GALAXY EVOLUTION AS
SEEN BY *HERSCHEL*

A Dissertation presented to

the Faculty of the Graduate School

at the University of Missouri

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

ZHIYUAN MA

Dr. Haojing Yan, Dissertation Supervisor

MAY 2018

The undersigned, appointed by the Dean of the Graduate School, have examined the Dissertation entitled:

GALAXY EVOLUTION AS

SEEN BY *HERSCHEL*

presented by Zhiyuan Ma,

a candidate for the degree of Doctor of Philosophy and hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Haojing Yan

Dr. Adam Helfer

Dr. Aigen Li

Dr. Angela Speck

Dr. Sergei Kopeikin

ACKNOWLEDGMENTS

It is not the end but start — to the stars, my destination.

I'd like to thank my advisor Haojing Yan for his tremendous help and support. His guidance helped me in all the time of research and writing of this thesis. I am also grateful to all of those with whom I have had the pleasure to work over the past few years, including my committee members, the professors in the department, and all the colleagues and friends. Thank you!

Follows are the official acknowledgements:

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The unprecedented *Herschel* data have revolutionized our view of the IR universe, providing new insights to galaxy formation and evolution. However, the *Herschel* data are still far away from being fully explored, mainly due to three reasons. First, in most of the *Herschel* wide fields, we are still lacking ancillary data even in the optical/near-IR, which prevents us from doing any follow-up analysis. Second, even in the few fields that have sufficient ancillary data in various archives, there is no homogeneous reduction such that the data can be readily used. Third, but not the least, the long-standing problem of poor spatial resolution in FIR/sub-mm observations severely limits our capability of counterpart identification, and hence creates tremendous obstacles in obtaining redshifts, modeling the spectral energy distributions, deriving star formation rates, etc.

In this thesis, I present our approach to attack these problems. My basic task is to establish a methodology that can be applied to all the *Herschel* wide-field survey data, with the science goal of constructing the largest, well-defined sample of high-redshift Ultra-Luminous InfarRed Galaxies (ULIRGs) whose optical-to-far-IR SEDs are accurately measured. We first conduct a comprehensive search of the available archival optical data in the public domain. Moreover, we have been conducting our own deep optical survey, Mizzou WIYN Survey, to observe those fields that lack full range optical data coverage. To facilitate the data reduction process, I have developed `APUS`, a pipeline building and management tool that offers great repeatability, scalability, and maintainability. With this tool, we can reduce all these data in a uniform manner and create deep optical mosaics. To combat the problem of poor spatial resolution in the *Herschel* data, we have developed a source de-blending algorithm, which I have implemented as a software tool called `CIDer`. Using this tool, we can extract the major contributors to the FIR fluxes based on the position priors from

the high-resolution optical data.

Treating all the *Herschel* wide-field data is obviously beyond the scope of an individual thesis. Therefore, my thesis mainly focuses on one field, namely the "First Look Survey" field. As the first study case, we compiled and reduced the full-range multi-wavelength data in this field, and constructed the panchromatic SEDs from optical to FIR using `CIDer`. The redshifts of the sources were found to be ranging from $z = 0$ to ~ 3. We derived the total IR luminosities by fitting modified black body model or starbursting templates, which are found to be ranges from 3×10^{10} to $\sim 10^{13} \, \mathrm{L_\odot}$ after taking into account the cut to ensure reliability of the `CIDer` result. On the L_{IR}–T plane, the sources show the similar turning over behavior as in our previous study Yan & Ma (2016), implies a limited size of the dust star forming sites.

Also presented in the thesis is our earlier work on the *Herschel*-detected SDSS quasars. The project was conducted before we have the `CIDer` to obtain the full-range SEDs and thereby the redshifts. Instead, we aimed at the quasars, which we can measure the redshift spectroscopically and to a high redshift. In the project, it is found that the *Herschel*-detected quasars are mostly ULIRGs, and are forming stars very actively. This implies the co-evolution of extreme AGNs and star formation.

Chapter 1

Introduction

1.1 Galaxy Formation and Evolution

In the field of extragalactic astronomy, two of the most fundamental questions to ask are: how the galaxies that we observe today were formed and assembled to their current form; and how the galaxy formation and evolution changed with respect to the cosmic time. The study presented in this thesis aims to contribute on this matter using the unprecedented *Herschel* data.

1.1.1 Gas, Dust, and Star Formation

Gas, dust, and stars are the most important baryonic components of galaxies. The key to study the galaxy formation and evolution is to understand the interplay of them and quantify the star formation.

Stars form in dense region of the gaseous material, commonly referred to as the (giant) molecular clouds (GMCs). GMCs are dense, typically have molecular hydrogen densities that are orders of magnitude above the density of ionized hydrogen in other regions of the interstellar medium (ISM). Within the molecular clouds, there are

also dusts, which are produced by dead and exploded stars in the previous generation.

Denser regions of the ISM are able to efficiently cool through molecular line emission. There is a higher chance for gas of higher density to have interaction between the hydrogen ions within. H_2 is believed to be formed from the recombination of two hydrogen atoms on the surface of a dust grain. Molecules like H_2 and CO can be collisionally excited then radiatively de-excited, taking away energy and causing the cloud to cool. As the cloud continues to gravitationally collapse, the cooling continues. This process could effectively cool the GMC to as low as $\sim 10\,\mathrm{K}$, which is far colder than other regions of the ISM.

As the collapse goes, the core will become opaque to the photons and energy is hence trapped within. This leads to eventually the nuclear fusion in the center, marking the birth of new stars. The stars will follow an evolutionary path that is dependent on their initial mass, and due their high surface temperature, most of the energy is emitted in the UV to optical wavelengths.

The star formation rate (SFR; measured in the unit of $\mathrm{M_\odot\,yr^{-1}}$) is used to quantify how fast the gas is turning into stars in a star forming site or galaxy. To measure the SFR, the most direct way is to count how many young stellar objects (YSOs) are there within a given molecular cloud. This, however, only works for star forming sites that are close to us because it requires resolved observation of the individual YSOs.

For measuring star formation rate in other galaxies, we have to rely on measuring the bulk of the light emitted during the star forming processes, which are referred to as the star formation rate indicators.

The most traditionally used SFR indicator is the UV light. The youngest stellar populations (O and B stars) emit the bulk of their energy in the restframe UV ($< 0.3\,\mathrm{\mu m}$). For a Kroupa stellar initial mass function (IMF), with constant star formation over a time scale of $> 100\,\mathrm{Myr}$, the non-ionizing UV ($912\,\mathrm{\mathring{A}} < \lambda < 3000\,\mathrm{\mathring{A}}$) stellar continuum can be converted to SFR with

$$\frac{\text{SFR}_{\text{UV}}}{\text{M}_\odot\,\text{yr}^{-1}} = 3 \times 10^{-47} \frac{\lambda L(\lambda)}{\text{\AA}\,\text{erg}\,\text{s}^{-1}}\,, \tag{1.1}$$

assuming a solar metallicity and starburst99 SED template (Leitherer et al., 1999).

In the early universe, the first batch of the stars formed from the collapsed gas clouds, and exploded soon after at supernovae (SNe), and the metals synthesized during the stars' lifetime and the SN events were returned to the surrounding environment, in the form of the dust. Dust plays an important role in studying the star formation because it is ubiquitous in the star forming sites and affects the measurement of star formation rate.

1.1.2 Dust Extinction

Due to the present of dust in the star forming sites, using UV light as the SFR indicator suffers from the dust extinction. The UV and optical light is scattered and absorbed by the enshrouding dust, causing the gas/dust mixture to heat up in regions near the new stars, and this eventually leads to emission of the reprocessed infrared light. The emission spectrum is largely blackbody-like (BB), and is commonly referred to as the "modified blackbody" (MBB) emission, which differs from the BB by an additional parameter called emissivity. Depending on the amount of surrounding dust, one may find, in some cases, all the UV light could be blocked and only the reprocessed IR light could be observed; or in more general situation, only part of the UV light is obscured thus both UV and IR emission could be detected.

Since essentially the IR light is from reprocessed UV light which traces the star formation, and the dust and gas co-exist in the star forming sites, we could use IR light as the probe to measure the SFR that is hidden by the dust. In Kennicutt

(1998), assuming a Chabrier IMF[1], the relation between SFR and the IR luminosity is given as

$$\frac{\text{SFR}_{\text{IR}}}{\text{M}_\odot \, \text{yr}^{-1}} = 1.0 \times 10^{-10} \frac{\text{L}_{IR}}{\text{L}_\odot} \,, \tag{1.2}$$

where L_{IR} is the bolometric IR luminosity, traditionally defined to be the luminosity density integrated from 8 to 1000 µm.

From the theoretic point of view, one can also estimate the effect of the dust on the observed spectral energy distribution (SED) for a star forming site/galaxy, given the knowledge of the relation between dust attenuation and the wavelengths. The relation is referred to as the dust extinction laws or reddening laws. One of the most used dust extinction law in extragalactic context is proposed in Calzetti (1997). With the dust extinction law, we can make "corrections" to the measured stellar light so that obscured star formation could be taken in to account, with the absent of measurements in the IR wavelengths.

1.1.3   Cosmic Star Formation History

Looking at a broader picture, one question to ask is how the galaxy formation and evolution in the whole universe changes with respect to the cosmic time. Staring from 1990's, a number of large deep optical surveys such as the Canada-France-Redshift Survey (CFRS; Lilly et al., 1995) and the *Hubble Space Telescope* (*HST*) greatly enlarged the number of galaxies at cosmic distance out to $z \sim 5$. Lilly et al. (1996) and Madau et al. (1996) made the first attempts to connect the luminosity densities over a wide redshift range, which led to the construction of the famous "Lilly-Madau" plot that shows the evolution of the global star formation rate density (GSFRD) of the universe as a function of redshift (top panel of Figure 1.1). The star formation

---

[1]The conversion would be a factor of ~ 1.7 higher if using a Salpeter IMF, which was adopted in Kennicutt (1998).
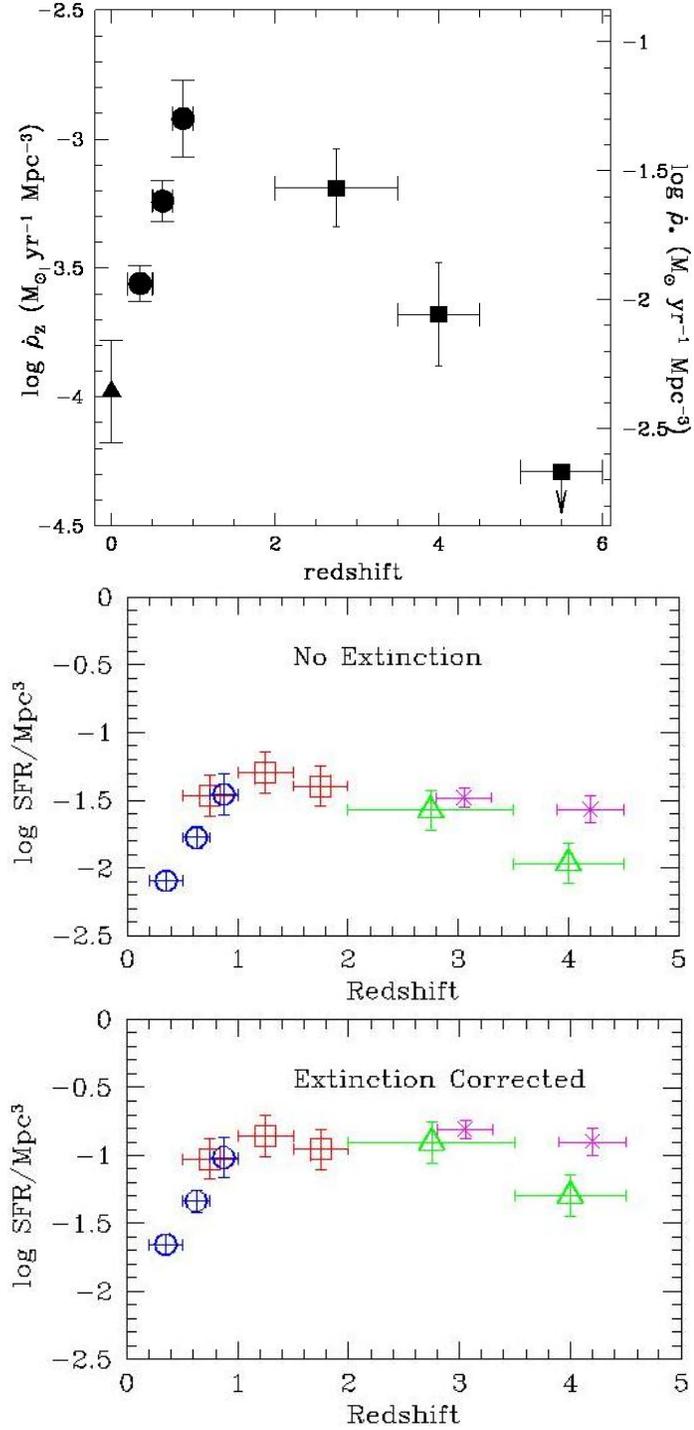
Figure 1.1: Different versions of the "Lilly-Madau" plot. Top: the original version presented in Madau (1998); Middle: a version that incorporates a much larger sample by Steidel et al. (1999); Bottom: also from Steidel et al. (1999), showing the importance of the corrections for extinction.

activity is shown to peak at $z \sim 1$–2 and to decline on both sides. The most intriguing feature of the original version of Lilly-Madau plot is the rapid increase in the GSFRD from $\langle z \rangle \simeq 4$ to $\langle z \rangle \simeq 2.75$. This result appears to be consistent with what is observed for the space density of luminous, high-redshift QSOs (e.g. Schmidt et al., 1995; Kennefick et al., 1995). However, other studies did not find evidence for such a clear change in the UV luminosity density over this redshift range (e.g., Sawicki et al., 1997).

One of the most important arguments made to reconcile the discrepancy was to incorporate the dust extinction correction to the measured SFR, because even a moderate amount of dust can effectively block the UV light emitted from newly formed stars. In an attempt to better estimate the SFR, Tresse & Maddox (1998) have calculated the extinction-corrected $H\alpha$ luminosity density at $z \sim 0.2$, and the result is consistent with that based on the UV-continuum at $z \sim 0.35$ (Lilly et al., 1996) if an extinction of 1 mag is assumed in the UV. Another study of more distant galaxies indicates that the situation might be similar at $z \sim 1$ (Glazebrook et al., 1998). Using a multi-wavelength approach including IR, sub-mm, and radio observations, Flores et al. (1999) put an upward correction for the $z \lesssim 1$ data values by an average factor of 2.9. At higher redshifts, Pettini et al. (1998) fit a typical extinction curve of the Small Magellanic Cloud (SMC) to the observed colors, and obtain a factor of 2.7 absorption at $1600\,\text{Å}$. Further improvement was made by a series of studies targeting at $z \sim 3$ galaxies, carried out by Steidel and his co-workers (see e.g., Steidel et al., 1996). They were able to obtain ~ 700 galaxies with confirmed redshifts that spans from $z = 2.5$ to 3.5. Their composite spectrum shows clear evidence of dust extinction with $E(B - V) \sim 0.3$. The correction for extinction based on a Calzetti reddening law (Calzetti, 1997), and assuming it being applicable to higher redshifts, leads to an almost flat trend of the SFR-z relation from $z \sim 4$–5 to ~ 1.5. The "Lilly-Madau plots" incorporating these new data and correction are shown in the

middle and bottom panels of Figure 1.1 (from Steidel et al., 1999).

These seminal works demonstrate that the correction of dust-obscured SFR is an important ingredient in understanding galaxy evolution at any redshifts. However, there are still large uncertainties in the measured SFRs, mainly due to the fact that the extinction correction depends sensitively on the form of the dust extinction law.

For example, Dickinson (1998) shows that the ratio of the mean corrected SFR to the uncorrected SFR for $z \sim 3$ galaxies was a factor 7 when assuming a Calzetti reddening law, but only a factor of 2 for an SMC reddening law. To address the problem, we have to directly measure the obscured star formation from dust emission in FIR/sub-mm wavelengths. However, observing at such wavelengths is challenging due to the enormous amount of technical difficulties. Nevertheless, remarkable progress has been made by a suite of facilities, such as the InfraRed Astronomy Satellite (IRAS), InfraRed Space Observatory (ISO), and Submillimetre Common-User Bolometer Array (SCUBA). The importance of dusty galaxy populations represented by Ultra Luminous Infrared Galaxies (ULIRGs) and Sub-Millimeter Galaxies (SMGs) are now widely recognized. In the next section, we will give a brief review of the development of our knowledge about the IR universe.

1.2 IR Astronomy in the pre-*Herschel* Era

Although astronomers had already known that the stars and galaxies emit light in all sorts of wavelength regimes including the infrared, direct observation was not possible until late in the 19th century, when IR detectors and efficient cooling systems can be manufactured. Most importantly, due to the fact that the atmosphere emits strong IR light, the most ideal place to conduct IR observations is the space, although ground-based observation could still be done at some "window" wavelengths.

1.2.1 IRAS and Discovery of ULIRGs

The first space-based observatory to perform a full-sky survey at infrared wavelengths was IRAS, the Infrared Astronomical Satellite (Neugebauer et al., 1984) launched in 1983 and covering wavelengths 12–100 µm. The data taken by the IRAS led to the discovery of a class of objects called ultra-luminous infrared galaxies, or ULIRGs (Houck et al., 1984, 1985; Aaronson & Olszewski, 1984). They are rare, local galaxies characterized by their exceptionally high IR luminosities ($L_{IR} > 1 \times 10^{12}\,\mathrm{L_\odot}$; integrated over restframe 8 to 1000 µm).

Follow-up studies suggested that the IR emission should be predominantly due to the re-radiation of star light processed by dust (see Lonsdale et al., 2006, for a review), and thus imply very high dust obscured star formation rates (SFR) of $> 100\text{-}1000\,\mathrm{M_\odot\,yr^{-1}}$ (using the conversion of Kennicutt, 1998). The later launched ISO satellite, with its better sensitivity, found many more IR sources at fainter flux level than IRAS probed, which implies that the universe must have many more IR luminous galaxies (Elbaz et al., 1999; Dole et al., 2001) in the past.

In the meantime, with the help of Cosmic Background Explorer (COBE) satellite, the Cosmic Infrared Background (CIB) was discovered (see Hauser & Dwek, 2001; Kashlinsky, 2005, for reviews). The intensity of the CIB is surprisingly high: it is comparable to (or maybe even exceed) the total optical light integrated to the limit as faint as could be probed by the HDF (Hauser et al., 1998). This results in a serious discrepancy when compared to the cosmic history of star formation using UV light as the SFR tracer (Madau et al., 1996). It requires at least a factor of two more star formation than was apparent in optical and UV surveys, highlighting the need of understand the star forming activity enshrouded by dust at $z \gtrsim 1$.

1.2.2 SCUBA and SMGs

A major advancement happened in the late 90s when a number of new instruments were commissioned, most significantly the SCUBA (Holland et al., 1999) and the Max-Plank Millimeter Bolometer Array (MAMBO; Kreysa et al., 1998), which led to the discovery of a large population of sub-mm bright ($S_{850} > 10\,\mathrm{mJy}$), optically faint sources (SMGs; Scott et al., 2002). These sources could directly account for around 50% of the CIB, and to near 100% with reasonable extrapolation (Barger et al., 1999; Blain et al., 1999). Also, the photometric redshift estimates from Bertoldi et al. (2000) and others (Webb et al., 2003; Fox et al., 2002; Borys et al., 2004) suggest that all the SCUBA sources should be at $z > 1$, and most at $z > 2$, implying that the SMGs are likely to be extreme ULIRGs at high redshifts.

The remarkable convergence of the studies on ULIRGs, SMGs, and CIB, and the underestimated GSFRD based on UV-optical, strongly suggests that the excessive dust emission in ULIGR/SMG is crucial in the galaxy assembly history, and that they must comprise a significant fraction of cosmic star formation rate density at $z \gtrsim 1$.

However, caution must be taken when attributing the bulk of IR emission of the ULIRGs to star forming activities. It has been noticed ever since the discovery of ULIRGs that a significant fraction of them, especially those very luminous ones, have optical signatures indicative of classic active galactic nuclei (AGN) activities (Carter, 1984; Sanders et al., 1988; Lawrence et al., 1999), which has led to a possible evolutionary scenario that ULIRGs are the prelude to quasars. Therefore, it is possible that AGN contribute non-negligible IR emission from ULIRGs. On the other hand, it is shown that large cold disks can be confused with compact warmer starbursts of very different bolometric luminosity (e.g. Kaviani et al., 2003). Due to the lack of direct redshift measurement, a much lower luminosity foreground cold disk or even a Galactic dust cloud could easily mimic a high-redshift ULIRGs (Lawrence, 2001).

Therefore, sub-mm surveys alone are not in the position of solving all the problems related to high-z ULIRGs.

The poor angular resolution at FIR/sub-mm of SCUBA and the optical faintness of most ULIRGs/SMGs add more complexity to the problem, as both factors make it extremely difficult to reliably identify optical counterparts for follow-up studies (see e.g., Dunlop et al., 2004). Nevertheless, a number of optical to NIR studies, with the help of high resolution radio interferometry and the FIR-radio relation, do show that the (radio-selected) SMGs are indeed high-z ULIRGs and both star formation and AGN could contribute to the IR luminosities (Chapman et al., 2005; Simpson et al., 2004).

1.2.3 *Spitzer Space Telescope*

Spitzer Space Telescope made further progress in this regard. The *Spitzer* is launched in 2003 and originally observed in the 3.5–160 μm wavelength range. The telescope comes with two imaging instruments, the Infrared Array Camera (IRAC) that works in 4 bands (channel 1–4): 3.6, 4.5, 5.8, and 8 μm, and the Multiband Imaging Photometer (MIPS) that works at 3 bands: 24, 70, and 160 μm. The coolant was exhausted in 2009 and since then *Spitzer* has operated in a warm-mode, only observing in wavelengths 3.6 and 4.5 μm with the IRAC as these wavelengths are less affected by the thermal emission of the telescope.

Results from IRAC and MIPS 24 μm imaging confirmed that SMGs have a wide range of observed mid-IR colors, and in most cases, the rest-frame mid-IR SEDs can be fitted with a starburst SED template, with the remainder being well fitted by a power-law AGN-dominated template. Photometric redshifts estimates constrained the redshift of these object to be $1 < z < 3.5$ (Egami et al., 2004; Frayer et al., 2004). Daddi et al. (2005a) reported the first *Spitzer* results from the Great Observatories Origins Deep Survey (GOODS) for $z \sim 2$ ULIRGs, suggesting that typical massive

($M^* \sim 10^{11}\,\mathrm{M_\odot}$) star forming galaxies at this redshift are ULIRGs, and that the number density of ULIRGs at $z = 2$ could be three orders of magnitude larger than the local ones. However, caution again should be taken when relating such ULIRGs to star formation activities at high-z. For example, Franceschini et al. (2005) and others (e.g., Martínez-Sansigre et al., 2005; Donley et al., 2005; Polletta et al., 2006) have discovered a population of heavily obscured AGN, some of which are Hyper-LIRGs ($L_{IR} > 10^{13}\,\mathrm{L_\odot}$) at $z > 2$, with AGN torus-dominated mid-IR SEDs (Polletta et al., 2006).

The ambiguity in the interpretation is at least partly due to the fact that the instruments have missed the rest frame far-infrared peak ($\sim 100\,\mathrm{\mu m}$) that dominates the spectral energy distributions (SED) of most IR galaxies. Extrapolation via either empirical or theoretical IR emission SED (e.g., Chary & Elbaz, 2001; Dale & Helou, 2002) is then needed to calculate the total energy output, which inevitably result in large uncertainties. Extrapolation from mid-IR is particularly dangerous, as AGN heating could dominate in the mid-IR regime. The issue is solved with the later launched telescope.

1.3   IR Universe as Seen by *Herschel*

In order to truly understand the nature of high-z ULIRGs and their role in galaxy evolution, the ultimate solution is to directly aim at the full FIR/sub-mm wavelength range to include the peak of IR emission (at around restframe $\sim 100\,\mathrm{\mu m}$), using instruments of much improved sensitivities and spatial resolution. The *Herschel* Space Telescope was the best that we have to date, and its rich data sets indeed enable us to study the galaxy formation and evolution through the IR universe to an unprecedented scale.

1.3.1 *Herschel Space Telescope*

Herschel, built and operated by the European Space Agency (ESA), was the largest infrared telescope ever launched. It has a single 3.5 m primary mirror. Two of its three instruments, namely, the Photodetector Array Camera and Spectrometer (PACS; Poglitsch et al., 2010) and the Spectral and Photometric Imaging REceiver (SPIRE; Griffin et al., 2010) were imaging survey instruments, and the other, Heterodyne Instrument for the Far-Infrared (HIFI; de Graauw et al., 2010), was a high resolution spectrometer.

HIFI HIFI covers the wavelength range of 240–625 µm and 157–213 µm with seven receivers at very high resolution, with an on-sky resolution of 13–40 arcsec (comparable with the SPIRE instrument's point source function's (PSF's) full-width half maximums (FWHMs)), and a spectral resolution of 0.02–0.7 km s^{-1}. The signal is processed with heterodyning, mixing the incoming signal with another produced by the telescope to create a signal at a lower and therefore more manageable frequency.

PACS The imaging unit of PACS covered a wavelength range of 60–160 µm in three bands (although observing only two at once, 160 µm with 70 µm or 100 m). Pixel sizes are 3.2, 3.2 and 6.4 arcsec with FWHM of 5.6, 6.8 and 11.3 arcsec, respectively. Comparing to *Spitzer*, the MIPS instrument has a FWHM of 18.6 arcsec at 70 µm, meaning PACS has a three-times improvement on the angular resolution of the sky at these wavelengths and less likely to suffer from source confusion.

SPIRE The SPIRE instrument is capable of taking images at three bands: 250, 350 and 500 µm. Standard image pixel sizes from the photometer are of 6, 10 and 14 arcsec with PSF FWHMs of 18.15, 25.15 and 36.3 arcsec. The FWHMs are

broader than those of SCUBA-2, although this is due to JCMT's larger mirror that was unfeasible for a space mission.

As shown in Figure 1.2, the photometry bands for PACS and SPIRE sample the peak of heated dust emission from $z = 0$ to 6. This, for the first time, has offered the capability to directly measure the total IR luminosity.
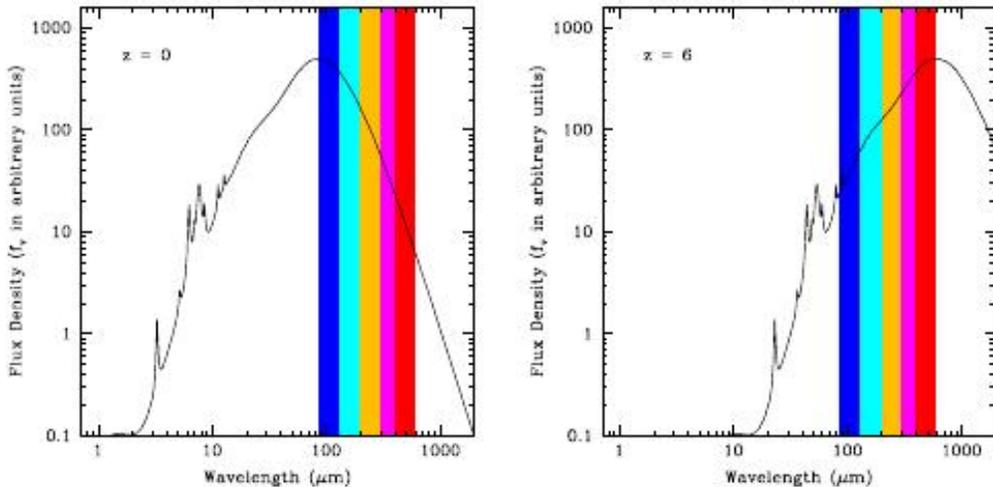
Figure 1.2: The five PACS + SPIRE FIR bands sample the broad peak of heated dust emission from $z = 0$ (left) to $z \gtrsim 6$ (right). The superposed spectra are based on an ULIRG model from Siebenmorgen & Krügel (SK07; 2007). The color stripes indicate the five bands: 100 (blue), 160 (cyan), 250 (brown), 350 (magenta) and 500 μm (red).

During its four years of operation, *Herschel* has produced a rich data set. Through the large extragalactic surveys, including the *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al., 2010), the *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al., 2012), the *Herschel* Stripe 82 Survey (HerS; Viero et al., 2014), the PACS Evolutionary Probe (PEP; Lutz et al., 2011), etc., *Herschel* have mapped the FIR/sub-mm universe in unprecedented scale and details (See § 2.1). In Figure 1.3, it shows an overview of the *Herschel* data in terms of the survey depth and area. A slew of new discoveries has been made using the *Herschel* data.
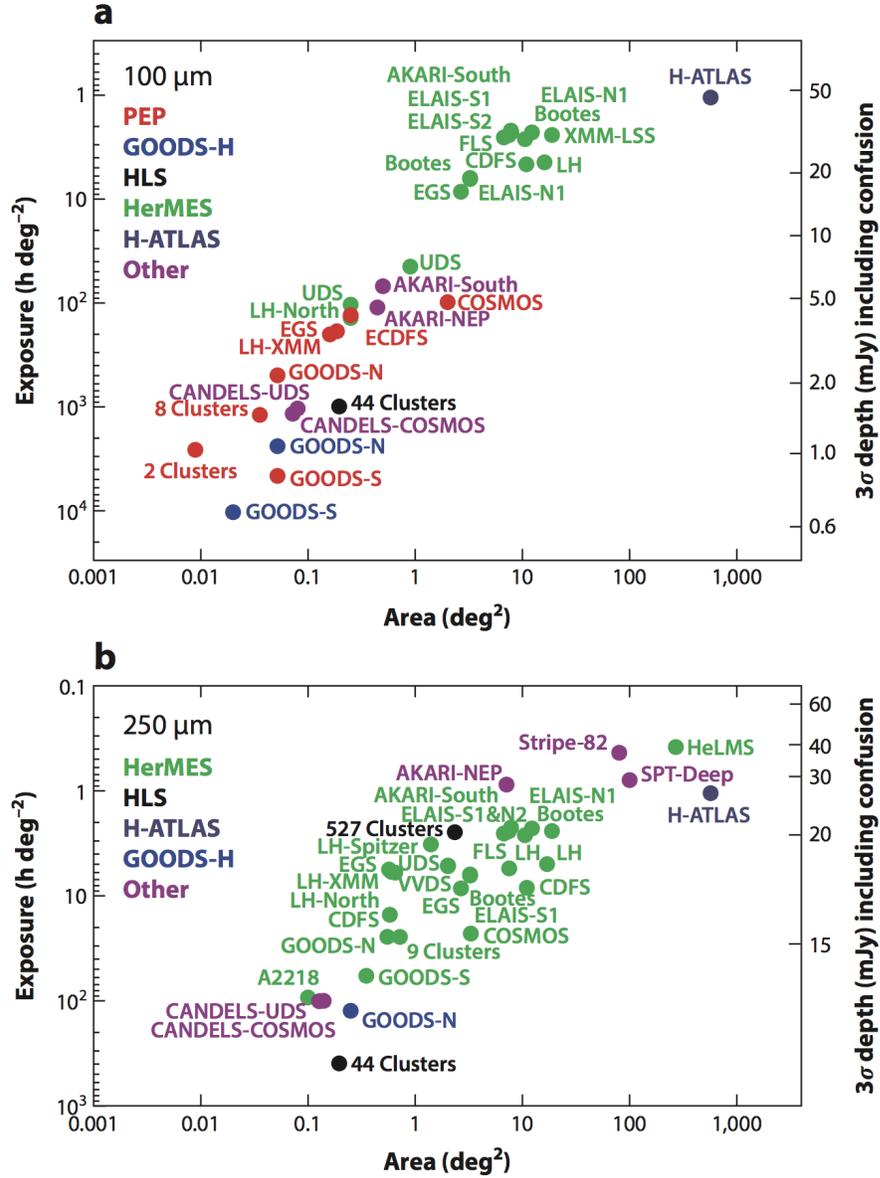
Figure 1.3: Figure 1 of Lutz (2014): Survey area and point source depth reached by some extragalactic surveys with (a) PACS and (b) SPIRE. Exposure is computed from total observing time (including overheads) and survey area. The point source depths (including confusion noise) shown on the right axis should hence be indicative only, because no attempt was made to capture detailed effects of different observing layouts in the various projects.

1.3.2 Highlights of *Herschel* Results

- A resolved view of the CIB. *Herschel* data helped in understanding of the CIB mentioned in § 1.2.1 and the "missing SFR" problem. In Figure 1.4, it summarizes the 70–850 μm CIB as directly measured by COBE and the contributions of galaxies as determined by *Spitzer*, *Herschel*, and ground-based sub-millimeter surveys. The extrapolation of measured number counts is consistent with the direct CIB measurements and provides smaller uncertainties at some wavelengths near the peak of the CIB. Combining the results, the total 8–1000 μm background is 27^{+7}_{-3} nW/m^2/sr (Béthermin et al., 2012). The CIB fraction accounted for by individually resolved galaxies peaks at 75% at 100 and 160 μm (Magnelli et al., 2013). The resolved fraction is lower at 70 μm, where *Herschel* could not integrate to its confusion limit (Berta et al., 2011). At 250–500 μm, 15% to 6% of the CIB is resolved at these wavelengths due to increasing source confusion (Béthermin et al., 2012). However, using stacking analysis, Béthermin et al. (2012) were able to attribute 73% to 59% of the CIB to known galaxies in this wavelength regime. As expected, the typical redshift of sources contributing to the CIB increases with wavelength (Figure 1.4, lower). Of the major CIB fraction that is directly resolved or retrieved by stacking, half of the emission at wavelengths 100, 160, 250, 350, and 500 μm is contributed by sources above $z = 0.75$, 1.00, 1.04, 1.20, and 1.25, respectively (Béthermin et al., 2012; Magnelli et al., 2013).

- Evolution of the infrared luminosity function (IRLF) and the cosmic star formation rate density (SFRD). Evolution of the infrared luminosity function and the infrared luminosity density encode the cosmic history of dusty star formation and are an essential part of measuring the total cosmic star formation history. With the *Herschel* data, we could derive a much more accurate version of the
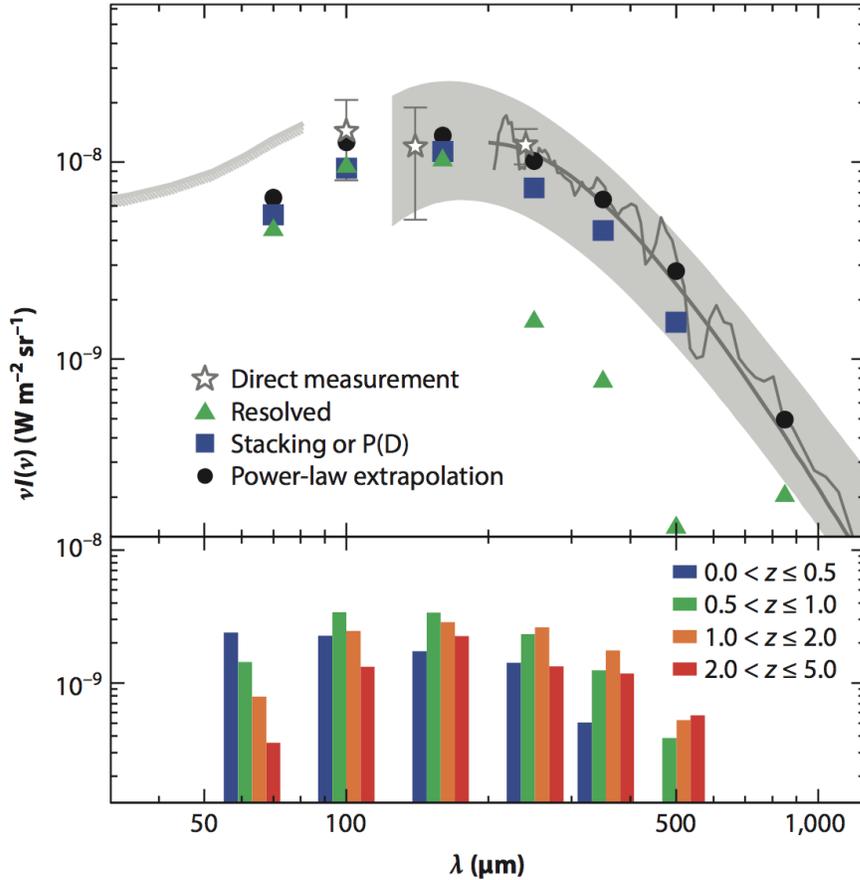
Figure 1.4: Figure 4 of Lutz (2014): (Upper) The cosmic infrared background (CIB) as seen by direct measurements and as resolved by *Herschel*. Direct measurements are shown in gray, including the results from the COBE Diffuse Infrared Background Experiment (DIRBE) (Dole et al., 2006), and the COBE Far Infrared Absolute Spectrophotometer (FIRAS) (Lagache et al., 1999; Fixsen et al., 1998). CIB contributions by resolved sources (green) are from Berta et al. (2011) and Magnelli et al. (2013) (70–160 μm), (Béthermin et al., 2012) (200–500 μm), and Zemcov et al. (2010) (850 μm). Stacking results (blue) are from Béthermin et al. (2010) (70 μm), Berta et al. (2011) (100 and 160 μm), and Béthermin et al. (2012) (250–500 μm). (Lower) Contributions of different redshift slices to the part of the CIB that is contained in resolved sources (PACS) and covered by stacking (SPIRE) (Berta et al., 2011; Béthermin et al., 2012; Magnelli et al., 2013).

