Distant star formation in the faint radio sky

Algera, H.S.B.

1 Introduction

Nothing became something
Turned itself inside out
Subatomic particles became atoms, became stars, became galaxies
– Jens Lekman on How We Met, The Long Version (2017)

1.1 The Beginning

At the time of writing, we live on planet Earth, the third planet around a relatively typical star we call the Sun. The Sun, like a few hundred billion other stars, resides in what we call a galaxy. In fact, this particular galaxy is so important to us that it is sometimes referred to as *the* Galaxy, although more typically it is known as the Milky Way. It may come as a surprise that until just about a century ago, it was still not known whether the Milky Way was the only galaxy in the Universe. It took new observations from American astronomer Edwin Hubble to conclusively show that the so-called "spiral nebulae" that had been observed over the years were, in fact, galaxies themselves, existing far beyond the outskirts of the Milky Way (Hubble 1926).

Much has happened since within astronomy. In the late 1920s, astronomers Georges Lemaître and Hubble[1] independently observed that distant galaxies all appear to move away from us, at a speed directly proportional to their distance (Lemaître 1927; Hubble 1929). From this observation, it follows that the Universe must be expanding, and in turn must have been smaller in the past. Indeed, the Universe can be traced back in time to a single point – known as the *Big Bang* – which has been reliably calculated to have occurred some 13.8 billion years ago (Planck Collaboration et al. 2020). Our current physical models are unable to fully explain nor

[1]Yes, the same Edwin Hubble – it turns out that breakthrough findings like these are a surefire way to get a space telescope named after you.

describe this beginning of the Universe, but begin to apply within the first second following the Big Bang. The particles that we know, and also the particles that we don't know, were formed.

At this point in time, the Universe was a hot soup consisting of ionized hydrogen and helium, and copious free electrons – the basic building blocks of all the visible matter around us. We have no direct observations of this early epoch, as the Universe remained far too dense for light to travel freely. Instead, all the light within the Universe would be constantly scattered around by the free electrons. As the Universe expanded, however, it gradually cooled down while its density dropped. The free electrons and ions were finally able to combine into neutral particles, and light could now freely traverse the Universe. Some of this light was able to make its way to Earth unimpeded, and is directly observable with our telescopes.

This *primordial* light provides the very first picture of the infant Universe, and is shown in Figure 1.1. It portrays what is known as the *Cosmic Microwave Background* (CMB), which is the low-energy radiation that permeates the entire Cosmos. Because the Universe has expanded substantially since the emission of the CMB, the wavelength of its photons has expanded with it. In turn, the CMB is much less energetic now than it was in the past, and radiates at a temperature of less than three degrees above absolute zero (Bennett et al. 2003). The CMB is therefore not visible with the human eye, but can instead be observed using dedicated radio telescopes.

One crucial observation about the CMB is that it is not perfectly smooth. Mind you, it is still really smooth, with any variations being on the order of just 0.01%, but it is not *perfectly* smooth. This has monumental consequences. Some regions are ever so slightly denser, and have ever so slightly more gravity than others. Over the course of a few hundred million years (e.g., Bromm & Yoshida 2011), more and more material will therefore slowly be pulled into these regions due to gravity. Here, the first galaxies form.

1.2 The Evolution of Galaxies

A *galaxy* is the collection of many millions or billions of stars amidst a tenuous reservoir of cosmic gas, dust and a mysterious substance known as dark matter, all of which is holding itself together via gravity. While a satisfactory definition in an astronomical sense, it is somewhat abstract. As a picture is known to speak a thousand words, some examples of galaxies

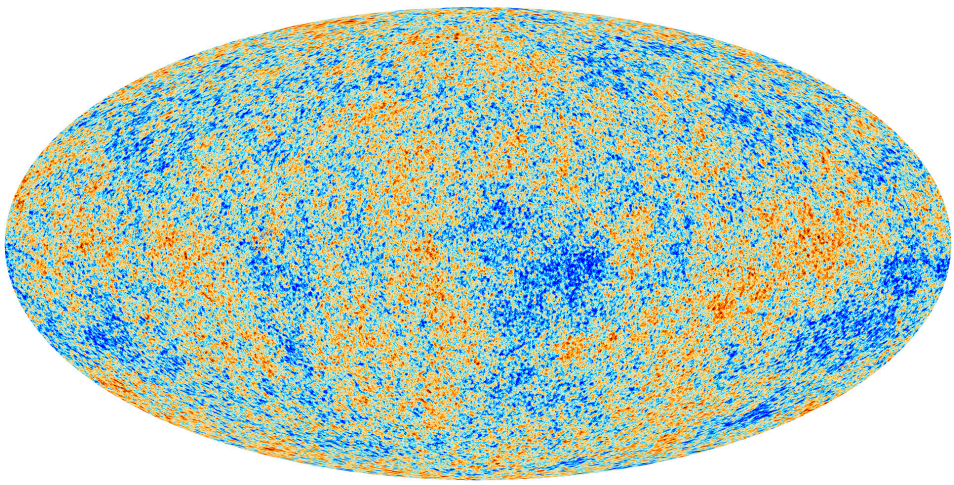
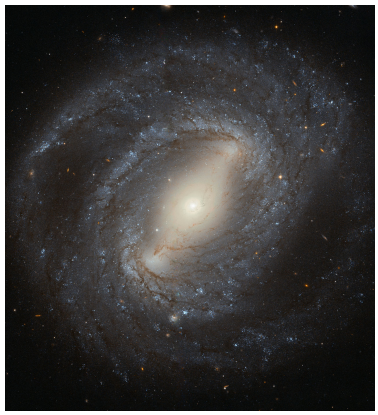


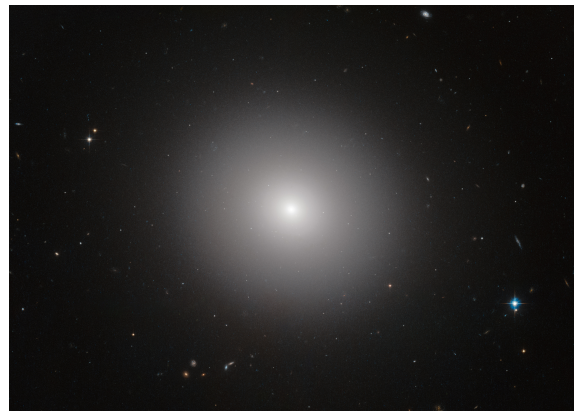
Figure 1.1: The cosmic microwave background, as observed by the *Planck* satellite. The CMB represents relic radiation from the Big Bang, and constitutes the earliest observable light from the Universe. The CMB is highly uniform, and the colour scale has been enhanced to show small deviations from this uniformity. Slightly overdense regions in the CMB give rise to future sites of galaxy formation. Credit: ESA and the Planck Collaboration.

are shown in Figure 1.2. Galaxies, clearly, come in a wide variety of shapes and sizes. Some galaxies, our Milky Way included, are flat spiral disks that appear blue when observed in visible light. Others appear as large, elliptical collections of stars with a relatively red colour. Even others are enshrouded in large amounts of dust, and are very difficult to detect at the wavelengths visible to the human eye. One of the main objectives in astronomy is to explain the large variety of galaxies that we observe, and to understand how galaxies may evolve from one type into another.

Understanding the evolution of galaxies, however, is complicated by one simple fact: a galaxy evolves very, very slowly. Individual stars that make up a galaxy, for example, have lifetimes from a few million years to many tens of billions of years – the latter being much longer than the current age of the Universe. In addition, our Sun is currently traveling through the Milky Way at a velocity of roughly 220 kilometers per second (Gunn et al. 1979). At this rate, it takes roughly 220 million years for the Sun to complete a single orbit through the Galaxy. All these timescales, from the evolution of individual stars to the motions of stars within galaxies, are much larger than a human lifespan. As a result, we cannot simply observe a galaxy for a few years, or even for a few decades or centuries, and infer how it evolves. Instead, we need to observe many galaxies at different



(a) Barred spiral galaxy NGC 4394.



(b) Elliptical galaxy IC 2006.

(c) Irregular galaxy IC 3583.

(d) The two colliding Antennae Galaxies.

Figure 1.2: Several examples of nearby galaxies, highlighting the diversity of the galaxy population. Our galaxy, the Milky Way, is a spiral galaxy likely not too dissimilar from the barred spiral shown in panel **a)**. Spiral galaxies such as the example shown here generally form stars within their spiral arms, which emit blue light. Elliptical galaxies, such as the example in panel **b)**, on the other hand, typically show very little star formation, and appear redder. Panel **c)** shows an irregular galaxy, with no clear discernible structure, while panel **d)** shows the chaotic collision of what previously were two normal spiral galaxies. Collisions such as these are often paired with a surge in star formation within the merging galaxy, as evidenced by the abundant blue star-forming regions within the Antennae (Credit: ESA/Hubble & NASA).

stages in their lives – young ones or old ones – and compare the way they look. All these individual galaxies form the puzzle pieces that have to be combined into a full theory of galaxy evolution.

Owing to the finite speed of light, finding the youngest galaxies in the Universe is equivalent to finding the most distant ones. Measuring accurate distances is crucial in astronomy, as our telescopes namely only provide information on the *apparent* brightness of the observed galaxy population. Only when distances are known, can their *intrinsic* brightness be inferred. In the field of galaxy evolution, distances are customarily expressed via the parameter *redshift*, notated as the letter z. As the Universe is constantly expanding, so is the wavelength of light emitted by distant galaxies. Since longer wavelengths appear redder to the human eye, a galaxy at a higher redshift is more distant. In particular, redshift is directly related to the size of the Universe: at a redshift $z = 1$, the Universe was half its current size, and the wavelength of light has in turn since expanded by a factor of two. At a redshift of $z = 2$, the Universe was a third its current size, while wavelengths emitted at this time are now observed to be three times larger, and so forth. We will later see that galaxies emit light across the full *electromagnetic spectrum*, including X-rays, ultraviolet and infrared light, and radio waves. As the light from more distant galaxies becomes increasingly redshifted, observations in, for example, the infrared part of the spectrum may in fact probe the visible light emitted by distant galaxies, or even their ultraviolet light in even more distant ones.

Since the size of the Universe increases with time as it expands, redshift is also often used to express the age of the Universe. At the present time, we define the redshift to be $z = 0$, while the Universe is roughly 13.8 billion years old. On this scale, the Big Bang, marking the beginning of the Universe, occurred at a redshift of infinity. The CMB was emitted around $z \approx 1100$ (Bennett et al. 2003), and the first stars and galaxies already formed a few hundred million years after the Big Bang, around a redshift of $z \sim 15 - 20$ (e.g., Bromm & Loeb 2004; Hashimoto et al. 2018). The formation of stars and the subsequent build-up of galaxies form the topic of the next Section.

1.2.1 Star Formation in Galaxies

When observed from a dark location on Earth, the Milky Way appears as a large, white band of stars stretched across the night sky. Stars are, indeed, the most readily visible constituents of galaxies, and understanding how

galaxies evolve therefore requires some understanding of the evolution of the individual stars that comprise them.

Stars form when clouds of dense molecular gas collapse under their own gravity. This process, like many in astronomy, takes a long time – on the order of a few million years (Krumholz & Tan 2007) – and as a result it is not possible to watch star formation happen in real time. In addition, stars generally do not form alone. Instead, massive nebulae of gas, such as the famous and well-studied Orion nebula, are large stellar nurseries forming a multitude of stars over the course of a few tens of millions of years (e.g., Murray 2011). Such a collection of recently formed stars is known as a *simple stellar population*, or SSP for short. The "simple" in this name stems from the fact that the stars are assumed to be born at roughly the same time, and also from the same gas, meaning they have the same initial composition (e.g., Bruzual 2010).

