Alma Mater Studiorum · University of Bologna

School of Science
Department of Physics and Astronomy
Master Degree in Astrophysics and Cosmology

Tracing Starbirth in the Shadows
Exploring Radio-selected NIR Dark Galaxies to study the Star Formation Rate Density

Supervisor:
Dr. Margherita Talia

Co-supervisors:
Prof. Francesca Pozzi
Dr. Fabrizio Gentile

Submitted by:
Matteo Sapori

Academic Year 2022/2023

Contents

Chapter 1

Abstract

The ongoing exploration of the Universe has brought a series of new questions, with one in particular standing out: how did galaxies form and evolve over cosmic times?
Despite the progress achieved in this research field, some key aspects are still open.
A Fundamental quantity in the study of galaxy formation and evolution is the star formation rate density (SFRD) that quantifies the star formation per unit of comoving volume as a function of redshift. This quantity has traditionally been investigated through optical and UV surveys. However, recent studies have opened new questions: are these surveys missing a population of dust-obscured optically faint galaxies? And, if so, what is their contribution to the SFRD and their role in the broader narrative of galaxy formation?
The advent of revolutionary tools such as the James Webb Space Telescope (JWST) and the Atacama Large Millimeter/submillimeter Array (ALMA) has highlighted how limited our knowledge in this field still is, especially regarding dust-obscured galaxies, which can be invisible optical/UV and near-infrared (NIR) observations.
This thesis places itself within this context. I focused on the search of highly dust-obscured star-forming galaxies (DSFGs) using a selection criterion based on radio-emission - a SFR tracer unaffected by dust absorption - drawing inspiration from the work of Talia et al. (2021) and follow-up studies such as Behiri et al. (2023) and Gentile et al. (2023). The search was performed within the J1030 field, a cosmological field of $\approx 30\times30$ arcmin^2 with a multiwavelength coverage from UV to sub/mm, including one of the deepest x-ray and 1.4 GHz observations of an extragalactic field to-date (Nanni et al. 2018; D'Amato et al. 2022). This study is the first research of this kind that has been carried out in this field.
To solve the problem of source blending created by the large Point Spread Function (PSF) of my observations in Mid Infrared (MIR) to sub/mm, I contributed to the development of a new tool for the deblending of NIR-Dark galaxies in MIR-to-FIR maps: PhoEBO (Photometric Extractor For Blended Objects; Gentile et al. 2023). Also, I applied a personally developed technique for the deblending of sub/mm bands by following those presented in Jin et al. (2018).
Once the galaxies have been selected, their characterization has been done through SED

fitting using the MAGPHYS+PHOTO-Z code (Battisti et al., 2019), revealing a population of DSFGs that extends from z=0 to the so-called 'cosmic noon' at z=2-3, with some galaxies reaching z>5. Most of these galaxies have stellar masses between 10^{10} and 10^{11} $M_{\odot}$. Interestingly, about 28% of sources with coverage in FIR/sub-mm bands have integrated (8μm to 1000μm) infrared luminosities similar to local ULIRGs, galaxies that experience exceptionally high rates of star formation (SFRs $>$ 100 $M_{\odot}$ yr^{-1}).

The relevance of this study becomes even more evident when considering the contribution of these galaxies to the SFRD, estimated to be between 3% at z≈2 and 13% at z≈3. These values, which do not account for the extrapolation to lower luminosities than the detection limits of radio observations, are only lower limits, thus indicating how optical and UV surveys may have missed a significant population of galaxies. Furthermore, the numerical density of these galaxies, varying between 10^{-6} Mpc^{-3} for z>3 and 10^{-5} Mpc^{-3} for 2<z<3, is not easily reproduced by simulations and partially aligns with that of massive quiescent galaxies at z≈2, of which DSFGs could be the progenitors.

Chapter 2

Scientific Context

2.1 Measuring Star Formation

Galaxies are complex systems emitting light across a broad wavelength range. The primary goal of this thesis is to leverage these data to study characteristics such as stellar mass (M_*), Star Formation Rate (SFR), and Star Formation History (SFH)(see Kennicutt Jr & Evans II 2012 for a comprehensive review). Effectively, all star formation tracers identify the formation of massive and young stars, as these contribute the most to the energy budget of stellar populations. Different wavelengths are sensitive to various processes occurring within galaxies and and to different ranges of stellar masses as well as the age of the population. For instance, Hα emission primarily originates from HII regions photoionized by massive O-type stars, typically with lifespans of less than 30 Myr. In contrast, UV rest-frame emission primarily originates from massive O, B, and A-type stars, thereby being sensitive to a wider range of masses.
Below, I summarize the main estimators of star formation, namely UV, Hα, infrared, and radio luminosities.

- **UV Emission**: for a typical Initial Mass Function (IMF) low-mass stars make up the majority of the total mass of all the stars in a given population. However, the integrated luminosity of the same population is dominated by the UV emission from massive and young stars. Since UV radiation directly traces the photosphere of these stars, and considering their short lifespans (less than 100 Myr), UV emission serves as the most direct tracer of recent SFR. In the literature, the emission at 1500 Å has been primarily used for studies in both the local Universe and at high redshift because it directly relates to the most massive stars. However, this approach has a limitation for local Universe studies, as it requires observations beyond Earth's atmosphere. On the other hand, studies at longer wavelengths (e.g., 2300 Å) are not subject to this limitation but are sensitive to a broader range of masses. This sensitivity complicates the conversion between luminosity and mass per year.
  Another major challenge when dealing with rest-frame UV observations is the severe

attenuation by interstellar dust. Dust extinction is more pronounced in the UV range, meaning that even a small amount of dust can systematically affect results if observed fluxes are not accurately corrected. If the intrinsic color of the emitting stellar population is known a priori, the FUV–NUV color or the UV spectral slope (commonly denoted as β) can be employed to estimate dust attenuation. Several calibrations for this purpose have been published (e.g. Treyer et al. 2007; Calzetti 2000). The effectiveness of these methods largely depends on the assumed nature of the dust grains, the dust extinction curve, and the complex interplay of geometry and scattering when averaged over a large physical region (see Calzetti 2000).

- **Hα Emission**: the Hα recombination line emission from hydrogen (656,281 nm) is a direct star formation indicator due to its close connection with HII regions in which it forms. These regions originate as a consequence of the UV emissions of massive stars with M $>$ 15 M$_\odot$ (see Kennicutt Jr & Evans II 2012) that ionize the molecular hydrogen in molecular clouds, where the star formation is happening. Thus, Hα acts as an almost instantaneous SFR indicator.
  The largest systematic errors affecting Hα-based SFRs are dust attenuation and sensitivity to the IMF in regions with low absolute SFRs, as when the SFR is low, the scarcity of massive stars can lead to an overestimate or underestimate of the SFR, depending on how the IMF is sampled in those specific regions. For regions with modest dust attenuations, the ratios of Balmer recombination lines (Balmer decrement) can be used to correct for dust, but it offers only approximate corrections for attenuation because of variations on scales smaller than the resolution of the observation (Kennicutt Jr & Evans II, 2012).
  Besides, warm ionized gas can also produce Hα emission, possibly accounting for 20–60% of the total Hα flux in local star-forming galaxies (Chevance et al., 2020).

- **Infrared (IR) Emission**: the energy absorbed by dust in the UV spectrum is reemitted in the Mid-Infrared (MIR) and Far-Infrared (FIR) bands, making IR emission another instrument for measuring star formation. Since most UV emission originates from stars, IR luminosity is often interpreted as directly proportional to the fraction of UV emission absorbed.
  Ideally, the IR luminosity of a galaxy can be derived by fitting models of dust emission (e.g. Draine 2005) and absorption to observations of data at different wavelengths, like Modified Black Body or complex physically motivated dust emission models (e.g. da Cunha et al. 2008). However, obtaining or analyzing these observations can be challenging: the main issue is that the majority of data available today are obtained through instruments that often lack the necessary sensitivity and resolution[1], especially when studying distant or faint galaxies. Additionally, the instrument's low reso-

[1]instruments like Hershel and SCUBA have been milestones in FIR/sub-mm observations but suffered from these limitations, see also Sec 2.2.2

lution makes it hard to accurately match the observed FIR emission with its source at other wavelengths, making follow-up studies with higher resolution essential (see Jin et al. 2018; Simpson et al. 2020). Additionally, the dust emission spectrum is complex and composed of multiple components at different temperatures; most of the dust mass fraction in a galaxy is in the form of cold dust (20-60 K), which predominantly contributes to the emission in FIR and sub-millimeter (50-1000 μm) ranges. The MIR spectrum (3 - 20 μm) is particularly complex, rich in features such as silicate absorption and Polycyclic Aromatic Hydrocarbons (PAH) emission (Draine et al., 2021).
Similar to other SFR indicators, dust emission is subject to significant systematic effects: just as UV and visible tracers fail to account for radiation attenuated by dust, infrared emission overlooks the starlight that is not absorbed by dust, as discussed by Hirashita et al. (2001), this "missing" unattenuated component varies significantly across different galaxy types. It ranges from virtually zero in dusty starburst galaxies to nearly 100% in dust-poor dwarf galaxies and metal-poor regions of more massive galaxies.
Dust heated by old stellar populations can also contribute to FIR emissions, with this effect being significant in old/evolved systems and/or with minimal star formation activity.
AGN can also contribute to MIR emissions due to dust emission from the torus and may even dominate the emission from the warm dust component inside the host galaxy. These last two cases are often a collateral effect of the fact that, typically, the conversions between IR luminosity and SFR assume a negligible contribution from AGNs and old stellar populations.

- **Radio Emission**: galaxies emit a continuous radio spectrum at centimeter wavelengths, primarily composed of two elements: a free-free component with a relatively flat spectrum, and a synchrotron component with a steeper spectrum. The latter dominates the integrated radio emission at frequencies lower than 5 GHz. The free-free component, which varies with ionizing luminosity and is marginally dependent on electrons temperature, can be isolated using multi-frequency radio measurements or high-frequency data. This isolation allows for an estimation of the SFR based on photoionization, avoiding the complexities associated with dust attenuation.
  At lower frequencies, the emission is mostly synchrotron radiation, generated by charged particles produced by supernovae; multiple observations confirm a strong correlation between this non-thermal emission and the FIR emission of galaxies, suggesting the use of synchrotron emission as a SFR indicator (Madau & Dickinson 2014, Dale & Helou 2002a).
  Current calibrations of the relation between radio continuum and SFR are based on FIR calibrations, exploiting the tight correlation between radio and IR (Dale & Helou, 2002a). This calibration is heavily influenced by frequency, given the steep nature of the synchrotron spectrum, and typically refers to 1.4 GHz.

A peculiar aspect emerges in analyzing low-dust-content galaxies: while IR-based SFR calibrations become less effective, the radio-IR correlation remains surprisingly tight and almost linear across the entire luminosity spectrum (see Molnár et al. 2021). This phenomenon can be explained by reduced dust opacity in lower-mass galaxies, which coincides with a decrease in synchrotron emission compared to other SFR indicators. This decrease in synchrotron luminosity implies a physical decline in emission per unit of SFR. While radio emission serves as an excellent tracer of star formation and is free from dust extinction, obtaining sufficiently deep observations to study ordinary star-forming galaxies at high redshifts often poses a challenge. Technological advancements in receivers, like those implemented in the Karl Jansky Very Large Array (JVLA), have further promoted the use of the radio continuum as a primary tool for identifying high-redshift star-forming galaxies and estimating their SFRs.
Like other tracers, radio emission has its drawbacks: AGN can contribute significantly to a galaxy's radio luminosity; as synchrotron emission, originating from relativistic jets, and free-free emission from hot gas and dust near the AGN, are key processes resulting from accretion activities. In such cases is necessary to utilize additional data across different spectral bands to differentiate the radio components associated with star formation from those originating from AGN activity.

2.1.1 The Star Formation Main Sequence

An established relation is the so-called star formation main sequence (SFMS) (Fig: 2.1), which represents an approximately linear correlation between the SFR and the stellar mass of star-forming galaxies (Speagle et al., 2014). This relation holds across a wide range of redshifts, both low (z<1; Brinchmann et al. 2004; Saintonge & Catinella 2022) and high (z $>$ 1; Elbaz et al. 2007), suggesting a possible universal nature of the star formation process in galaxies over much of cosmic time.
The SFMS is characterized by its tightness of about 0.2 dex (Speagle et al., 2014), and, in the literature, is expressed by the relation:

$$SFR \sim M_*^{\beta} \tag{2.1}$$

where the slope of SFMS (β), varies between 0.6 and 1, depending on the selection criteria used and the method by which the SFR and M_* values are derived, with some studies suggesting a flattening of the MS for high ($M_* > 10^{11}$ $M_\odot$) masses (Popesso et al., 2023). Furthermore, the normalization of the SFMS increases from z=0 to z ≈3.5 as $(1+z)^{3.5}$, after which it tends to flatten at higher redshifts (Tasca et al., 2015).
In addition to characterizing star-forming galaxies, the SFRMS provides a natural way to define starburst or quiescent galaxies. Starburst galaxies are those with star formation rates significantly higher than what the SFMS predicts for their stellar mass, indicating a relatively much higher level of star formation activity. Conversely, quiescent galaxies have SFRs much lower than expected based on their mass, indicating a subdued or halted star

formation process. To effectively identify these galaxies, it is useful to employ the Specific Star Formation Rate (sSFR), which is calculated as the ratio of SFR to M_*, serving as a metric for distinguishing between actively star-forming and quiescent galaxies. The small observed dispersion around the SFMS, coupled with the fact that the majority of star-forming galaxies reside on it, implies that most star formation in the Universe occurs in a steady state (Noeske et al., 2007). It also suggests that the fraction of a given star-forming galaxy's lifetime during which it lies significantly above the SFMS, possibly due to merger-induced starbursts, is small.

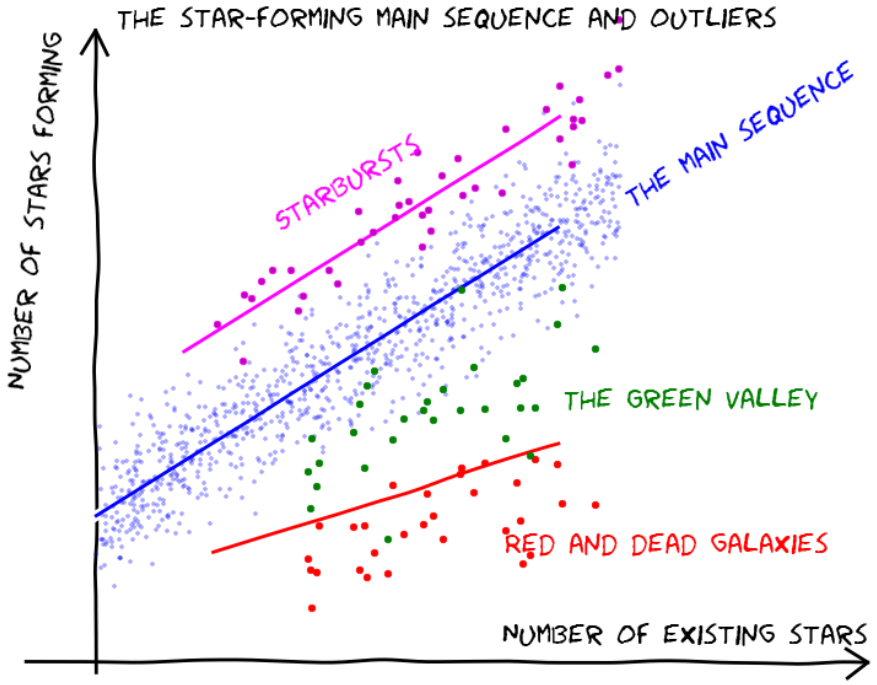


Figure 2.1: *SFR-M_* plane. The SFMS is highlighted in blue, galaxies with high star formation rates (starbursts) are in purple, quiescent galaxies with low star formation rates are in red, and the green valley, containing star-forming galaxies in the process of quenching, is depicted in green. Image taken from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS[2]) webpage.*

2.2 The Cosmic Star Formation Rate Density

The cosmic Star Formation Rate Density (SFRD) quantifies the average star formation rate per comoving volume of the Universe at a given redshift. The study of the evolution of this quantity is of paramount importance in theoretical models, as it represents a fundamental

[2] Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS)

constraint on the growth of stellar mass in galaxies throughout their evolution. estimating the SFRD across different epochs allows to trace the history of galaxy formation and evolution, providing crucial insights into the processes that have shaped the observable Universe. According to Madau & Dickinson (2014), who worked on a sample of optical/UV-selected galaxies, the SFRD has experienced a notable evolution over cosmic time. These changes can be traced back to the epoch of reionization, around redshift z = 8. The consensus among the scientific community is that the SFRD exhibited a phase of constant growth, scaling as $(1+z)^{-2.9}$ (black line in Fig 2.2), from z ≈ 8 up to about z = 2-3, this period, peaking around z ≈ 2 is often referred to as the "cosmic noon", a time when the Universe seems to have experienced the highest rates of star formation.

Following that, the SFRD began a decline phase, decreasing as $(1+z)^{2.7}$, continuing up to the present time. This trend signifies a gradual slowdown in the rate of star formation across the Universe, indicating a shift in the cosmic environment and galaxy evolution processes. The reasons behind this shift are complex and involve a variety of astrophysical processes, including the exhaustion of gas required for star formation in galaxies, the heating and ionization of intergalactic medium, and the evolution of galactic structures over time.

Currently, the study of the SFRD is crucial in astrophysics. It is essential for understanding galaxy formation and evolution mechanisms, imposing major constraints on the rate of stellar mass accretion in galaxies, and identifying the dominant star formation regime at specific redshifts.

However, the evolution of the SFRD at z > 3 faces significant limitations due to the lack of observations capable of probing such early cosmic epochs, resulting in large uncertainties. Before the advent of state-of-the-art instruments like the James Webb Space Telescope (JWST, Gardner et al. 2006) or the Atacama Large Millimeter/submillimeter Array (ALMA; Wootten & Thompson 2009), most high-redshift data was confined to UV observations (represented in Fig 2.2 by magenta pentagons from Bouwens et al. (2012), and black asterisks from Schenker et al. 2013). Further complicating matters is the current limited understanding of the evolution of dust content and temperature at high z (see Hwang et al. 2010; Jones & Stanway 2023), leading to potentially incorrect dust corrections. Therefore, it is necessary to have complementary observations across multiple bands to compare results, but, as previously mentioned, achieving this has not always been viable. The few estimates in the infrared in Fig 2.2 are derived from Gruppioni et al. (2013) (dark red hexagons), based on a study of sources observed with Herschel Space Observatory (Pilbratt et al., 2004) at 70, 100, and 160 μm.

It is also noteworthy that the redshift evolution of the SFRD closely mirrors that of the cosmic history of supermassive black hole accretion rate (BHAR), as depicted in Fig 2.3 (red line). This similitude suggests a connection between star formation and the accretion of Supermassive Black Holes (SMBHs). This correlation implies that the processes driving star formation and SMBH growth may be interconnected, possibly through feedback mechanisms in galaxies that regulate both star formation and black hole growth.

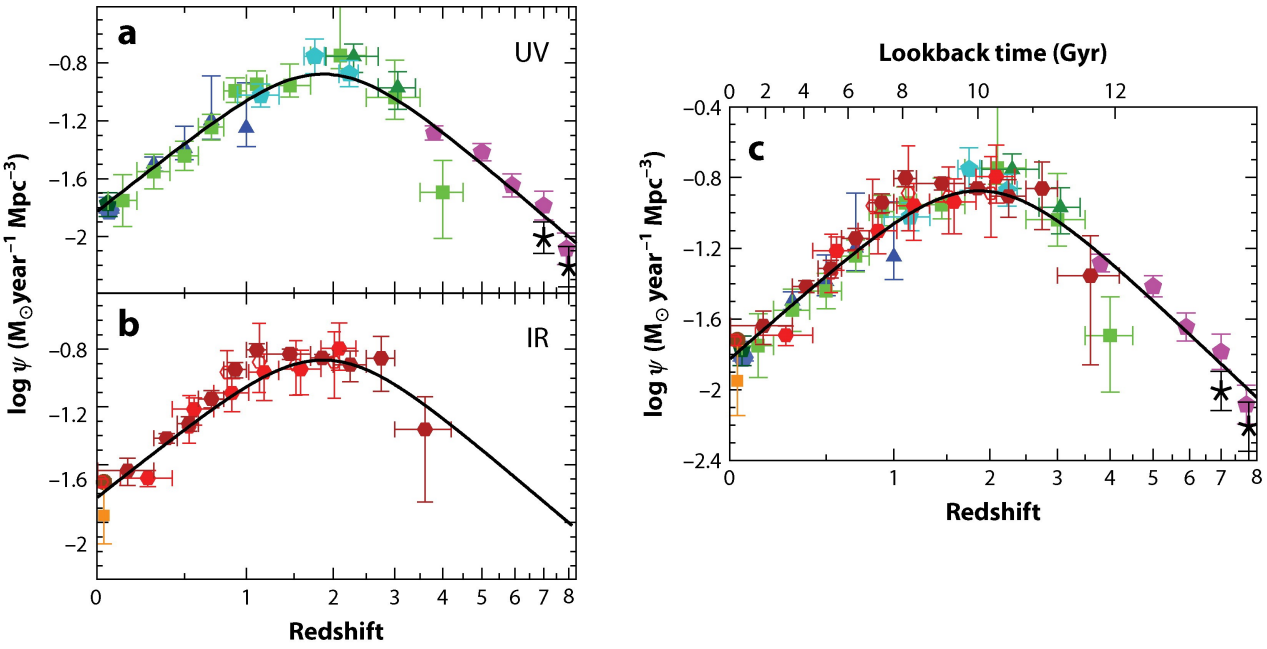


Figure 2.2: ***Right****: the best fit of SFRD as a function of redshift is represented by a black line. The colored points denote observations from both UV and FIR surveys used in the analysis.* ***Left****: The same fit, but separates the observations into UV (upper panel) and FIR (lower panel). Figures sourced from Madau & Dickinson (2014).*

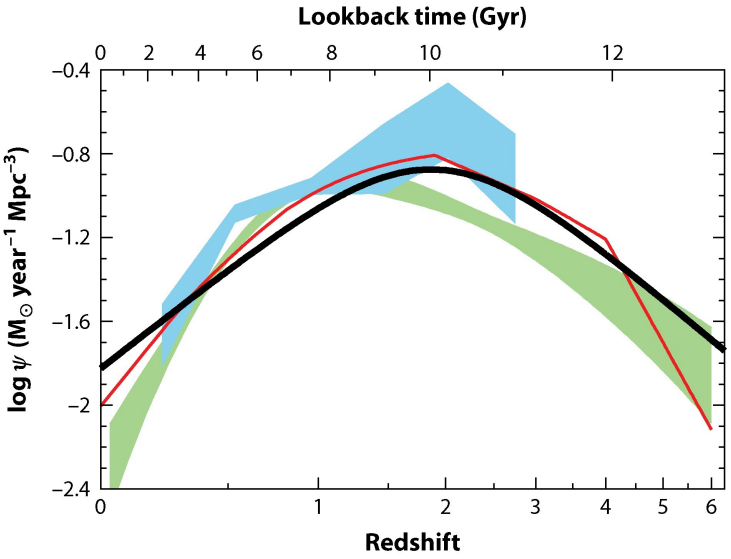


Figure 2.3: *Comparison between the SFRD and the evolution of SMBHs as a function of redshift. The Best fit of the SFRD is shown as a black line, and it is compared with the accretion history of Supermassive Black Holes (SMBHs). The blue band represents infrared observations (Delvecchio et al., 2014). The red line (Shankar et al., 2009) and the green band (Aird et al., 2010) represent observations in the X-ray band. Figure sourced from Madau & Dickinson (2014)*

2.2.1 SFRD Surveys limitations

SFRD between $0 < z < 1$ primarily suffers from the constraint that current measurements have been conducted over relatively small fields, covering only minor portions of the sky and small comoving volumes. Consequently, these measurements may be subject to cosmic variance[3] bias. Even fields like the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007), which extend over 2 square degrees, span less than 100 Mpc at $z < 1$ and may not be sufficient to enclose a large-scale cosmic structure. Furthermore, for such studies, precise distance measurements are crucial. However, this level of accuracy is often achievable only through spectroscopy, which typically can only be performed on small samples of galaxies at a time. This limitation reduces the ability to obtain such measurements for an entire field.

At $z > 1$, deep surveys are necessary to identify sources at lower luminosities and thus capture most of the star formation. In the UV, rest-frame emission at 1500 Å is easily detectable by ground-based telescopes up to $z = 2$, though it's important to remember that UV emission is biased by dust absorption. Direct measurements of dust emission, which peak in FIR or sub-mm bands depending on the redshift, are constrained to a few fields (like GOODS[4] and COSMOS) and, as explained in Madau & Dickinson (2014), often fail to reach the depth required to probe low luminosities at $z > 2$. Instruments like ALMA achieve great sensitivity and can observe dust continuum at high redshifts, but they can only cover small portions of the sky at a time.

2.2.2 State of the Art and Open Question

There are many unresolved questions regarding the efficiency of star formation in the early Universe (e.g., Behroozi et al. 2013; Stefanon et al. 2017) and the overall timeline for the buildup of stellar masses in galaxies.

At $z>3$, optical and UV rest-frame emissions are redshifted to Near-Infrared (NIR) bands. As a result, our understanding of the early Universe in optical/UV bands is limited to those star-forming galaxies selected through features such as the Lyman Break (LB). Such method proved effective in identifying high redshift Lyman Break Galaxies (LBGs) (Steidel et al. 1998; Bouwens et al. 2021, 2023a) but, most likely, leads to an incomplete census of galaxy populations, inevitably missing UV-faint galaxies such as obscured or quiescent galaxies (Wang et al. 2019; Valentino et al. 2020; Talia et al. 2021).

In the last decades, instruments like the Submillimetre Common-User Bolometer Array (SCUBA), and later Spitzer (Werner et al., 2004), Hershel (Pilbratt et al., 2010), and ALMA (Wootten & Thompson, 2009) have uncovered a population of galaxies that remained hidden in optical/UV surveys but are relatively bright in IR or sub-mm bands. (e.g. Smail

[3]Cosmic variance is the statistical fluctuation in observations of cosmic structures across different regions of the Universe, due to the uneven distribution of matter.

