ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN ASTROFISICA

Ciclo XXXVII

Settore concorsuale: Astronomia, Astrofisica, Fisica della terra e dei pianeti (02/C1)

Settore Scientifico Disciplinare: Astronomia e astrofisica (FIS/05)

Searching for high-redshift progenitors of massive galaxies

Presentata da:
Fabrizio Gentile

Coordinatore Dottorato
Prof. Andrea Miglio

Supervisore
Dott.ssa Margherita Talia

Co-supervisore
Prof. Andrea Cimatti

Esame Finale Anno 2025

Crescere, amare, cominciare a capire:
La nostra vita è un continuo cercare,
come un gabbiano che plana sul mare

-Traditional scout song

Abstract

This Thesis focuses on the process of star formation and its relationship with the formation and evolution of the most massive galaxies observed in the Universe. The main hypothesis behind this work is that the picture of the high-z Universe based on deep optical/near-infrared (NIR) surveys is potentially incomplete, since it misses most of the contribution from dusty sources that can be unveiled only with multi-wavelength data. In this Thesis, I follow two complementary approaches to support this hypothesis. In the initial chapters, I focus on the radio-selected (RS-) NIR-dark galaxies: a population of likely dusty star-forming galaxies selected at radio frequencies. I characterize these sources by taking advantage of the almost complete photometric coverage offered by the COSMOS field and involve in the analysis state-of-the-art facilities such as ALMA and the James Webb Space Telescope (JWST). The main result of this first part of the Thesis is the determination of the physical properties of the RS-NIRdark galaxies, picturing them as a population of highly dust-obscured ($A \sim 4$ mag), massive ($M_\star \sim 10^{10.5-11}$ $M_\odot$) and star-forming (SFR$\sim 300 - 500$ $M_\odot$ yr^{-1}) galaxies mainly located at $z \sim 3$ and beyond. These properties are then employed to estimate the expected contribution of these sources to the cosmic star formation rate density.

The second approach, followed in the fifth chapter of the Thesis, focuses on a class of extremely red sources unveiled in the first data collected with JWST and missed by old-generation facilities. The analysis presented in this work pictures them as extremely massive ($M_\star \sim 10^{11}$ $M_\odot$) sources at $z \sim 5 - 7$. The existence of such massive galaxies in the first Gyr of cosmic history suggests a much more efficient star formation in the early Universe than what is commonly derived by observing low-z galaxies.

These two approaches, combined, shed new light on the evolutionary paths that bring to the formation of massive galaxies in the low-z Universe.

Contents

Chapter 1

Introduction

One of the main open questions in modern astrophysics and cosmology is understanding how galaxies evolved from the initial perturbations of the density field after the Big Bang to the celestial objects we observe in our local Universe. The main process responsible for the growth of galaxies is star formation, i.e. the conversion of gas into stars as an effect of gravity (see e.g. the review by McKee & Ostriker 2007 and references therein). The main scientific goal of this Thesis is the study of the star formation across most of the cosmic history, to unveil how this process gave birth to the most massive galaxies that we observe in our Universe. The Thesis can be roughly divided in two parts: in the next three chapters, I will follow a statistical approach to this problem, by studying the cosmic Star Formation Rate Density (SFRD), while in Chapter 5 I will approach the problem from a different perspective, focusing on a class of extreme objects and on the constraints on the star formation efficiency that can be derived from their discovery.

In this Introduction, I will introduce the SFRD and the procedure followed for its estimation (Section 1.1). After discussing how decades of astrophysical research allowed us to obtain a quite consistent picture of the SFRD as a function of cosmic time, I will discuss how these results can be affected by possible biases arising from the presence of cosmic dust (Section 1.2). In the following part of the Introduction, I will discuss how the employment of different tracers of star formation and the inclusion of some additional sources can help us to obtain a more complete picture of the SFRD (Sections 1.3 and 1.4). Finally, I will present the main subjects of this Thesis (Section 1.5) and summarize its main scientific goals and aims (Section 1.6).

1.1 The cosmic star formation rate density

The cosmic Star Formation Rate Density (SFRD) quantifies the average amount of stellar mass that is created around the Universe in a given time and cosmic volume (usually, it is expressed as the number of solar masses created per year per Mpc^3). The detailed study of this quantity has a pivotal role in the field of the evolution of galaxies, since it is strongly related to the mass assembly of the Universe: the integral of the SFRD, as a function of redshift, gives the stellar mass density (SMD; see e.g. Weaver et al. 2023a; Shuntov et al. 2024). Moreover, its comparison with the molecular gas mass density and with the black hole accretion rate density can be

employed to study the efficiency of the star formation process (see e.g. Decarli et al. 2016 or Tacconi et al. 2020 and Combes 2018 for reviews) and its relation with the growth of super massive black holes (SMBHs; e.g. Vito et al. 2018), respectively.

The estimation of the SFRD relies on deep galaxy surveys, collecting large samples of objects in a given region of the sky above a certain limiting flux, and can be done at two levels.

At a base level, the SFRD can be computed by simply summing up the star-formation rate (SFR) values of all the galaxies detected in the field, in a given redshift interval, weighted by the maximum comoving volume occupied by each galaxy. This estimate, however, represents just a lower limit, because it only accounts for the contribution of the sources actually detected above a flux threshold in a given band.

At a higher level, the SFRD can be derived by computing and integrating the luminosity function (LF) at a SFR-sensitive wavelength, thus statistically accounting for those sources that are fainter than the survey limiting flux (see e.g. the seminal works by Schechter 1976 and Saunders et al. 1990, or Binggeli et al. 1988 for a review). This method, however, severely depends on the assumptions made on the shape of the LF.

1.1.1 The Hubble view of the cosmic SFRD

The current picture of the cosmic SFRD as a function of the cosmic time is the offspring of the joint effort of several collaborations in the last two decades. Most of them took advantage of the deep galaxy surveys conducted with the Hubble Space Telescope (HST) and other optical/near-infrared (NIR) facilities to assemble large samples of galaxies and constrain the behavior of the SFRD up to $z \sim 7$. All these efforts are summarized in the review by Madau & Dickinson (2014) and in the famous plot shown (in an updated version) in Figure 1.1, with the results based on deep optical/NIR surveys reported in green (Cucciati et al., 2012; Bouwens et al., 2012a,b), while the results achieved at longer wavelengths (mainly with the *Herschel* and *Spitzer* Space Telescopes; e.g. Gruppioni et al. 2013; Magnelli et al. 2013) are reported in red).

It is possible to notice how the vast majority of the data-points at redshifts higher than 3 are obtained by deep optical/NIR surveys. The reason for this resides in the rest-UV emission sampled by these sources in this redshift range, resulting particularly useful for two reasons. Firstly, it is well known that this regime of the spectral energy distribution (SED) of star-forming galaxies (SFGs) is dominated by the emission of young stellar populations (ages between 10 and 200 Myr for a standard Initial Mass Function – IMF – see e.g. Chabrier 2003). The large contribution of young stars to the UV luminosity makes it an excellent tracer of recent star-formation, with several well-established conversions linking this quantity to the SFR (see e.g. Kennicutt & Evans 2012 and references therein for the most common calibrations). Secondly, the un-extincted SED of a SFG presents a significant break at 912Å due to the absorption of the rest-UV stellar emission by the neutral interstellar medium (ISM) in the galaxy (the so-called Lyman break). The presence of this break allows us to easily assemble large samples of high-z galaxies through the "dropout" technique (the so-called Lyman Break Galaxies – LBGs – see e.g. the review by Giavalisco 2002 and references therein). The picture of the cosmic SFRD

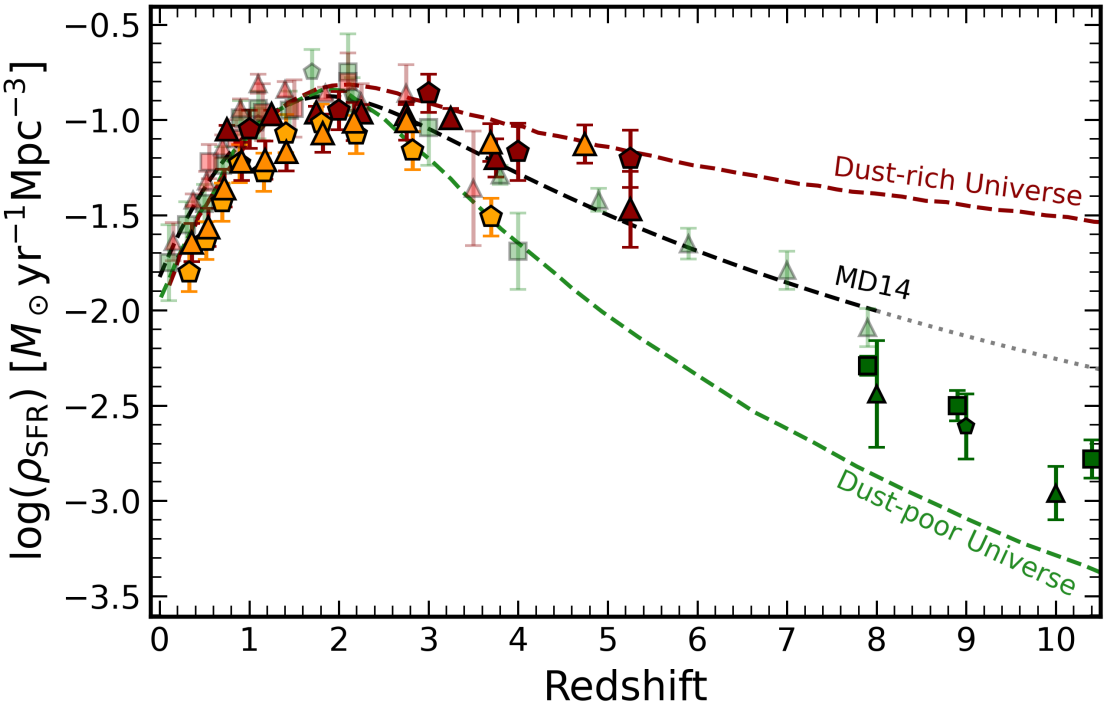

Figure 1.1: Updated version of the original plot from Madau & Dickinson (2014) representing the SFRD as a function of redshift. The original data-points employed by Madau & Dickinson (2014) for their fit (reported as the dashed black line) are shown as the shaded points. Those obtained by optical/NIR facilities (Cucciati et al., 2012; Bouwens et al., 2012a,b) are reported in green, while those obtained at longer wavelengths (Gruppioni et al., 2013; Magnelli et al., 2013) are reported in red. The same color-coding is employed for more recent studies, such as those obtained with JWST (solid green points) by Bouwens et al. (2023), Donnan et al. (2023), and Harikane et al. (2023) (triangles, squares, and pentagons, respectively). Similarly, the more recent studies at (sub)mm wavelengths by Gruppioni et al. (2020) and Traina et al. (2024) (solid red pentagons and triangles, respectively). Finally, the studies performed at radio wavelengths by Novak et al. (2017) and van der Vlugt et al. (2022) are reported as solid orange triangles and pentagons, respectively. The dotted gray line reports the extrapolated relation by Madau & Dickinson (2014) at $z > 8$.

presented by Madau & Dickinson (2014) and shown in Figure 1.1 is mainly based on these LBGs (especially at $z > 3$, where this selection technique is particularly useful; see Bouwens et al. 2012a,b) and can be summarized as a steep ($\propto (1+z)^{1.7}$) rising from $z \sim 0$ to $z \sim 2-3$ (the so-called "cosmic noon", where the SFRD reaches its maximum) followed by a steep ($\propto (1+z)^{-2.9}$) decline up to $z \sim 7$.

1.1.2 Going beyond: first results from JWST

This picture is basically confirmed by the new results achieved in the last years thanks to the James Webb Space Telescope (JWST). The unprecedented sensitivity achieved by this facility in the NIR allowed several collaborations to sample fainter rest-UV fluxes (i.e. reducing the range of luminosities that are extrapolated by LFs) and to collect galaxies at higher redshifts (up to $z \sim 14$; see the current record-holder galaxy whose spectroscopic confirmation is presented in Carniani et al. 2024). Even though most of these results are still based on photometric redshifts (uncertain and possibly prone to catastrophic failures, see e.g. some notable examples in Zavala et al. 2023 and Jin et al. 2024), the picture of the high-z SFRD is in quite good agreement with the overall trend suggested by Madau & Dickinson (2014) of a steep decline after the cosmic noon (see e.g. the results obtained by Bouwens et al. 2023; Donnan et al. 2023; Harikane et al. 2023 and reported in Figure 1.1 as solid green points), even though the slope suggested by these new studies seems to be steeper than what is expected by simply extrapolating the relation by Madau & Dickinson (2014) at higher redshifts. This difference could have important consequences on the efficiency of star formation in the early Universe, suggesting a more complex interplay between star formation and the ISM in the epoch of pre-reionization (see e.g. Barkana & Loeb, 2000) or on the shape of the IMF in the early Universe (see e.g. the discussion in Harikane et al. 2023 and references therein).

1.2 How accurate is the HST/JWST view?

Even though the data obtained with HST, JWST, and other optical/NIR facilities are overall in good agreement with the picture summarized by Madau & Dickinson (2014), in the last decades several clues started to suggest that their perspective on the galaxy evolution could be incomplete. One of the most interesting results concerns the recent discovery of a significant population of massive ($M_\star \sim 10^{11}$ $M_\odot$) and passive (sSFR$< 10^{-11}$ yr^{-1}) galaxies already in place at $z \sim 2-3$ (see e.g. Straatman et al. 2014; Schreiber et al. 2018a; Valentino et al. 2020a, but also de Graaff et al. 2024; Weibel et al. 2024 for more recent results obtained with JWST at even higher redshifts). The existence of such sources is puzzling for at least two orders of reason. Firstly, these galaxies are found to be quite common at their redshifts, with volume densities in the order of a few 10^{-5} Mpc^{-3}. This makes it hard to identify their progenitors, since the number density of the most massive LBGs (i.e. those galaxies whose evolution could bring to the formation of these sources) is at least one or two orders of magnitude lower (see e.g. the number densities presented in the review by Giavalisco 2002 and the discussion on this point by Toft et al. 2014 and Valentino et al. 2020a). A second crucial cause of tension is that

massive and passive galaxies are found just a few Gyrs after the Big Bang, therefore they should have assembled all their stellar mass in a very short amount of time, in a period of the cosmic history where the SFRD of the Universe was at least one order of magnitude lower than at the cosmic noon (see Figure 1.1).

A possible solution to these issues could reside in a selection bias due to the wavelength employed in the aforementioned studies on the SFRD. As described in the previous sections, most of the high-z samples of galaxies employed in these studies have been selected though optical/NIR facilities sampling, at $z > 3$, their rest-frame UV radiation, hence missing those sources that are either intrinsically UV-faint (i.e. less star-forming) or highly dust-attenuated. As mentioned in Section 1.1, the former can be accounted for statistically through the employment of a LF. Accounting for the latter is more challenging, since their properties could – in principle – differ significantly from those of the UV-bright population of galaxies (see e.g. Wang et al. 2019; Talia et al. 2021; Smail et al. 2021; Barrufet et al. 2023; Gottumukkala et al. 2023; McKinney et al. 2024). For this reason, as I will show in the remaining of this Introduction, several studies suggested that the inclusion of these dusty sources in the cosmic census could change our picture of the high-z SFRD (see e.g. Rowan-Robinson et al. 2016; Wang et al. 2019; Gruppioni et al. 2020; Talia et al. 2021; van der Vlugt et al. 2023) and shed light on the missing progenitors of the massive and passive galaxies at $z \sim 3$ (e.g. Toft et al. 2014; Valentino et al. 2020a).

1.2.1 The effect of cosmic dust

Dust is a fundamental component of galaxies. It is constituted by solid particles (grains) that are distributed in the interstellar medium (ISM), especially in the molecular and atomic gas clouds. Its presence in galaxies absorbs the stellar and nebular emission from the other components of the galaxy, making the observed fluxes fainter than the intrinsic ones. This effect is strongly dependent on the rest-frame wavelength of the emission, with bluer wavelengths being more extincted than the redder ones (see e.g. the review by Salim & Narayanan 2020 and references therein). The dependence of the dust attenuation on the wavelength is usually parametrized with the so-called dust attenuation laws (see e.g. Calzetti et al. 2000; Charlot & Fall 2000). The energy absorbed by the dust is then re-emitted thermally at longer wavelengths, with the peak of the emission being in the far-infrared (FIR; $\lambda \sim 15 - 1000$ μm) regime.

The main issue connected with the presence of dust in the observed galaxies is that the rest-UV emission cannot be directly employed to estimate the SFR, since the observed luminosity needs to be corrected for the dust extinction. The most common procedure consists in the estimation of the $A_{\rm v}$ parameter from the shape of the observed SED (through SED-fitting) or from spectroscopy (when available, e.g. through the so-called "Balmer decrement"), and in the consequent de-extinction of the SED through the assumption of an attenuation law. This procedure is somehow risky, since the common assumption is that these laws do not evolve with redshift (being normally calibrated on local or low-z galaxies) and that they are valid regardless of the environment. Several studies analyzing sources at $z > 4 - 5$ suggest that these assumptions could be an oversimplification (e.g. Fudamoto et al. 2020a,b) with important consequences on the estimation of the intrinsic properties of high-z

galaxies. These issues can be mitigated with the employment of different tracers of star formation.

1.3 Other tracers of star formation

In addition to the rest-UV luminosity, other observables can be used to estimate the SFR of a galaxy. In this Section, I will focus on two of the tracers of star formation that are employed in the rest of the Thesis, although acknowledging that others exist, such as the luminosity in the X-rays or that of the $H\alpha$ line. The interested reader can refer to the review by Kennicutt & Evans (2012) and to the references therein for a more detailed discussion on other tracers of star formation.

- **Infrared luminosity:** As described in the previous paragraph, dust absorbs most of the rest-UV radiation emitted by stars. That energy is then re-emitted thermally at longer wavelengths in the rest-frame IR. For this reason, the IR luminosity can be used as a tracer of star formation. Among the several possible calibrations available in the current literature, in the rest of the Thesis, I will employ that by Kennicutt & Evans (2012).

  The accuracy of IR luminosity as tracer of star formation is – however – dependent on the amount of stellar emission absorbed by interstellar dust. In the case of dusty starbursts, it is a good approximation to assume that all the UV radiation is re-emitted at longer wavelengths, making it an optimal tracer. However, in less star-forming or less dusty objects (e.g. low-metallicity dwarf galaxies) the IR luminosity can miss a significant amount of star formation, making this tracer only complementary to the (dust-uncorrected) UV luminosity (see e.g. Wuyts et al. 2011; Talia et al. 2015). Since the main objects of this Thesis are dusty starburst galaxies (see Section 1.5), the choice of this tracer is well justified.

- **Radio luminosity:** Radio photons can trace star formation since they are produced as synchrotron emission by accelerated electrons in *supernovae* and as free-free emission in HII regions. Even though a calibration immediately linking radio luminosity and SFR from first principles has never been established, the radio emission in star forming galaxies is found to tightly correlate with IR luminosity through the so-called IR-radio correlation (e.g. Molnár et al. 2021). Therefore, the radio luminosity can be employed as a star formation tracer by assuming this correlation and the SFR/$L_{\rm IR}$ relation (see e.g. Novak et al. 2017).

  The main drawback in the use of this tracer is represented by the possible contribution of nuclear emission to the radio luminosity of the analyzed galaxies. It is well known, in fact, that AGN can emit at radio frequencies due to the presence of significant amounts of electrons that are accelerated by the intense magnetic fields originated in the proximity of the super massive black hole (see e.g. the review by Tadhunter 2016 and references therein). As I will describe in the next Sections, this issue can be partly overcome when multi-wavelength data are available for the galaxies of interest.

As I will show in the following parts of this Introduction, the employment of other tracers of star formation is crucial to fully understand our Universe, as nearly half of the total star formation taking place in the cosmos is obscured by dust, as shown by several studies analyzing the (sub)mm cosmic background (see e.g. the review by Hauser & Dwek 2001 and reference therein).

1.3.1 Dusty star-forming galaxies

The more significant the amount of dust in galaxies, the more important it is to rely on multiple tracers of star formation to avoid missing important pieces of information. This requirement becomes crucial when dealing with the so-called "dusty star-forming galaxies" (DSFGs; see e.g. the review by Casey et al. 2014 and references therein) which are, to simplify, star-forming galaxies hosting significant amounts of dust. Collecting these galaxies is not easy, since their nature is hard to unveil with optical/NIR data alone: bright tracers of star-formation at longer wavelengths are needed. Depending on the tracer, we can classify different sub-populations of DSFGs.

1.3.2 (sub)mm selection

The first method to pick up DSFGs relies on detections at FIR/(sub)mm wavelengths. This selection is the most straightforward, because a significant emission in this regime is an optimal tracer of the dust thermal emission due to ongoing star formation, and it is hardly reproducible by other processes taking place in galaxies. This explains the important role of the (sub)mm-selected galaxies (also known as sub-millimetre galaxies; SMGs) in the history of the study of DSFGs, that started with the first facilities able to observe the FIR/(sub)mm sky (see e.g. the initial studies by Rowan-Robinson et al. 1991 and Irwin et al. 1998) and went on with the assembly of the first statistically significant samples of these sources (e.g. Smail et al. 1997; Hughes et al. 1998; Barger et al. 1998).

The (sub)mm selection is also incredibly efficient thanks to the strongly negative k-correction of SMGs at (sub)mm wavelengths. Taking as a reference Figure 1.2, it is possible to notice how redshift has an almost negligible effect on the (sub)mm flux of a DSFG. This property allows high-z dusty galaxies to be detected even in relatively shallow (sub)mm maps, such as those presented in some notable studies by Smail et al. (1997); Hughes et al. (1998), and Barger et al. (1998) .

The main drawback of the FIR/(sub)mm selection is merely observational and it is related to the facilities able to observe the sky at these wavelengths. Old generation instruments such as the SCUBA (*Submillimetre Common-User Bolometer Array*) bolometer equipped on the James Clerk Maxwell Telescope (JCMT) or the PACS (*Photodetector Array Camera and Spectrometer*) and SPIRE (*Spectral and Photometric Imaging Receiver*) cameras equipped on the *Herschel* Space Telescope, are characterized by low spatial resolutions (up to tens of arcsec) and sensitivities. The effect of these properties is twofold. On the one hand, the low spatial resolution makes it incredibly hard to associate the correct multi-wavelength counterpart to each FIR/(sub)mm source. This issue makes the estimation of the physical properties (including the redshift) of each source incredibly challenging (see e.g. Danner-

bauer et al. 2004). On the other hand, the low sensitivity of old-generation facilities biases the collected samples towards the most extreme objects able to produce the brightest (sub)mm fluxes (see e.g. Figure 1.3). These galaxies are generally characterized by high values of the SFR (up to $\sim 10^3$ $M_\odot$ yr^{-1}; see e.g. Swinbank et al. 2014; da Cunha et al. 2015), not necessarily representative of the bulk of the DSFG population.

These issues could be solved – in principle – with state-of-the-art (sub)mm interferometers such as the *Atacama Large Millimetre Array* (ALMA) or the *NOrtern Extended Millimetre Array* (NOEMA), characterized by higher spatial resolutions (up to 0.1" in the most extended ALMA configurations) and sensitivities. The downside of these instruments, however, is represented by their small field of view (FOV) and low mapping speed, making these telescopes unsuitable to perform large and deep blank surveys of the sky, although the total covered area is increasing with each observational cycle (see e.g. Dunlop et al. 2017; Franco et al. 2018; Casey et al. 2021, and Long et al. 2024). This issue well explains the low number of homogeneously selected samples of DSFGs collected by these instruments (see the number counts in the aforementioned references). Nevertheless, these new facilities can be used in combination with old-generation ones to perform follow-ups at higher spatial resolution of regions with bright and large (sub)mm emission and accurately measure the position and flux of DSFGs (see some noteworthy examples in Simpson et al. 2019; Dudzevičiūtė et al. 2021 and McKinney et al. 2024). This procedure increases the number of detected DSFGs, but makes it more complex to determine their selection function, since the higher resolution and sensitivity of the new facilities can unveil the presence of multiple (and fainter) sources contributing to the same low-resolution (sub)mm signal. A final caveat potentially affecting the (sub)mm selection of DSFGs is that it relies on the still-unclear high-z properties of dust: a correlation between dust temperature and redshift, as suggested e.g. by Schreiber et al. (2018b), Liang et al. (2019), and Faisst et al. (2020), could produce selection effects in the samples collected at (sub)mm wavelengths. More in detail, an increasing dust temperature would move the peak of the dust emission to shorter wavelengths, producing fainter fluxes at the usual wavelengths (~ 870 μm) where SMGs are collected.

An overview of the physical properties of SMGs can be obtained by looking at one of the largest samples of these sources in the current literature: the 707 galaxies collected in the AS2UDS survey by Dudzevičiūtė et al. (2021) consisting in the ALMA follow-up of SCUBA-2 sources with $S_{850\mu m} > 3.6$ mJy (and unveiling sources up to $S_{870\mu m} \sim 0.6$ mJy). From this study, we can see how the redshift distribution of SMGs peaks around $z \sim 2.5 - 3$ and how — on average — these sources represent a population of massive ($M_\star \sim 10^{11}$ $M_\odot$) and highly star-forming (SFR ~ 300 $M_\odot$ yr^{-1}) galaxies (Figure 1.3).

1.3.3 MIR selection

An alternative way of selecting DSFGs relies on the detection of Extremely Red Objects (EROs; see e.g. Huang et al. 2011; Wang et al. 2016, 2019), based on the difference in color between the NIR and MIR emission. DSFGs can be captured by this selection since the combination of ongoing star formation and dust obscuration

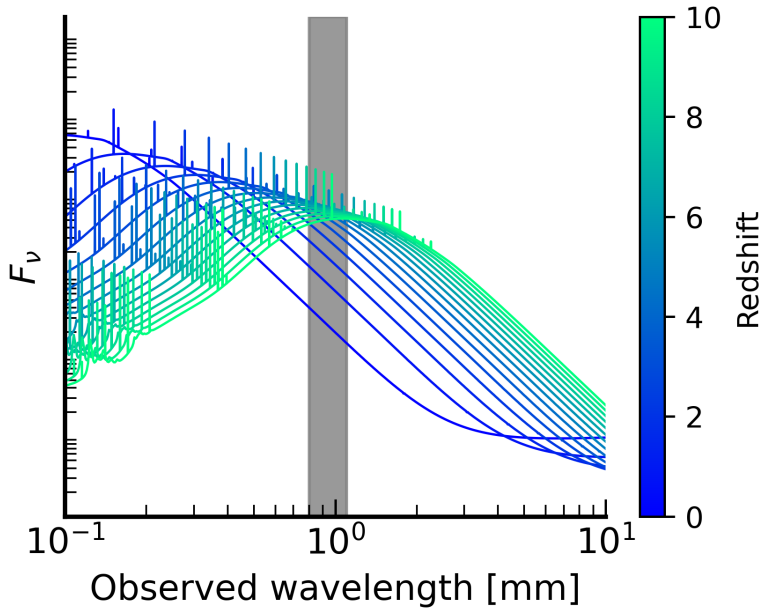

Figure 1.2: Sketch presenting the strongly negative k-correction of SMGs at (sub)mm wavelengths. The plot shows the FIR SED of a DSFG at different redshift. It is possible to notice how the flux at 870 μm (commonly employed for the selection of SMGs) increases from $z \sim 0$ to $z \sim 4$ and stays almost constant up to $z \sim 10$. This property explains why the (sub)mm selection is one of the most common methods for finding DSFGs.

can produce a red optical/NIR slope that at high-z ($z > 3$) is sampled by filters in the the two aforementioned regimes. For instance, the so-called H-dropouts by Wang et al. (2019) were selected as galaxies with $H - [4.5] > 2.1$ mag. This selection takes advantage of the higher sensitivity and spatial resolution of NIR and MIR facilities with respect to (sub)mm ones. For this reason, it enables the collection of samples of less extreme DSFGs (i.e. less massive and star-forming; see Wang et al. 2019 and Figure 1.3). The same color-selection has been employed by some recent JWST-based studies to collect analogous samples of sources up to fainter [4.5] fluxes (Barrufet et al., 2023; Pérez-González et al., 2024; Gottumukkala et al., 2023).

The main drawback of this selection consists in its inability to collect a pure sample of DFSGs. Indeed, there are other populations of galaxies that reproduce the NIR-to-MIR red colors without necessarily hosting significant dust-obscured star formation. Several studies presenting detailed analysis of these sources assessed how a significant fraction (up to $\sim 20\%$) of these samples can be represented by passive galaxies or dust-reddened AGN (Pérez-González et al., 2024; Barrufet et al., 2024). True DSFGs can be singled out thanks to additional data such as a (sub)mm detection (as in Wang et al. 2019) or with a rest-optical spectrum showing the presence of significant $H\alpha$ emission (as in Barrufet et al. 2024), to trace ongoing star-formation activity.

As before, we can have an overview of the physical properties of MIR-selected DSFGs by looking at large samples of these sources. In this case, I focus on the 148 galaxies collected by Gottumukkala et al. (2023) in the JWST survey CEERS

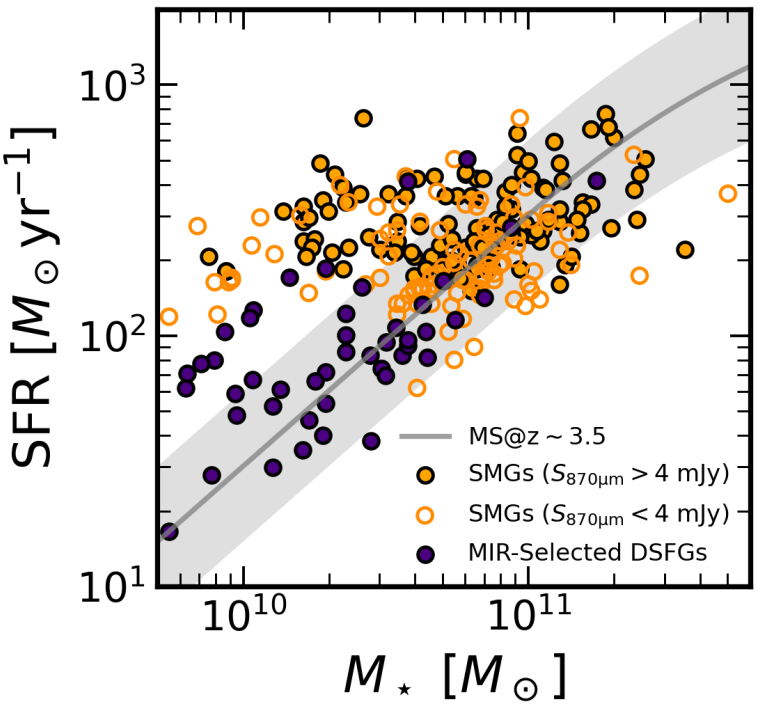

Figure 1.3: Stellar masses and SFR of $3 < z < 4$ DSFGs selected via their (sub)mm or MIR emission (orange and indigo points, respectively). The MIR-selected galaxies are those studied in Gottumukkala et al. (2023), while the (sub)mm-selected ones are those reported in Dudzevičiūtė et al. (2021). SMGs are divided in two classes depending on their flux at 870 μm. The sources with $S_{870\mu m} > 4.2$ (i.e. those compatible with the sensitivity of old-generation surveys such as Weiß et al. 2013) are reported as full points, while the empty markers indicate the fainter sources unveiled with ALMA follow-ups. The gray line and the shaded area report the main sequence of star-forming galaxies at $z \sim 3.5$ by Schreiber et al. (2015) and its 1σ scatter.

(Finkelstein et al., 2023). For this study, we can see how the redshift distribution of these sources peaks at slightly higher redshift with respect to the SMGs ($z \sim 3.13$) and how this selection includes — on average — less massive ($M_\star \sim 10^{10.1}$ M$_\odot$) and less star-forming (SFR ~ 40 M$_\odot$ yr^{-1}) galaxies (Figure 1.3).

1.4 The multi-wavelength picture of cosmic SFRD

From a theoretical perspective, the impact on the SFRD of the employment of different star formation tracers and of the selection of high-z galaxies at wavelengths other then UV-to-NIR, strongly depends on the amount of dust in the high-z Universe. Considering, for instance, the models presented by Casey et al. (2018) (Figure 1.1), the SFRD at $z \sim 4.5$ of a dust-rich Universe could be up to one order of magnitude higher than what would be observed in a dust-poor one (and even more at higher redshifts).

As it can be appreciated from Figure 1.1, most of the optically/NIR-based stud-

ies employed by Madau & Dickinson (2014) to fit their SFRD(z) relation well agree with the prediction of the dust-poor Universe. It is crucial to underline, however, that the surveys employed in these studies sampled rest-UV radiation of the analyzed galaxies. Therefore, their selection was naturally biased towards dust-poor sources (not necessarily representative of the overall population of high-z sources) and potentially missed significant fractions of their star formation if that happened in dust-obscured environments.

If we consider studies based at longer wavelengths, the picture becomes more complex. Most studies based at lower redshift are unanimous in confirming the trend presented in Madau & Dickinson (2014) until $z \sim 3$, while at higher redshifts the results start to diverge.

Some studies such as Gruppioni et al. (2020) or Novak et al. (2017) (red pentagons and orange triangles in Figure 1.1) — selecting galaxies at (sub)mm or radio frequencies — found a much flatter behavior of the SFRD at $3 < z < 4.5$ than what obtained by Madau & Dickinson (2014). The same trend is not found in other studies such as those by Traina et al. (2024) or van der Vlugt et al. (2022) (red triangles and orange pentagons in Figure 1.1), finding high-z behaviors more in agreement with that reported by Madau & Dickinson (2014) even when other tracers are employed. Other contradictory results have been presented by other studies inferring the SFR from the infrared luminosity of SFGs observed in the ALMA large programs ALPINE (Béthermin et al., 2020; Faisst et al., 2020; Le Fèvre et al., 2020) and REBELS (Bouwens et al., 2022). These studies include those by Khusanova et al. (2021) (finding a SFRD at $z > 4$ higher than what reported by Madau & Dickinson 2014), Algera et al. (2023) (finding a SFRD at $z \sim 7$ compatible with Madau & Dickinson (2014)), and Barrufet et al. (2023) (SFRD at $z \sim 7$ significantly lower than Madau & Dickinson (2014)). It is crucial to underline, however, that these results are still based on UV-selected galaxies and — therefore — still biased towards dust-poor objects.

A possible point of contact between these families of results is represented by the analysis discussed in Zavala et al. (2021) and based on the first results from the MORA survey (see Casey et al. 2021, but also Long et al. 2024). In this study, the authors employ the number counts at 2 mm to suggest a transition from an Universe dominated by dust-unobscured star formation to another dominated by dust-obscured one around $z \sim 4$.

1.5 This Thesis: a radio selection for DSFGs

The main hypothesis behind this Thesis is that a more precise picture of the high-z Universe can only be achieved by focusing on all the galaxies that have been missed by HST and other old-generation optical/NIR facilities. In the next three chapters, I aim at assembling a statistically significant sample of homogeneously-selected DSFGs, in order to unveil their contribution to the cosmic SFRD. The main subject of these chapters are the "radio-selected NIRdark galaxies", hereafter RS-NIRdark (see the initial studies by Chapman et al. 2001, 2002 and the more recent by Talia et al. 2021; Enia et al. 2022; Behiri et al. 2023). Together with my group, I broadly define these sources as radio-detected galaxies with a lacking counterpart at optical/NIR wavelengths. The principles guiding this selection are

that bright radio fluxes can trace ongoing star formation (once properly accounting for the possible contribution by nuclear activity), while the requirement of faint (down to non-detections) NIR fluxes should limit the collected sample to the most dust-obscured and high-z sources. This selection takes advantage of the capabilities of modern radio interferometers (high sensitivity and spatial resolution coupled with large FOVs and high mapping speed) to solve some of the issues affecting (sub)mm-selected samples. Moreover, the employment of radio-emission as the tracer of star formation makes the selection robust against possible biases related to the properties of dust. Compared with the MIR selection, the radio-based one is less prone to contamination by passive sources, since it is, by construction, biased towards actively star-forming galaxies.

The drawback of this selection is represented by the possible contribution of nuclear activity to the radio emission, that needs to be taken into account in order to properly estimate the SFR of the detected sources. Moreover, the radio frequencies have a positive k-correction, limiting the possibilities of finding SFGs at very high redshift ($z > 5$).

The initial studies on these sources seem to support the initial hypothesis on the nature of the RS-NIRdark galaxies as likely DSFGs, estimating a total contribution to the cosmic SFRD at $z > 4.5$ up to $\sim 40\%$ of the estimate by Madau & Dickinson (2014) (see Talia et al. 2021; Enia et al. 2022; Behiri et al. 2023).

1.6 Main goals and structure of this Thesis

The first three chapters of this Thesis focus on the search and characterization of RS-NIRdark galaxies and on estimating their contribution to the cosmic SFRD and to the evolution of the most massive galaxies in the Universe. This Thesis continues the series of studies started with Talia et al. (2021), Enia et al. (2022), and Behiri et al. (2023) and overcomes one of the main limitations of these studies: the lack of a proper deblending of the low-resolution maps in the MIR regime, limiting the analyzed samples to the most isolated sources. As I will show in the rest of the Thesis, the development of a deblending pipeline allows me to analyze the full sample of galaxies. Moreover, the estimated contribution of the RS-NIRdark galaxies to the cosmic SFRD by Talia et al. (2021) and Behiri et al. (2023) only represents a lower limit, since it only account for the sources detected in the analyzed survey. The ultimate goal of this Thesis will be to statistically account for the fainter sources through estimation of the first radio luminosity function of the RS-NIRdark galaxies.

All the analyses presented here are based on the data collected in the Cosmic Evolution Survey (COSMOS) field (Scoville et al., 2007; Koekemoer et al., 2007). The choice of this field allows me to take advantage of one of the largest (~ 2 deg^2) and deepest ($\sigma = 2.3$ μJy beam^{-1}) radio surveys currently on the scene (the VLA-COSMOS large program; see Smolčić et al. 2017 and Section 2.3 for a more detailed discussion) to perform the initial selection. Moreover, since COSMOS is one of the most famous extra-galactic fields in modern astronomy, it has been observed by several of the main telescopes in the last two decades, offering an almost complete photometric coverage spanning from the radio to the X-rays (see e.g. Laigle et al. 2016; Weaver et al. 2022 and references therein). Finally, the central region of the COSMOS field (~ 0.54 deg^2) has been observed by the new JWST during

the COSMOS-Web survey (Casey et al., 2023), providing deep and high-resolution NIR and MIR data for this area of the sky. These data are extensively used in the fourth chapter of the Thesis to complement the initial analysis on the radio-selected NIRdark galaxies. The same data are the crucial ingredient for the analysis presented in Chapter 5, describing the discovery and analysis of some extremely massive galaxies in the COSMOS field.

This Thesis is structured as follows. Chapter 2 describes the initial selection of galaxies in the COSMOS field, the assembling of the full photometric catalog and the first assessment of photo-z, physical properties, and number densities. Chapter 3 focuses on the first follow-up at (sub)mm wavelengths performed with ALMA on a pilot sample of nine galaxies. This chapter presents the first spectroscopic confirmation of the photo-z through (sub)mm spectroscopy, the estimation of the molecular gas mass inside the targets, and a simple evolutionary model linking our sources to the massive and passive galaxies at $z \sim 2-3$. Chapter 4 describes the improvements in the analysis achieved with the new JWST COSMOS-Web data. The main result of the chapter is the first estimation of the radio luminosity function for my sources and their contribution to the total cosmic SFRD, and a thorough comparison with other notable populations of DSFGs with JWST data. Chapter 5 describes the analysis of a pilot sample of two extremely red and massive galaxies at $z \sim 5-7$ discovered in the COSMOS-Web field. This chapter, although characterized by a slight change in the topic, with respect to the previous chapters, nicely complements the main goal of the Thesis, by exploring alternative pathways for the formation of lower-z massive and passive galaxies. Finally, the main conclusions of the Thesis, along with the future perspectives of this work, are presented in Chapter 6.

Throughout this Thesis, I will assume a standard flat ΛCDM cosmology with the parameters $[h, \Omega_{\rm M}, \Omega_{\Lambda}] = [0.7, 0.3, 0.7]$. I will also assume a Chabrier (2003) IMF and the AB photometric system (Oke & Gunn, 1983).

Chapter 2

A radio selection for DSFGs: Initial analysis

2.1 Introduction to the chapter

This chapter describes my initial selection of RS-NIRdark galaxies in the COSMOS field, the assembly of the full photometric catalog, and the first estimation of the physical properties through SED fitting. It is structured as follows. In Section 2.2, I introduce the sample selection. In Section 2.3, I describe the maps employed to extract the photometry in the optical-to-MIR regime and the additional photometry retrieved at longer wavelengths. In Section 2.4, I present the new `PhoEBO` pipeline (Photometry Extractor for Blended Objects) used to extract the photometry and its validation on simulated data. Section 2.5 is focused on the SED-fitting procedure employed to assess the photo-zs and the physical properties of the galaxies in my sample. Furthermore, in Section 2.6 I analyze in detail the possible AGN contribution. In Section 2.7, I discuss what the results presented in this chapter tell us about the nature of the RS-NIRdark galaxies and compare them with previous analogous studies in the current literature. Finally, I summarize the main conclusions of the chapter in Section 2.8.