The notion of an SSP is very useful in extragalactic astronomy, because in distant galaxies we cannot observe single stars even with powerful telescopes. Instead, what is typically observed is the combined light of all the stars within many individual SSPs. It is therefore illuminating to first examine an SSP in its purest form – without any intervening dust or other complexities – in some detail. The light emitted by an SSP is most readily characterized through its *spectral energy distribution*, which is a measure of the total energy emitted as a function of wavelength. An example spectral energy distribution of an SSP at various ages is shown in Figure 1.3. The youngest SSP shown has an age of 10 million years, while the oldest one has an age of 10 billion years. The difference between these SSPs is stark, in particular at short wavelengths (indicated by the Greek letter λ). At a wavelength of 0.1 μm, which lies in the ultraviolet part of the spectrum, the young SSP is nearly a million times brighter than its aged counterpart. However, at longer wavelengths, beyond roughly one micron, the difference is reduced to less than a factor of one hundred.

This difference in energy emitted between young and old SSPs is directly related to the types of stars that comprise these populations, as the lifetime of any star is strongly dependent on its mass. Massive stars use up their available fuel much more quickly than their low-mass counterparts, and as such live much shorter lives. For example, a star of 30 solar masses has a typical lifetime of only 10 million years. However, a star such as the Sun, which by definition consists of one solar mass, lives roughly 10 billion years, or one thousand times as long. This means that, if one observes a young SSP, the most massive stars are still alive. However, older SSPs will

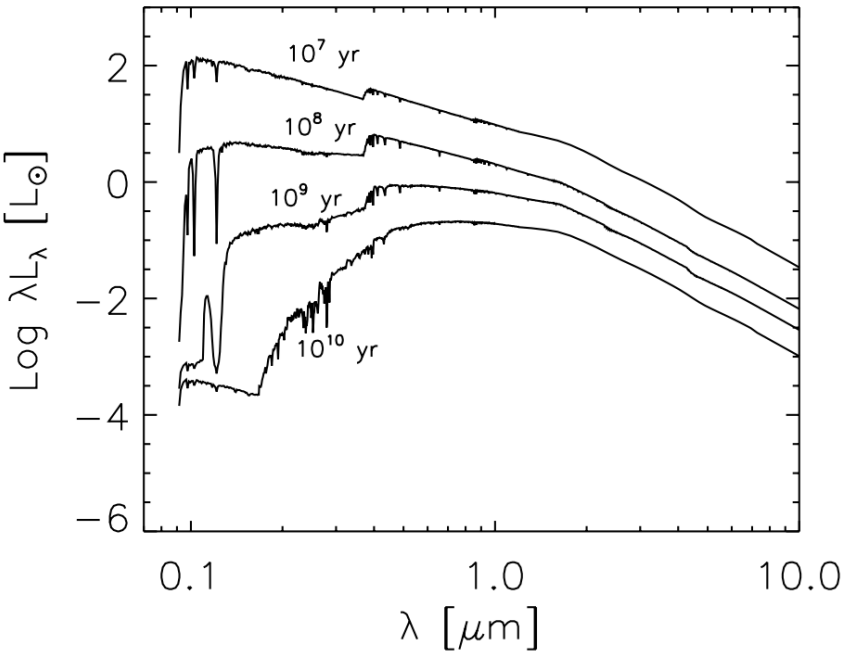


Figure 1.3: Spectral energy distributions of four simple stellar populations of various ages. The vertical axis shows the energy output in solar luminosities, as a function of wavelength in micron. Younger SSPs emit significantly more ultraviolet radiation than their aged counterparts. Figure taken from Takagi et al. (2003).

only have low-mass stars left – the massive ones will have ended their lives in powerful *supernova* explosions.

Because massive stars burn through their fuel very quickly, they are also significantly hotter than low-mass stars such as the Sun. Therefore, massive stars emit light at more energetic wavelengths, such as in the ultraviolet part of the spectrum, whereas low-mass stars are mostly limited to the less energetic optical and infrared wavelength regimes.[2] Young SSPs will therefore emit copious amounts of energetic ultraviolet radiation, whereas older ones will emit much smaller amounts. This fact can be inverted: if one measures a lot of ultraviolet radiation coming from a galaxy, this

[2]The wavelength where a star emits most of its energy is inversely proportional to the temperature of a star – a fact known as Wien's displacement law. However, all stars emit radiation across a broad wavelength range. This is part of the reason why, even for a not particularly massive star like the Sun, wearing sunscreen is recommended.

must mean that it harbors a young stellar population.[3] In fact, the precise amount of ultraviolet radiation holds valuable clues on how many young, luminous stars exist. In turn, ultraviolet observations can measure a galaxy's recent *star formation rate*.

1.2.2 Star Formation Rates

The star formation rate (SFR) of a galaxy is a fundamental quantity describing its evolution. Expressed in $M_\odot\,\mathrm{yr}^{-1}$, the SFR describes how quickly a galaxy is forming stars over time. The Milky Way, for example, has an SFR of roughly $1 - 2\,M_\odot\,\mathrm{yr}^{-1}$ (Licquia & Newman 2015). This number should not be directly interpreted as one or two Sun-like stars forming in the Milky Way each year. As previously mentioned, the formation of a single star is a lengthy process, and instead the SFR is a measure of the *average* yearly mass formed in stars within some recent history. We will see later on that this "recent history" can be anywhere from a few million years to more than a hundred million years for the various different ways astronomers use to measure star formation rates.

One important topic that we have until now neglected is the actual *composition* of a simple stellar population. If we can only measure the combined light of an SSP in distant galaxies, can we actually say something about how many stars have formed, and how massive these stars are? In the distant Universe, the answer is actually most generally "no". Instead, astronomers use observations of nearby SSPs, where it is still possible to observe and count individual stars, and adopt these results for galaxies in the early Universe. The number of stars that form at a given stellar mass is a distribution known as the *initial mass function*, or IMF (e.g., Salpeter 1955). A variety of determinations for the IMF exist in the astronomical literature, though they all share a common feat: low-mass stars are abundant, while massive stars are rare. This fact has an interesting consequence. Simple stellar populations are most readily identified by the high-mass stars they produce, as these are short lived and produce strong ultraviolet radiation, while most of the actual stellar mass that has formed is hidden in low-mass stars.

Given an IMF, and a model for the light produced by an SSP, the total observed light at ultraviolet wavelengths L_{UV} can directly be used to in-

[3] The critical reader will note that this (incorrectly) assumes no other mechanisms are able to generate significant ultraviolet radiation in galaxies. We discuss this in Section 1.5.

fer a galaxy's star formation rate ($\mathrm{SFR_{UV}}$). We will not concern ourselves here with the explicit conversion factor between these quantities, but simply state they are linearly related (e.g., Kennicutt 1998):

$$\mathrm{SFR_{UV}} \propto L_{\mathrm{UV}} \,. \tag{1.1}$$

Some of the most powerful telescopes, such as the *Hubble Space Telescope*, are able to observe at ultraviolet wavelengths, and have been used to observe UV radiation from galaxies within the first billion years of the Universe (e.g., Bouwens et al. 2016).

One important caveat, however, affects the accuracy with which ultraviolet radiation can be used as a tracer of star formation in galaxies. A galaxy, namely, is more than just a collection of simple stellar populations. In addition to stars, galaxies contain potentially large amounts of gas and dust within them, and any radiation emanating from an SSP may encounter these contents along its way to our telescopes.[4] Dust, in particular, has a significant impact on ultraviolet and optical light. Upon absorbing energetic radiation from stars, dust will be heated and subsequently radiate its energy at longer wavelengths, generally in the infrared and (sub)millimeter parts of the spectrum.

Dust itself is, in fact, a byproduct of star formation. It is predominantly produced by supernovae and a particular late evolutionary stage of low to intermediate mass stars known as the *AGB phase* (Michałowski 2015). Since the most massive stars will explode as supernovae in only a few million years, dust buildup in galaxies can happen quite rapidly, and dust has been shown to exist in some of the earliest galaxies that have formed within a billion years after the Big Bang (Bouwens et al. 2021; Schouws et al. 2021).

Since dust is capable of attenuating ultraviolet light, star formation rates inferred from ultraviolet emission may be strongly underestimated in dusty galaxies. In this case, we refer to the missing star-formation as being *dust obscured*. Various studies (e.g., Whitaker et al. 2017) have indicated that most galaxies with star formation rates in excess of $\mathrm{SFR} \gtrsim 10\,M_\odot\,\mathrm{yr}^{-1}$ have over half of their star formation obscured by dust. For a complete picture of galaxy formation, it is therefore crucial to understand the prevalence and effects of dust in galaxies.

[4] When astronomers talk about "dust", they refer to mostly carbon- and silicon-based pebbles with a typical size of, roughly, a nanometer to a micron (Draine 2003).

The shift of optical and ultraviolet light to longer wavelengths due to dust is occasionally referred to as *dust reddening*. The presence of dust in galaxies can therefore be inferred observationally as a deficit in emission at short ("blue") wavelengths, or – typically more readily – by observing significant emission at "red" wavelengths, including the infrared and submillimeter regimes. Owing to the brightness of newly formed massive stars, galaxies that are actively star-forming will have most of their dust heated by these young stellar populations. As a result, the observed total dust emission from galaxies is a direct measure of the attenuated ultraviolet luminosity of young stars, and hence measures a galaxy's (dust-obscured) star formation rate. As for ultraviolet SFRs (Equation 1.1), star formation rate and infrared luminosity $L_{\rm IR}$ are directly related through a linear calibration (Kennicutt 1998):

$$\mathrm{SFR}_{\rm IR} \propto L_{\rm IR} \tag{1.2}$$

If a galaxy contains both obscured and unobscured star formation, its total star formation rate is given by $\mathrm{SFR} = \mathrm{SFR}_{\rm IR} + \mathrm{SFR}_{\rm UV}$, that is, by summing the infrared and ultraviolet contributions. As we will discuss in the next Section, dust-obscured star formation forms a crucial part of any full theory of galaxy evolution.

1.2.3 Dusty Star-forming Galaxies

When the *Infrared Astronomical Satellite* (IRAS) was launched in 1983, and performed an all-sky survey at infrared wavelengths, it quickly made a surprising discovery (e.g., Soifer et al. 1987). The satellite, operating at wavelengths between $12-100\,\mu$m (Neugebauer et al. 1984), observed thousands of bright galaxies that had not previously been seen in optical surveys. These galaxies were among the most actively star-forming hitherto observed, but were so dust-obscured that nearly all of their stellar light was attenuated and re-emitted at infrared wavelengths. While certainly an interesting population, the bright dust-obscured galaxies observed by IRAS – designated (Ultra) Luminous Infrared Galaxies; (U)LIRGs for short – make up only a small fraction of the local galaxy population, and only a few percent of the total amount of star formation happening nearby (Sanders & Mirabel 1996).

However, around the turn of the 21st century, infrared telescopes had become powerful enough that more distant galaxies could be probed. Using

the *James Clerk Maxwell Telescope*, operating at a wavelength of $850\,\mu$m, Smail et al. (1997) presented the first evidence that bright dust-obscured starburst galaxies were more prevalent in the early Universe, which was quickly confirmed by subsequent studies (see Blain et al. 2002 for a review). Recent estimates indicate that more than half of all star formation from redshift $z \sim 4$ to the present day is obscured by dust (e.g., Zavala et al. 2021), with a significant fraction due to bright dust-obscured starbursts. Owing to their discovery at $850\,\mu$m – a much longer wavelength than probed with IRAS – these distant starbursts were dubbed *submillimeter galaxies* (SMGs). Many thousands of SMGs have been observed to date, with the brightest among them reaching star formation rates in excess of $1000\,M_\odot\,\mathrm{yr}^{-1}$. The physical conditions in these galaxies that lead to such extravagant star formation rates remain a topic of much research in astronomy (see e.g., Hodge & da Cunha 2020 for a review).