[4]Great Observatories Origins Deep Survey, Dickinson et al. 2003

et al. 1997; Williams et al. 2019; Gruppioni et al. 2020; Wang et al. 2019, Smail et al. 2021; Xiao et al. 2023).

These galaxies have very red Spectral Energy Distributions (SEDs) and remain undetected even in the deepest Hubble Space Telescope (HST) surveys, hence the nik-name Hubble-dark Galaxies.

Thanks to its exceptional sensitivity, JWST can observe the optical/UV emission of dust-obscured galaxies at z>3 and beyond. Optical/UV studies enhanced by JWST's investigations have extended our capacity to measure the SFRD up to z $\approx$ 7-8, with research efforts pushing this boundary to z $>$ 10 (Harikane et al. 2022; Adams et al. 2023; Harikane et al. 2024; Barrufet et al. 2023; Donnan et al. 2023) enabling a deeper understanding of the evolution of star formation up to the epoch of reionization, suggesting the presence of a considerable build-up of dust-mass already in place at that cosmic times. Despite these advancements, these same investigations may be biased by the observing band (i.e., the rest-frame UV) which is highly affected by dust obscuration.

Studies such as Bouwens et al. (2023b) have examined LBGs with JWST up z $\approx$ 15 (see Fig 2.4). The SFRD values derived from their work are in agreement with previous estimates (Bouwens et al. 2015 and Oesch et al. 2018) but show higher values at z $>$ 12 than those of different works (Donnan et al. 2023; Harikane et al. 2023, and Finkelstein et al. 2022) highlighting the fact that a consensus on this topic is yet be reached.

The scientific landscape in recent years has highlighted the emergence of two galaxy populations whose contribution to the cosmic evolution of galaxies as well as of the SFRD is still poorly constrained:

The first consists of massive (M $\approx 10^{11}$ $M_{\odot}$) passive galaxies (sSFR $< 10^{-11}$ yr^{-1}) that are extremely compact at z $>$ 2 (now spettroscopically confirmed up to z $\approx$ 4.5 ,Carnall et al. 2023). The number density of these galaxies challenges our existing cosmological models and theories on galaxy evolution (see Cimatti et al. 2008; Straatman et al. 2014) with simulations struggling to predict observed number densities, high masses, stellar densities, and low levels of star formation (Cecchi et al. 2019; Valentino et al. 2020). Their high stellar density and quiescent nature at such high redshifts may result from a starburst event, likely triggered by the rapid collapse of a large amount of gas at z $>$ 3-4. This hypothesis aligns with findings from galactic archaeology in the local Universe, which suggest that massive and passive galaxies formed over timescales shorter than 1 Gyr (Thomas et al., 2005), under the assumption that this evolutionary path may be still valid at earlier cosmic times. However, identifying progenitors for this quiescent galaxy population is challenging when directly comparing them to typical Star-Forming Galaxies at z $\approx$ 4; as the cosmic SFRD derived from optical/UV studies at that epoch was at least an order of magnitude lower than during the cosmic noon, and, the observed number density of optical/UV selected galaxies at these redshifts, mainly Lyman Break Galaxies, is found to be one or two orders of magnitude lower than that of massive and passive galaxies at z $\approx$ 3 (Toft et al. 2014; Marrone et al. 2018; Valentino et al. 2020).

The second emerging galaxy population that has recently garnered significant attention is

that of the Submillimeter Galaxies (SMGs), consisting mainly of massive galaxies with very high SFRs (SFRs $>$ 100 $M_\odot$ yr^{-1}, see Swinbank et al. 2014; Gruppioni et al. 2020); for some of these galaxies spectroscopic redshift of $z > 4$ have been confirmed (Daddi et al. 2009; Jin et al. 2022; Gentile et al. 2024).

SMGs are among the best candidates for the progenitors of passive galaxies at $z \approx 3$ thanks to their sizes, masses, and number densities (Tacconi et al. 2006; Toft et al. 2014; Valentino et al. 2020). Still, the definition of SMGs is based on observational criteria, possibly hiding an inherent diversity within the population it describes. Nonetheless, this classification overlaps with the physical definition of Dusty Star-Forming Galaxies (DSFGs) at the highest redshifts, capturing suitable candidates to be progenitors of massive quiescent galaxies that, at the same time, may play an important role in shaping the evolution of the SFRD.

To better contextualize concepts like SEDs and offer a comparative perspective, Fig 2.5 shows a comparison between the SED of a DSFG at $z = 2$, with an infrared luminosity (LIR) of $3x10^{12}$ $L_\odot$, as reported by Pope et al. (2008), with the templates of two local Ultra-Luminous Infrared Galaxies (ULIRGs).

Recent studies highlight the presence of a population of DSFGs so obscured to remain undetected even in the deepest current optical/NIR surveys (Wang et al. 2019; Talia et al. 2021; Enia et al. 2022; Gentile et al. 2023; Behiri et al. 2023; Barrufet et al. 2023). Determining the contribution of these elusive galaxies to the cosmic SFRD as well as exploring their potential as progenitors for passive galaxies at $z \approx$ 2-3, remains a subject of ongoing research.

There is no single method for identifying these galaxies: leveraging SEDs features as the Lyman Break can be done only with current state-of-the-art instruments like JWST (Bouwens et al., 2023b), as this feature is often too dimmed by dust obscuration. Studies like Wang et al. (2019) identify such galaxies in FIR bands as sources lacking an opt/UV counterpart. Other studies such as Gruppioni et al. (2020) have serendipitously identified these galaxies in ALMA pointings (Fig 2.4), while others like Talia et al. (2021), Gentile et al. (2023), Behiri et al. (2023) and Enia et al. (2022) have demonstrated that selection methods based on radio emission, which is free from dust bias and readily applicable to wide surveys, are effective in identifying extremely obscured DSFGs.

Theoretical studies, focusing on modern hydrodynamic simulations to comprehend the role of obscured star formation, and its evolution across different cosmic epochs, suggest that obscured star formation might even dominate the SFR budget at $z > 4$ (Zimmerman et al., 2024), once again confirming the importance of understanding the role of dust-obscured galaxies in cosmic evolution.

Nevertheless, it is still possible that the early cosmic star formation and stellar mass densities are dominated by less extreme, DSFGs or quiescent galaxies, which have been missing from rest-frame UV data sets. To confirm these possibilities new studies using dust-unbiased tracers such as radio observations, along with the deep and sensitive observations of the James Webb Space Telescope (JWST), are crucial.

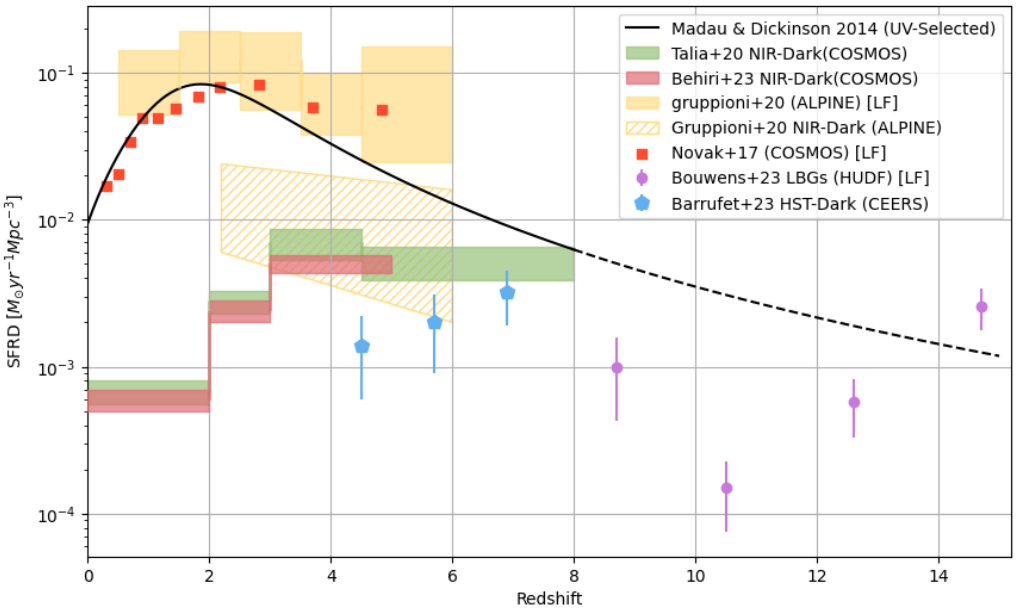


Figure 2.4: *Illustration of the evolution of cosmic SFRD and its various contributions. Results marked in the legend with 'LF' mean that are obtained through the construction of a luminosity function. In purple dots, the results from Bouwens et al. (2023b) for their sample of LBGs selected up to z=15 via JWST observations. In blue, observations from Barrufet et al. (2023) of a sample of H-Dark galaxies observed with JWST. In green and red, respectively, the contributions from the populations of radio-selected NIR-Dark DSFGs identified by Talia et al. (2021) and Behiri et al. (2023) in the COSMOS field. In yellow, the data from Gruppioni et al. (2020) from the ALPINE[5] survey. In red squares, the findings reported by Novak et al. (2017) for their sample of star-forming galaxies selected at 3 GHz in the COSMOS field. In black the Madau & Dickinson (2014) relation. The dotted part of the line indicates an extrapolation of the relation.*

[5] ALMA Large Program to Investigate C+ at Early Times (ALPINE)

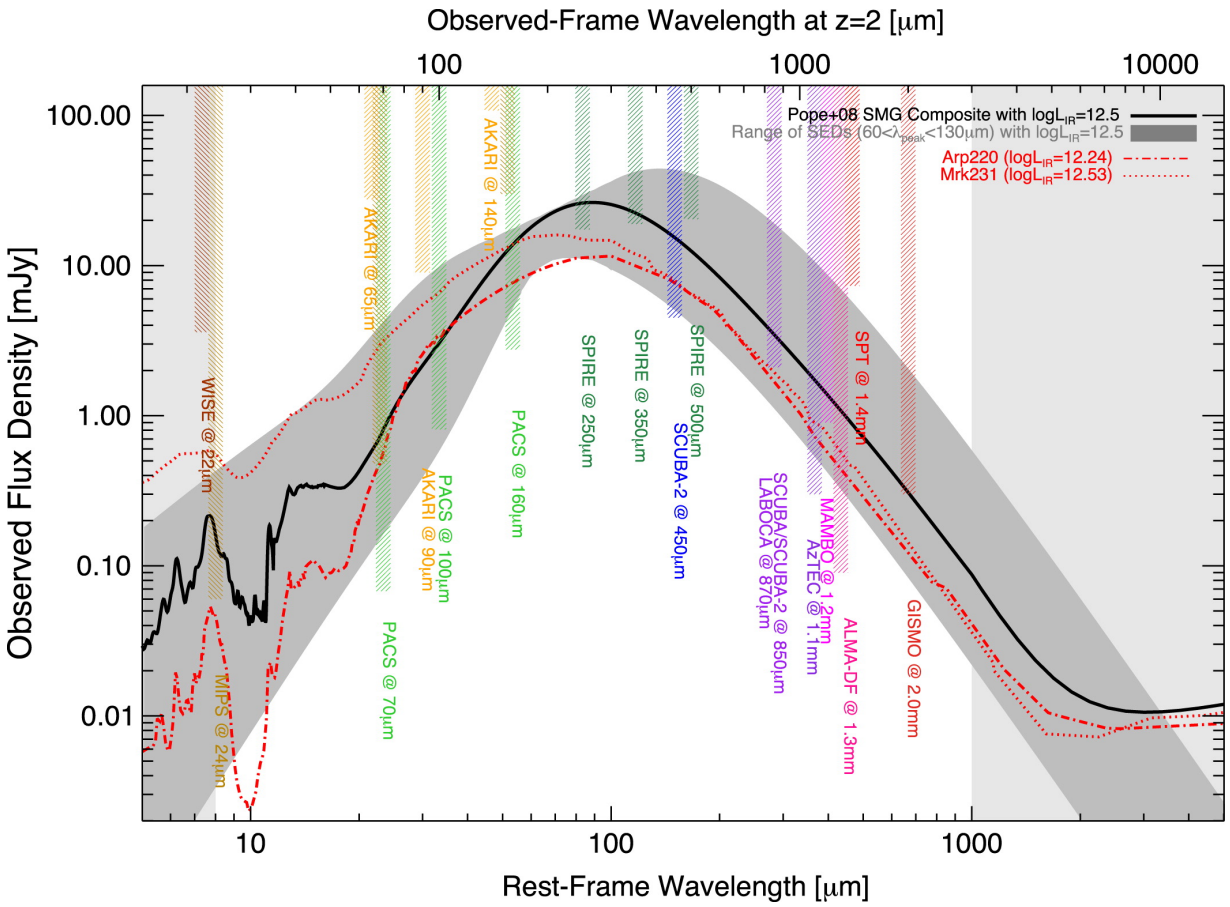


Figure 2.5: *Composit SED of a SMGs at $z = 2$ with a star formation rate of approximately 500 $M_\odot/yr$ (represented in black, Pope et al. (2008)). This is superimposed on a gray area indicating the potential variability in SED profiles for galaxies with a constant infrared luminosity of 10^{12} $L_\odot$, with peak SED wavelengths spanning 70–130 μm (equivalent to temperatures of about 30–58K). The diagram further includes vertical shaded areas that denote the sensitivity ranges of various far-infrared and sub-mm observatories, with the lowest flux density value of each colored area reflecting the 3σ detection threshold for the respective instrument. The two red dotted lines represent the two templates from the two local ULIRGs Arp 220 and Mrk 231. Image taken from Casey et al. (2014)*

2.3 Motivations and Aims of the Thesis

This thesis places itself within the previously outlined scientific context, aiming at understanding and identifying DSFGs within the J1030 field: an extragalactic field of $\approx 30\times30$ arcmin^2 that has been observed from UV to sub/mm by the teams at INAF[6]/OAS[7] (Gilli et al., 2019). The main goal of this work is the identification of these DSFGs through a radio selection-based approach, drawing inspiration from studies such as Talia et al. (2021) and subsequent works like Behiri et al. (2023) and Gentile et al. (2023). The final aim of this work is to assess the contribution of such galaxies to the cosmic SFRD, placing them within the broader context of galactic formation, and evaluating their role as potential progenitors of the massive quiescent galaxies already mentioned.
In particular, Gentile et al. (2023) developed a specific pipeline, PhoEBO (Photometric Extractor for Blended Objects), dedicated to the deblending of NIR-Dark galaxies in the MIR Spitzer/IRAC bands. This study aims to extend the applicability of this pipeline, adapting it to operate effectively also in the FIR bands, as well as in sky areas not previously analyzed.
The deblending of FIR/sub-mm maps of the J1030 field is done via an integrated approach that combines PhoEBO with methods derived from Jin et al. (2018), allowing for precise photometric measurements in these bands as well.
The characterization of the identified galaxies is performed through SED fitting using the MAGPHYS+PHOTO-Z software (Battisti et al., 2019).
To mitigate the risk of contamination of the selected galaxy sample by AGNs, several measures have been adopted, including cross-checking with X-band data, verifying potential radio excesses relative to infrared luminosity, and morphological analysis at 1.4 GHz. These precautions are fundamental to ensuring the reliability and validity of the obtained results.
This thesis is organized as follows: In Chapter 3 I discuss The data used. In Chapter 4 I describe the methodology for extracting accurate flux measurements when using PhoEBO and the deblending technique of sub/mm SPIRE bands. In Chapter 5 I present the results, discuss them, and see how they stack up against earlier studies before concluding. Lastly, Chapter 6 provides a recap of all the work done in this thesis and looks ahead to future directions.
Trhuoght this thesis I used AB magnitudes (Oke & Gunn, 1983), employed a Chabrier (2003) and assumed a ΛCDM cosmology with $H_0 = 67.66$ [km s^{-1} Mpc^{-1}] and (Ω_{tot}, Ω_Λ, Ω_m) = (1, 0.7, 0.3). (Planck Collaboration et al., 2020)

[6]Istituto Nazionale di AstroFisica
[7]Osservatorio di Astrofisica e Scienza dello spazio

Chapter 3

Data Description

This chapter provides an overview of the data utilized in this thesis: it consists primarily of a set of images across various bands at different wavelengths (see Tab 3.3), as spectroscopic data are not included.
Two photometric catalogs enhance the dataset, the first of which includes a range of observations across the U, B, V, R, Y, J, H, Ks, IRAC1, and IRAC2 bands, is comprehensive of 14730 sources and built upon priors in the Ks band[1], henceforth it will be referred to as *Ks-selected full-band catalog*. The second catalog is represented by that associated with the Karl G. Jansky Very Large Array (JVLA) data derived from D'Amato et al. (2022) (Sec 3.1.2).

3.1 The J1030 field

The study around SDSS J1030+0524[2], a QSO at a redshift of z=6.31 (Fig. 3.1) with an estimated mass of $M \approx 1 \times 10^9\ M_\odot$, was born in 2012 as part of a LBT[3] follow-up campaign for high-redshift quasars (Morselli et al., 2014). By 2023, the J1030 team had accumulated a wealth of data from a variety of instruments, including the JVLA, JWST, ALMA, and the Chandra X-ray Observatory (which conducted the deepest X-ray observation of a z>6 quasar to date) and can exhibit over 50 scientific publications. SDSS J1030+0524 (hereafter referred to as J1030) is notably important because it may offer the most convincing evidence yet of a large-scale structure surrounding a SMBH that formed more than 12 billion years ago. The extensive multi-wavelength observational coverage of this area has uncovered an overdensity of galaxies sharing the same redshift as the quasar, spreading across several megaparsecs in diameter. Therefore, the J1030 field serves as an exceptional setting for exploring the development of distant black holes and galaxies.

[1] Priors in Ks map are selected with an absolute threshold of $\approx 0.36\ \mu$Jy
[2] J1030 Web Page
[3] Large Binocular Telescope (LBT; Hill et al. 2012)

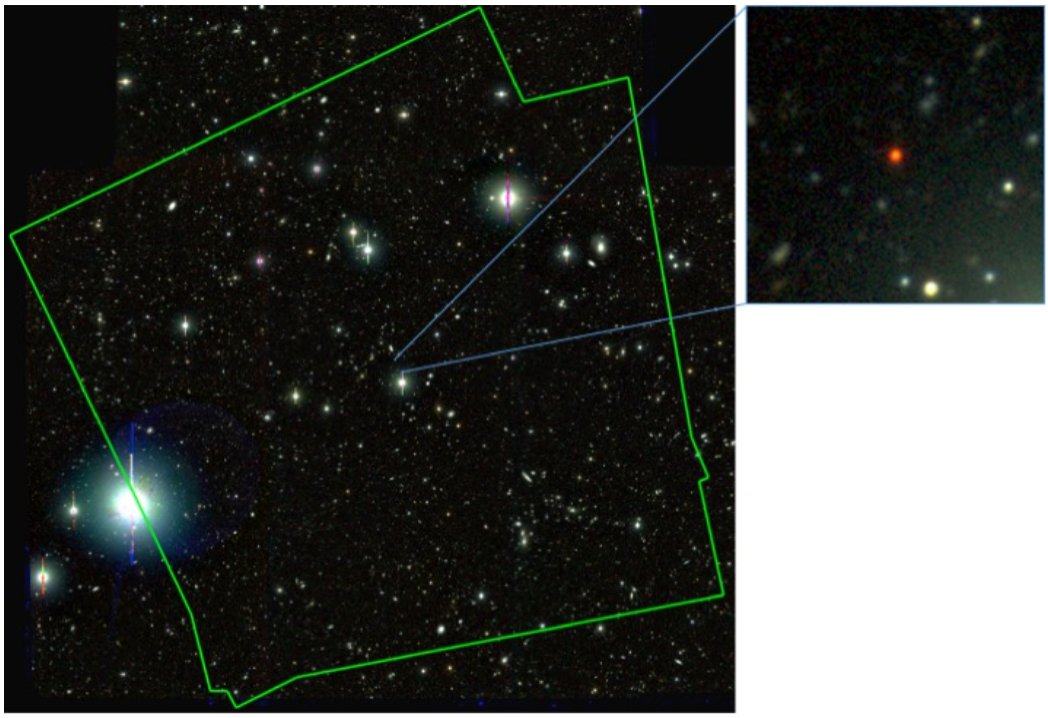


Figure 3.1: *LBT color image of the field around SDSS J1030+0524. The image is 25×25 arcmin^2, i.e. 8×8 Mpc^2 at z=6.31. The region covered by deep Chandra data is shown in green. The zoom shows the quasar at the center of the field. Image taken from J1030 webpage.*

3.1.1 Optical and Near-Infrared Data

As part of the Multiwavelength Survey by Yale-Chile (MUSYC; Gawiser et al. 2006) survey, the J1030 field underwent extensive deep optical observations (UBVRIz; Blanc et al. 2008) covering an area of 30' x 30' using the Cerro Tololo Inter-American Observatory (CTIO). The resulting scientific images constitute the optical data used in this study. The NIR data were collected from two complementary campaigns/observations: the MUSYC deep NIR survey (J and H bands, Quadri et al. 2007) and a series of WIRCAM/CFHT [4] deep observations (Y, J and Ks band,Balmaverde et al. 2017). The former was specifically targeted at a 10'x10' region at the center of the field, achieving a depth of 23 mag at a 5σ level. By contrast, the WIRCAM observations cover a larger area of 25'x25' with a depth reaching up to 24 mag at a 5σ level.

[4] Wide-field InfraRed Camera (WIRCAM), Canada-France-Hawaii-Telescope (CFHT)

3.1.2 Radio Data

The radio data employed in my analysis are sourced from D'Amato et al. (2022), in which they conducted observations on the central quasar of the field using the JVLA instrument in the L band ($\nu = 1.4$ GHz). A single pointing, with a Full Width at Half Maximum (FWHM) of 30', covers the entire field. Notably, the observation reaches its maximum depth at the field's center, yielding a 5σ sensitivity of ~ 12.5 μJy, thus making it one of the deepest radio observations so far at 1.4GHz (Fig 3.2 for reference).
Together with the scientific images, D'Amato et al. (2022) also provided a catalog of detected sources comprehensive of 1489 entries detected at Signal to Noise ratio (S/N) > 5.

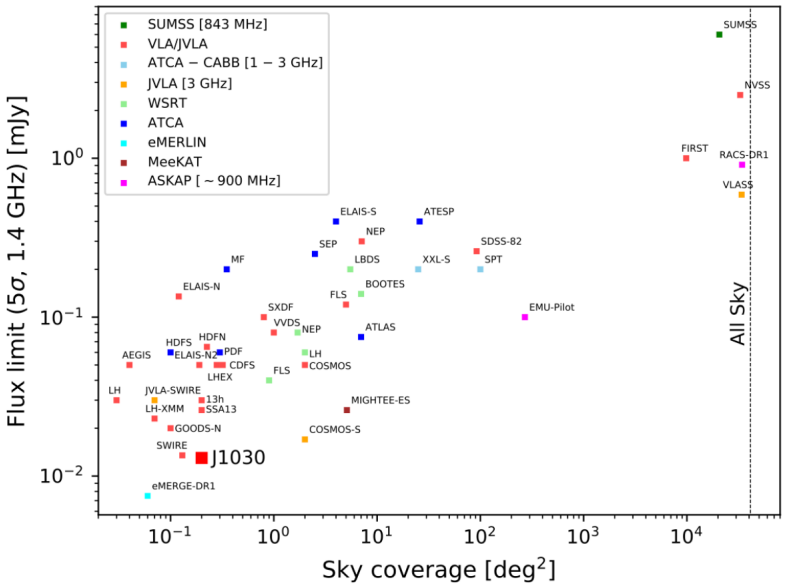


Figure 3.2: *Sky coverage vs. 5σ sensitivity at 1.4 GHz for a collection of available radio surveys to date including J1030. Image taken from D'Amato et al. (2022).*

3.1.3 Mid-Infrared Data

The data from the Spitzer Space Telescope, particularly from the InfraRed Array Camera (IRAC) and MIPS (Multi-Band Imaging Photometer for Spitzer) instruments, represent the primary resources available for J1030 in the MIR bands.

- **IRAC ch1 and ch2**: Spitzer IRAC channel 1 (3.6 μm) and 2 (4.5 μm)data used in this study derive from Annunziatella et al. (2018). Their publication includes a set of reduced images for all IRAC ch1 and ch2 data associated with the MUSYCS1030 field rearranged into two final mosaics of $\sim$ 35' x 35' that cover the entire field.
  It's important to highlight a difference from the native IRAC mosaic pixel scale, which traditionally refers to a dimension of 0.6 arcseconds per pixel, as the final IRAC data and coverage maps have been resampled to a smaller pixel scale value of 0.267" per

pixel. Both mosaics reach their maximum depth at the center, with a 5σ level of 22.5 mag. Together with the scientific map, a photometric catalog of the identified sources was provided, which was subsequently integrated into the J1030 Ks-selected full band photometric catalog (Sec 3) using a cross-match procedure based on coordinates, where possible, to link an IRAC flux to the corresponding Ks-selected prior. Following this procedure, for 9562 out of the 14730 sources in the Ks-selected catalog, a flux in either IRAC1 or IRAC2 has been successfully associated. By definition, the target sources investigated in this study are not present within this catalog.

- **IRAC ch3 and ch4**: The scientific mosaics of IRAC channel 3 and IRAC channel 4 are taken from the IRSA archives[5] and cover an area of $\sim$ 19'x19', thus smaller than that for ch1 and ch2, with a maximum depth of 20.5 AB mag at 5σ. Maps are provided in units of Mjy/str, however, unlike those of IRAC 1 and 2, they are sampled at the native resolution of an IRAC mosaic, with a pixel scale of 0.6" per pixel.

- **MIPS Data**: The 24 μm Spitzer data from the MIPS instrument were sourced from the IRSA archives. These consist of a series of observations that cover an area of approximately 10' x 16' and are centered on the QSO at the center of the field. They reach a limiting magnitude of 19.5 at 5σ and are characterized by a pixel angular scale of 2.45' per pixel.