Most of the analysis presented in this chapter was presented in the paper "*Illuminating the Dark Side of Cosmic Star Formation III: Building the largest homogeneous sample of Radio-Selected Dusty Star-Forming Galaxies in COSMOS with PhoEBO*" by Gentile et al. (2024a), published in The Astrophysical Journal (Volume 962, Issue 1, id.26, 16 pp.).

2.2 Sample selection

To assemble a large sample of candidate DSFGs with complete photometry from the optical to the radio, together with my collaborators, I focus on the galaxies in the COSMOS field. We collect our sources by performing a selection analogous to that employed by Talia et al. (2021) and summarized here:

1. We start from the VLA-COSMOS 3GHz Large Project catalog (Smolčić et al., 2017). We select 8850 radio sources with an S/N$>$ 5.5 ($S_{3\mathrm{GHz}} > 12.65$ μJy) over the full ~ 2 deg^2 coverage of the survey. This cut allows us to assemble

a sample of radio-detected galaxies, limiting the likely contamination of fake sources to 0.4% (Smolčić et al., 2017).

2. To limit the expected presence of galaxies where nuclear activity significantly contribute to the radio luminosity (common in radio-selected catalogs; see e.g. Bonzini et al. 2013 and Novak et al. 2018), we remove from the sample all the sources flagged as "multi-component" in the initial catalog. We underline that this flag is the result of a visual inspection of the sources in the full catalog by Smolčić et al. (2017). Therefore — at this stage — we cannot exclude that some multi-component radio sources are still present in our sample (see the discussion in Section 2.6).

3. To take advantage of the multi-wavelength coverage of the COSMOS field, we exclude from the sample all the radio sources lying outside the UltraVISTA survey footprint (Laigle et al., 2016). This further limits the sample to 5982 galaxies and the effective area mapped by this study to 1.38 deg^2.

4. Finally, we cross-match the resulting sample with the two versions of the COSMOS2020 catalog (the "`Classic`" and "`Farmer`"; see Weaver et al. 2022), removing all the sources with a match within 0.7". The final sample is composed of 323 galaxies.

The last step aims at including in the catalog only sources without a significant counterpart in the optical/NIR bands. However, it is crucial to underline how the source detection in the COSMOS2020 catalog is performed on the χ^2-image produced through the `SWarp` software (Bertin et al., 2002) by combining the maps in the i and z band from the Subaru telescope and the Y, J, H and Ks bands from the VISTA telescope (see Weaver et al., 2022). The flux in the generic pixel I of this image is computed as a weighted average of the fluxes f_i in the different photometric bands, with the weights w_i provided by the uncertainty maps:

$$I = \sqrt{\frac{\sum_{i=1}^{N} w_i f_i^2}{N}} \tag{2.1}$$

Given this definition, we cannot exclude that some of the sources in our catalog could have a significant counterpart in a few individual NIR bands. Moreover, through a visual inspection of the χ^2-map employed by Weaver et al. (2022) for detecting the sources in the COSMOS2020, we notice that 60 galaxies of our sample have a detection in that image with a S/N higher than 1.5σ (i.e. the threshold used in Weaver et al. 2022). Some of these sources ($\sim 10\%$) are inside regions masked in the COSMOS2020 that were employed in the COSMOS2015, while most of the lasting sources are close to a bright companion, therefore it is likely that they were not properly deblended in the COSMOS2020. Since these galaxies are not included in the COSMOS2020, they should be part of our sample of RS-NIRdark galaxies. However, due to the different photometry, we expect their properties to differ from the rest of the sample. Therefore, we do not consider these sources in all the statistical analyses performed in this chapter aimed at characterizing our population of galaxies: for all these studies, we will focus on the lasting 263 galaxies. Finally, we underline that the S/N cut applied to the radio catalog excludes from the sample all the faintest

radio-sources. Although this step could potentially cause the lack of significantly high-z galaxies (see e.g. the source described in Casey et al. 2019), it is necessary to exclude a high contamination by fake sources (this rate would increase up to 25% just including the sources with $5 < \mathrm{S/N} < 5.5$, see Smolčić et al. 2017). Both these issues will be discussed in greater detail in Chapter 4.

A significant difference between this sample selection and those employed in the previous papers on analogous sources (Talia et al. 2021 and Behiri et al. 2023) resides in the last step: while the previous works selected a sample of 476 RS-NIRdark galaxies without a counterpart in the COSMOS2015 catalog (Laigle et al., 2016), we employ the updated (and deeper) COSMOS2020 catalog (Weaver et al., 2022). This step allows us to exclude from the sample 153 galaxies un-detected in the COSMOS2015 but revealed by COSMOS2020. More in detail, 16 sources were only included in the "`Farmer`" catalog, 21 only in the "`Classic`", and 116 in both the catalogs. Finally, improving the works by Talia et al. (2021) and Behiri et al. (2023), we do not employ any additional selection aimed at avoiding sources with a bright contaminant in the vicinity, but we analyze the entire sample, performing a more accurate photometry extraction in presence of contamination or source blending.

2.3 Data

2.3.1 Analyzed maps

The analysis of DSFGs requires the most comprehensive wavelength coverage to account for all the physical processes taking place in these complex objects (i.e. stellar emission, dust obscuration,thermal emission, and non-thermal processes). For this purpose, we extract accurate photometry in the optical-to-MIR regimes by analyzing the following maps:

- **Optical**: We analyze the maps produced with the Subaru telescope's Hyper Suprime Cam (HSC) during the Subaru Strategic Program (DR 3; Aihara et al. 2019), targeting the COSMOS field in the g, r, i, z and y bands.
- **NIR**: The photometry at NIR wavelengths is extracted from the DR4 maps of the UltraVISTA survey (McCracken et al., 2012) performed with the VISTA telescope in the Y, J, H and Ks bands.
- **MIR**: We analyze the maps produced with the *Infrared Array Camera* (IRAC) of the *Spitzer* space telescope. The maps analyzed in this study are the deepest ones made with this facility as part of the Cosmic Dawn Survey (Moneti et al., 2022), obtained by co-adding all the available exposures of the COSMOS field for each of the four channels of the IRAC camera.

Further details on the maps employed in the photometry extraction can be found in Table 2.1.

Table 2.1: Main properties of the maps analyzed in Section 2.3.Part of the data are reproduced from Weaver et al. (2022) with author's permission.

Instrument /Telescope	Band	$\lambda^{(a)}$ (Å)	$\Delta\lambda^{(b)}$ (Å)	Depth$^{(c)}$ (2")	Corr. Fact.$^{(d)}$	PSF FWHM$^{(e)}$ (")
Hyper-Suprime Cam/Subaru	g	4847	1383	28.1	1.4	0.79
	r	6219	1547	27.8	1.4	0.75
	i	7699	1471	27.6	1.5	0.61
	z	8894	766	27.2	1.4	0.68
	y	9761	786	26.5	1.4	0.68
VIRCAM /VISTA	Y	10216	963	25.3/26.6$^{(f)}$	2.7/2.8$^{(f)}$	0.82
	J	12525	1718	25.2/26.4$^{(f)}$	2.5/2.7$^{(f)}$	0.79
	H	16466	2905	24.9/26.1$^{(f)}$	2.4/2.6$^{(f)}$	0.76
	Ks	21557	3074	25.3/25.7$^{(f)}$	2.4/2.4$^{(f)}$	0.75
IRAC /Spitzer	ch1	35686	7443	26.4	—	—
	ch2	45067	10119	26.3	—	—
	ch3	57788	14082	23.2	—	-
	ch4	79958	28796	23.1	—	-

Notes:
$^{(a)}$ Median λ of transmission curve
$^{(b)}$ FWHM of the transmission curve
$^{(c)}$ 3σ depths as reported by Weaver et al. (2022)
$^{(d)}$ Multiplicative correction for the photometric uncertainties (Weaver et al., 2022)
$^{(e)}$ Values taken from Aihara et al. (2019) and McCracken et al. (2012). For the IRAC maps we employ the PSF publicly available on the IRSA website.
$^{(f)}$ Two values for the deep/ultra-deep stripes in the UltraVISTA survey

2.3.2 Additional photometry

To analyze the dust emission in the FIR/(sub)mm and the non-thermal processes emitting at radio frequencies, we retrieve additional photometry for our sources by cross-matching our sample with other catalogs in the current literature:

- **FIR**: We cross-match our catalog with the 2020 version of the SuperDeblended catalog by Jin et al. (2018). Since the procedure followed in building that catalog employs as positional prior the sources in the VLA-COSMOS 3GHz Large Project, it is possible to retrieve FIR photometry obtained with *Spitzer* (24 μm; Le Floc'h et al. 2009), *Herschel* (100, 160, 200, 250, and 500 μm; Lutz et al. 2011), SCUBA-2 (850 μm; Cowie et al. 2017; Geach et al. 2017), AzTEC (1.1 mm; Aretxaga et al. 2011), and MAMBO (1.2 mm; Bertoldi et al. 2007) for all the galaxies in our sample. Further details on the employed maps and on their depth can be found in Jin et al. (2018).

- **(sub)mm**: We cross-match our sample with the catalog of the *Automated Mining of the ALMA Archive in the COSMOS Field* (A3COSMOS) survey (v.20200310; Liu et al. 2019). This catalog contains (sub)mm continuum fluxes for all the sources in the COSMOS field observed with ALMA. Since the coverage of the A3COSMOS survey is not uniform, only a tiny percentage of our RS-NIRdark sources has a counterpart in this catalog (32 galaxies within a matching radius of 1", $\sim 10\%$; Liu et al. 2019), with central bandwidth and depth strongly variable for different sources.

- **Radio**: Finally, we retrieve additional radio photometry by cross-matching our catalog with the public catalog by Schinnerer et al. (2010) containing 1.4 GHz photometry obtained during the VLA-COSMOS survey (77 galaxies; $\sim 24\%$ of the full sample). We also include radio photometry at 1.28 GHz from the MIGHTEE Early Science Data Release (170 sources within a matching radius of 8"; $\sim 53\%$ of the sample; Jarvis et al. 2016; Heywood et al. 2022).

2.4 `PhoEBO`: A new pipeline for photometry extraction

2.4.1 Description of the pipeline

Extracting accurate photometry for the sources in our catalog is a challenging task. The main limitation is represented by the possible presence of bright contaminants close to our (extremely faint or even undetected) galaxies in the optical/NIR bands. This issue becomes highly significant in the IRAC channels, where the large (up to FWHM= 2″ in ch3 and ch4) and irregular Point Spread Function (PSF) makes it almost impossible to blindly deblend multiple sources without priors on their positions and shapes.

Basic deblending algorithms such as that implemented in `sExtractor` (Bertin & Arnouts, 1996) require a minimum contrast between the different blended components inside the same cluster of pixel to recognize the presence of multiple sources.

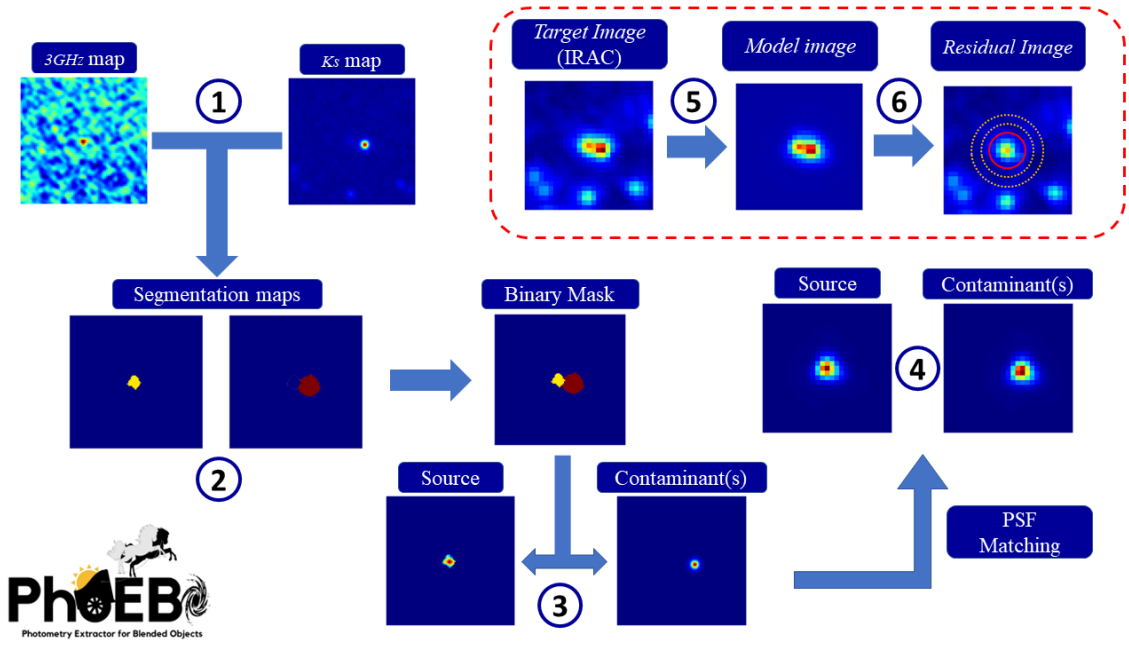

Figure 2.1: Scheme summarizing the workflow of the deblending algorithm `PhoEBO` employed to extract the optical/NIR/MIR photometry for the RS-NIRdark galaxies in the sample. Further details on the numbered steps and on the full procedure are given in Section 2.4.1.

This level of contrast is not reached in the IRAC bands, making these algorithms unfit to analyze our galaxies. More complex softwares such as `Tractor` (Lang et al., 2016) and `Farmer` (Weaver et al., 2022, 2023b), relying on parametric profile-fitting, can overcome this problem by extracting prior information on the position and shape of the different objects through a high-resolution image in which all the blended components are present and distinguishable at the same time (e.g. the χ^2-map employed by Weaver et al. 2022). As before, these techniques cannot be applied to our RS-NIRdark galaxies, since the only bands in which our galaxies are robustly detected — together with the contaminants — are the IRAC channels, where it is generally impossible to distinguish the different sources. Therefore, in order to build the photometric catalog, we developed a new deblending pipeline called `PhoEBO` (Photometry Extractor for Blended Objects)[1] relying on a slightly modified implementation of the method employed in several previous studies (e.g. Labbé et al., 2006; Endsley et al., 2021; Whitler et al., 2022). Apart from this Thesis, the `PhoEBO` pipeline has been already employed by Langan et al. (2024) and Sapori et al., (*in prep.*) to analyze analogous sources in the COSMOS and J1030 fields.

The whole procedure followed by the pipeline is summarized in Figure 2.1 and follows these steps:

1. We start from two high-resolution "*detection images*" and a low-resolution "*target image*" to be deblended. One *detection image* must contain information on the RS-NIRdark galaxy we want to analyze, the other on the contaminant sources. In our case, we choose the 3GHz and the Ks maps, respectively.

[1]The code is freely available in a `GitHub` repository: https://github.com/fab-gentile/PhoEBO

We underline that we minimize the risk of biases by employing two maps with a comparable PSF FWHM (0.75" and 0.78", respectively). An upgraded version of the software developed in this thesis called `PhoEBO+` and able to perform with images with different PSFs is presented in Sapori et al., (*in prep.*).

2. We employ the basic deblending algorithm included in the `Sep` library (Bertin & Arnouts, 1996; Barbary, 2016a) to identify the different sources in the two *detection images* and to produce the segmentation maps (i.e. a set of images assigning each pixel to one of the galaxies in the field). To include all the galaxies in the maps, we set a low detection threshold (2σ) and a minimum of 5 contiguous pixels in each detected source.

3. We combine the two segmentation maps and use their product as a binary mask to isolate the sources present in the *detection images* and produce a different image for each of them.

4. We convolve each image produced in the previous step with a matching kernel to homogenize the PSF with that of the *target image*. The matching kernel is produced with the "`photutils`" library (Bradley et al., 2020) using the ratio of Fourier transforms (Gordon et al., 2008; Aniano et al., 2011). The PSF of the *detection images* is modeled as a 2D Gaussian with a FWHM equal to the PSF FWHM of the Ks band reported in McCracken et al. (2012) and Weaver et al. (2022). The PSF of the optical/NIR bands are obtained in the same way (i.e. with Gaussians with fixed widths, see Table 2.1), while the PSFs of the four IRAC channels are downloaded from the IRSA archive[2]

5. We normalize all the resulting images and co-add them into a "*model image*". This one is resampled to match the pixel size of the *target image* (if needed) and then fitted to the *target image* by multiplying each normalized source by a free-parameter α_i. The fit is performed with the `Scipy` library (Virtanen et al., 2020), aiming at minimizing the χ^2 between the *model* and the *target image* .

6. We multiply all the components of the *model image* for the relative α_i obtained through the fitting procedure. Then, we subtract all the resulting images of the contaminants from the original *target image*. In doing so, we get a *residual image* containing only the source present in our sample.

7. Finally, we perform aperture photometry with `Photutils` on this residual image by employing a fixed diameter of 2 arcsec. The local background is computed in an annulus between 1.5 and 2 times the radius of the aperture and subtracted from the extracted counts. The counts are then converted to AB magnitudes and to micro-Jansky through the photometric zeropoints employed in Weaver et al. (2022), corrected for the systematic offset reported in Weaver et al. (2022). This last correction is needed to account for the systematic mismatch between the spec-zs and the photo-zs computed in the COSMOS field (Laigle et al., 2016; Weaver et al., 2022).

[2] https://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/

The whole procedure described above allows us to estimate the fluxes in all the NIR and IRAC bands reported in Section 2.3.1. As prescribed by the IRAC Instrument Handbook, we correct the fluxes measured through aperture photometry by the factors reported in the IRAC Instrument Handbook[3]. For the optical bands — with a smaller PSF FWHM than the radio map at 3GHz and the *Ks* map — we subtract the contaminants with `PhoEBO` by using the HSC-*i* band as a *detection image*. Finally, we estimate the photometric uncertainties as:

$$\Delta F = \sqrt{\sum_{i \in A} \sigma_i^2} \tag{2.2}$$

where the σ are obtained from the weight maps and the sum is extended to all the pixels in the aperture employed in the estimation of the flux. As prescribed by Weaver et al. (2022), we correct the photometric uncertainties in the optical/NIR bands with the multiplicative factors reported in Table 2.1. This step is required to account for the expected under-estimation of the uncertainties through Equation 2.2 due to the presence of correlated noise in the analyzed maps (e.g. Leauthaud et al., 2007). Since the weight maps are not affected by the subtraction of the contaminants, we underline that the procedure followed here to estimate the photometric uncertainties is totally consistent with that employed by Weaver et al. (2022) for the optical/NIR bands. Regarding the IRAC bands, we can compare our uncertainties with those by Weaver et al. (2022) by running `PhoEBO` on the 153 RS-NIRdark galaxies selected by Talia et al. (2021) and excluded from this study being revealed in the COSMOS2020 (see Section 2.2). We obtain that the median ratio between our uncertainties and those included in the "Classic" COSMOS2020 is in the order of ~ 3. This result can be explained by the likely underestimation of the IRAC uncertainties found by Weaver et al. (2022) in the COSMOS2020 catalog.

2.4.2 Validation of `PhoEBO`

Thanks to the small difference in the PSFs, the photometry extraction performed by `PhoEBO` in the optical/NIR bands does not differ significantly from that performed by "classic" algorithms such as `sExtractor` (Bertin & Arnouts, 1996). Nevertheless, the extraction in the IRAC channels is quite different. Therefore, in order to obtain reliable results on the physical properties extracted from the photometric catalog, we need to validate the performances of the pipeline in the MIR regime. We validate the results of the `PhoEBO` pipeline performing extensive simulations of blended galaxies in the four IRAC channels. The simulation procedure recalls the philosophy discussed in Section 2.4.1:

1. We simulate two noise maps through a random Gaussian generator, requiring an rms compatible with the 3GHz images observed in the VLA-COSMOS survey at 3 GHz and with the *Ks* band images observed in the UltraVISTA survey (conservatively, we employ the sensitivity reached in the deep stripes).

2. We simulate a radio source and a NIR-bright contaminant. Both the galaxies are simulated as 2D Gaussians on the noise maps generated in the previous

[3] https://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/iracinstrumenthandbook/

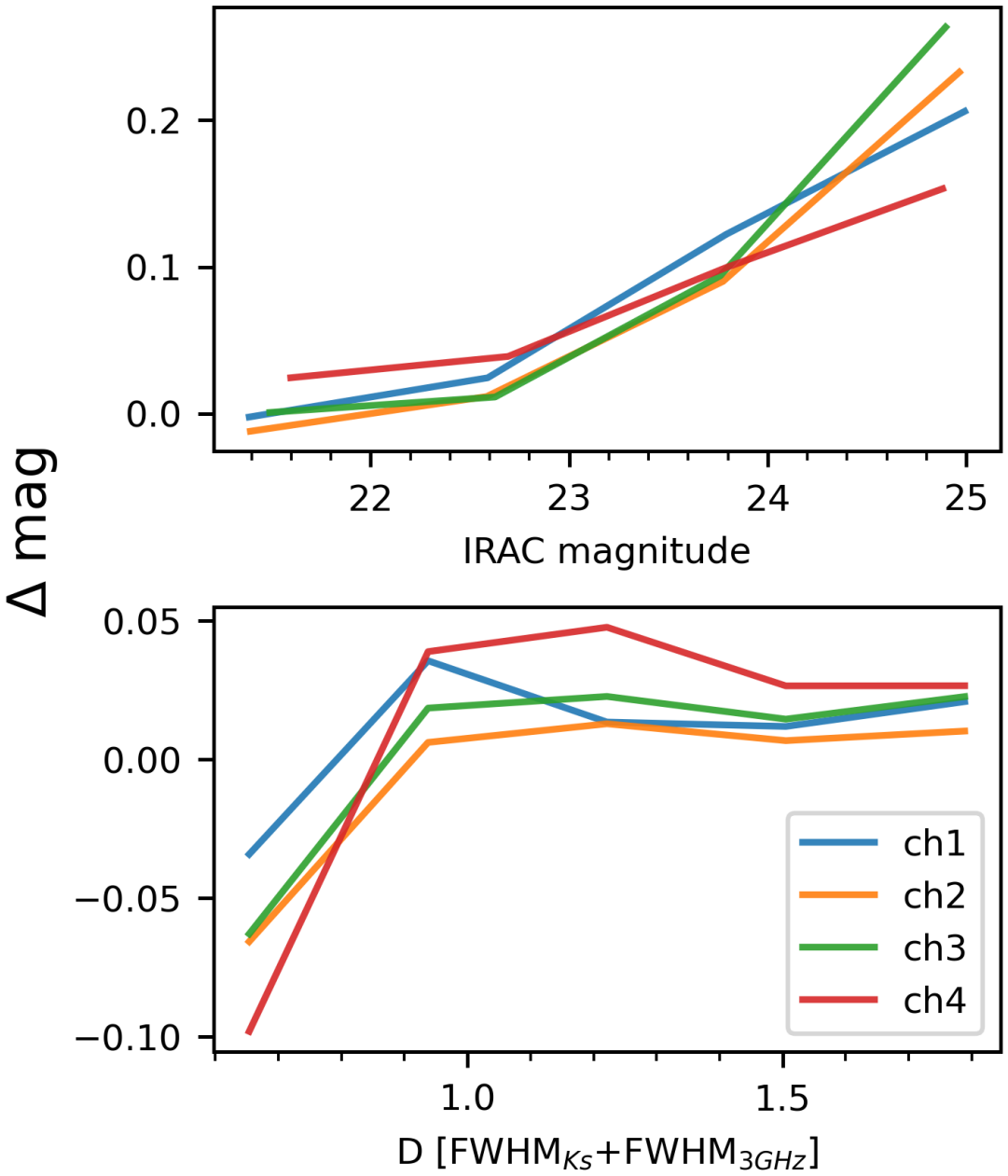

Figure 2.2: Accuracy of the `PhoEBO` pipeline as a function of the IRAC flux of the RS-NIRdark galaxies and of the angular distance between these sources and the blended contaminant (a proxy for the blending between the sources, once parameterized through the sum of the FWHMs of the two sources). These results are obtained by applying the `PhoEBO` pipeline to the set of simulated images described in Section 2.4.2.

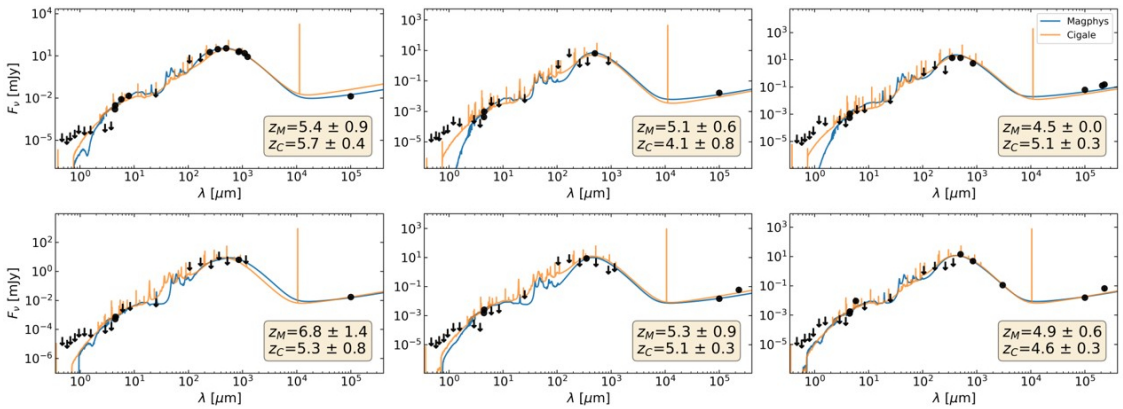

Figure 2.3: Some examples of SEDs of high-z candidates fitted with `Magphys` (blue solid lines; da Cunha et al. 2008; Battisti et al. 2019) and `Cigale` (orange solid lines; Boquien et al. 2019). The photometry is represented by the black dots (for the detections at $S/N > 3$) and the black arrows (for the upper limits). The yellow boxes report the photo-zs computed with the two codes and the relative uncertainties.

step. Since the vast majority of the RS-NIRdark galaxies are unresolved at 3 GHz, we choose a FWHM=0.7" for the radio sources (i.e. the width of the synthesized beam in the VLA-COSMOS survey). On the contrary, to account for the presence of partially resolved contaminants, the FWHMs of the Gaussians in the Ks band are uniformly sampled in the range [0.7,1.4] arcsec (i.e. between one and two PSF FWHM). The normalization of the two Gaussians are chosen to obtain a S/N uniformly sampled in the range [5.5,8] and [3,8], for the radio and Ks images, respectively. The radio sources are placed in the center of each image, while the contaminants are allowed to space in the range [0.7,1.4]arcsec from the center. This range allows us to study the accuracy of the algorithm as a function of the blending between the two sources. The lower limit on the distance is chosen as 0.7", since we recall that — due to the selection described in Section 2.2 — all the NIR sources with a separation lower than this threshold are considered NIR counterparts of the radio signal.

3. We convolve each of these Gaussians with a set of matching kernels computed as prescribed in Section 2.4.1 to obtain the sources as observed in each of the four IRAC channels.

4. We rescale the flux of the convolved radio source to obtain an integrated flux in the range [21,M] mag, where M is the limiting magnitude in each IRAC channel. Consequently, we rescale the flux of the convolved contaminant to obtain a flux ratio uniformly sampled in the range [0.1,1] (with the radio source brighter than the other galaxy).

5. Finally, we co-add the two images and add Gaussian noise with the same rms expected in the four IRAC channels.

To explore the whole parameter space of the randomly sampled quantities, we

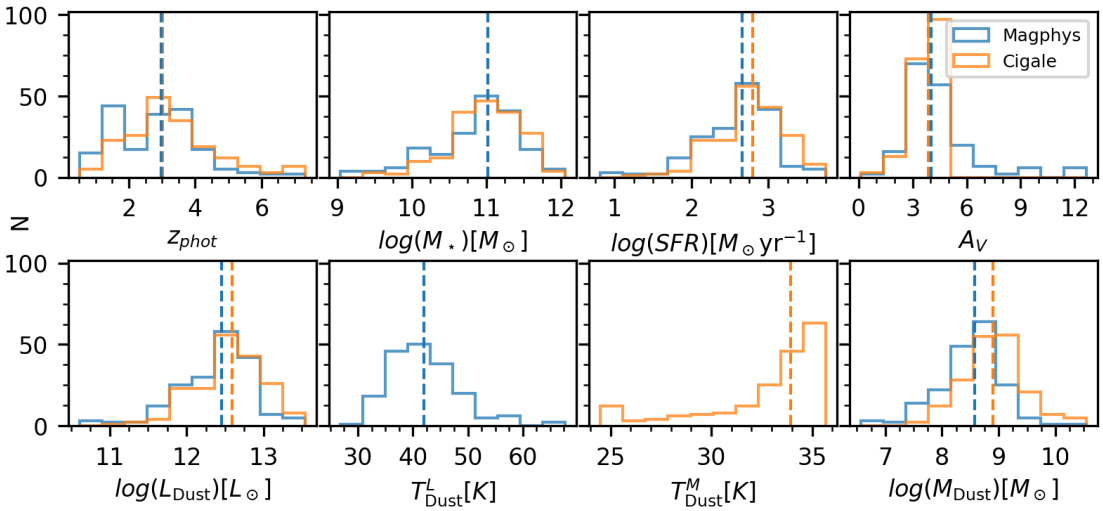

Figure 2.4: Distribution of the photo-zs and of the main physical properties(in order: stellar mass, star formation rate, extinction, dust luminosity, dust temperature — both luminosity- and mass-weighted — and dust mass) of the RS-NIRdark as estimated by SED-fitting with `Magphys+photo-z` (da Cunha et al., 2015; Battisti et al., 2019) and `Cigale` (Boquien et al., 2019) in blue and orange, respectively. The dashed lines of the same colors mark the median values of the distributions. Further details are given in Section 2.5.

simulate $\sim 10^3$ images. Then, we run the `PhoEBO` pipeline on the IRAC-like images, employing as *detection images* the high-resolution data in the radio and Ks bands. In order to assess the performances of the pipeline, we compare the fluxes reported by `PhoEBO` with those obtained by performing a standard aperture photometry on the isolated IRAC-like images (i.e. those obtained in point 3 of the simulation procedure, before adding the contaminant).

Computing the Δmag on the whole dataset, we obtain a median(Δmag)$\sim$0.03 in all the IRAC channels, without any significant difference between the different bands. Similarly, we obtain a std(Δmag)$\sim$0.15 in all the channels. Moreover, the employment of simulations allows us to estimate the Δmag as a function of the different parameters employed during the simulation procedure (Figure 2.2). We find interesting — albeit expected — trends with the IRAC flux and with the angular separation between the high-resolution images, with lower accuracy achieved on more blended sources and IRAC-fainter RS-NIRdark galaxies.

2.5 Physical properties from SED-Fitting

To determine the nature of the RS-NIRdark galaxies, we need to assess their photo-zs and physical properties. We do this through an SED-fitting procedure. To test the robustness of our results against the characteristics of different codes, we base our analysis on two algorithms: `Magphys+photo-z` (da Cunha et al., 2008; Battisti et al., 2019) and `Cigale` (Boquien et al., 2019).

2.5.1 SED-fitting with `Magphys`

`Magphys` (da Cunha et al., 2008) is a physically-motivated SED-fitting code based on the energy balance between stellar attenuation and thermal dust emission. The software estimates the physical properties of a galaxy by comparing its optical-to-radio broad-band photometry with more than a million templates, including stellar emission, dust attenuation, thermal dust emission and non-thermal radio emission. The stellar emission is considered by combining the Single Stellar Populations (SSPs) by Bruzual & Charlot (2003) with an exponentially declining Star Formation History (SFH) with random bursts of star formation superimposed on the continuum. The dust attenuation is included as prescribed by Charlot & Fall (2000), with the addition of a 2175 Å feature accounting for the young stars born in denser clouds. A three-component model accounts for thermal dust emission. These components include the "hot" dust (i.e. warmed up by young stars in the birth clouds), the "cold" dust (i.e., present in the diffuse ISM), and the characteristic emission by the Polycyclic Aromatic Hydrocarbures (PAHs). The first two components are modeled as modified gray bodies, while the PAH emission is modeled with an empirical template (see da Cunha et al. 2008 for details). The code computes the dust temperature ($T_{\rm d}$) as the luminosity-average of these three components. Finally, `Magphys` includes the radio emission from star formation as prescribed by da Cunha et al. (2015). All the ranges of the free parameters employed in the templates can be found in da Cunha et al. (2008) and da Cunha et al. (2015). In this work, we use the `photo-z` version of `Magphys` (Battisti et al., 2019), able to estimate the photometric redshift together with the physical properties of the analyzed galaxies.

2.5.2 SED-fitting with `Cigale`

To avoid possible biases arising from the use of a single SED-fitting code, we involve the software `Cigale` (Boquien et al., 2019) in the analysis. This code is based on a similar energy-balance principle as `Magphys` but allows a larger customization of the libraries employed in building the templates. In this chapter, we start by using a setup as close as possible to `Magphys` to achieve consistent results. The SSPs, SFH, dust attenuation and radio emission are the same as discussed in Section 2.5.1. A significant difference between the two codes resides in the treatment of the thermal dust emission. `Cigale` does not have a single model including both the gray-body thermal emission and the PAH emission as `Magphys`. The choice is limited to the models by Draine & Li 2007 and their updated version by Draine et al. 2014 parameterized through the intensity of the radiation field and including the PAHs emission. A second possibility would be the analytical model by Casey (2012) parameterized through the dust temperature but not including the PAHs. Since our photometric catalog includes a point at 24μm from the SuperDeblended catalog, sampling the typical PAHs emission at $z \sim 3$ and since this feature is generally crucial for determining a robust redshift, we decide to employ the Draine et al. (2014) model. We compute the mass-weighted dust temperature starting from the intensity of the radiation field reported in output by `Cigale` following the framework described in Draine & Li (2007) and Draine et al. (2014).

Table 2.2: Comparison between the median properties estimated by `Magphys` and `Cigale`

Property	`Magphys`		`Cigale`		Unit
	Median	σ	Median	σ	
$z_{\rm phot}$	2.96 ± 0.04	1.2	3.02 ± 0.04	1.3	
$\log(M_\star)$	11.01 ± 0.02	0.61	11.01 ± 0.02	0.44	$M_\odot$
$\log(SFR)$	2.67 ± 0.02	0.48	2.79 ± 0.04	0.45	$M_\odot \mathrm{yr}^{-1}$
A_v	4.01 ± 0.04	1.3	3.89 ± 0.05	0.31	mag
$\log(L_{\rm Dust})$	12.45 ± 0.02	0.48	12.59 ± 0.02	0.44	$L_\odot$
$T^L_{\rm Dust}$	42.05 ± 0.2	5.8	–		K
$T^M_{\rm Dust}$	–		33.9 ± 0.2	2.7	K
$\log(M_{\rm Dust})$	8.57 ± 0.02	0.5	8.9 ± 0.02	0.46	$M_\odot$

Note: The uncertainties on the median properties are estimated as the Median Absolute Deviation: $MAD = 1.482 \times \mathrm{median}(|x_i - \mathrm{median}(x_i)|)$ (Hoaglin et al., 1983) divided by $\sqrt{N}$, where N is the number of galaxies in the sample. For each quantity, we also report the dispersion computed as half the symmetrized interval between the 16th and the 84th percentiles.

2.5.3 SED-fitting results and comparison between the codes

The two codes are applied to the photometric catalog presented in Section 2.4. Before running the codes, we correct the photometry in the catalog for the galactic extinction. We employ the dust maps by Lenz et al. (2017) and the extinction law by Fitzpatrick & Massa (2007). This correction is performed through the python libraries `DustMaps` (Green, 2018) and `Extinction` (Barbary, 2016b). To account for the possible biases in the photometry extraction (see Section 2.4.2), to account for the known under-estimation of the photometric uncertainties by the `Photutils` library employed in `PhoEBO` (see e.g Leauthaud et al., 2007; Laigle et al., 2016; Weaver et al., 2022) and to allow the SED-fitting codes to explore a wider region in the photometry space, we add in quadrature 0.15 mag to the photometric uncertainties included in the catalog for the optical/NIR/MIR bands, following Laigle et al. (2016) and Weaver et al. (2022). The output of the SED-fitting codes `Magphys` and `Cigale` are summarized in Figure 2.4, while the median values of the photo-zs and the main physical properties for both codes are reported in Table 2.2. In generating the histograms and in computing the medians, we exclude from the sample 57 sources that could host an AGN (see the discussion in Section 2.6), in order to avoid possible biases coming from the SED-fitting performed with templates not including nuclear activity. Similarly, we do not include in our analysis 10 sources with no other robust detections than in the radio (the so-called "type 0" of Behiri et al. 2023, see that study for the possible models of these sources) for which the properties estimated through SED-fitting would be highly unreliable. The good convergence of the SED-fitting procedure is ensured by the median reduced $\chi^2 < 2.5$ obtained by both codes on the analyzed sample. Since for our galaxies the constraints on

stellar population in the optical/NIR regimes are limited — in the majority of the cases — to upper limits, we compute the SFR starting from the infrared luminosity through the relation by Kennicutt & Evans (2012) rescaled to a Chabrier (2003) IMF. This quantity is more robustly estimated, since $\sim 60\%$ of the galaxies in the sample have a detection at FIR/(sub)mm wavelengths. However, we must underline how this quantity is more unconstrained for the other $\sim$40% of the sources in the sample without these detections. Nevertheless, since the two codes employed in the SED-fitting procedure rely on the energy balance and include the radio fluxes, some constraints on the infrared luminosity can also come by the modeled dust attenuation and by the radio luminosity (da Cunha et al., 2015; Battisti et al., 2019; Boquien et al., 2019). Finally, we notice that no significant discrepancy is visible when comparing the SFR estimated through the $L_{\rm IR}$ and the $L_{1.4\rm GHz}$ (see e.g. Kennicutt & Evans 2012; Novak et al. 2017) for the galaxies with FIR/(sub)mm detections and those without. This result can be partly explained by the fact that `Magphys` assumes a constant radio-infrared correlation with a $q_{\rm TIR}$ centered on 2.34 and with a 1σ dispersion of 0.25 (da Cunha et al., 2015; Battisti et al., 2019). This value is broadly compatible with the median $q_{TIR} = 2.1 \pm 0.3$ computed for the galaxies in our sample (see Section 2.6.3).

As it can be seen, the results of the two software for what concerns the photo-zs, and most of the physical properties are broadly compatible. The slight difference between the outputs for these quantities can be explained by some minor differences between the codes (e.g. the treatment of the upper limits, representing the majority of the constraints in the optical/NIR regime for our galaxies; see e.g. Battisti et al. 2019 and Boquien et al. 2019). Major differences hold for the dust mass and temperature. As discussed in Sections 2.5.1 and 2.5.2, `Magphys` and `Cigale` report the luminosity- ($T^{L}_{\rm Dust}$) and mass-weighted ($T^{M}_{\rm Dust}$) dust temperature, respectively. Since these quantities weight differently the hot and cold components of the dust (see e.g. Liang et al., 2019; Sommovigo et al., 2020), it is not surprising the difference in the two distributions reported in Figure 2.4. Moreover, this difference in the two codes can also explain the discrepancy in the two estimates of the dust mass, being this quantity computed starting from the dust temperature. In the following, for consistency with the previous papers on analogous source (Talia et al., 2021; Behiri et al., 2023), we will consider the results obtained by `Magphys`. A final note concerns the accuracy of our photo-zs. Unfortunately, due to the elusive nature of the sources in our sample, it is not easy to retrieve spectroscopic redshifts for our RS-NIRdark galaxies from the current literature. One spectroscopic redshift can be found in Jin et al. (2022), where they targeted one of our galaxies with a spectral scan performed with the ALMA and NOEMA interferometers. For our galaxy RSN-436 (ID 3117 in Jin et al. 2022), they obtained $z_{\rm spec} = 3.545$, which is in good agreement with our estimates ($z_{\rm M} = 3.3 \pm 0.3$ and $z_{\rm C} = 3.4 \pm 0.2$). A significant improvement of the spectroscopic coverage of our sample is the follow-up program performed with ALMA described in Chapter 3.