1.2.4 The Star-forming Population

At this point, it is illuminating to take a step back, and place these dusty galaxies in a broader context. In fact, a simple yet revealing illustration of the overall galaxy population requires only two parameters, one of which we have previously introduced: star formation rate and *stellar mass*. These two quantities are naturally related: while the star formation rate is a measure of a galaxy's current evolution, the stellar mass represents its past evolution. Indeed, stellar mass may simply be understood as the integral of a galaxy's star formation rate across time.[5]

Figure 1.4 shows a sample of galaxies in the nearby Universe on the star formation rate versus stellar mass plane. The galaxy population roughly divides into three different subsets: the bulk of the galaxies fall onto a (nearly) linear relation between star formation rate and stellar mass, known as the *star formation main sequence* (e.g., Brinchmann et al. 2004; Noeske et al. 2007). The main sequence constitutes a tight correlation with a scatter of roughly $0.3\,$dex, and has been shown to already be in place at high redshift (at least $z \sim 4-5$; Speagle et al. 2014; Schreiber et al. 2015). In addition, the normalization of the main sequence increases with increasing redshift, implying that – at fixed stellar mass – typical galaxies were more strongly star-forming in the past.

[5]This is somewhat of an oversimplification, as a fraction of the mass formed during star formation is later returned to the interstellar medium. As such, the true stellar mass of a galaxy is different from the total mass formed in stars (e.g., Bruzual & Charlot 2003).

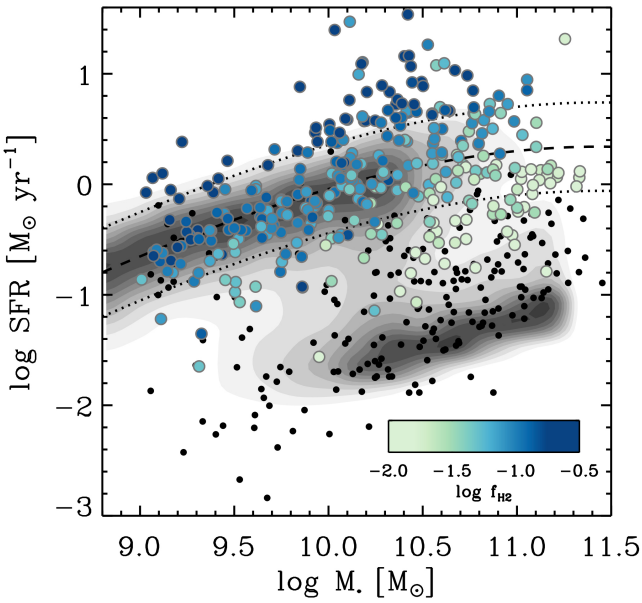


Figure 1.4: Galaxies (from the xCOLD GASS survey) on the star formation rate versus stellar mass plane. At fixed stellar mass, galaxies can have a variety of star-formation rates: most galaxies lie on the star formation main sequence, where – at least at low and intermediate masses – stellar mass and SFR are roughly linearly related. A small fraction of galaxies has star formation rates in strong excess with respect to the main sequence; these form the starburst population. In contrast, a fraction of predominantly massive galaxies shows very little star formation, and makes up the quiescent population. Galaxies in this figure are colored by their molecular gas contents, which constitutes the fuel for star formation. High gas fractions typically correspond to higher star formation rates. (Figure credit Amélie Saintonge, based on data from Saintonge et al. (2016); `http://www.star.ucl.ac.uk/~amelie/research.html`).

Not all galaxies, however, lie on the main sequence. A fraction of galaxies contains very little ongoing star formation, despite being relatively massive. This population is known as the *quiescent population*. Quiescent galaxies must have had significant star formation in the past in order to assemble their current stellar mass, though their star formation has since mostly ceased by the as of yet not fully understood mechanism of *quenching*. Quenching, most likely, involves either expelling or heating of gas within galaxies, which in turn inhibits future star formation (e.g., Kereš et al. 2005; Croton et al. 2006). In order to shed further light on the precise mechanisms and timescales involved in quenching, significant obser-

vational efforts are focused on detecting quiescent galaxies in the early Universe, where their number densities were significantly smaller compared to the present day (e.g., Ilbert et al. 2013).

In contrast to the quenched population, some galaxies have significantly higher star formation rates than would be expected based on the main sequence; this highly star-forming population constitutes what are known as *starburst galaxies*. Starburst activity may be triggered via galaxy mergers, as is commonly observed for local starbursts (Sanders & Mirabel 1996), or alternatively by an increase in their cold gas contents, or an increased efficiency of star formation from the available gas (Santini et al. 2014; Elbaz et al. 2018; Liu et al. 2019). Starburst galaxies overlap in part with the submillimeter galaxies introduced earlier, as high star formation is accompanied by significant dust build-up, although some SMGs may fall onto the upper end of the main sequence and may not formally be classified as starbursts (e.g., Michałowski et al. 2012; Birkin et al. 2021).[6] Nearly all massive galaxies are thought to go through one or multiple starburst phases (e.g., Dudzevičiūtė et al. 2020), and, similarly, less massive galaxies may also often rise above or dip below the main sequence (Sparre et al. 2015; Matthee & Schaye 2019) As such, while the star-forming main sequence may be regarded as an average trajectory for galaxy evolution, it should not be regarded as an evolutionary sequence for any individual galaxy.

Having established an inventory of the full galaxy population, the next few Sections will focus on their multi-wavelength emission, beginning with the longest wavelengths.

1.3 Radio Emission in Galaxies

A large fraction of this thesis concerns itself with a wavelength regime we have yet to explore: the radio window. Radio waves from an astronomical origin were originally discovered in 1931 by American engineer and physicist Karl Jansky, using a radio telescope he himself constructed. The legacy of Jansky's work remains reflected in the unit of flux density that is ubiquitously used in radio astronomy, the *Jansky*, defined as $1\,\mathrm{Jy} = 10^{-26}\,\mathrm{W\,m^{-2}\,Hz^{-1}}$. In addition, the famous collection of linked radio telescopes near Socorro, New Mexico in the United States was renamed the *Karl G. Jansky* Very Large Array in his honor in 2012.

[6] It must be noted that stellar masses tend to be quite uncertain for dust-obscured populations such as SMGs, see for example Michałowski et al. (2012).

Radio emission in galaxies can be produced in a variety of ways. Here, continuing our discussion from the previous sections, we will focus on radio emission originating from star formation, while we discuss other origins in Section 1.5.

As we have established in Section 1.2.2, measuring star formation rates in distant galaxies typically revolves around probing the emission originating from young stars. While some stars do emit radio waves, and in fact the Sun is the brightest object in the radio sky due to its proximity to us, direct stellar radio emission is not used as a tracer of star formation. Most stars are, indeed, radio-faint. Using radio observations from the Very Large Array across more than half of the entire sky, supplemented by similar all-sky observations at optical wavelengths, Kimball et al. (2009) did not conclusively detect any radio emission from a stellar origin. Instead, the radio emission produced by star-forming galaxies stems from two main sources: supernovae shocks and ionized plasma. The radio waves emitted by these sources are produced by, respectively, *synchrotron emission* and *free-free emission*.

1.3.1 Synchrotron Emission

Synchrotron radiation is produced by the acceleration of charged particles in a magnetic field. The exact origin of magnetic fields in galaxies remains debated (e.g., Beck 2012), and it is unclear whether they are generated in the early Universe, or instead arise later through astrophysical phenomena. Nevertheless, magnetic fields are ubiquitous in galaxies, and have been directly observed in nearby galaxies through polarized emission at various wavelengths, as well as Zeeman splitting (Robishaw et al. 2008). Magnetic fields further appear to be particularly strong in local starburst galaxies (McBride et al. 2014, 2015), where they may attain strengths of several milliGauss (mG). More typical spiral galaxies have magnetic field strengths that are generally on the order of $10\,\mu$G (Beck 2012; Tabatabaei et al. 2017).

The charged particles generating synchrotron emission are created in the shock waves that accompany the violent death of massive stars ($\gtrsim 8\,M_\odot$) in supernova explosions. 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 ($\lesssim 30\,$Myr; Bressan et al. 2002). However, from a theoretical point of view, it is far from trivial to use radio observations as a means of determining star formation rates, as this process involves

understanding many key ingredients, such as the physics of supernovae, the acceleration of cosmic rays in their remnants, the propagation of such cosmic rays through the host galaxy, and the magnetic field in which they radiate their energy (e.g., Condon 1992).

However, nature lends her helping hand. Early observations by Van der Kruit (1971, 1973) already indicated the existence of a correlation between the infrared and radio luminosities of galaxies, a result which was soon extended to larger galaxy samples (De Jong et al. 1985; Helou et al. 1985). Since then, this so-called *far-infrared/radio correlation* has been shown to hold across many orders of magnitude in galaxy luminosity, from small dwarf galaxies to bright starbursting systems (e.g., Yun et al. 2001; Bell 2003). In turn, this surprisingly robust and tight correlation undoubtedly masks a plethora of underlying physical phenomena, which we explore in the following Section.

1.3.2 The Far-infrared/radio Correlation

The Theory

For the far-infrared/radio correlation to exist, naively these two statements must hold: 1) a large and roughly fixed fraction of star formation within galaxies is dust-obscured, and 2) the supernova-induced cosmic rays radiate a roughly fixed fraction of their energy through synchrotron emission (e.g., Völk 1989; Lisenfeld et al. 1996). This second condition, in turn, encompasses two related effects: cosmic rays should not escape the galaxy prior to losing the bulk of their energy, and the magnetic fields in galaxies should be sufficient that synchrotron emission contributes significantly to cosmic ray energy loss. As we will outline below, these conditions are expected to be fulfilled in relatively massive galaxies, which form the main focus of this thesis. In low-mass galaxies, however, these conditions are not necessarily expected to hold (e.g., Murphy 2009a; Lacki et al. 2010). A significant fraction of star-formation in low-mass systems may not be enshrouded in dust (Whitaker et al. 2017), resulting in infrared emission not being a robust star formation rate indicator in these galaxies. However, in what Bell (2003) dub "a conspiracy", radio synchrotron emission similarly fails to account for all star formation in low-mass galaxies, by virtue of cosmic rays being able to escape these typically smaller systems prior to losing all their initial energy. In turn, the far-infrared/radio correlation appears to be preserved even in faint, low-mass galaxies (Bell 2003; Delvecchio et al. 2021). In what follows, however, we will focus mostly

on massive galaxies, as these constitute the general population detected in high-redshift radio surveys (Section 1.6).

As a start, let us examine the nature of synchrotron emission in galaxies in some detail. The synchrotron *spectrum* of a typical star-forming galaxy, defined as its flux density as a function of frequency, is often well-characterized by a single power-law that may be written as $S_\nu \propto \nu^{\alpha_{\rm NT}}$. Observations indicate that $\alpha_{\rm NT} \approx -0.85$ (e.g., Niklas et al. 1997; Murphy et al. 2012; Tabatabaei et al. 2017), for a typical cosmic ray energy distribution (e.g., Lacki et al. 2010). In addition, this value appears to be typical for both normal star-forming galaxies and high-redshift starbursts, albeit with substantial scatter (e.g., Ibar et al. 2009; Thomson et al. 2014; Smolčić et al. 2017b; Gim et al. 2019).[7]

However, it was quickly recognized that the strong magnetic fields in starburst galaxies imply very short *cooling times* for cosmic rays, given that the timescale for synchrotron emission $\tau_{\rm syn}$ depends strongly on the magnetic field strength ($\tau_{\rm syn} \propto B^{-3/2}$; Thompson et al. 2006). Cooled synchrotron emission is expected to produce a spectral slope steeper than the canonical $\alpha_{\rm NT} \approx -0.85$, potentially reaching values of $\alpha_{\rm NT} \approx -1.3$ (Kardashev 1962). However, such steep spectra are not typically observed in star-forming galaxies at GHz frequencies.

