

---

ALMA MATER STUDIORUM  
UNIVERSITÀ DEGLI STUDI DI BOLOGNA

---

DIPARTIMENTO DI FISICA E ASTRONOMIA  
Corso di Dottorato di Ricerca in Astronomia  
CICLO XXVII

Tracing AGN accretion  
and star formation in *Herschel* galaxies

Dottorando  
**Ivan Delvecchio**

Relatore  
**Dott. Francesca Pozzi**

Correlatori:  
**Dott. Carlotta Gruppioni**  
**Prof. Andrea Cimatti**  
**Dott. Giovanni Zamorani**

Coordinatore  
**Prof. L. Moscardini**

Esame finale anno 2014

---

Settore Concorsuale: 02/C1 – Astronomia, Astrofisica, Fisica della Terra e dei Pianeti  
Settore Scientifico-Disciplinare: FIS/05 – Astronomia e Astrofisica



*To my Family  
and Arabel*



# Contents

<b>Abstract</b>	<b>I</b>
<b>1 AGN vs star formation: the current evolutionary picture</b>	<b>1</b>
1.1 Active galactic nuclei: a panchromatic overview . . . . .	2
1.1.1 Super massive black hole . . . . .	2
1.1.2 Accretion disc . . . . .	3
1.1.3 Broad and narrow line regions . . . . .	4
1.1.4 Classification of AGN . . . . .	4
1.1.5 The Unified model of AGN . . . . .	6
1.1.6 Challenging the Unified Model . . . . .	8
1.1.7 X-ray observations of AGN hosts . . . . .	9
1.1.8 Optical/UV continuum and emission lines . . . . .	11
1.1.9 Infrared observations of AGN hosts . . . . .	13
1.2 Evidences for AGN/Galaxy co-evolution . . . . .	18
1.3 What triggers SMBH activity? . . . . .	19
1.4 The variety of AGN host galaxies . . . . .	21
1.5 Evolutionary scheme of AGN/galaxy evolution . . . . .	22
<b>2 The <i>Herschel</i> view of AGN/galaxy evolution</b>	<b>27</b>
2.1 The road to <i>Herschel</i> . . . . .	27
2.2 <i>Herschel</i> surveys and observing strategy . . . . .	28
2.3 Data handling . . . . .	32
2.4 An unbiased tracer of galaxy star formation . . . . .	34
2.4.1 Total (8–1000 $\mu\text{m}$ ) IR emission from dust . . . . .	35
2.4.2 Measuring dust-insensitive SFRs . . . . .	35
2.5 <i>Herschel</i> insights on the evolution of AGN and star forming galaxies . . . . .	37
2.5.1 Two modes of star formation . . . . .	37
2.5.2 A census of the cosmic star formation history . . . . .	38
2.5.3 Hunting for heavily obscured AGN . . . . .	39
2.5.4 Does AGN activity suppress or correlate with galaxy SF? . . . . .	40

<b>3</b>	<b>SED-fitting decomposition</b>	<b>45</b>
3.1	The parent sample . . . . .	46
3.1.1	Multi-wavelength identification . . . . .	46
3.1.2	Spectroscopic and photometric redshifts . . . . .	47
3.2	Fitting broad-band SED <sub>s</sub> . . . . .	49
3.2.1	SED-fitting with <b>MAGPHYS</b> . . . . .	49
3.2.2	Fitting with <b>MAGPHYS + AGN</b> . . . . .	50
3.3	AGN Bolometric Luminosity and Bolometric corrections . . . . .	54
3.4	Testing the AGN incidence . . . . .	55
3.5	AGN detection rate and evolution with $L_{1-1000}$ . . . . .	59
3.6	Comparison with AGN diagnostics . . . . .	60
3.6.1	Comparison with mid-IR colour-colour AGN selection . . . . .	60
3.6.2	Comparison with X-ray AGN samples . . . . .	62
3.6.3	AGN luminosity in CT-AGN candidates . . . . .	63
3.6.4	Piercing through the dust: a candidate CT-QSO missed by deep X-ray observations . . . . .	66
<b>4</b>	<b>Black Hole accretion history from <i>Herschel</i></b>	<b>69</b>
4.1	Introduction . . . . .	69
4.2	Sample selection and multi-wavelength catalogues . . . . .	70
4.3	AGN Bolometric Luminosity Function . . . . .	72
4.3.1	Method . . . . .	72
4.3.2	“Observed”AGN bolometric LF . . . . .	73
4.3.3	Evolution of the best-fit AGN bolometric LF . . . . .	75
4.4	SMBH growth across cosmic time . . . . .	77
4.4.1	Comparison with previous results . . . . .	79
<b>5</b>	<b>Mapping the AGN accretion in the SFR–<math>M_{\star}</math> plane</b>	<b>83</b>
5.1	Introduction . . . . .	83
5.2	<i>Herschel</i> dataset and X-ray counterparts . . . . .	85
5.2.1	Cross-match with X-rays . . . . .	85
5.3	The SFR– $M_{\star}$ plane . . . . .	86
5.4	X-ray analysis . . . . .	87
5.4.1	Stacking on X-ray maps . . . . .	88
5.4.2	Mean X-ray luminosity of <i>Herschel</i> sources . . . . .	88
5.4.3	Subtraction of X-ray emission from star formation . . . . .	89
5.4.4	Correction for nuclear obscuration . . . . .	90
5.5	Results . . . . .	91
5.5.1	Average AGN accretion rate . . . . .	91
5.5.2	Correlation with galaxy properties . . . . .	95
5.5.3	Comparison with previous studies . . . . .	101
5.6	Discussion . . . . .	103

<b>6</b>	<b>Concluding remarks and future prospects</b>	<b>107</b>
6.1	Decomposition of broad-band SEDs . . . . .	107
6.2	AGN accretion history since $z \sim 3$ . . . . .	109
6.3	The connection between AGN accretion and galaxy properties . . . . .	110
6.4	Future prospects . . . . .	111
	<b>Bibliografia</b>	<b>113</b>
	<b>Appendices</b>	<b>133</b>
<b>A</b>	<b>Incompleteness in accretion luminosity</b>	<b>135</b>
<b>B</b>	<b>Details on X-ray stacking</b>	<b>139</b>
<b>C</b>	<b>Statistics in the SFR–<math>M_*</math> plane</b>	<b>141</b>
<b>D</b>	<b>Maximum Likelihood function</b>	<b>145</b>





# Abstract

In this Thesis we investigate the integrated properties of Active Galactic Nuclei (AGN) and the mutual relationships with their host galaxies in large extragalactic samples of sources selected in the infrared (IR) with the *Herschel* satellite.

An increasing number of studies highlights the great relevance of AGN in the framework of galaxy evolution. In the local Universe, the mass of local super massive black holes (SMBHs) have been found to correlate with the stellar bulge masses of their hosts. This relation is generally interpreted as the outcome of a common, long-lasting interplay between SMBHs and galaxy growth over cosmic time. The question of what could be the detailed mechanisms shaping this connection is still open.

Since both AGN and galaxy emission processes are detected over a wide range wavelengths (from radio to X-rays), it has become clear that a multi-wavelength approach is mandatory to better understand the physical and evolutionary properties of these sources. Studies based on optical, ultraviolet and X-ray observations of AGN host galaxies can be severely affected by dust and gas attenuation and can provide only an incomplete perspective of the galaxies.

Indeed, a significant fraction of the stellar light might be extinguished by interstellar dust clouds, to be subsequently re-emitted at mid-IR ( $\lambda \sim 3\text{--}12 \mu\text{m}$ ) and far-IR ( $\lambda \sim 100 \mu\text{m}$ ) wavelengths. At the same time, also most of the AGN emission that originates from the central engine can be hidden by circumnuclear (on  $<10$  pc scales) dust and re-emitted in the mid-IR ( $\lambda \sim 10\text{--}30 \mu\text{m}$ ) regime.

For these reasons, it is now commonly agreed that IR observations provide a complementary view of AGN and star formation in galaxies, particularly in highly obscured environments. This relevance also holds in a more cosmological context, since both observations and theoretical models support a scenario where galaxy star formation activity was significantly more obscured in the past ( $z \sim 1\text{--}3$ ), rather than in the present Universe. This further emphasizes the complementary perspective provided by IR observations to studies performed at other wavelengths.

In this Thesis, we study the physical properties and the cosmic evolution of AGN and their host galaxies since  $z \sim 3$ . Our analysis exploits samples of star forming galaxies detected with *Herschel* at far-IR wavelengths (from 70 up to 500  $\mu\text{m}$ ) in different extragalactic surveys, such as COSMOS and the deep GOODS (South and North) fields. The *Herschel*-based selection is sensitive to the peak of interstellar dust emission (rest-frame around 100  $\mu\text{m}$ ) up to  $z \sim 3$ . This allows us to measure with unprecedented accuracy the global star formation content in galaxies, without uncertain extrapolations based on local

templates. However, an additional contribution to the total IR energy budget can arise from warmer dust heated by the AGN rather than by stellar processes. *Spitzer* data, together with *Herschel* capabilities, are paramount to reach a complete assessment of the dust content at different temperatures across the IR domain. In addition, the broad-band photometric coverage available in COSMOS and the GOODS fields, allows us to implement *Herschel* and *Spitzer* photometry with multi-wavelength ancillary data, spanning the electromagnetic spectrum down to the optical/ultraviolet regime.

We use this panchromatic approach to characterize the observed spectral energy distributions (SEDs) of each *Herschel* galaxy. A multicomponent SED-fitting decomposition is performed to decouple the emission due to star formation from that due to AGN accretion, and to estimate both host-galaxy parameters (such as stellar mass,  $M_{\star}$ , and star formation rate, SFR), and nuclear intrinsic bolometric luminosities. A statistical analysis has been developed to evaluate the relative incidence of the AGN component in each observed SED, and to distinguish between galaxies with or without a significant AGN contribution within our sample.

We use the individual estimates of AGN bolometric luminosity obtained through SED-fitting decomposition to reconstruct the redshift evolution of the AGN bolometric luminosity function since  $z \sim 3$ . The resulting trends are used to estimate the overall AGN accretion rate density at different cosmic epochs and to trace the first ever estimate of the AGN accretion history from an IR survey.

Later on, we focus our study on the connection between AGN accretion and integrated galaxy properties. The main scientific goal that we want to address is to measure the level of AGN accretion as a function of different galaxy properties of their host galaxies. We cross-match our *Herschel* sample with X-ray data taken from *Chandra* observations across the same fields mentioned above. The average X-ray luminosity (by combining both X-ray detected and stacked sources) is used as a proxy to estimate the average AGN accretion rate of *Herschel* galaxies. Iterating this analysis in different bins of SFR,  $M_{\star}$  and specific SFR (sSFR), we analyse the relationships of AGN accretion with galaxy properties in the SFR– $M_{\star}$  plane and at different cosmic epochs. Finally, we infer what is the parameter that best correlates with AGN accretion, comparing our results with previous studies and discussing their physical implications in the context of current scenarios of AGN/galaxy evolution.

This Thesis is structured as follows.

1. Chapter 1 summarizes the main physical and observational properties of AGN with respect to their hosts. Widely used diagnostics are described to identify “active” and “inactive” galaxies, according to different selection wavelengths. We introduce the reader to the most widely accepted AGN/galaxy evolutionary pictures of the last decade, intentionally focusing on the state of the art before the advent of the *Herschel* satellite.
2. We introduce the main *Herschel* surveys in Chapter 2, including the cross-match with multi-wavelength data-set. A discussion on the recent progresses made possible

by *Herschel* is also presented, highlighting the main discoveries and challenges that motivate the study carried on in this Thesis.

3. A detailed analysis of the SED-fitting decomposition performed in this Thesis is provided in Chapter 3. We also present the analysis done to evaluate the statistical significance of the AGN emission component in each observed SED. Later on, we compare our results with those obtained from independent AGN diagnostics. Finally, we describe some useful applications of this powerful technique, such as hunting for highly obscured AGN.
4. In Chapter 4, for each *Herschel* galaxy hosting an AGN, we show how we use the AGN bolometric luminosity to trace the AGN accretion history since  $z \sim 3$ . Our derivation is compared with previous estimates from the literature inferred from independent selection criteria (e.g. X-rays).
5. In Chapter 5 we investigate the mutual relationships of AGN accretion with integrated galaxy properties in a sample of *Herschel*-selected galaxies at  $0 < z \leq 2.5$ . In particular, we map the average AGN accretion rate in the SFR– $M_\star$  plane, as a function of SFR,  $M_\star$  and offset with respect to the “star-forming main-sequence”.
6. We present our concluding remarks and future prospects in Chapter 6.



# AGN vs star formation: the current evolutionary picture

During their lifecycle, galaxies may cross some “active” phases in which transient processes take place, producing great amounts of energy. These episodes may include gas accretion onto the central super massive black hole (SMBH), which is then revealed as an active galactic nucleus (AGN). On the other hand, cold molecular gas clouds may collapse and generate new stellar populations: these episodes are generally referred as “star formation activity”. It is currently believed that these two processes are intimately connected, despite the largely different spatial scales at which they operate. Still debated is the role of AGN activity in shaping the integrated properties of the host galaxies. On the one hand, theoretical models and hydrodynamical simulations require a powerful AGN phase as “deus ex machina” to stop subsequent star formation in the host. This scenario is also supported by several observations of massive AGN-driven outflows that could potentially shut down the galaxy star formation. On the other hand, studies of statistical samples of AGN host galaxies argue that the vast majority of known AGN are placed in star forming galaxies, showing somewhat enhanced star formation activity with respect to non-AGN hosts. These apparently conflicting statements reflect the fact that the physical mechanisms responsible for triggering and quenching these emission processes are still poorly understood.

Though this Thesis does not provide a conclusive explanation of these issues, we proceed to investigate the integrated properties of AGN and their hosts through a panchromatic approach. Indeed, both AGN and galaxy emission processes cover a wide range of the electromagnetic spectrum (from radio to hard X-rays). Therefore, exploiting multi-wavelength data is crucial to understand the physical and evolutionary properties of AGN host galaxies. In this Chapter we describe the panchromatic properties of galaxies hosting an AGN, introducing the reader to the current AGN/galaxy evolutionary picture inferred from optical, X-ray and infrared (IR) studies of these objects. We stress that the state-of-the-art picture discussed in the next Sections will intentionally overlook recent studies based on *Herschel* observations, as their contribution will be properly addressed in Chapter 2 and in the rest of this Thesis.

## 1.1 Active galactic nuclei: a panchromatic overview

“Active” galaxies<sup>1</sup> are characterized by the fact that a significant fraction of their bolometric luminosity does not arise from emission processes related to star-formation activity. Detailed studies of large samples of AGN started in the early 1960s, when the first quasi-stellar objects (hereafter Quasar or QSOs) were discovered. Though the detailed mechanisms that are the source of the powerful nuclear activity are still debated (Rees 1978; Lodato & Natarajan 2007; Devecchi & Volonteri 2009; Volonteri & Begelman 2010; Ball et al. 2011; Natarajan 2011), it is believed that it is mainly due to mass accretion onto a SMBH (Salpeter 1964; Lynden-Bell 1969; Shakura & Sunyaev 1973; Soltan 1982; Rees 1984), allowing them to be revealed as AGN.

In the following we briefly describe the inner structure of a typical AGN (Section 1.1.1 and 1.1.2). Later on, we recall widely used nomenclatures to classify AGN (Section 1.1.4), and discuss different multi-wavelength diagnostics that are commonly adopted to identify active galaxies (Section 1.1.7, 1.1.8 and 1.1.9).

### 1.1.1 Super massive black hole

The central engine of an AGN is believed to be a SMBH. The term “black hole” originally indicates a region of the space-time whose escape velocity exceeds the speed of light  $c$ . These extreme objects are described in the theory of general relativity through three parameters: mass ( $M_{\text{bh}}$ ), charge ( $Q$ ) and angular momentum ( $J$ ). The best evidence for the existence of a SMBH can be found in the centre of the Milky Way galaxy, where the mass of the Sgr A\* source was established to be  $\approx 4.4 \times 10^6 M_{\odot}$  (Genzel et al. 1997). The mass of SMBHs spans a wide range, from  $10^6$  up to  $10^{10} M_{\odot}$  (Lynden-Bell 1969; Lauer et al. 2007). Interestingly, some extremely massive BHs, as large as  $10^9 M_{\odot}$ , have been found even at redshift  $z \sim 7$  (Mortlock et al. 2011), where the Universe was less than 700 Myr old (assuming a standard flat cosmology<sup>2</sup>). This raises important questions related to the formation and the mass assembly of SMBHs since early epochs (e.g. Volonteri & Silk 2014 and references therein). Before growing their mass, there must be some pre-existing BH seeds. Largely recognized scenarios propose the BH seeds to generate from either remnants of Pop III stars (Haiman & Loeb 2001), or direct halo gas collapse (Begelman et al. 2006). The first hypothesis allows to get  $\approx 10^3 M_{\odot}$  seeds, while the latter one ends up with  $\approx 10^6 M_{\odot}$  BHs. Later on, these seeds need to increase their initial mass to become SMBHs. Plausible mechanisms postulate cold gas flow or hot gas accretion via major mergers. Current predictions of the SMBH mass function suggest that SMBHs are nearly ubiquitous in galaxies, but only a small fraction of them ( $\approx 10\%$ , e.g. Soltan 1982, Merloni et al. 2004) are accreting, while the rest are quiescent.

Here below we introduce a few simple concepts on gas accretion onto a SMBH that may benefit the following discussion (see Peterson 1997 for a comprehensive review).

<sup>1</sup>In this Thesis the terms “active” and “inactive” will be used in a broad sense to identify galaxies with or without AGN, respectively.

<sup>2</sup>In this Thesis we assume a flat cosmology with  $\Omega_m = 0.30$ ,  $\Omega_{\Lambda} = 0.70$ , and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

*Gravitational radius* ( $R_g$ ) – The size of a non-rotating black hole with mass  $M_{\text{bh}}$  is:

$$R_g = \frac{GM_{\text{bh}}}{c^2} \approx 1.48 \frac{M_{\text{bh}}}{10^8 M_\odot} 10^{13} \text{cm} \quad (1.1)$$

The gravitational radius  $R_g$  measures the last stable orbit of a BH, that is the minimum distance at which a photon (or whatever particle) is allowed to escape the potential well of the SMBH.

*Matter-to-radiation conversion efficiency* ( $\epsilon$ ) – It represents the fraction of gravitational energy involved in the gas accretion that is translated into photons. This parameter generally ranges between 5% and 10% (e.g. Marconi et al. 2004), but may formally reach 42% in case of maximally rotating (i.e. Kerr) BHs ( $J=1$ ).

*Eddington luminosity* ( $L_{\text{edd}}$ ) – It represents the maximum luminosity allowed for an object with mass  $M_{\text{bh}}$  that is powered by spherical accretion<sup>3</sup>. The Eddington luminosity is dependent on  $M_{\text{bh}}$  only, according to the expression:

$$L_{\text{edd}} \sim 1.26 \times 10^{38} \left( \frac{M_{\text{bh}}}{M_\odot} \right) [\text{erg s}^{-1}] \quad (1.2)$$

which implies a maximum threshold of the accretion rate  $\dot{M}_{\text{edd}}$ :

$$\dot{M}_{\text{edd}} = \frac{L_{\text{edd}}}{\epsilon c^2} \sim 2.2 \left( \frac{M_{\text{bh}}}{10^8 M_\odot} \right) \left( \frac{\epsilon}{0.1} \right)^{-1} [M_\odot \text{ yr}^{-1}] \quad (1.3)$$

There is a general consensus that most of the SMBH growth takes place during the AGN phase (Sanders et al. 1988). The inflowing material is arranged in a disk structure that surrounds the SMBH.

### 1.1.2 Accretion disc

The most efficient accretion processes are associated with accretion discs, which can provide the necessary mechanisms to transfer angular momentum outward and track the gas inward onto the vicinity of the BH. The first accretion disc model has been developed by Shakura & Sunyaev (1973) and is widely known as “optically thick geometrically thin accretion disk”. This simplified representation predicts the following radial profile of the gas temperature in the disc:

$$T(r) \propto \left( \frac{M_{\text{bh}}}{10^8 M_\odot} \right)^{-1/4} \left[ \frac{\dot{M}_{\text{bh}}}{\dot{M}_{\text{edd}}} \right]^{1/4} \cdot \left[ \frac{R}{R_g} \right]^{-3/4} \quad (1.4)$$

which returns  $T(r = 6 R_g) \simeq 10^5 \text{K}$  if assuming a BH mass as high as  $10^8 M_\odot$ . The optical and ultraviolet (UV) radiation emitted by the accretion disc is generally interpreted as a multi-colour black-body, therefore a superimposition of different Planckian curves, each

<sup>3</sup>The Eddington luminosity becomes a physical threshold only in the (not realistic) case of steady-state spherical accretion of fully ionized plasma, in a completely optically thin regime.

one corresponding to a given gas temperature. The following expression describes the monochromatic surface brightness of a black body at a given temperature  $T$ :

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \cdot \frac{1}{e^{h\nu/kT} - 1} \quad (1.5)$$

The (radiative) bolometric output of the AGN is:

$$L_{\text{bol,AGN}} = \epsilon \dot{M}_{\text{bh}} c^2 \text{ [erg s}^{-1}\text{]} \quad (1.6)$$

and can reach up to  $10^{48}$  erg s<sup>-1</sup>, covering 10 orders of magnitude in wavelength.

### 1.1.3 Broad and narrow line regions

Part of the radiation field originated from the accretion disc may pass through excited or ionized gas clouds ( $T \approx 10^{3-4}$  K) located from pc to kpc distances from the center (e.g. Peterson et al. 2004). These regions are commonly known as "broad line regions" (BLRs) and "narrow line regions" (NLRs).

- BLRs are clumpy structures made of ionized ( $T \sim \text{few } 10^4$  K) gas clouds. They are characterized by very broad ( $10^3 < \text{FWHM} < 10^4$  km s<sup>-1</sup>) emission lines, unambiguously ascribed to large scale motions of high electron density gas ( $n_e \sim 10^9$ – $10^{10}$  cm<sup>-3</sup>) in the vicinity (0.1–1 pc scales) of the SMBH. The broad lines emitted in these regions are mainly due to recombination or de-excitation of photo-ionized atoms, while thermal broadening has a negligible effect. Common features include lines from the Balmer and Lyman series (see Section 1.1.8).
- NLRs are constituted by clouds of ionized gas ( $T \sim 10^4$  K) emitting narrower lines ( $\text{FWHM} < 10^3$  km s<sup>-1</sup>, both permitted and forbidden<sup>4</sup>) compared to BLR. They are commonly associated to gas at lower densities ( $n_e \sim 10^4$ – $10^5$  cm<sup>-3</sup>) orbiting at larger distances (from 100 pc to kpc scales) from the SMBH. The typical morphology of NLRs has been found to resemble two opposite ionizing cones centered into the nuclear source (e.g. Wilson et al. 1985).

### 1.1.4 Classification of AGN

In this Section we briefly recall a few important distinctions between classes of AGN, based on historical studies of optical and radio selected samples of AGN. As usual, on the basis of the properties considered it is also possible to associate the same object to different classes. We refer the interested reader to Peterson (1997) for a comprehensive review of the AGN taxonomy.

#### Radio galaxies

Radio-galaxies are usually identified with massive elliptical galaxies and characterized by the presence of radio-emitting jets and lobes extending up to Mpc scales from the central

<sup>4</sup>The terms "permitted" and "forbidden" lines refer to typical timescale at which the transition producing a given line occurs, with respect to the Hubble time.



region. Different classifications have been proposed in the past decades to distinguish between radio-emitting sources, mostly based on radio power and morphological properties. Fanaroff & Riley (1974, FR hereafter) claimed that radio-luminous AGN consist of two apparently distinct types of sources, referred to as FR types I and II, with prominent differences in both radio morphology and luminosity.

FR I have  $L_{1.4} \lesssim 10^{25} \text{ W Hz}^{-1}$ . Their radio morphology shows core dominated structures, with symmetric radio jets.

FR II have  $L_{1.4} \gtrsim 10^{25} \text{ W Hz}^{-1}$ . They are edge-brightened with highly collimated large-scale jets.

An alternative way of classifying radio AGN is based on the level of excitation of their optical emission lines (Hine & Longair 1979; Laing et al. 1994). Smolčić et al. (2009) distinguished radio galaxies between high-excitation and low-excitation radio AGN (HERAGN and LERAGN, respectively). They found that LERAGN are more common at low radio luminosities: almost all FR I radio galaxies are LERAGN, while optical hosts of the most powerful radio sources, i.e., FR IIs, usually have strong emission lines.

## Radio quasars

Historically, the term “quasar” indicates a point-like source (“quasi-stellar radio-source”, or “quasi-stellar object”, QSO), and it is currently associated to extremely luminous objects. In the radio, they are classified according to the radio loudness parameter,  $R = F_r/F_o$ , that is defined as the ratio between radio (5 GHz) and optical (B-band at 4400Å) flux densities. The nominal threshold is set to  $R \approx 10$  (Kellermann et al. 1989).

*Radio-loud quasars*, are characterized by a strong radio emission and an unresolved luminous nucleus ( $M_V < -23$ ). Their optical spectra show very broad ( $\text{FWHM} > 10^3 \text{ km s}^{-1}$ ) emission lines.

*Radio-quiet quasars*, constituting 90% of the global quasar population, show optical spectra similar to radio-loud quasars, but intrinsically lower radio power. This is the reason why they are considered as scaled-down analogs of radio-loud quasars in radio power.

## Optical classification

*Type-1 AGN* are characterized by the presence of both (permitted) broad ( $\text{FWHM} > 10^3 \text{ km s}^{-1}$ ) and (permitted and forbidden) narrow emission lines in their optical/UV spectra. This class includes Seyfert-1 galaxies (Seyfert 1943; Khachikian & Weedman 1974) and the brighter type-1 QSOs, that are distinguished on the basis of their intrinsic power (i.e.  $M_V < -23$ ). As a consequence, Seyfert-1s appear as a quasar-like nucleus, but with the host galaxy clearly detectable (Peterson 1997). X-ray spectra of type-1 AGN typically show prominent soft-excess at  $\approx 1 \text{ keV}$  likely originated from the accretion disc, superimposed to a Comptonization power-law.

*Type-2 AGN* show only narrow lines ( $\text{FWHM} \leq 10^3 \text{ km s}^{-1}$ ) in their optical/UV spectra, both permitted (e.g. Balmer series) and forbidden (e.g. [O III], [Ne III] and [O II]), with similar strengths compared to type-1 objects. The QSO-to-Seyfert dichotomy still separates type-2 QSOs from Seyfert-2 on the basis of their intrinsic luminosity. X-ray spectra of type-2 AGN are flatter (i.e. harder) than those observed in type-1 AGN.

### Blazars

Blazars identify a class of highly polarized and/or rapidly variable AGN ( $t \sim \text{days}$ ), including:

*BL Lacertae* objects, which are very bright sources in the radio band, typically found in elliptical galaxies. They show weak (or absent) emission and absorption lines in the optical spectra. They are highly variable in optical, radio and X-ray bands ( $t \sim \text{days}$ ).

*Optically Violent Variables* (OVVs), similar to the BL Lacertae objects, except for the presence of broad emission lines in their optical spectra, although fainter than those found in QSOs.

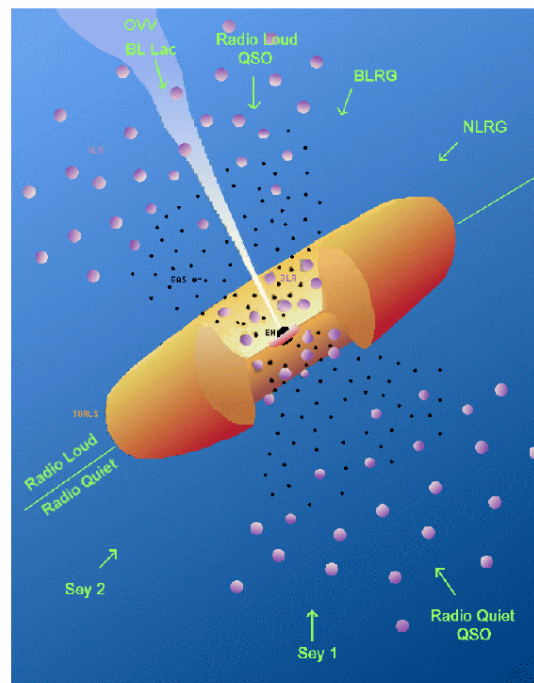
### LINERs

*Low Ionization Nuclear Emission-line Regions* (LINERs) identify a class of active galaxies with intrinsically low nuclear luminosity, but showing strong, low-ionization, emission lines in their optical spectra. LINERs share most of the properties shown by Seyfert galaxies, except for the higher intensity of narrow lines with lower level of ionization.

#### 1.1.5 The Unified model of AGN

Several authors have discussed the need for a self-consistent picture to explain most of the observational differences mentioned in Section 1.1.4 that come out of different classifications. Some formulations of the “unified model of AGN” have been presented by Antonucci (1984), Antonucci & Miller (1985), Urry & Padovani (1995) for optically-identified AGN (but see also Barthel 1989 for an equivalent unification in the radio among radio galaxies and quasars). The main argument is that Seyfert-1 and Seyfert-2 galaxies are intrinsically the same sources, but seen under different viewing angles with respect to the line of sight. As a consequence, the observational differences are purely ascribed to orientation effects relative to an obscuring medium absorbing part of the nuclear emission. The first strong argument in support of this scenario was presented by Antonucci & Miller (1985), who surprisingly discovered the presence of broad optical lines in the polarized spectrum of the Seyfert-2 galaxy NGC 1068. Broad lines had never been observed before in type-2 AGN: in the case of NGC 1068, the presence of these lines in the polarized spectrum was attributed to scattering of free electrons.

As shown in Fig. 1.1, a toroidal structure (of  $\sim$  a few to 100 pc in size) was the simplest configuration that could be postulated to be able to obscure the BLR, but not the NLR, and to explain the differences between type-1 and type-2 AGN simply in terms of orientation angle.



**Figure 1.1:** Spatial structure and components of a typical AGN (M. Polletta’s PhD Thesis, Polletta96, adapted from Urry & Padovani95), according to the Unified model of AGN (Section 1.1.5).

Studies based on IR and X-ray observations of AGN have argued that the “dusty torus” should be made of cold gas and dust to explain the attenuation of the optical/UV and soft (0.5–2 keV) X-ray emission. Indeed, hard ( $>2$  keV) X-ray observations of AGN have detected most of local type-2 AGN that were missed in the soft-band (0.5–2 keV), where photo-electric absorption from cold gas is more important. X-ray spectral analysis of these objects have also shown very strong Fe  $K\alpha$  line at rest 6.4 keV with large equivalent widths ( $EW \approx 500$  eV, Matt et al. 1999), interpreted as due to Compton reflection from a neutral and optically thick medium. On the other hand, studies based on mid-IR (MIR) spectra of local Seyfert-2 galaxies (Siebenmorgen et al. 2005; Netzer et al. 2007) have shown compelling evidence for the presence of optically thick<sup>5</sup> dust along the line of sight, through the detection of the 9.7  $\mu\text{m}$  silicate feature in absorption. The depth of the silicate feature is observed to correlate (although with large scatter) with the hydrogen column density ( $N_{\text{H}}$ ) measured from X-ray spectra, suggesting that cold gas absorbing X-rays and circumnuclear dust might be co-planar. A more detailed description of the radiation processes that are necessary to interpret dust emission in the IR is provided in Section 1.1.9.

<sup>5</sup>This means that the optical depth measured at visual wavelengths  $\tau_V \geq 1$ .

## Torus geometry

The torus geometry that was first implemented in models of AGN tori (e.g. Barvainis 1987; Pier & Krolik 1992; Granato & Danese 1994; Efstathiou & Rowan-Robinson 1995) was a “smooth” structure, meaning that the planar distribution of circumnuclear dust grains was assumed axisymmetric and their properties (temperature, density) had to follow monotonic radial profiles. On the other hand, more complex and realistic models (Nenkova et al. 2002, 2008; Schartmann et al. 2005, 2008) have proposed “clumpy” structures, which arrange dust grains in several clouds with substructures along the line of sight.

Observations based on NIR and MIR interferometry have provided the *direct* imaging of the torus structure (Swain et al. 2003; Wittkowski et al. 2004; Jaffe et al. 2004). These studies have proved unambiguously that AGN tori indeed exist, and that their geometry is consistent with a clumpy structure. Moreover, the fact that the  $9.7 \mu\text{m}$  silicate feature is observed also in emission (typically in Seyferts 1, but also in a few Seyferts 2, Siebenmorgen et al. 2005), is predicted if assuming a clumpy absorbing medium.

Nonetheless, recent studies (e.g. Feltre et al. 2012, 2013) comparing smooth vs clumpy models to multi-wavelength dataset (both spectroscopic and photometric) in non-local ( $z > 0.1$ ) AGN have claimed that none of these two configurations is significantly favoured with respect to the other.

## Torus size

Conversely to the original torus size postulated in AGN models ( $\leq 100$  pc, see Pier & Krolik 1992; Granato & Danese 1994), recent observations have constrained the torus size to be much more compact than expected before. Estimates obtained with the MID-infrared interferometric Instrument (MIDI) mounted on the Very Large Telescope Interferometer (VLTI) and the VLT Imager and Spectrometer for mid Infra Red (VISIR) camera, have suggested the outer radius  $R_{\text{out}}$  to be equal or less than 10 pc. A few benchmark cases include: the Circinus galaxy ( $R_{\text{out}} < 1$  pc, Tristram et al. 2007); NGC 1068 ( $R_{\text{out}} \sim 1\text{--}2$  pc, Jaffe et al. 2004; Raban et al. 2009); Centaurus A ( $R_{\text{out}} < 0.3$  pc, Meisenheimer et al. 2007); NGC 3281 ( $R_{\text{out}} < 11$  pc, Sales et al. 2011). This compilation has been recently extended to other 23 nearby galaxies (Burtscher et al. 2013), showing comparable sizes ( $0.1 < R_{\text{out}} < 10$  pc). Further constraints on the torus geometry have been inferred through high-resolution spectroscopy at  $\sim 12 \mu\text{m}$  obtained with the VISIR spectrograph. These observations have allowed to constrain the inner radius  $R_{\text{in}}$  of the torus (Kishimoto et al. 2007; Hönig et al. 2010) for about 20 Seyfert galaxies, suggesting  $R_{\text{in}}$  measurements of the order of  $\approx 0.1$  pc, which imply  $R_{\text{out}}/R_{\text{in}}$  of the order of 20–30. We stress that these measurements are currently available for nearby Seyferts, but they do not necessarily hold in high-redshift quasars.

### 1.1.6 Challenging the Unified Model

The unified model of AGN has been debated in the recent literature, as several observations appear to be in conflict with some fundamental predictions of the unified model, e.g. questioning that type-1 and type-2 AGN are intrinsically the same sources seen at different

viewing angles. Several studies are suggesting that a self-consistent picture of the AGN taxonomy does not depend only on orientation effects, but also on the intrinsic luminosity of the source and on the redshift.

X-ray based studies claim that nuclear obscuration is a strong function of nuclear luminosity (Ueda et al. 2003; Steffen et al. 2003; La Franca et al. 2005; Hasinger 2008; Burlon et al. 2011; Lusso et al. 2013). The covering factor<sup>6</sup> of the torus may play an important role. While the original unified model postulates the torus covering factor not to change among type-1 and type-2 AGN, recent studies have shown a clear dependence on the intrinsic luminosity of the source. This interesting finding has been interpreted within the context of the so-called “receding torus” scenario (Simpson 2005). This model predicts a luminosity-dependent covering factor, due to the recession of the sublimation radius  $R_{\text{subl}}$  of the torus. Circumnuclear dust recedes in presence of a more luminous heating source, that sublimates circumnuclear dust at increasing radii, following the relation  $L_{\text{bol,AGN}} \propto R_{\text{subl}}^{0.5}$  (Lawrence 1991).

While such luminosity-dependence is well established from recent X-ray observations, its possible evolution with redshift is still debated (La Franca et al. 2005; Hasinger 2008; Hiroi et al. 2012; Merloni et al. 2014). Observations of low-luminosity AGN (LLAGN,  $L_{\text{bol,AGN}} < 10^{42} \text{ erg s}^{-1}$ ) seem to confirm the trend predicted by the receding torus model. Indeed, the fraction of unobscured sources has been found to decrease at low AGN accretion rates (Nicastro et al. 2003; Hopkins et al. 2009). A possible explanation for the absence of broad lines could be the disappearance of the BLR at low AGN accretion rates (e.g. Elitzur & Shlosman 2006).

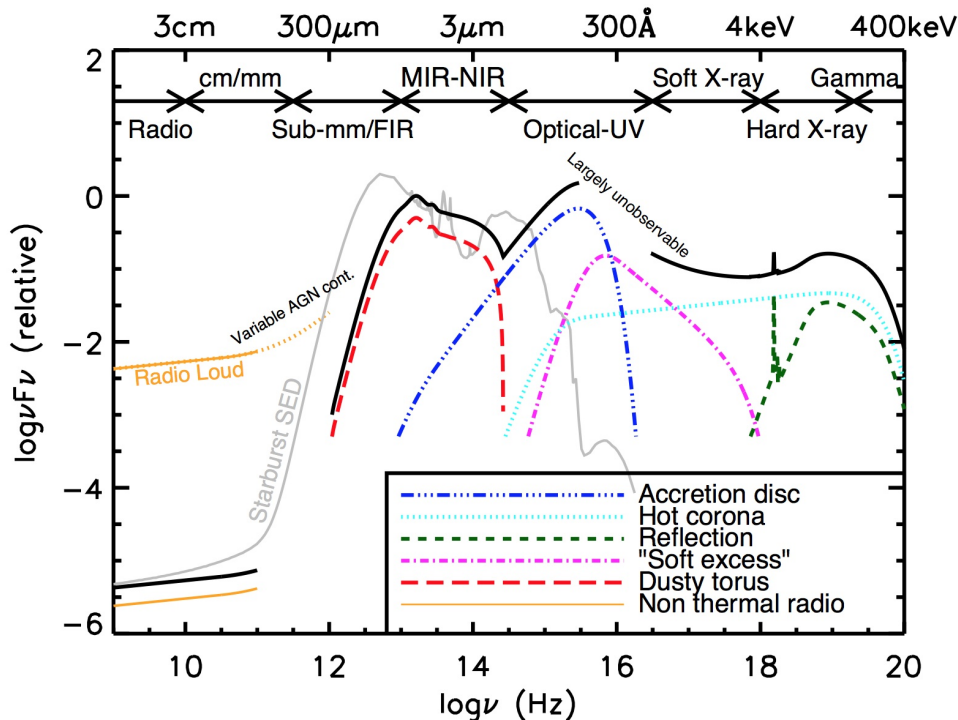
In the near future, the Atacama Large Millimeter Array (ALMA) is expected to address these issues through high-angular resolution observations in the sub/mm regime. Indeed, ALMA will be able to detect emission from cold ( $T \sim 10 \text{ K}$ ) dust clumps in AGN tori (Maiolino 2008), therefore enabling to discriminate between smooth and clumpy models.

In the light of these studies, it has become clear that the structure of a typical AGN (see Fig. 1.1) is complex and includes different components. Therefore, a multi-wavelength study of AGN emission is necessary to interpret their physical properties. In the following paragraphs, we describe panchromatic diagnostics and spectral features that allow to characterize the spectral energy distributions (SEDs) of AGN host galaxies. Fig. 1.2 shows a broad-band SED of a radio-quiet QSO (black solid and coloured dashed lines). Different colours mark specific emission processes that arise from different spatial regions of the AGN, as shown in Fig. 1.1. The light-grey curve represents a typical spectrum of a starburst galaxy (M82, taken from Silva et al. 1998) for comparison.

### 1.1.7 X-ray observations of AGN hosts

AGN are nearly ubiquitous X-ray emitters (Elvis et al. 1978). *Chandra* and XMM-Newton observations at 0.5-10 keV have truly revolutionized our current view of AGN, by revealing up to  $\approx 7000 \text{ AGN deg}^{-2}$  in the deepest pointings (at flux limits of  $\approx 2.5 \times 10^{-17} \text{ erg s}^{-1}$

<sup>6</sup>The covering factor is the fraction of the entire solid angle that is occupied by the torus structure.



**Figure 1.2:** Schematic representation of a multi-wavelength AGN SED, from radio to  $\gamma$ -rays (black solid line). Different colours mark specific emission processes that arise from different spatial regions. The light-grey line identifies a typical spectrum of a starburst galaxy for comparison. Taken from C. Harrison’s webpage: <http://astro.dur.ac.uk/~cpnc25/research.html>.

$\text{cm}^2$ , see Bauer et al. 2004). Fig. 1.2 highlights that the AGN X-ray spectrum shows several features. While the soft ( $<2$  keV) X-ray spectrum is dominated by the thermal emission of the accretion disc (also referred as “soft-excess”), non-thermal processes are necessary to reproduce the observed hard (2-100 keV) X-ray spectrum. Hard X-ray emission is attributed to “seed” photons emitted from the accretion disc (or from outer regions), subsequently up-scattered through Inverse Compton (cyan dotted line) from the electrons placed in the surrounding “hot corona”. This process reproduces the observed power-law spectrum with a cut-off at  $\approx 100$  keV:

$$F_\nu \propto \nu^{1-\Gamma}, \quad (1.7)$$

where  $F_\nu$  is the monochromatic X-ray flux and  $\Gamma$  is the photon index, that is on average  $\Gamma = 1.8 \pm 0.2$  (Tozzi et al. 2006; Mainieri et al. 2007). In addition, photo-electric absorption may drop down the X-ray emission at softer energies, producing a flatter observed X-ray spectrum. The comparison between the observed and the expected photon index is widely adopted to estimate the hydrogen column density ( $N_{\text{H}}$ ) of the source along the line of sight (e.g. Luo et al. 2008; Xue et al. 2010). Part of the radiation reprocessed by the “hot corona” may be back-scattered to the accretion disc or to neutral gas regions, producing other features such as the  $K\alpha$  fluorescence Fe line (at rest 6.4 keV) and the Compton reflection bump (peaking at  $\approx 30$  keV, green dashed line of Fig. 1.2).

X-ray observations provide reliable identification of AGN, since they are weakly affected

by the host galaxy light. Typical AGN diagnostics that are used to identify AGN in X-ray selected samples are described in Xue et al. (2011), where a source is classified as AGN if it satisfies at least one of the following criteria:

1. the full-band X-ray luminosity  $L(0.5\text{--}8\text{ keV})$  is larger than a typical threshold, such as  $3 \times 10^{42}\text{ erg s}^{-1}$ . This value is safely unaffected by the potential X-ray emission due to star formation (i.e. X-ray binaries);
2. the X-ray-to-optical flux ratio  $\log(F_X/F_R) > -1$  in one X-ray band at least;
3. the observed photon index  $\Gamma \leq 1.0$ , which identifies an obscured AGN;
4. the full-band X-ray luminosity exceeds by a factor of  $\geq 3$  the average X-ray emission expected from an “inactive” starburst galaxy (Alexander et al. 2005);
5. the presence in optical/UV spectra of broad ( $\text{FWHM} > 10^3\text{ km s}^{-1}$ ) optical emission or absorption lines (see Section 1.1.8).

These AGN diagnostics provide robust indicators of AGN activity. However, as stated by the authors, these criteria might fail in identifying heavily obscured AGN, for which a complementary approach is necessary (see Section 1.1.9 and 3.6.4).

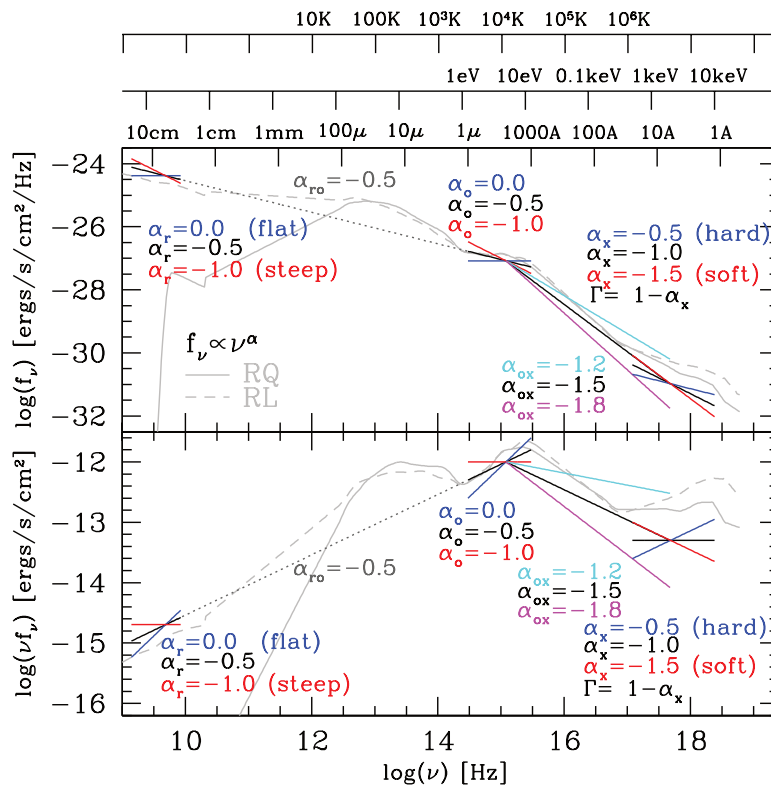
### 1.1.8 Optical/UV continuum and emission lines

The continuum emission observed in the UV spectrum of type-1 quasars is generally called “big-blue bump”, or “UV-excess (Shields 1978), which is interpreted as thermal radiation from the accretion disc, with gas temperatures around  $10^5\text{ K}$  (Malkan & Sargent 1982). This emission component irradiates the bulk of the bolometric output of the AGN, even outshining the entire luminosity of the galaxy in quasar hosts. As a consequence, the shape and intensity of the optical continuum has been extensively used as an effective selection criterion to identify relatively bright and unobscured QSOs (Richards et al. 2009). Fig. 1.3 from Richards et al. (2006) shows the mean SEDs of radio-loud (solid line) and radio-quiet (dashed line) quasars (from Elvis et al. 1994). Different coloured lines highlight spectral indices at different wavelengths, according to the sign convention adopted in Eq. 1.1.7, with  $\alpha = \Gamma - 1$ . This plot shows that a noticeable scatter exists among intrinsic quasar SEDs. For instance, the relative amount of X-ray and optical emission, that is parameterized by the X-ray-to-optical slope  $\alpha_{\text{ox}}$  (Zamorani et al. 1981),

$$\alpha_{\text{ox}} = -\log\left(\frac{L_{2\text{keV}}/L_{2500}}{\nu_{2\text{keV}}/\nu_{2500}}\right) = -\log\left(\frac{L_{2\text{keV}}/L_{2500}}{2.605}\right) \quad (1.8)$$

typically ranges between 1.2 and 1.8, with a mean value of  $\approx 1.5$ .