2.6 AGN contribution

While in the Introduction to this thesis we discussed the possible biases affecting the FIR and (sub)mm selection of DSFGs, this section focuses on the main bias

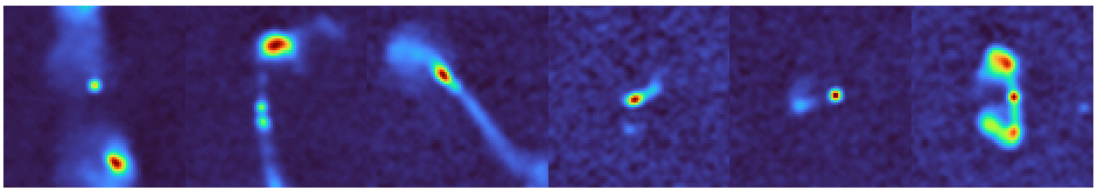

Figure 2.5: Some examples of the sources marked as likely AGN in the final catalog because of their morphology during the visual inspection of the radio maps at 3GHz. The postage have a 15 arcsec side. Further details are given in Section 2.6.

possibly affecting our radio selection: the presence of a significant contribution of nuclear activity by Active Galactic Nuclei (AGN) to the radio luminosity of our galaxies. For instance, Bonzini et al. (2013); Novak et al. (2017) and Enia et al. (2022), analyzing samples of radio-detected galaxies selected at 1.4 GHz, reported significant fractions of AGN in their catalogs, spanning from 10% (for the faintest radio fluxes) to 100% (for the radio-brightest sources). Fortunately, our selection procedure focuses on galaxies without a significant optical/NIR counterpart. This step allows us to remove from the sample all the brightest AGN. This selection, however, does not allow us to remove from the sample the obscured AGN, where a significant amount of dust absorbs the optical/NIR emission (see e.g. the review by Hickox & Alexander 2018 and references therein). The presence of these sources in our sample can be unveiled by searching for AGN tracers at longer wavelengths (namely the IR and radio regimes), or in the X-ray, taking advantage of the broad wavelength coverage of our sample offered by the several facilities that observed the COSMOS field in the last decades.

Estimating the fraction of obscured AGN and properly accounting for their presence is crucial for two main reasons. Firstly, the presence of an extra IR component due to a dusty torus surrounding the AGN and/or a radio excess due to nuclear activity can bias the SFR when obtained from the IR/radio luminosity through the relation by Kennicutt & Evans (2012). Similarly, the employment of galaxy-only templates could bias all the other physical properties (and the photo-zs) estimated through SED-fitting. Secondly, the contamination affecting the selection procedure has to be considered in the determination of the statistical properties of the sample (e.g. the luminosity function and the contribution to the SFRD; see Chapter 4).

2.6.1 Visual inspection

The first selection of galaxies hosting a radio-powerful AGN in our sample is performed through a visual inspection of the 3GHz radio maps. This procedure aims at selecting all the galaxies with radio morphologies suggesting the presence of AGN (e.g. radio blobs and relativistic jets). This visual inspection allows us to mark as likely AGN 17 galaxies. Some examples of these sources are reported in Figure 2.5. We underline that — although this selection is 100% pure (since non-active galaxies do not have these peculiar radio morphologies) — it is rather far from complete since most of the galaxies in the sample are unresolved at the 0.7" resolution of the 3GHz maps by Smolčić et al. (2017). As an additional check, 9 RS-NIRdark

galaxies are also part of the catalog by Vardoulaki et al. (2021) containing AGN in the COSMOS field selected through visual inspection thanks to their morphology in the same radio maps employed for our selection. Cross-matching our sample with this catalog, we find out that all the 9 sources are classified by AGN by both the selections. This results strengthen the reliability of our classification.

2.6.2 X-ray stacking

Another test to unveil the presence of AGN in our sample can be performed in the X-ray regime. We conduct two complementary tests to characterize the sample of RS-NIRdark galaxies as a whole and the individual sources.

- **Sample characterization:** The first test is based on the median X-ray flux computed on all the galaxies in our sample and on the possibility of explaining it by only invoking the star formation. We follow this procedure:

  1. We extract the X-ray flux for each RS-NIRdark galaxy through aperture photometry on the event file in the range [0.5,7] keV produced by the Chandra telescope in the C-COSMOS (Elvis et al., 2009) and COSMOS legacy (Civano et al., 2016) surveys. We perform the extraction through the `dmextract` function of the `CIAO` library (Fruscione et al., 2006), employing circular apertures centered on the radio position of each source. We choose the radius of each aperture by assuming that each source is not resolved (a reasonable hypothesis for possible AGN) and employing the radius used in Elvis et al. (2009) and Civano et al. (2016) to extract the flux of the X-ray source closest to the considered galaxy. More in detail, we employ the median radius (considering all the observations in which the source fell) corresponding to an area encompassing 90% of the PSF at the X-ray closest source. The local background is computed and subtracted from each galaxy through an annulus with an outer radius 1.5 times the circular aperture.

  2. Once the net counts are obtained, these are converted to X-ray luminosities through the `PIMMS`[4] software by assuming a power-law model with a slope of $\Gamma = 1.8$ (i.e. that expected for possible AGN) and a galactic $n_H = 1.7 \times 10^{20}$ cm^{-2} (Civano et al., 2016).

  3. Finally, we convert the luminosities in fluxes by assuming the photometric redshifts computed by `Magphys` (Section 2.5) and these in SFR through the empiric relation by Ranalli et al. (2003).

  Considering the median SFR obtained by these X-ray fluxes, we obtain a value of $\log(SFR_X) = (2.24 \pm 0.02)\ M_\odot\ \mathrm{yr}^{-1}$, slightly lower than that obtained through the FIR flux ($\log(SFR_{IR}) = (2.38 \pm 0.02)\ M_\odot\ \mathrm{yr}^{-1}$. This result confirms the lack of strong un-obscured AGN activity in our sample.

- **Source characterization:** If the analysis performed in the previous paragraph ensures that the bulk of the sample of RS-NIRdark is composed of SFGs, it does not give insights on the presence of single AGN within the total sample. To address this point, we perform an additional analysis. We cross-match

[4] https://heasarc.gsfc.nasa.gov/docs/software/tools/pimms.html

the two catalogs of X-ray sources in the COSMOS field by Elvis et al. (2009) and Civano et al. (2016) with our sample by employing as matching radii the positional uncertainty included in these catalogs (for the X-ray sources) and 0.7" (for the NIRdark galaxies). Given the relatively shallow depth of the X-ray coverage in the COSMOS field, all these sources have an X-ray luminosity higher than 10^{42} erg s^{-1}, therefore, we can assume that their X-ray flux is largely due to the presence of an AGN (see e.g. Hickox & Alexander 2018). Once the list of RS-NIRdark galaxies with a possible X-ray counterpart is obtained, we visually inspect the NIR and radio maps at 3GHz. We obtain two main cases:

1. The positional uncertainty on the X-ray source includes only the RS-NIRdark galaxy. In this case, we can safely assume that the X-ray signal is produced by the galaxy in our sample, suggesting the presence of an AGN in that galaxy. The 3 sources in this class are marked in the photometric catalog with an appropriate flag.

2. The positional uncertainty of the X-ray source includes both an RS-NIRdark galaxy and a NIR-bright galaxy. In this case, we cannot unambiguously associate the X-ray signal to one of these galaxies. The 15 NIR-dark galaxies in this class are marked in the catalog with a different flag accounting for the possibility of hosting an AGN.

2.6.3 $q_{\rm TIR}$ analysis

A standard method to identify AGN relies on the so-called infrared-radio correlation. It is well established that the radio luminosity measured at 1.4 GHz and the infrared luminosity measured in the range [8;1000] μm are tightly ($\sigma \sim 0.16$ dex; e.g. Molnár et al. 2021) correlated in star-forming galaxies (see e.g. de Jong et al. 1985; Helou et al. 1985). This correlation is generally measured through the $q_{\rm TIR}$ parameter defined as (see e.g. Helou et al. 1985; Yun et al. 2001):

$$q_{\rm TIR} = \log\left(\frac{L_{\rm IR}[W]}{3.75 \times 10^{12}{\rm Hz}}\right) - \log\left(\frac{L_{1.4{\rm GHz}}}{{\rm W\,Hz^{-1}}}\right) \tag{2.3}$$

and mainly arises because of the connection between the star formation, the radio synchrotron emission in star forming regions and the thermal FIR emission of dust in the same regions. Several authors studied the possible evolution of the $q_{\rm TIR}$ parameter with cosmic time or possible correlations with other physical quantities. In this work, we explore two of the main studies on this point. Delhaize et al. (2017) found a possible evolution with the redshift through the relation

$$q_{\rm TIR}(z) = (2.88 \pm 0.03) \times (1+z)^{(-0.19 \pm 0.01)} \tag{2.4}$$

On the contrary, more recently, Delvecchio et al. (2021) suggested a $q_{\rm TIR}$ almost constant with the redshift, but strongly dependent on the stellar mass of the hosting galaxy through the relation

$$q_{\rm TIR}(M_\star, z) = (2.646 \pm 0.024) \times A^{(-0.023 \pm 0.008)} - B \times (0.148 \pm 0.013) \tag{2.5}$$

where $A = (1 + z)$ and $B = (\log(M/M_{\odot}) - 10)$.

We compute the radio luminosity at 1.4 GHz for the galaxies with a counterpart in the 1.4 GHz catalog by Schinnerer et al. (2010) through the relation

$$L_{1.4\mathrm{GHz}} = \frac{4\pi D_L^2(z)}{(1+z)^{1+\alpha}} S_{1.4\mathrm{GHz}} \tag{2.6}$$

by employing the photometric redshifts estimated by `Magphys` and the spectral slope α computed through the flux densities reported in the 3 GHz and 1.4 GHz catalogs. For the galaxies without a counterpart in the 1.4 GHz sample, we evaluate the 1.4 GHz flux starting from the 3GHz flux density through the relation

$$L_{1.4\mathrm{GHz}} = \frac{4\pi D_L^2(z)}{(1+z)^{1+\alpha}} \left(\frac{1.4\mathrm{GHz}}{3\mathrm{GHz}}\right)^{\alpha} S_{3\mathrm{GHz}} \tag{2.7}$$

and assuming a spectral slope of $\alpha = -0.7$ (generally considered for star-forming galaxies; see e.g. Novak et al. 2017).

Once we applied Equations 2.4 and 2.5 to our sample (Figure 2.6), marking as likely AGN the sources distant more than 3σ from the relations (see e.g. Delvecchio et al., 2017; Enia et al., 2022), we obtain that 40 sources are classified as AGN following Delhaize et al. (2017) and 57 following Delvecchio et al. (2021). We include the results from both the tests in the final catalog. However, given the high uncertainties affecting our photo-zs and our stellar masses, in the following we consider as likely AGNs only the 37 sources classified as AGN according to both the relations. A final interesting remark concerns the overall distribution of the q_{TIR} as a function of the photometric redshift. As visible in Figure 2.6, the bulk of the population of the RS-NIRdark galaxies have a median q_{TIR} compatible with that expected at the median redshift of the sample from the relation by Delhaize et al. (2017) (2.1 ± 0.3 and 2.21 ± 0.05, respectively). However, the distribution of these values at low redshift appear to significantly differ from that expected from the relation. This discrepancy can be explained with the different selection operated in this study with respect to Delhaize et al. (2017). In particular, our selection of NIR-dark sources is expected to produce a sample of more extremely-obscured sources at low-z to account for the extinction of redder rest-frame bands. Therefore, we do not expect in this regime the properties of our (incomplete) sample to completely resemble those of the total population of radio-selected galaxies analysed in Delhaize et al. (2017).

2.6.4 SED decomposition

A further sign of AGN hosted in our galaxies can be found at MIR wavelengths. It arises from the presence of a dusty torus surrounding the supermassive black hole and heated up from the high-energy radiation coming from there. Trying to model the SED with a template not accounting for this additional component generally produces a best-fitting SED under-estimating all the MIR fluxes (see e.g. Hickox & Alexander, 2018). This problem can be solved by adding a torus component to the templates fitted to the galaxy photometry. We perform this test with `Cigale` (`Magphys` does not allow the addition of an AGN component). The overall setup is the same as discussed in the Section 2.5.2, but adding the dusty torus component as

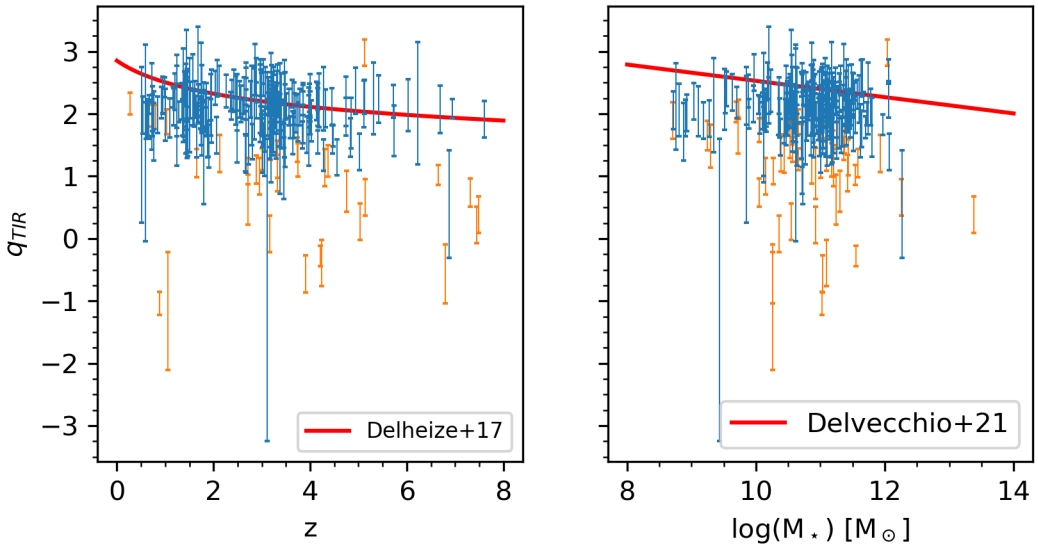

Figure 2.6: Behavior of the $q_{\rm TIR}$ as a function of the redshift (left panel) and stellar mass (right panel). The blue solid lines show the relation by Delhaize et al. (2017) and Delvecchio et al. (2021), respectively. The objects distant more than 3σ from each relation are highlighted in red.

modeled by Fritz et al. (2006). The model's parameters are the same as employed in Donevski et al. (2020) in modeling a sample of likely DSFGs. The main parameter describing the effect of the AGN on the modeled SED is the AGN fraction $f_{\rm AGN}$, defined as the ratio between the torus luminosity and the dust luminosity in the range $[5, 40]$ μm. Defining as likely AGN all the galaxies with a $f_{\rm AGN} > 10\%$, we mark 11 sources among the entire sample of RS-NIRdark galaxies. We underline that the median value of $f_{\rm AGN}$ computed on the whole sample is compatible with zero, suggesting an overall small contribution of AGN in the RS-NIRdark galaxies. However, we underline that — due to the limited coverage of the MIR regime in our sample — also this estimation should be considered as a lower limit on the actual AGN contribution in our galaxies.

2.6.5 Final remarks on AGN contamination

The sample of likely AGN reported by the different methods have some overlap, as shown in Figure 2.7. Considering all the galaxies marked as possible AGN by at least one method, we obtain a sample of 64 sources ($\sim 23\%$ of the full sample). Excluding the sources with an uncertain X-ray counterpart (i.e. those at point 2 in Section 2.6.2) and with no other indication of AGN activity from the other methods, we obtain 57 sources ($\sim 20\%$ of the sample).

As a final remark, we can compare the estimated fraction of RS-NIRdark galaxies hosting an AGN with a theoretical prediction based on analogous works present in the literature. Bonzini et al. (2013) and Novak et al. (2018), working on radio-selected catalogs of galaxies, reported in their work the fraction of AGN in the analysis of their sources as a function of the radio flux at 1.4 GHz. Considering these

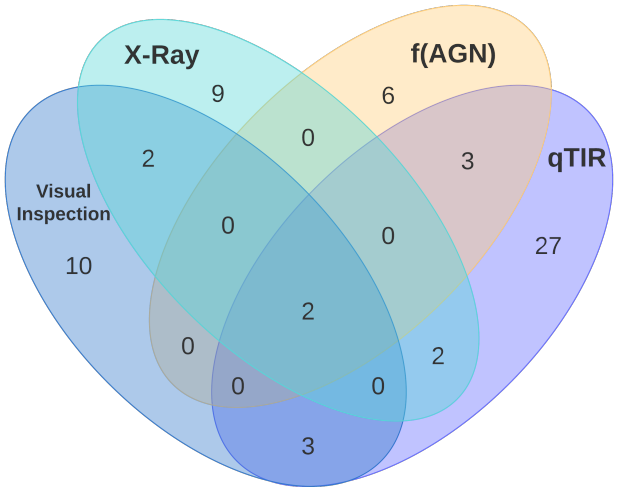

Figure 2.7: Diagram reporting the number of likely AGN unveiled by the different methods described in Section 2.6 and the relative overlap. Further details in Section 2.6.

relations and integrating them on our[5] flux distribution at 1.4 GHz, we estimated an expected fraction of AGN in our sample lower than 28.3% (further divided in 27% of radio-quiet AGN and 1.3% of radio-loud AGN). We underline that this last estimation represents an upper limit on the expected number of AGN in our sample, since we expect that focusing on radio sources without an optical/NIR counterpart contributes to exclude a significant fraction of these objects.

2.7 Discussion

2.7.1 Analysis of the physical properties

The results shown in Figure 2.4 and Table 2.2, allow us to infer some general properties of the RS-NIRdark galaxies:

- The high value of the median stellar attenuation $A_V \sim 4$ mag confirms the initial hypothesis that our selection can provide a sample of highly-obscured galaxies.
- The high value of the infrared luminosity ($L_{IR} > 10^{12} L_{\odot}$) and — hence — the high median SFR (~ 450 $M_{\odot} yr^{-1}$) suggests that our galaxies are actively forming stars (as expected from a radio selection). To determine the nature of the RS-NIRdark galaxies, we need to compare the SFRs obtained through

[5]Here, we employ the radio fluxes at 1.4GHz provided by Schinnerer et al. (2010). For the sources without a counterpart we conservatively assume a flux equal to the sensitivity of the survey

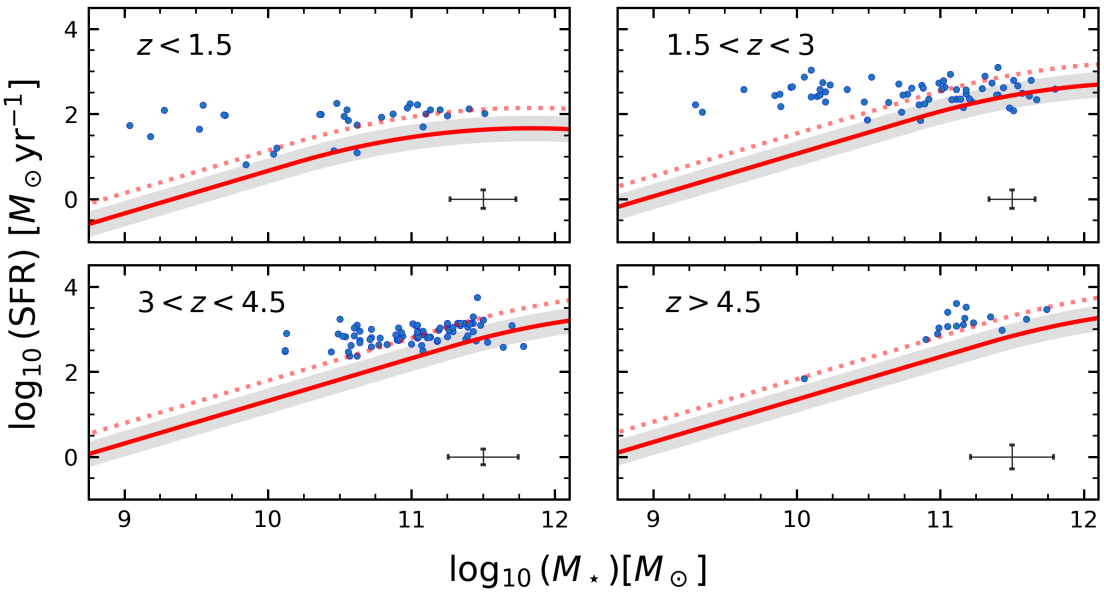

Figure 2.8: Comparison between the SFRs and stellar masses of the RS-NIRdark galaxies (as computed by `Magphys`) and those expected from the main sequence of the star-forming galaxies (red solid line; Schreiber et al. 2015 rescaled to a Chabrier 2003 IMF). The gray shaded area represents the $1\sigma = 0.3$ dex scatter, while the red dotted line reports our threshold for selecting star-burst galaxies (i.e. three times the SFR expected from MS galaxies). It is possible to notice how the vast majority of the galaxies in the sample lie above the main-sequence line. The median uncertainty on the SFR and on the stellar mass is reported in the lower corner of each panel. Further details are given in Section 2.7.1.

the SED-fitting with those expected from main sequence galaxies in the same redshift range (e.g. Schreiber et al., 2015). As shown in Figure 2.8, most galaxies lie above the main sequence line at all the redshifts. Moreover, when computing the SFR/SFR_{MS} (i.e. the ratio between the SFR and that expected from a main sequence galaxy of the same mass and in the same redshift bin), we obtain that more than 50% of the RS-NIRdark galaxies have a $SFR/SFR_{MS} > 3$, being candidate starburst galaxies.

- An additional result, strictly connected to the previous one, concerns the high median stellar mass estimated through SED-fitting ($M_\star \sim 10^{11} M_\odot$). This quantity — albeit quite uncertain due to the weak constraints in the optical/NIR wavelengths — suggest that the RS-NIRdark galaxies are a population of massive star-forming galaxies, with a high-z tail suitable for playing a significant role in the evolution of the massive and passive galaxies at $z \sim 3$ (see e.g. Straatman et al., 2014; Schreiber et al., 2018a; Valentino et al., 2020a)

All these findings, once combined, confirm the initial hypothesis that the RS-NIRdark galaxies represent a significant population of high-z DSFGs.

2.7.2 High-z sources

The redshift distribution of the RS-NIRdark galaxies is worth a deeper discussion. The median redshift around $z \sim 3$ tells us that we are looking at a population whose bulk is located at the so-called *cosmic noon*. However, the presence of a significant tail of high-z sources (namely, 99 galaxies at $z > 3$ and 17 galaxies at $z > 4.5$, once excluded the possible AGNs selected in Section 2.6.) can provide some insights on the possible evolutionary path of these sources.

The main result concerns the number density of these sources. We compute this quantity through the $V_{\rm max}$ method (Schmidt, 1968), considering the galaxies located at $z > 3.5$ and that could — therefore — play a role in the evolution of the massive galaxies at $z \sim 3.5$. We account for the uncertainties in the photo-zs and in the radio fluxes through a MonteCarlo integration. Specifically, we perform a large number of realizations (~ 500) of the total number density, sampling each time and for each source a couple of values for the redshift and for the radio flux from their distribution (namely, the $p(z)$ computed by `Magphys` and the values and uncertainties reported in Smolčić et al. 2017). At the end of this procedure, we consider as our number density the median value of the distribution and as its uncertainty the symmetrized interval between the 16th and the 84th percentiles. We obtain a number density of $n = (3.3 \pm 0.9) \times 10^{-6}$ Mpc^{-3} for the galaxies at $z > 3.5$. This quantity is by a factor 6 lower than those computed by Straatman et al. (2014), Schreiber et al. (2018a), and Valentino et al. (2020a) for the passive and massive galaxies at $z \sim 3.5$ ($\sim 2 \times 10^{-5}$ Mpc^{-3}). It is important to notice — however — that the our estimation must be considered as a lower limit on the actual number density of the RS-NIRdark galaxies, since it is not corrected for the expected *duty cycle* of the galaxies and for the incompleteness of the selection. This issue will be discussed in detail in Chapter 4. Moreover, when looking at the stellar masses of our sources, we obtain that the RS-NIRdark galaxies located at $z > 3.5$ have a median stellar mass of $\log(M_\star) = 11.0 M_\odot$ with a 1σ dispersion of 0.45 dex. Comparing this quantity

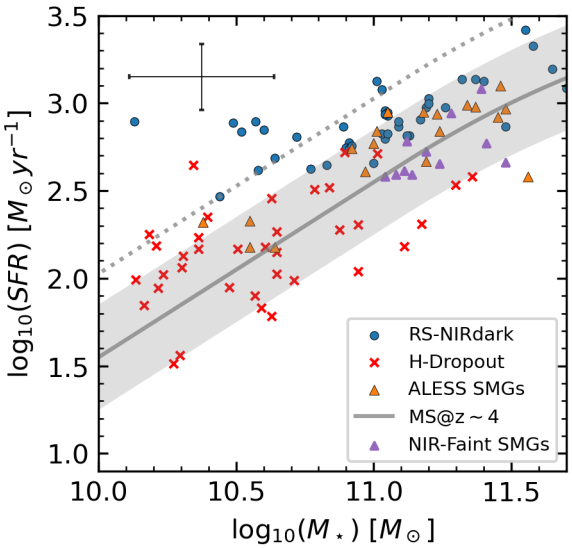

Figure 2.9: Comparison between the SFRs and stellar masses of several populations of DSFGs. The different symbols show the RS-NIRdark galaxies (blue dots), the H-dropout (Wang et al. 2019; red crosses), the SMGs found in ALESS (da Cunha et al. 2015; orange triangles) and the NIR-faint SMGs discovered by Smail et al. (2021) in the UKIDSS Ultra Deep Survey (purple triangles). The solid gray line represents the main sequence at $z \sim 4$ (Schreiber et al. 2015 rescaled to a Chabrier 2003 IMF) and the gray shaded area its $1\sigma = 0.3$ dex scatter. The dotted gray line reports our threshold to define starburst galaxies (i.e. three times the SFR of a main sequence galaxy). The median uncertainty on the SFR and on the stellar mass is reported in the lower corner of each panel. For consistency with the cited studies, we only report galaxies with $3.5 < z < 4.5$. Further details are given in Section 2.7.1

.

with the expected properties of the progenitors of the massive galaxies at $z \sim 3.5$ (see e.g. the forecasts by Valentino et al. 2020a), we can notice how the low-mass end of those objects cannot be formed by the galaxies in our sample. This result also can be used to explain the difference between the number densities of the two populations.

2.7.3 Comparison with the literature

An interesting final point to discuss concerns the RS-NIRdark galaxies and their possible overlap with the other notable populations of DSFGs described in the Introduction to this Thesis:

- **H-dropout:** The first population is composed of the H-dropout galaxies (Wang et al., 2019). These galaxies are selected as H-dark sources in the CANDELS survey (the limiting magnitude at 5σ in the H-band is 27mag; Grogin et al. 2011; Koekemoer et al. 2011) with a counterpart in the second channel of the IRAC camera from the SEDS survey ([4.5]<24 mag; 80% completeness limit; Ashby et al. 2013). When considering the full photometric

catalog of 323 RS-NIRdark galaxies, we obtain that only 266 sources satisfy the selection criteria exposed in Wang et al. (2019), the others being too faint at 4.5 μm to be selected with the cut on the [4.5] flux performed by Wang et al. (2019). This result indicates that a non-negligible fraction ($\sim$ 20%) of the RS-NIRdark galaxies would not be selected as H-dropout. On the contrary, when examining the 18 H-dropout galaxies selected by Wang et al. (2019) in the COSMOS field, we find that only 2 sources have a significant ($S/N > 5.5$) radio counterpart at 3 GHz in the catalog by Smolčić et al. (2017). These two results, when combined, ensures that the two selections of RS-NIRdark galaxies and H-dropout are different, with just some sources belonging to both.

- **SMGs:** The second population is composed of the so-called sub-millimeter galaxies (SMGs). Since this definition can be applied to all the sources detected in a (sub)mm survey, it strongly depends on the considered instrument's sensitivity. To obtain results comparable with those reported by Wang et al. (2019), we consider the ALESS survey (Swinbank et al., 2014) targeting bright galaxies with the SCUBA bolometer (i.e. galaxies with a flux density at 870μm $S_{870} > 4.2$ mJy). We can estimate the (sub)mm flux density of the RS-NIRdark galaxies at 870μm through the best-fitting SED provided by `Magphys`. In doing so, we obtain that only 76 galaxies have $S_{870} > 4.2$ mJy, being consistent with the selection by Swinbank et al. (2014). On the other side, Thomson et al. (2014) pointed out that only $\sim$70% of SMGs are radio-bright galaxies, while Gruppioni et al. (2020) pointed out that only a tiny percentage of SMGs are NIR-dark galaxies. As before, these results suggest that the RS-NIRdark galaxies and SMGs are two different populations of galaxies with just some sources in common.

Finally, we can compare the physical properties estimated by Wang et al. (2019) and da Cunha et al. (2015) for the population of H-dropout and SMGs with those presented in this study. Figure 2.9 shows the comparison between the SFR and stellar mass of the RS-NIRdark galaxies in the redshift range $3.5 < z < 4.5$ (i.e. around the median redshift $z \sim 4$ reported by Wang et al. 2019 for the H-dropout) with those reported by da Cunha et al. (2015) and Wang et al. (2019) and with the main sequence at this redshift (Schreiber et al. 2018b rescaled to a Chabrier 2003 IMF). For completeness, we also include in the plot the sample of "NIR-faint" ($Ks > 23.9$) SMGs selected in COSMOS by Smail et al. (2021).

We can notice that the RS-NIRdark galaxies are — on average — more star-forming than the H-Dropout ($\Delta \log(\mathrm{SFR}/\mathrm{M}_\odot\ \mathrm{yr}^{-1})\sim0.6$), with a median SFR comparable with the SMGs ($\Delta \log(\mathrm{SFR}/\mathrm{M}_\odot\ \mathrm{yr}^{-1})\sim0.02$), but with a significant tail of less star-forming sources. This result can be explained by the radio-selection, cutting out most of the sources below the main sequence (belonging mainly to the H-dropout). Additionally, we can see how the RS-NIRdark galaxies cover most of the mass range covered by the other selections, being on average more massive than the H-dropout and less massive than the SMGs. This result is in agreement with the discussed overlap with both the other populations.

2.8 Chapter summary

In this chapter, I presented the first panchromatic study of 263 Radio-Selected NIRdark galaxies discovered in the COSMOS field updating the selection by Talia et al. (2021).

- The development of a new deblending tool (`PhoEBO`: Photometry Extractor for Blended Objects) allowed my collaborators and me to extract accurate photometry in the optical-to-MIR regime, even for the sources with a close bright contaminant. This procedure, in particular, allowed us to analyze a wider sample of galaxies missed in previous studies on the RS-NIRdark galaxies in the COSMOS field (Talia et al., 2021; Behiri et al., 2023).

- The complete photometric catalog[6] has been employed to estimate the photo-zs and physical properties of all the galaxies in the sample through an SED-fitting procedure performed with two complementary codes (`Magphys` and `Cigale`). The results obtained with these algorithms confirmed the initial hypothesis that the RS-NIRdark galaxies are a population of starburst DSFGs, lying above the main sequence in all the redshift bins and with a significant amount of dust absorbing their optical/NIR emission.

- Moreover, by studying in detail the redshift distribution of the galaxies in the sample and their number density, we obtain precious insights on the possible evolutionary path of these sources, collecting significant clues that the RS-NIRdark galaxies could play a key role in the evolution of the massive and passive galaxies discovered at $z \sim 3$.

- In addition, the analysis of the multi-wavelength counterpart in all the wavelength regimes (from the X-rays to the radio) allowed us to estimate the possible AGN contribution in our sample, obtaining that nuclear activity could contribute to the radio luminosity of $\sim 20\%$ of our sources.

- Finally, through a comparison with other populations of DSFGs, we confirmed that the radio-selection produces a population of galaxies with different physical properties with respect to the SMGs (e.g. da Cunha et al., 2015) and the H-Dropout (Wang et al., 2019).

[6]The complete catalog, together with the other materials supporting the findings described in this chapter are available on the website of the collaboration: https://sites.google.com/inaf.it/rsnirdark/

Chapter 3

A radio selection for DSFGs: Insights from ALMA

3.1 Introduction to the chapter

In this chapter, I present the first follow-up at (sub)mm wavelengths for a pilot sample of the RS-NIRdark presented in the previous chapter. It is structured as follows. In Section 3.2, I introduce my targets, and the new ALMA observations (the ancillary photometry is the same as presented in the previous chapter). In Section 3.3, I describe the analysis of the ALMA cubes, the identification of any bright emission line in the targets, and my modeling of the spectroscopic redshifts. Moreover, I present some initial insights into the ISM kinematics and derive the physical properties of this pilot sample of galaxies through SED fitting. Then, in Section 3.4, I discuss my results, estimate the gas mass in these sources, and forecast a possible evolutionary path for them. Finally, I draw the main conclusions of the chapter in Section 3.5.

Most of the analysis in this Chapter was also presented in "*Dark progenitors and massive descendants: A first ALMA perspective of radio-selected near-IR-dark galaxies in the COSMOS field*" by Gentile et al. (2024b), published in Astronomy & Astrophysics (Volume 687, id.A288, 18 pp.).

3.2 Data

3.2.1 ALMA observations and data reduction

The main focus of this study is on the observations carried out by ALMA during its cycle 8 as a part of the observing program 2021.1.01467.S (PI: M. Talia). These observations consist of a spectroscopic follow-up at millimeter wavelengths for a pilot sample of nine RS-NIRdark galaxies (Table 3.1). These sources were initially selected in the COSMOS field by Talia et al. (2021) among those located in the high-z tail of the redshift distribution (photo-$z > 4.5$), with the best-sampled SEDs (i.e., with at least one significant detection at $S/N > 3$ at FIR or (sub)mm wavelengths), and with a reliable SED fitting. The selection performed in Talia et al. (2021) is analogous to that described in Section 2.2, but required the non-detection in the

COSMOS2015 catalog (Laigle et al., 2016). The targets with a counterpart in the deeper COSMOS2020 catalog (and therefore not included in the sample analysed in Chapter 2) are highlighted with an appropriate flag in Table 3.1.[1]

The main scientific goal of the observing program was to assess the spectroscopic redshifts of the nine targets. Therefore, the program required a spectral setup covering the whole band 3 of ALMA (i.e., all the frequencies between ∼84 and ∼115 GHz). This setup, analogous to the setups employed in similar studies in the current literature (see, e.g., Walter et al. 2016; Jin et al. 2019, 2022; Cox et al. 2023), ensures that at least one line of the CO and [CI] transitions should be detected for almost all the redshifts in the range $0 < z < 8$. Moreover, this spectral scan provides the possible detection of two lines for most of the redshifts higher than 3, allowing an unambiguous determination of the spec-z (see Figure 3.1). To cover the whole band 3 with ALMA, five settings are required. By estimating the integrated fluxes of the expected CO and [CI] lines observable in our setup, the proposal required a sensitivity of 0.32 mJy/beam per stacked channel for a total of 27h of ALMA observing time. The observations were performed in service mode between March and September 2022, when the interferometer was in its C-4/C-3 configurations (i.e., with baselines between 15 and 500/784 m, an expected beam size of 1.4”/0.92”, and a maximum recoverable scale of 16.2”/11.2”). The calibration was performed by the Alma Regional Center through the ALMA standard pipeline. After the calibrated measurement sets were obtained, I — together with my collaborators — merged the multiple observations through the *Common Astronomy Software Applications* package (`CASA` v6.1; CASA Team et al. 2022). Finally, to achieve the sensitivity originally requested in the proposal, we resampled the native spectral resolution to obtain ∼ 0.02 GHz channels (∼ 50 km/s at the reference frequency of 100 GHz). After a first cleaning (performed with the task `tclean` and employing natural weighting to maximize the sensitivity; see Högbom 1974), we verified that the median rms across the stacked channels was 0.2 mJy/beam (i.e., slightly better than requested in the original proposal), increasing toward higher frequencies due to the decreasing transmissivity of the ALMA band 3 up to 0.4 mJy/beam. The various settings have some overlap in frequency, and therefore, these (overlapping) ranges have a better rms. The beam shape is quite uniform across the channels: it can be modeled as an ellipse with a half-power beam width equal to 1.45” x 1.31” and a position angle of $\sim -70°$.

[1] Since these sources were detected in the COSMOS2020 catalog, they are not part of the sample analyzed in Chapter 2. For these sources, we performed the same analysis (photometry extraction and SED-fitting) as described in that chapter.

Table 3.1: Main observational properties of the nine targets.

ID	ID (Talia+21)	RA [*hh:mm:ss*]	Dec [*dd:mm:ss*]	C20	Ks [mag]	IRAC Ch 2 [mag]
RSN-41	COSMOSVLA3-49	09:58:17.869	+02:30:38.305	-	> 25.7	22.47 ± 0.04
RSN-84	COSMOSVLA3-106	09:58:43.440	+02:45:18.135	-	> 25.8	23.21 ± 0.07
RSN-121	COSMOSVLA3-152	09:59:14.234	+02:35:26.432	✓	25.0 ± 0.3	22.54 ± 0.02
RSN-182	COSMOSVLA3-225	09:59:46.699	+02:48:41.215	-	> 25.0	23.29 ± 0.05
RSN-235	COSMOSVLA3-291	10:00:09.550	+01:42:50.834	-	> 26.2	23.40 ± 0.04
RSN-247	COSMOSVLA3-308	10:00:23.785	+01:41:59.273	✓	24.9 ± 0.3	23.02 ± 0.03
RSN-298	COSMOSVLA3-370	10:00:57.970	+01:38:26.078	-	> 25.7	23.43 ± 0.04
RSN-361	COSMOSVLA3-442	10:01:28.390	+02:21:27.857	✓	25.0 ± 0.3	23.84 ± 0.07
RSN-456	COSMOSVLA3-576	10:02:48.219	+02:24:30.629	-	> 25.7	22.97 ± 0.06

Note: For each galaxy, we report the ID employed in Chapter 2 and throughout this chapter. For completeness, we also report the original ID employed by Talia et al. (2021), which can be used to retrieve the observations from the ALMA science archive. Finally, we report the coordinates (of the radio counterpart, i.e., the counterpart with the higher spatial resolution) and a flag signaling whether each source has a counterpart in the COSMOS2020 catalog (Weaver et al., 2022).

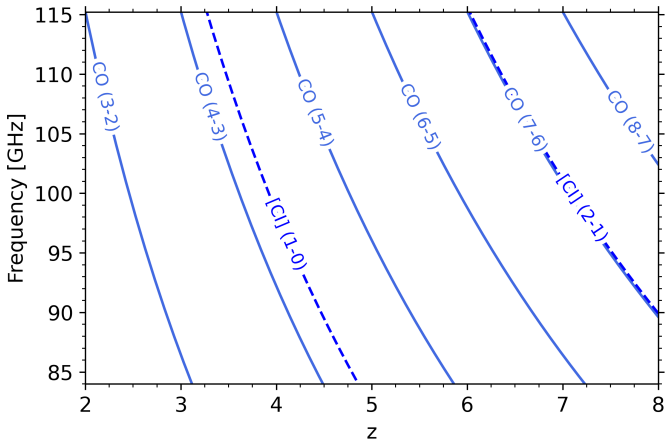

Figure 3.1: Spectral setup adopted for the ALMA observations. This configuration allowed us to observe at least one line of the CO and [CI] transition for all the redshifts below 8. For most of the redshifts higher than 3, two lines are detectable at the observed frequencies.

3.3 Analysis of the datacubes

3.3.1 Continuum images

The first analysis performed on the calibrated MSs consists of the production of a continuum image (see Figure 3.2), which is useful to study the properties of dust in our targets. This procedure is performed through the `CASA` task `tclean` in multi-frequency synthesis (mfs) mode after masking any bright line that could contaminate the continuum emission. To maximize the sensitivity of the cleaned image, we employ a natural weighting throughout this procedure. To estimate the continuum fluxes, we perform aperture photometry with `CARTA` (Comrie et al., 2021), for which we employ an aperture corresponding to the 2σ contour of the continuum image. We verify that this estimate is compatible within the estimated uncertainties with the flux estimated through a 2D profile-fitting performed with the `CASA` task `imfit`. The results of this procedure are reported in Table 3.2. We obtain that six out of nine targets are robustly detected (S/N>3) in the continuum images.

3.3.2 Line identification and reliability

The emission lines inside our datacubes are unveiled through a line-finding algorithm analogous to those employed in several previous studies (e.g., Daddi et al., 2015; Walter et al., 2016; Coogan et al., 2018; Puglisi et al., 2019; Jin et al., 2019, 2022) and summarized here.

1. We obtain a continuum-subtracted MS for each source through the CASA task `uvcontsub`. We model the continuum as a first-grade polynomial whose slope is fit in the whole frequency range covered by our observations after masking any bright line that could contaminate the continuum estimation.

Table 3.2: Continuum fluxes for the RS-NIRdark galaxies analysed in this study.

ID	$S_{3\mathrm{mm}}$ [mJy]	$S_{850\mu\mathrm{m}}$ [mJy]	Other flux [mJy]
41	(0.04 ± 0.01)	(2.9 ± 0.3)	-
84	(0.11 ± 0.03)	$(\mathbf{7 \pm 2})$	0.873 mm: (4.6 ± 0.8)
121	(0.09 ± 0.03)	(3.9 ± 0.4)	-
182	(0.14 ± 0.02)	$(\mathbf{5 \pm 1})$	-
235	(0.18 ± 0.04)	$(\mathbf{5 \pm 1})$	-
247	(0.09 ± 0.01)	$(\mathbf{5 \pm 1})$	-
298	< 0.036	(3.3 ± 0.3)	-
361	< 0.036	(3.3 ± 0.4)	1.249 mm: (2.0 ± 0.4)
456	< 0.036	(1.9 ± 0.2)	-

Note: The values at $\lambda = 3$ mm are obtained through aperture photometry on the continuum images obtained in Section 3.3.1. We consider a source as robustly detected with a S/N>3, and therefore, we report a 3σ upper limit for the undetected targets. The value at 850 μm is retrieved from the SuperDeblended catalog (Jin et al., 2018) for the sources at S/N¿3 (reported in bold), while for the other galaxies, we report the best-fitting flux at 850 μm computed with `Cigale`. For two sources, we also report an additional (sub)mm flux measured with ALMA through cross-matching with the A3COSMOS catalog (Liu et al., 2019).

