Far-Infrared constraints on the ongoing Star Formation of distant galaxies hosting Active Galactic Nuclei

FLORA STANLEY

STANLEY, FLORA (2016) Far-Infrared constraints on the ongoing Star Formation of distant galaxies hosting Active Galactic Nuclei. Doctoral thesis, Durham University.

Far-Infrared constraints on the ongoing Star Formation of distant galaxies hosting Active Galactic Nuclei

Flora Stanley

Abstract

There have been a multitude of observational findings supporting a co-evolution of the central BH and its host galaxy. This co-evolution is expected to be driven by a connection between the two growing mechanisms of active galactic nuclei (AGN; accretion onto the central super-massive black hole), and star formation. In an effort to find more direct evidence for a connection between the two mechanisms of AGN and star formation I investigate the star formation rates (SFRs) of galaxies hosting X-ray, optical, and radio AGN. For the analysis of these samples I have developed methods to calculate the mean and individual SFRs of distant AGN and improve on previous work by simultaneously taking into account redshift and mass when interpreting the results. I use infrared photometry from *WISE*, *Spitzer* and *Herschel* to decompose the infrared (IR) spectral energy distributions (SEDs) into AGN and star formation components and undertake careful treatment of the upper limits in the SED analyses and average SFR calculations. Using these methods I have calculated the mean SFRs of ∼2000 X-ray and ∼3000 optically selected AGN spanning the AGN luminosity range of $10^{43} < L_{\rm AGN} < 5 \times 10^{47}$ erg s^{-1} at redshifts of $0.2 < z < 2.5$. I have established that the mean SFR as a function of AGN luminosity shows a flat relationship with a slight positive trend at the highest AGN luminosities. Crucially, I show that the positive trends observed here, and in previous work, are strongly influenced by the internal redshift and mass dependencies of average SFRs observed in the global galaxy population. Indeed, the AGN studied in this thesis show mean SFRs consistent to the main sequence of star-forming galaxies, indicating that on average they tend to live in star-forming galaxies. Using new sensitive sub-mm photometry from the ALMA interferometer for a sample ∼100 X–ray AGN I obtain up to 10 times improvement on the SFR constraints compared to those obtained using the archival photometry. Furthermore, I demonstrate that combining deep 870μm observations with mid-infrared photometry is a useful diagnostic for AGN identification.

Far-Infrared constraints on the ongoing Star Formation of distant galaxies hosting Active Galactic Nuclei

Flora Stanley

A thesis presented in accordance with the regulations for
admittance to the degree of Doctor of Philosophy

Durham
University

Centre for Extragalactic Astronomy
Department of Physics
University of Durham
United Kingdom

July 2016

Dedicated to

My mother

Declaration

The work described in this thesis was undertaken between 10/2012 and 07/2016 while the author was a research student under the supervision of David M. Alexander, Christopher M. Harrison, and Anthony M. Swinbank in the Department of Physics at the University of Durham. No part of this thesis has been submitted for any other degree at the University of Durham or any other university.

Portions of this work have appeared in the following paper:

Chapter 3: **"A remarkably flat relationship between star formation rate and AGN luminosity for distant X-ray AGN"**
Stanley, F., Harrison, C. M., Alexander, D. M., Swinbank, A. M., Aird, J. A., Del Moro, A., Hickox, R. C., Mullaney, J. R., MNRAS, 453, 591 (2015)

The work described in Chapter 3 was performed in collaboration with the co-authors of the above mentioned paper. The work described in Chapter 4 was performed in collaboration with Harrison, C. M., Alexander, D. M., Rosario, D. J., Wang, L., and members of the *Herschel*-ATLAS team. The work described in Chapter 5 was performed in collaboration with Harrison, C. M., Alexander, D. M., Rosario, D. J., and Scholtz, J. J.. Chapter 3 uses deblended *Herschel* catalogues created by A. M. Swinbank. Chapter 5 is based on the results of observing proposals led by J. R. Mullaney and D. M. Alexander. The data-reduction of the ALMA observations of Chapter 5 was done by J. M. Simpson. All of the text has been written by the author.

In addition to the work presented in this thesis, the author has also been involved in, the following work during the period of their PhD:

"ALMA Observations of a z 3.1 Protocluster: Star Formation from Active Galactic Nuclei and Lyman-Alpha Blobs in an Overdense Environment"
Alexander, D.M., Simpson, J.M., Harrison. C.M., Mullaney, J.R., Smail, I., Geach, J.E., Hickox, R.C., Hine, N.K., Karim, A., Kubo, M., Lehmer, B.D., Matsuda, Y., Rosario, D.J., Stanley, F., Swinbank, A.M., Umehata, H., Yamada, T., MNRAS, in press (arXiv:1601.00682)

"ALMA resolves extended star formation in high-z AGN host galaxies"
C. M. Harrison, J. M. Simpson, F. Stanley, D. M. Alexander, E. Daddi, J. R. Mullaney, M. Pannella, D. J. Rosario, I. Smail, MNRAS, 457, L122 (2016)

"Mid-infrared luminous quasars in the GOODS-Herschel fields: a large population of heavily-obscured, Compton-thick quasars at z$\sim$2"

A. Del Moro, D. M. Alexander, F. E. Bauer, E. Daddi, D. D. Kocevski, D. H. McIntosh, F. Stanley,W. N. Brandt, D. Elbaz, C. M. Harrison, B. Luo, J. R. Mullaney, Y. Q. Xue, MNRAS, 456, 2105 (2015).

"ALMA and Herschel reveal that X-ray-selected AGN and main-sequence galaxies have different star formation rate distributions"
Mullaney, J. R., Alexander, D. M., Aird, J., Bernhard, E., Daddi, E., Del Moro, A., Dickinson, M., Elbaz, D., Harrison, C. M., Juneau, S., Liu, D., Pannella, M., Rosario, D., Santini, P., Sargent, M., Schreiber, C., Simpson, J., Stanley, F., MNRAS, 453, L83 (2015)

"Herschel reveals a molecular outflow in a z = 2.3 ULIRG"
George, R. D., Ivison, R. J., Smail, Ian, Swinbank, A. M., Hopwood, R., Stanley, F., Swinyard, B. M., Valtchanov, I., van der Werf, P. P., MNRAS, 442, 1877 (2014)

Acknowledgements

I would like to thank Chris Harrison and Dave Alexander for being my supervisors and supporting and continuously helping improve my research and knowledge-base throughout these past years. I have greatly enjoyed working with both, as well as the whole group of excellent scientists that Dave manages to attract.

I would also like to thank my collaborators Agnese Del Moro, David Rosario, James Mullaney, James Aird, and Mark Swinbank, who have assisted in producing the research presented in this thesis, and have also helped in training and developing my skill-sets.

Acknowledgement is also due to the Faculty of Science Durham Doctoral Scholarship Committee for awarding me the Durham Doctoral Scholarship, thus enabling my postgraduate studies at Durham University.

I feel an important part of the past few years have been the invaluable friends I have made. Special thanks to Mathilde, Ben, George, James, Sree, Niki, Alice, Tamsyn, Ady, Sownak, Violeta, Michelle, as well as current and past office mates, and my two besties Sofia and Sofia all the way back in Greece, for being such great friends, making good times great, and those tough times of post-grad blues bearable.

Finally, I feel most grateful to my mother that has supported me at every step in my career development and life decisions, and was the one to plant the researcher bug in me.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1 Overview

It is now widely accepted that all massive galaxies host a central supermassive black hole (BH) and that all have gone through an active phase of intense growth as active galactic nuclei (AGN). The energy output of AGN makes them some of the most powerful sources in the Universe, with the potential to have a significant impact on the properties of their host galaxies. Indeed, the amount of energy injected to the interstellar medium (ISM) can cause a reduction in the available cold gas (either by heating or ejection) and hence potentially reduce the levels of star formation (this process is typically termed "AGN feedback"). Models of galaxy evolution have found it necessary to use AGN feedback as a means of regulating the star formation of galaxies to reproduce the properties of galaxies in the local Universe. However, observations of the host galaxies of AGN have come to a variety of results, with no clear and unambiguous evidence on the impact of AGN feedback on the star formation of galaxies. The aims of this introductory chapter are to provide a brief introduction and motivation to the research presented in this thesis. In this thesis I investigate the star formation rate (SFR) of galaxies hosting AGN identified from their X-ray, optical, and radio emission, and explore the relationship between the SFR and the AGN luminosity. To understand the connection between the two processes of AGN and star formation, it is important to also understand them individually. In this chapter I give a brief background of the star-forming properties and evolution of galaxies, the growth of their central BHs through the phases of AGN activity, and the evidence for a connection between AGN activity and star formation.

Throughout this introduction I use the following definitions for different wavelength ranges. UV (ultraviolet; $\sim$0.01–0.4μm), optical ($\sim$0.4–0.74μm), IR (infrared; 8–1000μm), NIR (near-infrared; $\sim$0.75–5μm), MIR (mid-infrared; $\sim$5–40μm), FIR (far-infrared; $\sim$40–500μm), and Sub-mm (sub-millimeter; $\sim$500–1200μm) wavebands.

1.2 Star formation

It was only during the early 20^{th} century that it became clear to astronomers that the observed “spiral nebulae” (Rosse 1850) composed of stars (found from early spectroscopic analysis, e.g. Huggins & Miller 1864; Scheiner 1899), were indeed extragalactic (i.e., not part of the Milky way; Hubble 1925, 1929). The research of “extragalactic nebulae”, now referred to as galaxies, has become one of the major branches of astronomy, with numerous studies on their properties spanning from the local Universe to the most distant galaxies.

One of the drivers of galaxy evolution is star formation. The rate at which star formation occurs is called the star formation rate (SFR), and is measured in solar masses ($M_\odot$) per year. Star formation occurs when a dense cold/molecular gas cloud collapses under its own gravity, and forms a group or cluster of stars. For a cloud of molecular gas to collapse the gravitational force has to exceed the gas pressure (in the case of stable gas; Jeans instability) and the shear force (in cases where the gas is in a differentially rotating disk; Toomre’s criterion). By necessity the star forming process is dependent on the availability of cold/molecular gas in the galaxy, which itself is dependent on a number of processes that can reduce it (by using, expelling or heating the cold gas; e.g., star formation itself, super-novae explosions, AGN outflows) or increase it (e.g., gas inflow through filaments, and gas-rich mergers).

The star formation history of a galaxy, as well as its current SFR, defines many of the key properties, such as the stellar mass (M_*) and levels of cold gas available. To understand the evolution of galaxies it is vital to measure the SFRs of galaxies across time.

1.2.1 Measuring the SFR of galaxies

Star formation tracers can be split in two groups, direct tracers of the emission from young stars, and indirect tracers that measure on the re-radiated emission from dusty star-forming regions. A direct tracer of the emission from young stellar populations is the UV emission of the young massive OB stars, and covers timescales of 0–100Myr, with higher frequency UV light tracing shorter timescales (e.g., Hao et al. 2011; Murphy et al. 2011). The X-ray band (typically 0.2–10keV) tracing the emission of X-ray binaries, supernovae, and supernovae remnants, can also be used as a measure of recent star formation (0–100Myr; e.g., Ranalli et al. 2003) Due to the nature of high-frequency UV and optical emission however, it can easily suffer from obscuration due to the surrounding dust in star forming galaxies. The absorbed emission from dust enshrouding the star-forming regions is re-emitted at IR wavelengths. As a result, the IR continuum is a great tracer of obscured star formation, but provides a time-averaged star formation measurement (tracing timescales of 10–100Myrs), and can only successfully recover the SFR of dusty galaxies (see review by Kennicutt & Evans 2012). Indeed, when the first deep IR surveys began it became evident that most of the star formation emission at higher redshifts is emitted in the IR by dusty star forming galaxies, and is the main contributor to the cosmic IR background (see review by Lutz 2014). To quantify the contribution between obscured and unobscured star formation, a study by Burgarella et al. (2013) have investigated the relative ratios of the UV and IR luminosity densities (calculated by integrating over the luminosity functions). They found that in the local universe the IR luminosity density is a factor of $\sim$4 above the UV luminosity density, and by redshift of $\sim$1.2 the IR luminosity density is more than an order of magnitude larger than that of the UV. Consequently, to successfully trace the majority of star formation at high redshifts requires sensitive IR observations.

Figure 1.1 shows two examples of galaxy SEDs, one with ongoing star formation and one with none or minimal star formation. In red is a dusty star forming galaxy (characteristic of star forming galaxies at higher redshifts) that peaks at rest-frame FIR wavelengths of $\sim$60μm, and is dominated by the reprocessed UV emission of young stars that has been absorbed by dust and reradiated in the FIR. In green is an Elliptical galaxy with no current star formation that peaks at optical/NIR wavelengths and is dominated by the emission from older stellar populations. To retrieve the majority of the star formation in the predominant population of star-forming galaxies at higher redshifts, we need to observe the rest-frame far-IR (FIR; $\lambda \sim 40 - 500\,\mu$m) that covers the peak of the reprocessed emission.

The FIR as a tracer of SFR

The cosmic FIR background was first discovered in 1996 by Puget et al. (1996) at wavelengths $\gtrsim$200μm with the Cosmic Background Explorer (COBE), and later extended over 70–850μm (e.g. Lagache et al. 1999; Dole et al. 2006; Berta et al. 2011; Magnelli et al. 2013). The shape of the CIB can be fitted with a modified blackbody (e.g., Fixsen et al. 1998) that peaks at $\sim$150–200μm. The emission is now known to be due to high-redshift dusty galaxies, with a large fraction of the emission being transmitted in the FIR. It was later shown that 50% of the energy produced by star formation throughout cosmic time has been emitted in the IR wavelengths (e.g. Dole et al. 2006; reviews by Madau & Dickinson 2014, and Lutz 2014).

The FIR waveband has been explored with numerous space observatories. The Infrared Astronomical Satellite (IRAS; Neugebauer et al. 1984) had bands extending out to 100μm and detected more than 250,000 IR sources in the local Universe and up to redshifts of $\sim$0.1 (e.g., Olnon et al. 1986). The Infrared Space Observatory (ISO) reached up to 240μm and provided the first deep survey in both MIR and FIR, from which the strong evolution of the IR luminosity up to z$\sim$1 was found (see review by Genzel & Cesarsky 2000; Lutz 2014). The Spitzer Space Telescope (*Spitzer*) reaches up to FIR wavelengths of 160μm, and is still one of the most important telescopes for deep field MIR photometry. However, both the ISO and *Spitzer* telescopes are strongly limited in their FIR observations by the source confusion, due to the blending of faint sources, caused by their relatively small mirrors of 60–80cm diameter.

With the launch of the Herschel Space Observatory (*Herschel*; Pilbratt et al. 2010), and its successful operation until 2013, astronomers were finally able to reach the peak of the FIR emission that dominates in the SEDs of star-forming galaxies at high sensitivity. On board *Herschel*, there were two instruments, the Photoconductor Array Camera and Spectrometer (PACS; Poglitsch et al. 2010) with bands centred at 70, 100, and 160μm, and the Spectroscopic and Photometric Imaging Receiver (SPIRE; Griffin et al. 2010) with bands centred at 250, 350, 500μm. Combined, the Herschel instruments cover a broad range of FIR wavelengths, and combine high sensitivity and large aperture (3.5m diameter mirror), with an aim to reduce confusion levels compared to progenitor observatories. *Herschel* provided deep survey field observations that have allowed us to resolve the CIB (e.g., Berta et al. 2011; Magnelli et al. 2013; Béthermin et al. 2012), and place constraints on the star-forming properties of distant galaxies (see review by Lutz 2014) .

1.2.2 The SFR of galaxies across cosmic time

A multitude of studies of extragalactic star formation at a range of wavelengths have traced the SFR density (i.e., the total integrated SFR of the observable universe at a given redshift averaged over the corresponding volume, in units of M$_\odot$/yr/Mpc^3) up to redshifts of $\sim$ 10 (e.g., Rodighiero et al. 2010; Karim et al. 2011; Cucciati et al. 2012; Burgarella et al. 2013 see also Madau & Dickinson 2014). The SFR density increases with redshift, with a peak at $z \sim 2$ and then a decline at $z \sim 3$ (see Fig. 1.2). An increase is also observed for the galactic molecular gas fractions increasing from 10% of the mass of local galaxies (e.g., Leroy et al. 2008) to 60% at z$\sim$3 (e.g., Tacconi et al. 2006; Daddi et al. 2010; Geach et al. 2011).

The total IR luminosity function in the local universe, is well fit by a Schechter (1976) function, with a break at IR luminosities of $L_{IR} \simeq 10^{11} L_\odot$, that increases rapidly with redshift (Le Floc'h et al. 2005; Gruppioni et al. 2010, Vaccari et al. 2010, Dye et al. 2010, Eales et al. 2010b, Lapi et al. 2011). Luminous IR galaxies with IR luminosities of $L_{IR} > 10^{11} L_\odot$ (known as LIRGs) make up 2%–5% of galaxies of the local IR luminosity function (Sanders & Mirabel 1996), but as redshift increases so does their number and for redshifts $z > 1$ they become the dominant contributor to the IR luminosity function (e.g. Chary & Elbaz 2001; Le Floc'h et al. 2005), with even more luminous galaxies ($L_{IR} > 10^{12} L_\odot$ known as ultra-luminous IR galaxies; ULIRGs) becoming dominant at redshifts $z > 2$ (e.g. Blain et al. 1999; Magnelli et al. 2013; Casey et al. 2013).

Dust properties of the average galaxy at different redshifts show that high-z LIRGs and ULIRGs have similar features to the normal star-forming galaxies at $z = 0$ (e.g., SF region sizes, the equivalent width of polycyclic aromatic (PAH) lines, FIR SED shapes;

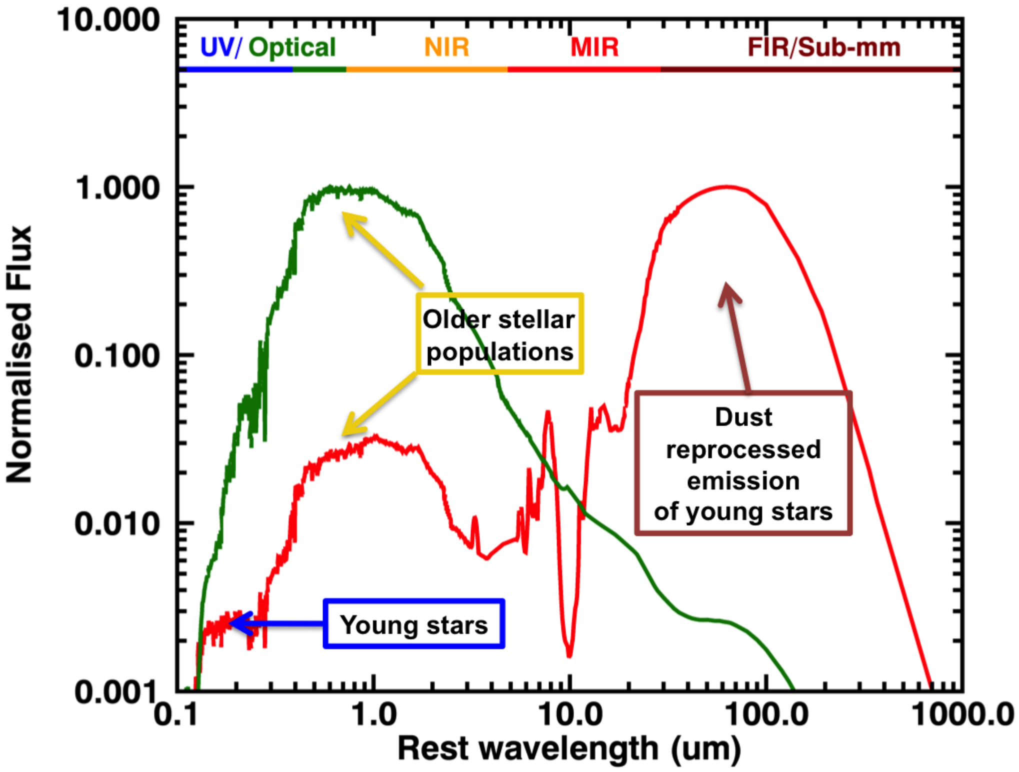


Figure 1.1: Example SEDs of a dusty, highly star-forming galaxy (red) as opposed to a massive elliptical with minimal levels of dust in its ISM (green). Both SEDs are normalised to their peak emission, for the comparison of their relative shapes. The dusty star forming galaxy SED is dominated by the emission from massive young stars (direct at UV/optical and reprocessed in MIR–FIR wavelengths) and shows a peak of emission in the FIR regime. The SED of the elliptical galaxy is dominated by older stellar populations emitting in the optical and NIR waveband, and has no signatures of significant ongoing star formation. For galaxies at higher redshifts, where the majority of star formation is taking place in dusty galaxies, it is necessary to use FIR observations to successfully retrieve the majority of the ongoing star formation. The templates were taken from the SWIRE template library, but both examples were created by Silva et al. (1998)

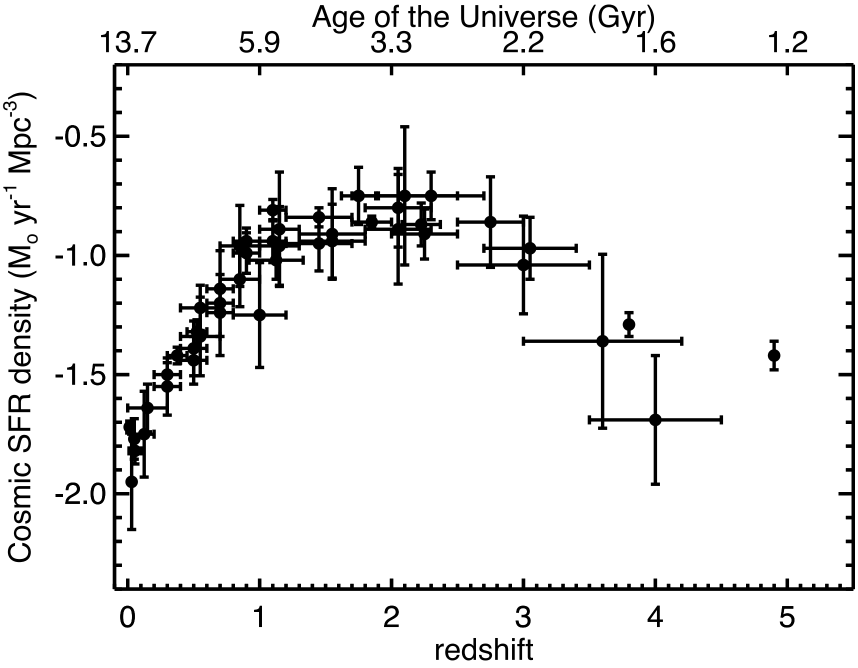


Figure 1.2: Volume averaged cosmic SFR density as a function of redshift, from the compilation by Madau & Dickinson (2014) (see references there-in), demonstrating the evolution of star formation throughout cosmic time. The star-forming activity of galaxies peaks at redshifts of $z \sim 2$, where we also observe high fractions of molecular gas in galaxies, up to 60% of the galaxy masses (e.g, Geach et al. 2011; Béthermin et al. 2014).

see Lutz 2014 and references there-in). Furthermore, there is a tight correlation between the SFRs and M_* of galaxies, that remains up to redshifts of ~4 (this is often called the "main sequence" of star forming galaxies; e.g., Noeske et al. 2007; Daddi et al. 2007a; Rodighiero et al. 2010; Elbaz et al. 2011; Schreiber et al. 2015). The tightness of the correlation suggests a universal star formation inducing process, such as stable secular star-forming histories (e.g., Noeske et al. 2007), rather than a stochastic external trigger such as galaxy mergers (e.g., Conselice et al. 2003). Figure 1.3 shows recent IR constraints on the main sequence of star-forming galaxies as a function of redshift, and it highlights the dependence of SFR values on the mass and redshift of the galaxies, something that can introduce a bias if not taken into account.

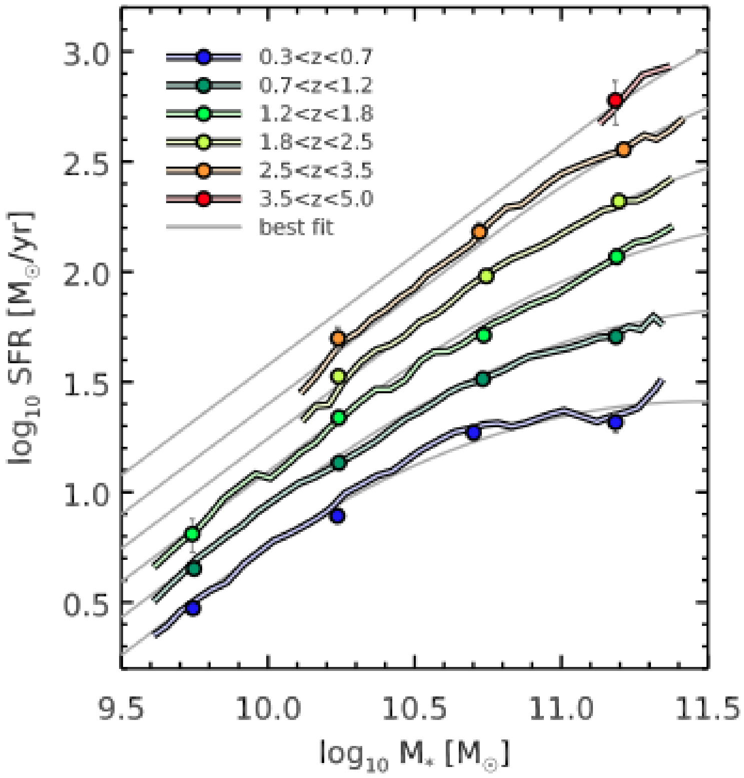


Figure 1.3: Recent IR constraints on the main sequence of star-forming galaxies by Schreiber et al. (2015). Plotted are mean SFR versus stellar mass (M_*) in bins of redshift. The term of "main sequence" was first used by Noeske et al. (2007) when finding this tight correlation between UV to MIR derived SFRs and M_*, up to redshifts of $z \lesssim 1$, and was extended to the FIR using *Herschel* observations (e.g., Rodighiero et al. 2010; Elbaz et al. 2011). The tight correlation and its increase with redshift has been seen since the first of these studies (e.g., Noeske et al. 2007), and indicates the dependence of SFR values on the stellar mass and redshift of the galaxies.

1.3 Active Galactic Nuclei: Growing black holes at centres of galaxies

In the late 20th century super-massive black holes (BHs; with masses of $\mathrm{M_{BH}} \sim 10^5 - 10^{10} M_\odot$) where successfully identified in the centres of a multitude of galaxies in the local Universe (e.g., Sargent et al. 1978; Tonry 1984; Dressler & Richstone 1988; Harms et al. 1994; review by Kormendy & Richstone 1995), and are now believed to reside in all massive galaxies. Around the same time a consensus was reached for the mechanism behind the extreme optical and radio properties found in a small fraction of galaxies (e.g., strong and broad high-excitation emission-line profiles, as well as the presence of forbidden lines, Slipher 1917, Seyfert 1943; extremely bright radio sources in association with optical point sources, Matthews 1963). After extensive debate and research, it became apparent that a mechanism of mass accretion onto a BH was the most likely origin of these extreme sources in the Universe (Salpeter 1964; Zel'dovich & Podurets 1964; Lynden-Bell 1969; Rees 1984). These sources are now known as Active Galactic Nuclei (AGN). AGN can be subdivided into classes or types depending on their observed properties and wavelength at which they have been identified (see 1.3.2).

1.3.1 Mechanics of BH growth

Most BH growth occurs during AGN phases[1] where material (e.g. cold gas, dust) has been funnelled to the centre of the galaxy and lies within the gravitational potential of the BH, forming a disk surrounding the inner $\simeq$pc scales (see section 2.1 in Alexander & Hickox 2012). Due to the shear stress within the disk, the accreting material heats and emits high-energy photons. The overall accretion luminosity can be calculated from the following equation:

$$L_{acc} = \in \dot{m} c^2 \tag{1.3.1}$$

---

[1] Although AGN phases can describe the growth of BHs it can not describe its origin. Three candidate mechanisms for the formation of the original BHs (or BH seeds) are that (a) they are remnants of the first generation (population III) of stars, (b) they are a result of direct collapse of primordial gas clouds, (c) they are a result of runaway collisions in dense stellar clusters (see reviews Volonteri 2010; Alexander & Hickox 2012). An additional driver of BH growth is through BH mergers, although to date it has been observed only for stellar mass black holes (Abbott et al. 2016).

where $\in$ is the mass–energy efficiency conversion, $\dot{m}$ is the mass accretion rate, and c is the speed of light. The value of $\in$ is canonically assumed to be $\simeq 0.1$, but the value can range between 0.05–0.42 depending on the spin of the BH (e.g., Kerr 1963; Thorne 1974).

There is a physical limit on the luminosity that the AGN can emit before radiation pressure overcomes the gravitational force of the BH and stops the accretion. This is referred to as the Eddington luminosity and can be calculated assuming the spherical accretion of fully ionised hydrogen onto a BH of mass M_{BH}. The Eddington luminosity is that for which the radiation pressure is equal to the force of gravity, and can be expressed as:

$$L_{Edd} = \frac{4\pi G M_{BH} m_p c}{\sigma_T} \simeq 1.3 \times 10^{38} (M_{BH}/M_{\odot}) \text{ erg s}^{-1} \tag{1.3.2}$$

where m_p is the proton mass, G the gravitational constant, and σ_T the Thompson cross-section for electrons. Due to the fact that accretion is happening in a disk rather than spherical geometry (assumed for the above equation), it is possible for AGN to achieve luminosities a factor of a few above the Eddington limit.

A useful quantity for defining how powerful an AGN is relative to the Eddington luminosity (and therefore the mass of the BH) is the Eddington ratio:

$$\lambda_{Edd} = \frac{L_{acc}}{L_{Edd}} \propto \frac{\dot{M}}{M_{BH}} \tag{1.3.3}$$

1.3.2 Tracing the AGN power output

Figure 1.4 shows a schematic of the standard model for AGN (Antonucci 1993; Urry & Padovani 1995; Netzer 2015). The standard model was developed to explain the observed spectral properties of different types of AGN and their average SEDs, and is a good starting point for understanding the emission produced by AGN. The SED of AGN covers a broad range of wavelengths, from hard X-rays to radio. A schematic for a typical AGN SED is shown in Figure 1.5.

Surrounding the BH on sub–pc scales there is a rotational dominated accretion flow (accretion disk) that emits in the optical and UV (blue curve in Fig. 1.5). At scales of 0.1–10pc an axisymmetric dusty structure (the torus) encircles the accretion disk and BH. The structure of the torus is a matter of ongoing research, but is currently widely believed to be clumpy (e.g., Hönig et al. 2006; Nenkova et al. 2008). The optical/UV emission from the accretion disk that penetrates the torus will be absorbed by the dust and re-radiated in the IR wavelengths (e.g., Pier & Krolik 1992, Rowan-Robinson 1995; Polletta et al. 2000; red curve in Fig. 1.5). A portion of the accretion disk photons will go through the hot corona situated above and below the accretion disk, and due to inverse-Compton scattering produce powerful X-ray continua (cyan curve in Fig. 1.5). The X-ray photons may also reflect off the torus and/or the accretion disk to produce an additional reflection component in the X-ray region (e.g., George & Fabian 1991; green curve in Fig. 1.5). An additional component of the X-ray emission is what is referred to as "soft excess", emission at the soft X-ray frequencies that exceeds what is expected by the simple accretion disk model, and the origin of which is still under research (see Done et al. 2012 and references there-in). The emission of AGN at radio wavelengths varies between sources. Luminous radio galaxies (radio-loud) have radio emission produced by relativistic jets, launched close to the accretion disk, as well as from extended radio structures. But radio weak AGN (usually classified as radio-quiet) tend to show a more compact origin of the radio emission, the nature of which is not clear but can be attributed to supernovae, compact radio jets and a corona close to the accretion disk (see e.g., Smith et al. 1998; Polletta et al. 2000; Laor & Behar 2008; and references there-in).

The emission of an AGN in all these wavebands has led to a variety of identification methods, with each wavelength finding different populations of AGN with small overlap between them. This is due to different properties of the AGN and/or the line of sight at which the AGN is observed.

In this section I give a brief overview of some of the different identification approaches for AGN, but note that in this thesis we primarily select AGN on their optical and X-ray properties.

X-ray identification of AGN

X-ray observations provide one of the cleanest selections of AGN, suffering little from the presence of obscuration and contamination from the host galaxy. This is due to the fact that AGN emission in the X-ray is much stronger than that of the host galaxy and at energies of $>$2keV it can penetrate material of high column densities, of $N_H \sim 10^{22.5} cm^{-2}$,

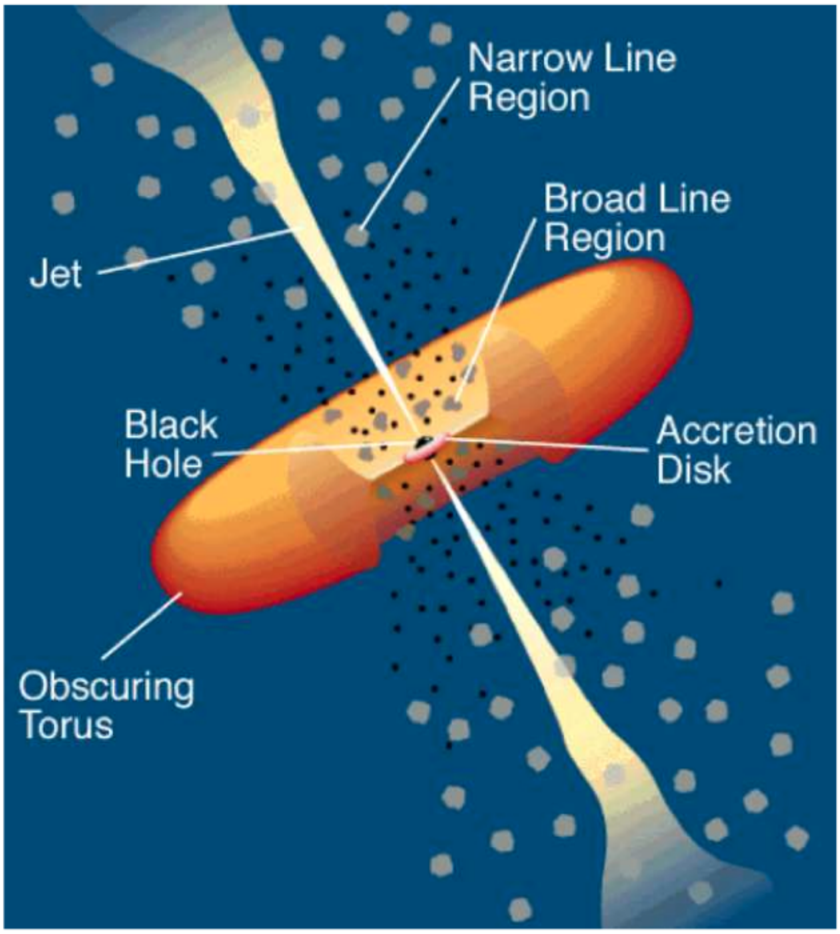


Figure 1.4: Schematic of the central structure of AGN, adapted from Urry & Padovani (1995). The central BH is surrounded by an accretion disk, which in turn is surrounded by a geometrically and optically thick dusty structure called the “torus”, now believed to be of a clumpy nature (e.g., Hönig et al. 2006; Nenkova et al. 2008). Two distinct emission line regions are present above and below the accretion disk: (1) the broad-line region (BLR) that is under the gravitational influence of the BH and (2) the narrow-line region (NLR) that is extended over much larger scales. Radio jets may also be launched from close to the accretion disk. The observed emission from the AGN depends on the inclination angle of the torus with respect to the observer. For example, the obscuration by the dusty torus will prevent the optical emission from the accretion disk and BLR from being directly observable, while the emission from the NLR, torus is less dependent on the line of sight (see also resent reviews of Heckman & Best 2014, and Netzer 2015).

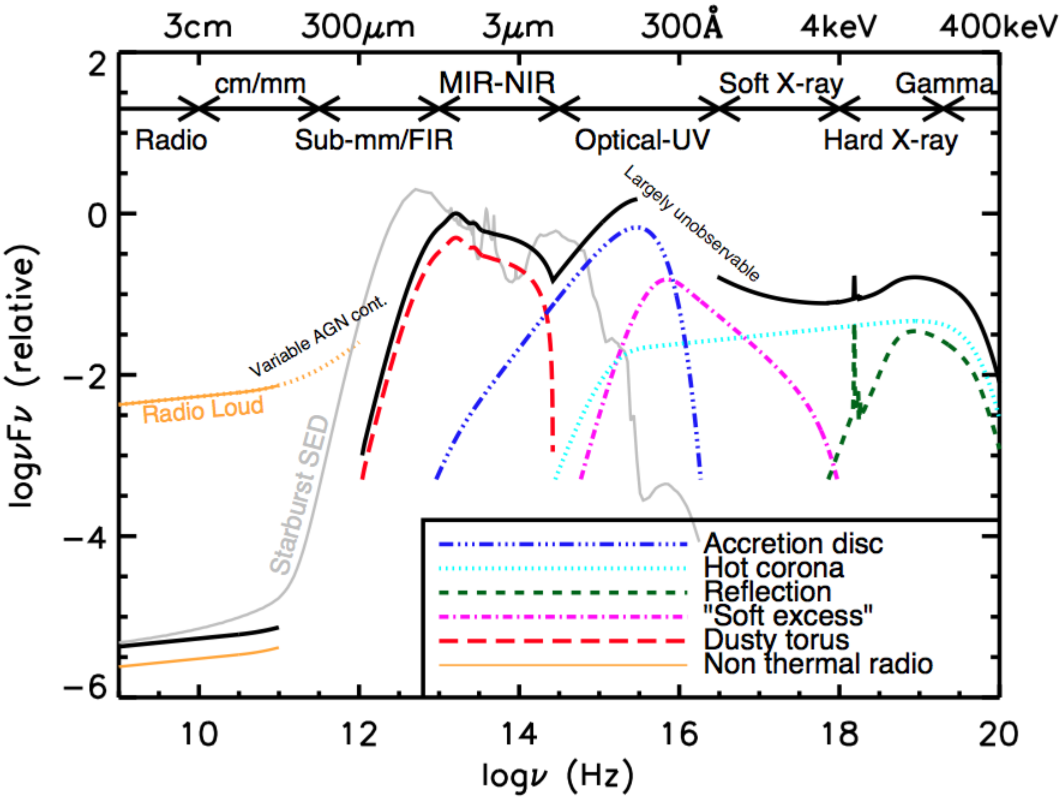


Figure 1.5: A schematic representation of an AGN SED. The shape of the SED is loosely based on the observed SEDs of radio-quiet quasars (e.g., Elvis et al. 1994; Richards et al. 2006). The black solid curve represents the total SED and the various coloured curves (with an arbitrary offset) represent the individual components (see Section 1.3.2). The SED of an AGN in the mm–FIR regime is widely believed to have a minimal contribution (to an overall galaxy SED) compared to star formation, except in the most intrinsically luminous quasars and powerful radio-loud AGN. The primary emission from the AGN accretion disk peaks in the UV region. Radio-loud AGN have radio emission that can be several orders of magnitude higher than radio-quiet AGN (shown with the labelled orange line). Also shown is an example radio–UV SED of a star-forming galaxy (grey curve; the SED is of M82 taken from the GRASIL library; Silva et al. 1998). Image credit: C. M. Harrison

increasing with increasing energies (e.g. Wilms et al. 2000; Mushotzky et al. 2000; Tozzi et al. 2006; Gilli et al. 2007).

The two most sensitive X-ray observatories at frequencies up to $<$10keV operating today are the Chandra X-ray Observatory (Chandra; e.g. Weisskopf 1999), and the X-ray Multi-Mirror Mission (XMM-Newton; e.g. Jansen et al. 2001), with X-ray energy bands of 0.3-8 keV and 0.2-12 keV, respectively. Using these observatories there have been numerous deep field surveys that provide large samples of AGN over a wide range of redshifts, with excellent source positions and point spread functions, and benefiting from a near-complete selection for distant AGN, except for the most heavily obscured systems (e.g., Brandt et al. 2001; Giacconi et al. 2002; Alexander et al. 2003b; Luo et al. 2008; Xue et al. 2011; see review by Brandt & Alexander 2015).

For the selection of AGN in the X-ray it is common to use an X-ray hard band (2–8keV or 2–10keV) luminosity cut of $\mathrm{L_{HB}} \gtrsim 10^{42}$ erg s^{-1}. This selection method is used to select the AGN samples for Chapters 3 & 5. The majority of sources with such luminosities are AGN, with the majority of star-forming galaxies having $\mathrm{L_{HB}} \lesssim 10^{42}$ erg s^{-1} (e.g., Alexander et al. 2005; Laird et al. 2010; Lehmer et al. 2010; Symeonidis et al. 2011; Wang et al. 2013).

Optical identification of AGN

Optically luminous AGN, were first detected as star-like objects with unique spectra displaying broad emission lines in the optical, and UV excess emission. These sources are known as quasi-stellar-objects (QSOs) and are some of the most luminous AGN, outshining their hosts in the UV and optical wavelengths.

The selection of large numbers of AGN in the optical waveband is facilitated by multiple large surveys, such as the Sloan Digital Sky Survey (SDSS; Abazajian et al. 2009), 2 Micron All-Sky Survey (2MASS; Skrutskie et al. 2006), 2 Degree Field Galaxy Redshift Survey (2dF; Colless et al. 2001), that due to their sheer size provide numerous detections of QSOs of up to redshifts of $z \sim 6$. SDSS alone has produced a large homogeneous sample of optical quasars, the most luminous unobscured AGN (type-1), with more than 150,000 quasars (e.g. Richards et al. 2002, 2004; Schneider et al. 2007, 2010; Pâris et al. 2014). The optical selection of quasars is mostly based on optical colour, using for example the *ugriz* bands in SDSS surveys (e.g., Richards et al. 2002), and can achieve high levels of completeness for optical quasar samples at AGN bolometric luminosities of $L_{AGN}>10^{45}$ erg s^{-1}. We use such samples of optically identified QSO samples in Chapter 4. However, optical colour selection fails to pick up the obscured AGN population (type-2), and low luminosity/accretion rate AGN, that have dominant host galaxy contributions to their optical colours (e.g. Hopkins et al. 2009).

IR identification of AGN

As described earlier, a fraction of the accretion disk photons are absorbed by the gas and dust enshrouding the disk. This fraction can be large depending on the torus properties, making it difficult for even the X-ray emission to be observed (at column densities of $N_H > 10^{24.5} cm^{-2}$; e.g., Treister et al. 2004; Daddi et al. 2007b; Alexander et al. 2008). The absorbed emission is reradiated in the rest-frame NIR–MIR and longer wavelengths (see red curve in Fig. 1.5; e.g., Sanders et al. 1989; Netzer et al. 2007; Mullaney et al. 2011; Mor & Netzer 2012). The shape of the AGN emission SED in the IR is different to that of stellar and star-forming emission (see Fig. 1.5), which means that NIR and MIR colours can be used to select AGN, as the emission from the AGN populates a part of the colour space well separated from the other sources of MIR emission (e.g., Lacy et al. 2004; Stern et al. 2005; Donley et al. 2007). Based on this separation and available photometry from *Spitzer*/IRAC, and the Wide-Field Infrared Survey Explorer (WISE; Wright et al. 2010), a multitude of AGN selection criteria have been developed (e.g., Stern et al. 2005; Donley et al. 2012; Mateos et al. 2012; Stern et al. 2012; Eisenhardt et al. 2012; Assef et al. 2013). Although these methods have low contamination from non-AGN galaxies, they tend to miss a number of AGN where the IR emission due to the AGN is diluted by the IR emission due to star formation in the galaxy (e.g., Cardamone et al. 2008; Assef et al. 2013; Kirkpatrick et al. 2013). Consequently, IR colour selection provides a non-uniform selection of AGN.

Radio identification of AGN

Identification of luminous AGN through their radio emission (usually at 1.4GHz and 5GHz) can be very powerful, as most radio luminous sources are AGN, with rapidly star-forming galaxies only contributing significantly to the lower luminosities ($L_{1.4GHz} < 10^{24}\,W\,Hz^{-1}$). Radio-luminous AGN represent an extreme state of the AGN phase, with large luminosities ($L_{1.4GHz} > 10^{24}\,W\,Hz^{-1}$) due to core emission from the AGN and large-scale radio jets and lobes. Only 10% of AGN are luminous in the radio ($L_{1.4GHz} > 10^{24}\,W\,Hz^{-1}$), and the vast majority of this population are not identified in the optical (see Alexander & Hickox 2012). So even though radio surveys only sample a small fraction of the overall AGN population, they are still essential to identify these extreme states of AGN.

Radio-luminous AGN tend to be divided into two classes depending on their optical spectral properties, low-excitation (Liners; LERGs) and high-excitation (quasars; HERGs). Both populations are found across the full range of radio luminosities, but dominate at different luminosities. The low-excitation radio AGN (hereafter LERGs) typically have low-accretion rates and are the dominant population at modest radio luminosities ($L_{1.4GHz} < 10^{26}\,W\,Hz^{-1}$; e.g., Best & Heckman 2012), while high-excitation radio AGN (hereafter HERGs) have high accretion rates and the highest radio luminosities ($L_{1.4GHz} > 10^{26}\,W\,Hz^{-1}$; e.g., Best & Heckman 2012; see also Alexander & Hickox 2012).

1.3.3 BH growth across cosmic time

The first evidence of the higher levels of BH growth in the distant universe came with the first optical QSO surveys that found that at redshifts of $z \sim 1-2$ they were much more common than at $z \sim 0$ (e.g. Schmitt 1968; Schmidt & Green 1983). This was later confirmed by more detailed studies of the luminosity density of AGN from different selections (e.g., optical, X-ray, radio). The most complete AGN samples, selected in the X-rays, show a luminosity density that peaks at redshifts of $z \sim 1$, with a steep drop-off to lower redshifts and a more gradual drop-off at higher redshifts (see Fig. 1.6; e.g. Fiore et al. 2003; Shankar 2009; Barger & Cowie 2005; Hasinger et al. 2005; Silverman et al. 2008; Brusa et al. 2009; Aird et al. 2010, 2015). The luminosity density as a function of redshift is also dependent on AGN luminosity (see Fig. 1.6; e.g., Aird et al. 2010, 2015), with AGN of $10^{42} < \frac{L_{2-10keV}}{\mathrm{erg\,s^{-1}}} < 10^{43}$ being predominant at $z \sim 0.5$, $10^{43} < \frac{L_{2-10keV}}{\mathrm{erg\,s^{-1}}} < 10^{44}$ being predominant at $z \sim 1.5$, $10^{44} < \frac{L_{2-10keV}}{\mathrm{erg\,s^{-1}}} < 10^{45}$ being predominant at $z \sim 2$, and

$10^{45} < \frac{L_{2-10keV}}{\mathrm{erg\ s^{-1}}} < 10^{46}$ being predominant at $z \sim 3$.

The overall shape of the X-ray luminosity density with redshift is also traced by other AGN populations, such as the optical QSOs (e.g., Richards et al. 2006; Croom et al. 2009; Assef et al. 2011), the radio selected AGN (e.g., Rigby et al. 2008; Smolčić et al. 2009), and obscured AGN (e.g., La Franca et al. 2005; Ballantyne et al. 2006; Treister & Urry 2006; Hasinger 2008).

Overall, the majority (∼95–99%) of the integrated BH growth occurs at $z > 0.1$ (e.g., Marconi et al. 2004; Shankar et al. 2004; Shankar 2009; Hopkins et al. 2007; Merloni & Heinz 2008; Aird et al. 2010, 2015), in luminous AGN of $L_{2-10keV} \simeq 10^{43} - 10^{45}$ erg s^{-1}, that cover the knee of the AGN luminosity function (e.g. Hopkins et al. 2007; Aird et al. 2010, 2015), and are radio-quiet (See also section 3.3 of Alexander & Hickox 2012).

1.4 A connection between the two processes of AGN and star formation

Both AGN and star formation are strongly dependent on the cold gas supply within the galaxy, so we may expect at least a first order connection through their dependence on the same fuel supply, although AGN only depend on the gas supply in and around the accretion scale. Studies of the local and distant galaxies show evidence for such a connection between AGN activity and star formation, with observational results supporting a co-evolution of galaxies and their central BHs (see section 1.4.1). However, the size scales of AGN activity (sub-pc scales) and star formation (over kpc scales) are very different, which has lead to suggestions that any tight connection between them must be due to one process regulating the other. Indeed, for models of galaxy evolution to reproduce the observable properties of galaxies, they find it necessary for AGN to have a more regulatory role over the gas within the inter-stellar medium (ISM) by heating or expelling large amounts of cold gas. Consequently, they will also have a regulatory role on themselves and the star formation of their host galaxies (see Alexander & Hickox 2012, Fabian 2012, and Kormendy & Ho 2013 for recent reviews). There is now a significant amount of research seeking more direct evidence for a connection between the AGN and star formation of galaxies. In this thesis I investigate the relationship between the star formation

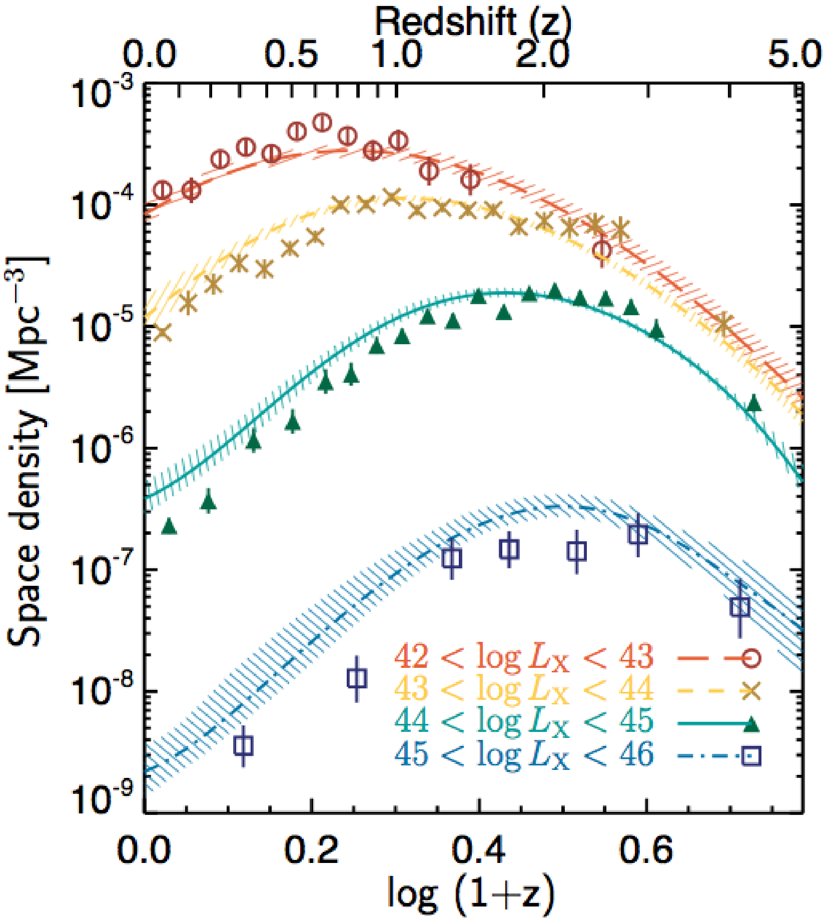


Figure 1.6: Total space density of AGN for different ranges of Hard-Band (2–7keV) X-ray luminosity based from Aird et al. (2015). The data points are taken from Miyaji et al. (2015), and the shaded region is the model parametrisation of Aird et al. (2015). This shows that higher luminosity AGN peak in terms of their space density at higher redshifts than the lower luminosity AGN. Furthermore, at the redshifts coinciding to those of the peak of the SFR density ($z \sim 2$) the luminosity density is dominated by moderate to high X-ray luminosity ($10^{44} < \frac{L_{2-7keV}}{\mathrm{erg\ s^{-1}}} < 10^{46}$) AGN.

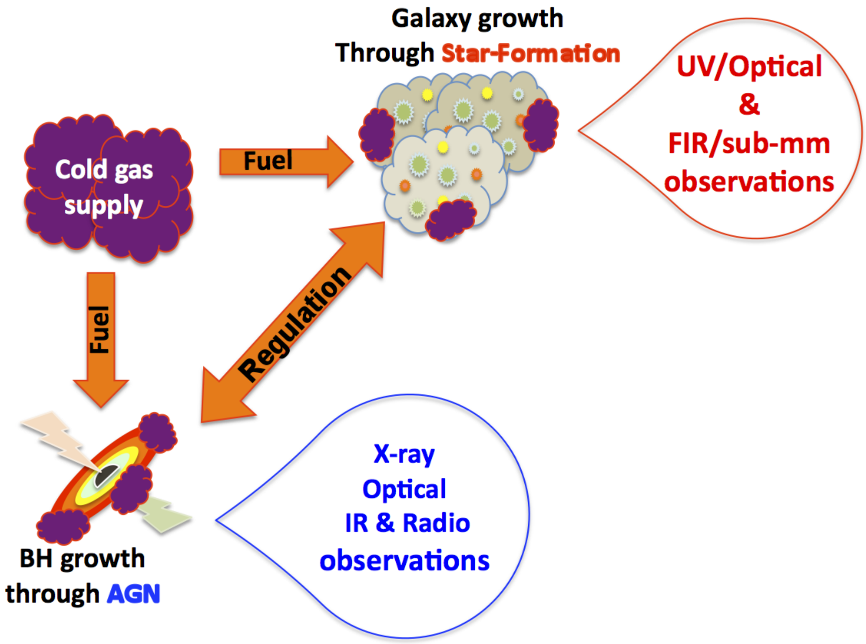


Figure 1.7: A schematic diagram to demonstrate the connection between the two processes of BH growth through AGN activity and galaxy growth through star formation. Both processes are fuelled by the cold gas supply of the galaxy, giving them an inherent first order connection. However, due to the difference in the scales of the BH and the galaxy ($<$ 0.1pc compared to kpc scales) and based on results from computer simulations of galaxy evolution, we would expect a more interactive relationship between the two processes. Studies trying to identify such a connection use a range of wavelengths for identifying and measuring the AGN and star-forming activity, outlined above (see sections 1.2.1 and 1.3.2).

rate (SFR) and AGN luminosity.

1.4.1 The co-evolution of BHs and Galaxies

There are several pieces of empirical evidence for at least a broad connection between AGN activity and star formation. For example, the tight correlation observed between the mass of the central BH and the galaxy spheroid for galaxies in the local universe (Fig. 1.8; as well as other properties of the bulge such as the velocity dispersion, and the bulge luminosity; e.g., Kormendy & Richstone 1995, Magorrian et al. 1998), serves as archaeological evidence for a connection between the growth of the BH (through mass accretion, where it becomes visible as an AGN), and the growth of the galaxy (through star formation). Additionally, observations of AGN have found that the volume average of

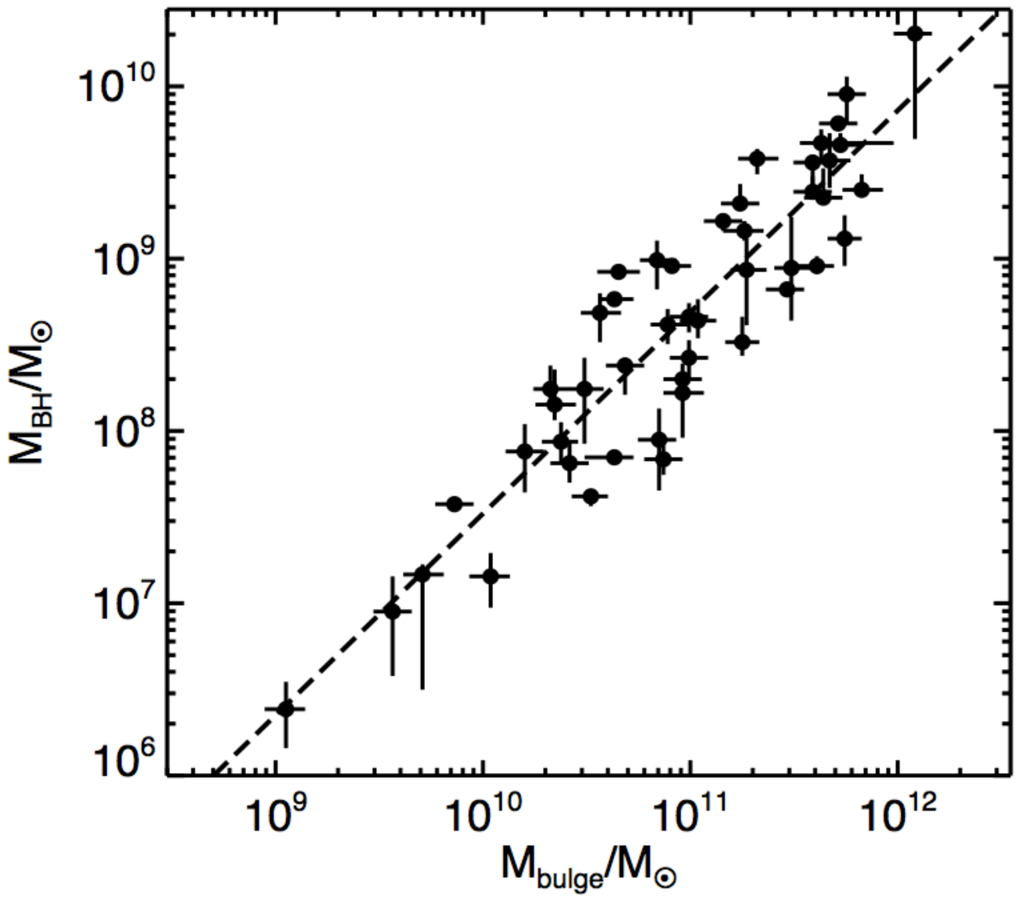


Figure 1.8: BH mass (M_{BH}) as a function of bulge mass (M_{bulge}) for a sample of ellipticals and classical bulges, from Kormendy & Ho (2013). The correlation between M_{BH} and M_{bulge} is suggestive of a co-evolution between galaxies and their central BHs. Since the AGN phase is the result of active BH growth and star formation is a driver of galaxy growth, we would also expect at least a broad connection between these two properties.

the BH mass accretion rate tracks that of the SFR, with a normalisation of ∼3–4 orders of magnitude, up to redshifts of $z \sim 2$ (see Fig. 1.9; e.g., Heckman et al. 2004; Merloni et al. 2004; Aird et al. 2010) suggesting a co-evolution of AGN and star formation. Despite how significant these results may appear, they only provide indirect evidence for a relationship between AGN activity and star formation and cannot place strong constraints on the form of the relationship.

The most successful models of galaxy formation and evolution require AGN activity (via "AGN feedback") to explain many of the puzzling properties of local massive galaxies and the intergalactic medium (IGM): red colours, steep luminosity function, black hole to spheroid relationship, metal enrichment of the IGM (see Cattaneo et al. 2009, Alexander

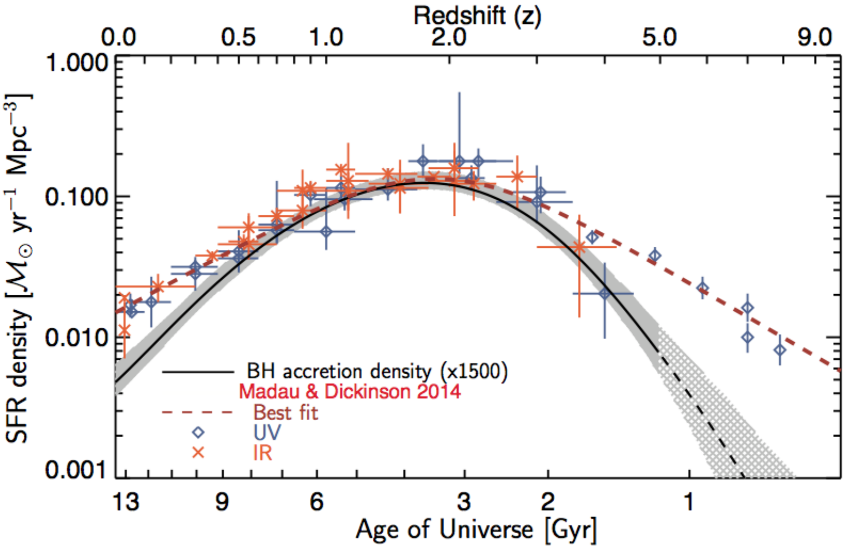


Figure 1.9: Volume averaged cosmic SFR density as a function of redshift, with a representative model to the observed evolution of AGN space density (grey region) scaled by ≃1500 (taken from Aird et al. 2015). The data points are from the compilation by Madau & Dickinson (2014) (see references there-in) and the model of AGN space density has been derived by Aird et al. (2015) from observational data (see Fig. 1.6).