It is worth mentioning that the observed slope of the optical/UV continuum could be significantly altered by attenuation due to both interstellar dust of the host galaxy and intrinsic obscuration of the source. In addition, except for bright quasars, the host galaxy light might significantly contaminate the estimate of the AGN bolometric luminosity at these wavelengths. Because of these aspects, optical/UV selection of AGN is used only for very bright and unobscured quasars.

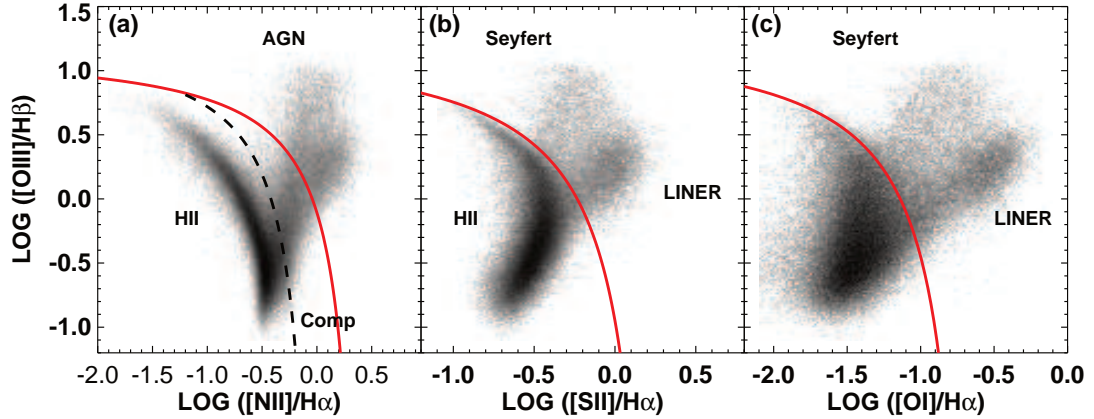


**Figure 1.3:** Quasar SED diagnostic plot. The same SEDs are shown in monochromatic flux (*top panel*) and integrated flux (*bottom panel*), as a function of frequency, wavelength, temperature (of the emitting medium, in case of thermal radiative process) and energy. Light-grey curves are the Elvis et al. (1994) radio-quiet (solid) and radio-loud (dashed) mean SEDs. The colored lines indicate typical spectral indices in the radio, optical, and X-ray using the same sign convention. See the text for details.

Physically motivated diagnostics often rely on optical spectroscopy to identify AGN. As mentioned in Section 1.1.3, very broad emission lines emitted by the BLR enable to unambiguously identify an AGN. Nonetheless, in most cases these broad lines are not detectable from optical spectra (due to obscuration, low-quality spectra, intrinsically low nuclear luminosity). This is the reason why several optically-based AGN diagnostics rely on narrow line spectroscopy. However, individual narrow lines themselves do not allow to distinguish between active and inactive galaxies, as the ionizing continuum may originate from both star formation and AGN activity.

A step forward was made by Baldwin et al. (1981), who introduced the widely-known BPT diagram. They suggested to use emission line ratios between lines with a small separation in wavelength to minimize the effect of differential extinction. Among the most useful emission line ratios there are:  $[\text{O III}]/[\text{O II}]$  that is sensitive to the ionization level, and  $[\text{O I}]/\text{H}\alpha$  that measures the hardness of the radiation field. The combination of these line ratios provides a powerful tool to distinguish between inactive and active galaxies. Later on, Kewley et al. (2006) applied this technique to Sloan Digital Sky Survey (SDSS) galaxies, complemented with other emission line ratios, such as  $[\text{O III}]/\text{H}\beta$ ,  $[\text{N II}]/\text{H}\alpha$  and  $[\text{S II}]/\text{H}\alpha$  (see Fig. 1.4). These additional features have refined and extended the original





**Figure 1.4:** BPT diagrams to identify pure star forming galaxies and AGN hosts on the basis of their rest-frame optical emission lines (Kewley et al. 2006). Black dots are SDSS galaxies with  $S/N > 3$ . The red lines mark the Kewley & Dopita (2002) extreme starburst line; while the dashed line identifies H II-region like galaxies, according to Kewley et al. (2004).

BPT technique to studies of higher redshift and lower luminous AGN, including LINERs.

Nevertheless, these emission lines diagnostic are very time demanding and need typical AGN-like spectral features, that are available only for relatively bright and nearby objects. In addition, optical spectroscopy can detect the above-mentioned narrow lines up to  $z \sim 0.5$ , reaching at most  $z \sim 1$  objects if supplemented with near-infrared (NIR) spectra. Other diagrams, using only lines in the blue part of the optical spectra (e.g. [O II], [O III],  $H\beta$ ) have been proposed as alternative diagnostics for higher redshift ( $z \lesssim 1$ ) galaxies (e.g. Rola et al. 1997).

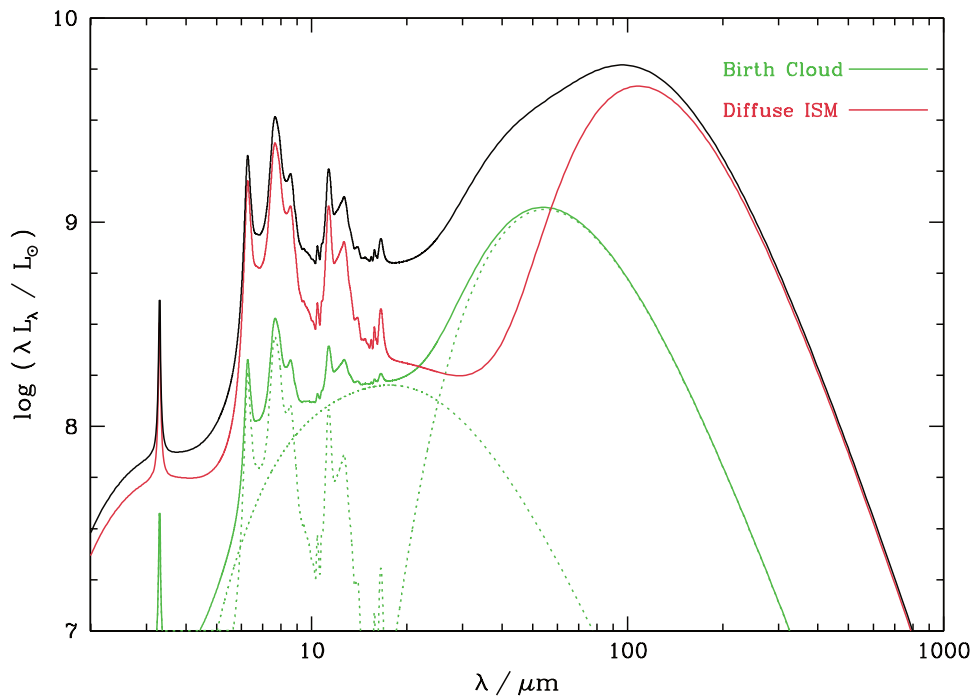
### 1.1.9 Infrared observations of AGN hosts

In this Section we briefly discuss the primary emission mechanisms that contribute to the global IR emission of star forming AGN hosts. Later on, we proceed to discuss some commonly used IR-based diagnostics to identify AGN. Additional aspects of IR SEDs of star forming galaxies, both with and without an AGN component, will be described in more detail in Section 2.4.1 and 3.2.

#### Modelling the IR SED in star forming galaxies

The IR spectral window covers the range 1–1000  $\mu\text{m}$ . The IR domain can be split in NIR (1–3  $\mu\text{m}$ ), MIR (3–30  $\mu\text{m}$ ) and far-IR (FIR, 30–1000  $\mu\text{m}$ ).

While the emission at NIR wavelengths is mostly originated from evolved ( $>1$  Gyr) stellar populations, MIR and FIR observations are more sensitive to reprocessed radiation from dust grains. In galaxies with on-going star-formation, young stellar populations are generally enshrouded in dense dust clouds (also called "birth clouds"), in turn surrounded more diffuse inter-stellar medium (ISM). Studies of the Milky Way and nearby galaxies in the Local Group have allowed to constrain the physical and chemical properties of dust



**Figure 1.5:** Example of IR SED template of a typical star forming galaxy (taken from da Cunha et al. 2008). The black curve indicates the integrated IR emission in the range  $2 < \lambda < 1000 \mu\text{m}$ , which is the sum of the emission components marked with the red and green curves. In particular, the red curve shows the contribution of dust emission placed in the ISM. The green solid curve refers to the dust emission from the stellar birth clouds, in turn decomposed (dashed curves) between PAHs, hot MIR continuum and warm dust in thermal equilibrium.

grains, and to calibrate models to interpret the origin of dust emission (e.g. da Cunha et al. 2008; Noll et al. 2009; Casey et al. 2012, see also the comprehensive review by Casey et al. 2014). Here we describe the simple model presented by da Cunha et al. (2008) (see Section 3.2.1), that reproduces the total IR (1–1000  $\mu\text{m}$ ) emission in star forming galaxy SEDs as the sum of different emission components.

(i) polycyclic aromatic hydrocarbons (PAHs, Leger & Puget 1984; Puget & Leger 1989; Allamandola et al. 1989) produce very strong emission features at 3.3, 6.2, 7.7, 8.6, 11.3 and 12.7  $\mu\text{m}$ . These molecules are probably placed in the outskirts of HII regions, where they can still survive the ionizing radiation field from young stars (Cesarsky et al. 1996).

(ii) MIR continuum emission due to hot ( $T \sim 100\text{-}200 \text{ K}$ ) and small dust grains (typical sizes reach 0.01  $\mu\text{m}$ ), that are stochastically heated by the absorption of a single UV photon (Sellgren 1984). They reproduce a smooth continuum emission at MIR wavelengths ( $3 < \lambda < 40 \mu\text{m}$ ).

(iii) FIR continuum emission from colder ( $T \sim 15\text{-}60 \text{ K}$ ) dust grains in thermal equilibrium. This emission component typically contributes for most of the total IR luminosity. The FIR continuum is typically modelled using two types of dust grains: warm grains placed

both in birth clouds and spread in the ISM, with temperatures of  $T \sim 30\text{--}60$  K; cold grains only placed in the ISM and with temperatures of  $T \sim 15\text{--}25$  K.

The most simplistic way of modelling dust emission is the blackbody fit (Eq. 1.5), which is simply a function of the temperature. However, a more realistic parametrization, that accounts for the non-perfect reflection and non-uniform opacity of galaxy’s dust, describes the flux density  $S_\nu$  as a grey-body spectrum (or “modified black-body”) of the form:

$$S_\nu \propto (1 - e^{-\tau_\nu}) B_\nu(T) = \frac{(1 - e^{-\tau_\nu}) \nu^3}{e^{h\nu/kT} - 1} \quad (1.9)$$

where  $S_\nu$  is in units of  $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$ .  $B_\nu(T)$  is the perfect black-body curve, which applies under the limit  $\tau_\nu \rightarrow +\infty$ . The optical depth  $\tau_\nu$  scales as  $\nu^\beta$  (Draine & Allaf-Akbari 2006), where  $\beta$  is the spectral emissivity index, and typically ranges between 1 and 2.5 (Casey et al. 2011).

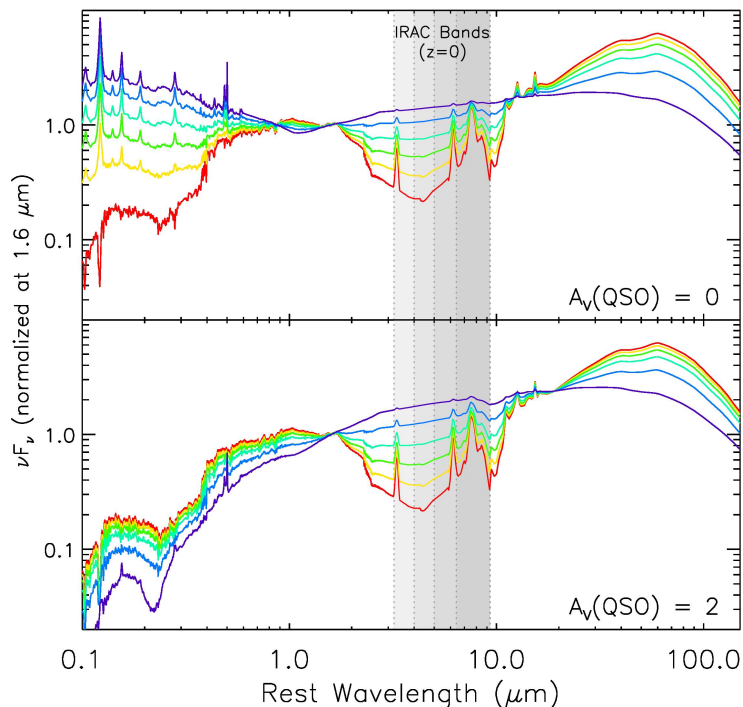
Fig. 1.5 shows a typical IR SED template in the range  $2 < \lambda < 1000 \mu\text{m}$ , computed by considering the emission components mentioned above. The black curve represents the integrated IR emission of the galaxy, that is divided in different sub-components. The red curve shows the IR contribution of the diffuse ISM, from both PAHs (at  $\lambda \sim 3\text{--}10 \mu\text{m}$ ) and cold dust emission, which peaks at  $\lambda \approx 100 \mu\text{m}$  and dominates the IR energy output of the galaxy. The green solid curve refers to warmer dust emission coming from the stellar birth clouds, while green dashed curves show the breakdown of this contribution in different components: PAHs (peaking at  $\lambda \sim 7.7 \mu\text{m}$ ), hot MIR continuum (peaking at  $\lambda \sim 20 \mu\text{m}$ ) and warm dust (peaking at  $\lambda \sim 50 \mu\text{m}$ ) in thermal equilibrium.

In active galaxies, an additional dust emission component has to be taken into account in modelling the IR SEDs. Indeed, the thermal emission originated from the accretion disc, peaking at optical/UV wavelengths and rapidly decreasing at  $\lambda > 1 \mu\text{m}$  (Richards et al. 2006), can be partially (around 30%, e.g. Risaliti & Elvis 2004; Treister et al. 2008) reprocessed by warm ( $T \sim 300\text{--}1000$  K) dust grains placed at circumnuclear ( $\approx$  a few pc) distances from the SMBH (Glikman et al. 2006), and re-emitted from 3 to  $100 \mu\text{m}$  as a grey-body spectrum peaking at  $\lambda \sim 10\text{--}20 \mu\text{m}$  (Nenkova et al. 2002; Fritz et al. 2006).

### IR observations and diagnostics to identify AGN

The complex IR SED of (active) galaxies has motivated several studies with the purpose of developing MIR-based diagnostics that could be used to identify AGN. Here we briefly discuss the AGN selection criteria proposed by Donley et al. (2012) and the comparison with previous MIR diagnostics.

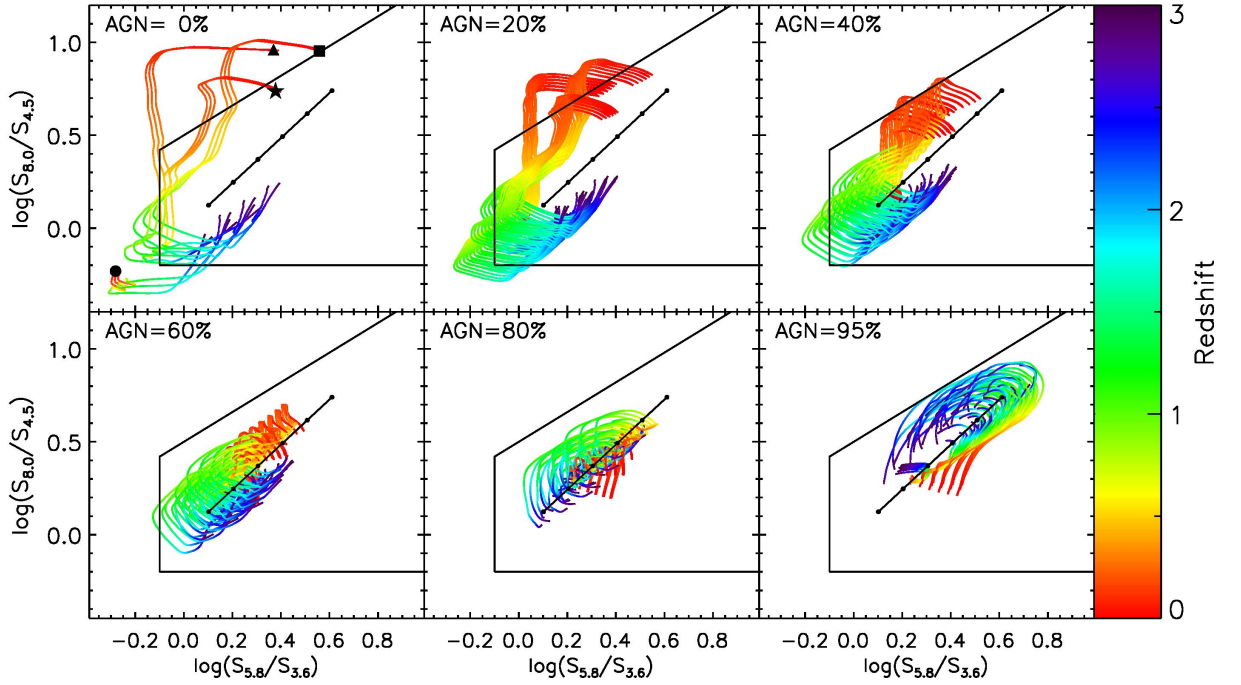
Fig. 1.6 shows composite (AGN+starburst) SEDs constructed by Donley et al. (2012), using stellar templates by Dale & Helou (2002), the starburst M82 and the QSO1 templates from the library of Polletta et al. (2007), scaled to give 1– $10 \mu\text{m}$  AGN contributions of 0% (red) to 95% (purple). The final SEDs have been normalized at  $1.6 \mu\text{m}$ . Upper and lower panels show the same templates but changing the extinction of the QSO1 model from  $A_V=0$  to  $A_V=2$ . While luminous unobscured and obscured AGN have very different



**Figure 1.6:** Composite SEDs from Donley et al. (2012), obtained by combining the QSO1 and M82 templates of Polletta et al. (2007, other curves) with the stellar templates taken from Dale & Helou (2002) to build self-consistent UV-to-FIR SEDs. The same AGN+starburst template, normalized to  $1.6 \mu\text{m}$ , is scaled to give a  $1\text{--}10 \mu\text{m}$  AGN contribution from 0% (red) to 95% (purple). Upper and lower panels show the same templates but changing the extinction of the QSO1 model from  $A_V=0$  to  $A_V=2$ . The four IRAC bands at  $z=0$  are shaded.

optical/UV SEDs, they are expected to show a so-called “mid-IR excess” irrespectively of the level of nuclear obscuration (Daddi et al. 2007; Fiore et al. 2009). The detection of this extra hot-dust component is a widely used diagnostic to isolate AGN from inactive star-forming galaxies. NIR emission due to stellar light generally peaks (in  $\nu F_\nu$ ) at rest-frame  $1.6 \mu\text{m}$ , decreasing at longer wavelengths, as shown in Fig. 1.6. Most of IR-based criteria to identify active galaxies use the *InfraRed Array Camera* (IRAC) wavebands at  $3.6$ ,  $4.5$ ,  $5.8$  and  $8.0 \mu\text{m}$ , as well as the *Multiband Imaging Photometer for Spitzer* (MIPS) at  $24 \mu\text{m}$ .

Fiore et al. (2008, 2009) used these filters to unveil a large population of Compton-Thick AGN candidates. They studied the X-ray properties of a sample of galaxies with high MIR-to-optical flux ratio ( $F_{24}/F_R > 1000$ ) and extremely red colours ( $R - K > 4.5$ ). Interestingly, only 25–30% of  $F_{24} > 0.55$  mJy sources had a counterpart in the *Chandra*-COSMOS catalogue (Elvis et al. 2009; Civano et al. 2012). X-ray stacking confirmed that most of them ( $\approx 80\%$ ) were consistent with being intrinsically obscured AGN (Fiore et al. 2009; Treister et al. 2009). This technique is very powerful in chasing highly obscured AGN, but is efficient for relatively bright sources in the MIR. Another major issue related to IR-based AGN selection techniques is that star forming galaxies can be very luminous at IR wavelengths and can outshine the central emission, affecting the reliability of the

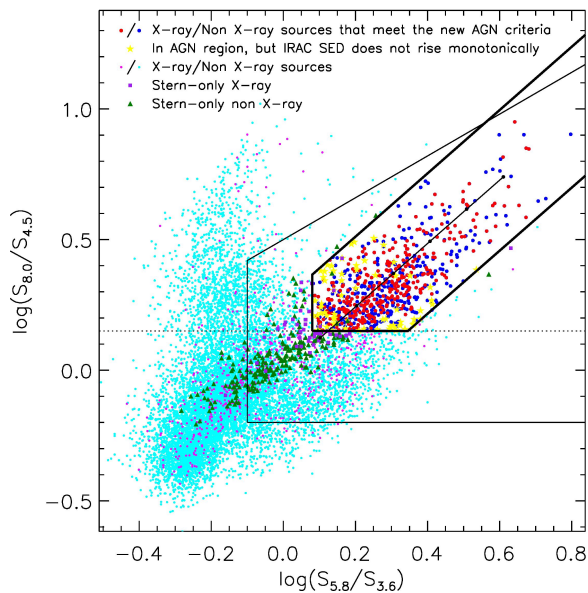


**Figure 1.7:** Tracks of different galaxy template SEDs with colour coded redshift. Different panels simulate how SED templates with increasing 1–10  $\mu\text{m}$  AGN emission move across the IRAC colour space with increasing redshift.

IR-based classification (Cardamone et al. 2008).

Other IR-based AGN diagnostics adopt IRAC colours to identify powerful and MIR-dominated AGN hosts (Lacy et al. 2004, 2007; Stern et al. 2005). Fig. 1.7 (taken from Donley et al. 2012) shows the wedge in the colour-colour diagram traced by Lacy et al. (2007) to identify AGN. Different panels mark the location of MIR colours of galaxy SEDs (shown in Fig. 1.6) with increasing AGN emission. Coloured lines simulate the tracks of each SED template as a function of redshift (in colour-coding). This plot shows that luminous AGN occupy a well-defined region in the IRAC colour space. However, also star-forming galaxy templates (first panel) fall in the same selection wedge (Lacy et al. 2004, 2007, Stern et al. 2005). This contamination issue gets worse with increasing redshift, since the galaxy stellar emission is shifted into the IRAC bands, producing similar IRAC colours to those of AGN at lower redshifts.

Donley et al. (2012) revisited the previous AGN selection criteria to minimize the contamination from “pure starbursts” without an AGN. Fig. 1.8 shows the Lacy et al. (2007) wedge (thin solid line) superimposed to the locus identified by Donley et al. (2012) in the IRAC colour space (thick solid line). Among the AGN so identified by Lacy et al. (2007) criterion, the fraction of them entering the Donley wedge is only  $\approx 17\%$ , but the contamination from non-AGN sources is strongly reduced. This conservative diagnostic identifies about 75% of AGN with QSO-like luminosities ( $L_X > 10^{44} \text{ erg s}^{-1}$ ) and also unveils highly obscured sources up to  $z \sim 3$ , but is less complete at lower X-ray luminosities compared to Lacy et al. (2007).



**Figure 1.8:** Revised AGN selection criteria by Donley et al. (2012) (thick solid lines) and Lacy et al. (2007) (thin solid lines). Small points represent IRAC-selected galaxies, both X-ray detected and undetected. Different colours mark IRAC subsamples that eventually meet either Stern et al. (2005) or Donley et al. (2012) criteria.

## 1.2 Evidences for AGN/Galaxy co-evolution

In the last 20 years, extensive broad-band observations of AGN host galaxies have greatly improved our understanding of AGN and galaxy evolution. It is now well-established that AGN activity and star-formation in their hosts are related processes. This interplay has been proved by a number of evidences.

1. Several studies highlighted empirical correlations between black hole mass and galaxy properties: galaxy stellar mass bulge  $M_{\text{bulge}}$ , stellar velocity dispersion  $\sigma_*$ , dark matter halos (e.g. Magorrian et al. 1998; Gebhardt et al. 2000; Ferrarese 2002; Gültekin et al. 2009). According to Fabian & Iwasawa (1999), this relationships could be reasonably expected, in case an active SMBH is growing at its Eddington limit, balancing the outward radiative force with the inward gravitational force. The relatively small scatter ( $1\sigma \approx 0.3$  dex) of these relations suggests that a *causal* connection exists between fueling mechanisms of AGN accretion and star-formation in the galaxy bulge (but see e.g. Jahnke & Macciò 2011 for an alternative interpretation).

In addition, several studies (Decarli et al. 2010; Merloni et al. 2010) argue in favour of a positive evolution of this relationship with redshift, suggesting that  $M_{\text{BH}}/M_{\text{bulge}} \propto (1+z)^{0.3}$ . This trend would predict over-massive BHs at high redshift, compared to their hosts. However, theoretical implications are currently limited by observational biases and uncertainties on host/BH mass measurements (e.g. Shen & Kelly 2010; Volonteri & Stark 2011).

2. Observations have suggested a tight AGN/starburst connection through the “down-

sizing” scenario, first proposed to explain the evolution of the observed galaxy luminosity functions (Lilly et al. 1996; Madau et al. 1996; Cowie et al. 1996, Bell et al. 2005; Juneau et al. 2005; Bundy et al. 2006; Wall et al. 2008). The same argument has been validated for AGN: most luminous SMBHs accrete earlier, faster and with more extreme accretion rates than less luminous ones (Ueda et al. 2003; Fiore et al. 2003; Le Floch et al. 2005; Barger & Cowie 2005; Pérez-González et al. 2005; Hasinger et al. 2005; Silverman et al. 2008). The similarity between AGN and galaxies’ anti-hierarchical growth is also reflected to the parallel evolution between volume-averaged cosmic star-formation history and black hole accretion history, both peaking at  $z \sim 2$  and declining towards the local Universe (Boyle & Terlevich 1998; Merloni et al. 2004; Marconi et al. 2004; Silverman et al. 2008; Shankar et al. 2009; Madau & Dickinson 2014).

3. Unlike most semi-analytical models predicted, observations have shown that the local galaxy stellar luminosity function and the dark halo mass function have different shapes, with the former showing a break at low and high luminosities. This required some mechanisms to reduce the efficiency of star-formation at high and low masses. Stellar feedback from Supernovae can reproduce the inefficient star-formation at low-luminosity regimes, removing large amounts of gas (Bolatto et al. 2013; Cicone et al. 2014b; Sakamoto et al. 2014). Nonetheless, semi-analytical models have shown that stellar processes on their own struggle to reproduce the break at high galaxy luminosities ( $L \sim 10^{12} L_{\odot}$ ), as well as the observed distributions of galaxy properties, without adding a further source of energy (Benson et al. 2003; Bower et al. 2006; Schawinski et al. 2006; Gabor et al. 2011). Still open remains the question whether such extra injection of heat might come from the central active black hole (known as “AGN-driven feedback”) or from stellar processes (e.g. winds or Supernovae).

Despite the mutual dependence of galaxy star-formation and AGN activity on cold gas supply, the co-evolution scenario is challenging, since the vastly different spatial scales at which star-formation (many kpc) and AGN (sub-pc) operate. In Section 1.5 we mention the most recognized mechanisms that have been proposed to explain the observed AGN/host correlations, while in Section 1.4 we summarize the main AGN host galaxy properties derived from observations of statistical samples.

### 1.3 What triggers SMBH activity?

Semi-analytical models suggest that the observed correlations between AGN activity and galaxy properties might result from a long-lasting self-regulation process between the SMBH and its host. It is well established that both AGN accretion and star-formation in galaxies required supplies of cold gas, but the detailed mechanisms responsible for tracking the gas inward onto the SMBH are still poorly understood. Two main scenarios have been suggested, as discussed in the following.

- **Major mergers** are violent interactions between two or more colliding galaxies.

Stellar orbits are disrupted, cold gas is compressed and funnelled to the common center of mass. Major mergers between gas rich galaxies are thought to activate rapid episodes of SMBH accretion close to the Eddington limit. These processes are used as a means to explain the rapid growth of the most luminous Quasars at high redshift ( $z > 3$ ).

- **Secular evolution** postulates stochastic accretion episodes, due to internal and external processes that can remove significant amounts of angular momentum and track the gas inward onto the SMBH (Kormendy & Kennicutt 2004). Internal processes encompass gas instabilities, galaxy bars and large-scale gravitational torques, while external processes may include galaxy interactions (i.e. minor mergers).

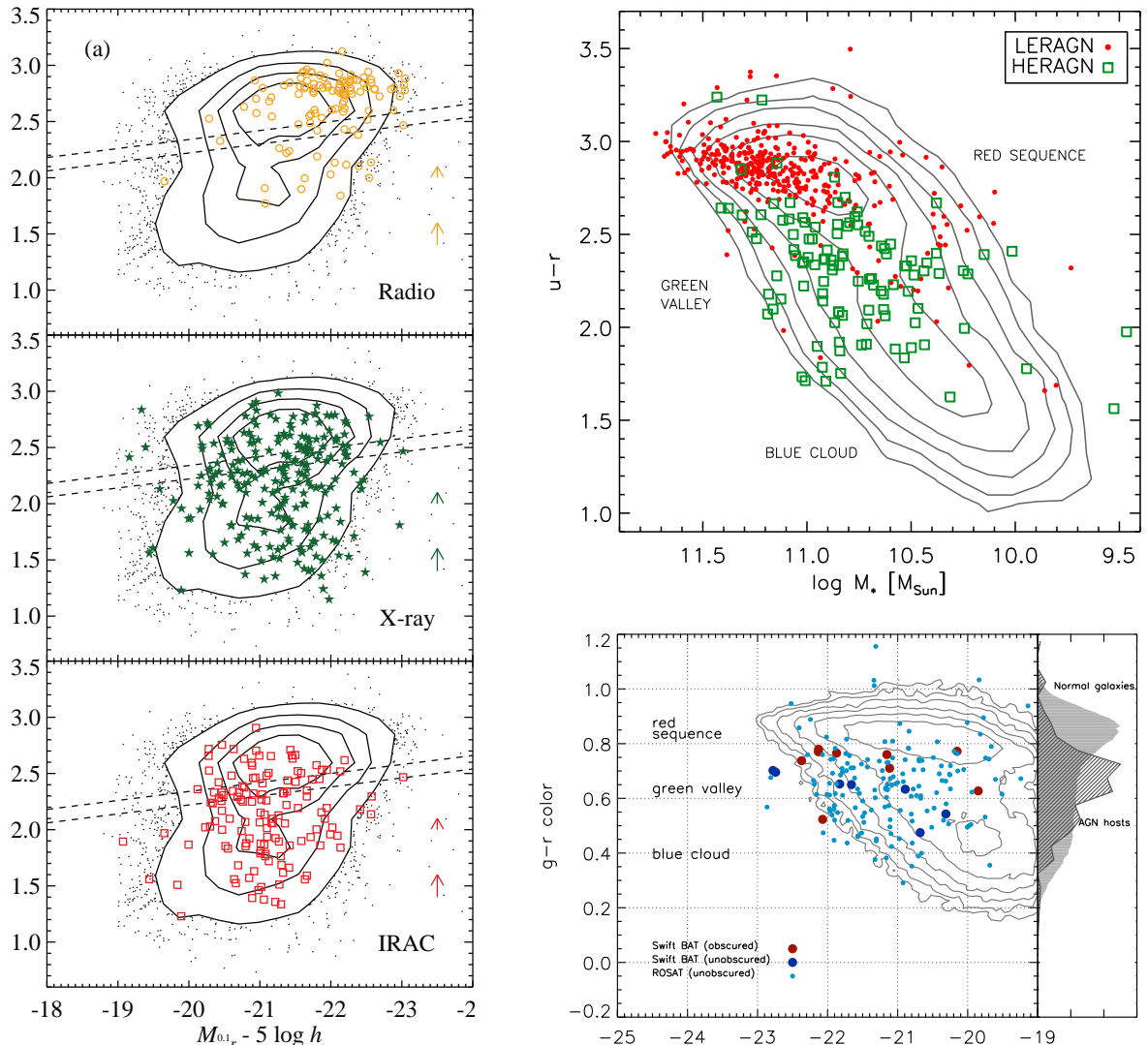
Major mergers were first postulated by Sanders et al. (1988), claiming that these violent interactions between gas rich galaxies might trigger the central SMBH close to the Eddington limit. This scenario is well supported by hydro-dynamical simulations (e.g. Di Matteo et al. 2005; Hopkins et al. 2006; Menci et al. 2008; Hopkins & Quataert 2010a) and it is widely recognized to reproduce galaxy colours, the high-luminosity break in the local galaxy luminosity function and the  $M_{\text{BH}}-M_{\text{bulge}}$  relation in the nearby Universe (see comprehensive review by Alexander & Hickox 2012).

Nonetheless, recent observations are challenging this merger-driven picture, and supporting the predominance of the stochastic accretion scenario. For instance, extensive morphological studies of X-ray selected samples of AGN have shown that AGN hosts do not appear to be merging/interacting more frequently than control samples of inactive galaxies, either at  $z \sim 1$  or  $z \sim 2$  (e.g. Grogin et al. 2005; Gabor et al. 2009; Cisternas et al. 2011; Schawinski et al. 2011, 2012; Kocevski et al. 2012; Simmons et al. 2012).

This picture suggests that, whatever the mechanisms that trigger distant SMBH growth, they also trigger distant galaxy growth (e.g. Cisternas et al. 2011; Kocevski et al. 2011; Schawinski et al. 2011). However, most of the galaxies in these studies are associated to intermediate disturbance classes, characterized by irregular (i.e. clumpy) morphologies and/or disturbed disks, which are hard to distinguish from mergers, if using broad-band morphological information only. Furthermore, Kocevski et al. (2012) claimed that a significant fraction of major mergers might be hidden by nuclear obscuration. New techniques (e.g. SINS/zC-SINF and KMOS-3D, Mancini et al. 2011; Förster Schreiber et al. 2011, 2014; Wisnioski et al. 2014) using near-IR integral field spectroscopy (IFS) to obtain rest-frame optical imaging and 3D-spectroscopy will be essential to resolve the internal kinematics of these sources, therefore providing new insights on fueling mechanisms of AGN activity.

Actually, the most accepted scenario is that merger-driven AGN accretion is the dominant mode at high AGN luminosities ( $L_{\text{bol,AGN}} > 10^{12} L_{\odot}$ , Hopkins & Hernquist 2006; Hasinger 2008; Kartaltepe et al. 2010; Hopkins et al. 2013). On the other hand, gradually varying processes, such as internal secular evolution (i.e. large-scale disc instabilities and bars), external secular evolution (galaxy interactions) and inflow of recycled cold gas from old stellar populations (Ciotti & Ostriker 2007; Ciotti et al. 2010), are expected to domi-





**Figure 1.9:** Left panel: colors VS R-band luminosity of AGN host galaxies selected with different criteria (Hickox et al. 2009). Top-right panel: same colour-mass space for radio AGN classified as either HERAGN or LERAGN (Smolčić et al. 2009, see Section 1.1.4). Bottom-right panel: Colour-mass plane for X-ray AGN (Schawinski et al. 2009). The general consensus is that AGN in radiatively efficient states are mostly found in blue cloud and green valley galaxies, while low-Eddington ( $\lambda_{\text{EDD}} < 10^{-3}$ ) AGN populate the red sequence.

nate at lower AGN luminosities (Kormendy & Kennicutt 2004; García-Burillo et al. 2005; Simões Lopes et al. 2007; Hopkins & Quataert 2010b; Storchi-Bergmann 2010).

## 1.4 The variety of AGN host galaxies

Galaxies show a well established bimodality in the optical colour-magnitude diagram (CMD, Strateva et al. 2001), that separates the “red sequence” of luminous, passively evolving galaxies, from the “blue cloud” of less luminous, star-forming systems. In this paragraph we briefly summarize the most relevant properties of AGN host galaxies in-

ferred from studies of X-ray, radio and IR-selected AGN.

Fig. 1.9 shows the location of AGN host galaxies in the CMD according to different selection criteria (taken from Hickox et al. 2009). Radio-selected AGN (top-left panel) AGN host galaxies are preferentially found in the red sequence, with spheroidal-like morphologies (Dunlop et al. 2003; McLure et al. 2004; Herbert et al. 2011) and hosted in very massive hosts ( $M_{\star} \sim 10^{11-12} M_{\odot}$ ). X-ray AGN host galaxies mostly lie in the “green valley”, with intermediate optical colours between blue and red galaxies (e.g., Nandra et al. 2007; Georgakakis et al. 2008; Silverman et al. 2008; Hickox et al. 2009). They show similar morphologies with respect to those inferred from a control sample of mass-matched inactive galaxies (Pierce et al. 2007; Gabor et al. 2009; Georgakakis et al. 2009) and their hosts have stellar mass  $M_{\star} \approx 0.3-3 \times 10^{11} M_{\odot}$  (Alonso-Herrero et al. 2008; Bundy et al. 2008; Brusa et al. 2009; Hickox et al. 2009; Xue et al. 2010). IR-AGN selected with IRAC-based colour-colour criteria show an optical colour distribution similar to that of X-ray AGN, though more pronounced towards the blue cloud and hosted in less massive galaxies with disky-like morphology (Hickox et al. 2009). This suggests that, on average, IR-selected AGN are preferentially hosted in dusty galaxies with on-going star-formation.

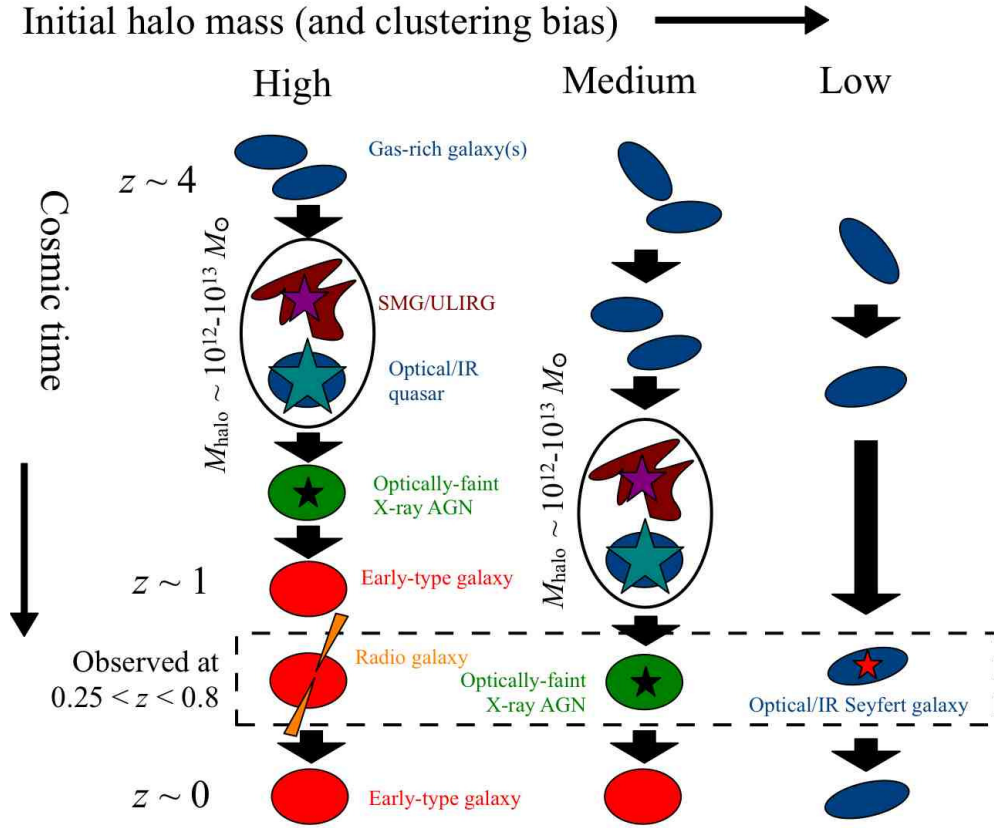
Studies focusing on clustering have highlighted systematically different AGN accretion rate distributions as a function of the environment (see Fig. 1.9, left). Radio AGN have relatively low Eddington ratios<sup>7</sup> ( $\lambda_{\text{EDD}} < 10^{-3}$ ) and typically live in strongly clustered, luminous red galaxies. In addition, Smolčić et al. (2009) found radio AGN with high-excitation optical spectra preferentially live in the green valley, while those with low-excitation spectra commonly lie on the red sequence (Fig. 1.9, top right). Conversely, IR-selected AGN have higher Eddington ratios ( $\lambda_{\text{EDD}} > 10^{-2}$ ) and live in more weakly clustered environments (Hickox et al. 2009). X-ray AGN have intermediate to high  $\lambda_{\text{EDD}}$  and reside preferentially found in overdense regions similarly to the most massive elliptical galaxies (e. g., Georgakakis et al. 2007; Silverman et al. 2009; Hickox et al. 2009; Coil et al. 2009).

These results indicate that samples of AGN selected with different criteria represent intrinsically separate populations, with systematic difference in their colours, environments and AGN accretion modes. In Section 1.5 we discuss some challenging models to reconcile all the pieces of the puzzle within a single, self-consistent picture of AGN/galaxy evolution.

## 1.5 Evolutionary scheme of AGN/galaxy evolution

Observations of high- $z$  galaxies (Elmegreen et al. 2007), complemented with numerical models (Hopkins et al. 2009) suggest that the first galaxies in the early Universe were, on average, more gas-rich than today’s systems and mainly supported by rotation. Following a hierarchical evolution, galaxies (as well as their seed BHs) have become progressively more massive with cosmic time (likely through minor mergers), and a similar growth occurs for their dark matter halos. Since most QSOs are generally found in halos with mass  $M_{\text{h}} \sim 10^{12} M_{\odot}$ , regardless of redshift, this value has been interpreted as a sort of critical

<sup>7</sup>The Eddington ratio  $\lambda_{\text{EDD}}$  is simply defined as the fraction of the emitted bolometric luminosity compared to its Eddington limit



**Figure 1.10:** Cartoon adapted from Hickox et al. (2009) to illustrate possible evolutionary schemes of AGN and their hosts through the cosmic time, according to the initial dark matter halo mass.

halo mass. More massive halos are expected to grow faster than less massive ones, therefore reaching the critical  $\approx 10^{12} M_{\odot}$  at earlier times. On this basis, a plethora of AGN/galaxy evolutionary models have been suggested. Here below we discuss the model first proposed by Hickox et al. (2009), which is widely recognized and extensively supported by panchromatic observations of AGN host galaxies (see also Hopkins et al. 2008; Hasinger 2008). As mentioned above, the cosmic growth of the most massive halos (and galaxies) follows different stages with respect to less massive ones. Fig. 1.10 provides a schematic view of the various evolutionary stages as a function of the initial halo mass  $M_{\text{h}}$ .

The cosmic evolution of the most massive halos ( $M_{\text{h}} > 10^{12} M_{\odot}$ , see left/middle schemes) encompasses four different stages.

1. rapid and efficient ( $\lambda_{\text{EDD}} \gtrsim 10^{-1}$ ) SMBH accretion is triggered. Possible mechanisms include major merger of gas rich galaxies, while also disk instabilities or recycled gas from evolved stars might play a role. Whatever the mechanisms triggering QSO accretion, they also switch on rapid star-formation in the host galaxy. This short-lasting ( $\sim 10^8$  yr, e.g. Hopkins et al. 2008) phase is characterized by the presence of powerful star-formation enshrouded in dense dusty environments and fueled by huge

amounts of cold gas/dust. These systems are commonly identified as sub-millimeter galaxies (SMGs, Sanders & Mirabel 1996). Observations of high-redshift SMGs or Ultra Luminous InfraRed Galaxies (ULIRGs) often show disturbed kinematics consistent with galaxy mergers (Alexander et al. 2003, 2005; Tacconi et al. 2008; Engel et al. 2010; Laird et al. 2010). The majority of them have X-ray properties consistent with luminous X-ray AGN, most of which (*approx*80%) are heavily obscured ( $N_H > 10^{23} \text{ cm}^{-2}$ , Alexander et al. 2005).