Lilly-Madau diagram, taking into account the dust obscured SFR, as shown in Figure 1.5. As it illustrates, most of the energy from star-forming galaxies at $0 < z < 2$ is absorbed and re-emitted by the dust in IR, and only a small fraction is un-obscured UV light. The gap between the UV and IR measurements

increases with redshift out to $z \sim 2$ and then may start falling off. From the IR luminosity function (Figure 1.6), we could find that the contribution of LIRG and ULIRG luminosity bins to the luminosity density increases from 5% and 0.4% in the local Universe to 31% and 50%, respectively, at $z \sim 2$ (Magnelli et al., 2013).
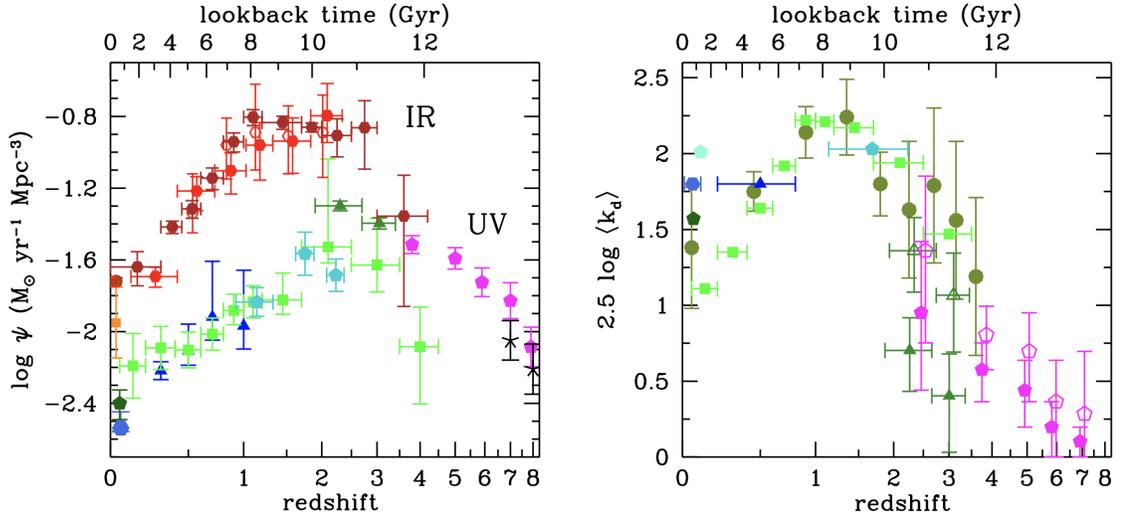
Figure 1.5: Figure 8 of Madau & Dickinson (2014): (Left panel) SFR densities in the FUV (uncorrected for dust attenuation) and in the FIR with direct measurement from *Herschel*. A Salpeter IMF is assumed in converting the luminosities to the SFRs. (Right panel) Mean dust attenuation in magnitudes as a function of redshift. Most of the data points shown are based on ultraviolet spectral slopes or stellar population model fitting.

1.3.3   Synergy with ALMA and VLA

Despite the great success of *Herschel*, the still large beam size (18 arcsec in 250 μm) gives rise to severe source confusion problem and imposes great difficulties in identifying the counterparts of the IR sources in the optical to near-IR (with typical FWHM of $< 1''$) (see § 1.4 for a review).

Fortunately, the recently commissioned interferometric facilities, the radio Jansky Very Large Array (JVLA) and the Atacama Large Millimeter/sub-millimeter Array
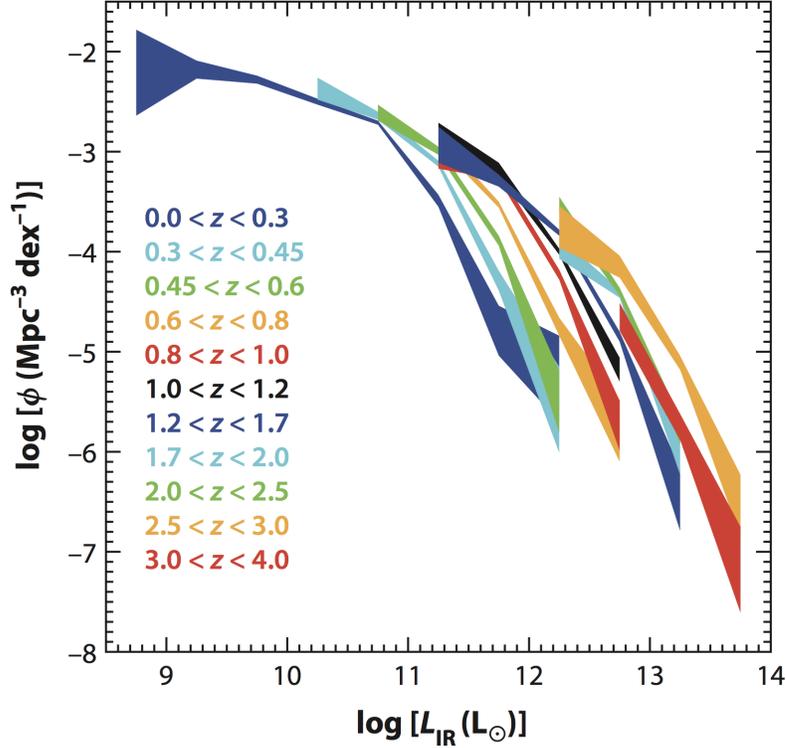
Figure 1.6: Evolution of the total infrared (8–1000 μm) luminosity function, from the work of Gruppioni et al. (2013). The width of the band indicates the 1σ poissonian error.

(ALMA), are able to deliver sub-arcsec spatial resolution in sub-mm/mm wavelength, which could help pin-point the location of *Herschel* sources as well as provide morphological information. Some of the recent work are Hodge et al. (2013) for the ALESS project, Mullaney et al. (2015) for *Herschel* selected AGNs, Bussmann et al. (2015) for lensed galaxies in the HerMES survey.

ALMA and JVLA are great in studying the details of a small sample of selected sources. However, due to the limitation of the interferometry, surveys that covers a large region of the sky are not feasible. On the other hands, a great area of the sky has been scanned and mapped by *Herschel*, and we do need the ancillary data to remedy the counterpart identification problem before we can fully unleash its great power. We *need* more affordable way of dealing with the source blending issue, and

in the next section, a review is given on this matter.

1.4   The Haunting Problem of Source Blending in FIR/sub-mm/mm Observations

Due to the diffraction limit, the FWHM of point sources detected in the IR are in general much larger than those shown up in the optical bands. Sources that are close-by are smeared and show up in the image as a single blob of light, which is known as the source blending or source confusion problem.

The source confusion results in the confusion noise, which is attributed to the uncertainties from the un-resolved sources in the beam of a point source. Measuring accurate flux densities for sources in astronomical images dominated by great confusion noise is the greatest obstacle to scientific analysis of data at far-IR regime, which is where *Herschel* operates. Deep exposures no longer increase the sensitivity due to the confusion, because of the blended signal from numerous, unresolved, faint sources forming a non-negligible fraction of the observed telescope background.

On the other hand, the large beam size imposes difficulties in interpreting the observed fluxes. It is often the case that within the beam size of one single far-IR source, there are tens, if not hundreds, of sources showing up in the deep optical image of the same patch of the sky. It is then critical to correctly identify those sources that contribute to the observed IR light (referred to as source de-blending), so that we could use the optical to near-IR spectral energy distribution (SED) or spectrum to obtain the redshift, which is essential to model the dust emission in the far-IR and measure the dust obscured SFR. Also, we could use the SED fitting technique to infer the properties of the galaxy such as stellar mass and age. Moreover, with more powerful imaging unit like *HST* we can study the morphology. Combining all these information is crucial to a coherent picture of the galaxy formation and evolution

history.

Over the decades, along with the advance of the IR astronomy, a suite of treatments and workarounds have been developed to deal with this haunting problem.

1.4.1   The Explored: Blind Source Extraction

Source extraction refers to the step of constructing a source catalog with the fluxes measured from the observed imaging data.

The IRAS point source catalogue was constructed by running a square wave filter on the data and the detection was done by looking for the signal peaks on the filtered map. To get the flux, each peak was compared to an idealized peak expected from a noiseless source (Beichmann, 1985). The IRAS Faint Source Survey was able to go deeper than the previous point-source catalogues, by using a filter that matches with the shape of the point source (Moshir et al., 1992).

At the time of the later IR missions including ISO and *Spitzer*, the SExtractor (Bertin & Arnouts, 1996) became the most widely used tool and was used to create catalogs for those data. In brief, SExtractor is best suited for performs blind source detection and measures the fluxes on large optical surveys. It first uses a thresholding algorithm on the input image (with optional filtering) to construct a segmentation map. Then each segment is examined by a de-blending algorithm, which basically searches fluctuations across the area of pixels of that segment. Depending on the configuration, the segment could be split to multiple smaller components. The photometry can be estimated in a number of ways. In the case of ISO and *Spitzer* IRAC catalog, the aperture photometry was used, which is done by adding fluxes assigned to pixels within a defined radius around the center of the source.

Later on, the *Spitzer* team developed APEX, a source extracting software that makes use of the PSF fitting technique. It is the recommended tool to create catalogs for the *Spitzer* images, especially for the MIPS bands. The source extraction is done

in two steps as usual, the detection and the extraction. In the detection step, a source list is created by a thresholding algorithm, and in the second step, a fitting is done by using the point response function (PRF) of the map. A critical step in controlling the reliability of the extracted flux is the estimates of the map background and the noise. APEX provided a number of options to fine-tune the process according to different situations.

In summary, these tools, despite their varying implementation, have been successful in delivering accurate catalogs based on the science map. However, they all suffer the same issue, especially for the images with large beam size (approximately the FWHM of the PSF or PRF), that is, the detection and the flux measurement are based on the same map. Therefore, objects that are smeared by or blended with nearby companions are often misidentified as one single source, results in systematically higher fluxes and a deficit of faint objects on the source count.

One approach to solve this problem is to make use of ancillary data to provide additional information (priors).

1.4.2   The Underexplored: Source Extraction with Prior

Entering to the realm of *Herschel* data, the fact that they are taken at much longer wavelength regimes makes the beam size much larger than the *Spitzer*. Especially, we are most interested in the cold dust emission that rises from the star forming activity, and the bulk of the IR flux would peak at around restframe $100\,\mu$m. This means for redshift $z \sim 2$ sources, at which time the dust obscured SFR is considered to be dominating the total cosmic SFRD, the peak falls to the wavelength range of the three SPIRE bands, 250, 350 and $500\,\mu$m. In these bands, unfortunately, the typical beam sizes are 18.15, 25.15 and 36.3 arcsec, which are much larger than the *Spitzer*.

This places great obstacle to the interpretation of the observed IR sources on those maps. There are in general two directions to go to deal with the problem:

(1) **Nominal blind extraction** is done regardless of the underlying truth on the 250 μm map, and the source list is used to extract fluxes from the other two maps with longer wavelengths. This approach is essentially a compromise that it generates fluxes with guaranteed consistency among the SPIRE bands, while deferring the treatment of the source multiplicity issue to the next level analysis (cross matching or interpretation). This is the original and *de facto* treatment adopted in the literatures. To make use of the catalog generated this way for studying the SEDs of the galaxies, one has to pay close attention when cross matching with other catalogs created from maps of higher spatial resolution, such as in optical or near-infrared. Advanced matching techniques have been developed in order to provide robust source identification (ID). For example, Smith et al. (2011) used likelihood ratio (LR) technique to match the SDSS catalog with the H-ATLAS (for description of the project, see § 2.1.2) science demonstration phase (SDP) data. It is essentially a statistical method that could assign a probability ("reliability") to each potential match, and thus distinguish robust counterparts from chance alignments with background sources. The LR method relies on knowledge of the intrinsic positional uncertainty of the sources as well as the magnitude distributions of true counterparts and background sources. The advantage of this method is that it is effective in create a *subsample* of sources with robust ID, but the disadvantage is that we have no mean to decompose the fluxes of the sources that are flagged with low reliability, i.e., the ones that are likely to have multiple counterparts in the high resolution maps.

(2) **Extraction with Prior**. To create a catalog that both have correct *Herschel* fluxes assigned and have as many as possible sources included, one has to rely on sources extraction methods that make use of additional information (prior). The *Spitzer* MIPS 24 μm positions are the most commonly used prior in the literatures for extracting *Herschel* fluxes, because it both has OK angular resolution (6 arcsec) and is the closest to the *Herschel* wavelengths.

As an example, Magnelli et al. (2013) used the following source extraction scheme for the PEP/GOODS-*Herschel* program: They started from the positions of the *Spitzer* IRAC 3.6 μm sources (FWHM $\sim 1.6''$) from the GOODS *Spitzer* Legacy Program (PI: M. Dickinson), and use them to extract sources in the MIPS 24 μm maps (Magnelli et al., 2009). Then, the 24 μm detected sources are used as priors to extract the sources in the PACS maps. The main advantage of this approach is that it could deals with the blending issues encountered in dense fields, and provides a straightforward association between IRAC, MIPS and PACS sources. This approach requires the assumption that all sources present in the PACS images have already been detected at 24 μm, which was tested to be largely true, with a missing rate of a few percent.

Using of the MIPS 24 μm data as the prior, despite its effectiveness, can only be done in the fields that had observed by the *Spitzer* during its cryogenic phase. Ever since 2009, we have no longer the instrument that could probe similar wavelength range. This is a severe limitation, as the area of the fields that have MIPS 24 μm coverage is only a small fraction. In the field without the MIPS 24 μm data, we have to rely on alternative methods or strategies.

As mentioned in the previous section, the new generation interferometry instruments including ALMA and JVLA are able to observe in the sub-mm to mm regime to an angular resolution of sub-arcsec level. And most importantly, the wavelength range is close enough to that of the *Herschel*. This renders the ALMA/VLA the best "source of truth" for pin-pointing the counterparts of the IR sources found on the *Herschel* maps. As an example, Hodge et al. (2013) conducted an ALMA survey of sub-millimeter galaxies in the Extended Chandra Deep Field South (ECDFS) field, one of the field mapped by the HerMES project. As expected, the high resolution of the ALMA allowed them resolve sources that were previously blended and accurately identify the origin of the far-IR/sub-mm light. The observation shows that at least

35% (possibly up to 50%) of the detected sources have been resolved into multiple sources, and the multiplicity increases as the increase of the total, blended flux. The remarkable result shows that indeed we are missing a great fraction of true IR sources if we only study the non-blended cases. Also, if we are interested in constructing the IR luminosity function, we must take into account the effect of the source multiplicity.

The limitation of using the interferometry data as the prior is obvious: it is super expensive to conduct large scale surveys using interferometry. For example, HeLMS is the largest continuous field ever observed by *Herschel*. It has an area of $\sim 270\,\mathrm{deg}^2$. It is by no means possible that all the *Herschel* sources in this region could be surveyed using either of the ALMA or JVLA. For the wide-field *Herschel* data, we have to seek alternative methods that (1) does not rely on the MIPS 24 μm as the bridge for counterpart ID; (2) less costly so that it is feasible for large datasets.

In our previous work, Yan et al. (2014), we proposed a new technique that meets the aforementioned requirement, that is to use directly the optical/near-IR positions as the prior. As presented in Yan et al. (2014), the algorithm worked well for the sample of selected optical faint, far-infrared bright *Herschel* sources in the CANDELS field. However, at that time, the algorithm was only in its proto-type stage, and was executed manually.

The key contribution of this thesis is that we have fully developed and automated the algorithm, as well as implemented it as a software tool named `CIDer`. Using our major contributor identification tool, we are able to carry out a comprehensive study of the infrared luminous sources found in the wide *Herschel* survey fields, by constructing the largest, reliable IR source catalogs (high-z ULIRGs) with panchromatic photometry from optical to far-infrared.

1.5 Thesis Overview

Given the context described in the previous sections, the scope of this thesis is summarized as follows:

This thesis project studies the formation and evolution of extremely dusty galaxies at $z \sim 1$–6, taking advantage of the unprecedented far infrared (FIR) data sets from the *Herschel* wide field surveys. We aim at compiling the largest high-z ULIRG sample, based on accurate measurement of the total IR luminosities using FIR bands, with the help of our own major contributor identification algorithm.

The *Herschel* Space Observatory generated a rich FIR/sub-mm data set of unprecedented sensitivity and spatial coverage during its lifetime, and for the first time in history, it allows us to explore FIR/sub-mm universe at $z > 1$ on a large scale that is comparable to those in optical/near-IR. In the meantime, owe to the great synergy with interferometric facilities like ALMA and JVLA, we are in the great position to explore the far-IR/sub-mm universe, which will help fully address the open questions about the nature of ULIRGs and their role in galaxy assembly history, and to lead to transformational discoveries on the dusty side of galaxy formation.

The major challenge of the project is the persisting problem of poor angular resolution in far-IR/sub-mm that the *Herschel* still suffered. Recent studies have shown that a large fraction of seemingly single FIR/sub-mm sources are actually made of multiple objects that may or may not be associated. Therefore, to correctly construct the ULIRG sample, one should take this multiplicity into account. In Yan et al. (2014) paper, we proposed a de-blending technique for high-S/N *Herschel* sources using optical data as the position priors, and this technique is now fully developed and automated (as the core content of this thesis). This approach is particularly important for the fields where no higher-resolution mid-IR data are available as the proxies, and will provide a much less expensive alternative to doing interferometry at sub-mm/mm, where a survey machine does not yet exist. Combining it with the

large pool of ancillary optical/near-IR data, the proposed project will maximize the scientific return of the precious *Herschel* wide field survey data.

Another challenge, which is of equal importance, is that we need to have the multi-wavelength data to begin with, among which those from deep optical imaging surveys are the key component. Not only are the optical (to near-IR) positions the inputs to the counterpart identification of the *Herschel* sources, but also we have to rely on the photometric redshifts to model the far-IR fluxes and derive the SFR, which can only be obtained from fitting accurate, well sampled optical to near-IR SEDs (see § 5). For the well-studied, legacy extragalactic fields, such data already exists (e.g., the CANDELS project, see § 2.2.1), but for a large fraction of the wide fields (e.g., those HerMES fields that are dubbed as level > 5, see § 2.2.2), either we do not have the observations/data, or the data are in their raw form which are needed to be compiled and reduced in a homogeneous way so that accurate multi-wavelength SEDs can be constructed. Over the course of the project, we managed the task by (1) conducting our own deep optical imaging observation, the Mizzou WIYN survey (MizzouWINS), to complement the existing data; (2) developing an efficient pipeline building and management tool, APUS, and creating our own data reduction pipelines based on it, highly optimized for putting together data that were taken over long period of time and at varying conditions. Using the data both from the data archives and from our own MizzouWINS, and the data reduction tools that we created, we are able to produce high-quality mosaics and catalogs in the optical bands for the *Herschel* wide fields, which lays down a solid foundation for the science that we derive from the study.

To summarize, we aim at constructing the largest, well-defined sample of high-z ULIRGs, whose full wavelength range (optical-to-FIR) SEDs are accurately derived. The entire project is divided to two stages. The first stage is to compile and reduce the multi-wavelength data, and the second stage is to carry out the source de-blending

and construct SEDs that range from optical to FIR. The main science goals include: (1) a comprehensive and robust census of ULIRG at $z \simeq$ 1-3; (2) study of ULIRGs at $z \gtrsim$ 4; (3) ULIRG luminosity function and the contribution to the global star formation rate density (SFRD); (4) applying to new ULIRG diagnostics.

The main body of the thesis is to present the methods/tools that we developed to carry out the study, as well as the results that we get for the work in the First Look Survey (FLS) field, one of the wide fields surveyed by the HerMES project. Also, the FLS field is the best case that we currently have to demonstrate how the new data we obtained at WIYN played a critical role in supplementing the existing, archival data.

In Chapter 2, we describe, both in general and for the FLS field as a case study, the multi-wavelength datasets that are available to date. In Chapter 3, we describe the software package APUS that we developed to aid the optical data reduction tasks. The pipelines that we created for reducing the optical data, including PostCalib, the one for the MizzouWIYNS data, in the FLS field are also presented in this chapter. In Chapter 4, we present the major contributor identification algorithm as described in Yan et al. (2014), and describe the software implementation CIDer. In Chapter 5, we introduce the SED fitting and modeling techniques that we use to analysis the constructed catalog. In Chapter 6 we show and discuss the results we get for the SFR in FLS field. Also included as part of this thesis is our study on the SDSS quasars detected by *Herschel*, which is given in Chapter 7. The results have played an important role in directing our research over the years. Finally, we summarize the main results in Chapter 8.

Chapter 2

Data

In this chapter, we summarize the data we used in this thesis project. We also describe the data reduction procedures used to create the *Spitzer* IRAC and MIPS maps and catalogs.

For the multi-wavelength optical data, we developed a pipeline building and management tool APUS, and created our own pipelines to perform the imaging data reduction. The details will be presented in the next chapter (Chapter 3).

2.1 Wide-field *Herschel* Surveys

During its 3-year lifetime, *Herschel* carried out a total of 19 776 hours of science observations, among which 32 % are guaranteed time (GT), and the rest are open time (OT). Per the science goals stated by the *Herschel*, A number of "Key Programs" (KP) in the form of large spatial and spectral surveys are executed and constituted a significant fraction ($\sim 50\%$) of the available time of the overall mission (Pilbratt et al., 2010).

The Herschel Multi-tiered Extragalactic Survey (HerMES, Oliver et al., 2012), is

the largest GT KP that provides a legacy survey of star-forming galaxies over the wavelengths at which the cold dust emission peaks; and the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS, Eales et al., 2010) is the largest OT KP, which aims to map $550 \deg^2$ at five wavelengths from 110–$500\,\mu m$ in order to constraint on the star formation rates of many thousands of quasars through far-infrared and sub-mm photometry.

2.1.1   HerMES

HerMES survey was initially designed to take ~ 900 hours and cover an area of $100\deg^2$ with the SPIRE and PACS instruments. In 2011, an additional $270\deg^2$, HerMES Large-Mode Survey (HeLMS) field was added to the program for studying large-scale structure and the bright end of the number counts. The broad science goal of the survey was to provide insight into galaxy evolution using far-infrared wavelengths.

The survey fields are chosen to vary in sizes and depths, which is referred to as the "wedding cake" design, as shown in Figure 2.1 and Figure 2.2. These fields were selected to sample a broad range of redshifts and luminosities of galaxies. The smaller fields are observed to a greater depth to find fainter galaxies, and the wider, shallower fields aim at finding the brighter, rarer objects. In general, the survey was designed to find 75 galaxies in a luminosity-redshift bin with $\Delta \log L/L_\odot = 0.5$ and $\Delta z = 0.2$, for redshifts range $0 \leq z \leq 3$. At $250\,\mu m$, the luminosity depths are approximated be $2 \times 10^{12}\,L_\odot$ at $z = 2$ for deep fields like COSMOS, and $1 \times 10^{13}\,L_\odot$ in the shallowest HeLMS. Most of these fields are recognized as the "legacy fields" due to the availability of rich multi-wavelength data obtained from other instruments spanning from optical to near-to-mid-IR, to enable detailed, comprehensive studies.

The data from the *Herschel* telescope were processed by the HerMES team using their data reduction pipelines. The high-level data products including maps and
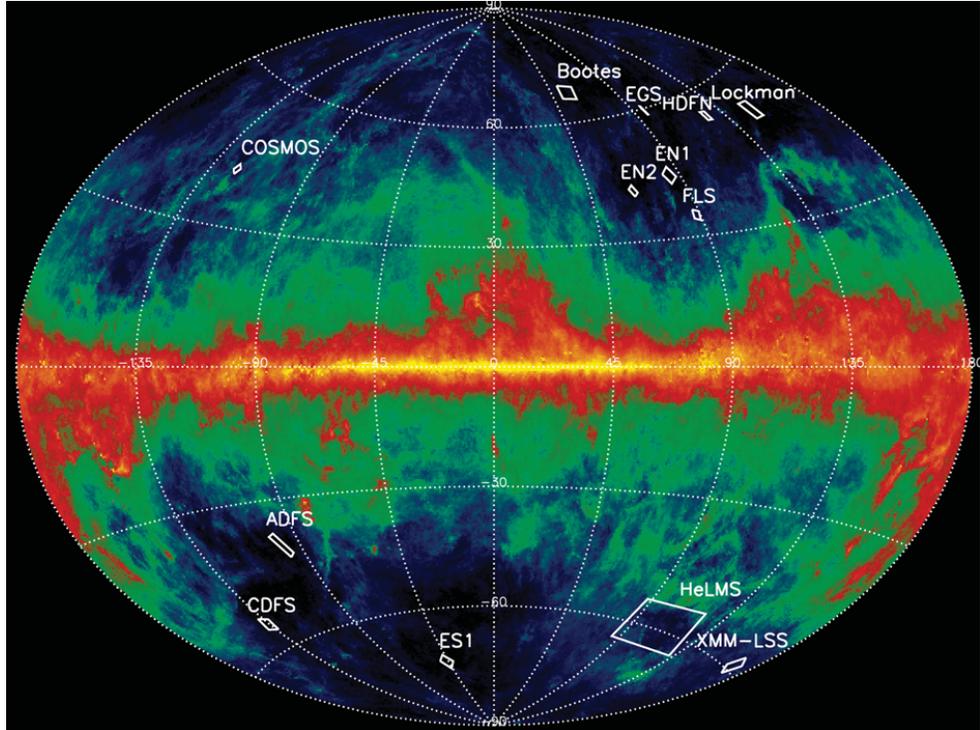
Figure 2.1: Map of dust emission from the Galaxy, with HerMES fields over-plotted as presented in Figure 4 of Oliver et al. (2012). The image is the 100 μm, COBE-normalized, IRAS map of extended emission (Schlegel et al., 1998), plotted in Galactic coordinates. The sky brightness is color coded with regions of very low Galactic emission appearing black and the Galactic plane yellow. In addition to the fields marked with the white boxes, HerMES has also observed 12 known clusters which are not shown.

catalogs were made public through the first data release (DR1) on April 3th, 2012, which includes *Herschel* SPIRE (250, 350, and 500 μm) sky maps and object catalogs for part of the survey fields. And later in October 2013, those for the rest of the fields excluding L7-HeLMS were released through data release 2 (DR2). In July 2015, through the support of the Herschel Extragalactic Legacy Project (HELP)[1], the HerMES third data release (DR3) was announced, which includes the SPIRE maps for the HeLMS field, as well as the cross-identification catalogs based on *Spitzer* MIPS 24 μm source positions. The fourth data release (DR4), which is also the latest

---

[1]HELP is a European Commission Research Executive Agency funded project under the SP1-Cooperation, Collaborative project, Small or medium-scale focused research project, FP7–SPACE–2013–1 scheme.
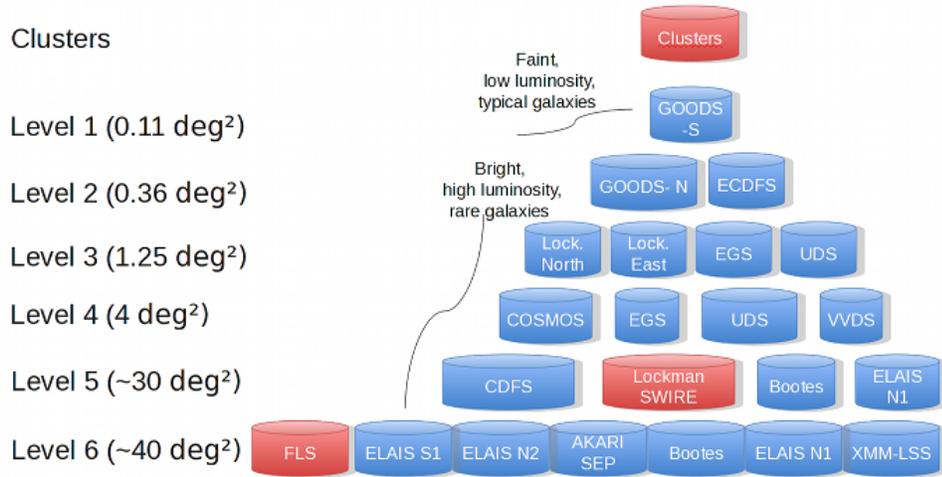
Figure 2.2: The hierarchical structure ("wedding cake") design of the HerMES survey. The clusters are observed to the most depth, and from L1 to L6, the fields go with larger areas but shallower coverages. The shallowest HeLMS field (L7) is not shown in the plot. This figure is adopted from the HerMES project site http://hedam.lam.fr/HerMES/index/survey.

data release from the official HerMES team, was made available on July 2016. The DR4 includes the SPIRE maps built from the complete set of observations made by *Herschel* for the HerMES program, and most notably, includes the SPIRE maps of the largest contiguous extragalactic field observed by *Herschel* obtained by combining the observations of HeLMS (P.I. Marco Viero), HerS (P.I. Marco Viero), and XMM-LSS (part of HerMES). The DR4 provides all-band catalogs extracted on blind $250\,\mu m$ positions for all the HerMES fields, as well as those based on *Spitzer* MIPS $24\,\mu m$ extracted by the HELP team. It is also the first release of the PACS maps for the HerMES fields. Note that DR4 is not a complete replacement of DR3, therefore DR3 and DR4 should be used together.

In this work, we make use of the maps and catalogs from DR3/DR4 and we briefly describe how these data products are produced. The details of the data reduction can be found in Oliver et al. (2012) for field definition and survey design, Smith et al. (2012), Roseboom et al. (2010), and Wang et al. (2014) for the catalogs, and Levenson et al. (2010), Viero et al. (2013) for the image maps. All the data products

are available online at HerDaM[2].

Maps The raw data from *Herschel* observations are made of multiple "scans" and are referred to as "time-ordered data" (TOD). The TOD are processed using HIPE (Herschel Interactive Processing Environment; Ott, 2010) trough a number of steps including glitch detection/repair, instrumental correction, signal linearization and flux calibration, and bolometer time response correction, etc. The processed data (referred to as level 1 or timeline data) are fed into a customized software package SMAP (Viero et al., 2013; Levenson et al., 2010) to produce combined sky maps. Each produced map (in multi-extension-fits format) contains a science image in unit of $Jy\,beam^{-1}$, a noise map in the same unit, a coverage map in unit of second. For DR4 (SMAP v6.0) maps, they contain all the SPIRE observations made for HerMES field. See Table 2.1 for an overview of all the maps available in HerMES DR3/DR4.

Catalogs For this thesis, we are interested in the distant, point sources detected by the *Herschel*. The methods and techniques used to generate the HerMES point sources catalogs are described in detail in Roseboom et al. (2010, DE-SPHOT), and the characteristics of the catalogs produced for DR1 and DR2 are described extensively in Smith et al. (2012) and Wang et al. (2014). In brief, due to the coarse spatial resolution, the source confusion is severe and it is critical to optimize the source extraction algorithm to ensure consistent fluxes across multiple wavelengths. DESPHOT makes use of simultaneous PSF fitting technique across images of different bands with a common positional prior to derive band-merged catalogs. Starting from DR3, the HELP team has been developing more advanced methods to solve the problem, and the algorithm is implemented in `XID+` (Hurley, 2017), and is used to produce the catalogs in the HerMES DR4. Two sets of positional priors are used in DR4: (1) Blind extracted $250\,\mu m$

---

[2] http://hedam.lam.fr

sources ("xID250" catalogs); (2) MIPS 24 μm sources ("MIPS24" catalogs). In this work, we mainly make use of the xID250 catalogs, and we present our own treatment (CIDer; see § 4) for the method that we use to construct the band-merged FIR catalogs. See Table 2.1 for an overview of all the catalogs available in HerMES DR3/DR4.

Table 2.1: Summary of HerMES DR3/DR4 data products (excluding the clusters)

Level[a]	Name	Area (deg^2)	xID250	MIPS24
L1/L1/L5	GOODS-S/ECDFS/ CDFS-SWIRE	0.35/0.58/11.39	Full	Partial
L2/L5	COSMOS/COSMOS-HerMES	2.82/4.38	Full	Partial
L2	GOODS-North	0.55	Full	Full
L3/L5	EGS/EGS-HerMES	0.60/2.67	Full	Partial
L3/L3/L5	Lockman-North/ Lockman-East-ROSAT/ Lockman-SWIRE	0.65/1.40/17.37	Full	Partial
L4/L4/L6	UDS/VVDS/XMM-LSS-SWIRE	2.02/2.02/18.87	Full	Partial
L5/L6	Boötes	3.25/10.57	Full	Full
L5/L6	ELAIS-N1	3.25/12.28	Full	Full
L6	ELAIS-S1-SWIRE	7.86	Full	Full
L6	ADFS	7.47	Full	Full
L6	FLS	6.71	Full	Full
L7	HeLMS	280	N/A	N/A

NOTE— The information is compiled from the HerMES official website http://hedam.lam.fr/HerMES/index/download.

[a] Each row of the table represents one "nested" field that consists of multiple sub-fields. The levels, names, and areas are listed separately.

2.1.2   H-ATLAS

The *Herschel* Astrophysical Terahertz Large Area Survey (the *Herschel* ATLAS or H-ATLAS) is an OT KP project that surveys $570\,\mathrm{deg}^2$ of the sky in five photometric bands using PACS and SPIRE (Eales et al., 2010). Compared with HerMES, it is $\gtrsim 8$ times larger in the coverage, and the main scientific goal of the H-ATLAS is

to explore the dust masses and dust-obscured star formation for a large number of nearby galaxies. However, using the H-ATLAS data we can also look for bright, distant IR sources.

The fields for H-ATLAS were chosen to be the best-studied large area fields at high Galactic latitudes. To maximize the legacy value of the survey, the fields are chosen to have minimal foreground Galactic debris, and have multi-wavelength data observed from other projects. Figure 2.3 shows an overview of the fields. In brief, The North Galactic Plane (NGP) is a rectangular block measuring 15 degrees by 10 degrees centered on RA=199.5, Dec=29 and rotated by 8 degrees clockwise. The Coma clusters is located within this field. The three fields on the equator are designed to overlap with the Galaxy And Mass Assembly (GAMA) survey (Driver et al., 2009), which is an optical spectroscopic survey of $\sim 300\,000$ galaxies that covers an area of $\sim 286\,\mathrm{deg}^2$ on the sky. The H-ATLAS GAMA fields are at RA 9 hrs, 12 hrs and 14.5 hrs and each covers about $12\,\mathrm{deg}$ in RA and $3\,\mathrm{deg}$ in Dec. The only field in the southern sky, South Galactic Plane (SGP), measures approximately $290\,\mathrm{deg}^2$ and it is also covered by a number of large imaging/spectroscopic surveys.

The first data release (DR1) of the H-ATLAS project was made in 2016 and the detailed description of the data processing procedures is presented in Valiante et al. (2016). The DR1 includes maps and catalogs of the three GAMA fields. Soon after that, the second data release, which provides maps and catalogs for the NGP and SGP fields, are made public by the survey team. The maps are described in Smith et al. (2017), and the catalogs are described in Maddox et al. (2018, in prep.). All the data products are available online at the official H-ATLAS project site[3]. Here we briefly review how the H-ATLAS data products are created.

First a local background is The maps are then

Maps Similar to HerMES, the H-ATLAS team uses HIPE to convert the observed

---

[3] http://www.h-atlas.org/public-data/download
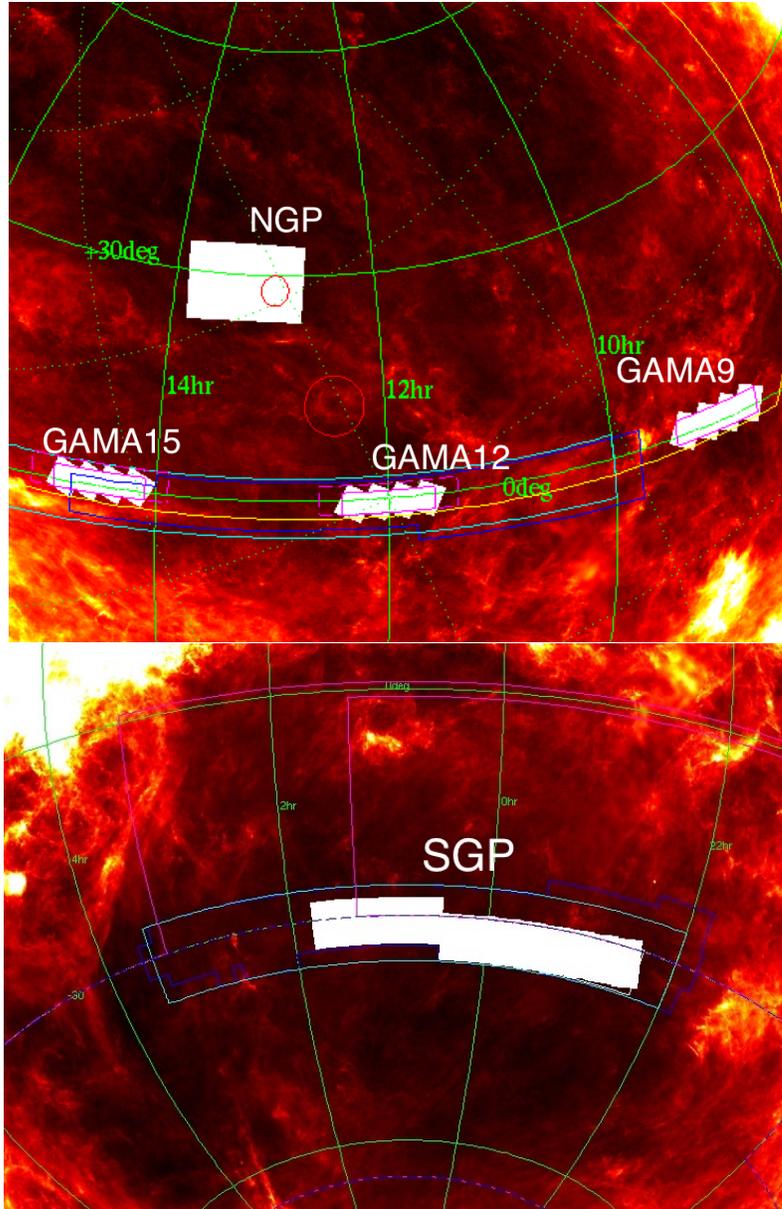
Figure 2.3: Top: H-ATLAS fields in the northern sky. The North Galactic Plane (NGP): A rectangular block measuring 15 degrees by 10 degrees centered on RA=199.5, Dec=29 and rotated by 8 degrees clockwise. The Coma cluster is within this field (denoted by the red circle). Three GAMA fields: Designed to overlap with the Galaxy And Mass Assembly (GAMA) survey. These fields are at RA 9 hrs, 12 hrs and 14.5 hrs and each covers about 12 degrees in RA and 3 degrees in Dec. Bottom: South Galactic Plane (SGP) is the only field in the southern sky and it covers approximately $290 \deg^2$. The figures are adapted from the H-ATLAS official site http://www.h-atlas.org/survey/fields.