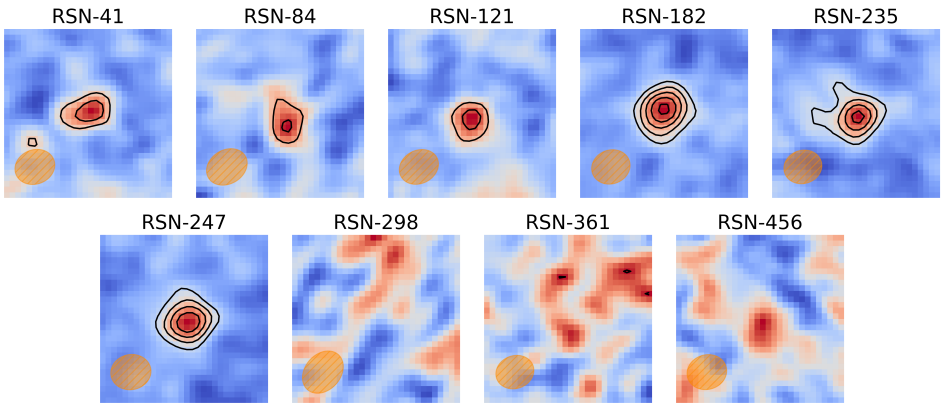

Figure 3.2: Continuum maps of the nine targets. The black contours are in steps of 2σ starting from 3σ. All the images have a 7.5" side, while the synthetized beam is reported in the lower left corner of each image.

2. We perform the imaging of the continuum-subtracted visibilities through the CASA task `tclean`. We employ a natural weighting to maximize the sensitivity of the cleaned images.

3. We convolve the cleaned datacube with a series of boxcar kernels with variable widths between one and 13 channels (i.e., between 60 and 780 km/s at a representative frequency of 100 GHz).

4. For each convolved datacube, we produce an S/N cube by dividing each channel by the relative rms. This quantity is computed through a sigma clipping performed on the inner region of the primary beam to avoid possible biases due to significant emission and higher noise far from the phase center.

5. Finally, we extract an S/N spectrum for each convolved datacube through the Python library `Interferopy` (Boogaard et al., 2021). We employ as extracting region the 2σ contour of the zeroth-moment map obtained on the continuum-subtracted datacube through the CASA task `immoments`. We underline that a significant overlap is visible between this map and the 3GHz one, ensuring that the millimeter emission can be safely associated to our targets.

This procedure results in a list of possible lines, with the related S/N and full width at zero intensity (FWZI)[2]. However, given the nature of the noise in interferometric data, one normally needs to establish the reliability of each line $R = 1-p$ (where p is the probability of a spurious detection). Since several methods exist to compute this quantity for interferometric data, we follow two complementary approaches. Firstly, following Jin et al. (2019), we compute the spurious probability of each line as

$$p(\mathrm{S/N}) = 1 - R_0(\mathrm{S/N})^{N_{\mathrm{Eff}}}, \tag{3.1}$$

where R_0 is the reliability expected in the Gaussian case (which therefore approaches unity toward higher S/N), and N_{Eff} is the number of effective searches. Through an extensive series of simulations, Jin et al. (2019) estimated that this quantity can be approximated through the relation

$$N_{\mathrm{Eff}} \sim 10\frac{N_{\mathrm{total,ch}}}{N_{\mathrm{line,ch}}}N_{\mathrm{Line,ch}}^{0.58}\log\frac{N_{\mathrm{line,ch}}^{\mathrm{max}}}{N_{\mathrm{line,ch}}^{\mathrm{min}}}, \tag{3.2}$$

where $N_{\mathrm{total,ch}}$ is the total number of channels inside the datacube, $N_{\mathrm{line,ch}}$ is the number of channels in which the line is detected, and $N_{\mathrm{line,ch}}^{\mathrm{max}}$ and $N_{\mathrm{line,ch}}^{\mathrm{min}}$ are the minimum and maximum width of the boxcar kernels employed during the line search (see point 4), respectively. While this approach is based on simulations and does not depend on the properties of the actual cube, it relies on the hypothesis that the noise at the phase center of our observations is approximately Gaussian (a reasonable assumption given the almost complete uv coverage generally produced by ALMA). We also compute the reliability of each line following Walter et al. (2016), through

[2]The FWZI is an immediate product of the procedure employed to identify the lines in the extracted spectra since it corresponds to the binning maximizing the S/N of the line. However, in Table 3.3, we report the more common full width at half maximum (FWHM) obtained through the Gaussian modeling described in Section 3.3.4

the `FindClumps` algorithm as implemented in the Python library `Interferopy`. This method estimates the reliability as

$$R(\mathrm{S/N}) = 1 - \frac{N_P(\mathrm{S/N})}{N_N(\mathrm{S/N})} \tag{3.3}$$

where N_P and N_N are the number of positive and negative peaks in the whole datacube in a given S/N bin, respectively. This approach does not rely on any assumption about the nature of the noise in our datacubes, but it could be biased — in principle — by the small statistics affecting the number of pixels in our observations.

Both the procedures described here allow us to identify at least one bright (S/N$>$ 6) line (see Table 3.3) in all targets. The S/N of all the detected lines is high, and we are therefore able to estimate a spurious probability lower than 10^{-6} for all of them, following Jin et al. (2019). Similarly, because their S/N is higher than every negative peak in the analyzed datacubes, we can estimate for all of them a 100% reliability following Walter et al. (2016). By producing the zeroth moment of each (continuum-subtracted) line through the `CASA` task `immoments`, we obtain the maps reported in the insets in Figure 3.3. Moreover, by performing aperture photometry with `CARTA` on the 2σ contour of these maps, we measure the integrated line fluxes reported in Table 3.4.

Table 3.3: Lines detected in our targets following the procedure discussed in Section 3.3.2 and related models (Section 3.3.3).

ID	Line 1	Freq (GHz)	FWHM (km/s)	S/N	Line 2	Freq (GHz)	S/N	$z_{\rm spec}$	$z_{\rm phot}$ (G23)	$z_{\rm phot}$ (T21)
RSN-84	CO(4-3)	102.05	584 ± 96	11.23	[CI](1-0)	108.91	5.39	3.518	4.5 ± 0.3	5.2 ± 0.9
RSN-121	CO(4-3)	106.64	683 ± 77	12.26	[CI](1-0)	113.74	4.89	3.323	3.3 ± 0.2	5.2 ± 0.1
RSN-235	CO(5-4)	101.95	481 ± 63	13.24	[CI](1-0)	87.07	5.68	4.652	4.5 ± 0.5	5.1 ± 0.5
RSN-298	CO(4-3)	89.53	408 ± 93	8.44	CO(5-4)	111.94	4.93	4.150	3.5 ± 0.3	5.4 ± 0.8
RSN-361	CO(5-4)	103.19	349 ± 81	6.30	[CI](1-0)[(a)]	88.13	3.47	4.585	4.3 ± 0.5	5.1 ± 0.9
RSN-41	CO(3-2)	91.50	627 ± 60	12.18	-	-	-	*2.779*	3.0 ± 0.2	6.9 ± 0.1
RSN-182	CO(5-4)	100.31	441 ± 89	9.59	-	-	-	*4.745*	4.6 ± 0.2	5.9 ± 0.9
RSN-247	CO(3-2)	88.02	905 ± 285	6.79	-	-	-	*2.929*	3.4 ± 0.4	5.0 ± 0.1
RSN-456	CO(3-2)	86.74	497 ± 71	7.22	-	-	-	*2.987*	2.8 ± 0.2	6.8 ± 0.7

Note: In the first tier of galaxies, the two lines identified in the spectrum are employed to estimate the spectroscopic redshift. For the galaxies in the second tier (i.e., in those with a single line identified), we assume the spec-z as the redshift allowed by the visible line with the best agreement with the photometry. For each galaxy, we also report the photometric redshift computed by Talia et al. (2021) and that computed through SED fitting with `Cigale` as in Chapter 2 after adding the 3 mm continuum point.

[(a)] Tentative second line; see the discussion in Section 3.3.3

3.3.3 Redshift estimation

After detecting the different lines in our datacubes, we estimate the spectroscopic redshift of our sources following Jin et al. (2019). To do this, we consider the line with the highest S/N in each cube (i.e., the line with the highest reliability) and model it as each of the CO transitions that should be visible in the redshift range $0 < z < 8$ (see Figure 3.1). For all redshifts higher than 3, for most of which a second line is expected, we search for a detection at the expected frequency in the line list produced in Section 3.3.2. Through the S/N of each detection, we compute the reliability of the tentative second lines through Equation 3.1. It is crucial to underline that for the second line, the number of effective searches ($N_{\rm Eff}$) is much lower than those employed in Section 3.3.2. In this case, we do not perform an active search of the line throughout the whole spectrum, but we only analyze the frequencies allowed by the first line. Therefore, the $N_{\rm Eff}$ just corresponds to the number of possible CO transitions with which we can model the first line. For each redshift, we finally estimate a joint spurious probability given by the product between the spurious probability of the first and second line. The redshift with the highest reliability is assumed to be the spectroscopic redshift of our sources.

This approach is sufficient for all the galaxies in which two lines are robustly detected. However, for four of our targets, no second line is detected at a sufficiently high S/N. For these galaxies, we assess the redshift based on additional information from the photometry. After the continuum datapoint at 3 mm (Table 3.2) is added to the photometric catalog presented in Chapter 2 (or an upper limit for the galaxies that are undetected in the continuum images), we perform an SED fitting through the code `Cigale` (Boquien et al. 2019) using the same models and setups presented in Section 3.3.5. We fix the redshifts to all the spec-z allowed by the single line identified in the spectrum. Hence, we assume the redshift with the best agreement between the modeled SED and the photometry (i.e., the redshift with the lowest χ^2) as the final value for the spec-z. It is interesting to notice that — for one of the remaining galaxies (RSN-182) — the redshift estimated through this procedure falls in a small frequency range at $z \sim 4.7$ where a single line (CO(5-4)) is expected. Similarly, three targets (RSN-41, RSN-247, and RSN-456) have a redshift lower than 3, where according to our spectral setup no second line is expected to be observed. We also report a tentative second line in RSN-361 at $\nu = 88.13$ GHz. Even though this detection falls exactly where the [CI](1-0) line would be expected for a galaxy at $z = 4.585$, the low $S/N \sim 3.5$ and the spatial offset with the robustly detected line at $\nu = 103.19$ GHz make the line identification unsure. Finally, we underline that the continuum image (see Figure 3.2) shows a quite irregular morphology for this source, suggesting the a possible major merger (which might explain the spatial offset of the tentative [CI](1-0) line). It is important to note that these redshifts based on a single detected line are clearly more uncertain than those relying on a double detection. We cannot exclude that a fainter line (e.g., a [CI](1-0)) would be observed in our frequency range with deeper observations. For instance, given the redshift estimated through CO(4-3) and CO(5-4) in RSN-298, we would expect a [CI](1-0) line at $\nu = 95.57$ GHz in that source, even though nothing is detected at that frequency at S/N higher than 1σ. For the other galaxies where a single line is detected, however, these solutions would be disfavored by the photometry.

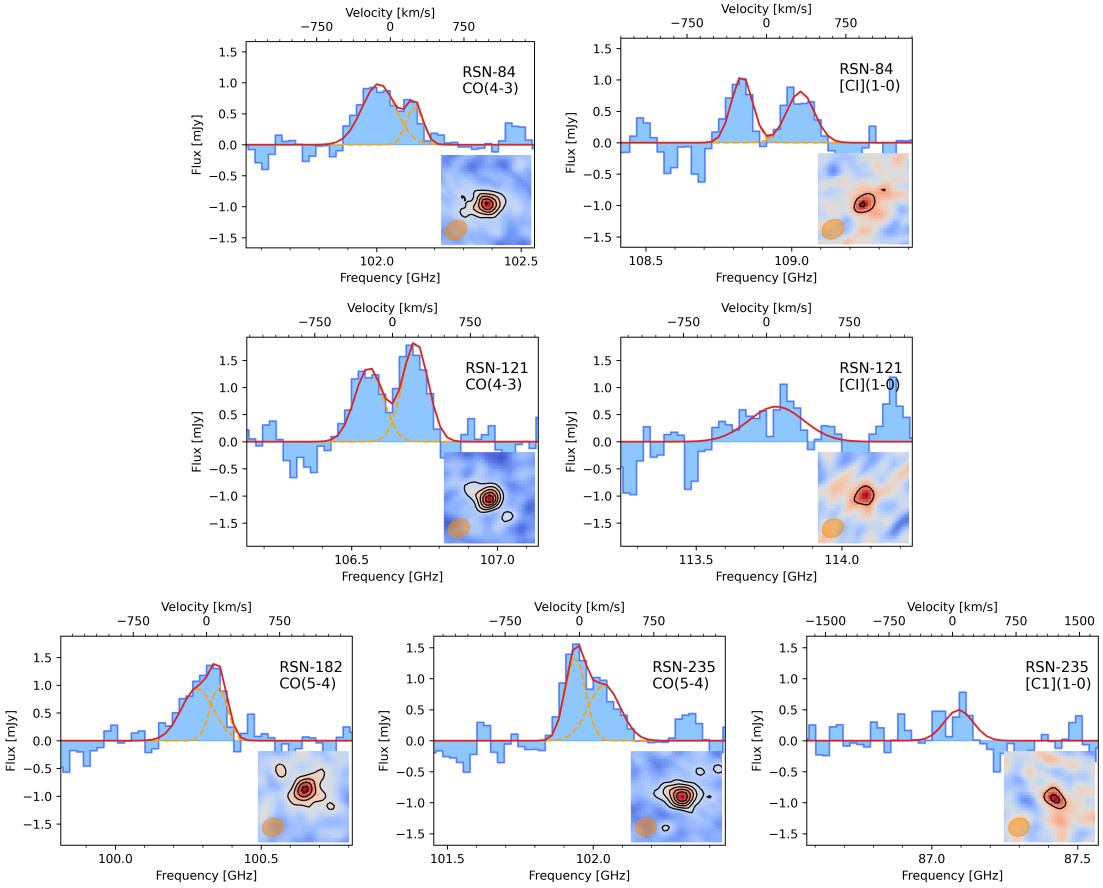

Figure 3.3: Spectrum of the various lines identified through the procedure discussed in Sections 3.3.2 and 3.3.3 in our targets. To increase the visibility of the lines, we resampled the original spectral resolution employed in the study up to $\sim$ 180 km/s. For each line, we report in the upper right corner the ID of the galaxy and our modeling as CO or [CI] transitions. The insets show the moment zero of each line (7.5" side) centered on the radio position measured from the 3 GHz maps, with the contours in steps of 2σ starting from 3σ. In each line, we also report in red the Gaussian modeling with one or two components as described in Section 3.3.4. In the lines modeled with a double Gaussian, we also show the two subcomponents with a dashed orange line.

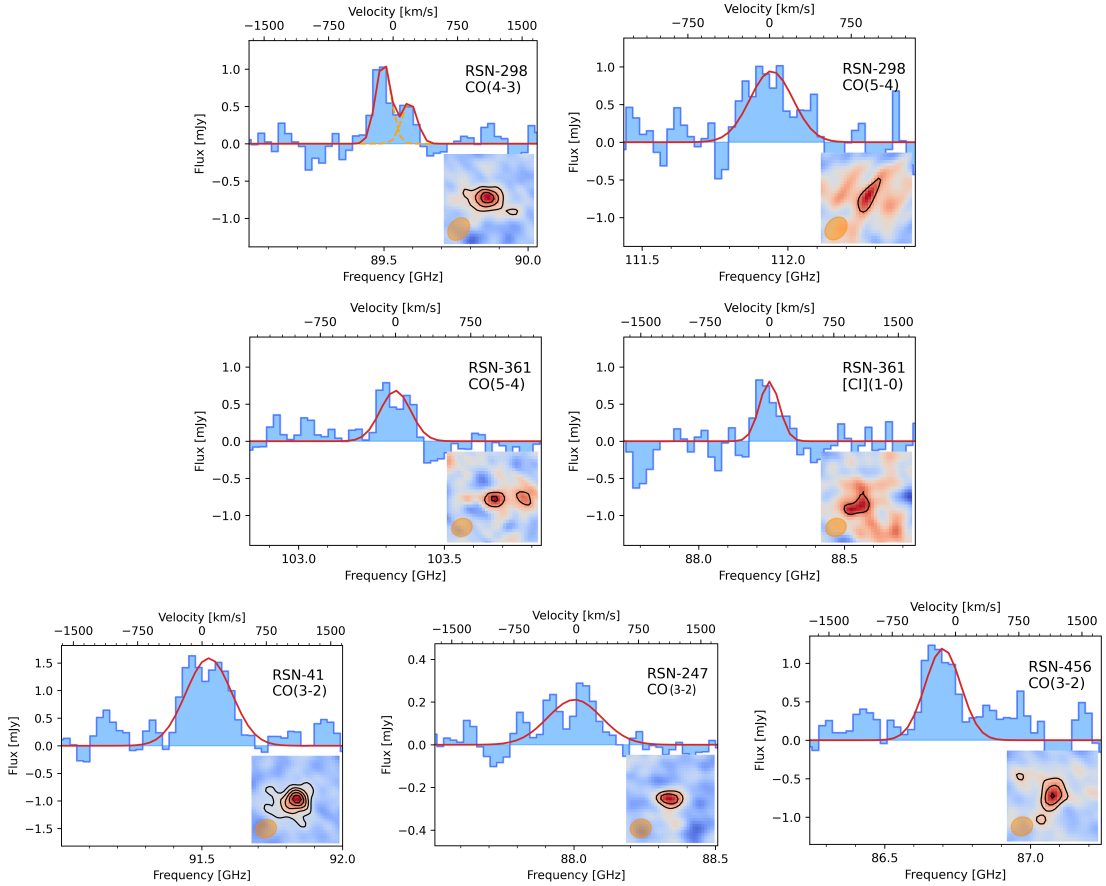

Figure 3.3: *(Continue)*

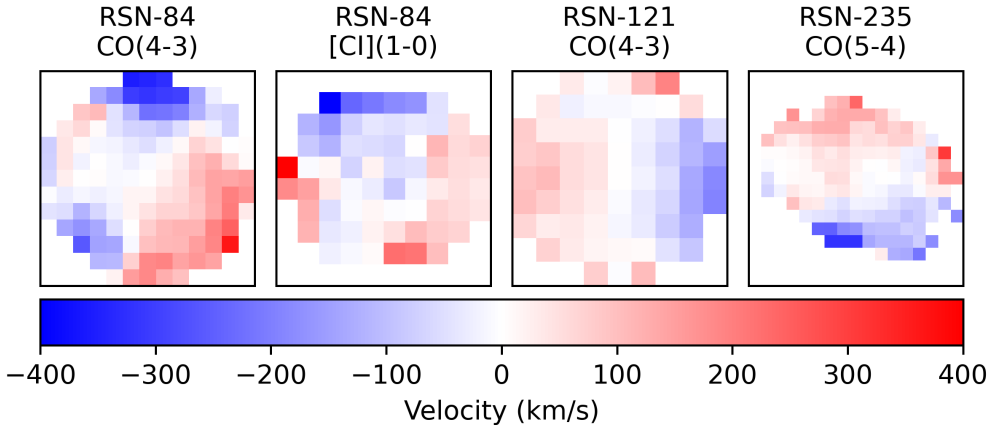

Figure 3.4: Moment-one maps (2.5" × 2.5") of the lines detected in three of the targets. According to our modeling, the maps show clear evidence of a rotating structure or a late stage of a merger.

Table 3.4: Integrated fluxes of the lines detected in our sources.

ID	$I_{CO(3-2)}$ [Jy km/s]	$I_{CO(4-3)}$ [Jy km/s]	$I_{CO(5-4)}$ [Jy km/s]	$I_{[CI](1-0)}$ [Jy km/s]
41	1.06 ± 0.07	-	-	-
84	-	0.55 ± 0.04	-	0.53 ± 0.08
121	-	1.03 ± 0.04	-	0.40 ± 0.07
182	-	-	0.61 ± 0.05	-
235	-	-	0.71 ± 0.04	0.21 ± 0.04
247	0.40 ± 0.06	-	-	-
298	-	0.37 ± 0.03	0.47 ± 0.06	-
361	-	-	0.31 ± 0.04	0.27 ± 0.04
456	0.63 ± 0.03	-	-	-

Note: All the values are obtained through aperture photometry on the zeroth moment of each line with an aperture equal to the 2σ contour.

Table 3.5: Physical properties for our sources derived through SED fitting.

ID	$\log(M_\star)$ $[M_\odot]$	$\log(SFR)$ $[M_\odot\ yr^{-1}]$	$\log(L_{IR})$ $[L_\odot]$	A_V [mag]	$\log(M_{H_2}^{CO})$ $[M_\odot]$	$\log(M_{H_2}^{[CI]})$ $[M_\odot]$	τ_D [a] [Myr]
RSN-41	11.0 ± 0.1	2.72 ± 0.06	12.56 ± 0.06	4.9 ± 0.2	11.10 ± 0.06	-	242 ± 50
RSN-84	11.0 ± 0.1	2.76 ± 0.03	12.57 ± 0.03	4.8 ± 0.2	10.86 ± 0.08	11.17 ± 0.06	258 ± 53
RSN-121	11.0 ± 0.1	2.65 ± 0.06	12.47 ± 0.06	4.2 ± 0.2	11.11 ± 0.07	11.01 ± 0.07	229 ± 60
RSN-182	11.1 ± 0.1	2.92 ± 0.03	12.72 ± 0.03	3.8 ± 0.2	11.02 ± 0.08	-	126 ± 27
RSN-235	11.3 ± 0.1	3.04 ± 0.03	12.83 ± 0.04	4.5 ± 0.2	11.07 ± 0.08	10.96 ± 0.08	83 ± 18
RSN-247	10.7 ± 0.2	2.57 ± 0.04	12.46 ± 0.04	4.6 ± 0.4	10.71 ± 0.09	-	140 ± 31
RSN-298	11.0 ± 0.1	2.69 ± 0.06	12.58 ± 0.06	4.2 ± 0.3	10.81 ± 0.08	-	132 ± 30
RSN-361	10.8 ± 0.3	2.79 ± 0.03	12.61 ± 0.03	3.9 ± 0.6	10.7 ± 0.1	11.06 ± 0.06	186 ± 34
RSN-456	10.8 ± 0.2	2.41 ± 0.07	12.37 ± 0.07	4.6 ± 0.3	10.93 ± 0.06	-	288 ± 63
Median [b]	11.0 ± 0.1	2.6 ± 0.1	12.6 ± 0.1	4.5 ± 0.4	10.9 ± 0.2	11.03 ± 0.07	263 ± 86
G24 [b]	11.0 ± 0.4	2.8 ± 0.4	12.6 ± 0.4	3.9 ± 0.3	–	–	–

Note: The SED fitting is performed with the code `Cigale` (Boquien et al., 2019) after the assumed spectroscopic redshifts were retrieved in Section 3.3.3. The last two columns report the molecular gas mass and the depletion time as estimated in Section 3.4.3. The last two rows report the median value computed for the galaxies here and for the whole sample analyzed in Chapter 2 with `Cigale`.

[a] Estimated from the $M_{H_2}^{[CI]}$ when available.

[b] The uncertainties on the median quantities are reported as half the interval between the 84th and the 16th percentile.

3.3.4 Initial insights into the ISM kinematics

As shown in Table 3.3, most of our targets have quite broad lines (FWHMs of several hundreds of km/s). This result is familiar for high-z DSFGs and ULIRGs in general (see, e.g., Jin et al. 2019, 2022; Cox et al. 2023), and it is generally explained through an ISM that is much more turbulent than what is commonly observed in local galaxies. In this study, however, we perform a more detailed analysis of some targets because the FWHM of the lines is much larger than the spectral resolution requested in our observation. This property allows us to infer some initial insights into the ISM kinematics inside our galaxies. As shown in Figure 3.3, most of the lines observed in our targets have a peculiar morphology that suggests the possible presence of two peaks in the observed line spectrum. This result could be explained by a kinematically decoupled component such as in a disk or in the later stage of a merger. To decide in a statistically motivated way whether our lines should be modeled with a single or double component, we perform a test hypothesis. In our case, the null hypothesis consists of modeling the line with a single Gaussian, while the alternative hypothesis consists of a modeling with two Gaussians. We employ two nested models (with four and seven free parameters, since we allow a residual continuum component), and perform a partial F-test (e.g., Bevington & Robinson, 2003) employing a threshold of 0.05 for the level of significance to reject the null hypothesis. Considering only the highest S/N line in each spectrum, we obtain that a double component is statistically significant for five out of nine targets ($\sim 55\%$). For most galaxies in which two lines are detected (with the notable exception of RSN-84), the lower S/N of the second line prevents us from concluding that the additional component is statistically required for a correct modeling. For all the lines for which the double model is statistically motivated, we report the best-fitting parameters in Table 3.6. A comparison of the fraction of double-peaked lines with other similar studies in the current literature presenting spectroscopic follow-up of SMGs at (sub)mm wavelengths shows that our percentage is higher than the $\sim 30\%$ reported by Bothwell et al. (2013) (detecting CO emission in a large sample of 32 SMGs in the redshift range $1.2 < z < 4.1$) and Aravena et al. (2016) (observing 17 lensed DSFGs at $2.5 < z < 5.7$). and the $\sim 40\%$ of double-peaked profiled reported by Birkin et al. (2021) (studying 61 ALMA-detected SMGs). Unfortunately, the limited size of our sample prevents us from unambiguously establishing whether this difference is due to the different selections or a consequence of the different S/N achieved by the different observations. The two components with different velocities in our galaxies can be explained with a rotating structure or as the signature of the late stage of a major merger. These hypotheses are also strengthened by the first moments of the CO and [CI] lines within three of our targets (those with the highest S/N in the CO-lines, RSN-84, RSN-121, and RSN-235. In RSN-84, the same structure is visible in the CO(4-3) and [CI](1-0) lines; see Figure 3.4). Unfortunately, the spatial resolution of our observations is not sufficient to distinguish between the two possible models (i.e., a disk or a merger). Similarly, the coarse spatial and spectral resolution prevents us from performing a proper modeling of the possible disk (see, e.g., Di Teodoro & Fraternali 2015; Roman-Oliveira et al. 2023). Constraining the deprojected velocity of the gas and its velocity dispersion would allow us to determine whether the structure is stable. This result would be of crucial importance to

Table 3.6: Best-fitting parameters for the Gaussian modeling when two components are employed, as described in Section 3.3.4.

ID	Δv [km/s]	FWHM (red) [km/s]	FWHM (blue) [km/s]
84 - CO(4-3)	(366 ± 61)	(191 ± 99)	(401 ± 114)
84 - [CI](1-0)	(552 ± 39)	(234 ± 51)	(305 ± 77)
121	(449 ± 33)	(289 ± 47)	(319 ± 68)
182	(219 ± 321)	(387 ± 440)	(203 ± 214)
235	(290 ± 132)	(376 ± 222)	(233 ± 78)
298	(302 ± 58)	(208 ± 129)	(208 ± 66)

Note: The columns report the IDs of the galaxies, the velocity offset between the two Gaussian components, and the FWHM of each component.

constrain some of the evolutionary models of massive galaxies (see Section 3.4.4).

3.3.5 SED fitting

After we assessed the spectroscopic redshift of our sources, we estimate their physical properties through an SED fitting with the code `Cigale` (Boquien et al., 2019), by using the same libraries and setup as discussed in Section 2.5.2. The results of the SED fitting and all the physical properties estimated with `Cigale` are summarized in Figure 3.5 and in Table 3.5. Through the SED fitting, we estimate the stellar mass ($M_\star$), the infrared luminosity ($L_{\rm IR}$), and the dust attenuation ($A_{\rm v}$). As in Chapter 2, we estimate the SFR from the $L_{\rm IR}$ through the relation by Kennicutt & Evans (2012), rescaled to a Chabrier (2003) IMF. Interestingly, `Cigale` reports a $f_{\rm AGN} = 0$ for all our targets, and we can therefore safely conclude that the photometry of our galaxies is correctly reproduced without a dusty torus component. Another indication for the lack of a strong AGN contribution in our sample comes from the estimation of the $q_{\rm TIR}$ from the infrared luminosity (computed through the SED fitting) and the radio flux (see, e.g., Helou et al., 1985). The latter was converted into a 1.4 GHz radio luminosity through the spectroscopic redshift and the radio slope measured through the radio fluxes at 3 GHz and 1.4 or 1.28 GHz. For the galaxies without a second radio detection (RSN-41, RSN-235, and RSN-298), we assume the median slope computed for the rest of the sample. We obtain that all the $q_{\rm TIR}$ are in the range [2.45,2.55], which agrees well with what is commonly measured for star-forming galaxies (e.g., Yun et al., 2001).

3.4 Results and discussion

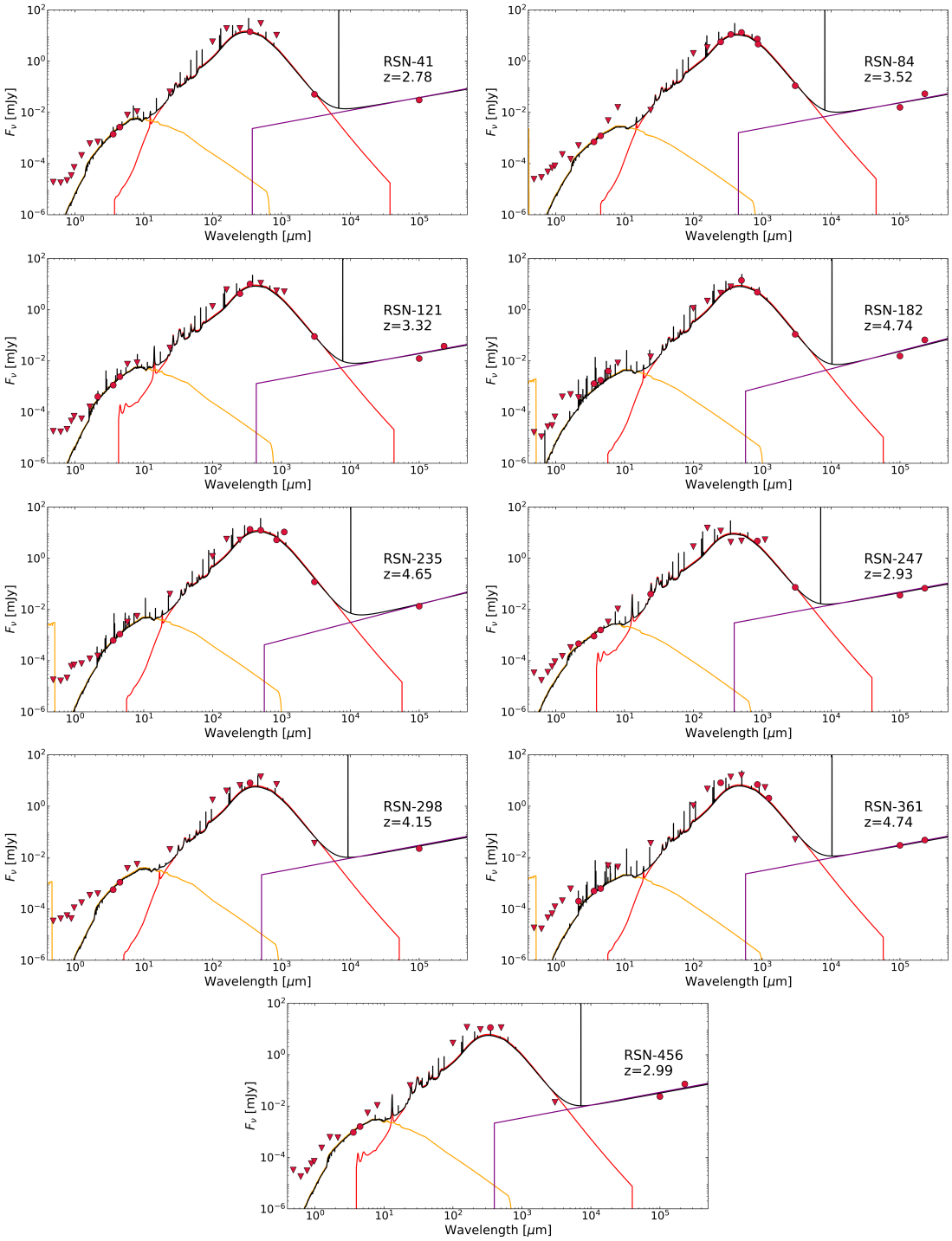

Figure 3.5: Best-fitting SEDs of our targets as computed by `Cigale` (Boquien et al., 2019). The different emissions in the galaxies are color-coded. Attenuated stellar emission is reported as an orange line. The dust emission is reported in red, and the radio emission is reported in purple. Finally, the solid black line shows the best-fitting SED. The observed photometry is reported as red points (detections with S/N>3) and triangles (upper limits).

3.4.1 Analysis of the spec-z

As visible in Table 3.3, the nine galaxies analyzed in this paper have a spec-z between 2.8 and 4.7, with a median value of $\sim$ 3.52. By comparing these values with the photometric redshifts estimated in Chapter 2 (when the continuum point at 3 mm obtained in Section 3.3.1 is added), we note that the agreement is quite good (see Figure 3.6). More quantitatively, we can measure the accuracy of our photo-z as

$$\text{median}\left(\frac{|z_{\text{phot}} - z_{\text{spec}}|}{1 + z_{\text{spec}}}\right) = 0.05 \tag{3.4}$$

when considering all the spec-z assessed in Section 3.3.3. This result validates the procedure followed in the previous chapter for estimating the photometric redshift of the RS-NIRdark galaxies, and it is quite encouraging for future follow-ups for the high-z candidates reported there. Finally, we underline that because for three of our galaxies with at least one line, we have a counterpart in the COSMOS2020 catalog (see Table 3.1), we can retrieve three photo-z from that catalog ($z = 3.6 \pm 0.3$, $z = 2.8 \pm 0.3$, and 4.6 ± 0.3 for RSN-84, RSN-247, and RSN-361, respectively). These quantities were computed by Weaver et al. (2022) with the two SED-fitting codes `Eazy` (Brammer et al., 2008) and `LePhare` (Arnouts et al., 1999; Ilbert et al., 2006) on the optical and NIR bands included in the COSMOS2020 catalog and on the first two channels of IRAC, and they agree perfectly with the spectroscopic redshifts estimated through (sub)mm spectroscopy (see Table 3.3). A last interesting comparison can be performed with the photometric redshifts computed by Talia et al. (2021) that were employed to select the targets for these ALMA observations. Unfortunately, most of the sources are located at a lower redshift than expected from that study (see Table 3.3). This difference can be explained by the several improvements in the photometry extraction and in the photo-z estimation employed in Chapter 2 with respect to Talia et al. (2021). We expect that most of the differences arise because of the new deblending procedure based on the `PhoEBO` algorithm (allowing us to better extract the photometry from low-resolution maps such as the IRAC maps), the deeper IRAC images employed in Chapter 2, and the more stringent upper limits employed in the photometric bands without detections.

3.4.2 Analysis of the physical properties

The results obtained through SED fitting in Section 3.3.5 allow us to confirm one of the main results established in Chapter 2 where we assessed that the RS-NIRdark galaxy selection produces a sample of star-bursting DSFGs. However, since this result was based on an SED fitting in which the redshift was a free parameter, the quantities estimated through this procedure were affected by significant uncertainties due to the several degeneracies between the shape of the SED and the redshift. By assuming the spec-z measured through our ALMA observations, we can decrease the uncertainty on these quantities. First of all, the median properties estimated with the improved SED fitting are broadly compatible with those estimated in Chapter 2 for the whole sample (see also the results discussed in Talia et al. 2021 and Behiri et al. 2023 regarding a smaller subset of the same sample). We underline, however, that since the galaxies in the proposal were selected from those with at least one

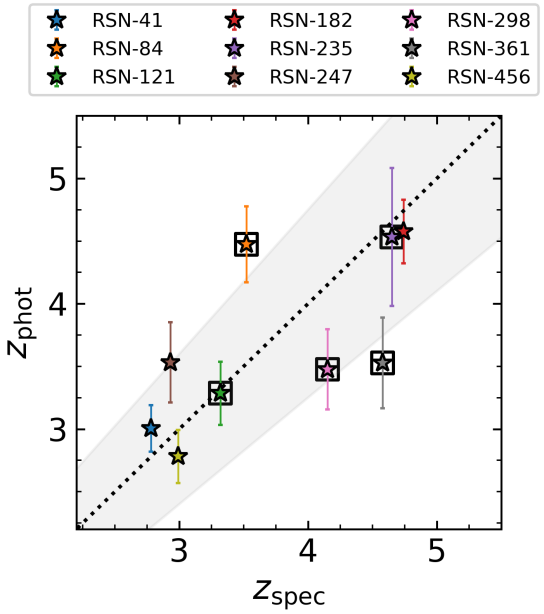

Figure 3.6: Comparison of the photometric redshift estimated in Chapter 2 and the spectroscopic redshifts measured in this study. The one-to-one relation is reported as the dotted black line, and the shaded gray area shows the galaxies with $|\Delta z|/(1+z) < 0.15$. The galaxies with a spec-z obtained from the modeling of two lines are highlighted with an additional box.

secure detection in the FIR or (sub)mm regimes, we do not expect the median properties of our sample to be necessarily similar to those of the whole sample of RS-NIRdark sources. A second interesting comparison between this chapter and the previous one resides in the comparison of the SFR and stellar mass computed through the new SED fitting and the main sequence of the star-forming galaxies. As shown in Figure 3.7, most of the targets are still located above the main sequence by Schreiber et al. (2015), close to the star-bursting regime (i.e., galaxies with an SFR at least three times higher than what is expected from a main-sequence source with the same mass and in the same redshift bin). We underline, however, that the location on the main sequence strongly relies on the estimated stellar mass, which is quite uncertain for our targets because the rest-frame optical continuum is highly obscured by the dust. Nevertheless, the employment of `Cigale`, relying on the energy balance principle between dust absorption and emission, allows us to obtain some indirect constraint on the dust extinction from the infrared and radio coverage. More stringent constraints on the stellar mass will be provided for most of the galaxies in the whole sample of RS-NIRdark galaxies in COSMOS by the deep NIR imaging provided by JWST as part of the COSMOS-Web survey (see Chapter 4). Figure 3.7 also reports for comparison the location of other populations of dark DSFGs in the stellar mass versus SFR plane: the H-dropouts by Wang et al. (2019) and the NIRfaint SMGs by Smail et al. (2021). While the overlap between our RS-NIRdark galaxies and the H-dropouts has already been studied in the previous chapter, the availability of the new (sub)mm data allows us to study more quantitatively how many of our sources would be selected with the

criteria employed by Smail et al. (2021). We recall that these sources are part of a sample of 707 SMGs collected in the Ultra Deep Survey by re-imaging a sample of galaxies initially detected in the 850 μm maps produced with the SCUBA-2 camera (Geach et al., 2017) with ALMA (in band 7, i.e., at a representative frequency of 870 μm; Stach et al. 2019). The sample studied by Smail et al. (2021) contains all[3] the sources with $Ks > 25.3$ mag (at 5 σ). Since the Ks limiting magnitudes in the UDS and in COSMOS are similar, we study the overlap between the two populations by comparing the (sub)mm flux at 850 μm. For our sources, only RSN-84 has an ALMA flux at the same frequency (see Table 3.2). Three of the other galaxies have analogous fluxes from the deblending of the SCUBA2 maps (Jin et al., 2018), while the others have only upper limits (S/N$< 3\sigma$, since the uncertainties in the SuperDeblended catalog also account for the deblending procedure). For these sources, we employ the best-fitting fluxes at 850 μm computed with `Cigale`. On the other hand, the sources in Smail et al. (2021) were initially selected for having an S/N$> 4\sigma$ (equivalent to $S_{850\mu m} \geq 3.8$ mJy) in the original SCUBA-2 maps. However, the higher resolution achieved by ALMA in the Stach et al. (2019) follow-up allowed the discovery of multiple fainter sources that contribute to the original sources detected by SCUBA2 up to $S_{850\mu m} > 1$ mJy and with a median value of 3.8 ± 0.3 mJy. Therefore, even though four out of nine sources in our sample would not have been selected by the original SCUBA2 survey because they are too faint for the limited sensitivity of that instrument (see Table 3.2), all of them would have been detected by the deeper ALMA follow-up.