& Hickox 2012, and Fabian 2012 for reviews). The key attribute of the AGN in these models is the injection of significant energy into the interstellar medium (ISM), which inhibits or suppresses star formation (SF) by either heating the ISM or ejecting the gas out of the host galaxy through outflows (e.g., Di Matteo et al. 2005; Hopkins et al. 2005; Debuhr et al. 2012; Schaye et al. 2015; Lacey et al. 2015).

To better understand the connection between AGN activity and star formation, and to investigate the potential impact that AGN may have on the star formation of their host galaxy, it is necessary to individually constrain the SFRs and AGN luminosities in AGN and galaxy samples, and investigate the relationship of the two processes.

1.4.2 What is the relationship between star formation and AGN activity?

There have been numerous studies that have concentrated on constraining the SFR and AGN luminosity properties, and to investigate the relationship between star formation and AGN activity. There are two ways most studies go about this, (a) select an AGN sample through the optical, X-ray, IR, and/or radio, and constrain the average SFRs of the sample, or (b) select samples of star-forming galaxies and constrain the average AGN luminosities of the sample. In this section I give a brief overview of the latest results using method (a) for samples selected through X-ray, optical, and radio properties, which is also the approach I followed in this thesis. However, its worth noting that the observed relationship between the AGN luminosity and SFR from the two methods tend to differ due to the fact that in (a) we are averaging over the more variable quantity (i.e., the AGN luminosity) than in (b), which is discussed in more detail in Chapter 3.

A number of studies of X-ray AGN have used FIR observations to measure their star-forming properties (e.g., Lutz et al. 2010; Shao et al. 2010; Mainieri et al. 2011; Mullaney et al. 2012a; Rovilos et al. 2012; Santini et al. 2012; Harrison et al. 2012; Rosario et al. 2013c,a; Lanzuisi et al. 2015). The main results shown by these studies are that: (1) the average star formation rates ($\langle$SFR$\rangle$) of AGN track the increase with redshift found for the overall star forming galaxy population; (2) the $\langle$SFR$\rangle$ of AGN are higher than those of the overall galaxy population (i.e., when including quiescent galaxies); and (3) the specific SFRs (i.e., the ratio of SFR over stellar mass, which serves as a measure of the relative growth rate of the galaxy) of AGN are in quantitative agreement with those of star forming galaxies. The majority of the current studies also find no correlation between the AGN luminosity and $\langle$SFR$\rangle$ for moderate luminosity AGN (X-ray luminosities of $\mathrm{L_{2-8keV}} \lesssim 10^{44}$ erg s^{-1}; e.g., Lutz et al. 2010; Shao et al. 2010; Mullaney et al. 2012a; Rovilos et al. 2012; Harrison et al. 2012). However, there is significant disagreement in the results for high luminosity AGN ($\mathrm{L_{2-8keV}} \gtrsim 10^{44}$ erg s^{-1}; see Fig. 1.10), with studies finding either an increasing (e.g., Lutz et al. 2010; Rovilos et al. 2012; Santini et al. 2012), decreasing (e.g., Page et al. 2012; Barger et al. 2014), or broadly flat relationship of $\langle$SFR$\rangle$ with X-ray luminosity (e.g., Harrison et al. 2012; Rosario et al. 2012; Azadi et al. 2015). The origin of the differences in the conclusions of these studies has been shown to be at least partly driven by low source statistics for high luminosity AGN (see Harrison et al. 2012) that can strongly affect single-band derivations of the SFR (see section 2.6). In this thesis I present work that has overcome the limitations of these previous studies (Chapter 3; Stanley et al. 2015). Not only do I concentrate on large samples of AGN, but but I also take into account several other effects often neglected in previous studies (see section 1.5).

The hosts of optically selected, unobscured QSOs also show signs of ongoing star formation (e.g., Jahnke et al. 2004; Trump et al. 2013; Rosario et al. 2013b). The SFRs of QSOs can range between a few $M_{\odot}$/yr at the local universe to 100s–1000s $M_{\odot}$/yr at higher redshifts (e.g., Schweitzer et al. 2006; Netzer et al. 2007, Omont et al. 2003). FIR studies of the $\langle \mathrm{SFR} \rangle$ as a function of AGN luminosity of unobscured QSOs find consistent results to those of X-ray AGN (e.g., Rosario et al. 2013b; Netzer et al. 2015), and appear to lie above (but not significantly) the main sequence of the star-forming population (e.g., Rosario et al. 2013b).

When looking at radio-luminous AGN, representing a potentially high-impact stage of AGN feedback, there is evidence that low accretion rate systems at low redshift ($z < 0.5$) are responsible for regulating the inflow of cool gas in massive galaxies and massive galaxy clusters via the heating from radio jets and lobes (e.g., Best et al. 2005; McNamara & Nulsen 2012; Heckman & Best 2014). However, at higher redshifts it has been found that radio-luminous AGN can reside in strongly star-forming galaxies. There are positive trends between $\langle \mathrm{SFR} \rangle$ with increasing redshift (e.g., Reuland et al. 2004; Magliocchetti et al. 2014; Kalfountzou et al. 2014; Podigachoski et al. 2015), broadly tracing that found for both the star-forming galaxy and X-ray AGN populations (e.g., Mullaney et al. 2012a; Rosario et al. 2012). The $\langle \mathrm{SFR} \rangle$ of the overall radio-luminous AGN show no significant difference to radio-weak AGN (e.g., Kalfountzou et al. 2014), but when split into the two populations of HERGs and LERGs, LERGs tend to show lower SFRs to the HERGs and the radio-weak AGN (e.g., Gürkan et al. 2015). However, both optical QSOs and radio-luminous AGN studies also suffer from the uncertainties of single-band derivations of the SFR as well as effects of redshift and stellar mass that have not been taken into account (see section 1.5). In Chapter 4 I build on the previous work and use improved methods to estimate the SFRs of a large sample of optical unobscured QSOs, as well as their radio-luminous sub-sample.

1.5 Thesis overview

The overall aim of this thesis is to place accurate constraints on the average SFRs of AGN samples covering the redshift range of the peak in AGN and SF activity, and a wide range

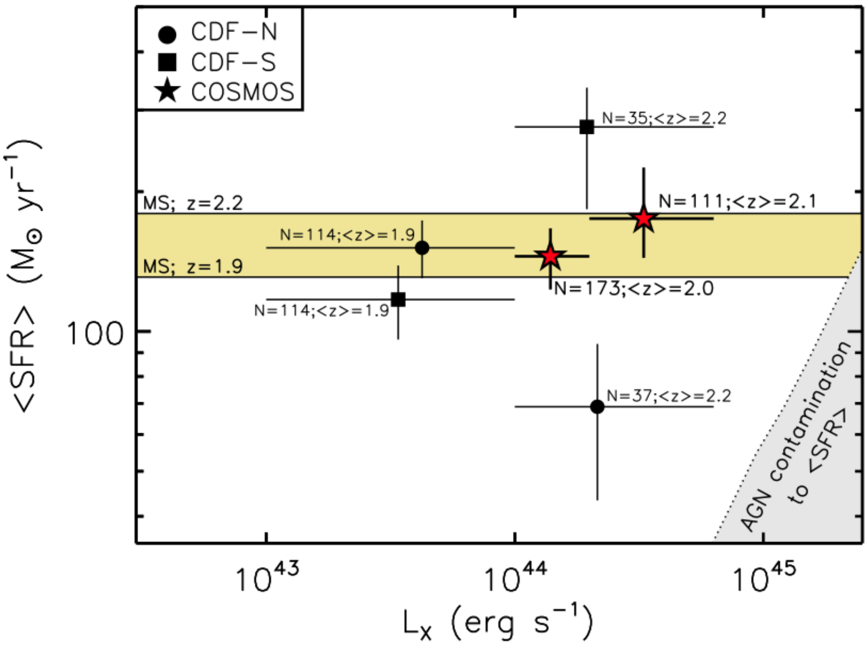


Figure 1.10: The mean SFR (from 250μm photometry) as a function of X-ray luminosity of AGN in three fields, from the review of Harrison (2014). Each bin has the corresponding number of sources and mean redshift indicated. The solid horizontal lines show the SFRs for main sequence star-forming galaxies of comparable masses at two redshifts (z =1.9 and z =2.2; Elbaz et al. (2011)). This figure serves as a summary of studies on X-ray AGN, preceding this thesis, of the SFRs of AGN in the three most used deep fields covered by both deep X-ray *Chandra* and deep FIR *Herschel* observations. Overall, the mean SFRs of X-ray AGN is in agreement with what is expected for normal star-forming galaxies; however there is a discrepancy between different fields at the high X-ray luminosities.

in AGN luminosity, to investigate the nature of the connection between the AGN and star-forming activity of galaxies.

Most of the studies on the $\langle\mathrm{SFR}\rangle$ of distant AGN preceding this thesis, suffer from a variety of limitations, which affect the accuracy of $\langle\mathrm{SFR}\rangle$ measurements, such as: (1) the small number of sources, which can lead to large statistical uncertainties, particularly at high AGN luminosities; (2) high levels of source confusion at FIR wavelengths, which can cause an overestimation of source fluxes; (3) use of a single FIR band from which to derive SFRs, which will result in large uncertainties on the $\langle\mathrm{SFR}\rangle$ and will not take into account possible contamination of the SFR measurements from the AGN. In section 2.5 we go further in analysing these limitations. Furthermore, in many cases the results are interpreted without taking into account the redshift and stellar mass dependencies of SFRs. Overlooking these dependencies can lead to ambiguous results.

In my methods I strive to overcome the limitations faced by previous studies, and place the most accurate constraints currently possible on the SFRs of AGN. I make use of X-ray and optical photometry to select large and representative samples of AGN over the redshift range of $0.2< z <2.5$, and spanning the bolometric AGN luminosity range of $10^{43} < \frac{\mathrm{L_{AGN}}}{\mathrm{erg\,s^{-1}}} < 5\times10^{47}$. To derive the SFRs of my AGN samples I take advantage of the FIR photometry from deep *Herschel* surveys and ALMA observations, in combination with MIR photometry from *Spitzer* or WISE, and perform SED fitting and decomposition to remove contamination from the AGN. Furthermore, when interpreting my results I take into account the redshift and stellar mass dependencies on SFRs discussed above.

Throughout this thesis I assume $\mathrm{H_0} = 71\mathrm{km\,s^{-1}}$, $\Omega_\mathrm{M} = 0.27$, $\Omega_\Lambda = 0.73$, and a Chabrier (2003) initial mass function (IMF). When estimating a SFR from the IR luminosity I use the Kennicutt (1998) relation corrected for a Chabrier IMF, $SFR = 2.65\times10^{-44}\,L_{8-1000\mu\mathrm{m}}$.

The chapters included in this thesis are:

- **Chapter 2: Methods**

  In this chapter I describe the methods followed in this thesis for the estimation of individual and mean SFRs for large samples of AGN. An important part of the analysis followed is the SED fitting of IR photometry, and its decomposition into an AGN and star formation component. This provides an IR luminosity due solely to star formation without the contamination from AGN emission. To estimate mean

SFRs two different methods were followed, depending on the sample studied. One uses the Kaplan-Meier product limit estimator to estimate the mean SFR from individual measurements, including upper limits. The other uses stacking techniques in the FIR to calculate the mean flux densities of different bands, and combining them to create a composite SED, which is then decomposed into the AGN and star formation components, to calculate the mean SFR.

- **Chapter 3: A remarkably flat relation between the star formation rate and AGN luminosity for distant X-ray AGN**

  In this chapter I investigate the relationship of the mean SFR as a function of redshift and AGN luminosity for a sample of ~2000 X-ray detected AGN covering the redshifts of $0.2 < z < 2.5$, and three orders of magnitude in X-ray luminosity ($\mathrm{L_{2-8keV}} = 10^{42} - 10^{45.5}\ \mathrm{erg\,s^{-1}}$). Using photometry from *Spitzer* and *Herschel* I use the SED fitting, and Kaplan-Meier product limit estimator, described in Chapter 2, to estimate the mean SFR as a function of redshift and AGN luminosity.

- **Chapter 4: The mean SFRs of luminous unobscured QSOs: searching for evidence of suppressed or enhanced star formation**

  In this chapter I investigate a sample of optically selected QSOs from the SDSS survey within the *Herschel*–ATLAS fields, and a radio-luminous sub-sample defined using FIRST. The sample includes more than 3000 sources covering a wide range of redshift ($z = 0.2$–2.5), and concentrating on high bolometric AGN luminosities of $\mathrm{L_{AGN}} = 10^{45}$–$10^{48}\ \mathrm{erg\ s^{-1}}$. Using *Herschel* and WISE photometry and the stacking, and SED fitting processes described in Chapter 2, I place constraints on the mean star formation rates (SFRs) of unobscured QSOs.

- **Chapter 5: Achieving deeper constraints on the SFRs of X-ray AGN with ALMA 870μm observations**

  In this chapter I investigate the improvements to the SFR measurements when including deep 870μm ALMA photometry. The sample consists of 109 X-ray AGN observed with deep ALMA Band-7 continuum observations, with moderate to high luminosities of $\mathrm{L_{2-8keV}} = 10^{42} - 10^{45}\ \mathrm{erg\,s^{-1}}$, at redshifts $z > 1$. Combined with MIR and FIR photometry from *Spitzer* and *Herschel*, and using the SED fitting methods described in Chapter 2, I quantify the improvements that deep ALMA photometry can achieve in comparison to what we have previously achieved with only *Spitzer* and *Herschel* photometry. Additionally, I demonstrate the utility of the $F_{870\mu m}/F_{24\mu m}$ ratio as a function of redshift, as a parameter space for the identification of AGN and testing of SED templates.

- **Chapter 6: Conclusions and Future work**

  In this chapter I give a summary of the work presented in this thesis, and the key results of each experiment. I also describe current and future experiments based on the methods and results of this thesis.

CHAPTER 2

Methods

2.1 Introduction

This thesis uses infrared (IR) photometric data, covering the wavelengths of 3.6 – 870μm, to measure the SFRs of large samples of galaxies hosting AGN selected from their X-ray, optical, or radio emission. As both star formation and AGN emission can contribute to the observed IR emission of the galaxy, it is important to be able to successfully disentangle the two. An important factor in the following analysis is the available photometry at FIR wavelengths. I describe the MIR and FIR photometry used in our methods in section 2.2. In section 2.3 I describe the IR SED fitting methods used in this thesis to decompose the IR emission into AGN and star formation components. In section 2.4 I describe the method used for the estimation of the mean when the sample includes upper limits. In section 2.5 I describe an alternative method of stacking the IR photometry to provide average IR flux measurements. Finally, in section 2.6 I demonstrate the uncertainties of FIR single-band derivations of the SFR.

2.2 MIR–FIR photometry

As described in detail in Chapter 1 the FIR wavelengths trace the bulk of the reprocessed emission of obscured star-forming regions, making the extent and depth of the available FIR photometry an important factor in the methods followed throughout this thesis. The best available FIR photometry for large and deep extragalactic surveys comes from the instruments aboard the Herschel Space Observatory (*Herschel*; Pilbratt et al. 2010). The Photoconductor Array Camera and Spectrometer (PACS; Poglitsch et al. 2010) with bands centred at 70μm, 100μm, and 160μm, and the Spectral and Photometric Imaging Receiver (SPIRE; Griffin et al. 2010) with bands centred at 250μm, 350μm, and 500μm, have been used to map multiple fields including various legacy survey fields, such as COSMOS and the Chandra Deep Fields (CDF).

Due to the low resolution in the SPIRE bands, with PSFs of $\sim$20–30arcsec, deep surveys suffer from high levels of blending of faint sources ("source confusion"), with single pixels potentially having significant contributions from multiple sources. This results in two types of noise in the *Herschel* maps, the typical instrumental noise, and the additional confusion noise which is highly correlated between pixels. For these reasons, detection limits are comparatively high, and at $z \gtrsim 0.5$ only luminous galaxies will be detected (see section 2.6). To overcome the low fraction of detections, studies of AGN samples that include *Herschel* non-detections, tend to use stacking at one of the FIR bands to constrain the mean SFRs. However, stacking and single band derivations of the mean SFR can be influenced by several causes of uncertainty (see section 2.6 for a discussion), such as the wide range in IR SEDs, the possible AGN contamination, and the contamination due to the confusion of sources in the *Herschel* maps. With the methods followed in this thesis I try to overcome these sources of uncertainty. By using catalogues from deblended maps, and SED fitting to decompose the IR emission, I remove, as best possible, the impact of confusion, and AGN contamination, respectively (see section 2.3; Chapter 3 & 5). By performing individual SED fitting I have a set of measurements and upper limits calculated in a consistent manner, and I combine them using the Kaplan-Maier estimator to calculate the mean SFRs even in the presence of a large fraction of upper limits (see section 2.4; Chapter 3). When stacking is used in this thesis, the effects of confusion noise are carefully accounted for. Furthermore, by stacking in multiple bands any AGN contribution to the IR emission can be removed before calculating SFRs. (section 2.5; Chapter 4).

In this thesis I also take advantage of recent observations with the Atacama Large Millimeter/submillimeter Array (ALMA). The ALMA observatory can achieve sensitivities down to 0.1mJy/beam, achieving more than a magnitude lower sensitivities than the *Herschel* bands. With the sensitivities achieved by ALMA it is now possible to easily detect galaxies and AGN at redshifts above $z \sim 1$ at lower fluxes than that possible with previous FIR/Sub-mm observatories. I combine ALMA 870μm photometry with that of *Herschel* to achieve better constraints on the IR emission of typical to faint AGN host galaxies at $z > 1$ in Chapter 5.

To successfully remove the AGN contribution to the IR emission I require good quality photometry covering the peak of the IR emission of the AGN at rest-frame wavelengths of $\sim$10–40μm (see Fig. 2.2). For the analysis followed in this thesis I take advantage of archival photometric catalogues from the Spitzer Space Telescope (*Spitzer*;Werner et al. 2004) deep field observations (in Chapter 3 & 5), and the All-Sky source catalogue of the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010; in Chapter 4). In this thesis I specifically use observations done with the *Spitzer* IRAC and MIPS instruments, at the four bands of IRAC at 3.6, 4.5, 5.8, and 8μm, and the MIPS 24μm band. Observations with IRAC and MIPS–24μm are ideal for deep surveys as they reach sensitivities of a few 10s of μJy, and so can detect faint sources over the redshifts of interest for this thesis. WISE consists of four bands at 3.4, 4.6, 12, and 22μm that reach down to sensitivities of 0.08–6mJy respectively. The all-sky survey provides great source statistics on more rare objects, like luminous QSOs for which I take advantage in Chapter 4.

2.3 IR SED fitting with the removal of the AGN component

The spectral energy distribution (SED; i.e. the energy emitted as a function of wavelength, frequency, or energy) of a galaxy serves as a diagnostic of its emission processes and intrinsic properties. The shape of the SED is governed by the physical processes occurring in the galaxy, and different wavelength ranges probe different processes that contribute to the overall emission. For example, at the UV-optical wavelength range the dominant contributors to the galaxy's emission are the stars (depending on the stellar population) and the AGN (depending on how powerful it is; i.e. quasars), while at MIR and FIR wavelengths the main contributors are the reprocessed emission from the AGN and star-forming regions (depending on the abundance and spatial distribution of the dust in the galaxy).

When studying the star-forming properties of AGN at IR wavelengths it is important to take into account the full IR SED and also the possible contribution from an AGN component. As has been suggested by previous studies (e.g., Netzer et al. 2007; Mullaney et al. 2011; Del Moro et al. 2013; Delvecchio et al. 2014) it is important to apply SED decomposition methods at IR wavelengths to constrain the strength of the AGN and SF components. The above is also demonstrated in Chapter 4, where it is found that a strong AGN contribution that is not taken into account when interpreting IR data can significantly alter the star formation measurements.

Due to the lack of uniformly distributed photometry across the IR SED of distant galaxies and AGN it is impossible to accurately characterise the SEDs of each source. Instead, we rely on empirically or model derived SED shapes which we fit to the available IR photometry. Within this thesis, I make use of a set of empirically derived SEDs to model the IR emission due to star formation (from now on referred to as SF templates), and the IR emission due to the AGN (from now on referred to as AGN templates).

2.3.1 The empirically defined SEDs and how they were constructed:

The SF and AGN templates used in my SED fitting analysis were originally constructed in Mullaney et al. (2011) and later extended by Del Moro et al. (2013). In the following paragraphs I give a general overview of how these templates were constructed.

The primary SF templates were defined using nearby star-forming galaxies with available *Spitzer*-IRS spectra (covering the wavelengths ∼6–35μm) and IRAS photometry (at wavelengths 12, 25, 60, and 100μm). The sample was selected to uniformly sample the IRAS colour-colour space of IR bright galaxies. Using the *IRAS Revised Bright Galaxy Survey* (RBGS; Sanders et al. 2003) to define the IRAS colour-colour parameter space of IR bright galaxies, Mullaney et al. (2011) chose 10 star-forming galaxies from the sample of Brandl et al. (2006) covering a wide range of the RBGS sample, and added 4 complementary sources with colder IRAS colours (see Fig. 3 of Mullaney et al. 2011). The sample selection is representative of SED shapes of star-forming galaxies with $L_{IR} = 10^{10} - 10^{12} L_{\odot}$. These galaxies were then grouped in terms of their overall MIR SED shape and the relative strength of their PAH features to produce the SF templates. In total 5 groups of galaxies were identified and composite SF templates were produced for each group of galaxies, covering the wavelength range of 6–1000μm. To extend the templates beyond the IRAS–100μm photometry Mullaney et al. (2011) extrap-

SF template	Galaxy	$\log(L_{IR}/L_{\odot})$
SF1	NGC 1667	10.96
	NGC 5734	11.06
	NGC 6286	11.32
	NGC 7590	10.16
SF2	NGC 7252	10.77
SF3	Mrk 52	10.25
	NGC 4818	9.75
	NGC 7714	10.72
SF4	NGC 1222	10.60
	NGC 3256	11.56
	NGC 4194	11.06
SF5	NGC 520	10.91
	NGC 660	10.49
	NGC 2623	11.54
SF6	Arp220	12.30

Table 2.1: The galaxy SEDs used to create the five SF templates from Mullaney et al. (2011), and the additional Arp220 template added in this thesis. The final SEDs are shown in Fig. 2.1 with red curves.

olated the average SEDs assuming a modified blackbody function (i.e. $F = F_{\nu}^{BB}\nu^{\beta}$, where F_{ν}^{BB} is the blackbody-specific flux, ν is photon frequency, and $\beta = 1.5$).

In this thesis I use the 5 templates described above, as extended by Del Moro et al. (2013) to cover the wide wavelength range of 3 – 10^5 μm, in addition to the SED of Arp220 from Silva et al. (1998) as an additional SF SED covering cases of extremely dusty and luminous sources. In Table 2.1 I list the different SF templates and the galaxies used to construct them, and in Figure 2.1 I plot the SF templates in comparison to the set of the model defined SED templates from Dale & Helou (2002), and the two empirically defined high-redshift templates of Kirkpatrick et al. (2012). The SF templates used in this thesis successfully cover a wide range of SED shapes, including the majority of shapes of model templates, as well as covering the shapes of higher redshift galaxies.

To construct the AGN templates Mullaney et al. (2011) used a sample of 25 Swift-BAT AGN with IRS spectra (covering the wavelengths $\sim$6–35μm) that showed characteristics of the power-law continuum of the AGN overpowering the PAH lines due to star-forming emission. From the IRS spectra of these AGN dominated galaxies, it was evident that there is an underlying power-law continuum of spectral index $0.7 \leq \alpha_1 \leq 2.7$ with a break

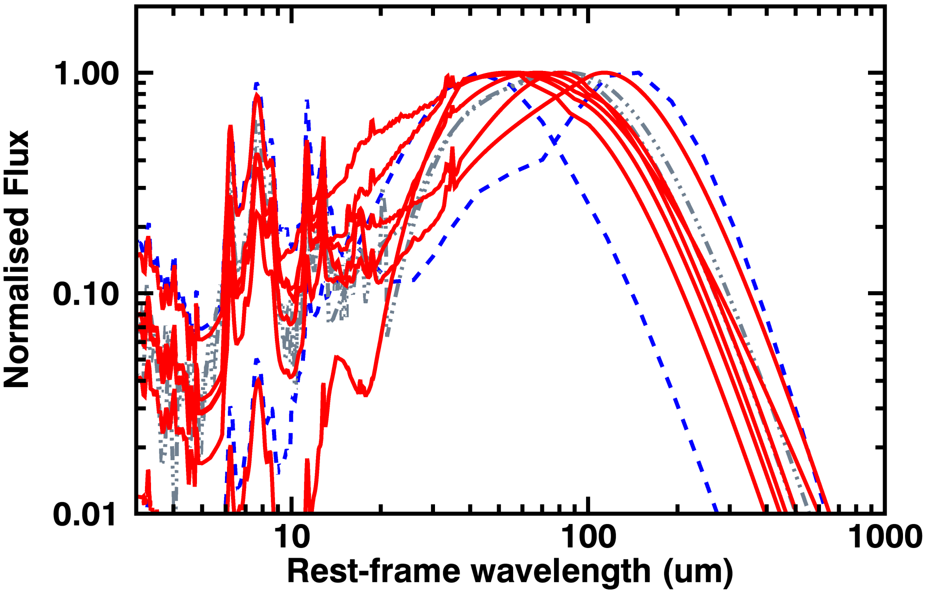


Figure 2.1: The six empirical SF templates used in our SED fitting process (red curves) in comparison to the model defined templates of Dale & Helou (2002) (blue dashed curves giving the two extremes of the set of templates), and empirically derived templates for galaxies at redshift 1 and 2 from Kirkpatrick et al. (2012) (dark grey dot-dashed curves). The SEDs used in my analysis cover a wide range in shapes, covering both the high-redshift templates of Kirkpatrick et al. (2012), as well as a wide range of the model defined templates of Dale & Helou (2002).

at $15 \leq \lambda_{brk} \leq 20\mu$m, after which the SED has a lower spectral index of $0 \leq \alpha_2 \leq 1.5$. To extend the intrinsic AGN SED to longer wavelengths, Mullaney et al. (2011) extrapolated to the FIR wavelengths using the IRAS 60μm and 100μm bands and assuming that the intrinsic AGN SED falls as a modified black body beyond a given wavelength λ_{BB}, which is left free to range between 20, and 100μm. However, at FIR wavelengths there can be strong contribution from star forming emission, even if a source is AGN dominated in the MIR. To remove any star-forming contribution to the final AGN templates, the photometry of the sample of AGN dominated galaxies was fitted by a combination of the previously defined SF templates and the intrinsic AGN templates with λ_{BB} free within the range specified above. From the resulting best-fits a sample of AGN templates was defined. In this thesis I make use of the parametrisation of the mean intrinsic AGN SED of the whole sample, which gives $\alpha_1 = 1.8$, $\alpha_2 = 0.2$, $\lambda_{brk} = 19\mu$m, and $\lambda_{BB} = 40\mu$m.

In Fig. 2.2 I show the SEDs constructed by Mullaney et al. (2011) in comparison to SEDs derived for QSO samples by Richards et al. (2006), Mor & Netzer (2012), and Symeonidis et al. (2016). Both Mor & Netzer (2012) and Symeonidis et al. (2016) took a similar approach to that of Mullaney et al. (2011), where an estimate of the emission due to star formation is removed from the total SED to retrieve the shape of the AGN template. The main difference between the definitions of the Mor & Netzer (2012) templates to the ones defined by Mullaney et al. (2011) is the fact that in Mor & Netzer (2012) they derive the template of only the torus emission removing additional AGN emission from the narrow and broad-line regions. The earlier Richards et al. (2006) templates of QSOs was defined using photometric data covering the radio to X-ray regimes for a sample of optically selected, broad-line QSOs, and shows a similar drop-off at long wavelengths as the templates of Mullaney et al. (2011). The only template set with significant difference at long wavelengths to the set of Mullaney et al. (2011) are those of Symeonidis et al. (2016), where the template has a more gradual drop-off at long wavelengths. However, in Chapter 5 we demonstrate that the shape of this template is not appropriate to describe the majority of AGN studied here, by using the observed MIR–Sub-mm colours.

In this thesis the objective has been to remove all of the AGN contribution from the IR emission of the galaxy, for this reason between the Mullaney et al. (2011) and the Mor & Netzer (2012) templates the former is preferred.

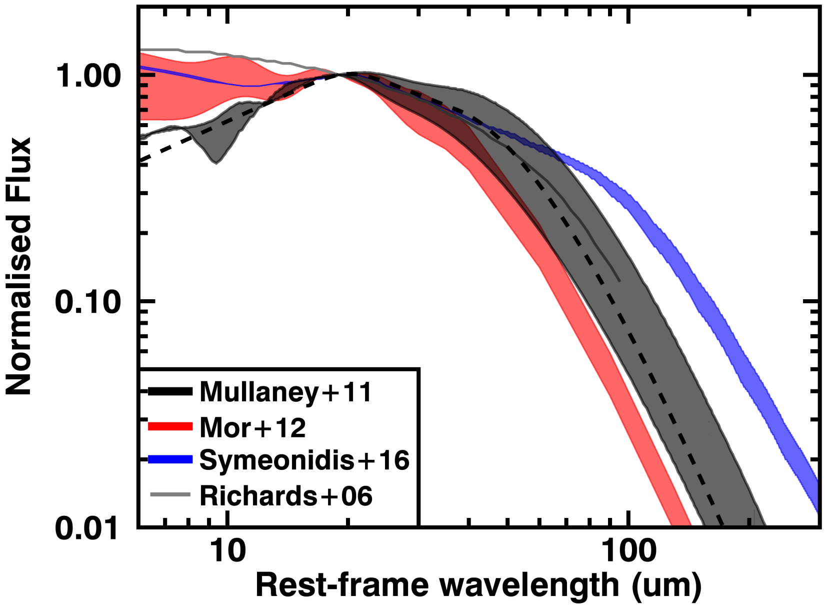


Figure 2.2: The AGN templates constructed by Mullaney et al. (2011) for X-ray identified AGN of moderate luminosities (black region), with the black dashed line showing the mean AGN template used in the SED fitting analysis throughout this thesis. Also shown for comparison are other templates derived by Mor & Netzer (2012) (red region), Richards et al. (2006) (grey curve), and Symeonidis et al. (2016) (blue region). All templates are normalised at 19μm. The empirical template used in this work lies in between the Richards et al. (2006) and Mor & Netzer (2012) templates at the long wavelengths, with a similar drop off slope. However at the shorter wavelengths, the template shows a steep rise, while all of the QSO derived ones show a more flat slope. In later chapters of this thesis, I test different AGN templates, and how they can affect the results presented.

2.3.2 The SED fitting process

With the set of AGN and SF templates described above I can simultaneously fit the AGN and SF IR SED components to the available MIR and FIR photometry to constrain the IR luminosity due to star formation ($L_{\mathrm{IR,SF}}$) in the galaxy without contamination from the AGN. Although the SEDs were created using nearby galaxies, the samples used were such that they covered the same IR luminosities observed for higher redshift typical star-forming galaxies. Furthermore, when comparing to the Kirkpatrick et al. (2012) the SED shapes are similar to those redshift 1 and 2 galaxy templates. Consequently, the SF templates used in this thesis are appropriate for analysing the IR SEDs of higher redshift AGN host galaxies.

To fit and decompose the IR SEDs of our respective samples I use a combination of MIR and FIR/Sub-mm photometry, covering the wavelength range of 8 – 500μm in science Chapters 3 & 4, and 8 – 870μm in science Chapter 5.

I fit to the photometry chosen in each case using the publicly available DECOMPIR code of Mullaney et al. (2011). The code uses χ^2 minimisation to converge on the best fit solution. As input it requires the wavelength, flux density, and error arrays, as well as the redshift of the source. The algorithm shifts the observed frame SED to rest-frame wavelengths, and uses the response curves of the photometric bands corresponding to the given photometry, to integrate over and accurately fit to the photometry. The code also allows the input of any SF template, specified on input, but automatically creates the AGN template based on specified parameters describing the shape of the template (see section 2.2.1, and section 4.2 of Mullaney et al. 2011). We fix the parameters of the AGN template to those values describing the mean empirical AGN template discussed in section 2.2.1. We use the default option of DECOMPIR of only allowing for the normalisation of the given templates to vary during the fit, this allows for a maximum of two free parameters, the normalisation of the SF and AGN template respectively. Additionally, I have further developed DECOMPIR to include options for fixed normalisation of the SF or AGN templates, the option of a maximum normalisation for the SED, and the option to choose no AGN or no SF component for the fit.

As described above, the fitting process that is used has two free parameters, the normalisation of the SF template and the normalisation of the AGN template, and so we need at least three photometric measurements (detections) to be able to confidently distinguish between an IR SED that requires an AGN component and one that does not. Additionally, as the AGN component can dominate the SED at rest-frame MIR wavelengths, with a turnover from the AGN to the SF component (where the two components cross-over) at $\lambda \sim 70$–80μm, we require at least one FIR data point at $\lambda > 80\mu$m. Consequently, the types of fits I perform are dependent on the available photometry of each source, and are outlined below:

1. For a source with detections in more than two bands, with at least one at rest-frame wavelengths longer than 80 μm, I am able to fit both the AGN and SF templates simultaneously. Using the measured photometry to fit to, and the photometric upper limits as upper bounds on the SED, I perform fits for two possible cases:

    (a) Fit all six SF templates individually (SF-only)

    (b) For the same SF templates, fit with the addition of the AGN component (AGN+SF)

    The above results in a set of 12 individual fits, which will later be evaluated to chose the best-fitting solution.

2. For a source with not enough photometry to confidently constrain the SED, I perform the following:

    (a) In the case of a source with no photometric detections, I use the upper limits to constrain an upper limit on the SF templates, and the AGN template individually.

    (b) For the cases with both photometric detections and upper limits, I use both the limits and the detections for the calculation of the upper limit on the SF and AGN emission.

    (c) In the case where the upper limit of all the SF templates (calculated as described in 2b) is more than 5σ below the MIR photometry, I allow an AGN component to be fit to the MIR photometry, and define a range for the AGN luminosity. The range is between the AGN contributing all of the MIR emission and the AGN emission if the SF template is right at the maximum possible value taking into account the photometric upper limits.

From the resulting fitted SEDs I determine a best-fitting solution using the Bayesian Information Criteria (BIC; Schwarz 1978) which allows the objective comparison of different non-nested models with a fixed data set, and is defined as:

$$\mathrm{BIC} = -2 \times \ln \mathrm{L} + \mathrm{k} \times \ln \mathrm{N} \tag{2.3.1}$$

where L is the maximum likelihood, k is the number of free parameters, and N the number of data points. This method penalises against models with extra free parameters counter-balancing the fact that a model with more free parameters will in general fit the data better, irrespective of the relevance of the parameters. This is an improvement over a simple $\Delta\chi^2$ test or a maximum likelihood comparison that would tend to favour the model with more free parameters. For each source the BIC value is calculated for all of the different fits. The best fitting model will be the one which minimises the BIC value, its absolute value being irrelevant; however for one model to be significantly better than the others it needs to have a difference in BIC value of $\Delta\mathrm{BIC} \geq 2$. If $\Delta\mathrm{BIC} \leq 2$ then both models are considered equally valid (e.g. see section 2 of Liddle 2004; Mukherjee et al. 1998). Our final best fit solution is the one with the lowest BIC value; however, I only accept the AGN+SF fit as best fit over the fit with only the SF component, if the inclusion of the AGN reduces the BIC value by ≥ 2. From the best fit SEDs I then measure the integrated 8 – 1000μm IR luminosity of the star formation component ($\mathrm{L}_{\mathrm{IR,SF}}$). In cases of multiple fits having BIC values equal to the minimum BIC value, I consider them equally valid and take the mean of their derived $\mathrm{L}_{\mathrm{IR,SF}}$. As errors on the chosen $\mathrm{L}_{\mathrm{IR,SF}}$ value from the best fit, I take the range of $\mathrm{L}_{\mathrm{IR,SF}}$ values from the fits with BIC values with $\Delta\mathrm{BIC} \leq 2$ to the best fit.

In the cases where I can only define upper limits on the fits, I derive the upper limit on all the SF templates and conservatively take the maximum value of the resulting $\mathrm{L}_{\mathrm{IR,SF}}$ as the upper limit value for the source. The advantage of our method is the fact that I calculate individual $\mathrm{L}_{\mathrm{IR,SF}}$ measurements and upper limits through the same SED-fitting process and photometry.

2.4 Average SFRs: A Survival Analysis Technique

One major limitation in FIR studies is the fact that even with deep surveys, there will always be a fraction of the sample that will not be detected in a given waveband and only an upper limit on its photometry can be reliably calculated. This tends to affect high redshift surveys the most, with large fractions of the galaxy population only having upper limit constraints on their FIR photometry.

Astronomy is not the only research subject suffering from the restriction imposed by limits and non-detections. A multitude of actuarial sciences, such as clinical research, also suffer from such limitations. For example, a study on patients receiving cancer treatment lasting 5 years will have patients that are successfully treated within the timescale of the study, patients that died or decided to leave the studied group, and patients that even though they remained in the group for the full extent of the study they were not successfully treated by the end of it. Consequently, estimating the success rate of the treatment is challenging. To address such issues there is a whole branch of statistics, known as "survival analysis" or "lifetime data", where the inclusion of limits for the estimation of the underlying distribution function is explored. With such methods it becomes possible to use the combination of limits and measured values, and under certain assumptions, estimate the properties of the underlying distribution.

In this thesis I make use of a non-parametric approach to estimating the mean properties of our sample of $L_{IR,SF}$ measurements and upper limits. The Kaplan-Meier product limit estimator was first derived by Kaplan & Meier (1958) in a study of cancer treatments, and in Feigelson & Nelson (1985) and Schmitt (1985) it was introduced as a tool that could also be used in astronomical research. The Kaplan-Meier (KM) estimator is a non-parametric estimator, i.e. does not assume a specific type of the underlying distribution. To be able to use the KM estimator correctly the samples I apply it to need to follow certain principals. The main principal for using the KM estimator is that the sample follows a pattern of "random censorship". Censorship refers to the presence of limits. In statistical studies left and right censoring correspond to upper and lower limits respectively, and I will follow those terms in the rest of this section. Random censorship refers to a state where the censoring variables (i.e., limits) "are mutually independent, identically distributed, and independent of the true measurements " (Feigelson & Nelson

1985).

Due to the nature of actuarial sciences, the samples suffer more from right censorship rather than the left censorship that affects astronomical samples. However, it is simple to transform a left censored sample to a sample with right censorship. For the sake of simplicity I use this transformation for the derivation of the KM estimator, following Feigelson & Nelson (1985).

For a sample of n sources with $\{\chi_i^L\}_{i=1}^n$ values, where L denotes the left censorship of the sample, true measurements are defined as T_i^L, and upper limits as A_i^L. We also define the flag $\delta_i^L = 1$ or 0, when the value is a measurement or upper limit respectively. We assume that all true measurements, T_i^L, have been randomly selected from the same population with a distribution function of

$$F^L(t) = P(T^L \leq t) \tag{2.4.2}$$

where P is the probability function of the event $T^L \leq t$.

We can transform the left censored sample to an equivalent right censored sample, by taking the constant $M \geq$ max($\{\chi_i^L\}_{i=1}^n$) and subtract from it the values of our original sample. The new sample is then χ_i = M - χ_i^L, with $T_i = M - T_i^L, A_i = M - A_i^L$, and $\delta_i = \delta_i^L$. The distribution function of the transformed true measurements can now be described as $F(t) = P(T \leq t)$, with t being the original values of the distribution from which T where selected. By doing this transformation I am able to use the Kaplan-Meier estimator to calculate the underlying distribution function for a given sample with left censorship, even though it was originally defined for a right censored sample.

Kaplan & Meier (1958) derived a non-parametric maximum-likelihood-type estimator of $F(t)$, which is usually stated as the survivor function $S(t)$, where

$$\begin{aligned} S(t) &\equiv P(T \geq t) \qquad (2.4.3)\\ &= P(T > t) + P(T = t) \qquad (2.4.4)\\ &= 1 - F(t) + P(T = t) \qquad (2.4.5)\\ & \qquad (2.4.6) \end{aligned}$$

For a sample of distinct and sorted values $\{\chi_i\}_{i=1}^{n}$, and $\chi_0 \equiv 0$, the probability function will be described by

$$P_i \equiv P[T \geq \chi_{(i+1)} | T \geq \chi_{(i)}] \tag{2.4.7}$$

i.e., the conditional probability that T is at least $\chi_{(i+1)}$ given that it is at least $\chi_{(i)}$. And so, the survivor function can now be expressed as:

$$\begin{aligned} S(\chi_{(j)}) &= P[T \geq \chi_{(j)}] && (2.4.8) \\ &= \prod_{i=0}^{j-1} P[T \geq \chi_{(i+1)} | T \geq \chi_{(i)}] && (2.4.9) \\ &= \prod_{i=0}^{j-1} P_i && (2.4.10) \\ & && (2.4.11) \end{aligned}$$

The estimator of P can be defined as:

$$\hat{P}_i = \begin{cases} 1 - \frac{1}{n-i+1} & \text{, when } \chi_{(i)} \text{ is a true value} \\ 1 & \text{, when } \chi_{(i)} \text{ is a limit} \end{cases}$$

$$\Rightarrow \hat{P}_i = [1 - \frac{1}{n-i+1}]^{\delta(i)}$$

which is based on the two facts that (a) if χ_i is a true value, then there are $n-i+1$ true values $\{\chi_j\}$ equal to or larger than $\chi_{(i)}$, with only one of them ($\chi_{(i)}$) not being as large as $\chi_{(i+1)}$; and (b) if $\chi_{(i)}$ is a censored value, all the true values $\{\chi_{(j)}\}$ are equal to or larger than $\chi_{(i+1)}$.

The estimator of S is then:

$$\hat{S}_i = \prod_{i=1}^{j-1} [1 - \frac{1}{n-i+1}]^{\delta(i)} \tag{2.4.12}$$

And because this is a nonincreasing step function, which only jumps at true values, it can be expressed as:

$$\hat{S}(t) = \begin{cases} \prod_{j,\chi_{(i)}<t} [1 - \frac{1}{n-i+1}]^{\delta_{(i)}} & \text{, when t} > \chi'_{(1)} \\ 1 & \text{, when t} \leq \chi_{(1)} \end{cases} \tag{2.4.13}$$

For a more realistic sample, where not all values are distinct, we define a sub-sample of all the distinct values, $\{\chi'_i\}_{i=1}^{L}$. Then we can break the ties between a true value and an upper limit that have the same value, by assuming that the upper limit value is larger. We then define the parameters:

$$n_j = \#\{k, \chi_k \geq \chi'_{(j)}\} \tag{2.4.14}$$

$$d_j = \#\{k, \chi_k = \chi'_{(j)}\} \tag{2.4.15}$$

$$\tag{2.4.16}$$

and now the survivor function becomes:

$$\hat{S}(t) = \begin{cases} \prod_{j,\chi'_{(j)}<t} [1 - \frac{d_j}{n_j}]^{\delta'_{(j)}} & \text{, when t} > \chi'_{(1)} \\ 1 & \text{, when t} \leq \chi_{(1)} \end{cases} \tag{2.4.17}$$

Once the survivor function is defined it can then be transformed back to the equivalent for our original sample that includes upper limits. The estimator is: $F^L(t) = \hat{S}(M - t)$.

The mean of the distribution can then be defined as:

$$\hat{\mu} = \int_0^{\inf} \hat{S}(x)dx = \sum_{j=1}^{r+1} \hat{S}(x'_j)[x'_(j) - x'_{(j-1)}] \tag{2.4.18}$$

$$\hat{\mu}^L = M - \hat{\mu} \tag{2.4.19}$$

assuming that $x'_{(0)} = 0$ and $x'_{(r+1)} = \inf$. In the case where the minimum value is an upper limit I assume that it is a true value, to overcome the fact that μ would then tend to infinity

To demonstrate how important it is to include sources with only upper limit constraints, so as to not bias the sample, we calculate the average SFR for only the SFR detections and compare to the average of the whole sample. I select a sample of X-ray AGN from Chapter 3, for which I have performed the SED fitting methods described in section 2.2.2, covering the redshift range of $z = 1-3$[1], I estimate the mean SFRs in bins of

---

[1] Since this plot aims only to demonstrate a comparison of two methods, I do not worry about the wide range of redshift.

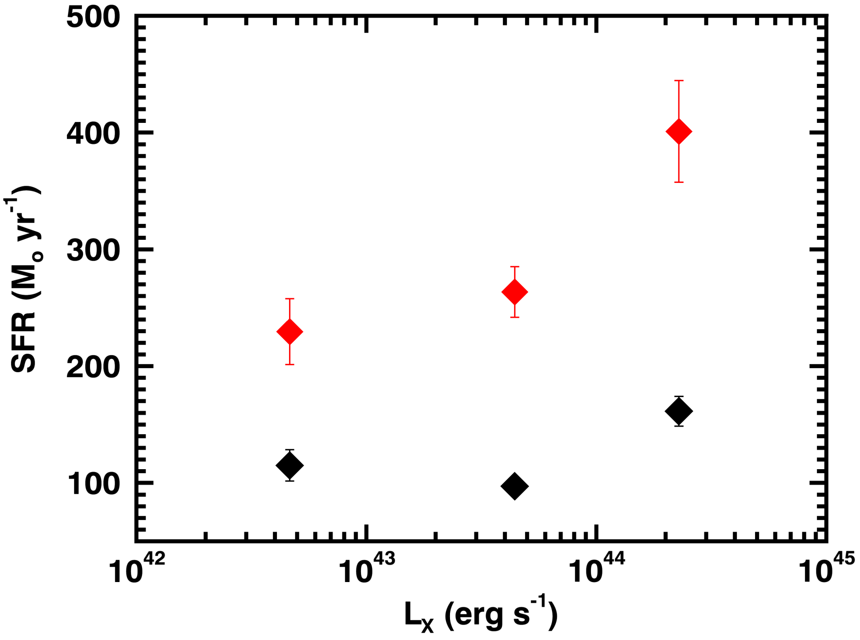


Figure 2.3: Plot of the average SFR for three L_X bins for AGN at $z = 1 - 3$. Here I compare the average SFR results when including (black) and not including (red) the SFR upper limits. The means have been calculated using the Kaplan-Meier estimator, after performing SED fitting to their IR photometry, and converting from $L_{IR,SF}$ to SFR using the Kennicutt (1998) relation for a Chabrier IMF. The errors on the means are calculated using the bootstrap technique. The difference between the two measurements demonstrates how high the bias can be if the non-detections are ignored.

X-ray hard-band (2–8keV) luminosity, with and without the inclusion of the upper limits using the Kaplan-Meier estimator. In Fig. 2.3 I compare the results. In red are the results of just the sources with a SFR measurement, and in black the results when including the whole sample. Just looking at the part of the sample with SFR measurements will give a significantly different result to including the whole sample. Being able to account for the whole population rather than a biased sample of only detections, is important to be able to place reliable constraints on the mean properties of the entire population. The Kaplan-Meier estimator, allows for this, and provides an alternative to the more commonly used stacking techniques, while also providing the key advantage of decomposing each individual IR SED into the AGN and SF components.

2.5 Average SFRs: Multi-wavelength Stacking and composite SEDs

Another method used in this thesis is the stacking of FIR photometry from *Herschel* PACS and SPIRE observations. When studying large samples of sources, the majority of which are undetected in a given FIR band, stacking analyses are typically used to calculate the average photometric properties. With stacking it becomes possible to define the mean photometry of a sample, even if a large fraction of individual sources do not have sufficient flux to be considered a detection. In this section we discuss the uncertainties of stacking FIR photometry, such as the blending of sources due to large beam sizes, and only using a single band for deriving the SFRs. In Chapter 4, I stack in multiple FIR bands and in combination to MIR photometry I perform composite SED fitting to remove the contribution from IR AGN emission to SFR measurements.

For each band I take cut-outs of the maps centred at the source positions. These cut-outs have been re-gridded to a 1” pixel scale so as to maximise the accuracy of the centring. We then create an image cube that holds all the cut-outs of the sample centred at each source position. We collapse this cube to a single mean image for the whole sample (see Fig. 2.4), and use it to calculate the mean flux density of the stacked “source”.

When extracting the mean flux density of the stacked “source ” from the stacked images of PACS bands at 100μm, and 160μm, I follow the suggested procedure by Valiante et al. (in prep) for a point source. I integrate the flux density (in units of Jy/pixel) within an aperture of 3” radius and use the recommended corrections of 2.63, and 3.57, for 100μm and 160μm respectively, to retrieve the full flux density of the source.

When performing the stacking procedure with the SPIRE photometry at 250μm, 350μm, and 500μm, the process is slightly simpler as the photometry in these maps is given in units of Jy/beam. This means that the flux density at each pixel has been integrated over the beam-size of the instrument, and so I only need to read off the flux from the central pixel corresponding to the central “source” position.

To determine the error on the stacked mean I used the bootstrap technique. We randomly pick sources from our image cube and take their mean. The process is repeated 1000-10000 times and from each stacked image the stacked “source” flux density is ex-

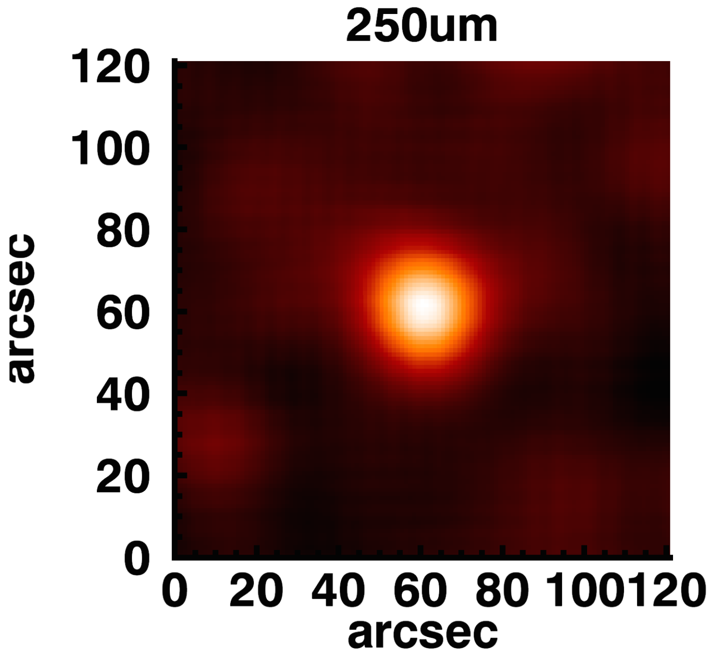


Figure 2.4: Example of a mean stacked image for SPIRE–250μm photometry. Examples for all bands are shown in Chapter 4.

tracted following the methods outlined above. From the distribution of flux density values I determine the 16th and 68th percentiles, which are used to calculate the 1σ errors on the stacked mean (see Fig. 2.5).

Due to the underlying source confusion noise present in the SPIRE maps it is possible that a stack of a few sources could produce a mean flux that is just the consequence of stacking noise. To make sure our results do not suffer from the inclusion of random values I perform a series of stacks at random positions on the map to calculate the background noise. We perform 10000 random stacks including the same number of random positions as the number of sources included in each source stack. From the series of random stacks a distribution of stacked background noise flux densities is produced (see Fig. 2.6). This distribution usually has the form of a Gaussian with a positive heavy tail due to the confusion background and can be offset from zero due to the stacking of the confusion background. I use the 99th percentile of the distribution (red dashed line in Fig. 2.6) to evaluate if a stacked flux density is significant, and not just stacked noise. To remove the contamination from background source confusion from our stacked source fluxes, we remove the offset of the random-stack distribution from the stacked source flux.

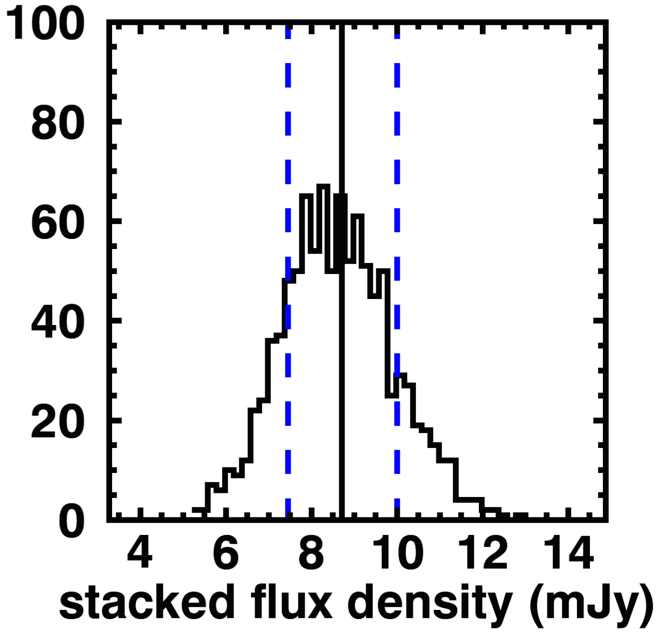


Figure 2.5: Example of the distribution of values after performing the bootstrap technique for the SPIRE–250μm band. the blue dashed lines correspond to the 16th and 84th percentiles, which are taken as the 1σ error on the stacked flux density. Examples for all bands are shown in Chapter 4.

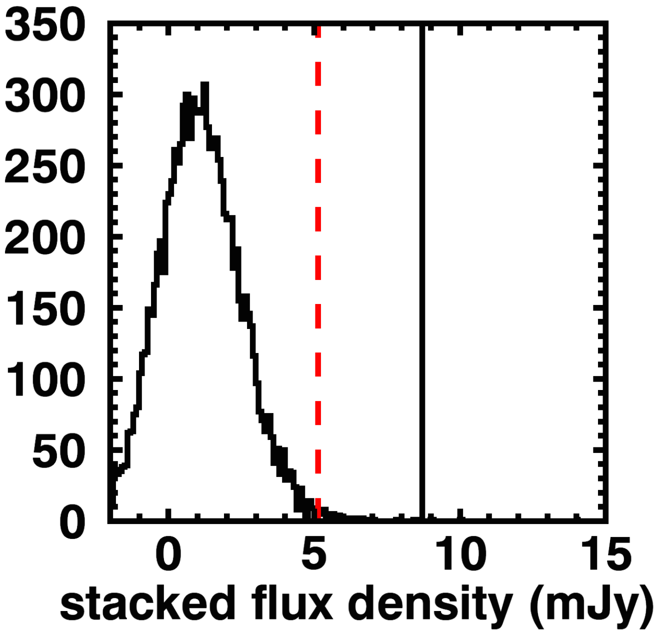


Figure 2.6: An example of the resulting distribution of background noise flux densities from random stacking for the SPIRE–250μm band. The red dashed line corresponds to the 99th percentile of the distribution, and the solid black line is the flux density of a stacked source. Examples for all bands are shown in Chapter 4.

This method is mainly used in Chapter 4, where I also show example distributions of both random-stacks and the bootstrap method.

With the combination of the mean fluxes in both MIR and FIR bands, I perform SED fitting to decompose the IR emission (as described in section 2.3) and remove the AGN contribution from the composite SEDs in Chapter 4.

2.6 A comparison of SFR measurement methods: single band derivation vs multi-wavelength SED fitting

Most FIR studies of the star-forming properties of AGN use single band derivations of the L_{IR}, to calculate the SFR from the galaxy. The preferred bands are usually those of PACS–100μm and PACS–160μm (e.g., Mullaney et al. 2012a, Rosario et al. 2012), and SPIRE–250μm (e.g., Page et al. 2012, Harrison et al. 2012, Kalfountzou et al. 2014).

However, the calculation of L_{IR} from a single band will cause uncertainty in the results, as the shapes of the IR SEDs of galaxies can vary substantially, and from single-band photometry it is not possible to reliably remove the AGN contamination from the L_{IR}.

To demonstrate the uncertainty in deriving the L_{IR} from a single band, I have used the 5 SF templates from Mullaney et al. (2011) (see section 2.3), normalized to a flux density of 1 at the different MIR-FIR bands commonly used, i.e, MIPS–24μm, PACS–100μm, PACS–160μm, and SPIRE–250μm. In Figure 2.7 I show the range of L_{IR} measurements from a single band derivation as a function of redshift and waveband. The spread of results from a single band derivation of the SFR is dependent on the redshift of the sources, because the observed frame wavelength will probe different parts of the SEDs, and the SF templates tend to vary less close to the peak of the SED when compared to shorter and longer wavelengths. Each band shows a minimum range at redshifts where the observed band covers the region where the SF templates are peaking, i.e. rest frame wavelengths of $\sim$60 – 120μm. Once off the peak a single band can give a wide range of results of up to a factor of $\sim$3–4, in the FIR bands. However, it is evident when comparing to the range covered by the 24μm band, that the FIR is still a better choice for the estimation of the SFR than the MIR wavelengths. In Figure 2.8, I show the typical depth of deep-field *Herschel* surveys in SFR measurements as a function of redshift. The SFRs are calculated using the detection limit for sources in the COSMOS field (from the deblended catalogues used in Chapter 3), and normalising the templates to the limiting flux density for all available MIR–FIR wavebands (i.e., MIPS–24μm, PACS–100μm, PACS–160μm, SPIRE–250μm, SPIRE–350μm, SPIRE–500μm) at the different redshifts. Past a redshift of $z \sim$0.5 the MIPS and *Herschel* bands start to only detect the most luminous star-forming galaxies.

To demonstrate the extent that AGN contamination can boost the estimates of SFR measurements I estimate how much of the estimated SFR could actually be from the AGN, by using the X-ray hard-band luminosities and our AGN template. We use the Stern (2015) relation to convert the X-ray, 2–8keV, luminosity to the 6um luminosity from the AGN. We then normalise the AGN templates used in our SED fitting analysis to the corresponding rest-frame 6μm flux and calculate the flux due to the AGN at the rest-frame wavelengths of 60, 100, 120μm. We then fit the SF templates to the fluxes estimated for each wavelength, which results in an erroneous SFR due to misinterpreting

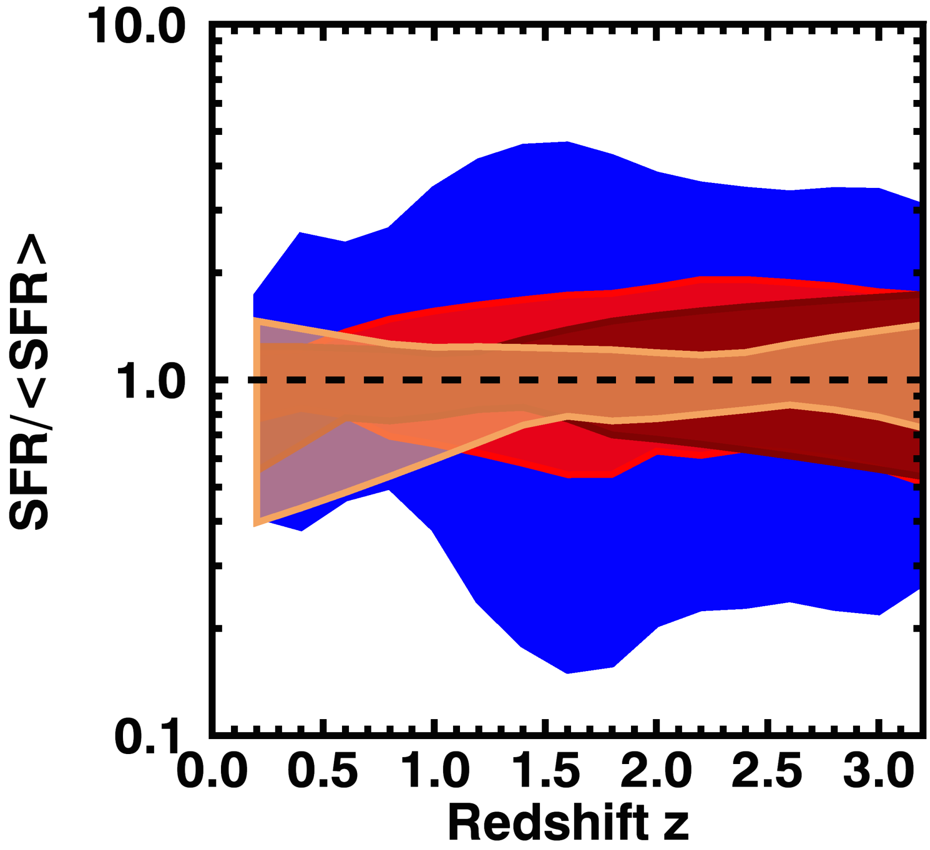


Figure 2.7: The range of SFR measurements from using single-band derivation from the 5 SF templates (see section 2.2), as a function of redshift. MIPS-24μm in blue, PACS-100μm in red, PACS-160μm in brown, and SPIRE-250μm in orange. The bands tracing the peak of the FIR emission give the smallest range of results, with PACS-160μm performing best at z $<$ 1.5 and SPIRE-250μm at z $>$ 1.5.