2. subsequently, this rapid AGN accretion reaches its Eddington limit and crosses a very short-lasting ( $\sim 10^7$  yr) “blow-out phase”, where large amounts of gas flowing onto the central BH get efficiently converted to ionizing photons, therefore swept away up to kpc-scales. Possible scenarios invoke powerful winds driven by the AGN, that should appear as unobscured, as a means for removing large amounts of cold gas or to heat it up. This so-called “Quasar mode” is required to regulate the SMBH mass in order to explain the not runaway growth of today’s SMBHs. This scenario is supported by hydro-dynamical simulations, since the radiation pressure of the central SMBH might prevent further episodes of galaxy star-formation (Croton et al. 2006; Hopkins & Quataert 2010a), tracking the host galaxy from the blue cloud to the red sequence of the CMD, therefore naturally reproducing the observed bimodality in galaxy colours. Testing this paradigm is challenging, as this is supposed to be a quick transition from highly accreting to fading AGN. Some studies finding “smoking guns” of AGN-driven feedback, such as massive, large scale ( $>10$  kpc) molecular outflows (Feruglio et al. 2010, Cicone et al. 2012, 2014a), might also explain how the star-formation of the galaxy has been “quenched”.
3. Soon after the blow-out phase, the lack of cold gas supply makes the accretion onto the central SMBH less radiatively efficient ( $\lambda_{\text{EDD}} \lesssim 10^{-2}$ ) and progressively dominated by X-ray emission. For the same reason, AGN host galaxies at this stage generally show co-existing young and old stellar populations, producing “green-like” integrated galaxy colours. This phase is supposed to last  $\approx 10^9$  yr to explain the fact that most X-ray selected AGN hosts are found to have intermediate optical colours.
4. Accretion onto the central SMBH becomes radiatively inefficient ( $\lambda_{\text{EDD}} \lesssim 10^{-3}$ ). To justify the lack of massive ( $M_\star > 3 \times 10^{10} M_\odot$ ) galaxies in the local Universe, some mechanism preventing the gas in the halo to cool down and collapse must be in place. This so-called “radio-mode” is required to last much more time than the previous phases, in order to maintain the galaxy “clean” and explain the formation of red and dead spheroidal galaxies in the local Universe. Models of mechanical feedback invoke either powerful radio jets driven by the central SMBH, or stellar-driven winds, or both phenomena (e.g. Hopkins et al. 2008, Hopkins & Quataert 2011).

As mentioned above, dark matter halos with lower initial mass reach the critical halo mass at later times. Therefore, galaxies residing in the least massive halos should have never hosted luminous Quasars, as illustrated in Fig. 1.10 (right scheme). These galaxies are

believed to smoothly increase their stellar mass and SMBH mass through stochastic fueling of cold gas (e.g. Hopkins & Hernquist 2006) driven by minor-mergers and/or funnelled into the central SMBH via galaxy-scale processes (e.g. galaxy bars). According to the “downsizing” scenario, the number density of moderately luminous AGN ( $L_X < 10^{44}$  erg s $^{-1}$ ) is expected to increase at low redshift, therefore supporting this scenario. Nevertheless, hunting for LLAGN is challenging, since their emission is generally outshined from the host galaxy, unlike for the upper tail of their accretion rate distribution, which make them detectable in moderately deep X-ray surveys.

Evolutionary models based on these prescriptions have successfully reproduced, in broad terms, luminosities, colors, and clustering of galaxies and AGN (e.g., Bower et al. 2006; Croton et al. 2006; Hopkins et al. 2008; Merloni & Heinz 2008a).

On the other hand, AGN activity could result also in positive feedback (Silk & Norman 2009, Silk et al. 2013), therefore triggering new star formation episodes. This alternative picture is motivated by the evidence for AGN triggering on star-formation by compressing dense gas clouds. This over-pressure caused by the AGN radiation field may exceed the dynamical pressure that keeps gas clouds bound (e.g. through shocks). As predicted by the Kennicutt-Schmidt law, larger cold gas surface density implies larger SFR density, which is linked to on-going star-formation. This positive feedback hypothesis is motivated by recent simulations (e.g. Silk 2005; Zubovas et al. 2013) and also supported by observations of jet-induced outflow in radio galaxies showing enhancement of star-formation (Morganti et al. 2007; Gaibler et al. 2011; Tadhunter et al. 2014). In line with these arguments, Hasinger (2008) discussed a more complex scenario where the “quenching phase” likely operated by AGN-driven feedback might be followed by sporadic accretion episodes. Possible inflows of cold gas might fuel and rejuvenate the central SMBH, as well as new star-formation in the galaxy. According to this picture, the “green valley” identifies a subsequent stage with respect to the “red sequence”, where stochastic, low-Eddington ( $\lambda_{\text{EDD}} < 10^{-2}$ ) accretion triggers SMBHs as LLAGN. This picture provides an alternative scenario, predicting the intermediate optical colours effectively found in most X-ray AGN host galaxies.

All these scenarios reproduce equivalently well several observables (clustering, optical colours, Eddington ratios) of AGN hosts, whereas discriminating among them is tricky, as well as to determine whether AGN are aftermaths or precursors to star-formation, or both.

Though it is today well established that the evolution of active SMBHs is connected to the evolution of their hosts, the underlying mechanisms explaining their mutual connection are still poorly constrained. In this Chapter we have discussed the physical and observational properties of AGN, including different diagnostics to identify AGN from “inactive” galaxies. We have also argued that a panchromatic study of this class of objects is mandatory to constrain their properties and their cosmic evolution. As mentioned in Section 1.1, in the present Chapter we have intentionally overlooked the recent results achieved from *Herschel*-based studies in the framework of AGN/galaxy evolution, since their contribution will be mostly emphasized in Chapter 2 and in the rest of this PhD Thesis.



# The *Herschel* view of AGN/galaxy evolution

In this Chapter we discuss the relevance of IR observations in the framework of AGN/galaxy evolution, primarily focusing on the profound advances made possible by exploiting data taken from the *Herschel* satellite. In the first part we introduce *Herschel* and the most relevant surveys in the context of this Thesis. Next we provide a selection of recent *Herschel*-based results that have improved our knowledge of AGN/Galaxy evolution and also motivate the subsequent work accomplished in this Thesis.

## 2.1 The road to *Herschel*

It has now become clear that extragalactic surveys in the FIR represent a key tool for understanding galaxy formation and evolution. Since star formation typically occurs in dusty molecular clouds, dust is responsible for attenuating most of the optical/UV starlight generated in the Universe and subsequently re-emitted in the IR domain (i.e. Genzel & Cesarsky 2000). For these reasons, FIR data are essential to reconstruct how galaxies have evolved since their formation epoch.

About 30 years ago, the InfraRed Astronomical Satellite (IRAS, Neugebauer et al. 1984) provided the first MIR and FIR observations (at 12, 24, 60 and 100  $\mu\text{m}$ ) of local star-forming galaxies and AGN (Sanders et al. 1986; Soifer et al. 1987), essential to construct the first templates to model the galaxy/AGN IR emission (e.g. Rowan-Robinson & Crawford 1989). The Infrared Space Observatory (ISO, Kessler et al. 1996) provided the first IR observations of non-local galaxies, reaching up to  $z \sim 1-1.5$  and obtained MIR spectra for modelling different dust emission components in the range between 2.5 and 200  $\mu\text{m}$  (Efstathiou & Rowan-Robinson 1995; Moorwood 1999). ISOCAM observations at 15  $\mu\text{m}$  first discovered an excess in MIR number counts (Aussel et al. 1999; Gruppioni et al. 2002) with respect to the predictions for a non evolving population. This supported an intrinsic evolution with redshift of the energy density released by star-forming galaxies (e.g. Franceschini et al. 2001). First IR-based derivations of the star formation rate density (SFRD) obtained with ISO showed a good agreement with those inferred from the optical (e.g. Madau et al.



**Figure 2.1:** Artistic impression of the *Herschel* satellite. Credit: ESA.

1996), at least up to  $z \sim 1$ . A higher redshift, the latter showed a steep decrease, in tension with the flat trend derived from ISO data. This discrepancy suggested that average dust attenuation in galaxies was stronger at earlier times rather than in the present epoch (e.g. Flores et al. 1998). Further investigations were addressed by *Spitzer* (Werner et al. 2004) through observations at NIR and MIR wavelengths, that allowed for the first time to characterize different populations of star-forming galaxies and to trace the evolution of the SFRD up to  $z \sim 3$ . These studies corroborated previous results, finding a peak at  $z \sim 1$ , followed by a flattening up to  $z \sim 2.5-3$  (Pérez-González et al. 2005; Caputi et al. 2007; Reddy et al. 2008).

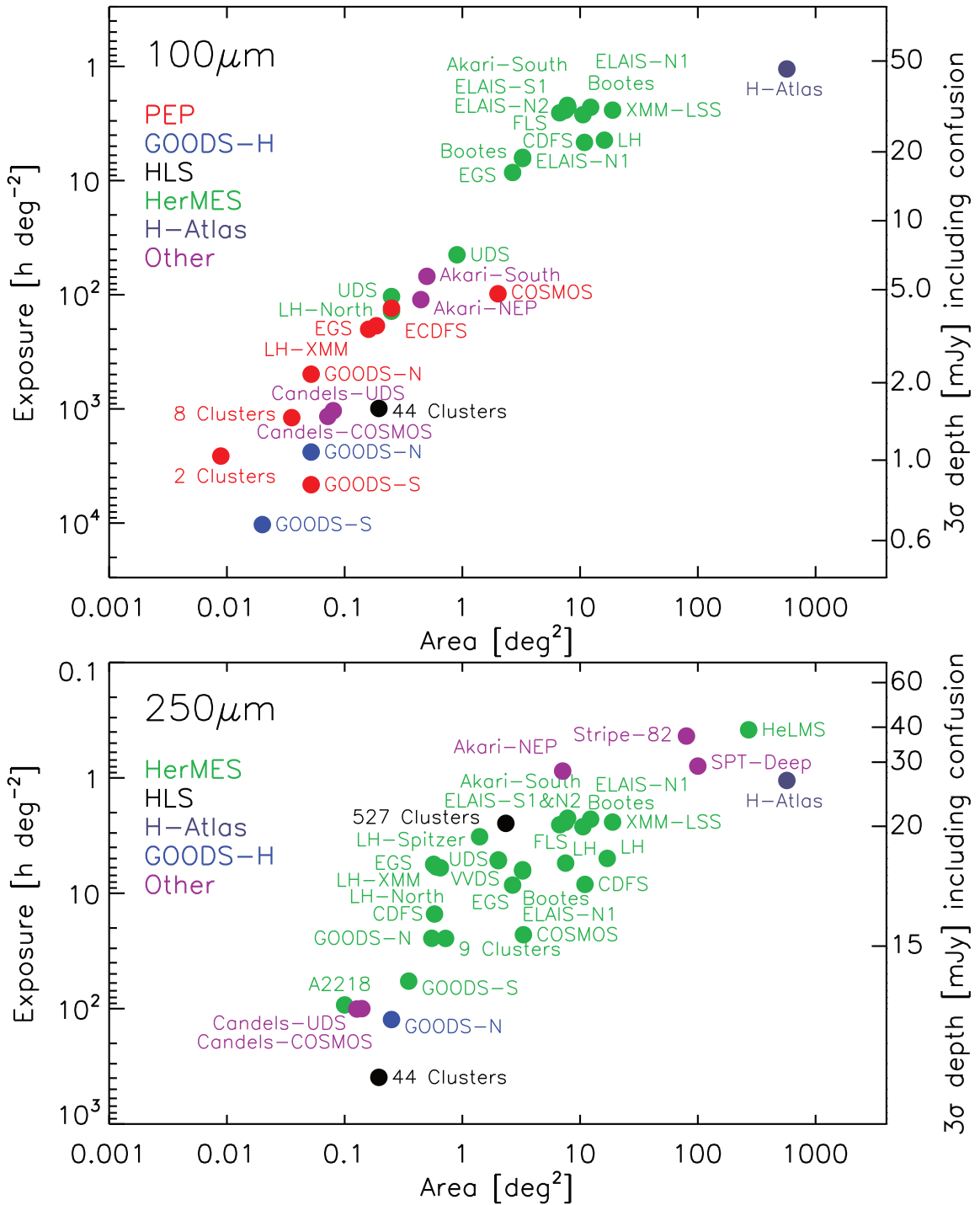
## 2.2 *Herschel* surveys and observing strategy

The *Herschel Space Observatory*<sup>1</sup> (Pilbratt et al. 2010), equipped with a 3.5 meter diameter reflecting telescope and instruments cooled to close to absolute zero, it is the only space facility ever developed to cover the FIR part of the spectrum (55–672  $\mu\text{m}$ ). Moreover, thanks to the large size of the primary mirror (3.5 rather than 0.7 meters of ISO and *Spitzer*), *Herschel* observations enabled to reach unprecedented sensitivities in the FIR. During its operation time (from May 14<sup>th</sup> 2009 to April 29<sup>th</sup> 2013), *Herschel* has been designed to provide broad band imaging in six bands centered at 70, 100, 160, 250, 350 and 500  $\mu\text{m}$  and spectroscopy over the entire *Herschel* coverage, using the following instruments.

- The *Photodetector Array Camera and Spectrometer* (PACS, Poglitsch et al. 2010): a camera and a slow to medium resolution spectrometer for wavelengths in the range between 55 and 210  $\mu\text{m}$ . Totally it includes four detector arrays: two bolometer arrays dedicated to wide-band photometry and two Ge:Ga photoconductor arrays employed exclusively for spectroscopy.

<sup>1</sup><http://herschel.esac.esa.int/>, <http://herschel.cf.ac.uk/>





**Figure 2.2:** Survey area and point source depth (including confusion noise) reached by most extragalactic surveys with PACS (top panel) at 100  $\mu\text{m}$  and SPIRE (bottom panel) at 250  $\mu\text{m}$ . Adapted from Lutz (2014).

- The *Spectral and Photometric Imaging REceiver* (SPIRE, Griffin et al. 2010), a camera and a low to medium resolution spectrometer complementing PACS for wavelengths in the range between 194 and 672  $\mu\text{m}$ . It employs an imaging photometer

and a Fourier Transform Spectrometer (FTS), both using bolometer detector arrays, which include three arrays dedicated to photometry and two for spectroscopy.

- The *Heterodyne Instrument for the Far Infrared* (HIFI, de Graauw et al. 2010), is a very high resolution heterodyne spectrometer covering the pass-bands at 157–212  $\mu\text{m}$  and 240–612  $\mu\text{m}$ .

These instruments are devoted to address a number of scientific objectives: studying the molecular composition of planetary atmospheres, comets and the interstellar medium, as well as to investigate the formation of galaxies in the early Universe and their subsequent evolution. This Thesis focuses on this latter point by exploiting photometric surveys covered by the PACS (70, 100 and 160  $\mu\text{m}$ ) and SPIRE (250, 350 and 500  $\mu\text{m}$ ) instruments. Fig. 2.2 gives an overview of area and depth of guaranteed time and open time *Herschel* key programmes. Here below we list some of the major projects providing data used in this Thesis.

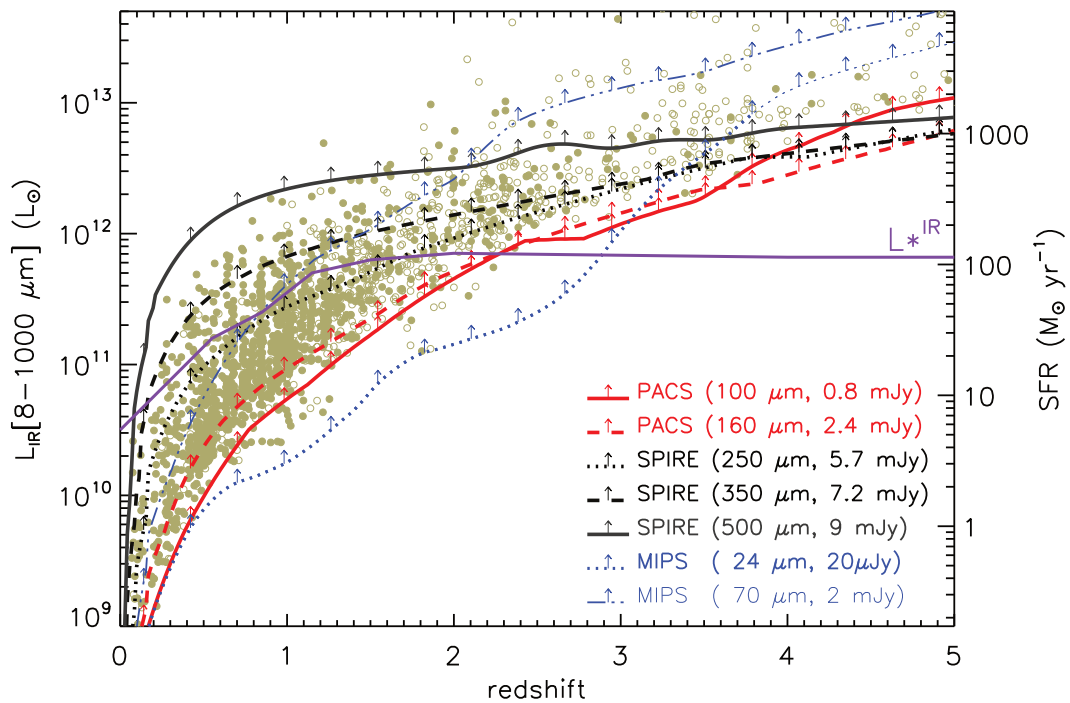
- the PACS Evolutionary Probe<sup>2</sup> (PEP, Lutz et al. 2011) is a guaranteed PACS survey covering popular multi-wavelength fields, such as GOODS-South, GOODS-North, COSMOS, Extended Groth Strip (EGS), Extended Chandra Deep Field South (ECDFS), Lockman Hole and two galaxy clusters. Sky areas covered by PACS are also complemented with SPIRE observations taken from the HerMES survey (see below). This project totally covers 2.7 deg<sup>2</sup>, equivalent to 550 hours of observing time.
- the *Herschel* Multi-Tiered Extragalactic Survey<sup>3</sup> (HerMES, Oliver et al. 2012) is guaranteed survey structured as a wedding cake. Observations are done with both SPIRE and PACS instruments, surveying  $\approx 70$  deg<sup>2</sup> for a total of 900 hours observing time.
- the Great Observatories Origins Deep Survey-*Herschel*<sup>4</sup> (GOODS-H, Elbaz et al. 2011) is an open-time survey, providing with deep PACS and SPIRE observations of the GOODS-North and ultra-deep PACS coverage of part of the GOODS-South.

PEP and HerMES large key programmes are structured as “wedding cakes” (i.e., with large area shallow surveys and smaller pencil beam deep surveys) and include many widely studied blank and lensed extragalactic fields. In the following we will focus on the GOODS-South (GOODS-S), the GOODS-North (GOODS-N), which are also coordinated with more recent GOODS-H observations, and the COSMOS field. The sky area covered by PEP/HerMES observations is about 2 deg<sup>2</sup> in COSMOS, while in each GOODS field it reaches around 300 arcmin<sup>2</sup>. As it will be discussed in Section 4.2, the sub-area covered by multi-wavelength ancillary data in each GOODS field is limited to about 190 arcmin<sup>2</sup> (17×11 arcmin<sup>2</sup>). In addition, GOODS-*Herschel* observations cover the entire GOODS-N and only partially the GOODS-S (the inner  $\approx 100$  arcmin<sup>2</sup>). Magnelli et al. (2013) have

<sup>2</sup><http://www.mpe.mpg.de/ir/Research/PEP/index.php>

<sup>3</sup><http://hermes.sussex.ac.uk/content/hermes-project>

<sup>4</sup><http://hedam.oamp.fr/GOODS-Herschel/>



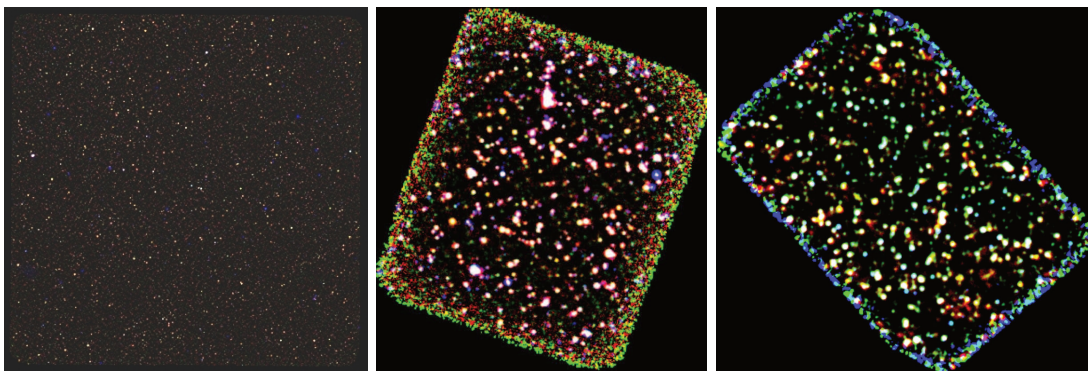
**Figure 2.3:** GOODS-H detection limits and total IR luminosities of the *Herschel* sources as a function of redshift (filled dots: spectroscopic, open dots: photometric) taken from Elbaz et al. (2011). The right axis is the SFR derived from the Kennicutt (1998) scaling factor using a Salpeter (1955) IMF. These limits were computed assuming the local library of template SEDs of Chary & Elbaz (2001). For comparison, the Spitzer MIPS-24  $\mu\text{m}$  detection limit is represented as well as the knee of the total IR luminosity function, as derived by Magnelli et al. (2009, 2010).

combined PEP observations at 100 and 160  $\mu\text{m}$  obtained in the GOODS-S and GOODS-N fields with deeper GOODS-H observations. The final PEP/GOODS-H catalogues cover the entire  $11 \times 17 \text{ arcmin}^2$  in both fields. In particular, PEP observations in the GOODS-S cover  $11 \times 17 \text{ arcmin}^2$ , and are replaced by the deeper GOODS-H data only in the inner  $10 \times 10 \text{ arcmin}^2$ . In the GOODS-N, new GOODS-H data are available over the whole  $11 \times 17 \text{ arcmin}^2$  area.

The different sensitivities reached in these fields (see Table 2.1) allow us to probe wide and complementary ranges of infrared luminosities up to  $z \sim 3$ . The combination of PACS and SPIRE observations is paramount to probe directly the rest-frame peak of interstellar dust emission at  $\approx 100 \mu\text{m}$  over a wide range of redshifts. Due to the fact that the peak wavelength is not dramatically dependent on the adopted SED template, intrinsic IR luminosity and redshift (Symeonidis et al. 2013; Lee et al. 2013), this causes a positive  $K$ -correction to enter filters at longer wavelengths, thus gaining some extra-IR flux with increasing redshift. With the sensitivity limits of GOODS-H, Fig.2.3 shows that the minimum  $L_{\text{IR}}$  detectable by *Herschel* in the local Universe is much lower at shorter wavelengths. Moving at higher redshifts, longer wavelengths (e.g. SPIRE) become comparably sensitive to shorter ones (e.g. PACS) with increasing redshift. In the following we provide a short description of the *Herschel*-selected catalogues in the above-mentioned

**Table 2.1:** Main properties of the *Herschel* surveys used in this study. We list PSF beam, central point and  $3\sigma$  flux density limits for each field used in this Thesis, spanning from 24 to 500  $\mu\text{m}$ . “...” indicate that observations in a given band are not available.

		24 $\mu\text{m}$	70 $\mu\text{m}$	100 $\mu\text{m}$	160 $\mu\text{m}$	250 $\mu\text{m}$	350 $\mu\text{m}$	500 $\mu\text{m}$
FWHM (arcsec)		5.7	5.8	6.7	11.0	18.1	24.9	36.6
Field	Central point	Depth ( $3\sigma$ )						
		$\mu\text{Jy}$	mJy	mJy	mJy	mJy	mJy	mJy
GOODS-S	03 <sup>h</sup> 32 <sup>m</sup> , -27° 48'	20	1.1	1.2	2.4	7.8	9.5	12.1
GOODS-H (South)	03 <sup>h</sup> 32 <sup>m</sup> , -27° 48'	20	...	0.8	2.4	...	...	...
GOODS-N	12 <sup>h</sup> 36 <sup>m</sup> , +62° 14'	21	...	3.0	5.7	9.2	12	12.1
GOODS-H (North)	12 <sup>h</sup> 36 <sup>m</sup> , +62° 14'	21	...	1.1	2.7	5.7	7.2	9.0
COSMOS	00 <sup>h</sup> 00 <sup>m</sup> , +02° 12'	45	...	5.0	10.2	8.1	10.7	15.4



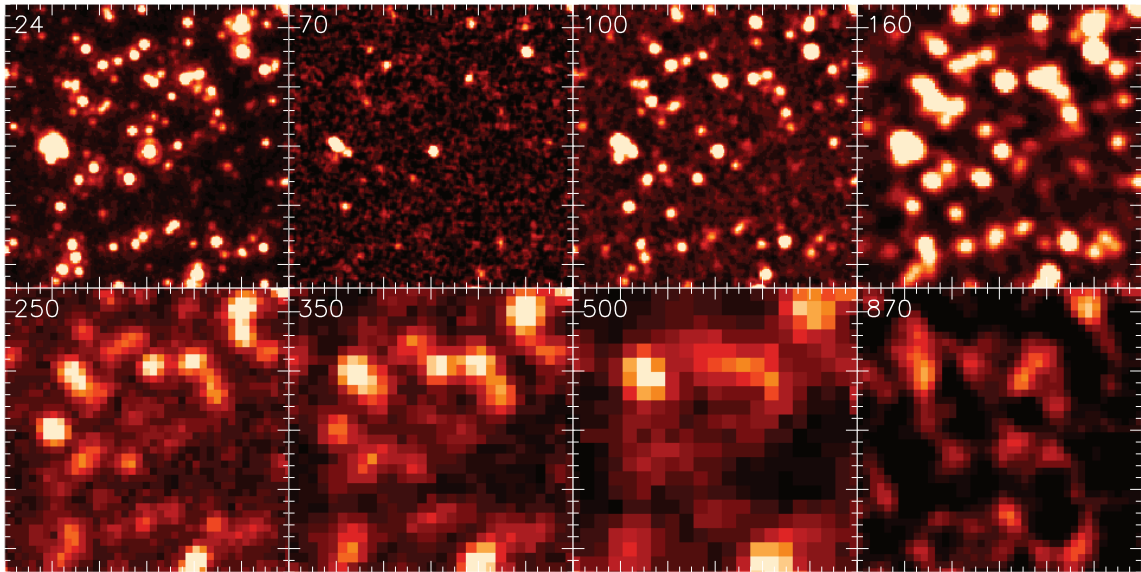
**Figure 2.4:** RGB composite images of COSMOS (left) and GOODS-H South (middle) observed with *Spitzer*/MIPS at 24  $\mu\text{m}$  (blue) and with *Herschel*/PACS at 100  $\mu\text{m}$  (green) and 160  $\mu\text{m}$  (red). The composite image of the GOODS-H North (right) includes observations with PACS at 100  $\mu\text{m}$  (blue) and 160  $\mu\text{m}$  (green) and with SPIRE at 250  $\mu\text{m}$  (red).

fields. We refer the interested reader to the official PEP and HerMES releases, publicly<sup>5</sup> available at [http://www.mpe.mpg.de/ir/Research/PEP/public\\_data\\_releases.php](http://www.mpe.mpg.de/ir/Research/PEP/public_data_releases.php) and <http://hedam.lam.fr/HerMES/index/download>, respectively.

## 2.3 Data handling

The best suited observing mode for covering *Herschel* surveys is the scan map, where the telescope moves along parallel paths, scanning also along orthogonal lines to ensure a uniform coverage of the entire area. This strategy ensures relatively flat background maps. Before the extraction of source counts, an accurate determination of the total noise level is required, particularly of the “instrumental noise” and the “confusion noise”. We refer the reader to Lutz et al. (2011), Oliver et al. (2012) and Magnelli et al. (2013) for a detailed description of the observing strategy adopted in the PEP, HerMES and PEP/GOODS-H

<sup>5</sup>The HerMES Data Release 2 for the GOODS and COSMOS fields includes the bright ( $>5\sigma$ ) HerMES catalogues only, while the full version ( $>3\sigma$ ) is not currently accessible to the public.



**Figure 2.5:** Postage stamp images of the same  $4 \times 4$  arcmin<sup>2</sup> region in the Hubble Ultra-Deep Field North (HUDF-N). Data are from the GOODS project ( $24 \mu\text{m}$ ), the combined PEP and GOODS-*Herschel* data ( $70$ - $160 \mu\text{m}$ ), HerMES ( $250$ - $500 \mu\text{m}$ ), and the groundbased LESS survey ( $870 \mu\text{m}$ , Weiß et al. (2009)). Adapted from Lutz (2014).

surveys, respectively.

### Instrumental and confusion noise

The instrumental noise is linked to the imperfect detector capabilities of identifying a “real source” from a “spurious signal”. This source of uncertainty can be accounted for through Monte Carlo simulations, but is generally not as much relevant as the confusion noise. This latter relates to the (stochastic) fluctuations of the background sky brightness below which sources cannot be detected individually. The confusion limit is dependent on the telescope aperture, observation wavelength and the position on the sky. Due to the limited telescope diameter compared to the observed wavelength, these fluctuations may particularly affect the total noise budget in FIR and sub-mm observations. As a consequence, SPIRE observations get more rapidly limited by confusion than those obtained by PACS. In Fig. 2.5 we show some postage stamp images of the same sky area taken at different wavelengths, from  $24$  to  $870 \mu\text{m}$ . Looking at those related to *Herschel* filters (all but the upper-left and the bottom-right panels), these images show an increasing degradation with observed wavelength. Nguyen et al. (2010) measured the confusion noise from SPIRE observations taken in the GOODS-N. They estimated a  $3\sigma$  confusion noise of  $14.4$ ,  $16.5$  and  $18.3$  mJy at  $250$ ,  $350$  and  $500 \mu\text{m}$ , respectively. Berta et al. (2011) and Magnelli et al. (2013) measured a confusion noise of  $0.6$  and  $2.0$  mJy ( $3\sigma$ ) at  $100$  and  $160 \mu\text{m}$ , respectively. GOODS-H observations in both South and North deep fields are always limited by confusion, except for PACS- $100 \mu\text{m}$  in the GOODS-N (Elbaz et al. 2011). SPIRE sensitivities reported in Table 2.1 in the GOODS-S and the GOODS-N include also confusion noise<sup>6</sup>.

<sup>6</sup>We caution that SPIRE sensitivities listed in Table 2.1 for GOODS-H (North) observations apply only to the sub-sample of “clean” galaxies, located in isolated regions of the  $24 \mu\text{m}$  maps, so they should be

### Flux extraction

Due to confusion issues, SPIRE flux densities at 250, 350 and 500  $\mu\text{m}$  were derived with a point spread function (PSF) fitting analysis guided using the position of sources detected in deep 24  $\mu\text{m}$  observations from the MIPS onboard *Spitzer* (Magnelli et al. 2009). The underlying assumption is that all the sources detected at the reference (longer) wavelength have been observed at the prior (shorter) wavelength. This is true for almost all (>95%) MIPS-24  $\mu\text{m}$  sources, since *Spitzer* observations are much deeper than *Herschel* ones with respect to an average MIR-to-FIR SED (Magdis et al. 2011). The “prior” extraction significantly reduced blending issues and provides reliable association between MIPS and SPIRE sources (Roseboom et al. 2010). An alternative approach is called “blind” extraction, which allows to extract flux densities using a blind PSF-fitting<sup>7</sup>, i.e., without taking as prior information the expected positions of sources. The advantage is that this method can be used regardless of whether a source is detected or not at shorter wavelength, but it is more affected by blending issues. However, checks have been made by Lutz et al. (2011) and Magnelli et al. (2013) in the PEP survey, reporting a satisfactory agreement between these two techniques.

In all the fields studied in this Thesis (i.e. the GOODS-S, GOODS-N and COSMOS), we make use of the “blind” versions of the PACS catalogues, while SPIRE fluxes are taken from the “prior” catalogues centered on MIPS-24  $\mu\text{m}$  positions.

### Multi-wavelength cross-match

The cross-match with photometry at different wavelengths can be accomplished through various approaches. The neighbour method depends on the angular distance between the reference source and its presumable counterpart observed at another wavelength. The counterpart is selected to have the minimum projected distance from the source of interest. A more suitable method is the likelihood ratio algorithm (Sutherland & Saunders 1992; Ciliegi et al. 2001), which selects the most probable counterpart for a given source, according to a statistical analysis that accounts for the magnitude distribution of the background sources close to the presumable counterpart, their angular distances, and the local background levels. A more detailed description of the multi-wavelength identification with ancillary photometry will be presented in Section 3.1.1.

## 2.4 An unbiased tracer of galaxy star formation

In this Section we focus on the *Herschel* capabilities of measuring the total IR budget related to galaxy star formation, which is essential to calibrate dust-insensitive tracers of SFR.

---

considered as underestimates. See Elbaz et al. (2011) for details.

<sup>7</sup>The PSF-fitting is performed by using the Starfinder tool (Diolaiti et al. 2000).

### 2.4.1 Total (8–1000 $\mu\text{m}$ ) IR emission from dust

Probing the total IR luminosity ( $L_{\text{IR}}$  computed in the rest-frame 8–1000  $\mu\text{m}$ ) of galaxies before the advent of the *Herschel* satellite was not trivial. Previous estimates of  $L_{\text{IR}}$  had to rely on large corrections for dust attenuation (Meurer et al. 1999), or near-IR SED fitting (Polletta et al. 2007), or MIPS-24  $\mu\text{m}$  photometry (Elbaz et al. 2007), or a combination of these methods. Subsequently, Elbaz et al. (2010) demonstrated that the required extrapolations of  $L_{\text{IR}}$  from the MIPS-24  $\mu\text{m}$  passband, based on the assumption that the IR SEDs of star forming galaxies remained the same at all epochs, provided good  $L_{\text{IR}}$  determinations ( $1\sigma$  dispersion around 0.15 dex) up to  $z\sim 1.3$ . At higher redshift, these methods required strong and uncertain extrapolations that were calibrated locally, but insufficiently tested at  $z>1.3$ .

Studies focused on *Spitzer*  $z\sim 2$  ULIRGs (Papovich et al. 2007; Daddi et al. 2007), subsequently confirmed by *Herschel* (Elbaz et al. 2010, 2011; Nordon et al. 2012), pointed out that estimates of SFR based on extrapolations of 24  $\mu\text{m}$  photometry from local templates were overestimated by a factor of  $\approx 5$ . This discrepancy has been observed in galaxy samples at  $z\sim 2$ , where the potential rest-frame emission from PAHs enters the MIPS-24  $\mu\text{m}$  filter. Possible explanations of this “mid-IR excess” invoked an intrinsic evolution of the MIR-to-FIR SEDs of galaxies, or the presence of extra-dust emission due to warmer (few 100 K) circum-nuclear dust heated by an AGN. However, thorough analyses of InfraRed Spectrograph (IRS) spectra of  $z\sim 2$  ULIRGs (Nordon et al. 2012) have disclaimed AGN activity from explaining this discrepancy.

More recently, using deep GOODS-H observations in the FIR, Elbaz et al. (2011) argued that no intrinsic SED evolution was responsible for explaining the observed discrepancy. They found that  $z\sim 2$  ULIRGs are simply analogs of local “normal” galaxies, scaled up in IR luminosity. On the other hand, these studies have shown that galaxies on the star forming “main-sequence” have more extended star formation, and exhibit intrinsically different SEDs compared to compact “starbursts” that lie above the main-sequence. This result suggests that the most appropriate parameter to distinguish among galaxy SEDs is the offset from the star forming main-sequence. We refer the reader to Section 2.5.1 and 5.1 for more details about this distinction.

### 2.4.2 Measuring dust-insensitive SFRs

In the context of galaxy formation and evolution, “calorimetric” rest-frame FIR measurements obtained with *Herschel* allow to get direct and reliable measurements of the galaxy SFR. At FIR wavelengths, the emission from galaxies is generally dominated by dust grains in thermal equilibrium at low temperatures ( $T\sim 20$  K). The grain temperatures depend sensitively on the intensity of the interstellar radiation field. This is the reason why the peak of the FIR SED of a galaxy is a good indicator of its current SFR.

As a consequence, the measurement of the SFR can be inferred by integrating  $L_{\text{IR}}$  through commonly adopted conversion factors. One of the most widely used is that introduced by Kennicutt (1998):  $\text{SFR} [M_{\odot} \text{ yr}^{-1}] = 10^{-10} (L_{\text{IR}}/L_{\odot})$ , scaled to a Chabrier (2003) initial mass function (IMF).

Whatever relation is used to translate  $L_{\text{IR}}$  to SFR, it typically relies on some basic assumptions. Below we demonstrate that most of them are justified in studies dealing with samples of *Herschel* galaxies.

1. Typical  $L_{\text{IR}}$ -to-SFR relations assume that the dust luminosity is totally attributed to young stars. However, the fraction of dust heating from young stars varies by a large factor among galaxies. While in bulges of local spirals with low sSFR, this fraction can reach  $\sim 90\%$ , in starburst galaxies nearly 100% of the dust heating arises from young ( $< 1$  Gyr old) stars (Kennicutt & Evans 2012). This implies that such an assumption is safely applicable to samples of *Herschel* galaxies.
2. Translating  $L_{\text{IR}}$  to SFR assumes that the global star formation of the galaxy is totally obscured by dust. Indeed, the estimate of SFR computed by integrating the total IR luminosity of a galaxy template is a tracer of the obscured star formation. Short wavelength tracers (e.g. UV slope,  $\text{H}\alpha$ , see Domínguez Sánchez et al. 2014) have been used to quantify the escaping fraction, in order to evaluate the fraction of obscured vs unobscured star formation. For instance, Magnelli et al. (2013) compared the evolution of the obscured vs unobscured SFRD at different cosmic epochs. They found that the unobscured SFR density accounts for about  $\sim 25\%$ ,  $\sim 12\%$  and  $\sim 17\%$  of the total SFRD of the Universe at  $z \sim 0$ ,  $z \sim 1$  and  $z \sim 2$ , respectively. The difference is even larger if considering massive ( $M_{\star} > 10^{11} M_{\odot}$ ) star forming galaxies detected by *Herschel*, where the obscured-to-unobscured SFR can reach typical values of 5–10.
3. No AGN heating of dust is usually taken into account. *Herschel* studies of X-ray AGN hosts have demonstrated that in more than 90% of PACS-selected galaxies, FIR fluxes are dominated by emission related to star formation in the host (Shao et al. 2010; Mullaney et al. 2012b; Rosario et al. 2012, 2013). As a consequence, this assumption is believed not to affect the derived SFRs. However, only decoupling the AGN emission from that related to the host in the IR domain enables to remove any potential systematics in measuring SFRs. We refer to Section 3.2 for a detailed description of this approach.

The validity of these assumptions make *Herschel* data particularly powerful to measure the global SFR in galaxies. A number of studies confirm that, as long as one of the either PACS or SPIRE bands is near the star formation FIR peak, the derived IR luminosities are solid within 0.1–0.2 dex uncertainties (Elbaz et al. 2011, Nordon et al. 2012, Berta et al. 2013a, Magnelli et al. 2013). It is now well established that *Herschel*-based SFRs are the most robust ones, followed by a hierarchy of alternative estimates, spanning (in order of decreasing SFR) from MIR, NIR SED-fitting and finally to UV-corrected measurements.

Moreover, the robustness of *Herschel* measurements has allowed also to establish important correlations with radio luminosity (Ivison et al. 2010; Magnelli et al. 2014; Magliocchetti et al. 2014) and X-ray luminosity (Symeonidis et al. 2014), that are commonly applied as alternative tracers of the galaxy SFR.



## 2.5 *Herschel* insights on the evolution of AGN and star forming galaxies

In this Section, we describe the innovative aspects and the scientific impact of the *Herschel* satellite. The combination between unprecedented sensitivity and reduced beam size at FIR wavelengths, makes *Herschel* a unique tool to measure the IR emission of star-forming galaxies and investigate the evolution of the cosmic star formation history (CSFH) up to  $z \gtrsim 3$ . In this Section we describe recent results based on *Herschel* data which have significantly improved our knowledge of galaxy properties and their evolution across the cosmic time. These pieces of work introduce several aspects that will be analysed in the context of this Thesis.

### 2.5.1 Two modes of star formation

First suggested by studies based on *Spitzer* photometry (Brinchmann et al. 2004; Noeske et al. 2007; Daddi et al. 2007; Elbaz et al. 2007), star-forming galaxies up to  $z \sim 2$  have been shown to follow a remarkably tight ( $1\sigma$  scatter  $\approx 0.3$  dex) relation between SFR and  $M_*$ , frequently called “star formation main-sequence” (MS hereafter). This relation shows a positive trend of the type  $\text{SFR} \propto M_*^\alpha$ , with  $\alpha \sim 0.6-1$ , and its normalization evolves with redshift as  $\text{SFR}/M_* \propto (1+z)^{2.5-3}$ , at least up to  $z \sim 2$ . This relation has been interpreted as the evidence that most of galaxies evolve in steady-state, spending much of their lifetime getting fueled by continuous inflows of cold gas and/or other secular processes (Dekel et al. 2009; Ciotti et al. 2010). Starburst galaxies stand out from the MS relation, while passive systems populate the region below the MS, closely resembling the distribution of galaxies in the colour- mass diagram into blue cloud and red sequence (Section 1.4).

More recent *Herschel*-based studies (Elbaz et al. 2011) have confirmed the previous trends, uncovering the evolution of the MS relation at least up to  $z \sim 3$  (Magdis et al. 2013). The combination between *Herschel*-selected samples and the existing star forming BzK<sup>8</sup> galaxies with lower SFRs has allowed to investigate for the first time the contribution of starbursts to the cosmic SFRD. Rodighiero et al. (2011) estimated that starburst galaxies, defined to have  $> 4$  times larger SFRs compared to the MS rate at a fixed  $M_*$  (see Fig. 1 of Rodighiero et al. 2011), only contribute for  $\approx 10\%$  to the global SFRD. By applying a smoother parametrization (a double gaussian rather than a single threshold) to distinguish among these two populations, Sargent et al. (2012) have validated previous findings, claiming that these outliers contribute up to 20% of the global SFRD.

In summary, studies based on *Herschel* data have shed light on the cosmic incidence of these two-fold galaxy populations in tracing the global star formation history since more than 10 Gyr ago. A relatively steady state rate in disks galaxies, which defines the star formation main-sequence, and a starburst mode in outliers of this relation. As mentioned in Section 1.5, powerful starburst galaxies are generally interpreted as regulated by major mergers, especially in the local Universe. However, only direct measurements of

<sup>8</sup>The BzK technique has been introduced by Daddi et al. (2004). Requiring  $\text{BzK} = (z - K)_{\text{AB}} - (B - z)_{\text{AB}} > -0.2$  enables to select highly complete samples of star forming galaxies in the redshift range  $1.4 \lesssim z \lesssim 2.5$ .

cold molecular gas content inside these galaxies, such as CO line observations at mm wavelengths with ALMA and/or IRAM (see comprehensive review by Carilli & Walter 2013) can help us testing the merger-driven paradigm for these outliers and finally establish the primary mechanism responsible for fueling galaxies evolving steadily on the MS or starbursting well above it.

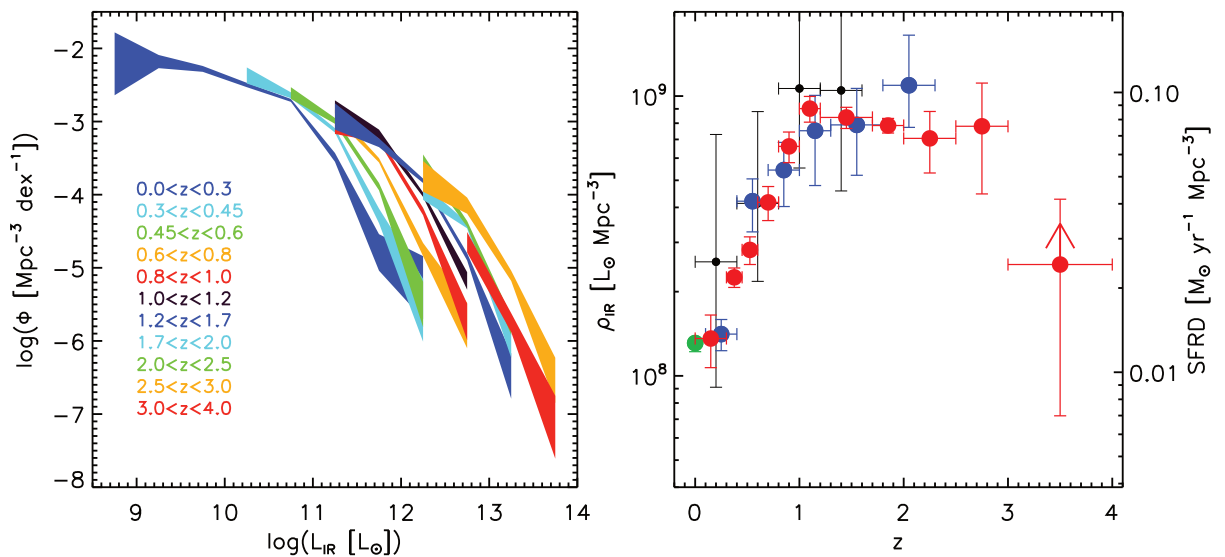
### 2.5.2 A census of the cosmic star formation history

A comprehensive understanding of the global evolution of the SFRD through cosmic time requires to apply different, but complementary methods to take a census of both obscured and unobscured star formation in galaxies.