3.4.3 Gas mass and depletion time

It is well known that the [CI](1-0) and the CO(1-0) lines can be employed as good tracers of the molecular gas inside galaxies (e.g., Papadopoulos et al., 2004; Valentino et al., 2020b; Gururajan et al., 2023). For four galaxies in our sample (see Table 3.3), we observed the [CI](1-0) line and — therefore — directly estimate the molecular gas mass in our objects by employing the relation by Papadopoulos et al. (2004),

$$M(H_2)^{[CI]} = 1375.8 \times 10^{-12} \frac{D_L^2 I_{[CI](1-0)}}{(1+z) A_{10} Q_{10} X_{CI}} [M_\odot] \tag{3.5}$$

, where D_L is the luminosity distance of our target expressed in Mpc, $I_{\rm CO}$ is the integrated line flux, and $A_{10} = 0.793 \times 10^{-7}$ s^{-1} is the Einstein coefficient. Q_{10} and X_{10} are the [CI] excitation factor and the [CI]/H_2 abundance ratio, respectively. For these quantities, we employ literature standard values of $X_{CI} = 3 \times 10^{-5}$ and $Q_{10} = 0.6$ (Papadopoulos et al., 2004; Bothwell et al., 2017). Through Equation 3.5, we estimate the gas masses for RSN-84, RSN-121, RSN-235, and RSN-361 reported in Table 3.5. For all the other targets in which we did not detect the [CI](1-0) line, we derive the gas mass from the CO(1-0) line through the relation

$$M(H_2)^{\rm CO} = 3.25 \times 10^7 \alpha_{\rm CO} I_{\rm CO} \nu_{\rm obs}^{-2} D_L^2 (1+z)^{-3} \tag{3.6}$$

[3]The original sample of NIR-faint sources by Smail et al. (2021) would contain another 50 sources that are not included in that study because the photometry at optical and NIR frequencies is contaminated. We assume that the 30 galaxies analyzed by Smail et al. (2021) are representative of the full sample

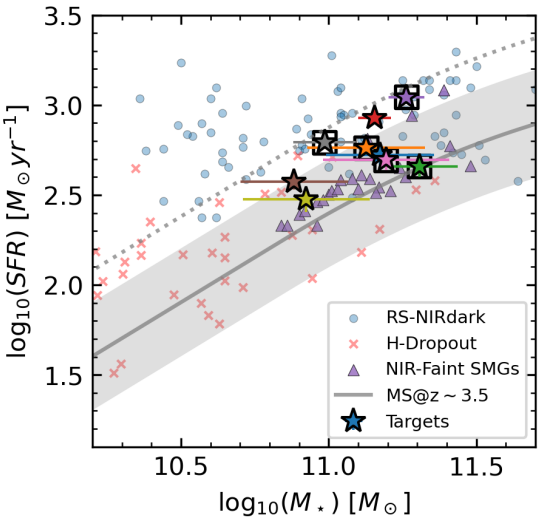

Figure 3.7: Comparison of the physical properties estimated through `Cigale` and the main sequence of star-forming galaxies by Schreiber et al. (2015) (solid gray line; the shaded area is its $1\sigma = 0.3$ dex its scatter). The targets are reported as colored stars, with the same color-code as employed in Figure 3.6. The dotted gray line reports our threshold for identifying the star-bursting galaxies (i.e., those whose SFR is higher than three times that expected from a main-sequence galaxy). For reference, we also report the location of the RS-NIRdark galaxies around $z \sim 3.5$ studied in Chapter 2, the NIR-faint SMGs by Smail et al. (2021), and the H-dropout galaxies by Wang et al. (2019)

.

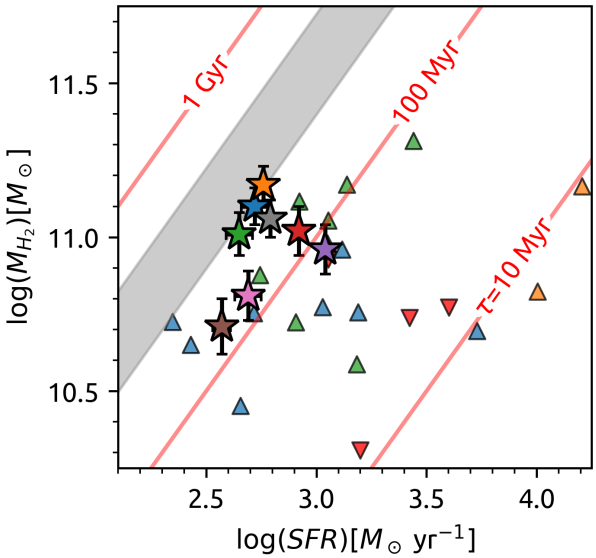

Figure 3.8: Depletion time and star formation rate for our targets. The RS-NIRdark galaxies are reported as colored stars, following the same color-code as in Figure 3.6. The colored triangles are other populations of SMGs, namely those collected by Bothwell et al. (2017), Cañameras et al. (2018), and Walter et al. (2011), reported in blue, orange, and green, respectively. We also report some confirmed QSOs from the same studies as reversed red triangles. The shaded gray area reports the depletion time expected from main-sequence galaxies at $z \sim 3.5$ following the relation $\tau_{\rm D} = 1.5(1+z)^{\alpha}$ found by Saintonge et al. (2013), with α spanning from -1.0 (Davé et al., 2012) to -1.5 (Magnelli et al., 2013), rescaled to a Chabrier (2003) IMF.

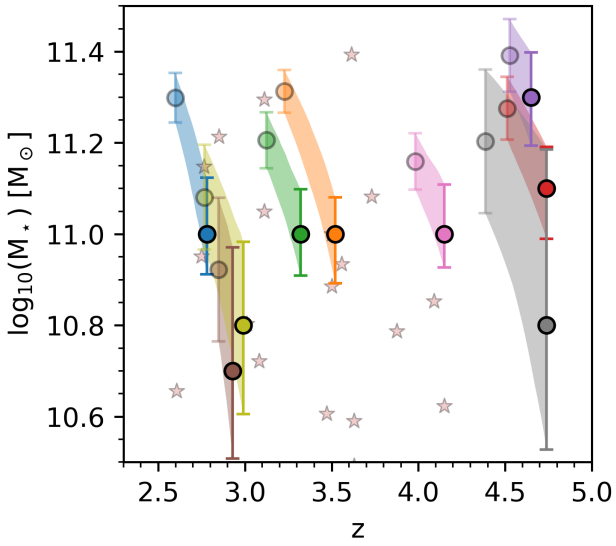

Figure 3.9: Possible evolutionary paths of our targets, assuming a simple evolutionary model with a constant star formation and a final stage in which all the molecular gas has been transformed into stars. The uncertainties are considered through a Monte Carlo integration. The color map is the same as employed for Figure 3.6. For reference, the shaded red stars report the stellar mass and the redshifts for the massive and passive galaxies at $z \sim 3$ discovered by Schreiber et al. (2018a).

(see, e.g., Bolatto et al. 2013 and references therein). It is well known that the value of $\alpha_{\rm CO}$ is highly uncertain and strongly depends on the specific property of each galaxy. We choose a literature value of 0.8 $M_{\odot}$ [K km s^{-1} pc^{-2}]$^{-1}$, which is usually employed for star-bursting galaxies (see, e.g., Bolatto et al. 2013; Gururajan et al. 2023). In order to use Equation 3.6, we rescale the measured fluxes of our CO lines to the CO(1-0) transition by assuming a CO spectral line energy distribution (SLED). Several studies highlighted that the CO-SLED of galaxies is strongly affected by AGN (see, e.g., Vallini et al., 2019) and evolves with redshift (e.g., Boogaard et al., 2020). Since the previous test performed on our targets by Talia et al. (2021) and in Chapter 2 (together with the null AGN fraction obtained through SED fitting with `Cigale`; Section 3.3.5) excluded strong nuclear activity, we choose the CO-SLED obtained by Bothwell et al. (2013) for a sample of DSFGs in the redshift range $1.2-4.1$ (i.e., compatible with the spec-z of our targets). We employ $R_{31} = 2.3 \pm 0.3$, $R_{41} = 3.0 \pm 0.4$, and $R_{51} = 3.8 \pm 0.7$, where R_{n1} is the ratio of the integrated line flux in the nth CO transition and the CO(1-0). The gas mass obtained through these relations is reported in Table 3.5. For the galaxies with both a CO and a [CI] line, we report both estimates of the gas mass. However, since the values estimated from the [CI](1-0) line rely on fewer assumptions than those based on the CO lines (i.e., they do not rely on the assumed CO-SLED, even though they still depend on the conversion factor between the [CI](1-0), as uncertain as the $\alpha_{\rm CO}$ value), we employ these gas masses in the following analyses. The information about the gas content of our galaxies can be combined with the SFR estimated in Section 3.3.5 to assess the depletion time of our galaxies. This quantity is defined as the amount of

time in which each object would transform its whole gas mass into stars assuming a constant SFR:

$$\tau_{\rm D} = \frac{M(H_2)}{\rm SFR}. \tag{3.7}$$

For our galaxies, we obtain the depletion times reported in the last column of Table 3.5. The values for the galaxies in our sample range from 80 to 300 Myr. These quantities can be compared with other populations of galaxies in the current literature, as shown in Figure 3.8. Firstly, the depletion time of our targets can be compared with that expected from main-sequence galaxies. Saintonge et al. (2013) reported a $\tau_{\rm D}$ that evolved with redshift as $\tau_{\rm D} = 1.5(1+z)^{\alpha}$, with several collaborations finding different values for the exponent, spanning from $\alpha = -1.0$ (Davé et al., 2012) to $\alpha = -1.5$ (Magnelli et al., 2013). Our targets have a shorter depletion time than main-sequence galaxies. This result represents a further confirmation (independent of the more uncertain stellar mass) of the star-bursting nature of our sources. Secondly, we can compare the gas mass and the SFR of our RS-NIRdark galaxies with several SMGs in the current literature (those analyzed by Walter et al. 2011, Bothwell et al. 2017, and Cañameras et al. 2018). For all these sources, we retrieve the infrared luminosity and the [CI](1-0) line fluxes from the study by Valentino et al. (2020b). Following the relations by Kennicutt & Evans (2012) and Papadopoulos et al. (2004), we estimate the SFR and gas mass in a consistent way with those derived from our targets. We obtain that our RS-NIRdark galaxies are — on average — more gas rich than the SMGs analyzed in those studies and located in the low-SFR tail of their distribution and therefore have a longer depletion time.

3.4.4 Possible evolutionary path

The gas mass and depletion times estimated in Section 3.4.3 allow us to forecast a possible evolutionary path for our sources. We employ the same simplistic model that was used to define the depletion time: we assume that the SFR remains constant inside our sources until all the gas mass is converted into stars. Clearly, this model does not account for any quenching mechanism (e.g., due to possible AGN feedback; see, e.g., Fabian 2012) or gas accretion from the intergalactic medium (e.g., Sancisi et al., 2008). With this model, we assume that our galaxies evolve from an initial state characterized by $z_0 = z_{\rm spec}$ and $M_{\star,0} = M_{\star}$ to a final state with $z_f = z_{\rm spec} - \Delta z$ and $M_{\star,f} = M_{\star} + M_{\rm H_2}$, where Δz is the difference in redshift elapsed during the depletion time. The results of this simplistic model applied to our targets are shown in Figure 3.9, where we consider all the uncertainties on the involved quantities through a Monte Carlo integration. We compare the final stage of our galaxies with the redshift and stellar mass of the massive and passive galaxies discovered by Schreiber et al. (2018a) at $z \sim 3$ noticing a significant overlap between the two populations. This result, combined with the number densities estimated in Chapter 2 (see also Talia et al. 2021 and Behiri et al. 2023), suggests that the RS-NIRdark galaxies could represent a significant fraction of the progenitors of the massive and passive galaxies at $z \sim 3$. In the broader context of galaxy evolution studies, this result confirms the idea that the dust-obscured galaxies play a significant role in the evolution of the most massive galaxies in the Universe (see,

e.g., Casey et al., 2014; Toft et al., 2014; Valentino et al., 2020a). In addition, the components with different velocities detected in our targets in Section 3.3.4 could support two possible evolutionary scenarios for our galaxies. On the one hand, they could be the signature of a significant fraction of major mergers within our sample. This result would confirm the so-called *merger-driven* scenario for the formation of massive galaxies, where these objects are formed through a series of major mergers (see, e.g., Hopkins et al. 2008b,a). On the other hand, the double component in our galaxies could be due to the presence of a rotating disk. In this case, this result would support the so-called *in situ* scenario, where the build-up of massive galaxies occurs via the rapid compaction of a gaseous outer disk triggering a significant burst of star formation, and via the subsequent stellar and AGN feedback processes quenching it within a relatively short timescale (< 1 Gyr; see e.g Lapi et al. 2014; Pantoni et al. 2019).

3.5 Chapter summary

In this chapter, I presented the first spectroscopic follow-up at millimeter wavelengths for a pilot sample of nine RS-NIRdark galaxies in the COSMOS field. Through a new series of ALMA observations, my collaborators and I identified at least one bright emission line in all the targets and two lines in five out of nine objects. From the analysis of the new ALMA data, we obtained the results listed below.

- Modeling the detected lines as CO and [CI] transitions, we estimated a spectroscopic redshift for all the galaxies in our sample. These values agree well with those estimated through SED fitting in Chapter 2.

- The new availability of a spectroscopic redshift allowed us to decrease the degeneracies in the SED-fitting procedure. This improved SED fitting confirmed one of the main results for the RS-NIRdark galaxies: they represent a population of highly obscured ($A_{\rm v} \sim 4$), massive ($M_\star \sim 10^{11} M_\odot$), and star-forming (SFR$\sim 500$ M$_\odot$ yr^{-1}) galaxies.

- The same improved SED fitting, together with preexisting data, allowed us to estimate the flux of our sources at 870 μm. This value is analogous to that reported by Smail et al. (2021) for their sample of NIR-faint SMGs, suggesting that our radio selection is able to provide a similar population of DSFGs as those obtained from (sub)mm selections. A similar conclusion was reached by observing the overlap between the physical properties (stellar mass and SFR) computed through SED fitting for the two samples.

- The good spectral resolution of the new ALMA observations allowed us to assess a high fraction ($\sim 55\%$) of double-peaked profiles in the lines detected in our targets. We explained this result with the possible presence of a rotating structure within our galaxies or with major mergers in our sample. High-resolution follow-up with ALMA or JWST are needed to distinguish between these two possibilities.

- Based on the CO and [CI] lines detected in our targets, we estimated the gas mass and the depletion time of our galaxies. These results allowed us to forecast a possible evolutionary path for our objects. This strongly suggests that the RS-NIRdark galaxies might represent a significant fraction of the progenitors of the massive and passive galaxies at $z \sim 3$ and that they are excellent probes for testing galaxy evolution models.

Chapter 4

A radio selection for DSFGs: The impact of JWST

4.1 Introduction to the chapter

This chapter describes the improved analysis of my RS-NIRdark galaxies taking advantage of the new JWST data of the COSMOS field obtained as part of the COSMOS-Web survey (Casey et al., 2023). It is structured as follows. In Section 4.2, I describe the data and the updated sample selection. In Section 4.3, I estimate the physical properties of my targets through SED fitting and the possible AGN contribution. In Section 4.4 I estimate the luminosity function of my sources and in Section 4.5 I use this result to compute how much these galaxies contribute to the cosmic SFRD. I discuss my main results in Section 4.6 and — finally — I present my conclusions in Section 4.7.

Most of the analysis presented in this chapter is also described in "*Going deeper into the dark with COSMOS-Web: JWST unveils the total contribution of Radio-Selected NIRfaint galaxies to the cosmic Star Formation Rate Density*" by Gentile et al., (*subm.*), currently under review in Astronomy & Astrophysics.

4.2 Data

4.2.1 JWST photometry

The JWST photometry for the sources described in this chapter comes from the COSMOS-Web survey (GO #1727, PIs Kartaltepe & Casey; Casey et al. 2023), a cycle 1 treasury program consisting in the NIRCam and MIRI imaging of the COSMOS field. The program includes a contiguous NIRCam mosaic covering the central region of the field ($\sim$0.54 deg^2) in the four filters F115W, F150W (the short-wavelength filters; SW hereafter), F277W, and F444W (long-wavelength; LW). In parallel with the NIRCam mosaic, COSMOS-Web also includes a MIRI mosaic covering a total (non-contiguous) area of 0.19 deg^2 in the F770W filter. A full description of the COSMOS-Web program can be found in Casey et al. (2023), while the data reduction procedure is described in detail in M. Franco et al., (*in prep.*).

In this work, I use the NIRCam and MIRI photometry extracted with `SourceXtractor++`

(SE++; Bertin et al. 2020; Kümmel et al. 2020). This software performs the detection on a positive-truncated χ^2 combination of the four NIRCam filters PSF-homogenized to F444W. Each source is then modeled with a Sérsic profile (Sérsic, 1963). The best-fitting model is then convolved with the PSFs of the five (NIRCam and MIRI) scientific maps in order to extract the fluxes and the related uncertainties. A full description of the procedure followed to build the COSMOS-Web catalogs can be found in M. Shuntov, L. Paquereau et al., (*in prep.*).

4.2.2 Radio data

The radio data analysed in this chapter are the same described in Chapter 2 and come from the VLA-COSMOS Large Program (Smolčić et al., 2017). To perform the sample selection, together with my collaborators, I cross-match the photometric catalog of COSMOS-Web with the public catalog by Smolčić et al. (2017) with a matching radius of 0.7" (as in Chapter 2). This procedure gives in output a sample of 3196 galaxies. However, given the better spatial resolution of NIRCam than that achieved with the VLA, 139 radio sources have multiple NIRCam galaxies falling in the matching radius, that we exclude from the final sample.

4.2.3 Sample selection

We perform a sample selection resembling those previously employed in Chapter 2 (see also Talia et al. 2021 and Behiri et al. 2023) to collect the initial sample of RS-NIRdark galaxies in the COSMOS field, but taking advantage of the new JWST photometry coming from the COSMOS-Web survey.

We start from the parent sample assembled in Section 4.2.2 and we select our galaxies through the following criteria:

$$\begin{cases} S_{3\mathrm{GHz}} & > 11.5\ \mu\mathrm{Jy} \\ F150W & > 26.1\ \mathrm{mag} \end{cases} \tag{4.1}$$

Clearly — by construction — all the galaxies in the final sample also need to be detected in the detection image employed to assemble the COMSOS-Web catalog.

The first criterion requires that the sources are robustly ($S/N > 5$) detected at radio frequencies, while the latter requires that they would not be detected in the NIR filter H ($\lambda \sim 1.6\ \mu$m) at the 3σ depth of the COSMOS2020 catalog (2" aperture; see Weaver et al. 2022). The full sample includes 127 galaxies in the 0.54 deg^2 observed in the COSMOS-Web survey (55 with MIRI coverage, $\sim 32\%$ of the full sample, 44 with a robust detection at S/N>3).

Compared with the previous criteria employed in Chapter 2, we include in our sample sources with lower values of the S/N of the radio flux. The availability of a strong prior such as the NIRCam imaging — indeed — allows us to relax this criterion, potentially including in our sample some higher-z sources previously missed due to their faintness at radio frequencies (e.g. the spectroscopically-confirmed source at $z \sim 5$ reported by Jin et al. 2019, with a reported S/N at 3 GHz of 5.2).

Moreover, in Chapter 2, we required the lack of counterpart in the COSMOS2020 catalog, whose detection was performed on a χ^2-image including the four VISTA

filters Y, J, H, and Ks and the two optical filters i and z from HSC (Weaver et al., 2022). However, the low resolution of these images did not completely allow a counterpart-matching for some optically-bright galaxies blended with nearby sources (see the discussion in Section 2.3). With these improved criteria and with the high-resolution of the NIRCam maps, we aim at selecting a less contaminated sample of "NIRdark" galaxies. Since these sources are now detected at the new depths reached by NIRCam, in the following we will refer to our galaxies as radio-selected NIRfaint (RS-NIRfaint).

4.2.4 Ancillary data

Our targets are located in the COSMOS field. Therefore, the almost complete photometric coverage presented in Section 2.3.1 is available for them. More in detail, we employ in the analysis described in this chapter the ancillary data listed below.

- **Optical-to-MIR**: The `SE++` software employed to extract the NIRCam and MIRI photometry can also be applied to other ancillary data, by fitting each source with the parametric model computed on the NIRCam maps (once convolved with the PSF of low-resolution maps). These additional data include:

  1. The optical maps obtained during the Subaru Strategic Program (SSP DR3; Aihara et al. 2019) performed with the HyperSupreme Cam (HSC) mounted on the Subaru telescope and those obtained with the Advanced Camera for Surveys (ACS) equipped on HST (Koekemoer et al., 2007).
  2. The NIR data obtained during the UltraVISTA survey (DR6; McCracken et al. 2012) performed with the VIRCAM instrument of the VISTA telescope.
  3. The MIRI data obtained as part of the Cosmic Dawn Survey of the COSMOS field (Moneti et al., 2022) performed with the Infrared Array Camera (IRAC) equipped on the *Spitzer* Space Telescope.

  The footprint of all these surveys overlap with the full COSMOS-Web area. In order to cover wavelength ranges not included in our NIRCam and MIRI photometry, we include in our catalog the data in the eight filters $g, r, i, z, y, F814W, Ks$ and in the first channel of IRAC ($\lambda \sim 3.8$ μm). The depths and additional details on the employed maps can be found in Weaver et al. (2022) and in M. Shuntov, L. Paquereau et al., (*in prep.*).

- **FIR/(sub)mm:** We cross-match our catalog with the most updated version of the super-deblended catalog (Jin et al. 2018; Jin et al., *in prep.*), including deblended photometry from the PACS and SPIRE instruments equipped on the *Herschel* space telescope that observed the COSMOS field during the surveys described in Lutz et al. (2011) and Oliver et al. (2012). From the same catalog, we also retrieve photometry at 870 μm from the deblending of the SCUBA-2 maps obtained during the S2COSMOS survey by Simpson et al. (2019). Since the super-deblended employs the 3 GHz sources of the VLA-COSMOS survey as priors, all our galaxies have an entry in that catalog. More in detail, 72 sources in our sample have at least one detection at S/N>3 in at

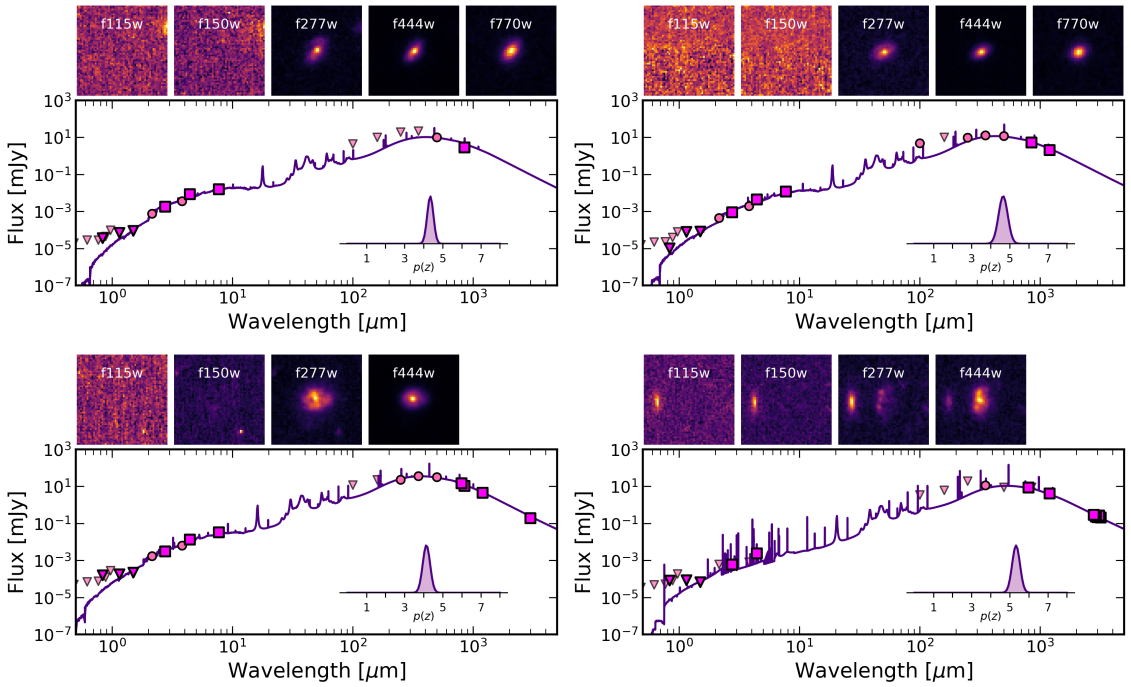

Figure 4.1: Some examples of the SED fitting performed with `Cigale` on the galaxies included in the highest redshift bin. The squares indicate the photometric points from JWST, A3COSMOS and the 870μm from the deblending of the S2COSMOS maps by Simpson et al. (2019). The circular points report the ancillary photometry from HSC, VISTA, Spitzer, and from the super-deblending of the Herschel Maps. The inset shows the Gaussianized $p(z)$ computed with `Cigale`. The top row shows the cutouts (3 arcsec side) in the NIRCam and MIRI filters of COSMOS-Web.

least one PACS or SPIRE filter. Similarly, 31 objects are detected at S/N>3 in the 870 μm maps. All the other objects have upper limits. We also obtain ALMA photometry in the (sub)mm range for 38 of our sources ($\sim$ 30% of the sample) by cross-matching with the most recent catalog (v20220606) from the Automated mining of the public ALMA Archive in the COSMOS field (A3COSMOS; Liu et al. 2019; Adscheid et al. 2024).

- **Radio:** Finally, we retrieve data at radio frequencies (3, 1.4, and 1.28 GHz) by cross-matching with the public catalogs from the VLA-COSMOS large program (Smolčić et al., 2017; Schinnerer et al., 2007) and the early data release of the MIGHTEE survey performed with MeerKAT (Jarvis et al., 2016). Given the lower resolution of the MIGHTEE maps ($\sim 8''$), we only consider the 1.28 GHz radio fluxes of 84 isolated sources (i.e. without another 3 GHz object within 8").

4.3 SED fitting

Our SED fitting is performed with `Cigale` (Boquien et al. 2019), with the same setups described in Section 2.5.2. Differently from the previous chapters — however

— we do not include the radio data in the fitting. These data are employed later (see Section 4.3.3) to estimate the AGN fraction in a different way from what done in the previous chapters. The good convergence of the SED-fitting procedure is ensured by the distribution of the reduced χ^2, with a median value of $\langle\chi^2_\nu\rangle = 1.1$ and 95% of the sample with a $\chi^2_\nu < 5$. Some examples of the SED-fitting output are reported in Figure 4.1.

4.3.1 Photometric redshifts

The distribution of the photometric redshifts for our sources is reported in Figure 4.2. The distribution is quite peaked at $\langle z\rangle \sim 3.6$, with a 1σ dispersion (given as the symmetrized interval between the 16th and the 84th percentile) of 0.8. The availability of the new NIRCam and MIRI photometry allows us to sample the rest-frame optical/NIR emission of our galaxies, reducing the uncertainty on the photometric redshifts. The median $\delta z/(1+z)$ for our galaxies is 0.08, nearly half of what we would obtain by removing the JWST photometry from the SED fitting performed with `Cigale` (0.15). The spectroscopic coverage of our sample is not sufficient to allow a proper testing of our photometric redshift. However, thanks to the collection of spectroscopic redshift in COSMOS (A. A. Khostovan et al., *in prep.*), we found four sources in our sample with a spec-z. These are ID20010161 ($z_{\rm spec} = 5.051$) from Jin et al. (2019) (see also Chapter 5)) and AS2COS0002 ($z_{\rm spec} = 4.600$), AS2COS0011 ($z_{\rm spec} = 4.783$), and AS2COS0014 ($z_{\rm spec} = 2.920$) from Chen et al. (2022). These spectroscopic redshifts are well recovered by our SED fitting, with the first three objects having a discrepancy lower than 2σ and only the last one having a spec-z at 5σ from the photo-z. The median value (computed on this small sample) of the $|\Delta z|/(1+z)$ is 0.09.

4.3.2 Physical properties

With `Cigale`, we also estimate the stellar masses and dust attenuation for our galaxies. Since we do not have any constraints on the rest-frame UV SED of our sources, we do not use the SFR computed through SED fitting for our analysis. Instead, we compute the SFR from the radio luminosity — after accounting for the possible AGN contribution — as described in Section 4.3.3. The results of this procedure are summarized in Table 4.1 and in Figure 4.3. These distributions picture the RS-NIRfaint galaxies as a population of highly dust-obscured ($\langle A_{\rm v}\rangle \sim 3.5$ mag), massive ($\langle M_\star\rangle \sim 10^{10.8}$ M$_\odot$) and star-forming galaxies ($\langle$SFR$\rangle \sim 300$ M$_\odot$ yr^{-1}) located at $\langle z\rangle \sim 3.6$. Figure 4.3 also shows for reference the same properties computed for other notable populations of dusty star-forming galaxies (i.e. those selected with JWST by Gottumukkala et al. 2023 and with SCUBA/ALMA by McKinney et al. 2024). A full comparison between the RS-NIRfaint galaxies and these other samples will be the focus of Section 4.6.1.

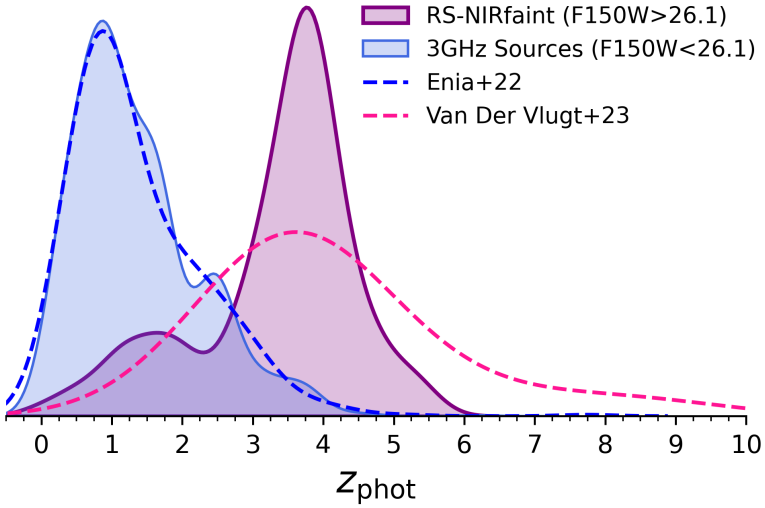

Figure 4.2: Distribution of the photometric redshifts of our RS-NIRfaint galaxies in COSMOS-Web. Our sources are reported as the purple solid line, while the complementary sample of 3 GHz objects with F150W>26.1 mag is reported as the blue solid line. For reference, we also show the photo-z computed by Enia et al. (2022) on their sample of radio sources (with optical/NIR counterparts) in the GOODS-N field (blue dashed line) and those estimated by van der Vlugt et al. (2023) for their sample of optically/NIR-faint galaxies in their deeper COSMOS-XS survey (dashed pink line).

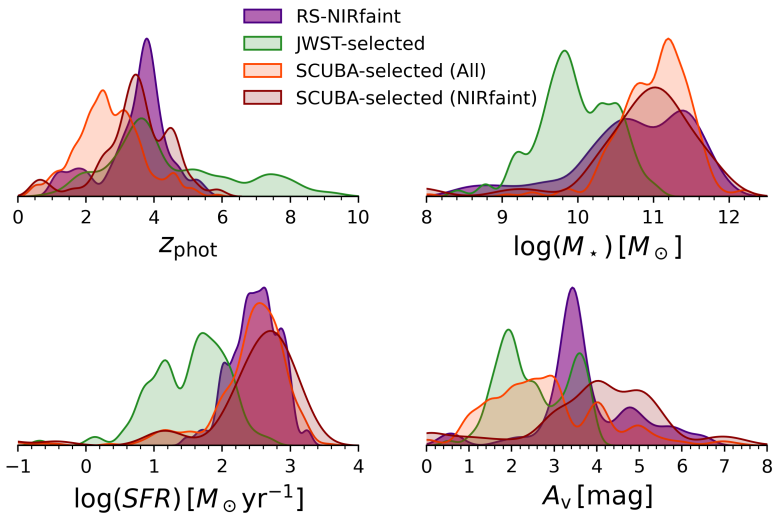

Figure 4.3: Main properties (photometric redshift, stellar mass, SFR, and dust extinction) of our RS-NIRfaint galaxies. We show — for comparison — the same properties computed by Gottumukkala et al. (2023) and McKinney et al. (2024) for their sample of JWST-selected and SCUBA-selected DSFGs as the green and orange dashed lines, respectively. To allow a fair comparison, we only include the sources by Gottumukkala et al. (2023) with $F150W > 26.1$ mag and exclude the sources flagged as possible AGN.

Table 4.1: Median properties of our RS-NIRfaint galaxies

Property	Median	σ	Unit
$z_{\rm phot}$	3.6	0.8	-
$\log(M_\star)$	10.8	0.7	$M_\odot$
log(SFR)	2.5	0.4	$M_\odot$ yr^{-1}
$A_{\rm v}$	3.5	0.9	mag

Note:
The σ of each distribution is reported as the symmetrized interval between the 84th and the 16th percentiles.

4.3.3 AGN contribution

Since we are dealing with galaxies selected at radio frequencies, we need to consider that part of that emission could be due to nuclear activity and not to star formation. For doing so, we follow the procedure established by Ceraj et al. (2018) to measure the so-called "AGN fraction" ($f_{\rm AGN}$), quantifying the AGN contribution to the radio luminosity.

We start by computing the infrared luminosity ($L_{\rm IR}$) of our sources by integrating the best-fitting SED given in output by `Cigale` in the range [8,1000] μm. The $L_{\rm IR}$ is then employed to compute the $q_{\rm TIR}$ parameter (quantifying the ratio between infrared and radio emission) for our galaxies as:

$$q_{\rm TIR} = \log\left(\frac{L_{\rm IR}[W]}{3.75\times10^{12}{\rm Hz}}\right) - \log\left(\frac{L_{\rm 1.4GHz,tot}}{\rm WHz^{-1}}\right) \tag{4.2}$$

where $L_{\rm 1.4,tot}$ is the radio luminosity at 1.4 GHz computed from the radio flux at 3 GHz. The conversion between these quantities is done by computing the radio slope between 1.4 (or 1.28) GHz and 3 GHz (for the galaxies with a measured flux at these frequencies, see Section 4.2.4) or by assuming a fixed slope of $\alpha = -0.7$ (commonly employed for star-forming galaxies, see e.g. Novak et al. 2017).

The $q_{\rm TIR}$ of star-forming galaxies is found to correlate with redshift. A possible parametrization is that found by Delhaize et al. (2017)[1] as:

$$q_{\rm TIR}(z) = (2.88\pm0.03)(1+z)^{-0.19\pm0.01} \tag{4.3}$$

We can take advantage of this relation to estimate the AGN fraction as

$$f_{\rm AGN}(q,z) = 10^{q-\bar{q}(z)} \tag{4.4}$$

where q is the $q_{\rm TIR}$ measured for our galaxy and $\bar{q}(z)$ is the same quantity expected for a galaxy at the same redshift following the relation by Delhaize et al. (2017). Since in that study the $q_{\rm TIR}(z)$ relation is found to have an intrinsic scatter of 0.26,

[1]We underline that the analysis described in Delhaize et al. (2017) is based on the same radio survey employed to select the galaxies in this study.

we assign a $f_{\rm AGN} = 0$ to all the sources with $q - \bar{q} < 0.26$ (i.e. those compatible within 1σ with the relation expected for star-forming galaxies).

As visible in Figure 4.4, most of the galaxies in our sample have low values of the AGN fraction ($f_{\rm AGN} < 0.3$) and just 13 sources ($\sim 10\%$ of the full sample) are dominated by the AGN emission ($f_{\rm AGN} > 0.9$; see e.g. Enia et al. 2022) and — therefore — are removed from the sample. For all the other sources, we can account for the AGN contribution by computing the radio luminosity due to star formation as

$$L_{\rm 1.4GHz} = L_{\rm 1.4GHz,tot}(1 - f_{\rm AGN}) \tag{4.5}$$

This value is then used to compute the SFR following Novak et al. (2017):

$$\mathrm{SFR}[M_\odot\,\mathrm{yr}^{-1}] = 10^{-24} 10^{q(z)} \frac{L_{\rm 1.4GHz}}{\mathrm{W\,Hz^{-1}}} \tag{4.6}$$

by assuming the same $q_{\rm TIR}(z)$ relation found by Delhaize et al. (2017) and presented in Equation 4.3.

4.4 Radio Luminosity Function

We compute the radio luminosity function of our RS-NIRfaint galaxies by dividing our sample in three equally-populated redshift bins, covering the ranges $2.5 - 3.3$, $3.3 - 3.8$, and $3.8 - 5.5$. In each redshift bin, we divide the galaxies in several bins of (AGN-corrected) radio luminosity with a fixed width of 0.35 dex. To better sample the range of luminosities we employ bins that are overlapping by half of their width. In each combined redshift-luminosity bin, we compute the radio luminosity function following the $1/V_{\rm max}$ method (Schmidt, 1968) as:

$$\Phi(L, z) = \frac{1}{\Delta \log L} \sum_i \frac{1}{V_{\rm max,i}} \tag{4.7}$$

where $\Delta \log L$ is the log-amplitude of the luminosity bin and the sum extends to all the galaxies in that bin. For each galaxy, $V_{\rm max}$ is computed as

$$V_{\rm max} = \sum_{z=z_{\rm min}}^{z_{\rm max}} [V(z + \Delta z) - V(z)]\, C(z) \tag{4.8}$$

where we employ a fixed width of the redshift shells of $\Delta z = 0.05$. The $C(z)$ term is the completeness function and it can be written as the product of two terms:

$$C(z) = C_{\rm A} C_{\rm 3GHz}(z) \tag{4.9}$$

The first one accounts for the limited area covered by the COSMOS-Web survey. In our case it can be written as the ratio between the area of the survey and that of the whole sky:

$$C_{\rm A} = \frac{0.54\ \mathrm{deg}^2}{41253\ \mathrm{deg}^2} \tag{4.10}$$

The second term accounts for the incompleteness of the radio survey where our galaxies are selected. The completeness function of the 3 GHz survey performed

during the VLA-COSMOS large program is reported in Smolčić et al. (2017) for resolved and un-resolved sources. Since in our sample we have both kinds of sources, for each flux we consider the average between the two values.

The sum in Equation 4.8 extends between two values $z_{\rm min}$ and $z_{\rm max}$ indicating the minimum and maximum redshift in which our target would be included in our sample according to our selection criteria. As described in Section 4.2.3, our galaxies must have a radio S/N> 5 at 3 GHz and be faint in the F150W filter of NIRCam[2].

As shown in Figure 4.5, these properties are affected by the redshift of each object: at higher redshifts, our galaxy becomes fainter in the radio, up to a certain $z_{\rm 3GHz}$ where it becomes too faint to satisfy our S/N cut. Similarly, moving from high to low redshift, the F150W filter starts to sample less dust-obscured region of the SED, until a certain $z_{\rm NIR}$ where the galaxy becomes too bright to satisfy our magnitude cut ($F150W > 26.1$ mag). Both these effects are accounted for in the definition of $z_{\rm min}$ and $z_{\rm max}$ as:

$$\begin{cases} z_{\rm min} &= \max\left(z_{\rm min,bin}, z_{\rm NIR}\right) \\ z_{\rm max} &= \min\left(z_{\rm max,bin}, z_{\rm 3GHz}\right) \end{cases} \tag{4.11}$$

where $z_{\rm min,bin}$ and $z_{\rm max,bin}$ indicate the lower and upper limits of the considered redshift bin. The values $z_{\rm NIR}$ and $z_{\rm 3GHz}$ are estimated for each galaxy by redshifting the best-fitting rest-frame SED on a grid with a step of $\Delta z = 0.01$ and measuring the expected fluxes at 3 GHz and in F150W filter of NIRCam at each redshift.

To account for the uncertainties in the photo-z and radio fluxes at 3 GHz of our sources, we perform a Monte-Carlo integration by extracting 5000 random values from the probability distributions of these quantities. For the redshifts, we directly sample the Gaussianized $p(z)$ computed by `Cigale`, while for the radio fluxes we extract values from a Gaussian distribution centered on the mean value of the radio flux of each source and standard deviation equal to the related uncertainty. The value of the LF in each bin of redshift and radio luminosity is computed as the median value of all the iterations. Similarly, we assume as the upper and lower uncertainties the 16th and 84th percentiles, respectively. To account for the likely underestimation of the uncertainties in the less populated bins — due to the contribution of Poissonian uncertainties affecting low number counts — we correct these values in the bins with less than five galaxies with the 1σ confidence intervals reported in Gehrels (1986). Finally, we underline that our uncertainties do not include any contribution from cosmic variance. The values of the radio LF and the relative uncertainties are shown in Figure 4.6.

4.4.1 Modified Schechter function

In each redshift bin, we fit the LF of our sources with a modified Schechter function (Saunders et al., 1990), a common choice for radio and (sub)mm luminosity functions (see e.g. Novak et al. 2017; Gruppioni et al. 2020; Enia et al. 2022; van der Vlugt

[2]According to our selection, our targets must also have a counterpart in the COSMOS-Web catalogs. However, since all our galaxies are robustly detected in the LW filters of NIRCam, we neglect this additional constraint.