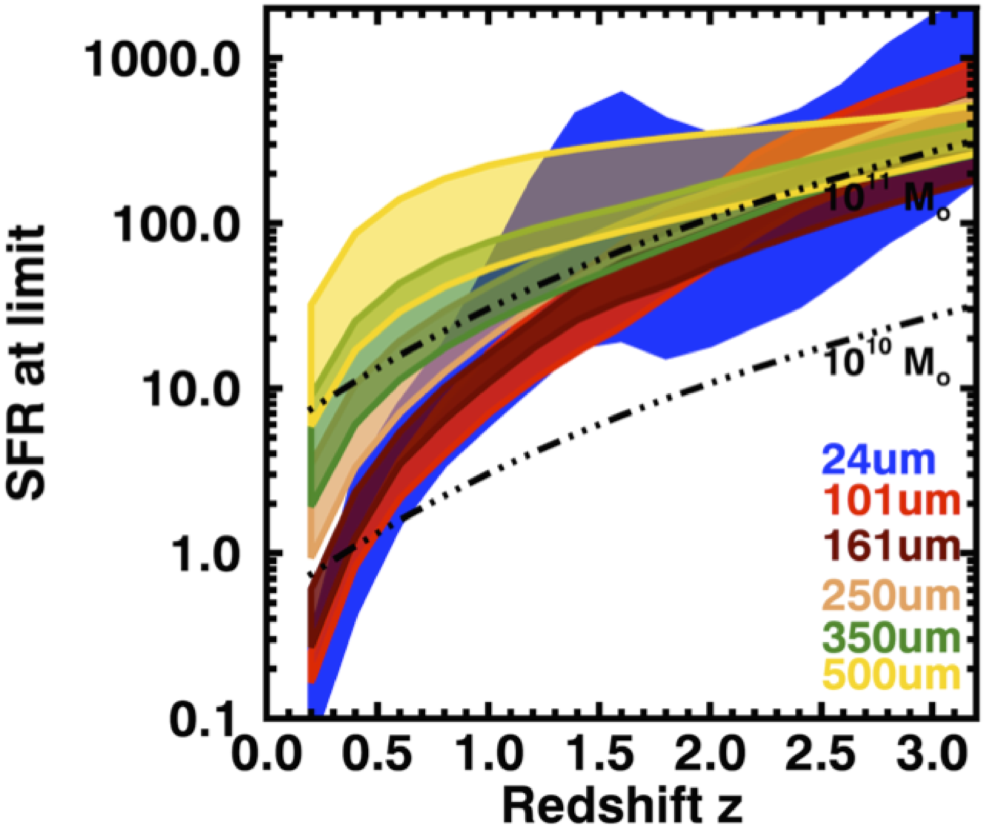


Figure 2.8: The range in minimum SFRs achievable in different wavebands, calculated by normalising the 5 SF templates to the set of 3σ flux limits for the different bands in the COSMOS field, as a function of redshift. The limits of each band are $f_{24} = 0.05$mJy, $f_{100} = 1.2$mJy, $f_{160} = 2.7$mJy, $f_{250} = 7$mJy (from the deblended catalogues used in Chapter 3). Also plotted is the SFR of main sequence star-forming galaxies with stellar masses of M = 10^{10} and 10^{11} M$_{\odot}$ (e.g., Elbaz et al. 2011), the typical masses of galaxies hosting AGN. It is evident that past a redshift of z ~0.5 the MIPS and *Herschel* bands start to only detect the most luminous end of the galaxy population. Consequently, when trying to understand the star-forming properties of a representative sample of AGN, it becomes important to use methods that allows for the inclusion of undetected sources in the analysis.

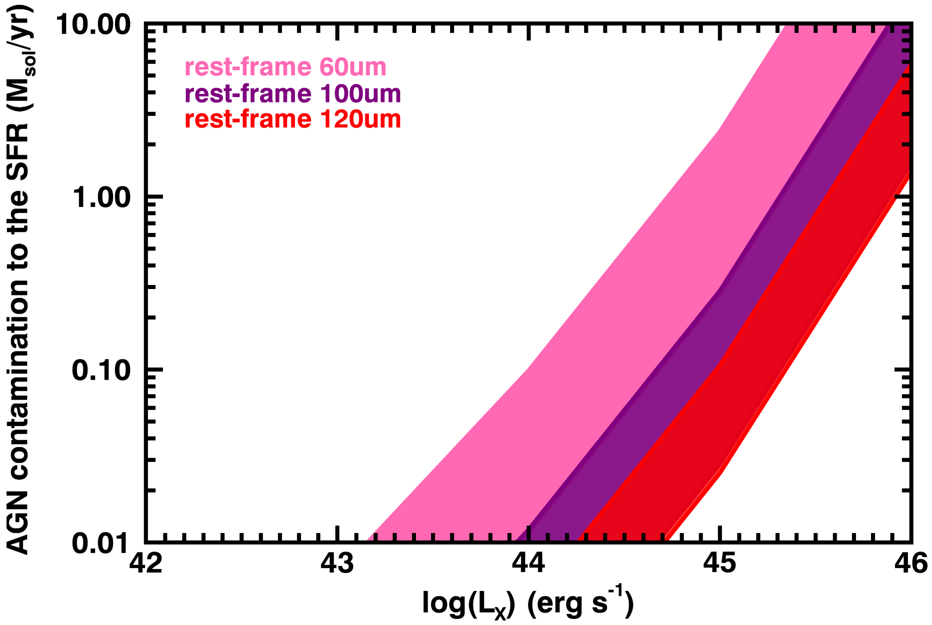


Figure 2.9: The estimated AGN contamination to measurements of the SFR from single bands that correspond to a rest-frame wavelength of 60, 100, or 120μm, as a function of X-ray (2–8keV) luminosity.

the AGN emission for star forming emission. In Figure 2.9, I plot the falsely identified SFR values that are due to the emission from the AGN. AGN contamination will only significantly affect the results for high luminosity X-ray AGN ($L_X > 10^{44}$ erg s^{-1}) in star-forming galaxies with SFRs of a few to a few 10s of $M_\odot$/yr. From the three rest-frame wavelengths explored, as expected the shortest wavelength (60μm) SFR measurements suffers the most from contamination by the AGN. When comparing to Figure 2.8, it is evident that 60μm derivations of the SFR for AGN in galaxies of stellar masses at $10^{10} M_\odot$ up to redshifts of $\sim$3, can be significantly affected by AGN contamination.

A demonstration of the uncertainties of using single-band stacked SFR measurements

A point of contention between FIR studies of X-ray AGN samples, has been the difference in reported trends by different studies (see Fig. 1.10). The majority of the studies that have estimated the mean SFR as a function of AGN luminosity have used one of the three fields of CDF-N (448 arcmin^2 area), CDF-S (464 arcmin^2 area), or COSMOS (2 deg^2 area). Depending on which field the studies explored they found different results, with studies looking at CDF-N (e.g., Page et al. 2012) found decreased mean SFRs at high $L_{2-8\mathrm{keV}}$, studies in CDF-S found increased mean SFRs at high $L_{2-8\mathrm{keV}}$ (e.g., Rovilos et al. 2012), and studies in COSMOS finding constant mean SFRs out to high X-ray luminosities (e.g., Rosario et al. 2012). Harrison et al. (2012) suggested that the discrepancy of results is due to sample variance and small number statistics that affect the small area CDF-N and CDF-S fields.

To investigate what is the driver of the differences in the results for these three fields, I compared the single-band stacking at 250μm, adopted in the majority of studies, with the approach followed in my thesis, of SED fitting and calculating the mean with the K-M estimator. I do this on each of the three fields separately. I take the X-ray AGN with $L_{2-8\mathrm{keV}} > 10^{42}$ erg s^{-1} in the three fields of GOODS-N (in CDF-N), GOODS-S (in CDF-S), and COSMOS, for which I have access to *Herschel* catalogues created by deblending of the FIR maps, and so suffer less from the source confusion that could be boosting the fluxes. The photometry used is presented in detail in Chapter 3. In Fig. 2.10 I show the results of the two different approaches for the calculation of the SFR in comparison to each other. The results for each field are presented in different symbols, with the light colours representing the results from stacking and darker colours the results of my methods. When using the stacking method at 250μm, I find similar results to the previous studies at the different fields, with GOODS-S showing an increase of SFRs and GOODS-N a decrease at the high AGN luminosity bin. The results for the COSMOS field also show an increase with AGN luminosity, but the bins also show significant increase in their mean redshifts that will be driving this trend. When using my methods of SED fitting to deblended photometry and removing the AGN component the results change with all three fields now in agreement in the highest luminosity bin. I demonstrate that once taking into account of the AGN contamination, and using higher quality photometry from decomposition of the FIR images, the difference between the fields disappears. After visually inspecting the maps at the positions of the X-ray AGN in the highest luminosity bin, I found that in the case of CDF-S there were 2 bright neighbouring sources close to the X-ray AGN positions. These could be boosting the 250μm flux density at the close-by

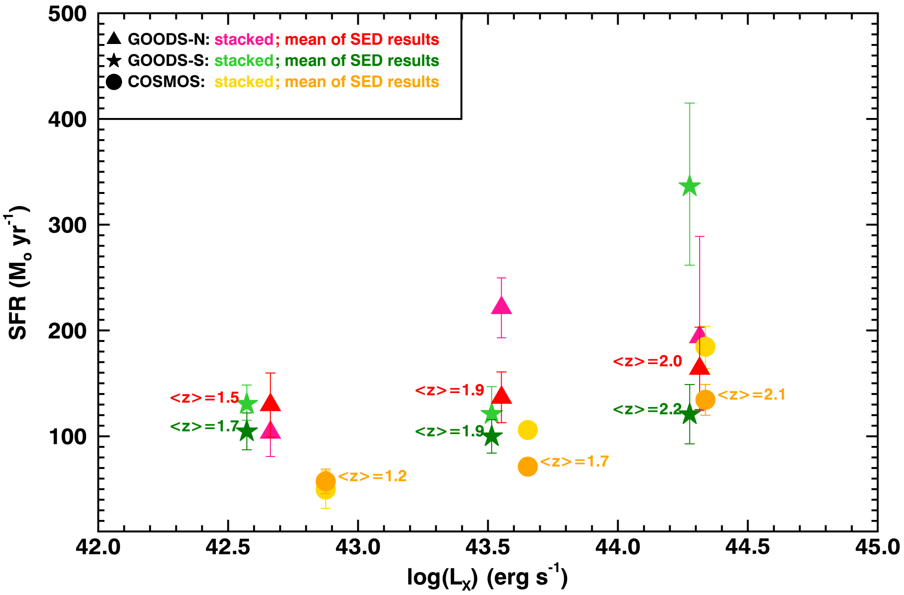


Figure 2.10: Mean SFR measurements using SED fitting and the K-M method compared to stacking at 250μm for X-ray AGN at redshifts z =0–3 in the three fields of GOODS-N, GOODS-S, and C-COSMOS. When stacking I find a similar discrepancy between the three fields as that shown in previous studies (e.g., Harrison et al. 2012). However, when I perform SED fitting to deblended FIR photometry and the K-M method this discrepancy disappears and the resulting SFRs are consistent across all three fields.

positions when stacking is used, leading to the enhanced SFR values. This problem is avoided when using deblended photometric catalogues and SED fitting.

In conclusion, stacking single-band photometry as an estimate of the mean SFR is more sensitive to field-to-field variations due to confusion noise, and can suffer from AGN contamination.

2.7 Discussion

In this chapter I have described the methods followed throughout this thesis for the calculation of individual and mean SFRs from IR photometry. IR SED fitting and decomposition of the AGN and SF SED counterparts are a major part in the analysis of all of the science chapters in this thesis, both in estimating individual SFR measurements and upper limits (Chapter 3 & 5), and for creating composite SEDs and estimating the mean SFRs

(Chapter 4). I have followed two different methods to calculate mean SFRs, depending on depths of the *Herschel* surveys used to define the FIR photometry.

The two methods for the estimation of the mean SFR are: (a) individual IR SED fitting and decomposition into the AGN and SF components, and the combination of the SFR measurements and upper limits in the estimation of the mean using the KM method; and (b) stacking in the multiple FIR wavebands to estimate the mean stacked flux density in all of the FIR bands, combining with the available MIR photometry, and performing IR SED fitting and decomposition to the stacked "source" photometry.

In both cases the result is a mean SFR, for which the AGN IR emission has been accounted for. The advantage of case (a) is that by taking the individual source photometry we can use higher quality processed data, such as deblended source catalogues, and analyse all sources individually. However, even for deep *Herschel* surveys such as GOODS-H and COSMOS, $\sim 70\%$ of sources are still undetected when probing distant sources ($z > 0.5$). This means that even though we can still use the KM-method when studying these fields, it becomes impossible for wider-field and shallower depth surveys, where the fraction of undetected sources will be even larger. In the wider field and shallower depth surveys we have to adopt case (b). As discussed earlier in this chapter, stacking as a method for the estimation of the mean SFR of AGN can suffer from multiple uncertainties. In my methods for the calculation of the SFR through stacking I try to overcome the problems faced by previous studies. By taking account of the confusion background when stacking in the FIR bands, and fitting to the stacked photometry to remove the AGN counterpart of the IR emission, I calculate mean SFRs without the uncertainty of having artificially increased values from contamination from blended sources in the FIR maps, and/or contamination from AGN emission.

Both methods for the estimation of the mean SFRs of galaxies hosting AGN have been developed to provide the best possible SFR estimates based on the available IR photometry.

CHAPTER 3

A remarkably flat relationship between the average star formation rate and AGN luminosity for distant X-ray AGN

Abstract

In this chapter we investigate the relationship between the star formation rate (SFR) and AGN luminosity ($L_{\rm AGN}$) for $\sim$2000 X-ray detected AGN. The AGN span over three orders of magnitude in X-ray luminosity ($10^{42} < L_{2-8\rm keV} < 10^{45.5}$ erg s^{-1}) and are in the redshift range $z = 0.2 – 2.5$. Using infrared (IR) photometry (8 – 500 μm), including deblended *Spitzer* and *Herschel* images and taking into account photometric upper limits, we decompose the IR spectral energy distributions into AGN and star formation components. Using the IR luminosities due to star formation, we investigate the average SFRs as a function of redshift and AGN luminosity. In agreement with previous studies, we find a strong evolution of the average SFR with redshift, tracking the observed evolution of the overall star forming galaxy population. However, we find that the relationship between the average SFR and AGN luminosity is broadly flat at all redshifts and across all the AGN luminosities investigated; in comparison to previous studies, we find less scatter amongst the average SFRs across the wide range of AGN luminosities investigated. By comparing to empirical models, we argue that the observed flat relationship is due to short timescale variations in AGN luminosity, driven by changes in the mass accretion rate, which wash out any underlying correlations between SFR and $L_{\rm AGN}$. Furthermore, we show that the exact form of the predicted relationship between SFR and AGN luminosity (and its normalisation) is highly sensitive to the assumed intrinsic Eddington ratio distribution.

3.1 Introduction

One of the key outstanding problems in studies of galaxy evolution is understanding the connection between active galactic nuclei (AGN) and star formation. Both AGN activity and star formation are predominately dependent on the availability of a cold gas supply from the galaxy, as it is the fuel of both processes, and therefore a first order connection between these two processes may be expected. However the scales of AGN activity and star formation are very different, which has lead to suggestions that any tight connection between them must be due to one process regulating the other (see Alexander & Hickox 2012, Fabian 2012, and Kormendy & Ho 2013 for recent reviews).

There are several pieces of empirical evidence for at least a broad connection between AGN activity and star formation. For example, the tight correlation observed between the mass of the super-massive black hole (SMBH) and the galaxy spheroid for galaxies in the local universe (e.g., Kormendy & Richstone 1995, Magorrian et al. 1998), serves as archaeological evidence of a connection between the growth of the SMBH (through mass accretion, where it becomes visible as AGN activity), and the growth of the galaxy (through star formation). Additionally, observations of AGN have found that the volume average of the SMBH mass accretion rate tracks that of the star formation rate (SFR), within $\sim$3–4 orders of magnitude, up to redshifts of $z \sim 2$ (e.g., Heckman et al. 2004; Merloni et al. 2004; Aird et al. 2010) suggesting a co-evolution of AGN and star forming activity. Despite how significant these results may appear, they only provide indirect evidence for a relationship between AGN activity and star formation and cannot place strong constraints on the form of the relationship.

To acquire more direct evidence on the form of the relationship between AGN activity and star formation requires sensitive measurements of the AGN and star forming luminosities of individual galaxies. X-ray and far–infrared (FIR; $\lambda = 30 - 500\,\mu$m) observations are ideal for quantifying the amount of AGN and star formation activity, respectively. A key advantage of X-ray observations, specifically in the hard band (e.g., 2 – 8 keV), over other tracers of AGN activity, is that they are not greatly affected by the presence of obscuration and contamination effects from the host galaxy (see sections 1 and 2 of Brandt & Alexander 2015 for more details of the use of the X-ray as an AGN tracer). A key advantage of FIR observations, as a measurement of star formation, is that they trace the peak of the obscured emission from star forming regions surrounded by cold gas and dust. Even though the FIR provides an indirect tracer of star formation, a significant advantage over more direct tracers, such as the UV and optical emission from the young massive stars, is that it does not suffer significantly from obscuration (e.g. Kennicutt 1998; Calzetti et al. 2007; Calzetti et al. 2010; see also section 2.2 in Lutz 2014). Indeed, as shown by Domínguez Sánchez et al. (2014a), for luminous infrared galaxies (FIR luminosities of $L_{\rm FIR} \gtrsim 10^{44}$ erg s^{-1}) more than 75% of the total emission due to star formation is produced at FIR wavelengths, a fraction that increases at higher $L_{\rm FIR}$.[1] However, the AGN can also contribute to the FIR luminosity due to the thermal re-radiation of obscuring dust from the surrounding torus (e.g. Antonucci 1993). Hence, for the most reliable measurements of the star formation it is important to apply decomposition methods of the AGN and star formation components at infrared wavelengths (e.g. Netzer et al. 2007; Mullaney et al. 2011; Del Moro et al. 2013; Delvecchio et al. 2014).

A number of studies have used X-ray and FIR observations to understand the connection between distant AGN activity and star formation by measuring the mean SFRs of AGN and star forming galaxy samples (e.g., Lutz et al. 2010; Shao et al. 2010; Mainieri et al. 2011; Mullaney et al. 2012a; Rovilos et al. 2012; Santini et al. 2012; Harrison et al. 2012; Rosario et al. 2013c,a; Lanzuisi et al. 2015). The main results shown by these studies are that: (1) the average star formation rates ($\langle$SFR$\rangle$) of AGN track the increase with redshift found for the overall star forming galaxy population; (2) the $\langle$SFR$\rangle$ of AGN are higher than those of the overall galaxy population (i.e., when including quiescent galaxies); and (3) the specific SFRs (i.e., the ratio of SFR over stellar mass, which serves as a measure of the relative growthrate of the galaxy) of AGN are in quantitative agreement with those of star forming galaxies. The majority of the current studies also find no correlation between the AGN luminosity and $\langle$SFR$\rangle$ for moderate luminosity AGN (X-ray luminosities of $L_{2-8\rm keV} \lesssim 10^{44}$ erg s^{-1}; e.g., Lutz et al. 2010; Shao et al. 2010; Mullaney et al. 2012a; Rovilos et al. 2012; Harrison et al. 2012). However, there are significant disagreements in the results for high luminosity AGN ($L_{2-8\rm keV} \gtrsim 10^{44}$ erg s^{-1}). There are

---

[1] We note that for less luminous infrared galaxies ($L_{\rm FIR} \lesssim 10^{44}$ erg s^{-1}) Domínguez Sánchez et al. (2014a) find that the FIR emission accounts for ~50% of the total emission due to star formation. However in this work we find that our galaxies have average $L_{\rm FIR} \gtrsim 10^{44}$ erg s^{-1} and so the majority of the star formation is expected to be produced at FIR wavelengths.

studies arguing that the $\langle\mathrm{SFR}\rangle$ increases at high AGN luminosities (e.g., Lutz et al. (2010); Rovilos et al. 2012; Santini et al. 2012), a result that seems in agreement with the concept of AGN and star formation activity being connected due to their mutual dependence on the cold gas supply in the galaxy. Other studies have argued that the SFR decreases at high AGN luminosities (e.g., Page et al. 2012; Barger et al. 2014), potentially suggesting that the AGN may be responsible for reducing or even quenching the ongoing star formation (a result also inferred by some simulations of galaxy evolution; e.g., Di Matteo et al. 2005; Hopkins et al. 2005; Debuhr et al. 2012). There are also studies arguing that $\langle\mathrm{SFR}\rangle$ remains constant up to high AGN luminosities (i.e., a broadly flat relationship; e.g., Harrison et al. 2012; Rosario et al. 2012; Azadi et al. 2015), extending the trend seen for moderate luminosity AGN. Nevertheless, the difference in the conclusions of such studies could be attributed to the low source statistics for high luminosity AGN, and strong field to field variations (e.g., Harrison et al. 2012). For example, Harrison et al. (2012) demonstrated that when using a large high luminosity AGN sample the broadly flat relationship between $\langle\mathrm{SFR}\rangle$ and AGN luminosity found for moderate luminosity AGN continues to high luminosities, with no clear evidence for either a positive or negative correlation (see also Harrison 2014 for a recent review).

To first order a flat relationship between $\langle\mathrm{SFR}\rangle$ and AGN luminosity can seem surprising, since it appears to suggest the lack of a connection between AGN activity and star formation. However, Hickox et al. (2014) have shown that a true underlying correlation between AGN luminosity and $\langle\mathrm{SFR}\rangle$ can be masked if the AGN varies significantly (i.e., by more than an order of magnitude) on much shorter timescales than the star formation across the galaxy. In fact, observational studies such as Rafferty et al. (2011), Mullaney et al. (2012b), Chen et al. (2013), Delvecchio et al. (2014), and Rodighiero et al. (2015) have shown that when the average AGN luminosity is calculated as a function of SFR (i.e., taking the average of the more variable quantity as a function of the more stable quantity) a positive relationship is found, suggesting that AGN activity and star formation are correlated on long timescales. Studies using small scale hydrodynamical simulations of SMBH growth (e.g., Gabor & Bournaud 2013; Volonteri et al. 2015) have indeed suggested that AGN activity can vary by a typical factor of ~ 100 over $\sim$Myr timescales, which results in a flat relationship between $\langle\mathrm{SFR}\rangle$ and AGN luminosity over a wide range of AGN luminosity. These studies therefore demonstrate that the relationship between AGN luminosity and $\langle\mathrm{SFR}\rangle$ can potentially place constraints on the variability of mass accretion onto the SMBH in galaxies. However, to date, the observational constraints of the $\langle\mathrm{SFR}\rangle$ of AGN as a function of AGN luminosity and redshift have lacked the accuracy to be able to distinguish between the different SMBH mass accretion models.

Most of the current studies on the $\langle\mathrm{SFR}\rangle$ of distant X-ray AGN suffer from a variety of limitations, which affect the accuracy of $\langle\mathrm{SFR}\rangle$ measurements, such as: (1) small number of sources, which can lead to large statistical uncertainties, particularly at high AGN luminosities; (2) high levels of source confusion at FIR wavelengths, which can cause the overestimation of the flux; (3) use of a single FIR band from which to derive SFRs, which will result in large uncertainties on the $\langle\mathrm{SFR}\rangle$ and will not take into account possible contamination of the SFR measurements from the AGN; (4) neglect of the information that can be obtained from the photometric upper limits of the FIR undetected AGN, which make up the majority of the distant AGN in X-ray samples (this final point is not applicable for studies that use stacking analyses).

In this work we aim to overcome the limitations outlined above by exploiting a large sample of X-ray detected AGN with deep and extensive multi-wavelength data, for which we perform spectral energy distribution (SED) fitting on a source by source basis, and measure the SFR for each source in our sample. We use deblended FIR photometry from *Herschel*, which provides the best constraints on the FIR fluxes of individual sources by reducing the contamination due to blended and confused sources, the most significant drawback of the *Herschel* field maps. Furthermore, we make use of the photometric upper limits in the fitting procedure to better constrain the SED templates and SFRs. We finally calculate the $\langle\mathrm{SFR}\rangle$ values as a function of X-ray luminosity, with the inclusion of sources with only upper limit constraints using survival analysis techniques (e.g., Feigelson & Nelson 1985, Schmitt 1985). Our methods ensure the use of all available data (i.e. photometric detections and upper limits, SFR measurements and upper limits) to provide improved $\langle\mathrm{SFR}\rangle$ as a function of X-ray luminosity and redshift. In Section 3.2 we outline the photometric catalogues used in this work, as well as the choice of redshift and the choice of matching radii between photometric positions. In Section 3.3 we analyse our methods of SED-fitting as well as the calculation of the average IR luminosity due to star formation ($\langle L_{IR,SF} \rangle$). Finally, in Section 3.4 we present and discuss our results. In our analysis we use $H_0 = 71 km\,s^{-1}$, $\Omega_M = 0.27$, $\Omega_\Lambda = 0.73$ and assume a Chabrier (2003) initial mass function (IMF).

3.2 AGN Sample, IR photometry and Redshifts

In this work we use the available Mid–IR (MIR; $\lambda \approx 3-30$) to Far–IR (FIR; $\lambda \approx 30-500\,\mu$m) photometric data to constrain the average SFRs of a large sample of X-ray detected AGN over the redshift range $z \approx 0.2$ – 2.5. To construct a large sample of X-ray detected AGN we make use of three fields with deep X-ray observations: (1) *Chandra* Deep Field North (CDF-N; Alexander et al. 2003a), (2) *Chandra* Deep Field South (CDF-S; Xue et al. 2011), and (3) a combination of *Chandra*-COSMOS (C-COSMOS; Elvis et al. 2009) and *XMM*-COSMOS (Cappelluti et al. 2009). To construct our final AGN sample we obtain the MIR and FIR photometry from observations of the X-ray deep fields made with the *Spitzer* (Werner et al. 2004) and *Herschel* (Pilbratt et al. 2010) space observatories. The recent *Herschel* observational programs PEP/GOODS-*H* (Lutz et al. 2011; Elbaz et al. 2011) and HerMES (Oliver et al. 2012) in the three fields of GOODS-N, GOODS-S, and COSMOS, covering the wavelength range of 70 – 500μm are our main source of the FIR photometry (details in §3.2.2). We therefore restrict the CDF-N and CDF-S X-ray catalogues to these regions with sensitive MIR–FIR coverage, i.e. the GOODS-N and GOODS-S with areas of 187 arcmin^2 each, but use the full 2 deg^2 of COSMOS. In total these areas cover 3609 X-ray sources. Figure 3.1 shows the X-ray sources in GOODS-N, GOODS-S, C-COSMOS, and *XMM*-COSMOS in the $L_{2-8keV} - z$ plane.

In the following subsections we describe our sample selection and the catalogues used for the sample. In §3.2.1 we present the X-ray observations used to define our AGN sample and to determine their X-ray luminosities. In §3.2.2 we present the MIR and FIR photometric catalogues used to constrain the SFRs of the AGN hosts via SED fitting. In §3.2.3 we describe the method of matching the X-ray sources to the MIR and FIR catalogues and the redshift counterparts.

3.2.1 X-ray Data

To select the sample of AGN for our study we use the publicly available X-ray catalogues for the CDF-N (Alexander et al. 2003a), CDF-S (Xue et al. 2010) and COSMOS (Elvis et al. 2009; Cappelluti et al. 2009) fields, restricted to the areas covered by PEP/GOODS-*H* and HerMES observations as described above (see Figure 3.1). For the COSMOS field we use the C-COSMOS X-ray catalogue as the primary sample, while for the sources over the larger region, not covered by *Chandra*, we use the *XMM*-COSMOS catalogue. Rest-frame, hard band 2 – 8 keV luminosities were calculated following Alexander et al. (2003b) with the equation:

$$\mathrm{L}_{2-8\mathrm{keV}} = 4\pi \times \mathrm{D}_{\mathrm{L}}^{2} \times \mathrm{F}_{2-8\mathrm{keV}} \times (1+\mathrm{z})^{(\Gamma-2)} \tag{3.2.1}$$

where $\mathrm{F}_{2-8\mathrm{keV}}$ is the observed X-ray hard band flux (2–8 keV), D_{L} is the luminosity distance, z is the redshift (see Section 3.2.3), and Γ is the photon index used for k-corrections, which was fixed to a standard value of $\Gamma = 1.9$ (e.g., Nandra & Pounds 1994). Although the hard band observations in CDF-N and CDF-S are in the energy range of 2–8 keV, the C-COSMOS and *XMM*-COSMOS catalogues report hard band fluxes of the energy range of 2–10 keV. To convert the 2–10 keV to 2–8 keV fluxes we assume $\Gamma = 1.9$ which yields a conversion factor of 0.85.

For the 20% of X-ray sources in our final sample (see below) not detected in the hard band we used the full band of 0.5 – 8 keV (or the soft band of 0.5 – 2 keV if undetected in the full band) to estimate the hard band flux. We estimated the hard-band flux assuming a $\Gamma = 1.4$ spectral slope, unless this provided a measurement greater than the hard-band upper limit, in which case we assumed a $\Gamma = 2.3$ spectral slope; the assumed range in spectral slope is motivated by the range observed in AGN (e.g., Nandra & Pounds 1994; George et al. 2000). Overall, with this procedure, the hard-band fluxes were estimated assuming $\Gamma = 1.4$ for 19% and $\Gamma = 2.3$ for 1% of sources in our sample (see Figure 3.1).

3.2.2 Mid-IR & Far-IR Data

To measure the SFRs of our AGN sample we need to reliably constrain the IR luminosity due to star formation and remove any contribution from the AGN. To do this we need data covering both the MIR and FIR wavelengths for each source in our sample (e.g., Mullaney et al. 2011). We exploit available photometry in the wavelength range of 8μm – 500μm, provided by observations carried out by: *Spitzer*-IRAC at 8μm; *Spitzer*-IRS at 16μm; *Spitzer*-MIPS at 24μm, 70μm; *Herschel*-PACS at 70, 100, 160μm; and *Herschel*-SPIRE at 250, 350, 500μm. One of the advantages of our study over several previous studies, is the use of catalogues of deblended FIR *Herschel* images (details below). The deblending of sources in the PACS and SPIRE observations allows us to overcome the blending and confusion issues encountered in dense fields that can lead to an overestimation of the flux densities (e.g., Oliver et al. 2012; Magnelli et al. 2013). It also ensures the direct association between the measured FIR flux densities and the sources used as priors in the deblending process. In addition to this, we also make sure that we have a reliable photometric upper limit for sources not detected in the FIR. This enables us to constrain the star forming galaxy templates and gain an upper limit on the IR luminosity due to star formation, as we describe in §3.3.1.

The MIPS 24μm photometric catalogues that we use were created by Magnelli et al. (2013). These catalogues are made by simultaneous PSF fitting to the prior positions of 3.6μm sources. The catalogues were limited to a 3σ detection limit at 24μm going down to 20μJy in GOODS-N and GOODS-S, and 50μJy in COSMOS. The PACS 70μm, 100μm and 160μm catalogues were also created by Magnelli et al. (2013) using the MIPS 24μm detected sources, described above, as the priors for the deblending of the PACS maps. Only sources with at least a 3σ detection at MIPS 24μm were used as priors and the resulting PACS catalogues were also limited to a 3σ detection limit. [2] The SPIRE 250μm, 350μm, and 500μm catalogues were created following the method described in Swinbank et al. (2014), again using these MIPS 24μm positions as priors to deblend the SPIRE maps.

Although both the PACS and SPIRE catalogues have been produced in the same way, Magnelli et al. (2013) do not provide flux upper limits. In order to keep the priored FIR catalogues consistent with each other, we calculate upper limits for the non-detections

---

[2] The PACS catalogues for GOODS-N and GOODS-S are published in Magnelli et al. (2013). The catalogue for COSMOS was created in the same way and is available online (http://www.mpe.mpg.de/ir/Research/PEP/DR1).

Field	AGN	with spec-z	with 24μm
GOODS-N	177	98	137
GOODS-S	209	128	154
COSMOS	1753	914	1151
Total	2139	1140	1442

Table 3.1: Number of X-ray detected AGN in our parent sample ($L_{2-8\text{keV}} > 10^{42}$ erg/s; $z = 0.2$–2.5) in each field, as well as the number of sources with a spectroscopic redshift and the number of sources with a 24 μm counterpart.

in the PACS catalogues of Magnelli et al. (2013) in a similar way to the upper limit calculation performed for the SPIRE priored catalogues of Swinbank et al. (2014). This was done by performing aperture photometry at thousands of random positions in the PACS residual maps and taking the 99.7th percentile of the distribution of the measured flux densities as the 3σ upper limit on the nondetections. To account for the fact that the deblending is more uncertain in regions of luminous sources, we calculated these 3σ upper limits as a function of the pixel values in the original maps (see Swinbank et al. 2014). Consequently, this approach results in upper limits being higher for non-detected sources that lie near a bright source, when compared to non-detected sources in blank areas of the maps.

Due to the fact that we are using MIPS 24μm priored catalogues for the FIR photometry of our sources, any undetected at 24μm will not have FIR measurements in the published catalogues. Therefore for the 24μm undetected sources, we extracted the FIR photometry at the optical counterpart positions following the method described in Swinbank et al. (2014). Overall there are only 23 sources that are undetected at 24μm but have FIR counterparts, making up a very small fraction of our overall sample.

In the MIR bands we also use the catalogues of *Spitzer*–IRAC 8μm observations as described in Wang et al. (2010), Damen et al. (2011), and Sanders et al. (2007), for GOODS-N, GOODS-S, and COSMOS, respectively, as well as *Spitzer*–IRS 16μm from Teplitz et al. (2011) for GOODS-N and GOODS-S. Since all the IRAC catalogues have their detections determined by the 3.6μm maps, and the 16μm catalogues have been produced with the use of 3.6μm priors, they are all consistent with the deblended PACS and SPIRE catalogues described above.

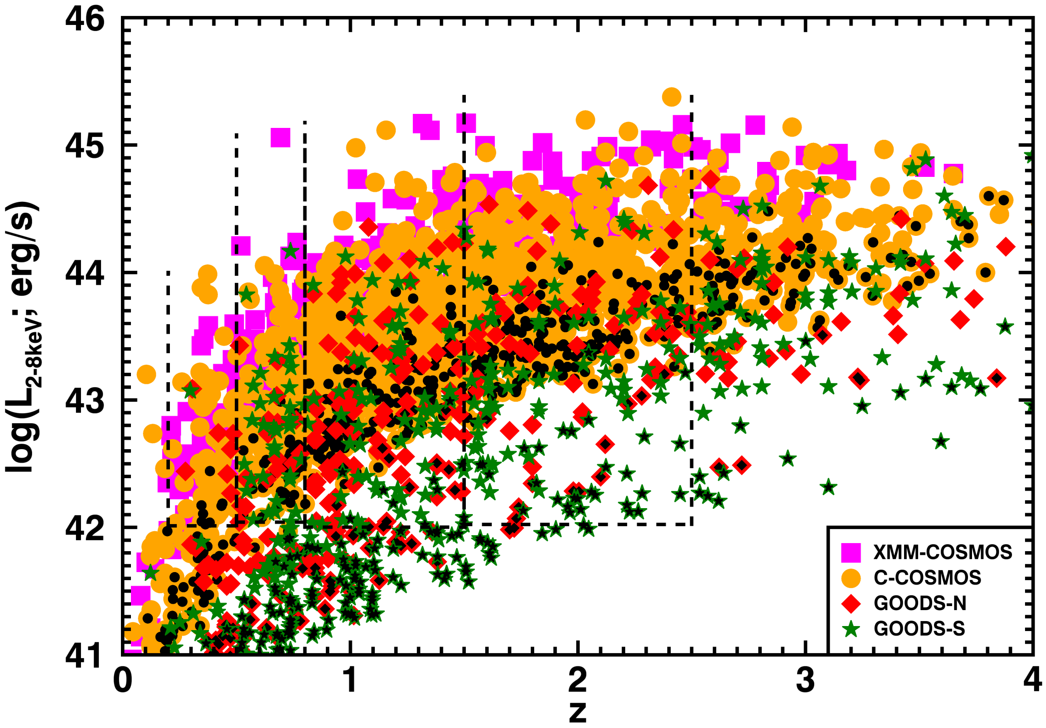


Figure 3.1: X-ray (2-8 keV) luminosity ($L_{2-8\mathrm{keV}}$) versus redshift (z) for the X-ray sources in the GOODS-N, GOODS-S, and COSMOS regions described in §3.2.1. Black centers indicate the X-ray sources without a direct hard-band detection (as described in §3.2.1). The vertical dashed lines indicate the 4 redshift ranges used in this study. The lower X-ray luminosity threshold ($L_{2-8\mathrm{keV}} > 10^{42}$ erg/s) used to define our AGN sample is shown with the horizontal dashed line. The combination of the three fields enables us to explore the SFRs of AGN over three orders of magnitude in AGN luminosity.

3.2.3 Redshifts and catalogue matching

For our SED fitting analysis (see §3.3.1) we need matched catalogues containing X-ray fluxes, MIR-FIR photometric flux densities, and redshifts. To obtain the appropriate counterparts for each X-ray source, we matched the catalogues starting with the X-ray catalogues described in §3.2.1. We first match the positions of the optical counterparts of the X-ray sources to the MIPS 24μm positions in the catalogues of Magnelli et al. (2013).[3] To choose the matching radii between catalogues we measure the number of total matches as a function of radius and estimate the fraction of spurious matches for each matching radius. The matching radius of the X-ray to the MIPS 24μm catalogue for GOODS-N and GOODS-S was 0.8”, while for C-COSMOS and *XMM*-COSMOS it was 1”. This matching radius was chosen to maximise the number of matches while minimising the number of spurious matches, with a ratio of spurious to true matches of 1%. Due to the way that the FIR catalogues were deblended, each MIPS 24μm detected source also has a corresponding photometric measurement or flux upper limit for PACS 70μm, 100μm, 160μm and SPIRE 250μm, 350μm, 500μm (see §3.2.2). For the sources not matched to a MIPS 24μm counterpart we use the FIR data extracted at their optical counterpart positions, as described in §3.2.2. We then match to the IRAC, and to the IRS 16μm catalogues for the two GOODS fields (see §3.2.2) using the same method.

A necessity for this analysis are the redshifts of the X-ray sources. To allocate the redshift counterpart of the sources in GOODS-S and C-COSMOS we make use of the spectroscopic and photometric redshift compilation by Xue et al. (2011) and Civano et al. (2012), respectively. We also added redshifts from Teplitz et al. (2011) for sources in GOODS-S when necessary. For the sources in GOODS-N we created our own compilation using catalogues of spectroscopic redshifts from Barger et al. (2008) and Teplitz et al. (2011) and photometric redshifts from Wirth et al. (2004) and Pannella et al. (2009). Overall we obtained redshifts for 91.4% of the X-ray sources.

In total there are 3297 X-ray sources covered by *Chandra*, *XMM*, and PEP/GOODS-*H* observations with a redshift (see Figure 3.1). For this study we restrict this sample to redshifts of $z = 0.2 – 2.5$ and a luminosity range of $\mathrm{L}_{2-8\mathrm{keV}} > 10^{42}$ erg s^{-1}, resulting in

---

[3] For the X-ray catalogues of CDF-S and C-COSMOS the optical counterparts are provided by Xue et al. (2010) and Elvis et al. (2009). For the sources in CDF-N we use the catalogue of Barger et al. (2008).

our parent sample of 2139 AGN. Of the parent sample 53.3% have spectroscopic redshifts and 67.4% are detected at MIPS-24μm (see Table 3.1 for a summary of the three fields).

3.3 Data analysis

In this study we are interested in measuring the mean SFRs of galaxies, hosting an X-ray detected AGN, as a function of the AGN luminosity and redshift. We use multi-band IR photometry, including photometric upper limits, to perform SED fitting for all 2139 X-ray detected AGN in our parent sample (see §3.2.3; Figure 3.1). For each source we decompose the contribution of AGN activity and star formation to the overall SED. This allows us to measure the IR luminosity due to star formation ($L_{\rm IR,SF}$), the key quantity for this study, which we can use as a proxy for SFR (e.g., Kennicutt 1998, Calzetti et al. 2007, Calzetti et al. 2010). In §3.3.1 we outline the SED fitting procedure and describe the calculation of $L_{\rm IR,SF}$. In §3.3.2 we describe the method that we follow for the calculation of the average $L_{\rm IR,SF}$ as a function of $L_{2-8\rm keV}$ (our tracer of the AGN luminosity) for the whole sample, where we include both direct $L_{\rm IR,SF}$ measurements and upper limits. The calculation of these values thus allows us to investigate how SFR relates to AGN luminosity (Section 3.4).

3.3.1 SED fitting procedure

To calculate individual $L_{\rm IR,SF}$ values for our sample we perform SED fitting to the MIR and FIR photometry. In these bands there could be a contribution from both AGN and star formation, with emission from the AGN peaking at MIR wavelengths and dropping off at the FIR wavelengths (e.g., Netzer et al. 2007; Mullaney et al. 2011). Those factors make it important to decompose the contribution from both star formation and AGN to the overall SED so as to avoid an overestimation of the SFR measurement. In Figure 3.2 we give example SED fits to demonstrate our procedure.

To fit and decompose the IR SED of our sources we develop the publicly available DecompIR code of Mullaney et al. (2011), and use the 8 – 500μm data and upper limits described in §3.2.2. We use a set of empirical templates that consist of the mean AGN template and the five star forming galaxy templates originally defined in Mullaney et al.

(2011), and extended by Del Moro et al. (2013) to cover the wide wavelength range of 3 – 10^5 μm. We also include the Arp220 galaxy template from Silva et al. (1998) which serves as a sixth template to ensure that we are also covering the possibility of extremely dusty star forming systems. The advantage of using a few, but representative, templates to fit the data is that we can avoid the degeneracy in the fitting procedure caused by a large number of templates. Furthermore as many of our sources have limited photometric detections (with only one or two data points), it is sensible to keep the number of free parameters as small as possible. We note that the set of star forming galaxy templates described above covers a broad range of empirical shapes, including the large template library of Chary & Elbaz (2001), as shown in Figure 2 of Del Moro et al. (2013), and the templates described by Kirkpatrick et al. (2012).

In our fitting procedure the only free parameters of the fit are the normalisation of the star forming galaxy and AGN templates. Since there are two free parameters in the fit we require that the source has at least three photometric detections to simultaneously fit the AGN and star forming galaxy templates. When we have less than three photometric detections we can only derive upper limits on $\mathrm{L_{IR,SF}}$, as we cannot constrain the AGN contribution (see below).

When a source is detected in three or more photometric bands we perform a series of fits following the method of Del Moro et al. (2013). We fit the data in two steps: firstly we fit using each of the six star forming galaxy templates separately without including the AGN component, and secondly we fit again with each of the star forming galaxy templates in combination with the AGN template. We fit to the photometric flux density detections, but use the available flux density upper limits to eliminate the fits which are above any of the upper limits. This procedure results in a maximum of twelve models (the six star forming galaxy templates without an AGN and the six star forming galaxy templates with an AGN) to chose from.

To determine the best fitting solution of the twelve possibilities described above, we use the Bayesian Information Criteria (BIC; Schwarz 1978) which allows the objective comparison of different non-nested models with a fixed data set, and is defined as:

$$\mathrm{BIC} = -2 \times \mathrm{lnL} + \mathrm{k} \times \mathrm{lnN} \tag{3.3.2}$$

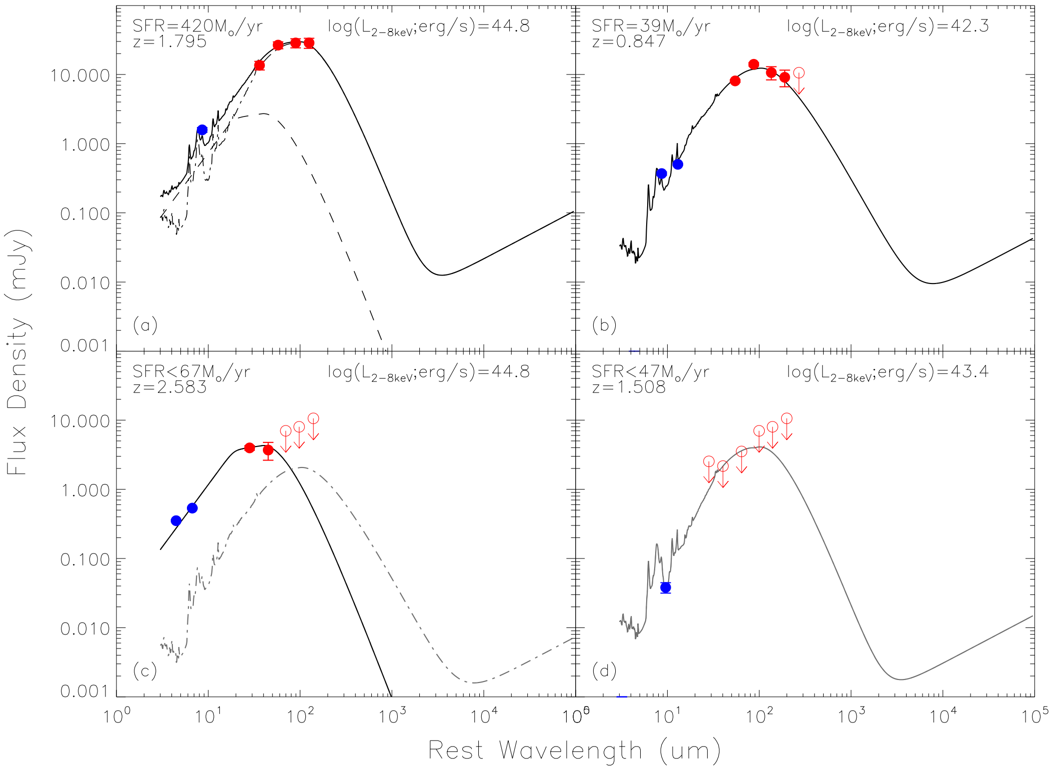


Figure 3.2: Examples of the four types of best fitting SED solutions. (a) A galaxy where the best fit (solid curve) is the combination of AGN (dashed) and star forming galaxy (dot-dashed curve) templates. (b) A galaxy where the best fit is that of a star forming galaxy template alone, with no AGN contribution. (c) A galaxy where the best fit solution is an AGN (solid curve) with no star formation contribution, in this case we calculated an upper limit on the star forming component shown by the grey dot-dashed curve. (d) A galaxy with only one photometric detection where we can only calculate an upper limit for the star forming galaxy templates, as shown here by the grey curve. In all four cases the blue data points are from *Spitzer* observations, while red data points are from *Herschel* observations. The filled circles are measured flux densities, while the empty circles with an arrow are the flux density upper limits. For each example we also give the SFR, X-ray luminosity, and redshift of the source. The wavelengths have been shifted to the rest frame, but the flux densities are in the observed frame.

where L is the maximum likelihood, k is the number of free parameters, and N the number of data points. This method penalises against models with extra free parameters counterbalancing the fact that a model with more free parameters can fit the data better, irrespective of the relevance of the parameters. This is an improvement over a simple $\Delta\chi^2$ test or a maximum likelihood comparison that would tend to favour the model with more free parameters. For each source the BIC value is calculated for all of the different fits. The best fitting model will be the one which minimises the BIC value, its absolute value being irrelevant; however for one model to be significantly better than the others it needs to have a difference in BIC value of $\Delta\mathrm{BIC} \geq 2$. If $\Delta\mathrm{BIC} \leq 2$ then both models are considered equally valid (e.g. Liddle 2004). Our final best fit solution is the one with the lowest BIC value; however we only accept the AGN component as significant if the inclusion of it reduces the BIC value by ≥ 2. In Figure 3.2(a) we show a best fit SED that includes the AGN and star formation component, and in Figure 3.2(b) a best fit SED with only the star formation component. From the best fit SEDs we then measure the integrated $8-1000\,\mu$m IR luminosity of the star formation component ($\mathrm{L_{IR,SF}}$). Furthermore, if multiple fits have BIC values equal to the minimum BIC value, we consider them equally valid and take the average of their derived $\mathrm{L_{IR,SF}}$.

For sources detected in fewer than three photometric bands we can only calculate upper limits on $\mathrm{L_{IR,SF}}$, due to the insufficient degrees of freedom to calculate the AGN contribution to the IR luminosity. To calculate the upper limits of the normalisation of each star forming galaxy template we increase the normalisation of each template until it reaches one of the 3σ upper limits, or exceeds the 3σ uncertainty of a data point. We take the star forming galaxy template with the highest upper limit of $\mathrm{L_{IR,SF}}$ as our conservative upper limit for that source (e.g. see Figure 3.2(d)). Using the same method we also derive upper limits on the star formation contribution for sources where the best fit is fully dominated by the AGN (e.g. see Figure 3.2(c)).

Due to the limited photometry and quality of the data, our procedure is not expected to significantly detect an AGN component in the IR SEDs of all sources. Indeed, the detection of the AGN component in the MIR will be dependent on the relative ratio of $\mathrm{L_{IR,SF}}$ over the IR luminosity due to the AGN ($\mathrm{L_{IR,AGN}}$); for example, a source with a high ratio of $\mathrm{L_{IR,SF}}$ over $\mathrm{L_{IR,AGN}}$ will not show strong evidence of an AGN component in its IR SED (e.g., see Appendix A of Del Moro et al. 2013). However, we note that if we force an AGN component to be present in the IR SEDs of each of our sources, our results of mean $L_{IR,SF}$ in bins of X-ray luminosity and redshift (see §3.3.2) only change within a $\sim$5% level, which is smaller than the uncertainty of the mean $L_{IR,SF}$ results presented in §3.4.1. We also verified that our results were not sensitive to the choice of AGN template that we used by refitting sources with two different AGN templates. One template is representative of low luminosity AGN, while the other template is representative of high luminosity AGN, as provided by Mullaney et al. (2011). The first template is "colder" than that used in our main analysis, with less emission in the MIR and extended emission to the FIR wavelengths, and the second template is "hotter", with most emission occurring at MIR wavelengths and a steep drop-off in the FIR (in agreement with the mean empirical templates of Quasars in the FIR; e.g. Netzer et al. 2007). Between them, these two templates, encompass most clumpy-torus models (see Fig. 7 in Mullaney et al. 2011). In both cases our results of mean $L_{IR,SF}$ in bins of X-ray luminosity and redshift (see §3.3.2) only change within a $\sim$10% level, which again is smaller than the uncertainty in the mean $L_{IR,SF}$ results presented in §3.4.1.

Using our SED fitting approach we have a sample of 2139 AGN with individual measurements (including upper limits) of $L_{IR,SF}$. From our results for the whole sample there are 263 fits that required a significant AGN component in addition to star formation, 274 fits that required only the star forming galaxy template, and for 1602 sources only upper limits on the star formation component could be derived due to limited photometry.

3.3.2 Calculating average source properties

For this study we aim to constrain the average star formation rates of our X-ray AGN sample as a function of redshift and X-ray luminosity. A challenge for all studies using *Herschel* FIR photometry is the low detection rate of individual sources (e.g., Mullaney et al. 2012a). In our sample we can only place upper limit constraints on the $L_{IR,SF}$ for many of our sources, i.e. 1612 out of the 2139 (75.4%) sources in our sample, due to the limited photometry or because they are AGN dominated. In our study we have placed conservative upper limits on the $L_{IR,SF}$ for the AGN for which it was not possible to directly identify the star formation component (see §3.3.1). In order to not bias our study to only the FIR bright sources we study the average properties of the whole X-ray selected AGN sample by using a Survival Analysis technique (e.g., Feigelson & Nelson 1985, Schmitt 1985) to calculate the mean IR luminosities with the inclusion of all of the upper limits (details below).

We divide our sample in to four redshift ranges, $z = 0.2 – 0.5$, $0.5 – 0.8$, $0.8 – 1.5$, and $1.5 – 2.5$. For each redshift range we also divide the sample in to L_{2-8keV} bins determined such that they included ≈40 sources in each bin. To ensure that all of the sources within the redshift range are included we allow the number to vary slightly, resulting in bins of 40 – 43 sources. For each $L_{2-8keV} – z$ bin we calculate the mean IR luminosity due to star formation ($\langle L_{IR,SF} \rangle$; see §3.3.1) and mean X-ray luminosity ($\langle L_{2-8keV} \rangle$; see §3.2.1). To calculate the $\langle L_{IR,SF} \rangle$ values, with the inclusion of upper limits, we use the Kaplan-Meier product limit estimator (Kaplan & Meier 1958), a non-parametric maximum-likelihood-type estimator of the distribution function. We use the formula as described in Feigelson & Nelson (1985) for the estimation of the mean of a sample including the upper limit values. The advantage of this method is that it does not assume an underlying distribution. We will refer to this method as the K-M method for the rest of this paper.

The main requirement for the use of the K-M method, is for the upper limit values to be randomly distributed among the measured values and independent of them. Due to the different types of upper limits that result from our fitting procedure (see §3.3.1) the upper limits on $L_{IR,SF}$ are indeed random enough for the use of this method. [4] Furthermore, a K-S test on our $L_{2-8keV} – z$ bins, with a probability threshold of 1%, shows no evidence of the distributions of upper limits and measured values being drawn from different distributions. This method also requires that the lowest $L_{IR,SF}$ value in each bin is a measurement and not an upper limit. For the 12 bins where this is not the case we follow the popular procedure amongst studies using this method, and assume that the lowest value is a measurement (e.g., Feigelson & Nelson 1985, Zhong & Hess 2009). These 12 bins

---

[4] Our SED fitting procedure provides upper limits for the cases where a source is: (a) MIR – FIR undetected; (b) MIR – FIR detected, but in less than three bands; (c) AGN dominated, i.e. the fit doesn't require any contribution from the SF templates. In the case of (a) the upper limits are calculated by constraining the SF templates to the flux upper limits, while in the cases of (b) and (c) they are calculated by constraining the SF templates to the 3σ flux errors or the flux upper limits. The fact that bright IR sources can meet the criteria of (b) and (c), in combination with the spatial dependence of the FIR flux upper limits, helps drive the similarity between the distributions of the $L_{IR,SF}$ upper limits and measurements.

are randomly distributed with $L_{2-8\mathrm{keV}}$ and redshift (see Table 3.2) , and therefore do not affect our conclusions on the trends of $\langle L_{\mathrm{IR,SF}} \rangle$ with redshift, and $L_{2-8\mathrm{keV}}$.

Feigelson & Nelson (1985) use the K-M method to estimate means with up to a censorship (i.e., the fraction of upper limits) of 90%, but argue that there can be a significant bias in such cases. Additionally, a study by Zhong & Hess (2009) estimating the bias of this method for a wide range of distribution types, find that the estimated means are within a factor of 2 for up to 80–90% censorship levels. In our work we have imposed a limit of 90% censorship on our bins, and have discarded 7 bins with greater censorship. The median censorship level amongst the remaining 45 bins we have used in our analysis is ∼73%, with 11 of them having a censorship of 80–90% (see Table 3.2). For the calculation of the uncertainty on the mean we use the bootstrap technique, for which we take 10000 random samplings in each bin and recalculate the mean. We then take the 16th and 84th percentiles of the overall distribution as the 1σ errors. As discussed above, bins of high censorship levels could suffer from additional uncertainties of a factor of $\lesssim 2$. However, when comparing to the results of the stacking procedure, we find that the two methods are consistent (see Appendix), and hence, we do not have concerns about the high censorship levels in our bins causing a significant systematic bias.

We show our final results of $\langle L_{\mathrm{IR,SF}} \rangle$ as a function of $L_{2-8\mathrm{keV}}$ in Figure 3.3. In our plots, throughout Section 3.4, we also include axes of SFR and AGN bolometric luminosity (L_{AGN}) to help interpret the $L_{\mathrm{IR,SF}}$ and $L_{2-8\mathrm{keV}}$ measurements. We calculate L_{AGN} from $L_{2-8\mathrm{keV}}$ by using the luminosity dependent relation of Stern (2015) to convert the $L_{2-8\mathrm{keV}}$ to an AGN 6μm luminosity density. We then multiply this by a factor of 8 to convert the 6μm luminosity density to L_{AGN} (following Richards et al. 2006). The SFRs were calculated from the $\langle L_{\mathrm{IR,SF}} \rangle$ with the use of the Kennicutt (1998) relation corrected to a Chabrier IMF (Chabrier 2003).

3.4 Results & Discussion

In this section we present our results and explore the form of the relationship between the average SFR, $\langle L_{\mathrm{IR,SF}} \rangle$, and X-ray luminosity, $L_{2-8\mathrm{keV}}$, for our sample of 2139 X-ray detected AGN (see Section 3.2) . In section 3.4.1 we present our results of average

SFR (calculated from $\langle L_{IR,SF}\rangle$) as a function of X-ray (and bolometric) AGN luminosity for four redshift ranges within $z = 0.2 - 2.5$. In section 3.4.2 we compare the SFR of the AGN to those of the overall star forming galaxy population, for a subsample of our sources with reliable host-galaxy masses. In section 3.4.3 we compare our results to the predictions from two empirical models that connect AGN activity to star formation.

3.4.1 Mean star formation rate as a function of X-ray luminosity

The main focus of this chapter is to determine the form of the relationship between the average SFR and AGN X-ray luminosity over 4 redshift ranges. The results of our analysis as described in §3.3.2 are presented in Figure 3.3 and Table 3.2. In Figure 3.3 the data are colour coded by redshift where each point is the mean of ≈40 sources and the error bars correspond to the 1σ of the bootstrap errors (see §3.3.2). We find that the $\langle L_{IR,SF}\rangle$ (and hence $\langle SFR\rangle$) increases with redshift, by a factor of ~3 between each redshift range, in agreement with both the observed evolution found for normal star forming galaxies (e.g. Elbaz et al. 2011; Schreiber et al. 2015) and previous studies on AGN populations (e.g. Shao et al. 2010; Rovilos et al. 2012; Rosario et al. 2012; Mullaney et al. 2012a). However for the individual redshift ranges we find no strong correlation between $\langle L_{IR,SF}\rangle$ and L_{2-8keV}, a result inconsistent with that suggested by some other studies which have reported a rise or fall of $\langle L_{IR,SF}\rangle$ at high X-ray luminosities (e.g., Lutz et al. 2010; Page et al. 2012; Rovilos et al. 2012; although see Harrison et al. 2012).