TODs to level 1 timeline data. One difference is that H-ATLAS corrected glitches in the bolometer and thermistor data using a different technique (sigma-kappa deglitcher) from the standard module (wavelet deglitcher as used in Her-MES) according to the tests done in Davies et al. (2010). Moreover, visual inspection instead of automatic detection was used to identify rare jumps in the bolometer timelines. Further treatments of the timelines include noise reduction, filtering, etc., are described in great details in Valiante et al. (2016) and Smith et al. (2017). To create maps, the level 1 data are binned according to the pixelization process and averaged. The SPIRE maps are binned using HIPE's naive map-maker at pixel sizes of 6, 8 and 12 arcsec. Pixels without samples are flagged and have no flux value. This is different from HerMES, in which median pixel values are used rather than the mean. Also, the HerMES $350\,\mu\mathrm{m}$ maps have a pixel size of 8.3 arcsec.

Catalogs The H-ATLAS team developed their own source extraction tool, the Multiband Algorithm for source Detection and eXtraction (MADX; Maddox, in prep.) to detect point sources. A detailed description of the algorithm, as well as extensive tests of performance through simulation, is presented in § 5 of Valiante et al. (2016). In general, the first step is to use `Nebuliser` to remove the diffuse Galactic dust emission from all the maps, followed by convolving the maps with matched filters created based on the noise properties. The detection is done by finding the peaks on the so called "variance" map created from the noise map in the $250\,\mu\mathrm{m}$. In the last step, the fluxes are measured on each of the maps at the detected positions using an iterative fit-and-subtract procedure. Fluxes of the brighter sources are determined first and subtracted from the map. At the end, a band-merged catalog is produced for each of the survey field.

2.1.3   Notes on the *Herschel* Catalogs

In this subsection, several important aspects of the *Herschel* sources catalogs generated by the official team are highlighted.

Impact of the Positional Prior Used For both the HerMES and H-ATLAS data products, it provides the band-merged catalogs based on the source detected on the 250 μm maps. The fluxes are determined by fitting the instrument point response function (PRF) at the detection position to the observed science images. The disadvantage of this approach is that we will miss sources that are very faint at 250 μm, but bright at other wavelengths. A number of studies have been done to systematically search for such "drop-outs" using specifically designed algorithms. For example, Rigby et al. (2011) investigated the alternative "red prior" method in which the three images are combined to create a detection map that is sensitive to sources that are bright at 350 and 500 μm wavelengths. From the comparison between the generated catalogs, it is concluded that at the detection limit of $5\,\sigma$ level in 250 μm, $7\,\%$ of the sources at 350 μm and $12\,\%$ at 500 μm will be missed.

It is important to note that due to the large beam size, it is always almost the case that one single *Herschel* sources corresponds to multiple, individual sources found in maps of much higher spatial resolution, e.g., in optical and near-infrared. Therefore, in the ideal case, these "true" positions shall be used as the prior for extracting the *Herschel* fluxes. However, this is not a trivial task due to that fact that *Herschel* maps does not carry detailed enough information to break the degeneracies between close-by sources. The HELP project has done a substantial amount of the work in this regard, developed `XID+` (Hurley, 2017), a tool that uses a probabilistic Bayesian framework to provide a natural framework in which prior information could be included. It uses the Bayesian

inference tool, Stan, to obtain the full posterior probability distribution on flux estimates. Using this method, they produced *Herschel* catalogs for the *Spitzer* MIPS 24 µm sources. In this work, we propose our own treatment to deal with the source blending issue. The algorithm and the implementation is described in in § 4.

Confusion Noise Due to high degree of source confusion on the *Herschel* (especially the SPIRE) maps, it is critical to quantify the confusion noise and include it in the error estimates of the source catalogs. For both HerMES and H-ATLAS, this is done by extensive end-to-end simulation. The first step is to create a pseudo sources catalog based on some source count model (e.g., HerMES used Béthermin et al. (2010)). The sources in this catalog is then placed on the sky and simulated *Herschel* maps are created. Photometry is then performed on these maps to get the measured fluxes of the input sources. One can estimate the confusion noise based on the comparison of the measurements with the true fluxes. For HerMES catalogs, they assume a constant confusion noise over a single map. For example, the SPIRE maps of GOODS-North field have confusion noise of 5.6, 7.4, and 7.7 mJy for 250, 350, and 500 µm band, respectively. For H-ATLAS, however, the confusion noises adopted are dependent on the fluxes of the sources, e.g., for 250 µm catalog:

$$\sigma^2_{conf250} = \min\left(0.0049, F/5.6\right)^2 + 0.00253^2 \,, \tag{2.1}$$

where the flux density F is given in Jy. The detailed discussion is presented in Valiante et al. (2016).

Astrometry Accuracy The astrometry accuracy is one of the most important properties for a source catalog. In the map-making step, the astrometry calibration is done when converting the raw observed data to the level 1 timelines

using HIPE (Smith et al., 2012). It is found that for the H-ATLAS maps, the typical positional offset is about one to few arcseconds when compared with SDSS (Valiante et al., 2016), being consistent with the instrumental limit (Pilbratt et al., 2010). At the catalog level, it appears that the source positions are dependent to the signal-to-noise ratio of the sources. Ivison et al. (2007) found that the $1\,\sigma$ uncertainty of the R.A. and Dec. follows

$$\Delta\alpha = \Delta\delta = \frac{0.6\theta}{\text{S/N}}\,, \tag{2.2}$$

where θ is the beam size (e.g., $18.15''$ for HerMES $250\,\mu\text{m}$ map).

2.2 Overview of Multi-wavelength Data for Wide-field Herschel Surveys

The multi-wavelength datasets are critical to make interpretations of the *Herschel* data. In this section we review such datasets available to-date that we could make use of to exploit the science behind the *Herschel* maps and catalogs, and an overview is given in Table 2.2 for the well-studied legacy fields, and Table 2.3 for the ones that still lack complete wavelength coverage.

2.2.1 Multi-wavelength Data in the Extragalactic Legacy Fields

The *Herschel* fields listed in Table 2.2 are considered to be the legacy fields and there exist rich datasets that range in wavelengths from X-ray to radio. These fields are the primary fields for a number of well-known, high-significance projects including CANDELS, LESS, COSMOS, AEGIS, UKIDSS, GAMA, etc.

The Cosmic Assembly Near-IR Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011) is designed to document the first third of

Table 2.2: Multi-wavelength data available to the well-studied *Herschel* legacy fields

Field Name	Area (deg²)	Projects
GOODS-S/	0.04/	CANDELS[a], *Spitzer*[b], LESS/ALESS/zLESS[c]
ECDFS	0.35	
COSMOS	2.38	CANDELS, COSMOS[d], *Spitzer*, S2CLS[e]
GOODS-N	0.55	CANDELS, *Spitzer*, VLA-GOODSN, S2CLS
EGS	0.6	CANDELS, AEGIS[f], *Spitzer*, S2CLS
UDS	0.8	CANDELS, UKIDSS UDS[g], *Spitzer* S2CLS
H-ATLAS GAMA	126	GAMA[h]

[a,c,d,e,f,g] See § 2.2.1 for a brief overview of the projects.

[b] The entries with *Spitzer* all have both IRAC and MIPS data available.

[h] See § 2.1.2 for references to the GAMA field.

galactic evolution from $z = 8$ to 1.5 via deep imaging of more than $250\,000$ galaxies with *HST* WFC3/IR and ACS. Five premier multi-wavelength sky regions are selected; each has multi-wavelength data from *Spitzer* and other facilities, and has extensive spectroscopy of the brighter galaxies. The fields are GOODS-N, GOODS-S, COSMOS, UDS, and EGS, within all of which deep (Level 1 to 3) *Herschel* observations through HerMES are available. The survey is designed to yield statistically robust and complete samples of galaxies down to $10^9\,\mathrm{M_\odot}$ out to $z \sim 8$. In addition to the HerMES data, a number of separate, *Herschel* open time projects are also available, including GOODS-Herschel (PI: D. Elbaz), and CANDELS-Herschel (PI: M. Dickinson).

All of the CANDELS fields have their own legacy surveys. COSMOS field has been observed at all accessible wavelengths from the X-ray to the radio with most of the other major space-based (*Spitzer*, GALEX, XMM, Chandra, NuStar) and ground based telescopes (Keck, Subaru, Very Large Array (VLA), European Southern Observatory Very Large Telescope (ESO-VLT), United Kingdom Infrared Telescope (UKIRT), The National Optical Astronomical Observatory (NOAO) Badde and Blanco telescopes, the Canada France Hawaii Telescope (CFHT), and others) (Scoville et al., 2007). For EGS, all-wavelength data have been obtained by the All-

Table 2.3: *Herschel* fields with partial coverage of multi-wavelengths data

Field	Area (deg^2)	Optical Coverage (deg^2)	*Spitzer* IRAC Ch-1234	*Spitzer* MIPS 24 μm
Boötes	10.57	9.0	YYYY	Y
CDFS-SWIRE	11.39	7.1	YYYY	Y
EGS-HerMES	2.67	1.1	YYYY	Y
ELAIS-N1	3.25	2.8	YYYY	Y
Lockman-SWIRE	7.37	2.0	YYYY	Y
ADFS	7.47	5.9	NNNN	N
ELAIS-S1-SWIRE	7.86	5.1	YYYY	Y
FLS	6.71	5.8	YYYY	Y
XMM-LSS-SWIRE	18.87	17.0	YYYY	Y
HeLMS	280	∼100	YYNN	N

NOTE—The optical data coverage is based on the available public data in various archives (NOAO, CFHT, Subaru, DECam, and ESO) and from our previous follow-up programs. The Lockman-SWIRE field does have optical data from the ESO public imaging survey, however those are too shallow to be useful for this work. The FLS field only has R-band coverage.

Wavelength Extended Groth Strip International Survey (AEGIS) project, including deep imaging data from X-ray to radio and optical spectroscopy over $1 \deg^2$ of the sky, with the aim of studying the panchromatic properties of galaxies over the last half of the Hubble time (Davis et al., 2007). The UDS field is the primary target of the UKIDSS project(Lawrence et al., 2007). UKIDSS is the next generation near-infrared sky survey, the successor to 2MASS. It is designed to survey 7500 square degrees of the northern sky, extending over both high and low Galactic latitudes in JHK bands. For the GOODS-South field, We have additional sub-mm/mm data available through the LESS/ALESS/zLESS project(Weiß et al., 2009; Hodge et al., 2013). These data, due to the high spatial resolution, are of great help to pin-point the true counterparts of the *Herschel* sources, and serve as the great test-field for any source deblending algorithm.

The rich multi-wavelength datasets available in these fields have enable a slew of studies about the IR luminous universe (see § 1). However, it is noted that these fields, being the best-studied *Herschel* fields, only are of a small fraction of the entire

Herschel footprints. In majority of the fields ever observed by *Herschel*, there are still a lot of blank to fill.

2.2.2 Lack of Multi-wavelength Data in the Wide Herschel Fields

Only a small fraction of the *Herschel* survey fields has rich ancillary data across the entire wavelength regime. For most of the HerMES fields, in particular, those assigned with low level number (level $>= 5$), only a limited fraction of optical-to-near-IR data are available, as shown in Table 2.3. Without the optical data, it would be impossible to identify the counterparts of the infrared sources found on the *Herschel* maps, infer their properties, and draw any conclusion about the star formation buried behind the dust.

On the bright side, there are quite a few fields that we have *Spitzer* IRAC Channel 1–4 (3–8 μm) and MIPS 24 μm data, which will greatly facilitate the analysis of the stellar populations as well as help derive robust estimate of the total IR luminosity.

As shown in Table 2.3, we searched the data archives of the large ground-based telescope/instruments including NOAO, CFHT, Subaru, ESO, and DECam, and found that the existence optical coverage already allows us to carry out studies of the IR sources to a scale that supersede the existing, small area legacy surveys. For example, in the Boötes field, 9.0 out of the $10.57 \deg^2$ have already been covered by deep optical imaging from the NDWFS survey by NOAO, in Bw, R and I bands. In the FLS field, 5.8 out of $6.7 \deg^2$ of the data have optical coverage from the CFHT MegaCam in u, g, r, i and z bands. This basically motivated us to start this project of comprehensively studying the *Herschel* sources, as we have already a great deal of data that no one has explored before.

We decided to carry-out the project for the FLS field first. The multi-wavelength data we compile for the FLS field is described in § 2.4. And the results are presented

in Chapter 6.

Also started was our campaign of observations using the WIYN telescope, motivated by the fact that in a big fraction of the fields, there are already some optical data, but not the full coverage in the optical bands (typically, u, g, r, i, z-bands as in SDSS). Having full coverage in the optical is very important in terms of using SED fitting technique to obtain the photometric redshifts. Our observations using WIYN telescope, namely, the Mizzou WIYN Survey, aims at filling in these gaps, which will greatly enlarge the area of *Herschel* fields with full optical coverage. The details of the survey are described in the next section.

2.3   Fill the Gaps: Mizzou WIYN Survey

In 2014, MU joined the WIYN consortium as an operational partner at 5% level for five years (2014B to 2019A). Thanks to this timely investment, we now have guaranteed access to the $3.5\,\mathrm{m}$ WIYN telescope and its suite of instruments. The telescope locates at the best site on top of the Kitt Peak.

To take advantage of this opportunity, we started our campaign (Mizzou WIYN Survey; MizzouWINS) of deep optical observations targeting, within the *Herschel* wide-fields, at the "gaps" where some, but not of the full wavelength range, of the optical imaging data have already been taken by other facilities. MizzouWINS is designed to take at least five years to complete, and upon finishing, the full-range optical coverage will increase from ~ 47.0 to $\sim 120.8\,\mathrm{deg}^2$.

The MizzouWINS makes use of the One Degree Imager (ODI) instrument. ODI is the newest instrument at WIYN. It utilizes both WIYN's one degree field of view and excellent image quality. The focal plane of the optical imager is sampled with 0.11 arcsec pixels, or up to 1 Gigapixels in total. At the current moment, 30 out of a total possible 64 OTAs comprise the central observing area, subtending an area of
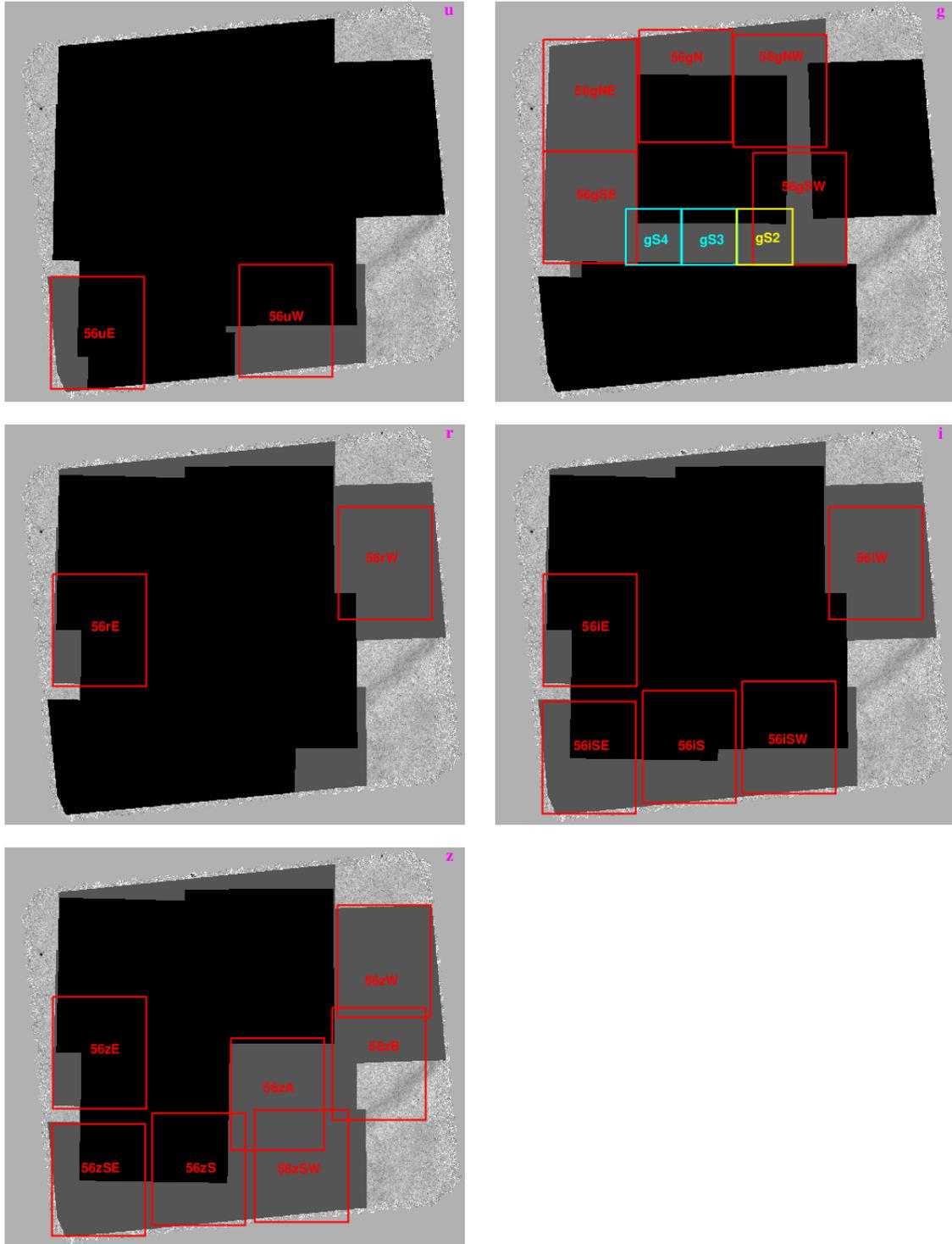
Figure 2.4: Overview of the footprints of the MizzouWINS with in FLS field. The smaller boxes in the g-band are the footprints for pODI, before the upgrading to ODI 5×6 at 2015 summer.

approximately 4048 arcmin^2.

As an example, Figure 2.4 shows the survey design of MizzouWINS for the FLS field.

2.4 First Look Survey Field as a Case Study

In this section, we describe the multi-wavelength data we used for studying the star formation in the FLS field. The *Spitzer* data reduction is also described here.

For the optical, we developed our own tool and machinery for the data reduction, and the details are presented in in the next chapter (Chapter 3).

2.4.1 Introduction to FLS Field: *Spitzer* FLS

Soon after the Science Verification phase of the *Spitzer* at the end of November 2003, the *Spitzer* Science Center (SSC) executed the First Look Survey (FLS) as one of the first science tasks during nominal operations. This program, own to its scientific/operational significance, was scheduled as the inaugural Director's Discretionary Time (DDT) program. Three fields are observed: The Extragalactic FLS (XFLS), the Galactic FLS, and the Ecliptic FLS, among which the first one (XFLS) is observed as level 6 survey of the HerMES program. In this work, we are only interested in the XFLS and hereafter we simply refer to this field as FLS.

The main goals of the surveys were (1) to detect enough extragalactic sources at unexplored sensitivity levels in order to gain insight about the source count; (2) to characterize the dominant source populations with both MIPS and IRAC data from *Spitzer*, plus ancillary data at optical, near-IR, and radio wavelengths, and (3) to explore the cirrus foreground at moderately high Galactic latitudes and its effect on point-source detectability.

The field was observed by the imaging instruments IRAC and MIPS on-board

Spitzer in all of the bands, including IRAC channel 1 to 4 and MIPS 24, 70 and 160 μm. The survey was designed to have 9 astronomical observation requests (AORs) covering the main field with area of $3.8\,\mathrm{deg}^2$ centered on R.A. (J2000.0) $17^h18^m00^s$, decl. (J2000.0) $+59°45'00''$ in a 3×3 grid in array coordinates. There were also additional 6 AORs targeting at the so-called verification area (a $0.25\,\mathrm{deg}^2$ area within) scheduled.

2.4.2 *Spitzer* Data and the Data Reduction

The early release (partial) of the FLS data was done at January 2004, to give general observers real data to assist in planning of their Cycle 1 proposals rather than basing extragalactic observations on highly extrapolated models. At 2004 May, the full pipeline-processed data set was released along with the launch of the *Spitzer* archive. The data released were processed by the S10.5 version of the Spitzer Science Center (SSC) pipeline, described in the IRAC Instrument Handbook[4] and MIPS Instrument Handbook[5]. The details of the data processing and characteristics of the produced maps and catalogs are described in Lacy et al. (2005) for IRAC, Fadda et al. (2006) for MIPS 24 μm and Frayer et al. (2006) for MIPS 70 and 160 μm. In brief, the produced catalogs were able to reach limiting fluxes at $5\,\sigma$ of 20, 25, 100, and 100 μJy for IRAC channel 1 to 4, and 0.3 mJy for MIPS 24 μm.

These data products, however, are no longer optimal for our purpose of studying the properties of the FLS-HerMES sources, due to the following reasons: (1) We aims at deriving the properties of the FIR sources as many as possible from the *Herschel* maps, therefore we need the deepest possible ancillary data; (2) To get the most reliable estimate of the IR luminosities and the dust properties, we would need

---

[4]http://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/iracinstrumenthandbook/
[5]http://irsa.ipac.caltech.edu/data/SPITZER/docs/mips/mipsinstrumenthandbook/

Table 2.4: Coordinates used to obtain the AORs in FLS-HerMES from the SHA

RA	Dec	Radius
254.40646	61.123221	1.5°
259.12498	59.409551	1.5°
254.50543	59.311205	1.5°
259.12466	61.148268	1.5°
263.65849	61.022318	1.5°
263.18818	59.078729	1.5°
263.12451	57.466322	1.5°
258.5836	57.478211	1.5°
254.86536	57.508883	1.5°

accurate photometry in all bands to construct the SEDs; (3) The aforementioned data products were created back in 2005, therefore observations that were made after year 2005, until to-date, were not included in the constructed mosaics; furthermore, the pipeline has evolved greatly since the commissioning of *Spitzer*, therefore we should use the latest pipeline to create the maps and sources catalogs.

For our purpose, we have to retrieve the most complete datasets available from the archive, and perform our own data reduction using the most updated pipelines.

Starting at Fall 2010, all the *Spitzer* data are managed by the Spitzer Heritage Archive[6] (SHA) interface. To obtain *all* the AORs that are available to-date and within the footprint of the FLS-HerMES field, we executed search in the archive at the coordinates with radii as listed in Table 2.4.

As of Jun 2016, this gives 4193 AORs in total for both IRAC and MIPS. An overview of the footprints of all the AORs returned is shown in Figure 2.4.2. All these AORs are processed by the SSC using the latest pipeline version (18.18) at the time of the retrieval. As expected, the number of AORs available at the present is much more than that of the original FLS survey carried out over ten years ago. In the next two subsections, we describe how these AORs (after removing those are completely outside of the FLS-HerMES footprint) are reduced to create the most

---

[6]http://sha.ipac.caltech.edu/applications/Spitzer/SHA/

updated maps and catalogs for the FLS field.
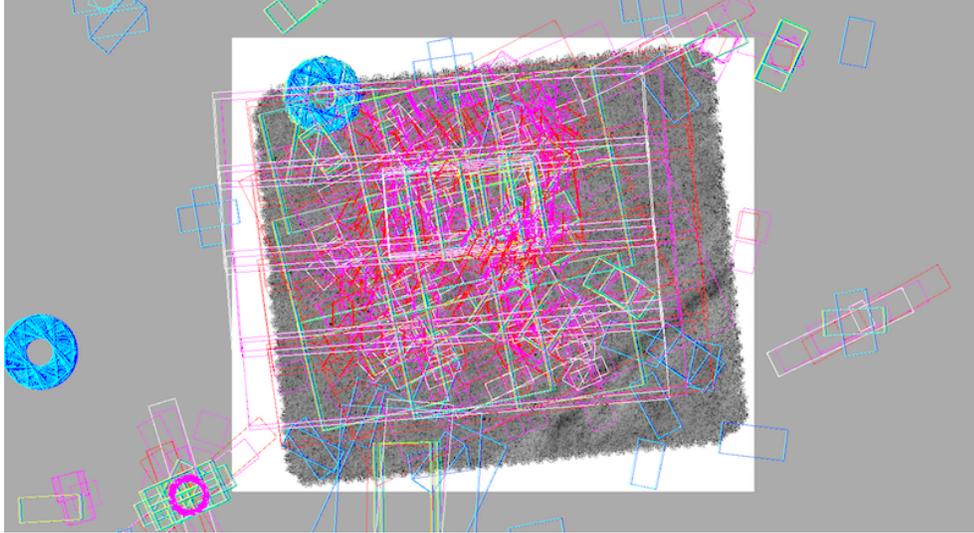
Figure 2.5: Overview of the footprints of the AORs covering the FLS-HerMES field. Each rectangle represents one AOR of certain band, denoted by the colors: blue, cyan, green, and yellow are for the IRAC channel 1–4, and magenta, red, and white are for the MIPS 24, 70, and 160 µm bands, respectively.

Reduction for both the IRAC and MIPS images are done using the official data reduction tool MOPEX. We followed the recipes that are described in the *Spitzer* Data Analysis Cookbook[7].

For the IRAC maps, we removed the AORs that belongs to the regularly visiting calibration filed (shown up as the blue-ish rings to the Northeast of the field), because the repeated, super-deep observations around that area greatly affect the noise estimate of the map. For the MIPS, we only included the AORs that were scheduled as "scan" mode. If not, the smaller AORs will render the final mosaic very patchy, imposes difficulties in extracting catalogs with reliable flux measurements.

We run the official data reduction software MOPEX (MOsaicking and Point source EXtraction) to create mosaics from the Basic Calibrated Data (BCD) and Corrected BCD (CBCD) retrieved from the SHA for the IRAC AORs. In total 95 AORs are included as the input.

---

[7]https://irsa.ipac.caltech.edu/data/SPITZER/docs/dataanalysistools/cookbook/

For the MIPS point source catalog, we used APEX, the official tool to perform PSF fitting photometry. For the IRAC data, we instead adopted the catalog generated from SExtractor, which is tested to provide more accurate error estimates.

2.4.3   Optical Data in FLS

CFHT MegaCam

We looked into the Canadian Astronomy Data Centre (CADC) to obtain the latest available broad-band optical observations targeting at FLS field using the MegaCam instrument at Canada-France-Hawaii Telescope (CFHT).

We executed search in the archive at the coordinates with radii as the same as in Table 2.4. The footprint of exposures returned by the search is shown in Figure 2.6.
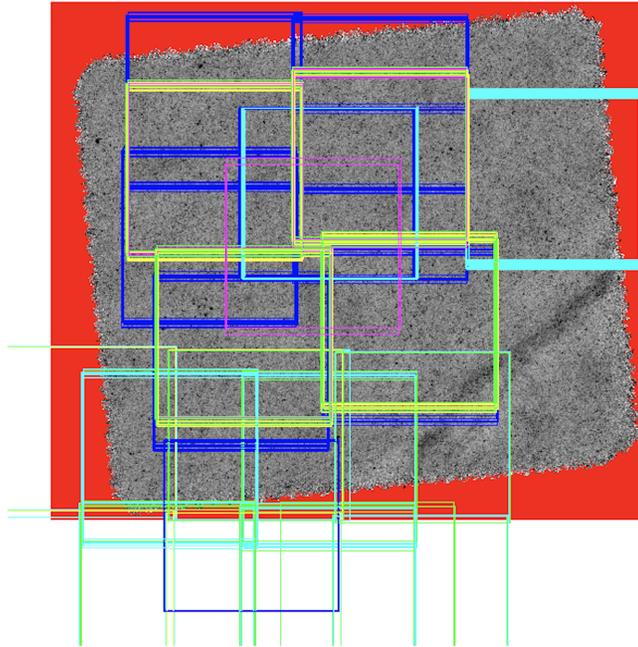
Figure 2.6: The overview of the exposures in FLS field, found in the CADC archive. The colors blue, cyan, green, yellow, and magenta are corresponding to the u, g, r, i, and z-band data, respectively.

To combine the individual observations and create the mosaics, we discard those exposures that have too short exposure time ($\texttt{EXPTIME} < 100s$) in order to improve

the efficiency and final image quality (because those short exposures tend to vary more in quality or condition). The data reduction is described in Chapter 3.

MizzouWINS

FLS field is one of the primary fields that the MizzouWINS aims at covering. The design of the survey is shown in Figure 2.4.

As of the time of the writing, we have completed 4 observation runs that observed the FLS field, the 2014B, 2015B-Oct, 2015B-Nov, 2017B-Sep. As it turned out, the data we took during the 2015B semester was not usable due to an unexpected issue with the instrument. In this project, we include the data taken during 2014B, in total 4 good pointings of the pODI (3×3 ODIs) in the g-band. The detailed data reduction is described in Chapter 3.

Chapter 3

APUS and the Optical Data Reduction Pipelines

The optical data play the most important role in this project, as they (1) are the inputs to the counterpart identification of the *Herschel* sources; (2) are essential to get the photometric redshifts to model the far-IR fluxes and derive the SFR, which can only be done by fitting accurate, well sampled optical to near-IR SEDs (see § 5).

In this chapter, we describe the machinery that we developed and used to reduce the optical imaging data for the *Herschel* wide-field surveys.

3.1 APUS: A Pipeline Building and Management Framework

APUS (Astronomy Pipeline Using ruffuS) package is a pipeline building and management framework written in Python, based on the Ruffus package (Goodstadt, 2010). Our motivation of creating APUS is to aid the data reduction of this thesis work, and provide a repeatable, scalable, and maintainable work flow to the task. The main features of APUS are summarized as follows:

- `APUS` acts as the glue to existing data reduction toolkits/recipes. As mentioned in the previous chapter, we searched the various data archives for the raw optical images in the *Herschel* survey fields. This results in a dataset that consists of data from a great variety of instruments/telescopes/projects, each of which has its own data management system, data calibration procedures, and data reduction toolkits. This imposes difficulties for us to compile all the data available and create master data products in a consistent and maintainable manner. For example, the best practice of reducing the CFHT data is to use the `AstrOmatic` software suite (https://www.astromatic.net), including `SExtractor` for source extraction, `SCAMP` for astrometry and photometry calibration, `SWarp` for creating stacked images or mosaics. The `AstrOmatic` tools are designed to work under command-line, configured using a text file, taking input from and producing output to the local disk. To the contrary, the ODI data that we took from our own survey, MizzouWINS, are processed first by the pipeline named `QuickReduce` built in PPA portal https://portal.odi.iu.edu, which is written in Python, and at the time we obtained the data, the PPA pipeline did not have a mature enough routine for creating stacked images that met our requirements for the photometric accuracy. This inhomogeneity in the data reduction *status quo* involved in this project means that we would have to manage multiple set of parameters/configurations/environments to run the various tools and the connection of the inputs to outputs has to be maintained properly. To make the situation worse, because the optical data reduction usually involves a number of steps such as flat fielding, astrometry calibration, and photometry calibration, etc., it is often the case that for some of the steps, we would like to follow the instruments' best practice, but for other steps, we might need our own special treatment to meet our data quality goal. The great complexity involved essentially motivated us to create `APUS`, which is designed

to be a framework that could centralize the use of different tools, providing a unified and uniform mean to manage the parameters/configurations/environments. What APUS does is that it provides a descriptive language system to "describe" the relation of each of the data reduction steps (could be using any tools) and defines the flow of the data involved. Each task on its own could be executed normally if run from outside of APUS, but when run from within the APUS, the relations among tasks are parsed, and the data are processed and passed between consecutive steps (across the barrier of different tools). This capability greatly reduces the overhead of managing different softwares and their inter-communications, and offers a centralized control over the data processing steps.

- APUS provides repeatable, scalable, and maintainable work flow. Another characteristic of the data reduction task involved in this thesis project is that we are working with a compilation of all available data from the archive, which is not a fixed set as a regular survey would have. For example, in the First Look Survey field, we have the MegaCam data taken using the CFHT, hosted on the CADC archive. The data are mostly from the various PI projects that targeted around the FLS field region over the past years, which subjects to the proprietary period of 13 months after the observation semester. This means that, we are expected to have more and more data available as the time goes. In fact, our initial poll of the FLS data from the CADC resulted in only a 60% of all the currently available i-band data. Also, for the MizzouWINS, the survey was designed to last for 5+ years, and due to the visibility, the observations in one field has to be break up and scheduled into different semesters/years. This means that the data can only be available in batch. With this in mind in designing APUS, we made use of the Ruffus package, which provides functionalities to checkpoint/store the state of execution of the pipeline, and supports a "depen-

dency" resolution scheme, that is, when additional data is added to the pipeline working directory, only the parts of the data that are needed to get updated will get updated, and the rest will be skipped. This fits well to our situation, as we are expected to update the data product by adding new batch of data, however, we would like to skip those parts that are already up-to-date. This feature greatly saved the time that would have been wasted on re-processing that already done result, and most importantly, all these managements are done automatically behind the scene of APUS. As a consequence, the pipeline also carries high repeatability, because using APUS, all the settings/configurations, as well as the data inputs from multiple batches/versions are managed automatically with minimum human interference, therefore, it is easy to copy over the entire pipeline protocol to other machines or computation facilities, and the results would be identical. We made use of this feature in this project, by defining, tuning, and debugging on our local machines on a subset of our entire data, and we run the pipeline on the Lewis high performance cluster available at University of Missouri to make use of the large number of cluster nodes for parallelized computation.

Using APUS, we created the pipelines for reducing the CFHT MegaCam and the WIYN ODI data, which are described in the next two sections, respectively.

3.2 CFHT Data Reduction Pipeline

We retrieved the CFHT MegaCam calibrated individual exposures from the CADC, and produced deep mosaics using the data reduction pipeline that we developed. The pipeline is implemented under the framework of APUS, consists of a suite of scripts/-tools that each performs a single step in the reduction. Most of the customized codes are written in Python, and the AstrOmatic softwares SExtractor, SCAMP and
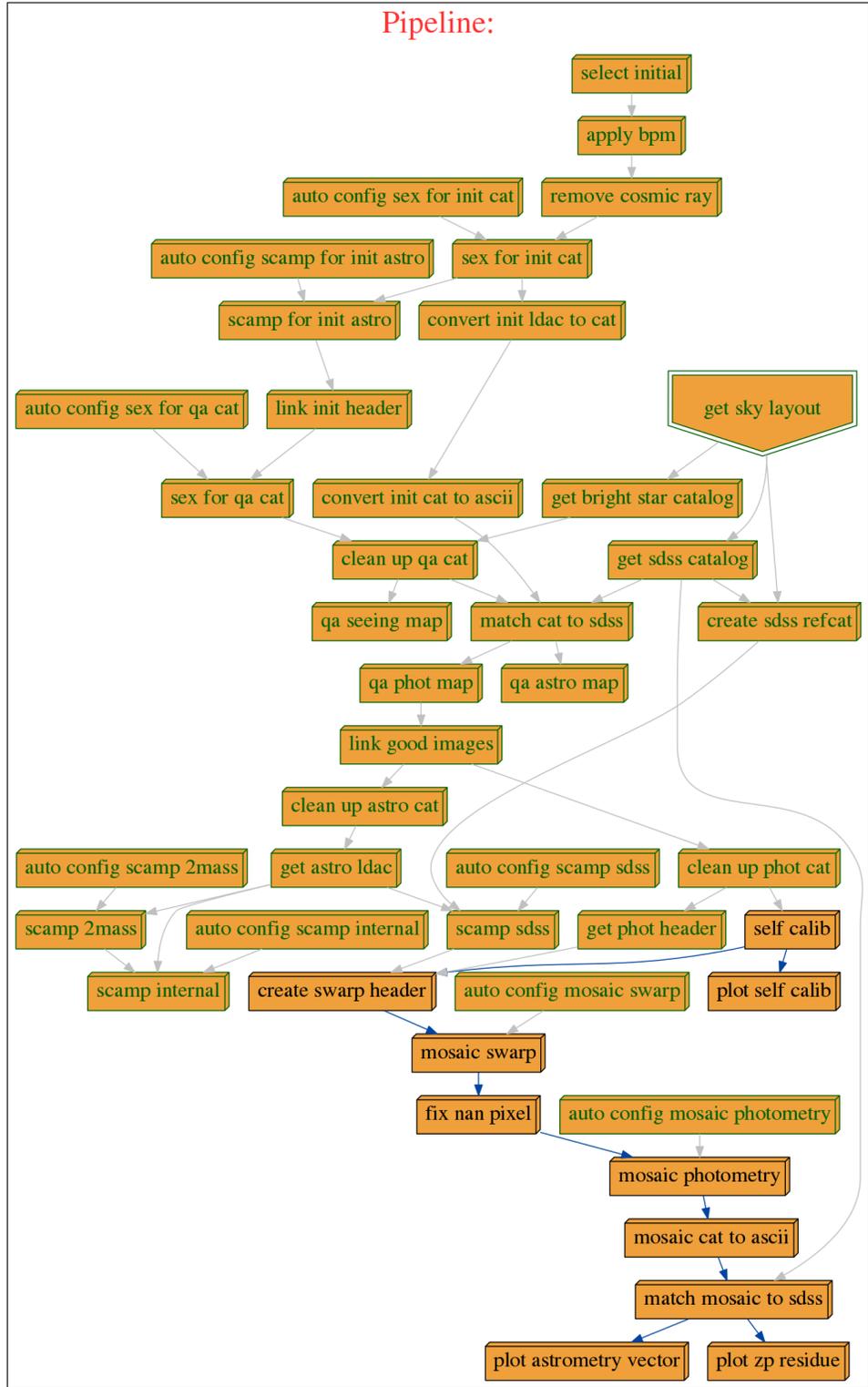
Figure 3.1: The flowchart of the data reduction pipeline for the CFHT.