As a result, it became apparent that the physics underpinning the far-infrared/radio correlation may be more complex than previously thought. (Thompson et al. 2006) proposed that, in dense starburst environments, other effects were at play that could flatten the synchrotron spectrum from its cooled value back to the canonical $\alpha_{\rm NT} \approx -0.85$. These two effects are *ionization cooling* and *bremsstrahlung*. Ionization cooling involves cosmic rays colliding with the ambient interstellar medium in galaxies (generally atomic or molecular hydrogen), whereby they lose energy. Such collisions are most efficient for relatively low-energy cosmic rays, whereas synchrotron cooling is most effective for their high-energy counterparts. As such, Thompson et al. (2006) argued these two processes could balance each other and in turn reconstruct what appears to be a typical uncooled synchrotron spectrum. Bremsstrahlung losses, which similarly occur from the interaction of relativistic cosmic rays with the neutral interstellar medium, have no such energy dependence, but serve as an additional process competing with synchrotron cooling.

[7] Local ULIRGs, however, may show somewhat shallower spectra, see e.g., Clemens et al. (2008) and the discussion in **Chapter three**.

With the introduction of additional mechanisms of energy loss, however, a new problem arises. If the cosmic rays that drive the far-infrared/radio correlation lose *too much* of their energy through processes other than synchrotron emission, the correlation should disappear. Lacki et al. (2010) carefully model the relative importance of synchrotron, ionization and bremsstrahlung losses, in addition to *inverse Compton cooling* involving the energy loss of cosmic rays upon their interaction with the local radiation field. They determine that, while synchrotron emission is indeed suppressed in starburst galaxies due to the various competing processes, collisions of cosmic rays with the interstellar medium should generate *secondary cosmic rays*, such as secondary electrons and positrons. These newly created charged particles will then, in turn, continue to emit synchrotron emission, thereby augmenting the synchrotron radiation produced by the primary cosmic rays. This careful balancing of energy loss mechanisms with the additional synchrotron radiation produced by secondary cosmic rays, in turn, is expected to preserve the far-infrared/radio correlation in massive starbursts.[8]

As radio telescopes were becoming capable of probing more distant galaxy populations, it became paramount to additionally assess the nature of the far-infrared/radio correlation in the high-redshift Universe. As the cosmic microwave background was more energetic in the past, inverse Compton losses on the CMB should become increasingly dominant further back in cosmic time (e.g., Murphy 2009a). Eventually, inverse Compton cooling is therefore expected to result in a breakdown of the far-infrared/radio correlation (Schleicher & Beck 2013), although the exact redshift where this breakdown should occur depends on the magnetic field strength of individual galaxies. Starburst galaxies, fortified by their strong magnetic fields, are expected to be most resilient to this breakdown (Murphy 2009a). In addition, Lacki et al. (2010) argue that their strong magnetic should further increase their synchrotron luminosities, as these will enhance the radiative losses of the less energetic, yet more numerous, cosmic ray population. In turn, starburst galaxies form powerful laboratories for studies of the far-infrared/radio correlation in the early Universe. Indeed, Lacki & Thompson (2010) argue that "puffy" starbursts – being bulkier than their local, more compact starbursting counterparts – may, in fact, show enhanced radio emission at high redshift due to their high star formation

[8] In fact, the prediction that the correlation is preserved in starbursts was *also* labeled a conspiracy by Lacki et al. (2010).

rates. We provide strong evidence for this prediction in **Chapter three**.

The Observations

While a fully satisfactory theoretical interpretation may remain somewhat elusive, the importance of the far-infrared/radio correlation in observational studies cannot be overstated. Since infrared emission from galaxies constitutes a probe of dust-obscured star formation, this correlation provides the recipe for converting radio emission into star formation rates. As a result, the precise nature of the far-infrared/radio correlation is widely explored also in the observational realm, and generally involves defining the 'q' parameter first introduced by Helou et al. (1985):

$$q_{\rm IR} = \log_{10}\left(\frac{L_{\rm IR}}{3.75\times10^{12}\,{\rm W}}\right) - \log_{10}\left(\frac{L_{1.4}}{{\rm W\,Hz^{-1}}}\right)\,. \tag{1.3}$$

Parameter $q_{\rm IR}$, as such, constitutes the logarithmic difference of the infrared and radio luminosity of a galaxy, with an appropriate normalization to turn it into a dimensionless quantity. The radio luminosity is defined at a frequency of $1.4\,$GHz, which has historically been a popular frequency for performing radio surveys. In the local Universe, $q_{\rm IR}$ attains a typical value of $q_{\rm IR} = 2.64$, with a modest scatter of roughly $0.2-0.3\,$dex (e.g., Yun et al. 2001; Bell 2003). Given a value of $q_{\rm IR}$, star formation rates may be determined from radio synchrotron emission via

$$\mathrm{SFR}_{\rm radio} \propto L_{1.4} \propto 10^{q_{\rm IR}} L_{\rm IR}\,. \tag{1.4}$$

The importance of $q_{\rm IR}$ for determining radio-based star formation rates has lead to significant interest in exploring its potential dependence on either redshift, or galaxy physical parameters from an observational perspective. Early work by Garrett (2002) in the well-studied *Hubble Deep Field North* indicated the correlation seemed to persist out to $z\sim1$, albeit for a small sample of just 20 galaxies. Ivison et al. (2010b), used observations from the *Herschel Space Telescope* to probe the correlation out to $z\sim2$, and found a slight redshift-dependence, which was corroborated by subsequent studies (e.g., Thomson et al. 2014; Magnelli et al. 2015; Delhaize et al. 2017). Most intriguingly, this apparent redshift evolution contradicted the evolution predicted by theoretical studies: instead of high-redshift galaxies becoming radio-dim due to inverse Compton losses on the cosmic microwave

background, galaxies in the early Universe appeared to become increasingly radio-bright.

One caveat that perhaps remained underappreciated in these early studies of the far-infrared/radio correlation at high redshift (though see Sargent et al. 2010), was that selection effects may be a source of bias. For example, observations at far-infrared wavelengths benefit from what is known as the *negative k-correction*: while naively distant galaxies should appear dimmer than their more close by counterparts, this effect is compensated for by virtue of probing shorter rest-frame wavelengths in these more redshifted systems, where dusty galaxies tend to be brighter (Casey et al. 2014). In turn, observations at far-infrared wavelengths (e.g., $850\,\mu$m) have a rather uniform sensitivity to star formation at high redshift ($z \sim 1.5 - 6$). Radio observations, however, suffer from a *positive k-correction*, whereby more distant galaxies become increasingly faint quickly (Section 1.3.4). If not precisely accounted for, this leads to biases whereby at high redshift only the brightest radio sources are detected. This, in turn, will induce artificial evolution in the correlation, consistent with the evolution seen in observations (Sargent et al. 2010).

In addition, radio surveys may probe emission from a non-star-forming origin. Radio emission produced by supermassive holes in the center of galaxies – the topic of Section 1.5 – may form a source of contamination, and bias studies of the far-infrared/radio correlation. Upon carefully weeding out such contaminants, Molnár et al. (2018) argue that the correlation does not evolve with redshift, at least out to $z \sim 1.5$. Furthermore, if the far-infrared/radio correlation is non-linear, this too can induce artificial redshift evolution (Basu et al. 2015; Molnár et al. 2021).

Puzzled by an apparent (lack of) redshift evolution, and armed with deep multiwavelength ancillary observations, recent studies have begun exploring alternative dependencies of the far-infrared/radio correlation. In a detailed analysis, Delvecchio et al. (2021) find that the correlation likely depends on stellar mass, while any redshift dependence is secondary (see also Smith et al. 2020). Such evolution, however, remains difficult to explain from a theoretical point of view, and is in apparent disagreement with the models from Lacki et al. (2010) and Lacki & Thompson (2010). On the other hand, some recent studies of the sizes of radio galaxies support their predictions of secondary cosmic ray generation in galaxies (e.g., Varenius et al. 2016; Thomson et al. 2019). As such, despite its importance, the far-infrared/radio correlation remains a rather nebulous manifestation of

a variety of competing physical processes – a conspiracy, if you will.

1.3.3 Free-free Emission

While the mechanism generating synchrotron radiation in galaxies has proven to be quite complex, the opposite is true for the second process generating emission at radio frequencies: *free-free emission*. Free-free emission is produced during an interaction between charged particles, most commonly between free electrons and protons. The resulting free-free luminosity, in turn, is directly proportional to the rate of such interactions, and hence depends on the square of the ionized particle density ($N_{\mathrm{FF}} \propto n_e n_p \propto n_e^2$, where n_e and n_p are, respectively, the electron and proton density). As a result, the bulk of the free-free emission in galaxies is produced in dense, ionized regions.

Upon the formation of a sufficiently massive star from an initial cloud of dense molecular gas, the stellar radiation will proceed to ionize its surroundings. The resulting dense ionized regions, known as H II regions, produce the majority of free-free emission in galaxies. The lifetime of such H II regions is short, typically on the order of a few million years, as the cloud is rapidly dispersed by the young stars it harbors (Chevance et al. 2020). As a result, free-free emission provides a measure of massive ($\gtrsim 10 - 15\,M_\odot$; Kennicutt & Evans 2012; Byler et al. 2017) star formation on a timescale much shorter than ultraviolet, infrared or radio synchrotron emission.

Similar to the aforementioned tracers, free-free emission has been calibrated as a tracer of star-formation, as given by Murphy et al. (2011):

$$\mathrm{SFR}_{\mathrm{FF}} \propto T_e^{-0.45} \nu^{0.10} L_{\mathrm{FF}}(\nu)\,. \tag{1.5}$$

Here $L_{\mathrm{FF}}(\nu)$ is the free-free luminosity at a frequency ν and T_e is the typical electron temperature of the H II regions, upon which the star formation rate weakly depends.

The spectrum of free-free emission is remarkably simple. An analytical treatment of the Coulomb interactions in an ionized plasma, such as an H II region, shows that free-free emission should give rise to flat spectrum, though some minor quantummechanical corrections are required for a complete description. Taking these into account, the spectrum of free-free emission may be written as $S_\nu \propto \nu^{\alpha_{\mathrm{FF}}}$, where $\alpha_{\mathrm{FF}} = -0.10$ is the known,

fixed spectral index (e.g., Condon 1992; Murphy et al. 2012). The combination of free-free and synchrotron emission then provides a description of the *radio spectrum* of star-forming galaxies.