We find that our results are in general agreement to those studies that stack the FIR data to derive SFRs using large numbers of sources (e.g., Harrison et al. 2012, Rosario et al. 2012); however our results have reduced scatter and reduced uncertainties on the AGN contribution to the IR luminosity. We look in more detail at how our results compare to those of stacking in the Appendix of this paper. Additionally, we compare our results directly to those of Rosario et al. (2012), who explore the average 60μm luminosity ($\nu L_{60\mu m}$) values (as a tracer of SFR) in the same redshift ranges as our study, by stacking *Herschel*–PACS data. We use the average difference between $\nu L_{60\mu m}$ and $L_{IR,SF}$ from our SED fitting results, $L_{IR,SF} / \nu L_{60\mu m} = 2.2$, to convert the results of Rosario et al. (2012) to $L_{IR,SF}$. In Figure 3.4 we plot our results in comparison to those of Rosario et al. (2012) (hollow black symbols) and find broad agreement with our results both as

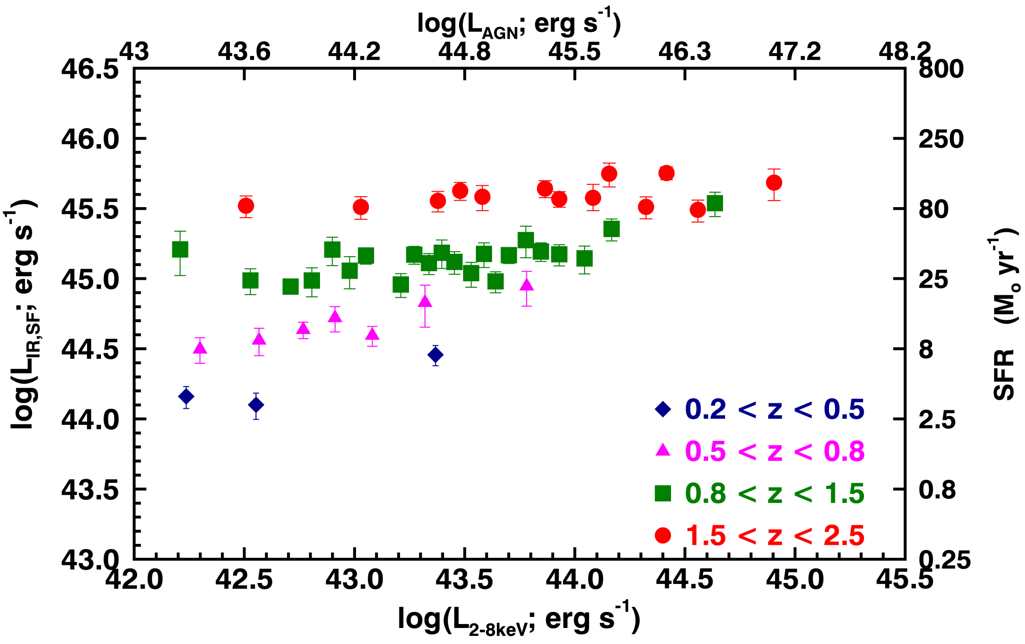


Figure 3.3: Mean IR luminosity due to star formation, $\langle L_{IR,SF} \rangle$, as a function of X-ray luminosity, $\langle L_{2-8keV} \rangle$, for four redshift ranges. Each L_{2-8keV} bin has ∼40 sources. We also give the corresponding SFR values using the Kennicutt (1998) relation corrected to a Chabrier IMF (Chabrier 2003), and the bolometric AGN luminosity L_{AGN} calculated from L_{2-8keV} using the luminosity dependent relation of Stern (2015). The errors on the $\langle L_{IR,SF} \rangle$ are calculated using the bootstrap analysis as described in §3.3.2 (see also §3.3.2 for a discussion on the additional uncertainties).

$\langle z\rangle$	$\langle L_{2-8keV}\rangle$	$\langle L_{AGN}\rangle$	Censorship	$\langle L_{IR,SF}\rangle$	$\langle SFR\rangle$	flag
	erg s^{-1}	erg s^{-1}	%	erg s^{-1}	$M_\odot/yr^{-1}$	
0.38	1.7×10^{42}	1.9×10^{43}	70	$1.4^{+0.3}_{-0.3}\times10^{44}$	4^{+1}_{-1}	1
0.36	3.6×10^{42}	4.5×10^{43}	72	$1.3^{+0.3}_{-0.3}\times10^{44}$	3^{+1}_{-1}	0
0.39	2.3×10^{43}	6.9×10^{44}	54	$2.9^{+0.5}_{-0.5}\times10^{44}$	8^{+1}_{-1}	0
0.65	2.0×10^{42}	2.3×10^{43}	70	$3.1^{+0.6}_{-0.7}\times10^{44}$	8^{+2}_{-2}	0
0.68	3.7×10^{42}	4.6×10^{43}	79	$3.6^{+0.8}_{-0.8}\times10^{44}$	10^{+2}_{-2}	1
0.66	5.9×10^{42}	7.9×10^{43}	74	$4.3^{+0.6}_{-0.6}\times10^{44}$	11^{+2}_{-2}	1
0.68	8.2×10^{42}	1.2×10^{44}	67	$5.2^{+1.1}_{-1.1}\times10^{44}$	14^{+3}_{-3}	0
0.67	1.2×10^{43}	1.9×10^{44}	70	$3.9^{+0.6}_{-0.6}\times10^{44}$	10^{+2}_{-2}	0
0.68	2.1×10^{43}	3.9×10^{44}	67	$6.8^{+2.2}_{-2.2}\times10^{44}$	18^{+6}_{-6}	0
0.67	6.0×10^{43}	1.8×10^{45}	61	$8.8^{+2.4}_{-2.4}\times10^{44}$	23^{+6}_{-7}	0
1.11	1.6×10^{42}	1.8×10^{43}	48	$1.6^{+0.6}_{-0.6}\times10^{45}$	43^{+15}_{-15}	0
1.04	3.4×10^{42}	4.2×10^{43}	75	$9.7^{+2.0}_{-2.0}\times10^{44}$	26^{+5}_{-5}	0
1.02	5.1×10^{42}	6.8×10^{43}	70	$8.8^{+1.0}_{-1.0}\times10^{44}$	23^{+3}_{-3}	0
1.0	6.4×10^{42}	8.8×10^{43}	68	$9.7^{+2.2}_{-2.3}\times10^{44}$	26^{+6}_{-6}	0
1.1	7.9×10^{42}	1.1×10^{44}	65	$1.6^{+0.4}_{-0.4}\times10^{45}$	43^{+10}_{-10}	0
1.1	9.5×10^{42}	1.4×10^{44}	73	$1.1^{+0.3}_{-0.3}\times10^{45}$	30^{+8}_{-8}	0
1.09	1.1×10^{43}	1.8×10^{44}	73	$1.5^{+0.2}_{-0.2}\times10^{45}$	39^{+5}_{-5}	1
1.07	1.6×10^{43}	2.8×10^{44}	88	$9.0^{+1.8}_{-1.8}\times10^{44}$	24^{+5}_{-5}	0
1.15	1.9×10^{43}	3.3×10^{44}	78	$1.5^{+2.2}_{-2.2}\times10^{45}$	39^{+6}_{-6}	1
1.13	2.2×10^{43}	4.1×10^{44}	78	$1.3^{+0.2}_{-0.2}\times10^{45}$	34^{+6}_{-6}	1
1.14	2.5×10^{43}	4.9×10^{44}	75	$1.5^{+0.4}_{-0.3}\times10^{45}$	40^{+9}_{-9}	0
1.14	2.8×10^{43}	5.8×10^{44}	68	$1.3^{+0.2}_{-0.2}\times10^{45}$	35^{+6}_{-6}	0
1.17	3.4×10^{43}	7.3×10^{44}	88	$1.1^{+0.2}_{-0.2}\times10^{45}$	29^{+6}_{-6}	1
1.14	3.9×10^{43}	8.7×10^{44}	73	$1.5^{+0.3}_{-0.3}\times10^{45}$	40^{+8}_{-8}	0
1.11	4.4×10^{43}	1.0×10^{45}	80	$9.6^{+1.6}_{-1.6}\times10^{44}$	25^{+4}_{-4}	0
1.13	5.0×10^{43}	1.2×10^{45}	65	$1.5^{+0.2}_{-0.2}\times10^{45}$	39^{+5}_{-5}	0
1.14	6.0×10^{43}	1.6×10^{45}	75	$1.9^{+0.5}_{-0.5}\times10^{45}$	50^{+13}_{-13}	0
1.19	7.0×10^{43}	2.0×10^{45}	85	$1.6^{+0.2}_{-0.2}\times10^{45}$	41^{+6}_{-6}	1
1.16	8.5×10^{43}	2.6×10^{45}	68	$1.5^{+0.3}_{-0.3}\times10^{45}$	39^{+7}_{-7}	1
1.2	1.1×10^{44}	3.9×10^{45}	78	$1.4^{+0.3}_{-0.3}\times10^{45}$	37^{+7}_{-7}	0
1.2	1.5×10^{44}	5.9×10^{45}	58	$2.3^{+0.4}_{-0.4}\times10^{45}$	60^{+11}_{-11}	0
1.31	4.3×10^{44}	4.5×10^{46}	68	$3.5^{+0.7}_{-0.7}\times10^{45}$	91^{+18}_{-18}	0

Table 3.2: The average redshift, X-ray luminosity, AGN bolometric luminosity, IR luminosity due to star formation, and SFR, for the data presented in Figure 3.3. The errors on the $\langle L_{IR,SF}\rangle$ are calculated using the bootstrap analysis (see §3.3.2). We also provide the censorship level of each bin, and a flag indicating when the minimum value of the sources in that bin is an upper limit (when the flag has a value of 1), which can result to an extra uncertainty on the $\langle L_{IR,SF}\rangle$ (see §3.3.2).

$\langle z \rangle$	$\langle L_{2-8\mathrm{keV}} \rangle$	$\langle L_{\mathrm{AGN}} \rangle$	Censorship	$\langle L_{\mathrm{IR,SF}} \rangle$	$\langle \mathrm{SFR} \rangle$	flag
1.88	3.2×10^{42}	4.0×10^{43}	61	$3.3^{+0.6}_{-0.6}\times10^{45}$	88^{+15}_{-16}	0
1.83	1.1×10^{43}	1.7×10^{44}	73	$3.2^{+0.6}_{-0.6}\times10^{45}$	86^{+16}_{-16}	0
1.86	2.4×10^{43}	4.6×10^{44}	85	$3.6^{+0.6}_{-0.6}\times10^{45}$	94^{+16}_{-16}	1
1.9	3.0×10^{43}	6.3×10^{44}	76	$4.2^{+0.6}_{-0.6}\times10^{45}$	112^{+16}_{-16}	0
1.88	3.8×10^{43}	8.5×10^{44}	81	$3.8^{+0.8}_{-0.8}\times10^{45}$	101^{+21}_{-21}	0
2.02	7.3×10^{43}	2.1×10^{45}	83	$4.4^{+0.6}_{-0.6}\times10^{45}$	116^{+16}_{-16}	1
1.94	8.5×10^{43}	2.6×10^{45}	78	$3.7^{+0.5}_{-0.5}\times10^{45}$	98^{+12}_{-12}	0
1.95	1.2×10^{44}	4.4×10^{45}	85	$3.8^{+0.9}_{-0.7}\times10^{45}$	100^{+25}_{-19}	0
1.89	1.4×10^{44}	5.7×10^{45}	71	$5.6^{+1.1}_{-1.1}\times10^{45}$	148^{+28}_{-29}	0
2.01	2.1×10^{44}	1.0×10^{46}	81	$3.2^{+0.6}_{-0.6}\times10^{45}$	86^{+15}_{-15}	0
1.94	2.6×10^{44}	1.4×10^{46}	76	$5.7^{+0.6}_{-0.6}\times10^{45}$	150^{+16}_{-16}	0
1.91	3.6×10^{44}	2.5×10^{46}	85	$3.1^{+0.5}_{-0.6}\times10^{45}$	82^{+14}_{-15}	0
2.09	8.0×10^{44}	1.2×10^{47}	83	$4.8^{+1.2}_{-1.2}\times10^{45}$	127^{+32}_{-33}	1

Table 3.3: Table 3.2 continued.

a function of redshift and $L_{2-8\mathrm{keV}}$, although we have more $L_{2-8\mathrm{keV}}$ bins and our results show less scatter. To compare to the highest $L_{2-8\mathrm{keV}}$ bins of Rosario et al. (2012) at the redshift ranges of $z = 0.8 - 1.5$ and $z = 1.5 - 2.5$ we calculate the $\langle L_{\mathrm{IR,SF}} \rangle$ for the five highest $L_{2-8\mathrm{keV}}$ sources in our study in both of these redshift ranges (plotted in Figure 3.4 with solid black symbols). We find that our highest $L_{2-8\mathrm{keV}}$ sources are in agreement with those of Rosario et al. (2012); however, due to the very small number of sources in these bins (5 – 23 sources across both studies), we do not interpret them any further.

To asses the contribution of the upper limits on the overall mean, we take an extreme scenario where all upper limits are assumed to correspond to zero values. We find that $\langle L_{\mathrm{IR,SF}} \rangle$ can drop by 0.2 dex (factor of 1.6) at $0.2 < z < 0.5$, by 0.3 dex (factor of 2) at $0.5 < z < 0.8$ and $0.8 < z < 1.5$, and by 0.4 dex (factor of 2.5) at the highest redshift range of $1.5 < z < 2.5$. However, we note that the form of the observed *flat relationship* of $\langle L_{\mathrm{IR,SF}} \rangle$ with $L_{2-8\mathrm{keV}}$ (Figure 3.3) shows little to no change for all redshift ranges, in this extreme scenario.

To test whether our results are consistent with a flat trend of $\langle L_{\mathrm{IR,SF}} \rangle$ with $L_{2-8\mathrm{keV}}$ we show in Figure 3.4, as a horizontal grey line, the mean $\langle L_{\mathrm{IR,SF}} \rangle$ for each redshift range. Across all redshifts the data lie within a factor of 2 of the mean. However, we find that the $\langle L_{\mathrm{IR,SF}} \rangle$ values of the most luminous AGN for all of the redshift ranges at $z < 1.5$

are systematically above the overall mean. To quantify the deviation between the $\langle L_{IR,SF} \rangle$ of the high L_{2-8keV} bins to the rest of the data we make two fits; one to the two highest L_{2-8keV} bins (with the exception of $z = 0.2 – 0.5$ where we use only the highest L_{2-8keV} bin); and one to the rest of the luminosity bins in the same redshift range (see the grey dashed lines of Figure 3.4). We find an increase in $\langle L_{IR,SF} \rangle$ by a factor of $\sim$2 for the highest L_{2-8keV} when compared to the lower L_{2-8keV} bins in each of the redshift ranges with $z < 1.5$. For $z = 1.5 – 2.5$ there is no significant difference in $\langle L_{IR,SF} \rangle$ between the highest and lowest L_{2-8keV} that we cover. We note that the systematic increase of $\langle L_{IR,SF} \rangle$ at high L_{2-8keV} values observed in the redshift ranges of $z < 1.5$ does not correspond to a systematic increase of the redshifts at high L_{2-8keV} values (see Table 3.2). Thus the modest trends observed at the high L_{2-8keV} are not driven by redshift. We investigate the observed trends further in §3.4.3.

3.4.2 Comparing to the average SFRs of the overall star forming galaxy population

Here we explore whether X-ray AGN have SFRs that are consistent with being selected from the overall star forming galaxy population. We compare the average SFRs of the AGN to the observed relationship between SFR, redshift, and stellar mass (M_*) of normal star forming galaxies, which is defined as the “main sequence” of star forming galaxies (e.g., Noeske et al. 2007; Elbaz et al. 2011; Schreiber et al. 2015; Speagle et al. 2014). To make this comparison we require stellar masses for the AGN in our sample. We use the stellar masses from Ilbert et al. (2013) for the sources in the C-COSMOS area. Since their analysis did not take into account of a possible AGN component to the rest–frame UV to near–IR SEDs, we applied a colour cut to exclude sources for which there is likely to be significant AGN contamination to the SED. We only include AGN with rest frame colours $U - V > 1$ and $V - J > 1$ based on the analyses of Mullaney et al. (2012b). This results in a subsample of primarily moderate luminosity AGN ($L_{2-8keV} \lesssim 10^{44}$ erg s^{-1}) making up $\sim$40% of the parent sample at $z < 1.5$, but only 26% of the parent sample at $z = 1.5 – 2.5$. For these sources, with a reliable M_*, we calculate the $\langle L_{IR,SF} \rangle$ as described in §3.3.2. Due to the reduced number of sources with masses we can no longer use bins

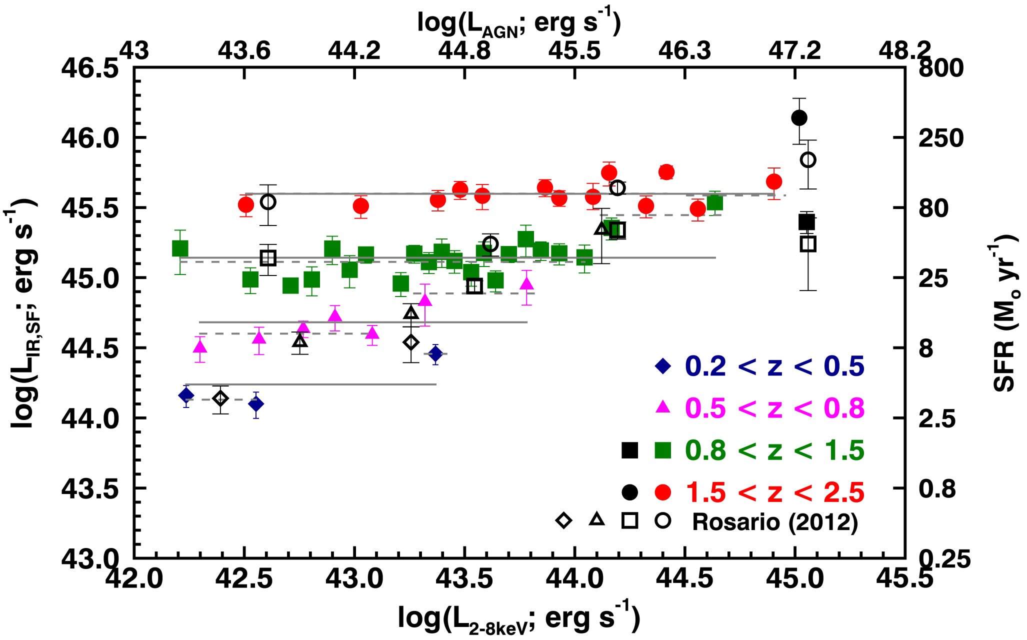


Figure 3.4: $\langle L_{IR,SF} \rangle$ as a function of $\langle L_{2-8keV} \rangle$, as plotted in Figure 3.3 (also to be referred to for axis definitions). The horizontal grey lines indicate the overall mean $L_{IR,SF}$ across all of the L_{2-8keV} bins for each redshift range. The dashed grey lines indicate the mean $L_{IR,SF}$ for (1) the one or two highest L_{2-8keV} bins and (2) the lower L_{2-8keV} bins for each redshift range (see §3.4.1). The black hollow symbols are the stacking results of Rosario et al. (2012), and the black filled symbols are bins of the highest L_{2-8keV} sources from our study (we note that there are very few sources in these bins for both studies; see §3.4.1). Our results are broadly consistent with a flat relationship; however, for the redshift ranges with $z < 1.5$ the highest L_{2-8keV} bins are systematically a factor of ~2 higher than the mean $L_{IR,SF}$.

of $\approx$40 sources and we therefore reduce the number of sources required in each bin to 25. We show the $\langle \mathrm{L_{IR,SF}} \rangle$ as a function of $\mathrm{L_{2-8keV}}$, for the sub-sample with reliable $\mathrm{M_*}$ values, in Figure 3.5. We note that this sub-sample have $\langle \mathrm{L_{IR,SF}} \rangle$ values consistent with the whole parent sample (see Figure 3.5), with the exception of the z = 1.5 – 2.5 range which appear to be systematically higher.

We use the mean redshift and mean $\mathrm{M_*}$ of each bin in Figure 3.5 to calculate the expected range in $\mathrm{L_{IR,SF}}$ for "main sequence" galaxies using Equation 9 of Schreiber et al. (2015). The shaded regions, colour-coded by redshift, correspond to the range of $\mathrm{L_{IR,SF}}$ covered by the main sequence galaxies at the mean redshift and mean $\mathrm{M_*}$ of the sources in each bin; i.e. a scatter of 2 around the mean results from Schreiber et al. (2015). We also find that these results are the same if we use the Elbaz et al. (2011) definition of the "main sequence". We find that, for this sample of X-ray AGN with $\mathrm{L_{2-8keV}} \lesssim 10^{44}$ erg s^{-1}, the $\langle \mathrm{L_{IR,SF}} \rangle$ in all redshift ranges with $z < 1.5$ are consistent with that of star forming galaxies of the same mean redshift and mass. This result agrees with the results of previous studies (e.g., Mullaney et al. 2012a, Harrison et al. 2012, Bongiorno et al. 2012, Lanzuisi et al. 2015). However, for the redshift range of z = 1.5 – 2.5 the $\langle \mathrm{L_{IR,SF}} \rangle$ is systematically at the higher end of the $\mathrm{L_{IR,SF}}$ region covered by "main sequence" galaxies, which may be due, in part, to a bias due to the fact that only 26% of the parent sample at those redshifts have reliable masses, and these have systematically higher $\langle \mathrm{L_{IR,SF}} \rangle$ values than the parent population (see Figure 3.5).

3.4.3 Comparing to empirical models

As shown in Figure 3.4, the trend of $\langle \mathrm{L_{IR,SF}} \rangle$ ($\langle$SFR$\rangle$) with $\mathrm{L_{2-8keV}}$ ($\mathrm{L_{AGN}}$) is broadly consistent with being flat. This result may initially seem in disagreement with the results of studies such as Rafferty et al. (2011), Mullaney et al. (2012b), Chen et al. (2013), Delvecchio et al. (2014), and Rodighiero et al. (2015), which find a correlation between the average $\mathrm{L_{AGN}}$ and SFR of star forming galaxies. However, these studies start with a parent population of *galaxies* for which they calculate the average $\mathrm{L_{AGN}}$, while in this study we start with a population of *AGN* for which we calculate the average SFR. It has been suggested that the variability of AGN, taking place on smaller timescales to that of star formation, could flatten any intrinsic correlation between the SFR and the $\mathrm{L_{AGN}}$ when

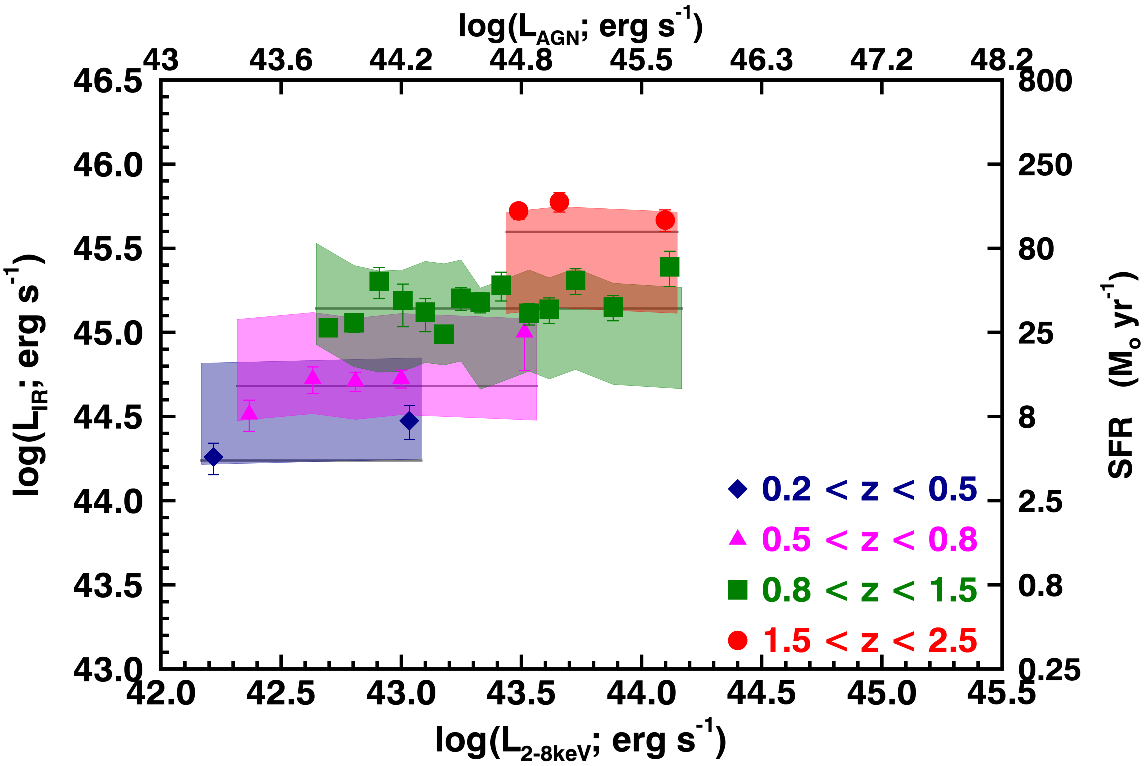


Figure 3.5: $\langle L_{IR,SF} \rangle$ as a function of $\langle L_{2-8keV} \rangle$ for the subsample of sources that have a reliable stellar mass (M_*) measurement in Ilbert et al. (2013) (see §1.3; also see Figure 3.3 for the axis definitions). The solid lines are the means for each redshift range of the whole parent sample (see Figure 3.4). The shaded regions correspond to the expected range in $L_{IR,SF}$ for the overall star forming galaxy population at the mean redshift and mean M_* of each bin as defined by Schreiber et al. (2015). For all redshift ranges the $\langle L_{IR,SF} \rangle$ values of the AGN appear to be consistent with normal star forming galaxies.

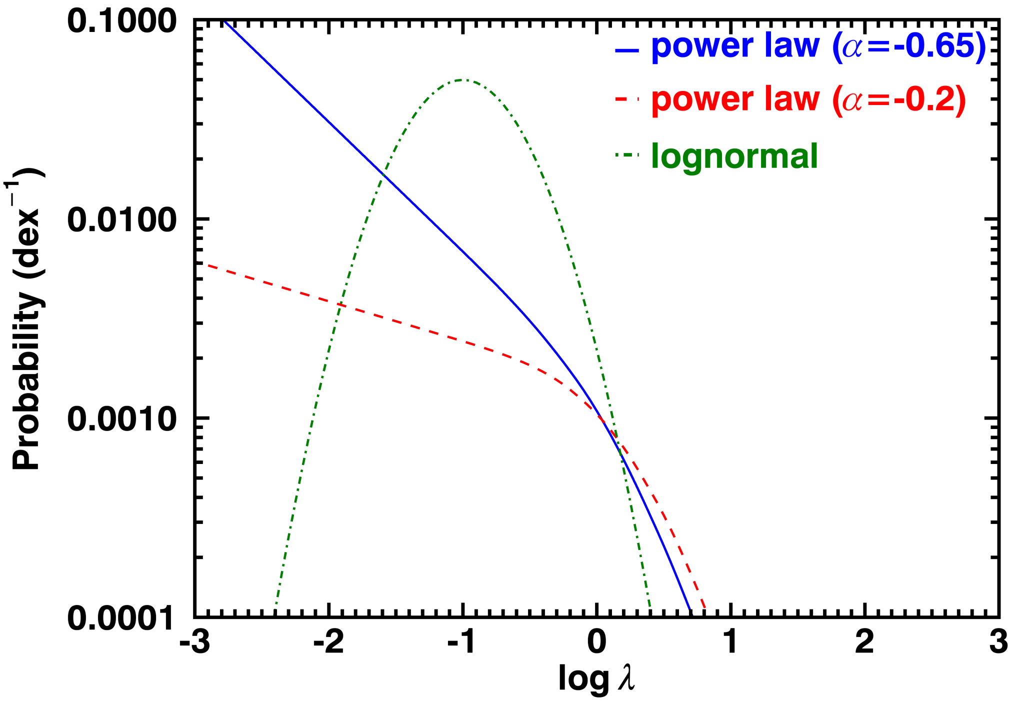
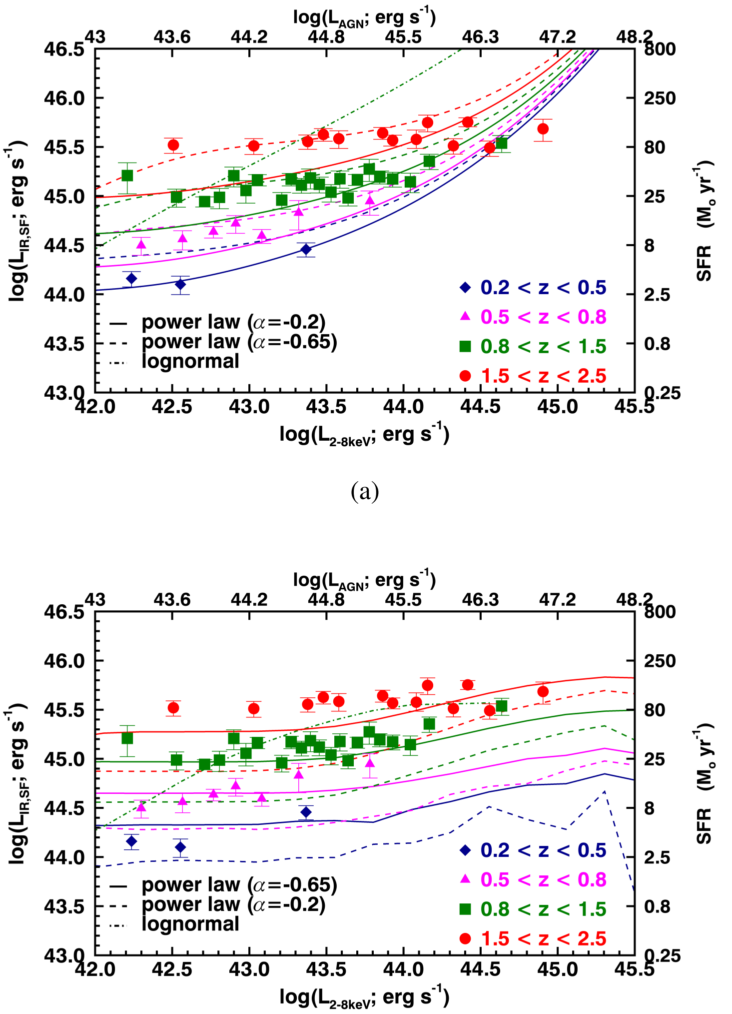


Figure 3.6: The probability distribution of the Eddington ratio (λ) for the three cases assumed in Figure 3.7 (i.e. two broken power law distributions with a faint end slope of $\alpha = -0.65$ and $\alpha = -0.2$, and a lognormal distribution with 0.4 dex dispersion; see §2.2 in Aird et al. 2013). This also serves as a *schematic representation* of the three distributions assumed for the Hickox et al. (2014) model, assuming that the shape of the distributions represent the variability function of individual AGN (see Section 3.4.3 and Section 2 of Hickox et al. 2014).

not averaging over the most variable quantity (i.e. by taking the average L_{AGN} over bins of SFR; e.g., Hickox et al. 2014). To assess what could be the driver of the flat relationship that we observe, and if indeed it is AGN variability that is driving its shape, we compare to two empirical "toy-models" that predict the $\langle L_{IR,SF} \rangle$ as a function of L_{AGN}. Firstly that of Hickox et al. (2014) and secondly, a model based on Aird et al. (2013) (also see Caplar et al. 2014 for a similar model).

The empirical "toy-model" presented in Hickox et al. (2014) assumes that SFR is correlated to L_{AGN} when averaged over timescales of the order of 100 Myr. To create the SFR distribution of the galaxies in their model, they assume the redshift dependent IR luminosity function (LF) from Gruppioni et al. (2013). In the model, the individual AGN



(a)

(b)

Figure 3.7: $\langle L_{\mathrm{IR,SF}}\rangle$ as a function of $\langle L_{2-8\mathrm{keV}}\rangle$ in four redshift ranges compared to model tracks from (a) Hickox et al. (2014) and (b) the extended Aird et al. (2013) model (see Figure 3.3 for the axis definitions). The solid lines in both cases show the predictions of the models with their originally assumed Eddington ratio distribution. From the two models the one of Aird et al. (2013), which is based solely on observational data, is in better agreement to our results; however both models demonstrate how the flatness of the trends in our results are likely to be a consequence of the assumed Eddington ratio distribution, or AGN variability. We also investigate how different the trends are when assuming different Eddington ratio distributions in the two models (i.e. two broken power-law distributions with a faint end slope of $\alpha = -0.65$ and $\alpha = -0.2$ respectively, and a lognormal distribution for which we only show the tracks for $0.8 < z < 1.5$ to avoid confusion; see Figure 3.6). The different assumed Eddington ratio distributions show significant differences in the predicted trends. See §3.4.3 for more details.

are allowed to vary on short time scales on the basis of an assumed $\mathrm{L_{AGN}}$ /$\langle \mathrm{L_{AGN}} \rangle$ distribution, which serves as a tracer of the Eddington ratio distribution of individual AGN in the absence of black hole masses (see details in Hickox et al. 2014). The fiducial model assumes that the distribution of $\mathrm{L_{AGN}}$ /$\langle \mathrm{L_{AGN}} \rangle$ has the form of a Schechter function (broken power law form) with a faint end slope of $\alpha = -0.2$ and a cutoff luminosity of $\mathrm{L_{cut}} = 100\ \langle \mathrm{L_{AGN}} \rangle$ (see the dashed red curve in Figure 3.6 for a schematic of this distribution). The model can then predict the average SFR as a function of instantaneous (i.e. observed) $\mathrm{L_{AGN}}$ of a large population of simulated AGN. We ran the model for the four redshift ranges of this study and plot the results in Figure 3.7(a) with solid tracks. The model successfully reproduces an increase of the $\langle \mathrm{L_{IR,SF}} \rangle$ with redshift, for a fixed range in $\mathrm{L_{AGN}}$, [5] and is in good agreement with the data at $z = 0.2$ – 0.5; however, it fails to reproduce the trends observed for the higher redshift ranges. In particular, the normalisation of the predicted trends are too low compared to our data and the rise of $\langle \mathrm{L_{IR,SF}} \rangle$ with $\mathrm{L_{AGN}}$ is much steeper than that observed. The steepness of the predicted $\langle \mathrm{L_{IR,SF}} \rangle$ trends at the highest $\mathrm{L_{AGN}}$ could be a result of the enforced correlation between SFR and the long term $\langle \mathrm{L_{AGN}} \rangle$, or could be caused by the lack of an explicit Eddington limit in the model but rather a cut-off limit at high $\mathrm{L_{AGN}}$/SFR ratios (see Hickox et al. 2014 for details). We investigate how the predicted relationship varies with different variability prescriptions later on in this section.

The second empirical "toy-model" that we have compared to is based on Aird et al. (2013), which we extended to make predictions for the relationship between AGN luminosity and star formation. This model uses the observed redshift dependent stellar mass function (SMF) of galaxies (from Moustakas et al., 2013) in combination with the probability function of a galaxy of a given stellar mass and redshift hosting an AGN, based on measurements in Aird et al. (2012) for $z \lesssim 1$. This model predicts the distribution of stellar masses, for which they correct to BH masses assuming $\mathrm{M_{BH}} = 0.002 \times \mathrm{M_*}$ based on Marconi & Hunt (2003), as a function of X-ray luminosity. In contrast to the Hickox

---

[5] The increase of the $\langle \mathrm{L_{IR,SF}} \rangle$ with redshift, for a fixed range in $\mathrm{L_{AGN}}$, could initially seem contradictory to the model's original assumption of a correlation of SFR and the long term averaged $\mathrm{L_{AGN}}$. However, even though the increase of $\langle \mathrm{L_{IR,SF}} \rangle$ will be accompanied by an increase in the long term averaged $\mathrm{L_{AGN}}$, there is not a significant difference in the range of *instantaneous* $\mathrm{L_{AGN}}$, across the simulated population, which is the quantity we effectively observe for an X-ray AGN sample.

et al. (2014) model they use an Eddington ratio distribution in the form of a broken power-law function with the faint end slope being steeper with $\alpha = -0.65$ (see the blue curve in Figure 3.6 for a schematic of this distribution). [6] In order to compare to our results we have extended the model to cover the same redshift range as that of our sample and convert the predictions of stellar mass to predictions of SFR. To achieve this we adopt the measurements of the SMF by Ilbert et al. (2013) at $z = 1 - 2.5$ as an extension of the Moustakas et al. (2013) SMF up to $z = 1$, and extrapolate the redshift-dependence of the probability of a galaxy hosting an AGN from Aird et al. (2012) to $z > 1$ (which is consistent with the $z > 1$ measurements from Bongiorno et al. 2012). Furthermore, we make the assumption that all of the AGN are hosted by normal star forming galaxies that lie on the "main sequence" as derived by Elbaz et al. (2011), which is motivated by the results of our study (see §3.4.2). [7] We convert from the model predicted stellar masses to SFRs, allowing for a scatter of 0.3 dex in SFR around the "main sequence" relation. In Figure 3.7(b) we present the resulting predictions of $\langle \mathrm{L_{IR,SF}} \rangle$ as a function of $\mathrm{L_{AGN}}$, plotted with solid lines, in comparison to our results for each of the four redshift ranges. The predicted trends of the mean SFR in this case are flat for a wide range of $\mathrm{L_{AGN}}$, similar to our data, with a slight rise in $\langle \mathrm{L_{IR,SF}} \rangle$ at high $\mathrm{L_{AGN}}$ (i.e., $\mathrm{L_{AGN}} \gtrsim 10^{45}$ erg s^{-1}). On the basis of this modified Aird et al. (2013) model, the slight rise of $\langle \mathrm{L_{IR,SF}} \rangle$ observed in our data (see §3.4.1) may be driven by a small increase in the average masses of the galaxies hosting very luminous AGN.

To first order, the data are better described by the extended Aird et al. (2013) model than the Hickox et al. (2014) model; see the solid tracks in Figure 3.7(b) compared to those in Figure 3.7(a). However, since the two models have assumed different Eddington ratio distributions (or, equivalently, $\mathrm{L_{AGN}} / \langle \mathrm{L_{AGN}} \rangle$ for the Hickox et al. 2014 model) we also explore how sensitive the results are to this assumption. We therefore also ran the models with a series of three different Eddington ratio distributions to understand how sensitive the predicted trends of $\langle \mathrm{L_{IR,SF}} \rangle$ with $\langle \mathrm{L_{AGN}} \rangle$ are on the assumed Eddington ratio

---

[6] We note that Aird et al. (2013) use an observed specific accretion rate distribution (i.e., $\mathrm{L_{AGN}} / \mathrm{M_*}$) which they convert to an Eddington ratio distribution.

[7] We note that there is evidence in optical studies of X-ray AGN, such as Azadi et al. (2015), that a small fraction of these AGN are hosted by non star forming galaxies; however, Azadi et al. (2015) find that these AGN appear to form a minority of the population and therefore we do not expect them to significantly affect our mean SFRs.

distribution. We used (1) a broken power-law with $\alpha = -0.2$ (i.e, the fiducial distribution assumed by Hickox et al. 2014); (2) a broken power-law with $\alpha = -0.65$ (i.e, the fiducial distribution assumed by Aird et al. 2013); and (3) a narrow lognormal distribution with a dispersion of $\sim$0.4 dex centred at an Eddington ratio of ~ 0.06, as defined by Kauffmann & Heckman (2009) for nearby AGN residing in star forming galaxies. These three distributions can be seen in Figure 3.6.

In Figure 3.7(a)&(b) we show the three sets of tracks which correspond to the resulting trends of $\langle \mathrm{L_{IR,SF}} \rangle$ with $\langle \mathrm{L_{AGN}} \rangle$ for the different assumptions of the Eddington ratio distributions. A clear correlation between $\langle \mathrm{L_{IR,SF}} \rangle$ and $\langle \mathrm{L_{AGN}} \rangle$ is predicted for the lognormal distribution while, by comparison, the power-law models predict a much flatter relationship. With a change of power-law slope from $\alpha = -0.2$ to $\alpha = -0.65$, the normalisation of the model tracks increase and the trend becomes flatter. The different shapes of the model tracks are driven by the relative difference between the low Eddington ratio slope and the slope of the low-mass end of the galaxy SMF (i.e., for $M < M_{*}$, $\alpha \sim 0$). The predicted correlation between $\langle \mathrm{L_{IR,SF}} \rangle$ and $\langle \mathrm{L_{AGN}} \rangle$ for the lognormal distribution is due to the narrow range of probable Eddington ratios. For the assumptions behind our models when assuming the lognormal distribution, most of the AGN are accreting at a broadly similar Eddington ratio and therefore an increase in $\mathrm{L_{AGN}}$ is predominantly due to an increase in stellar mass (and hence SFR since we assume the main sequence of star-forming galaxies). By contrast, the steep low-Eddington ratio slope for the power-law models, when compared to the low-mass end slope of the galaxy SMF, allows for a broad range of Eddington ratios across a narrow range in stellar mass; i.e., there is a higher probability for an AGN of a given luminosity to be hosted in a high-mass galaxy with a low Eddington ratio than a low-mass galaxy with a high Eddington ratio. Indeed, on the basis of the extended Aird et al. (2013) model, the population of low-to-moderate luminosity AGN ($\mathrm{L_{AGN}} \lesssim 10^{45}$ erg s^{-1}) predominantly reside in galaxies of similar stellar mass ($M_{*} \sim 10^{10.5-11} M_{\odot}$), and thus similar SFRs, but with a wide range of possible Eddington ratios.

Overall, our results suggest that the observed trends of $\langle \mathrm{SFR} \rangle$ – $\mathrm{L_{AGN}}$ are due to AGN being highly variable and residing, on average, in normal star forming galaxies. Similar results have also been found by hydrodynamical simulations that show that AGN variability can cause a flat trend between L_{AGN} and SFR (e.g., Gabor & Bournaud 2013; Volonteri et al. 2015). The Eddington ratio distributions of AGN are typically constructed to describe a population of AGN. However, as adopted in our models, they can also be understood as the distribution of Eddington ratios for an individual AGN over time, and hence could be used as a variability prescription of the AGN (as originally adopted in Hickox et al. 2014). As is clear from Figure 3.7(a)&(b), the choice of Eddington ratio distribution plays a major role in the form of the predicted $\langle \mathrm{SFR} \rangle$ – L_{AGN} relationship. For example, our results are much better described with the use of a broken power-law Eddington ratio distribution with a faint end slope of $\alpha = -0.65$, than with a narrow lognormal Eddington ratio distribution, which predicts a qualitatively different $\langle \mathrm{SFR} \rangle$ – L_{AGN} relationship to that found from our data. Thus, the $\langle \mathrm{SFR} \rangle$ (or $\langle L_{IR,SF} \rangle$) – L_{AGN} plane can be a useful diagnostic tool for placing constraints on the intrinsic Eddington ratio distribution of AGN (also see Veale et al. 2014).

3.5 Conclusions

We have created a large sample of X-ray detected AGN with FIR coverage and individual SFR measurements. Our sample has a total of 2139 AGN at redshifts of $z = 0.2$–2.5, with $10^{42} < L_{2-8keV} < 10^{45.5}$ erg s^{-1}. Using the available photometry from 8–500μm we have performed individual SED fitting to all of the sources in our sample, and measure the IR luminosity due to star formation, $L_{IR,SF}$.

Our analysis has a number of key advantages over many previous studies: (a) the use of deblended source catalogues for the FIR photometry, which ensures better constraints on the flux density measurements and eliminates the overestimation due to blending and confusion of sources (see §3.2.2); (b) the use of photometric upper limits in the SED fitting analysis, which achieve better constraints on the fitted SEDs (see §3.3.1); (c) the decomposition of the AGN and star formation contributions to the FIR emission, which provides values of $L_{IR,SF}$ that are not contaminated by the AGN (see §3.3.1); (d) the calculation of upper limits on $L_{IR,SF}$ when the data are insufficient to identify the star forming component directly (i.e., not enough photometric data points, poor S/N data, or dominant AGN component), which allows us to estimate the $\langle L_{IR,SF} \rangle$ for all the sources in our sample avoiding the bias that could be caused by removing these sources (see §3.3.2).

With the $L_{IR,SF}$ measurements for each source we derived the mean $L_{IR,SF}$ values ($\langle L_{IR,SF}\rangle$; a proxy of the $\langle SFR\rangle$) as a function of L_{2-8keV} (a proxy of the AGN luminosity; L_{AGN}) in bins of ∼40 sources, for the redshift ranges of 0.2 – 0.5, 0.5 – 0.8, 0.8 – 1.5, and 1.5 – 2.5. In comparison to previous studies, our results show less scatter amongst $\langle L_{IR,SF}\rangle$ across the wide range of L_{2-8keV} investigated in this study. Overall we found that:

1. The $\langle SFR\rangle$ increases by more than an order of magnitude from redshifts of 0.2 – 0.5 to 1.5 – 2.5, in agreement with previous studies on the redshift evolution of the SFR for the general star forming galaxy population. See §3.4.1.

2. For each redshift range the $\langle SFR\rangle$ shows no strong dependence on AGN luminosity; however we note that for the redshift ranges of $z \leq 1.5$ the highest L_{AGN} systems have $\langle SFR\rangle$ values that are systematically higher than those of lower L_{AGN} systems by a factor of ≈2. See §3.4.1.

3. For the ∼40% of the sources within the COSMOS area with reliable stellar masses, we compare their $\langle SFR\rangle$ to the "Main Sequence" of the overall star forming galaxy population. The X-ray AGN, at all redshift ranges, have $\langle SFR\rangle$ that are consistent with normal star forming galaxies at the same redshifts and masses. Due to a lack of secure masses for the high L_{AGN} systems in our sample this result is restricted to moderate AGN luminosities (i.e., $L_{2-8keV}\lesssim 10^{44.2}$ erg s^{-1} or $L_{AGN}\lesssim 10^{45.5}$ erg s^{-1}). See §1.3.

4. To qualitatively understand the flat relationship between the $\langle SFR\rangle$ and L_{AGN} we compared to two empirical "toy-models" that make predictions for this relation: Hickox et al. (2014) and an extended version of Aird et al. (2013). These models take mock galaxy populations and assign them with SFR values based on observed distributions, and instantaneous L_{AGN} values based on an assumed Eddington ratio distribution. We find that the flat relationship seen in our data could be due to short timescale variations in the mass accretion rates, which, in combination with the relative shapes of the Eddington ratio distribution and the galaxy SMF, can wash out the long term relationship between $\langle SFR\rangle$ and L_{AGN}. See §3.4.3

5. We find that the predicted ⟨SFR⟩ – L_{AGN} relationship is sensitive to the assumed Eddington ratio distribution. For example, both models predict a relatively flat relationship over all redshift ranges, assuming an Eddington ratio distribution of a broken power-law form with a faint end slope of $\alpha = -0.65$, whilst with a log-normal distribution the predicted trends are too steep to be consistent with our data. Therefore, the observed ⟨SFR⟩ – L_{AGN} relationship appears to be a sensitive diagnostic of the intrinsic Eddington ratio distribution of AGN. See §3.4.3.

3.6 APPENDIX:

Comparison of the K-M method to the stacking analysis method

In this Appendix we compare our results using our SED fitting analysis and the K-M method that we used in this work (see Section 3.3), to those we would obtain with stacking analysis, a method commonly used in similar studies of star-forming and AGN galaxy samples.

Following the method of Harrison et al. (2012) we stacked the SPIRE-250μm maps at the X-ray positions of the sources of our sample in C-COSMOS. We use the C-COSMOS sample since it makes up most of our overall sample and avoids issues that can arise when combining stacks of different fields with different depths.

We bin the sample in bins of $L_{2-8\mathrm{keV}}$ and redshift containing ~40 sources each, in the same way as described in §3.3.2 for the K-M method (in the redshift range of $z = 0.2$ – 0.5 we use ~30 sources to allow for more than one bin). We show the stacking results in Figure 3.8, in comparison with the overall means of the K-M method results for each redshift range, as well as the results of Rosario et al. (2012). We find that our main results are consistent with the results we obtain when using the stacking analysis, and that both methods are in agreement with the results of Rosario et al. (2012).

This comparison demonstrates that our method for calculating the mean produces results consistent with the popular method of stacking in the FIR. However, our method produces less scatter amongst bins, as well as smaller uncertainties on the mean values. This is likely due to the use of deblended FIR photometry, and the removal of AGN contamination, in our analysis, which are effects not taken into when stacking.

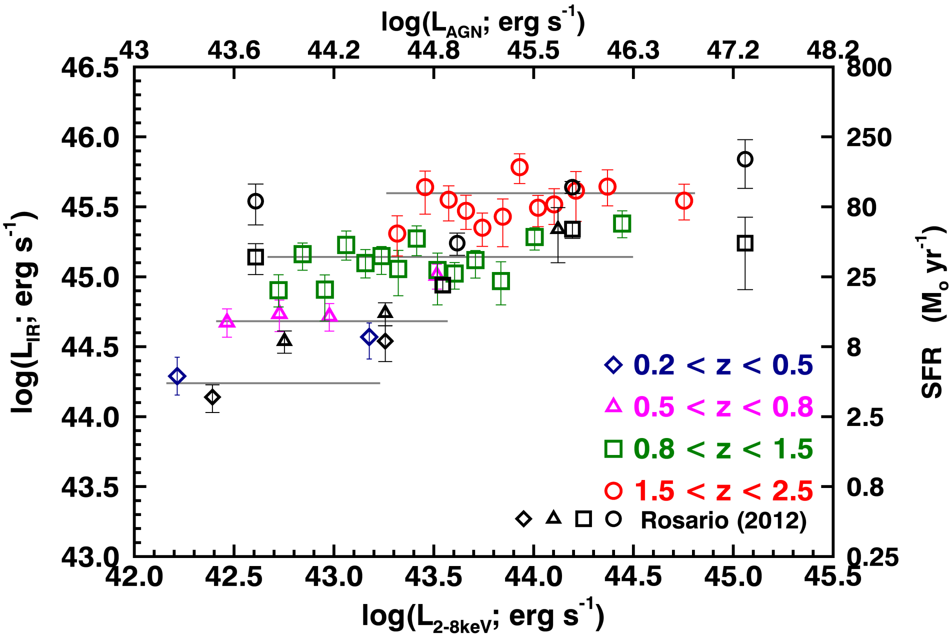


Figure 3.8: $\langle L_{IR,SF} \rangle$ as a function of $\langle L_{2-8keV} \rangle$ when stacking the SPIRE data at 250μm for the sources of C-COSMOS see Figure 3.3 for the axis definitions. We compare these results to the overall K-M means of our SED results (grey lines; see §3.4.1), and the results of Rosario et al. (2012). We find that our results are consistent with those obtained using the stacking analysis, however the K-M method's results produce less scatter (see Figure 3.4).

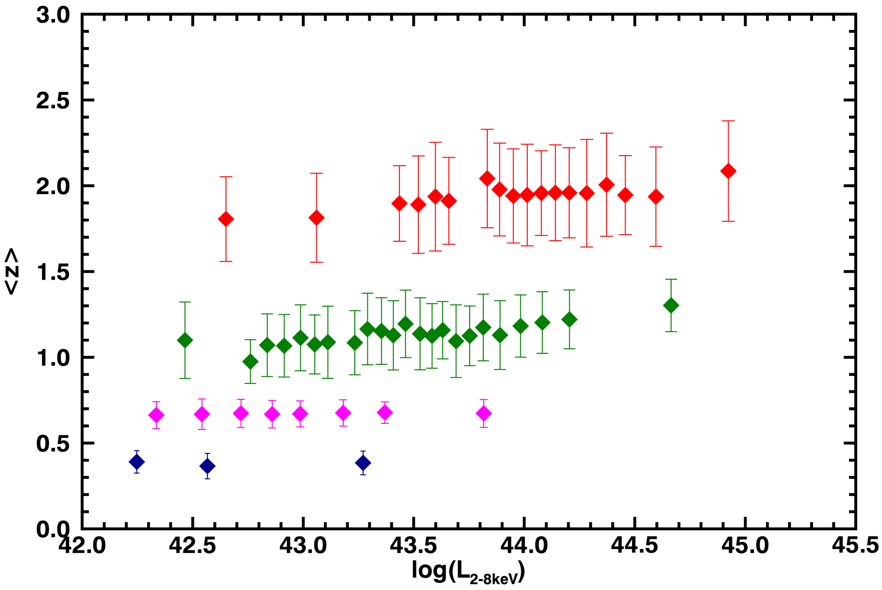


Figure 3.9: Average redshift in each of the $L_{2-8\mathrm{keV}} - \mathrm{z}$ bins of Figure 3.3. The data points are colour-coded for different redshift ranges, with blue corresponding to $0.2 < z < 0.5$, magenta to $0.5 < z < 0.8$, green to $0.8 < z < 1.5$, and red to $1.5 < z < 2.5$. The average bin redshifts are consistent within each redshift range.

The mean redshift of the $L_{2-8\mathrm{keV}}$ –z bins as a function of $L_{2-8\mathrm{keV}}$

For each $L_{2-8\mathrm{keV}}$ –z bin in our sample we calculate the average redshifts to investigate if the slight rise of average SFR at high $L_{2-8\mathrm{keV}}$ is driven by redshift dependencies. We find that the average redshifts are consistent with being constant throughout each redshift range, confirming that the slight rise of SFR at the high $L_{2-8\mathrm{keV}}$ bins is not due to a redshift increase.

AGN luminosities in the galaxies of our sample with the highest SFRs

We explore whether the most powerful star forming galaxies of our sample are hosting luminous AGN ($L_{2-8\mathrm{keV}} > 10^{44}\,\mathrm{erg\,s^{-1}}$ or if they only host AGN of low to moderate luminosities ($L_{2-8\mathrm{keV}} = 10^{42} - 10^{44}\,\mathrm{erg\,s^{-1}}$). Studies such as Alexander et al. (2005) and Page et al. (2012) have found that luminous FIR/Sub-mm sources preferentially host moderate luminosity AGN and find few luminous FIR sources with high luminosity AGN. These

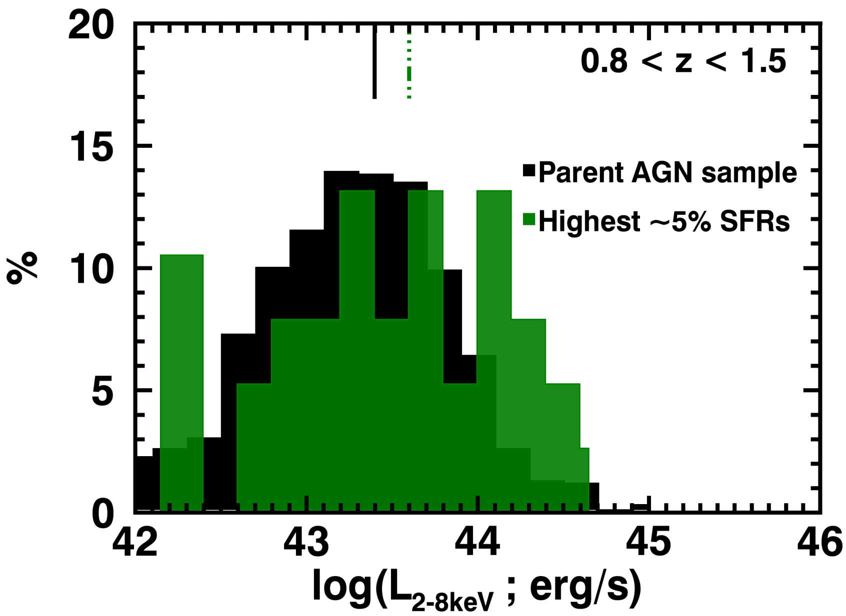


Figure 3.10: X-ray luminosity distribution of the AGN sample at 0.8 $<$ z $<$ 1.5 (Black) and the top ∼5% of sources with the highest SFRs (Green) at the same redshift. The medians are shown with a black solid and a green dashed line respectively. We find that for $L_{2-8\mathrm{keV}}>10^{44}$ erg s^{-1} the fraction of powerful SF galaxies is 3.7 times higher than for $L_{2-8\mathrm{keV}}<10^{44}$ erg s^{-1}, indicating that the most powerful SF galaxies preferentially host luminous AGN.

results have been interpreted by Page et al. (2012) as evidence for suppression of star formation by luminous AGN, a result in apparent disagreement to the results presented in this work where we find that the average SFRs are increased by a factor of 2 at the high $L_{2-8\mathrm{keV}}$ compared to the lower.

To understand this difference of results between the two methods we chose the highest $\sim 5\%$ of the $L_{\mathrm{IR,SF}}$ values in the redshift range of $z = 0.8 - 1.5$ (38 out of 918 sources) where we are complete in high SFRs (i.e. there are no upper limits). With this selection we create a sub-sample of sources of extreme star forming galaxies hosting an X-ray detected AGN. We then compare the distribution of the X-ray luminosity of this sub-sample to that of the parent AGN sample within this redshift range.

As shown in Figure 3.10, the distribution of highly star forming AGN galaxies is wide and covers the same region of $L_{2-8\mathrm{keV}}$ as the parent sample. The fraction of AGN with the highest 5% of SFRs increases by a factor of 3.7 at the luminosity range of $L_{2-8\mathrm{keV}} =$

10^{44-45} erg s^{-1} compared to $L_{2-8\mathrm{keV}}<10^{44}$ erg s^{-1}. This result is not consistent with the deficit of luminous AGN in highly star forming galaxies that has been observed by previous studies, but coincides with the rise of $L_{\mathrm{IR,SF}}$ seen at the highest $L_{2-8\mathrm{keV}}$ bin of this redshift range in Figure 3.3 and could be a result of mass dependencies. When performing a two sample K-S test we find that the two $L_{2-8\mathrm{keV}}$ distributions have a probability of 2.8% of being the same, i.e. they are different at a 2σ level.

CHAPTER 4

The mean SFRs of luminous unobscured QSOs: searching for evidence of suppressed or enhanced star formation

Abstract

We use *Herschel* and WISE photometry to investigate the mean star formation rates (SFRs) in the hosts of optically selected QSOs from the SDSS survey within the *Herschel*–ATLAS fields, and a radio-luminous sub-sample defined using FIRST. Our sample includes more than 3000 sources covering a wide range of redshift (z =0.2–2.5), and with high bolometric AGN luminosities of $\mathrm{L_{AGN}}$=10^{45}–10^{48} erg $\mathrm{s^{-1}}$. We stacked PACS and SPIRE photometry of the sources in our sample in redshift and $\mathrm{L_{AGN}}$ bins. We use the stacked PACS and SPIRE fluxes in combination with mean WISE fluxes for each bin, and perform SED fitting to estimate the mean IR luminosity due to star formation after removing the contamination from the AGN. We find that the mean SFRs show a weak positive trend with AGN luminosity, which is likely to be driven by the stellar masses of the galaxies. We find that the mean SFRs of our QSO sample are consistent with the main sequence of star-forming galaxies. Furthermore, we compare to a sample of X-ray AGN and find that the mean SFRs of our QSO sample are consistent with those of X-ray AGN. When investigating the mean SFRs of the radio-luminous QSOs in our sample we find that their mean SFRs show a flat trend with radio luminosity, and are consistent with living in normal star-forming galaxies on the main sequence.

4.1 Introduction

In this chapter we aim to investigate the mean SFRs of the most luminous AGN, with $L_{\rm AGN} > 10^{45}$ erg s^{-1}, building on the results of Chapter 3 for galaxies at intermediate to high redshifts, $0.2 < z < 2.5$.

As discussed in Chapter 3, FIR studies of X-ray selected AGN that reach moderate to high AGN luminosities ($L_{\rm AGN} < 10^{45-46}$ erg s^{-1}) find that the mean SFRs as a function of AGN luminosity show flat trends independently of redshift, up to a redshift of $z \sim 3$ (e.g. Mullaney et al. 2012a, Harrison et al. 2012, Rosario et al. 2012, Azadi et al. 2015; Stanley et al. 2015; see also Chapter 1). In Stanley et al. (2015) (Chapter 3) we demonstrated how the flat trends can be reproduced by empirical models that assume a connection between the AGN activity and star formation (Aird et al. 2013; Hickox et al. 2014). The models assume that the AGN activity is a stochastic process, with the probability of an AGN at a given luminosity defined by the observed Eddington ratio distribution (e.g., Aird et al. 2012), and the SFR is defined by the stellar mass distribution and the main sequence of star forming galaxies (e.g., Schreiber et al. 2015), based on which they can estimate the mean SFR trends as a function of AGN luminosity. From this comparison it became apparent that the primary driver of the flat trends is the fact that a wide range of stellar mass galaxies (and hence a wide range in SFRs) can host AGN of moderate luminosities ($L_{\rm AGN} < 10^{45}$ erg s^{-1}), and when averaged this large scatter will result in similar mean values as a function of AGN luminosity.

However, when moving to the highest AGN luminosities of $L_{\rm AGN} > 10^{45}$ erg s^{-1}, the range in the stellar masses of galaxies able to host such high accretion rates narrows. Large accretion rates require large amounts of available gas and so will tend to live in more massive galaxies. Consequently, one might expect that if there is an impact of the AGN energy output on the SFR of its host galaxy (see section 1.4) it will be traceable at the most powerful AGN, such as Quasi-Stellar Objects (QSOs) with $L_{\rm AGN} > 10^{45}$ erg s^{-1}, and luminous radio-AGN with $L_{\rm 1.4GHz} > 10^{24}$ W Hz^{-1}.

FIR studies of optically selected QSOs at $z \gtrsim 0.2$ are finding that they tend to live in galaxies with ongoing star formation (e.g. Kalfountzou et al. 2014, Netzer et al. 2015, Gürkan et al. 2015) at levels consistent with those of star-forming galaxies (e.g., Rosario et al. 2013b). When looking at the mean SFR as a function of the bolometric AGN luminosity from the optical some studies argue for a positive correlation between them (e.g. Rosario et al. 2013b; Kalfountzou et al. 2014; Gürkan et al. 2015). However, the positive trend of Rosario et al. (2013b) is attributed to the mass dependencies of SFR (see below).

