Astrophysics Group
Department of Physics
Imperial College London

Multi-wavelength Investigations of *Herschel*-selected Dusty Star-Forming Galaxies

Yifan Ding

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy to Imperial College London
October 2024

Abstract

Dusty Star-Forming Galaxies (DSFGs) are found to dominate the cosmic star formation rate density at $z > 1$ and are much more populous at $z \sim 2$ than locally. Interstellar dust, as a by-product of star formation, absorbs about half of the UV and optical emission from stars and star formation and re-emit as infrared radiation. Therefore, dust-enshrouded DSFGs have star formation rates of tens to thousands of solar masses per year, leading to their extreme infrared luminosity of $\sim 10^{12-14} L_{\odot}$. Moreover, DSFGs are likely progenitors of current-day elliptical galaxies which are commonly found in the cores of massive galaxy clusters. They are thus believed to trace massive overdensities, especially those at $z > 4$. However, high redshift DSFGs and the environments they reside in still remain poorly understood, partially due to the lack of observations of high redshift DSFGs and their associated environments. This thesis presents a series of results which aim to better understand the nature of high redshift DSFGs and the environments they reside in. We first present 1.28GHz radio observations of three *Herschel-Planck* selected candidate DSFG protoclusters. With accurate positions acquired from the radio observations, we estimate the photometric redshifts of *Herschel* sources in the three fields, finding that there are potentially $z > 1$ overdensities aligning with two of the candidate protoclusters. We further present deep infrared observations of 13 *Herschel-Planck* selected candidate DSFG protoclusters. By constructing multi-wavelength photometry for sources extracted from our observations, we find that 6 of the 13 candidate DSFG protoclusters align with overdensities at $4 \leq z \leq 5$, suggesting they are real DSFG protoclusters at high redshifts. We additionally investigate the nature of 19 candidate $z > 4$ DSFGs, which are known as 500 μm risers as they have rising SEDs from 250 to 500 μm. We build multi-band photometry for these sources and derive their physical properties by fitting SED templates. We find a mean redshift of $< z >= 3.874 \pm 0.233$ for our sample with 10 of the 19 sources at $2.6 < z < 3.7$. The mean SFR and stellar mass of our sample are $4107 \pm 800\ M_{\odot} yr^{-1}$ and $< M_{*} >= 2.297 \pm 0.459 \times 10^{12}\ M_{\odot}$, with $L_{IR} > 10^{13}\ L_{\odot}$ for nearly all sources. Their SFR-stellar mass relation implies that either $z < 3.8$ 500 μm risers are starbursts which are more intense than typical star-forming galaxies at similar redshifts, or their redshifts are underestimated. Finally, we discuss potential follow-up studies based on the work begun here, and the outlook of DSFG studies in the future.

Acknowledgements

It has been quite a journey since I started my PhD programme during the pandemic. For all these years as a PhD student, I have been helped by so many people that it will need an entire thesis to credit all of them fully. Nevertheless, I would like to express my deepest gratitude to several people who are special to me.

The very first person that I'd like to thank is undoubtedly my supervisor, Dr. Dave Clements. Since the first time I met Dave, I have been constantly encouraged by his huge enthusiasm towards astronomy, supported by his superb expertise, and comforted by his endless patience and understanding. It will never be possible for me to complete this journey without Dave. Also, I would like to acknowledge my fellow PhD students in the Infrared group, Joe, Ayushi, and Wei, whose comments, answers, questions, and discussions have greatly helped me in both the works in this thesis and my understanding of the field. To those who started this journey with me: Alfie, Ethan, and Ellie, we all started in a quite peculiar fashion because of the pandemic, thank you for sharing your feelings, laughs, quizzes, journal clubs, secret seminars, and Christmas parties which have helped me settled in the Astro group. Of course, I also wish to thank all the other members of the Astro group who have created such a welcoming environment and made these four years a special and enjoyable experience.

A special acknowledgement must go to my fiancée, Renren Chen, who also started her PhD four years ago. I am most fortunate to have her company along this journey. For these four years, we have shared our experiences and thoughts as new PhD students, discussing scientific topics, and helping each other's works. Also, in our daily life, we have spent so much time together joking, cooking, travelling, playing badminton, and countless other things. Her love, compassion, and support have been such invaluable parts of this journey that I simply cannot imagine a PhD life without her any more.

Finally, I would like to thank my father Jiakui, and my mother Yonghong. It has never been an easy decision to support my study in a foreign country, but they have never questioned me for once. For all these years, they have shown an incredible level of understanding and support and provided comfort and guidance when I needed them the most. I hope you know that I will never achieve what I have achieved today without you, and it has been such a great comfort for me to know that you are always at home waiting for me.

Dedication

To my father and mother,

For your love and continuous support,

As my guiding light in the Universe.

'We live on an island surrounded by a sea of ignorance. As our island of knowledge grows, so does the shore of our ignorance.'

-John Archibald Wheeler

All of the content within this thesis is the result of the author's own work. Any work that is not original to the author is correctly referenced and attributed to its original source.

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Contents

Abstract **i**

Acknowledgements **iii**

1 Introduction **1**

1.1 A simple introduction to cosmology 1

1.2 Infrared astronomy 3

1.2.1 The Cosmic Infrared Background 3

1.2.2 Local infrared-luminous galaxies 5

1.3 Observations of the FIR sky 8

1.3.1 Atmospheric transmission 8

1.3.2 Single dish observations 9

1.3.3 Interferometry 10

1.3.4 The very negative K-correction 13

1.3.5 Instrumentation 15

1.4 Current Understanding Of Dusty Star Forming Galaxies 19

1.4.1 Number counts 20

1.4.2 Multiplicity of DSFG 22

1.4.3 The redshift distribution of DSFG 24

1.4.4 The luminosity functions of DSFGs 26

1.4.5 DSFG protoclusters 29

1.4.6 Radio identification of DSFGs 33

1.5 Summary of the following chapters 36

2 Herschel protocluster candidates: MeerKAT 1.28GHz observation 39

2.1 Introduction 39

2.2 Data 43

2.2.1 Target Selection 43

2.2.2 *Herschel* Data 43

2.2.3 Ancillary Data 47

2.2.4 MeerKAT 1.28GHz Data 47

2.3 Source Extraction 48

2.3.1 Extraction Software 48

2.3.2 Extraction Completeness & Reliability 50

2.3.3 MeerKAT 1.28GHz Source counts 55

2.4 *Herschel* Identifications 61

2.4.1 Multiplicity of *Herschel* Sources 64

2.4.2 The Far-IR Radio Correlation 68

2.4.3 Herschel Colours 74

2.5 Using the FIRC for *Herschel* cross-IDs 76

2.6 Accurate Optical/IR Cross-IDs of *Herschel* sources 81

2.6.1 Photometric Redshift Estimations 84

2.7 Conclusion 91

3 Deep Optical/NIR Observations of Candidate DSFG Protoclusters 94

3.1 Introduction & Motivation 94

3.2 Instrumentation and Observation 96

3.3 Data reduction 101

3.4 Source extraction & Photometry calibration 108

3.5 Complementary Optical/NIR data 117

3.5.1 DESI 117

3.5.2 KiDS & VIKING 118

3.5.3 DES 118

3.5.4 SHARKS 119

3.5.5 2dFGRS 119

3.5.6 Cross identification 119

3.6 Photometric redshift estimation 123

3.6.1 Validation of photometric redshifts 127

3.7 Evidence for DSFG protoclusters 128

3.8 Limitations of photo-z analysis 132

3.9 Conclusions 134

4 The Nature of 500μm Risers **136**

4.1 Introduction 136

4.2 Sample Selection 141

4.3 Photometry 142

4.3.1 Dark Energy Survey 145

4.3.2 DESI Legacy Imaging Survey 145

4.3.3 Sloan Digital Sky Survey 145

4.3.4 COSMOS2020 146

4.3.5 DeepDrill 146

4.3.6 HFLS_3 147

4.4 Photometric Redshifts 148

4.5 CIGALE SED fitting 153

4.6 Validating CIGALE fits 156

4.7 Physical Properties 157

4.7.1 IR Luminosity 159

4.7.2 SFR and Stellar Mass 160

4.7.3 The impact of AGN on star formation 164

4.8 Discussion 164

4.8.1 Reliability of Photometric Redshifts and SED Fittings 164

4.8.2 Gravitational Lensing in 500 μm Risers 166

4.8.3 Cosmos_9 167

4.9 Conclusions 168

5 Conclusion **172**

5.1 Summary of Thesis Achievements 172

5.2 Future Work 174

5.2.1 Identification of DSFG protoclusters with high-quality radio data 174

5.2.2 DSFGs and their environments in the optical/NIR 175

5.3 Concluding remarks 176

Bibliography **179**

List of Tables

2.1 Names, positions, associated secondary calibrators, and total on-source time for each of the protocluster candidate target fields. 48

2.2 The 1.28GHz source counts in this work 58

3.1 Names, positions, EFOSC2/SOFI filter used and EFOSC2/SOFI observation date for each candidate protocluster observation 99

3.2 Key SOURCE EXTRACTOR parameters used to extract sources from our SOFI images . 113

3.3 Summary of the photometry available for each target and the source of the data . . . 122

3.4 Density contrast of each SOFI target field at $4 < z < 5$ 130

4.1 The position and integrated FIR/sub-mm photometry for the target sources 142

4.2 List of optical and NIR/MIR surveys used to build the multi-wavelength photometry in this chapter. 147

4.3 List of photometric redshifts of our sources given by EAZY, MMPZ and CIGALE . . 150

4.4 CIGALE modules and parameters used in this work 155

4.5 IR luminosities, stellar luminosities, SFR, stellar masses, dust masses and AGN fractions of our targets derived by CIGALE 159

List of Figures

1.1 Cosmic bolometric background energy density as a function of wavelength 4

1.2 A schematic diagram outlining the evolutionary path of a current epoch massive elliptical galaxy 7

1.3 The atmosphere transmission as a function of wavelength as seen from Mauna Kea, Hawaii 9

1.4 A simple diagram of a two antenna interferometer 11

1.5 The observed flux densities vs. redshift for a typical IR bright galaxy 16

1.6 Differential number counts at 70, 100, 170, 250, 350μm and 1.1mm from various studies 21

1.7 The photometric redshift distribution of 850 μmselected SMGs 25

1.8 Measured integrated IR luminosity functions in four redshift bins 28

2.1 The MeerKAT 1.28GHz image of G014 presented in this work 44

2.2 An example of a faint source missed by the first run of PyBDSF 49

2.3 (a) Extraction completeness as a function of 1.28GHz flux. (b) The reliability of the source extraction with SExtractor and PyBDSF. (c) Flux distributions of sources extracted from both the original maps and the flux-inverted maps. (d) Average total flux of injected sources recovered by SExtractor and PyBDSF vs. average total flux of sources injected into the map in Jy 51

2.4 An example of a faint source detected by SExtractor but not PyBDSF 52

2.5 MeerKAT image cutouts 53

2.6 The average area of injected sources (in $arcsecond^2$) extracted by SExtractor and PyBDSF as a function of flux density 56

2.7 Brightness-weighted differential source counts as a function of flux 59

2.8 Number of sources in G257 per steradian as a function of flux 60

2.9 Brightness-weighted differential 1.28GHz source counts in field centres 61

2.10 Number of *Herschel* sources with real MeerKAT counterparts / missed MeerKAT counterparts / random identification as a function of matching radius 63

2.11 The distribution of the number of MeerKAT matched to *Herschel* sources 66

2.12 Flux density distributions of *Herschel* sources with/without 1.28GHz MeerKAT counterparts 68

2.13 $q_{250\mu m}$ against the 250 μm flux density for all sources in the three maps 70

2.14 $q_{250\mu m}$ against the 250 μm flux density for all sources, central sources differentiated . 71

2.15 $q_{250\mu m}$ against the 250 μm flux density for *Herschel* sources with only one 1.28GHz counterparts and *Herschel* sources with multiple 1.28GHz counterparts 73

2.16 *Herschel* S_{250}/S_{350} and S_{350}/S_{500} colours of our MeerKAT sources 74

2.17 *Herschel* S_{250}/S_{350} and S_{350}/S_{500} colours of our MeerKAT sources, divided into subsamples 75

2.18 The $q_{250\mu m}$ against the 250μm flux density for *Herschel* sources matched to simulated MeerKAT sources, with a matching radius of 7.5″ 77

2.19 The $q_{250\mu m}$ against the 250μm flux density for *Herschel* sources matched to simulated MeerKAT sources, with a matching radius of 17.6″ 78

2.20 Two examples of MeerKAT data helping *Herschel* identifications with SHARKS sources 81

2.21 Colour-colour diagram between Ks, 250μm and 1.28GHz 82

2.22 The KiDS DR4 catalogue footprint in G014 85

2.23 $10' \times 10'$ cutouts of the centres of our MeerKAT 1.28GHz images, and positionally matched KiDS DR4 r band images and H-ATLAS *Herschel* 250μm images 86

2.24 Photometric redshift distributions of *Herschel* sources identified in both MeerKAT and KiDS/VIKING in the three observed fields 88

2.25 Photometric redshift distributions of all KiDS sources identified in the three observed fields 88

2.26 Density contrast maps of each observed field 90

3.1 The ESO 3.6 m New Technology Telescope 97

3.2 An example of the flat-field correction 104

3.3 An example bad pixel mask 106

3.4 An example of background subtraction 107

3.5 The final science images of all targets in every observed filter band 109

3.6 The average magnitude differences for ZP derived from 2MASS catalogue magnitudes 116

3.7 The AB magnitude distributions of sources in all target fields 116

3.8 DESI-SOFI matching percentage as a function of matching radius 121

3.9 An example of a SOFI J-band detected source with no apparent r-band counterpart . 122

3.10 Comparison of photo-z and spec-z for sources in the blank field 125

3.11 Median of the EAZY photo-z uncertainty of SOFI-detected sources as a function of photo-z 126

3.12 Photometric redshift distributions of sources in each target field 129

3.13 *Herschel* S_{250}/S_{350} and S_{350}/S_{500} colours of *Herschel* sources inside the 6 SOFI target fields 131

4.1 Examples of the SMA maps of sources presented in this chapter 143

4.2 Multi-wavelength cutout image of XMM_M7 as an example of our cross-identifications 144

4.3 Comparison between spectroscopic redshifts and photometric redshifts derived by EAZY, MMPZ and CIGALE, for sources with spectroscopically confirmed redshifts . 151

4.4 Redshift distribution of our 19 targets 152

4.5 CIGALE best-fit SEDs for XMM_M7, HeLMS_783, HFLS_3 and Bootes_33 as examples 156

4.6 Comparison between the results of the original run and the mock run 158

4.7 The total IR luminosities of our sample as a function of redshift without correction for lensing 160

4.8 SFR vs.stellar mass in different redshift bins 161

4.9 SFR vs.stellar mass for all our targets 163

4.10 SFR as a function of redshift 165

4.11 CIGALE best-fit SEDs for Cosmos_9 168

4.12 The optical i band image of Cosmos_9 from the HSC Public Data Release 3 Ultra Deep image (Aihara et al., 2022) 169

Chapter 1

Introduction

1.1 A simple introduction to cosmology

Since the first day of the human civilisation, countless attempts have been made to understand the Universe we live in. Based on these remarkable efforts, our current understanding is that the Universe is flat, homogenous and isotropic on large scales, and it is expanding at an accelerating rate (e.g., Hubble 1929; Riess et al. 1998; Ade et al. 2016; Aghanim et al. 2020). Numerous hypotheses have been proposed to incorporate all observations of the Universe. Among these models, the one that is now mostly accepted and widely used is the Λ-CDM model. This model simply describes the Universe as a result of three components. First of all is an hypothetical form of energy known as dark energy which is closely associated with the cosmological constant Λ. It is thought to be responsible for the accelerating expansion of the Universe. Then, there is a theoretical substance called dark matter, with "CDM" in the name of the model stands for cold (non-relativistic) dark matter. Dark matter is thought to weakly interact with other particles only through gravity and probably the weak force, a characteristic known as collisionless, hence its name "dark" matter. The last component is the baryonic matter which corresponds to matter that consists of protons and neutrons. The recent 2020 data release of the Planck Collaboration (Aghanim et al., 2020) measures that dark energy contributes $\sim 68.5\%$ of the energy density of the Universe, while dark matter and baryonic matter together make up $\sim 31.5\%$ of the energy density of the baryonic Universe.

According to the Λ-CDM model[1], the Universe originated ∼ 13.8 billion years ago (Aghanim et al., 2020) after the Big Bang. Initially, the Universe was hot, uniform, dense and compact, filled with a plasma of relativistic particles. This was immediately followed by a period of rapid expansion and cooling, called inflation. About 2-3 minutes after the Big Bang, the Universe was cool enough for particles such as protons and neutrons to decouple from the hot relativistic plasma, allowing them to combine into the first elements such as Helium and Lithium. With further cooling, this was followed by the combination of protons and electrons to form electronically neutral hydrogen atoms, a process known as recombination. As most free electrons in the hot plasma were gradually captured by protons to form hydrogen atoms, photons in the hot plasma were no longer bounded to the plasma by electrons' Thomson scattering and began to travel freely in the Universe, which are now observed as the cosmic microwave background (CMB). The CMB is found to have very small anisotropies, which are thought to originate from small perturbations of energy density in the Universe before the inflation. These small density fluctuations were later enlarged by inflation and became overdense regions. Such overdensities became increasingly dense by attracting nearby matter, while also expanded spatially due to the expansion of the Universe. This resulted in a slower expansion rate of the overdense regions than that of the Universe. Eventually, the densities of these regions reached a critical value such that they started to collapsed inward under their own gravity, breaking away from the universal expansion. The dark matter particles in these collapsing regions were not able to cool down by emitting photons, as dark matter only interacts via gravity and probably the weak force, therefore they eventually virialised into dark matter halos. The baryons, on the other hand, were collisional and were able to cool and further collapse into proto-stellar cores. Nuclear fusion was ignited by pressure in such cores, becoming the seeds of the first stars and thus the first galaxies. In the Λ-CDM model, the largest density perturbations in the primordial Universe give rise to the present days galaxy clusters (e.g., Kravtsov and Borgani 2012).

Despite the remarkable simplicity and accuracy of the Λ-CDM model, there are still some problems which cannot be explained by the current model such as the Hubble tension, the Horizon and Flatness problems (e.g., Guth 1981; Bullock and Boylan-Kolchin 2017; Perivolaropoulos and Skara 2022), and more crucially, the lack of direct evidence of both dark energy and dark matter. Better understanding

[1]The majority of the following descriptions of the evolution of the Universe comes from Frenk and White (2012); Peebles (2012); Primack (2012); Deruelle and Uzan (2018) and references therein.

of the evolution of galaxies and galaxy clusters will pose more constraints on the Λ-CDM model and eventually improve the understanding of our Universe. This is the ultimate goal of the studies in this thesis.

1.2 Infrared astronomy

1.2.1 The Cosmic Infrared Background

Over the last century, optical and ultraviolet (UV) observations were the main focus of astronomy. It was, however, not until the early 1990s that astronomers realised that the Universe has comparable amounts of radiation in the far-infrared (FIR) and submillimetre (submm) wavelengths since the launch of the COsmic Background Explorer (COBE), equipped with the Far-InfraRed Absolute Spectrophotometer (FIRAS) and The Diffuse Infrared Background Experiment (DIRBE), which detected the diffuse cosmic infrared background radiation (CIB, Puget et al. 1996; Fixsen et al. 1998; Hauser et al. 1998; Cooray 2016). The CIB covers the 10-1000 μm wavelength range with an energy density of $30 \pm 10\ nWm^{-2}sr^{-1}$, while its optical counterpart, the cosmic optical background which radiates between 0.1 and 5 μm (COB, Cooray 2016), has a similar energy density of $24 \pm 4\ nWm^{-2}sr^{-1}$. Together, they account for roughly 5% of the total bolometric background energy density, with the well-studied CMB being responsible for the other 95% (Cooray 2016, Figure 1.1). In our current understanding, optical/UV radiation primarily originates from direct stellar emission and active galactic nuclei (AGN), while the infrared (IR) fluxes result mainly from interstellar dust absorbing optical/UV radiation and re-emitting these energies at mid/far-infrared wavelengths, with AGN contributing 10% in both wavelength ranges (Hauser and Dwek 2001; Andrews et al. 2018). Dust in the local universe is found to consist of silicates and carbonaceous grains, with polycyclic aromatic hydro-carbons (PAHs) containing up to 15-20% of the carbon in the interstellar medium (ISM) (Draine 2003a; Zubko et al. 2004). This dust has a peak absorption at roughly 100 *nm* in the rest frame, with photons with wavelengths larger than 300 *nm* have a greater chance to be scattered by dust than absorbed (Draine 2003a). The fact that the CIB and COB have comparable energy densities implies that around half of the stellar emission is absorbed and re-emitted in the infrared by dust over the entire history of the

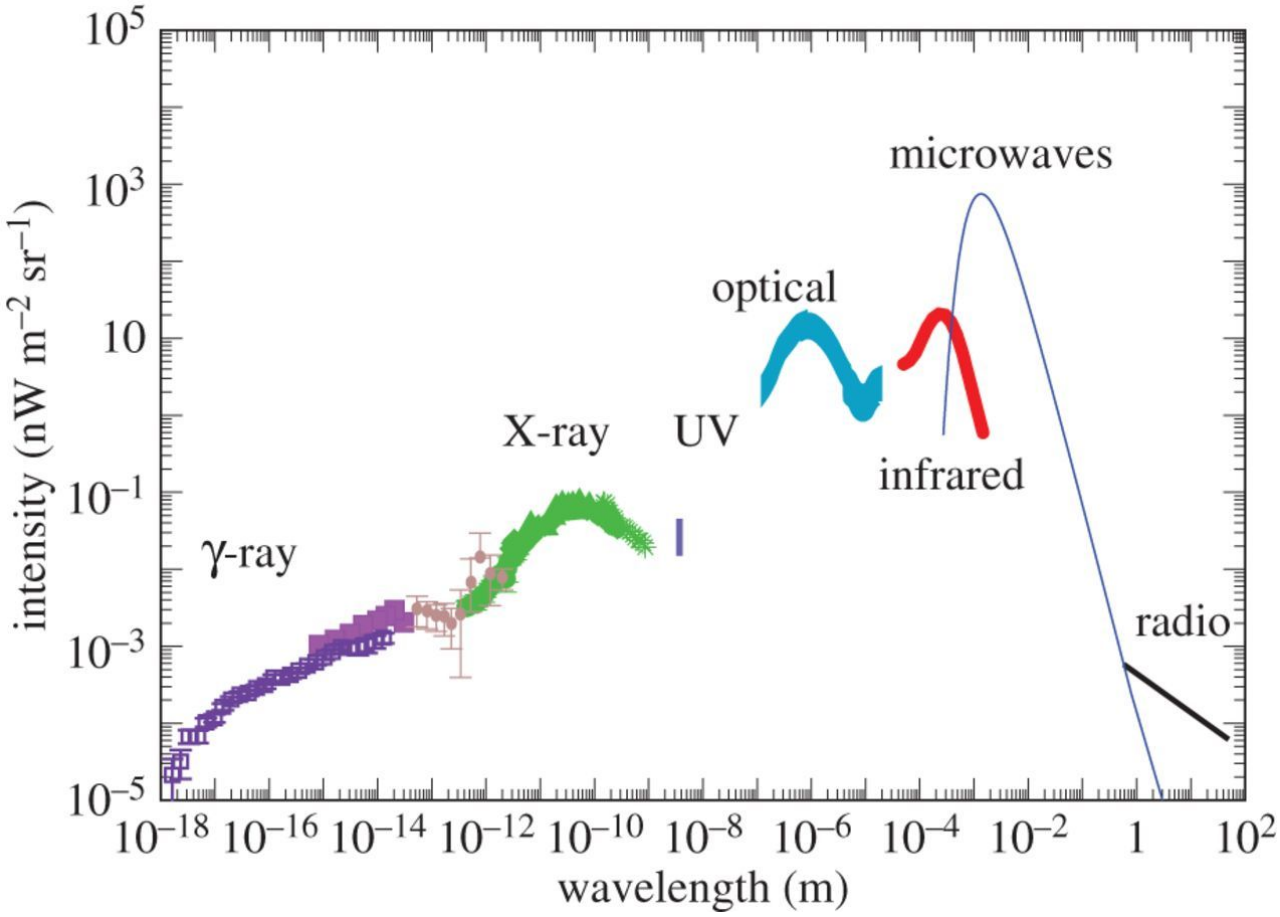


Figure 1.1: Cosmic bolometric background energy density as a function of wavelength. This figure is adopted from Cooray (2016). As shown by the figure, while the microwave background contribution to the cosmic background is the greatest ($\sim$ 95%), the COB and CIB together take account for around 5% of the cosmic background with roughly equal contributions.

Universe.

However, observations of the local Universe find a very different scenario: the optical/UV energy density is three times larger than the IR energy density (Soifer and Neugebauer 1991). This in fact implies that the IR emission must have a much larger contribution in the high redshift Universe. In addition, the slope of the CIB at submm wavelength is especially shallow, with $I_\nu \propto \nu^{1.4}$ (Gispert et al. 2000) where I_ν is the intensity per unit frequency, ν is the frequency. In contrast, the Rayleigh-Jeans tail of blackbody radiation gives $I_\nu \propto \nu^2$, and variations on such blackbodies are often used to model the dust thermal emission from individual galaxies (Blain et al. 2002; Casey 2012). This means that sources which are responsible for the CIB peak emission at 100 μm are a different population to sources responsible for submm and mm emission, which are likely at higher redshifts.

Combining the facts stated above, the IR emission of the Universe must have been evolving faster than optical/UV emission since $z = 1$, and it must also have remained constant from $z = 1$ to at least $z = 3$ (Gispert et al. 2000; Lagache et al. 2005). As we have discussed above, the IR radiation is related to the dust absorption of optical/UV emission, and the optical/UV emission originates from star formation activity (since only the short-life, newly-formed massive stars can emit significant amount of optical/UV radiation) which is then reprocessed into IR radiation by dust. This, taken together with the increase of cosmic IR emission with redshift, suggests that dusty star-forming galaxies (DSFGs) probably have played an important role in the cosmic star-formation history at $z > 1$. Some studies have shown that nearly half of star-formation activity is invisible to pure optical and UV observations (e.g. Lonsdale et al. 2006)

It is then natural to investigate the properties of these IR bright DSFGs in order to get a more complete picture of the star-formation history of the Universe. To achieve this, we need samples of FIR/submm bright galaxies that are faint in the optical/UV and located at $z > 1$.

1.2.2 Local infrared-luminous galaxies

Before we discuss the high-redshift DSFG population, it is necessary to have some knowledge of local Luminous InfraRed Galaxies (LIRGs), which are defined as galaxies with their infrared luminosity (defined as the integrated rest frame luminosity of a source between 8 and 1000 μm) $L_{IR} > 10^{11} L_{\odot}$ (e.g., Rowan-Robinson 2000; Lonsdale et al. 2006). They are local analogues to the high-redshift DSFGs, and hence are excellent samples to investigate the possible physical properties of DSFGs. This local infrared bright galaxy population was first discovered by the *InfraRed Astronomy Satellite* (IRAS, Neugebauer et al. 1984). Among various studies of the IRAS extragalactic sources, the most well-studied samples include the Revised Bright Galaxy Sample (RBGS, Sanders et al. 2003) which contains 629 extremely IR bright sources, the 1.2 Jy IRAS Redshift Survey (Strauss et al., 1992; Fisher et al., 1995) which provides redshifts of 5320 60 μmgalaxies, and the IRAS Point Source Catalog Redshift (PSCz) Survey (Saunders et al., 2000) which contains 18351 IRAS sources with flux $S_{60\,\mu\mathrm{m}} > 0.595$ Jy. The Infrared Space Observatory (ISO) satellite with ISOPHOT aboard provided extra populations of local IR bright galaxies (Lemke et al. 1996). The majority of these

samples mentioned above are LIRGs with infrared luminosities $10^{11} < L_{IR} < 10^{12} L_{\odot}$, with a small fraction being Ultra Luminous InfraRed Galaxies (ULIRGs) with $10^{12} < L_{IR} < 10^{13} L_{\odot}$. Detailed investigation (including morphologies, molecular gas and dust emission etc.) of these samples show that most galaxies with $L_{IR} > 10^{11.5} L_{\odot}$ or star formation rates (SFRs) larger than 50 $M_{\odot} yr^{-1}$ are major mergers of two or more equal-mass galaxies (Sanders and Mirabel 1996; Clements et al. 1996). They also appear to be rich in molecular gas, with molecular gas mass estimated between 10^9 – 10^{10} $M_{\odot}$ (Sanders et al. 1991).

Putting things together, the local (U)LIRGs seem to be a set of IR bright major mergers with very high SFRs which are powered by enormous amounts of gas drifting towards the core of merging galaxies during the merger process and forming stars. With these observations, it is now widely accepted that (U)LIRGs are part of the evolutionary path of extremely luminous systems towards the massive elliptical galaxies observed today (Hopkins et al. 2008; Casey et al. 2014). Figure 1.2 shows this hypothetical evolutionary path. In this hypothesis, (U)LIRGs are at a stage where two gas-rich disk galaxies collide with each other, resulting in an extreme star formation event by rapid compression and cooling of gas. This collision and the following intense star formation process produce numerous dust grains, which will then absorb rest-frame optical/UV radiation emitted by newly-formed stars and re-emit them at IR/submm wavelengths. Galaxies at this stage hence appear as ULIRGs/LIRGs to us. This (U)LIRG phase is thought to be relatively short in duration, as there is no extra supply to the gas reservoir while the SFR is high. At a certain point, the cores of the two galaxies will eventually merge, forming a new AGN which is further fuelled by in-falling material in its accretion disk. Before the depletion of gas, the galaxy looks like a bright AGN (or quasar), but when it finally runs out of gas, no stars will form and it becomes an elliptical galaxy which is characterised by a virialised state and an old stellar population (Hopkins et al. 2008; Casey et al. 2014).

Although this theory of the evolution of local IR bright galaxies looks very convincing, its significance is diminished by the fact that (U)LIRGs only make up a very small fraction of the local galaxy population. Therefore they probably contribute little to the cosmic star formation rate density at $z = 0$. However, as discussed in the previous section, IR luminous galaxies should be far more numerous in the early Universe. This may imply that the cosmic star formation in the early Universe mainly comes

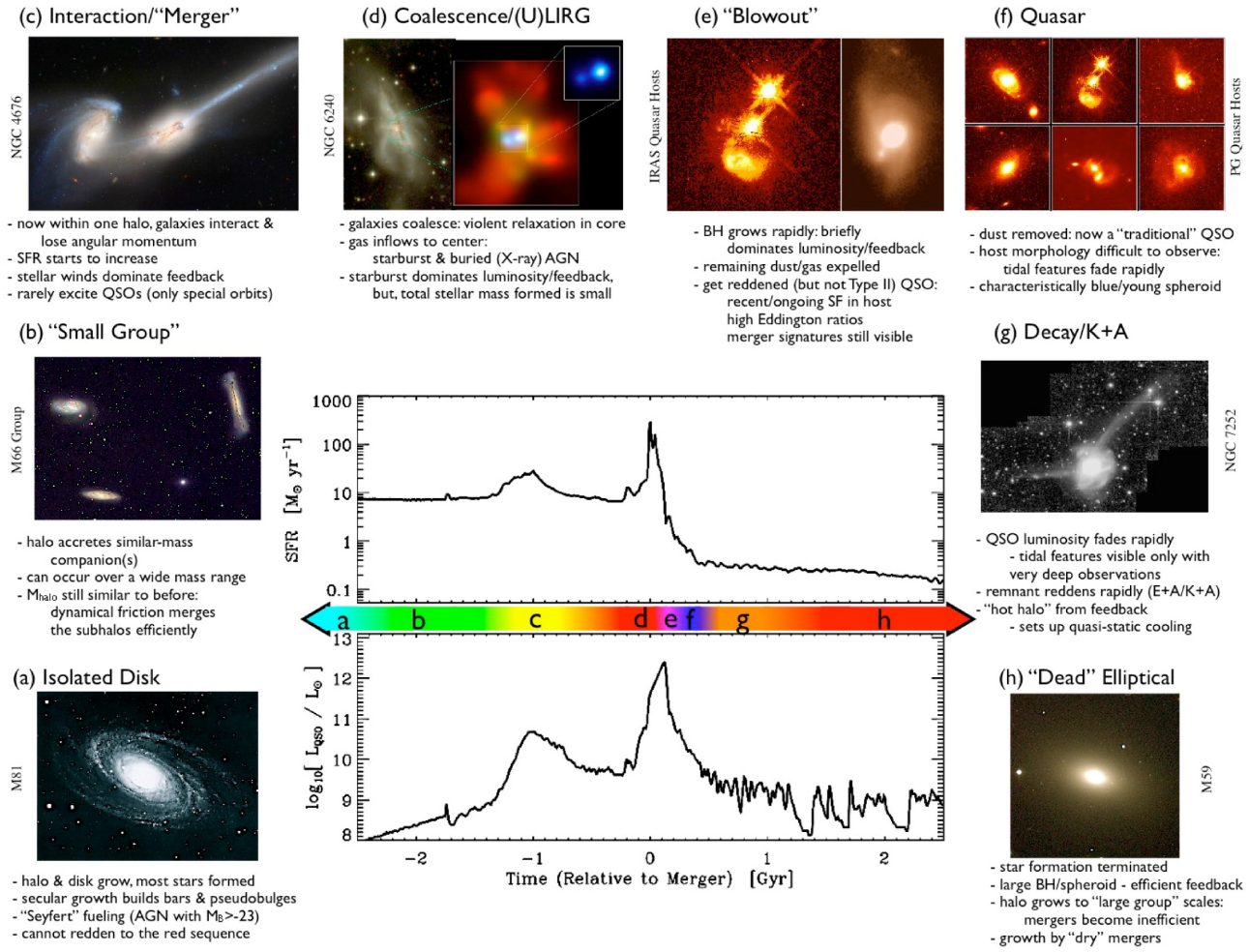


Figure 1.2: A schematic diagram outlining the evolutionary path of a current epoch massive elliptical galaxy. This figure is adopted from Hopkins et al. (2008).

from major mergers, or there is another unknown evolutionary path for these galaxies. Thus, naturally, studies of local (U)LIRGs turn towards investigations of high-redshift DSFGs.

1.3 Observations of the FIR sky

The dusty nature of DSFGs makes them very faint or even invisible in the optical/NIR, as the stellar emission and AGN emission of galaxies are absorbed by dust and re-emitted into the IR, especially the FIR and submm wavelength where the redshifted dust emission peaks. Therefore, observations in the FIR and submm regime are vital to DSFG studies. In this section, we outline the major advantages and challenges of FIR/submm observations.

1.3.1 Atmospheric transmission

Since the early 20^{th} century, it has been known that the atmosphere is almost completely opaque in the infrared, mainly due to the absorption of water vapour plus a number of other molecules such as CO_2 and O_3 in the atmosphere (e.g., Gebbie et al. 1951; Howard et al. 1956a,b; Taylor and Yates 1957). In particular, the FIR and submm suffer severely from the absorption via the rotational transitions of water vapour. Therefore, ground-based infrared observatories are usually sited in dry places where the precipitable water vapour (PWV) is constantly very low. Examples of such places are the summit of Mauna Kea in Hawaii where the James Clerk Maxwell Telescope (JCMT) and the Submillimeter Array (SMA) are based, and the Atacama desert in Chile where the Atacama Large Millimetre/submillimeter Array (ALMA) is based. Figure 1.3 shows the atmospheric transmission curve as a function of wavelength at the summit of Mauna Kea. It can be seen that there are only a few narrow spectral windows in the FIR which are free of absorption, and the $850\mu m$ atmospheric window is the most accessible wavelength for most situations. However, a typical DSFG has its FIR emission peaking at about $100\mu m$ in the rest frame (Casey et al., 2014). Finding this peak in the observed frame will thus provide a strong constraint on the photometric redshift estimation of the galaxy. This means that for DSFGs at $z < 8$ observations in the $850\mu m$ band will miss the peak of their FIR emission.

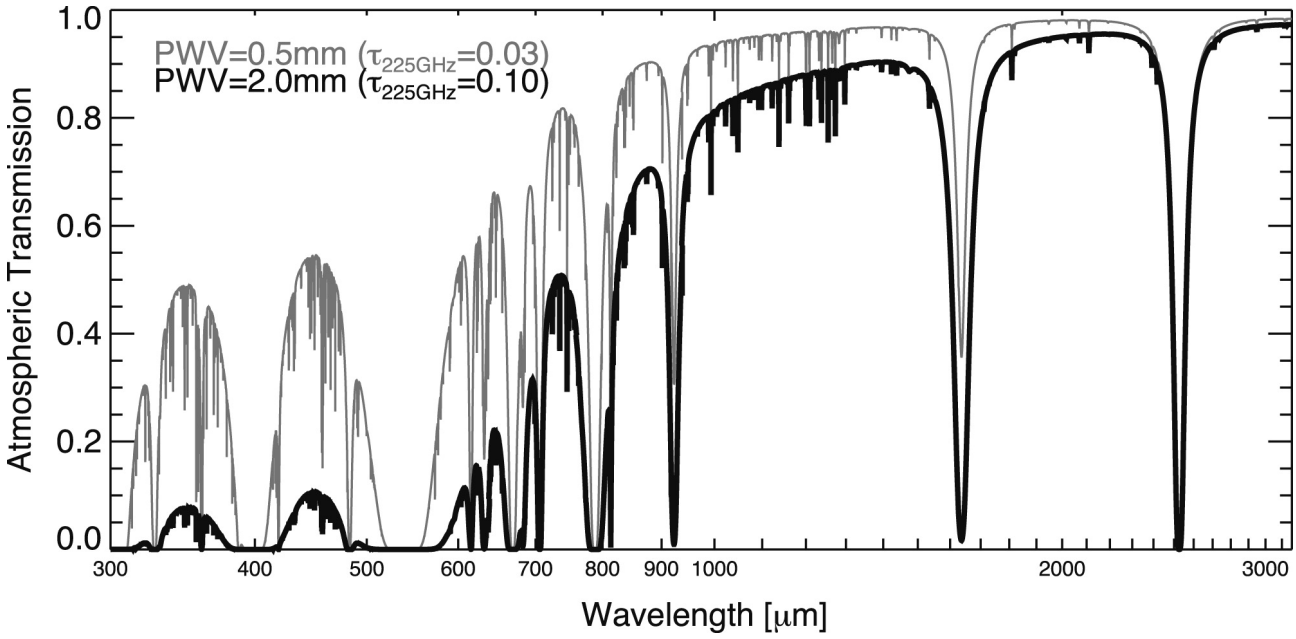


Figure 1.3: The atmosphere transmission as a function of wavelength as seen from Mauna Kea, Hawaii. Two PWV levels at 2.0mm and 0.5mm, corresponding to two different optical depth τ, are plotted. This figure is adopted from Casey et al. (2014)

This limitation on ground-based FIR/sbumm observations strongly constrained early DSFG studies, and was not resolved until the launch of the space-based *Spitzer* and especially *Herschel* missions.

1.3.2 Single dish observations

Beside the limitation brought by atmospheric opacity, there is another outstanding difficulty in FIR/submm astronomy which is the limitation of image resolution of single dish observations. For a typical single dish telescope with a circular aperture, the minimum angular distance between two resolved nearby point sources is defined by the Airy disk as:

$$\theta = 1.22\frac{\lambda}{D} \tag{1.1}$$

where θ is the minimum angular distance between two resolved sources (i.e. the angular resolution), λ is the observation wavelength and D is the diameter of the telescope's circular aperture. From this equation, it can be seen that observations at longer wavelengths will have lower angular resolution for the same aperture diameter, and sources that can be resolved at shorter wavelength may be seen as one unresolved source at longer wavelength. This downgrade of angular resolution has two direct

consequences in FIR/submm observations.

Firstly, the worse angular resolution means that there may be a number of unresolved sources inside the large beamsize of an instrument, or small-scale structures in the infrared cirrus emission. The fluxes of these unresolved sources or structures will dominate the background sky brightness of an observation taken by this instrument. The fluctuations of the background sky brightness below which sources cannot be detected individually is commonly known as the confusion noise (e.g., Scheuer 1974; Condon 1974; Dole et al. 2004). Because the number of sources inside a beam is entirely stochastic and the confusion noise comes from real sources, the confusion noise will vary from regions to regions and cannot be reduced like instrumental noise. There is no effective method to fully take account for the confusion noise. This inevitable confusion noise thus sets a hard detection limit, below which it is impossible to detect individual sources regardless the integration time.

Secondly, multiple faint objects within the large beam may be blended into one single source with extremely high apparent IR luminosity. This phenomenon of IR objects is known as multiplicity, and it makes the cross-identification of FIR/submm sources in the optical/NIR very challenging. Due to the high number density of optical/NIR sources, it is common to find several optical/NIR sources overlapping with a FIR/submm source. It is thus hard to know which optical/NIR sources are responsible for the FIR/submm emission and the contributions of individual counterparts are difficult to determine, or there may even be undetected optical/NIR sources which contribute to the FIR/submm flux. The multiplicity of FIR/submm sources is still a topic under intense discussion, more details will be discussed in Section 1.4.2.

1.3.3 Interferometry

As discussed in the previous section, FIR/submm studies are greatly limited by the poor angular resolution of typical single dish instruments. The most direct solution to this problem is to build a telescope with a larger aperture, in other words, a larger dish. However, it is very difficult to achieve, due to the limitations of current technology. An alternative solution to increase the size of the aperture is to use interferometers, which are arrays of antennas observing the target simultaneously to achieve

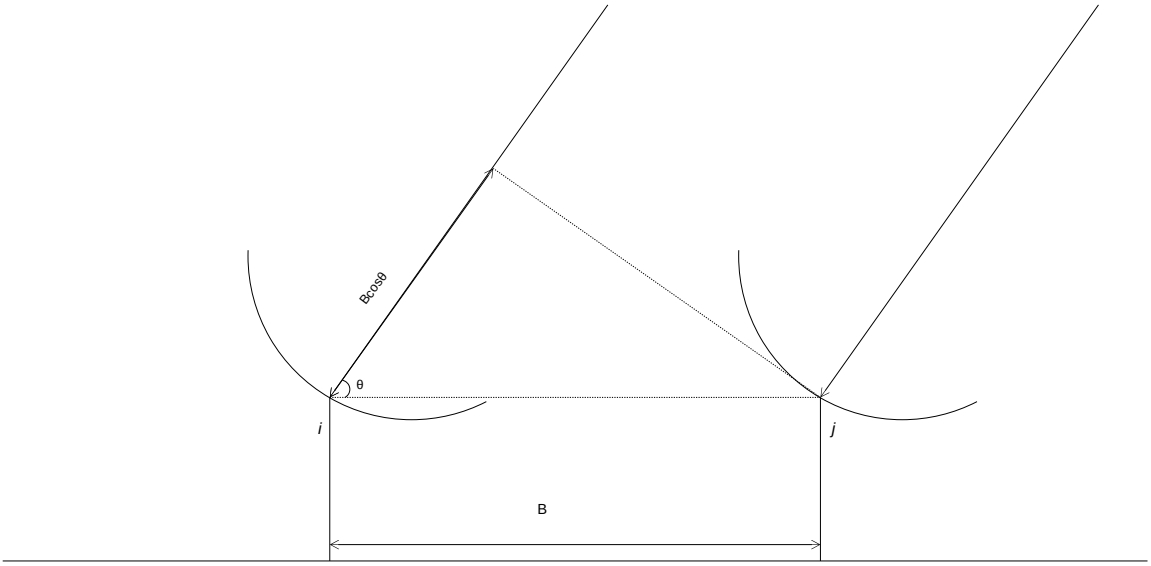


Figure 1.4: A simple diagram of a two antenna interferometer. The two antennas i and j is separated by the baseline B and pointing towards the sky with an angle of θ above the horizon.

very high angular resolutions. As such, each pair of antennas will act as an aperture with a diameter as large as the distance between the two antennas (known as the baseline), the angular resolution of each antenna pairs is then

$$\theta = 1.22\frac{\lambda}{B} \tag{1.2}$$

where B is the baseline. As it is technically easier to place two antennas with a large separation distance than to build a single telescope dish of the same size, interferometers can provide much higher angular resolutions at longer wavelength. A significant part of this thesis uses interferometric data, therefore we will very briefly introduce the basic concept of interferometry in this section[2].

Imagine the simplest case of two antennas i and j, as shown in Figure 1.4. Both of them point to the sky in the same direction with an angle of θ above the horizon. Light emitted by an astronomical object will arrive at the antennas as plane waves. As the two antennas are separated at a distance B, antenna j will detect the light of this object before the other, and the same light have to travel $Bcos\theta$ more distance to arrive antenna i. This will result in a time difference between the signals of the two

[2]The bulk of this section is based on "Interferometry and synthesis in radio astronomy" (Thompson et al., 2017) and references therein.

antennas, given by $B\cos\theta/c$ where c is the speed of light. These signals are then passed to a device called a correlator, where their phase difference is measured. Essentially, what is measured in the correlator is the interference pattern produced by the two antennas which is conventionally named as the "visibility". The visibility is directly related to the source brightness, as it is the 2-dimensional Fourier transform of the sky in the source plane. The visibility plane is commonly referred as the uv plane, where u and v correspond to the projected baseline length of each antenna pairs, in terms of the number of wavelengths, in the east-west and north-south directions respectively. In principle, if we can sample every point in the uv plane, we can synthesise a single dish instrument with a large aperture of size (u_{max}, v_{max}). However, in reality there will be only a finite number of antennas in a interferometric array, so it is impossible to fully sample the uv plane. The most direct way to increase uv coverage of an array is to add more antennas, as each pair of antennas will provide one baseline and each baseline corresponds to two sample points in the uv plane. Therefore, for N antennas, there will be $N(N-1)$ sample points in the uv plane at a time. Beside adding more antennas, the Earth's rotation can help to increase the uv coverage as well. For observations with a long integration time, the antennas will be moving relative to astronomical objects, due to the rotation of the Earth, leading to larger uv coverage. In addition, some interferometers have a reconfigurable antenna layout such that they can change the baselines before or during observations, achieving larger uv coverage. For an interferometer with N antennas, its angular resolution is $\theta \propto \lambda/B_{max}$ where λ is the wavelength and B_{max} is the longest baseline, while its maximum angular scale is $\theta \propto \lambda/B_{min}$ where B_{min} is the shortest baseline. In addition, its field of view (FOV) is $FOV \propto \lambda/D$ where D is the diameter of the individual antenna.

In practice, an interferometer observes the uv plane by multiplying the sampling pattern by the visibility. The Fourier transform of this product is the convolution of the Fourier transform of the sampling pattern and the visibility. The Fourier transform of the visibility is the true sky brightness distribution, while the Fourier transform of the sampling pattern is the beam shape of the interferometer, also known as the "dirty beam". In other words, the raw image of an interferometer is the image of the true sky convolved with the dirty beam, this image is conventionally called the "dirty image" or "dirty map". Since the uv plane will never be completely sampled by the interferometer, regions that are not sampled in the uv plane will be dominated by the dirty beam in the image plane, resulting in

numerous sidelobe artefacts in the dirty image. In order to retrieve the true sky brightness distribution, we need to deconvolve it from the dirty beam. This effectively means extrapolating regions that are not sampled in the *uv* plane from sampled regions. While there are several algorithms which are developed for this, the most widely-used one is the CLEAN algorithm (Högbom, 1974) which is also used for the reduction of some data in this thesis. With the dirty map and the dirty beam as inputs, the CLEAN algorithm firstly detects the brightest source in the dirty image, then copies this source to a new blank image, called the CLEANed image, with a scaled dirty beam being subtracted from the source extracted from the dirty image. This process is then repeated to add more sources to the CLEANed image, until a certain number of iterations is reached or there is no further source that can be detected above a defined threshold. Lastly, the CLEANed image is convolved with a model of the "clean beam" which is typically an elliptical Gaussian fitted to the main lobe of the dirty beam, and the residual of the dirty image is added to the CLEANed map. As such, most sidelobe artefacts are removed to produce a reasonable restoration of the true sky image.

The effects of sidelobes can also be reduced by changing the weighting of the sampling function of the visibility. There are three commonly used weighting schemes, Natural weighting, Uniform weighting and Briggs weighting. The Natural weighting weights each sample in the *uv* plane by the inverse of their noise variance. It provides the maximum point source sensitivity but at the cost of lower angular resolution. Uniform weighting weights the sample points of the *uv* plane inversely proportional to their local density, such that the sum of the sample weights in a region is equal to a constant. In this way, longer baselines will have larger weights as they provide sample points at a larger *uv* scale, hence leading to higher angular resolution and lower sensitivity. Briggs weighting provides a compromise between the Natural and Uniform weighting. It is based on the Uniform weighitng but it will give lower weights to *uv* samples with low Natural weights. The Briggs weighing scheme has a "robust" parameter which can be adjusted to change the resulting angular resolution and point source sensitivity.

1.3.4 The very negative K-correction

The commission of the Submillimeter Common User Bolometric Array (SCUBA) in 1997 opened a new era of research in the high-z Universe by providing a statistically large sample of unlensed

high-redshift FIR/submm bright sources for the first time. There were, however, great challenges in studying these sources' physical properties at that time because of the large beamsize of SCUBA (around 15" at 850μm) and thus the difficulty in cross-identification in other wavelengths (Ivison et al. 1998). Nonetheless, these submm sources were already believed to exist in the $z \geqslant 2$ Universe (e.g., Clements et al. 1992), this is because they all have nearly constant observed flux densities with increasing redshift, which is called a very negative K-correction.

A K-correction is used to convert the observed-frame absolute magnitude/flux of a redshifted object to its rest-frame magnitude/flux. It depends only on the intrinsic shape of a galaxy's spectral energy distribution (SED). Because of the convention in the sign of magnitude, the K-correction is referred as "positive" if the observed flux density decreases with increasing redshift, and "negative" if it increases with redshift (Hogg et al. 2002; Casey et al. 2014).

As mentioned in previous sections, galaxies' FIR/submm emission primarily originates from dust thermal emission which can be well approximated by a modified blackbody function peaking at rest-frame wavelengths 100μm, and the long-wavelength part of the spectrum is the Rayleigh-Jeans regime (Blain et al. 2002; Bianchi 2013). Under this approximation, the observed flux density S_ν is expressed as:

$$S_\nu \propto \frac{\nu^\beta B_\nu}{D_L^2} \tag{1.3}$$

where B_ν is the Planck function at frequency ν, β is the dust emissivity index which is usually between 1 and 3, and D_L is the luminosity distance. Note that beyond 3 mm wavelength, dust's thermal emission will become insignificant and a combination of synchrotron and free-free emission will dominate (Casey et al. 2014).

In the Rayleigh-Jeans approximation, which is applicable to the long-wavelength end of the FIR, submm and millimetre, the Planck function can be simplified to $B_\nu \propto \nu^2$, then the observed flux density becomes

$$S_\nu \propto \frac{\nu^2 \nu^\beta}{D_L^2} \propto \frac{\nu^{2+\beta}}{D_L^2} \tag{1.4}$$

Recalling the relationship between observed frequency and rest-frame frequency: $\nu_{rest} = \nu_{obs}/(1+z)$,

we find

$$S_\nu \propto \frac{\nu_{obs}^{2+\beta}}{D_L^2} \propto \frac{(\nu_{rest}(1+z))^{2+\beta}}{D_L^2} \tag{1.5}$$

At $0.5 < z < 3$, D_L can be approximated as $D_L \propto (1+z)^2$ (Casey et al. 2014), therefore

$$S_\nu \propto \frac{(\nu_{rest}(1+z))^{2+\beta}}{D_L^2} \propto \frac{\nu_{rest}^{2+\beta}(1+z)^{2+\beta}}{(1+z)^4} \propto (1+z)^{\beta-2} \tag{1.6}$$

Studies have showed that $\beta = 1.5 - 2.0$ is a reasonable estimation for dusty galaxies (Casey 2012; Bianchi 2013; Casey et al. 2014). This then leads to a roughly constant $S_\nu(z)$. In other words, in the wide redshift range $1 < z < 8$ where the Rayleigh-Jeans approximation and the luminosity distance approximation apply, the observed flux density will be roughly constant (or even slightly increasing), which makes the K-correction for submm sources very negative.

Figure 1.5 describes the relationship between the observed flux density and redshift in a series of wavelength bands. It clearly shows the very negative K-correction between $850\mu m$ and 2 mm, while it becomes less negative in the $250 - 500\mu m$ FIR range, and eventually turns positive at shorter wavelengths.

1.3.5 Instrumentation

Over the last few decades, a large number of both ground-based and space-based IR/submm telescopes were built and launched, producing much data for high-z DSFG studies. However, it should be noted that observations in most mid-infrared (MIR)/FIR wavelengths are strongly limited for ground-based telescopes by the opacity of the Earth's atmosphere, as shown in Figure 1.3. Mainly caused by water vapour absorption, the atmosphere only provides a few narrow windows occurring at 350, 450, 770, 870 μm and two relatively broad windows at 1 mm and 2 mm. Meanwhile, although there is no such limitation for space-based telescopes, they are also constrained by our capability to send instruments to space as well as the high cost in building and maintaining space-based telescopes, resulting in generally worse resolution. Knowing these limitations in instrumentation is crucial to understanding high-z DSFG studies carried out in the last few decades.

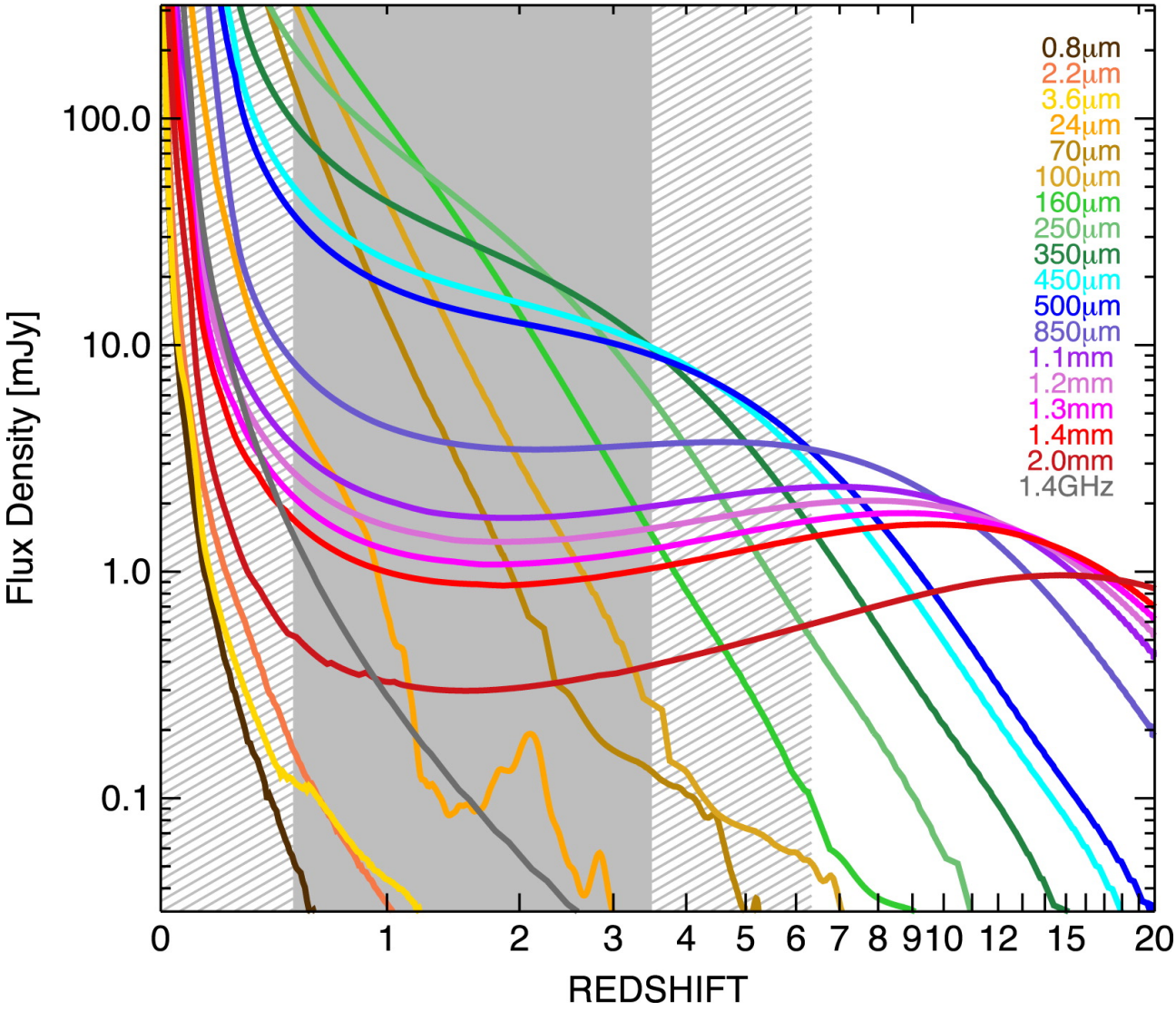


Figure 1.5: The observed flux densities vs. redshift for a typical IR bright galaxy with its IR luminosity being normalised to $10^{12.5} L_{\odot}$. This figure is adopted from Casey et al. (2014). It displays the nearly-constant flux densities in submm and mm wavelengths over a wide range of redshift. The flux densities are generated from a template DSFG SED of fixed luminosity (Pope et al. 2008). The grey area highlights the redshift range in which most of the DSFGs are discovered, while the shaded regions represent redshift ranges where such IR bright galaxies are rare.

In the rest of this section, we will review several key telescopes/instruments which are extensively used in IR/submm astronomy in this thesis. There are other facilities and instruments that are equally crucial in infrared and submm astronomy, but for briefness we will not discuss them here. In addition, a brief introduction to the radio telescope MeerKAT will be included, as it is closely relevant to our FIR-radio study.

Spitzer The Spitzer Space Telescope (Werner et al. 2004) was the third space telescope dedicated to infrared astronomy after IRAS and ISO. It was launched in 2003 and retired in early 2020. There were three scientific instruments on board, the Infrared Array Camera (IRAC; Fazio et al. 2004) which had four near-IR bands at 3.6, 4.5, 5.8 and 8.0 μm, the Multi-band Imaging Photometer for Spitzer (MIPS; Rieke et al. 2004) which had three channels at 24, 70 and 160 μm, and the InfraRed Spectrograph (IRS; Houck et al. 2004) which worked between 10 and 40 μm. In 2009, the coolant on Spitzer was exhausted. It then continued to operate in only the two shortest IRAC bands until its retirement.

MIPS is particularly important for early high-redshift galaxy studies in the MIR. Its beamsizes are $6''$, $18''$ and $41''$ at 24, 70 and 160 μm respectively, with the 24 μm channel being the most sensitive and extensively used. MIPS has covered a large area of the sky, allowing astronomers to build up statistically significant populations of high-z MIR luminous dusty galaxies for the first time.

The Sub-Millimeter Array The Sub-Millimeter Array (SMA; Ho et al. 2004), which officially began its operations in late 2003, is an interferometer for submm observations consisting of eight 6m-diameter antennas. It is located at Mauna Kea, Hawaii, with baselines ranging from 16 to 508 m leading to angular resolutions of between $0.1''$ and a few arcseconds, and working at wavelength between 735 μm to 1.545 mm. The SMA is one of the first instruments to produce submm and mm images with sub-arcsecond angular resolution. With this feature, it has been used to resolve submm galaxies (SMGs) (e.g. Chen et al. 2011; Riechers et al. 2013).

Herschel The *Herschel* Space Observatory (Pilbratt et al. 2010) was launched in May of 2009 and operated until April 2013, carrying three scientific instruments: the Photodetector Array Camera and Spectrometer (PACS, Poglitsch et al. 2010), the Spectral and Photometric Imaging REceiver (SPIRE,

Griffin et al. 2010), the Heterodyne Instrument for the Far Infrared (HIFI, De Graauw et al. 2010). PACS consisted of an integral field spectrometer and an imaging photometer, with the spectrometer operated between $50 - 210\mu m$, and the photometer simultaneously imaging in 70 or 100 μm and 160 μm with beamsizes of 5", 7" and 12" respectively. SPIRE was also made up of a spectrometer and an imaging photometer. The photometer operated simultaneously in three channels at 250, 350 and 500 μm with beamsizes of 18", 26" and 36" respectively. HIFI was a high resolution spectrometer working between $157 - 212\mu m$ and $240 - 625\mu m$.

Although PACS had smaller beamsizes and less confusion noise than SPIRE as it operated at lower wavelengths, PACS was much less efficient in mapping large areas of sky compared to SPIRE. For this reason, SPIRE data are more widely utilised for high-redshift DSFG studies. The major data of SPIRE for this thesis comes from the two largest surveys, the *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al. 2012), and the *Herschel*-ATLAS Survey (H-ATLAS; Eales et al. 2010).

SCUBA-2 The Submillimeter Common-User Bolometer Array 2 (SCUBA-2; Holland et al. 2013), the second generation bolometer array for the 15m diameter JCMT, is a 10,000 pixel bolometer camera operating simultaneously at 450 and 850 μm with beamsizes of 9.8" and 14.6" respectively. Since its commissioning in 2011, SCUBA-2 has dedicated plenty of its observing time to blank-field extragalactic mapping at both 450 and 850 μm, with a mapping speed 100-150 times faster than its predecessor SCUBA. Some of the large surveys using SCUBA-2 include the SCUBA-2 Cosmology Legacy Survey (S2CLS, Geach et al. 2017), and the SCUBA-2 Ultra Deep Imaging EAO Survey (STUDIES, Wang et al. 2017).

ALMA The Atacama Large Millimeter/submillimeter Array (ALMA; Wootten and Thompson 2009) is an interferometer array consisting of 66 antennas in the Atacama Desert of northern Chile. The telescope began scientific observations in late 2011, and became fully operational since March 2013. It operates between 0.32 and 3.6 mm, with baselines ranging from 150 m to 16 km. The main array of ALMA is composed of 50 antennas, each with 12 m diameter, and this main array is complemented by a compact array of four 12 m antennas and twelve 7 m diameter antennas. With a primary beamsize of 17", the maximum angular resolution can reach 0.015". ALMA has been extensively used to detect

atomic and molecular lines, image protostars and protoplanetary disks and resolve extended structures in both the local and high-z Universe (e.g. Wang et al. 2013, 2021).

MeerKAT The MeerKAT radio telescope (Jonas 2009), originally known as the Karoo Array Telescope located in the Karoo desert in South Africa, is a precursor to the upcoming Square Kilometre Array (SKA; Carilli and Rawlings 2004). Launched in 2018, it consists of 64 13.5m-diameter antennas with baselines ranging from 29m to 8km, and a working frequency between 580 MHz and 15 GHz. 48 of the antennas are concentrated in an inner region with a shortest baseline of 29 m and a longest baseline of 1km, while an outer region holds the rest of the antennas with a maximum base line of 8km. Being the most sensitive cm-wavelength interferometer array in the south hemisphere, its scientific scope covers a wide range of astronomy, including identification of FIR/submm bright DSFGs which are generally difficult to detect in the optical/NIR (see Section 1.4.6 and Chapter 2).

1.4 Current Understanding Of Dusty Star Forming Galaxies

As discussed in Section 1.2.2, the contribution of local (U)LIRGs is not sufficient to explain the observed CIB intensity and spectral shape, despite their extreme infrared luminosities and SFRs. In the late 1990s, a large number of submm bright galaxies ($> 1\ mJy$ at 850 μm) were discovered, with only a few of them having optical/NIR counterparts identified (Smail et al. 1997, 2002). Follow-up analysis including number counts and redshift distributions found that these SMGs/DSFGs were 100-1000 times more populous than local (U)LIRGs, and the majority of them probably lay at $z > 1$. Their high redshifts imply their extreme FIR luminosities (defined as integrated rest frame luminosity between 42.5 and 122.5 μm) $L_{FIR} > 10^{12} L_{\odot}$ or even $> 10^{13} L_{\odot}$ in some cases, implying SFRs of $> 100 - 1000\ M_{\odot} yr^{-1}$. Many studies have found that these DSFGs had a larger submm emission size of around 5 kpc than that of local ULIRGs, whose submm emission normally only extends hundreds of parsec (e.g., Engel et al. 2010; Hodge et al. 2016). This extended submm emission of these DSFGs also appears to be smooth, suggesting most star-forming activities are hosted in a smooth disk rather than in clumps (Gullberg et al. 2018). In addition, DSFGs at $z \approx 2$ are observed to have both a much higher gas mass and a higher gas to stellar mass ratio than local (U)LIRGs for a fixed FIR luminosity

(e.g., Tacconi et al. 2008, 2010). For these reasons, DSFGs are thought to be scaled-up versions of local (U)LIRGs, and there is a high possibility that these DSFGs dominate the cosmic SFR density at $z \geqslant 2$, with nearly 60% of the CIB intensity at 850 μm provided by them (Smail et al. 2002).

In the following sections, we will review our understanding of the high-redshift DSFGs, and outline the current problems and challenges that we are facing with them.

1.4.1 Number counts

One of the most fundamental characteristics of any galaxy population is its number count as a function of flux density at a given wavelength. Because of the limitations of the atmospheric transmission, as discussed in Section 1.3.1, there are only a small number of spectral windows available for observations in the FIR/submm. Hence only a limited number of wavelengths have accurate DSFG number count measurements. Early data from ISO observations provides number counts at $70 - 170\mu m$, which were later updated by *Spitzer* MIPS and *Herschel* PACS observations with better constraints (e.g., Kawara et al. 2004; Berta et al. 2011; Magnelli et al. 2013). Accurate number counts at $250, 350$ and $500\mu m$ are measured by *Herschel* SPIRE and BLAST observations (e.g., Patanchon et al. 2009; Clements et al. 2010b; Glenn et al. 2010; Béthermin et al. 2010, 2012; Valiante et al. 2016), with the addition of $450\mu m$ number counts from SCUBA-2 surveys (e.g., Geach et al. 2013; Zavala et al. 2016; Wang et al. 2017). The $850 - 870\mu m$ number counts are probed by large surveys using SCUBA and SCUBA-2, together with additional LABOCA and ALMA data (e.g., Knudsen et al. 2008; Chen et al. 2013; Karim et al. 2013; Simpson et al. 2015; Geach et al. 2017; Michałowski et al. 2017). More recently, number counts of DSFGs at millimetre wavelengths are also investigated with instruments such as ALMA, AzTEC and the SPT (e.g., Scott et al. 2010; Hatsukade et al. 2013; Aravena et al. 2016; Umehata et al. 2017; Zavala et al. 2018a; Everett et al. 2020; Gómez-Guijarro et al. 2022). Figure 1.6 shows examples of differential number counts at $70\mu m - 1.1mm$.

It is found that gravitationally lensed DSFGs contribute a great fraction of the bright-end number counts at wavelengths above 500 μm. There are two potential reasons. First, FIR/submm number counts are steep, which means that the number of galaxies decreases sharply as the flux density increases.

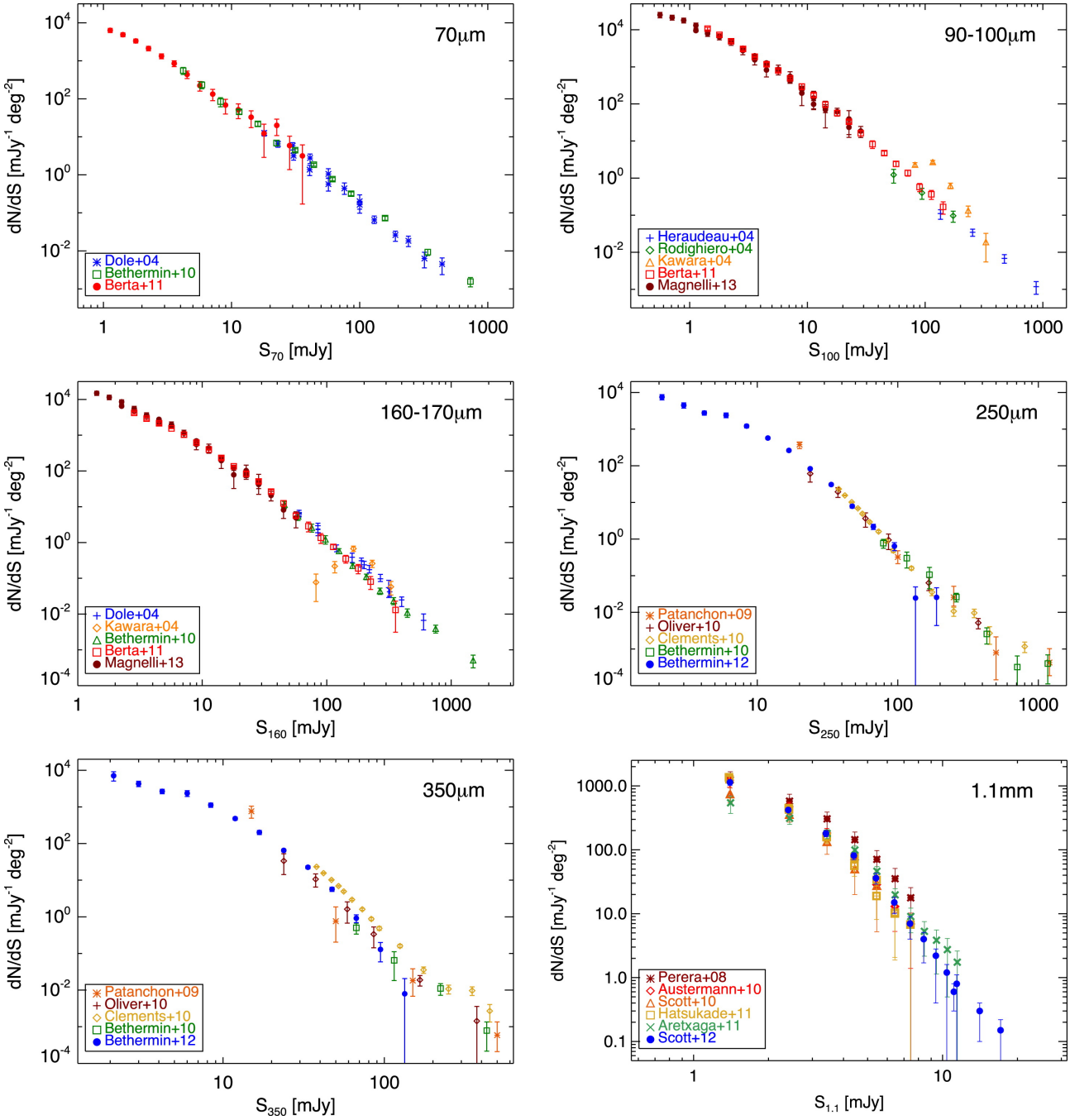


Figure 1.6: Differential number counts at 70, 100, 170, 250, 350μm and 1.1mm from various studies. This figure is adopted from Casey et al. (2014).

Gravitational lensing of sources at the steep faint end moves them to the bright end and flattens the bright-end number counts, thus a large fraction of bright FIR/submm galaxies are lensed DSFGs (e.g., Blain 1996). Second, due to the negative K-correction as discussed in Section 1.3.4, higher redshift galaxies can be observed at longer wavelengths (e.g., Weiß et al. 2013). At wavelengths shorter than 500 μm, the bright-end number counts are dominated by low redshift star-forming galaxies and the contribution from lensed DSFGs is not significant (e.g., Negrello et al. 2007; Wardlow et al. 2012). Theoretical predictions and cosmological simulations suggest that, after accounting for local galaxies and AGNs, all sources at $S_{500\ \mu\mathrm{m}} > 80\ mJy$ should be gravitationally lensed DSFGs (Negrello et al., 2007; Béthermin et al., 2017). While there is no observational evidence of a 100% lensed fraction, some studies find that over 90% of bright sources are lensed DSFGs at both 500 μm(e.g., Negrello et al. 2010; Wardlow et al. 2012; Negrello et al. 2016) and 1.4 *mm* (e.g., Vieira et al. 2010, 2013; Mocanu et al. 2013). Overall, gravitationally lensed DSFGs are of great importance to high redshift DSFG studies, as they account for the majority of bright-end FIR/submm number counts at wavelengths above 500 μm, and enable studies of high redshift DSFGs most of which are too faint to be observed without lensing.

1.4.2 Multiplicity of DSFG

One of the main constraints in early DSFG studies comes from the multiplicity of DSFGs, in other words, the number of sources whose emission is blended into a form of single-source emission. For confusion-limited submm surveys, sources' flux densities have little dependence on redshift because of the very negative K-correction, thus it is highly possible for two neighbouring, relatively faint submm sources to appear as one line-of-sight source, even if they are not physically associated (e.g., Negrello et al. 2005; Hodge et al. 2013; García-Vergara et al. 2020). This indeed poses great challenges for studies involving very bright DSFGs, which may eventually resolve into multiple faint counterparts when examined at higher spatial resolution (e.g., Oteo et al. 2017; Greenslade et al. 2020; Cairns et al. 2023). This is not an issue for optical sources, as flux densities of optical sources will evolve substantially with redshift, plus optical observations typically have much higher resolution. Also, as an unknown fraction of bright DSFGs may break into multiple fainter sources, it is thus difficult for

theoretical models to determine the genuine number count of DSFGs. A large number of studies have been carried out to understand the multiplicity of DSFGs (e.g., Tacconi et al. 2006, 2010; Simpson et al. 2015; Stach et al. 2018, 2019; Greenslade et al. 2020; Cairns et al. 2023)

Most of the DSFG multiplicity studies rely on high-resolution interferometric follow-up observations of DSFGs detected via single-dish FIR/submm instruments. The construction of the SMA significantly help such studies. For, example Wang et al. (2010) utilised the SMA to investigate the morphologies of two SMGs. It turned out that both consisted of multiple counterparts. Following this study, Barger et al. (2012) found roughly 25% of their 16 SMG sample resolved into multiple sources as well. Greenslade et al. (2020) and Cairns et al. (2023) present SMA follow-up on *Herschel*-selected DFSGs that potentially lie at $z \geq 4$, finding that $\sim 35\%$ of their bright ($S_{500\mu m} > 60mJy$) sample and $\sim 60\%$ of their faint ($S500 < 60mJy$) sample are multiples. This already suggests that a relatively large fraction of detected SMGs are blends. These studies were pushed further by the commissioning of ALMA. Hodge et al. (2013) and Karim et al. (2013) estimated that 30-50% of a sample of ≥ 100 high-redshift SMGs resolved into multiple counterparts, with all $S_{870} > 10mJy$ sources being blends. Simpson et al. (2015) obtained 52 SMGs with ALMA from a sample of 30 single dish SMGs, and suggested $61^{+19}_{-15}\%$ single dish SMGs will resolve into multiple counterparts at higher spatial resolution. Stach et al. (2018) re-observed 606 single dish SMGs with ALMA, finding that $44^{+16}_{-14}\%$ of $S_{850} > 9mJy$ sources break into multiple sources. In summary, it is clear that, at least for bright ($S_{850} > 9mJy$) SMGs, a significant fraction are multiple.

The question then follows: are the multiple counterparts of a SMG physically associated? Simpson et al. (2015) find that components of a single dish source are generally associated by comparison to blank field counts. Photometric redshift analysis in Stach et al. (2018) agrees with this conclusion. CO observations of 12 SMGs (Tacconi et al. 2008; Engel et al. 2010) suggests that all the galaxies in the sample are either ongoing mergers or compact dense galaxies which are considered as late stage mergers, meaning that SMGs are likely to be high-redshift analogues to local (U)LIRGs. To the contrary, models from Hayward et al. (2013) and Cowley et al. (2015) indicate multiple sources resolved from single dish SMGs are generally not associated, hence are not merger companions. Moreover, Wardlow et al. (2018), using ALMA, examined the redshifts of 6 single dish SMGs which

resolve into multiple counterparts via CO emission line detections. While 61 ± 18% of their samples probably have no physical associations, they found that 3 of their 6 targets have faint CO emitting companions at close redshifts to the source. This may imply that these SMGs are still part of mergers, though their companions are not submm luminous.

In conclusion, for now it is only clear that most single dish DSFGs (or more broadly speaking SMGs) are blends of multiple sources. However, whether these blended sources are physically associated or not is still under debate, as evidence exists for both sides. The majority of detected DSFGs display signs of ongoing/late stage merger, though the chances are that the merger companions are submm faint in some cases (Wardlow et al. 2018). Overall, the hypothesis that DSFGs are merger induced starbursts at high redshifts, like (U)LIRGs in the local Universe, is so far the most plausible theory to explain current observations of DSFGs.

1.4.3 The redshift distribution of DSFG

Although DSFGs are thought to generally lie at high redshifts, accurate determination of their redshift is significantly restricted by the fact that DSFGs have severe extinction in their optical and UV spectrum where most commonly-used emission-line redshift indicators are located. The large beamsizes in FIR/submm observations also complicate multi-wavelength counterpart identifications of DSFGs which are usually used for SED fitting and hence redshift estimation. Since their discovery, numerous attempts have been made to constrain the redshift distribution of high-z DSFGs. The first and so far one of the most well-studied samples of high-redshift DSFGs are the 850μm selected SMGs. Chapman et al. (2005) identified 1.4 GHz counterparts to 73 out of 150 SCUBA selected SMGs and thus obtained their accurate coordinates. They then spectroscopically confirmed redshifts of these SMGs, resulting in a median redshift of $z = 2.2$ and interquartile range of $z = 1.7 - 2.8$. Later on, Wardlow et al. (2011) derived photometric redshifts from 17-band optical to mid-infrared photometry of 78 robust radio, Spitzer 24 μm and IRAC counterparts to 72 of the 126 SMGs selected at 870 μm by LABOCA (Siringo et al. 2009). They found a median redshift of $z = 2.2 \pm 0.1$, and an interquartile range between $z = 1.8$ and $z = 2.7$, which is in perfect agreement with the results in Chapman et al. (2005). Moreover, Smith et al. (2017a) applied statistical techniques to a sample of 761 850μm

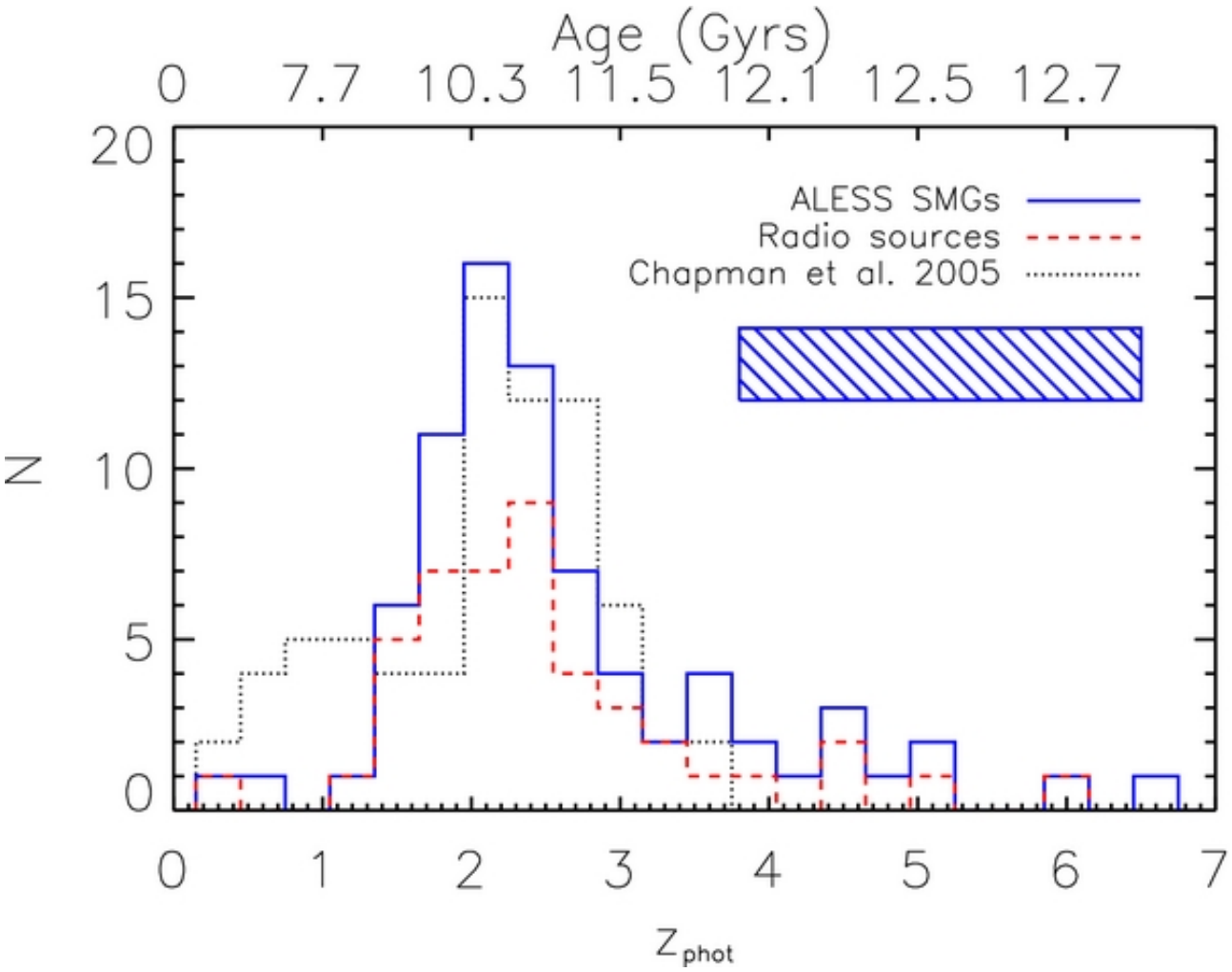


Figure 1.7: The photometric redshift distribution of 850 μmselected SMGs. This figure is adopted from Simpson et al. (2014). The blue solid lines are the Simpson et al. (2014) ALESS SMG samples, while the red dashed line represents their sources with radio flux densities $S_{1.4GHz} > 40\mu Jy$, The black dotted line indicates the Chapman et al. (2005) sample. The hatched box represents missing histogram of the Simpson et al. (2014) sources due to 19 SMGs with insufficient photometry to derive photometric redshifts. They pointed out that these 19 sources should have redshifts > 3.

