UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

VNIVERSITATIS HAFNIENSIS
FACVLTATIS NATVRALIS

CHARTING THE STELLAR, DUST AND GAS CONTENT OF GALAXIES

A Panchromatic View at the Galaxy Evolution Across Cosmic Time

Dissertation submitted for the degree of

Philosophiæ Doctor

to the PhD School of The Faculty of Science, University of Copenhagen

on May 22, 2022, by

Vasily Kokorev

Supervisor: *Professor Georgios Magdis*

CHARTING THE STELLAR, DUST AND GAS CONTENT OF GALAXIES

A Panchromatic View at the Galaxy Evolution Across Cosmic Time

The Universe is change, life is opinion.

— Marcus Aurelius

Abstract

Since time immemorial the humanity has gazed upon the heavens in fascination. Generations upon generations have attempted to uncover the mysteries of the cosmos and understand the inner-workings of this complex mechanism. At first, the luminous bands of light in the night sky belonging to our home the Milky Way galaxy, have been revealed to be trillions of stars, not unlike our Sun. Soon after, by observing the faint and diffuse nebulae barely visible in the night sky, we have come to understand that a great multitude of galaxies similar to our own exist in the vast expanse that we call the Universe. The light which travels to us from these distant galaxies, has a finite speed, thus allowing us to look far not only in space, but also time. By observing the Universe at different distances, we can therefore reconstruct the history of the galaxy formation by looking at various snapshots spread out in time. In my thesis, I primarily studied the interstellar medium, which consists of two primary ingredients - dust, which attenuates and dims the stellar light, and gas, the fuel for star formation.

First, I developed an algorithm - `Stardust`, which combines the information from light, emitted by galaxies at different wavelengths, in order to understand their physical properties. These properties include total masses of all stars, dust and gas, fraction of light contributed by the active galactic nuclei, as well as rate of formation of new stars. By applying my algorithm to the deepest and most complete optical and infrared galaxy catalogues to date I have conducted a thorough examination of the evolution of both the dust and gas - to stellar mass ratios over a course of 12 billion years of cosmic history. The scaling relations which I recover, not only agree with the existing studies, but also push our knowledge of dust and gas mass evolution to higher redshifts. In addition to that, my study is one of the most statistically robust and complete investigations of the ISM evolution to date and can provide a perfect test-bed for simulations. During this analysis I have also discovered a population of galaxies with massive gas mass reservoirs, the so called 'gas-giants', which defy the current observational and theoretical predictions.

Second, I have constructed a set of *HST* and *Spitzer* mosaics, and photometric catalogues, covering the ALMA lensing cluster survey (ALCS) fields. My work presents, for the first time, a uniformly combined and reprocessed set of archival *HST*/*Spitzer* observations covering 33 lensed cluster fields in Hubble Frontier, RELICS and CLASH, and currently contains the deepest and most

complete photometry in those fields. These catalogues will serve as an important tool in aiding the search of the sub-mm galaxies in future ALMA surveys, as well as follow ups of the "*HST* dark" - IRAC sources. The combination of photometry from two telescopes allows us to place better constraints on photometric redshifts and stellar masses of these galaxies, thus giving us an opportunity to identify high-redshift ($z > 7$) candidates for spectroscopic follow ups and answer the important questions regarding the epoch of reionisation and formation of first galaxies.

Finally, I conducted a preliminary assessment of source detection feasibility for the MIRI component of the upcoming upcoming *JWST* Cycle I GO COSMOS - Webb program. By simulating the SEDs and modeling the MIRI mosaics I predict that the survey is expected to recover a significant proportion of optically detected sources. However more work is required to correctly estimate the number and properties of optically dark and high-z galaxies. These sets of simulations will be used to optimise critical source detection algorithms, and aid in the interpretation and reduction of the data.

The work presented in this thesis, including the peripheral studies that I have contributed to, will allow us to expanding our understanding of the evolution of dusty star forming galaxies. Moreover, the tools and the catalogues produced during this thesis are made available to the wider scientific community, and can be used to conduct the search and analysis of new and unique galaxy populations.

Resumé

Siden tidernes morgen har menneskeheden stirret fascineret op på himlen. Generationer efter generationer har forsøgt at afdække kosmos dets mysterier og forstå dets komplekse mekanismer. Tidligt blev det lysende bånd på nattehimlen, der tilhører vores hjem, Mælkevejen, afsløret som billioner af stjerner, ikke ulig vores Sol. Ved at observere de svage og diffuse tåger, der anes på nattehimlen, forstod vi senere, at der findes et væld af galakser, der ligner vores egen, i det vældige rum, som vi kalder Universet. Lyset, som rejser til os fra disse fjerne galakser, har en begrænset hastighed, hvilket gør os i stand til at se langt, ikke kun i rummet, men også i tiden. Gennem observationer af Universet i forskellige afstande kan vi derfor rekonstruere historien om galaksernes dannelse ved at se på forskellige øjebliksbilleder i tiden. I min afhandling har jeg primært studeret det interstellare medium, som består af to primære ingredienser — støv, som absorberer og formørker stjernelyset, og gas, brændstoffet til stjernedannelse.

For første har jeg udviklet en algoritme — `Stardust`, som kombinerer informationen fra lyset, der udsendes af galakser ved forskellige bølgelængder, for at forstå deres fysiske egenskaber. Disse egenskaber omfatter de totale masser af alle stjerner, støv og gas; andelen af lyset, som de aktive galaksekerner bidrager med; samt hastigheden af dannelsen af nye stjerner. Ved at anvende min algoritme på de dybeste og mest komplette optiske og infrarøde galaksekataloger til dato har jeg gennemført en grundig analyse af udviklingen af både støv- og gas-til-stjernemasseforholdet henover 12 milliarder års kosmisk historie. De skaleringsforhold jeg finder stemmer ikke kun overens med eksisterende undersøgelser, men udvider også vores viden om støv- og gasmasseudvikling til højere rødforskydninger. Derudover er mit studie en af de mest statistisk robuste og komplette undersøgelser af udviklingen af det interstellare medium til dato og kan give et perfekt udgangspunkt for computersimuleringer. I løbet af denne analyse har jeg desuden opdaget en population af galakser med meget store reservoirer af gas, de såkaldte 'gas-giganter', som trodser de nuværende observationelle og teoretiske forudsigelser.

For det andet har jeg konstrueret et sæt *HST*- og *Spitzer*-mosaikker og fotometriske kataloger, der dækker ALMA lensing cluster survey (ALCS)-felterne. Mit arbejde præsenterer for første gang et ensartet kombineret og gen-processeret sæt af *HST*/*Spitzer*-arkivobservationer, der dækker 33 tyngdelinse-forstørrede felter med galaksehobe i Hubble Frontier, RELICS og CLASH, og

indeholder den for tiden dybeste og mest komplette fotometri i disse felter. Disse kataloger vil tjene som et vigtigt værktøj til at hjælpe med at søge efter sub-mm-galakser i fremtidige ALMA-observationer, såvel som opfølgninger af "*HST* dark"-IRAC-kilderne. Kombinationen af fotometri fra to teleskoper tillader os at sætte bedre begrænsninger på fotometriske rødforskydninger og stjernemasser af disse galakser, hvilket giver os mulighed for at identificere højrødforskydnings-kandidater ($z > 7$) til spektroskopiske opfølgninger og besvare vigtige spørgsmål vedrørende reioniserings-epoken og dannelse af de første galakser.

Endelig foretog jeg en foreløbig vurdering af muligheden for at gennemføre kildedetektioner af MIRI-delen af det kommende *JWST* Cycle I GO-program COSMOS - Webb. Ved at simulere SED'erne og modellere MIRI-mosaikkerne forudsiger jeg, at undersøgelsen forventes at genskabe en betydelig andel af optisk detekterede kilder. Mere arbejde er dog nødvendigt for korrekt at estimere antallet og egenskaberne af optisk mørke og høj-rødforskydnings-galakser. Disse sæt simuleringer vil blive brugt til at optimere kritiske kildedetekteringsalgoritmer og hjælpe med fortolkning og reduktion af dataene.

Det arbejde, der præsenteres i denne afhandling, herunder flere parallelle studier som jeg har bidraget til, vil give os mulighed for at udvide vores forståelse af udviklingen af støvfyldte, stjernedannende galakser. Desuden er værktøjerne og katalogerne, der er produceret i løbet af denne afhandling, gjort offentligt tilgængelige for det bredere videnskabelige samfund og kan bruges til at udføre søgning og analyse af nye og unikke galaksepopulationer.

Acknowledgements

First of all I would like to offer my deepest gratitude to my parents and family. Thank you for all the support you have given me during this journey. I would not be where I am without your constant and continuing care and patience. Thank you to all of my dear friends, who were so supportive and encouraged me to keep going over all of these years. Without you, my sanity would have surely perished. You are all great.

I would like to say a big thank you to my PhD supervisor Georgios. For you guidance and great patience over the years, for helping me to become and overcome. The academic world can be chaotic and confusing sometimes, so thank you for introducing me to all the wonderful teams of international scientists and collaborations. Thank you Sune for bringing this amazing team together and allowing me to be a part of it. My gratitude to Gabe, for introducing me to the world of optical astronomy and helping me overcome its many technical challenges. Thank you Charles, for all the great discussions and picking my brain from time to time. I would like to thank Caitlin for hosting me at UT Austin, for her patience, support and research that contributed to this thesis. Thank you for giving me the opportunity to observe the first galaxies with Keck, on top of Mauna Kea in Hawaii, and a big thank you to the Keck support scientists for hosting us.

I would also like to extend a special gratitude to the wonderful people at DAWN, where I spent the last few years. My sincere gratitude to all my colleagues for their inspiring research and a hundreds of hours spent chatting by the coffee machine. Thank you to the wonderful Guarn for holding this centre together and putting out all the fires. To my brother in arms John, thank you for all the scientific discussions and whisky fueled fun times. Thank you Francesco for all your help and support. Thank you Seiji for your encouraging advice and great ramen. Thank you to the old guard, Carlos, Isabella and Mikkel, for imparting invaluable knowledge and passing us the torch.

Finally, I would like to thank the Assessment Committee Members, Professor Andrew Bunker, Professor James Dunlop and Professor Charlotte Mason for expressing interest in my research and finding the time to schedule and censor the dissertation and the defence.

Contents

CHAPTER 1

Introduction

Within the contemporary cosmological and astronomical context, the Universe has manifested itself as a result of a Big Bang. This model predicts that all energy in the Universe initially existed in a state of extreme temperature and density. This fundamental cosmological theory came as a result of independent discoveries by George Lemaître and Edwin Hubble, who noted a linear correlation between the distance and the radial velocity of observed extra-galactic nebulae (Lemaître 1927, 1931; Hubble 1929; Hubble & Humason 1931) – the Universe was expanding. Two decades prior to the discovery of cosmic expansion, Albert Einstein has extended the works on the classical gravitation by Sir Isaac Newton (Newton 1687) and presented his theory of general relativity which described gravity within the geometrical context of space and time (Einstein 1916). Soon after, Alexander Friedmann calculated the expanding solutions to Einstein's field equations (Friedmann 1922). Combining the theories of general relativity, solutions to the field equations and the observational evidence of expansion, it was possible to extrapolate back in time and describe the Universe to exist as an infinitely dense singularity, where our laws of physics and understanding of concepts of time and space lose meaning.

1.1 Inflation, Synthesis of First Nucleons, and Structure Formation in the Early Universe

The accidental discovery of the cosmic microwave background (CMB) by Arno Penzias and Robert Wilson (Penzias & Wilson 1965) and its subsequent high resolution follow-ups with the Cosmic Background Explorer (COBE) (Smoot et al. 1992) have revealed an extremely uniform distribution of matter and radiation, with only minute fluctuations present. Later denoted as the Copernican or the cosmological principle, the Universe appeared to be isotropic and homogeneous for all possible observers, in all directions. In this scenario a single model of expansion was not adequate to explain the observed uniformity, as the matter in the Universe would no longer be in causal contact with each other, and thus achieving a thermal equilibrium would not be possible. Two decades after the discovery of the CMB, an alternative mechanism to explain expansion at very early times has been proposed by several theoretical physicists in the field, which aimed to alleviate the conflict between theory and observations (Albrecht & Steinhardt 1982; Guth 1981; Linde 1982, 1983). This new model introduced a period of inflation, where rapid exponential expansion occurred in the first $\sim 10^{-35}$ s of the lifetime of the Universe. As a result, effectively "freezing" in place the moment immediately following the Big Bang when all the matter in the Universe was in causal contact.

Following inflation, the matter in the Universe cooled down and allowed for the leptons and quarks to fuse into first protons, or ionised hydrogen. This process would later go on to form more complex atoms, helium and lithium, as described by a model of chemical production during the primordial nucleosynthesis was, which was first explored by Ralph Alpher, Hans Bethe, and George Gamow (Alpher et al. 1948). Further expansion and cooling of the Universe has allowed for electrons to recombine with protons, forming neutral hydrogen. This period, known as the recombination resulted in the Universe transitioning from fully ionised, to a neutral state, thus allowing photons to escape for the first time. The light escaping during the epoch of recombination, some 370,000 after the Big Bang, is exactly what has been observed as the CMB.

As discussed previously, the CMB appears to be homogeneous and isotropic, with its temperature being measured at 2.7 K, with minor anisotropies, on the scale of $\sim 10^{-5}$ K (Planck Collaboration et al. 2020). These perturbations are thought to arise from the quantum fluctuations immediately following the Big Bang, which have been frozen in time by inflation. While seemingly insignificant, these perturbations provide key conditions, which result in the subsequent structure formation. The small perturbations on a highly homogeneous and isotropic Universe would be eventually gravitationally pulled together, forming slight overdensities, and this manifest themselves into a cosmic web. Within these nodes, clouds of gas collapse under their own gravity, overcoming any other forces that support them, with subsequent fragmentation and further collapse resulting in the emergence of the first celestial bodies.

The first objects that formed in the early Universe would emit extreme amounts of radiation, sufficient enough to ionise the existing neutral hydrogen back into a hot plasma. This time period is called "reionisation" and started around 150 Myr after the Big Bang (Gunn & Peterson 1965). This however did not result in the Universe becoming opaque to photons again, as the cosmic expansion has already caused the matter in the Universe to become diffuse. This as a result has reduced the probability of interaction between photons and electrons, thus leaving the Universe transparent to radiation, just as after recombination.

Currently the most commonly utilised and accepted cosmological picture of energy and matter distribution is the Lambda cold dark dark matter model (ΛCDM). In this context the Universe is geometrically flat, primarily consisting of the dark energy (Λ), and the cold dark matter (CDM). In combination both of these components constitute $\sim 97\%$ of the total energy budget, however are yet to be understood. Although the first discussion about the "dark bodies" in the Milky Way galaxy were started by Lord Kelvin in 1884, the first evidence of non-luminous exotic matter which interacts only through gravity was not presented until 1922 by Jacobus Kapteyn, and then later confirmed by Knut Lundmark, Jan Oort and finally set into stone by Fritz Zwicky (Kapteyn 1922; Lundmark 1930; Oort 1932; Zwicky 1933). Despite the discussions regarding the origin and composition of the Universe still being a contentious topic among contemporary scholars, the ΛCDM model has been able to adequately explain the observed data and contains significant predictive power when coupled with theoretical models and simulations. This model has allowed scientists to explain the existence of the CMB, the large-scale structure and clustering of galaxies, the currently observed abundance of elements, and the accelerating cosmic expansion.

Observations of the CMB by COBE, WMAP and most recently Planck have constrained the cosmological parameters into a concordance cosmology model that we use today (see Figure 1.1). This picture outlines the matter density $\Omega_m = 0.3153 \pm 0.0073$, of which only one sixths corresponds to ordinary matter; and the cosmological constant, or the dark energy density is $\Omega_\Lambda = 0.6847 \pm 0.0073$. In a Universe with no curvature ($k \sim 10^{-9}$), the total density $\Omega_{\rm tot} = 1$, which means that the contribution from radiation density $\Omega_{\rm r} \sim 10^{-5}$ is negligible. With these parameters the Universe is currently expanding at a rate of $H_0 = 67.74 \pm 0.46$ km s^{-1} Mpc^{-1} and has an age of 13.799 Gyr (Planck Collaboration et al. 2020).

1.2 The Great Debate and Emergence of "Island Universes"

The majority of 'standard', baryonic matter in the local Universe exists in the form of warm/hot gas in the intergalactic space, however some is also condensed into collections of stars, planets, and interstellar gas and dust which are bound to each other gravitationally. These systems are called "galaxies", originating from the Greek word "$\gamma\alpha\lambda\alpha\xi\iota\alpha\varsigma$", meaning "milky", owed to interstellar gas and dust giving these systems such an appearance. Galaxies vary a lot in terms of their mass, and can contain anywhere from tens of millions to hundreds of trillions stars. Physical processes that govern these complex systems, as well as their shapes, and spatial location within the Universe are also extremely diverse. However, the majority of galaxies are thought to contain a super-massive black hole in their centres, and will be located within galaxy clusters, at the nodes and filaments of the cosmic web. The two existing scenarios which attempt to describe formation of galaxies are the top-down and the bottom-up models. In the former case, galaxies arose as a result of segmentation and gravitational collapse of large perturbations. The latter, and

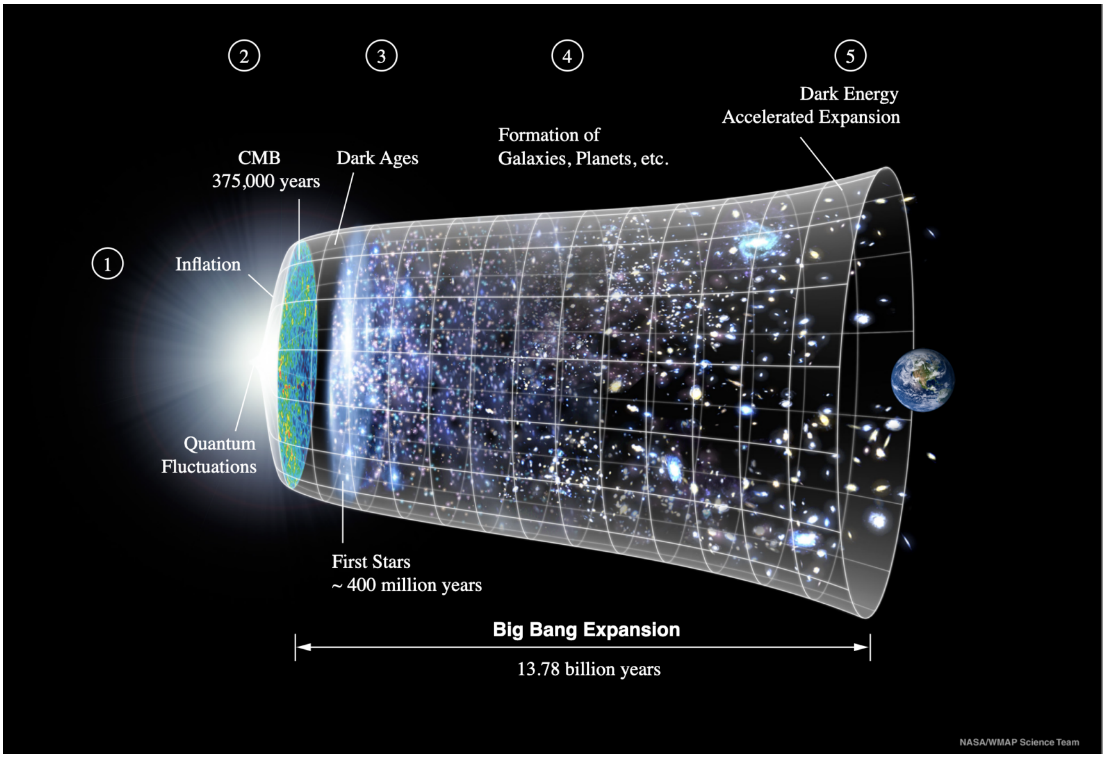

Figure 1.1: Dissected timeline of the Universe, from its birth during the Big Bang, through inflation, recombination and subsequent formation of structure. Based on the standard cosmological evolutionary model. Credit: NASA/WMAP Science Team.

more recent, scenario however argues that galaxies manifested themselves as a result of a build up of small scale matter anisotropies. The intergalactic space is not empty, and is filled with low density, but dominant in terms of mass, amounts of gas called the intergalactic medium (IGM), with densities ranging around a few atoms per cubic metre.

Our modern, bottom-up, understanding of galaxies and their formation however, has not always been as such, and even the concept of existence of extragalactic structures was at some point under heavy scrutiny. The first conceptualisation of galaxies began to arise during the times of the ancient Greece, by the greatest minds of that time, Democritus and Aristotle, believing that stars, or emission thereof are responsible for the streaks of bright light seen in the night sky. Arabian and Persian scholars also noted that it was, at that time, impossible to measure the distance to these light filaments by using parallax, and as such attributed that light as one being very distant. It was not until the onset of 17th century, when Galileo Galilei studied that light, and confirmed it to consist of the myriad of very faint stars. As such, the first galaxy to be ever detected, our own Milky Way, was observed. Identification of other galaxies followed during that time or soon after, with Andromeda and Triangulum galaxies and Large and Small Magellanic clouds being visible with the naked eye.

Extragalactic origins of these galaxies were not immediately apparent, and two different

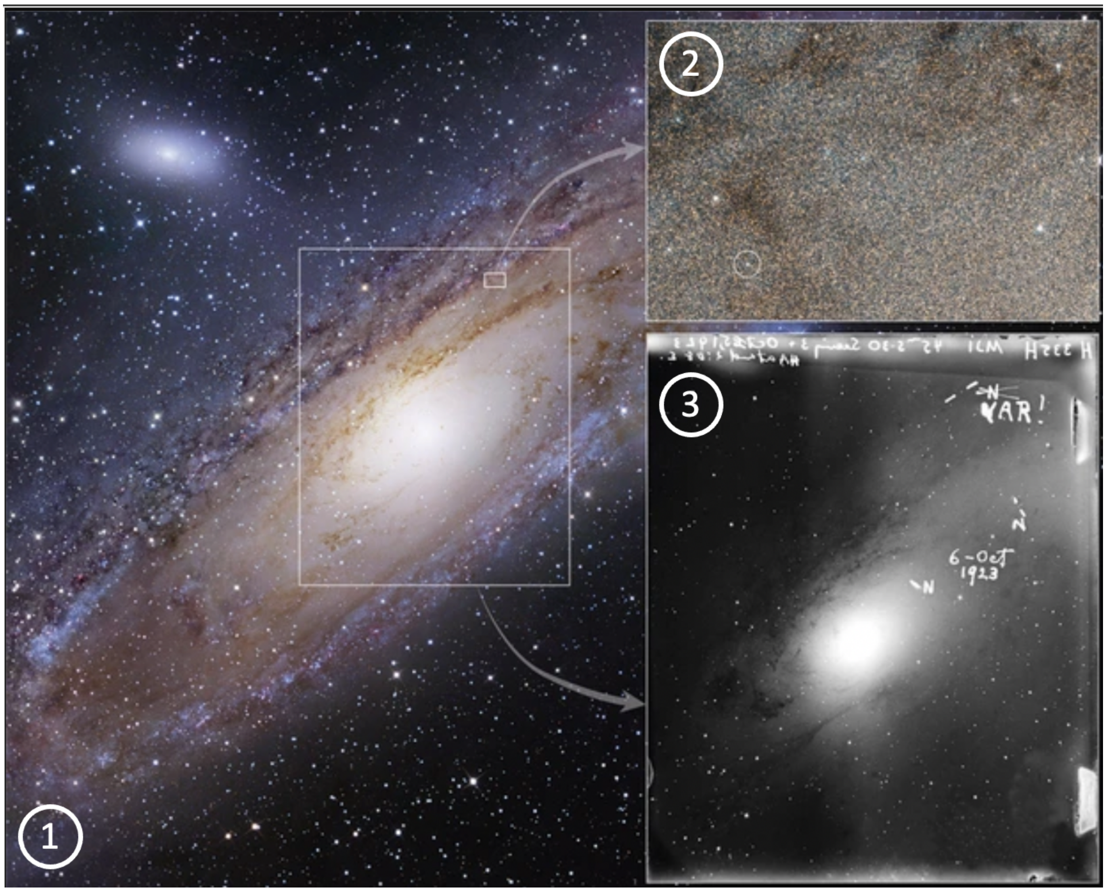

Figure 1.2: Clockwise from bottom left: 1) High resolution optical image of the Andromeda Galaxy. The location of the Cepheid star observed by Hubble in 1923 is circled on the original and a zoom-in in panel 2); 3) The original glass plate image taken by Hubble. Credit: NASA, ESA, and the Hubble Heritage Team (STScI/AURA); Illustration: NASA, ESA, and Z. Levay (STScI).

schools of thought quickly formed in an attempt to describe these distant nebulae. The event called the Great Debate between fellow astronomers Harlow Shapley and Heber Curtis took place in the year of 1920. Based on the arguments which involved the rate of novae and the rotational velocities within the "Great Spiral Nebula" in the constellation Andromeda, being seemingly decoupled from our own Milky Way (Webb 1999). It was clear then, that the Andromeda nebula was in fact a separate galaxy, its own "island universe", a term coined by the 18th century philosopher Immanuel Kant. This notion was later cemented by Edwin Hubble, Henrietta Leavitt and Vesto Slipher, who confirmed the distance to Andromeda galaxy to be well beyond the scale of the Milky Way by measuring a set of "standard candles", Cepheid stars (see Figure 1.2) of known brightness (Leavitt 1908; Slipher 1917; Hubble 1929).

Presenting extragalactic distances in terms of regular units of measure quickly becomes impractical. In the local Universe, we can still understand and measure separation between galaxies in terms of millions of years it takes light to travel between them, however as we move outside of our galactic neighbourhood, these simply become too large to comprehend. As a direct consequence of that, contemporary astronomers normally measure and report distances to galaxies in terms of their "redshift" - z. Redshift is a measure of how much the frequency or wavelength of a light beam emitted by a source, has shifted from its "rest" frame to the "observed" frame, as measured by instruments in the Solar system. This (in non-relativistic case) can be defined as:

$$\frac{\lambda_{\text{observed}} - \lambda_{\text{emitted}}}{\lambda_{\text{emitted}}} = \frac{v}{c} = z, \tag{1.1}$$

where λ is the wavelength, v is the recessional velocity, and c is the speed of light. Although galaxy rotation, or proper motion can also shift radiation away from its original wavelength, the primary source of redshift is the cosmic expansion, as described by the Hubble–Lemaître law. In fact, if we combine this law with the definition of redshift in terms of wavelength "shift", we will obtain a direct linear correlation between redshift and distance:

$$D = \frac{v}{H_0} = \frac{c\,z}{H_0}, \tag{1.2}$$

where H_0 is the Hubble constant, and D is the proper distance. For some context, the furthest galaxy that has been confirmed at the moment of writing this thesis, has been observed to be at $z = 11.09$ (Oesch et al. 2016), some 400 Myr after the Big Bang, the recombination, which gave rise to the CMB, occurred at $z \sim 1, 100$, and reonisation happened at $z \sim 6 - 8$.

1.3 Galaxy Formation and Morphology

Even before the full understanding of the origin and distance to the extragalactic nebulae has become clear, two unambiguously prominent galaxy shapes stood out. The first one being the elliptical, that can be spherical or slightly elongated, with largely random motion of stars. This galaxy type is also more yellowish or red in colour. The second type is a disk or spiral like shape, which is most recognisable by a flat disk, or a bar of uniformly rotating stars and gas. This latter shape also appears to be more white or blue. The first galaxy classification sequence, the so called "tuning fork" (see Figure 1.3) has been compiled by Edwin Hubble. This diagram also assumed that, in an evolutionary sense, ellipticals were "early" - type galaxies, and spirals were "late" type, which later on turned out to be incorrect (Hubble 1936). This galaxy classification was later improved upon, by Gérard de Vaucouleurs and Allan Sandage, who introduced more complexity to the "fork" and added new subdivisions to the existing types. These classifications remain in use to this day, although the methodology has largely evolved to rely on computational techniques.

Given our understanding of cosmology and conditions in the early Universe, the first galaxies formed as a result of baryonic matter, primarily hydrogen gas clouds, collapsing within the

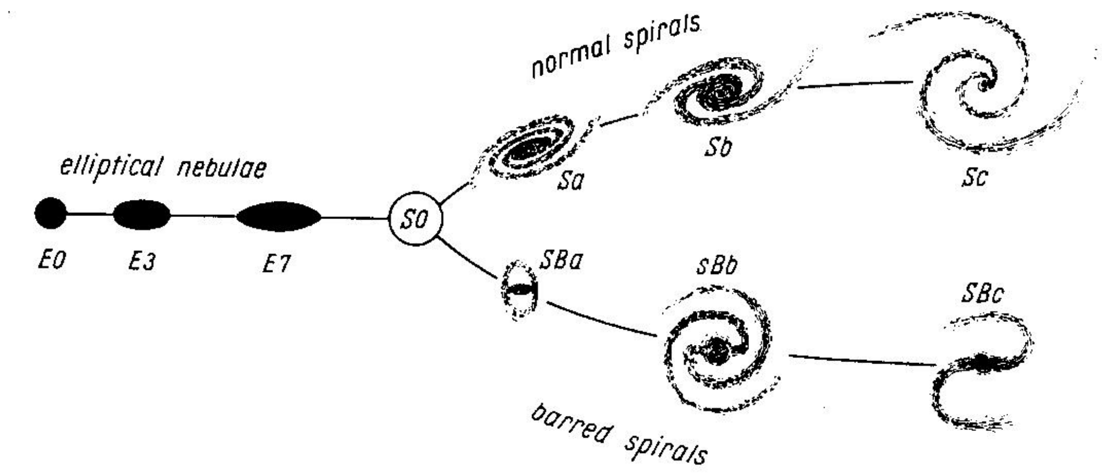

Figure 1.3: The original Hubble tuning fork diagram, as published in the "Realm of the Nebulae" (Hubble 1936). An early attempt to classify the complex array of galaxy morphology.

cold dark matter overdensities, or halos. In the early 70s, Vera Rubin noted that the rotational velocity of observed gas and stars in galaxies, does not seem to follow the expected trends given the predicted mass and distance from the centre, indicating the presence of some other unseen gravitationally interacting matter (Rubin 1983). These observations laid foundation towards models predicting that all galaxies formed in such dark matter halos, namely the two-stage theory for galaxy formation and clustering, presented in White & Rees (1978). The first stage involves the gradual build-up of dark matter halos, while the second stage involves the the gravitational collapse of baryonic matter into the potential wells of dark matter halos. As that matter collapses, the gas clouds start to fragment, and begin further collapsing into an array of smaller overdensities. In most cases, the gravitational collapse of these individual fragments does not continue indefinitely, and upon reaching hydrostatic equilibrium they would turn into stars.

Depending on the initial temperature and angular momentum of the initial gas cloud infalling into a dark matter halo, stars can begin forming at different time intervals. In cases where star formation is slow, the infalling gas has the time to dissipate the excess angular momentum, and virialise into a disk supported by gravitational potential and initial kinetic energy. This process results in galaxies that have a more disk - like or spiral appearance. Alternatively, if the gas collapse and fragmentation are quicker, the newly formed stars have few mechanism to dissipate their original angular momentum, resulting in random stellar motion within the system, giving rise to what we know as the elliptical galaxies (Eggen et al. 1962). The above model explains how elliptical galaxies can form in secular, isolated scenarios. The current scientific consensus however, is that all galaxies initially start off as disks, and later on merge (White 1978, 1979; Gerhard 1981). With the formation of elliptical galaxies being a direct consequence of such intergalactic collisions, that disturb ordered rotation (Toomre & Toomre 1972). In fact the existence of massive dark matter halos increases the frequency of such collisions, thus making formation of elliptical galaxies more probable (Efstathiou & Silk 1983).

1.4 Star Formation and Interstellar Dust

The study of galaxy formation and evolution both at high-redshift and locally still pose a major challenge for modern day astronomy. The processes that govern star formation can be separated into three major components, the collapse of gas clouds and their conversion into stars, the accretion and replenishment of gas from the IGM, and the interactions between galaxies. Galaxy mergers play a major role in rejuvenation or possibly cessation of star forming activity, gas replenishment and redistribution, as well as feeding the central super-massive black holes (SMBH) or active galactic nuclei (AGN). As discussed in the previous section, galaxy collisions can also transform an ordered spiral or disk - like morphology into a more random ellipsoidal shape.

One of the main mechanisms driving galaxy evolution is the interaction between the interstellar medium (ISM), primarily consisting of gas and dust, and the radiation field induced by stellar activity. In this context, dust poses challenges in the detection of the ultraviolet (UV) and optical emission of galaxies and in the interpretation of these observations in terms of physical properties. Dust is a fundamental building block of galaxies in the Universe. Interstellar dust grains consist in sub-micron and micron sized particles within the ISM. These have been observed and documented around the same time as our understanding of extragalactic nebulae has started taking shape (Trumpler 1930). Dust particles primarily manifest themselves via the reddening of stellar radiation, where short wavelength, high energy light being absorbed and then subsequently re-emitted at longer wavelengths, thus reddening starlight. The amount of total energy absorbed will depend on the initial wavelength of the radiation, with higher energy emission being absorbed at a higher rate, compared to low energy (see Figure 1.4) (Calzetti et al. 2000; Calzetti 2001)

Nominally the presence of dust has been considered a detriment to extragalactic observations and estimate of physical parameters, with various modelling and simulations of interstellar grains being conducted in an attempt to understand and correct for the dust extinction. Since then, the wealth of early infrared surveys, notably the ones performed with the Infrared Astronomical Satellite (IRAS), have positioned dust as an essential part of the evolution and growth of galaxies, and as an important tracer of star formation and ISM in the far-IR (FIR) observations. Moreover, dust is now being commonly used as a tracer of ISM composition, mass, temperature and structure (e.g. Compiégne 2010; Bernard et al. 2010; Miville-Deschênes et al. 2010; Magdis et al. 2012a; Scoville et al. 2016; Narayanan et al. 2011, 2012; Narayanan & Davé 2012 and many others). Dust can also assist with cooling of gas clumps, via the dust-gas collisions, and act as a catalyst that aids the gas to transition from the atomic to molecular form, a necessary step in the star formation

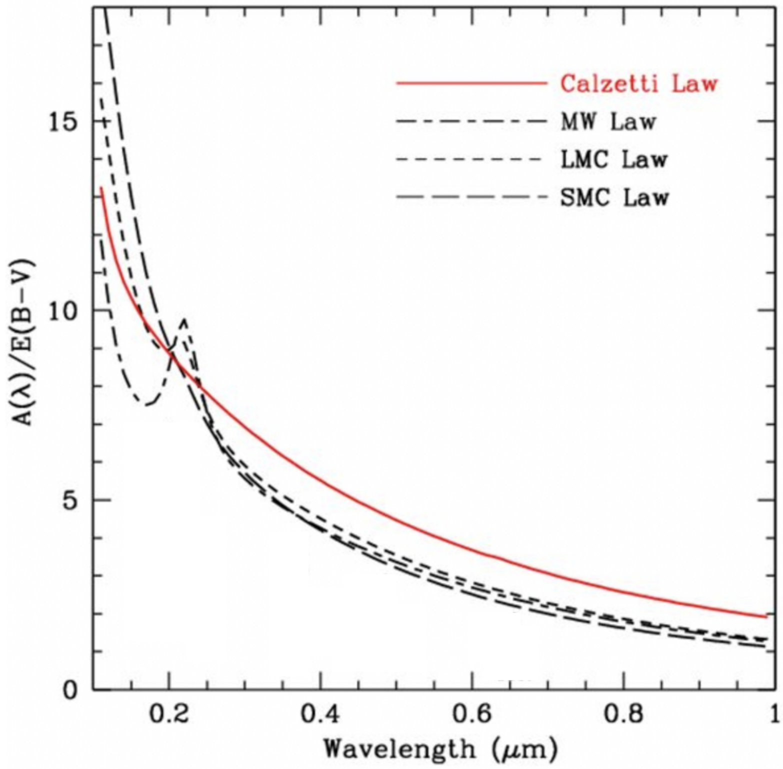

Figure 1.4: Extinction factor $R_V \equiv A(\lambda)/E(B-V)$ as a function of wavelength, estimated for the Small and Large Magellanic Clouds (SMC and LMC) and the Milky Way (MW). Figure adopted from (Calzetti 2001) and the "First Light in the Universe" slides by Richard Ellis (2006).

activity.

The exact origins and evolution of cosmic dust are still not understood in their entirety. Interstellar dust grains are intricately tied to the conditions of the ISM and follow a complex feedback loop of condensation, growth, ejection and destruction. The three major dust production mechanisms are the asymptotic giant branch (AGB) stars, supernovae, and grain growth within dense molecular clouds. In the first case, the low and intermediate mass AGB stars which are at the end of their life cycle rapidly lose mass due to the stellar winds and inject vast amounts of dust into the ISM (Matsuura et al. 2009). In the second case, the Type II core collapse supernovae will rapidly eject up to 10 $\%$ of their initial mass as dust (Arnett et al. 1989), however the radiation and winds produced during these massive explosions also has a potential in destabilising and destroying the existing dust. In the latter scenario the increase of the dust mass in the ISM is a manifestation of the slow grain growth, via the accretion of metals on the gas phase onto the grains (Dwek 1998; Zhukovska et al. 2008; Inoue 2011; Asano et al. 2013). Dust can be also be destroyed through its incorporation into stars, a process known as astration. Owed both to simulations and observational studies of dust evolution, we now know that these processes can become more or less dominant, depending on redshift. In these scenarios, galaxies at high-z

will form dust via condensation of the AGB and supernovae ejecta, while at low-z grain growth becomes more prominent (Triani et al. 2020).

1.5 Evolution of Galaxies

During the course of the last twenty years the evolutionary scenario which emerged from the analysis and processing of, first only the UV/optical, and then FIR/millimeter data, has lead to a discovery that a majority of star-forming galaxies (SFGs) follow a tight correlation between their star formation rate (SFR) and the stellar mass (M_*) (see Figure 1.5). This relation, the so called main sequence (MS) holds for all galaxies in a given time or redshift interval, with the normalisation factor - specific SFR (sSFR = SFR/M_*), increasing up to $z \sim 4$ (e.g. Brinchmann et al. 2004; Daddi et al. 2007, 2010; Magdis et al. 2010; Elbaz et al. 2011; Sargent et al. 2014; Schreiber et al. 2015) (see also Figure 1.6). Robust estimates of both SFR and M_*, and thus parametrisation of the MS above $z = 4$, are complicated by a combination of the lack of depth of the current surveys, and the lack of the wavelength regime which allows to calculate these parameters. The recent launch of the *James Webb Space Telescope - JWST*, and the wealth of galaxy surveys performed in the first few years of its operation are however capable of substantially improving our estimates of M_* at $z \sim 8 - 9$, and thus allowing us to examine whether the current paradigm of star formation still holds in the early Universe.

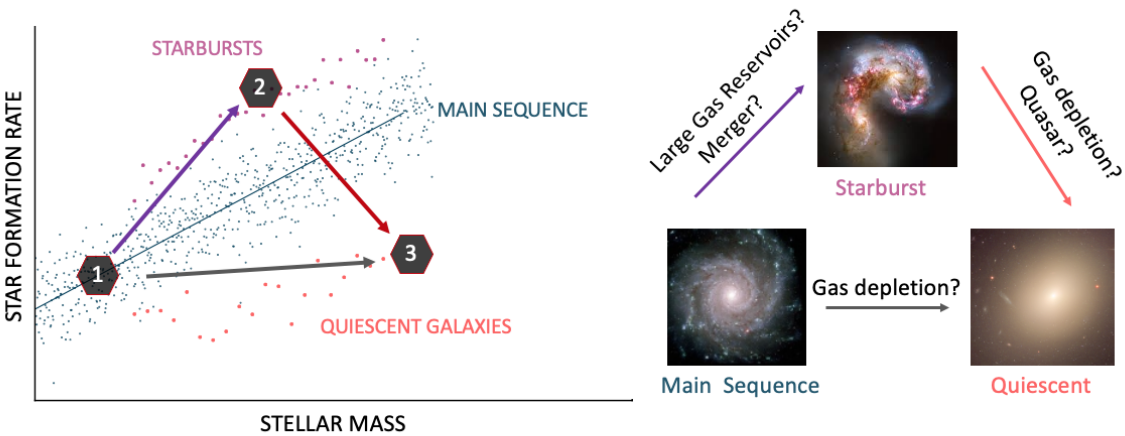

Figure 1.5: The Lives (and Deaths) of Galaxies. What processes are responsible for increase and cessation of star formation activity? **Left:** The relationship between star-formation rate and stellar mass. Main-sequence galaxies, which comprise a majority of sources at a given redshift interval, are shown in blue. Outliers above the MS that display elevated SFRs - Starburst galaxies are shown in purple. Quiescent galaxies that display little to no star formation are shown in red. **Right:** A predicted evolutionary sequence of SFGs. Galaxy images credit: NASA.

Blind surveys reveal that "normal" or MS galaxies constitute $\sim 95\%$ of the total star-formoing galaxy population at any given time, with the rest of the sources either being elevated above the MS (Rodighiero et al. 2011). The so called "starburst" galaxies (SBs) are located in the scatter

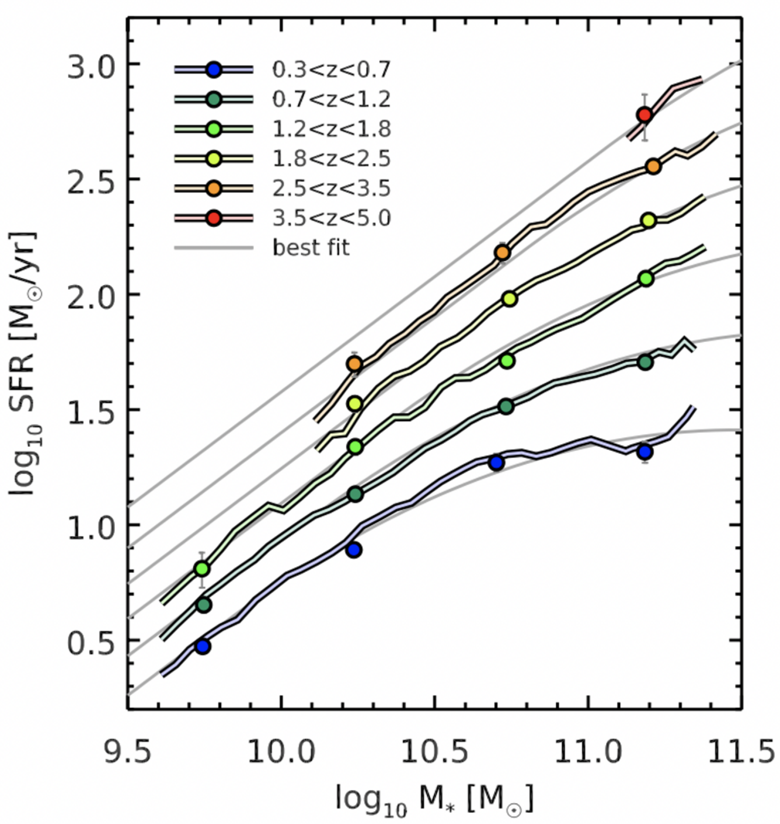

Figure 1.6: The evolution of the MS with cosmic time. Coloured points and lines shown the observed relations and a given redshift, while the grey solid lines are best fit lines to each. Figure adopted from Schreiber et al. (2015).

above the MS and display SFRs $\sim 4\times$ higher than objects at the same M_* and redshift. Starburst undergo an accelerated mode of star-formation, thought to be brought about by major galaxy mergers, during which the ISM is enriched by gas which then quickly compresses and cools to form stars (Sanders & Mirabel 1996). In stark contrast, the quiescent galaxies (QGs) located below the MS, form very few to no stars, given their already assembled masses, and are characterised by their red colours and elliptical morphology. These sources are thought to have reached the end of their life-cycle by exhausting all the available fuel for star formation, or quenched as a result of morphological effects, AGN feedback, halo quenching, gas stripping, harassment by cluster members and strangulation (Peng et al. 2015; Man & Belli 2018).

When we look at starburst galaxies in the nearby Universe, we see that they are relatively rare, and contribute very little to the integrated SFR. Morphologically, these systems appear to have undergone or currently experiencing a major merger event, and also show signs of massive extinction in the UV-optical, hinting at substantial dust obscuration (Sanders & Mirabel 1996; Rodighiero et al. 2011). Historically, starburst galaxies in the local Universe have been separated into different classes based on their observed total FIR luminosity in the 8 - 1000 μm range ($L_{\rm IR}$). These are the luminous and ultra-luminous infrared galaxies (LIRGs and ULIRGS), simply denoting objects that exceed $\log_{10}(L_{\rm IR}/L_{\odot}) > 11$ or 12, respectively. Since a vast majority of radiation emitted by dust is reprocessed starlight, the $L_{\rm IR}$ is a direct tracer of star formation, one can compute SFR $[M_{\odot}\ {\rm yr}^{-1}] \sim 10^{-10} \times L_{\rm IR}[L_{\odot}]$ (Kennicutt 1998), for a Chabrier (2003) initial mass function (IMF). As we peer deeper into the Universe, the amount of LIRGs and ULIRGs appears to increase with redshift, and at $z \sim 2-3$ these become commonplace, with the majority of the MS being populated by sources which would be considered extremely star forming at $z = 0$ (Sargent et al. 2014; Schreiber et al. 2015). This realisation can then lead us to believe that, low - z starbursts can serve as analogues of the dominant galaxy populations in the early Universe, and that the majority of star formation at high - z can potentially be caused by mergers. Naturally, the real picture is vastly more complex, and there is a vast body of both simulations and observational studies attempting to piece together these evolutionary scenarios.

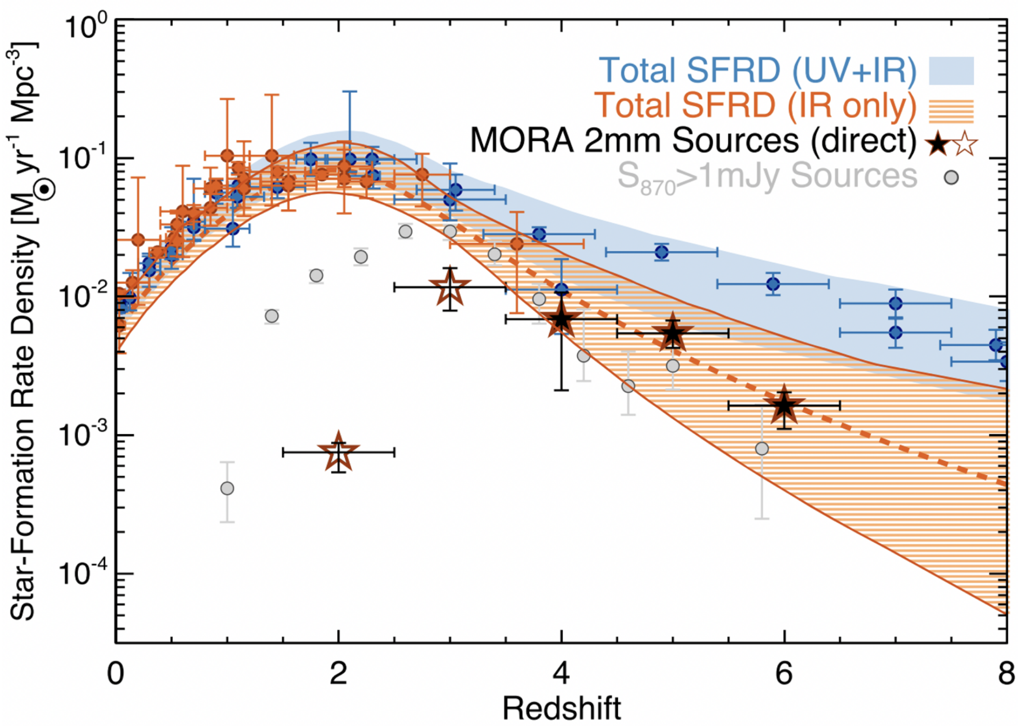

Figure 1.7: A comprehensive overview of SFRD evolution across cosmic time as calculated from different SFR tracers. Literature values from Madau & Dickinson (2014) for UV+IR, and IR only SFR measurements are shown as blue and orange points respectively. Shaded orange region shows SFRD estimates from the IR data from Zavala et al. (2021). Grey points show SFRD estimates from the *SCUBA2* sub-mm data (Dudzevičiūtė et al. 2020). MORA 2mm survey with *ALMA* results are shown as either filled black or open orange stars (Casey et al. 2021). This compilation highlights that UV - selected galaxies seem to dominate the SFRD at high-z. Figure adopted from Casey et al. (2021).

In the last decade large FIR surveys (Le Floc'h et al. 2005, 2009) have uncovered that the star formation activity in Universe was more vigorous in the past, than it is now. More specifically, observations of SFR density (SFRD; Figure 1.7) evolution present us with a Universe where the star formation has already reached its peak at $z \sim 2-3$ and is now declining (e.g. Madau

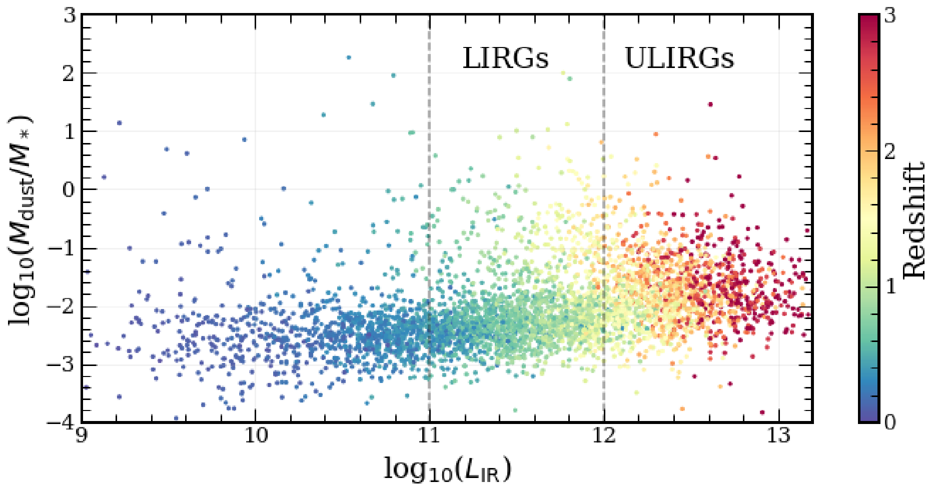

Figure 1.8: Evolution of the dust to stellar mass ratio as a function of total infrared luminosity. Galaxies shown have been collected from the COSMOS catalogue (Laigle et al. 2016; Jin et al. 2018). Colour of the points denotes redshift, while the dashed vertical lines show where the LIRGs and ULIRGs would be located based on their luminosities.

& Dickinson 2014; Casey et al. 2021). Coupled with our knowledge of the MS evolution up to $z \sim 4$ the galaxy mass assembly appears to be a largely stable process, during which galaxies gradually grow their stellar masses and deplete the available gas mass reservoirs (Elbaz et al. 2011; Sargent et al. 2014). In this framework starburst episodes appear to be playing only a minor role. Large scale cosmological and galaxy simulations hint that the compression, consumption and replenishment of available molecular hydrogen is what drives the tightness of the MS, and cause only minor oscillations within the 0.5 dex scatter in the sSFR (Tacchella et al. 2016). Given the above, the reasons which cause galaxies to move above the MS towards the mode of enhanced star formation are poorly understood. The contemporary literature envisions two possible scenarios, first starburst simply have more available fuel for star formation, and second the processes which govern starbursts allow for the existing fuel to be utilised more efficiently. One thing is clear, however, if we want to answer these fundamental questions, a detailed and thorough understanding of the M_*, $L_{\rm IR}$, dust mass ($M_{\rm dust}$), gas mass ($M_{\rm gas}$), and the overall conditions of the ISM hold the key.

The elevation of SFRD towards $z \sim 2$ and the increasing normalisation of the MS (sSFR) with time, as well as the similarities between local (U)LIRGS and high-z MS galaxies mirror the increase of the dust to stellar mass ratios ($f_{\rm dust} \equiv M_{\rm dust}/M_*$; Figure 1.8) within SFGs (e.g. Daddi et al. 2010; Geach et al. 2011; Magdis et al. 2012a,b; Tacconi et al. 2018; Liu et al. 2019a,b; Kokorev et al. 2021). Since $M_{\rm gas}$ is intrinsically linked to $M_{\rm dust}$ the increase in one would also mean the increase in the other (e.g. Mannucci et al. 2010; Magdis et al. 2012a; Scoville et al. 2016), and therefore the gas to stellar mass ratio ($f_{\rm gas} \equiv M_{\rm gas}/M_*$) also appears to grow as we move to higher redshifts.

The results in Figure 1.7 show that the contribution of dust obscured SFR, to the total SFRD seems to lessen as we move into the early Universe. More specifically, the MORA project (Casey et al. 2021) suggests that at high-z SFR is dominated by unobscured SFR, primarily through UV - selected galaxies. This might imply that that the early Universe was less dusty, which would align with our predictions regarding the time-scale and mechanisms of dust production. This, however, is still highly debatable, as the blind ALMA studies of galaxies at $z \sim 2-6$ (Gruppioni et al. 2020) show that contribution of optically dark sources to the total SFRD is significant, and is higher than that of optical or UV derived SFR. Ultimately, describing the dusty obscured SFR comes down to our ability to recover high-z dusty galaxies, such as '*HST*' - dark, or optically dark sources (Wang et al. 2016; Fudamoto et al. 2021). In general these objects are difficult to locate, as they lack UV-optical-NIR priors, and are sometimes too faint even in the FIR, to be detected by anything other than ALMA. The significance of optically dark sources and their contribution to the SFRD will be discussed in greater detail in the next sections.

In order to understand and reliably describe the complexity behind various star formation regimes we depend on robust estimates of SFR, SFE, $f_{\rm dust}$ and $f_{\rm gas}$. These in return rely on other galaxy properties, such as $L_{\rm IR}$, $M_{\rm dust}$ and M_*. The question then remains, how do we measure these parameters?

1.6 Deciphering Galaxy Emission

Galaxies, and the matter which surrounds them, will produce radiation spanning a range of all wavelengths/frequencies. Detecting and measuring this light, can therefore aid is in understanding various galaxy components, how they interact with one another, and how they change with time. A considerable amount of what we know about galaxy composition and evolution comes from modelling spectral energy distributions (SEDs), or measuring the intensities of various spectral emission and absorption lines. Using these two methods, we can compute a given set of physical parameters, and use them to derive and establish scaling relations. It is also imperative to note that tracing the observed galaxy properties and their evolution with cosmic time is tied to the wavelength coverage, and also the selection function that we inadvertently apply when measuring light emitted by galaxies.

The approach of SED fitting relies on finding template combinations that are the best representation of the photometric data. These templates vary vastly in their complexity, from simple black - body emission curves, reference galaxy spectra, and finally sophisticated, physically motivated, systems that aim to describe entire stellar populations and ISM which surrounds them. One of the major advantages of SED fitting, is its relative simplicity. In terms of data, obtaining broad-band photometry, a type most commonly utilised for this method, does not require significant integration times. Therefore, in contrast to galaxy spectra, it can be done for thousands of galaxies within a reasonable time-frame. However, a major drawback of SED fitting is having to rely on limited amount of models, which might describe the data well, but be otherwise invalid or return nonphysical results in regions where data is not available, and we have to extrapolate. Limited wavelength resolution can also be a significant issue, as no no absorption or emission features can be resolved with broad-band filters.

SED fitting can assist us with elucidating various galaxy properties such as redshift (Brammer et al. 2008; Arnouts et al. 1999; Ilbert et al. 2006), dust obscured and unobscured stellar light (Burgarella et al. 2005; Noll et al. 2009; Boquien et al. 2019; Berta et al. 2013; Battisti et al. 2019; Kokorev et al. 2021), and AGN properties (Mullaney et al. 2011; Delhaize et al. 2017; Calistro Rivera et al. 2016; Delvecchio et al. 2020). Different parts of the SED will produce radiation that corresponds to different galaxy components. Observations of the UV - optical regime $\lambda \sim 1500 - 2800$ Å will contain information about the young stellar populations, and thus can also be used as an (dust-attenuated) SFR indicator. Moving redwards, the near-infrared (NIR) light from $\lambda \sim 8000 - 30000$ Å is emitted by the bulk of the evolved stellar population, and thus is used in calculating the M_{*}. This also implies the existence of an IMF where the mass budget is dominated by low-mass stars. Beyond that, within the mid-infrared (MIR) and FIR regimes, dust heated by either AGN accretion or reprocessed stellar light will describe total contribution of the AGN towards the infrared luminosity, the $L_{\rm IR}$ (and thus SFR), as well as various ISM properties such as $M_{\rm dust}$ and $M_{\rm gas}$. The MIR also contains the polycyclic aromatic hydrocarbon (PAH) line features, which can also act as important cooling mechanisms. Finally the radio emission observed in galaxies will be a result of bremsstrahlung and the synchrotron radiation, with the latter being caused by an AGN jet, or supernovae.

1.7 Deriving Stellar Population Properties

The core of this thesis focuses specifically on the evolution of the ISM, however the full evolutionary picture would be incomplete without a thorough understanding of how the UV - stellar emission impacts the derived scaling relations. To this end, and prior to describing how the FIR/mm regime can be modelled, it would be worth to spend some time discussing how the stellar population parameters are derived.

Over the last few decades, a considerable amount of resources has been spent of extracting information from galaxy SEDs, by utilising the data from far - UV (FUV) to the FIR. During the same time, a substantial amount of progress has been made in development of the synthesis models of stellar evolution. These rely on the our theoretical knowledge of stellar evolution and can describe and constrain the properties of various stellar types, depending on their age and metallicity. These stellar population synthesis (SPS) models later on became standard when modelling UV - optical galaxy SEDs (see e.g. Charlot & Bruzual A 1991; Bruzual A. & Charlot 1993; Bressan et al. 1994; Worthey 1994; Fioc & Rocca-Volmerange 1997; Leitherer et al. 1999; Vazdekis 1999).

Most commonly SPS models contain within them a full description of a composite stellar population (CSP) (Conroy 2013). Obtaining a CSP, however, requires a few steps (Figure 1.9) which involve understanding of both stellar physics and the ISM. The beginning point all SPS templates consists in a simple stellar population (SSP), which characterises a single co-evolving population of stars, at a fixed metallicity and element abundance. To construct an SSP three components are required - the initial mass function (IMF), libraries containing description of stellar spectra and the evolutionary path of stars - the isochrones.

Isochrones denote the evolutionary tracks of stars that have similar metallicity and age on the evolutionary path - the Hertzprung - Russel (HR) diagram. Integrated over a range of all physically feasible stellar masses ($\sim 0.1 - 100$ $M_{\odot}$) these would produce a complete picture of stellar evolution in both mass and time. Generally, combining isochrones for SPS models can be complicated, due to the former spanning a wide range of ages and metallicites. There is also an issue of stellar physics itself, such as rotation, mass loss and binary systems. Modern SPS libraries take these effects into account, however some uncertainties still remain.

The stellar IMF is one of the most extensively studied properties of galaxies (Salpeter 1955; Scalo 1986, 1998; Kroupa 2001; Chabrier 2003). It describes the initial configuration and distribution of stellar masses, with respect to the galaxy MS. For the purpose of constructing the SPS models, the IMF aids in understanding how the luminosity evolves for a given stellar population, the relation between mass and light, and also effects the SEDs of both simple and composite stellar populations.

Finally, the stellar spectral libraries consist of stellar SEDs, which are either computed theoretically, by taking into account temperatures, metal abundances and gravities, or simply observed empirically, for a given stellar type. By knowing what is the initial distribution of stars (the IMF), and how they evolve (isochrones), the stellar spectra will give us an idea of how this particular population will look at different stages of its evolution. The combination of all three parts will result in an SSP.

We can then use simple stellar populations as construction blocks to describe a more complex array and diversity of stellar distributions. The key difference between CSPs and SSPs is that the former contain populations of varying ages, as denoted by their star formation hisotries (SFH), metallicities, that depend on element abundance, and finally the effects of dust attenuation are taken into account.

Star formation histories can be assumed to have different degrees of complexity. With the

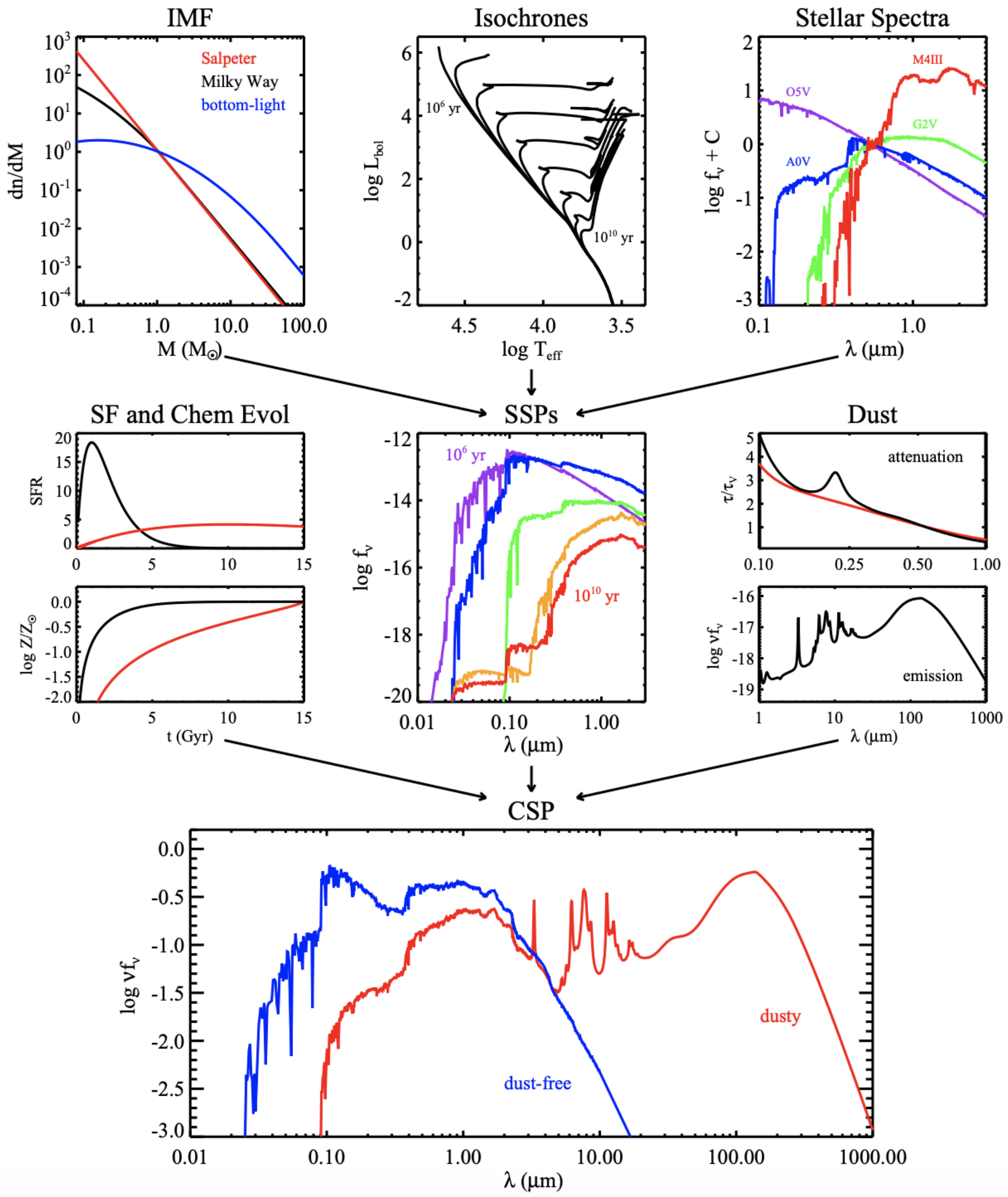

Figure 1.9: A simple layour of the SPS technique. **Top panel:** The construction of a single SSP: the IMF, stellar isochrones, and the libraries of stellar spectra. **Middle panel:** Ingredients of composite stellar populations (CSP) involve combining SSPs with the star formation histories and dust attenuation laws. **Bottom panel:** An SED of a final CSP. Blue colour shows the unatennuated stellar light, while red line displays the light reprocessed by dust. Figure adopted from Conroy (2013).

most commonly utilised model being the exponentially declining model, where SFR decays by a factor of e in some time τ, since the burst at some initial point t. This rather simple model is a result of assuming a closed system where SFR is directly proportional to the density of gas, a relation which is commonly known as the Kennicutt - Schmidt law (Schmidt 1959; Kennicutt 1989, 1998). Alternatively, and especially for the high redshift Universe, forms of SFH which rise slowly at early times and then decay at late times can also be used, but their universality is disputed

(Maraston et al. 2010; Papovich et al. 2011).

As briefly mentioned in the previous section, dust grains, locked within the ISM, will obscure the stellar light through absorption and scattering. The amount by which the starlight can be extinct can be determined from comparing the observed, attenuated stellar light, to the predicted, dust free, spectrum of a given stellar type or stellar population. The energy absorbed by dust is however not lost, and is subsequently re-emitted at a longer wavelength, with galaxy SEDs above 10 μm, being generally dominated by dust grain emission. The observed FIR emission spectrum can be then understood as a range of black bodies, in the form of dust grains. Varying sizes of these grains, their total mass, and different degrees of interstellar radiation to which they are exposed to, will then determine the shape of the FIR SED, its the total luminosity and the temperature. Similarly to the SFH the complexity of attenuation and dust emission models can vary from simple attenuation laws (e.g. Calzetti et al. 2000; Calzetti 2001) to more elaborate templates where the sizes and chemical composition of dust grains is taken into account (Weingartner & Draine 2001; Draine et al. 2007, 2014).

1.7.1 Energy Balance

In principle stellar population synthesis models can be used to describe and fit data that covers the entire wavelength range of galaxy SEDs from UV, through optical, NIR and finally FIR. In this case the energy balance approach would have to be applied, which dictates that all of the energy lost in the stellar emission part of the SED, must be re-emitted by dust in the IR regime. Indeed the majority of the SED fitting routines rely on the energy balance approach, for example CIGALE (Burgarella et al. 2005; Noll et al. 2009; Boquien et al. 2019) and MAGPHYS (Battisti et al. 2019) are the ones being utilised the most when performing panchromatic SED analysis. However, most recent high-resolution FIR observations with the Atacama Large Millimeter/submillimeter Array (ALMA), in conjunction with resolved optical data from the Hubble Space Telescope (*HST*), show that energy balance does not apply universally. In Figure 1.10, a set of cutouts of high-z SFGs display a spatial disparity between UV/optical and millimetre emission (Hodge et al. 2016). Moreover, an increasing number of luminous FIR galaxies has been found to lack or have very faint optical/NIR emission (e.g. Wang et al. 2016; Jin et al. 2019; Casey et al. 2019; Franco et al. 2020), that could additionally introduce technical challenges to the widely adopted energy balance approach.

1.8 Calculating FIR properties

The FIR physical properties of galaxies can be derived either directly from SPS templates, assuming the energy balance holds, or alternatively from models of black-body radiation or more

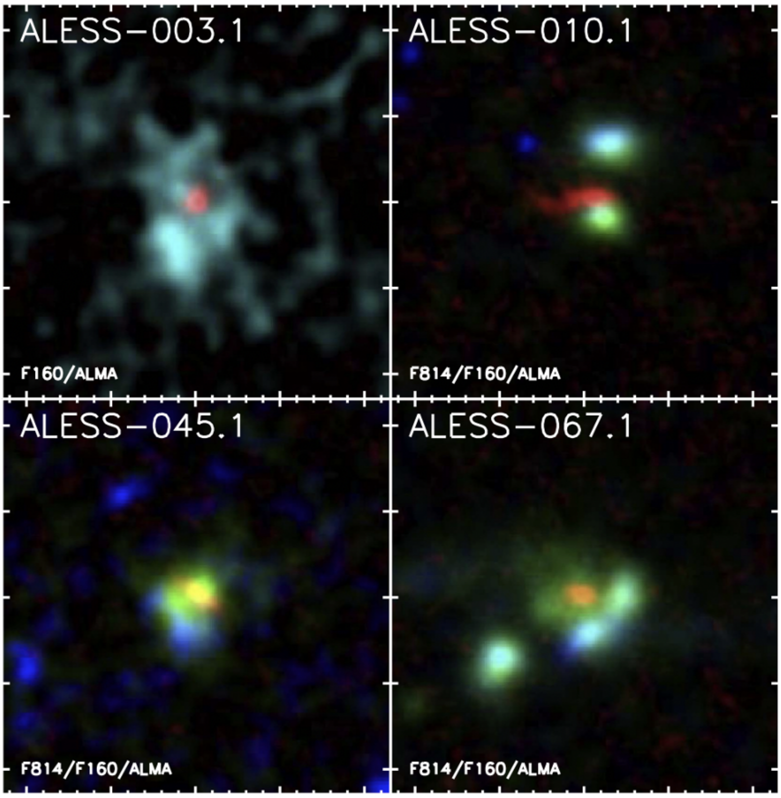

Figure 1.10: A combination of *HST* and ALMA data for a sample of high-z dusty SFGs. The colours correspond to HST F814W (green), HST F160W (blue) and ALMA 870 μm (red). The disparity between the spatial extent of stellar and dust continua might induce unwanted effects if energy balance is assumed when analysing the data. Figure adopted from Hodge et al. (2016).

sophisticated physically motivated dust grain recipes. These parameters will include L_{IR}, M_{dust}, as well as the dust temperature (T_{dust}), intensity of the radiation field induced by the stellar activity, and, models permitting, dust grain distribution.

1.8.1 Modified Black-Body

The first, and still most commonly used approach to FIR SED fitting is a black - body function. This assumption relies on dust grains to be (locally) thermalised, meaning they can be described in bulk as a single entity. Naturally, the temperature of interstellar dust can cover a wide range of temperatures, necessitating a functional form combining multiple such black bodies - the modified black body (MBB) function. For a general case this function is given as:

$$S_\nu \propto \frac{(1 - e^{-(\frac{\lambda_0}{\lambda})^\beta})(\frac{c}{\lambda})^3}{e^{hc/\lambda k_B T_{\mathrm{dust}}} - 1}, \tag{1.3}$$

where h and k_B are the Planck and Boltzmann constant respectively, λ_0 is the characteristic wavelength at which optical depth is unity, β is the emissivity index and S_ν is the flux density.

Given specific morphologies and dust densities, stellar light can be absorbed and re-emitted multiple times as it exits the dust cloud. In this case the galaxy emission, or SED is considered "optically thick", implying that the observed emission is "colder" than it actually is. Therefore λ_0 can be considered as a cutoff point, and for all values of λ below it, the emission becomes optically thick. If we assume that the SED is only optically thick for some $\lambda_0 << \lambda$, the above form can be simplified to the optically thin MBB:

$$S_\nu \propto \frac{(\frac{c}{\lambda})^{3+\beta}}{e^{hc/\lambda k_B T_{\rm dust}} - 1}. \tag{1.4}$$

From the best fit one can then calculate the FIR luminosity by integrating the MBB over the 8 - 1000 μm range, and the dust mass by using:

$$M_{\rm dust,MBB} = \frac{S_\nu D_{\rm L}^2}{(1+z)\kappa_\nu B_\nu(T_{\rm dust})}, \tag{1.5}$$

where $D_{\rm L}^2$ is the luminosity distance, κ_ν is the dust mass absorption coefficient and B_ν is the Planck function at the best fit $T_{\rm dust}$.