3.1.4 Far-Infrared Data

J1030 has been observed in FIR and sub-mm bands by the Photoconductor Array Camera and Spectrometer (PACS) and Spectral and Photometric Imaging Receiver (SPIRE) instruments onboard the Hershel Space Observatory. As for the Spitzer data, PACS images at 100 μm (green channel) and 160 μm (red channel) were obtained from the IRSA archives (further details available in Leipski et al. 2014). These consist of two maps that cover the same area of approximately 4'×7' in size and have a maximum depth of about 10 mJy at a 5σ level.
Similarly, the Herschel data from the SPIRE instrument at 250, 350, and 500 μm were obtained from the IRSA archives and cover an area of approximately 16'×16' (see also Leipski et al. 2014). Each of the three observations has a similar maximum depth of around 30 mJy at 5σ.
The characteristic FWHM of the observations are reported in Table 3.1.

3.1.5 X-ray Data

The data between 0.5-7 KeV are from the X-ray catalog presented in Nanni et al. (2020), where they conducted an observation of the central region, over a total area of 335 arcmin^2. From the catalog, the extraction regions were derived and then used in my work. Notably,

[5] IRSA Archives

the observation associated with the extraction regions has an average depth of 1.7x10^{-16} erg/cm^2/s obtained over an integration time of 479 ks, effectively making J1030 field the fifth deepest extragalactic X-ray survey to date.

Table 3.1: *Summary of key specifications for NIR to RADIO data including central wavelengths, pixel scales, geometric FWHM in arcseconds, and Beam Areas (SPIRE beam areas are those reported in SPIRE Hershel Handbook[6]). Optical and NIR data before the Ks band are not included.*

Band	Central λ	Pixel Scale	FWHM (*arcsec*)	Beam Area ($arcesc^2$)
NIR Ks	2.2 μm	0.304	0.45	-
IRAC ch1	3.6 μm	0.267	1.66	-
IRAC ch2	4.5 μm	0.267	1.72	-
IRAC ch3	5.8 μm	0.6	1.88	-
IRAC ch4	8.1 μm	0.6	1.98	-
MIPS 24	24 μm	2.45	6.1	-
PACS 100	100 μm	1.6	6.7	-
PACS 160	160 μm	3.2	12	-
SPIRE250	250 μm	6.0	17.9	469.35
SPIRE350	350 μm	10.0	24.2	831.27
SPIRE500	500 μm	14.0	35.4	1804.31
RADIO 1.4 GHz	21 *cm*	0.3	1.49/1.15	21.52

3.2 Radio Ks-Undetected Initial Catalog

The initial catalog of radio-selected sources was assembled through a cross-matching procedure between the radio catalog (D'Amato et al., 2022) and the *Ks-selected full-band catalog* using a cross-match radius of 1". This process resulted in the identification of 79 radio sources that do not have counterparts in the Ks catalog, these will be referred to as *target sources* henceforth. A similar selection criterion has been previously employed successfully in similar studies (Talia et al. 2021, Gentile et al. 2023, Enia et al. 2022, Behiri et al. 2023) and it has proven to be effective for identifying highly obscured, massive, and star-forming galaxies.

The positions of the *target sources* and wavelength coverage are illustrated in Figure 3.3.

[6] SPIRE Handbook

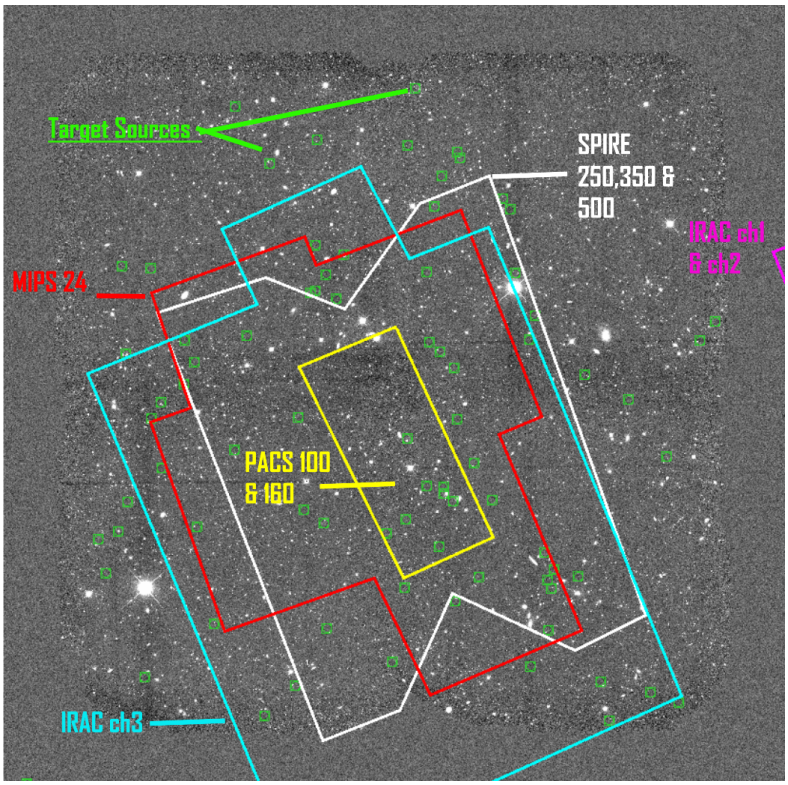


Figure 3.3: *Image of J1030 at 2.2 μm (Ks-band) with the 'target sources' marked by green squares. The coverage of different instruments is delineated in multiple colors: MIPS 24 (red), PACS 100 & 160 (yellow), SPIRE 250, 350, 500 (white), and IRAC channels 1 & 2 (violet) and channel 3 (cyan)*

3.3 Data Background Subctraction

The preparation process for the IRAC ch3, ch4 and MIPS 24 μm images is an important step and worth a deeper discussion.

The previous reduction process for these maps did not effectively remove the background, requiring additional work to make them suitable for my analysis. As indicated in Table 3.2, the median of the background for the three maps deviates from the expected value of 0 (in units of the maps, see also Fig 4.7a; 4.7b) for a background-subtracted image. Furthermore, upon closer examination of the scientific map of IRAC ch4, it becomes apparent that the background is not uniform throughout the entire image but instead varies with position. This immediately rules out a simple sigma-clip subtraction method, where sources are first identified and masked, and the background is obtained by averaging the remaining pixels. While this approach is fast, it fails to account for potential background variations across the map, as it subtracts a constant value assumed to be valid for the entire image. Instead, I used a background interpolation technique provided by the photutils[7] library in Astropy. This approach generates a 2D background map by interpolating between user-defined box

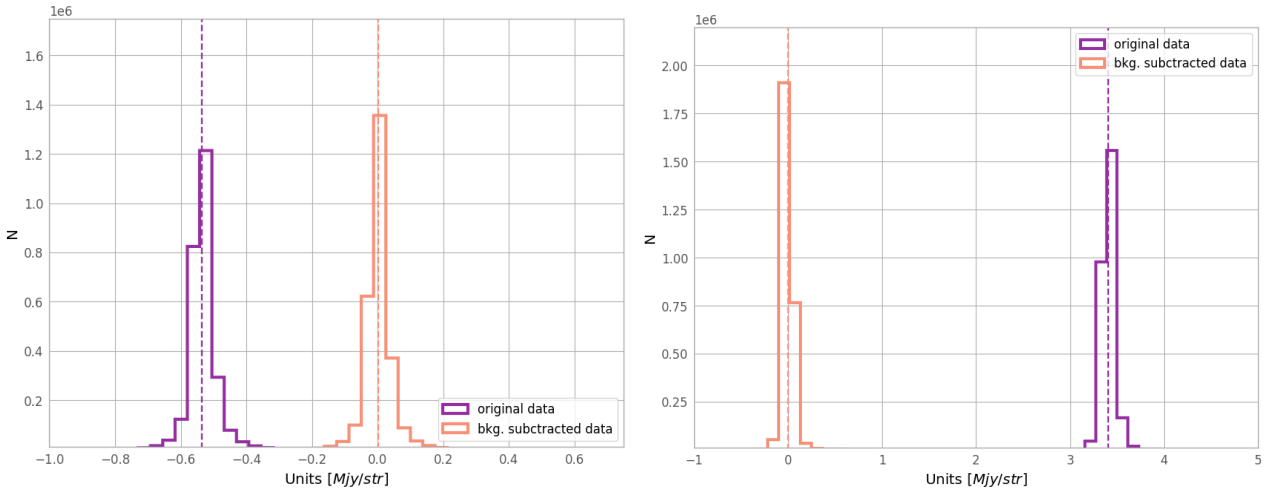
[7] Photutils Astropy affiliated package

sizes to preserve the total flux. The size of these boxes should be chosen such that the amount of background -i.e., the number of unmasked pixels- is statistically significant but not excessively large, ensuring that the number of boxes is sufficient to describe the entire map.

For the IRAC scientific maps, I selected a box size of 30 by 30 pixels while, for the MIPS map, the grid size was fixed to 25 by 25 pixels. For each of the boxes into which the maps are partitioned, the background is computed using an implementation of the method used in SourceExtractor (Bertin & Arnouts, 1996); the background within each mesh is determined as $(2.5 \cdot median) - (1.5 \cdot mean)$, however, if the ratio of (*mean* - *median*) to the standard deviation (*std*) is greater than 0.3, then the *median* value is employed instead. The final 2D map of the background obtained at the end of the process is then subtracted, the results of the procedure are reported in Tab 3.2, Fig 3.5 and 3.4.

Table 3.2: *Table of median map values for IRAC ch3, ch4 and MIPS 24 μm before (left column) and after (right column) the background subtraction procedure.*

Band	Background-Median [Mjy/sr]	Bkg-Median (Subtracted) [MJy/sr]
IRAC3	-0.544	0.003
IRAC4	3.427	0.002
MIPS24	4.473	0.001



(a) ***(b)***

Figure 3.4: ***Left*** *Comparative Analysis of Original and Background-Subtracted for IRAC ch3. The histogram illustrates the intensity distribution measured in units of the map, with the original data represented by the purple bars, while the orange bars depict the data after background subtraction, with the median value (dashed line) close to zero, indicative of an effective reduction of the background noise.* ***Right*** *the same procedure of background subtraction applied to IRAC ch4.*

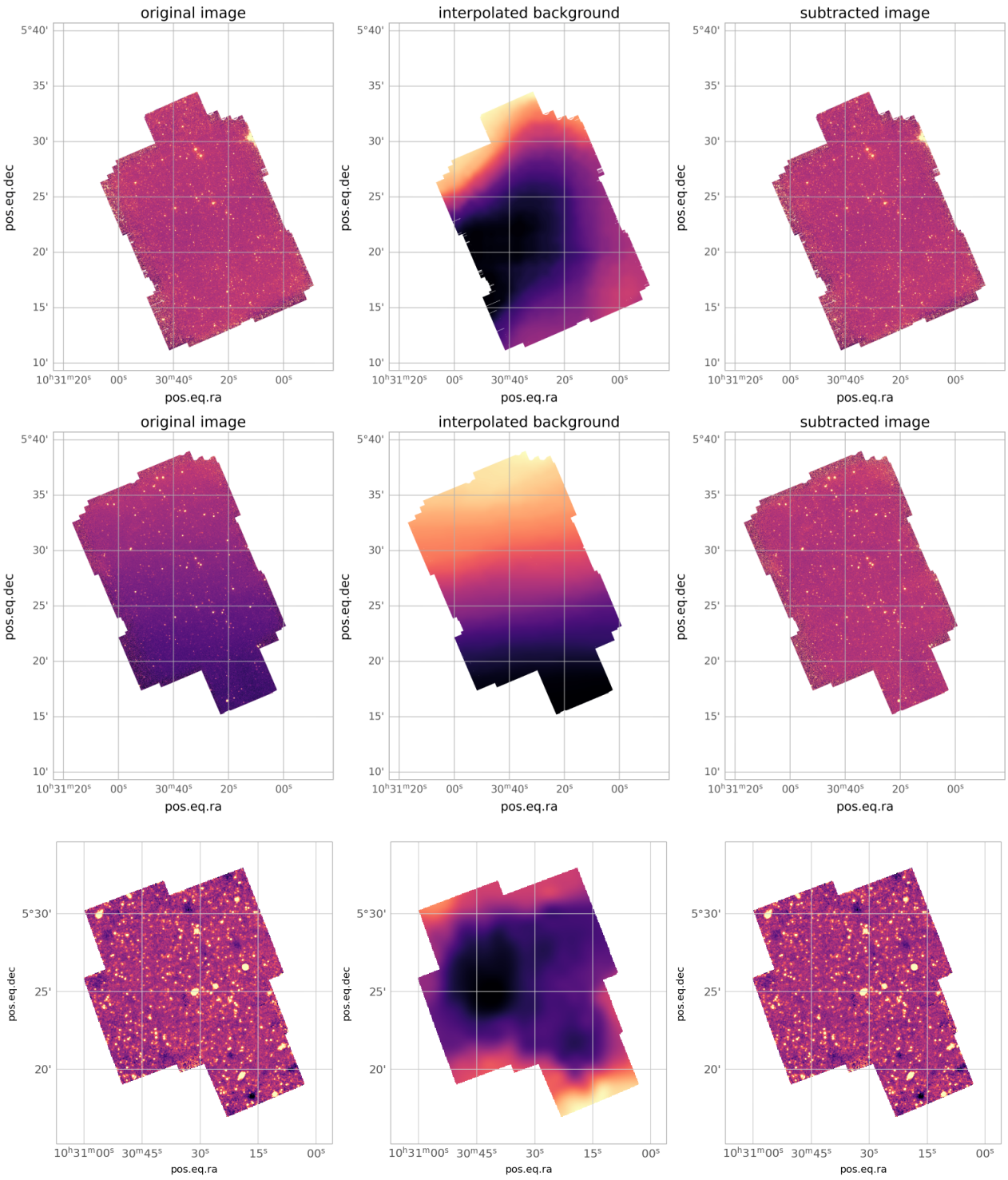


Figure 3.5: ***Top****: Sequential Stages of Background Subtraction for IRAC ch3. The left panel displays the original image, the center panel presents the interpolated background model, and the right panel shows the resulting image after the background model has been subtracted from the original data.* ***Center , Bottom****: The same but for IRAC ch4 and MIPS24.*

3.4 Comparison with COSMOS Data

This study is closely related to the research previously conducted in the COSMOS field by Talia et al. (2021) and subsequent works of Behiri et al. (2023) and Gentile et al. (2023). Consequently, this chapter presents a quick comparison between data used in my work and those employed in Gentile et al. (2023), and Behiri et al. (2023), which are the latest in the series, working with data from the 2020 COSMOS catalog (Weaver et al., 2022). In particular, this thesis draws its methodology from Gentile et al. (2023) in which the optical to MIR photometry was performed ex-novo using the first implementation of the same pipeline used in this thesis. Among the main differences, it is noteworthy that the J1030 maps are generally more shallow, especially in the optical and NIR bands where the difference can be more than a factor of 5 for the limit flux at 1σ. Conversely, the 1.4 GHz observations, are more shallow in COSMOS. However, it should also be considered that COSMOS JVLA data exhibits a more uniform average depth, essentially constant across the entire field, while for J1030, the maximum depth is achieved at the center of the field.

Table 3.3: *Flux Limits at 1σ for COSMOS and J1030 fields, COSMOS data depths from optical to IRAC 4 are sourced from Weaver et al. (2022) while from MIPS to SPIRE from Jin et al. (2018).*
** - The first value is the flux at 3GHz while that in parenthesis is the equivalent rescaled at 1.4 GHz assuming a radio spectral index of -0.7.*

BAND	COSMOS	J1030	Reference J1030
U band	0.015 uJy	0.043 uJy	Blanc et al. (2008)
B band	0.015 uJy	0.074 uJy	Blanc et al. (2008)
V band	0.027 uJy	0.031 uJy	Blanc et al. (2008)
R band	0.017 uJy	0.084 uJy	Blanc et al. (2008)
Y band	0.028 uJy	0.182 uJy	Blanc et al. (2008)
J band	0.033 uJy	0.182 uJy	Blanc et al. (2008)
H band	0.043 uJy	0.450 uJy	Blanc et al. (2008)
Ks band	0.063 uJy	0.182 uJy	Gilli et al. (2019)
Spitzer/IRAC 1	0.033 uJy	0.726 uJy	Annunziatella et al. (2018)
Spitzer/IRAC 2	0.036 uJy	0.726 uJy	Annunziatella et al. (2018)
Spitzer/IRAC 3	0.630 uJy	4.581 uJy	IRSA archives
Spitzer/IRAC 4	0.690 uJy	4.581 uJy	IRSA archives
Spitzer/MIPS 24 μm	10.22 uJy	11.51 uJy	IRSA archives
PACS 100 μm	1.440 mJy	2.00 mJy	Leipski et al. (2014)
PACS 160 μm	3.550 mJy	5.55 mJy	Leipski et al. (2014)
SPIRE 250 μm	1.770 mJy	$\sim$ 6.0 mJy	Leipski et al. (2014)
SPIRE 350 μm	2.680 mJy	$\sim$ 6.0 mJy	Leipski et al. (2014)
SPIRE 500 μm	2.910 mJy	$\sim$ 6.0 mJy	Leipski et al. (2014)
1.4 GHz	10.00 uJy	2.44 uJy	D'Amato et al. (2022)
3 GHz*	2.89 , (4.927) uJy	- , (2.44) uJy	D'Amato et al. (2022)

Chapter 4

Data Analysis

4.1 PhoEBO

Standard aperture photometry, while reliable for extracting fluxes of isolated sources, presents limitations when applied to closely separated or partially overlapping sources, often as a side-effect of the poor resolving capabilities of some instruments. In such cases, alternative methods may be necessary for accurate flux measurements. This becomes particularly problematic with low-resolution data, especially for observations at long wavelengths, where the image's PSF is large (even tens of arcseconds). As a result, sources that would otherwise be resolved in some bands appear merged (blended) in others (see Fig 4.1), leading to contamination of measured fluxes by secondary sources, and introducing potential inaccuracies when performing photometry. The procedure used to split the total flux and correctly assign it to the respective sources in cases where they appear partially overlapping is called 'deblending'. Standard deblending algorithms, such as the one implemented in *Source-Extractor* (Bertin & Arnouts, 1996), rely on minimal contrast between blended components in subsequent pixels to separate distinct elements, but this level of contrast is often unattainable in the MID to FIR bands, making them unsuitable for my analysis. To address this issue, advanced software tools like *Tractor* (Lang et al., 2016) and *The Farmer* (Weaver et al., 2019) use parametric profile-fitting techniques, leveraging high-resolution images as priors for the position and shape of sources. However, these techniques cannot be applied to NIR-Dark galaxies, which are only robustly detected in IRAC channels. This is the main reason why PhoEBO (Gentile et al., 2023) was developed. PhoEBO, short for "Photometry Extractor for Blended objects", is a Python-based pipeline designed to deblend MID to FIR images of NIR-dark galaxies. Firstly developed by (Gentile et al., 2023) it was initially employed to deblend IRAC sources in the COSMOS field. The version of PhoEBO employed in this study differs from the original implementation described in Gentile et al. (2023): this iteration has been specifically developed to extend the capabilities of the original code beyond its initial purposes. Such customization has enabled the use of PhoEBO for all the photometric analysis of sources in the J1030 field.

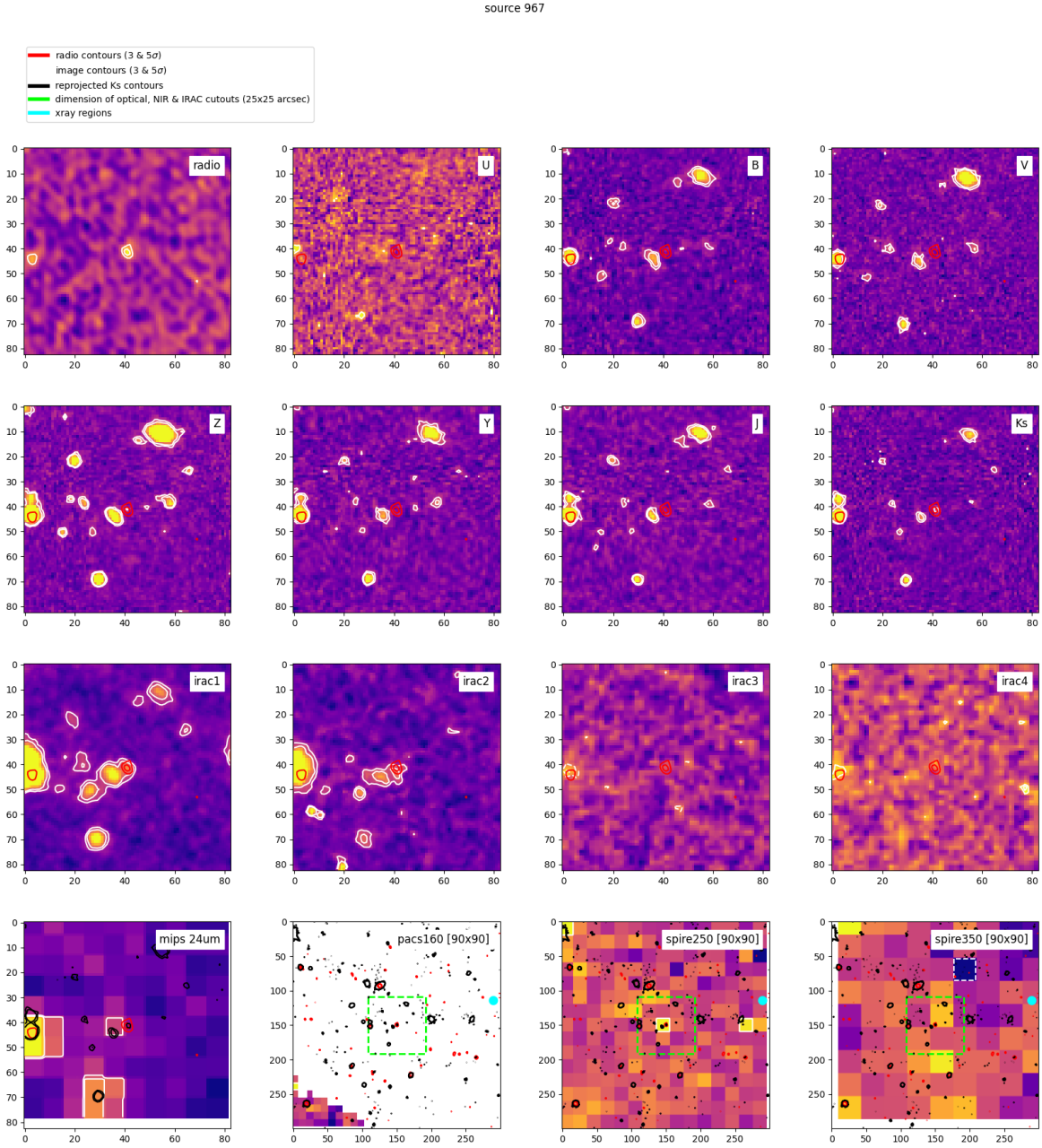


Figure 4.1: *Cutouts centered on source N 967 (radio catalog ID) in the bands from UV to 350 μm. For each cutout, contours at 3σ and 5σ are drawn in white, while radio contours at 3σ and 5σ are depicted in red. In the bands from 160 to 350 μm, cutouts sizes change as a function of the band, shorter wavelength bands have dimensions of 25'×25', and longer bands have sizes of 90'×90'. The green dashed box represents the size of the Ks cutout. Additionally, black contours for Ks at 3σ and 5σ are included from 24 μm to 350 μm. Blue circles represent Xray regions marking Xray sources*

4.1.1 PhoEBO general features

In this section, I will discuss the general workflow when using PhoEBO; for a more complete pipeline documentation, refer to the appendix A.
To ensure PhoEBO operates correctly, it's essential to include specific information about both potential contaminants and the 'target sources' beforehand. This information is stored in different scientific maps: as traced by my selection criteria (Sec 3.2), the 1.4 GHz radio map provides information on the target sources. Meanwhile, potential contaminants are identified in the Ks map. Finally, the map that requires deblending (such as the IRAC maps) shows both the target and the contaminants. Together, the radio and Ks images form the two *detection maps*.
Once this information is provided to PhoEBO, the analysis proceeds following these steps:

1. In the initial stage of the process, PhoEBO handles the loading and management of data for the target sources and contaminants. Subsequently, a cutout is generated around the target source, and the specific area for analysis is established.

2. In this phase, the Python implementation of the deblending algorithm present in *sExtracor*[1] (*SEP*[2]; Barbary 2016) is employed to detect sources within the two detection maps. To ensure the inclusion of all sources, an absolute threshold of 2σ and a minimum of 6 contiguous pixels above the threshold are used for each identified source. Simultaneously, the corresponding segmentation maps are generated. These segmentation maps comprise a collection of images associating each pixel with its respective source via a unique integer value (pixels not affiliated with any source are designated with the value 0). The segmentation map of contaminants is subsequently reorganized into a data cube, with the number of slices equal to the number of identified contaminants. Each slice contains the normalized segmentation map (also referred to as *binary mask*) of an individual isolated contaminant.

3. Each of the normalized segmentation maps is employed to generate a second data cube containing isolated contaminant sources. This is achieved through a straightforward multiplication procedure as pixels in the segmentation map associated with a source are assigned a value of 1, while all other pixels are set to 0. The same process is applied to the radio detection map.

4. The data cube of isolated contaminants and the isolated source undergo a PSF-matching procedure. Following this, models are constructed and normalized to be utilized in the subsequent optimization phase. To achieve this, each isolated source is convolved with the corresponding convolution kernel generated by PhoEBO from the input PSF models. These kernels are obtained using the python *Photutils*[3] library

[1] Source-Extractor, Bertin & Arnouts 1996
[2] A python wrapper to Source-Extractor (Bertin & Arnouts, 1996).
[3] Photutils Astropy affiliated package

that generates a matching kernel between two PSFs using a procedure based on the ratio of Fourier transforms (see Gordon et al. 2008; Aniano et al. 2011).