SCUBA-2 SMGs, detecting excess galaxies around the positions of those SMGs in optical and NIR catalogues. Assuming the optical/NIR redshifts were representative of the 870μm-selected galaxies, a redshift median of $z = 2.05 \pm 0.03$ was derived with 68% of them between $1.07 < z < 3.06$. Results from these studies show that most SMGs are distributed at around $1 < z < 3$, which is often referred as "cosmic noon" representing the most intense star-forming period (see Madau and Dickinson 2014 for a detailed review), whilst there are only a few observations of higher-redshift DSFGs. Figure 1.7 describes the photometric redshift distribution of several well-studied 850 μmselected samples.

Apart from the 850μm selected galaxies, large surveys conducted by Herschel-SPIRE (especially Her-

MES and H-ATLAS) have provided sufficient FIR samples for follow-up analysis including redshift determination. However, redshift estimations for SPIRE selected sources suffer heavily from uncertainties because of SPIRE's large beamsizes and hence low spatial resolution. Nonetheless, several attempts have been made. Using a sample of 1686 350 μm detected sources in the early science phase of H-ATLAS (Eales et al. 2010), Amblard et al. (2010) fitted PACS and SPIRE flux densities to isothermal modified blackbody SED models with varying parameters (e.g. dust temperature, redshift and dust spectral indices). A median redshift of $z = 2.2 \pm 0.6$ was obtained, with no high redshift ($z > 3.5$) source seen in the 350 μm selected sources. Casey et al. (2012a) identified 24 μm and 1.4 GHz counterparts of 767 SPIRE selected source at 250, 350 and 500 μm to locate their accurate positions, followed with spectroscopic observations, finding a peak redshift of $z = 0.85$ with 731 sources at $z < 2$ and a tail of sources out to $z \approx 5$. They also intentionally select a sub-sample of 350 μm detected sources to compare to results in Amblard et al. (2010), but there was significant disagreement between the redshift distributions of the samples in the two studies.

While redshifts of generic SPIRE selected galaxies are not well constrained, some subsets have drawn more attention. Among these better-studied sub-samples of SPIRE selected sources, a particular one is called the "500 μm risers", which is one of the main research targets of this thesis. We will discuss them in Chapter 4 with more details.

1.4.4 The luminosity functions of DSFGs

Knowing that most DSFGs are ULIRGs with $L_{FIR} > 10^{12} L_{\odot}$ with redshifts at $z \approx 2$, we can then probe the evolution of the integrated IR luminosity function of DSFGs with respect to redshift, which represents the obscured star formation rate density. A luminosity function is written as $\phi(L)dL$, describing the volume density of galaxies with a given luminosity L ($\phi(L)$) between L and $L + dL$. It should be noticed that knowledge of luminosity functions beyond $z = 4$ is very limited by sample selection effects and biases, as well as incompleteness because of the lack of data on high-z DSFGs.

Le Floc'h et al. (2005) firstly examined the integrated infrared luminosity function out to $z \approx 1$ by extrapolating the Spitzer 24 μm flux densities to an integrated IR luminosity. They find that the

integrated IR luminosity functions evolves $\propto (1+z)^{3.9\pm0.4}$, while the UV luminosity function (which represents the unobscured star formation rate density) over the same redshift range goes as $\propto (1+z)^{2.5}$. More work on Spitzer data was carried out by Caputi et al. (2007) and Magnelli et al. (2011) which all agreed with Le Floc'h et al. (2005) reasonably well. These results, together with those in other similar studies, suggest that dust-obscured star formation probably dominates the cosmic star formation rate density (CSFRD) at $z \approx 2$ (e.g., Chapman et al. 2005; Wardlow et al. 2011; Barger et al. 2012; Lim et al. 2020; Rodríguez-Puebla et al. 2020; Traina et al. 2024).

Meanwhile, similar studies with Herschel data are more advantageous than those using Spitzer data, as PACS and SPIRE directly observed at the peak wavelengths of DSFG emission. Such Herschel-based studies (e.g., Casey et al. 2012a,b; Magnelli et al. 2013; Gruppioni et al. 2013), while in excellent agreement with previous results, have now pushed the DSFG integrated luminosity function all the way to $z \approx 4$. The IR luminosity function appears to evolve strongly out to $z = 1$, then continues to moderately increase to $z \approx 3$. With SMGs dominating the SFR density at $1 < z < 3$, this implies an increase in both the number density and the luminosity of DSFGs in this redshift range. Koprowski et al. (2017) further extended the luminosity function to $z \approx 5$ using SCUBA-2 and ALMA, but they found that the high-z evolution of the IR luminosity function consists of a positive luminosity evolution coupled with negative density evolution with increasing redshift, leading to a conclusion that the cosmic star formation density is dominated by optical/UV star formation instead of IR luminous star formation at $z > 4$. Conversely, several studies (e.g., Rowan-Robinson et al. 2016; Novak et al. 2017; Yan et al. 2020) show that the integrated IR luminosity function remains roughly constant out to $z \approx 6$. A plot summarising how the IR luminosity function evolves with redshift up to $z \sim 3.6$ is adopted from Casey et al. (2014) in Figure 1.8.

It still remains unclear how the IR luminosity function evolves at high redshifts because of various difficulties in constraining the physical properties of DSFGs, and more importantly the lack of $z > 4$ samples of DSFGs. Casey et al. (2018) compared two models of IR luminosity function evolution, one where the DSFG contribution to the total cosmic star formation at $z > 4$ is insignificant, and one where the DSFG contribution is dominant larger than 90%. Their conclusion is that current data are insufficient to rule out either of these two models. Therefore, the nature of evolution of the

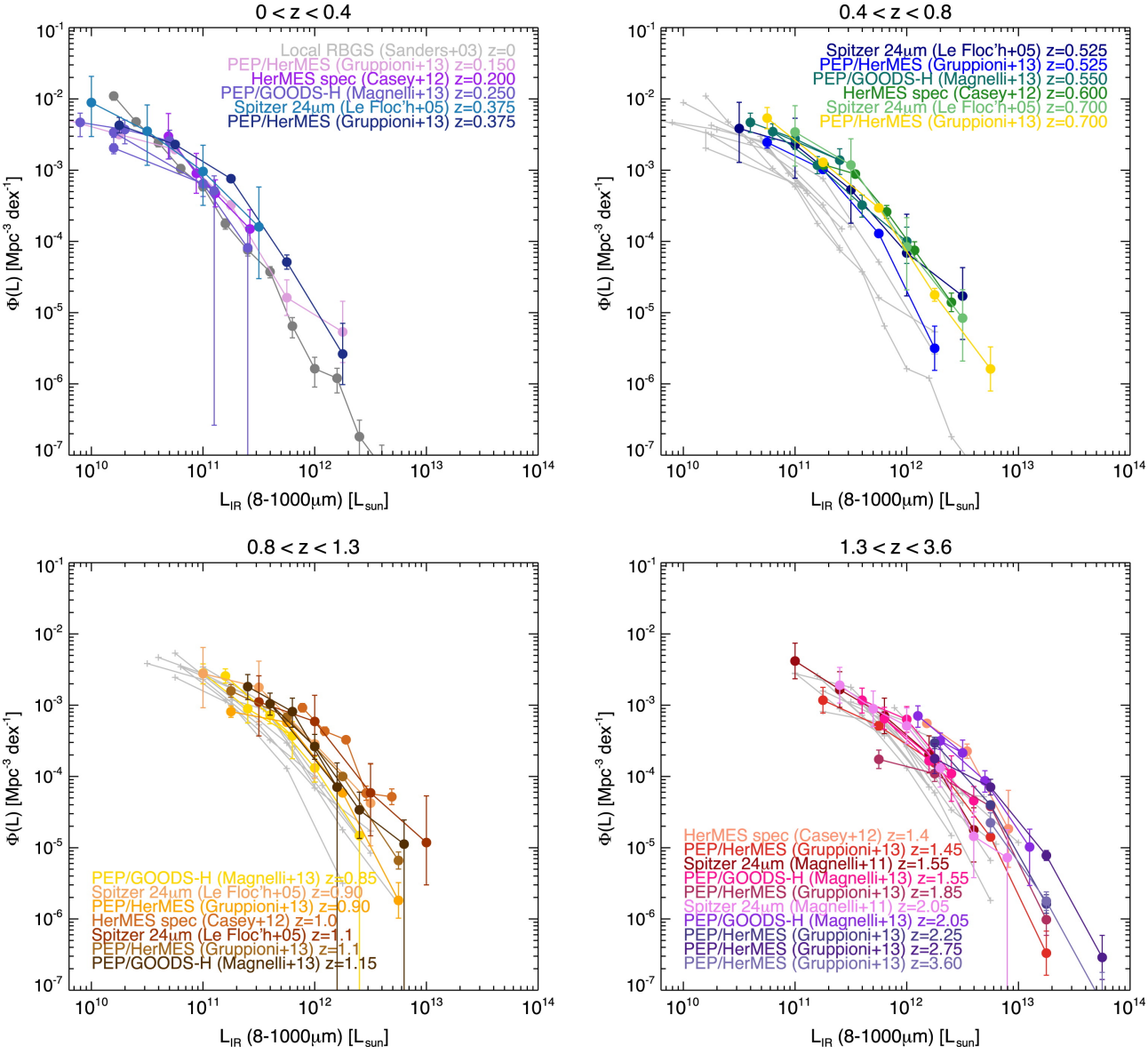


Figure 1.8: Measured integrated IR luminosity functions in four redshift bins. The corresponding redshift bin is shown in the title of each panel. This figure is adopted from Casey et al. (2014). The results of various studies are shown in different colours, including Sanders et al. (2003); Le Floc'h et al. (2005); Magnelli et al. (2011); Casey et al. (2012a); Magnelli et al. (2013) and Gruppioni et al. (2013)

IR luminosity function at $z > 4$ is yet to be determined. Nonetheless, some recent studies utilising ALMA have began to shed light on the high-redshift regime of the IR luminosity function. Both Gruppioni et al. (2020) and Zavala et al. (2021) carry out deep ALMA observations on DSFGs at millimetre wavelengths. They find similar results of a flattening of the IR luminosity function at the faint tend. However, while Gruppioni et al. (2020) find little evolution in the dust-obscured star formation density between $z = 2$ and $z = 6$, Zavala et al. (2021) suggest a reduced CSFRD contribution from dust-obscured star formation at $z > 4$, which is estimated to contribute $\sim 20 - 25\%$ of the total CSFRD at $z > 6$. Even though there are some discrepancies between these ALMA studies of the IR luminosity function, their results show that the CSFRD is very likely to be dominated by DSFGs at $1 < z < 4$.

1.4.5 DSFG protoclusters

Protoclusters are the progenitors of today's massive galaxy clusters. If we define clusters as virialised objects with total masses $\geq 10^{14} M_{\odot}$, the definition of a protocluster then naturally becomes a structure which will eventually collapse into a $M \geq 10^{14} M_{\odot}$ galaxy cluster at $z = 0$ (Overzier, 2016). Such massive structures can not only place strong constraints on our current cosmology models, but can also provide key information on galaxy formation and evolution theories (e.g., Harrison and Coles 2012; Casey 2016). In the case of $z = 0$ galaxy clusters with typical mass of $M \geq 10^{14} M_{\odot}$ and virial temperature of intracluster medium (ICM) of $\sim 10^{7} - 10^{8} K$, their emission is dominated by X-ray emission from the hot ICM. Such hot ICM contains ionised hydrogen, helium and other metals such as iron, the X-ray emission of the hot ICM primarily comes from free-free bremsstrahlung radiation of metals (Sparke and Gallagher, 2000). Moreover, the cores of local galaxy clusters are found to contain red, quiescent elliptical galaxies with little star formation activity (e.g. Dressler 1980; Bamford et al. 2009; Kravtsov and Borgani 2012). Therefore, galaxy clusters at $z \leq 1.5$ can be efficiently discovered by searching for X-rays, red sequences of quiescent ellipticals, and also the Sunyaev-Zeldovich (SZ) effect which is the spectral distortion of the CMB caused by the inverse Compton scattering of CMB photons by relativistic electrons in the ICM. These methods have already been proven to be very successful in identifying low-redshift clusters (e.g., Kravtsov and Borgani

2012; Overzier 2016), but they become much less effective in searching earlier stages of clusters at higher redshifts, i.e. protoclusters. This is partly due to they sit at higher redshifts and hence longer distances and lower flux densities. But more importantly, protoclusters are not virialised to form hot X-ray emitting ICM, thus they can not be found by looking for X-rays or the SZ effect. Also, galaxies in the cores of protoclusters are not evolved into red quiescent galaxies and are still actively forming stars, so protoclusters cannot be discovered by searching for red sequences of quiescent galaxies. This makes the search for protoclusters very challenging, and a wide variety of galaxy populations are used as signposts of protoclusters. These include $H\alpha$ emitters (HAEs, e.g. Shimakawa et al. 2018a,b; Zheng et al. 2021), Lyman α emitters (LAEs, e.g. Chiang et al. 2015; Harikane et al. 2019), Lyman Break galaxies (LBGs, e.g. Toshikawa et al. 2016, 2018), Balmer Break galaxies (BBGs, e.g. Shi et al. 2019, 2020), high-redshift radio galaxies (HzRGs, e.g. Overzier et al. 2006; Cooke et al. 2014), radio-detected AGN (e.g., Hatch et al. 2014; Onoue et al. 2018; Chapman et al. 2024). and SMGs/DSFGs (e.g., Capak et al. 2011; Oteo et al. 2018; Miller et al. 2018)

Among all the aforementioned galaxy populations, DSFGs are one of the most promising tracers of high-redshift protoclusters. In the local Universe, galaxy clusters are found to consist of massive quiescent elliptical galaxies with very low SFR, while active star-forming galaxies are often found in less dense environments (e.g., Dressler 1980; Balogh et al. 1998; Goto et al. 2003). This implies that member galaxies of the progenitors of the present day galaxy clusters, i.e. protoclusters, must have rapid star-formation activity at higher redshifts. As discussed in Sections 1.2.2 and 1.4.2, currently the most accepted theory for DSFG evolution leads to local massive quiescent elliptical galaxies most of which are located in the cores of galaxy clusters. A typical DSFG with $L_{FIR} = 10^{12.5} L_{\odot}$ can develop a stellar mass comparable to that of the Milky Way ($\approx 6 \times 10^{10} M_{\odot}$) in a mere 100 Myr with its extreme SFR, which is 100 times faster than the Milky way (Licquia et al. 2015). High-redshift DSFGs are also found to have large dynamical masses and gas masses, both at the scale around $10^{10-11} M_{\odot}$ (e.g. Tacconi et al. 2008, 2010; Bothwell et al. 2013). Spatial clustering of $1.5 < z < 3$ ULIRGs also suggest that they are likely star-forming sources which will become sources that reside in $z = 0$ clusters (Farrah et al., 2006). With their extreme SFR and gas masses, it is thus reasonable to assume them to be the progenitors of the most massive local ellipticals. In particular, Toft et al. (2014) study 18 compact quiescent SMGs with spectroscopic redshifts between $3 < z < 6$. By comparing their

physical properties such as SFR and stellar masses to compact quiescent galaxies at $z = 2$, they show that these high redshift SMGs are consistent with being the progenitors of $z = 2$ quiescent galaxies. The fact that these massive DSFGs are highly likely to be major mergers, as also discussed in Section 1.4.2, suggests they may lie in large overdensities, and evidence is found to support this speculation (e.g., Daddi et al. 2009b,a; Clements et al. 2016; Wilkinson et al. 2017, Oteo et al. 2018; Dudzevičiūtė et al. 2020).

Given the above evidence, it is therefore reasonable to think that DSFGs may be a good tracer of the formation of massive elliptical galaxies and hence galaxy protoclusters. Evidence has been found showing that high-redshift DSFGs do indeed trace dense protocluster/cluster regions. The local SFR-density relation, which finds that denser environments have significantly less star formation activity, is found to be reversed at higher redshifts (Elbaz et al. 2007; Bamford et al. 2009; Lemaux et al. 2022). Capak et al. (2011) discovered a massive protocluster at $z \approx 5.3$, consisting of DSFGs. Using *Herschel* SPIRE observations, Clements et al. (2014) found that four *Planck* sources are in fact galaxy clusters with overdensities of DSFGs at $0.8 < z < 2.3$. Clements et al. (2016) found an overdensity of SMGs around a *Planck*-detected $z = 3.26$ gravitationally lensed galaxy, in both *Herschel* SPIRE, LABOCA and SCUBA-2 data. Oteo et al. (2018) spectroscopically confirmed a $z = 4.002$ protocluster, whose core has an extremely red *Herschel* colour, with at least 10 DSFGs lying in this protocluster. Long et al. (2020) and Ivison et al. (2020) carried out follow-up observations on this protocluster, finding that it may have a larger spatial scale with more DSFG members. A similar $z = 4.3$ protocluster core is reported by Miller et al. (2018), comprised of at least 14 starbursting DSFGs. Follow-up observations reveal many more DSFG members with extreme SFRs in this protocluster (Hill et al., 2020; Rotermund et al., 2021; Hill et al., 2022). Combining *Herschel* and *Planck* data, Greenslade et al. (2018) identify 27 candidate DSFG protoclusters, SCUBA-2 follow-up observations on 13 of these candidates by Cheng et al. (2019) find that 9 of them exhibit overdensities with photometric redshifts at $z > 2$. More recently, the results of semi-analytic modelling by Araya-Araya et al. (2024) find a higher fraction of SMGs in the cores of protoclusters than in both the outskirts of protoclusters and the field, with the SMG fraction in the latter two regions being similar. All the above studies suggest that DSFGs can trace high redshift overdensities well. However, this is also surprising, as the time-scales on which these galaxies are bright in the submm are predicted to be very short (around

100 Myr), therefore detecting multiple physically associated submm bright sources will either imply an unknown large-scale starburst triggering mechanism (Hung et al., 2016), or that the duration of the star-burst phase is longer than expected (Falgarone et al., 2017). If DSFGs are indeed excellent tracers of massive overdensities in the distant Universe, it provides an effective method to discover protoclusters. The extreme brightness and rarity of high-z DSFGs, which make them relatively easy to be identified in large surveys, will be perfect to efficiently flag protoclusters.

However, there are other studies showing that DSFGs do not trace the underlying dark matter distribution, hence the most massive overdensities. Chapman et al. (2009) carry out observations of a SMG cluster at $z \approx 1.99$, finding that SMGs could probe a wider range of environments, instead of just tracing the progenitors of massive clusters. Miller et al. (2015) use a mock SMG catalogue derived from a cosmological simulation to assess how well SMGs can trace the underlying dark matter structure. They find that SMGs are poor tracers of dark matter overdensities because of the quenching of star-formation activity in SMGs at $z \leq 2.5$, and only $\sim 25\%$ of the most massive overdensities contain ≥ 1 SMG. Some apparent overdensities of SMGs are in fact dominated by Poisson noise because of the rarity of SMGs. A more recent study by García-Vergara et al. (2020) use ALMA to observe single-dish-detected SMGs, finding that the multiplicity of single-dish SMGs will artificially boost the overdensities derived from them. As there are a significant fraction of single-dish SMGs that are comprised of multiple fainter sources at random redshifts, some overdensities of SMGs at a given redshift may in fact be a larger number of fainter SMGs at different redshifts. They estimate that at $1 < z < 3$ SMGs with $S_{870\mu m} \geq 1.2mJy$ would not trace massive overdensities, and only the brightest ones with $S_{870\mu m} \geq 5 - 6mJy$ might be associated with massive structures. Furthermore, recent works have found protoclusters which are mainly comprised of quiescent galaxies at $2.77 < z < 3.37$ (McConachie et al., 2022; Ito et al., 2023), against the assumption that most members of protoclusters should be active star-forming galaxies.

Although no decisive evidence has been found yet to reach a conclusion on this topic, there are some further observations which may provide further insights into the association between DSFGs and protoclusters. While overdensities of DSFGs are found in the cores of some protoclusters (Capak et al., 2011; Miller et al., 2018; Oteo et al., 2018), DSFGs are also found to primarily locate outside

the cores of protoclusters. Observations of the $z = 2.16$ Spiderweb protocluster (Dannerbauer et al., 2014; Pérez-Martínez et al., 2023) show that at least eight SMGs are distributed around the core of this protocluster which consists of HAEs. Smail et al. (2014) study a $z = 1.62$ cluster with SCUBA-2 observations, finding 31 starbursting SMGs surrounding the core of this cluster which hosts a population of already passive red galaxies. Similarly, Zhang et al. (2022) find overdensities of DSFGs in the outskirts of two HAE traced protoclusters at $z \approx 2.2$. These discoveries imply a consistent hypothesis of the association between DSFG and galaxy cluster formation. At $z > 2.5$, some massive protoclusters have intense star-formation in their cores which will be traced by DSFGs. At redshift around $1.5 < z < 2.5$, the star-forming activities in the cores of some protoclusters have quenched, leaving a population of quiescent galaxies in their cores. Meanwhile in the outskirts of these protoclusters, there is still on-going star-formation activity, or star-formation activity is triggered in these regions and galaxies become FIR/submm bright. These SMGs will continue to form stars and eventually collapse to core and become part of the present day clusters. Even though this theoretical picture is consistent with most DSFG observations, it raises further questions. For example, what is the mechanism behind the quenching/triggering of star-formation in protoclusters at $1.5 < z < 2.5$? What is the spatial scale of such quenching/triggering of star-formation, and hence the scale of protoclusters? Even if DSFGs do trace protoclusters, how can we select a broad and unbiased sample of protoclusters such that we will not be biased towards a particular stage of cluster evolution? Without sufficient observational constraints, such questions can not be easily answered.

Overall, the environment of high-redshift DSFG and their relation to high-redshift protoclusters are still unclear. We can only be certain that DSFGs are occasionally associated with massive galaxy clusters in the distant Universe. Without further evidence, whether they always trace the most high-redshift massive overdensities is still unclear.

1.4.6 Radio identification of DSFGs

Radio observations have been widely used to study DSFGs due to several unique advantages. First, radio emission will not be absorbed or scattered by dust, allowing direct observation of most dust-obscured objects. Second, as most radio observations are obtained with interferometers, they can

achieve much higher angular resolutions and sensitivities compared to typical FIR/submm observations, hence allowing precise cross-identification of DSFG in other wavelength with accurate positions. Lastly, there is a well-known empirical correlation between the FIR emission and radio emission of a galaxy, called the FIR-radio correlation (FIRC), which can be used to identify radio counterparts of DSFGs and rule out radio-loud AGN. This surprisingly tight correlation between galaxies' FIR emission and their radio emission has been known since the late 1970s, first proposed by Van der Kruit (1971) then systematically measured by Helou et al. (1985). The FIR radiation is believed to come from thermal re-radiation of dust-enshrouded regions while the radio radiation is thought to be synchrotron emission from cosmic ray electrons accelerated in supernova remnants (e.g., Condon 1992). This naturally explains the observed correlation as the massive stars that eventually end as supernovae are those stars which substantially heat up their surrounding gas and dust (Harwit and Pacini 1975).

Numerous studies have been carried out on the FIRC (e.g., Helou et al. 1985; De Jong et al. 1985; Yun and Carilli 2002; Jarvis et al. 2010; Ivison et al. 2010a,b; Molnár et al. 2021; Delvecchio et al. 2021; Giulietti et al. 2022). The FIRC is often expressed by the parameter q_{IR} defined as (e.g., Helou et al. 1985; Ivison et al. 2010a):

$$q_{IR} = \log\left(\frac{S_{IR}}{3.75 \times 10^{12}\ Wm^{-2}}\right) - \log\left(\frac{S_{1.4GHz}}{Wm^{-2}Hz^{-1}}\right) \tag{1.7}$$

where S_{IR} is the integrated IR flux density between rest frame $8 - 1000\mu m$, $S_{1.4GHz}$ is the rest frame 1.4 GHz flux density. Radio flux density which is obtained in other radio frequencies can be converted to observed-frame 1.4 GHz flux density by $S_{1.4GHz} = S_\nu(1.4GHz/\nu \times (1+z))^{\alpha-1}$. Whether or not the FIRC evolves with redshift is still an open question. Yun et al. (2001) measured a value of $q_{IR} = 2.34 \pm 0.72$ in local starburst galaxies, while Magnelli et al. (2010) got a slightly lower $q_{IR} = 2.17 \pm 0.19$ with a sample of SMGs and optically faint radio galaxies. Given the high-redshift characteristics of DSFGs, this brings the interesting possibility that the FIRC may evolve with redshift. Ivison et al. (2010a) and Ivison et al. (2010b) used large samples of Balloon-borne Large-Aperture Submillimeter Telescope (BLAST; Devlin et al. 2009) and Herschel selected star-forming galaxies, and obtained a shallow redshift evolution of $q_{IR} \propto (1+z)^{-0.26\pm0.07}$ and $q_{IR} \propto (1+z)^{-0.15\pm0.03}$ respectively.

Additionally, a slightly steeper redshift evolution of $q_{IR} \propto (1+z)^{-0.30}$ out to $z \approx 2$ was evaluated by Casey et al. (2012a), using a sample of Herschel-selected DSFGs with spectroscopic redshifts. However, as Ivison et al. (2010b) has pointed out, the studies mentioned above all used IR-selected samples instead of a combination of radio- and IR-selected samples, which will result in intrinsic selection biases. Some researchers have argued that there will be no redshift evolution for the FIRC if all selection biases are carefully taken account (e.g. Sargent et al. 2010; Jarvis et al. 2010).

With all the advantages described above, it is relatively easy to find radio identifications of DSFGs, which provides a fairly quick and simple way to obtain accurate positions. Using accurate positions of their radio counterparts, DSFGs can be identified at other wavelength much more easily for spectroscopic follow-up to study their physical properties (e.g., Chapman et al. 2005; Ivison et al. 2007; Casey et al. 2012a; An et al. 2018). However, there are also disadvantages in matching radio sources to DSFGs. The major problem of this is that a fraction of SMGs/DSFGs have no radio identification. The radio matching percentage of FIR/submm sources depends on the quality of both radio and FIR/submm data as well as some other factors. For example, Chapman et al. (2003) and Ivison et al. (2007) find their radio matching rates of SMGs are around $60-70\%$, while Ivison et al. (2010b) has a $\sim 50\%$ radio identification rate for their *Herschel* sample. Barger et al. (2012) shows that all of their 16 SMA-detected SMGs have secure radio counterparts, whereas Ivison et al. (2010a) and Jarvis et al. (2010) find that only $\sim 10\%$ of their FIR samples have robust radio IDs. This lack of radio counterparts for some SMGs/DSFGs could be a result of these sources lying at higher redshifts. As radio observations do not benefit from a very negative K-correction unlike the case of FIR/submm observations discussed in Section 1.3.4, sources at $z \geq 3.5$ are very faint in the radio and hence are hard to detect. This suggests the possibility that SMGs with no radio counterpart lie at $z \geq 3.5$. However, this is further complicated by other factors, such as the multiplicity of DSFGs (see Section 1.4.2) and the depths of the observations used which could also lead to a lower radio-FIR/submm matching rate. Therefore, the non-detection of SMGs in the radio alone cannot be used as a criterion to select high-redshift sources, and it only hampers the effectiveness of radio identifications in FIR/submm studies. Nonetheless, the situation is improving as next-generation radio telescopes are already in use or coming online soon (e.g., ASKAP, MeerKAT, SKA). These instruments will be able to observe the radio sky with unprecedented angular resolution, sensitivity and mapping speed, hence identifying

most FIR/submm sources in the radio.

In summary, thanks to the FIRC and interferometry, radio observations can greatly help identifications of SMGs/DSFGs at other wavelengths and hence accurate determination of their physical properties. Its effectiveness is, however, hindered by the fact that a fraction of DSFGs is not detected in the radio. With the upcoming next-generation radio telescopes, more DSFGs will be seen in radio observations, providing greater help to DSFG studies.

1.5 Summary of the following chapters

The aim of this thesis is to study *Herschel*-selected high-redshift DSFGs and the environment they reside in. Multi-wavelength observations of individual DSFGs and DSFG protocluster candidates at potential high redshifts are used to investigate their nature. The work in this thesis thus improves our understanding of the properties of high redshift DSFGs and DSFG protoclusters, as well as providing promising targets for future studies.

In Chapter 2, we present MeerKAT 1.28GHz radio imaging of 3 *Herschel* candidate protoclusters, with a synthesised beam size of $\sim 7.5'' \times 6.6''$ and a central thermal noise down to $4.35 \mu Jy/beam$. Our source counts are consistent with other radio counts with no evidence of overdensities. Around 95% of the *Herschel* sources have 1.28GHz IDs. Using the *Herschel* $250\mu m$ primary beam size as the searching radius, we find that 54.2% of *Herschel* sources have multiple 1.28GHz IDs. With accurate MeerKAT positions, we cross-match our *Herschel* sources to optical/NIR data followed by photometric redshift estimations. By removing $z < 1$ sources, the density contrasts of two of the candidate protoclusters increase, suggestive of them being real protoclusters at $z > 1$. There is also potentially a $0.9 < z < 1.2$ overdensity associated with one candidate protocluster. Our average FIR-radio correlation coefficient $q_{250\mu m}$ is 2.33 ± 0.26. Adding $q_{250\mu m}$ as a new constraint, the probability of finding chance-aligned sources is reduced by a factor of ~ 6, but with the risk of discarding true identifications of radio-loud or radio-quiet sources. In summary, photometric redshifts from radio-optical cross-identifications have provided some tentative evidence of overdensities aligning with two of the candidate protoclusters.

In Chapter 3, we present deep SOFI J, H, K_sobservations of 13 candidate DSFG protoclusters selected from Greenslade et al. (2018). Raw SOFI images are reduced using IRAF, Python package Astropy CCDPROC and ASTROALIGN. Sources are extracted from the reduced images with Source Extractor and their photometry is calibrated to the 2MASS photometry. There are in total 2238 sources in all fields, reaching 5σ detection down to $J = 23.07 \pm 0.21$, $H = 23.12 \pm 0.33$ and $K_s = 21.17 \pm 0.13$ mag. By cross-matching our SOFI sources to other ancillary data, we construct 7-9 band photometry in the optical/NIR for all sources. Photometric redshifts of sources are computed via the photometric redshift code EAZY, with estimated average photo-z uncertainty to be ~ 0.09. Comparison between the photo-z distribution of sources in our target fields and the photo-z distribution of sources in a randomly selected blank field reveals that there are overdensities at $4 \leq z \leq 5$ in six of our thirteen target fields. This provide evidence that these six candidate protoclusters are real DSFG protoclusters lying at high redshifts, making them promising targets for future spectroscopic follow-up observations.

In Chapter 4, we carry out multi-wavelength cross-identifications of 18 SMA single-detection $500\mu m$ risers plus one two-detection source, selected from Greenslade et al. (2020). With multi-band photometry obtained, photometric redshifts of the selected sample are estimated using three different codes, EAZY (Brammer et al. 2008), MMPZ (Casey 2020) and CIGALE (Boquien et al. 2019). We find that MMPZ provides the most reliable estimations but with large uncertainties. Overall, we find a mean redshift of $< z >= 3.874 \pm 0.233$ for our 500 μmriser sample with single SMA detection. 10 of the 19 sources have $2.6 < z < 3.7$, lower than the $z > 4$ expectation. We then compute SEDs of the target objects with CIGALE using known spectroscopic redshifts and MMPZ redshifts, deriving their physical properties. The mean SFR and stellar mass of our sample are $4107 \pm 800\ M_{\odot} yr^{-1}$ and $< M_* >= 2.297 \pm 0.459 \times 10^{12}\ M_{\odot}$, with $L_{IR} > 10^{13}\ L_{\odot}$ for nearly all sources. The IR luminosities of our 500 μm risers suggest that they are ULIRGs, while their SFR-stellar mass relation implies that either $z < 3.8$ 500 μm risers are starbursts which are more intense than typical star-forming galaxies at similar redshifts, or their redshifts are underestimated. We also find that the fraction of AGN contribution has no significant impact on the SFR of our sample, consistent with previous studies on X-ray selected AGN of SMGs.

Chapter 5 includes a summary our the results in this thesis, describes the outlook for future DSFG

studies and suggest further potential follow-up observations. Throughout this thesis, unless otherwise stated, AB magnitude is used and we assume a standard cosmology, with $H_0 = 70.0\ kms^{-1}Mpc^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$.

Chapter 2

Herschel protocluster candidates: MeerKAT 1.28GHz observation

The majority of the content within this chapter is based upon a paper submitted to MNRAS

2.1 Introduction

As discussed in Section 1.4.5 of Chapter 1, high density regions at $z > 2$ that will become today's galaxy clusters are important in understanding the evolution of massive galaxies and large scale structure. Being progenitors of present day galaxy clusters, these regions, which are called protoclusters, are still poorly understood due to a lack of observations. Searches for $z < 1$ galaxy clusters typically look for X-ray emission from the hot intracluster medium (ICM) of galaxy clusters, the Sunyaev-Zeldovich (SZ) effect caused by electron scattering of CMB photons in the ICM, and sequences of red quiescent elliptical galaxies which are often found in the cores of clusters (e.g., Kravtsov and Borgani 2012; Overzier 2016). However, none of these methods are effective in protocluster searches, because of their high redshift and lack of hot ICM. Furthermore, galaxies in the cores of protoclusters which have got to evolved into the quiescent ellipticals are actively forming stars, hence protoclusters can not be found by searching for sequences of red quiescent ellipticals. Currently, the only way of finding protoclusters is to look for physical overdensities of galaxies at high redshifts. Such overdensities are

rare and hard to detect due to their high redshifts, therefore a number of different galaxy populations are used as signposts of protoclusters, as discussed in Section 1.4.5 of Chapter 1.

As the progenitors of massive elliptical galaxies at $z = 0$, which are often found in the cores of galaxy clusters (e.g., Hopkins et al. 2008; Toft et al. 2014), DSFGs are thought to trace high-redshift protoclusters. Indeed, evidence has been found that a significant number of protoclusters are made up of DSFGs, suggesting DSFGs are a good tracer of protoclusters (e.g., Daddi et al. 2009b; Capak et al. 2011; Clements et al. 2014; Casey et al. 2015; Clements et al. 2016; Greenslade et al. 2018; Miller et al. 2018; Oteo et al. 2018; Hill et al. 2022). Nonetheless, some studies have also shown that DSFGs do not always trace overdensities at high redshifts. The apparent correlation between DSFGs and high-redshift overdensities could be a result of large Poisson noise of DSFGs due to their rarity and blending of multiple faint sources lying at different redshifts in single-dish observations (e.g., Chapman et al. 2009; Miller et al. 2015; García-Vergara et al. 2020). Whether DSFGs trace protoclusters currently remains an open question. More observations of protoclusters in various stages of evolution are therefore needed to answer this question. Typical angular scales of DSFG-rich protoclusters vary between a few arcseconds (e.g., Capak et al. 2011; Oteo et al. 2018) to tens of arcminutes (e.g., Casey et al. 2015; Muldrew et al. 2015), and the number density of DSFGs is also very low compared to optical/NIR galaxies. Therefore, large-area FIR/submm surveys with an effective method to look for such overdensities in large datasets are vital for searching DSFG overdensities.

A method of searching for protoclusters of dusty starbursts using *Planck* and *Herschel* imaging data in the FIR and submm was first proposed by Negrello et al. (2005). Due to the large beam sizes of single-dish FIR/submm instruments, a bright DSFG may in fact be a blend of multiple fainter sources clustered inside the instrument beam. Then, further higher-resolution observations at similar wavelength can then be used to resolve this apparently single DSFG into individual sources and confirm the overdense nature of this region. The origin of such overdensities could be real structures such as protoclusters, or mere chance alignments, or even several less populated real structures aligning along the line of sight (Flores-Cacho et al., 2016). This method of selecting DSFG overdensities were later successfully demonstrated, with a number of candidate protoclusters as well as real protoclusters confirmed in this way (e.g., Clements et al. 2014; Flores-Cacho et al. 2016; Miller et al. 2018; Greenslade et al. 2018;

Martinache et al. 2018).

Using *Herschel* data, Clements et al. (2014) examined 16 *Planck* sources in the Early Release Compact Source Catalogue of *Planck* (ERCSC; Piacentini et al. 2011), finding four candidate protoclusters at $z \sim 1 - 2.3$. Following Clements et al. (2014), Greenslade et al. (2018) combine the full the *Planck* Catalogue of Compact Sources (PCCS; Piacentini et al. 2011; Ade et al. 2014, 2016) and large *Herschel* surveys to search for potential protoclusters. Compact sources are first extracted from Planck observations at 857GHz ($350\mu m$) and 545GHz ($550\mu m$) at $4.67'$ resolution, then, using *Herschel*/SPIRE (250, 350, $500\mu m$) images from large surveys such as H-ATLAS (Eales et al. 2010) and HerMES (Oliver et al. 2012), these *Planck* compact sources are resolved to $18'' - 36''$ resolution. Using archival catalogues and visual inspection, *Planck* sources that might originate from overdensities of multiple dusty starbursts while also being overdense in *Herschel*/SPIRE sources $> 3\sigma$ are selected as candidate protoclusters. Greenslade et al. (2018) eventually identify 27 candidate protoclusters which are estimated to lie at $z > 1$. They also have total estimated star-formation rates (SFRs) of $> 10,000\ M_{\odot}/yr$. With such high SFRs, a significant fraction of their cluster stellar mass is assembled in this phase. Thus these candidate protoclusters likely probe an important epoch of cluster formation when rapid stellar mass assembly takes place. However, with their exact nature unknown, the poor resolution and large beam size of the *Herschel* SPIRE data make it difficult to determine the exact locations (uncertainty $> 18''$) of these dusty starburst protocluster galaxies for optical/NIR identification and spectroscopic follow-up.

In this chapter, we continue the work of Greenslade et al. (2018) by carrying out MeerKAT (Jonas 2009) 1.28GHz radio imaging observations centred on 3 of the 27 DSFG protocluster candidates identified by Greenslade et al. (2018), to permit their further identification in optical/NIR data. As mentioned in Section 1.4.6 in Chapter 1, radio identification of FIR/submm objects is a fairly simple and reliable method to obtain accurate locations, and they have been widely used to study DSFGs (e.g., Ivison et al. 2007, Chapin et al. 2009, Casey et al. 2012a, Roseboom et al. 2012, Hodge et al. 2013, An et al. 2018). There are several major advantages in cross-identifying DSFGs in the radio: (a) dust is transparent at radio wavelengths so radio observations will trace the cosmic star formation rate density unbiased by dust emission and absorption; (b) high-resolution radio interferometry observations will

greatly help the cross-identification of multi-wavelength counterparts of *Herschel* sources with precise position measurements, as well as investigating the radio morphologies of DSFGs; (c) in optical/NIR images, each *Herschel* SPIRE source (beam size of $\sim 18'' - 36''$) might overlap multiple optical/NIR sources. This is a difficult situation if we want to do optical/NIR spectroscopy because the counterpart of the FIR source is uncertain. Thanks to the well-known correlation between a galaxy's FIR and radio emission (e.g., Helou et al. 1985, De Jong et al. 1985), we can use MeerKAT to obtain locations down to $\approx 1''$ rms and the optical/near-infrared counterpart is thus unambiguous; (d) comparing a source's radio flux to its FIR flux can help identify the type of source that contributes to the high FIR flux in our candidate protoclusters, e.g. lensed objects, line-of-sight projections, enhanced starbursts, and radio-loud AGNs. The primary goal of this study is to acquire accurate MeerKAT radio identifications of three *Herschel* candidate protoclusters. This will significantly simplify identifications of *Herschel* sources in the optical/NIR, hence confirming if the candidates are real or not by fitting SEDs and redshifts to the sources. Furthermore, we will be able to study the stellar mass assembly during their dusty starburst phase and also the impact of environment on the starburst protocluster galaxies (e.g. AGN feedback). More broadly, this work demonstrates how high-quality radio data can help FIR studies, as next-generation radio telescopes such as the Square Kilometre Array (SKA; Dewdney et al. 2009) are coming closer to reality.

This chapter is structured as follows. In Section 2.2 we introduce the data used in this work. Section 2.3 describes the source extraction process for our MeerKAT images, and how we account for the extraction completeness and reliability. Source counts extracted from our MeerKAT images are presented in Section 2.3.3, along with the investigation of potential overdensities in these images. In Section 2.4 and Section 2.4.1 we present the cross-identifications (cross-IDs) of *Herschel* sources with MeerKAT sources and investigate the multiplicity of *Herschel* sources seen by MeerKAT. Analysis of the FIR-radio correlation and *Herschel* colours are shown in Section 2.4.2 and 2.4.3. In Section 2.5, we discuss how high-quality radio data can help to improve the robustness of *Herschel* identifications in the radio. We show how high quality MeerKAT data help the identification of *Herschel* sources in optical/NIR data in Section 2.6, followed by cross-IDs to KiDS and VIKING data which allows photometric redshift estimations in Section 2.6.1. Section 2.7 concludes the chapter.

2.2 Data

2.2.1 Target Selection

The three protocluster candidates studied in this chapter are chosen from the 27 protocluster candidates selected by Greenslade et al. (2018). We briefly outline their selection method here.

By cross-matching *Planck* compact source catalogue sources with *Herschel* catalogues using the 4.63 arcmin *Planck* beam size at 857GHz as the matching radius, Greenslade et al. (2018) identifies 4445 *Herschel* sources matched with 431 *Planck* compact sources. They then exclude *Planck* sources whose counterparts are not potential protoclusters, such as local galaxies, stars or Galactic cirrus. For the remaining *Planck* sources, Greenslade et al. (2018) counts the number of 250, 350, and 500 μm sources with fluxes > 25.4 mJy that lie within 4.63 arcmin of the *Planck* position, with the flux limit chosen to compare to published number counts. Any object with a 3σ overdensity in any of the three SPIRE bands is classed as a candidate protocluster of galaxies. This results in 27 candidates found.

Among these 27 candidate protoclusters, the 3 targets selected for this work have the most complete archival multi-band photometric data for further studies on the nature of DSFG protoclusters. The names of the three selected targets are PLCKERC857 G014.99-59.64, PLCKERC857 G017.86-68.67 and PLCKERC857 G257.09-87.10. For convenience, the three observed fields that contain the three candidate protoclusters are referred as G014, G017 and G257 hereafter.

2.2.2 *Herschel* Data

The candidate protoclusters are selected from the H-ATLAS DR2 (Smith et al. 2017b, Maddox et al. 2018) South Galactic Pole (SGP) data. The *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS) is the largest single key project carried out in open time with the *Herschel Space Observatory* (Pilbratt et al. 2010). It is a survey of 660 deg^2 with the PACS and SPIRE cameras in five photometric bands: 100, 160, 250, 350, and 500 μm. Data release 2 of H-ATLAS covers the SGP fields, and the total 1σ map sensitivity is 5.7, 6.0, and 7.3 mJy at 250, 350, and 500μm, respectively. The SGP field,

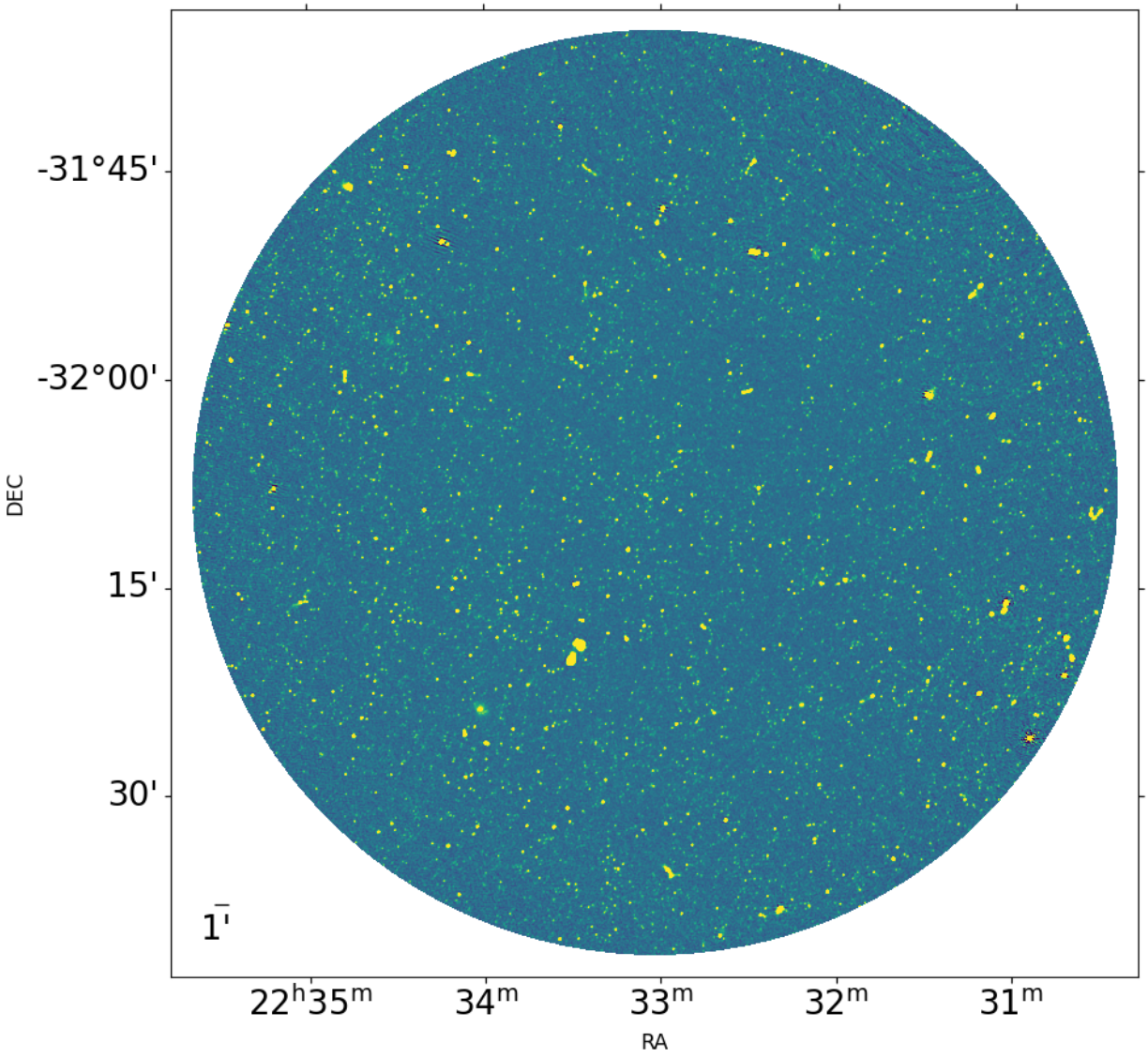


Figure 2.1: The MeerKAT 1.28GHz images presented in this work. From the first to the last image: G014, G017, and G257. The right ascension (RA) and the declination (DEC) in each image are shown by the x-axis and the y-axis respectively. A scale bar of 1′ is shown in the bottom left corner of each panel.

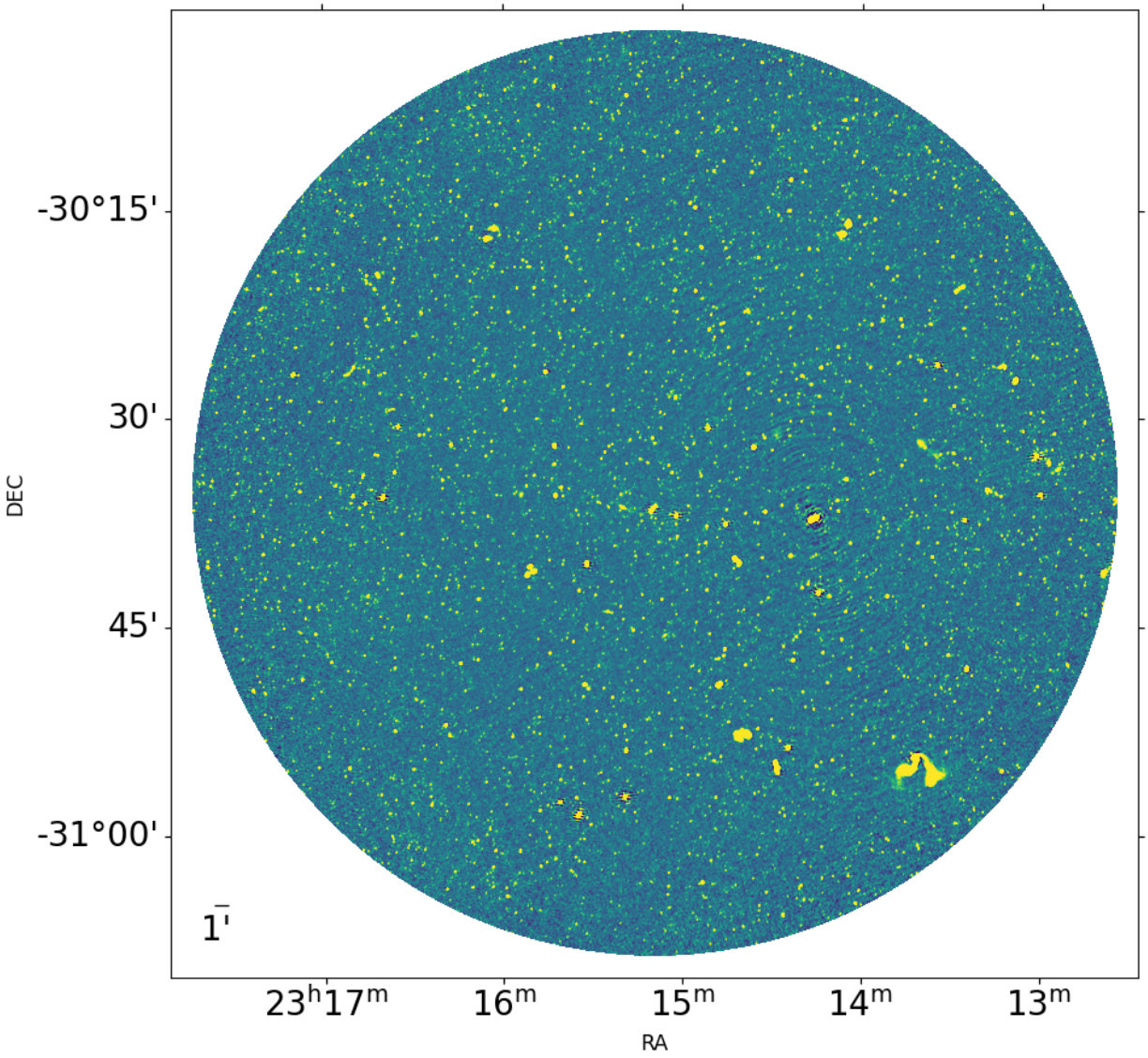


Figure 2.1 continued, G017.

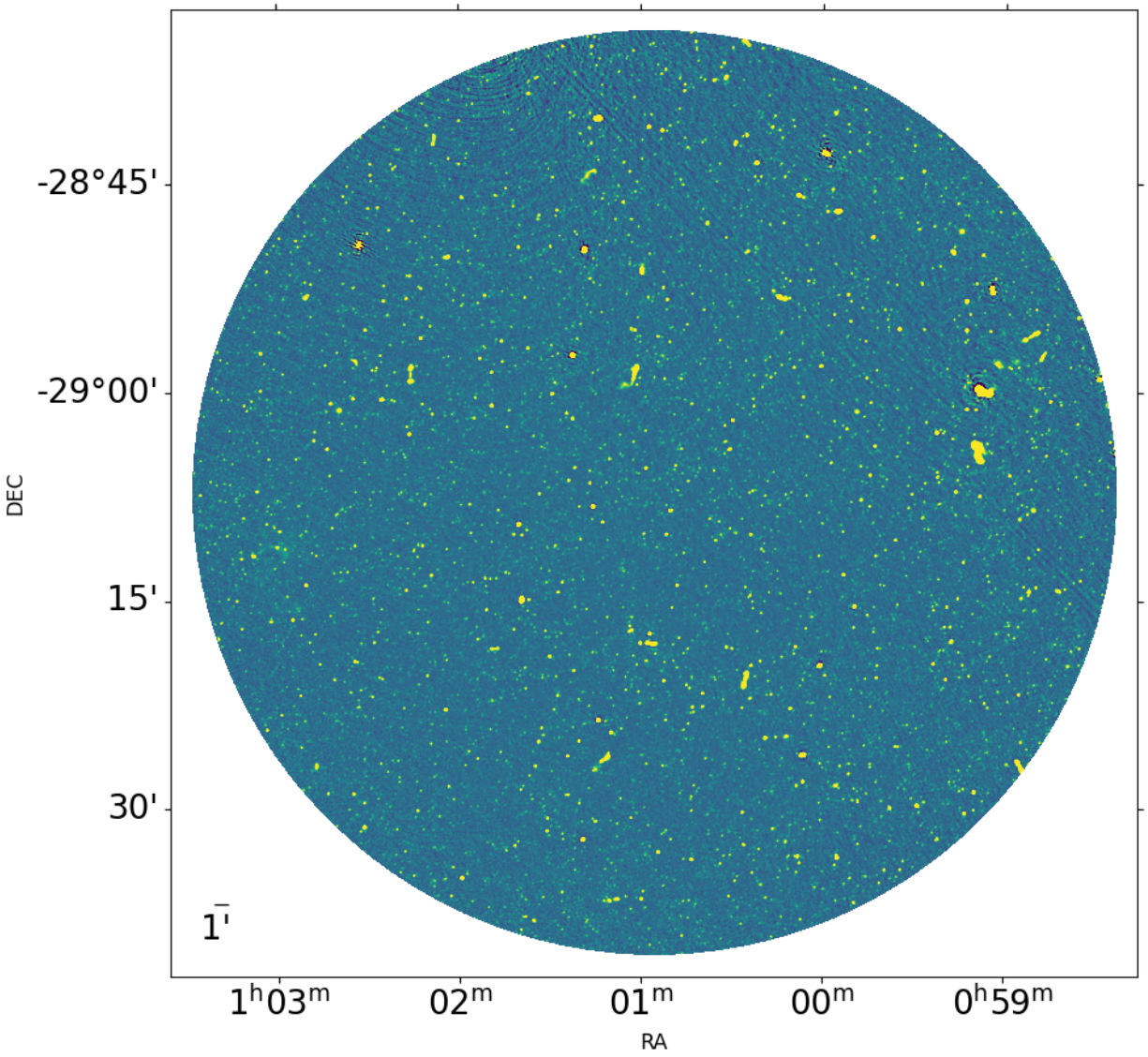


Figure 2.1 continued, G257.

which includes our MeerKAT maps, is centred approximately at right ascension $0h6m$ and declination $-32°44'$(J2000) and has an area of 317.6 deg^2. The catalogues of FIR and submillimeter sources for the SGP field contains 193,527 sources detected at more than 4σ in any of the 250, 350 and 500 μm bands. The source detection process is based on the 250 μm map.

2.2.3 Ancillary Data

The SGP field is covered by many large surveys which provide extensive data for further studies. These include the 2-Degree-Field Galaxy Redshift Survey (2dF, Colless et al. 2001), the VST Kilo-Degree Survey (KIDS, de Jong et al. 2013) which is an optical imaging survey in u, g, r, i, and the VISTA Kilo-Degree Infrared Galaxy Survey (VIKING, Edge et al. 2013) which is a NIR imaging survey in Z, Y, J, H and K_s. The H-ATLAS DR2 also has some overlap with the Dark Energy Survey (DES, Abbott et al. 2018) which is a deep optical survey in u, g, r, i, z, Y, as well as with the recently-released ESO public survey Southern H-ATLAS Regions Ks-band Survey (SHARKS, Dannerbauer et al. 2022).

2.2.4 MeerKAT 1.28GHz Data

The three protocluster candidate fields were observed[1] with MeerKAT's L-band receivers (856–1712 MHz) between 12 and 15 June 2019, reaching a 5σ depth of $\sim 16\mu Jy$. Details of the observations are provided in Table 2.1. The FWHM primary beam size is $\sim 34'$ at 1.28GHz, covering an area of ~ 0.971 deg^2 for each image. Figure 2.1 displays our MeerKAT observations of the three fields.

The MeerKAT data used in this chapter is reduced by our collaborator Professor Ian Heywood. The data were retrieved from the MeerKAT archive[2] in Measurement Set format, averaged down in frequency to 1024 channels prior to transfer. Initial flagging of the data was followed by the determination of bandpass, delay and time-dependent complex gain solutions from scans of the primary (PKS B1934-638 in all cases) and secondary calibrators, using the CASA package casa2022. The resulting gain solutions were applied to the target scans, which were then flagged using TRICOLOUR. This implements

[1]Project code SCI-20190418-LL-01
[2]`https://archive.sarao.ac.za`

Table 2.1: Names, positions, associated secondary calibrators, and total on-source time for each of the protocluster candidate target fields.

Name	RA (J2000)	Dec (J2000)	Secondary calibrator	Time on-source (h)	Synthesised beam size	Central thermal noise ($\mu Jy/beam$)
G014.99−59.64	$22^h33^m02^s.16$	$-32°08'16''.8$	J2152−2828	3.75	$7.5'' \times 6.6''$	4.35
G257.09−87.10	$01^h00^m55^s.68$	$-29°07'19''.2$	J0025−2602	4	$7.6'' \times 7.0''$	5.04
G017.86−68.67	$23^h15^m09^s.36$	$-30°35'27''.6$	J2341−3506	3.5	$7.5'' \times 6.3''$	5.62

a MeerKAT-optimised version of the SUMTHRESHOLD algorithm (Offringa et al., 2010), accelerated using DASK-MS (Perkins, 2022).

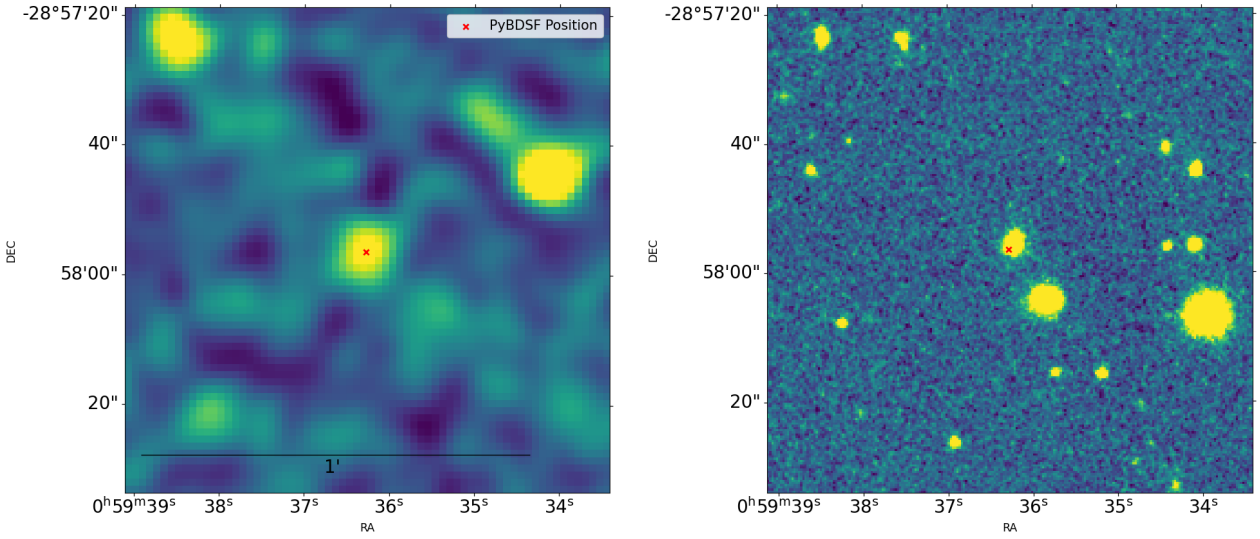
A single round of phase and delay self-calibration is performed, by imaging using WSCLEAN (Offringa et al., 2014) with the deconvolution constrained to regions of genuine emission, and deriving the self-calibration solutions using CUBICAL (Kenyon et al., 2018). Following this the data were subjected to a round of direction-dependent calibration, with imaging performed using DDFACET (Tasse et al., 2018, 2023), and directional gain solutions derived for ∼10 user-defined directions using KILLMS (Smirnov and Tasse, 2015; Tasse, 2023). The final image was primary beam corrected by dividing it by an azimuthally-averaged model of the Stokes I beam derived using the KATBEAM[3] package. Full details of the reduction process are provided by Heywood et al. (2022). The scripts that performed the reduction are available online for general use (Heywood, 2020). All processing was done using the *ilifu* cluster[4] in Cape Town, and all of the packages mentioned above were containerised using SINGULARITY (Kurtzer et al., 2017).

2.3 Source Extraction

2.3.1 Extraction Software

To improve the robustness of the final catalogues as well as comparing different source extraction tools, sources are extracted from the MeerKAT maps with both the radio source extraction tool PyBDSF (the Python Blob Detector and Source Finder, Mohan and Rafferty 2015) and the SExtractor (Source Extractor, Bertin and Arnouts 1996).It should be noted that SExtractor is not designed for radio images

[3] https://github.com/ska-sa/katbeam
[4] https://www.ilifu.ac.za/



(a) MeerKAT

(b) SHARKS

Figure 2.2: An example of a faint source missed by the first run of PyBDSF. MeerKAT 1.28GHz image (a) and SHARKS Ks image (b) are positionally matched and shown, with RA as x-axis and DEC as y-axis. As can be seen from the SHARKS Ks image, there is a source at the centre of the images. However, this faint source is not detected by the first run of PyBDSF. The second run of PyBDSF on the residual map of the first run successfully picks up this source.

and it may find unphysical sources smaller than the synthesised beam.

PyBDSF first calculates the background rms and mean. Regions where fluxes are 3σ above the mean flux of the image and peak fluxes 5σ above the mean are fitted with Gaussians then extracted as valid sources. Visual inspection finds that a single run of PyBDSF extraction on our images leaves a significant number of faint sources undetected, an example of such cases is shown in Figure 2.2. Therefore, to increase the extraction completeness, we have two PyBDSF runs for each map: the original MeerKAT image is used for the first run, then a second extraction run is carried out on the residual map of the first run. The default flagging options are used to exclude residuals from the first run being detected as sources. Visual inspection of the second extraction run does not find any false detection originating from first run residuals. The two resulting catalogues are then concatenated into one final catalogue. The final PyBDSF catalogues of the G014, G017 and G257 MeerKAT images contain 5925, 5415 and 5795 sources respectively.

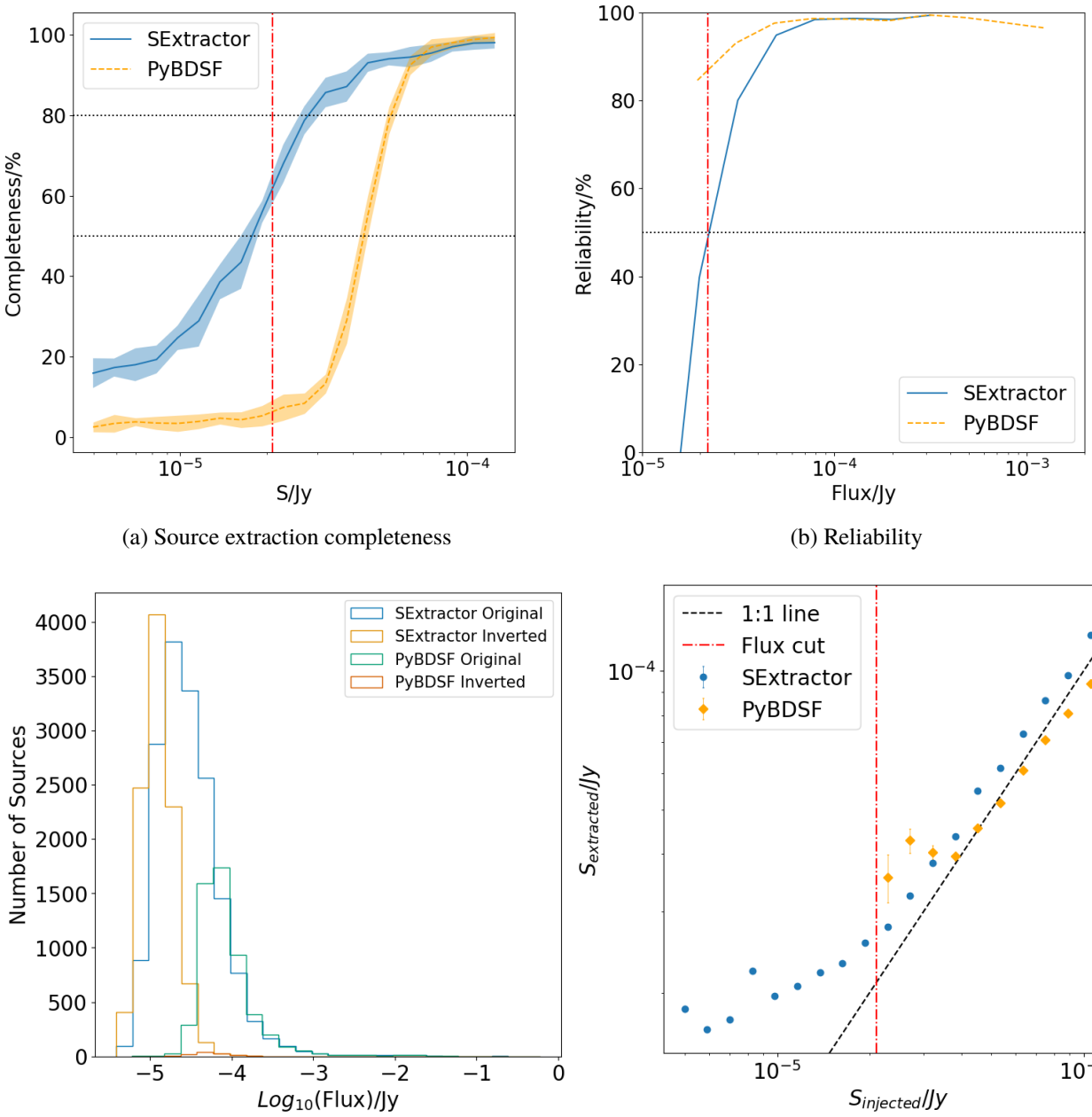
SExtractor uses a different approach. Two passes are made through the data. During the first pass, a

model of the sky background is built, and several global statistics are computed. During the second pass, image pixels are background-subtracted, filtered and segmented on-the-fly. Detections are then deblended, pruned (or, "CLEANed"), and enter the measurement phase. Finally, the measured quantities which are defined by the user are written to the output catalogue, after cross-matching with an optional input list. We set the single-pixel detection threshold to be $\sim 1.8\sigma$ with the minimum number of connected pixels above the threshold set to 5, which results in a $\geq 4\sigma$ final detection limit for a valid source. The default settings are used for the other parameters. The final SExtractor catalogues of the G014, G017 and G257 MeerKAT maps contain 16403, 16557 and 16339 sources respectively. For the measured total flux densities of the detected sources, we use FLUX_BEST from the SExtractor catalogue. In most cases, FLUX_BEST uses the fluxes measured by integrating pixel values within an adaptively scaled aperture, but it will use the corrected isophotal fluxes if the contribution of other nearby sources exceeds 10%. The catalogued fluxes are in units of the image pixel which is $Jy/beam$. We have converted these to total fluxes in Jy by:

$$Total\ flux(Jy) = Flux(Jy/beam) \times \frac{Pixel\ area(deg^2)}{Synthesised\ beam\ area(deg^2)} \tag{2.1}$$

2.3.2 Extraction Completeness & Reliability

By injecting fake sources into the original MeerKAT maps and re-extracting sources following the same extraction procedures described in the last section, we quantify the extraction completeness of both PyBDSF and SExtractor as shown in Figure 2.3a. The fake sources are compact 2-D Gaussians which are randomly distributed on the maps, the size of the Gaussians is comparable to the general size of real sources, most of which have FWHMs of $\sim 6'' - 12''$. 100 sources with the same total flux are injected and re-extracted, then the same procedure is repeated with increasing total flux until the completeness reaches 100%. Injecting more than 100 fake objects into our images will lead to simulated sources confused with real sources, therefore at each flux level we iterate the injection process 10 times, then take the average completeness as the final number to produce a robust measure of completeness. To avoid overlapping between injected sources and real MeerKAT sources, all simulated sources are set to be at least $6''$ away from any real source.



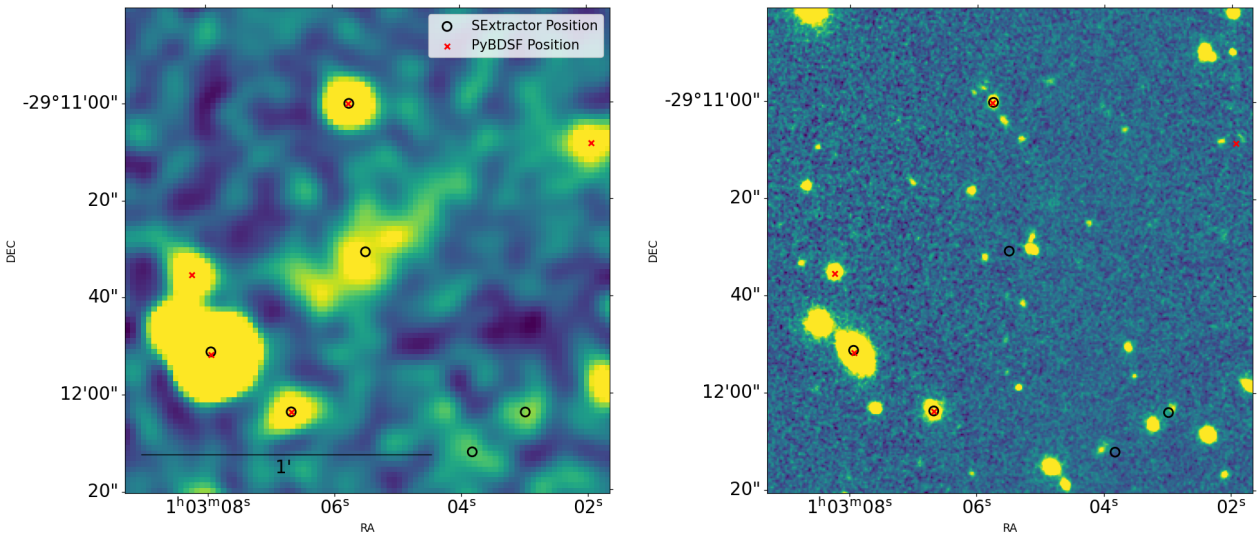
(a) Source extraction completeness

(b) Reliability

(c) Flux distributions

(d) SExtractor-PyBDSF flux comparison

Figure 2.3: (a) Extraction completeness as a function of 1.28GHz flux. The shaded areas show the respective standard deviation, the black dotted lines indicate 50% and 80% completeness, and the red dashed-dotted line shows the $S = 0.021 mJy$ flux cut. It can be seen that SExtractor (blue solid line) has better completeness at fainter fluxes. SExtractor reaches a 50% completeness at $S \approx 0.017\ mJy$ while PyBDSF (orange dashed line) reaches a 50% completeness at $S \approx 0.035\ mJy$ with smaller uncertainties. (b) The reliability of the source extraction with SExtractor and PyBDSF. The blue solid line shows the result from SExtractor and the orange dashed line shows the reliability of PyBDSF. The black dotted line indicates 50% reliability and the red dashed-dotted line shows the $S = 0.021 mJy$ flux cut. PyBDSF barely picks up any spurious objects, whereas over half of sources detected by SExtractor are likely spurious at fluxes lower than $\sim 0.021\ mJy$. (c) Flux distributions of sources extracted from both the original maps and the flux-inverted maps. (d) Average total flux of injected sources recovered by SExtractor and PyBDSF vs. average total flux of sources injected into the map in Jy. The blue dots are SExtractor fluxes and the yellow diamonds are PyBDSF fluxes. The black dashed line is the 1:1 line, the red dashed-dotted line is the $S = 0.021\ mJy$ flux cut. At $S > 0.021\ mJy$ SExtractor slightly overestimates the flux density of a source, while PyBDSF only have reliable flux density measurements at $S > 0.030\ mJy$ and it increasingly underestimates source flux densities as the flux level goes higher. At $S > 0.021\ mJy$, a 1.42 linear proportionality is found between PyBDSF flux measurements and SExtractor flux measurements.



(a) MeerKAT

(b) SHARKS

Figure 2.4: An example of a faint source detected by SExtractor but not PyBDSF. MeerKAT 1.28GHz image (a) and SHARKS Ks image (b) are positionally matched and shown, with RA as x-axis and DEC as y-axis. As can be seen from the SHARKS Ks image, there is a faint source right at the centre of the images. However, this faint source is only detected by Source Extractor while ignored by PyBDSF. PyBDSF only detects the much brighter sources in this region.

It can be seen that SExtractor has higher completeness at fainter fluxes, as it reaches 50% and 80% completeness at $S \approx 0.018\ mJy$ and $S \approx 0.027\ mJy$ while PyBDSF has the same completenesses at $S \approx 0.036\ mJy$ and $S \approx 0.048\ mJy$, although PyBDSF's completeness has smaller variations. This indicates that SExtractor can extract sources at $0.018 \leq S \leq 0.036\ mJy$ which are difficult for PyBDSF to detect. An example is shown in Figure 2.4.

To assess the reliability of the two extractors, in other words what fraction, as a function of flux, of the detected sources are real objects instead of noise, we invert our images by multiplying the original flux maps by -1, then re-extract catalogues with SExtractor and PyBDSF using the same parameters. The resulting flux distributions are then compared to the original distributions, as shown by Figure 2.3c. Objects detected in the flux-inverted images are mostly spurious sources induced by noise or random fluctuations. Therefore measuring the number of objects in the flux-inverted images will provide an estimate of how many spurious sources our source extractors detect at different flux levels in real

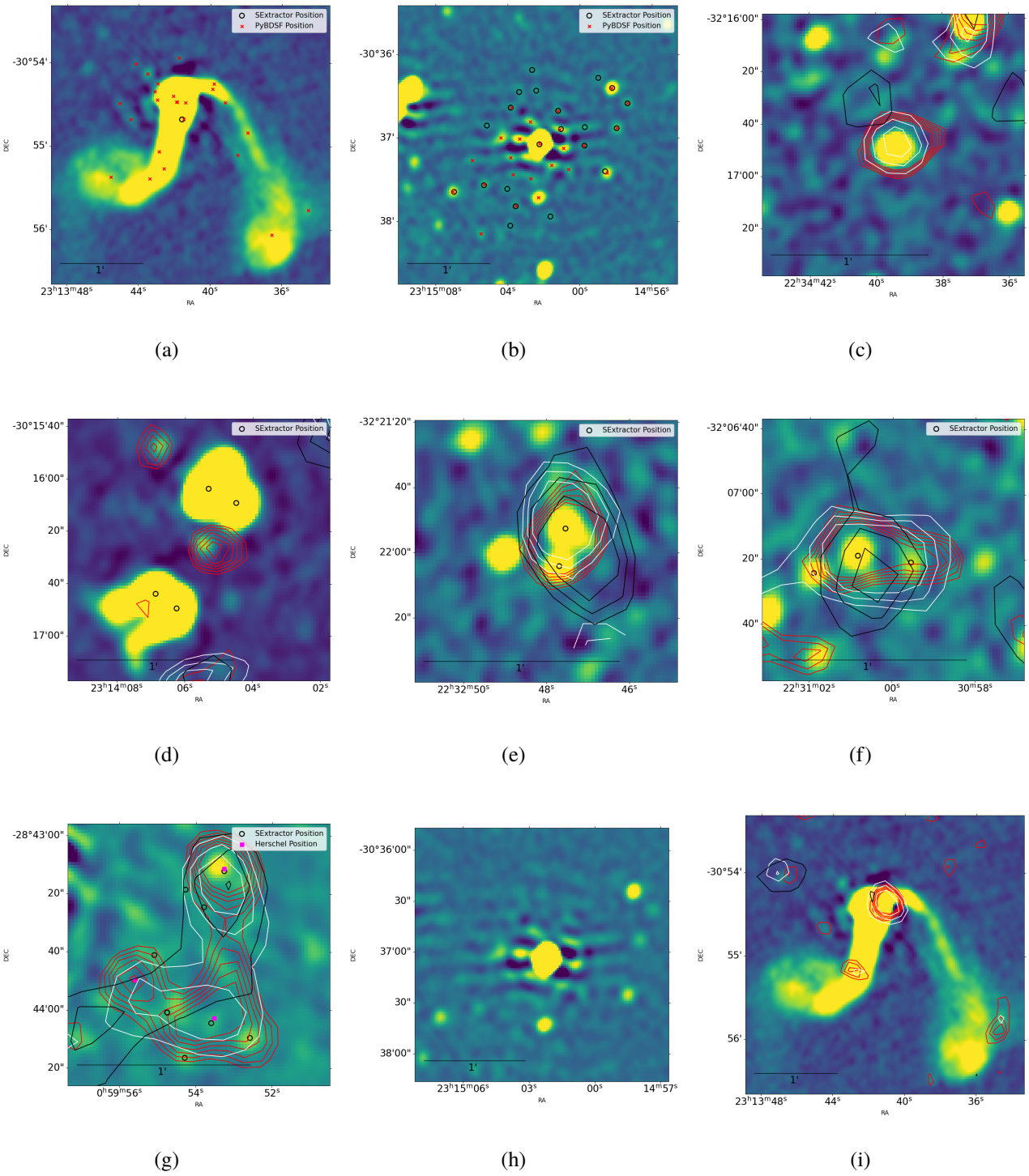


Figure 2.5: MeerKAT image cutouts. The background images are the 1.28GHz MeerKAT observations, the red contour illustrates the $5-10\sigma$ *Herschel* $250\mu m$ flux density, the white and black contours are the $3-5\sigma$ *Herschel* 350 and $500\mu m$ flux density respectively. A scale bar of $1'$ is shown in every cutout. Positions of SExtractor detected MeerKAT sources are shown in black circles, PyBDSF positions are shown in red crosses and *Herschel* positions are shown in magenta squares. (a) A bright source with extended morphology. (b) A bright source with strong beam artefacts. (c) An example of the *Herschel*-MeerKAT cross-identification. (d) A true *Herschel*-MeerKAT cross-identification not being matched due to this source's extended morphology. The centre of this 1.28GHz source with two jet-like structures clearly coincides with the *Herschel* contour, but SExtractor does not recognise its centre as an individual source. (e) A *Herschel* source with two MeerKAT counterparts. (f) A *Herschel* source with three MeerKAT counterparts. (g) A *Herschel* source with apparent MeerKAT multiplicity of 7. There are actually 3 *Herschel* sources close to each other, overlapping with 7 faint MeerKAT sources. The true number of counterparts of each *Herschel* sources will be $1-4$. (h) The MeerKAT sources with the lowest $q_{250\mu m}$. It is a very bright radio sources with beam artefacts around it. (i) A radio-loud source with low $q_{250\mu m}$. Note that this is the same source shown in (a). This radio source is associated with two jets orthogonal to each other. A *Herschel* source coincides with the centre of this extended radio structure, suggestive of AGN activity.

images. We define the reliability as:

$$Reliability = \frac{N_{real} - N_{flux-inverted}}{N_{real}} \tag{2.2}$$

where N_{real} and $N_{flux-inverted}$ are the number of sources extracted from real images and flux-inverted images respectively. Figure 2.3b shows the reliabilities of the two extractors as a function of flux. For PyBDSF, very few spurious objects are detected as real sources, showing a very high source extraction reliability. For SExtractor, spurious objects contribute over half the number counts at fluxes lower than $\sim 0.021\ mJy$ as indicated by the red dashed line, which means sources extracted by SExtractor with $S \leq 0.021\ mJy$ cannot be considered reliable. Comparing Figure 2.3a&b, while PyBDSF has better extraction reliability than SExtractor at any flux, SExtractor has much better completeness at fainter fluxes. SExtractor can detect sources down to $S = 0.021\ mJy$ with reasonable reliability and a completeness of $\sim 60\%$, meanwhile PyBDSF can only reach a completeness of $\sim 7\%$ at this flux. Therefore, at this flux level, the fraction of real detected objects is far higher in the SExtractor based catalogue, even with the lower reliability.