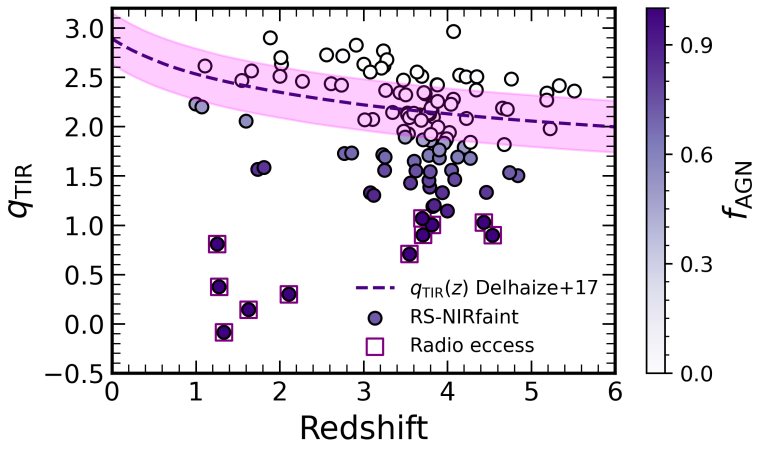

Figure 4.4: $q_{\rm TIR}$ parameter for our galaxies as a function of the redshift. The dashed line reports the $q_{\rm TIR}(z)$ relation by Delhaize et al. (2017), with the shaded area indicating the intrinsic scatter of the relation. Our galaxies are color-coded for their AGN fraction, measuring the contribution of nuclear activity to the radio luminosity. Galaxies classified as "radio-excess" (i.e. with $f_{\rm AGN} > 0.9$) are surrounded by a red square and removed from the sample.

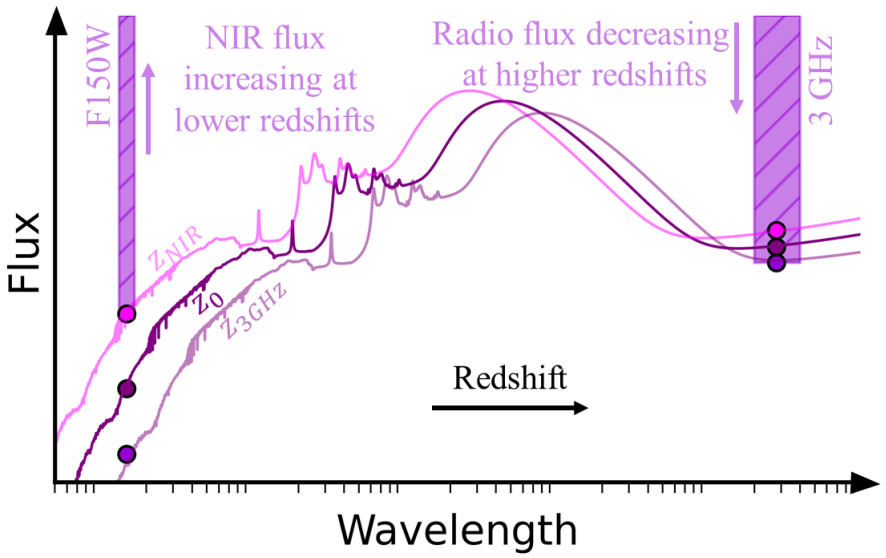

Figure 4.5: Sketch showing the physical meaning of the $z_{\rm min}$ and $z_{\rm max}$ values employed in the estimation of the LF. These values account for the redshift range in which each target could be found according to our selection criteria. See the definitions of $z_{\rm NIR}$ and $z_{\rm 3GHz}$ in Section 4.4

.

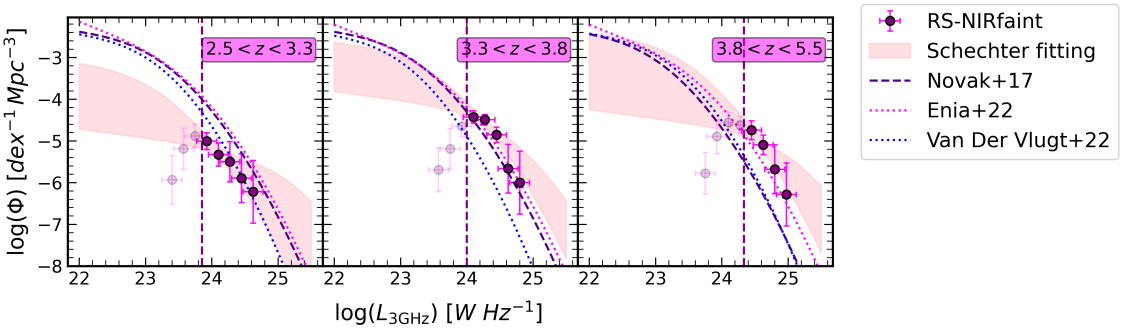

Figure 4.6: Radio luminosity function at 3 GHz of our RS-NIRfaint galaxies. The points show the values computed as described in Section 4.4, while the dashed vertical lines indicate the minimum radio luminosity observable in a given redshift bin given the sensitivity of our survey. All the points corresponding to fainter luminosities (shaded points) are not included in the fitting. The shaded areas report the fitting of the modified Schechter function, with its 1 σ uncertainty. For reference, we also report the radio luminosity function (converted to 3 GHz by assuming a standard radio slope of $\alpha = -0.7$) estimated by Novak et al. (2017), Enia et al. (2022), and van der Vlugt et al. (2022) on their samples of radio NIR-bright galaxies.

et al. 2022; Traina et al. 2024);

$$\Phi(L) = \Phi_\star \left(\frac{L}{L_\star}\right)^{1-\alpha} \exp\left[-\frac{1}{2\sigma^2}\log^2\left(1+\frac{L}{L_\star}\right)\right] \tag{4.12}$$

Due to the limited number of data-points available for fitting and the lack of constraints on the faint-end of the LF, we fix the two slope factors (α and σ) to the values found by Novak et al. (2017) in their analysis of the sample of radio sources (with optical/NIR counterparts) in the VLA-COSMOS large program. These values are $\sigma = 0.63$ and $\alpha = 1.22$. We underline that the latter value is also in agreement with that found by van der Vlugt et al. (2022) in the deeper COSMOS-XS survey, better sampling the faint end of the radio LF. Possible caveats related to this choice are discussed in Section 4.5.2.

We fit the two remaining free parameters ($\Phi_\star$ and $L_\star$, giving the normalization and the "knee" of the LF, respectively) with a Monte-Carlo Markov Chain (MCMC) performed with the Python library `emcee` (Foreman-Mackey et al., 2013). We adopt a flat prior on the two parameters: $\log(\Phi_\star) \in [-6, -2]$ Mpc^{-3} dex^{-1} and $\log(L_\star) \in [20, 26]$ W Hz^{-1}. We only include in the fitting procedure the points with $L > L_{\rm min}$, where $L_{\rm min}$ represents the minimum radio luminosity observable in a redshift bin given the 5σ sensitivity of our radio survey. Moreover, since the LF is computed in overlapping bins of radio luminosity, we only include in the fitting half of the data-points to obtain un-correlated uncertainties.

The best-fitting parameters of the modified Schechter function — for each redshift bin — are reported in Table 4.2, while the fitted function (with its 1σ confidence interval) is reported as the shaded pink area in Figure 4.6.

Table 4.2: Best-fitting parameters of the modified Schechter function fitted to our radio luminosity function at 3 GHz. As a reference, we report the same values obtained by Enia et al. (2022) in the GOODS-N field considering NIR-bright galaxies. Their luminosities (originally at 1.4 GHz) are converted assuming a standard radio slope $\alpha = -0.7$.

	This work		Enia+22	
z	$\log(\Phi_\star)$ [Mpc^{-3} dex^{-1}]	$\log(L_\star)$ [W Hz^{-1}]	$\log(\Phi_\star)$ [Mpc^{-3} dex^{-1}]	$\log(L_\star)$ [W Hz^{-1}]
$2.5 < z < 3.3$	$-4.4^{+0.8}_{-1.1}$	$23.1^{+0.7}_{-1.2}$	$-2.6^{+0.5}_{-0.5}$	$22.7^{+0.3}_{-0.3}$
$3.3 < z < 3.8$	$-3.5^{+0.6}_{-0.8}$	$23.1^{+0.4}_{-0.6}$	$-2.7^{+0.4}_{-0.5}$	$22.8^{+0.3}_{-0.2}$
$3.8 < z < 5.5$	$-3.6^{+1.1}_{-1.0}$	$23.2^{+0.5}_{-1.0}$		

Table 4.3: Contribution of the RS-NIRfaint galaxies to the cosmic SFRD. These values are obtained by integrating the LF in the range of radio luminosity covered by our observations ("observed") and in the full range of luminosities ("extrapolated").

	$\log(\rho_{\rm SFR})$ [$\times 10^{-3}$ M$_\odot$ yr^{-1} Mpc^{-3}]	
z	(Observed)	(Extrapolated)
$2.5 < z < 3.3$	$1.3^{+0.7}_{-0.4}$	6^{+14}_{-3}
$3.3 < z < 3.8$	8^{+2}_{-2}	30^{+34}_{-9}
$3.8 < z < 5.5$	4^{+2}_{-1}	30^{+58}_{-17}

4.5 Contribution to the cosmic SFRD

In each redshift bin, we compute the contribution of the RS-NIRfaint galaxies to the total cosmic SFRD by integrating the analytic expression of the modified Schechter function weighted for the SFR related to each radio luminosity:

$$\mathrm{SFRD}(z) = \int_{L_{\mathrm{min}}}^{L_{\mathrm{max}}} \Phi(L, z)\, \mathrm{SFR}(L, z)\, d \log L \tag{4.13}$$

where $\mathrm{SFR}(L, z)$ is the expression reported in Equation 4.6 and z is the mean redshift of each bin. In this integral, we consider the full posterior distribution of the parameters in the modified Schechter function obtained from the MCMC in Section 4.4.1 in order to compute the uncertainties on the SFRD.

We perform two different integrals. The first one only includes the radio luminosities between the L_{min} related to the sensitivity of our radio survey and the maximum luminosity observed in each redshfit bin. The second extends over the full range of radio luminosities ($0 \rightarrow +\infty$). While the first integral estimates the contribution of the galaxies that are actually observed in our surveys to the cosmic SFRD, the second one includes in this estimates the galaxies that are not observed in our survey but that we expect by extrapolating the radio LF at higher and lower luminosities. We underline that — given these definitions — the first value should be interpreted as a lower limit on the actual contribution of the RS-NIRfaint galaxies to the cosmic SFRD. The results of these integrals are summarized in Table 4.3 and in Figure 4.7. It is possible to notice how our estimates have large uncertainties. The reason for this can be found in the poor constraints on the $L_{\star}$ parameter of the LF. Being our galaxies unable to trace the faint end of the LF, our analysis can only pose a robust upper limit on the location of the "knee". Deeper radio data would be needed to better constrain this value (and, consequently, the contribution to the SFRD).

4.5.1 How much do the RS-NIRfaint galaxies contribute to the cosmic SFRD?

The results obtained in Section 4.5 indicate an increasing contribution of the RS-NIRfaint galaxies to the cosmic SFRD from $z \sim 3$ to $z \sim 3.5$ and then decreasing until $z \sim 4.5$. This result agrees with previous studies focusing on NIRdark/faint radio sources at $z > 2.5$. Our "observed" results are consistent (even though moderately lower at $z \sim 3$) with those obtained by Behiri et al. (2023), analyzing the same kind of sources in the COSMOS field (even though with a slightly different selection, see Section 4.2.3). In that study, no extrapolation was performed outside the observed range of luminosities, therefore it is not surprising that our "extrapolated" results are significantly higher. Similarly, our estimation of the SFRD at $z \sim 3$ agrees well with the lower limit presented in Enia et al. (2022) when no extrapolation is performed.

Our "extrapolated" results can be compared with the upper limit provided by Enia et al. (2022) (analyzing NIRdark galaxies in the GOODS-N field) and with the estimates by van der Vlugt et al. (2023) (analyzing analogous sources in the deep

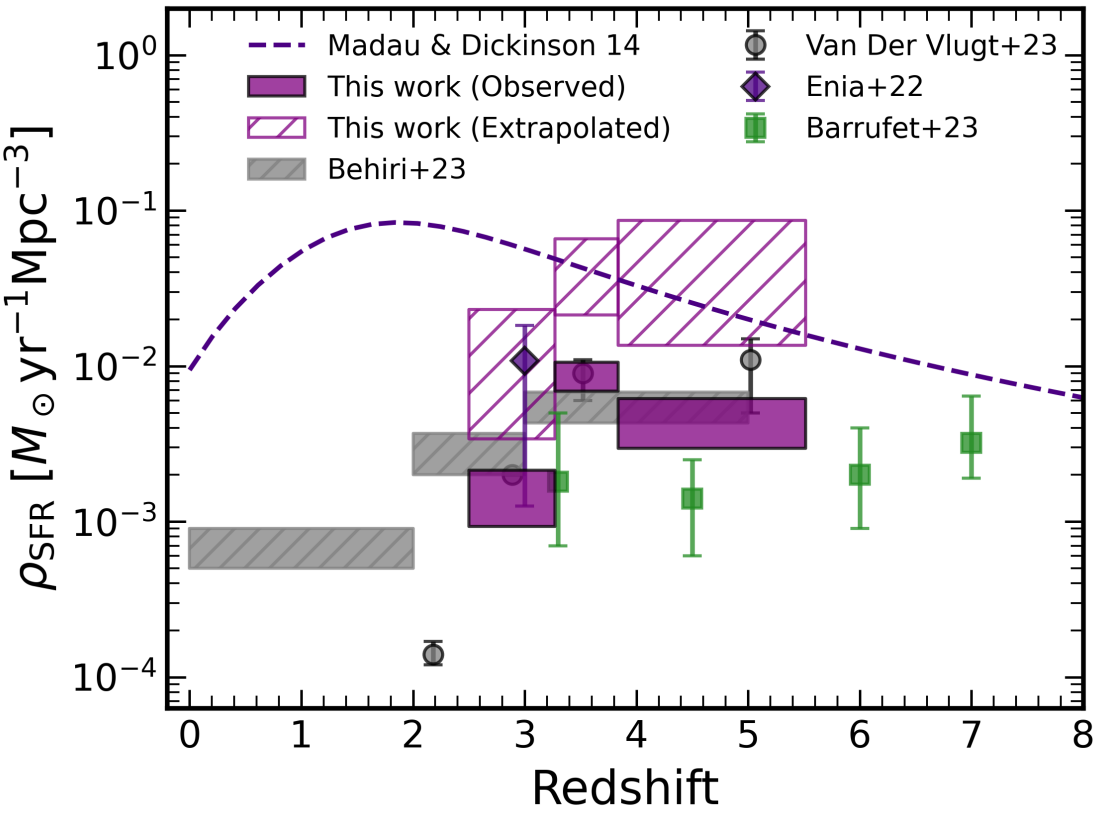

Figure 4.7: Cosmic Star Formation Rate Density. The values obtained by integrating the luminosity function of our RS-NIRfaint galaxies are reported in purple. The values obtained by integrating the observed range of luminosities are shown with full squares, while the empty boxes report the values obtained by integrating the LF on the full range of radio luminosities. The figure shows, for reference, the SFRD computed by Madau & Dickinson (2014) on "optically-bright" galaxies (dashed indigo line), and those obtained by Behiri et al. (2023), van der Vlugt et al. (2023), Enia et al. (2022), and Barrufet et al. (2023) (grey boxes, purple diamond, grey points, and green squares, respectively) for their "NIRfaint galaxies" selected in the radio and with JWST.

COSMOS-XS survey). Both these studies estimated the radio LF and computed the contribution to the cosmic SFRD by integrating it in the full range of radio luminosities ($0 \rightarrow \infty$). It is important to notice, however, that the three studies estimate the contribution to the total SFRD in different ways. The other two studies, indeed, focused on smaller fields compared with COSMOS-Web (0.05 deg^2 in GOODS-N and 0.1 deg^2 in COSMOS-XS, instead of the 0.54 deg^2 covered by our survey). This difference produced significantly smaller samples of NIRdark galaxies available for the analysis (9 sources for Enia et al. 2022 and 20 for van der Vlugt et al. 2023, against the $\sim$ 120 analysed here). To overcome this issue, Enia et al. (2022) fixed three parameters in the Schechter fitting of the LF (σ, α, and $L_\star$) to those obtained for the NIR-bright population, leaving only the normalization ($\Phi_\star$) as a free parameter of the fitting. Likewise, van der Vlugt et al. (2023) computed the LF on the full sample of radio sources twice: once including the NIRdark galaxies and once excluding them: the contribution of these sources was then found by subtracting the two inferred SFRDs.

Our results at $z \sim 3.5$ agree well with the upper limit at the same redshift by Enia et al. (2022). Similarly, our result in the highest redshift bin ($z \sim 4.5$) is compatible with the estimate by van der Vlugt et al. (2023). Their results are more in tension with ours at $z \sim 3$ and $z \sim 3.5$, with our estimates being up to one order of magnitude higher. A possible explanation of this discrepancy could reside in the accuracy of the photo-z (our photometry includes the new NIRCam and MIRI photometry, that could result in more accurate redshift estimates; see e.g. Barrufet et al. 2024).

4.5.2 Possible caveats of our analysis

Our analysis is not immune to possible biases. The most common one is an inaccuracy in the photometric redshifts estimated through SED fitting. Even with the new constraints obtained thanks to NIRCam and MIRI, the small number of photometric detections makes our estimates uncertain. We reduced the possible impact of this issue on our results by employing in our analysis the full Gaussianized $p(z)$ given in output by `Cigale`, resulting in larger uncertainties on our LF and SFRD. This issue could be solved in the future with a spectroscopic follow-up of our galaxies (analogous to that performed by Barrufet et al. 2024 with JWST, or those performed by Jin et al. 2019; Chen et al. 2022; Jin et al. 2022 or in Chapter 3 with ALMA).

Another caveat (more related to our radio selection) resides in the possible AGN contribution not unveiled by our $q_{\rm TIR}$ analysis. An overestimated radio luminosity due to star-formation could still bias our LF and SFR, producing a different behavior of the SFRD. This issue can be partially solved with a more complete photometric coverage at MIR wavelengths, where the thermal emission of a hot dusty torus surrounding the AGN should be visible (see e.g. Hickox & Alexander 2018 for a review and Chien et al. 2024 for a recent application with MIRI data).

Finally, our analysis (especially the "extrapolated" value of the SFRD) relies on a strong assumption about the shape of the LF. In our fitting procedure, we fixed the two parameters of the modified Schechter function α and σ to those obtained by Novak et al. (2017) for their sample of radio sources in the COSMOS field (i.e. the parent sample of our selection). This choice, however, assumes that the selection

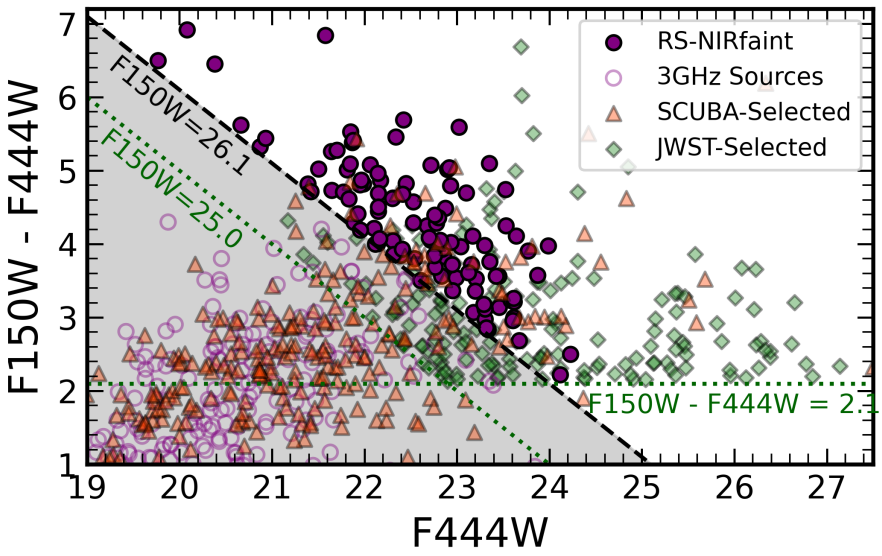

Figure 4.8: Our RS-NIRfaint galaxies in the F150W-F444W vs F444W color-magnitude plot. We report, for comparison, the JWST-selected DSFGs by Gottumukkala et al. (2023) and the SCUBA-selected ones by McKinney et al. (2024). Even if we do not pose any constrain in the F150W-F444W color, we obtain that all our sources are above the threshold adopted by Gottumukkala et al. 2023, identifying our galaxies as a sub-population of the JWST-selected DSFGs. Similarly, part of the SCUBA-selected DSFGs satisfy our "NIRfaint" criterion.

of more dust-obscured sources is not altering these parameters. This assumption could be incorrect, since we expect the number of NIRfaint sources to be lower in the low-SFR regime (e.g. for a lower number of *supernovae* and — consequently — a potential lower production of cosmic dust) and higher in the high-SFR one (see e.g. Whitaker et al. 2017 and A. Traina et al., *in prep.*). This issue could be improved with a deeper radio survey (e.g. van der Vlugt et al. 2022) giving us more information about the faint-end slope of the LF for our sources.

4.6 Discussion

4.6.1 What is the effect of the selection for NIRfaint sources?

To interpret the results presented in the previous sections, it is useful to compare our population of galaxies with other samples studied in the current literature. The first one is the total population of radio-detected sources, of which the RS-NIRfaint galaxies are a sub-population with $F150W > 26.1$ mag. One of the main results of this comparison is shown in Figure 4.2, where we report the photo-z estimated for our galaxies and those computed by M. Shuntov, L. Paquereau et al., (*in prep.*) for the other 3 GHz sources in the COMSOS-Web survey with $F150W > 26.1$ mag. It is immediately possible to see that our galaxies are — on average — located at higher redshifts with respect to the rest of the population. An analogous result has

been shown by van der Vlugt et al. (2023) for their NIRdark galaxies selected in the deep radio survey COSMOS-XS and by Talia et al. (2021) and Behiri et al. (2023) in their previous analysis of RS-NIRdark galaxies in the COSMOS field. The same idea that NIRfaint DSFGs represent the high-redshift end of their parent distributions is also present in other studies selecting DSFGs at other wavelengths (see e.g. the studies on NIRfaint sub-millimeter galaxies by Simpson et al. 2014; Franco et al. 2018; Smail et al. 2021; McKinney et al. 2024).

The different distribution of the photo-z also affects the estimated luminosity functions. Here, we consider those by Novak et al. (2017), Enia et al. (2022), and van der Vlugt et al. (2022) focusing on NIR-bright galaxies. These estimates are reported for reference in Figure 4.6, while the best-fitting parameters of the modified Schechter function found by Enia et al. (2022) are reported in Table 4.2. It is possible to notice how the normalization of the luminosity function for our NIRfaint galaxies is on average $\sim 1.5-2$ dex lower than what is observed for optically-bright sources. The luminosity of the knee, instead, seems to be higher for our sources, even though compatible with the values from Enia et al. (2022) within the (large) uncertainties.

The first result is not unexpected, since we are dealing with a population of heavily dust-obscured sources less common than NIR-bright galaxies (see e.g. the comparison between the number of 3 GHz sources in COSMOS-Web and those in our sample in Section 4.2). The second result is more interesting, since it suggests that the high dust obscuration could positively correlate with the radio luminosity (and — therefore — with the SFR). This result is not completely new, since several previous studies (e.g. Whitaker et al. 2017 and Traina et al., *in prep.*) highlighted how the dust-obscured star formation could dominate the high-SFR end of the star formation rate function, at least until the cosmic noon. The main consequence of the higher turn-off luminosity can be seen in Figure 4.6, where is visible how the NIRfaint galaxies could dominate the LF in its bright end (even though, again, still compatible in most cases within the estimated uncertainties). The higher volume densities of our sources in the bright end also explains the large contribution to the SFRD when we integrate the LF on the whole range of radio luminosities to compute the SFRD (Figure 4.7).

4.6.2 What is the effect of the radio selection?

Once confirmed that the radio selection is able to produce a sample of DSFGs, we want to analyse how this method compares with analogous strategies present in the current literature. In this section, we will focus on two main approaches: the selection by Barrufet et al. (2023) and Gottumukkala et al. (2023) based on the JWST colors, and that by McKinney et al. (2024) based on the (sub)mm detection.

JWST-selected DSFGs

We firstly compare our sources with those collected by Barrufet et al. (2023) and Gottumukkala et al. (2023) in the Cosmic Evolution Early Release Science (CEERS; Finkelstein et al. 2023). These studies mimic the DSFG selection initially performed by Wang et al. (2019) with the so-called "H-dropout" (i.e. galaxies selected through their H - [4.5] colors) by taking advantage of the new JWST photometry. More in

detail, this selection couples the F150W-F444W> 2.1 mag color cut with a magnitude cut in the F150W filter (F150W> 25.0 mag in Gottumukkala et al. 2023). As visible in Figure 4.8, our RS-NIRfaint galaxies represent a sub-sample of the sources analysed by Gottumukkala et al. (2023), satisfying — by construction — the same criteria about faintness in F150W, having comparable F150W-F444W colors, but with the additional requirement of the radio detection. By comparing these samples, we can study how the radio selection affects the retrieved population of galaxies. Since the selection by Gottumukkala et al. (2023) includes fainter sources in the F150W filter, in the following we will focus on their sources satisfying our F150W>26.1 criterion.

The first notable difference is in the projected sky density of the two populations 6000 ± 77 deg^{-2} for the sources by Gottumukkala et al. (2023) and 205 ± 14 deg^{-2} for our RS-NIRfaint galaxies[3]. Then, by looking at the physical properties summarized in Figure 4.3, we can see that the distributions of the photo-z are quite similar until $z \sim 4$, but with different behavior at higher redshifts. This result is not surprising given the much stronger k-correction in the radio than in the F444W filter of NIRCam (see e.g. the shape of the SEDs in Figure 4.1), biasing our sample towards lower-z sources.

Another notable difference consists in the reported values of SFR. As visible in Figure 4.3, our RS-NIRfaint galaxies are ~ 1.5 dex more star-forming than the sources analysed by Barrufet et al. (2023) and Gottumukkala et al. (2023). On the one hand, this result is not surprising since the need for a radio detection excludes from our sample all the objects with low values of SFR. As noted by Wang et al. (2019) and verified by Barrufet et al. (2024) with a spectroscopic follow-up of some JWST-selected objects, the red NIRCam colors alone are not enough to select DSFGs, since the same colors can be seen in quiescent galaxies. This ambiguity in the true nature of these sources can be solved with additional data from JWST (mainly spectra, as in Barrufet et al. 2024), or a more complete MIRI coverage (as in Pérez-González et al. 2023), ALMA (as in Wang et al. 2019) or in the radio regime (as for our sources).

If the radio selection well explains why our sample lacks low-SFR galaxies, it does not explain why the high-SFR sources are not present in the sample by Gottumukkala et al. (2023). A possible explanation resides in the photometric coverage of their sample. At the redshifts covered by their observations, no constraints on the rest-frame UV is available for the SED fitting, making quite uncertain the value of the SFR obtained through it. Similarly, we expect most of the star formation in these NIRfaint galaxies to be dust-obscured and — therefore — hard to constrain without some information on the FIR or radio emission, producing a likely underestimation of the SFR (see, e.g., some examples in Xiao et al. 2023).

The higher values of the SFR also explain one of the results visible in Figure 4.7: even though the RS-NIRfaint galaxies are ~ 1.5 dex less common than the JWST-selected sources, the higher values of SFR (~ 1.5 dex) still produce compatible contribution to the cosmic SFRD for the "observed" sources. The values obtained by extrapolating at higher and lower radio luminosities are obviously higher than those reported by Barrufet et al. (2023), since their estimate only accounts for the

[3]For both estimates, we assume a Poissonian uncertainty $\sim \sqrt{N}$ on the observed number counts and do not correct for incompleteness

detected sources.

SCUBA-selected DSFGs

Another possibility to select DSFGs relies on the detection at (sub)mm wavelengths of the bright thermal emission by warm dust heated by ongoing star formation (see e.g. Casey et al. 2014 for a review). Galaxies selected through this procedure are commonly known as sub-millimeter galaxies (SMGs). Since this definition strongly relies on the depth of the observations employed to select these objects, in this study we only focus on the sources described in McKinney et al. (2024), acknowledging that some differences might arise when other samples (detected in deeper or shallower surveys; see e.g. da Cunha et al. 2015; Dudzevičiūtė et al. 2021) are considered. These 289 galaxies were initially selected in the COSMOS field as sources with $S_{870\mu\mathrm{m}} > 2$ mJy in the SCUBA-2 observations performed during the S2COSMOS survey (Simpson et al., 2019) and then followed-up with ALMA (unveiling in some cases the presence of multiple fainter sources contributing to the total flux of the SCUBA-2 objects). The SCUBA-Dive program described in McKinney et al. (2024) performs a deeper analysis of these sources taking advantage of the new JWST data coming from the COSMOS-Web survey.

Looking at the color-magnitude plot in Figure 4.8, we can note only a partial overlap between the SCUBA-selected galaxies and our RS-NIRfaint ones, since most of the former are brighter (i.e. not NIRfaint) in the F150W filter. Since both samples are selected in the COSMOS field, we can cross-matching the two catalogs, finding only a partial overlap of 27 objects. We underline that the SCUBA-Dive programs only focuses on the SCUBA-2 sources with previous ALMA data, therefore some of our galaxies could still be included in that sample with additional data. To overcome this issue, we analyze the best-fitting SEDs computed with `Cigale` for our galaxies, obtaining that only 82 ($\sim 65\%$) of our sources would have a $S_{850\mu m} > 2$ mJy, being consistent with the initial cut employed on the SCUBA-2 maps by McKinney et al. (2024).

By comparing the physical properties estimated through SED fitting (Figure 4.3), we can see that the distributions of stellar masses and SFR are quite similar. More significant differences hold for the photo-z (with our galaxies located — on average — at higher redshifts: $\langle z \rangle \sim 3.6$ for our sources and $\langle z \rangle \sim 2.6$ for those in SCUBA-Dive) and for the dust attenuation (with our sources being — on average — more dust-obscured; $\langle A_\mathrm{v} \rangle \sim 3.5$ mag against $\langle A_\mathrm{v} \rangle \sim 2.5$ mag). Both results are easily explained by the NIRfaint selection, biasing the sample towards more dust-obscured and higher-z objects (see also the discussion in Section 4.6.1).

A further confirmation of this result can be found in the comparison with the SCUBA-Dive sources satisfying the same magnitude cut as our RS-NIRfaint galaxies ($F150W > 26.1$ mag). As visible from Figure 4.3, these sources have analogous properties to our galaxies, indicating that the two selections are able to identify the same population of objects with different procedures.

4.7 Chapter summary

In this chapter, I presented the first analysis taking advantage of JWST data of RS-NIRfaint sources in the COSMOS field. These sources are defined as radio-detected sources with a counterpart at NIR wavelengths (unveiled with the deep NIRCam observations) fainter than the common depths reached by ground-based facilities. Together with my collaborators, I obtained the following results:

- The physical properties estimated through SED fitting indicate that the RS-NIRfaint galaxies are a population of highly dust-obscured ($\langle A_{\rm v} \rangle \sim 3.5$ mag), massive ($\langle M_\star \rangle \sim 10^{10.8}$ M$_\odot$) and star-forming galaxies ($\langle {\rm SFR} \rangle \sim 300$ M$_\odot$ yr^{-1}) located at $\langle z \rangle \sim 3.6$.

- Estimating the radio luminosity function of our sources and fitting it with a modified Schechter function, we find that its normalization ($\Phi_\star$) is $\sim 1.5-2$ dex lower than that computed on radio-selected NIR-bright galaxies (e.g. Novak et al. 2017; Enia et al. 2022; van der Vlugt et al. 2022), indicating that — as expected — DSFGs with faint optical/NIR counterparts are a rare population.

- Interestingly, the knee of the LF for our sources is brighter than for NIR-bright sources (even though compatible within the large estimated uncertainties). This result suggests that the contribution of NIRfaint sources is negligible in the low-SFR end of the star formation rate function, while it becomes dominant in the high-SFR end, at least until $z \sim 4.5$. This result confirms what has been found in previous studies focusing on the ratio between obscured and unobscured star formation in high-z galaxies.

- By integrating the LF of our sources in the range of radio luminosities covered by our observations, we put a lower limit on their contribution to the cosmic SFRD. Our result shows an increasing contribution from $z \sim 3$ to $z \sim 3.5$ and then decreasing until $z \sim 4.5$.

- By integrating the LF in the full range of radio luminosities ($0 \rightarrow \infty$; i.e. extrapolating at higher and lower radio luminosities than what is actually observed), we increase the contribution to the total cosmic SFRD, reaching the same level of the NIR-bright galaxies analysed by Madau & Dickinson (2014).

- When compared with the JWST selection of DSFGs carried out by Barrufet et al. (2023) and Gottumukkala et al. (2023), we obtain that our radio selection generally misses the sources located at higher redshifts ($z > 5.5$), as expected given the positive k-correction affecting the radio emission. Moreover, our sources are ~ 1.5 dex more rare, and ~ 1.5 dex more star-forming. The main consequence of these differences, is that their contribution to the cosmic SFRD (not extrapolating at higher and lower radio luminosities) is compatible with that estimated by Barrufet et al. (2023) up to $z \sim 5$.

- Finally, we compare our sources with the SCUBA-selected ones found in the COSMOS-Web survey by McKinney et al. (2024), obtaining that our galaxies

have analogous properties to their sources in terms of stellar mass and SFR. Our RS-NIRfaint galaxies are — on average — located at higher redshifts and more dust-obscured, as expected for our requirement about the faintness at NIR wavelengths. Comparing our sources with the SCUBA-selected galaxies satisfying the same $F150W > 26.1$ mag requirement, we obtain more similar distributions of photo-z and $A_{\rm v}$.

Chapter 5

Going beyond: Looking for massive galaxies with JWST

5.1 Introduction to the chapter

In this chapter, I present a complementary work to those based on the RS-NIRdark/faint galaxies described in the previous chapters. The main goal is still the research for the progenitors of the most massive galaxies observed in the local Universe and constraining the high-z behavior of star formation. Nevertheless, in this chapter I follow a different approach. While earlier in this Thesis I focused on a full population of galaxies whose evolutionary path could bring to the formation of the massive and passive galaxies at $z \sim 3$ and at $z \sim 0$, here I focus on some extremely massive galaxies at $z \sim 5-7$ discovered in the first JWST data of the COSMOS-Web survey. These galaxies are tightly related to the DSFGs presented in the other chapters (one is also included in the sample studied in Chapter 4) and represent likely progenitors of the most massive galaxies observed in our Universe. Moreover, as shown in this chapter, the existence of these extreme sources is puzzling for several reasons, potentially giving some insights on the efficiency of star formation in the early Universe. The chapter is structured as follows. After a brief scientific introduction on the first significant results obtained by JWST in the field of galaxy evolution, I present the two galaxies that are the subjects of this chapter and the data available for them (Section 5.3). Then, in Section 5.4 and Section 5.5, I describe the SED-fitting and morphological analysis performed on these sources and — in Section 5.6 — the possible consequences of our results. Finally, I present my conclusions in Section 5.7.

Most of the analysis presented here is also included in "*Not-so-little Red Dots: Two massive and dusty starbursts at $z \sim 5-7$ pushing the limits of star formation discovered by JWST in the COSMOS-Web survey*" by Gentile et al. (2024c), published in The Astrophysical Journal Letters (Volume 973, Issue 1, id.L2, 8 pp.).

5.2 Scientific introduction

Since its launch, JWST has dramatically changed our perspective on the high-z Universe. The discovery of a significant population of highly red and compact sources

at $z \gtrsim 5$ ("little red dots"; LRDs henceforth) is puzzling in their origin. These objects have been found in most of the first JWST surveys through their color in the long-wavelength (LW) filters of NIRCam (mainly F277W and F444W; see e.g. Labbé et al. 2023a,b; Akins et al. 2023; Matthee et al. 2023; Kokorev et al. 2024; Pérez-González et al. 2024). These analyses have shown that LRDs are predominately at $z \gtrsim 5$ (with few lower-z analogs; see e.g. the low fraction of Extremely Red Objects at $z < 5$ reported by Barro et al. 2023), are spatially compact, and some have a seemingly decoupled blue SED rising into the rest-UV (e.g. Labbé et al. 2023a,b; Kokorev et al. 2024; Pérez-González et al. 2024). Moreover, the first spectroscopic follow-ups showed a significant presence of broad emission lines in their spectra, suggesting the presence of Active Galactic Nuclei (AGN) in their cores with a high $M_{\rm BH}/M_\star$ ratio (Kocevski et al., 2023; Furtak et al., 2023; Greene et al., 2023; Matthee et al., 2023).

Still unclear is the main physical process responsible for the red continuum: is it highly-obscured stellar emission (e.g. Labbé et al. 2023a; Akins et al. 2023; Xiao et al. 2023; Pérez-González et al. 2024; Williams et al. 2023) or reddened radiation emitted by the AGN (e.g. Labbé et al. 2023b; Greene et al. 2023; Kokorev et al. 2024)?

Modeling the spectral energy distribution (SED) of these galaxies with stellar-only templates, one can obtain near "Universe-breaking" stellar masses (up to $\sim 10^{11}$ $M_\odot$ at $z \sim 8$; see e.g. Labbé et al. 2023a). These values were in tension with the current cosmological model (e.g. Boylan-Kolchin, 2023), unless (after some downward revisions) we assume that the star formation in the high-z Universe is much more efficient than what is commonly observed at lower-z (see e.g. Xiao et al. 2023). Of course, assuming that the reddened emission comes from an AGN accretion disk lowers the stellar mass estimates, but a new problem arises when the number density of these LRDs is compared with the quasar luminosity function (e.g. Shen et al., 2020), as LRDs are 100-1000 times more common, implying an overabundant population of massive black holes at early times (Greene et al., 2023).

One of the main reasons for the uncertainty about the dominant continuum emission in LRDs is their compactness, not allowing us to completely disentangle the nuclear and stellar contribution. For this reason, in this chapter, I set out to find sources that fulfill the color selection of LRDs but have unambiguously extended morphologies with no embedded point source emission. By virtue of their spatial extent, one can attribute their continuum to stellar origin and – therefore – obtain AGN-unbiased estimation of the stellar masses. In this chapter, I report the discovery of two candidate massive ($M\star \sim 10^{11}$ $M_\odot$) galaxies (one of which already has a spectroscopic redshift from the literature) in the COSMOS-Web survey.

5.3 Data

5.3.1 JWST Photometry from COSMOS-Web

The NIRCam and MIRI photometry for these sources comes from the COSMOS-Web survey already presented in Chapter 4. The space-based photometry employed here is extracted with `sourceXtractor++` (`SE++` in the following; Bertin et al. 2020;

Kümmel et al. 2020), a model-based tool for extracting photometry from datasets with different spatial resolution. Detection is conducted on a χ^2-image (Szalay et al., 1999) generated by co-adding the four NIRCam filter maps. A double Sérsic profile (Sérsic, 1963) is fitted to the four NIRCam filter images simultaneously. Further details about the COSMOS-Web catalogs will be included in Shuntov, Paquereau et al., (*in prep.*).

By default, `SE++` photometric uncertainties do not include Poisson noise from the background (see e.g. the discussion in Akins et al. 2023 and Casey et al. 2024). Therefore, we measure the background noise in empty circular apertures and add it in quadrature to the uncertainties estimated by `SE++`, following Casey et al. (2024).

5.3.2 Initial sample selection

The goal of this work is to find spatially-resolved sources with analogous colors to the LRDs (e.g. Labbé et al. 2023a,b; Akins et al. 2023; Kokorev et al. 2024; Pérez-González et al. 2024). Therefore, the initial sample selection consists of a color cut to collect all the reddest sources in the survey area. We adopt the same criterion as in Barro et al. (2023) and analogous to that employed by Akins et al. (2023):

$$F277W - F444W > 1.5\,\mathrm{mag} \tag{5.1}$$

To reduce the possible contamination by fake sources and focus only on the robustly detected galaxies, we couple the color cut with three additional criteria:

$$\begin{cases} F444W < 27.5\,\mathrm{mag} \\ S/N(F277W) > 2 \\ S/N(F444W) > 4 \end{cases} \tag{5.2}$$

We do not include any S/N requirement in F115W and F150W, since we expect most of the high-z or highly dust-obscured sources to dropout at these wavelengths. Moreover, the depth of COSMOS-Web in the short-wavelength (SW) filters is slightly shallower those of other surveys where LRDs have been detected (see e.g. Bezanson et al. 2022; Finkelstein et al. 2023). Therefore, we do not expect significant emission in these filters even with a similar SED. Moreover, as noted by Barro et al. (2023), the inclusion of the additional constrain on the blue colors of LRDs seems to bias the sample towards high-z objects. Without this criterion, we aim at focusing on the full redshift distribution of these sources. In addition, Akins et al. (2024) – also selecting (unresolved) LRDs in the COSMOS-Web survey with the same color cut employed here – showed how this selection can produce a sample of galaxies with properties analogous to those obtained by adding constraints on the blue shape of the SED. All these criteria produce an initial sample of $\sim$ 450 sources, after a first vetting to remove imaging artifacts.