5. In the optimization step, the models of contaminants and the source are multiplied by a series of multiplicative factors. The objective is to minimize the chi-square between the observed data and the complete model (comprising contaminants and source). If the optimization process terminates successfully, the models of contaminants (and only the models of contaminants) are subtracted from the target map. At the end of this procedure, the target source appears as a residual in the subtracted map.
Best-fit flux values for the contaminants are saved in a separate file.

6. In the final step, aperture photometry is conducted on the residual map. Fluxes and magnitudes are computed based on the ZeroPoint specified by the user. Concurrently, error estimations are derived as:

$$\Delta F = \sqrt{\sum_{i \in A} \sigma_i^2} \tag{4.1}$$

where the sum is extended to all the pixels in the aperture used for flux measurement and σ are obtained from the weight map. Noteworthy, by default, PhoeBO treats as upper limits the errors computed at 1σ level.

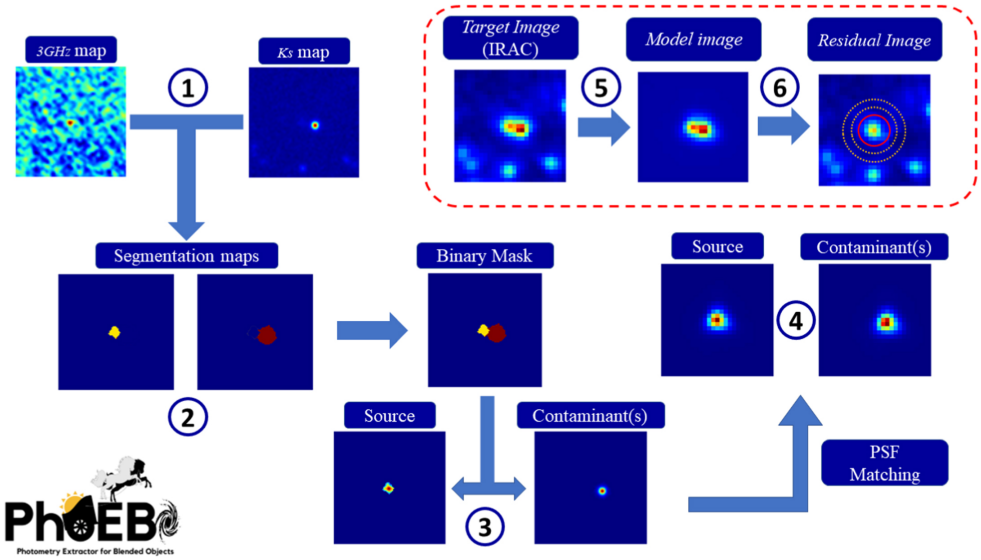


Figure 4.2: *Scheme of the main logical steps involved in the operation of PhoEBO. The scheme is taken Gentile et al. (2023).*

4.2 Building The MIR-to-FIR Photometric Catalogue with PhoEBO

A substantial part of this thesis has been directed toward the codebase of PhoEBO, extending its capabilities to longer wavelength bands and other quality-of-life improvements (see Appendix B). In its current implementation, I employed it to derive the photometry of all sources analyzed in this study, target sources and contaminants. The accuracy of the pipeline in extracting photometry for the targets (my NIR-dark candidates) has been previously assessed through extensive simulation in Gentile et al. (2023) and will not be revisited in this chapter. Nonetheless, as part of my efforts to expand PhoEBO's capabilities, this chapter will provide an overview of the reliability of the code in accurately extracting photometry from all sources within the analysis area.
PhoEBO's approach to modeling contaminants and the sources themselves has been extended, thus allowing for the fitting of all sources within the analysis area to obtain flux measurements. This method offers two key benefits:

1. The models, being constructed from the high-resolution segmentation map of prior sources, enable PhoEBO to fit and extract photometry for extended sources.

2. Model optimization addresses the blending issue for all considered sources, thereby yielding more precise measurements than typical aperture photometry-based programs would achieve.

It is important to acknowledge that, while effective, this approach is not without its limitations.
The efficacy of PhoEBO, especially in the aspects of source deblending and segmentation, is highly dependent on the user-defined configuration parameters. For example, in scenarios where sExtractor (Bertin & Arnouts, 1996) erroneously aggregates distinct galaxies into a single group, PhoEBO, constrained by its initial inputs, is unable to subsequently deblend them.
Sometimes, PhoEBO might report flux values that are much higher than they should be. This may happen with sources that are near the edge of the area being analyzed; these sources are usually discarded because they are too close to the edge, and it might not be possible to measure all of their flux accurately. However, if there is another source close to the one that was ignored, PhoEBO might not have a way to correctly measure the flux of both sources separately. As a result, it might mistakenly add some of the flux from the ignored source to the nearby one. I have verified through visual inspection that this issue affects only a small fraction of sources.

4.2.1 IRAC photometry with PhoEBO

IRAC ch1 and ch2

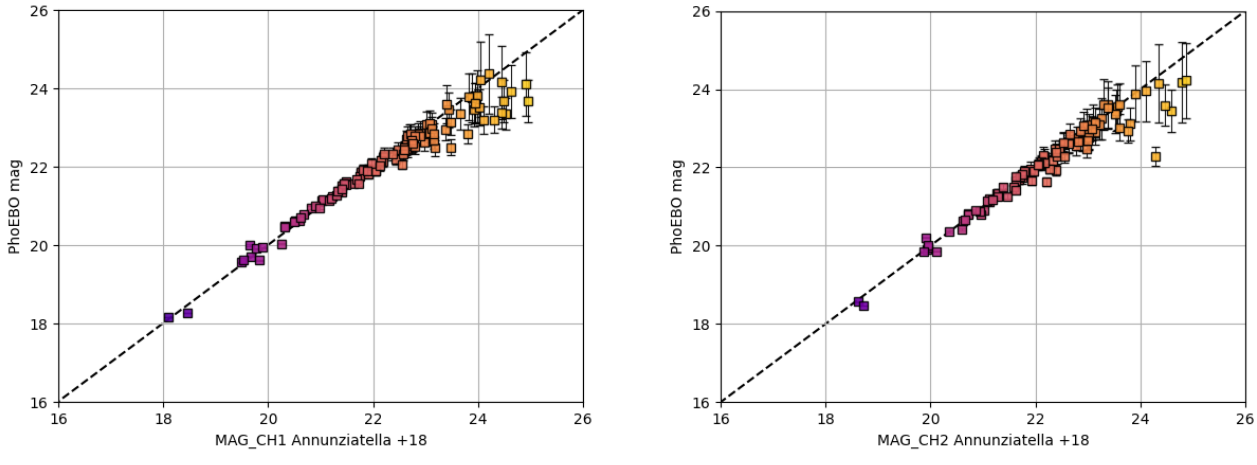
At the time of writing, photometric catalogs for the J1030 field in bands beyond IRAC ch2 are not present, thus the effectiveness of PhoEBO in accurate flux extraction was benchmarked against the deblended catalog for IRAC 1 and 2, as provided by Annunziatella et al. (2018).

for the comparison, photometry was carried out using PhoEBO near the central region of the maps, in an area of 110 square arcseconds and centered around one of my target sources. To ensure accuracy, all sources within 10 arcseconds from the outer perimeter were omitted from the analysis, effectively narrowing the effective area to 100 arcsec^2.

The results, illustrated in Figures 4.3a and 4.3b, demonstrate the alignment between the pipeline's flux best-fit values and those cataloged, with a Mean Absolute Error (MAE) of 0.20 mag for IRAC1 and 0.18 mag for IRAC2 when comparing the results from PhoEBO to those reported by Annunziatella et al. (2018). Minor deviations are noted for very faint sources, close to the detection limit of the map. The lack of significant differences between the two sets of data indicates a strong agreement between PhoEBO's outcomes and the results previously established.

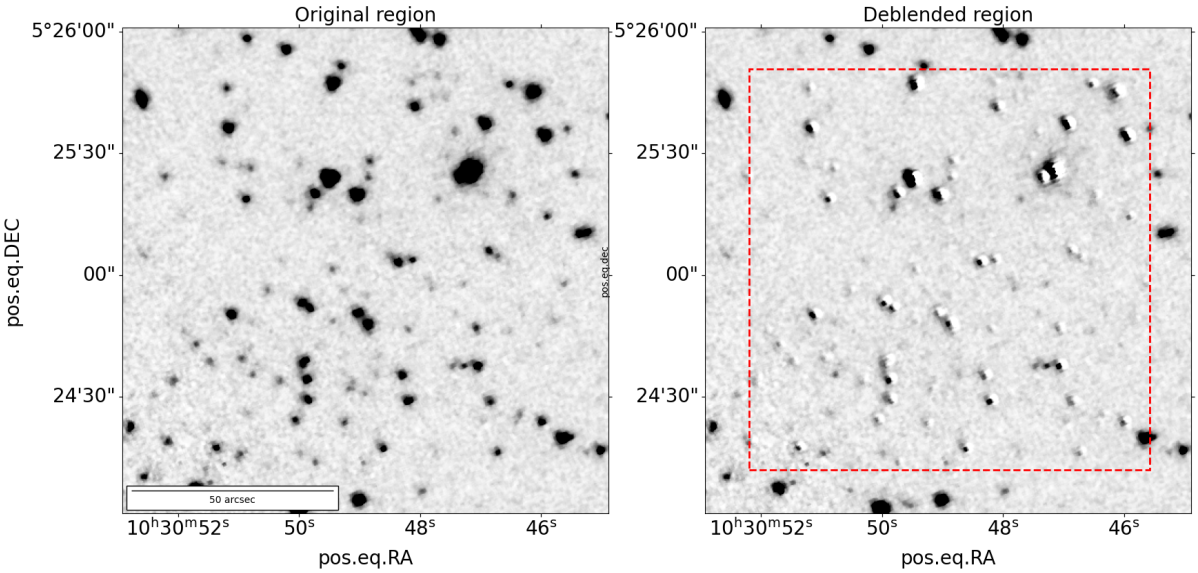
The region and the corresponding residuals from which these photometric results were obtained are visible in Fig 4.4.

To correct minor astrometric errors and enhance the quality of the fit, I allowed PhoEBO to freely fit the position of the source's centroid as a free parameter, with a maximum shift allowance of 1.5 pixels (approximately 0.4 arcseconds). The difference in residuals between allowing PhoEBO to fit the centroid of the source and when the position is fixed is visible in Fig 4.4a.

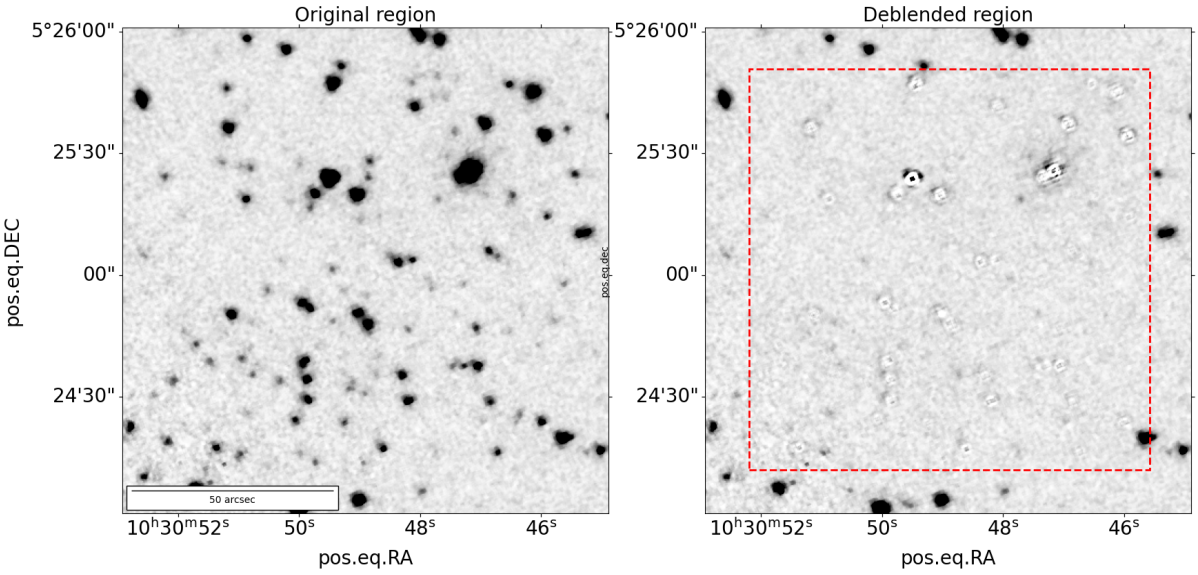


(a) *(b)*

Figure 4.3: ***Left*** *Comparison between the photometry obtained from PhoEBO and that by Annunziatella et al. (2018), for the IRAC ch1 band.* ***Right*** *as for (a) but for IRAC ch2*



(a)



(b)

Figure 4.4: *(**a**)**Left**: an IRAC ch2 band cutout measuring 110x110 arcseconds.* ***Right**: the residuals from PhoEBO without centroid fitting. The dashed red square represents the actual area of 100"x100" where PhoEBO has fitted the sources.*
*(**b**) as for (**a**) but with centroid fitting.*

Proper choice of PSF

Like many source fitting programs, PhoEBO is dependent on user inputs, by far, the most influential being the user-provided PSF. The precision of the outcomes is influenced by the dependability of the PSF and, by extension, the convolution kernel that the pipeline generates. A marginal error in the central slope of the PSF's light profile can result in substantial systematic deviations in the flux measurements. Furthermore, the fitting algorithm employed operates by minimizing residuals through pixel summation or subtraction, therefore, if the PSF profile is inaccurately sampled, it will consistently lead to distinctive residual patterns, manifesting as concentric positive and negative rings (e.g. Fig 4.5). These patterns in the residuals are indicative of the discrepancies between the observed data and the model predictions, and their presence serves as a sign of an imprecise PSF profile.
This behavior is especially pronounced in IRAC images, which are characterized by a highly variable, typically triangular PSF. Moreover, the creation of mosaics through the stacking of different observations further modifies the characteristic PSF of the final image.
This limitation can be mitigated through tools like PSF-Extractor (PSFEx; Bertin 2013), which generates a model of the image's PSF by analyzing isolated point sources and then provides it to the end user. PSFEx's approach to PSF modeling is effective in addressing the issue of local PSF variability, particularly in complex imaging scenarios like those with IRAC, however, a key constraint of such extraction methods is their requirement for a statistically significant number of sources to construct a representative model. This requirement limits the applicability of the method, particularly in cases where sufficient point sources are not available or the map is particularly noisy.
In my research, this limitation meant that I could derive a PSF using PSFEx only for IRAC bands 1 and 2 (Fig 4.6).

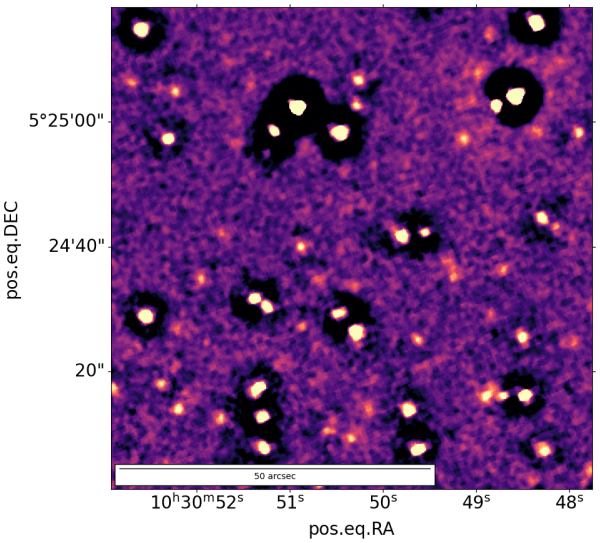


Figure 4.5: *Zoom over the center of the region in Fig 4.4 showing the aftermath of a source-subtraction phase where the residual artifacts take on a distinctive ring-like pattern, indicating an imperfect subtraction due to bad PSF modeling.*

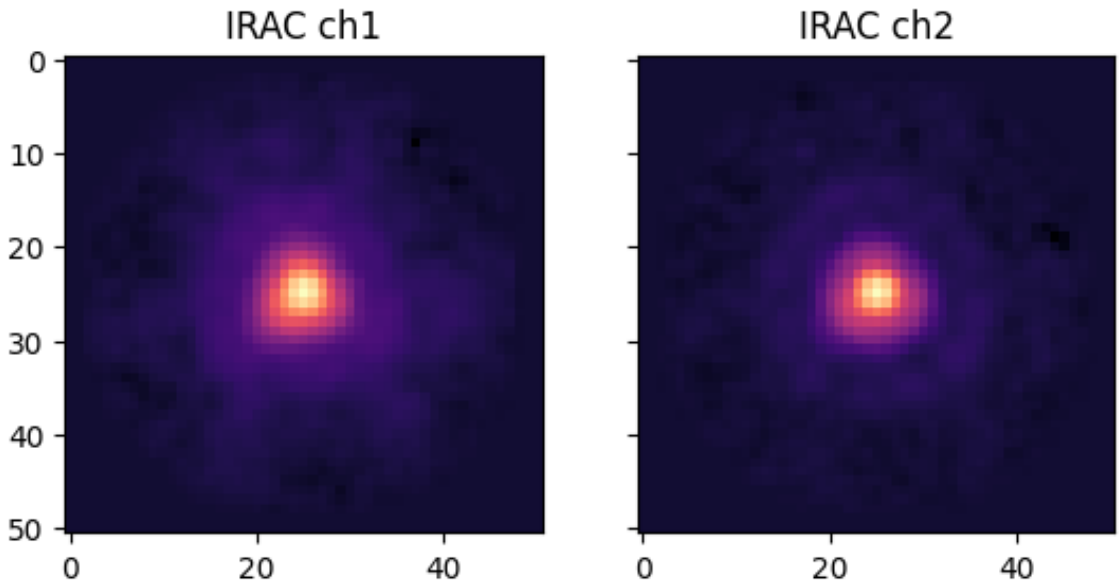
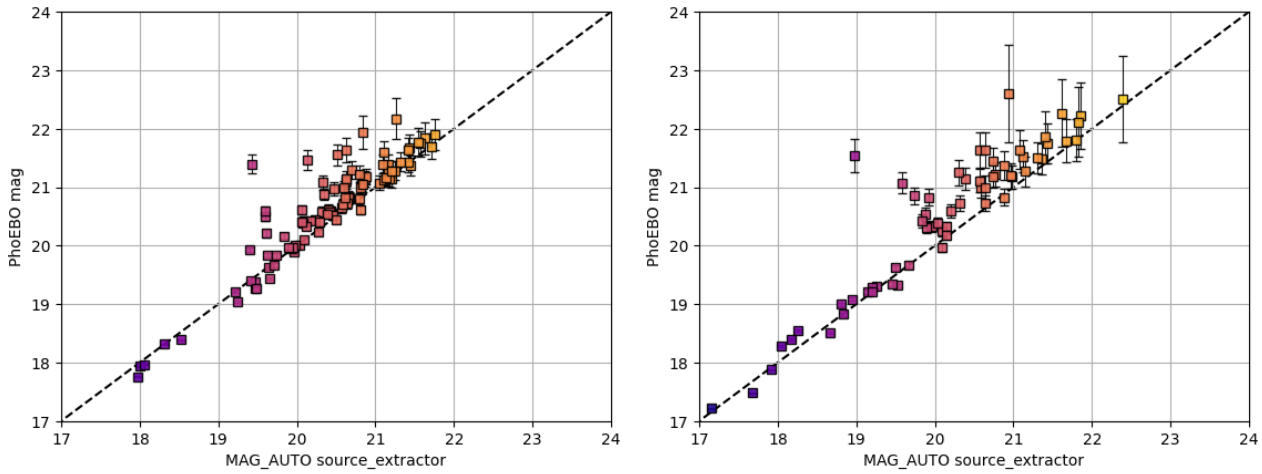


Figure 4.6: ***Left*** *Effective PSF for IRAC ch1 extracted using PSFEx.* ***Right*** *Effective PSF extracted for IRAC ch2.*

IRAC ch3 and ch4

As previously outlined, a complete deblended catalog for the IRAC ch3 and ch4 maps is currently unavailable. Nevertheless, a key objective of this study is to conduct a comparative analysis of PhoEBO's performance within these bands. To achieve this, I utilized the full implementation of sExtractor (Bertin & Arnouts, 1996) on a carefully selected cutout from the IRAC ch3 and ch4 maps near the map's center, away from the edges, as the influence of the borders significantly increases the map's rms and may affect the sExtractor algorithm, causing it to identify an anomalous number of sources despite careful selection of convolution and clean parameters.
Photometric data for the same map region were acquired using both sExtractor and PhoEBO. It is important to acknowledge that in this case sExtractor is not used directly to verify PhoEBO's reliability because it doesn't employ prior-based deblending like PhoEBO does. Hence, I don't anticipate an agreement comparable to what I obtained for IRAC ch1 and ch2. Consequently, I expect sExtractor to yield generally higher flux measurements than PhoEBO, though it should produce similar results for isolated and/or bright sources.
The comparison results can be seen in Fig 4.7a and 4.7b. As predicted, bright source fluxes align well across both methods (all sources with mag < 19 show a difference of only 0.22 mag between PhoEBO and sExtractor). Furthermore, a pattern emerges where sExtractor often reports higher flux values compared to PhoEBO. This is likely due to sExtractor's susceptibility to likely blended sources. This observation reinforces the expectation that sExtractor, lacking PhoEBO's deblending capability, tends to overestimate flux for potentially blended

sources.



(a) *(b)*

Figure 4.7: ***Left*** *comparison between the photometric results obtained from PhoEBO and those reported by Source-extractor for IRAC ch3* ***Right*** *as for (a) but for IRAc ch4.*

4.2.2 MIPS 24 μm Photometry with PhoEBO

I used PhoEBO also to perform photometry at 24 μm. Spizer/MIPS 24 μm PSF size, with his FWHM of 4.5", is larger than the intrinsic size of most sources (except for rare cases of very low-redshift galaxies), making PhoEBO's analysis close to a PSF fitting procedure. The shape of the PSF used was directly derived from the empirical one calculated and made available in the Spitzer archives[4].

The photometry was performed based on IRAC ch1 and ch2 priors at fixed RA, DEC allowing PhoEBO to fit the centroid position of the sources as a free parameter with a maximum allowed shift of 1 pixel (2.45 arcseconds). Sources fitted simultaneously are forced to share the same shift to preserve the complex morphology of extended groups [5].

The ability to fit the relative positions of sources allows for the correction of minor astrometric errors, simultaneously improving the quality of the fit. Additionally, PhoEBO automatically reruns the source search/fitting procedure until a certain threshold is reached (i.e. No more sources are detected following the criteria in sec 4.1.1, or the maximum number of iterations has been exceeded), significantly reducing the workload for residual analysis and potential flux reassignment.

As previously done with the IRAC photometry, the analyzed area is 110' × 110', with all sources within a 10-arcsecond border excluded from the fit. This results in an actual 'effective' area of 100' × 100'.

[4] Spitzer Archives

[5] Refer to the Appendix B to see how PhoEBO determines the source groups

To better showcase the capabilities of PhoEBO, I show in Fig 4.8 an example covering a larger area of 200' × 200', with an effective area of 180' × 180', showing the residuals after the subtraction of the fitted sources.

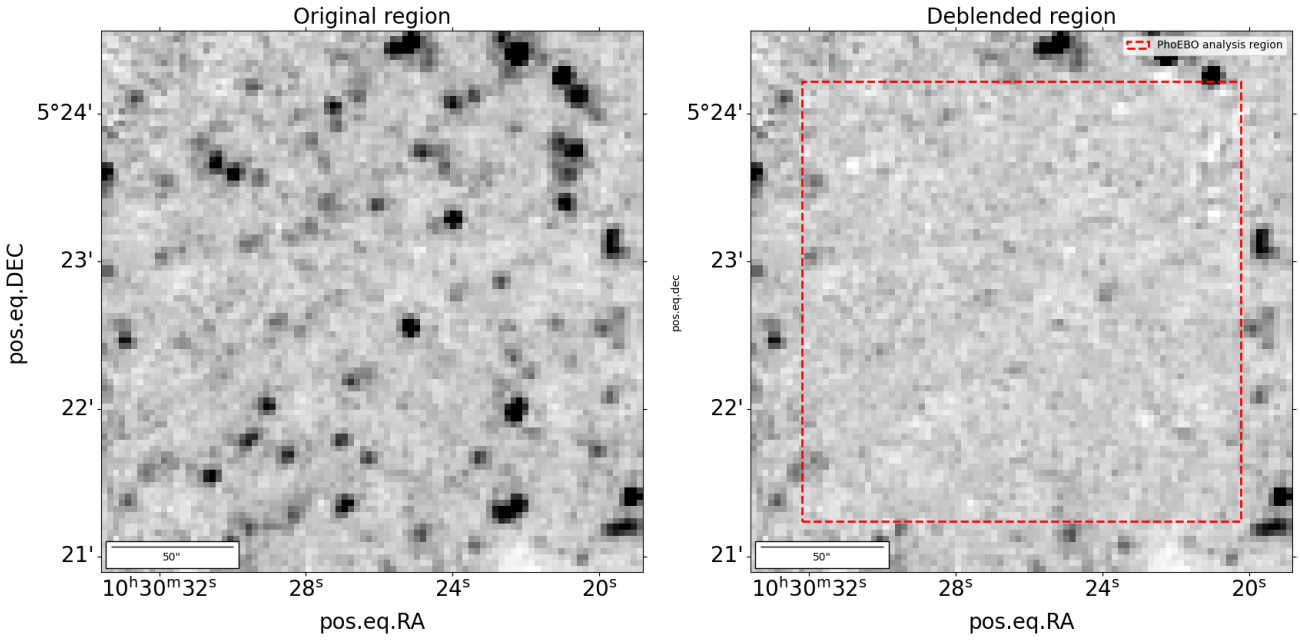


Figure 4.8: ***Left****: region of 200'x200' centered on one of my sources.* ***right****, residuals as produced by PhEBO in the same region. The actual area of analysis is outlined in dashed red.*

4.2.3 PACS 100 and 160 μm Photometry with PhoEBO

Due to the strong correlation between IR luminosity and luminosity at 24 μm (e.g. Dale & Helou 2002b; Dale & Helou 2002a), it is common in Herschel bands deblending to use priors at 24 μm as a starting point (see Hurley et al. 2017; Jin et al. 2018). A common method to obtain these priors is through 'blind extraction'.