To evaluate how well the two codes measure flux densities, we compare the average total fluxes of injected fake sources recovered by SExtractor and PyBDSF, as a function of injected flux. Figure 2.3d shows the results. It can be seen that SExtractor will slightly overestimate source flux densities but its flux density measurements are consistently reasonable at $S > 0.021\ mJy$, while PyBDSF have better flux recovery at $S > 0.030\ mJy$ but it will increasingly underestimate source flux densities as the flux level goes higher. Fitting a linear regression model to the SExtractor and PyBDSF extracted fluxes, we find a 1.42 proportionality at $S > 0.021\ mJy$, meaning that the flux density of a source measured by SExtractor is 1.42 times higher than that measured by PyBDSF on average. This offset between the measured flux densities is likely to be a result of different detection algorithms implemented by the software. PyBDSF fits Gaussians to regions that are 3σ above the mean flux of the image with peaks above 5σ, then measures the total fluxes within the fitted Gaussians, while SExtractor identifies regions which contain at least 5 connected pixels that are above 1.8σ significance then estimates the total fluxes by integrating pixel values within an adaptively scaled elliptical aperture. Therefore, a possible explanation of why SExtractor measures a higher total flux is that for each detected source, the

elliptical aperture used by SExtractor will include more pixels than the Gaussians fitted by PyBDSF, leading to a higher measured flux. Figure 2.6 shows the average sizes of simulated sources extracted by SExtractor and PyBDSF as a function of flux density. The average size of SExtractor sources decreases as the flux density increases, while the average size of PyBDSF detections remains roughly constant after the flux cut. This agrees well with the trend seen in Figure 2.3d, which is supportive of the hypothesis that SExtractor is measuring a larger area per source. Furthermore, Hale et al. (2019) finds that PyBDSF has an excess of positive residuals, suggesting that the Gaussian models of PyBDSF are under-fitting the sources in the field and hence slightly underestimating the true fluxes. After visual examination, we find that for bright sources with extended morphologies or obvious artefacts, SExtractor generally tends to blend them with nearby smaller sources or noise, boosting the measured fluxes, whereas PyBDSF is more likely to pick up small flux variations inside or around extended objects as individual sources. Figure 2.5a&b illustrates two example bright sources and the positions of objects identified by SExtractor and PyBDSF around the bright sources. As can be seen in Figure 2.5a, for a bright source with extended morphology, PyBDSF has identified numerous objects inside and around it, whereas SExtractor has taken the entire area as one large source. For a bright source with beam artefacts as in Figure 2.5b, PyBDSF has picked up several small blobs induced by the synthesised beam, while SExtractor has only identified one or two such artefacts as sources. In short, the origin of the flux measurement offset between the two source finders is likely due to the detection algorithms implemented by the two codes.

In conclusion, we decide to use the SExtractor catalogue as our final catalogue, with a flux cut at 0.021 mJy to minimise spurious objects in the catalogue while achieving a higher completeness. After this flux cut, there are 11016, 11509 and 13332 sources in the G014 G107 and G257 maps respectively.

2.3.3 MeerKAT 1.28GHz Source counts

One of the main goals of our MeerKAT observations is to aid the identification of DSFG protocluster candidates selected from *Herschel* data. To look for traces of protoclusters in the observed area, we calculate the brightness-weighted differential number counts $S^2n(S)$ (i.e. Euclidean normalised

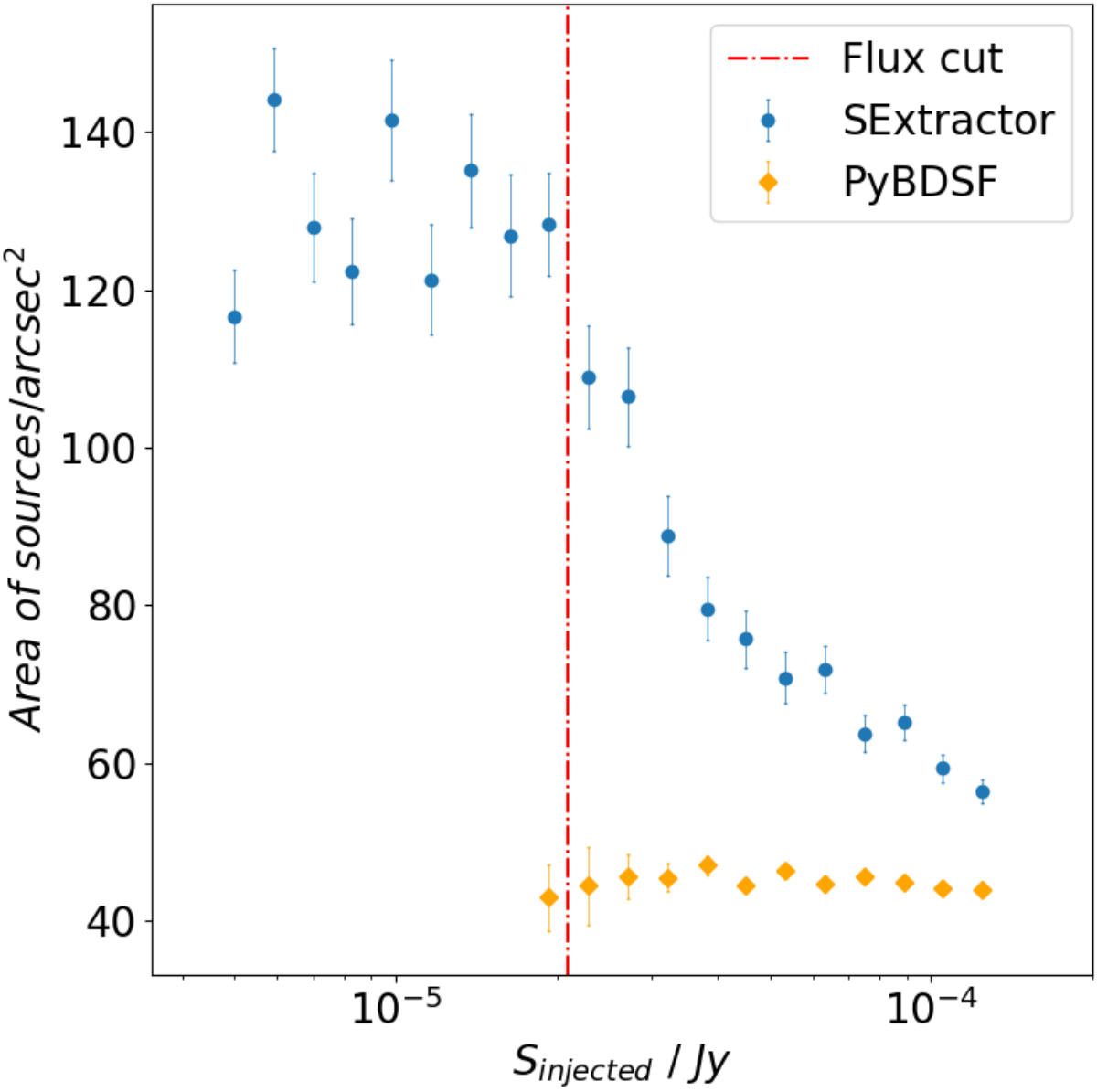


Figure 2.6: The average area of injected sources (in $arcsecond^2$) extracted by SExtractor (blue dots) and PyBDSF (orange diamonds) as a function of flux density. The red dashed-dotted line shows the $S = 0.021\ mJy$ cut. The average area of SExtractor source decreases as the flux density increases, while the area of PyBDSF detections is roughly a constant after the flux cut. This agrees well with the flux density offset between the two source finders seen in Figure 2.3d, supporting the hypothesis that SExtractor is measuring a slightly larger area per source than PyBDSF does.

number counts) for our extracted sources with a flux range between $0.021\ mJy \leq S < 1\ Jy$, where S is the flux density of the source, and n(S) is the the number of sources per steradian with flux densities between S and $S + dS$ defined as:

$$n(s) = \frac{N}{A \times dS} \tag{2.3}$$

where N is the number of sources between S and $S + dS$ and A is the total area which is $\sim 0.971 deg^2$ for each image. We count the extracted sources in bins of width 0.2 dex, centred on 1.28GHz flux densities i.e. $log_{10}[S(Jy)] = -4.6, -4.4, -4.2, -4.0, -3.8, ..., 0.0$, then compare our results with source counts at similar frequencies in the literature. Corrections based on the completeness and reliability obtained in Section 2.3.2 are applied to our number counts. To compare with literature results which use PyBDSF, we also apply the 1.42 SExtractor-PyBDSF flux correction to our number counts to make them comparable. Table 2.2 shows our 1.28GHz source counts. The uncertainties in the differential counts are the quadratic sum of the Poisson uncertainties in samples of size N in each bin.

To choose data for comparison, it is important to select data as similar to ours as possible. This makes MIGHTEE data (Jarvis et al. 2016; Heywood et al. 2022) a very good choice, as MIGHTEE uses the same instrument, MeerKAT, to survey the COSMOS and XMM-LSS fields, and its data are reduced in a similar manner using the same software, CASA (McMullin et al. 2007). The official source counts of MIGHTEE are presented in Hale et al. (2023). We also download the MIGHTEE robust 0.0 science images in both COSMOS and XMM-LSS fields, then use SExtractor to extract sources from the two images following the same extraction steps described in Section 2.3.1. We combine the number counts in both COSMOS and XMM-LSS fields, then compare these number counts to our source counts. Figure 2.7 shows our 1.28GHz counts together with other 1.4GHz radio source counts from Hale et al. (2023), and the number counts extracted from MIGHTEE (Heywood et al. 2022) COSMOS and XMM-LSS images following the same extraction processes used in this work. It can be seen that compared to the number counts in Hale et al. (2023), the 1.28GHz source counts of this work have an excess at $S \approx 0.1\ mJy$ in G257. At this flux both the completeness and reliability of SExtractor are $> 95\%$, therefore this excess of source counts is not likely caused by over correction. In addition, the PyBDSF number counts also show the same excess at the same flux of $\sim 0.1\ mJy$. Since

Table 2.2: The 1.28GHz source counts in this work, corresponding to Figure 2.7. For each flux density bin of width 0.2 dex, the bin centre $log_{10}[S]$, raw source counts N in the bin, the integral counts $N_{integral}$ (i.e. cumulative raw source counts) and the uncorrected brightness-weighted counts $S^2n(s)$ in the bin is shown. The brightness-weighted counts corrected for completeness, reliability and PyBDSF-SExtractor flux ratio $S^2n(s)_{corrected}$ is also displayed.

$log_{10}[S/Jy]$	N	$N_{integral}$	$S^2n(s)/Jysr^{-1}$	$S^2n(s)_{corrected}/Jysr^{-1}$
-4.65	7031	7031	381.9 ± 4.6	575.7 ± 5.6
-4.45	10136	17167	872.6 ± 8.7	853.2 ± 8.6
-4.25	8150	25317	1112.0 ± 12.3	964.3 ± 11.5
-4.05	5196	30513	1123.6 ± 15.6	908.3 ± 14.0
-3.85	2716	33229	930.8 ± 17.9	699.6 ± 15.5
-3.65	1220	34449	662.7 ± 19.0	509.7 ± 16.6
-3.45	584	35033	502.8 ± 20.8	382.9 ± 18.2
-3.25	303	35336	413.4 ± 23.8	325.8 ± 21.1
-3.05	163	35499	352.5 ± 27.6	280.3 ± 24.6
-2.85	111	35610	380.4 ± 36.1	331.5 ± 33.7
-2.65	62	35672	336.8 ± 42.8	335.8 ± 42.7
-2.45	52	35724	447.7 ± 62.1	352.0 ± 55.0
-2.25	33	35757	450.3 ± 78.4	353.8 ± 69.5
-2.05	25	35782	540.6 ± 108.1	496.0 ± 103.6
-1.85	25	35807	856.8 ± 171.4	820.2 ± 167.7
-1.65	16	35823	869.1 ± 217.3	758.3 ± 203.0
-1.45	9	35832	774.8 ± 258.3	600.9 ± 227.5
-1.25	4	35836	545.8 ± 272.9	1088.5 ± 385.4
-1.05	6	35842	1297.5 ± 529.7	215.6 ± 215.9
-0.85	2	35844	685.4 ± 484.7	1025.3 ± 592.8
-0.65	2	35846	1086.4 ± 768.2	541.7 ± 542.4
-0.45	0	35846	0.0 ± 0.0	0.0 ± 0.0
-0.25	1	35847	1395.3 ± 1395.3	1360.6 ± 1362.5

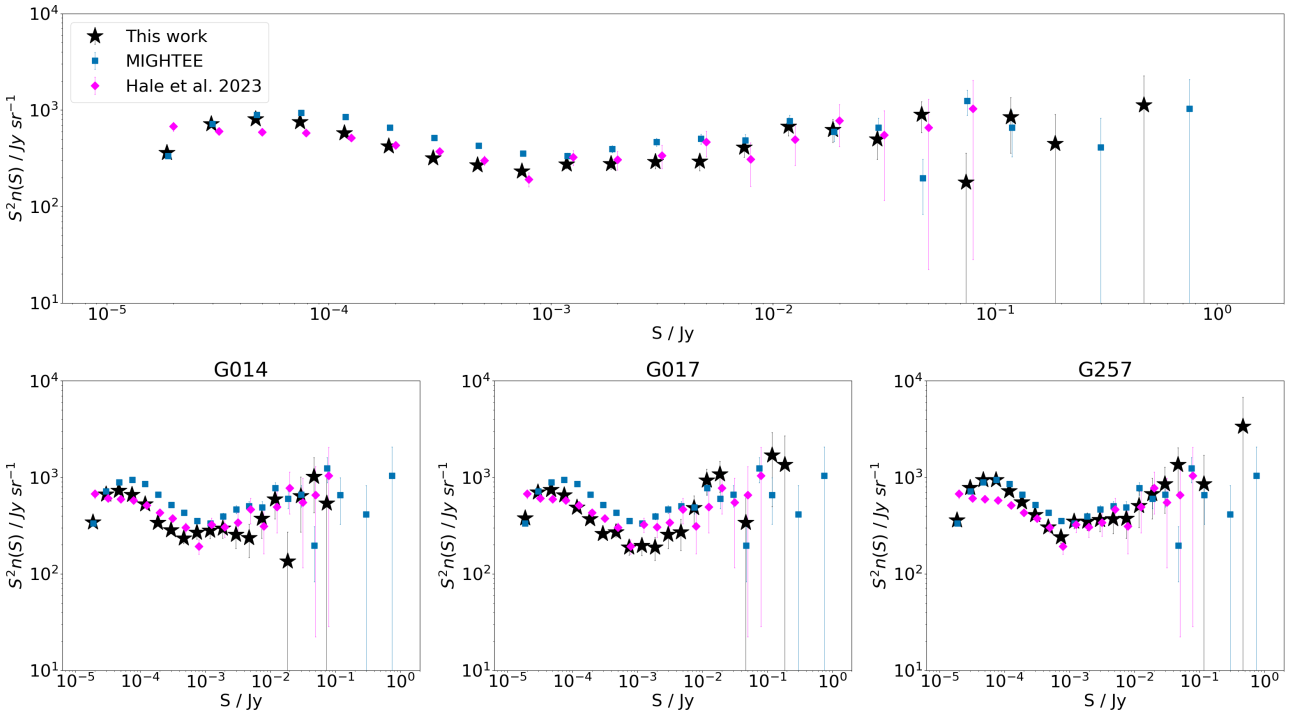


Figure 2.7: Brightness-weighted differential source counts as a function of flux. The black stars are our MeerKAT 1.28GHz counts, the blue squares are number counts extracted from MIGHTEE (Heywood et al. 2022) COSMOS and XMM-LSS images following our extraction processes and the magenta diamonds are the source counts in Hale et al. (2023). The top panel shows our total counts by integrating counts from the 3 maps, while the bottom 3 panels show the source counts from individual maps, G014, G017 and G257 from left to right. Comparing our number counts to those in Hale et al. (2023), we can see an excess at the flux of $\sim 0.1\ mJy$ in G257, suggesting the possibility of real overdensities associated with the candidate protoclusters. However, if we compare our number counts to the counts extracted from MIGHTEE data using the same extraction procedures as us, the excess at $S \approx 0.1\ mJY$ now disappears, showing that it is probably a systematic effect caused by different data processing methods.

our images are centred at 3 DSFG protocluster candidates, this excess in number counts suggest the possibility of a real overdensity in G257. However, comparing to the re-extracted MIGHTEE counts, this excess is no longer seen. We additionally compare the raw number counts per steradian in G257 to the raw counts of Hale et al. (2023) and MIGHTEE re-extraction, as the raw number counts are less sensitive to uncertainties. Figure 2.8 shows the results. It can be seen that in this case the excess at $S \approx 0.1\ mJy$ seen in our G257 brightness-weighted differential number counts compared to that in Hale et al. (2023) is hardly visible. This means that this excess is probably a result of the differences in data processing such as extraction completeness or extraction parameters used, instead of being evidence of real overdensity.

We then compare the source counts in the central $4.63'$ *Planck* beam, where the candidate protoclusters

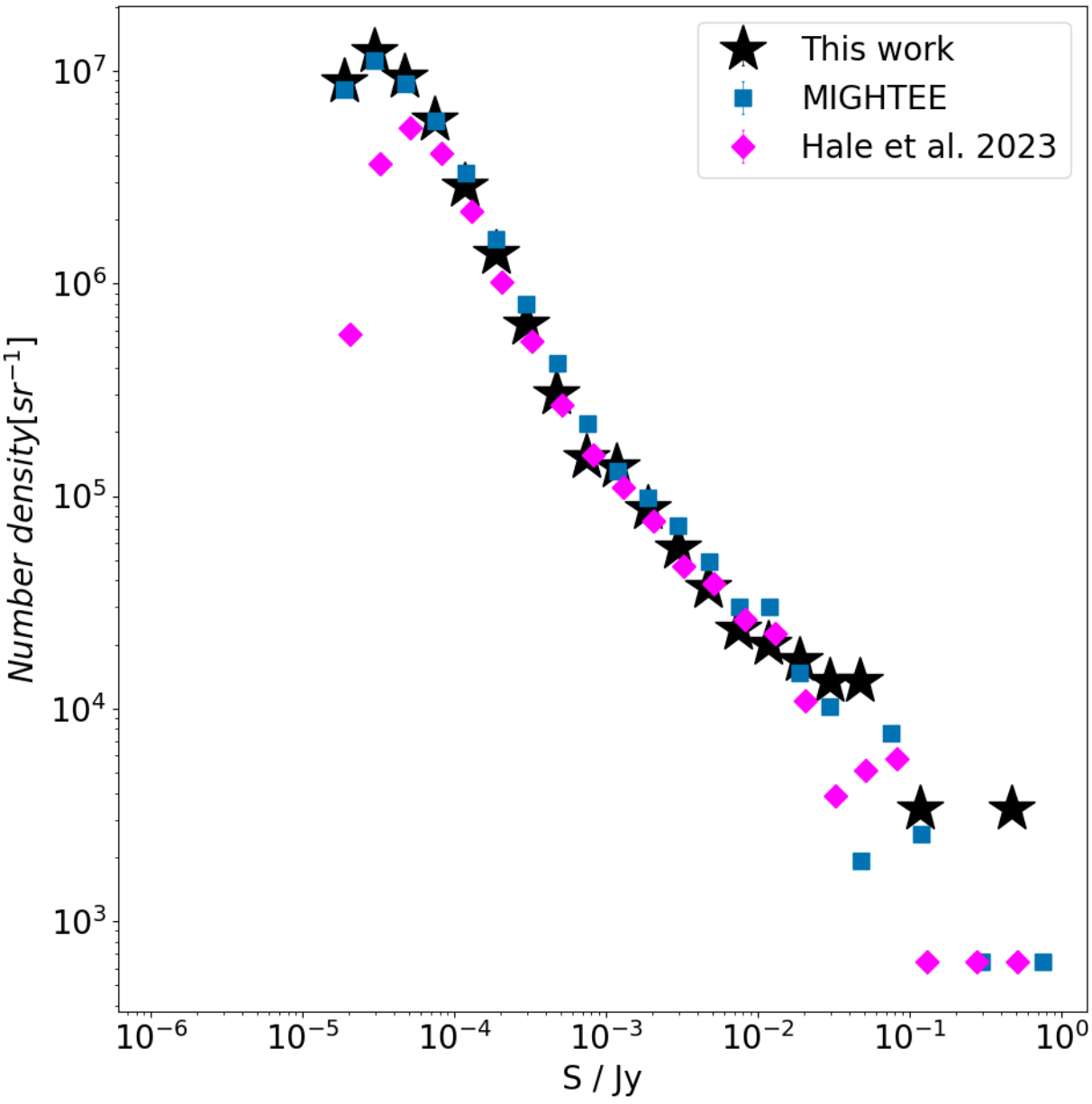


Figure 2.8: Number of sources in G257 per steradian as a function of flux. The black stars are our MeerKAT 1.28GHz counts in G257, the blue squares are number counts extracted from MIGHTEE (Heywood et al. 2022) COSMOS and XMM-LSS images following our extraction processes and the magenta diamonds are the source counts in Hale et al. (2023). The error bars are too small to be seen. The excess at $S \approx 0.1\ mJy$ seen in Figure 2.7 is no longer visible in this case, proving it just being a result of the differences in data processing.

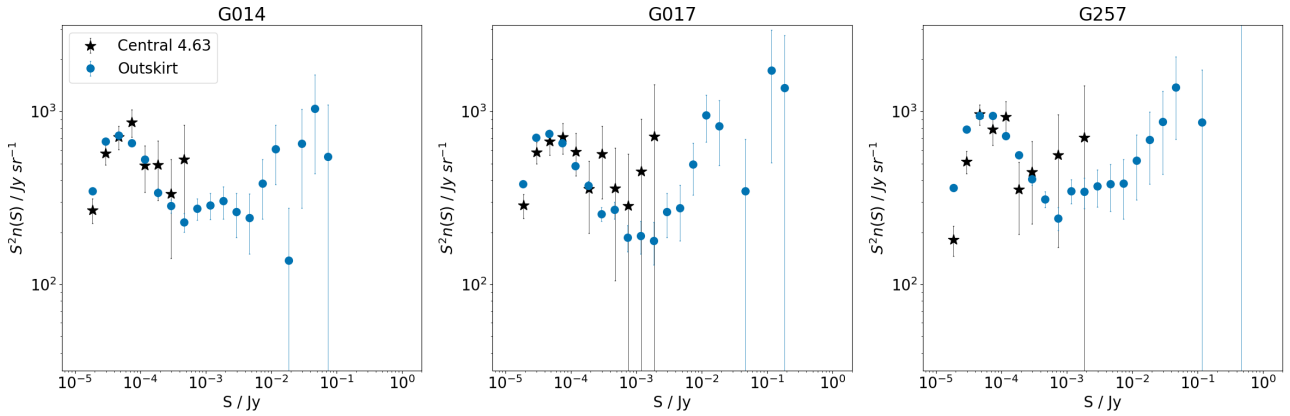


Figure 2.9: Brightness-weighted differential 1.28GHz source counts as a function of flux. The black stars are source counts in the central 4.63′ region, while the blue circles are counts outside the central region.

are located, to the source counts outside the *Planck* beam which gives the local field counts. Figure 2.9 shows the resulting counts. There is no excess seen at the faint flux end. At higher fluxes the number count errors become very large due to the very small number of bright sources in the field centres. By calculating the density contrast as follows:

$$\delta = \frac{n_{centre}}{n_{field}} - 1 \tag{2.4}$$

where n_{centre} is the number density in the centre and n_{field} is the number density in the entire field, we have $\delta = -0.08 \pm 0.07$, $\delta = -0.14 \pm 0.06$ and $\delta = -0.25 \pm 0.05$ for G014, G017 and G257 respectively. This means that there is no radio source overdensity in the central regions compared to other areas.

In summary, the source counts of the MeerKAT images presented in this work are consistent with other radio counts at similar frequencies. There is no direct sign of overdensity in our source counts.

2.4 *Herschel* Identifications

Due to the finite size of the telescope, *Herschel* has large beam sizes ($\geq 18''$). It is therefore hard to obtain accurate positions of *Herschel* sources without the help of high-resolution observations at other wavelengths. However, FIR-bright sources are generally heavily dust-obscured objects which are difficult to observe with high-resolution optical/UV instruments as dust strongly absorbs at these

wavelengths. In this scenario, high-resolution radio observations become an appropriate solution, as: (1) dust is transparent at radio wavelength; (2) radio interferometers such as MeerKAT usually have good angular resolution; (3) there is a known correlation between star-forming galaxies' FIR emission and their radio emission (e.g. De Jong et al. 1985, see Section 2.4.2 for more discussion), which will help rule out random radio sources overlapping with *Herschel* sources by chance. With more accurate positions of *Herschel* sources determined by radio observations, it is much easier to find their cross-identifications in other wavelengths and hence further multi-wavelength studies.

One of the most important factors in the identification is the matching radius applied. A matching radius too small will result in missing many true counterparts, while a radius too large will include irrelevant background sources. In both cases the reliability of the cross-identification will be reduced. To determine the best matching radius, we investigate the number of *Herschel* sources which have 1.28GHz counterparts as a function of matching radius. The results are illustrated in Figure 2.10. Also plotted in Figure 2.10 is the number of *Herschel* sources with no counterpart. Both numbers have a sharp gradient at matching radius $\leq\sim 7.5''$ that levels out at a much steadier rate. This indicates that using matching radii smaller than $7.5''$ will miss a large fraction of true cross-identifications. This is further supported by a sky simulation reproducing the number counts shown in Section 2.5. The simulated catalogue has the same flux density distribution as the real catalogue, but the source positions are completely random (see Section 2.5 for a similar simulation with more details). The number of *Herschel* sources matched to the simulated sources as a function of searching radius is also shown in Figure 2.10. This shows that the number of random identifications starts to rise rapidly beyond searching radius $\sim 7''$. Therefore we select $7.5''$ as our final matching radius and carry out cross-identification.

There are 1914 *Herschel* sources in total which are inside our MeerKAT observed fields, 1817 of them have 1.28GHz counterparts, resulting in a matching percentage of $\sim 94.9\%$. The number of matched sources in G014, G017 and G257 is 419, 599 and 799 respectively. Figure 2.5c shows an example of a cross-matched source. It should be noted that only 18.3% of our matched *Herschel* sources have $> 3\sigma$ detections in all 3 SPIRE bands, with 99.2%, 60.8% and 19.5% of the *Herschel* sources having $> 3\sigma$ detections in 250, 350 and $500\mu m$ band respectively.

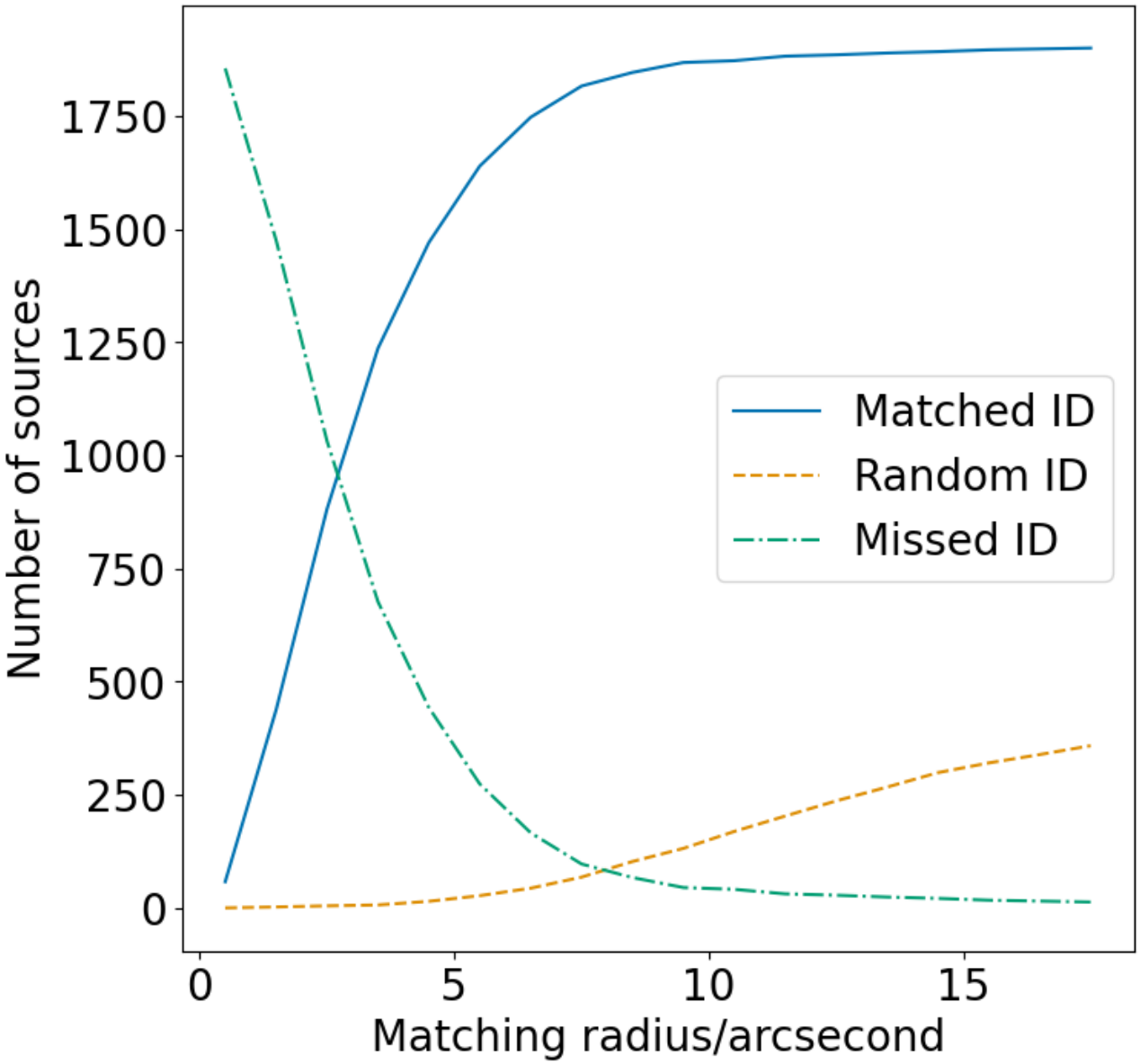


Figure 2.10: Number of *Herschel* sources with real MeerKAT counterparts (blue solid line) / missed MeerKAT counterparts (green dashed-dotted line) / random identification (yellow dashed line) as a function of matching radius. At $\sim 7.5''$ the gradient of the increase in number of matched sources significantly reduces while the number of random IDs starts to rise, this means that using matching radii smaller than $7.5''$ will miss a large part of true cross-identifications, while using matching radii larger than $7.5''$ will include more background sources aligned by chance. Using a matching radius of $7.5''$ will maximise the number of true IDs while minimising the number of missed IDs and random IDs.

Visual inspection of the *Herschel* sources without MeerKAT counterparts reveals that some of them do overlap with 1.28GHz objects, but their 1.28GHz counterparts are either close to other bright sources, and hence ignored or blended with nearby bright sources by SExtractor, or they have extended morphologies and the separations between the central positions assigned by SExtractor and *Herschel* are larger than the matching radius. Figure 2.5d shows an example.

To further evaluate the robustness of our 1.28GHz identifications of *Herschel* sources, we estimate the Poisson probability of an identification being a chance alignment (e.g., Downes et al. 1986, Biggs et al. 2011). Assuming random distribution of 1.28GHz sources, the expected number of sources within a radius r is:

$$\mu = \pi r^2 n \tag{2.5}$$

where n is the blank field number density of sources brighter than S. Then the Poisson probability of one random object within r is:

$$P = 1 - e^{-\mu}. \tag{2.6}$$

Using the number density brighter than $S = 0.021\ mJy$ from Hale et al. (2023) and our search radius of $7.5''$ for r, the probability that one 1.28GHz source lines up with a *Herschel* source by coincidence is 7.33%, in other words the probability of it being a good identification is $(1 - P) = 92.67\%$. Given that we have a a matching percentage of $\sim$ 94.9% with some genuine cross-IDs not counted due to their extended morphologies, this leads to the conclusion that $\gtrapprox$ 90% of *Herschel* sources will have one or more true 1.28GHz counterparts at $S > 0.021\ mJy$.

2.4.1 Multiplicity of *Herschel* Sources

Since the *Herschel* SPIRE beam sizes are much larger than the beam size of our MeerKAT images, there is a chance that a *Herschel* source has multiple corresponding MeerKAT sources. It is known that most single dish DSFGs are blends of multiple submillimeter sources (e.g., Ivison et al. 2007, Wang et al. 2010, Karim et al. 2013, Stach et al. 2018, Greenslade et al. 2020). One of the most popular hypotheses to explain this multiplicity of DSFGs is that most DSFGs are ongoing/late-stage mergers (e.g. Hopkins et al. 2008, Wardlow et al. 2018), but whether or not the multiple components

of any given DSFG are truly physically associated is unclear. CO observations of 12 submillimetre galaxies (SMGs) (Tacconi et al. 2008; Engel et al. 2010) suggests that all the galaxies in the sample are either ongoing mergers or compact dense galaxies which are considered as late stage mergers, meaning that SMGs are likely to be high-redshift analogues to local (U)LIRGs. To the contrary, models from Hayward et al. (2013) and Cowley et al. (2015) indicate that multiple sources resolved from single dish SMGs are generally not associated, hence are not merger companions. In conclusion, for now it is only clear that most single dish DSFGs (or more broadly speaking SMGs) are blends of multiple sources. Here, we utilise our 1.28GHz data to investigate whether the multiplicity of DSFGs can still be seen at this wavelength, and if it can provide any insight into the nature of DSFGs.

We re-match our MeerKAT sources to the H-ATLAS DR2 catalogue, using a matching radius of 17.6″ which is the primary beam size of the *Herschel* SPIRE 250μm channel. This matching radius is used to search for all potential MeerKAT counterparts of *Herschel* sources, especially for unresolved sources. However, using a matching radius as large as 17.6″ will boost the risk of random 1.28GHz sources being mis-identified to 34.2%. We will call this new catalogue the multiplicity catalogue for convenience in the following sections. Figure 2.5e&f illustrates some *Herschel* sources matched with multiple MeerKAT counterparts. We find that 1885 *Herschel* sources have MeerKAT counterparts at this matching radius and 1038 out of the 1885 *Herschel* sources have 2 or more radio counterparts. This results in 54.2% of *Herschel* sources with multiple 1.28GHz IDs. The majority of the 1038 *Herschel* matched sources have $2-4$ components with 698, 270 and 59 of them having two, three and four components respectively, while the maximum number of counterparts of a single *Herschel* source is 7, as shown by Figure 2.11. Visual examination of *Herschel* sources with more than 4 MeerKAT counterparts reveals that all of them are actually two or more close *Herschel* sources acting as one large object during cross-matching, leading to their apparent multiplicities larger than 4. The actual number of counterparts of these sources will be 1-4. An example of this is shown in Figure 2.5g.

To further investigate the nature of *Herschel* sources with 1.28GHz multiplicity, we carry out two-sample Kolmogorov–Smirnov tests (hereafter, K-S tests) to examine whether or not the flux density distribution of *Herschel* sources with multiple 1.28GHz components is intrinsically different to sources without such multiplicity. The two-sample K-S test is used to test whether two underlying one-

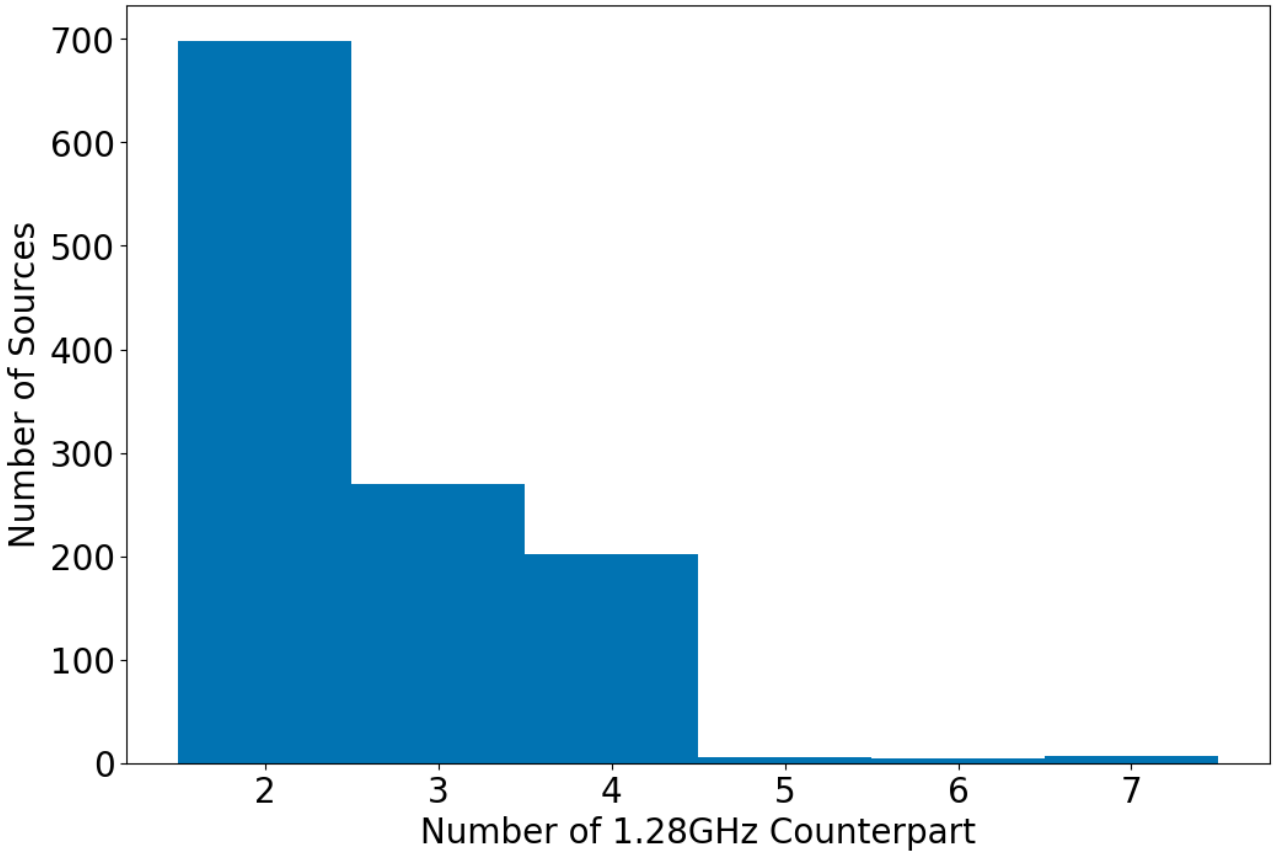


Figure 2.11: The distribution of the number of MeerKAT matched to *Herschel* sources. Most *Herschel* matched sources have 2 – 4 MeerKAT counterparts, while the highest number of radio components of a single source is 7. Visual inspection suggests all *Herschel* sources with more than 4 radio components are in fact several *Herschel* sources lying close to each other. The true number of MeerKAT counterparts should be 1-4, see Figure 2.5g

dimensional probability distributions differ. If there is a true difference between the flux density distributions of the two populations, it implies that multiple 1.28GHz components of a *Herschel* sources are likely to be associated instead of being mere chance-alignments. For this analysis, *Herschel* detections below 3σ are removed for each SPIRE band.

Figure 2.12 illustrates the flux density distributions of *Herschel* sources with/without 1.28GHz MeerKAT multiplicity, with the associated p-values of the K-S tests shown in the corresponding panels. Statistically, we can be over 99.9% confident that *Herschel* sources with/without MeerKAT multiplicity have different flux distributions at $250\mu m$, but in the other two SPIRE bands the confidence level that the two populations have different flux distributions drops to $\sim$ 92.3% and $\sim$ 33.4%. In other words, the two-sample K-S test results of our sample suggest that the difference between the flux density distributions of *Herschel* sources with and without multiple 1.28GHz counterparts is only significant at $250\mu m$. The average $250\mu m$ fluxes and standard deviations of *Herschel* sources with/without MeerKAT multiplicity are 42.2 ± 16.1 and 47.5 ± 23.1 respectively. One possible explanation of why the $250\mu m$ distributions of the two populations are different is that there are too few *Herschel* sources without multiplicity detected at 350 and $500\mu m$, which can also be seen from Figure 2.12. Since the $250\mu m$ band has the best sensitivity and resolution of all 3 SPIRE bands while the $500\mu m$ band has the worst, there will be more faint sources without multiple components detected in the $250\mu m$ band compared to the other two bands. In the other two bands, bright sources which are more likely to be blends of multiple fainter sources will be dominant and faint sources with no multiplicity will be too few to reflect their genuine flux distribution. This also explains the increasing trend of p-value from 250 to $500\mu m$, as the sensitivity of the band decreases. Another possible explanation is that, since we are using a larger matching radius here, hence higher percentage of chance alignments, and there are far more faint $250\mu m$ sources than bright ones, there will be more chances for MeerKAT sources to randomly match a faint rather than bright SPIRE sources, leading to a fainter $250\mu m$ flux distribution of sources with multiple MeerKAT IDs.

Overall, our results indicate that *Herschel* sources with multiple 1.28GHz MeerKAT components have a different $250\mu m$ flux density distribution to *Herschel* sources with just one MeerKAT counterpart, while at 350 and $500\mu m$ there are insufficient *Herschel* sources to make a statistically significant

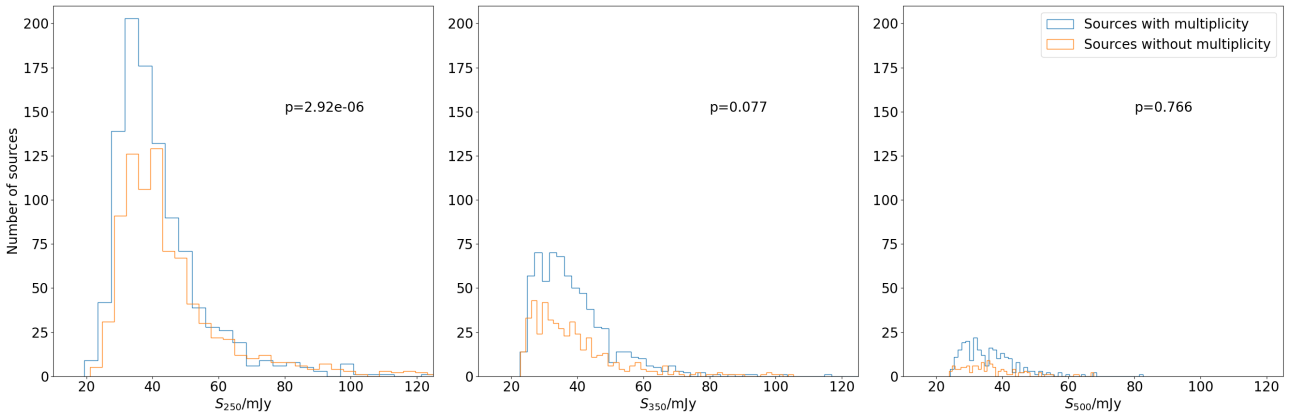


Figure 2.12: Flux density distributions of *Herschel* sources with/without 1.28GHz MeerKAT counterparts. From left to right are the distributions in 250, 350 and 500μm. The resulting p-values of the two-sample K-S tests are shown. For a confidence level above 95%, only the 250μm distributions of the two populations are statistically different.

comparison. This suggests further evidence that the multiple 1.28GHz counterparts of a *Herschel* source are associated instead of being a chance-alignment. However, this could also just be a bias due to lack of data or inaccurate determinations of multiple radio counterparts to *Herschel* sources at large matching radius. We need wider area radio surveys to cover more 350 and 500μm sources and other techniques to achieve better cross-IDs at larger searching radius to draw further conclusions.

2.4.2 The Far-IR Radio Correlation

A surprisingly tight correlation between galaxies' FIR emission and their radio emission was discovered in the late 1970s, first proposed by Van der Kruit (1971). The FIR radiation is believed to come from thermal re-radiation of dust-enshrouded regions while the radio radiation is thought to be synchrotron emission from cosmic ray electrons accelerated in supernova remnants. This naturally explains the observed correlation as the massive stars that eventually end as supernovae are those stars which substantially heat up their surrounding gas and dust (Harwit and Pacini 1975). Numerous studies have been carried out on the Far-Infrared Radio Correlation (FIRC; e.g., Helou et al. 1985; De Jong et al. 1985; Yun and Carilli 2002; Jarvis et al. 2010; Molnár et al. 2021).

The FIRC has many useful applications. It can be used to estimate L_{IR} with radio data when no IR data are available, or it can be used to complement or constrain the FIR SEDs when FIR data are inadequate (e.g., Yun and Carilli 2002). Likewise, the FIRC can also be used to study the relation

between obscured star formation and galaxies' radio emission, as $SFR \propto L_{IR} \propto L_{1.4GHz}$.

Following Ivison et al. (2010a), we calculate the monochromatic $q_{250\mu m}$ to investigate the FIRC as:

$$q_{250\mu m} = log_{10}\frac{S_{250\mu m}}{S_{1.28GHz}}$$

where $S_{250\mu m}$ is the $250\mu m$ flux density of a source and $S_{1.28GHz}$ is its 1.28GHz flux density. Using the *Herschel*-MeerKAT matched catalogue with 7.5″ matching radius, the $q_{250\mu m}$ as a function of $250\mu m$ flux is shown in Figure 2.13. It should be noted that the $q_{250\mu m}$ calculated here is derived from observed frame quantities due to the lack of accurate redshift information for most sources. Therefore our results are expected to have a redshift dependency which will not be seen in samples with K-corrections applied (e.g., Ivison et al. 2010a). The mean value of our $q_{250\mu m}$ is 2.33 ± 0.26, which is consistent with the mean value of 2.26 ± 0.35 in Ivison et al. (2010a). This further indicates that our *Herschel*-MeerKAT cross-identifications are of good reliability. Moreover, using the EAZY photometric redshifts estimated in Section 2.6.1, we divide our sample into two subsamples. One of the two subsamples includes sources with EAZY photometric redshifts larger than one or with no photometric redshift assigned, and the other includes sources with EAZY redshift lower than one. The mean $q_{250\mu m}$ are 2.34 ± 0.28 and 2.29 ± 0.22 for $z_{EAZY} > 1$ and $z_{EAZY} < 1$ sources respectively. There is no significant difference between the two subsamples. In Figure 2.14, we additionally compare the $q_{250\mu m}$ of all sources within the 4.63′ radius circles in the field centres to the other sources, as they are potential members of the candidate protoclusters. The average $q_{250\mu m}$ is 2.31 ± 0.21 which is consistent with the overall value. It should be noted that the source with the lowest $q_{250\mu m}$ is in the centre of G017, it is a very bright radio source with some beam artefacts around it, see Figure 2.5h.

We also compare the far-IR radio correlation between *Herschel* sources with and without 1.28GHz multiplicity utilising our multiplicity catalogue. Since we cannot determine the FIR flux contributions of each radio component of a given *Herschel* source, the $q_{250\mu m}$ of a *Herschel* source with multiple counterparts is calculated by integrating the radio flux densities of all its 1.28GHz components. The results are shown in Figure 2.15. The average $q_{250\mu m}$ values are 2.14 ± 0.22 and 2.24 ± 0.34 for *Herschel* sources with and without multiplicity respectively. Again, we perform a K-S test to examine whether there is a statistically significant difference between the two populations. The p-value of this K-S test

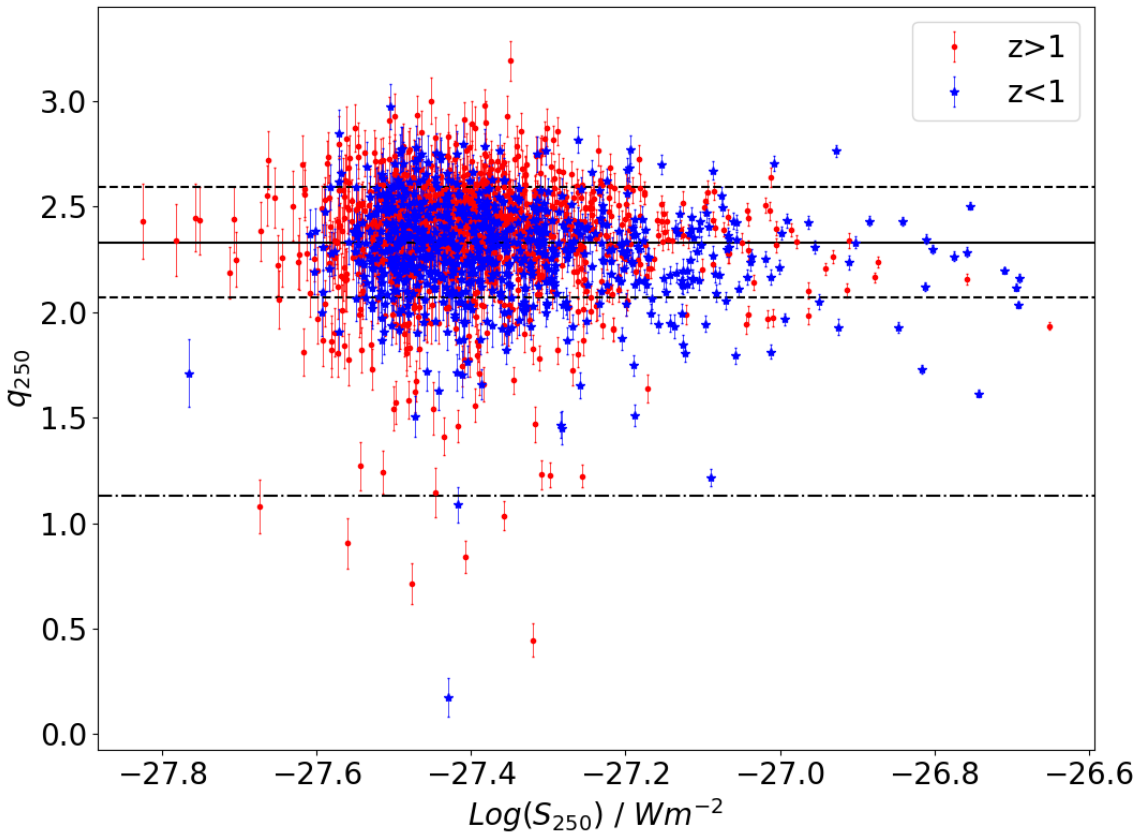


Figure 2.13: $q_{250\mu m}$ against the 250 μm flux density for all sources in the three maps. The red dots are sources with EAZY photometric redshift larger than one or with no photometric redshift assigned, the blue stars are sources with $z_{EAZY} < 1$. The solid line is the mean value of $q_{250\mu m}$ with dashed lines being its 1σ standard deviation. The dotted-dash line is the lower 5σ clipping boundary. The mean $q_{250\mu m}$ value is 2.33 ± 0.26 for all sources, it is comparable to the results in Ivison et al. (2010a) which is $\sim 2.26 \pm 0.35$ with relatively small dispersion.

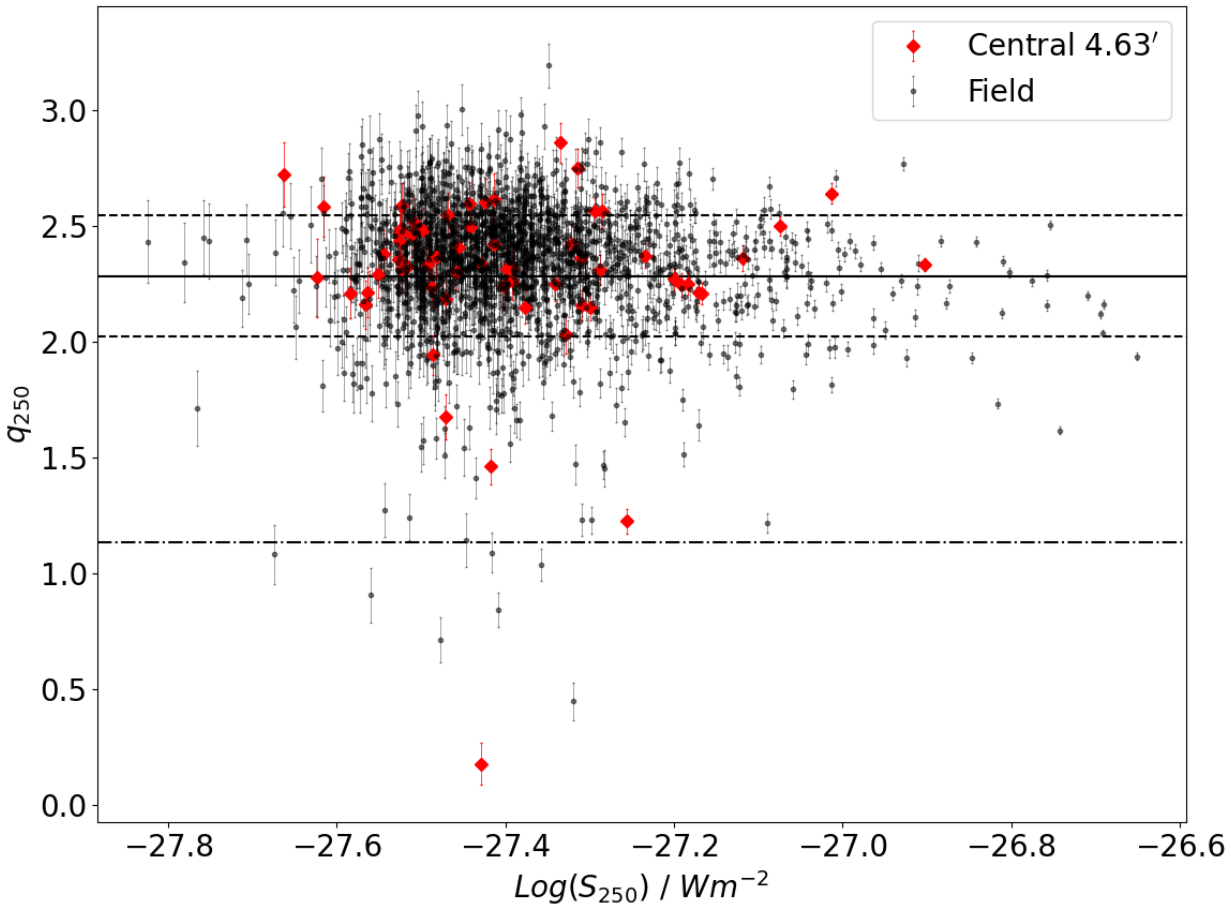


Figure 2.14: $q_{250\mu m}$ against the 250 μm flux density for all sources in the three maps. The red diamonds are sources within the 4.63′ radius circles in the field centres, the black dots are the remaining sources in the fields. The solid line is the mean value of $q_{250\mu m}$ for all sources with dashed lines being its 1σ standard deviation. The dotted-dash line is the lower 5σ clipping boundary. There is no apparent difference between the distributions of the two populations, as the average $q_{250\mu m}$ for central sources 2.31 ± 0.21 is consistent with the overall $q_{250\mu m}$ of 2.33 ± 0.26.

is 2.86×10^{-34} which means it is very likely that the $q_{250\mu m}$ distributions of the two populations are different. *Herschel* sources with 1.28GHz multiplicity show a lower $q_{250\mu m}$ on average with smaller scatter. A possible cause of this difference in $q_{250\mu m}$ is that we derive the $q_{250\mu m}$ of *Herschel* sources with multiple radio components by adding up all the radio counterparts' flux densities, which may lead to higher 1.28GHz flux densities and hence lower $q_{250\mu m}$. As already mentioned in Section 2.4.1, this could also be a consequence of the fact that the probability of random alignments is higher in the multiplicity catalogue and there are more faint $250\mu m$ sources than bright ones. In this case if a considerable fraction of matched sources are just chance alignments and these radio sources are not significant FIR emitters, it would explain the lower $q_{250\mu m}$ of *Herschel* sources with multiple counterparts seen here. Whether this difference is real or just a bias can only be investigated by further studies with accurate determinations of each 1.28GHz components' contributions to the *Herschel* flux densities.

Sources with low $q_{250\mu m}$ have much larger radio fluxes than expected from their FIR fluxes. In other words they are radio-loud and radio-loudness is a typical feature of AGN activity. We flag sources with very low $q_{250\mu m}$ by performing sigma-clipping to the $q_{250\mu m}$ in all catalogues. Sources with $q_{250\mu m}$ 5σ smaller than the mean are excluded, then the average $q_{250\mu m}$ and the standard deviation are recalculated. We iterate this process to the point where no further source is excluded. The resulting $q_{250\mu m}$ for the $7.5''$ catalogue is 2.33 ± 0.24, with 8 sources being flagged out. For the multiplicity catalogue, the updated $q_{250\mu m}$ are 2.15 ± 0.21 and 2.26 ± 0.29 for *Herschel* sources with and without multiplicity with 5 and 8 sources being flagged out respectively. After the sigma-clipping, the average values of $q_{250\mu m}$ for all catalogues do not change significantly, while the 1σ standard deviations become smaller in all cases as an expected consequence of sigma-clipping. In both Figure 2.13 and Figure 2.15 we can see there are sources lying outside the lower 5σ boundaries. Visual examination of these sources finds that some of them are sources with interesting morphologies such as head-tail sources or jets which are more likely to be true identifications with potential AGN activity. Figure 2.5i shows an example with extended radio structure coinciding with a *Herschel* source, suggestive of a possible AGN with radio jets heavily enshrouded by dust. This demonstrates that $q_{250\mu m}$ could be an effective selection criterion for radio-loud AGN searches in future studies. Additionally, an interesting observation is that in Figure 2.13, the majority of radio-loud sources have photometric

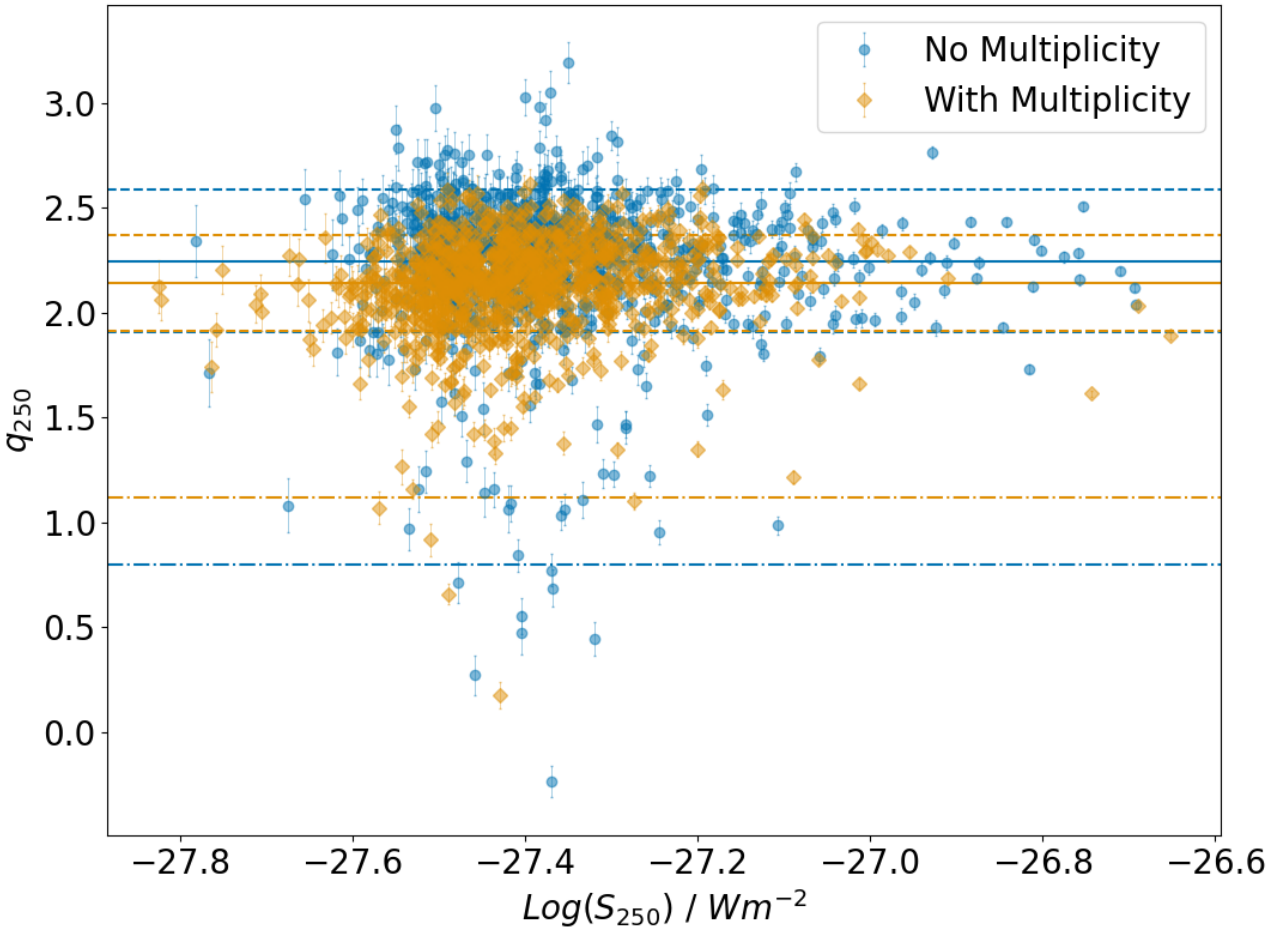


Figure 2.15: $q_{250\mu m}$ against the 250 μm flux density for *Herschel* sources with only one 1.28GHz counterparts (blue dots) and *Herschel* sources with multiple 1.28GHz counterparts (orange diamonds). The solid lines are the mean values of $q_{250\mu m}$ of the two populations, with dashed lines being the corresponding 1σ standard deviations. The dotted-dash lines are the respective lower 5σ clipping boundaries.The average $q_{250\mu m}$ values for the two populations are 2.14 ± 0.22 and 2.24 ± 0.34 for *Herschel* sources with and without multiplicity respectively.

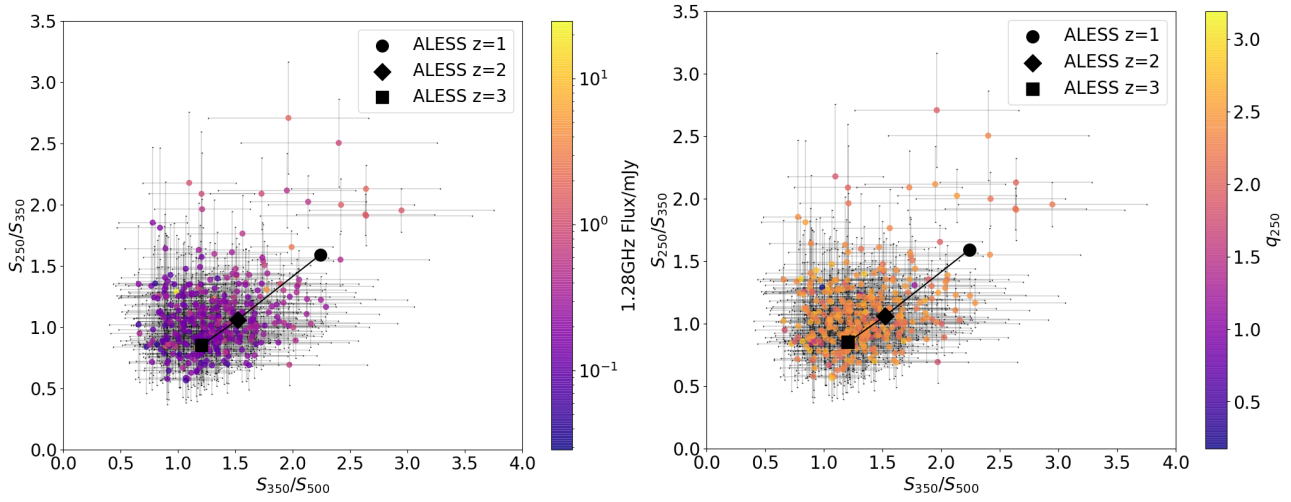
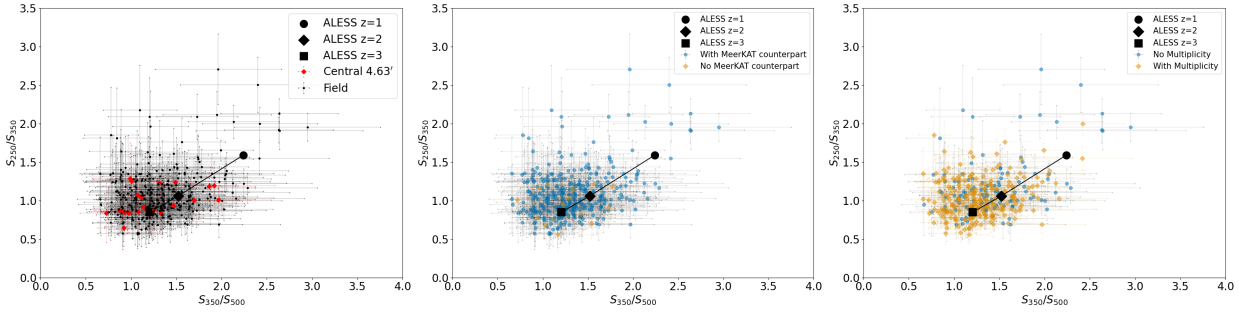


Figure 2.16: *Herschel* S_{250}/S_{350} and S_{350}/S_{500} colours of our MeerKAT sources. The *Herschel* colours of the ALESS template SED (da Cunha et al. 2015) at redshift of 1,2 and 3 are also shown as black dot, diamond and square respectively. Most sources reside in regions indicated as redshift 2-3. (a) *Herschel* colour-colour diagram, colour-coded by their 1.28GHz MeerKAT fluxes. No obvious correlation between the 1.28GHz flux and the *Herschel* colours can be observed. (b) The same diagram, colour-coded by their FIRC factor $q_{250\mu m}$, there is also no apparent correlation between the $q_{250\mu m}$ and the *Herschel* colours.

redshift larger than one or have no optical/NIR IDs to provide the photometry needed for redshift estimations. This further suggests that these sources are heavily dust-enshrouded AGN as they are faint in the optical/NIR.

2.4.3 Herschel Colours

We next examine the *Herschel* SPIRE colours of our sources. These are commonly used as redshift indicators for submillimeter sources (e.g. Clements et al. 2014, Dowell et al. 2014, Ivison et al. 2016, Greenslade et al. 2018). It is widely accepted that sources whose SEDs peak at longer wavelengths tend to lie at higher redshifts (Casey et al. 2014). Therefore sources with SEDs peaking at 250, 350 and 500μm are likely to have progressively higher redshifts. We investigate the *Herschel* SPIRE colours of our MeerKAT sources to obtain a rough estimation of their redshifts, and to see if their *Herschel* colours have any correlation with their 1.28GHz fluxes. Again, *Herschel* detections below 3σ in any of the 3 SPIRE bands are removed.



(a) (b) (c)

Figure 2.17: Similar figure to Figure 2.16, but the sample is divided into (a) Sources within the central 4.63′ *Planck* beam, marked as red diamonds, and the other sources in the fields as black dots; (b) *Herschel* sources with and without MeerKAT counterparts, blue dots and orange diamonds respectively; and (c) *Herschel* sources with and without MeerKAT multiplicity, blue dots and orange diamonds respectively.

In Figure 2.16a we plot S_{250}/S_{350} against S_{350}/S_{500} for our sources, colour-coded by their corresponding 1.28GHz MeerKAT fluxes. The ALESS SED for a template starburst galaxy (da Cunha et al. 2015) is used as a benchmark, and its *Herschel* colours at various redshifts are displayed in the figure. As can be seen, most sources lie in regions suggesting redshifts between 2 and 3, while there is no obvious correlation between the 1.28GHz flux and *Herschel* colours. Figure 2.16b shows a similar diagram, but colour-coded with the monochromatic FIRC coefficient $q_{250\mu m}$. Again, there is no apparent correlation between the $q_{250\mu m}$ and the *Herschel* colours.

In addition, we differentiate our sample into two sub-samples by: (a) sources inside/outside the 4.63′ *Planck* beam at the field centres; (b) whether the *Herschel* sources have MeerKAT counterparts; (c) whether the *Herschel* sources have multiple or single MeerKAT counterparts. The *Herschel* colours of these sub-samples are shown in Figure 2.17. Note that we use the multiplicity catalogue in Figure 2.17c.

The average SPIRE colours of the various subsamples in Figure 2.17 are as follows: (a) the mean S_{250}/S_{350} and S_{350}/S_{500} are 1.17 ± 0.09 and 1.34 ± 0.18 for sources in the centre, while the mean S_{250}/S_{350} and S_{350}/S_{500} are 1.26 ± 0.02 and 1.37 ± 0.04 for the other sources in the field. Most central sources locate at $z \approx 2-3$ region with a slightly lower S_{250}/S_{350}; (b) the mean S_{250}/S_{350} and S_{350}/S_{500} are 1.27 ± 0.01 and 1.33 ± 0.02 for sources with MeerKAT ID and 1.10 ± 0.05 and 1.23 ± 0.10 for sources with no MeerKAT ID. Both colours for sources with no MeerKAT ID are lower, the difference

in S_{250}/S_{350} is over 3σ; (c) the meanS_{250}/S_{350} and S_{350}/S_{500} are 1.40 ± 0.02 and 1.44 ± 0.04 for sources with single MeerKAT ID and 1.17 ± 0.01 and 1.28 ± 0.02 for sources with multiple MeerKAT IDs. In this case both colours for sources with multiple MeerKAT IDs are significantly lower.

We further apply K-S tests to examine if there is a statistical difference between these sub-samples. The p-values of the K-S tests between *Herschel* sources inside the $4.63'$ *Planck* beam in field centres and the others in the fields are 0.238 and 0.995 for S_{250}/S_{350} and S_{350}/S_{500} respectively, meaning that there is likely no difference between these two sub-samples. Similarly for sources with and without MeerKAT ID, the p-values are 0.556 and 0.999 for S_{250}/S_{350} and S_{350}/S_{500} respectively, with no statistical difference between the SPIRE colours of the two population. Meanwhile, the p-values of the K-S tests between *Herschel* sources with MeerKAT multiplicity and those without multiplicity are 4.27×10^{-14} and 9.67×10^{-5} for S_{250}/S_{350} and S_{350}/S_{500} respectively, indicating a difference between these two populations. However, this is expected since the *Herschel* flux density distributions between these two populations are different as discussed in Section 2.4.1. Therefore as the difference in *Herschel* colour distributions is just a direct result of the different flux density distributions, no further conclusion can be drawn.

Despite there being no strong correlation between *Herschel* colours and MeerKAT fluxes or $q_{250\mu m}$, the *Herschel* colours suggest that most of our sources have redshift larger than 1 which is as expected from Greenslade et al. (2018). Nevertheless, estimating the redshift of sources from their *Herschel* colours can only provide rough indications with large uncertainties. Later in Section 2.6.1, we carry out photometric redshift estimations using accurate optical to IR cross-IDs, to further investigate the properties of our sample.

2.5 Using the FIRC for *Herschel* cross-IDs

We have performed cross-identifications of *Herschel* sources with our MeerKAT 1.28GHz catalogue as presented in Section 2.4. However, we only use the Poisson probability of chance alignment as the measure of the cross-identification's robustness. There is still the possibility of contamination in our cross-IDs due to random 1.28GHz sources aligned with *Herschel* objects by chance, especially for the

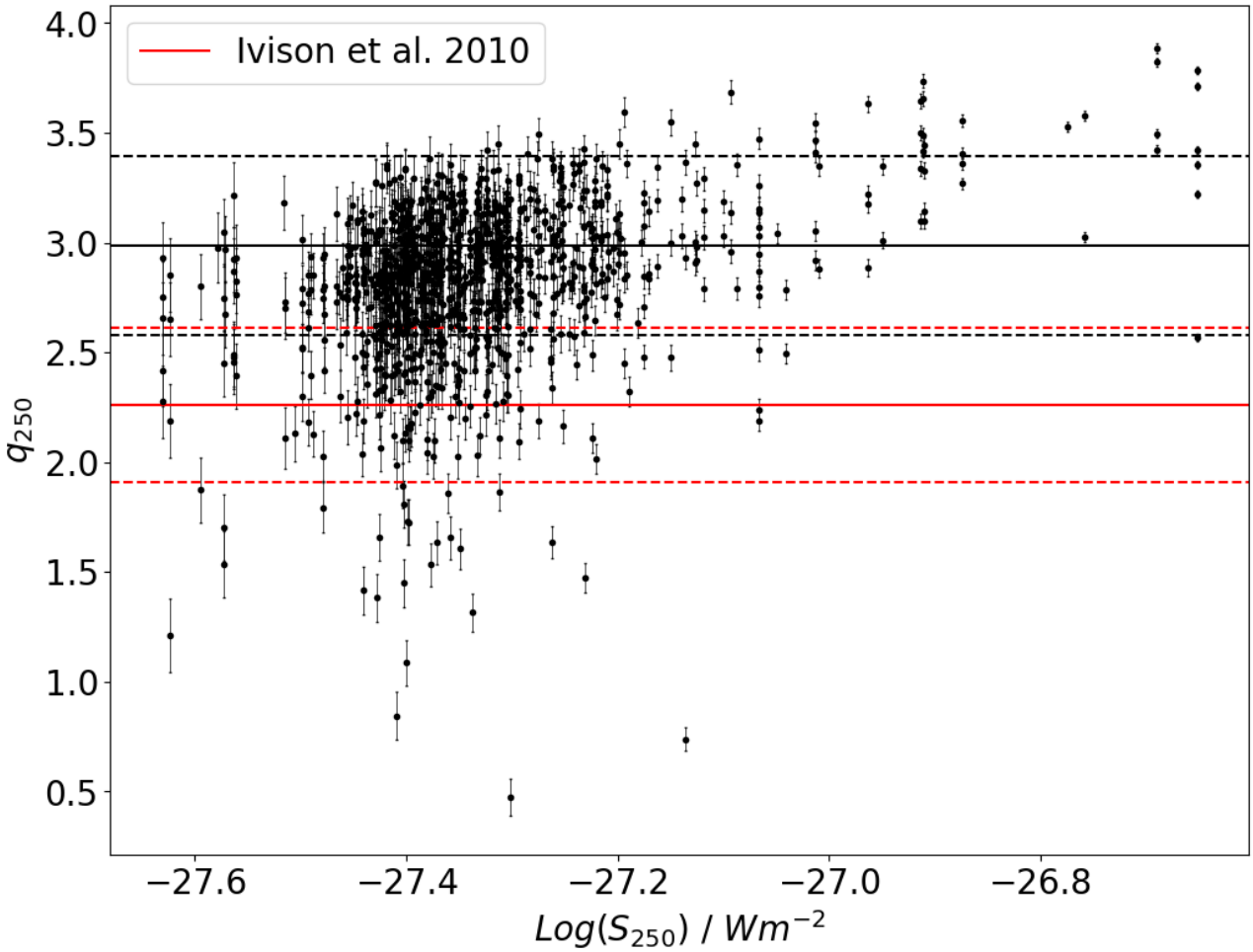


Figure 2.18: The $q_{250\mu m}$ against the $250\mu m$ flux density for *Herschel* sources matched to simulated MeerKAT sources, with a matching radius of $7.5''$. The average $q_{250\mu m}$ of these sources is 2.99 ± 0.41, the black solid line shows the mean $q_{250\mu m}$ of these sources and the dashed lines show their 1σ scatter. The mean and 1σ scatter of $q_{250\mu m}$ in Ivison et al. (2010a) are shown as red solid line and dashed lines.