This sample includes a large sample of spatially-compact LRDs (Akins et al., 2024). To focus only on the resolved objects, we perform an initial profile fitting on the F444W maps (i.e., those in which we expect our galaxies to have the highest S/N), forcing the modeling with an unresolved point source (in doing so, we employ the empirical PSF computed on the scientific mosaics through the software `PSFEx` and the standard routines included in the `Astropy` library; Bertin 2011). We exclude

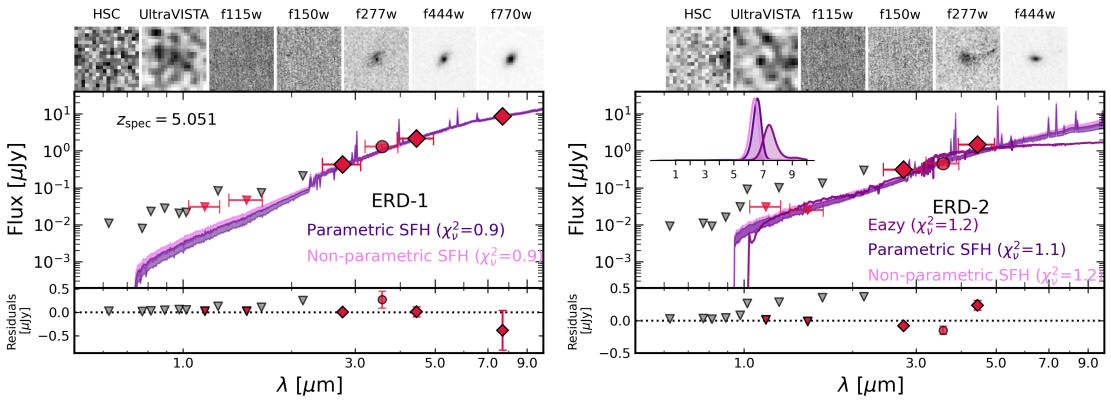

Figure 5.1: Cutouts (3" × 3") and best fit spectral energy distributions for ERD-1 and ERD-2. Points have $S/N > 3$, while downward-pointing triangles represent 2σ upper limits. Photometry from JWST is in red, while other constrains are shown in gray. The first two cutout panels represent stacked images of the g, r, i, z, y bands of the HyperSupremeCam and the Y, J, H, Ks filters of the VISTA telescope, respectively. The inset on the right panel shows the redshift probability distributions for ERD-2 estimated by `Bagpipes` (with both SFHs) and `Eazy`. Shaded regions mark 16th-84th percentile confidence intervals on SED fits. The lower panels report the normalized residuals computed on the `Bagpipes` SED with parametric SFH.

all galaxies that are well-described as an unresolved point source (i.e. with a reduced $\chi^2 \sim 1$ and no significant residual once subtracted the unresolved model). The final sample of galaxies consists of 61 objects. The whole sample – including new objects found in the full 0.54 deg^2 – will be described in a forthcoming paper (Gentile et al., *in prep.*), and here we focus on the two unique sources among the 61, selected for being particularly high redshift with high stellar mass estimates (both of which are described in detail below). Both galaxies, which we named ERD-1 and ERD-2 (for "extended red dots"), have $F277W - F444W \sim 1.75$, just slightly lower than the color cut by Akins et al. (2023), but still significantly redder than the nominal LRD selection (with $F277W - F444W > 1$ mag; see e.g. Labbé et al. 2023a; Greene et al. 2023; Pérez-González et al. 2024).

5.3.3 Ancillary data

We extract additional photometry for our sources, using the wealth of other ground/space imaging in COSMOS. We extract fluxes in fixed circular apertures from the datasets in Weaver et al. (2022). These include the Subaru data (HyperSuprimeCam Subaru Strategic Program, HSC SSP DR3; Aihara et al. 2019), VISTA data (UltraVISTA survey DR5; McCracken et al. 2012), *Spitzer Space Telescope* data (Cosmic Dawn Survey; Moneti et al. 2022), and HST data (Koekemoer et al., 2007). We use a fixed aperture radius of 1", well-matched to seeing limits of the ground based-data. For the *Spitzer* data, we only include the photometric point at 3.6 μm as the second channel is superseded by F444W. We calculate aperture corrections by convolving the best-fitting model by `SE++` with the IRAC PSF and calculating what fraction of the total flux is missed.

5.3.4 ALMA data on ERD-1

ERD-1 has a reported spectroscopic redshift $z_{\rm spec} = 5.051$ in the literature from Jin et al. (2019) thanks to an ALMA spectral scan. They find a single line at 95.2 GHz (5.97σ). The *lack* of other lines identified in the broad frequency range lead to the conclusion the line is CO(5-4). As noted in Section 5.4, our new JWST photometry is inconsistent with lower redshift explanations (e.g. CO(4-3) at $z = 3.84$ or CO(3-2) at $z = 2.63$). Jin et al. (2019) measure a (sub)mm flux density at 3 mm of $S_{3\rm mm} = (0.115 \pm 0.008)$ mJy after removing the spectral line. ERD-1 is also detected at $\lambda = 2$ mm in the Ex-MORA survey (Long et al., 2024), the extended version of the original MORA survey (Casey et al., 2021; Zavala et al., 2021), with $S_{2\rm mm} = (0.48 \pm 0.13)$ mJy. We also fit the (sub)mm and FIR photometry (including *Herschel*; Lutz et al. 2011; Oliver et al. 2012) with the `mmpz` code (Casey, 2020), determining that the probability of a redshift solution at $z < 5$ is lower than 9%, consistent with the spec-z measurement.

5.3.5 Radio data for ERD-1

A radio detection at 3GHz is available for ERD-1 from the VLA-COSMOS large program (Smolčić et al., 2017), with $S_{3\rm GHz} = (11 \pm 2)$ μJy. ERD-1 is not detected in the 1.4 GHz maps (Schinnerer et al., 2007), with $S_{1.4\rm GHz} = (25 \pm 19)$ μJy or a 2σ limit of < 63 μJy. This limit implies a radio slope $S \sim \nu^{\alpha}$ with $\alpha > -1.1$.

5.4 SED fitting

We derive photometric redshifts with `Eazy` (Brammer et al., 2008) for both galaxies, using the standard set of templates `tweak_fsps_QSF_12_v3` (Conroy & Gunn, 2010) coupled with those generated by Larson et al. (2023), and by allowing the redshift to vary in the range [0,10]. We independently estimate the photo-z and physical properties with `Bagpipes` (Carnall et al., 2018) using the stellar models from Bruzual & Charlot (2003), leaving the metallicity as a free parameter in the range $[0.1, 1] Z_{\odot}$, invoking a delayed exponentially declining star formation history (SFH) with a uniform prior on the e-folding time and on the total stellar mass formed in the ranges [0.3,10] Gyr and [10^5,10^{13}] $M_{\odot}$. Following Franco et al. (2023), we use a slightly different implementation of the SFH contained by default in `Bagpipes`, parametrizing the age of the main stellar population as a function of the Hubble time at a given redshift. We also add a recent instantaneous burst of star formation to the "secular" SFH taking place in the last t Myr, with a uniform prior on t and on the total stellar mass formed in the burst in the ranges [10,100] and [0,10^{13}] $M_{\odot}$. The synthetic SEDs by `Bagpipes` also include the emission lines modeled by `Cloudy` (Ferland et al., 2017) assuming a fixed ionization parameter of $\log(U) = -2$. To account for the possible dust emission that could be visible in the MIRI filter at high redshift, we also include the models by Draine & Li (2007) with uniform priors on the mass fraction of the Polycyclic Aromatic Hydrocarbons ($q_{\rm PAH} \in [1, 4]$) and on the lower limit of starlight intensity distribution ($U_{\rm min} \in [0.1, 25]$). The γ factor is fixed at the value $\gamma = 0.02$. Finally, SEDs are dust-extincted assuming a Calzetti

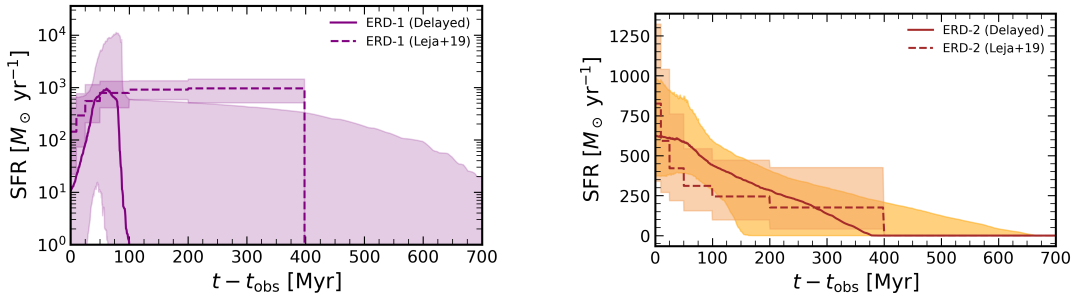

Figure 5.2: Star formation histories computed for our targets with `Bagpipes` by assuming a standard delayed exponentially declining SFH (solid lines) and the non-parametric one by Leja et al. (2019). In both panels, the shaded area reports the 16th-84th percentile of the posterior distribution.

et al. (2000) law with a uniform prior on $A_{\rm v}$ in the range [0,5] mag and a flat redshift prior from [0,10].

As a final test, we explore the possible existence of biases due to the chosen parametrization of the star formation history (SFH) of `Bagpipes` invoking a non-parametric SFH with the continuity prior as prescribed by Leja et al. (2019). We compute the $\Delta \log(SFR)$ in six age bins (0-10 Myr, 10-25 Myr, 25-50 Myr, 50-100 Myr, 100-200 Myr, and 200-400 Myr) adopting a prior shaped as a $t-$distribution with $\sigma = 0.3$ and $\nu = 2$ degrees of freedom. All the other priors are the same as for the parametric SFH. The output is summarized in Table 5.1, while the best-fitting SEDs are shown in Figure 5.1. The inferred SFHs can be found in Figure 5.2.

Table 5.1: Properties of our galaxies estimated through SED fitting. The table reports, for each objects, the two photometric redshifts computed by `Bagpipes` ($z_{\rm BG}$) and `Eazy` ($z_{\rm EZ}$). The stellar mass ($M_\star$) and dust attenuation ($A_{\rm v}$) are estimated with `Bagpipes` with both a parametric (P) and non-parametric (NP) SFH.

Name	RA (J2000)	DEC (J2000)	$z_{\rm spec}$ –	$z_{\rm BG}^{\rm P}$ –	$z_{\rm BG}^{\rm NP}$ –	$z_{\rm EZ}$ –	$\log(M_\star^{\rm P})$ $[M_\odot]$	$A_{\rm v}^{\rm P}$ [mag]	$\log(M_\star^{\rm NP})$ $[M_\odot]$	$A_{\rm v}^{\rm NP}$ [mag]
ERD-1	10:00:47.088	+02:10:16.680	5.051	$5.4^{+0.5}_{-0.4}$	$5.3^{+0.4}_{-0.4}$	$5.3^{+0.4}_{-0.4}$	$11.32^{+0.07}_{-0.15}$	$3.8^{+0.1}_{-0.2}$	$11.42^{+0.05}_{-0.07}$	$3.8^{+0.2}_{-0.2}$
ERD-2	10:01:07.368	+01:52:00.840	–	$6.7^{+0.3}_{-0.3}$	$6.6^{+0.2}_{-0.2}$	$7.3^{+0.3}_{-0.3}$	$11.2^{+0.1}_{-0.2}$	$2.9^{+0.3}_{-0.3}$	$11.0^{+0.1}_{-0.2}$	$3.0^{+0.2}_{-0.4}$

5.4.1 Estimated physical properties

Seen on Figure 5.1, the two galaxies have similar best-fit SEDs, characterized by a red continuum with a Balmer break and strong emission lines. Indeed, a red continuum with Balmer break explains the high stellar masses and dust extinction (because Balmer breaks imply older stellar populations that have a higher mass-to-light ratio). Moreover, the presence of strong emission lines suggests that these sources are actively forming stars. In addition, the presence of the Balmer break helps secure the photometric redshifts of our sources: the break between the Ks band and the F277W filter (for ERD-1) and between F277W and IRAC channel 1 (for ERD-2) produces the two photometric redshifts of $z \sim 5$ and $z \sim 6.5$ for the two galaxies, respectively.

As visible from Table 5.1, the photo-z estimated with `Eazy` and `Bagpipes` are in good agreement with each other (and, for ERD-1, just slightly higher than the spec-z measured by Jin et al. 2019, albeit compatible within the uncertainties). Since it is known, we fix the redshift of ERD-1 to the spectroscopic redshift. Moreover, the physical properties estimated with the non-parametric SFH are in good agreement with those computed by assuming the more standard parametric SFH. In the following, we will assume the latter as reference values for both galaxies.

For ERD-1, we find a stellar mass of $\log(M_\star) = (11.32^{+0.07}_{-0.15})$ $M_\odot$ and a dust attenuation of $A_\mathrm{v} = (3.8^{+0.1}_{-0.2})$ mag. ERD-2 has $\log(M_\star) = (11.2^{+0.1}_{-0.2})$ $M_\odot$ and $A_\mathrm{v} = (2.9^{+0.3}_{-0.3})$ mag. The mass for ERD-1 is more accurate than that of ERD-2 thanks to the spec-z and the MIRI/F770W detection (tracing the rest-frame NIR wavelengths and – therefore – the bulk of the stellar mass, less affected by the presence of dust; see e.g. Papovich et al. 2023; Wang et al. 2024a). The availability of a MIRI constraint for ERD-2 could, in principle, decrease the inferred stellar mass. Akins et al. (2024) measures offset as high as ~ 1 dex for LRDs with and without the additional MIRI point. However, this effect is primarily due to an overestimated age for the main stellar population (Akins et al., 2024). Since we find a young stellar population for ERD-2 ($\lesssim 400$ Myr; see Figure 5.2), we do not expect this effect to dramatically change our results. To better quantify this possibility, we run an additional test by artificially adding a MIRI/F770W point to the ERD-2 photometry and assuming a ratio [F770W]/[F444W] in the range [1,5], broadly consistent with measured LRD colors at similar redshifts (for instance, the same value for ERD-1 is $\sim$4). By fitting again the new photometry, we obtain a decrease of $\sim$0.4 dex in the inferred stellar mass with a completely flat SED. The offset rapidly decreases with higher F770W fluxes and becomes negligible with [F770W]/[F444W]$\sim$3. This result suggests that the 0.4 dex decrease should be interpreted as the maximum effect of the additional MIRI constraint on the stellar mass of ERD-2. Nevertheless, a MIRI follow-up of this galaxy would be crucial to properly estimate this value. Finally, following Daddi et al. (2010), we estimate the dynamical mass of ERD-1 from the FWHM of the CO line (~ 850 km s^{-1}; Jin et al. 2019) and the physical size computed in Section 5.5). The obtained value of $\log(M_\mathrm{Dyn}) = (11.3 \pm 0.2)$ $M_\odot$ is in good agreement with the inferred stellar mass.

Since neither galaxy is detected in F115W, we have a weak limit on the rest-frame UV flux. Hence, the values reported by the SED-fitting codes ($\log(SFR) = 3.4^{+0.1}_{-0.6}$ $M_\odot$ yr^{-1} and $\log(SFR) = 2.8^{+0.2}_{-0.2}$ $M_\odot$ yr^{-1} for ERD-1 and ERD-2, respectively)

are largely unconstrained. However, since ERD-1 has multiple (sub)mm and radio detections – qualifying it for classic DSFGs selection (see e.g Casey et al. 2014) – and a spectroscopic redshift reducing the possible degeneracies, we estimate the obscured SFR. More specifically, by assuming a standard radio slope of $\alpha = -0.7$ (star-forming galaxies and in good agreement with the upper limit given in Section 5.3.5), we estimate a radio luminosity of $L_{1.4\mathrm{GHz}} = (3.1 \pm 0.6) \times 10^{24}$ W Hz^{-1}. Assuming that all the radio emission is produced by star formation, we use Kennicutt & Evans (2012) to estimate a $\mathrm{SFR}_{1.4\mathrm{GHz}} = (1.6 \pm 0.3) \times 10^3$ M$_\odot$ yr^{-1}. This estimate is similar to that computed via the infrared SED in Jin et al. (2019): $L_{\mathrm{IR}} = (6 \pm 1) \times 10^{12}$ $L_\odot$, implying a $\mathrm{SFR}_{\mathrm{IR}} = (1.0 \pm 0.3) \times 10^3$ M$_\odot$ yr^{-1}. We measure the FIR-to-radio luminosity ratio as $q_{\mathrm{IR}} = 2.3 \pm 0.2$ (see e.g. Helou et al., 1985), in good agreement with q_{IR} of star-forming galaxies (e.g Yun et al., 2001). The high SFR also explains the presence of the strong nebular emission lines visible in the best-fit SEDs. All the above-mentioned properties for ERD-1 suggest – together with the ALMA detections – that this galaxy belongs to the high-mass end of the dusty star-forming galaxy population (see e.g. Casey et al. 2014; Wang et al. 2019), something that is not always true of galaxies with red JWST colors (see e.g. Barrufet et al. 2024).

5.5 Morphological Analysis

We use `Imfit` (Erwin, 2015) to perform a two-dimensional profile fitting of our galaxies in F444W, the highest S/N filter, using a single Sérsic model, leaving the centroid, axis ratio, position angle, Sérsic index, effective radius, and flux as free parameters. Our resulting models describe the data well, with reduced $\chi^2 \sim 1$ (see Figure 5.3). From the profile fitting, we obtain that ERD-1 and ERD-2 have effective radii of (1.87 ± 0.02) kpc and (1.43 ± 0.02) kpc respectively, and that both are well described by a Sérsic profile with $n \sim 1$.

We also fit the data to a combined model with both an unresolved point source and Sérsic profile. To limit the additional free parameters, we fix the unresolved centroid to the center of the Sérsic model, and leave the flux as a free parameter. The best-fitting model always converges towards a zero-flux point source. If we fix the contribution from the point source to be at least $\sim 20\%$ of the total integrated flux (see e.g. the right panel in Figure 5.3), we obtain a significantly worse fit, clearly indicating no contribution from an unresolved component.

5.6 Discussion

5.6.1 Estimating the stellar baryon fraction

ΛCDM cosmology uniquely determines a certain mass function of dark matter halos at a given redshift (e.g. Sheth & Tormen, 1999). However, converting this halo mass to a stellar mass is a more challenging task, since baryonic processes are complex and not completely understood, especially at high-z. Under the oversimplified hypothesis that all the baryons in halos are converted in stars (an assumption that is most certainly incorrect), we can derive a stringent upper limit on the galaxies' possible masses at each redshift in a given cosmological volume (see e.g. Steinhardt et al.

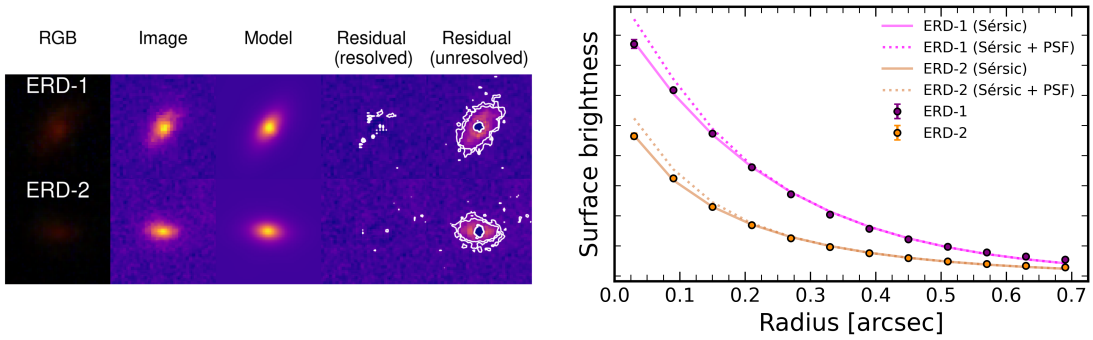

Figure 5.3: *Left panel:* RGB cutouts and F444W profile-fitting of ERD-1 and ERD-2 in 2" × 2" cutouts through `Imfit`. The last two panels show the residuals with the Sérsic model and an unresolved point source. We overplot the contours at 3 and 5 σ. A single Sérsic profile correctly reproduces the brightness profile of our objects without the need for an additional unresolved component. *Right panel:* Surface brightness profile (in F444W) modeled with a single Sérsic profile and with the addition of an unresolved component. Again, the Sérsic profile well reproduces the observations.

2016; Behroozi & Silk 2018; Boylan-Kolchin 2023) and compare our estimated stellar masses with these limits.

Here we follow the analysis of Boylan-Kolchin (2023)[1]. We first consider the halo mass function by Sheth & Tormen (1999) and multiply it by the cosmic baryon fraction $f_{\rm b} \sim 0.16$ (Planck Collaboration et al., 2020) to obtain a relationship between the maximal stellar mass and volume density: $M_\star = M_{\rm H} f_{\rm b} \epsilon_\star$. This value depends on $\epsilon_\star$, the stellar baryon fraction, describing the integrated history of the star formation efficiency in a given galaxy.

Considering the redshift range spanned by our galaxies ($5 < z < 7$) and the sky area analyzed here from COSMOS-Web (0.28 deg^2), we estimate that our survey covers a total comoving volume of 7.6×10^{11} Mpc^3. In Figure 5.4, we compare our sources' stellar masses with the maximum mass allowed with redshift in our survey volume. To fit within ΛCDM, the two galaxies presented in this chapter would require a (minimum) $\epsilon_\star = 0.24^{+0.02}_{-0.08}$ and $\epsilon_\star = 0.5^{+0.2}_{-0.2}$ (for ERD-1 and ERD-2, respectively) under the hypothesis that they are the only galaxies in the cosmic volume explored so far by COMSOS-Web with these masses at these redshifts. Clearly, the possible identification of additional sources with analogous masses (e.g. the rest of the sample presented in Section 5.3.2) would increase the inferred $\epsilon_\star$ values. This point will be addressed in more detail in a forthcoming paper (Gentile et al., *in prep.*)

[1] We use the github code referenced therein, provided at this link: https://github.com/mrbk/JWST_MstarDensity/tree/main

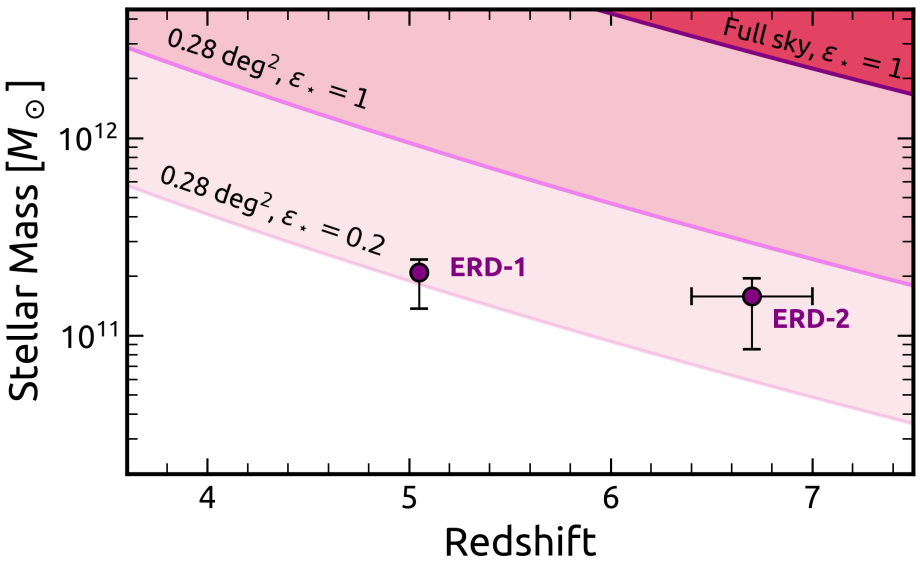

Figure 5.4: ERD-1 and ERD-2 in the stellar mass vs. redshift plane. The lines correspond to different number densities: 1.3×10^{-12} Mpc^{-3} (i.e. a single galaxy in the whole sky; dark violet) and 1.9×10^{-7} Mpc^{-3} (i.e. one source in the cosmic volume explored so-far by COSMOS-Web; dark red) with a (unrealistic) stellar baryon fraction of $\epsilon_\star = 1$. The lower (light pink) line scales down the second using a more reasonable efficiency of $\epsilon_\star = 0.2$.

5.6.2 Implication for star formation efficiency

Star formation is known to be very inefficient, with several self-regulating processes limiting the efficiency at small scales (see e.g. Kennicutt 1998; Ostriker & Shetty 2011). Observational studies at lower redshifts ($z \lesssim 2-3$) have an implied strong upper limit on $\epsilon_\star \ll 0.2$ (e.g. Mandelbaum et al., 2006; Behroozi et al., 2010; Kennicutt & Evans, 2012; Moster et al., 2013; Wechsler & Tinker, 2018; Shuntov et al., 2022), with some (e.g. Behroozi et al. 2010 and Moster et al. 2013) notably expecting a decreased integrated star formation efficiency towards higher redshifts and others finding a more constant behavior until at least $z \sim 4$ (e.g. Behroozi et al. 2013). Our values of $\epsilon_\star$ are significantly higher than these studies expect, suggesting star formation pathways much more efficient at high-z than what is observed in the local Universe. This conclusion is strengthened by similar $\epsilon_\star$ reported by Xiao et al. (2023) for their red galaxy sample (with spectroscopic redshifts) from the FRESCO survey (Oesch et al., 2023), suggesting $\epsilon_\star \sim 0.5$, or with those reported by Wang et al. (2024b) for a sample of MIRI-detected massive galaxies, with a moderately lower $\epsilon_\star \sim 0.3$.

We underline, however, that the estimated $\epsilon_\star$ are also the product of some assumptions that – if proved wrong – could strongly affect the measured values. Firstly, the halo mass function is a consequence of the assumed ΛCDM cosmology. Some alternative models (e.g. early dark energy; see e.g. Karwal & Kamionkowski 2016, Poulin et al. 2018) forecast the early formation of more massive halos (e.g. Klypin et al., 2021). In this case, halos are more massive, and a larger baryonic mass would be available for star formation.

Secondly, the stellar masses obtained in this chapter assume a Chabrier (2003) IMF. Several works argue that this could be an oversimplification, due to the dependence of the IMF on the metallicity or the temperature of molecular clouds (see e.g. Low & Lynden-Bell 1976; Davé 2008). A top-heavy IMF could produce lower stellar masses (with a factor $\gtrsim 0.3$ dex; see e.g. Woodrum et al. 2023; Wang et al. 2024a) and give slightly lower values of $\epsilon_\star$.

Moreover, the possible contribution of reddened AGN emission to the continuum can significantly alter the stellar mass estimation (even with a spectroscopically confirmed Balmer break; see e.g. Wang et al. 2024b). This hypothesis is ruled out for our galaxies thanks to their extended morphology, ensuring that the main contributor to the red continuum is the stellar emission.

Finally, even excluding the contribution of AGN, the presence of strong emission lines (due to intense star formation) can produce an overestimation of the stellar masses of galaxies (see e.g. Schaerer & de Barros 2009). Unfortunately, the contribution of these lines is hard to constrain with photometry alone. Therefore, a spectroscopic follow-up with JWST would be necessary to further confirm the inferred physical properties of our targets.

While some simulations suggest that very efficient star formation may be possible in dense giant molecular clouds (e.g. Torrey et al. 2017; Grudić et al. 2018; Dekel et al. 2023), the extended morphology in these cases seems somewhat disjoint from such an explanation. Further observations of the mass budget in such systems e.g. from spatially resolved observations of [CII] dynamics from ALMA, may be crucial to deciphering the underlying puzzle.

5.7 Chapter summary

In this chapter, I presented two candidate massive, extended and highly dust-obscured galaxies: ERD-1 and ERD-2. These sources have extremely red ($F277W - F444W \sim 1.7$ mag) color, mimicking the selection of "little red dots" (e.g. Akins et al., 2023; Barro et al., 2023). However, these galaxies have an extended morphology, ruling out the possibility that the continuum is dominated by reddened nuclear activity (unlike the majority of LRDs) thus ensuring that the stellar emission is responsible for the observed red colors.

Together with my collaborators, I have employed SED fitting using two different codes (`Eazy` and `Bagpipes`) yielding photo-z estimates of ERD-2 of $z \sim 6.7$, while a spec-z of 5.051 of ERD-1 (Jin et al., 2019) is already published. Both the best-fitting SEDs show a red continuum with a Balmer break combined with emission lines. These features suggest a moderately aged stellar population with high stellar masses ($M_\star \sim 10^{11}$ $M_\odot$) and dust extinction ($A_\mathrm{v} > 2.5$ mag), together with an ongoing episode of significant star formation.

The high stellar masses, once placed in the cosmological context through comparison with the halo mass function, suggest that our galaxies could have a much higher stellar baryon fraction ($\epsilon_\star \gtrsim 0.2$) than what is commonly observed in lower-redshift sources.

Possible explanations of this result are a top-heavier IMF in the high-z Universe, a much more efficient star formation in the early stages of the cosmic history than

has been constrained by observations at low-z, or an earlier formation of massive halos than what predicted by ΛCDM.

These possible consequences increase the need for spectroscopic follow-up of these interesting systems, to confirm the photo-z (for ERD-2), confirm the presence of the Balmer breaks suggested by the photometry (e.g. Wang et al. 2024b), and to accurately measure the dynamical masses of our galaxies, which is fully possible using JWST directly or ALMA.

Chapter 6

Conclusions

The main goal of this Thesis was to study the behavior of the cosmic SFRD in the $z > 3$ Universe and identify the progenitors of the massive and passive galaxies observed at $z \sim 3$ and in the local Universe. The works presented in this Thesis followed two complementary approaches. On the one hand — in Chapters 2, 3, and 4 — I studied the population of radio-selected NIRdark/faint galaxies, presenting their physical properties, likely evolutionary path, and total contribution to the cosmic SFRD. On the other hand, in Chapter 5, I took advantage of the new JWST data of the COSMOS field as part of the COSMOS-Web survey to present the discovery of two extremely red galaxies that — according to my analysis — could be as massive as $\sim 10^{11}$ $M_{\odot}$ at $z \sim 5-7$, representing a challenge for the current picture of star formation in the early Universe. In this last chapter of the Thesis, I summarize the main findings of my PhD and the future perspective of this field of research.

6.1 Main results of this thesis

One of the common threads connecting the several works presented in this Thesis is the more complete picture of the Universe that we get every time a new piece of information is added to our analysis. Starting from the Introduction, I showed how the picture of the high-z SFRD based solely on dust-biased tracers could be partial and miss significant pieces of information. For this reason, in Chapter 2, I presented a complementary dust-unbiased selection based at radio frequencies. The following chapters showed how the employment of state-of-the-art facilities such as ALMA and JWST helped me to better characterize my galaxies. More in detail, the ALMA data described in Chapter 3 allowed me to measure the spectroscopic redshift of a pilot sample of RS-NIRdark galaxies, estimate their molecular gas content, and forecast their likely evolutionary path. Similarly, the deeper and high-resolution NIR data obtained with JWST (and described in Chapter 4) allowed me to increase the purity of the selection, obtain better constraints on the photo-zs and physical properties of my galaxies and properly constrain their radio luminosity function and their total contribution to the cosmic SFRD. The same JWST data were crucial for the analysis presented in Chapter 5 to unveil the presence of two extremely red galaxies, missed by previous optical/NIR facilities, and to constrain their physical properties.

This plethora of data available in the COSMOS field allowed me to obtain several

scientific results that represent the main product of this Thesis. I summarize them here.

- First of all, I introduced the radio-selected NIRdark galaxies: a population of sources defined as galaxies detected at radio frequencies but lacking a counterpart in deep surveys conducted with old-generation facilities such as the VISTA telescope. The advent of the James Webb Space Telescope allowed me to update this definition, changing the requirement of a missing detection at NIR wavelengths with that of a faint flux in the same regime. The initial hypothesis on these sources — formulated by Talia et al. (2021) (see also the first results by Chapman et al. 2001, 2002, 2004) — is that this selection could be able to to produce a sample of likely DSFGs. This hypothesis was justified by the fact that radio emission can trace star formation (once properly accounted for the possible contribution of nuclear activity to the radio luminosity) and that the faint NIR counterpart could suggest a significant amount of dust in the analysed sources.

- The results presented in Chapter 2, 3, and 4 confirm this initial hypothesis, picturing the RS-NIRdark/faint galaxies as a population of highly dust-obscured ($A_{\rm v}$ ∼3-4 mag), massive ($M_\star \sim 10^{10.5-11}$ $M_\odot$), and star-forming (SFR∼300-500 $M_\odot$ yr^{-1}) galaxies mainly located at $z > 2.5 - 3$. A few minor differences hold between the properties estimated in Chapter 2 and Chapter 4, but they are easily explained by the updated (and more pure) selection performed in Chapter 4 and with the deeper NIR data offered by JWST reducing the uncertainties due to the missing constraints at these wavelengths affecting the initial analysis described in Chapter 2.

- These properties are — on average — confirmed by the improved analysis described in Chapter 3 based on the ALMA data collected for a pilot sample of nine RS-NIRdark galaxies. The availability of a spectroscopic redshift estimated through (sub)mm emission lines confirmed the photo-z estimated in Chapter 2. Moreover, the possibility to reduce the degeneracies in the SED-fitting procedure by fixing the redshift value, also allowed me to better constrain the physical properties of our targets, confirming the overall picture of the RS-NIRdark galaxies as a population of DSFGs at $z > 3$.

- Concerning the likely evolutionary path of RS-NIRdark galaxies, the results achieved in this thesis confirm the hypothesis presented, for instance, by Toft et al. (2014) and Valentino et al. (2020a), that a significant percentage of the high-z progenitors of the massive and passive galaxies discovered at $z \sim 3$ by Straatman et al. (2014) and Schreiber et al. (2018a) could be significantly dust obscured. The number densities (uncorrected for the completeness of the radio survey and not accounting for the likely *duty cycle* of these sources) estimated in Chapter 2 for the high-z ($z > 3.5$) tail of the RS-NIRdark galaxies is in the order of a few 10^{-6} Mpc^{-3}, about a factor 6x lower than what is measured by Straatman et al. (2014), Schreiber et al. (2018a), and Valentino et al. (2020a). Moreover, the comparison between the mass distribution of my sources with that reported by these other studies show that the RS-NIRdark galaxies could account for the formation of the high-mass tail of their distribution (that —

commonly — is the most challenging to produce with lower-mass NIR-bright sources; see e.g. the discussion in Toft et al. 2014).

- This evolutionary scenario is strengthen by the analysis described in Chapter 3 involving the ALMA data. The molecular gas mass estimated starting from the CO/[CI] emission lines detected in the (sub)mm spectra of the nine pilot targets can be combined with the SFR obtained through the FIR luminosity to estimate their depletion time. Being these sources in the starburst regime (see Chapters 2 and Chapter 3), even in the presence of significant gas reservoirs, they are able to deplete all the molecular gas in short (< 300 Myr) amounts of time, producing massive ($M_\star \sim 10^{11}$ $M_\odot$) galaxies before $z \sim 3$.

- Another interesting clue that the impact of NIR-dark (or NIR-faint, after the advent of JWST) galaxies on our picture of the high-z Universe could be significant is represented by the results presented in Chapter 4, taking advantage of the new JWST coverage of the COSMOS field. Here, the main result is that the radio-selected NIRfaint galaxies could dominate the bright-end of the radio luminosity function. This result could suggest a positive correlation between the star formation of galaxies and their dust obscuration, strengthening what has been suggested by several observational and theoretical works (e.g. Whitaker et al. 2017; Khusanova et al. 2021; Traina et al., *in prep.*).

- The main consequence of the previous point is the estimated contribution of RS-NIRdark/faint galaxies to the cosmic SFRD. Only accounting for the sources observed in our survey, this contribution could be as high as $(8 \pm 2) \times 10^{-3}$ $M_\odot$ yr $^{-1}$ Mpc^{-3} at $z \sim 3.5$, and up to a factor 3x higher if fainter and brighter sources are included by integrating the radio luminosity function outside the observed range of observed luminosity.

- Another result concerns the relationship between the RS-NIRdark/faint galaxies and the other populations of DSFGs presented in the Introduction (Section 1.3.1). This comparison has been described several times across the Thesis, taking advantage of the new data described in each chapter. The overall result is that the RS-NIRdark/faint galaxies are a sub-population of the MIR-selected DSFGs, even though some of their MIR fluxes can be too faint for being detected by old-generation facilities such as *Spitzer*, requiring the higher sensitivity of JWST for being completely detected (see Sections 2.7.3 and 4.6.1). Nevertheless, the RS-NIRdark/faint galaxies represent the high-mass and high-SFR tail of the MIR-selected galaxies, with a selection more robust to the contamination by passive galaxies reported by Wang et al. (2019); Barrufet et al. (2024) and Pérez-González et al. (2024).

- Similarly, the RS-NIRdark/faint galaxies significantly overlap with the classical (sub)mm selection. The difference here is clearly determined by the depth of the (sub)mm observations. As shown in Chapters 2, 3, and 4, the (sub)mm flux of my sources are — on average — fainter than those normally observed by old-generation facilities such as SCUBA-2 ($S_{870\mu m} \sim 4$ mJy). Nevertheless, their fluxes significantly overlap with those sampled by the ALMA follow-ups of sources observed with old-generation facilities. This result well explains why

— on average — their physical properties (namely, stellar mass and SFR) are lower than what observed for the classical SMGs.

Overall, the achieved results on the RS-NIRdark/faint galaxies in the COSMOS field strongly suggest that the picture of the high-z Universe only based on HST and other old-generation optical/NIR facilities is far from being complete. The same suggestion comes from the results presented in Chapter 5 and summarized here:

- The discovery of two extremely massive ($M_\star \sim 10^{11}$ M$_\odot$) already in place at $z \sim 5-7$ is in a slight tension with the halo mass function derived from the ΛCDM cosmology. Without invoking a dramatic change of paradigm in our cosmological model, this result suggests that the overall star formation efficiency in the early Universe could be much higher than what suggested by more local observations.
- Moreover, the existence of galaxies presenting extremely red colors without being dominated by reddened AGN emission contributes to the scientific debate about the true nature of the so-called Little Red Dots. More in detail, the possibility that a part of this ubiquitous population of red sources unveiled by the new JWST data is actually composed by extremely massive galaxies strengthen the idea that the star formation could have been much more efficient in the early Universe than in the more local one.

6.2 Future perspectives and follow-up opportunities

Once assessed the main scientific results achieved in this Thesis, I focus this last section on the discussion of the future perspective of this work.

- **Follow-up opportunities for the RS-NIRdark/faint galaxies:** The overall Thesis presented the case for collecting new data on the RS-NIRdark galaxies. Due to the elusive nature of these sources, the follow-up opportunities clearly involve (sub)mm facilities like ALMA and NOEMA, plus JWST.
  - **ALMA/NOEMA:** For the first family of facilities, the main idea is to extend the pilot study described in Chapter 3 to the full sample of RS-NIRdark/faint galaxies in COSMOS (or — to save observational time — to the high-z tail of the overall distribution). The main goal would be confirming the photo-z and — consequently — the number density at $z > 3.5$ to confirm the evolutionary scenario presenting my sources as a significant fraction of the progenitors of the massive and passive galaxies at $z \sim 2.5-3$. The spectral setup presented in Chapter 3 is ideal for this task, since the detection of multiple CO/[CI] transitions allows an unambiguous determination of the spec-z and the simultaneous estimate of the molecular gas mass inside the targets. As shown in Section 3.4.4, this quantity has a pivotal role in the forecast of the possible evolutionary scenario. An alternative setup would include the observation of the [CII] line, at the same time a good tracer of star formation (e.g. De Looze et al.

2014; Schaerer et al. 2020) and molecular gas (e.g. Zanella et al. 2018). This approach is the same as employed by the REBELS collaboration (Bouwens et al., 2022) in their ALMA follow-up of UV-selected galaxies: it has the main advantage of being less time consuming than the observations of the full ALMA band 3 and of targeting a brighter line than the CO/[CI] ones. This approach is however less robust to catastrophic failures of the photo-z since it relies on a single line detection that could be wrongly modeled as an [OIII] line or an hihgh-J CO transition.

Tightly related to these observations, another interesting follow-up opportunity is represented by high-resolution observations with ALMA of bright lines (e.g. [CII]) in spectroscopically-confirmed targets. These observations are clearly more time-consuming, but can allow a kinematic modeling of the observed galaxies. As described in Chapter 3, the potential discovery of kinematically-decoupled components in my galaxies, or the presence of a kinematically-cold disk, could be employed to constrain galaxy evolution scenarios such as the *in situ* or the *merger driven* ones (see e.g. Lapi et al. 2018; Pantoni et al. 2019; Hopkins et al. 2008b,a).

– **JWST:** Regarding follow-ups with the James Webb Space Telescope, here the main advancement would consist in spectroscopic observations at NIR wavelengths (with the NIRSpec instrument). The possibility to detect in my galaxies the rest-optical emission lines such as $H\alpha$, $H\beta$, or [OIII] would translate in the ability of measuring the spec-z and constrain some of the physical properties now only estimated through SED fitting. More in detail, the simultaneous observations of $H\alpha$ and $H\beta$ would allow me to estimate the dust attenuation $A_{\rm v}$ through the Balmer decrement and — therefore — reduce the uncertainties on the stellar mass of my galaxies. Similarly, the detection of the [OIII] line would give me the chance of estimating the metallicity of my galaxies and — consequently — match their chemical evolution with that observed in the local massive and passive galaxies. Moreover, the possible detection of broad lines and the possibility of producing a BPT diagram will give me the opportunity of unveiling the presence of AGNs not discovered through the procedures employed in the previous chapters.