A reliable way of measuring the obscured CSFH is by looking at the evolution of the IR luminosity function and luminosity density. Pre-*Herschel* studies required to extrapolate locally determined galaxy templates up to the total IR regime, and were also significantly limited in terms of  $L_{\text{IR}}$  range spanned by observations. A great improvement has been achieved through the combination of PEP/HerMES and GOODS-H surveys covered by FIR observations (Gruppioni et al. 2013; Magnelli et al. 2013). Employing *Herschel* photometry from both PACS and SPIRE instruments allows to sample the rest-frame peak of star formation at least up to  $z \sim 3$ . In addition, combining rest-frame FIR observations in deep GOODS fields, as well as in the shallower and wider ECDFS and COSMOS, makes possible to span an unprecedented range in  $L_{\text{IR}}$ , with highly reliable measurements of SFR. Gruppioni et al. (2013) used PACS-selected galaxies at  $160 \mu\text{m}$  supplemented with SPIRE observations up to  $500 \mu\text{m}$  to derive the total IR luminosity function up to  $z \sim 4$  (due to the selection at rest-frame  $\lambda < 200 \mu\text{m}$ , the highest redshift bin may be a mild lower limit), while Magnelli et al. (2013) exploited the deepest PACS data from the combined PEP/GOODS-H observations. Fig. 2.6 shows (left panel) the evolution of the total IR luminosity function derived from Gruppioni et al. (2013), as well as the corresponding dust-obscured SFRD (right panel), including also consistent derivations obtained by Magnelli et al. (2013).

As expected given the significant overlap in terms of sample selection and statistical approaches, both studies claim a combined luminosity and density evolution of the total IR luminosity function with increasing redshift. This finding implies a strong evolution of the obscured SFRD up to  $z \sim 1.1$ , reaching  $\approx 6$  times higher values with respect to the local determination from Sanders et al. (2003). At higher redshift, the global SFRD flattens up to  $z \sim 3$ , where incompleteness effects prevent accurate determinations. Importantly, these works have also quantified the increasing importance of ULIRGs with redshift to the total IR luminosity density, passing from 0.4% at  $z \sim 0$  to 50% at  $z \sim 2$ . On the other hand, the SFRD due to luminous infrared galaxies (LIRGs,  $L < 10^{11} L_{\odot}$ ) peaks at  $z \sim 1$ .

An estimate of the unobscured part of the global SFRD has been derived from Cucciati et al. (2012) through the rest-UV luminosity function. Burgarella et al. (2013) compared their derivation with the obscured SFRD from Gruppioni et al. (2013). The IR-to-UV luminosity density ratio provides at any epoch the average level of dust obscuration of the underlying galaxy populations. This ratio is found to increase by a factor of  $\approx 3$  from the



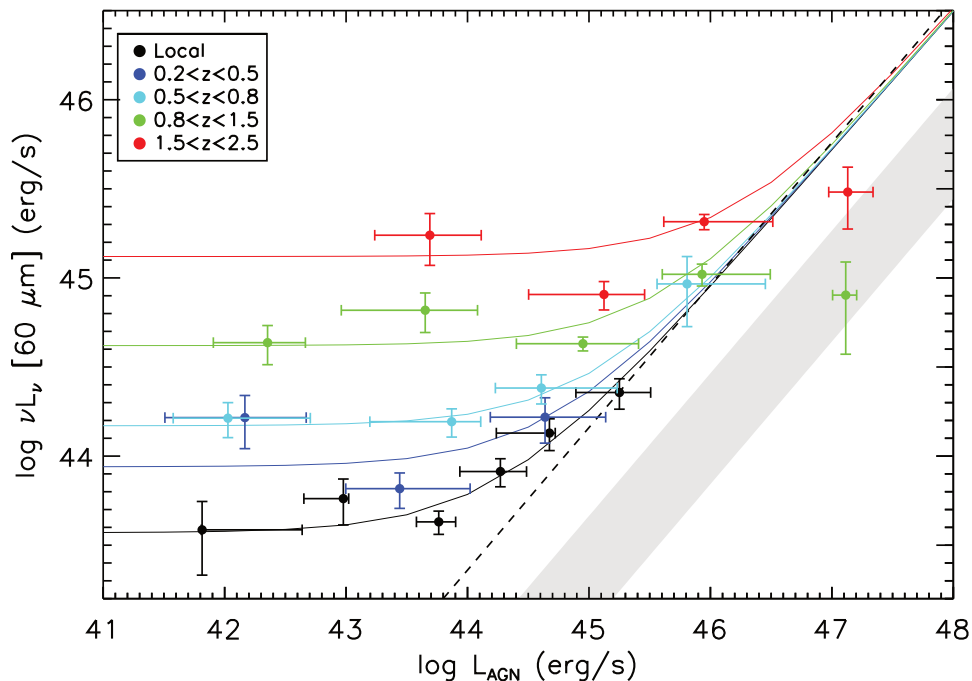
**Figure 2.6:** (*Left panel*): cosmic evolution of the total IR luminosity function, as derived from Gruppioni et al. (2013). (*Right panel*): evolution of the obscured SFRD obtained from Gruppioni et al. (2013, red), Magnelli et al. (2013, blue), Casey et al. (2012, black) and the local measurement given by Sanders et al. (2003, green). Taken from Lutz (2014).

local Universe to  $z \sim 1.2$ , decreasing at higher redshifts. This finding is interpreted as due to an intrinsic evolution of the average dust attenuation in galaxies with a peak around  $z \sim 1.2$ , suggesting that dust is dominant at  $1 < z < 2$ , but less important at high ( $z > 4$ ) redshift.

### 2.5.3 Hunting for heavily obscured AGN

The heavily obscured ( $N_{\text{H}} > 10^{23} \text{ cm}^{-2}$ ) AGN population holds a great relevance in the context of AGN/Galaxy evolution. According to several evolutionary models (see Section 1.5), heavily obscured AGN might witness a key stage of AGN/Galaxy evolution, where both the central SMBH and its host galaxy grow rapidly, enshrouded in dense environments of dust and gas, probably fueled by major mergers (e.g. Hopkins et al. 2008). In the extreme case of optically thick ( $N_{\text{H}} > 10^{24} \text{ cm}^{-2}$ ) gas clouds, these systems are not accessible even by the deepest X-ray surveys (Brandt & Hasinger 2005a) and are commonly called Compton Thick AGN. Detecting large samples of CT-AGN would allow us to constrain the AGN duty-cycle and test several models to explain the empirical evidences of AGN/Galaxy co-evolution observed in the nearby Universe (Section 1.2).

It has become clear that a panchromatic approach is invaluable to detect this elusive AGN population. Some progresses have been made by performing stacking on *Chandra* X-ray images at the positions of luminous  $z \sim 2$  BzK galaxies (Daddi et al. 2007) or MIR-excess sources (Fiore et al. 2009). An alternative approach to identify “bona-fide” CT-AGN candidates is through optical spectroscopy (e.g. [Ne V]; Gilli et al. 2010; Vignali et al. 2010; Mignoli et al. 2013) or MIR spectroscopy (Bauer et al. 2010; Alonso-Herrero et al. 2012). However, identifying heavily obscured AGN remains a considerable challenge, since significant gas and dust absorption might originate from both the host galaxy and the



**Figure 2.7:** Relation between AGN luminosity and mean FIR luminosity computed at rest-frame  $60 \mu\text{m}$ , as taken by Rosario et al. (2012). Black points are taken from Swift-BAT observations of nearby hard X-ray AGN, while coloured points identify X-ray selected AGN in the GOODS and COSMOS, binned along both axes and split in 5 different redshift bins. The grey shaded area identifies the locus in the  $L_{\text{IR}}$ -to- $L_{\text{AGN}}$  plane corresponding to AGN dominated SEDs. The black dashed line marks the Netzer (2009) relation ( $L_{\text{IR}} \propto L_{\text{AGN}}^{0.8}$ ), while the coloured lines are the functional fits to the mean measurements in various redshift bins.

AGN itself.

A significant improvement can be achieved through *Herschel* data. Despite *Herschel*-PACS/SPIRE photometry alone is not much sensitive to the presence of a possible warm dust component heated by an obscured AGN (peaking around  $10\text{--}20 \mu\text{m}$ ), a combination with *Spitzer* data would significantly improve the accuracy of AGN and galaxy SEDs, therefore allowing the AGN emission component to be isolated (Juneau et al. 2013; Berta et al. 2013b). Several works have been applying this panchromatic approach (e.g. Fritz et al. 2006; Mullaney et al. 2011; Lusso et al. 2012; Pozzi et al. 2012; Del Moro et al. 2013; Feltre et al. 2013), including that presented in this Thesis (see Chapter 3.2).

#### 2.5.4 Does AGN activity suppress or correlate with galaxy SF?

The *Herschel* capabilities of measuring the galaxy SFR are particularly useful to derive robust estimates of SFR in AGN hosts. As discussed in Section 2.4, this implicitly assumes that the rest-frame FIR spectrum is weakly affected by a possible AGN emission component. Despite the large scatter in star-forming galaxies and AGN SEDs published in the literature (e.g. Polletta et al. 2007; Mullaney et al. 2011; Berta et al. 2013b), all of them consistently support the FIR as the spectral domain where the contrast between galaxy and AGN emission stand out most clearly. To test this hypothesis, one possible approach is to

decompose the broad-band AGN+host SEDs (e.g. Netzer et al. 2007; Mullaney et al. 2011; Mor & Netzer 2012), combining multi-wavelength photometry with *Spitzer*-IRS spectra in the MIR regime, to disentangle dust emission due to heating from either star formation or AGN activity. An alternative approach that can be applied to larger statistical samples (where IRS spectra are not available) is to compare rest-frame FIR colours of ( $z < 3$ ) active galaxies to those taken from a control sample of inactive galaxies at similar redshift and  $M_*$  (e.g. Hatziminaoglou et al. 2010; Rosario et al. 2012), selected on the basis of their X-ray luminosity. All these studies point out that the average FIR content of AGN hosts arises primarily from the host galaxy emission. This is the main reason why several works exploit *Herschel* data to investigate the link between star formation and AGN accretion in active galaxies.

### Studies from optical and X-ray selected samples

Different studies of X-ray selected samples of luminous ( $L_X > 10^{44}$  erg s $^{-1}$ ) AGN indicated the presence of a correlation with the global SFR of their hosts. Netzer (2009) analysed low-redshift ( $z < 0.2$ ) optically-selected type-2 AGN, finding that  $SFR \propto L_{\text{bol,AGN}}^{0.8}$  over four orders of magnitude in AGN luminosity. They interpreted this correlation as driven by galaxy major mergers, as proposed by Lutz et al. (2008) for bright QSOs at  $z \sim 2$ . The improved sensitivity and statistics inferred by *Herschel* data allowed to uncover characteristic trends between mean SFR and AGN accretion in less luminous AGN and in less star-forming hosts (Shao et al. 2010; Rosario et al. 2012). Fig. 2.7 shows the relation between AGN luminosity and rest-frame FIR luminosity  $\nu L_\nu$  at rest 60  $\mu\text{m}$  (related to galaxy SFR) studied by Rosario et al. (2012). They considered a sample of X-ray selected AGN at  $0.2 < z < 2.5$  in the GOODS and COSMOS fields, both detected and stacked on *Herschel*-PACS maps. The observed relations are found to depend on both luminosity and redshift. At  $z < 0.8$ , they identified a positive correlation for luminous AGN ( $L_{\text{AGN}} \sim 10^{45}$  erg s $^{-1}$ ), validating previous analyses of local hard X-ray AGN from Swift-BAT (Cusumano 2009). On the other hand, less luminous AGN at the same redshift are weakly dependent on the mean galaxy SFR. The authors interpreted this observation as a sign of an increasing relevance of major-mergers in driving both the growth of SMBHs and global star formation in their hosts at high AGN luminosities. At  $0.8 < z < 2.5$  the mean SFR steeply evolves with redshift, but with no significant dependence on the AGN luminosity, suggesting that the role of mergers is less important at these epochs. This lack of trend confirms independent results found in moderately luminous ( $L_X < 10^{44}$  erg s $^{-1}$ ) X-ray selected AGN (e.g. Shao et al. 2010; Mullaney et al. 2012a; Rovilos et al. 2012). These results support a two-fold mechanism scheme to explain the observed trends of AGN accretion with galaxy star formation, where galaxy mergers dominate in the QSO regime, and secular processes regulate SMBH growth at lower AGN luminosity.

As reported in Section 2.5.1, it is generally believed that star formation in galaxies on and above the MS relation is triggered and fueled through different mechanisms. It is relevant to ask whether these two-fold schemes, for both SMBHs and their star-forming hosts, are somewhat connected one to the other. Moreover, this implicitly relates to the question

of whether AGN accretion quenches or triggers star formation in galaxies. Enlightening results have been obtained by analysing *Herschel*-based SFRs in samples of X-ray selected ( $L_X < 10^{44}$  erg s $^{-1}$ ) AGN hosts.

1. The average SFR is enhanced in X-ray AGN hosts with respect to a control sample of mass-matched inactive galaxies at the same epoch (Santini et al. 2012), including both quiescent and star-forming systems.
2. The same analysis leads to consistent mean SFRs if comparing to a reference sample of mass-matched MS galaxies only (Mullaney et al. 2012b; Santini et al. 2012; Rosario et al. 2013).
3. The fraction of X-ray AGN hosts at  $0.5 < z < 2$  detected by *Herschel*-PACS observations is systematically higher than what obtained from a control sample of inactive galaxies (Rosario et al. 2013), suggesting that X-ray AGN hosts are preferentially found in star-forming systems, therefore less likely to be “quenched” than inactive galaxies.
4. Most ( $\approx 80\%$ ) of X-ray selected AGN hosts lie on the main-sequence, while the rest of the sample are placed roughly half above and half below the MS relation (Mullaney et al. 2011). This suggests that active galaxies are preferentially hosted in star-forming systems (Rosario et al. 2013).
5. The comparison between mean SFRs of obscured AGN are consistent with those inferred for unobscured AGN, on the basis of X-ray and optical spectra (Rosario et al. 2012; Rovilos et al. 2012; Merloni et al. 2014).

Though these studies do not allow to establish to what extent the AGN activity is connected to global star formation, they strongly argue against an overall suppression of star formation in X-ray selected AGN hosts (Harrison et al. 2012). These findings are further supported by morphological studies of AGN hosts that did not find any evidence for an increasing merger fraction in AGN hosts compared to a reference sample of inactive galaxies (e.g. Cisternas et al. 2011; see Section 1.5).

These trends apply to large statistical samples, therefore may still leave some room for quenched SF in a fraction of luminous AGN hosts. Compelling *Herschel*-based evidence for possible AGN-driven quenching has been suggested in some massive radio-excess AGN (Del Moro et al. 2013), that showed lower SFRs compared to inactive galaxies at the same redshift. In addition, sub-mm CO observations of fast ( $\approx 1000$  km s $^{-1}$ ) and massive (outflow rates up to  $\approx 1200 M_\odot$  yr $^{-1}$ ) molecular outflows (e.g. Feruglio et al. 2010; Sturm et al. 2011; Veilleux et al. 2013; Cicone et al. 2014b) might be considered interesting cases of “smoking-gun” (perhaps AGN-driven) quenching, where the cold gas is rapidly ( $10^6$ – $10^8$  yr) swept outward but not yet depleted.

### Studies from IR selected samples

The above-mentioned *Herschel*-based studies have significantly reinforced the idea that most of the SMBH growth takes place in star-forming systems. This argument has been

reversing the way of investigating the AGN/Galaxy evolution, e.g. by selecting samples of star-forming galaxies and probing their mean AGN accretion activity through X-ray observations.

Studies exploiting *Herschel*-selected galaxies at  $z < 1$  have found that SFR of AGN hosts correlates with the average  $\dot{M}_{\text{bh}}$  (Rafferty et al. 2011; Chen et al. 2013). Mullaney et al. (2012a) analysed a sample of star forming BzK-selected galaxies at  $z \sim 1$  and  $z \sim 2$  in the GOODS-South field, finding that SFR and mean AGN accretion correlate linearly with galaxy  $M_{\star}$  in both epochs. These trends are consistent with the existence of the so-called ‘‘AGN main-sequence’’ in which SFR and AGN activity are tightly linked one to the other since  $z \sim 2$ . Though this simple picture fits well the similarity between CSFH and SMBH accretion history (Madau & Dickinson 2014), this interplay is not probed for individual galaxies, as reported also from studies of X-ray selected AGN hosts (e.g. Rosario et al. 2012).

A possible explanation (Mullaney et al. 2012a; Hickox et al. 2014) is that the AGN main-sequence is ‘‘hidden’’ behind different variability timescales between accretion onto the central BH ( $\approx 10^5\text{--}10^6$  yr) and the global SFR of its host ( $\approx 10^8$  yr). Indeed, deriving properties from samples of AGN implies that the resulting AGN accretion rates are averaged over galaxy evolution timescales, which are much longer compared to AGN activity (Aird et al. 2012). Since the distribution of AGN Eddington ratios does not depend significantly on the properties of AGN hosts (Bongiorno et al. 2012), studying the average AGN luminosity of a population allows to smooth over the variations of individual sources, studying the long-term  $\dot{M}_{\text{bh}}$  in any sample of galaxies.

In light of the above-mentioned studies exploiting IR and X-ray observations of  $z \lesssim 2$  AGN hosts, it has becoming progressively clear that AGN are preferentially hosted in galaxies with on-going star formation and that are mostly placed on the star forming main-sequence. Detail mechanisms for fueling AGN accretion and galaxy star formation are not yet well constrained. However, it is generally believed that major mergers might play some role in low-redshift, high-luminosity ( $L_{\text{X}} > 10^{44}$  erg  $\text{s}^{-1}$ ) AGN, while secular processes and rapid AGN variability may explain the lack of correlation between SFR and AGN luminosity in moderately luminous ( $L_{\text{X}} < 10^{44}$  erg  $\text{s}^{-1}$ ) AGN.

According to these findings, studying AGN accretion and star formation processes in *Herschel*-selected galaxies allows to investigate the interplay between these two emission components without strong biases against obscured star formation or obscured AGN activity. In this Thesis, we take advantage of FIR data taken from *Herschel* observations of large samples of star forming galaxies, combined with multi-wavelength ancillary data down to the optical/UV. This allows us to infer robust measurements of the total IR budget for each source, paving the way to isolate the possible AGN emission from that related to the host galaxy (see Chapter 3).





## SED-fitting decomposition

In this Chapter we present our broad-band SED-fitting decomposition analysis, that is devoted to isolate the AGN energy contribution from that related to the host galaxy across the electromagnetic spectrum. Distinguishing between these two emission components is extremely useful for many applications. The most robust and related to the present work are:

1. to derive integrated galaxy properties, such as SFR and  $M_*$ , both corrected for the contribution due to the presence of a possible AGN (Section 3.2);
2. to calculate AGN bolometric luminosities for both unobscured and obscured AGN (Section 3.3 and 3.6.2);
3. to identify star forming AGN hosts from “inactive” galaxies, on the basis of the statistical incidence of the AGN emission component across the observed SED (Section 3.4).

FIR photometry taken from PACS and SPIRE instruments onboard *Herschel* gives an invaluable contribution to assess the star formation content of each galaxy, since FIR data are sensitive to the star formation rest-frame peak over a wide range of redshifts. However, an additional contribution can arise from warmer nuclear dust heated by material accreting onto the SMBH rather than stellar processes. The combined *Spitzer* and *Herschel* capabilities are paramount to reach a complete assessment of the dust content at different temperatures stored across the IR domain. This is the reason why a broad-band photometric coverage down to the optical/UV regime is necessary to fully characterise each observed SED and to constrain galaxy templates towards a meaningful best-fit solution.

In Section 3.1 we describe the sample of *Herschel*-selected galaxies that we have exploited from PEP, HerMES and GOODS-H surveys in the GOODS-S, GOODS-N and COSMOS fields (see Section 2.3). Here we focus on the multi-wavelength identification and redshift measurements of our sources. Subsequently, the SED-fitting decomposition technique is presented and discussed in Section 3.2 and 3.3. We performed a statistical analysis to classify *Herschel* sources between AGN host galaxies or “inactive” galaxies (Section 3.4), investigating also how the fraction of sources identified as AGN evolves with

redshift and IR luminosity (Section 3.5). Different tests have been made in Section 3.6 to compare AGN luminosities and SED-fitting classification with those inferred from independent AGN diagnostics. Finally, the use of SED-fitting decomposition to search for heavily obscured AGN is discussed in Section 3.6.3 and 3.6.4.

### 3.1 The parent sample

Our parent sample exploits *Herschel*-PACS (Data Release 1) and SPIRE observations in the GOODS-N ( $\approx 190$  arcmin<sup>2</sup>), GOODS-S ( $\approx 190$  arcmin<sup>2</sup>) and COSMOS (about 2 deg<sup>2</sup>) fields. In the following we present our *Herschel*-selected sample and briefly mention the cross-match with multi-wavelength ancillary photometry, referring the reader to Lutz et al. (2011); Berta et al. (2011); Oliver et al. (2012) and Magnelli et al. (2013) for a detailed explanation of data reduction and construction of multi-wavelength catalogues.

#### 3.1.1 Multi-wavelength identification

In both GOODS fields, the *Herschel*-selected sample includes all sources with  $\geq 3\sigma$  flux density in at least one PACS band. FIR data are taken from the blind catalogues described in Magnelli et al. (2013) that combine the data of PEP (Lutz et al. 2011) and GOODS-H (Elbaz et al. 2011) surveys. In the COSMOS field, we consider all sources with  $\geq 3\sigma$  flux density in at least one PACS band. In addition, our sample also includes SPIRE-250  $\mu\text{m}$  selected galaxies (Oliver et al. 2012) without a counterpart in PACS observations. We exploited this double selection only in COSMOS, since SPIRE-250  $\mu\text{m}$  observations in this field reach similar sensitivities to those achieved by PACS (see Section 2.3). We also checked that the redshift distribution and the average  $L_{\text{IR}}$  obtained from SPIRE-selected sources without PACS detection closely resembles that of PACS-selected galaxies. As a consequence, the inclusion of SPIRE-selected sources in COSMOS allows roughly to double our sample of star forming galaxies in this area without introducing a significant bias in our analysis.

The cross-match between FIR data and the extensive broad-band photometry available in all these fields has been accomplished via a likelihood ratio algorithm (Sutherland & Saunders 1992; Ciliegi et al. 2001). FIR-to-MIR associations have been made by starting from the longest PACS wavelength available, and progressively matching with PACS detections at shorter wavelengths, down to MIPS-24  $\mu\text{m}$  (see Lutz et al. 2011). In all fields, MIPS-24  $\mu\text{m}$  positions have been used as priors to extract SPIRE fluxes up to 500  $\mu\text{m}$  (Roseboom et al. 2012).

In both GOODS fields, the cross-match from MIR to optical/UV wavelengths is described in detail by Berta et al. (2010, 2011) and Magnelli et al. (2013). Briefly, MIPS-24  $\mu\text{m}$  fluxes have been cross-matched with *Spitzer*-IRS fluxes at 16  $\mu\text{m}$  (when available) and *Spitzer*-IRAC detections down to 3.6  $\mu\text{m}$ , using the latter positions as priors. PSF-matched photometry at shorter wavelengths connects IRAC fluxes with those at NIR ( $J$ ,  $H$ ,  $K_s$ ), optical ( $b$ ,  $v$ ,  $i$  and  $z$  bands with HST-ACS) and UV (WFI 3500 and 3800Å in the GOODS-S;  $U$ -band with CFHT in the GOODS-N) wavelengths, collecting up to 21 and 19

pass-bands in the GOODS-S and GOODS-N, respectively. These catalogues provide with 892 (GOODS-S) and 850 (GOODS-N) FIR-selected sources with broad-band photometry down to the optical/UV.

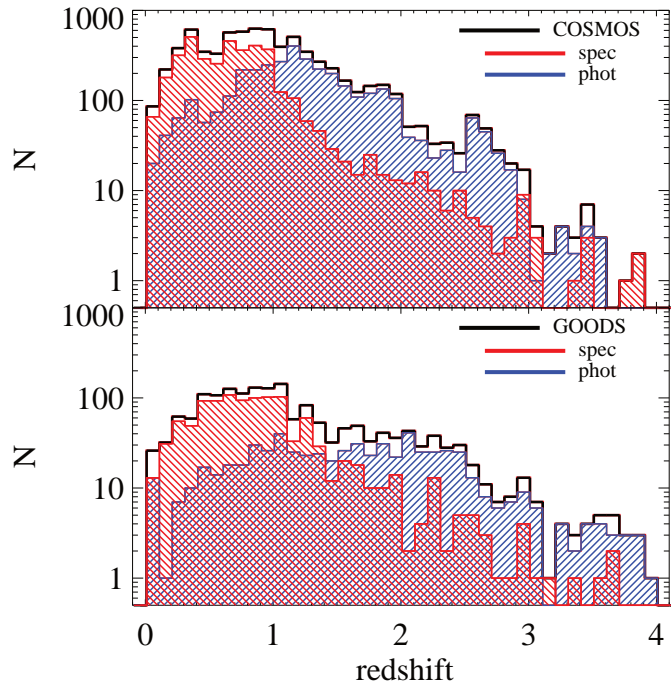
An extensive ancillary photometric is available also in COSMOS, where MIPS-24  $\mu\text{m}$  fluxes of PEP/HerMES sources have been cross-matched to the optically-based catalogues (Capak et al. 2007; Ilbert et al. 2009), via a nearest neighbour method. These associations allow to collect up to 19 pass-bands, spanning MIR (MIPS-24  $\mu\text{m}$  and IRAC bands), NIR ( $J$ ,  $K_s$  bands from CFHT), as well as optical/UV ( $U$ ,  $b$ ,  $g$ ,  $v$ ,  $r$ ,  $i$ ,  $z$  imaging taken from CFHT) wavelengths. Totally, the number of FIR sources in COSMOS with either PACS or SPIRE-250  $\mu\text{m}$  detections is about 17000. This sample has been limited to the sub-area covered by *Chandra*-COSMOS (Elvis et al. 2009; Civano et al. 2012) observations (i.e. 0.9 instead of 2  $\text{deg}^2$ ), since our sample will be cross-matched with available X-ray data in the common area (see Section 5.2). This cut leads to a total number of 7272 sources.

### 3.1.2 Spectroscopic and photometric redshifts

We collect a total number of 8964 *Herschel*-selected galaxies across all fields, out of which 8948 at  $z \leq 4$ . Extensive redshift compilations are publicly available in COSMOS and in the GOODS-S/N fields. In Fig. 4.1 the redshift distributions for the GOODS and COSMOS samples are shown, distinguishing between spectroscopic (red) and photometric (blue) measurements. We defer the reader to Berta et al. (2011) for a careful description of redshift catalogues. Here below we provide a list of references of redshift measurements. The redshift completeness for our *Herschel* sources with counterpart in optical/near-infrared catalogs reaches 100% in all fields.

- GOODS-S: we extended the original spectroscopic sample presented by Grazian et al. 2006 and Santini et al. 2009 in the GOODS-Multi-wavelength Southern Infrared Catalog (GOODS-MUSIC) using redshifts from public spectroscopic database<sup>1</sup> including more than 3000 sources (Cristiani et al. 2000; Croom et al. 2001; Bunker et al. 2003; Dickinson et al. 2004; Stanway et al. 2004a,b; Strolger et al. 2004; Szokoly et al. 2004; van der Wel et al. 2004; Doherty et al. 2005; Le Fèvre et al. 2005; Mignoli et al. 2005; Vanzella et al. 2008; Popesso et al. 2009; Santini et al. 2009; Balestra et al. 2010; Cooper et al. 2012; Kurk et al. 2013). We used this database to enrich the GOODS-MUSIC spectroscopic catalogue, replacing photometric redshifts with the up-to-date spectroscopic ones in the case of “secure” redshift determination (according to the scale adopted by Balestra et al. (2010)). The final spectroscopic fraction reaches 67%. In this field, uncertainties on photometric redshifts have been estimated by Grazian et al. (2006), who determined an average absolute scatter  $\langle |\Delta z / (1 + z)| \rangle = 0.045$  over the whole redshift range  $0 < z < 6$ .
- GOODS-N: photometric redshifts have been derived by Berta et al. (2011) by using the EAZY (Brammer et al. 2008) code, as described in Wuyts et al. (2011a). In

<sup>1</sup>Publicly available at: <http://www.eso.org/sci/activities/garching/projects/goods/MasterSpectroscopy.html>



**Figure 3.1:** Redshift distribution of 8948 *Herschel* galaxies in the COSMOS (top panel) and in the GOODS-S/N (bottom panel) fields at  $z \leq 4$ . Spectroscopic and photometric redshifts are shown in red and blue, respectively, while the black line is the sum of the two. Note that the scale of the  $y$ -axis is logarithmic.

addition, Berta et al. (2011) integrated photometric measurements by adding spectroscopic redshifts from Barger et al. (2008) for about 64% of the sample. Photometric redshift uncertainties were inferred by Berta et al. (2011), who estimated a median absolute deviation ( $\text{MAD}^2$ ) of  $\approx 0.04$ , while the expected fraction of outliers, defined as objects having  $\langle |\Delta z / (1 + z)| \rangle \geq 0.2$ , is as high as  $\sim 6\%$ .

- COSMOS: we used photometric redshifts from Ilbert et al. (2010) and Wuyts et al. (2011a). For *Herschel* sources that are detected by *Chandra* observations (see Section 5.2.1), we cross-matched our dataset with photometric redshifts presented by Salvato et al. (2011), which are more suitable for AGN dominated sources. We retrieved spectroscopic measurements from the  $z$ COSMOS survey by Lilly et al. (2007, 2009), both the public  $z$ COSMOS-bright and the proprietary  $z$ COSMOS-deep data base. We also browsed the most recent public spectroscopic surveys, replacing our photometric redshifts with spectroscopic ones in case they are flagged as “secure”: Ahn et al. (2014), from Data Release 10 of the Sloan Digital Sky Survey (SDSS); Coil et al. (2011) from the PRISM Multi-object Survey (PRIMUS) catalogue; Trump et al. (2009) from the COSMOS-Magellan spectroscopic catalogue. Globally, about 50% of the COSMOS FIR-selected sample have a spectroscopic redshift. Ilbert et al. (2010) evaluated the  $\langle |\Delta z / (1 + z)| \rangle$  of photometric redshifts to range from 0.008 (for

<sup>2</sup> $\text{MAD}(x) = \text{median}(|(x) - \text{median}(x)|)$ , where  $x$  represents the difference between spectroscopic and photometric redshift.

$I < 22.5$ ) to 0.053 (for  $24 < I < 25$ ), while the more up-to-date photometric redshifts provided by Berta et al. (2011) have a MAD around 0.01. The expected fraction of outliers in this database reaches  $\sim 2\%$ .

## 3.2 Fitting broad-band SEDs

In this section we present the SED-fitting analysis performed for the *Herschel* sample described above. Each observed SED has been fitted making use of the MAGPHYS code<sup>3</sup> (da Cunha et al. 2008) and a modified version of this code presented by Berta et al. 2013b).

### 3.2.1 SED-fitting with MAGPHYS

MAGPHYS is a public code which uses physically motivated templates to reproduce the observed SEDs from the ultraviolet to sub-mm wavelengths. It has been developed especially to fit star-forming galaxy SEDs, so that it is well suited for our purposes (see Smith et al. 2012 for an extensive application of MAGPHYS on *Herschel* data). The MAGPHYS code is easily applicable to any modern galaxy survey: indeed, the model technique is particularly versatile for constraining galaxy SEDs in a wide range of star formation histories, since the input stellar synthesis template library can be adjusted accordingly. Through broad-band SED-fitting, the MAGPHYS code enables to interpret in a consistent way the UV, optical and IR SEDs of galaxies. Indeed, the basic recipe that MAGPHYS relies on is a self-consistency between the energy originating from stars, which is partially absorbed by dust, and the infrared light due to re-emission by the dust itself. Briefly, the stellar and dust component are computed as follows (see da Cunha et al. 2008 for details).

- The Bruzual & Charlot (2003) library of templates is used to compute the integrated light produced by stars in galaxies. These models predict the spectral evolution of stellar populations at  $\lambda \leq 10 \mu\text{m}$  and ages between  $10^5$  and  $10^{10}$  yr, starting from a Chabrier (2003) IMF and combining parameters representative of different metallicities, star formation histories and dust contents.
- Each stellar template is modified by the effect of dust extinction, which is parametrised as the cumulative attenuation produced by interstellar medium (ISM) and dusty molecular clouds, according to the angle-averaged model of Charlot & Fall 2000.
- The global dust emission is ascribed to three main constituents of interstellar dust: polycyclic aromatic hydrocarbons (PAHs, in the range  $3\text{--}12\mu\text{m}$ ); mid-IR continuum from hot (130–250 K) dust grains; and thermal emission, both from warm (30–60 K) and cold (15–25 K) dust grains in thermal equilibrium, according to da Cunha et al. (2008) prescriptions.

The analysis of each single observed spectral energy distribution with MAGPHYS starts by building up a large library of templates at the same redshift of the source. Given the set of parameters incorporated in the code, as well as the wide range of values adopted by

<sup>3</sup>MAGPHYS can be retrieved at <http://www.iap.fr/magphys/magphys/MAGPHYS.html>

each of them, in the end 50 000 stellar and 50 000 independent dust emission models are computed for each source. Only about 30% of the IR templates can be combined with each stellar model, because of the underlying assumption that all the absorbed UV and optical stellar radiation has to be accounted for in the integrated infrared emission. As a result of such selection, typically  $10^9$  different self-consistent (i.e. stellar+dust) model combinations are created to reproduce each observed SED from the UV to the FIR. Then, each model combination is compared to the observed SED and the stellar+dust model normalization is rigidly scaled to fit the data as well as possible.

### Uncertainties on output parameters

A Bayesian approach is implemented to build a marginalised likelihood distribution (or probability distribution function, PDF) for each parameter. This function is given by the distribution of values of a given parameter across the whole library of templates. This statistical analysis accounts for the degeneracy level in the parameter space, which occurs when a relatively small set of data points is compared with such a large range of models.

According to the *Bayes Theorem*, it is possible to write the probability  $P(M/D)$  to select the model  $M$  as the best-fit to the observed data  $D$  as:

$$P(M/D) = P(M) \times P(D/M) \quad (3.1)$$

where  $P(M)$  is the frequency of sampling the model  $M$ , while  $P(D/M)$  is the “likelihood function” that weights the goodness of the fit obtained from the model  $M$  on the basis of the resulting  $\chi^2$ . This function is generally parametrised as  $e^{-\chi^2/2}$ .

### 3.2.2 Fitting with MAGPHYS + AGN

MAGPHYS does not consider a potential AGN contribution, but only star formation processes are involved in building the SEDs. Since we are interested in probing the possible AGN contribution to the observed SEDs, we make use of a modified version of the code (Berta et al. 2013b), which accounts also for a possible nuclear heating source.

The code by Berta et al. (2013b) includes the AGN component and differs from the original MAGPHYS in the model normalization approach. The stellar+dust normalization is not scaled to fit the data, but is allowed to float within a range of normalization values (the same range adopted by the original version of MAGPHYS). For each normalization, the stellar+dust model is subtracted from the observed photometry, and the AGN emission is then added to reproduce what is left out from this subtraction. Thanks to the MAGPHYS structure, the dust emission remains self-consistently connected to the stellar optical component. This approach results in a simultaneous 3-component model, where the quest for the best-fit solution is still based on  $\chi^2$  minimization.

### AGN library of templates

The AGN templates implemented by Berta et al. (2013b) in the MAGPHYS code have been developed by Fritz et al. (2006) and Feltre et al. (2012). Their library of templates includes around 2376 AGN models, each one analysed for ten different orientation angles

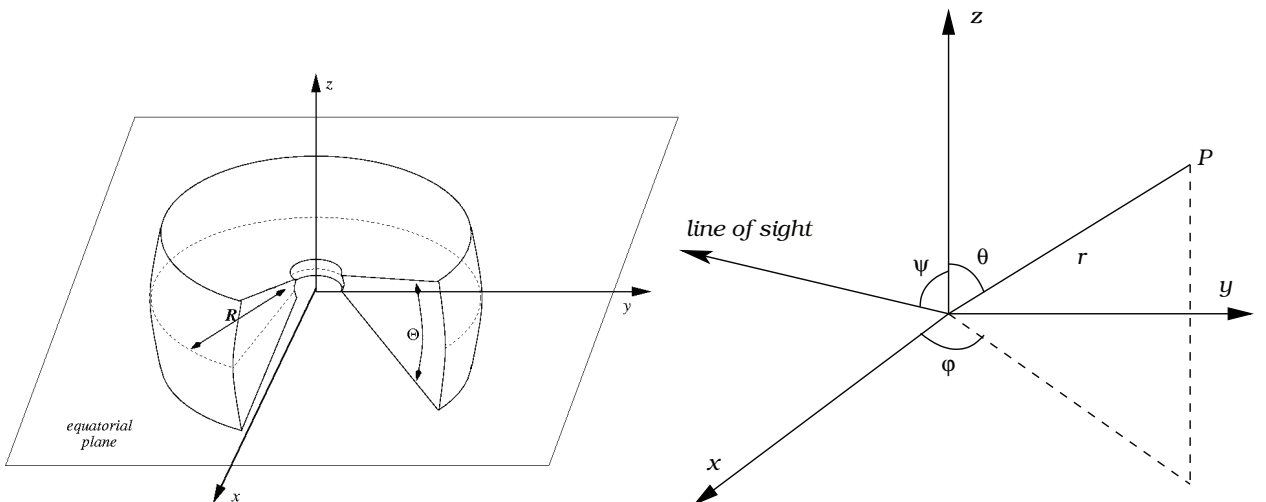
with respect to the line of sight,  $\theta$ , which is uniformly distributed from edge-on to face-on (for a total of 23760 AGN templates). This set of templates is widely used in the literature, since it is physically motivated and suitable in reproducing a wide variety of AGN features, such as hot circumnuclear dust in mid-IR dominated SEDs (e.g. Mrk3, see Mullaney et al. 2011) and the silicate features at 9.7 and 18  $\mu\text{m}$  (see Feltre et al. 2012). Each AGN template includes both the contribution from the central engine (i.e. accretion disc) and the reprocessed emission from the surrounding dusty torus.

The intrinsic luminosity of the central AGN,  $\lambda L_\lambda$ , is modeled as a broken power law with four wavelength-dependent spectral indexes:

$$\lambda L_\lambda \propto \begin{cases} \lambda^{1.2} & \text{if } 0.001 \leq \lambda < 0.03 \text{ } [\mu\text{m}] \\ \lambda^0 & \text{if } 0.03 \leq \lambda < 0.125 \text{ } [\mu\text{m}] \\ \lambda^{-1} & \text{if } 0.125 \leq \lambda < 20.0 \text{ } [\mu\text{m}] \\ \lambda^{-3} & \text{if } \lambda \geq 20.0 \text{ } [\mu\text{m}] \end{cases} \quad (3.2)$$

The adopted spectral indices are motivated by the comparison with observations.

1.  $0.001 \leq \lambda < 0.03 \text{ } \mu\text{m}$ : the spectral index is chosen to match theoretical models for accretion discs built by Hubeny et al. (2000);
2.  $0.03 \leq \lambda < 0.125 \text{ } \mu\text{m}$ : the adopted value comes from the comparison with observed optical/NIR slope for Radio Quiet AGN (Zheng et al. 1997);
3.  $0.125 \leq \lambda < 20.0 \text{ } \mu\text{m}$ : this slope is determined from the analysis of broad-band SEDs in SDSS galaxies (Hatziminaoglou et al. 2008);
4. At longer wavelengths, the emission scales as a Planck spectrum in the Rayleigh-Jeans regime.



**Figure 3.2:** Schematic representation of the smooth torus geometry used by Fritz et al. (2006) and Feltre et al. (2012).

The torus model has a flared-disc geometry (Efstathiou & Rowan-Robinson 1995; van Bemmell & Dullemond 2003) depicted as two concentric spheres which set the inner and

the outer radii, without the polar cones (Fig.3.2, right). The inner radius  $R_{\min}$ , generally identified as the dust sublimation radius, depends on the accretion luminosity  $L_{\text{bol,AGN}}$ <sup>4</sup> ascribed to the central engine, according to the following expression:

$$R_{\min} \simeq 1.3 \sqrt{L_{\text{bol,AGN}} \cdot T_{1500}^{-2.8}} [\text{pc}] \quad (3.3)$$

where  $L_{\text{AGN}}$  is in units of  $10^{46} \text{ erg s}^{-1}$  and  $T_{1500}$  corresponds to the dust grains sublimation temperature, in units of 1500 K (Barvainis 1987). The chemical composition of dust grains is a mixture of silicates and graphite. Emission from silicates is necessary to reproduce the spectral feature at  $9.7 \mu\text{m}$ . Graphite is used to reproduce circum-nuclear dust emission close to the sublimation radius, with temperatures up to 1500 K.

The dusty structure here considered is thought to be a “smooth” torus (Fig.3.2, left). This implies that the mixture of dust grains follows monothonic density profile, both on the equatorial plane and from different lines of sight  $\theta$ . Each torus model is univocally identified by a set of six different parameters: the outer to inner radius ratio  $R_{\text{out}}/R_{\text{inn}}$ ; the opening angle  $\Theta$ ; the equatorial optical depth at  $9.7 \mu\text{m}$   $\tau_{9.7}$ ; the radial and height slopes of the density profile ( $\beta$ ,  $\gamma$ ); and the viewing angle  $\theta$  of the line of sight.

During the AGN fitting procedure, 100 different AGN templates are randomly selected for each source and for each stellar+dust realization. This allows to optimise the random sampling by exploring a wider range of the AGN parameter space for each source.

By solving the radiative transfer equation for a smooth dusty structure irradiated by the accretion disc, a bolometric correction (BC) value is available for each AGN template. Such information allows us to calculate the AGN bolometric luminosity ( $L_{\text{bol,AGN}}$ ) directly from the infrared (1–1000  $\mu\text{m}$  rest frame) luminosity of the best-fit AGN model, according to the following expression:

$$k_{\text{bol}} = \frac{L_{\text{bol,AGN}}}{L_{1-1000}^{\text{agn}}} \quad (3.4)$$

Our  $L_{\text{bol,AGN}}$  estimates are derived without double counting the dust-reprocessed contribution arising from the accretion disc, since the radiative transfer recipe performs a self-consistent re-distribution of the input energy at all wavelengths, according to the adopted set of templates.

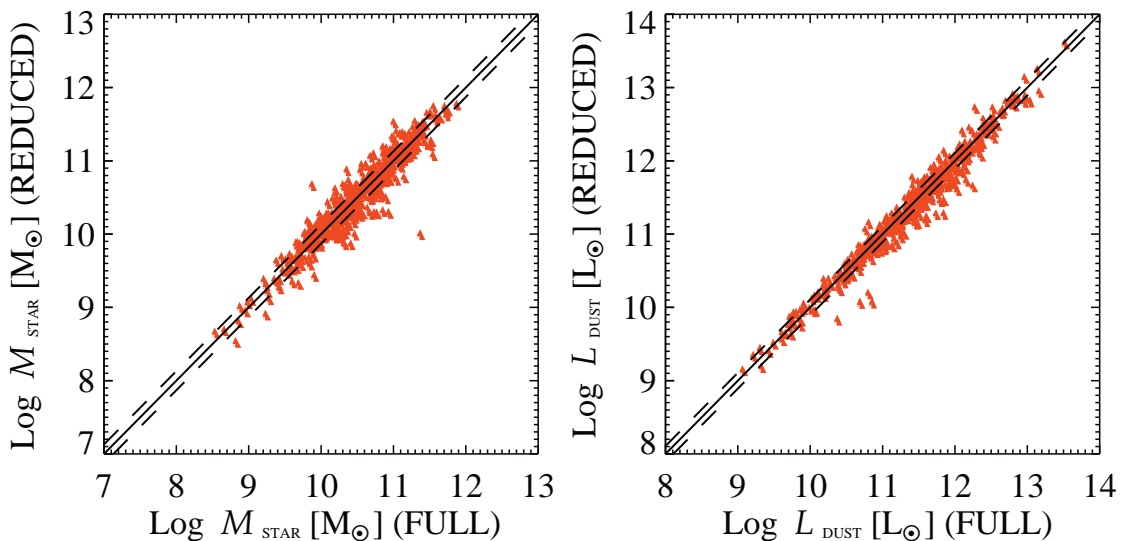
### Reduced library of templates

The set of possible model combinations included in the original version of MAGPHYS is very large (around  $10^9$  possible configurations, see Section 3.2.1), as well as the set of physical parameters that the stellar+dust emission models rely on. The range of values spanned by each parameter is physically reasonable, but the computational time taken by the code strongly increases when the AGN component is added to the SED-fitting procedure. For these reasons we take advantage of a reduced grid of stellar+dust templates, as described by Berta et al. (2013b).

The reduced grid has been built up by randomly selecting 1000 stellar models and 1000 dust emission models for each stellar template, giving a total of  $\sim 3 \times 10^5$  self-consistent

<sup>4</sup>The terms “accretion” and “bolometric” AGN luminosity adopted throughout this Thesis are assumed to have the same physical meaning.





**Figure 3.3:** Comparison between the values obtained with the full (*x-axes*) and reduced (*y-axes*) MAGPHYS library of templates for three different integrated parameters: stellar mass and infrared (8–1000  $\mu\text{m}$ ) luminosity due to star formation. The full grid uses all the available templates of galaxy SEDs encoded in MAGPHYS, whereas the reduced grid accounts for a small fraction of them resulting from a random selection (see text for details). Red triangles are the best-fit values of each parameter as taken from the GOODS-S sample. The black solid line is the bisector, while the dashed lines represents the  $\pm 1\sigma$  confidence regions. The resulting  $1\sigma$  uncertainties related to these parameters are 0.13 and 0.11 dex, respectively.

templates for each source (after the 30% cut required by the energy balance argument). Such selection uniformly spans the parameter space covered by the original library, without introducing any significant bias in the adopted grid. In Fig. 3.3 we compare the values of some integrated physical quantities (stellar mass and IR luminosity) obtained using the full and the reduced grid of MAGPHYS<sup>5</sup>. The plots show that mostly the integrated physical parameters inferred by adopting the full and the reduced grids are reasonably consistent.