`SWarp` are used whenever possible to follow the general practices. `Topcat/Stilts` is used to manipulate the catalogs when performance is the concern, otherwise the relevant Python Astropy or Numpy functionalities are used. For an overview of the entire pipeline, see Figure 3.1, in which each node represents one task, and the arrows denote the dependency between tasks. In general, the data reduction steps can be summarized as follows:

- Post-processing the calibrated data. The calibrated images downloaded from the data archive suffers a number of artifacts that have to be treated properly in order to move to the next step, and the most important ones include the bad pixels and the cosmic rays. The CFHT instrumentation team has been maintaining and updating a set of bad pixel masks. To determine the appropriate mask to use, we look for the epoch at which each individual exposure was taken. To reject the cosmic ray, we make use of the `LACosmic` algorithm (Laplacian Cosmic Ray Identification; van Dokkum, 2001). The method is very efficient in detecting cosmic rays of arbitrary shapes and sizes, while distinguishing between under-sampled point sources and cosmic rays.

- Preparing ancillary data. In this step, we retrieve the reference catalogs used for the next steps. We query, for each individual frame, the MAST (The Mikulski Archive for Space Telescopes; https://archive.stsci.edu/index.html) archive for the guide star catalog, which serves the purpose of rejecting bright stars, and the SDSS archive for the latest SDSS catalogs, which is used for the astrometry and photometry calibration.

- Astrometry calibration. The astrometry calibration is done by using `SCAMP`. An initial source extraction is done on the artifact-free image, and `SCAMP` compares this catalog with the SDSS reference catalog to find the positional offsets between the matched sources. A focal plane model is assumed and the offsets

are used to fit the model parameters, which is known as the geometric distortion terms in the associated world coordinate system (WCS) of the image. An updated header file is generated for each image that contains the updated (corrected) WCS, to be used by the mosaicking step performed later.

- Photometric calibration. The photometric calibration is done by comparing the magnitudes of the matched sources used for the astrometry correction. Because of the fact that the CFHT filters are different from the SDSS ones in their response curves, the broad-band photometry is dependent on the exact shape of the spectrum of the particular source. While we do not have knowledge to the spectra of the sources, the common practice is to use the first-order approximation, which goes with the form in our case:

$$m_{SDSS} - m_{CFHT} = k * c_{SDSS} + m \,, \tag{3.1}$$

where m_{CFHT} is the magnitudes as measured on the image, m_{SDSS} is the magnitude in the reference catalog for the same band, c_{SDSS} is the color (the current band to a closest band) in the reference catalog, k is the color term, and m is the constant offset commonly referred to as magnitude zeropoint. To perform the calibration, we fit the above relation to the measured offsets, and determines the best k and m. We note that the fitting is done simultaneously on a set of images that share the same color term k but with different m. Due to the fact that we use SDSS magnitudes as the reference, the calibrated magnitudes are in the SDSS filter system, but not in the CFHT.

- Quality assurance. It is inevitable that some of the exposures are of low-quality due to either instrumental failure or bad weather. To prevent such data from getting into our final data product, we implemented the data quality assurance (QA) steps in our pipeline. Two rounds of QAs are done. The first QA

is done automatically during/after the calibration steps, during which images that shows signatures of outlier properties such as seeing, FWHM, ellipticity, or photometric zeropoint are identified. The second QA is optional, which involves manually going over the catalogs and diagnostic plots, looking for any unexpected images. With these means, we are able to reject the outlier images and create the mosaics using the best input datasets.

- Re-calibrating. The astrometry and photometry calibration are run again using the best input sets to ensure that we have the parameters determined to the best accuracy from the calibration steps.

- Mosaicking and source extraction. The mosaicking is done using SWarp, in which step the images are resampled to a common WCS grid and combined. The source catalog is created from the mosaic using SExtractor. As a final assess of the quality, diagnostic plots showing the properties of the constructed map and the catalog are created.

3.3 PostCalib: An Improved ODI Data Reduction Pipeline

The combined mosaics and catalogs from the MizzouWINS are created using PostCalib. We designed PostCalib to be a standalone tool, again using the APUB framework, that could help create science-ready mosaics from a collection of WIYN ODI images, not only limited to our MizzouWINS data. PostCalib features a full suite of command line switches/arguments to allow the users control how the pipeline is run, and it uses an ASCII configuration table to define the flow of the data.

As mentioned before, the repeatable nature of the APUS pipeline fits great to that of the MizzouWINS survey, as it is often the case that observations of the same field

are scheduled at different semesters, therefore we need to run the pipeline whenever new batch of observations are done. Starting at 2014B, over the past 4 years we have been taken deep optical images targeting the wide *Herschel* fields including FLS, Boötes, Lockman-Hole, and XMM-LSS, and this resulted in an enormous amount of data. APUS's great maintainability has been of great help in organizing the data files and we managed creating high quality stacked maps and deep catalogs in a timely manner as the survey goes.

PostCalib is designed to work as the next step (post-processing) of the WIYN ODI data reduction pipeline QuickReduce (QR). PostCalib typically takes flat-fielded and WCS-corrected images produced from the QR, performs various collection-based operations such as fringe/pupil ghost removal, photometry calibration, image stacking, and source extraction. Figure 3.2 shows the flowchart of all the process steps that the input data have to go through. A walk through of the data reduction steps is as follows:

- Bad pixel mask. The bad pixels on the calibrated data that we download from the PPA are not masked by default therefore we have to apply the mask on our own. We set the pixels of the OTA (the entire image consists of multiple individual OTAs) used for guiding to all NAN; We also make use of the region files provided in QR[1] to mask out the bad pixels. Additional bad pixel region files are also created in some occasions to mask out some non-common artifacts.

- Fringe removal. Fringes are diffraction patterns formed due to uneven thickness of the CCD, and it has to be removed before the image stacking. We group the exposures of same target (by name; it may also be done by grouping by observation date, spanning multiple targets, which will result in a fringe template of higher SNR), and median-combine them to get the raw fringe, using

---

[1]https://github.com/WIYN-ODI/QuickReduce/tree/master/mastercals/bpm/odi_5x6

the following equation

$$I_{comb} = \underset{i}{\texttt{Median}} \left(\frac{I^i}{\texttt{Mode}(I^i)} - 1 \right) , \qquad (3.2)$$

where the mode is computed as $3 \times median - 2 \times mean$, and the superscript ^{i} identifies an individual image within a group. Here we decompose the raw fringe to high and low frequency components, in order to get rid of any large scale background variation, (i.e., the low-frequency component, which is *not* part of the additive fringe pattern and really should be treated in the illumination correction step), and to obtain the real, pure fringe template. The decomposition is done *separately* for each OTA, by assigning a grid of sampling anchors (separation of 500 pixels), and interpolating using 2D spline function. The interpolated OTA data array is the low-frequency component, and the high-frequency component is obtained by subtracting the low-frequency component from the raw fringe. For each masked image, we determine the fringe pattern scaling factor and subtract the scaled **raw** fringe image from the masked image to get the fringe-free image. To determine the scaling factor, we make use of the "fringe vector" region files in QR[2], which contains sets of vectors that starting from dark spot of the fringe pointing to bright spot. The amplitudes of the fringe pattern sampled at each vector is calculated by taking the difference of the medians of the 5×5 pixels vicinity of the head and tail. We measure the amplitudes on both the masked image and the high frequency pure fringe image for a given vector, and divide them to get the scaling factor of this vector. The final scaling factor is calculated by doing a 3σ clipped mean of the scaling factors of all the vectors.

- Preparing initial source catalog and reference catalog. For each input image,

---

[2]https://github.com/WIYN-ODI/QuickReduce/tree/master/mastercals/fringevectors/odi_5x6

we query the online databases to retrieve the reference catalogs and store them locally. We mainly use two databases: SDSS and GSC. The SDSS catalog is downloaded using Python `astroquey.SDSS` module (the latest DR), with a slightly modified version of the query SQL in QR[3], and is used for both astrometric and photometric calibrations. The GSC catalog is queried directly using the http API, which is mainly used for masking very bright objects on science images. The catalogs are then collated based on target name, producing merged reference catalogs for each target. The SDSS catalog contains RA, Dec, and the photometry of SDSS u, g, r, i, z bands, with their uncertainties. The GSC catalog contains RA, Dec, and photometry of three photographic bands Fpg, Jpg, and Npg. We adopt the Fpg magnitude for identifying very bright source. For the sources in GSC that do not have Fpg magnitude cataloged, we estimate one by doing extrapolation using the magnitudes in other bands.

In the first step, we run SExtractor twice on each science frame to get two flavors of raw source catalogs: one is tuned to detect as many real sources as possible for the photometric calibration, and the other is to get non-saturated, bright sources for the astrometric calibration.

For the photometry, we use both Kron aperture (i.e. MAG_AUTO) and fixed apertures to derive magnitudes. The fixed aperture sizes are based on the relevant configuration files in QR[4].

We examine the raw source catalog to reject the sources that are likely false detections or have photometry that are not reliable. First, we check the detector position of the sources and remove those that are located less than 20 pixels to the edge of the cells or the masked regions. Second, we compare the sky position of the raw sources to the GSC catalog, to check if they are affected by

---

[3]https://github.com/WIYN-ODI/QuickReduce/blob/master/query_all_sdss.py
[4]https://github.com/WIYN-ODI/QuickReduce/tree/master/config

any very bright (saturated and blooming) sources nearby. The rejection radius is a function of the Fpg magnitude of the bright source, whose detailed form is determined empirically. Third, we impose magnitude limits to remove the sources that are either too faint or too bright. The end product of this step is cleaned source catalogs that (hopefully) contain minimum number of sources with skewed photometry.

Note that our pipeline is set up such that it is easy to switch to other reference catalogs. For example, 2MASS or Gaia data could be used for calibrating astrometry for regions that the SDSS does not cover. For simplicity, we just assume that we use SDSS catalogs as the reference.

The cleaned source catalogs are cross-matched with the SDSS reference catalog to produce a zeropoint catalog for photometric calibration purpose. The matching radius is 1.2 arcsec. During the match, we only keep the sources with SNR > 10 in the SDSS catalog.

- Photometric calibration. We developed a self-calibration formalism to determine the zeropoints and color terms of individual exposures.

  The procedure starts by merging a number of zeropoint catalogs for a certain target, or for a certain night's observation. Each entry in the merged catalog has the information of:

  – source properties extracted from ODI image, e.g. X_IMAGE, Y_IMAEG, MAG_AUTO, MAGERR_AUTO

  – source properties from SDSS reference catalog, e.g., RA, DEC, u, g, r, i, z magnitudes, etc.

  – catalog information, including IMAGE_ID (from which individual image this source is detected), and OTA_ID (from which ODI this source is detected)

The goal of the photometric calibration is to determine a color term slope (k) and a magnitude zeropoint (b) so that the measured instrument magnitudes can be converted to the magnitudes in the reference catalog. Our self-calibration procedure is based on the following assumptions about the aforementioned parameters. The notation X_i^x denotes a quantity X (could be k, b, or any magnitude, etc.) at OTA of OTA_ID=x on image of IMAGE_ID=i:

- The color slope k is constant across images and OTAs

$$\forall\ i,\ x, \quad k_i^x = k \tag{3.3}$$

- The ZP offset of a given OTA b^x w.r.t. a reference OTA (without generality, say x=0) is constant across images

$$\forall\ i, \quad b_i^x - b_i^0 = b^x - b^0 \tag{3.4}$$

- The airmass/transparency term of a given image c_i is constant across all OTAs

$$\forall\ x, \quad c_i^x = c_i \tag{3.5}$$

Therefore, the calibrated magnitude on the input images can be calculated with

$$\mathsf{MAG}_{cal,i}^x = \mathsf{MAG}_{ins,i}^x + k \cdot \mathsf{COLOR}_{ref,i}^x + b^x + c_i\,, \tag{3.6}$$

where

$$\mathsf{MAG}_{ins} \quad \text{is measured instrumental magnitude,}$$

$$\mathsf{MAG}_{cal} \quad \text{is calibrated magnitude, and}$$

$$\mathsf{MAG}_{ref}/\mathsf{COLOR}_{ref} \quad \text{is magnitude/color from reference catalog.}$$

The parameters k, b^x, and c_i are determined by minimizing the following function:

$$\chi^2 = \sum_{\text{all sources}} \frac{(\mathsf{MAG}_{ref,i}^x - \mathsf{MAG}_{cal,i}^x)^2}{\mathsf{MAGERR}_{ref,i}^{x}{}^2 + \mathsf{MAGERR}_{cal,i}^{x}{}^2} \, , \tag{3.7}$$

where the colors used during the computation for different bands are

```
odi_u: sdss u - g
odi_g: sdss g - r
odi_r: sdss r - i
odi_i: sdss i - z
odi_z: sdss i - z .
```

We use Python package `lmfit` for minimizing the above function to get the best-fit parameters, which are then used to calculate the FLXSCALE keywords corresponding to magnitude zeropoint of 25 for each OTA for each image to be used by the mosaicking routine SWarp. The keyword values are stored in SWarp-compatible ASCII header files.

- Astrometric calibration. We use SCAMP for astrometric calibration, i.e., determining the geometric distortion. The inputs are the bright raw catalog extracted from exposures corresponded to a single target. If multiple targets are close to each other, the procedure can also run them all together.

This step is the single most vulnerable link of the entire pipeline. A detailed discussion of possible tricks would be beyond of the scope of this document. In short, it is usually the case that one need to first derive a good initial focal plane model by setting the following parameters in SCAMP

```
MOSAIC_TYPE      FIX_FOCALPLANE
HEADER_TYPE      FOCAL_PLANE
```

and then use the output focal plane model (the ".ahead" file) as the global header to run SCAMP again with

```
MOSAIC_TYPE      SAME_CRVAL
HEADER_TYPE      NORMAL
```

in order to get sufficiently good source matching and focal plane solution. If fine-tuned successfully, the SCAMP run can usually result in a highly optimal correction.

- Quality assurance. This step creates overall reports regarding the quality of each input image, including seeing, astrometry error, photometry error, source count, limiting magnitude (depth), etc. These QA reports are tabulated as a single table and will be queried in the mosaicking stage to decide what strategies to be used in creating the final stack.

- Mosaicking. Once we have the photometric calibration information (zp_hdr) and astrometric calibration information (astro_hdr) in place, we run SWarp to create the stacked image and weight map.

  Optionally, one can determining how many stacks are to be created and what images to go to which stack based on the QA report generated from the calibration phase. For example, one might want to stack only the images with good seeing.

One also can control whether or not to use the astrometric calibration obtained from the SCAMP run, as it might not be as good as the QR determined solution if not fine-tuned.

The re-sampling and co-adding are controlled by various parameters set in the SWarp configuration file, including projection plane, mosaic size, combining algorithm, etc.

The results of this step include both science mosaics and weight maps that comprising the coverage information for estimating the magnitude errors.

Note that by default, the mosaics produced by SWarp have zero value filled in for those pixels that are of weight zero, i.e., bad pixels. We replace these pixels with a value of NAN.

- Source extraction and final diagnostics. We perform the final source extraction using SExtractor on the stacked images and measure the quality of the stacks. The parameters of the SExtractor run are tuned to pick up as many source as possible. The magnitude zeropoint is set to be 25, matching the targeting magnitude zeropoint of the calculated FLXSCALE keywords from photometric calibration step. The resultant catalog is matched against the SDSS reference catalog to generate a quality verification catalog, for assessing the goodness of the photometry and astrometry.

To compare the photometry to SDSS, one also need to fold in the color term derived from the self-calibration process, to transform the filter system from ODI to SDSS. A diagnostic plot is generated in this manner showing the mean and dispersion of the magnitude differences between the mosaic catalog and the reference catalog.

To get a sense of the astrometry accuracy, we simply generate a vector plot showing the on sky offset between the extracted sources and their SDSS coun-

terparts.

A histogram of the distribution of magnitude of all the extracted sources is also generated to give a sense of how deep the mosaic goes.

Some additional notes to the `PostCalib` pipeline are:

1. Not all the steps are necessary to get a science-ready mosaic. For example, the fringe removal is only for i band and z which have fringes present.

2. The quality of the self-calibration procedure depends on how **clean** the source catalogs are. Some trial-and-error tuning of the magnitude limits during the cleaning step may be necessary. Also, due to different sensitivities in different bands, the optimum limiting magnitudes are different. Another caveat of the cleaned catalog is that the close-to-very-bright-object rejection could fail either due to proper motion of the bright source, or due to the failure of the empirical rejection radius algorithm. To deal with that, our pipeline generates diagnostic plots of the sky positions of the sources for each cleaned catalog for visual inspection and approval.

3. The magnitude zeropoint of the individual image is set to 25.0 in our self-calibration process, and SWarp is configured to use this zeropoint in creating the stack. However, we have observed some small offset between the actual zeropoint and the expected value of 25.0. SWarp was known to have this small caveat, and probably this still has not yet been fixed. Therefore, one should pay attention to this problem and have it measured and fixed by comparing to the reference catalog.
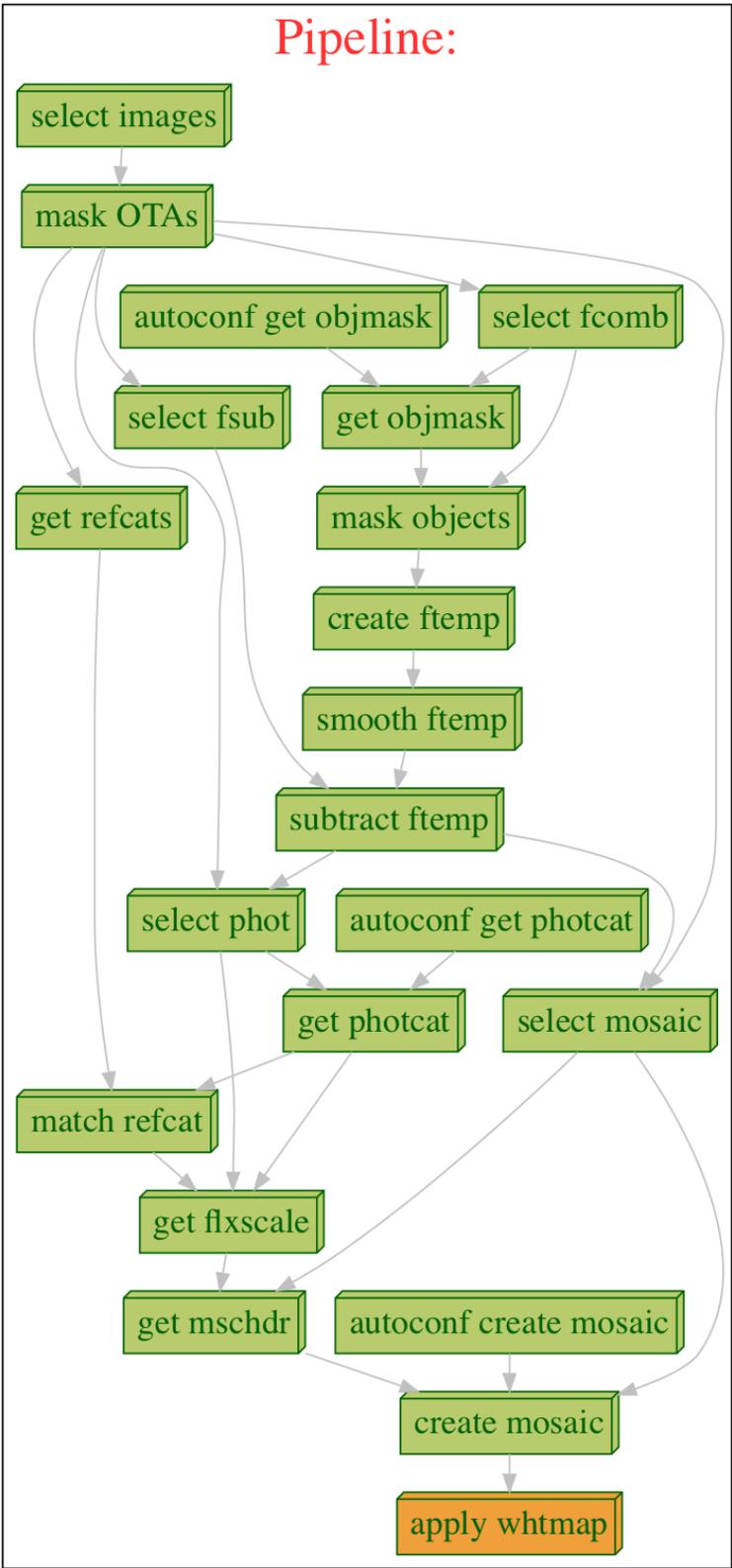
Figure 3.2: The flowchart of the `PostCalib` pipeline.

Chapter 4

CIDer: *Herschel* Source De-blending with Optical Priors

As stressed in Chapter 1, dealing with the source confusion of the *Herschel* data at FIR wavelengths is critical but technically challenging.

During this thesis work, we fully develop and implement the source de-blending method as described in our pilot study as presented in Yan et al. (2014, hereafter Y14). This source de-blending method is unique in that we directly use high-resolution optical or near-IR (when available) images for the position priors, which allows us to accomplish the true counterpart identification and the flux extraction in one step. This is different from other approaches using the mid-IR data such as the MIPS 24 μm (beam FWHM $\sim 7.5''$) for proxies, which still suffer from the blending problem in the first place. Our unique method eliminates any intermediate steps and the ambiguities that still exist in the mid-IR proxy approach, and hence is superior.

4.1 Sketch of the Algorithm

In the initial Y14 paper, we used the GALFIT package developed by Peng et al. (2002) to do the decomposition. While its wide usage by the community is mostly to study galaxy morphologies, GALFIT has a straightforward capability of fitting the PSF at multiple, fixed locations. This well suits our cases, because all the potential components are effectively point sources at the angular resolutions of the *Herschel* instruments.

The algorithm presented in Y14 is as follows.

Ideally, the source deblending procedure starts with a given list of all possible high-resolution sources in the low-resolution beam. Then the candidate will go through an "automatically iterative" (AI) step, where the position priors of all the potential components are fit by GALFIT simultaneously. The objects that have their derived fluxes smaller than the associated flux errors are deemed to be negligible and hence are removed from the input list for the next round. The simultaneous fit is repeated using the cleaned input list, until no negligible objects are left. This procedure usually will converge after 2 to 3 iterations. Among the survived components, some could have fitted fluxes an order of magnitude smaller than the others. When this happens, a new fit is performed using an input list that contains only the major components. If the residual image is of the same quality (by visual examination) as from the previous round, these major components are deemed as the only contributors and those less important ones are ignored. Otherwise we keep the results from the previous round.

However, experiences from Y14 show that in a significant number of cases, we are not able to perform the AI step, due to failure of the GALFIT algorithm. This happens when the input list of candidate contributors includes too many sources that actually do not contribute to the low-resolution flux. This will make the iterative χ^2-minimization process exercised by GALFIT fail to converge or even crash due to singularity encountered when inverting the matrices. There is no better solution to

this except that we fine tune our input candidate list to be as close as possible to the genuine contributors of the low-resolution flux. We managed doing this in Y14 by introducing another strategy, namely, the "trial-and-error" (TAE) step. The process is highly interactive. The TAE step starts with the most possible counterpart within the low-resolution beam and we fit for only this object in the first round. We then check the residual map and see if there are any residuals left at the positions of any other objects in the candidate list. If yes, these objects are added to the fitting list, one at a time, and the fit iterates until reaching the best result possible. The iteration stops when adding more objects either only produces negligible contributors or starts to fail the AI step.

Over the course of this project, we have made significant improvement/change to the procedures to deal with the various problem mentioned above. The most important change is that we implemented a more robust fitting and minimization strategy/routine so that it could always converge, thus only the AI step is needed.

The essence of the algorithm can be described as follows: (1) maintaining a list of sources as input and doing the PSF fitting iteratively; (2) In each iteration, assess and modify the input list according to the fitting results; (3) Stop when a satisfactory solution is found.

4.2   Implementation

We implement the algorithm using Python as the software package `CIDer` (Counterpart IDEntification Routine).

The recommended input file format is ASDF (Advanced Scientific Data Format; Greenfield et al., 2015), which handles better when it comes to a collection of heterogeneous data, although the traditional FITS and/or ASCII files are also supported (in which case a separate YAML file is needed to provide additional information on

the data files).

A typical configuration file for `CIDer` is as follows:

```
# YMAL Configuration file for CIDer
cider:
    xid250:
        type: segmentation
        catalog: spire_xid250
        # fwhm: 18"
        pre_mask: "f250 > 6"
        post_mask: "f250 > 30"
        sort: 'f250'
        image: spire250
    hst_her:
        type: fit
        segm: xid250
        h_cat: f160w_aless
        # h_cat_fwhm: 3"
        h_map: acs_z
```

The entries under key `xid250` defines the protocol for the "segmentation" step, and the entries under key `hst_her` gives instruction on the contribution identification (CID) process. For the data-specifying entries, such as `h_map`, `h_cat`, etc., the value of the entries should be set to the data keys defined in the input dataset, specifying the location of the actual data files. In the case that ASDF file is used as input, the data key definitions are stored along with the data. If separate input data files are used, an YAML file that provides the data key definitions has to be supplied.

The first thing the program does once executed is to prepare the data structure:

```
# prepare the data
1. setup working directory:
    + project.cider
2. link over all images and catalogs with the data keys (in this example
```

```
, map_0, map_1, etc.)
 # list of images and catalogs
     maps: [map_0, map_1, ...]
     cats: [cat_0, cat_1, ...]
 # list of cats for segmentation (low-res image)
     l_cats: [l_cat_0, l_cat_1, ...]
 # list of cats to be used as input position
     h_cats: [h_cat_0, h_cat_1, ...]
 # list of cats to be used as masks
     m_cats: [m_cat_0, m_cat_1, ...]
 # list of images for cid
     l_maps: [l_map_0, l_map_1, ...]
 # list of images for flux extraction
     f_maps: [f_map_0, f_map_1, ...]
```

The next step after the data is prepared is the "segmentation". In contrary to what the name suggests, the "segmentation" is done at the catalog level, rather than on the image. The purpose of the "segmentation" is to break the entire low resolution map (i.e., *Herschel* map for this work) to individual "segements", according to the distributions of sources in the blind extraction catalog from the low resolution map (i.e., l_cat). Upon finishing this step, the large map is broken into multiple segments, and due to the grouping algorithm, there is minimum blending between "segments", therefore the CID can be done within each segment separately. The full procedure is described as follows:

```
# "segmentation"
for each l_cat source:
        * find the close companoin d < 1.5 * FWHM, create a group for
            all the sources save them as lcat
        * find any sources d > 1.5 * FWHM && d < 2.0 * FWHM
            of all the grp members, save that as vicinity
        * compute a circle r that enclose all the group member
```

73

```
        * generate roi of R = r + 3 * FWHM
# after "segmentation" is done
for each roi: we will have three cats:
        * l_cat  # to have their counterparts identified
        * a_cat  # other sources that have d > 1.5 and d < 2.0
        * b_cat  # rest of l_cat in R
```

The next step is to execute the CID routine for each segment:

```
# prepare h_cat
    get a list of h_cat with d < 1.5
    we should exclude sources that have d < 0.5 in a_cat
    group soruces that are too close d < h_grpthresh * fwhm


# prepare image stubs
    the cutout is a suquared region R [-R:R, -R:R]


# fitting
    do the initial fitting
# iterate
    do the AI step until converge
# extract fluxes
    do the fitting (extraction) for each f_map
```

One important treatment that we introduced to remedy the "crashing" problem seen in Y14 is that sources in the high resolution catalog (i.e., the optical catalog) shall be grouped if they are too close, defined by the parameter h_grpthresh. Sources that are closer than this distance are highly degenerated, therefore will fail in the minimization. This is found to be able to greatly improve the robustness of the fitting overall. Once the contributors are found, the fluxes are extracted from the designated low-resolution maps (f_maps).

To enable advanced control of the input list, we introduce the m_cats. For

example, in the fields where we have *Spitzer* MIPS 24 μm data available, we can use the MIPS 24 μm sources positions as the `m_cat`. This will basically reject all the sources that are not within the vicinity of the MIPS sources, by assuming that the MIPS catalog is deep enough to detect all the genuine *Herschel* contributors.

Chapter 5

SED Fitting

The spectral energy distributions (SEDs) are of the most important information that we could get for a distant galaxy. And the SED fitting technique has been used universally for getting the physics properties including redshift, stellar mass, age, etc.

In this thesis work, we construct SEDs for the high-z ULIRGs that span from optical to far-infrared, and we perform SED fitting for both the stellar population sampled by the optical to near-IR, and the dust emission spans from mid- to far-IR. Most importantly, the fitting in the optical-to-NIR will help determine the photometric redshift, which is the keystone to analyze the FIR SED in the absent of spectroscopic data (the case as in this project).

In general, the SED fitting involves the following steps: (1) constructing the input SED from observations of the same target in multiple broad photometric bands; (2) generating a mesh of model SEDs of the same bands from an SED library, constructed empirically and/or theoretically. The generated model SEDs span a certain volume of parameter space configured according to the need. In the context of galaxy optical-to-NIR SEDs, the parameters often include all or part of the follows: redshift, stellar mass, age, metallicity, star forming history, extinction law, dust attenuation, etc.; (3) the input (observed) SED is compared against the model SEDs to find the best-fit by

minimizing the chi-square. The output includes the value of best-fit parameters and the uncertainties, as well as the χ^2 value, which may be used to assess the goodness of the fit.

5.1 Optical-NIR SED Modeling: Photometric Redshift and Stellar Population

The stars and/or the AGN is responsible for most of the optical to near-IR light observed for galaxies.

In the case that the interested galaxy does not host a strong AGN, the most commonly used SED library is the stellar population synthesis model proposed in Bruzual & Charlot (2003, hereafter BC03). The model is built upon the idea that the galaxy light is resembled by all the stars within. In the simplest scenario, simple stellar population (SSP), which is an idealized single-age, single-abundance ensemble of stars whose distribution in mass depends on both the initial distribution and the assumed age of the ensemble, the total flux could be found by using the locus of stars with the same age, called an isochrone, in the Hertzsprung-Russel diagram and integrating the spectra of all stars along the isochrone. The spectrum (flux emitted per unit frequency per unit mass), L_ν, of a SSP of mass M, age t, and metallicity Z is given by the sum of the individual stars:

$$\mathbb{L}_\nu(t, Z) = \int_M \phi(M)_{t,Z} \, L_\nu^*(M, t, Z) \mathrm{d}M \, , \tag{5.1}$$

where ϕ is computed from the initial mass function and L_ν^* is the spectrum of the individual stars. This equation could be generalized to the case where there are multiple (may be discrete or continuous) of star forming events, i.e., a more complex star forming history. Such model is referred to as the composite stellar population

(CSP). One commonly used star forming history goes as an exponentially declining curve with respect to the age of the galaxy, characterized by a time constant τ.

In this work, we use the BC03 and adopt a library of SSP plus CSPs with exponentially decreasing SFH of timescales τ ranging from $1\,\mathrm{Myr}$ to $20\,\mathrm{Gyr}$. All the libraries are based on the Chabrier IMF and the solar metallicity. The models are allowed to be reddened by dust following the Calzetti's reddening law, with the attenuation in V band (A_V) ranges from 0 to $4\,\mathrm{mag}$.

Over the last decades, a great number of softwares were developed to perform the fitting task. Despite the different design preferences, implementations, or user interfaces, they all perform reliably with reasonable setting and tuning, and the results differ marginally. In our pioneer work of Yan et al. (2014), we used `Hyperz`(Bolzonella et al., 2000, 2011) to fit the optical to near-IR SEDs of the CANDELS sources in order to get the photometric redshift, and used the companion tool `Hyperzmass` to derive other physical quantities such as the stellar mass (M), the age (T), etc. For the fitting on our FLS IR major contributor catalogs, we use `LePHARE` developed by Arnouts & Ilbert (2011). `LePHARE` works very similar to `Hyperz/Hyperzmass`, and besides the quantities that we could get from the latter, `LePHARE` has a few more features. First, it has the option of adding emission lines to the model SEDs. However, according to our test, this has little impact on the measured photometric redshifts. Second, `LePHARE` is designed to separate the step of constructing the model SEDs from the actual fitting stage. In the case that we have a large set of observed SEDs but a fixed set of bands, the generated model SEDs could be re-used and save a lot of computation time.

5.2 Mid-to-far-IR SED Modeling: IR Luminosity and Dust Properties

From the mid-to-far infrared SED, we could get insights about the dust properties, and most importantly, measure the total IR Luminosity, which quantifies, in cases where the IR light is due to star forming, the star formation rate. The thermal emission over the full IR regime can be viewed as the collective result of all heated dust components of various temperatures, and the FIR part is dominated by the coldest component.

Two kinds of models are used to characterize the IR SEDs, one is the analytical modified black body spectrum (MBB), the other is the empirical or theoretical templates derived from detailed studies of the nearby IR sources.

5.2.1 Modified Black Body Model

To fit the modified black body model to our constructed IR SEDs, we follow the recipe described in Casey (2012). The paper also comes with an IDL implementation of the recipe, named `cmcirsed`[1] The program could produce estimates of the IR luminosity, the dust temperature, and the dust mass.

The FIR emission due to MBB can be written as

$$S(\lambda) = N_{mbb} \frac{(1 - e^{-(\frac{\lambda_0}{\lambda})^\beta})(\frac{c}{\lambda})^3}{e^{hc/(\lambda k T_{mbb})} - 1} \, , \tag{5.2}$$

where T_{mbb} is the characteristic temperature of the MBB, N_{mbb} is the scaling factor that is related to the intrinsic luminosity, β is the emissivity, and λ_0 is the reference wavelength where the opacity is unity.

For the data in the HerMES wide field surveys, they typically have five bands available, from PACS and SPIRE combined. However, it is generally the case that

---

[1] http://herschel.uci.edu/cmcasey/sedfitting.html

the three SPIRE bands, 250, 350 and 500 μm, are the important ones, due to the fact that the cold dust emission peaks around this wavelength range for the objects with redshift $z \sim 2$. For the PACS bands, either the official team did not produce a sufficiently reliable catalog as reference (for HerMES), or it is limited by the sensitivity due to the survey design (for H-ATLAS). This leaves us only three input data points in the most general case, therefore we have to limit the number of degrees of freedom to produce reasonable fitting result. In our work on the *Herschel* detected SDSS quasars, for which we faced the same circumstance, we showed that the results from the fitting is trustworthy (see Chapter 7 for the detail).

To limit the degrees of freedom, we adopted the default emissivity of $\beta = 1.5$, which is the value typically assumed for cold dust (Casey, 2012). By default, `cmcirsed` sets $\lambda_0 = 200$ μm. We adopted $\lambda_0 = 100$ μm, following Draine (2006). While the exact choice of λ_0 only marginally affects the estimates of the total IR luminosity and the dust mass, it will significantly impact the estimate of the dust temperature.

Taking a close look at Equation 5.2, the term $1 - e^{-(\frac{\lambda_0}{\lambda})^{\beta}}$ modifies the black body term in that it includes a wavelength-dependent optical depth $\tau(\lambda)$, which is assumed to follow a power-law of $\tau(\lambda) = (\frac{\lambda_0}{\lambda})^{\beta}$. The parameter λ_0 is the wavelength where the optical depth is unity.