Parameters derived with the MBB, however have a few important caveats. Firstly, while the emission from the warm and cold ($T_{\rm dust} \sim 40K$) dust will dominate the FIR component of the SED, and indeed can be approximated by the MBB function, the same can not be said about the hot ($T_{\rm dust} > 100K$) dust. This is why, the MBB is only applicable at wavelength $> 40\mu$m, with the hot dust component most commonly being approximated by a power law - like function (e.g. see the extensive discussion in Casey et al. 2012). The other important point to note is exactly what dust temperature is being calculated by the MBB. In this case the $T_{\rm dust}$ is the luminosity weighted dust temperature - $\langle T_{\rm dust}\rangle_L$, and is largely determined by the peak of the FIR emission rather than the temperature of the star forming (cold and molecular) gas. The latter can then be inferred from the mass weighted temperature $\langle T_{\rm dust}\rangle_M$, instead. The discussions regarding the physical significance of the luminosity and mass weighted $T_{\rm dust}$ are still ongoing, however it is important to keep this distinction in mind when MBB fits are used to elucidate the ISM mass, rather than luminosity (see Appendix 2 of Scoville et al. 2016 for a more in depth discussion).

For galaxies with poor sampling of the FIR space, the optically thin approximation is generally more suitable due to having fewer free parameters, however when the sampling is sufficient, deciding on whether the thin or thick form of MBB is most applicable for a given galaxy can be complicated. If the galaxy is indeed optically thick, the $T_{\rm dust}$ derived by the optically thin approximation would be colder than it actually is, leading to erroneously low $L_{\rm IR}$ and increased $M_{\rm dust}$. In addition to that, deciding on which model is more applicable is not possible beyond the standard goodness of fit tests. Some literature works however, have noted that the dust temperatures of high-z dusty galaxies appear to deviate from the predicted evolutionary trends, unless the optically thick MBB is used (Béthermin et al. 2015; Cortzen et al. 2020).

1.8.2 Draine & Li 2007 templates

Physically motivated dust models is another commonly used technique when deriving FIR properties of galaxies. There exists a wide variety of dust models, which combine the grain size, distribution and chemical composition with stellar activity induced radiation field, in order to produce the FIR emission (Desert et al. 1990; Silva et al. 1998; Devriendt et al. 1999; Popescu et al. 2000; Dale et al. 2001; Jonsson 2006; Piovan et al. 2006; Draine et al. 2007; Groves et al. 2008; Popescu et al. 2011). From these, the most commonly used set of models are provided in the work by Draine et al. (2007), and later on expanded upon in Draine et al. (2014).

These templates describe the dust within ISM as a mix of carbonaceous and amorphous silicate grains, with their sizes corresponding to the extinction laws measured in Milky Way, SMC and LMC. One other key feature is the treatment of hot and warm dust, and the presence of PAH emission features. The fraction of PAHs is denoted by the $q_{\rm PAH}$ parameter. Within these models, dust is being treated to be optically thin throughout the FIR regime, and is separated into the two distinct components, the photo-dissociation regions (PDRs), and the diffuse ISM. The fraction of dust molecules within PDRs is given as γ. The PDRs are exposed to stellar radiation of varying intensities ($U_{\rm min}$ to $U_{\rm max}$) , while the diffuse ISM is assumed to be irradiated by a constant radiation field ($U_{\rm min}$). By combining these parameters, it is also possible to compute the average radiation field intensity ($\langle U \rangle$), which can act as a proxy for the luminosity weighted dust temperature. From the best fit, the $L_{\rm IR}$ is computed in the same way as for MBB, and the $M_{\rm dust}$ is encoded into the normalisation of the templates. A more detailed description of these templates, their applications and caveats with regard to the derived physical parameters will be presented in the next chapter.

1.9 Molecular Gas Content

Estimating the molecular gas mass is imperative in elucidating the depletion times, SFE and gas fractions, which in return are required to unravel the great variety of galaxy environments. Molecular hydrogen, a fuel for star formation, is neutral and cold. While dense gas can be heated mechanically (collisions, turbulence, shocks), by X-rays and cosmic rays, and surfaces of molecular clouds can be energised by UV photons, these would result in molecular hydrogen being dissociated, thus preventing it from forming stars. Low radiation fields are only able to induce the vibrational and rotational transitions, and therefore produce observable emission lines. These, however, are very faint, and are only observable directly within the Milky Way and nearby galaxies (Sternberg & Dalgarno 1989), making low, intermediate and high redshift studies impossible to conduct. This necessitates an alternative approach, wherein gas is not estimated directly, but through a set of tracers.

1.9.1 Gas from Dust

A few of the common approaches to derive the gas mass involve using the dust mass or Rayleigh-Jeans (RJ) continuum emission. Compared to observing emission lines, these do not require significant integration times, and can be carried out relatively inexpensively for statistically significant samples of galaxies at all redshifts.

The first approach relies on converting dust mass to gas mass, via the gas-to-dust ratio ($\delta_{\rm GDR}$). This ratio is anti-correlated with the metal abundance in the ISM, the metallicity, and has been studied both locally and at intermediate redshift (Leroy et al. 2011a; Magdis et al. 2012a). The consistency and robustness of this method largely depends on two factors, how well we can constrain the dust mass, and how well do we know the metallicity (Z). This takes the following functional form:

$$M_{\rm gas} \equiv M_{\rm HI} + M_{\rm H_2} = \delta_{\rm GDR}(Z) \times M_{\rm dust}, \tag{1.6}$$

where $M_{\rm HI}$ is the amount of neutral atomic hydrogen. It is also important to point out that, the presence of HI as one of the components of total gas mass represents a major caveat towards robust determination of molecular gas content. While when we move towards intermediate to high redshifts HI to H_2 ratio quickly becomes rather negligible (Obreschkow & Rawlings 2009), the same can not be said locally. A sufficient sampling of the rest-frame RJ emission is absolutely imperative to obtain a robust estimate of the $M_{\rm dust}$, which generally means that detections at or beyond ~ 160 μm rest-frame, are required (Berta et al. 2016; Kokorev et al. 2021). Estimating metallicity, however, is not as straightforward. This is especially true for starburst galaxies, and high-z systems, where the standard array of metallicity tracers, based on the rest-frame optical observations, might not be available, or is simply too faint, due to dust obscuration and redshift effects. Moreover, as briefly mentioned in Section 1.4, dust production and destruction mechanism also evolve with redshift, which might also alter the behaviour of $\delta_{\rm GDR}$ at high-z. Finally, even for normal MS galaxies the $\delta_{\rm GDR}$ - Z relation is also not universal, and appears to change at low values of Z (Rémy-Ruyer et al. 2014; Capak et al. 2015). Despite the apparent complexity and a wide array of possible uncertainties, gas masses derived with this method are largely consistent with other tracers, such as molecular lines, to roughly within a factor of two.

The second method to get $M_{\rm gas}$ is to use a single photometric flux density measurement from the RJ - tail, yielding specifically $M_{\rm ISM}$, rather than $M_{\rm gas}$. The feasibility of using a single measurement of the FIR continuum to calculate the gas mass was initially discussed in Eales et al. (2012); Magdis et al. (2013) (and references therein), and finally refined and verified in Scoville et al. (2014, 2016). The method relies on the assumption that optically thin dust emission at rest -

frame 850μm, at some fixed mass-weighted $T_{\rm dust}$ is linearly correlated with the dust mass, and therefore can act as a reliable predictor of the ISM mass. Initially the relation was calibrated from the local SFGs and ULIRGS, and subsequently sub-millimetre galaxies (SMGs) at $z \sim 2$, and then further expanded using more statistically significant samples of galaxies (Hughes et al. 2017). The ISM mass is obtained via:

$$\alpha_{850} \equiv \frac{L_{850}}{M_{\rm ISM}} = \kappa_{\rm ISM}(850\mu{\rm m})\, T_{\rm dust}\, \times (6.7 \pm 1.7) \times 10^{19} {\rm erg\, s^{-1} Hz^{-1}}\, M_{\odot}^{-1}, \qquad (1.7)$$

where α_{850} is the light to ISM mass ratio at 850 μm, L_{ν} is the specific luminosity at 850 μm and $\kappa_{\rm ISM}$ (850 μm) is ISM dust opacity. Following the approach from Scoville et al. (2016), the dust temperature is normally fixed at 25 K, and while it is generally applicable for MS galaxies at intermediate redshifts, recent examinations of the mass weighted dust temperature of extreme starbursts and their high-z analogues shows that $\langle T_{\rm dust}\rangle_M$ appears to be increasing (Elbaz et al. 2011; Magnelli et al. 2014; Béthermin et al. 2015; Schreiber et al. 2018). Therefore, similarly to the dust mass method, the monochromatic RJ technique becomes highly uncertain with increasing redshift, and also at very low values of Z (e.g. see Liang et al. 2018; Privon et al. 2018).

To conclude, both methods offer moderately robust and more importantly time efficient ways to compute gas/ISM masses. These have been calibrated with the emission lone based estimates of the molecular gas mass, and remain consistent up to at least 0.3 dex (Scoville et al. 2016; Hughes et al. 2017; Saintonge et al. 2018; Kaasinen et al. 2019), both locally and up to $z \sim 3$. The existence of these methods allows to observe, and study large galaxy populations and conduct studies of gas mass evolution for statistically significant samples. However, both techniques heavily rely on a number of underlying assumptions, such as the metallicity or dust temperature, that result in these methods becoming highly uncertain for extremely star-forming and/or high-z systems.

1.9.2 Emission Lines

Alternatively to using dust and RJ continuum emission it is also possible to trace molecular hydrogen content by using emission lines. As mentioned before, cold H_2 by itself does not emit any detectable radiation. Temperatures in excess of 100 K are required to excite the gas, which is significantly larger than the observed $\langle T_{\rm dust}\rangle_L$ and $\langle T_{\rm dust}\rangle_M$ of giant molecular clouds (GMCs) in which star formation takes place.

Carbon Monoxide

The most common emission line which is being used to trace molecular gas is the transition from the first excited state to the ground level of the carbon monoxide molecule ^{12}CO, J = $1 \rightarrow 0$, (CO(1-0) hereafter). Within the ISM, CO is the second most abundant molecule, after molecular hydrogen, which is a result of cooling and combination of oxygen and carbon atoms inside molecular clouds. The CO(1-0) line is located at $\nu_0 = 115.27$ GHz and is a result of collisional and radiative excitation (Carilli & Walter 2013). Energies which are required to excite CO are relatively low, with excitation temperature $T_{\rm exc} \sim 6K$, low enough not to dissociate the molecule itself. Critical density of CO is also low ($\sim 2.1 \times 10^3$ cm^{-3}) , further aided by high optical depth, making it easily detectable within GMCs by ground - based observatories. Emissions from CO have been observed to be co-spatial with H_2 (Bolatto et al. 2013), implying that the molecular hydrogen density is proportional to CO(1-0) luminosity. Indeed, since its first detection (Rickard et al. 1975), CO(1-0) has been extensively observed by modern day radio telescopes - ALMA, Northern Extended Millimeter Array (NOEMA), and the Very Large Array (VLA). This has enabled CO(1-0) luminosity and molecular gas studies at a wide array of redshifts and ISM conditions (Carilli & Walter 2013; Tacconi et al. 2013; Liu et al. 2019b) and helped calibrate alternative gas mass tracers. It is possible to convert the line luminosity of CO(1-0) to the molecular gas mass as follows:

$$M_{H_2} = \alpha_{\rm CO}\, L'_{\rm CO}, \tag{1.8}$$

where the $\alpha_{\rm CO}$ is the conversion factor (Bolatto et al. 2013). Although $\alpha_{\rm CO}$ is well calibrated for the Milky Way, local and low-z galaxies, similarly to the dust methods, the conversion factor varies substantially with Z, distance to the main sequence, gas morphology and density, and the gas kinematics (Bolatto et al. 2013; Casey et al. 2014). For example, in SB galaxies, that generally have higher Z, the $\alpha_{\rm CO}$ is ~ 0.8 [M$_\odot$ (K km s^{-1}pc^2)$^{-1}$], while normal galaxies at the same redshift would have a conversion factor that is 6 times higher (Daddi et al. 2010; Magdis et al. 2011). At the same time, dynamical mass studies of high-z, dust obscured SMGs, report $\alpha_{\rm CO}$ values which are similar to local SB galaxies (Tacconi et al. 2010; Magdis et al. 2011; Hodge et al. 2012). Moreover, CO(1-0) transition can be too faint to be detected in appreciable integration times at high-z. While with sufficiently sampled spectral line energy distributions (SLEDs) is is also possible to use higher - J lines, which are brighter, e.g. CO(2-1) and CO(3-2), it is also possible to obtain $M_{\rm H_2}$, SLEDs of SBs are also highly uncertain (Dannerbauer et al. 2009; Daddi et al. 2015; Cassata et al. 2020). Finally, theoretical (Wolfire et al. 2010; McKee & Krumholz 2010; Smith et al. 2014; Glover & Smith 2016) and observational (Allen et al. 2012; Langer et al. 2014; Pineda et al. 2017) studies hint at the existence of sizable reservoirs of CO - dark gas, which can make up to ~ 50 % of total $M_{\rm H_2}$ reservoirs.

Neutral Carbon

In recent years, the neutral carbon [CI] (^{3}P$_1$ -> ^{3}P$_0$) transition ([CI] hereafter), has emerged as an alternative tracer of gas mass. Located at $\nu_0 = 492.16$ GHz, the [CI] emission is located in a unique position which allows us to use the same facilities as for CO, yet estimate the molecular gas mass at higher redshifts. Moreover, regardless of redshift, [CI] has been argued for as a better tracer of molecular gas, especially in extreme starbursts and galaxies with high cosmic ray incidence, where in both of these cases the CO molecule would be easily dissociated (Papadopoulos et al. 2004; Bisbas et al. 2015, 2017; Papadopoulos et al. 2018). In contrast to the CO emission, that originates from inside the GMCs, [CI] was thought to be emitted only from the outer envelope of ISM, which suffers from extreme radiation fields originating from young stars, the so called photodissociation regions (PDRs). That view has, however, changed over the years, with several early works on [CI] suggesting that its emission is in fact closely related and associated with that of CO, with the two sharing similar critical densities (Ojha et al. 2001; Weiß et al. 2003). Observationally, [CI] was detected in molecular clouds of the Galactic disk, the galactic centres, and other nearby galaxies. This has further confirmed that CI is closely associated with CO(1-0), effectively this implies that [CI] is emitted from volumes comparable to CO(1-0), and similar $T_{\rm exc}$ (Ikeda et al. 2002; Israel 2005; Kamenetzky et al. 2014), and thus is capable of tracing the same amount of molecular gas. Moving to high$z$ the available samples are primarily restricted to strongly star-bursting SMGs, radio-galaxies, and quasars (QSOs), and high-z populations with poorly constrained stellar masses and optical properties (e.g. Weiß et al. 2003; Walter et al. 2011; Alaghband-Zadeh et al. 2013; Bothwell et al. 2017). Recent [CI] surveys in $z \sim 1$ have confirmed the feasibility of detecting [CI] in normal MS galaxies, with moderate SFRs, and also have shown that MS galaxies exhibit brighter [CI] line luminosity, per unit $L_{\rm IR}$, compared to their SB counterparts (Valentino et al. 2018). Similarly to CO the total $M_{\rm gas}$ can be computed as:

$$X[\mathrm{CI}] = M_{\rm gas}/L'_{\mathrm{CI}[1-0]}, \tag{1.9}$$

where $X[\mathrm{CI}]$ is the conversion factor. In contrast to CO, it has been noted that $X[\mathrm{CI}]$ is only marginally affected by radiation field intensities and metallicities. Furthermore, [CI] is optically thin reaching deeper into the molecular clouds than CO lines. While $X[\mathrm{CI}]$ is not free from uncertainties, all this this establishing [CI] as a novel and robust gas mass tracer at a variety of redshifts, and galactic environments.

Singly Ionised Carbon

The singly ionised state of carbon - CII, (or sometimes C+) $^2\mathrm{P}_{3/2} \rightarrow {}^2\mathrm{P}_{1/2}$ transition is located at a wavelength of 158 μm ($\sim$ 19 THz) and is another auspicious method to examine the conditions of the molecular gas in nearby and distant galaxies (Carilli & Walter 2013). The CII method is not as widespread as either CO or [CI], and has generally been used to trace SFR (De Looze et al. 2011) rather than $M_{\rm gas}$. Being one of the brightest fine structure lines produced by the ISM of SFGs, more luminous than both CO and CI, the ionisation potential of CII is still below that one of hydrogen, meaning that CII can be emitted from inside the neutral ISM, GMCs and ionized gas at the same time. More recent studies have suggested that similarly to [CI], the CII emission might originate from the outer envelopes of GMCs, within the PDRs (Stacey et al. 1991; Sargsyan et al. 2012; Rigopoulou et al. 2014; Cormier et al. 2015; Díaz-Santos et al. 2017; Croxall et al. 2017). Observations of CII in the Milky Way (Pineda et al. 2013; Velusamy & Langer 2014) and simulations (Vallini et al. 2015; Accurso et al. 2017; Olsen et al. 2017) have displayed that a vast majority ($\sim 70\%$) of the CII luminosity was coming from the molecular gas. Carbon can not be dissociated like the CO molecule, and thus similarly to the [CI] line, the CII emission can potentially trace the elusive "CO - dark" molecular gas in high cosmic ray rate environments, such as starbursts. Recent extragalactic observations of CII in $z = 2$ MS galaxies (Zanella et al. 2018) have revealed a moderately tight correlation of 0.3 dex between the CII luminosity and molecular gas mass, similarly to [CI], the $L'_{\rm CII}/L_{\rm IR}$ ratio seems to be higher in MS galaxies than SBs. The CII to $M_{\rm H_2}$ conversion factor - $\alpha_{\rm CII}$ was not found to deviate significantly with metallicity, depletion time (1/SFE) or redshift. This positions CII as yet another (together with [CI]) potentially convenient way to trace the gas mass in unusual, low-metallicity or high radiation field environments at all redshifts. However, there are still some major caveats related to the CII method, for example recent work by Heintz et al. (2021) claims that HI, rather than H_2 is being traced. This would become especially prominent at very high-z, where the contrast with the CMB would now allow to easily trace cold molecular gas.

1.9.3 Gas Mass Evolution

We are now able to combine the various gas mass tracers, as well as semi-analytical galaxy models into a single and coherent picture, (see Figure 1.11) and examine the evolutionary path of molecular gas (Liu et al. 2019b). Moving into the past, this picture implies that the molecular gas density slowly rises from $z \sim 0 - 3$, peaking at around $z \sim 2.5$, thus lagging slightly behind, but still mirroring the epoch where the most extreme episodes of star formation took place (Carilli & Walter 2013). Gas mass density then slowly decays as we move past the reionisaiton and closer to the Big Bang. It is important to note that while the methods that were used to calculate the molecular gas density are different, they all largely agree rather well up to $z \sim 3$. The evolutionary picture however becomes more uncertain at high-z, with rather large differences not just between CO and dust methods, but even among the dust methods themselves. Naturally, the differences between the functional forms chosen could play a role, however the prominent source of uncertainty is fact that the number of galaxies with secure detections of molecular gas beyond $z = 4$ are very few. From examining the SFRD evolution in Figure 1.7 we noted that it seems that the star formation in early Universe is dominated by the relatively dust-free, bright UV - galaxies (Casey et al. 2021). This could explain the discrepancy and uncertainty when gas masses are obtained from either RJ - continuum or dust. However, more and more works note the existence

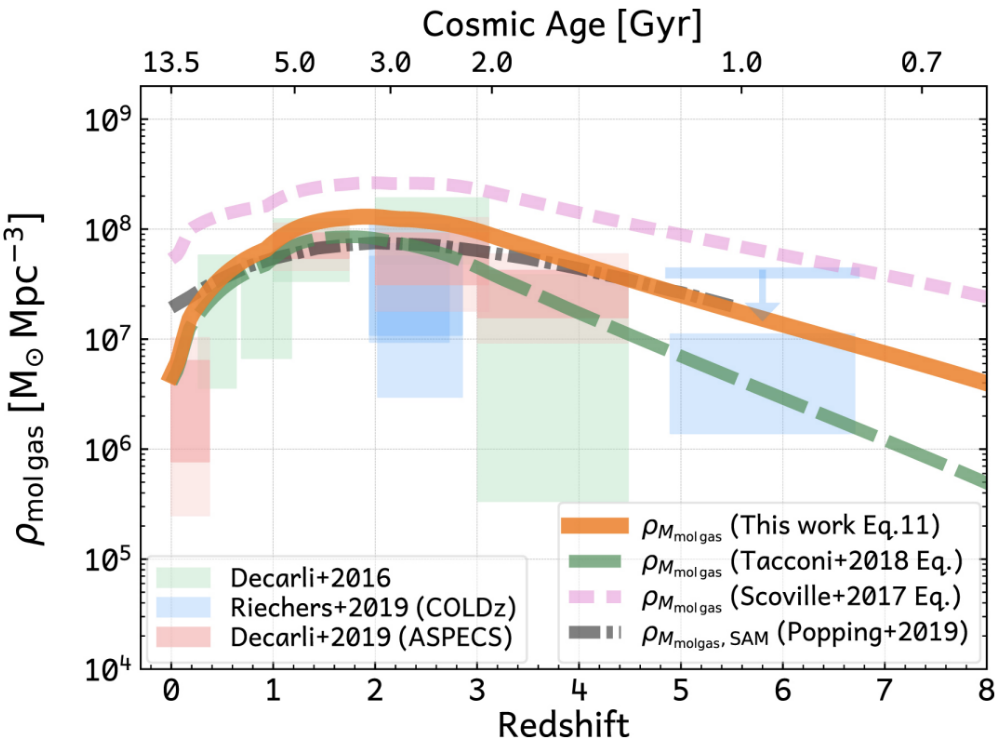

Figure 1.11: Compilation of the evolution of the molecular gas mass density from different tracers with cosmic time and redshift. The orange line shows a compilation of ALMA results from Liu et al. (2019a). Blind field studies of CO are shown in green (Decarli et al. 2016), blue (Riechers et al. 2019) and red (Decarli et al. 2019). Integrated gas mass densities computed from the evolutionary scaling relations derived by Tacconi et al. (2018) (dashed green line) and Scoville et al. (2017) (dashed pink line). Simulations from the semi - analytical model examined by Popping et al. (2019) are displayed as a dashed grey line. Figure adopted from Liu et al. (2019a).

of dusty, UV - optically dark sources, (Wang et al. 2016; Gruppioni et al. 2020; Fudamoto et al. 2021), that seem to contribute a substantial amount towards SFRD. Since a vast majority of studies presented in Figure 1.11 focus on optically detected sources, could it be possible that, similarly to SFRD, there exists a population of sources, hidden from sight, that can have a significant contribution to the gas mass density of the Universe at high-z? To answer this question, and complete the picture of the gas mass evolution in the early Universe, thorough examinations of the faint end of the cosmic IR background and IR luminosity function are required. These sources have been largely inaccessible as of recent, however it seems that blind surveys with the state of the art sub-mm and mm facilities, such as ALMA, are capable of filling the gap in resolution and sensitivity, and aid us in completing the full picture of ISM evolution.

1.10 Examining Unique Galaxy Populations

Our ability to draw correct conclusions about the visible Universe is hinged upon observing complete samples of galaxies, from the brightest, extremely star-forming local galaxies, to the faintest high-z systems. Voiding that, invites a risk of the derived scaling relations, mass and luminosity functions to be biased and results in diminishing our ability to make predictions. Since the time of the first successful deep surveys with *HST*, the size and complexity of galaxy surveys has grown exponentially, owing to the advent of new facilities, particularly in the IR-sub-mm regime, with *Spitzer Space Telescope* (Spitzer), *Herschel Space Observatory* (*Herschel*), and ALMA. However, there are still some questions that remain unanswered. The primary concerns are the objects which currently have not been identified in an appreciable capacity, these include high-z systems, both UV - bright, and more evolved dusty systems, optically dark sources, as well as faint sub-mm galaxies at $z > 4$.

1.10.1 Optically-selected High-redshift Galaxies

Ohe ability to identify high-z galaxies and measure their properties is instrumental towards our understanding of the processes that drive galaxy evolution. These properties include redshift, stellar and gas masses, star formation rates, and metallicities. These efforts were, however, complicated by the lack of reliable techniques which allow us to estimate galaxy redshifts, in absence of detailed photometric coverage, or spectral line observations. However, certain features of galaxy SEDs allow us to identify certain galaxy types, even when their photometric coverage is limited. One such feature is the Lyman break, or the Lyman limit, located at $\lambda = 912\,\text{Å}$. In the context of UV - optical astronomy, the Lyman discontinuity is a telling sign, which enables us to identify galaxies at high-z. This method relies on the fact that radiation with wavelength lower than the Lyman limit is almost completely absorbed by the neutral ISM. Galaxies identified with this technique are called Lyman break galaxies (LBGs), and will appear to be very faint, or completely disappear out of sight when observed at rest-frame blue-ward of the Lyman limit. Within the context of SEDs, this feature will manifest itself as a sudden drop in observed flux density. Since the position of the Lyman break is fixed in the rest frame, objects that are further away, will have their Lyman break position shifted redwards, thus allowing for simple redshift estimates. This technique is relatively cheap, and only relies on using two broad-band filters, and is this very effective way of identifying high-z galaxies (Steidel et al. 1996).

The Lyman break identification technique has been used to find thousands of, what at that time were considered, high-z galaxies (e.g. Shapley et al. 2003), and is still widely utilised today. Most recently it has helped to identify the most distant galaxy ever to be observed at $z = 11$ (Oesch et al. 2016). The detected LBGs appear to be highly star forming, typically blue and disk-like, indicating relatively young systems (see e.g. Pettini et al. 2001; Shapley et al. 2001; Giavalisco 2002), and can potentially act as progenitors of low-z massive elliptical galaxies. The major source of uncertainty for the Lyman break detected galaxies, is what is known as the break

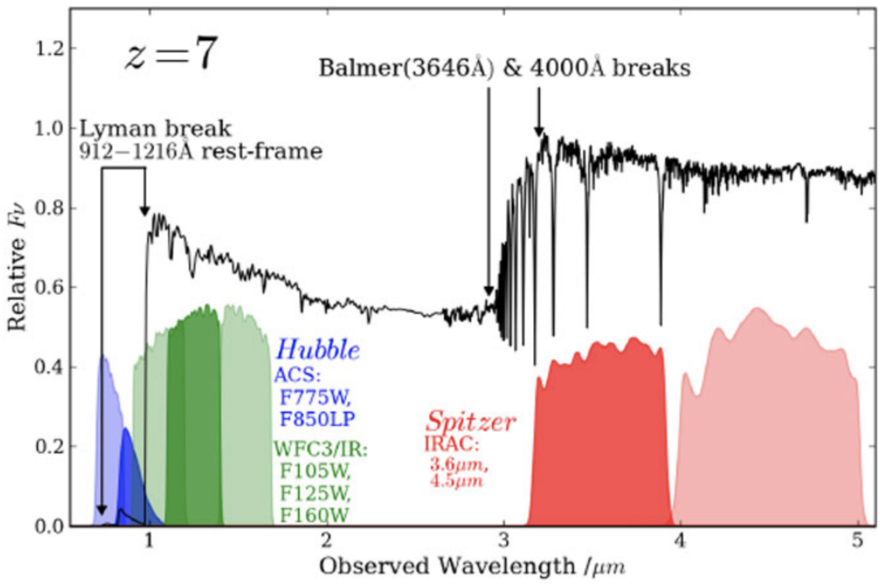

Figure 1.12: An example of an SED of a young star-forming galaxy at $z = 7$, with two distinct breaks. The Lyman break is a result of neutral gas absorbing light with wavelengths leftwards of the Lyman break. The Balmer breaks result from the bound-free absorption within the hydrogen atom. The key UV and NIR filters of *HST* are shown as shaded blue and green regions, highlighting the theoretical required coverage to successfully identify such a galaxy. The *Spitzer*/IRAC broad-band coverage is shown as shaded red regions, covering the Balmer break. Most notably, the gap between the *HST* and *Spitzer* filters, which would be instrumental in identifying of $z > 10$ galaxies will not be covered by any space based observatory, until the successful calibration of *JWST*. Figure credit: James S. Dunlop, "The First Galaxies", Vol 396 (2013).

confusion. With limited photometric coverage it sometimes becomes impossible to distinguish between the Lyman break at $\lambda = 912\,\text{Å}$ and the Balmer break, which is located at a longer wavelength of $\lambda = 3645\,\text{Å}$. Moreover, the dust absorption itself can mimic a break-like feature, in fact most of the contaminants among the high-z LBGs are low-z dusty objects. Despite these uncertainties, however the Lyman break method was shown to be effective at identifying high numbers of potential high-z candidates, which can later on be used to guide future spectroscopic follow-ups as well as photometric redshift template fitting routines.

1.10.2 Optically Dark Galaxies

In Section 1.5 "*HST* - dark" objects were briefly introduced, alongside the notion that the majority of the SFRD contribution at $z > 3$ is derived from objects detected in rest - frame UV (Casey et al. 2021), with very few of these objects containing significant amounts of dust. By that time, however, almost 2 billion years after the Big Bang, one would expect the most massive galaxies to already evolve and produce large amounts of dust. In return, this could potentially mean the the UV - selected population is not a complete representation of the Universe at $z > 3$.

Recent detections of optically - dark sources (Wang et al. 2016, 2019; Gruppioni et al. 2020), argue that the existence of such systems can potentially create obstacles in our understanding

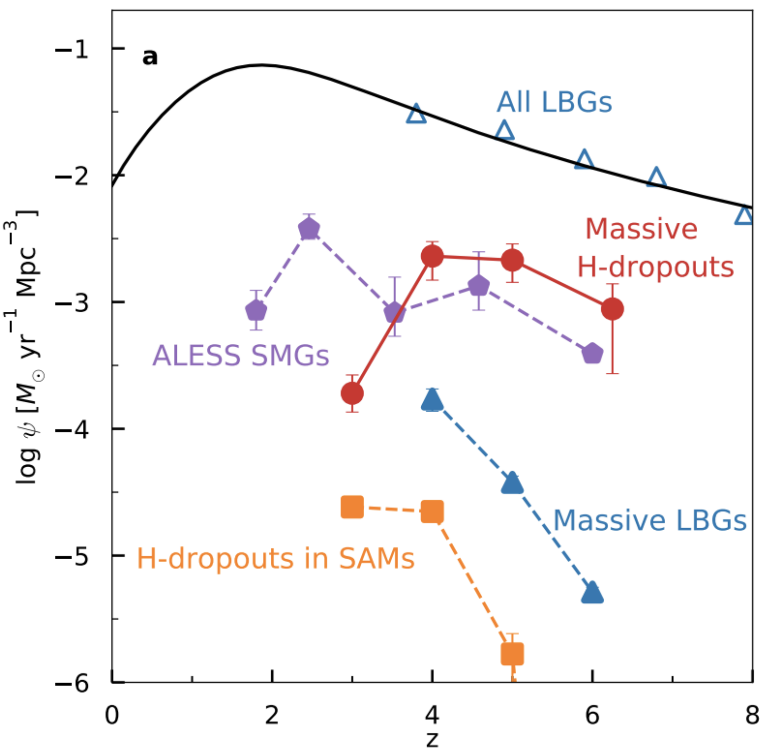

Figure 1.13: The contribution of optically dark sources to the total SFRD. The solid black line represents the current paradigm of the cosmic SFRD history, based on the Lyman break (LBGs) identified objects. Contribution of massive LBGs is shown as the blue triangles. The optically dark population is shown as the red circles. Massive SMGs are shown as purple pentagons, these are dusty and star formation, yet insufficient to explain the full picture at $z > 3$. Orange squares show the predicted contribution of the optically dark sources based on the semi analytical models. Figure adopted from Wang et al. (2019).

of the true numbers of massive galaxies and SFRD in the early Universe. These sources appear to contribute ten times more towards SFRD, than similarly bright galaxies with a rest-frame UV detection, and reside within centres of the most massive matter overdensities (see Figure 1.13). Implying that the optically dark galaxies that we see at $z > 3$ could potentially be progenitors of bright central galaxies (BCGs) within galaxy clusters. In addition to that, the existence of high number of massive, evolved, and dusty sources only a few billion years after the Big Bang challenges our current understanding of galaxy formation.

Detection and identification of these objects can be a significant challenge, even when sufficient IR coverage is available. As mentioned before they are completely invisible in UV - optical and NIR light. The brightest, most massive optically dark sources can be detected by *Spitzer* IRAC, however poor resolution can result in blending and erroneous assignment of flux if nearby objects are present. Generally, the existing *Herschel* maps do not reach the required sensitivity to produce robust photometry at $z > 3$, with only a limited sample of secure detections. The lack of the UV - optical continuum emission in "*HST* - dark" makes it impossible to directly compute the stellar mass, and (optically derived) metallicity. It also renders any photometric redshift technique unfeasible, with the spectral line scan remaining as the only option to reliably secure redshifts. As a result the only reliable way to observe such objects are blind field studies with ALMA or NOEMA. Later on in this thesis, I will discuss the detection and analysis of a few promising candidates for the optically dark galaxies within the ALMA lensing cluster survey (ALCS) fields.

1.10.3 Faint Sub-mm Galaxies

A complete understanding and characterisation of the cosmic infrared background (CIB) has existed as a significant challenge for modern astronomy, since the time of its first conceptualisation. Significant amounts of energy are emitted in the IR, and were found to be similar to those of the cosmic optical background. The previous section introduced high-z LBGs, in the context of largely unexplored galaxy populations. These objects appear to have moderately high SFRs, as evidenced by their bright UV-optical rest frame luminosities, thus seemingly lacking any significant dust content. In Section 1.4, the connection between star formation and large amount of dust has been discussed, this in conjunction with the existence of optically dark objects shows the importance of the IR to mm/submm observations, implying that large portions of the cosmic SFRD are dust obscured (Lagache et al. 2005). High redshift deep surveys with sub-mm facilities, beginning with Submillimeter Common-User Bolometer Array (SCUBA; Holland et al. 1999), and later on using ground and space based single dish telescope like James Clark Maxwell Telescope (JCMT) and *Herschel* have identified bright ($f_{1\mathrm{mm}}$ >1 mJy) sub-mm galaxies (SMGs). Most of these have been associated with major merger events, and display extremely high SFRs, and high densities of molecular gas in their cores (Greve et al. 2005; Tacconi et al. 2006, 2008; Michałowski et al. 2012; Wiklind et al. 2014) Naturally, these objects are individually fascinating, however their contribution to the CIB is relatively insignificant. Therefore significant portions of star formation and molecular gas are either partially or completely missing from the current evolutionary picture. To alleviate this, ALMA has been used extensively to uncover populations of faint mm/submm galaxies ($f_{1\mathrm{mm}} \sim 0.02 - 1$ mJy), which are much fainter compared to the "standard" SMGs (Hatsukade et al. 2013, 2016; Ono et al. 2014; Carniani et al. 2015; Oteo et al. 2016; Fujimoto et al. 2016; Aravena et al. 2016; Dunlop et al. 2017; González-López et al. 2017; Franco et al. 2018). Despite numerous observations of faint sub-mm emitters, the origins of the CIB largely remain shrouded in mystery. This is further corroborated by the discrepancy of the mm source number counts in the faint-end slope of the CIB, as shown in Figure 1.14.

More specifically, in their work Fujimoto et al. (2016), show that faint mm/sub-mm sources contribute up to 25 % to the total CIB. This is similar to the other studies that report identifying objects adding up to 60 - 100 % of the total CIB contribution in the $f_{1\mathrm{mm}} \sim 0.02$ mJy regime (Carniani et al. 2015; Hatsukade et al. 2016; Aravena et al. 2016). Less optimistically, recent ALMA observations of even fainter (< 0.013 mJy) sub-mm galaxies, shows that this population is

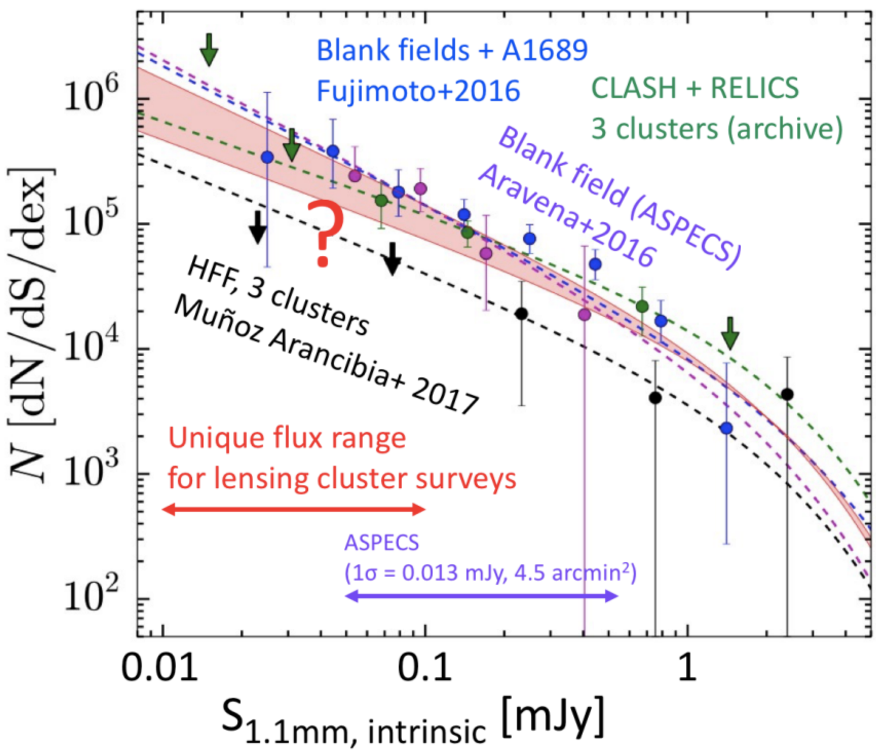

Figure 1.14: ALMA source counts in the blank (non - lensed) fields, and lensing clusters. The faint mm regime below 0.1 mJy reveals a significant $\sim$ 1 dex discrepancy to the source counts. The red line represents the best fit Schechter function to all available data. Figure credit: Kotaro Konho, ALCS collaboration.

only $\sim 15\%$ complete. This again, naturally begs the question as to how many objects are we missing exactly, what is their contribution to the cosmic SFRD, as well as the total molecular gas mass budget, especially as high-z. It appears that currently only ALMA is capable at resolving this problem, and the most promising and effective way to accomplish this task is to observe a sufficiently large number of lensing clusters.

1.10.4 Gravitational Lensing

Deep observations of the early Universe with *HST* have consistently furthered our understanding of galaxy formation and evolution. Studies of *HST* deep fields have manifested a number of methods to confirm galaxy redshifts which are now standardised. These include the previously discussed Lyman break dropout selection, photometric redshifts, that depend on SPS template fitting to multi-wavelength photometry, and the detection of high-z transients. However, as we move towards $z > 4$, obtaining reliable multi-band photometry becomes increasingly more complicated, and the populations that we can identify with the Lyman break method become biased towards the brightest and most massive galaxies. At the moment of writing of this thesis, the *James Webb Space Telescope* (*JWST*) has already launched, however extensive surveys of the

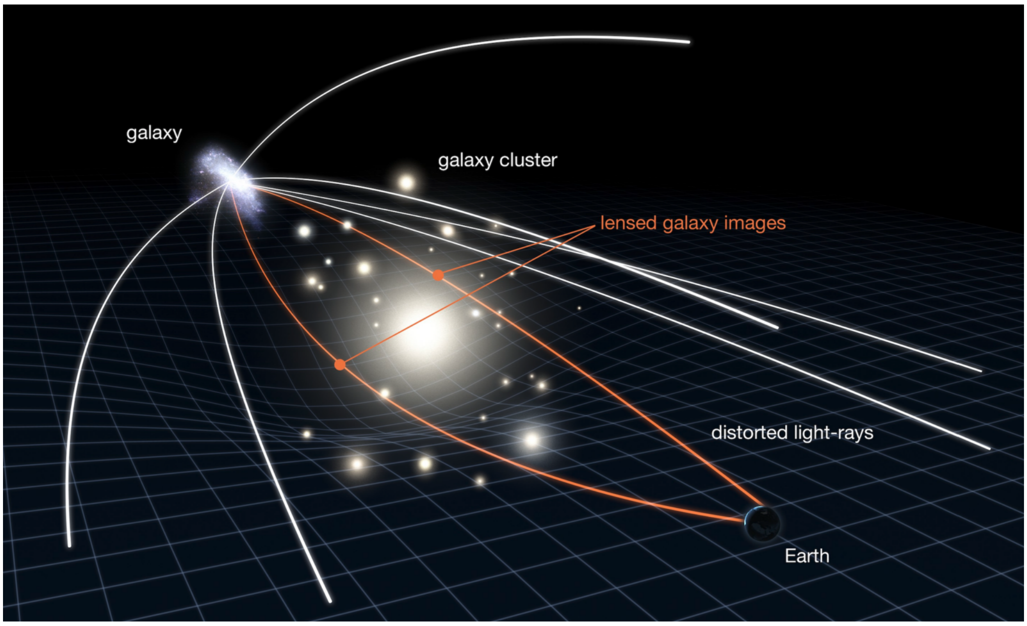

Figure 1.15: An illustration of the gravitational lensing phenomenon. Massive galaxy cluster between the target galaxy and Earth, bends spacetime, thus altering the light geodesics. Radiation from the target galaxy no longer follows a straight line and is bent, thus amplifying the amount detected of the detected light. Figure credit: NASA, ESA L. Calcada.

sky, at the required sensitivity are still years away. As such, the progress towards unraveling the mysteries of the first few billion years after the Big Bang, still falls on the shoulders of *HST*. Without any new servicing missions, or planned technical updates to the *HST*, is it still possible to push the redshift frontier even further?

Several of the *HST* and *Spitzer* projects proposed a new strategy, to observe farther and fainter galaxies. This approach relies on using galaxy clusters, as massive space telescopes, to amplify the light from galaxies very far away. Massive galaxy clusters, are one of the heaviest structures in the observable Universe. Located at the knots of the cosmic web, these collections of galaxies have the potential to bend space-time through their sheer mass. This phenomenon is called gravitational lensing, and is a direct consequence of theory of general relativity (Einstein 1908, 1907). As can be seen in Figure 1.15, distortions in the fabric of space-time can form highly effective lenses, and, if the conditions are right, allow to magnify the light behind them (Kneib & Natarajan 2011). The amount of magnification can range from a factor of a few, arcminutes away from BCGs, to a factors of hundreds or even thousands, along the critical curves, where the geometry of the lens is almost perfect. Using these lenses, within strongly magnified fields, in conjunction with space based and ground based observatories with sufficient resolution can allow us to probe intrinsically faint, and high-z galaxies, at much shorter integration times, provided they fall inside high magnification regions.

The advantages of utilising gravitational lenses have already been demonstrated with *HST* within the Hubble Frontier Fields (HFF; Lotz et al. 2017, see Figure 1.16), the Cluster Lensing and Supernova Survey (CLASH; Postman et al. 2012), and the Reionization Lensing Cluster Survey (RELICS; Coe et al. 2019). For example, only in a course of a few orbits, multiple $z > 9$ candidates have been observed to be magnified by the CLASH clusters (Zheng et al. 2012; Coe et al. 2013; Bouwens et al. 2014). Lensing does not appear to only benefit the UV - optical observations, as demonstrated by the observation of [CII] line in a very distant galaxy at $z = 6.08$ (Fujimoto et al. 2021).

Detecting intrinsically faint or very distant galaxies in lensed fields is unfortunately not a straightforward process, and is mostly down to luck. Specific lens configuration, such as mass, mass distribution and redshift, are required to detect galaxies at particular distances away from us, therefore in order to find unique distant objects, one has to rely on wide area observations covering the entire field, and its surroundings. Large scale surveys, covering tens of arcmin^2 would be impossible to conduct with *JWST* due to its limited field of view, so until the lunch of Euclid space telescope, identification of highly magnified objects would still be limited to *HST* and *ALMA*. Another good source to find lensed objects would be to look at the extremely bright tail of either the *Herschel* or *Planck* sources. These facilities can therefore serve as a powerful tool, and guide future follow up observations with *JWST* and other upcoming telescopes.

1.11 Thesis Outline

This thesis mainly focuses on the construction and analysis of large multi-wavelength galaxy catalogues, and the thorough examination of various galaxy populations, through the data across the entirety of the electromagnetic spectrum. The methods and long standing questions presented in the previous sections also lay a groundwork towards the analysis and the motivations behind the work which will be presented in the next sections. The following two chapters contain two articles, with the first one already published in the peer-reviewed scientific journal, and the second one awaiting submission.

In Chapter 2, I present the motivations and methodology behind a novel SED fitting algorithm - `Stardust`, as well as the results of its first application to the deepest and largest FIR galaxy catalogues to date. I start by discussing the caveats of analysing large galaxy catalogues, from the point of SED fitting, and the necessary steps that should be undertaken in order to ensure a robust and unbiased extraction of various galaxy properties. Following that, I provide a brief description of the software and its technical capabilities. I combine the available literature data about the existing SED fitting routines and discuss the motivation behind developing `Stardust`. The first application and results from the code are then discussed. I will report the emerging trends that capture the evolution of the far-IR physical properties of star forming galaxies over

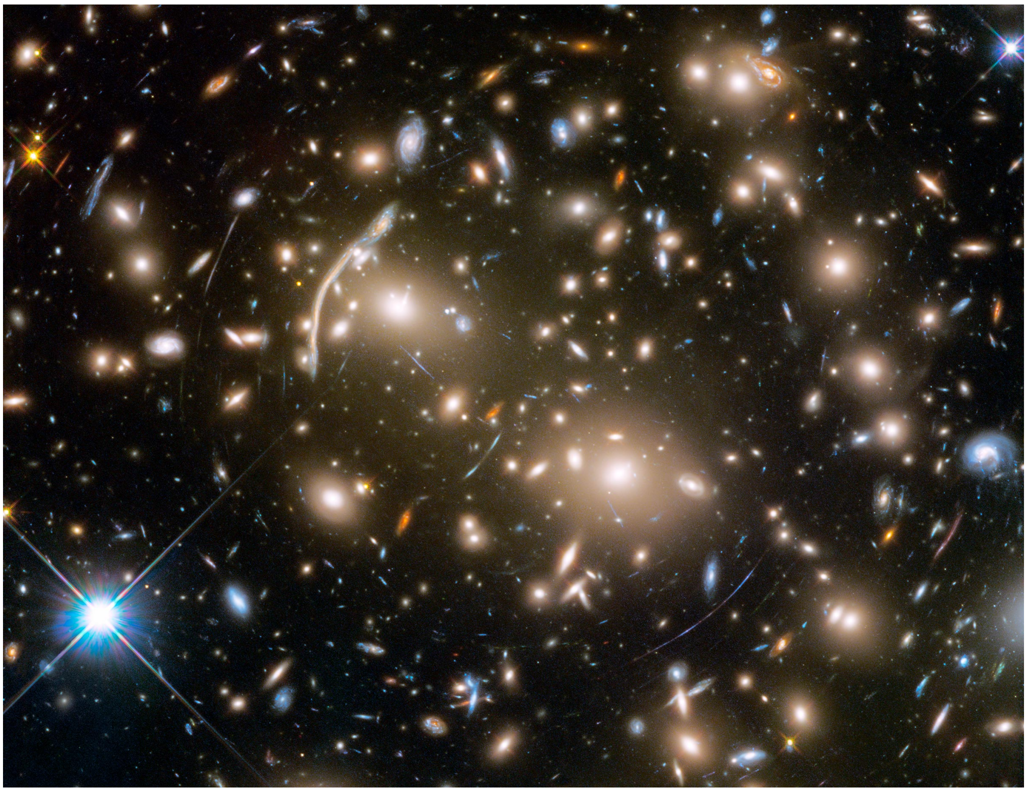

Figure 1.16: Galaxy cluster Abell 370, located within the Hubble Frontier Fields as seen on a composite *HST* image. Extended arcs surrounding the central clump of galaxies, are in fact distorted images of galaxies behind the cluster, for which the light has been bent by extreme gravitational fields of the galaxy cluster. These objects would be impossible to detect directly, however they are greatly magnified thanks to this astrophysical phenomenon. Figure credit: NASA, ESA, and J. Lotz and the HFF Team (STScI).

the last 12 billion years, including the evolution of the dust to stellar mass ratio, of the dust mass function, and of the gas fraction. This study presents one of the deepest and most complete examinations of $f_{\rm dust}$ and $f_{\rm gas}$ evolution, and can be used to guide further telescope observations, as well as simulations. The derived scaling relations are then used for an indirect computation of the dust mass function (DMF) from the stellar mass function (SMF) for the COSMOS field. Finally I will focus on a discovery and the discussion regarding a possible origin of unique population of "gas-giants", a sub-sample of galaxies within our catalogues with unusually high $f_{\rm dust}$ and $f_{\rm gas}$.

In Chapter 3, I will shift focus from SED fitting and will discuss the uniform processing of the archival *Hubble Space Telescope* and *Spitzer*/IRAC data in the ALMA lensing cluster survey (ALCS) fields. This includes both the creation of new mosaics from all archival exposures, and the extraction of photometry. The final catalogues contain 218,000 sources, covering a combined area of 690 arcmin^2. These catalogs will serve as an important tool in aiding the search of the sub-mm galaxies in future ALMA surveys, as well as follow ups of the *HST* dark - IRAC sources. Coupled with the available *HST* photometry the addition of the 3.6 and 4.5 μm bands it will allow us to place a better constraint on photometric redshifts and stellar masses of these objects, thus giving us an opportunity to identify high-redshift candidates for spectroscopic follow ups and answer the important questions regarding the epoch of reionisation and formation of first galaxies.

In Chapter 4, I will discuss my most recent project, which focuses on predictions for the COSMOS - Webb survey. This survey consists in 200 hours of observing time, over 0.6 square degrees—with the Near-Infrared Camera NIRCam plus a smaller area with the Mid-Infrared Instrument (MIRI), an area roughly equivalent to the size of the full Moon. Within NIRCam it will result in multi-band photometry for half a million sources, and tens of thousands of potential MIRI detections. My work focuses on multiple aspects of the MIRI part of COSMOS - Webb, mainly on predicting how many objects will end up being observed, and their properties. Using the existing multi-wavelength catalogues in the COSMOS field, in conjunction with SED fitting codes such as `EAZY` and `Stardust`, I predict the flux in the F770W filter of MIRI at 7.7 μm, for around 200,000 objects falling within the NIRCam area. I then conduct a Monte-Carlo simulation, by randomly positioning MIRI pointings within the NIRCam map, to calculate exactly how many objects, above the sensitivity limit of the survey we should expect to see. Finally, I model all the galaxies within the catalogue, and produce a full simulated image of the sky, as it would be seen by MIRI/F770W. This simulated image can then be used to guide the photometry extraction algorithm, which later on would be applied to the actual observations.

I finalise my doctorate thesis in Chapter 6, where I summarise my work over the last few years, and present the ongoing and potential future research directions.

CHAPTER 2

Deriving Galaxy Properties Through Their Spectral Energy Distributions

This chapter contains the following article:
"The Evolving Interstellar Medium of Star-Forming Galaxies, as traced by Stardust".

Published in The Astrophysical Journal, Volume 921, Issue 1, November 2021

Authors: *Vasily Kokorev, Georgios E. Magdis, Iary Davidzon, Gabriel Brammer, Francesco Valentino, Emanuele Daddi, Laure Ciesla, Daizhong Liu, Shuowen Jin, Isabella Cortzen, Ivan Delvecchio, Carlos Gómez-Guijarro, Mark Sargent, Sune Toft & John R. Weaver*

2.1 Abstract

In this work I analyse the far-infrared properties of $\sim$5,000 star-forming galaxies at $z < 4.5$, drawn from the deepest, super-deblended catalogues in the GOODS-N and COSMOS fields. I develop a novel panchromatic SED fitting algorithm, `Stardust`, that models the emission from stars, AGN, and infrared dust emission, without relying on energy balance assumptions. My code provides robust estimates of the UV-optical and FIR physical parameters, such as the stellar mass (M_*), dust mass ($M_{\rm dust}$), infrared luminosities ($L_{\rm IR}$) arising from AGN and star formation activity, and the average intensity of the interstellar radiation field ($\langle U \rangle$). Through a set of simulations I quantify the completeness of the data in terms of $M_{\rm dust}$, $L_{\rm IR}$ and $\langle U \rangle$, and subsequently characterise the distribution and evolution of these parameters with redshift. I focus on the dust-to-stellar mass ratio ($f_{\rm dust}$), which I parametrise as a function of cosmic age, stellar mass, and specific star formation rate. The $f_{\rm dust}$ is found to increase by a factor of 10 from $z = 0$ to $z = 2$ and appears to remain flat at higher$-z$, mirroring the evolution of the gas fraction. I also find a growing fraction of warm to cold dust with increasing distance from the main sequence, indicative of more intense interstellar radiation fields, higher star formation efficiencies and more compact star forming regions for starburst galaxies. Finally, I construct the dust mass functions (DMF) of star-forming galaxies up to $z = 1$ by transforming the stellar mass function to DMF through the scaling relations derived here. The evolution of $f_{\rm dust}$ and the recovered DMFs are in good agreement with the theoretical predictions of the Horizon-AGN and IllustrisTNG simulations.

2.2 Introduction

One of the main mechanisms driving galaxy evolution is the interaction between the interstellar medium (ISM), primarily consisting of gas and dust, and the radiation field induced by stellar activity. In this context, dust poses challenges in the detection of the UV/optical emission of galaxies and in the interpretation of these observations in terms of physical properties (e.g. star formation rate (SFR), stellar mass M_*, etc), but also is an important tracer of star formation and ISM in the FIR observations. At the same time, dust shields cold molecular hydrogen from ionising photons, and facilitates the collapse of molecular gas and subsequent star formation (Goldsmith 2001; Krumholz et al. 2011; Narayanan et al. 2011, 2012; Narayanan & Davé 2012). As such, dust plays a critical role in the life cycle of galaxies and offers observational signatures regarding their past evolutionary stages.

The impressive variety of infrared and millimetre facilities commissioned in the last few decades have propelled the extragalactic ISM studies at an ever increasing number, redshift, and detail. Indeed, the enormous observational efforts manifested by the large far-IR/mm imaging and spectroscopic surveys such as PEP (Lutz et al. 2011), GOODS-Herschel (Elbaz et al. 2011), PHIBBS (Tacconi et al. 2013), S2CLS (Geach et al. 2017) and many others (e.g. Oliver et al. 2012; Magnelli et al. 2013; Walter et al. 2016; Saintonge et al. 2017; Dunlop et al. 2017; Maddox et al. 2018; Franco et al. 2020; Valentino et al. 2020; Béthermin et al. 2020; Reiter et al. 2020, for a review see Carilli & Walter 2013 and Hodge & da Cunha 2020) have yielded a wealth of multi-wavelength data sets and have advanced our understanding of galaxy evolution through scaling relations that have been used to guide simulations and theoretical models (e.g. Dekel et al. 2009; Popping et al. 2014; Narayanan et al. 2015; Lagos et al. 2015, 2020; Popping et al. 2017; Davé et al. 2017, 2020).

In the evolutionary picture that is emerging from the analysis of the FIR/mm surveys, the majority of star-forming galaxies (SFGs) follow a tight relation – the Main Sequence (MS) – between the SFR and M_* with an increasing normalisation factor (specific star-formation rate, sSFR= SFR/M_*) at least up to $z \sim 4$ (Brinchmann et al. 2004; Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007; Salim et al. 2007; Chen et al. 2009; Pannella et al. 2009; Santini et al. 2009; Daddi et al. 2010; Elbaz et al. 2010; Oliver et al. 2010; Magdis et al. 2010; Elbaz et al. 2011; Karim et al. 2011; Rodighiero et al. 2011; Shim et al. 2011; Lee et al. 2012; Reddy et al. 2012; Salmi et al. 2012; Whitaker et al. 2012; Zahid et al. 2012; Kashino et al. 2013; Moustakas et al. 2013; Rodighiero et al. 2014; Sargent et al. 2014; Steinhardt et al. 2014; Sobral et al. 2014; Speagle et al. 2014; Whitaker et al. 2014; Lee et al. 2015; Schreiber et al. 2015; Shivaei et al. 2015; Tasca et al. 2015; Whitaker et al. 2015; Erfanianfar et al. 2016; Kurczynski et al. 2016; Santini et al. 2017; Pearson et al. 2018; Leslie et al. 2020). This elevation of sSFR with lookback-time, which broadly mirrors the overall increase of the star-formation rate density of the Universe from $z = 0$ to $z = 2 - 3$, is also followed by a similar rise in the gas fraction ($f_{\rm gas}$ = $M_{\rm gas}/M_*$) of SFGs (e.g. Daddi et al. 2010; Geach et al. 2011; Magdis et al. 2012a,b; Tacconi et al. 2013; Liu et al. 2019a,b). Nevertheless, for fixed M_*, the increase in star-formation efficiencies (SFE=SFR/$M_{\rm gas}$) surpasses that in $M_{\rm gas}$, resulting in higher SFR, an activity that coupled with the observed M_*– size evolution (e.g. van der Wel et al. 2014) instils warmer luminosity-weighted dust temperatures of the ISM as a function of redshift (e.g. Magdis et al. 2012a, 2017; Magnelli et al. 2014; Béthermin et al. 2015; Casey et al. 2018; Schreiber et al. 2018; Liang et al. 2019; Cortzen et al. 2020).

Although the purity of the MS as a measure of the evolutionary stage of a galaxy has recently been challenged (e.g. Elbaz et al. 2018; Puglisi et al. 2019; Valentino et al. 2020), it appears that the majority of SFGs grow along the MS by secularly converting (and hence depleting) their gas mass reservoirs into stars (e.g. Daddi et al. 2010; Davé et al. 2012; Lilly et al. 2013; Tacchella et al. 2016), with a high degree of uniformity in the properties of their ISM (at fixed redshift). On the other hand, galaxies above the MS (starbursts; hereafter SBs) are primarily characterised by elevated sSFR, SFE and dust temperature ($T_{\rm d}$) with respect to the average star-forming population at their corresponding redshift (e.g. Daddi et al. 2010; Magdis et al. 2012a; Magnelli et al. 2014; Scoville et al. 2017; Tacconi et al. 2018; Silverman et al. 2018). Galaxies below the MS are mainly post starbursts or quenched systems with low levels of star-formation activity, low gas fractions and cold $T_{\rm d}$ (e.g. Williams et al. 2020; Magdis et al. 2021).

At the core of the aforementioned results is the robustness of the derived FIR properties, namely the total infrared luminosity ($L_{\rm IR}$), the dust mass ($M_{\rm dust}$), the mean intensity of the radiation field ($\langle U \rangle \propto L_{\rm IR}/M_{\rm dust}$) and $T_{\rm d}$. These quantities and their emerging evolution with redshift rely on the availability of FIR/mm data and on selection effects. In this regard, while the ongoing ALMA observations are quickly filling the gap in resolution and sensitivity between the available UV/optical/NIR (sub-arcsecond) data and the coarse resolution of the confusion-limited SCUBA2 and SPIRE/surveys, the vast majority of the SFG samples with available FIR photometry are still restricted to the latter. Thus, FIR studies are still largely focusing either on the FIR luminous and most massive SFGs, on limited and possibly non-homogeneous or biased ALMA samples, or on stacking techniques.

Moreover, the derived measurements of $M_{\rm dust}$ and $T_{\rm d}$ heavily rely on the adopted models and fitting techniques (e.g. Dale et al. 2014; Berta et al. 2016). Indeed, without a coherent and homogeneous treatment of the data sets it is impossible to overcome systematic effects that could distort the observed trends. On top of that, recent high-resolution observations with ALMA indicate that the UV/optical and millimetre emission of some of high-z SFGs are spatially distinct (e.g. Hodge et al. 2016), posing challenges to the widely adopted energy balance assumption that is inherent in most multi-wavelength fitting codes. Similarly, there is an ever increasing number of IR bright yet optically faint/dark sources (e.g. Wang et al. 2016; Jin et al. 2019; Casey et al. 2019; Franco et al. 2020) that an energy balance approach would have technical difficulties to accommodate.

Finally, while many studies have focused on $f_{\rm gas}$, the evolution of the dust fraction ($f_{\rm dust}$ = $M_{\rm dust}$/M$_*$) and the dust mass function (DMF) with redshift have not been scrutinised to the same extent (Dunne et al. 2011; Magdis et al. 2012a; Santini et al. 2014; Tan et al. 2014; Béthermin et al. 2015; Driver et al. 2018; Magnelli et al. 2020; Donevski et al. 2020). Given that the use of $M_{\rm dust}$ as a proxy of M$_{\rm gas}$ either through the metallicity dependent dust-to-gas mass ratio technique (e.g. Leroy et al. 2011b; Eales et al. 2012; Magdis et al. 2011, 2012a; Berta et al. 2016; Tacconi et al. 2018) or (indirectly) through the monochromatic flux densities in the Rayleigh-Jeans (R-J) tail of the SED (e.g. Groves et al. 2015; Scoville et al. 2017) has gained momentum, a proper investigation of the evolution of $f_{\rm dust}$ and DMF with redshift is necessary, and remains to be done. More importantly, the $f_{\rm dust}$, the DMFs and in general the life cycle of dust are key in our understanding of metal enrichment processes and dust production mechanisms. These derived properties are also critical parameters of semi-analytical and analytical models that couple the evolution of stars, metals, and gas (Lacey et al. 2016; Popping et al. 2017; Imara et al. 2018; Cousin et al. 2019; Vijayan et al. 2019; Lagos et al. 2019; Pantoni et al. 2019), as well as of cosmological simulations that also trace the dark matter component (Hayward et al. 2013; Narayanan et al. 2015; McKinnon et al. 2017; Davé et al. 2019; Aoyama et al. 2019).

These considerations provide motivation for a coherent and homogeneous analysis of the full population of IR galaxies that are detected in the recently constructed, state-of-the-art 'Super-Deblended' FIR catalogues in two of the most extensively studied cosmological fields, the Great Observatories Origins Deep Survey North (GOODS-N; Dickinson et al. 2003) and the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007a). These catalogues are built using the 'Super-Deblending' technique (Liu et al. 2018; Jin et al. 2018) that allows prior-based source extraction from highly confused *Herschel* and *SCUBA+AzTEC* maps, yielding robust UV to radio photometry for thousands of individually detected galaxies. To model the observed SEDs I built and make publicly available a novel, time efficient and panchromatic SED fitting algorithm that I use to infer and explore the evolution and the variations of IR properties of SFGs ($M_{\rm dust}$, $T_{\rm d}$, $\langle U \rangle$, $f_{\rm dust}$, $f_{\rm gas}$, DMF) out to $z \sim 4$ and compare those to recent theoretical predictions. The catalogues with the derived FIR parameters for the full sample are also publicly released.

This chapter is organised as follows. In Section 3.3 I describe the data sets used in this work, while Section 2.4 introduces my SED fitting algorithm. In Section 2.5 I perform various simulations to determine the robustness of our sample, as well as the limiting $M_{\rm dust}$. Section 2.6 presents my physical estimates for each galaxy in the sample, and their evolution with z. In Section 2.7 I analyse the evolution of $f_{\rm dust}$ and calculate the DMF through the conversion of the stellar mass function (SMF). In Section 2.8 I constrain the evolution of $f_{\rm gas}$ and compare it to the literature results. I discuss the implications that my findings have in Section 2.9, and present my main conclusions and summary in Section 2.10.

Throughout this paper I assume a flat ΛCDM cosmology with $\Omega_{\rm m,0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$ and $\rm H_0 = 70$ km s^{-1} Mpc^{-1}, and a Chabrier (2003) initial mass function (IMF).

2.3 Panchromatic catalogues and sample selection

2.3.1 GOODS-N 'Super-Deblended' Catalogue

First I consider the 'Super-Deblended' photometric catalogue (hereafter SDC1) constructed by Liu et al. (2018) using the FIR and sub-mm images in GOODS-N. These images come from the *Herschel* Space Observatory (PACS and SPIRE instruments, see Elbaz et al. 2011; Magnelli et al. 2013) and the ground-based facilities SCUBA-2 (850 μm; Geach et al. 2017) and AzTEC+MAMBO (1.16 mm; Penner et al. 2011).

Several novelties introduced in SDC1 are particularly useful for my analysis. First, detections from deep Spitzer IRAC and VLA 20 cm (Owen 2018) are used as a prior for the positions of the blended FIR/sub-mm sources. Second, the SED information from shorter-wavelength photometry is also used as a prior for subtracting lower redshift sources. This substantially decreases blending degeneracies and allows for a robust photometry extraction of sources at longer wavelengths. Moreover, the authors estimated more realistic photometry uncertainties for each photometric measurement with extensive Monte Carlo simulations. These improvements allow for deeper detection limits and statistically reliable estimates (both measurements and uncertainties) in the FIR+mm bands.

The SDC1 contains 3,306 priors in total, including over 1,000 FIR+mm detections. All sources have photometric redshifts and stellar masses inferred by EAZY (Brammer et al. 2008), and FAST (Kriek et al. 2009) respectively, based on the 3D-HST UV-near-IR (Skelton et al. 2014) and Pannella et al. 2015 GOODS-N catalogues. Following Liu et al. (2018), I extend the photometric coverage of the published SDC1 catalogue to shorter wavelengths by cross-matching with the 3D-HST UV-near-IR (Skelton et al. 2014) and Pannella et al. 2015 GOODS-N catalogues. Approximately half of the objects within the catalogue are spectroscopically confirmed. However, as mentioned by the authors, the outer perimeter of the GOODS-N area contains objects with high instrumental noise in the 24 μm prior image that may impair the extraction process. I therefore choose to limit our analysis to the central 134 arcmin^2 with reliable photometry (flag `goodArea` = 1 in SDC1). This reduced my final sample to 2,344 objects.

2.3.2 COSMOS 'Super-Deblended' Catalogue

I supplement my study with the 'Super-Deblended' catalogue (SDC2 hereafter) presented in Jin et al. (2018). The catalogue covers 1.7 deg^2 in COSMOS, in the same bands as in SDC1, plus additional MAMBO data at 1.2 mm (Bertoldi et al. 2007).

In practice, the deblending methodology remains identical to that adopted in SDC1, with one primary difference: an additional step selecting a highly complete sample of priors in K_s band from the UltraVista catalogues (McCracken et al. 2012). The resulting 24 μm detections are then combined with the mass-limited sample of K_s sources in order to fit the remaining bands in the catalogue.

The final input dataset contains 195,107 priors, with 13% of them having spectroscopic confirmation. The authors highlight that only 11,220 objects are in fact detected over the 100-1200 μm range. Similarly to GOODS-N, I impose the `goodArea = 1` flag to only include sources that are present in the UltraVista Data Release 4 area (McCracken et al. 2012). I note that for their catalogue, Jin et al. (2018) used a combination of Laigle et al. (2016) and Muzzin et al. (2013) (M13 hereafter) catalogues. The M13 catalogue has an advantage in that it does not completely remove the sources around saturated stars, which has a positive effect on the number

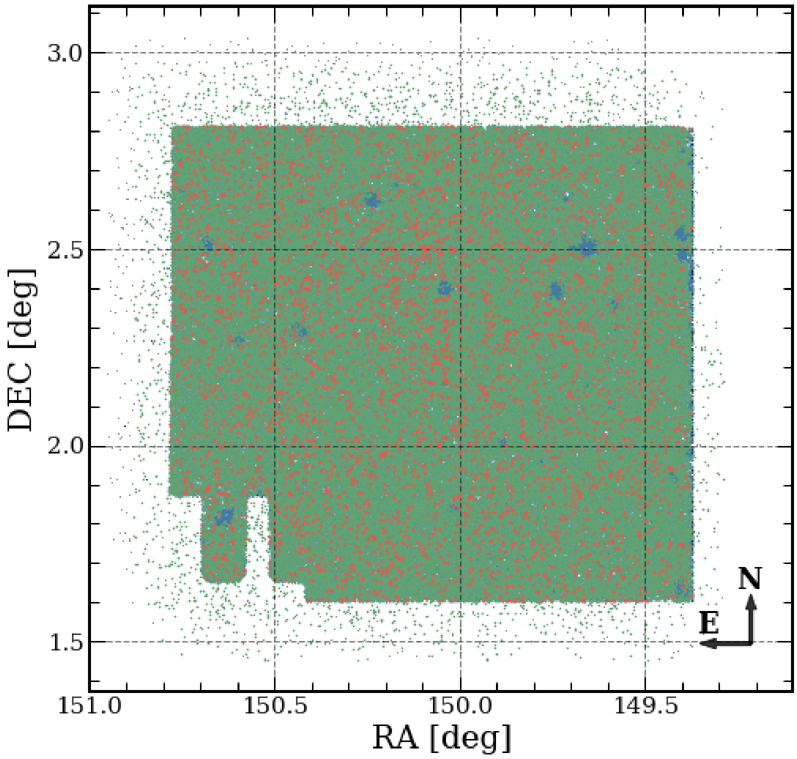

Figure 2.1: COSMOS sky map. The blue regions and black points represent sources from Muzzin et al. (2013) with `goodArea` = 1 and 0 respectively (according to the quality flag in Jin et al. 2018 catalogue). The Laigle et al. (2016) coverage is shown in green. Sources that I use in my analysis are indicated in red.

counts, however the reduced quality of the photometry could lead to unreliable estimates for photometric redshifts ($z_{\rm phot}$), as well as any parameters extracted by fitting optical templates. To be on the safe side, my analysis of SDC2 will only focus on the Laigle et al. 2016 sources, which narrows down the number of objects in the input catalogue to 186,549 and the total area to 1.38 deg^2. The COSMOS 2015 catalogue (Laigle et al. 2016), also comes with a plethora of UV-optical photometry, spanning an additional $\sim$ 20 bands, as well as photometric redshifts and stellar mass estimates by `LePhare` (Arnouts et al. 1999; Ilbert et al. 2006). I exploit these data by cross-matching the same COSMOS 2015 UV-optical photometry that was used to derive M_* and redshift to SDC2, thus extending the photometric coverage. In total, the merged catalogues consist in $\sim$ 40 bands.