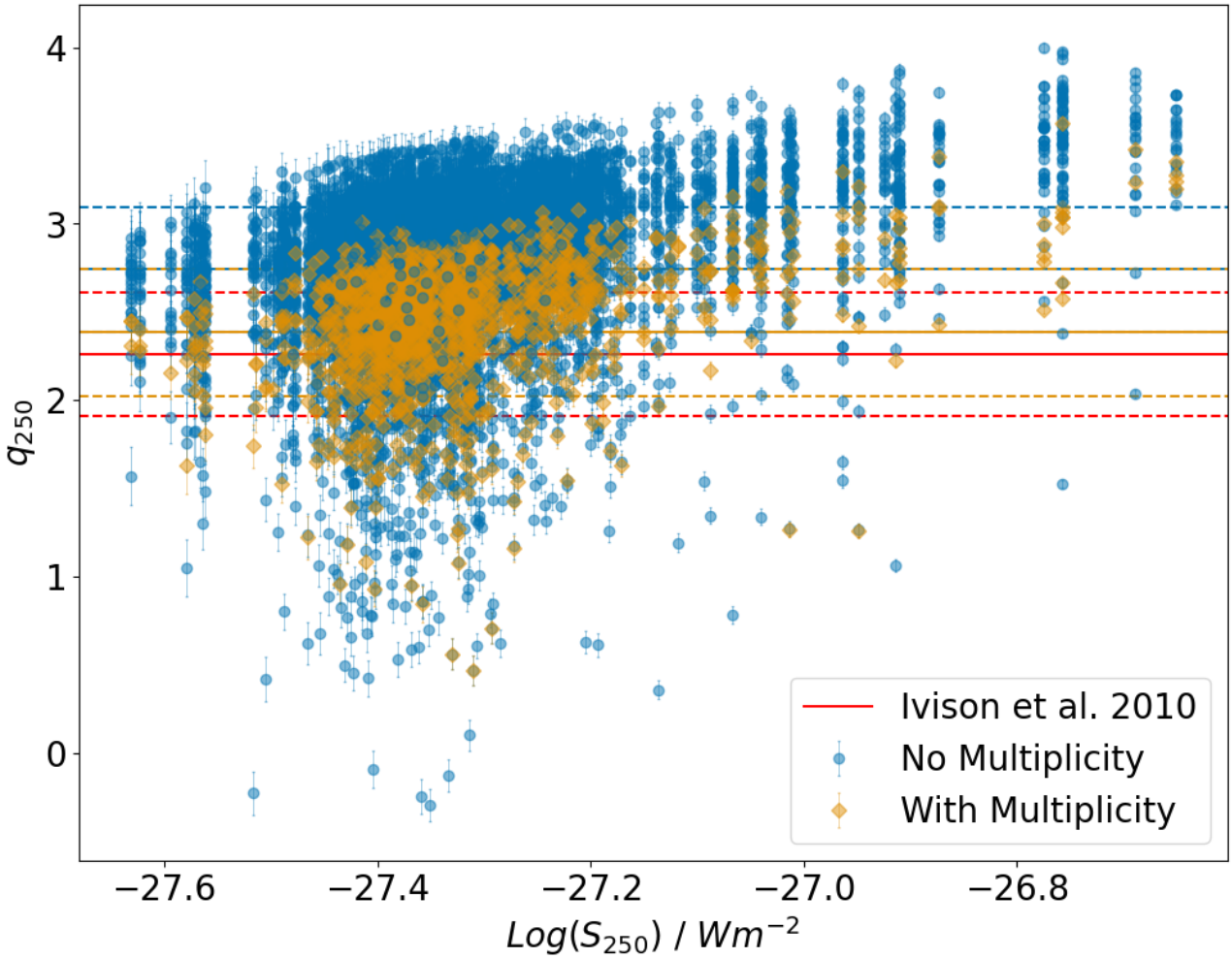


Figure 2.19: The $q_{250\mu m}$ against the $250\mu m$ flux density for *Herschel* sources matched to simulated MeerKAT sources, with a matching radius of 17.6″. The average $q_{250\mu m}$ for sources with multiple counterparts (orange diamonds) and without multiple counterparts (blue dots) are 2.38 ± 0.36 and 2.74 ± 0.35 respectively. The solid line shows the mean value and the dashed lines show the 1σ scatter. The mean and 1σ scatter of $q_{250\mu m}$ in Ivison et al. (2010a) are shown as red solid line and dashed lines.

cross-IDs used for the multiplicity study where a larger matching radius of 17.6″ is used. Here, we explore the feasibility of using the FIRC to improve the quality of our cross-IDs without introducing other more complicated methods.

We create a simulated catalogue in the G014 map, reproducing the Hale et al. (2023) blank field number counts in flux density bins ranging from 0.016 mJy to 630 mJy. In each flux bin, the flux density of a simulated source is randomly assigned from the flux density range of this bin. To eliminate the possibility of simulated sources coinciding with real MeerKAT sources, all simulated sources are at least 6″ (roughly the synthesised beam size of MeerKAT) away from any real sources. The simulated catalogue is then matched to the H-ATLAS catalogue with matching radius of 7.5″ and 17.6″, corresponding to our cross-ID catalogue and multiplicity catalogue respectively. Since the simulated sources are randomly located on the sky, any source that matches a *Herschel* source should be a purely chance alignment, and the FIRC coefficient $q_{250\mu m}$ of it should be very different to the Ivison et al. (2010a) result. The same procedure is then carried out 100 times to ensure robust statistics.

There are 438 *Herschel* sources inside the G014 map, on average 11 ± 3 of them have counterparts within the 7.5″ search radius on average, leading to a matching percentage of 2.60 ± 0.70%. The $q_{250\mu m}$ of those *Herschel* sources with identifications in the simulated catalogue are calculated and displayed in Figure 2.18. The average $q_{250\mu m}$ of these sources is 2.99 ± 0.41. The mean $q_{250\mu m}$ of these chance coincidences is higher than both 2.26 ± 0.35 in Ivison et al. (2010a) and 2.33 ± 0.26 from our MeerKAT images, but only 18.0 ± 12.0% of these source have $q_{250\mu m}$ values inside the 1σ scatter of the $q_{250\mu m}$ in Ivison et al. (2010a). Since all these sources are randomly distributed on the sky and should not have any physical association with real *Herschel* sources, this much higher $q_{250\mu m}$ is likely a reflection of the blank field source counts. There are many more faint radio sources than bright ones. These faint radio sources have higher possibility to coincide with a *Herschel* source, leading to higher $q_{250\mu m}$. This implies that only 18.0 ± 12.0% of chance-alignment sources will have $q_{250\mu m}$ values comparable to true cross-IDs. Given the 7.33 ± 0.04% Poisson probability of chance alignments, the probability of random MeerKAT sources matched to *Herschel* sources having comparable $q_{250\mu m}$ is only 1.35 ± 0.86%, which is reduced by a factor of ~ 6.

Similar procedures are carried out for the 17.6″ matching radius catalogue. There are now 26.4±2.50% of the 438 *Herschel* sources in G014 that have matched simulated sources, 11.2 ± 2.90% of these matched sources have multiple counterparts. 92.7 ± 3.72% of *Herschel* sources with multiple radio counterparts have 2 components, 6.80±0.75% sources have 3 components and 0.46±0.19% have four. In total, the percentage of *Herschel* sources that have multiple aligned-by-chance counterparts is only 3.00 ± 0.70%, which implies that our identifications of *Herschel* sources with multiple counterparts are reliable. The $q_{250\mu m}$ of *Herschel* sources with matched simulated sources are presented in Figure 2.19. The average $q_{250\mu m}$ for sources with multiple counterparts and without multiple counterparts are 2.38 ± 0.36 and 2.74 ± 0.35 respectively. Sources without multiple counterparts have higher $q_{250\mu m}$ than both our results from real MeerKAT sources and the results in Ivison et al. (2010a), while sources with multiplicity have comparable values. The fact that sources without multiplicity here still have higher mean $q_{250\mu m}$ supports our hypothesis that the difference between $q_{250\mu m}$ of *Herschel* sources with and without multiplicity is a bias due to our calculation method as discussed in Section 2.4.2, which may also explain why the $q_{250\mu m}$ of sources with multiplicity are comparable to real values. 25.0±4.0% of matched sources have $q_{250\mu m}$ located inside the 1σ scatter of the $q_{250\mu m}$ in Ivison et al. (2010a), indicating this fraction of random radio sources matched to *Herschel* sources by chance will have $q_{250\mu m}$ comparable to real cross-IDs. Combining this with the 34.2 ± 0.2% Poisson probability of chance alignment at 17.6″ matching radius, the probability of finding a chance-associated radio counterpart will be increased to 8.47 ± 1.52%.

It is thus fair to say that using the FIRC coefficient $q_{250\mu m}$ as a constraint will indeed improve the reliability of *Herschel*-MeerKAT cross-identification. It can also be used to decide the most likely counterpart(s) of a *Herschel* source if there are multiple radio sources coinciding with it. However, it should be noted again that the FIRC coefficient $q_{250\mu m}$ used here is not K-corrected due to the lack of redshift information. Therefore it is dependent on redshifts of the sample and the significance of using $q_{250\mu m}$ to improve cross-IDs will depend on redshifts as well. Moreover, using a $q_{250\mu m}$ cut will raise the risk of discarding true identifications with unusual q, for example radio-loud sources as shown in Figure 2.5i. Without more data, it is difficult to determine whether these sources with unusual q_{250} are correct identifications or not.

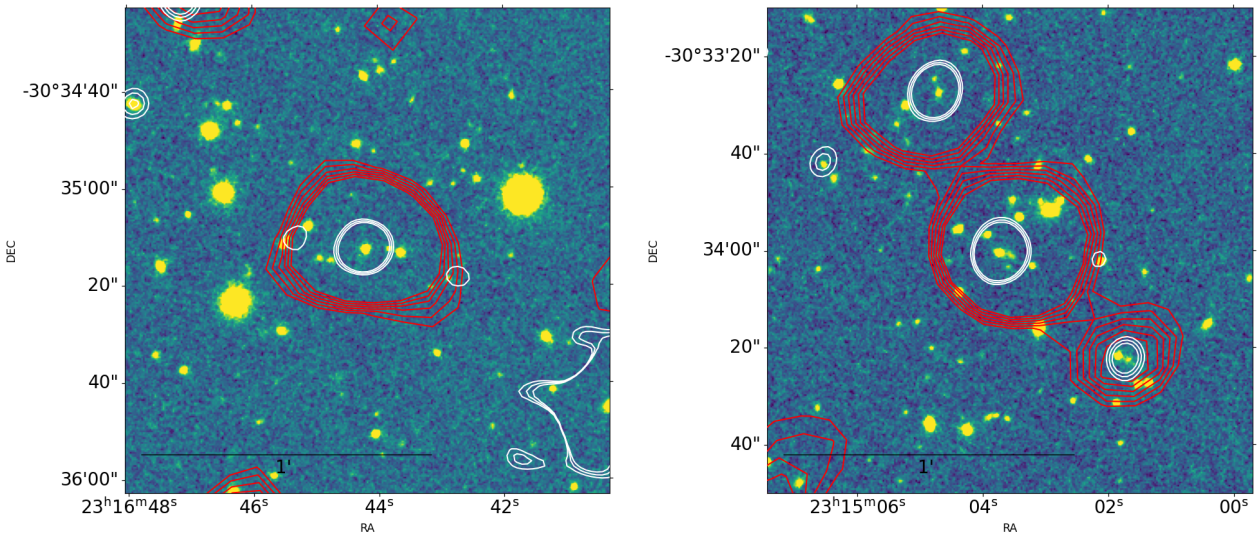


Figure 2.20: Two examples of MeerKAT data helping *Herschel* identifications with SHARKS sources. The background images are the SHARKS Ks data, the red contours show the $5-10\sigma$ *Herschel* $250\mu m$ flux density, the white contour shows the $3-5\sigma$ MeerKAT 1.28GHz flux density. In the left panel, the large *Herschel* source has overlapped with multiple faint SHARKS sources, which prevents direct Ks identification of this source. However, with the help of MeerKAT data, where there is only one source coincides with the *Herschel* source, the SHARKS identification is pinpointed. Similarly in the right panel, 3 *Herschel* sources are connected to each other and cover a lot of faint SHARKS objects, but there is one MeerKAT counterpart for each of the *Herschel* sources, which greatly helps locating the potential true SHARKS identifications.

Overall, using the FIRC coefficient as a constraint to *Herschel*-radio cross identifications potentially provides a relatively simple method to improve the reliability of the cross-IDs. However, care must be taken as it will also exclude potential good identifications of radio-loud sources (or radio-quiet sources with large q), combining this method with other techniques such as visual examination of excluded sources can be a more reliable way of using it. More investigations should be done before one can safely use the FIRC to improve *Herschel* identification.

2.6 Accurate Optical/IR Cross-IDs of *Herschel* sources

The identification of *Herschel* sources in optical/NIR data is difficult to implement, as optical/NIR surveys such as DES and SHARKS generally have much better angular resolution and higher source number densities, resulting in multiple optical/NIR sources overlapping with each *Herschel* source.

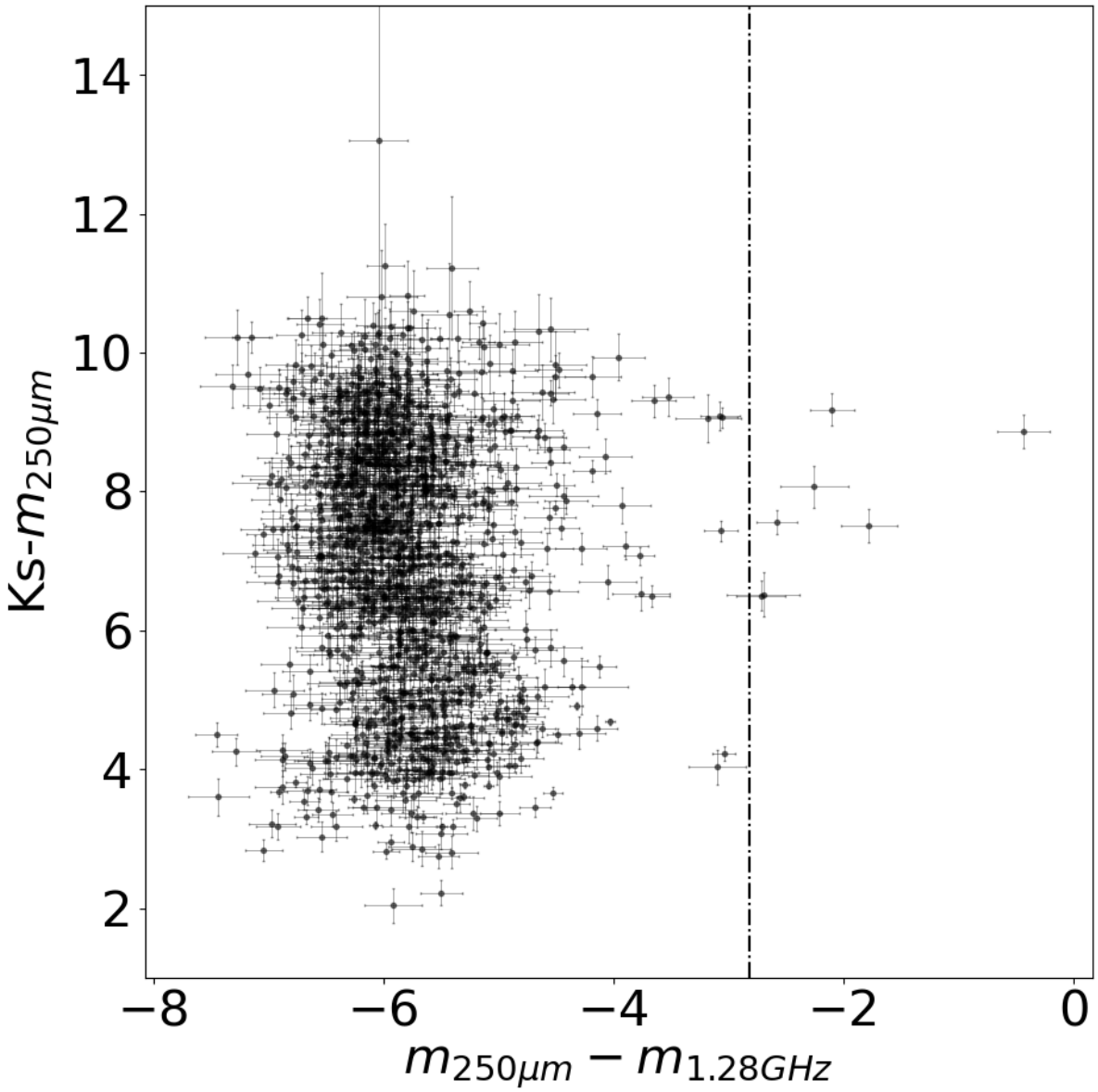


Figure 2.21: Colour-colour diagram between Ks, 250μm and 1.28GHz. The y-axis is the Ks magnitude minus 250μm magnitude, the x-axis is 250μm magnitude minus 1.28GHz magnitude. The 250 μm and 1.28GHz magnitudes are converted from the corresponding flux densities. The dashed-dotted line shows the same 5σ boundary of the FIR-radio colour as in Figure 2.13, sources lie to the right of this boundary are considered as radio-loud.

Moreover, *Herschel* sources are generally dusty objects which means they may be very faint in the optical/NIR, making their optical/NIR counterparts harder to find. Conventionally, *Herschel* cross-identifications with optical/NIR data are carried out using methods such as likelihood ratio to determine the most likely counterpart between several optical/NIR objects overlapping with one *Herschel* source (e.g., Smith et al. 2011). However, all these difficulties no longer exist for *Herschel*-radio cross-identification, since dust is transparent to radio emission and the FIRC can be used to determine the most likely counterpart of a *Herschel* object. With the help of high-quality radio data such as the MeerKAT data presented in this work, we can obtain accurate positions of *Herschel* sources which will significantly help the cross-identification of these *Herschel* sources in optical/NIR data. Here, we make use of the recently released SHARKS Ks data to demonstrate the power of radio data in *Herschel* identification. Two examples are illustrated in Figure 2.20. As can be seen, *Herschel* sources generally coincide with many faint SHARKS Ks sources. With the help of MeerKAT data, the true SHARKS identification of *Herschel* sources can be either directly pinpointed or limited to a small number of potential counterparts. Precise identifications will enable more multi-wavelength analysis, such as accurate photometric redshift estimates to search for evidence of protoclusters. A quick demonstration of such multi-wavelength analysis utilising the SHARKS Ks data is shown in Figure 2.21, where the colour-colour relations between Ks, 250μm and 1.28GHz are illustrated, the 250 μm and 1.28GHz flux densities are converted to AB magnitudes for comparison. There are some sources to the right of the figure, separate from most of the others. This means that they have both a high 250μm to Ks flux ratio and a high 1.28GHz to 250μm flux ratio. A high FIR-NIR flux ratio implies that they may have low star formation rates, while objects with high radio-FIR flux ratios are likely to be AGNs as discussed in Section 2.4.2. Therefore, these sources could potentially be low SFR AGNs, which are rare objects requiring more samples to be studied. Accurate identifications of *Herschel* sources in optical/NIR with the help of radio data can provide a quick and easy way to select such rare objects from large surveys.

Despite its power, there are still some problems in using MeerKAT data for *Herschel* identifications in the optical/NIR. First, the average synthesised beam size of MeerKAT is significantly larger than those in optical/NIR surveys. Thus a MeerKAT source may encompass several faint optical/NIR objects, as shown in Figure 2.20. Secondly, some faint *Herschel* sources match no MeerKAT counterpart or

only a tentative MeerKAT detection, and hence cannot be identified in optical/NIR data using this method. The next-generation radio telescopes such as the SKA will improve these issues with their better resolutions and sensitivities.

2.6.1 Photometric Redshift Estimations

Using the accurate positions of *Herschel* sources obtained from MeerKAT IDs, we cross-match our sources to several optical/NIR catalogues and then estimate their photometric redshifts, followed by analysis to further examine if the three DSFG protocluster candidates are real.

We cross-match our sources to the KiDS data release 4 (DR4) catalogue (Kuijken et al. 2019) which is an optical survey in $ugri$ bands. KiDS DR4 additionally includes aperture-matched $ZYJHK_s$ photometry from their partner NIR survey VIKING (Edge et al. 2013), resulting in a 9-band optical-NIR photometric catalogue. KiDS DR4 covers the entire areas of G017 and G257. Over half of G014 including the most of the central area is covered, with a tile in the eastern edge not included. Figure 2.22 illustrates the KiDS DR4 footprint in G014.

Following a similar procedure in Section 2.4, we choose a matching radius of $2''$ to cross-match our *Herschel*-matched MeerKAT sources to the KiDS sources, which leads to 105, 350, and 498 cross-matched sources in G014, G017, and G257 respectively. Excluding G014 since it is not fully covered by KiDS and VIKING, about 58.4% and 62.3% of *Herschel*-MeerKAT matched sources in G017 and G257 are identified, leading to a total matching rate of $\sim$ 60.7%. Figure 2.23 shows the $10' \times 10'$ centres of our MeerKAT images, together with matched KiDS r band images and *Herschel* 250μm images. As shown by the cutout images, a significant fraction of MeerKAT-*Herschel* matched sources have no KiDS r band counterpart which demonstrates the $\sim$ 60.7% matching rate found.

With the 9-band optical-NIR photometry, we then use the SED template fitting code EAZY (Brammer et al. 2008) to estimate the photometric redshifts of our sample. The default EAZY_v1.3 galaxy template set is used. This template set is designed to incorporate nebular emission lines and also includes a dusty galaxy SED template and an extremely blue SED with strong line emission. EAZY allows all the templates to be included in the final results using fits of N-linear combinations. No

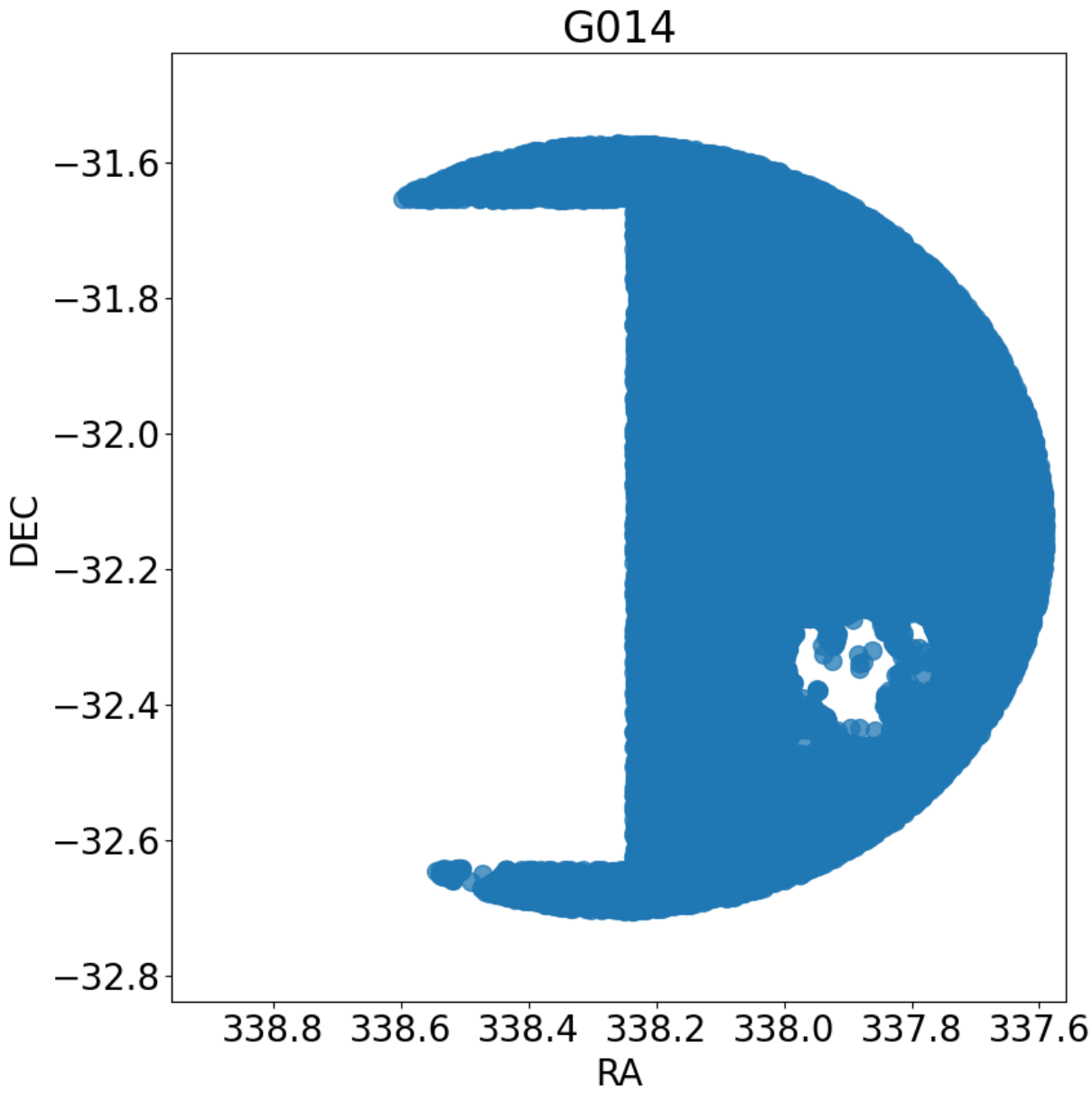


Figure 2.22: The KiDS DR4 catalogue footprint in G014. Most of the central region is covered, a tile in the east is not covered. There is also a hole in the KiDS catalogue in this region, which is the result of a very bright star.

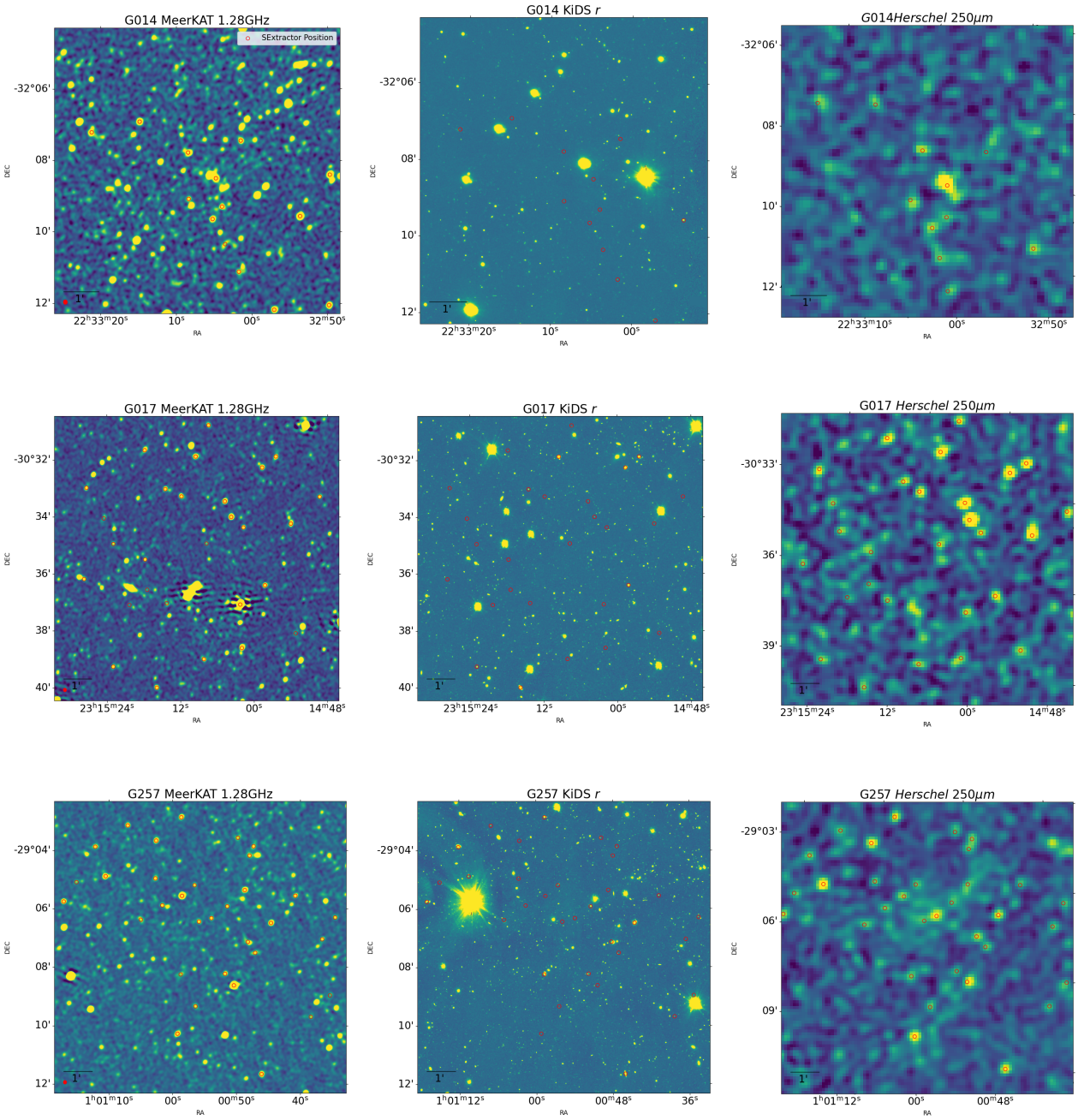


Figure 2.23: $10' \times 10'$ cutouts of the centres of our MeerKAT 1.28GHz images (left column), and positionally matched KiDS DR4 *r* band images (middle column) and H-ATLAS *Herschel* 250μm images (right column). From top to bottom is G014, G017 and G257 respectively. The red circles show the SExtractor positions of MeerKAT-*Herschel* matched sources. The MeerKAT synthesised beam shape in each field is shown as red solid ellipse in the bottom left of each MeerKAT cutout. A $1'$ scale bar is also shown in the bottom left corner of each cutout. Cutouts of G014 are slightly smaller with a size of $10' \times 8'$ as G014 is not fully covered by KiDS DR4. It can be seen that while most *Herschel* sources have clear MeerKAT counterparts, a significant fraction of MeerKAT-*Herschel* matched sources have no KiDS *r* band counterpart.

magnitude-dependent redshift prior is applied to our fit. The average 68% confidence interval of our EAZY redshift estimation is 0.174. By cross-matching our KiDS-matched sample to the 2dF spectroscopic catalogues, we are able to evaluate the accuracy of our EAZY redshifts. There are in total 119 of our sample with 2dF redshifts. We calculate the photometric redshift deviation as $|\Delta z/(1+z_{spec})|$ where $\Delta z = z_{phot} - z_{spec}$. An average of 0.0383 is found. Following Brammer et al. (2008), we further calculate the normalised median absolute deviation σ_{NMAD}, which is less sensitive to outliers, as:

$$\sigma_{NMAD} = 1.48 \times median\left(\left|\frac{\Delta z - median(\Delta z)}{1+z_{spec}}\right|\right) \tag{2.7}$$

The average σ_{NMAD} of the 119 sources is 0.178. This shows that the accuracy of our EAZY redshifts are reasonably good.

We then examine the redshift distributions of sources within the 4.63′ radius circles at the centres of the three observed fields where the candidate protoclusters are. Figure 2.24 shows the redshift distributions of sources in the centres as well as the redshift distributions of all sources in the fields. There are 3, 11, and 14 sources with optical-IR IDs in the centres of G014, G017 and G257, while the numbers of all *Herschel*-MeerKAT matched sources in the same areas are 17, 24, and 26. The fraction of $z \geq 1$ sources is significantly larger in the centres in comparison to the overall distribution for G017 and G257, which is 45.0% and 57.1% respectively, while the same fraction for the entire field of G017 and G257 is 24.0% and 24.5%. In G257, there are 5 sources at $0.9 < z < 1.2$, leading to a density contrast of 2.6 ± 1.7 at this redshift range. This is potentially a sign of a $z \approx 1$ overdensity associated with the candidate protocluster. However, the numbers of sources in the centres are too small to draw any statistically robust conclusion directly from them. Excluding G014, about 50.0% of *Herschel*-MeerKAT matched sources are not seen in KiDS and VIKING optical/NIR data, implying these sources have very faint optical/NIR counterparts and potentially lie at high redshifts. Given the $\sim 60.7\%$ identification rate in the entire fields of G017 and G257, the lower identification rate ($\sim 50\%$) in the field centres suggests that there may be more high redshift *Herschel* sources in the centres than in the whole fields with their optical/NIR counterparts too faint to be detected in KiDS/VIKING.

Since a DSFG protocluster may also have underlying non-starbursting members associated with it, overdensities seen in optical/NIR data along the line of sight of our candidate protoclusters will also

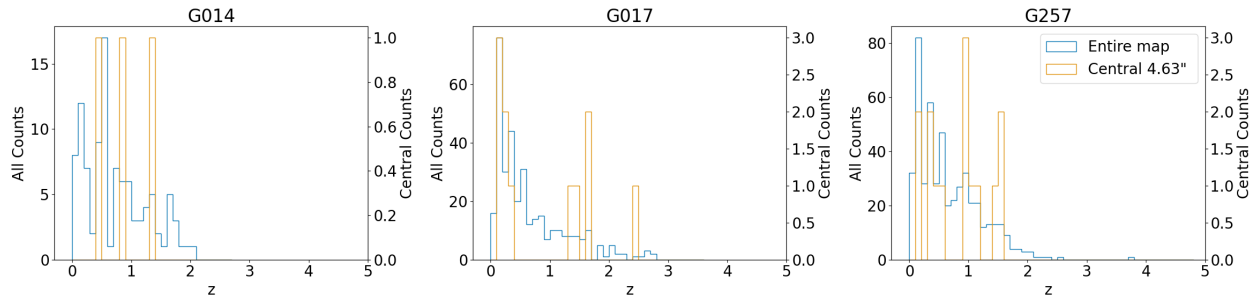


Figure 2.24: Photometric redshift distributions of *Herschel* sources identified in both MeerKAT and KiDS/VIKING in the three observed fields. The left y-axis shows the total number of sources in the field, while the right y-axis shows the number of sources in the 4.63′ radius centre of the field. A bin size of 0.1 is used. The blue histogram shows the redshift distribution of sources in the entire map, while the orange histogram shows the distribution of sources inside the 4.63′ radius circle in the field centre. The fraction of $z \geq 1$ sources is significantly larger in the centres in comparison to the overall distribution for G017 and G257. In G257, there are 5 sources at $0.9 < z < 1.2$, leading to a density contrast of 2.6 ± 1.7 at this redshift range which is potentially a sign of overdensity.

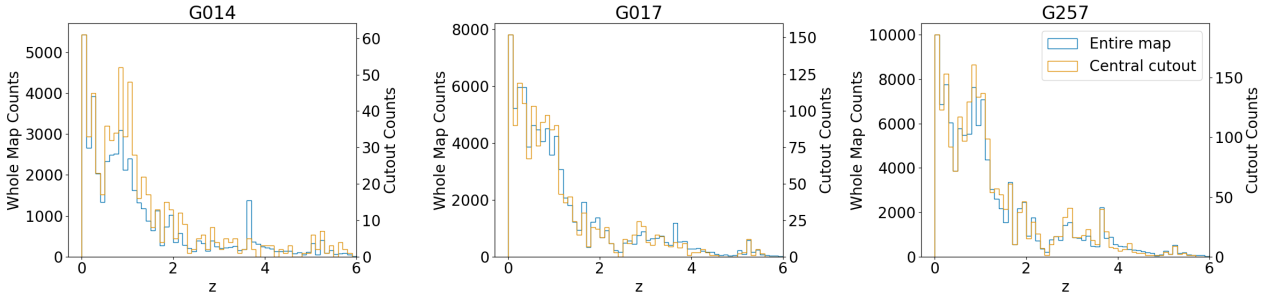


Figure 2.25: Photometric redshift distributions of all KiDS sources identified in the three observed fields. The left y-axis shows the total number of sources in the field, while the right y-axis shows the number of sources in the 4.63′ radius centre of the field. A bin size of 0.1 is used. The blue histogram shows the redshift distribution of sources in the entire map, while the orange histogram shows the distribution of sources inside the 4.63′ radius circle in the field centre. No apparent sign of overdensity of optical sources is seen in any of the field.

be evidence of real protoclusters. We additionally check the redshift distributions of all KiDS sources in the three fields, without matching them to either *Herschel* or MeerKAT objects. The result is shown in Figure 2.25. The redshift distribution of central sources are basically identical to the overall distributions, showing no apparent sign of overdensity of optical sources. Nonetheless, as there are a large ratio of *Herschel* sources not being detected in both KiDS/VIKING, deeper optical/NIR observations may reveal more traces of overdensities in the observed fields at higher redshifts and provide further evidence of the candidate protoclusters.

Greenslade et al. (2018) estimates that most member galaxies of the candidate protoclusters should lie

at $z > 1$ and the optical/IR survey data used here is probably not deep enough to detect most sources at $z > 1$ (over 50% KiDS sources have $z_{phot} < 1$, see Figure 11. in Kuijken et al. 2019). This means that the majority of *Herschel* sources identified in KiDS/VIKING are unlikely to be associated with our candidate protoclusters. By excluding *Herschel* sources with photometric redshift lower than 1, we can reduce low-redshift contaminations and better isolate sources in higher redshift overdensities. Figure 2.26 shows the density contrast maps before and after the removal. The maps are created by calculating the local density contrast of KiDS-*Herschel*-MeerKAT matched sources within a 4.63′ radius circle centred at every location across the image. At the centre of each field, the density contrast δ is 1.1 ± 0.5, 1.0 ± 0.4, and 0.64 ± 0.34 for G014, G017 and G257 before removing any source, whereas $\delta(z > 1)$ become 1.2 ± 0.6, 1.8 ± 0.7, and 1.1 ± 0.5 after removing sources with EAZY redshift lower than 1. It can be seen that at $z > 1$ the central regions appear to have at least twice the number density as the surrounding fields and the significance of the central density contrasts increase after the removal, which supports the existence of $z > 1$ protoclusters in the centres of observed fields. Note that the change of density contrast in the centre of G014 after removing low redshift sources is less significant. It is likely because of the incomplete coverage of KiDS/VIKING in this field, thus the number of sources with photometric redshift estimations in this field is smaller and the effect of removing $z < 1$ sources is less obvious.

With the accurate positions of *Herschel* sources provided by MeerKAT, we will be able to obtain more IDs from future deeper optical/NIR data of high redshift *Herschel* sources coming from e.g. James Webb Space Telescope (JWST; Gardner et al. 2006), Euclid (Laureijs et al., 2011) and Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019), and hence identify any overdensity associated with our candidate protoclusters. As the next step of this work, we have recently carried out deep optical/NIR observations on a number of DSFG protocluster candidates including the 3 studied here. Future work using our MeerKAT data to identify optical/NIR *Herschel* counterparts in this new data is underway.

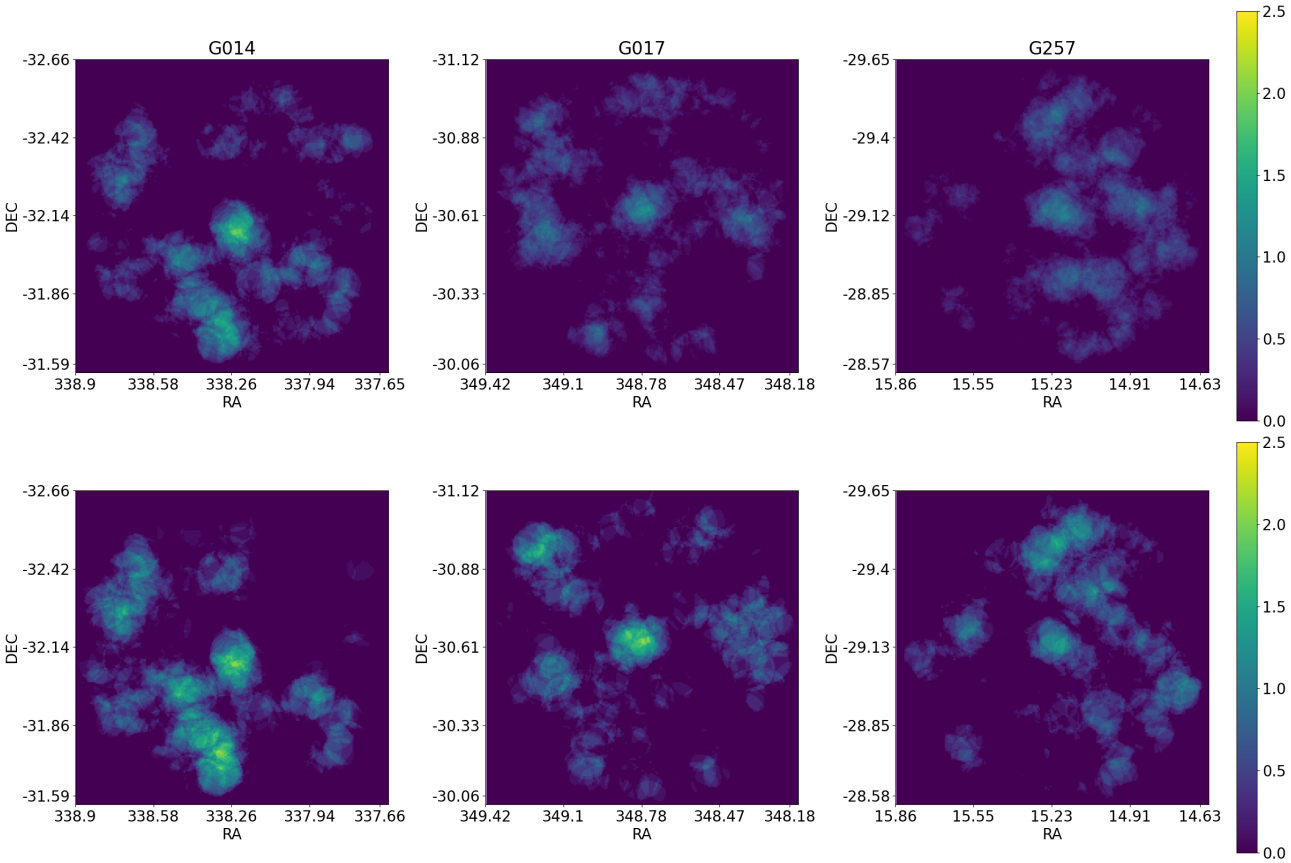


Figure 2.26: Density contrast maps of each observed field, with RA and DEC shown on x and y axes. For every pixel in these maps, the colour represents the local density contrast of KiDS-MeerKAT-*Herschel* matched sources within a 4.63′ radius circle centred at this pixel. The colour scale is the same for all panels. The top row shows the three fields before removing sources with EAZY redshift lower than 1, while the bottom row shows the results after removing $z < 1$ sources. The density contrasts δ are 1.1 ± 0.5, 1.0 ± 0.4 and 0.64 ± 0.34 for G014, G017 and G257 before removing any sources, whereas $\delta(z > 1)$ become 1.2 ± 0.6, 1.8 ± 0.7 and 1.1 ± 0.5 after the removal. The density contrasts in the centres of G017 and G257 have significantly increased after removing low redshift sources, suggestive of $z > 1$ protoclusters in the centres of observed fields. The change in G014 is less significant which is likely due to the incomplete coverage of optical/NIR photometry in this field.

2.7 Conclusion

In this chapter, we present MeerKAT 1.28GHz radio imaging observations of 3 candidate DSFG protoclusters selected from Greenslade et al. (2018). Sources are extracted using SExtractor, and then matched to the H-ATLAS DR2 in the SGP field. We do not find direct evidence of *Herschel* protoclusters from our MeerKAT data alone, but with the help of accurate MeerKAT positions, the optical/NIR IDs and photometric redshifts of our *Herschel* sources show evidence of real protoclusters in G017 and G257 at $z > 1$. More detailed results and conclusions are summarised as follows:

(a) Our 1.28GHz source counts are comparable to other literature results. There is an excess in our G257 counts at $S \approx 0.1\ mJy$ which is suggestive of an overdensity. By directly extracting sources from the MIGHTEE data (Heywood et al. 2022) using the same extraction procedures as in this work, we find that this excess disappears. It is likely caused by difference in source extraction and source count corrections applied. Source counts just in the central $4.63'$ radius regions where the candidate protoclusters lie do not reveal any evidence of overdensity either.

(b) *Herschel* sources are cross-matched to our MeerKAT sources using a matching radius of $7.5''$. The Poisson probability of chance association contaminants is 7.33% at this searching radius. Approximately 95% of the *Herschel* sources in our observed regions have a MeerKAT counterpart. Visual inspection of *Herschel* sources without MeerKAT counterparts reveal that most of them overlap with an extended 1.28GHz source, but the positions of these extended radio sources as assigned by SExtractor are $> 7.5''$ away from the *Herschel* sources. Correcting for these sources with extended morphologies leads to a near 100 percent matching rate.

(c) To study the multiplicity of *Herschel* sources in 1.28GHz data, we also carry out another cross-identification of *Herschel* sources to our MeerKAT sources, using a larger matching radius of $17.6''$, which is the beam size of the *Herschel* SPIRE $250\mu m$ band. We find 1885 out of the 1914 *Herschel* sources have MeerKAT counterparts at this matching radius with 1038 out of the 1885 *Herschel* sources having 2 or more radio components. The $250\mu m$ flux distributions of *Herschel* sources with and without multiplicity are statistically different, but the same difference is not shown in the 350 and $500\mu m$ flux distributions. We suspect that this may be a result of insufficient faint sources in the 350

and 500μm bands due to worse sensitivities. The distribution of $q_{250\mu m}$ of *Herschel* sources with and without multiplicity are also different, but this is likely a bias induced by our method of calculating $q_{250\mu m}$ of *Herschel* sources with multiple 1.28GHz components.

(d) The FIRC is investigated by calculating the monochromatic FIRC coefficient $q_{250\mu m}$ of our sample. The average $q_{250\mu m}$ of our sources is 2.33 ± 0.26, which is comparable to that in Ivison et al. (2010a). Using $q_{250\mu m}$ as a constraint to the *Herschel*-Radio identification can greatly reduce the probability of contaminations due to chance alignments by a factor of ~ 6 for the $7.5''$ matching radius catalogue, but with the risk of ignoring potential good identifications of unusual sources such as radio-loud AGNs.

(e) The *Herschel* SPIRE colours of sources in the central regions, which are potential members of the candidate protoclusters, indicate that their redshifts should be at least larger than 1. No significant correlation between the *Herschel* colours and the 1.28GHz flux density or $q_{250\mu m}$ is found.

(f) We demonstrate that MeerKAT cross-IDs can significantly help *Herschel* source identifications in optical/NIR data. We cross-match of our *Herschel* sources with the public KiDS/VIKING data using MeerKAT positions, then carry out photometric redshift estimations with the SED fitting code EAZY. About 60.7% of *Herschel* sources are identified but only $\sim$ 50.0% of *Herschel* sources in the central $4.63'$ radius regions have optical/NIR IDs, which implies that there are more higher redshift *Herschel* sources with very faint optical/NIR counterparts in the centres of the fields. The redshift distribution of sources in the centre of G257 shows a unique redshift peak at $0.9 < z < 1.2$, corresponding to a group of 5 sources with a density contrast of $\delta = 2.6 \pm 1.7$. This could be evidence of the candidate protocluster in G257 being real protocluster, but the robustness of this conclusion is hindered by small number statistics. By removing sources with EAZY redshift lower than 1, the density contrasts of the $4.63'$ centres of G017 and G257 increase from $\delta = 1.0 \pm 0.4$ and 0.64 ± 0.34 to $\delta = 1.6 \pm 0.6$ and 1.0 ± 0.5, implying there are overdensities in these regions at $z > 1$. With deeper MeerKAT observations or more powerful radio telescopes such as the SKA coming in the future, radio observations will be very powerful in identifying FIR/submm sources with optical/NIR data. This will be the future of optical/NIR IDs in FIR studies. Future work using our MeerKAT data to identify optical/NIR *Herschel* counterparts is underway and will be discussed elsewhere.

Future studies can expand the work in this chapter by combining more *Herschel* and radio observations.

For example, the complete MIGHTEE survey will cover an area of $\sim$ 20 deg^2 to μJy sensitivity. Combining future MIGHTEE data with existing *Herschel* data will significantly increase the number of *Herschel* sources with precise radio IDs, allowing cross-IDs in the optical/NIR and hence photometric redshifts. The large number of *Herschel* sources with photometric redshifts will then enable further studies of these FIR bright sources, such as the evolution of their luminosity function with redshift. Future SKA data, together with higher quality optical/NIR data from new instruments such as JWST, Euclid and LSST, will further provide more accurate positions of *Herschel* sources and dramatically boost the number of *Herschel* IDs in the radio and optical/NIR.

Chapter 3

Deep Optical/NIR Observations of Candidate DSFG Protoclusters

3.1 Introduction & Motivation

As discussed in Section 1.4.5 of Chapter 1 as well as in Section 2.1 of Chapter 2, protoclusters play a vital role in galaxy formation and evolution. However, protoclusters at high redshifts, especially those at $z > 4$, are still poorly understood due to the lack of observations. This is mainly because there is no reliable, efficient method to search for high redshift protoclusters. At the present, the only reliable way of detecting protoclusters is to look for physical overdensities of galaxies which are likely to form clusters at $z \approx 0$. DSFGs are thought to be good tracers of such overdensities as they are likely ancestors of present-day giant ellipticals which are often found in galaxy clusters (e.g., Stevens et al. 2003; Capak et al. 2011; Casey et al. 2014; Toft et al. 2014; Oteo et al. 2018; Miller et al. 2018; Hill et al. 2022). However, there are also some studies showing that SMGs/DSFGs are in fact poor tracers of protoclusters (e.g., Chapman et al. 2009; Miller et al. 2015; García-Vergara et al. 2020). Whether or not DSFGs trace high redshift protoclusters remains a topic of debate today, therefore more observations of protoclusters at various stages of evolution are needed to reach a decisive conclusion.

As already described in Section 2.1, Negrello et al. (2005) first proposed a method to search for DSFG

overdensities, utilising the multiplicity of DSFGs seen in large-area, low-resolution FIR/submm surveys such as those carried out by *Planck* and *Herschel*. This method of selecting DSFG overdensities was later successfully demonstrated, with a number of candidate protoclusters as well as real protoclusters confirmed in this way (e.g., Clements et al. 2014; Flores-Cacho et al. 2016; Miller et al. 2018; Greenslade et al. 2018; Martinache et al. 2018). In particular, using *Herschel* data, Clements et al. (2014) examined 16 *Planck* sources in the Early Release Compact Source Catalogue of *Planck* (ERCSC; Piacentini et al. 2011), finding four candidate protoclusters at $z \sim 1 - 2.3$. Greenslade et al. (2018) extended the work of Clements et al. (2014) by combining the full release of the *Planck* Catalogue of Compact Sources (PCCS; Piacentini et al. 2011; Ade et al. 2014, 2016) and large *Herschel* surveys, and identified 27 candidate protoclusters with potential redshifts $z > 1$. SCUBA-2 follow-up observations of 13 of these candidates by Cheng et al. (2019) found that 9 of them exhibit overdensities with photometric redshifts at $z > 2$. Following these previous works, in this chapter we present deep optical/NIR observations of 13 candidate protoclusters selected from Greenslade et al. 2018, with the European Southern Observatory (ESO) Faint Object Spectrograph and Camera version 2 (EFOSC2; Buzzoni et al. 1984) and Son of ISSAC (SOFI; Moorwood et al. 1998). We were awarded two observing nights for SOFI and three nights for EFOSC2 in ESO period 111 (Program ID: 111.250G, PI: Yifan Ding). The 13 targets are selected from Greenslade et al. (2018) by their availability to EFOSC2 and SOFI during the proposed observing semester. Such observations allow us to obtain accurate photometric redshifts of galaxies in these candidates and thus determine if they may be physically associated in a protocluster. These observations also enable us to assess the morphology of member galaxies, including those detected in the FIR/submm as well as the non-starbursting galaxies that must make up the bulk of the underlying cluster galaxy population. Based on these observations, future spectroscopic follow-up of these sources can be carried out to provide complete analyses of these systems and determine their state of virialisation. It should be noted that, due to the non-ideal weather conditions during the EFOSC2 observations, most targets were not successfully observed with EFOSC2 and EFOSC2 images of the remaining targets have not reached the desired depths. Therefore, we will not include the EFOSC2 data in this chapter, only the SOFI observations are presented followed by analysis based on these data and archival photometry.

In Section 3.2, we introduce the instruments used and the observations taken in detail. Section 3.3

describes how the raw data is reduced into final science images. We describe our source extraction and photometry calibration processes in Section 3.4. In Section 3.5, we outline the complementary data used in this chapter and the cross-identification of SOFI sources with this data. We describe the estimation of photometric redshifts and the evaluation of the photometric redshift accuracy in Section 3.6. We analyse the photometric redshifts of our target sources, finding six candidate overdensities at high redshift aligning with the candidate protoclusters. We discuss the limitations of our photo-z analysis in Section 3.8 and conclude this chapter in Section 3.9.

3.2 Instrumentation and Observation

We select 13 candidate protoclusters from Greenslade et al. (2018) to be observed with the New Technology Telescope (NTT) at the ESO La Silla Observatory (Figure 3.1). Targets are selected based on their availabilities during the observation dates at the observatory. The NTT is a 3.6m diameter telescope which saw its first light in March 1989. It is the first telescope featuring flexible, controlled mirrors, a technology named active optics (Wilson et al., 1987; Noethe et al., 1988; Wilson et al., 1989, 1991). It has three optics mirrors, the primary mirror is actively controlled to preserve its figure at all telescope positions. The secondary mirror position is also actively controlled in three directions. The telescope chamber is ventilated by a system of flaps which optimize the air flow across the NTT minimizing the dome and mirror seeing, and a combination of water-cooling system and reflective insulation is used to control the temperature of the telescope. The optimized airflow, the thermal controls, and the active optics provide excellent image quality at the NTT. There are two Nasmyth focus platforms to mount two sets of instruments simultaneously, the third mirror of the telescope is used to switch between the two platforms. There are two instruments mounted at the NTT, one is the optical instrument EFOSC2 for low resolution spectroscopy and imaging, the other one is the infrared spectrograph and imaging camera SOFI. It should be noted that SOFI has been decommissioned since August 2023. The SOFI observations presented here were the last observation run for SOFI. The two instruments will be briefly described below.

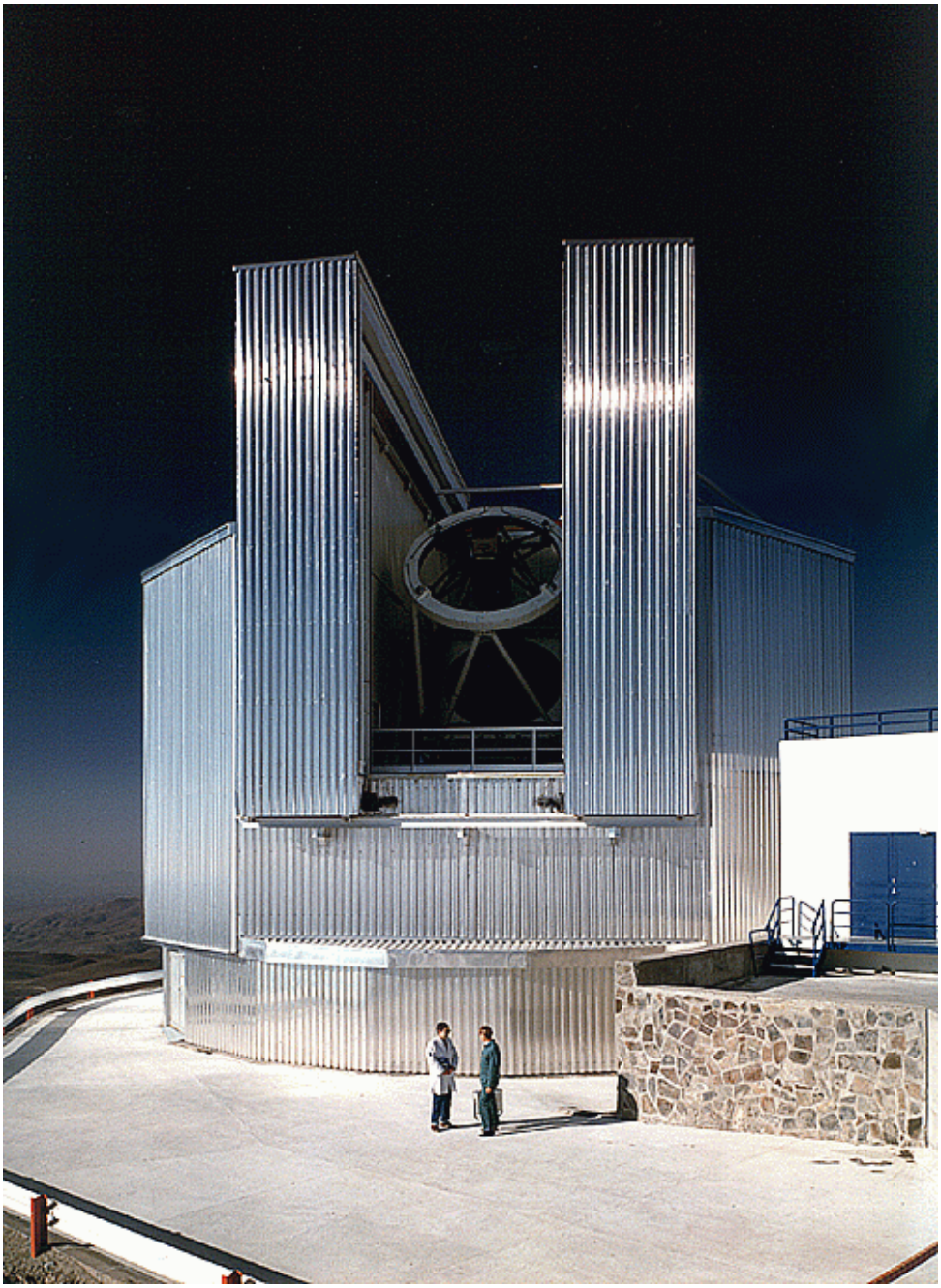


Figure 3.1: The ESO 3.6 m New Technology Telescope. Image credit: ESO/C.Madsen (`https://www.eso.org/public/images/eso0912b/`)

EFOSC2; Buzzoni et al. (1984) EFOSC2 is a focal reducer using multi–layer coated all–transmission optics, providing both imaging and spectroscopy observation modes. More specifically, EFOSC2 offers seven different observing modes, three for imaging and four for spectroscopy including polarimetry. These modes are achieved using different combinations of the optical components which can be inserted into the instrument. We will only focus on the imaging mode here as it is the one used for our observations. The field of view of EFOSC2 is $\sim 4.1' \times 4.1'$ with a pixel scale of $0.12''$. The basic set of EFOSC2 imaging filters contains 15 filters, including 8 broad band filters (Bessel $UBVR$ filters and Gunn $griz$ filters) and 7 narrow band filters ranging from $\sim$ 444.5 to $\sim$ 673.0 nm. EFOSC2 was designed as an upgraded version of EFOSC1 to be mounted at the NTT. In 1997 EFOSC1 was decommissioned and EFOSC2 was moved to the 3.6m telescope at La Silla Observatory where it was offered until 2007. In April 2008, EFOSC2 was mounted on the Nasmyth B focus of the NTT until now.

SOFI; Moorwood et al. (1998) SOFI was mounted on the Nasmyth A focus of the NTT. A mask wheel consists of a variety of masks enables several imaging and spectroscopy modes. Again we will focus on the imaging modes used for our observations. SOFI offers imaging at two objective modes, in Large Field Objective imaging mode has a field of view of $4.92' \times 4.92'$ with a pixel scale of $0.288''$, the Small Field Objective imaging mode provides a field of view of $4.66' \times 4.66'$ with a pixel scale of $0.273''$. Imaging can be done through the standard IR broad band filters: J, J_s, H, K_s , and a Z filter which peaks at 0.9 microns, and many narrow band filters. SOFI was installed at the NTT in 1997 and eventually decommissioned at the end of August 2023.

The 13 candidate protoclusters selected as targets have a wide range of locations in the southern hemisphere, therefore they are covered by various public optical/NIR surveys. For example, there are 7 candidates covered by DES (Abbott et al., 2018, 2021) in $grizY$ bands, 11 candidates covered by near-IR SHARKS Ks data (Dannerbauer et al., 2022) and 3 candidates observed at 1.28GHz radio wavelength with MeerKAT as mentioned in Chapter 2. Our EFOSC2/SOFI observations were designed such that all 13 candidate protoclusters would eventually have deep 8 band optical/NIR photometry, including the optical $Ugriz$ filter bands and NIR JH & K_s bands. Table 3.1 shows the coordinates of the 13 targets, filter bands in which they were observed with EFOSC2 and SOFI and

Table 3.1: Names, positions, EFOSC2/SOFI filter used and EFOSC2/SOFI observation date for each candidate protocluster observation. Targets which were not observed by EFOSC2 are shown as "_" in the EFOSC2 observation date column.

Name	RA(J2000)	DEC(J2000)	EFOSC2 Filter Used	SOFI Filter Used	EFOSC2 Observation Date	SOFI Observation Date
G007	$23^h01^m27^s.840$	-35°06′10″.800	*Ugriz*	*JH*	16/17 Aug 2023	13/14 Aug 2023
G014.92	$22^h26^m32^s.400$	-32°10′37″.200	*Ugriz*	*JH*	_	13/14 Aug 2023
G014.99	$22^h33^m02^s.400$	-32°08′20″.400	*Ugriz*	*JH*	16/17 Aug 2023	13/14 Aug 2023
G015	$22^h36^m20^s.160$	-31°54′07″.200	*Ugriz*	*JH*	17/18 Aug 2023	13/14 Aug 2023
G017	$23^h15^m09^s.600$	-30°35′27″.600	*Ugriz*	*JH*	16/17 Aug 2023	14/15 Aug 2023
G019	$22^h29^m14^s.640$	-29°40′12″.000	*Ugriz*	*JH*	_	14/15 Aug 2023
G222	$04^h47^m59^s.040$	-52°38′34″.800	*U*	JHK_s	_	14/15 Aug 2023
G239	$01^h41^m19^s.920$	-31°47′09″.600	*U*	*JH*	_	13/14 Aug 2023
G252	$01^h06^m59^s.760$	-29°54′36″.000	*U*	*JH*	17/18 Aug 2023	14/15 Aug 2023
G257	$01^h00^m55^s.920$	-29°07′19″.200	*U*	*JH*	17/18 Aug 2023	14/15 Aug 2023
G260	$03^h32^m11^s.040$	-27°07′01″.200	*U*	JHK_s	_	14/15 Aug 2023
G261	$04^h40^m40^s.320$	-53°44′02″.400	*U*	JHK_s	17/18 Aug 2023	14/15 Aug 2023
G354	$00^h12^m11^s.760$	-33°13′40″.800	*U*	JHK_s	_	13/14 Aug 2023

their EFOSC2/SOFI observation dates. It should be noted that G014.99, G017 and G257 are the same candidate protoclusters studied in Chapter 2.

Our SOFI observations were carried out from 13^{th} to 14^{th} August 2023. Flat field images of the telescope dome in J, H & K_s filters were taken in the afternoon before each observing night. At the beginning of each observing night, three standard stars in the vicinity of the science targets were observed in J, H and K_s in the default jittered observation mode, with a 10 second exposure time for one frame and 5-6 exposure frames in total for each filter, resulting in a $50-60$s integration time for each standard star in each band. For the target observation, we chose to use the Large Field Imaging mode of SOFI as it was the only mode SOFI provided at that time. In order to maximise the quality of observed images in the allocated telescope time, the observing strategy needs to be carefully planned. For a single exposure, the detector integration time (DIT) should be as small as possible such that bright objects and sky background will not saturate the detector array. The total integration time of this exposure can be controlled by the number of DITs (NDIT). On the other hand, in order to maximise the S/N ratio, the sky intensity in a single frame should be sufficiently high that the sky background noise will dominate over the detector read-out noise, such that the resulting images will be limited only by the background noise which can later be subtracted. Furthermore, shorter DITs will significantly increase the total observation time, as each DIT has some fixed overheads such as detector read-out and shorter DIT means more time spent on such overheads to reach the proposed total integration time. After choosing a suitable DIT, the NDIT can be determined from the desired integration time of one exposure. In optical/NIR imaging, a commonly used observing technique for a long integration time

is to have a number of exposures jitter around the target coordinate with small position offsets. In this way, observers can avoid sky background saturating the detector and create a good sky image without object residuals, remove the array cosmetic defects and allow cosmic rays hitting instrument detectors to be dealt with. The final image can be obtained by combining all the individual jittered image frames. The SOFI user manual advises a 2 minutes integration time for a single exposure, therefore based on our desired 5σ depth we eventually choose to use jitter imaging mode with a combination of $DIT = 30s, 20s, 10s$, $NDIT = 4, 6, 12$ and number of exposures $NEXPO = 9, 12, 36$ for J, H and K_s observations respectively. These result in the total integration time ($DIT \times NDIT \times NEXPO$) of $1080s$, $1440s$ and $4320s$ for J, H and K_s to reach our proposed depths of $J = 22.4$, $H = 22.4$ and $K_s = 22.7$. All of our observations were carried out using the SOFI AutoJitter observing template. The sky conditions during the first observing night ($13/14^{th} Aug$) was clear, with an average humidity of 4% and an average seeing of $0.74''$ for most of the night. The weather conditions during the second night ($14/15^{th} Aug$) were not ideal, with thin to thick cirrus and mild wind over the night. The average humidity and seeing of this night are 10.2% and $1.27''$. No dark or bias frame was taken for our SOFI observations, as the darks will be naturally removed when the sky background is removed. As the NTT tracks very well over most of the sky, guiding is not necessary for observations with short exposure times, as is the case here, and therefore it was not used. All 13 targets were successfully observed with SOFI and with the full proposed on-source time executed.

The EFOSC2 observations were carried out from 15^{th} to 18^{th} August 2023. Similar to the SOFI observations, dome flat field images in $Ugriz$ bands were taken in the afternoon of each day, and several standard stars near to the science targets were observed in the $UBVri$ filter bands at the beginning of each night. For the imaging of the selected targets, we again used the jitter imaging mode for accurate photometry as well as reducing the effect of cosmic ray hits. In comparison to IR imaging, the optical sky has a much lower background level due to better atmospheric transmission, therefore a single optical exposure can tolerate longer integration times and it is not necessary to divide a single exposure into a number of sub-integrations as in the case of SOFI observations. To reach the desired image depths, we observed the selected targets with $450s, 100s, 90s, 85s$ and $70s$ exposure time for a single frame in $Ugriz$ respectively and 10 exposures in each filter. The total integration time for each target is hence $4500s$, $1000s$, $900s$, $850s$ and $700s$ in $Ugriz$. All of our observations were carried out

using the EFOSC2 ImageJit observing template. Unfortunately, the weather conditions during the three EFOSC2 observation nights were very bad and severely affected the observations. For the first night ($15/16^{th}Aug$), the telescope dome was forced to close for the entire night due to strong winds. In the night of $16/17^{th}Aug$, the telescope was closed again in the first half of the night due to overcast sky and strong wind. It reopened later and targets G014.99, G017 and G007 were observed with thick cirrus to cloudy sky conditions, the average humidity was 29.8% and the seeing was varying from 1.64″ to 2.55″. On the third observation night ($17/18^{th}Aug$), the sky had thin cirrus to thick clouds with humidity of 25% – 35% and seeing of 1.87″ – 2.44″. The telescope was forced to close during the observation of G252 due to heavy clouds and reopened an hour later. G015, G252, G257 and G261 were observed on the last night. Overall, only 7 of the 13 targets were observed with EFOSC2 under very poor sky conditions. The average FWHM seeing of these observations ($\sim 2 - 3''$) was significantly higher than our desired 0.8″ seeing condition. The final image quality and desired depths cannot be reached in any band. For this reason, we will focus on the SOFI data in the rest of this chapter. All data or results mentioned in the following sections refer to SOFI.

3.3 Data reduction

All the raw data and calibration frames were obtained from the ESO Science Archive Facility[1]. Our SOFI data is reduced using a combination of IRAF (Image Reduction and Analysis Facility, Tody 1986) and Astropy CCDPROC package (Craig et al., 2023). IRAF is a general purpose software system for the reduction and analysis of astronomical data and has been widely used for image reduction. CCDPROC is an Astropy (Price-Whelan et al., 2022) affiliated package for basic data reductions of images which provides the essential tools for processing of astronomical images in a Python framework.

A raw astronomical image produced from a detector as in the case of SOFI is essentially a 2-dimensional array of values called Analogue Digital Units (ADU) or counts, which come from the detector chip converting the analogue voltage in each pixel to a numerical electron count. Ideally, there should be only counts generated by photons hitting the detector via any sort of the photoelectric effect. However,

[1] https://archive.eso.org/

in reality, there are several other sources of counts in a pixel contaminating the counts related to the actual astronomical objects. Imaging data reduction is therefore basically a process of removing all these unwanted values. In principle, the sources of count contaminations consist of: (a) an offset voltage called bias which is applied to the detector chip to avoid negative counts during readout. The values of the bias will vary across the chip and there can be small variation in the bias level over time; (b) an electric current in the chip called dark current, induced by the thermal motion of electrons. Dark currents vary over time and strongly depend on the temperature of the detector; (c) the background light in the night sky which enters the camera and increases the counts. Typical sources of the sky background are scattered light or OH line emission from the atmosphere, local light sources and moonlight; (d) cosmic rays which hit the detector and generates charge that is converted to counts; (e) the intrinsic read noise to the electronics of the instrument. It is not possible to completely remove this read noise as it is intrinsic to the detector, but there are ways to minimise it. After removing counts from all the sources described above, there should only be counts from astronomical objects remaining. In summary, a raw image is:

$$Raw\ Image = Target + Sky + Bias + Dark + Noise \tag{3.1}$$

However, even after removing all other sources of counts, the remaining counts are still not a correct measurement of the light from the astronomical objects because there are some additional effects affecting the measurement. First, it is very common to see a dimming near the corners of an image caused by the optical system to which the detector is attached, a phenomenon called vignetting. Then, it is also common to have dust on optical elements such as filters or lenses over the chip, reducing the amount of light reaching the detector. Lastly, the sensitivity to light which affects the counts read from the chip is not uniform across the chip, thus light sensitivity varies from pixel to pixel. The effects of vignetting, dust and non-uniform sensitivity are often evaluated by taking an image of an illuminated blank field which is conventionally called a flat field. Since vignetting, dust and non-uniform sensitivity reduce the counts from light reaching the detector, they are therefore affecting the counts of light only, in a multiplicative way. Therefore, the components of a raw image become:

$$Raw\ Image = Flat \times (Target + Sky) + Bias + Dark + Noise \tag{3.2}$$

The goal of data reduction is to extract the final counts from the science target, this is can be derived from Equation 3.2 as follows:

$$Target + Noise = \frac{Raw\ Image + Bias + Dark}{Flat} - Sky \tag{3.3}$$

As already mentioned, it is impossible to remove read noise since it is intrinsic to the detector chip and is completely random. In the case of SOFI, the bias, which is usually measured with a zero second exposure with the camera shutter closed, cannot be measured as the IR detector can not make a zero second exposure. Therefore in this context the bias is an indistinguishable part of the dark currents. The SOFI detector has a very low dark current of $20e^{-}/hour$ and it is generally ignorable in most cases. Moreover, removing the sky background will naturally remove the effects of dark current as well. It is hence not necessary to explicitly calibrate for the bias and darks in our case, and the task of our data reduction now becomes:

$$Target + Noise = \frac{Raw\ Image}{Flat} - Sky \tag{3.4}$$

The first step of our SOFI data reduction is thus to obtain an image of a uniformly illuminated flat surface in each broad band filter. For each filter, a set of flat field images of the dome of the telescope illuminated by a lamp were taken before each observation night. Each set of flat field images includes four images of the dome with the lamp on and off, followed by the same four images but with a mask applied to the detector array. In this way, the thermal component of the signal which is independent of the flat-field lamp intensity can be estimated and removed. These sets of flat field images are combined into one flat field frame for each filter band using an IRAF script SPECIAL_FLAT.CL provided by the SOFI archive[2]. However, the combined dome flats are still likely to contain large scale variations as the illumination of the dome can never perfectly represent the illumination of the actual sky. Moreover, the illumination of the dome can change over time due to variations in the flat-field lamp as well as the temperature of the dome. The solution to this problem is to create an illumination surface via observing a standard star in a rectangular grid of 9-16 observations across the field of view. Calibrating this observation with the uncorrected flats and fitting a plane to the fluxes of the star will produce a

[2]http://www.eso.org/sci/facilities/lasilla/instruments/sofi/tools.html

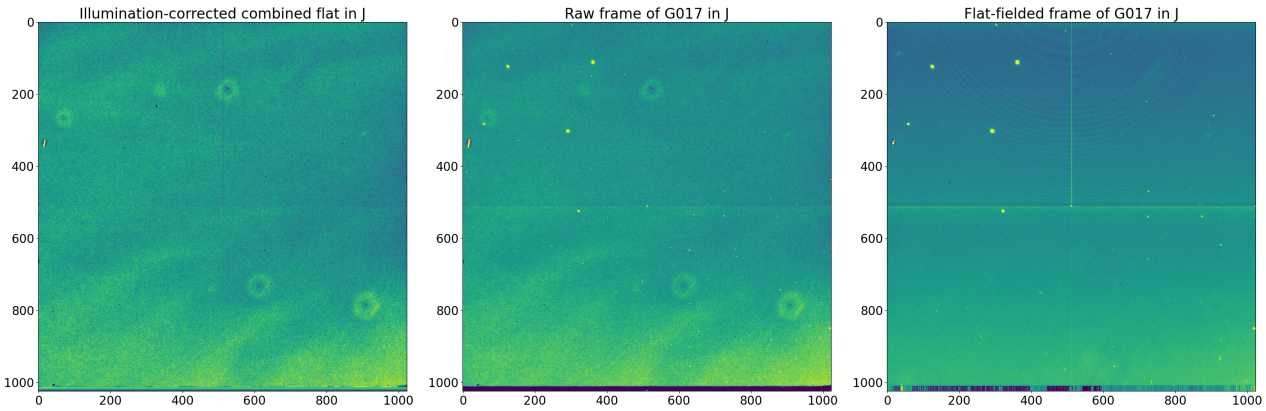


Figure 3.2: An example of the flat-field correction. The images are scaled with counts, with black to white as faint to bright. The x and y axes show the pixel column and row numbers. **Left**: Illumination-corrected combined flat-field image in *J* band. **Middle**: The raw image of one exposure of G017 in *J*. **Right**: The same frame of G017 shown in the middle panel, with flat-field correction applied. It can be clearly seen that features seen in the flat field image are removed from the raw frame after the correction. It should be noted that a ripple pattern can be seen in the flat-fielded image, the actual variations of pixel values caused by this pattern are insignificant and will be further removed after combining jittered images.

normalised surface of illumination. Such illumination correction surfaces for all broad band filters are prepared by the observatory and are available in the SOFI archive. Multiplying the illumination correction surfaces by the combined flat field images in the corresponding filter bands yields the final illumination-corrected flats. Using the FLAT_CORRECT method in the CCDPROC package, we applied flat-field correction to every raw exposure frame with the final flats. Figure 3.2 demonstrates the effect of the flat-field correction. It should be noted that a ripple pattern can be seen in the flat-fielded image, the origin of this pattern is not clear, but the actual variations of pixel values caused by this pattern are insignificant (at the level of $\sim$ 10 ADUs) and will be further removed after combining jittered images.

The next step of the data reduction is removing bad pixels from every image frame. Bad pixels generally consist of: (a) hot pixels, which are pixels with much larger dark current than others and their responses are usually not proportional to exposure time; (b) dead pixels, which respond much less to light than others or have no response to light due to defects on the detector chip; (c) frame pixels, which are masked by the edges of the focal plane field mask; (d) pixels hit by cosmic rays during integration, which will produce hot pixels with varying positions in each exposure. The SOFI detector array only has $\sim$ 0.1% of bad pixels, effectively $\sim$ 1000 bad pixels. The observatory has measured and produced a bad pixel mask which is available on the SOFI archive. This archival mask includes

all bad pixels described above except those induced by cosmic rays. We run the CCDMASK method in the CCDPROC package on two flat frames to detect any extra bad pixels that are not included in the archival mask. The mask produced by CCDMASK is then combined with the archival mask. To detect and remove cosmic rays in our science images, we use the COSMICRAY_LACOSMIC method in CCDPROC. As its name has suggested, this method use the LACosmic method (Van Dokkum, 2001) to detect cosmic rays, and CCDPROC provides a wrapper of another Astropy affiliated package Astro-SCRAPPY (McCully et al., 2018) to implement it. For each science exposure, cosmic rays are detected and summarised into a pixel mask, which is then also merged into the combined bad pixel mask to create a unique final mask for each science frame. We then utilise the IRAF FIXPIX task to incorporate science images and their corresponding bad pixel mask to remove all the bad pixels. Figure 3.3 shows a final bad pixel mask as an example.

One of the most critical step in our data reduction is the sky background subtraction, as it imposes strong influences on the final photometry. There are two basic strategies for sky subtraction of jittered observations. The first one is to create a single sky image, then subtract it from each science frame. The sky image can be obtained by scaling a certain number of images that were taken close in time to a common median and combining them into one image. This sky image should be scaled to have the same median as the science image from which it will be subtracted, and at last carry out the subtraction. The number of images used to create the sky background can be freely chosen as long as it gives a reasonable sky image. The second sky subtraction strategy is, for a series of jittered images taken on a target, subtract the median (or average) of the previous few and the next few (usually 1-3) frames from each image. This strategy is effectively a special case of the first one. In this case the sky image is constructed from the previous few and the next few frames of each science frame. The benefit of this method is that the sky image created is a more accurate representation of the sky background of each image, and it also takes into account sky variation on a short time scale. However, since the sky image is constructed only from images of the same area of the sky, there will be negative residuals of astronomical objects in the images remaining on the background-subtracted images. If the observed field is crowded, these negative residuals will cause serious problems for the final science products. Given that the number of sources in our observed fields is not small, plus one of the goals of this work is to obtain accurate photometry of observed sources, we choose to use the first sky background strategy

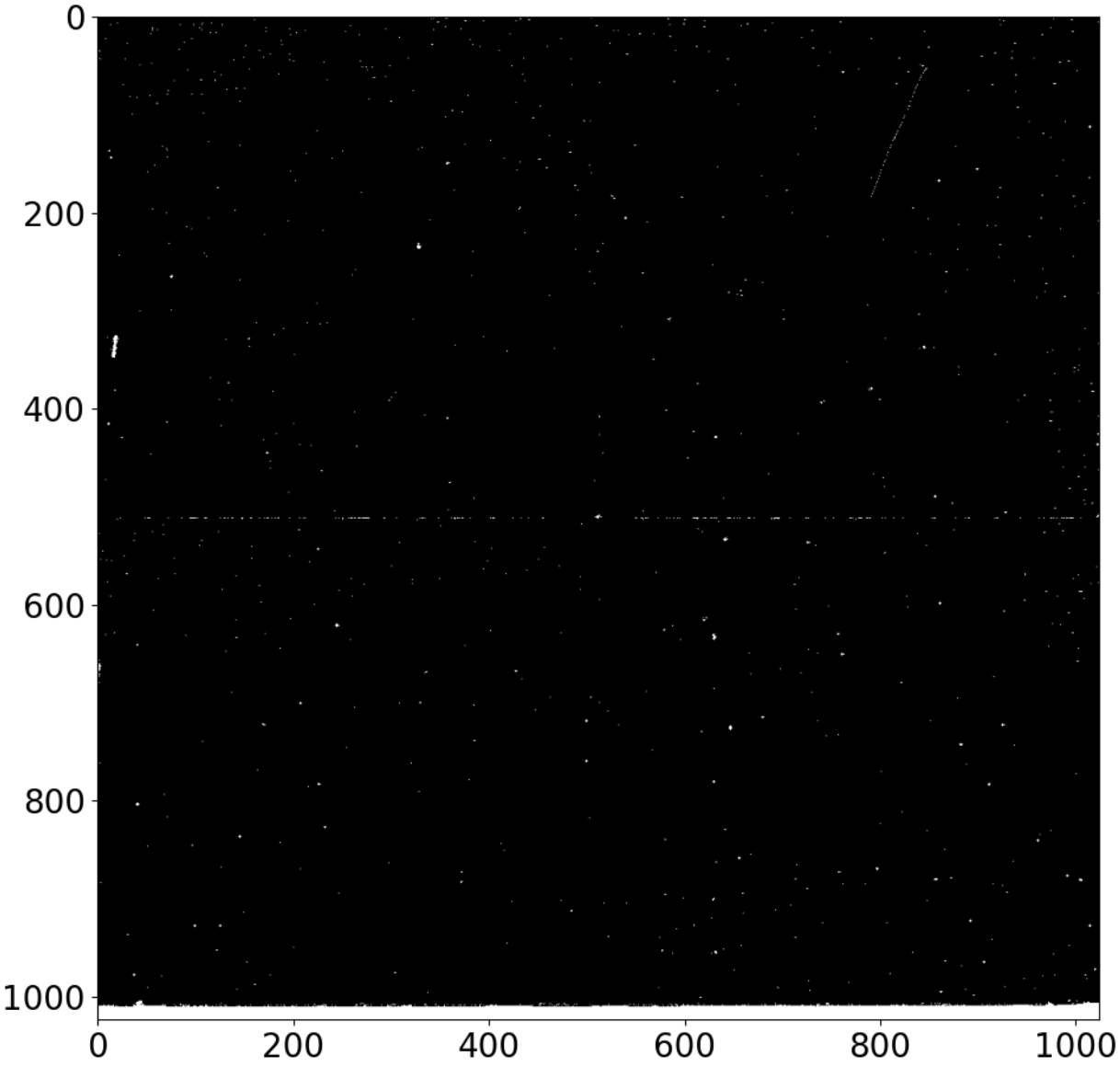


Figure 3.3: An example bad pixel mask used for the same frame of G017 in J shown in Figure 3.2, with bad pixels shown as white. The x and y axes show the pixel column and row numbers.