When looking at the radio-AGN population, FIR studies find that at $z \gtrsim 0.2$, their hosts have ongoing star formation (e.g. Karouzos et al. 2014; Kalfountzou et al. 2014; Gürkan et al. 2015; Magliocchetti et al. 2014), with evidence of enhanced star formation at redshifts of $z \gtrsim 1$ (e.g. Magliocchetti et al. 2016). Radio-AGN are also showing evidence of a positive trend of mean SFRs with both radio AGN luminosity (e.g., Karouzos et al. 2014), and optically derived AGN bolometric luminosity (Kalfountzou et al. 2014; Gürkan et al. 2015). Furthermore, there is evidence that radio-AGN can show significant differences in their SFR levels if split in the two populations of low-excitation and high-excitation radio galaxies (LERGs and HERGs respectively; Best & Heckman 2012 and references therein). Although both seem to follow the positive trends observed for the full sample, LERGs tend to show lower SFRs to HERGs (e.g. Hardcastle et al. 2013; Gürkan et al. 2015).

A key limitation of the majority of previous studies is that they have not simultaneously taken into account the observed stellar mass and redshift dependencies of SFR in the global galaxy population. The global galaxy SFR average increases with increasing redshift up to redshift of ∼2–3 where we observe the peak of star formation. Furthermore, at redshifts of $z \gtrsim 0.2$ there is a well known stellar mass dependency of the SFR, for actively star-forming systems, which is called the main sequence of star-forming galaxies (e.g., Noeske et al. 2007; Elbaz et al. 2007; Whitaker et al. 2012; Schreiber et al. 2015). These effects could be driving the observed correlations of the SFR with AGN luminosity, and need to be simultaneously taken into account when investigating such trends. An additional source of uncertainty in studies on the SFRs of galaxies hosting AGN activity, is the fact that such powerful AGN could be contributing significantly to the FIR luminosities observed (e.g. Drouart et al. 2014), something that has repeatedly not been taken into account from previous studies. Not removing the potential AGN contamination to the FIR photometry used to derive SFRs can cause an artificial boost in SFR values.

In this work, we aim to define the mean SFRs of more than 3000 optically selected

QSOs at 10^{45} <L_{AGN}< 10^{48} erg s^{-1}, and a sub-sample of 258 radio-luminous QSOs of $L_{1.4GHz}$>10^{24} W Hz^{-1}, over the redshift range of 0.2< z <2.5. We will compare our results to the normal star-forming galaxies of the same epoch, and expand the $\langle SFR \rangle$ – L_{AGN} plane of Stanley et al. (2015) (see Chapter 3) to higher AGN luminosities. In our analysis we will simultaneously take into account of both redshift, and stellar mass dependencies, and remove AGN contamination from the IR luminosity.

4.2 Sample & Data used

The aim of this work is to constrain the mean SFRs as a function of AGN bolometric luminosity, reaching very high luminosities ($L_{AGN}\sim 10^{48}$ erg s^{-1}; see Fig. 4.1), as well as investigating dependencies of the mean SFRs on the presence of a radio-luminous AGN.

Far-IR (FIR) photometry is one of the best tracers of star formation, as it traces the peak of the dust-reprocessed emission from star-forming regions (e.g. Kennicutt 1998; Calzetti et al. 2010; Domínguez Sánchez et al. 2014b). We use FIR data from the *Herschel*-ATLAS observational program (H-ATLAS; Eales et al. 2010a; section 4.2.2) that covered the fields of GAMA09, GAMA12, and GAMMA15 in its Phase 1, and the north and south galactic poles (NGP, and SGP respectively) in its Phase 2 observations. The *Herschel*-ATLAS fields benefit from multi-wavelength coverage, with excellent optical (SDSS; section 4.2.1), MIR and FIR photometry (WISE and *Herschel*; section 4.2.2), and radio observations (FIRST; section 4.2.3). We use the available data to draw a sample of optically selected QSOs from the SDSS survey, determine a radio-luminous sub-sample of QSOs using the FIRST survey, and define their SFRs using the WISE and *Herschel* observations. As we only study the fields that have overlap with the SDSS survey area, we exclude the SGP field.

4.2.1 Optical/SDSS QSOs

To define our QSO sample we use the publicly available SDSS data release 7 (DR7) QSO catalogue as presented in Shen et al. (2011) (see also Schneider et al. 2010 for original selection of QSOs). To provide a measurement of the power of the QSOs we use the AGN bolometric luminosity as given in this catalogue. This luminosity has been derived from

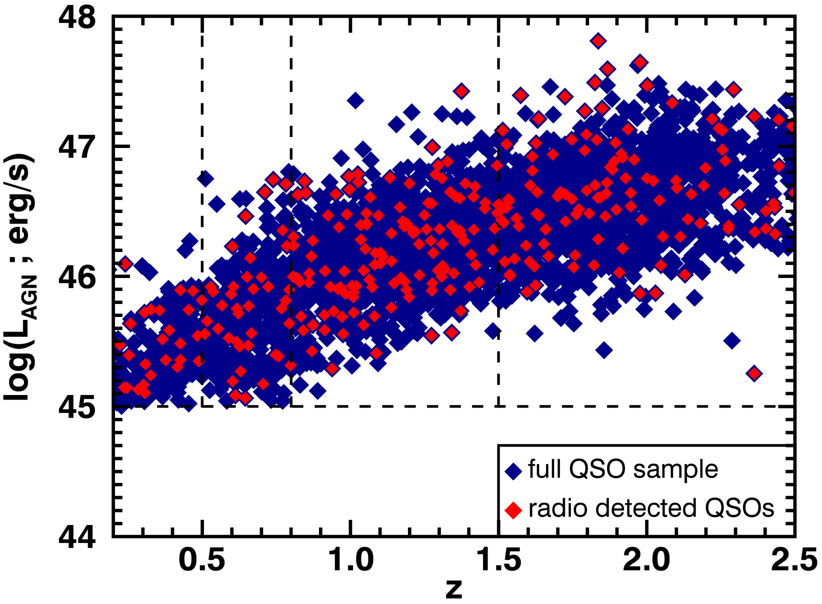


Figure 4.1: AGN bolometric luminosity (L_{AGN}) versus redshift (z) for the full QSO sample from SDSS DR7 covered by H-ATLAS in the NGP, GAMA9, GAMA12, and GAMA15 fields. The vertical dashed lines indicate the redshift ranges taken in our analysis, and the horizontal dashed line shows the L_{AGN} cut that defines the sample (see section 4.2.1). In red we highlight the radio detected sources from the FIRST radio catalogue (see section 4.2.3). Within the redshift range of interest (z =0.2–2.5) there are a total of 3026 optically selected QSOs.

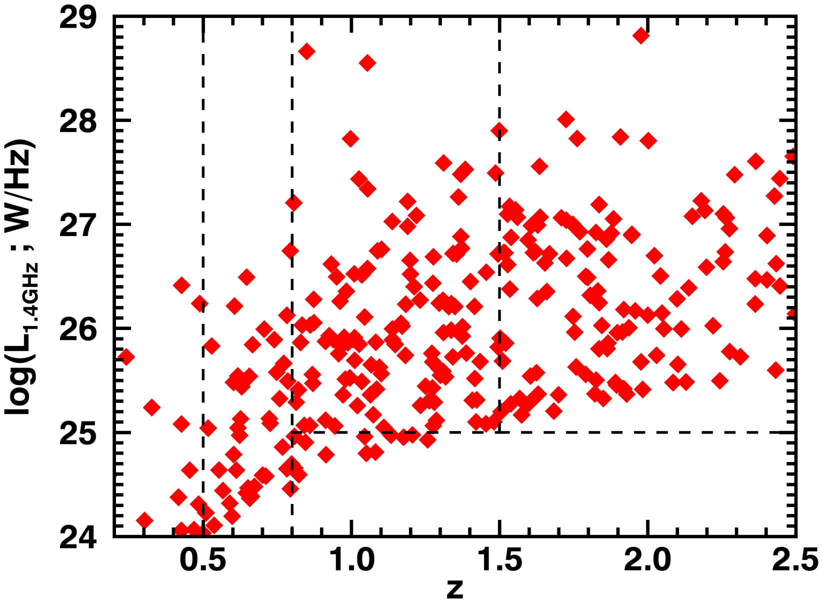


Figure 4.2: Radio luminosity from the FIRST survey ($L_{1.4GHz}$) versus redshift (z), for the full QSO sample from SDSS DR7 covered by H-ATLAS in the NGP, GAMA9, GAMA12, and GAMA15 fields that is radio detected. The vertical dashed lines indicate the redshift ranges taken in our analysis, and the horizontal dashed lines show the $L_{1.4GHz}$ limits used to define sources as radio-luminous. A total of 258 are classified as radio-luminous within the redshift range of interest (z =0.2–2.5; see section 4.2.3).

L_{5100A}, L_{3000A}, and L_{1350A}, for sources at redshifts of z <0.7, 0.7$\leq z$ <1.9, and $z \geq$1.9 respectively, using the spectral fits and bolometric corrections from the composite SED in Richards et al. (2006) ($BC_{5100A} = 9.26$, $BC_{3000A} = 5.15$ and $BC_{1350A} = 3.81$; see Shen et al. 2011). We constrain the sample of QSOs within the regions covered by H-ATLAS, removing sources that are close to the image boundaries. All the QSOs of our sample have a bolometric luminosities of $L_{AGN} \gtrsim 10^{45}$ erg s^{-1} (see Fig. 1).

We also make use of the virial BH masses (M_{BH}) estimates from Shen et al. (2011), from which we estimate the stellar masses (see section 4.4.2 and Eg. 4.4.4). The M_{BH} have been calculated using the FWHM of H_β, Mg_{II}, and C_{IV} lines (see section 3 of Shen et al. 2011). Specifically, the M_{BH} is estimated from H_β for sources with redshifts of z <0.7, from Mg_{II} for sources with 0.7$< z \leq$1.9, and from C_{IV} for sources with z >1.9.

This study looks at sources with redshifts z = 0.2–2.5, and includes a total of 3026 QSOs, with BH masses ranging within $10^7 < M_{BH} < 10^{11} M_\odot$.

4.2.2 Mid-infrared and Far-infrared photometry

For our analysis we make use of the psf-smoothed and background subtracted PACS and SPIRE image products provided by the H-ATLAS team (Valiante et al. in prep) for the four fields of GAMA09 (54 deg^2), GAMA12 (54 deg^2), GAMA15 (54 deg^2), and NGP (150 deg^2) that overlap with the SDSS survey. Detailed information on the construction of the images is presented in Valiante et al. (in prep). The images used in our analysis have had the large scale background subtracted (i.e., the cirrus emission), and each pixel contains the best estimate of the flux density of a point source at that position, making them ideal for stacking analysis. In addition to the images there are also noise maps available that provide the sum of the instrumental and confusion noise at each position (details in Valiante et al. in prep).

To define the MIR properties of our sample we use the WISE all-sky survey (Wright et al. 2010; catalogue available at: http://vizier.u-strasbg.fr/viz-bin/VizieR-3?-source=II/311/wise). Using a radius of 1 arcsec we match to the optical positions of our QSO sample described in §4.2.1, with a spurious match fraction of ∼0.4%. We find that more than 90% of our sources have a WISE counterpart. For our analysis we use the W3 and W4 bands at 12μm and 22μm respectively.

4.2.3 Radio data and classification

To determine the radio luminosities of our QSO sample we use the FIRST radio catalogue (Becker et al. 1995) that covers the full sky area observed by SDSS, to a sensitivity of 1mJy. To define our radio luminous QSO sub-sample we matched the SDSS QSO catalogue to the FIRST catalogue using a 2" radius, to minimise the number of spurious matches, with a resulting spurious match fraction of ∼2.5%. We calculate the 1.4GHz luminosity ($L_{1.4GHz}$) from the catalogued flux densities, using the following equation:

$$L_{1.4GHz} = 4\pi D^2 F_{1.4GHz}(1+z)^{-(1-\alpha)} \tag{4.2.1}$$

where D is the luminosity distance, $F_{1.4GHZ}$ is the catalogued flux density, and assuming $f_\nu \propto \nu^{-\alpha}$ with a spectral index of α =0.8. In Figure 2 we plot the radio luminosity of the detected sources as a function of redshift.

We classify sources as radio-luminous AGN, using a luminosity lower limit cut of $L_{1.4GHz} > 10^{24}$W Hz^{-1} for z <0.8, and $L_{1.4GHz} > 10^{25}$W Hz^{-1} for z >0.8 (see Fig. 2). Based on work from McAlpine et al. (2013), Magliocchetti et al. (2016) argue that the radio luminosity beyond which the radio emission is dominated by the AGN evolves with redshift up to a redshift of $z \sim 1.8$, after which it remains constant at $L_{1.4GHz,limit} = 10^{23.5}$W Hz^{-1}. Our luminosity cut is always higher in comparison, meaning that we are selecting only AGN-powered radio sources, and that we are selecting the most powerful of radio AGN. Furthermore, in Figure 4.10 of section 4.4.3 we demonstrate how the radio luminosities of this sample are >1–3 orders of magnitude higher than the radio luminosities predicted from the IR luminosities due to star-formation.

Within the redshift range studied here (z = 0.2–2.5), there are 258 QSOs classified as radio-luminous.

4.3 Analysis

In this study we want to measure the average SFRs of optical QSOs as a function of their bolometric luminosity and redshift, using multiwavelength photometry covering the MIR–FIR, and taking into account of the AGN emission. With the sample of QSOs explored in this study we can extend the SFR – L_{AGN} plane of Stanley et al. (2015) by an order of magnitude in AGN luminosity, with 3026 sources covering the luminosities of $L_{AGN}=10^{45}$–10^{48} erg s^{-1}. Following Stanley et al. (2015), we have divided our sample in four redshift ranges, z = 0.2–0.5, 0.5–0.8, 0.8–1.5, and 1.5–2.5, which then are split in L_{AGN} bins of roughly equal number of sources (80–100 sources; see Table 4.1). For each z–L_{AGN} bin we performed stacking analysis in the *Herschel* PACS and SPIRE bands to estimate the mean 100μm, 160μm, 250μm, 350μm, and 500μm fluxes (section 4.3.1). We also calculate the mean 12μm and 22μm WISE fluxes (section 4.3.2), and mean bolometric AGN luminosities from the optical (see section 4.2.2). We then used the mean fluxes of each bin to perform composite SED fitting to decompose the IR luminosity into the AGN and star formation contributions (section 4.3.3). The combination of the multi-wavelength stacking and SED fitting, provides constraints on the mean IR luminosity due to star formation free from the possible AGN contamination, and the uncertainties of monochromatic estimations (see Chapter 2, section 2.6).

4.3.1 Stacking Herschel photometry

In this section we describe the methods followed to calculate the mean stacked flux for each z–L_{AGN} bin in our analysis. For each bin we perform a weighted mean stack of the H-ATLAS PACS-100μm, 160μm, and SPIRE-250μm, 350μm, and 500μm images at the optical positions of the SDSS QSOs. In all cases we regrid the images to pixels of 1arcsec, so as to have more accurate positioning. We used the noise maps to define the weighting on the mean, by taking the inverse of the noise as the weight; this also accounts for the fact that instrumental noise changes within the maps. The equation for the weighted mean is:

$$\langle x \rangle = \frac{\sum_0^n x_i \times w_i}{\sum_0^n w_i} \tag{4.3.2}$$

where x in our case is the flux density of each pixel of stacked image, x_i is the flux density of the equivalent pixel at all images used in the stack, and w_i is the inverted flux density at the equivalent pixel of the noise map.

From the mean stacked image (see Fig. 4.3 & 4.4) we estimate the mean flux density of the bin in two different ways depending on if its a PACS or SPIRE image. For the

PACS stacks we estimate the mean flux density by integrating the flux within an aperture of 3 arcsec radius and using the recommended aperture corrections of 2.63, and 3.57, for 100μm and 160μm respectively (Valiante et al. in prep). For the SPIRE stacks we take the flux density of the central pixel, as the units are flux density per beam.

To ensure that a stacked flux density measurement is significant, and not just stacked noise, we perform random stacks within the image. Random stacks are stacks that are done to a number of random positions on the map. Because each bin includes a different number of sources from each field, we perform random stacks for each bin individually, and define the number of random positions to be taken from each field to be the same to that of the real source stack of the bin. We perform 10000 random stacks of the maps to create a distribution of randomly stacked values. Examples of the resulting random stack distributions for all the bands are shown in Figures 4.3 & 4.4. The resulting random stack distributions for the SPIRE bands are not centered on zero, but are positively offset by typical values of 1.3mJy, 2mJy, and 0.5mJy, for the 250μm, 350μm, and 500μm respectively. The offset is caused by the fact that random stacks will include positive flux from the confused background (i.e., blending of faint sources) as well as bright sources on or near the random position. Additionally, the images have a non-zero mean that will also be contributing to the observed offset. We fit a Gaussian plus a positive tail to each random stack distribution. From the Gaussian fit we have derived the σ of the distribution. We use the 3σ of the distribution, with an added offset corresponding to the offset of the distribution from zero, as our detection limit. If a source is above the defined limit, then it is a detection and we use its absolute value, if it is below the limit we take an upper limit equal to the 3σ value of the random stack distribution.

Finally, as the offset of the random stack distribution reflects a boosting in flux flux density from the confusion background as well as random neighbouring sources, that will also affect our stacked fluxes of individual bins, we remove this offset from the stacked flux density in all bands for all z–L_{AGN} bins. An additional cause of uncertainty is the fact that QSOs are well known for their clustering (e.g., White et al. 2012 and references therein). In Wang et al. (2015) it was found that due to clustering of other dusty star-forming galaxies around optical QSOs there is a $\sim$ 8–13% contamination to the 250μm–500μm flux density, respectively. We test this for the bins in our sample by taking the

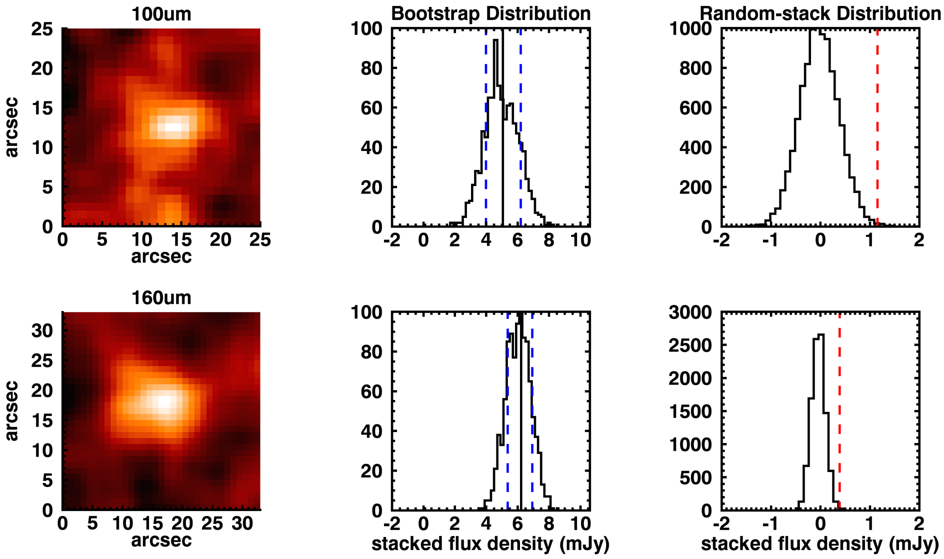


Figure 4.3: Examples of our stacking procedure for the PACS bands. First shown are the stacked images in 100μm and 160μm, followed by the bootstrap and random-stack distributions. The bootstrap distribution is a result of randomly re-sampling the sources in the stacks and estimating the stacked mean flux density 1000 times. The mean flux density of the bin is shown with the black line, and in blue dashed lines we show the 16th and 84th percentiles that correspond to the 1σ error on the mean. The random-stack distribution is produced by stacking at random positions in the images, the number of which is defined by the number of sources in the bin. The 99.5th percentile ($\sim 3\sigma$; red dashed line) is the limit we use to define if a stacked flux density is significant (see section 4.3.1).

radial light profile of the stacked images and find that the amount of contamination is equivalent to the offsets found within the random stack distributions of our bins. Therefore, our background subtraction method accounts for any potential boosting due to clustering effects. This analyses and its results are discussed in more detain in the Appendix of this chapter.

The errors on the mean fluxes are estimated using the bootstrap technique. We perform 1000 re-samplings for each bin and calculate the mean flux density of each. From the resulting distribution of mean flux densities we can define the 1σ errors by taking the taking the 16th and 84th percentiles (see examples in Fig. 4.3 & 4.4).

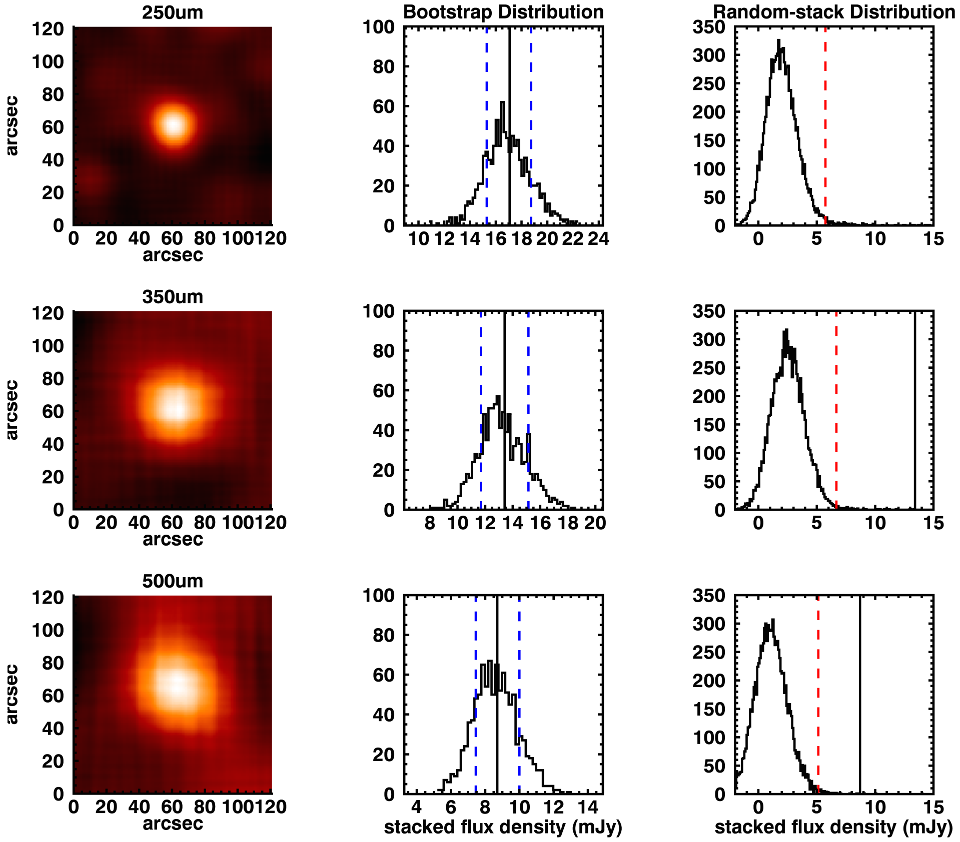


Figure 4.4: Examples of our stacking procedure for the SPIRE bands. First shown are the stacked images in 250μm, 350μm and 500μm, followed by the bootstrap, and the random-stack distributions. The bootstrap distribution is a result of randomly re-sampling the sources in the bin and estimating the stacked mean flux density 1000 times. The mean flux density of the bin is shown with the black line, and in blue dashed lines we show the 16th and 84th percentiles that correspond to the 1σ error on the mean. The random-stack distribution is produced by stacking the images at random positions, the number of which is defined by the number of sources in the bin. The 3σ percentile (red dashed line) is the limit we use to define if a stacked flux density is significant (see section 4.3.1). The black line corresponds to the mean stacked flux density of the bin. In the case of the 250μm stack the mean stacked flux density is larger than the range plotted with a value of 17mJy.

4.3.2 Mean flux densities of the WISE counterparts

For each z–L_{AGN} bin of our sample we took the mean flux densities at 12μm and 22μm for the sources with a WISE counterpart. On average ∼90% of the sources in each bin have a WISE counterpart. The fraction of upper limits in the z–L_{AGN} bins has a median of 1.3% and 32% in the 12 and 22μm bands respectively. When present, the limits show a random enough distribution amongst the measured flux densities, to allow us to use the non-parametric Kaplan-Meier estimator for the calculation of the mean of each bin, including both upper limits and measured flux densities (see Chapter 2 for detailed explanation/discussion; Stanley et al. 2015; Feigelson & Nelson 1985 for more details). We use this method for the estimation of the mean WISE fluxes in each bin of our sample. We chose this method over stacking the WISE photometry, as the source extraction that has been performed by the WISE team has taken into account of instrumental effects (Wright et al. 2010), providing good quality photometry. Furthermore, the almost complete detection rate makes stacking unnecessary. To test the uncertainties on our estimations, we take two extreme cases, where all the upper limit sources are given a value of 0, and where all upper limit sources are assumed detections at that limit. We find that the range between the two is less than 0.15mJy in 12μm band, and less than 2mJy in the 22μm band, and the K-M mean always lies within the range of these values. For this reason we trust that the K-M method is giving realistic results. We use bootstrap re-sampling to estimate the 1σ errors on the means.

4.3.3 Composite SED fitting

In Fig. 4.5 we show how the *Herschel* bands cover the peak of the star-forming templates in the redshifts of interest, making them essential for the estimation of the SFRs. However, the AGN could also be contributing to the FIR fluxes of each bin, especially at higher redshifts (see Fig. 4.5). For this reason we perform SED fitting to the WISE-12μm and 22μm, PACS-100μm, 160μm, and SPIRE-250μm, 350μm, and 500μm mean flux densities of each bin, and decompose the AGN and star formation contributions to the IR luminosity.

We follow the methods described in Stanley et al. (2015) (Chapter 3; described in

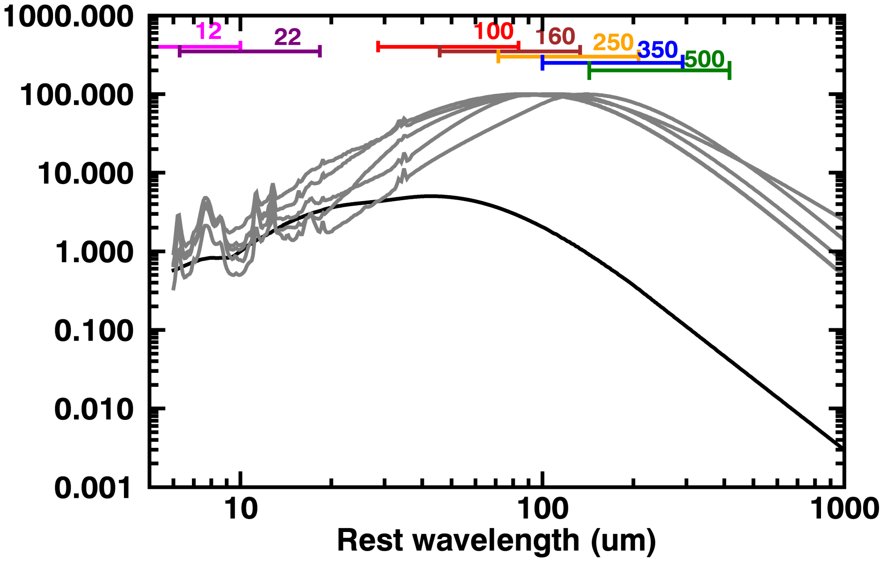


Figure 4.5: The IR SED templates of SF galaxies (grey curves), as well as the AGN template (black curve), used in this study. In coloured horizontal lines we show the range of the SED covered by the W3 (12μm) and W4(22μm) WISE bands and the five FIR *Herschel* bands of 100, 160, 250, 350, and 500μm at redshifts of 0.2–2.5. The 250μm band covers the peak region for the full redshift range of this study. However, at high redshifts (z >1) it nears the peak of the AGN SED, and hence it could suffer from significant contamination from AGN emission.

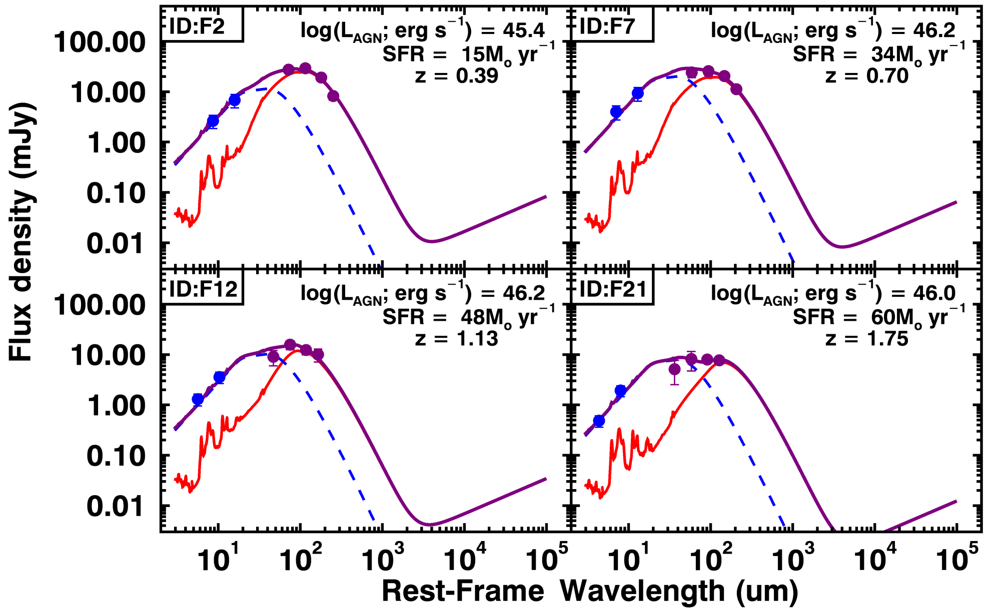


Figure 4.6: Examples of the SED fits from four L_{AGN}–z bins in our sample. The blue data points correspond to the WISE bands W3 and W4, while the purple data points correspond to the PACS 100μm, 160μm, and SPIRE 250μm, 350μm, and 500μm bands. The purple curve is the full SED, the blue dashed curve is the AGN template, and the red solid curve is the star-forming template. The AGN emission can significantly contribute to the PACS and SPIRE bands, especially at high redshifts.

detail in Chapter 2), which we briefly outlined here. We simultaneously fit an AGN template and a set of star-forming templates, and only allow the normalisation of the star-forming and AGN template as free parameters of the fit. The set of star-forming templates includes the five originally defined in Mullaney et al. (2011), and extended by Del Moro et al. (2013) to cover the wider wavelength range of 3–10^5 μm, as well as the Arp220 galaxy template from Silva et al. (1998) (see Fig. 4.5). For each bin we perform two sets of SED fitting, one using only the six different star-forming templates, and the other using the combination of the AGN and star-forming templates. Using the BIC parameter (Bayesian Information Criteria; Schwarz 1978) to compare the two sets of fits, we determine if a fit requires the AGN component, and find that all of our bins require the presence of the AGN counterpart in their IR SEDs. The fit with the minimum BIC value is taken to be the best-fitting result.

From the resulting best-fit SEDs we calculate the mean IR luminosity due to star formation of each bin, $\langle \mathrm{L_{IR,SF}} \rangle$, by integrating the SF component over 8–1000μm. The same is also done to estimate the mean IR (8–1000μm) luminosity of the AGN ($\langle \mathrm{L_{IR,AGN}} \rangle$) of each bin. For the error on the $\langle \mathrm{L_{IR,SF}} \rangle$, and $\langle \mathrm{L_{IR,AGN}} \rangle$, we propagate the error on the fit, and the range of luminosities of the fits within $\Delta\mathrm{BIC} = \mathrm{BIC} - \mathrm{BIC_{min}} \leq 2$. We chose to do the former as the fits that follow the above criteria, can be argued to be equally good fits (e.g., Liddle 2004).

Examples of best-fit SEDs for bins at the four different redshift ranges investigated in this thesis are given in Figure 4.6. The resulting best-fit SEDs for all the bins are shown in the Appendix of this chapter. We can see from Figure 4.6 that as we move towards higher redshifts the strong AGN component, present in all our fits, becomes dominant in the FIR bands. Indeed, as we show in section 4.4.1 the AGN is contributing up to 60% of the 250μm flux at redshifts of z ~2. Furthermore, we observe a tendency at redshifts of z >0.8 for the SED templates to be higher than the PACS-100μm photometry. This is predominant at the redshifts where the AGN is dominating at that wavelength, and so indicates that the shape of our AGN template should be hotter, and/or have a steeper drop-off at long wavelengths. The template of Mor & Netzer (2012), derived from a QSO sample with similar methods to Mullaney et al. (2011), has a steeper drop-off at shorter wavelengths (see Fig. 2.2 in Chapter 2). We repeat our SED fitting procedure using the mean Mor & Netzer (2012) AGN template, and find that over all the results on the $\langle \mathrm{L_{IR,SF}} \rangle$ do not change significantly, with a maximum increase in $\langle \mathrm{L_{IR,SF}} \rangle$ of a factor of $\sim$1.2, and so we argue that our choice of AGN template is sufficiently reliable.

To further test if our AGN template is suitable for this sample of QSOs, we compare the bolometric AGN luminosity derived from our fitted AGN components to that derived from the optical. To do this we use the 6μm luminosity of the fitted AGN components of our bins, and convert to an AGN bolometric luminosity with a bolometric correction factor of 8 (following Richards et al. 2006). We compare the IR derived bolometric AGN luminosity to that derived from the optical photometry and find that they are consistent within a factor of $\sim$1.5. Consequently, we trust that the AGN template that we use is reliable for this sample.

4.4 Results & Discussion

The aim of this study has been to constrain the mean SFRs of an optically selected sample of QSOs that have full *Herschel* coverage within the H-ATLAS survey. Initially, we compare our results of mean SFRs from our composite SEDs, to those from a monochromatic derivation at 250μm (section 4.4.1). We investigate the SFR properties of our full QSO sample (section 4.4.2), and of a radio-luminous sub-sample (section 4.4.3), and compare to the main sequence of star-forming galaxies.

4.4.1 Multi-band SED fitting VS single band derivation

A common method of previous studies estimating the SFRs of QSOs, is using stacking at observed frame 250μm from which the IR luminosity is derived. In this section we compare our results from the multi-wavelength composite SED fitting to the single band 250μm derivation, where we do not take into account the contribution from the AGN. To derive the average IR luminosities (integrated over 8–1000μm) from the 250μm stacked fluxes, we normalise the 6 SF galaxy templates that we used in our SED fitting method (see Fig. 4.5, and section 4.3.3), to the mean flux density at 250μm, and take the mean of the resulting $\langle \mathrm{L_{IR,SF}} \rangle$ of the 6 star formation templates.

In Fig. 4.7 we compare the results of the two methods described above, the mean IR

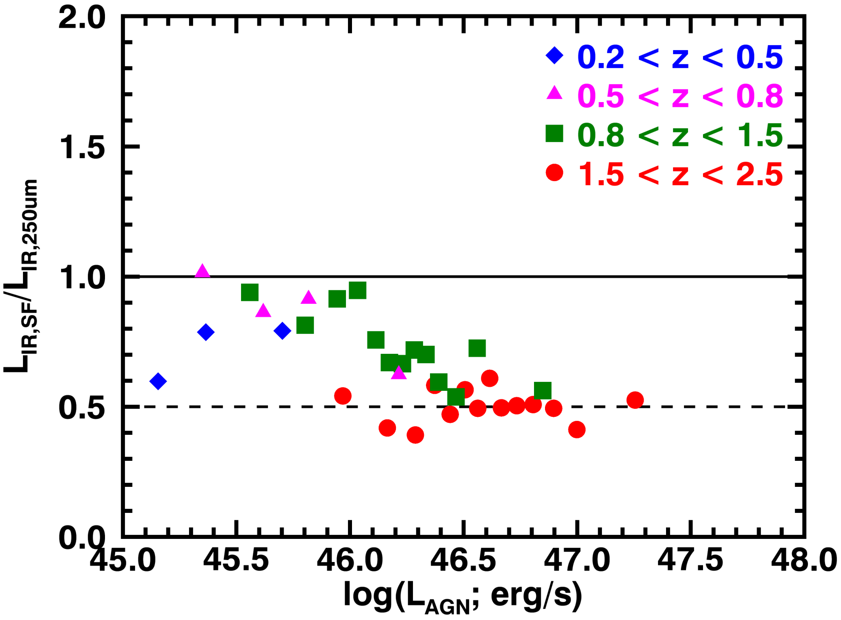


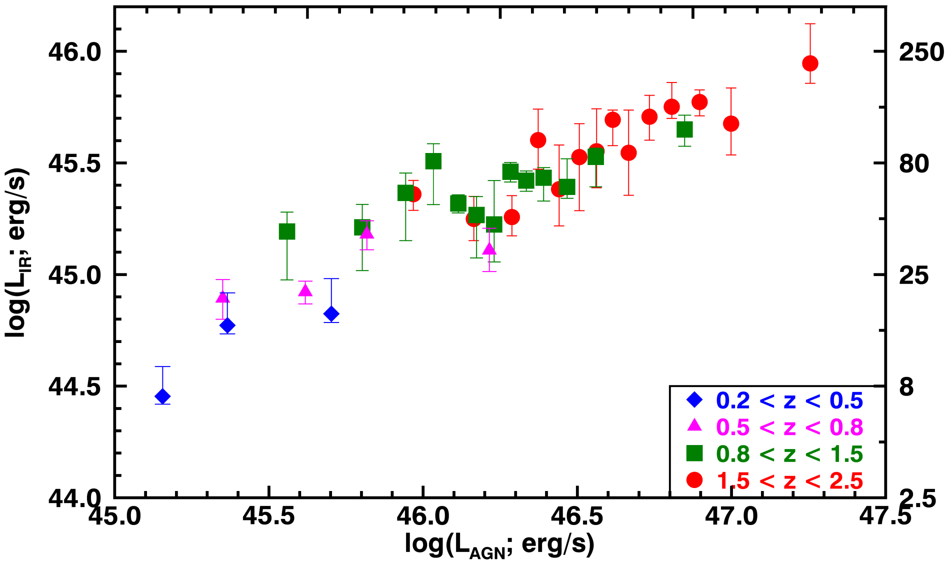
Figure 4.7: A comparison of the resulting mean IR luminosity due to star formation ($\langle L_{\rm IR,SF} \rangle$) from our composite SED fitting method, compared to the single-band derivation of the IR luminosity from the 250μm band ($L_{\rm IR,250\mu m}$). The solid line corresponds to the 1–1 line, and the dashed line is a factor of two offset from that. We find the 250μm band starts to be strongly contaminated by the AGN emission for high luminosity AGN ($L_{\rm AGN}$>10^{46} erg s^{-1}) and at high redhift (z >1.5).

luminosity derived from observed frame 250μm photometry, and the multi-wavelength SED fitting and decomposition method followed in our analyses. We find that for redshifts of $z \lesssim 0.5$ a single-band derivation from the 250μm band is not affected significantly by the AGN, with a median offset of a factor of 1.2. At redshifts of $z \gtrsim 0.5$ we see a more luminosity dependent effect, with the 250μm band being affected by the AGN by an increasing factor with AGN luminosity, reaching up to 50% AGN contribution at the highest luminosities ($L_{\rm AGN}$>10^{46} erg s^{-1}). At higher redshifts, z >1.5 AGN emission in the 250μm band consistently contributes ~50% and can reach up to 60% (see Fig. 4.6 SEDs of F12&F21).

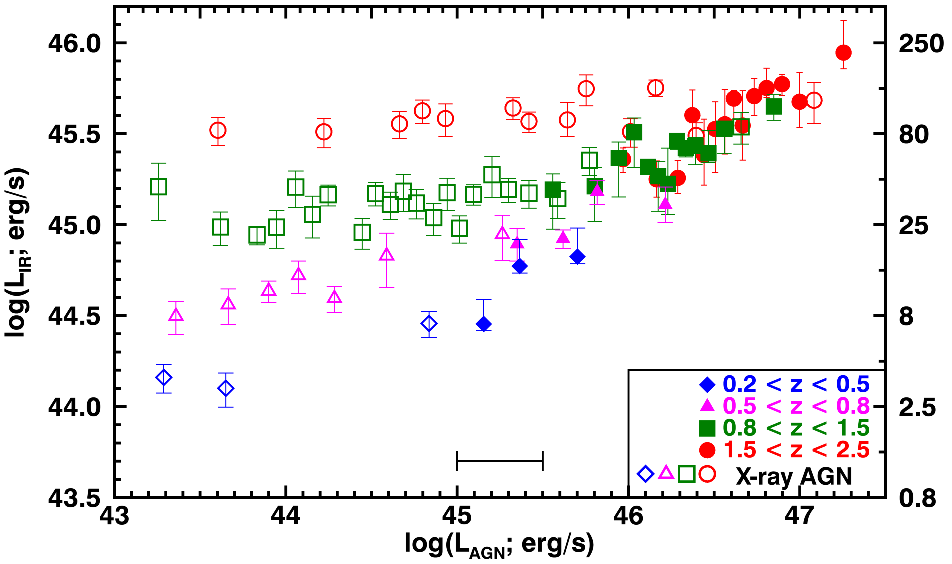
4.4.2 The mean SFRs of optical QSOs as a function of the bolometric AGN luminosity

As mentioned in section 3, we split our sample in bins of redshift and $\mathrm{L_{AGN}}$, for which we then estimate the mean $\mathrm{L_{IR,SF}}$ ($\langle \mathrm{L_{IR,SF}} \rangle$) through multi-wavelength stacking and SED fitting and decomposition into AGN and star-forming components. To take into account the known redshift evolution of the SFR (e.g., Schreiber et al. 2015) we split the QSO sample in redshift ranges of $z = 0.2$–0.5, 0.5–0.8, 0.8–1.5, and 1.5–2.5. For each redshift range we then split in luminosity bins of roughly the same number, ranging within ∼80–100 (see Table 4.1). From fitting the IR SED for each bin we then have a measure of the mean IR luminosity due to star formation ($\langle \mathrm{L_{IR,SF}} \rangle$) from integrating the SF component of the fit over 8–1000μm. In Fig. 4.8(a) we present our results on $\langle \mathrm{L_{IR,SF}} \rangle$ as a function of $\mathrm{L_{AGN}}$ and redshift. We see a positive trend of the $\langle \mathrm{L_{IR,SF}} \rangle$ as a function of $\mathrm{L_{AGN}}$ of more than an order of magnitude, something also observed in previous studies (e.g., Rosario et al. 2013b; Kalfountzou et al. 2014; Karouzos et al. 2014; Gürkan et al. 2015). However, when splitting in redshift ranges, we find that the observed trend is largely due to the redshift evolution of typical SFR values. Within each redshift range we still see a slight positive trend of $\langle \mathrm{L_{IR,SF}} \rangle$ with $\mathrm{L_{AGN}}$ in all redshift ranges, with the factor of increase ranging from ∼1.6–6.3 (0.2–0.8dex), with the highest redshift range of 1.5 – 2.5 showing the largest increase with $\mathrm{L_{AGN}}$. Even though we have divided our sample in redshift ranges, there could still be redshift dependencies within each redshift range that are driving the observed trends. To examine if the trend within each redshift range is driven by redshift dependencies, we look at the mean redshift of each bin as a function of $\mathrm{L_{AGN}}$ for each redshift range. We find that the mean redshift of the bins can not be the primary driver of the observed trends, as the increase in redshift among the bins corresponds to only a factor of $\sim 1.1 - 1.2$ in mean SFR, while the observed increase in $\langle \mathrm{L_{IR,SF}} \rangle$ is within factors of $\sim 1.6 - 6.3$, as mentioned above.

In our previous work (Stanley et al. 2015; Chapter 3) we constrained the $\langle \mathrm{L_{IR,SF}} \rangle$ for a sample of X-ray AGN in bins of redshift and $\mathrm{L_{AGN}}$. The sample of X-ray AGN covers 3 orders of magnitude in $\mathrm{L_{AGN}}$ of both moderate and high luminosity AGN ($10^{43} < \mathrm{L_{AGN}} < 5 \times 10^{47}$ erg s^{-1}). The sample of high luminosity optical QSOs in this work is ideal to



(a)



(b)

Figure 4.8: (a)$\langle L_{IR,SF} \rangle$ as a function of AGN bolometric luminosity (L_{AGN}). The coloured filled symbols show the results for the full QSO sample in L_{AGN}–z bins. (b) $\langle L_{IR,SF} \rangle$ as a function of AGN bolometric luminosity (L_{AGN}) for the combination of the X-ray AGN sample from Stanley et al. (2015) and the current sample of optical QSOs. The two samples are complementary to each other, and together cover 3–4 orders of magnitude in L_{AGN}. We give the relative uncertainty between the L_{AGN} values of the two samples as an error on the x-axis, stemming from the fact that the X-ray AGN sample has L_{AGN} values derived from X-ray photometry, while for the QSO sample it has been derived from optical photometry. We attribute this slight trend to increasing stellar masses across the bins (see Fig. 4.9)

ID (a)	N (b)	$\langle z \rangle$ (c)	$\langle M_{BH} \rangle$ ($M_\odot$) (d)	$\langle L_{AGN,bol} \rangle$ ($\mathrm{erg\,s^{-1}}$) (e)	$\langle L_{IR,SF} \rangle$ ($\mathrm{erg\,s^{-1}}$) (f)
F1	83	$0.321^{+0.078}_{-0.075}$	$0.37^{+0.25}_{-0.30} \times 10^9$	$1.43^{+0.40}_{-0.48} \times 10^{45}$	$2.85^{+1.03}_{-0.22} \times 10^{44}$
F2	80	$0.394^{+0.067}_{-0.076}$	$2.95^{+2.08}_{-2.34} \times 10^8$	$2.32^{+0.67}_{-0.46} \times 10^{45}$	$0.59^{+0.24}_{-0.05} \times 10^{45}$
F3	88	$0.410^{+0.059}_{-0.054}$	$0.46^{+0.09}_{-0.36} \times 10^9$	$0.50^{+1.37}_{-0.20} \times 10^{46}$	$0.67^{+0.29}_{-0.06} \times 10^{45}$
F4	89	$0.635^{+0.107}_{-0.082}$	$0.47^{+0.25}_{-0.36} \times 10^9$	$0.42^{+0.10}_{-0.10} \times 10^{46}$	$0.84^{+0.10}_{-0.10} \times 10^{45}$
F5	94	$0.640^{+0.114}_{-0.102}$	$0.33^{+0.24}_{-0.25} \times 10^9$	$2.24^{+0.96}_{-1.61} \times 10^{45}$	$0.78^{+0.17}_{-0.15} \times 10^{45}$
F6	94	$0.670^{+0.098}_{-0.089}$	$0.66^{+0.48}_{-0.47} \times 10^9$	$0.66^{+0.17}_{-0.14} \times 10^{46}$	$1.52^{+0.22}_{-0.22} \times 10^{45}$
F7	96	$0.697^{+0.076}_{-0.066}$	$1.12^{+0.60}_{-0.84} \times 10^9$	$1.64^{+4.73}_{-0.80} \times 10^{46}$	$1.29^{+0.33}_{-0.26} \times 10^{45}$
F8	85	$0.989^{+0.161}_{-0.156}$	$0.61^{+0.40}_{-0.46} \times 10^9$	$0.36^{+0.14}_{-0.22} \times 10^{46}$	$1.56^{+0.34}_{-0.61} \times 10^{45}$
F9	90	$1.080^{+0.216}_{-0.181}$	$0.74^{+0.48}_{-0.43} \times 10^9$	$1.08^{+0.12}_{-0.11} \times 10^{46}$	$0.32^{+0.06}_{-0.12} \times 10^{46}$
F10	89	$1.100^{+0.248}_{-0.236}$	$0.80^{+0.39}_{-0.55} \times 10^9$	$0.88^{+0.10}_{-0.11} \times 10^{46}$	$2.08^{+0.19}_{-0.19} \times 10^{45}$
F11	86	$1.104^{+0.255}_{-0.202}$	$1.03^{+0.44}_{-0.65} \times 10^9$	$1.30^{+0.10}_{-0.10} \times 10^{46}$	$2.32^{+0.52}_{-0.90} \times 10^{45}$
F12	82	$1.132^{+0.185}_{-0.166}$	$1.01^{+0.69}_{-0.64} \times 10^9$	$1.49^{+0.10}_{-0.09} \times 10^{46}$	$1.85^{+0.39}_{-0.66} \times 10^{45}$
F13	86	$1.133^{+0.210}_{-0.189}$	$0.63^{+0.54}_{-0.42} \times 10^9$	$0.63^{+0.13}_{-0.13} \times 10^{46}$	$1.63^{+0.44}_{-0.59} \times 10^{45}$
F14	84	$1.157^{+0.203}_{-0.191}$	$0.97^{+0.60}_{-0.60} \times 10^9$	$1.70^{+0.11}_{-0.11} \times 10^{46}$	$1.68^{+0.96}_{-0.54} \times 10^{45}$
F15	82	$1.175^{+0.181}_{-0.202}$	$1.21^{+0.90}_{-0.79} \times 10^9$	$1.92^{+0.10}_{-0.11} \times 10^{46}$	$2.89^{+0.29}_{-0.29} \times 10^{45}$
F16	89	$1.223^{+0.151}_{-0.140}$	$1.14^{+0.60}_{-0.65} \times 10^9$	$2.16^{+0.14}_{-0.14} \times 10^{46}$	$2.63^{+0.28}_{-0.28} \times 10^{45}$
F17	85	$1.245^{+0.202}_{-0.204}$	$1.39^{+0.67}_{-0.76} \times 10^9$	$2.93^{+0.23}_{-0.33} \times 10^{46}$	$0.34^{+0.03}_{-0.09} \times 10^{46}$
F18	87	$1.254^{+0.201}_{-0.187}$	$1.74^{+0.71}_{-0.87} \times 10^9$	$0.36^{+0.05}_{-0.05} \times 10^{47}$	$2.47^{+0.83}_{-0.27} \times 10^{45}$
F19	99	$1.272^{+0.188}_{-0.229}$	$2.34^{+1.49}_{-1.37} \times 10^9$	$0.70^{+1.95}_{-0.29} \times 10^{47}$	$0.45^{+0.07}_{-0.07} \times 10^{46}$
F20	87	$1.273^{+0.200}_{-0.173}$	$1.29^{+0.71}_{-0.79} \times 10^9$	$2.46^{+0.14}_{-0.15} \times 10^{46}$	$2.71^{+0.30}_{-0.58} \times 10^{45}$
F21	86	$1.750^{+0.158}_{-0.205}$	$1.02^{+0.74}_{-0.72} \times 10^9$	$0.93^{+0.26}_{-0.66} \times 10^{46}$	$2.29^{+0.35}_{-0.35} \times 10^{45}$
F22	90	$1.776^{+0.209}_{-0.217}$	$1.63^{+0.98}_{-1.04} \times 10^9$	$0.37^{+0.02}_{-0.02} \times 10^{47}$	$0.36^{+0.20}_{-0.11} \times 10^{46}$
F23	91	$1.777^{+0.233}_{-0.217}$	$1.69^{+1.39}_{-1.12} \times 10^9$	$2.76^{+0.20}_{-0.23} \times 10^{46}$	$2.41^{+1.40}_{-0.76} \times 10^{45}$
F24	97	$1.782^{+0.254}_{-0.216}$	$1.59^{+1.04}_{-1.02} \times 10^9$	$0.32^{+0.02}_{-0.02} \times 10^{47}$	$0.34^{+0.14}_{-0.14} \times 10^{46}$
F25	88	$1.785^{+0.330}_{-0.240}$	$1.84^{+0.98}_{-1.25} \times 10^9$	$2.36^{+0.16}_{-0.16} \times 10^{46}$	$0.40^{+0.15}_{-0.10} \times 10^{46}$
F26	90	$1.847^{+0.237}_{-0.236}$	$1.22^{+0.80}_{-0.86} \times 10^9$	$1.46^{+0.26}_{-0.26} \times 10^{46}$	$1.81^{+0.45}_{-0.32} \times 10^{45}$
F27	93	$1.853^{+0.207}_{-0.251}$	$1.97^{+1.29}_{-1.34} \times 10^9$	$0.41^{+0.02}_{-0.03} \times 10^{47}$	$1.78^{+0.47}_{-0.36} \times 10^{45}$
F28	93	$1.854^{+0.330}_{-0.299}$	$1.13^{+0.77}_{-0.80} \times 10^9$	$1.94^{+0.25}_{-0.22} \times 10^{46}$	$0.49^{+0.05}_{-0.12} \times 10^{46}$
F29	88	$1.859^{+0.256}_{-0.262}$	$2.16^{+1.30}_{-1.49} \times 10^9$	$0.46^{+0.04}_{-0.03} \times 10^{47}$	$0.35^{+0.19}_{-0.12} \times 10^{46}$
F30	80	$1.879^{+0.224}_{-0.257}$	$2.12^{+1.33}_{-1.21} \times 10^9$	$0.54^{+0.05}_{-0.04} \times 10^{47}$	$0.51^{+0.13}_{-0.11} \times 10^{46}$
F31	93	$1.911^{+0.240}_{-0.244}$	$2.30^{+0.78}_{-1.22} \times 10^9$	$0.64^{+0.06}_{-0.05} \times 10^{47}$	$0.56^{+0.16}_{-0.06} \times 10^{46}$
F32	89	$2.015^{+0.236}_{-0.289}$	$2.61^{+0.93}_{-1.61} \times 10^9$	$0.79^{+0.09}_{-0.08} \times 10^{47}$	$0.59^{+0.08}_{-0.08} \times 10^{46}$
F33	99	$2.053^{+0.258}_{-0.246}$	$0.48^{+0.28}_{-0.28} \times 10^{10}$	$1.80^{+13.16}_{-0.66} \times 10^{47}$	$0.88^{+0.45}_{-0.16} \times 10^{46}$
F34	94	$2.058^{+0.299}_{-0.310}$	$0.37^{+0.19}_{-0.22} \times 10^{10}$	$1.00^{+0.14}_{-0.12} \times 10^{47}$	$0.47^{+0.21}_{-0.13} \times 10^{46}$

Table 4.1: Table of the properties of each bin in our sample of optical QSOs. (a) The ID of the bin that corresponds to the SEDs presented in the Appendix. (b) The number of sources in each bin. (c) The mean redshift of each bin. (d) The mean BH mass of each bin. The errors in (c) and (d) correspond to the 16th to the 84th percentiles of the values in each bin. (e) The mean AGN bolometric luminosity of each bin, as errors we give the full range of values covered in each bin. (f) The mean IR luminosity due to star formation of each bin, the errors are defined by the combination of the error on the fit and the range of $\langle L_{IR,SF} \rangle$ values from other templates that had good SED fits (see section 4.3.3).

extend and fill-in the $\langle L_{IR,SF} \rangle$–$L_{AGN}$ plane as defined in Stanley et al. (2015) in the highest L_{AGN} region, and the search for systematic differences between the two population of AGN. In Fig. 4.8(b) we plot the $\langle L_{IR,SF} \rangle$ as a function of L_{AGN} for both the X-ray AGN and optical QSOs extending the $\langle L_{IR,SF} \rangle$–$L_{AGN}$ plane to 4 orders of magnitude. Where there is overlap between the X-ray selected AGN sample of Stanley et al. (2015) and our current sample of optical QSOs, we see a good agreement in $\langle L_{IR,SF} \rangle$ values. [1]

Comparing to the main sequence of star-forming galaxies

The main sequence of star-forming galaxies is defined from the observed correlation between SFR and stellar mass, and has been found to evolve with redshift (e.g., Noeske et al. 2007; Elbaz et al. 2011; Schreiber et al. 2015). As well as providing a general description of the SFRs of the general star-forming population, the equation of the main sequence also provides a measure of the known dependencies of the SFR of a galaxy to its stellar mass and redshift. By comparing our results to the mean SFRs of the main sequence galaxies with the same redshift and stellar masses as our sample, we can simultaneously see if the observed trends of the $\langle L_{IR,SF} \rangle$–$L_{AGN}$ plane are driven by stellar mass dependencies, and if our QSO sample shows systematic differences to the general star-forming population (that does not host QSOs). In Stanley et al. (2015) we showed that when taking into account the stellar masses and redshifts of the X-ray AGN sample, their SFRs are consistent with the main sequence star-forming galaxies.

We make use of two equations for defining the mean IR luminosity of main sequence galaxies ($\langle L_{IR,MS} \rangle$) of each bin based on the BH mass and redshift distribution of the sources included. We use Eq. 9 of Schreiber et al. (2015) to calculate the $L_{IR,MS}$:

$$\log_{10}(SFR_{MS}[M_\odot/yr]) = m - 0.5 + 1.5r - 0.3[max(0, m - 0.36 - 2.5r)]^2 \qquad (4.4.3)$$

---

[1] We note that there is a relative uncertainty ∼0.56 dex between the AGN bolometric luminosities of the X-ray AGN and optical QSO samples, as they have been derived in different ways from different photometry. However, despite the uncertainty on comparing these samples will not be affected the observed trends significantly. They uncertainty has been estimated using the 2XMM to SDSS DR7 cross-correlated catalogue from Pineau et al. (2011). We take the X-ray hard band flux and calculate a bolometric luminosity, and compare to the bolometric luminosity from the optical. We take the ratio of the two, and find that there is a median offset of 3.6 (or 0.56dex).

where $r = \log_{10}(1+z)$, $m = \log_{10}(M_{\star}/10^9 M_{\odot})$, and $\mathrm{L_{IR,MS}} = \mathrm{SFR_{MS}}/4.5\mathrm{d}-44$ (as Schreiber et al. 2015 assume a Salpeter IMF). The 1σ scatter of the relation is +/- 0.3dex and remains for at least up to a redshift of ~4 (Schreiber et al. 2015). As can be seen in the above equation, to estimate the $\langle \mathrm{L_{IR,MS}} \rangle$ we need a measurement of the stellar masses of our sample. As our sample is consisted by QSOs, where the QSO emission overpowers that of the host galaxy in the optical, it is not possible to use SED fitting methods to the optical photometry to calculate reliable stellar masses, and only BH masses are available (see section 4.2.1). Consequently, we derive stellar mass estimates from the available BH masses. To do this we use Eq. 10 from Kormendy & Ho (2013) derived from local galaxies, to convert the mean BH masses of each bin to stellar masses, and follow the common practice of assuming $\mathrm{M}_* \sim M_{bulge}$.

$$\frac{M_{BH}}{10^9 M_{\odot}} = 0.49^{+0.06}_{-0.05} \left(\frac{M_{bulge}}{10^{11} M_{\odot}} \right)^{1.16 \pm 0.08} \tag{4.4.4}$$

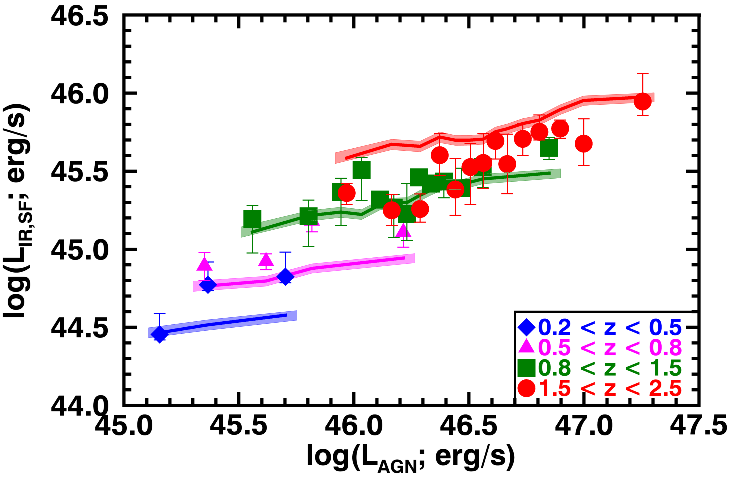
.

To establish if our optical QSO sample has $\langle \mathrm{L_{IR,SF}} \rangle$ values consistent with the general star-forming population, we test if the derived $\langle \mathrm{L_{IR,SF}} \rangle$ are consistent with the hosts of our optical QSOs being a randomly selected sample from the main sequence of star-forming galaxies. To do this we follow a similar approach to Rosario et al. (2013b) for the definition of a representative sample of main sequence galaxies for the properties of our sample, we extend the comparison of the star-forming properties of the QSOs to the main sequence, to higher AGN bolometric luminosities, and lower redshifts.