For posterity, in Figure 2.1 I present the UltraVista Data Release 2 map (Laigle et al. 2016;

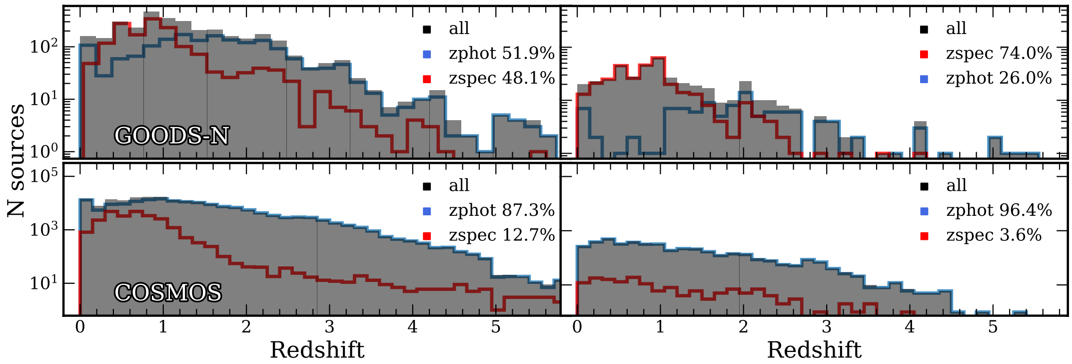

Figure 2.2: Redshift distributions of the sources considered in the present work. Both the photometric and spectroscopic redshifts were taken from the corresponding SDCs. The left and right panels show the redshift distribution of the original, full catalogue and of the final sample that meets the selection criteria described in Section 2.3.3, respectively.

Davidzon et al. 2017) (L16 area), with the regions where the star subtraction took place being clearly identified. In addition to that, on the West border of the survey there exists a number of sources falling outside of the L16 area. These only have UltraVista coverage and lack additional *Subaru* photometry, which could affect the reliability of the $z_{\rm phot}$.

2.3.3 Sample Selection

The primary parameters that we can derive from observing the rest-frame FIR emission are the total infrared luminosity ($L_{\rm IR}$), the dust mass ($M_{\rm dust}$), and $T_{\rm d}$ or equivalently the intensity of interstellar radiation field ($\langle U \rangle$) in the Draine & Li (2007) dust model. To obtain robust estimates for these parameters, an adequate multi-wavelength sampling of a galaxy's SED is required. As such, constraining the IR-peak is necessary for a robust $L_{\rm IR}$ estimate, while detections in the long wavelength regime (Rayleigh-Jeans) are imperative to capture the emission from cold dust.

With these considerations in mind, after the initial cleaning of the catalogues described above, I perform the sample selection based on the following requirements:

- Detection at a S/N>3 significance in at least three FIR to sub-mm bands from 100 μm to 1.2 mm [1].

- Available $z_{\rm phot}$ (or $z_{\rm spec}$) and M_* estimates inferred by UV to near-IR photometry

After the quality cuts and the selection criteria, I was left with 4,331 objects in COSMOS and 585 sources in GOODS-N, which constitute my sample. I have additionally identified 75 objects

[1]Three bands are also required to reduce fitting degeneracies

within SDC2, that fulfil my FIR detection criteria, but despite having a well sampled UV-optical photometry, do not have either $z_{\rm phot}$ or M_{*} estimates. I subsequently fit these sources with EAZY (Brammer et al. 2008), extract their $z_{\rm phot}$, as well as other UV-optical properties, and add them to my final catalogue. This brings the total amount of COSMOS sources to 4,406. The $z_{\rm phot}$ and $z_{\rm spec}$ distributions of the final sample in the two fields are presented in Figure 2.2 (right). As I will discuss later, a third criterion requesting at least one detection at $\lambda_{\rm rest} > 150\,\mu$m will be imposed in order to define a sub-sample with robust $M_{\rm dust}$ estimates.

2.4 SED fitting

2.4.1 The Inventory of Available SED fitting routines

Prior to providing the description of my methodology, I believe it is important to outline and present a brief introduction of the available SED fitting codes that deal with a three component fitting approach, namely combining the optical, AGN and dust emissions. These include the well established and widely used energy-balance routines such as `CIGALE` (Burgarella et al. 2005; Noll et al. 2009; Boquien et al. 2019), `MAGPHYS` (da Cunha et al. 2008; Battisti et al. 2019) as well as its AGN template extension presented in Chang et al. 2017. These have inspired more novel and sophisticated approaches that optimise the template libraries to achieve significant improvements in computational speeds - `SED3FIT` (Berta et al. 2013), or adopt MCMC methods when extracting best fit parameters such as AGN`fitter` (Calistro Rivera et al. 2016) and `Prospector`-α (Leja et al. 2018). These efforts are not just limited to published software packages, with many authors implementing their own routines for a panchromatic model analysis (e.g. see Feltre et al. 2013; Symeonidis et al. 2013).

2.4.2 Basic Description of the `Stardust` Fitting Code

To model the extensive photometric coverage of the galaxies in my sample I develop a new, panchromatic SED fitting tool: `Stardust`. The code performs a multi-component fit that linearly combines stellar libraries with AGN torus templates and IR models of dust emission arising from star-formation (SF-IR). This approach, which is very similar to that presented in (Liu et al. 2021), has a number of key differences compared to the currently existing SED fitting codes such as `MAGPHYS`, `CIGALE` and `SED3FIT`.

First, the three components (stellar, AGN and SF-IR) are fit simultaneously yet independently from each other, without assuming an energy balance between the absorbed UV/optical radiation and the IR emission. The energy balance approach relies on the assumption that fitted stellar and dust emissions are co-spatial, i.e. the process of UV absorption and subsequent re-emission at IR wavelength happen in the same environment (da Cunha et al. 2008, e.g. see their Section

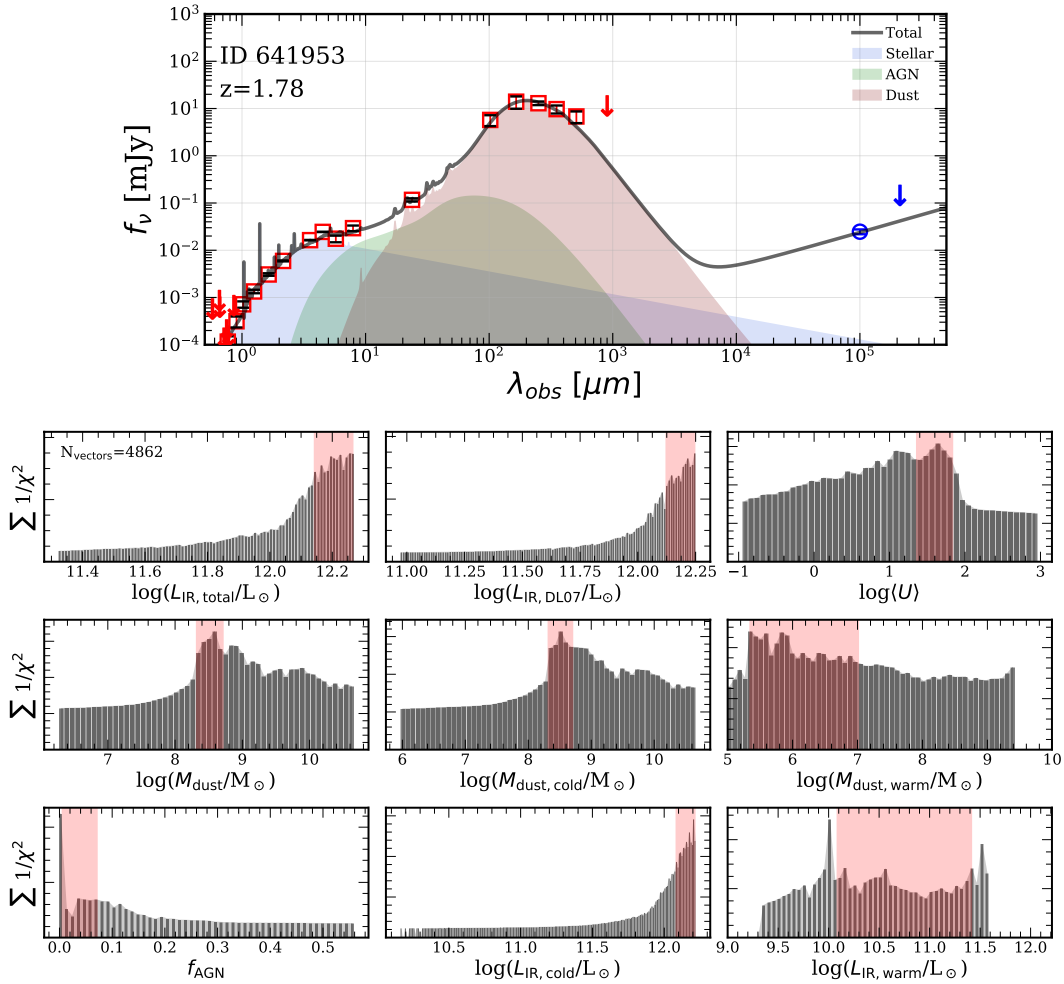

Figure 2.3: Top: Example of an observed and best fit SED, as obtained with `Stardust` code for a $z_{\rm phot}$ = 1.78 galaxy (ID641953) drawn from the SDC2 sample. The squares and circles represent the S/N > 3 photometric detections while 3σ upper limits are shown as downward arrows. Red symbols represent the SDC2 photometry that was used in the fitting routine, while blue symbols show the radio measurements at 1.4 GHz and 3 GHz that were not included in the fit. Instead, the radio part of the model SED is based on the $L_{\rm IR}$ - $L_{1.4\,\rm GHz}$ relation of Delvecchio et al. (2020). The shaded red, green and blue regions, correspond to the dust, AGN and stellar components respectively. A linear combination of all three is given by a solid black line. **Bottom:** The χ^2 distributions for the main derived parameters. The shaded red areas enclose solutions for which $\Delta\chi^2_\nu = 1$.

2.1 and 2.2). However, the detected stellar and dust distributions within a galaxy are not always physically connected. Resolved observations of high-z dusty-SFGs (DSFGs) (Simpson et al. 2015; Gómez-Guijarro et al. 2018; Franco et al. 2018; Hodge et al. 2016, 2019; Kaasinen et al. 2020) have revealed spatial offsets between the extent of the dust and stellar emitting sizes of high$-z$ DSFGs, with the former being on average more compact (e.g. Chen et al. 2017; Tadaki et al. 2017;

Calistro Rivera et al. 2018; Cochrane et al. 2021). While energy balance is anticipated to apply universally, the aperture and sensitivity limitations of the observations cast a concern on how reliably I can bring these components together in the same panchromatic fit. These observations are also theoretically supported by radiative transfer codes (e.g. SKIRT, Cochrane et al. 2019) and hydrodynamical simulations (e.g. IllustrisTNG, Pillepich et al. 2019; Popping et al. 2021).

Moreover, the detection of 'HST-dark' galaxies, i.e. sources that are undetected in the UV/optical bands, and thus do not have the photometry to constrain neither dust absorption nor stellar emission, but are bright in the IRAC and FIR/mm bands (Wang et al. 2016; Franco et al. 2018), pose another technical challenge to the correct application of the energy balance approach. Finally, the manner in which dust and stellar emission are connected, by assuming a single dust attenuation law and dust composition, can have a significant impact on derived parameters, as these recipes have been shown not to apply universally (Buat et al. 2019). In summary, while the premise of the energy balance routines is undoubtedly elegant and in most cases physically motivated, I choose to use independent stellar, AGN, and dust components to better focus on the dust properties themselves.

Furthering this complex picture, it is important to note the presence of cold diffuse dust, that is being heated by an older stellar population, rather than an ongoing star-formation activity, (see Boquien et al. 2011; Bendo et al. 2012; Galametz et al. 2014; Hayward et al. 2014). However, when dealing with non-resolved observations, the diffuse dust and RJ-tail emissions are highly degenerate, and as such none of the aforementioned codes, nor `Stardust`, utilise these templates.

The other advantage of `Stardust` is related to the χ^2 minimisation approach to select the best fit models. Instead of finding the solution from a vast library of pre-compiled templates, I devise an optimisation method to find the best linear combination of a much smaller set of 'basic' templates (similar to eigenvectors in principal component analysis). This is the same approach adopted in the photometric redshift fitting code EAZY (Brammer et al. 2008). In my case, the basic templates are divided in three classes and the linear combination includes a sum of templates from these classes. The models used to create these templates are the following:

- Stellar library. I incorporate an updated version of the Stellar Population Synthesis (SPS) models described in Brammer et al. (2008). Although UV and optical photometry is not always available, the inclusion of the stellar component in the code is important in the NIR regime. In particular, it allows to better constrain the AGN contribution (see Figure A.1 and Section A.1). This stellar library represents an optimised basis-set, where the non-negative linear combinations of models can be considered to be the "principal components" of a much larger parent template catalogue (see Brammer et al. 2008 and Blanton & Roweis 2007).

- AGN library. I adopt empirically derived templates from Mullaney et al. (2011) describing AGN intrinsic emission from 6 to 100 μm [2]. I include both high- and low-luminosity templates (total of 2). Since these can be linearly combined, I do not include the median luminosity AGN template.

- Infrared library. It consists in 4,862 Draine & Li 2007 (DL07 hereafter) templates, with the additional updates from Draine et al. 2014 (also see Aniano et al. 2020). These models describe the contribution from warm dust and polycyclic aromatic hydrocarbon (PAH) features in the photo-dissociation regions (PDR), together with cold dust in the diffuse part of the ISM. I consider[3] a wide array of values for the minimum radiation field ($U_{\rm min}$) in the $0.1 < U_{\rm min} < 50$ range, as well as the fraction of the total dust mass locked in PAHs ($q_{\rm PAH}$) between 0 and 10 %. I have fixed $U_{\rm max} = 10^6$ and $\alpha = 2$, as described in Magdis et al. 2012a. These templates are not linearly combined within their class, the algorithm instead chooses a single best-fit DL07 template.

- Radio continuum. Data points in radio are not considered by my fitting routine, however they can be used a posteriori to quantify possible radio excess and further confirm the presence of an AGN, if needed. My radio model is based on the radio-FIR correlation, described in Delvecchio et al. (2020), with a spectral index of -0.75.

More details on the characteristics of the templates and the motivation for selecting them are provided in Section A.1. With such a configuration, fitting a single object (including the computation of the uncertainties) with `Stardust` takes less than 10 seconds[4], i.e. a factor of 8-10 faster than software like `CIGALE` (see Section A.3), based on large pre-compiled template sets. If I choose to pre-compile all of the templates, instead of linearly combining them, the resulting model library would contain millions of possible SEDs, with computation time increasing by a significant amount. The code is also highly parallelised, which allows it to run on multiple threads simultaneously, thus achieving significant computation speed improvements on modern CPUs.

2.4.3 Configuration of the Code

For each object, the input consists of measured flux densities, their corresponding uncertainties and a redshift estimate.[5] The user must then choose the corresponding filter curves from the pre-compiled set, or upload their own. The individual template components can be switched

[2]Note: These templates do not account for X-ray selected QSOs. The flexible nature of the code however allows these templates to be manually injected if necessary.

[3]The modular structure of the code allows the user to decide which DL07 templates to use.

[4]Tested on a i9-8950HK CPU.

[5]It is also possible to manually define a rectangular filter at a desired wavelength, for cases where the filter transmission curve is not easily obtainable, e.g. ALMA.

off and on as an additional user input. The algorithm then outputs the best-fit FIR as well as AGN properties of the source. If the photometry is available, the UV-optical parameters are also produced. These can be summarised as follows:

- The total infrared luminosity integrated over the SF-IR+AGN templates ($L_{\mathrm{IR,total}}$)[6], the total infrared luminosity associated with star-formation ($L_{\mathrm{IR,DL07}}$) and the relative contribution of the PDR component (f_{PDR}) to $L_{\mathrm{IR,DL07}}$.
- The bolometric IR luminosity of the AGN (L_{AGN}) and its fractional contribution (f_{AGN}) to the total IR energy budget[7].
- The total dust mass (M_{dust}), the warm dust mass component heated by PDRs ($M_{\mathrm{dust}}^{\mathrm{warm}}$), the fraction of the total dust mass heated by PDRs (γ), the cold dust mass component ($M_{\mathrm{dust}}^{\mathrm{cold}}$) in the diffuse ISM, and the fraction of the total dust mass locked in PAHs q_{PAH}.
- The intensity of the radiation field at which the diffuse ISM is exposed to (U_{min}), which is a proxy of the mass-weighted T_{d} of the ISM, and the mean radiation field intensity ($\langle U \rangle$), which is a proxy of the luminosity-weighted T_{d}.
- The stellar mass (M_*), star-formation history (SFH), $E(B-V)$ and the unobscured SFR, if there is available optical photometry.

Figure 2.3 presents an example fit to one of the COSMOS galaxies, chosen for its prominent AGN contribution. The top panel of the Figure shows the data points and the best-fit model, with different colours for the four components listed above; the bottom panel displays the χ^2 distributions of all relevant IR quantities.

2.4.4 Derivation of Uncertainties

In order to estimate the errors associated with the derived quantities during the fitting procedure of `Stardust`, I consider two main sources of uncertainty; one concerning the linear combination coefficients of the best-fit optimisation and one linked to the broad-band photometric data.

To quantify the linear combination uncertainty, I re-sample the best solution coefficients. A covariance matrix is first created by considering all of the templates that went into the best-fit solution. I draw the coefficients from a multivariate normal distribution whose median are given by the coefficient of the best-solution vector and the standard deviation is computed from the diagonalised covariance matrix. This is done 10^4 times to provide a good balance between robustness of the error estimates and computational speed. I recompute all the relevant FIR

[6] In this work I use terms L_{IR} and $L_{\mathrm{IR,total}}$ interchangeably.

[7] The quality of the photometry in this work does not allow us to distinguish bolometric AGN contributions below $\sim 0.5\%$, and thus the non-zero entries below that threshold are treated here as zero.

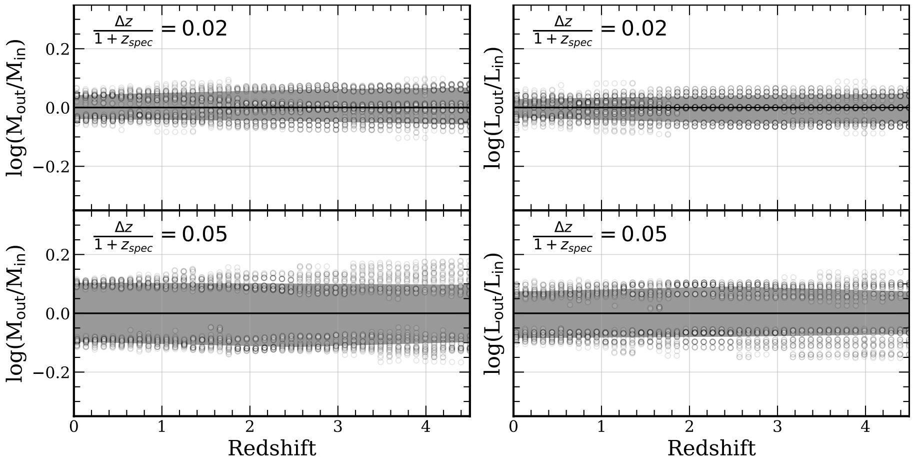

Figure 2.4: Effect of the $z_{\rm phot}$ uncertainty in the derivation of $M_{\rm dust}$ (left panel) and $L_{\rm IR}$ (right panel) assuming $\epsilon = 0.02$ (top) and $\epsilon = 0.05$ (bottom). Red circles represent the ratio of the output to input quantities from my simulations, as inferred by shifting the fitted redshift by $\pm\Delta z$. The shaded regions cover the 68% confidence interval and the solid black line indicates a ratio equal to unity.

properties for each realisation of the routine. From resultant distributions I then define the lower and upper uncertainty as the 16th and 84th percentile confidence interval respectively. However, given the fact that only a single solution with a single coefficient is considered in the IR, the final uncertainty on the FIR properties is underestimated.

I therefore also consider the observational uncertainty, that is primarily driven by the quality of the photometric data. I compute it by considering all possible solutions from my template library which fall within the 68% confidence interval range of the best-fit. This would correspond to a region where the solutions fall within $\Delta\chi^2_\nu = 1$, where $\Delta\chi^2_\nu = \chi^2_{\nu,i} - \chi^2_{\nu,\rm best}$, since the non-diagonal terms of the template covariance matrix are not zero [8]. I show these as red shaded areas on the bottom panel of Figure 2.3. The observational uncertainty is then derived as simply the width of the shaded region.

The final errors are computed as a quadrature sum of the linear combination uncertainty and the observational uncertainty.

2.4.5 The Effect of Photometric Redshift Uncertainty

To explore and quantify how the uncertainty in $z_{\rm phot}$ propagates into the estimates of $L_{\rm IR}$ and $M_{\rm dust}$, I built mock IR SEDs of 1,200 galaxies utilising a suite of 0.1< $U_{\rm min} < 50$ DL07 models and place them uniformly in the $0.03 < z_{\rm true} < 5$ redshift range. Thus, each mock SED is

[8] See sections 15.1 and 15.6 of Press et al. (1986) and Avni (1976). Note that $\Delta\chi^2_\nu = 1$ only applies when marginalised over all other parameters.

characterised by a set of pre-defined $L_{\rm IR,in}$, $M_{\rm dust,in}$, and $z_{\rm true}$ values. I then infer synthetic broad-band photometry in all IR bands available in *SDC2* (24-1100 μm) for each simulated galaxy. Since I am interested in the effect of the $z_{\rm phot}$ uncertainty on the derived FIR properties, to minimize any other possible sources of error (e.g. photometric uncertainty, poor photometric sampling of the SED) I adopt S/N = 5 in all bands and add to the photometric data set the monochromatic flux density of the template at 2.2 mm. I then fit the synthetic photometry of each galaxy by fixing the redshift first to $z_{-} = z_{\rm true} - \Delta z$, and then to $z_{+} = z_{\rm true} + \Delta z$ where Δz=ϵ (1+$z_{\rm true}$). For the purposes of this work, and based on the $z_{\rm phot}$ accuracy of the COSMOS field ($\epsilon = 0.005 - 0.03$ as quantified in L16) I first adopt $\epsilon = 0.02$ and then repeat the analysis for an even more conservative case with $\epsilon = 0.05$. The comparison between the extracted $L_{\rm IR,out}$ and $M_{\rm dust,out}$ to the input values for each simulation is presented in Figure 2.4.

My analysis suggests that the effects Δz has in the derivation of the FIR properties is not negligible, even for the idealised case of detailed (24 μm – 2.2 mm) and high quality (S/N = 5) photometric coverage. Indeed, I find that a typical value of $\epsilon = 0.02$ ($\epsilon = 0.05$), introduces an extra scatter of $\sim$12% (25%) and $\sim$17% (35%) in the derived $L_{\rm IR}$ and $M_{\rm dust}$, that remains rather constant with redshift (at least out to $z = 4$). At the same time, I also find that a symmetric Δz, as the one adopted in my simulations, does not inflict a noticeable systematic offset in the extracted FIR quantities (< 0.05 dex).

Based on these results I update the uncertainties of the inferred FIR properties of my $z_{\rm phot}$-sample (`ztype`=0, see Table A.2) by adding, in quadrature, the extra error arising from a symmetric Δz (assuming $\epsilon = 0.02$ for all sources) to the error budget inferred by the SED fitting procedure (photometry+model). While my correction is based on an average value of $\epsilon = 0.02$, I note that for catastrophic $z_{\rm phot}$ failures ($\epsilon \sim 0.15$) I find a systematic offset of $\lesssim$30% between the input vs output $M_{\rm dust}$, while the $L_{\rm IR}$ ratios remain uniformly scattered around unity.

2.5 Completeness and Systematics

By construction, the photometric catalogues considered in this work combine observations that span $\sim$4 orders of magnitude in wavelength range, have different sensitivity limits and are differently affected by source blending and confusion. The fact that I choose to draw my sample based on the criteria described in section Section 3.3, rather than selecting galaxies detected in a single band (i.e. a flux limited sample), necessitates a series of simulations in order to quantify possible biases, systematic effects, as well as the completeness of my sample in terms of $M_{\rm dust}$ and $L_{\rm IR}$.

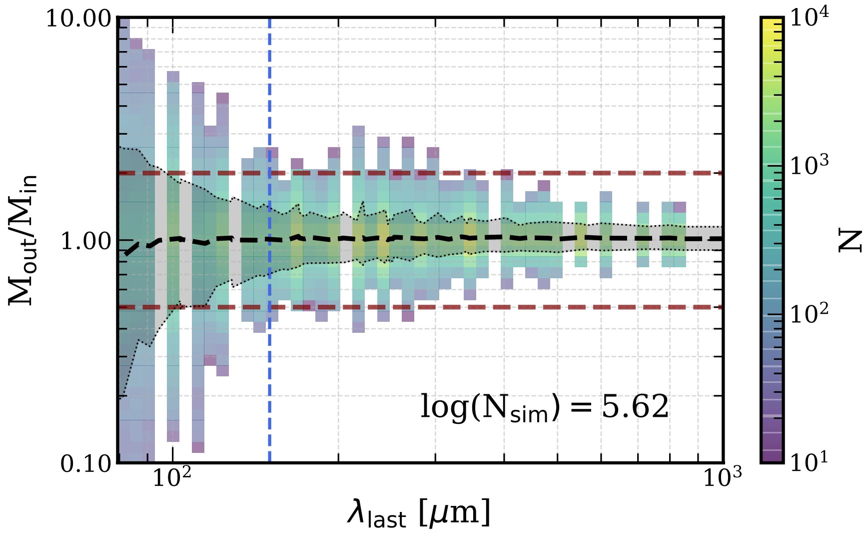

Figure 2.5: Accuracy of the inferred $M_{\rm dust}$ estimates as a function of (rest-frame) $\lambda_{\rm last}$, parametrised by the ratio of the output to input $M_{\rm dust}$ in my simulations. The colour coding illustrates the density of the data points. The dashed black line and the shaded grey area denote the median and the 16$^{\rm th}$ and 84$^{\rm th}$ percentile confidence intervals respectively. The dashed maroon lines represent the value where the $M_{\rm dust}$ ratio is 0.5 and 2. The vertical blue line at 150 μm denotes the $\lambda_{\rm last}$ onwards where for ≥68% of the simulated galaxies the accuracy of the recovered $M_{\rm dust}$ is ≥70%. The quantisation along the x axis is a consequence of the step-size in redshift, alongside with the available observed bands.

2.5.1 The effect of $\lambda_{\rm last}$ Cutoff

It has been well established that for a robust modelling of the $M_{\rm dust}$, at least one photometric data point long-wards of the FIR peak, i.e. into the R-J tail of the SED, is necessary (e.g. Draine et al. 2007; Magdis et al. 2012a; Berta et al. 2016; Scoville et al. 2017). Here, I attempt to quantify how the accuracy of the derived $M_{\rm dust}$ estimates varies as a function of the rest-frame wavelength of the last available detection in conjunction with the selection criteria of my sample (i.e. at least 3 detections at 24 μm $< \lambda \leq 1200\,\mu$m). For this, I perform the following simulations.

I start by building mock IR SEDs of fixed $M_{\rm dust}$ and $L_{\rm IR}$, with $0.1 < U_{\rm min} \leq 50$, $0 \leq \gamma \leq 0.5$ and $q_{\rm PAH}$ = 2.8%, using the DL07 library and place them at $0.01 \leq z \leq 4.5$ with a step of Δz = 0.05. After all of the models are created, synthetic photometry is performed by convolving each mock redshifted SED with a filter transmission curve. I consider all filters redder than MIPS 24 μm available in SDC2, for a total of 9 bands and set all recovered fluxes to a S/N=3 significance level. At each redshift the algorithm calculates the rest-frame wavelength for each available band, producing a grid of possible rest-frame wavelengths of the last detection ($\lambda_{\rm last}$) after accounting for my selection criterion that requests the availability (and the detection) in at least two bluer bands. For each $\lambda_{\rm last}$ it then randomly selects two additional bands bluewards of $\lambda_{\rm last}$, producing a final set of three photometric data points. I then fit each set, and 50 permutations by varying the original fluxes by 10 %, with my code to extract $M_{\rm dust}$ and $L_{\rm IR}$ estimates. This procedure is then repeated for all mock SEDs and all acceptable values of $\lambda_{\rm last}$ in each redshift.

The accuracy with which I can recover $M_{\rm dust}$ estimates for each $\lambda_{\rm last}$ is then quantified by the scatter of the output to input $M_{\rm dust}$ ratio presented in Figure 2.5. As expected, I find a decreasing scatter in $M_{\rm dust,out}/M_{\rm dust,in}$ at longer $\lambda_{\rm last}$, that drops from a factor of $\sim$2 (for 68% of the simulated galaxies) at $\lambda_{\rm last} = 100\,\mu$m to a factor of $\sim$1.1 at $\lambda_{\rm last} = 400\,\mu$m.

Based on these results, I choose to define the sub-sample of '$M_{\rm dust}$-*robust*' galaxies, for which at least one detection at $\lambda_{\rm last} \geq 150\ \mu$m is available. This threshold was chosen as an optimal compromise between the number of the rejected sources and the precision of the derived $M_{\rm dust}$ that for $\lambda_{\rm last} \geq 150\,\mu$m is $\geq$70%. Indeed, while $\lambda_{\rm last} \geq 150\,\mu$m is evidently not deep into the R-J, it seems that the addition of the two extra data points blueward of $\lambda_{\rm rest} = 150\,\mu$m are adequate to anchor the general shape, and eventually the $M_{\rm dust}$, of the templates.

As a sanity check for the effectiveness of my criterion, I cross match the '$M_{\rm dust}$-*robust*' sample with the A3COSMOS ALMA photometric catalogue presented in Liu et al. 2019a and refit the 233 galaxies that both have in common, adding this time the extra ALMA data point to the input photometry. The comparison of the inferred $M_{\rm dust,A3}$ to my $M_{\rm dust}$ estimates yields a very good agreement between the two estimates with a median log($M_{\rm dust,A3}/M_{\rm dust}$) $\approx -0.04 \pm 0.06$, further supporting my analysis. Nevertheless, I do identify a handful of sources for which the addition of the ALMA data point results in lower $M_{\rm dust}$ estimates by a factor of $\geq$3. An inspection of the SEDs of these extreme outliers reveals that the discrepancy originates either from possible catastrophic blending of the SPIRE 500 μm photometry or alternatively from over-resolved ALMA photometry [9].

The emerging '$M_{\rm dust}$-*robust*' sample consists of 3,312 sources drawn from the same M_* and redshift distributions as the originally 4,991 selected galaxies. Finally, I note that my simulations operate under the assumption that the DL07 models are a good representation for the FIR emission of the real galaxies. Variations in dust composition, dust emissivity and dust absorption coefficients that could result in systematic offsets in the inferred $M_{\rm dust}$ (e.g. Magdis et al. 2012a; Dale et al. 2012; Berta et al. 2016; Scoville et al. 2017) cannot be modelled with my approach. As is the case for any other $M_{\rm dust}$ analysis - the relative rather than the face value estimates bare more physical significance.

[9]The SEDs of the most extreme outliers can be retrieved here https://github.com/VasilyKokorev/sdc_ir_properties/.

2.5.2 Limiting M_{dust} and L_{IR}

I now attempt to compute the completeness threshold of my sample in terms of M_{dust} and L_{IR} as a function of redshift. Again, I build a grid of mock IR SED in the $z = 0 - 5$ range using the same approach as described above. However, instead of considering the full range of possible $\langle U \rangle \propto L_{\mathrm{IR}}/M_{\mathrm{dust}}$ values of the DL07 models, the constructed templates this time follow the $\langle U \rangle$ – z relation of MS galaxies presented in Béthermin et al. (2015). At each redshift I then create a grid of templates normalised to log(M_{dust}/M$_{\odot}$) = $6 - 10$ in steps of 0.1.

As before, the templates are used to produce synthetic photometry for each template in all bands available in the SDC2 catalogue. For each band I then adopt an rms based on the depth of the corresponding survey at each wavelength in the COSMOS field (Jin et al. 2018), and impose the same selection criteria to the mock photometry as those applied to the real catalogues. Following Section 2.5.1 I also request that the simulated sources have $\lambda_{\mathrm{last}} \geq 150\,\mu$m. If a galaxy of given M_{dust} fulfils these criteria at a given redshift, the algorithm moves to a lower M_{dust} until the object is rejected by my selection. By following these steps at different redshifts I thus obtain a limiting M_{dust} as a function of z, that I coin lim(M_{dust})(z). In the process I also consider the scatter of the $\langle U \rangle$ – z relation of Béthermin et al. 2015, in order to account for the variation of $\langle U \rangle$ among MS galaxies at a given redshift. The derived lim(M_{dust})(z) can then be converted to lim(L_{IR})(z), via:

$$\langle U \rangle = \frac{L_{\mathrm{IR,DL07}}}{125\, M_{\mathrm{dust}}}, \tag{2.1}$$

as described in Draine & Li (2007).

I also repeat my simulation for SBs, by fixing $\langle U \rangle = 40$ (e.g. Magdis et al. 2012a; Tan et al. 2014; Béthermin et al. 2015). I note however that the $\langle U \rangle$ of SBs can vary substantially to lower or higher values (e.g. Magdis et al. 2012a, 2017; Schreiber et al. 2018; Jin et al. 2019; Cortzen et al. 2020). Therefore, the chosen $\langle U \rangle$ = 40 is only representative, but not necessarily unique. Nevertheless, $\langle U \rangle < 40$ templates are represented in the simulation of the MS galaxies while galaxies with $\langle U \rangle > 40$ are rather rare.

The results of my simulations are presented in Figure 2.6, where I show the derived M_{dust} and L_{IR} as a function of z for the whole sample, along with the evolution of lim(M_{dust}) and lim(L_{IR}). I see that the limiting M_{dust} increases towards high-z, peaking at $z \sim 2$ and remains flat afterwards, signifying that the balance between cosmological dimming and negative K - correction is achieved beyond that point. Following the black line one could infer the M_{dust} threshold above which my sample should be 100 % complete, assuming an MS-like population of galaxies. However, since my sample is not limited to MS galaxies I naturally also find sources that fall below my limiting M_{dust} track. As I will discuss later, these are predominately starbursting galaxies that exhibit elevated $\langle U \rangle$ with respect to the MS. The secondary lim(M_{dust}) trend for an SB-like population

Table 2.1. Properties of the galaxy sample selected in Section 3.3. With the exception of redshift, the other quantities are derived via SED fitting. Values are presented as the median and a double sided 68% confidence interval.

	COSMOS	GOODS-N	All
# of Galaxies	4,406	585	4,991
z	$0.88^{+1.09}_{-0.57}$	$1.01^{+1.02}_{-0.53}$	$0.90^{+1.08}_{-0.58}$
$L_{\mathrm{IR,total}}$ ^{a} [10^{12} $\mathrm{L}_\odot$]	$0.45^{+2.17}_{-0.40}$	$0.35^{+1.60}_{-0.28}$	$0.44^{+2.09}_{-0.38}$
$L_{\mathrm{IR,DL07}}$ ^{b} [10^{12} $\mathrm{L}_\odot$]	$0.44^{+2.10}_{-0.39}$	$0.34^{+1.53}_{-0.27}$	$0.42^{+2.04}_{-0.37}$
SFR^c [$\mathrm{M}_\odot$ yr^{-1}]	$43.96^{+209.93}_{-38.88}$	$33.84^{+152.74}_{-27.02}$	$42.30^{+203.88}_{-37.06}$
M_{dust} [10^8 $\mathrm{M}_\odot$]	$2.73^{+16.66}_{-2.08}$	$2.60^{+10.34}_{-1.76}$	$2.71^{+15.71}_{-2.04}$
M_* ^{d} [10^{10} $\mathrm{M}_\odot$]	$4.17^{+5.83}_{-2.89}$	$3.56^{+6.03}_{-2.32}$	$4.07^{+5.93}_{-2.72}$
f_{AGN} [%]	$0.86^{+2.44}_{-0.60}$	$2.60^{+4.74}_{-2.14}$	$0.93^{+3.06}_{-0.67}$
$\langle U \rangle$	$10.12^{+29.55}_{-7.72}$	$9.38^{+29.82}_{-6.95}$	$10.00^{+29.68}_{-7.58}$

a Computed over a linear combination of AGN+DL07 best fit templates.
b Only considering the best fit DL07 template.
c Computed from $L_{\mathrm{IR,DL07}}$.
d Taken from the parent catalogue.

displays that with the SDC2 detection limits it is possible to detect a low M_{dust} object only if it is also a starburst.

I also find a similar trend for MS galaxies when considering the evolution of lim(L_{IR}). In this case, however I do not observe a flattening at $z \sim 2$ and the trend continues to rise into the early Universe. The balance between cosmological dimming and the negative K - correction is not being achieved here, since the wavelengths that are required to reliably constrain the L_{IR} are positioned to the left of the FIR peak.

Admittedly, depths of FIR surveys are not the only limiting factors of sample selection. A requirement to have a photometric redshift and M_*, would mean that the optical photometry has to be sufficiently sampled, to allow such an analysis. Moreover, the deblending procedure itself goes through various selection stages, including both brightness and mass cuts. Various IR studies (Wang et al. 2016; Franco et al. 2018) have revealed substantial populations of optically

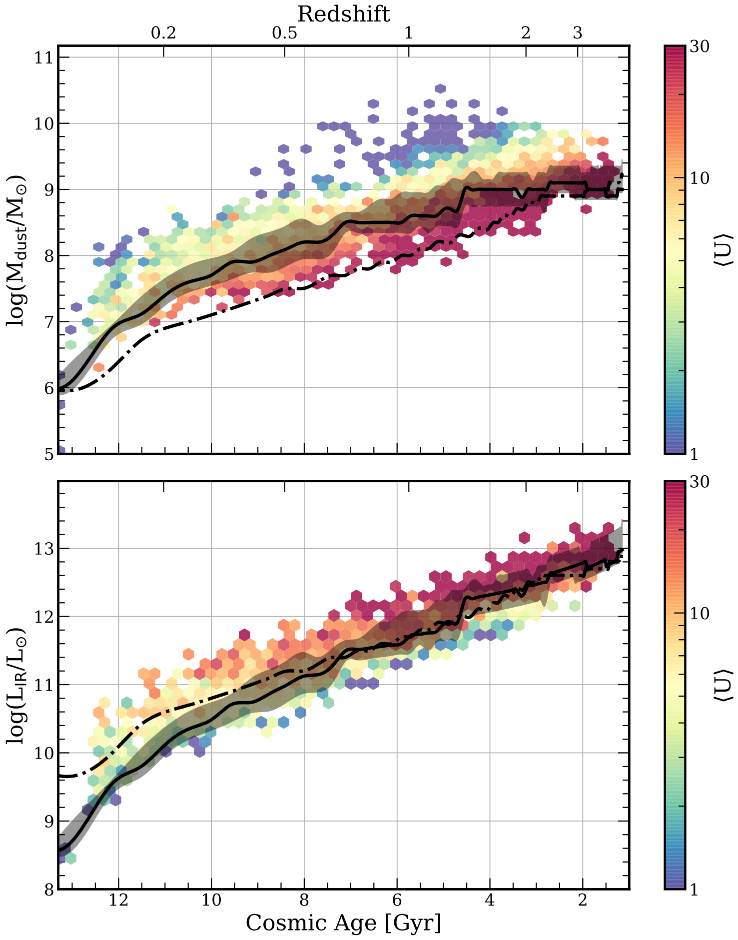

Figure 2.6: Simulated evolution of the lim($M_{\rm dust}$)(z) (top) and lim($L_{\rm IR}$)(z) (bottom), as described in Section 2.5.2. The black line represents the derived trend for MS galaxies and the dashed-dotted line shows the same relations for SBs ($\langle U \rangle \sim 40$). The shaded regions define the 16$^{\rm th}$ and 84$^{\rm th}$ percentile confidence intervals, based on the scatter of the $\langle U \rangle - z$ relation from Béthermin et al. (2015). The hexagonal bins contain the inferred parameters of the '$M_{\rm dust}$-*robust*' sample (i.e. at least one detection at $\lambda \geq 150\,\mu$m), colour coded by mean radiation field intensity $\langle U \rangle$.

dark sources at $2 < z < 4$, that are otherwise bright in IRAC and FIR bands, which would be unintentionally excluded from my analysis. Moreover, even at low-z, objects that are faint in the K-band would also be missed. Indeed, a combination of these factors creates significant obstacles in my completeness analysis, I address this in more detail in Section 2.7.2.

2.6 Far-infrared Properties of GOODS-N and COSMOS galaxies

Using my newly developed code presented in Section 2.4, I extract the FIR and UV-optical properties for all 4,991 galaxies from the SDC1 and SDC2 that meet my selection criteria as listed

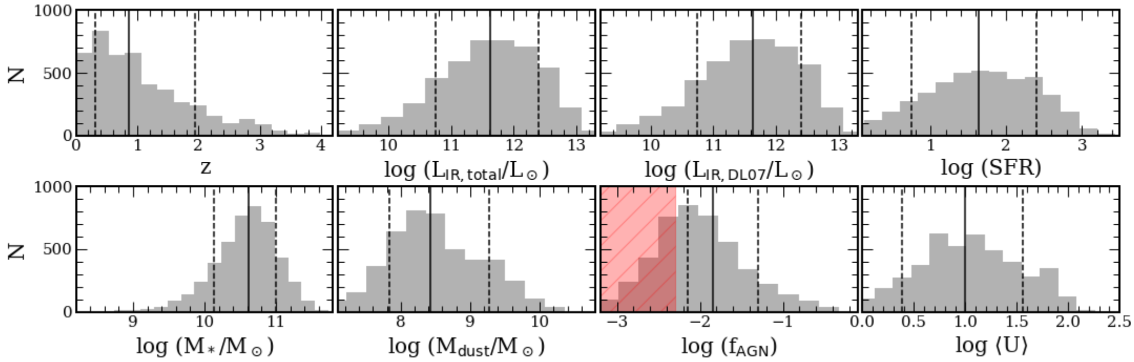

Figure 2.7: Distribution of the inferred IR properties of the COSMOS and GOODS-N samples. With the exception of z and M_*, these properties are the output of my SED fitting code (see Section 2.4). The solid black and dashed lines represent the median and the 68% confidence interval respectively. The hatched red region on the $f_{\rm AGN}$ histogram highlights the range where estimates are not reliable (i.e., $f_{\rm AGN} < 0.005$).

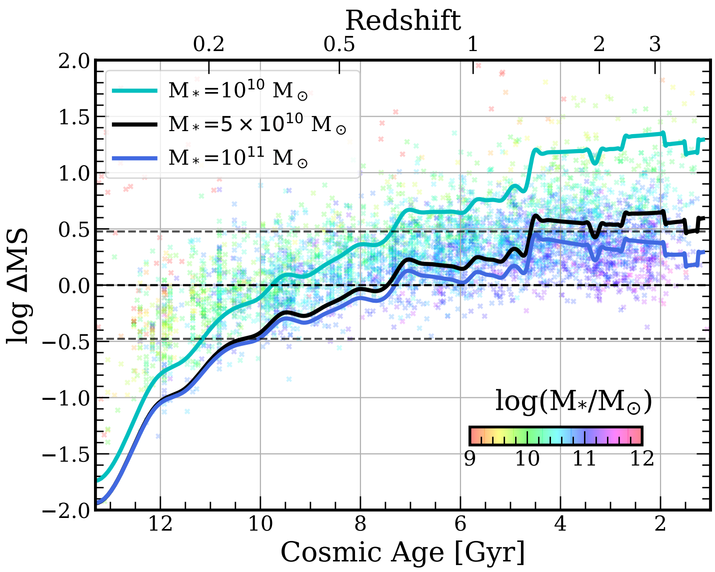

Figure 2.8: Position of my sources with respect to the MS as a function of cosmic age and redshift. Points are colour coded according to the M_*. The dashed black and gray lines denote the MS and its 0.5 dex scatter. The solid coloured lines correspond to the ΔMS detection limit as computed based on the inferred $lim(L_{\rm IR})$ (see Section 2.5.2), and assuming $M_* = 10^{10}$, $5{\times}10^{10}$ and 10^{11} $M_\odot$.

in Section 3.3. Moreover, since both input catalogues contain M_* estimates provided by either `LePhare`, EAZY or FAST, I am able to carry out a comparison of these M_* to the ones derived by `Stardust`. I find that the stellar masses are consistent with one another and direct the reader to Section A.2 for a more detailed comparison between the two methods, as well as to EAZY

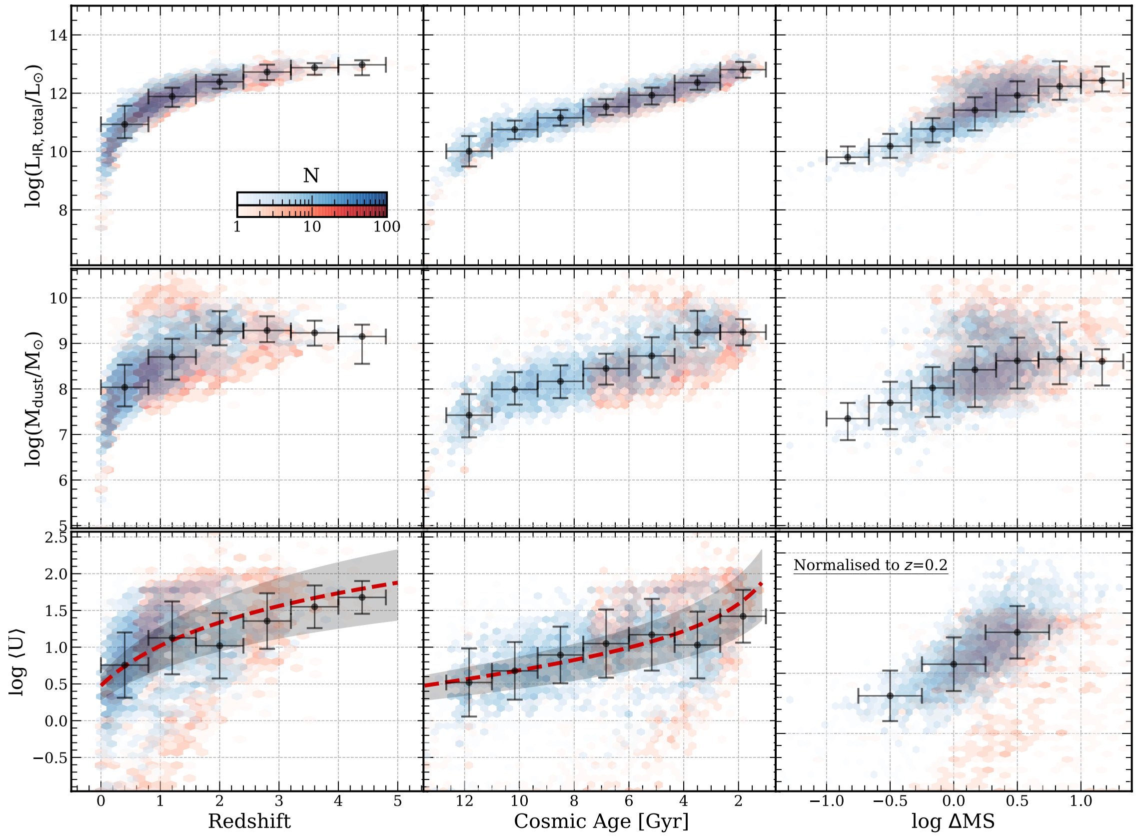

Figure 2.9: Evolution of general FIR properties, as computed with `Stardust`, as a function of z, cosmic age and ΔMS. The hexagonal bins are normalised by the number count, and contain the 'M_{dust}-*robust*' sample in blue, and the objects that were removed after the quality cut in red. For the 'M_{dust}-*robust*' I show the binned median points, with their y-uncertainty corresponding to the 16$^{\mathrm{th}}$ and 84$^{\mathrm{th}}$ percentile intervals and x-uncertainty to the bin width. The dashed red lines and shaded regions correspond to the $\langle U \rangle - z$ relation for MS galaxies from Béthermin et al. (2015).

derived M_*. Despite the similarities between the available and derived M_*, in my subsequent analysis I will utilise the M_* from the parent catalogue, unless it is specified otherwise. This is done to preserve the original mass cuts described in Liu et al. (2018) and Jin et al. (2018), and therefore the mass completeness and homogeneity of the catalogue.

In total, out of 4,991 sources, there are 21 that I consider to be catastrophic fits ($\chi^2_\nu > 100$), these only comprise 0.4 % of the entire output catalogue and are subsequently removed. The average χ^2 per degree of freedom of the entire dataset was computed to be equal to 0.98. The distribution of the FIR properties of the whole sample, their medians and associated uncertainties are presented in Figure 2.7 and summarised in Table 2.1. The catalogue containing the extracted FIR properties is described in the Appendix and is publicly available along with the best fit SED for each object.

I also calculate the position of the galaxies in my sample with respect to the MS, by converting the AGN-free $L_{\mathrm{IR,DL07}}$ estimates to SFR (Kennicutt 1998) and using the functional form of the main-sequence as presented in Schreiber et al. (2015), accounting for the fact that they use a Salpeter (1955) IMF. The distribution of ΔMS= SFR/SFR$_{\mathrm{MS}}$ as a function of redshift and stellar mass is presented in Figure 2.8. I define the boundary between the star-forming and quiescent galaxies at logΔMS=-0.5 dex and between MS and SBs at logΔMS=0.5 dex, which in linear space corresponds to $\times 3$ below/above the MS. Quite naturally, for decreasing M$_*$ and increasing redshift, my sample is progressively restricted to galaxies that lie above the MS. This is shown by the tracks in Figure 2.8 that indicate the limiting ΔMS for fixed M$_*$ as a function of redshift that is reached by my data, after converting the inferred $lim(L_{\mathrm{IR}})$ to lim(SFR). Nevertheless, I find that the majority of my sources are classified as MS galaxies (69%), with the remaining objects either considered to be undergoing a phase of 'bursty' star-formation (26%) or being passive galaxies (5%).

As a final sanity check I additionally fit the same sources with CIGALE, by utilising DL07 models and similar sets of optical and AGN templates. I find that the output parameters as derived from the two codes are in good agreement and defer the reader to Section A.3 for a more detailed comparison.

2.6.1 The 'M_{dust}-*robust*' Sample

Now I focus on the FIR properties of the 'M_{dust}-*robust*' galaxies described in Section 2.5.1, that should represent the most reliable sample for the exploration of the dependency of the L_{IR}, M_{dust}, and $\langle U \rangle$ on redshift, cosmic age, and ΔMS. The emerging results are presented in Figure 2.9, where for completeness and to facilitate comparisons, I also include the inferred properties of the full sample.

Both L_{IR} and M_{dust} are found to increase smoothly as a function of ΔMS. At the same time, I also find that for MS galaxies $\langle U \rangle$ evolves as $(3.2 \pm 1.3) \times (1+z)^{1.2\pm0.3}$, which is in excellent agreement with the stacking analysis of Béthermin et al. (2015). The fact that the individually detected galaxies appear to follow the same $\langle U \rangle$ - z relation as the much deeper stacked ensembles, reinforces the notion that the adopted 'M_{dust}-*robust*' sub-sample does not introduce a significant bias towards colder objects.

Since $\langle U \rangle$ is proportional to $L_{\mathrm{IR,DL07}}/M_{\mathrm{dust}}$, and also a proxy for T_{d}, my analysis provides further evidence that dust in MS galaxies becomes warmer towards higher redshifts, a trend that has already been recovered in previous studies (most of them based on stacking analysis, see e.g. Magnelli et al. 2014; Schreiber et al. 2015; Davidzon et al. 2018). Similarly, these data also confirm a progressive increase of $\langle U \rangle$ (or T_{dust}) with an increasing elevation above the MS (e.g. Magdis et al. 2017; Jin et al. 2019).

It is worth noticing that the full sample follows the same general trends albeit with a considerably larger scatter ($\sim \times 2$) in $M_{\rm dust}$ and $\langle U \rangle$. The reduced scatter for the '$M_{\rm dust}$-*robust*' sub-sample is driven by the imposed $\lambda_{\rm last} \geq 150\,\mu$m selection criteria that primarily removes the locus of sources with very cold fitting solutions ($\langle U \rangle \lesssim 1$). I highlight that the rejection of these objects should not introduce a bias in my sample since such low $\langle U \rangle$ values are more indicative of poor photometric coverage/quality (lack of available data point in the R-J) rather than of realistic, extremely cold ISM conditions. However, I note that not all the extremely cold solutions have been removed from the '$M_{\rm dust}$-*robust*' sample by my selection, as ~200 objects with $\langle U \rangle < 1$ meet the $\lambda_{\rm last} > 150\,\mu$m criterion. These can be easily identified in the $\langle U \rangle - z$ plot and as the outliers populating the secondary blue cloud of points in the $M_{\rm dust} - \Delta$MS plot in Figure 2.9. As I will discuss later, these could be sources with unreliable $z_{\rm phot}$ estimates, failures of the deblending in the SPIRE bands, or, more interestingly, gas giants or very compact galaxies with optically thick FIR emission.

2.6.2 Cold vs Warm Dust

The SED decomposition introduced in Section 2.4 can also provide constrains on the relative contribution of the warm (PDR, $L_{\rm IR}^{\rm warm}$, $M_{\rm dust}^{\rm warm}$) and cold (diffuse, $L_{\rm IR}^{\rm cold}$, $M_{\rm dust}^{\rm cold}$) ISM components to the total $L_{\rm IR}$ output and the total $M_{\rm dust}$ budget of the galaxies in my sample. In particular, it is worth investigating if and how the relative contribution of the components varies as a function of ΔMS. Indeed, if SBs are experiencing elevated star-formation activity per surface area (Elbaz et al. 2011, 2018; Valentino et al. 2020) then one would expect to see an increased fraction of $L_{\rm IR}$ (and $M_{\rm dust}$) originating (and being heated) from the 'PDR' component, where the radiation intensity ranges from $U_{\rm min}$ to $U_{\rm max}$ (Draine & Li 2007).

In Figure 2.10 I plot the inferred properties of the warm and cold ISM components as a function of ΔMS. I find that for a fixed $L_{\rm IR}$ (or equally SFR), SBs tend to have lower amounts of $M_{\rm dust}^{\rm cold}$, with $M_{\rm dust}^{\rm cold}/L_{\rm IR}$ showing a tight anti-correlation with ΔMS. The same, however, does not apply to $M_{\rm dust}^{\rm warm}/L_{\rm IR}$, which exhibits a significantly larger scatter and only a very weak dependence on ΔMS. This is a consequence of an increasing fraction of warm to cold $M_{\rm dust}$ and $L_{\rm IR}$ between MS and SBs galaxies (Figure 2.10).

The observed trends suggest that compared to MS galaxies, SBs have a larger fraction of their total $M_{\rm dust}$ exposed to the intense stellar radiation fields of the PDRs, in agreement with expectations discussed above. My result could indeed reflect an increase in the compactness of the star-formation activity for increasing distance from the MS as suggested in recent high resolution studies. Finally, under the assumption that $M_{\rm dust}^{\rm cold}$ is proportional to $M_{\rm gas}$, and $L_{\rm IR}$ is proportional to SFR, my results point to enhanced star-formation efficiencies and shorter gas depletion time scales for sources residing above the MS, as already reported in the literature (e.g. Tacconi et al.

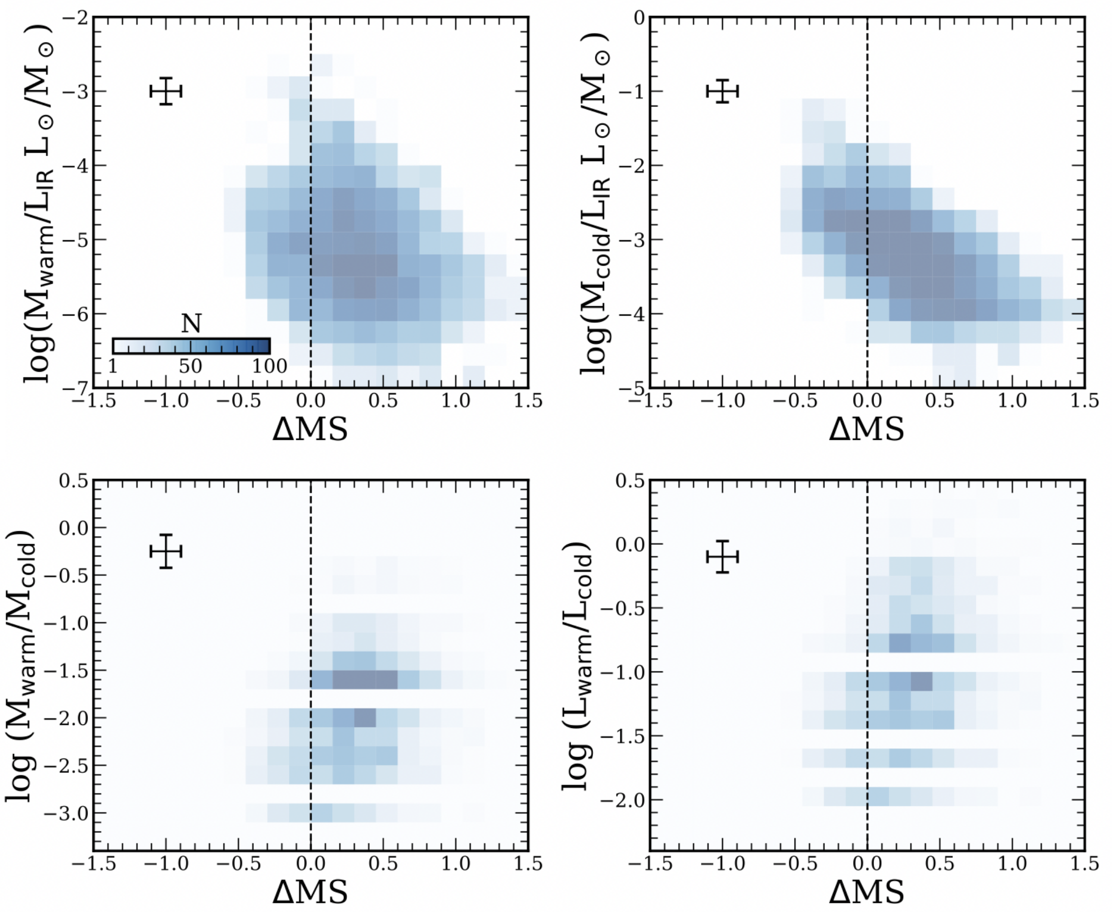

Figure 2.10: Comparison of the properties of the warm (PDR) and cold (diffuse) dust components of the ISM as a function of ΔMS. From left to right the panels show the warm and cold dust mass components weighted by the total $L_{\rm IR}$, the fraction of warm to cold $M_{\rm dust}$ and the fraction of 'PDR' to diffuse ISM IR output. The bins are coloured based on number density of the data points. I show a typical uncertainty on the plotted parameters in the top left corner of each panel.

2010, 2020; Daddi et al. 2010; Magdis et al. 2012a, 2017; Sargent et al. 2014; Silverman et al. 2018).

2.7 Dust to Stellar Mass Relation and Dust Mass Functions

As discussed in Section 2.2, constrains on the evolution of $f_{\rm dust}$ and the DMF are key towards a better understanding of dust production and destruction mechanisms at different epochs. Within this context, I explore how the current data set traces the evolution of $f_{\rm dust}$ and use it to characterise the DMFs at various redshifts.

2.7.1 The Evolution of the Dust Mass Fraction

To infer the evolution of $f_{\rm dust}$ I adopted the formula described in Liu et al. (2019b), which parametrises $f_{\rm dust}$ in terms of z, M_* and ΔMS. Compared to more simple log-space linear fitting models (e.g. Scoville et al. 2017), this formulation recovers trends that are more physically meaningful and also explores how these parameters are covariant and degenerate with each other in a multi-dimensional fitting space. As an initial check, I performed a Spearman correlation test and found $M_{\rm dust}$ to be mildly correlated with logΔMS ($\rho = 0.40$) and strongly correlated with M_* and $t_{\rm age}$ ($\rho = 0.63$ and -0.80 respectively). I consider the following functional form:

$$
\begin{aligned}
\log(f_{\rm dust}) =& (a_0 + a_1 \log(M_*/10^{10} M_\odot)) \log\Delta\mathrm{MS} \\
& + b \log(M_*/10^{10} M_\odot)) \\
& + (c_0 + c_1 \log(M_*/10^{10} M_\odot)) \, t_{\rm age}(z) \\
& + d
\end{aligned}
\tag{2.2}
$$

where $t_{\rm age}$ is the cosmic age at a given redshift in Gyr and M_* is the stellar mass in $\mathrm{M}_\odot$. For fitting I use the `Python` package - `scipy.optimize.curve_fit`, which finds the solution based on the least-squares method. I also consider how the extreme outliers can affect my results and thus only fit the medians in a given redshift bin. The best fit values are as follows:

$$
\begin{aligned}
a_0 &= -0.90 & a_1 &= 1.31 \\
b &= -0.98 & c_0 &= -0.23 \\
c_1 &= 0.11 & d &= -0.64,
\end{aligned}
$$

with the uncertainties being computed from the covariance matrix. I then used the functional form given by Equation 2.2 to re-normalise all galaxies to lie on the main sequence (ΔMS = 1), and $M_* = 5 \times 10^{10}$ $\mathrm{M}_\odot$, in order to directly compare with my best-fit function in 2 dimensions.

The normalised data and the best fit relation presented in Figure 2.11 are in very good agreement with the collection of similar trends drawn from the literature (e.g. Scoville et al. 2017; Tacconi et al. 2018; Liu et al. 2019b; Magnelli et al. 2020). I note that any apparent discrepancies between the slope and the normalisation of the recovered relation to that of Scoville et al. (2017) and Magnelli et al. (2020) are driven by the model parametrisation, as for the latter the multi-variable functional forms do not consider the covariance between the fitted values.

In Figure 2.12 I also show how the one dimensional relation between $M_{\rm dust}$ and M_* compares to the multi-dimensional fit within six redshift bins. Since mt analysis can be affected by the completeness of my sample in terms of M_* and $M_{\rm dust}$, I also consider the underlying selection

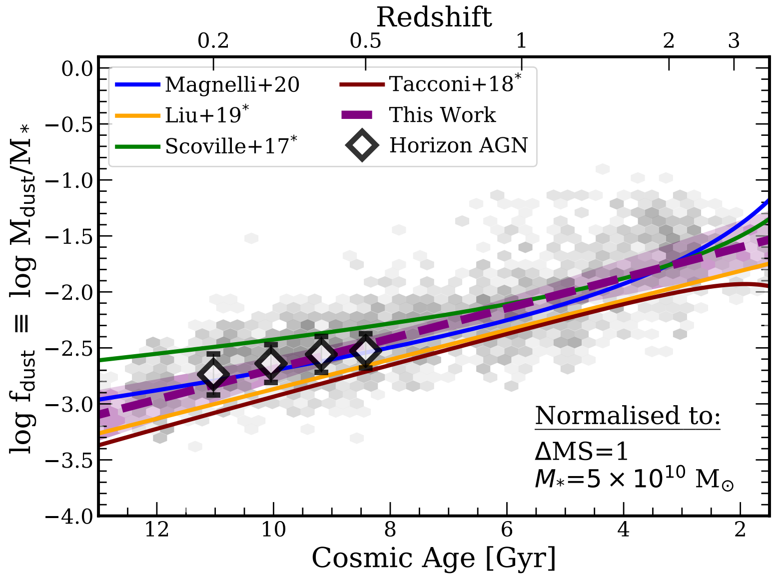

Figure 2.11: Derived relations for $f_{\rm dust}$ as a function of $z/t_{\rm age}$. The dashed purple line shows the fit to the data, while solid coloured lines display literature results. The shaded purple region denotes the 16$^{\rm th}$ and 84$^{\rm th}$ percentile confidence intervals of my fit. Starred labels denote literature calculations where a direct comparison was not available and a $\delta_{\rm GDR} = 100$ was assumed. The grey hexbins contain the data from '$M_{\rm dust}$-*robust*' sample, and are normalised by the number count. Both the data and the derived relations have been re-scaled to ΔMS=1 and $M_* = 5 \times 10^{10}$ M$_\odot$. White diamonds show median positions of the Horizon AGN star-forming galaxies at that redshift, normalised in the same way as my data.

effects based on my lim($M_{\rm dust}$)(z) derivation and assuming that my catalogue is complete at M$_*$ $> 10^{10}$ M$_\odot$. I find that for a fixed redshift range, the inferred $M_{\rm dust}-$ M$_*$ relation shadows the multi-parameter fit up to $z \sim 1.1$, in moderate agreement with the trends reported in Liu et al. (2019b) and Magnelli et al. (2020). However, the increasing incompleteness and the low number statistics do not allow me to extend the analysis to higher redshifts.

Based on these results I will limit the subsequent DMF analysis to the $0.2 < z < 1.1$ range.

2.7.2 Dust Mass Functions

After obtaining the functional form for $f_{\rm dust}(t_{\rm age}, M_*, \Delta{\rm MS})$, I am now in a position to examine the shape and the evolution of the DMF with redshift. For this I will restrict my analysis to the COSMOS sample (SDC2), as there is enough statistics or coverage of the GOODS-N field to reliably constrain the DMF.

The 'Super-Deblending' procedure that went into producing the SDC2 catalogue creates a significant obstacle when attempting to consider all of the incompleteness effects of the sample. The objects that end up in my sample go through several selection stages, both before and after the deblending procedure. These include both the brightness and mass cuts of the parent catalogue (Laigle et al. 2016), the availability of infrared coverage (Jin et al. 2018) and the selection criteria imposed in my study. As such, it is not possible to robustly assess the properties and the number of objects that end up being 'lost'. Therefore, I select an alternative approach in computing the DMF for my bjects, namely utilising the derived $f_{\rm dust}$ parametrisation along with the available SMFs in the literature. Later in this section, I also attempt to account for the incompleteness effects and compare the SMF-derived DMF to the observed number density of galaxies per $M_{\rm dust}$ bin.

DMF from SMF

For my analysis I adopt the SMF computed by Davidzon et al. (2017), which covers the entire COSMOS field and the $z = 0.2 - 4$ range. Their mass function already accounts for the Eddington bias, so I do not need to consider any additional corrections. Since the vast majority of galaxies in my sample are star-forming, I adopt the derived parameters for the 'active' SMF only.

The galaxy mass function is normally expressed as a Schechter function (Schechter 1976), which in logarithmic form can be written as:

$$\begin{aligned}\Phi(\log M)d(\log M) = \ln(10)\exp(-M/M^*)\\ \times\Phi^*(-M/M^*)^{(\alpha+1)}d(\log M_*)\end{aligned} \tag{2.3}$$

where α is the slope of the faint-end, Φ^* is the normalisation and M^* is the characteristic mass, indicating the position of the 'knee' of the Schechter function. To convert the SMF to DMF, I first postulate that the number of galaxies in a given redshift bin is the same, regardless of whether one integrates over M_* or $M_{\rm dust}$, namely:

$$\int_0^\infty \Phi(\log M_*)d(\log M_*) = \int_0^\infty \Phi(\log M_{\rm dust})d(\log M_{\rm dust}),$$

the integrands can be re-arranged to obtain:

$$\Phi(\log M_{\rm dust}) = \Phi(\log M_*)d(\log M_*)/d(\log M_{dust}).$$

I then differentiate Equation 2.2, to obtain $d(\log M_*)/d(\log M_{\rm dust}) = (b + c_1 \times t_{\rm age} + 1)^{-1}$. In order to perform the final conversion I also transform all the M_* bins into $M_{\rm dust}$ bins, by inverting my formulation of $f_{\rm dust}$, taken at $\Delta{\rm MS} = 1$.

Accounting for the Eddington Bias

Before comparing to the real data it is important to note that, while calculating the SMF-derived DMF, the Eddington bias, induced by the $f_{\rm dust}$ scatter should be taken into account. Since I am directly employing the Davidzon et al. SMF, where the Eddington bias has already been corrected, using the best-fit $f_{\rm dust}$ relation to convert SMF to DMF will only reproduce the median trend and will not properly account for the full dynamic range of observed $M_{\rm dust}$. To alleviate this, I rely on the work by Loveday et al. (1992), that showed that the Eddington bias manifests itself as a Gaussian, whose width is equal to the scatter of the variable of interest, convolved to the mass function. I have thus utilised an approach similar to that used in Davidzon et al. (2017) for the SMF and Beeston et al. (2018) for the DMF, where they successfully deconvolve their observed mass functions by using the scatter of the observed variable. As such, within each redshift bin I consider the standard deviation of the $f_{\rm dust}$ in a logarithmic space. I then use this scatter to create a simple Gaussian that is centred at zero, and then convolve it with my SMF-derived DMF. This allows me to better take into account the scatter of my data, and thus produce a more realistic mass function. In conclusion, I have indirectly produced two versions of SMF-derived DMFs, with and without the scatter.

Comparison to the Observed Number Density

Now I would like to compare the SMF-derived DMF to the observed number density of galaxies in the '$M_{\rm dust}$-*robust*' sample. To this end, I first apply the widely used non-parametric 1/$V_{\rm max}$ method to correct for the Malmquist bias of my sample (Schmidt 1968). This method relies on assigning the $V_{\rm max}$ to each redshift bin, based on the detection limits of the survey. Effectively, this correction accounts for the fact that in a given volume, a brightness limited survey is more likely to pick up the bright sources, while missing faint galaxies, that would populate the low-$M_{\rm dust}$ end. I explicitly highlight that the $V_{\rm max}$ correction only accounts for the FIR flux rms cuts, and not the selection criteria outlined in Section 2.3.3.

To calculate the $V_{\rm max}$ I use the prescription from Weigel et al. (2016), which provides a volume correction for each individual source. As a first step, and for a given redshift bin, I split my sources into 0.4 dex bins in the log($M_{\rm dust}$/M$_\odot$) = 6 − 11 interval. Given that the median uncertainty on the $M_{\rm dust}$ is ~ 0.3 dex, the following bin spacing will ensure that there is very little to no cross-contamination between mass bins. The $V_{\rm max}, i$, where i denotes an individual galaxy, can be then calculated as:

$$V_{\rm max} = \frac{A}{3} \left(d_c \left(z_{\rm max,i}\right)^3 - d_c \left(z_{\rm min,i}\right)^3\right), \tag{2.4}$$

where d_c is the comoving distance and A is the area, which in our case is equal to 1.38 deg^2. Following Weigel et al. (2016), the $z_{\rm min,i}$ is given simply by the lower boundary of the bin. The

$z_{\rm max,i}$ on the other hand can be calculated empirically, either through detection limits of individual bands or by considering a limiting mass of the survey. It however cannot exceed the maximum redshift of the bin.

To obtain the $z_{\rm max}$, I consider the best fit SEDs for my sources, and the rms of the parent catalogues in order to redshfit the sources to the point where they no longer fulfil my selection criteria as outlined in Section 3.3. Using this method I however found that, for an overwhelming majority of sources, the computed $z_{\rm max,i}$ exceeds the upper boundary of the bin they belong to. Therefore, the $V_{\rm max}$ correction that I apply becomes effectively bound between the lower and the upper redshift of the bin. I find that this method works best in the lowest $(0.2 < z < 0.5)$ redshift bin, with the $\langle V/V_{\rm max}\rangle$ test returning a value of 0.47. The remaining two redshift bins are significantly incomplete, with the ratio returning 0.83 and 1.38 respectively.