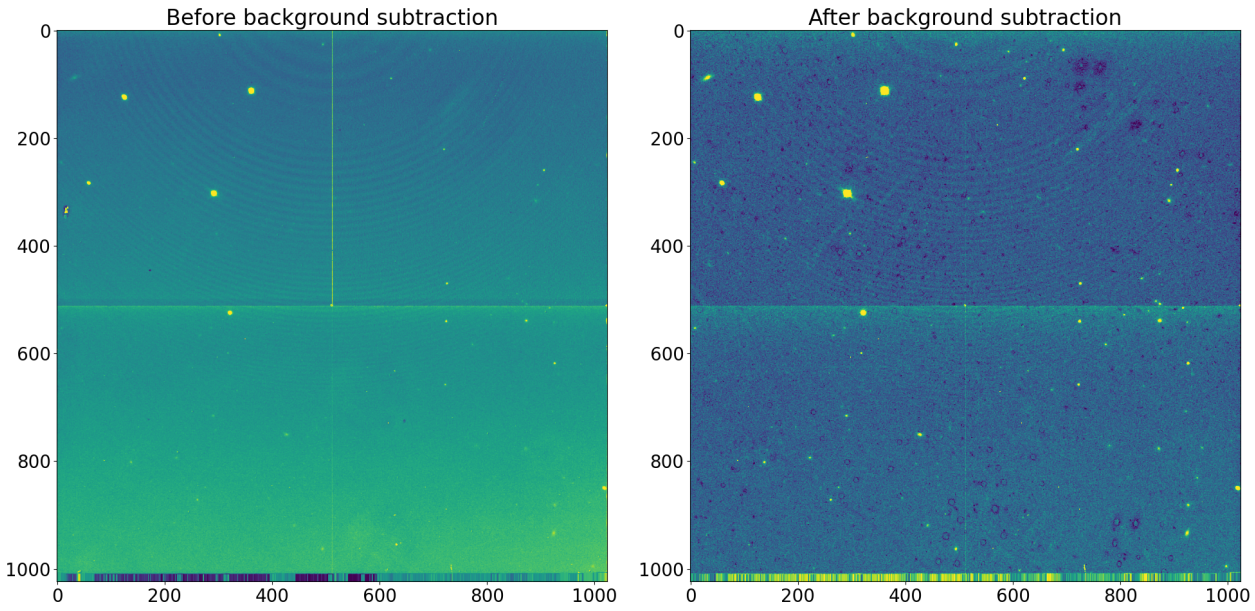


Figure 3.4: An example of background subtraction. The images are scaled with counts, with black to white as faint to bright. The x and y axes show the pixel column and row numbers. **Left**: The same flat-fielded frame of G017 in J shown in Figure 3.2, with bad pixels removed. **Right**: The same frame of G017 shown in the left panel, with sky background subtracted. The background level is significantly reduced and the SNR of all sources increases.

for our data reduction. For each night's observations, all flat-fielded, bad-pixel-removed images in each filter band are median combined to a master sky image, using IRAF task IMCOMBINE with pixels whose values are more than 3σ from the median being rejected. The corresponding master sky image is scaled to each image's median and then removed from that image using CCDPROC. Figure 3.4 shows an example of this background subtraction. It can be seen that the sky subtraction has significantly lowered the background level and increased the signal-to-noise ratios (SNR) of observed sources. Note that there are some ring-like structures seen in the background-subtracted image. This is due to the sub-structures in the master sky image used for the subtraction. Similar to the ripple pattern seen in the flat-fielded image, the effects of these rings are trivial (at the level of $\sim 10 - 30$ ADUs) and can be further reduced by image combination later.

After the sky background is subtracted from all images, the final step of the data reduction is to align and combine all the jittered images for each target. For image alignment, we make use of the Python package ASTROALIGN (Beroiz et al., 2020) which aligns two astronomical images by finding 3 similar sources above a user-defined detection threshold to form triangles in both images and estimating the affine transformation between them. All background-subtracted images of each target in each filter

band are aligned using ASTROALIGN to the first of them, with a low detection threshold of 2σ as the SNRs of some faint sources may be very low in individual frames. The images of a few targets in certain filter bands which cannot be automatically aligned by ASTROALIGN, because of an inadequate number of sources detected above the 2σ detection threshold. These include J band images of G252, H images of G015, G257 and G252, and K_s images of G222 G260 and G261. For each set of these images, we visually pick up a couple of sources and manually measure the pixel offsets of these sources' positions between the first exposure and the others, then align these images to the first one using the measured pixel offsets with IRAF task IMSHIFT. The interpolation parameter in IMSHIFT is set to linear to conserve the flux during the shifting, and pixels with no value created by the shift are set to -10^6. Lastly, for all targets, the aligned images are median combined using the IRAF task IMCOMBINE with a low-boundary threshold of -10^5 to remove pixels which are set to a large negative value during shifting. The combined images are used as the final science images. We note that the last 5 K_s frames of G222, in time order, are not used for image combination as they are fully saturated probably because they were observed near sunrise and ruined by the dawn light. Figure 3.5 displays the final science images of all targets in every observed filter band. For images aligned by ASTROALIGN, clear margins can be seen at the edges of these images. This is because we cannot assign a large negative value to pixels created by the shift with ASTROALIGN, as we have done with IMSHIFT, therefore the pixels at the edges are not excluded by IMCOMBINE. Nevertheless, these margins have no effect on the photometry and can be safely ignored, and the effective area of target fields excluding these margins is $\sim 4.6' \times 4.6'$.

3.4 Source extraction & Photometry calibration

With SOFI images of all targets reduced, we extract sources from their images using SOURCE EXTRACTOR (Bertin and Arnouts, 1996). Table 3.2 shows the key SOURCE EXTRACTOR parameters used for the extraction. These parameters are chosen to extract sources with signal-to-noise ratios larger than 5σ. The source lists created by SOURCE EXTRACTOR consist of the J2000 coordinates of each source, the integrated count and related errors measured by summing pixel values within an adaptively scaled aperture (Kron, 1980), and the associated Kron radius of the adaptive aperture applied for each

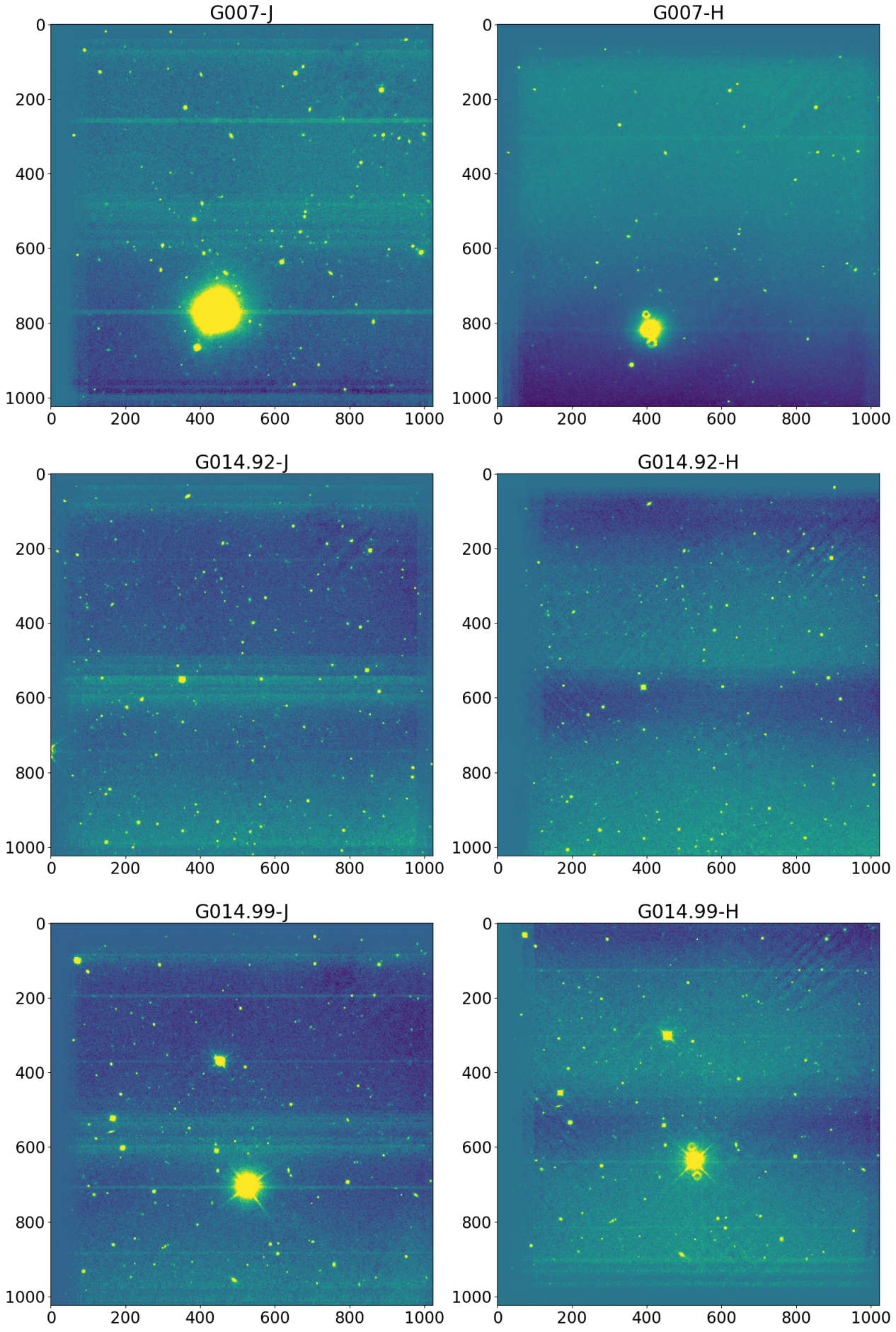


Figure 3.5: The final science images of all targets in every observed filter band, with the corresponding target and filter name shown in the title of each panel. The images are scaled with counts, with black to white as faint to bright. The x and y axes show the pixel column and row numbers.

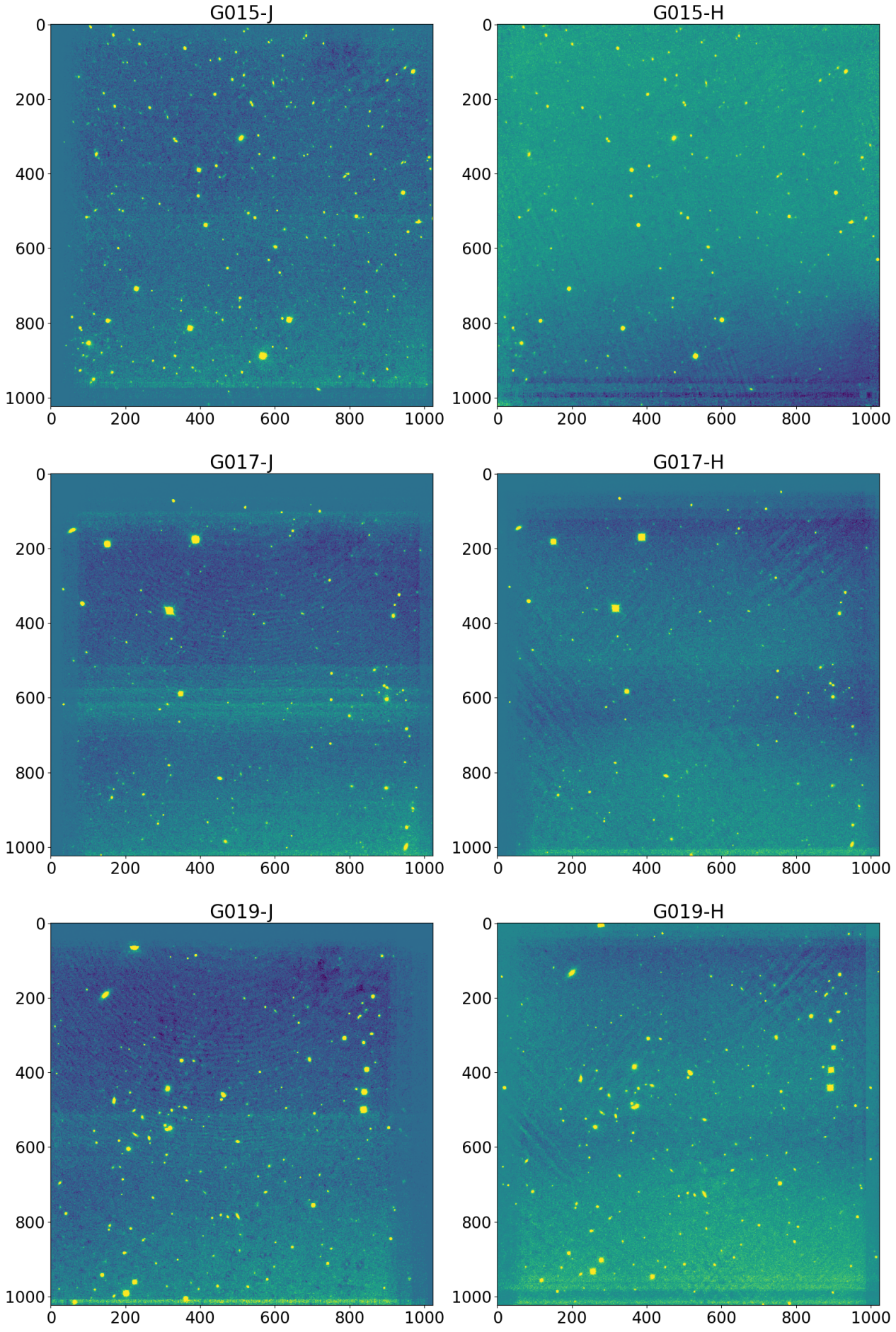


Figure 3.5 continued.

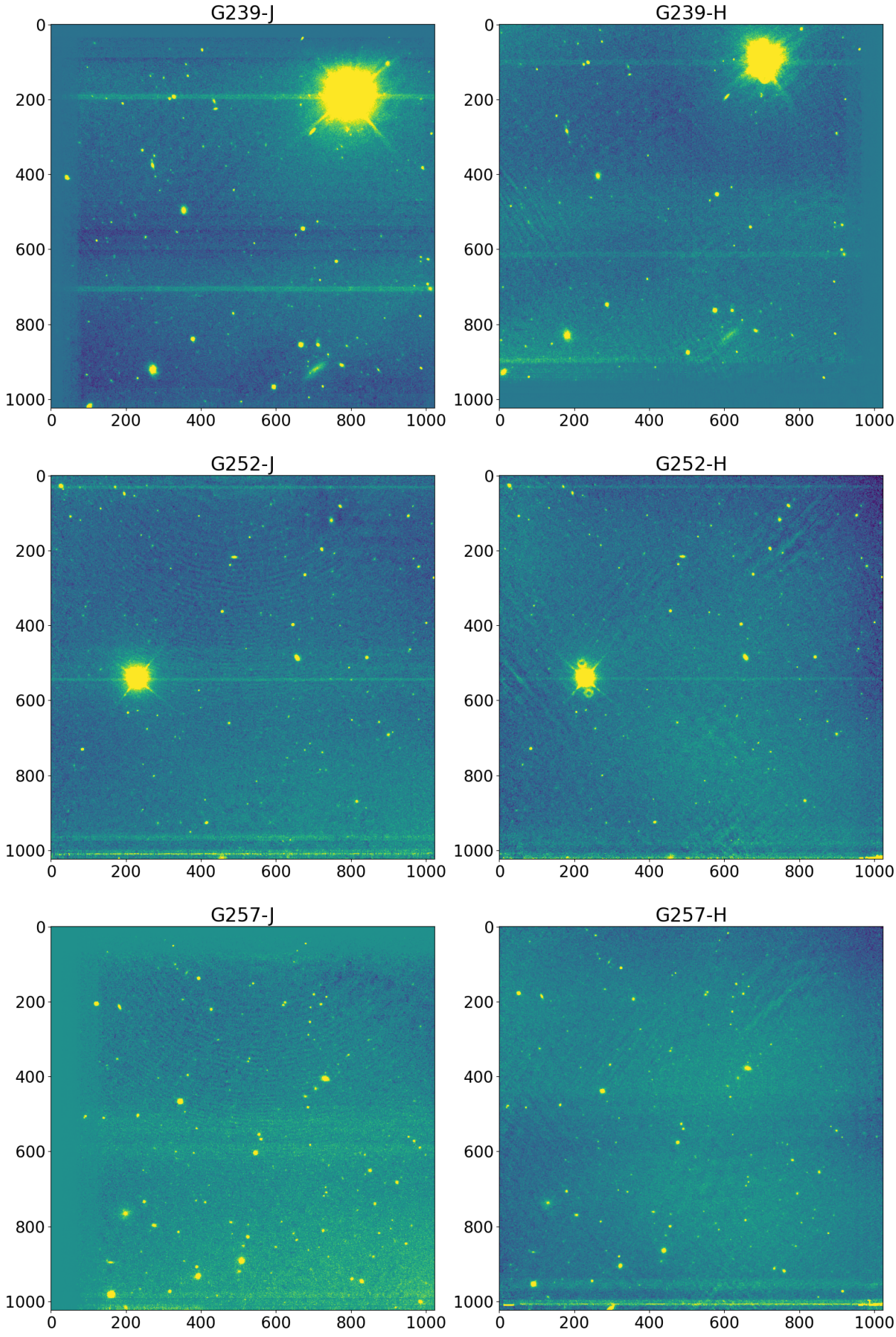


Figure 3.5 continued.

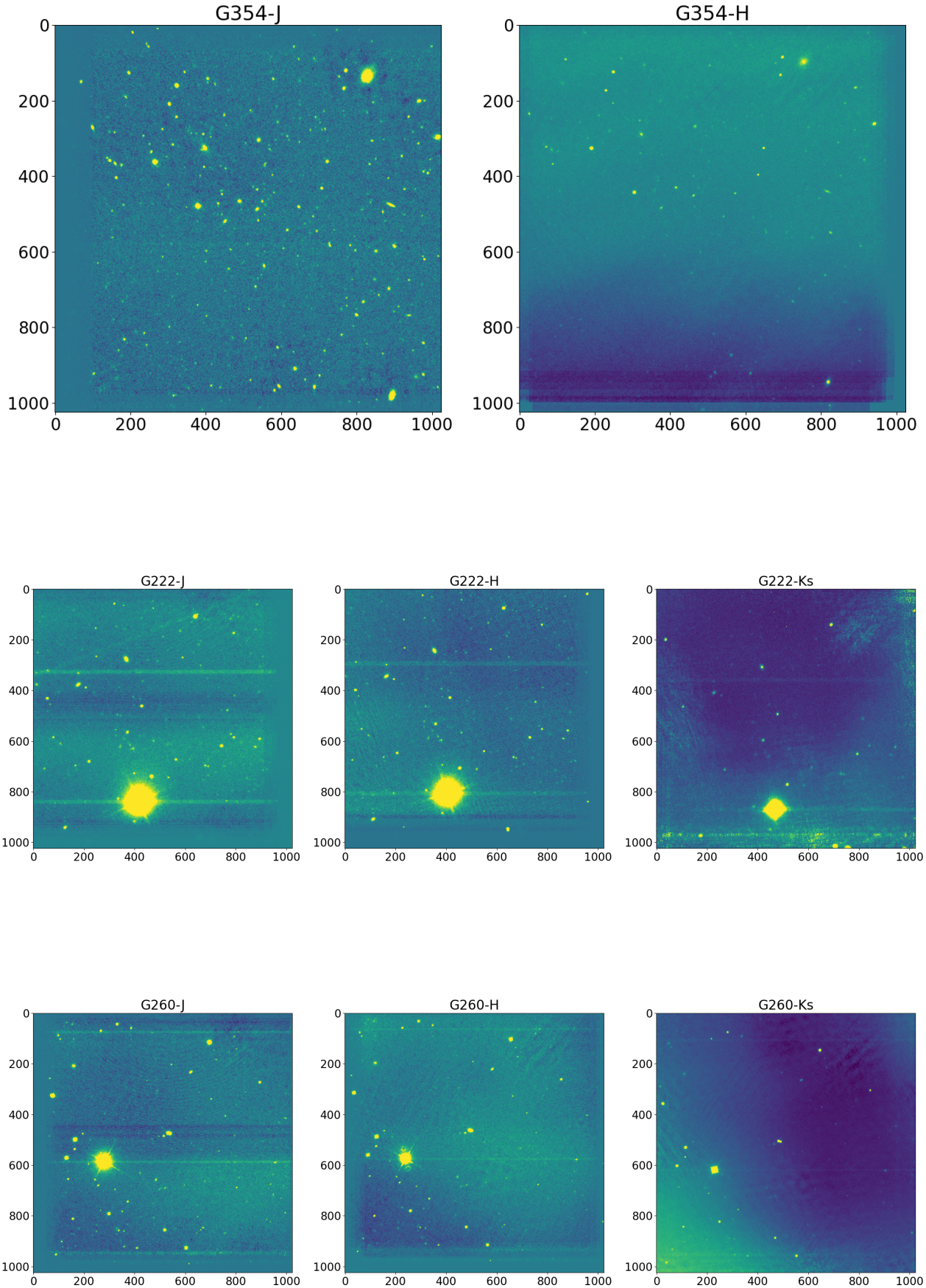


Figure 3.5 continued.

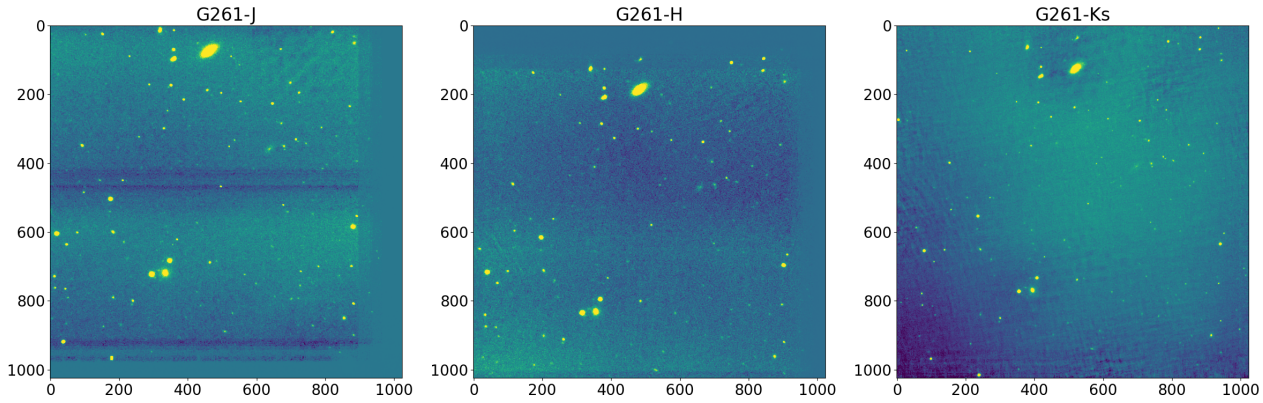


Figure 3.5 continued.

source. In rare cases where nearby sources contribute over 10% to the photometry of a source, the corrected isophotal flux is used for this source. For each target, the extracted catalogues in J and H are cross-matched with a matching radius of $1''$ to exclude spurious detections, the K_s catalogue is then also matched to this $J - H$ matched catalogue to form a final combined catalogue containing photometry in all observed bands. In this way, source detections are based on J and H images only as the K_s image depth is much shallower than the other two bands, due to the high background level at this wavelength, so some of the faintest sources may not be seen in K_s images. Sources not detected in K_s are flagged and have their K_s photometry as blank in the final catalogues. There are in total 2238 sources detected in all target fields and 317 of them detected in K_s

Table 3.2: Key SOURCE EXTRACTOR parameters used to extract sources from our SOFI images. Default values are used for all the other parameters.

Parameter	Value
DETECT_TYPE	CCD
DETECT_MINAREA	5
THRESH_TYPE	RELATIVE
DETECT_THRESH	2.236
ANALYSIS_THRESH	2.236
FILTER	N

To make the extracted signal of an object comparable to parameters in other datasets, we need to convert the counts into magnitudes in a photometric system. Firstly, the total counts of an object are converted into an instrumental magnitude m_{ins} following the equation:

$$m_{ins} = -2.5 log_{10} \frac{F}{t_{exp}} \tag{3.5}$$

where F is the total raw counts of the object in ADUs (or counts) and t_{exp} is the exposure time in seconds. This equation is a special case of the more general magnitude equation $m_1 - m_2 = -2.5log_{10}\frac{F_1}{F_2}$ by assuming the second object has zero magnitude and a flux of unity. The instrumental magnitude calculated in this way is dependent on the telescope instrumentation and thus it is meaningless to compare two such magnitudes obtained from different instruments. Next, the instrumental magnitude which is measured from the ground is corrected for the atmospheric extinction. To reliably compare the magnitudes of two objects on the sky, this is an essential correction as the amount of light absorbed by the atmosphere depends on the angle of the observations above the horizon, which are likely to be different for any two objects. The corrected instrumental magnitude m_{ins0} is derived as:

$$m_{ins0} = m_{ins} - \chi k \tag{3.6}$$

where k is the extinction coefficient of the filter used for the observation, χ is the airmass of the observation which is a measure of the light path length through the atmosphere. For our SOFI observations, the extinction coefficients k for J, H, K_s are 0.07 ± 0.02, 0.02 ± 0.01 and 0.05 ± 0.01 respectively, as measured by the La Silla observatory and available in the SOFI archive. The airmass of each target's observations is automatically recorded during observation and stored in the headers of every image. The final step is to turn the atmospheric extinction corrected instrumental magnitude into a magnitude in a standard photometric system. To achieve this, it is necessary to observe some photometric standard stars with known magnitudes in a photometric system, then derive the zero point (ZP) magnitude for this photometric system from the standard star observations. Following the same step outlined above, the extinction corrected instrumental magnitude of a standard star $m_{ins0std}$ should firstly be calculated, then the ZP magnitude m_{zp} can be derived as:

$$m_{zp} = m_{std} - m_{ins0std} \tag{3.7}$$

where m_{std} is the catalogue magnitude of this standard star. The magnitude calibrated to the photometric system of any observed objects can then be calculated by simply adding the ZP magnitude to their m_{ins0}:

$$m = m_{ins0} + m_{zp} \tag{3.8}$$

It should be noted that each filter will have a different ZP magnitude. Although we have conducted standard star observations at the start of each observing night, the sky conditions varied significantly during the observing nights. Therefore, the standard star observations do not accurately represent the sky conditions of our observations, and cannot provide ZPs for accurate photometric calibration, especially for our K_s observations which were carried out at the end of each observing night. We thus use an alternative method of obtaining a reliable ZP, which is to directly derive ZPs from the Two Micron All Sky Survey (2MASS; Cutri et al. (2003)) catalogue. The 2MASS ZP magnitude can be easily calculated by:

$$m_{zp} = m_{2MASS} - m_{ins0} \tag{3.9}$$

where m_{2MASS} is the 2MASS catalogue magnitude of a source and m_{ins0} is the atmospheric extinction corrected instrumental magnitude of the same source extracted from our SOFI images. SOFI sources in all target fields are cross-matched to the 2MASS catalogue with a matching radius of $2''$. There are 29 2MASS sources matched to SOFI sources in G222, G260 and G261 which are the three target fields with K_s observations. For each of these 29 matched 2MASS sources, we calculate the ZP magnitude from its catalogue magnitude, then derive the apparent magnitudes of all 29 sources and calculate the average magnitude difference between the derived magnitudes and the catalogue magnitudes, the results are shown in Figure 3.6. As can be seen, except for a few outliers, most ZPs have low average magnitude differences of $\sim 0.1 - 0.2mag$ with uncertainties of $\sim 0.1 - 0.2mag$ in all bands. This means that using ZPs calculated from 2MASS sources, accurate apparent magnitude can be reliably calculated. We calculate the average ZP magnitude of all 2MASS ZPs, then exclude ZPs that are 2σ away from the mean. We iterate this process to the point where no further ZP is excluded, the clipping boundary is shown as black dashed-dotted line in Figure 3.6. The average of the remaining ZPs is then used to calculate the apparent Vega magnitudes of all sources in all observed target fields, which are recorded in the final catalogues of all targets. Additionally, AB magnitudes of all sources are recorded in the final catalogues as well, which are converted from sources' Vega magnitude by adding an offset of 0.916, 1.366 and $1.827mag$ to the Vega magnitudes for J, H and K_s respectively (Hewett et al., 2006; González-Fernández et al., 2018). All magnitudes refer to the AB magnitudes hereafter as AB magnitudes are currently more widely used. Figure 3.7 shows the magnitude distributions of sources in all target fields in the three filter bands. The faintest magnitudes detected in J, H and K_s

are 23.07 ± 0.21, 23.12 ± 0.33 and 21.17 ± 0.13 respectively. This means that we have exceeded our proposed image depths of $22.4mag$ in both J and H, but have not reached the desired depth in K_s probably because of the non-ideal weather conditions and dawn light contamination during the K_s imaging.

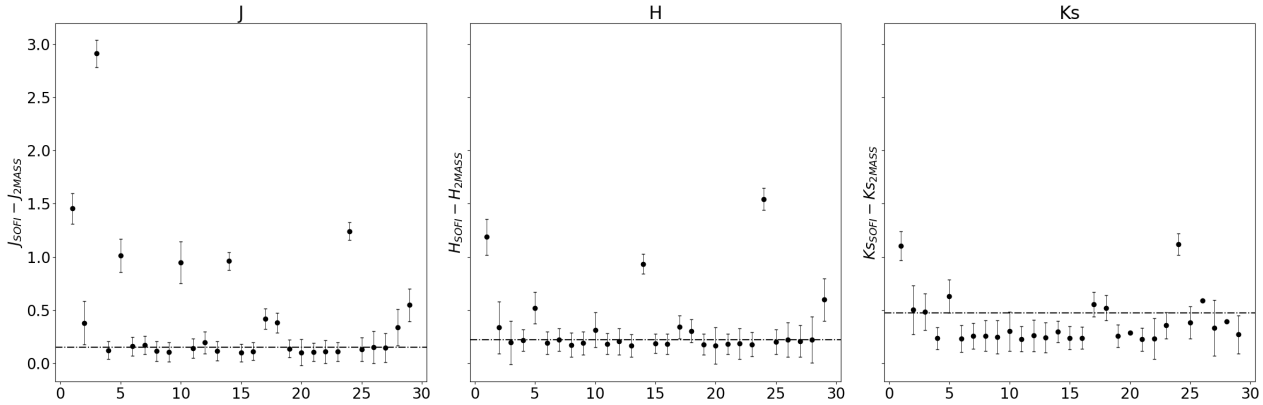


Figure 3.6: The average magnitude differences for ZP derived from 2MASS catalogue magnitudes, in J, H and K_s from left to right. The magnitude difference in the corresponding band is on the y-axis, and the x-axis shows the arbitrary indices of the 29 2MASS sources used to calculate ZPs. As can be seen, except for a few outliers, most ZPs result in magnitude differences of $\sim 0.1 - 0.2mag$ with uncertainties of $\sim 0.1 - 0.2mag$ in all bands. The 2σ clipping boundary is shown as a black dashed-dotted line in each panel.

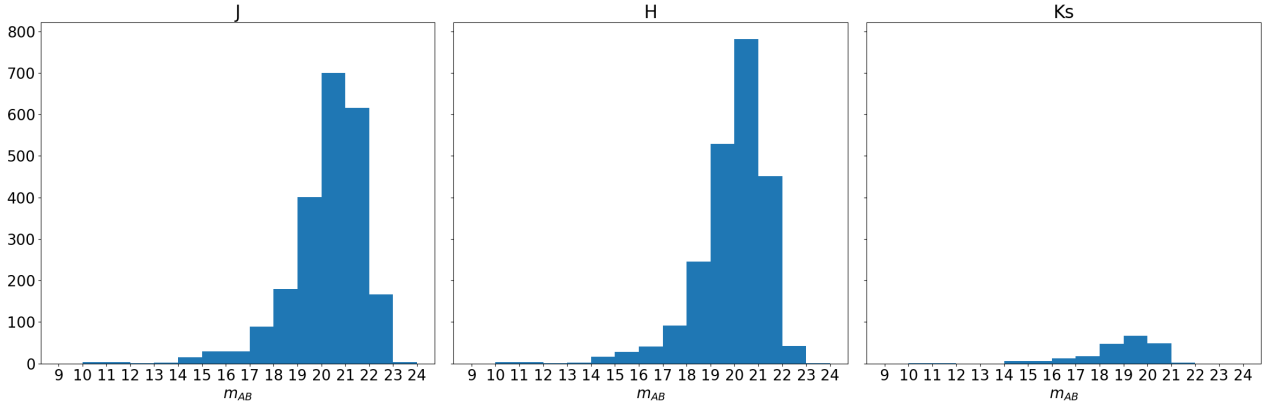


Figure 3.7: The AB magnitude distributions of sources in all target fields, in J, H and K_s band from left to right.

Overall, we have extracted 2238 sources in all target fields and 317 of them have K_s detections. ZPs are derived from catalogued 2MASS sources to calculate the magnitudes of all sources. We have 5σ detection down to $J = 23.07 \pm 0.21$, $H = 23.12 \pm 0.33$ and $K_s = 21.17 \pm 0.13$.

3.5 Complementary Optical/NIR data

The SOFI data alone is not enough to strongly constrain the properties of our candidate protoclusters, so complementary data at other wavelength are required to investigate the nature of our targets. In the absence of deep, complete spectroscopic data which will provide precise redshift information, we make use of public deep imaging surveys in the optical/NIR to characterise our targets. In this section, we briefly describe the complementary data used in this chapter.

3.5.1 DESI

The Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Survey (Dey et al., 2019) is a combination of three public projects, the Dark Energy Camera Legacy Survey (DECaLS; Dey et al. 2019), the Beijing–Arizona Sky Survey (BASS; Zou et al. 2017) and the Mayall z-band Legacy Survey (MzLS; Dey et al. 2019). These three projects jointly image $\simeq 14,000\ deg^2$ of the extragalactic sky visible from the northern hemisphere in three optical bands (g, r, and z) using telescopes at the Kitt Peak National Observatory and the Cerro Tololo Inter-American Observatory (CTIO) in northern Chile. The Legacy Survey catalogue includes photometry from the grz optical bands with the addition of forced-photometry in four mid-infrared bands (at 3.4, 4.6, 12, and 22 μm) observed by the Wide-field Infrared Survey Explorer (*WISE*; Wright et al. 2010). In this chapter, we use the latest data release 10 (DR10) of the Legacy Surveys. In DR10, DESI additionally includes new imaging data in the southern sky (declination $\leq 32.375°$), collected by the Dark Energy Camera (DECam; Flaugher et al. 2015) which is a 570 Megapixel camera with a 3 deg^2 field-of-view installed at the prime focus of the Blanco 4-m telescope at the CTIO. This newly-added DECam data provides i band photometry on top of g, r, z band observations, with median 5σ point source AB magnitude limits of $\sim 25.0, 24.5, 24.3, 23.7$ in g, r, i, z respectively.

3.5.2 KiDS & VIKING

The Kilo-Degree Survey (KiDS; de Jong et al. 2013) is an on-going wide-field optical survey in four broad band filters u, g, r, i. It is carried out with the 268-million pixel camera OmegaCAM (Kuijken et al., 2011) mounted at the VLT Survey Telescope (VST; Capaccioli and Schipani 2011) at ESO Paranal observatory in Chile. KiDS is designed as a cosmology survey to measure the effect of weak lensing. The latest public-available data release of KiDS is the fourth data release (DR4; Kuijken et al. 2019), covering $\sim$ 1350 square degree of the sky. In particular, KiDS DR4 incorporates data from the VISTA Kilo-degree Infrared Galaxy survey (VIKING; Edge et al. (2013)) which observes the same patch of the sky as KiDS in five broad bands, namely Z, Y, J, H, K_s. Together, KiDS DR4 provides a nine-band photometry catalogue with mean 5σ limiting magnitudes of $u = 24.23$, $g = 25.12$, $r = 25.02$, $i = 23.68$, $Z = 23.1$, $Y = 22.3$, $J = 22.1$, $H = 21.5$ and $K_s = 21.2$. With the nine-band photometry, the KiDS DR4 catalogue also includes photometric redshifts estimated by the Bayesian Photometric redshift code (BPZ; Benitez 2000).

3.5.3 DES

The Dark Energy Survey (DES; Dark Energy Survey Collaboration: Fermilab and Flaugher 2005) is a wide-field optical/NIR imaging survey in the g, r, i, z, Y bands, covering $\sim$ 5000 deg^2 of the sky near the south Galactic pole. DES is designed to study cosmic acceleration and the nature of dark energy by measuring four parameters: weak gravitational lensing, the distances to Type Ia supernovae, the large-scale clustering of galaxies including baryon acoustic oscillation, and galaxy cluster counts. DES uses the 570 Megapixel DECam mounted at the Blanco 4-m telescope at CTIO. In 2018, the DES collaboration team published the first major data release of DES (DR1) (Abbott et al., 2018) which contained data taken in the first three years of DES science operations. The latest second DES data release (DR2; Abbott et al. 2021) contains all processed data of DES including reprocessed DR1 data. The median magnitude limits for a 1.″95 diameter aperture at signal-to-noise ratio = 10 are $g = 24.7$, $r = 24.4$, $i = 23.8$, $z = 23.1$, and $Y = 21.7$ mag.

3.5.4 SHARKS

The Southern H-ATLAS Regions in K_s-band Survey (SHARKS; Dannerbauer et al. 2022) is a deep K_s band imaging survey using the VISTA InfraRed CAMera (VIRCAM, Dalton et al. 2006) mounted at the 4 metre NIR-optimised Visible and Infrared Survey Telescope for Astronomy (VISTA, Sutherland et al. 2015). It is designed to provide deep K_s data to detect $z > 1$ extragalactic sources discovered by the *Herschel*-ATLAS (H-ATLAS, Eales et al. 2010). It covers $\sim 300\ deg^2$ of sky in large parts of the SGP, GAMA-12h and GAMA-15h fields, reaching a 5σ magnitude limit of $K_s = 22.7$ mag for a $2''$ aperture and a median seeing of $1''$. The first public data release of SHARKS (Dannerbauer et al., 2022) contains 5% of the total collected data, covering an area of $19.24\ deg^2$.

3.5.5 2dFGRS

The 2-degree Field Galaxy Redshift Survey (2dFGRS or 2dF; Colless et al. 2001) uses the 2dF multi-fibre spectrograph on the Anglo-Australian Telescope (Lewis et al., 2002), which allows simultaneous observations over a 2 degree diameter field. 2dFGRS covers two large areas, one is around the SGP spanning $80° \times 15°$, the other is close to the celestial equator in the northern Galactic hemisphere spanning $75° \times 10°$. 2dFGRS has $\sim 250,000$ target galaxies which are selected to have extinction-corrected Vega magnitude $b_J \leq 19.45$. Redshifts are measured from spectra covering $3600 - 8000$ Å reaching a median redshift depth of $z = 0.11$.

3.5.6 Cross identification

Table 3.3 summarises the photometry band data available for each of our target fields and the source of the data. It should be noted that WISE photometry is not included despite that it is provided by the DESI catalogue, because we find that adding the WISE photometry will significantly lower the quality of estimated photometric redshifts of our sources (see Section 3.8 for more details). Of the 13 targets, 8 have nine-band photometry of $ugrizYJHK_s$, 4 have $grizYJHK_s$ eight-band photometry, and only G015 has seven-band $grizJHK_s$ photometry as it is not observed by either KiDS or DES.

To build the aforementioned multi-band photometry, we positionally match our SOFI sources to the corresponding optical/NIR surveys described above. Since all of our target fields are covered by DESI, we choose to use DESI's $griz$ photometry over KiDS or DES to make the final photometry as homogenous as possible. The search radius used for the cross-identification has a significant impact to our results, as a radius too small will miss some true IDs while a search radius too large will result in false identifications due to chance alignments. We examine the matching rate of true SOFI-DESI identification and random chance alignments as a function of matching radius, shown in Figure 3.8. The percentage of random identifications is calculated by creating a simulated DESI catalogue covering one of the SOFI fields which reproduces the true DESI number counts, then this simulated catalogue and the SOFI catalogue of this field is matched. Any identification found in this case is thus a result of chance alignments. As can be seen in Figure 3.8, the percentage of true IDs rapidly increases to $\sim 90\%$ at the matching radius of about $0.6''$, then starts to increase with the matching radius at a much slower rate. In the meantime, the number of random IDs is about 0 at a matching radius $< 2''$, then gradually rises as the matching radius goes up. This indicates that using a search radius larger than $2''$ will include some false identifications, and using a search radius smaller than $1''$ will miss part of true IDs. As the percentages of true IDs at $1''$ and $2''$ search radius are nearly the same, we eventually choose $1.5''$ as our matching radius to avoid any possible random ID while including most true IDs.

For the 2238 SOFI sources, 1890 of them remain after cross-identification, resulting in an overall matching rate of $\sim 84.5\%$. We visually examine the SOFI sources which are not identified in the ancillary data, finding that the majority of these sources are true SOFI sources with no apparent optical counterpart. An example of such a source is shown in Figure 3.9. This source has an apparent magnitude of $J = 20.36 \pm 0.10$ and $H = 19.85 \pm 0.10$. The non-detection of these SOFI J-selected sources means that they are redder than other sources, implying that they lie at higher redshifts or consist of older stellar population. Nonetheless, the lack of optical photometry of these sources limits our ability to assess their photometric redshifts and other properties. Deeper optical/NIR observations are needed to better understand them.

We further assess the quality of our cross-identification by calculating the Poisson probability of an

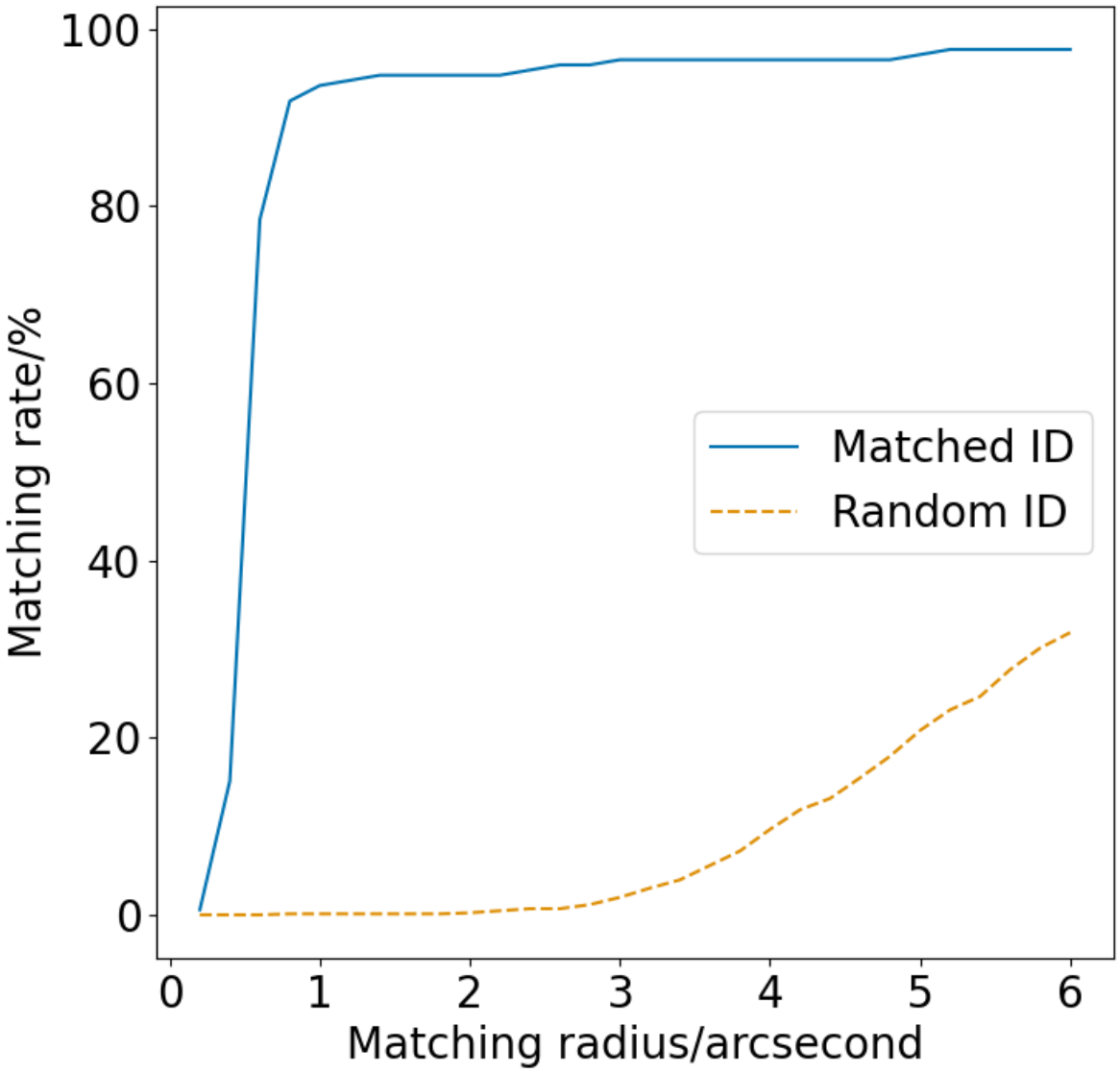


Figure 3.8: DESI-SOFI matching percentage as a function of matching radius. The blue solid line shows the percentage of SOFI sources with real DESI counterparts, the orange dashed-line shows the percentage of SOFI sources with random simulated DESI counterparts. As can be seen, while the matching rate becomes > 90% at matching radii larger than 1″, at a matching radius of ∼ 2″ the number of random IDs starts to increase. This indicates that using a matching radius smaller than 1″ will miss a large fraction of true cross-identifications, while using a matching radius larger than 2″ will include more chance alignments. Therefore, we choose to use a matching radius of 1.5″ to minimise the number of random IDs and maximise the number of true IDs.

Table 3.3: Summary of the photometry available for each target and the source of the data. Each row of the table shows the name of the target, the photometric bands available, and public surveys as the source of the photometry with photometric bands provided by them in brackets. Note that the "SOFI" in the Data source column refers to the SOFI data collected and reduced in this chapter.

Name	Photometric band	Data source
G007	$ugrizYJHK_s$	KiDS (u, Y), DESI $(griz)$, SOFI (J, H) SHARKS (K_s)
G014.92	$ugrizYJHK_s$	KiDS (u, Y), DESI $(griz)$, SOFI (J, H), VIKING (K_s)
G014.99	$grizYJHK_s$	DESI $(griz)$, SOFI (J, H), VIKING (Y, K_s)
G015	$grizJHK_s$	DESI $(griz)$, SOFI (J, H), SHARKS (K_s)
G017	$ugrizYJHK_s$	KiDS (u, Y), DESI $(griz)$, SOFI (J, H) SHARKS (K_s)
G019	$ugrizYJHK_s$	KiDS (u, Y), DESI $(griz)$, SOFI (J, H), VIKING (K_s)
G222	$grizYJHK_s$	DESI $(griz)$, DES (Y), SOFI (J, H, K_s)
G239	$grizYJHK_s$	DESI $(griz)$, DES (Y), SOFI (J, H), SHARKS (K_s)
G252	$ugrizYJHK_s$	KiDS (u, Y), DESI $(griz)$, SOFI (J, H) SHARKS (K_s)
G257	$ugrizYJHK_s$	KiDS (u, Y), DESI $(griz)$, SOFI (J, H) SHARKS (K_s)
G260	$ugrizYJHK_s$	KiDS (u, Y), DESI $(griz)$, SOFI (J, H, K_s)
G261	$grizYJHK_s$	DESI $(griz)$, DES (Y), SOFI (J, H, K_s)
G354	$ugrizYJHK_s$	KiDS (u, Y), DESI $(griz)$, SOFI (J, H) SHARKS (K_s)

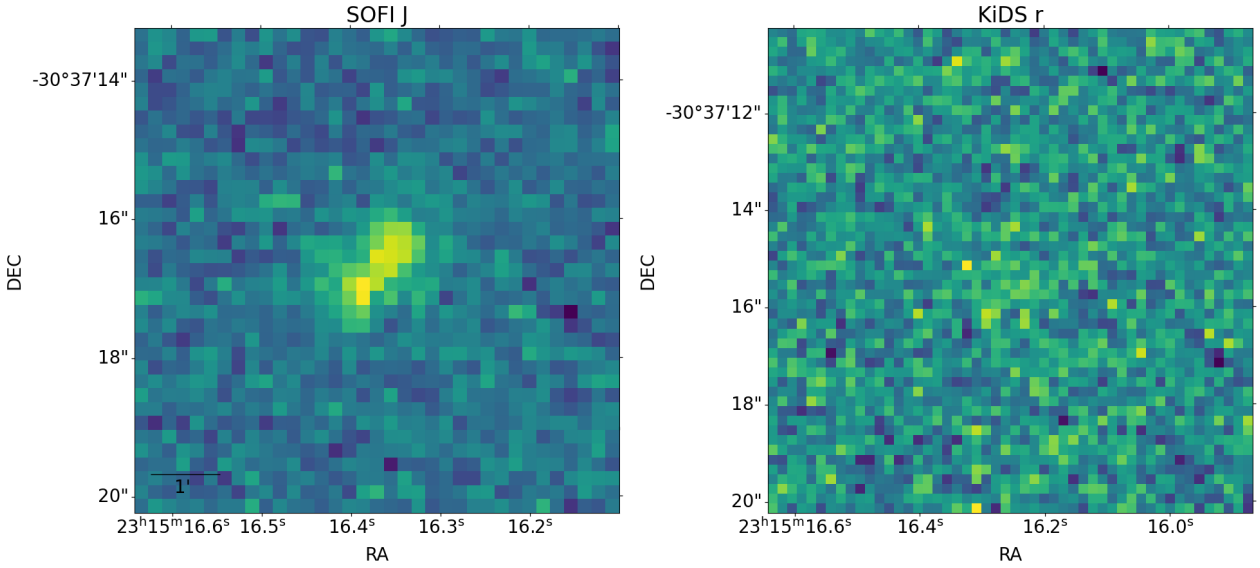


Figure 3.9: An example of a SOFI J-band detected source with no apparent r-band counterpart. The left panel shows the SOFI J cutout image of an object in G017, the right panel shows the KiDS r cutout of the same object. A scale bar of $1''$ is shown in the bottom left of the left panel. The apparent magnitude of this source is $J = 20.36 \pm 0.10$ and $H = 19.85 \pm 0.10$.

identification being a chance alignment (e.g., Downes et al. 1986, Biggs et al. 2011). Assuming random distribution of sources in a given wavelength, the expected number of sources in a circle of radius r is:

$$\mu = \pi r^2 n \tag{3.10}$$

where n is the number density of sources. Then the Poisson probability of one random object located in this circle is:

$$P = 1 - e^{-\mu} \tag{3.11}$$

Using the DESI number density calculated from a randomly selected DESI area of $\sim 0.16\ deg^2$ and our search radius of $1.5''$, the Poisson probability that one SOFI source coincides with a random DESI source is 3.72%. Given the very low Poisson probability of chance alignment, our cross-identification is thus unlikely to be significantly contaminated by chance alignments.

3.6 Photometric redshift estimation

With multi-wavelength photometry, we then proceed to estimate the photometric redshifts (photo-z) of all sources in our target fields. Recently, the popular choices of photometric redshift codes are SED template fitting codes, such as BPZ (Benitez, 2000), EAZY (Brammer et al., 2008), BEAGLE (Chevallard and Charlot, 2016) and CIGALE (Boquien et al., 2019). Such codes fit SED templates to user-provided photometric points of sample sources and derive their best-fit SEDs, then infer their physical properties including redshift from the fitted SEDs. Of the four template fitting codes mentioned above, BPZ and EAZY are specifically designed for photometric redshift estimations, while BEAGLE and CIGALE are optimised to derive not only redshift but also other physical properties such as stellar mass and star formation history. This makes BEAGLE and CIGALE much slower than the other two codes. As the primary goal of this chapter is to obtain accurate photometric redshift to investigate the nature of the selected candidate protoclusters, BPZ and EAZY are the more efficient choices for us. Both BPZ and EAZY have been widely used in galaxy studies including large surveys. For example, BPZ is used in KiDS and EAZY is used in the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007; Weaver et al. 2022). There are also studies which specifically investigate the

difference between several photometric redshift codes including both BPZ and EAZY, finding no significant difference between the quality of BPZ and that of EAZY. For example, Brinchmann et al. (2017) compare the performances of BPZ, EAZY and BEAGLE using precise spectroscopic redshifts for 1227 galaxies from the Multi-Unit Spectroscopic Explorer (MUSE; Bacon et al. 2010) integral field spectrograph, down to $F775W = 30\ mag$ at $z < 6$. With the ultra-deep sample, they find little difference in the performances of the three codes. Salmon et al. (2020) examine the bright $6 \leq z \leq 8$ candidate high-redshift galaxies found in the Reionization Lensing Cluster Survey (RELICS; Coe et al. 2019). They use BPZ and EAZY independently to estimate redshifts of their sample, finding that the results of the two codes agree with each other very well for galaxies at $z < 5.5$, while EAZY has a slightly better performance for faint galaxies at $z > 5.5$. Overall, both BPZ and EAZY can provide good quality photometric redshifts, and there is no significant difference in the performance of the two codes at $z < 5.5$. We eventually choose to use EAZY in this chapter, as it has better performance in handling potential $z > 5.5$ sources.

EAZY provides a number of template sets as options. We use the default EAZY_v1.3 galaxy template set as it is demonstrated to give the best performance in general (e.g., Brinchmann et al. 2017). This template set consists of 9 galaxy SED templates including a dusty SED (Bruzual and Charlot, 2003) and the SED of a high-redshift low-metallicity galaxy (Erb et al., 2010). EAZY will then simultaneously combine all templates to fit the photometry of sources in the input catalogue, and calculate their redshift probability distributions, the peaks of which are the estimated redshifts. Using this template set and a built-in apparent K-band magnitude prior, we derive the photometric redshifts of our SOFI detected sources. Additionally, we randomly select a patch of the sky as a blank field as a comparison to our SOFI target fields. The area of this blank field is ~ 3.2 square degree. Since only KiDS DR4 has a comparable number of photometric bands and similar depths in these bands to our SOFI catalogues, we use the KiDS DR4 $ugriZYHJK_s$ photometry to calculate the photometric redshifts of sources in this blank field with the same EAZY settings as mentioned above.

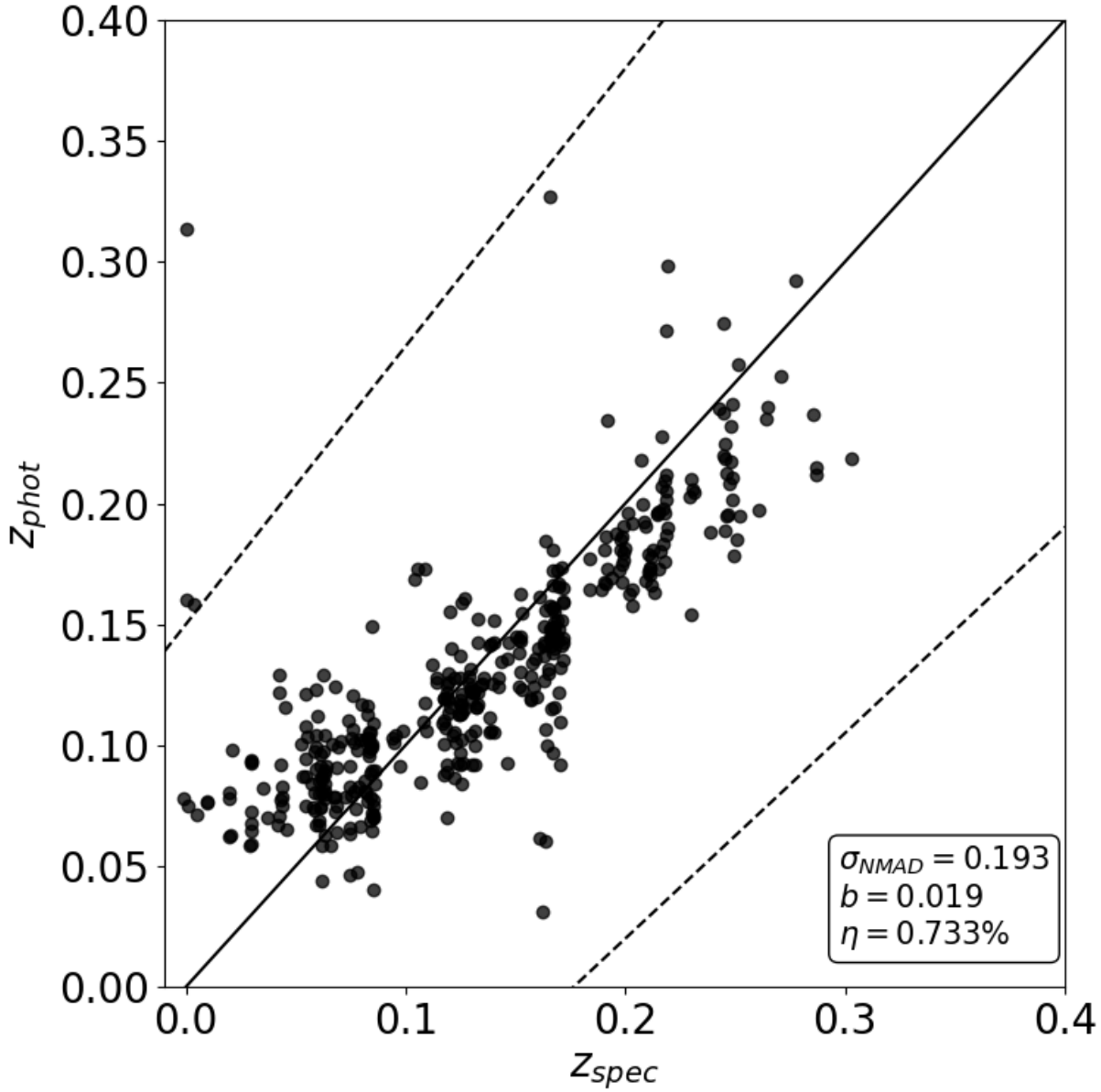


Figure 3.10: Comparison of photo-z and spec-z for sources in the blank field. The blank solid line is the one-to-one line, the two dashed lines indicate $\Delta z/(1+z) = \pm 0.15$. The overall NMAD, bias and the fraction of outliers of are shown in the bottom right panel.

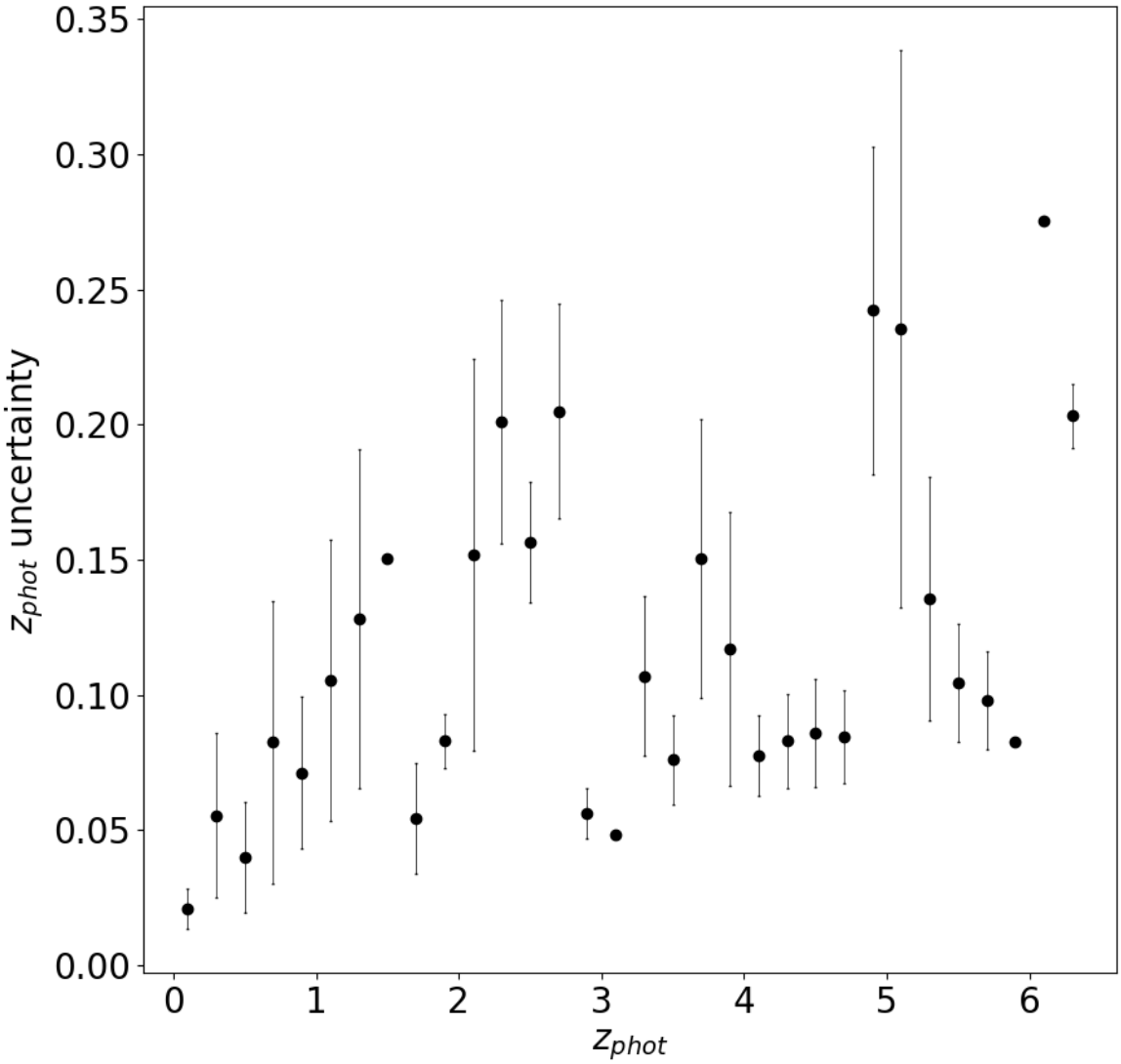


Figure 3.11: Median of the EAZY photo-z uncertainty of SOFI-detected sources as a function of photo-z. Sources are grouped into redshift bins with 0.2 increment. The average photo-z uncertainty of all sources is ~ 0.09. For data points with no error bar, there is only one sample in the corresponding redshift bins, therefore their error bars cannot be calculated.

3.6.1 Validation of photometric redshifts

Before analysing the results of the photometric redshift estimation, we first need to quantify the accuracy of the derived photo-z. This is commonly assessed by comparing the derived photo-z of sources with known spectroscopic redshifts (spec-z) to their spec-z, then calculating the normalised median absolute deviation (NMAD) of the photo-z, the bias of the photo-z and the fraction of photo-z outliers. Following Brammer et al. (2008), NMAD is defined as:

$$\sigma_{NMAD} = 1.48 \times median\left(\left|\frac{\Delta z - median(\Delta z)}{1 + z_{spec}}\right|\right) \tag{3.12}$$

where Δz is defined as $z_{phot} - z_{spec}$. As σ_{NMAD} is parametrised by median Δz, it is insensitive to catastrophic outliers. The bias of the derived photo-z is defined as:

$$b = |\Delta z|/(1 + z_{spec}) \tag{3.13}$$

which represents the difference between a galaxy's photo-z and its spec-z, normalised by its spec-z. Lastly, the fraction of outlier η is defined as the fraction of sources with $b > 0.15$.

Calculations of all these parameters require reliable spectroscopic redshifts. The only large spectroscopic redshift survey that covers both our target fields and the footprint of KiDS DR4 is 2dFGRS. However, due to the relatively shallow depth of 2dFGRS, there are only 9 of our SOFI sources which are observed by 2dFGRS. The small number of sample makes it difficult to have a reliable assessment of the photo-z quality of our SOFI-detected sources. Therefore, we use the photo-z of sources in the blank field to quantify the accuracy of our EAZY redshifts. Since the photometric redshifts of blank field sources are derived from KiDS DR4 photometry which has comparable quality to that of our SOFI catalogues, their quality should approximate the photo-z accuracy of target sources. There are 409 sources in the blank field that have 2dF redshifts. Figure 3.10 shows the comparison between the photo-z and the corresponding spec-z. For the blank field sources, we have achieved $\sigma_{NMAD} = 0.193$, a median bias $b = 0.019$ and a fraction of outliers $\eta = 0.733\%$. However, it should be noted that due to the shallow depth of 2dFGRS, sources which have 2dF redshifts are bright sources at low redshift ($< z > \approx 0.133$), therefore the accuracy of EAZY redshifts for sources lying at higher redshifts is still

unclear. Nonetheless, we can still estimate the precision of EAZY-derived redshifts at high-redshift by examining the 68% confidence interval of the redshift probability density distributions. We define our photo-z uncertainty as

$$\frac{(z_{phot} - z_{phot}^{min}) + (z_{phot}^{max} - z_{phot})}{2} \tag{3.14}$$

where z_{phot}^{min} and z_{phot}^{max} is the lower and upper boundary of the 68% confidence interval respectively. Figure 3.11 shows the median of the photo-z uncertainty as a function of redshift. As can be seen, all sources have their photo-z uncertainty lower than 0.30 at any redshift, with the majority of them having photo-z uncertainty between 0.00 and 0.10, and the average photo-z uncertainty of all sources is $\sim$ 0.09. Therefore, the difference between our EAZY photo-z and the true redshift of a source is likely to be smaller than 0.10 in most cases. In general, this indicates that we have achieved decent photo-z accuracy for our photo-z estimation.

3.7 Evidence for DSFG protoclusters

We then compare the photometric redshift distributions of sources in our target fields to that of sources in the blank field, as shown in Figure 3.12. As can be seen from the figure, the photo-z distributions of all target fields have an major spike at $z \approx 1$ which is also seen in the blank field photo-z distribution, this is likely the reflection of the overall redshift distribution of general sources combined with observational selection effects. However, some of the target fields have additional redshift spikes unlike the blank field distribution at redshifts roughly between 4 and 5. Such redshift spikes suggest overdensities of sources lying in this redshift range aligning with our candidate protoclusters. We quantify the overdensity by computing the density contrast between the target field and the blank field, which is defined as (e.g., Greenslade et al. 2018):

$$\delta(z) = \frac{n_{target}(z)}{n_{field}(z)} - 1 \tag{3.15}$$

where $n_{target}(z)$ and $n_{field}(z)$ is the number density of the target field and the blank field within a given redshift range respectively. Table 3.4 shows the density contrast at $4 \leq z \leq 5$ of each SOFI target field. A target field with density contrast $\delta > 1$ means that it is at least twice as dense as the

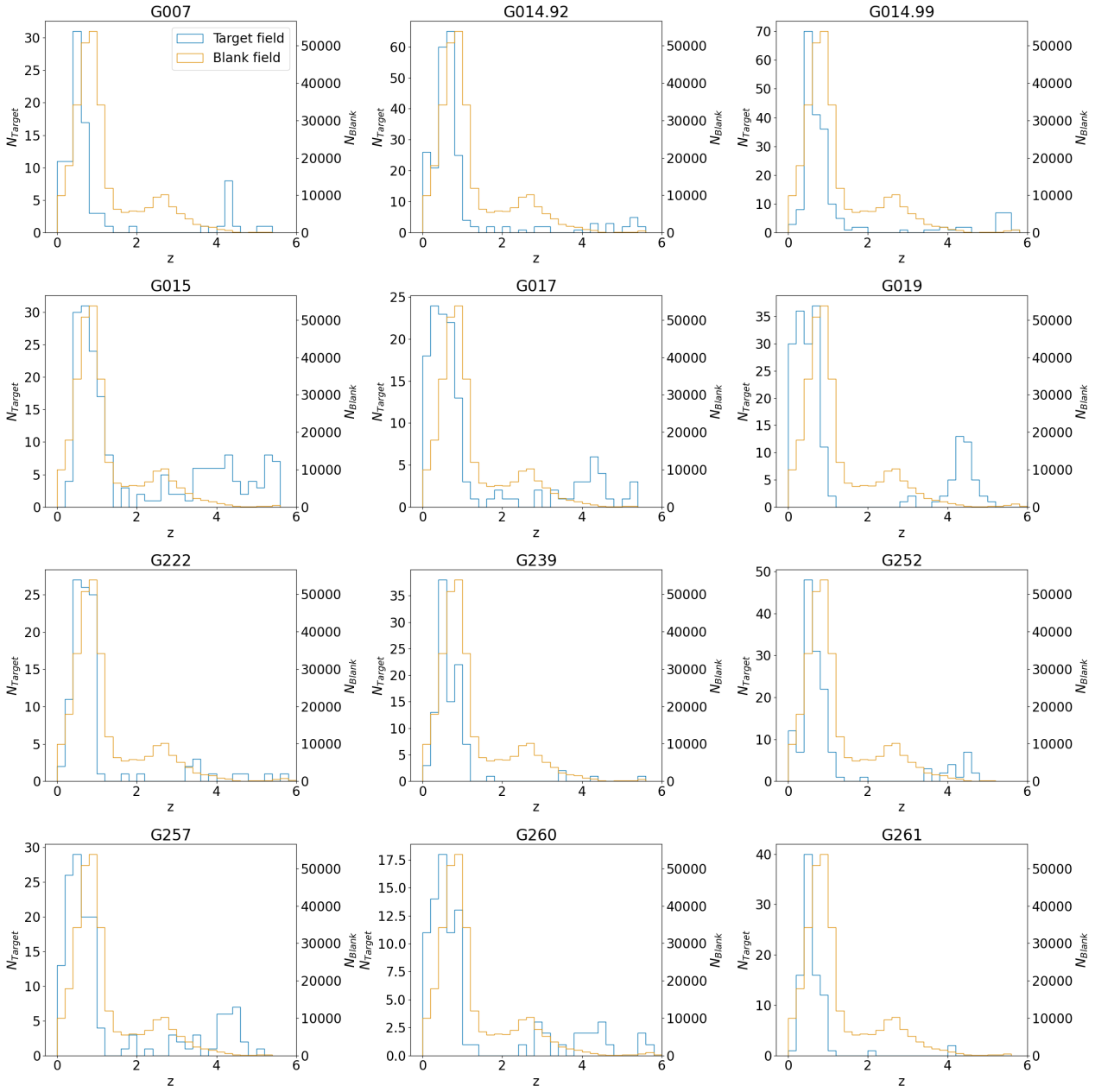


Figure 3.12: Photometric redshift distributions of sources in each target field. Each panel corresponds to one of the SOFI target fields, with the name of the field displayed in the panel title. In each panel, the blue histogram is the photo-z distribution of sources in the SOFI target field, the orange histogram shows the photo-z distribution of blank field sources. The left y-axis shows the number of sources in the SOFI target field and the right y-axis shows the number of sources in the blank field. While all target fields and the blank field have a $z \approx 1$ spike which may reflect the overall redshift distribution combined with selection effects, some target fields have redshift spikes at $z \approx 4-5$ unlike the blank field distribution, implying potential high-redshift overdensities of sources which may be associated with DSFG protoclusters.

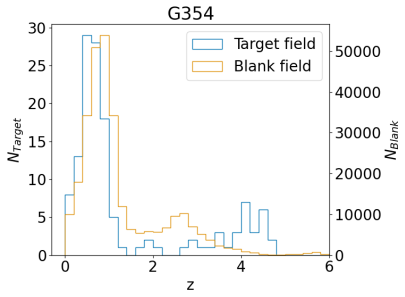


Figure 3.12 continued.

Table 3.4: Density contrast of each SOFI target field at $4 < z < 5$. The uncertainty of the density contrast is derived from the Poisson error of the number density of each field.

Name	$\delta(4 < z < 5)$
G007	1.3 ± 0.7
G014.92	0.5 ± 0.6
G014.99	0.2 ± 0.5
G015	4.2 ± 1.1
G017	2.4 ± 0.9
G019	8.1 ± 1.5
G222	-0.6 ± 0.3
G239	-0.7 ± 0.3
G252	2.1 ± 0.8
G257	4.4 ± 1.2
G260	1.1 ± 0.7
G261	-0.6 ± 0.3
G354	3.4 ± 1.0

blank field, thus considered to contain an overdensity which is potentially associated with the targeted candidate protocluster. Taking the uncertainties into account, 6 out of the 13 target fields have density contrasts larger than 1, with the highest density contrast being 8.1 ± 1.5 in G019. We additionally check the SPIRE colours of all *Herschel* sources inside the six SOFI fields with large density contrasts, as shown in Figure 3.13. As can be seen from Figure 3.13, most *Herschel* sources in the 6 SOFI fields have SPIRE colours comparable to an average ALESS SMG (da Cunha et al., 2015) at $z \sim 4$. This is roughly consistent with the redshift of the $4 \leq z \leq 5$ density contrasts seen in these fields, further suggesting there are potential structures at $z \sim 4 - 5$ in these 6 fields.

In conclusion, by comparing the photometric redshift distribution between the SOFI target fields and a randomly selected blank field, we have found candidate overdensities at $4 \leq z \leq 5$ in six of our

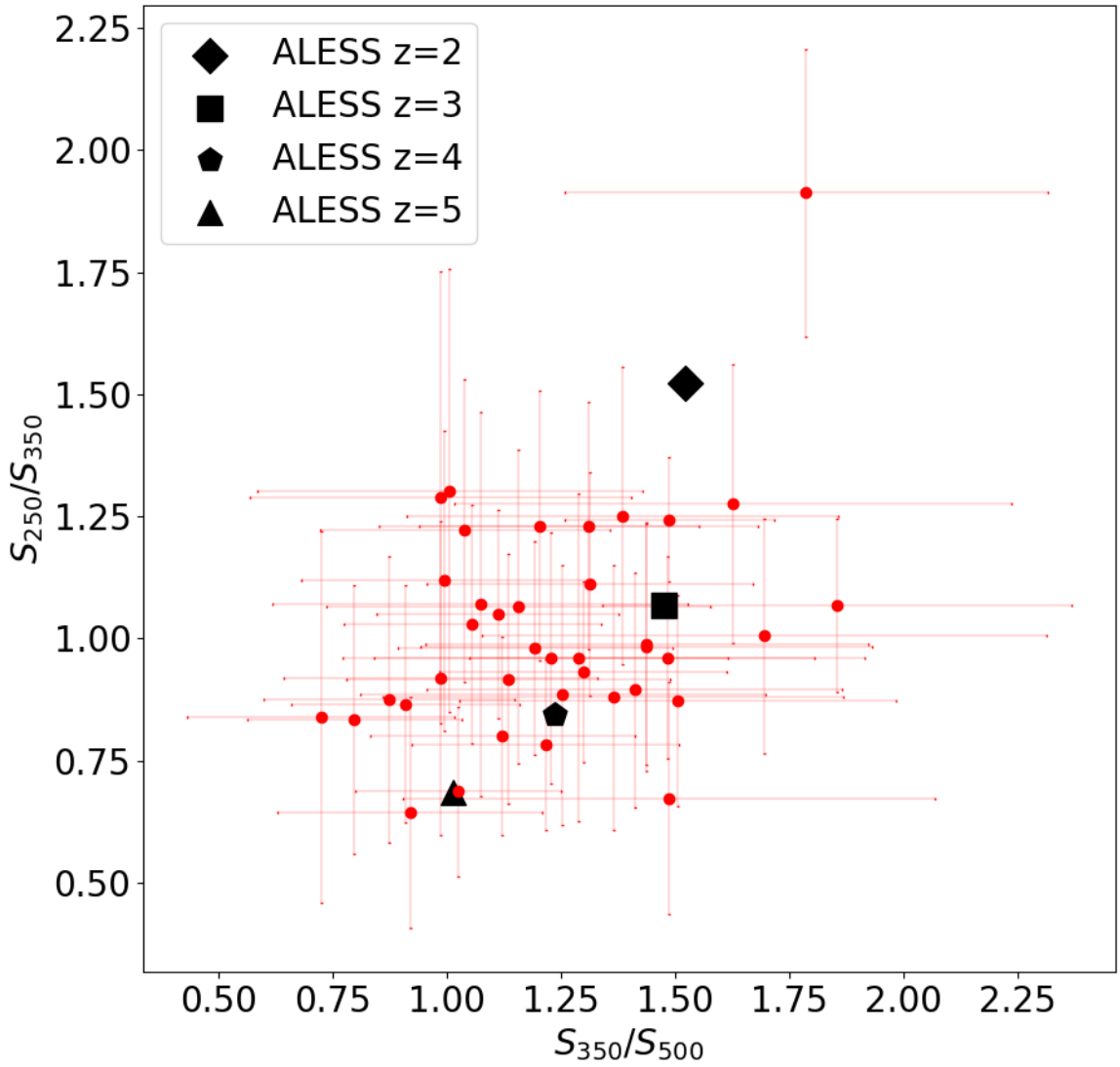


Figure 3.13: *Herschel* S_{250}/S_{350} and S_{350}/S_{500} colours of *Herschel* sources inside the 6 SOFI target fields which contain density contrasts at $z \sim 4-5$. The *Herschel* colours of the ALESS template SED (da Cunha et al. 2015) at redshift of 2, 3, 4 and 5 are also shown as black diamond, square, pentagon and triangle respectively. Most *Herschel* sources in the SOFI fields have colours roughly comparable with the ALESS colour at $z = 4$, indicating they are likely to be associated with the $4 \leq z \leq 5$ candidate overdensities.

thirteen SOFI target fields. The SPIRE colours of *Herschel* sources in these 6 fields also roughly indicates that they lie at $z \sim 4$. As our SOFI fields aim towards candidate DSFG protoclusters, these overdensities aligning with the candidate protoclusters are potential evidence of the candidates being real protoclusters at redshift higher than 4. With future spectroscopic follow-up on these candidate protoclusters, we can confirm their properties which will greatly boost the number of DSFG protoclusters at $z > 4$.

3.8 Limitations of photo-z analysis

The major limitation of our analysis is that its results solely depend on the photometric redshifts of sources, which are definitely not robust enough to be used as true redshifts of sources. Despite that we have already achieved decent photo-z accuracy, there is still potential to improve the photo-z quality to a higher level. This can be done by using higher quality photometry than the photometry used in this chapter. As our EFOSC2 data has not reached the desired quality due to non-ideal weather conditions, we have to combine public survey data in the optical/NIR with our deep SOFI data to construct multi-band photometry for photometric redshift estimations. Most public surveys are slightly shallower than our desired optical image depths in most filter bands. This will result in missing some faint SOFI sources in the optical/NIR data which is supported by the $\sim 84\%$ overall matching rate of SOFI sources to the complementary data used. In addition, the quality of our SOFI data is also varying between different target fields because of changes in weather conditions during the observations. Such a downgrade of the photometry will lead to poorer quality photometric redshifts, as demonstrated by previous photo-z studies (e.g., Dahlen et al. 2013; Bezanson et al. 2016; Brinchmann et al. 2017). Therefore, in future studies we must use higher quality photometry, either obtained from individual observations or from future public surveys such as those carried out via JWST, Euclid and Vera C. Rubin Observatory LSST, to compute more accurate photometric redshifts.