The reason that T_{mbb} is dependent on λ_0 is that the "modifying" term varies significantly when different λ_0 is adopted. As an example, let us consider a FIR SED that peaks at restframe $\lambda = 100$ μm. The "modifying" term is 0.63 at this peak wavelength when $\lambda_0 = 100$ μm, but becomes 0.94 when $\lambda_0 = 200$ μm. Therefore, in order to fit the peak flux density, the black body term will need to be smaller in the case of $\lambda_0 = 100$ μm, which means that the derived T_{mbb} must be smaller. For further demonstration, we generated a series of MBB models of varying T_{mbb} using $\lambda_0 = 100$ μm, convolved them with the SPIRE band response curves, and then fitted the simulated photometry using the MBB models with $\lambda_0 = 200$ μm. A representative

case is given in Figure 5.1, where the simulated objects are at $z = 2$ and all have $L_{IR}^{mbb} = 10^{12}\,\mathrm{L_\odot}$. The left panel shows the relation of the derived T_{mbb} values and the input values. The peak temperature T_{peak}, on the other hand, will not be affected significantly because the fitting procedure, regardless of the choice of λ_0, will always find the model that best matches the given SED, and therefore the peak wavelength and the peak temperature based on the Wien's displacement law, will not change much.
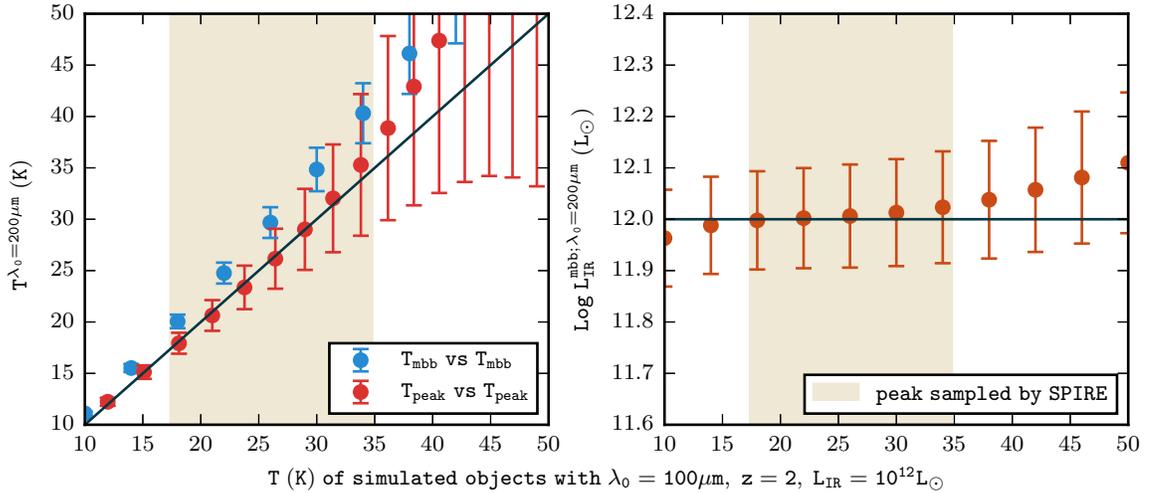
Figure 5.1: Impact of the different choices of λ_0 to the dust temperature and the IR luminosity estimates, using a representative case where the input MBB models using $\lambda_0 = 100\,\mu\mathrm{m}$ are at $z = 2$ and have $L_{IR}^{mbb} = 10^{12}\,\mathrm{L_\odot}$. The left panel shows the comparison of the output T_{mbb} (blue symbols) and T_{peak} (red symbols) estimates, derived using the MBB fit with $\lambda_0 = 200\,\mu\mathrm{m}$, to the input temperature. Similarly, the right panel shows the comparison of L_{IR}^{mbb} estimates to the input value. In both panels, the shaded region indicates the temperature range within which the peak region of the MBB spectra are well sampled by the SPIRE bands.

Similarly, the choice of λ_0 has little impact to the derived L_{IR}^{mbb} values because the shape of the best-fit model is governed by the observed FIR SED. When its peak region is well sampled by the three SPIRE bands, the small differences in the fitted SED beyond the peak will only have a small contribution to L_{IR}^{mbb} and thus have little influence. The right panel of Figure 5.1 shows the comparison of L_{IR}^{mbb} based on the same simulation.
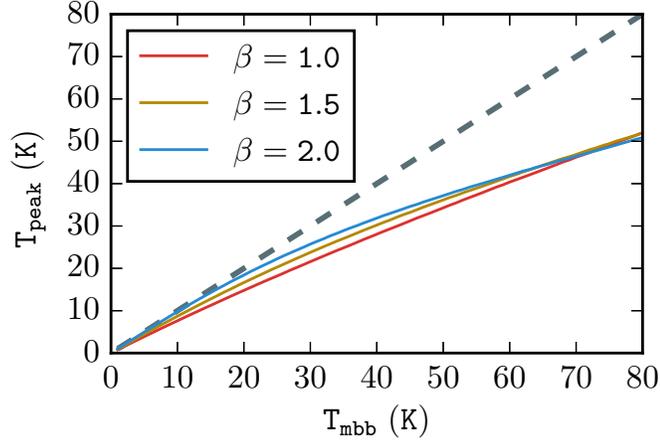
Figure 5.2: Relation between the temperature of the modified blackbody spectrum T_{mbb} and the temperature inferred from the Wein's displacement law T_{peak}, for the three cases when the emissivity β is 1.0, 1.5 and 2.0, respectively. The dashed line represents the equality if these two quantities were the same.

We note that the above form is for general opacity. In the optical thin case, at $\lambda \gg \lambda_0$, the term $(1 - e^{-(\frac{\lambda_0}{\lambda})^{\beta}})$ reduces to $(\frac{\lambda_0}{\lambda})^{\beta}$, which is often adopted in the submm/mm regime. Throughout this work, we used the general opacity form as in Equation (5.2).

Recalling that the FIR emission is dominated by the cold-dust component, we can obtain the total IR luminosity of this component as

$$L_{IR}^{(cd)} \equiv L_{IR}^{mbb} \equiv \int_{8\mu m}^{1000\mu m} \mathcal{L}^{mbb}(\lambda)\mathrm{d}\lambda \tag{5.3}$$

by integrating the best-fit model $\mathcal{L}^{mbb}(\lambda)$ from 8 to $1000\,\mu m$.

We emphasize that $L_{IR}^{(cd)}$ (here represented by the MBB fit) is the luminosity of the cold-dust component over the entire IR range (8–$1000\,\mu m$). While the bulk of its emission is in the FIR regime, this cold-dust component emits beyond FIR and thus it is necessary to integrate over 8–$1000\,\mu m$ to capture its total IR luminosity. Note that this is *not* the total IR luminosity (L_{IR}) of the galaxy that includes the contributions of all dust components over 8–$1000\,\mu m$. In other words, we have $L_{IR}^{(cd)} < L_{IR}$.

The best-fit temperature from `cmcirsed` was taken as the temperature of the cold dust, i.e., $T_{dust} \equiv T_{mbb}$. We also calculated the "peak" temperature inferred from Wien's displacement law, which is given by

$$T_{peak} = b/\lambda_{peak} , \qquad (5.4)$$

where the coefficient $b = 2.898 \times 10^3 \, \mu\text{m K}$. As compared to T_{mbb}, T_{peak} is less sensitive to the specifics of the dust emission models in use and therefore could be a better proxy to the dust temperature when comparing results derived based on different types of templates.

For this project, instead of using the IDL implementation `cmcirsed`, we implemented the algorithm in Python on our own, as a submodule of a general purpose SED fitting package `sedfit` (author: Zhiyuan Ma). Our implementation has a few more features, among which the most important one is that it could compute the equivalent emitting area of the IR source, inspired by our *Herschel*-detected SDSS quasar project (see Chapter 7). This has led to our sequel work on the limited sizes of dusty starbursting regions at high redshift (Yan & Ma, 2016).

In brief, as shown in § 7.5.2, we realized that the T_{mbb}-$L_{IR}^{(cd)}$ trend revealed in Figure 7.8 is not due to the selection effect, but rather the underlying physical mechanism. The location of an object on the T-L_{IR} plane is governed by the intensity of the MBB as expressed in Equation (5.2). To obtain the total power radiated from the cold-dust component of the object, i.e., its total luminosity, one should integrate the MBB function $I_{mbb}(\lambda)$ over wavelength and over the solid angle ("Ω") that its surface area ("A") subtends, and then multiply by this surface area:

$$
\begin{aligned}
L &= A \cdot \int_0^\infty I_{mbb}(\lambda) \, \mathrm{d}\lambda \int \mathrm{d}\Omega \\
&= A \cdot \frac{2\pi k^4}{h^3 c^2} T^4 \int_0^\infty \frac{1 - \mathrm{e}^{-(\tau u)^\beta}}{1 - \mathrm{e}^{-1}} \frac{u^3}{\mathrm{e}^u - 1} \mathrm{d}u ,
\end{aligned} \qquad (5.5)
$$

where we make the substitution of $u = \frac{hc}{\lambda kT}$, and set $\tau \equiv \frac{\lambda_0 kT}{hc}$. We also write T_{mbb} as T for the sake of convenience. If there were no the modified term $(1 - e^{-(\tau u)^\beta})$ to the Planck's blackbody function, the integral above would be $\pi^4/15$ and Equation (5.5) would reduce to the Stefan-Boltzmann law, i.e., $L = A\sigma T^4$, where $\sigma = \frac{2\pi^5 K^4}{15h^3 c^2}$ is the Stefan-Boltzmann constant. For simplicity, we introduce

$$\tilde{\sigma}(T, \beta, \lambda_0) = \frac{2\pi k^4}{h^3 c^2} \int_0^\infty \frac{1 - e^{-(\tau u)^\beta}}{1 - e^{-1}} \frac{u^3}{e^u - 1} du \,, \tag{5.6}$$

which is the equivalent of the Stefan-Boltzmann constant in the MBB case. Note that $\tilde{\sigma}$ is dependent of T because $\tau \equiv \frac{\lambda_0 kT}{hc}$ is involved.

In practice, we can take the approximation that the integral in Equation (5.5) is only over the conventional total IR regime of $S(\lambda)$ from 8 to 1000 µm as in Equation (5.3), and that $L \approx L_{IR}^{(cd)}$. In other words, we have $L_{IR}^{(cd)} = A\tilde{\sigma} T_{mbb}^4$. For simplicity, we shall refer this relation as the "MBB S-B equivalent".

Assuming spherical symmetry, we can define an "effective radius", R_{eff}, such that $A = 4\pi R_{eff}^2$. In other words, this is to imagine that all the star-forming regions within the galaxy are combined together and that the sum can be approximated by an effective sphere with the radius of R_{eff}. We thus have

$$L_{IR}^{(cd)} = 4\pi R_{eff}^2 \tilde{\sigma} T_{mbb}^4 \,. \tag{5.7}$$

For T_{mbb} within the range of interest ($\sim$ 10–100 K), the deviation of $\tilde{\sigma}$ from σ is within a factor of two, and can be well approximated as

$$\tilde{\sigma}(T)/\sigma = 10^{-3}(-2.98T^{1.5} + 45.05T - 118.52)$$

for our choice of $\lambda_0 = 100$ μm and $\beta = 1.5$.

Following the recipe in Casey (2012), we can estimate the dust mass from the

MBB fitting, although extra care has to be taken when using it. This is because the quantity is a strong dependent of the dust temperature ($M_d \propto L_{IR}T^{-5}$; see Casey 2012). Applying a nominal gas-to-dust ratio, the gas mass (M_{gas}) can also be estimated. We adopt the nominal Milky Way gas-to-dust-mass ratio of 140 (e.g. Draine et al., 2007).

5.2.2 Starburst Templates

An alternative way to derive the IR Luminosity is to fit the observed SED to a set of IR SED templates created empirically or theoretically for starbursting galaxies. We used three different flavors of such templates, namely, the theoretical model of Siebenmorgen & Krügel (2007, hereafter SK07; 7220 templates), and the empirical templates of Chary & Elbaz (2001, hereafter CE01; 105 templates) and Dale & Helou (2002, hereafter DH02; 64 templates). The best-fit template is chosen by minimizing the χ-square:

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{f_{obs} - f_{th}}{\sigma_{obs}} \right)^2. \tag{5.8}$$

The advantage of using the starburst templates, compared with the analytical MBB model, is that it also works to a reasonable degree for SEDs that have only two or even one photometry available in the (far) infrared. The error of the L_{IR} derived this way could be estimated by taking the difference between the L_{IR} value of the best-fit template and that of the one of the second smallest χ^2. As compared to the formal likelihood method, this simple approach has the advantage that it works consistently when the parameter space is discrete, and that it includes the possible systematic errors intrinsic to the template set.

Comparing with the single temperature MBB fitting, we note that L_{IR}^{SB} thus derived represents the total IR luminosity of the galaxy. If using the starburst models is appropriate, we should have $L_{IR}^{SB} > L_{IR}^{mbb}$, where $L_{IR}^{mbb} \equiv L_{IR}^{(cd)}$.

We also implemented the template fitting as a submodule of the `sedfit`.

5.3   From SED to SFR

Our goal is to compute the star formation rate for the high-z ULIRGs in the constructed *Herschel* major contributor samples. As introduced in Chapter 1, the dust plays a more and more important role in obscuring the stellar light from the young stars in the galaxies as the redshift goes from 0 to ~ 2. In order to get the most accurate SFR estimate, we have to take into account that in the most general cases we could have both obscured and un-obscured stellar population.

From the optical to near-IR SED fitting, we could obtain the instantaneous SFR from the best-fit BC03 template (SFR_{fit}), and by definition this value naturally takes into account the effect of the dust extinction.

From the FIR SED fitting, we obtain the estimate of the total IR luminosities, which is a well-established SFR indicator, assuming that the IR flux is caused by star forming activities. The equation to convert from L_{IR} to SFR reads:

$$\frac{\text{SFR}_{\text{IR}}}{\text{M}_\odot \, \text{yr}^{-1}} = 1.0 \times 10^{-10} \frac{\text{L}_{IR}}{\text{L}_\odot} \,, \tag{5.9}$$

assuming a Chabrier IMF (Kennicutt, 1998).

The total SFR would then be the combination of the two parts. However, there is a caveat here: The measured L_{IR} includes not only the IR emission from the region completely blocked by dust (L_{blk}) but also the contribution from the exposed region where part of the UV/optical light is reprocessed and re-emitted as IR (L_{ext}). This means that, the observed $L_{IR} = L_{blk} + L_{ext}$. In other word, SFR_{IR} derived from the Kennicutt's law would include part of the light that contributes to SFR_{fit}. To deal with this complexity, we calculate L_{ext} by integrating the difference between the

reddened and the de-reddened spectra from the best-fit BC03 model and assuming that this amount of light is completely re-radiated in the FIR. We then use L_{ext} to compute SFR_{ext} following the Kennicutt's conversion. This way we will get the total $SFR = SFR_{fit} - SFR_{ext} + SFR_{IR}$.

Chapter 6

Star Formation in First Look Survey Field: Results

In this chapter, we present the results that we get for the FLS field as a case study.

6.1 High-level Data Products

Using the tool set that we developed, including `APUS`, `PostCalib`, `CIDer`, we produce a set of high-level data products for the FLS field. We note that these data products, including the deep optical to mid-IR mosaics and the associated catalogs, are the most up-to-date ones of their own kind, and are of high significance to allow a broad range of sciences as they will soon be made public.

6.1.1 Mosaics

We created mosaics for the optical bands u, g, r, i, and z from the archival CFHT MegaCam and the MizzouWINS data.

As examples, in Figure 6.1 and Figure 6.2, we show the CFHT g-band mosaic and the pODI g-band mosaic, respectively.
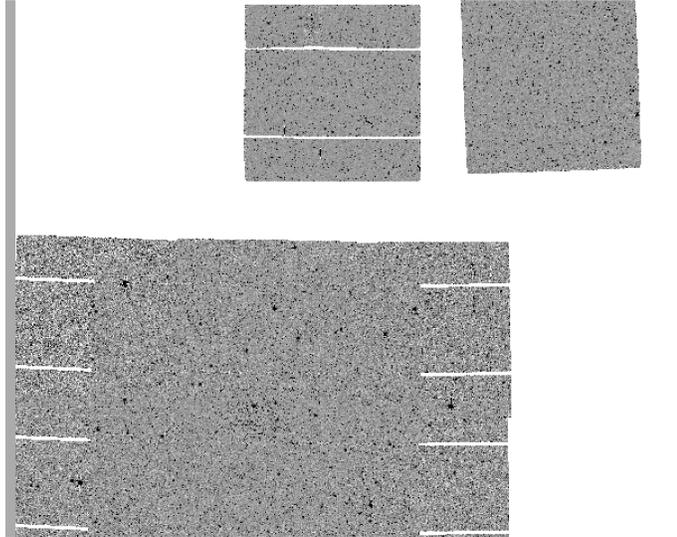
Figure 6.1: The mosaic created using the CFHT MegaCam data. The image is in g-band.
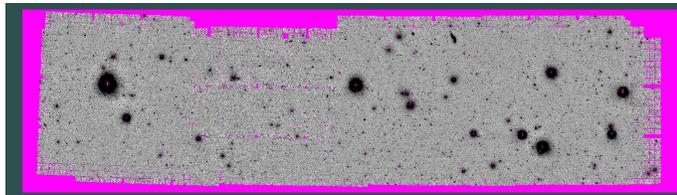
Figure 6.2: The mosaic created using the MizzouWINS data. The image is in g-band.

Using the *Spitzer* data, we created the most up-to-date mosaics combining all the available AORs. Snapshots of the IRAC ch1 map and the MIPS 24 μm map are shown in Figure 6.3 and Figure 6.4.

6.1.2   Catalogs

The optical band-merged catalog is created by matching of blind extracted catalogs in all the optical bands. We note that this is different from the approach of matched aperture photometry, which is widely used in similar context. The reasons are two folds. First, our mosaics are created from a large number of exposures taken in different surveys or PI projects, across a long period of time. This highly heterogeneous nature gives rise to a larger variability in the source FWHM across the entire large
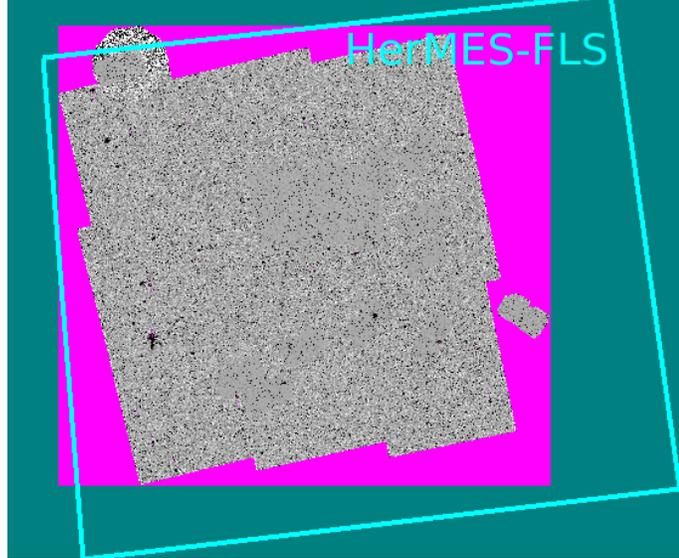
Figure 6.3: The mosaic created using the *Spitzer* IRAC (3.6 μm) data. The cyan outline indicates the footprint of the HerMES FLS map.

mosaic. As an example, Figure 6.5 shows an overview of the FWHM (color coded) of the sources found on the CFHT g-band map. The peak to peak variation could be as large as 0.7 to 1.2 arcsec. This results in the fact that for a given source, the PSF could vary largely across different bands, therefore the nominal matched aperture photometry will not work the best directly. One way to remedy the problem is to perform the so-called PSF matching technique, in which the images are convolved with specifically constructed kernels so that they share the common PSF after the transformation. This however, means that we have to de-grade the quality of the images that are taken under better seeing conditions. Also, the PSF matching is very computationally expensive. It would take too long for our large optical data set to get processed. Second, the matched aperture photometry requires picking a detection band, which is not a trivial task for our scenario, because the footprints in different bands are quite different.

The created catalog contains 214 720 sources that have at least photometry in 4 bands, among which 58 768 have all 5 bands covered. The sources in the catalog span an area of $\sim 5.6 \, \mathrm{deg}^2$ in the FLS field. The photometric accuracy is typically
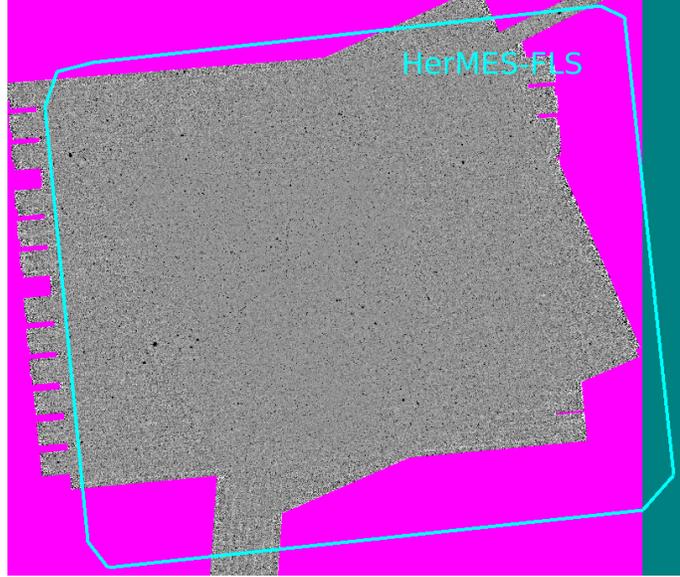
Figure 6.4: The mosaic created using the *Spitzer* MIPS 24 μm data. The cyan outline indicates the footprint of the HerMES FLS map.

2.3% for the CFHT data, and for the WIYN ODI data, it is $\lesssim 4\%$. As an example, Figure 6.6 shows the dispersion of the extracted magnitudes compared with the SDSS photometry.

For the IRAC Ch1 and Ch2 catalogs, we made use of the matched aperture photometry, and we apply an $SNR \geq 3$ cut. In total, the Ch1 & 2 catalog has 212726 sources. The IRAC catalog is matched using a radius of 1.2 arcsec to the optical band-merged catalog, creating the optical to NIR catalog. Note that the sources that have both optical and IRAC photometry spans an area of $\sim 3.8 \, \mathrm{deg}^2$.

6.2 Optical to FIR SED

We run `CIDer` to put together the optical (-to-NIR) band-merged catalog, with the HerMES DR 3/4 catalogs. In total the HerMES xID250 catalog contains 16694 sources, among which 5900 are of SNR ≥ 5. A match to the optical catalog results in a number of 160 258 sources that are candidates to the *Herschel* flux contributors,
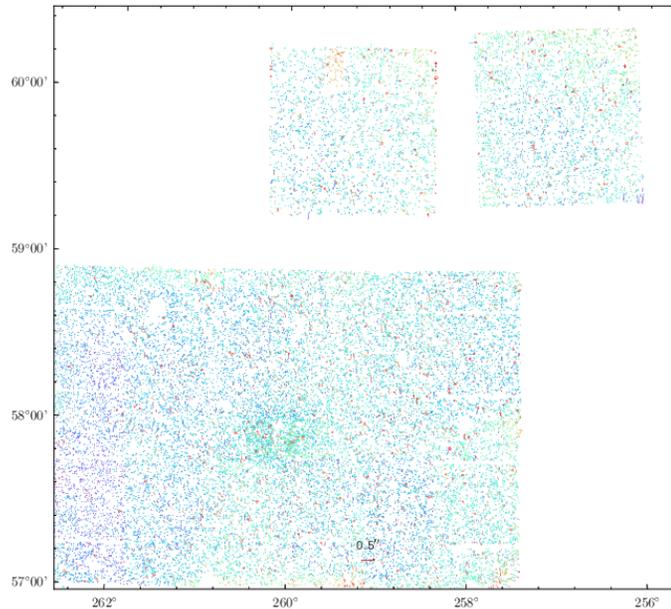
Figure 6.5: Overview of the FWHM of the sources found on the CFHT g-band map. Small values are shown in purple and larger values are in red.

among which 33 888 have at 4 optical bands' photometry, for which we could determine the photometric redshifts in the next SED fitting step. After running `CIDer`, 13 143 sources are found to be contributors to the *Herschel* fluxes. We note that because we have MIPS data available in FLS, we make use of the 24 µm positions to reject optical sources that are not in the vicinity of the MIPS sources. This may not be done for other fields in which such data are not available. For the effect of MIPS 24 µm data to the `CIDer` results, see § 6.3.

6.2.1 Photometric Redshift

We obtain the photometric redshifts for the sources in the our full-range SED catalog (hereafter CID catalog). The distribution of the redshifts is shown in Figure 6.7. The source in the CID catalog span a redshift range of 0 to ~ 3, and for the subsample that at least contributes 50% of the blended *Herschel* flux, the distribution is similar in general, with some decrease in number for the $1 < z < 2$ range.
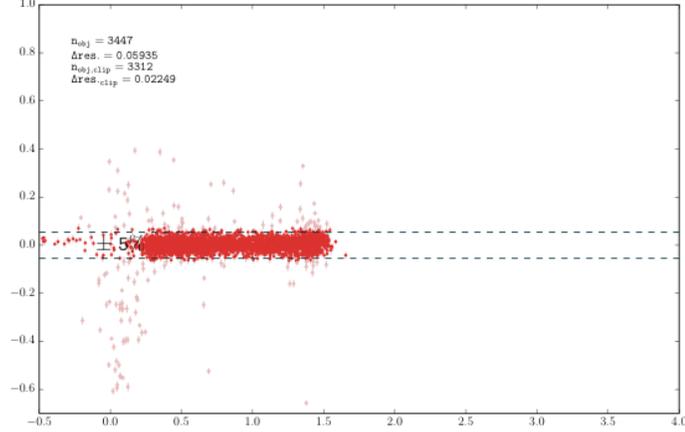
Figure 6.6: The dispersion of the extracted magnitudes compared with the SDSS photometry. The x-axis is the SDSS color (g-r), and the y-axis is the magnitude offset.
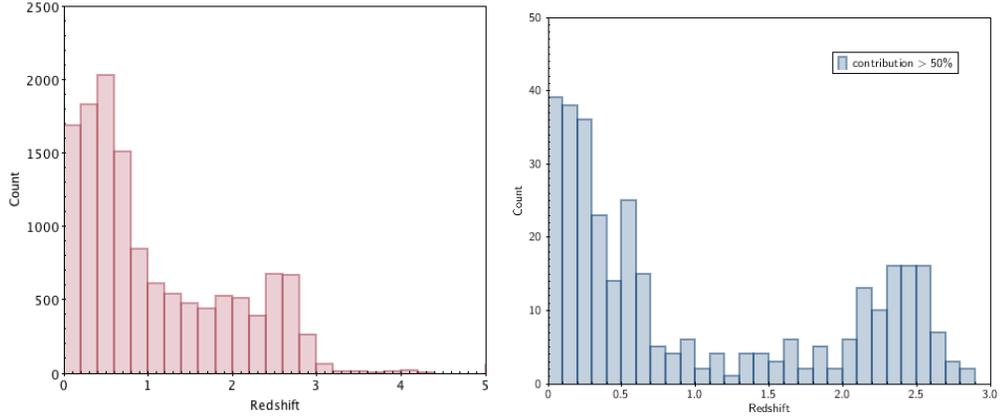
Figure 6.7: The redshift distribution in our CID catalog. The left panel shows that for the entire CID catalog, and the right panel is for those with contribution $\geq 50\%$ to the blended FIR flux.

To assess the reliability of our photo-z result, we made use of the spectroscopic redshifts measurements in the literature: (1) FLS MMT/Hectospec Spectroscopic Catalog by Papovich et al. (2006); (2) XFLS WIYN/Hydra Spectroscope Catalog by Marleau et al. (2007). The comparison is shown in Figure 6.8. The results show that, in general, out photometric redshift is in agreement with the Spectroscopic redshifts, with a standard deviation of about 0.1 and a systematic offset of -0.08. Most importantly, a fraction of sources suffers the "catastrophic failure" problem, which resulted in much higher redshifts than the spectroscopic ones. The reason
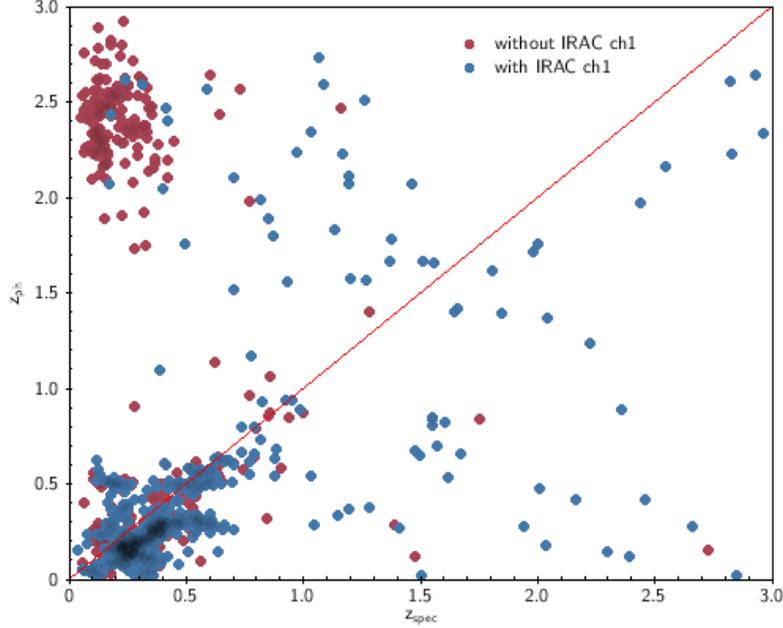
Figure 6.8: The comparison of our photo-z to the spectroscopic redshifts of a sub-sample of sources in our CID catalog. The red points are for those without IRAC data, and the blue points are for those with IRAC data.

for such failure is the degeneracy in the fitting, which could be solved by add more bands to the input SEDs. In our case, we found that, by incorporating the *Spitzer* IRAC data, we are able to obtain good estimate of the photo-z for the sources with catastrophic failures. However, this will shrink our sample size, due to the fact that not all the optical sources have the IRAC data available.

6.2.2 Total IR Luminosity, Dust Properties, and L_{IR}–T Relation

Using the photometric redshifts and FIR fluxes assigned to each optical source in the CID catalog, we derive the IR Luminosity L_{IR} as well as the dust properties such as dust temperature T. The measured L_{IR} ranges from $\sim 3 \times 10^{10}$ to $\sim 10^{13}\,\mathrm{L_\odot}$, where the lower limit is due to the SNR cut in the *Herschel* data for reliable `CIDer`.

Figure 6.9 shows an example of the SED fitting and the results for a random source in the CID catalog. The source is found to be at $z_{ph} = 0.55$, and the measured

$L_{IR} = 1.6 \times 10^{11} \, L_{\odot}$, $T = 24.6 \, \text{K}$. Using the Kennicutt's law, we can infer the SFR that is due to the IR emission is about $16 \, M_{\odot} \, \text{yr}^{-1}$ for this source, assuming a Chabrier IMF.
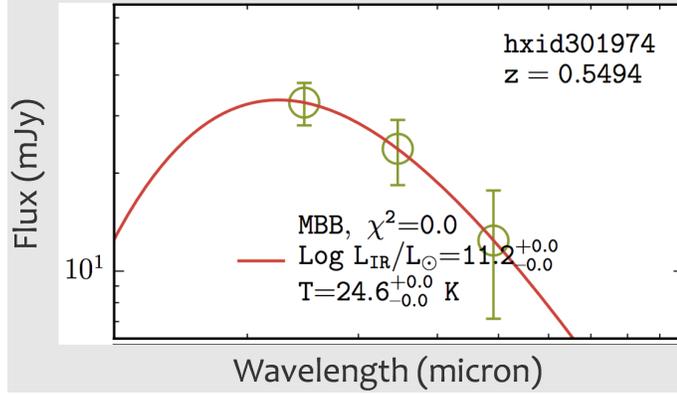
Figure 6.9: An example of the SED fitting and the results for a random source in the CID catalog. The source is at $z_{ph} = 0.55$, and the measured $L_{IR} = 1.6 \times 10^{11} \, L_{\odot}$, $T = 24.6 \, \text{K}$.

With the measured L_{IR} and the dust temperatures, we plot the source distributions on the L_{IR}–T plane, as shown in Figure 6.10. We can see the similar turn-over point at around $L_{IR} = 1 \times 10^{12} \, L_{\odot}$, after which the IR light emitting size, if attributed to the star formation, stop increasing and the trend follows that of the Stefan-Boltzmann law.

6.2.3   Contribution to Global Star Formation Rate Density

With the catalog constructed, we are now at the position to explore the sciences. One of the most interesting topics that we are able to discuss is the contribution of the sources to the global star formation rate density. To perform the calculation, we need a comprehensive test of the catalog for selection effects and any biases. In the near future, using the catalog and the measurements that we got, we will be able to plot the total SFR density as a function of the redshift, contributing data points to the Lilly-Madau diagram.
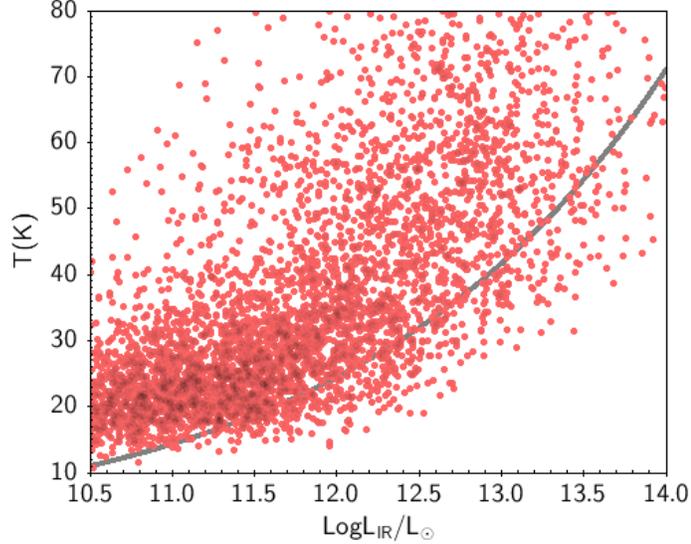
Figure 6.10: The sources in the CID catalog, plotted on the L_{IR}–T plane. The gray curve shows the Stefan-Boltzmann law in the modified black body case with a fixed emitting area.

6.3   Discussion: Test on the Performance of `CIDer`

Our work is based on the source de-blending method `CIDer`, and in this section, we present our test on its performance using the sub-mm/mm/radio data with high spatial resolution. We selected two *Herschel* fields that have such data available: COSMOS and GOODS-South.

In the COSMOS field, we have deep 1.4 GHz data obtained from the VLA-COSMOS project (Schinnerer et al., 2004). It is to date the largest radio deep field to the depth of 7–15 μJy/beam and has angular resolution of 1.5 arcsec. For the optical to near-IR data, we used the catalog from Ilbert et al. (2013), which contains 339 384 sources. We ran `CIDer` using the optical-to-NIR catalog and the HerMES xid250 catalog (in total 31 848 sources), and produced the CID catalog for the COSMOS field, which contains 15 732 sources that are contributing to the IR light. We then matched it with the VLA radio catalog to see if our identified major contributor(s) for a *Herschel* source are in line with the radio source(s) with the same beam. Out of the 2864 VLA sources, 1884 sources are found to be close to the major contributors

identified, among which 935 are found to be contributing more than 50% of the IR fluxes.

We note that as in FLS field, we run the `CIDer` with the help of the MIPS 24 μm data. We also test the performance without the MIPS data, which is the case for a number of the *Herschel* wide fields including HeLMS. The results show that, without MIPS data, the number of matched sources between the VLA catalog and the CID catalog dropped to 1302, and among them, 433 are found to be contributing more than 50% of the IR fluxes. The implies that the MIPS 24 μm data are indeed very important in bridging the optical to the far-IR, in agreement with the previous studies that made use of the MIPS data as priors.

Chapter 7

Co-evolution of Extreme Star Formation and Quasars: Hints from *Herschel* and the Sloan Digital Sky Survey

7.1 Summary

Using the public data from the *Herschel* wide field surveys, we study the far-infrared properties of optical-selected quasars from the Sloan Digital Sky Survey. Within the common area of $\sim 172 \, \mathrm{deg}^2$, we have identified the far-infrared counterparts for 354 quasars, among which 134 are highly secure detections in the *Herschel* $250 \, \mu\mathrm{m}$ band (signal-to-noise ratios ≥ 5). This sample is the largest far-infrared quasar sample of its kind, and spans a wide redshift range of $0.14 \leq z \leq 4.7$. Their far-infrared spectral energy distributions, which are due to the cold dust components within the host galaxies, are consistent with being heated by active star formation. In most cases ($\gtrsim 80\%$), their total infrared luminosities as inferred from only their far-infrared emissions ($L_{IR}^{(cd)}$) already exceed $10^{12} \, \mathrm{L}_\odot$, and thus these objects qualify

as ultra-luminous infrared galaxies. There is no correlation between $L_{IR}^{(cd)}$ and the absolute magnitudes, the black hole masses or the X-ray luminosities of the quasars, which further support that their far-infrared emissions are not due to their active galactic nuclei. A large fraction of these objects ($\gtrsim$ 50–60%) have star formation rates $\gtrsim 300 \, \mathrm{M_\odot \, yr^{-1}}$. Such extreme starbursts among optical quasars, however, is only a few per cent. This fraction varies with redshift, and peaks at around $z \approx 2$. Among the entire sample, 136 objects have secure estimates of their cold-dust temperatures (T), and we find that there is a dramatic increasing trend of T with increasing $L_{IR}^{(cd)}$. We interpret this trend as the envelope of the general distribution of infrared galaxies on the (T, $L_{IR}^{(cd)}$) plane.

7.2 Introduction

It has been noticed ever since the discovery of the ULIRGs by IRAS, that a significant fraction of them, especially those very luminous ones, have optical signatures indicative of classic AGN activities. For examples, Carter (1984) notes that ten among a sample of 13 IRAS sources with 60 μm flux density $f_{60} \geq 1.2 \, \mathrm{Jy}$ are Seyfert galaxies. Sanders et al. (1988) show that ten galaxies among the 324 sources with $f_{60} \geq 5.4 \, \mathrm{Jy}$ from the IRAS Bright Galaxy Survey are ULIRGs and that they all have a mixture of starburst and AGN signatures, which has led them to propose an evolutionary scenario that ULIRGs are the prelude to quasars. In their redshift survey of the IRAS $S_{60} > 0.6 \, \mathrm{Jy}$ galaxy sample, Lawrence et al. (1999) have found that $\sim 20\%$ of their 95 ULIRGs are AGN. The consensus is that ULIRGs that have "warm" IR colors, i.e., whose emissions tend to peak at restframe mid-IR (MIR) rather than far-IR (FIR), generally host (optical) AGN, which could be the main energy sources that power their strong IR emissions (e.g., de Grijp et al., 1985; Osterbrock & De Robertis, 1985; Kim & Sanders, 1998). On the other hand, "cold" ULIRGs that have their IR emis-

sions peak at FIR should be mostly powered by starbursts (e.g., Elston et al., 1985; Heckman et al., 1987).