Among the other sources of incompleteness, as discussed in Section 2.7.2, here I can attempt to account for lost sources due to the sensitivity limits of the survey and failures in the deblending procedure. I therefore multiply my points by the loss fraction in each redshift bin that is computed as the ratio between sources in my catalogue over the sources that have `SN-IR`> 1 [10] in the parent catalogue. The `SN-IR` parameter, described in greater detail in Liu et al. (2018) and Jin et al. (2018), and references therein, considers a combination of FIR bands starting with 100 μm. I thus expect that in this context, my `SN-IR` threshold can indicate whether a galaxy is intrinsically dusty.

The comparison of the SMF-derived DMF with and without the Eddington bias taken into account, along with the observed volume-weighted number density of galaxies, derived as described above, for three redshift bins, is presented in Figure 2.13. I find that the DMF without the Eddington bias is insufficient to reproduce the observed dynamic range of $M_{\rm dust}$, particularly in the higher redshift bins, exactly as I have predicted in Equation 2.7.2. On the other hand, the SMF-derived DMF with the artificially induced bias, through the $f_{\rm dust}$ scatter, is in good agreement with the data in the high-mass end, further highlighting the necessity of accounting for the observational biases when inferring relations (i.e. $M_* - M_{\rm dust}$ in our case) through the observed mass (or luminosity) functions. Although my model underpredicts the high-mass end data in the higher redshift bin, both still agree within the error bars. At the same time though, in the low-mass regime my data significantly underestimate the number density of galaxies mirroring the incompleteness of my sample in this $M_{\rm dust}$ regime. It is worth noticing though that the turnover of the observed data perfectly coincides with the independent estimates of lim($M_{\rm dust}$), offering an indirect validation of my simulations presented in Section 2.5.2 . For the analysis in the next sections, I adopt the SMF-derived DMF, which has the Eddington bias corrected out, as the final result against which I will compare previous observationally driven

[10] $\texttt{SN-IR}^2 = \sum_\lambda (\mathrm{S/N})^2_\lambda$, with $\lambda \geq 100\ \mu$m.

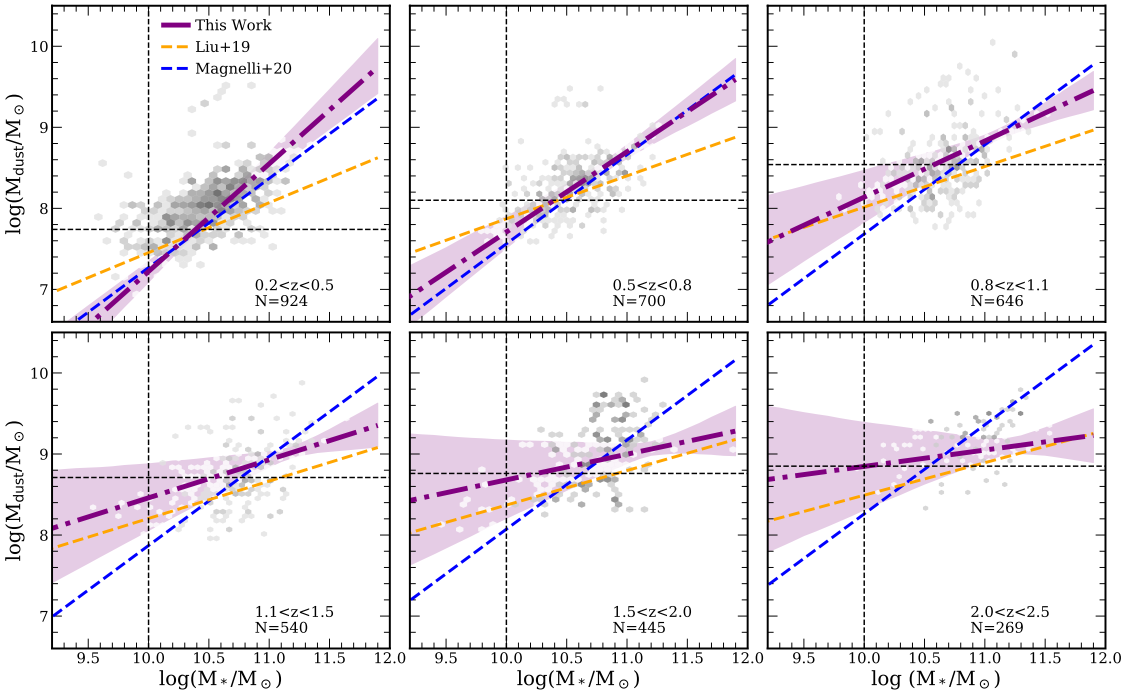

Figure 2.12: $M_{\rm dust}$ as a function of M_* in six redshift bins. The dashed-dotted purple line represents my best fit from Equation 2.2 that was collapsed to a single dimension, with ΔMS=1 and $z = \langle z_{\rm bin} \rangle$. The shaded purple regions denote the 16$^{\rm th}$ and 84$^{\rm th}$ percentile confidence intervals of my fit. The orange and the blue lines show the relations derived in Liu et al. (2019b) and Magnelli et al. (2020), respectively. The dashed black lines represent the detection limits of the original catalogue, in M_* (vertical), and the lim($M_{\rm dust}$) that I compute in Section 2.5.2 (horizontal).

DMFs and theoretical predictions [11].

2.8 Gas Content of Star-Forming Galaxies

The inferred $M_{\rm dust}$ estimates can be used as an invaluable proxy of the gas mass ($M_{\rm gas}$). To this end, I adopt the metallicity dependent gas-to-dust mass ratio $\delta_{\rm GDR}$ technique, that takes advantage of the relatively tight anti-correlation between the gas-phase metallicity, and the $\delta_{\rm GDR}$ of galaxies, both in the local Universe and at high$-z$ (see e.g. Leroy et al. 2011b; Magdis et al. 2012a; Rémy-Ruyer et al. 2014; Genzel et al. 2015). For a source with known metallicity (Z) and $M_{\rm dust}$, one can estimate the amount of $M_{\rm gas}$ via the following relation:

$$M_{\rm gas} = \delta_{\rm GDR}(Z)\, M_{\rm dust}, \tag{2.5}$$

[11] Tables containing the DMFs can be accessed here:
https://github.com/VasilyKokorev/sdc_ir_properties.

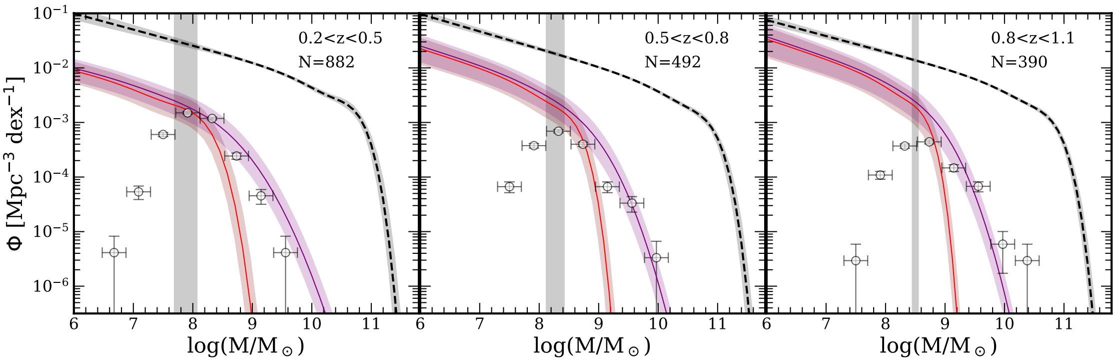

Figure 2.13: Derived DMFs in the $0.2 < z < 0.5$, $0.5 < z < 0.8$, and $0.8 < z < 1.1$ ranges. The dashed black line represents the original SMF for active galaxies from Davidzon et al. (2017). The purple and red lines are the DMFs obtained by converting the SMF, with and without the Eddington bias applied, respectively. The blue points are the DMF calculated directly from my data. The shaded rectangular area highlights the theoretical prediction for lim($M_{\rm dust}$) in that redshift interval.

where $\rm M_{gas}$ corresponds to $M_{\rm H_2} + M_{\rm HI}$, i.e. the sum of the atomic and molecular hydrogen.

Given the absence of direct metallicity measurements for my sample, I adopt the fundamental metallicity relation (FMR) of Mannucci et al. (2010). In particular, I use the $\rm M_*$ and SFR estimates as inputs to the FMR, and obtain metallicities calibrated for the Kewley & Dopita (2002) (KD02) photoionisation models. These metallicities are subsequently converted to the Pettini & Pagel (2004) (PP04 N2) scale following Kewley & Ellison (2008). I then estimate the $\delta_{\rm GDR}$ of each galaxy through the $\delta_{\rm GDR} - Z$ relation of Magdis et al. (2012a), given as:

$$\log(\delta_{\rm GDR}) = (10.54 \pm 1.00) - [12 + \log({\rm O/H})]^{(0.99 \pm 0.12)} \tag{2.6}$$

and subsequently derive $\rm M_{gas}$ through Equation 2.5, for all the sources in SDC1 and SDC2 catalogues. I propagate the uncertainties on $\rm M_{gas}$ by taking into account the uncertainty on $M_{\rm dust}$ and combining it with the typical scatter of 0.2 dex on the $\delta_{\rm GDR} - Z$ relation (Magdis et al. 2012b). These inferred $M_{\rm gas}$ estimates with associated uncertainties are included in the released catalogue.

2.8.1 Gas to Stellar Mass Relation

Similarly to $f_{\rm dust}$, I also explore the dependence of $f_{\rm gas}$ on cosmic age, ΔMS and M_*. I utilise the same multi-parameter fitting function as before (see Equation 2.2), and focus on the '$M_{\rm dust}$-*robust*' sample. I calculate the Spearman rank correlation between the variables and find $M_{\rm gas}$ to be mildly correlated with logΔMS and M_* ($\rho = 0.47$ and 0.53 respectively) and strongly negatively correlated with $t_{\rm age}$ ($\rho = -0.82$). The fitting procedure yields the following best-fit parameters:

$$a_0 = -0.73 \qquad a_1 = 1.16$$
$$b = -1.02 \qquad c_0 = -0.20$$
$$c_1 = 0.09 \qquad d = 1.39.$$

The best-fit $f_{\rm dust}$-$t_{\rm age}$ (or redshift) relation along with my data, both normalised to ΔMS = 1 and M_* = 5$\times$10^{10} M$_\odot$, are presented in Figure 2.14 (top). Similar to previous studies (Scoville et al. 2017; Tacconi et al. 2018; Liu et al. 2019b; Magnelli et al. 2020), I find a sharp increase in the gas fraction up to $z = 2$, followed by a milder evolution at higher redshifts, a change that is noticeable only in the $f_{\rm dust}-z$ parameter space. I note however that due to poor statistics and lack of spectroscopic redshifts, my data cannot reliably constrain the high-z evolution of $f_{\rm gas}$ at $z > 2$. I also detect a population of sources that display significantly elevated gas reservoirs for their redshift (log($f_{\rm gas}$)> 0.5). Some of those objects have either only a $z_{\rm phot}$ estimate available, or appear to be blended in the SPIRE bands and therefore could have an erroneously large $M_{\rm dust}$, and subsequently $M_{\rm gas}$ estimate assigned to them. However, among these I do identify some sources with spectroscopic redshifts, that are also 'clean'/isolated in the IR maps. In particular I find ~ 40 such sources with log($f_{\rm gas}$)> 0.5) and six with log($f_{\rm gas}$)> 1 that, as I discuss later, I coin as 'gas-giants'.

2.8.2 Evolution of Depletion Time

Finally, I focus my attention on the depletion time $\tau_{\rm depl} = M_{\rm gas}/{\rm SFR} = 1/{\rm SFE}$. I employ the same fitting technique as before, and explore the evolution of $\tau_{\rm depl}$ with cosmic age, ΔMS and M_* in a multi-dimensional parameter space, for the '$M_{\rm dust}$-*robust*' sample. The $\tau_{\rm depl}$ correlates mildly with age and ΔMS (Spearman $\rho = -0.46$ and -0.58, respectively) and weakly with M_* ($\rho = -0.23$). The best fit parameters are as follows:

$$a_0 = -1.65 \qquad a_1 = 1.42$$
$$b = -0.95 \qquad c_0 = -0.02$$
$$c_1 = 0.12 \qquad d = 0.37.$$

I show the best-fit $\tau_{\rm depl}$-$t_{\rm age}$ relation along with my data, both normalised to ΔMS = 1 and M_* = 5$\times$10^{10} M$_\odot$, are presented in Figure 2.14 (bottom). In line with the previous studies of $\tau_{\rm depl}$ by Scoville et al. 2017; Tacconi et al. 2018; Liu et al. 2019b, I recover a relatively weak decrease of depletion time (or increase in SFE) with redshift.

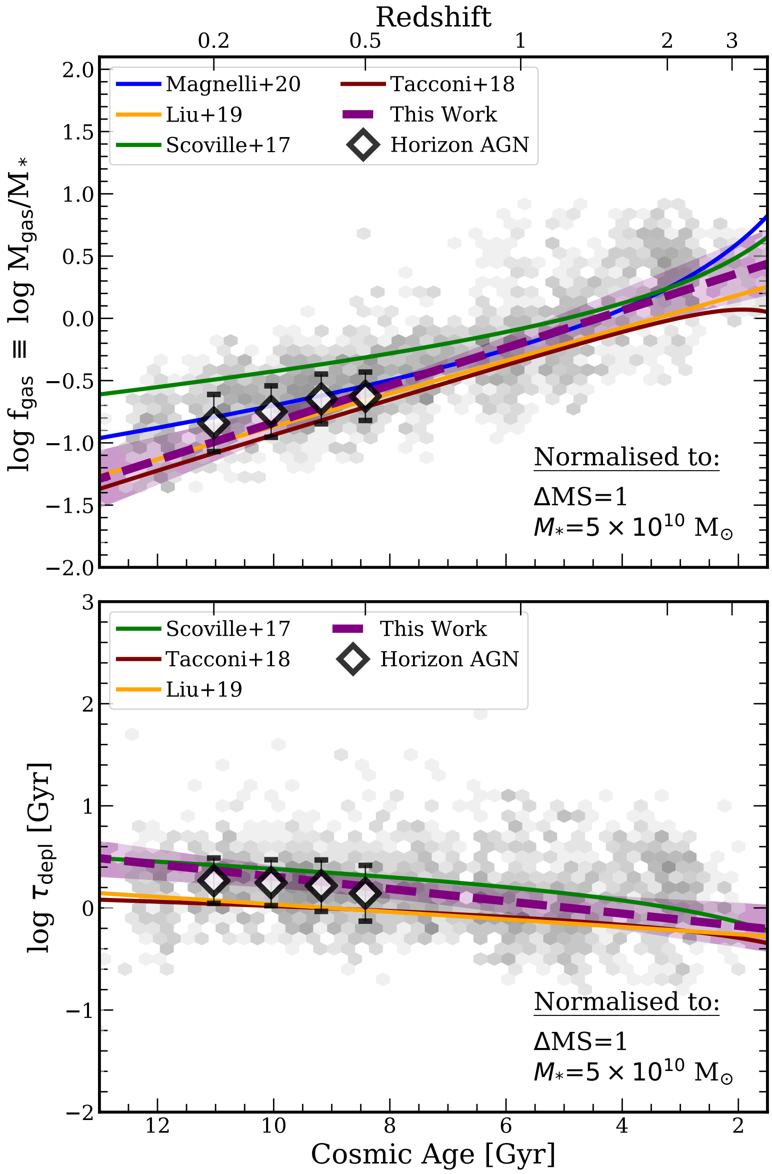

Figure 2.14: Derived relations for $f_{\rm gas}$ (top) and $\tau_{\rm depl}$ (bottom) as a function of $z/t_{\rm age}$. The dashed purple line shows the fit to my data, while solid coloured lines display literature results. The shaded purple region denotes the 16$^{\rm th}$ and 84$^{\rm th}$ percentile confidence intervals of my fit. The grey hexbins contain the data from '$M_{\rm dust}$-*robust*' sample, and are normalised by the number count. Both the data and the derived relations have been re-scaled to ΔMS=1 and $M_* = 5 \times 10^{10}$ $M_\odot$. White diamonds show median positions of the Horizon AGN star-forming galaxies at that redshift, normalised in the same way as my data.

2.9 Discussion

2.9.1 On the Dust and Gas Scaling Relations

The recovered trends between $M_{\rm dust}$, M_* and $M_{\rm gas}$ and their evolution with redshift offer a test-bed against which theoretical and previous observationally driven studies can be compared to. As shown in Figure 2.11 and Figure 2.14, my analysis yields $f_{\rm dust}$ and $f_{\rm gas}$ evolutionary tracks consistent with those presented in Tacconi et al. (2018); Liu et al. (2019b) and to a smaller degree to those reported in Scoville et al. (2017); Magnelli et al. (2020). As discussed earlier, the mild tension between the latter works and my results could be primarily attributed to the choice of the fitting function.

The $f_{\rm dust}$ and $f_{\rm gas}$ in my sample of SFGs increase rapidly from $z = 0$ to $z = 1$, peak around $z \sim 2 - 3$, and then remain roughly constant. It is however a point of contention whether the latter is driven by actual physical processes or is a consequence of the scarcity of data at $z > 2$. It it also worth mentioning that my analysis points towards a milder evolution of $f_{\rm dust}$ (-0.8 dex) compared to $f_{\rm gas}$ (-1.3 dex) from $z = 2$ to $z = 0$ with the latter dropping $\sim 3\times$ faster. This is aligned with the evolution of $\rho_{\rm dust}$ and $\rho_{\rm gas}$ derived by the ALMA stacking analysis of Magnelli et al. (2020), and could in fact reflect the evolution of metallicity, and thus of $\delta_{\rm GDR}$, for fixed M_* towards lower redshifts.

At the same time, the decrease of $M_{\rm dust}$ with decreasing redshift, for fixed M_*, can be attributed to either the destruction of dust grains by interstellar radiation fields, or their incorporation into the stellar population. This is discussed in more detail in Donevski et al. (2020), where they also report a decreasing $f_{\rm dust}$ from earlier cosmic age to the present epoch. I note that the observed trend could also mirror the overall decline in the SFRD in the Universe from $z = 2$ to the present day, that points towards lower star-formation activity and thus lower dust production at later cosmic epochs. Finally, at a fixed redshift, both $f_{\rm dust}$ and $f_{\rm gas}$ decrease as a function of M_* (as indicated by a negative value of the fitting parameter b, see Equation 2.2), in line with previous studies (e.g. Magdis et al. 2012a; Magnelli et al. 2020 and references therein).

In addition to observational studies, I can also compare my results to theoretical predictions. To this end, I consider the Horizon AGN (HAGN) hydrodynamical simulations in the COSMOS field (Dubois et al. 2014; Laigle et al. 2019) and draw a sample of SFGs (ΔMS>0.3) in the $z = 0.2 - 0.5$ range, selected to meet the M_* completeness of the COSMOS 2015 survey and which fall within a simulation box of 143 Mpc per side (Dubois et al. 2014). To measure ΔMS for each galaxy I considered the M_* and the 100 Myr averaged SFR from the simulations. Also, since $M_{\rm dust}$ is not an explicit paramater of HAGN galaxies, I used a constant $\delta_{\rm GDR} = 100$ to convert the $\mathrm{M}_{\rm gas}$ values, as derived from the simulations, to $M_{\rm dust}$. I then use the $\mathrm{M}_{\rm gas}$, M_*, and $M_{\rm dust}$ of the simulated galaxies to infer $f_{\rm gas}$ and $f_{\rm dust}$. The median values and their scatter, re-normalised to MS (ΔMS = 1) and $M_* = 5 \times 10^{10}$ $\mathrm{M}_\odot$ in four redshift bins, are presented and compared to the real data in Figure 2.11 and Figure 2.14. I find a good agreement between the theoretical predictions and my observationally driven trends (in the $< 0.2 < z < 0.5$ range at least) indicating that the HAGN simulation can successfully reproduce the baryonic components of the galaxies and its evolution with redshift. Conversely, the agreement of my results with both theoretical and observational studies, provides an extra indirect validation for the performance of my new SED fitting code.

2.9.2 On the Evolution of Depletion Time

As with $f_{\rm dust}$ and $f_{\rm gas}$, my recovered trends, that connect $\tau_{\rm depl}$ to redshift, ΔMS and M_*, show similar behaviour to the ones presented in Scoville et al. 2017, and to a lesser extent Tacconi et al. 2018; Liu et al. 2019b. The dependence of $\tau_{\rm depl}$ on M_* is relatively weak across all studies, however, similarly to Liu et al., I find that the depletion time for high-mass galaxies increases from early cosmic ages towards present times, while low-mass galaxies display an opposite trend of decreasing $\tau_{\rm depl}$ with cosmic age. As discussed in Liu et al. (2019b) and Hodge & da Cunha (2020), this could be a signature of downsizing, meaning that more massive galaxies evolve at earlier times.

During my analysis I find $\tau_{\rm depl} \sim (1+z)^{-1.07}$, which is more reflective of the scaling relations derived in Scoville et al. (2017) ($\tau_{\rm depl} \sim (1+z)^{-1.04}$), rather than weaker dependencies ($\tau_{\rm depl} \sim (1+z)^{-0.62}$ and $(1+z)^{-0.58}$) found by Tacconi et al. (2018) and Liu et al. (2019b) respectively. As expected, and in line with the literature results, I also find that galaxies above the MS (at a fixed M_* and z), form stars with a much higher efficiency (lower $\tau_{\rm depl}$), than their MS counterparts, with $\tau_{\rm depl} \sim \Delta{\rm MS}^{-1.68}$. I would also like to caution the reader and highlight the fact that Tacconi et al. and Scoville et al. use functional forms that are different from mine, when fitting for evolution of $\tau_{\rm depl}$. For example, Tacconi et al. consider additional dependence on the effective radius $R_{\rm e}$, which might inadvertently carry some redshift dependence. As such the fitted exponents are not necessarily directly comparable. The differences between evolutionary trends could also be attributed to the different samples used (see e.g. Hodge & da Cunha 2020).

Presumably, the existence of these outliers can be explained by an increased SFE, which results from major-merger events (see e.g. Scoville et al. 2017; Cibinel et al. 2019). In fact, galaxies that lie above the MS are also found to have increased gas fractions (Dekel et al. 2009; Tacconi et al. 2020), which is attributed to a more efficient gas accretion from the cosmic web, but the enhanced gas reservoirs are still not large enough to explain significantly enhanced sSFR. The debates regarding the exact reason, which results in an onset of a SB - like mode of star-formation, are still ongoing, however it seems very likely that it is a combination of both increased gas fractions and enhanced SFE. I find that my sample supports this notion, with galaxies above the MS having both large gas reservoirs with median log($f_{\rm gas}$)=0.15, meaning that gas mass reservoirs take up $\sim 59\%$ of the total baryonic matter, and also relatively short depletion times of ~ 400 Myr. My $M_{\rm gas}$ values were however derived with a general FMR, assuming solar-like metallicities. This, however, might not be applicable for SBs, which can display elevated metallicities due to the increased sSFR. In fact it has been shown (see e.g. Silverman et al. 2015), if SBs had super-solar metallicites, it would

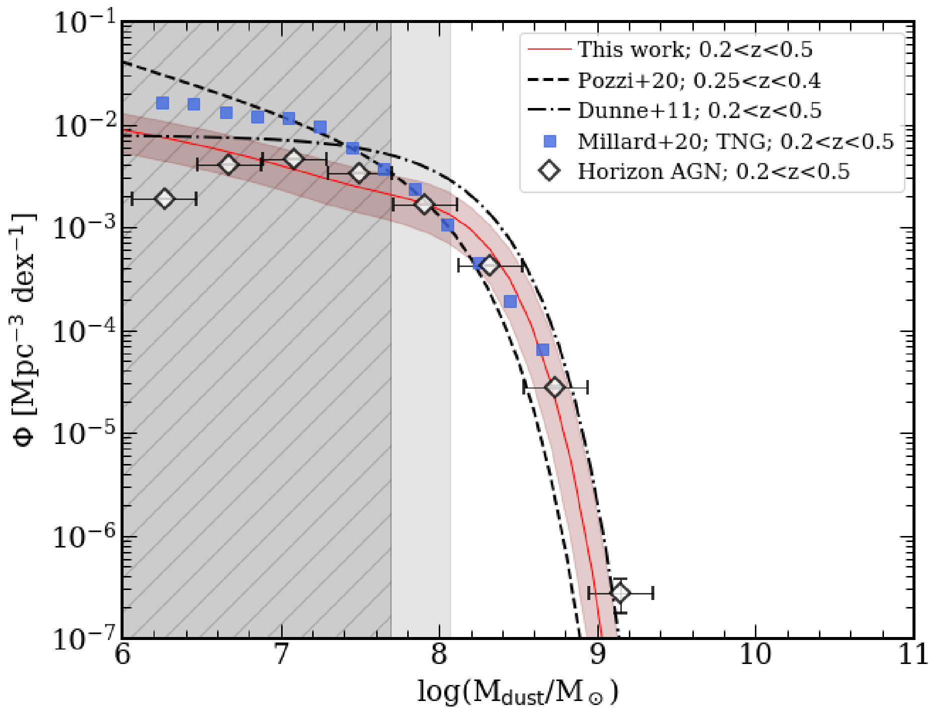

Figure 2.15: A compilation of the theoretical and observationally derived DMFs in the $0.2 < z < 0.5$ range. The Davidzon et al. (2017) SMF, converted to DMF, is shown as the solid red line, with the shaded area corresponding to the 1σ uncertainty. The dashed and dash-dotted black lines correspond to the DMFs of Pozzi et al. (2020) and Dunne et al. (2011) respectively, rescaled to $\kappa_{250\mu\mathrm{m}} = 0.51\ \mathrm{m}^2\ \mathrm{kg}^{-1}$. The white diamonds and the blue squares depict the theoretical predictions of the HAGN and TNG simulations from Dubois et al. (2014) and Millard et al. (2020) respectively. The grey shaded region highlights the M_{dust} regime below the lim(M_{dust}) of my sample, as derived in Figure 2.6. The hatched region denotes the M_{dust} regime where my sample becomes severely limited, i.e. $> 1\sigma$ below lim(M_{dust}).

drive down δ_{GDR} together with f_{gas}, and in turn result in increased SFE, thus implying that only the SFE is responsible for galaxies being elevated above the MS.

2.9.3 On the DMFs and the Theoretical Predictions

With the derived DMF in hand, I am also in position to bring together my findings with those presented in previous observationally driven studies and provide a direct comparison to the theoretical predictions as inferred by recent simulations. For these purposes, I focus on the $0.2 < z < 0.5$ redshift interval that contains the majority of my objects and offers the most robust statistical analysis. These results are shown in Figure 2.15.

I first compare my DMF to that presented in Pozzi et al. (2020), based on a PACS-160 μm selected sample of SFG. In Figure 2.15, the two DMFs appear to be in tension both in the high and the low-mass end, with my compilation over-predicting the number density of galaxies with high $M_{\rm dust}$ and under-predicting that of less dusty sources. The discrepancy between the two DMFs can be attributed to the choice of fitting methods/templates, the adopted κ_ν to infer $M_{\rm dust}$ as well as to selection effects. For example, the DL07 templates adopt a $\kappa_{250\mu{\rm m}}$ of 0.51 $\rm m^2\ kg^{-1}$, while the analysis in Pozzi et al. uses $\kappa_{250\mu{\rm m}} = 0.4\ {\rm m^2\ kg^{-1}}$, which would result in a 0.1 dex smaller $M_{\rm dust}$ estimates. It is also important to point out that Pozzi et al. compared their MBB SED fitting to MAGPHYS, finding that their MBB method recovers systematically lower $M_{\rm dust}$. Indeed, choice of the fitting methodology can induce up to a factor of two difference in derived $M_{\rm dust}$ (see e.g. Magdis et al. 2013 and an in-depth comparison in Berta et al. 2016). Moreover, for a flux limited survey, the mere selection at $\lambda_{\rm obs}$=160 μm could introduce a bias towards warmer sources that for fixed $L_{\rm IR}$ have lower $M_{\rm dust}$ and which could explain the small number density of sources with log($M_{\rm dust}$/M$_\odot$) > 8. While it is not possible to correct for the effects of selection and broader SED fitting methodology, I have rescaled the Pozzi et al. DMF to have the same $\kappa_{250\mu{\rm m}}$ as was adopted in my analysis.

I also compare my results to Dunne et al. (2011), that computed a DMF based on a sample of 250 μm selected galaxies. For my comparison, I have rescaled their DMF by -0.24 dex, to account for the difference in κ_ν. Contrary to Pozzi et al. (2020), I now find that Dunne et al. (2011) over-predicts the number density of dusty galaxies at high dust masses. This again can be understood in terms of selection effects since the 250 μm selection could bias the sample towards cold sources and thus to higher $M_{\rm dust}$ values (again, for a flux limited survey). While my criterion for at least one detection at $\lambda_{\rm rest} > 150\,\mu$m could be perceived as similar to a 250 μm selection at $z \sim 0.3$, I note that the requirement for two extra detections at $\lambda_{\rm rest} < 150\,\mu$m, and the super-deblendend catalogues, that allow for the detection of fainter than the nominal confusion noise in the SPIRE bands, ease any bias towards either intrinsically cold or warm objects. This is further supported by the fact that my SMF-derived DMF, where the Eddington bias has already been corrected, falls directly between the calculations from Pozzi et al. and Dunne et al. (Figure 2.15). In conjunction with the derivation of a $\langle U \rangle - z$ relation that is in excellent agreement with the stacking analysis of Béthermin et al. (2015), this suggests that the careful treatment of selection criteria and of the detection limits of the parent sample has allowed me to gain a unique and unbiased perspective on the evolution of dust properties of COSMOS galaxies.

Finally, I compliment my analysis by comparing my DMF to the theoretical predictions of the HAGN and IllustrisTNG simulations (Millard et al. 2020). In order to produce a HAGN DMF, I define a simulated sample following the procedure described above, bin the galaxies in 0.4 dex intervals of $M_{\rm dust}$, and normalise by the volume of the simulation $(4 \times (142)^3\ {\rm Mpc}^3)$. For the

IllustrisTNG simulation, Millard et al. (2020) consider multiple TNG100 snapshots in a box size of 106 Mpc per side, comparable to the HAGN simulated subset presented earlier. The TNG-DMF is constructed through the $M_{\rm dust}$ values of the simulated galaxies, derived in post-processing through a fixed dust-to-metals ratio of 0.5.

Unlike real data, simulations do not suffer from observational bias, and as such should be compared to DMF derived from the SMF, without adding the Eddington bias. As shown in Figure 2.15, both HAGN and IllustrisTNG are in excellent agreement with the high-mass end of my SMF-derived DMF. Notably, the HAGN-DMF is also consistent with my results at the low-mass end down to $M_{\rm dust} \sim 10^7$ M$_\odot$. One may recall, that for simplicity, when converting $M_{\rm gas}$ to $M_{\rm dust}$ for the HAGN sample, I considered a universal $\delta_{\rm GDR} = 100$. However, for sources with lower M$_*$ ($< 10^8$ $M_\odot$) and thus with sub-solar metallicities, a larger $\delta_{\rm DGR}$ ($\sim$ 150) is probably more applicable (Rémy-Ruyer et al. 2014). This would translate into a $\times 1.5$ downward correction for the low-mass HAGN bins, bringing them into exact agreement with my DMF down to $M_{\rm dust} \approx 10^7$ M$_\odot$. I note that this $M_{\rm dust}$, assuming an average $M_{\rm gas}/M_{\rm dust} \approx$ 100, corresponds to the M$_*$ completeness limit of the simulation ($M_* \approx 10^9$ M$_\odot$). Therefore, the observed decline of the number density of the HAGN galaxies at $M_{\rm dust} \leq 10^7$ M$_\odot$ is fully consistent with the expectations.

In comparison to the TNG-DMF though, I predict a factor of $\times 2.5$ fewer objects at the low-mass end. This tension could arise from the incompleteness of my sample at the low-mass end, that leaves the slope of the $M_{\rm dust} - M_*$ relation at $M_{\rm dust} < 5 \times 10^7$ M$_\odot$ largely unconstrained. I am thus unable to ascertain whether this discrepancy is caused by the limitations of the sample, or whether the TNG simulations over-predict the number density of the galaxies in the low-mass end.

Put together, these comparisons indicate that, at least down to $M_{\rm dust} \approx 5 \times 10^7$ M$_\odot$, my $M_{\rm dust}$−M$_*$ $-z$ relation and the resulting DMFs are robust and fully consistent with the theoretical expectations.

2.9.4 Population of Gas Giants

As briefly discussed in Section 2.8, during my analysis I identified some extreme outliers from the average $f_{\rm dust}$ and $f_{\rm gas}$ evolutionary trends (Figure 2.11 and Figure 2.14), and which typically have log($f_{\rm gas}$) > 0.5, i.e. their gas mass reservoir takes $\sim 75\%$ of their baryonic matter. Since z_{phot} could be a major source of uncertainty in both M$_*$ and M$_{\rm gas}$, before looking further into this population of 'gas-giants', I first narrow down my sample to spectroscopically confirmed sources. I then examined the individual SEDs and the cut-out images of the remaining sources in order to identify either poor coverage of the FIR peak or blending issues that could result in erroneously large M$_{\rm gas}$ estimates. With the above considerations, I was left with 41 objects whose extreme

$f_{\rm gas}$ can only be explained by gigantic $\mathrm{M_{gas}}$ reservoirs. This population spans a wide range in redshift ($0.21 < z < 4.05$, $\langle z \rangle = 1.34$), with $9.0 < \log(M_*/\mathrm{M}_\odot) < 11.3$, $\langle \log(M_*/\mathrm{M}_\odot) \rangle = 10.3$ and $0.11 < \Delta\mathrm{MS} < 14.2$, $\langle \Delta\mathrm{MS} \rangle = 1.8$. The best-fit SEDs of two such objects are presented in Figure 2.16 and Figure 2.17 . I also note that these two sources are otherwise unremarkable, and have what can be considered 'typical' values for the $\log(M_*) \sim 10.7$, and also do not appear to be strong SBs (ΔMS=3.8 and 2.2, respectively for 10041706 and 10100707). Furthermore, the cutouts presented in Figure 2.16 and Figure 2.17 indicate that these sources do not appear to be blended, therefore the only unusual characteristic that they possess, seems to be an elevated $M_{\rm gas}$.

A possible explanation for the very high $M_{\rm dust}$ and subsequently $\mathrm{M_{gas}}$ estimates for these galaxies, other than an extremely low $\delta_{\rm GDR}$, could be an optically thick FIR emission. In this scenario, the attenuation of the emission in the Wien part of the spectrum makes the galaxy *appear* cold, leading to an overestimate of its true $M_{\rm dust}$ (e.g. Jin et al. 2019; Cortzen et al. 2020). Since the DL07 models assume that the galaxy is optically thin at $\lambda_{\rm rest} > 1\,\mu$m, to test this scenario I employed Modified Black Body (MBB) models of general opacity, leaving the effective wavelength ($\lambda_{\rm eff}$) at which $\tau = 1$ as a free parameter (e.g. see Casey et al. 2012 for the functional form). I fixed the Rayleigh-Jeans slope to $\beta = 1.8$, and only fit the available photometry of each source at $\lambda_{\rm rest} > 40\,\mu$m. Due to the large number of free parameters in this model, I further limit our sample to sources with five or more IR detections. Out of the 41 'gas giants', I am thus able to constrain $\lambda_{\rm eff}$ for only 19. The distribution of the inferred $\lambda_{\rm eff}$ values is presented in Figure 2.18.

I find that the vast majority of these objects have $\lambda_{\rm eff} > 100\,\mu$m, which implies that these galaxies could be optically thick in the FIR. The unusually high $f_{\rm dust}$ and $f_{\rm gas}$, can therefore be incorrect simply due to the optical depth effects. Indeed, a comparison between the inferred $M_{\rm dust}$ estimates for an optically thin and optically thick case, yields an average ratio of $\sim \times 1.8$ for my sample. However, while this correction would reduce the average $f_{\rm gas}$ of the 'gas-giants' from $\langle \log(f_{\rm gas}) \rangle = 0.72$ to $\langle \log(f_{\rm gas}) \rangle = 0.47$, this is still substantially larger with respect to the average population of SFGs. Finally, the real $f_{\rm gas}$ could in fact be lower if the M_* of the sources is underestimated, a scenario that is indeed in line with a dusty, optically thick ISM.

To understand whether these objects do indeed host unusually high gas mass reservoirs and shed light on their nature, additional observations, with either ALMA or NOEMA, of $M_{\rm gas}$ tracers (e.g. CO and CI) are necessary.

2.10 Summary

In this work I present an in-depth analysis of the evolution of the FIR properties of SFGs by studying a large sample of sources drawn from the publicly available infrared catalogues in the GOODS-N and COSMOS fields (Liu et al. 2018; Jin et al. 2018). Both catalogues are constructed

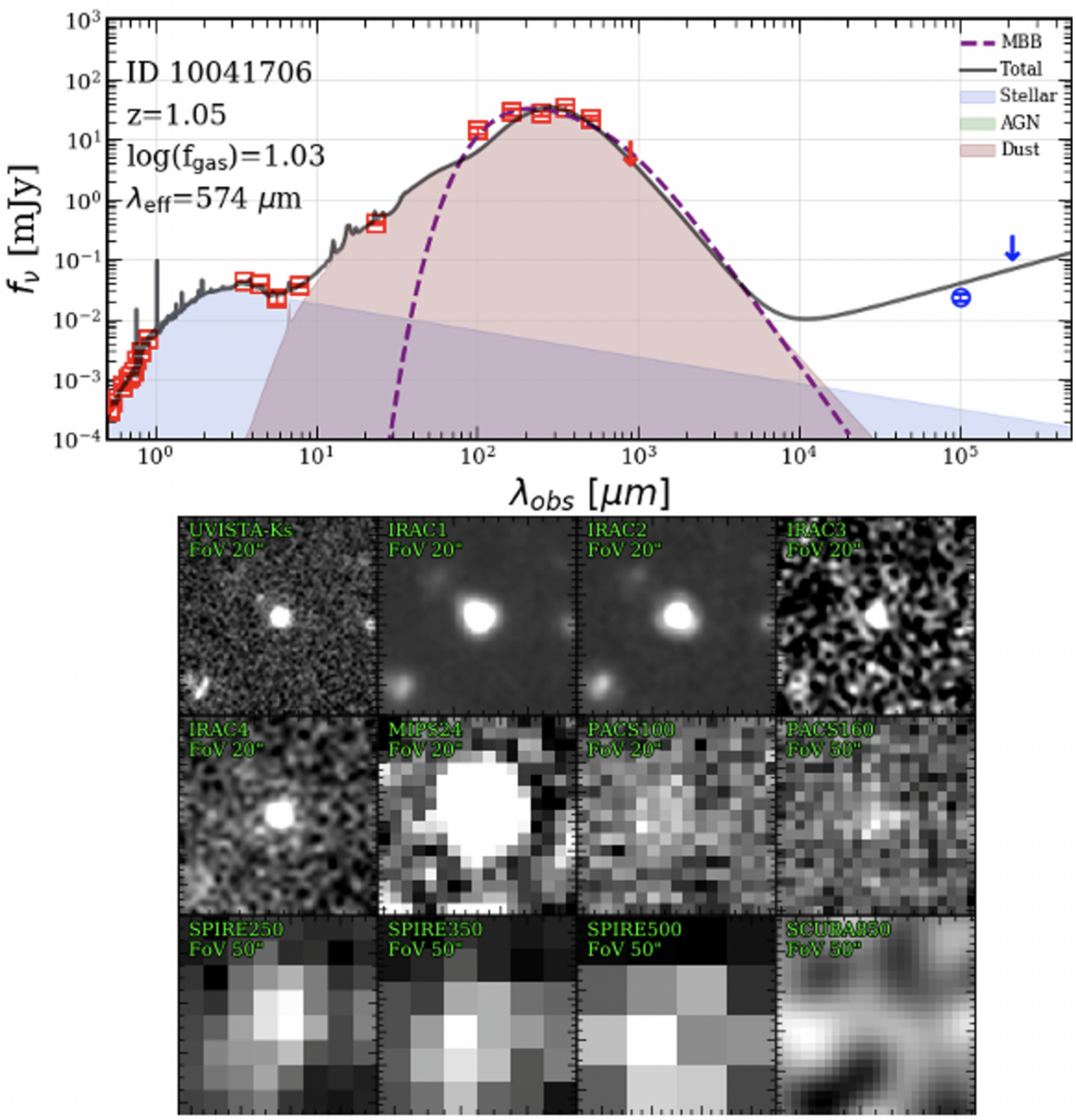

Figure 2.16: Top:. Photometry and best-fit SEDs for the 'gas-giant' ($\log(f_{\rm gas}) > 0.5$) at $z_{\rm spec}$ = 1.05. Colour coding and symbols are the same as in Figure 2.3, with the addition of a dashed purple line that shows the best-fit optically thick MBB. The $\lambda_{\rm eff}$ (in rest-frame) at which the SED becomes optically thick ($\tau = 1$) is displayed in the panels. **Bottom:** NIR-FIR cutouts of this galaxy. The cutout sizes range from $20''$ in the NIR-MIR range to $50''$ in FIR.

based on a novel 'Super-Deblending' technique that allows prior-based photometry in the highly confused *Herschel* and *SCUBA+AzTEC* maps.

In the process, I developed a new panchromatic SED fitting algorithm - `Stardust` - to fit a linear combination of stellar, AGN and infrared (star-forming) templates, in an attempt to perform a coherent, systematic and homogeneous analysis in the two fields. My fitting tool has two key

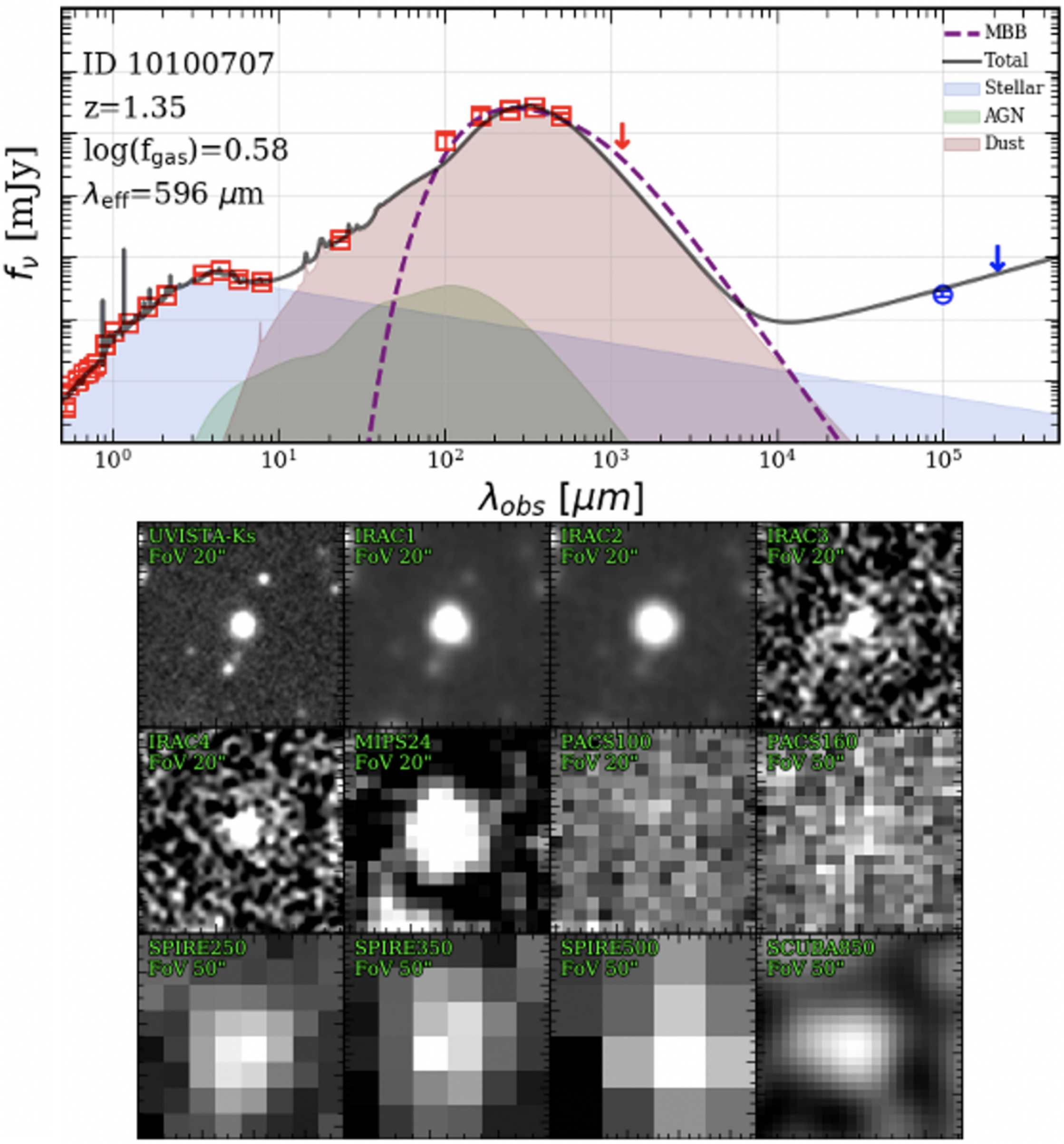

Figure 2.17: Photometry and best-fit SED for the 'gas-giant' ($\log(f_{\rm gas}) > 0.5$) at $z_{\rm spec}$ = 1.35. Colour coding and symbols are the same as in Figure 2.16

advantages. Firstly, the best-fit model is a set of coefficients rather than a single template, thus it does not rely on iterating through thousands of possible template combinations, speeding up the fitting process by a factor of ~ 10 compared to other multi-wavelength fitting available codes. Secondly, the fitting process does not impose energy balance between absorption in the UV/optical and emission in the IR, treating the stellar and the dust emission components independently. As such, it is very relevant for sources where the stellar and dust emission are not co-spatial. The

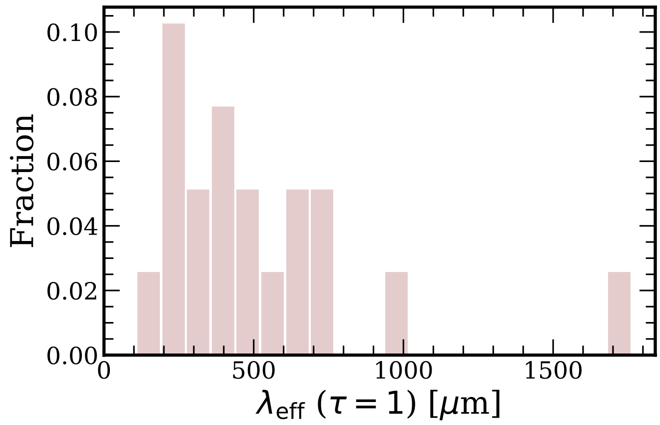

Figure 2.18: Distribution of $\lambda_{\rm eff}$, below which the emission of the 'gas-giants' in our sample becomes optically thick ($\tau = 1$), as inferred by MBB models of general opacity.

code itself is also highly modular, allows for user input templates, and is publicly available.

A first product of this new software is a multi-parameter catalogue that contains the FIR properties of $\sim 5,000$ infrared bright galaxies in GOODS-N and COSMOS. The extracted parameters, their uncertainties and the matched photometry from the original 'Super-Deblended' catalogues are released and are also publicly available [12]. The list of output best-fit parameters and the structure of the released catalogue can be found in Table A.2.

I subsequently used the extracted parameters to explore the evolution of the FIR properties of SFGs and recover scaling relations, aided by a careful set of simulations that quantify the underlying selection effects and biases of my sample in terms of limiting $M_{\rm dust}$, $L_{\rm IR}$ and $\langle U \rangle$. In particular, I parametrised the $f_{\rm dust}$ of galaxies as a function of their cosmic age, M_*, and ΔMS. The median $f_{\rm dust}$ is found to increase by a factor of $\times 10$ from $z = 0$ to $z = 2$ with a mild, if any, evolution at higher z. Through the metallicity dependent $\delta_{\rm GDR}$ technique, I also derive the evolution of $f_{\rm gas}$ and find it to be consistent with previous observational studies, as well as with theoretical predictions.

Furthermore, I constructed the DMF up to $z = 1$ by converting the SMF of SFGs to a DMF, through the evolution of $f_{\rm dust}$ and its scatter, as parametrised in our study. A comparison of the derived DMFs to the theoretical predictions of the HAGN and TNG100 simulations in the $0 < z < 0.5$ range reveals an excellent agreement down to a limiting $M_{\rm dust} \sim 5 \times 10^7\, \rm M_\odot$, where due to poor statistics I cannot adequately constrain the $M_{\rm dust}$-M$_*$ relation.

Finally, I identified a population of SFGs with extreme log($f_{\rm gas}$) > 0.5, that I coin as 'gas giants'. The $f_{\rm gas}$ excess of these galaxies compared to the average SFG population persists even

[12] Tables containing the DMFs can be accessed here:
https://github.com/VasilyKokorev/sdc_ir_properties

when opacity effects in the FIR emission are taken into account. Follow-up observations targeting alternative M_{gas} tracers are necessary to confirm the extreme nature of these systems.

Some further remarks that I would like to emphasise:

- The effect of the photo$-z$ uncertainty in the derivation of $M_{\rm dust}$(and $L_{\rm IR}$) is not negligible and should be accounted for. I find that a photo$-z$ uncertainty of Δz/(1+$z_{\rm spec}$) ~ 0.02, characteristic of fields like GOODS-N and COSMOS, introduces an extra 20% of scatter in the derivation of $M_{\rm dust}$ and $L_{\rm IR}$.

- As already discussed in the literature, the uncertainty in the derivation of $M_{\rm dust}$ increases substantially in the absence of a data point in the R-J tail ($\lambda_{\rm rest} > 150\,\mu$m). However, the presence of three data points in the mid-to-FIR could securely constrain $M_{\rm dust}$ within a factor of ~ 0.3 dex, even if the last available data point is at $\lambda_{\rm rest} \approx 150\,\mu$m.

- When using the $M_{\rm dust}$ – M_* – z scaling relations to convert SMF to DMF (or similarly to gas mass functions), the scatter of the relations used for the transformation should be taken into account for a proper comparison to the data. Similarly, any attempt to derive scaling relations between two (or more) parameters through the comparison of mass (or luminosity) functions inferred through the modelling of the observed number densities should entail a proper consideration of the scatter of the parameters in question.

- Both the warm $M_{\rm dust}$ and the warm IR emission arising from the PDRs are increasing with respect to the cold $M_{\rm dust}$ and cold dust emission as we move above the MS, indicative of more compact/active star-forming activity. Subsequently, the clear and relatively tight trend of decreasing $M_{\rm dust}^{\rm cold}$ (for fixed $L_{\rm IR}$) with ΔMS is less pronounced for $M_{\rm dust}^{\rm warm}$. This enforces the overall picture where SBs are characterised by higher star-formation efficiencies and with a larger fraction of their $M_{\rm dust}$ being exposed to more intense radiation fields.

CHAPTER 3

Massive Cosmic Lenses Reveal Unique Galaxy Populations

This chapter contains an advanced draft of the following article:
"A Uniform Processing of HST and IRAC Photometry Within 33 Lensed Fields Covered by ALMA".

To be submitted to The Astrophysical Journal Supplemental Series, May 2022

Authors: *Vasily Kokorev, Gabriel Brammer, Seiji Fujimoto, Georgios E. Magdis, Francesco Valentino, Pascal Oesch, Kotaro Konho, Iary Davidzon, Sune Toft, John R. Weaver and others*

3.1 Abstract

In this work I present a set of multi-wavelength mosaics and photometric catalogues in the ALMA lensing cluster survey (ALCS) fields. The catalogues were built by reprocessing the archival data from the CHArGE compilation, taken by the *Hubble Space Telescope* (*HST*) in the RELICS, CLASH and Hubble Frontier Fields. Following a similar procedure, I have reconstructed the *Spitzer* IRAC 3.6 and 4.5 μm mosaics, by utilising all the available archival IRSA/SHA exposures. Source detection and photometry for *HST* data was done by creating a weighted master image from all the available *HST* ACS/WFC and WFC3/IR filters, and then extracting flux within various apertures. To alleviate the effect of blending in such a crowded region, I have modelled the *Spitzer* photometry by convolving the *HST* detection image with the *Spitzer* PSF using a novel `golfir` algorithm. The final catalogues contain 218,000 sources, covering a combined area of 690 arcmin^2. These catalogues will serve as an important tool in aiding the search of the sub-mm galaxies in future ALMA surveys, as well as follow ups of the *HST* dark - IRAC sources. Coupled with the available *HST* photometry the addition of the 3.6 and 4.5 μm bands will allow us to place a better constraint on photometric redshifts and stellar masses of these objects, thus giving us an opportunity to identify high-redshift candidates for spectroscopic follow ups and answer the important questions regarding the epoch of reionisation and formation of first galaxies.

3.2 Introduction

The emergence of large multi-wavelength photometric surveys has allowed us to conduct detailed studies of galaxy formation and evolution across cosmic time by observing statistically significant population of galaxies. In particular, the investment of thousands of orbits of *Hubble Space Telescope* (*HST*) and *Spitzer Space Telescope* (*Spitzer*) time has cemented their unprecedented imaging legacy and enabled us to revolutionise our understanding of both observational cosmology and galaxy evolution. Compared to ground based observations, the data taken in space do not suffer from the distortion and absorption effects of the atmosphere, thus achieving remarkable sensitivity and wavelength coverage. For example, these unique capabilities allowed us to capture the accelerating expansion of the Universe (Riess et al. 2004), and have helped demonstrate that the majority of star formation took place within a relatively short time span, in the epoch at $1 < z < 3$ (see e.g. Hopkins & Beacom 2006; Bouwens et al. 2007).

More recently, the advantages of space-based observations have become particularly pronounced in the search for high redshift galaxies, with the combined efforts of the very sensitive WFC3/IR camera onboard *HST* and the ultra deep *Spitzer*/IRAC imaging. The remarkable wavelength coverage of these instruments has helped us push the observational frontier to the beginning of the cosmic epoch of reionisation at $z \sim 7-8$, some 700 Myr from the Big Bang, and beyond toward the epoch beyond $z \sim 10$, where the formation of the first galaxies has taken place. A number of large deep extragalactic blank field surveys has now led to a discovery of significant and statistically meaningful number of galaxies at $z \sim 7-8$ (e.g. McLure et al. 2013; Finkelstein et al. 2015; Bouwens et al. 2015), an ever growing sample of $z \sim 9-11$ candidates (Ellis et al. 2013; Oesch et al. 2013, 2014; Bouwens et al. 2016; Calvi et al. 2016), and even the most distant galaxy discovered to-date at $z = 11.1$ (Oesch et al. 2016). The most staggering and impactful discoveries of high-z galaxies have however been made within lensing cluster fields, which include Hubble Frontier Fields (HFF; Lotz et al. 2017), the Reionization Lensing Cluster Survey (RELICS; Coe et al. 2019), and the Cluster Lensing And Supernova survey with Hubble (CLASH; Postman et al. 2012). All three combine the power of *HST* and *Spitzer* observations and a strong gravitational lensing potential of massive galaxy clusters, to produce the deepest available observations of high-z galaxies lensed by clusters ever obtained (e.g. see Zheng et al. 2012; Coe et al. 2013; Bradley et al. 2014; Schmidt et al. 2014; Zitrin et al. 2014; Infante et al. 2015; Ishigaki et al. 2015; Kawamata et al. 2015; Oesch et al. 2015; McLeod et al. 2015; Hashimoto et al. 2018; Hoag et al. 2018; Strait et al. 2020).

The redshift estimates of these objects still largely rely on spectral energy distribution (SED) fitting of broad-band restframe UV+optical photometry from *HST*/WFC3 and *Spitzer*/IRAC, and spectroscopically confirmed samples of high-z candidates remain limited. The SED fitting photometric redshift technique is largely leveraged on the correct identifications of either the Lyman or Balmer breaks, at 912 Å and 3640 Å respectively, in the stellar continuum. At $z \sim 9-10$ the *Spitzer*/IRAC targets the $\sim 3000-4000$ Å rest frame continuum and as such can greatly aid in removing the low-redshift interlopers from the high-z samples. Moreover, even for spectroscopically confirmed objects the *Spitzer* observations are essential for conducting robust measurements of the stellar population parameters, such as stellar mass (M_*), dust attenuated star formation rate (SFR), and extinction (A_V) (González et al. 2011; Ryan et al. 2014; Salmon et al. 2015). The existing data has already lead to implications that the first star-formation has taken place ~ 250 Myr after the Big Bang (Hashimoto et al. 2018).

Some questions regarding star formation, are however yet to be answered. From the rest-frame UV and FIR observations conducted during the last decade (Le Floc'h et al. 2005, 2009) we know that star formation activity in Universe was more substantial in the past, than it is now. Tracing the evolution of the volume-weighted SFR - the SFR density (SFRD) with time, presents us with a view of the Universe where the star formation has already reached its peak at $z \sim 2-3$ and is now declining (e.g. Madau & Dickinson 2014; Casey et al. 2021). Studies of the SFRD also show a growing disparity between the contribution of dust obscured SFR, measured from the IR data, and the unobscured SFR, measured from the UV - optical data. On one hand, observations with the

Atacama Large Millimeter/submillimeter Array (ALMA), such as the MORA project (Casey et al. 2021), suggest that, as we move towards higher-z, the SFR becomes dominated by unobscured SFR, primarily through UV - selected galaxies. This, in return, might imply that the early Universe was less dusty, aligning with some of our predictions regarding the time-scale and mechanisms of dust production. On the other hand, blind ALMA studies of galaxies at $z \sim 2 - 6$ (Wang et al. 2016, 2019; Gruppioni et al. 2020) reveal a population of optically dark, dusty sources, which contribute ten times more towards the SFRD, than similarly bright galaxies with a rest-frame UV detection, and reside within centres of the most massive matter overdensities. The ubiquity of such systems, can potentially create obstacles in our understanding of the true numbers of massive galaxies and SFRD in the early Universe. In addition to that, the existence of high number of massive, evolved, and dusty sources only a few billion years after the Big Bang challenges our current understanding of galaxy formation. Moreover, there are implications that the optically dark galaxies that we see at $z > 3$ could also potentially act as the progenitors of bright central galaxies (BCGs) within galaxy clusters. As a result, our ability to correctly recover the total SFR comes down to the detection of high-z dusty galaxies, such as "*HST* - dark", or optically dark sources (Wang et al. 2016; Fudamoto et al. 2021).

The complete dust obscuration of UV-optical emission, makes the detection and identification of these objects a significant challenge, even when sufficient MIR and FIR coverage are available. The brightest, most massive optically dark sources can be detected by *Spitzer*/IRAC, however poor resolution can result in blending and erroneous assignment of flux if nearby objects are present. The lack of the UV - optical continuum emission in "*HST* - dark" galaxies makes it impossible to directly compute the stellar mass, and (optically derived) metallicity. It also renders any photometric redshift technique unfeasible, with the spectral line scan remaining as the only option to reliably secure redshifts. As a result the only reliable way to observe such objects are blind field studies with ALMA or the NOrthern Extended Millimeter Array (NOEMA).

ALMA in particular, has been the primary tool driving the discovery of new faint sub-mm galaxies, ($S_{1.2\mathrm{mm}} \sim 0.02 - 1$ mJy), which are substantially fainter compared to the traditionally observed sub-mm galaxies (SMGs) (see e.g. Hatsukade et al. 2013, 2016; Ono et al. 2014; Carniani et al. 2015; Oteo et al. 2016; Fujimoto et al. 2016; Aravena et al. 2016; Dunlop et al. 2017; González-López et al. 2017; Franco et al. 2018). Despite such depths, the origin of the cosmic infrared background (CIB), and therefore the majority of SFRD still remains hidden, with the significant discrepancy between the fain-end of mm number counts. Over the last few years, ALMA observations of faint 1.2 mm sources have been able to derive ALMA mm counts down to depths of ~ 0.02 mJy (Carniani et al. 2015; Fujimoto et al. 2016; Hatsukade et al. 2016; Aravena et al. 2016), and even further down to 0.013 mJy, in the Hubble Frontier Fields (Muñoz Arancibia et al. 2018). All of these studies have managed to roughly account, and discretize, between 15 –

50 % of total CIB contribution. It has quickly become apparent that deeper ALMA observations are absolutely imperative in order to separate and resolve the remained of the CIB into discrete sources, in order to study their individual properties. The most efficient way to complete this puzzle is to observe a sufficiently large number of lensing clusters using ALMA.

The ALMA Lensing Cluster Survey (ALCS) aims to do exactly that. At the moment, it is the largest, by area, among other ALMA surveys targeting clusters of galaxies. Combined with previous ALMA observations, the survey covers a total of 33 massive galaxy clusters. The primary goals of ALCS are: (1) Examine the faint-end slope of 1.2 mm source counts, (2) Explore the types of galaxies responsible for the faint sub-mm emission in lensed fields, (3) Study the redshift distribution of the faint sub-mm population, (4) Constraining the ionised carbon (CII) and carbon monoxide (CO) luminosity functions from the observed spectral lines, (5) Calculate the mm-properties of different types of star-forming galaxies, and perform stacking analysis in case of non-detections.

All of these objectives however require additional data, in order to derive photometric redshifts, and relevant physical parameters. Fortunately, all ALCS clusters have been previously imaged with *HST*, and `Spitzer`/IRAC, enabling accurate positions and other quantities derived from the photometry. In this work I describe the uniform processing of all archival *HST*, and `Spitzer`/IRAC mosaics, covering the ALCS. I perform careful aperture photometry of all *HST* sources, and use them as priors to model and fit the fluxes for the sources in the blended IRAC maps. The final images and catalogues will then act as a powerful tool to establish better constraints on photometric redshifts and physical properties of these objects. In addition, this will allows us to identify high-redshift candidates for spectroscopic follow ups and answer the important questions regarding the epoch of reionisation and formation of first galaxies.

I provide combined catalogues which include photometry, photometric redshifts, and stellar population properties, for each field included in ALCS (similarly to the ASTRODEEP collaboration Merlin et al. 2016; Di Criscienzo et al. 2017 and the HFF-DeepSpace catalogues Shipley et al. 2018, albeit following a different methodology). Moreover, the public release of my data is complemented by all the new *HST*/ACS, *HST*/WFC3 and IRAC mosaics, including detection images, models, residuals and segmentation maps.

This chapter is organised as follows. In Section 3.3 I list the data sets used in this work, and describe the creation of new *Spitzer*/IRAC mosaics. In Section 3.4 I describe the high- and low-resolution photometry algorithms. In Section 3.5 I describe the catalogue format, ALMA counterparts, and quality flags. Section 3.6 presents the quality and consistency check for my catalogue. In Section 3.7 I describe the spectral energy distribution fitting of the extracted photometry, resultant photometric redshifts, stellar population parameters, and rest-frame colours. Finally my main conclusions and summary are given in Section 3.8.

Throughout this paper I assume a flat ΛCDM cosmology with $\Omega_{\rm m,0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$ and $\rm H_0 = 70$ km s^{-1} Mpc^{-1}, and a Chabrier (2003) initial mass function (IMF). All magnitudes are expressed in the AB system (Oke 1974).

3.3 Data Sets

The ALCS is a Cycle 6 ALMA large program (Project ID: 2018.1.00035.L; PI: K. Kohno) targeting 33 lensing cluster fields in Band 6 ($\bar{\lambda} = 1.25$ mm/$\bar{\nu} = 243$ GHz). In total ALCS covers an area of 134 arcmin^2, with a depth of 70 μJy (1σ). The sample is designed to be contained within the best-studied massive clusters also imaged in *HST* programs. More specifically, ALCS includes 5 clusters from the Hubble Frontier Fields (HFF; Lotz et al. 2017), 16 clusters from the Reionization Lensing Cluster Survey (RELICS; Coe et al. 2019), and 12 clusters from The Cluster Lensing And Supernova survey with Hubble (CLASH; Postman et al. 2012). The observations aim to provide an in-depth look of the high magnification regions within the cluster fields, and, in particular, target dust-continuum-selected and line-emitting high-z galaxies. The main science goal of the survey is to examine the faint-end slope of the 1.2 mm source counts, to provide the best estimate for the cosmic infrared background (CIB) at that wavelength. The typical galaxies contributing to the CIB at 1.2 mm are intrinsically faint (see e.g. Fujimoto et al. 2016), however in conjunction with the rich *HST* and *Spitzer*/IRAC data-sets covering the field, the survey aims to reveal the fundamental physical properties of the $S_{\rm 1.2mm} < 0.1$ mJy galaxies, such as stellar masses and IR based SFRs.

A vast majority of objects within ALCS are only continuum detected, ruling out redshift constraints from spectra. However, the rich UV-optical and NIR treasury data already collected within these cluster fields, will allow for the derivation of photometric redshifts for a vast majority of faint sub-mm galaxies. This will allow for the derivation of the dust-based FIR luminosity functions, and give upper limits on its evolution at $z > 3$, something that has been out of reach for $Spitzer$ and $Herschel$. The redshifts measurements for faint dusty galaxies are expected to provide a significant constraint on, or perhaps even challenge the current interstellar dust formation models (see e.g. Mancini et al. 2015; Mancuso et al. 2016; Wang et al. 2017; Aoyama et al. 2017), especially if a high numbers of $z > 7$ dusty SFGs are found (Watson et al. 2015; Knudsen et al. 2016).

The ALCS is also expected to detect the ionised carbon (CII) and carbon monoxide (CO) emission lines, facilitating the studies of the ISM for a unique sample of faint galaxies magnified by lensing clusters. Moreover, even for non-detections the mm-properties of various classes of SFGs can be extracted through stacking, facilitated by the presence of both *HST* and IRAC priors within these lensed fields. For example, the dust emission properties in LBGs above z=3 are poorly understood (e.g. see Reddy et al. 2010; Bouwens et al. 2016), with more recent ALMA observations suggesting that the dust continuum emission of $z \sim 5-6$ is fainter than that of local starbursts (Capak et al. 2015). In conjunction with the wealth of ancillary *HST* and *Spitzer* data, one can not only identify LBGs, but also conduct a dust stacking analysis by using the deep ALMA observations collected by ALCS.

An additional emphasis of the survey is to detect and characterise the magnified ALMA continuum sources, without *HST* counterparts, i.e. the intrinsically faint, "HST-dark", ALMA sources. The majority of these objects also have faint counterparts in the $Spitzer$/IRAC bands, with the measured IRAC to 1.2 mm flux ratios pointing towards these sources being either distant $z = 4-6$ galaxies, or massive bCG progenitors at $z \sim 4$, similar to the objects presented in Wang et al. (2016, 2019) and Gruppioni et al. (2020).

3.3.1 *HST* Imaging

The combined *HST* dataset has been processed with the `grizli` pipeline, which creates filter mosaics for all ACS, WFC3/UVIS and WFC3/IR exposures that cover a given area of the sky (e.g., an ALCS field). The overlapping exposures are then broken into discrete "visit" associations, where the grouping is done for a given filter for data that were taken in a single target acquisition. These associations generally share the same spacecraft orientation and zodiacal sky background.

Within a single visit, all exposures are aligned to each other using high S/N sources detected in them, allowing the relative `x` and `y` coordinates between exposures to shift, until the best match is found. These are analogous to `DrizzlePac TweakShifts` and are generally a fairly small fraction of a pixel for dither offsets within a single orbit and a few tenths of a pixel between subsequent orbits that share the same initial target acquisition.

A source catalogue is created from a preliminary mosaic generated from the visit exposures and aligned (shift, rotation, scale) to an astrometric reference catalogue. Generally this is PanSTARRS DR1 as it is well aligned to the GAIA DR2, but has a higher source density than the bright GAIA stars alone.

The final fine alignment is performed simultaneously optimising a) alignment between all of the individual visit catalogues, and b) GAIA DR2 stars with proper motions projected to each visit observation epoch. This ensures robust internal alignment of the *HST* images for matched-aperture photometry and the final absolute astrometric precision is generally < 100 mas.

A pedestal sky background of each exposure is estimated in the AstroDrizzle preparation of each visit association. A smooth background is subtracted from each visit mosaic to remove gradients that can then appear as sharp discontinuities in the final combined filter mosaics. This background is estimated with the `sep` (Barbary 2016), a non-executable software extension of the `SourceExtractor` (Bertin & Arnouts 1996) with `BACK_FILTERSIZE = 3` and `BACK_SIZE = 32 arcsec`. While the background estimation includes a mask for detected sources, it can include extended structure for very large, bright galaxies and inter-cluster light (ICL) in the ALCS cluster fields.

Final rectified mosaics combining all exposures in each available filter are created with `AstroDrizzle`. All WFC3/IR mosaics are created with 0.1" pixels, while the ACS and UVIS optical/UV images are drizzled with 0.05" pixels on a grid that subsamples the IR mosaic 2×2. Both optical and IR mosaics are drizzled with `pixfrac = 0.33`. The `sci` (science) and `wht` (inverse variance weights) mosaics are provided for each filter.

While this approach might not necessarily result in the best reconstruction of the undersampled *HST* PSFs, but the larger pixels ensure more uniform weights across the diversity of dither coverage across the survey fields (e.g., hundreds of exposures for the Frontier Fields but as few as 2–4 exposures for some pre-RELICS filters). The larger pixels and small `pixfrac` result in lower correlated noise between adjacent pixels, and therefore the inverse variance maps are more reliable estimate of the pixel variances for, e.g., aperture photometry.

The units of the filter mosaics are in electrons/s, with the photometric calibration to cgs units provided in the PHOTFLAM (f_λ) and PHOTFNU (f_ν) header keywords of each mosaic. I additionally provide PSF models for each IR filter using the effective PSF models described in Anderson & King (2000).

3.3.2 IRAC imaging

I begin by collecting all the Basic Calibrated Data (BCD) exposures (pBCD), from the *Spitzer* SHA archive for each field in the ALCS. These only include *Spitzer* IRAC 3.6 and 4.5 μm, as the fields of interest do not have a uniform coverage in 5.8 and 8 μm. Similarly to how I process the *HST* CHArGE data, I perform relative alignment of all exposures in the IRAC Astronomical Observation Requests (AOR), and then align AOR mosaics to GAIA DR2. A larger size of IRAC FoV is generally beneficial, as it allows for a sufficient number of reference sources to be used for alignment.

The IRAC background is removed by creating a master background image for each AOR with detected sources masked. Finally, the complete IRAC mosaics are aligned to the *HST* pixel grid and drizzled with 0."5 pixels and the pixel fraction parameter `pixfrac` set to 0.2.

With full knowledge of the individual pBCD exposures that contribute at any location in the final mosaic, I can generate robust models of the IRAC PSF that fully account for the diversity of depth and detector position angle across the mosaic. These position-dependent PSFs are used for the IRAC model-based photometry, and can be re-generated using `golfir`.