1.3.4 The Radio Spectrum

Having introduced the two main ingredients that make up radio emission in star-forming galaxies, as well as their frequency dependencies, we may now examine their relative contributions to the radio spectrum. To this end, let us introduce the *thermal fraction*, which quantifies the contribution of thermal free-free emission to the total radio emission (e.g., Murphy et al. 2012):[9]

$$f_\nu^{\mathrm{th}} = \frac{S_\nu^{\mathrm{FF}}}{S_\nu^{\mathrm{FF}} + S_\nu^{\mathrm{NT}}} \,. \tag{1.6}$$

Given the different spectral indices of free-free and synchrotron emission, the thermal fraction naturally depends on frequency. Indeed, the thermal fraction is expected to increase towards higher frequencies, due to the flat spectrum of free-free emission. What remains, then, is to anchor the thermal fraction at some reference frequency ν_0, from whereon it may be scaled to other frequencies. Observations of local galaxies indicate a thermal fraction of roughly 10% at a frequency of 1.4 GHz (Condon 1992; Tabatabaei et al. 2017). The total radio spectrum may then be written as the combination of synchrotron and free-free emission, and is of the form

$$S_\nu = S_{\nu_0} \left[\left(1 - f_{\nu_0}^{\mathrm{th}}\right) \left(\frac{\nu}{\nu_0}\right)^{\alpha_{\mathrm{NT}}} + f_{\nu_0}^{\mathrm{th}} \left(\frac{\nu}{\nu_0}\right)^{-0.1} \right] . \tag{1.7}$$

A model of the radio spectrum, adopting a thermal fraction of $f_{\mathrm{th}}(1.4\,\mathrm{GHz}) = 0.10$ and $\alpha_{\mathrm{NT}} = -0.85$, is shown in Figure 1.5. As a result of the dominant synchrotron contribution at GHz frequencies, the overall spectral index (i.e., averaging free-free and synchrotron emission) of the radio spectrum is typically $\alpha \sim -0.70$ in this regime, which is a value that is commonly adopted in the literature (Condon 1992; Smolčić et al. 2017b).

[9] A brief elaboration on the subscripts used: NT stands for non-thermal, as synchrotron as synchrotron emission does not involve particles with a thermal (i.e., Maxwell-Boltzmann) distribution of velocities. Free-free emission, denoted by subscript FF, on the other hand, is a thermal process. This nomenclature will be used interchangeably throughout this thesis.

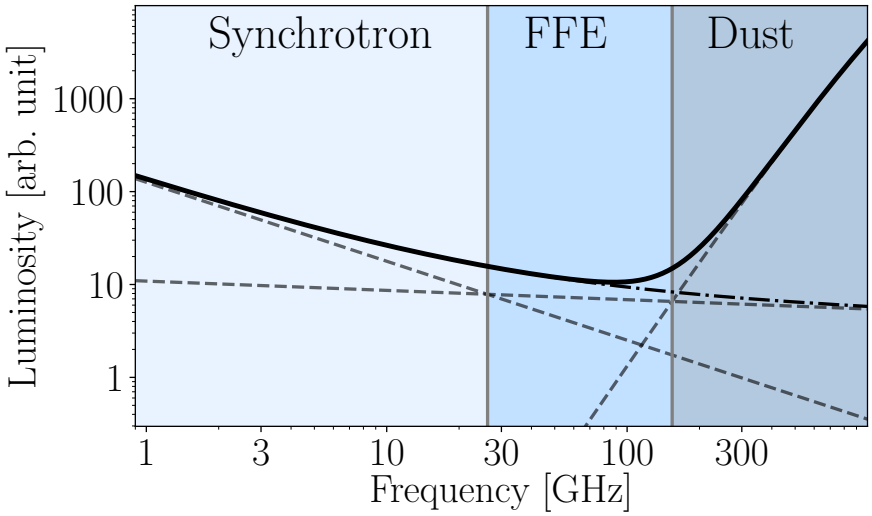


Figure 1.5: A simple model for the radio spectrum of a star-forming galaxy. At low frequencies, the steep synchrotron component is dominant, while at $\nu \gtrsim 30$ GHz, free-free emission overtakes the synchrotron contribution. At even higher frequencies, $\nu \gtrsim 200$ GHz, dust overtakes free-free emission. In this thesis, when referring to the radio spectrum, we consider the region where free-free and synchrotron dominate the emission, unless stated otherwise.

Around $\nu \sim 30$ GHz, however, the thermal fraction reaches upwards of 50%, and free-free emission takes over as the main emission mechanism. Eventually, around $\nu \gtrsim 200$ GHz, thermal emission from dust grains overtakes free-free emission as we start probing the millimeter regime. In reality, there is considerable scatter about these values, which may depend on a variety of factors, such as the precise synchrotron slope, the age of the starburst and the star-formation history, and the dust mass and temperature of the galaxy. A more detailed exploration of the radio spectrum of star-forming galaxies is provided in **Chapters 4 & 5**.

It is important to note that the model of the radio spectrum presented in Equation 1.7, combining just free-free and synchrotron emission, is a simplification in its own right. The radio spectrum typically does not extend as a power law to the lowest frequencies ($\nu \lesssim 1$ GHz), where it instead flattens or turns over due to a non-negligible optical depth (Condon 1992). In addition, older starbursts will show a deficit in synchrotron emission at high

frequencies due to spectral ageing (e.g., Thomson et al. 2019). However, more complex prescriptions of the radio spectrum naturally require more and better data, which is generally not available for distant galaxies. As a result, the radio spectrum is often approximated by either a single power law, or by the combination of two power laws describing free-free and synchrotron emission.

1.4 The Spectral Energy Distribution of Galaxies

In the previous sections, we have introduced a variety of ways galaxies produce light. At ultraviolet and optical wavelengths, we measure the direct emission from stars, while at infrared wavelengths this is predominantly starlight that has been reprocessed by cold dust. Radio wavelengths, instead, measure either supernovae through synchrotron emission, or dense ionized gas through free-free emission.[10]

A typical spectral energy distribution of a (dusty) star-forming galaxy, spanning ultraviolet to radio wavelengths, is shown in Figure 1.6. Of course, this figure is just a single example that cannot possibly encompass the variety of the full galaxy population. More dust-rich galaxies will be fainter in the optical and ultraviolet regimes, while their emission at (far-)infrared wavelengths is boosted. On the other hand, galaxies that have little ongoing star formation will predominantly shine at optical and near-infrared wavelengths, by virtue of their older stellar populations.

1.4.1 An Inventory of Star formation Rate Tracers

Having introduced a plethora of star formation rate indicators (and having omitted a plethora more), we here discuss them jointly and compare the advantages and disadvantages of each.

Ultraviolet Emission

Ultraviolet emission is produced directly by relatively young stars and traces star formation across a timescale of $\sim 10 - 200\,\mathrm{Myr}$ (Kennicutt & Evans 2012), corresponding to the lifetime of stars that produce significant UV emission. For some of the earliest known galaxies, this is the only star formation rate indicator available, owing to the sensitivity of telescopes such

[10] For now, we have conveniently left X-ray emission out of our discussion. X-rays, too, may be generated by star formation (e.g., Symeonidis et al. 2014), but are often of a non-star-forming origin – see Section 1.5 for more details.

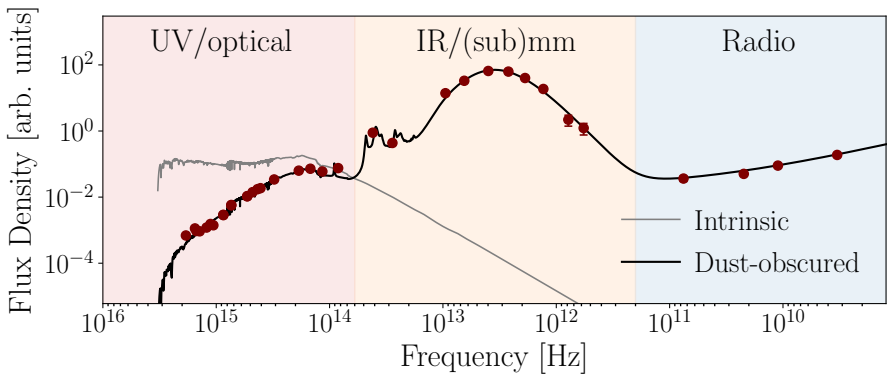


Figure 1.6: Example spectral energy distribution of a (dusty) galaxy. The rough extent of the various multiwavelength regimes in their colloquial astronomical definition are indicated. The grey line shows the intrinsic stellar spectrum, while the black line indicates the dust-reprocessed emission. The red circles represent a variety of multiwavelength measurements, and are shown to indicate the extent of the typical sampling of a galaxy's spectral energy distribution.

as the *Hubble Space Telescope*. However, in the absence of additional long wavelength information, ultraviolet star formation rates will always constitute a lower limit on the true star formation rate, as dust attenuation reduces a galaxy's UV luminosity. As infrared observations are generally not available, average dust corrections are often assumed to incorporate any contribution from dust-obscured star formation (e.g., Oesch et al. 2013), though these are inherently uncertain for individual galaxies.

In addition, while targeting the rest-frame ultraviolet emission in distant galaxies provides a powerful way of detecting and characterizing these faint systems, this emission will be redshifted to longer wavelengths due to the expansion of the Universe. In fact, the emission at wavelengths shortward of the Lyman α transition, ($\lambda < 1216\,$Å) is generally fully extincted by neutral hydrogen both within the galaxy and along the line of sight towards us (e.g., Steidel et al. 1996). In turn, optical and near-infrared telescopes are required to study the UV properties of the earliest galaxies.

Infrared emission

Infrared emission is emitted when dust in galaxies reprocesses the light from massive stars. Infrared emission, as such, traces dust-obscured star

formation across a timescale of ~ 100 Myr (Kennicutt & Evans 2012) – similar to the timescale of ultraviolet emission. Infrared emission, however, is not a direct tracer of star formation, as it does not directly probe stellar light, and relies on the presence of dust within the galaxy. From a theoretical perspective, this complicates its use as a star formation tracer, as any older stellar populations may also contribute to the heating of dust. Nevertheless, for highly star-forming galaxies, the contribution of older stellar populations to the infrared luminosity is generally negligible ($\lesssim 10\%$; Kennicutt & Evans 2012).

From a practical point of view, measuring the infrared emission in distant galaxies is complicated by the limited sensitivity and resolution of most infrared and submillimeter telescopes. The *resolution* of a telescope defines the smallest detail it can observe. Any objects smaller than this smallest observable size will be *unresolved*, and therefore blended together. The resolution with which a telescope can observe a galaxy, or any object for that matter, is strongly wavelength-dependent. It is typically measured in arcseconds, and depends on the ratio λ/D, where λ is the observing wavelength, and D the diameter of the telescope.[11] In order to obtain the same resolution, an infrared telescope must therefore be significantly larger than an ultraviolet telescope. In practice, this is generally not feasible, and as a result, most infrared-based telescopes have a relatively modest resolution, much larger than the size of individual galaxies at cosmological distances. As an example, the *Herschel Space Observatory* has a resolution spanning $5'' - 36''$ depending on the wavelength used for observing. The James Clerk Maxwell Telescope has a similarly modest resolution of $15''$ at a wavelength of $850\,\mu$m. The *Hubble Space Telescope*, on the other hand, can attain a resolution as high as $0''.05$. For comparison, the typical size of a distant galaxy is on the order of one arcsecond – the precise value of course varying with distance, galaxy type, as well as with observing wavelength (e.g., van der Wel et al. 2014; Gullberg et al. 2019; Jiménez-Andrade et al. 2019). Infrared telescopes are therefore oftentimes unable to separate closely neighboring galaxies, which remain blended together. This, in turn, makes it significantly more difficult to observe distant galaxies, and infer their dust-obscured star formation rates. In addition, infrared-based telescopes are generally limited by either the cooling of their instruments, which itself may provide thermal emission at infrared wavelengths, or by the bright, omnipresent infrared background light (e.g., Farrah et al. 2019). These factors, in turn, complicate probing the earliest galaxies at infrared

[11] An arcsecond is defined as $1/3600^{\text{th}}$ of a degree.

wavelengths. One notable exception to either of these limitations is the *Atacama Large Millimeter Array* (ALMA), which can attain both a high resolution and sensitivity. Within the last decade, ALMA has indeed revolutionized infrared-based studies of distant star formation, and has provided the first window on dust at the highest redshifts (Hodge & da Cunha 2020).