As described in Section 3.2.2, each AGN template is defined by the combination of six physical parameters. We limit the torus library of templates by restricting the multi-dimensional volume of the parameter space. In particular, we removed the largest optical depth value at 9.7  $\mu\text{m}$  ( $\tau_{9.7}=10$ , see e.g. Pier & Krolik 1992) and the most extended geometries ( $R_{\text{out}}/R_{\text{inn}} > 100$ ), as no evidence for such large structures has yet been found (Jaffe et al. 2004; Tristram et al. (2007, 2009).

After this cut, the torus grid includes 504 templates, each one computed at 10 different lines of sight. As mentioned before, a random selection of 100 / 5040 AGN configurations is performed for each source and stellar+dust model combination. This reduced library spans several values of the ratio  $R_{\text{out}}/R_{\text{inn}}$  (10, 30, 60 and 100), the opening angle  $\Theta$  ( $40^\circ$ ,  $100^\circ$  and  $140^\circ$ ), the equatorial optical depth  $\tau_{9.7}$  (0.1, 0.3, 0.6, 1, 2, 3 and 6) and the slopes of the density profile ( $-1$ ,  $-0.5$  and  $0$  for  $\beta$ ;  $0$  and  $6$  for  $\gamma$ ). The resulting library turns out to be similar to the one used by Pozzi et al. (2012) and Hatziminaoglou et al. (2008,

<sup>5</sup>Since the SED-fitting procedure using the full MAGPHYS library of templates is very time demanding, this check has been limited only to sources in the GOODS-S.

2009), which we refer to for a detailed description of the degeneracy level drawn by torus parameters.

Since the only crucial parameter in the following analysis is the AGN bolometric correction ( $k_{\text{bol}}$ ), used to convert the total infrared AGN emission into AGN bolometric luminosity, we show in Fig. 3.4 how much our restrictions modify the shape of the distribution of the bolometric corrections. In particular, the reduced grid of AGN templates reaches  $k_{\text{bol}}$  values around 60, while the full grid samples up to  $k_{\text{bol}} \sim 150$ . The largest values are mostly representative of the most obscured AGN ( $\tau_{9.7}=10$ ). Despite the existing cut in  $k_{\text{bol}}$ , we point out that the missed values (i.e.  $k_{\text{bol}} > 60$ ) constitute a negligible number of AGN templates with respect to the full grid ( $\sim 0.3\%$ ).

### 3.3 AGN Bolometric Luminosity and Bolometric corrections

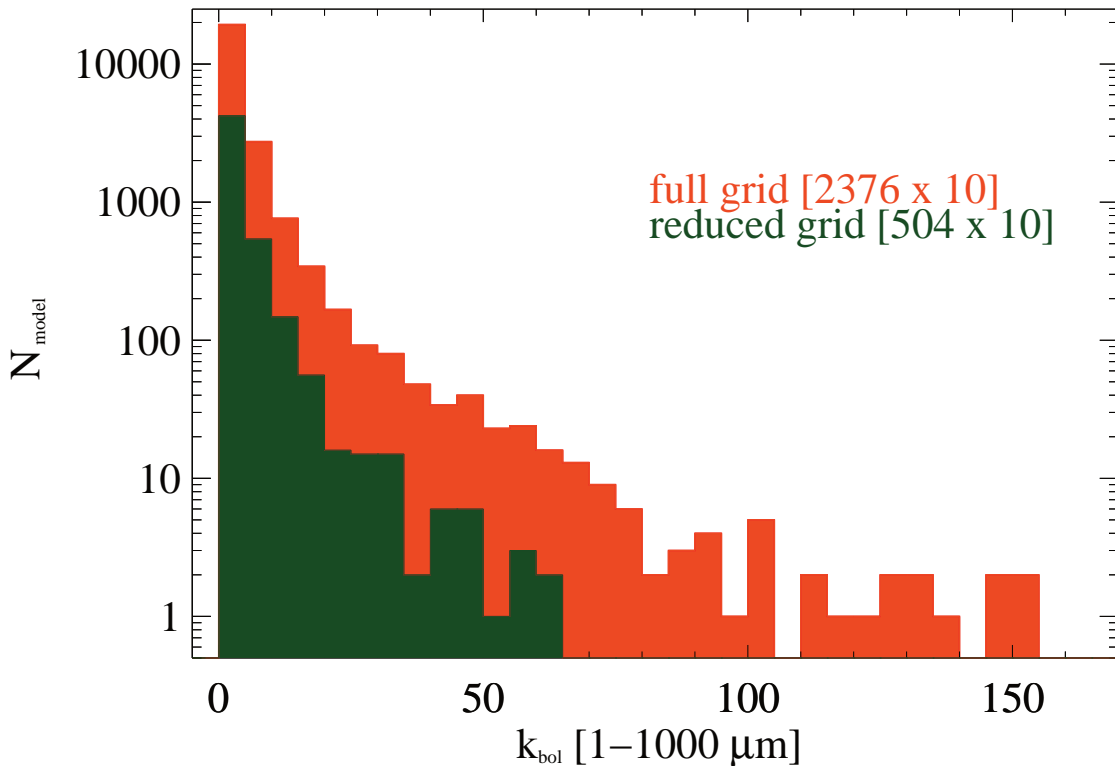
For every observed SED, a PDF is created for each parameter to trace its confidence range. As mentioned before, among the physical parameters returned by the code, we have focused on the AGN bolometric luminosity only (computed by accounting for the rest-frame 1-1000  $\mu\text{m}$  AGN emission and adopting its related bolometric correction).

We stress that each  $L_{\text{bol,AGN}}$  value is not properly a bolometric luminosity, since the input (i.e. accretion) energy of the central engine is limited to the range  $10^{-3} - 10^3 \mu\text{m}$  (i.e.  $10^{-1} - 10^{-7}$  keV). This means that the X-ray emission is supposed to be negligible in terms of input energy. Such an assumption relies on the fact that large  $k_{\text{bol}}$  values have been found in hard X-rays, of the order of 20–30 (Risaliti & Elvis 2004; Marconi et al. 2004; Hopkins et al. 2007; Pozzi et al. 2007; Vasudevan & Fabian 2009; Lusso et al. 2012), increasing as a function of  $L_{\text{bol,AGN}}$ .

In Fig. 3.5 we show a comparison between distributions of bolometric corrections computed in various bands. On the one hand, Hopkins et al. (2007) derived their  $k_{\text{bol}}$  values by assuming different amounts of obscuration applied to an average, unobscured, type-1 AGN SED (Richards et al. 2006). They found  $k_{\text{bol}}$  to depend on the input AGN luminosity, in agreement with other works (e.g. Marconi et al. 2004; Lusso et al. 2012).

We derive the distribution of  $k_{\text{bol}}$  by integrating the best-fit AGN model of each source (if it fulfills the criterion imposed by the Fisher test, as explained in Section 3.4) over the whole IR domain (1–1000  $\mu\text{m}$ , rest-frame). Each black circle represents the median  $k_{\text{bol}}$  value on different luminosity bins ( $= 0.3$  dex wide). Error bars correspond to  $\pm 1\sigma$  uncertainties and are centered on the median value of the respective bin. We find that  $k_{\text{bol}}$  does not change significantly with increasing AGN bolometric luminosity  $L_{\text{bol,AGN}}$ . This result could suggest that the fraction of the input energy coming from the central engine which is absorbed by the obscuring structure and re-emitted in the IR does not change significantly as a function of  $L_{\text{bol,AGN}}$ .

The average bolometric corrections derived from the IR are much lower than those derived from X-ray measurements (see Fig. 3.5). Indeed, the distribution of 1–1000  $\mu\text{m}$  bolometric corrections has a median value of  $\sim 3.8$  over the whole range of  $L_{\text{bol,AGN}}$  covered by our sample, ranging on average from 2 to 5.



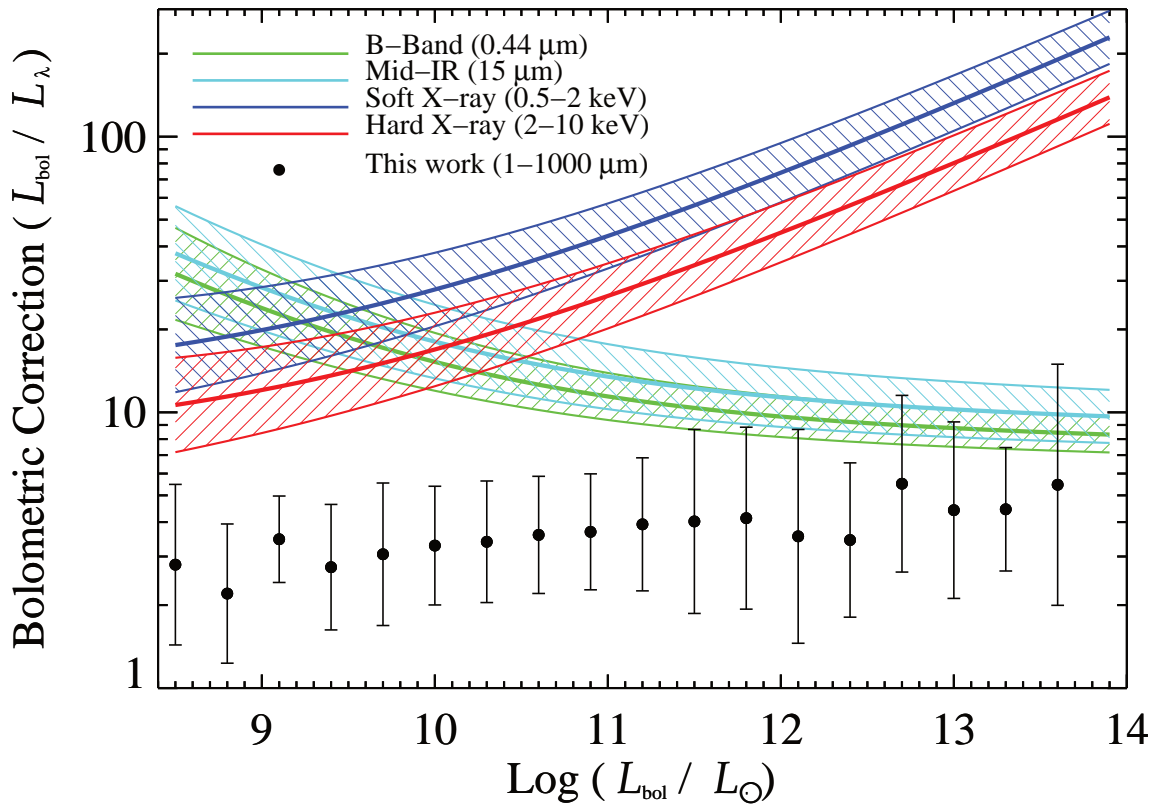
**Figure 3.4:** Bolometric correction (1–1000  $\mu\text{m}$ , rest-frame) distributions to turn the infrared AGN luminosity into bolometric luminosity. The red area traces the full grid of the library by Feltre et al. (2012) and Fritz et al. (2006). The green area represents the reduced library that is adopted in this work.

### 3.4 Testing the AGN incidence

In order to evaluate the impact of an additional AGN component on a “pure” star-forming galaxy SED, for each source we run both MAGPHYS and MAGPHYS+AGN. As mentioned in Section 3.2.2, a reduced grid has been adopted in both fitting procedures. First, we run the code version including the AGN and during the random selection of stellar and dust templates their template identification numbers are registered. Once the SED-fitting with MAGPHYS+AGN is accomplished, we run the MAGPHYS code for each source by using the same set of stellar+dust templates previously extracted for the same source. This implementation allows us to test if the fit provided by similar star formation templates improves when accounting for an additional nuclear emission component.

In Fig. 3.6 three different examples of best-fits are shown as a result of this SED decomposition. The integrated emission has been decomposed into three emission contributions, according to the recipe previously described. The expected stellar light (green line) is partially absorbed by a mixture of dust grains and re-distributed in the MIR/FIR domain (blue line). The AGN emission is the red curve. The respective  $\chi^2$  values are then used to estimate the significance of the improvement of the fit occurring when a nuclear component is added.

Around 98% of the *Herschel* sample fitted with the MAGPHYS+AGN code includes an



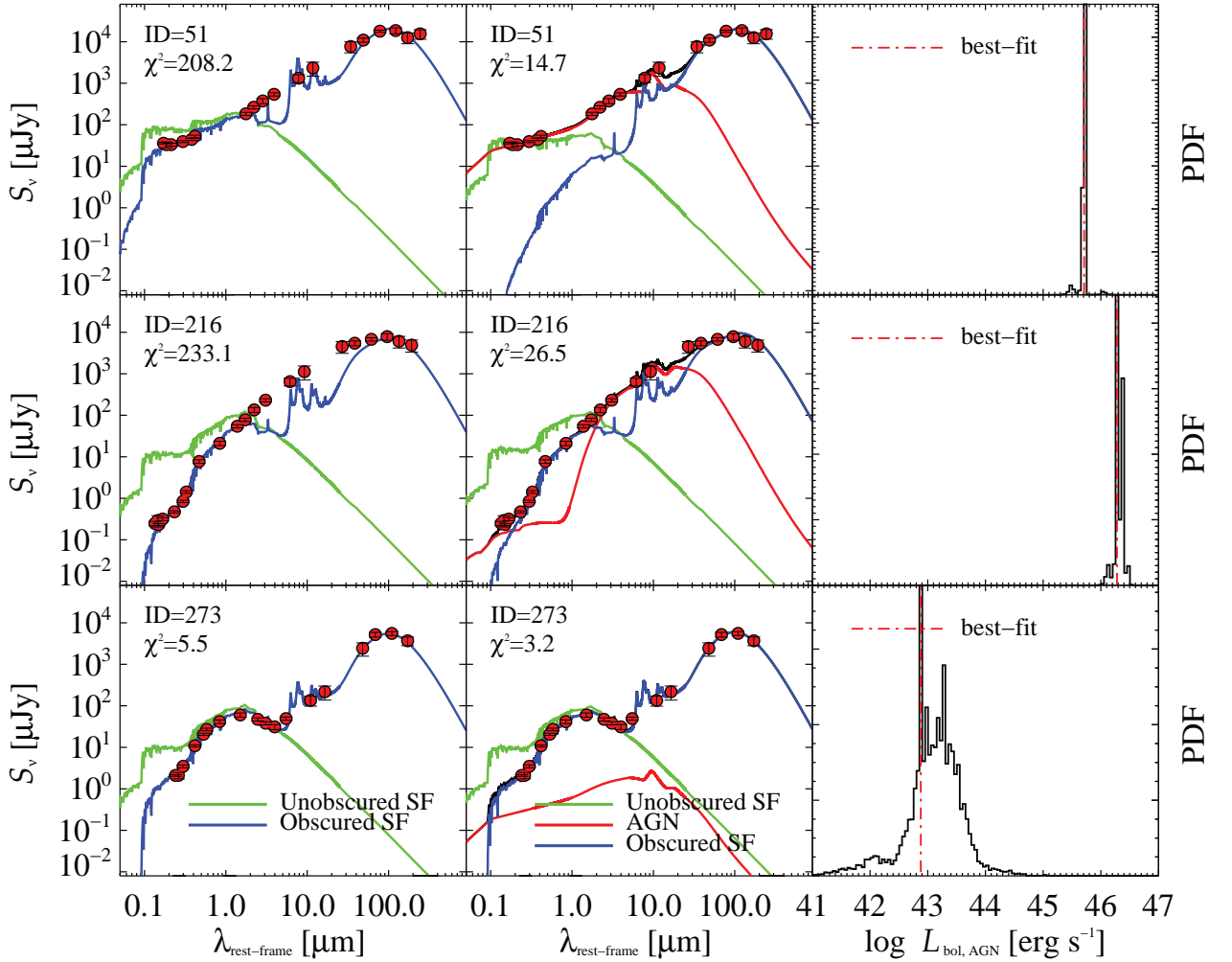
**Figure 3.5:** Distributions of AGN bolometric corrections. The regions traced by solid lines represent various trends of  $k_{\text{bol}}$  computed in different bands by Hopkins et al. (2007). Black circles (with  $\pm 1\sigma$  uncertainties) come from the sample analysed in this work, only including sources with relevant signatures of nuclear activity (see Section 3.4 for details).

AGN component in the best-fit model. However, the addition of the AGN component does not always significantly improve the best-fit. For instance, the last source shown in Fig. 3.6 turns out to exhibit negligible signatures of AGN activity over the entire wavelength range. To pinpoint the significance of the AGN component on the global SED we apply a Fisher test ( $F$ -test) to the two SED-fitting runs, as presented by Bevington & Robinson (2003):

$$F_{\text{test}} = \frac{\chi_{\text{no agn}}^2 - \chi_{\text{agn}}^2}{\chi_{\nu, \text{agn}}^2} \geq F_{\text{threshold}}(\text{CL}) \quad (3.5)$$

with  $\chi_{\text{no agn}}^2$  and  $\chi_{\text{agn}}^2$  being the  $\chi^2$  values of the best-fit without and with the AGN, respectively, whereas  $\chi_{\nu, \text{agn}}^2$  is the reduced  $\chi^2$  value referred to the best-fit with an AGN component. If this ratio exceeds a given threshold, then the source of interest is included in the following analysis, otherwise it is ruled out and a null AGN contribution is assumed for its specific SED. The threshold value depends on the number of degrees of freedom as well as on the required confidence level (CL).

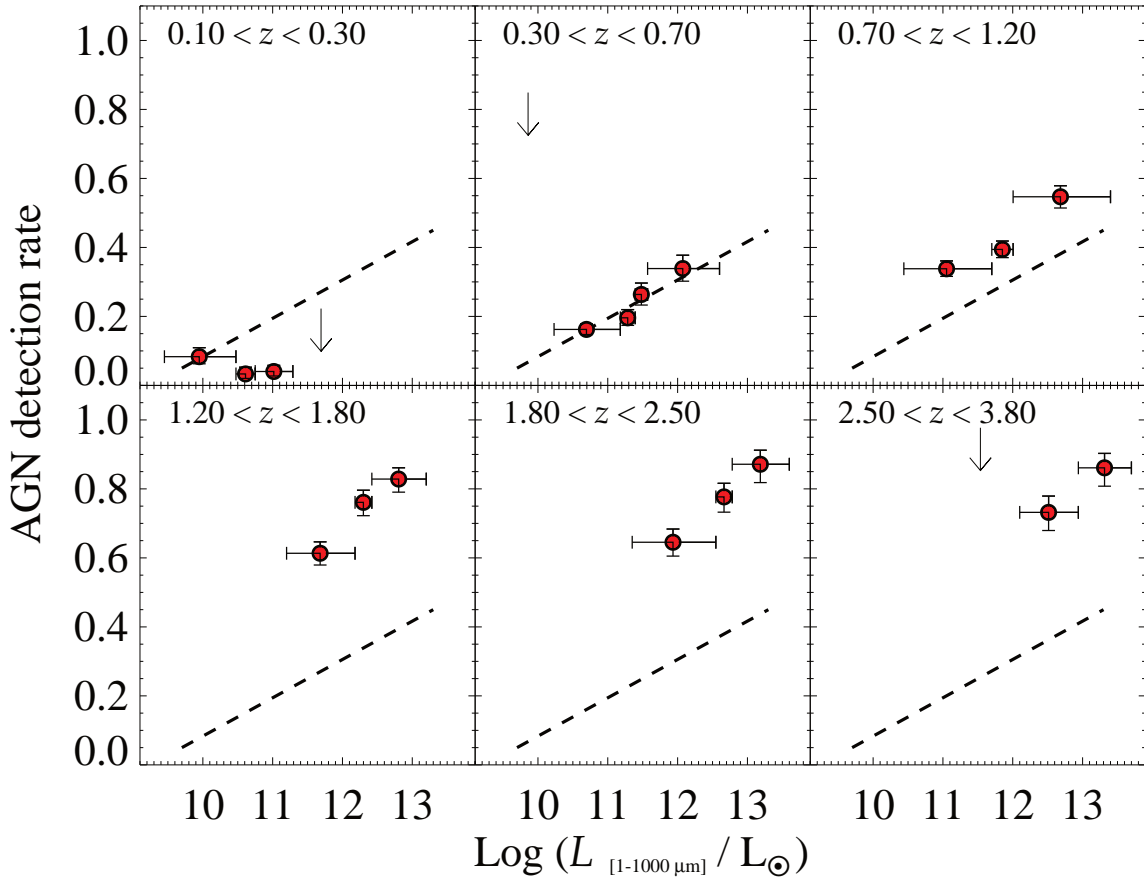
The overall fraction of  $\geq 99\%$  significant AGN reaches 37%, while it increases to 45% and 55% at  $\geq 95\%$  and  $\geq 90\%$  confidence levels, respectively. We label as AGN only those sources satisfying the  $F$ -test at the  $\geq 99\%$  confidence level, as this threshold is



**Figure 3.6:** Three examples of SED decomposition performed without AGN (*left panels*) and with AGN (*central panels*). The corresponding PDF of the AGN bolometric luminosity is shown in the right panels. The red circles represent the observed data points. The green line shows the integrated unextinguished emission originating from stars. The blue solid line represents the star formation contribution for dust absorption, partially redistributed across the MIR/FIR range in a self-consistent way. The red line reproduces the AGN contribution and incorporates both the accretion disc and the torus emission, according to the adopted library of templates. The sources shown in the upper and middle panels are classified as AGN at  $\geq 99\%$  confidence, while one example of source with no significant AGN contribution is shown in the bottom panel.

fairly conservative and reliable in terms of AGN contribution to the observed SED. In Fig. 3.6, the upper and middle panels represent sources with a  $\geq 99\%$  significant AGN contribution, while that shown in the bottom panel has a no significant AGN contribution, hence classified as a “purely star-forming galaxy”. However, we stress that the average fraction of IR luminosity ascribed to an AGN component is around 10%, even for sources classified as “AGN” at  $>99\%$  CL. This implies that the total IR luminosity, as well as the SFR of the host galaxy, is not significantly affected by the presence of the AGN.

We have verified that the star-forming galaxy SEDs adopted in the code were physically motivated. For instance, we found that our best-fit galaxy SEDs well reproduce the



**Figure 3.7:** AGN detection rate in the *Herschel*-selected sample, as a function of the total IR (1–1000  $\mu\text{m}$ ) luminosity and redshift. In each redshift bin, the sample has been split into similarly populated luminosity bins. Error bars along the  $y$ -axis correspond to  $1\sigma$  confidence limits, while downward arrows show  $2\sigma$  upper limits. The dashed line is a linear fit to the data of the second redshift bin to highlight the evolution of the AGN detection rate with redshift.

“IR8” relation, presented by Elbaz et al. (2011), within  $\pm 1\sigma$  uncertainty. This finding also removes possible systematics related to the treatment of PAH emission component with respect to the total IR luminosity. We also checked the potential degeneracy between AGN and Starburst emission components when fitting the rest-frame mid-IR ( $3 < \lambda/\mu\text{m} < 8$ ) data points. Indeed, we note that the inter-stellar medium (ISM) emission in the MAGPHYS star-forming SEDs includes dust and PAH emission down to 3  $\mu\text{m}$  (see da Cunha et al. 2008). An excess above the pure stellar continuum will thus not be automatically ascribed to AGN. Besides, given the redshift distribution of our sources in Fig. 4.1, the rest mid-IR continuum is sampled at least by two filters in almost all observed SEDs (four IRAC bands, MIPS-24  $\mu\text{m}$ , and IRS-16  $\mu\text{m}$  in GOODS-S). Such a rich spectral coverage allows us to better constrain the rest mid-IR emission through the available templates.

### 3.5 AGN detection rate and evolution with $L_{1-1000}$

The fraction of sources fulfilling the  $F$ -test at  $>99\%$  confidence is as high as 37%. Our finding settles in an open debate concerning the fraction of AGN which are detected through independent selection methods. In the following we check the consistency between our results and the most recent estimates inferred from the analysis of different IR galaxy samples.

In the local Universe, Yuan et al. (2010) found the percentage of sources hosting an optically identified AGN to depend on the total infrared luminosity (8-1000  $\mu\text{m}$ )<sup>6</sup>. This fraction ranges from 20–30% up to  $\sim 60\%$  in local LIRGs ( $L_{\text{IR}} \sim 10^{11} L_{\odot}$ ) and ULIRGs, respectively. Consistent results have been obtained by Imanishi et al. (2010) and Alonso-Herrero et al. (2012) through *Akari*-IRC and *Spitzer*-IRS spectroscopy of local LIRGs and ULIRGs. These authors independently found that the AGN detection rate increases with increasing  $L_{\text{IR}}$ , from 25% (LIRGs) up to 70% (brightest ULIRGs), strengthening the previous results of Lutz et al. (1998) based on ISO data (from  $\sim 15\%$  at  $L_{\text{IR}} < 2 \times 10^{12} L_{\odot}$  up to 50% for brighter ULIRGs). The spectral decomposition has also been useful to effectively identify AGN: through detection of IR fine-structure lines, PAH equivalent widths and mid-IR continuum, several works (e.g. Sajina et al. 2009; Nardini et al. 2010; Sajina et al. 2012) claimed the presence of a significant AGN in 60–70% of local and distant ULIRGs, as well as evident signatures of coexisting star formation.

At higher redshifts, the situation is more controversial. Olsen et al. (2013) analysed a mass-selected sample at  $z \sim 2$  within the *Chandra Deep Field-South* (CDF-S) and ended up with a high fraction (43–65%) of star-forming galaxies likely hosting an AGN, according to the X-ray classification taken from Xue et al. (2011). Pozzi et al. (2012) studied a sample of 24 ULIRGs at  $z \sim 2$  (see Fadda et al. 2010) within the GOODS-South field and selected to have faint 24  $\mu\text{m}$  flux densities ( $0.14 < S_{24\mu\text{m}}/\text{mJy} < 0.55$ ). They ended up with a smaller fraction ( $\sim 35\%$ ) of ULIRGs featuring signatures of AGN activity and claimed that their IR luminosity emission is dominated by starburst processes.

We split our *Herschel* sample into six redshift bins ( $0.1 \leq z < 0.3$ ;  $0.3 \leq z < 0.7$ ;  $0.7 \leq z < 1.2$ ;  $1.2 \leq z < 1.8$ ;  $1.8 \leq z < 2.5$  and  $2.5 \leq z < 3.8$ ) and into different IR luminosities. We check for the evolution of the AGN detection rate with total (AGN+starburst) 1–1000  $\mu\text{m}$  luminosity  $L_{1-1000}$  at different redshifts, as shown in Fig. 3.7. Error bars along the  $y$ -axis show the  $\pm 1\sigma$  uncertainty (following Gehrels 1986 and Gerke et al. 2007), while the error bars along the  $x$ -axis set the bin width. Downward arrows represent  $2\sigma$  upper limits to the expected AGN detection rate when only “inactive” galaxies are detected in that bin. We find an overall increase of the AGN detection rate as a function of  $L_{1-1000}$ , except in the lowest redshift bin. At  $z < 0.3$  the AGN hosts are mainly “normal” ( $L_{1-1000} \leq 10^{11} L_{\odot}$ ) star-forming galaxies. At this redshifts, the average AGN detection rate is  $\sim 5\text{--}10\%$  and we do not find any clear trend with 1–1000  $\mu\text{m}$  luminosity. This is likely due to the fact that the comoving volume covered by COSMOS is not large enough to detect sources with

<sup>6</sup>For historical reasons, throughout the Thesis the expression “IR luminosity” (or  $L_{\text{IR}}$ ) will refer to the luminosity integrated over the spectral range 8–1000  $\mu\text{m}$ . If calculated in a different range (e.g. 1–1000  $\mu\text{m}$ ) it will be explicitied.

$$L_{1-1000} \geq 10^{12} L_{\odot}.$$

At higher redshift ( $0.3 < z < 3.8$ ), the AGN detection rate does increase with increasing  $L_{1-1000}$ , consistently with results from previous studies. This finding might suggest that the probability to find AGN in FIR-selected galaxies increases as a function of  $L_{1-1000}$ . Besides the trend with  $L_{1-1000}$ , we point out the evidence for redshift evolution (see Fig. 3.7). Indeed, the AGN detection rate of  $L_{1-1000} \sim 10^{11} L_{\odot}$  does increase from 5 to 30% up to  $z \sim 1$ ; galaxies with  $10^{11} \leq L_{1-1000}/L_{\odot} \leq 10^{12}$  host AGN activity with a probability rising from 25% at  $z \sim 0.5$  to 60% at  $z \sim 2$ . Brighter ( $L_{1-1000} \geq 10^{12} L_{\odot}$ ) IR galaxies show an AGN detection rate rising from 55% at  $z \sim 1$  to 70–80% at  $z \sim 2-3$ . We stress that the observed trends are not significantly affected by selection biases. While the observed increase of IR luminosity with redshift reflects our *Herschel*-based selection, we note that the fraction of 1–1000  $\mu\text{m}$  luminosity related to the AGN is generally negligible (around a few per cent), therefore poorly related to AGN activity. This check makes us reasonably confident that the selection at 160  $\mu\text{m}$  does not affect the overall increase of the AGN detection rate with redshift.

## 3.6 Comparison with AGN diagnostics

In this Section we test the robustness of the SED-fitting decomposition to estimate the AGN contribution in FIR-selected galaxies. We compare our findings with those obtained from independent AGN diagnostics taken from the literature.

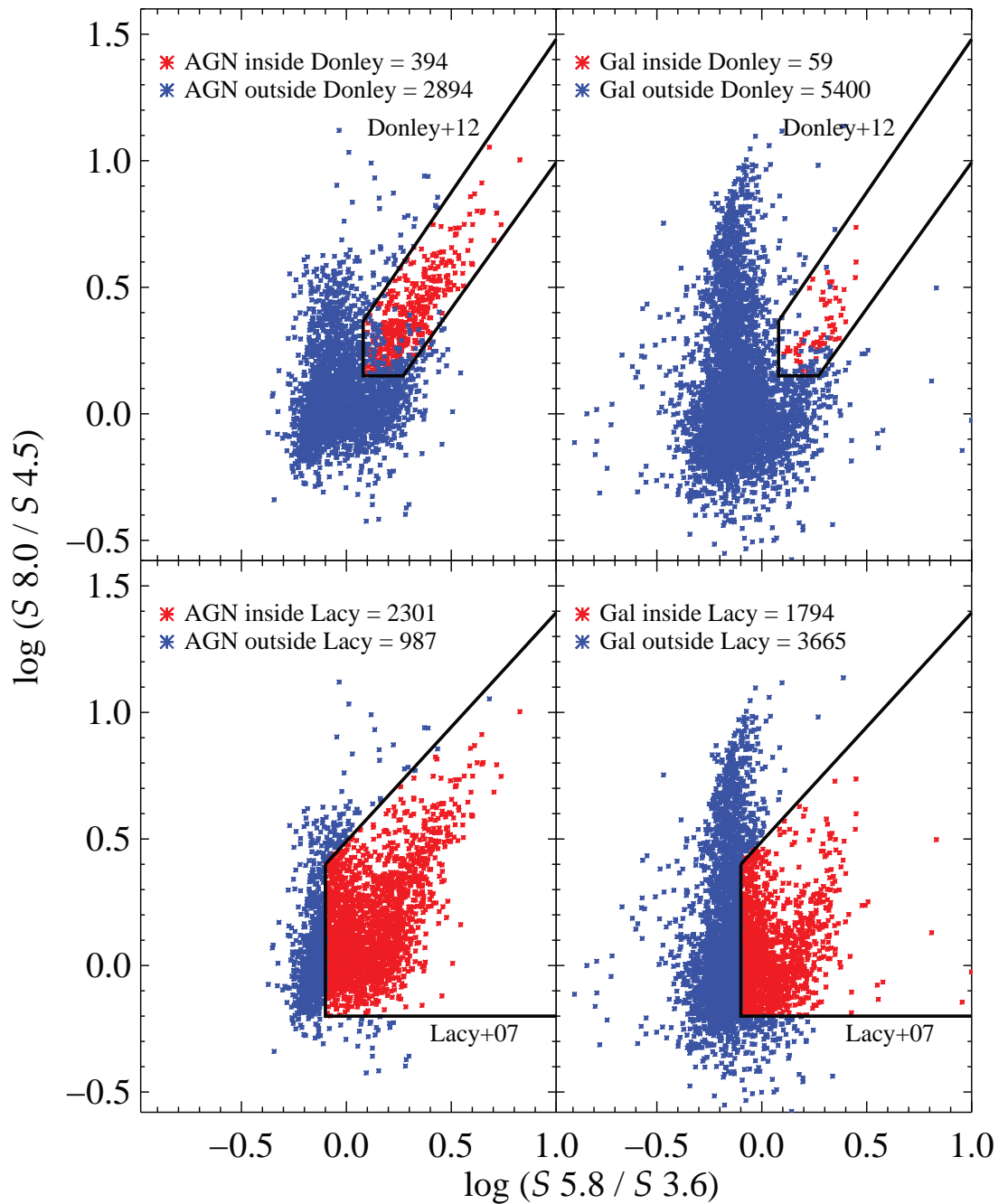
### 3.6.1 Comparison with mid-IR colour-colour AGN selection

The mid-IR colour-colour selection has been first introduced by Lacy et al. (2004, 2007) by making use of *Spitzer*-IRAC colours to isolate the AGN candidates, both obscured and unobscured. However, as already stated by Lacy et al. (2007), a sample of so-called ‘‘AGN candidates’’ might suffer from a significant contamination from ‘‘pure’’ galaxy SEDs with similar mid-IR colours. More recently, Donley et al. (2012) revisited the previous method to provide a more reliable AGN selection criterion, although it was less complete at low X-ray luminosities ( $L_{[0.5-8]} < 10^{44} \text{ erg s}^{-1}$ ). In Fig. 3.8 we compare our classification with that derived from these mid-IR colour-colour diagnostics. We report the comparison with Donley et al. (2012) in the upper panels and with Lacy et al. (2007) in the bottom ones. Left and right panels refer to our AGN and galaxy samples, respectively. Red asterisks mark the population of sources ‘‘inside’’ the corresponding wedge, whereas blue asterisks mark sources not satisfying the colour-colour criterion.

As shown in Fig. 3.8, some sources, although inside the black solid wedge of Donley et al. (2012), are labelled as ‘‘outside’’. This is why, in addition to the requirement of being inside the wedge, the IRAC colours of each source have to fulfill further conditions on IRAC fluxes at the same time to satisfy the Donley criteria. <sup>7</sup>

<sup>7</sup>Given the monochromatic fluxes ( $S_{\lambda}$ ) at 3.6, 4.5, 5.8 and 8.0  $\mu\text{m}$ , the criteria established by Donley et al. (2012) set the following additional conditions:  $S_{3.6} < S_{4.5} < S_{5.8} < S_{8.0}$  at the same time, which make the black solid wedge a multi-dimensional region.





**Figure 3.8:** Comparison with AGN colour-colour selection criteria. *Upper panels:* comparison with Donley et al. (2012). *Bottom panels:* comparison with Lacy et al. (2007). Left and right panels refer to our *Herschel* sources classified as AGN and galaxies, respectively. Sources satisfying the corresponding criteria are marked with red asterisks, otherwise with blue asterisks.

The colour-colour wedge defined by the Donley criteria (black solid line) includes  $\sim 12\%$  of our AGN sample and only  $\sim 1\%$  of our galaxy sample. According to our classification, we confirm that Donley et al. (2012) selection is very efficient in avoiding contamination from purely star-forming galaxies. The overlapping fraction of AGN obviously increases when considering the wider Lacy et al. (2007) region: we find that 70% of our AGN and

33% of our galaxies are inside the wedge.

We checked the average integrated properties of sources classified as AGN, falling either inside or outside the colour-colour AGN selection criteria. As expected, objects satisfying the Donley criterion are mostly AGN-dominated systems in the mid-IR (both obscured and unobscured in the optical) and populate the highest luminosity tail of different luminosity distributions. Indeed, typical values for the 1–1000  $\mu\text{m}$  AGN luminosity reach  $\sim 10^{12}L_{\odot}$ , as well as for the 1–1000  $\mu\text{m}$  luminosity due to star formation. Moving towards the area delimited by the Lacy et al. (2007) criteria and outside the Donley et al. (2012) selection, the typical properties become less and less extreme, with Seyfert-like SEDs (i.e. dominated by the host galaxy light in the optical/NIR domain) being more common than QSO-like ones.

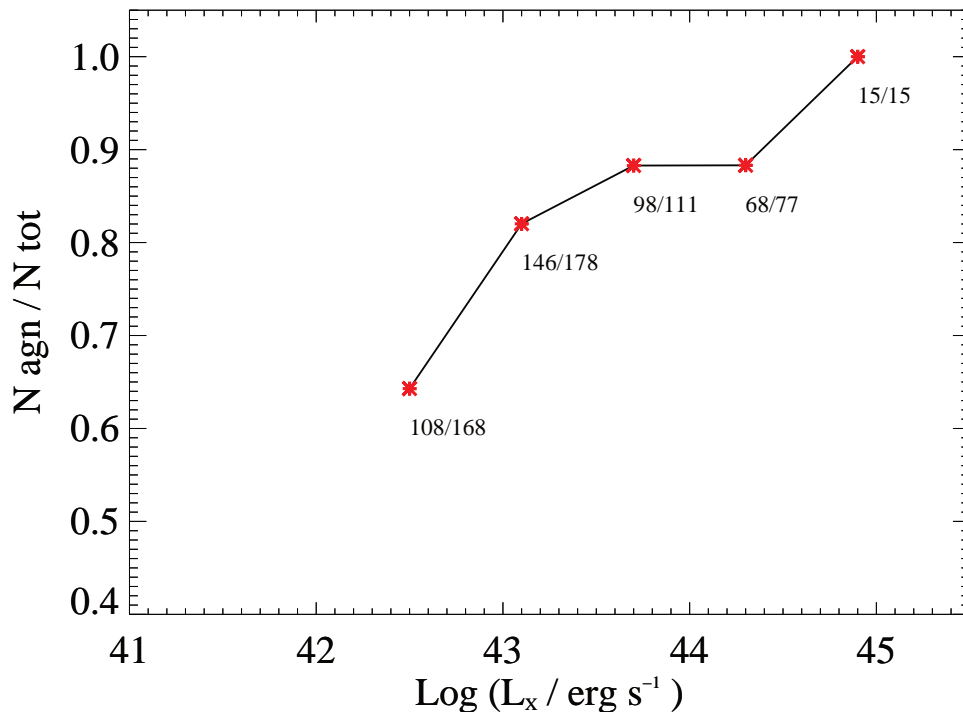
In summary, we find the colour-colour criterion by Lacy et al. (2007) to be in reasonably good agreement with our classification and to represent an acceptable compromise between reliability and completeness. In addition, the Lacy wedge effectively rules out sources with 1–1000  $\mu\text{m}$  SF luminosity  $L_{1-1000} < 10^{11}L_{\odot}$ , where the AGN contribution to the IR becomes negligible ( $\sim$  few %) and also the probability to pick up AGN is generally lower (see Section 3.5 and Fig. 3.7).

*Herschel* sources identified as AGN through SED decomposition lie in different regions in the mid-IR colour-colour diagram, compared to those classified as “inactive” galaxies. This finding, together with the fraction of our AGN sample fitting with the Donley and Lacy criteria discussed above, further supports the robustness of our classification.

### 3.6.2 Comparison with X-ray AGN samples

Another check to verify the reliability and completeness of the SED-fitting decomposition is to compare our IR classification with that inferred from X-ray observations. The fraction of *Chandra* sources that are detected by *Herschel* in the same fields reaches  $\approx 20\%$  for the full X-ray sample. We isolate X-ray AGN from sources with X-ray emission consistent with being “inactive” galaxies, on the basis of their observed X-ray luminosity  $L_X$  (in the rest-frame 0.5–8 keV). We apply a conservative threshold of  $L_X \geq 3 \times 10^{42} \text{ erg s}^{-1}$  to select “bona-fide” X-ray AGN (e.g. Xue et al. 2011)<sup>8</sup>. Fig. 3.9 illustrates the fraction of X-ray AGN classified as IR-AGN from SED-fitting, as a function of  $L_X$ . The trend increases monotonically, ranging from  $\approx 80\%$  at  $L_X \sim 10^{43} \text{ ergs}^{-1}$  to 100% at  $L_X \sim 10^{45} \text{ ergs}^{-1}$ . This is due to the fact that X-ray luminous ( $L_X > 10^{44} \text{ ergs}^{-1}$ ) AGN typically show up in the optical (for type-1 sources) and/or in the MIR domain, therefore their identification becomes more secure and less dependent on the host galaxy luminosity. This implies that the reliability of this technique is higher for more luminous AGN. More tricky is the identification of less luminous AGN, where the bolometric output of the SMBH may be significantly affected by the host galaxy light. Indeed, as shown in Fig. 3.9, about 1/3 of X-ray AGN at  $L_X \sim 3 \times 10^{42} \text{ erg s}^{-1}$  are classified as “inactive” galaxies. This argues that the reliability of the SED-fitting decomposition is a function of the AGN-to-galaxy

<sup>8</sup>This threshold can be safely applied to the vast majority of star forming galaxies, but is not suitable for powerful starbursts with  $\text{SFR} > 500 M_{\odot} \text{ yr}^{-1}$  (e.g. Mineo et al. 2014; Symeonidis et al. 2014), which correspond to  $\sim 1\%$  of our *Herschel* sample.



**Figure 3.9:** Fraction of X-ray selected AGN (detected also by *Herschel*) that are classified as AGN from SED-fitting decomposition, as a function of the 0.5–8 keV luminosity. Numbers are added for completeness.

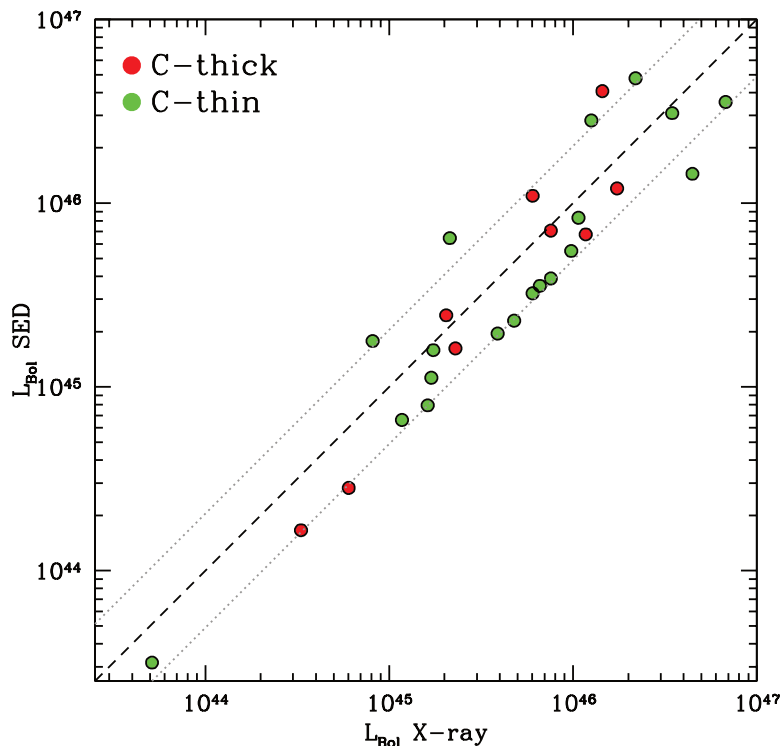
luminosity ratio. As a consequence, this technique does not provide with a *complete* sample of AGN (a more detailed analysis of the incompleteness of our IR-AGN sample will be given in Appendix A).

Nonetheless, SED-fitting decomposition can be very powerful in identifying heavily obscured AGN, even missed by deep X-ray surveys, and evaluating their bolometric luminosity  $L_{\text{bol,AGN}}$ . Here below we emphasize this important aspect by mentioning two different applications of our technique (see Sections 3.6.3 and 3.6.4 for details).

### 3.6.3 AGN luminosity in CT-AGN candidates

In a recent work, Lanzuisi et al. (2014) exploited X-ray data at 2–10 keV in the XMM-COSMOS survey (Hasinger et al. 2007, Cappelluti et al. 2009), which covers the entire 2 deg<sup>2</sup> of the COSMOS area. By performing a thorough X-ray spectral fitting analysis, they have identified a sample of 39 heavily obscured ( $N_{\text{H}} > 10^{23} \text{ cm}^{-2}$ ) AGN, including 10 CT-AGN candidates at  $0.1 < z < 2.5$ . For these 39 sources they derived reliable X-ray intrinsic (i.e. unobscured) luminosities, then converted into AGN bolometric luminosities after the application of X-ray bolometric correction  $k_{\text{bol,X}}$  (Lusso et al. 2012). This approach amplifies the uncertainties on the resulting AGN bolometric luminosity, since the obscuration-corrected X-ray luminosity is strongly dependent on the estimate of  $N_{\text{H}}$ , and typical  $k_{\text{bol,X}}$  values to apply for these sources range from  $\sim 10$  to  $\sim 100$ .