To address this limitation, I used again PhoEBO for the photometry, with the positions of the 24 μm sources obtained during the MIPS photometry phase (sec 4.2.2) as priors. All sources with S/N $>$ 1 were included and chosen as possible priors before the fit in PACS bands.

The PSFs employed for modeling the sources in these instances are Gaussian, with the FWHM equal to that specified in the Herschel Handbook and reported in Tab 3.1

4.3 'Superdeblending' of SPIRE Bands

Liu et al. (2018) developed a 'superdeblending' method to reliably extract FIR fluxes based on 24 μm prior in the GOODS-North field. A similar approach has been successfully employed in COSMOS (Jin et al., 2018). My work draws inspiration from them to extract accurate photometry for my sources in the FIR bands.
It is important to note that obtaining a deblended catalog in the FIR/sub-mm bands and studying the physical characteristics of sources that are not part of my sample is not the primary goal of this work, therefore, they are not covered in great detail.
The analysis of different sources varied due to diverse data coverage. The lack of uniform coverage across SPIRE, PACS, and MIPS bands meant that, while some sources can be analyzed with data from all these instruments, others were limited to data from bands at shorter wavelengths.

4.3.1 General workflow for SPIRE Superdeblending

The initial step involves selecting sources from my catalog based on specific criteria: having coverage in both MIPS 24 and SPIRE bands.
Out of a total of 79 initial sources, 36 met this requirement and thus form what is referred to as the *primary sample*. Of these 36 sources, only 6 have coverage in the PACS bands.
All remaining sources without coverage in SPIRE and MIPS bands form the *secondary sample*.
The analysis of the *primary sample* follows this scheme:

- **Creation of the initial catalog of 24 μm priors** derived from the photometric results of PhoEBO (see next section for more details about the procedure)
- **Flux prediction via SED fitting**: for each of the sources in the initial 24 μm catalog, the flux at 250 µm is predicted through SED fitting.
  For all sources that have coverage in the PACS bands, a stringent prior on the 250 μm flux, derived by data at 100 and 160 μm, is applied. In these cases, the flux predicted by the SED fitting is utilized.
  For sources lacking information in the PACS bands, the SED fitting process is limited to data available up to 24 μm. As tested, direct extrapolation from 24 to 250 μm can be imprecise, often leading to overestimated results. To address this, a second approach has been attempted for these sources, involving the construction of a median SED (see Appendix C).
- **Creation of Models and Optimization**: after obtaining flux values at 250 μm, I constructed models based on these predicted fluxes and optimized them. The fluxes obtained through this process then serve as my new photometric points.
- **Flux Prediction at 350, and 500 μm**: once the fluxes at 250 μm are obtained at

the end of the optimization step, a subsequent round of SED fitting is conducted to predict fluxes at 350 μm. Then the procedure is repeated for the 500 μm band.

4.3.2 'Superdeblending' Limitations and Source Optimization

The discrepancies between flux predictions from the two initial methods used for sources lacking PACS coverage (direct extrapolation from 24 μm and median SED) and actual observations often manifest as either overestimations or underestimations.
In galaxies modeled using the median SED, the mismatch could originate from two main factors:

1. The median SED that I derived might not be representative of the actual source analyzed.

2. The photometric redshift may not be correct. Tools like MAGPHYS + PHOTOZ require photometric data points in the FIR to impose stronger constraints on the photometric redshift of the source.

I encountered similar problems also for a small number of sources with PACS priors, in that cases inaccuracies might arise from suboptimal photometric coverage in the mid-infrared and FIR bands.
For each source in my sample, I compared the two methods previously described to predict the 250 μm flux. To do this, the RMS is calculated at a three-sigma level and then compared both before and following the subtraction of modeled sources using the fluxes derived (Fig 4.9). Special care was taken to mask my primary sources during this process to ensure they did not affect the RMS calculations. The analysis of the residuals map, however, highlits significant emission that was not accounted for, thus requiring the additional optimization step already introduce in Sec 4.3.1, whose results are shown in Fig 4.9.

4.3.3 Model Creation and Optimizarion

The first step of this procedure involves constructing models for both contaminants and sources, each paired with the optimal SED flux determined in the previous phase.
For those sources for which I was able to obtain photometric data at 100 and 160 μm using PhoEBO, thanks to a sufficiently stringent prior, the SED fitting managed to predict the 250 μm flux with reasonable accuracy for most of the cases. However, due to the limited depth of the map, the number of sources with a robust detection above 3σ is extremely constrained (less than 20 %).
In the case of sources lacking PACS coverage, I opted to use the fluxes predicted by the extrapolation from 24 μm since no statistical differences are present between median-SED and 24 μm extrapolation (see Fig 4.9).

All sources are approximated as Gaussian, with their FWHM corresponding to the values reported in Tab 3.1, consequently, the amplitudes are directly proportional to the predicted flux. Later, using these predicted fluxes as starting points, the total model -target sources plus contaminants- undergoes refinement through a chi-square minimization procedure.
In cases where the modeled fluxes for sources in this phase are less than the relative rms of the region, these are designated as upper limits. The fluxes, once adjusted, are then incorporated as photometric points for the relevant band, and the iteratively SED fitting process is repeated.
An example of this process is illustrated in Fig 4.10. The same method of model creation and subsequent re-optimization has also been applied to the SPIRE 350 and 500 μm bands, employing the same logical steps discussed previously.
In the analysis at 250 μm, only those sources demonstrating a S/N $>$ 1 in the preceding band were selected as priors for positions. For SPIRE 350, only sources exhibiting a flux at 250 μm (flux predicted by SED fitting) with an S/N greater than 1 were used as prior sources and so on (Fig 4.10).

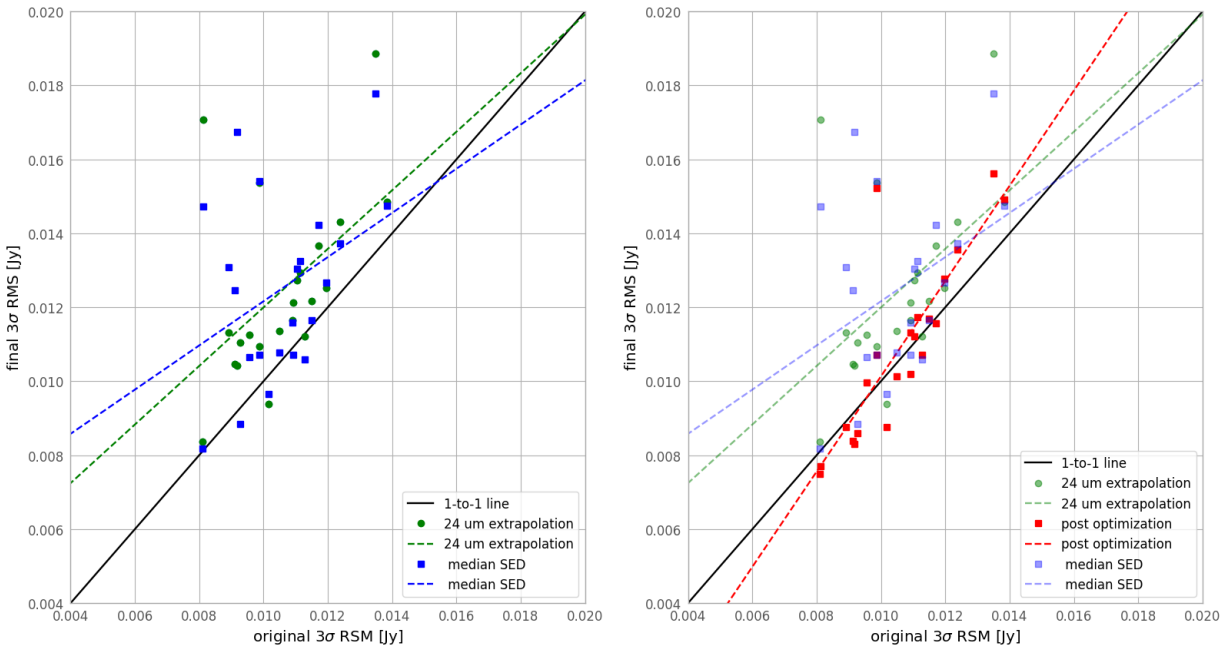
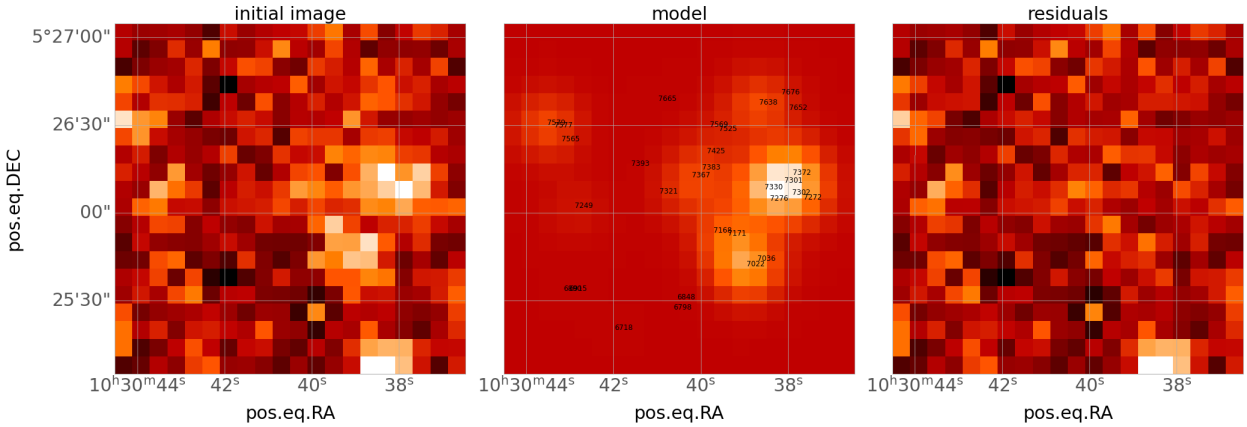
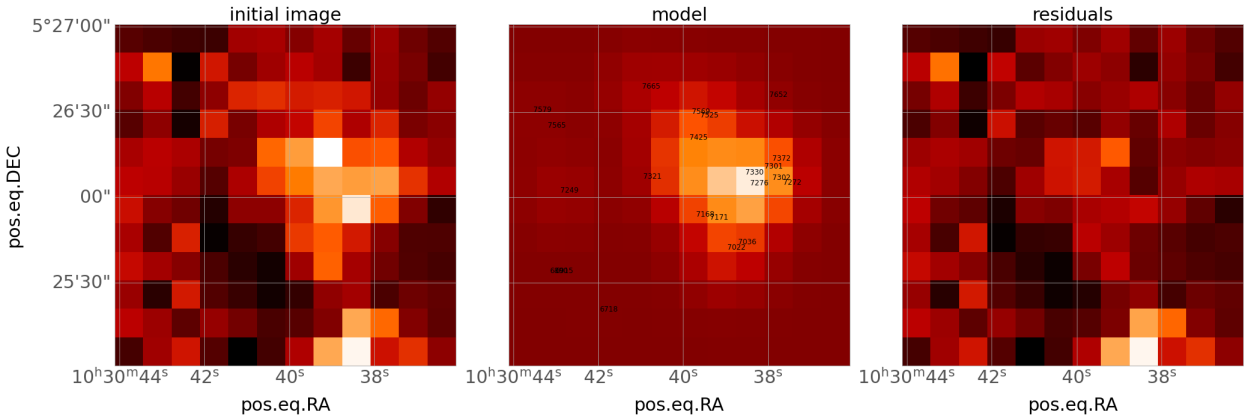


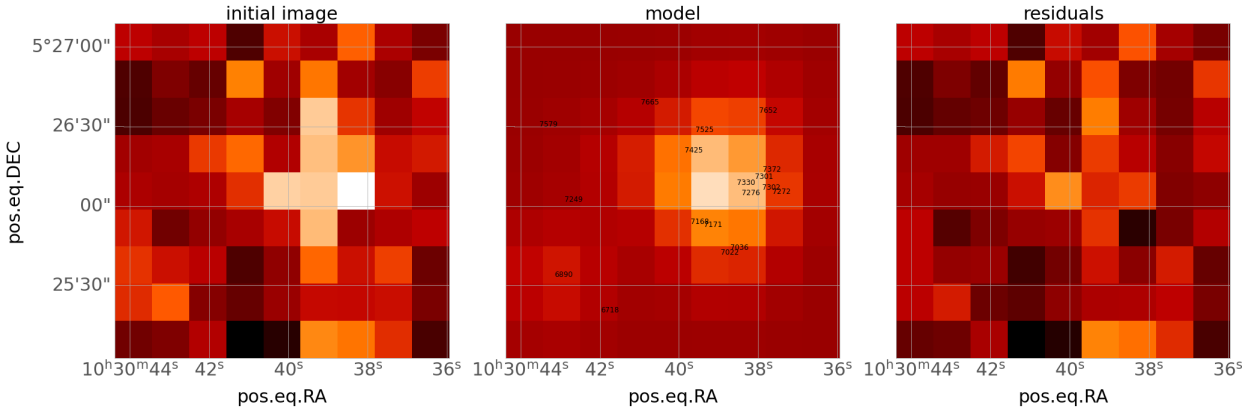
Figure 4.9: ***Left*** *RMS comparison for the primary sample of sources. The left panel displays the RMS values calculated before (original) and after (final) the application of two flux prediction methods: 24 μm extrapolation (green circles) and the median Spectral Energy Distribution (SED) approach (blue square); dashed lines represent the linear best-fits. The right panel extends the comparison to include the post-optimization method (red squares). In black the bisector line.*



(a)



(b)



(c)

Figure 4.10: *(**a**) **Left**: a cutout centered on source 508 (id: radio catalog). In the **center**: model created with best-fit fluxes which comprise both target source and contaminants, with the IDs of the sources used as priors on the positions for the contaminants shown in black. **right**: residuals after model subtraction. (**b**) and (**c**): Similar to (**a**), but for SPIRE 350 and 500 μm, respectively. It is possible to observe how the density of priors varies based on the selection criteria used.*

4.4 Deriving Physical properties and Photometric Redshifts through SED Fitting

After compiling the photometric catalogue from my sample, to obtain photometric redshifts and physical properties I performed SED fitting procedure using the MAGPHYS + PHOTOZ code (da Cunha et al. 2008; Battisti et al. 2019). MAGPHYS models emission across the entire UV-to-FIR range, assuming a balance of energy output between emission at UV-to-NIR wavelengths and that absorbed by dust, subsequently reemitted in the FIR. MAGPHYS measures Posterior Distribution Functions (PDF) of parameters by fitting observed photometry to a set of galaxy emission models derived from native libraries. These libraries consist of 50,000 stellar population spectra with star formation histories approximated by a delayed exponential function, incorporating superimposed random bursts (Bruzual & Charlot, 2003). These stellar population spectra are associated, via the energy balance criterion, with 50,000 dust emission Spectral Energy Distributions featuring two components at different temperatures (see Charlot & Fall 2000 da Cunha et al. 2008). In the MAGPHYS + PHOTOZ extension (Battisti et al., 2019), the code displaces these models across a broad redshift grid to infer the PDF of the galaxy photo-z.
In instances of non-detections, I choose to adopt the 1σ rms as upper limits. Notably, MAGPHYS does not enforce the SED model to strictly stay below the limit; instead, it treats upper limits as values with zero flux and an error equal to the upper limit value.

4.4.1 Stellar Component

For the analysis of the stellar component, the program employs the Bruzual & Charlot (2003) synthesis models for Single Stellar Populations (SSPs), adopting a Chabrier (2003) IMF and applying a uniform prior for metallicity that spans from 0.2 to 2 times that of the Sun. These models map the evolution of stellar populations ranging in age from 1×10^5 to 2×10^{10} years, offering a resolution of 0.003 μm across the wavelength spectrum from 0.32 μm to 0.95 μm, and extending up to 160 μm at a reduced resolution. The metallicity in these models varies between 0.02 $Z_\odot$ and 2 $Z_\odot$, with an instantaneous starburst assumed for the SFH, implying all stars were formed simultaneously in a single stellar formation event. The monochromatic luminosity of an SSP at a time t is calculated as follows:

$$L_\lambda^{em}(t) = \int_0^t \psi(t-t') l_\lambda^{SSP}(t', Z) exp[-\hat{\tau}_\lambda(t') dt'] \tag{4.2}$$

where $l_\lambda^{\rm SSP}(t', Z)$ is the monochromatic luminosity per unit mass emitted by an SSP of age t' and metallicity Z. $\hat{\tau}_\lambda(t')$ is the effective optical depth of dust absorption experienced by stars of age t'. Finally, $\psi(t-t')$ represents the evolution of the star formation rate over time

and is characterized by a delayed exponential function:

$$\psi(t) \sim \gamma^2(exp)(-\gamma t) \tag{4.3}$$

Another parameter worth discussing is the optical thickness $\hat{\tau}_\lambda(t')$, for which the Charlot & Fall (2000) model is used. A detailed description of this complicated parameter can be found in Charlot & Fall (2000), still, the model can be summarized as follows:
Young stars form in molecular clouds and their intense energy ionizes the gas around them, creating glowing regions known as HII regions. These regions emit light, both from the direct ionization process and from the broader, non-ionized light that dust particles in the surrounding gas absorb and re-emit. However, these molecular clouds don't last forever; they typically exist for about 10^7 years. The lifespan of these HII regions is dictated by how long the stars within can keep emitting enough energy to prevent the cloud from dissipating. If a star emits UV and visible light longer than the cloud lasts, then its light goes directly into the Interstellar Medium (ISM) and faces less blocking by dust. This means the light from such stars doesn't get as dimmed or 'extinct' as it travels through space.

4.4.2 Dust Emission

In MAGPHYS + PHOTO-Z the dust emission model follows the principles outlined in the basic version of MAGPHYS (da Cunha et al., 2008), which in turn builds on the model of Charlot & Fall (2000).

Molecular Clouds

The total emission from a molecular cloud is described as follows:

$$L_{BC;d} = (\xi_{PAH}^{BC} l_\lambda^{PAH} + \xi_{MIR}^{BC} l_\lambda^{MIR} + \xi_W^{BC} l_\lambda^{T_W^{BC}})(1 - f_\mu) L_d^{tot} \tag{4.4}$$

where $L_{tot,d}$ is the infrared luminosity re-radiated by dust, f_μ is the fraction of $LBC_{\lambda,d}$ due to the ISM, $l_{PAH,\lambda}$, $l_{MIR,\lambda}$, $l_{TBC,W,\lambda}$ are the normalized spectral distributions of PAHs, mid-infrared continuum, and warm dust in thermal equilibrium, and ϵ_{BC}^{PAH}, ϵ_{BC}^{MIR}, ϵ_{BC}^{W} are the fractions according to which the three components contribute to the cloud's total infrared luminosity.

Interstellar Medium

The dust emission due to the ambient medium is given by:

$$L_{\lambda;d}^{ISM} = (\xi_{PAH}^{ISM} l_\lambda^{PAH} + \xi_{MIR}^{ISM} l_\lambda^{MIR} + \xi_W^{ISM} l_\lambda^{T_W^{ISM}} + \xi_C^{ISM} l_\lambda^{T_C^{ISM}}) f_\mu L_d^{tot} \tag{4.5}$$

where l_λ^{PAH}, l_λ^{MIR}, $l_\lambda^{T_W^{ISM}}$, $l_\lambda^{T_C^{ISM}}$ are the normalized spectral distributions of PAHs, mid-infrared continuum, warm dust, and cold dust in thermal equilibrium. In the ambient interstellar medium, the proportions given by the variables ξ_{PAH}^{BC}, ξ_{MIR}^{BC}, ξ_W^{BC}, ξ_C^{ISM} are, unlike in the clouds, set according to the observed proportions in the emission of the Milky Way's cirrus.

4.4.3 Radio Component

The modeling of the radio component is based on two assumptions:
The logarithmic scale ratio between radio and infrared flux density, q_{tir} (see Chapter 5), is centered at a value of $q_{tir} = 2.34$, with a dispersion of $\sigma_{q_{tir}} = 0.25$. The slope and relative contributions for thermal and non-thermal emission are fixed.
Radio emission is described as the sum of a thermal component, $L_{\rm th}$, due to the free-free processes and fixed at 10% of the total at 20 cm, and a non-thermal component, $L_{\rm nth}$:

$$L_\nu^{th} \sim \nu^{-0.1} \;\; ; \;\; L_\nu^{non-th} \sim \nu^{-0.8} \tag{4.6}$$

It is crucial to emphasize that MAGPHYS + PHOTO-Z does not incorporate the ability to model AGN components and thus relies on the assumption that all radio emission (both thermal and non-thermal) primarily stems from star formation processes. Therefore, ensuring that the galaxy being analyzed does not contain excess radio emission is a critical step and will be discussed in more detail in the results chapter.

Chapter 5

Results

In this chapter, I present the results of my analysis. In order to derive the most robust conclusions on the physical properties of my Radio-Selected NIR-Dark sources and their contribution to the SFRD I decided to further refine my final sample with two additional criteria:
1) Having at least one detection (i.e. a photometric point that is not an upper limit) in both MIR and FIR bands.
2) A reduced chi-square (χ^2_ν) value in the sed fitting analysis falling within 3σ of the total sample distribution, a requirement that translated into χ^2_ν<5.
Following these criteria, my *primary sample* of 36 galaxies was distilled to a core group of 8, hereby designated as the *golden Sample*.

5.1 Physical Properties from SED fitting

The derivation of physical properties and photometric redshifts for my NIR-dark galaxies is conducted through SED fitting, employing the MAGPHYS + PHOTO-Z code (Battisti et al., 2019) as highlighted in the previous chapter. Following Gentile et al. (2023), who confirmed via extensive simulations a small bias by PhoEBO, an additional 0.15 magnitude was added to the photometric uncertainties in the results.
The median values alongside the distributions of physical properties for all samples are presented in Table 5.1 and Figure 5.1, respectively. The general goodness of fit of the procedure for the *primary sample* is highlighted by a median χ^2_ν value of 1.35. In contrast, the *secondary sample* exhibits a median χ^2_ν value of 0.55.
The median IR Luminosities (L_{Dust}) of 4.4×10^{11} and 3.5×10^{11} $L_\odot$ of the *primary* and *golden sample* respectively and the particularly high median values of dust extinction (A_V $\approx$ 6 mag) confirm the dusty nature of the samples. Also, the dust luminosity-weighted temperature estimated via SED fitting for both *primary* and *golden* samples has median values of $T_d \approx 40$ K, slightly higher than that estimated for main sequence galaxies by Schreiber et al. (2018) of $T_d \approx 35$ K, suggesting the possible starburst nature of my samples.

I do not report results on dust properties such as luminosity, mass, temperature, and extinction in the *secondary sample* because of the lack of FIR constraints.

Table 5.1: *Median physical properties estimated by SED fitting for primary, secondary samples, and golden samples. The uncertainties on the median properties are estimated using the Median Absolute Deviation (MAD) following Hoaglin et al. (1983) as MAD = 1.482×median(\|x_i-median(x_i)\|)/$\sqrt{N}$, where N is the number of galaxies in the sample.*

Property	Primary sample	Secondary sample	Golden sample	Units
$z_{\rm phot}$	1.60 ± 0.19	2.53 ± 0.21	1.63 ± 0.36	
M_*	$(7.6 \pm 1.6) \times 10^{10}$	$(2.2 \pm 0.35) \times 10^{11}$	$(7.1 \pm 2.9) \times 10^{10}$	$M_\odot$
SFR	178 ± 34	715 ± 108	169 ± 51	$M_\odot$ yr^{-1}
A_v	5.9 ± 0.44		6.2 ± 0.62	mag
$L_{\rm Dust}$	$(4.4 \pm 0.71) \times 10^{11}$		$(3.5 \pm 0.73) \times 10^{11}$	$L_\odot$
$T_{\rm Dust}$	40.4 ± 0.48		40.5 ± 0.68	K
$M_{\rm Dust}$	$(3.5 \pm 0.55) \times 10^{7}$		$(3.4 \pm 0.74) \times 10^{7}$	$M_\odot$

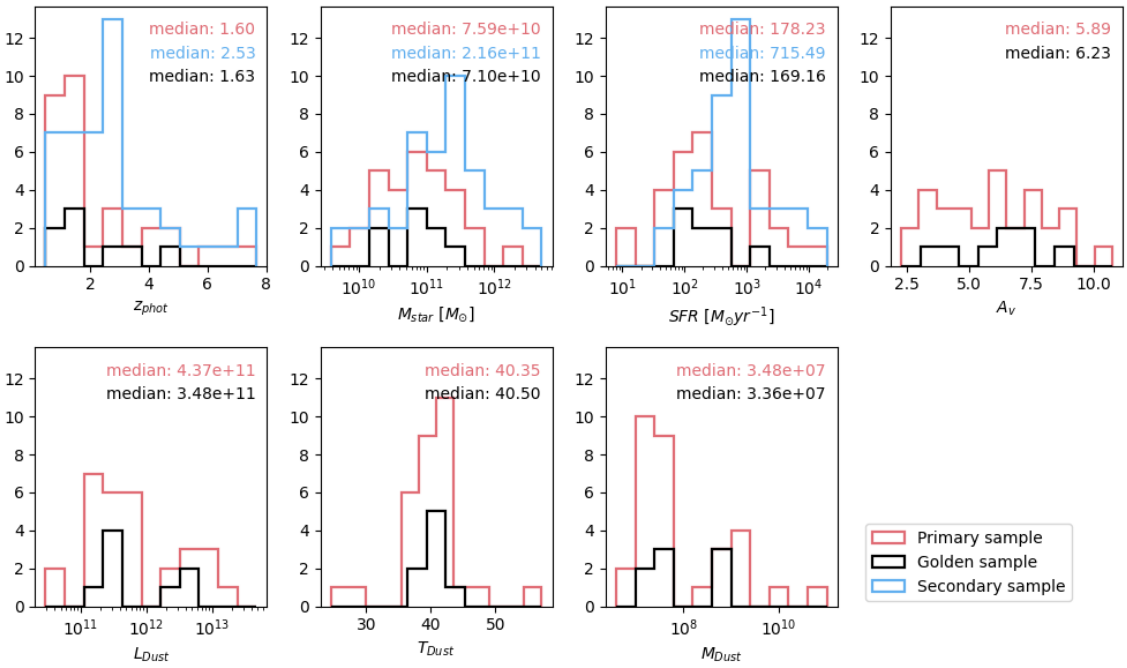


Figure 5.1: *Histograms of the main physical properties obtained through SED fitting with MAGPHYS + PHOTO-Z, with the primary sample indicated in red, the secondary sample in blue, and the golden sample in black. The median values are the same as those reported in Table 5.1.*

5.2 Redshift distribution

The comparison between the *primary* and *secondary* samples (see Fig 5.1), reveals a difference, with the two samples having two distinct median redshift values respectively of 1.6 and 2.5. Such difference may be a side effect of the fact that codes like MAGPHYS + PHOTO-Z use FIR points to place major constraints on the source's redshift, leveraging the distinct shape and features of the dust black-body emission. The absence of such data may introduce systematic uncertainties, highlighting the importance of multiwavelength analysis in SED fitting. The median value of the *primary sample*'s distribution is significantly lower than those reported in analogous research, for instance, studies like Behiri et al. (2023), focusing on radio-selected NIR-Dark galaxies in the COSMOS field, found their sample's median redshift values to be z = 3.3. I point out that the sample analyzed by Behiri et al. (2023) is defined as sources having at least one detection above 3σ in FIR/sub-mm bands, which is broadly consistent with my definition of the *golden sample*. Part of the difference may be traced back to the flux limits of the data used in their study. Essentially, since the flux limit determines the faintest object that can be detected within a survey, lower flux limits allow the detection of fainter, more distant objects. Behiri et al. (2023) operated with lower flux limits for Ks-band observations to those of J1030, but slightly fainter for radio observations (see Tab 3.3). This difference implies that galaxies at a lower redshift were less likely to be included in their samples due to the selection criteria based on the Ks-band observations, effectively trimming the lower-z tail from their datasets. In Fig 5.2 I compare the Empirical Cumulative Distribution Functions (ECDF) of the various samples: those of J1030 clearly show a higher percentage of low-z sources compared to Behiri et al. (2023). To perform a more meaningful comparison, I filtered the *primary sample* by removing all those galaxies whose best-fit flux (from SED fitting) in the Ks band was greater than the 3σ limit flux of the Ks map used in Behiri et al. (2023).