An interesting question then is whether any AGN ULIRGs, especially quasar ULIRGs, have starbursts that dominate their IR emissions. Quasars represent the most extreme process of supermassive black hole accretion, while starbursts are the most extreme process of star formation. It is expected that the interplay of these two extremes will have important consequences. This is made particularly important by the quasar evolutionary scenario of Sanders et al. (1988), as such objects could be the transitional type between non-quasar ULIRGs and "fully exposed" quasars. Sanders et al. themselves believe that AGN heating is the main mechanism for the strong IR emission of ULIRG quasars. In their discussion of PG quasar continuum distributions from UV to millimeter (mm), Sanders et al. (1989) further propose that a warped galactic disk (beyond the central $\sim 10\,\mathrm{pc}$ to a few kpc) heated by the central AGN can explain the entire range of IR emission from $5\,\mathrm{\mu m}$ to $1\,\mathrm{mm}$. However, Rowan-Robinson (1995) argues that this scheme could only be viable when the total luminosity in IR is comparable or less than that in UV-optical; instead, he has successfully modeled the IRAS-detected PG quasars by attributing their mid-IR emissions to AGN heating and the FIR emission to starburst heating, respectively. More analysis using larger samples from the Infrared Space Observatory (ISO) observations support this view. For example, Haas et al. (2003) have studied 64 PG quasars and have concluded that starburst heating is more likely the cause of the observed cold dust (~ 30–$50\,\mathrm{K}$) FIR emissions among the ULIRG part of the sample. However, they have also pointed out that AGN heating should be the main power source for those extreme ones that qualify as "Hyper-luminous Infrared Galaxies" (HyLIRG; usually defined by $L_{IR} > 10^{13}\,\mathrm{L_\odot}$).

There are also two pieces of important, albeit indirect, evidence supporting that the FIR emissions of quasar ULIRGs are likely due to starbursts. First, a significant fraction of such objects have a large amount of molecular gas (e.g., Sanders et al.,

1988; Evans et al., 2001; Scoville et al., 2003; Xia et al., 2012), which is a strong indicator of active star formation. Second, most quasar ULIRGs have polycyclic aromatic hydrocarbon (PAH) features, which are also strongly indicative of on-going star formation (e.g., Schweitzer et al., 2006; Shi et al., 2007; Hao et al., 2007; Netzer et al., 2007; Cao et al., 2008). While none of these are sufficient to assert that starbursts dominate the strong FIR continua of quasar ULIRGs, it is clear that they at least contribute significantly.

The picture above is largely based on the ULIRGs in the nearby universe where they can be studied in detail. It would not be surprising if any of it changes at high redshifts, as both quasars and ULIRGs evolve strongly. The number density of quasars rises rapidly from $z = 0$ to $z = 1$, and reaches the peak at $z \approx 2$–3 (e.g., Osmer, 2004). Similarly, while ULIRGs are very rare objects today, they are much more numerous in earlier epochs. The deep ISO surveys have revealed a large number of IR-luminous galaxies, among which $> 10\%$ are ULIRGs and many are at $z > 1$ (e.g., Rowan-Robinson et al., 2004). The discovery of the so-called "submillimeter (submm) galaxies" (SMGs) at 450 and 850 µm (see Blain et al. 2002 for a review) added a new population to the ULIRG family, as most of them are at $z \approx 2$–3 and have $L_{IR} > 10^{12} \, L_{\odot}$ dominated by the emission from cold dust. Furthermore, *Spitzer* observations suggest that high stellar mass ($> 10^{11} \, M_{\odot}$) and otherwise "normal" star-forming galaxies at $z \sim 2$ are likely all ULIRGs (e.g., Daddi et al., 2005b), which increases the ULIRG number density at high redshifts to a more dramatic level than expected. On the other hand, molecular gas has also been detected in quasar ULIRGs from $z \gtrsim 1$ to 6 (Solomon & Vanden Bout 2005, and the references therein; see also e.g., Wang et al. 2010, 2011a,b for the recent results at $z \sim 6$), lending support to the starburst-powered interpretation of the FIR emission of such objects at high redshifts as well.

If quasars and dust-enshrouded starbursts do co-exist, it is important to inves-

tigate their co-evolution, which would require a large sample over a wide redshift range. *Herschel* Space Observatory (Pilbratt et al., 2010) has offered an unprecedented opportunity to investigate this problem at the FIR wavelengths that were largely unexplored by the previous studies. The two photometry instruments, PACS and SPIRE, sample the peak of heated dust emission from $z \approx 0$ to 6 and beyond. There already have been a number of studies on the FIR emission of quasars using *Herschel* observations (e.g., Serjeant et al., 2010; Leipski et al., 2010, 2013; Dai et al., 2012; Netzer et al., 2014), however the current collection of quasars that have individual *Herschel* detections are still very scarce in number and few have spanned a sufficient redshift range (for examples, Leipski et al. (2013) present 11 objects at $z > 5$; Dai et al. (2012) include 32 objects at $0.5 \leq z \leq 3.6$; Netzer et al. (2014) report ten within a narrow window at $z \approx 4.8$).

In this project, we present a large sample of optical quasars that are detected by the *Herschel*, and provide our initial analysis of their FIR properties. The quasars are from the Sloan Digital Sky Survey (SDSS; York et al., 2000), and the FIR data are from the public releases of four major wide field surveys by *Herschel*, namely, the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al., 2010), the Herschel Multi-tiered Extragalactic Survey (HerMES; Oliver et al., 2012), the Herschel Stripe 82 Survey (Viero et al., 2014, HerS;), and the PACS Evolutionary Probe (PEP; Lutz et al., 2011). We describe the data and the sample construction in § 7.3, and present our analysis of the FIR dust emission in § 7.4. The implications of our results are detailed in § 7.5, and we conclude with a summary in § 7.6. The catalog of our sample is available as online data in its entirety. All quoted magnitudes are in the AB system. We adopt the following cosmological parameters throughout: $\Omega_M = 0.27$, $\Omega_\Lambda = 0.73$ and $H_0 = 71\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.
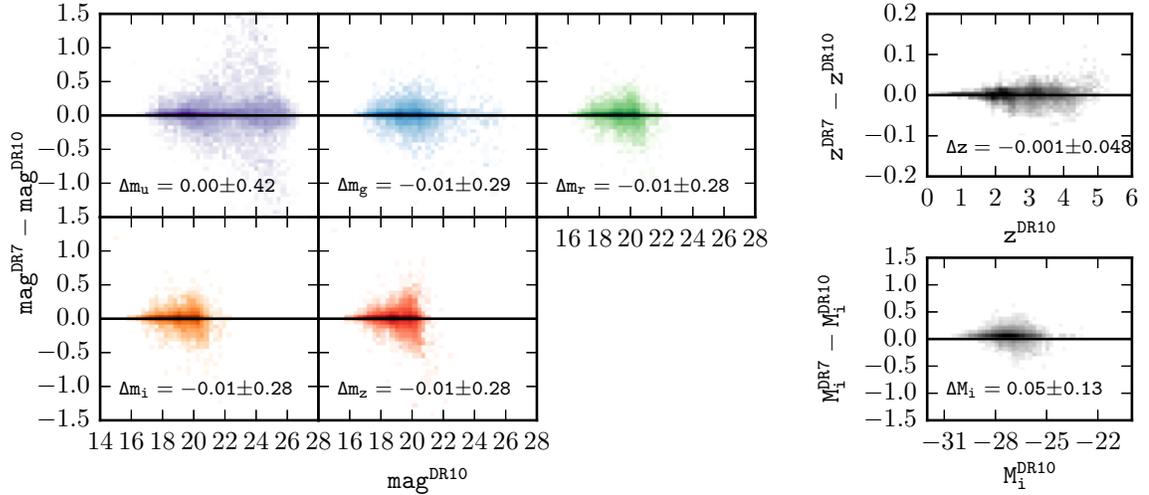
Figure 7.1: Comparison of the photometry, the redshifts and the absolute magnitudes from the DR7Q and the DR10Q for the overlapped population in these two quasar catalogs (16 356 objects in total). For this subset, we adopt the DR10Q values unless noted otherwise. For the sake of consistency, we use the absolute magnitudes k-corrected to $z = 2$ for the DR7Q quasars as well, adapting the values from the catalog presented in Shen et al. (2011).

7.3 Data Description and Sample Construction

In brief, we built our sample by searching for the counterparts of the SDSS quasars in the *Herschel* wide field survey data. For the sake of simplicity, hereafter we refer to these objects as "IR quasars". We describe below the data used in our study and the constructed IR quasar sample.

7.3.1 Parent Quasar Samples

The parent quasar samples that we used are based on the SDSS Data Release 7 and 10 quasar catalogs (hereafter DR7Q and DR10Q, respectively), which are summarized as follows.

DR7Q As detailed in Schneider et al. (2010), this quasar catalog is based on the SDSS DR7. It concludes the quasar survey in the SDSS-I and SDSS-II over $9380\,\mathrm{deg}^2$,
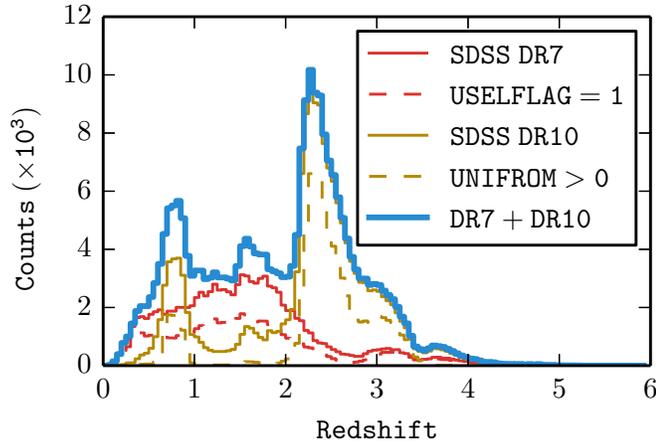
Figure 7.2: Redshift distribution of the SDSS DR7Q (red curves) and DR10Q catalogs (yellow curves). The solid ones are for the entire catalogs, while the dashed ones are for the "homogeneous" subsamples as described in Schneider et al. (2010) and Pâris et al. (2014).

and supersedes all previously released SDSS quasar catalogs. It includes 105 783 quasars between $z = 0.065$ and 5.46 (the median at $z = 1.49$), all with absolute i-band magnitudes (M_i) brighter than -22 mag.

DR10Q This quasar catalog is derived from the on-going Baryon Oscillation Spectroscopic Survey (BOSS) as part of the SDSS-III. While it was released in the SDSS DR10, its main target selection was based on the SDSS DR8. The detailed description of the catalog can be found in Pâris et al. (2014). It includes new quasars with $M_i[z = 2] < -20.5$ mag from the SDSS-III, where $M_i[z = 2]$ is the absolute magnitude k-corrected to $z = 2$ (for details, see Pâris et al., 2014). It also includes a large number of known quasars of similar characteristics (mostly from SDSS-I and II) that were re-observed by BOSS. In brief, the catalog contains 166 583 quasars over 6373 deg^2, with redshifts ranging from 0.05 to 5.86.

The quasars from these two catalogs are largely independent, however there are still 16 356 of them being duplicates, which we define as the ones falling within a matching radius of 0.4″. Figure 7.1 shows the comparison of the photometry, the

redshift measurements and the absolute magnitudes from these two catalogs for this overlapped population, which all agree reasonably well. For the sake of simplicity, we adopt the DR10Q values in this work for these duplicates unless noted otherwise.

In the end, we produced a merged sample of 256 010 unique quasars, which represents the largest optical quasar sample selected based on the most homogeneous data set over the widest area to date. The redshift distributions of DR7Q, DR10Q and the merged catalog are shown in Figure 7.2. For simplicity, hereafter we refer to the merged sample as the "SDSS quasar sample" and the quasars therein as the "SDSS quasars". We note that DR7Q and DR10Q are statistically different samples, and therefore any statistical results from this merged catalog should be inferred with caution. This is exacerbated by the fact that the quasars were not selected uniformly within either DR7Q or DR10Q, as detailed in Schneider et al. (2010) and Pâris et al. (2014), respectively. For example, while most of the quasars in DR7Q were selected using the algorithms as described in Richards et al. (2002) (with USELFLAG=1 in the catalog), a non-negligible fraction of them were selected early in the SDSS campaign when such algorithms had not yet been fully developed. DR10Q is even more non-uniform in this sense, because only about half of its quasars (called the "CORE" sample) were selected uniformly (with UNIFORM>0 in the catalog) through the XDQSO method (Bovy et al., 2011) and the other half (called the "BONUS" sample) were selected using various different methods (see Ross et al., 2012). Nevertheless, the non-uniformity does not affect our current work since we limit our study to the FIR properties of optical-selected quasars, whose being selected did not use any FIR information and thus should not favor or against any given FIR property.

7.3.2 *Herschel* Data

We used the HerMES, H-ATLAS, and HerS catalogs to match with the SDSS catalog. The details is presented in § 2.1. In this work, we adopted the band-merged catalogs
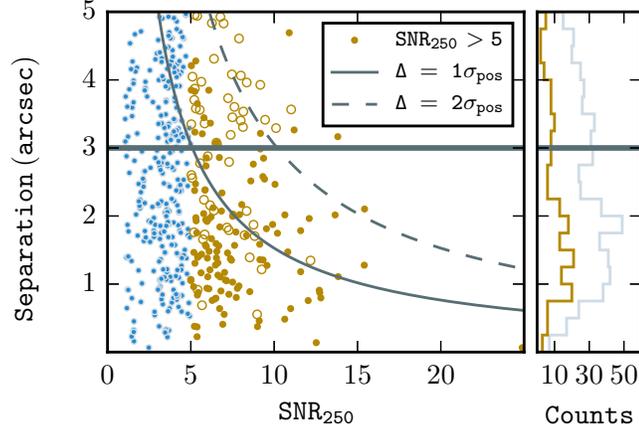
Figure 7.3: Distribution of the separation between the SDSS positions and the matched SPIRE 250 μm positions when expanding the matching radius to 5″. The left panel shows the separation as a function of the SNR of the 250 μm detection. The solid and the dashed lines are the theoretical $\sim 1\,\sigma$ and $\sim 2\,\sigma$ positional uncertainties (σ_{pos}). The circles in yellow color are the sources with $\mathrm{SNR}_{250} > 5$, while the light blue ones represent the rest. Among the yellow circles, the open and the solid ones are those deemed by our visual inspection to be "blended" (i.e., affected by source blending) and "clean" (free of source blending), respectively. The horizontal thick black line going through both panels indicates the adopted matching radius of 3″.

using the 250 μm positions (xID250).

7.3.3 *Herschel*-detected Quasars

Our IR quasar sample was derived by matching the positions of the SDSS quasars to those of the *Herschel* sources in the band-merged, 250 μm-based catalogs.

Matching Radius

The matching was performed using TOPCAT/STILTS[1]. We used a matching radius of 3″, which is justified below.

While the *Herschel* instruments have large beam sizes[2], the source centroids can

---

[1]TOPCAT http://www.starlink.ac.uk/topcat;
STILTS http://www.starlink.ac.uk/stilts

[2]The FWHM beam sizes are 18″, 25″ and 36″ for the SPIRE 250, 350 and 500 μm, respectively, and 6–7″ and 11–14″ for the PACS 100 and 160 μm, respectively.

still be determined to high accuracy. The positional uncertainty of a given source depends on its signal-to-noise ratio (SNR) (see, e.g., Ivison et al., 2007), which follows

$$\Delta\alpha = \Delta\delta = \frac{0.6\theta}{\text{SNR}}\,,\tag{7.1}$$

where $\Delta\alpha$ and $\Delta\delta$ are the nominal $1\,\sigma$ uncertainties of RA and Dec, respectively, and θ is the beam size. The SPIRE $250\,\mu\text{m}$ beam size is $\theta = 18''$, which means that the $1\,\sigma$ positional uncertainty of a given $250\,\mu\text{m}$ source is

$$\sigma_{pos} = \sqrt{(\Delta\alpha)^2 + (\Delta\delta)^2} = \frac{15.27''}{\text{SNR}}\,.\tag{7.2}$$

A matching radius of $3''$ thus corresponds to $\sim 1\,\sigma$ uncertainty for objects with $\text{SNR} > 5$, or $\sim 0.59\,\sigma$ for those with $\text{SNR} > 3$[3]. To further validate our choice, we performed a test using an enlarged matching radius of $5''$, and Figure 7.3 demonstrates the results. The separations between the $250\,\mu\text{m}$ and the SDSS positions versus the $250\,\mu\text{m}$ SNR are indeed consistent with the expectation from Equation (7.2), and the vast majority of the matches within $3''$ fall within the $1\sigma_{pos}$ curve. Furthermore, the separation shows a double-peak feature roughly divided at $3''$, indicating that the matches beyond this point are likely affected by other factors, such as the source blending problem (see § 1.4).

In summary, our conservative choice of the matching radius has resulted in a sample that is free of significant blending problem. This has also simplified the photometry in the SPIRE 350 and $500\,\mu\text{m}$ bands. While the source catalogs that we adopted from the survey teams were all derived using PSF-fitting based on the $250\,\mu\text{m}$ detections, the photometry in 350 and $500\,\mu\text{m}$ would still be prone to large

---

[3]We note that Smith et al. (2011) use a somewhat different relation between the astrometric uncertainty and the SNR, $\sigma_{pos} = 0.655\theta/\text{SNR}$, which means that $1\,\sigma$ uncertainty would be $2''.4$ for $\text{SNR} = 5$ in $250\,\mu\text{m}$. Our results shown in Figure 7.3 indicate that such a matching radius would be slightly too stringent, and therefore we adhered to our choice.

errors introduced by blending due to their much larger beam sizes ($25''$ and $36''$, respectively). Our relatively clean sample allows us to assume that the fluxes measured in these two redder bands are solely contributed by the source detected in $250\,\mu m$.

IR Quasar Sample Summary

Our *Herschel*-detected SDSS quasar sample, derived using a matching radius of $3''$ as described above, contains 354[4] objects in total. Most of our studies are based on a subsample of those, called the "SNR5" sample, which only includes 134 quasars that have SNR ≥ 5 in $250\,\mu m$ and thus is the sample of the most robust *Herschel* detections. Based on this SNR5 sample, we further imposed a cut in the $250\,\mu m$ flux density, $S_{250} \geq 56.6\,\mathrm{mJy}$, and formed a "bright" subsample of 82 objects for various discussions below. This flux density threshold was adopted based on the lowest flux density of the SNR5 objects in XMM-LSS-SWIRE field.

7.4 Dust Emission Modeling

We inferred the dust emission properties of the IR quasars by fitting their FIR SEDs, following the recipes outlined in § 5.2. The thermal emission over the full IR regime can be viewed as the collective result of all heated dust components of various temperatures, and the FIR part is dominated by the coldest component. In this project, we focus on this FIR part in the IR quasars and hence our conclusions are pertaining to their cold-dust components. We used two distinct types of models to fit the FIR SEDs, one being a single-temperature, modified blackbody (MBB) spectrum and the other being three different sets of starburst templates. Fitting an MBB spectrum to the FIR SED is always valid, regardless of the exact dust heating sources (i.e., due to photons from either star formation or AGN activity). However, it has the

---

[4]We note that there are 18 more matched objects from HerMES DR2 catalogs discarded due to $\mathrm{SNR}_{250} < 1$

drawback that the SED must be properly sampled in order to obtain well constrained results. Fitting starburst templates, on the other hand, is only appropriate if the FIR emission is dominated by heating from star formation, and the motivation of using these models was to test if the FIR emissions are consistent with being caused by star formation. These two types of SED fitting approach provide independent check to each other, and we will show later that they also lead to insights into the heating sources.

7.5   Results and Discussions

The catalog that contains all the major physical properties obtained from the IR SED fitting are published online and a summary of the information is given in Table 7.1. In this section, we highlight the most important results and their implications.

7.5.1   IR Luminosity

In this project, we perform two flavors of fittings: one is to fit a single temperature MBB model, and the other is to fit to the empirical IR SED templates.

The IR luminosities derived from the fitting are summarized here:

- L_{IR}: the general designation of the total IR luminosity over 8–1000 μm;

- $L_{IR}^{(cd)}$: the contribution of the cold-dust component to the total IR luminosity over 8–1000 μm, which is also referred to as the total IR luminosity of the cold-dust component;

- L_{IR}^{mbb}: the MBB best-fit to the FIR SED (as represented by the SPIRE data points) integrated over 8–1000 μm, and by definition $L_{IR}^{(cd)} \equiv L_{IR}^{mbb}$;
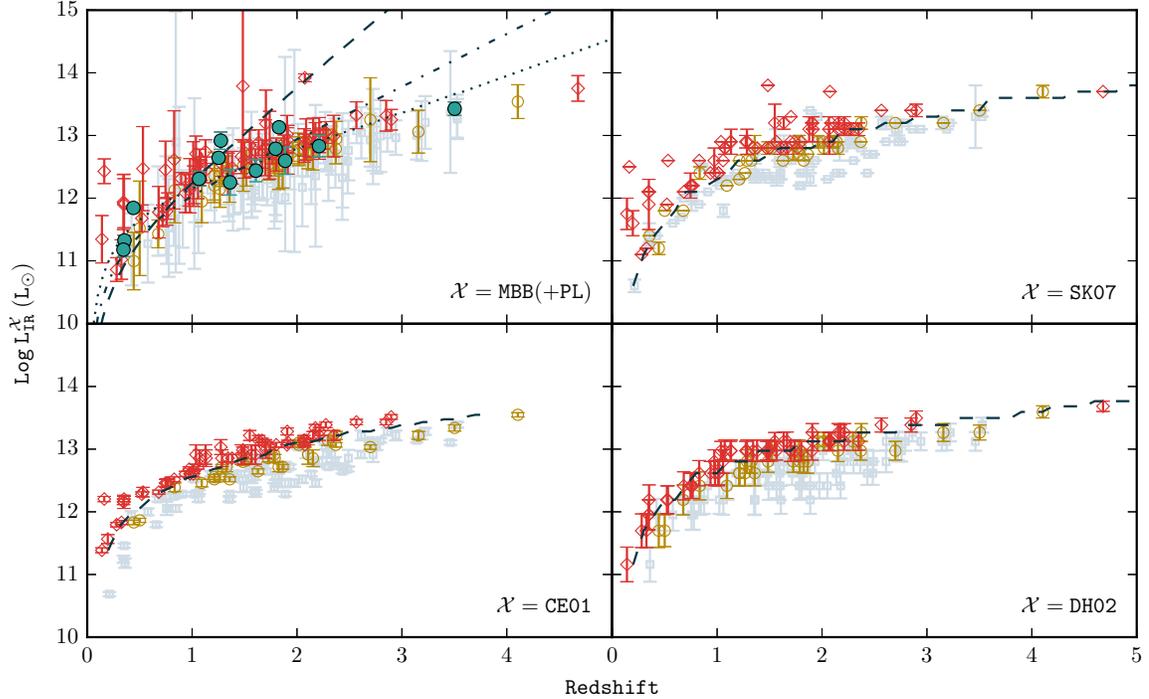
Figure 7.4: Derived IR luminosities ($L_{IR}^{\mathcal{X}}$) of the IR quasars in our sample, where "$\mathcal{X}$" denotes one of the four models in use, namely, MBB (+PL), SK07, CE01, and DH02. The errors between the MBB results and those from the starburst templates are not directly comparable because of the different approaches adopted in evaluating the errors. The colored symbols represent the SNR5 sample, while the grey squares represent the rest. Among the SNR5 objects, the red diamonds indicate those that are in the bright subsample ($S_{250} \geq 56.6\,\mathrm{mJy}$), while the yellow circles indicate those that are not. The dark green solid circles in the MBB panel are the objects with PACS data available and hence a power-law component was added to the MBB spectrum in the fitting. To illustrate the impact of the survey limit, the limits of $L_{IR}^{\mathcal{X}}$ corresponding to the fiducial flux density limit of $S_{250} = 56.6\,\mathrm{mJy}$ are shown as the dashed lines in the three panels for the starburst template fits (SK07, CE01 and DH02). In each of these cases, the limit is derived from the entire library by using the template with the lowest possible $L_{IR}^{\mathcal{X}}$ at a given redshift. In the panel for the MBB fit, the limits are given using three different T_{mbb} of 15, 25, and 35 K, shown as the dashed, the dot-dashed and the dotted lines, respectively.

- L_{IR}^{mbb+pl}: the MBB+PL best-fit to the FIR SED (as represented by the SPIRE and the PACS data points) integrated over 8–1000 μm, which is the measurement of L_{IR} using the MBB+PL model;

- L_{IR}^{SB}: the measurement of L_{IR} using the starburst models ("SB" is one of

110
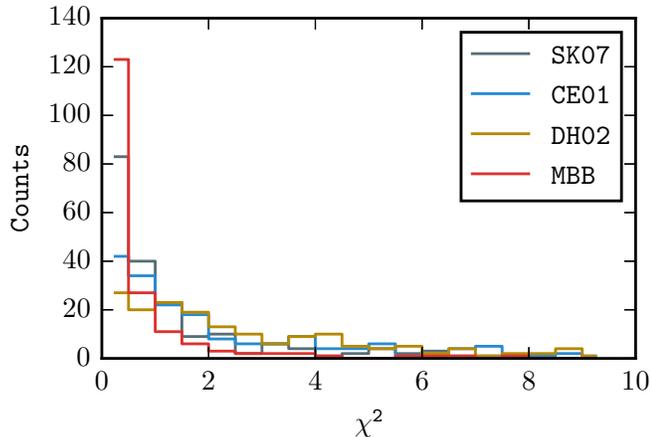
Figure 7.5: Distribution of χ^2 of the best-fit models for all the IR quasars in our sample, using the four different methods as detailed in the text.

"SK07", "CE01" and "DH02", depending on the model set in use), and effectively is the best-fit SB template integrated over 8–1000 μm.

For one of our purposes later, we also define the corresponding quantities in the FIR instead of over the entire IR range, i.e., by integrating over 60–1000 μm only. We designate these quantities with the subscript of "FIR", e.g., $L_{FIR}^{(cd)}$, L_{FIR}^{mbb}, L_{FIR}^{SB}, etc.

The derived L_{IR}^{mbb} and L_{IR}^{SB} (referred to as $L_{IR}^{\mathcal{X}}$) values of our IR quasars are shown in Figure 7.4, and the distributions of the best-fit χ^2 are shown in Figure 7.5, respectively. We note that deriving L_{IR}^{mbb} (and hence $L_{IR}^{(cd)}$) was not always possible because the MBB fit requires the FIR SED being well sampled by the three SPIRE bands, while obtaining L_{IR}^{SB} could always be done because of the nature of the method (see § 5.2.1 & 5.2.2). We also note that the SNR5 subsample, as expected, has the smallest errors in $L_{IR}^{\mathcal{X}}$. In addition, the majority of the objects outside of the SNR5 sample still have $\chi^2 \leq 10$ and thus are also deemed as having reliable $L_{IR}^{\mathcal{X}}$ measurements.

Regardless of the exact model set in use, the majority of our IR quasars have obtained good fits and the derived IR luminosities also agree to the extent that we expect. This is more clearly shown in the upper panels of Figure 7.6, where L_{IR}^{SB} are
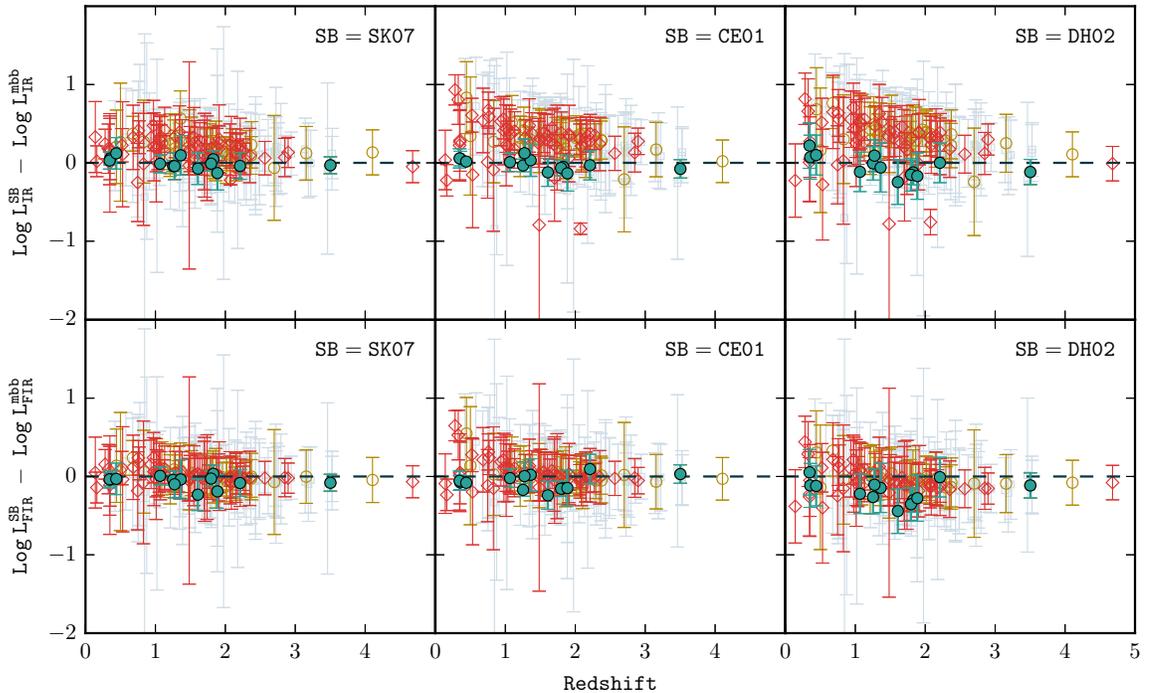
Figure 7.6: Comparisons of IR luminosities derived using the MBB model and those based on the three sets of starburst templates. The symbols are the same as in Figure 7.4. The upper panels show the comparisons to L_{IR}^{SB} (i.e., computed over the full IR range of 8–1000 μm), while the bottom panels show the comparison to L_{FIR}^{SB} (computed over the FIR regime of 60–1000 μm), where "SB" is one of SK07, CE01 and DH02. See § 7.5.1 for details.

compared against L_{IR}^{mbb}. On average, L_{IR}^{mbb} values are lower by 0.13, 0.23 and 0.25 dex as compared to L_{IR}^{SB} from the SK07, the CE01 and the DH02 models, respectively. As explained earlier, this is due to the fact that the MBB model only includes the cold-dust component ($L_{IR}^{mbb} \equiv L_{IR}^{(cd)}$), whereas the starburst models also include all other components of higher temperatures and thus give the total IR luminosity (L_{IR}). This is also demonstrated by the 13 objects that have PACS data (dark green points), for which we carried out the MBB+PL fit and thus obtained the total IR luminosity in the form of L_{IR}^{mbb+pl}. As one can see, there are no offsets between L_{IR}^{mbb+pl} and L_{IR}^{SB}.

To further demonstrate this point, the lower panel of Figure 7.6 shows similar comparisons as in the upper panel, but between L_{FIR}^{mbb} and L_{FIR}^{SB}. The agreements are excellent and no systematic offsets are found. This can be understood as fol-

112

lows. The emission from the hot-dust components should be minimal in the FIR regime, and hence integrating the starburst models in the FIR should only capture the contribution from the cold-dust component, i.e., $L_{FIR}^{SB} = L_{FIR}^{(cd)} = L_{FIR}^{mbb}$.

As we have been emphasizing, the MBB fit is independent of the heating source, while the SB fit being valid hinges upon the heating source being star formation. Therefore, the agreement between L_{FIR}^{mbb} and L_{FIR}^{SB} strongly suggests that the SB fits are valid and that the FIR emissions in these IR quasars are due to star formation. The corollary then is that L_{IR}^{mbb} ($\equiv L_{IR}^{(cd)}$) is due to star formation, because it is the total luminosity of the same cold-dust component that gives rise to the FIR emission. While L_{IR}^{mbb} could underestimate the true L_{IR} (see the top panel of Figure 7.6) by a factor of 1.35 (as compared to L_{IR}^{SK07}) to 1.70–1.78 (as compared to L_{IR}^{CE01} or L_{IR}^{DH02}), we prefer to be conservative due to the lack of observational constraints in the mid-IR for our entire sample.

Our IR quasars have $L_{IR}^{\mathcal{X}}$ values ranging from $\sim 10^{10.5}$ to $10^{13.8}\,\mathrm{L}_{\odot}$ (after discarding two objects whose SEDs are barely constrained). Most of them ($\gtrsim 80\%$) are ULIRGs ($L_{IR} > 10^{12}\,\mathrm{L}_{\odot}$), and some of them ($\gtrsim 15$–23%) are even HyLIRGs ($L_{IR} > 10^{13}\,\mathrm{L}_{\odot}$). As Figure 7.4 indicates, there is a trend of $L_{IR}^{\mathcal{X}}$ versus redshifts. Obviously, the lack of IR quasars with low $L_{IR}^{\mathcal{X}}$ at high redshifts is caused by the selection effect due to the survey limit. For illustration, Figure 7.4 shows the $L_{IR}^{\mathcal{X}}$ selection limit corresponding to a $250\,\mu\mathrm{m}$ flux density limit of $S_{250} = 56.6\,\mathrm{mJy}$ (which is what we adopted to select the bright subsample from the SNR5 sample). Interestingly, there seems to be a deficit of very luminous IR quasars at $z < 1$, which reflects a genuine IR luminosity evolution that is broadly consistent with the evolution of ULIRGs, i.e., there are more ULIRGs at $z > 1$ than at lower redshifts.

Our conclusion that the FIR emission of IR quasars are powered by star forming activity in dust-rich environments has also been suggested by previous studies at high redshifts (e.g., Wang et al., 2011b). If this is indeed the case, using the standard

113
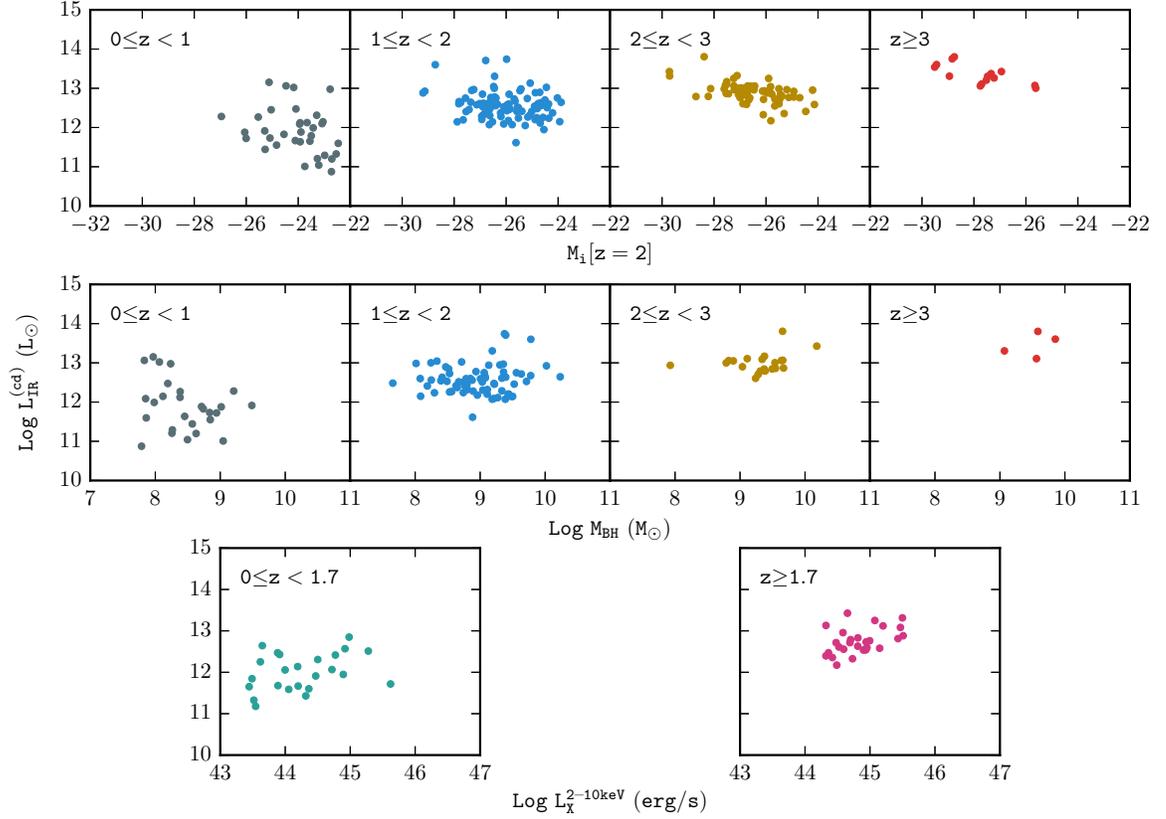
Figure 7.7: No correlation between $L_{IR}^{(cd)}$ and the absolute magnitudes of quasars (upper panels), their black hole masses (middle panels), or the X-ray luminosities (lower panels). The results are shown in four redshift bins as labeled, and the error bars are omitted for clarity.

L_{IR} to SFR conversion of Kennicutt (1998), i.e., $SFR_{IR} = 1.0 \times 10^{-10} L_{IR}/L_\odot$ for a Chabrier initial mass function (IMF)[5], the SFR of the HyLIRGs in our sample would be ~ 1.0–$6.3 \times 10^3 \, M_\odot \, yr^{-1}$[6].

There might be concerns whether such extreme SFRs are physical. The SFR could indeed be overestimated in two ways. First, one could argue that AGN heating is still an important contributor to $L_{IR}^{(cd)}$ and hence the SFR cannot be calculated without subtracting this contribution. While there is no viable model quantitatively showing that this could be the case (in fact, all available models assume the opposite), we

---

[5]The conversion would be a factor of ~ 1.7 higher if using a Salpeter IMF, which was adopted in Kennicutt (1998).