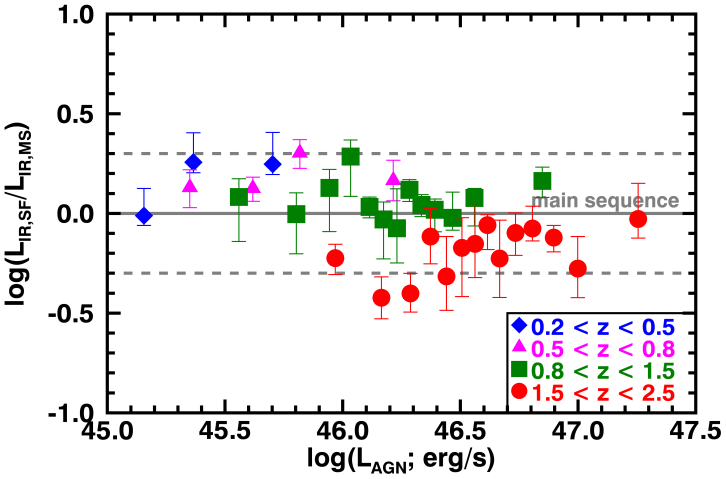
For each bin, we perform a Monte-Carlo estimation of the $\langle \mathrm{L_{IR,MS}} \rangle$ corresponding to the properties of the sources in the bin. Using Eq. 4.4.3, we define a distribution of possible stellar masses for each source based on their BH mass, and pick a random value from the distribution. The width of the stellar mass distribution includes both the errors in Eq. 4.4.3 and the error of the BH mass (provided by Shen et al. 2011; see section 4.2.1). Based on the chosen stellar mass, and the known redshift of the source we define a log-normal distribution of $\mathrm{L_{IR,MS}}$ values centered at the result from Eq. 4.4.1, with a σ of 0.3dex. We pick a random value from the distribution of $\mathrm{L_{IR,MS}}$. At the end, all sources within the bin will have a randomly picked $\mathrm{L_{IR,MS}}$ and we can take the $\langle \mathrm{L_{IR,MS}} \rangle$ of the bin. The above process is repeated 10,000 times for each bin, and results in a distribution of $\langle L_{\rm IR,MS} \rangle$ from which we can define the mean and 1σ range of the possible $\langle L_{\rm IR,MS} \rangle$ values for the respective bin.

In Fig. 4.9(a) & (b) we plot the results for $\langle L_{\rm IR,MS} \rangle$ in comparison to the $\langle L_{\rm IR,SF} \rangle$ of the QSO sample. In Fig. 4.9(a) we plot the $\langle L_{\rm IR,SF} \rangle$ as a function of $L_{\rm AGN}$ in comparison to the main sequence results for each redshift range. With coloured lines we show the mean, and the coloured shaded regions correspond to the 1σ range of the distribution $\langle L_{\rm IR,MS} \rangle$ of each bin from the MC simulation. Additionally, we take the ratio of the $\langle L_{\rm IR,SF} \rangle$ from our analysis over that of the main sequence ($\langle L_{\rm IR,MS} \rangle$). We show the $\langle L_{\rm IR,SF} \rangle / \langle L_{\rm IR,MS} \rangle$ ratio as a function of $L_{\rm AGN}$ in Fig. 4.9(b), where the errors combine the errors on both variables. With the line we show the 0 offset from the main sequence, while with the dashed lines we give the range covered by the scatter of the main sequence relation as defined by Schreiber et al. (2015). From these two figures we can see an evolution in the $\langle L_{\rm IR,SF} \rangle$ values of QSOs relative to those of the main sequence star-forming galaxies, as a function of redshift. At the highest redshift range of $1.5 < z < 2.5$ the $\langle L_{\rm IR,SF} \rangle$ values are systematically below the main sequence, with an average offset of a factor of 0.7 (or 1.42 if taking the inverse ratio). Moving to intermediate redshifts of $0.8 < z < 1.5$ the $\langle L_{\rm IR,SF} \rangle$ values become consistent with those of the main sequence, while at redshifts of $z < 0.8$ the $\langle L_{\rm IR,SF} \rangle$ values move above those of the main sequence by a factor of 1.5. Even though the means are not consistent within their errors, they are still consistent within the scatter of the main sequence (see Fig. 4.9(b)), and so are consistent within a factor of 2. One possible explanation for the redshift dependent offset from the main sequence could that the local M_{BH}–M_* relation evolves with redshift, while we have assumed remains the same throughout all the redshifts explored. Indeed, there have been studies with results indicating that the M_{BH} relative to the M_* of the host galaxy, moves off the relation to higher M_{BH} values with redshift (up to an order of magnitude at $z \sim 2$; e.g., Merloni et al. 2010; Drouart et al. 2014). If the above is true, then we have overestimated the M_* of the QSOs in our highest redshifts, and subsequently boosted the expected $\langle L_{\rm IR,MS} \rangle$. Overall, it is safe to say that the $\langle L_{\rm IR,SF} \rangle$ values of QSOs remain consistent with those of main sequence star-forming galaxies within a factor of 2.

As discussed above the assumptions made for the estimation of the M_* provide an



(a)



(b)

Figure 4.9: (a) $\langle L_{IR,SF}\rangle$ as a function of L_{AGN}. In coloured regions are the regions of $\langle L_{IR,MS}\rangle$ covered by the main sequence galaxies at each of the redshift ranges. The shaded regions reflect the bootstrap error on the $\langle L_{IR,MS}\rangle$ (see section 4.4.2 for details on defining $\langle L_{IR,MS}\rangle$ and the errors). (b) The ratio of the $\langle L_{IR,SF}\rangle$ of the QSOs over the $\langle L_{IR,MS}\rangle$ of the main sequence for galaxies of the same stellar mass and redshift, as a function of L_{AGN}. The errors on the ratio combine both the errors of the $\langle L_{IR,SF}\rangle$ and $\langle L_{IR,MS}\rangle$. The dashed lines indicate a factor of 2 offset on both sides of the line, characteristic of the error on the main sequence equation. There is an evident evolution of our QSO sample relative to the main sequence of star-forming galaxies, moving from having comparatively higher $\langle L_{IR,SF}\rangle$ values at low redshifts ($0.2< z <0.5$) to lower $\langle L_{IR,SF}\rangle$ values than the main sequence at the highest redshift range of our sample ($1.5< z <2.5$). This seeming evolution in the star formation of QSOs could instead be an evolution of the M_{BH}–M_* relation that we have not taken into account (see §4.2.1).

uncertainty on the normalisation of the main sequence tracks; however, there is no reason for that to affect the general trends. The positive trends observed in the $\langle L_{IR,SF}\rangle$ as a function of L_{AGN} seem to follow those of the $\langle L_{IR,MS}\rangle$ (see Fig. 4.9(a)), suggesting that the observed correlation is primarily driven by the stellar masses and redshifts of the QSOs.

4.4.3 The mean SFRs of Radio-luminous QSOs

In parallel to our analysis of the full sample of QSOs, we also analysed a sub-sample of radio-luminous QSOs selected based on a radio-luminosity ($L_{1.4GHz}$) cut (see Fig. 4.2 and section 4.2.3), in the redshift ranges of $0.5 < z < 0.8$, $0.8< z <1.5$, and $1.5< z <2.5$. As we show below, the radio-luminosities of our sample are at least an order of magnitude above those corresponding to the $\langle L_{IR,SF}\rangle$ of our bins, and so we are confident that these radio luminosities are dominated by the QSO and not the star formation emission. For each redshift range we split the sample in $L_{1.4GHz}$ bins of roughly equal numbers (see Table 4.2). Due to the limited number of sources we can only have two bins in each redshift range. For each bin we follow the procedures described in section 4.3, to estimate the $\langle L_{IR,SF}\rangle$.

In Fig. 4.10 we plot $\langle L_{IR,SF}\rangle$ as a function of $L_{1.4GHz}$ of each bin, colour-coded to the redshift range they belong in. We also plot the IR-radio correlation (e.g.,Magnelli et al. 2014; Pannella et al. 2015) multiplied by factors of 50, 500, and 5000, to demonstrate how the radio luminosities of our sample are a factor of $\sim$10–5000 above those corresponding to their $\langle L_{IR,SF}\rangle$ values. Even though we only have two luminosity bins in each redshift range, the $\langle L_{IR,SF}\rangle$ values as a function of $\langle L_{1.4GHz}\rangle$ are suggestive of a flat trend, further suggesting that the radio luminosity does not originate from the star formation in these systems and also indicating the lack of a direct relationship between the star formation emission of the galaxy and the radio-emission of the QSOs.

Following the same approach as for the full QSO sample, we estimate the expected IR luminosity of main sequence star-forming galaxies ($\langle L_{IR,MS}\rangle$) of the same redshift and stellar mass (estimated from the available M_{BH}) as our radio-luminous QSO sample, and compare to their $\langle L_{IR,SF}\rangle$. In Figure 4.11 we show the $\langle L_{IR,SF}\rangle/\langle L_{IR,MS}\rangle$ ratio as a function of L_{AGN}. We find that the radio-luminous QSOs have $\langle L_{IR,SF}\rangle$ values consistent with those of the main sequence within a factor of 2. Similar results were shown by Drouart et al. (2014) (following a similar SED fitting approach) for a smaller sample of 70 powerful radio-AGN, that found that for $z < 2.5$ the SFRs of radio-AGN are consistent with the main sequence of star-forming galaxies. Additionally a number of studies have argued for radio-AGN/QSOs living in star-forming galaxies up to redshifts of $\sim$5 (e.g., Drouart et al. 2014, Rees et al. 2016, Magliocchetti et al. 2016) following a variety of approaches.

We compare the radio-luminous QSOs to the overall QSO sample (as it is dominated by radio-quiet QSOs), and find that their $\langle L_{\mathrm{IR,SF}} \rangle$ results are broadly consistent within error, and show similar trends with redshift. In previous work comparing radio-loud to radio-quiet QSOs, for similar redshifts and L_{AGN}, Kalfountzou et al. (2014) also find them to be consistent, but for lower redshifts ($z < 0.6$), Gürkan et al. (2015) found that radio-luminous QSOs have lower SFRs than RQ-QSOs. Due to the limited number of radio-luminous sources we have not explored redshifts below $z < 0.5$, and so can not directly compare to the results of Gürkan et al. (2015). However, it is worth noting that the samples of radio-luminous QSOs presented here and in Gürkan et al. (2015) are dominated by different types of excitation levels. Our sample is dominated by high excitation radio galaxies (HERGs; Best & Heckman 2012) that are high-accretion-rate radio bright AGN, while the sample of Gürkan et al. (2015) is consisted predominantly by low excitation radio galaxies (LERGs; Best & Heckman 2012) that are low-accretion-rate radio-faint AGN. As HERGs and LERGs represent AGN populations of different modes and with different fuelling mechanisms (e.g., Hardcastle et al. 2007; Best & Heckman 2012 Heckman & Best 2014), the differences in their SFRs (e.g., Hardcastle et al. 2013; Gürkan et al. 2015) could be indicating that we need to investigate different AGN modes, to fully understand the impact of AGN on their hosts.

4.5 Conclusions

The aim of this work has been to constrain the mean SFRs of a sample of $z = 0.2$–2.5 QSOs with AGN bolometric luminosities of $10^{45} < L_{\mathrm{AGN}} < 10^{48}$ erg s^{-1}. We investigate the mean SFRs as a function of redshift and bolometric AGN luminosity of the whole sample, and a radio-luminous sub-sample with $L_{\mathrm{AGN}} > 10^{24}\,\mathrm{W\,Hz^{-1}}$. We combine the

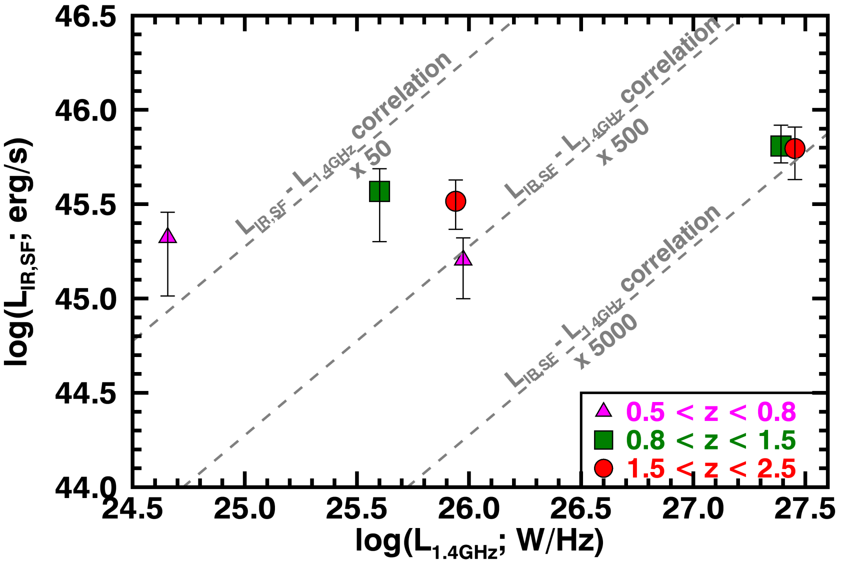


Figure 4.10: $\langle L_{\mathrm{IR,SF}} \rangle$ in bins of redshift and radio luminosity ($L_{1.4\mathrm{GHz}}$), as a function of $\langle L_{1.4\mathrm{GHz}} \rangle$. With dashed lines we indicate the IR-radio relation of star-forming galaxies increased by factors of 50–5000, to demonstrate that the radio luminosities of our sources cannot be attributed to their star formation. There is no evidence for a relation between $\langle L_{\mathrm{IR,SF}} \rangle$ and $L_{1.4\mathrm{GHz}}$, with the general trend being flat. However, we are limited by the number of sources and can only have two bins in each redshift range.

ID (a)	N (b)	$\langle z \rangle$ (c)	$\langle M_{BH} \rangle$ ($M_\odot$) (d)	$\langle L_{1.4GHz} \rangle$ ($\mathrm{W\,Hz^{-1}}$) (e)	$\langle L_{AGN,bol} \rangle$ ($\mathrm{erg\,s^{-1}}$) (f)	$\langle L_{IR,SF} \rangle$ ($\mathrm{erg\,s^{-1}}$) (g)
R1	17	$0.663^{+0.131}_{-0.074}$	$1.27^{+2.40}_{-1.07} \times 10^9$	$0.45^{+0.65}_{-0.30} \times 10^{25}$	$1.00^{+3.44}_{-0.84} \times 10^{46}$	$2.60^{+0.55}_{-1.30} \times 10^{45}$
R2	15	$0.710^{+0.076}_{-0.080}$	$1.52^{+2.99}_{-1.25} \times 10^9$	$0.94^{+4.62}_{-0.83} \times 10^{26}$	$1.22^{+1.67}_{-0.98} \times 10^{46}$	$1.00^{+0.17}_{-0.17} \times 10^{45}$
R3	53	$1.131^{+0.243}_{-0.187}$	$1.91^{+2.10}_{-1.56} \times 10^9$	$0.40^{+0.44}_{-0.28} \times 10^{26}$	$1.99^{+1.68}_{-1.45} \times 10^{46}$	$0.41^{+0.10}_{-0.22} \times 10^{46}$
R4	50	$1.180^{+0.192}_{-0.218}$	$1.52^{+0.97}_{-1.20} \times 10^9$	$2.46^{+43.36}_{-2.37} \times 10^{27}$	$3.05^{+1.60}_{-2.27} \times 10^{46}$	$0.52^{+0.11}_{-0.09} \times 10^{46}$
R5	49	$1.882^{+0.394}_{-0.321}$	$2.68^{+2.65}_{-1.98} \times 10^9$	$2.84^{+62.09}_{-2.55} \times 10^{27}$	$0.79^{+0.81}_{-0.60} \times 10^{47}$	$0.42^{+0.09}_{-0.06} \times 10^{46}$
R6	54	$1.913^{+0.308}_{-0.286}$	$2.28^{+1.12}_{-1.69} \times 10^9$	$0.87^{+1.68}_{-0.71} \times 10^{26}$	$0.58^{+0.49}_{-0.43} \times 10^{47}$	$3.12^{+1.48}_{-1.36} \times 10^{45}$

Table 4.2: Table of the properties of each bin in our sub-sample of radio-luminous QSOs. (a) The ID of the bin that corresponds to the set of SEDs presented in the Appendix. (b) The number of sources in each bin. (c) The mean redshift of each bin. (d) The mean BH mass of each bin. (e) The mean radio luminosity at 1.4GHz, with the errors being the range of values covered in each bin. (f) The mean AGN bolometric luminosity of each bin. The errors in (c), (d), (f) correspond to the 16th to the 84th percentiles of the values in each bin. (g) The mean IR luminosity due to star formation of each bin, the errors are defined by the combination of the error on the fit and the range of $\langle L_{\mathrm{IR,SF}} \rangle$ values from other templates that had good SED fits (see section 4.3.3).

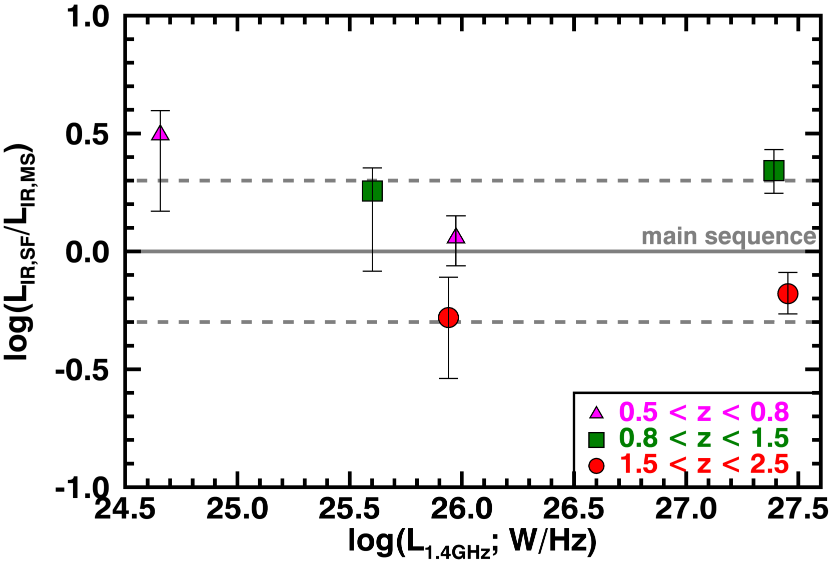


Figure 4.11: The ratio of the $\langle L_{IR,SF} \rangle$ of the radio-luminous QSOs over the $\langle L_{IR,MS} \rangle$ of the main sequence for galaxies of the same stellar mass and redshift, as a function of $L_{1.4GHz}$. We find that the $\langle L_{IR,SF} \rangle$ of radio-luminous QSOs are consistent with those of main sequence and starbursts galaxies for redshifts of $0.5 < z < 1.5$. At higher redshifts of $1.5 < z < 2.5$ the radio-luminous QSOs show lower values of $\langle L_{IR,SF} \rangle$ relative to the main sequence galaxies, in agreement with what we see for the overall sample.

five *Herschel* bands (100–500μm) of the H-ATLAS survey to the MIR bands (12 and 22μm) of WISE, and perform SED fitting to the mean fluxes of 34 $L_{\rm AGN}$–z bins of our full QSO sample, and 6 $L_{\rm 1.4GHz}$–z bins of the radio-luminous QSO sub-sample. We find that:

- It is important to take into account of AGN contamination in the FIR when calculating the SFRs of such a sample, especially at z >2.5 where the AGN can contribute up to 60% of the flux at observed frame 250μm
- The mean SFRs of the optical QSOs show a positive trend with AGN luminosity. However, we show that this trend is dominated by redshift and stellar mass dependencies.
- For redshifts of $0.2 < z < 0.5$, there is suggestive evidence that the mean SFRs at high AGN luminosities are elevated by a factor of ~1.5 compared to those of main sequence galaxies.
- We combine the results of our optical QSO sample to lower AGN luminosity X-ray selected AGN from Stanley et al. (2015), and find that the two samples show consistent mean SFRs.
- Based on stellar masses estimated from the available BH masses, the mean SFRs of the optical QSOs are consistent or higher than those of main sequence galaxies at redshifts up to $z \sim 1.5$. However, at the highest redshift range ($1.5 < z < 2.5$) of this study their mean SFRs values drop to below the main sequence, but are still consistent within the scatter. A possible interpretation is that this offset is driven by an evolution in the M_{BH}–M_* relation that we did not apply for our stellar mass estimates.
- The radio-luminous QSOs show consistent results to the overall optical QSO sample, and are consistent with the main sequence of star-forming galaxies within the scatter.

4.6 APPENDIX:

SED fits for all bins

In this Appendix section we present the best-fit SEDs for all bins in our sample. In Figure 5.7.1 we show the best-fits of each bin for our full QSO sample, with IDs that correspond to those of Table 4.1. In Figure 4.6.2 we show the best-fits for our radio-luminous sub-sample, with IDs that correspond to those of Table 4.2.

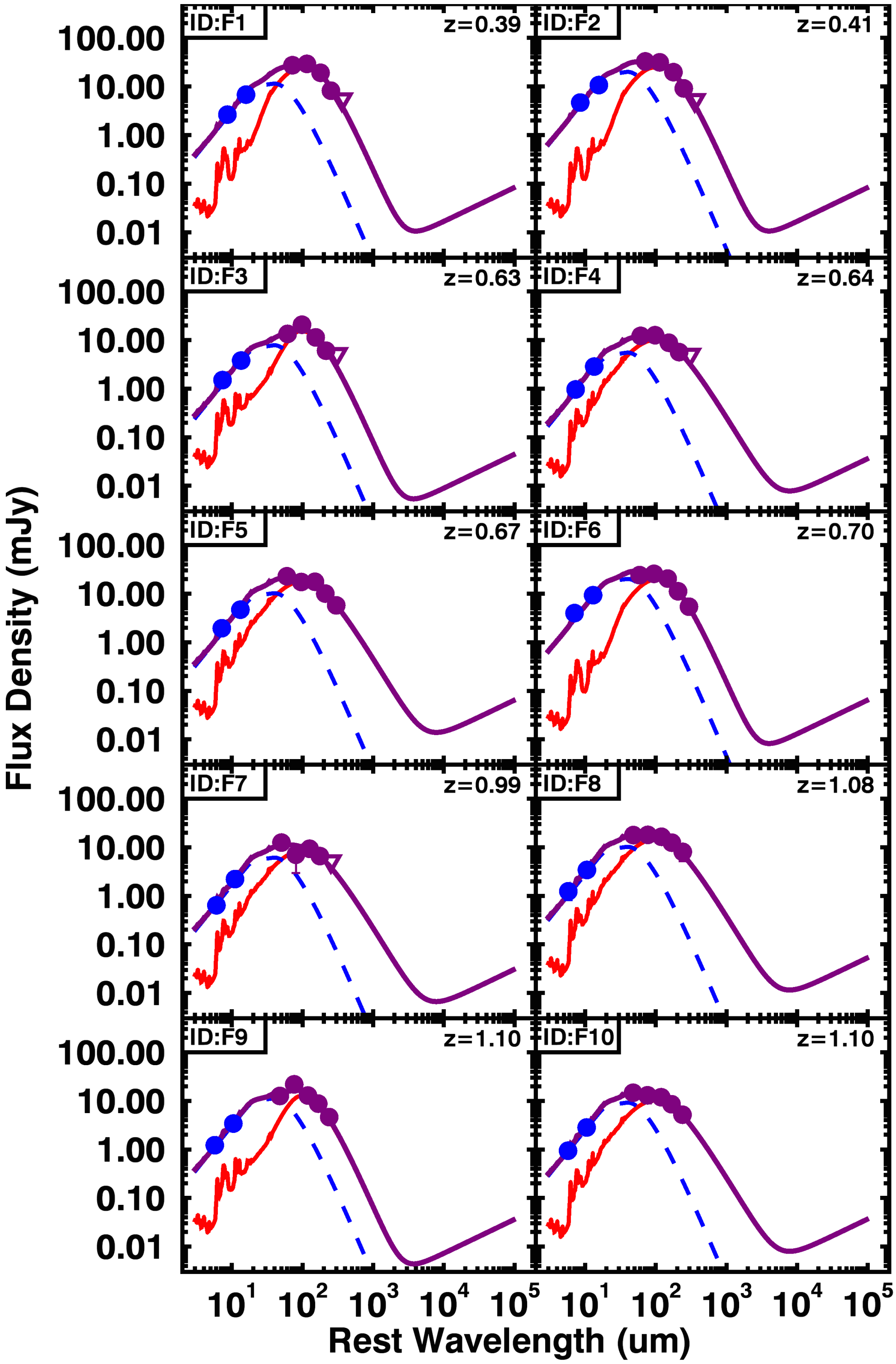


Figure 4.6.1: The best-fit SEDs for all the z–L_{AGN} bins of the QSO sample. The ID name corresponds to that of Table 4.1 for direct reference.

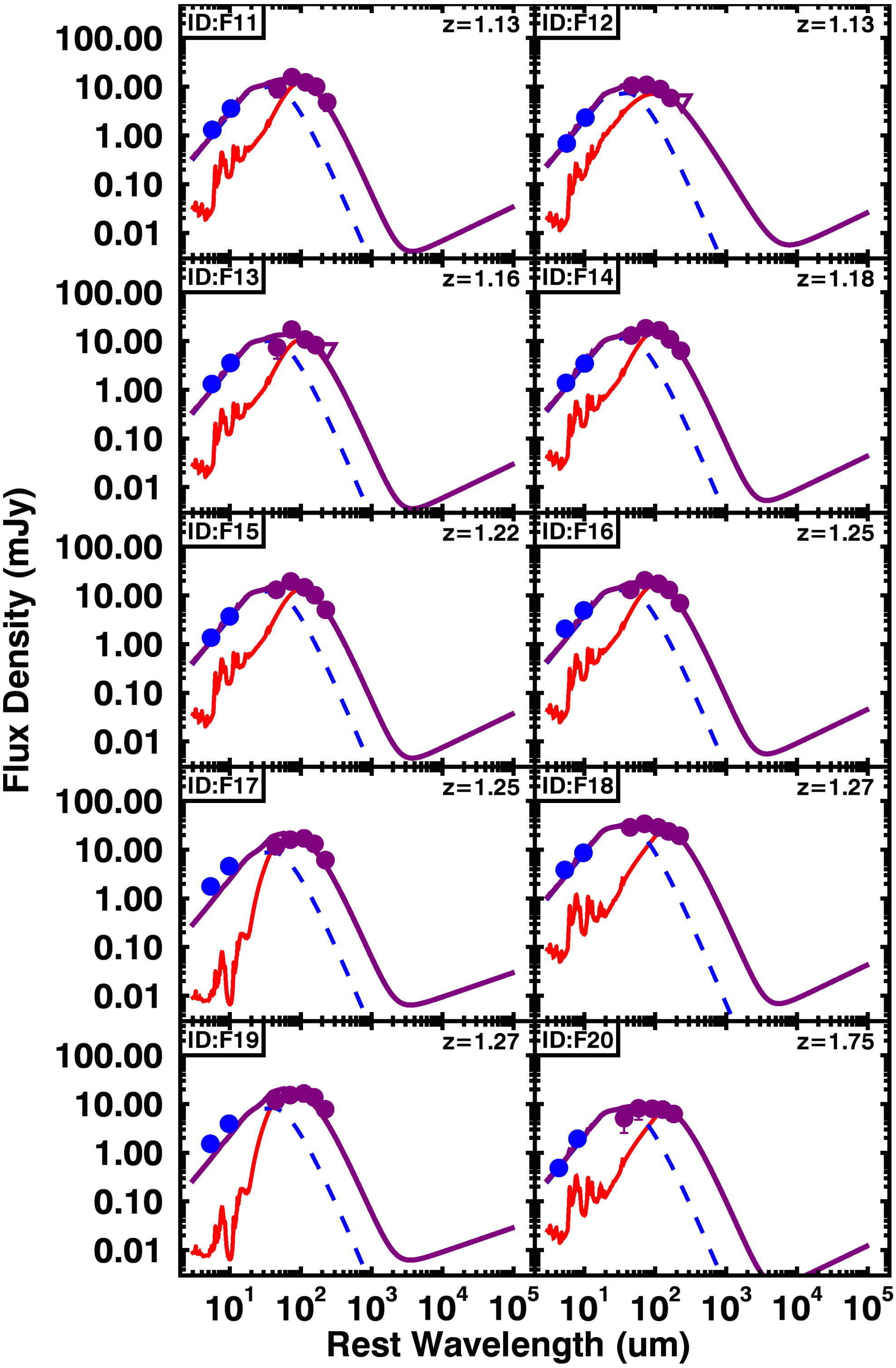


Figure 4.6.1: Continued

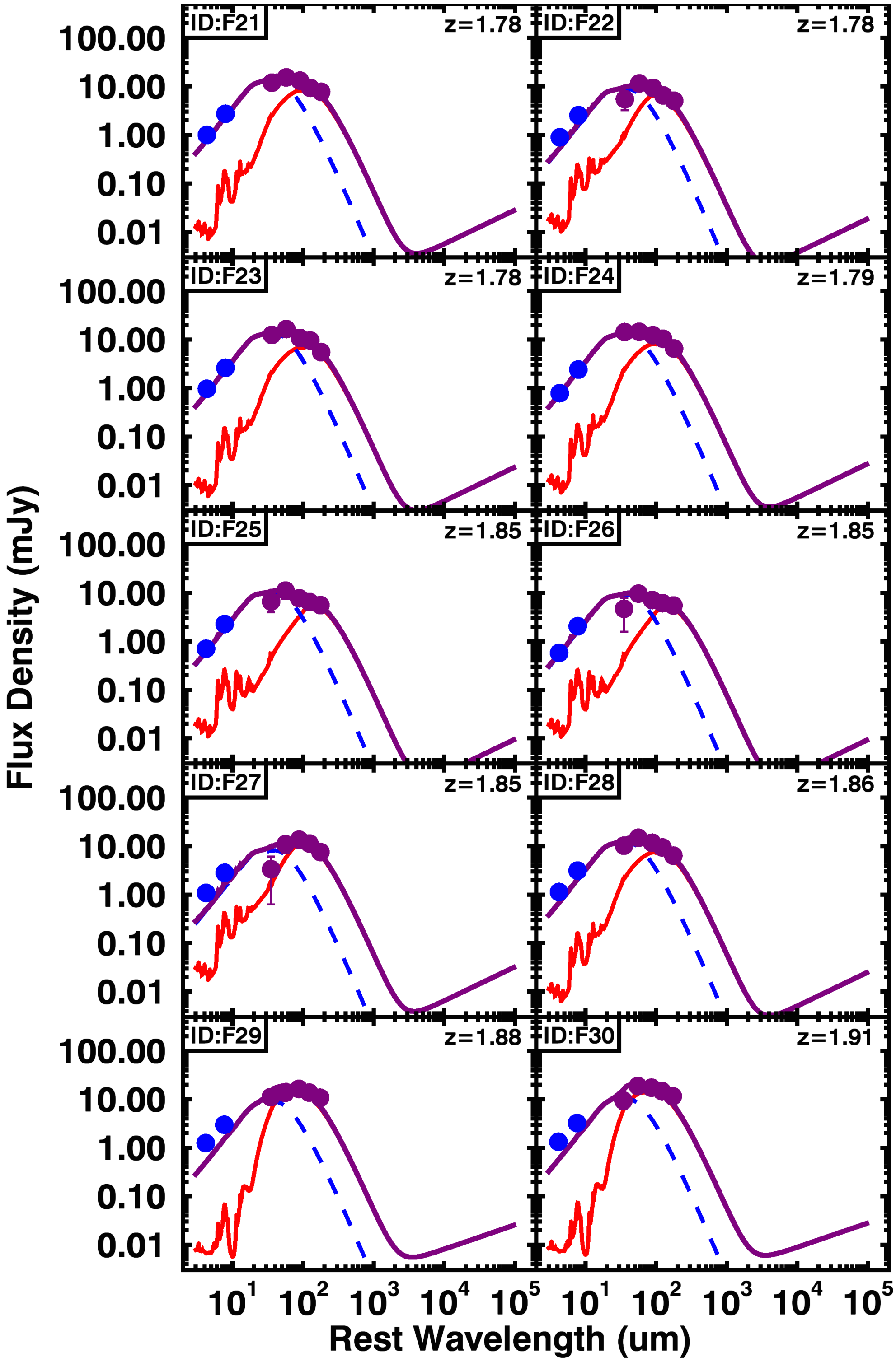


Figure 4.6.1: Continued

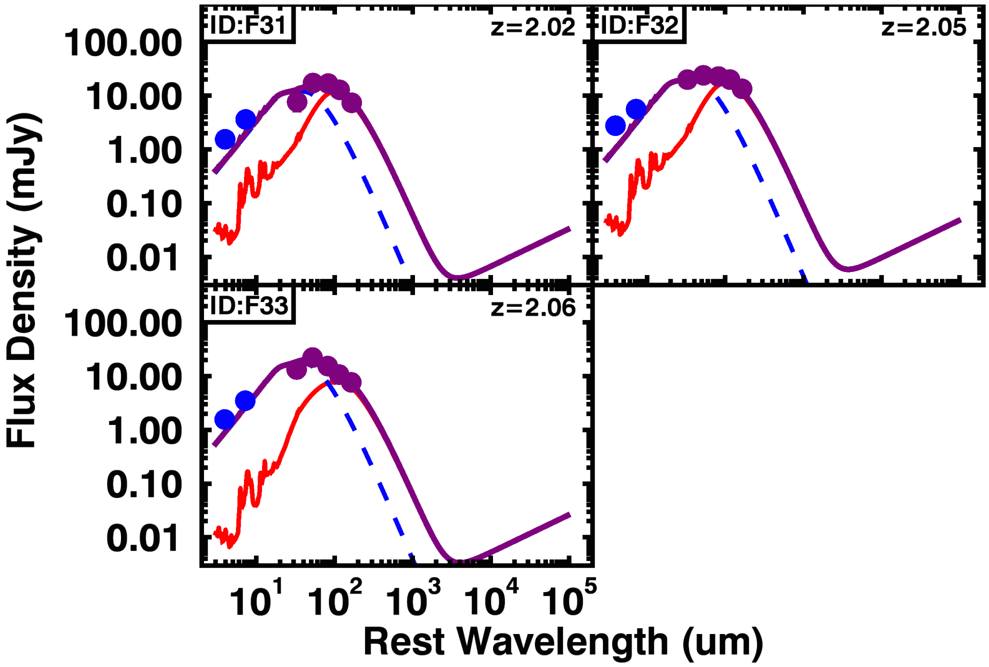


Figure 4.6.1: Continued

The radial light profile of SPIRE stacked sources

An additional cause for uncertainty in the SPIRE stacked flux density estimates is the possible boosting due to nearby sources. QSOs are well known for their clustering (e.g., White et al. 2012 and references therein), and in Wang et al. (2015) it was found that due to clustering of other dusty star-forming galaxies around optical QSOs there is a $\sim$ 8–13% contamination to the 250μm–500μm flux density, respectively. To take this possible source of contamination into account, we take the average flux density in annuli, and fit the flux as a function of radius from the center to a radius of $\sim$ 150". We use the SPIRE PSF (provided by H-ATLAS) convolved with itself, which corresponds to the images we are using, and a constant free to vary (see last panel in Fig. 4.6.3). [2] The factor of contamination calculated for each bin shows no dependency on redshift and AGN luminosity, and has a median of $\sim$11% at 250μm, 24% at 350μm, and 14% at

---

[2] To define the amount of contamination from nearby sources, we originally used a combination of the convolved PSF and a power-law of fixed slope. Due to the quality of the data we can not place a strong hold on the slope of power-law. For this reason we fitted with different fixed power-law slopes and chose to use the one with the lowest χ^2 values, which corresponds to a zero slope.

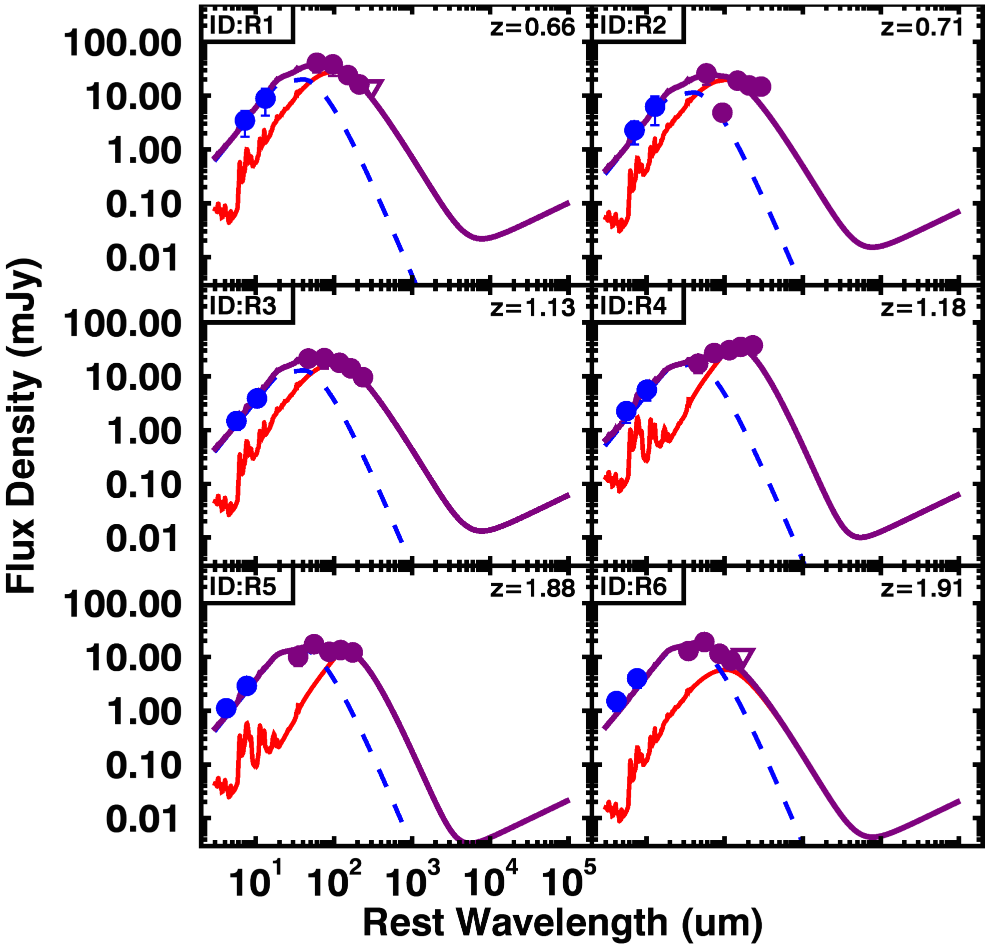


Figure 4.6.2: The best-fit SEDs for all the z–$L_{1.4GHz}$ bins of the radio-luminous QSO sample. The IDs correspond to those of Table 4.2 for direct reference.

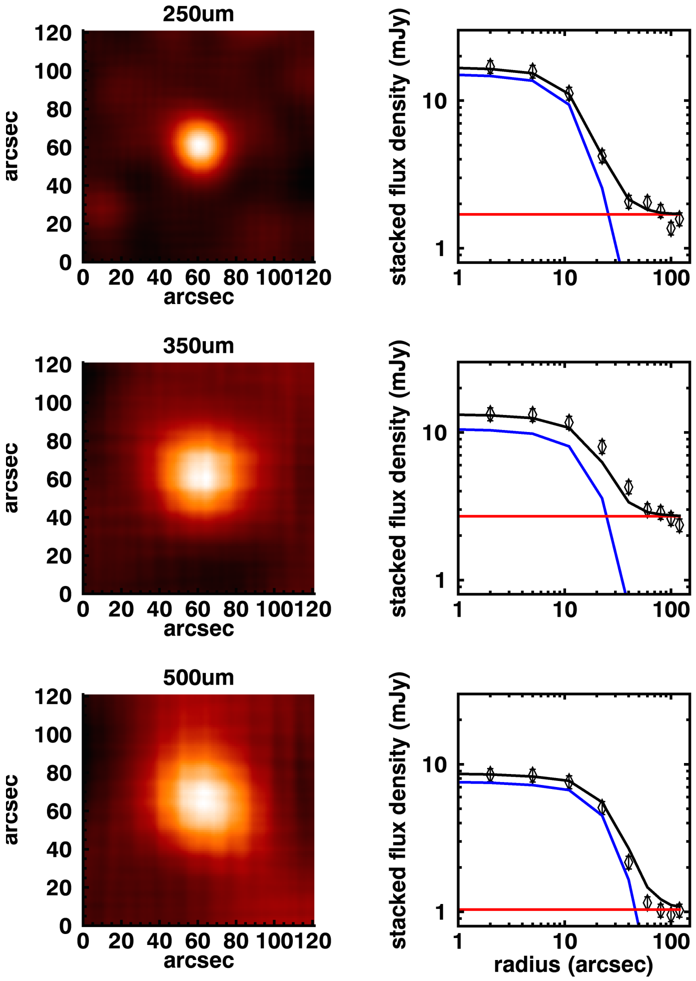


Figure 4.6.3: Examples of our radial light profile analysis for the three SPIRE bands. First shown are the stacked images in 250μm, 350μm and 500μm. In the second panel we show examples of the radial light profile of the stacked image used to estimate the contamination from bright neighbouring sources. The light profile is fitted with the convolved PSF (blue) for each band respectively, and a constant (red) fitted to the high end tail for the estimation of the contamination factor to the stacked flux density.

500μm. However, the absolute values of the contamination factor are equivalent to the offset we see in the random stack distribution.

CHAPTER 5

Achieving deeper constraints on the SFRs of X-ray AGN with ALMA 870μm observations

Abstract

In this chapter we study the SFRs of a sample of 109 X-ray AGN with moderate to high luminosities of $L_{2-8\rm keV} = 10^{42} - 10^{45}$ erg s^{-1}, at redshifts $z > 1$. The sample studied here is a sub-sample of 109 X-ray AGN from Stanley et al. (2015) (Chapter 3) that have been observed with deep ALMA Band-7 continuum observations. With the addition of the deep ALMA photometry to that of MIR and FIR photometry from *Spitzer* and *Herschel*, and using IR SED fitting, we can now place constraints on the SFRs up to $\sim$10 times lower than previously possible. This results in 35% of our observed sample now having a measured SFR from originally only having upper limits, and 67% of the sample now having upper limits improved by over a factor of 2 compared to previous constraints. Additionally, with the improved constraints on the star-forming emission of the galaxy achieved with the ALMA photometry, we can now identify a MIR AGN counterpart in 55% of our sample, where we could not previously. Finally, we explore the ($F_{870\mu\rm m}/F_{24\mu\rm m}$)–redshift space as as tool for the identification of MIR emitting AGN, and find that it can successfully retrieve all sources that had a MIR AGN counterpart identified through our SED fitting process. Furthermore, we test this space on two different complementary samples observed with ALMA 870μm, one of an extreme obscured AGN sample, and one of Submm bright star-forming galaxies, and find that for both samples it can successfully retrieve the presence of an AGN counterpart.

5.1 Introduction

As described in detail in Chapter 3, studies of the mean SFRs of moderate to powerful AGN, have consistently shown that on average AGN live in star-forming galaxies (e.g., Lutz et al. 2010; Shao et al. 2010; Mullaney et al. 2012a; Harrison et al. 2012; Rosario et al. 2012; Rosario et al. 2013b; Azadi et al. 2015). Furthermore, in Chapter 3 we investigated the trends between the mean SFRs and the X-ray AGN luminosity, for a sample of $\sim$2000 X-ray AGN, and found no strong evidence of a correlation or anti-correlation, with the mean SFRs showing flat trends as a function of AGN luminosity. However, these studies have only concentrated on the mean SFRs.

A major limitation of comparing the mean properties of a sample is the fact that even though two distributions can have the same mean, the distributions themselves can be significantly different (i.e., one can be broader than the other but centered at the same mean, or one could be of a more bimodal shape while the other has a Gaussian shape but the mean is still the same). A great advance in understanding the star-forming properties of X-ray AGN would therefore come from measuring their SFR distributions. Until now, a major restriction in accurately measuring the distributions of SFRs of high redshift ($z >$ 0.2) AGN samples, has been the limitations on the sensitivity of the available photometry. *Herschel* has provided the deepest field-survey observations in the FIR at 70–500μm, but even so the available surveys only detect the bright end of the galaxy population. In Figure 2.4 of Chapter 2 we show the SFRs corresponding to typical flux limits of the deepest Herschel surveys for all bands as a function of redshift. For $z \gtrsim$1 we can only detect star-bursting and/or massive star-forming galaxies. Therefore, we need even deeper observations in the FIR/Submm to directly constrain the SFRs of the typical population of galaxies and AGN at redshifts of $z \gtrsim$1.

In the past few years we have been able to use the new millimeter and submillimetre telescope, the ground-based interferometer Atacama Large Millimeter/submillimeter Array (ALMA). The ALMA observatory can achieve sensitivities down to 0.1mJy/beam and high resolution can go down to <0.1" (depending on the observational set-up), achieving more than a magnitude lower sensitivities than the *Herschel* bands at 100s of times better resolution. With the sensitivities achieved by ALMA it is now possible to easily detect and resolve galaxies and AGN at redshifts above $z \sim$1 at lower fluxes than that possible with previous FIR/Submm observatories, and place more accurate constraints on the SFRs of fainter galaxies and AGN and construct their full SFR distributions. This approach has been demonstrated previously in Mullaney et al. (2015), where they showed that with ALMA photometry it is possible to distinguish differences between the distribution of the SFRs of AGN in comparison to that of the overall population of star-forming galaxies.

With a sample of 24 X-ray AGN with deep ALMA 870μm observations, that covers the redshift range of $1.5 < z < 4$ and moderate X-ray AGN luminosities of $L_{2-8\mathrm{keV}} = 10^{42}$–$10^{44}$ erg s^{-1}, Mullaney et al. (2015) place constraints on the differences of the X-ray AGN sample to the overall star-forming galaxy population. The SFRs of the overall star-forming galaxy population can be described by the main sequence of star-forming galaxies, which is defined by the dependency of the SFR on the stellar mass and redshift of galaxies (e.g., Noeske et al. 2007; Elbaz et al. 2011; Schreiber et al. 2015; Speagle et al. 2014). By taking the distribution of the offset of the ALMA derived SFRs of the X-ray AGN to those of the main sequence of star-forming galaxies of the same stellar mass and redshift, they find that X-ray AGN show a different distribution to that of the main sequence star-forming galaxies. Even though the study suffers from having a limited number of sources, its results highlight the importance of constraining the distribution of SFRs rather than just the mean. In the study presented here, we build on the sample of Mullaney et al. (2015) with the observation of a larger sample of AGN covering higher luminosity AGN ($L_{2-8\mathrm{keV}} > 10^{44}$ erg s^{-1}).

In this chapter we present a sample of 109 X-ray AGN observed with ALMA in Band-7, covering the redshifts of $1 < z < 4.7$ and X-ray luminosities of $10^{42} < L_{2-8\mathrm{keV}} \leq 10^{45}$ erg s^{-1}. With combined photometry covering 8–870μm, we demonstrate the improvements possible with the addition of ALMA photometry, compared to the original constraints based on 8–500μm photometry, both in constraining the SFRs and in identifying the MIR emission of AGN. For simplicity, throughout this chapter we will refer to constraints on the IR SED from 8–500μm photometry as *Herschel* constraints, and from 8–870μm photometry as ALMA constraints, because it is the difference between the achieved sensitivities of the two observatories that influence the estimated SFRs. In Section 5.2 we present the sample used for this study, and give information on the ALMA observations. In Section 5.3 we present the IR SED fitting method. In Section 5.4 we present two ALMA observed samples used to fully test our methods. Finally, in Section 5.5 we demonstrate the improvements on constraining the SFRs and identifying the AGN component of the IR SEDs that ALMA provides, and in Section 5.6 we give a summary of our results.

5.2 Sample & Observations

In this chapter we present a sample of X-ray selected AGN that have been observed in two ALMA Band-7 programs during Cycle 1 and Cycle 2. The Cycle 1 sample was selected to have $L_{2-8\mathrm{keV}} > 10^{42}$ erg s^{-1} at redshifts of $1.5 < z < 3.2$ (see Mullaney et al. 2015; Harrison et al. 2016). The Cycle 2 sample was selected to uniformly sample the $L_{2-8\mathrm{keV}}$–z plane covering the redshifts of $1.5 < z < 3.2$ and $10^{43} < L_{2-8\mathrm{keV}} \lesssim 10^{45}$ erg s^{-1}. The sample was restricted within the areas covered by the *Herschel* observational programs PEP/GOODS-*H* (Lutz et al. 2011; Elbaz et al. 2011) and HerMES (Oliver et al. 2012) in the fields of GOODS-S, and COSMOS, that are our main sources of the FIR photometry covering the wavelengths of 70 – 500μm. In both ALMA programs the targeted sources were chosen to have insufficient *Herschel* photometry (i.e., undetected in most bands and with high upper limits) to successfully constrain the IR SED and decompose it to the star-forming and AGN components. Consequently, our sample consists of mostly *Herschel*, and sometimes *Spitzer*, undetected sources with poor SFR constraints. The selected sample of X-ray AGN is a sub-sample of that presented in Chapter 3. We make use of the photometric counterparts assigned to the X-ray AGN in that work for our analysis, in combination to the ALMA observations at 870μm.

Since the original definition of the samples for the ALMA programs, there have been new redshift catalogues of the CDF-S and C-COSMOS from Hsu et al. (2014) and Marchesi et al. (2016) respectively. We make use of the updated redshifts from the updated catalogues, and a number of sources no longer lie within the parameter space of the original selection, with 89% lying within, 3.7% lying below, and 7.3% lying above the original redshift selection.

In this chapter we analyse all of the X-ray AGN that were observed, including serendipitous detections within the primary beam. We restrain the sample to only sources with

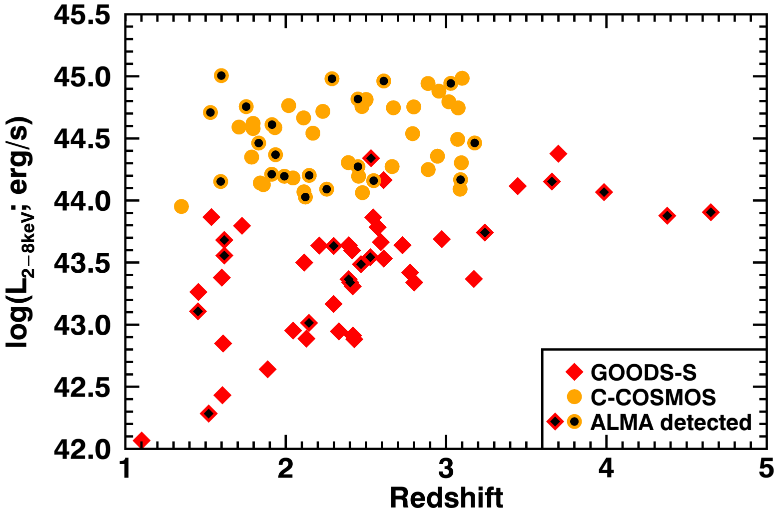


Figure 5.2.1: X-ray hard band (HB; 2–8keV) luminosity (L_{2-8keV}) as a function of redshift. Here we plot the 109 X-ray AGN observed with ALMA, including 101 originally targeted and 8 serendipitous detections. We highlight the ALMA detected sources with black centers.

$z > 1$, resulting in 109 X-ray AGN with ALMA observations, 101 originally targeted, and 8 serendipitously detected X-ray AGN. In Fig. 5.2.1 we plot the L_{2-8keV} as a function of redshift for the sample studied here, with updated redshift values, and highlight the ALMA Band-7 detected sources. In total there are 5 sources with $z < 1$ covered by the ALMA program, all in the field of GOODS-S, that are not included in the analysis of this chapter.

5.2.1 ALMA 870um observations

The sample of 109 X-ray AGN were observed during Cycle 1 and Cycle 2 with a bandwidth of 7.5GHz centered at 351GHz, with 55 sources in CDF-S and 54 sources in C-COSMOS.

The data were processed and imaged following the methods of Simpson et al. (2015); also see Harrison et al. (2016). We used the COMMON ASTRONOMY SOFTWARE APLICATION (CASA; version 4.4.0; McMullin et al 2007), and the CLEAN routine provided within CASA. The raw data was calibrated using the ALMA data reduction pipeline. The results were visually inspected, and when deemed necessary, the pipeline calibration process was repeated with additional data flagging. We cleaned the images by first creating "dirty" images. We then identify the sources with SNR$\geq$5, which we mask, and then repeat the cleaning process down to 1.5σ. We measure the noise in the resulting cleaned images, and repeat the above process around the sources with SNR$\geq$4. Finally, for the final cleaned images, we applied natural weighting and a Gaussian taper. The synthesized beams are of the size of $(0.8" - 0.9") \times 0.7"$, with noise levels of 0.1–0.8 mJy/beam in CDF-S [1], and 0.08–0.23 mJy/beam in C-COSMOS.

From the product images, Scholtz et al. (in prep) created catalogues of targeted sources and serendipitous detections. From the maps we extracted all the peaks of at least 2.5σ and matched to the optical counterparts from Hsu et al. (2014) and Marchesi et al. (2016) catalogues. To estimate the probability of spurious peaks being matched to our sources, we created negative maps (negatives of the original maps), from which we estimate the density of noise peaks at different SNR values. The surface density of noise peaks matched to the X-ray objects for SNR bins of 2.5 – 3, 3 – 4, and >4 are 0.01, 0.027, and 6×10^{-4} objects per arcsec^2, corresponding to 2.41, 0.89 and 0.052 overall spurious matches respectively, for the search radius of 0.5". For peaks of SNR $>$ 4 we increased the matching radius to 1" (corresponding to 0.24 expected spurious objects). Finally, the catalogue includes all targeted sources including spurious detections. We use an SNR>2.5 limit to define detected sources; however we note that the majority of our detected sources have SNR>3 ($\sim$73%; Scholtz et al. in prep). If a source remains undetected we chose to take a more conservative limit of 3$\times$RMS as the flux limit. In total we find that 40/109 (36.7%) of our sources are detected by ALMA.

5.2.2 MIR and FIR photometry

For our SED fitting analysis, we exploit available photometry in the wavelength range of 3.6 – 500μm, provided by observations carried out by: *Spitzer*-IRAC at 3.6–8μm; *Spitzer*-IRS at 16μm; *Spitzer*-MIPS at 24μm; *Herschel*-PACS at 70, 100, 160μm; and *Herschel*-SPIRE at 250, 350, 500μm, in addition to the ALMA photometry outlined above.

---

[1] The large noise levels of 0.8mJy/beam correspond to a sub-sample of targets that were observed at higher resolution than that requested (i.e, 0.3" instead of 1" resolution). Therefore, for these observations the images had to be heavily tapered to a resolution of $\sim$0.8", resulting in increased noise levels.

We make use of the catalogue of MIR-FIR counterparts of X-ray AGN created in Chapter 3. These were created based on the optical positions of X-ray AGN in CDF-S and C-COSMOS, with the advantage of priored FIR deblended catalogues; see Chapter 3.

The deblended catalogues consist of the MIPS 24μm, PACS 70μm, 100μm and 160μm photometric catalogues from Magnelli et al. (2013)[2], and SPIRE 250μm, 350μm, and 500μm photometric catalogues from Swinbank et al. (2014).

For the MIR bands we make use of the catalogues of *Spitzer*–IRAC 3.6–8μm observations as described in Damen et al. (2011), and Sanders et al. (2007), for GOODS-S, and COSMOS, respectively, as well as *Spitzer*–IRS 16μm from Teplitz et al. (2011) for GOODS-S. Since all the IRAC catalogues have their detections determined by the 3.6μm maps, and the 16μm catalogues have been produced with the use of 3.6μm priors, they are all consistent with the deblended PACS and SPIRE catalogues described above.

5.3 Method: IR SED fitting

We performed two sets of SED fitting following the method of Chapter 3, using photometry at 8–870μm. The first set includes only the star-forming galaxy templates in the fit, while the second set includes both the AGN and star-forming components. We fit to the photometric flux density detections, but also force the fits to not exceed any of the photometric flux density upper limits[3]. This procedure results in twelve fitted SEDs to chose from. We calculate the integrated 8–1000μm IR luminosity due to star-formation from the host galaxy ($L_{IR,SF}$) and due to the AGN ($L_{IR,AGN}$). To determine the best fitting solution of the twelve fitted SEDs, we use the Bayesian Information Criteria (BIC; Schwarz 1978) which allows the objective comparison of different non-nested models with a fixed data set (see section 2.3.2). The SED fit with the minimum BIC value is defined as the best fit; however, to establish if the SED of the source requires an AGN component the SED with the AGN component has to have a smaller BIC to that of the SED with no AGN com-

---

[2] The PACS catalogues for and GOODS-S are published in Magnelli et al. (2013). The catalogue for COSMOS was created in the same way and is available online (http://www.mpe.mpg.de/ir/Research/PEP/DR1).

[3] Note that this is different to the method in Chapter 3, where the fits were not forced to not exceed the upper limits, but rather the fits exceeding the limits where excluded in the picking of the best-fit.

ponent with a difference of ΔBIC>2. We outline five different cases of our SED fitting process below.

1. If we have more than two photometric detections and at least one is at the FIR range of the rest-frame SED (i.e. at wavelength greater than ~80μm), we then chose the fit with the minimum BIC value as our best fit. If multiple fits have the same value as the minimum BIC then we take the mean $L_{IR,SF}$, and $L_{IR,AGN}$ of those fits (e.g., Fig. 5.3.2(a)).

2. If a source is only detected in the MIPS–24μm and ALMA–870μm band, we use the comparative BIC values to decide if the IR SED requires the AGN component or not. We take the mean $L_{IR,SF}$, and $L_{IR,AGN}$ for the set of fits that best describe the SED (e.g. Fig. 5.3.2(b)).

3. If a source is only detected in the ALMA 870μm band we normalise the star-forming galaxy templates to the ALMA photometry and take the mean of the resulting $L_{IR,SF}$ for the full template region. We are confident that if the AGN was significantly contributing to the ALMA photometry, it would have been detected in the MIR at the depth of the MIPS-24μm photometry. Based on the shape of the AGN IR SED, if the AGN was right at the limit of the 24μm flux density (0.06mJy) it would emit $\sim6\times10^{-4}$–0.1mJy at 870μm from redshift 1 to 4.7 respectively. The highest contribution at redshift 4.7 only accounts for ~6% of the source emission at that redshift. For example, at a redshift of $z\sim3$ for the AGN to produce a flux density of 0.4mJy at 870μm the emission at 24μm would have to be ~4mJy, while the limit on the 24μm flux density is 0.06mJy; i.e., a 24μm undetected AGN at $z\sim3$ would contribute less than 2% to the total 870μm flux density (see Fig. 5.3.2(d)).

4. If a source has only MIR detections, or no detections at all, then we constrain an upper limit on the star-forming component using the limits and/or the 3σ error on the detections. We normalise all star-forming templates to the lowest value of the upper limits, and 3σ above the photometry if the source is detected in a given band. We then take the maximum $L_{IR,SF}$ of the range of normalised templates, as the upper limit. The same is done for the estimation of the upper limit on the IR luminosity of AGN component ($L_{IR,AGN}$).

5. If a source is detected in the MIR and the limit on the star-forming component is $>5\sigma$ below the observed frame 8–24μm photometry, then we can identify the presence of the AGN component. We find that in these cases we can measure the strength of the AGN component, and the $L_{IR,AGN}$ (e.g., Fig. 5.3.2(c))

Following this method, we have performed individual SED fitting for the whole sample of X-ray AGN studied here. To quantify the improvement on constraining the IR SED and SFRs achieved with the deep ALMA photometry, we fit our sample of X-ray AGN with and without the inclusion of the ALMA photometry.

5.4 Comparison samples

We make use of two other samples that have been observed with ALMA at 870μm, to test the effectiveness of our SED fitting approach and the ($F_{870\mu m}/F_{24\mu m}$)-redshift space (see section 5.5.2) in accurately identifying the AGN and star-forming components.

We chose two extreme samples representative of powerful star-forming galaxies, and powerful AGN dominating the IR emission.