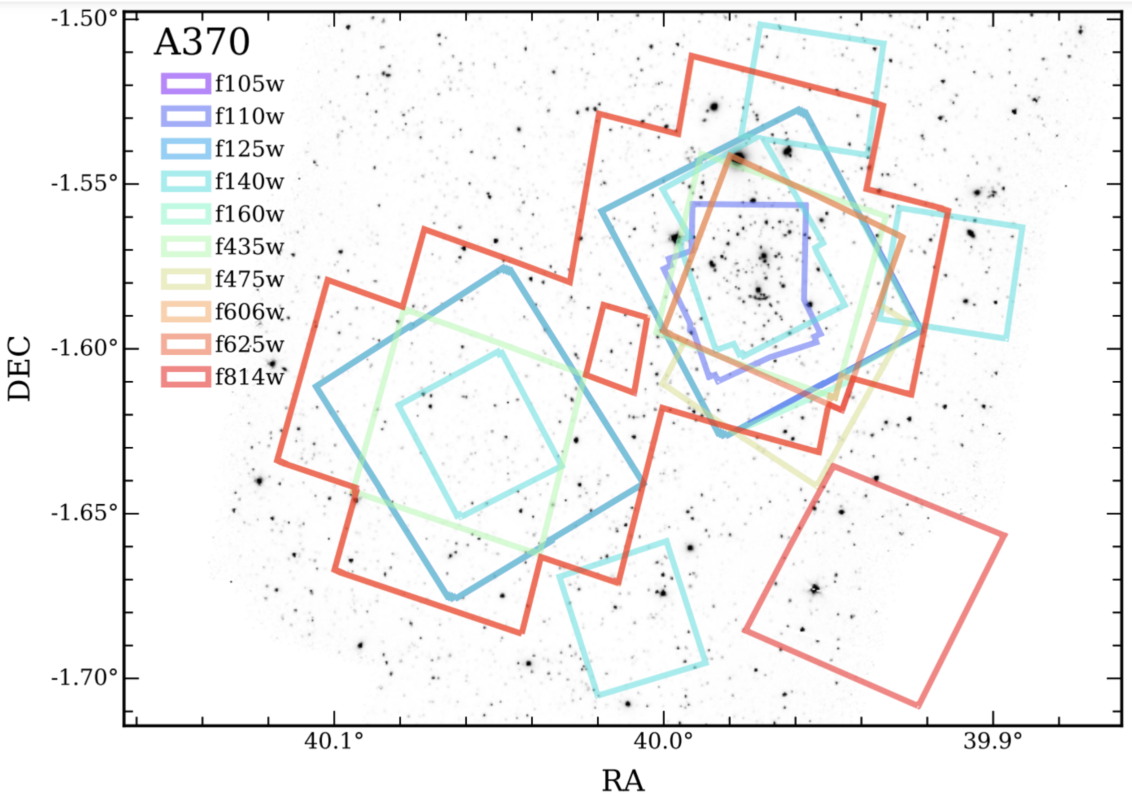

Figure 3.1: Layout of the *HST* and *Spitzer*/IRAC observations for the cluster and parallel observations of the Abell 370 field in the HFF. The footprints show of all the available coverage with *HST*, superimposed on the *Spitzer*/IRAC 3.6 μm data. North is up and east is to the left.

3.4 Photometry

During the generation of this catalogue, I have treated the highly resolved (HST/ACS and $WFC3$) data differently from the low-resolution $Spitzer/IRAC$ data.

3.4.1 *HST* photometry

I extract aperture photometry within (circular) aperture diameters 0.36, 0.5, 0.7, 1.0, 1.2, 1.5, 3.0 arcsec at the positions derived in the source detection as described above. I do not perform PSF-matching for any HST filters for the aperture measurements. This PSF-matching approach is quite common in the literature (e.g. see Shipley et al. 2018), however I feel that performing it on the images can result in substantial deleterious effects on the noise properties of the derived photometry. In particular, faint and/or dropout sources would be most affected, where there is no signal to use in matching. I therefore favour the approach here, primarily for simplicity and consistency across a wide variety of fields, with different noise properties, and defer tests on the aperture effects to ongoing work.

To compute aperture corrections I have defined the "total" *HST* flux within an elliptical Kron aperture determined by `sep`, as in `SourceExtractor`. However, I do not impose the lower limit of 3.5 on KRON_RADIUS typical with `SourceExtractor` as I find that in fact most derived values are actually lower than this threshold even for bright, well-measured sources. I do, however, impose a minimum circularized Kron aperture diameter of 0.7 arcsec, which is the favoured "color" aperture. I calculate a correction for flux outside of the Kron aperture using the PSF curves of growth (i.e., explicitly valid only for point sources). The "colour" aperture fluxes are therefore corrected by 1) `flux_auto / flux_aper` in the detection band and then 2) by the Kron aperture correction.

3.4.2 IRAC photometry

Calculating the photometry from low-resolution data, particularly in crowded regions containing galaxy clusters can be a difficult technical challenge. In order to correctly extract the flux in the redder bands the significant differences that exist between the *HST* data and the much lower resolution *Spitzer*/IRAC image data must be taken into account. Primarily this concerns the problem of blending, wherein a standard aperture photometry approach normally used for high-resolution data would be inadequate. To tackle these issues in my work, I use the Great Observatories Legacy Fields IR Analysis Tools or `golfir`, a set of tools developed to model *Spitzer*/IRAC and MIPS images based on high resolution templates from existing *HST* imaging, specifically in the context of the CHArGE data.

The method follows an approach utilized in similar lensing cluster catalogue works (see e.g. MOPHONGO; Skelton et al. 2014; Shipley et al. 2018 and T-PHOT; Merlin et al. 2015, 2016), and relies on using a high-resolution prior. I create this prior by combining all the available *HST* ACS/WFC and WFC3/IR filters, to produce a weighted mosaic, based on their corresponding inverse variance maps. I then use a convolution kernel to combine the detection image with the *IRAC* PSF and produce the low-resolution templates. The original IRAC image is then divided into square patches, of user defined size. These patches are also allowed to overlap, to allow for correct modelling of sources at the patch boundaries. For consistency I used a patch size of 1.2 arcmin with the 0.4 arcmin overlap for all the fields in my work. These parameters were chosen to achieve a balance between the quality of the final model image, and the available computational resources.

In order to improve the quality of extracted photometry I manually mask the brightest stars in a given field, prior to conducting the least square fit. I use the GAIA DR2 archive for the positions, and scale the size of the circular mask depending on the G-band magnitude of each star. I also mask all the pixels in the IRAC mosaic, for objects where the *HST* catalogue is brighter than AB =

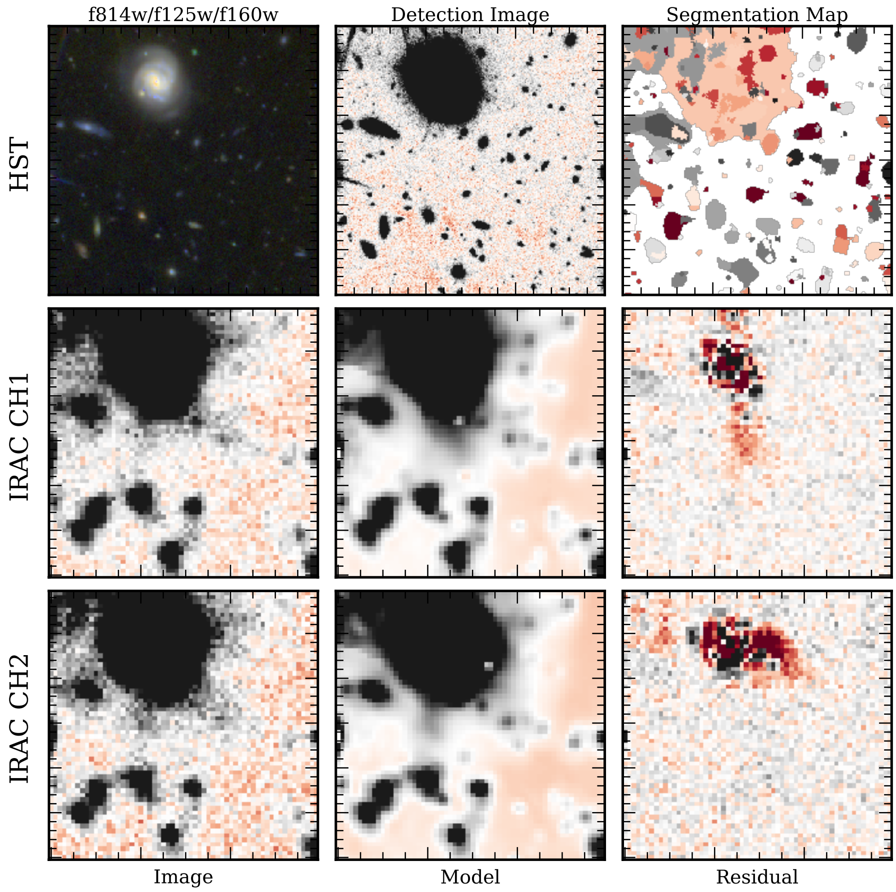

Figure 3.2: Results from the modelling procedure on the relatively crowded Abell 370 cluster field. The cutouts are 30" across, and were selected to show a wide variety of sources on the same image, i.e. the ones modelled purely with `golfir`, and the galaxy model refined by using GALFIT. **Top:** The *HST* RGB image created from the combination of F814W, F125W and F160W filters, the *HST* detection image and segmentation map. **Middle:** The original IRAC 3.6 μm science image, the `golfir` model and residual images. **Bottom:** Same as above, but for IRAC 4.5 μm filters. The colour scale has been adjusted to show the 3σ range for each image. Most notably, the final residual mosaic can also be used to extract additional IRAC photometry for sources without an HST counterpart. The images are oriented with north up, and east to the left.

15. This is done to avoid large residuals in the centers of bright stars that have not been manually masked. Additionally I mask all pixels where the IRAC S/N is above 80.

For the first model pass I generate IRAC model images for all objects in the *HST* detection catalog, brighter than AB = 24, by using the aforementioned convolved *HST* source cutouts, transformed into (position - dependent) IRAC PSFs. A least square fit of the low resolution IRAC cutouts is then performed to the real data, where the normalization acts as the only free parameter. I additionally derive any small residual shift between the reference *HST* and target IRAC mosaics, by using the generated IRAC model and mosaic images.

For the second model pass, I now focus on fainter *HST* galaxies, with AB < 27. IRAC models are then generated as before, with the least square normalisations now being adopted as the IRAC flux density measurement for each source. The diagonal of the covariance of the model design matrix is adopted as the photometric variance. An additional model refinement with GALFIT (Peng et al. 2002), is performed on sources for which the total SN is more than 50. For this I have used a single Sérsic model (Sérsic 1963). This often significantly improves the IRAC residuals, which is likely a combination of 1) an imperfect transformation between the HST and IRAC PSFs and 2) true morphological differences between F160W and the IRAC bands, for example the colour gradients. I note that I am not interested in the GALFIT parameters, but rather in the best empirical description of each IRAC morphological component. I do not adopt the "mag" of the GALFIT fit, but rather refit the model normalisations and covariances as in the previous steps, now using the GALFIT model cutouts in place of the *HST*-based models for the sources that have them. I show an example of the images processed with my pipeline in Figure 3.2.

As the IRAC fluxes are based on morphological model fits, I consider them to be on the same "total" scale as the aperture-corrected *HST* photometry. These are the IRAC fluxes that I will use for all my future data analysis. In addition to the model flux fits, I also perform a simple aperture photometry measurement on my images, by using a D = 3.''0 apertures, similar to the approach taken in Shipley et al. (2018). Using the IRAC PSF curves of growth I then correct the aperture fluxes into total fluxes. The model, aperture and aperture corrected flux measurements are all available in the final version of the catalog. In Table 3.1 I list all the names, coordinates and coverage areas of all 33 cluster fields covered by ALCS. I would also like to note that for the Hubble Frontier Fields, I do not make a distinction between the parallels or the cluster, and rather treat the entire mosaic, and objects within it, to be contained within a single field.

3.5 Catalogue

3.5.1 ALMA Photometry

To further complement my photometric catalogue, I have included an additional data entry containing either the measurement or the upper limit on the ALMA flux. To do this, I start by cross-matching my objects with the ALMA continuum catalogue (Fujimoto et al. 2022; in prep). This catalogue contains 180 sources, which have been selected with a S/N cut > 4. The total fluxes are computed as a peak count, after primary beam correction in the tapered map [mJy/beam]. If no peak is identified in the tapered map within a radius of 1.0 arcsec, the pixel count at the position of the source is used instead. I do not match these sources automatically, as this approach is inadequate for nearby or highly magnified sources with complex image plane morphology. In addition to the the astrometry difference, the size of the ALMA beam itself

Table 3.1: Lensed cluster fields covered by ALCS

Field	RA (h:m:s)	Dec (d:m:s)	Science Area† (arcmin^2)	3.6/4.5 μm Area (arcmin^2)	ALMA (1.2 mm) Area (arcmin^2)	Redshift z_{spec}
ALCS: Hubble Frontier Fields						
Abell S1063	22:48:44.40	-44:31:48.50	67.8	165.1	2.3	0.348
Abell 370	02:39:52.90	-01:34:36.50	74.4	197.9	3.3	0.375
MACSJ0416.10-2403	04:16:08.90	-24:04:28.70	68.1	191.3	2.3	0.396
Abell 2744	00:14:21.20	-30:23:50.10	64.7	172.3	2.7	0.308
MACSJ1149.5+2223	11:49:36.30	22:23:58.10	37.6	190.5	2.6	0.543
ALCS: RELICS						
RXCJ0032.1+1808	00:32:11.05	18:07:49.03	11.6	54.0	6.4	0.396
Abell 2537	23:08:22.23	-02:11:32.40	11.5	55.0	2.0	0.297
Abell 3192	03:58:53.07	-29:55:44.75	11.5	55.0	4.0	0.425
MACSJ0553.4-3342	05:53:23.09	-33:42:29.91	11.5	48.2	6.9	0.430
RXC J0600.1-2007	06:00:09.77	-20:08:08.94	11.7	57.6	5.8	0.460
RXC J0949.8+1707	09:49:50.89	17:07:15.27	11.6	55.4	2.6	0.383
MACSJ0257.1-2325	02:57:10.24	-23:26:11.84	13.3	57.4	1.7	0.505
Abell 2163	16:15:48.33	-06:07:36.70	22.0	97.7	1.5	0.203
PLCK G171.9-40.7	03:12:56.90	08:22:19.20	11.5	39.7	3.8	0.270
SMACSJ0723.3-7327	07:23:19.45	-73:27:15.63	12.2	50.7	1.6	0.390
MACSJ0035.4-2015	00:35:26.96	-20:15:40.31	12.0	58.2	2.3	0.352
MACSJ0417.5-1154	04:17:33.74	-11:54:22.64	11.7	42.1	5.0	0.443
MACSJ0159.8-0849	01:59:49.40	-08:50:00.00	11.5	57.9	2.2	0.405
ACT-CLJ0102-49151	01:03:00.01	-49:16:22.19	22.9	106.6	4.4	0.870
AbellS295	02:45:31.39	-53:02:24.85	11.8	48.5	3.2	0.300
RXC J2211.7-0350	22:11:45.92	-03:49:44.73	11.5	49.4	5.5	0.397
ALCS: CLASH						
Abell 383	02:48:03.36	-03:31:44.70	16.8	62.0	0.8	0.187
MACS1206.2-0847	12:06:12.28	-08:48:02.40	13.1	55.0	2.0	0.440
MACS1423.8+2404	14:23:47.76	24:04:40.50	15.2	59.9	1.2	0.545
MACS1931.8-2635	19:31:49.66	-26:34:34.00	13.1	49.1	1.8	0.352
RXJ 1347-1145	13:47:30.59	-11:45:10.10	12.6	46.8	2.5	0.451
MACS1311.0-0310	13:11:01.67	-03:10:39.50	13.2	52.8	0.9	0.494
MACS1115.9+0129	11:15:52.05	01:29:56.60	13.2	53.0	1.0	0.352
MACS0429.6-0253	04:29:36.10	-02:53:08.00	13.2	54.2	0.7	0.399
RXJ2129.7+0005	21:29:39.94	00:05:18.80	18.4	75.5	0.5	0.234
MACS0329.7-0211	03:29:41.68	-02:11:47.70	12.8	48.7	2.1	0.450
MACS2129.4-0741	21:29:26.21	-07:41:26.22	14.7	55.8	1.7	0.570
Abell 209	01:31:52.57	-13:36:38.80	15.1	54.1	0.7	0.206

† The "Science area" corresponds to the coverage of the detection band, as defined in Section 3.4

might result in erroneous counterpart assignment. Instead I manually examine cutouts for all 180 objects, and assign the $HST/Spitzer$ counterpart by hand, if such is available, to the ALMA detected object. In total, I find 145/180 matches during this procedure. These are flagged with $\texttt{alma\_coverage} = 2$ in the catalogue. From the 35 objects that do not have a match to my catalogue: 14 are "*HST*-dark" with clear detections in both IRAC channels; 8 are not visible in either HST or IRAC; 10 fall outside of the HST area, but have clear IRAC detections; 3 are too close to a star, or some other image artifact. The HST and IRAC "dark" objects are cataloged separately, and will be discussed in a later section.

Apart from the aforementioned 145 sources, the remaining objects in my catalogue do not have a reliably measured ALMA counterpart, however the ALMA continuum map at 1.2 mm still overlaps significantly with my mosaics, and as such it is still possible to extract the ALMA upper limits at the positions of my sources. For this I employ `photutils` (Bradley et al. 2020), a software package designed for quick and reliable aperture photometry. I begin the extraction of upper limits by isolating a sub-sample of sources in my catalogue that do not have an ALMA counterpart in the continuum catalogue, but still fall within an ALMA map. The objects that fall outside of ALMA coverage have been flagged with `alma_coverage` $= 0$. Within each field, I have used an approach similar to the original flux extraction of Fujimoto et al. (2022; in prep) and measured the flux from the central pixel on the ALMA map. As the uncertainty on the measured flux I adopted the noise level of the entire map (as described in Fujimoto et al. 2022, in prep), calibrated to the area of the beam. In total, this results in 30,586/217,958 ALMA upper limit measurements. While these are only 3σ upper limits, the addition of an extra constraint in the FIR will significantly aid and enhance the quality of panchromatic SED fitting. For example, these measurements can help to compute the upper limits on the dust mass ($M_{\rm dust}$), for objects up to $z \sim 7$, and the infrared luminosity $L_{\rm IR}$ above $z \sim 9$. This is due to the fact that the 1.2 mm stops sampling the rest-frame continuum above 150 μm at $z \sim 7$, which is required for robust calculations of the dust mass (e.g. see discussion in Berta et al. 2016 and Kokorev et al. 2021). In return, however, the 80-120 μm regime will then become available for 1.2 mm photometry at $z = 9$, which would allow to sample the peak of the FIR emission and allow to impose tighter constraints on both total infrared luminosity $L_{\rm IR}$ and the dust temperature - $T_{\rm dust}$. This naturally depends on how dusty, high-z galaxies actually are, and whether the commonly used optically thin dust approximation will apply to them.

3.5.2 Catalogue Description and Flags

I provide a full multi-wavelength photometric catalogue for each of the 33 clusters covered by ALCS, including the parallels for HFF. Combined, the catalogues contain aperture and total flux measurements for 218,000 sources. I list the description of the relevant columns of the catalogue in Table 3.2. All fluxes and uncertainties are in units of μJy, unless specified otherwise. For the *HST* data I include the aperture photometry, measured within $D = 0.''7$ arcsec, and the associated total fluxes. The IRAC measurements are provided as the `golfir` modeled IRAC photometry, as well as aperture (and aperture corrected) fluxes, measured within $D = 3.''0$. The cross-matched ALMA photometry is provided in units of μJy/beam. I also include an ALMA source flag column - `alma_coverage`, where 2 - indicates a detection, 1 - an upper limit and 0 - a lack of coverage. To help discern between the high and low quality photometry for sources in the catalog, I have added a number of flags which allow to uniformly select reliable samples of objects. For each source I compute how many pixels in a 3x3 square around the center have been masked or fall

Table 3.2: Description of the Relevant Photometric Catalogue Columns

Column Name	Units	Description
id	-	Object ID
ra	deg	Right Ascension
dec	deg	Declination
{filt}_flux_aper	uJy	*HST* D=0.″7 aperture flux
{filt}_err_aper	uJy	Uncertainty
{filt}_flux	uJy	Total *HST* flux density
{filt}_err	uJy	Uncertainty
irac_{ch}_flux	uJy	IRAC model flux
irac_{ch}_err	uJy	Uncertainty
irac_{ch}_flux_aper	uJy	IRAC D=3.0" aperture flux
irac_{ch}_err_aper	uJy	Uncertainty
alma_coverage	-	ALMA coverage flag
alma_flux	μJy/beam	ALMA flux at 1.2 mm
alma_err	μJy/beam	Uncertainty

outside the IRAC mosaic. This number is given in the `n_masked` column. If the central pixel itself is masked or missing, I flag that source with `bad_phot` = 1.

3.6 Quality and Consistency Verification

There exist a number of publicly available multi-wavelength photometric catalogues covering the ALCS fields. These include the catalogues in the HFF (Merlin et al. 2016; Di Criscienzo et al. 2017; Shipley et al. 2018), CLASH (Postman et al. 2012) and RELICS (Coe et al. 2019). Since my work focuses on both *HST* and IRAC photometry simultaneously, I would like to compare my results to the publicly available catalogues that have similar broad band *HST* coverage, and also process and include the IRAC bands. For this reason, I focus my comparison on HFF-DeepSpace and ASTRODEEP catalogues, by Shipley et al. (2018) and Merlin et al. (2016); Di Criscienzo et al. (2017) respectively. Despite this comparison being constrained only to the HFF, compared to the RELICS and CLASH, the Frontier fields generally have wider photometric coverage, and reach a greater depth, allowing us to carry out a uniform comparison for a large number of objects. The comparison will be shown on a per-filter basis, and also will include the difference between the derived colours.

3.6.1 HFF-DeepSpace

The HFF-DeepSpace multi-wavelength photometric catalogue is presented in Shipley et al. (2018). The catalogue focuses on the six Frontier fields, plus the associated parallels. My overlap with HFF-DeepSpace includes all cluster and parallel regions apart from MACS J0717.5+3745, which I will exclude from the comparison. The UV-optical photometric catalogue consists of up to 17 filters with *HST*/ACS and *HST*/WFC3, VLT/HAWK-I K_s filter, and the (post-cryogenic) IRAC 3.6

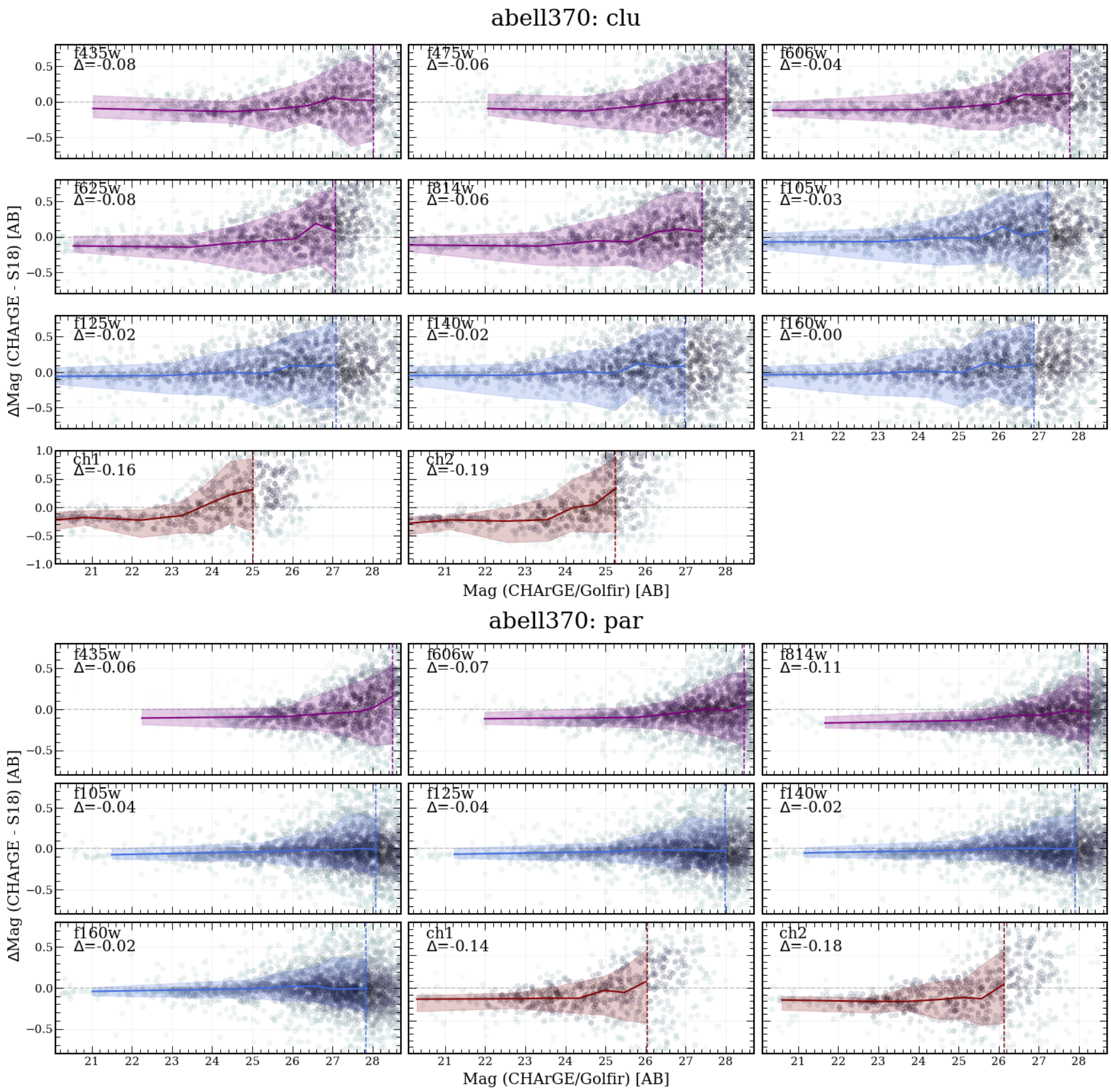

Figure 3.3: Comparison of *HST* and IRAC photometry between my catalogue and Shipley et al. 2018 for the cluster (top) and parallel (bottom) parts of Abell 370 field. I only compare the objects in the central part of the field. For more clarity the colours of shaded regions correspond to HST/ACS (purple), HST/WFC3 (blue), and $Spitzer$/IRAC (maroon). The difference in magnitude ΔMag is shown by scattered circles, which are coloured based on the density of sources around them. The overlaid solid lines correspond to the binned median, that are selected to contain the same amount of objects in each bin. The shaded envelope captures the 68% of points per magnitude bin. The vertical dashed lines correspond to the 1σ depth limit for each band. The median ΔMag for galaxies brighter than the depth limit is shown on each panel.

μm and 4.5 μm data, plus the archival IRAC 5.8 μm and 8 μm measurements, where available.

The HFF-DeepSpace methodology starts with recombination of all the background subtracted HST exposures from different epochs, and performing an initial image cleaning from artifacts and cosmic rays. After that, for each cluster field, they follow the Ferrarese et al. (2006) method to model

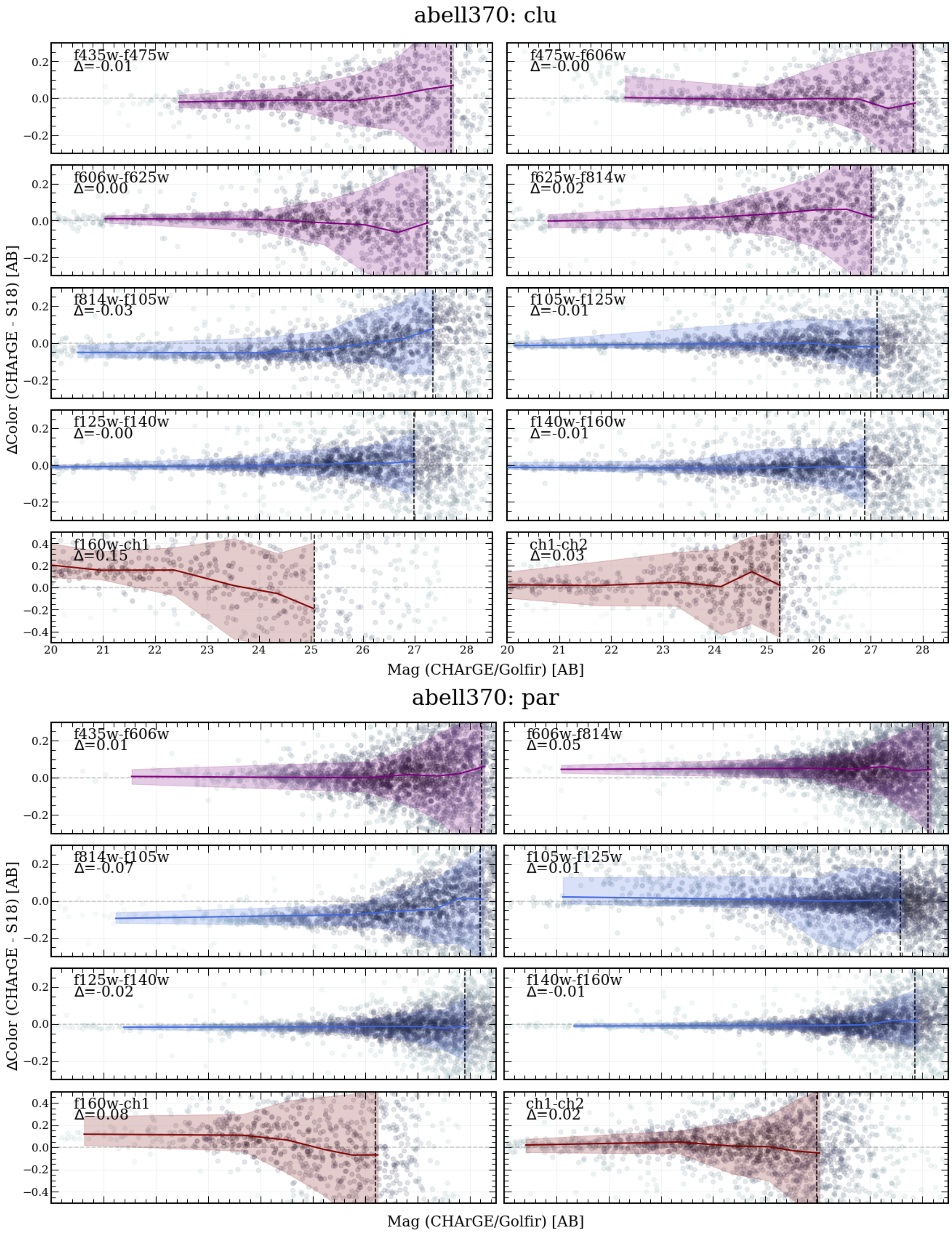

Figure 3.4: Comparison of *HST* and IRAC colours between my catalogue and Shipley et al. 2018, for the cluster (top) and parallel (bottom) parts of Abell 370 field. The colours and symbols are the same as in Figure 3.3.

and subtract out bCGs and the ICL from each field, performing additional background cleaning on the bCG - subtracted mosaics. This procedure allows to identify more objects magnified by the gravitational potential of the cluster members, and to assign the correct flux values to galaxies that are located close to the cluster on the sky.

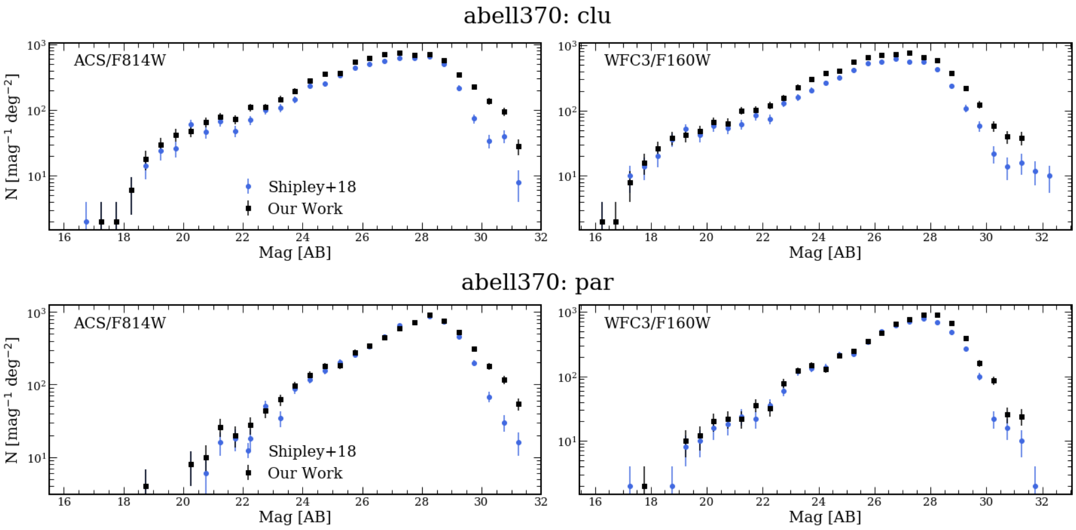

Figure 3.5: The area normalised number counts in F814W and F160W filters for the central part of the Abell 370 cluster and parallel fields. I show the number counts for my catalogue as black squares, while the Shipley et al. (2018) work is overplotted in blue circles.

After the bCGs have been subtracted, all shorter wavelength mosaics are then PSF matched to the F160W band. The mosaics for all available *HST* filters are then combined into a single weighted mean detection image. The source detection itself, is then performed on the resultant image by using `SExtractor`, generally following the methodology described in Skelton et al. (2014). The *HST* aperture photometry extraction within a diameter of 0.″7 is then performed on the detection and individual PSF matched *HST* images, and the correction to the total flux is calculated from the curves of growth.

Low resolution IRAC photometry is extracted by using MOPHONGO (Labbé et al. 2013, 2015), a code developed to process longer wavelength bands, specifically focusing on potentially blended objects. Similarly to `golfir` the high resolution detection image is convolved with low-resolution kernel, and used as a model to fit the IRAC photometry. However, these models are not used to extract the flux, but rather to correct for the possible contamination from the neighbouring sources, with the IRAC fluxes themselves being extracted from 3.″0 apertures. Additional flux corrections are performed by using the PSF curves of growth.

Apart from the aperture corrections, the measured fluxes of Shipley et al. (2018) have undergone a number of additional modifications with the aim of providing the best possible results during SED fitting. These include the zeropoint and the Milky Way extinction corrections. The values presented in my photometric catalogue do not incorporate such corrections, and therefore I had to de-apply them, as specified in the Shipley et al. catalogue documentation. Our knowledge of galaxy SEDs is still limited, especially as we move into the high-z Universe, thus preventing us from accurately calculating the zeropoint corrections (e.g. see Brown et al. 2014), with the template error function largely alleviating these effects, however not in their entirety (Brammer et al. 2008). Moreover, calculating extinction corrections would largely depend on the the adopted models, introducing unwanted shifts if not done carefully. Due to this, although both the zeropoint and MW corrections are largely inconsequential ($\mathcal{O} \sim 2\%$), I take great care to convert total fluxes back to their original values to carry out a robust and unbiased photometry comparison without these corrections. The fluxes presented in the Shipley et al. catalogue have been additionally normalised to AB=25, rather than AB=23.9, which I also take into account when comparing the results.

The total combined science area of the CHArGE/`golfir` *HST* and IRAC data in my catalogue, after removing masked regions, is equal to 312.6 arcmin^2, roughly double the size of the HFF-DeepSpace coverage of 136.7 arcmin^2. This results in a higher number of objects recovered in my work, which after removing potential spurious and masked sources totals to 125,947 across all five Frontier Fields, which is $\times 2.3$ larger than the `use_phot` HFF-DeepSpace sample which numbers 55,579 objects. To perform the photometry comparison between the two catalogues, I only focus on the regions where the two overlap. However, even for the parts where my data cover the same area, the CHArGE images include new exposures, not originally present in the Shipley et al. analysis, particularly on the edges of the old mosaic. Therefore I feel that for a fair photometry comparison blind matching of all available sources would be inappropriate, with the variation in depth potentially introducing some deleterious effects on the comparison. Instead, I focus on the central parts of both the cluster (clu) and parallel (par) regions of all five Frontier Fields, where the coverage is deepest and uniform for both catalogues. In addition, as mentioned in Section 3.4, I do not model or subtract the bCG when computing the photometry, therefore focusing only on the central area will allow us to contrast how bCG modelling affects the final measured fluxes. I have used a 1.''0 matching radius, and find that the median astrometric offset between the two catalogues in equal to ~ 0.12 arcsec, which I have corrected for.

In Figure 3.3 I show a comparison between ALCS CHArGe/`golfir` and HFF-DeepSpace catalogues for the Abell 370 field. I carry out the comparison for individual filters, and colours separately for the cluster and parallel parts, depending on the availability of the photometry. For *HST* photometry the agreement is largely good to excellent. The median offsets vary with instrument, from -0.06 mag in *HST*/ACS, to 0.02 mag in *HST*/WFC3 filters. My fluxes are consistently brighter, however this does not seem to stem from the lack of bCG subtraction, but rather PSF matching performed on all bands apart from F160W, where the median difference is zero. This notion is further reinforced by the fact that the flux disparity is larger for ACS filters, where the effects of PSF matching would be most noticeable. For the longer wavelength IRAC bands, I however note a much larger offset, of -0.17 mag, both for the cluster and the parallels. In an attempt to understand the discrepancy, I have first compared the science images themselves, by randomly placing 1,000 apertures in the mosaics. The extracted fluxes were different by at most ~ 0.05 mag, which does not explain the difference that I see in my comparison. I believe that the discrepancy in the IRAC bands can be attributed to the difference in chosen methodology, specifically the aperture to total flux correction. Comparison between broad band magnitudes for the remaining fields can be found in Appendix B.

Comparison of colours yields even better results for both *HST* and IRAC bands as it largely ignores any differences caused by PSF matching the individual bands, and on the aperture to total flux corrections. I show these results in Figure 3.4, again both for cluster and parallel parts of Abell 370. Apart from the F160W $-$ 3.6 μm colour, I find offsets of the order of -0.01 mag. Despite the difference in flux measurements, the agreement between colours indicates that measurements in both catalogues have been executed in a consistent way.

Finally, in Figure 3.5, I show the area weighted number of objects, as a function of their apparent magnitude. The source number counts is one of the basic tests to evaluate and characterise sensitivity-limited samples. The comparison is performed in the same "deep" areas of the Abell 370 field, for two representative ACS and WFC3 filters. I do not carry out the same analysis for the IRAC bands, as the flux offsets are too large to compare the number counts in a consistent manner. For this comparison I have only used the "clean" photometry, i.e. `use_phot=1` flag for HFF-DeepSpace and `n_masked=0` for my data. For each Frontier Field, my detection algorithm recovers roughly 30 % more sources, specifically in the faint end. A small fraction of these are expected to be spurious, or are a result of overly aggressive deblending of nearby galaxies. The number counts between both catalogues are however largely consistent, and show similar depths.

3.6.2 ASTRODEEP

For the next comparison I focus on the ASTRODEEP catalogue presented in Merlin et al. (2016), including the additional data release by Di Criscienzo et al. (2017). The catalogues present the data for the cluster and parallel pointings for four Frontier Fields, Abell-2744, MACSJ0416, MACSJ0717 and MACSJ1149, three of which lie in the ALCS area. The photometric coverage includes 10 bands, covering *HST*/ACS and WFC3, as well as HAWK-I K_s, and IRAC 3.6 μm and 4.5 μm data.

The methodology for the ASTRODEEP catalogue is largely similar to the one presented in Shipley et al. (2018). The ICL and bCGs are modelled and subtracted from the H-band image by using models from Ferrarese et al. (2006), in conjunction with `GALFIT`. The *HST* source detection is performed on a single image - the WFC3 F160W band by using `SExtractor` in both HOT+COLD modes (e.g. see Galametz et al. 2013). The authors then use a sequential approach to subtract the ICL and bCGs from all the other cluster images, by using the output of the previous `GALFIT` run on a redder band, as an initial guess for the bluer one. All of the images are then PSF - matched to the F160W filter, by using a convolution kernel. The final *HST* photometry

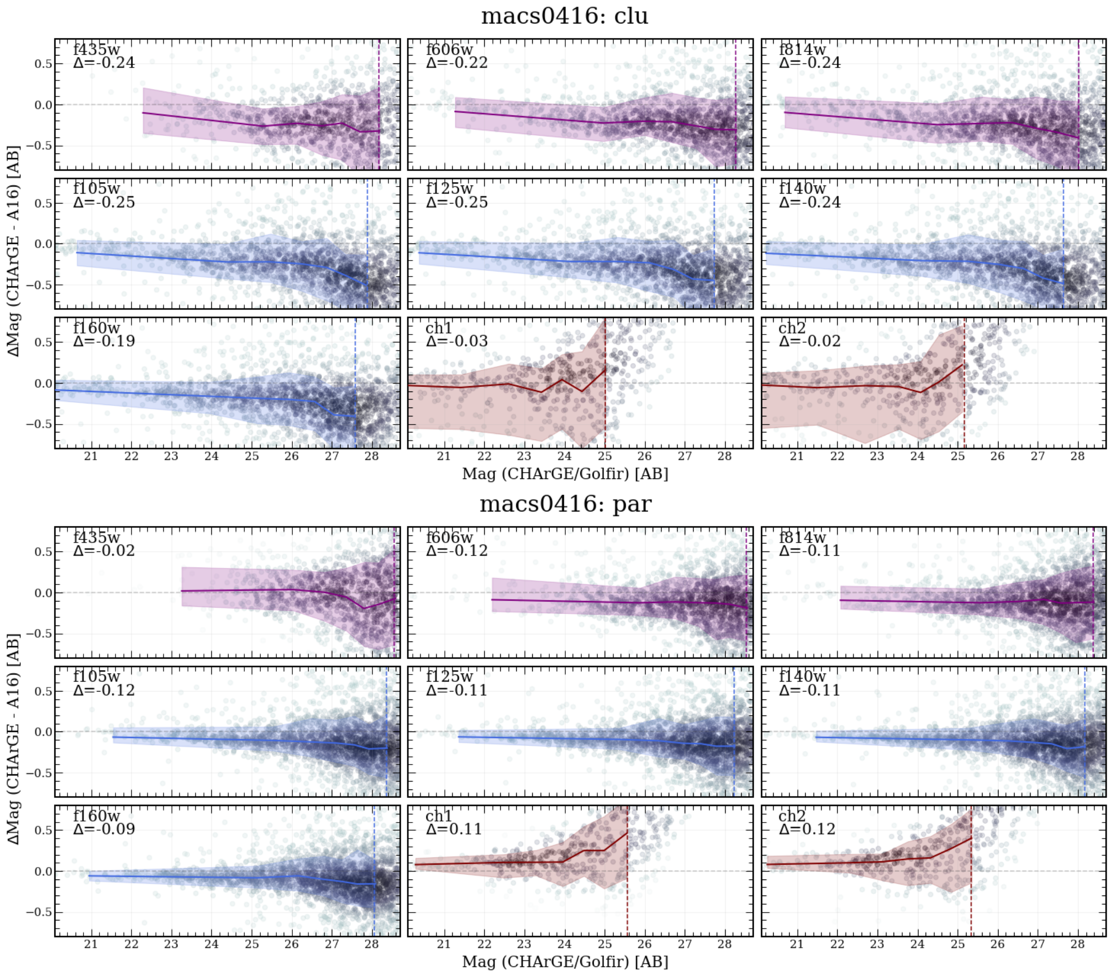

Figure 3.6: Comparison of *HST* and IRAC photometry between my catalogue and the ASTRODEEP data for the cluster (top) and parallel (bottom) parts of MACSJ0416 field. Symbols and colours are the same as in fig:deltamag.

is extracted with `SExtractor` running in dual mode, to measure the aperture and isophotal fluxes. Total fluxes in the F160W filter are computed from the `SExtractor` FLUX_AUTO parameter. For other bands the total flux is derived from the scaling of the detection band to all the relevant isophotal colours.

The low resolution IRAC photometric measurements in the ASTRODEEP catalogue follow a method similar to the one outlined in this work, and Shipley et al. (2018). The authors use `T-PHOT` (Merlin et al. 2015), which follows the same methodology as both MOPHONGO and `golfir`, by using high-resolution images, convolved with a low resolution PSFs to act as models for the photometric data.

As before, I only compare sources in the central parts of both cluster and parallel fields. For that I used a matching radius of 1."0. The median astrometric offset between the my galaxies and

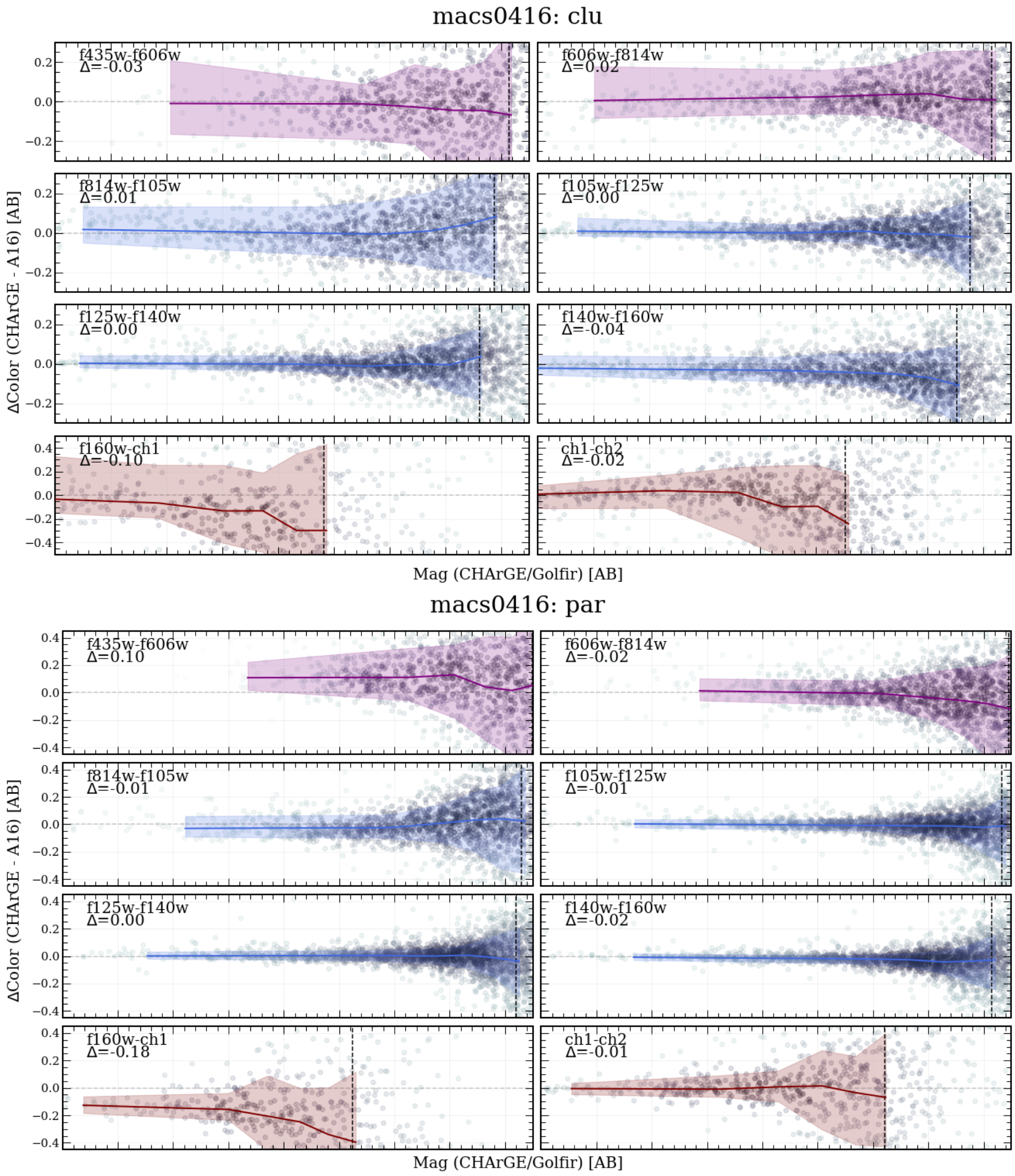

Figure 3.7: Comparison of *HST* and IRAC colours between my catalogue and the ASTRODEEP data for the cluster (top) and parallel (bottom) parts of MACSJ0416 field. The colours and symbols are the same as in Figure 3.3.

the ASTRODEEP catalogue is ~ 0.22 arcsec. In Figure 3.6 I show a comparison between ALCS CHArGe/`golfir` and ASTRODEEP catalogues for the MACS0416 field. For *HST* photometry the difference is quite substantial, with a median offset equal to -0.21 mag for the cluster and -0.11 mag for the parallel field, without dependence on the instrument. I suspect that the bCG subtraction is responsible for half of the offset, as the differences in flux are lessened in the parallel field. The fluxes presented in the ASTRODEEP catalogue are also extinction corrected, however I believe that this is not a major contributor to the difference that I find. If one ignores the potential effects

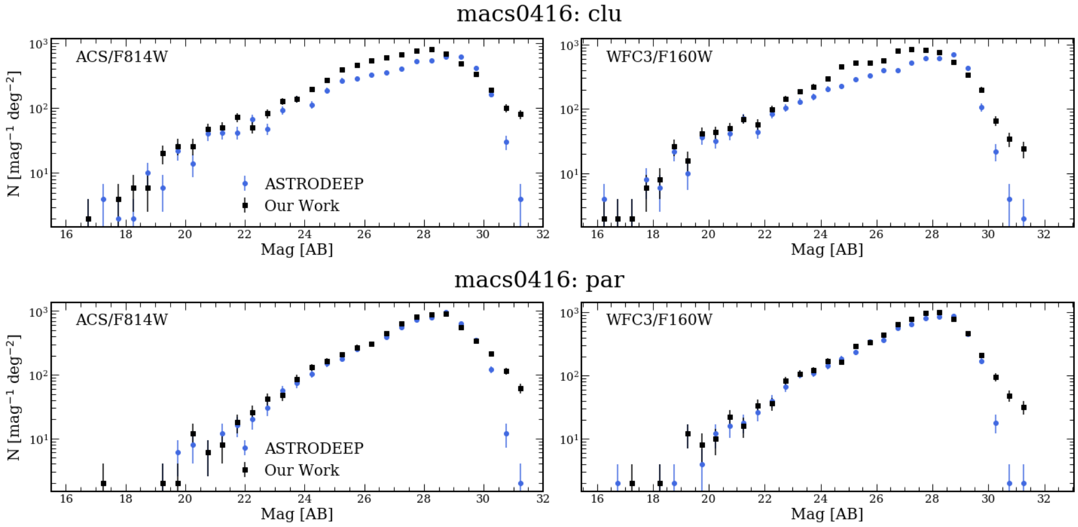

Figure 3.8: The area normalised number counts in F814W and F160W filters for the central part of the MACSJ0416 cluster and parallel fields. I show the number counts for my catalogue as black squares, while the ASTRODEEP data is overplotted in blue circles.

of bCG subtraction, in most cases the difference between the two catalogues is not a systematic shift, but rather an offset increasing at fainter magnitudes. The cause for the flux-dependent behavior of the offset likely originates from a different aperture used to extract the photometry, and the methodology to covert that flux to total. Unfortunately I could not carry out the exact flux comparison to the ASTRODEEP catalog, due to the unavailability of documentation regarding the aperture sizes used, and total flux corrections. In the IRAC 3.6 and 4.5 μm filters, however, I find a remarkable agreement, of ~ -0.01 mag, in both cluster and parallel fields. The colour comparison, again, yields consistent results, across all *HST* and IRAC filters. I find a median offset of -0.01 mag for all colours, apart from the F160W $-$ 3.6 μm colour.

I do not perform a number count comparison to the ASTRODEEP catalog, as the magnitude offset are too large to yield reasonable results. Notably, however, I again recover roughly 30 % more sources, in both cluster and parallel fields on the faint end.

3.7 Galaxy Properties

To derive the photometric redshifts, rest-frame colour, dust attenuation, and stellar population parameters I have used the updated Python version of EAZY (Brammer et al. 2008). EAZY is a photometric spectral energy distribution fitting code which is optimised to fit non-negative linear combinations of basis-set templates, rather than devising a solution from vast libraries of pre-compiled models. These templates are extracted from the Flexible Stellar Populations Synthesis (FSPS; Conroy et al. 2009) models, and then reduced to a set of 12, which are able to reproduce a much larger library, spanning a variety of dust attenuation, ages, mass-to-light ratios and star-formation history (SFH) properties (see Brammer et al. 2008 and Blanton & Roweis 2007). I have listed the properties for these individual templates in Table A.1. For each object in my multi-wavelength catalogue, EAZY integrates redshifted templates through a set of filters corresponding to the observed bands, and then finds the best combination for a given set of fluxes and associated uncertainties.

To improve the quality of the photometric redshift estimate, EAZY also implements a “template error function”, which is used to account any other miscellaneous uncertainties related to short-lived and otherwise unusual stellar populations, and the emission lines coupled to the properties of the ISM. In my analysis I have used a default template error function value of 0.2. To calculate the photometric redshifts and physical parameters with EAZY I have used total *HST* fluxes, calculated from the D=0.″7 aperture measurement, and the `golfir` IRAC models. These data are contained in "{`filt`}_flux" and "irac_{`ch`}_flux" columns, respectively.

3.7.1 Spectroscopic Redshift Catalogues

The ALCS fields have been covered by an wide range of spectroscopic surveys. I have examined the literature and compiled all of the available spectroscopic redshifts in order to assess the quality of my photometric redshifts. During this process I select only the redshifts that have the reliable quality flag, and choose the most recent source, if a galaxy is present in multiple catalogues. I have cross-matched all the available spectroscopic redshifts with objects in my catalogue, using a 0.″5 matching radius. These spectroscopic redshifts are included in my main catalogue, with a separate column providing the source where available.

For Hubble Frontier Fields I have used spectroscopic redshifts already compiled in Shipley et al. 2018, this includes Grism Lens-Amplified Survey from Space (GLASS) (Treu et al. 2015; Schmidt et al. 2014), and the spectral data presented in Smith et al. 2009; Owers et al. 2011; Ebeling et al. 2014; Jauzac et al. 2014; Richard et al. 2014; Balestra et al. 2016; Caminha et al. 2016; Diego et al. 2016; Grillo et al. 2016; Karman et al. 2016; Limousin et al. 2016; Lagattuta et al. 2017; Mahler et al. 2018. Where appropriate, I also have updated the spectroscopic redshifts with the most recent results from Richard et al. 2021. In total I recover 3,333 matches with my catalogue, in all five HFF fields, including the parallels. Please note however, that the number of spectroscopic redshifts in the parallel fields is severely limited. I however include them in my final comparison for posterity.

For CLASH I have compiled the most recent data from GLASS (Treu et al. 2015; Schmidt et al. 2014), the results from Caminha et al. 2019 and Richard et al. 2021. In total there are 1,241 matches with my catalogue, spanning ten out of twelve CLASH fields. Finally, for the RELICS spectroscopic redshifts, I have used the data from (Richard et al. 2021) for RXC0600, for which I have recovered 67 matches with my catalogues. The full list of fields and matches is outlined in Table 3.3.

3.7.2 Photometric Redshift Accuracy

To quantify the precision of my photo-z estimates I used the normalised median absolute deviation (NMAD, Hoaglin et al. 1983), defined as:

$$\sigma_{\mathrm{NMAD}} = 1.48 \times \mathrm{median}\left(\frac{|\Delta z - \mathrm{median}(\Delta z)|}{1 + z_{\mathrm{spec}}}\right). \tag{3.1}$$

This method in commonly used in the literature (e.g. see Shipley et al. 2018; Skelton et al. 2014), allowing for a quick and unbiased comparison of redshift quality between different catalogues, and is also less sensitive to outliers as described in Brammer et al. 2008. The outlier fraction η is given by $|\Delta z|/(1+z_{\mathrm{spec}}) > 0.15$, following the methodology described in Hildebrandt et al. 2012. In total I have carried out the comparison for 3,638 matched objects, spanning 16 fields, shown in Figure 3.9. I find that my redshift accuracy is generally good, with a σ_{NMAD} of 0.058, and 20.8 % of catastrophic failures.

For the $z_{\mathrm{phot}} - z_{\mathrm{spec}}$ comparison, I note the existence of over-densities located either at $z \sim 1$ or $z \sim 4$, for which $\Delta z \sim 3$. This is where a vast majority of my catastrophic outliers are located. These redshift discrepancies are caused by the mis-identification between the Lyman (912 Å) and the Balmer (3640 Å) breaks in the fitted SEDs. The manifestation of the break confusion is a consequence of degenerate behaviour of templates, when faced with either sparse or very faint photometry, and is particularly prominent when I compare large samples of photometric redshifts between different catalogues.

3.7.3 Gravitational Lensing Magnification

For objects within the cluster fields I compute and provide the lensing magnification factor (μ), which is based on the RA and Dec coordinates of the source in the detection band (i.e. the peak flux coordinate) and its redshift. Although a vast majority of sources in a given field only have a z_{phot} estimate, I use a z_{spec} where possible. Following the methodology described in Sun et al. (2022; submitted), I use the Zitrin-NFW lens models (Zitrin et al. 2013, 2015) for the HFF and CLASH clusters, and GLAFIC models (Oguri 2010; Okabe et al. 2020) for RELICS. These models consist of the mass surface density (κ) and weak lensing shear (γ) maps. I then compute the magnification by using:

$$\mu = \frac{1}{(1 - \kappa\,\beta)^2 - (\gamma\,\beta)^2}, \tag{3.2}$$

where β is the lensing depth, defined as $\beta = D_{\mathrm{ls}}/D_{\mathrm{s}}$, with D_{ls} being the angular diameter distance

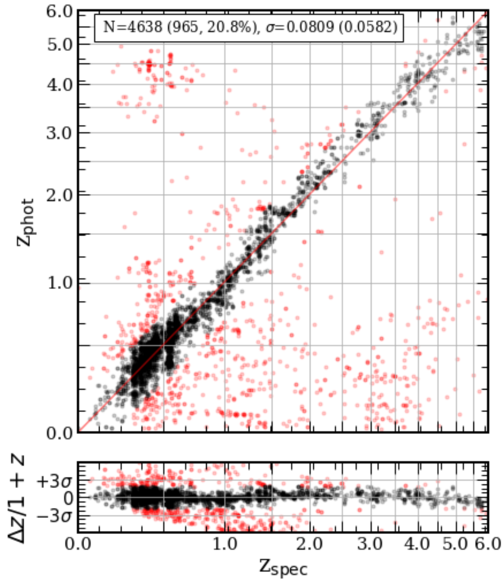

Figure 3.9: The comparison between photometric redshifts derived by EAZY and the spectroscopic redshift from the literature. I have carried out the comparison in all fields where the match was found, in total these include all five Frontier Fields, ten CLASH fields, and one RELICS field. Black and red circles denote galaxies below and above catastrophic limit of 0.15, respectively.

between the lens and the source, and D_{s} is the angular diameter distance to the source. Similarly to Rawle et al. 2016, if the source redshift z_{s} is below or within the cluster redshift z_{cl}, namely $z_{\mathrm{s}} \leq z_{\mathrm{cl}} + 0.1$, I set the magnification to unity. Sources that fall outside of the magnification maps for a given field, are not expected to be significantly affected by gravitational lensing. For these, I have set $\mu = 1$. Please note that I do not apply these lensing corrections to any fluxes in my catalogue.

3.7.4 Rest-Frame Colour Galaxy Classification

Comparison of galaxies at different redshifts often requires rest-frame, rather than observed frame, colours to be used. For each galaxy these are determined by assuming the best-fit EAZY template and its redshift, in order to calculate the rest-frame flux for a set of filters. The rest-frame colours are then computed by integrating a transmission curve for a given filter through the best fit template, as described in Brammer et al. 2011. In my final catalogues I provide the rest-frame fluxes for the most commonly used filters ($GALEX$ NUV, COSMOS r, Johnson U, B, V, and J).

The rest-frame fluxes can be used to assess the galaxy populations in each field by using the colour-colour analysis. Multiple previous studies have devised a variety of techniques to classify galaxies based on the broad-band photometry. One such prescription utilises the $U-V$ and $V-J$

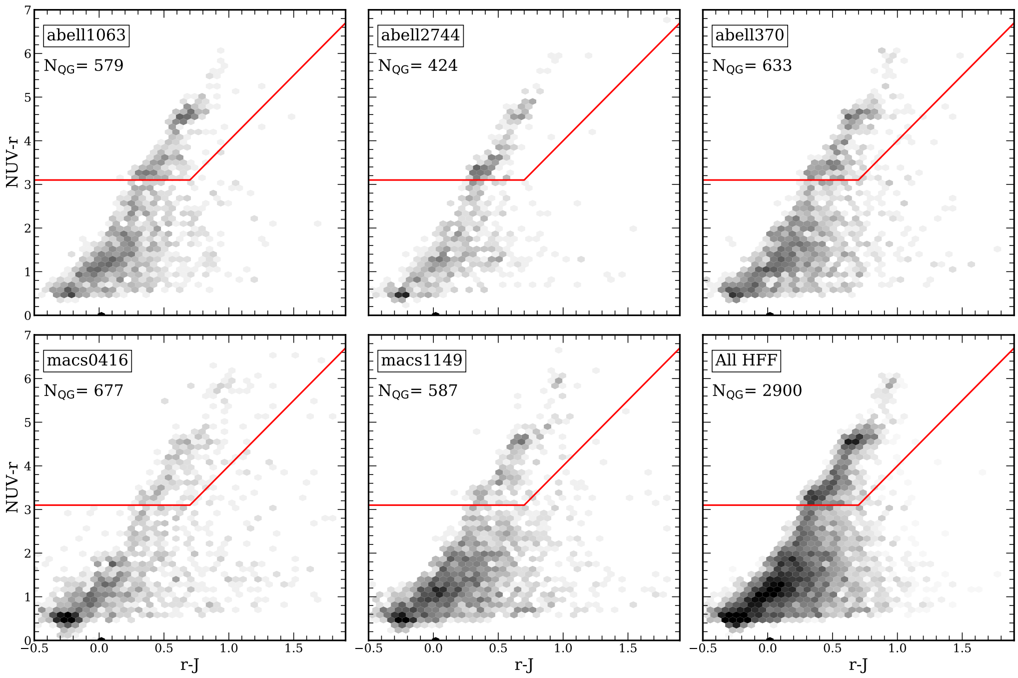

Figure 3.10: Classification of galaxies for the clusters in the Hubble Frontier Fields by using the $NUV - r$ and $r - J$ rest frame colours. The $NUV - r - J$ galaxy selection prescription from Ilbert et al. (2013) is shown as a solid red line. The number of quiescent galaxies is displayed within each plot. Here I have limited the selection to only include galaxies with `nusefilt`≥ 5.

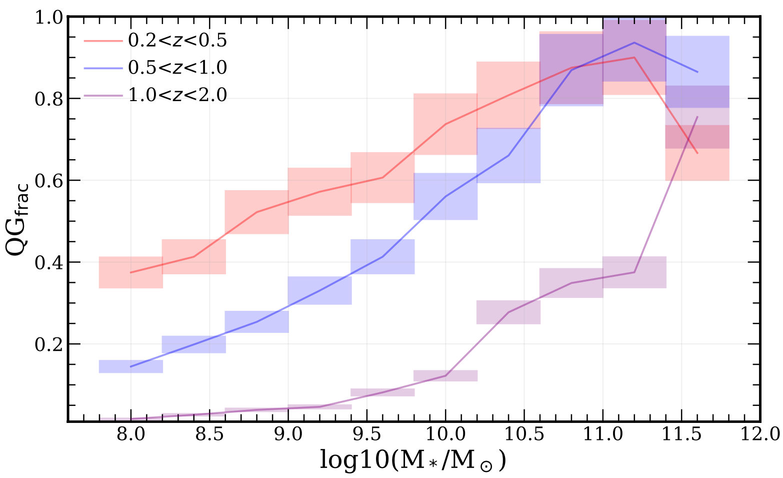

Figure 3.11: Fraction of quiescent galaxies in the Hubble Frontier Fields as predicted by $NUVrJ$ selection. I have separated my sample into stellar mass and redshift bins. The extent of the coloured rectangles corresponds to the uncertainty.

rest-frame colours (Labbé et al. 2005; Wuyts et al. 2007; Williams et al. 2009) to separate galaxies into quiescent, star-forming. Quiescent galaxies with low levels of star formation are red in the

$U - V$ regime and are easily distinguished from the similarly red (in $U - V$), dusty, star-forming galaxies, with the $V - J$ colour. An alternative method, using the $NUV - r - J$ colours instead, has been proposed by Ilbert et al. (2013) and Arnouts et al. (2013). While the UVJ selection is the most commonly used approach in the literature, the $NUVrJ$ method has some key advantages. The shorter wavelength NUV band is more sensitive both to the dust attenuation and emission from young stellar populations than the U band. Although the amount of quiescent galaxies at $z > 2$ is limited (see e.g. Ilbert et al. 2013; Muzzin et al. 2013; Davidzon et al. 2017), the rest-frame U band would already be shifted into near-IR bands, this would make the $NUVrJ$ selection technique more reliable at higher redshifts. In my work I adopt the $NUVrJ$ selection, however the U and V rest frame fluxes are also provided in the final catalogue. I display my selection in Figure 3.10, where the quiescent galaxies tend to be located in the upper-left corner of the diagram, with the boundaries defined in Ilbert et al. (2013):

$$NUV - r = \begin{cases} 3\,(r - J) + 1 & \text{for} \quad r - J > 0.7 \\ 3.1 & \text{for} \quad r - J < 0.7. \end{cases}$$

When comparing the UVJ and $NUVrJ$ colour classification methods, $\sim 70\%$ of objects selected with $NUVrJ$ overlap with UVJ quiescent candidates, varying slightly with the depth of the field of choice, and therefore its redshift distribution. Using $NUVrJ$ we find a relatively high fraction of quiescent galaxies as shown in Figure 3.11. The ALCS covers cluster fields, where the star formation is generally expected to be suppressed, especially at low redshift (see e.g. Boselli et al. 2016 and references therein). Therefore it is not at all surprising for us to recover high fractions of quiescent galaxies.

3.8 Conclusions

This paper describes the creation of *HST*+IRAC photometric and galaxy property catalogues within the 33 lensed cluster fields, covered by the 110 arcmin^2 ALCS survey. The mosaics and catalogues cover a combined area of $\sim$ 690 arcmin^2, in 33 ALCS fields, which include 5 Hubble Frontier, 16 RELICS, and 12 CLASH fields. The final catalogues number roughly 218,000 sources, which are covered by at most 12 *HST*/ACS,UVIS and WFC3 bands, plus the additional IRAC photometry at 3.6 and 4.5 μm. To process these data I have reprocessed and recombined all the available archival *HST* exposures, now combined into a single CHArGE dataset, as well as all available IRAC data covering the same fields. Each image has been aligned to the same highly precise Gaia DR2 reference frame, ensuring a robust internal alignment of the *HST* and IRAC images for matched-aperture photometry, with the final absolute astrometric precision generally being < 100 mas.

Table 3.3: Sources of Spectroscopic Redshift.

Field	Matches (# of Galaxies) height
HFF: cluster & parallel	
Abell 2744	525
Abell 370	513
MACSJ0416	1064
MACSJ1149	973
Abell S1063	395
CLASH	
MACS0329	101
MACS0429	3
MACS1115	60
MACS1206	360
MACS1311	65
RXJ_1347	135
MACS1423	149
MACS1931	126
MACS2129	303
RXJ2129	127
RELICS	
RXCJ0600	67

Table 3.4: Templates used for the EAZY fit.

Parameter	Value
Optical emission: Brammer et al. 2008[a]	
A_V (Calzetti et al. 2000)	[0.6,0.12,0.19,0.29,1.05,2.68, 0.11,0.36,0.98,1.54,1.97,2.96]
M/L_V	[0.38,0.76,1.68,4.01,6.45,44.48, 0.12,0.21,0.33,0.64,1.57,4.00]
$\log_{10}$(sSFR)	[-10.75,-11.37,-11.90,-12.53,-12.05, -12.47,-8.37,-8.60,-8.50,-8.57,-8.93,-8.90]

Please refer to Brammer et al. 2008 for a more detailed description of the creation and selection of these basis set templates. See Blanton & Roweis 2007 for a methodology regarding the SFH.

In my analysis, I have applied a consistent methodology in order to compute multi-wavelength photometry across all 33 fields. The `sep` software (Barbary 2016) is used to detect sources on a weighted master detection image from all available ACS/WFC and WFC3/IR filters of a given field. I do not PSF match the *HST* images and extract the photometry for each filter separately, with apertures of varying sizes, and then correct it to total fluxes by using curves of growth. I use a novel `golfir` algorithm, which relies on using the IRAC PSF - convolved high-resolution mosaics as a prior, to model and extract the IRAC photometry. Furthermore, I test the robustness of the derived photometry by comparing it to the publicly available HFF-DeepSpace (Shipley et al. 2018) and ASTRODEEP (Merlin et al. 2016; Di Criscienzo et al. 2017) catalogues in the Frontier Fields, and find consistent results, despite using a different approach.

To derive photometric redshifts and stellar population parameters, I use an SED fitting software EAZY (Brammer et al. 2008). For the fields where uniform spectroscopic data is available (16/33), I achieve an average NMAD of 0.059, for the photometric and spectroscopic redshifts across all fields, with an outlier fraction of 21 %. To accompany the photometric redshifts, I provide stellar masses, star formation rates, extinctions, and other stellar population parameters based on the observed photometry. I also compute and use the rest-frame NUVrJ colours to separate the detected galaxies into potentially star-forming and quiescent. I find that my lensed cluster sample data contains an increased fraction of QGs, compared to that of blind field observations. In addition to that, I use all the available magnification maps in the ALCS fields, and provide with my catalogue magnification factors, where possible. I manually cross-match my data with the ~ 140 SN> 4 ALMA photometric catalogues, and provide upper limit measurement for a further $\sim$ 30,000 sources.

These mosaics and catalogues, produced by the ALCS team, conclude one of the initial phases of the entire project, which, as outlined in Section 3.3, will focus on multiple aspects and properties of faint sub-mm sources. These catalogues will also facilitate the detection, and further examination of the optically dark galaxy populations (e.g. Wang et al. 2019). A few such, "*HST*-dark" galaxy candidates have already been found, and will be presented and discussed in Chapter 6. Moreover, after a careful examination of the EAZY derived redshifts, and SEDs I have identified and isolated a number of $z > 9$ LBGs candidates (see Chapter 6). As such these photometric catalogues can act as an important tool in designing future observations (e.g. with Keck/MOSFIRE, JWST, GMT) in an attempt to elucidate the key questions about the early onset of star formation, reionisation and assembly of first galaxies.

Both the *HST*/*Spitzer* mosaics, and photometric catalogues described in this work are publicly available in FITS format, through the following repository [1]. Alongside the photometric catalogues for each ALCS field, I include mosaics for all filters, detection images, segmentation maps, bright star masks, IRAC models and residuals. I also provide photometric redshifts and stellar population properties, as measured by EAZY, for each field. In the repository I also provide all the technical documentation regarding the source detection and modeling parameters, as well as notebooks to re-produce best fit EAZY SEDs.