Apart from rest-optical spectroscopy, the availability of imaging data at MIR wavelengths (e.g. through the MIRI instrument equipped on JWST) would reduce the uncertainties on the stellar masses in my galaxies (since at $\lambda > 4$ μm this instrument would sample the rest-NIR emission in my galaxies) and potentially unveil the presence of a hot dusty torus in the center of my galaxies allowing, again, the unveiling of AGNs in my sources.

- **Comparison with simulations:** A natural continuation of the work presented in this Thesis consists in the comparison of my predictions with hydrodynamical simulations and semi-analytic models. It is well known, indeed, that these simulations constantly under-predict the number density of massive dusty galaxies in the $z > 3$ Universe and — consequently — their contribution to the cosmic SFRD (see e.g. Henriques et al. 2015; Snyder et al. 2017;

Pillepich et al. 2018 or the discussion on this point in Wang et al. 2019). The number densities reported in this Thesis can be used to fine-tune these models, offering a new suit of observables that can be employed to refine the sub-grid models ruling the formation of dust in the simulated galaxies.

A second interesting point concerns the validation through simulations of some of the trends presented in this Thesis. More in detail, the possible positive correlation between SFR and dust obscuration unveiled with my sources can be easily studied with state-of-the-art zoom-in simulations able to give in output the dust distribution inside the simulated galaxies (see e.g. Davé et al. 2019; Pallottini et al. 2022; Hopkins et al. 2023).

Finally, simulations can be employed to study the past and future evolutionary path of my sources, giving us a perspective on the parts of their history that are not sampled by our observations. More in detail, once found in the simulated boxes some analogs for my sources, one can follow their past and future evolution to study the properties of the expected high-z progenitors and to confirm the evolutionary path presented in Chapter 3 that could bring to the formation of the massive and passive galaxies at $z \sim 3$. Finally, as shown in the previous chapters, my galaxies are expected to be found in the starburst regime of the main sequence. Therefore, one would expect a short *duty cycle* (i.e. a short percentage of their lifetime where they satisfy my selection criteria). The possibility of reconstructing their full evolutionary path would allow me to quantify the duty cycle and account for it in the determination of the statistical properties of the sample.

- **Theoretical works and the broader picture on the cosmic SFRD:** A crucial point that has been only partially solved by this Thesis is the overlap between the different selections of DSFGs based at different wavelengths. As a general result, in the previous chapters I stressed how the RS-NIRdark/faint galaxies have analogous properties of NIRdark/faint SMGs with (sub)mm fluxes — on average — lower than those normally sampled by old-generation (sub)mm facilities. Similarly, I showed how the radio selection collects a subsample of the MIR-selected DSFGs biased towards higher SFR values. Nevertheless, the relationship between physical properties (mass, dust attenuation, dust temperature, SFR, and redshift) and the selection with one method or another is still unclear. The availability of a single selection function would allow us to combine the different contribution to the SFRD obtained on studies following a single selection and unveil the total contribution of DSFGs to the cosmic SFRD. This is only possible through a theoretical study analyzing synthetic SEDs and their variability changing the physical properties of galaxies. This point is currently under investigation and will be the first future outcome of this Thesis (Gentile et al., *in prep.*, a).

- **Follow-up opportunities for "Extended red dots":** The final follow-up opportunity offered by this Thesis focuses on the extremely red sources discovered by JWST and presented in Chapter 5. Before discussing them, I recall that — in that chapter — I only presented two extremely massive sources out of a sample of ~ 100 galaxies found in the full footprint of COSMOS-Web

that will be analyzed in a forthcoming study (Gentile et al., *in prep.*, b). Since the scientific results presented in Chapter 5 concerns the comparison of the redshift and stellar masses with the halo mass function at those redshifts, the main follow-up should be focused on the measure of the spectroscopic redshifts of the sources in the sample and in the improvement of the stellar mass estimates. The first one, as already discussed for the RS-NIRdark/faint galaxies can be obtained with (sub)mm facilities such as ALMA or NOEMA through the detection of emission lines (mainly CO, [CI], or [CII]) or with rest-optical spectroscopy with JWST. The improvement on the stellar mass estimates can be achieved with imaging at longer wavelengths (i.e. with MIRI), or with the detection of the stellar continuum giving insights on the stellar populations inside the galaxies. Finally, the availability of spectroscopic data with high spectral resolution would allow the determination of the dynamical mass of the targets, automatically posing a stringent upper limit on the stellar mass of the galaxies in the sample.

Bibliography

Adscheid, S., Magnelli, B., Liu, D., et al. 2024, A&A, 685, A1, doi: 10.1051/0004-6361/202348407

Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, PASJ, 71, 114, doi: 10.1093/pasj/psz103

Akins, H. B., Casey, C. M., Allen, N., et al. 2023, ApJ, 956, 61, doi: 10.3847/1538-4357/acef21

Akins, H. B., Casey, C. M., Lambrides, E., et al. 2024, arXiv e-prints, arXiv:2406.10341, doi: 10.48550/arXiv.2406.10341

Algera, H. S. B., Inami, H., Oesch, P. A., et al. 2023, MNRAS, 518, 6142, doi: 10.1093/mnras/stac3195

Aniano, G., Draine, B. T., Gordon, K. D., & Sandstrom, K. 2011, PASP, 123, 1218, doi: 10.1086/662219

Aravena, M., Spilker, J. S., Bethermin, M., et al. 2016, MNRAS, 457, 4406, doi: 10.1093/mnras/stw275

Aretxaga, I., Wilson, G. W., Aguilar, E., et al. 2011, MNRAS, 415, 3831, doi: 10.1111/j.1365-2966.2011.18989.x

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540, doi: 10.1046/j.1365-8711.1999.02978.x

Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2013, ApJ, 769, 80, doi: 10.1088/0004-637X/769/1/80

Barbary, K. 2016a, Extinction v0.3.0, Zenodo, doi: 10.5281/zenodo.804967

—. 2016b, The Journal of Open Source Software, 1, 58, doi: 10.21105/joss.00058

Barger, A. J., Cowie, L. L., Sanders, D. B., et al. 1998, Nature, 394, 248, doi: 10.1038/28338

Barkana, R., & Loeb, A. 2000, ApJ, 539, 20, doi: 10.1086/309229

Barro, G., Perez-Gonzalez, P. G., Kocevski, D. D., et al. 2023, arXiv e-prints, arXiv:2305.14418, doi: 10.48550/arXiv.2305.14418

Barrufet, L., Oesch, P. A., Weibel, A., et al. 2023, MNRAS, 522, 449, doi: 10.1093/mnras/stad947

Barrufet, L., Oesch, P. A., Bouwens, R., et al. 2023, MNRAS, 522, 3926, doi: 10.1093/mnras/stad1259

Barrufet, L., Oesch, P., Marques-Chaves, R., et al. 2024, arXiv e-prints, arXiv:2404.08052, doi: 10.48550/arXiv.2404.08052

Battisti, A. J., da Cunha, E., Grasha, K., et al. 2019, ApJ, 882, 61, doi: 10.3847/1538-4357/ab345d

Behiri, M., Talia, M., Cimatti, A., et al. 2023, ApJ, 957, 63, doi: 10.3847/1538-4357/acf616

Behroozi, P., & Silk, J. 2018, MNRAS, 477, 5382, doi: 10.1093/mnras/sty945

Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379, doi: 10.1088/0004-637X/717/1/379

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 762, L31, doi: 10.1088/2041-8205/762/2/L31

Bertin, E. 2011, in Astronomical Society of the Pacific Conference Series, Vol. 442, Astronomical Data Analysis Software and Systems XX, ed. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, 435

Bertin, E., & Arnouts, S. 1996, åps, 117, 393, doi: 10.1051/aas:1996164

Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 228

Bertin, E., Schefer, M., Apostolakos, N., et al. 2020, in Astronomical Society of the Pacific Conference Series, Vol. 527, Astronomical Data Analysis Software and Systems XXIX, ed. R. Pizzo, E. R. Deul, J. D. Mol, J. de Plaa, & H. Verkouter, 461

Bertoldi, F., Carilli, C., Aravena, M., et al. 2007, ApJS, 172, 132, doi: 10.1086/520511

Béthermin, M., Fudamoto, Y., Ginolfi, M., et al. 2020, A&A, 643, A2, doi: 10.1051/0004-6361/202037649

Bevington, P. R., & Robinson, D. K. 2003, Data Reduction and Error Analysis for the Physical Sciences; 3rd Ed. (New York, NY: McGraw-Hill). https://cds.cern.ch/record/1305448

Bezanson, R., Labbe, I., Whitaker, K. E., et al. 2022, arXiv e-prints, arXiv:2212.04026, doi: 10.48550/arXiv.2212.04026

Binggeli, B., Sandage, A., & Tammann, G. A. 1988, ARA&A, 26, 509, doi: 10.1146/annurev.aa.26.090188.002453

Birkin, J. E., Weiss, A., Wardlow, J. L., et al. 2021, MNRAS, 501, 3926, doi: 10.1093/mnras/staa3862

Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207, doi: 10.1146/annurev-astro-082812-140944

Bonzini, M., Padovani, P., Mainieri, V., et al. 2013, MNRAS, 436, 3759, doi: 10.1093/mnras/stt1879

Boogaard, L., Meyer, R. A., & Novak, M. 2021, Interferopy: Analysing Datacubes from Radio-to-Submm Observations, doi: 10.5281/ZENODO.5775603

Boogaard, L. A., van der Werf, P., Weiss, A., et al. 2020, ApJ, 902, 109, doi: 10.3847/1538-4357/abb82f

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, åp, 622, A103, doi: 10.1051/0004-6361/201834156

Bothwell, M. S., Smail, I., Chapman, S. C., et al. 2013, MNRAS, 429, 3047, doi: 10.1093/mnras/sts562

Bothwell, M. S., Aguirre, J. E., Aravena, M., et al. 2017, MNRAS, 466, 2825, doi: 10.1093/mnras/stw3270

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012a, ApJ, 752, L5, doi: 10.1088/2041-8205/752/1/L5

—. 2012b, ApJ, 754, 83, doi: 10.1088/0004-637X/754/2/83

Bouwens, R. J., Smit, R., Schouws, S., et al. 2022, ApJ, 931, 160, doi: 10.3847/1538-4357/ac5a4a

Bouwens, R. J., Stefanon, M., Brammer, G., et al. 2023, MNRAS, 523, 1036, doi: 10.1093/mnras/stad1145

Boylan-Kolchin, M. 2023, Nature Astronomy, 7, 731, doi: 10.1038/s41550-023-01937-7

Bradley, L., Sipőcz, B., Robitaille, T., et al. 2020, Astropy/Photutils: 1.0.0, Zenodo, doi: 10.5281/zenodo.4044744

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503, doi: 10.1086/591786

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000, doi: 10.1046/j.1365-8711.2003.06897.x

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682, doi: 10.1086/308692

Cañameras, R., Yang, C., Nesvadba, N. P. H., et al. 2018, åp, 620, A61, doi: 10.1051/0004-6361/201833625

Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379, doi: 10.1093/mnras/sty2169

Carniani, S., Hainline, K., D'Eugenio, F., et al. 2024, Nature, 633, 318, doi: 10.1038/s41586-024-07860-9

CASA Team, Bean, B., Bhatnagar, S., et al. 2022, PASP, 134, 114501, doi: 10.1088/1538-3873/ac9642

Casey, C. M. 2012, MNRAS, 425, 3094, doi: 10.1111/j.1365-2966.2012.21455.x

—. 2020, ApJ, 900, 68, doi: 10.3847/1538-4357/aba528

Casey, C. M., Narayanan, D., & Cooray, A. 2014, Phys. Rep., 541, 45, doi: 10.1016/j.physrep.2014.02.009

Casey, C. M., Zavala, J. A., Spilker, J., et al. 2018, ApJ, 862, 77, doi: 10.3847/1538-4357/aac82d

Casey, C. M., Zavala, J. A., Aravena, M., et al. 2019, ApJ, 887, 55, doi: 10.3847/1538-4357/ab52ff

Casey, C. M., Zavala, J. A., Manning, S. M., et al. 2021, ApJ, 923, 215, doi: 10.3847/1538-4357/ac2eb4

Casey, C. M., Kartaltepe, J. S., Drakos, N. E., et al. 2023, ApJ, 954, 31, doi: 10.3847/1538-4357/acc2bc

Casey, C. M., Akins, H. B., Shuntov, M., et al. 2024, ApJ, 965, 98, doi: 10.3847/1538-4357/ad2075

Ceraj, L., Smolčić, V., Delvecchio, I., et al. 2018, A&A, 620, A192, doi: 10.1051/0004-6361/201833935

Chabrier, G. 2003, PASP, 115, 763, doi: 10.1086/376392

Chapman, S. C., Lewis, G. F., Scott, D., Borys, C., & Richards, E. 2002, ApJ, 570, 557, doi: 10.1086/339498

Chapman, S. C., Richards, E. A., Lewis, G. F., Wilson, G., & Barger, A. J. 2001, ApJ, 548, L147, doi: 10.1086/319111

Chapman, S. C., Smail, I., Blain, A. W., & Ivison, R. J. 2004, ApJ, 614, 671, doi: 10.1086/423833

Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718, doi: 10.1086/309250

Chen, C.-C., Liao, C.-L., Smail, I., et al. 2022, ApJ, 929, 159, doi: 10.3847/1538-4357/ac61df

Chien, T. C. C., Ling, C.-T., Goto, T., et al. 2024, MNRAS, doi: 10.1093/mnras/stae1550

Civano, F., Marchesi, S., Comastri, A., et al. 2016, ApJ, 819, 62, doi: 10.3847/0004-637X/819/1/62

Combes, F. 2018, A&A Rev., 26, 5, doi: 10.1007/s00159-018-0110-4

Comrie, A., Wang, K.-S., Hsu, S.-C., et al. 2021, CARTA: The Cube Analysis and Rendering Tool for Astronomy, Zenodo, doi: 10.5281/zenodo.4905459

Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833, doi: 10.1088/0004-637X/712/2/833

Coogan, R. T., Daddi, E., Sargent, M. T., et al. 2018, MNRAS, 479, 703, doi: 10.1093/mnras/sty1446

Cowie, L. L., Barger, A. J., Hsu, L. Y., et al. 2017, ApJ, 837, 139, doi: 10.3847/1538-4357/aa60bb

Cox, P., Neri, R., Berta, S., et al. 2023, åp, 678, A26, doi: 10.1051/0004-6361/202346801

Cucciati, O., Tresse, L., Ilbert, O., et al. 2012, A&A, 539, A31, doi: 10.1051/0004-6361/201118010

da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595, doi: 10.1111/j.1365-2966.2008.13535.x

da Cunha, E., Walter, F., Smail, I. R., et al. 2015, ApJ, 806, 110, doi: 10.1088/0004-637X/806/1/110

Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686, doi: 10.1088/0004-637X/713/1/686

Daddi, E., Dannerbauer, H., Liu, D., et al. 2015, åp, 577, A46, doi: 10.1051/0004-6361/201425043

Dannerbauer, H., Lehnert, M. D., Lutz, D., et al. 2004, ApJ, 606, 664, doi: 10.1086/383138

Davé, R. 2008, MNRAS, 385, 147, doi: 10.1111/j.1365-2966.2008.12866.x

Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827, doi: 10.1093/mnras/stz937

Davé, R., Finlator, K., & Oppenheimer, B. D. 2012, MNRAS, 421, 98, doi: 10.1111/j.1365-2966.2011.20148.x

de Graaff, A., Setton, D. J., Brammer, G., et al. 2024, Nature Astronomy, doi: 10.1038/s41550-024-02424-3

de Jong, T., Klein, U., Wielebinski, R., & Wunderlich, E. 1985, åp, 147, L6

De Looze, I., Cormier, D., Lebouteiller, V., et al. 2014, A&A, 568, A62, doi: 10.1051/0004-6361/201322489

Decarli, R., Walter, F., Aravena, M., et al. 2016, ApJ, 833, 69, doi: 10.3847/1538-4357/833/1/69

Dekel, A., Sarkar, K. C., Birnboim, Y., Mandelker, N., & Li, Z. 2023, MNRAS, 523, 3201, doi: 10.1093/mnras/stad1557

Delhaize, J., Smolčić, V., Delvecchio, I., et al. 2017, åp, 602, A4, doi: 10.1051/0004-6361/201629430

Delvecchio, I., Smolčić, V., Zamorani, G., et al. 2017, åp, 602, A3, doi: 10.1051/0004-6361/201629367

Delvecchio, I., Daddi, E., Sargent, M. T., et al. 2021, åp, 647, A123, doi: 10.1051/0004-6361/202039647

Di Teodoro, E. M., & Fraternali, F. 2015, MNRAS, 451, 3021, doi: 10.1093/mnras/stv1213

Donevski, D., Lapi, A., Małek, K., et al. 2020, åp, 644, A144, doi: 10.1051/0004-6361/202038405

Donnan, C. T., McLeod, D. J., Dunlop, J. S., et al. 2023, MNRAS, 518, 6011, doi: 10.1093/mnras/stac3472

Draine, B. T., & Li, A. 2007, ApJ, 657, 810, doi: 10.1086/511055

Draine, B. T., Aniano, G., Krause, O., et al. 2014, ApJ, 780, 172, doi: 10.1088/0004-637X/780/2/172

Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2021, MNRAS, 500, 942, doi: 10.1093/mnras/staa3285

Dunlop, J. S., McLure, R. J., Biggs, A. D., et al. 2017, MNRAS, 466, 861, doi: 10.1093/mnras/stw3088

Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJS, 184, 158, doi: 10.1088/0067-0049/184/1/158

Endsley, R., Stark, D. P., Charlot, S., et al. 2021, MNRAS, 502, 6044, doi: 10.1093/mnras/stab432

Enia, A., Talia, M., Pozzi, F., et al. 2022, ApJ, 927, 204, doi: 10.3847/1538-4357/ac51ca

Erwin, P. 2015, ApJ, 799, 226, doi: 10.1088/0004-637X/799/2/226

Fabian, A. C. 2012, ARA&A, 50, 455, doi: 10.1146/annurev-astro-081811-125521

Faisst, A. L., Fudamoto, Y., Oesch, P. A., et al. 2020, MNRAS, 498, 4192, doi: 10.1093/mnras/staa2545

Faisst, A. L., Schaerer, D., Lemaux, B. C., et al. 2020, ApJS, 247, 61, doi: 10.3847/1538-4365/ab7ccd

Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, Rev. Mexicana Astron. Astrofis., 53, 385, doi: 10.48550/arXiv.1705.10877

Finkelstein, S. L., Bagley, M. B., Ferguson, H. C., et al. 2023, ApJ, 946, L13, doi: 10.3847/2041-8213/acade4

Fitzpatrick, E. L., & Massa, D. 2007, ApJ, 663, 320, doi: 10.1086/518158

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Franco, M., Elbaz, D., Béthermin, M., et al. 2018, åp, 620, A152, doi: 10.1051/0004-6361/201832928

Franco, M., Akins, H. B., Casey, C. M., et al. 2023, arXiv e-prints, arXiv:2308.00751, doi: 10.48550/arXiv.2308.00751

Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, MNRAS, 366, 767, doi: 10.1111/j.1365-2966.2006.09866.x

Fruscione, A., McDowell, J. C., Allen, G. E., et al. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6270, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. D. R. Silva & R. E. Doxsey, 62701V, doi: 10.1117/12.671760

Fudamoto, Y., Oesch, P. A., Faisst, A., et al. 2020a, A&A, 643, A4, doi: 10.1051/0004-6361/202038163

—. 2020b, A&A, 643, A4, doi: 10.1051/0004-6361/202038163

Furtak, L. J., Labbé, I., Zitrin, A., et al. 2023, arXiv e-prints, arXiv:2308.05735, doi: 10.48550/arXiv.2308.05735

Geach, J. E., Dunlop, J. S., Halpern, M., et al. 2017, MNRAS, 465, 1789, doi: 10.1093/mnras/stw2721

Gehrels, N. 1986, ApJ, 303, 336, doi: 10.1086/164079

Gentile, F., Talia, M., Behiri, M., et al. 2024a, ApJ, 962, 26, doi: 10.3847/1538-4357/ad1519

Gentile, F., Talia, M., Daddi, E., et al. 2024b, A&A, 687, A288, doi: 10.1051/0004-6361/202348623

Gentile, F., Casey, C. M., Akins, H. B., et al. 2024c, ApJ, 973, L2, doi: 10.3847/2041-8213/ad738a

Giavalisco, M. 2002, ARA&A, 40, 579, doi: 10.1146/annurev.astro.40.121301.111837

Gordon, K. D., Engelbracht, C. W., Rieke, G. H., et al. 2008, ApJ, 682, 336, doi: 10.1086/589567

Gottumukkala, R., Barrufet, L., Oesch, P. A., et al. 2023, arXiv e-prints, arXiv:2310.03787, doi: 10.48550/arXiv.2310.03787

Green, G. M. 2018, The Journal of Open Source Software, 3, 695, doi: 10.21105/joss.00695

Greene, J. E., Labbe, I., Goulding, A. D., et al. 2023, arXiv e-prints, arXiv:2309.05714, doi: 10.48550/arXiv.2309.05714

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35, doi: 10.1088/0067-0049/197/2/35

Grudić, M. Y., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2018, MNRAS, 475, 3511, doi: 10.1093/mnras/sty035

Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, MNRAS, 432, 23, doi: 10.1093/mnras/stt308

Gruppioni, C., Béthermin, M., Loiacono, F., et al. 2020, åp, 643, A8, doi: 10.1051/0004-6361/202038487

Gururajan, G., B'ethermin, M., Sulzenauer, N., et al. 2023, åp, 676, A89, doi: 10.1051/0004-6361/202346449

Harikane, Y., Ouchi, M., Oguri, M., et al. 2023, ApJS, 265, 5, doi: 10.3847/1538-4365/acaaa9

Hauser, M. G., & Dwek, E. 2001, ARA&A, 39, 249, doi: 10.1146/annurev.astro.39.1.249

Helou, G., Soifer, B. T., & Rowan-Robinson, M. 1985, ApJ, 298, L7, doi: 10.1086/184556

Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, MNRAS, 451, 2663, doi: 10.1093/mnras/stv705

Heywood, I., Jarvis, M. J., Hale, C. L., et al. 2022, MNRAS, 509, 2150, doi: 10.1093/mnras/stab3021

Hickox, R. C., & Alexander, D. M. 2018, ARA&A, 56, 625, doi: 10.1146/annurev-astro-081817-051803

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. 1983, Understanding robust and exploratory data anlysis

Högbom, J. A. 1974, åps, 15, 417

Hopkins, P. F., Cox, T. J., Kereš, D., & Hernquist, L. 2008a, ApJS, 175, 390, doi: 10.1086/524363

Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008b, ApJS, 175, 356, doi: 10.1086/524362

Hopkins, P. F., Wetzel, A., Wheeler, C., et al. 2023, MNRAS, 519, 3154, doi: 10.1093/mnras/stac3489

Huang, J. S., Zheng, X. Z., Rigopoulou, D., et al. 2011, ApJ, 742, L13, doi: 10.1088/2041-8205/742/1/L13

Hughes, D. H., Serjeant, S., Dunlop, J., et al. 1998, Nature, 394, 241, doi: 10.1038/28328

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, åp, 457, 841, doi: 10.1051/0004-6361:20065138

Irwin, M. J., Ibata, R. A., Lewis, G. F., & Totten, E. J. 1998, ApJ, 505, 529, doi: 10.1086/306213

Jarvis, M., Taylor, R., Agudo, I., et al. 2016, in MeerKAT Science: On the Pathway to the SKA, 6

Jin, S., Daddi, E., Liu, D., et al. 2018, ApJ, 864, 56, doi: 10.3847/1538-4357/aad4af

Jin, S., Daddi, E., Magdis, G. E., et al. 2019, ApJ, 887, 144, doi: 10.3847/1538-4357/ab55d6

—. 2022, åp, 665, A3, doi: 10.1051/0004-6361/202243341

Jin, S., Sillassen, N. B., Hodge, J., et al. 2024, arXiv e-prints, arXiv:2407.07585, doi: 10.48550/arXiv.2407.07585

Karwal, T., & Kamionkowski, M. 2016, Phys. Rev. D, 94, 103523, doi: 10.1103/PhysRevD.94.103523

Kennicutt, Robert C., Jr. 1998, ARA&A, 36, 189, doi: 10.1146/annurev.astro.36.1.189

Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531, doi: 10.1146/annurev-astro-081811-125610

Khusanova, Y., Bethermin, M., Le Fèvre, O., et al. 2021, åp, 649, A152, doi: 10.1051/0004-6361/202038944

Klypin, A., Poulin, V., Prada, F., et al. 2021, MNRAS, 504, 769, doi: 10.1093/mnras/stab769

Kocevski, D. D., Onoue, M., Inayoshi, K., et al. 2023, ApJ, 954, L4, doi: 10.3847/2041-8213/ace5a0

Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, ApJS, 172, 196, doi: 10.1086/520086

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36, doi: 10.1088/0067-0049/197/2/36

Kokorev, V., Caputi, K. I., Greene, J. E., et al. 2024, arXiv e-prints, arXiv:2401.09981, doi: 10.48550/arXiv.2401.09981

Kümmel, M., Bertin, E., Schefer, M., et al. 2020, in Astronomical Society of the Pacific Conference Series, Vol. 527, Astronomical Data Analysis Software and Systems XXIX, ed. R. Pizzo, E. R. Deul, J. D. Mol, J. de Plaa, & H. Verkouter, 29

Labbé, I., Bouwens, R., Illingworth, G. D., & Franx, M. 2006, ApJ, 649, L67, doi: 10.1086/508512

Labbé, I., van Dokkum, P., Nelson, E., et al. 2023a, Nature, 616, 266, doi: 10.1038/s41586-023-05786-2

Labbé, I., Greene, J. E., Bezanson, R., et al. 2023b, arXiv e-prints, arXiv:2306.07320, doi: 10.48550/arXiv.2306.07320

Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24, doi: 10.3847/0067-0049/224/2/24

Lang, D., Hogg, D. W., & Mykytyn, D. 2016, The Tractor: Probabilistic Astronomical Source Detection and Measurement

Langan, I., Popping, G., Ginolfi, M., et al. 2024, A&A, 689, A283, doi: 10.1051/0004-6361/202449626

Lapi, A., Raimundo, S., Aversa, R., et al. 2014, ApJ, 782, 69, doi: 10.1088/0004-637X/782/2/69

Lapi, A., Pantoni, L., Zanisi, L., et al. 2018, ApJ, 857, 22, doi: 10.3847/1538-4357/aab6af

Larson, R. L., Hutchison, T. A., Bagley, M., et al. 2023, ApJ, 958, 141, doi: 10.3847/1538-4357/acfed4

Le Fèvre, O., Béthermin, M., Faisst, A., et al. 2020, A&A, 643, A1, doi: 10.1051/0004-6361/201936965

Le Floc'h, E., Aussel, H., Ilbert, O., et al. 2009, ApJ, 703, 222, doi: 10.1088/0004-637X/703/1/222

Leauthaud, A., Massey, R., Kneib, J.-P., et al. 2007, ApJS, 172, 219, doi: 10.1086/516598

Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019, ApJ, 876, 3, doi: 10.3847/1538-4357/ab133c

Lenz, D., Hensley, B. S., & Doré, O. 2017, ApJ, 846, 38, doi: 10.3847/1538-4357/aa84af

Liang, L., Feldmann, R., Kereš, D., et al. 2019, MNRAS, 489, 1397, doi: 10.1093/mnras/stz2134

Liu, D., Lang, P., Magnelli, B., et al. 2019, ApJS, 244, 40, doi: 10.3847/1538-4365/ab42da

Long, A. S., Casey, C. M., McKinney, J., et al. 2024, arXiv e-prints, arXiv:2408.14546, doi: 10.48550/arXiv.2408.14546

Low, C., & Lynden-Bell, D. 1976, MNRAS, 176, 367, doi: 10.1093/mnras/176.2.367

Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, åp, 532, A90, doi: 10.1051/0004-6361/201117107

Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415, doi: 10.1146/annurev-astro-081811-125615

Magnelli, B., Popesso, P., Berta, S., et al. 2013, åp, 553, A132, doi: 10.1051/0004-6361/201321371

Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006, MNRAS, 368, 715, doi: 10.1111/j.1365-2966.2006.10156.x

Matthee, J., Naidu, R. P., Brammer, G., et al. 2023, arXiv e-prints, arXiv:2306.05448, doi: 10.48550/arXiv.2306.05448

McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, åp, 544, A156, doi: 10.1051/0004-6361/201219507

McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565, doi: 10.1146/annurev.astro.45.051806.110602

McKinney, J., Casey, C. M., Long, A. S., et al. 2024, arXiv e-prints, arXiv:2408.08346, doi: 10.48550/arXiv.2408.08346

Molnár, D. C., Sargent, M. T., Leslie, S., et al. 2021, MNRAS, 504, 118, doi: 10.1093/mnras/stab746

Moneti, A., Euclid Collaboration, McCracken, H. J., et al. 2022, åp, 658, A126, doi: 10.1051/0004-6361/202142361

Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121, doi: 10.1093/mnras/sts261

Novak, M., Smolčić, V., Schinnerer, E., et al. 2018, åp, 614, A47, doi: 10.1051/0004-6361/201731635

Novak, M., Smolčić, V., Delhaize, J., et al. 2017, åp, 602, A5, doi: 10.1051/0004-6361/201629436

Oesch, P. A., Brammer, G., Naidu, R. P., et al. 2023, MNRAS, 525, 2864, doi: 10.1093/mnras/stad2411

Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713, doi: 10.1086/160817

Oliver, S. J., Bock, J., Altieri, B., et al. 2012, MNRAS, 424, 1614, doi: 10.1111/j.1365-2966.2012.20912.x

Ostriker, E. C., & Shetty, R. 2011, ApJ, 731, 41, doi: 10.1088/0004-637X/731/1/41

Pallottini, A., Ferrara, A., Gallerani, S., et al. 2022, MNRAS, 513, 5621, doi: 10.1093/mnras/stac1281

Pantoni, L., Lapi, A., Massardi, M., Goswami, S., & Danese, L. 2019, ApJ, 880, 129, doi: 10.3847/1538-4357/ab2adc

Papadopoulos, P. P., Thi, W. F., & Viti, S. 2004, MNRAS, 351, 147, doi: 10.1111/j.1365-2966.2004.07762.x

Papovich, C., Cole, J. W., Yang, G., et al. 2023, ApJ, 949, L18, doi: 10.3847/2041-8213/acc948

Pérez-González, P. G., Barro, G., Annunziatella, M., et al. 2023, The Astrophysical Journal, 946, L16, doi: 10.3847/2041-8213/acb3a5

Pérez-González, P. G., Barro, G., Rieke, G. H., et al. 2024, arXiv e-prints, arXiv:2401.08782, doi: 10.48550/arXiv.2401.08782

Pillepich, A., Springel, V., Nelson, D., et al. 2018, MNRAS, 473, 4077, doi: 10.1093/mnras/stx2656

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, åp, 641, A6, doi: 10.1051/0004-6361/201833910

Poulin, V., Boddy, K. K., Bird, S., & Kamionkowski, M. 2018, Phys. Rev. D, 97, 123504, doi: 10.1103/PhysRevD.97.123504

Puglisi, A., Daddi, E., Liu, D., et al. 2019, ApJ, 877, L23, doi: 10.3847/2041-8213/ab1f92

Ranalli, P., Comastri, A., & Setti, G. 2003, åp, 399, 39, doi: 10.1051/0004-6361:20021600

Roman-Oliveira, F., Fraternali, F., & Rizzo, F. 2023, MNRAS, 521, 1045, doi: 10.1093/mnras/stad530

Rowan-Robinson, M., Broadhurst, T., Lawrence, A., et al. 1991, Nature, 351, 719, doi: 10.1038/351719a0

Rowan-Robinson, M., Oliver, S., Wang, L., et al. 2016, MNRAS, 461, 1100, doi: 10.1093/mnras/stw1169

Saintonge, A., Lutz, D., Genzel, R., et al. 2013, ApJ, 778, 2, doi: 10.1088/0004-637X/778/1/2

Salim, S., & Narayanan, D. 2020, ARA&A, 58, 529, doi: 10.1146/annurev-astro-032620-021933

Sancisi, R., Fraternali, F., Oosterloo, T., & van der Hulst, T. 2008, åpr, 15, 189, doi: 10.1007/s00159-008-0010-0

Saunders, W., Rowan-Robinson, M., Lawrence, A., et al. 1990, MNRAS, 242, 318, doi: 10.1093/mnras/242.3.318

Schaerer, D., & de Barros, S. 2009, åp, 502, 423, doi: 10.1051/0004-6361/200911781

Schaerer, D., Ginolfi, M., Béthermin, M., et al. 2020, A&A, 643, A3, doi: 10.1051/0004-6361/202037617

Schechter, P. 1976, ApJ, 203, 297, doi: 10.1086/154079

Schinnerer, E., Smolčić, V., Carilli, C. L., et al. 2007, ApJS, 172, 46, doi: 10.1086/516587

Schinnerer, E., Sargent, M. T., Bondi, M., et al. 2010, ApJS, 188, 384, doi: 10.1088/0067-0049/188/2/384

Schmidt, M. 1968, ApJ, 151, 393, doi: 10.1086/149446

Schreiber, C., Elbaz, D., Pannella, M., et al. 2018a, åp, 609, A30, doi: 10.1051/0004-6361/201731506

Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, åp, 575, A74, doi: 10.1051/0004-6361/201425017

Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018b, åp, 618, A85, doi: 10.1051/0004-6361/201833070

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1, doi: 10.1086/516585

Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Shen, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2020, MNRAS, 495, 3252, doi: 10.1093/mnras/staa1381

Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119, doi: 10.1046/j.1365-8711.1999.02692.x

Shuntov, M., McCracken, H. J., Gavazzi, R., et al. 2022, åp, 664, A61, doi: 10.1051/0004-6361/202243136

Shuntov, M., Ilbert, O., Toft, S., et al. 2024, arXiv e-prints, arXiv:2410.08290, doi: 10.48550/arXiv.2410.08290

Simpson, J. M., Swinbank, A. M., Smail, I., et al. 2014, ApJ, 788, 125, doi: 10.1088/0004-637X/788/2/125

Simpson, J. M., Smail, I., Swinbank, A. M., et al. 2019, ApJ, 880, 43, doi: 10.3847/1538-4357/ab23ff

Smail, I., Ivison, R. J., & Blain, A. W. 1997, ApJ, 490, L5, doi: 10.1086/311017

Smail, I., Dudzevičiūtė, U., Stach, S. M., et al. 2021, MNRAS, 502, 3426, doi: 10.1093/mnras/stab283

Smolčić, V., Novak, M., Bondi, M., et al. 2017, åp, 602, A1, doi: 10.1051/0004-6361/201628704

Snyder, G. F., Lotz, J. M., Rodriguez-Gomez, V., et al. 2017, MNRAS, 468, 207, doi: 10.1093/mnras/stx487

Sommovigo, L., Ferrara, A., Pallottini, A., et al. 2020, MNRAS, 497, 956, doi: 10.1093/mnras/staa1959

Stach, S. M., Dudzevičiūtė, U., Smail, I., et al. 2019, MNRAS, 487, 4648, doi: 10.1093/mnras/stz1536

Steinhardt, Charles. L., Capak, P., Masters, D., & Speagle, J. S. 2016, ApJ, 824, 21, doi: 10.3847/0004-637X/824/1/21

Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, ApJ, 783, L14, doi: 10.1088/2041-8205/783/1/L14

Swinbank, A. M., Simpson, J. M., Smail, I., et al. 2014, MNRAS, 438, 1267, doi: 10.1093/mnras/stt2273

Szalay, A. S., Connolly, A. J., & Szokoly, G. P. 1999, AJ, 117, 68, doi: 10.1086/300689

Tacconi, L. J., Genzel, R., & Sternberg, A. 2020, ARA&A, 58, 157, doi: 10.1146/annurev-astro-082812-141034

Tadhunter, C. 2016, A&A Rev., 24, 10, doi: 10.1007/s00159-016-0094-x

Talia, M., Cimatti, A., Giulietti, M., et al. 2021, ApJ, 909, 23, doi: 10.3847/1538-4357/abd6e3

Talia, M., Cimatti, A., Pozzetti, L., et al. 2015, A&A, 582, A80, doi: 10.1051/0004-6361/201425430

Thomson, A. P., Ivison, R. J., Simpson, J. M., et al. 2014, MNRAS, 442, 577, doi: 10.1093/mnras/stu839

Toft, S., Smolčić, V., Magnelli, B., et al. 2014, ApJ, 782, 68, doi: 10.1088/0004-637X/782/2/68

Torrey, P., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2017, MNRAS, 467, 2301, doi: 10.1093/mnras/stx254

Traina, A., Gruppioni, C., Delvecchio, I., et al. 2024, åp, 681, A118, doi: 10.1051/0004-6361/202347048

Valentino, F., Tanaka, M., Davidzon, I., et al. 2020a, ApJ, 889, 93, doi: 10.3847/1538-4357/ab64dc

Valentino, F., Magdis, G. E., Daddi, E., et al. 2020b, ApJ, 890, 24, doi: 10.3847/1538-4357/ab6603

Vallini, L., Tielens, A. G. G. M., Pallottini, A., et al. 2019, MNRAS, 490, 4502, doi: 10.1093/mnras/stz2837

van der Vlugt, D., Hodge, J. A., Algera, H. S. B., et al. 2022, arXiv e-prints, arXiv:2204.04167

van der Vlugt, D., Hodge, J. A., Jin, S., et al. 2023, ApJ, 951, 131, doi: 10.3847/1538-4357/acd549

Vardoulaki, E., Jiménez Andrade, E. F., Delvecchio, I., et al. 2021, åp, 648, A102, doi: 10.1051/0004-6361/202039488

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Vito, F., Brandt, W. N., Yang, G., et al. 2018, MNRAS, 473, 2378, doi: 10.1093/mnras/stx2486

Walter, F., Weiß, A., Downes, D., Decarli, R., & Henkel, C. 2011, ApJ, 730, 18, doi: 10.1088/0004-637X/730/1/18

Walter, F., Decarli, R., Aravena, M., et al. 2016, ApJ, 833, 67, doi: 10.3847/1538-4357/833/1/67

Wang, B., Leja, J., Atek, H., et al. 2024a, ApJ, 963, 74, doi: 10.3847/1538-4357/ad187c

Wang, B., de Graaff, A., Davies, R. L., et al. 2024b, arXiv e-prints, arXiv:2403.02304, doi: 10.48550/arXiv.2403.02304

Wang, T., Elbaz, D., Schreiber, C., et al. 2016, ApJ, 816, 84, doi: 10.3847/0004-637X/816/2/84

Wang, T., Schreiber, C., Elbaz, D., et al. 2019, Nature, 572, 211, doi: 10.1038/s41586-019-1452-4

Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJS, 258, 11, doi: 10.3847/1538-4365/ac3078

Weaver, J. R., Davidzon, I., Toft, S., et al. 2023a, A&A, 677, A184, doi: 10.1051/0004-6361/202245581

Weaver, J. R., Zalesky, L., Kokorev, V., et al. 2023b, ApJS, 269, 20, doi: 10.3847/1538-4365/acf850

Wechsler, R. H., & Tinker, J. L. 2018, ARA&A, 56, 435, doi: 10.1146/annurev-astro-081817-051756

Weibel, A., de Graaff, A., Setton, D. J., et al. 2024, arXiv e-prints, arXiv:2409.03829, doi: 10.48550/arXiv.2409.03829

Weiß, A., De Breuck, C., Marrone, D. P., et al. 2013, ApJ, 767, 88, doi: 10.1088/0004-637X/767/1/88

Whitaker, K. E., Pope, A., Cybulski, R., et al. 2017, ApJ, 850, 208, doi: 10.3847/1538-4357/aa94ce

Whitler, L., Stark, D. P., Endsley, R., et al. 2022, arXiv e-prints, arXiv:2206.05315

Williams, C. C., Alberts, S., Ji, Z., et al. 2023, arXiv e-prints, arXiv:2311.07483, doi: 10.48550/arXiv.2311.07483

Woodrum, C., Rieke, M., Ji, Z., et al. 2023, arXiv e-prints, arXiv:2310.18464, doi: 10.48550/arXiv.2310.18464

Wuyts, S., Förster Schreiber, N. M., Lutz, D., et al. 2011, ApJ, 738, 106, doi: 10.1088/0004-637X/738/1/106

Xiao, M., Oesch, P., Elbaz, D., et al. 2023, arXiv e-prints, arXiv:2309.02492, doi: 10.48550/arXiv.2309.02492

Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, ApJ, 554, 803, doi: 10.1086/323145

Zanella, A., Daddi, E., Magdis, G., et al. 2018, MNRAS, 481, 1976, doi: 10.1093/mnras/sty2394

Zavala, J. A., Casey, C. M., Manning, S. M., et al. 2021, ApJ, 909, 165, doi: 10.3847/1538-4357/abdb27

Zavala, J. A., Buat, V., Casey, C. M., et al. 2023, ApJ, 943, L9, doi: 10.3847/2041-8213/acacfe