The first comparison sample is of highly obscured AGN. The sample consists of radio-powerful obscured AGN taken from Lonsdale et al. (2015), covering the redshifts $0.47 < z < 2.85$, and selected to have ultra-red WISE colours (i.e., lying significantly redward to the main WISE population in the (W1-W2) vs (W2-W3) colour space; where W1 corresponds to 3.4μm, W2 to 4.6μm, W3 to 12μm, and W4 to 22μm). Samples of sources selected to be the reddest sources in the WISE colour plane, have been revealed to be an IR-luminous population of high redshift obscured AGN (e.g., Eisenhardt et al. 2012; Bridge et al. 2013; Jones et al. 2014; Tsai et al. 2015) with IR luminosities likely dominated by the AGN emission (e.g., Jones et al. 2015). In total we use 41 (out of 49 sources) of this sample that were targeted with ALMA 870μm, with complementary WISE photometry, and 3 sources also have *Herschel* photometry, constrained to redshifts $z > 1$. The ALMA observations for this sample have a resolution of 0.5–1.2”, and reach noise levels of 0.3-0.6mJy/beam.

The second comparison sample consists of Submm bright galaxies (SMGs). SMGs represent the highly star-forming population of high redshifts of $z \sim 2$–3 (e.g., Blain et al.

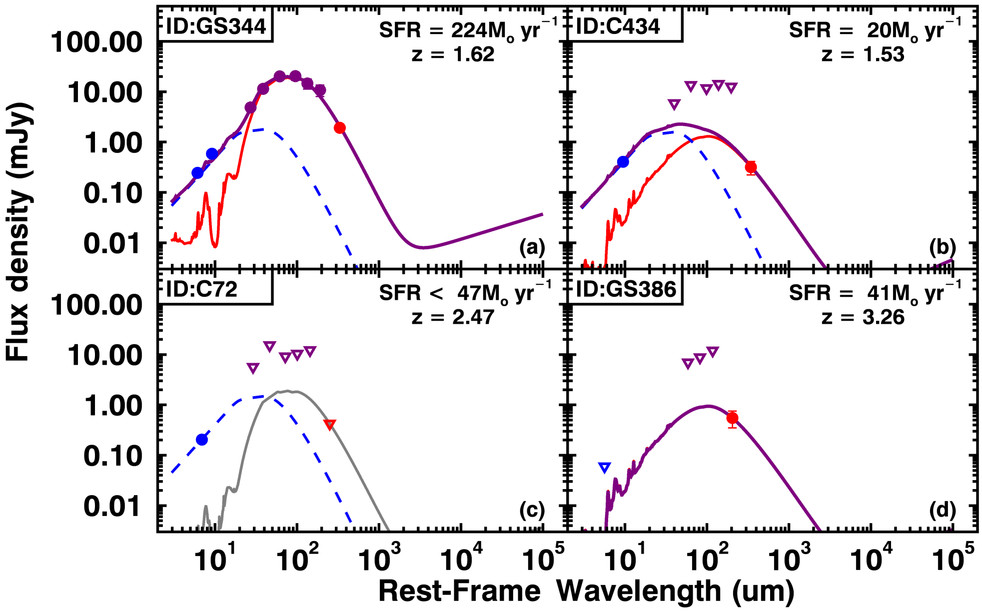


Figure 5.3.2: Examples of four different cases of SED fitting results. In all cases, the blue dashed curve is the AGN component, while the red solid curve is the star-forming component. The total combined SED is shown as a purple solid curve. The grey curves correspond to an upper limit constraint on the SF component. The photometry is colour-coded, with blue corresponding to *Spitzer*, purple to *Herschel* bands, and red to the ALMA photometry. The inverse triangles are upper limits on the flux density. (a) An example where we have full photometric coverage of the SED. In this case the ALMA photometry on the SED provides confidence in the SED templates used for our analysis (see case 1 in section 5.3). (b) An example where the source is only detected in MIPS-24μm and ALMA-870μm (see case 2 in section 5.3), and (c) an example of an ALMA undetected source that is only detected in the MIR (see case 5 in section 5.3). In both cases of (b) and (c) the deep ALMA photometry, allow us to constrain the star-forming component to a level that reveals the presence of an AGN component in the MIR . (d) An example were the source is undetected in all bands except for ALMA-870μm (see case 3 in section 5.3). In the last case we are confident all the emission is due to star-formation, as if there was any significant contribution from the AGN the source should be detected in the MIR, which is not the case.

2002; Wardlow et al. 2011; Casey et al. 2013), with typical IR luminosities of $L_{IR} \sim 10^{46}$ erg s^{-1} (e.g., Swinbank et al. 2014) dominated by the IR emission. The chosen sample of SMGs is taken from the ALMA-LESS survey (A-LESS; Hodge et al. 2013; Karim et al. 2013), including 122 sources over the redshift range of $0.4 < z < 7$ observed with ALMA 870μm during Cycle 0, with redshifts from Simpson et al. (2014), and MIR and FIR photometry from *Spitzer*-MIPS and *Herschel* (Swinbank et al. 2014). In total we use 113 sources of the sample constrained to redshifts of $1< z <5$ (covering a similar redshift range as our sample of X-ray AGN). For the ALMA observations of this sample the FWHM of the primary beam was 17.3”, and reach typical noise levels of 0.4–0.5 mJy/beam.

We follow the same SED fitting method on these two samples as for our X-ray AGN sample, using the available published photometry, and compare the results to our sample of X-ray AGN in section 5.5.2.

Overall, with our SED fitting procedure we have a $L_{IR,SF}$ measurement for 21/41 (51%) of the highly-obscured AGN with the rest having a well constrained upper limit. We identify an AGN component in 40/41 (97.6%) of the highly-obscured AGN with a minimum AGN contribution to the total IR luminosity of 50%, and with 22/41 (53.7%) of the sample having an AGN component that contributes $\gtrsim$90% of the IR luminosity. When looking at the SMG sample, our SED fitting process can constrain a $L_{IR,SF}$ measurement for the whole sample, and finds that all of the sample have IR emission dominated by the star-formation, with only 12/113 (10.6%) of the sources having an IR AGN counterpart. The IR luminosities due to star formation of these comparison samples cover the range of 6×10^{44}–10^{47} erg s^{-1} and 4×10^{44}–3.4×10^{46} erg s^{-1} for the obscured AGN and SMG samples, respectively.

5.5 Results & Discussion

In this section we present our results from the SED fitting of our sample, and explore the improvements on constraining the IR SEDs of AGN that ALMA photometry can provide. In §5.5.1 we present the new constraints on the $L_{IR,SF}$ and demonstrate the improvements in comparison to *Herschel* constraints. In §5.5.2 we demonstrate the improved ability to identify MIR emission from the AGN in our SED fitting method when combined with ALMA photometry, and explore the MIR-Submm colour of $F_{870\mu m}/F_{24\mu m}$ as an identifier of MIR AGN.

5.5.1 Improvements on $L_{IR,SF}$ constraints

To demonstrate how much better we can constrain the SFRs of our sample once we have ALMA photometry, we have performed the same SED fitting methods on the sample with and without the ALMA photometry. Here we show a comparison of the two sets of results on $L_{IR,SF}$.

In Figure 5.5.3 we show the $L_{IR,SF}$ when constrained using 8 – 500μm photometry (purple) and 8 – 870μm (red) photometry (i.e., without and with the ALMA photometry), as a function of redshift. For comparison we also plot the track for the mean $L_{IR,SF}$ of star-forming galaxies on the main sequence (e.g, Schreiber et al. 2015) for stellar masses of $M_* = 10^{11}$ $M_\odot$ (the median stellar mass for our sample; Scholz et al. in prep). As the sample was selected to be *Herschel* faint and undetected, the majority of the sources (107/109; 98%) originally only have upper limit constraints on their $L_{IR,SF}$ values. The ALMA photometry allows us to both measure the $\langle L_{IR,SF} \rangle$ of sources not possible previously (35% more measurements), and to also push the limits on $L_{IR,SF}$ values to significantly lower levels, up to a factor of 10 (see Fig. 5.5.4, and 5.5.3). For sources with sufficient *Herschel* constraints to measure $L_{IR,SF}$ (2/109) we find a mean change in $L_{IR,SF}$ when including the ALMA photometry of a factor of $\sim$1.3. The agreement of the ALMA photometry to the *Herschel* constraints provides extra confidence in our SED fitting approach and choice of templates, even in the absence of ALMA photometry. In total we now have $L_{IR,SF}$ measurements for 40/109 (36.7%) of the sources, and upper limit constraint for the remaining 69/109 (63.3%).

With the new constraints on the $L_{IR,SF}$ values, we find that sources that originally had upper limits above the expectations of a main sequence of star-forming galaxy of $M_* = 10^{11}$ $M_\odot$ are now significantly lower, with the majority of the observed sources lying a factor of $\sim$2–4 below.

In Figure 5.5.4 we show a histogram of the improvement on the upper limits of $L_{IR,SF}$ of our observed sources. The value plotted is the difference between the original $L_{IR,SF}$

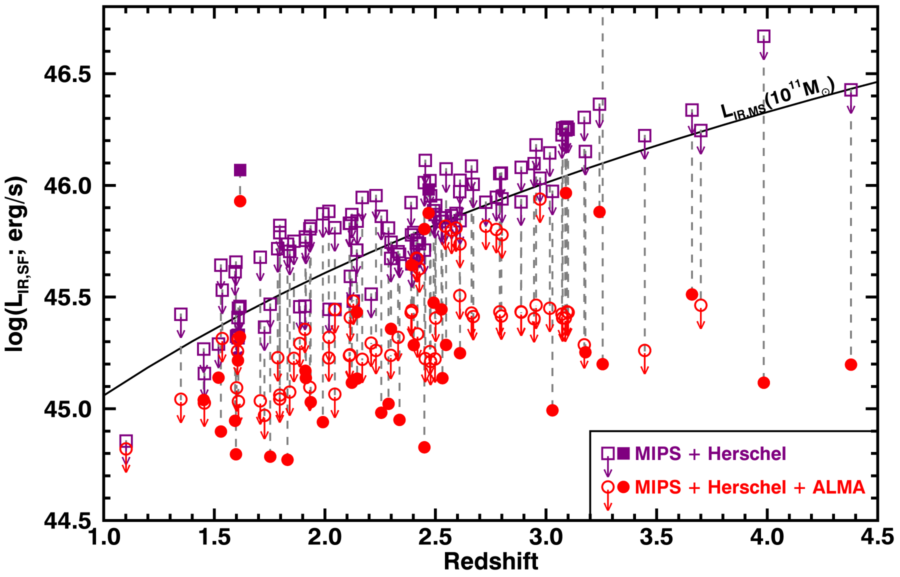


Figure 5.5.3: IR luminosity due to star formation ($L_{IR,SF}$) as a function of redshift for our sample before (purple) and after (red) the inclusion of deep ALMA photometry in our SED fitting. We find that for 73/109 (67%) of the sources the constraints of both measurements and upper limits on $L_{IR,SF}$ have decreased by more than a factor of 2, and 38/109 (35%) that originally only had upper limits now have measurements of $L_{IR,SF}$. We find a tendency for sources previously undetected by *Herschel* to have significantly lower $L_{IR,SF}$ values to those sources with sufficient *Herschel* constraints on the $L_{IR,SF}$ that populate similar and higher $L_{IR,SF}$ values to those of pre-ALMA upper limits (purple data points), indicating a non uniform coverage of the $L_{IR,SF}$ plane.

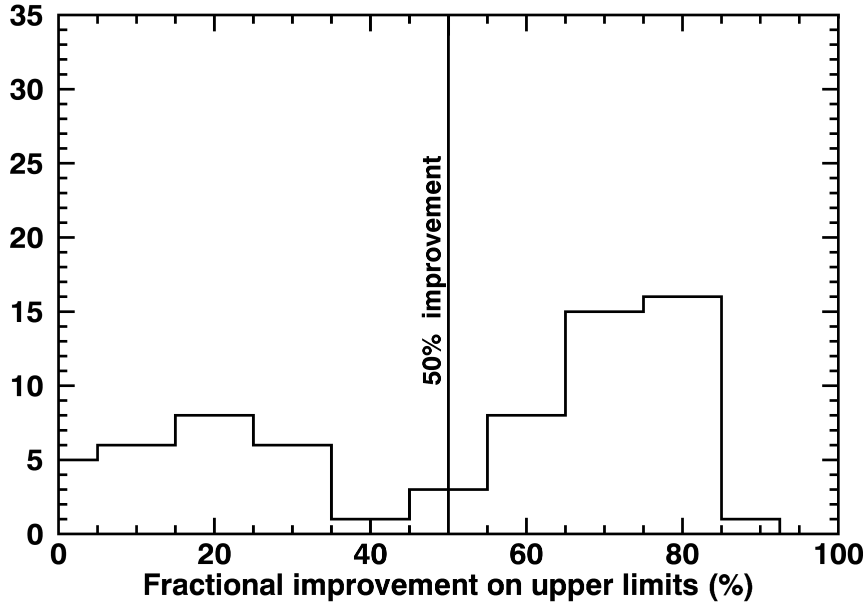


Figure 5.5.4: The histogram of the improvement in the constraints of the $L_{IR,SF}$ upper limits relative to their original constraints ($L_{IR,SF}{}^{orig}$), i.e. $(L_{IR,SF}{}^{orig} - L_{IR,SF}{}^{new}) / L_{IR,SF}{}^{orig}$. The majority of the upper limits in our sample (43/69; 62%) have new $L_{IR,SF}$ upper limits that have decreased by more than 50%.

values ($L_{IR,SF}{}^{orig}$) and the new $L_{IR,SF}$ values ($L_{IR,SF}{}^{new}$) over $L_{IR,SF}{}^{orig}$. It is immediately clear that more than half of our sample (67%) have SFR constraints that have changed by more than 50%. The apparent bi-modality in the improvements of the upper limit constraints is driven by the range of RMS values for our observations. For the subsample of sources observed with high resolution the resulting RMS of the ALMA maps is as high as 0.8mJy/beam, which results in a small improvement on the constraints of the $L_{IR,SF}$ upper limits.

We note that a full analysis of the individual SFRs and SFR distributions compared to the main sequence of star forming galaxies will be presented in Sholtz et al. (in prep) (discussed in Chapter 6).

5.5.2 The AGN IR emission

With the excellent constraints on the star formation component that the ALMA observations can provide, we are now able to better constrain the MIR emission of the AGN itself.

The shape of the star-forming IR SED, in combination with the constraints placed on it with the ALMA 870μm photometry, allows for the detection of a MIR excess, even when a source is undetected at 870μm. Indeed, where we could not identify a MIR AGN counterpart in any of our sources, we can now confidently identify a MIR AGN component in 54/109 (49.5%) of the ALMA observed sample, with AGN fractions down to 20% of the total IR (8–1000μm) luminosity. In table 5.1 we show the number of sources in our X-ray AGN sample, and two comparison samples described below, with different AGN fractions of the IR luminosity.

As many of our sources only have a 24μm and 870μm photometry, and the fact that in combination to our SED fitting, we can still successfully disentangle the two components of AGN and star formation emission, it is likely that the Submm–MIR colour can be used as an AGN identifier. For this reason we explore the ratio of flux densities at 870μm and 24μm ($F_{870\mu m}/F_{24\mu m}$) as a function of redshift for three samples, the X-ray AGN sample of this study, a highly obscured AGN sample, and a highly star-forming galaxy sample. To define the regions of the ($F_{870\mu m}/F_{24\mu m}$)-redshift plane dominated by purely AGN emission and by purely star-forming emission we use the star-forming templates of our SED fitting procedure, and the AGN templates of Mullaney et al. (2011), the mean of which is used in our SED fitting procedure. For means of comparison and for further exploring the ($F_{870\mu m}/F_{24\mu m}$)-redshift plane, we also include an additional two AGN templates, and an additional set of SF templates. We use the set of star-forming templates from Dale & Helou (2002) produced by a phenomenological model of star-forming galaxies, and the AGN templates of Mor & Netzer (2012) and Symeonidis et al. (2016) derived for samples of luminous QSOs. While the two sets of star-forming galaxy templates cover a similar range of parameter space, the AGN templates cover slightly different regions of the ($F_{870\mu m}/F_{24\mu m}$)-redshift plane, and so they can also be tested on this plane when compared to AGN dominated systems.

In Fig. 5.5.5 we plot the ($F_{870\mu m}/F_{24\mu m}$)-redshift plane, for our sample of X-ray AGN. We show the region of the parameter space covered by our star-forming templates (our set of templates in pink; and in grey the set of templates from Dale & Helou 2002), and the region covered by AGN templates (from Mullaney et al. 2011 in pink; Mor & Netzer 2012 in grey; and Symeonidis et al. 2016 in blue). We highlight the presence of

AGN fraction	X-ray AGN	highly-obscured AGN	A-LESS SMGs
<20%	55	1	103
$\gtrsim$20%	4		2
$\gtrsim$30%	16		3
$\gtrsim$50%	30	18	5
$\gtrsim$90%	4	22	
full sample	109	41	113

Table 5.1: The number of sources at different fractions of AGN contribution to the IR luminosity for three samples observed with ALMA at 870μm. Our X-ray AGN sample covers the full range of AGN strength in the IR, while the sample of highly-obscured AGN (Lonsdale et al. 2015) show strong AGN emission contributing $\gtrsim$50% of the IR luminosity. The additional comparison sample of SMGs (Swinbank et al. 2014), shows the expected low AGN contribution to the IR emission.

an AGN component in the best-fit result of each source, and the fraction of the total IR luminosity emitted by the AGN. As we would expect, our sample of X-ray AGN covers a wide range in the ($F_{870\mu m}/F_{24\mu m}$)-redshift space, with increasing AGN percentage when moving from the star-forming dominated region to the AGN dominated region. Overall, the ($F_{870\mu m}/F_{24\mu m}$)-redshift space shows a clear separation between the AGN and star-formation dominated regions, and successfully places sources in these regions. We also see a good separation amongst the sources at different AGN fractions. This agreement is not surprising of course, as we use the same templates used in our SED fitting analysis to also define the star formation and AGN dominated parameter space. However, we demonstrate below using samples of AGN and star formation dominated sources, that the adopted AGN and SF templates are indeed suitable.

To further test the ability of the ($F_{870\mu m}/F_{24\mu m}$)-redshift space to identify AGN and star formation, we use the two comparison samples described in section 5.4. Our comparison samples consist of the Lonsdale et al. (2015) sample of MIR and radio powerful AGN also observed with ALMA, and a sample of powerful star-forming SMGs from the A-LESS survey (e.g., Swinbank et al. 2014), described in section 5.4. All of these sources have been fitted following the SED fitting procedure outlined in section 5.3. From the resulting fits we can determine if the best-fitting solutions require an AGN, and the fraction of the IR emission due to the AGN. We plot the ($F_{870\mu m}/F_{24\mu m}$)-redshift space with the comparison samples in Figure 5.5.6, and highlight the sources where our SED fitting finds

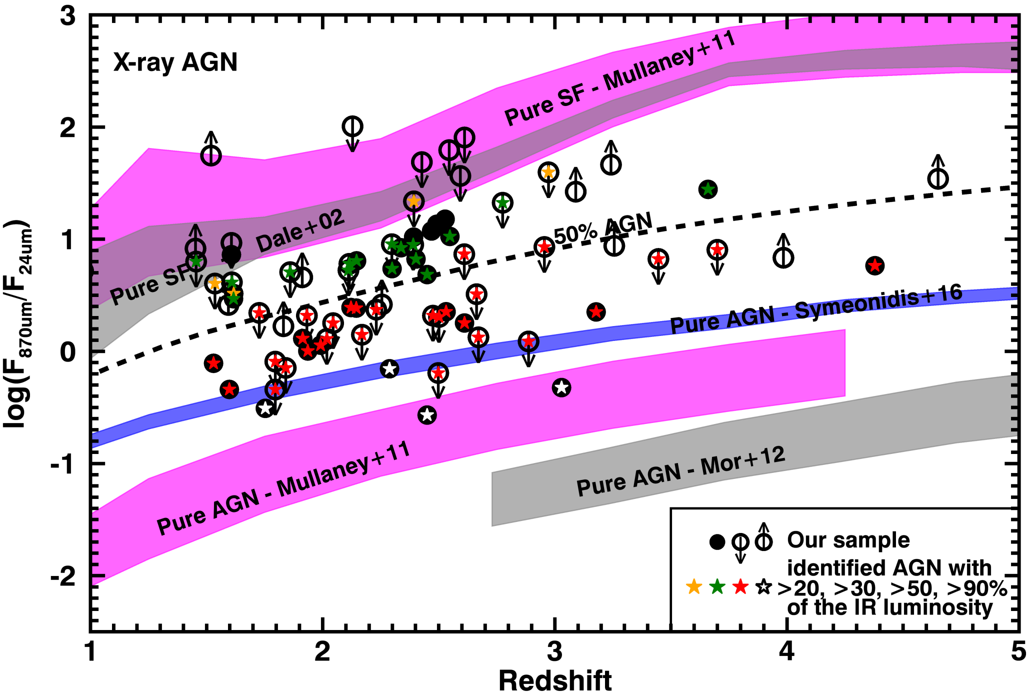


Figure 5.5.5: The $F_{870\mu m}/F_{24\mu m}$ ratio as a function of redshift for the ALMA observed sample of X-ray AGN. Sources for which our SED fitting finds an AGN component with more than 20% contribution to the IR emission, are highlighted by stars. The stars are colour-coded to the AGN fraction of the total IR emission. The star colours of yellow, green, red, and white, correspond to AGN fractions of $>20\%$, $\gtrsim 30\%$, $\gtrsim 50\%$, and $\gtrsim 90\%$ of the total IR emission respectively. We find that the $F_{870\mu m}/F_{24\mu m}$ ratio successfully recovers sources with an IR AGN counterpart.

an AGN with an indication of the fraction of the IR emission due to the AGN.

The ($F_{870\mu m}/F_{24\mu m}$)-redshift space successfully identifies all of the powerful AGN from Lonsdale et al. (2015). As this is a sample of AGN that dominate the IR emission, they are a good sample with which to test the accuracy of different AGN templates. We find that our AGN templates, are able to describe the majority of the AGN dominated systems with only 1/41 (2.4%) lying below our AGN colour space. The source below our AGN template is instead in the space of the Mor & Netzer (2012) templates that has a steeper long wavelength drop off in comparison to our AGN template. The colour range of the Symeonidis et al. (2016) AGN template fails to describe the full AGN sample, with 28/41 (68%) of the AGN dominated sources having lower $F_{870\mu m}/F_{24\mu m}$ fractions by an average factor of $\sim$2.

For the sample of SMGs the ($F_{870\mu m}/F_{24\mu m}$)-redshift space successfully places the majority of the sample within the space covered by our SF templates. The agreement between the colours of the SMGs and our templates is an additional indication for their suitability for our SED fitting analysis. Furthermore, the ($F_{870\mu m}/F_{24\mu m}$)-redshift space can still identify AGN counterparts, when present, for this extreme sample of star formation dominated galaxies. However, as is the case for all IR identification methods for AGN, it does miss out on sources with low AGN fractions, in galaxies with a much stronger star-forming component dominating the SED.

We have shown that the $F_{870\mu m}/F_{24\mu m}$ colour, and our SED fitting analysis, can identify even weak MIR AGN components. To understand how well our SED fitting approach, and the $F_{870\mu m}/F_{24\mu m}$ colour work in comparison to the most common IR AGN identification methods, we compare our results to MIR colour selection methods. We use the Donley et al. (2012) IRAC colour criteria for identifying MIR AGN. The Donley et al. (2012) criteria are the latest based on *Spitzer*–IRAC colours, and have the lowest contamination from non-AGN sources compared to previous IRAC selection criteria (e.g., Stern et al. 2005) We show this parameter space in Fig. 5.5.7 where we plot the 50 sources of our X-ray AGN sample detected in all four IRAC bands (46% of the sample). We find that with our SED fitting method we can successfully identify AGN (down to AGN fractions of 20%) in 23 out of 24 sources (95.8%) selected as AGN with the Donley et al. (2012) criteria. However, we do manage to recover more of the X-ray AGN sample than the

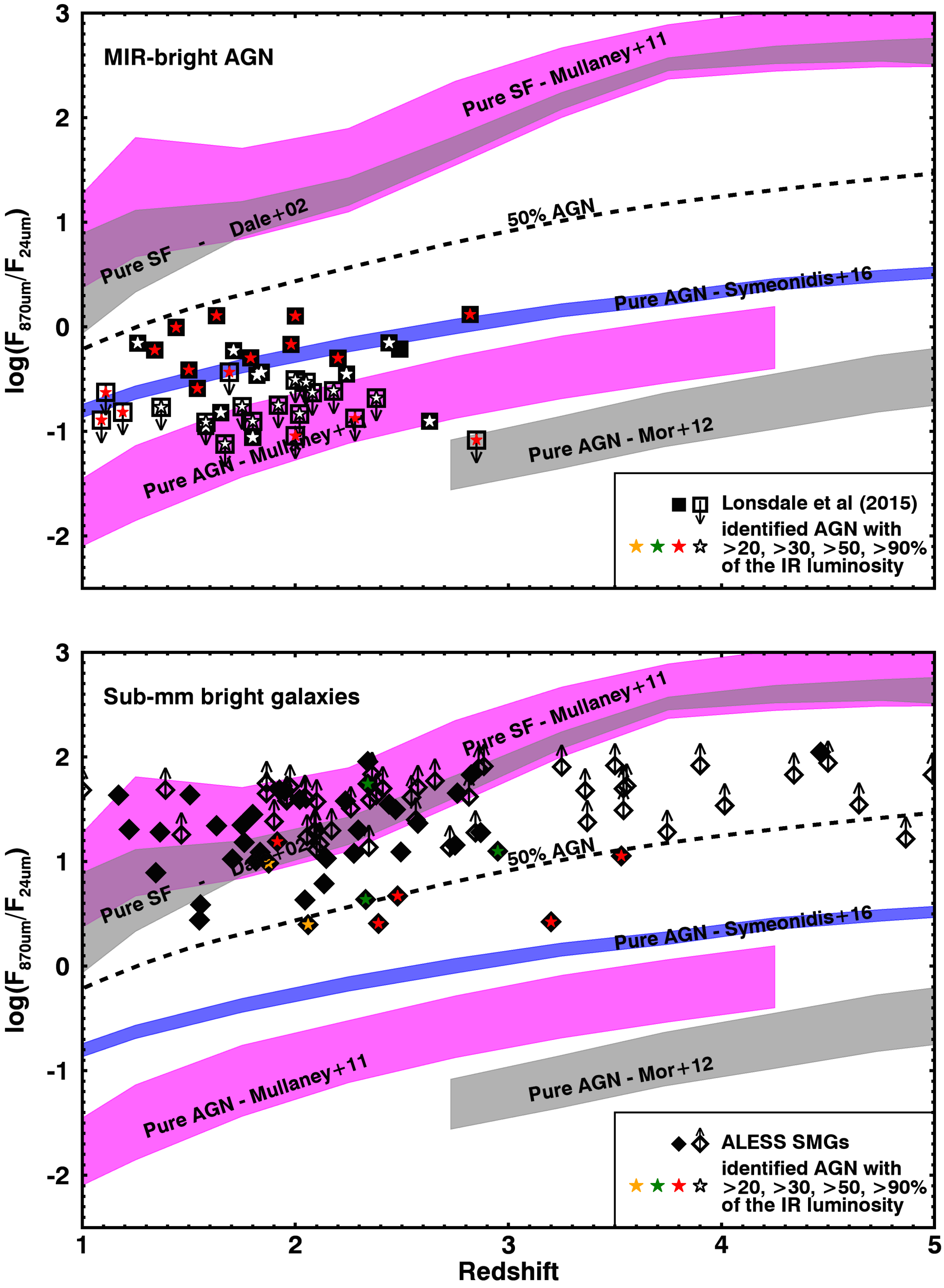


Figure 5.5.6: The $F_{870\mu m}/F_{24\mu m}$ ratio as a function of redshift for two different comparison samples observed with ALMA. Sources for which our SED fitting finds an AGN component with more than 20% contribution to the IR emission, are highlighted by stars, and colour-coded following Fig. 5.5.5. (*top*) The sample of extreme obscured, MIR and radio selected, AGN from Lonsdale et al (2015). (*bottom*) we show a sample of Submm sources from the A-LESS survey (Swinbank et al. 2014). We find that indeed the $F_{870\mu m}/F_{24\mu m}$ colour can successfully identify AGN with a wide range of AGN to star formation IR emission ratios.

AGN fraction	MIR selected	not selected
<20%	1	5
$\gtrsim$20%	1	10
$\gtrsim$30%	3	8
$\gtrsim$50%	14	3
$\gtrsim$90%	5	0
Total	24	26

Table 5.2: The number of X-ray AGN at different fractions of AGN contribution to the IR luminosity that would be selected and not selected by the IRAC colour selection criteria of Donley et al (2012). We look at the 50 sources of our sample detected in all four IRAC bands, and find that the Donley et al (2012) selection criteria successfully identify AGN contributing more than 50% of the IR emission, but misses the more moderate AGN (see also Fig. 5.5.7).

IRAC colour selection can, with 21 out of the 26 sources (80.7%) outside of the wedge with MIR AGN emission identified by our SED fitting method (see also Table 5.2).

We compare the properties of the sources in and out of the AGN wedge, and find that the sources in the wedge cover a similar range of redshifts as the ones outside the wedge; however, the sources in the wedge have overall higher $L_{2-8\text{keV}}$ values than the sources outside the wedge, with median values of 4.8×10^{44} erg s^{-1}, and 1.7×10^{44} respectively. The typical AGN fractions in the wedge are higher than those outside, with the majority of sources with AGN fractions of >50% lying within the wedge. We find that the IRAC colour criteria will successfully retrieve the majority (19/22; 86%) of MIR AGN with AGN fractions of >50%, and all (5/5) AGN dominated systems with AGN fractions of >90%. However, the IRAC colour criteria tend to miss out on AGN with smaller AGN fractions of 20–50%, finding only 4/22 (18%) sources (see Table 5.2).

5.6 Summary & Conclusions

We use deep ALMA observations to place constraints on the SFRs for a sample of 109 X-ray AGN that are faint or undetected in the *Herschel* bands. Our sample covers X-ray luminosities of $10^{42} < L_{2-8\text{keV}} < 10^{45}$ erg s^{-1} at redshifts of $z > 1$. 40/109 sources (36.7%) of our observed sample were detected at 870μm, but even though the majority are undetected the flux limit provided by ALMA is so low we can still place stronger

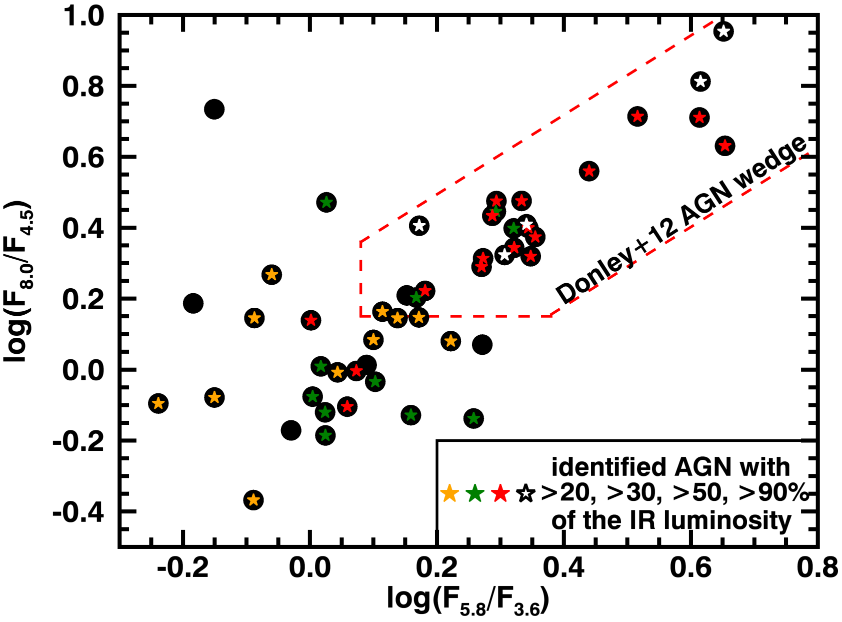


Figure 5.5.7: A colour-colour diagram based on IRAC photometry for our sample of X-ray AGN. The IRAC colour AGN selection criteria of Donley et al. (2012) are shown with the dashed lines. Only the 50 sources detected in all four IRAC bands are plotted (46% of the sample). Stars indicate that the best fitting solution from the SED-fitting requires an AGN and the colour-coding of the AGN fraction of the total IR emission is that of Fig. 5.5.5. We find that the $F_{870\mu m}/F_{24\mu m}$ colour retrieves a higher fraction of X-ray detected AGN than the MIR colour selection. However, the MIR selection is successful at retrieving all AGN dominated systems (i.e., AGN fraction of $\gtrsim$90%), as would be expected.

constraints on the SFR limit value than previously possible. We make use of the SED fitting methods of Stanley et al. (2015) (see Chapter 2) in combination with photometry at 8–870μm to fit and decompose the IR SED into AGN and star-forming components.

The results of our analyses are:

- We find that with the depths of our ALMA observations we can constrain SFR measurements and upper limits to up to a factor of $\sim$10 lower than previous constraints based on 8–500μm photometry. 38/109 (35%) of our observed sample now have a measured SFR from originally only having upper limits, and 73/109 (67%) of the total sample have new SFR constraints improved by over a factor of 2 compared to previously.

- With the excellent constraints at 870μm on the star-forming component of the IR SED, we are now able to also place stronger constraints on the IR emission of the AGN. Indeed, we can now identify an AGN component in 54/109 (49.5%) of our ALMA observed sample, with AGN fractions down to $\sim$20% of the total IR emission, where without the ALMA photometry we did not identify a MIR AGN component in any of the sources.

- We explore the parameter space of the flux density ratio of $F_{870\mu m}/F_{24\mu m}$ with redshift, and find that it can clearly identify the presence of MIR emission from the AGN. We suggest that this method could be developed as a tool for identifying AGN in future deep Submm and infrared surveys. To further test the $F_{870\mu m}/F_{24\mu m}$–redshift space we also use it on two different comparison samples representing the two extremes of AGN and star formation dominated IR emission. For both extremes the $F_{870\mu m}/F_{24\mu m}$–redshift space can successfully retrieve them as AGN dominated and star formation dominated systems respectively. Furthermore, we compare the parameter space covered by three sets of AGN templates, the AGN templates used in this thesis, those of Mor & Netzer (2012), and those of Symeonidis et al. (2016), to the $F_{870\mu m}/F_{24\mu m}$ of the AGN dominated sources. We find that our set of templates is suitable for the majority (97.7%) of the sample, while the Symeonidis et al. (2016) templates fail to describe the majority (68%) of the AGN dominated systems.

- We compare the number of AGN in our sample that would be MIR identified using the IRAC colour–colour AGN selection criteria of Donley et al. (2012), to that identified by our SED fitting analysis. For 50 sources in our sample, detected in all four IRAC bands, we find that 23 of the 24 sources identified by the colour criteria are also identified by our SED fitting, while in addition our SED fitting analysis retrieves another 21 out of the 26 sources that are not MIR AGN according to the colour criteria. In total our SED fitting analysis retrieves AGN in 44 out of the 50 sources (88%), while the IRAC colour criteria only retrieves AGN in 24/50 (48%). Furthermore, we find that the IRAC colour criteria will successfully retrieve all MIR AGN with a strong contribution to the total IR emission of the galaxy, but tends to miss out on AGN contributing <50%.

5.7 APPENDIX:

SED fits for our X-ray AGN sample

In this Appendix section we present the best-fit SEDs for all sources in our sample. In Figure 5.7.1 we show the best-fits of each source. With a blue dashed curve we show the AGN component, with a red solid curve we show the SF components, and with purple we show the full SED. In cases where we can only constrain an upper limit, the SF component is plotted with a grey curve. We have colour-coded the photometric data points, with blue corresponding to *Spitzer*, purple corresponding to *Herschel*, and with red corresponding to ALMA. With solid points we show the photometric detections, while with inverted triangles we show the photometric upper limits.

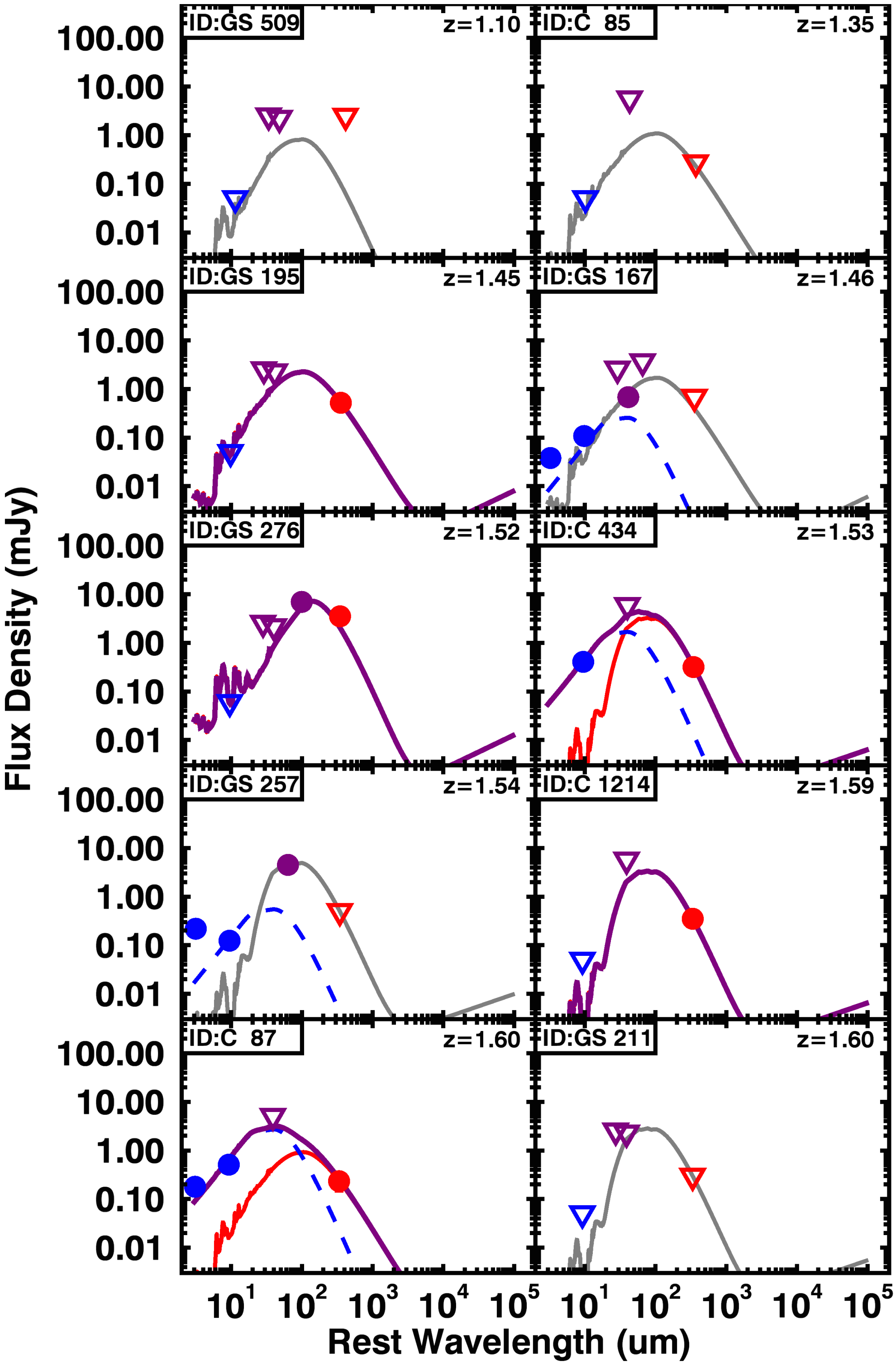


Figure 5.7.1: The best-fit SEDs for all the z–L_{AGN} bins of the QSO sample. the blue dashed curve is the AGN component, while the red solid curve is the star-forming component. The total combined SED is shown as a purple solid curve. The grey curves correspond to an upper limit constraint on the SF component. The photometry is colour-coded, with blue corresponding to *Spitzer*, purple to *Herschel* bands, and red to the ALMA photometry.