[1] https://github.com/dawn-cph/alcs-clusters

CHAPTER 4

Peering into the Unknown with *JWST*

This chapter contains the early condensed draft of the following article:
"Predictions for the JWST/MIRI 7.7 μm Observations for the COSMOS - Webb survey"*.*

To be submitted to The Astrophysical Journal

Authors: *Vasily Kokorev, Caitlin M. Casey, Georgios E. Magdis and others*

4.1 Abstract

The recent launch of James Webb Space Telescope (*JWST*), on December 25th 2021, has begun a new era of extragalactic astronomy. The unprecedented depth and wavelength coverage of its instruments is expected to create a long-lasting legacy, and enable us to uncover the mysteries behind the formation of first stars and galaxies. The question however remains, how many and what kinds of sources will *JWST* be able to detect? In this work, I present a mock catalogue, and a set of simulated images, which describe the number counts, magnitudes and the physical parameters of galaxies that would potentially be observed by the Mid-Infrared Instrument (MIRI) - F770W filter, during the COSMOS - Webb Cycle I GO program (PIs: J. Kartaltepe; C. Casey). The mock catalogues are built upon the existing wealth of multi-wavelength photometric data in the COSMOS field, utilising two different spectral energy distribution (SED) fitting approaches - EAZY and `Stardust` to simulate the expected flux density in the MIRI filter. The analysis indicates that, from the already detected galaxies in the COSMOS field, MIRI observations are expected to recover 30,000 - 50,000 sources at a 5σ level corresponding to the 24.7 mag AB, depending on the chosen simulation. From the predicted photometry, I model the full area included in the COSMOS - Webb survey, and create a set of mock images. These can be used to optimise critical source detection algorithms, and aid in the interpretation and reduction of the data.

4.2 Introduction

The last few decades of deep extragalactic surveys with the Hubble Space Telescope (*HST*), and Spitzer Space Telescope (*Spitzer*) have given rise to a paradigm shift in our understanding of galaxy formation and evolution. These observations have traced the full evolutionary picture of galaxies, from the local Universe, to the highest redshift objects currently known at $z \sim 11$ (Coe et al. 2013; Oesch et al. 2016). At the same time, ground based observatories, more specifically the W.M. Keck Observatory (*Keck*) and its *MOSFIRE* NIR multi-objects spectrograph, have mostly focused on the physical properties of galaxies around $z = 3$, at the peak of star-formation activity, with sparse detections of of high-z galaxies (Steidel et al. 2014; Kriek et al. 2015). At the moment, further advancement towards the high redshift frontier is mostly confined to the wavelength coverage of *HST* and bright IR line observations by *ALMA*, with *Spitzer* suffering from the low resolution and limited sensitivity, and the ground based observatories being hindered by the atmospheric effects.

The recently launched James Webb Space Telescope (*JWST*; Gardner et al. 2006) is currently undergoing initial calibrations, and is expected to be fully operational by the end of 2022. The telescope is expected break the redshifts records, and detect and study galaxies that are below the luminosity and mass limits currently observable by the existing facilities. In addition to its imaging capabilities, the instruments onboard *JWST* will enable us to obtain spectra for the faintest galaxies detected to date (e.g. Oesch et al. 2016). Despite that, the capabilities of *JWST* are limited. Most importantly, and unlike *HST*, the *JWST* is relatively inefficient for wide mosaic surveys, due to its limited field of view. To this end the state of the art and complex instruments onboard *JWST*, have to be leveraged in a way that optimises the time on the sky to the highest capacity. This, would therefore necessitate development of extremely efficient observational strategies, and the analysis tools to process the retrieved data. The telescope itself, has been built with that in mind. For example the multi-object spectroscopy with the Micro-Shutter Array (MSA) which is a part of the NIRSpec instrument, already allows modern astronomy to enter an epoch where a large amount of targets identified with *JWST* photometry can be followed up rapidly and simultaneously. Efficient organisation of photometric surveys is still however largely limited by survey design, and simulations of galaxies in targeted regions.

It is clear, that to handle these challenges, physically motivated simulated catalogues of *JWST* photometry are required, to efficiently design observations and data reduction. These simulations and mock catalogues, have the potential to test critical data processing tools, and aid with the future interpretation of the scientific data, specifically with regard to the "hidden" or missing populations, such as the optically - dark sources, and the faint sub-mm galaxies (e.g. Fujimoto et al. 2016; Wang et al. 2016; Gruppioni et al. 2020). A number of efforts have already concentrated on such mock data, in order to aid future observations (Williams et al. 2018). These rely purely on the existing scaling relations, stellar mass and luminosity functions, which might omit the complexity of galaxy populations which can be observed with *JWST*, specifically in regard to the high-z Universe, where the existing relations lose their predictive power.

On the other hand, *JWST* is also going to be used to target well studied fields, where large numbers of sources have already been detected. One such field is a part of the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007a). COSMOS itself began in 2003, as a 1.7 deg^2 mosaic observed by *HST*/ACS reaching a depth of 27.2 magnitudes in the i - band (Koekemoer et al. 2007; Scoville et al. 2007b). The COSMOS area was one of the most significant allocations of the *HST* time, and is to this date the largest area mapped by *Hubble*. The field, or parts of it have since been covered by deep observations by the majority of the other astronomical facilities, both in space and on the ground. In the optical to MIR regime this includes *Spitzer* (Sanders et al. 2007) as well as an array of ground based broad and narrow band observations with the Subaru Suprime and most recently Hyper-Suprime Cam (Aihara et al. 2018, 2019, 2022). In the FIR, the field coverage also includes *Herschel*, SCUBA2 (Lutz et al. 2011; Oliver et al. 2012; Cowie et al. 2017; Geach et al. 2017) and some extensive ALMA pointings (e.g. MORA, ex-MORA; Casey et al. 2021), plus the archival ALMA reprocessing projects like A3COSMOS (Liu et al. 2019a). The richness and depth of COSMOS observations has resulted in the field to take up a unique position in the landscape of modern astronomy, where the wealth of COSMOS photometric catalogues (Laigle et al. 2016; Jin et al. 2018; Weaver et al. 2022) has resulted in a number of significant discoveries in galaxy evolution and cosmology.

The lasting impact that the COSMOS field has had on the contemporary astronomy is expected to continue for many decades into the future. Given the combination of the depth, and the multi-wavelength coverage, COSMOS has become a perfect data-set for studying statistically significant samples of brightest and rarest objects, ranging from the elusive quiescent populations to extremely luminous starbursts at $z > 5$ and capturing the youngest galaxies at the end of reionisation. To further this legacy, a significant part of the field will be targeted by *JWST* NIRCam and MIRI observations, among others this includes the COSMOS - Webb Cycle I GO program. This work focuses specifically on predicting the observed fluxes, source distribution, mock images and the physical parameters which will be observed by the MIRI F770W filter, at 7.7 μm.

This chapter is organised as follows. In Section 4.3 I describe the COSMOS - Webb survey, which forms the basis of my work. In Section 4.4 I describe the photometric data-sets, and in Section 4.5 I present the SED fitting algorithms and methodology behind computing the expected MIRI flux in the F770W filter. In Section 4.7 I perform various simulations to determine the expected number of sources that would end up being observed, given the choice of SED fitting methodology. Section 4.7.1 presents the early work in modelling the COSMOS sources, within the expected coverage, and creating mock images. In Section 4.8 I present the preliminary testing of the source extraction algorithms on the mock images. I discuss the implications that our findings have and present my main conclusions and summary in Section 4.9.

In this work I adopt a flat ΛCDM cosmology with $\Omega_{\mathrm{m},0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$ and $\mathrm{H}_0 = 70$ km s^{-1} Mpc^{-1}. For all the templates used, and output parameters I have assumed a Chabrier (2003) initial mass function (IMF).

4.3 Survey Description

The COSMOS - Webb survey, is a 200 hour *JWST* Cycle I GO program (GO-1727), designed to map a contiguous area of the sky, equal to 0.6 deg^2, with 4 filters of NIRCam imaging, and additionally a smaller 0.2 deg^2 segment, contained within the NIRCam area, with MIRI, by leveraging the parallel observation capabilities of *JWST*. A preliminary survey layout, and its comparison to the original COSMOS area can be seen in Figure 4.1. The observations are expected to survey around half a million sources with NIRCam, and tens of thousands with MIRI, and will be the widest contiguous survey area observed in the first year of *JWST* operations. The COSMOS - Webb program design in itself, is unique, since *JWST* has not been designed as a survey telescope, and the majority of first year programs will focus on deep pencil-beam surveys which will image only small areas of the sky. One other notable large-area *JWST* program is Public Release IMaging for Extragalactic Research (PRIMER; PI: J. Dunlop), which will target the *HST* CANDELS Legacy Fields (COSMOS and UDS)[1]. Alternatively, Cycle I GTO and GO programs will also target a set of known objects, with the aim of confirming their redshifts, and extracting properties of their stellar populations. On the other hand, the large coverage of COSMOS - Webb can allow us to capture the large-scale structure at the beginning of the epoch of reionisation, as well as enable us to find the most distant, rare galaxies, a task that can not be easily accomplished with very narrow survey designs. Following the legacy of COSMOS, the COSMOS - Webb survey data will be made publicly available immediately after collection (at least for the raw data), allowing the wide astronomical community to leverage the unique capabilities of *Webb* and study galaxies at an unprecedented level of detail, across the entirety of the history of the Universe.

The COSMOS - Webb survey has three primary science goals. Firstly, it will detect and study thousands of early galaxies, during the epoch of reionisation ($z = 6 - 11$), to understand their distribution, physical conditions, and drivers of early star formation. Secondly, the depth and wavelength coverage of the survey will allow detections of thousands of quiescent galaxies at $z > 4$, and unravel the formation of the most massive galaxies in the Universe. Finally, these observations will aim to measure the stellar to halo mass relation, out to $z \sim 2.5$, and study how it changes with respect to star formation histories and galaxy morphology.

4.4 Data

To produce the simulated MIRI/F770W catalogue, number counts and mock images, I use a set of galaxies detected and characterised in the COSMOS 2020 catalogue (C2020 heareafter) (Weaver et al. 2022). The major improvements in C2020 compared to the previous iteration of the catalogue (Laigle et al. 2016) are the much deeper optical and NIR photometry, processed and extracted from the Subaru - Hyper Suprime Cam (HSC) (Aihara et al. 2018, 2019, 2022) and VISTA-VIRCAM surveys (McCracken et al. 2012; Moneti et al. 2019). Moreover, the updated catalogue contains a reprocessed set of all *Spitzer* data that covers the COSMOS field. In addition, all of the new and existing data have been re-aligned to the new astrometry from Gaia. The catalogue covers an impressive photometric range, spanning 29 bands in total, from from FUV/GALEX at 1526 Å to *Spitzer*/IRAC channel 4 at 8 μm. Photometry in these bands has been extracted by using two different methods. The `CLASSIC` verison catalogue utilises `SExtractor` (Bertin & Arnouts 1996) to detect sources, and then performs aperture photometry. The other version uses `sep` (Barbary 2016) to detect sources, and a newly developed FARMER (Weaver et al. 2022, in prep)

[1]PRIMER COSMOS area is fully contained within COSMOS - Webb coverage. Although the survey area is smaller, PRIMER is deeper and also uses different filters.

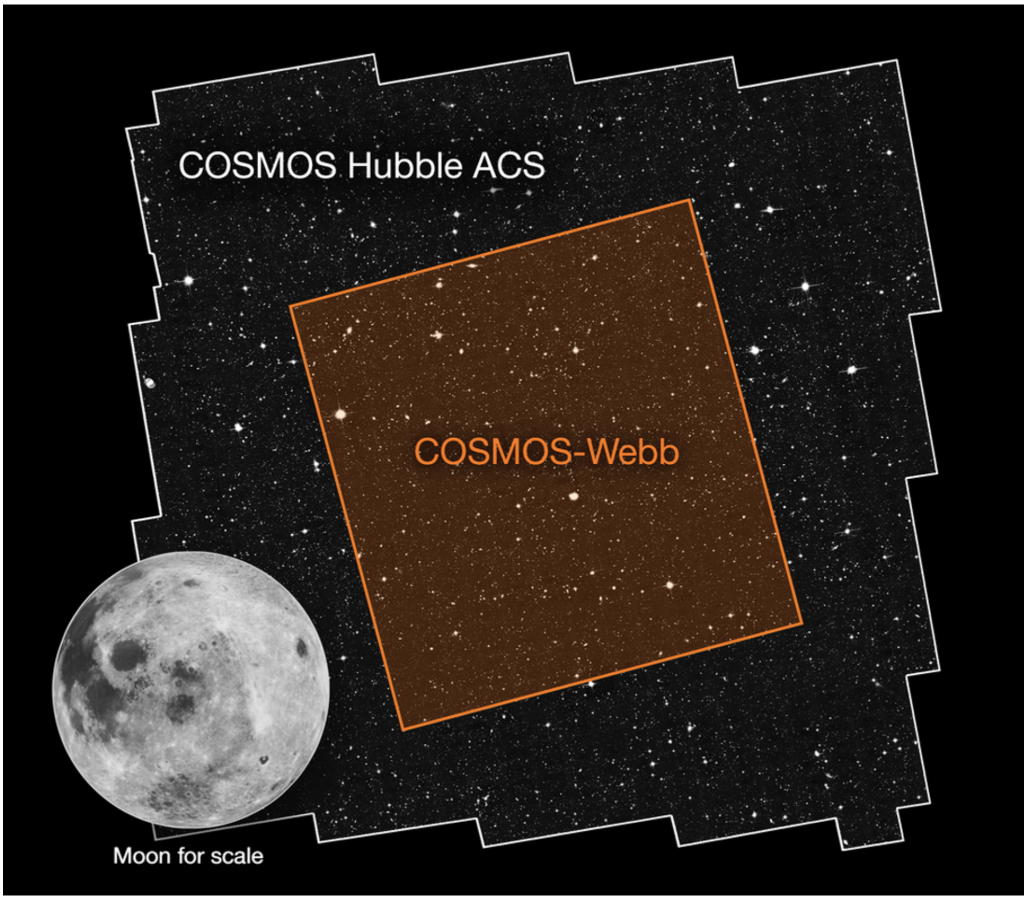

Figure 4.1: The full map of the COSMOS - Webb survey. The orange area shows the *JWST*/NIRCam observations, which will cover an area of 0.6 square degrees on the sky, roughly equivalent to the size of 3 full Moons. The mid - infrared coverage with MIRI is contained within the NIRCam area, and covers a smaller portion of the sky at 0.2 square degrees. The exact positions of MIRI exposures are still under development, and are therefore not displayed. In the background, a combined *HST*/ACS mosaic represents the original COSMOS area. Figure credit: COSMOS collaboration.

algorithm to calculate the photometry. Generally, the two catalogues are consistent with one another, however the FARMER catalogue allows for more model complexity. This in return might be useful when modelling nearby galaxies, which could be potentially resolved by MIRI. For this reason I have decided to utilise the FARMER version of C2020.

Both catalogues were additionally processed by using SED fitting codes EAZY (Brammer et al. 2008), and LePhare (Arnouts et al. 1999; Ilbert et al. 2006) in order to extract photometric redshifts (photo$z$) and physical parameters of the stellar population such as star formation rates (SFR) and stellar masses(M_*). For the purposes of our simulations I have adopted the EAZY derived photo$z$s and physical parameters. The full data-set contains $\sim 830,000$ objects, however, I focus my efforts only on the 0.6 deg^2 area contained within the NIRCam coverage of COSMOS - Webb. In combination with the flags to remove potential stars, this reduces the total number of analysed galaxies to $274,286$.

Over the last decade large parts of the COSMOS field have been covered by a multitude of IR and submm/mm observatories. Therefore, in order to improve my subsequent SED fitting analysis, I have further expanded the catalogue by including a set of MIR-FIR observations. These include MIPS 24 μm data from the COSMOS - *Spitzer* program (PI: D. Sanders; Le Floc'h et al. 2009), *Herschel*/PACS 100 and 160 μm data from the PEP (PI: D. Lutz; Lutz et al. 2011) and CANDELS-*Herschel* (PI: M. Dickinson) programs. The *Herschel*/SPIRE data at 200, 350 and 500 μm have been obtained during the Herschel Multi-tiered Extragalactic Survey (HerMES; PI: S.Oliver; Oliver et al. 2012). Finally, the JCMT/SCUBA2 850 μm images are from the S2CLS program (Cowie et al. 2017; Geach et al. 2017), and the AzTEC 1.1 mm data were described in Aretxaga et al. (2011). Although these data have already been processed and described in the "Super - deblended" analysis of the COSMOS field (Jin et al. 2018), the catalogue of Jin et al. is based on the COSMOS 2015 astrometry (Laigle et al. 2016), which makes blind cross-matching highly uncertain and might impact the quality of SED fitting. Generally, for the purposes of our work the FIR photometry is only required as a set of upper-limits, to avoid non-physical IR template solutions. To this end, and to eliminate any cross-matching uncertainties, I manually extract the flux densities from the above data-sets. For each galaxy in our sub-catalogue I find the corresponding central pixel value on each map, and record that as the final flux density. As the uncertainty on the measurement I use the 1σ limit, as presented in Jin et al. (2019) and Table 4.1. Where appropriate the units have been converted to either uJy/pix or uJy/beam.

Table 4.1: Properties of the infrared images

Instrument	Wavelength [μm]	1σ [mJy]
Spitzer/MIPS	24	0.01
Herschel/PACS	100	5.5
Herschel/PACS	160	4.2
Herschel/SPIRE	250	6.8
Herschel/SPIRE	350	7.5
Herschel/SPIRE	500	7.0
JCMT/SCUBA2	850	1.0
ASTE/AzTEC	1100	1.6

4.5 Modelling Galaxy SEDs

In order to construct the mock photometric catalogues, and simulated images, I chose to use the measured photometry from the C2020 catalogues (Weaver et al. 2022). I use the existing best-fit SEDs and physical parameters from the C2020 efforts, as well as calculate my own with a different SED fitting software. Based on the extracted flux densities, I perform a set of simulations to determine the expected number of objects that will fall within the MIRI exposures. The extracted flux densities and number counts are then used as a guide to create a set of mock images covering the full area of expected observations. To produce the synthetic photometry in the F770W filter, I have utilised two SED fitting approaches. Firstly, with EAZY (Brammer et al. 2008) and then with Stardust (Kokorev et al. 2021).

EAZY is a photometric spectral energy distribution fitting code which is optimised to fit non-negative linear combinations of basis-set templates, rather than devising a solution from vast libraries of pre-compiled models. These templates are extracted from the Flexible Stellar Populations Synthesis (FSPS; Conroy et al. 2009) models, and then reduced to a set of 12, which are able to reproduce a much larger library, spanning a variety of dust attenuation, ages, mass-to-light ratios and star-formation history (SFH) properties (see Brammer et al. 2008 and Blanton & Roweis 2007). Instead of using galaxy templates to model the MIR - FIR emission, EAZY uses the energy balance approach instead, similarly to, for example MAGPHYS (Battisti et al. 2019) and CIGALE (Burgarella et al. 2005; Noll et al. 2009; Boquien et al. 2019). For each object in a given multi-wavelength catalogue, EAZY integrates redshifted templates through a set of filters corresponding to our observed bands, and then finds the best combination for a given set of flux densities and associated uncertainties. While the C2020 catalogue already contains EAZY photometric redshifts, and physical parameters, my analysis would require best-fit SEDs in order to extract the predicted photometry at 7.7 μm. To perform this analysis I re-use the same configuration files and templates as in the original C2020 run. I do not include the additional MIR - FIR photometry which I manually extracted from the images, as EAZY templates do not cover the IR part of the spectrum. While re-fitting all sources I also verify that the this procedure returns exactly the same physical parameters as the original C2020 catalogue. Then I use the best-fit coefficients for all $274,286$ objects in our compilation and integrated the F770W filter band-pass through all of the SEDs in the observed frame.

Stardust (see Chapter 2) is a UV-optical-FIR SED fitting code, which uses the same computational principles as EAZY, by finding the best - fit coefficients from linear combinations of templates. The code utilises three separate sub-classes of templates, the UV-optical stellar emission, the MIR AGN heated dust, and the NIR-FIR dust. The stellar libraries consist of updated FSPS templates which were presented with the original EAZY software. The infrared library consists of a set of Draine & Li (2007) (DL07 hereafter) templates, with the additional updates from Draine et al. 2014 (also see Aniano et al. 2020). In brief, the DL07 templates separate dust into separate components. The first component described the contribution from warm dust and polycyclic aromatic hydrocarbon (PAH) features in the photo-dissociation regions (PDR), while the second one characterises the cold dust emission in the diffuse part of the ISM. Generally, the DL07 templates can be tweaked to a high degree of precision, however given the limited amount of NIR - FIR coverage in our data I adopt a simpler range of parameters. For each object in the catalogue, the minimum radiation field intensity ($U_{\rm min}$), a proxy for dust temperature, has been allowed to vary within the ranges predicted by the $\langle U \rangle$ - z evolutionary trends, as described in

Béthermin et al. (2015) and Kokorev et al. (2021). Although the effect of $U_{\rm min}$ parameter on 7.7 μm flux density is minimal, especially at higher redshifts, limiting the parameter grid rules out unnecessarily cold and unrealistic solutions, while also improving computational speeds. By far, the fraction of dust contained within PDRs (γ) and fraction of the total dust mass locked in PAHs ($q_{\rm PAH}$) are the largest contributors to the variation in F770W flux density. To constrain these parameters, I have used the works by Magdis et al. (2012a) and Kokorev et al. (2021), which have presented the distributions of both γ and $q_{\rm PAH}$ for statistically meaningful samples of SFGs. Both find that γ can be fixed at $\sim 2\%$, and the catalogue attached to Kokorev et al. (2021) shows that $q_{\rm PAH}$ rarely exceeds 3%. I have additionally fixed $U_{\rm max} = 10^6$ and $\alpha = 2$, as described in Magdis et al. 2012a. In this preliminary analysis, I have disabled the AGN - heated dust templates, to avoid introducing extra complexity to our simulations. The templates used in `Stardust` extend to the mm regime, therefore I incorporate the extracted MIR - FIR photometry into our fitting. The F770W flux density is then extracted from the best-fit templates using the same method as before.

4.6 Comparing the Derived Flux Densities

In this section I compare the 7.7 μm magnitudes extracted by using two different SED fitting approaches. I show the magnitude difference (ΔMag) between the two, and its evolution with redshift, in Figure 4.2. The 7.7 μm predicted by EAZY is consistently fainter than that of `Stardust`, with the median difference calculated at 0.63 magnitudes, corresponding to roughly a factor of two in flux density. This difference is not uniform, and is the highest at low-z, within the $0 < z < 0.5$ and $1 < z < 2$ ranges, peaking at $z \sim 0.1$ and $z \sim 1.5$. The overall difference is coming from the fact that in the low - z regime ($z \sim 0.1$), the F770W prediction is largely defined by the presence of PAH lines in the MIR. The IR emission in the EAZY templates is defined purely by energy balance, and does not contain the PAH features which are present in the Draine & Li templates utilised by `Stardust`. In fact, the disparity between flux densities which I report is consistent with the ionised PAH lines at 6.2, 7.7, and 8.6 μm. On the other hand, I believe that the "bump" seen at $z \sim 1.5$ in Figure 4.2, is caused primarily by the difference in PDR continuum, rather than PAH lines, since the Draine & Li. In Figure 4.3 I show a comparison between the EAZY and `Stardust` derived best fit SEDs, in the 1 - 20 μm regime. All of the galaxies shown have a significant ($> 5\sigma$) detection in IRAC 8μm, and therefore can act as a good benchmark for testing which approach is more appropriate for the F770W flux density modelling. From the figures it is apparent that the EAZY templates can not adequately model the continuum emission from the warm dust grains at $\lambda > 6$ μm. It is also worth noting that the neutral PAH line feature at 3.3 μm can also be partially responsible for the disparity, and in fact we see this line to be generally brighter for the Draine & Li templates in Figure 4.3. Generally, however the 3.3 μm line is poorly constrained and as such can not be modelled reliably for any galaxy template.

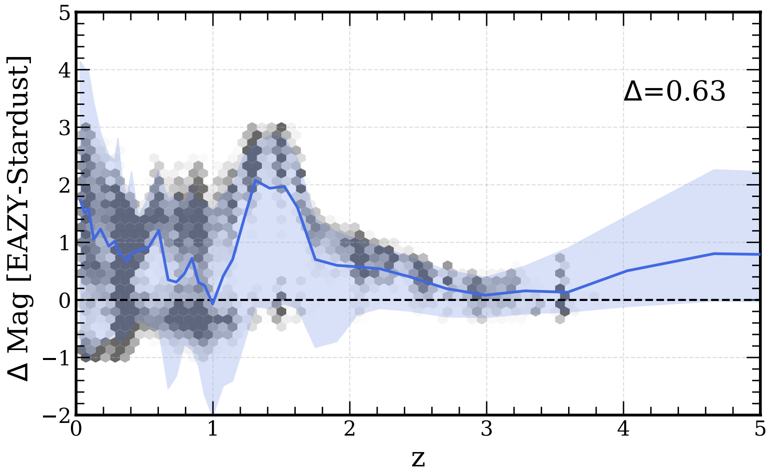

Figure 4.2: Magnitude difference between EAZY and `Stardust` predictions for the F770W magnitude, as a function of redshift. Individual sample of $\sim$ 270,000 galaxies is contained within the grey hexbins. The dashed grey line shows a 1:1 relation, the solid blue line traces the running median, and the shaded blue region contains the 68 % confidence interval.

At $z \sim 3$ the rest frame K - band shifts into 7.7 μm, and the flux density becomes dominated by stellar emission only. Since both SED fitting codes utilise very similar UV - stellar emission templates, the predictions are comparable within a factor of two at higher-z. Overall, I believe that the PAH emission would carry a significant portion of the MIR emission at 7.7 μm, especially at low-z range ($z < 2$), where the vast majority of C2020 - detected galaxies are located. In addition to that, the warm PDR dust continuum will be relevant up until $z \sim 3$. While the original GO proposal for COSMOS - Webb used EAZY predictions to calculate the expected number counts, it is clear that the energy balance approach is not capable in reproducing the necessary physical conditions at 7.7 μm for the vast majority of galaxies. Therefore in order to properly model the complexity of the MIR part of the spectrum, I decided to use the `Stardust` results in my simulations.

4.7 Predicting Galaxy Counts

The COSMOS - Webb survey will target a total of 0.2 deg^2 within 152 MIRI pointings. Each pointing includes a total of four 73.5" by 112.6" rectangular shaped exposures, with an area roughly equivalent to 2.3 arcmin^2 each. In order to maximise the signal, these exposures are arranged to overlap, which results in a pointing area of 4.1 arcmin^2, with an effective 5σ depth of 24.7 mag [AB] for point sources, as predicted by the *JWST* Exposure Time Calculator (ETC). The 152 pointings are then arranged into 8 contiguous stripes, within the NIRCam area. The positions of the MIRI pointings have not yet been finalised, however an approximate arrangement of the MIRI coverage can be seen in Figure 4.4.

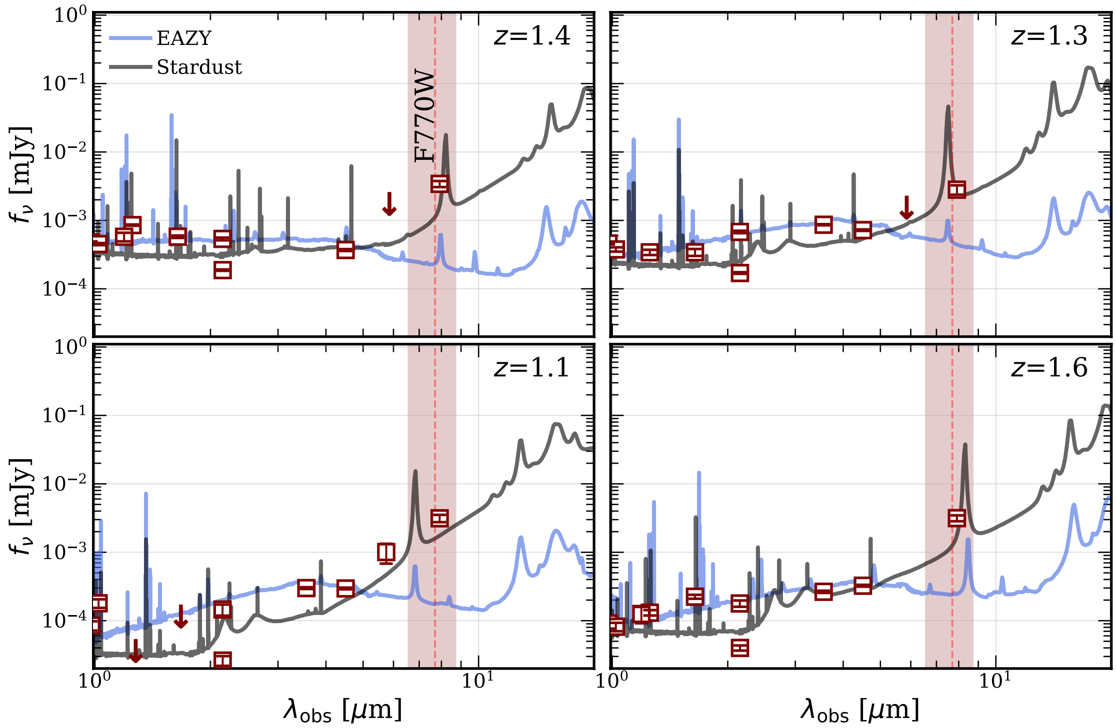

Figure 4.3: A comparison between the best fit SEDs from EAZY (blue line) and `Stardust` (black line), for a sample of $z \sim 1.5$ galaxies, selected to have secure IRAC Channel 4 detections. The open maroon squares show the measured photometry, while the shaded maroon region spans the $> 10\%$ throughput efficiency of the MIRI/F770W filter.

The next important step for this analysis would be to understand what number of sources is expected to fall into the MIRI areas. In order to calculate the average expected number of sources, I perform the following simulation. I begin by randomly placing 152 MIRI pointings within the NIRCam area, in a way that none of them overlap. A galaxy is then considered to be 'detected' if its total predicted flux density exceeds the 5σ rms of the survey. This assumes the simplest case, where all the emission from the source is concentrated within the central pixel. Realistically, of course, flux density will be distributed across the galaxy according to the chosen profile, and it is rather the surface brightness, than the total flux density that should be used to determine detection feasibility. In our case however, positions and (detection band) sizes of these galaxies are already known, which should allow prior based stacking to recover total flux densities. In my simulation I also resample the flux densities within their uncertainty, and re-calculate the number of detected sources, for each MIRI configuration. In total, the whole simulation is repeated 1000 times, for both EAZY and `Stardust` generated mock catalogues.

Using the EAZY mock catalogue I predict that a total of $28,327 \pm 192$ objects should be detected given the magnitude limit of COSMOS - Webb. This result is also consistent with the simulated result of the original proposal. In the case of the catalogue generated with `Stardust`,

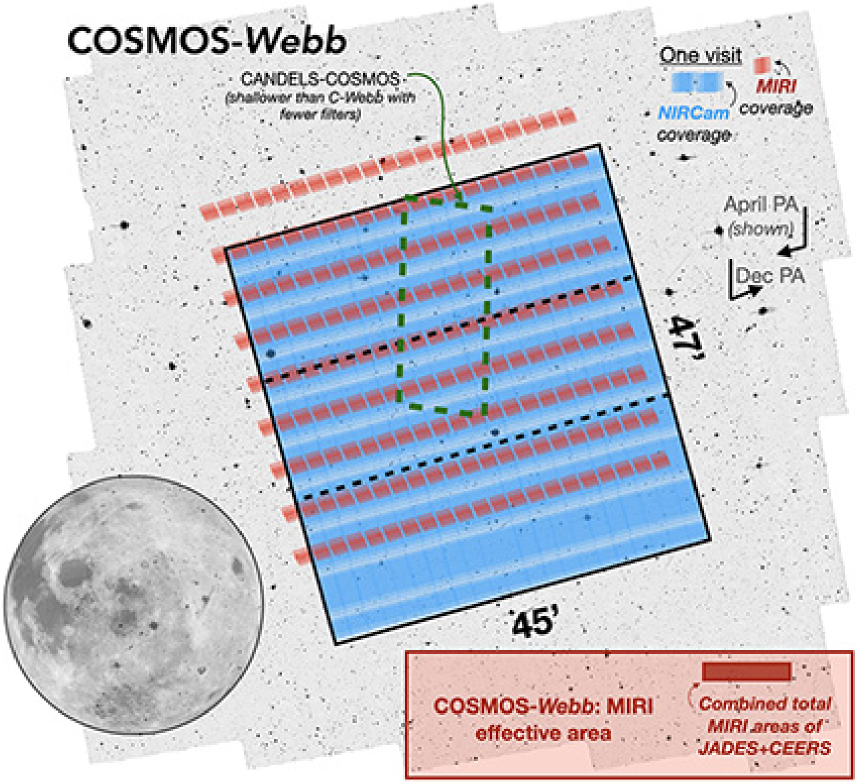

Figure 4.4: Preliminary positions for the MIRI pointings within COSMOS - Webb. Figure credit: COSMOS collaboration.

the total number of sources is $1.7\times$ higher, with a total of $49,263 \pm 331$. This result is somewhat unsurprising, given that `Stardust` predicts on average 0.5 mag brighter F770W magnitude.

In addition to simulating the number counts I can examine if different SED simulations introduce a bias towards what type of source is going to be detected by the survey. For both mock catalogues, and in each iteration of the simulation, I have also recorded exactly which galaxies are being detected. We then select $\sim 30,000\ (50,000)$ objects that have the largest amount of re-occurrence in the EAZY (`Stardust`) simulations. I present this comparison in Figure 4.5, for all the parameters (apart from magnitude) that are independent of the SED fitting methodology. Despite differences in the total flux density, both detected populations appear to have the same median effective radii ($R_{\rm eff}$), as given in the C2020 catalogue, and the total area (displayed in pixels; NPIX), at all redshifts. The detected galaxies derived from two methods however only overlap in about 50 % of cases. In Figure 4.6 I show the same comparison as before, but now only focusing on the objects "unique" to each simulation. Even in this case, however, the morphological properties are largely consistent with one another, and the only difference lie in exactly how much flux density is assigned to the F770W continuum by each SED fitting code.

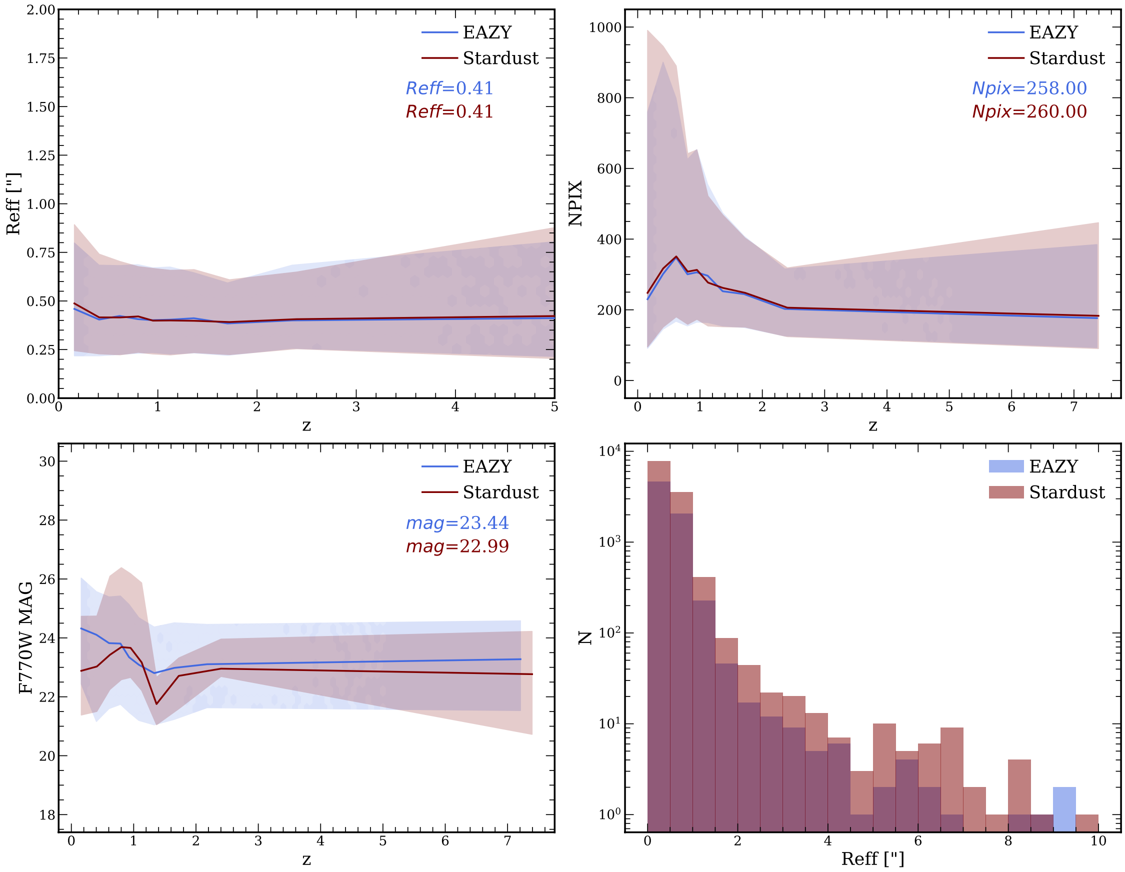

Figure 4.5: A comparison between all the sources for which the total F770W flux density exceeds 24.7 mag. Shown for both EAZY (blue) or `Stardust` (maroon) simulation. I examine the redshift evolution of the effective radius ($R_{\rm eff}$; top-left) ,the total area (displayed in pixels; top-right) and the F770W magnitude (bottom-left). The hexagonal bins in the background show the individual galaxy detections, while the solid coloured lines represent binned medians. The shaded regions correspond to the 68 % confidence interval. I also show how the effective radius is distributed in both cases (bottom-right). The median number of pixels belonging to each source - NPIX, is shown within each panel.

4.7.1 Generating Mock Images

The final step of my simulation is the creation of mock images covering the expected COSMOS - Webb MIRI area. For this, I first start by defining a rectangle containing the full NIRCam area, plus an additional padding of $\sim$ 6 arcmin on each side. As mentioned before, positions of MIRI exposures have not been finalised, and the additional padding is created to accommodate that. The final size of the image is 52.8 by 56.4 arcmin, with a total area of 0.83 deg^2. Based on the image size, an empty array is created to match the resolution of MIRI/F770W equal to 0."11/pix. A world coordinate system (WCS) grid is then assigned to the empty image. Source modelling has been performed by using the 2D Sérsic profile (Sérsic 1963) within `astropy` (Astropy Collaboration et al. 2013) Python package. Each model depends on 4 parameters, the effective radius ($R_{\rm eff}$), surface brightness at the effective radius $I_{\rm eff}$, ellipticity, and the Sérsic index - n. The C2020 uses a Python version of `SExtractor` (Bertin & Arnouts 1996) - `sep` (Barbary

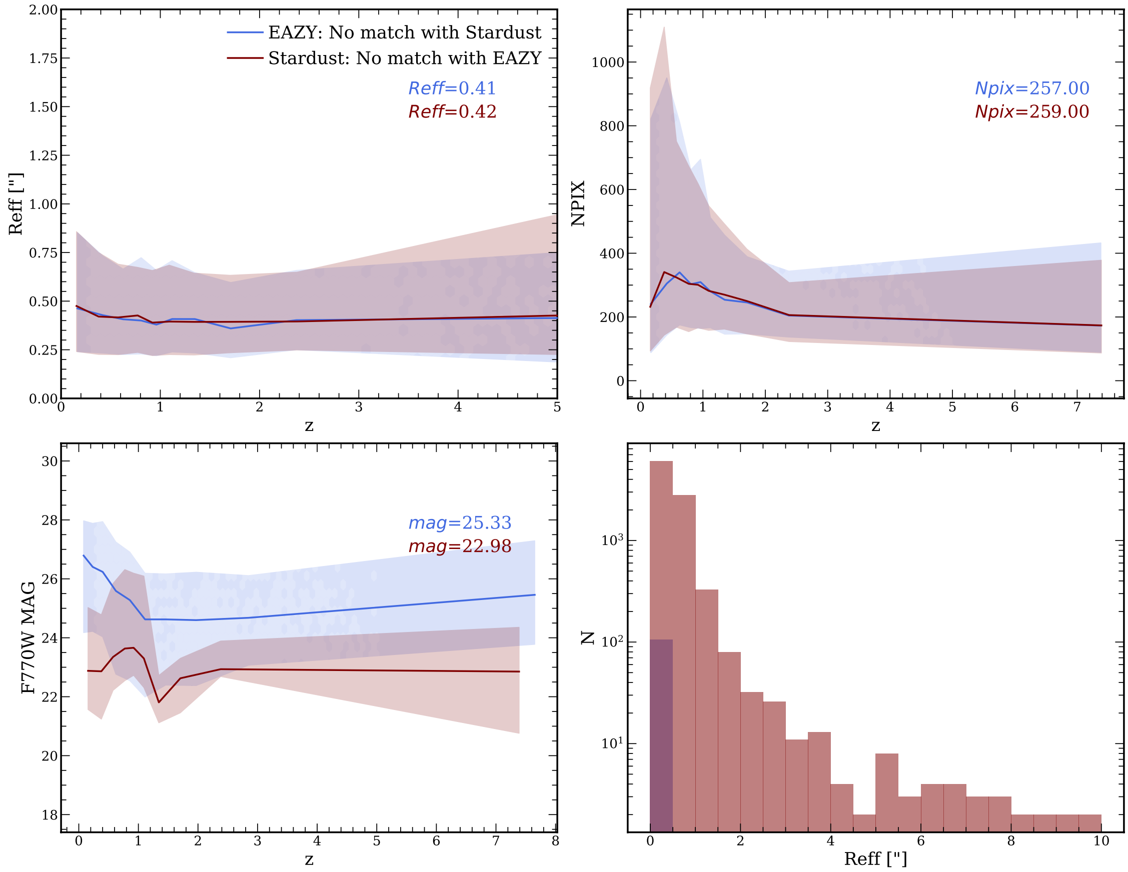

Figure 4.6: A comparison between only the unique sources detected given either EAZY (blue) or `Stardust` (maroon) simulation. Symbols and colours are the same as in Figure 4.5.

2016), in order to perform detection, and measure basic source sizes. The information contained in these basic parameters are the R_{eff} and ellipticity, as fitting for n quickly becomes degenerate for anything, but extremely resolved nearby galaxies. Instead, I use the Sérsic index distribution for the star-forming field galaxies presented in Allen et al. (2012). I approximate their distribution of Sérsic indices by a gamma function - $\Gamma(\alpha)$, with an α of 2. I draw a number from $\Gamma(2)$ distribution for each source in our mock catalogues, while also ensuring that n is physically meaningful, by limiting it between 0.5 and 10, as shown in Figure 4.7.

As mentioned before, the input model amplitude is assumed to be a surface brightness at R_{eff}, rather than the total flux density. One can convert between f_{tot} and I_{eff}, by integrating the Sérsic profile equation, given R_{eff} and n are also known. A faster alternative is to do it numerically, by creating models for each object in our catalogue, using the total flux density as an input, and then adding up all of the pixels in the model, which allows us to compute the $f_{\text{tot}}/I_{\text{eff}}$, for each object. By examining Figure 4.7 I see that the ratio is log-linear with R_{eff}, while also displaying a minor ~ 0.5 dex scatter due to different Sérsic indices. From this I can compute the correction factor from f_{tot} to I_{eff} by fitting a simple straight line. The correction, convolved with the n induced

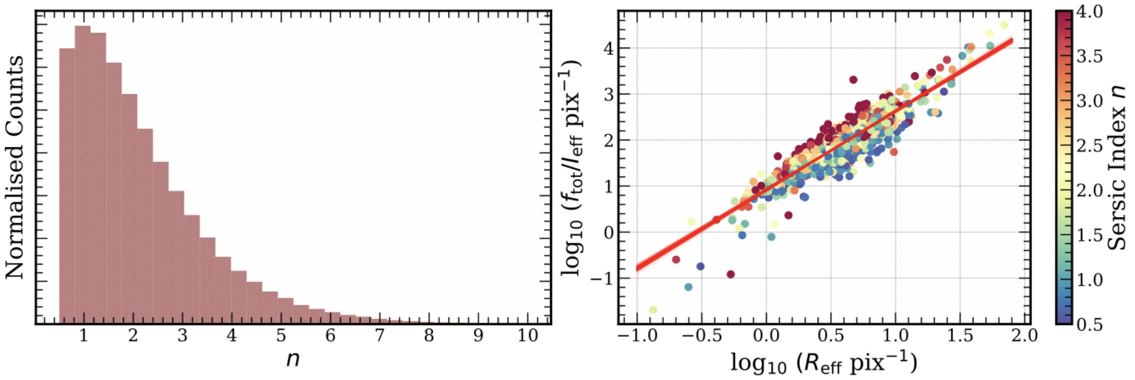

Figure 4.7: Left: Sérsic indices for galaxies in the sample, generated from the gamma function, based on the Allen et al. (2012) result. **Right:** Numerically derived correction between the total flux density ($f_{\rm tot}$) and the surface brightness at the effective radius ($I_{\rm eff}$) for a Sérsic profile, as a function of the effective radius. Point are colour coded by the Sérsic index. The solid red line with an envelope shows the best fit and the uncertainty of the $\log_{10}(f_{\rm tot}/I_{\rm eff}(R_{\rm eff})$ relation.

scatter can then be written as:

$$\log_{10}(f_{\rm tot}/I_{\rm eff}\,{\rm pix}^{-1}) = (1.73 \pm 0.02) \times \log_{10}(R_{\rm eff}\,{\rm pix}^{-1}) + (0.90 \pm 0.03). \quad (4.1)$$

With all the necessary parameters in hand, I can now start adding galaxy models to the empty image. This was done by using `photutils` (Bradley et al. 2020), an `astropy` based package which provides tools for modelling and extracting galaxy photometry. Adding models to the full 0.83 deg^2 image is computationally expensive and is generally not required as sources are rarely larger than a few tens of arcsec. To this end I additionally split our input image into 100 equally sized segments, as can be seen in Figure 4.8, and model sources within each slice separately. Each segment is a rectangle, 5.3 by 5.6 arcmin in size, with a total area of 0.08 deg^2. This size was chosen to ensure that very few galaxies are split in half by segmentation, and to minimise the effective time it take to add a single source to the image. This setting brings down the per source modelling time to ~ 0.5 s which is optimal for our current setup in terms of CPU time and RAM allocation, however it can be tweaked as necessary depending on available resources. In addition to that I flag every galaxy within $3 \times R_{\rm eff}$ from the slice/image boundary. In total there are only 1605 such objects, which constitutes less than 1% of the total sample.

I then recombine all 100 segments into the final image, and convolve the full map with the MIRI point spread function (PSF). The PSF was extracted from the SNR map, obtained from the online *JWST* ETC tool, by using the same detector setup as in the original COSMOS - Webb proposal. These include 45 groups per integration, 2 integrations per exposure and 2 exposures per specification, with the FASTR1 readout pattern. In the observing strategy I assume a 1."5 aperture, with the medium background configuration.

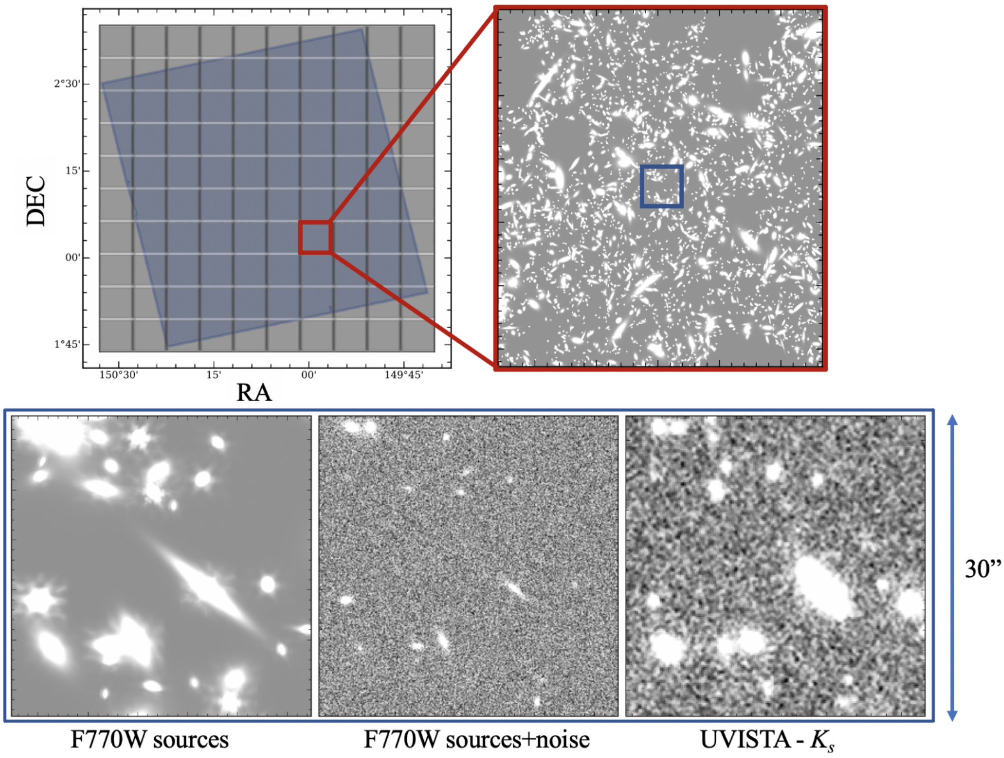

Figure 4.8: **Top Left:** Segments which make up the mock MIRI image. The blue area shows the NIRCam coverage. Large grey area is the full extent of our image, and the smaller squares are individual slices. **Top Right:** A zoom in on one of the segments showing the noiseless, PSF - convolved mock image. The empty areas on the map correspond to the original COSMOS 2020 star masks. **Bottom:** Cutout images (30") showing the F770W simulated image without noise, with noise, and the UVISTA - K_s band.

Although the formally cited survey 5σ rms is 24.7 mag AB, that does not correspond to the actual per pixel uncertainty, required to produce the background and noise added final image. To do that, I follow a similar procedure as we did for the PSF, and extract the ETC detector noise map in the units of e-/second/pixel. I convert that to the μJy/pixel by following the MIRI documentation, and compute the median, noise added background level at 2.7 ± 0.02 μJy/pixel. This is then added to the final, PSF convolved, image. Following this procedure, two sets of PSF convolved images are created for each catalogue, with and without noise. The example cutout, containing the sources constructed with `Stardust` predicted flux densities can be found in the right panel of Figure 4.8. The figure also shows how the real sources look like in the UltraVista K_s band filter. These images can be later on processed by a source extracting software e.g. `sep` to determine the most optimal source extraction parameters for the real images. To further complement the comparison between the simulation and the real data, I show a set of *Spitzer*/IRAC cutouts in all four filters in Figure 4.9.

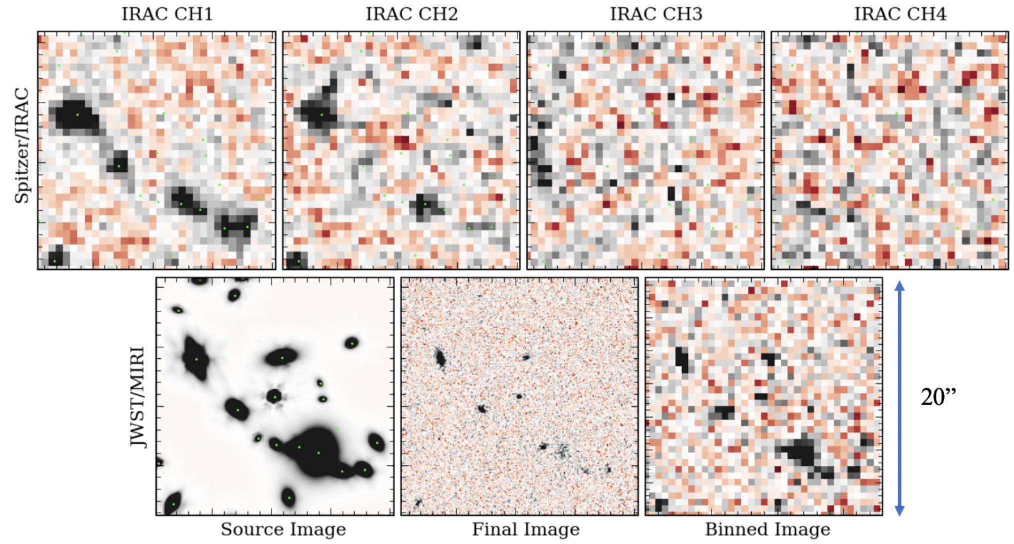

Figure 4.9: A set of 20" cutouts showing the observed galaxies look in IRAC CH1-CH4 images, and my simulation. **Top:** A set of IRAC images in 3.6, 4.5, 5.8 and 8μm. **Bottom:** The simulated MIRI F770W image without the added noise, with the noise, followed by a MIRI cutout binned to match IRAC 4 resolution.

4.8 Preliminary Source Detection Feasibility

Using the mock mosaics described in the previous section I conduct a preliminary source extraction with `sep` (Barbary 2016). I start by first calculating the background with `sep.background` and subtracting it from the image. Using the background subtracted image I then define a detection threshold as 1.0 of the rms, by using `sep.globalrms`. I set the minimum source area parameter `minarea` to 5, to avoid the detection of spurious sources. Additionally, we perform the image "cleaning" and deblending by setting the `clean_param` = 50, and `deblend_nthresh` to a default value of 32. Generally, however, the mock image is not crowded with an overlap fraction, i.e. the number of sources for which segments overlap with their neighbours is less than 5%.

A preliminary testing on the noise and background added image has resulted in $\sim$ 85% sources being recovered. In Figure 4.10 I show the completeness ($N_{\rm out}/N_{\rm tot}$), as a function of surface brightness, and the total 7.7 μm magnitude. I find that we are 100 $\%$ complete down to the surface brightness of 100 μJy/arcsecond^2, and a total F770W magnitude of 23. When comparing the input, and output total magnitudes, I find that that, on average, the flux density extracted by `sep` was $\sim$ 20% fainter than expected, which I predict is coming from the fact that `sep` apertures are smaller than the "total" size of the source. More testing is however required to find optimal parameters for the noise-convolved image.

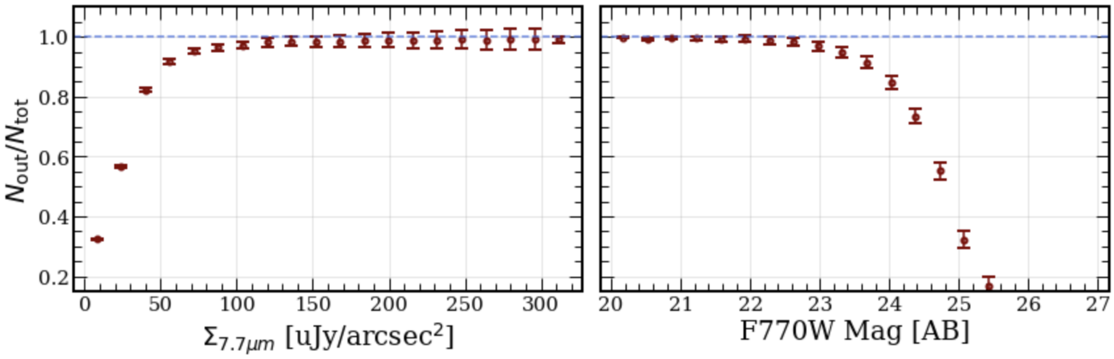

Figure 4.10: **Left:** The ratio between output and input sources, as a function of the surface brightness. The dashed blue line show the 100% completeness level. **Right:** The ratio between output and input sources, as a function of the total F770W magnitude.

4.9 Conclusions and Future Work

I have developed a set of mock photometric catalogues, and simulated MIRI/F770W mosaics to predict and guide the data reduction techniques for the upcoming COSMOS - Webb survey with *JWST*. Utilising the novel UV-optical-NIR catalogue in the COSMOS field, covering roughly a million sources, we have recovered the synthetic 7.7 μm photometry by using two different SED fitting approaches. Firstly with the UV-optical photometric redshift fitting code EAZY, which was also originally used to compute redshifts and the physical parameters for the C2020 catalogue, and then with `Stardust`, an algorithm that fits panchromatic photometry to sets of independently combined templates, including the MIR and FIR regimes. We find that in the $z < 1.5$ range, where PAH dominate the 7.7 μm observed photometry, `Stardust` predictions are, on average, 0.5 magnitudes brighter, than the ones computed with EAZY. The disparity between derived flux densities is a result of `Stardust` using the physically motivated MIR and FIR Draine & Li dust emission templates, which contain the PAH emission, as opposed to the energy balance approach taken by EAZY. Consequently, at higher redshifts, when optical bands are shifted into the F770W filter, the extracted flux densities are consistent with one another. By combining these mock catalogues with a number count simulation I compute that the expected amount of detected sources is ranging from $\sim$ 30,000 - 50,000, depending on the chosen SED fitting method.

From the derived flux densities, and original COSMOS 2020 source extraction parameters, I selected all the galaxies that are expected to be above the 1σ survey limit, and modelled them on a set of mock mosaics. The resultant, noise added, mosaics, coupled with preliminary source extraction testing with *sep*, however reveal that fewer sources than expected might be recovered. In comparison, previous NIR surveys conducted with *Spitzer*/IRAC had a resolution of 1.2"/pix for a comparable IRAC 8 μm filter. For MIRI, however the resolution is more than ten times higher, meaning that individual pixels in a resolved source are expected to be $\times$ 100 times fainter. It is not expected to be a major obstacle for COSMOS - Webb, as prior source positions and their extent on the sky are already known, therefore total flux density can be recovered via stacking or binning. Since this effect is especially prominent in the wings of disk galaxies, an adaptive binning strategy, e.g. Voronoi tessellation (Cappellari & Copin 2003), would be most appropriate to preserve at least some degree of spatial information, while still detecting individual pixels at an appreciable signal-to-noise. Similar adaptive binning strategies have already been used quite extensively for high resolution $Hubble$ imaging (e.g. Wuyts et al. 2012, 2013, 2014; Genzel et al. 2013; Lang et al. 2014; Tadaki et al. 2014; Chan et al. 2016). On the other hand, to detect objects which do not have any prior information the approach would be to bin the entire image. While it is rather unusual to discuss binning of NIR images, the new era of high resolution astronomy brought upon by $JWST$ will necessitate some changes in our approaches, this also highlight the importance of preliminary simulations and analysis.

In my future work I aim to optimise both the models and mock images themselves, as well as the source extraction parameters with e.g. `sep` to achieve the most optimal source recovery. In addition, and given the above discussion regarding the surface brightness limitations, we will conduct (prior-position based) adaptive and general image binning procedures to characterise the amount of information which can be extracted even in cases of initial non-detection. Moreover, I aim to make additions and improvements to the pipeline that I have developed for this work, by adding new SED fitting routines, and templates. At the moment, however, it is already fully customisable, and given a set of prior galaxy photometry, it can produce realistic flux density estimates and mock images for any filter/resolution covering a given field.

However, the COSMOS 2020 catalogue is complete only down to its selection limit, derived from the ground based optical bands, thus preventing detection of very distant and/or dusty galaxies. More specifically, within the area covered by MIRI one can expect to find $\sim$ 100 optically dark galaxies (e.g. see Wang et al. 2019), which can not be modelled by using any existing priors. To alleviate that, I aim to use the existing scaling relations describing the '*HST*-dark' sources, in conjunction with the phenomenological models (e.g. JAGUAR Williams et al. 2018) to explore the full dynamic range and potential of the MIRI observations in the COSMOS field. Tracing the optically obscured population is imperative towards understanding the contribution of DSFGS to the galaxy assembly beyond $z > 4$. This would become especially prominent, considering the upcoming MORA and its extension (Ex-MORA) large ALMA surveys at 2mm (Casey et al. 2021). The other important populations which is missing from our simulations are the high-z galaxies. To model these and predict the expected source counts I aim to use the recently updated FLARES simulations (Lovell et al. 2021; Wilkins et al. 2022). With the mock mosaics I aim to find and optimise the critical source detection and photometry extraction algorithms, in order to facilitate a rapid and robust data processing.

CHAPTER 5

Other Published Work

During the course of my PhD I have contributed to a series of projects that have led to publications in scientific peer-reviewed journals. Below I provide a list of articles to which I have made a significant contribution. The list also includes a short abstract for each, and a description of the work I have done for each article.

Title: A dusty compact object bridging galaxies and quasars at cosmic dawn

Journal: Nature, Volume 604, Issue 7905, p.261-265, (2022); doi:10.1038/s41586-022-04454-1

Authors: Fujimoto, S. ; Brammer, G. B. ; Watson, D. ; Magdis, G. E. ; **Kokorev, V.** ; Greve, T. R. ; Toft, S. ; Walter, F.; Valiante, R. ; Ginolfi, M. ; Schneider, R. ; Valentino, F. ; Colina, L. ; Vestergaard, M. ; Marques-Chaves, R. ; Fynbo, J. P. U. ; Krips, M. ; Steinhardt, C. L. ; Cortzen, I. ; Rizzo, F. ; Oesch, P. A.

Description: *This article describes the detection, and the physical parameters of a* $z = 7.2$ *galaxy - GNz7q. The source displays properties of both dusty starburst galaxy, and that of a UV compact quasar progenitor. The detection of GNz7q presents a unique case study of an antecedent to unobscured luminous quasars at late cosmic times. For this work I have contributed the* `Stardust` *SED fits, which were used to extract the IR properties of GNz7q. I also co-wrote the methodology section.*

Title: A^3COSMOS: A census on the molecular gas mass and extent of main-sequence galaxies across cosmic time

Journal: Astronomy & Astrophysics, Issue 660, A142 (2022); doi:10.1051/0004-6361/202142299

Authors: Wang, Tsan-Ming ; Magnelli, Benjamin ; Schinnerer, Eva ; Liu, Daizhong ; Aziz Modak, Ziad ; Faustino Jiménez-Andrade, Eric ; Karoumpis, Christos ; **Kokorev, Vasily** ; Bertoldi, Frank.

Description: *This work describes an innovate method which uses ALMA stacking in the* uv *- plane. The analysis is used to extract the mean mass and the extent of the molecular gas for a sample of main-sequence galaxies at* $0.4 < z < 3.6$*. In this work, I have contributed to the scientific discussion by providing* `Stardust` *fits to the photometry, and through general comments.*

Title: GOODS-ALMA 2.0: Compact star formation as a sign of self-regulation before quenching

Journal: Astronomy & Astrophysics, Volume 659, id.A196, 26 pp. (2022); doi:10.1051/0004-6361/202142352

Authors: Gómez-Guijarro, C. ; Elbaz, D. ; Xiao, M. ; **Kokorev, V. I.** ; Magdis, G. E. ; Magnelli, B. ; Daddi, E. ; Valentino, F. ; Sargent, M. T. ; Dickinson, M. ; Béthermin, M. ; Franco, M. ; Pope, A. ; Kalita, B. S. ; Ciesla, L. ; Demarco, R. ; Inami, H. ; Rujopakarn, W. ; Shu, X. ; Wang, T. ; Zhou, L. ; Alexander, D. M. ; Bournaud, F. ; Chary, R. ; Ferguson, H. C. ; Finkelstein, S. L. ; Giavalisco, M. ; Iono, D. ; Juneau, S. ; Kartaltepe, J. S. ; Lagache, G. ; Le Floc'h, E. ; Leiton, R. ; Leroy, L. ; Lin, L. ; Motohara, K. ; Mullaney, J. ; Okumura, K. ; Pannella, M. ; Papovich, C. ; Treister, E.

Description: *This work describes a follow up of a DSFG population at* $1 < z < 5$*. The authors find that the SFR in galaxies is sustained by both gas and the compression of star formation, which allows these objects to remain on the main-sequence longer, even if gas fractions are low. For this work I have contributed the* `Stardust` *SED fits. I also helped co-write the methodology and the discussion sections.*

Title: COSMOS2020: A panchromatic view of the Universe to $z \sim 10$ from two complementary catalogs

Journal: The Astrophysical Journal Supplement Series, Volume 258, Issue 1, id.11, 30 pp. (2022); doi:10.3847/1538-4365/ac3078

Authors: Weaver, J. R. ; Kauffmann, O. B. ; Ilbert, O. ; McCracken, H. J. ; Moneti, A. ; Toft, S. ; Brammer, G. ; Shuntov, M. ; Davidzon, I. ; Hsieh, B. C. ; Laigle, C. ; Anastasiou, A. ; Jespersen, C. K. ; Vinther, J. ; Capak, P. ; Casey, C. M. ; McPartland, C. J. R. ; Milvang-Jensen, B. ; Mobasher, B. ; Sanders, D. B. ; Zalesky, L. ; Arnouts, S. ; Aussel, H. ; Dunlop, J. S. ; Faisst, A. ; Franx, M. ; Furtak, L. J. ; Fynbo, J. P. U. ; Gould, K. M. L. ; Greve, T. R. ; Gwyn, S. ; Kartaltepe, J. S. ; Kashino, D. ; Koekemoer, A. M. ; **Kokorev, V.** ; Le Fèvre, O. ; Lilly, S. ; Masters, D. ; Magdis, G. ; Mehta, V. ; Peng, Y. ; Riechers, D. A. ; Salvato, M. ; Sawicki, M. ; Scarlata, C. ; Scoville, N. ; Shirley, R. ; Silverman, J. D. ; Sneppen, A. ; Smolcić, V. ; Steinhardt, C. ; Stern, D. ; Tanaka, M. ; Taniguchi, Y. ; Teplitz, H. I. ; Vaccari, M. ; Wang, W. -H. ; Zamorani, G.

Description: *This work presents the collection, processing and analysis of the novel imaging data in the COSMOS filed to produce a new multi wavelength photometric catalog. I contributed to this work by aiding in the scientific discussion, galaxy photometry extraction and providing comments for the entire manuscript.*

Title: Mapping Obscuration to Reionization with ALMA (MORA): 2mm Efficiently Selects the Highest-Redshift Obscured Galaxies

Journal: The Astrophysical Journal, Volume 923, Issue 2, id.215, 32 pp. (2021); doi:10.3847/1538-4357/ac2eb4

Authors: Casey, Caitlin M. ; Zavala, Jorge A. ; Manning, Sinclaire M. ; Aravena, Manuel ; Béthermin, Matthieu ; Caputi, Karina I. ; Champagne, Jaclyn B. ; Clements, David L. ; Drew, Patrick ; Finkelstein, Steven L. ; Fujimoto, Seiji ; Hayward, Christopher C. ; Dekel, Anton M. ; **Kokorev, Vasily** ; Lagos, Claudia del P. ; Long, Arianna S. ; Magdis, Georgios E. ; Man, Allison W. S. ; Mitsuhashi, Ikki ; Popping, Gergö ; Spilker, Justin ; Staguhn, Johannes ; Talia, Margherita ; Toft, Sune ; Treister, Ezequiel ; Weaver, John R. ; Yun, Min.

Description: *This work presents observations of 2 mm selected population of DSFGS at* $\langle z \rangle = 3.6$*. Detection and analysis of such objects represents an important step in estimating the total contribution dusty SFGs have to the total SFRD in the first few billion years of the lifetime of the Universe. I have contributed to the scientific discussion and conclusions part of the article.*

Title: The effect of active galactic nuclei on the cold interstellar medium in distant star-forming galaxies

Journal: Astronomy & Astrophysics, Volume 654, id.A165, 19 pp. (2021); doi:10.1051/0004-6361/202141417

Authors: Valentino, F. ; Daddi, E. ; Puglisi, A. ; Magdis, G. E. ; **Kokorev, V.** ; Liu, D. ; Madden, S. C. ; Gómez-Guijarro, C. ; Lee, M. -Y. ; Cortzen, I. ; Circosta, C. ; Delvecchio, I. ; Mullaney, J. R. ; Gao, Y. ; Gobat, R. ; Aravena, M. ; Jin, S. ; Fujimoto, S. ; Silverman, J. D. ; Dannerbauer, H.

Description: *This article describes an examination of the ISM properties for a sample of infrared detected* $z \sim 1.5$ *main-sequence and starburst galaxies, which also possess a significant AGN component. The authors find that AGN has minimal effects on IR derived properties such as dust and gas fractions, as well as SFE. For this article I have performed SED fitting with* `Stardust`*, contributed to the scientific discussion and conclusions sections, and provided general comments.*

Title: ALMA Lensing Cluster Survey: Bright [CII] 158 μm Lines from a Multiply Imaged Sub-L_* Galaxy at $z = 6.0719$

Journal: The Astrophysical Journal, Volume 911, Issue 2, id.99, 20 pp. (2021); doi:10.3847/1538-4357/abd7ec

Authors: Fujimoto, Seiji; Oguri, Masamune; Brammer, Gabriel; Yoshimura, Yuki; Laporte, Nicolas; González-López, Jorge; Caminha, Gabriel B.; Kohno, Kotaro; Zitrin, Adi; Richard, Johan; Ouchi, Masami; Bauer, Franz E.; Smail, Ian; Hatsukade, Bunyo; Ono, Yoshiaki; **Kokorev, Vasily**; Umehata, Hideki; Schaerer, Daniel ; Knudsen, Kirsten; Sun, Fengwu; Magdis, Georgios; Valentino, Francesco; Ao, Yiping; Toft, Sune; Dessauges-Zavadsky, Miroslava; Shimasaku, Kazuhiro; Caputi, Karina; Kusakabe, Haruka; Morokuma-Matsui, Kana; Shotaro, Kikuchihara; Egami, Eiichi; Lee, Minju M.; Rawle, Timothy; Espada, Daniel.

Description: *The authors present a detection of a bright [CII] emission in a $z = 6$ lensed LBG located in the ALCS field. The findings indicate that the relation between $L_{\rm CII}$ and SFR found in local galaxies is universal, and still holds at high$-z$. For this work I have fit the FIR photometry, and contributed to the scientific discussion section.*

Title: Submillimetre compactness as a critical dimension to understand the main sequence of star-forming galaxies

Journal: Monthly Notices of the Royal Astronomical Society, Volume 508, Issue 4, pp.5217-5238 (2021); doi:10.1093/mnras/stab2914

Authors: Puglisi, Annagrazia; Daddi, Emanuele; Valentino, Francesco; Magdis, Georgios; Liu, Daizhong; **Kokorev, Vasily**; Circosta, Chiara; Elbaz, David; Bournaud, Frederic; Gomez-Guijarro, Carlos; Jin, Shuowen; Madden, Suzanne; Sargent, Mark T.; Swinbank, Mark.

Description: *This work presents a study of the ISM for a sample of 77 SFGs at $z \sim 1.3$. The authors identify compact sub-mm size as a signature of a post-starburst MS population, following a merger driven starburst episode. I contributed to this article by performing* `Stardust` *fits to the FIR data.*

Title: The interstellar medium of quiescent galaxies and its evolution with time

Journal: Astronomy & Astrophysics, Volume 647, id.A33, 17 pp. (2021); doi:10.1051/0004-6361/202039280

Authors: Magdis, Georgios E.; Gobat, Raphael; Valentino, Francesco; Daddi, Emanuele; Zanella, Anita; **Kokorev, Vasily**; Toft, Sune; Jin, Shuowen; Whitaker, Katherine E.