A valid approach is to estimate the AGN bolometric luminosity from SED-fitting decomposition, using the IR luminosity due to the AGN is an indirect probe (through ab-

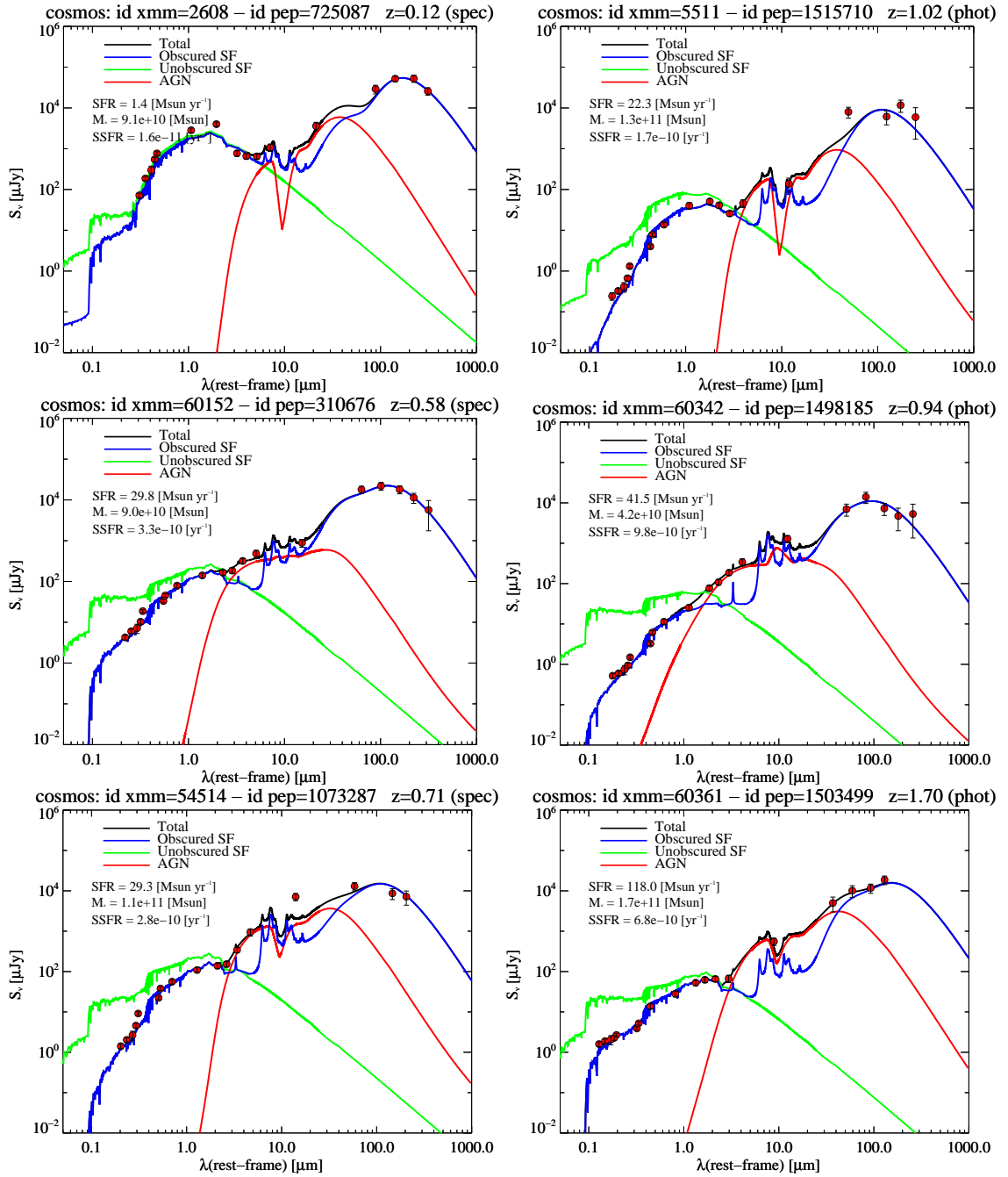


**Figure 3.10:** Comparison between X-ray based ( $x$ -axis) and SED-fitting based ( $y$ -axis) AGN bolometric luminosities, taken from Lanzuisi et al. (2014). Datapoints represent 19/29 C-thin (green) and 9/10 C-thick (red) candidates identified from X-ray spectral fitting. The dashed line is the bisector, while the dotted lines mark the  $1\sigma$  dispersion from the bisector, of the order of 0.3 dex.

sorption and isotropic re-emission) of the accretion disc optical/UV luminosity. In Fig. 3.10 we show the comparison between SED-based bolometric luminosities<sup>9</sup> ( $y$ -axis) and those inferred from X-ray analysis ( $x$ -axis), for 19 C-Thin ( $N_{\text{H}} < 10^{24} \text{ cm}^{-2}$ ) and 9 C-Thick ( $N_{\text{H}} > 10^{24} \text{ cm}^{-2}$ ) sources (the remaining sample did not show an AGN component in the best-fit). This comparison suggests that the two quantities are in good agreement ( $1\sigma$  scatter  $\approx 0.3$  dex), with no significant systematics. We stress that the contribution of the X-ray, obscuration-corrected luminosity to the  $L_{\text{bol,AGN}}$  inferred from SED fitting is very small (of the order of  $\approx 15\%$ ), and therefore these latter estimates are largely independent of the  $N_{\text{H}}$  measurements derived from the X-rays. We can interpret the good agreement found between the two quantities (within a factor of 2), as a strong confirmation of the reliability of the AGN bolometric luminosities inferred from SED-fitting decomposition.

Fig. 3.11 shows rest-frame SEDs of the 6 CT-AGN candidates by Lanzuisi et al. (2014) that are detected by *Herschel*. We note that in all cases the best-fit solution needs an AGN torus component, in order to reproduce the MIR photometric points. The best-fit AGN templates show in most cases a MIR SED dominated by the AGN emission component, with a  $9.7 \mu\text{m}$  silicate feature in absorption. Given the best-fit value of the  $\tau_{9.7}$ , the corresponding estimate of  $N_{\text{H}}$  predicted by a local dust-to-gas ratio (Shi et al. 2006) ranges

<sup>9</sup>Estimates for *Herschel* detected sources are taken from our analysis ( $\approx$  half of the datapoints shown in the plot), otherwise from Lusso et al. (2012).



**Figure 3.11:** SED-fitting decomposition 6 *Herschel*-detected out of 10 CT-AGN candidates identified by Lanzuisi et al. (2014). Coloured lines are as in Fig. 3.6. In each plot we report the estimates of SFR,  $M_*$  and sSFR inferred from the best-fit model.

from 23 to 23.8 in  $\log(N_{\text{H}})^{10}$ .

Importantly, we note that the fraction of CT-AGN candidates from Lanzuisi et al. (2014) that are detected by *Herschel* (6/10 = 60%) is much higher than the detected fraction for the full X-ray sample ( $\approx 20\%$ ) in the COSMOS field. We compared these

<sup>10</sup>We note that  $\log(N_{\text{H}})=23.8$  is the upper bound of the available  $\tau_{9.7}$  values (i.e.  $\tau_{9.7} = 6$ ), given our set of AGN templates, if assuming a local dust-to-gas ratio.

fractions also if applying a cut in (intrinsic) X-ray luminosity of  $L_X > 10^{44} \text{ erg}^{-1}$ , where the full X-ray sample is reasonably complete up to  $z \sim 2$ . We found that the difference remains noticeable (62% vs 19%). This might suggest that most X-ray selected CT-AGN are likely hosted in star forming galaxies, therefore that the amount of gas obscuring the AGN on circum-nuclear scales is perhaps connected to the global amount of dust in the host galaxy.

This study argues that the combined *Spitzer* and *Herschel* capabilities are paramount to disentangle the emission related to the AGN from the host galaxy light even for highly obscured, CT-AGN. Moreover, the large fraction of CT-AGN candidates observed by *Herschel* motivates the search for this elusive AGN population also from IR-selected galaxy samples.

### 3.6.4 Piercing through the dust: a candidate CT-QSO missed by deep X-ray observations

Given the small fraction (around 10%) of *Herschel*-selected galaxies that are detected in X-rays from *Chandra* observations, it is reasonable to ask whether a fraction of heavily obscured AGN that might be missed by X-ray observations could be detected by *Herschel*.

Starting from our *Herschel* sample, we searched for CT-AGN candidates in the COSMOS field<sup>11</sup>. Here below we list the criteria that we required for each source to enter our selection.

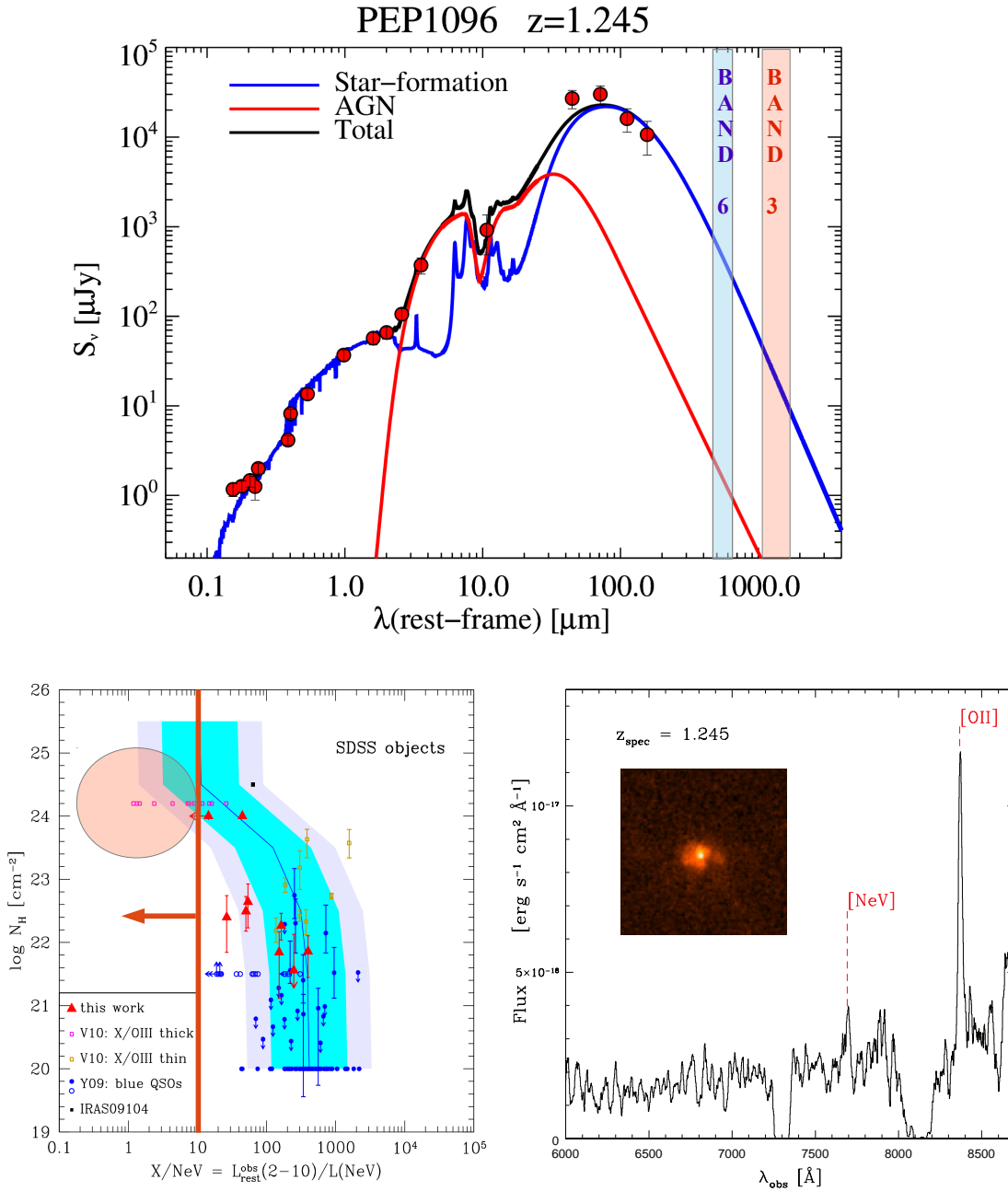
- to be classified as AGN (at  $\geq 99\%$  confidence level) from SED-fitting decomposition;
- to show a MIR (rest 3-30  $\mu\text{m}$ ) SED dominated by the AGN luminosity;
- undetected in the deepest 0.5  $\text{deg}^2$  covered by *Chandra*-COSMOS (C-COSMOS; Elvis et al. 2009, Civano et al. 2012) observations;
- the ratio  $L_{2-10, \text{intr}} / L_{2-10, \text{map}} > 50$  (Gilli et al. 2010).  $L_{2-10, \text{intr}}$  is the intrinsic 2–10 keV X-ray luminosity predicted by the AGN best-fit model and assuming a set of X-ray bolometric corrections from Marconi et al. (2004).  $L_{2-10, \text{map}}$  is the minimum X-ray luminosity detectable from the survey at a given redshift and flux limit.

Given these constraints, we ended up with 2 CT-QSO candidates<sup>12,13</sup>. Among them, we chose the only one with a secure spectroscopic redshift: “PEP1096” at  $z=1.245$  (Trump et al. 2007). Fig. 3.12 (upper panel) shows the decomposition of its broad-band SED, that seems to require a prominent AGN torus component in the best-fit. Interestingly, this object has been classified as AGN also from other diagnostics:

<sup>11</sup>We focused on the COSMOS fields to end up with candidates hosted in IR galaxies bright enough to be detected by ALMA at mm-wavelengths in a reasonable amount of observing time.

<sup>12</sup>The predicted  $L_{2-10, \text{intr}}$  for these sources is larger than  $10^{44.5} \text{ erg s}^{-1}$ , therefore suggesting their intrinsic bolometric power is similar to those of known QSOs.

<sup>13</sup>Our estimate of the surface density related to this specific subsample of CT-AGN is  $= 2 \text{ sources}/0.5 \text{ deg}^2$ . Despite the scarce statistics, this number is consistent with that predicted at the flux limit of the survey by current X-ray background (XRB) synthesis models (Gilli et al. 2007), that reaches  $4 \text{ sources}/\text{deg}^2$ .



1. Radio excess at 1.4 GHz with respect to radio emission expected from star-formation processes (Schinnerer et al. 2007; Del Moro et al. 2013).
2. It fits the “IR-AGN” wedge of the mid-IR colour-colour selection criterion defined by Donley et al. (2012).
3. The low X-ray/[Ne V] flux ratio. The [Ne V]3426 emission line has been identified from the optical spectrum (Fig. 3.12, bottom right panel). We estimated the X-ray flux limit for the source at  $3\sigma$  level. We ended up with a flux ratio  $X/[\text{NeV}] < 10$ . The region of the  $X/[\text{NeV}]$  vs  $N_{\text{H}}$  diagram covered by robust CT-AGN candidates (Fig. 3.12, bottom left panel) is defined by the red line and arrow ( $X/[\text{NeV}] < 15$ , Gilli et al. 2010). The low X-ray/[Ne V] strongly argues the presence of a high degree of nuclear obscuration.

These independent diagnostics suggest that our target is a robust CT-AGN candidate. Surprisingly, this object has not been detected in deep Chandra-COSMOS X-ray observations (with  $\sim 140$  ks exposure time at the source position). For this peculiar target, we will propose ALMA high-resolution ( $\approx 0.01''$ ) observations of both  $^{12}\text{CO}(2-1)$  and  $^{12}\text{CO}(5-4)$  line emission, falling in band 3 and 6, respectively (Fig. 3.12, upper panel). This will allow us to achieve the following objectives.

- (1) Measure the total molecular gas mass ( $M_{\text{gas}}$ ) from the  $^{12}\text{CO}(2-1)$  line.
- (2) Spatially resolve the molecular gas surface density traced by the  $^{12}\text{CO}(5-4)$  emission. This would allow us to constrain the degree of obscuration due to circum-nuclear dust and gas (on  $\approx 100$  pc scales), therefore to understand whether this is consistent with that required to explain the heavily obscured nature of this target, inferred from panchromatic diagnostics.



# Chapter 4

## Black Hole accretion history from *Herschel*

The joint *Herschel* and *Spitzer* capabilities of constraining the MIR-to-FIR observed SEDs are crucial to isolate the possible AGN emission component in galaxy SEDs. Our SED-fitting decomposition allows us to get an estimate of the AGN bolometric luminosity  $L_{\text{bol,AGN}}$ , as described in Section 3.3, for each source identified as AGN on the basis on the F-test (Section 3.4). In the following we make use of this approach to determine the evolution of the AGN bolometric luminosity function and to constrain the cosmic history of SMBH accretion since  $z \sim 3$ .

This Chapter is structured as follows. After a brief introduction (Section 4.1), we introduce the sample of *Herschel* galaxies used in this analysis in Section 4.2. In Section 4.3 we derive the evolution of the AGN bolometric luminosity function up to  $z \sim 3$ . This has been used to determine the SMBH accretion history (Section 4.4) over cosmic time, discussing and comparing our results with previous derivations from the literature.

### 4.1 Introduction

Recent attempts have been made to trace the SMBH growth over cosmic time (e.g. Marconi et al. 2004; Shankar et al. 2004; Hasinger et al. 2005; Hopkins et al. 2006; Hopkins et al. 2007; Merloni & Heinz 2008b; Aird et al. 2010; Fanidakis et al. 2012; Mullaney et al. 2012a), supporting the connection between black hole accretion density (BHAD) and SFD up to  $z \simeq 1-1.5$ . In particular, Mullaney et al. (2012a) found that the ratio SFD/BHAD remains constant at a value of  $\sim 1-2 \times 10^3$  over the redshift range  $0.5 < z < 2.5$ . In agreement with observations, semi-analytical models and hydro-dynamical simulations currently hint at an overall decrease of the BHAD from  $z \simeq 2$  down to the present day. At higher redshift the situation is more uncertain, since direct diagnostics for black hole mass estimates (e.g. reverberation, time variability, etc.) are no longer available, and indirect measurements need to be extrapolated up to these redshifts.

Beyond  $z \simeq 1-2$ , the AGN obscured growth is increasingly uncertain, since even the deepest X-ray surveys suffer from a decreasing completeness with increasing redshift and/or increasing obscuration (Brandt & Hasinger 2005b). While the dependence of the hydrogen column density  $N_{\text{H}}$  on the intrinsic AGN luminosity is currently well established (Lawrence

1991; Simpson 2005; Ueda et al. 2003; Steffen et al. 2003; Barger & Cowie 2005; Sazonov et al. 2007; Lusso et al. 2013), the evolution with cosmic time is still debated. On the one hand, a slight evolution with redshift has been suggested by some analyses (La Franca et al. 2005; Treister & Urry 2006; Ballantyne et al. 2006; Hasinger 2008; Della Ceca et al. 2008; Treister et al. 2009; Iwasawa et al. 2012); on the other hand, the space density of the obscured ( $\log[N_{\text{H}}/\text{cm}^{-2}] > 22$ ) AGN population seems to be consistent with a no evolution scenario, at least for moderately luminous ( $L_{\text{X}} < 10^{44} \text{ erg s}^{-1}$ ) X-ray AGN (Ueda et al. 2003; Akylas et al. 2006; Gilli et al. 2010; Vito et al. 2013).

AGN are ubiquitous X-ray emitters (Elvis et al. 1978) and X-ray surveys are only weakly affected by the X-ray host galaxy light. Nevertheless, recent works (Treister et al. 2012; Schawinski et al. 2012; Laird et al. 2010) claim that the host galaxies of high-redshift ( $z > 1-2$ ) AGN could have experienced intense bursts of star formation (e.g. SMGs and/or ULIRGs), likely hiding heavily obscured AGN with intrinsically high bolometric luminosities, but enshrouded in dense, cold, optically thick gas clouds (see Alexander et al. 2005). IR observations are only mildly affected by obscuration and can detect warm dust signatures for both unobscured and obscured AGN. This is the reason why IR observations offer a complementary perspective with respect to studies based on X-ray observations to constrain the obscured AGN growth across cosmic time. The main purpose of this work is to trace the evolution of the BHAD from an IR survey through a multi-wavelength analysis of galaxy SEDs.

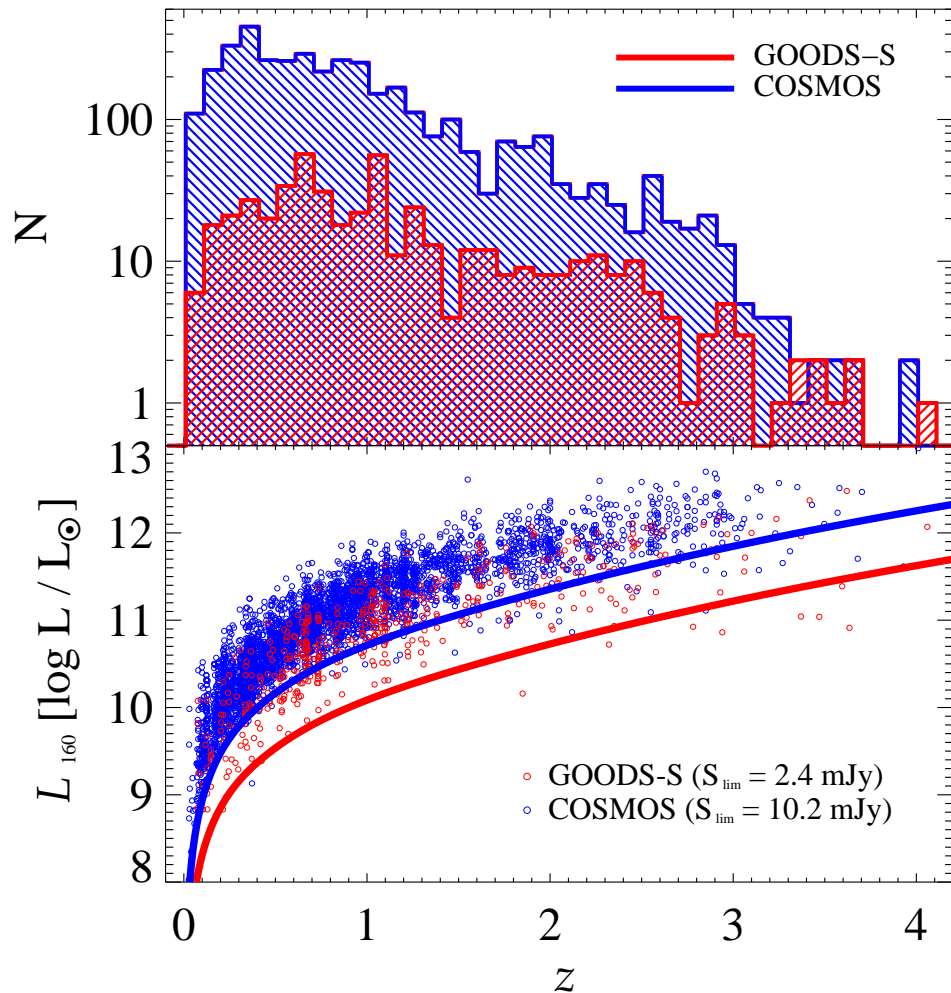
## 4.2 Sample selection and multi-wavelength catalogues

We exploit observations carried out within the PEP (Lutz et al. 2011) and the HerMES (Oliver et al. 2012) surveys in the GOODS-S and COSMOS fields. In the FIR domain, we take advantage of PACS photometry (100 and 160  $\mu\text{m}$ ; 70  $\mu\text{m}$  data are available for the GOODS-S field only). As a reference sample we use the PEP 160  $\mu\text{m}$ -selected blind catalogues (internal version 1.2 in GOODS-S and version 2.2 in COSMOS)<sup>1</sup>. The PEP sample comprises all sources with  $>3\sigma$  detection at 160  $\mu\text{m}$ , corresponding to nominal flux density limits of 2.4 and 10.2 mJy in the GOODS-S (782 objects) and COSMOS (5105 sources) fields, respectively.

The cross-match between PEP-selected sources and the multi-wavelength ancillary data has been executed using a likelihood ratio technique (Sutherland & Saunders 1992; Ciliegi et al. 2001), as done in Section 3.1.1 from PACS wavelengths down to MIPS-24  $\mu\text{m}$  detections. We exploited SPIRE data at 250, 350 and 500  $\mu\text{m}$ , if available. In both fields, the cross-match between SPIRE and 24  $\mu\text{m}$  is based MIPS-24  $\mu\text{m}$  positional priors, according to the method described in Roseboom et al. (2012).

In GOODS-S we have matched the PACS area in common with the inner region covered by multi-wavelength observations in UV, optical and near-IR with the GOODS-MUSIC cat-

<sup>1</sup>This catalogues were released earlier with respect to those described in Table 2.1 and in Section 3.1.1. At the time when this work has been accomplished, these were the most updated PEP catalogues available in the GOODS-S and in COSMOS. They reach the same flux density limits listed in Table 2.1 (without accounting for GOODS-H observations). See the text for details.



**Figure 4.1:** *Top panel:* redshift distribution of the  $160\ \mu\text{m}$  PEP sources identified in optical bands, in GOODS-S (*red dashed*) and COSMOS (*blue dashed*). Note that the scale of the  $y$ -axis is logarithmic. *Bottom panel:* infrared luminosity limit at  $160\ \mu\text{m}$  (rest-frame) as a function of redshift. Open circles represent the luminosity limits for GOODS-S (*red*) and COSMOS (*blue*) sources. The solid curves show the minimum luminosity (corresponding to the limiting flux) versus redshift for the two fields (GOODS-S in red, COSMOS in blue), obtained for the empirical SED of the source IRAS 20551–4250.

alogue (Grazian et al. 2006; Santini et al. 2009). The GOODS-MUSIC area ( $196\ \text{arcmin}^2$ ) includes 494 PACS- $160\ \mu\text{m}$  sources with optical/NIR identification.

In the COSMOS field, the common sky area covered by *Herschel* and ancillary observations at shorter wavelengths is around  $2\ \text{deg}^2$ . PACS- $160\ \mu\text{m}$  sources have been associated with the *Spitzer*-MIPS  $24\ \mu\text{m}$  catalogue (Le Floc'h et al. 2009). MIPS- $24\ \mu\text{m}$  positions have been cross-matched to the IRAC-selected catalogue by Ilbert et al. (2010), using a nearest neighbour method, and subsequently cross-matched to UV, optical and near-IR photometry taken from Capak et al. (2007) and Ilbert et al. (2009). By removing PEP sources within flagged areas of the optical-to-NIR catalogues, we are left with 4118 sources at  $160\ \mu\text{m}$  identified at shorter wavelengths. Around 96.5% of them have a MIPS  $24\ \mu\text{m}$  detection in the common area.

The fraction of PACS-selected sources having at least one SPIRE counterpart ( $> 3\sigma$ ) reaches 84% in the GOODS-S field (within the MUSIC area) and 86% in the COSMOS field. Overall, we collect photometry at a maximum of 21 (GOODS-S) and 19 (COSMOS) different wavebands, from UV to FIR wavelengths. The numbers of PEP-selected objects fulfilling the optical identification criterion are 494 and 4118, in the GOODS-S and COSMOS, respectively.

### Redshift distribution

In summary, the GOODS-S PEP sample includes 494 sources with 100% redshift completeness (71% spectroscopic and 29% photometric). The overall redshift completeness reached in COSMOS is as high as 93% (about 40% spectroscopic), corresponding to 3849 PEP sources.

The final sample of GOODS-S and COSMOS sources selected at  $160 \mu\text{m}$  and with a redshift estimate includes 4343 objects (494 for GOODS-S and 3849 for COSMOS). The global redshift distributions of our PACS- $160 \mu\text{m}$  selected sample is shown in Fig. 4.1 (top) for both the PEP fields. As expected, because of the higher flux limit, the COSMOS redshift distribution has a lower median redshift value ( $\langle z \rangle \sim 0.7$ ) than the GOODS-S ( $\langle z \rangle \sim 1$ ), although it also extends to high redshift ( $z > 3$ ). To highlight the range of infrared luminosities sampled at different redshifts, we show in Fig. 4.1 (bottom) the minimum  $160 \mu\text{m}$  luminosity versus redshift of the (rest-frame)  $160 \mu\text{m}$  luminosity. Red and blue circles mark GOODS-S and COSMOS sources, respectively. The solid curves represent the minimum luminosity (corresponding to the limiting flux) versus redshift for the two fields (GOODS-S in red, COSMOS in blue). This curve is purely illustrative and it is taken considering the SED template of *IRAS* 20551–4250, which is a local star-forming galaxy hosting an AGN (see Polletta et al. 2007).

## 4.3 AGN Bolometric Luminosity Function

In this section we investigate the evolution of the AGN bolometric luminosity function (LF) from the whole joint GOODS-S and COSMOS sample. Globally we collect 4343 objects, but just the fraction showing a substantial AGN contribution has been considered for the following analysis (37% at 99% confidence level). The sample of interest has been split into six different redshift bins (as done in Fig. 3.7) and 10 luminosity bins, from  $10^8$  up to  $10^{15} L_{\odot}$ , with a logarithmic bin-width equal to 0.7.

### 4.3.1 Method

The computation of the LF relies on the non-parametric  $1/V_{\text{max}}$  method introduced by Schmidt (1968), where  $V_{\text{max}}$  represents the maximum comoving volume that allows a source of a given luminosity to be observable above the flux limit of the survey. Since we consider sources coming from two different samples (GOODS-S and COSMOS), we use the method of Avni & Bahcall (1980), who extended the previous  $1/V_{\text{max}}$  method to coherently join and analyse simultaneously different samples. For each source of our sample we compute

the  $V_{\max}$  as follows:

$$V_{\max} = \int_{z_{\min}}^{z_{\max}} \frac{dV}{dz} \Omega(z) dz \quad (4.1)$$

with  $z_{\min}$  being the lower boundary of a given redshift bin and  $z_{\max}$  the minimum between the upper boundary of the redshift bin and the maximum redshift value available for the source of interest to be detected. Here  $\Omega(z)$  represents the effective area of visibility and is computed as the fraction of the geometrical projected sky area  $\Omega_{\text{geom}}$ , where the sensitivity of the instrument is enough to let the source be observable, according to the formula:

$$\Omega(z) = \Omega_{\text{geom}} \cdot f_c(z) , \quad (4.2)$$

where  $f_c(z)$  is the flux completeness correction obtained from simulations by Berta et al. (2010, 2011). For a given luminosity and redshift bin, we computed the luminosity function as follows:

$$\Phi(L, z) = \frac{1}{\Delta \log L} \sum_{i=1}^n \frac{1}{V_{\max,i}} \quad (4.3)$$

where  $n$  is the number of sources falling in the same luminosity bin, while  $\Delta \log L$  is equal to 0.7, as mentioned before. The COSMOS field has a 93% redshift completeness, regardless of PACS flux densities, so that we multiply  $\Phi(L, z)$  in COSMOS by a constant factor of 1.07, whereas the GOODS-S field is 100% complete in redshift and therefore no correction is needed. Moreover, we account for the fraction of the COSMOS sample without a MIPS-24  $\mu\text{m}$  counterpart inside the MIPS area. Indeed, in the context of a multi-component SED-fitting, the lack of the 24  $\mu\text{m}$  detection prevents the possible torus component from being constrained, resulting in low reliability best-fits. This corresponds to  $\sim 3.5\%$  of the PEP sample, roughly independent of redshift and AGN luminosity. We take into account their contribution by uniformly increasing  $\Phi(L, z)$  by a factor of 3.5%.

### 4.3.2 “Observed”AGN bolometric LF

In Fig. 4.2 we show our accretion LF in six different redshift bins. Error bars correspond to  $\pm 1\sigma$  Poissonian uncertainties (Marshall 1985), while in case of one single object in a given luminosity bin, we provide a 90% uncertainty following the recipe of Gehrels 1986. We computed separate LFs for GOODS-S and COSMOS (red and blue areas, respectively), and the total LF for the joint (GOODS-S + COSMOS) sample (black circles). As expected, in all redshift bins the GOODS-S field extends down to lower luminosities, whereas COSMOS typically samples higher luminosities.

The AGN bolometric LF shown here has been already corrected for incompleteness in accretion luminosity. We briefly describe our method as follows and refer the reader to the Appendix A for a more detailed explanation.

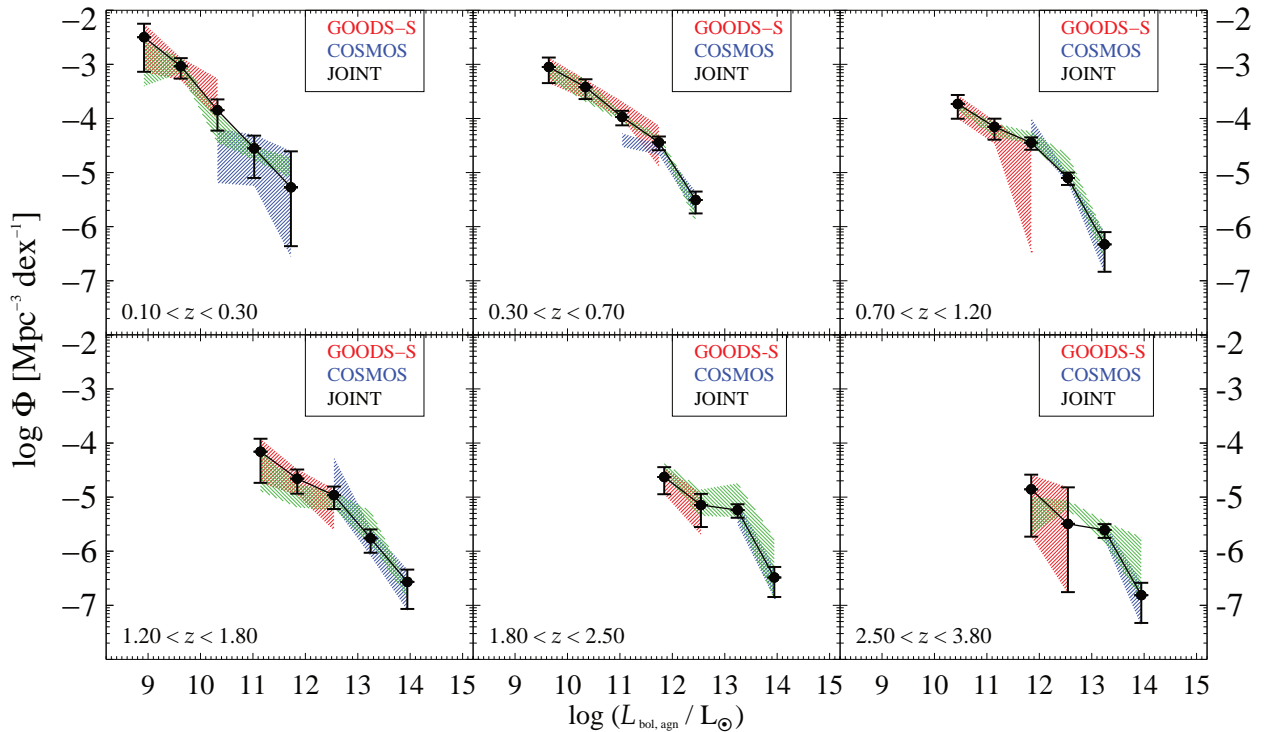
Our approach to account for the incompleteness in accretion luminosity follows that presented by Fontana et al. (2004), who built the stellar mass function correcting for the incompleteness in stellar mass, starting from a  $K$ -band selected sample. The same argument could be in principle generalised, under some assumptions, to deal with the evaluation of

the incompleteness affecting some unobservable quantity when selecting sources with any other observable quantity.

Focusing on our analysis, since a not negligible fraction of active galaxies might be missed by our FIR-based selection, it is necessary to quantify such missed AGN population. Starting from a FIR-selected sample, the incompleteness in accretion luminosity might be evaluated by looking at the distribution of the ratio between accretion flux  $S_{\text{accr}}$  (defined as the bolometric flux corresponding to a given  $L_{\text{bol,AGN}}$  and redshift) and  $S_{160}$ , traced by our AGN population. By assuming that this distribution does not change either with redshift or with *Herschel* flux density, it is possible to shift that down in  $S_{\text{accr}}$  and to virtually sample relatively high values of  $L_{\text{bol,AGN}}$  in weakly star-forming galaxies, not observable by *Herschel*. We iterate this shift down in  $S_{\text{accr}}$  as long as the expected number of missed AGN is equal to the number of observed AGN. The latter step identifies a threshold in accretion flux  $S_{\text{accr,lim}}$ , corresponding to a correction for incompleteness by a factor of two. The curve which parametrises the incompleteness is used to compute the maximum comoving volume where each source with  $S_{\text{accr}} \geq S_{\text{accr,lim}}$ , either detected or undetected by *Herschel*, is expected to be placed. We carried out this analysis separately for GOODS-S and COSMOS fields. As mentioned before, the AGN bolometric LF shown in Fig. 4.2 already incorporates the correction for incompleteness. This means that our LF is supposed to account for all active galaxies with  $S_{\text{accr}} \geq S_{\text{accr,lim}}$ , either above or below the *Herschel* detection limit. This allows us to remove the observational bias due to our FIR selection.

In the luminosity bins which are populated by both GOODS-S and COSMOS AGN at the same time, the two accretion luminosity functions are in reasonably good agreement one with the other. In the first redshift bin we find a large gap ( $\sim 1$  dex) in the connection point between the two surveys, but the very low statistics (only two objects per field) makes the two LF broadly consistent within  $2\sigma$ .

A Monte Carlo (MC) simulation has been performed to account for all the uncertainties in  $L_{\text{bol,AGN}}$  associated to the objects in the same redshift and luminosity bin. For each object belonging to a given  $(L_{\text{bol,AGN}}, z)$  bin, we randomly extract an individual  $L_{\text{bol,AGN}}$  value from the whole PDF distribution (obtained through SED-fitting decomposition, see Section 3.3), using the intrinsic shape of the PDF as weight. Then, using these random values of  $L_{\text{bol,AGN}}$ , we compute the LF for this simulated sample. By iterating the same MC simulation 100 times, we characterise a range of  $\Phi_{\text{MC}}(L, z)$  defining the uncertainty region associated to each LF datapoint  $\Phi(L, z)$ . In Fig. 4.2 we show our accretion LF (black circles, with associated Poissonian error bars) with its related  $\pm 1\sigma$  confidence region (green dashed areas), derived from MC simulation. The size of the uncertainty region is typically comparable to the Poissonian error bars, except for the less populated bins (large Poissonian error bars). The limits drawn by MC simulations have been used to provide an uncertainty range for the integrated LF (see Section 4.4).



**Figure 4.2:** Accretion luminosity function (LF) in different redshift bins for the PEP sample. Red and blue shaded areas trace the  $\pm 1\sigma$  Poissonian uncertainties separately for GOODS-S and COSMOS observed data points, respectively. The joint (GOODS-S and COSMOS) LF is represented by black circles, with  $\pm 1\sigma$  uncertainties (black vertical bars) related to the joint sample. The green dashed areas show the  $\pm 1\sigma$  range of  $\Phi_{MC}(L, z)$  obtained through Monte Carlo simulations (see text for details).

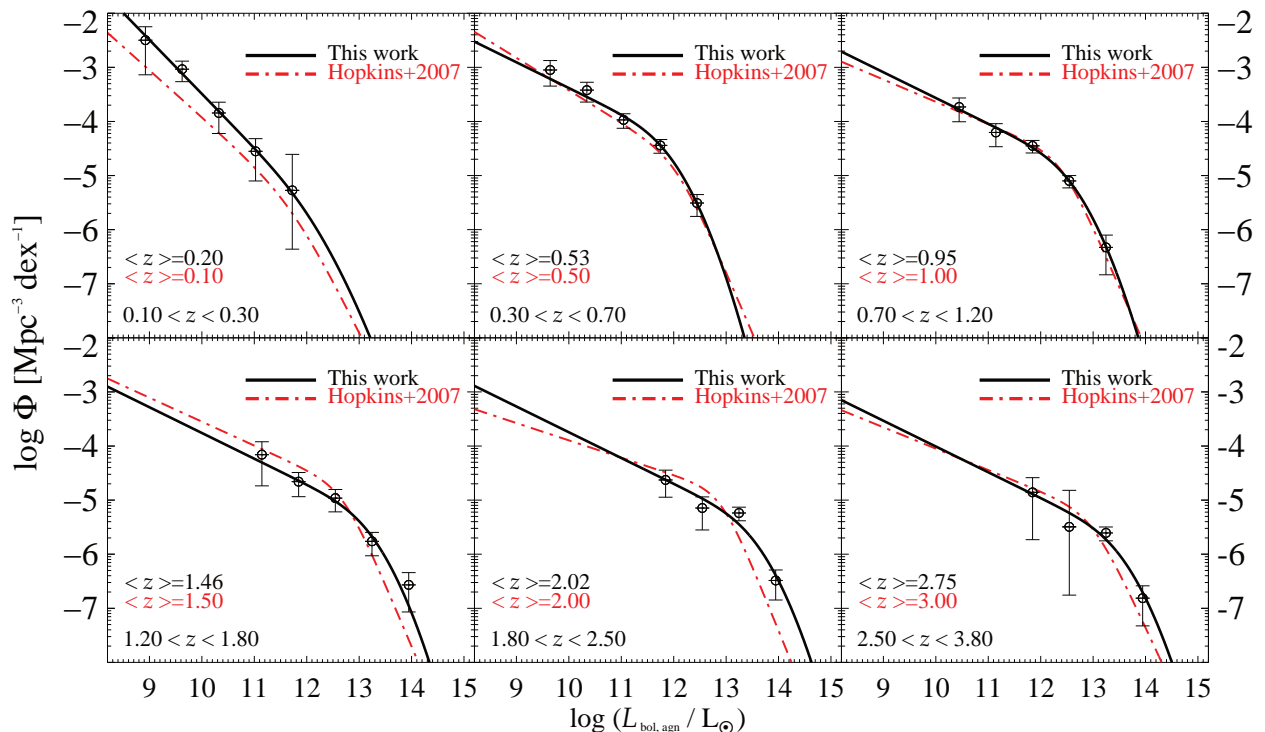
### 4.3.3 Evolution of the best-fit AGN bolometric LF

We fit our observed AGN bolometric LF making use of a modified version of the Schechter function (Saunders et al. 1990), as follows:

$$\Phi(L)d \log L = \Phi^* \left( \frac{L}{L^*} \right)^{1-\alpha} \exp \left[ -\frac{1}{2\sigma^2} \log_{10}^2 \left( 1 + \frac{L}{L^*} \right) \right] d \log L \quad (4.4)$$

which behaves as a power law for  $L \ll L^*$  and as a Gaussian in  $\log L$  for  $L \gg L^*$ . The couple of parameters  $(\Phi^*, L^*)$  represent the normalization and luminosity of the knee of the distribution, respectively. Two other parameters  $(\alpha, \sigma)$  are set to shape the low and high luminosity tails of the best-fit function.

We make use of a non-linear least square fitting routine to find the set of parameters that best reproduce the observed  $\Phi(L, z)$  through a modified Schechter function. In Fig. 4.3 the best-fit functions have been overplotted to the data. The black line represents our best-fit, while the red dot-dashed line is the Quasar Bolometric LF as derived by Hopkins et al. (2007, hereafter H07). They used a wide compilation of observed AGN luminosity functions, from the mid-IR through optical, soft and hard X-ray data to build up a self-consistent AGN bolometric luminosity function. H07 adopted a distribution of luminosity



**Figure 4.3:** Accretion LF (open circles with  $1\sigma$  Poissonian uncertainties) derived from PEP data. The black solid line represents our best-fit LF, whereas the red dash-dotted line is the best-fit curve of the AGN bolometric LF, as derived by Hopkins et al. (2007, hereafter H07). In each redshift bin we report the median redshift value of the AGN population, both for this work and for H07.

dependent bolometric corrections for each selection wave band, as well as a large baseline of intrinsic (i.e. corrected for obscuration) AGN SED shapes (from Richards et al. 2006; Steffen et al. 2006). However, the intrinsic SEDs of Richards et al. 2006 also include an additional IR ( $\lambda > 1 \mu\text{m}$ ) contribution arising from the torus reprocessed radiation. This IR excess (with respect to a blackbody emission in the Rayleigh-Jeans regime) has to be removed to avoid double counting the input energy of the accretion disc. As already done by Merloni & Heinz (2013), we independently evaluated such an excess to be of the order of 33%. As a consequence, all the AGN bolometric luminosities taken from H07 have been reduced by a factor  $\sim 0.33$  ( $\sim 0.18$  dex). We split our set of data points in order to sample similar mean redshifts to H07.

There are four functional parameters controlling the Schechter function, but not more than five luminosity bins are sampled per redshift bin (i.e. there is either 0 or 1 degree of freedom). As a consequence, not all the best-fit parameters could be sufficiently constrained at the same time. For this reason, we choose to relax the four functional parameters in the third redshift bin only, where both the faint and the bright-end of the LF are sampled. In the other bins, we fix  $\alpha$  and  $\sigma$  at the values obtained in the third redshift bin, except at  $z < 0.3$ , where the  $\alpha$  value was allowed to vary, since we observed a moderately steeper slope ( $\alpha = 1.97$ ) compared to the  $\alpha$  value adopted in the other bins ( $\alpha = 1.48$ ).



$z$ -bin	$\alpha$	$\sigma$	$\log_{10}(L^*/L_{\odot})$	$\log_{10} \Phi^*$ ( Mpc $^{-3}$ dex $^{-1}$ )
$0.1 \leq z < 0.3$	$1.97 \pm 0.25$	$0.54 \pm 0.31$	$10.91 \pm 1.96$	$-4.35 \pm 0.93$
$0.3 \leq z < 0.7$	$1.48 \pm 0.22$	$0.54 \pm 0.31$	$11.34 \pm 0.14$	$-4.03 \pm 0.31$
$0.7 \leq z < 1.2$	$1.48 \pm 0.22$	$0.54 \pm 0.31$	$11.96 \pm 0.15$	$-4.47 \pm 0.31$
$1.2 \leq z < 1.8$	$1.48 \pm 0.22$	$0.54 \pm 0.31$	$12.61 \pm 0.24$	$-5.00 \pm 0.55$
$1.8 \leq z < 2.5$	$1.48 \pm 0.22$	$0.54 \pm 0.31$	$12.96 \pm 0.18$	$-5.15 \pm 0.44$
$2.5 \leq z \leq 3.8$	$1.48 \pm 0.22$	$0.54 \pm 0.31$	$12.92 \pm 0.22$	$-5.40 \pm 0.61$

**Table 4.1:** List of best-fit parameters of the AGN bolometric LF with related  $\pm 1\sigma$  uncertainties.

This configuration implies that the shape of the Schechter function is defined in the first and third redshift bins and is allowed to scale rigidly along the (x,y) axes to fit the observed LF in the other redshift bins.