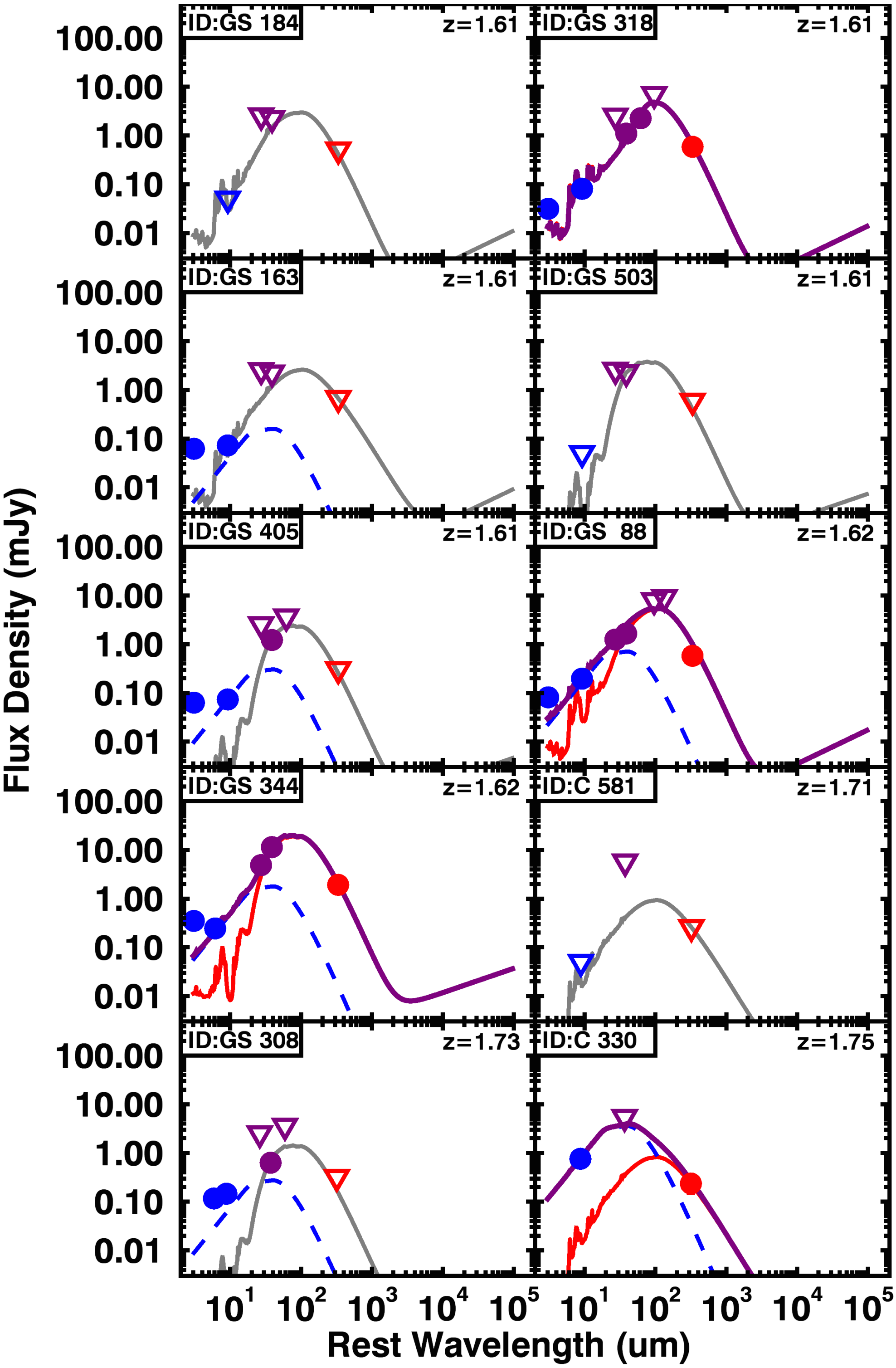


Figure 5.7.1: Continued

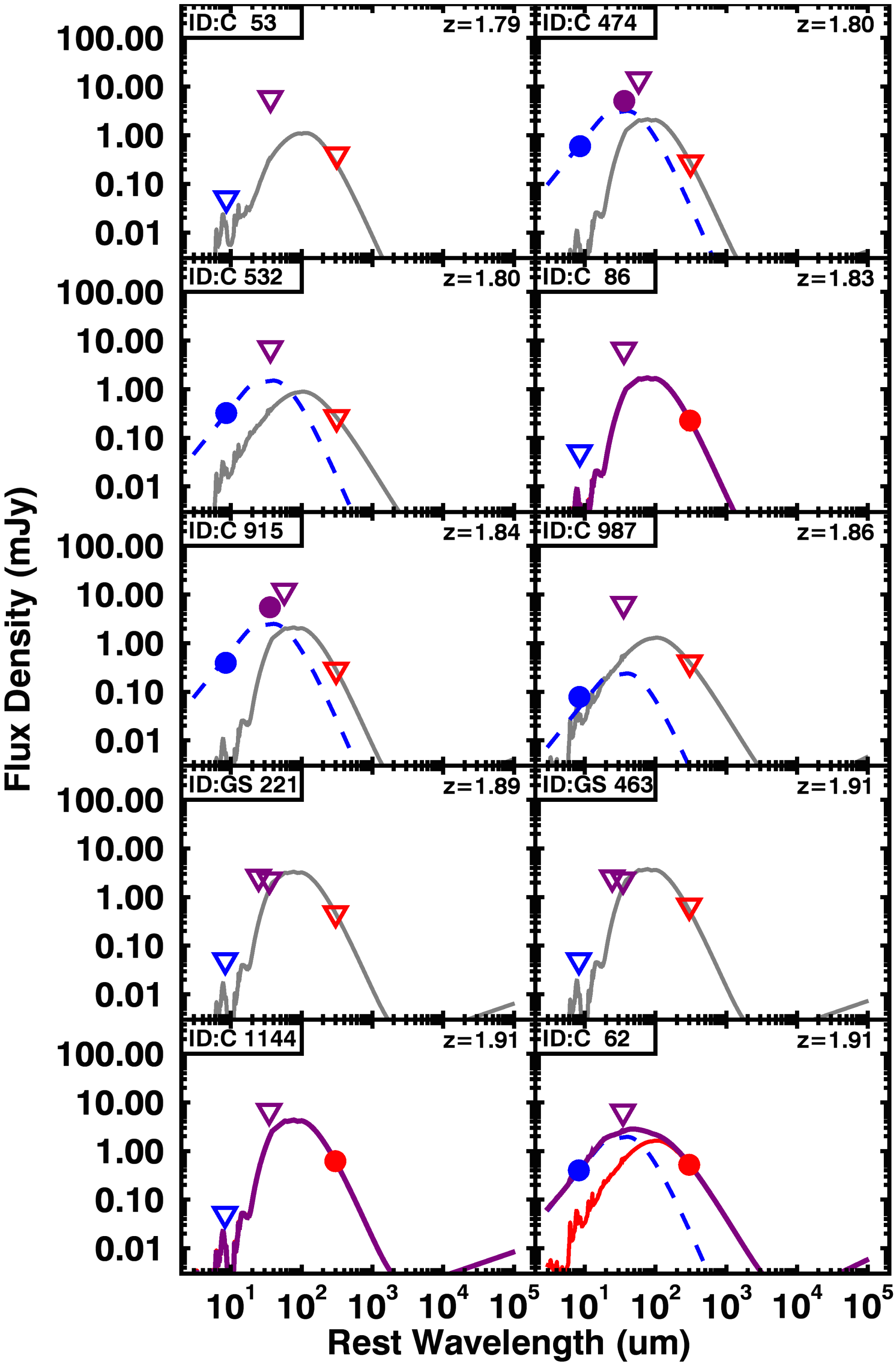


Figure 5.7.1: Continued

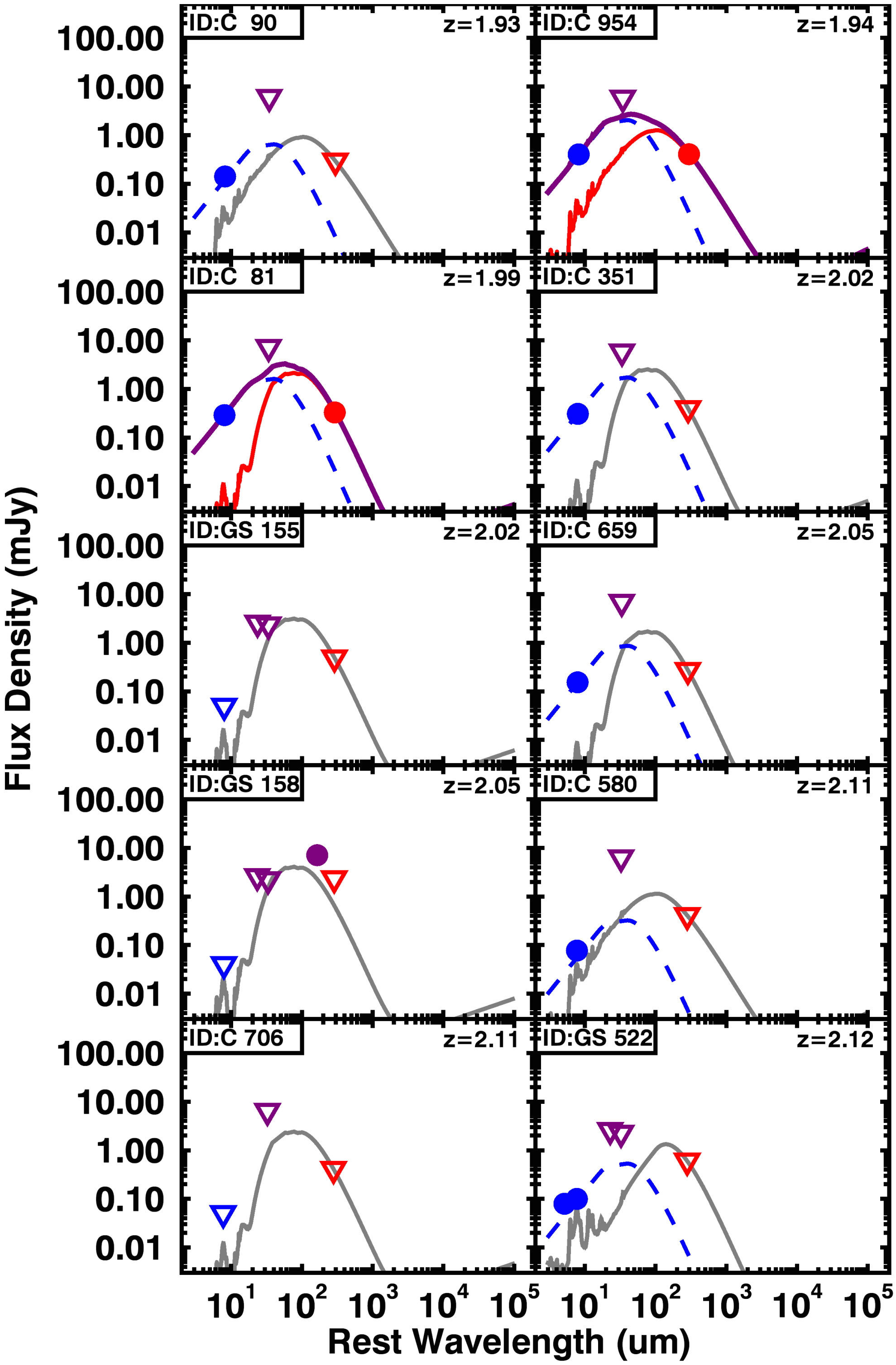


Figure 5.7.1: Continued

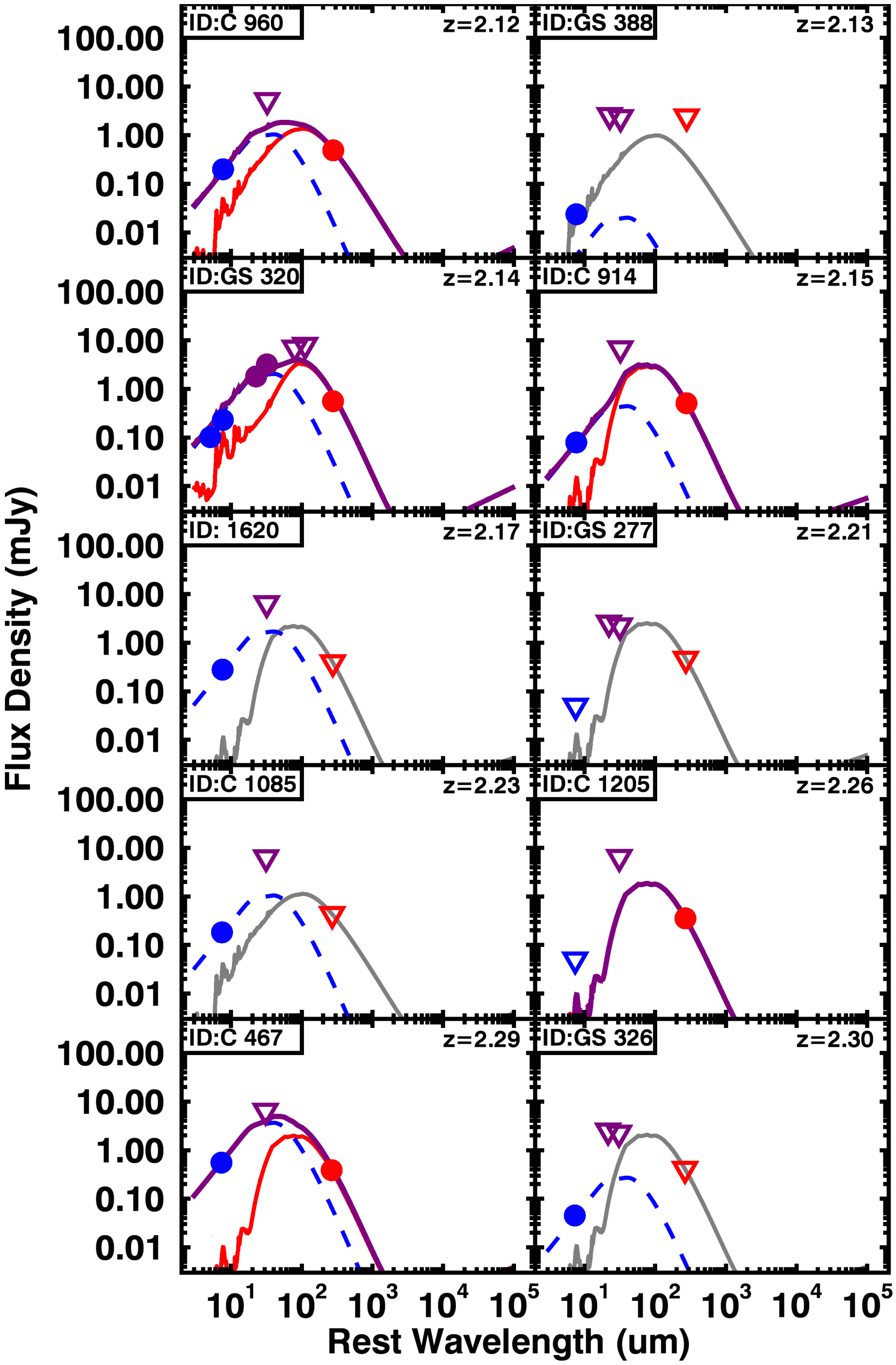


Figure 5.7.1: Continued

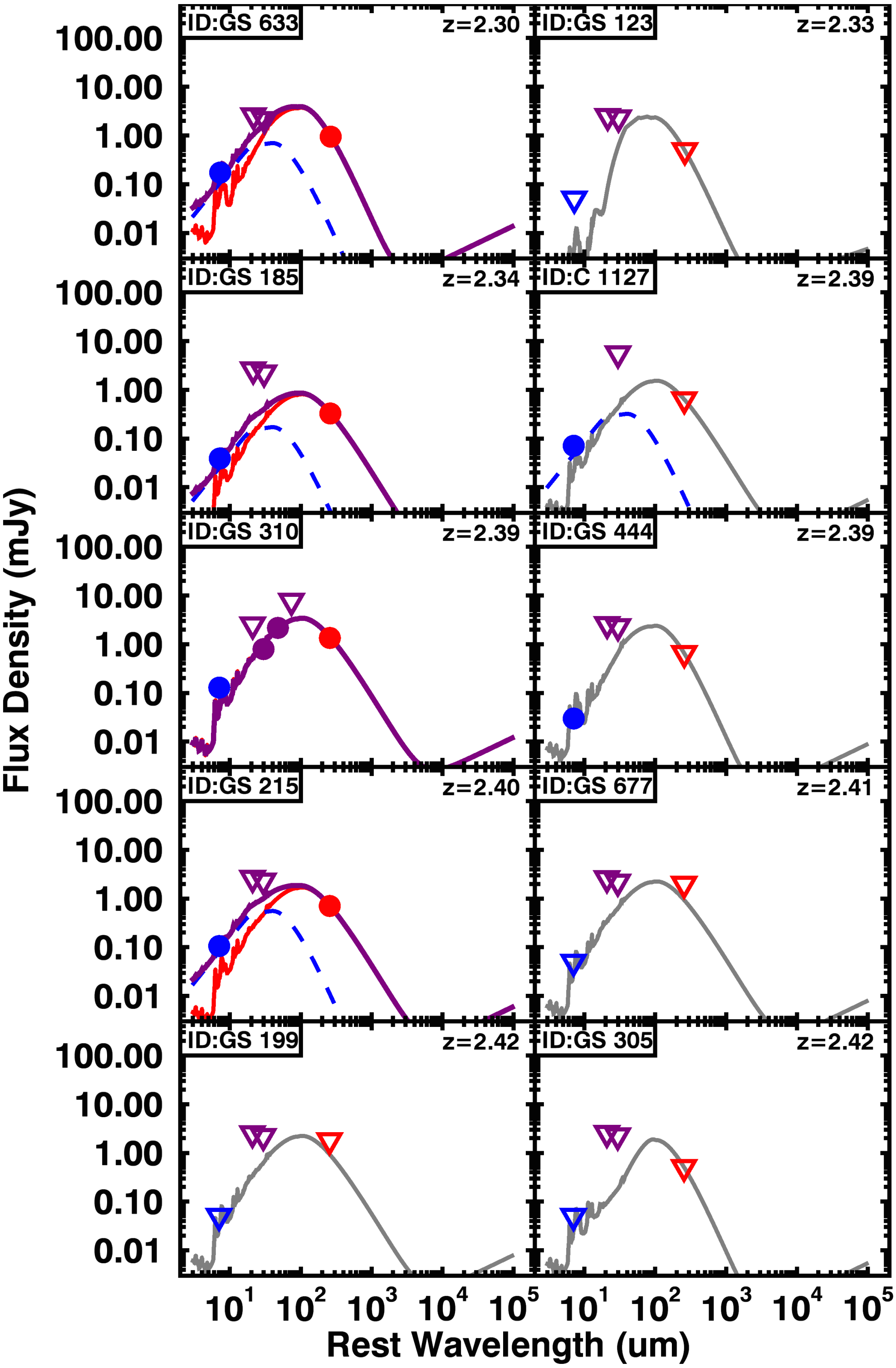


Figure 5.7.1: Continued

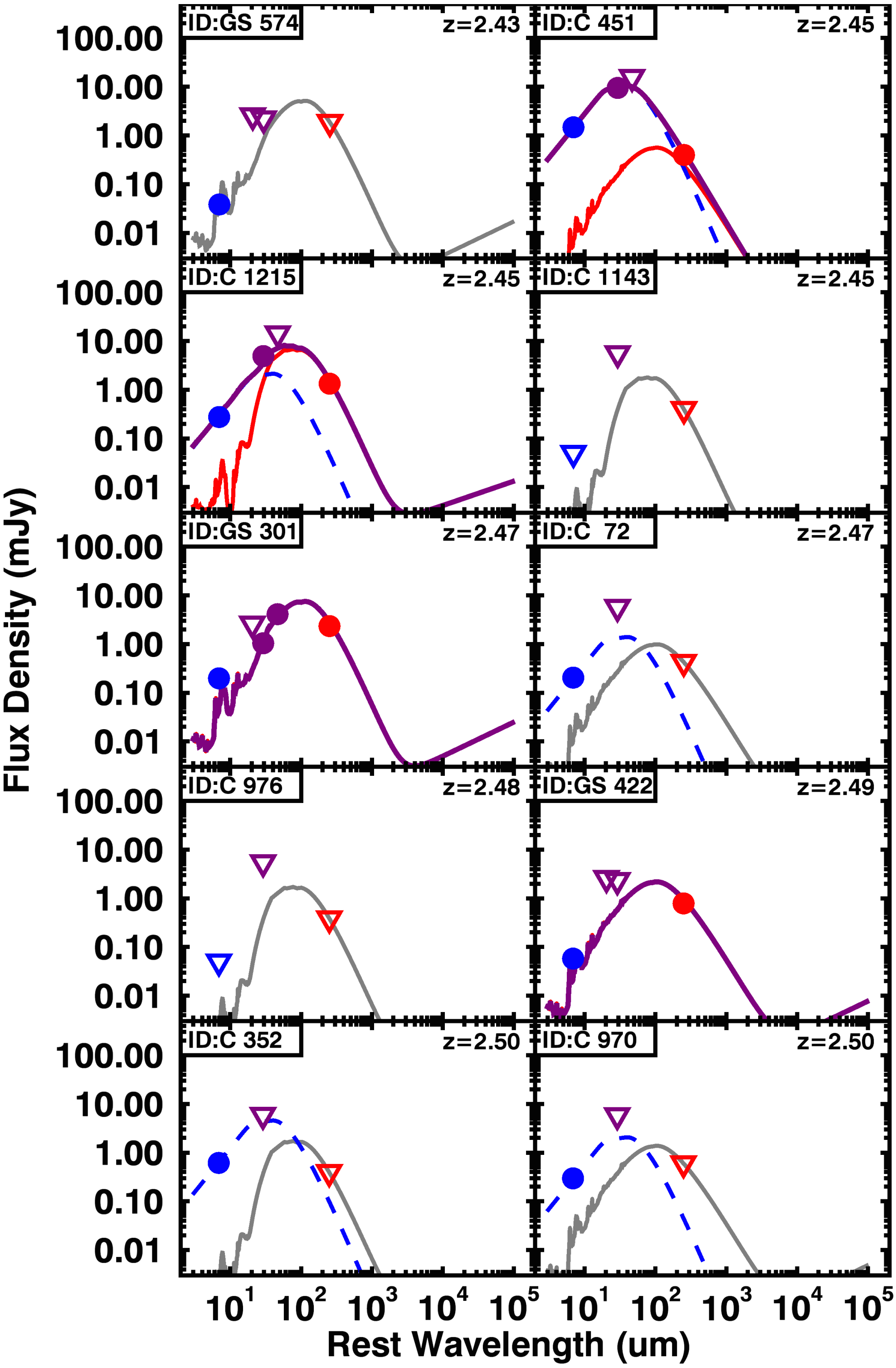


Figure 5.7.1: Continued

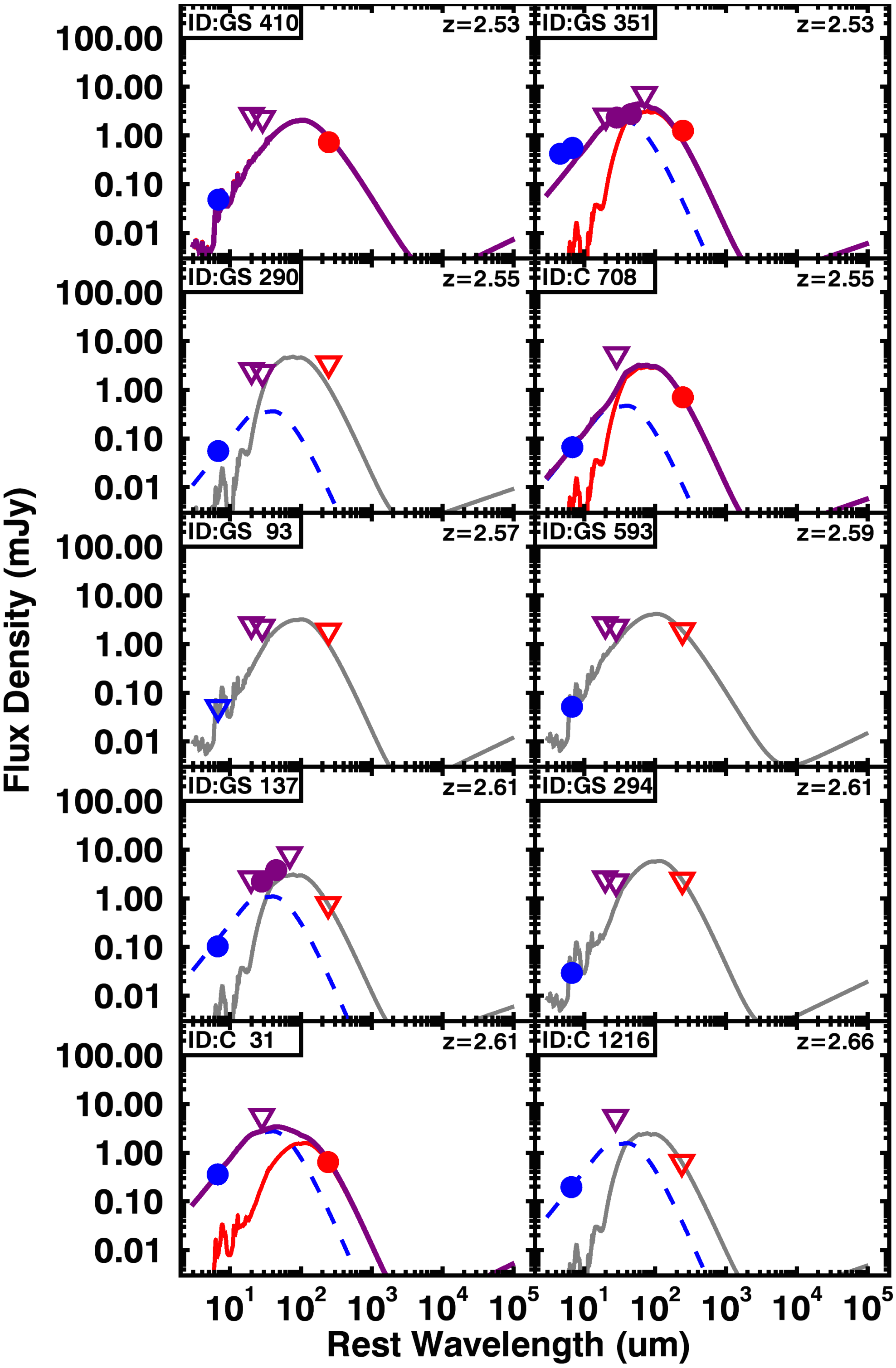


Figure 5.7.1: Continued

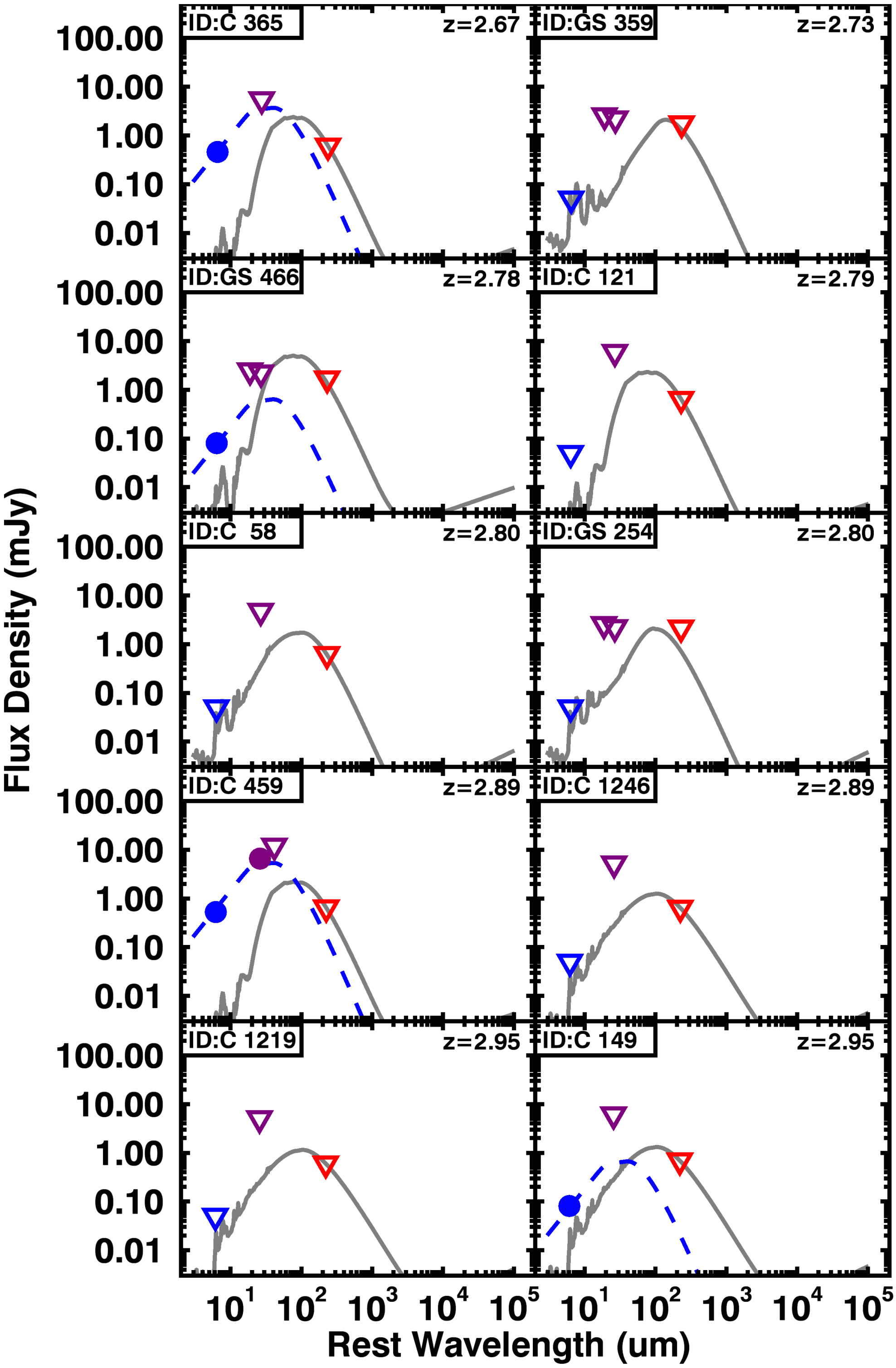


Figure 5.7.1: Continued

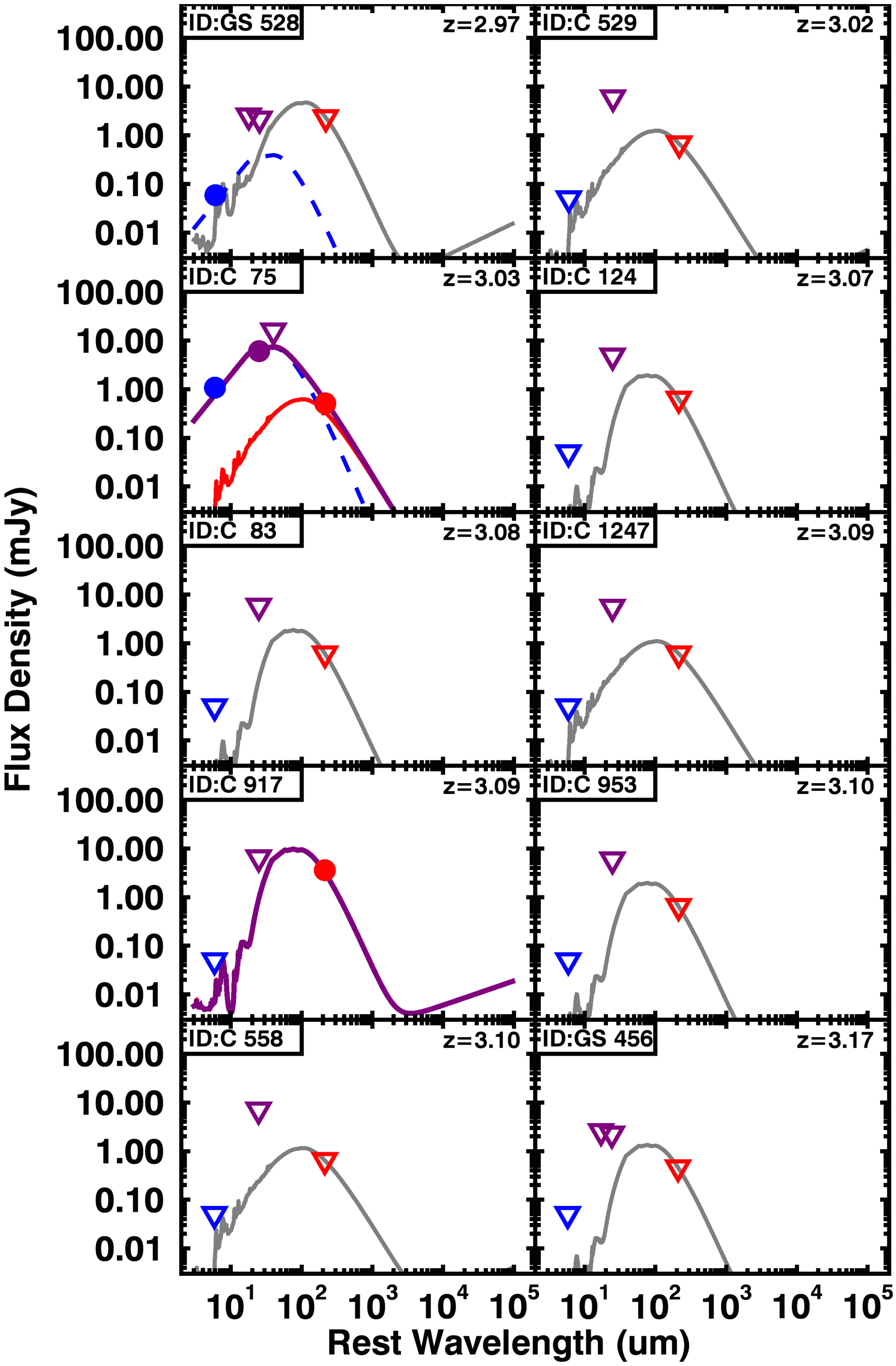


Figure 5.7.1: Continued

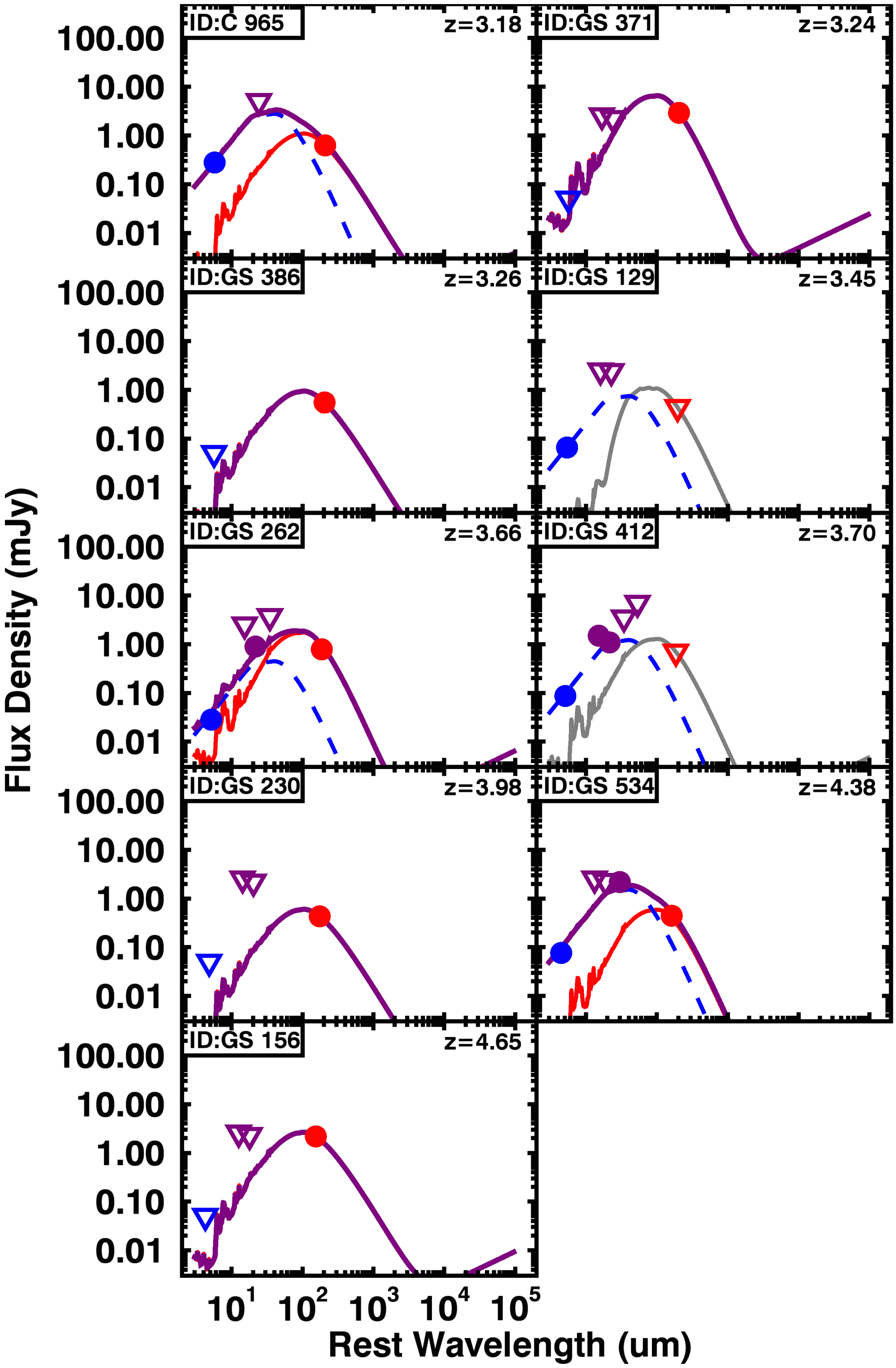


Figure 5.7.1: Continued

CHAPTER 6

Conclusions and future work

This thesis presents observational experiments to investigate the SFRs of galaxies hosting AGN To explore the connection between galaxy and BH growth. I have improved on previous work by developing new analyses, improving on source statistics and taking into account redshift, stellar mass, and AGN contamination effects simultaneously. In this chapter I summarise the main results of the presented work and also discuss ongoing and future work that aim to address the outstanding questions stemming from these results.

6.1 Summary of main results

There have been a multitude of observational findings supporting a co-evolution of central BHs and their host galaxies. This co-evolution is expected to be driven by a connection between the two growing mechanisms of AGN activity (accretion onto the BH), and star formation (see Chapter 1). In an effort to find more direct evidence for a connection between the two mechanisms of AGN and star formation I investigated the SFRs of galaxies hosting X-ray, optical, and radio AGN. For the analysis of these samples I developed new methods to calculate the mean and individual SFRs of these galaxies, and taking into account the AGN contamination. These methods are described in detail in Chapter 2 and include: (a) the SED fitting of IR photometry, with the inclusion of upper limit constraints on photometry, and the decomposition of the IR emission into an AGN and star formation component; (b) the use of the Kaplan-Meier method of calculating mean SFRs with the upper limit constraints on the individual SFR values included in the calculations; (c) stacking of multi-wavelength FIR photometry and the calculation of the mean SFRs from the SED fitting to the stacked photometry. In the following sections I summarise the results of three experiments done following these methods.

6.1.1 A remarkably flat relationship between the average star formation rate and AGN luminosity for distant X-ray AGN

In Chapter 3 we studied a sample of $\sim$2000 X-ray detected AGN spanning the redshifts of $0.2 < z < 2.5$ and covering AGN bolometric luminosities of $10^{43} < \frac{\mathrm{L_{AGN}}}{\mathrm{erg\ s^{-1}}} < 5 \times 10^{47}$. Using *Spitzer* and *Herschel* photometry over the wavelength range of 3.6–500μm, we performed SED fitting and decomposition of the star-forming and AGN components (see Chapter 2). Our analysis resulted in individual SFRs, with removed AGN contamination, and upper limit constraints where needed. We calculated the mean SFR in bins of $\mathrm{L_{AGN}}$ and redshift. We found a strong evolution of the average SFR with redshift, tracking the observed evolution of the overall star forming galaxy population. However, we found that the relationship between the average SFR and AGN luminosity is broadly flat at all redshifts and over all of the AGN luminosities explored. By comparing to empirical toy models, we argue that the observed flat relationship is due to short timescale variations in AGN luminosity relative to those of the SFR, which wash out any underlying correlations between SFR and $\mathrm{L_{AGN}}$. Furthermore, using two simple toy models we showed that the exact form of the predicted relationship between SFR and AGN luminosity (and its normalisation) is sensitive to the assumed intrinsic Eddington ratio distribution. Our results are well described by the empirical model of Aird et al. (2013) that assumes the best-fitting Eddington rate distribution of Aird et al. (2012) from the PRIMUS survey of active sources.

6.1.2 The mean SFRs of optically luminous QSOs

In Chapter 4 we studied a sample of $\sim$3000 optically luminous QSOs, over the redshift range of $0.2 < z < 2.5$ and covering the highest AGN luminosities of $10^{45} < \frac{\mathrm{L_{AGN}}}{\mathrm{erg\ s^{-1}}} < 5 \times 10^{47}$. With this sample we built on the work presented in Chapter 3, by increasing the number of AGN at $\mathrm{L_{AGN}} > 10^{45}$ erg s^{-1} by $\sim$10 times. We made use of the available SDSS DR-7 QSO catalogue to select a sample of QSOs within the regions overlapping with the *Herschel*-ATLAS survey. The H-ATLAS survey covers over 300deg^2 in the 100μm and 160μm bands of PACS, and the 250μm, 350μm, and 500μm bands of SPIRE. We make use of the available maps to stack in bins of redshift and $\mathrm{L_{AGN}}$. For each bin we took the mean stacked flux densities in all *Herschel* bands, and combined them to the available WISE photometry to perform IR SED fitting and decomposition. We found that all bins required an AGN component to the IR SED, and the contribution of AGN emission to the 250μm band rose from a few percent at $z \sim$0.3 to up to 60% at $z \gtrsim$0.5, with the contribution increasing with increasing $\mathrm{L_{AGN}}$. This demonstrates that caution should be taken when using $\leq$250μm photometry for single-band SFR indicators at these redshifts and AGN luminosities. The mean SFR as a function of AGN luminosity showed a strong positive trend. Crucially however, when split into redshift ranges the trend is weaker, with an increase of up to a factor of 3 across the AGN luminosities explored. Furthermore, we found that the modest trend between SFR and $\mathrm{L_{AGN}}$ in individual redshift ranges can be attributed to an increase of the average stellar mass with increasing $\mathrm{L_{AGN}}$. The mean SFRs of our QSO sample are consistent within a factor of 2 with normal star-forming galaxies of the same stellar mass and redshift, in agreement with what we found for the samples of X-ray AGN (see previous section). When investigating the properties of a sub-sample of radio-luminous QSOs ($\mathrm{L_{1.4GHz}} > 10^{24}\,\mathrm{W\,Hz^{-1}}$), we found no evidence for a dependence of the mean SFR on radio luminosity. Even in the case of these extreme radio-luminous sources ($10^{24} < \frac{\mathrm{L_{1.4GHz}}}{\mathrm{W Hz^{-1}}} < 10^{29}$), the mean SFRs are consistent with the hosts being normal star-forming galaxies. Our results appear in disagreement with models of galaxy evolution that argue that luminous radio activity impacts star formation. However, we investigated only the optically luminous QSOs that are radio bright, it is entirely possible that a sample of radio bright but optically faint AGN could be distinctly different.

6.1.3 Achieving deeper constrains on the SFRs of X-ray AGN with ALMA 870μm observations

In Chapter 5 we presented new ALMA Band-7 observations of 109 X-ray AGN spanning redshifts of $1 < z \lesssim 4.5$, and X-ray luminosities of $10^{42} < \frac{\mathrm{L_{2-8keV}}}{\mathrm{erg\,s^{-1}}} \lesssim 10^{45}$. We made use of the available photometry from *Spitzer* and *Herschel* observations in combination with the ALMA photometry at 870μm, and performed IR SED fitting. We demonstrated the power of the ALMA 870μm photometry in constraining the IR emission from star-formation, and in doing so helped constrain the IR AGN emission in the MIR. By comparing to the constraints possible with only the 8–500μm photometry, we found that the addition of 870μm photometry results in 35% of the sample with originally only upper limit constraints to now have a SFR measurement. Furthermore, we have achieved a reduction of upper limit values by up to a factor of 10. We explored the $F_{870\mu m}/F_{24\mu m}$–redshift plane as a diagnostic for AGN identification, and found that with our sample of X-ray AGN it can retrieve all AGN with an IR counterpart found by our SED fitting. Due to the fact that our sample is biased to FIR faint sources, we tested the $F_{870\mu m}/F_{24\mu m}$–redshift plane on two other published ALMA samples, a sample of SMGs (IR SED dominated by star formation emission), and radio-powerful obscured AGN (IR SED with strong AGN contribution). The $F_{870\mu m}/F_{24\mu m}$–redshift plane can successfully retrieve all powerful obscured AGN, and successfully places the majority of the SMGs in the star formation dominated region, while also identifying potential AGN counterparts in a fraction of the SMG sample. Consequently, the combination of 24μm and deep 870μm photometry can be used as a diagnostic for the identification of AGN emission in the MIR. We have also demonstrated that our choice of templates covers the observed $F_{870\mu m}/F_{24\mu m}$ colour space of these sub-samples, providing additional confidence in our SED fitting. In contrast, the AGN template of Symeonidis et al. (2016), which predicts a more significant AGN contribution at the FIR wavelengths than previous empirical AGN templates, such as those of Mullaney et al. (2011), and Mor & Netzer (2012), is inconsistent with the majority of the AGN dominated systems investigated here.

6.2 Overall summary and Conclusion

The work presented in this thesis has placed strong constraints on the mean SFRs of a large sample of AGN spanning a wide range of AGN bolometric luminosities and redshifts. The AGN samples in this thesis cover the redshift range of $0.2 < z < 2.5$, a range that samples the redshifts of the observed peak of AGN and star-forming activity (i.e., $z \approx 1$–2), where we could expect high levels of interaction between the two processes. The overall sample of AGN studied in this thesis cover the range of $10^{43} < \frac{L_{AGN}}{\mathrm{erg\,s^{-1}}} < 5 \times 10^{47}$ (Chapter 3 & 4), and have mean SFRs consistent with those of the general star-forming galaxy population (as represented by the main sequence of star-forming galaxies; e.g., Elbaz et al. 2011, Schreiber et al. 2015). Once we have taken into account of the redshift and mass dependencies, and the AGN contribution to the IR emission, the trends of the mean SFR as a function of AGN luminosity are consistent with being flat with evidence of a slight upturn at the highest AGN luminosities. The general form of the trends can be reproduced by empirical models that include short term variability of the AGN luminosity while assuming a long-term correlation of SFR with AGN luminosity. Subsequently a correlation of the two processes through their mutual dependence on the cold gas supply can be flattened due to the short-term variability of the AGN. Even though this suggests that the investigation of trends in the SFR–L_{AGN} plane is not ideal to investigate the interaction of the two activities, the results of Chapter 3 and simulation work by Veale et al. (2014) demonstrate that it can be used as a diagnostic to test the assumptions made in models of galaxy evolution, such as the underlying Eddington rate distribution. Reproducing the results presented in this thesis will also be an important test on cosmological simulations. On average, AGN show no evidence of suppressing the star-formation of their hosts when investigated through their mean SFRs.

We find that including sensitive ALMA photometry at 870μm in our SED fitting procedure provides significantly improved constraints on the SFRs, reducing upper limit constraints by up to an order of magnitude, in comparison to the SFR constraints when using only *Spitzer* and *Herschel* photometry (Chapter 5). Additionally, we have demonstrated how the combination of sensitive 870μm photometry from ALMA with MIPS-24μm can be used as an AGN identification tool.

6.3 Ongoing and Future work

The work in this thesis has defined a base for the understanding of the relationship between star formation and AGN luminosity and has resulted in a number of follow-up projects. In the next few paragraphs I outline ongoing and future work based on the methods and results presented in this thesis and the prospects of understanding the connection between star formation and AGN activity with new and future facilities and state of the art cosmological simulations of galaxy evolution.

6.3.1 Extending the SFR–$\mathrm{L_{AGN}}$ plane to lower redshifts

In Chapter 3 we found that the mean SFR of X-ray AGN follows a flat trend as a function of $\mathrm{L_{AGN}}$ when split into redshift ranges. For that work we concentrated on moderate to high redshifts of $0.2 < z < 2.5$, that include the epoch where AGN activity and star formation where at their peak.

In the local Universe however, extreme sources such as luminous AGN (or QSOs) are more of a rarity and have properties that characterise them as outliers to the general population. This leads to the question, what would the relation between the mean SFR and $\mathrm{L_{AGN}}$ look like at these low redshifts, and is it any different to that observed at the epoch of peak activity? The result would demonstrate any potential for an evolution of the observed trends with redshift.

The SWIFT-BAT telescope (Gehrels et al. 2004) provides an all-sky hard X-ray (14–195keV) survey, ideal for constructing a sample of local AGN almost completely unaffected by the presence of absorption. We have made use of the available catalogue from Shimizu et al. (2015) containing 313 SWIFT-BAT AGN selected at $z \lesssim 0.05$ from the 70-month catalogue (Baumgartner et al. 2013) with WISE and *Herschel* counterparts. The sample covers the AGN bolometric luminosity range of $10^{43} < \frac{\mathrm{L_{AGN}}}{\mathrm{erg\,s^{-1}}} < 5 \times 10^{45}$. Using the SED fitting method described in Chapter 2, we fit and decompose the IR SED into the AGN and star-forming components for all sources. More than 90% of the sample have a well constrained measurement on their SFR, a significant improvement over our higher redshift samples. Following Stanley et al. (2015; Chapter 3), the sample is split in bins of AGN luminosity, and for each bin the mean SFR is calculated (Murray, S. et al. in prep). From the preliminary results the flat trend of the mean SFRs as a function of AGN luminosity that is observed at higher redshifts (Chapter 3) is also seen for this sample of local X-ray AGN.

6.3.2 Constraining the faint end of the SFR distribution of AGN

With the new constraints on individual SFRs using ALMA observations presented in Chapter 5, we can start determining the distributions of SFRs of X-ray AGN. The mode and shape of the distribution can reveal possible differences between the AGN population and the overall star-forming galaxy population beyond the simple mean, as distributions of different shape can have similar means (i.e., Gaussian distributions of different widths but with the same center). A well constrained SFR distribution will also enable important tests on galaxy evolution models. Based on the SFR measurements of Chapter 5, J. Scholtz (in prep) will define the SFR distributions of AGN in two AGN luminosity ranges, following and developing further the methods of Mullaney et al. (2015). With the resulting distributions we will explore underlying differences in the SFR distribution of AGN at different luminosities, as well as differences to the overall star-forming galaxy population.

An important extension in this analyses when investigating the faint-end of the SFR distributions, will be the use of multi-wavelength SFR indicators. As the IR luminosity covers the re-emission of the obscured light of star-forming regions, and obscuration is possibly less significant in galaxies of low SFRs, it will be necessary to also have good constraints on the overall emission from both un-obscured and obscured star formation. Deep fields, such as COSMOS, have a plethora of multi-wavelength data available, including optical and rest-frame UV photometry. However, a number of challenges occur at these wavelengths, including varying sensitivity, and successfully disentangling the emission from: the AGN, the older stellar population and the ongoing star-forming emission.

In the era of the James Webb Space Telescope (JWST) we will be able to observe the rest-frame NIR and MIR emission of AGN and their host galaxies to lower sensitivities than previously achieved. By combining these observations with sensitive ALMA observations that can place strong constraints on the Submm end of the SED we will be able to disentangle the total emission due to star formation and that of the AGN. With such measurements it will be possible to finally constrain the full distributions of SFRs for AGN samples.

6.3.3 Testing models of galaxy evolution

In this thesis, and in the ongoing work described above, we have placed strong constraints on the trends of the SFR as a function of L_{AGN}, and in Chapter 3 we have demonstrated how empirical toy models can produce different predictions of these trends for different assumptions on the Eddington ratio (or accretion rate) distributions. Furthermore, Veale et al. (2014) also showed that the SFR–L_{AGN} plane is more sensitive than luminosity functions, to the initial conditions assumed, such as the Eddington ratio distribution. Overall, our results can be a useful tool for simulations of galaxy evolution to be tested.

Recently two major cosmological scale hydrodynamical models of galaxy evolution were released, the Illustris (Vogelsberger et al. 2014) and the EAGLE (Schaye et al. 2015) simulations. In these simulations a set of initial conditions and physical equations are set and the simulated universe evolves through cosmic time. Simulations such as these have the advantage of allowing us to trace individual or a set of galaxies throughout cosmic time, something particularly advantageous for understanding the interaction of the AGN with its host galaxy and how it changes with time. In both models the AGN feedback can regulate/quench the star-formation in the host, as well as the accretion on to the BH itself, mainly through ejecting and heating the cold-gas of the galaxy. By extracting the information for galaxies with active AGN at different redshifts in the simulations, it is possible to calculate the mean SFRs in bins of AGN luminosity, following the methods of this thesis. The results from the simulated AGN sample could then be compared to the observations as a function of mass and redshift. If the model reproduces the observed trends, we can use it to work backwards and better understand what drives this trend, and how the mechanism of AGN feedback can affect it, and what the underlying distribution looks like. Already, the results presented in Chapter 3 are being used to test the results of the EAGLE cosmological hydrodynamical simulation (McApline, S. et al. in prep). As we struggle to observationally constrain the quiescent fraction of AGN and how it compares proportionally to the general galaxy population (i.e, is the fraction of quiescent AGN hosts larger than the general quiescent fraction of galaxies), predictive results could help point to the right answer.

6.3.4 Radio powerful AGN and their impact on star formation

In this thesis we have placed strong constraints on the mean SFRs of the overall AGN population, and have shown that there is no evidence for an influence of the AGN on the average star-formation rates. However, it is possible for AGN going through extreme phases of activity to show different results, and so samples of extreme sources could show different SFRs. Extreme AGN samples can be selected through their optical, radio emission, and/or obscuration levels, by choosing the top few percent for each property .

I aim to perform a study on a sample of radio-luminous AGN, selected to have $L_{1.4GHz} \geq 10^{25}\,W\,Hz^{-1}$, from the VLA-COSMOS survey (Schinnerer et al. 2007, 2010), to constrain their star-forming properties following the methods I have developed so far, and place constraints on the star-forming properties of high-z radio luminous AGN. The COSMOS deep field benefits from multi-wavelength coverage that enables such a study, including *Spitzer* and deep *Herschel* observations. I have applied for time with the ALMA observatory to observe 41 such galaxies in the redshift range of $1.5< z <3.2$ at 870μm to a depth that will allow us to reach up to a factor of 10 lower constraints on SFRs (as demonstrated in Chapter 5) than we can achieve with just *Spitzer* and *Herschel* photometry. With the resulting SFRs we can construct SFR distributions, crucial to finding subtle differences between samples that can be concealed when using the simple mean (e.g., Mullaney et al. 2015; Scholtz in prep). This sample consists of both HERGs (High Excitation Radio Galaxies; dominant at luminosities of $L_{1.4GHz} > 10^{26}\,W\,Hz^{-1}$; e.g, Best & Heckman 2012), and LERGs (Low Excitation Radio Galaxies; dominant at luminosities of $L_{1.4GHz} < 10^{26}\,W\,Hz^{-1}$; e.g, Best & Heckman 2012). The two classes of AGN have characteristically different accretion rates, that could produce different types of feedback mechanisms, and so our sample will allow us to compare the impact the SFRs of AGN in different modes.

With the improved constraints on the SFR distributions of this sample, we will be able to compare to the general star-forming galaxy population, other extreme AGN samples such as that of Lonsdale et al. (2015) that consists of radio-powerful obscured AGN, and the general AGN population from X-ray and optical selections (Chapter 3, Chapter 4, Scholtz in prep).

6.4 Final remarks

In the past few decades our understanding of AGN has developed significantly, from perceiving them as a rare occurrences not representative of the general galaxy population, to be accepted as a crucial part in the evolution of massive galaxies. Studies of AGN are now in the forefront of extragalactic astronomy research, with both theoretical and observational studies trying to understand the nature and extent of the impact of AGN on the evolution of their host galaxies. Theoretical models on both cosmological and galactic scales predict that AGN eject a significant amount of energy into the ISM and the surrounding environment of the galaxy and regulate the amount of cold gas in the galaxy (and hence regulate the star formation of the galaxy). Evidence for the powerful output of AGN have been observed in the form of molecular gas outflows and radio jets, expelling and/or preventing gas from cooling. However, the predictions that AGN can regulate/quench the star formation in the host galaxy, are still to be confirmed by observational results (including this thesis). Fortunately, we are now entering a new era of observational and computational capabilities with observatories such as ALMA, JVLA, e-MERLIN, and LOFAR already in use, and JWST, and SKA in the future. With the new radio observatories we will be able to investigate the full range of radio AGN luminosities (down to fainter levels than previous surveys that mainly consist of the most luminous part of the population), as they are designed to observe down to faint fluxes at high spatial resolution. With ALMA in the Submm, and JWST in the NIR–MIR we will be able to extend the research of AGN and their host galaxies to even higher redshifts, and explore the evolution of galaxies and AGN from the very first galaxies until now.

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, APJS, 182, 543

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Physical Review Letters, 116, 241103

Aird, J., Coil, A. L., Georgakakis, A., et al. 2015, MNRAS, 451, 1892

Aird, J., Nandra, K., Laird, E. S., et al. 2010, MNRAS, 401, 2531

Aird, J., Coil, A. L., Moustakas, J., et al. 2012, ApJ, 746, 90

—. 2013, ApJ, 775, 41

Alexander, D. M., Bauer, F. E., Chapman, S. C., et al. 2005, ApJ, 632, 736

Alexander, D. M., & Hickox, R. C. 2012, NAR, 56, 93

Alexander, D. M., Bauer, F. E., Brandt, W. N., et al. 2003a, AJ, 126, 539

—. 2003b, AJ, 125, 383

Alexander, D. M., Chary, R.-R., Pope, A., et al. 2008, ApJ, 687, 835

Antonucci, R. 1993, ARAA, 31, 473

Assef, R. J., Kochanek, C. S., Ashby, M. L. N., et al. 2011, ApJ, 728, 56

Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, ApJ, 772, 26

Azadi, M., Aird, J., Coil, A. L., et al. 2015, ApJ, 806, 187

Ballantyne, D. R., Everett, J. E., & Murray, N. 2006, ApJ, 639, 740

Barger, A. J., & Cowie, L. L. 2005, ApJ, 635, 115

Barger, A. J., Cowie, L. L., Owen, F. N., et al. 2014, arXiv:1412.1148

Barger, A. J., Cowie, L. L., & Wang, W.-H. 2008, ApJ, 689, 687

Baumgartner, W. H., Tueller, J., Markwardt, C. B., et al. 2013, APJS, 207, 19

Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559

Berta, S., Magnelli, B., Nordon, R., et al. 2011, A&A, 532, A49

Best, P. N., & Heckman, T. M. 2012, MNRAS, 421, 1569

Best, P. N., Kauffmann, G., Heckman, T. M., et al. 2005, MNRAS, 362, 25

Béthermin, M., Le Floc'h, E., Ilbert, O., et al. 2012, A&A, 542, A58

Béthermin, M., Kilbinger, M., Daddi, E., et al. 2014, A&A, 567, A103

Blain, A. W., Smail, I., Ivison, R. J., & Kneib, J.-P. 1999, MNRAS, 302, 632

Blain, A. W., Smail, I., Ivison, R. J., Kneib, J.-P., & Frayer, D. T. 2002, , 369, 111

Bongiorno, A., Merloni, A., Brusa, M., et al. 2012, MNRAS, 427, 3103

Brandl, B. R., Bernard-Salas, J., Spoon, H. W. W., et al. 2006, ApJ, 653, 1129

Brandt, W. N., & Alexander, D. M. 2015, AAPR, 23, 1

Brandt, W. N., Alexander, D. M., Hornschemeier, A. E., et al. 2001, AJ, 122, 2810

Bridge, C. R., Blain, A., Borys, C. J. K., et al. 2013, ApJ, 769, 91

Brusa, M., Comastri, A., Gilli, R., et al. 2009, ApJ, 693, 8

Burgarella, D., Buat, V., Gruppioni, C., et al. 2013, A&A, 554, A70

Calzetti, D., Kennicutt, R. C., Engelbracht, C. W., et al. 2007, ApJ, 666, 870

Calzetti, D., Wu, S.-Y., Hong, S., et al. 2010, ApJ, 714, 1256

Caplar, N., Lilly, S., & Trakhtenbrot, B. 2014, arXiv:1411.3719

Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, A&A, 497, 635

Cardamone, C. N., Urry, C. M., Damen, M., et al. 2008, ApJ, 680, 130

Casey, C. M., Chen, C.-C., Cowie, L. L., et al. 2013, MNRAS, 436, 1919

Cattaneo, A., Faber, S. M., Binney, J., et al. 2009, Nature, 460, 213

Chabrier, G. 2003, PASP, 115, 763

Chary, R., & Elbaz, D. 2001, ApJ, 556, 562

Chen, C.-T. J., Hickox, R. C., Alberts, S., et al. 2013, ApJ, 773, 3

Civano, F., Elvis, M., Brusa, M., et al. 2012, APJS, 201, 30

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039

Conselice, C. J., Chapman, S. C., & Windhorst, R. A. 2003, ApJL, 596, L5

Croom, S. M., Richards, G. T., Shanks, T., et al. 2009, MNRAS, 399, 1755

Cucciati, O., Tresse, L., Ilbert, O., et al. 2012, A&A, 539, A31

Daddi, E., Dickinson, M., Morrison, G., et al. 2007a, ApJ, 670, 156

Daddi, E., Alexander, D. M., Dickinson, M., et al. 2007b, ApJ, 670, 173

Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686

Dale, D. A., & Helou, G. 2002, ApJ, 576, 159

Damen, M., Labbé, I., van Dokkum, P. G., et al. 2011, ApJ, 727, 1

Debuhr, J., Quataert, E., & Ma, C.-P. 2012, MNRAS, 420, 2221

Del Moro, A., Alexander, D. M., Mullaney, J. R., et al. 2013, A&A, 549, A59

Delvecchio, I., Gruppioni, C., Pozzi, F., et al. 2014, MNRAS, 439, 2736

Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604

Dole, H., Lagache, G., Puget, J.-L., et al. 2006, A&A, 451, 417

Domínguez Sánchez, H., Bongiovanni, A., Lara-López, M. A., et al. 2014a, MNRAS, 441, 2

—. 2014b, MNRAS, 441, 2

Done, C., Davis, S. W., Jin, C., Blaes, O., & Ward, M. 2012, MNRAS, 420, 1848

Donley, J. L., Rieke, G. H., Pérez-González, P. G., Rigby, J. R., & Alonso-Herrero, A. 2007, ApJ, 660, 167

Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142

Dressler, A., & Richstone, D. O. 1988, ApJ, 324, 701

Drouart, G., De Breuck, C., Vernet, J., et al. 2014, A&A, 566, A53

Dye, S., Dunne, L., Eales, S., et al. 2010, A&A, 518, L10

Eales, S., Dunne, L., Clements, D., et al. 2010a, PASP, 122, 499

Eales, S. A., Raymond, G., Roseboom, I. G., et al. 2010b, A&A, 518, L23

Eisenhardt, P. R. M., Wu, J., Tsai, C.-W., et al. 2012, ApJ, 755, 173

Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33

Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119

Elvis, M., Civano, F., & Vignali et al., C. 2009, APJS, 184, 158

Elvis, M., Wilkes, B. J., McDowell, J. C., et al. 1994, APJS, 95, 1

Fabian, A. C. 2012, ARAA, 50, 455

Feigelson, E. D., & Nelson, P. I. 1985, ApJ, 293, 192

Fiore, F., Brusa, M., Cocchia, F., et al. 2003, A&A, 409, 79

Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L., & Shafer, R. A. 1998, ApJ, 508, 123

Gabor, J. M., & Bournaud, F. 2013, MNRAS, 434, 606

Geach, J. E., Smail, I., Moran, S. M., et al. 2011, ApJL, 730, L19

Gehrels, N., Chincarini, G., Giommi, P., et al. 2004, ApJ, 611, 1005

Genzel, R., & Cesarsky, C. J. 2000, ARAA, 38, 761

George, I. M., Turner, T. J., Yaqoob, T., et al. 2000, ApJ, 531, 52

Giacconi, R., Zirm, A., Wang, J., et al. 2002, APJS, 139, 369

Gilli, R., Comastri, A., & Hasinger, G. 2007, A&A, 463, 79

Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, A&A, 518, L3

Gruppioni, C., Pozzi, F., Andreani, P., et al. 2010, A&A, 518, L27

Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, MNRAS, 436, 2875

Gürkan, G., Hardcastle, M. J., Jarvis, M. J., et al. 2015, MNRAS, 452, 3776

Hao, C.-N., Kennicutt, R. C., Johnson, B. D., et al. 2011, ApJ, 741, 124

Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2007, MNRAS, 376, 1849

Hardcastle, M. J., Ching, J. H. Y., Virdee, J. S., et al. 2013, MNRAS, 429, 2407

Harms, R. J., Ford, H. C., Tsvetanov, Z. I., et al. 1994, ApJL, 435, L35

Harrison, C. M. 2014, in IAU Symposium, Vol. 304, IAU Symposium, 284–290

Harrison, C. M., Alexander, D. M., Mullaney, J. R., et al. 2012, ApJL, 760, L15

Harrison, C. M., Simpson, J. M., Stanley, F., et al. 2016, MNRAS, 457, L122

Hasinger, G. 2008, A&A, 490, 905

Hasinger, G., Miyaji, T., & Schmidt, M. 2005, A&A, 441, 417

Heckman, T. M., & Best, P. N. 2014, ARAA, 52, 589

Heckman, T. M., Kauffmann, G., Brinchmann, J., et al. 2004, ApJ, 613, 109

Hickox, R. C., Mullaney, J. R., Alexander, D. M., et al. 2014, ApJ, 782, 9

Hodge, J. A., Karim, A., Smail, I., et al. 2013, ApJ, 768, 91

Hönig, S. F., Beckert, T., Ohnaka, K., & Weigelt, G. 2006, A&A, 452, 459

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2005, ApJ, 630, 716

Hopkins, P. F., Hickox, R., Quataert, E., & Hernquist, L. 2009, MNRAS, 398, 333

Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, ApJ, 654, 731

Hsu, L.-T., Salvato, M., Nandra, K., et al. 2014, ApJ, 796, 60

Hubble, E. P. 1925, ApJ, 62

—. 1929, ApJ, 69

Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55

Jahnke, K., Kuhlbrodt, B., & Wisotzki, L. 2004, MNRAS, 352, 399

Jansen, F., Lumb, D., Altieri, B., et al. 2001, A&A, 365, L1

Jones, S. F., Blain, A. W., Stern, D., et al. 2014, MNRAS, 443, 146

Jones, S. F., Blain, A. W., Lonsdale, C., et al. 2015, MNRAS, 448, 3325

Kalfountzou, E., Stevens, J. A., Jarvis, M. J., et al. 2014, MNRAS, 442, 1181

Kaplan, E. L., & Meier, P. 1958, Journal of the American Statistical Association, 53, pp. 457

Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, ApJ, 730, 61

Karim, A., Swinbank, A. M., Hodge, J. A., et al. 2013, MNRAS, 432, 2

Karouzos, M., Im, M., Trichas, M., et al. 2014, ApJ, 784, 137

Kauffmann, G., & Heckman, T. M. 2009, MNRAS, 397, 135

Kennicutt, Jr., R. C. 1998, ARAA, 36, 189

Kirkpatrick, A., Pope, A., Alexander, D. M., et al. 2012, ApJ, 759, 139

Kirkpatrick, A., Pope, A., Charmandaris, V., et al. 2013, ApJ, 763, 123

Kormendy, J., & Ho, L. C. 2013, ARAA, 51, 511

Kormendy, J., & Richstone, D. 1995, ARAA, 33, 581

La Franca, F., Fiore, F., Comastri, A., et al. 2005, ApJ, 635, 864

Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2015, ArXiv e-prints

Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, APJS, 154, 166

Lagache, G., Puget, J.-L., & Gispert, R. 1999, APSS, 269, 263

Laird, E. S., Nandra, K., Pope, A., & Scott, D. 2010, MNRAS, 401, 2763

Lanzuisi, G., Ranalli, P., Georgantopoulos, I., et al. 2015, A&A, 573, A137

Laor, A., & Behar, E. 2008, MNRAS, 390, 847

Lapi, A., González-Nuevo, J., Fan, L., et al. 2011, ApJ, 742, 24

Le Floc'h, E., Papovich, C., Dole, H., et al. 2005, ApJ, 632, 169

Lehmer, B. D., Alexander, D. M., Bauer, F. E., et al. 2010, ApJ, 724, 559

Leroy, A. K., Walter, F., Brinks, E., et al. 2008, AJ, 136, 2782

Liddle, A. R. 2004, MNRAS, 351, L49

Lonsdale, C. J., Lacy, M., Kimball, A. E., et al. 2015, ApJ, 813, 45

Luo, B., Bauer, F. E., Brandt, W. N., et al. 2008, APJS, 179, 19

Lutz, D. 2014, ARAA, 52, 373

Lutz, D., Mainieri, V., Rafferty, D., et al. 2010, ApJ, 712, 1287

Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, A&A, 532, A90

Lynden-Bell, D. 1969, Nature, 223, 690

Madau, P., & Dickinson, M. 2014, ARAA, 52, 415

Magliocchetti, M., Lutz, D., Santini, P., et al. 2016, MNRAS, 456, 431

Magliocchetti, M., Lutz, D., Rosario, D., et al. 2014, MNRAS, 442, 682

Magnelli, B., Popesso, P., Berta, S., et al. 2013, A&A, 553, A132

Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, A&A, 561, A86

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285

Mainieri, V., Bongiorno, A., Merloni, A., et al. 2011, A&A, 535, A80

Marchesi, S., Civano, F., Elvis, M., et al. 2016, ApJ, 817, 34

Marconi, A., & Hunt, L. K. 2003, ApJL, 589, L21

Marconi, A., Risaliti, G., Gilli, R., et al. 2004, MNRAS, 351, 169

Mateos, S., Alonso-Herrero, A., Carrera, F. J., et al. 2012, MNRAS, 426, 3271

Matthews, T. A. 1963, AJ, 68, 77

McAlpine, K., Jarvis, M. J., & Bonfield, D. G. 2013, MNRAS, 436, 1084

McNamara, B. R., & Nulsen, P. E. J. 2012, New Journal of Physics, 14, 055023

Merloni, A., & Heinz, S. 2008, MNRAS, 388, 1011

Merloni, A., Rudnick, G., & Di Matteo, T. 2004, MNRAS, 354, L37

Merloni, A., Bongiorno, A., Bolzonella, M., et al. 2010, ApJ, 708, 137

Miyaji, T., Hasinger, G., Salvato, M., et al. 2015, ApJ, 804, 104

Mor, R., & Netzer, H. 2012, MNRAS, 420, 526

Moustakas, J., Coil, A. L., Aird, J., et al. 2013, ApJ, 767, 50

Mukherjee, S., Feigelson, E. D., Jogesh Babu, G., et al. 1998, ApJ, 508, 314

Mullaney, J. R., Alexander, D. M., Goulding, A. D., & Hickox, R. C. 2011, MNRAS, 414, 1082

Mullaney, J. R., Pannella, M., Daddi, E., et al. 2012a, MNRAS, 419, 95

Mullaney, J. R., Daddi, E., Béthermin, M., et al. 2012b, ApJL, 753, L30

Mullaney, J. R., Alexander, D. M., Aird, J., et al. 2015, MNRAS, 453, L83

Murphy, E. J., Condon, J. J., Schinnerer, E., et al. 2011, ApJ, 737, 67

Mushotzky, R. F., Cowie, L. L., Barger, A. J., & Arnaud, K. A. 2000, Nature, 404, 459

Nandra, K., & Pounds, K. A. 1994, MNRAS, 268, 405

Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008, ApJ, 685, 160

Netzer, H. 2015, ARAA, 53, 365

Netzer, H., Lani, C., Nordon, R., et al. 2015, ArXiv e-prints

Netzer, H., Lutz, D., Schweitzer, M., et al. 2007, ApJ, 666, 806

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJL, 660, L43

Oliver, S. J., Bock, J., Altieri, B., et al. 2012, MNRAS, 424, 1614

Olnon, F. M., Raimond, E., Neugebauer, G., et al. 1986, AAPS, 65, 607

Omont, A., Beelen, A., Bertoldi, F., et al. 2003, A&A, 398, 857

Page, M. J., Symeonidis, M., Vieira, J. D., et al. 2012, Nature, 485, 213

Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, ApJL, 698, L116

Pannella, M., Elbaz, D., Daddi, E., et al. 2015, ApJ, 807, 141

Pâris, I., Petitjean, P., Aubourg, É., et al. 2014, A&A, 563, A54

Pier, E. A., & Krolik, J. H. 1992, ApJ, 401, 99

Pilbratt, G. L., Riedinger, J. R., Pas, T., et al. 2010, A&A, 518, L1

Pineau, F.-X., Motch, C., Carrera, F., et al. 2011, A&A, 527, A126

Podigachoski, P., Barthel, P., Haas, M., Leipski, C., & Wilkes, B. 2015, ApJL, 806, L11

Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, A&A, 518, L2

Polletta, M., Courvoisier, T. J.-L., Hooper, E. J., & Wilkes, B. J. 2000, A&A, 362, 75

Puget, J.-L., Abergel, A., Bernard, J.-P., et al. 1996, A&A, 308, L5

Rafferty, D. A., Brandt, W. N., Alexander, D. M., et al. 2011, ApJ, 742, 3

Ranalli, P., Comastri, A., & Setti, G. 2003, A&A, 399, 39

Rees, G. A., Spitler, L. R., Norris, R. P., et al. 2016, MNRAS, 455, 2731

Rees, M. J. 1984, ARAA, 22, 471

Reuland, M., Röttgering, H., van Breugel, W., & De Breuck, C. 2004, MNRAS, 353, 377

Richards, G. T., Fan, X., Newberg, H. J., et al. 2002, AJ, 123, 2945

Richards, G. T., Nichol, R. C., Gray, A. G., et al. 2004, APJS, 155, 257

Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, APJS, 166, 470

Rigby, E. E., Best, P. N., & Snellen, I. A. G. 2008, MNRAS, 385, 310

Rodighiero, G., Cimatti, A., Gruppioni, C., et al. 2010, A&A, 518, L25

Rodighiero, G., Brusa, M., Daddi, E., et al. 2015, ApJL, 800, L10

Rosario, D. J., Santini, P., Lutz, D., et al. 2012, A&A, 545, A45

—. 2013a, ApJ, 771, 63

Rosario, D. J., Trakhtenbrot, B., Lutz, D., et al. 2013b, A&A, 560, A72

Rosario, D. J., Mozena, M., Wuyts, S., et al. 2013c, ApJ, 763, 59

Rosse, T. E. O. 1850, Philosophical Transactions of the Royal Society of London Series I, 140, 499

Rovilos, E., Comastri, A., Gilli, R., et al. 2012, A&A, 546, A58

Rowan-Robinson, M. 1995, MNRAS, 272, 737

Salpeter, E. E. 1964, ApJ, 140, 796

Sanders, D. B., Mazzarella, J. M., Kim, D.-C., Surace, J. A., & Soifer, B. T. 2003, AJ, 126, 1607

Sanders, D. B., & Mirabel, I. F. 1996, ARAA, 34, 749

Sanders, D. B., Phinney, E. S., Neugebauer, G., Soifer, B. T., & Matthews, K. 1989, ApJ, 347, 29

Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, APJS, 172, 86

Santini, P., Rosario, D. J., Shao, L., et al. 2012, A&A, 540, A109

Sargent, W. L. W., Young, P. J., Lynds, C. R., et al. 1978, ApJ, 221, 731

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Schmidt, M., & Green, R. F. 1983, ApJ, 269, 352

Schmitt, J. H. M. M. 1985, ApJ, 293, 178

Schmitt, J. L. 1968, Nature, 218, 663

Schneider, D. P., Hall, P. B., Richards, G. T., et al. 2007, AJ, 134, 102

Schneider, D. P., Richards, G. T., Hall, P. B., et al. 2010, AJ, 139, 2360

Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74

Schwarz, G. 1978, The Annals of Statistics, 6, 461

Schweitzer, M., Lutz, D., Sturm, E., et al. 2006, ApJ, 649, 79

Seyfert, C. K. 1943, ApJ, 97, 28

Shankar, F. 2009, NAR, 53, 57

Shankar, F., Salucci, P., Granato, G. L., De Zotti, G., & Danese, L. 2004, MNRAS, 354, 1020

Shao, L., Lutz, D., Nordon, R., et al. 2010, A&A, 518, L26

Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, APJS, 194, 45

Shimizu, T. T., Mushotzky, R. F., Meléndez, M., Koss, M., & Rosario, D. J. 2015, MNRAS, 452, 1841

Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998, ApJ, 509, 103

Silverman, J. D., Green, P. J., Barkhouse, W. A., et al. 2008, ApJ, 679, 118

Simpson, J. M., Swinbank, A. M., Smail, I., et al. 2014, ApJ, 788, 125

Simpson, J. M., Smail, I., Swinbank, A. M., et al. 2015, ApJ, 807, 128

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Slipher, V. M. 1917, Popular Astronomy, 25, 36

Smith, D. A., Herter, T., & Haynes, M. P. 1998, ApJ, 494, 150

Smolčić, V., Zamorani, G., Schinnerer, E., et al. 2009, ApJ, 696, 24

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

Stanley, F., Harrison, C. M., Alexander, D. M., et al. 2015, MNRAS, 453, 591

Stern, D. 2015, ApJ, 807, 129

Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163

Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30

Swinbank, A. M., Simpson, J. M., Smail, I., et al. 2014, MNRAS, 438, 1267

Symeonidis, M., Giblin, B. M., Page, M. J., et al. 2016, MNRAS

Symeonidis et al., M. 2011, MNRAS, 417, 2239

Tacconi, L. J., Neri, R., Chapman, S. C., et al. 2006, ApJ, 640, 228

Teplitz, H. I., Chary, R., Elbaz, D., et al. 2011, AJ, 141, 1

Tonry, J. L. 1984, ApJL, 283, 27

Tozzi, P., Gilli, R., Mainieri, V., et al. 2006, A&A, 451, 457

Treister, E., & Urry, C. M. 2006, ApJL, 652, L79

Treister, E., Urry, C. M., Chatzichristou, E., et al. 2004, ApJ, 616, 123

Trump, J. R., Hsu, A. D., Fang, J. J., et al. 2013, ApJ, 763, 133

Tsai, C.-W., Eisenhardt, P. R. M., Wu, J., et al. 2015, ApJ, 805, 90

Urry, C. M., & Padovani, P. 1995, PASP, 107, 803

Vaccari, M., Marchetti, L., Franceschini, A., et al. 2010, A&A, 518, L20

Veale, M., White, M., & Conroy, C. 2014, MNRAS, 445, 1144

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, Nature, 509, 177

Volonteri, M. 2010, AAPR, 18, 279

Volonteri, M., Capelo, P. R., Netzer, H., et al. 2015, MNRAS, 449, 1470

Wang, L., Viero, M., Ross, N. P., et al. 2015, MNRAS, 449, 4476

Wang, S. X., Brandt, W. N., Luo, B., et al. 2013, ApJ, 778, 179

Wang, W.-H., Cowie, L. L., Barger, A. J., Keenan, R. C., & Ting, H.-C. 2010, APJS, 187, 251

Wardlow, J. L., Smail, I., Coppin, K. E. K., et al. 2011, MNRAS, 415, 1479

Weisskopf, M. C. 1999, in Bulletin of the American Astronomical Society, Vol. 31, American Astronomical Society Meeting Abstracts, 1514

Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, APJS, 154, 1

Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJL, 754, L29

White, M., Myers, A. D., Ross, N. P., et al. 2012, MNRAS, 424, 933

Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914

Wirth, G. D., Willmer, C. N. A., Amico, P., et al. 2004, AJ, 127, 3121

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868

Xue, Y. Q., Luo, B., Brandt, W. N., & Bauer et al., F. E. 2011, APJS, 195, 10

Xue, Y. Q., Brandt, W. N., Luo, B., et al. 2010, ApJ, 720, 368

Zel'dovich, Y. B., & Podurets, M. A. 1964, Soviet Physics Doklady, 9, 373

Zhong, M., & Hess, K. R. 2009, COBRA Preprint Series