Another common method to improve the quality of photo-z is to add more photometric bands. Photometric bands used in this chapter include $ugrizYJHK_s$ in the optical/NIR, therefore we can add MIR data such as *Spitzer* IRAC 3.6-8 μm and MIPS 24 μm observations, FIR/submm data such as

Herschel 250-500 μm and SCUBA-2 450 and 850 μm observations, and radio data such as those taken by MeerKAT. However, there are various difficulties in adding these extra data. First, there is no existing deep MIR data which covers most of our targets. Also, MIR data usually has worse angular resolution than data in the optical/NIR due to the increase of wavelength. Adding such data, especially data with shallow depth, may in fact decrease the performance of derived photometric redshifts of faint objects. For example, Brinchmann et al. (2017) have demonstrates that adding IRAC photometry for faint galaxies actually worsens EAZY photo-z quality. DESI DR10 also includes WISE photometry in the MIR which is much shallower than the other photometry used in this chapter. This provides an opportunity to test the change of photo-z quality with WISE photometry added. The resulting photo-z accuracy is much worse, with $\sigma_{NMAD} \approx 1.8$ and $b \approx 2.7$. This dramatic decrease in photo-z performance demonstrates that adding low-quality MIR photometry is likely to significantly decrease EAZY photo-z performance. Second, the large beamsizes of FIR/submm instruments make it difficult to cross-match optical/NIR sources and FIR/submm sources, and FIR/submm bright sources are usually very faint in the optical/NIR which makes them more difficult to identify in the optical/NIR, as discussed in Section 1.4 of Chapter 1. Last, commonly used photometric redshift codes such as EAZY and BPZ do not allow direct inclusion of FIR/submm and radio data. Other template fitting codes such as CIGALE do allow the addition of FIR/submm and radio data, but they are optimised to derive physical properties of galaxies and are thus much less time-efficient in obtaining photo-z. For the above reasons, we do not include photometry in these wavelengths in this chapter, but future follow-up studies can obtain more data with better quality to improve the accuracy of photo-z.

Nonetheless, due to the lack of spectroscopically confirmed redshifts, we cannot obtain a reliable estimation of the photo-z accuracy of our target sources at a wider redshift range. Even if one can obtain very good average photo-z accuracy for a large sample of sources, it is still possible that photo-z estimations of some target sources fail catastrophically, producing photometric redshifts that significantly deviate from the true redshifts of the sources. Such catastrophic outliers may strongly change the final results, depending on the size or the nature of the sample. A recent example of this is a strongly lensed DSFG named as COSBO-7. Ling et al. (2024) identify the NIR/MIR counterpart of this source in JWST data and derive a $z_{phot} > 7$ for this source using four different SED fitting codes. This is further supported by ALMA archival single-line detection which is consistent with

CO(7-6) emission at $z = 7.46$. These make COSBO-7 a very promising $z > 7$ lensed DSFG with an regular disk morphology. However, later ALMA spectroscopic follow-up by Jin et al. (2024) confirm its redshift to be $z = 2.625$, revealing it as a more common main-sequence galaxy. Jin et al. (2024) suggest that the possible reasons for this catastrophic failure of photo-z are the underestimation of photometric uncertainty at 0.9 μm, and the lack of photometry at wavelengths > 20 μm. This example demonstrates that subtle changes in the photometry may result in vast differences in the resulting photo-z, and care should be taken with photo-z and pseudo-spectroscopic redshifts based on single line detections. Overall, photo-z can only provide evidence of promising candidate DSFG protoclusters. Spectroscopy follow-up is essential to securely confirm the nature of these candidates.

3.9 Conclusions

In this chapter, we present deep SOFI NIR observations in J, H and K_s of 13 candidate DSFG protoclusters selected from *Planck* and *Herschel* data in Greenslade et al. (2018). EFOSC2 optical observations of the same targets were also taken together with the SOFI data, but the EFOSC2 observations are severely downgraded due to bad weather conditions and therefore they are not presented here. Raw SOFI images are reduced using a combination of IRAF, the Astropy CCDPROC package and the Python package ASTROALIGN. The data reduction steps include flat-field correction and illumination correction, bad pixel removal, sky background subtraction and image alignment and combination. Sources are then extracted from the final science images using Source Extractor and their photometry is calibrated using zero points derived from 2MASS photometry. Catalogues of J, H and K_s images are combined into final catalogues based on the J-band source positions, leading to 2238 sources in all fields. We have 5σ detection down to $J = 23.07 \pm 0.21$, $H = 23.12 \pm 0.33$ and $K_s = 21.17 \pm 0.13$ mag.

We then cross-match our SOFI sources to complementary public survey data in the optical/NIR, resulting in 7-9 band multi-wavelength photometry for all sources. Photometric redshifts of all sources are estimated via the template fitting code EAZY. For comparison, we also randomly select a blank field and compute the photo-z of sources in this blank field. We have achieved $\sigma_{NMAD} = 0.193$,

bias $b = 0.019$ and $\eta = 0.733\%$ for the photo-z of blank field sources at low redshift. The average photometric redshift uncertainty of target sources at any redshift is ~ 0.09. These indicate that we have achieved decent photo-z accuracy for our photo-z estimations. By comparing the photo-z distribution of the blank field and that of our target fields, we have found candidate overdensities of $4 \leq z \leq 5$ sources aligning with 6 of our candidate protoclusters. The SPIRE colours of *Herschel* sources in these 6 fields also roughly indicates that they lie at $z \sim 4$. All the above evidence suggest that these candidate protoclusters are real protoclusters at high redshifts.

Our results further prove the feasibility of combining *Planck* and *Herschel* data to identify promising candidate DSFG protoclusters. Follow-up spectroscopic observations can then be carried out to confirm their nature. Apart from spectroscopic follow-up, the work presented in this chapter can be further expanded by including more candidate protoclusters selected from large *Herschel* surveys and the all-sky *Planck* survey. Also, the absence of deep EFOSC2 optical observations has hindered our ability to detect more sources with better photometry, hence limited us in discovering more overdensities at higher redshifts aligning with the candidate protoclusters. With future deep optical/NIR data, such as data collected by JWST, Euclid and LSST, more accurate photometric redshifts can be obtained to reliably identify more candidate DSFG protoclusters at even higher redshifts.

Chapter 4

The Nature of 500μm Risers

4.1 Introduction

In the early 1990s, the launch of the COsmic Background Explorer (COBE), equipped with the Far-InfraRed Absolute Spectrophotometer (FIRAS) and The Diffuse Infrared Background Experiment (DIRBE), detected the diffuse Cosmic Infrared Background radiation (CIB; Puget et al. 1996; Fixsen et al. 1998; Hauser et al. 1998; Cooray 2016). The CIB was found to have energy density comparable to the Cosmic Optical Background (COB; Cooray 2016). As interstellar dust mainly absorbs in the UV/optical and emits in the IR, the fact that the CIB and COB have comparable energy densities implies that around half of the stellar emission is absorbed and re-emited in the IR by dust over the entire history of the Universe.

The commissioning of the Submillimeter Common User Bolometric Array (SCUBA) in 1997 opened a new era of research in the high-z Universe by providing a statistically large sample of unlensed high-redshift FIR/submm bright sources for the first time. The 850 μm surveys with SCUBA demonstrated that the most intense high-redshift star-forming galaxies are heavily obscured by dust, which makes them largely invisible in optical/UV images. These dusty star-forming galaxies (DSFGs) have been extensively studied since then. Most of these DSFGs are found to have moderate redshifts of $z \approx 2-3$ (Chapman et al., 2005; Wardlow et al., 2011; Simpson et al., 2014; Smith et al., 2017a; Stach et al., 2019), large stellar masses ($M_* \approx 10^{10} - 10^{11} M_\odot$; Swinbank et al. 2004; Hainline et al. 2011;

Michałowski et al. 2012; Gruppioni et al. 2013; Simpson et al. 2014; Dudzevičiūtė et al. 2020; Pantoni et al. 2021), extreme IR luminosity ($L_{IR} \approx 10^{11} - 10^{13} L_{\odot}$; Magnelli et al. 2011; Swinbank et al. 2014; Gruppioni et al. 2013; MacKenzie et al. 2017; Michałowski et al. 2017; Rowan-Robinson et al. 2018; Hill et al. 2018; Cheng et al. 2019; Greenslade et al. 2020) and large dust and gas masses ($M_d \approx 10^7 - 10^9 M_{\odot}$, $M_{gas} \approx 10^{10} - 10^{11} M_{\odot}$, Smail et al. 2002; Greve et al. 2005; Tacconi et al. 2006, 2008; Clements et al. 2010a; Bothwell et al. 2013; Dudzevičiūtė et al. 2020; Pantoni et al. 2021). With such extreme properties, DSFGs generally have SFR of $\approx 10^2 - 10^3 M_{\odot} yr^{-1}$ and are thus likely to dominate the Cosmic Star Formation Rate Density (CSFDR) at $z \approx 2$ at which time the CSFRD peaks (also know as the Cosmic Noon). Therefore, understanding of the DSFG population will provide great insight into a critical stage in galaxy formation and evolution (Casey et al., 2014).

Despite extensive studies of the DSFG population, the nature of high redshift DSFGs, especially those that lie at $z > 4$, remains poorly understood. This is mainly due to the lack of large and homogeneous samples of the high redshift DSFGs discovered, as well as inadequate spectroscopically confirmed samples. Following the launch of the *Herschel Space Observatory* (Pilbratt et al., 2010) in 2010, an unexpectedly large number of potential high-redshift DSFGs was discovered. This sample is selected to have a rising SED from the 250 μm SPIRE band to the 500 μm band (i.e. $S_{250} < S_{350} < S_{500}$), hence their name 500 μm risers. The rest frame SED of a typical DSFG peaks at $\approx 100 \mu m$, therefore theoretically DSFGs with a rising SED at 500 μm will lie at $z > 4$ (e.g., Dowell et al. 2014; Asboth et al. 2016; Ivison et al. 2016). Dowell et al. (2014) fitted modified blackbodies to 38 HerMES-selected (Oliver et al., 2012) 500 μm risers, finding a redshift distribution peaking at $z = 4.7$, with spectroscopic redshifts for 5 of these sources supporting this redshift distribution. They also estimated the number density of bright ($S_{500} \geq 30 mJy$) sources to be $3.3 \pm 0.8\ deg^{-2}$. If these sources truly lie at $z > 4$, they would have IR luminosity $L_{IR} > 10^{13} L_{\odot}$ and corresponding SFR $> 1000 M_{\odot} yr^{-1}$, thus being the most extreme DSFGs known to date. However, only a handful of these 500 μm risers have spectroscopically confirmed redshifts (e.g. Riechers et al. 2013; Dowell et al. 2014; Ivison et al. 2016; Fudamoto et al. 2017; Zavala et al. 2018b), and their physical properties are not well studied.

The extreme properties of such DSFGs bring vast challenges to our current understanding of galaxy formation and evolution. Numerous attempts have been made to reproduce the observed DSFG/SMG

population with simulations using theoretical cosmology models and other observational constraints, but none of these have succeeded so far. Some early studies tried to use semi-analytic models (SAMs) with simple radiative transfer models or empirical galaxy SED templates to recover the observed SMG number counts, but these studies only found number counts that were $\sim 20-30$ lower than the observed counts (Granato et al., 2000; Fontanot et al., 2007; Somerville et al., 2012). There are other studies which have successfully reproduced the observed SMG number counts, but they failed to predict the physical properties of the observed SMGs such as redshift distribution, stellar mass or SFR (e.g., Granato et al. 2004; Baugh et al. 2005; Shimizu et al. 2012). Simulations focusing on individual systems have provided insights into the origin and formation of SMGs (e.g., Narayanan et al. 2015), but such studies cannot explain the observed number of SMGs. More recently, large-volume cosmological hydrodynamic simulations have been used to recreate the observed SMG number counts and properties, but these simulations generally have underestimated the observed SMG number counts. For example, McAlpine et al. (2019) make use of EAGLE (Schaye et al., 2015; Crain et al., 2015) to reproduce a selected sample of SMGs which have $z \geq 1$ with mock sub-mm flux $S_{850\mu m} \geq 1\ mJy$. While they have recovered most of the physical properties of the sample reasonably well, their predicted number counts of SMGs with $S_{850\mu m} \geq 3\ mJy$ are significantly lower than the real counts (Wang et al., 2019). A similar analysis carried out by Lovell et al. (2021), using SIMBA (Davé et al., 2019), results in underestimated SMG number counts that are at most 10 times lower than the observed counts. Also, Hayward et al. (2021) carry out similar simulations with IllustrisTNG (Pillepich et al., 2018; Nelson et al., 2018; Naiman et al., 2018; Marinacci et al., 2018; Springel et al., 2018), and their predicted SMG number counts are also significantly less than the observed values by more than 1 dex at $S_{850\mu m} \geq 4\ mJy$.

It has been known for a long time that a significant fraction of 500μm risers are blends of multiple fainter objects, but the precise fraction is still not clear. Béthermin et al. (2017) find via simulations that most 500 μm risers are blends of multiple galaxies, with the brightest one contributing 60% on average to the total 500 μm flux density of a SPIRE source. On top of that, they find nearly all sources with $S_{500} > 60\ mJy$ are strongly lensed single galaxies. Oteo et al. (2017) perform follow-up observations with ALMA on 500 μm risers selected from Ivison et al. (2016), finding that 16 of their 44 single dish sources are found to be blends, with several sources resolved into 3 or

more counterparts, and in particular one *Herschel* source splits into 5 galaxies. Moreover, 18 of their sources are gravitationally lensed, but they do not specifically investigate the lensing effect on $S_{500} > 60\,mJy$ sources, most of which should be strongly lensed according to Béthermin et al. (2017). The relatively low fraction of blends in Oteo et al. (2017) agrees with the 500 μm riser counts from *Herschel* to some extent, but their sample size is too small to draw statistically solid conclusions, plus the small primary beamsize means that there may be sources lying outside the small ALMA beam that also contribute to the *Herschel* fluxes. Montaña et al. (2021) obtain Large Millimetre Telescope (LMT)/AzTEC 1.1mm observations of 93 bright 500 μm risers with an effective angular resolution of 9.5″, 40 of them have single detections, 8 have multiple detections while 45 of them have no detection. This leads to a $\sim$ 9% multiplicity fraction, but some of their non-detections and single detections are potential multiple systems made of intrinsically fainter galaxies with flux densities just below their 4σ detection threshold. Taking account for these sources will boost the fraction of multiples to $\sim$ 23%. In the most extreme scenario where all non-detections are considered as potential multiple detections, their multiplicity fraction becomes $>$ 50%. Greenslade et al. (2020) carry out SMA observations on thirty-four 500 μm risers selected from the *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al. 2012). They find that 18 of their 34 targets resolve into single sources, 4 resolve into two individual sources and the other 12 targets have no SMA detection. By taking the ALESS average SED (da Cunha et al., 2015) at redshift between 4 and 6, Greenslade et al. (2020) predict that these 12 undetected sources may consist of $\geq$ 3 fainter DSFGs separated by 2 – 20″ (the SMA and SPIRE beamsizes respectively) so that their flux densities are below the detection threshold of the SMA. Assuming all non-detections are blends of multiple sources, Greenslade et al. (2020) conclude that $\sim$ 35% of bright 500 μm risers ($S_{500\mu\mathrm{m}} \geq 60\,mJy$) and $\sim$ 60% $S_{500\mu\mathrm{m}} < 60\,mJy$ faint 500 μm risers should be blends. Cairns et al. (2023) carry out follow-up SCUBA-2 and deep SMA observations on four of the non-detections in Greenslade et al. (2020), finding that for two of the non-detections, each of them breaks up into two faint sources in SMA images and are extended in the SCUBA-2 images. One of their targets resolves into a single source in both SMA and SCUBA-2 images. The remaining target also resolves into a single SMA source, but has two partially resolved sources in its SCUBA-2 observation, one of which has no SMA counterpart. The results of Cairns et al. (2023) support the multiplicity fraction of Greenslade et al. (2020). 500 μm risers with high multiplicity are natural

candidates for being high-redshift DSFG protoclusters which are poorly understood as only a limited number of them have been discovered so far (Miller et al., 2018; Oteo et al., 2018).

Despite the importance of 500 μm risers in constraining models of galaxy evolution and formation, there are only a small number of studies on the nature of large homogeneous samples of 500 μm risers, thus many of the properties of 500 μm risers are not well understood. Although the selection criteria of 500 μm risers are designed to pick out intrinsically red, FIR luminous, high-redshift DSFGs, it has been pointed out that it will also include some multiple systems at moderate redshift with their *Herschel* SPIRE colours being artificially reddened by blending. As the *Herschel* SPIRE beamsize increases from 17.6″ in the 250 μm band to 35.2″ in the 500 μm band, multiple nearby *Herschel* sources are more likely to be blended into one brighter source at longer wavelengths, leading to an extreme red SPIRE colour. Ma et al. (2019) show that about 20% of their selected 500 μm risers would not be 500 μm risers after deblending. Similarly, $\sim$ 25% of the 500 μm riser samples in Montaña et al. (2021) would not pass the criteria after taking account of the effects of blending. Furthermore, Montaña et al. (2021) find that the 500 μm risers with multiplicity have a higher fraction of low photometric redshift ($z_{phot} < 3$) sources ($\sim$ 27%) than those with single detections ($\sim$ 10%). Cairns et al. (2023) also find that all of the four 500 μm risers, which are not detected by SMA in Greenslade et al. (2020), have photometric redshifts of $2 \lessapprox z < 3$. Donevski et al. (2018) further added that confusion and instrumental noise can also increase the number of 500 μm risers detected by *Herschel*.

It is therefore essential to understand the nature of 500 μm risers before they can be extensively used in future studies. To better understand the physical properties of 500 μm risers, we carry out multi-wavelength cross-identifications of 18 SMA single-detection 500μm risers plus one two-detection source, selected from Greenslade et al. (2020). With multi-band photometry obtained, photometric redshifts of the selected sample are estimated using three different codes, EAZY (Brammer et al., 2008), MMPZ (Casey, 2020) and CIGALE (Boquien et al., 2019). We then compute SEDs of the target objects with CIGALE using known spectroscopic redshifts and MMPZ redshifts, deriving their physical properties such as L_{IR}, SFR and stellar mass.

This chapter is organised as follows. In Section 4.2 we describe our sample selection. In Section 4.3, the photometry used for multi-wavelength cross-identification and SED fitting is presented. Section

4.4 describes the photometric redshift estimations of our sample. Section 4.5 and 4.6 outline the use of CIGALE SED fitting and the validation of its quality. Derived physical properties are shown and discussed in Section 4.7. Some limitations and potential improvements in our results are discussed in Section 4.8. Finally, Section 4.9 concludes this chapter and discusses future directions.

4.2 Sample Selection

Greenslade et al. (2020) have selected thirty-four 500 μm risers from the HerMES survey based on their *Herschel*-SPIRE fluxes and colours (i.e. $S_{250\mu m} < S_{350\mu m} < S_{500\mu m}$). All sources are detected by *Herschel* at $> 4\sigma$, with an average 500 μm flux of $67 \pm 29\ mJy$. It should be noted that Greenslade et al. (2020)'s sample is heterogeneous in the sense that there is no single consistent flux or colour criterion used. Greenslade et al. (2020) also note that recent refinements to the data reduction change the flux densities of their sample in the latest maps and catalogues, such that 7 of the original 34 sources cannot be strictly classified as 500 μm risers. These sources are still included nonetheless, as investigations by Donevski et al. (2018) and Béthermin et al. (2017) have shown that there can be a significant contamination of selected 500 μm risers by flux boosted sources with $S_{500\mu m} \leq S_{350\mu m}$. As mentioned in Section 4.1, Greenslade et al. (2020) have carried out deep observations of their sample with the SMA at frequencies around either 345GHz or 265GHz, finding that 4 of the selected sources resolved into two individual components, 18 have single detections and the remaining 12 have no apparent detection. In this chapter, the 18 Greenslade et al. (2020) sources with single SMA detection are selected for multi-wavelength cross-identifications, and hence precise SED fitting and derivation of their physical properties. In addition, Lock_2 is also selected despite the fact that it breaks into 2 sources in 270GHz maps, but only one of these two sources has a counterpart in 345GHz maps. The position and photometry of the 345GHz source is used in the following analyses, as it is the one with robust detection at both frequencies and can be treated as a single source. The other source in Lock_2 is brighter in 270GHz, and is thus probably a blend of multiple faint sources below the detection limit as discussed in Greenslade et al. (2020). It should also be noted that, both Lock_102 and HFLS_3 are also studied by other literatures, they are found to be gravitationally lensed with magnification factor $\mu = 12.5 \pm 1.2$ and $\mu = 2.2 \pm 0.3$ respectively (Wardlow et al., 2012; Riechers et al., 2013; Cooray

Source	RA (J2000)	DEC (J2000)	$S_{250\mu m}$ (mJy)	$S_{350\mu m}$ (mJy)	$S_{500\mu m}$ (mJy)	$S_{870\mu m}$ (mJy)	$S_{1.13mm}$ (mJy)
XMM_M7	02:26:44.73	-03:25:10.90	39.7 ± 7.4	60.3 ± 8.2	57.6 ± 9.3	13.3 ± 2.55	_
XMM_M2	02:25:15.38	-02:47:08.10	44.5 ± 8.7	54.1 ± 9.9	67.6 ± 9.4	14.4 ± 2.51	_
XMM_26	02:25:44.78	-02:59:14.30	45.7 ± 11.9	67.8 ± 9.6	81.6 ± 8.8	_	7.62 ± 1.92
HeLMS_11	00:29:36.06	+02:07:13.05	83.9 ± 15.6	89.6 ± 14.9	110.4 ± 17.5	_	10 ± 2.07
HeLMS_12	00:52:58.52	+06:13:18.80	76.9 ± 16.7	122.5 ± 16.6	138.7 ± 17.3	_	38.6 ± 3.8
HeLMS_28	00:44:10.27	+01:18:22.00	91.2 ± 14.9	137.8 ± 14.2	160.5 ± 16.9	_	34.4 ± 3.75
HeLMS_32	23:24:34.59	-05:24:55.78	36.0 ± 16.0	61.5 ± 14.8	72.4 ± 16.9	_	12.6 ± 1.9
HeLMS_34	00:22:20.79	-01:55:20.70	66.7 ± 14.7	90.7 ± 14.4	106.7 ± 16.0	_	26 ± 3.25
HeLMS_62	00:55:17.18	+02:04:01.20	36.0 ± 27.4	30.6 ± 17.0	41.0 ± 17.1	_	11.9 ± 2.91
HeLMS_783	00:41:30.21	-02:47:54.11	42.0 ± 15.0	46.6 ± 15.0	61.0 ± 17.0	_	7.81 ± 1.72
Cosmos_9	10:02:48.04	+02:33:19.98	20.6 ± 6.1	32.9 ± 5.9	47.6 ± 6.9	_	3.62 ± 0.66*
LSW_28	11:01:27.81	+56:19:34.60	41.0 ± 11.8	65.1 ± 11.9	66.9 ± 13.2	_	19.2 ± 1.71
Lock_2	10:53:10.92	+56:42:06.84	41.4 ± 7.7	67.0 ± 8.0	65.1 ± 8.1	13.1 ± 3.0	5.68 ± 0.94
Lock_5	10:51:31.89	+56:36:19.7	17.5 ± 8.3	38.1 ± 9.1	44.1 ± 8.9	14.2 ± 2.92	_
Lock_102	10:40:50.64	+56:06:53.84	64.0 ± 6.3	105.5 ± 6.7	122.7 ± 7.1	72.1 ± 2.99†	_
Bootes_33	14:31:32.00	+34:21:09.88	50.0 ± 8.1	60.9 ± 8.7	57.0 ± 8.0	_	6.23 ± 1.08
HFLS_1	17:08:17.67	+58:28:40.10	51.6 ± 9.8	82.4 ± 10.3	82.9 ± 9.7	_	15.8 ± 2.87
HFLS_3	17:06:47.69	+58:46:23.88	15.2 ± 6.7	33.4 ± 8.0	52.8 ± 9.3	24.6 ± 1.79	18.4 ± 1.36
HFLS_5	17:20:49.49	+59:46:27.12	24.8 ± 8.4	53.0 ± 9.2	45.8 ± 9.1	_	13.3 ± 1.32

*Cosmos_9 was observed by SMA at 231GHz instead of 265GHz, the flux density here is $S_{1.3mm}$.
†Lock_102 was observed by SMA at 302GHz instead of 345GHz, the flux density here is $S_{993\mu m}$.

Table 4.1: The position and integrated FIR/sub-mm photometry for the target sources, taken from Greenslade et al. (2020). The SMA 345GHz and 265GHz flux densities are shown as $S_{870\mu m}$ and $S_{1.13mm}$ respectively

et al., 2014; Riechers et al., 2020). The positions, *Herschel* SPIRE photometry and SMA photometry for these sources are presented in Table 4.1. Figure 4.1. illustrates the SMA images of some example sources.

4.3 Photometry

Due to the wide range of sky locations of our targets, the optical, NIR and MIR photometry used in this chapter is selected from various databases with heterogeneous photometric band coverage and depth. This section describes the photometry used in our cross-identifications and SED fitting. Positional cross-matching of our targets to various databases is performed with a matching radius of 3" to their respective SMA positions, followed by visual examinations. A summary of photometry used for each target is listed in Table 4.2. An example of our cross-identifications is illustrated in Figure 4.2.

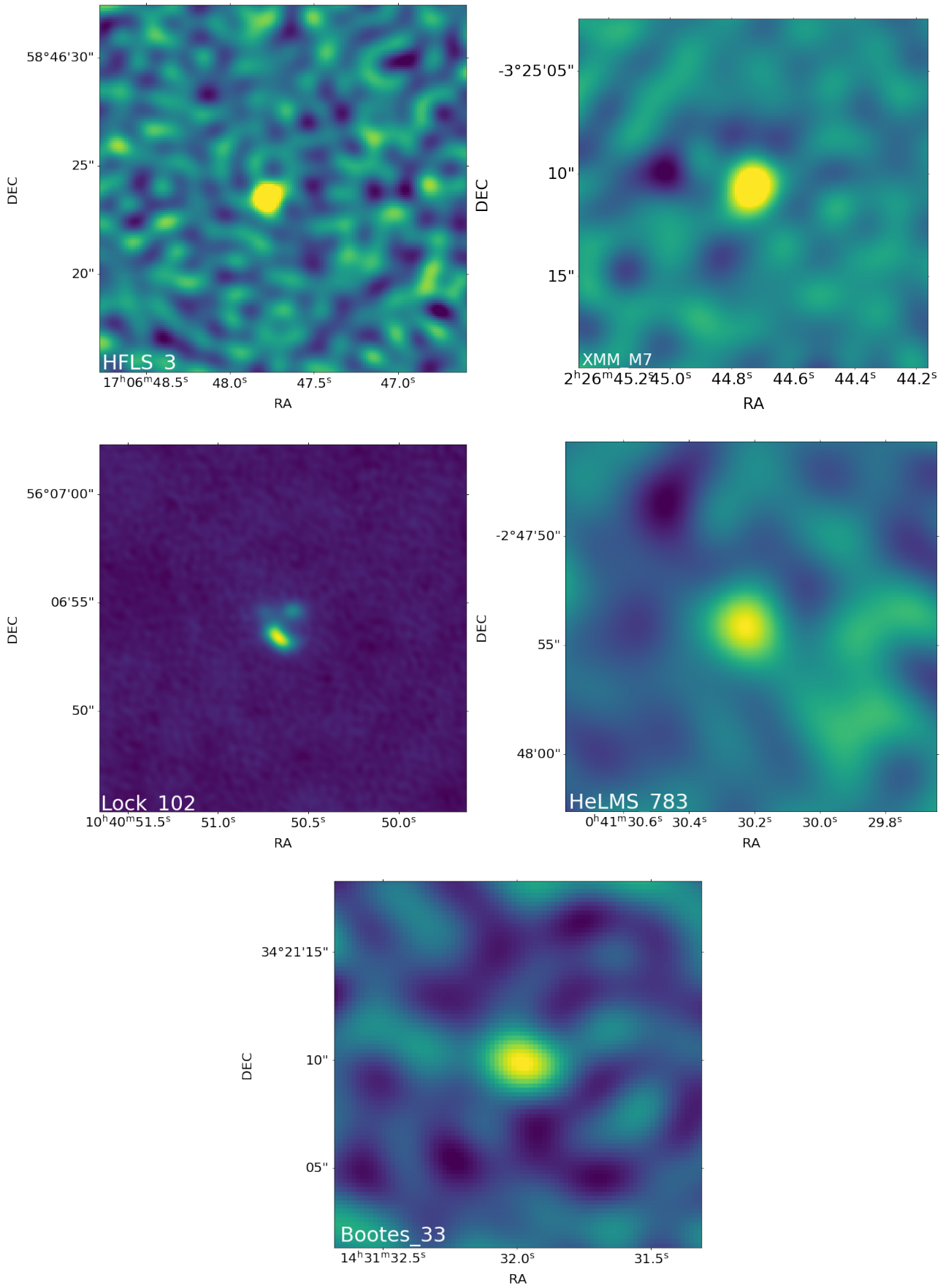


Figure 4.1: Examples of the SMA maps of sources presented in this chapter. The name of the source is shown in bottom left of each panel. Lock_102 is known to be strongly lensed with a magnification $\mu = 12.5 \pm 1.2$ (Riechers et al., 2020), HFLS_3 is also known to be weakly lensed with a magnification $\mu = 2.2 \pm 0.3$ (Cooray et al., 2014)

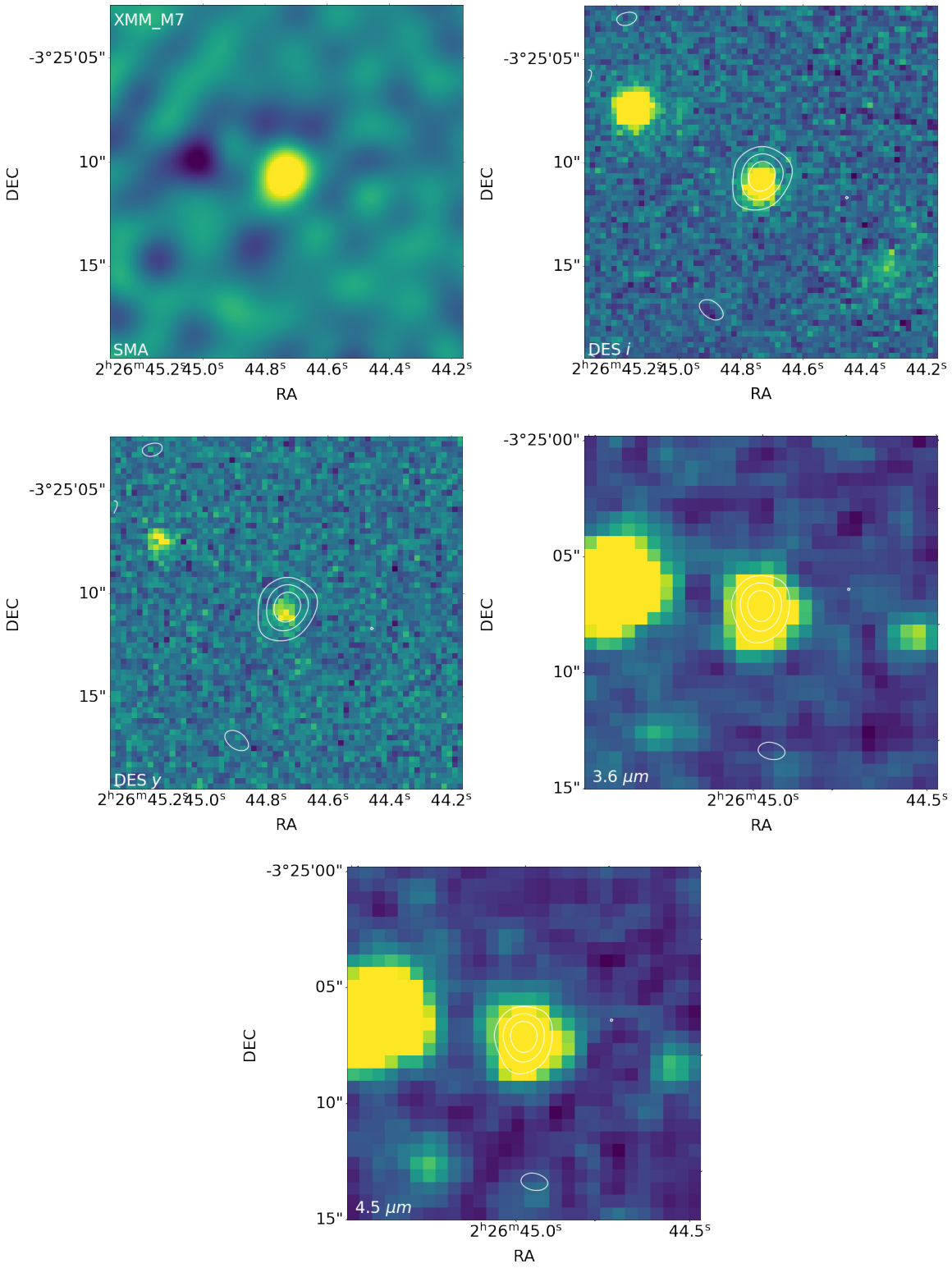


Figure 4.2: Multi-wavelength cutout image of XMM_M7 as an example of our cross-identifications. Images listed are (from top to bottom, left to right): SMA 345 GHz, DES *i* band, DES *y* band, *Spitzer* 3.6μm and *Spitzer* 4.5μm. The white SMA contour of XMM_M7 is overlaid.

4.3.1 Dark Energy Survey

7 of our targets are covered by the Dark Energy Survey data release 1 (DES DR1; Abbott et al. 2018). DES is a deep wide area optical/NIR survey covering $\sim$ 5000 deg^2 of the southern Galactic cap in five broad photometric bands, $grizY$. DES uses the Dark Energy Camera (DECam; Flaugher et al. 2015), a 570 Megapixel camera with a 3 deg^2 field-of-view installed at the prime focus of the Blanco 4-m telescope at the Cerro Tololo Inter-American Observatory (CTIO) in northern Chile. The median coadded catalogue depth for a 1.95" diameter aperture at S/N = 10 is $g = 24.33, r = 24.08, i = 23.44, z = 22.69$, and $Y = 21.44$ mag. Seven of our targets were observed by DES.

4.3.2 DESI Legacy Imaging Survey

The Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Survey (Dey et al., 2019), hereafter the Legacy Survey or DESI, are a combination of three public projects, the Dark Energy Camera Legacy Survey (DECaLS; Dey et al. 2019), the Beijing–Arizona Sky Survey (BASS; Zou et al. 2017) and the the Mayall z-band Legacy Survey (MzLS; Dey et al. 2019). These three projects jointly image $\simeq 14,000$ deg^2 of the extragalactic sky visible from the northern hemisphere in three optical bands (g, r, and z) using telescopes at the Kitt Peak National Observatory and the CTIO. The Legacy Survey catalogue includes photometry from the grz optical bands and from four mid-infrared bands (at 3.4, 4.6, 12, and 22 μm) observed by the Wide-field Infrared Survey Explorer (*WISE*; Wright et al. 2010) satellite during its full operational lifetime. The latest data release of the Legacy Surveys is the tenth public data release. Two of our 500 μm risers have counterparts in DECaLS data which reaches 5σ depths of $g = 23.95$, $r = 23.54$, $z = 22.50$ for point sources in individual images.

4.3.3 Sloan Digital Sky Survey

The Sloan Digital Sky Survey (SDSS; York et al. 2000) is a large area imaging survey in five broad optical bands $ugriz$, using the 2.5m Sloan Foundation Telescope (Gunn et al., 2006) at the Apache Point Observatory (APO). In this chapter, the latest sixteenth data release from SDSS (DR16)

is used. SDSS DR16 covers a large sky area of 14,555 deg^2, reaching median 5σ depths of $u = 22.15$, $g = 23.13$, $r = 22.70$, $i = 22.20$, $z = 20.71$. Two of our targets are identified in SDSS DR16.

4.3.4 COSMOS2020

One of our targets lies in the COSMOS field. COSMOS2020 (Weaver et al., 2022) is the latest release of the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007) catalogue. COSMOS2020 comprises several new deep optical and NIR surveys in the ~ 2 deg^2 COSMOS field, such as the ultra-deep optical data from the Hyper Suprime-Cam (HSC) Subaru Strategic Program (SSP) Public Data Release 2 (Aihara et al., 2019), and the new UltraVISTA data release 4 (McCracken et al., 2012; Moneti et al., 2019) in the NIR. *Spitzer* data from the Cosmic Dawn Survey (Moneti et al., 2022) is included in COSMOS2020. This consists of all IRAC data taken in the COSMOS field up to the end of the mission in January 2020. Photometric redshifts are computed for all sources in the COSMOS2020 catalogue utilizing two independent photometric redshift codes, *Le Phare* (Arnouts et al., 2002; Ilbert et al., 2006) and EAZY (Brammer et al., 2008).

4.3.5 DeepDrill

The *Spitzer* DeepDrill survey (Lacy et al., 2021) is designed to observe three of the four Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019) Deep Drilling Fields at 3.6μm and 4.5μm. The DeepDrill survey covers the footprint of LSST in the Extended Chandra Deep Field–South (ECDFS) field, the ELAIS-S1 field (ES1), and the XMM-Large-Scale Structure Survey field (XMM-LSS), with a approximate 5σ point-source depth of 23.1 mag in both IRAC1 and IRAC2 bands. One of our target is covered by the DeepDrill survey.

Source	Optical Data Origin	NIR/MIR Data Origin	Photometric Band Used
XMM_M7	DES	XMM_DeepDrill	$grizY$, $IRAC1$, $IRAC2$
XMM_M2	DES	_	$gizY$
XMM_26	DES	_	$grizY$
HeLMS_11	DECaLS	WISE(DECaLS)	grz, $WISE1$, $WISE2$, $WISE3$, $WISE4$
HeLMS_12	_	_	_
HeLMS_28	DES	_	$griz$
HeLMS_32	DECaLS	WISE(DECaLS)	grz, $WISE1$, $WISE2$, $WISE3$
HeLMS_34	DES	_	$grizY$
HeLMS_62	DES	_	$grizY$
HeLMS_783	DES	_	$grizY$
Cosmos_9	COSMOS2020	COSMOS2020	$ugrizy$, $IRAC2$
LSW_28	_	_	_
Lock_2	_	_	_
Lock_5	SDSS*	_	$ugriz$
Lock_102	_	_	_
Bootes_33	SDSS	_	$ugriz$
HFLS1	_	_	_
HFLS3	_	Riechers et al. (2013), Cooray et al. (2014)	WFC3/F125W, WFC3/F160W, Keck/NIRC2 K_s $IRAC1$, $IRAC2$
HFLS5	_	_	_

*The SDSS counterpart of Lock_5 is 3.28" away from its SMA position, slightly larger than the 3" matching radius.

Table 4.2: List of optical and NIR/MIR surveys used to build the multi-wavelength photometry in this chapter.

4.3.6 HFLS_3

HFLS_3 is a special case, as it is well-studied by Riechers et al. (2013). It is included in our sample as a benchmark to assess the qualities of the various analyses performed in this chapter. Riechers et al. (2013) spectroscopically confirmed a redshift of 6.34 for HFLS_3, along with detailed multi-wavelength photometry. Although they carried out deep optical observations on HFLS_3, no significant emission was detected at the position of HFLS_3. Later, Cooray et al. (2014) acquired Hubble/WFC3 and Advanced Camera for Surveys (ACS) imaging data of HFLS_3, finding no robust detection of the source in the ACS F625W and F814W filter even with deep limiting magnitudes of 27.01 and 28.20 mag respectively. Instead, they did obtained robust detection of HFLS_3 in the NIR. This is consistent with total absorption due to a Gunn-Peterson trough at observed-frame $< 892\ nm$, and significant dust obscuration longward of the Lyman break. Riechers et al. (2013) presented *Spitzer* IRAC1 and IRAC2 data of HFLS_3, Cooray et al. (2014) added Hubble Space Telescope (HST) WFC3/F125W, WFC3/F160W and Keck/NIRC2 K_s IR data of the source. These data are used in the following analyses.

4.4 Photometric Redshifts

With sufficient multi-wavelength photometry, we estimate photometric redshifts for our targets using three different tools, namely EAZY , MMpz (Casey, 2020) and CIGALE (Boquien et al., 2019).

EAZY is a photometric redshift code designed to produce high-quality redshifts over $0 < z < 4$, though it can still be used for higher redshifts with slightly worse quality. The code can fit linear combinations of galaxy SED templates, and it includes optional flux and redshift-based priors. Its default template set, as well as the default functional forms of the priors, are not based on (usually highly biased) spectroscopic samples, but on semi-analytical models. The photometric redshifts of our sources were estimated by EAZY using only optical to MIR photometry in a redshift range of $0.1 < z < 6$. Note that although HFLS_3 has a spectroscopic redshift of $z = 6.34$ which is larger than 6, we do not estimate its photometric redshift with EAZY due to the lack of optical photometry.

MMpz is another photometric redshift code specialised in providing far-infrared (FIR) and submillimeter/millimetre photometric redshifts for dusty galaxies which lack good redshift constraints in the optical/NIR. It takes into account the intrinsic breadth in galaxies' rest-frame SEDs, which reflects a variety of dust temperatures. It also uses both the galaxies' flux densities and their FIR/mm colours to determine the best underlying redshift constraints. Photometric redshifts of our sources are estimated by MMpz using the FIR to millimetre photometry without other constrains.

CIGALE, which stands for Code Investigating GALaxy Emission, is an SED-fitting code which models the FUV to radio spectrum of galaxies and estimates their physical properties such as redshift, star formation rate, attenuation, dust luminosity, stellar mass and others. To compute the spectral models, CIGALE builds composite stellar populations from simple stellar populations combined with highly flexible star formation histories, calculates the emission from gas ionised by massive stars, and attenuates both the stars and the ionised gas with a highly flexible attenuation curve. AGN contributions can also be taken into account during model computation. A large grid of models is then fitted to the data and the physical properties are estimated through the analysis of the likelihood distribution. We computed the SEDs of our sample with CIGALE using all photometry available to estimate their photometric redshifts between $1 < z < 6.5$.

Table 4.3. shows a summary of our photometric redshift estimations with respective errors. With EAZY, the default setting is used except for the redshift range being set to $0.1 < z < 6$. MMPZ has few parameters to modify, so the default setting is implemented. CIGALE has many choices of modules and parameters to compute the models. For this work we use the sfh2exp, bc03, nebular, dustatt_modified_CF00, dl2014, fritz2006, redshifting modules. More detailed descriptions and explanations of the choice of CIGALE modules and parameters are included in Section 4.5. To assess the quality of the photometric redshifts given by these 3 tools, we compare the derived photometric redshifts to the spectroscopic redshifts for the 8 sources with known spectroscopic redshifts, as shown in Figure 4.3. The names of sources with spectroscopic redshifts and the original literatures providing these redshifts are as follow: HFLS_1,3,5 and Lock_102 from Riechers et al. (2013), Lock_5 from Dowell et al. (2014), HeLMS_34 from Asboth et al. (2016), and HeLMS_12 and HeLMS_28 from Riechers et al. (2021). As one can notice from Figure 4.3, photometric redshifts estimated by MMPZ are generally consistent with the spectroscopic redshifts within $\Delta z/(1+z) = \pm 0.2$, while CIGALE redshifts have a worse performance with larger errors. Due to the lack of optical photometry, only three sources with spectroscopic redshifts have EAZY redshifts, they also have the worst accuracy among the three sets of photometric redshifts. Therefore, we conclude that in our case MMPZ gives the most reliable photometric redshift estimations although the uncertainties are significant. The redshifts of our sample hereafter in this chapter refer to the MMPZ redshifts for sources without known spectroscopic redshifts, and the spectroscopic redshifts for sources with known spec-z. Figure 4.4 shows the redshift distribution of our sample. The mean redshift of our sample is $< z >= 3.874 \pm 0.989$ and the median is 3.680 which is slightly lower than the expected $z > 4$ for 500 μm risers, but it is higher than the average redshift of $850\mu m$ selected sources with $< z > \approx 2.5 - 2.6$ (e.g. Chapman et al. 2005; Dudzevičiūtė et al. 2020; Shim et al. 2022). Of the 19 targets, 10 have redshift lower than 4, within the range of $2.6 < z < 3.7$. Nonetheless, the redshifts of our sample with single SMA detections are in general higher than those of the sample with multiplicity studied by Cairns et al. (2023) ($2 \lessapprox z < 3$).

Source	z_{EAZY}	z_{MMpz}	z_{CIGALE}	z_{spec}
XMM_M7	$3.195^{+0.187}_{-0.180}$	$2.875^{+0.673}_{-0.600}$	2.242 ± 0.566	_
XMM_M2	$3.305^{+1.075}_{-1.087}$	$2.975^{+0.846}_{-0.850}$	2.486 ± 0.756	_
XMM_26	$2.725^{+0.666}_{-0.704}$	$3.075^{+0.889}_{-0.900}$	2.099 ± 0.496	_
HeLMS_11	$3.745^{+0.523}_{-0.494}$	$2.725^{+0.793}_{-0.750}$	1.328 ± 0.459	_
HeLMS_12	_	$4.625^{+0.535}_{-0.550}$	4.303 ± 1.115	4.370
HeLMS_28	$2.682^{+0.771}_{-0.797}$	$4.175^{+0.578}_{-0.550}$	2.351 ± 0.543	4.17
HeLMS_32	$5.285^{+0.309}_{-0.279}$	$3.675^{+0.542}_{-0.550}$	5.565 ± 0.762	_
HeLMS_34	$2.509^{+0.589}_{-0.721}$	$4.275^{+0.564}_{-0.550}$	2.392 ± 0.301	5.16
HeLMS_62	$3.138^{+0.092}_{-0.090}$	$4.425^{+0.712}_{-0.750}$	4.446 ± 0.070	_
HeLMS_783	$3.369^{+0.417}_{-0.298}$	$3.125^{+0.606}_{-0.650}$	4.045 ± 0.187	_
Cosmos_9	$1.996^{+0.176}_{-0.197}$	$3.375^{+0.846}_{-0.700}$	1.002 ± 0.015	_
Lock_2	_	$2.875^{+0.883}_{-0.700}$	2.416 ± 0.776	_
Lock_5	$3.649^{+0.461}_{-0.466}$	$3.625^{+0.558}_{-0.650}$	4.329 ± 0.062	3.36
Lock_102	_	$6.575^{+0.555}_{-0.700}$	4.742 ± 1.134	5.29
LSW_28	_	$4.425^{+0.507}_{-0.500}$	4.332 ± 1.104	_
Bootes_33	$3.367^{+0.874}_{-0.441}$	$2.625^{+0.522}_{-0.500}$	4.348 ± 0.147	_
HFLS_1	_	$3.675^{+0.424}_{-0.600}$	3.565 ± 1.079	4.29
HFLS_3	_	$5.375^{+1.106}_{-0.950}$	5.190 ± 0.896	6.34
HFLS_5	_	$4.125^{+0.589}_{-0.600}$	4.292 ± 1.112	4.44

Table 4.3: List of photometric redshifts of our sources given by EAZY, MMpz and CIGALE. Sources without optical photometry have no EAZY redshifts. Errors of EAZY and MMpz redshifts are given by the 68% credible interval, while CIGALE gives 1σ errors for photometric redshifts estimated. The spectroscopic redshifts, if known, is given in the last column. The names of sources with spectroscopic redshifts and the original literatures providing these redshifts are as follow: HFLS_1,3,5 and Lock_102 from Riechers et al. (2013); Lock_5 from Dowell et al. (2014); HeLMS_34 from Asboth et al. (2016); HeLMS_12 and HeLMS_28 from Riechers et al. (2021).

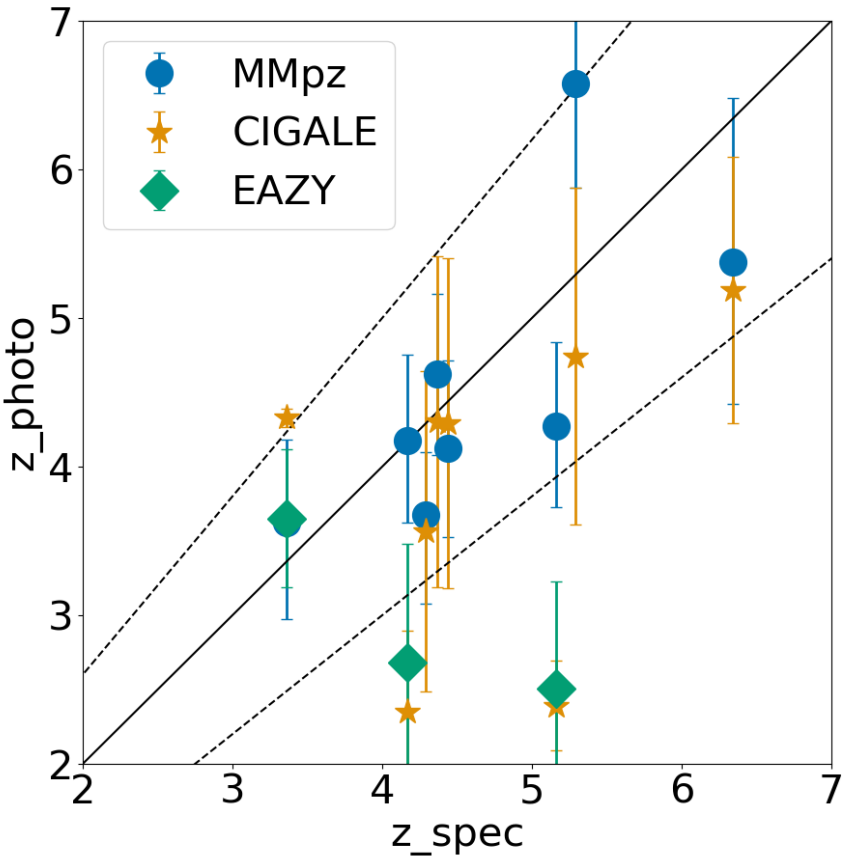


Figure 4.3: Comparison between spectroscopic redshifts and photometric redshifts derived by EAZY, MMPZ and CIGALE, for sources with spectroscopically confirmed redshifts. The solid line is the 100% accurate photometric redshift criteria, the dashed lines indicate $\frac{\Delta z}{(1+z)} = \pm 0.2$. As can be seen, MMPZ redshifts generally have better accuracy with smaller errors than CIGALE redshifts, while only 3 sources with spectroscopic redshifts have EAZY photo-z due to the lack of optical photometry.

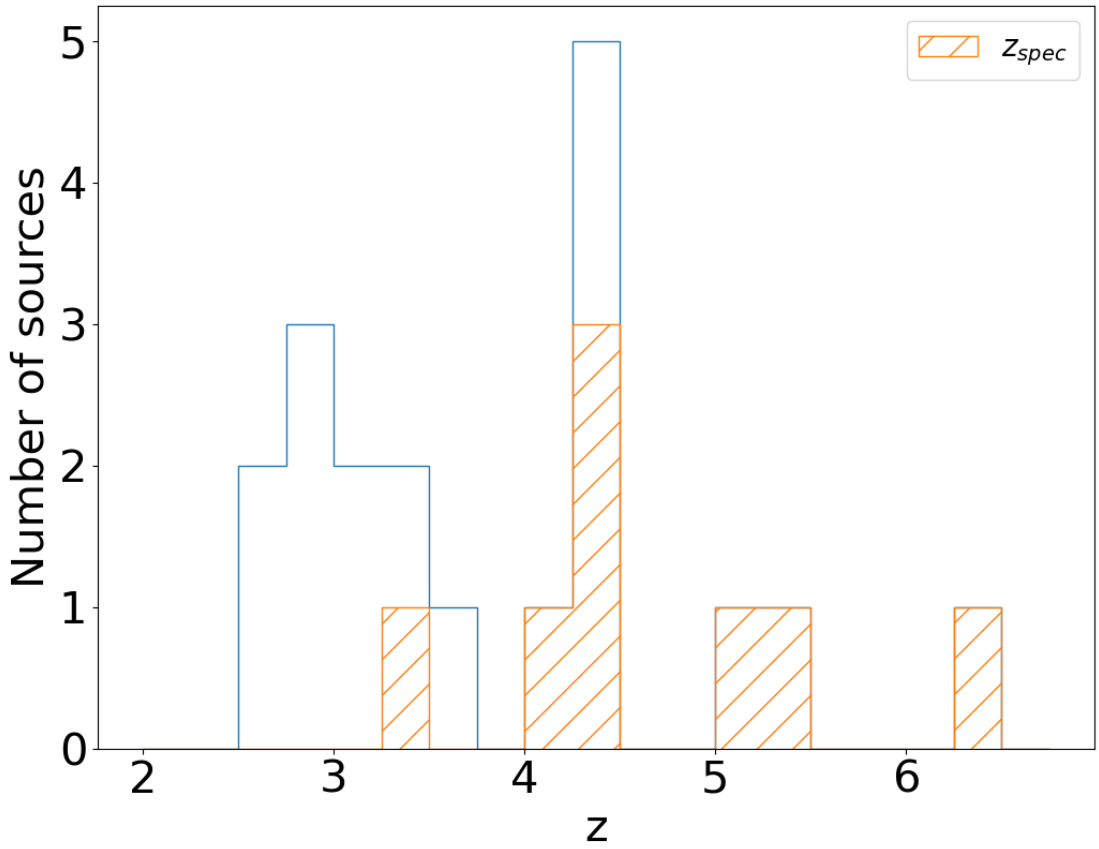


Figure 4.4: Redshift distribution of our 19 targets. The blue histogram shows the overall distribution, while the hatched orange histogram shows the redshift distribution of the 8 sources with spectroscopic redshifts. For all the other sources with no spectroscopic redshift, MMPZ redshifts are used. The average redshift of our sample is $< z >= 3.874 \pm 0.233$

4.5 CIGALE SED fitting

Various physical properties of our sources are calculated by CIGALE via SED fitting. To decide what CIGALE modules and parameters to use for more accurate derived physical properties, we refer to previous studies which use CIGALE to study FIR-selected galaxies (Barrufet et al., 2020) and submm-selected galaxies (Shim et al., 2022). Michałowski et al. (2014) has demonstrated that, for SMGs, a star formation history (SFH) with two decaying exponential components provides the most accurate derived stellar masses. This corresponds to the sfh2exp SFH module in CIGALE. With this module, one of the exponentials models the long-term star formation history which has produced the majority of the stellar mass, and the other exponential models the most recent burst of star formation. This model provides a simple but versatile way to fit the SFH of various galaxies. For example it can model the classical single exponential SFH by eliminating the second star burst exponential, or it can fit rising exponentials by setting the e-folding times of the stellar populations to a negative value. A wide range of stellar ages, e-folding time and stellar mass fraction of the late burst population is used to allow both moderate and intense starbursts to be computed.

The next module to be considered is the single stellar population library, which is used to calculate the intrinsic stellar spectrum. We adopt the Bruzual and Charlot (2003) library (bc03 in CIGALE) together with a Chabrier Initial Mass Function (Chabrier, 2003) as they are widely used in similar studies. The default 0.02 solar metallicity is used.

CIGALE includes a nebular module which takes account of lines and continuum emission originating from ionised gas surrounding massive stars. Such nebular emission provides insight into recent star formation and has a strong impact on SED modelling. The nebular module is the only nebular template adopted by CIGALE and is based on the empirical templates of Inoue (2011). The default parameters of this module are used as these give the best results as shown by Barrufet et al. (2020).

Dust attenuation/absorption of UV/optical emission is one of the key factors in fitting the SEDs of FIR-selected galaxies such as the 500 μm risers in this study. As demonstrated by both Barrufet et al. (2020) and Shim et al. (2022), the extended Charlot and Fall (2000) attenuation law module (dustatt_modified_CF00) appears to best fit the SEDs of FIR-selected galaxies. This module models

two different attenuation curves, one for the birth cloud of young stars and the other for the interstellar medium (ISM). In this case, the radiation of young stars is attenuated by both the birth cloud and the ISM, while stars older than 10 Myr only suffer from the attenuation of the ISM. A wide range of V-band attenuation coefficients in the ISM, A_V[1], is used. Furthermore, we use the dl2014 module (Draine and Li, 2007; Draine et al., 2013) to model MIR/FIR PAH emission as dust emission. This module involves two components of dust emission. One models diffuse dust emission from heating of the main stellar population, the others takes account of dust coupled to star-forming regions. As pointed out by Shim et al. (2022), this makes dl2014 a better option to fit SMG SEDs compared to other dust emission modules. Again, a large range of parameters is used for this module to allow a decent degree of freedom for the fitting.

Lastly, AGN contributions are computed with the fritz2006 module (Fritz et al., 2006). This models AGN emission as three components, including the radiation of the primary source in the dusty torus, the emission scattered by the torus and the emission of the torus. The angle between the AGN equatorial axis and the line of sight is set to range from 0.001 (type-1 AGN) to 89.999 (type-2 AGN), and the AGN fraction is set to range from 0.0 to 0.9.

The CIGALE modules and relevant parameters are summaried in Table 4.4. Since the photometric redshifts given by CIGALE are generally less accurate than those given by MMPz as Figure 4.3. shows, it is thus possible that the physical properties of our sources derived by CIGALE are less reliable as well when leaving redshift as a free parameter. Therefore, we performed another CIGALE run, using the known spectroscopic redshifts and the MMPz redshifts as input parameters for sources with and without spectroscopic redshifts respectively, other settings remained the same as described in Section 4.4 and Table 4.4. The resulting best-fitting SEDs of this run have lower χ^2 values for all sources compared to the run with no redshift constraint, proving the higher reliability of this run. Figure 4.5 shows the fitted SEDs for some example sources. It should be noted that, although CIGALE will determine the best-fit SED for each object, the derived physical properties of each object are not directly inferred from its best-fit SED. CIGALE will assign weights to all SEDs fitted to the target based on the fitting quality, then derive the probability distribution function for each physical property. Then

[1] A_V is defined as $A_V = m_V - m_{V,0}$ where m_V is the V-band apparent magnitude of a given galaxy, and $m_{V,0}$ is the intrinsic, unattenuated V-band magnitude of the galaxy.

Model and input parameters	Value	Comment
Star-formation history:	sfh2exp	
τ_{main}	1000, 3000, 5000	e-folding time of the main stellar population model [Myr]
τ_{burst}	30, 100, 9000	e-folding time of the late starburst population model in [Myr]
f_{burst}	0.0, 0.01, 0.1, 0.3, 0.5	Mass fraction of the late burst population
age	500, 1000, 3000, 5000, 10000	Age of the main stellar population in the galaxy [Myr]
$burst\_age$	50, 100, 300	Age of the late burst [Myr]
SFR_0	1.0	Value of SFR at $t = 0$ $[M_\odot/yr]$
Stellar population:	bc03	
IMF	1	Initial mass function: 0 (Salpeter) or 1 (Chabrier)
$metallicity$	0.02	Metallicity
$separation\_age$	10	Age [Myr] of the separation between the young and the old star population
Nebular emission:	nebular	
$logU$	−2.0	Ionisation parameter
Dust attenuation:	dustatt_modified_CF00	
Av_{ISM}	0.3, 0.6, 1.0, 1.6, 2.3, 3.0, 3.8, 5.0	V-band attenuation in the interstellar medium
Dust emission:	DL2014	
q_{PAH}	0.47, 1.12, 2.50, 3.90	Mass fraction of PAH
U_{min}	1.0, 5.0, 10.0, 25.0, 40.0	Minimum radiation field
AGN emission:	fritz2006	
τ	1.0, 6.0	Optical depth at 9.7 μm
$opening\_angle$	100	Full opening angle of the dust torus
ψ	0.001, 60.100, 89.990	Angle between equatorial axis and line of sight
f_{AGN}	0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9	AGN fraction

Table 4.4: CIGALE modules and parameters used in this work. Default values are used for parameters not listed in this table.

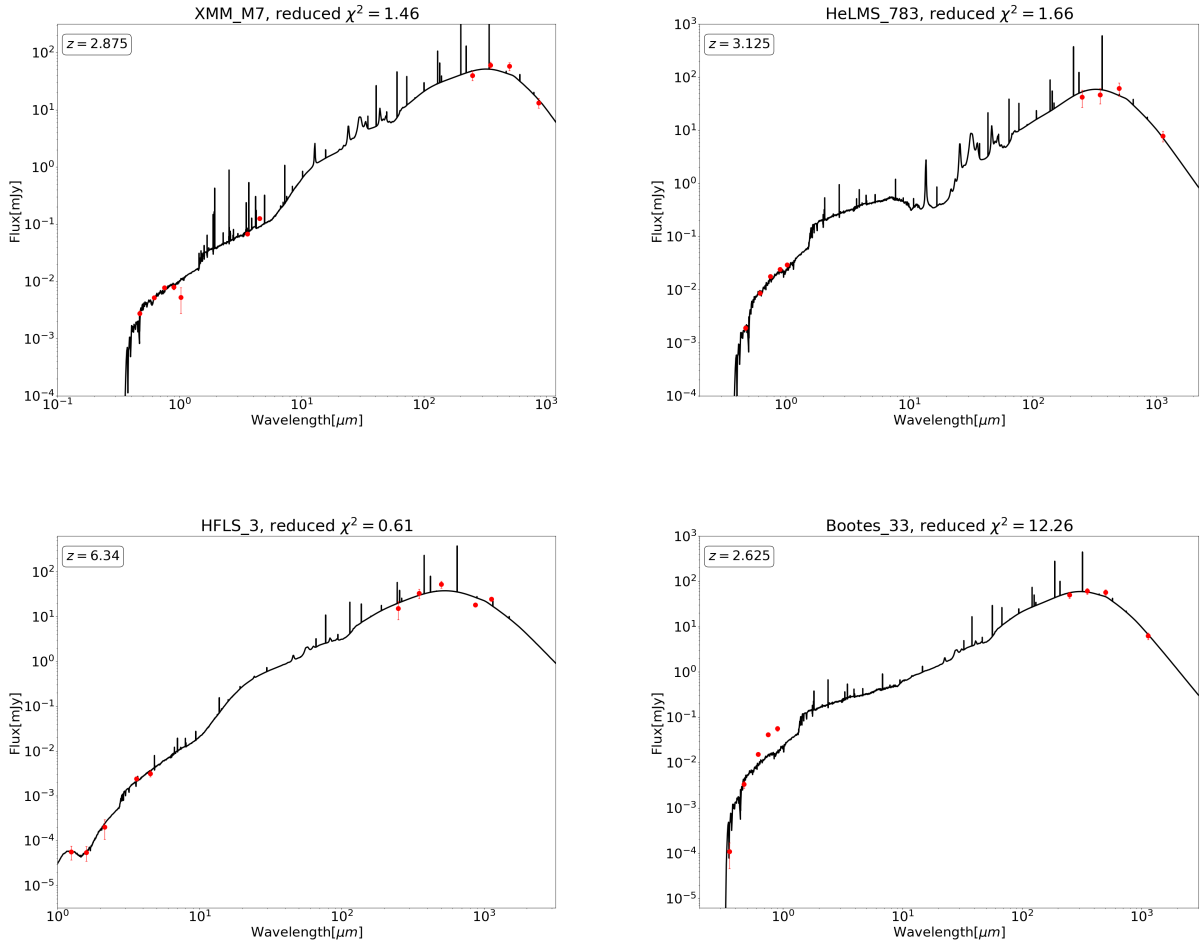


Figure 4.5: CIGALE best-fit SEDs for XMM_M7, HeLMS_783, HFLS_3 and Bootes_33 as examples. The red dots are the observed fluxes with associated uncertainties, their redshifts are shown in the top left corner. The reduced χ^2 values of these SEDs are also shown. The number of photometric points used in the SED fitting is varying for different targets. For the SED fitting of all sources, their redshifts are fixed to their MMPZ redshifts and spectroscopic redshifts for sources with and without spectroscopic redshifts respectively.

the final value and associated error of each property are calculated based on the likelihood-weighted mean and standard deviation.

4.6 Validating CIGALE fits

We need to validate the reliability of our targets' physical properties derived by CIGALE. The structure of the CIGALE algorithm provides a way to test if the physical properties can be reliably retrieved in a self-consistent way (Boquien et al., 2019). The idea is to construct a mock catalogue by using

photometric points obtained from the best-fit SED for each object, run CIGALE with this mock catalogue with the exact same setting as the previous run, then compare physical properties derived in this run to those derived in the original run which are already known. If the physical properties derived in the two separate runs are similar, it demonstrates that the estimations of these parameters are at least reliable in a self-consistent sense. The photometric points in the mock catalogue are obtained from the best fits in the first run, with noise injected to simulate new observations.

Figure 4.6 shows the results of this mock test. The square of the linear correlation coefficient r^2 of each parameter is shown in the title of its panel. For SFR, IR luminosity, stellar luminosity and dust mass, the r^2 is very close to one which indicates these parameters are reliably estimated by CIGALE. The discrepancy between the original and mock values of stellar mass and gas mass are larger than the aforementioned four parameters, with lower r^2 of around 0.75. This means that there is at least a small fraction of sources whose stellar masses and gas masses are less reliably derived. AGN fraction has the worst reliability with $r^2 = 0.60$. This is probably due to the lack of MIR photometry for most of our sample which corresponds to the peak of the AGN dust torus emission. Nonetheless, the original and mock values of AGN fraction are still positively correlated, showing that its estimation is reliable to some extent and can be used in some rough analysis or classification of subsamples.

In short, the results of the mock test show that SFR, IR luminosity, stellar luminosity and dust mass are reliably estimated by CIGALE. While stellar mass and gas mass have slightly lower reliability but still acceptable, care should be taken when using them in analysis. It should also be noted that the results of this mock test are dependent on the breadth and quality of the photometric coverage, such that it may produce larger discrepancy for objects with poorer photometry.

4.7 Physical Properties

The CIGALE derived physical properties are shown in Table 4.5. We also include physical properties of HFLS_3 given by Riechers et al. (2013) in Table 4.5 as a benchmark. It can be seen that our L_{IR}, M_{Dust} and SFR of HFLS_3 roughly agree with those in Riechers et al. (2013), with slightly higher means. Meanwhile, our stellar mass of HFLS_3 is two orders of magnitude higher than the

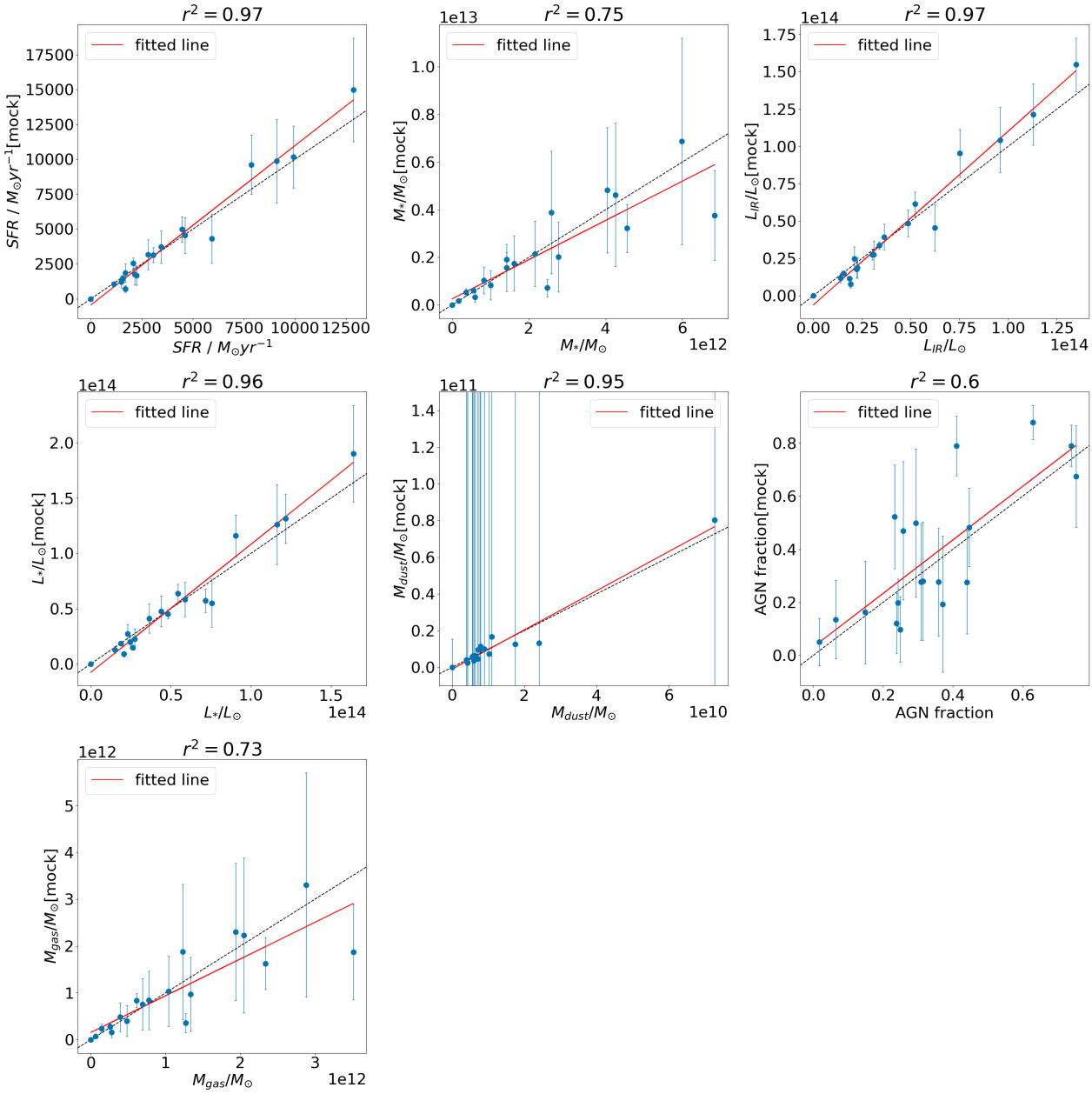


Figure 4.6: Comparison between the original (x-axis) and the mock (y-axis) values of SFR, stellar mass (M_*), IR luminosity (L_{IR}), stellar luminosity (L_*), dust mass (M_{dust}), AGN fraction (f_{AGN}) and gas mass (M_{gas}) from top left to bottom right. The blue dots indicate our sources, the red solid line is the fitted linear correlation and the black dashed line is the one-to-one line. The square of linear correlation coefficient of each parameter is shown in the title of the corresponding panel. SFR, IR luminosity, stellar luminosity and dust mass can be reliably measured by CIGALE, while estimations of stellar mass and gas mass are slightly less robust but still acceptable. Measurements of AGN fraction have the worst consistency, therefore care should be taken when using this parameters.

Source	$L_{IR}/L_{\odot}$	$L_*/L_{\odot}$	SFR/$M_{\odot}yr^{-1}$	$M_*/M_{\odot}$	$M_{Dust}/M_{\odot}$	f_{AGN}
XMM_M7	$1.546 \pm 0.187 \times 10^{13}$	$1.869 \pm 0.192 \times 10^{13}$	1565 ± 283	$5.600 \pm 0.836 \times 10^{11}$	$6.464 \pm 2.215 \times 10^{9}$	0.447 ± 0.181
XMM_M2	$2.096 \pm 0.929 \times 10^{13}$	$2.299 \pm 0.982 \times 10^{13}$	1695 ± 718	$8.266 \pm 5.692 \times 10^{11}$	$5.557 \pm 2.222 \times 10^{9}$	0.371 ± 0.314
XMM_26	$1.910 \pm 0.841 \times 10^{13}$	$2.061 \pm 0.895 \times 10^{13}$	1690 ± 795	$5.997 \pm 4.606 \times 10^{11}$	$6.060 \pm 3.247 \times 10^{9}$	0.630 ± 0.249
HeLMS_11	$1.390 \pm 0.214 \times 10^{13}$	$1.514 \pm 0.222 \times 10^{13}$	1489 ± 338	$1.684 \pm 0.384 \times 10^{11}$	$7.620 \pm 3.389 \times 10^{9}$	0.740 ± 0.078
HeLMS_12	$9.590 \pm 2.072 \times 10^{13}$	$1.162 \pm 0.338 \times 10^{14}$	9101 ± 2797	$4.254 \pm 2.795 \times 10^{12}$	$2.409 \pm 0.740 \times 10^{10}$	0.309 ± 0.215
HeLMS_28	$1.128 \pm 0.204 \times 10^{14}$	$1.217 \pm 0.221 \times 10^{14}$	9919 ± 2339	$4.047 \pm 2.508 \times 10^{12}$	$1.749 \pm 0.555 \times 10^{10}$	0.149 ± 0.192
HeLMS_32	$2.225 \pm 0.672 \times 10^{13}$	$2.448 \pm 0.694 \times 10^{13}$	2230 ± 857	$3.617 \pm 0.824 \times 10^{11}$	$1.021 \pm 0.399 \times 10^{10}$	0.410 ± 0.269
HeLMS_34	$7.518 \pm 1.982 \times 10^{13}$	$9.048 \pm 0.225 \times 10^{13}$	7859 ± 2385	$2.588 \pm 2.008 \times 10^{12}$	$8.906 \pm 3.615 \times 10^{9}$	0.440 ± 0.205
HeLMS_62	$3.093 \pm 1.184 \times 10^{13}$	$3.653 \pm 1.440 \times 10^{13}$	2807 ± 1144	$1.424 \pm 1.021 \times 10^{12}$	$7.845 \pm 6.451 \times 10^{9}$	0.242 ± 0.098
HeLMS_783	$3.388 \pm 0.169 \times 10^{13}$	$4.808 \pm 0.240 \times 10^{13}$	2066 ± 103	$4.548 \pm 0.268 \times 10^{12}$	$3.970 \pm 1.378 \times 10^{9}$	0.064 ± 0.093
Cosmos_9	$6.838 \pm 7.683 \times 10^{9}$	$7.162 \pm 7.884 \times 10^{9}$	0.595 ± 0.706	$2.079 \pm 2.448 \times 10^{8}$	$5.442 \pm 12.664 \times 10^{6}$	0.754 ± 0.149
LSW_28	$4.866 \pm 0.141 \times 10^{13}$	$5.887 \pm 2.066 \times 10^{13}$	4606 ± 1684	$2.159 \pm 1.498 \times 10^{12}$	$1.095 \pm 0.424 \times 10^{10}$	0.360 ± 0.246
Lock_2	$2.255 \pm 0.541 \times 10^{13}$	$2.746 \pm 0.856 \times 10^{13}$	2154 ± 705	$1.000 \pm 0.672 \times 10^{12}$	$4.256 \pm 1.264 \times 10^{9}$	0.294 ± 0.231
Lock_5	$3.022 \pm 0.151 \times 10^{13}$	$7.168 \pm 0.358 \times 10^{13}$	3050 ± 153	$6.845 \pm 0.355 \times 10^{12}$	$4.281 \pm 1.993 \times 10^{9}$	0.017 ± 0.045
Lock_102	$1.348 \pm 0.210 \times 10^{14}$	$1.641 \pm 0.408 \times 10^{14}$	12865 ± 3434	$5.984 \pm 3.805 \times 10^{12}$	$7.271 \pm 11.734 \times 10^{10}$	0.238 ± 0.167
Bootes_33	$1.862 \pm 0.262 \times 10^{13}$	$2.631 \pm 0.253 \times 10^{13}$	1119 ± 139	$2.475 \pm 0.196 \times 10^{12}$	$5.685 \pm 1.338 \times 10^{9}$	0.234 ± 0.182
HFLS_1	$6.238 \pm 1.164 \times 10^{13}$	$7.562 \pm 2.039 \times 10^{13}$	5923 ± 1704	$2.766 \pm 1.784 \times 10^{12}$	$7.120 \pm 2.015 \times 10^{9}$	0.258 ± 0.209
HFLS_3	$5.229 \pm 0.886 \times 10^{13}$	$5.435 \pm 0.913 \times 10^{13}$	4459 ± 925	$1.419 \pm 0.291 \times 10^{12}$	$6.0.9 \pm 2.466 \times 10^{9}$	0.249 ± 0.151
HFLS_3 (Riechers et al., 2013)	$2.86^{+0.32}_{-0.31} \times 10^{13}$	_	2900	$\sim 3.7 \times 10^{10}$	$1.31^{+0.32}_{-0.30} \times 10^{9}$	_
HFLS_5	$3.639 \pm 0.979 \times 10^{13}$	$4.394 \pm 1.456 \times 10^{13}$	3434 ± 1192	$1.615 \pm 1.098 \times 10^{12}$	$7.120 \pm 2.729 \times 10^{9}$	0.314 ± 0.236
Lock_102*	$1.078 \pm 0.168 \times 10^{13}$	$1.313 \pm 0.326 \times 10^{13}$	1029 ± 275	$4.787 \pm 3.044 \times 10^{11}$	$5.817 \pm 9.387 \times 10^{9}$	0.238 ± 0.167
HFLS_3*	$2.377 \pm 0.403 \times 10^{13}$	$2.470 \pm 0.415 \times 10^{13}$	2027 ± 420	$6.451 \pm 1.323 \times 10^{11}$	$2.745 \pm 1.121 \times 10^{9}$	0.249 ± 0.151

*: lensing corrected

Table 4.5: IR luminosities, stellar luminosities, SFR, stellar masses, dust masses and AGN fractions of our targets derived by CIGALE, values in the upper part of the table are uncorrected for lensing magnification. The physical properties of HFLS_3 measured and derived by Riechers et al. (2013) are also listed here for comparison. Note that the physical properties of Cosmos_9 are inconsistent with the others, see Section 4.8.3 for more discussions on this source. Lock_102 is known to be strongly lensed with a magnification factor $\mu = 12.5 \pm 1.2$ (Riechers et al., 2020)), and HFLS_3 is known to be weakly lensed with $\mu = 2.2 \pm 0.3$ (Cooray et al., 2014). The lower part of the table displays our derived properties of Lock_102 and HFLS_3 corrected for lensing.

literature's stellar mass. The disagreement between our derived properties of HFLS_3 and those of Riechers et al. (2013) may be a result of the difference in SED fitting techniques implemented and the lack of optical photometry, we discuss this later in Section 4.7.2. The lensing corrected physical properties of two sources which are known to be lensed, Lock_102 and HFLS_3, are also shown in Table 4.5. Some sources such as HeLMS_12 and HeLMS_28 also have very high derived properties, implying that they could be gravitationally lensed as well.

4.7.1 IR Luminosity

CIGALE estimates the dust luminosities of our sources, which can approximate the total IR luminosities for DSFGs (Draine, 2003b). Figure 4.7. shows the total IR luminosity distribution of our sample as a function of redshift. Note that the L_{IR} shown in Figure 4.7 is not lensing corrected, as for most of our samples we have no knowledge if they are lensed or not, and correction of lensing for the two lensed sources (Lock_102 and HFLS_3) does not significantly change the overall results as well. The

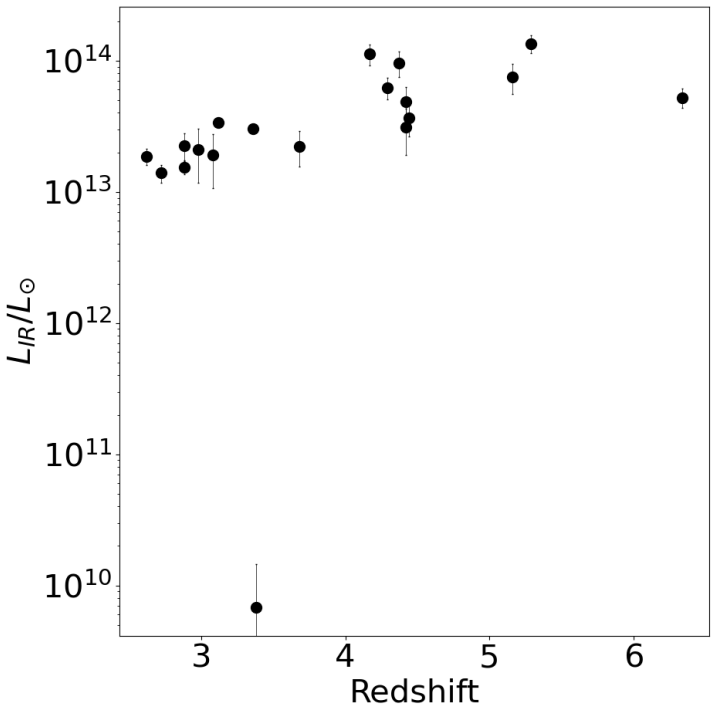


Figure 4.7: The total IR luminosities of our sample as a function of redshift without correction for lensing. Most sources have $L_{IR} > 10^{13}$ comparable with local ULIRGs. The outlier near the bottom of the figure with large errors is Cosmos_9.

total IR luminosities of the sources are in a range between 10^{13} $L_{\odot} < L_{IR} < 10^{14.5}$ $L_{\odot}$ (except one outlier), consistent with the expectation that 500μm risers are ULIRGs with $L_{IR} > 10^{12}$ $L_{\odot}$. There is no significant trend between the IR luminosity and the redshift. As one can notice, there is one outlier near the bottom of the figure which is Cosmos_9. Cosmos_9 has a CIGALE SED fit that failed to match its FIR photometry points, making it an outlier in every analysis. We will discuss this source in more detail in Section 4.8.3. There is no apparent correlation between L_{LR} and redshift, as expected from the very negative K-correction of SMGs.