Description: *In this article the authors present a characterisation of the FIR properties for a stacked sample of quiescent galaxies. The findings reveal that the dust and gas fractions in QGs steeply rise out to $z = 1$ and remain flat out to at least $z = 2$. In addition to that QGs were found to have gas fractions three times lower than those of normal star-forming galaxies at similar redshifts, and were also systematically colder. I have provided the data for the gas fraction evolution of normal galaxies, as well as contributed to the discussion and conclusion parts of this work.*

Title: CO emission in distant galaxies on and above the main sequence

Journal: Astronomy & Astrophysics, Volume 641, id.A155, 21 pp. (2020); doi:10.1051/0004-6361/202038322

Authors: Valentino, F.; Daddi, E.; Puglisi, A.; Magdis, G. E.; Liu, D.; **Kokorev, V.**; Cortzen, I.; Madden, S. ; Aravena, M.; Gómez-Guijarro, C.; Lee, M. -Y.; Le Floc'h, E.; Gao, Y.; Gobat, R.; Bournaud, F.; Dannerbauer, H.; Jin, S.; Dickinson, M. E.; Kartaltepe, J.; Sanders, D.

Description: *The authors present ALMA observations of the CO line, detected in few tens of infrared-selected galaxies at $1.1 < z < 1.7$. The analysis adds evidence towards starburst galaxies having a higher excitation of the spectral energy distribution, mainly affected by the both large SFRs and compact sizes. I contributed to this article by performing `Stardust` fits to the photometry.*

Title: The Properties of the Interstellar Medium of Galaxies across Time as Traced by the Neutral Atomic Carbon [CI] ***Journal:*** The Astrophysical Journal, Volume 890, Issue 1, id.24, 24 pp. (2020); doi:10.3847/1538-4357/ab6603

Authors: Valentino, Francesco; Magdis, Georgios E.; Daddi, Emanuele; Liu, Daizhong; Aravena, Manuel; Bournaud, Frédéric; Cortzen, Isabella; Gao, Yu; Jin, Shuowen; Juneau, Stéphanie; Kartaltepe, Jeyhan S.; **Kokorev, Vasily**; Lee, Min-Young; Madden, Suzanne C.; Narayanan, Desika; Popping, Gergö; Puglisi, Annagrazia.

Description: *The authors present ALMA observations of the [CI] and CO lines for 30 MS galaxies. The line ratios were fit to a set photodissociation models , which revealed systematically large densities and radiation field intensities in sub-mm galaxies, compared to main-sequence and local objects. I have contributed to the scientific discussion section, as well as provided general comments for the manuscript.*

CHAPTER 6

Ongoing and Future Work

6.1 Gas Giants

During my analysis of the COSMOS 2015 (Laigle et al. 2016; Jin et al. 2018; Kokorev et al. 2021) optical+FIR catalogue I identified a number of extreme outliers from the average $f_{\rm dust}$ and $f_{\rm gas}$ evolutionary trends (see Figure 6.1). These typically have $\log(f_{\rm gas}) > 0.5$, i.e. their gas mass reservoir takes up $\sim 75\%$ of total baryonic matter budget. Individual SEDs of these objects have been examined along with the cut-out images in order to identify either bad or insufficient optical photometry - resulting in an incorrect M_* estimate, poor coverage of the FIR peak or blending issues that could result in erroneously large $M_{\rm gas}$ estimates. With the above considerations, we are left with ~ 50 objects which are shown in red on Figure 6.1.

These sources are in stark contrast when compared to the expected evolutionary tracks, with the 'gas-giant' sample being elevated by at least a factor of 3$\times$ in both $f_{\rm dust}$ and $f_{\rm gas}$ compared to median empirical trends, and appearing as $\sim 5\sigma$ outliers to the parent sample. When contrasted with simulations the result are even more puzzling, with $f_{\rm dust}$ deviation of a factor 5$\times$ and a staggering factor of 10$\times$ in $f_{\rm gas}$, when parametrised to the same SFR/SFR$_{\rm MS}$ and M_*. The possible explanations for the very high $M_{\rm dust}$ and subsequently $M_{\rm gas}$ estimates, could be optically thick FIR emission, general lack and inconsistency in the deblended FIR photometry, or simply an erroneous gas-to-dust ratio ($\delta_{\rm GDR}$). All three of these hypotheses have been thoroughly tested, via simulations, and alternative SED fitting techniques, but failed to demote a significant fraction of the "gas giants". Thus the only possible explanation for these extreme $f_{\rm dust}$ ($f_{\rm gas}$) are the gigantic $M_{\rm dust}$ ($M_{\rm gas}$) reservoirs, which I want to probe for a pilot sample of galaxies, shown in blue in Figure 6.1.

Significantly elevated depletion times $\tau_{\rm depl}$ (Figure 6.1), could indicate that these objects are 'failed starbursts', where despite having seemingly massive $M_{\rm gas}$ reservoirs the star-formation mode is otherwise unremarkable. An alternative hypothesis is then suggested by the FLARES

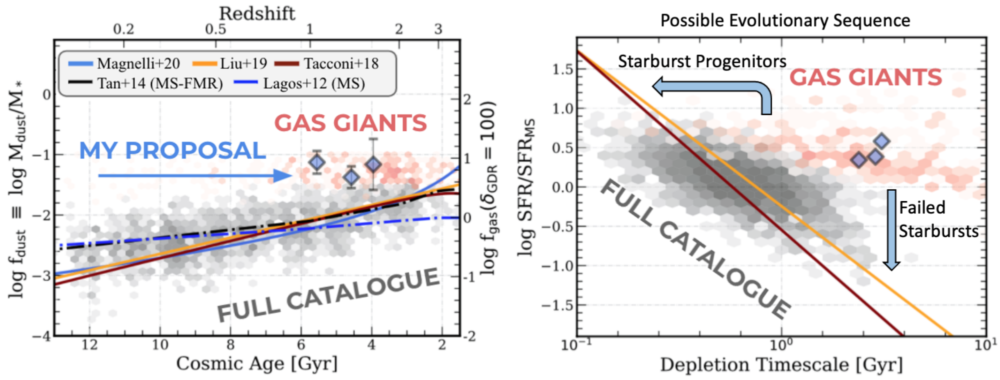

Figure 6.1: Top: Evolution of $f_{\rm dust}$ and $f_{\rm gas}$ as a function of z/cosmic age. Solid coloured lines correspond to empirical evolution by Tacconi et al. (2018); Liu et al. (2019a); Magnelli et al. (2020). The dashed-dotted lines show theoretical predictions of Lagos et al. (2012) and Tan et al. (2014). Both empirical and theoretical trends are normalised to SFR/SFR$_{\rm MS}$=2.4 and M_*=$10^{10.7}$ as our proposed sample, where possible. The grey hexbins contain the full sample, while red hexbins show positions of the gas-giants. The sample from my NOEMA proposal is marked by blue hexagons. **Bottom:** A possible evolutionary sequence for the "gas-giants", as suggested by FLARES simulations.

simulations (Lovell et al. 2021), wherein galaxies with significantly elevated gas fractions exist at $z \sim 4$ and then, following a gas reservoir destabilising episode, likely caused by AGN activity, transition into the star bursting phase by $z \sim 2-3$.

6.1.1 A Robust Sample of Gas Giants

To act as a pilot study of these unusual systems a robust sub-sample of these 'gas-giants' has been selected, which fulfil the following criteria: **1)** Secure FIR coverage with ≥ 3 continuum detections at 3σ level, and at least a single detection > 150 μm rest-frame for a robust $M_{\rm dust}$ and $M_{\rm gas}$ estimates, **2)** Target needs to be isolated, to facilitate a secure M_* estimate and avoid blending issues in the FIR, **3)** Since z_{phot} could be a major source of uncertainty in both M_* and $M_{\rm gas}$, before looking further into this population of 'gas-giants', the sample was narrowed to only include spectroscopically confirmed sources. Moreover, it is important to point out that these sources have what can be considered 'typical' values for the log(M_*) $\sim$ 10.6, and also do not appear to be strong SBs with median SFR/SFR$_{\rm MS}$ $\sim$ 2.4, thus making them even more unique and puzzling. The best-fit SEDs along with cutouts for the proposed sample are shown in Figure 6.2

The choice of CO[J=2$\rightarrow$1] over CO[J=1$\rightarrow$0] is driven by the fact that, while CO[J=1$\rightarrow$0] might be accessible with e.g. VLA observations, the high frequencies of the lines would make it very expensive. On the other hand the remarkable technical capabilities of NOEMA, allow us to use CO[J=2$\rightarrow$1] and recover the necessary observations in a matter of a few hours. While more luminous, higher J lines e.g. CO[J=3 $\rightarrow$ 2] do not offer a substantial decrease in the integration time to achieve the same S/N, due to a fall in sensitivity at higher frequencies. Moreover, CO[J=2$\rightarrow$1] is an excellent tracer of the total gas mass, that only requires a minor excitation correction (e.g. Daddi et al. 2015). For the dust-based $M_{\rm gas}$ estimates we will employ the dust to gas mass ratio

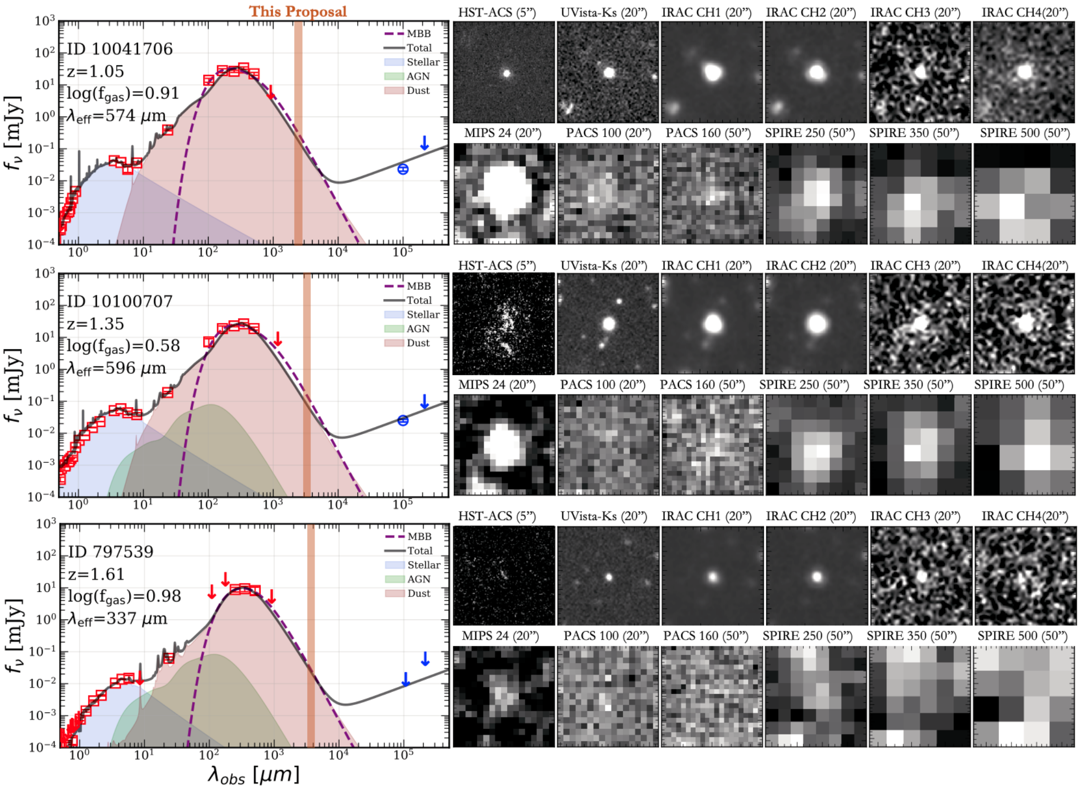

Figure 6.2: Three 'gas-giant' candidates at $\langle \mathbf{z} \rangle = \mathbf{1.3}$. **Left:** Best fit multi-component template spectrum (grey line and shaded areas) + an MBB model fit (purple line), the observed photometry (red squares), including upper limits (red arrows) are also displayed. The wavelength of the proposed observations is highlighted by the shaded orange region. Radio measurements are also overlaid in blue. The SED is separated into three components - stellar emission (blue), AGN emission (green), dust (red). **Right:** Optical-NIR-FIR cutouts of these objects, ranging from 5" in the optical-MIR range to 50" in FIR.

($\delta_{\rm GDR} - Z$, e.g. Magdis et al. 2012a) and the R$-$J monochromatic flux density method (e.g. Scoville et al. 2016, 2017). The final goal of this proposal is to obtain an independent gas mass tracer in the form of CO[J=2$\rightarrow$1] + the underlying R-J continuum at 1.2 mm rest-frame to unambiguously confirm the presence of extreme gas reservoirs in our proposed sample of gas giants.

A pilot subsample of the most robust 'gas-giant' candidates has been accepted to be observed by NOEMA. The data has already been taken, and I currently await its reception, after which I aim to quickly reduce and analyse it to determine whether the $M_{\rm dust}$ derived gas masses are indeed correct. If this is the case, the next aim will be to apply for more follow-up observations, with ALMA and NOEMA, targeting alternative $M_{\rm gas}$ tracers, as well as continuum deep in R-J tail, as it would be required to understand these sources. Discerning why these objects can proceed directly to the quenched mode is key towards understanding cessation of star formation through gas stabilisation, and not just as a consequence of gas depletion or quasar activity (see Man & Belli 2018).

6.2 Exploring the Products of ALCS Catalogue

6.2.1 Optically Dark Galaxies

By studying the ALCS catalogues and mosaics I have identified a population of objects that are detected in IRAC, but lack any counterpart in the UV to the H-band. More specifically, out of 180 S/N>4 ALMA detected objects, there are a total of 35 which do not have a counterpart in the CHArGE/`golfir` catalog. Further examination of the cutouts in all available filters revealed that, 3/35 objects are located next to a bright star or the edge of the map, and 10/35 fall outside *HST* footprint. In total, we were left with 22 confirmed '*HST*-dark' candidates, with secure ALMA detections. Within this sample, only 14 have a secure IRAC detection, while the remaining 8 are 'IRAC-dark'.

The lensed fields that comprise ALCS contain wealth of ancillary data, which also includes multiple *Herschel* exposures which overlap with our sources. To take advantage of the ancillary data, we have used the photometric catalog provided in the upcoming paper from the ALCS team (Sun et al. 2022, submitted), where *Herschel* fluxes have been measured for all 180 S/N>4 ALMA detected galaxies. The *Herschel* measurements are not deep enough to detect the majority of our galaxies in all filters, however marginal detections and upper limits can still provide strong constraints during SED fitting.

Optically dark objects have been shown to contribute ten times more than equivalently massive UV-bright galaxies to the cosmic SFR density at $z>3$ (Franco et al. 2018; Wang et al. 2019), and as such constitute excellent candidate progenitors of the most massive galaxies presently located in the center of massive groups and galaxy clusters. While the '*HST*-dark' galaxies pose another technical challenge to the correct application of the energy balance approach, the `Stardust` SED fitting software, described in Chapter 2, was built specifically for this purpose, and can assist in performing a robust extraction of infrared luminosities and dust masses.

The IRAC photometry extraction method, presented in Chapter 3, depends on the availability of an *HST* prior. Due to the sources being optically dark however, an optimal model photometry with `golfir` is not possible. In order to extract the IRAC fluxes we have placed D=3."0 apertures on the IRAC `golfir` residual images, centred on the coordinates of ALMA sources. The aperture to total flux conversion was then derived from the IRAC instrument handbook. Using all of the available IRAC, *Herschel* and ALMA 1.2 mm photometry we have fit all 14 robust '*HST*-dark' candidates. For the purposes of this analysis we did not use the DL07 IR template library, due to the lack in the photometric coverage. Instead, a template of local ULIRG Arp 220 (Rowan-Robinson & Crawford 1989), and a combined best-fit template for the all 707 AS2UDS SMGs, presented in Dudzevičiūtė et al. (2020), were used. From the best-fit SEDs we have extracted the total infrared luminosities, and the IR - based photometric redshifts. In Figure 6.3, I present a set of cutouts, and

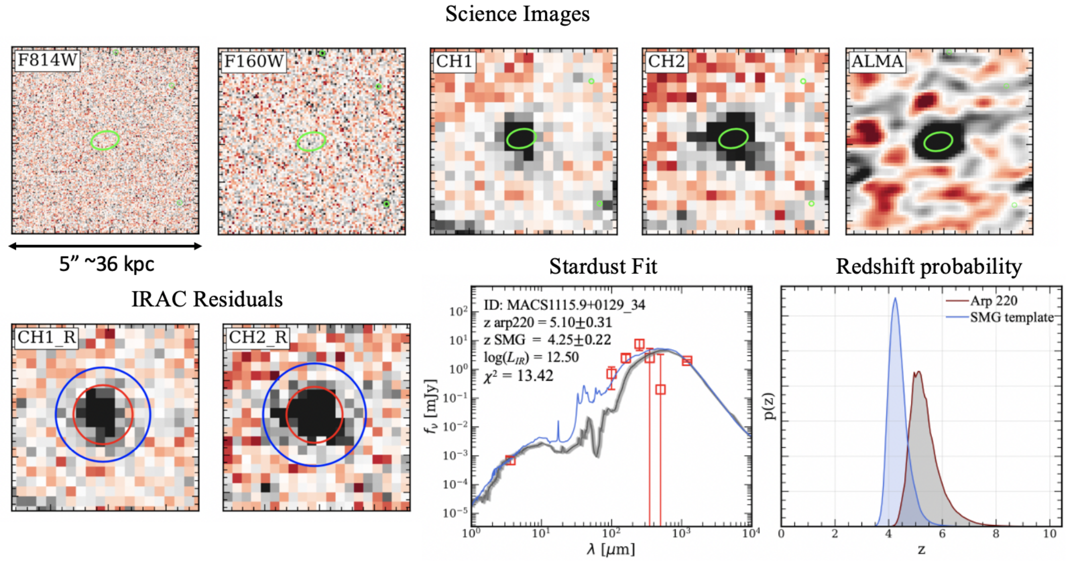

Figure 6.3: Top: Science images showing the *HST*/F814W and F160W filters, IRAC 3.6 μm and 4.5 μm, plus ALMA 1.2 mm continuum. The green ellipse corresponds to the ALMA beam size. **Bottom Left:** IRAC residual images from `golfir` on which aperture photometry was performed. Red circle denotes the D =3."0 apertures, while the blue circle shows the annulus from which we have computed the uncertainty. **Bottom Right:** Best-fit `Stardust` SEDs to the IRAC, *Herschel* and ALMA photometry. We also show the redshift probability distributions ($p(z)$) which arise from using two different templates.

a `Stardust` best-fit SED to an optically dark source with the largest number of detections in the sample. From the SED fitting I find that the 14 objects have a moderately tight distribution of $L_{\rm IR}$ - and redshifts. Based on the χ^2 analysis I also find that the majority of our galaxies are better fit by the Dudzevičiūtė et al. (2020) template, rather a local ULIRG. I find that the sample has a median log($L_{\rm IR}/L_{\odot}$)=12.43$^{+0.25}_{-0.10}$ and a median photometric IR redshift of zphot$_{\rm IR}$=3.23$^{+0.76}_{-1.11}$, with uncertainties derived from the 16th and 84th percentiles.

Similar objects have been detected and described before, for example in their work Wang et al. (2019), present a sample of 39 H-band dropouts, which have been detected by ALMA at $S_{870\mu{\rm m}}$>870 mJy. The best fit SED fit to the stacked sample reveals an average log($L_{\rm IR}/L_{\odot}$)=12.34 $\pm$ 0.07, with a median $z \sim$ 3.5, which overlaps well with our sample. The results of Wang et al., predict that the number density of optically dark galaxies is approximately equal to 5.3×10^2 deg^{-1}, therefore given the total ALMA coverage of ALCS of 134 arcmin^2, we should therefore expect to find $\sim$ 19 such sources. The recovered 14 objects align well with the predicted number densities, with the discrepancy possibly arising due to the fact that ALCS is non-contiguous, and is thus more prone to the effects of the cosmic variance. The fact that these objects are detected by ALMA securely ensures that the majority of these *HST* dropouts are dusty SFGs, at high-z. This picture is consistent with their IRAC and *Herschel* photometry. Admittedly, the photometric redshifts derived by Stardust are highly uncertain, as the optical photometry is entirely missing.

The first step towards investigating the properties of these objects would be a secure redshift confirmation. This can be achieved relatively cheaply with, for example, a line scan with either ALMA or NOEMA. For that we can target the CI[2-1] and CO (7-6) simultaneously, at $\lambda_{\mathrm{obs}} \sim 1400$ μm, with two tunings of ALMA band 5/6. These objects are sufficiently bright to achieve a satisfactory SN threshold in relatively short times, especially for Herschel detected objects like MACS1115.9+0129_34 in Figure 6.3. As introduced in Chapter 1, M_{gas} , f_{gas} and especially ρ_{gas} and beyond $z > 2.5$ are notoriously difficult quantities to constrain. At high-z, these parameters have mostly been measured for bright, optically detected objects, which can potentially omit the substantial contribution towards the total dust/gas mass density coming from the optically obscured galaxies. Moreover, metallicity is most commonly computed from the optical lines, which are unavailable for this unique set of galaxies. However, recent works have revealed that it is also possible to derive metallicity from the IR lines instead (Rigopoulou et al. 2018). The recent analysis focusing on the metallicity measurement in ULIRGS (Chartab et al. 2022), reveals a ~ 0.3 dex under-estimate of metallicites derived form optical lines, compared to the FIR method. Therefore if we believe that optically dark galaxies are elevated above the main sequence, deriving the metallicity with IR lines, such as [OIII] 52 μm, [OIII]88 μm, and [NIII]57 μm would be preferable to the optical method, even if they were available.

The results of Wang et al. (2019) suggest the density of ~ 530 such objects per square degree, indicating that our understanding of massive SFGs in the early Universe is incomplete. With that, the recent launch of *JWST* offers some additional opportunities to investigate *HST* dark populations. The spectroscopic follow-ups with the NIRCAM and MIRI instruments on-board *Webb* will offer a whole new window on the population of H-band dropouts, and provide us with further insights regarding their nature.

6.2.2 High-z Follow Ups

In addition to the sample of H-band dropouts within ALCS, I will focus on studying a list of newly identified $z = 8 - 10$ candidates. These objects are completely novel, and have been discovered on top of the ones that have already been found in these fields (e.g. see Laporte et al. 2017; Hashimoto et al. 2018; Tamura et al. 2019; Strait et al. 2021; Laporte et al. 2021; Fujimoto et al. 2021). The redshifts have been calculated via SED fitting and after a careful examination of both the cutouts and best fit solutions I have isolated a robust sample that can form a basis for future proposals with JWST and ALMA. One example of a $z = 9.17$ candidate is shown in Figure 6.4. With these I will be able to study the gas and dust evolution at the frontier of the epoch of reionisation. The available ALMA maps in ALCS cover the 1.2 mm emission, which at $z = 8 - 10$ translates into the $\sim$ 100-130 μm rest frame regime. If the T_{dust} - z evolution trends continue into the high-z Universe (e.g. Béthermin et al. 2015; Cortzen et al. 2020), the peak of dust emission should shift towards $\sim 70 - 80$ μm in the rest-frame. This means, that

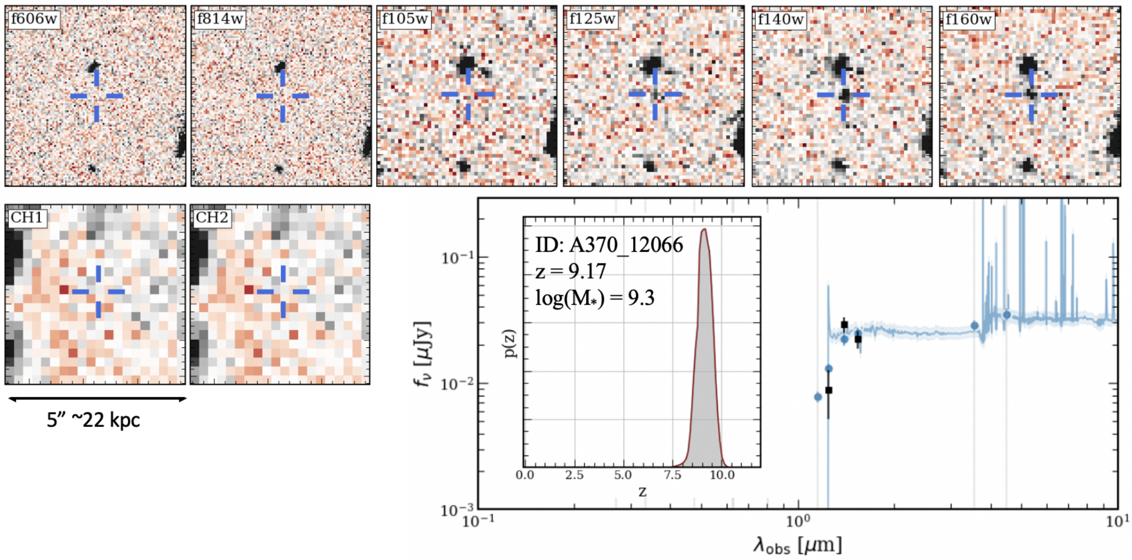

Figure 6.4: Candidate LBG in the Abell 370 cluster of the Hubble Frontier Fields. **Top:** The 5" cutouts of the *HST*/ACS and *HST*/WFC3 filters. A clear dropout in visible in the F105W filter, indicating a z∼9 galaxy. **Bottom Left:** IRAC 3.6 μm and 4.5 μm cutouts. The absence of the source in the *Spitzer* images helps to identify the galaxy as a dropouts, rather than an dust obscured low-z source. **Bottom Right:** Best fit EAZY SED and the redshift probability distribution.

depending on the redshift of our LBG candidate, the 1.2 mm ALMA upper limits can be potentially used to provide constraints on the upper limits of either $L_{\rm IR}$ or the dust/gas mass at high-z. To confirm the redshift for these LBGs, I aim to follow up this population with the ALMA band 5, specifically targeting the [CII] emission line at $\lambda_{\rm rest}$ 158 μm. Observations of [CII] should allow for independent verification of both the dust obscured SFRs, as well as gas masses (Zanella et al. 2018), provided that the low-z trends still hold at high-z.

6.2.3 Red Quiescent Objects

The study of the dust and gas content of massive red galaxies, responsible for the bending of the main sequence, is instrumental to understand how these objects transition from the peak of their activity towards quiescence. In particular, a growing body of evidence shows that quiescent galaxies at $z > 1.5$ might retain significant dust reservoirs after quenching (e.g. Ciesla et al. 2021; Magdis et al. 2021; Whitaker et al. 2021). For this I will aim to characterise the far-IR emission of a mass-complete sample of massive *HST*+IRAC-selected "red" galaxies individually and by stacking) in order to (i) broadly distinguish strongly starbursting galaxies from bona fide quiescent objects; (ii) cross-correlate the presence of substantial dust pockets with galaxy properties (stellar mass, SED-based SFRs, star formation histories, environment); (iii) derive dust (and gas) masses for the quiescent objects, constraining the quenching mechanisms (gas expulsion vs gas stabilisation). Thus bringing us full circle back to the "gas-giants".

CHAPTER 7

Conclusion and Final Remarks

In this final chapter I would like to present a summarised view of the main findings and conclusions from my thesis. The majority of my work is focused around the studies of the interstellar medium, its interaction and effects on the galaxy environments, and its evolution with time.

In Chapter 2 I have introduced a novel SED fitting algorithm - `Stardust`. The creation of this tool arose from the necessity to independently, yet simultaneously, process the UV-optical and IR photometry, without having to rely on generalised assumptions regarding the distribution of dust and stars in galaxies. I then applied my code to the deepest available photometric catalogues in the GOODS-N and COSMOS field, and recover properties of both the stellar populations and the ISM for $\sim 5,000$ galaxies, spanning a wide redshift range $0 < z < 5$. My analysis has revealed that both dust and gas to stellar mass ratios steeply rise up to $z = 2$, mirroring the SFRD evolution found in the literature (Madau & Dickinson 2014; Casey et al. 2021), and then remain roughly flat. I find that the recovered trends agree with the existing studies, thus verifying the reliability of `Stardust`. Moreover, the scaling relations which I recovered are one of the most statistically significant examinations of the ISM evolution of SFGs to date, and can be used to guide simulations and telescope proposals.

In addition to the tool and the catalogue containing galaxy properties I also report a discovery of a new galaxy population, whose extreme gas mass reservoirs seem to defy the current observational and theoretical predictions. These objects could have a potentially significant impact on the way in which we look at the cessation of star formation activity through gas stabilisation, and not just as a consequence of gas depletion or quasar activity (see e.g. Man & Belli 2018), however follow up line observations are imperative to confirm these extreme gas masses.

In Chapter 3 I present a uniform reprocessing of all the available archival *HST* and *Spitzer* data covering the ALCS fields. ALCS is an ALMA large survey, which covers 33 of the best studied lensed fields at 1.2 mm, over an area of 134 arcmin^2. The finalised catalogues that I present contain the deepest and most complete photometry of those fields to date, spanning 15 bands from UV to the mm regime for $\sim 218,000$ sources. In addition to that, the wealth of the photometric coverage achieved through the combination of the data from *HST* and *Spitzer* allowed me to robustly constrain the stellar population properties of these objects, such as their photometric redshifts, M_* and SFRs. The release of these data products, in conjunction with their physical parameters derived via SED fitting, will provide the scientific community with a powerful tool to examine the nature of faint sub–mm galaxies and line emitters, both in the field and within the clusters themselves. Additionally, through the redshift distribution, SFR, stellar and ISM masses of sources that only ALMA can locate and study, we will probe and constrain the evolution of the molecular gas mass density up to the peak epoch of cosmic star formation.

My catalogues will also aid in the identification and further examination of the optically dark populations (see e.g. Wang et al. 2019), which host the majority of star formation at $z > 4$ and high-z galaxies. In fact, with the aid of `Stardust` SED fitting I have already identified a few such "*HST* - dark" galaxy candidates, which I aim to follow up with the state-of-the-art facilities such as Keck/MOSFIRE, ALMA and subsequently *JWST*. Moreover, a careful examination of the EAZY derived photometric redshifts, and SEDs, yields a significant number of potential $z > 9$ LBGs candidates, which I also aim to target in my future telescope proposals in an attempt to elucidate the key questions about the early onset of star formation, reionisation and assembly of first galaxies.

In Chapter 4 I present my ongoing work aimed at understanding and assessing the detection feasibility of various galaxy populations with *JWST*/MIRI as a part of the COSMOS - Webb program. I aim to provide the scientific community with a set of catalogues containing simulated photometry, number counts and mock images of the entire area covered by MIRI. The mock catalogues are built upon the existing wealth of multi-wavelength photometric data in the COSMOS field, utilising two different spectral energy distribution (SED) fitting approaches - EAZY and `Stardust` to simulate the expected fluxes. I predict that, from the already detected galaxies in the COSMOS field, MIRI observations are expected to recover 30,000 - 50,000 sources at a 5σ level corresponding to the 24.7 mag AB, depending on the chosen simulation. From the predicted photometry, I modelled the full area included in the COSMOS - Webb survey, and created a set of mock images. However, the COSMOS 2020 catalogue is complete only down to its selection limit, derived from the ground based optical bands, thus preventing detection of very distant and/or dusty galaxies. More specifically, within the area covered by MIRI we expect to find $\sim$ 100 optically dark galaxies (e.g. see Wang et al. 2019), which can not be modelled by using any existing priors. To alleviate that, I aim to use the existing scaling relations describing the '*HST*-dark' sources, in conjunction with the phenomenological models (e.g. JAGUAR Williams et al. 2018) to explore the full dynamic range and potential of the MIRI observations in the COSMOS field. Tracing the optically obscured population is imperative towards understanding the contribution of DSFGS to the galaxy assembly beyond $z > 4$. This would become especially prominent, considering the upcoming MORA and its extension (Ex-MORA) large ALMA surveys at 2mm (Casey et al. 2021). With the mock mosaics we aim to find and optimise the critical source detection and photometry extraction algorithms, in order to facilitate a rapid and robust data processing.

In my future research plan outlined in Chapter 6 I aim to establish a complete and coherent picture of the star-formation activity through cosmic time, assisted by the work outlined in the previous chapters. To achieve this I will address the following questions: How do the fractions of gas and dust of galaxies evolve with time? How does this evolutionary picture change when switching from one mode of star formation to another? Why do some galaxies form stars more efficiently that the others? Why do galaxies die? To address these questions I will look at galaxies in all three stages of their evolution.

I will begin by exploring a tantalising population of "gas-giants, as outlined in Chapter 2. These objects are located on the main-sequence, however display significantly elevated depletion timescales, given their SFRs and M_*. To this end, I have successfully applied for a pilot proposal with NOEMA, targeting the CO(2-1) line in a sample of 'giants'. I will aim to quickly reduce these data, and if indeed the gas masses are confirmed to be true, I will propose for follow-up observations targeting more objects, and also alternative $M_{\rm gas}$ tracers, as well as continuum deep in R-J tail, as it would be required uncover the nature of these mysterious sources.

Following my work in Chapter 3, I will explore the emergent products of the ALCS catalogues. Moving above the main sequence towards more extreme star - formation, I will examine the potential "*HST*-dark" sources. The objects that I find are similar to the ones already reported in the literature (Franco et al. 2018; Wang et al. 2019), which have been shown to contribute ten times more to the SFRD at $z > 3$, than similarly massive UV-bright galaxies. These therefore can be potentially thought of as candidate antecedents to the massive ellipticals, currently located at the centres of massive galaxy clusters. Moving towards the quenched regime, I will study the dust and gas content of massive red galaxies detected in ALCS. The existing wealth of the optical+NIR+mm data will be instrumental towards understanding how these objects transition from the peak of their activity towards quiescence. This is especially important, given the growing body of evidence (Ciesla et al. 2021; Magdis et al. 2021; Whitaker et al. 2021), showing that quenched galaxies at $z > 1.5$ seem to retain their gas mass reservoirs, potentially linking these red objects to the "gas-giants". Finally, I aim to follow up and characterise the LBG candidates at $z > 9$.

The work presented in this thesis contributes to expanding our understanding of the evolution of dust and gas fraction evolution in star-forming galaxies from now and up to $z = 5$. The tools, images, catalogues and simulations produced during my work are made available to the wider scientific community, and can be used to conduct the search and analysis of new and unique galaxy populations.

List of Publications

Source (ADS) (May 22, 2022)

published: 16 — first author: 1 — citations: 250

16. Sun, F. et al., *ALMA Lensing Cluster Survey: ALMA-Herschel Joint Study of Lensed Dusty Star-Forming Galaxies across $z \simeq 0.5-6$* ApJ, Accepted 2022 (arXiv:2204.07187)
15. Wang, T. et al., *A^3COSMOS: A census on the molecular gas mass and extent of main-sequence galaxies across cosmic time* A&A, 660:142, 2022 (arXiv:2201.12070)
14. Fujimoto, S. et al., *A dusty, compact object bridging galaxies and quasars at cosmic dawn* **Nature**, (arXiv:2204.06393)
13. Gómez-Guijarro, C. et al., *GOODS-ALMA 2.0: Compact star formation as a sign of self-regulation before quenching* A&A, 659:196, 2022 (arXiv:2201.02633)
12. Weaver, J. et al., *COSMOS2020: A panchromatic view of the Universe to $z \sim 10$ from two complementary catalogs* ApJS, 258, 1:11, 2022 (arXiv:2110.13923)
11. Casey, C. et al., *Mapping Obscuration to Reionization with ALMA (MORA): 2mm Efficiently Selects the Highest-Redshift Obscured Galaxies* ApJ, 923 215, 2021 (arXiv:2110.06930)
10. **Kokorev** et al., *The Evolving Interstellar Medium of Star-Forming Galaxies, as traced by Stardust* ApJ, 921 40, 2021 (arXiv:2109.06209)
9. Valentino et al., *The effect of active galactic nuclei on the cold interstellar medium in distant star-forming galaxies* A&A, 654, 165:19, 2021 (arXiv:2109.03842)
8. Fujimoto et al., *ALMA Lensing Cluster Survey: Bright [CII] 158 μm Lines from a Multiply Imaged Sub-L_* Galaxy at $z = 6.0719$* ApJ, 911, 99, 20, 2021 (arXiv:2101.01937)
7. Puglisi et al., *Sub-millimetre compactness as a critical dimension to understand the Main Sequence of star-forming galaxies* MNRAS, 508, 4:5217-5238, 2021 (arXiv:2103.12035)

6. Magdis et al., *The interstellar medium of quiescent galaxies and its evolution with time* A&A, 647, 33, 2021 (arXiv:2101.04700)

5. Valentino et al., *CO emission in distant galaxies on and above the main sequence* A&A, 641, 155, 2020 (arXiv:2006.12521)

4. Steinhardt et al., *The BUFFALO HST Survey* ApJS, 247, 64, 2020 (arXiv:2001.09999)

3. Valentino et al., *The Properties of the Interstellar Medium of Galaxies across Time as Traced by the Neutral Atomic Carbon [CI]* ApJ, 890, 24, 2020 (arXiv:2001.01734)

2. Cortzen et al., *Deceptively cold dust in the massive starburst galaxy GN20 at $z \sim 4$* A&A, 634, 14, 2020 (arXiv:2002.02974)

1. Cibinel et al., *Early- and late-stage mergers among main sequence and starburst galaxies at $0.2 \leq z \leq 2$*, MNRAS, 485, 5631, 2019 (arXiv:1809.00715)

APPENDIX A

Appendix A

A.1 SED Fitting

A.1.1 Draine & Li (2007) Templates

In our fitting routine we utilise the dust models of Draine & Li (2007), with the updated opacity from Draine et al. 2014. These models aim for a robust and physically motivated approach to SED fitting in both MIR and FIR, as well as allow us to calculate the amount of luminous dust.

The description of the dust locked in the interstellar medium is one of a mixture of carbonaceous and amorphous silicate grains, with their sizes and distributions following the extinction law in the Milky Way, Large Magellanic Cloud (LMC) and the bar region of the Small Magellanic Cloud (SMC). The carbonaceous grains behave similarly to the polycyclic aromatic hydrocarbon (PAH) molecules, with their properties given by the PAH index $q_{\rm PAH}$, that is defined as a fraction of dust mass locked into the PAH grains.

The models provide a bimodal description of the environments containing the interstellar dust: the diffuse ISM and the PDRs. The bulk of the dust mass is thought to be located in the cold and diffuse part of the interstellar medium, that is being heated by a radiation field of a constant intensity $U_{\rm min}$. A smaller proportion of the mass budget described by the γ index is exposed to a gradient of radiation intensities ranging from $U_{\rm min}$ to $U_{\rm max}$, and is supposedly located in the warmer PDRs. Although these warm regions normally contain only a small fraction of the total dust mass, they can make a substantial contribution to the luminosity in the mid-IR SEDs. As described by DL07, the infinitesimal proportion of $dM_{\rm dust}$ exposed to radiation fields between U and $U + dU$ can be modelled by a power law distribution, and in the case of the diffuse ISM where $U_{\rm min} = U_{\rm max}$, by a Kronecker δ-function. This leads to the following description:

$$dM_{\rm dust} = (1-\gamma)\,\delta(U - U_{\rm min}) + \gamma\, M_{\rm dust}\frac{\alpha - 1}{U_{\rm min}^{1-\alpha} - U_{\rm max}^{1-\alpha}} \qquad \text{(A.1)}$$

with $U_{\rm min} \leq U_{\rm max}$ and $\alpha \neq 1$. The parameter - γ - is the fraction of dust mass locked into the high starlight intensity regions described by the power-law, α gives the distribution of radiation intensities in the PDRs and $M_{\rm dust}$ as the total dust mass.

The methods described in DL07 allow us to compute a distribution of temperatures for all particles: ones that are small so their size makes them susceptible to the effects of quantised heating, and the larger ones where the steady-state temperatures dictated by the stellar radiation and radiative cooling equilibrium take hold. One can compute an averaged IR emission for a given grain type by first considering their temperature distribution and cross sections, and then sum everything up to obtain the specific mass weighted power that is being radiated by the dust exposed to starlight of intensity U. By integrating these numerical recipes from $U_{\rm min}$ to $U_{\rm max}$, one can obtain the power per unit frequency per unit mass $p_\nu(q_{\rm PAH}, U_{\rm min}, U_{\rm max}, \alpha)$.

In line with DL07, one can then model the galaxy spectral energy distribution as a linear combination of the diffuse ISM and the PDRs. This can be written as follows:

$$j_\nu = (1-\gamma)\, j_\nu[U_{\rm min}, U_{\rm max}] + j_\nu[U_{\rm min}, U_{\rm max}, \alpha] \tag{A.2}$$

where j_ν is the emissivity per hydrogen nucleon. If one now considers a galaxy at some distance $D_{\rm L}(z)$, the received flux density can be written as:

$$f_\nu = \frac{M_{\rm H}}{m_{\rm H}} \frac{j_\nu}{D_{\rm L}^2(z)}. \tag{A.3}$$

Since j_ν is the quantity contained in the DL07 models, the normalisation extracted from the fitting represents the total number of hydrogen nucleons, and can be then converted to the luminous dust mass.

The total luminosity contained in both dust components can be written as:

$$L_{\rm dust} = \langle U \rangle P_0 M_{\rm dust} \tag{A.4}$$

with $\langle U \rangle$ representing a mean intensity of the radiation field, given by:

$$\langle U \rangle = (1-\gamma)\, U_{\rm min} + \frac{\gamma \ln\left(U_{\rm max}/U_{\rm min}\right)}{U_{\rm min}^{-1} - U_{\rm max}^{-1}}, \tag{A.5}$$

for $\alpha = 2$ and P_0 denoting the power absorbed per unit dust mass in a radiation field of intensity $U = 1$.

In principle, one could think of these models as having six effective free parameters - $q_{\rm PAH}$, $U_{\rm min}$, $U_{\rm max}$, α, γ and $M_{\rm dust}$- acting as the normalisation as described above. It has been shown in Draine et al. (2007) that the parameter space is insensitive to the adopted dust model (MW, LMC and SMC) and the values of α and $U_{\rm max}$. It is possible thus to recover a wide range of properties of various SEDs by fixing these to $\alpha = 2$ and $U_{\rm max} = 10^6$. The values of $U_{\rm min}$ below 0.7 correspond to temperatures below 15 K, and while it is expected to find very few systems that exhibit this behaviour, we have decided not to limit the range of $U_{\rm min}$ and allow it to vary between 0.1 and 50, to capture even the most extreme cases. It has been shown that at least in the case of local galaxies in the Spitzer Nearby Galaxy Survey (SINGS), the $U_{\rm min}$ can be limited between $0.7 \leq U_{\rm min} \leq 25$, and a MW - like dust model can be adopted to limit $q_{\rm PAH}$ between 0.004 and 0.046. Incorporating the use of the optimised set of parameters, which has been done in similar studies (e.g., Magdis et al. 2012a; Magnelli et al. 2012; Santini et al. 2014), might have a positive effect on computational speeds, however we need to consider that this reduced parameter space has only been robustly verified for nearby solar-like metallicity populations of galaxies and might otherwise risk under(over)estimating the dust masses for extremely cold(warm) systems.

We can thus extract the following physical parameters from the fit - γ, $q_{\rm PAH}$, $U_{\rm min}$. The dust mass is simply computed from the normalisation, while the $L_{\rm IR}$ is L_ν integrated over the 8 - 1000 μm range. As an additional parameter, we can obtain $\langle U \rangle$ by either utilising the Equation A.5, or alternatively, as prescribed by Magdis et al. (2012a), we can use Equation A.4, where we set $L_{\rm dust} = L_{\rm IR}$ and $P_0 \approx 125$.

A.1.2 Removing AGN contamination

Active galactic nuclei (AGN) can have a significant impact on the ISM of galaxies that host them. They possess an ability to halt star-formation by heating up the gas and dust or completely quenching the galaxy by stripping away its fuel. Under the common assumption (Antonucci 1993; Urry & Padovani 1995; Tristram et al. 2007), AGN are surrounded by dusty tori, that similarly to the ISM dust, can absorb the UV/optical light from the AGN and re-radiate it at redder wavelengths, normally peaking in the mid-infrared (MIR) regime at 20-50 μm. It has also been shown that for select extreme cases (Mullaney et al. 2011), the AGN emission dominates the SED of a galaxy even at 60 μm, which presents a new challenge when calculating an infrared luminosity of a source. Infrared derived star-formation rate estimates rely on a robust understanding of the $L_{\rm IR}$. Therefore, it is imperative to separate the energy contributions from hot dust in the ISM and a possible AGN.

In order to account for the effects of the IR contamination by AGN when calculating $L_{\rm IR}$, as well as to identify all the possible systems that might contain an active nucleus in our sample, we have decided to adopt a set of AGN templates from Mullaney et al. (2011) (M11). These templates have been empirically derived by assuming a modified blackbody function and fitting it to a set of Swift-BAT AGNs as well as IRAS spectra. The obtained models describe intrinsic AGN emissions in the range spanning from MIR to FIR (6-100 μm). In this case, a typical AGN SED could be thought of as a broken power law at $\leq 40 \mu m$, that rapidly vanishes when moving above 40 μm.

The average intrinsic AGN emission can be described as follows:

$$F_\nu = \begin{cases} \lambda^{1.8} & \text{at} \quad 6\,\mu\text{m} < \lambda < 19\,\mu\text{m} \\ \lambda^{\alpha} & \text{at} \quad 19\,\mu\text{m} < \lambda < 40\,\mu\text{m} \\ \nu^{1.5} F_\nu^{BB} & \text{at} \quad \lambda > 40\,\mu\text{m} \end{cases} \tag{A.6}$$

where ${F_\nu}^{BB}$ is the modified blackbody function and α is the spectral index. In our procedure we utilise the high and the low luminosity templates, with the $\alpha = 0.0$ and 0.4 respectively. In addition to that, we allow for a linear combination between the two, with varying coefficients, thus expanding the existing template space.

A.1.3 Stellar emission component

The contribution of stellar emission in the observed NIR bands, such as $Spitzer$ IRAC 1-4, can be quite significant, especially when we move to higher z. Therefore, if no libraries representing the luminosity from dust attenuated stellar light are available, one might either underestimate the slope of the AGN or overestimate its normalisation. This can lead to erroneously assigning more luminosity to the AGN component in the MIR and therefore underestimating the L_{IR}.

To avoid this, we additionally incorporate a library of stellar emission models, which are an updated version of the templates described by Brammer et al. (2008) (GB08). These are based on the stellar population synthesis models (Conroy et al. 2009), that were optimised for deep optical-NIR broadband surveys. The models form a basis set of a larger library, and were derived by using the 'non-negative matrix factorization' algorithm that was described in Blanton & Roweis (2007). The method attempts to reproduce the full library of templates, by finding a non-negative linear combination of a smaller number of models. These can be considered as the 'principal component' blueprint of the larger catalogue. In total there are 12 optical SEDs, which include both dust attenuated and non-attenuated starlight. We have incorporated these models into our fitting routine, and if the UV-optical photometry is available, this allows us to extract properties such as M_*, SFR, $E(B-V)$ and SFH. If no UV-optical data is available, the addition of these templates is still useful, as they can account for the excess flux in the rest-frame NIR bands, in conjunction with AGN and IR templates.

To test how our fits behave without the stellar component, we have isolated ~ 100 objects with $z > 2$, so that our bluest available band traces $\lambda_{\text{rest}} \sim 1\,\mu$m, where the contribution from stellar emission becomes non-negligible. We then exclude GB08 templates and re-fit our objects. This results in two outcomes that can be seen in Figure A.1. We find that, by removing the additional component, we tend to either overestimate the AGN contribution (blue squares), with the median being ~ 1.5, or alternatively erroneously assign a galaxy to contain an AGN (red

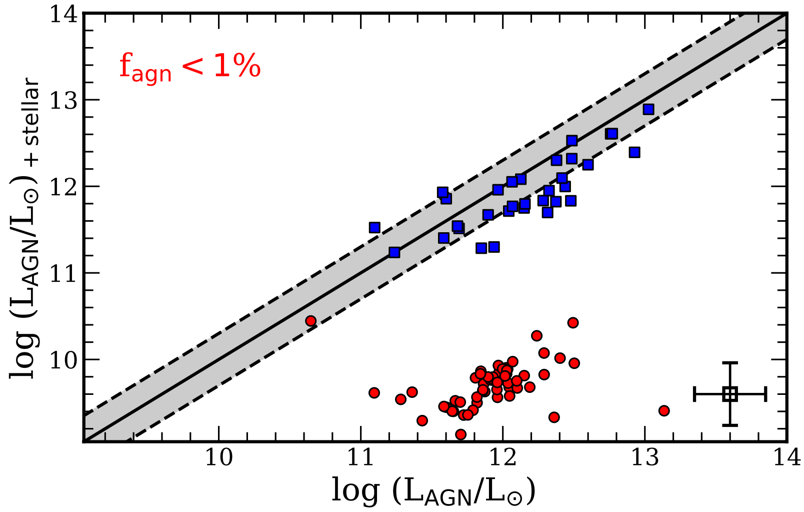

Figure A.1: A comparison between fits that incorporate the stellar component, and those that do not for $z > 2$ galaxies. Blue squares and red circles represent objects which were assigned $f_{\rm AGN} > 0.01$ and $f_{\rm AGN} < 0.01$ respectively, by the means of our three component fit. The median uncertainty for both quantities is in the bottom right corner. The black solid line represents the 1:1 relation and the grey regions cover the 0.3 dex range.

points).

A.1.4 Bringing it all together

In order to model an SED of a galaxy and extract the physical parameters, we first transport all three components - the stellar, the dusty torus AGN and the infrared dust emission to a common wavelength grid spanning the range from 10^{-4} to 10^5 μm. Our method relies on a linear combination of these models, thus making it imperative for them to share a common range, so that co-adding them is made possible. In certain cases where this grid falls outside the original range of the template, we extrapolate bluewards and redwards by using a steeply declining power law. This was done to ensure that the resultant galaxy emission is continuous without any sudden breaks, which could interfere with the fitting, where one of the templates has ran out of range. These added power laws do not introduce any additional emission, as the flux density contribution from them is orders of magnitude lower than that of the original template. We then redshift the wavelength grid on a per galaxy basis and normalise the templates to ensure that they are not separated by tens of orders of magnitude. In Figure A.2 we show all the templates used in `Stardust`, normalised to the K-band, and in Table A.1 we list all the relevant parameters of the models.

Subsequently, we perform a synthetic photometry on all three components separately in all

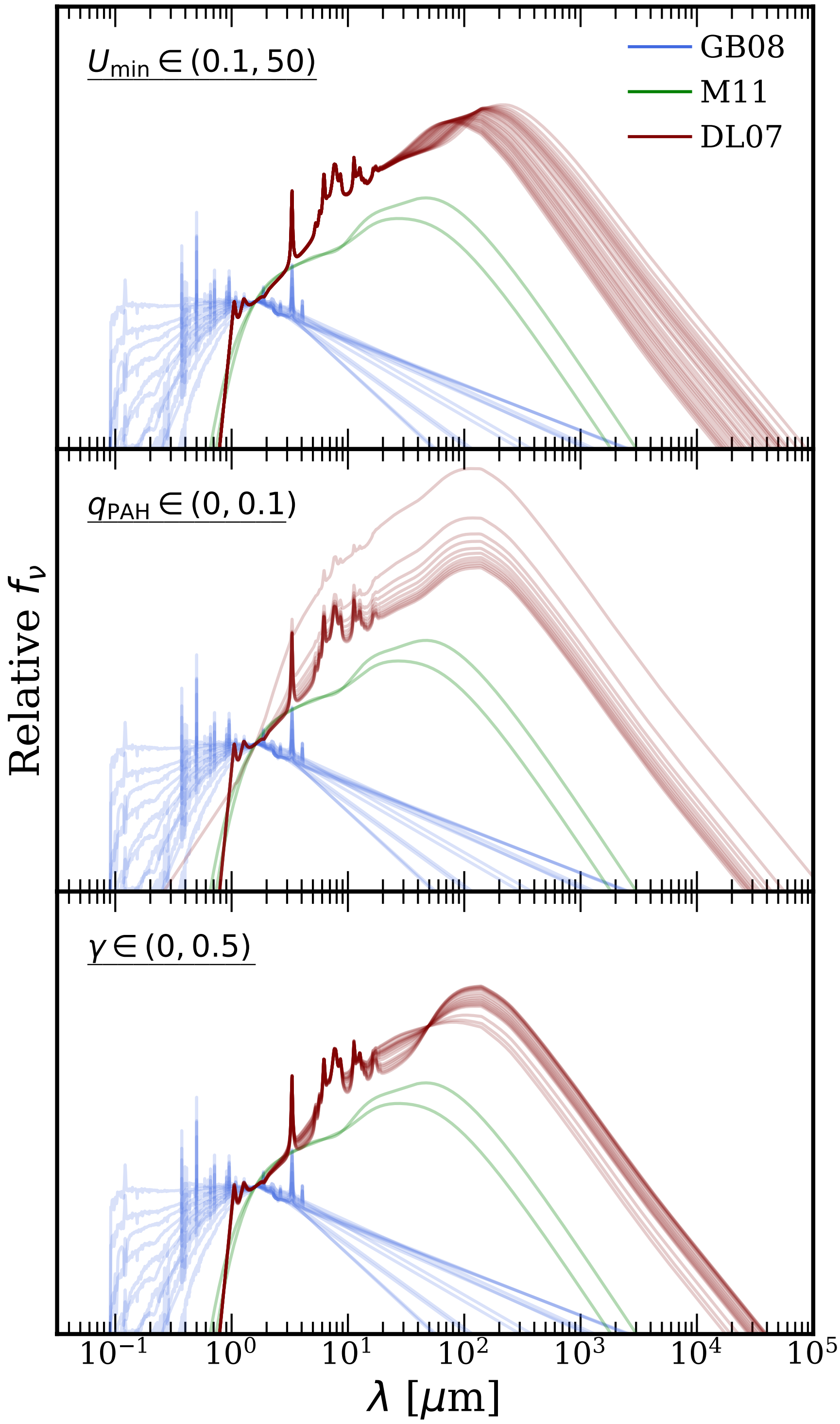

Figure A.2: A compilation of all template flavours used in `Stardust`. Colour-coding is blue, green and maroon for Brammer et al. 2008 UV-Optical SPS models, Mullaney et al. 2011 AGN models, and Draine et al. 2007 IR dust models respectively. We show the variations in U_{min}, q_{PAH}, and γ separately on each panel, while other parameters are fixed. For visualisation purposes all templates shown here have been normalised to the K-band.

Table A.1: Template parameters used for `Stardust` fit.

Parameter	Value
Dust emission: DL07 (Updated in Draine et al. 2014)	
U_{min}	$[0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8,$ $1.0, 1.7, 2.0, 3.0, 4.0, 5.0, 7.0, 8.0,$ $10.0, 12.0, 15.0, 20.0, 25.0]$, 26 values
q_{PAH}	$[0; 0.1]$, 11 values in steps of 0.1
γ	$[0, 0.001, 0.0025, 0.005, 0.0075]$ $+[0.01; 0.1]$, 9 values in steps of 0.01 $+[0.2, 0.35, 0.5]$, 17 values in total
Optical emission: Brammer et al. 2008[a]	
A_V (Calzetti et al. 2000)	[0.6,0.12,0.19,0.29,1.05,2.68, 0.11,0.36,0.98,1.54,1.97,2.96]
M/L_V	[0.38,0.76,1.68,4.01,6.45,44.48, 0.12,0.21,0.33,0.64,1.57,4.00]
$\log_{10}$(sSFR)	[-10.75,-11.37,-11.90,-12.53,-12.05, -12.47,-8.37,-8.60,-8.50,-8.57,-8.93,-8.90]

Please refer to Brammer et al. 2008 for a more detailed description of the creation and selection of these basis set templates. See Blanton & Roweis 2007 for a methodology regarding the SFH.

observed bands where data is available. This is done by convolution of the filter transmission curves with the model SEDs. The resultant synthetic fluxes for each template and available observed band are then all combined into a two-dimensional matrix and passed onto the non-negative least square (nnls) algorithm in the python `scipy.optimize.nnls` package, that finds the best solution vector for that object. The nnls is a simple minimisation algorithm that in our case takes the form of,

$$\min ||A_{b,t}x^t - y^b||_2 \quad \text{where} \quad x \geq 0, \tag{A.7}$$

with A being the uncertainty weighted template matrix, y as the signal to noise for each band respectively, x the solution vector and $||.||_2$ signifying the Eucledian norm. Instead of finding the best fitting template, the algorithm computes the best-fit coefficients x, with $x \geq 0$. The number of simultaneously fit templates is user defined. For our purposes we chose to fit all 12 GB08 and 2 M11 templates at the same time, each time combining them with a single DL07 template, and iterating through all the possible combinations, and then finding the best fit. Following the Equation A.7, this is done by building the matrix A, for each band b, and each template t, where the first 14 templates in the matrix remain fixed, while the last element is being continuously replaced with a new DL07 model and looped over. In the end we obtain $4,862$ possible best-fit vectors - x^t, one per each DL07 model. The final result is then extracted from the χ^2 distribution of all best-fit solutions. The advantage of this method is in avoiding progressively looping over all possible template combinations to find the best solution, and instead only choosing to loop through DL07 templates. This approach significantly reduces the amount of required computational resources.

The resultant solution vector encodes the individual contributions from each template to the total emitted flux in each band. These are then added together to return the best fit solution. These three component split allows us to predict exactly how much each component contributes to the source's $L_{\rm IR}$, thus allowing us to differentiate between the AGN and the warm ISM dust emissions. In addition to that, the normalisation of DL07 templates leads us directly to the number of hydrogen nucleons, linking it to dust mass via Equation A.4. The radio data points are not being considered by our fitting routine, however we add a power-law radio slope with a spectral index of -0.75, as described by the FIR-radio correlation in Delvecchio et al. (2020), mainly for visualisation purposes, but also to detect the existence of the AGN radio excess.

A.1.5 Calculating the IR properties

To derive luminosity estimates, we integrate the three component summed SED in the $8-1000\,\mu$m. This gives us the $L_{\rm IR,tot}$ that contains within itself the energy emitted by the ISM dust and AGN torus, if present. The contribution of stellar emission at $\lambda_{\rm rest} > 8\,\mu$m is negligible, therefore we do not go through an additional step of subtracting those models. We then integrate the best fit template with just the AGN contribution in it, to obtain the $L_{\rm AGN}$ and $L_{\rm IR} = L_{\rm IR,tot} - L_{\rm AGN}$. We also compute the $f_{\rm AGN} = L_{\rm AGN}/L_{\rm IR,tot}$ to estimate how strongly the infrared SED of a galaxy is contaminated by AGN activity. In addition to that, it allows us to separate our sample into objects that have an active nucleus and those that do not. The conditions of the ISM in these different environments may vary quite significantly and would affect the extracted scaling relations if not treated correctly.

The normalisation of the DL07 models returns the number of hydrogen nucleons as one of its free parameters, from which we can obtain the fiducial $M_{\rm H}$. We then compute $M_{\rm dust}$ by converting this quantity assuming a fixed gas-to-dust ratio of $\delta_{GDR} = 100$, as prescribed in DL07. It is important to note that this ratio is encoded into the models and does not represent an actual physically meaningful conversion factor. We also compute a $T_{\rm dust}$ proxy in the form of the average radiation field intensity $\langle U \rangle$, from Equation A.4, by assuming a $P_0 = 125$. We note however that this quantity represents the luminosity-weighted dust temperature and has little to no bearing on the temperature of the cold dust or gas (see e.g. Scoville et al. 2016, 2017).

A.2 Stellar Mass Comparison

To test the robustness of `Stardust` - derived M_* estimates we start by first comparing them to the M_* as given by the original catalogue. The SDC2 takes its M_* directly from the COSMOS 2015 catalogue (Laigle et al. 2016), which uses SED fitting code `LePhare` (Arnouts et al. 1999; Ilbert et al. 2006) to constrain both the photometric redshift, M_*, as well as a host of other parameters.

Similarly to our approach, LePhare relies on Bruzual & Charlot (2003) SSP libraries to fit galaxy SEDs. We begin our fitting procedure by carefully correcting all the COSMOS 2015 aperture fluxes to total fluxes, as well as correcting for the MW extinction, as prescribed in Laigle et al. (2016). We then cross-match these sources to SDC2, by using the K-band, and fit the entirety of available 36 bands with our code. As a sanity check, we additionally run the same photometric catalogue through EAZY, albeit stopping at IRAC 2.

We present this comparison in Figure A.3. The M_* given by the parent catalogue and the ones derived by Stardust derived M_*, agree very well, with Stardust on average under-predicting the M_* by ~ 0.01 dex. We attribute a considerable 0.31 dex scatter to the fact that in their LePhare fit, Laigle et al. have used an iterative procedure, that involved correcting the observed fluxes in order to match the colours of the model library.

More reassuringly, we find a good correlation when comparing our M_* to EAZY, with a median offset of -0.03 dex and a minor 0.15 dex scatter being most likely induced by the fact that with Stardust we fit the entire available spectrum from UV to FIR, as opposed to just UV-optical, with EAZY.

A.3 Comparison with CIGALE

In order to better understand how our independent linear combination approach compares to the energy balance method, we have fit our sources with CIGALE. For this we have utilised a set of simple stellar population (SSP) models from Bruzual & Charlot (2003) for the non-obscured stellar light, the revised version of the Draine & Li (2007) templates for the obscured stellar light, reprocessed by dust, and Fritz et al. (2006) for the AGN contribution. The attenuation law that we considered was described in Calzetti et al. (2000). For the SSP templates, we have assumed a single delayed SFH. These CIGALE fits should be treated as a 'basic' first-pass approach, due to the computational limitations necessitating a constrained range for the template parameters. Ideally, such an analysis would require a flexible SFH, in order to obtain better SFR, as well as a wider range of parameters, both for the DL07 and Fritz et al. templates.

We show the comparison between the two methods in Figure A.4. There is a very good agreement between the $M_{\rm dust}$ derived with our code and CIGALE, with the difference having a mean of 0.09 dex and median of 0.02 dex. The derived values of $L_{\rm IR}$ are however in tension, with a mean of 0.20 dex and median of 0.11 dex. We attribute the significant outliers ($> 1\sigma$) to cases where the energy balance method in CIGALE has failed to account for the extra FIR flux. We also compare the Stardust and CIGALE computed M_*, and find that the two agree within 0.1 dex, albeit with a significant 0.3 dex scatter. As we have already discussed in Section 2.4, in certain environments the stellar and the dust emission could be spatially disconnected, thus the energy balance might not be the best physically motivated option. In addition, when dealing

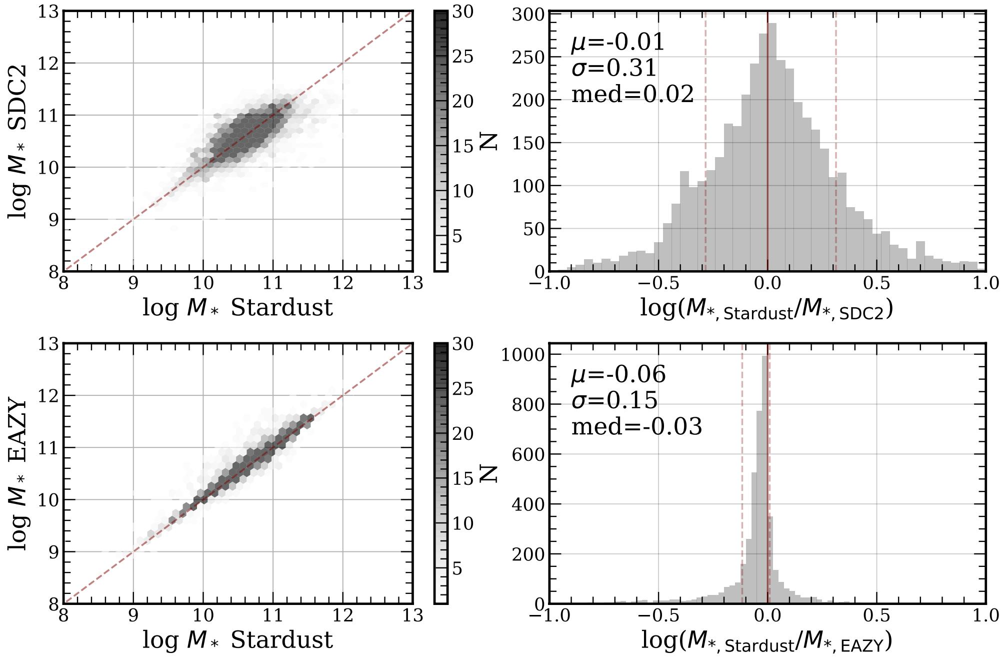

Figure A.3: Comparison of the Stardust derived M_* vs SDC2 (top) and EAZY (bottom). The dashed maroon line represents a 1:1 relation. The solid and dashed maroon lines represent a 1:1 relation and the 68 % confidence interval, respectively.

with extreme sources, the Calzetti et al. attenuation law might not allow the energy balance approach to account for all IR flux (e.g. see Buat et al. 2019). The above, however, are not the only explanations, as the identification/matching problems as well as IR flux extractions could also play a part in creating this tension between our results and CIGALE (e.g. see Małek et al. 2018).

When directly comparing computation times, it is important to note that, CIGALE fits sources within redshift blocks, where it pre-compiles a set of models first and then estimates the best-fit parameters, while Stardust fits sources sequentially. As such, despite both methods being parallelised, it is difficult to achieve a fair comparison between the two. Within a single redshift block, that numbers 288 objects, CIGALE has computed 50×10^6 models, and found the best fit in about 2.5 hours. Due to how the linear combination is performed within Stardust, defining an exact number of models attempted is not possible. However, considering that the 12 optical templates have been constructed as a basis set of $\sim 3,000$ models described in Brammer et al. (2008), combining that with 2 AGN templates and $\sim$ 4,800 DL07 models, results in roughly 30×10^6 total effective model combinations. Our code then takes 14 minutes in total to fit the same 288 objects, which is approximately 11 times faster than CIGALE, for the same number of CPU cores.

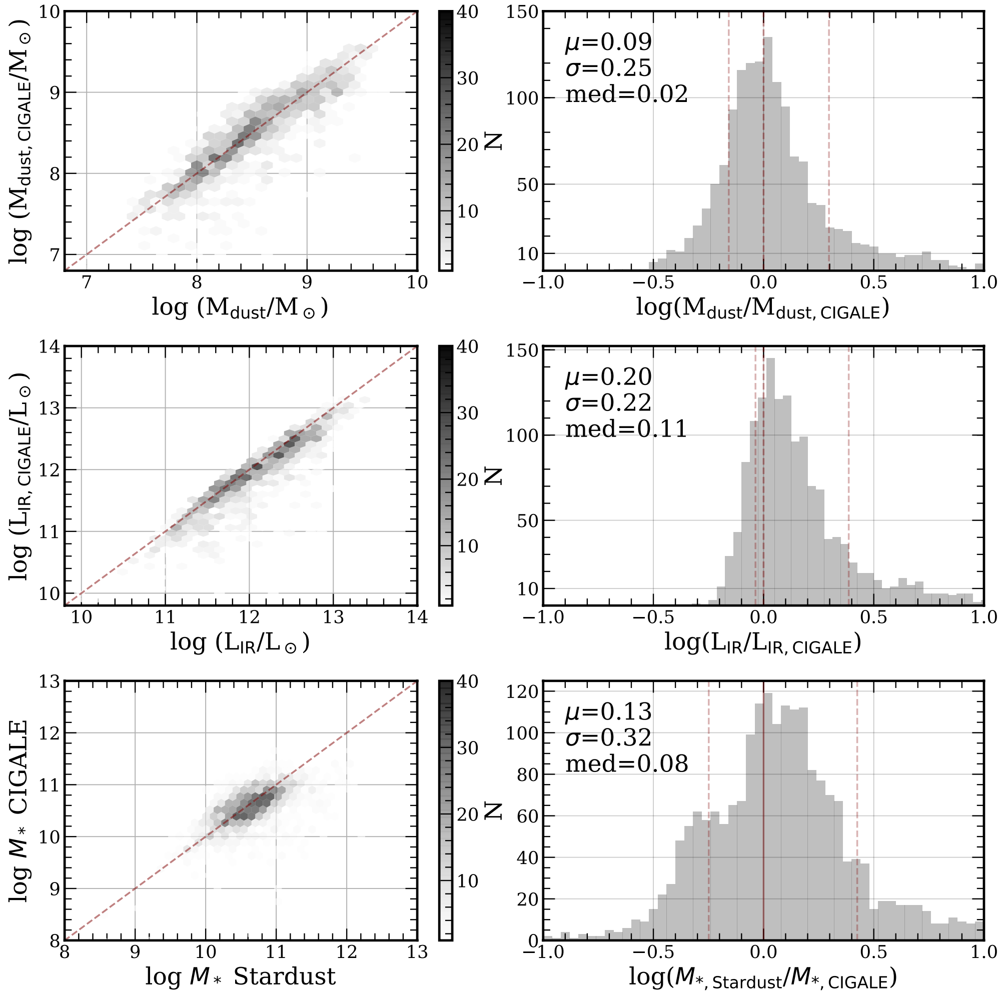

Figure A.4: Comparison of the derived M_{dust} (top) and L_{IR} (bottom) between `Stardust` and CIGALE. The solid and dashed maroon lines represent a 1:1 relation and the 68 % confidence interval, respectively.

Table A.2: Structure of the best-fit catalogue

Column Name	Units	Description
ID	-	ID of the object from the parent catalogue
RA	deg	The right ascension coordinate, as given in the parent catalogue
DEC	deg	The declination, as given in the parent catalogue
Area	-	Same as goodArea flag in the parent catalogues
z	-	Redshift used for fitting
ztype	-	Redshift type, 1 for spectroscopic, 0 for photometric, and 2 for $z_{\rm phot}$ from EAZY
LIR_total	$L_\odot$	Total FIR luminosity, obtained as the sum between AGN and DL07 components
eLIR_total	$L_\odot$	Uncertainty on the total FIR luminosity
Lagn_total	$L_\odot$	AGN luminosity
eLagn_total	$L_\odot$	Uncertainty on the AGN luminosity
Lir_draine	$L_\odot$	Luminosity given by the DL07 template
eLir_draine	$L_\odot$	Uncertainty on the luminosity given by the DL07 template
MD	$M_\odot$	Dust mass as predicted by the best-fit DL07 template
eMD	$M_\odot$	Uncertainty on the dust mass
deltaGDR	-	Gas-to-dust ratio
MG	$M_\odot$	Gas mass computed from $\delta_{\rm GDR}$ and $M_{\rm dust}$
eMG	$M_\odot$	The uncertainty on gas mass
Mstar	$M_\odot$	Stellar mass, equal to the one in the original catalogue
lastdet	μm	Last band that has a S/N>3 detection ($\lambda_{\rm last}$). Given in rest frame.
chi2	-	The χ^2 of the best-fit coefficients
f_agn	-	AGN fraction, given as $L_{\rm AGN}/L_{\rm IR,total}$
efagn	-	Uncertainty on the AGN fraction
fgas	-	Gas fraction computed as $M_{\rm gas}/M_{\rm dust}$
fgas_FMR	-	Gas fraction computed assuming $\delta_{\rm GDR}=100$
Umin	-	Best fit $U_{\rm min}$
gamma	-	Best fit γ
qpah	-	Index of the best fit $q_{\rm PAH}$ value
U	-	Average radiation field intensity $\langle U\rangle$
sU	-	Uncertainty on $\langle U\rangle$
deltaMS	-	SFR/SFR$_{\rm MS}$ where SFR=$L_{\rm IR,draine}\times 10^{-10}$ and SFR$_{\rm MS}$ is given by Schreiber et al. 2015
e_deltaMS	-	Uncertainty on ΔMS

Stardust also returns M_*, $A_{\rm V}$ and UV-optical SFR, these are not included in the release version of the catalogue, but are available upon request.
We have used the 7970d55 (28/06/21) version of Stardust to produce these catalogues.