[6]The most conservative SFR estimates would be using L_{FIR}^{mbb} (over 60 to 1000 μm) instead of L_{IR}^{mbb} (over 8 to 1000 μm), which would reduce the SFR values by a factor of ~ 1.5.

cannot yet assert that this is impossible. Our argument of $L_{FIR}^{mbb} = L_{FIR}^{SB}$ presented earlier is only a necessary condition that $L_{IR}^{(cd)}$ is due to the heating from star formation but not a sufficient one, and therefore we cannot rule out such a possibility based on this argument alone. However, we can demonstrate that AGN heating is very unlikely dominant in $L_{IR}^{(cd)}$. If it is dominant, it is expected that $L_{IR}^{(cd)}$ should be positively correlated with the AGN activity, i.e., the stronger the AGN is, the larger $L_{IR}^{(cd)}$ should be. Figure 7.7 shows $L_{IR}^{(cd)}$ versus the quasar absolute i-band magnitude (normalized to $z = 2$, adapted from Shen et al. (2011) and Pâris et al. (2014) for DR7 and DR10, respectively, and all are based on PSF magnitudes after the Galactic extinction correction) in four redshift bins. Apparently, no such a correlation can be seen. In the lowest redshift bin, the distribution of the objects is completely chaotic. In the bins at higher and higher redshifts, we tend to see only those objects that are more and more IR luminous, which is simply due to the selection effect in a flux-limited survey. Even among these the most luminous ones, no correlation among $L_{IR}^{(cd)}$ and M_i can be vouched for. Furthermore, Figure 7.7 also shows $L_{IR}^{(cd)}$ versus the back hole mass (M_{BH}) for the quasars that have these estimates (taken from Shen et al. (2011) for the DR7Q quasars). Similarly, no correlation exists. Finally, the bottom panels of Figure 7.7 show $L_{IR}^{(cd)}$ versus the hard-band X-ray luminosity in the restframe 2–10 keV ($L_X^{2-10\,keV}$) for a limited number of quasars that we can derive this quantity based on the data available in the literature[7]. As there are only 40 such quasars, we split them into two redshift bins, $0 < z \leq 1.7$ and $z > 1.7$, respectively, so that each bin receives approximately the same number of objects for statistics (19 and 21 objects, respectively). They all have $L_X^{2-10\,keV} > 10^{43}$ erg/s, which is well above the conventional X-ray AGN selection threshold of 10^{42} erg/s, above which the

---

[7]$L_X^{2-10\,keV}$ were derived based on the data from the Chandra Source Catalog Release 1 (Evans et al., 2010) and the 3XMM-DR5 catalog (Rosen et al., 2015). Briefly, a power-law in the form of $I_\nu \propto \nu^{-\alpha}$ was fit to the flux densities at different energy bands, and the total energy in restframe 2–10 keV was calculated by integrating the best-fit power-law over this energy range. The best-fit α has a median of ~ 0.7, which correspond to the photon index $\Gamma \sim 1.7$.

X-ray luminosity is believed to be predominantly due to AGN. Therefore, $L_X^{2-10\,\mathrm{keV}}$ is a strong indicator of the AGN activity. Again, no correlation between $L_{IR}^{(cd)}$ and $L_X^{2-10\,\mathrm{keV}}$ can be seen. This is also very consistent with the recent results of Symeonidis et al. (2014) and Azadi et al. (2015) in the similar $L_X^{2-10\,\mathrm{keV}}$ regime.

Therefore, while we do not have direct evidence to assert that AGN has no contribution to $L_{IR}^{(cd)}$, we do have evidence (albeit still indirect) against that AGN contribution can be dominant. This further strengths our conclusion of $L_{IR}^{(cd)}$ being due to star formation based on the earlier argument of $L_{FIR}^{mbb} = L_{FIR}^{SB}$.

The other possibility is that the most luminous objects are actually gravitationally lensed, which means that their intrinsic luminosities must be lower and so are their SFR estimates. Currently, we do not have further data to address this issue. In §7.5.2, however, we will show that it is also unlikely that the most luminous objects are predominantly the result of lensing.

7.5.2 Relation between Dust Temperature and IR Luminosity

Here we examine the T_{dust}-$L_{IR}^{(cd)}$ relation of IR quasars. For simplicity, we use T_{mbb} in this discussion, and only include the 134 objects that have good estimates of both $L_{IR}^{(cd)}$ ($\chi^2 < 10$) and T_{mbb} ("SNR"> 3). This is shown in Figure 7.8, where the crosses are the individual objects and the red circles represent the average at a given $L_{IR}^{(cd)}$ (step-size of 0.2 dex in Log $L_{IR}^{(cd)}$). We also plot the mean result from Symeonidis et al. (2013, light green circles), who have analyzed a sample of IR luminous ($L_{IR} > 10^{10}\,\mathrm{L_\odot}$) galaxies at $0.1 < z < 2$ using the deep PACS and SPIRE data in the COSMOS, the GOODS-N and the GOODS-S fields. Symeonidis et al. (2013) find that their L_{IR}-T relation[8] has only a modest increasing trend towards high luminosities, and their interpretation is that the increase of L_{IR} is caused by an increase in the dust mass

---

[8]We note that Symeonidis et al. (2013) adopt $\lambda_0 = 100\,\mu\mathrm{m}$ and $\beta = 1.5$ as we do here.
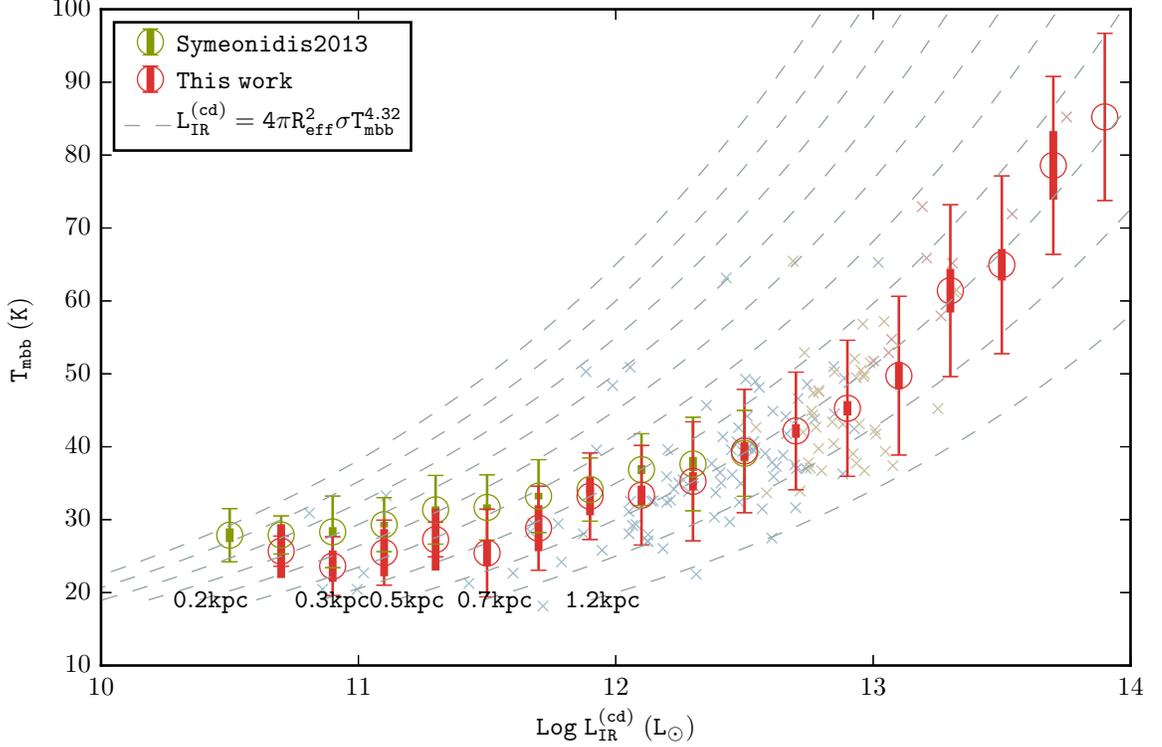
Figure 7.8: Strong T_{mbb}-$L_{IR}^{(cd)}$ relation as inferred from our IR quasars that have reliable T_{mbb} and $L_{IR}^{(cd)}$ measurements. The light grey crosses represent the individual quasars (134 objects in total), while the big red circles show their mean T_{mbb} values in $L_{IR}^{(cd)}$ bins of 0.2 dex. The thick error bars on these red circles are the standard deviations around the mean values, while the thin error bars are the summed (in quadrature) errors of all objects in each bin. For comparison, the results from (Symeonidis et al., 2013) are shown as the green symbols. The dashed lines represent the MBB equivalent of Stefan-Boltzmann law as derived based on Equation (5.2), using a family of R_{eff} values as labeled.

and/or the IR emitting radius rather than by an increase in the intensity of the dust heating radiation field.

However, the picture is rather different for our sample because our IR quasars span a much wider range in luminosity. Our T_{mbb}-$L_{IR}^{(cd)}$ relation agrees reasonably with that of Symeonidis et al. (2013) in the overlapped, low luminosity range ($L_{IR}^{(cd)} \lesssim 10^{12}\,L_{\odot}$), and then dramatically rises to higher luminosities. The existence of such a relation is against the possibility that our sample could be significantly affected by gravitational lensing, because the magnification cannot be correlated with the dust temperature

(see §7.5.1).

We argue that this T_{mbb}-$L_{IR}^{(cd)}$ relation cannot be attributed to the selection effect of our sample. To demonstrate this point, we simulated a large number of objects of different T_{mbb} and $L_{IR}^{(cd)}$ over the redshift range of our sample, and recovered them using various selection criteria in T_{mbb} and S_{250}. Figure 7.9 shows the results in four redshift bins, for two different T_{mbb} "SNR" thresholds of three and five, and two different S_{250} thresholds of 1 and 50 mJy, respectively. The key points can be summarized as follows. First, adopting a higher T_{mbb} "SNR" (e.g., $T_{mbb} > 5T_{mbb}^{err}$ instead of $T_{mbb} > 3T_{mbb}^{err}$) would be against objects with high T_{mbb}. Second, adopting a higher S_{250} threshold (e.g., $S_{250} > 50$ mJy instead of $S_{250} > 1$ mJy) would be against objects with low $L_{IR}^{(cd)}$. To reiterate, our current work adopts $T_{mbb} > 3T_{mbb}^{err}$. While our sample does not have a uniform S_{250} threshold due to the varying survey limits in different fields, our objects all have $S_{250} > 50$ mJy. From Figure 7.9, one can see that the objects that have highest probability of being selected by our criteria would be those with higher $L_{IR}^{(cd)}$ and lower T_{mbb} than our data points, however such objects are not presented in our sample, i.e., there is a genuine lack of such objects.

Before discussing the lack of objects with (high $L_{IR}^{(cd)}$, low T_{mbb}), let us first understand the increasing trend of T_{mbb} with increasing $L_{IR}^{(cd)}$. Recall that for a perfect black body, Stefan-Boltzmann law states that $L = 4\pi R^2 \sigma T^4$. Motivated by this, we integrated Equation (5.2) and found that the equivalent for a general opacity MBB should follow $L_{IR}^{(cd)} = 4\pi R_{eff}^2 \sigma T_{mbb}^{4.32}$, where R_{eff} of a given galaxy should be interpreted as the effective radius of the equivalent FIR emitting region if we combine together *all* its dust-enshrouded star-forming regions. As shown in Figure 7.8, our data points can be explained by this relation with a family of R_{eff}. At $L_{IR}^{(cd)} \lesssim 10^{12}$ L$_\odot$, the increasing in $L_{IR}^{(cd)}$ is mostly dominated by the increasing in R_{eff}, which range from ~ 0.1 to 0.5 kpc, and the increasing in T_{mbb} only plays a modest role. This is consistent with the suggestion of Symeonidis et al. (2013) as summarized earlier. Naturally, R_{eff}
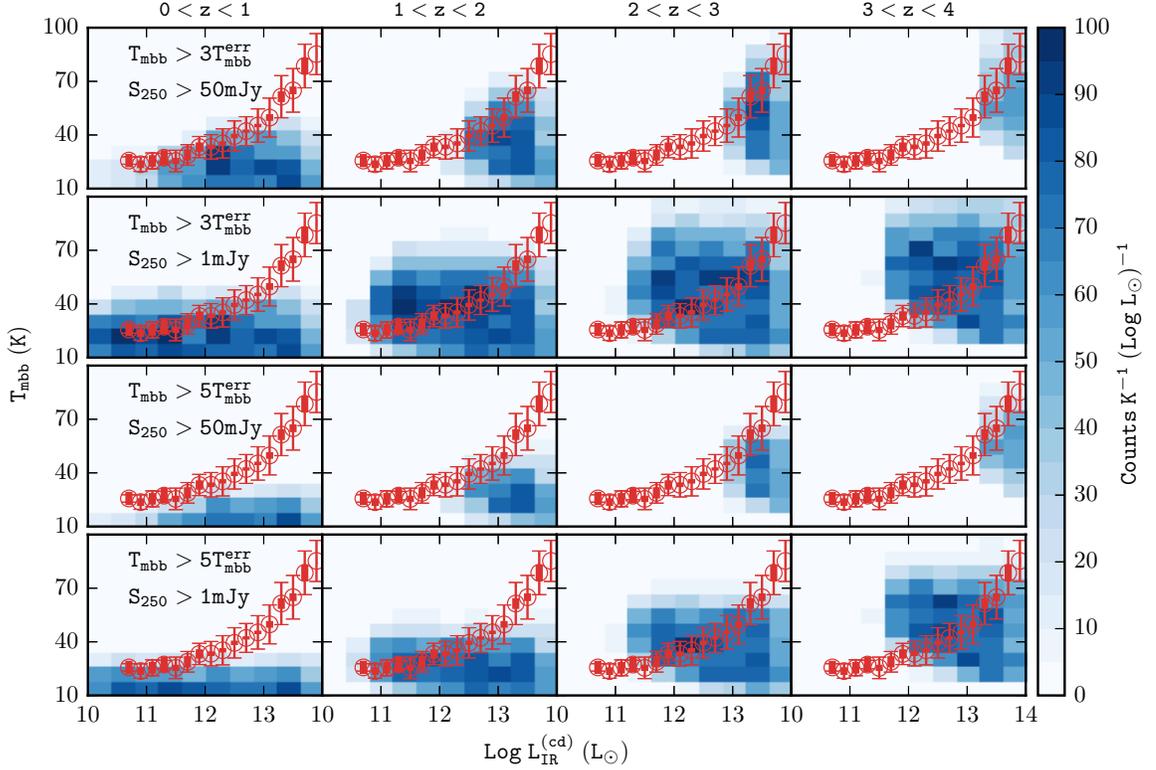
Figure 7.9: Simulation results showing that the observed T_{mbb}-$L_{IR}^{(cd)}$ relation is not due to selection effect. As labeled in each panel, the results are given in four redshift bins (in column) for the combinations of two T_{mbb} thresholds and two S_{250} thresholds (in row). The red symbols are the same T_{mbb}-$L_{IR}^{(cd)}$ relation as in Figure 7.8, while the blue blocks in the background represent the densities of the simulated objects recovered by the labeled selection thresholds (the coding of the color depth is shown to the right). Under our current sample selection (the first row), the highest densities of the simulated objects all occur below the T_{mbb}-$L_{IR}^{(cd)}$ relation, i.e., in the regions of higher $L_{IR}^{(cd)}$ and lower T_{mbb} than those defined by the T_{mbb}-$L_{IR}^{(cd)}$ relation. In other words, while the objects of higher $L_{IR}^{(cd)}$ and lower T_{mbb} would have the highest probability of being included by our selection, they do not present in our sample, i.e., there is a genuine lack of such objects in the universe. On the other hand, our selection (mostly the S_{250} threshold due to the survey limit) is against the objects of higher T_{mbb}. Therefore, it is most likely that the T_{mbb}-$L_{IR}^{(cd)}$ relation, discovered among IR quasars, is the envelope of the general distribution of IR galaxies.

cannot be increased indefinitely because the sizes of galaxies are finite. Our data suggest that R_{eff} reaches its maximum of ~ 0.5–0.7 kpc at the ULIRG luminosity. At $L_{IR}^{(cd)} \gtrsim 10^{12}$ L$_\odot$, the increasing in $L_{IR}^{(cd)}$ is taken over by the increasing of T_{mbb}, which can be due to more intense radiation field caused by more intense starburst

activity. The lack of (high $L_{IR}^{(cd)}$, low T_{mbb}) objects thus is the result of the limit in R_{eff}.

This also suggests that the T_{mbb}-$L_{IR}^{(cd)}$ relation as seen in Figure 7.8 for IR quasars is an envelope of the general distribution of IR-luminous objects on the $(T_{mbb}, L_{IR}^{(cd)})$ plane. In other words, for a given $L_{IR}^{(cd)}$, the dust temperature reached in IR quasars is the lowest among all possibilities in IR-luminous objects. In fact, the data points of Symeonidis et al. (2013) indeed are above ours in Figure 7.8, which is perfectly consistent with our interpretation. However, we do not have an explanation on why this envelope manifests itself in IR quasars.

### 7.5.3 Dust Mass and Gas Mass
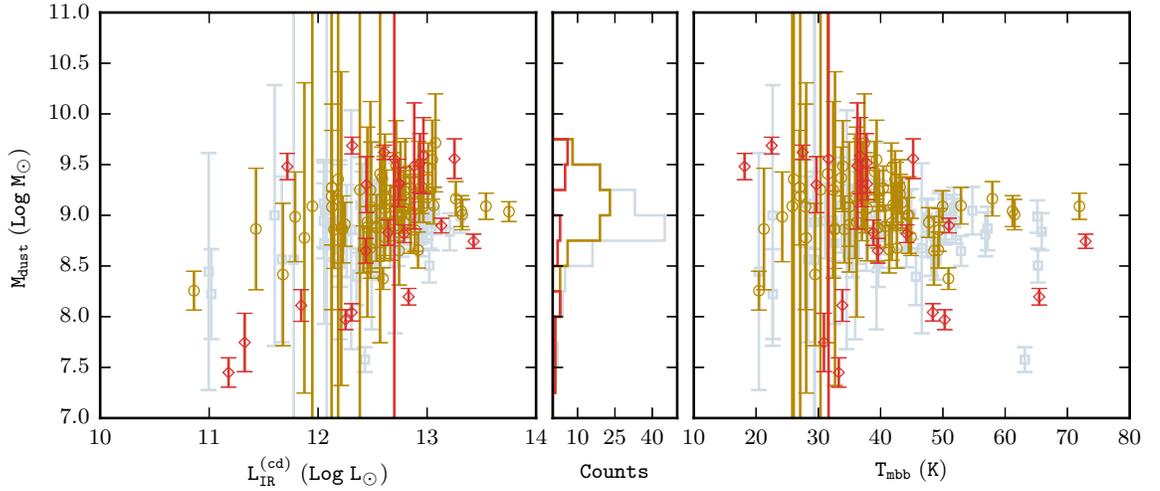
Figure 7.10: Distribution of dust mass (M_{dust}) of the IR quasars with respect to $L_{IR}^{(cd)}$ (left panel) and T_{mbb} (right panel). The middle panel shows the histogram. The gray squares, yellow circles, and red diamonds are for the objects with $5 > T_{mbb}/T_{mbb}^{err} > 3$, $10 > T_{mbb}/T_{mbb}^{err} > 5$, and $T_{mbb}/T_{mbb}^{err} > 10$, respectively. The gray, yellow, and red histograms in the middle panel are for the objects with $T_{mbb}/T_{mbb}^{err} > 3$, 5, and 10, respectively.

The MBB fits also resulted in estimates of dust mass (hereafter M_{dust}), whose distribution is shown in Figure 7.10 with respect to $L_{IR}^{(cd)}$ and T_{mbb}. As the calculation
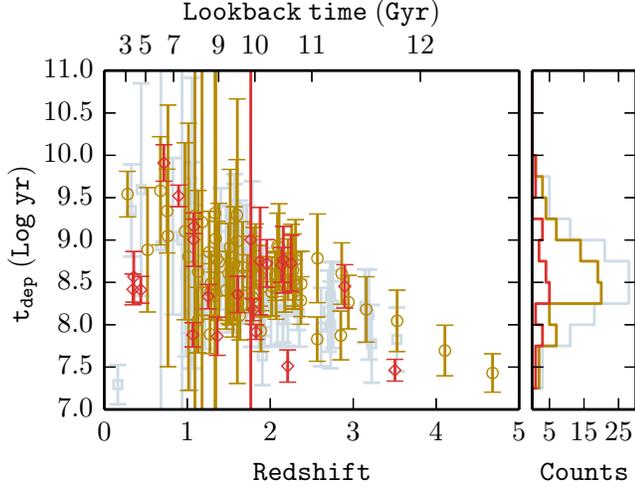
Figure 7.11: Distribution of the gas depletion time t_{dep} for the IR quasars. The left panel shows the distribution with respect to redshifts and look-back time. Some objects have t_{dep} larger than the look-back time, which could mean that they would still have plenty of gas left when evolve 0to today. The middle panel shows the histogram of t_{dep}. The symbols are the same as in Figure 7.10.

of M_{dust} is strongly affected by T_{mbb} ($M_{dust} \propto L_{IR}^{(cd)} T_{mbb}^{-5}$; see Casey 2012), again only those with $T_{mbb}/T_{mbb}^{err} \geq 3$ are included in the plot. The distribution of M_{dust} peaks at $\sim 10^{8.75}$–$10^{9.5}$ $M_\odot$, which seems to be higher than the dust contents of ULIRGs at these redshifts in general (see, e.g., Yan et al., 2014). However, considering that a good fraction of our objects have $L_{IR}^{(cd)} > 10^{13}$ $L_\odot$, such high dust masses probably are not surprising. Interestingly, for the objects in our sample, Figure 7.10 also suggests that M_{dust} is almost a flat distribution of T_{mbb}. This can be explained by the observed T_{mbb}-$L_{IR}^{(cd)}$ relation of our IR quasars, which can be approximated by $L_{IR}^{(cd)} \propto T_{mbb}^\alpha$, where $\alpha \sim 4$–5. This means that for *this group of objects* we should observe a flat distribution of M_{dust} with respect to T_{mbb}, which is exactly what Figure 7.10 shows. Therefore, our results are self-consistent.

Adopting a nominal gas-to-dust ratio of 140, we obtain the gas masses of these objects, which peak at $M_{gas} \sim (0.8$–$4.4) \times 10^{11} M_\odot$. This indicates that the IR quasar host galaxies are very rich in gas. If they could turn all their gas into stars in

their current ULIRG phase, they would grow substantially in stellar masses. In fact, the added stars alone would amount to the stellar masses of typical giant elliptical galaxies in the local universe, which are to the order of $10^{11}\,\mathrm{M}_\odot$. Figure 7.11 shows the time scale, t_{dep}, that their host galaxies would deplete the gas reservoir if they would keep forming stars at the rates as seen in their current ULIRG phase. Most of these objects have $t_{dep} \lesssim 560\,\mathrm{Myr}$ ($\mathrm{Log}(t_{dep}) \lesssim 8.75$), which are broadly consistent with the duration of ULIRGs and therefore would suggest that they could indeed turn all their gas reservoir in their current ULIRG phase. However, there are a few objects (22, among which 5 are among those of the most secure T_{mbb} estimates) that have $t_{dep} > 1\,\mathrm{Gyr}$ (among which one has $t_{dep} > 5\,\mathrm{Gyr}$). It is unclear whether such objects would be able to keep their extreme SFRs over such a long period.

7.5.4 Contribution to Luminosity Density

Here we investigate the contribution of IR quasars to the IR luminosity density ($\rho_{LIR}^{SB;QSO}$). This is done by adding L_{IR}^{SB} of the IR quasars in a given redshift bin and then dividing by the total survey volume in this bin. We do not intend to correct for the incompleteness imposed by the *Herschel* survey limits, and thus what we can obtain would only be a strict lower limit of the contribution from the optical quasar population.

While our sample is the largest one possible at this stage, its size is still rather limited when being divided into subsamples by redshifts, and thus the exact step-sizes in the redshift domain will slightly affect the detailed results. For this reason, we adopted two types of binning in redshifts, one being a uniform division at a step-size of $\Delta z = 0.5$ and the other being equal volume ($1.46 \times 10^9\,\mathrm{Mpc}^3$) in the successive bins. The results are shown in Figure 7.12. While the detailed features are somewhat different in these two schemes, the overall characteristics are the same.

In general, IR quasars among optical quasar population only contribute a very
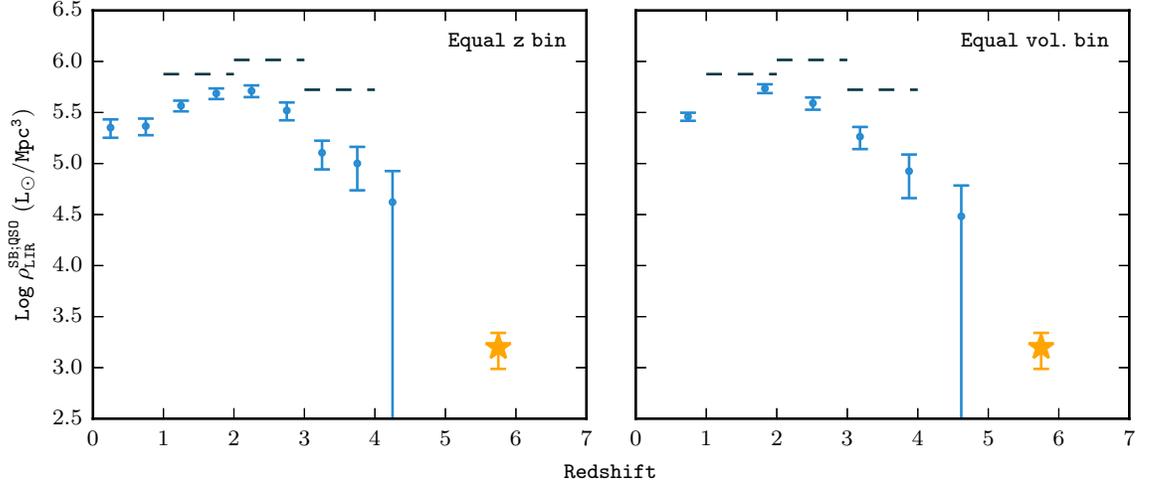
Figure 7.12: Contribution of IR quasars to the IR luminosity density as a function of redshift, for the cases of equal redshift binning (left) and equal volume binning (right). The IR luminosities used here are based on the starburst models (i.e., L_{IR}^{SB}). The error bars take into account both the errors in the derived L_{IR}^{SB} values and the Poissonian uncertainties in the number counts. The yellow asterisk denotes the result derived using the high-z IR quasar sample of Leipski et al. (2014). The horizontal dashed lines are the contribution of the *Herschel*-undetected quasars based on the stacking analysis (Ma & Yan, 2015)

small fraction to the total IR luminosity density. One recent example to compare to is the study of Viero et al. (2013), where the authors use a K-band selected galaxy sample to perform a similar analysis in the FIR and find that the IR luminosity density produced by galaxies peaks around $z = 1$–2 at $\sim 10^{8.9}\,\mathrm{L_\odot\,Mpc^{-3}}$. As Figure 7.12 shows, dust emission due to star forming activity of IR quasars only contribute $\sim 0.06\%$ of this amount.

It is well known that optical quasar number density peaks at around $z \approx 2$–3 (see e.g., Osmer 2004 for review). Using the SDSS DR3 quasars, Richards et al. (2006) find the peak at $z \approx 2.5$ when integrating the i-band luminosity function to $M_i[z = 2] = -27.6\,\mathrm{mag}$. Jiang et al. (2006) use a faint quasar sample in an area of $3.9\,\mathrm{deg}^2$ within the SDSS Stripe 82 and find that the peak would shift to $z \approx 2$ if integrating to $M_g = -22.5\,\mathrm{mag}$ (see also Ross et al., 2013). The UV luminosity density (and hence the global SFR density) of normal galaxies rises sharply from

$z = 0$ to higher redshifts, also peaks at $z \approx 2$, but keeps flat out to $z \approx 4$–5 (see e.g., Hopkins & Beacom, 2006). It is thus intriguing that $\rho_{LIR}^{SB;QSO}$ peaks at around the same redshift, which suggests that this could be the result of the co-evolution of extreme star formation and quasars.

7.6 Conclusion

In this work, we combined the SDSS quasars from DR7 and DR10 and searched for their FIR counterparts in the $172 \deg^2$ *Herschel* wide survey fields where the high-level *Herschel* maps have been made public by the relevant survey teams. From the total of 6170 SDSS quasars within the *Herschel* field coverage, we assembled a sample of 354 quasars that are detected in the *Herschel* SPIRE data, 134 of which are highly secure, SNR ≥ 5 sources in $250\,\mu$m. As we used a stringent matching criterion, the contamination due to source blending is minimal. This IR quasar sample spans a wide redshift range of $0.14 \leq z \leq 4.7$, and is the largest sample of optical quasars that have FIR detections.

To investigate their properties, we analyzed their FIR SEDs using a modified black body model (MBB) as well as three different sets of starburst (SB) templates (SK07, CE01 and DH02). By focusing on the FIR emission, we confine our discussion mainly to the cold-dust component of IR quasars, as the FIR emission is predominantly due to this component. Our conclusions are summarized below.

- The results based on the MBB model (independent of the heating source) are consistent with those based on the SB models (legitimate only when the heating source is from star formation), which strongly suggests that the IR luminosity of the cold-dust component in these IR quasars ($L_{IR}^{(cd)}$) is mainly due to the heating of star formation rather than AGN. This is further strengthened by the additional supporting evidence that there is no positive correlation between

$L_{IR}^{(cd)}$ and the absolute magnitudes or the black hole masses of the IR quasars, both of the latter being indicators of the strength of AGN. $L_{IR}^{(cd)}$, derived in this work as L_{IR}^{mbb} based on the SPIRE photometry, underestimates the total IR luminosity because it does not include the contribution from other warmer dust components. However, as it is very unlikely being significantly contaminated by the AGN heating, this quantity is more preferable in inferring the SFR of the host galaxies.

- The derived $L_{IR}^{(cd)}$ values, adopting the more conservative ones based on the MBB fitting results, range from $10^{10.8}$ to $10^{13.8}\,\mathrm{L_\odot}$ (after discarding two objects whose SEDs are barely constrained), with $\sim 80\%$ being ULIRG ($> 10^{12}\,\mathrm{L_\odot}$) and $\sim 15\%$ being HyLIRG ($> 10^{13}\,\mathrm{L_\odot}$). There is a general trend that $L_{IR}^{(cd)}$ increases at larger redshifts, which is mostly due to the selection effect in our flux-limited sample. However, there is a lack of high $L_{IR}^{(cd)}$ objects at the lowest redshifts ($z < 1$), which is broadly consistent with the picture that ULIRGs are scarce in the low-redshift universe.

- The IR quasars with well constrained dust temperatures allow us to investigate the T-$L_{IR}^{(cd)}$ relation over a wide dynamic range in $L_{IR}^{(cd)}$. We find that there is a dramatic increase of dust temperature at $L_{IR}^{(cd)} > 10^{12}\,\mathrm{L_\odot}$. Through simulations, we show that this trend cannot be due to the selection effect of our sample. Instead, this trend, which holds for IR quasars, seems to be the envelope of the general distribution of IR objects on the $(T, L_{IR}^{(cd)})$ plane. At the low luminosity end along this envelope, the increasing of $L_{IR}^{(cd)}$ is largely due to the increasing in the effective radius of the heated region (or equivalently, the enclosed dust mass). The behavior of the trend shows that the size of the heated region cannot be arbitrarily increased, and any further increasing of $L_{IR}^{(cd)}$ must be largely driven by the increased heating (i.e., more intense star formation rate

per unit volume).

- The SFR values inferred from $L_{IR}^{(cd)}$ range from 6.3 to 6310 $M_\odot\,yr^{-1}$ (for a Chabrier IMF; the values would be 1.7× higher if using a Salpeter IMF). From the dust mass derived via MBB fitting, and using a nominal gas-to-dust ratio of 140, we have inferred the gas mass for the IR quasars, most of which being within $\sim (0.8\text{–}4.4) \times 10^{11} M_\odot$. Given the SFR and the gas mass, for most of the objects in our sample, the time scale that their host galaxies would deplete the gas reservoir is $t_{dep} \lesssim 560\,\mathrm{Myr}$, while a few of them could have $t_{dep} > 1\,\mathrm{Gyr}$.

- IR quasars only contribute a very small fraction to the total IR luminosity density. This contribution also peaks at around $z \approx 2\text{–}3$.

Table 7.1: Summary of the online data table

Column	Name	Description
1	SDSS_IAU_name	SDSS quasar IAU name
2	Herschel_IAU_name	IAU name of the Herschel counterpart as in the original catalogs
3	str_id	Object ID assigned in this work
4	RA_250	Herschel RA (J2000)
5	Dec_250	Herschel Dec (J2000)
6	F250	$250\,\mu m$ flux (mJy)
7	E250	$250\,\mu m$ flux error (mJy)
8	F350	$350\,\mu m$ flux (mJy)
9	E350	$350\,\mu m$ flux error (mJy)
10	F500	$500\,\mu m$ flux (mJy)
11	E500	$500\,\mu m$ flux (mJy)
12	is_SNR5	1: SNR in $250\,\mu m >= 5$; 0: SNR in $250\,\mu m < 5$
13	RA_100	$100\,\mu m$ RA (J2000)
14	Dec_100	$100\,\mu m$ Dec (J2000)
15	F100	$100\,\mu m$ flux (mJy)
16	E100	$100\,\mu m$ flux error (mJy)
17	RA_160	$160\,\mu m$ RA (J2000)
18	Dec_160	$160\,\mu m$ Dec (J2000)
19	F160	$160\,\mu m$ flux (mJy)
20	E160	$160\,\mu m$ flux error (mJy)
21	RA_DR7	SDSS DR7 RA (J2000)
22	Dec_DR7	SDSS DR7 Dec (J2000)
23	z_DR7	SDSS DR7 redshift
24	Mi_DR7	SDSS DR7 i'-band absolute magnitude (at z=0)
25	RA_DR10	SDSS DR10 RA (J2000)
26	Dec_DR10	SDSS DR10 Dec (J2000)
27	z_DR10	SDSS DR10 redshift
28	Mi_DR10	SDSS DR10 i'-band absolute magnitude (at z=2)
29	LOG_MBH	Black hole mass measured by Shen et al. (2011) ($LogM_\odot$)
30	LOG_MBHERR	Error of the black hole mass
31	LOG_LX	Hard-band X-ray luminosity in the restframe 2–10 keV (erg/s)
32	GAMMA	Photon index of the X-ray SED used to derive X-ray luminosity
33	XRAY_REF	References of the X-ray data. CSC: Evans et al. (2010); 3XMM: Rosen et al. (2015)
34	LOG_LIR_MBB	$L_{IR}^{(cd)}$ measured using MBB fitting ($LogL_\odot$), integrated over 8–1000 μm
35	LOG_LIRERR_MBB	Error of the measured $L_{IR}^{(cd)}$
36	CHISQ_MBB	χ^2 of the MBB fitting
37	TMBB	Derived black body temperature of the best fit model (K)
38	TMBBERR	Error of T_{MBB} (K)
39	TPEAK	Derived peak temperature of the best fit model (K)
40	LOG_MDUST	Dust mass ($LogM_\odot$)
41	LOG_MDUSTERR	Error of dust mass
42	LOG_SFR	SFR using Kennicutt (1998), modified for Chabrier IMF ($LogM_\odot/yr$)
43	LOG_MGAS	Gas mass converted from dust mass (assuming gas-to-mass-ratio of 140)
44	LOG_TDEP	Gas depletion time scale (Logyr)
45	LOG_LIR_SK07	L_{IR} measured using SK07 templates ($LogL_\odot$)
46	LOG_LIRERR_SK07	Error of the SK07 L_{IR}
47	CHISQ_SK07	χ^2 of the SK07 fitting
48	LOG_LIR_CE01	L_{IR} measured using CE01 templates ($LogL_\odot$)
49	LOG_LIRERR_CE01	Error of the CE01 L_{IR}
50	CHISQ_CE01	χ^2 of the CE01 fitting
51	LOG_LIR_DH02	L_{IR} measured using DH02 templates ($LogL_\odot$)
52	LOG_LIRERR_DH02	Error of the DH02 L_{IR}
53	CHISQ_DH02	χ^2 of the DH02 fitting

Chapter 8

Conclusion

8.1 Summary

In this work, we study the galaxy formation and evolution using the *Herschel* data. Using the source de-blending tool that we developed, combined with the multi-wavelength data available, we conduct a comprehensive study of the IR luminous sources in the *Herschel* wide survey fields. The main achievements of this project are summarized as follows:

- We developed a suite of data reduction tools/pipelines, including `APUS`, the pipeline building and management framework, `PostCalib`, the post-processing pipeline for preparing science-ready mosaics for the WIYN ODI instrument, as well as a pipeline for reducing the CFHT MegaCam imaging data. With these tools, we created the most up-to-date, deepest mosaics in the optical bands, from compiling the existing archival data in a variety of data archives. These high-level data products, along with the tools to produce them, are of high significance: (1) The interpretation of the *Herschel* data needs deep optical images and catalogs to identify counterparts, measure the photometric redshifts, and

128

infer the stellar properties; (2) Such deep optical data and tools can be used in a wider context for research that is not related to *Herschel*.