4.7.2 SFR and Stellar Mass

SFR and stellar mass are crucial properties to understand galactic evolution, as they can indicate whether a galaxy is in a growing stage with rapid star formation activity, or in a quiet and mature stage with low SFR but high stellar mass. A well-known correlation exists between star-forming galaxies' SFR and stellar masses called the star-forming main sequence, which evolves with redshift

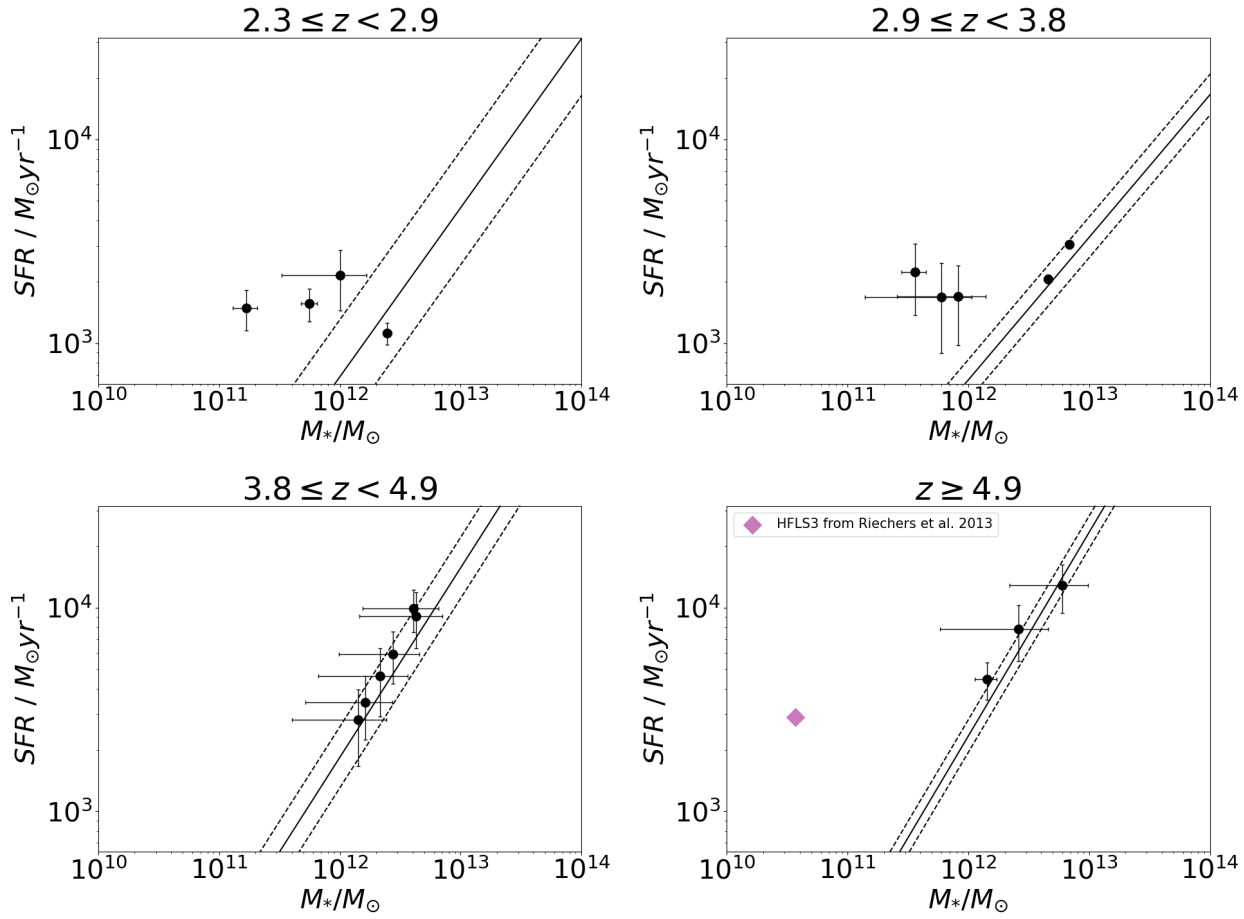


Figure 4.8: SFR vs.stellar mass in different redshift bins, uncorrected for lensing. Respective star-forming main sequences from Pearson et al. (2018) are plotted as solid lines at each redshift bin, while the dashed lines show their 1σ scatters. At redshift bins $2.3 \leq z < 2.9$ and $2.9 \leq z < 3.8$ most of our sources are located above the main sequence, while in the other redshift bins, the SFR and stellar masses of our sample match the main sequence. The purple diamond is the SFR$/M_*$ of HFLS_3 given by Riechers et al. (2013). Cosmos_9 is excluded.

(e.g. Speagle et al. 2014; Pearson et al. 2018). Figure 4.8. presents the relationship between SFR and stellar mass for our targets, divided into four redshift bins. Pearson et al. (2018) traces the main sequence over $0.2 \leq z \leq 6$ using the latest de-blended *Herschel* photometry. The main sequence derived by them is shown in each bin for comparison. The mean SFR of our sample is $4107 \pm 800\ M_{\odot} yr^{-1}$, and the mean stellar mass $< M_* >= 2.297 \pm 0.459 \times 10^{12}\ M_{\odot}$. Except for one outlier at the $2.9 \leq z < 3.8$ bin which is Cosmos_9, all sources have their SFR$> 10^3\ M_{\odot} yr^{-1}$ and stellar mass $M_* > 10^{11}\ M_{\odot}$ which is consistent with the prediction that 500μm risers are massive galaxies with intense star-forming activities, or strongly lensed sources. At $2.3 \leq z < 2.9$, three of the four sources lie above the main sequence, implying they are forming stars more rapidly than the main sequence predictions; similarly, at $2.9 \leq z < 3.8$, 3 of our targets in this bin have higher SFR than expected, while the other two sources are consistent with the main sequence. Cosmos_9 has a redshift in this range with very low SFR and stellar mass, it is not shown in the figure. At redshift bins $3.8 \leq z < 4.9$ and $z \geq 4.8$, the SFR and stellar masses of all our targets match the main sequence. Also plotted as a purple diamond is HFLS_3 in the $z \geq 4.8$ panel, with the SFR and stellar mass in Riechers et al. (2013). It is well above the main sequence, indicating it as a very intense starburst. Note that our derived stellar mass of HFLS_3 is two orders of magnitude higher than that in Riechers et al. (2013), thus our results of HFLS_3 lie on the main sequence. This could be a result of different strategies and data used to derive its physical properties. While Riechers et al. (2013) also use CIGALE to estimate the stellar mass of HFLS_3, they do not specifically describe their choices of CIGALE modules and parameters and are thus likely to be at least partially different to ours, which could lead to very different results. The lack of optical photometry will further enlarge the difference in stellar masses produced by the two methods, as the optical photometry is a strong constraint of stellar mass.

The fact that sources lying above the main sequence are mainly located at redshift $z < 3.8$ implies that either: (a) most 500μm risers which lie at $z < 3.8$, especially those lie at $z < 2.9$, are more intense starbursts than typical star-forming galaxies; (b) we have underestimated the redshifts of these sources, this can be better seen in Figure 4.9 where sources with photometric redshift $z < 3.8$ are better matched to the $z \geq 3.8$ main sequences. Since we expect 500μm risers to lie at $z > 4$ from their SED shapes, the latter case should be more likely to be true. This is partially supported by the fact that 7 of the 8 spectroscopically confirmed sources in our 500 μm riser sample lie at $z > 4$.

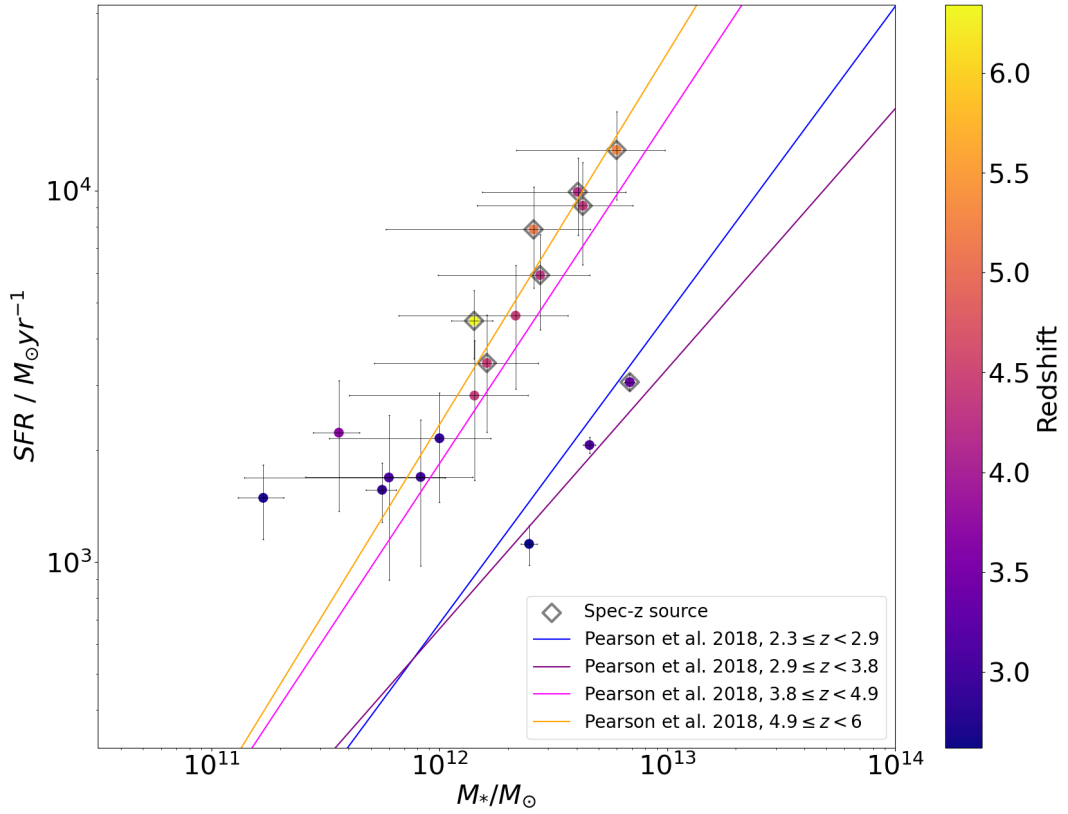


Figure 4.9: SFR vs.stellar mass for all our targets, Cosmos_9 ignored. Colours of the markers represent their redshifts, sources with spectroscopic redshifts are marked with grey diamonds. The main sequences from Pearson et al. (2018) are shown as solid lines, with their colours roughly corresponding to the redshift colour scheme of the markers. As can be seen, most sources are matched to the $3.8 \leq z < 4.9$ and $4.9 \leq z < 6$ main sequences, implying either these sources are more intense starbursts, or they lie at higher redshift than their photometric redshifts estimated here.

This requires more consistent and complete photometry for better photometric redshift estimations, or spectroscopic follow-up observations to confirm their redshifts. However, the possibility that $500\mu m$ risers are more extreme starbursts cannot be ruled out, as even if sources which lie above their main sequences are truly at $z > 3.8$, some of them still lie above the high-z main sequences as can be seen in Figure 4.9, indicating intense star formation activities.

4.7.3 The impact of AGN on star formation

We investigate the effect of AGN on the SFR of our 500 μm risers. As mentioned in Section 4.6, our AGN fractions derived by CIGALE are less reliable compared to other parameters, therefore here we only use f_{AGN} to classify our sample into 'high' and 'low' AGN fraction subsamples. We adopt a f_{AGN} = 0.3 threshold following Serjeant et al. (2010). Sources with $f_{AGN} \geq 0.3$ are the high AGN fraction sample and vice versa. Figure 4.10 shows the SFR of our 500 μm risers as a function of redshift, divided by their AGN fractions. The average SFR for the high AGN fraction sample and the low AGN fraction sample is 3367 ± 941 and 4929 ± 1331 respectively. Given the large scatters, there is no significant correlation between the AGN fraction and SFR, implying that SFR of 500 μm risers will not be strongly influenced by AGN luminosities. This agrees with the results of Barrufet et al. (2020) and Shim et al. (2022), and is also consistent with studies on X-ray selected AGN which do not find any correlation between X-ray luminosity and SFR of SMGs. (Mullaney et al., 2012; Stanley et al., 2015; Ramasawmy et al., 2019).

4.8 Discussion

4.8.1 Reliability of Photometric Redshifts and SED Fittings

Although the multi-wavelength data used in this chapter lead to much more accurate photometric redshifts and derived physical properties of our targets than those estimated by Greenslade et al. (2020), the lack of certain photometric bands for some targets and heterogeneity of the photometry used for each target limit the reliability of the results. Around 36.8% of our sources have no optical identification and ~ 73.7% have no NIR/MIR photometry. For sources with optical and NIR/MIR counterparts, the origins of the photometry come from various databases such as DES and SDSS with different photometric band coverage and depths. The lack of optical/NIR/MIR data will strongly affect the quality of photometric redshift estimations and SED fits, as these wavelengths cover the Lyman break for high redshift star-forming galaxies which is an important feature for redshift estimation (e.g. Burgarella et al. 2011). Therefore, we need to dig wider and deeper into public-available

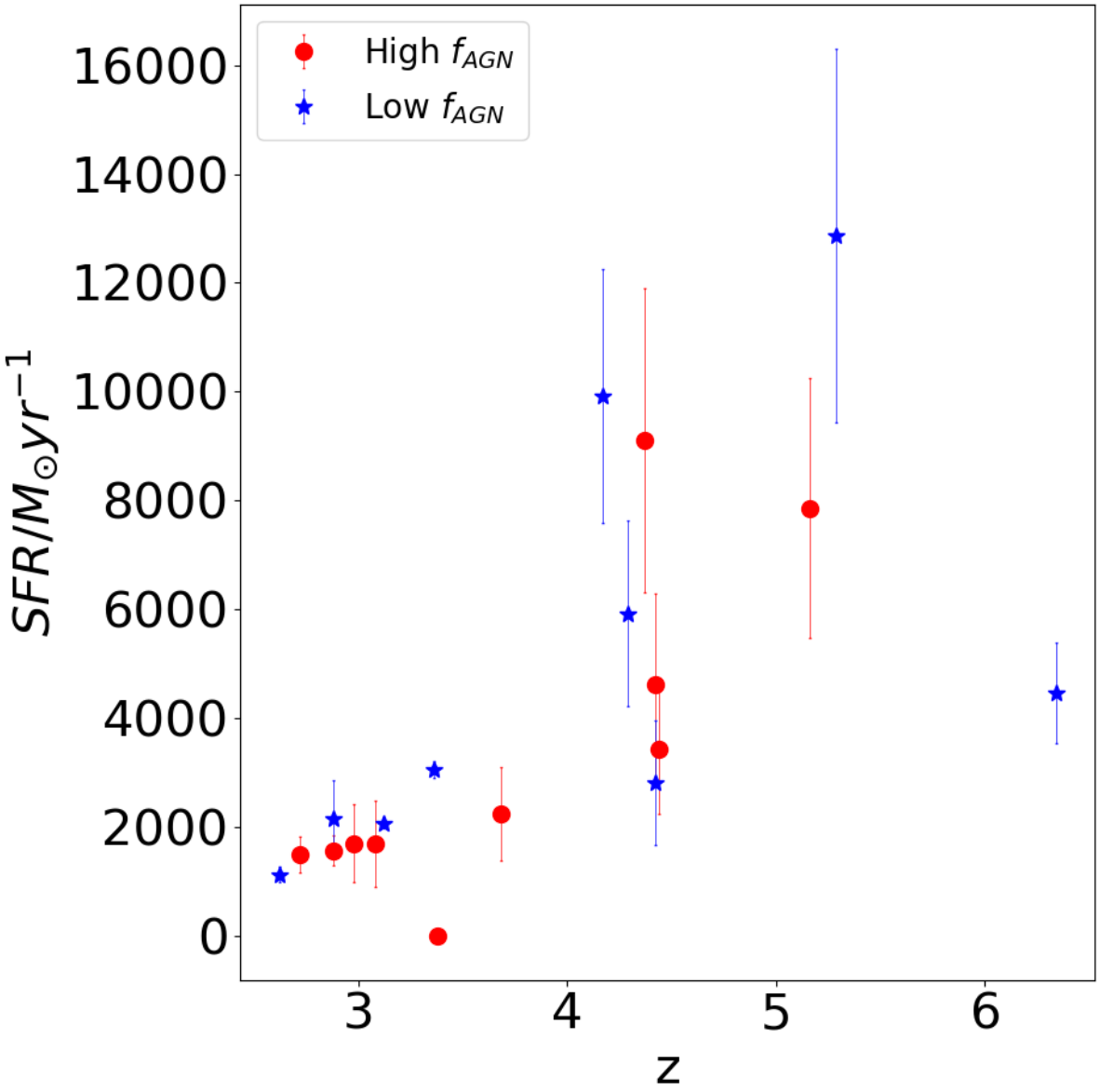


Figure 4.10: SFR as a function of redshift. The red dots are sources with high AGN fraction and blue stars are sources with low AGN fraction. No apparent correlation is seen between SFR and AGN fraction.

databases to find better photometry. However, it is also hard to find optical/NIR data for all our 500μm risers, as they may be invisible or extremely faint in these wavelengths due to heavy dust obscuration. For example, Riechers et al. (2013) carried out dedicated deep optical observation on HFLS_3, and found no significant detection. This thus imposes great difficulty in finding complete and consistent optical/NIR photometry for our sources. This may also explain why MMPz, which is designed to deal with FIR/submm bright star-forming sources without optical/NIR data, gives the best redshift estimations.

As the next step to improve the quality of our results, a possible choice is to include radio data in our cross-identifications. Radio emission from star-forming galaxies is known to have a correlation with their FIR emission (e.g., Helou et al. 1985; De Jong et al. 1985) which can help to constrain the SEDs fitted to our targets. With the next generation radio telescope Square Kilometre Array becoming closer and closer to reality, high quality radio data will play a crucial role in FIR/submm studies.

4.8.2 Gravitational Lensing in 500 μm Risers

Béthermin et al. (2017) predicts via simulations that all sources with $S_{500} > 60\,mJy$ are strongly lensed. Greenslade et al. (2020) only have extended array observations for HFLS_1, HFLS_3, HFLS_5, Lock_5 and Lock_102, which means that only these 5 sources were observed with resolution high enough to examine potential features of lensing. Lock_102 is known to be strongly lensed with a magnification $\mu = 12.5 \pm 1.2$ (Riechers et al., 2020), which was confirmed by the SMA observation (top right panel in Figure 4.1.). HFLS_3 is also known to be weakly lensed with a magnification $\mu = 2.2 \pm 0.3$ (Cooray et al., 2014). Greenslade et al. (2020) also points out that there are signs of extension seen in HFLS_1 and Lock_5, implying possible weak lensing. Oteo et al. (2017) finds that 40% of their 500 μm risers show some lensing signatures with ALMA high-resolution observations. Both works find that the lensing fraction in $S_{500} > 60\ mJy$ 500 μm risers is significantly lower than Béthermin et al. (2017)'s prediction.

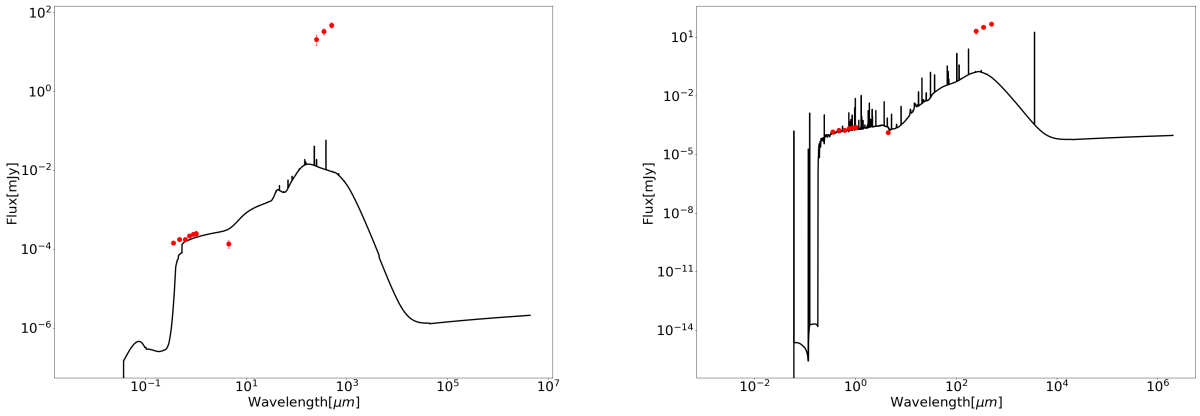
Although the effect of gravitational lensing in 500 μm risers is not extensively investigated in this chapter, the derived physical properties of some of our sources, such as HeLMS_28 which has

$L_{IR} > 10^{14} M_{\odot}$, indicate that these sources are potentially strongly lensed (Table 4.5). We further investigate if there are any potential lensed sources by re-running CIGALE using only the optical/NIR photometry of our sample. The physical properties derived in this CIGALE run suggest that XMM_M2, HeLMS_32, HeLMS_783, Lock_5 and Bootes_33 may be strongly lensed, as they have extremely large stellar mass of $M_* > 10^{13} M_{\odot}$ and IR luminosity of $L_{IR} > 10^{14} M_{\odot}$. It is definitely an intriguing direction to explore in the future with high-resolution observations.

4.8.3 Cosmos_9

As mentioned in previous sections, the derived physical properties of Cosmos_9 are very different to those of the others: it has a $L_{IR} \approx 7 \times 10^9\ L_{\odot}$ which is ~ 4 orders of magnitude smaller than the other sources' IR luminosities (Figure 4.7.); its SFR and stellar mass are 0.6 $M_{\odot} yr^{-1}$ and $1.7 \times 10^8\ M_{\odot}$ while the others have SFR and stellar mass of $\sim 10^{3-4}\ M_{\odot} yr^{-1}$ and $\sim 10^{11-13}\ M_{\odot}$ (Figure 4.8.). These peculiar numbers are probably results of a failed SED fit given by CIGALE, as Figure 4.11a. illustrated. CIGALE failed to construct an SED to match the FIR photometry points, leading to unexpected results.

An initial guess at the reason for this failure is that there is a foreground optical source coinciding with the FIR source, thus the optical/NIR misguides CIGALE to fit a low-redshift, non-star-forming SED to Cosmos_9. This is partially proven by the CIGALE best-fit SED of Cosmos_9 when redshift is left as a free parameter, shown in Figure 4.11b. When redshift is set to be a free parameter, CIGALE builds an SED which matches the optical/NIR photometry better but fails to fit the FIR data. The photometric redshift estimated with this SED is 1.002 ± 0.015. The COSMOS2020 catalogue includes its own EAZY photo-z estimations, the redshift of Cosmos_9 derived by them is 1.054, fairly consistent with our CIGALE estimation. Meanwhile, MMPZ, which only takes FIR/submm data, estimates that Cosmos_9 lies at $z = 3.375$. All these suggest that the optical/NIR counterpart of Cosmos_9 could be a different source in front of our SMA detection, and the FIR/submm flux density of the background source could be lensed by this foreground optical source, hence using its photometry will indeed result in a low-redshift SED.



(a) best-fit SED with MMPZ redshift

(b) best-fit SED with redshift as a free parameter

Figure 4.11: CIGALE best-fit SEDs for Cosmos_9, with the observed fluxes and uncertainties shown as red dots.(a) best-fit SED with MMPZ redshift provided; (b) best-fit SED with redshift as a free parameter, the photometric redshift estimated with this SED is 1.002 ± 0.015. As can be seen, in either case, CIGALE failed to construct an SED to match the photometry points.

Visual examination of Cosmos_9 does not provide sufficient evidence to determine whether the cross-identification of Cosmos_9 is correct, as illustrated in Figure 4.12. The angular distance between the centre of the SMA detection and the centre of the HSC source is $\sim$ 1.42″, and there is no apparent gravitational lensing feature in the SMA map. It is therefore not clear whether the optical source is truly a different foreground object, and if so, whether Cosmos_9 is lensed by the foreground source. We may require further observations to clarify the nature of Cosmos_9, such as deeper and higher-resolution submm/mm images, or spectroscopically confirming its redshift.

4.9 Conclusions

In this chapter, we utilise public databases to perform optical to MIR cross-identifications of 19 500μm risers detected by the SMA as single sources, selected from Greenslade et al. (2020). Based on the multi-wavelength photometry built from the cross-identifications, EAZY, MMPZ and CIGALE were used to estimate photometric redshifts of these 19 sources, we find that MMPZ provides the most reliable estimations but with large uncertainties. Overall, we find a mean redshift of $< z >= 3.874 \pm 0.233$ for our 500 μm riser sample with single SMA detection. 10 of the 19 sources have $2.6 < z < 3.7$,

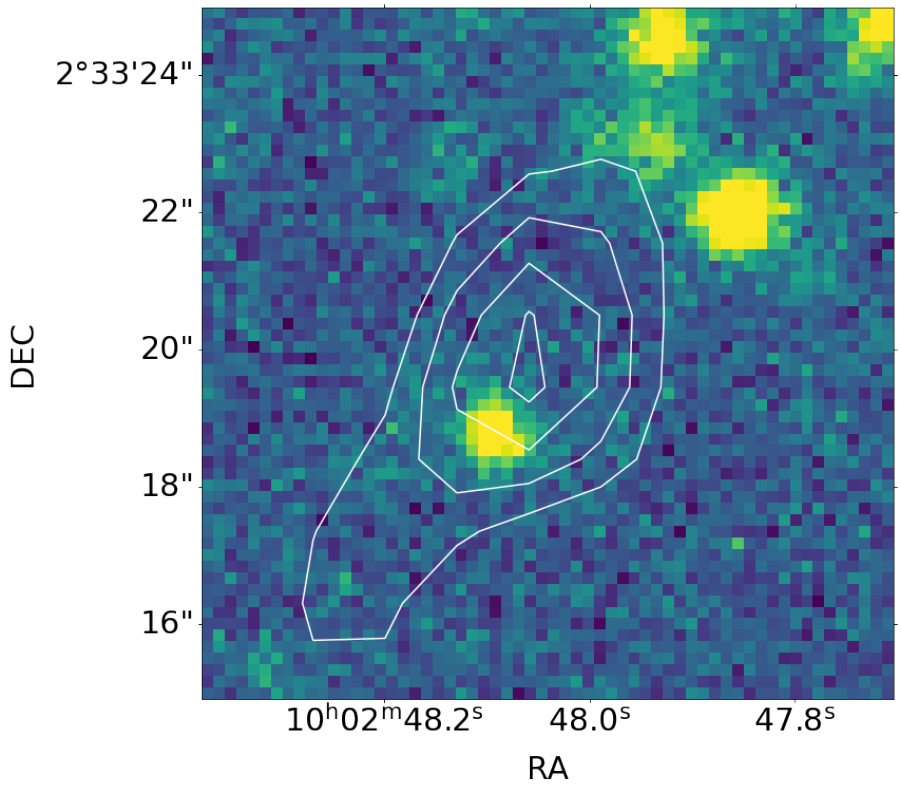


Figure 4.12: The optical *i* band image of Cosmos_9 from the HSC Public Data Release 3 Ultra Deep image (Aihara et al., 2022), overlaid with the white SMA 230.4 GHz contour. The angular distance between the centre of the SMA detection and the centre of the HSC contour is $\sim 1.42''$. No apparent gravitational lensing feature is found.

lower than the $z > 4$ expectation. By inputting known spectroscopic redshifts and MMPZ photometric redshifts, we carry out SED fitting to target objects with CIGALE, deriving their physical properties such as L_{IR}, SFR and stellar mass. The mean SFR and stellar mass of our sample are $4107 \pm 800\ M_{\odot} yr^{-1}$ and $< M_* >= 2.297 \pm 0.459 \times 10^{12}\ M_{\odot}$, with $L_{IR} > 10^{13}\ L_{\odot}$ for nearly all sources. The IR luminosities of our 500 μm risers suggest that they are ULIRGs, while their SFR-stellar mass relation implies that either $z < 3.8$ 500 μm risers are starbursts which are more intense than typical star-forming galaxies at similar redshifts, or their redshifts are underestimated. We also find that the fraction of AGN contribution has no significant impact on the SFR of our sample, consistent with previous studies on X-ray selected AGN of SMGs.

An unusual case is Cosmos_9 whose derived physical properties are atypical of star-forming galaxies. We suspect that there is a foreground optical/NIR object coinciding with the FIR source which could be lensed by the optical/NIR source, thus the optical/NIR photometry of this foreground object yields unusual results. Future high-resolution observations may be needed to verify whether this is truly the case.

This study is limited by the relatively incomplete and heterogeneous multi-band photometry, due to the wide range of sky locations of our sample and the fact that the optical/NIR counterparts of high-redshift DSFGs are generally too faint to be detected for most large area surveys. As the next step, we can exploit more future public survey data to construct a multi-band photometry of 500 μm risers that is as complete and homogenous as possible. In particular, new data from JWST, Euclid and Vera C. Rubin Observatory LSST, which have outstanding image depths in the optical/NIR, will significantly help the identification of optically faint 500 μm risers. Another possible improvement is to include radio data. As radio emission from star-forming galaxies is known to have a correlation with their FIR emission, it will thus be more likely to find counterparts of our 500μm risers in radio data, which can help to constrain the SEDs fitted to our targets. Radio identifications of our targets will also significantly help us to find their optical/NIR counterparts, as discussed in Chapter 2. Also, more 500 μm risers can be added to our sample, such as those selected in Oteo et al. (2017). Building a larger sample will minimise the effect of outliers and make the results and conclusions more statistically significant. Lastly, spectroscopic observations of more 500 μm risers will enable precise determination of their

redshifts, which will greatly improve the quality of SED fitting and the physical properties derived.

Chapter 5

Conclusion

5.1 Summary of Thesis Achievements

The goal of this thesis is to investigate the physical properties of DSFGs, the environments associated with them, and ultimately to enhance our understanding of the role of DSFGs in galaxy formation and evolution.

In Chapter 2, we present MeerKAT 1.28GHz radio imaging of 3 *Herschel* candidate protoclusters, with a synthesised beam size of $\sim 7.5'' \times 6.6''$ and a central thermal noise down to $4.35 \mu Jy/beam$. Our source counts are consistent with other radio counts with no evidence of overdensities, source counts just in the central $4.63'$ radius regions where the candidate protoclusters lie do not reveal any evidence of overdensity either. Around 95% of the *Herschel* sources have 1.28GHz IDs within a search radius of $7.5''$. Using the *Herschel* $250 \mu m$ primary beam size ($17.6''$) as the searching radius, we find 54.2% of *Herschel* sources have multiple 1.28GHz IDs. In addition, the 250 μm flux distributions of *Herschel* sources with and without multiplicity are found to be statistically different. Our average FIR-radio correlation coefficient $q_{250\mu m}$ is 2.33 ± 0.26, consistent with other literature. Adding $q_{250\mu m}$ as a new constraint, the probability of finding chance-aligned sources is reduced by a factor of ~ 6, but with the risk of discarding true identifications of radio-loud/quiet sources. With accurate MeerKAT positions, we cross-match our *Herschel* sources to optical/NIR data followed by photometric redshift estimations via EAZY. About 60.7% of *Herschel*-MeerKAT matched sources are identified in the optical/NIR,

but only $\sim$ 50.0% of *Herschel*-MeerKAT matched sources in the central 4.63′ radius regions have optical/NIR IDs. This implies that there are more higher redshift *Herschel* sources with very faint opitcal/NIR counterparts in the centres of the fields. By removing $z < 1$ sources, the density contrasts of two of the candidate protoclusters increase, suggestive of them being real protoclusters at $z > 1$. There is also potentially a $0.9 < z < 1.2$ overdensity associated with one candidate protocluster. In summary, photometric redshifts from radio-optical cross-identifications have provided some tentative evidence of overdensities aligning with two of the candidate protoclusters.

In Chapter 3, we present deep SOFI NIR images in J, H and K_s filters of 13 candidate DSFG protoclusters selected from *Planck* and *Herschel* data by Greenslade et al. (2018). Raw data is reduced into science images by using a combination of IRAF and Python packages CCDPROC and ASTROALIGN. Sources are extracted from the final science images with Source Extractor, followed by photometry calibration, resulting in a total of 2238 sources with a 5σ detection limit down to $J = 23.07 \pm 0.21$, $H = 23.12 \pm 0.33$ and $K_s = 21.17 \pm 0.13$ mag. We then cross-identify the extracted SOFI sources with some large public surveys in the optical/NIR, building 7-9 band photometry for all sources. With this photometry, we derive the photometric redshifts of these sources via the template fitting code EAZY, reaching an approximate photo-z accuracy of 0.09. By comparing the photometric redshift distributions of our SOFI sources to that of blank field sources, we find that 6 of our 13 targeted candidate protoclusters align with overdensities at $4 \leq z \leq 5$, suggesting that they are real DSFG protoclusters at high redshifts.

In Chapter 4, we investigate the physical properties of nineteen 500 μm risers which are defined as sources with increasing flux densities in the three *Herschel* SPIRE bands. Such red *Herschel* colours suggests that they lie at $z > 4$, making them candidate extreme DSFGs. The selected 19 sources have single detections in high-resolution SMA observations. Based on their SMA positions, we cross-match these sources to public databases in the optical to MIR. For sources without known spectroscopic redshifts, we used the FIR/submm-optimised photometric redshift code MMPZ to estimate their redshifts. We find a mean redshift of $< z >= 3.874 \pm 0.233$ with 10 of the 19 sources at $2.6 < z < 3.7$ and the other 9 sources at $z > 4$. Using MMPZ photometric redshifts and available spectroscopic redshifts, we fit SED templates to our sample with CIGALE then derive their physical properties

based on the fitted SEDs. The mean SFR and stellar mass of our sample are $4107 \pm 800\ M_{\odot} yr^{-1}$ and $< M_{*} >= 2.297 \pm 0.459 \times 10^{12}\ M_{\odot}$, with $L_{IR} > 10^{13}\ L_{\odot}$ for nearly all sources. The SFR-stellar mass relation of our sample suggests that either $z < 3.8$ 500 μm risers are starbursts that are forming stars more intensely than typical star-forming galaxies at similar redshifts, or their redshifts are underestimated. We also find that there is no significant correlation between the AGN fractions and SFR of sample, implying that SFR of 500 μm risers will not be strongly influenced by AGN luminosities, consistent with previous studies of X-ray selected AGN in SMGs. One of our sample, Cosmos_9 has an unsuccessful SED fit with derived physical properties that are uncharacteristic of DSFGs. We suspect that this source is a foreground optical source aligning with it and potentially lenses the FIR/submm emission of the background object, thus the optical to MIR photometry does not match to its FIR/submm photometry and causes the SED fitting failure. Visual examination does not reveal any significant evidence to determine whether there is a FIR/submm source lensed by a optical foreground source, further observations are required to understand the nature of Cosmos_9.

5.2 Future Work

5.2.1 Identification of DSFG protoclusters with high-quality radio data

The transparency of interstellar dust at radio wavelengths and the tight FIR-radio correlation allow easy identifications of FIR/submm bright DSFGs in the radio. With accurate positions of DSFGs determined by their radio counterparts, further cross-identifications of these DSFGs can be more easily carried out in the optical/NIR and thus enable multi-wavelength analysis of these DSFGs such as SED fitting and photometric redshift estimation. Our work in Chapter 2 has already demonstrated the potential of utilising high-quality radio data in identifying DSFG protoclusters, where we present deep MeerKAT 1.28GHz observations of three candidate DSFG protoclusters selected from *Herschel* and *Planck* data. Future work can expand the work in Chapter 2 by collecting deep radio observations of more candidate DSFG protoclusters and DSFGs in general, which will greatly increase the number of promising candidate protoclusters for spectroscopic follow-up. In addition, deeper radio observations will allow us to find more robust radio identifications of FIR/submm sources. While the *Herschel*-

MeerKAT matching rate in Chapter 2 is sufficiently high ($\sim 95\%$), there are still some *Herschel* sources with no MeerKAT counterpart. Moreover, MeerKAT sources are extracted to have larger than 4σ signal-to-noise ratio, so the resulting catalogue is completeness and reliability limited to only contain sources at $S_{1.28GHz} \geq 21\mu Jy$. With deeper radio observations, we will be able to find more reliable counterparts of FIR/submm sources at much fainter flux density levels, allowing us to achieve a even higher FIR-radio identification rate and thus obtain a better picture of DSFGs and DSFG protoclusters in the radio. The next-generation radio telescope SKA will start seeing its first light by the end of this decade. At 1.4 GHz, it is expected to reach a maximum angular resolution of $0.4''$ and a sensitivity of $2\mu Jy/beam$ with a $\sim 60'$ diameter field of view for a one hour observation (Braun et al., 2019). Compared to the angular resolutions of $\sim 6''$ and the sensitivity of $\sim 5\mu Jy/beam$ of our ~ 5 hours MeerKAT observations, future SKA observations of candidate DSFG protoclusters will provide not only more complete identifications of DSFGs, but also more accurate positions of DSFGs for better cross-identifications in the optical/NIR.

5.2.2 DSFGs and their environments in the optical/NIR

In Chapter 2, Chapter 3 and Chapter 4, we heavily rely on the optical/NIR photometry of DSFGs to investigate them and the environments they reside in. However, high-redshift DSFGs are very difficult to detect in the optical/NIR as they are dusty and optical/NIR observations do not benefit from a very negative K-correction as the case of FIR/submm observations. As such, most optical/NIR survey datasets that are currently available are either not deep enough to detect counterparts of most high-redshift DSFGs, or they only cover a small area which limits our ability to build photometry of statistically large samples of DSFGs. For example, in Chapter 2, we use KiDS data to construct photometry of candidate DSFG protocluster members as KiDS is the deepest optical/NIR survey that observes our target fields. But KiDS is still not deep enough as only $\sim 60\%$ of *Herschel*-MeerKAT matched sources are identified in KiDS, and one of the three target fields is not fully covered by KiDS. With future deeper optical/NIR observations coming from, for example, JWST, Euclid an Vera C. Rubin Observatory LSST, more DSFGs will be identified in the optical/NIR which will provide better indication of the nature of the candidate protoclusters. Similarly, in Chapter 4, some of our 500 μm

risers have no optical/NIR IDs due to the lack of deep optical/NIR data. If there is deep optical/NIR photometry available, we will be able to improve the SEDs fitted to such sources and increase the accuracy of their derived physical properties. Therefore, a major focus of future studies of DSFGs is to obtain and utilise deeper optical/NIR photometry. Fortunately, the recent launches of JWST and Euclid have already started to shed light on high-redshift DSFG studies (Chen et al., 2022; Gillman et al., 2023; Zavala et al., 2023; McKinney et al., 2023). In particular, the large area surveys carried out by Euclid and LSST will be ideal for DSFG studies similar to those presented in this thesis.

On the other hand, spectroscopic observation is still the only method to obtain robust redshift of any astronomical object in the foreseeable future. Therefore, following the work in this thesis, future deep spectroscopic observations are essential to acquire the redshifts of sources aligned with candidate protoclusters to confirm their nature, as well as the redshifts of 500 μm risers to better constrain their derived properties. Such observations can be carried out with JWST, NOEMA and ALMA in the future.

5.3 Concluding remarks

For the last few decades, our knowledge of the DSFG/SMG population and its role in galaxy formation and evolution has been dramatically improved. In our current understanding, half of the UV/optical emission produced by star formation is absorbed and re-emitted in the infrared by interstellar dust. Galaxies hosting intense star formation are thus dusty and extremely IR bright, and such galaxies are found to dominate at $z = 2 - 3$ where the cosmic star formation rate density peaks. These DSFGs are thought to be progenitors of present day massive elliptical galaxies which are found in galaxy clusters, therefore they may probe the formation sites of galaxy clusters, structures commonly known as protoclusters. However, there are still many open questions about DSFGs, such as what is their role in galaxy evolution at higher redshifts, what is the mechanism behind the triggering/quenching of the intense star formation in them, and what is the nature of the environments in which they reside. These questions remain difficult to answer due to the large beamsizes of FIR/submm single-dish instruments causing confusion of faint objects which limit our ability to detect fainter sources, while

higher-resolution interferometers typically have fields of view too small to carry out large area surveys. As such, searching of potential targets from large surveys, followed by targeted high-resolution deeper observations will still be an important method for future DSFG studies. While there is no further space-based FIR/submm mission planned for at least the next decade, the recent advance of optical/NIR and radio instrumentation such as JWST, Euclid, LSST and SKA will open a new era of DSFG studies as such instruments will provide much deeper photometry than any current data. DSFG studies in the next decade are thus likely to focus on combining data from these new instruments with data from ground-based FIR/submm telescopes such as the SMA, SCUBA-2, NOEMA and ALMA, as well as legacy data from *Herschel* and *Spitzer*.

Acknowledgements

Herschel is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. The H-ATLAS website is `http://www.h-atlas.org/`. This work made extensive use of the Starlink Table/VOTable Processing Software, TOPCAT (Taylor, 2005). This research has made use of "Aladin sky atlas" developed at CDS, Strasbourg Observatory, France. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The MeerKAT telescope is operated by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation. The Herschel-ATLAS is a project with Herschel, which is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. The H-ATLAS website is http://www.h-atlas.org/. This work is based on data products created from observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programme 198.A-2006. This research is based on observations made with ESO Telescopes at the La Silla Paranal Observatory under programme IDs 177.A-3016, 177.A-3017, 177.A-3018 and 179.A-2004, and on data products produced by the KiDS consortium. The KiDS production team acknowledges support from: Deutsche Forschungsgemeinschaft, ERC, NOVA and NWO-M grants; Target; the University of Padova, and the University Federico II (Naples). For the creation of the data used in this work, the SHARKS team at the Instituto de Astrofísica de Canarias has been financial supported by the Spanish Ministry of Science, Innovation and Universities (MICIU) under grant AYA2017-84061-P, co-financed by FEDER (European Regional Development Funds), by the Spanish Space Research Program "Participation in the NISP instrument and preparation

for the science of EUCLID" (ESP2017-84272-C2-1-R) and by the ACIISI, Consejería de Economía, Conocimiento y Empleo del Gobierno de Canarias and the European Regional Development Fund (ERDF) under grant with reference PROID2020010107. We thank the support of the Wide-Field Astronomy Unit for testing and parallelising the mosaic process and preparing the releases. The work of the Wide-Field Astronomy Unit is funded by the UK Science and Technology Facilities Council through grant ST/T002956/1. This work is based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programme(s) 111.250G. NOIRLab IRAF is distributed by the Community Science and Data Center at NSF NOIRLab, which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the U.S. National Science Foundation. This research made use of Astropy, a community-developed core Python package for Astronomy (Robitaille et al., 2013; Price-Whelan et al., 2018). This research made use of ccdproc, an Astropy package for image reduction (Craig et al. 2023). This work is based on the Legacy Surveys which consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID: 2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing-Arizona Sky Survey (BASS; NOAO Prop. ID: 2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Prop. ID: 2016A-0453; PI: Arjun Dey). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory (LBNL). The Legacy Surveys project is honoured to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

Bibliography

Abbott, T., Abdalla, F., Allam, S., Amara, A., Annis, J., Asorey, J., Avila, S., Ballester, O., Banerji, M., Barkhouse, W., et al. (2018). The dark energy survey: Data release 1. *The Astrophysical Journal Supplement Series*, 239(2):18.

Abbott, T., Adamów, M., Aguena, M., Allam, S., Amon, A., Annis, J., Avila, S., Bacon, D., Banerji, M., Bechtol, K., et al. (2021). The dark energy survey data release 2. *The Astrophysical Journal Supplement Series*, 255(2):20.

Ade, P. A., Aghanim, N., Argüeso, F., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A., et al. (2014). Planck 2013 results. xxviii. the planck catalogue of compact sources. *Astronomy & Astrophysics*, 571:A28.

Ade, P. A., Aumont, J., Baccigalupi, C., Banday, A., Barreiro, R., Bartolo, N., Basak, S., Battaglia, P., Battaner, E., Benabed, K., et al. (2016). Planck 2015 results-iii. lfi systematic uncertainties. *Astronomy & Astrophysics*, 594:A3.

Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R., Bartolo, N., Basak, S., et al. (2020). Planck 2018 results-vi. cosmological parameters. *Astronomy & Astrophysics*, 641:A6.

Aihara, H., AlSayyad, Y., Ando, M., Armstrong, R., Bosch, J., Egami, E., Furusawa, H., Furusawa, J., Goulding, A., Harikane, Y., et al. (2019). Second data release of the hyper suprime-cam subaru strategic program. *Publications of the Astronomical Society of Japan*, 71(6):114.

Aihara, H., AlSayyad, Y., Ando, M., Armstrong, R., Bosch, J., Egami, E., Furusawa, H., Furusawa,

J., Harasawa, S., Harikane, Y., et al. (2022). Third data release of the hyper suprime-cam subaru strategic program. *Publications of the Astronomical Society of Japan*, 74(2):247–272.

Amblard, A., Cooray, A., Serra, P., Temi, P., Barton, E., Negrello, M., Auld, R., Baes, M., Baldry, I., Bamford, S., et al. (2010). Herschel-atlas: Dust temperature and redshift distribution of spire and pacs detected sources using submillimetre colours. *Astronomy & Astrophysics*, 518:L9.

An, F. X., Stach, S., Smail, I., Swinbank, A., Almaini, O., Simpson, C., Hartley, W., Maltby, D., Ivison, R., Arumugam, V., et al. (2018). A machine-learning method for identifying multiwavelength counterparts of submillimeter galaxies: training and testing using as2uds and aless. *The Astrophysical Journal*, 862(2):101.

Andrews, S. K., Driver, S. P., Davies, L., Lagos, C. d. P., and Robotham, A. S. (2018). Modelling the cosmic spectral energy distribution and extragalactic background light over all time. *Monthly Notices of the Royal Astronomical Society*, 474(1):898–916.

Aravena, M., Decarli, R., Walter, F., Da Cunha, E., Bauer, F., Carilli, C., Daddi, E., Elbaz, D., Ivison, R. J., Riechers, D., et al. (2016). The alma spectroscopic survey in the hubble ultra deep field: continuum number counts, resolved 1.2 mm extragalactic background, and properties of the faintest dusty star-forming galaxies. *The Astrophysical Journal*, 833(1):68.

Araya-Araya, P., Cochrane, R. K., Hayward, C. C., Yates, R. M., Sodré Jr, L., Vicentin, M. C., Rennehan, D., Overzier, R., and van Daalen, M. (2024). Modelling the multi-wavelength detection of protoclusters. i: An excess of submillimetre galaxies in protocluster cores. *arXiv preprint arXiv:2408.00062*.

Arnouts, S., Moscardini, L., Vanzella, E., Colombi, S., Cristiani, S., Fontana, A., Giallongo, E., Matarrese, S., and Saracco, P. (2002). Measuring the redshift evolution of clustering: the hubble deep field south. *Monthly Notices of the Royal Astronomical Society*, 329(2):355–366.

Asboth, V., Conley, A., Sayers, J., Béthermin, M., Chapman, S., Clements, D., Cooray, A., Dannerbauer, H., Farrah, D., Glenn, J., et al. (2016). Hermes: a search for high-redshift dusty galaxies in the hermes large mode survey–catalogue, number counts and early results. *Monthly Notices of the Royal Astronomical Society*, 462(2):1989–2000.

Bacon, R., Accardo, M., Adjali, L., Anwand, H., Bauer, S., Biswas, I., Blaizot, J., Boudon, D., Brau-Nogue, S., Brinchmann, J., et al. (2010). The muse second-generation vlt instrument. In *Ground-based and Airborne Instrumentation for Astronomy III*, volume 7735, pages 131–139. SPIE.

Balogh, M. L., Schade, D., Morris, S. L., Yee, H. K., Carlberg, R. G., and Ellingson, E. (1998). The dependence of cluster galaxy star formation rates on the global environment. *The Astrophysical Journal*, 504(2):L75.

Bamford, S. P., Nichol, R. C., Baldry, I. K., Land, K., Lintott, C. J., Schawinski, K., Slosar, A., Szalay, A. S., Thomas, D., Torki, M., et al. (2009). Galaxy zoo: the dependence of morphology and colour on environment. *Monthly Notices of the Royal Astronomical Society*, 393(4):1324–1352.

Barger, A., Wang, W.-H., Cowie, L., Owen, F., Chen, C.-C., and Williams, J. (2012). Precise identifications of submillimeter galaxies: measuring the history of massive star-forming galaxies to z> 5. *The Astrophysical Journal*, 761(2):89.

Barrufet, L., Pearson, C., Serjeant, S., Małek, K., Baronchelli, I., Campos-Varillas, M., White, G., Valtchanov, I., Matsuhara, H., Conversi, L., et al. (2020). A high redshift population of galaxies at the north ecliptic pole-unveiling the main sequence of dusty galaxies. *Astronomy & Astrophysics*, 641:A129.

Baugh, C. M., Lacey, C., Frenk, C., Granato, G., Silva, L., Bressan, A., Benson, A., and Cole, S. (2005). Can the faint submillimetre galaxies be explained in the λ cold dark matter model? *Monthly Notices of the Royal Astronomical Society*, 356(3):1191–1200.

Benitez, N. (2000). Bayesian photometric redshift estimation. *The Astrophysical Journal*, 536(2):571.

Beroiz, M., Cabral, J., and Sanchez, B. (2020). Astroalign: A python module for astronomical image registration. *Astronomy and Computing*, 32:100384.

Berta, S., Magnelli, B., Nordon, R., Lutz, D., Wuyts, S., Altieri, B., Andreani, P., Aussel, H., Castaneda, H., Cepa, J., et al. (2011). Building the cosmic infrared background brick by brick with herschel/pep. *Astronomy & Astrophysics*, 532:A49.

Bertin, E. and Arnouts, S. (1996). Sextractor: Software for source extraction. *Astronomy and astrophysics supplement series*, 117(2):393–404.

Béthermin, M., Dole, H., Cousin, M., and Bavouzet, N. (2010). Submillimeter number counts at 250 μm, 350 μm and 500 μm in blast data. *Astronomy & Astrophysics*, 516:A43.

Béthermin, M., Le Floc'h, E., Ilbert, O., Conley, A., Lagache, G., Amblard, A., Arumugam, V., Aussel, H., Berta, S., Bock, J., et al. (2012). Hermes: deep number counts at 250 μm, 350 μm and 500 μm in the cosmos and goods-n fields and the build-up of the cosmic infrared background. *Astronomy & Astrophysics*, 542:A58.

Béthermin, M., Wu, H.-Y., Lagache, G., Davidzon, I., Ponthieu, N., Cousin, M., Wang, L., Doré, O., Daddi, E., and Lapi, A. (2017). The impact of clustering and angular resolution on far-infrared and millimeter continuum observations. *Astronomy & Astrophysics*, 607:A89.

Bezanson, R., Wake, D. A., Brammer, G. B., Van Dokkum, P. G., Franx, M., Labbé, I., Leja, J., Momcheva, I. G., Nelson, E. J., Quadri, R. F., et al. (2016). Leveraging 3d-hst grism redshifts to quantify photometric redshift performance. *The Astrophysical Journal*, 822(1):30.

Bianchi, S. (2013). Vindicating single-t modified blackbody fits to herschel seds. *Astronomy & Astrophysics*, 552:A89.

Biggs, A., Ivison, R., Ibar, E., Wardlow, J., Dannerbauer, H., Smail, I., Walter, F., Weiß, A., Chapman, S., Coppin, K., et al. (2011). The laboca survey of the extended chandra deep field-south–radio and mid-infrared counterparts to submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 413(4):2314–2338.

Blain, A. (1996). Galaxy—galaxy gravitational lensing in the millimetre/submillimetre waveband. *Monthly Notices of the Royal Astronomical Society*, 283(4):1340–1348.

Blain, A. W., Smail, I., Ivison, R., Kneib, J.-P., and Frayer, D. T. (2002). Submillimeter galaxies. *Physics Reports*, 369(2):111–176.

Boquien, M., Burgarella, D., Roehlly, Y., Buat, V., Ciesla, L., Corre, D., Inoue, A., and Salas, H. (2019). Cigale: a python code investigating galaxy emission. *Astronomy & Astrophysics*, 622:A103.

Bothwell, M., Smail, I., Chapman, S., Genzel, R., Ivison, R., Tacconi, L., Alaghband-Zadeh, S., Bertoldi, F., Blain, A., Casey, C., et al. (2013). A survey of molecular gas in luminous sub-millimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 429(4):3047–3067.

Brammer, G. B., van Dokkum, P. G., and Coppi, P. (2008). Eazy: a fast, public photometric redshift code. *The Astrophysical Journal*, 686(2):1503.

Braun, R., Bonaldi, A., Bourke, T., Keane, E., and Wagg, J. (2019). Anticipated performance of the square kilometre array–phase 1 (ska1). *arXiv preprint arXiv:1912.12699*.

Brinchmann, J., Inami, H., Bacon, R., Contini, T., Maseda, M., Chevallard, J., Bouché, N., Boogaard, L., Carollo, M., Charlot, S., et al. (2017). The muse hubble ultra deep field survey-iii. testing photometric redshifts to 30th magnitude. *Astronomy & Astrophysics*, 608:A3.

Bruzual, G. and Charlot, S. (2003). Stellar population synthesis at the resolution of 2003. *Monthly Notices of the Royal Astronomical Society*, 344(4):1000–1028.

Bullock, J. S. and Boylan-Kolchin, M. (2017). Small-scale challenges to the λ cdm paradigm. *Annual Review of Astronomy and Astrophysics*, 55(1):343–387.

Burgarella, D., Heinis, S., Magdis, G., Auld, R., Blain, A., Bock, J., Brisbin, D., Buat, V., Chanial, P., Clements, D., et al. (2011). Hermes: Lyman break galaxies individually detected at 0.7< z< 2.0 in goods-n with herschel/spire. *The Astrophysical Journal Letters*, 734(1):L12.

Buzzoni, B., Delabre, B., Dekker, H., Dodorico, S., Enard, D., Focardi, P., Gustafsson, B., Nees, W., Paureau, J., and Reiss, R. (1984). The eso faint object spectrograph and camera (efosc). *ESO Messenger (ISSN 0722-6691), Dec. 1984, p. 9-13.*, 38:9–13.

Cairns, J., Clements, D., Greenslade, J., Petitpas, G., Cheng, T., Ding, Y., Parmar, A., Pérez-Fournon, I., and Riechers, D. (2023). The nature of 500 micron risers–ii. multiplicities and environments of sub-mm faint dusty star-forming galaxies. *Monthly Notices of the Royal Astronomical Society*, 519(1):709–728.

Capaccioli, M. and Schipani, P. (2011). The vlt survey telescope opens to the sky: history of a commissioning. *The Messenger*, 146(2-7):27.

Capak, P. L., Riechers, D., Scoville, N. Z., Carilli, C., Cox, P., Neri, R., Robertson, B., Salvato, M., Schinnerer, E., Yan, L., et al. (2011). A massive protocluster of galaxies at a redshift of z 5.3. *Nature*, 470(7333):233–235.

Caputi, K., Lagache, G., Yan, L., Dole, H., Bavouzet, N., Le Floc'h, E., Choi, P., Helou, G., and Reddy, N. (2007). The infrared luminosity function of galaxies at redshifts z= 1 and z˜ 2 in the goods fields. *The Astrophysical Journal*, 660(1):97.

Carilli, C. and Rawlings, S. (2004). Science with the square kilometer array: Motivation, key science projects, standards and assumptions. *arXiv preprint astro-ph/0409274*.

Casey, C., Berta, S., Béthermin, M., Bock, J., Bridge, C., Budynkiewicz, J., Burgarella, D., Chapin, E., Chapman, S., Clements, D., et al. (2012a). A redshift survey of herschel far-infrared selected starbursts and implications for obscured star formation. *The Astrophysical Journal*, 761(2):140.

Casey, C., Berta, S., Béthermin, M., Bock, J., Bridge, C., Burgarella, D., Chapin, E., Chapman, S., Clements, D., Conley, A., et al. (2012b). A population of z> 2 far-infrared herschel-spire-selected starbursts. *The Astrophysical Journal*, 761(2):139.

Casey, C., Cooray, A., Capak, P., Fu, H., Kovac, K., Lilly, S., Sanders, D., Scoville, N. Z., and Treister, E. (2015). A massive, distant proto-cluster at z= 2.47 caught in a phase of rapid formation? *The Astrophysical Journal Letters*, 808(2):L33.

Casey, C. M. (2012). Far-infrared spectral energy distribution fitting for galaxies near and far. *Monthly Notices of the Royal Astronomical Society*, 425(4):3094–3103.

Casey, C. M. (2016). The ubiquity of coeval starbursts in massive galaxy cluster progenitors. *The Astrophysical Journal*, 824(1):36.

Casey, C. M. (2020). Far-infrared photometric redshifts: A new approach to a highly uncertain enterprise. *The Astrophysical Journal*, 900(1):68.

Casey, C. M., Narayanan, D., and Cooray, A. (2014). Dusty star-forming galaxies at high redshift. *Physics Reports*, 541(2):45–161.

Casey, C. M., Zavala, J. A., Spilker, J., da Cunha, E., Hodge, J., Hung, C.-L., Staguhn, J., Finkelstein, S. L., and Drew, P. (2018). The brightest galaxies in the dark ages: Galaxies' dust continuum emission during the reionization era. *The Astrophysical Journal*, 862(1):77.

Chabrier, G. (2003). The galactic disk mass function: reconciliation of the hubble space telescope and nearby determinations. *The Astrophysical Journal*, 586(2):L133.

Chapin, E. L., Pope, A., Scott, D., Aretxaga, I., Austermann, J. E., Chary, R.-R., Coppin, K., Halpern, M., Hughes, D. H., Lowenthal, J. D., et al. (2009). An aztec 1.1 mm survey of the goods-n field–ii. multiwavelength identifications and redshift distribution. *Monthly Notices of the Royal Astronomical Society*, 398(4):1793–1808.

Chapman, S., Blain, A., Ibata, R., Ivison, R., Smail, I., and Morrison, G. (2009). Do submillimeter galaxies really trace the most massive dark-matter halos? discovery of a high-z cluster in a highly active phase of evolution. *The Astrophysical Journal*, 691(1):560.

Chapman, S. C., Blain, A., Ivison, R., and Smail, I. R. (2003). A median redshift of 2.4 for galaxies bright at submillimetre wavelengths. *Nature*, 422(6933):695–698.

Chapman, S. C., Blain, A., Smail, I., and Ivison, R. (2005). A redshift survey of the submillimeter galaxy population. *The Astrophysical Journal*, 622(2):772.

Chapman, S. C., Hill, R., Aravena, M., Archipley, M., Babul, A., Burgoyne, J., Canning, R. E., Deane, R. P., De Breuck, C., Gonzalez, A. H., et al. (2024). Brightest cluster galaxy formation in the z= 4.3 protocluster spt 2349-56: discovery of a radio-loud active galactic nucleus. *The Astrophysical Journal*, 961(1):120.

Charlot, S. and Fall, S. M. (2000). A simple model for the absorption of starlight by dust in galaxies. *The Astrophysical Journal*, 539(2):718.

Chen, C.-C., Cowie, L. L., Barger, A. J., Casey, C. M., Lee, N., Sanders, D. B., Wang, W.-H., and Williams, J. P. (2013). Resolving the cosmic far-infrared background at 450 and 850 μm with scuba-2. *The Astrophysical Journal*, 776(2):131.

Chen, C.-C., Cowie, L. L., Wang, W.-H., Barger, A. J., and Williams, J. P. (2011). Submillimeter sources behind the massive lensing clusters a370 and a2390. *The Astrophysical Journal*, 733(1):64.

Chen, C.-C., Gao, Z.-K., Hsu, Q.-N., Liao, C.-L., Ling, Y.-H., Lo, C.-M., Smail, I., Wang, W.-H., and Wang, Y.-J. (2022). Jwst sneaks a peek at the stellar morphology of z 2 submillimeter galaxies: Bulge formation at cosmic noon. *The Astrophysical Journal Letters*, 939(1):L7.

Cheng, T., Clements, D., Greenslade, J., Cairns, J., Andreani, P., Bremer, M., Conversi, L., Cooray, A., Dannerbauer, H., De Zotti, G., et al. (2019). Scuba-2 observations of candidate starbursting protoclusters selected by planck and herschel-spire. *Monthly Notices of the Royal Astronomical Society*, 490(3):3840–3859.

Chevallard, J. and Charlot, S. (2016). Modelling and interpreting spectral energy distributions of galaxies with beagle. *Monthly Notices of the Royal Astronomical Society*, 462(2):1415–1443.

Chiang, Y.-K., Overzier, R. A., Gebhardt, K., Finkelstein, S. L., Chiang, C.-T., Hill, G. J., Blanc, G. A., Drory, N., Chonis, T. S., Zeimann, G. R., et al. (2015). Surveying galaxy proto-clusters in emission: A large-scale structure at z= 2.44 and the outlook for hetdex. *The Astrophysical Journal*, 808(1):37.

Clements, D., Braglia, F., Hyde, A., Pérez-Fournon, I., Bock, J., Cava, A., Chapman, S., Conley, A., Cooray, A., Farrah, D., et al. (2014). Herschel multitiered extragalactic survey: clusters of dusty galaxies uncovered by herschel and planck. *Monthly Notices of the Royal Astronomical Society*, 439(2):1193–1211.

Clements, D., Dunne, L., and Eales, S. (2010a). The submillimetre properties of ultraluminous infrared galaxies. *Monthly Notices of the Royal Astronomical Society*, 403(1):274–286.

Clements, D., Sutherland, W., McMahon, R., and Saunders, W. (1996). Optical imaging of ultraluminous iras galaxies: how many are mergers? *Monthly Notices of the Royal Astronomical Society*, 279(2):477–497.

Clements, D. L., Braglia, F., Petitpas, G., Greenslade, J., Cooray, A., Valiante, E., De Zotti, G., O'Halloran, B., Holdship, J., Morris, B., et al. (2016). H-atlas: a candidate high redshift clus-

ter/protocluster of star-forming galaxies. *Monthly Notices of the Royal Astronomical Society*, 461(2):1719–1733.

Clements, D. L., Rigby, E., Maddox, S., Dunne, L., Mortier, A., Pearson, C., Amblard, A., Auld, R., Baes, M., Bonfield, D., et al. (2010b). Herschel-atlas: Extragalactic number counts from 250 to 500 microns. *Astronomy & Astrophysics*, 518:L8.

Clements, D. L., Rowan-Robinson, M., Lawrence, A., Broadhurst, T., and McMahon, R. (1992). Submillimetre observations of the z=2.286 IRAS galaxy 10214+4724. *Monthly Notices of the Royal Astronomical Society*, 256:35P–37.

Coe, D., Salmon, B., Bradač, M., Bradley, L. D., Sharon, K., Zitrin, A., Acebron, A., Cerny, C., Cibirka, N., Strait, V., et al. (2019). Relics: reionization lensing cluster survey. *The Astrophysical Journal*, 884(1):85.

Colless, M., Dalton, G., Maddox, S., Sutherland, W., Norberg, P., Cole, S., Bland-Hawthorn, J., Bridges, T., Cannon, R., Collins, C., et al. (2001). The 2df galaxy redshift survey: spectra and redshifts. *Monthly Notices of the Royal Astronomical Society*, 328(4):1039–1063.

Condon, J. (1992). Radio emission from normal galaxies. *In: Annual review of astronomy and astrophysics. Vol. 30 (A93-25826 09-90), p. 575-611.*, 30:575–611.

Condon, J. J. (1974). Confusion and Flux-Density Error Distributions. *The Astrophysical Journal*, 188:279–286.

Cooke, E., Hatch, N. A., Muldrew, S., Rigby, E., and Kurk, J. (2014). A z= 2.5 protocluster associated with the radio galaxy mrc 2104- 242: star formation and differing mass functions in dense environments. *Monthly Notices of the Royal Astronomical Society*, 440(4):3262–3274.

Cooray, A. (2016). Extragalactic background light measurements and applications. *Royal Society Open Science*, 3(3):150555.

Cooray, A., Calanog, J., Wardlow, J. L., Bock, J., Bridge, C., Burgarella, D., Bussmann, R., Casey, C., Clements, D., Conley, A., et al. (2014). Hermes: the rest-frame uv emission and a lensing model for the z= 6.34 luminous dusty starburst galaxy hfls3. *The Astrophysical Journal*, 790(1):40.

Cowley, W. I., Lacey, C. G., Baugh, C. M., and Cole, S. (2015). Simulated observations of sub-millimetre galaxies: the impact of single-dish resolution and field variance. *Monthly Notices of the Royal Astronomical Society*, 446(2):1784–1798.

Craig, M., Crawford, S., Seifert, M., Robitaille, T., Sipőcz, B., Walawender, J., Crawford, S., Vinícius, Z., Ninan, J. P., Droettboom, M., Bowers, T. E., Youn, J., Gondhalekar, Y., Tollerud, E., Lim, P. L., Bray, E. M., Bach, Y. P., stottsco, Janga, V. R., walkerna22, Günther, H. M., Rol, E., A., J., Bradley, L., Price-Whelan, A., Deil, C., Ryon, J., Lee, K., Barbary, K., and Weiner, B. (2023). astropy/ccdproc: 2.4.1.

Crain, R. A., Schaye, J., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J. C., et al. (2015). The eagle simulations of galaxy formation: calibration of subgrid physics and model variations. *Monthly Notices of the Royal Astronomical Society*, 450(2):1937–1961.

Cutri, R., Skrutskie, M., Van Dyk, S., Beichman, C., Carpenter, J., Chester, T., Cambresy, L., Evans, T., Fowler, J., Gizis, J., et al. (2003). 2mass all sky catalog of point sources. *The IRSA 2MASS All-Sky Point Source Catalog*.

da Cunha, E., Walter, F., Smail, I., Swinbank, A., Simpson, J., Decarli, R., Hodge, J., Weiss, A., Van der Werf, P., Bertoldi, F., et al. (2015). An alma survey of sub-millimeter galaxies in the extended chandra deep field south: physical properties derived from ultraviolet-to-radio modeling. *The Astrophysical Journal*, 806(1):110.

Daddi, E., Dannerbauer, H., Krips, M., Walter, F., Dickinson, M., Elbaz, D., and Morrison, G. (2009a). A co emission line from the optical and near-ir undetected submillimeter galaxy gn10. *The Astrophysical Journal Letters*, 695(2):L176.

Daddi, E., Dannerbauer, H., Stern, D., Dickinson, M., Morrison, G., Elbaz, D., Giavalisco, M., Mancini, C., Pope, A., and Spinrad, H. (2009b). Two bright submillimeter galaxies in a z= 4.05 protocluster in goods-north, and accurate radio-infrared photometric redshifts. *The astrophysical journal*, 694(2):1517.

Dahlen, T., Mobasher, B., Faber, S. M., Ferguson, H. C., Barro, G., Finkelstein, S. L., Finlator, K., Fontana, A., Gruetzbauch, R., Johnson, S., et al. (2013). A critical assessment of photometric redshift methods: a candels investigation. *The Astrophysical Journal*, 775(2):93.

Dalton, G., Caldwell, M., Ward, A., Whalley, M., Woodhouse, G., Edeson, R., Clark, P., Beard, S., Gallie, A., Todd, S., et al. (2006). The vista infrared camera. In *Ground-based and Airborne Instrumentation for Astronomy*, volume 6269, page 62690X. International Society for Optics and Photonics.

Dannerbauer, H., Carnero, A., Cross, N., and Gutierrez, C. (2022). Exploring the universe via the wide, deep near-infrared imaging eso public survey sharks. *The Messenger*, 187:12–15.

Dannerbauer, H., Kurk, J., De Breuck, C., Wylezalek, D., Santos, J., Koyama, Y., Seymour, N., Tanaka, M., Hatch, N., Altieri, B., et al. (2014). An excess of dusty starbursts related to the spiderweb galaxy. *Astronomy & Astrophysics*, 570:A55.

Dark Energy Survey Collaboration: Fermilab, University of Illinois at Urbana-Champaign, U. o. C. L. B. N. L. C.-T. I.-A. O. and Flaugher, B. (2005). The dark energy survey. *International Journal of Modern Physics A*, 20(14):3121–3123.

Davé, R., Anglés-Alcázar, D., Narayanan, D., Li, Q., Rafieferantsoa, M. H., and Appleby, S. (2019). Simba: Cosmological simulations with black hole growth and feedback. *Monthly Notices of the Royal Astronomical Society*, 486(2):2827–2849.

De Graauw, T., Helmich, F., Phillips, T., Stutzki, J., Caux, E., Whyborn, N., Dieleman, P., Roelfsema, P., Aarts, H., Assendorp, R., et al. (2010). The herschel-heterodyne instrument for the far-infrared (hifi). *Astronomy & Astrophysics*, 518:L6.

de Jong, J. T., Verdoes Kleijn, G. A., Kuijken, K. H., Valentijn, E. A., KiDS, and Consortiums, A.-W. (2013). The kilo-degree survey. *Experimental Astronomy*, 35:25–44.

De Jong, T., Klein, U., Wielebinski, R., and Wunderlichº, E. (1985). from spiral galaxies: a close correlation. *Astron. Astrophys*, 147:L6–L9.

Delvecchio, I., Daddi, E., Sargent, M., Jarvis, M., Elbaz, D., Jin, S., Liu, D., Whittam, I., Algera, H., Carraro, R., et al. (2021). The infrared-radio correlation of star-forming galaxies is strongly mstar-dependent but nearly redshift-invariant since z 4. *Astronomy & astrophysics*, 647:A123.

Deruelle, N. and Uzan, J.-P. (2018). *Relativity in Modern Physics*. Oxford University Press.

Devlin, M. J., Ade, P. A., Aretxaga, I., Bock, J. J., Chapin, E. L., Griffin, M., Gundersen, J. O., Halpern, M., Hargrave, P. C., Hughes, D. H., et al. (2009). Over half of the far-infrared background light comes from galaxies at z> 1.2. *Nature*, 458(7239):737–739.

Dewdney, P. E., Hall, P. J., Schilizzi, R. T., and Lazio, T. J. L. W. (2009). The Square Kilometre Array. *IEEE Proceedings*, 97(8):1482–1496.

Dey, A., Schlegel, D. J., Lang, D., Blum, R., Burleigh, K., Fan, X., Findlay, J. R., Finkbeiner, D., Herrera, D., Juneau, S., et al. (2019). Overview of the desi legacy imaging surveys. *The Astronomical Journal*, 157(5):168.

Dole, H., Rieke, G. H., Lagache, G., Puget, J. L., Alonso-Herrero, A., Bai, L., Blaylock, M., Egami, E., Engelbracht, C. W., Gordon, K. D., Hines, D. C., Kelly, D. M., Le Floc'h, E., Misselt, K. A., Morrison, J. E., Muzerolle, J., Papovich, C., Pérez-González, P. G., Rieke, M. J., Rigby, J. R., Neugebauer, G., Stansberry, J. A., Su, K. Y. L., Young, E. T., Beichman, C. A., and Richards, P. L. (2004). Confusion of Extragalactic Sources in the Mid- and Far-Infrared: Spitzer and Beyond. *Astrophysical Journal Supplement*, 154(1):93–96.

Donevski, D., Buat, V., Boone, F., Pappalardo, C., Bethermin, M., Schreiber, C., Mazyed, F., Alvarez-Marquez, J., and Duivenvoorden, S. (2018). Towards a census of high-redshift dusty galaxies with herschel-a selection of "500 μm-risers". *Astronomy & Astrophysics*, 614:A33.

Dowell, C. D., Conley, A., Glenn, J., Arumugam, V., Asboth, V., Aussel, H., Bertoldi, F., Béthermin, M., Bock, J., Boselli, A., et al. (2014). Hermes: Candidate high-redshift galaxies discovered with herschel/spire. *The Astrophysical Journal*, 780(1):75.

Downes, A., Peacock, J., Savage, A., and Carrie, D. (1986). The parkes selected regions: powerful radio galaxies and quasars at high redshifts. *Monthly Notices of the Royal Astronomical Society*, 218(1):31–62.

Draine, B. (2003a). Interstellar dust grains. ara&a 41, 241–289 (2003). *arXiv preprint astro-ph/0304489*, 8.

Draine, B., Aniano, G., Krause, O., Groves, B., Sandstrom, K., Braun, R., Leroy, A., Klaas, U., Linz, H., Rix, H.-W., et al. (2013). Andromeda's dust. *The Astrophysical Journal*, 780(2):172.

Draine, B. T. (2003b). Interstellar dust grains. *Annual Review of Astronomy and Astrophysics*, 41(1):241–289.

Draine, B. T. and Li, A. (2007). Infrared emission from interstellar dust. iv. the silicate-graphite-pah model in the post-spitzer era. *The Astrophysical Journal*, 657(2):810.

Dressler, A. (1980). Galaxy morphology in rich clusters-implications for the formation and evolution of galaxies. *Astrophysical Journal, Part 1, vol. 236, Mar. 1, 1980, p. 351-365.*, 236:351–365.

Dudzevičiūtė, U., Smail, I., Swinbank, A., Stach, S., Almaini, O., da Cunha, E., An, F. X., Arumugam, V., Birkin, J., Blain, A., et al. (2020). An alma survey of the scuba-2 cls uds field: physical properties of 707 sub-millimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 494(3):3828–3860.

Eales, S., Dunne, L., Clements, D., Cooray, A., De Zotti, G., Dye, S., Ivison, R., Jarvis, M., Lagache, G., Maddox, S., et al. (2010). The herschel atlas. *Publications of the Astronomical Society of the Pacific*, 122(891):499.

Edge, A., Sutherland, W., Kuijken, K., Driver, S., McMahon, R., Eales, S., and Emerson, J. P. (2013). The vista kilo-degree infrared galaxy (viking) survey: bridging the gap between low and high redshift. *The Messenger*, 154:32–34.

Elbaz, D., Daddi, E., Le Borgne, D., Dickinson, M., Alexander, D., Chary, R.-R., Starck, J.-L., Brandt, W., Kitzbichler, M., MacDonald, E., et al. (2007). The reversal of the star formation-density relation in the distant universe. *Astronomy & Astrophysics*, 468(1):33–48.

Engel, H., Tacconi, L., Davies, R., Neri, R., Smail, I., Chapman, S., Genzel, R., Cox, P., Greve, T., Ivison, R., et al. (2010). Most submillimeter galaxies are major mergers. *The Astrophysical Journal*, 724(1):233.

Erb, D. K., Pettini, M., Shapley, A. E., Steidel, C. C., Law, D. R., and Reddy, N. A. (2010). Physical conditions in a young, unreddened, low-metallicity galaxy at high redshift. *The Astrophysical Journal*, 719(2):1168.

Everett, W., Zhang, L., Crawford, T., Vieira, J., Aravena, M., Archipley, M., Austermann, J., Benson, B., Bleem, L., Carlstrom, J., et al. (2020). Millimeter-wave point sources from the 2500 square degree spt-sz survey: catalog and population statistics. *The Astrophysical Journal*, 900(1):55.

Falgarone, E., Zwaan, M., Godard, B., Bergin, E., Ivison, R., Andreani, P., Bournaud, F., Bussmann, R., Elbaz, D., Omont, A., et al. (2017). Large turbulent reservoirs of cold molecular gas around high-redshift starburst galaxies. *Nature*, 548(7668):430–433.

Farrah, D., Lonsdale, C., Borys, C., Fang, F., Waddington, I., Oliver, S., Rowan-Robinson, M., Babbedge, T., Shupe, D., Polletta, M., et al. (2006). The spatial clustering of ultraluminous infrared galaxies over 1.5< z< 3. *The Astrophysical Journal*, 641(1):L17.

Fazio, G., Hora, J., Allen, L., Ashby, M., Barmby, P., Deutsch, L., Huang, J.-S., Kleiner, S., Marengo, M., Megeath, S., et al. (2004). The infrared array camera (irac) for the spitzer space telescope. *The Astrophysical Journal Supplement Series*, 154(1):10.

Fisher, K. B., Huchra, J. P., Strauss, M. A., Davis, M., Yahil, A., and Schlegel, D. (1995). The IRAS 1.2 Jy Survey: Redshift Data. *Astrophysical Journal Supplement*, 100:69.

Fixsen, D., Dwek, E., Mather, J. C., Bennett, C., and Shafer, R. (1998). The spectrum of the extragalactic far-infrared background from the cobe* firas observations. *The Astrophysical Journal*, 508(1):123.

Flaugher, B., Diehl, H., Honscheid, K., Abbott, T., Alvarez, O., Angstadt, R., Annis, J., Antonik, M., Ballester, O., Beaufore, L., et al. (2015). The dark energy camera. *The Astronomical Journal*, 150(5):150.

Flores-Cacho, I., Pierini, D., Soucail, G., Montier, L., Dole, H., Pointecouteau, E., Pelló, R., Le Floc'h, E., Nesvadba, N., Lagache, G., et al. (2016). Multi-wavelength characterisation of z˜ 2 clustered, dusty star-forming galaxies discovered by planck. *Astronomy & Astrophysics*, 585:A54.

Fontanot, F., Monaco, P., Silva, L., and Grazian, A. (2007). Reproducing the assembly of massive galaxies within the hierarchical cosmogony. *Monthly Notices of the Royal Astronomical Society*, 382(2):903–914.

Frenk, C. S. and White, S. D. (2012). Dark matter and cosmic structure. *Annalen der Physik*, 524(9-10):507–534.

Fritz, J., Franceschini, A., and Hatziminaoglou, E. (2006). Revisiting the infrared spectra of active galactic nuclei with a new torus emission model. *Monthly Notices of the Royal Astronomical Society*, 366(3):767–786.

Fudamoto, Y., Ivison, R., Oteo, I., Krips, M., Zhang, Z.-Y., Weiss, A., Dannerbauer, H., Omont, A., Chapman, S., Christensen, L., et al. (2017). The most distant, luminous, dusty star-forming galaxies: redshifts from noema and alma spectral scans. *Monthly Notices of the Royal Astronomical Society*, 472(2):2028–2041.

García-Vergara, C., Hodge, J., Hennawi, J. F., Weiss, A., Wardlow, J., Myers, A. D., and Hickox, R. (2020). The clustering of submillimeter galaxies detected with alma. *The Astrophysical Journal*, 904(1):2.

Gardner, J. P., Mather, J. C., Clampin, M., Doyon, R., Greenhouse, M. A., Hammel, H. B., Hutchings, J. B., Jakobsen, P., Lilly, S. J., Long, K. S., et al. (2006). The james webb space telescope. *Space Science Reviews*, 123:485–606.