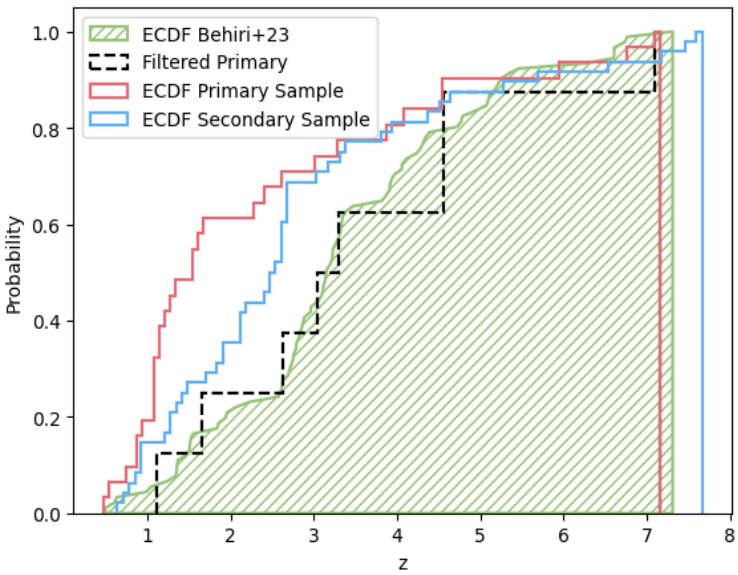


Figure 5.2: *Empirical cumulative distributions for the redshift distributions of sources from the primary sample (red), secondary sample (blue), and Behiri et al. (2023) sample (green). The dashed black line represents the cumulative distribution of the primary sample after filtering to match the depth of the COSMOS Ks map to that in Behiri et al. (2023).*

5.3 SFR from radio emission and AGN contribution

The radio selection of sources is inherently subject to the potential inclusion of AGN given the dual nature of radio emissions, which may originate from both star formation processes and accretion activities. This is highlighted by studies such as Novak et al. (2017), Enia et al. (2022), Gentile et al. (2023), which, within their respective samples of radio-selected galaxies, identified an AGN fraction of approximately 10% to 20%.

The selection criterion for my sample, of galaxies lacking an optical/NIR counterpart, effectively reduces the likelihood of including the most luminous AGN. However, this strategy does not address the challenge posed by obscured AGN, which may elude detection in optical/NIR wavelengths due to dust absorption.

To verify the presence of such obscured AGN within the sample, a multi-wavelength investigation was undertaken. The first phase of this check employed the X-ray data presented in Sec 3.1.5, to determine whether any galaxies could be safely associated with an X-ray source within a 1.5" radius. This examination yielded to no such associations. Subsequently, a detailed visual analysis of radio morphologies was conducted for each source, focusing on the identification of jet-like structures indicative of AGN activity, such as radio lobes or relativistic jets. This step resulted in no sources being excluded based on their radio morphology.

Assuming that my radio sources are galaxies with star formation heavily obscured by dust, is possible to deduce several crucial properties to understand their nature and their role in cosmic evolution. As already described in Sec 2.1, the radio luminosity is a direct indicator of the SFR (Kennicutt Jr & Evans II, 2012) in the absence of significant radio emissions from AGN. In the presence of an AGN, corrections would be needed to differentiate between radio emissions linked to star formation and those originating from the latter. Multicomponent SED-fitting codes that incorporate the contribution of AGN can indicate their possible presence (e.g. SED3FIT, Berta et al. 2013). Unfortunately, in the case of my study, this was not possible due to the limited number of photometric points in the MIR, as these programs look for patterns such as power-law emissions in MIR bands (Donley et al., 2012) which arise from the overlay of black-body emissions from dust at varying temperatures (T $\approx$ 100-500 K) coming from the AGN dusty torus, or complex features characteristic of multicomponent emissions (e.g. Fritz et al. 2006).

Employing the 1.4 GHz flux and assuming a power-law spectrum in the form $S_\nu \sim \nu^\alpha$, where S_ν represents the monochromatic flux and α is the spectral index with $\alpha = -0.7$ generally considered for star-forming galaxies (Novak et al., 2017). Once the cross-correlation with the X-ray data and the check of the radio morphology have indicated that the radio emission is predominantly linked to SFR, I have computed the 1.4 GHz luminosity in the galaxy's rest frame as indicated by Novak et al. (2017).

$$L_{1.4\mathrm{GHz}} = \frac{4\pi D_L^2}{(1+z)^{1+\alpha}} S_\nu \qquad [\mathrm{W\ Hz^{-1}}] \tag{5.1}$$

Where D_L is the luminosity distance at redshift z, $(1+z)^{-(1+\alpha)} = K(z)$ is the K-correction and S_ν is the observed flux at 1.4 HGz.
Once the radio luminosity is known, it is possible to calculate the SFR using the relation from Kennicutt Jr & Evans II (2012):

$$\log SFR = \log L - \log C \qquad [M_\odot \ yr^{-1}] \tag{5.2}$$

where L is the rest-frame luminosity and C is a calibration constant with $\log(C) = 28.20$, the same used in Kennicutt Jr & Evans II (2012).
Similarly, by knowing the maximum depth of the survey and using eq 5.1, one can trace the minimum luminosity as a function of the redshift that a source should exhibit at 1.4 GHz to be detected (indicated by the green dashed line in Fig 5.3). The results presented refer to a Chabrier (2003) IMF, whereas Kennicutt Jr & Evans II (2012) employed a Kroupa (2001) IMF. The transition from one IMF to another was executed following the same conversions used in Madau & Dickinson (2014)[1].
The results of the SFR values calculated following Eq. 5.2 are reported in Table 5.1, and Figures 5.3 and 5.1; values are in the range 10^1-10^4 $M_\odot$ yr^{-1} with the median values for the *primary sample* of SFR $\approx$ 180 $M_\odot$ yr^{-1}, and SFR $\approx$ 700 $M_\odot$ yr^{-1} for the *secondary sample*. Such difference may be a consequence of the redshift distributions, as the latter has a median redshift value much higher than the former, following eq 5.2 this would result in higher SFRs values.
Observed Luminosited values are in the range 10^{22} W $Hz^{-1} < L_{1.4\ GHz} < 10^{25}$ W Hz^{-1}, in line with those of high-redshift Ultra Luminous Infra-Red Galaxies (ULIRGs), as reported by Harrison et al. (2012).
Nonetheless, interesting is the identification of an object whose radio luminosity exceeds the commonly accepted threshold for classification as an FRII/radio-loud galaxy (Fanaroff & Riley, 1974), specifically, greater than 10^{26} W Hz^{-1}. That source is highlighted in Fig 5.3 by a dotted blue square; its estimated IR luminosity, based on radio emissions, is among the highest ($L_{IR} > 10^{13}$ $L_\odot$). This outcome aligns with expectations since MAGPHYS + PHOTO-Z incorporates radio data when available to predict IR luminosity. Nonetheless, this particular source, part of the *secondary sample* (thus not reported in Fig 5.5), not only lacks FIR data but also has no other photometric data beyond the radio emission useful for SED modeling. Specifically, for this source, the analysis has yielded a photometric redshift with a high and narrow probability at z=7. This potentially inaccurate redshift could explain the source's exceptionally high luminosity as calculated through the eq 5.1, and by extension, its derived IR luminosity.

[1]Following Madau & Dickinson (2014), to convert SFRs from a Kroupa (2001) IMF to a Chabrier (2003) IMF, one must multiply the first by a factor 0.96

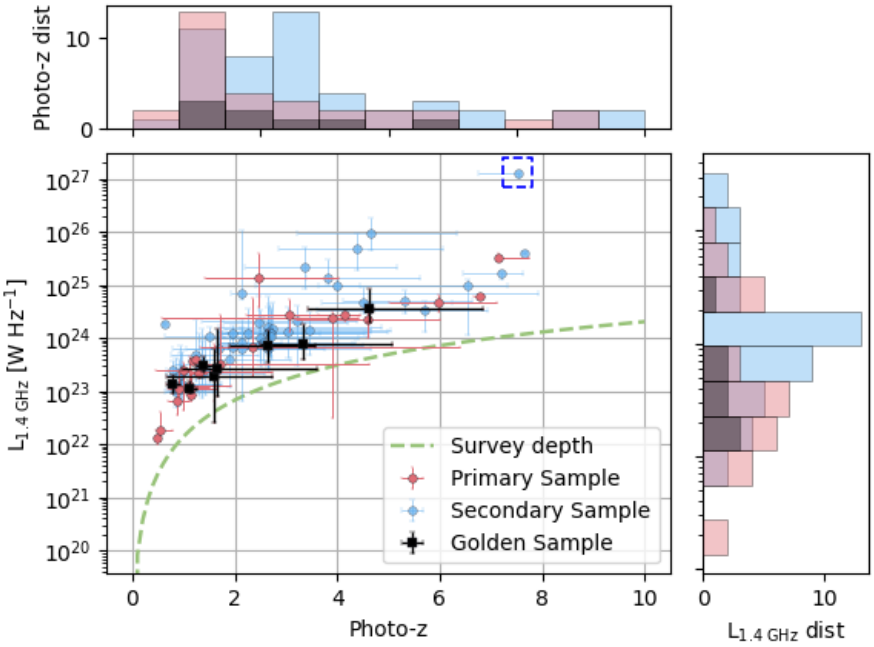


Figure 5.3: *Evolution of the rest-frame luminosity at 1.4 GHz calculated using Eq. 5.2. The minimum survey luminosity given its flux limit is shown in green and, for this plot, is represented by the max depth of the survey (assumed to be constant over the map). Like the other quantities, errors are taken from the 16th and 84th percentiles of the probability distribution of the parameter as reported from MAGPHYS + PHOTO-Z. The single source with $L_{1.4\ GHz} > 10^{26}$ W Hz^{-1} is marked by a blue square.*

5.4 Stellar Masses

The redshift and M_{star} (M_*) distributions can be observed in Fig 5.1, while their median values are reported in Table 5.1. Here, I analyze the evolution of these parameters together and how the two relate to each other.

In Fig 5.4, the M_*-z plane is presented along with the corresponding distributions; as in previous cases, the values reported are the median values from the MAGPHYS + PHOTO-z PDF. The stellar masses range goes from 3.8×10^9 $M_\odot$ to 1.1×10^{13} $M_\odot$, with the 16th and 84th percentiles respectively at 1.2×10^{11} $M_\odot$ and 6.1×10^{11} $M_\odot$.

Behroozi et al. (2013) estimated galaxy stellar masses to range from 10^9 to 10^{11} M$\odot$ at z $\approx$ 2. This aligns my sample's median value with literature expectations.

It is evident from Fig 5.4, that many of the galaxies with mass $M_*>10^{12}$ $M_\odot$ coincide with the high redshift tail, with median values at z $>$ 3 of 3.5×10^{11} $M_\odot$ and 6.0×10^{11} $M_\odot$ for the *primary* and *secondary sample* respectively.

Stellar Masses as high as 10×10^{12} $M_\odot$ at z$>$6 are inconsistent with the studies mentioned above (e.g. Behroozi et al. 2013) and, in extreme cases, challenging to contextualize within

current cosmological models. It is more likely that MAGPHYS+PHOTO-Z overestimates the stellar masses for these galaxies. In the case of the *secondary sample*, such a scenario is supported by the median χ^2_ν of all galaxies with $M_* > 10^{12}$ $M_\odot$ at $z > 3$ of 14.28 and a median number of photometric points in the SED fitting of 2 (non considering the radio detection and the upper limits), thus indicating a poorly sampled SED.

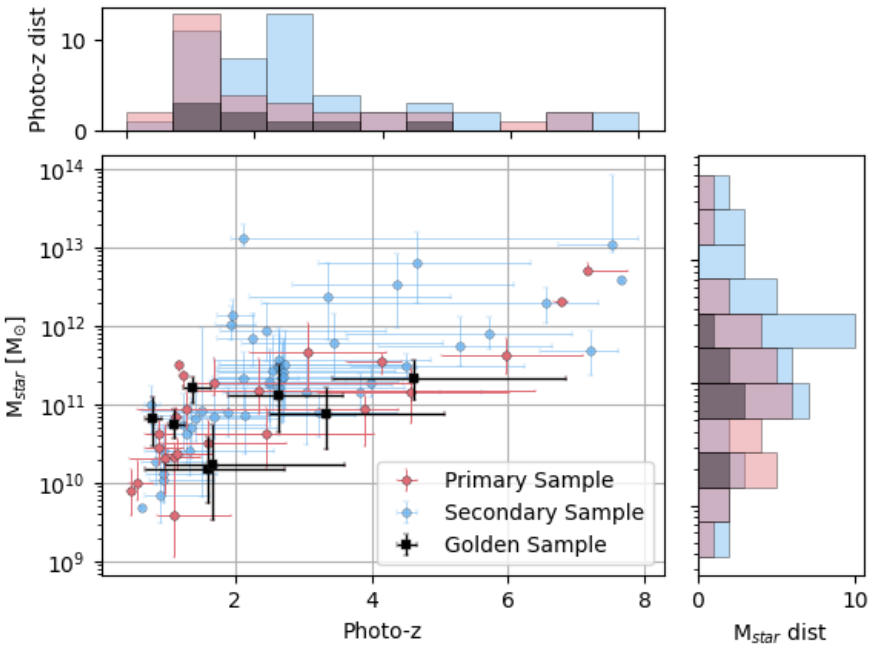


Figure 5.4: *Distribution and corresponding histograms for stellar masses and photometric redshifts. Both results are obtained from the maximum likelihood values from MAGPHYS+PHOTO-Z, with errors derived from the 16th and 84th percentiles of the distributions.*

5.5 Radio-IR correalation

An important relation for star-forming galaxies is that between radio luminosity at 1.4 GHz and infrared luminosity in the range 8-1000 μm (see Fig 5.5), known as the InfraRed-Radio Correlation (IRRC). This correlation is commonly used for estimating distances and temperatures of galaxies emitting in the sub-millimeter at high redshifts, but it also plays a crucial role in calibrating radio luminosities to estimate the SFRs. Despite being an empirical relation, the IRRC shows a remarkably consistent scatter of about 0.16 dex (Molnár et al., 2021).

The distribution of integrated infrared luminosity between 8 and 1000 μm, of the *primary sample* is within a range of $10^{10} < L_{IR} < 10^{13}$ $L_\odot$, with a median value $4.37{\times}10^{11}$ $L_\odot$. Luminosities exceeding $L_{IR} > 10^{12}$ $L_\odot$ are typically associated with ULIRGs, and interest-

ingly, about 28% of the *primary sample*, amounting to 10 candidates, qualify as such.
As already outlined in Sec 2.1 the infrared luminosity originates from the thermal re-emission of dust grains heated by the absorption of opt/UV photons from young, massive stars. However, the exact nature of how infrared emission is related to radio emission, as well as how the IRRC evolves with redshift, remains unclear (e.g. Delhaize et al. 2017).
This correlation is usually measured through the q_{tir} parameter defined as (e.g. Yun et al. 2001):

$$q_{tir} = \log\left(\frac{L_{\rm IR}}{3.75 \times 10^{12}\,{\rm Hz}}\right) - \log\left(\frac{L_{1.4\,{\rm GHz}}}{\rm W\ Hz^{-1}}\right) \tag{5.3}$$

In the local universe, the calculated mean value of q_{tir} is 2.64 ± 0.02 (Bell et al., 2003) while for my *primary sample*, the median q_{tir} value is 2.20 ± 0.045[2].
The evolution of q_{tir} as a function of cosmic time is still a subject of ongoing research; Delhaize et al. (2017) in their VLA-COSMOS study, identified the following relation for the evolution of q_{tir}:

$$q_{tir} = (2.88 \pm 0.03)(1+z)^{-0.19\pm0.01} \tag{5.4}$$

Similar to that in Magnelli et al. (2015) and Calistro Rivera et al. (2017), with the latter focusing on radio-selected star-forming galaxies.
The q_{tir} equation is a useful tool for identifying potential AGN by detecting an excess in radio luminosity (see Delhaize et al. 2017; Delvecchio et al. 2017). I defined possible AGN as sources more distant than 2σ from the relation 5.4 based on their calculated radio luminosity using eq 5.2. Out of the 36 sources, none were marked as a potential AGN.

5.6 The Star Formation Main Sequence

To improve the characterization of my NIR-Dark galaxies, I constructed the M_*-SFR plane (previously introduced in Sec 2.1.1). Various studies have already demonstrated a tight correlation between these two quantities (see Speagle et al. 2014; Saintonge & Catinella 2022; Popesso et al. 2023) and how this relation evolves over cosmic time. Therefore, understanding how galaxies in this study position themselves within the M_*-SFR plane can significantly assist in understanding the ongoing star formation regime and to better understand whether they are normal star-forming galaxies or starbursts.
In this context, I have opted to employ the relation describing the evolution of the Star Formation Main Sequence as found by Speagle et al. (2014), based on a Kroupa (2001) IMF but here corrected for a Chabrier (2003) IMF:

$$\log{\rm SFR}(M_*, t) = (0.84 - 0.026 \cdot t)\log M_* - (6.51 - 0.11 \cdot t) \tag{5.5}$$

[2]Uncertanties for the median calculated using the Median Absolute Deviation (Hoaglin et al., 1983)

where t represents the age of the universe in Gyr, and M_* is the stellar mass in units of $M_\odot$. The results are shown in Figure 5.7, where Equation 5.5 is illustrated by the solid green line. The intrinsic scatter of the SFMS of approximately 0.2 dex has also been highlighted (Speagle et al., 2014). Additionally, the starburst line (dashed green line) is depicted at four times the values of the Main Sequence (e.g. Rodighiero et al. 2011).
I divided the sample into four redshift bins (see Tab 5.2, 5.3); the SFR values presented are derived from Equation 5.2. However, a key factor in obtaining accurate estimates is the determination of the source's redshift, which, for the *secondary sample*, could be significantly affected by systematic errors due to the lack of FIR constraints.
Across each redshift bin, are present galaxies exceeding the starburst line.

Table 5.2: *Median values of SFRs $[M_\odot\ yr^{-1}]$ for the three sub-samples across different redshift. Uncertainties are computed with the MAD as for Tab 5.1.*

z bins	Primary sample	Secondary sample	Golden sample
$z < 1$	52.5 ± 20.4	80.8 ± 13.1	78.4 ± 0.0
$1 < z < 2.5$	139.6 ± 25.8	394.7 ± 80.0	136.5 ± 24.2
$2.5 < z < 4.5$	1452.7 ± 118.8	788.4 ± 73.0	444.2 ± 7.3
$z > 4.5$	2739.3 ± 663.5	7752.0 ± 2827.3	2174.4 ± 0.0

Table 5.3: *Median Stellar Mass $[M_\odot]$ for the three sub-samples across different redshift. Uncertainties are computed with the MAD as for Tab 5.1.*

z bins	Primary sample	Secondary sample	Golden sample
$z < 1$	$(2.5 \pm 0.9) \times 10^{10}$	$(1.3 \pm 0.3) \times 10^{10}$	6.6×10^{10}
$1 < z < 2.5$	$(5.6 \pm 1.5) \times 10^{10}$	$(8.1 \pm 1.7) \times 10^{10}$	$(3.7 \pm 1.5) \times 10^{10}$
$2.5 < z < 4.5$	$(1.3 \pm 0.4) \times 10^{11}$	$(2.6 \pm 0.3) \times 10^{11}$	$(1.0 \pm 0.3) \times 10^{11}$
$z > 4.5$	$(4.2 \pm 1.8) \times 10^{11}$	$(1.4 \pm 0.5) \times 10^{12}$	2.1×10^{11}

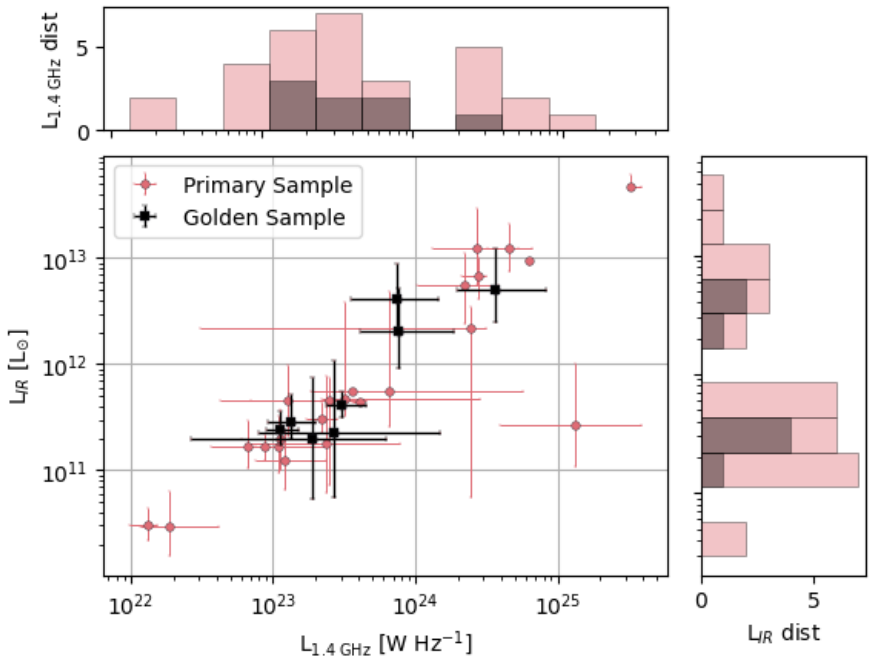


Figure 5.5: *IRRC for the primary (red) and golden samples (black) of galaxies as in the analysis of infrared luminosities and other dust properties the secondary sample has been excluded. As for the other results, best-fit values are taken from the maximum likelihood values from MAGPHYS+PHOTO-Z while errors are represented by the 16th and 84th percentiles.*

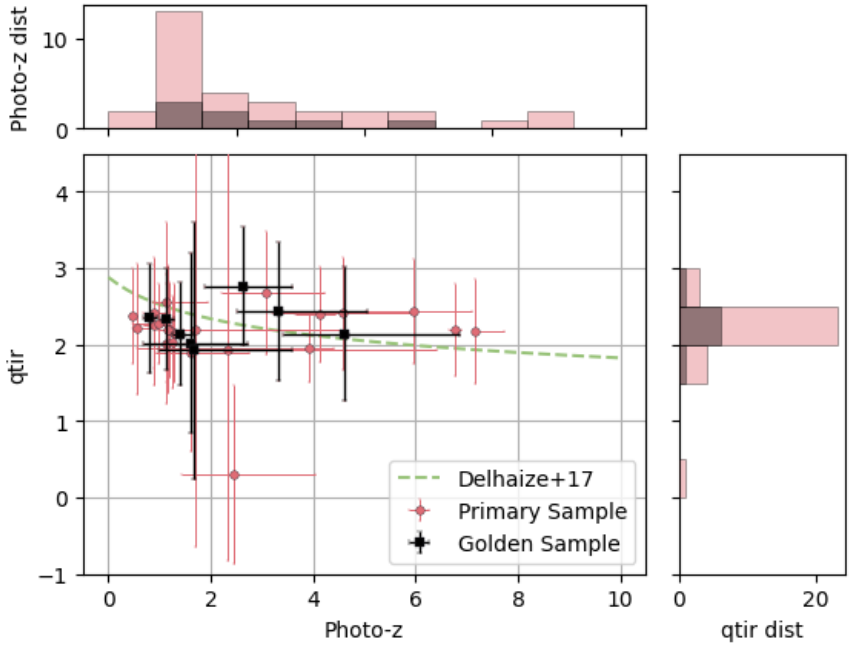


Figure 5.6: *q_{tir} parameter as a function of redshift including best-fit values and errors for the primary and golden samples (as the secondary sample has been excluded in the analysis of dust properties). An estimated fit from Delhaize et al. (2017) (eq 5.4) has also been added represented by the green dashed line.*

