6 new stations in the Netherlands and new international stations. LOFAR 2.0 will also allow for simultaneous observations with high-band and low-band antennas enabling to combine information from both antennas (Edler et al. 2021).