Radio Synchrotron Emission

Radio Synchrotron Emission is produced by supernova-induced shocks, and can be used as a tracer of star formation through its tight correlation with the infrared luminosities of galaxies (Section 1.3). Given the long wavelength nature of synchrotron emission, it traces star formation in a dust-unbiased manner. However, as it relies on stellar deaths, synchrotron emission is a highly indirect tracer of star formation, lagging behind other tracers by $\sim 30\,\mathrm{Myr}$ (Bressan et al. 2002). However, when star formation has ceased, it lingers for a similar timescale as ultraviolet and infrared emission – roughly $\sim 100\,\mathrm{Myr}$. As the radio and infrared luminosities of galaxies are known to correlate, there is no obvious benefit in using the less direct synchrotron emission as a star formation rate tracer from a theoretical point of view. In practice, however, radio telescopes are generally more sensitive than infrared/submillimeter telescopes, and as such they are able to probe more distant galaxies. This is due to the powerful technique of *radio interferometry*. Instead of having a single radio telescope observe the sky, a radio interferometer is a collection of telescopes all working together. Having a large number of individual telescopes increases the area across which radio waves are collected, and as such allows for the detection of fainter objects. In addition, the resolution of a radio interferometer is not dependent on the size of the individual radio dishes, but instead on the *separation* between its dishes. The Very Large Array (VLA), for example, consists of 27 radio dishes, each with a diameter of 25 meter. However, the maximum separation between the dishes can be up to 36 kilometers – far larger than the size of any single telescope!

Powerful interferometers capable of probing dust emission exist as well, such as the aforementioned ALMA operating at (sub)millimeter wavelengths. However, compared to radio telescopes such as the VLA, submillimeter interferometers have a much smaller field of view, owing to their shorter observing wavelengths. From a practical standpoint, therefore, radio synchrotron emission is a very powerful and efficient tracer of star formation in the early Universe, being capable of probing a large number of faint galax-

ies in feasible amounts of observing time.

The main drawback of using synchrotron emission to probe distant star formation, however, is the somewhat nebulous theoretical underpinning of the far-infrared/radio correlation on which it relies, as discussed in Section 1.3.2. It remains debated if, or to what extent, the correlation is modified in distant galaxies, with a variety of recent studies indicating that the correlation may evolve with redshift (Magnelli et al. 2015; Delhaize et al. 2017) or galaxy physical parameters such as stellar mass (e.g., Delvecchio et al. 2021). In **Chapter Three**, we further provide evidence that the far-infrared/radio correlation depends on the physical conditions within galaxies, by studying the correlation for a large sample of highly star-forming dusty galaxies.

Radio Free-free Emission

Radio free-free emission is emitted by the ionized bubbles created by young, massive stars. The short lifetime of these H II regions ensures free-free emission traces star formation on a near-instantaneous timescale of $\sim 3 - 10$ Myr (Kennicutt & Evans 2012). Similar to synchrotron radiation, free-free emission is a dust-unbiased tracer of star formation. Its key benefits, however, are its direct nature as a star formation rate tracer, and the fact that free-free emission can be observed with powerful radio telescopes providing high resolution observations. In theory, free-free emission is therefore one of the best star formation rate tracers available for distant galaxies.

In practice, however, free-free emission has remained mostly unexplored in the early Universe, with observations by and large having been limited to local galaxies. This is a result of the relative faintness of free-free emission compared to synchrotron emission (Figure 1.5), and the high-frequency nature of the former. High radio frequencies are more difficult to observe due to more stringent weather constraints, and also provide a significantly smaller field of view than low-frequency observations. In addition, the precise contribution of free-free emission to the measured radio flux density – the thermal fraction introduced in Section 1.3 – depends on a variety of factors, such as the star formation history of the galaxy (Bressan et al. 2002) and the non-thermal synchrotron slope. As such, detecting free-free emission generally requires deep, multi-frequency observations designed to disentangle the free-free component from the synchrotron emission which dominates the radio spectrum at low frequencies (e.g., Tabatabaei et al. 2017). In turn, observations aimed at detecting free-free emission are expensive. Nevertheless, free-free emission is set to be a key tracer of distant

star formation when improved radio facilities come online in the (near) future, warranting its exploration already with current telescopes. **Chapters four and five** present two pioneering studies of radio free-free emission in the early Universe, and discuss its potential as a star formation rate tracer in further detail.

1.4.2 The Cosmic Star Formation Rate Density

Armed with this variety of star formation rate indicators, astronomers are able to measure the *cosmic star formation rate density* (SFRD) – one of the most fundamental quantities in the field of galaxy evolution. The cosmic star formation rate density is a measure of the *average* star formation rate across all galaxies, at a given moment in cosmic history. A compilation of various ultraviolet and infrared-based measurements of the SFRD from Madau & Dickinson (2014) is shown in Figure 1.7. The ultraviolet observations have been corrected for extinction by dust, though these corrections are – particularly at high redshift – very uncertain.

In the early Universe, the SFRD rapidly rose, peaking around a redshift of $z \sim 2$ (10 Gyr ago). The SFRD subsequently entered a decline towards the present day, which is expected to continue into the future (Walter et al. 2020). Besides accurately measuring the shape of the SFRD, it is also crucial to understand why its shape is as observed. Since stars form out of clouds of molecular gas, parallel efforts are underway to measure the cosmic molecular gas density, which have determined a similar evolution across cosmic time (Decarli et al. 2019; Riechers et al. 2019).

The wealth of constraints on the SFRD compiled in Figure 1.7 may suggest that the history of cosmic star formation is, for the most part, fully understood. This, however, is not the case. At high redshift ($z \gtrsim 3$), nearly all constraints on cosmic star formation are from ultraviolet observations, which require the aforementioned uncertain dust corrections. In addition, these observations are by construction incomplete, as they miss galaxies that are (nearly) fully dust-obscured. Recent studies of the SFRD at far-infrared and radio wavelengths have hinted towards an excess in star formation at $z \gtrsim 3$ compared to the Madau & Dickinson (2014) compilation (e.g., Novak et al. 2017; Gruppioni et al. 2020), which may indicate the early Universe being significantly more dust-rich than previously thought (e.g., Casey et al. 2018). Several recent detections of large samples of dust-obscured galaxies at very high redshift ($z \gtrsim 6$) with ALMA (e.g., Bouwens et al. 2021) will provide a much improved census of dust at early cosmic epochs in the near future. In addition, there is substantial room

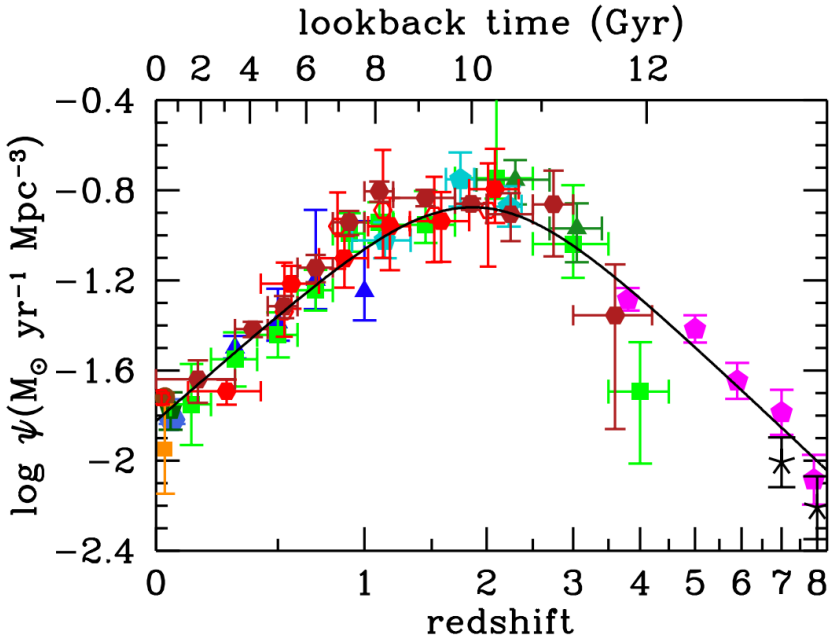


Figure 1.7: A compilation of the cosmic star formation rate density by Madau & Dickinson (2014), with the various colours representing different studies. The vertical axis shows the SFRD parameterized by the Greek letter ψ, as a function of redshift (lower x-axis) or time (upper x-axis). Starting with the formation of the first galaxies at high redshift ($z > 10$), the SFRD rises rapidly up to $z \sim 2$. Since then, the SFRD has been on the decline.

for improvement also in the way star formation rates are measured, as the canonical tracers of star formation suffer from various uncertainties (Section 1.4.1). In **Chapter five**, we therefore place the first constraints on the cosmic star formation rate density using the powerful radio free-free emission as a tracer.

1.5 Active Galaxy Nuclei

In the previous sections, we have discussed how star formation imprints its signature on the spectral energy distribution of galaxies. However, in a fraction of galaxies this emission of star-forming origin is outshined – potentially by orders of magnitude – by an *active galactic nucleus* (AGN).

In the early 1960s, Dutch astronomer Maarten Schmidt took an optical spectrum of a star that appeared to coincide with the bright radio source 3C 273 that had been observed a few years prior with an array of radio telescopes located near Cambridge, England. In a surprising find, Schmidt observed the spectral lines to have shifted from their intrinsic wavelength, placing the "star" at a redshift of $z = 0.158$; a distance of several hundred megaparsecs (Schmidt 1963). This vast distance, in turn, suggested 3C 273 to be an extremely luminous object, while its "quasi-stellar" nature implied it was compact, with the emission being generated within its nuclear region. It was quickly postulated (e.g., Lynden-Bell 1969) that accretion of material onto a supermassive black hole, hosted in the center of galaxies, could provide this enormous luminosity.

It has since become well-accepted that nearly all galaxies host such massive black holes in their centers (Richstone et al. 1998), with those subject to significant accretion emanating as active galactic nuclei. Most galaxies are thought to go through multiple AGN phases (e.g., Kormendy & Ho 2013), though only when the galaxy is active at the time of observations will it be classified as an AGN. From the perspective of galaxy evolution, AGN play an important role. The cosmic star formation history is seen to correlate strongly with the cosmically averaged accretion rate of AGN (Shankar et al. 2009), which is indicative of gas fueling both star formation and the central black hole. Furthermore, the existence of the so-called Magorrian relation (Magorrian et al. 1998), which describes the correlation between the mass of central galaxy bulges and that of their supermassive black holes, indicates some form of co-evolution between the black holes and the galaxy as a whole, despite the vastly different physical scales involved. These findings are generally attributed to feedback processes from the central AGN, via the launch powerful winds or jets that may trigger or alternatively impede star formation in galaxies (e.g., Best et al. 2006; Wang et al. 2010).

The wavelength at which AGN emission can be observed may vary from galaxy to galaxy. In particular, Heckman & Best (2014) distinguish two types of AGN with different observational signatures, related to their particular accretion mechanisms. *Radiative-mode* AGN are characterized by significant accretion – generally in excess of 1% of the Eddington luminosity – and are fed by a surrounding accretion disk.[12] Radiative-mode AGN

[12]The Eddington luminosity scales linearly with the mass of the central black hole, and is given by $L_{\rm Edd} = 3 \times 10^4 \, (M_\bullet / M_\odot) \; L_\odot$, while the accretion luminosity simply follows from

generally emit strong X-ray and ultraviolet emission, originating from the accretion disk and surrounding hot corona (Brandt & Alexander 2015). In addition, such AGN are thought to be surrounded by a clumpy, dusty torus, which, upon being heated by the central black hole, provides additional emission at mid-infrared wavelengths (e.g., Lacy et al. 2004).