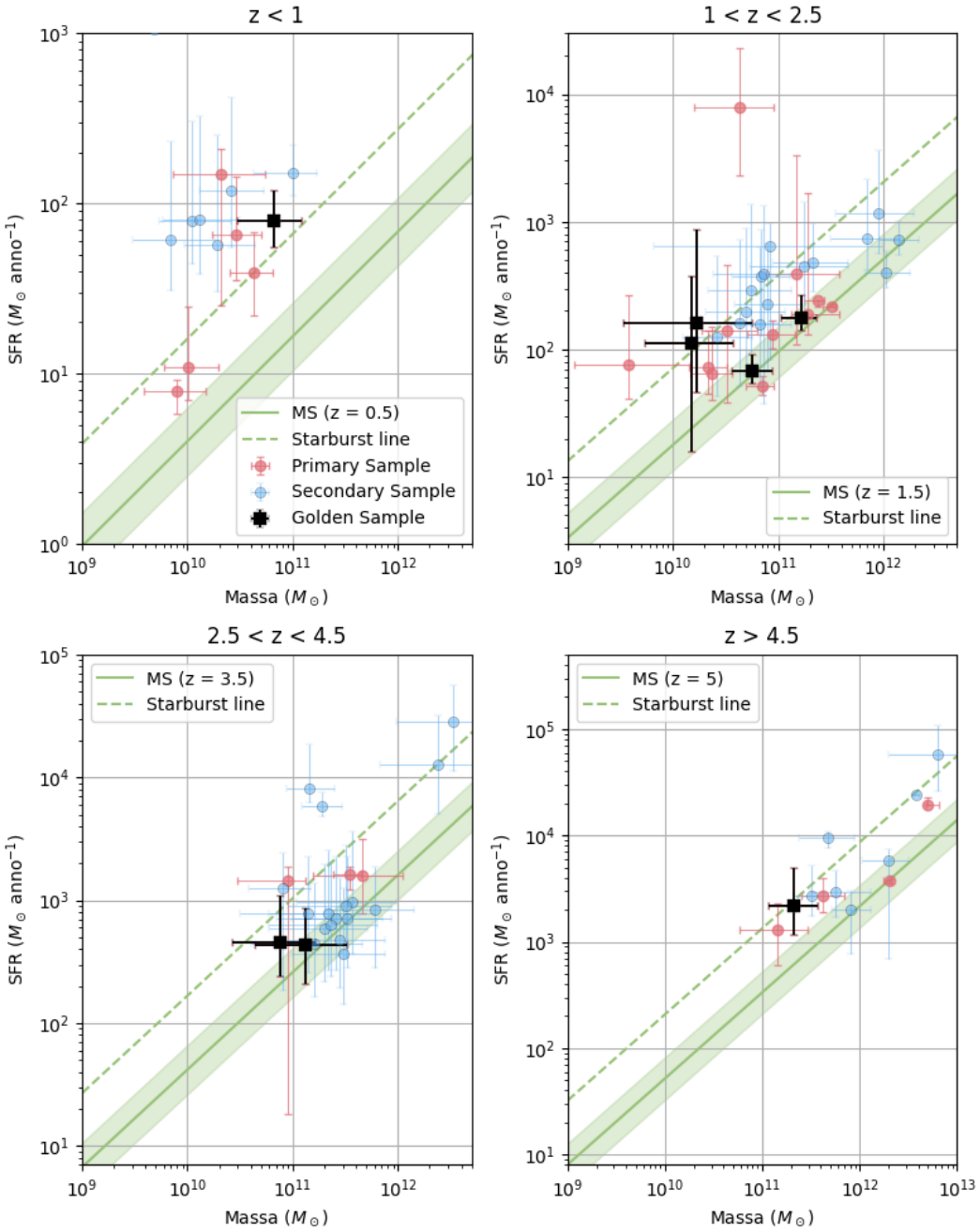


Figure 5.7: *SFMS in the four redshift bins (Tab 5.2), following the color scheme proposed in the other plots, the secondary sample is in red, the primary sample in blue, and the golden sample is highlighted in black. The SFMS and its intrinsic scatter of 0.2 dex are shown in solid green and follow Speagle et al. (2014). The starburst line, defined as four times the main sequence, is illustrated in dashed green.*

5.7 Contribution to the SFRD

One of the primary goals of this work is to determine the relative contribution of radio-selected NIR-Dark galaxies to the SFRD. In this analysis, only the *golden sample* was used, having the best-constrained z-phot and SFR, under the assumption that is representative of the whole population. For the SFRD and source number densities, the $1/V_{max}$ method (Schmidt, 1968) is employed. This nonparametric method determines the density of sources based on their maximum observable comoving volume V_{max}, thereby eliminating dependences on any prior analytical assumptions.
The selected redshift bins are the three in the interval $0 < z < 4.5$ (see Tab: 5.4), for each bin the V_{max} has been computed as:

$$V_{\max} = \sum_{z=z_{\min}}^{z_{\max}} [V(z + \Delta z) - V(z)] \cdot \frac{C_A}{C_I} \tag{5.6}$$

Where the sum is performed over spherical shells of comoving volume separated by a redshift step of $\Delta z = 0.05$ between z_{min}, the lower boundary of the considered redshift bin, and z_{max}, the smaller value between the upper boundary of the bin and the maximum redshift at which the source would be observed given its radio luminosity (eq 5.2) and the maximum depth of the survey.
The corrective terms C_A & C_I in eq 5.6 account for the effective area of the observation and the incompleteness of the *golden Sample* relative to the *primary*:

$$C_A = \frac{A_{\text{eff}}}{41253\,\text{deg}^2} \tag{5.7}$$

Where A_{eff} represents the effective area of the SPIRE observations, $A_{\text{eff}} = 0.048$ deg^2, and 41253 deg^2 is the total area of the sky. The C_I parameter accounts for the incompleteness of the sample and is defined as the ratio between the number of sources that are in the *golden sample* and the total number of sources in the *primary sample*:. Here I assume that the global properties of the *golden sample* reflect those of the *primary sample* both in terms of redshift distribution and SFRs.

$$C_I = \frac{N_{golden}}{N_{primary}} = \frac{8}{36} \tag{5.8}$$

The SFRD in each bin is then calculated as:

$$SFRD = \sum_{i \in bin} \frac{SFR_i}{V_{max,i}} \qquad [\text{M}_\odot\ \text{yr}^{-1}\ \text{Mpc}^{-3}] \tag{5.9}$$

The values obtained for each bin are presented in Fig 5.8 and Tab 5.4, they were derived by taking the median over 1000 realizations drawing random samples from the redshift probability distributions of each galaxy, while the lower and upper limits were determined from

the 16th and 84th percentiles of the final distribution, respectively.
The SFRD values for z>1 found in this study align with those reported by Talia et al. (2021), Behiri et al. (2023) for a radio-selected sample in the COSMOS field, as well as the sample of NIR-Dark SMGs from Gruppioni et al. (2020) who worked in the ALPINE fields. My galaxies contribute between 3% and 13%[3] to the SFRD as calculated by Madau & Dickinson (2014). However, it is important to note that the SFRD values computed in this research represent lower bounds to the true values because I did not extrapolate to lower luminosities. Such an extrapolation could otherwise be achieved by constructing a luminosity function, a step not feasible in this instance due to the limited statistical sample available.

Table 5.4: *SFRD values for each redshift bin over 1000 realizations. Uncertainties are calculated as the 16th and 84th percentiles of the distribution. The last column indicates the percentage ratio between the predicted SFRD values from Madau & Dickinson (2014) relation and my median values at the same redshift.*

z bin	**SFRD** [$M_\odot$ yr^{-1} Mpc^{-3}]	**Number Density** [Mpc^{-3}]	% Madau
0-1	$2.03^{+0.77}_{-0.92} \times 10^{-3}$	$4.75^{+2.37}_{-2.37} \times 10^{-5}$	7.0 %
1-2.5	$2.64^{+1.18}_{-1.23} \times 10^{-3}$	$1.74^{+0.61}_{-0.67} \times 10^{-5}$	3.25 %
2.5-4.5	$2.80^{+7.07}_{-1.02} \times 10^{-3}$	$7.85^{+3.92}_{-3.91} \times 10^{-6}$	13.3 %
>4.5			

5.7.1 DSFGs as possible progenitors for quiescent Galaxies

Why are DSFGs so important? As pointed out by Toft et al. (2014), the number density of passive quiescent galaxies with $M > 10^{11}$ $M_\odot$ at $z \approx 2$ is of the order of $n \approx 10^{-5}$ Mpc^{-3}, which is 10 to 100 times greater than that of LBGs with similar mass. Studies on SMGs identify in this galaxy population the likely progenitors of the quiescent galaxies in question. However, since the definition of SMGs is purely observational, often the selection criteria need to be corrected for their duty cycle, as the SMG phase is likely short-lived. DSFGs, at least in part, may belong to the same population as SMGs but can be observed also when the starburst phase is not at its peak. Consequently, DSFGs can represent the missing link to explain the observed population of massive quiescent galaxies. The number density found for my galaxies at $z > 3$ is $n = 8.20^{+0.57}_{-4.10} \times 10^{-6}$ Mpc^{-3}, which is lower than previous results of radio-selected DSFGS of $n \approx 10^{-5}$ Mpc^{-3}(Talia et al. 2021; Behiri et al. 2023; Gentile et al. 2024). However, this discrepancy might be a consequence of bias effects due to cosmic variance given the small angular dimension of the analyzed area (0.05 deg^2 compared to 1.38 deg^2 for Talia et al. 2021 and Behiri et al. 2023). In any case, even the lowest estimate

[3]The values are computed by averaging those calculated using the relation from Madau & Dickinson (2014) in the same redshift bins.

is still an order of magnitude greater than those predicted by semi-analytical models or simulations for massive ($\log(\frac{M_*}{M_\odot}) > 10.5$) dusty galaxies (see Valentino et al. 2020). Interestingly, the numerical density values at z≈3 are similar to those of passive quiescent galaxies identified at z≈2, been both of the order of 10^{-5} Mpc^{-3} (see also Straatman et al. 2014; Schreiber et al. 2018). Unfortunately, my sample does not extend to high enough redshifts to comment on the population of quiescent galaxies at z > 3, whose progenitors should be identified at z > 4-5.

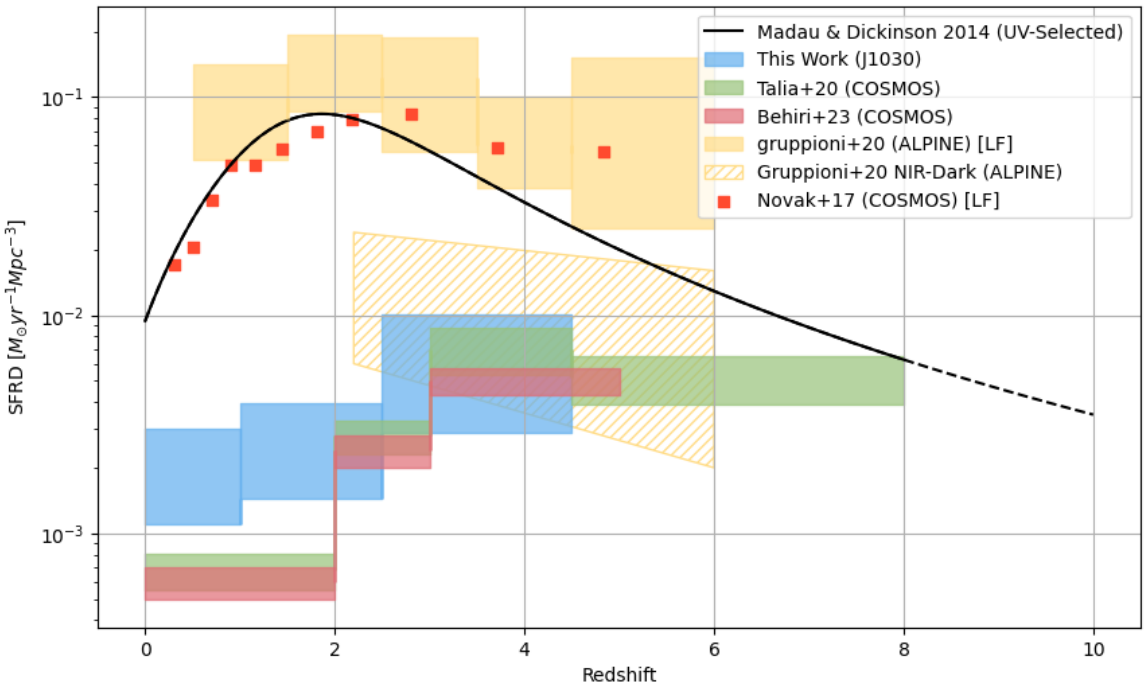


Figure 5.8: *SFRD values found in this study (blue) and others. Values in the legend marked with 'LF' mean that the results from that study have been obtained via an integration of a luminosity function. In yellow results from Gruppioni et al. (2020) who identied NIR-dark SMGs in the ALPINE survey. In green and red respectively Talia et al. (2021) and Behiri et al. (2023) with their samples of radio-selected DSFGs in the COSMOS field. The red squares are from Novak et al. (2017) with their sample of radio-selected star-forming galaxies in the COSMOS field. In black the SFRD evolution relation as found by Madau & Dickinson (2014), the dotted part of the line indicates an extrapolation of the relation.*

Chapter 6

Conclusions and Future Perspectives

This thesis focuses on the study of a population of radio-selected DSFGs within the J1030 field and their contribution to the SFRD as well as their possible role as progenitors of quiescent galaxies at z = 2. Below, the main steps and results of this work are summarized:

- Starting from the radio catalog (D'Amato et al., 2022) and the multiband catalog of J1030, I selected 79 galaxies with a radio flux at 1.4 GHz $>$ 12.5 μJy and absence of NIR-Ks counterparts at a limit flux at 2.2 μm of 0.36 μJy. The initial sample was divided into two main subsamples based on their coverage in FIR/sub-mm bands (*primary* and *secondary sample*).

- I adapted the PhoEBO pipeline (Gentile et al., 2023) and used it for accurate photometric measurements for all the sources, including the blended ones, up to 160 μm.

- I designed and applied a deblending procedure of the FIR/sub-mm bands from 250 to 500 μm, based on that described in Jin et al. (2018) and Liu et al. (2018) and on SED fitting.

- I characterized the two main samples using SED-fitting with the code MAGPHYS+PHOTO-Z (Battisti et al., 2019), which allowed for the extraction of the main physical characteristics and photometric redshifts. Further screening of the *primary sample* led to the construction of the *golden sample*, containing galaxies with the strongest observational constraints and the most robustly estimated parameters.

- Finally, I studied the contribution of this population of galaxies to the cosmic SFRD by applying the $1/V_{max}$ non-parametric technique. The results suggest a lower limit contribution to the SFRD, as calculated by Madau & Dickinson (2014), from 3% at z $\approx$ 2 to 13% at z $\approx$ 3, highlighting how, so far, opt/UV survey may have missed a crucial galaxy population.

- The number densities of the sources analyzed are in fair agreement with similar studies (Talia et al. 2021, Behiri et al. 2023, Wang et al. 2019), suggesting that DSFGs

may, indeed, be the possible missing link to explain the presence of massive quiescent galaxies observed at $z > 2$. Still, the high number of DSFGs found can not be currently explained by models/simulations.

The role of future observations

A key role in the future developments of this work will be played by groundbreaking observations. The adopted methodology for identifying DSFG galaxies by using the synergy of radio and NIR data has proven effective on different scales and holds potential for future missions like the NIR satellite Euclid[1], launched in July of 2023, and the future radio interferometer SKA[2] (Square Kilometre Array). These studies will explore a much larger area of the sky than this study has.

Moreover, to achieve a more accurate characterization of the identified galaxies, the use of state-of-the-art instruments like ALMA is crucial. Operating in the mm and sub-mm spectrum, ALMA is perfect for adding essential photometric points to this as well as future studies, allowing to accurately constrain dust properties such as mass, temperature, and luminosity. Moreover, through spectroscopic follow-ups, ALMA can determine the redshift of dusty galaxies up to $z = 6$ and beyond, for example, by observing the CO and/or C+ transitions.

Finally, another facility that will play another key role in the characterization of DSFGs is the James Webb Space Telescope[3] (JWST), whose unprecedented sensitivity proved enough to observe these "dark" galaxies that would otherwise remain hidden in the NIR. Its capabilities for deep photometric and spectroscopic observation are crucial for imposing more stringent constraints on the stellar masses of DSFGs, potentially unlocking the secrets of DSFGs and their pivotal role in the cosmic saga.

[1] Euclid mission Web page
[2] Square Kilometer Array Web page
[3] James Webb Space Telescope Web page

Appendix A

PhoEBO Documentation

This appendix focuses on the desciption of PhoEBO, a pipeline developed by Gentile et al. (2023) for the deblending of NIR-Dark galaxies, and updated in this work.

A.1 Main structure

PhoEBO is a Python-based pipeline primarily oriented toward the object-oriented paradigm. This means that data are internally contained within user-defined "classes" and all subsequent steps are invoked through their associated functions, known as methods. This approach has been adopted to ensure the code's maximum flexibility, ease of use, and ease of update.
The main class is '**Image**,' and the entire process of analysis and optimization goes through the methods associated with it. However, to provide a user experience that is as convenient and intuitive as possible, PhoEBO incorporates two subclasses ('**Band** and **Target**) that handle the process of data preparation.

A.1.1 Class: Band

The 'Band' class contains the necessary information for PhoEBO to perform operations on the detection maps[1], such as creating segmentation maps, generating convolution kernels, and managing the World Coordinate System (WCS)[2] for source coordinates.
Similar to a standard Python class, the 'Band' class requires specific 'Keys' that must be specified by the user.
The full list of keys for this class, both mandatory and not, are listed below

- file-path The system path to the scientific map file.
- error-path The system path to the weight map file.

[1]The detection maps are the radio 1.4 GHz and the Ks images
[2]World Coordinate Systems (WCSs) are used to convert from one set of coordinates to another through geometric transformations, like to align the pixels in an image with positions on the celestial sphere.

- px-scale The pixel scale in arcseconds per pixel of the scientific map.
- ra, dec The position of the target source, expressed in RA-DEC J2000 coordinates.
- psf-file The system path to the PSF file (if available) if '.fit' format.
  Note: PhoEBO can also generate PSFs assuming a Gaussian profile if the user does not specify the 'psf-file' key, making it optional. In this case, the Gaussian PSF constructed has an FWHM equal to the 'fwhm' key.
- psf-scale The pixel scale at which the PSF is registered (this allows to work also with oversampled PSFs).
  Note: this is necessary only if the user has specified a PSF file, with the default value being equal to the pixel scale of the scientific map.
- fwhm A Python list containing information about the full width at half maximum (FWHM) of the Point Spread Function (PSF) for both major and minor axes, with units expressed in arcseconds.
- psf-theta Specifies the angle, in degrees counterclockwise, of the PSF. This is optional, with the default value set to 0 deg.
- size The size in arcseconds of the region over which PhoEBO will operate.
- size-align The size of the region used by PhoEBO for image alignment.To be specified only if the user wants PhoEBO to try to align the detection and the target image.
- wcs-radio A boolean variable that specifies if the radio map has an axis for the polarization info or not.
- show An optional key that can be used to display on-screen the regions that will subsequently be used by PhoEBO for data analysis. This may be useful for initial data exploration.

Below are two examples of 'Band' class usage directly taken from my work. The first example includes information regarding the Ks-band detection map, while the second refers to the radio-band detection map. It also demonstrates the usage of optional keys, such as 'psf-file.'

```
Ks =  Band(
            file_path = 'path to Ks map',
            error_path = 'path to Ks error map',
            px_scale = 0.304,
            ra = ra,
            dec  = dec,
            fwhm = [0.45,0.45],
```

```
            psf_file = 'path to psf fit file',
            psf_scale = 0.304,
            show = False,
            size = (25,25),
            size_align = (60,60)
            )
```

```
radio = Band(
            file_path = 'path to radio map',
            error_path = 'path to radio error map',
            px_scale = 0.3,
            ra = ra,
            dec  = dec,
            fwhm=[1.45,1.15],
            psf_theta = 30,
            show = False,
            size = (25,25),
            wcs_radio = True,
            size_align = (60,60)
            )
```

A.1.2 Class: Target

The 'Target' class inherits all the previously mentioned information from the 'Band' class and adds further details useful for the deblending procedure, including:

- zp The observation Zero-Point of the target map.
- channel This key is used for naming the folders where generated files will be saved during the process. Setting the 'channel' key to one of the following: ch1, ch2, ch3, ch4 automates the process of calculating aperture corrections (if specified) for the IRAC bands using the user-specified aperture radii and angle values. If no aperture corrections are needed, changing the 'channel' name to any desired string will suffice. The name of the folder will also match the user-specified channel name, where PhoEBO saves the results.

```
irac = Target(
            file_path = 'path to target image map',
            error_path =   'path to target rms map',
            zp = 18.5263,
            px_scale = 0.6,
```

```
6               ra = ra,
7               dec = dec,
8               channel='ch3',
9               fwhm = [1.88, 1.88],
10              psf_file = 'path to target psf file',
11              psf_scale=0.305
12              show=False,
13              size = (60,60),
14              size_align = (60,60)
15              )
```

A.1.3 Class: Image

The 'Image' class and its respective methods are the heart of PhoEBO. It takes the information and data prepared previously by the 'Target' and 'Band' classes and performs the user-specified analyses.
To continue the example of using PhoEBO, after defining the classes 'irac,' 'radio,' and 'Ks' with their respective information, they are passed to the 'Image' class, which requires the following keys:

- target An instance of the 'Target' class.
- detection-radio An instance of the 'Band' class, containing the first of the two necessary detection maps. In this case the radio map.
- detection-Ks An instance of the 'Band' class, containing the second of the two detection maps. In this case the Ks map.
- alligment Specifies whether PhoEBO should attempt astrometric alignment of the detection maps with the target.
- prior With the 'prior' key, it is possible to provide PhoEBO in advance with a list of source positions. This can be useful for expediting calculations (skipping the subsequent convolution phase for all sources not belonging to the prior list) and for focusing optimization efforts on specific targets. The required input is a path to a file (the format can be FITS, ASCII, or CSV) with three columns: the first two columns represent 'RA' and 'DEC,' which are the right ascension and declination of the sources, respectively. The third column is labeled 'flag,' where 1 indicates that the source should not be considered, while 0 signifies that the source should be considered for optimization. An output file of sources optimized by PhoEBO can be used as a prior file for subsequent optimizations.
- n-core Set the number of CPU cores to use. If is not specified, the default value is set to 1, indicating that it will run the calculations on a single core.

In addition to these mandatory keys, the 'Image' class also includes entirely optional keys that allow the user to have better control over the program work-flow:

- cmap Specifies the color map that PhoEBO will use for all plots.
- id If provided, the ID specified by the user in this key will be used for naming the output files.
- show This key is used for displaying all the maps on which PhoEBO will perform subsequent analyses. If the 'alignment' key is set to 'True,' it will also display the residual transformation map if such transformation is found.

```
test = Image(
            target=irac,
            detection_radio=radio,
            detection_Ks=Ks,
            id=ID,
            alligment=True,
            n_core = 4,
            show=False,
            cmap='magma'
            )
```

Methods

The fundamental steps of data analysis are carried out through the use of the methods associated with the 'Image' class.
In its current implementation, the order in which these methods are called through their respective functions is one and only one, and it is consistent with the order presented in this guide.

Method 1: generate-kernels

This function generates the convolution kernels required for the convolution phase. Firstly, all the PSFs are rescaled to the same pixel scale as the contaminants detection map (in this example the Ks band map) while preserving their angular dimensions. This process generates arrays of different sizes. The next step involves cropping these PSFs to obtain final arrays, all of the same dimensions, and registered to the same pixel scale.
The prepared PSFs are used to create matching kernels, which are then normalized and cropped, ensuring that the geometric center of the PSF is positioned at the array's center.
The 'create-kernels' method has the following keys:

- alpha: The alpha parameter that describes the width of the window function used during the creation of the kernels[3].
- show: A key to display the original PSFs and their corresponding matching kernels on-screen.

Here is an example:

```
test.create_kernels(
                        alpha=0.35,
                        show=False
                        )
```

Method 2: generate-seg-maps

This method generates segmentation maps from the contaminants detection and source detection provided to the 'Image' class. For the segmentation map generation, PhoEBO utilizes the Python implementation of the 'Source Extractor' program (SEP, Barbary 2016). The segmentation maps are processed differently based on whether they are derived from the radio detection map or Ks detection maps. In the first case, only the central source in the map is considered (represented by the target source), and the others are discarded. In the second case, the segmentation map is decomposed into a data cube, where each slice contains the binary mask of individual contaminants.
From the data cube, all sources that extend beyond the limits specified by the user in the 'border' key are removed.
The keys for this method are as follows:

- threshold: Specifies the absolute treshold for source detection to pass to Source-Extractor.
- npix: Specifies the minimum number of contiguous pixels to consider a detection as a source to pass to Source-Extractor.
- border: Specifies the dimension (in percentage) of the edge relative to the detection map beyond which a source is discarded. For example, a value of 10 means that all sources within a distance less than 10% of the size of the x-axis of the detection map will be discarded.
- threshold-radio: The number of standard deviations beyond which a pixel is considered a detection or not in the radio map to pass to Source-Extractor. This key is optional, with the default value set to 5.