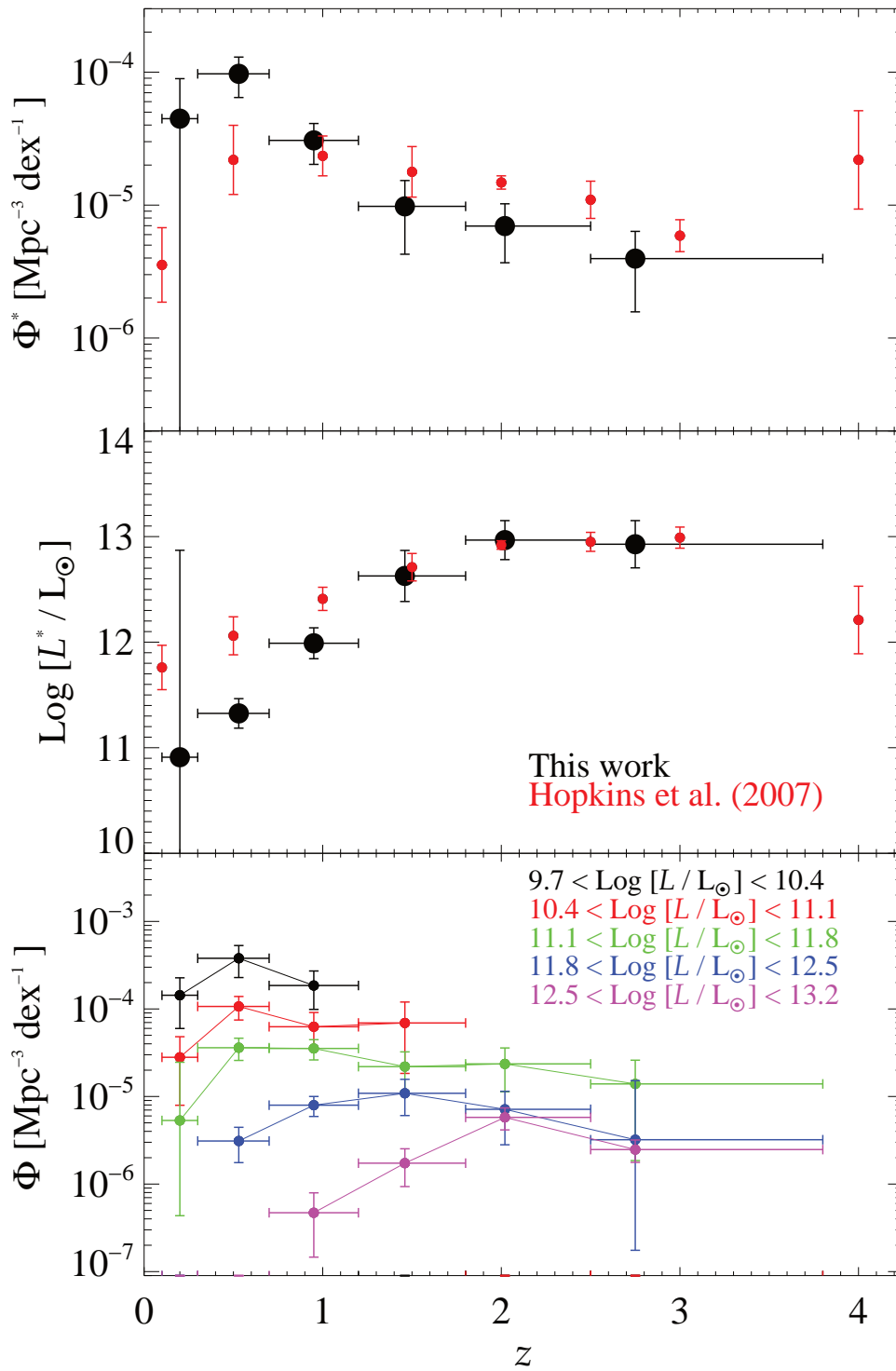
In Table 4.1 we list all the best-fit parameters achieved through the least squares fitting, together with their related  $\pm 1\sigma$  errors. Since only one object populates the highest luminosity bin of the first redshift slice, the resulting best-fit does not allow us to constrain either  $\Phi^*$  or  $L^*$  (see Table 4.1).

Despite the uncertainties characterising the first  $z$ -bin, there is evidence in favour of an evolution of the parameters ( $\Phi^*$ ,  $L^*$ ) as a function of redshift. As show in Fig. 4.4 (top panel), the normalization value has a maximum around  $z \sim 0.5$ , then it decreases to  $z \sim 3$ . We compare our trends with those obtained by H07, finding a similar evolution of  $\Phi^*$  with redshift. In the middle panel, the knee of the modified Schechter LF increases with redshift, at least up to  $z \sim 2$ . Comparing the LF of our work with Table 2 of H07<sup>2</sup>, the knee of the LF is systematically at lower luminosities that what was inferred by H07 in the respective redshift bins, except at  $z > 1.5$ , where our  $L^*$  estimates are consistent with H07. This is likely due to the fact that we are biased against early-type (and/or weakly star-forming) galaxies, which are known to host the most luminous AGN in the local Universe. Letting the number density of early-type galaxies decrease towards higher redshift, one might also expect that fewer and fewer passive galaxies host an active SMBH, as predicted by the major merger evolutionary scenario (e.g. Sanders et al. 1988). This would imply that the bulk of the SMBH growth would be progressively dominated by late-type galaxies, so that at higher redshift our  $L^*$  values become consistent with those found by H07. In Fig. 4.4 (bottom panel) we show the evolution with redshift of the AGN number density, for different luminosity bins. As already pointed out from several works (e.g. Hasinger et al. 2005; Silverman et al. 2008) based on X-ray data, we confirm an anti-hierarchical growth for SMBHs harboured in star-forming galaxies selected by *Herschel*, with the most luminous ones evolving faster and peaking at higher redshift.

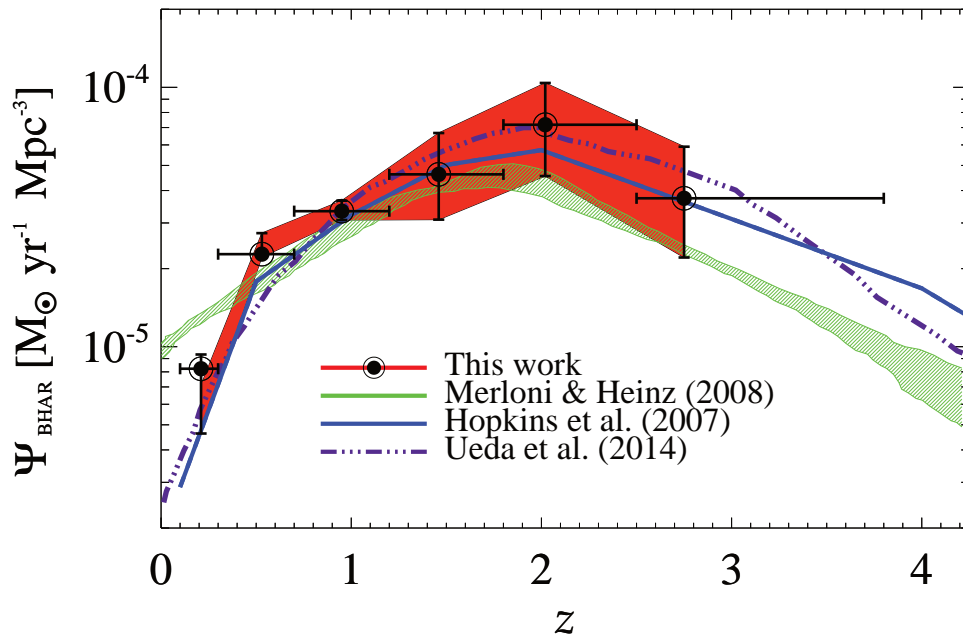
## 4.4 SMBH growth across cosmic time

Given the estimated LF, we are able to derive the Black Hole Accretion Rate Density (BHAD or  $\Psi_{\text{bhar}}$ ) over cosmic time, from  $z \sim 3$  down to the present epoch. This quantity is fundamental for characterising the effective growth of AGN and is defined by the following

<sup>2</sup>Note that the  $L^*$  values derived by H07 have been corrected in this work by a factor of  $\sim 0.18$  dex.



**Figure 4.4:** *Top panel:* evolution with redshift of the normalization value  $\Phi^*$ , from this work (black circles) and from H07 (red circles), taken from the respective best-fit LF. *Middle panel:* evolution of the knee luminosity of the LF. Error bars correspond to the best-fit uncertainty in each redshift bin. *Bottom panel:* AGN number density of different luminosity bins as a function of redshift.



**Figure 4.5:** Black Hole Accretion Rate Density estimate from the AGN bolometric LF as a function of redshift (black circles). The red shaded area shows the  $\pm 1\sigma$  uncertainty region. Previous estimates from different selection wavelengths (from Merloni & Heinz 2008b, H07 and Ueda et al. 2014) are reported for comparison.

expression:

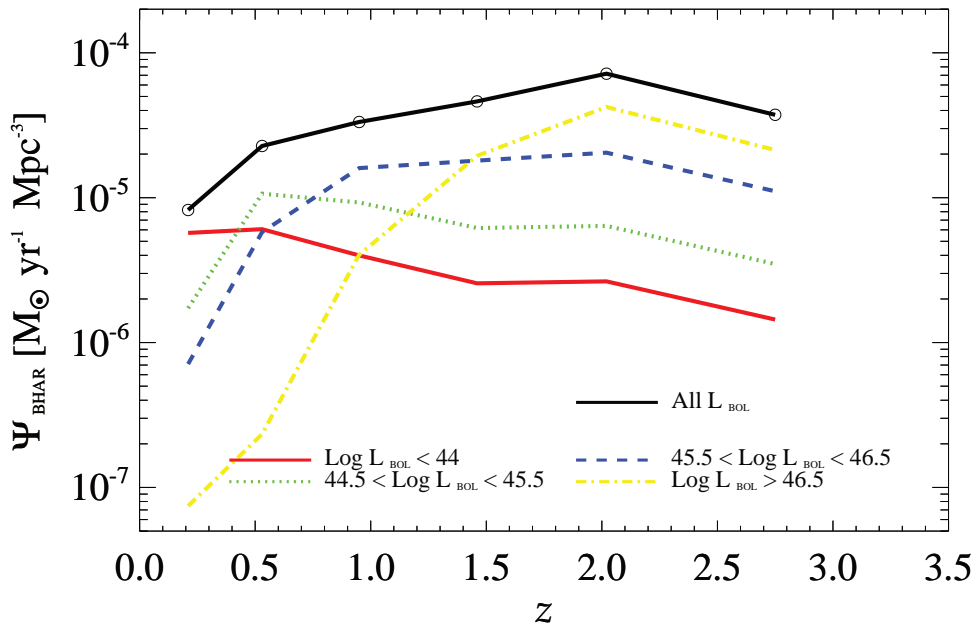
$$\Psi_{\text{bhar}}(z) = \int_0^{\infty} \frac{1 - \epsilon_{\text{rad}}}{\epsilon_{\text{rad}} c^2} L_{\text{bol,AGN}} \phi(L_{\text{bol,AGN}}) d \log L_{\text{bol,AGN}} \quad (4.5)$$

where all the ingredients are already available. The only parameter that still needs to be adopted is the radiative efficiency  $\epsilon_{\text{rad}}$ , which is set to the canonical value of 0.1, and assumed constant with redshift and intrinsic AGN luminosity (e.g. Marconi et al. 2004). The factor  $(1 - \epsilon_{\text{rad}})$  is necessary to account for the fact that not all the accreting mass falls into the SMBH.

#### 4.4.1 Comparison with previous results

In Fig. 4.5 we report our BHAD estimate compared with previous findings from the literature.

Merloni & Heinz (2008b hereafter MH08) estimated the BHAD from the evolution of the hard X-ray LF taken by Silverman et al. (2008), taking a distribution of column densities  $N_{\text{H}}$  (Ueda et al. 2003) and a set of luminosity-dependent X-ray bolometric corrections from Marconi et al. (2004). By solving the continuity equation for the SMBH mass function (assuming its local value from Shankar et al. 2009) and combining that with the evolution of the hard X-ray LF, they were able to trace the BHAD also as a function of black hole mass and accretion rate. It is worth mentioning that MH08 show in their Fig. 4 that most of the local ( $z \sim 0.1$ ) SMBH growth is hidden in radiatively inefficient accreting systems, whose released energy is dominated by kinetic feedback rather than radiative losses. Such “silent” (Eddington ratio  $\lambda_{\text{Edd}} < 3 \times 10^{-2}$ ) AGN population probably does not enter our



**Figure 4.6:** Evolution with redshift of the BHAD, as a function of different AGN population. The black solid line is the total curve shown in Fig. 4.5, while the coloured lines mark different ranges of AGN bolometric luminosity.

sample and could be responsible for the discrepancy (by a factor of about 1.7) shown in Fig. 4.5 in the local BHAD between MH08 and the other estimates reported in the plot. As mentioned in Section 3.3, H07 built up an observed quasar LF by consistently connecting different LFs, each one computed in a single wave band.

Ueda et al. (2014, U14 hereafter) used the most updated compilation of X-ray luminosity functions by collecting samples of X-ray selected AGN from different surveys (*Swift*/BAT, MAXI, ASCA, *XMM-Newton*, *Chandra* and *ROSAT*). They derived the evolution of the AGN bolometric luminosity function by assuming a set of X-ray bolometric corrections from H07 and the AGN absorption function ( $N_{\text{H}}$  distribution) from Ueda et al. (2003). The overall BHAD inferred by U14 is in good agreement with our derivation, showing a peak at  $z \sim 2$  and declining at lower and higher redshift.

To guarantee a coherent comparison between our study and other works taken from the literature, we checked that MH08, H07 and U14 have accounted for the fraction of missed AGN because of their own selection effects. MH08 estimated the fraction of highly-obscured ( $N_{\text{H}} > 1.5 \times 10^{24} \text{ cm}^{-2}$ ) AGN from the X-ray background (XRB) synthesis models by Gilli et al. (2007). In particular, they calculated the expected fraction between observed and unabsorbed 2–10 keV X-ray luminosity, by assuming a canonical AGN X-ray spectrum (absorbed power-law with photon index  $\Gamma=1.8$ ; Tozzi et al. (2006), and as a function of obscuration in the range  $21 < \log(N_{\text{H}}/\text{cm}^{-2}) < 24$ . For  $\log(N_{\text{H}}/\text{cm}^{-2}) > 24$  they assumed that the ratio between observed and intrinsic X-ray luminosity was of the order of 2% (Gilli et al. 2007).

H07 estimated the incompleteness fraction with a set of luminosity dependent  $N_{\text{H}}$  distributions. First, they collected a wide set of luminosity-dependent AGN spectral shapes

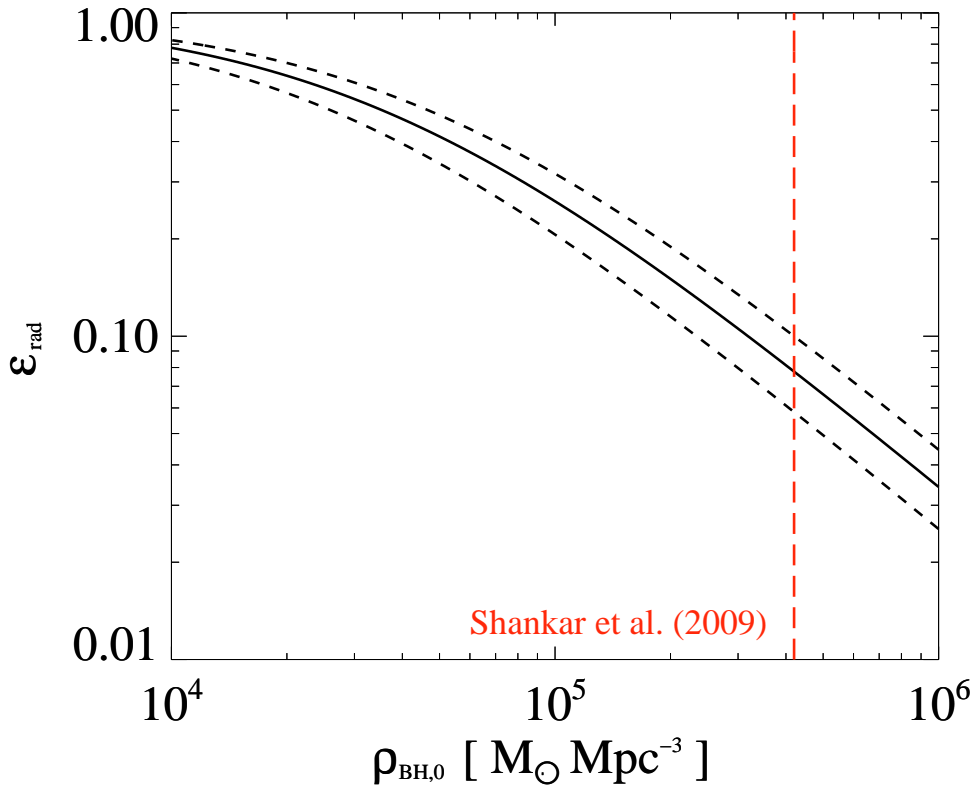
and calculated the bolometric correction distributions, in case of no obscuration. In addition, they implemented three possible models for obscuration. For each parametrization, H07 calculated the expected amount of extinction in X-rays, optical and mid-IR. Finally, through a convolution between AGN templates, bolometric correction distributions and  $N_{\text{H}}$  distributions, they were able to predict the incompleteness fraction of AGN in each band.

U14 have modelled the AGN absorption function through a likelihood ratio method to reproduce the count-rate distribution of both Compton-Thick and Compton-Thin AGN populations. The obscured AGN fraction has been taken into account on the basis of a hard X-ray ( $>15$  keV) selected AGN sample in the local Universe, for which detailed spectral information was available.

Our estimates (black circles) have been derived by integrating the best-fit curve of the LF (already corrected for incompleteness, as explained in Section 4.3.2), down to  $10^8 L_{\odot}$ , in each redshift bin. The red shaded area represents the  $\pm 1\sigma$  uncertainty region. The latter has been computed by accounting for the 100 different integrated  $\Phi_{\text{MC}}(L, z)$  values and plotting their cumulative distribution at 16 and 84 percentiles in each redshift bin. The BHAD seems to evolve quickly from  $z > 3$  to  $z \sim 2$ , where it shows a peak, then decreases towards the present epoch. Despite our AGN sample having been selected and analysed independently of the previous ones, the overall trend is consistent with others taken from the literature, as shown in Fig. 4.4. This is reassuring, since both our estimate and that obtained by H07 and MH08 have been already corrected for their own incompleteness effects and should be in fair agreement as a proof of their mutual consistency. All the estimates are plotted for the canonical case  $\epsilon_{\text{rad}} = 0.1$ . A different value of  $\epsilon_{\text{rad}}$  simply results in a change of the BHAD normalization.

Fig. 4.6 shows the evolution with redshift of the total BHAD (black solid line), splitted as a function of AGN bolometric luminosity. The coloured lines mark different ranges in  $\log L_{\text{bol,AGN}}$ . The integrated curve peaks at  $z \sim 2$ , where  $L_{\text{bol,AGN}} > 10^{46}$  erg  $\text{s}^{-1}$  make the largest contribution. Lower AGN bolometric luminosities peak at lower redshifts, as a consequence of the downsizing behaviour already observed in Fig. 4.4 (bottom panel). As found by different studies (e.g. Silverman et al. 2008) this further suggests that most of the SMBH growth at  $z \sim 1-2$  has been made by Quasars, while in the nearby Universe ( $z < 0.5$ ) Seyfert-like luminosities ( $L_{\text{bol,AGN}} < 10^{45}$  erg  $\text{s}^{-1}$ ) dominate the overall SMBH growth.

An important constraint that allows us to test the consistency of the integrated BHAD( $z$ ) is to deal with Soltan's argument, i.e. deriving the local BH mass density and comparing it to the estimate achieved from observations. This represents the overall energy density released by SMBHs during cosmic time, down to  $z = 0$ . One of the most recent estimates comes from Shankar et al. (2009), as mentioned before. Tuning the integrated BHAD evolution in order to reproduce the observed  $\rho_{\text{bh},0}$  is a way to check the range spanned by the radiative efficiency  $\epsilon_{\text{rad}}$ . From the standard relativistic accretion theory it is known that  $\epsilon_{\text{rad}} = 0.06-0.20$  (Novikov & Thorne 1973). We integrate the BHAD over the redshift range from  $z=6$  to  $z=10^{-5}$  (but taking  $z=\infty$  as upper boundary would change our result by a factor  $< 1\%$ ). By assuming  $\epsilon_{\text{rad}} = 0.1$ , as shown in Fig. 4.7, our local BH mass density



**Figure 4.7:** SMBH radiative efficiency  $\epsilon_{\text{rad}}$  as a function of the local BH mass density predicted by this work (black solid line, with  $\pm 1\sigma$  uncertainty shown by the dashed lines). The vertical red dashed line is the estimate of the observed local BH mass density by Shankar et al. (2009).

is

$$\rho_{\text{bh},0} = 3.1_{-0.8}^{+1.0} \times 10^5 \text{ M}_{\odot} \text{ Mpc}^{-3} \quad (4.6)$$

which is consistent with previous estimates. Indeed, for a typical matter-to-radiation conversion efficiency  $\epsilon_{\text{rad}} = 0.1$ , Yu & Tremaine (2002) found  $\rho_{\text{bh},0} = 2.9 \pm 0.5 \times 10^5 \text{ M}_{\odot} \text{ Mpc}^{-3}$ ; Marconi et al. (2004) reported  $\rho_{\text{bh},0} = (4.6_{-1.4}^{+1.9}) \times 10^5 \text{ M}_{\odot} \text{ Mpc}^{-3}$ ; Shankar et al. (2009) achieved a value of  $\rho_{\text{bh},0} = (4.2_{-1.0}^{+1.2}) \times 10^5 \text{ M}_{\odot} \text{ Mpc}^{-3}$ . All these measurements come from integration of the SMBH mass function by assuming  $\epsilon_{\text{rad}} = 0.1$ .

To fully reconcile the SMBH growth derived here with the observed local BH mass density (e.g. the one from Shankar et al. 2009), we need to adopt  $\epsilon_{\text{rad}} = 0.076_{-0.018}^{+0.023}$ , which is in broad agreement with previous measurements. Later on, Merloni et al. (2004) found  $0.04 < \epsilon_{\text{rad}} < 0.12$  and H07 report  $\rho_{\text{bh},0} = 4.81_{-0.99}^{+1.24} \times 10^5 \text{ M}_{\odot} \text{ Mpc}^{-3}$ , which results in a slightly higher radiative efficiency ( $\epsilon_{\text{rad}} \sim 0.1$ ). In addition, MH08 ended up with an intermediate value ( $\sim 0.07$ ), while U14 found  $\epsilon_{\text{rad}} \sim 0.08_{-0.009}^{+0.013}$ , which is in good agreement with our results.

# Mapping the AGN accretion in the SFR– $M_{\star}$ plane

In this Chapter we study the relation of AGN accretion, SFR and  $M_{\star}$  using a sample of about 8600 star forming galaxies up to  $z=2.5$ , selected with *Herschel* imaging in the GOODS-S, GOODS-N and COSMOS fields. The main purpose is to map the average AGN accretion rate in the SFR– $M_{\star}$  plane at different cosmic epochs, to infer the best predictor of AGN activity in star forming galaxies observed with *Herschel*.

The Chapter is structured as follows. We recall previous studies focusing on the AGN/galaxy connection in Section 5.1. We briefly present our *Herschel*-selected galaxies (Section 5.2), and place them in the SFR– $M_{\star}$  plane at  $0 < z \leq 2.5$  (Section 5.3). We derive average X-ray luminosities for *Herschel* galaxies (both detected and undetected in X-rays) in Section 5.4, which are then used as a proxy to measure the average AGN accretion rate. We present the observed relationships between AGN accretion and galaxy properties in Section 5.5 and discuss the implications of this work in Section 5.6.

## 5.1 Introduction

Studies exploiting *Spitzer* data, and subsequently corroborated with *Herschel* (see Section 2.5.1), have highlighted that more than 95% of star forming galaxies follow the well-known MS relation between SFR and  $M_{\star}$ , from the local Universe up to  $z \sim 3$  (Noeske et al. 2007; Daddi et al. 2007; Magdis et al. 2010; Elbaz et al. 2011; Whitaker et al. 2012; Schreiber et al. 2014). This trend is currently thought to reflect a large duty cycle of steady star formation in galaxies, fueled by a continuous gas inflow (Dekel et al. 2009; Ciotti et al. 2010). The most massive galaxies have larger gas reservoirs (Tacconi et al. 2013) and thus higher SFR. On the other hand, starbursts are outliers with respect to this relation, showing  $> 4$  times larger SFRs compared to the MS rate at a fixed  $M_{\star}$  (e.g. Rodighiero et al. 2011). This distinction is supported by several studies, claiming a systematic variation of several galaxy properties with increasing offset from the MS: more compact structures (Elbaz et al. 2011; Wuyts et al. 2011b), warmer interstellar dust (Magnelli et al. 2014), higher gas-to- $M_{\star}$  ratio (e.g. Gao & Solomon 2004), larger Far-to-Mid Infrared flux ratio (Nordon et al. 2010,

2012) and higher star formation efficiency (SFE<sup>1</sup>, Daddi et al. 2010b; Genzel et al. 2010). Whether the transition from MS to starburst galaxies is steady or discontinuous remains an open question.

As discussed in Section 2.5.4, studies based on X-ray selected samples of AGN found a similar two fold-scheme, where AGN with  $L_X < 10^{44}$  erg s<sup>-1</sup> show a weak correlation with SFR, while bright Quasars seem to follow a positive correlation with SFR at least up to  $z \sim 1$ , probably driven by major mergers (Lutz et al. 2008; Netzer 2009; Shao et al. 2010; Lutz et al. 2010; Rosario et al. 2012). At  $z \sim 2$  such a correlation seems weak or even absent (Rosario et al. 2012; Harrison et al. 2012). However, compelling evidences of AGN-driven feedback (e.g. Farrah et al. 2012) and the apparently controversial inverted correlation found in smaller samples of luminous  $z \sim 2$  AGN (Page et al. 2012), caution that our current picture of AGN/galaxy coevolution is still incomplete, and dependent on sample statistics and selection biases.

In contrast to the studies of AGN hosts and to reach a comprehensive understanding of the cosmic SMBH growth, several recent studies try to take a census of AGN accretion history on the basis of SFR and/or mass-selected samples of galaxies. Unlike the weak or absent correlation of star formation and black hole accretion rate (BHAR<sup>2</sup>) for individual AGN hosts, there is clear correlation of *average* BHAR and key properties of galaxy samples. Positive and close to linear correlation is found between average BHAR and mean stellar mass at various redshifts (Mullaney et al. 2012a) as well as with SFR (Rafferty et al. 2011; Chen et al. 2013).

These apparently contradictory results have been interpreted by Hickox et al. (2014) as due to different variability time scales between nuclear activity and global star formation. In this context, all episodes of star formation are accompanied by SMBH growth, but only when smoothing over the variations of individual sources, the average properties of AGN and their hosts show a mutual evolution, as stated by Mullaney et al. (2012a) and Chen et al. (2013).

Nevertheless, it is worth mentioning that both these latter studies derived average trends by binning their parent samples in either  $M_*$  or SFR. Since their selection techniques were mostly sensitive to main-sequence galaxies, in principle the resulting correlations found with average BHAR might be primarily due to one parameter, but reflected into a correlation with the other one, simply because of the main-sequence relation that holds between the two. To break this degeneracy and investigate in detail the role of AGN accretion in the context of galaxy evolution, it is necessary to split the sample as a function of both SFR and  $M_*$  and study the evolution of the average AGN accretion properties in the SFR– $M_*$  plane at different redshifts.

<sup>1</sup>SFE is defined as the ratio between SFR and cold gas mass.

<sup>2</sup>The terms “AGN rate”, “AGN accretion rate”, BHAR and  $\langle \dot{M}_{\text{bh}} \rangle$  adopted in the text are assumed to have the same physical meaning.



## 5.2 *Herschel* dataset and X-ray counterparts

Our sample selection exploits *Herschel*-PACS and SPIRE observations in the GOODS-N, GOODS-S and COSMOS fields. We refer the reader to Section 3.1 for a detailed description of the parent sample and multi-wavelength identifications at shorter wavelengths. Totally we collect 892, 850 and about 17000 *Herschel* galaxies in the GOODS-S, GOODS-N and COSMOS sky areas, respectively.

### 5.2.1 Cross-match with X-rays

We used optical/near-infrared counterpart positions to cross-match our *Herschel*-selected sample with available X-ray data from *Chandra* observations in COSMOS and in the GOODS fields.

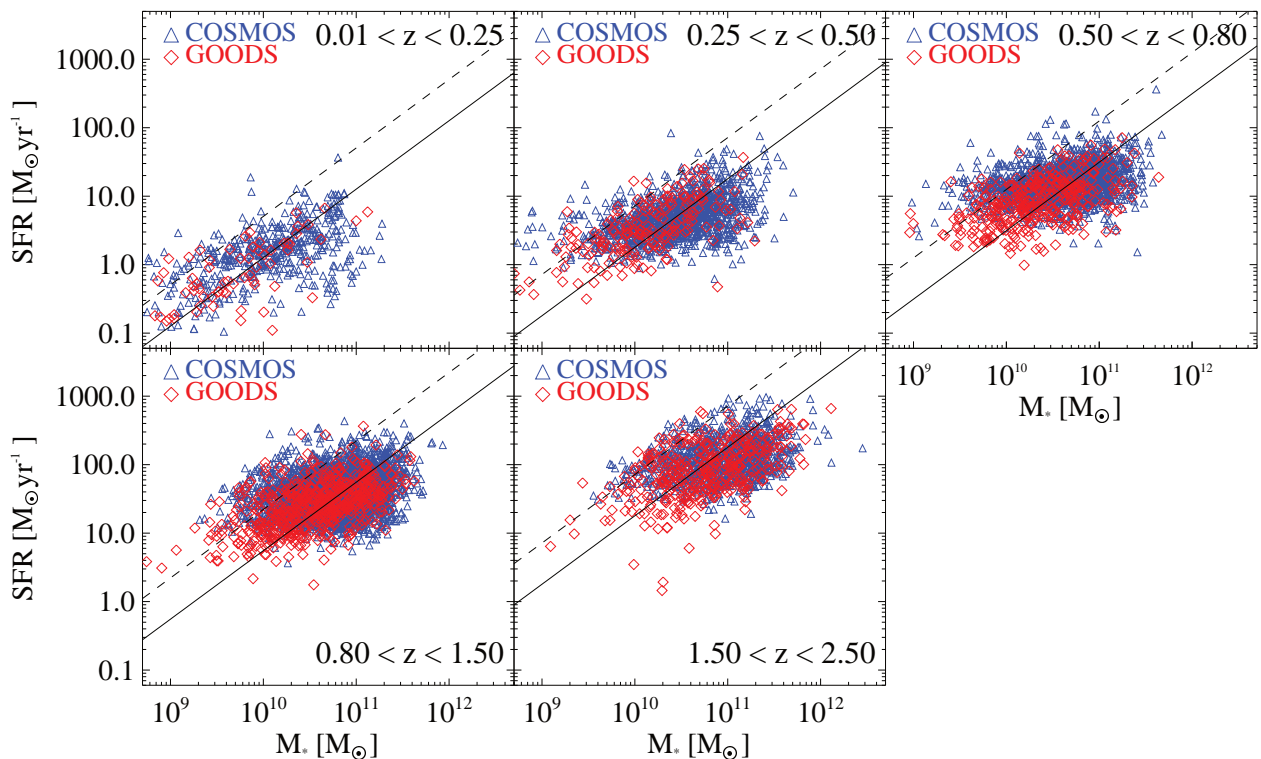
- In the GOODS-South, we use the 4-Ms *Chandra*-Deep Field South (CDF-S) observations (Xue et al. 2011). The X-ray catalogue provides count rates and observed fluxes for each source in different bands: soft (0.5–2 keV), hard (2–8 keV) and full (0.5–8 keV).
- In the GOODS-North, X-ray data are available from the 2-Ms *Chandra*-Deep Field North (CDF-N) observations (Alexander et al. 2003).
- In the COSMOS field, 200-ks observations from the *Chandra*-COSMOS (C-COSMOS, Elvis et al. 2009; Civano et al. 2012) survey are publicly available but cover only about 0.9 deg<sup>2</sup> instead of  $\sim 2$  deg<sup>2</sup> scanned by *Herschel*. Consequently, we limit our parent sample to the common sky area, which implies a cut of  $\sim 55\%$  of the original population (from 17000 to 7272 FIR-selected galaxies).

After having collected as many spectroscopic redshifts as possible (see Section 3.1.2), we cut our original sample at redshift  $z \leq 2.5$ , since above this threshold poor statistics would affect the significance of our results<sup>3</sup>. The overall fraction of spectroscopic redshifts is larger than 60% at  $z \leq 1.5$ , while it decreases to 20% at  $z \sim 2$ . We stress that, if limiting our sample to spectroscopic redshifts only, all the results would remain consistent within  $1\sigma$  uncertainty with those that will be presented and discussed in Sections 5.5 and 5.6.

This cut in redshift leads to 829, 804 and 7011 sources in the GOODS-S, GOODS-N and COSMOS, respectively, for a total number of 8644 FIR-selected galaxies across all fields. The cross-match with X-ray detections has been performed by means of a nearest neighbour algorithm, by assuming 1 arcsec as matching radius between optical positions. The number of X-ray detections reaches 212/829 in the GOODS-S, 134/804 in the GOODS-N and 448/7011 in the COSMOS field.

As highlighted in previous studies (e.g. Rosario et al. 2012), we confirm that the fraction of *Herschel* sources detected in X-rays is generally small and is a function of

<sup>3</sup>In the following, our sample will be split in bins of SFR,  $M_*$  and redshift. Despite our large parent sample, this multi-dimensional binning significantly reduces our statistics in each SFR- $M_*$ - $z$  slice. This is the main reason why we choose to limit ourselves to sources at  $z \leq 2.5$



**Figure 5.1:** Individual *Herschel*-selected sources as a function of SFR,  $M_*$  and redshift. Red and blue symbols refer to GOODS-S/N and COSMOS sources, respectively. The black solid line at any redshift represents the main-sequence defined by Elbaz et al. (2011). The black dashed line marks  $4 \times$  higher sSFR as a threshold between main and off-sequence galaxies.

field, depending on the relative depth of X-ray/IR observations. The fraction of *Herschel* sources detected by *Chandra* is, on average, less than 10%. In particular, it reaches 26% in the GOODS-S, 17% in the GOODS-N and only  $\sim 6\%$  in COSMOS. In all the fields, we have taken the observed X-ray fluxes in each band (soft, hard and full) from the publicly available catalogues.

We derived the average X-ray properties for the rest of the sample by performing a stacking analysis on X-ray maps (see Section 5.4.1 and Appendix B).

### 5.3 The SFR– $M_*$ plane

Given the wealth of photometric datapoints available in all fields from the UV to the sub-mm, in Section 3.2 we have performed a broad-band SED decomposition to derive  $M_*$  and SFR for the entire sample. Each observed SED has been fitted with the MAGPHYS code (da Cunha et al. 2008), as well as with a modified version of MAGPHYS adapted to include an AGN component (Berta et al. 2013b). The best-fit obtained with the AGN is preferred if the resulting  $\chi^2$  value significantly decreases (at  $\geq 99\%$  confidence level, on the basis of a Fisher test, see Section 3.4) compared to the fit without the AGN. In this case, SFRs and  $M_*$  estimates are taken from the fit with AGN, otherwise from the original MAGPHYS code. However, even if the AGN component is required in the best-fit, we stress

that its contribution to the galaxy IR luminosity is mostly not dominant, amounting to  $\lesssim 10\%$  for the majority of our sources (see Section 3.4).

The SFR has been derived by converting the total IR (rest 8-1000  $\mu\text{m}$ ) luminosity taken from the best-fit galaxy SED (i.e. corrected for a possible AGN emission) using the conversion from Kennicutt (1998), scaled to a Chabrier 2003 IMF. The  $M_\star$  is derived from the SED decomposition itself, which allows to get robust measurements also for type-1 AGN, as most of them show near-IR ( $\approx 1 \mu\text{m}$ ) emission dominated by the host galaxy light (e.g. Bongiorno et al. 2012). We have checked that our estimates of  $M_\star$  for optically identified type-1 AGN are consistent within the uncertainties ( $\approx 0.3$  dex) with those presented by Bongiorno et al. (2012), with no systematics.

The sample has been split in five different redshift bins:  $0.01 \leq z < 0.25$ ,  $0.25 \leq z < 0.50$ ,  $0.50 \leq z < 0.80$ ,  $0.80 \leq z < 1.50$  and  $1.50 \leq z < 2.50$ . We place our galaxies on the SFR– $M_\star$  plane in Fig. 5.1, marking with different colours GOODS (red) and COSMOS (blue) sources. Given that PACS flux density limits in the GOODS fields are about 5 times lower than in COSMOS, *Herschel* observations in the GOODS fields may detect fainter IR galaxies (i.e. lower SFR) compared to observations in COSMOS at the same redshift. Since our *Herschel*-based selection is sensitive to the most star forming galaxies in each field, the wedge traced by the observed galaxy distribution in the SFR– $M_\star$  is relatively flat (see Rodighiero et al. 2011) compared to the linear main-sequence relation defined by Elbaz et al. (2011). The MS evolution with redshift is parametrized by Elbaz et al. 2011 as  $\text{sSFR}_{\text{ms}} = 26 \times t_{\text{cosmic}}^{-2.2} [\text{Gyr}^{-1}]$ , where  $t_{\text{cosmic}}$  is the cosmic time (in Gyr) starting from the Big Bang. Our selection also includes off-sequence galaxies, with  $\text{sSFR} > 4 \times \text{sSFR}_{\text{ms}}$  ( $\sim 0.6$  dex), corresponding to symbols above the black dashed lines. In each redshift slice, the sample has been split in both SFR and  $M_\star$ , taking  $0.5 \times 0.5 \text{ dex}^2$  bins<sup>4</sup>. Instead of keeping a fixed binning configuration in SFR and  $M_\star$  at all redshifts, we decided to make our bins centered on the main-sequence relation corresponding to each bin of  $M_\star$  and redshift<sup>5</sup>. This arrangement is preferable, since it allows us to better highlight any potential systematics in terms of average AGN properties between main and off-sequence galaxies.

## 5.4 X-ray analysis

In this Section we present the X-ray analysis of our *Herschel* selected sample. For sources detected in X-ray ( $\sim 10\%$ ) we get count rates, fluxes and observed luminosities in different bands from the above-mentioned catalogues (Section 5.2.1). X-ray undetected sources represent most ( $\sim 90\%$ ) of the *Herschel* sample studied in this work. To derive their average X-ray properties, we performed a stacking analysis on X-ray maps (Section 5.4.1), in the observed soft, hard and full bands. Subsequently, we characterize the average X-ray properties of *Herschel* galaxies in each bin of SFR,  $M_\star$  and redshift by grouping together X-ray detections and stacks (Section 5.4.2). Then we subtract the X-ray flux expected

<sup>4</sup>We note that the typical uncertainty on individual SFR and  $M_\star$  measurements is of the order of 0.2–0.3 dex, so significantly smaller than the bin-width.

<sup>5</sup>We checked also for alternative binning configurations (i.e. either different shapes for individual bins, or different placement on the SFR– $M_\star$  plane) but this did not produce any significant impact on our results.

from star formation (Section 5.4.3) in single X-ray bands and correct the remaining X-ray emission for the nuclear obscuration (Section 5.4.4) to derive the intrinsic (unobscured) mean AGN X-ray luminosity (rest-frame). Using widely-adopted conversion factors, we consider the final nuclear X-ray luminosity as a proxy to estimate the mean AGN accretion rate in each bin of SFR,  $M_*$  and redshift (see Section 5.5).

#### 5.4.1 Stacking on X-ray maps

Here we summarize the main steps concerning the X-ray stacking, but refer the reader to Appendix B for a detailed description. We combined all *Herschel* sources undetected in X-rays from both GOODS and COSMOS fields and we grouped them together as a function of SFR,  $M_*$  and redshift. After a proper masking of all X-ray detected sources which could potentially affect the stacked signals, we piled up single cutouts of X-ray undetected sources in the same bin of SFR,  $M_*$  and redshift, centered on their optical coordinates. For each stacked object, we defined regions from which we extract source and background photons, in order to derive background-subtracted (i.e. net) photon counts. Finally, we needed to correct for differential sensitivities between various X-ray fields, so we normalized the resulting net photon counts of each object by the corresponding effective (i.e. corrected for instrumental effects) exposure time, which provides exposure-corrected mean count rates in each observed X-ray band<sup>6</sup>. To convert count rates into observed fluxes, we assumed a power-law spectrum with  $\Gamma=1.4$ , based on empirical spectrum of the X-ray background (e.g. Gilli et al. 2007).

#### 5.4.2 Mean X-ray luminosity of *Herschel* sources

To get the overall X-ray properties of our *Herschel*-selected galaxies, in each bin of SFR,  $M_*$  and redshift we collect both X-ray detected and undetected sources. Assuming that the bin includes  $N$  sources,  $m$  detected and  $n$  undetected in X-rays, we compute a weighted linear average of their X-ray fluxes, according to the following formula:

$$\langle S_{\text{bin}} \rangle = \frac{n \times S_{\text{stack}} + \sum_{i=1}^m S_i}{N} \quad (5.1)$$

where  $S_{\text{stack}}$  is the mean X-ray flux obtained by stacking  $n$  undetected sources and  $S_i$  is the individual flux measured for the  $i$ -th detected source within the same bin. We stress that the above-mentioned expression is a weighted linear average, therefore appropriate to investigate *mean* properties (e.g. mean X-ray luminosity) of the underlying galaxy population, but we caution that it does not necessarily represent the most *probable* value of the observed distribution of a given parameter.

As a sanity check, we constructed and analysed the maximum likelihood (ML) function to estimate the most probable X-ray flux (or luminosity) taking into account both detected and undetected sources. The detailed description of this alternative method is given in Appendix D. We found that Eq. 5.1 and the ML function provide consistent estimates of the average X-ray flux for our *Herschel* sample. This suggests that the average X-ray flux

<sup>6</sup>We note that our results are in good agreement, within the uncertainties, with those obtained using CSTACK (<http://cstack.ucsd.edu/cstack/>, developed by T. Miyaji) on the same X-ray maps.

given in Eq. 5.1 is really representative of the overall X-ray population. In addition, the consistency with the ML-based analysis validates the assumption that X-ray detected and undetected sources are part of the same underlying population.

Considering that  $\approx 90\%$  of *Herschel* galaxies are not X-ray detected, as well as the fact that X-ray detections are about 50 times brighter than stacks, in most bins X-ray detected and stacked sources provide comparable contributions to the mean X-ray flux. Rest-frame average X-ray luminosities in each band are derived from mean X-ray fluxes by assuming a power-law spectrum with intrinsic slope  $\Gamma=1.9$  (Tozzi et al. 2006; Mainieri et al. 2007) and no obscuration. We performed this calculation as a function of SFR,  $M_\star$  and redshift.

### 5.4.3 Subtraction of X-ray emission from star formation

Using X-rays to investigate the level of AGN activity in the SFR– $M_\star$  plane requires the subtraction of X-rays from other processes in the host galaxy, especially those related to star formation. This subtraction will allow us to apply the correction for nuclear obscuration (Section 5.4.4) only to the AGN-related X-ray emission rather than the total (i.e. AGN + galaxy) one. It is known that X-ray observations provide a relatively clean selection of accreting SMBHs, but a fraction of the X-ray emission (e.g. in *Chandra* Deep Fields, CDFs) may also come from X-ray binaries and the hot interstellar medium (e.g. Mineo et al. 2012a,b). A single threshold (i.e.  $L_X = 3 \times 10^{42}$  erg s $^{-1}$ ) in X-ray luminosity is not a proper way of classifying this galaxy population, since it is known that X-ray emission from star forming galaxies shows a positive correlation with SFR. Previous works (e.g. Ranalli et al. 2003) calibrated this relation in the local Universe, while recent studies (Vattakunnel et al. 2012; Mineo et al. 2014; Symeonidis et al. 2014) uncovered this relation up to  $z \sim 1.5$ .

Though all these relations are calibrated through independent analyses and selection techniques, they provide reasonably consistent (within a factor of 2) estimates of the X-ray emission expected from star formation. However, we prefer to use the relation by Symeonidis et al. 2014, since they also exploited *Herschel* data and performed stacking on X-ray maps to better characterize the average  $L_X$ –SFR correlation in inactive (i.e. non-AGN) SFR-selected galaxies. They found a quasi linear relation holding at  $1 < \text{SFR} < 1000$   $M_\odot \text{ yr}^{-1}$  and not evolving significantly with redshift up to  $z \sim 1.5$ . We scaled their relation to a Chabrier 2003 IMF and converted to our rest-frame soft (0.5–2 keV) and hard (2–8 keV) X-ray bands:

$$L_X^{\text{soft}} [\text{erg s}^{-1}] = 2.04 \times 10^{39} \text{ SFR} [M_\odot \text{ yr}^{-1}] \quad (5.2)$$

$$L_X^{\text{hard}} [\text{erg s}^{-1}] = 5.13 \times 10^{39} \text{ SFR} [M_\odot \text{ yr}^{-1}] \quad (5.3)$$

In each bin we subtracted the X-ray luminosity due to star formation from the average X-ray luminosity (Section 5.4.2), both in soft and hard band. However, we note that this subtraction does not impact significantly the previous values, as the typical non-AGN contribution arising from the joint (X-ray detected + stacked) sample is less than 10%.