- To increase the area of the *Herschel* fields in which full wavelength range of optical data are available, we started our own deep optical survey, MizzouWINS. The survey started at 2014 fall, and we have obtained great amount of data in the *Herschel* wide fields including FLS, Boötes, Lockman Hole, and XMM-LSS. With the help of `APUS` and `PostCalib`, the data reduction are managed in a timely and efficient manner as the survey goes.

- We developed `CIDer`, a tool that performs source de-blending using optical priors for the *Herschel* data. This provided us the capability to explore the wide-field data obtained by the largest *Herschel* surveys including HerMES and H-ATLAS, and conduct comprehensive study to the IR luminous sources found in the *Herschel* maps that have (only) optical data available.

- Making use of our tool sets, we studied the ULIRGs and the star formation in the First Look Survey field. We produced high level data products in the optical bands using the CFHT MegaCam data as well as the data from MizzouWINS. In the near- to mid-IR, using the data from the *Spitzer* archive, we created the most up-to-date mosaics and catalogs, superseding those created 10+ years ago. Combining these rich multi-wavelength datasets, using `CIDer` we constructed a catalog for the major contributors of the *Herschel* fluxes with panchromatic SEDs that spans from optical to FIR. We analyzed the SEDs and measured the photometric redshifts, and consequently the total IR luminosities and dust properties. This allow us to infer the dust obscured star formation rate, and contribute to the global star formation rated density plot. The L_{IR}–T relation for our sources agrees with our previous result about a limited emission size for the dust star forming regions.

8.2 Future Work

There are wide possibilities to go along and beyond this project, which are summarized as follows:

- The further development of `CIDer`. Although we have demonstrated that `CIDer` are effective in solving the haunting problem of source de-blending problem for observations that have low spatial resolution like the FIR or sub-mm, it still suffers the problem of degeneracy between very close sources. To break the degeneracy, additional information, other than the position, has to be considered. One possibility is to make use of the redshift, but how exactly this could be done is still unclear.

- The study in other *Herschel* fields. This is straightforward, as we have already had all the necessary tools and data, and we will apply the same analysis as we did for the FLS field to the other wide *Herschel* fields, including Boötes, Lockman Hole, and XMM-LSS.

- Detailed study of selected samples of the *Herschel* ULIRGs. From our ULIRG catalog, we are able to select those sources that present interesting features or are representative to a certain population. This serves as a great pool of sources to more detailed follow-up observations using the interferometers such as ALMA. Also, with the launch of *JWST*, soon will the view of the early universe change. The study on the *Herschel* data are of great value in guidance of our research in the *JWST* era.

Bibliography

Aaronson, M., & Olszewski, E. W. 1984, Optical counterparts of unidentified IRAS point sources Infrared luminous galaxies, Nature, 309, 414, doi: `10.1038/309414a0`

Arnouts, S., & Ilbert, O. 2011, LePHARE: Photometric Analysis for Redshift Estimate, Astrophysics Source Code Library. `http://ascl.net/1108.009`. `http://adsabs.harvard.edu/abs/2011ascl.soft08009A`

Azadi, M., Aird, J., Coil, A. L., et al. 2015, PRIMUS: The Relationship between Star Formation and AGN Accretion, ApJ, 806, 187, doi: `10.1088/0004-637X/806/2/187`

Barger, A. J., Cowie, L. L., & Sanders, D. B. 1999, Resolving the Submillimeter Background: The 850 Micron Galaxy Counts, ApJL, 518, L5, doi: `10.1086/312054`

Beichmann, C. A. 1985, Infrared Astronomical Satellite (IRAS) catalogs and atlases. Explanatory supplement. `http://adsabs.harvard.edu/abs/1985iasi.book.....B`

Berta, S., Magnelli, B., Nordon, R., et al. 2011, Building the cosmic infrared background brick by brick with Herschel/PEP, A&A, 532, A49, doi: `10.1051/0004-6361/201116844`

Bertin, E., & Arnouts, S. 1996, SExtractor: Software for source extraction., A&AS, 117, 393, doi: 10.1051/aas:1996164

Bertoldi, F., Carilli, C. L., Menten, K. M., et al. 2000, Three high-redshift millimeter sources and their radio and near-infrared identifications, A&A, 360, 92

Béthermin, M., Dole, H., Cousin, M., & Bavouzet, N. 2010, Submillimeter number counts at 250 μm, 350 μm and 500 μm in BLAST data, A&A, 516, A43, doi: 10.1051/0004-6361/200913910

Béthermin, M., Le Floc'h, E., Ilbert, O., 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, A&A, 542, A58, doi: 10.1051/0004-6361/201118698

Blain, A. W., Kneib, J.-P., Ivison, R. J., & Smail, I. 1999, Deep Counts of Submillimeter Galaxies, ApJL, 512, L87, doi: 10.1086/311879

Blain, A. W., Smail, I., Ivison, R. J., Kneib, J.-P., & Frayer, D. T. 2002, Submillimeter galaxies, PhR, 369, 111, doi: 10.1016/S0370-1573(02)00134-5

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, Photometric redshifts based on standard SED fitting procedures, A&A, 363, 476

—. 2011, Hyperz: Photometric Redshift Code, Astrophysics Source Code Library. http://ascl.net/1108.010. http://adsabs.harvard.edu/abs/2011ascl.soft08010B

Borys, C., Scott, D., Chapman, S., et al. 2004, The Hubble Deep Field North SCUBA Super-map - II. Multiwavelength properties, MNRAS, 355, 485, doi: 10.1111/j.1365-2966.2004.08335.x

Bovy, J., Hennawi, J. F., Hogg, D. W., et al. 2011, Think Outside the Color Box: Probabilistic Target Selection and the SDSS-XDQSO Quasar Targeting Catalog, ApJ, 729, 141, doi: 10.1088/0004-637X/729/2/141

Bruzual, G., & Charlot, S. 2003, Stellar population synthesis at the resolution of 2003, MNRAS, 344, 1000, doi: 10.1046/j.1365-8711.2003.06897.x

Bussmann, R. S., Riechers, D., Fialkov, A., et al. 2015, HerMES: ALMA Imaging of Herschel-selected Dusty Star-forming Galaxies, ApJ, 812, 43, doi: 10.1088/0004-637X/812/1/43

Calzetti, D. 1997, Reddening and Star Formation in Starburst Galaxies, AJ, 113, 162, doi: 10.1086/118242

Cao, C., Xia, X. Y., Wu, H., et al. 2008, Mid-Infrared spectroscopic properties of ultra-luminous infrared quasars, MNRAS, 390, 336, doi: 10.1111/j.1365-2966.2008.13747.x

Carter, D. 1984, The nature of the IRAS 'warm' galaxies, Astronomy Express, 1, 61

Casey, C. M. 2012, Far-infrared spectral energy distribution fitting for galaxies near and far, MNRAS, 425, 3094, doi: 10.1111/j.1365-2966.2012.21455.x

Chapman, S. C., Blain, A. W., Smail, I., & Ivison, R. J. 2005, A Redshift Survey of the Submillimeter Galaxy Population, ApJ, 622, 772, doi: 10.1086/428082

Chary, R., & Elbaz, D. 2001, Interpreting the Cosmic Infrared Background: Constraints on the Evolution of the Dust-enshrouded Star Formation Rate, ApJ, 556, 562, doi: 10.1086/321609

Daddi, E., Renzini, A., Pirzkal, N., et al. 2005a, Passively Evolving Early-Type Galaxies at $1.4 < z < 2.5$ in the Hubble Ultra Deep Field, ApJ, 626, 680, doi: 10.1086/430104

Daddi, E., Dickinson, M., Chary, R., et al. 2005b, The Population of BzK-selected ULIRGs at z ~ 2, ApJL, 631, L13, doi: 10.1086/496918

Dai, Y. S., Bergeron, J., Elvis, M., et al. 2012, A Population of Dust-rich Quasars at z ~ 1.5, ApJ, 753, 33, doi: 10.1088/0004-637X/753/1/33

Dale, D. A., & Helou, G. 2002, The Infrared Spectral Energy Distribution of Normal Star-forming Galaxies: Calibration at Far-Infrared and Submillimeter Wavelengths, ApJ, 576, 159, doi: 10.1086/341632

Davies, J. I., Baes, M., Bendo, G. J., et al. 2010, The Herschel Virgo Cluster Survey. I. Luminosity function, A&A, 518, L48, doi: 10.1051/0004-6361/201014571

Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, The All-Wavelength Extended Groth Strip International Survey (AEGIS) Data Sets, ApJL, 660, L1, doi: 10.1086/517931

de Graauw, T., Helmich, F. P., Phillips, T. G., et al. 2010, The Herschel-Heterodyne Instrument for the Far-Infrared (HIFI), A&A, 518, L6, doi: 10.1051/0004-6361/201014698

de Grijp, M. H. K., Miley, G. K., Lub, J., & de Jong, T. 1985, Infrared Seyferts - A new population of active galaxies?, Nature, 314, 240, doi: 10.1038/314240a0

Dickinson, M. 1998, Color-Selected High Redshift Galaxies and the HDF, in The Hubble Deep Field, ed. M. Livio, S. M. Fall, & P. Madau, 219

Dole, H., Gispert, R., Lagache, G., et al. 2001, FIRBACK: III. Catalog, source counts, and cosmological implications of the 170 mu m ISO, A&A, 372, 364, doi: 10.1051/0004-6361:20010449

Dole, H., Lagache, G., Puget, J.-L., et al. 2006, The cosmic infrared background

resolved by Spitzer. Contributions of mid-infrared galaxies to the far-infrared background, A&A, 451, 417, doi: 10.1051/0004-6361:20054446

Donley, J. L., Rieke, G. H., Rigby, J. R., & Pérez-González, P. G. 2005, Unveiling a Population of AGNs Not Detected in X-Rays, ApJ, 634, 169, doi: 10.1086/491668

Draine, B. T. 2006, On the Submillimeter Opacity of Protoplanetary Disks, ApJ, 636, 1114, doi: 10.1086/498130

Draine, B. T., Dale, D. A., Bendo, G., et al. 2007, Dust Masses, PAH Abundances, and Starlight Intensities in the SINGS Galaxy Sample, ApJ, 663, 866, doi: 10.1086/518306

Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, GAMA: towards a physical understanding of galaxy formation, Astronomy and Geophysics, 50, 5.12, doi: 10.1111/j.1468-4004.2009.50512.x

Dunlop, J. S., McLure, R. J., Yamada, T., et al. 2004, Discovery of the galaxy counterpart of HDF 850.1, the brightest submillimetre source in the Hubble Deep Field, MNRAS, 350, 769, doi: 10.1111/j.1365-2966.2004.07700.x

Eales, S., Dunne, L., Clements, D., et al. 2010, The Herschel ATLAS, PASP, 122, 499, doi: 10.1086/653086

Egami, E., Dole, H., Huang, J.-S., et al. 2004, Spitzer Observations of the SCUBA/VLA Sources in the Lockman Hole: Star Formation History of Infrared-Luminous Galaxies, ApJS, 154, 130, doi: 10.1086/423322

Elbaz, D., Cesarsky, C. J., Fadda, D., et al. 1999, Source counts from the 15 mu m ISOCAM Deep Surveys, A&A, 351, L37

Elston, R., Cornell, M. E., & Lebofsky, M. J. 1985, The properties of far-infrared luminous galaxies. I Spectroscopic and near-infrared observations, ApJ, 296, 106, doi: 10.1086/163423

Evans, A. S., Frayer, D. T., Surace, J. A., & Sanders, D. B. 2001, Molecular Gas in Infrared-Excess, Optically Selected and the Quasars Connection with Infrared-Luminous Galaxies, AJ, 121, 1893, doi: 10.1086/319972

Evans, I. N., Primini, F. A., Glotfelty, K. J., et al. 2010, The Chandra Source Catalog, ApJS, 189, 37, doi: 10.1088/0067-0049/189/1/37

Fadda, D., Marleau, F. R., Storrie-Lombardi, L. J., et al. 2006, The Spitzer Space Telescope Extragalactic First Look Survey: 24 μm Data Reduction, Catalog, and Source Identification, AJ, 131, 2859, doi: 10.1086/504034

Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L., & Shafer, R. A. 1998, The Spectrum of the Extragalactic Far-Infrared Background from the COBE FIRAS Observations, ApJ, 508, 123, doi: 10.1086/306383

Flores, H., Hammer, F., Thuan, T. X., et al. 1999, 15 Micron Infrared Space Observatory Observations of the 1415+52 Canada-France Redshift Survey Field: The Cosmic Star Formation Rate as Derived from Deep Ultraviolet, Optical, Mid-Infrared, and Radio Photometry, ApJ, 517, 148, doi: 10.1086/307172

Fox, M. J., Efstathiou, A., Rowan-Robinson, M., et al. 2002, The SCUBA 8-mJy survey - II. Multiwavelength analysis of bright submillimetre sources, MNRAS, 331, 839, doi: 10.1046/j.1365-8711.2002.05111.x

Franceschini, A., Manners, J., Polletta, M. d. C., et al. 2005, A Complete Multi-wavelength Characterization of Faint Chandra X-Ray Sources Seen in the Spitzer Wide-Area Infrared Extragalactic (SWIRE) Survey, AJ, 129, 2074, doi: 10.1086/428004

Frayer, D. T., Chapman, S. C., Yan, L., et al. 2004, Infrared Properties of Radio-selected Submillimeter Galaxies in the Spitzer First Look Survey Verification Field, ApJS, 154, 137, doi: 10.1086/423250

Frayer, D. T., Fadda, D., Yan, L., et al. 2006, Spitzer 70 and 160 μm Observations of the Extragalactic First Look Survey, AJ, 131, 250, doi: 10.1086/498690

Glazebrook, K., Abraham, R., Santiago, B., Ellis, R., & Griffiths, R. 1998, The physical parameters of the evolving population of faint galaxies, MNRAS, 297, 885, doi: 10.1046/j.1365-8711.1998.01585.x

Goodstadt, L. 2010, Ruffus: a lightweight Python library for computational pipelines, Bioinformatics, 26, 2778

Greenfield, P., Droettboom, M., & Bray, E. 2015, ASDF: A new data format for astronomy, Astronomy and Computing, 12, 240, doi: 10.1016/j.ascom.2015.06.004

Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, The Herschel-SPIRE instrument and its in-flight performance, A&A, 518, L3, doi: 10.1051/0004-6361/201014519

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey, ApJS, 197, 35, doi: 10.1088/0067-0049/197/2/35

Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, The Herschel PEP/HerMES luminosity function - I. Probing the evolution of PACS selected Galaxies to z 4, MNRAS, 432, 23, doi: 10.1093/mnras/stt308

Haas, M., Klaas, U., Müller, S. A. H., et al. 2003, The ISO view of Palomar-Green quasars, A&A, 402, 87, doi: 10.1051/0004-6361:20030110

Hao, L., Weedman, D. W., Spoon, H. W. W., et al. 2007, The Distribution of Silicate Strength in Spitzer Spectra of AGNs and ULIRGs, ApJL, 655, L77, doi: 10.1086/511973

Hauser, M. G., & Dwek, E. 2001, The Cosmic Infrared Background: Measurements and Implications, ARA&A, 39, 249, doi: 10.1146/annurev.astro.39.1.249

Hauser, M. G., Arendt, R. G., Kelsall, T., et al. 1998, The COBE Diffuse Infrared Background Experiment Search for the Cosmic Infrared Background. I. Limits and Detections, ApJ, 508, 25, doi: 10.1086/306379

Heckman, T. M., Armus, L., & Miley, G. K. 1987, Evidence for large-scale winds from starburst galaxies. II - an optical investigation of powerful far-infrared galaxies, AJ, 93, 276, doi: 10.1086/114310

Hodge, J. A., Karim, A., Smail, I., et al. 2013, An ALMA Survey of Submillimeter Galaxies in the Extended Chandra Deep Field South: Source Catalog and Multiplicity, ApJ, 768, 91, doi: 10.1088/0004-637X/768/1/91

Holland, W. S., Robson, E. I., Gear, W. K., et al. 1999, SCUBA: a common-user submillimetre camera operating on the James Clerk Maxwell Telescope, MNRAS, 303, 659, doi: 10.1046/j.1365-8711.1999.02111.x

Hopkins, A. M., & Beacom, J. F. 2006, On the Normalization of the Cosmic Star Formation History, ApJ, 651, 142, doi: 10.1086/506610

Houck, J. R., Schneider, D. P., Danielson, G. E., et al. 1985, Unidentified IRAS sources - Ultrahigh-luminosity galaxies, ApJL, 290, L5, doi: 10.1086/184431

Houck, J. R., Soifer, B. T., Neugebauer, G., et al. 1984, Unidentified point sources in the IRAS minisurvey, ApJL, 278, L63, doi: 10.1086/184224

Hurley, P. 2017, XID+: Next generation XID development, Astrophysics Source Code Library. http://ascl.net/1704.012. http://adsabs.harvard.edu/abs/2017ascl.soft04012H

Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, Mass assembly in quiescent and star-forming galaxies since z 4 from UltraVISTA, A&A, 556, A55, doi: 10.1051/0004-6361/201321100

Ivison, R. J., Greve, T. R., Dunlop, J. S., et al. 2007, The SCUBA HAlf Degree Extragalactic Survey - III. Identification of radio and mid-infrared counterparts to submillimetre galaxies, MNRAS, 380, 199, doi: 10.1111/j.1365-2966.2007.12044.x

Jiang, L., Fan, X., Cool, R. J., et al. 2006, A Spectroscopic Survey of Faint Quasars in the SDSS Deep Stripe. I. Preliminary Results from the Co-added Catalog, AJ, 131, 2788, doi: 10.1086/503745

Kashlinsky, A. 2005, Cosmic infrared background and early galaxy evolution [review article], PhR, 409, 361, doi: 10.1016/j.physrep.2004.12.005

Kaviani, A., Haehnelt, M. G., & Kauffmann, G. 2003, Modelling SCUBA sources in a ΛCDM cosmology: hot starbursts or cold extended galactic dust?, MNRAS, 340, 739, doi: 10.1046/j.1365-8711.2003.06318.x

Kennefick, J. D., Djorgovski, S. G., & de Carvalho, R. R. 1995, The Luminosity Function of z4 Quasars from the Second Palomar Sky Survey, AJ, 110, 2553, doi: 10.1086/117711

Kennicutt, Jr., R. C. 1998, Star Formation in Galaxies Along the Hubble Sequence, ARA&A, 36, 189, doi: 10.1146/annurev.astro.36.1.189

Kim, D.-C., & Sanders, D. B. 1998, The IRAS 1 Jy Survey of Ultraluminous Infrared Galaxies. I. The Sample and Luminosity Function, ApJS, 119, 41, doi: 10.1086/313148

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey—The Hubble Space Telescope Observations, Imaging Data Products, and Mosaics, ApJS, 197, 36, doi: 10.1088/0067-0049/197/2/36

Kreysa, E., Gemuend, H.-P., Gromke, J., et al. 1998, Bolometer array development at the Max-Planck-Institut fuer Radioastronomie, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 3357, Advanced Technology MMW, Radio, and Terahertz Telescopes, ed. T. G. Phillips, 319–325

Lacy, M., Wilson, G., Masci, F., et al. 2005, The Infrared Array Camera Component of the Spitzer Space Telescope Extragalactic First Look Survey, ApJS, 161, 41, doi: 10.1086/432894

Lagache, G., Abergel, A., Boulanger, F., Désert, F. X., & Puget, J.-L. 1999, First detection of the warm ionised medium dust emission. Implication for the cosmic far-infrared background, A&A, 344, 322

Lawrence, A. 2001, Blank-field submm sources, failed stars and the dark matter, MNRAS, 323, 147, doi: 10.1046/j.1365-8711.2001.04116.x

Lawrence, A., Rowan-Robinson, M., Ellis, R. S., et al. 1999, The QDOT all-sky IRAS galaxy redshift survey, MNRAS, 308, 897, doi: 10.1046/j.1365-8711.1999.02593.x

Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, The UKIRT Infrared Deep Sky Survey (UKIDSS), MNRAS, 379, 1599, doi: 10.1111/j.1365-2966.2007.12040.x

Leipski, C., Meisenheimer, K., Klaas, U., et al. 2010, Herschel-PACS far-infrared photometry of two z 4 quasars, A&A, 518, L34, doi: 10.1051/0004-6361/201014718

Leipski, C., Meisenheimer, K., Walter, F., et al. 2013, Complete Infrared Spectral Energy Distributions of Millimeter Detected Quasars at z 5, ApJ, 772, 103, doi: 10.1088/0004-637X/772/2/103

—. 2014, Spectral Energy Distributions of QSOs at z 5: Common Active Galactic Nucleus-heated Dust and Occasionally Strong Star-formation, ApJ, 785, 154, doi: 10.1088/0004-637X/785/2/154

Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, Starburst99: Synthesis Models for Galaxies with Active Star Formation, ApJS, 123, 3, doi: 10.1086/313233

Levenson, L., Marsden, G., Zemcov, M., et al. 2010, HerMES: SPIRE Science Demonstration Phase maps†‡, MNRAS, 409, 83, doi: 10.1111/j.1365-2966.2010.17771.x

Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996, The Canada-France Redshift Survey: The Luminosity Density and Star Formation History of the Universe to Z approximately 1, ApJL, 460, L1, doi: 10.1086/309975

Lilly, S. J., Tresse, L., Hammer, F., Crampton, D., & Le Fevre, O. 1995, The Canada-France Redshift Survey. VI. Evolution of the Galaxy Luminosity Function to Z approximately 1, ApJ, 455, 108, doi: 10.1086/176560

Lonsdale, C. J., Farrah, D., & Smith, H. E. 2006, in Astrophysics Update 2, ed. J. W. Mason (Springer Berlin Heidelberg), 285–336. http://adsabs.harvard.edu/abs/2006asup.book..285L

Lutz, D. 2014, Far-Infrared Surveys of Galaxy Evolution, ARA&A, 52, 373, doi: 10.1146/annurev-astro-081913-035953

Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, PACS Evolutionary Probe (PEP) - A Herschel key program, A&A, 532, A90, doi: 10.1051/0004-6361/201117107

Ma, Z., & Yan, H. 2015, Co-evolution of Extreme Star Formation and Quasars: Hints from Herschel and the Sloan Digital Sky Survey, ApJ, 811, 58, doi: 10.1088/0004-637x/811/1/58

Madau, P. 1998, After the Dark Ages: the Evolution of Luminous Sources at z5, ArXiv Astrophysics e-prints

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

Madau, P., Ferguson, H. C., Dickinson, M. E., et al. 1996, High-redshift galaxies in the Hubble Deep Field: colour selection and star formation history to z˜4, MNRAS, 283, 1388

Magnelli, B., Elbaz, D., Chary, R. R., et al. 2009, The 0.4 z 1.3 star formation history of the Universe as viewed in the far-infrared, A&A, 496, 57, doi: 10.1051/0004-6361:200811443

Magnelli, B., Popesso, P., Berta, S., et al. 2013, The deepest Herschel-PACS far-infrared survey: number counts and infrared luminosity functions from combined PEP/GOODS-H observations, A&A, 553, A132, doi: 10.1051/0004-6361/201321371

Marleau, F. R., Fadda, D., Appleton, P. N., et al. 2007, Spectroscopic Survey of 1.4 GHz and 24 μm Sources in the Spitzer First Look Survey with WIYN Hydra, ApJ, 663, 218, doi: 10.1086/518114

Martínez-Sansigre, A., Rawlings, S., Lacy, M., et al. 2005, The obscuration by dust of most of the growth of supermassive black holes, Nature, 436, 666, doi: 10.1038/nature03829

Moshir, M., Kopman, G., & Conrow, T. A. O. 1992, IRAS Faint Source Survey, Explanatory supplement version 2. http://adsabs.harvard.edu/abs/1992ifss.book.....M

Mullaney, J. R., Alexander, D. M., Aird, J., et al. 2015, ALMA and Herschel reveal that X-ray-selected AGN and main-sequence galaxies have different star formation rate distributions, MNRAS, 453, L83, doi: 10.1093/mnrasl/slv110

Netzer, H., Mor, R., Trakhtenbrot, B., Shemmer, O., & Lira, P. 2014, Star Formation and Black Hole Growth at z ˜= 4.8, ApJ, 791, 34, doi: 10.1088/0004-637X/791/1/34

Netzer, H., Lutz, D., Schweitzer, M., et al. 2007, Spitzer Quasar and ULIRG Evolution Study (QUEST). II. The Spectral Energy Distributions of Palomar-Green Quasars, ApJ, 666, 806, doi: 10.1086/520716

Neugebauer, G., Habing, H. J., van Duinen, R., et al. 1984, The Infrared Astronomical Satellite (IRAS) mission, ApJL, 278, L1, doi: 10.1086/184209

Oliver, S. J., Bock, J., Altieri, B., et al. 2012, The Herschel Multi-tiered Extragalactic Survey: HerMES, MNRAS, 424, 1614, doi: 10.1111/j.1365-2966.2012.20912.x

Osmer, P. S. 2004, The Evolution of Quasars, Coevolution of Black Holes and Galaxies, 324

Osterbrock, D. E., & De Robertis, M. M. 1985, Optical spectra of IRAS 'warm' galaxies, PASP, 97, 1129, doi: 10.1086/131676

Ott, S. 2010, The Herschel Data Processing System — HIPE and Pipelines — Up and Running Since the Start of the Mission, in Astronomical Society of the Pacific Conference Series, Vol. 434, Astronomical Data Analysis Software and Systems XIX, ed. Y. Mizumoto, K.-I. Morita, & M. Ohishi, 139. http://adsabs.harvard.edu/abs/2010ASPC..434..139O

Papovich, C., Cool, R., Eisenstein, D., et al. 2006, An MMT Hectospec Redshift Survey of 24 μm Sources in the Spitzer First Look Survey, AJ, 132, 231, doi: 10.1086/504598

Pâris, I., Petitjean, P., Aubourg, É., et al. 2014, The Sloan Digital Sky Survey quasar catalog: tenth data release, A&A, 563, A54, doi: 10.1051/0004-6361/201322691

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, Detailed Structural Decomposition of Galaxy Images, AJ, 124, 266, doi: 10.1086/340952

Pettini, M., Kellogg, M., Steidel, C. C., et al. 1998, Infrared Observations of Nebular Emission Lines from Galaxies at Z ˜= 3, ApJ, 508, 539, doi: 10.1086/306431

Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, Herschel Space Observatory. An ESA facility for far-infrared and submillimetre astronomy, A&A, 518, L1, doi: 10.1051/0004-6361/201014759

Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, The Photodetector Array Camera and Spectrometer (PACS) on the Herschel Space Observatory, A&A, 518, L2, doi: 10.1051/0004-6361/201014535

Polletta, M. d. C., Wilkes, B. J., Siana, B., et al. 2006, Chandra and Spitzer Unveil Heavily Obscured Quasars in the Chandra/SWIRE Survey, ApJ, 642, 673, doi: 10.1086/500821

Richards, G. T., Fan, X., Newberg, H. J., et al. 2002, Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Quasar Sample, AJ, 123, 2945, doi: 10.1086/340187

Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, The Sloan Digital Sky Survey Quasar Survey: Quasar Luminosity Function from Data Release 3, AJ, 131, 2766, doi: 10.1086/503559

Rigby, E. E., Maddox, S. J., Dunne, L., et al. 2011, Herschel-ATLAS: first data release of the Science Demonstration Phase source catalogues, MNRAS, 415, 2336, doi: 10.1111/j.1365-2966.2011.18864.x

Roseboom, I. G., Oliver, S. J., Kunz, M., et al. 2010, The Herschel Multi-Tiered Extragalactic Survey: source extraction and cross-identifications in confusion-dominated SPIRE images, MNRAS, 409, 48, doi: 10.1111/j.1365-2966.2010.17634.x

Rosen, S. R., Webb, N. A., Watson, M. G., et al. 2015, The XMM-Newton serendipitous survey. VII. The third XMM-Newton serendipitous source catalogue, ArXiv e-prints. https://arxiv.org/abs/1504.07051

Ross, N. P., Myers, A. D., Sheldon, E. S., et al. 2012, The SDSS-III Baryon Oscillation Spectroscopic Survey: Quasar Target Selection for Data Release Nine, ApJS, 199, 3, doi: 10.1088/0067-0049/199/1/3

Ross, N. P., McGreer, I. D., White, M., et al. 2013, The SDSS-III Baryon Oscillation Spectroscopic Survey: The Quasar Luminosity Function from Data Release Nine, ApJ, 773, 14, doi: 10.1088/0004-637X/773/1/14

Rowan-Robinson, M. 1995, A new model for the infrared emission of quasars, MNRAS, 272, 737

Rowan-Robinson, M., Lari, C., Perez-Fournon, I., et al. 2004, The European Large-Area ISO Survey (ELAIS): the final band-merged catalogue, MNRAS, 351, 1290, doi: 10.1111/j.1365-2966.2004.07868.x

Sanders, D. B., Phinney, E. S., Neugebauer, G., Soifer, B. T., & Matthews, K. 1989, Continuum energy distribution of quasars - Shapes and origins, ApJ, 347, 29, doi: 10.1086/168094

Sanders, D. B., Soifer, B. T., Elias, J. H., et al. 1988, Ultraluminous infrared galaxies and the origin of quasars, ApJ, 325, 74, doi: 10.1086/165983

Sawicki, M. J., Lin, H., & Yee, H. K. C. 1997, Evolution of the Galaxy Population Based on Photometric Redshifts in the Hubble Deep Field, AJ, 113, 1, doi: 10.1086/118231

Schinnerer, E., Carilli, C. L., Scoville, N. Z., et al. 2004, The VLA-COSMOS Survey. I. Radio Identifications from the Pilot Project, AJ, 128, 1974, doi: 10.1086/424860

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds, ApJ, 500, 525, doi: 10.1086/305772

Schmidt, M., Schneider, D. P., & Gunn, J. E. 1995, Spectrscopic CCD Surveys for Quasars at Large Redshift.IV.Evolution of the Luminosity Function from Quasars Detected by Their Lyman-Alpha Emission, AJ, 110, 68, doi: 10.1086/117497

Schneider, D. P., Richards, G. T., Hall, P. B., et al. 2010, The Sloan Digital Sky Survey Quasar Catalog. V. Seventh Data Release, AJ, 139, 2360, doi: 10.1088/0004-6256/139/6/2360

Schweitzer, M., Lutz, D., Sturm, E., et al. 2006, Spitzer Quasar and ULIRG Evolution

Study (QUEST). I. The Origin of the Far-Infrared Continuum of QSOs, ApJ, 649, 79, doi: 10.1086/506510

Scott, S. E., Fox, M. J., Dunlop, J. S., et al. 2002, The SCUBA 8-mJy survey - I. Submillimetre maps, sources and number counts, MNRAS, 331, 817, doi: 10.1046/j.1365-8711.2002.05193.x

Scoville, N., Aussel, H., Brusa, M., et al. 2007, The Cosmic Evolution Survey (COS-MOS): Overview, ApJS, 172, 1, doi: 10.1086/516585

Scoville, N. Z., Frayer, D. T., Schinnerer, E., & Christopher, M. 2003, The Host Galaxies of Optically Bright Quasi-stellar Objects: Molecular Gas in z=0.1 Palomar-Green Quasi-stellar Objects, ApJL, 585, L105, doi: 10.1086/374544

Serjeant, S., Bertoldi, F., Blain, A. W., et al. 2010, Herschel ATLAS: The cosmic star formation history of quasar host galaxies, A&A, 518, L7, doi: 10.1051/0004-6361/201014565

Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, A Catalog of Quasar Properties from Sloan Digital Sky Survey Data Release 7, ApJS, 194, 45, doi: 10.1088/0067-0049/194/2/45

Shi, Y., Ogle, P., Rieke, G. H., et al. 2007, Aromatic Features in AGNs: Star-forming Infrared Luminosity Function of AGN Host Galaxies, ApJ, 669, 841, doi: 10.1086/521594

Siebenmorgen, R., & Krügel, E. 2007, Dust in starburst nuclei and ULIRGs. SED models for observers, A&A, 461, 445, doi: 10.1051/0004-6361:20065700

Simpson, C., Dunlop, J. S., Eales, S. A., et al. 2004, Deep near-infrared spectroscopy of submillimetre-selected galaxies, MNRAS, 353, 179, doi: 10.1111/j.1365-2966.2004.08054.x

Smith, A. J., Wang, L., Oliver, S. J., et al. 2012, HerMES: point source catalogues from deep Herschel-SPIRE observations, MNRAS, 419, 377, doi: `10.1111/j.1365-2966.2011.19709.x`

Smith, D. J. B., Dunne, L., Maddox, S. J., et al. 2011, Herschel-ATLAS: counterparts from the ultraviolet-near-infrared in the science demonstration phase catalogue, MNRAS, 416, 857, doi: `10.1111/j.1365-2966.2011.18827.x`

Smith, M. W. L., Ibar, E., Maddox, S. J., et al. 2017, 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, ApJS, 233, 26, doi: `10.3847/1538-4365/aa9b35`

Solomon, P. M., & Vanden Bout, P. A. 2005, Molecular Gas at High Redshift, ARA&A, 43, 677, doi: `10.1146/annurev.astro.43.051804.102221`

Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, Lyman-Break Galaxies at z~4 and the Evolution of the Ultraviolet Luminosity Density at High Redshift, ApJ, 519, 1, doi: `10.1086/307363`

Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, Spectroscopic Confirmation of a Population of Normal Star-forming Galaxies at Redshifts Z 3, ApJL, 462, L17, doi: `10.1086/310029`

Symeonidis, M., Vaccari, M., Berta, S., et al. 2013, The Herschel census of infrared SEDs through cosmic time, MNRAS, 431, 2317, doi: `10.1093/mnras/stt330`

Symeonidis, M., Georgakakis, A., Page, M. J., et al. 2014, Linking the X-ray and infrared properties of star-forming galaxies at z 1.5, MNRAS, 443, 3728, doi: `10.1093/mnras/stu1441`

Tresse, L., & Maddox, S. J. 1998, The H alpha Luminosity Function and Star Formation Rate at Z approximately 0.2, ApJ, 495, 691, doi: `10.1086/305331`

Valiante, E., Smith, M. W. L., Eales, S., et al. 2016, The Herschel-ATLAS data release 1 - I. Maps, catalogues and number counts, MNRAS, 462, 3146, doi: `10.1093/mnras/stw1806`

van Dokkum, P. G. 2001, Cosmic-Ray Rejection by Laplacian Edge Detection, PASP, 113, 1420, doi: `10.1086/323894`

Viero, M. P., Moncelsi, L., Quadri, R. F., et al. 2013, HerMES: The Contribution to the Cosmic Infrared Background from Galaxies Selected by Mass and Redshift, ApJ, 779, 32, doi: `10.1088/0004-637X/779/1/32`

Viero, M. P., Asboth, V., Roseboom, I. G., et al. 2014, The Herschel Stripe 82 Survey (HerS): Maps and Early Catalog, ApJS, 210, 22, doi: `10.1088/0067-0049/210/2/22`

Wang, L., Viero, M., Clarke, C., et al. 2014, HerMES: point source catalogues from Herschel-SPIRE observations II, MNRAS, 444, 2870, doi: `10.1093/mnras/stu1569`

Wang, R., Carilli, C. L., Neri, R., et al. 2010, Molecular Gas in z ~ 6 Quasar Host Galaxies, ApJ, 714, 699, doi: `10.1088/0004-637X/714/1/699`

Wang, R., Wagg, J., Carilli, C. L., et al. 2011a, CO (2-1) Line Emission in Redshift 6 Quasar Host Galaxies, ApJL, 739, L34, doi: `10.1088/2041-8205/739/1/L34`

—. 2011b, Far-infrared and Molecular CO Emission from the Host Galaxies of Faint Quasars at z ~ 6, AJ, 142, 101, doi: `10.1088/0004-6256/142/4/101`

Webb, T. M. A., Lilly, S. J., Clements, D. L., et al. 2003, The Canada-UK Deep

Submillimeter Survey. VII. Optical and Near-Infrared Identifications for the 14 Hour Field, ApJ, 597, 680, doi: 10.1086/381217

Weiß, A., Kovács, A., Coppin, K., et al. 2009, The Large Apex Bolometer Camera Survey of the Extended Chandra Deep Field South, ApJ, 707, 1201, doi: 10.1088/0004-637X/707/2/1201

Xia, X. Y., Gao, Y., Hao, C.-N., et al. 2012, Molecular Gas in Infrared Ultraluminous QSO Hosts, ApJ, 750, 92, doi: 10.1088/0004-637X/750/2/92

Yan, H., & Ma, Z. 2016, From the L_{IR}-T Relation to the Limited Sizes of Dusty Starbursting Regions at High Redshifts, ApJL, 820, L16, doi: 10.3847/2041-8205/820/1/L16

Yan, H., Stefanon, M., Ma, Z., et al. 2014, Optical-faint, Far-infrared-bright Herschel Sources in the CANDELS Fields: Ultra-luminous Infrared Galaxies at z 1 and the Effect of Source Blending, ApJS, 213, 2, doi: 10.1088/0067-0049/213/1/2

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, The Sloan Digital Sky Survey: Technical Summary, AJ, 120, 1579, doi: 10.1086/301513

Zemcov, M., Blain, A., Halpern, M., & Levenson, L. 2010, Contribution of Lensed SCUBA Galaxies to the Cosmic Infrared Background, ApJ, 721, 424, doi: 10.1088/0004-637X/721/1/424

VITA

I was born in a small town (Shanyin, Shanxi) in China, and got my junior high school finished there. I spent three years in Tianyuan, Shanxi to finish my senior high school, and started my seven years' journey in Nankai University, located in Tianjin, to complete my bachelor's and master's degree in physics. In year 2012, I was admitted to Department of Physics and Astronomy, University of Missouri-Columbia, as a PhD student working with Dr. Haojing Yan on galaxy formation and evolution.

Now is 2018, and it is the time to *start*, with my PhD in hand, with all the preparation in mind, and with all the hope and ambition in me.