[3] PSF matching in Photutils

- show: A key to display the results of this step on-screen.

```
test.generate_seg_map(
                      treshold = 2,
                      npix = 5,
                      treshold-radio = 5,
                      show=False,
                      border=1
                      )
```

Method 3: generate-models

This method employs the isolated contaminants obtained from the 'generate-seg-map' step and convolves them with the kernels generated in the 'create-kernels' step. Subsequently, after convolution, the contaminants are reprojected into the frame of the 'target' image. This 'Image' class method does not possess any user-specified keys.

```
test.generate_model()
```

Method 4: optimize-models

This function represents the core of PhoEBO, where all results from previous steps are used for deblending the 'target.'

- sigma: Represents the absolute treshold beyond which PhoEBO will search for residual sources in the 'target' map in subsequent iterations.
- n-iter: Specifies the maximum number of iterations PhoEBO should perform during optimization.
- exclude-center: This key allows the user to specify whether to consider the optimization of sources located at a certain distance from the center of the target image.
- group-radius: This key indicates the maximum separation distance in units of FWHM at which two or more sources are considered isolated. Non-isolated sources are optimized together. The default value for this parameter is 2.5 times the provided Full Width at Half Maximum (FWHM).

- max-shift: Specify if PhoEBO should fit the centroid as a free parameter for each source. The units are expressed in pixels and should contain the negative shift and positive one, where a positive shift is in the direction right/top while the negative is left/bottom. If there's no desire to attempt fitting the center of the source the maximum shift should be set to [0,0].

Example Usage:

```
test.optimize_model(
                    sigma = 3,
                    n_iter = 3,
                    exclude_center = False,
                    group_radius = 2,
                    max-shift = [-1,1]
                    )
```

Method 5: extract-flux

The final step involves extracting photometry from the residual map. This is the only step in PhoEBO that returns results to the user. These results include:
1. AB magnitude and associated error.
2. Flux values and flux error measured from the optimized map, reported in μJy.

The keys that the user can specify in this step are:

- radius: A list of three values, respectively representing the central aperture radius and the inner and outer radii of the annulus for background in arcseconds.
- correction: Allows the user to choose whether to apply aperture corrections and, if so, which multiplicative factor to use.
- save-cont: Allows the user to select whether to save a file containing information obtained from the optimization of contaminants. The file includes columns for various data, such as position, flux, magnitude, and flags for each source.
- name-cont: An optional parameter to specify the name of the contaminant file.
- show: A key to plot apertures and flux values on-screen together with the residual map.

Example Usage:

```
table = test.extract_flux(
                          radius = [3, 6, 10],
                          show = False,
                          correction = None,
                          save_cont = True,
                          name_cont = 'cont'
                          )
```

```
RA              # The position of contaminants expressed in RA J2000
DEC             # The position of contaminants expressed in DEC J2000
x               # The position of contaminants expressed in pixels on the
                x-axis of the target map
y               # The position of contaminants expressed in pixels on the
                y-axis of the target map
flux-native     # Best-fit flux in native units of the map
mag             # Best-fit AB magnitude
flux            # Best-fit flux in uJy
flux_err        # Error on the best-fit flux in uJy
mag_err         # Error on the magnitude
flag            # Source flag (0 = accepted,
                              1 = S/N < 1,
                              3 = source & prior discarded (can happen because the source
                              has bad pixels in it),
                              5 = prior discarded (can happen because no sources where
                              associated to that prior),
                              8 = prior is outside the analysis region)
```

A.1.4 Plotting Results

In its current implementation, PhoEBO includes a standard routine for result visualization, which is controlled through the 'plot-results' method. This method automatically generates and saves two PNG images that are useful for obtaining an immediate overview of the results provided by the optimization process.
Like all the methods described so far, below is an example of how to use the method:

```
test.plot_results(show = False,
                  cmap = 'magma',
                  save_fits = True,
                  fits_name = 'psf',
                  interpolation = None)
```

Tips and Tricks

Here are some recommendations for the proper functioning of PhoEBO.

1. The dimensions of the analysis regions and those dedicated to image alignment must match, otherwise the code will raise an error. This is necessary because many of PhoEBO's internal operations are linked to arrays (pixels) operations, and arrays of different dimensions do not overlap, leading to errors.

2. The central coordinates of the analysis and alignment regions must be the same for Band and Target (as specified by the KEYs ra, dec).
   It is possible that, if the target map is significantly smaller than the detection map, it may not be possible to extract an analysis region. In this case, an error will occur in the pipeline.

3. PhoEBO excels in bands typically affected by broad Point Spread Functions (PSF), such as mid-IR and FIR. This is because PhoEBO operates under the crucial (yet reasonable) assumption that the apparent modification to the intrinsic shape of the image due to the change in the observational band is smaller than that generated by the change in PSF. As a result, PhoEBO assumes that the intrinsic shape of the source in the target band is the same as that recorded in the detection band. This allows for the fitting of extended sources and those with complex profiles.

Appendix B

PhoEBO; Differences with the original code

During my thesis research, I had to adapt PhoEBO for uses beyond its original design, specifically to analyze previously unexplored spectral bands. This effort was driven by the nature of my data and by the will to test PhoEBO's versatility in new analytical scenarios. To achieve this, I integrated novel functionalities into the pipeline's existing codebase and developed supplementary modules to facilitate the analysis in these additional spectral bands.

1. **Implementation of Band and Target Classes**: In its original version, PhoEBO required that data be prepared and organized in a specific manner: for each source, a folder named like the target source's ID was needed, and cutouts centered on the source had to be saved within this folder, one for each band, in a precise order, and with correct information included in the header. This procedure, which had to be carried out by the user, was prone to errors, and often difficult to identify. To simplify data management and reduce the likelihood of errors, two classes (Band and Target) have been implemented. Their primary function is to prepare the data while requiring as few inputs as possible from the user. Moreover, this approach offers greater flexibility, allowing for the change of various parameters for different sources.

2. **Extension over longer wavelengths**: One of the most significant enhancements introduced is PhoEBO's ability to operate at wavelengths beyond those for which it was originally designed, extending its capabilities from IRAC bands to PACS bands. This advancement was made possible by modifications to PhoEBO's optimization engine, which now allows the pipeline to use priors based on the positions of the sources. This was achieved by integrating information from various spectral bands, including the detection band, significantly improving the versatility and effectiveness of the analysis.

3. **source fitting**: To enable PhoEBO to optimize sources in large analysis areas, which can contain dozens or even hundreds of sources, it was necessary to implement a different optimization logic than the original one. The previous strategy involved the simultaneous fitting of all sources within the analysis area. This approach is manageable when the number of models and, consequently, free parameters is limited, thus facilitating the convergence of the optimization process. However, optimizing a large number of models in parallel significantly increases the number of free parameters (such as the multiplicative factors of the models' flux and the position of the centroids of the membership groups), making it difficult to explore such a vast parametric space and increasing the risk of parameter degeneration. The solution identified to overcome this challenge was relatively simple: instead of optimizing all sources at the same time, the process was divided, distinguishing which sources could be adapted individually and which required joint optimization, thus forming groups. This segmentation was made possible by analyzing the morphological characteristics of the sources identified as priors and the information on the PSF size of the target band.

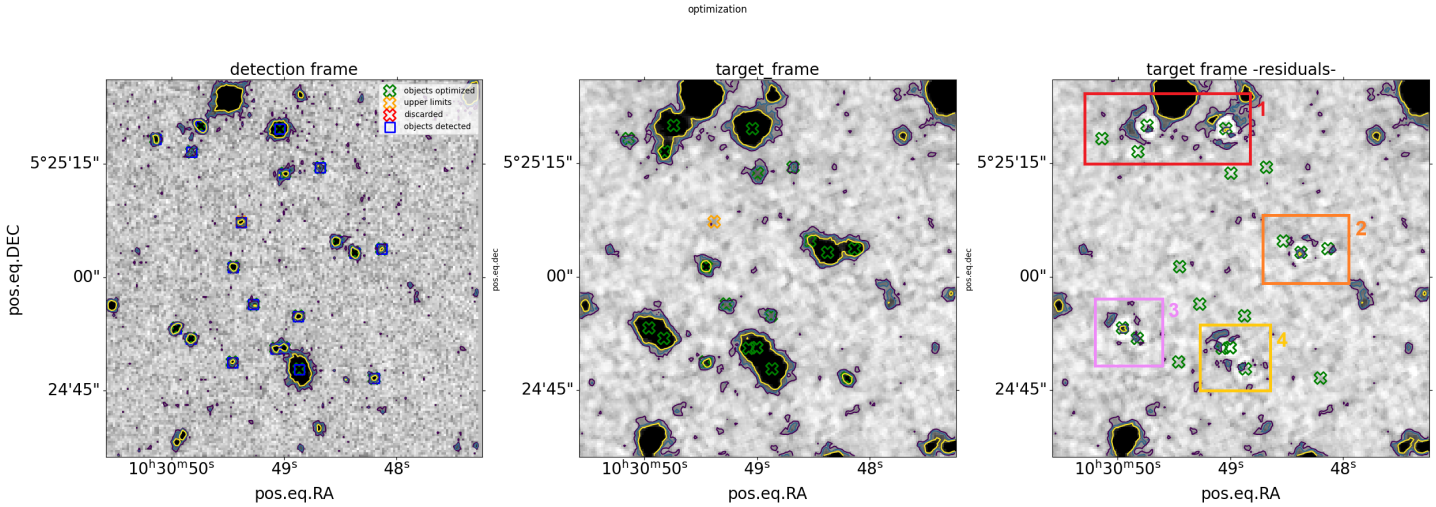


Figure B.1: *In the* ***left panel****: a cutout of a specific analysis region in PhoEBO's detection map (Ks) measuring 50″ × 50″, with sources identified on the map highlighted by blue squares. At the* ***center****: the same region for the target map (IRAC2), where sources marked with green crosses have been optimized and successfully subtracted, and those marked with orange indicate that PhoEBO was only able to determine an upper limit for the flux. On the* ***right****: the residual map with overplotted the groups of sources in various colors. Sources not enclosed within rectangles have been optimized individually. For each panel, contours at 2, 3, and 5σ are plotted, with the contours for the residual map calculated based on the rms of the original map.*

4. **Image Alignement**: PhoEBO has been enhanced to allow for the re-registration of analysis regions within the same frame. This process leverages the Astroalign library (Beroiz et al. 2020), which is capable of finding transformations for aligning images even with differing PSFs. The library was utilized to align the detection and target frames with each other. However, the transformation is applied exclusively to the detection frame to avoid possible issues of flux conservation in the target frame.
   Since Astroalign is limited to linear rigid transformations, it may fail when the misalignment between the two frames is more complex, with different components (like a rotational one); so the ability to fit the centroid of the sources for each fitting group has been implemented, physically shifting the models with sub-pixel precision.

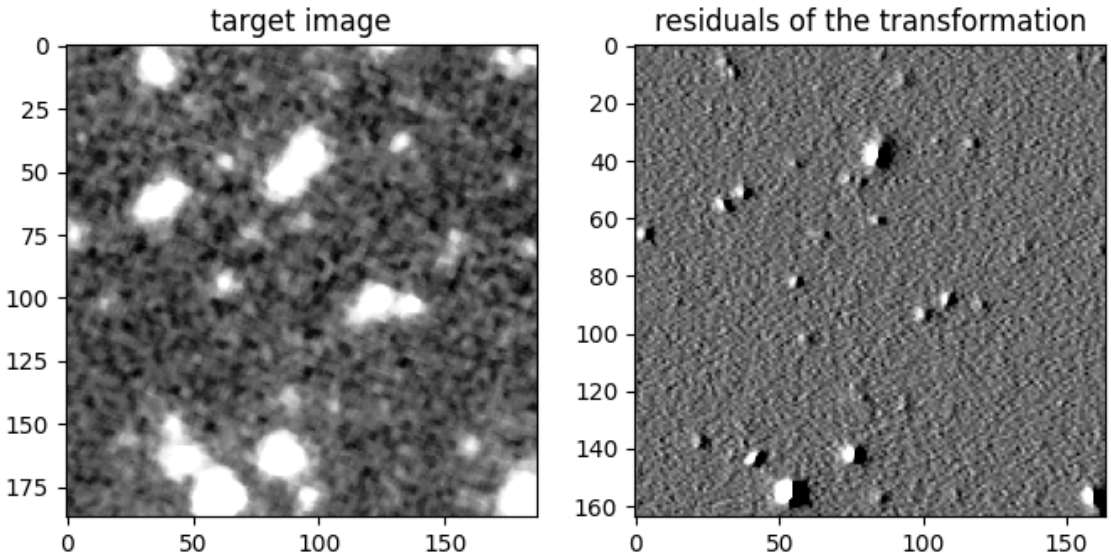


***Figure B.2: left**: Cutout of the analysis area of PhoEBO for IRAc ch2 band. **right**: the residual after the process of image alignment obtained as a subtraction between the original map and the same map after the transformation from Beroiz et al. (2020) is applied clearing showing the presence of a systematic shift. This example illustrates the target frame rather than the detection frame, to emphasize how the alignment issue primarily concerns the target.*

Appendix C

Sources Without PACS prior and Construction of Median SED

For all those sources for which data at 100 and 160 μm are not available (due to the limited coverage of the PACS observations), SED fitting for predicting the 250 μm flux is necessarily confined to data up to 24 μm. This translates in the absence of significant constraints in the far infrared range, and, due to this limitation, direct extrapolation of the flux is frequently incorrect.
To try to address this issue and obtain a more representative estimation of the SPIRE fluxes in line with the observations, I have extracted fluxes from the median SED (Fig C.1) that I constructed based on all those sources for which PhoEBO obtained flux measurements at 100 and 160 μm.
The underlying assumption of this approach is that the statistical characteristics of the sources do not vary significantly within the coverage area of SPIRE bands. Therefore, once the median SED is constructed, the flux at 250 μm can be extrapolated from it.

To construct the SED I followed these steps:

1. Red-shifting to Rest Frame: the best-fit SEDs are redshifted to bring them to the rest frame using, for each object, its photometric redshift
2. Interpolation on a common wavelength grid: The various templates, now in the rest frame, are interpolated onto a common wavelength grid.
3. Construction of the Median SED: The median SED is then constructed by taking the median of the values on this common grid.

Once the median SED is constructed, I used it to derive the fluxes at 250 μm by following these steps:

1. Redshift the Median SED: The first step involves redshifting the median SED to the best-fit redshift of the source considered.

2. Calculate the Predicted 250 μm flux: the predicted flux at 250 μm from the median SED is computed by averaging the values within the wavelength range over which the SPIRE 250 instrument operates.

3. Normalization to the Observed 24 μm Flux: The estimated flux is then rescaled by considering the ratio between the observed 24 μm of the considered source and the 24 μm flux derived from the median SED.

4. Model Subtraction: As the last step, models for contaminant sources are created and subtracted.

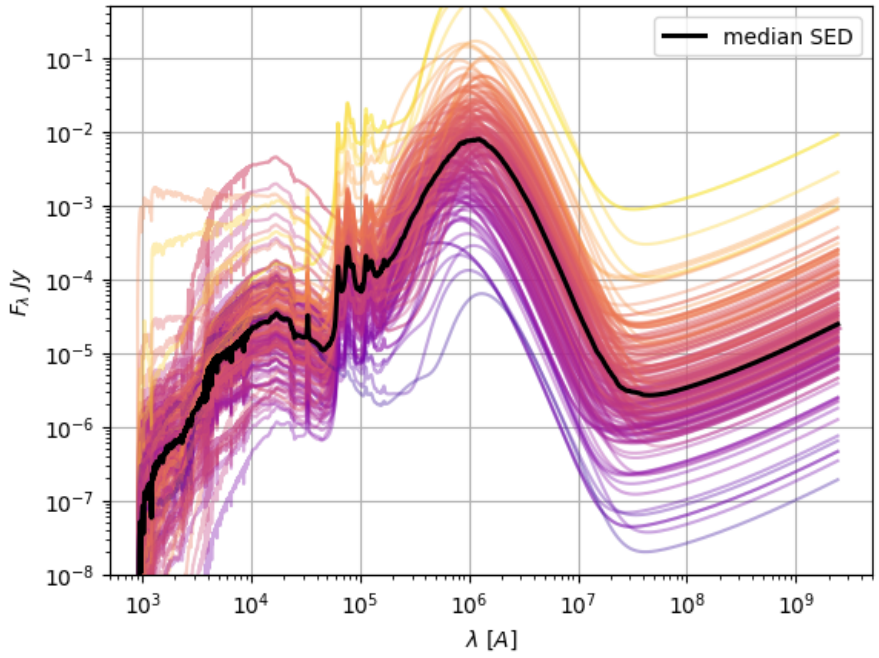


Figure C.1: *In black, the median SED obtained after the procedure described in this Section*

Bibliography

Adams N. J., et al., 2023, , 518, 4755

Aird J., et al., 2010, , 401, 2531

Aniano G., Draine B. T., Gordon K. D., Sandstrom K., 2011, , 123, 1218

Annunziatella M., et al., 2018, , 130, 124501

Balmaverde B., et al., 2017, Astronomy & Astrophysics, 606, A23

Barbary K., 2016, Journal of Open Source Software, 1, 58

Barrufet L., et al., 2023, , 522, 449

Battisti A. J., et al., 2019, , 882, 61

Behiri M., et al., 2023, , 957, 63

Behroozi P. S., Marchesini D., Wechsler R. H., Muzzin A., Papovich C., Stefanon M., 2013, , 777, L10

Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, , 149, 289

Beroiz M., Cabral J. B., Sanchez B., 2020, Astronomy and Computing, 32, 100384

Berta S., et al., 2013, , 551, A100

Bertin E., 2013, PSFEx: Point Spread Function Extractor, Astrophysics Source Code Library, record ascl:1301.001 (ascl:1301.001)

Bertin E., Arnouts S., 1996, , 117, 393

Blanc G. A., et al., 2008, , 681, 1099

Bouwens R. J., et al., 2012, , 754, 83

Bouwens R. J., et al., 2015, , 803, 34

Bouwens R. J., et al., 2021, , 162, 47

Bouwens R., Illingworth G., Oesch P., Stefanon M., Naidu R., van Leeuwen I., Magee D., 2023a, , 523, 1009

Bouwens R. J., et al., 2023b, , 523, 1036

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, , 351, 1151

Bruzual G., Charlot S., 2003, , 344, 1000

Calistro Rivera G., et al., 2017, , 469, 3468

Calzetti D., 2000, A UV Imaging Survey of IR-Bright Star- Forming Galaxies, HST Proposal ID 8721. Cycle 9

Carnall A. C., et al., 2023, , 619, 716

Casey C. M., Narayanan D., Cooray A., 2014, , 541, 45

Cecchi R., Bolzonella M., Cimatti A., Girelli G., 2019, , 880, L14

Chabrier G., 2003, , 115, 763

Charlot S., Fall S. M., 2000, , 539, 718

Chevance M., et al., 2020, , 493, 2872

Cimatti A., et al., 2008, , 482, 21

Daddi E., et al., 2009, , 694, 1517

Dale D. A., Helou G., 2002a, , 576, 159

Dale D. A., Helou G., 2002b, , 576, 159

Delhaize J., et al., 2017, , 602, A4

Delvecchio I., et al., 2014, , 439, 2736

Delvecchio I., et al., 2017, , 602, A3

Dickinson M., Giavalisco M., GOODS Team 2003, in Bender R., Renzini A., eds, The Mass of Galaxies at Low and High Redshift. p. 324 (arXiv:astro-ph/0204213), doi:10.1007/10899892_78

Donley J., et al., 2012, in American Astronomical Society Meeting Abstracts #219. p. 225.07

Donnan C. T., McLeod D. J., McLure R. J., Dunlop J. S., Carnall A. C., Cullen F., Magee D., 2023, , 520, 4554

Draine B. T., 2005, in Wilson A., ed., ESA Special Publication Vol. 577, ESA Special Publication. pp 251–255

Draine B. T., Li A., Hensley B. S., Hunt L. K., Sandstrom K., Smith J. D. T., 2021, , 917, 3

D'Amato Q., et al., 2022, Astronomy & Astrophysics, 668, A133

Elbaz D., et al., 2007, , 468, 33

Enia A., et al., 2022, The Astrophysical Journal, 927, 204

Fanaroff B. L., Riley J. M., 1974, , 167, 31P

Finkelstein S. L., et al., 2022, , 928, 52

Fritz J., Franceschini A., Hatziminaoglou E., 2006, , 366, 767

Gardner J. P., et al., 2006, , 123, 485

Gawiser E., et al., 2006, The Astrophysical Journal Supplement Series, 162, 1

Gentile F., et al., 2023, arXiv e-prints, p. arXiv:2312.05305

Gentile F., et al., 2024, arXiv e-prints, p. arXiv:2402.05994

Gilli R., et al., 2019, Astronomy and Astrophysics, 632, A26

Gordon K. D., Engelbracht C. W., Rieke G. H., Misselt K. A., Smith J. D. T., Kennicutt Robert C. J., 2008, , 682, 336

Gruppioni C., et al., 2013, , 436, 2875

Gruppioni C., et al., 2020, Astronomy & Astrophysics, 643, A8

Harikane Y., et al., 2022, , 259, 20

Harikane Y., et al., 2023, , 265, 5

Harikane Y., Nakajima K., Ouchi M., Umeda H., Isobe Y., Ono Y., Xu Y., Zhang Y., 2024, , 960, 56

Harrison C. M., et al., 2012, , 426, 1073

Hill J. M., Green R. F., Ashby D. S., Brynnel J. G., Cushing N. J., Little J. K., Slagle J. H., Wagner R. M., 2012, in Stepp L. M., Gilmozzi R., Hall H. J., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 8444, Ground-based and Airborne Telescopes IV. p. 84441A, doi:10.1117/12.926636

Hirashita H., Inoue A. K., Kamaya H., Shibai H., 2001, , 366, 83

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

Hurley P. D., et al., 2017, , 464, 885

Hwang H. S., et al., 2010, , 409, 75

Jin S., et al., 2018, , 864, 56

Jin S., et al., 2022, , 665, A3

Jones G. T., Stanway E. R., 2023, , 525, 5720

Kennicutt Jr R. C., Evans II N. J., 2012, Annual Review of Astronomy and Astrophysics, 50, 531

Kroupa P., 2001, , 322, 231

Lang D., Hogg D. W., Mykytyn D., 2016, The Tractor: Probabilistic astronomical source detection and measurement, Astrophysics Source Code Library, record ascl:1604.008 (ascl:1604.008)

Leipski C., et al., 2014, The Astrophysical Journal, 785, 154

Liu D., et al., 2018, , 853, 172

Madau P., Dickinson M., 2014, Annual Review of Astronomy and Astrophysics, 52, 415

Magnelli B., et al., 2015, , 573, A45

Marrone D. P., et al., 2018, , 553, 51

Molnár D. C., et al., 2021, , 504, 118

Morselli L., et al., 2014, , 568, A1

Nanni R., et al., 2018, , 614, A121

Nanni R., et al., 2020, , 637, A52

Noeske K. G., et al., 2007, , 660, L43

Novak M., et al., 2017, Astronomy & Astrophysics, 602, A5

Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, , 855, 105

Oke J. B., Gunn J. E., 1983, , 266, 713

Pilbratt G. L., Prusti T., Heras A. M., Leeks S., Marston A. P., Vavrek R., 2004, in American Astronomical Society Meeting Abstracts #204. p. 81.01

Pilbratt G. L., et al., 2010, , 518, L1

Planck Collaboration et al., 2020, , 641, A6

Pope A., et al., 2008, , 675, 1171

Popesso P., et al., 2023, , 519, 1526

Quadri R., et al., 2007, The Astronomical Journal, 134, 1103

Rodighiero G., et al., 2011, , 739, L40

Saintonge A., Catinella B., 2022, , 60, 319

Schenker M. A., et al., 2013, , 768, 196

Schmidt M., 1968, , 151, 393

Schreiber C., et al., 2018, , 618, A85

Scoville N., et al., 2007, The Astrophysical Journal Supplement Series, 172, 1

Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, , 690, 20

Simpson J. M., et al., 2020, , 495, 3409

Smail I., Ivison R. J., Blain A. W., 1997, The Astrophysical Journal, 490, L5

Smail I., et al., 2021, , 502, 3426

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, , 214, 15

Stefanon M., Bouwens R. J., Labbé I., Muzzin A., Marchesini D., Oesch P., Gonzalez V., 2017, , 843, 36

Steidel C. C., Adelberger K. L., Dickinson M., Giavalisco M., Pettini M., 1998, arXiv e-prints, pp astro–ph/9812167

Straatman C. M. S., et al., 2014, , 783, L14

Swinbank A. M., et al., 2014, , 438, 1267

Tacconi L. J., et al., 2006, , 640, 228

Talia M., Cimatti A., Giulietti M., Zamorani G., Bethermin M., Faisst A., Fèvre O. L., Smolçić V., 2021, The Astrophysical Journal, 909, 23

Tasca L. A. M., et al., 2015, , 581, A54

Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, , 621, 673

Toft S., et al., 2014, , 782, 68

Treyer M., et al., 2007, , 173, 256

Valentino F., et al., 2020, , 889, 93

Wang T., et al., 2019, , 572, 211

Weaver J., Toft S., Davidzon I., Capak P., McCracken H., 2019, in The Art of Measuring Galaxy Physical Properties. p. 9, doi:10.5281/zenodo.3554205

Weaver J. R., et al., 2022, The Astrophysical Journal Supplement Series, 258, 11

Werner M. W., et al., 2004, , 154, 1

Williams C. C., et al., 2019, , 884, 154

Wootten A., Thompson A. R., 2009, IEEE Proceedings, 97, 1463

Xiao M., et al., 2023, arXiv e-prints, p. arXiv:2309.02492

Yun M. S., Reddy N. A., Condon J. J., 2001, , 554, 803

Zimmerman D. T., Narayanan D., Whitaker K. E., Davè R., 2024, arXiv e-prints, p. arXiv:2401.06719

da Cunha E., Charlot S., Elbaz D., 2008, Monthly Notices of the Royal Astronomical Society, 388, 1595

IRSA Archives: https://irsa.ipac.caltech.edu/frontpage

Photutils Astropy affiliated package https://photutils.readthedocs.io/en/stable/index.html