APPENDIX B

Appendix B

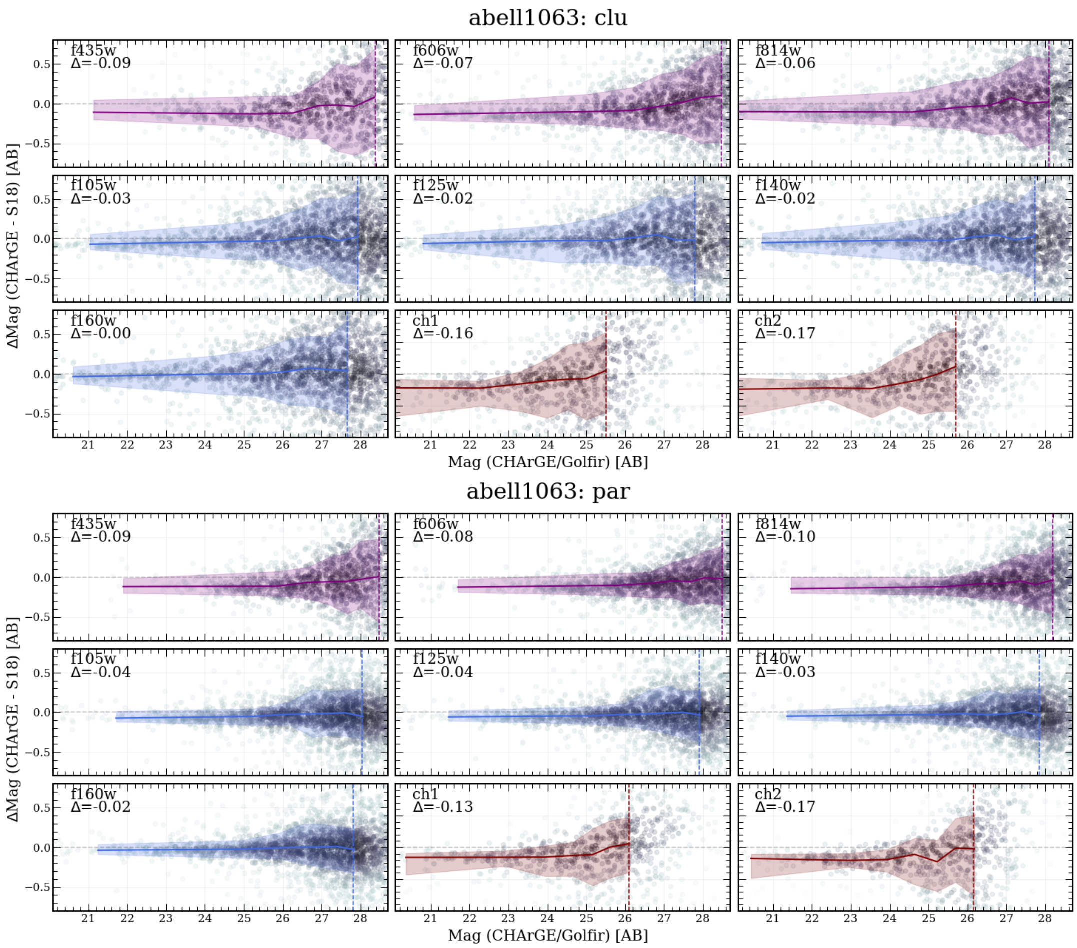

Figure B.1: Comparison of *HST* and IRAC photometry between our catalog and Shipley et al. 2018 for the cluster (top) and parallel (bottom) parts of Abell 1063 field. Symbols and colors are the same as in Figure 3.3.

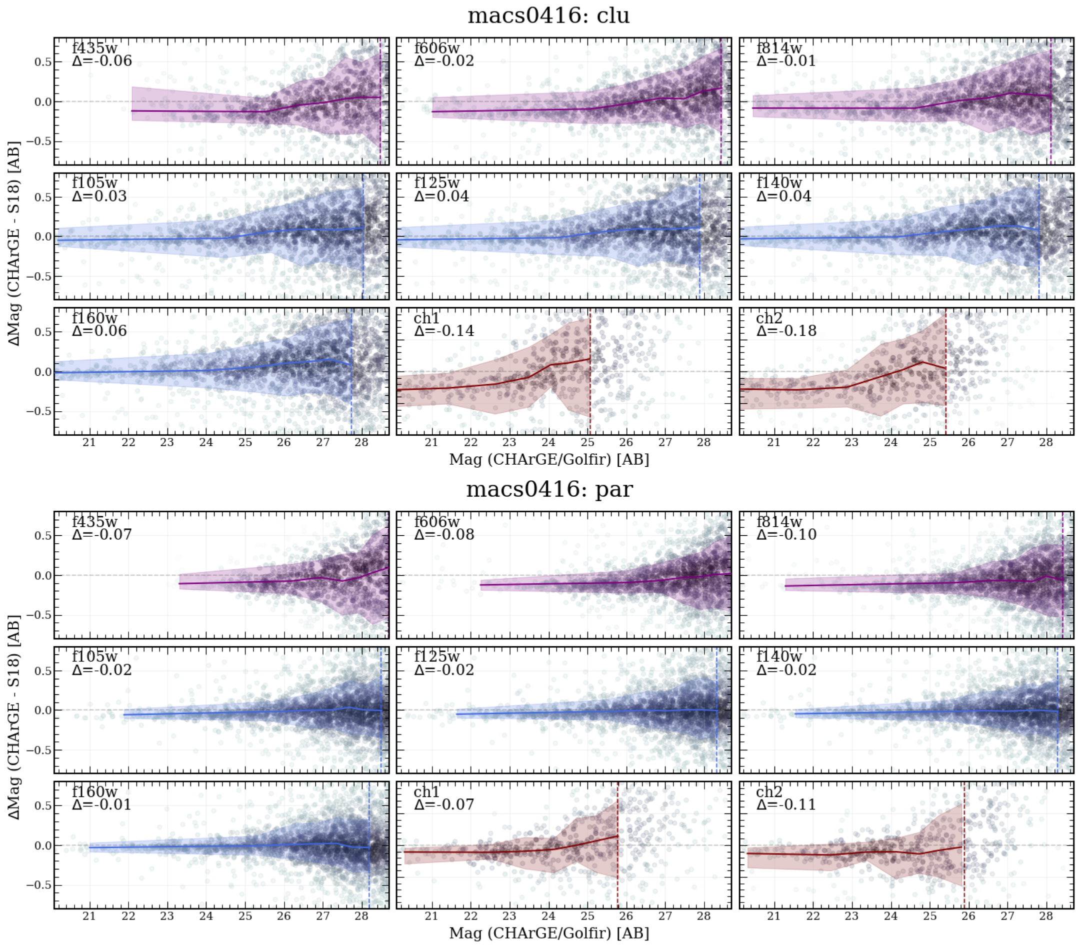

Figure B.2: Comparison of *HST* and IRAC photometry between our catalog and Shipley et al. 2018 for the cluster (top) and parallel (bottom) parts of MACS0416 field. Symbols and colors are the same as in Figure 3.3.

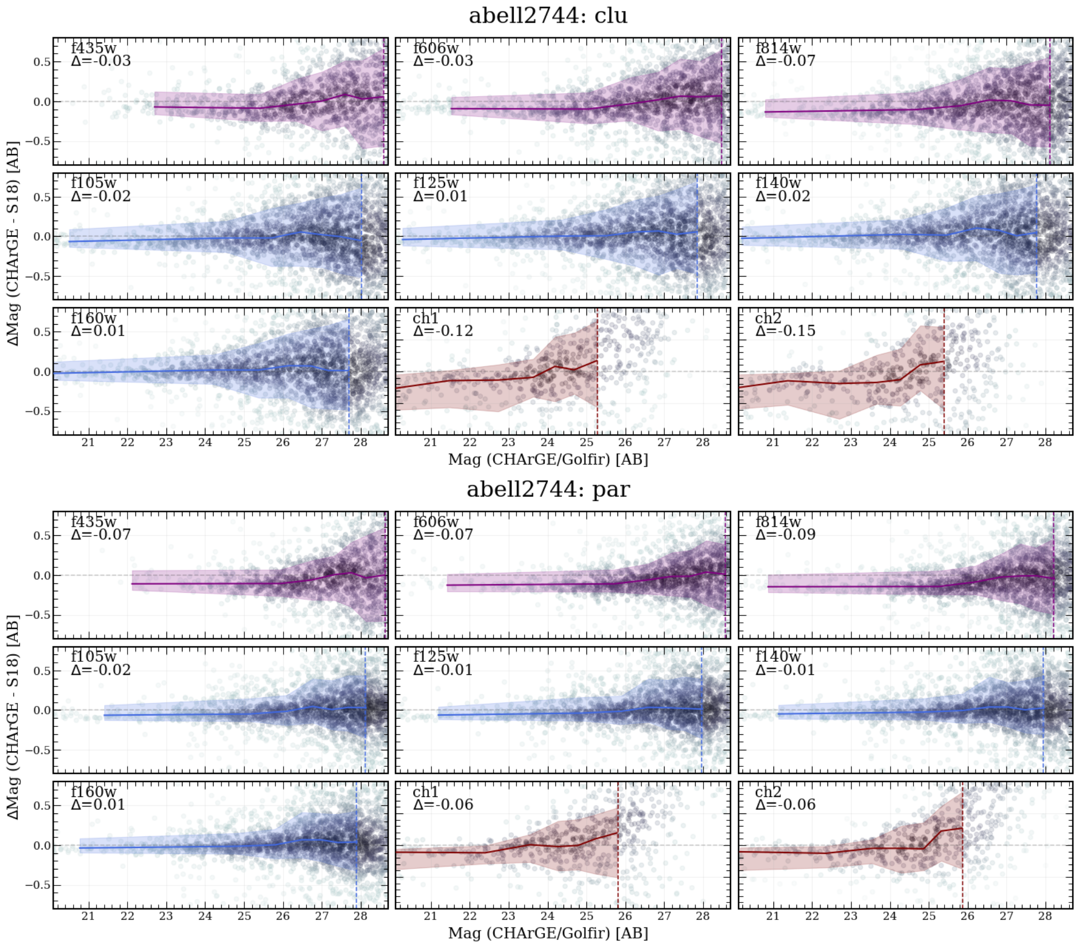

Figure B.3: Comparison of *HST* and IRAC photometry between our catalog and Shipley et al. 2018 for the cluster (top) and parallel (bottom) parts of Abell 2744 field. Symbols and colors are the same as in Figure 3.3.

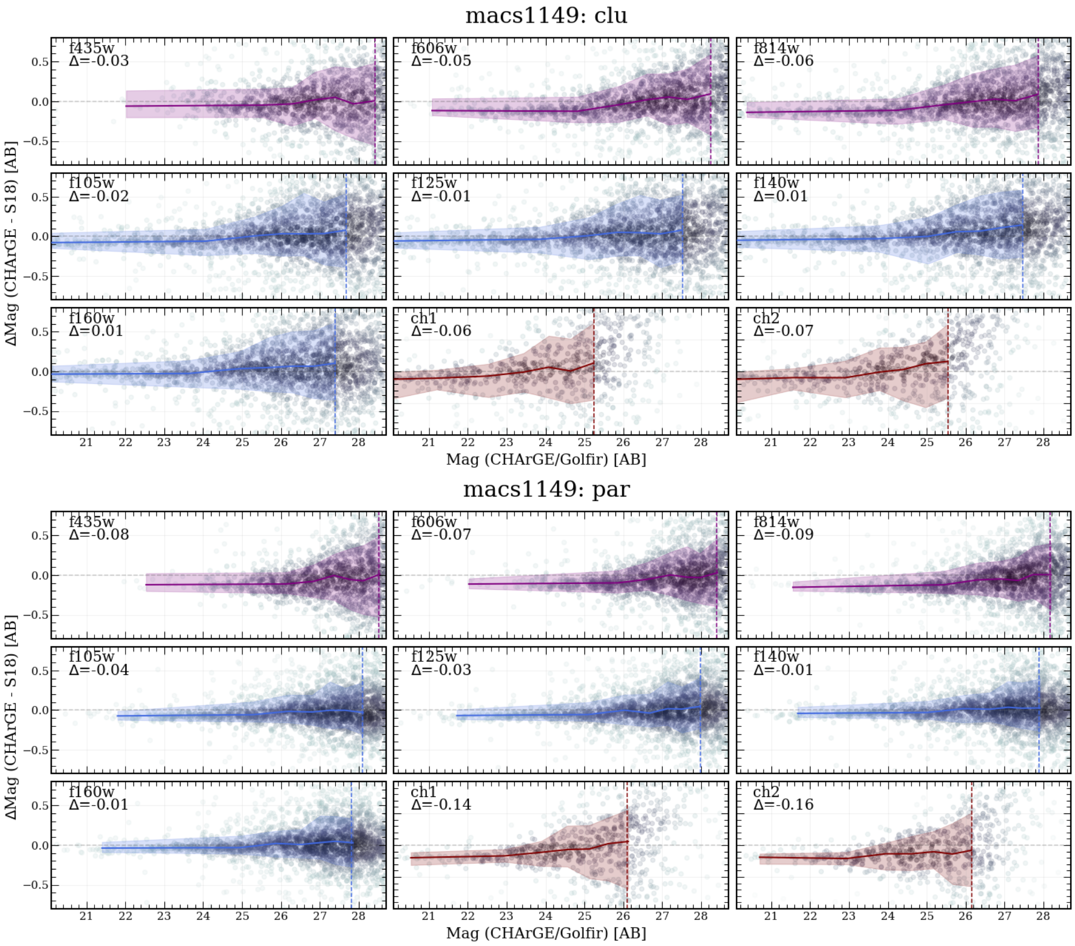

Figure B.4: Comparison of *HST* and IRAC photometry between our catalog and Shipley et al. 2018 for the cluster (top) and parallel (bottom) parts of MACS1149 field. Symbols and colors are the same as in Figure 3.3.

Bibliography

Accurso, G., Saintonge, A., Catinella, B., et al. 2017, MNRAS, 470, 4750, doi: 10.1093/mnras/stx1556

Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S8, doi: 10.1093/pasj/psx081

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

—. 2022, PASJ, 74, 247, doi: 10.1093/pasj/psab122

Alaghband-Zadeh, S., Chapman, S. C., Swinbank, A. M., et al. 2013, MNRAS, 435, 1493, doi: 10.1093/mnras/stt1390

Albrecht, A., & Steinhardt, P. J. 1982, Phys. Rev. Lett., 48, 1220, doi: 10.1103/PhysRevLett.48.1220

Allen, R. J., Ivette Rodríguez, M., Black, J. H., & Booth, R. S. 2012, AJ, 143, 97, doi: 10.1088/0004-6256/143/4/97

Alpher, R. A., Bethe, H., & Gamow, G. 1948, Physical Review, 73, 803, doi: 10.1103/PhysRev.73.803

Anderson, J., & King, I. R. 2000, PASP, 112, 1360, doi: 10.1086/316632

Aniano, G., Draine, B. T., Hunt, L. K., et al. 2020, ApJ, 889, 150, doi: 10.3847/1538-4357/ab5fdb

Antonucci, R. 1993, ARA&A, 31, 473, doi: 10.1146/annurev.aa.31.090193.002353

Aoyama, S., Hou, K.-C., Shimizu, I., et al. 2017, MNRAS, 466, 105, doi: 10.1093/mnras/stw3061

Aoyama, S., Hirashita, H., Lim, C.-F., et al. 2019, MNRAS, 484, 1852, doi: 10.1093/mnras/stz021

Aravena, M., Decarli, R., Walter, F., et al. 2016, ApJ, 833, 68, doi: 10.3847/1538-4357/833/1/68

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

Arnett, W. D., Bahcall, J. N., Kirshner, R. P., & Woosley, S. E. 1989, ARA&A, 27, 629, doi: 10.1146/annurev.aa.27.090189.003213

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

Arnouts, S., Le Floc'h, E., Chevallard, J., et al. 2013, A&A, 558, A67, doi: 10.1051/0004-6361/201321768

Asano, R. S., Takeuchi, T. T., Hirashita, H., & Inoue, A. K. 2013, Earth, Planets and Space, 65, 213, doi: 10.5047/eps.2012.04.014

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Avni, Y. 1976, ApJ, 210, 642, doi: 10.1086/154870

Balestra, I., Mercurio, A., Sartoris, B., et al. 2016, ApJS, 224, 33, doi: 10.3847/0067-0049/224/2/33

Barbary, K. 2016, Journal of Open Source Software, 1, 58, doi: 10.21105/joss.00058

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

Beeston, R. A., Wright, A. H., Maddox, S., et al. 2018, MNRAS, 479, 1077, doi: 10.1093/mnras/sty1460

Bendo, G. J., Galliano, F., & Madden, S. C. 2012, MNRAS, 423, 197, doi: 10.1111/j.1365-2966.2012.20784.x

Bernard, J. P., Paradis, D., Marshall, D. J., et al. 2010, A&A, 518, L88, doi: 10.1051/0004-6361/201014540

Berta, S., Lutz, D., Genzel, R., Förster-Schreiber, N. M., & Tacconi, L. J. 2016, A&A, 587, A73, doi: 10.1051/0004-6361/201527746

Berta, S., Lutz, D., Santini, P., et al. 2013, A&A, 551, A100, doi: 10.1051/0004-6361/201220859

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

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

Béthermin, M., Daddi, E., Magdis, G., et al. 2015, A&A, 573, A113, doi: 10.1051/0004-6361/201425031

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

Bisbas, T. G., Papadopoulos, P. P., & Viti, S. 2015, ApJ, 803, 37, doi: 10.1088/0004-637X/803/1/37

Bisbas, T. G., van Dishoeck, E. F., Papadopoulos, P. P., et al. 2017, ApJ, 839, 90, doi: 10.3847/1538-4357/aa696d

Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734, doi: 10.1086/510127

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

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

Boquien, M., Calzetti, D., Combes, F., et al. 2011, AJ, 142, 111, doi: 10.1088/0004-6256/142/4/111

Boselli, A., Roehlly, Y., Fossati, M., et al. 2016, A&A, 596, A11, doi: 10.1051/0004-6361/201629221

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

Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928, doi: 10.1086/521811

Bouwens, R. J., Bradley, L., Zitrin, A., et al. 2014, ApJ, 795, 126, doi: 10.1088/0004-637X/795/2/126

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34, doi: 10.1088/0004-637X/803/1/34

Bouwens, R. J., Oesch, P. A., Labbé, I., et al. 2016, ApJ, 830, 67, doi: 10.3847/0004-637X/830/2/67

Bradley, L., Sipőcz, B., Robitaille, T., et al. 2020, astropy/photutils: 1.0.0, 1.0.0, Zenodo, doi: 10.5281/zenodo.4044744. https://doi.org/10.5281/zenodo.4044744

Bradley, L. D., Zitrin, A., Coe, D., et al. 2014, ApJ, 792, 76, doi: 10.1088/0004-637X/792/1/76

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

Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2011, ApJ, 739, 24, doi: 10.1088/0004-637X/739/1/24

Bressan, A., Chiosi, C., & Fagotto, F. 1994, ApJS, 94, 63, doi: 10.1086/192073

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151, doi: 10.1111/j.1365-2966.2004.07881.x

Brown, M. J. I., Moustakas, J., Smith, J. D. T., et al. 2014, ApJS, 212, 18, doi: 10.1088/0067-0049/212/2/18

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

Bruzual A., G., & Charlot, S. 1993, ApJ, 405, 538, doi: 10.1086/172385

Buat, V., Ciesla, L., Boquien, M., Małek, K., & Burgarella, D. 2019, A&A, 632, A79, doi: 10.1051/0004-6361/201936643

Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413, doi: 10.1111/j.1365-2966.2005.09131.x

Calistro Rivera, G., Lusso, E., Hennawi, J. F., & Hogg, D. W. 2016, ApJ, 833, 98, doi: 10.3847/1538-4357/833/1/98

Calistro Rivera, G., Hodge, J. A., Smail, I., et al. 2018, ApJ, 863, 56, doi: 10.3847/1538-4357/aacffa

Calvi, V., Trenti, M., Stiavelli, M., et al. 2016, ApJ, 817, 120, doi: 10.3847/0004-637X/817/2/120

Calzetti, D. 2001, PASP, 113, 1449, doi: 10.1086/324269

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

Caminha, G. B., Grillo, C., Rosati, P., et al. 2016, A&A, 587, A80, doi: 10.1051/0004-6361/201527670

Caminha, G. B., Rosati, P., Grillo, C., et al. 2019, A&A, 632, A36, doi: 10.1051/0004-6361/201935454

Capak, P. L., Carilli, C., Jones, G., et al. 2015, Nature, 522, 455, doi: 10.1038/nature14500

Cappellari, M., & Copin, Y. 2003, MNRAS, 342, 345, doi: 10.1046/j.1365-8711.2003.06541.x

Carilli, C. L., & Walter, F. 2013, ARA&A, 51, 105, doi: 10.1146/annurev-astro-082812-140953

Carniani, S., Maiolino, R., De Zotti, G., et al. 2015, A&A, 584, A78, doi: 10.1051/0004-6361/201525780

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

Casey, C. M., Berta, S., Béthermin, M., et al. 2012, ApJ, 761, 140, doi: 10.1088/0004-637X/761/2/140

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

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

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

Cassata, P., Liu, D., Groves, B., et al. 2020, ApJ, 891, 83, doi: 10.3847/1538-4357/ab7452

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

Chan, J. C. C., Beifiori, A., Mendel, J. T., et al. 2016, MNRAS, 458, 3181, doi: 10.1093/mnras/stw502

Chang, Y.-Y., Le Floc'h, E., Juneau, S., et al. 2017, MNRAS, 466, L103, doi: 10.1093/mnrasl/slw247

Charlot, S., & Bruzual A, G. 1991, ApJ, 367, 126, doi: 10.1086/169608

Chartab, N., Cooray, A., Ma, J., et al. 2022, arXiv e-prints, arXiv:2201.07478. https://arxiv.org/abs/2201.07478

Chen, C.-C., Hodge, J. A., Smail, I., et al. 2017, ApJ, 846, 108, doi: 10.3847/1538-4357/aa863a

Chen, Y.-M., Wild, V., Kauffmann, G., et al. 2009, MNRAS, 393, 406, doi: 10.1111/j.1365-2966.2008.14247.x

Cibinel, A., Daddi, E., Sargent, M. T., et al. 2019, MNRAS, 485, 5631, doi: 10.1093/mnras/stz690

Ciesla, L., Buat, V., Boquien, M., et al. 2021, A&A, 653, A6, doi: 10.1051/0004-6361/202140762

Cochrane, R. K., Hayward, C. C., Anglés-Alcázar, D., et al. 2019, MNRAS, 488, 1779, doi: 10.1093/mnras/stz1736

Cochrane, R. K., Best, P. N., Smail, I., et al. 2021, MNRAS, 503, 2622, doi: 10.1093/mnras/stab467

Coe, D., Zitrin, A., Carrasco, M., et al. 2013, ApJ, 762, 32, doi: 10.1088/0004-637X/762/1/32

Coe, D., Salmon, B., Bradač, M., et al. 2019, ApJ, 884, 85, doi: 10.3847/1538-4357/ab412b

Compiégne, M. 2010, in Astronomical Society of the Pacific Conference Series, Vol. 438, The Dynamic Interstellar Medium: A Celebration of the Canadian Galactic Plane Survey, ed. R. Kothes, T. L. Landecker, & A. G. Willis, 55

Conroy, C. 2013, ARA&A, 51, 393, doi: 10.1146/annurev-astro-082812-141017

Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486, doi: 10.1088/0004-637X/699/1/486

Cormier, D., Madden, S. C., Lebouteiller, V., et al. 2015, A&A, 578, A53, doi: 10.1051/0004-6361/201425207

Cortzen, I., Magdis, G. E., Valentino, F., et al. 2020, A&A, 634, L14, doi: 10.1051/0004-6361/201937217

Cousin, M., Buat, V., Lagache, G., & Bethermin, M. 2019, A&A, 627, A132, doi: 10.1051/0004-6361/201834674

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

Croxall, K. V., Smith, J. D., Pellegrini, E., et al. 2017, ApJ, 845, 96, doi: 10.3847/1538-4357/aa8035

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

Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156, doi: 10.1086/521818

Daddi, E., Elbaz, D., Walter, F., et al. 2010, ApJl, 714, L118, doi: 10.1088/2041-8205/714/1/L118

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

Dale, D. A., Helou, G., Contursi, A., Silbermann, N. A., & Kolhatkar, S. 2001, ApJ, 549, 215, doi: 10.1086/319077

Dale, D. A., Helou, G., Magdis, G. E., et al. 2014, ApJ, 784, 83, doi: 10.1088/0004-637X/784/1/83

Dale, D. A., Aniano, G., Engelbracht, C. W., et al. 2012, ApJ, 745, 95, doi: 10.1088/0004-637X/745/1/95

Dannerbauer, H., Daddi, E., Riechers, D. A., et al. 2009, ApJ, 698, L178, doi: 10.1088/0004-637X/698/2/L178

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

Davé, R., Crain, R. A., Stevens, A. R. H., et al. 2020, MNRAS, 497, 146, doi: 10.1093/mnras/staa1894

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

Davé, R., Rafieferantsoa, M. H., Thompson, R. J., & Hopkins, P. F. 2017, MNRAS, 467, 115, doi: 10.1093/mnras/stx108

Davidzon, I., Ilbert, O., Faisst, A. L., Sparre, M., & Capak, P. L. 2018, ApJ, 852, 107, doi: 10.3847/1538-4357/aaa19e

Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, A&A, 605, A70, doi: 10.1051/0004-6361/201730419

De Looze, I., Baes, M., Bendo, G. J., Cortese, L., & Fritz, J. 2011, MNRAS, 416, 2712, doi: 10.1111/j.1365-2966.2011.19223.x

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

Decarli, R., Walter, F., Gónzalez-López, J., et al. 2019, ApJ, 882, 138, doi: 10.3847/1538-4357/ab30fe

Dekel, A., Sari, R., & Ceverino, D. 2009, ApJ, 703, 785, doi: 10.1088/0004-637X/703/1/785

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

Delvecchio, I., Daddi, E., Sargent, M. T., et al. 2020, arXiv e-prints, arXiv:2010.05510. https://arxiv.org/abs/2010.05510

Desert, F. X., Boulanger, F., & Puget, J. L. 1990, A&A, 500, 313

Devriendt, J. E. G., Guiderdoni, B., & Sadat, R. 1999, A&A, 350, 381. https://arxiv.org/abs/astro-ph/9906332

Di Criscienzo, M., Merlin, E., Castellano, M., et al. 2017, A&A, 607, A30, doi: 10.1051/0004-6361/201731172

Díaz-Santos, T., Armus, L., Charmandaris, V., et al. 2017, ApJ, 846, 32, doi: 10.3847/1538-4357/aa81d7

Dickinson, M., Giavalisco, M., & GOODS Team. 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini, 324

Diego, J. M., Broadhurst, T., Wong, J., et al. 2016, MNRAS, 459, 3447, doi: 10.1093/mnras/stw865

Donevski, D., Lapi, A., Małek, K., et al. 2020, arXiv e-prints, arXiv:2008.09995. https://arxiv.org/abs/2008.09995

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

Draine, B. T., Dale, D. A., Bendo, G., et al. 2007, ApJ, 663, 866, doi: 10.1086/518306

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

Driver, S. P., Andrews, S. K., da Cunha, E., et al. 2018, MNRAS, 475, 2891, doi: 10.1093/mnras/stx2728

Dubois, Y., Pichon, C., Welker, C., et al. 2014, MNRAS, 444, 1453, doi: 10.1093/mnras/stu1227

Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2020, MNRAS, 494, 3828, doi: 10.1093/mnras/staa769

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

Dunne, L., Gomez, H. L., da Cunha, E., et al. 2011, MNRAS, 417, 1510, doi: 10.1111/j.1365-2966.2011.19363.x

Dwek, E. 1998, ApJ, 501, 643, doi: 10.1086/305829

Eales, S., Smith, M. W. L., Auld, R., et al. 2012, ApJ, 761, 168, doi: 10.1088/0004-637X/761/2/168

Ebeling, H., Ma, C.-J., & Barrett, E. 2014, ApJS, 211, 21, doi: 10.1088/0067-0049/211/2/21

Efstathiou, G., & Silk, J. 1983, Fund. Cosmic Phys., 9, 1

Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, ApJ, 136, 748, doi: 10.1086/147433

Einstein, A. 1907, Annalen der Physik, 328, 371, doi: 10.1002/andp.19073280713

—. 1908, Jahrbuch der Radioaktivität und Elektronik, 4, 411

—. 1916, Annalen der Physik, 354, 769, doi: 10.1002/andp.19163540702

Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33, doi: 10.1051/0004-6361:20077525

Elbaz, D., Hwang, H. S., Magnelli, B., et al. 2010, A&A, 518, L29, doi: 10.1051/0004-6361/201014687

Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119, doi: 10.1051/0004-6361/201117239

Elbaz, D., Leiton, R., Nagar, N., et al. 2018, A&A, 616, A110, doi: 10.1051/0004-6361/201732370

Ellis, R. S., McLure, R. J., Dunlop, J. S., et al. 2013, ApJ, 763, L7, doi: 10.1088/2041-8205/763/1/L7

Erfanianfar, G., Popesso, P., Finoguenov, A., et al. 2016, MNRAS, 455, 2839, doi: 10.1093/mnras/stv2485

Feltre, A., Hatziminaoglou, E., Hernán-Caballero, A., et al. 2013, MNRAS, 434, 2426, doi: 10.1093/mnras/stt1177

Ferrarese, L., Côté, P., Jordán, A., et al. 2006, ApJS, 164, 334, doi: 10.1086/501350

Finkelstein, S. L., Ryan, Russell E., J., Papovich, C., et al. 2015, ApJ, 810, 71, doi: 10.1088/0004-637X/810/1/71

Fioc, M., & Rocca-Volmerange, B. 1997, A&A, 500, 507. https://arxiv.org/abs/astro-ph/9707017

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

Franco, M., Elbaz, D., Zhou, L., et al. 2020, A&A, 643, A30, doi: 10.1051/0004-6361/202038312

Friedmann, A. 1922, Zeitschrift fur Physik, 10, 377, doi: 10.1007/BF01332580

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

Fudamoto, Y., Oesch, P. A., Schouws, S., et al. 2021, Nature, 597, 489, doi: 10.1038/s41586-021-03846-z

Fujimoto, S., Ouchi, M., Ono, Y., et al. 2016, ApJS, 222, 1, doi: 10.3847/0067-0049/222/1/1

Fujimoto, S., Oguri, M., Brammer, G., et al. 2021, ApJ, 911, 99, doi: 10.3847/1538-4357/abd7ec

Galametz, A., Grazian, A., Fontana, A., et al. 2013, ApJS, 206, 10, doi: 10.1088/0067-0049/206/2/10

Galametz, M., Albrecht, M., Kennicutt, R., et al. 2014, MNRAS, 439, 2542, doi: 10.1093/mnras/stu113

Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, Space Sci. Rev., 123, 485, doi: 10.1007/s11214-006-8315-7

Geach, J. E., Smail, I., Moran, S. M., et al. 2011, ApJl, 730, L19, doi: 10.1088/2041-8205/730/2/L19

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

Genzel, R., Tacconi, L. J., Kurk, J., et al. 2013, ApJ, 773, 68, doi: 10.1088/0004-637X/773/1/68

Genzel, R., Tacconi, L. J., Lutz, D., et al. 2015, ApJ, 800, 20, doi: 10.1088/0004-637X/800/1/20

Gerhard, O. E. 1981, MNRAS, 197, 179, doi: 10.1093/mnras/197.1.179

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

Glover, S. C. O., & Smith, R. J. 2016, MNRAS, 462, 3011, doi: 10.1093/mnras/stw1879

Goldsmith, P. F. 2001, ApJ, 557, 736, doi: 10.1086/322255

Gómez-Guijarro, C., Toft, S., Karim, A., et al. 2018, ApJ, 856, 121, doi: 10.3847/1538-4357/aab206

González, V., Labbé, I., Bouwens, R. J., et al. 2011, ApJ, 735, L34, doi: 10.1088/2041-8205/735/2/L34

González-López, J., Bauer, F. E., Romero-Cañizales, C., et al. 2017, A&A, 597, A41, doi: 10.1051/0004-6361/201628806

Greve, T. R., Bertoldi, F., Smail, I., et al. 2005, MNRAS, 359, 1165, doi: 10.1111/j.1365-2966.2005.08979.x

Grillo, C., Karman, W., Suyu, S. H., et al. 2016, ApJ, 822, 78, doi: 10.3847/0004-637X/822/2/78

Groves, B., Dopita, M. A., Sutherland, R. S., et al. 2008, ApJS, 176, 438, doi: 10.1086/528711

Groves, B. A., Schinnerer, E., Leroy, A., et al. 2015, ApJ, 799, 96, doi: 10.1088/0004-637X/799/1/96

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

Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633, doi: 10.1086/148444

Guth, A. H. 1981, Phys. Rev. D, 23, 347, doi: 10.1103/PhysRevD.23.347

Hashimoto, T., Laporte, N., Mawatari, K., et al. 2018, Nature, 557, 392, doi: 10.1038/s41586-018-0117-z

Hatsukade, B., Ohta, K., Seko, A., Yabe, K., & Akiyama, M. 2013, ApJ, 769, L27, doi: 10.1088/2041-8205/769/2/L27

Hatsukade, B., Kohno, K., Umehata, H., et al. 2016, PASJ, 68, 36, doi: 10.1093/pasj/psw026

Hayward, C. C., Narayanan, D., Kereš, D., et al. 2013, MNRAS, 428, 2529, doi: 10.1093/mnras/sts222

Hayward, C. C., Lanz, L., Ashby, M. L. N., et al. 2014, MNRAS, 445, 1598, doi: 10.1093/mnras/stu1843

Heintz, K. E., Watson, D., Oesch, P. A., Narayanan, D., & Madden, S. C. 2021, ApJ, 922, 147, doi: 10.3847/1538-4357/ac2231

Hildebrandt, H., Erben, T., Kuijken, K., et al. 2012, MNRAS, 421, 2355, doi: 10.1111/j.1365-2966.2012.20468.x

Hoag, A., Bradač, M., Brammer, G., et al. 2018, ApJ, 854, 39, doi: 10.3847/1538-4357/aaa9c2

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

Hodge, J. A., Carilli, C. L., Walter, F., et al. 2012, ApJ, 760, 11, doi: 10.1088/0004-637X/760/1/11

Hodge, J. A., & da Cunha, E. 2020, Royal Society Open Science, 7, 200556, doi: 10.1098/rsos.200556

Hodge, J. A., Swinbank, A. M., Simpson, J. M., et al. 2016, ApJ, 833, 103, doi: 10.3847/1538-4357/833/1/103

Hodge, J. A., Smail, I., Walter, F., et al. 2019, ApJ, 876, 130, doi: 10.3847/1538-4357/ab1846

Holland, W. S., Robson, E. I., Gear, W. K., et al. 1999, MNRAS, 303, 659, doi: 10.1046/j.1365-8711.1999.02111.x

Hopkins, A. M., & Beacom, J. F. 2006, ApJ, 651, 142, doi: 10.1086/506610

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168, doi: 10.1073/pnas.15.3.168

Hubble, E., & Humason, M. L. 1931, ApJ, 74, 43, doi: 10.1086/143323

Hubble, E. P. 1936, Realm of the Nebulae

Hughes, T. M., Ibar, E., Villanueva, V., et al. 2017, MNRAS, 468, L103, doi: 10.1093/mnrasl/slx033

Ikeda, M., Oka, T., Tatematsu, K., Sekimoto, Y., & Yamamoto, S. 2002, ApJS, 139, 467, doi: 10.1086/338761

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

Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55, doi: 10.1051/0004-6361/201321100

Imara, N., Loeb, A., Johnson, B. D., Conroy, C., & Behroozi, P. 2018, ApJ, 854, 36, doi: 10.3847/1538-4357/aaa3f0

Infante, L., Zheng, W., Laporte, N., et al. 2015, ApJ, 815, 18, doi: 10.1088/0004-637X/815/1/18

Inoue, A. K. 2011, Earth, Planets and Space, 63, 1027, doi: 10.5047/eps.2011.02.013

Ishigaki, M., Kawamata, R., Ouchi, M., et al. 2015, ApJ, 799, 12, doi: 10.1088/0004-637X/799/1/12

Israel, F. P. 2005, A&A, 438, 855, doi: 10.1051/0004-6361:20042237

Jauzac, M., Clément, B., Limousin, M., et al. 2014, MNRAS, 443, 1549, doi: 10.1093/mnras/stu1355

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

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

Jonsson, P. 2006, MNRAS, 372, 2, doi: 10.1111/j.1365-2966.2006.10884.x

Kaasinen, M., Scoville, N., Walter, F., et al. 2019, ApJ, 880, 15, doi: 10.3847/1538-4357/ab253b

Kaasinen, M., Walter, F., Novak, M., et al. 2020, ApJ, 899, 37, doi: 10.3847/1538-4357/aba438

Kamenetzky, J., Rangwala, N., Glenn, J., Maloney, P. R., & Conley, A. 2014, ApJ, 795, 174, doi: 10.1088/0004-637X/795/2/174

Kapteyn, J. C. 1922, ApJ, 55, 302, doi: 10.1086/142670

Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, ApJ, 730, 61, doi: 10.1088/0004-637X/730/2/61

Karman, W., Grillo, C., Balestra, I., et al. 2016, A&A, 585, A27, doi: 10.1051/0004-6361/201527443

Kashino, D., Silverman, J. D., Rodighiero, G., et al. 2013, ApJl, 777, L8, doi: 10.1088/2041-8205/777/1/L8

Kawamata, R., Ishigaki, M., Shimasaku, K., Oguri, M., & Ouchi, M. 2015, ApJ, 804, 103, doi: 10.1088/0004-637X/804/2/103

Kennicutt, Robert C., J. 1989, ApJ, 344, 685, doi: 10.1086/167834

—. 1998, ARA&A, 36, 189, doi: 10.1146/annurev.astro.36.1.189

Kewley, L. J., & Dopita, M. A. 2002, ApJS, 142, 35, doi: 10.1086/341326

Kewley, L. J., & Ellison, S. L. 2008, ApJ, 681, 1183, doi: 10.1086/587500

Kneib, J.-P., & Natarajan, P. 2011, A&A Rev., 19, 47, doi: 10.1007/s00159-011-0047-3

Knudsen, K. K., Richard, J., Kneib, J.-P., et al. 2016, MNRAS, 462, L6, doi: 10.1093/mnrasl/slw114

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

Kokorev, V. I., Magdis, G. E., Davidzon, I., et al. 2021, ApJ, 921, 40, doi: 10.3847/1538-4357/ac18ce

Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221, doi: 10.1088/0004-637X/700/1/221

Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, ApJS, 218, 15, doi: 10.1088/0067-0049/218/2/15

Kroupa, P. 2001, MNRAS, 322, 231, doi: 10.1046/j.1365-8711.2001.04022.x

Krumholz, M. R., Klein, R. I., & McKee, C. F. 2011, ApJ, 740, 74, doi: 10.1088/0004-637X/740/2/74

Kurczynski, P., Gawiser, E., Acquaviva, V., et al. 2016, ApJl, 820, L1, doi: 10.3847/2041-8205/820/1/L1

Labbé, I., Huang, J., Franx, M., et al. 2005, ApJ, 624, L81, doi: 10.1086/430700

Labbé, I., Oesch, P. A., Bouwens, R. J., et al. 2013, ApJ, 777, L19, doi: 10.1088/2041-8205/777/2/L19

Labbé, I., Oesch, P. A., Illingworth, G. D., et al. 2015, ApJS, 221, 23, doi: 10.1088/0067-0049/221/2/23

Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2016, MNRAS, 462, 3854, doi: 10.1093/mnras/stw1888

Lagache, G., Puget, J.-L., & Dole, H. 2005, ARA&A, 43, 727, doi: 10.1146/annurev.astro.43.072103.150606

Lagattuta, D. J., Richard, J., Clément, B., et al. 2017, MNRAS, 469, 3946, doi: 10.1093/mnras/stx1079

Lagos, C. d. P., Bayet, E., Baugh, C. M., et al. 2012, MNRAS, 426, 2142, doi: 10.1111/j.1365-2966.2012.21905.x

Lagos, C. d. P., da Cunha, E., Robotham, A. S. G., et al. 2020, MNRAS, 499, 1948, doi: 10.1093/mnras/staa2861

Lagos, C. d. P., Crain, R. A., Schaye, J., et al. 2015, MNRAS, 452, 3815, doi: 10.1093/mnras/stv1488

Lagos, C. d. P., Robotham, A. S. G., Trayford, J. W., et al. 2019, MNRAS, 489, 4196, doi: 10.1093/mnras/stz2427

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

Laigle, C., Davidzon, I., Ilbert, O., et al. 2019, MNRAS, 486, 5104, doi: 10.1093/mnras/stz1054

Lang, P., Wuyts, S., Somerville, R. S., et al. 2014, ApJ, 788, 11, doi: 10.1088/0004-637X/788/1/11

Langer, W. D., Velusamy, T., Pineda, J. L., Willacy, K., & Goldsmith, P. F. 2014, A&A, 561, A122, doi: 10.1051/0004-6361/201322406

Laporte, N., Ellis, R. S., Boone, F., et al. 2017, ApJ, 837, L21, doi: 10.3847/2041-8213/aa62aa

Laporte, N., Zitrin, A., Ellis, R. S., et al. 2021, MNRAS, 505, 4838, doi: 10.1093/mnras/stab191

Le Floc'h, E., Papovich, C., Dole, H., et al. 2005, ApJ, 632, 169, doi: 10.1086/432789

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

Leavitt, H. S. 1908, Annals of Harvard College Observatory, 60, 87

Lee, K.-S., Ferguson, H. C., Wiklind, T., et al. 2012, ApJ, 752, 66, doi: 10.1088/0004-637X/752/1/66

Lee, N., Sanders, D. B., Casey, C. M., et al. 2015, ApJ, 801, 80, doi: 10.1088/0004-637X/801/2/80

Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3, doi: 10.1086/313233

Leja, J., Johnson, B. D., Conroy, C., & van Dokkum, P. 2018, ApJ, 854, 62, doi: 10.3847/1538-4357/aaa8db

Lemaître, G. 1927, Annales de la Société Scientifique de Bruxelles, 47, 49

—. 1931, MNRAS, 91, 483, doi: 10.1093/mnras/91.5.483

Leroy, A. K., Bolatto, A., Gordon, K., et al. 2011a, ApJ, 737, 12, doi: 10.1088/0004-637X/737/1/12

—. 2011b, ApJ, 737, 12, doi: 10.1088/0004-637X/737/1/12

Leslie, S. K., Schinnerer, E., Liu, D., et al. 2020, ApJ, 899, 58, doi: 10.3847/1538-4357/aba044

Liang, L., Feldmann, R., Faucher-Giguère, C.-A., et al. 2018, MNRAS, 478, L83, doi: 10.1093/mnrasl/sly071

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

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119, doi: 10.1088/0004-637X/772/2/119

Limousin, M., Richard, J., Jullo, E., et al. 2016, A&A, 588, A99, doi: 10.1051/0004-6361/201527638

Linde, A. D. 1982, Physics Letters B, 108, 389, doi: 10.1016/0370-2693(82)91219-9

—. 1983, Physics Letters B, 129, 177, doi: 10.1016/0370-2693(83)90837-7

Liu, D., Daddi, E., Dickinson, M., et al. 2018, ApJ, 853, 172, doi: 10.3847/1538-4357/aaa600

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

Liu, D., Schinnerer, E., Groves, B., et al. 2019b, ApJ, 887, 235, doi: 10.3847/1538-4357/ab578d

Liu, D., Daddi, E., Schinnerer, E., et al. 2021, arXiv e-prints, arXiv:2101.06646. https://arxiv.org/abs/2101.06646

Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97, doi: 10.3847/1538-4357/837/1/97

Loveday, J., Peterson, B. A., Efstathiou, G., & Maddox, S. J. 1992, ApJ, 390, 338, doi: 10.1086/171284

Lovell, C. C., Vijayan, A. P., Thomas, P. A., et al. 2021, MNRAS, 500, 2127, doi: 10.1093/mnras/staa3360

Lundmark, K. 1930, Meddelanden fran Lunds Astronomiska Observatorium Serie I, 125, 1

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

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

Maddox, S. J., Valiante, E., Cigan, P., et al. 2018, ApJS, 236, 30, doi: 10.3847/1538-4365/aab8fc

Magdis, G. E., Elbaz, D., Daddi, E., et al. 2010, ApJ, 714, 1740, doi: 10.1088/0004-637X/714/2/1740

Magdis, G. E., Daddi, E., Elbaz, D., et al. 2011, ApJl, 740, L15, doi: 10.1088/2041-8205/740/1/L15

Magdis, G. E., Daddi, E., Béthermin, M., et al. 2012a, ApJ, 760, 6, doi: 10.1088/0004-637X/760/1/6

Magdis, G. E., Daddi, E., Sargent, M., et al. 2012b, ApJl, 758, L9, doi: 10.1088/2041-8205/758/1/L9

Magdis, G. E., Rigopoulou, D., Helou, G., et al. 2013, A&A, 558, A136, doi: 10.1051/0004-6361/201322226

Magdis, G. E., Rigopoulou, D., Daddi, E., et al. 2017, A&A, 603, A93, doi: 10.1051/0004-6361/201731037

Magdis, G. E., Gobat, R., Valentino, F., et al. 2021, arXiv e-prints, arXiv:2101.04700. https://arxiv.org/abs/2101.04700

Magnelli, B., Lutz, D., Santini, P., et al. 2012, A&A, 539, A155, doi: 10.1051/0004-6361/201118312

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

Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, A&A, 561, A86, doi: 10.1051/0004-6361/201322217

Magnelli, B., Boogaard, L., Decarli, R., et al. 2020, ApJ, 892, 66, doi: 10.3847/1538-4357/ab7897

Mahler, G., Richard, J., Clément, B., et al. 2018, MNRAS, 473, 663, doi: 10.1093/mnras/stx1971

Małek, K., Buat, V., Roehlly, Y., et al. 2018, A&A, 620, A50, doi: 10.1051/0004-6361/201833131

Man, A., & Belli, S. 2018, Nature Astronomy, 2, 695, doi: 10.1038/s41550-018-0558-1

Mancini, M., Schneider, R., Graziani, L., et al. 2015, MNRAS, 451, L70, doi: 10.1093/mnrasl/slv070

Mancuso, C., Lapi, A., Shi, J., et al. 2016, ApJ, 823, 128, doi: 10.3847/0004-637X/823/2/128

Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, MNRAS, 408, 2115, doi: 10.1111/j.1365-2966.2010.17291.x

Maraston, C., Pforr, J., Renzini, A., et al. 2010, MNRAS, 407, 830, doi: 10.1111/j.1365-2966.2010.16973.x

Matsuura, M., Barlow, M. J., Zijlstra, A. A., et al. 2009, MNRAS, 396, 918, doi: 10.1111/j.1365-2966.2009.14743.x

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

McKee, C. F., & Krumholz, M. R. 2010, ApJ, 709, 308, doi: 10.1088/0004-637X/709/1/308

McKinnon, R., Torrey, P., Vogelsberger, M., Hayward, C. C., & Marinacci, F. 2017, MNRAS, 468, 1505, doi: 10.1093/mnras/stx467

McLeod, D. J., McLure, R. J., Dunlop, J. S., et al. 2015, MNRAS, 450, 3032, doi: 10.1093/mnras/stv780

McLure, R. J., Dunlop, J. S., Bowler, R. A. A., et al. 2013, MNRAS, 432, 2696, doi: 10.1093/mnras/stt627

Merlin, E., Fontana, A., Ferguson, H. C., et al. 2015, A&A, 582, A15, doi: 10.1051/0004-6361/201526471

Merlin, E., Bourne, N., Castellano, M., et al. 2016, A&A, 595, A97, doi: 10.1051/0004-6361/201628751

Michałowski, M. J., Dunlop, J. S., Cirasuolo, M., et al. 2012, A&A, 541, A85, doi: 10.1051/0004-6361/201016308

Millard, J. S., Eales, S. A., Smith, M. W. L., et al. 2020, MNRAS, 494, 293, doi: 10.1093/mnras/staa609

Miville-Deschênes, M. A., Martin, P. G., Abergel, A., et al. 2010, A&A, 518, L104, doi: 10.1051/0004-6361/201014678

Moneti, A., McCracken, H. J., Rouberol, S., et al. 2019, The Fourth UltraVISTA data release, ESO

Moustakas, J., Coil, A. L., Aird, J., et al. 2013, ApJ, 767, 50, doi: 10.1088/0004-637X/767/1/50

Muñoz Arancibia, A. M., González-López, J., Ibar, E., et al. 2018, A&A, 620, A125, doi: 10.1051/0004-6361/201732442

Mullaney, J. R., Alexander, D. M., Goulding, A. D., & Hickox, R. C. 2011, MNRAS, 414, 1082, doi: 10.1111/j.1365-2966.2011.18448.x

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18, doi: 10.1088/0004-637X/777/1/18

Narayanan, D., & Davé, R. 2012, MNRAS, 423, 3601, doi: 10.1111/j.1365-2966.2012.21159.x

Narayanan, D., Krumholz, M., Ostriker, E. C., & Hernquist, L. 2011, MNRAS, 418, 664, doi: 10.1111/j.1365-2966.2011.19516.x

Narayanan, D., Krumholz, M. R., Ostriker, E. C., & Hernquist, L. 2012, MNRAS, 421, 3127, doi: 10.1111/j.1365-2966.2012.20536.x

Narayanan, D., Turk, M., Feldmann, R., et al. 2015, Nature, 525, 496, doi: 10.1038/nature15383

Newton, I. 1687, Philosophiae Naturalis Principia Mathematica. Auctore Js. Newton, doi: 10.3931/e-rara-440

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJl, 660, L43, doi: 10.1086/517926

Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793, doi: 10.1051/0004-6361/200912497

Obreschkow, D., & Rawlings, S. 2009, ApJ, 696, L129, doi: 10.1088/0004-637X/696/2/L129

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2015, ApJ, 808, 104, doi: 10.1088/0004-637X/808/1/104

—. 2013, ApJ, 773, 75, doi: 10.1088/0004-637X/773/1/75

—. 2014, ApJ, 786, 108, doi: 10.1088/0004-637X/786/2/108

Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, ApJ, 819, 129, doi: 10.3847/0004-637X/819/2/129

Oguri, M. 2010, PASJ, 62, 1017, doi: 10.1093/pasj/62.4.1017

Ojha, R., Stark, A. A., Hsieh, H. H., et al. 2001, ApJ, 548, 253, doi: 10.1086/318693

Okabe, T., Oguri, M., Peirani, S., et al. 2020, MNRAS, 496, 2591, doi: 10.1093/mnras/staa1479

Oke, J. B. 1974, ApJS, 27, 21, doi: 10.1086/190287

Oliver, S., Frost, M., Farrah, D., et al. 2010, MNRAS, 405, 2279, doi: 10.1111/j.1365-2966.2010.16643.x

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

Olsen, K., Greve, T. R., Narayanan, D., et al. 2017, ApJ, 846, 105, doi: 10.3847/1538-4357/aa86b4

Ono, Y., Ouchi, M., Kurono, Y., & Momose, R. 2014, ApJ, 795, 5, doi: 10.1088/0004-637X/795/1/5

Oort, J. H. 1932, Bull. Astron. Inst. Netherlands, 6, 249

Oteo, I., Zwaan, M. A., Ivison, R. J., Smail, I., & Biggs, A. D. 2016, ApJ, 822, 36, doi: 10.3847/0004-637X/822/1/36

Owen, F. N. 2018, ApJS, 235, 34, doi: 10.3847/1538-4365/aab4a1

Owers, M. S., Randall, S. W., Nulsen, P. E. J., et al. 2011, ApJ, 728, 27, doi: 10.1088/0004-637X/728/1/27

Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, ApJl, 698, L116, doi: 10.1088/0004-637X/698/2/L116

Pannella, M., Elbaz, D., Daddi, E., et al. 2015, ApJ, 807, 141, doi: 10.1088/0004-637X/807/2/141

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

Papadopoulos, P. P., Bisbas, T. G., & Zhang, Z.-Y. 2018, MNRAS, 478, 1716, doi: 10.1093/mnras/sty1077

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

Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, MNRAS, 412, 1123, doi: 10.1111/j.1365-2966.2010.17965.x

Pearson, W. J., Wang, L., Hurley, P. D., et al. 2018, A&A, 615, A146, doi: 10.1051/0004-6361/201832821

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266, doi: 10.1086/340952

Peng, Y., Maiolino, R., & Cochrane, R. 2015, Nature, 521, 192, doi: 10.1038/nature14439

Penner, K., Pope, A., Chapin, E. L., et al. 2011, MNRAS, 410, 2749, doi: 10.1111/j.1365-2966.2010.17650.x

Penzias, A. A., & Wilson, R. W. 1965, ApJ, 142, 419, doi: 10.1086/148307

Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59, doi: 10.1111/j.1365-2966.2004.07591.x

Pettini, M., Shapley, A. E., Steidel, C. C., et al. 2001, ApJ, 554, 981, doi: 10.1086/321403

Pillepich, A., Nelson, D., Springel, V., et al. 2019, MNRAS, 490, 3196, doi: 10.1093/mnras/stz2338

Pineda, J. L., Langer, W. D., Velusamy, T., & Goldsmith, P. F. 2013, A&A, 554, A103, doi: 10.1051/0004-6361/201321188

Pineda, J. L., Langer, W. D., Goldsmith, P. F., et al. 2017, ApJ, 839, 107, doi: 10.3847/1538-4357/aa683a

Piovan, L., Tantalo, R., & Chiosi, C. 2006, MNRAS, 366, 923, doi: 10.1111/j.1365-2966.2005.09732.x

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

Popescu, C. C., Misiriotis, A., Kylafis, N. D., Tuffs, R. J., & Fischera, J. 2000, A&A, 362, 138. https://arxiv.org/abs/astro-ph/0008098

Popescu, C. C., Tuffs, R. J., Dopita, M. A., et al. 2011, A&A, 527, A109, doi: 10.1051/0004-6361/201015217

Popping, G., Narayanan, D., Somerville, R. S., Faisst, A. L., & Krumholz, M. R. 2019, MNRAS, 482, 4906, doi: 10.1093/mnras/sty2969

Popping, G., Somerville, R. S., & Galametz, M. 2017, MNRAS, 471, 3152, doi: 10.1093/mnras/stx1545

Popping, G., Somerville, R. S., & Trager, S. C. 2014, MNRAS, 442, 2398, doi: 10.1093/mnras/stu991

Popping, G., Pillepich, A., Calistro Rivera, G., et al. 2021, arXiv e-prints, arXiv:2101.12218. https://arxiv.org/abs/2101.12218

Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25, doi: 10.1088/0067-0049/199/2/25

Pozzi, F., Calura, F., Zamorani, G., et al. 2020, MNRAS, 491, 5073, doi: 10.1093/mnras/stz2724

Press, W. H., Flannery, B. P., & Teukolsky, S. A. 1986, Numerical recipes. The art of scientific computing (Cambridge University Press)

Privon, G. C., Narayanan, D., & Davé, R. 2018, ApJ, 867, 102, doi: 10.3847/1538-4357/aae485

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

Rawle, T. D., Altieri, B., Egami, E., et al. 2016, MNRAS, 459, 1626, doi: 10.1093/mnras/stw712

Reddy, N., Dickinson, M., Elbaz, D., et al. 2012, ApJ, 744, 154, doi: 10.1088/0004-637X/744/2/154

Reddy, N. A., Erb, D. K., Pettini, M., Steidel, C. C., & Shapley, A. E. 2010, ApJ, 712, 1070, doi: 10.1088/0004-637X/712/2/1070

Reiter, M., Guzmán, A. E., Haworth, T. J., et al. 2020, MNRAS, 496, 394, doi: 10.1093/mnras/staa1504

Rémy-Ruyer, A., Madden, S. C., Galliano, F., et al. 2014, A&A, 563, A31, doi: 10.1051/0004-6361/201322803

Richard, J., Jauzac, M., Limousin, M., et al. 2014, MNRAS, 444, 268, doi: 10.1093/mnras/stu1395

Richard, J., Claeyssens, A., Lagattuta, D., et al. 2021, A&A, 646, A83, doi: 10.1051/0004-6361/202039462

Rickard, L. J., Palmer, P., Morris, M., Zuckerman, B., & Turner, B. E. 1975, ApJ, 199, L75, doi: 10.1086/181852

Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7, doi: 10.3847/1538-4357/aafc27

Riess, A. G., Strolger, L.-G., Tonry, J., et al. 2004, ApJ, 607, 665, doi: 10.1086/383612

Rigopoulou, D., Pereira-Santaella, M., Magdis, G. E., et al. 2018, MNRAS, 473, 20, doi: 10.1093/mnras/stx2311

Rigopoulou, D., Hopwood, R., Magdis, G. E., et al. 2014, ApJ, 781, L15, doi: 10.1088/2041-8205/781/1/L15

Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, ApJl, 739, L40, doi: 10.1088/2041-8205/739/2/L40

Rodighiero, G., Renzini, A., Daddi, E., et al. 2014, MNRAS, 443, 19, doi: 10.1093/mnras/stu1110

Rowan-Robinson, M., & Crawford, J. 1989, MNRAS, 238, 523, doi: 10.1093/mnras/238.2.523

Rubin, V. C. 1983, Scientific American, 248, 96, doi: 10.1038/scientificamerican0683-96

Ryan, R. E., J., Gonzalez, A. H., Lemaux, B. C., et al. 2014, ApJ, 786, L4, doi: 10.1088/2041-8205/786/1/L4

Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, ApJS, 233, 22, doi: 10.3847/1538-4365/aa97e0

Saintonge, A., Wilson, C. D., Xiao, T., et al. 2018, MNRAS, 481, 3497, doi: 10.1093/mnras/sty2499

Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267, doi: 10.1086/519218

Salmi, F., Daddi, E., Elbaz, D., et al. 2012, ApJl, 754, L14, doi: 10.1088/2041-8205/754/1/L14

Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, ApJ, 799, 183, doi: 10.1088/0004-637X/799/2/183

Salpeter, E. E. 1955, ApJ, 121, 161, doi: 10.1086/145971

Sanders, D. B., & Mirabel, I. F. 1996, ARA&A, 34, 749, doi: 10.1146/annurev.astro.34.1.749

Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86, doi: 10.1086/517885

Santini, P., Fontana, A., Grazian, A., et al. 2009, A&A, 504, 751, doi: 10.1051/0004-6361/200811434

Santini, P., Maiolino, R., Magnelli, B., et al. 2014, A&A, 562, A30, doi: 10.1051/0004-6361/201322835

Santini, P., Fontana, A., Castellano, M., et al. 2017, ApJ, 847, 76, doi: 10.3847/1538-4357/aa8874

Sargent, M. T., Daddi, E., Béthermin, M., et al. 2014, ApJ, 793, 19, doi: 10.1088/0004-637X/793/1/19

Sargsyan, L., Lebouteiller, V., Weedman, D., et al. 2012, ApJ, 755, 171, doi: 10.1088/0004-637X/755/2/171

Scalo, J. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 142, The Stellar Initial Mass Function (38th Herstmonceux Conference), ed. G. Gilmore & D. Howell, 201

Scalo, J. M. 1986, Fund. Cosmic Phys., 11, 1

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

Schmidt, K. B., Treu, T., Brammer, G. B., et al. 2014, ApJ, 782, L36, doi: 10.1088/2041-8205/782/2/L36

Schmidt, M. 1959, ApJ, 129, 243, doi: 10.1086/146614

—. 1968, ApJ, 151, 393, doi: 10.1086/149446

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

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

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

Scoville, N., Abraham, R. G., Aussel, H., et al. 2007b, ApJS, 172, 38, doi: 10.1086/516580

Scoville, N., Aussel, H., Sheth, K., et al. 2014, ApJ, 783, 84, doi: 10.1088/0004-637X/783/2/84

Scoville, N., Sheth, K., Aussel, H., et al. 2016, ApJ, 820, 83, doi: 10.3847/0004-637X/820/2/83

Scoville, N., Lee, N., Vanden Bout, P., et al. 2017, ApJ, 837, 150, doi: 10.3847/1538-4357/aa61a0

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

Shapley, A. E., Steidel, C. C., Adelberger, K. L., et al. 2001, ApJ, 562, 95, doi: 10.1086/323432

Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65, doi: 10.1086/373922

Shim, H., Chary, R.-R., Dickinson, M., et al. 2011, ApJ, 738, 69, doi: 10.1088/0004-637X/738/1/69

Shipley, H. V., Lange-Vagle, D., Marchesini, D., et al. 2018, ApJS, 235, 14, doi: 10.3847/1538-4365/aaacce

Shivaei, I., Reddy, N. A., Steidel, C. C., & Shapley, A. E. 2015, ApJ, 804, 149, doi: 10.1088/0004-637X/804/2/149

Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998, ApJ, 509, 103, doi: 10.1086/306476

Silverman, J. D., Daddi, E., Rodighiero, G., et al. 2015, ApJl, 812, L23, doi: 10.1088/2041-8205/812/2/L23

Silverman, J. D., Rujopakarn, W., Daddi, E., et al. 2018, ApJ, 867, 92, doi: 10.3847/1538-4357/aae25e

Simpson, J. M., Smail, I., Swinbank, A. M., et al. 2015, ApJ, 799, 81, doi: 10.1088/0004-637X/799/1/81

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24, doi: 10.1088/0067-0049/214/2/24

Slipher, V. M. 1917, Proceedings of the American Philosophical Society, 56, 403

Smith, G. P., Ebeling, H., Limousin, M., et al. 2009, ApJ, 707, L163, doi: 10.1088/0004-637X/707/2/L163

Smith, R. J., Glover, S. C. O., Clark, P. C., Klessen, R. S., & Springel, V. 2014, MNRAS, 441, 1628, doi: 10.1093/mnras/stu616

Smoot, G. F., Bennett, C. L., Kogut, A., et al. 1992, ApJ, 396, L1, doi: 10.1086/186504

Sobral, D., Best, P. N., Smail, I., et al. 2014, MNRAS, 437, 3516, doi: 10.1093/mnras/stt2159

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15, doi: 10.1088/0067-0049/214/2/15

Stacey, G. J., Townes, C. H., Poglitsch, A., et al. 1991, ApJ, 382, L37, doi: 10.1086/186208

Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJ, 462, L17, doi: 10.1086/310029

Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, ApJ, 795, 165, doi: 10.1088/0004-637X/795/2/165

Steinhardt, C. L., Speagle, J. S., Capak, P., et al. 2014, ApJl, 791, L25, doi: 10.1088/2041-8205/791/2/L25

Sternberg, A., & Dalgarno, A. 1989, ApJ, 338, 197, doi: 10.1086/167193

Strait, V., Bradač, M., Coe, D., et al. 2020, ApJ, 888, 124, doi: 10.3847/1538-4357/ab5daf

—. 2021, ApJ, 910, 135, doi: 10.3847/1538-4357/abe533

Symeonidis, M., Vaccari, M., Berta, S., et al. 2013, MNRAS, 431, 2317, doi: 10.1093/mnras/stt330

Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, MNRAS, 457, 2790, doi: 10.1093/mnras/stw131

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

Tacconi, L. J., Neri, R., Chapman, S. C., et al. 2006, ApJ, 640, 228, doi: 10.1086/499933

Tacconi, L. J., Genzel, R., Smail, I., et al. 2008, ApJ, 680, 246, doi: 10.1086/587168

Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, Nature, 463, 781, doi: 10.1038/nature08773

Tacconi, L. J., Neri, R., Genzel, R., et al. 2013, ApJ, 768, 74, doi: 10.1088/0004-637X/768/1/74

Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, ApJ, 853, 179, doi: 10.3847/1538-4357/aaa4b4

Tadaki, K.-i., Kodama, T., Tanaka, I., et al. 2014, ApJ, 780, 77, doi: 10.1088/0004-637X/780/1/77

Tadaki, K.-i., Genzel, R., Kodama, T., et al. 2017, ApJ, 834, 135, doi: 10.3847/1538-4357/834/2/135

Tamura, Y., Mawatari, K., Hashimoto, T., et al. 2019, ApJ, 874, 27, doi: 10.3847/1538-4357/ab0374

Tan, Q., Daddi, E., Magdis, G., et al. 2014, A&A, 569, A98, doi: 10.1051/0004-6361/201423905

Tasca, L. A. M., Le Fèvre, O., Hathi, N. P., et al. 2015, A&A, 581, A54, doi: 10.1051/0004-6361/201425379

Toomre, A., & Toomre, J. 1972, ApJ, 178, 623, doi: 10.1086/151823

Treu, T., Schmidt, K. B., Brammer, G. B., et al. 2015, ApJ, 812, 114, doi: 10.1088/0004-637X/812/2/114

Triani, D. P., Sinha, M., Croton, D. J., Pacifici, C., & Dwek, E. 2020, MNRAS, 493, 2490, doi: 10.1093/mnras/staa446

Tristram, K. R. W., Meisenheimer, K., Jaffe, W., et al. 2007, A&A, 474, 837, doi: 10.1051/0004-6361:20078369

Trumpler, R. J. 1930, PASP, 42, 214, doi: 10.1086/124039

Urry, C. M., & Padovani, P. 1995, PASP, 107, 803, doi: 10.1086/133630

Valentino, F., Magdis, G. E., Daddi, E., et al. 2018, ApJ, 869, 27, doi: 10.3847/1538-4357/aaeb88

—. 2020, ApJ, 890, 24, doi: 10.3847/1538-4357/ab6603

Vallini, L., Gallerani, S., Ferrara, A., Pallottini, A., & Yue, B. 2015, ApJ, 813, 36, doi: 10.1088/0004-637X/813/1/36 van der Wel, A., Chang, Y.-Y., Bell, E. F., et al. 2014, ApJl, 792, L6, doi: 10.1088/2041-8205/792/1/L6

Vazdekis, A. 1999, ApJ, 513, 224, doi: 10.1086/306843

Velusamy, T., & Langer, W. D. 2014, A&A, 572, A45, doi: 10.1051/0004-6361/201424350

Vijayan, A. P., Clay, S. J., Thomas, P. A., et al. 2019, MNRAS, 489, 4072, doi: 10.1093/mnras/stz1948

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

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

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

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

Wang, W.-C., Hirashita, H., & Hou, K.-C. 2017, MNRAS, 465, 3475, doi: 10.1093/mnras/stw2966

Watson, D., Christensen, L., Knudsen, K. K., et al. 2015, Nature, 519, 327, doi: 10.1038/nature14164

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

Webb, S. 1999, Measuring the universe : the cosmological distance ladder

Weigel, A. K., Schawinski, K., & Bruderer, C. 2016, MNRAS, 459, 2150, doi: 10.1093/mnras/stw756

Weingartner, J. C., & Draine, B. T. 2001, ApJ, 563, 842, doi: 10.1086/324035

Weiß, A., Henkel, C., Downes, D., & Walter, F. 2003, A&A, 409, L41, doi: 10.1051/0004-6361:20031337

Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJl, 754, L29, doi: 10.1088/2041-8205/754/2/L29

Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104, doi: 10.1088/0004-637X/795/2/104

Whitaker, K. E., Franx, M., Bezanson, R., et al. 2015, ApJl, 811, L12, doi: 10.1088/2041-8205/811/1/L12

Whitaker, K. E., Williams, C. C., Mowla, L., et al. 2021, Nature, 597, 485, doi: 10.1038/s41586-021-03806-7

White, S. D. M. 1978, MNRAS, 184, 185, doi: 10.1093/mnras/184.2.185

—. 1979, MNRAS, 189, 831, doi: 10.1093/mnras/189.4.831

White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341, doi: 10.1093/mnras/183.3.341

Wiklind, T., Conselice, C. J., Dahlen, T., et al. 2014, ApJ, 785, 111, doi: 10.1088/0004-637X/785/2/111

Wilkins, S. M., Vijayan, A. P., Lovell, C. C., et al. 2022, arXiv e-prints, arXiv:2204.09431. https://arxiv.org/abs/2204.09431

Williams, C. C., Curtis-Lake, E., Hainline, K. N., et al. 2018, ApJS, 236, 33, doi: 10.3847/1538-4365/aabcbb

Williams, C. C., Spilker, J. S., Whitaker, K. E., et al. 2020, arXiv e-prints, arXiv:2012.01433. https://arxiv.org/abs/2012.01433

Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, ApJ, 691, 1879, doi: 10.1088/0004-637X/691/2/1879

Wolfire, M. G., Hollenbach, D., & McKee, C. F. 2010, ApJ, 716, 1191, doi: 10.1088/0004-637X/716/2/1191

Worthey, G. 1994, ApJS, 95, 107, doi: 10.1086/192096

Wuyts, E., Rigby, J. R., Gladders, M. D., & Sharon, K. 2014, ApJ, 781, 61, doi: 10.1088/0004-637X/781/2/61

Wuyts, S., Labbé, I., Franx, M., et al. 2007, ApJ, 655, 51, doi: 10.1086/509708

Wuyts, S., Förster Schreiber, N. M., Genzel, R., et al. 2012, ApJ, 753, 114, doi: 10.1088/0004-637X/753/2/114

Wuyts, S., Förster Schreiber, N. M., Nelson, E. J., et al. 2013, ApJ, 779, 135, doi: 10.1088/0004-637X/779/2/135

Zahid, H. J., Bresolin, F., Kewley, L. J., Coil, A. L., & Davé, R. 2012, ApJ, 750, 120, doi: 10.1088/0004-637X/750/2/120

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

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

Zheng, W., Postman, M., Zitrin, A., et al. 2012, Nature, 489, 406, doi: 10.1038/nature11446

Zhukovska, S., Gail, H. P., & Trieloff, M. 2008, A&A, 479, 453, doi: 10.1051/0004-6361:20077789

Zitrin, A., Meneghetti, M., Umetsu, K., et al. 2013, ApJ, 762, L30, doi: 10.1088/2041-8205/762/2/L30

Zitrin, A., Zheng, W., Broadhurst, T., et al. 2014, ApJ, 793, L12, doi: 10.1088/2041-8205/793/1/L12

Zitrin, A., Fabris, A., Merten, J., et al. 2015, ApJ, 801, 44, doi: 10.1088/0004-637X/801/1/44

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110