Jet-mode AGN, on the other hand, are thought to be fueled not directly by an accretion disk, but via an accretion flow dominated by advection (Heckman & Best 2014). Their accretion rates tend to be lower than those of radiative-mode AGN, being less than a percent of the Eddington rate. In addition, a dusty torus is often absent in jet-mode AGN, and instead these sources are most readily characterized by the radio synchrotron emission they produce, generally in the form of powerful radio jets extending far beyond the outskirts of the galaxy.

Some overlap between these two classes of sources exists, however, as a fraction of AGN emit strongly in both the radio and X-ray regimes (Ceraj et al. 2018; Delvecchio et al. 2018). In turn, various classifications of AGN remain common in astronomical literature, with Padovani et al. (2017) going as far as describing this categorization as an "AGN zoo". Radio astronomers, for example, tend to distinguish two main AGN populations: radio-loud and radio-quiet AGN, which, as their names imply, differ based on their observed level of nuclear radio emission. Radio-loud AGN, in turn, may exhibit a variety of morphologies, leading to the often-used Fanaroff & Riley (1974) classification scheme. Among the most powerful AGN are the *quasars*, which have been detected already as far back as $z \sim 7$ (e.g., Wang et al. 2019). However, only roughly 10% of quasars appear to be radio-loud (e.g., Bañados et al. 2015). In turn, it is clear that a detailed study of the full AGN population involves a multiwavelength perspective, which constitutes the topic of **Chapter two**.

1.6 The Radio Revolution

The single most straightforward thing one can do upon the completion of a new radio survey, is determining the distribution of radio sources as a function of flux density. Perhaps somewhat surprisingly, even this simple counting exercise can have far-reaching implications. The number of sources one expects to measure in a radio survey is namely a function of the *cosmology* of the Universe, and depends, for example, on whether the Uni-

Einstein's famous mass-energy equivalence, with an added efficiency factor $\epsilon \sim 0.1$ (King & Pounds 2015), via $L_{\rm acc} = \epsilon \times \dot{M}_\bullet c^2$. Here, $\dot{M}_\bullet$ is the mass accretion rate of the black hole.

verse is static or evolving (Kellermann & Wall 1987). Back when the first dedicated radio interferometers saw extensive use in astronomy, building upon the rapid development of radar technology during the Second World War, the *steady state theory* of the Universe remained a popular concept. This model involves a Universe that, while expanding, is non-evolving, and has no beginning or end (Bondi & Gold 1948; Hoyle 1948). While the discovery of the cosmic microwave background (Dicke et al. 1965) proved to be the definitive nail in the coffin for the steady state model, radio number counts had already provided strong evidence of its inconsistency with observations (Priester 1958; Ryle & Clarke 1961).

Returning to the present day, the cosmology of the Universe has naturally been constrained to an accuracy beyond what radio number counts can provide. Nevertheless, it remains customary to determine the number counts for any large new radio survey, even if only to ensure the (rather technical process of) data reduction was performed successfully. As the capabilities of radio telescopes continue to improve, new regimes of faint galaxies are beginning to be unraveled – the *faint radio sky*.[13]

1.6.1 The Faint Radio Sky

Observational astronomy, of course, is a field that is constantly driven by technological advancements. As telescopes become bigger, or expand their observable wavelength ranges, new discoveries arise without exception. In that regard, the present is an exciting time for any radio astronomer. Indeed, at the heart of this thesis lies the "revolution" that is currently taking place within radio astronomy: the Very Large Array – the telescope that has without a doubt been most instrumental to this thesis – was constructed in the 1970s, but received a major upgrade in 2012 enhancing its sensitivity by an order of magnitude (Kellermann et al. 2020). Just a few years ago, the *Giant Meterwave Radio Telescope* (GMRT) in India, which saw its first light in 1995, received an upgrade of a similar magnitude. With the aim of probing even longer wavelengths than the GMRT, the *Low Frequency Array* (LOFAR) was constructed in The Netherlands, taking its first observations of the radio sky in 2010. In the decade following, LOFAR has expanded into a multitude of other European countries, further enhancing both its sensitivity and resolution.

[13] What any one astronomer refers to as "faint" is, of course, somewhat subjective. Furthermore, the notion of faintness will even change over time. Some quantitative discussion in the following section is warranted.

The observations provided by this radio revolution allow for significant advances in our understanding of the faint radio sky. As outlined in Section 1.5, the bulk of the bright radio population is composed of active galactic nuclei (e.g., Condon 1984), which can be observed out to cosmological distances. Star-forming galaxies, however, are significantly fainter and require more powerful radio telescopes to be detected at early cosmic times. Owing to the on average increased brightness of radio sources at low observing frequencies, we begin by discussing recent progress in radio observations of star-forming galaxies in the GHz frequency regime.

Quantifying the Faint Radio Sky

One of the first radio surveys capable of probing the *microjansky* radio population over a relatively large area was undertaken in the extragalactic COSMOS field (Schinnerer et al. 2007, 2010). Carried out with the VLA at 1.4 GHz, the VLA-COSMOS survey reached a sensitivity of, roughly $10 - 15\,\mu$Jy across an area of two square degrees. At these depths, bright AGN, with radio powers of $L_{1.4} = 10^{25}\,\mathrm{W\,Hz^{-1}}$, are detectable out to high redshift ($z \sim 5$). However, even with these deep (and only decade-old) observations, just the very brightest star-forming galaxies – the tip of the iceberg – were detectable beyond $z \gtrsim 1$.

The vastly improved capabilities of the current VLA are illustrated by a comparison of the pre-upgrade VLA-COSMOS survey with a more recent example: the 3 GHz COSMOS-XS survey (Figure 1.8; van der Vlugt et al. 2021; Chapter 2).[14] While the former was limited to probing ULIRGs and HyLIRGs, with star formation rates in excess of $\mathrm{SFR} = 100\,M_\odot\,\mathrm{yr^{-1}}$ and $\mathrm{SFR} = 1000\,M_\odot\,\mathrm{yr^{-1}}$, respectively, the COSMOS-XS survey is capable of probing such galaxies out to $z \sim 3$ and $z \gg 5$. [15]

In addition, as outlined in Section 1.2.4, the normalization of the star forming main sequence increases rapidly with redshift, such that at a fixed stellar mass, high-redshift galaxies are more strongly star-forming. In turn, it is illustrative to focus on a galaxy at a constant mass of $M_\star = 10^{10.5}\,M_\odot$, corresponding to roughly half the mass of the Milky Way (Licquia & New-

[14] While carried out at 3 GHz, we converted the relevant COSMOS-XS flux densities to 1.4 GHz with a spectral index of $\alpha = -0.70$; Section 1.3.4.

[15] However, note that the 1.4 GHz VLA-COSMOS survey targets the full COSMOS field across 23 separate pointings (Schinnerer et al. 2007), while COSMOS-XS focuses on a single, deep pointing at 3 GHz. Nevertheless, recent observations with the upgraded VLA across the full field at 3 GHz (Smolčić et al. 2017b) improve upon the depth at 1.4 GHz by a factor of $3\times$, when adopting $\alpha = -0.70$.

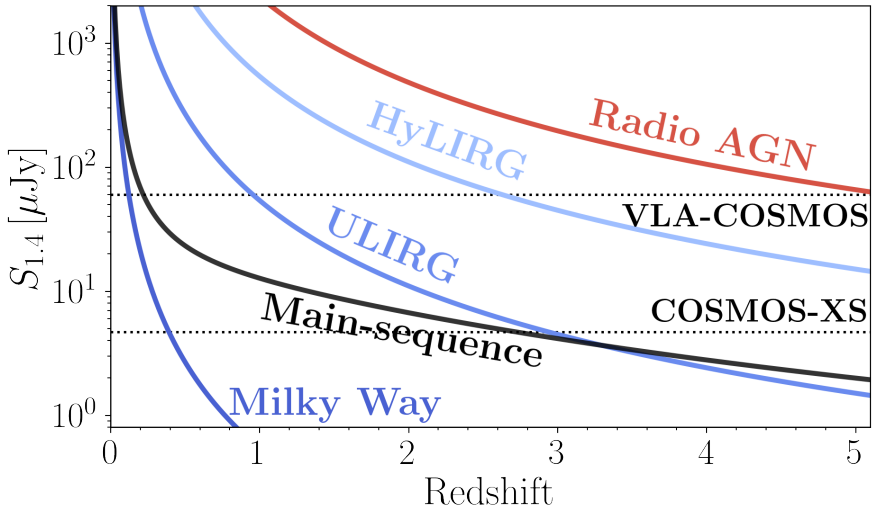


Figure 1.8: The typical 1.4 GHz flux density of various types of radio sources as a function of redshift. A luminous AGN ($L_{1.4} = 10^{25}\,\mathrm{W\,Hz^{-1}}$) remains brighter than $S_{1.4} \gtrsim 100\,\mu$Jy out to $z \sim 4$, and is therefore detectable even in relatively shallow radio surveys, such as the 1.4 GHz VLA-COSMOS survey (5σ detection limit of $S_{1.4} \approx 60\,\mu$Jy). However, Milky Way-like galaxies are virtually undetectable in such surveys, and even ULIRGs (SFR $= 100\,M_\odot\,\mathrm{yr^{-1}}$) can only be observed out to $z \lesssim 1$. Deeper surveys such as COSMOS-XS ($5\sigma \approx 2.5\,\mu$Jy; Chapter 2), made possible by the 2012 upgrade of the Very Large Array, can probe ULIRGs out to $z \sim 3$ and HyLIRGs (SFR $= 1000\,M_\odot\,\mathrm{yr^{-1}}$) across the entire history of the Universe. In addition, due to the rapid evolution of the star forming main sequence with redshift, COSMOS-XS allows for the detection of $M_\star = 10^{10.5}\,M_\odot$ main sequence galaxies out to $z \sim 3$ (black line), while such galaxies are virtually undetectable in the VLA-COSMOS observations.

man 2015; black line in Figure 1.8). The star formation rate of such a galaxy is expected to be similar to that of a ULIRG at $z \gtrsim 3$ (Speagle et al. 2014), while being significantly lower at the present day. In turn, even at this relatively high stellar mass, main-sequence galaxies are virtually undetectable in previous radio surveys at all but the lowest redshifts. The improved sensitivity of the *Karl G. Jansky* Very Large Array, however, provides the ability of detecting such main sequence star-forming galaxies out to $z \sim 3$, or roughly across 80% of cosmic history. As such, the current "revolution" allows for the detection of *representative* galaxies at high redshift with radio telescopes.

1.6.2 The Faint Radio Population

Having established that deep radio surveys are capable of probing faint star-forming galaxies in the early Universe, the question remains to what extent it is indeed truly star formation being probed. Answering this question requires the combination of deep radio observations with equally sensitive observations across the full electromagnetic spectrum, in order to distinguish between star formation and AGN activity. While initially it was thought that star-forming galaxies started dominating the radio sky at flux densities below ~ 1 mJy at 1.4 GHz, subsequent studies (e.g., Bonzini et al. 2013; Padovani et al. 2015) indicated that both radio-quiet and radio-loud AGN still made up a relatively large fraction of the sub-mJy sky.