Geach, J., Dunlop, J., Halpern, M., Smail, I., Van der Werf, P., Alexander, D., Almaini, O., Aretxaga, I., Arumugam, V., Asboth, V., et al. (2017). The scuba-2 cosmology legacy survey: 850 μm maps, catalogues and number counts. *Monthly Notices of the Royal Astronomical Society*, 465(2):1789–1806.

Geach, J. E., Chapin, E., Coppin, K., Dunlop, J., Halpern, M., Smail, I., Werf, P. v. d., Serjeant, S., Farrah, D., Roseboom, I., et al. (2013). The scuba-2 cosmology legacy survey: blank-field number counts of 450-μm-selected galaxies and their contribution to the cosmic infrared background. *Monthly Notices of the Royal Astronomical Society*, 432(1):53–61.

Gebbie, H., Harding, W., Hilsum, C., Pryce, A., and Roberts, V. (1951). Atmospheric transmission in the 1 to 14 μ region. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 206(1084):87–107.

Gillman, S., Gullberg, B., Brammer, G., Vijayan, A. P., Lee, M., Blánquez, D., Brinch, M., Greve, T. R., Jermann, I., Jin, S., et al. (2023). Sub-millimetre galaxies with webb-near-infrared counterparts and multi-wavelength morphology. *Astronomy & Astrophysics*, 676:A26.

Gispert, R., Lagache, G., and Puget, J. (2000). Implications of the cosmic infrared background for light production and the star formation history in the universe. *arXiv preprint astro-ph/0005554*.

Giulietti, M., Massardi, M., Lapi, A., Bonato, M., Enia, A., Negrello, M., D'Amato, Q., Behiri, M., and De Zotti, G. (2022). The far-infrared/radio correlation for a sample of strongly lensed dusty star-forming galaxies detected by herschel. *Monthly Notices of the Royal Astronomical Society*, 511(1):1408–1419.

Glenn, J., Conley, A., Béthermin, M., Altieri, B., Amblard, A., Arumugam, V., Aussel, H., Babbedge, T., Blain, A., Bock, J., et al. (2010). Hermes: deep galaxy number counts from a p (d) fluctuation analysis of spire science demonstration phase observations. *Monthly Notices of the Royal Astronomical Society*, 409(1):109–121.

Gómez-Guijarro, C., Elbaz, D., Xiao, M., Béthermin, M., Franco, M., Magnelli, B., Daddi, E., Dickinson, M., Demarco, R., Inami, H., et al. (2022). Goods-alma 2.0: Source catalog, number counts, and prevailing compact sizes in 1.1 mm galaxies. *Astronomy & Astrophysics*, 658:A43.

González-Fernández, C., Hodgkin, S. T., Irwin, M. J., González-Solares, E., Koposov, S. E., Lewis, J. R., Emerson, J. P., Hewett, P. C., Yoldaş, A. K., and Riello, M. (2018). The vista zyjhks photometric system: calibration from 2mass. *Monthly Notices of the Royal Astronomical Society*, 474(4):5459–5478.

Goto, T., Yamauchi, C., Fujita, Y., Okamura, S., Sekiguchi, M., Smail, I., Bernardi, M., and Gomez, P. L. (2003). The morphology—density relation in the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 346(2):601–614.

Granato, G. L., De Zotti, G., Silva, L., Bressan, A., and Danese, L. (2004). A physical model for the coevolution of qsos and their spheroidal hosts. *The Astrophysical Journal*, 600(2):580.

Granato, G. L., Lacey, C. G., Silva, L., Bressan, A., Baugh, C., Cole, S., and Frenk, C. (2000). The infrared side of galaxy formation. i. the local universe in the semianalytical framework. *The Astrophysical Journal*, 542(2):710.

Greenslade, J., Clements, D., Cheng, T., De Zotti, G., Scott, D., Valiante, E., Eales, S., Bremer, M., Dannerbauer, H., Birkinshaw, M., et al. (2018). Candidate high-z protoclusters among the planck compact sources, as revealed by herschel–spire. *Monthly Notices of the Royal Astronomical Society*, 476(3):3336–3359.

Greenslade, J., Clements, D., Petitpas, G., Asboth, V., Conley, A., Pérez-Fournon, I., and Riechers, D. (2020). The nature of 500 micron risers i: Sma observations. *Monthly Notices of the Royal Astronomical Society*, 496(2):2315–2333.

Greve, T. R., Bertoldi, F., Smail, I., Neri, R., Chapman, S., Blain, A., Ivison, R., Genzel, R., Omont, A., Cox, P., et al. (2005). An interferometric co survey of luminous submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 359(3):1165–1183.

Griffin, M. J., Abergel, A., Abreu, A., Ade, P. A., André, P., Augueres, J.-L., Babbedge, T., Bae, Y., Baillie, T., Baluteau, J.-P., et al. (2010). The herschel-spire instrument and its in-flight performance. *Astronomy & Astrophysics*, 518:L3.

Gruppioni, C., Béthermin, M., Loiacono, F., Le Fèvre, O., Capak, P., Cassata, P., Faisst, A., Schaerer, D., Silverman, J., Yan, L., et al. (2020). The alpine-alma [cii] survey-the nature, luminosity function, and star formation history of dusty galaxies up to z 6. *Astronomy & Astrophysics*, 643:A8.

Gruppioni, C., Pozzi, F., Rodighiero, G., Delvecchio, I., Berta, S., Pozzetti, L., Zamorani, G., Andreani, P., Cimatti, A., Ilbert, O., et al. (2013). The herschel pep/hermes luminosity function–i. probing the evolution of pacs selected galaxies to z= 4. *Monthly Notices of the Royal Astronomical Society*, 432(1):23–52.

Gullberg, B., Swinbank, A. M., Smail, I., Biggs, A., Bertoldi, F., De Breuck, C., Chapman, S., Chen,

C.-C., Cooke, E., Coppin, K., et al. (2018). The dust and [c ii] morphologies of redshift 4.5 submillimeter galaxies at 200 pc resolution: the absence of large clumps in the interstellar medium at high-redshift. *The Astrophysical Journal*, 859(1):12.

Gunn, J. E., Siegmund, W. A., Mannery, E. J., Owen, R. E., Hull, C. L., Leger, R. F., Carey, L. N., Knapp, G. R., York, D. G., Boroski, W. N., et al. (2006). The 2.5 m telescope of the sloan digital sky survey. *The Astronomical Journal*, 131(4):2332.

Guth, A. H. (1981). Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23(2):347.

Hainline, L. J., Blain, A., Smail, I., Alexander, D., Armus, L., Chapman, S., and Ivison, R. (2011). The stellar mass content of submillimeter-selected galaxies. *The Astrophysical Journal*, 740(2):96.

Hale, C., Robotham, A., Davies, L., Jarvis, M., Driver, S. P., and Heywood, I. (2019). Radio source extraction with profound. *Monthly Notices of the Royal Astronomical Society*, 487(3):3971–3989.

Hale, C., Whittam, I., Jarvis, M., Best, P., Thomas, N., Heywood, I., Prescott, M., Adams, N., Afonso, J., An, F., et al. (2023). Mightee: deep 1.4 ghz source counts and the sky temperature contribution of star-forming galaxies and active galactic nuclei. *Monthly Notices of the Royal Astronomical Society*, 520(2):2668–2691.

Harikane, Y., Ouchi, M., Ono, Y., Fujimoto, S., Donevski, D., Shibuya, T., Faisst, A. L., Goto, T., Hatsukade, B., Kashikawa, N., et al. (2019). Silverrush. viii. spectroscopic identifications of early large-scale structures with protoclusters over 200 mpc at z 6–7: Strong associations of dusty star-forming galaxies. *The Astrophysical Journal*, 883(2):142.

Harrison, I. and Coles, P. (2012). Testing cosmology with extreme galaxy clusters. *Monthly Notices of the Royal Astronomical Society: Letters*, 421(1):L19–L23.

Harwit, M. and Pacini, F. (1975). Infrared galaxies-evolutionary stages of massive star formation. *The Astrophysical Journal*, 200:L127–L129.

Hatch, N. A., Wylezalek, D., Kurk, J., Stern, D., De Breuck, C., Jarvis, M., Galametz, A., Gonzalez,

A., Hartley, W. G., Mortlock, A., et al. (2014). Why z> 1 radio-loud galaxies are commonly located in protoclusters. *Monthly Notices of the Royal Astronomical Society*, 445(1):280–289.

Hatsukade, B., Ohta, K., Seko, A., Yabe, K., and Akiyama, M. (2013). Faint end of 1.3 mm number counts revealed by alma. *The Astrophysical Journal Letters*, 769(2):L27.

Hauser, M. G., Arendt, R. G., Kelsall, T., Dwek, E., Odegard, N., Weiland, J. L., Freudenreich, H. T., Reach, W. T., Silverberg, R. F., Moseley, S. H., Pei, Y. C., Lubin, P., Mather, J. C., Shafer, R. A., Smoot, G. F., Weiss, R., Wilkinson, D. T., and Wright, E. L. (1998). The COBE Diffuse Infrared Background Experiment Search for the Cosmic Infrared Background. I. Limits and Detections. *The Astrophysical Journal*, 508(1):25–43.

Hauser, M. G. and Dwek, E. (2001). The cosmic infrared background: measurements and implications. *Annual Review of Astronomy and Astrophysics*, 39(1):249–307.

Hayward, C. C., Behroozi, P. S., Somerville, R. S., Primack, J. R., Moreno, J., and Wechsler, R. H. (2013). Spatially unassociated galaxies contribute significantly to the blended submillimetre galaxy population: predictions for follow-up observations of alma sources. *Monthly Notices of the Royal Astronomical Society*, 434(3):2572–2581.

Hayward, C. C., Sparre, M., Chapman, S. C., Hernquist, L., Nelson, D., Pakmor, R., Pillepich, A., Springel, V., Torrey, P., Vogelsberger, M., et al. (2021). Submillimetre galaxies in cosmological hydrodynamical simulations–an opportunity for constraining feedback models. *Monthly Notices of the Royal Astronomical Society*, 502(2):2922–2933.

Helou, G., Soifer, B., and Rowan-Robinson, M. (1985). Thermal infrared and nonthermal radio-remarkable correlation in disks of galaxies. *The Astrophysical Journal*, 298:L7–L11.

Hewett, P. C., Warren, S. J., Leggett, S. K., and Hodgkin, S. T. (2006). The ukirt infrared deep sky survey zy jhk photometric system: passbands and synthetic colours. *Monthly Notices of the Royal Astronomical Society*, 367(2):454–468.

Heywood, I. (2020). oxkat: Semi-automated imaging of MeerKAT observations. Astrophysics Source Code Library, record ascl:2009.003.

Heywood, I., Jarvis, M., Hale, C., Whittam, I., Bester, H., Hugo, B., Kenyon, J., Prescott, M., Smirnov, O. M., Tasse, C., et al. (2022). Mightee: total intensity radio continuum imaging and the cosmos/xmm-lss early science fields. *Monthly Notices of the Royal Astronomical Society*, 509(2):2150–2168.

Heywood, I., Jarvis, M. J., Hale, C. L., Whittam, I. H., Bester, H. L., Hugo, B., Kenyon, J. S., Prescott, M., Smirnov, O. M., Tasse, C., Afonso, J. M., Best, P. N., Collier, J. D., Deane, R. P., Frank, B. S., Hardcastle, M. J., Knowles, K., Maddox, N., Murphy, E. J., Prandoni, I., Randriamampandry, S. M., Santos, M. G., Sekhar, S., Tabatabaei, F., Taylor, A. R., and Thorat, K. (2022). MIGHTEE: total intensity radio continuum imaging and the COSMOS/XMM-LSS Early Science fields. *Monthly Notices of the Royal Astronomical Society*, 509(2):2150–2168.

Hill, R., Chapman, S., Phadke, K. A., Aravena, M., Archipley, M., Ashby, M. L., Béthermin, M., Canning, R. E., Gonzalez, A., Greve, T. R., et al. (2022). Rapid build-up of the stellar content in the protocluster core spt2349- 56 at z= 4.3. *Monthly Notices of the Royal Astronomical Society*, 512(3):4352–4377.

Hill, R., Chapman, S., Scott, D., Apostolovski, Y., Aravena, M., Béthermin, M., Bradford, C., Canning, R. E., De Breuck, C., Dong, C., et al. (2020). Megaparsec-scale structure around the protocluster core spt2349–56 at z= 4.3. *Monthly Notices of the Royal Astronomical Society*, 495(3):3124–3159.

Hill, R., Chapman, S. C., Scott, D., Petitpas, G., Smail, I., Chapin, E. L., Gurwell, M. A., Perry, R., Blain, A. W., Bremer, M. N., et al. (2018). High-resolution sma imaging of bright submillimetre sources from the scuba-2 cosmology legacy survey. *Monthly Notices of the Royal Astronomical Society*, 477(2):2042–2067.

Ho, P. T., Moran, J. M., and Lo, K. Y. (2004). The submillimeter array. *The Astrophysical Journal Letters*, 616(1):L1.

Hodge, J., Karim, A., Smail, I., Swinbank, A., Walter, F., Biggs, A., Ivison, R. J., Weiss, A., Alexander, D., Bertoldi, F., et al. (2013). An alma survey of submillimeter galaxies in the extended chandra deep field south: source catalog and multiplicity. *The Astrophysical Journal*, 768(1):91.

Hodge, J., Swinbank, A., Simpson, J., Smail, I., Walter, F., Alexander, D., Bertoldi, F., Biggs, A., Brandt, W., Chapman, S., et al. (2016). Kiloparsec-scale dust disks in high-redshift luminous submillimeter galaxies. *The Astrophysical Journal*, 833(1):103.

Högbom, J. (1974). Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement, Vol. 15, p. 417*, 15:417.

Hogg, D. W., Baldry, I. K., Blanton, M. R., and Eisenstein, D. J. (2002). The k correction. *arXiv preprint astro-ph/0210394*.

Holland, W., Bintley, D., Chapin, E., Chrysostomou, A., Davis, G., Dempsey, J., Duncan, W., Fich, M., Friberg, P., Halpern, M., et al. (2013). Scuba-2: the 10 000 pixel bolometer camera on the james clerk maxwell telescope. *Monthly Notices of the Royal Astronomical Society*, 430(4):2513–2533.

Hopkins, P. F., Cox, T. J., Kereš, D., and Hernquist, L. (2008). A cosmological framework for the co-evolution of quasars, supermassive black holes, and elliptical galaxies. ii. formation of red ellipticals. *The Astrophysical Journal Supplement Series*, 175(2):390.

Houck, J. R., Roellig, T. L., Van Cleve, J., Forrest, W. J., Herter, T., Lawrence, C. R., Matthews, K., Reitsema, H. J., Soifer, B. T., Watson, D. M., et al. (2004). The infrared spectrograph*(irs) on the spitzer space telescope. *The Astrophysical Journal Supplement Series*, 154(1):18.

Howard, J., Burch, D., and Williams, D. (1956a). Infrared transmission of synthetic atmospheres.* ii. absorption by carbon dioxide. *JOSA*, 46(4):237–241.

Howard, J., Burch, D., and Williams, D. (1956b). Infrared transmission of synthetic atmospheres.* iii. absorption by water vapor. *JOSA*, 46(4):242–245.

Hubble, E. (1929). A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the national academy of sciences*, 15(3):168–173.

Hung, C.-L., Casey, C. M., Chiang, Y.-K., Capak, P. L., Cowley, M. J., Darvish, B., Kacprzak, G. G., Kovač, K., Lilly, S. J., Nanayakkara, T., et al. (2016). Large-scale structure around a z= 2.1 cluster. *The Astrophysical Journal*, 826(2):130.

Ilbert, O., Arnouts, S., McCracken, H., Bolzonella, M., Bertin, E., Le Fevre, O., Mellier, Y., Zamorani, G., Pello, R., Iovino, A., et al. (2006). Accurate photometric redshifts for the cfht legacy survey calibrated using the vimos vlt deep survey. *Astronomy & Astrophysics*, 457(3):841–856.

Inoue, A. K. (2011). Rest-frame ultraviolet-to-optical spectral characteristics of extremely metal-poor and metal-free galaxies. *Monthly Notices of the Royal Astronomical Society*, 415(3):2920–2931.

Ito, K., Tanaka, M., Valentino, F., Toft, S., Brammer, G., Gould, K. M., Ilbert, O., Kashikawa, N., Kubo, M., Liang, Y., et al. (2023). Cosmos2020: Discovery of a protocluster of massive quiescent galaxies at z= 2.77. *The Astrophysical Journal Letters*, 945(1):L9.

Ivezić, Ž., Kahn, S. M., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., Alonso, D., AlSayyad, Y., Anderson, S. F., Andrew, J., et al. (2019). Lsst: from science drivers to reference design and anticipated data products. *The Astrophysical Journal*, 873(2):111.

Ivison, R., Alexander, D. M., Biggs, A. D., Brandt, W., Chapin, E. L., Coppin, K. E., Devlin, M. J., Dickinson, M., Dunlop, J., Dye, S., et al. (2010a). Blast: the far-infrared/radio correlation in distant galaxies. *Monthly Notices of the Royal Astronomical Society*, 402(1):245–258.

Ivison, R., Biggs, A., Bremer, M., Arumugam, V., and Dunne, L. (2020). Alma unveils wider environment of distant red protocluster core. *Monthly Notices of the Royal Astronomical Society*, 496(4):4358–4365.

Ivison, R., Greve, T., Dunlop, J., Peacock, J., Egami, E., Smail, I., Ibar, E., Van Kampen, E., Aretxaga, I., Babbedge, T., et al. (2007). The scuba half degree extragalactic survey–iii. identification of radio and mid-infrared counterparts to submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 380(1):199–228.

Ivison, R., Lewis, A., Weiss, A., Arumugam, V., Simpson, J., Holland, W., Maddox, S., Dunne, L., Valiante, E., Van der Werf, P., et al. (2016). The space density of luminous dusty star-forming galaxies at z> 4: Scuba-2 and laboca imaging of ultrared galaxies from herschel-atlas. *The Astrophysical Journal*, 832(1):78.

Ivison, R., Magnelli, B., Ibar, E., Andreani, P., Elbaz, D., Altieri, B., Amblard, A., Arumugam,

V., Auld, R., Aussel, H., et al. (2010b). The far-infrared/radio correlation as probed by herschel. *Astronomy & Astrophysics*, 518:L31.

Ivison, R., Smail, I., Le Borgne, J.-F., Blain, A., Kneib, J.-P., Bezecourt, J., Kerr, T., and Davies, J. (1998). A hyperluminous galaxy at z= 2.8 found in a deep submillimetre survey. *Monthly Notices of the Royal Astronomical Society*, 298(2):583–593.

Jarvis, M., Taylor, R., Agudo, I., Allison, J. R., Deane, R. P., Frank, B., Gupta, N., Heywood, I., Maddox, N., McAlpine, K., Santos, M., Scaife, A. M. M., Vaccari, M., Zwart, J. T. L., Adams, E., Bacon, D. J., Baker, A. J., Bassett, B. A., Best, P. N., Beswick, R., Blyth, S., Brown, M. L., Bruggen, M., Cluver, M., Colafrancesco, S., Cotter, G., Cress, C., Davé, R., Ferrari, C., Hardcastle, M. J., Hale, C. L., Harrison, I., Hatfield, P. W., Klockner, H. R., Kolwa, S., Malefahlo, E., Marubini, T., Mauch, T., Moodley, K., Morganti, R., Norris, R. P., Peters, J. A., Prandoni, I., Prescott, M., Oliver, S., Oozeer, N., Rottgering, H. J. A., Seymour, N., Simpson, C., Smirnov, O., and Smith, D. J. B. (2016). The MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE) Survey. In *MeerKAT Science: On the Pathway to the SKA*, page 6.

Jarvis, M. J., Smith, D., Bonfield, D. G., Hardcastle, M., Falder, J., Stevens, J., Ivison, R. J., Auld, R., Baes, M., Baldry, I., et al. (2010). Herschel-atlas: the far-infrared–radio correlation at z< 0.5. *Monthly Notices of the Royal Astronomical Society*, 409(1):92–101.

Jin, S., Sillassen, N. B., Hodge, J., Magdis, G. E., Casey, C., Rizzo, F., Koekemoer, A. M., Valentino, F., Kokorev, V., Magnelli, B., et al. (2024). A photo-z cautionary tale: Redshift confirmation of cosbo-7 at $z = 2.625$. *arXiv preprint arXiv:2407.07585*.

Jonas, J. L. (2009). Meerkat—the south african array with composite dishes and wide-band single pixel feeds. *Proceedings of the IEEE*, 97(8):1522–1530.

Karim, A., Swinbank, A., Hodge, J., Smail, I., Walter, F., Biggs, A., Simpson, J., Danielson, A., Alexander, D., Bertoldi, F., et al. (2013). An alma survey of submillimetre galaxies in the extended chandra deep field south: high-resolution 870 μm source counts. *Monthly Notices of the Royal Astronomical Society*, 432(1):2–9.

Kawara, K., Matsuhara, H., Okuda, H., Taniguchi, Y., Sato, Y., Sofue, Y., Wakamatsu, K., Oyabu, S., Sanders, D., and Cowie, L. (2004). Iso deep far-infrared survey in the "lockman hole"-iii. catalogs and source counts at 90 & 170 μm. *Astronomy & Astrophysics*, 413(3):843–859.

Kenyon, J. S., Smirnov, O. M., Grobler, T. L., and Perkins, S. J. (2018). CUBICAL - fast radio interferometric calibration suite exploiting complex optimization. *Monthly Notices of the Royal Astronomical Society*, 478(2):2399–2415.

Knudsen, K., van Der Werf, P., and Kneib, J.-P. (2008). Probing the submillimetre number counts at f 850 μm< 2 mjy. *Monthly Notices of the Royal Astronomical Society*, 384(4):1611–1626.

Koprowski, M., Dunlop, J. S., Michałowski, M. J., Coppin, K., Geach, J., McLure, R. J., Scott, D., and van der Werf, P. (2017). The evolving far-ir galaxy luminosity function and dust-obscured star formation rate density out to. *Monthly Notices of the Royal Astronomical Society*, 471(4):4155–4169.

Kravtsov, A. V. and Borgani, S. (2012). Formation of galaxy clusters. *Annual Review of Astronomy and Astrophysics*, 50:353–409.

Kron, R. G. (1980). Photometry of a complete sample of faint galaxies. *Astrophysical Journal Supplement Series, vol. 43, June 1980, p. 305-325. Research supported by the University of California*, 43:305–325.

Kuijken, K. et al. (2011). Omegacam: Eso's newest imager. *The Messenger*, 146(8).

Kuijken, K., Heymans, C., Dvornik, A., Hildebrandt, H., de Jong, J., Wright, A., Erben, T., Bilicki, M., Giblin, B., Shan, H.-Y., et al. (2019). The fourth data release of the kilo-degree survey: ugri imaging and nine-band optical-ir photometry over 1000 square degrees. *Astronomy & Astrophysics*, 625:A2.

Kuijken, K., Heymans, C., Dvornik, A., Hildebrandt, H., de Jong, J. T. A., Wright, A. H., Erben, T., Bilicki, M., Giblin, B., Shan, H. Y., Getman, F., Grado, A., Hoekstra, H., Miller, L., Napolitano, N., Paolilo, M., Radovich, M., Schneider, P., Sutherland, W., Tewes, M., Tortora, C., Valentijn, E. A., and Verdoes Kleijn, G. A. (2019). The fourth data release of the Kilo-Degree Survey: ugri imaging and nine-band optical-IR photometry over 1000 square degrees. *Astronomy and Astrophysics*, 625:A2.

Kurtzer, G. M., Sochat, V., and Bauer, M. W. (2017). Singularity: Scientific containers for mobility of compute. *PLoS ONE*, 12(5):e0177459.

Lacy, M., Surace, J. A., Farrah, D., Nyland, K., Afonso, J., Brandt, W. N., Clements, D., Lagos, C., Maraston, C., Pforr, J., et al. (2021). A spitzer survey of deep drilling fields to be targeted by the vera c. rubin observatory legacy survey of space and time. *Monthly Notices of the Royal Astronomical Society*, 501(1):892–910.

Lagache, G., Puget, J.-L., and Dole, H. (2005). Dusty infrared galaxies: sources of the cosmic infrared background. *Annu. Rev. Astron. Astrophys.*, 43:727–768.

Laureijs, R., Amiaux, J., Arduini, S., Augueres, J. L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., Pasian, F., Percival, W., Rhodes, J., Saavedra Criado, G., Sauvage, M., Scaramella, R., Valenziano, L., Warren, S., Bender, R., Castander, F., Cimatti, A., Le Fèvre, O., Kurki-Suonio, H., Levi, M., Lilje, P., Meylan, G., Nichol, R., Pedersen, K., Popa, V., Rebolo Lopez, R., Rix, H. W., Rottgering, H., Zeilinger, W., Grupp, F., Hudelot, P., Massey, R., Meneghetti, M., Miller, L., Paltani, S., Paulin-Henriksson, S., Pires, S., Saxton, C., Schrabback, T., Seidel, G., Walsh, J., Aghanim, N., Amendola, L., Bartlett, J., Baccigalupi, C., Beaulieu, J. P., Benabed, K., Cuby, J. G., Elbaz, D., Fosalba, P., Gavazzi, G., Helmi, A., Hook, I., Irwin, M., Kneib, J. P., Kunz, M., Mannucci, F., Moscardini, L., Tao, C., Teyssier, R., Weller, J., Zamorani, G., Zapatero Osorio, M. R., Boulade, O., Foumond, J. J., Di Giorgio, A., Guttridge, P., James, A., Kemp, M., Martignac, J., Spencer, A., Walton, D., Blümchen, T., Bonoli, C., Bortoletto, F., Cerna, C., Corcione, L., Fabron, C., Jahnke, K., Ligori, S., Madrid, F., Martin, L., Morgante, G., Pamplona, T., Prieto, E., Riva, M., Toledo, R., Trifoglio, M., Zerbi, F., Abdalla, F., Douspis, M., Grenet, C., Borgani, S., Bouwens, R., Courbin, F., Delouis, J. M., Dubath, P., Fontana, A., Frailis, M., Grazian, A., Koppenhöfer, J., Mansutti, O., Melchior, M., Mignoli, M., Mohr, J., Neissner, C., Noddle, K., Poncet, M., Scodeggio, M., Serrano, S., Shane, N., Starck, J. L., Surace, C., Taylor, A., Verdoes-Kleijn, G., Vuerli, C., Williams, O. R., Zacchei, A., Altieri, B., Escudero Sanz, I., Kohley, R., Oosterbroek, T., Astier, P., Bacon, D., Bardelli, S., Baugh, C., Bellagamba, F., Benoist, C., Bianchi, D., Biviano, A., Branchini, E., Carbone, C., Cardone, V., Clements, D., Colombi, S.,

Conselice, C., Cresci, G., Deacon, N., Dunlop, J., Fedeli, C., Fontanot, F., Franzetti, P., Giocoli, C., Garcia-Bellido, J., Gow, J., Heavens, A., Hewett, P., Heymans, C., Holland, A., Huang, Z., Ilbert, O., Joachimi, B., Jennins, E., Kerins, E., Kiessling, A., Kirk, D., Kotak, R., Krause, O., Lahav, O., van Leeuwen, F., Lesgourgues, J., Lombardi, M., Magliocchetti, M., Maguire, K., Majerotto, E., Maoli, R., Marulli, F., Maurogordato, S., McCracken, H., McLure, R., Melchiorri, A., Merson, A., Moresco, M., Nonino, M., Norberg, P., Peacock, J., Pello, R., Penny, M., Pettorino, V., Di Porto, C., Pozzetti, L., Quercellini, C., Radovich, M., Rassat, A., Roche, N., Ronayette, S., Rossetti, E., Sartoris, B., Schneider, P., Semboloni, E., Serjeant, S., Simpson, F., Skordis, C., Smadja, G., Smartt, S., Spano, P., Spiro, S., Sullivan, M., Tilquin, A., Trotta, R., Verde, L., Wang, Y., Williger, G., Zhao, G., Zoubian, J., and Zucca, E. (2011). Euclid Definition Study Report. *arXiv e-prints*, page arXiv:1110.3193.

Le Floc'h, E., Papovich, C., Dole, H., Bell, E. F., Lagache, G., Rieke, G. H., Egami, E., Pérez-González, P. G., Alonso-Herrero, A., Rieke, M. J., et al. (2005). Infrared luminosity functions from the chandra deep field-south: the spitzer view on the history of dusty star formation at 0< z< 1. *The Astrophysical Journal*, 632(1):169.

Lemaux, B., Cucciati, O., Le Fèvre, O., Zamorani, G., Lubin, L., Hathi, N., Ilbert, O., Pelliccia, D., Amorín, R., Bardelli, S., et al. (2022). The vimos ultra deep survey: The reversal of the star-formation rate- density relation at 2< z< 5. *Astronomy & Astrophysics*, 662:A33.

Lemke, D., Klaas, U., Abolins, J., Abraham, P., Acosta-Pulido, J., Bogun, S., Castaneda, H., Cornwall, L., Drury, L., Gabriel, C., et al. (1996). Isophot-capabilities and performance. *Astronomy and Astrophysics*, 315:L64–L70.

Lewis, I. J., Cannon, R. D., Taylor, K., Glazebrook, K., Bailey, J., Baldry, I., Barton, J., Bridges, T., Dalton, G., Farrell, T., et al. (2002). The anglo-australian observatory 2df facility. *Monthly Notices of the Royal Astronomical Society*, 333(2):279–298.

Licquia, T. C., Newman, J. A., and Brinchmann, J. (2015). Unveiling the milky way: a new technique for determining the optical color and luminosity of our galaxy. *The Astrophysical Journal*, 809(1):96.

Lim, C.-F., Wang, W.-H., Smail, I., Scott, D., Chen, C.-C., Chang, Y.-Y., Simpson, J. M., Toba, Y., Shu,

X., Clements, D., et al. (2020). Scuba-2 ultra deep imaging eao survey (studies). iii. multiwavelength properties, luminosity functions, and preliminary source catalog of 450 μm selected galaxies. *The Astrophysical Journal*, 889(2):80.

Ling, C., Sun, B., Cheng, C., Li, N., Ma, Z., and Yan, H. (2024). A strongly lensed dusty starburst of an intrinsic disk morphology at photometric redshift of 7.7. *arXiv preprint arXiv:2402.16649*.

Long, A. S., Cooray, A., Ma, J., Casey, C. M., Wardlow, J. L., Nayyeri, H., Ivison, R., Farrah, D., and Dannerbauer, H. (2020). Emergence of an ultrared, ultramassive galaxy cluster core at z= 4. *The Astrophysical Journal*, 898(2):133.

Lonsdale, C. J., Farrah, D., and Smith, H. E. (2006). Ultraluminous infrared galaxies. *Astrophysics Update 2*, pages 285–336.

Lovell, C. C., Geach, J. E., Davé, R., Narayanan, D., and Li, Q. (2021). Reproducing submillimetre galaxy number counts with cosmological hydrodynamic simulations. *Monthly Notices of the Royal Astronomical Society*, 502(1):772–793.

Ma, J., Cooray, A., Nayyeri, H., Brown, A., Ghotbi, N., Ivison, R., Oteo, I., Duivenvoorden, S., Greenslade, J., Clements, D., et al. (2019). Spitzer catalog of herschel-selected ultrared dusty star-forming galaxies. *The Astrophysical Journal Supplement Series*, 244(2):30.

MacKenzie, T. P., Scott, D., Bianconi, M., Clements, D. L., Dole, H. A., Flores-Cacho, I., Guery, D., Kneissl, R., Lagache, G., Marleau, F. R., et al. (2017). Scuba-2 follow-up of herschel-spire observed planck overdensities. *Monthly Notices of the Royal Astronomical Society*, 468(4):4006–4017.

Madau, P. and Dickinson, M. (2014). Cosmic star-formation history. *Annual Review of Astronomy and Astrophysics*, 52:415–486.

Maddox, S. J., Valiante, E., Cigan, P., Dunne, L., Eales, S., Smith, M. W. L., Dye, S., Furlanetto, C., Ibar, E., De Zotti, G., et al. (2018). The herschel-atlas data release 2. paper ii. catalogs of far-infrared and submillimeter sources in the fields at the south and north galactic poles. *The Astrophysical Journal Supplement Series*, 236(2):30.

Magnelli, B., Elbaz, D., Chary, R., Dickinson, M., Le Borgne, D., Frayer, D., and Willmer, C. (2011). Evolution of the dusty infrared luminosity function from z= 0 to z= 2.3 using observations from spitzer. *Astronomy & Astrophysics*, 528:A35.

Magnelli, B., Lutz, D., Berta, S., Altieri, B., Andreani, P., Aussel, H., Castaneda, H., Cava, A., Cepa, J., Cimatti, A., et al. (2010). Far-infrared properties of submillimeter and optically faint radio galaxies. *Astronomy & Astrophysics*, 518:L28.

Magnelli, B., Popesso, P., Berta, S., Pozzi, F., Elbaz, D., Lutz, D., Dickinson, M., Altieri, B., Andreani, P., Aussel, H., et al. (2013). The deepest herschel-pacs far-infrared survey: number counts and infrared luminosity functions from combined pep/goods-h observations. *Astronomy & Astrophysics*, 553:A132.

Marinacci, F., Vogelsberger, M., Pakmor, R., Torrey, P., Springel, V., Hernquist, L., Nelson, D., Weinberger, R., Pillepich, A., Naiman, J., and Genel, S. (2018). First results from the IllustrisTNG simulations: radio haloes and magnetic fields. *Monthly Notices of the Royal Astronomical Society*, 480(4):5113–5139.

Martinache, C., Rettura, A., Dole, H., Lehnert, M., Frye, B., Altieri, B., Beelen, A., Béthermin, M., Le Floc'h, E., Giard, M., et al. (2018). Spitzer planck herschel infrared cluster (spheric) survey: Candidate galaxy clusters at 1.3< z< 3 selected by high star-formation rate. *Astronomy & Astrophysics*, 620:A198.

McAlpine, S., Smail, I., Bower, R. G., Swinbank, A., Trayford, J. W., Theuns, T., Baes, M., Camps, P., Crain, R. A., and Schaye, J. (2019). The nature of submillimetre and highly star-forming galaxies in the eagle simulation. *Monthly Notices of the Royal Astronomical Society*, 488(2):2440–2454.

McConachie, I., Wilson, G., Forrest, B., Marsan, Z. C., Muzzin, A., Cooper, M., Annunziatella, M., Marchesini, D., Chan, J. C., Gomez, P., et al. (2022). Spectroscopic confirmation of a protocluster at z= 3.37 with a high fraction of quiescent galaxies. *The Astrophysical Journal*, 926(1):37.

McCracken, H., Milvang-Jensen, B., Dunlop, J., Franx, M., Fynbo, J., Le Fèvre, O., Holt, J., Caputi, K., Goranova, Y., Buitrago, F., et al. (2012). Ultravista: a new ultra-deep near-infrared survey in cosmos. *Astronomy & Astrophysics*, 544:A156.

McCully, C., Crawford, S., Kovacs, G., Tollerud, E., Betts, E., Bradley, L., Craig, M., Turner, J., Streicher, O., Sipocz, B., Robitaille, T., and Deil, C. (2018). astropy/astroscrappy: v1.0.5 zenodo release.

McKinney, J., Manning, S. M., Cooper, O. R., Long, A. S., Akins, H., Casey, C. M., Faisst, A. L., Franco, M., Hayward, C. C., Lambrides, E., et al. (2023). A near-infrared-faint, far-infrared-luminous dusty galaxy at z 5 in cosmos-web. *The Astrophysical Journal*, 956(2):72.

McMullin, J., Waters, B., Schiebel, D., Young, W., Golap, K., Shaw, R. A., Hill, F., and Bell, D. J. (2007). Asp conf. ser. vol. 376, astronomical data analysis software and systems xvi.

Michałowski, M. J., Dunlop, J., Koprowski, M., Cirasuolo, M., Geach, J., Bowler, R., Mortlock, A., Caputi, K., Aretxaga, I., Arumugam, V., et al. (2017). The scuba-2 cosmology legacy survey: the nature of bright submm galaxies from 2 deg2 of 850-μm imaging. *Monthly Notices of the Royal Astronomical Society*, 469(1):492–515.

Michałowski, M. J., Dunlop, J. S., Cirasuolo, M., Hjorth, J., Hayward, C. C., and Watson, D. (2012). The stellar masses and specific star-formation rates of submillimetre galaxies. *Astronomy & Astrophysics*, 541:A85.

Michałowski, M. J., Hayward, C. C., Dunlop, J. S., Bruce, V. A., Cirasuolo, M., Cullen, F., and Hernquist, L. (2014). Determining the stellar masses of submillimetre galaxies: the critical importance of star formation histories. *Astronomy & Astrophysics*, 571:A75.

Miller, T., Chapman, S., Aravena, M., Ashby, M., Hayward, C., Vieira, J., Weiß, A., Babul, A., Béthermin, M., Bradford, C., et al. (2018). A massive core for a cluster of galaxies at a redshift of 4.3. *Nature*, 556(7702):469–472.

Miller, T. B., Hayward, C. C., Chapman, S. C., and Behroozi, P. S. (2015). The bias of the submillimetre galaxy population: Smgs are poor tracers of the most-massive structures in the z 2 universe. *Monthly Notices of the Royal Astronomical Society*, 452(1):878–883.

Mocanu, L., Crawford, T., Vieira, J., Aird, K., Aravena, M., Austermann, J., Benson, B., Béthermin, M., Bleem, L., Bothwell, M., et al. (2013). Extragalactic millimeter-wave point-source catalog,

number counts and statistics from 771 deg2 of the spt-sz survey. *The Astrophysical Journal*, 779(1):61.

Mohan, N. and Rafferty, D. (2015). Pybdsf: Python blob detection and source finder. *Astrophysics Source Code Library*, pages ascl–1502.

Molnár, D. C., Sargent, M. T., Leslie, S., Magnelli, B., Schinnerer, E., Zamorani, G., Delhaize, J., Smolčić, V., Tisanić, K., and Vardoulaki, E. (2021). The non-linear infrared-radio correlation of low-z galaxies: implications for redshift evolution, a new radio sfr recipe, and how to minimize selection bias. *Monthly Notices of the Royal Astronomical Society*, 504(1):118–145.

Moneti, A., Mccracken, H., Gozaliasl, G., Keihänen, E., Kirkpatrick, C., Kurki-Suonio, H., Lindholm, V., Väliviita, J., Collaboration, E., et al. (2022). Euclid preparation: Xvii. cosmic dawn survey: Spitzer space telescope observations of the euclid deep fields and calibration fields. *Astronomy & Astrophysics*.

Moneti, A., McCracken, H., Rouberol, S., et al. (2019). The fourth ultravista data release.

Montaña, A., Zavala, J., Aretxaga, I., Hughes, D., Ivison, R., Pope, A., Sánchez-Argüelles, D., Wilson, G., Yun, M., Cantua, O., et al. (2021). Early science with the large millimeter telescope: a 1.1 mm aztec survey of red-herschel dusty star-forming galaxies. *Monthly Notices of the Royal Astronomical Society*, 505(4):5260–5282.

Moorwood, A., Cuby, J.-G., and Lidman, C. (1998). Sofi sees first light at the ntt. *The Messenger, vol. 91, p. 9-13*, 91:9–13.

Muldrew, S. I., Hatch, N. A., and Cooke, E. A. (2015). What are protoclusters?–defining high-redshift galaxy clusters and protoclusters. *Monthly Notices of the Royal Astronomical Society*, 452(3):2528–2539.

Mullaney, J., Pannella, M., Daddi, E., Alexander, D., Elbaz, D., Hickox, R., Bournaud, F., Altieri, B., Aussel, H., Coia, D., et al. (2012). Goods-herschel: the far-infrared view of star formation in active galactic nucleus host galaxies since z= 3. *Monthly Notices of the Royal Astronomical Society*, 419(1):95–115.

Naiman, J. P., Pillepich, A., Springel, V., Ramirez-Ruiz, E., Torrey, P., Vogelsberger, M., Pakmor, R., Nelson, D., Marinacci, F., Hernquist, L., Weinberger, R., and Genel, S. (2018). First results from the IllustrisTNG simulations: a tale of two elements - chemical evolution of magnesium and europium. *Monthly Notices of the Royal Astronomical Society*, 477(1):1206–1224.

Narayanan, D., Turk, M., Feldmann, R., Robitaille, T., Hopkins, P., Thompson, R., Hayward, C., Ball, D., Faucher-Giguere, C.-A., and Kereš, D. (2015). The formation of submillimetre-bright galaxies from gas infall over a billion years. *Nature*, 525(7570):496–499.

Negrello, M., Amber, S., Amvrosiadis, A. e., Cai, Z.-Y., Lapi, A., Gonzalez-Nuevo, J., De Zotti, G., Furlanetto, C., Maddox, S., Allen, M., et al. (2016). The herschel-atlas: a sample of 500μm-selected lensed galaxies over 600 square degrees. *Monthly Notices of the Royal Astronomical Society*, page stw2911.

Negrello, M., González-Nuevo, J., Magliocchetti, M., Moscardini, L., De Zotti, G., Toffolatti, L., and Danese, L. (2005). Effect of clustering on extragalactic source counts with low-resolution instruments. *Monthly Notices of the Royal Astronomical Society*, 358(3):869–874.

Negrello, M., Hopwood, R., De Zotti, G., Cooray, A., Verma, A., Bock, J., Frayer, D., Gurwell, M., Omont, A., Neri, R., et al. (2010). The detection of a population of submillimeter-bright, strongly lensed galaxies. *science*, 330(6005):800–804.

Negrello, M., Perrotta, F., González, J. G.-N., Silva, L., De Zotti, G., Granato, G., Baccigalupi, C., and Danese, L. (2007). Astrophysical and cosmological information from large-scale submillimetre surveys of extragalactic sources. *Monthly Notices of the Royal Astronomical Society*, 377(4):1557–1568.

Nelson, D., Pillepich, A., Springel, V., Weinberger, R., Hernquist, L., Pakmor, R., Genel, S., Torrey, P., Vogelsberger, M., Kauffmann, G., Marinacci, F., and Naiman, J. (2018). First results from the IllustrisTNG simulations: the galaxy colour bimodality. *Monthly Notices of the Royal Astronomical Society*, 475(1):624–647.

Neugebauer, G., Habing, H., Van Duinen, R., Aumann, H., Baud, B., Beichman, C., Beintema, D.,

Boggess, N., Clegg, P., De Jong, T., et al. (1984). The infrared astronomical satellite (iras) mission. *The Astrophysical Journal*, 278:L1–L6.

Noethe, L., Franza, F., Giordano, P., Wilson, R., Citterio, O., Conti, G., and Mattaini, E. (1988). Active optics ii. results of an experiment with a thin 1 m test mirror. *Journal of modern optics*, 35(9):1427–1457.

Novak, M., Smolčić, V., Delhaize, J., Delvecchio, I., Zamorani, G., Baran, N., Bondi, M., Capak, P., Carilli, C., Ciliegi, P., et al. (2017). The vla-cosmos 3 ghz large project: Cosmic star formation history since z˜ 5. *Astronomy & Astrophysics*, 602:A5.

Offringa, A. R., de Bruyn, A. G., Biehl, M., Zaroubi, S., Bernardi, G., and Pandey, V. N. (2010). Post-correlation radio frequency interference classification methods. *Monthly Notices of the Royal Astronomical Society*, 405(1):155–167.

Offringa, A. R., McKinley, B., Hurley-Walker, N., Briggs, F. H., Wayth, R. B., Kaplan, D. L., Bell, M. E., Feng, L., Neben, A. R., Hughes, J. D., Rhee, J., Murphy, T., Bhat, N. D. R., Bernardi, G., Bowman, J. D., Cappallo, R. J., Corey, B. E., Deshpand e, A. A., Emrich, D., Ewall-Wice, A., Gaensler, B. M., Goeke, R., Greenhill, L. J., Hazelton, B. J., Hindson, L., Johnston-Hollitt, M., Jacobs, D. C., Kasper, J. C., Kratzenberg, E., Lenc, E., Lonsdale, C. J., Lynch, M. J., McWhirter, S. R., Mitchell, D. A., Morales, M. F., Morgan, E., Kudryavtseva, N., Oberoi, D., Ord, S. M., Pindor, B., Procopio, P., Prabu, T., Riding, J., Roshi, D. A., Shankar, N. U., Srivani, K. S., Subrahmanyan, R., Tingay, S. J., Waterson, M., Webster, R. L., Whitney, A. R., Williams, A., and Williams, C. L. (2014). WSCLEAN: an implementation of a fast, generic wide-field imager for radio astronomy. *Monthly Notices of the Royal Astronomical Society*, 444(1):606–619.

Oliver, S., Bock, J., Altieri, B., Amblard, A., Arumugam, V., Aussel, H., Babbedge, T., Beelen, A., Béthermin, M., Blain, A., et al. (2012). The herschel multi-tiered extragalactic survey: Hermes. *Monthly Notices of the Royal Astronomical Society*, 424(3):1614–1635.

Onoue, M., Kashikawa, N., Uchiyama, H., Akiyama, M., Harikane, Y., Imanishi, M., Komiyama, Y., Matsuoka, Y., Nagao, T., Nishizawa, A. J., et al. (2018). Enhancement of galaxy overdensity

around quasar pairs at z< 3.6 based on the hyper suprime-cam subaru strategic program survey. *Publications of the Astronomical Society of Japan*, 70(SP1):S31.

Oteo, I., Ivison, R., Dunne, L., Manilla-Robles, A., Maddox, S., Lewis, A., De Zotti, G., Bremer, M., Clements, D., Cooray, A., et al. (2018). An extreme protocluster of luminous dusty starbursts in the early universe. *The Astrophysical Journal*, 856(1):72.

Oteo, I., Ivison, R., Negrello, M., Smail, I., Pérez-Fournon, I., Bremer, M., De Zotti, G., Eales, S., Farrah, D., Temi, P., et al. (2017). Witnessing the birth of the red sequence: the physical scale and morphology of dust emission in hyper-luminous starbursts in the early universe. *arXiv preprint arXiv:1709.04191*.

Overzier, R. A. (2016). The realm of the galaxy protoclusters. *The Astronomy and Astrophysics Review*, 24(1):1–57.

Overzier, R. A., Miley, G., Bouwens, R., Cross, N., Zirm, A., Benítez, N., Blakeslee, J., Clampin, M., Demarco, R., Ford, H., et al. (2006). Clustering of star-forming galaxies near a radio galaxy at z= 5.2. *The Astrophysical Journal*, 637(1):58.

Pantoni, L., Lapi, A., Massardi, M., Donevski, D., Bressan, A., Silva, L., Pozzi, F., Vignali, C., Talia, M., Cimatti, A., et al. (2021). Unveiling the nature of 11 dusty star-forming galaxies at the peak of cosmic star formation history. *Monthly Notices of the Royal Astronomical Society*, 504(1):928–950.

Patanchon, G., Ade, P. A., Bock, J. J., Chapin, E. L., Devlin, M. J., Dicker, S. R., Griffin, M., Gundersen, J. O., Halpern, M., Hargrave, P. C., et al. (2009). Submillimeter number counts from statistical analysis of blast maps. *The Astrophysical Journal*, 707(2):1750.

Pearson, W., Wang, L., Hurley, P., Małek, K., Buat, V., Burgarella, D., Farrah, D., Oliver, S., Smith, D., and van der Tak, F. (2018). Main sequence of star forming galaxies beyond the herschel confusion limit. *Astronomy & Astrophysics*, 615:A146.

Peebles, P. (2012). Seeing cosmology grow. *Annual Review of Astronomy and Astrophysics*, 50(1):1–28.

Pérez-Martínez, J. M., Dannerbauer, H., Kodama, T., Koyama, Y., Shimakawa, R., Suzuki, T. L., Calvi, R., Chen, Z., Daikuhara, K., Hatch, N. A., et al. (2023). Signs of environmental effects on star-forming galaxies in the spiderweb protocluster at z= 2.16. *Monthly Notices of the Royal Astronomical Society*, 518(2):1707–1734.

Perivolaropoulos, L. and Skara, F. (2022). Challenges for λcdm: An update. *New Astronomy Reviews*, 95:101659.

Perkins, S. (2022). Distributed Streaming Radio Astronomy Reduction with Dask. In Ruiz, J. E., Pierfedereci, F., and Teuben, P., editors, *Astronomical Society of the Pacific Conference Series*, volume 532 of *Astronomical Society of the Pacific Conference Series*, page 337.

Piacentini, F., Reach, W., Savini, G., and Smoot, G. (2011). Planck early results. vii. the early release compact source catalogue. *A&A*, 536:A7.

Pilbratt, G., Riedinger, J., Passvogel, T., Crone, G., Doyle, D., Gageur, U., Heras, A., Jewell, C., Metcalfe, L., Ott, S., et al. (2010). Herschel space observatory-an esa facility for far-infrared and submillimetre astronomy. *Astronomy & Astrophysics*, 518:L1.

Pillepich, A., Nelson, D., Hernquist, L., Springel, V., Pakmor, R., Torrey, P., Weinberger, R., Genel, S., Naiman, J. P., Marinacci, F., and Vogelsberger, M. (2018). First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies. *Monthly Notices of the Royal Astronomical Society*, 475(1):648–675.

Poglitsch, A., Waelkens, C., Bauer, O. H., Cepa, J., Feuchtgruber, H., Henning, T., Van Hoof, C., Kerschbaum, F., Krause, O., Renotte, E., et al. (2010). The herschel photodetector array camera & spectrometer (pacs): design and in-flight operation and scientific performance. *38th COSPAR Scientific Assembly*, 38:13.

Pope, A., Chary, R.-R., Alexander, D. M., Armus, L., Dickinson, M., Elbaz, D., Frayer, D., Scott, D., and Teplitz, H. (2008). Mid-infrared spectral diagnosis of submillimeter galaxies. *The Astrophysical Journal*, 675(2):1171.

Price-Whelan, A. M., Lim, P. L., Earl, N., Starkman, N., Bradley, L., Shupe, D. L., Patil, A. A., Corrales, L., Brasseur, C., Nöthe, M., et al. (2022). The astropy project: sustaining and growing a

community-oriented open-source project and the latest major release (v5. 0) of the core package. *The Astrophysical Journal*, 935(2):167.

Price-Whelan, A. M., Sipőcz, B., Günther, H., Lim, P., Crawford, S., Conseil, S., Shupe, D., Craig, M., Dencheva, N., Ginsburg, A., et al. (2018). The astropy project: Building an open-science project and status of the v2. 0 core package. *The Astronomical Journal*, 156(3):123.

Primack, J. R. (2012). Triumphs and tribulations of λcdm, the double dark theory. *Annalen der Physik*, 524(9-10):535–544.

Puget, J.-L., Abergel, A., Bernard, J.-P., Boulanger, F., Burton, W., Desert, F.-X., and Hartmann, D. (1996). Tentative detection of a cosmic far-infrared background with cobe. *Astronomy and Astrophysics*, 308:L5.

Ramasawmy, J., Stevens, J., Martin, G., and Geach, J. E. (2019). A flat trend of star formation rate with x-ray luminosity of galaxies hosting agn in the scuba-2 cosmology legacy survey. *Monthly Notices of the Royal Astronomical Society*, 486(3):4320–4333.

Riechers, D. A., Bradford, C., Clements, D., Dowell, C., Pérez-Fournon, I., Ivison, R., Bridge, C., Conley, A., Fu, H., Vieira, J., et al. (2013). A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34. *Nature*, 496(7445):329–333.

Riechers, D. A., Cooray, A., Pérez-Fournon, I., and Neri, R. (2021). The gadot galaxy survey: Dense gas and feedback in herschel-selected starburst galaxies at redshifts 2 to 6. *The Astrophysical Journal*, 913(2):141.

Riechers, D. A., Hodge, J. A., Pavesi, R., Daddi, E., Decarli, R., Ivison, R. J., Sharon, C. E., Smail, I., Walter, F., Aravena, M., et al. (2020). Coldz: A high space density of massive dusty starburst galaxies 1 billion years after the big bang. *The Astrophysical Journal*, 895(2):81.

Rieke, G. H., Young, E., Engelbracht, C., Kelly, D., Low, F., Haller, E., Beeman, J., Gordon, K., Stansberry, J., Misselt, K., et al. (2004). The multiband imaging photometer for spitzer (mips). *The Astrophysical Journal Supplement Series*, 154(1):25.

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., et al. (1998). Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The astronomical journal*, 116(3):1009.

Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., et al. (2013). Astropy: A community python package for astronomy. *Astronomy & Astrophysics*, 558:A33.

Rodríguez-Puebla, A., Avila-Reese, V., Cano-Díaz, M., Faber, S., Primack, J. R., Franco, J., Aretxaga, I., and Santiago-Mayoral, E. (2020). The star-forming main sequence and the contribution of dust-obscured star formation since z 4 from the far-uv+ ir luminosity functions. *The Astrophysical Journal*, 905(2):171.

Roseboom, I., Ivison, R. J., Greve, T., Amblard, A., Arumugam, V., Auld, R., Aussel, H., Bethermin, M., Blain, A., Bock, J., et al. (2012). The herschel multi-tiered extragalactic survey: Spire–mm photometric redshifts. *Monthly Notices of the Royal Astronomical Society*, 419(4):2758–2773.

Rotermund, K., Chapman, S., Phadke, K., Hill, R., Pass, E., Aravena, M., Ashby, M., Babul, A., Béthermin, M., Canning, R., et al. (2021). Optical and near-infrared observations of the spt2349-56 proto-cluster core at z= 4.3. *Monthly Notices of the Royal Astronomical Society*, 502(2):1797–1815.

Rowan-Robinson, M. (2000). Hyperluminous infrared galaxies. *Monthly Notices of the Royal Astronomical Society*, 316(4):885–900.

Rowan-Robinson, M., Oliver, S., Wang, L., Farrah, D., Clements, D. L., Gruppioni, C., Marchetti, L., Rigopoulou, D., and Vaccari, M. (2016). The star formation rate density from z= 1 to 6. *Monthly Notices of the Royal Astronomical Society*, 461(1):1100–1111.

Rowan-Robinson, M., Wang, L., Farrah, D., Rigopoulou, D., Gruppioni, C., Vaccari, M., Marchetti, L., Clements, D. L., and Pearson, W. J. (2018). Extreme submillimetre starburst galaxies. *Astronomy & Astrophysics*, 619:A169.

Salmon, B., Coe, D., Bradley, L., Bouwens, R., Bradač, M., Huang, K.-H., Oesch, P. A., Stark, D., Sharon, K., Trenti, M., et al. (2020). Relics: The reionization lensing cluster survey and the brightest high-z galaxies. *The Astrophysical Journal*, 889(2):189.

Sanders, D., Mazzarella, J., Kim, D.-C., Surace, J., and Soifer, B. (2003). The iras revised bright galaxy sample. *The Astronomical Journal*, 126(4):1607.

Sanders, D. and Mirabel, I. (1996). Luminous infrared galaxies. *Annual Review of Astronomy and Astrophysics*, 34(1):749–792.

Sanders, D., Scoville, N., and Soifer, B. (1991). Molecular gas in luminous infrared galaxies. *The Astrophysical Journal*, 370:158–171.

Sargent, M. T., Schinnerer, E., Murphy, E., Carilli, C., Helou, G., Aussel, H., Le Floc'h, E., Frayer, D., Ilbert, O., Oesch, P., et al. (2010). No evolution in the ir–radio relation for ir-luminous galaxies at z< 2 in the cosmos field. *The Astrophysical Journal Letters*, 714(2):L190.

Saunders, W., Sutherland, W. J., Maddox, S. J., Keeble, O., Oliver, S. J., Rowan-Robinson, M., McMahon, R. G., Efstathiou, G. P., Tadros, H., White, S. D. M., Frenk, C. S., Carramiñana, A., and Hawkins, M. R. S. (2000). The PSCz catalogue. *Monthly Notices of the Royal Astronomical Society*, 317(1):55–63.

Schaye, J., Crain, R. A., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I., Helly, J. C., et al. (2015). The eagle project: simulating the evolution and assembly of galaxies and their environments. *Monthly Notices of the Royal Astronomical Society*, 446(1):521–554.

Scheuer, P. A. G. (1974). Fluctuations in the X-ray background. *Monthly Notices of the Royal Astronomical Society*, 166:329–338.

Scott, K., Yun, M., Wilson, G. W., Austermann, J., Aguilar, E., Aretxaga, I., Ezawa, H., Ferrusca, D., Hatsukade, B., Hughes, D., et al. (2010). Deep 1.1 mm-wavelength imaging of the goods-s field by aztec/aste–i. source catalogue and number counts. *Monthly Notices of the Royal Astronomical Society*, 405(4):2260–2278.

Scoville, N., Aussel, H., Brusa, M., Capak, P., Carollo, C. M., Elvis, M., Giavalisco, M., Guzzo, L., Hasinger, G., Impey, C., et al. (2007). The cosmic evolution survey (cosmos): overview. *The Astrophysical Journal Supplement Series*, 172(1):1.

Serjeant, S., Negrello, M., Pearson, C., Mortier, A., Austermann, J., Aretxaga, I., Clements, D., Chapman, S., Dye, S., Dunlop, J., et al. (2010). The agn fraction of submm-selected galaxies and contributions to the submm/mm-wave extragalactic background light. *Astronomy & Astrophysics*, 514:A10.

Shi, K., Lee, K.-S., Dey, A., Huang, Y., Malavasi, N., Hung, C.-L., Inami, H., Ashby, M., Duncan, K., Xue, R., et al. (2019). A census of galaxy constituents in a coma progenitor observed at z> 3. *The Astrophysical Journal*, 871(1):83.

Shi, K., Toshikawa, J., Cai, Z., Lee, K.-S., and Fang, T. (2020). A detailed study of massive galaxies in a protocluster at z= 3.13. *The Astrophysical Journal*, 899(1):79.

Shim, H., Lee, D., Kim, Y., Scott, D., Serjeant, S., Ao, Y., Barrufet, L., Chapman, S. C., Clements, D. L., Conselice, C. J., et al. (2022). Multiwavelength properties of 850-μm selected sources from the north ecliptic pole scuba-2 survey. *Monthly Notices of the Royal Astronomical Society*, 514(2):2915–2935.

Shimakawa, R., Kodama, T., Hayashi, M., Prochaska, J. X., Tanaka, I., Cai, Z., Suzuki, T. L., Tadaki, K.-i., and Koyama, Y. (2018a). Mahalo deep cluster survey i. accelerated and enhanced galaxy formation in the densest regions of a protocluster at z= 2.5. *Monthly Notices of the Royal Astronomical Society*, 473(2):1977–1999.

Shimakawa, R., Koyama, Y., Röttgering, H. J., Kodama, T., Hayashi, M., Hatch, N. A., Dannerbauer, H., Tanaka, I., Tadaki, K.-i., Suzuki, T. L., et al. (2018b). Mahalo deep cluster survey ii. characterizing massive forming galaxies in the spiderweb protocluster at z= 2.2. *Monthly Notices of the Royal Astronomical Society*, 481(4):5630–5650.

Shimizu, I., Yoshida, N., and Okamoto, T. (2012). Submillimetre galaxies in cosmological hydrodynamic simulations: source number counts and the spatial clustering. *Monthly Notices of the Royal Astronomical Society*, 427(4):2866–2875.

Simpson, J., Smail, I., Swinbank, A., Chapman, S., Geach, J., Ivison, R., Thomson, A., Aretxaga, I., Blain, A. W., Cowley, W., et al. (2015). The scuba-2 cosmology legacy survey: Alma resolves the bright-end of the sub-millimeter number counts. *The Astrophysical Journal*, 807(2):128.

Simpson, J., Swinbank, A., Smail, I., Alexander, D., Brandt, W., Bertoldi, F., de Breuck, C., Chapman, S. C., Coppin, K., da Cunha, E., et al. (2014). An alma survey of submillimeter galaxies in the extended chandra deep field south: the redshift distribution and evolution of submillimeter galaxies. *The Astrophysical Journal*, 788(2):125.

Siringo, G., Kreysa, E., Kovács, A., Schuller, F., Weiß, A., Esch, W., Gemünd, H.-P., Jethava, N., Lundershausen, G., Colin, A., et al. (2009). The large apex bolometer camera laboca. *Astronomy & Astrophysics*, 497(3):945–962.

Smail, I., Geach, J. E., Swinbank, A. M., Tadaki, K., Arumugam, V., Hartley, W., Almaini, O., Bremer, M. N., Chapin, E., Chapman, S., et al. (2014). The scuba-2 cosmology legacy survey: Ultraluminous star-forming galaxies in a z= 1.6 cluster. *The Astrophysical Journal*, 782(1):19.

Smail, I., Ivison, R., and Blain, A. (1997). A deep submillimeter survey of lensing clusters: A new window on galaxy formation and evolution. *The Astrophysical Journal Letters*, 490(1):L5.

Smail, I., Ivison, R., Blain, A., and Kneib, J.-P. (2002). The nature of faint submillimetre-selected galaxies. *Monthly Notices of the Royal Astronomical Society*, 331(2):495–520.

Smirnov, O. M. and Tasse, C. (2015). Radio interferometric gain calibration as a complex optimization problem. *Monthly Notices of the Royal Astronomical Society*, 449(3):2668–2684.

Smith, D., Dunne, L., Maddox, S. J., Eales, S., Bonfield, D., Jarvis, M., Sutherland, W., Fleuren, S., Rigby, E., Thompson, M., et al. (2011). Herschel–atlas: counterparts from the ultraviolet–near-infrared in the science demonstration phase catalogue. *Monthly Notices of the Royal Astronomical Society*, 416(2):857–872.

Smith, D., Hayward, C., Jarvis, M., and Simpson, C. (2017a). A complete distribution of redshifts for submillimetre galaxies in the scuba-2 cosmology legacy survey uds field. *Monthly Notices of the Royal Astronomical Society*, 471(2):2453–2462.

Smith, M. W., Ibar, E., Maddox, S. J., Valiante, E., Dunne, L., Eales, S., Dye, S., Furlanetto, C., Bourne, N., Cigan, P., et al. (2017b). The herschel–atlas data release 2, paper i. submillimeter and far-infrared images of the south and north galactic poles: the largest herschel survey of the extragalactic sky. *The Astrophysical Journal Supplement Series*, 233(2):26.

Soifer, B. and Neugebauer, G. (1991). The properties of infrared galaxies in the local universe. *Astronomical Journal*, 101(2):354–361.

Somerville, R. S., Gilmore, R. C., Primack, J. R., and Domínguez, A. (2012). Galaxy properties from the ultraviolet to the far-infrared: λ cold dark matter models confront observations. *Monthly Notices of the Royal Astronomical Society*, 423(3):1992–2015.

Sparke, L. S. and Gallagher, J. S. (2000). *Galaxies in the universe: an introduction*. Cambridge University Press.

Speagle, J. S., Steinhardt, C. L., Capak, P. L., and Silverman, J. D. (2014). A highly consistent framework for the evolution of the star-forming "main sequence" from z 0–6. *The Astrophysical Journal Supplement Series*, 214(2):15.

Springel, V., Pakmor, R., Pillepich, A., Weinberger, R., Nelson, D., Hernquist, L., Vogelsberger, M., Genel, S., Torrey, P., Marinacci, F., and Naiman, J. (2018). First results from the IllustrisTNG simulations: matter and galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, 475(1):676–698.

Stach, S. M., Dudzevičiūtė, U., Smail, I., Swinbank, A., Geach, J., Simpson, J., An, F. X., Almaini, O., Arumugam, V., Blain, A. W., et al. (2019). An alma survey of the scuba-2 cosmology legacy survey ukidss/uds field: source catalogue and properties. *Monthly Notices of the Royal Astronomical Society*, 487(4):4648–4668.

Stach, S. M., Smail, I., Swinbank, A., Simpson, J., Geach, J., An, F. X., Almaini, O., Arumugam, V., Blain, A., Chapman, S., et al. (2018). An alma survey of the scuba-2 cosmology legacy survey ukidss/uds field: number counts of submillimeter galaxies. *The Astrophysical Journal*, 860(2):161.

Stanley, F., Harrison, C., Alexander, D., Swinbank, A., Aird, J., Del Moro, A., Hickox, R., and Mullaney, J. (2015). A remarkably flat relationship between the average star formation rate and agn luminosity for distant x-ray agn. *Monthly Notices of the Royal Astronomical Society*, 453(1):591–604.

Stevens, J. A., Ivison, R., Dunlop, J., Smail, I. R., Percival, W., Hughes, D., Röttgering, H.,

Van Breugel, W., and Reuland, M. (2003). The formation of cluster elliptical galaxies as revealed by extensive star formation. *Nature*, 425(6955):264–267.

Strauss, M. A., Huchra, J. P., Davis, M., Yahil, A., Fisher, K. B., and Tonry, J. (1992). A Redshift Survey of IRAS Galaxies. VII. The Infrared and Redshift Data for the 1.936 Jansky Sample. *Astrophysical Journal Supplement*, 83:29.

Sutherland, W., Emerson, J., Dalton, G., Atad-Ettedgui, E., Beard, S., Bennett, R., Bezawada, N., Born, A., Caldwell, M., Clark, P., et al. (2015). The visible and infrared survey telescope for astronomy (vista): design, technical overview, and performance. *Astronomy & Astrophysics*, 575:A25.

Swinbank, A., Simpson, J., Smail, I., Harrison, C., Hodge, J., Karim, A., Walter, F., Alexander, D., Brandt, W., de Breuck, C., et al. (2014). An alma survey of sub-millimetre galaxies in the extended chandra deep field south: the far-infrared properties of smgs. *Monthly Notices of the Royal Astronomical Society*, 438(2):1267–1287.

Swinbank, A., Smail, I., Chapman, S. C., Blain, A. W., Ivison, R. J., and Keel, W. C. (2004). The rest-frame optical spectra of scuba galaxies. *The Astrophysical Journal*, 617(1):64.

Tacconi, L., Genzel, R., Neri, R., Cox, P., Cooper, M., Shapiro, K., Bolatto, A., Bouché, N., Bournaud, F., Burkert, A., et al. (2010). High molecular gas fractions in normal massive star-forming galaxies in the young universe. *Nature*, 463(7282):781–784.

Tacconi, L., Genzel, R., Smail, I., Neri, R., Chapman, S., Ivison, R., Blain, A., Cox, P., Omont, A., Bertoldi, F., et al. (2008). Submillimeter galaxies at z˜ 2: Evidence for major mergers and constraints on lifetimes, imf, and co-h2 conversion factor. *The Astrophysical Journal*, 680(1):246.

Tacconi, L. J., Neri, R., Chapman, S., Genzel, R., Smail, I., Ivison, R., Bertoldi, F., Blain, A., Cox, P., Greve, T., et al. (2006). High-resolution millimeter imaging of submillimeter galaxies. *The Astrophysical Journal*, 640(1):228.

Tasse, C. (2023). killMS: Direction-dependent radio interferometric calibration package. Astrophysics Source Code Library, record ascl:2305.005.

Tasse, C., Hugo, B., Mirmont, M., Smirnov, O., Atemkeng, M., Bester, L., Hardcastle, M. J., Lakhoo, R., Perkins, S., and Shimwell, T. (2018). Faceting for direction-dependent spectral deconvolution. *Astronomy & Astrophysics*, 611:A87.

Tasse, C., Hugo, B., Mirmont, M., Smirnov, O., Atemkeng, M., Bester, L., Hardcastle, M. J., Lakhoo, R., Perkins, S., and Shimwell, T. (2023). DDFacet: Facet-based radio imaging package. Astrophysics Source Code Library, record ascl:2305.008.

Taylor, J. H. and Yates, H. W. (1957). Atmospheric transmission in the infrared. *JOSA*, 47(3):223–226.

Taylor, M. B. (2005). Topcat & stil: Starlink table/votable processing software. In *Astronomical data analysis software and systems XIV*, volume 347, page 29.

Thompson, A. R., Moran, J. M., and Swenson, G. W. (2017). *Interferometry and synthesis in radio astronomy*. Springer Nature.

Tody, D. (1986). The iraf data reduction and analysis system. In *Instrumentation in astronomy VI*, volume 627, pages 733–748. SPIE.

Toft, S., Smolčić, V., Magnelli, B., Karim, A., Zirm, A., Michalowski, M., Capak, P., Sheth, K., Schawinski, K., Krogager, J.-K., et al. (2014). Submillimeter galaxies as progenitors of compact quiescent galaxies. *The Astrophysical Journal*, 782(2):68.

Toshikawa, J., Kashikawa, N., Overzier, R., Malkan, M. A., Furusawa, H., Ishikawa, S., Onoue, M., Ota, K., Tanaka, M., Niino, Y., et al. (2016). A systematic survey of protoclusters at z 3–6 in the cfhtls deep fields. *The Astrophysical Journal*, 826(2):114.

Toshikawa, J., Uchiyama, H., Kashikawa, N., Ouchi, M., Overzier, R., Ono, Y., Harikane, Y., Ishikawa, S., Kodama, T., Matsuda, Y., et al. (2018). Goldrush. iii. a systematic search for protoclusters at z 4 based on the> 100 deg2 area. *Publications of the Astronomical Society of Japan*, 70(SP1):S12.

Traina, A., Gruppioni, C., Delvecchio, I., Calura, F., Bisigello, L., Feltre, A., Magnelli, B., Schinnerer, E., Liu, D., Adscheid, S., et al. (2024). A3cosmos: The infrared luminosity function and dust-obscured star formation rate density at 0.5< z< 6. *Astronomy & Astrophysics*, 681:A118.

Umehata, H., Tamura, Y., Kohno, K., Ivison, R., Smail, I., Hatsukade, B., Nakanishi, K., Kato, Y., Ikarashi, S., Matsuda, Y., et al. (2017). Alma deep field in ssa22: source catalog and number counts. *The Astrophysical Journal*, 835(1):98.

Valiante, E., Smith, M., Eales, S., Maddox, S., Ibar, E., Hopwood, R., Dunne, L., Cigan, P., Dye, S., Pascale, E., et al. (2016). The herschel-atlas data release 1–i. maps, catalogues and number counts. *Monthly Notices of the Royal Astronomical Society*, 462(3):3146–3179.

Van der Kruit, P. (1971). Observations of core sources in seyfert and normal galaxies with the westerbork synthesis radio telescope at 1415 mhz. *Astronomy and Astrophysics*, 15:110–122.

Van Dokkum, P. G. (2001). Cosmic-ray rejection by laplacian edge detection. *Publications of the Astronomical Society of the Pacific*, 113(789):1420.

Vieira, J., Crawford, T., Switzer, E., Ade, P. A., Aird, K., Ashby, M., Benson, B., Bleem, L., Brodwin, M., Carlstrom, J., et al. (2010). Extragalactic millimeter-wave sources in south pole telescope survey data: source counts, catalog, and statistics for an 87 square-degree field. *The Astrophysical Journal*, 719(1):763.

Vieira, J., Marrone, D., Chapman, S., De Breuck, C., Hezaveh, Y., Weiβ, A., Aguirre, J., Aird, K., Aravena, M., Ashby, M., et al. (2013). Dusty starburst galaxies in the early universe as revealed by gravitational lensing. *Nature*, 495(7441):344–347.

Wang, F., Yang, J., Fan, X., Hennawi, J. F., Barth, A. J., Banados, E., Bian, F., Boutsia, K., Connor, T., Davies, F. B., et al. (2021). A luminous quasar at redshift 7.642. *The Astrophysical Journal Letters*, 907(1):L1.

Wang, L., Pearson, W., Cowley, W., Trayford, J., Béthermin, M., Gruppioni, C., Hurley, P., and Michałowski, M. (2019). Multi-wavelength de-blended herschel view of the statistical properties of dusty star-forming galaxies across cosmic time. *Astronomy & Astrophysics*, 624:A98.

Wang, R., Wagg, J., Carilli, C. L., Walter, F., Lentati, L., Fan, X., Riechers, D. A., Bertoldi, F., Narayanan, D., Strauss, M. A., et al. (2013). Star formation and gas kinematics of quasar host galaxies at z 6: new insights from alma. *The Astrophysical Journal*, 773(1):44.

Wang, W.-H., Cowie, L. L., Barger, A. J., and Williams, J. P. (2010). Sma* observations of goods 850–11 and goods 850–13: First examples of multiple submillimeter sources resolved by an interferometer. *The Astrophysical Journal Letters*, 726(2):L18.

Wang, W.-H., Lin, W.-C., Lim, C.-F., Smail, I., Chapman, S. C., Zheng, X. Z., Shim, H., Kodama, T., Almaini, O., Ao, Y., et al. (2017). Scuba-2 ultra deep imaging eao survey (studies): faint-end counts at 450 μm. *The Astrophysical Journal*, 850(1):37.

Wardlow, J. L., Cooray, A., De Bernardis, F., Amblard, A., Arumugam, V., Aussel, H., Baker, A., Béthermin, M., Blundell, R., Bock, J., et al. (2012). Hermes: candidate gravitationally lensed galaxies and lensing statistics at submillimeter wavelengths. *The Astrophysical Journal*, 762(1):59.

Wardlow, J. L., Simpson, J., Smail, I., Swinbank, A., Blain, A., Brandt, W., Chapman, S., Chen, C.-C., Cooke, E., Dannerbauer, H., et al. (2018). An alma survey of co in submillimetre galaxies: companions, triggering, and the environment in blended sources. *Monthly Notices of the Royal Astronomical Society*, 479(3):3879–3891.

Wardlow, J. L., Smail, I., Coppin, K., Alexander, D., Brandt, W., Danielson, A., Luo, B., Swinbank, A., Walter, F., Weiß, A., et al. (2011). The laboca survey of the extended chandra deep field-south: a photometric redshift survey of submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 415(2):1479–1508.

Weaver, J. R., Kauffmann, O., Ilbert, O., McCracken, H., Moneti, A., Toft, S., Brammer, G., Shuntov, M., Davidzon, I., Hsieh, B., et al. (2022). Cosmos2020: A panchromatic view of the universe to z 10 from two complementary catalogs. *The Astrophysical Journal Supplement Series*, 258(1):11.

Weiß, A., De Breuck, C., Marrone, D., Vieira, J., Aguirre, J., Aird, K., Aravena, M., Ashby, M., Bayliss, M., Benson, B., et al. (2013). Alma redshifts of millimeter-selected galaxies from the spt survey: the redshift distribution of dusty star-forming galaxies. *The Astrophysical Journal*, 767(1):88.

Werner, M. W., Roellig, T., Low, F., Rieke, G. H., Rieke, M., Hoffmann, W., Young, E., Houck, J., Brandl, B., Fazio, G., et al. (2004). The spitzer space telescope mission. *The Astrophysical Journal Supplement Series*, 154(1):1.

Wilkinson, A., Almaini, O., Chen, C.-C., Smail, I., Arumugam, V., Blain, A., Chapin, E. L., Chapman, S. C., Conselice, C. J., Cowley, W. I., et al. (2017). The scuba-2 cosmology legacy survey: the clustering of submillimetre galaxies in the ukidss uds field. *Monthly Notices of the Royal Astronomical Society*, 464(2):1380–1392.

Wilson, R., Franza, F., Giordano, P., Noethe, L., and Tarenghi, M. (1989). Active optics: Iii. final results with the 1 m test mirror and ntt 3· 58 m primary in the workshop. *Journal of modern optics*, 36(11):1415–1425.

Wilson, R., Franza, F., Noethe, L., and Andreoni, G. (1991). Active optics: Iv. set-up and performance of the optics of the eso new technology telescope (ntt) in the observatory. *Journal of modern optics*, 38(2):219–243.

Wilson, R. N., Franza, F., and Noethe, L. (1987). Active optics. I. A system for optimizing the optical quality and reducing the costs of large telescopes. *Journal of Modern Optics*, 34(4):485–509.

Wootten, A. and Thompson, A. R. (2009). The atacama large millimeter/submillimeter array. *Proceedings of the IEEE*, 97(8):1463–1471.

Wright, E. L., Eisenhardt, P. R., Mainzer, A. K., Ressler, M. E., Cutri, R. M., Jarrett, T., Kirkpatrick, J. D., Padgett, D., McMillan, R. S., Skrutskie, M., et al. (2010). The wide-field infrared survey explorer (wise): mission description and initial on-orbit performance. *The Astronomical Journal*, 140(6):1868.

Yan, H., Ma, Z., Huang, J.-S., and Fan, L. (2020). A systematic search for the reddest far-infrared and submillimeter galaxies: Revealing dust-embedded starbursts at high redshifts. *The Astrophysical Journal Supplement Series*, 249(1):1.

York, D. G., Adelman, J., Anderson Jr, J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J., Barkhouser, R., Bastian, S., Berman, E., et al. (2000). The sloan digital sky survey: Technical summary. *The Astronomical Journal*, 120(3):1579.

Yun, M. S. and Carilli, C. (2002). Radio-to-far-infrared spectral energy distribution and photometric redshifts for dusty starburst galaxies. *The Astrophysical Journal*, 568(1):88.

Yun, M. S., Reddy, N. A., and Condon, J. (2001). Radio properties of infrared-selected galaxies in the iras 2 jy sample. *The Astrophysical Journal*, 554(2):803.

Zavala, J., Aretxaga, I., Geach, J., Hughes, D., Birkinshaw, M., Chapin, E., Chapman, S., Chen, C.-C., Clements, D., Dunlop, J., et al. (2016). The scuba-2 cosmology legacy survey: The egs deep field i–deep number counts and the redshift distribution of the recovered cosmic infrared background at 450 and 850 μm. *Monthly Notices of the Royal Astronomical Society*, page stw2630.

Zavala, J., Casey, C., Manning, S., Aravena, M., Bethermin, M., Caputi, K., Clements, D., Da Cunha, E., Drew, P., Finkelstein, S., et al. (2021). The evolution of the ir luminosity function and dust-obscured star formation over the past 13 billion years. *The Astrophysical Journal*, 909(2):165.

Zavala, J. A., Buat, V., Casey, C. M., Finkelstein, S. L., Burgarella, D., Bagley, M. B., Ciesla, L., Daddi, E., Dickinson, M., Ferguson, H. C., et al. (2023). Dusty starbursts masquerading as ultra-high redshift galaxies in jwst ceers observations. *The Astrophysical journal letters*, 943(2):L9.

Zavala, J. A., Casey, C. M., da Cunha, E., Spilker, J., Staguhn, J., Hodge, J., and Drew, P. (2018a). Constraining the volume density of dusty star-forming galaxies through the first 3 mm number counts from alma. *The Astrophysical Journal*, 869(1):71.

Zavala, J. A., Montaña, A., Hughes, D. H., Yun, M. S., Ivison, R., Valiante, E., Wilner, D., Spilker, J., Aretxaga, I., Eales, S., et al. (2018b). A dusty star-forming galaxy at z= 6 revealed by strong gravitational lensing. *Nature Astronomy*, 2(1):56–62.

Zhang, Y., Zheng, X. Z., Shi, D. D., Gao, Y., Dannerbauer, H., An, F. X., Shu, X., Gao, Z.-K., Wang, W.-H., Wang, X., et al. (2022). Submillimetre galaxies in two massive protoclusters at z= 2.24: witnessing the enrichment of extreme starbursts in the outskirts of hae density peaks. *Monthly Notices of the Royal Astronomical Society*, 512(4):4893–4908.

Zheng, X. Z., Cai, Z., An, F. X., Fan, X., and Shi, D. D. (2021). Mammoth: confirmation of two massive galaxy overdensities at z= 2.24 with hα emitters. *Monthly Notices of the Royal Astronomical Society*, 500(4):4354–4364.

Zou, H., Zhou, X., Fan, X., Zhang, T., Zhou, Z., Nie, J., Peng, X., McGreer, I., Jiang, L., Dey, A.,

et al. (2017). Project overview of the beijing–arizona sky survey. *Publications of the Astronomical Society of the Pacific*, 129(976):064101.

Zubko, V., Dwek, E., and Arendt, R. G. (2004). Interstellar dust models consistent with extinction, emission, and abundance constraints. *The Astrophysical Journal Supplement Series*, 152(2):211.