As multiwavelength data are crucial for the classification of the radio population, an efficient strategy entails targeting well-studied extragalactic fields with deep radio observations. The aforementioned two square-degree COSMOS field, indeed, has seen tremendous efforts from radio and non-radio astronomers alike (e.g., Scoville et al. 2007; Laigle et al. 2016 and references therein). Recently, sensitive 3 GHz observations from Smolčić et al. (2017b) have mapped the entire field down to a detection limit of $\sim 12\,\mu$Jy, with Smolčić et al. (2017a) subsequently providing a detailed view on the incidence of AGN in the radio sky at flux densities down to only a few tens of microjanskies. **Chapter two** introduces the Very Large Array COSMOS-XS survey, which targets a subset of the COSMOS field at two frequencies – 3 and 10 GHz – in a traditional *wedding cake* design: while limited to a smaller area, the COMSOS-XS 3 GHz observations improve upon the depth of the existing 3 GHz data by a factor of five. This combination of large-area but shallower observations with deep data across a smaller field of view is common in radio astronomy as it provides a detailed view of the full radio population.

In addition to this observational perspective, theoretical models for the radio sky have seen rapid development as well, generally with the aim of providing predictions for next-generation radio facilities. These simulations, by their very nature, consist of a complex mixture of observationally determined number densities of the various radio populations, imprinted upon the large scale structure of the Universe as provided by theoretical N-body simulations. Some of the earliest simulations of the radio sky were presented by Wilman et al. (2008), while a recent, improved set of simulations is provided by Bonaldi et al. (2019). We briefly discuss the ingredients of such simulations below, and explicitly compare them to the observed ra-

dio sky in Chapters 2 & 4.

The Bonaldi et al. (2019) simulations, for example, rely on the *Planck Millennium* dark matter-only simulations from Baugh et al. (2019). These provide a realistic framework of the structure of the Universe, allowing for an accurate description of the spatial distribution and clustering of radio sources. In these simulations, individual dark matter haloes are identified and subsequently populated by either star-forming galaxies, or (radio-loud) active galactic nuclei. While Wilman et al. (2008) treat radio-loud and -quiet AGN separately, by modelling the latter based on the observed X-ray luminosity function, Bonaldi et al. (2019) opt to include radio-quiet AGN as a subset of the star-forming population. Observationally, it remains unclear whether the radio emission in the high-redshift radio-quiet AGN population is (significantly) enhanced by the nuclear activity, or whether it is produced mainly through star formation (e.g., Panessa et al. 2019 and references therein; see also Chapter two).

Utilizing observational evidence that radio-loud AGN are typically hosted in massive galaxies (e.g., Sabater et al. 2019), as well as the probabilistic *abundance matching* technique used to assign galaxies of a given stellar mass to a typical halo mass, Bonaldi et al. (2019) subsequently populate the dark matter haloes with a realistic AGN population. This, in turn, is built upon the observed luminosity functions of the various types of radio AGN and their observed redshift evolution (e.g., the models from Massardi et al. 2010). This, generally, requires extrapolating the luminosity functions beyond the redshift ranges where they were constrained observationally, in turn inducing some intrinsic uncertainties into the simulated radio skies.

Modelling of star-forming galaxies similarly builds upon existing observational constraints. Wilman et al. (2008) adopt the measured 1.4 GHz luminosity function by Yun et al. (2001), thereby extrapolating it both to faint radio sources and to high redshift. On the other hand, Bonaldi et al. (2019) use the known (albeit uncertain) correlation between radio emission and star-formation rate to predict radio luminosity functions from the cosmic star formation rate density. In turn, such semi-empirical modelling provides direct predictions of star-forming galaxies and AGN down to ~ 1 nJy at 1.4 GHz; beyond what is currently in reach of radio telescopes. In addition, upon assuming spectral index distributions for the various radio populations, these simulations are capable of providing a view of the full radio sky across a wide range of frequencies (e.g., 150 MHz - 20 GHz in the case of Bonaldi et al. 2019).

What remains a clear gap in our understanding, however, is the nature of the faint high-frequency radio sky. Detailed studies of the composition of the high-frequency ($\gtrsim$ 30 GHz) radio populations in the microjansky regime have not been performed, for the simple reason that these depths have been unattainable for radio telescopes until the advent of the upgraded Very Large Array in 2012. Yet, even the sensitive upgraded VLA requires significant amounts of observing time to probe distant star-forming galaxies at high frequencies, owing to the typical faintness of this population (Section 1.3.4). Nevertheless, this regime is of considerable interest, as it harbors radio free-free emission as a powerful tracer of star formation.

At present, high-frequency detections of star-forming galaxies have remained limited to individual bright sources, potentially aided by gravitational lensing (e.g., Thomson et al. 2012; Wagg et al. 2014; Huynh et al. 2017). In addition, as the upcoming Square Kilometer Array Phase-2 will only provide frequency coverage up to 20 GHz, theoretical models of the radio sky have similarly not attempted to explore the higher frequency regimes. Furthermore, this in turn implies that, until the advent of the next generation Very Large Array sometime in the 2030s or 2040s, the current VLA will remain the go-to facility for exploring the high-frequency radio sky.

Some progress has been made at intermediate frequencies, with a recent study by Murphy et al. (2017), carried out at 10 GHz, providing some initial constraints on free-free emission in the early Universe. However, their observations still remained limited to the radio frequencies dominated by synchrotron emission, and lacked the ancillary data to robustly isolate the free-free component. Combined observations spanning both the low and high frequencies are therefore crucial to fully explore free-free emission in the early Universe, and form the topic of **Chapters 4 & 5**.

1.7 This Thesis

The ongoing revolution in radio astronomy and the accompanying wealth of new data are opening up a new window on galaxy evolution. This thesis includes four studies of the galaxy population in the faint radio sky, with the overarching aim of establishing how radio emission may be used to probe their star formation activity. The structure of this thesis is as follows. We begin by studying the composition of the faint radio population, separating it into star-forming galaxies and AGN (Chapter Two). In Chapter Three, we combine the low-frequency radio emission of dusty star-forming galaxies

with their infrared properties in order to explore the elusive origins of the far-infrared/radio correlation. In Chapters Four and Five, we turn to the high-frequency radio emission in star-forming galaxies, and perform a pioneering study of radio free-free emission in the early Universe. We briefly expand upon these studies below.

Chapter Two: With the advent of new, deep radio observations, the natural question arises what powers the radio emission of the faintest galaxy populations. We investigate the composition of the radio population identified in the Very Large Array COSMOS-XS survey, dividing it into star-forming galaxies and active galactic nuclei based on a variety of multiwavelength tracers of AGN activity. We find that the incidence of radio AGN decreases tremendously among the faint radio population, which is therefore predominantly powered through star formation.

Chapter Three: While the far-infrared/radio correlation provides the crucial recipe for measuring star formation rates via radio emission, its origins remain nebulous. We investigate the nature of the correlation for a large sample of dust-obscured star-forming galaxies at high redshift. These galaxies form an ideal laboratory for studies of the far-infrared/radio correlation, as they can be observed in an unbiased manner. We find no evidence for redshift evolution in the correlation, unlike previous (more biased) samples in the literature. In addition, we find that dusty star-forming galaxies are offset from the local correlation, being radio-bright by a factor of three. We argue this offset is the result of the extreme physical conditions present in such starbursting systems.

Chapter Four: Free-free emission constitutes an ideal probe of star formation activity in galaxies, being both a dust-unbiased and direct tracer. However, free-free emission is overshadowed by the brighter synchrotron emission at low radio frequencies, and hence requires expensive high-frequency observations to be robustly detected. We perform the first survey of free-free emission at high redshift, making use of deep 34 GHz observations from the COLDz survey. We detect eighteen galaxies in these observations, including seven that are likely powered through star formation. We determine free-free star formation rates for this star-forming subsample, and show that these are in agreement with canonical tracers of star formation.

Chapter Five: Left wanting more than just the seven star-forming galaxies from Chapter Four, we push the available radio data to their limits via a multi-frequency stacking analysis on optical/infrared-selected galax-

ies. This analysis is aimed at constraining the shape of the radio spectrum of faint, star-forming galaxies, and detecting their typical level of free-free emission. We find a systematic deficit of high frequency emission in the stacked radio spectra, which we interpret as the ageing of the synchrotron component. We additionally provide the first constraints on the cosmic star formation rate density with radio free-free emission, and find these to be in agreement with canonical constraints.

1.8 The Future

As the "radio revolution" is showing no signs of winding down, the future is, well, *faint*. There is significant room even for the current Very Large Array to extend the studies of free-free emission presented in this thesis. While increasing the depth and/or area covered at the highest frequencies ($\gtrsim 30$ GHz) may be prohibitively expensive, opportunities exist at slightly lower frequencies. As an example, we recently extended the ultra-deep 3 and 10 GHz COSMOS-XS survey to 1.4 GHz, allowing for an unprecedented view of the faint star-forming population with sufficient frequency coverage for free-free emission to be robustly disentangled from the low-frequency synchrotron component.[16] This intermediate frequency regime may, indeed, constitute the ideal compromise between field of view, typical source brightness and frequency lever arm for studies of this powerful star formation tracer. With increased sample sizes, free-free emission cannot only be established as a tracer of star formation, it can be used as a means of *calibrating* other tracers.

Perhaps the one tracer that benefits particularly from such calibrations, is radio synchrotron emission. While the galaxy samples used for studying this tracer through the far-infrared/radio correlation grow increasingly larger, an alternative approach is to investigate in detail how the correlation depends on the physical conditions in individual galaxies. Firmly establishing the correlation across varying physical scales in both starburst and main-sequence galaxies may provide key insights into its puzzling origins. The current VLA is already capable of providing such detailed tests in bright, gravitationally lensed systems, which will provide a benchmark for future facilities.

[16] A brief introduction to these new observations is given in Chapter Five, though future works will fully explore these exciting data.

Any outlook towards the future is certainly incomplete without mentioning the *Square Kilometer Array*. While not extending the frequency coverage of the VLA, the SKA will provide an unprecedented survey speed through its combination of field of view and sensitivity. This, in turn, allows for detailed observations of the microjansky galaxy population at low frequencies across much larger areas than presently feasible. In addition, SKA *deep fields* will probe down to sensitivities of several hundred nanojanskies, and reveal the ever fainter galaxy populations. Parallel efforts by non-radio facilities will subsequently be able to characterize these populations in exceptional detail.

The high-frequency capabilities of the SKA should, however, not be underestimated. Reaching up to 15 GHz in Phase-1 – with the aim of being expanded to 20 GHz in the more distant future – SKA surveys of free-free emission are all but unthinkable. As an example, deep surveys with the high-frequency bands of SKA1-Mid, such as Band 5b with a central frequency of 12.5 GHz (Braun et al. 2019), will probe the free-free-dominated regime in star-forming galaxies at $z \gtrsim 1$. Even modestly star-forming galaxies, with star formation rates of only $10\,M_{\odot}\,\mathrm{yr}^{-1}$ at $z = 1$, will be detectable in just 20 hr of observing time. As a result, a COSMOS-XS-like survey can be undertaken with the SKA in just a fraction of the observing time currently required. For detailed observing strategies and time estimates in SKA-based studies of free-free emission, we refer the reader to Chapter four.

The true *game changer*, however, will be the next-generation VLA, which will revolutionize our understanding of the radio spectra of faint, distant galaxies by virtue of its impressive sensitivity and frequency coverage. The ngVLA is expected to span an impressive range of 1.2 - 116 GHz, with a bandwidth reaching up to 20 GHz (Selina et al. 2018). In turn, large-area surveys of free-free emission will become, perhaps, the most powerful tool of measuring star formation in the early Universe.