#### 5.4.4 Correction for nuclear obscuration

To derive the intrinsic AGN luminosity in each bin of SFR,  $M_\star$  and redshift, we consider the AGN-related X-ray emission derived in Section 5.4.3 and follow the approach developed by Xue et al. (2010), who used the hardness ratio (HR) as a proxy to estimate the nuclear obscuration. The hardness ratio is defined as:

$$\text{HR} = (\text{CR}_{\text{hard}} - \text{CR}_{\text{soft}}) / (\text{CR}_{\text{hard}} + \text{CR}_{\text{soft}}) \quad (5.4)$$

where  $\text{CR}_{\text{hard}}$  and  $\text{CR}_{\text{soft}}$  represent the (exposure-corrected) count rates in hard and soft-bands, respectively.

Given that we are not allowed to constrain the HR for each individual galaxy in our sample, we correct the mean AGN X-ray emission of all *Herschel* galaxies (both detected and undetected in X-rays, see Eq. 5.1) for an average level of obscuration. We make the simple assumption that the mean HR calculated after the subtraction in soft and hard X-ray bands (Eqs. 5.2, 5.3) is representative of the galaxy population in the same bin of SFR,  $M_\star$  and redshift.

We parametrized the effect of nuclear obscuration by assuming a single power-law X-ray spectrum, with intrinsic photon index  $\Gamma = 1.9$  (model `waxpwa` in XSPEC, Arnaud 1996) and accounting for absorption, both Galactic and intrinsic to the AGN. However, since the hardness ratio deals with photon counts instead of fluxes, one needs to convolve the intrinsic model with the instrument response curve.<sup>7</sup> After performing this convolution, we built a set of simulated spectra that predict the observed hardness ratio as a function of hydrogen column density ( $N_{\text{H}}$ ) and redshift. We calculated the observed hardness ratio from AGN count rates obtained in Section 5.4.2 and selected the spectral model (i.e. intrinsic column density) best reproducing the observed HR at the mean redshift of the underlying galaxy population in the same bin of SFR and  $M_\star$ . From the observed-frame, absorption-corrected fluxes in the 0.5–8 keV band ( $S_{[0.5-8],\text{int}}$ ), we calculated the rest-frame, intrinsic full-band X-ray luminosity as follows:

$$L_{[0.5-8],\text{int}} = 4\pi D_{\text{L}}^2 S_{[0.5-8],\text{int}} (1+z)^{\Gamma-2} \quad (5.5)$$

where  $D_{\text{L}}$  is the luminosity distance corresponding to the mean redshift of the *Herschel* population in a given bin, while the intrinsic photon index  $\Gamma$  is set to 1.9.

As a sanity check, we also compared our obscuration-corrected full-band (0.5–8 keV) luminosities with those presented by Xue et al. (2011) in the *Chandra*-Deep field South (CDF-S) 4-Ms catalogue and found an excellent agreement, as expected, given that we followed similar approaches. We found that obscuration level does not significantly affect the average AGN X-ray luminosity, as the typical correction factor is about 1.3. The obscuration-corrected X-ray luminosities estimated through hardness ratio are generally consistent within a factor of  $\sim 30\%$  with more precise measurements from spectral-fitting analysis (Xue et al. 2011). Nevertheless, as argued by Xue et al. (2011) the level of obscuration might be affected by strong uncertainties in case of highly obscured AGN. In

<sup>7</sup>The instrument response curve has been corrected for several instrumental effects: vignetting, exposure time variations, energy-dependent efficiency, variation of effective area with time of observations.

addition, we compared our predictions based on HR with spectral measurements of  $\sim 400$  *Chandra*-COSMOS AGN presented by Lanzuisi et al. (2013). We ended up with reasonably small scatter (about 0.2 dex) in (0.5–8 keV) X-ray luminosities, even for sources classified as highly obscured ( $N_{\text{H}} > 10^{23} \text{ cm}^{-2}$ ) AGN from X-ray spectral-fitting. Given these sanity checks and the relatively small average corrections for obscuration obtained for our sample, this might suggest that the average intrinsic X-ray luminosity of *Herschel*-selected galaxies does not arise primarily from highly obscured AGN. However, a thorough X-ray spectral analysis of these sources would be beyond the scope of this work, and will be performed in a future study.

### Uncertainty on the intrinsic $L_{\text{X}}$

We evaluated the uncertainty on the intrinsic X-ray luminosity  $L_{[0.5-8],\text{int}}$  by performing a bootstrapping analysis. This technique provides reliable error bars, especially in case a small fraction of the objects dominate the signal.

Suppose there are  $N$  objects populating a given bin of SFR,  $M_{\star}$  and redshift, out of which  $m$  are detected in X-rays, while  $n$  are not detected. We selected at random  $N$  objects from the same bin, allowing duplication of the same source. This random extraction likely leads to different numbers of detections ( $m'$ ) and non-detections ( $n'$ ). We combined the photon counts together from new  $n'$  sources to derive stacked count rates and fluxes. Then we applied Eq. 5.1 (see Section 5.4.2) to get the average X-ray flux representative of this random realization. With 1000 iterations of this calculation, we obtained the distribution of average count rates and fluxes of the galaxy sample. The 16<sup>th</sup> and 84<sup>th</sup> percentiles of the final distribution set the  $1\sigma$  lower and upper bounds on the measured X-ray flux.

To evaluate the uncertainty on the obscuration-corrected  $L_{\text{X}}$ , we iterated 1000 times the same analysis described in Sections 5.4.3 and 5.4.4, once for each random realization. This approach returns the final distribution of the intrinsic X-ray luminosities, with  $\pm 1\sigma$  error bars estimated through bootstrapping. Since this technique would provide quite large error bars (i.e. almost unconstrained fluxes and luminosities) in case of poor statistics, we have required a minimum number of sources in each bin ( $\geq 15$ , regardless of the number of detections and non-detections).

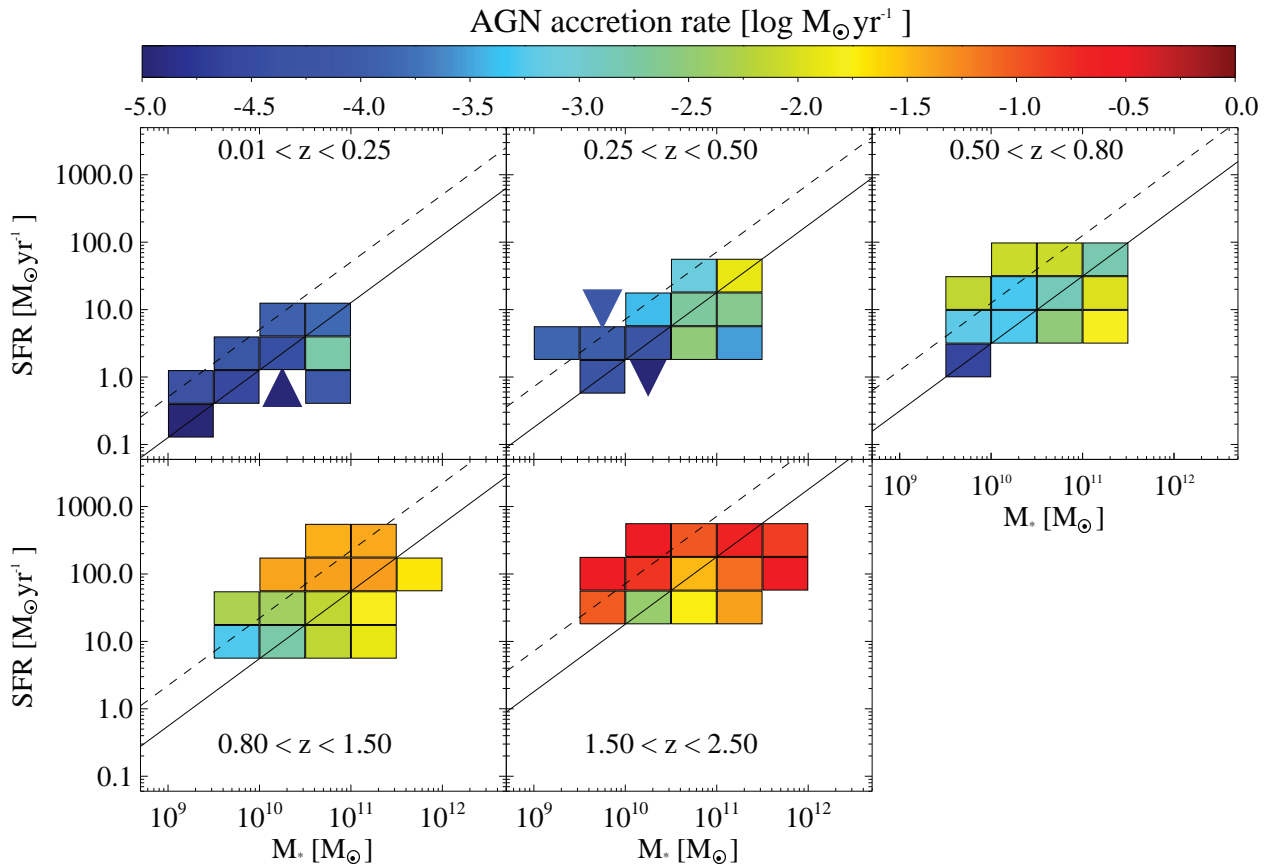
## 5.5 Results

We used the intrinsic X-ray luminosity as a proxy for mapping the average AGN accretion rate in the SFR– $M_{\star}$  plane and studying its correlation with integrated galaxy properties.

### 5.5.1 Average AGN accretion rate

Obscuration-corrected X-ray luminosities in each bin have been turned into average bolometric AGN luminosities  $\langle L_{\text{bol}} \rangle$  by assuming a set of luminosity-dependent X-ray bolometric corrections from Marconi et al. (2004)<sup>8</sup>. For simplicity, we have taken a constant

<sup>8</sup>We remark that if taking a fixed bolometric correction value of 22.4 (as done by Mullaney et al. 2012a and Chen et al. 2013), the AGN bolometric luminosities would be larger by a factor of about 2, for  $L_{\text{X}} < 10^{44}$



**Figure 5.2:** Distribution of the average AGN accretion rate in the SFR– $M_*$  plane at  $0.01 \leq z \leq 2.5$ . The bins in SFR and  $M_*$  are arranged to sample the main-sequence relation (black solid line), which evolves with redshift. Upward and downward triangles set  $1\sigma$  lower and upper limits, respectively, on the average AGN accretion rate. The black dashed line divides main-sequence from off-sequence galaxies.

conversion factor from average AGN bolometric luminosity  $\langle L_{\text{bol}} \rangle$  (in  $\text{erg s}^{-1}$ ) to average SMBH accretion rate  $\langle \dot{M}_{\text{bh}}(M_*, \text{SFR}, z) \rangle$  (in  $M_{\odot} \text{yr}^{-1}$ ), according to the formula (see Alexander & Hickox 2012):

$$\langle \dot{M}_{\text{bh}}(M_*, \text{SFR}, z) \rangle = 0.15 \frac{\epsilon}{0.1} \frac{\langle L_{\text{bol}}(M_*, \text{SFR}, z) \rangle}{10^{45}} \quad (5.6)$$

where the matter-to-radiation conversion efficiency  $\epsilon$  is assumed to be 10% (e.g. Marconi et al. 2004). Table C.1 lists redshift, SFR,  $M_*$  and accretion rates for all bins involved, as well as the number of sources entering each bin. In Fig. 5.2 we show the SFR– $M_*$  plane at different redshifts with colour coded average AGN accretion rate. The black solid line is the main-sequence (MS) relation, whose normalization increases with redshift, preserving the same linear slope (Elbaz et al. 2011). Coloured bins include at least 15 sources, either detected or undetected in X-rays. A few upward and downward triangles set  $1\sigma$  lower and upper limits, respectively. They replace the formal values in case the correction for obscuration is not applicable. In particular, as the expected non-AGN contribution was

$\text{erg s}^{-1}$ .



subtracted from the average X-ray luminosities obtained in Section 5.4.2, in some bins this subtraction left solely the hard (soft) X-ray emission, which returned a lower (upper) limit in the final intrinsic X-ray luminosity. In addition, an upper limit is also imposed in case the observed X-ray spectrum is softer than any spectral model with intrinsic  $\Gamma=1.9$ <sup>9</sup>. This is plausible, since this slope represents just a mean value among the overall distribution of AGN X-ray spectra, whose typical scatter is around 0.2 (e.g. Piconcelli et al. 2005).

We note that our SFR- $M_{\star}$ - $z$  grid follows the evolution with redshift of the MS. Even at our large source statistics, uncertainties in BHAR remain noticeable (about 0.4 dex on average, see Table C.1), also due to the fact that in some bins a few, intrinsically luminous X-ray detections might dominate the global signal, as stated in Equation 5.1. Trends in BHAR are indicated within the panels for each redshift slice, and differences in BHAR between local low SFR and distant high SFR sources are obvious. We proceed to an investigation which galaxy property best correlates with these BHAR differences.

---

<sup>9</sup>The softest X-ray spectrum that is reproducible with intrinsic slope  $\Gamma=1.9$  corresponds to  $HR = -0.47$ .

**Table 5.1:** Best-fit parameters returned by fitting a linear relation in the log–log space, as a function of redshift, between  $\langle \dot{M}_{\text{BH}} \rangle$  and galaxy properties ( $x$ -parameter). Values in brackets set the  $1\sigma$  uncertainty on the related parameters. The Spearman’s rank correlation coefficient  $\rho$  set the strength of the correlation, while  $P(\text{correlation})$  represents the significance of its deviation from zero, considering the number of points  $N$  in each panel. In columns 5 and 6 their estimates have been derived from a linear regression fit, while in columns 7 and 8 we used a partial-correlation fitting (see text for details).

(1) $z$ -bin	(2) $x$ -parameter	(3) $\alpha$	(4) $\beta$	(5) $\rho$	(6) $P(\text{correlation})$	(7) partial $\rho$	(8) partial $P(\text{correlation})$	(9) $N$
$0.01 \leq z < 0.25$	SFR [ $M_\odot \text{ yr}^{-1}$ ]	$0.54 (\pm 0.27)$	$-4.31 (\pm 0.12)$	0.68	95.8%	0.39	83.0%	9
$0.25 \leq z < 0.50$	SFR [ $M_\odot \text{ yr}^{-1}$ ]	$1.06 (\pm 0.24)$	$-4.16 (\pm 0.38)$	0.75	99.1%	0.49	92.6%	11
$0.50 \leq z < 0.80$	SFR [ $M_\odot \text{ yr}^{-1}$ ]	$1.05 (\pm 0.52)$	$-4.02 (\pm 0.64)$	0.60	96.1%	0.40	89.9%	12
$0.80 \leq z < 1.50$	SFR [ $M_\odot \text{ yr}^{-1}$ ]	$1.44 (\pm 0.30)$	$-4.29 (\pm 0.46)$	0.80	99.9%	0.71	99.6%	14
$1.50 \leq z \leq 2.50$	SFR [ $M_\odot \text{ yr}^{-1}$ ]	$1.13 (\pm 0.38)$	$-3.42 (\pm 0.78)$	0.68	99.0%	0.68	99.2%	13
$0.01 \leq z < 0.25$	$M_*$ [ $M_\odot$ ]	$0.44 (\pm 0.21)$	$-8.63 (\pm 2.17)$	0.73	97.5%	0.52	91.7%	9
$0.25 \leq z < 0.50$	$M_*$ [ $M_\odot$ ]	$0.76 (\pm 0.33)$	$-11.41 (\pm 3.51)$	0.75	99.1%	0.52	93.8%	11
$0.50 \leq z < 0.80$	$M_*$ [ $M_\odot$ ]	$1.11 (\pm 0.34)$	$-14.59 (\pm 3.66)$	0.58	95.2%	0.41	89.9%	12
$0.80 \leq z < 1.50$	$M_*$ [ $M_\odot$ ]	$1.07 (\pm 0.18)$	$-13.58 (\pm 1.96)$	0.65	98.8%	0.43	93.0%	14
$1.50 \leq z \leq 2.50$	$M_*$ [ $M_\odot$ ]	$0.25 (\pm 0.27)$	$-3.89 (\pm 2.94)$	0.23	56.3%	-0.17	30.0%	13
$0.01 \leq z < 0.25$	sSFR/sSFR <sub>ms</sub>	$-0.09 (\pm 0.27)$	$-4.17 (\pm 0.13)$	-0.02	33.2%	/	/	9
$0.25 \leq z < 0.50$	sSFR/sSFR <sub>ms</sub>	$-0.18 (\pm 0.42)$	$-3.33 (\pm 0.21)$	-0.17	43.6%	/	/	11
$0.50 \leq z < 0.80$	sSFR/sSFR <sub>ms</sub>	$-0.53 (\pm 0.50)$	$-2.74 (\pm 0.17)$	0.03	36.7%	/	/	12
$0.80 \leq z < 1.50$	sSFR/sSFR <sub>ms</sub>	$-0.53 (\pm 0.40)$	$-2.13 (\pm 0.15)$	0.19	54.6%	/	/	14
$1.50 \leq z \leq 2.50$	sSFR/sSFR <sub>ms</sub>	$0.10 (\pm 0.31)$	$-1.10 (\pm 0.15)$	0.40	83.8%	/	/	13

### 5.5.2 Correlation with galaxy properties

We explore here the observed trends between  $\langle \dot{M}_{\text{bh}} \rangle$  and various galaxy properties: SFR,  $M_{\star}$  and offset from the MS<sup>10</sup>. Because of the evolution with cosmic time of SFRs and BHARs, we fit the data separately for each redshift bin, by assuming a linear trend in the log–log space through:

$$\log(\langle \dot{M}_{\text{bh}} \rangle) = \alpha \times \log(x) + \beta \quad (5.7)$$

where  $\alpha$  and  $\beta$  represent slope and intercept, respectively, while  $x$  is the corresponding independent variable. We have used the IDL routine `IMSL_MULTIREGRESS.PRO`, which performs a linear regression fit considering error bars and returns best-fit intercept and slope with related  $1\sigma$  uncertainties. The linear best-fits obtained in each redshift bin are summarized in Table 5.1.

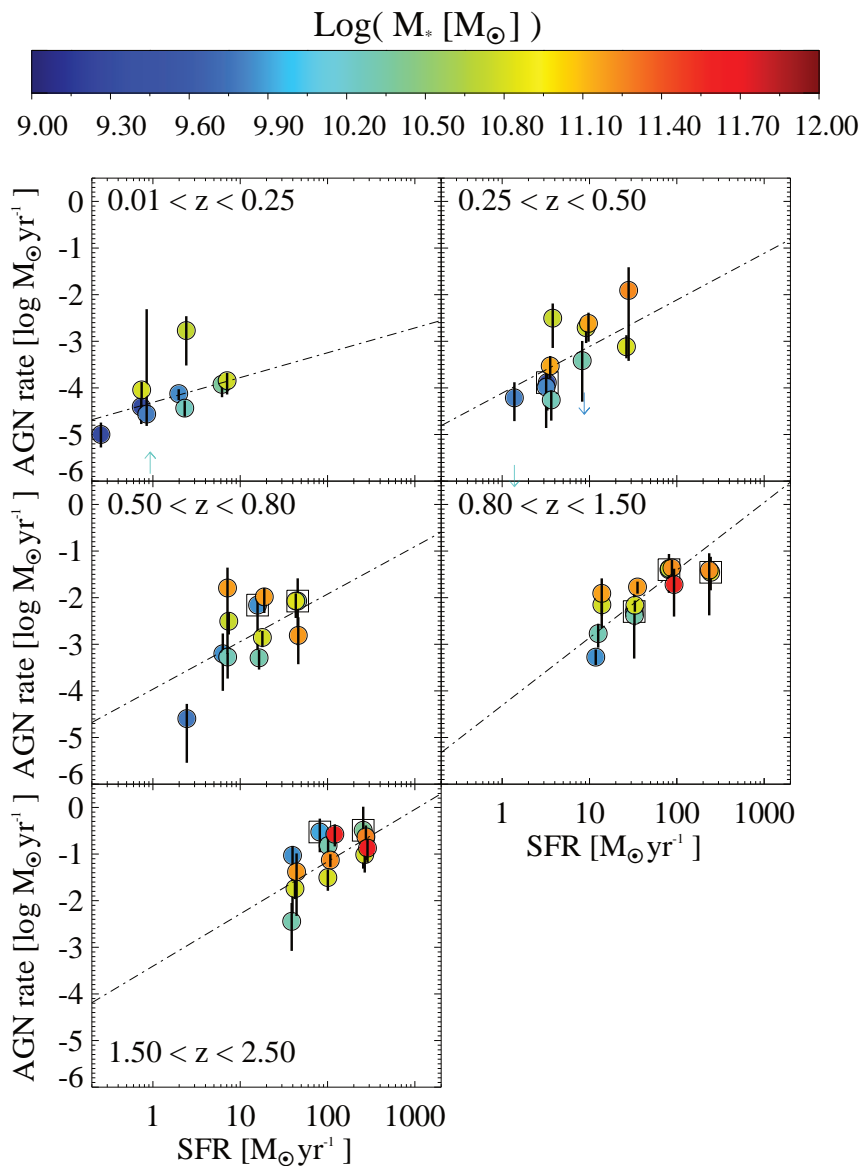
#### Spearman’s rank

To evaluate the significance of the observed trends, we used the Spearman’s rank correlation coefficient  $\rho$ , that indicates the strength of any observed correlation. However, the Spearman’s test considers the observed data points as “exact” and does not take into account their possible error bars. To estimate the most probable correlation coefficient given the error bars, we used the procedure detailed in Santos-Sanz et al. (2012) (see their Appendix B.2 for details). Briefly, they generated 1000 samples of data points, building each synthetic dataset from its associated Gaussian distribution function, where error bars correspond to one standard deviation. This Monte Carlo bootstrap analysis provides a distribution of Spearman’s rank correlation coefficients, where its median value gives the weighted Spearman’s coefficient that represents the most likely correlation to the observed data points. We used the IDL routine `R_CORRELATE.PRO` to derive the average value of Spearman’s  $\rho$ : in addition, this function provides the significance of its deviation from zero  $P(\text{correlation})$  of the observed trend, considering the number of points populating the correlation. These parameters are used to compare different relationships and infer the strength of the observed correlations. We applied this analysis when studying the relationship between average AGN accretion rate and SFR,  $M_{\star}$  and MS–offset in each redshift bin (see Table 5.1).

#### Correlation with SFR and $M_{\star}$

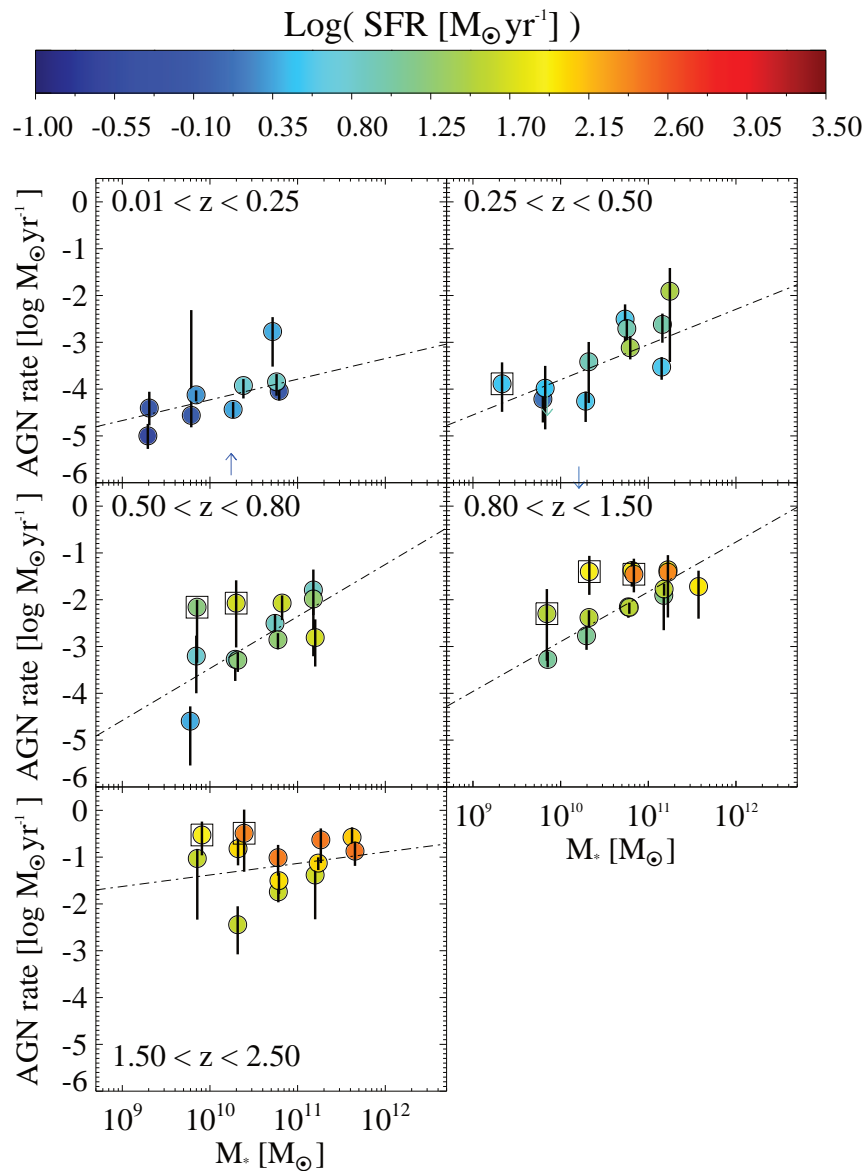
In Fig. 5.3, the relation between  $\langle \dot{M}_{\text{bh}} \rangle$  and SFR is shown in different redshift bins with colour coded  $M_{\star}$ . Circles mark the same bins already shown in Fig. 5.2, but re-arranged to plot the  $\langle \dot{M}_{\text{bh}} \rangle$  on the *y-axis*. Error bars of each bin mark the  $\pm 1\sigma$  uncertainties in  $\langle \dot{M}_{\text{bh}} \rangle$ , obtained by applying a Monte Carlo bootstrapping (see Section 5.4.4) to the obscuration-corrected  $L_X$  and rescaling to  $\langle \dot{M}_{\text{bh}} \rangle$  as described in Section 5.5.1. The black dashed dotted line in each *z*-bin represents the best-fit linear relation parametrized in Eq. 5.7, whose parameters are listed in Table 5.1. The best-fit slope ( $\alpha$ ) suggests that the correlation with

<sup>10</sup>As the adopted MS relation shows a linear trend at all redshifts, the offset from the MS becomes simply the ratio between the sSFR and that corresponding to the main-sequence sSFR<sub>ms</sub>, for a given bin of  $M_{\star}$  and redshift.



**Figure 5.3:** Average AGN accretion rate vs SFR, at different redshifts. The colour coded bar is the  $M_*$ , while individual black dashed-dotted lines set the best-fit linear relation inferred in each  $z$ -bin, whose coefficients are listed in Table 5.1. Bins representing off-sequence galaxies are enclosed in black squares.

SFR is consistent with a linear relation, at  $z > 0.25$ . Consistent trends are found in Fig. 5.4 as a function of  $M_*$ . In particular, we note that in both figures the best-fit relation at  $0.01 \leq z < 0.25$  is flatter than linear. This is likely due to the fact that the comoving volume covered by COSMOS is not large enough to detect ULIRGs hosting powerful Quasars. At  $z > 0.25$  both relationships show clear trends with nearly linear slopes and relatively high significance ( $> 2\sigma$  level). Focusing on individual redshift slices, the comparison between Spearman’s rank coefficients in Table 5.1 (columns 5 and 6) shows that the strength of the observed correlations with either SFR and  $M_*$  are comparable at  $z < 0.8$ , while at higher redshifts the correlation with SFR becomes progressively more significant. Especially at



**Figure 5.4:** Average AGN accretion rate vs  $M_*$ , at different redshifts. The colour coded bar is the SFR, while individual black dashed-dotted lines set the best-fit linear relation inferred in each  $z$ -bin, whose coefficients are listed in Table 5.1. Bins representing off-sequence galaxies are enclosed in black squares.

$z \sim 2$ , these coefficients significantly favour the correlation with SFR, showing a nearly flat trend between  $\langle \dot{M}_{\text{bh}} \rangle$  and  $M_*$ .

Another way of illustrating this point is by looking at the  $z > 0.5$  panels in Fig. 5.3 and 5.4: Fig. 5.4 shows that off-sequence galaxies (highlighted with black squares) have higher average BHAR, compared to mass-matched main-sequence galaxies. In contrast, at a given SFR in Fig. 5.3 the BHAR for off-sequence galaxies does not stand out of those for main-sequence galaxies. Overall, this may indicate that at  $z > 0.5$  the BHAR is primarily dependent on SFR.

Given the well known main-sequence relation between SFR and  $M_*$  at all redshifts

considered here, it is necessary to carry this further and simultaneously investigate the dependence of  $\langle \dot{M}_{\text{bh}} \rangle$  on both SFR and  $M_*$ . We addressed this issue by performing a partial-correlation fitting analysis. We used the IDL routine `R_CORRELATE.PRO` to evaluate the Spearman’s  $\rho$  related to each couple of parameters and then we combined them according to the following expression:

$$\rho_{\text{abc}} = \frac{\rho_{\text{ab}} - \rho_{\text{ac}}\rho_{\text{bc}}}{\sqrt{(1 - \rho_{\text{ac}}^2)(1 - \rho_{\text{bc}}^2)}} \quad (5.8)$$

which returns the partial correlation between  $a$  and  $b$  adjusted for  $c$ . Assuming that  $(a, b, c) = (\text{SFR}, \langle \dot{M}_{\text{bh}} \rangle, M_*)$ , then  $\rho_{\text{abc}}$  would represent the partial correlation between SFR and  $\langle \dot{M}_{\text{bh}} \rangle$ , removing the concomitant dependence on  $M_*$ . As done in Section 5.5.2, we performed again a Monte Carlo bootstrap analysis to provide a distribution of Spearman’s coefficients from partial-correlation analysis, taking the median value of that distribution as the most probable to reproduce our data<sup>11</sup>. The latter  $\rho$  values obtained from partial-correlation fitting (hereafter “partial  $\rho$ ’ ’), together with their corresponding significance levels, are listed in Table 5.1 (columns 7 and 8).

The comparison between partial  $\rho$  values and those obtained from a rank correlation of individual parameters shows similar trends. This suggests that the main-sequence relation between SFR and  $M_*$  does not significantly affect the correlations with the average BHAR. The observed trends show that a correlation with  $M_*$  is slightly preferable than with SFR at  $z < 0.5$ , while at higher redshifts the correlation with SFR becomes progressively more significant. However, we evaluated the difference between SFR and  $M_*$ , in terms of resulting partial-correlation coefficients, to be less than  $1\sigma$  at  $z < 1$  and of the order of  $2.2\sigma$  at  $1.5 < z < 2.5$ . This may suggest that for our current sampling of the SFR– $M_*$  plane, the overall evolution of the average BHAR is best represented through a joint dependence on both SFR and  $M_*$  at all redshifts. A difference is detected only in the last  $z$ -bin, where the trend with SFR is fairly preferable with respect to that with  $M_*$ .

We accomplished a multiple-correlation fitting in the log space to provide a simple analytic expression of the simultaneous dependence of  $\langle \dot{M}_{\text{bh}} \rangle$  on both SFR and  $M_*$  as a function of redshift. According to the following parametrization:

$$\log(\langle \dot{M}_{\text{bh}} \rangle) \propto \gamma \log(\text{SFR}) + \delta \log(M_*) \quad (5.9)$$

we defined  $\gamma$  and  $\delta$  as the slopes of SFR and  $M_*$ , respectively, and listed their values in Table 5.2 for each  $z$ -bin. At  $z < 1$ , the slopes related to SFR and  $M_*$  are always consistent (within  $1\sigma$  uncertainty), meaning that both parameters are required to trace the evolution of the average BHAR. As expected from our previous analyses, at  $z \sim 2$  the correlation with  $\langle \dot{M}_{\text{bh}} \rangle$  is primarily driven by the SFR.

Given that we are studying *Herschel*-selected galaxies, our sample is likely biased towards the most star forming galaxies at a given redshift. Since this selection effect could

<sup>11</sup>If taking the mean value instead of the median would change the resulting  $\rho$  by around 2–3%. Moreover, we note that, if the observed trend is poorly significant, the most probable  $\rho$  value might prefer a mildly negative rather than a mildly positive correlation (see Table 5.1), or viceversa.

**Table 5.2:** Best-fit parameters obtained in each  $z$ -bin through multiple-correlation fitting procedure.  $\gamma$  is the slope of SFR, while  $\delta$  refers to  $M_\star$ . Numbers in brackets represent their  $1\sigma$  uncertainties. The last column provides with the reduced  $\chi^2$  values returned from the multiple linear fitting.

$z$ -bin	$\gamma$ (SFR)	$\delta$ ( $M_\star$ )	$\chi_\nu^2$
$0.01 \leq z < 0.25$	0.38 ( $\pm 0.20$ )	0.31 ( $\pm 0.16$ )	1.90
$0.25 \leq z < 0.50$	0.78 ( $\pm 0.32$ )	0.52 ( $\pm 0.24$ )	1.78
$0.50 \leq z < 0.80$	0.48 ( $\pm 0.36$ )	0.92 ( $\pm 0.29$ )	1.93
$0.80 \leq z < 1.50$	0.85 ( $\pm 0.22$ )	0.73 ( $\pm 0.16$ )	0.89
$1.50 \leq z \leq 2.50$	1.16 ( $\pm 0.33$ )	-0.04 ( $\pm 0.18$ )	1.86

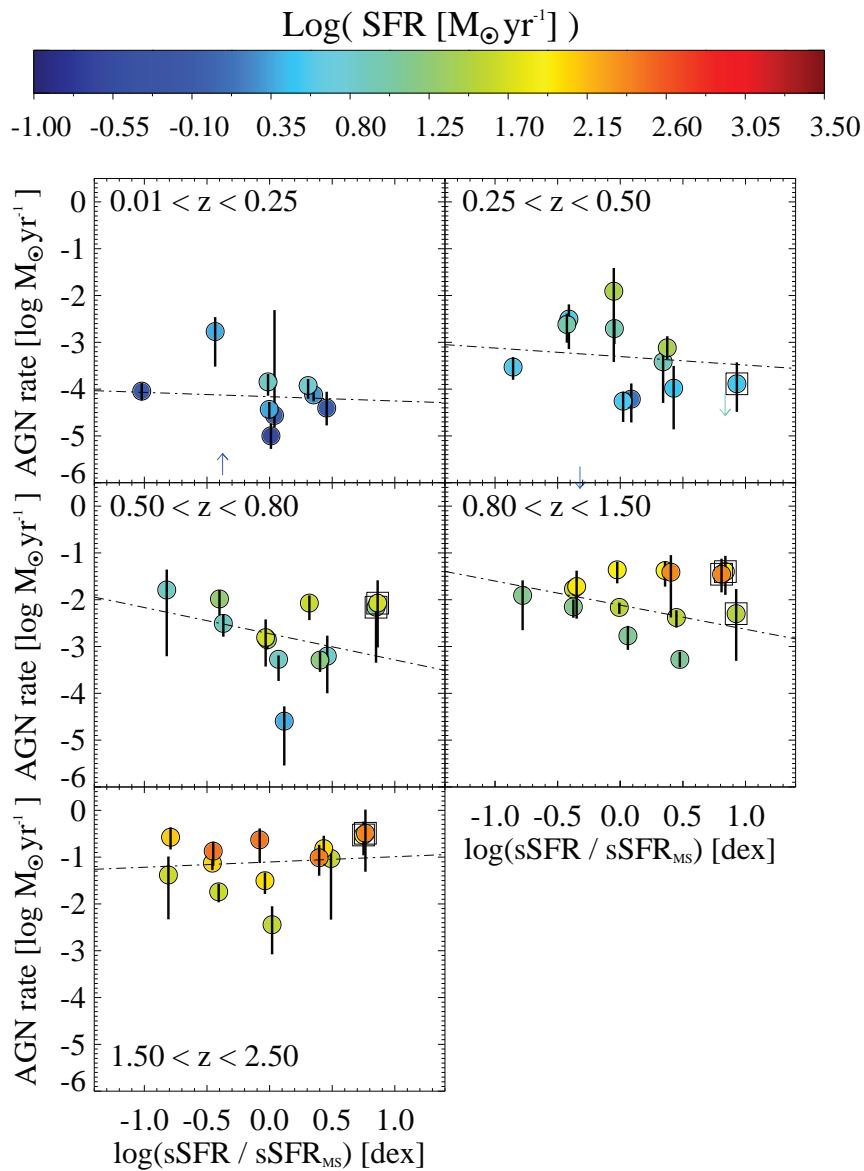
potentially bias our results, we tried to evaluate the implications that incompleteness effects might have on this study. We repeated the same analysis by limiting our parent sample to *Herschel* galaxies with FIR flux (in either PACS or SPIRE bands) larger than the flux corresponding to 80% completeness level. This cut strongly reduces our statistics (about a factor of three), as well as the range of galaxy properties we are allowed to probe, but it allows us to test the validity of our findings in a reasonably complete sample of star-forming galaxies. We accomplished the same analysis previously described and evaluated the strength of the observed correlations between average BHAR, SFR and  $M_\star$ , on the basis of a quite complete sample of star forming galaxies. We ended up with consistent trends (within 1 *sigma* uncertainty) in all redshift bins, thus the overall correlations between  $\langle \dot{M}_{\text{bh}} \rangle$  and galaxy properties are in agreement with our previous analysis. In addition, we checked that the observed trends obtained separately in GOODS fields and in COSMOS provide consistent results.

This general agreement might suggest that the relationships presented in this Section are safely valid among star forming galaxies in a wide range of SFRs and redshifts. A deeper investigation will be feasible through forthcoming *Chandra* observations of the entire COSMOS field (2 deg<sup>2</sup> instead of 0.9 deg<sup>2</sup>) from the COSMOS Legacy Survey (PI: F. Civano). In a future work, we plan to enrich the statistics and further increase the significance of the results presented in this work.

### Correlation with offset from the MS

The relation between  $\langle \dot{M}_{\text{bh}} \rangle$  and offset from the MS is shown in Fig. 5.5 and the best-fit parameters obtained through a linear regression fitting are listed in Table 5.1. It is evident that the average BHAR does not depend significantly on the MS offset. At all redshifts, the best-fit slope is consistent with a flat or slightly declining relation. Our analysis based on Spearman's  $\rho$  coefficients rejects with high probability any possible trend, suggesting a very weak, or absent, correlation at any redshift here considered<sup>12</sup>. This lack of correlation with MS offset might be likely induced by the fact that correlations between  $\langle \dot{M}_{\text{bh}} \rangle$  and SFR and  $M_\star$  show consistent slopes. At  $z \sim 2$ , given the stronger dependence of  $\langle \dot{M}_{\text{bh}} \rangle$  on the SFR, the resulting slope with MS offset is not negative, at odds with previous  $z$ -bins,

<sup>12</sup>We have verified that our results remain if using the MS definition of Whitaker et al. (2012) or Schreiber et al. (2014), instead of the one proposed by Elbaz et al. (2011).



**Figure 5.5:** Average AGN accretion rate vs offset from the MS, at different redshifts. The colour coded bar is the SFR, while individual black dashed-dotted lines set the best-fit linear relation inferred in each  $z$ -bin, whose coefficients are listed in Table 5.1. Bins representing off-sequence galaxies are enclosed in black squares.

though it remains very weak.

While we have argued above, on the basis of Fig. 5.4, that at a given  $M_*$  high sSFR (above MS) objects tend to show larger BHAR, this is washed out in Fig 5.5 by the significant  $M_*$  range covered in each redshift bin, in combination with the MS relation and the link of BHAR to SFR (and  $M_*$ ). In addition, we have looked into possible trends of BHAR/SFR with MS offset and found no significant one at any redshift.

The top panel of Fig. 5.6 shows the distribution of *Herschel* galaxies (both detected and undetected in X-rays) in bins of SFR,  $M_*$  and  $\langle \dot{M}_{\text{bh}} \rangle$ , with a  $\langle \dot{M}_{\text{bh}} \rangle$  colour coding, at  $0.01 \leq z \leq 2.50$ . To mitigate the selection effects in SFR, here we limit ourselves to the



subsample with  $>80\%$  completeness, which provides a reasonable trade-off between number of sources and completeness of the sample. The central and bottom panels show the projections of the previous plot on the  $\text{SFR}-\langle\dot{M}_{\text{bh}}\rangle$  plane and  $M_{\star}-\langle\dot{M}_{\text{bh}}\rangle$  plane, respectively, with a redshift colour coding. The observed correlations shown in these panels suggest that the trend with SFR is steeper and more significant than with  $M_{\star}$ , which may support a primary dependence of the BHAR on SFR, though almost indiscernible from  $M_{\star}$  within single redshift slices. We also looked into possible trends of the BHAR/SFR with offset from the MS and found no significant one at any redshift. In Section 5.6 we discuss this lack of trends in the framework of classical merger scenarios.

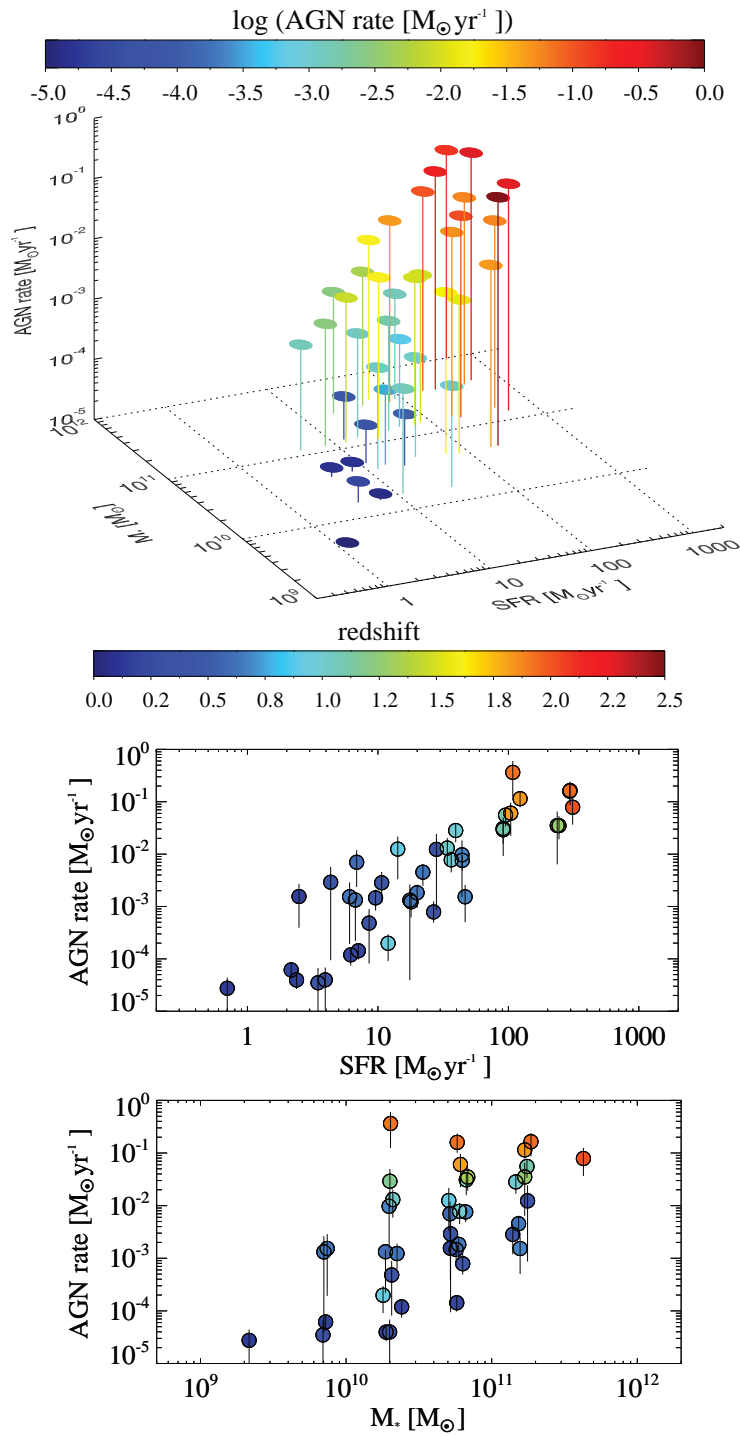
### 5.5.3 Comparison with previous studies

We compare the relationships between AGN accretion and galaxy properties obtained from our analysis with those found by Mullaney et al. (2012a, M12 hereafter) and Chen et al. (2013, C13 hereafter). M12 claimed a linear relation at  $z\sim 2$  between  $\langle\dot{M}_{\text{bh}}\rangle$  and  $M_{\star}$ , while our Fig. 5.4 shows a weak trend. We checked whether the lack of trend may be due to our sample selection. Our parent sample has been cut to match the selection criteria adopted by M12. We included main-sequence galaxies only (i.e.  $\text{sSFR} \leq 4 \times \text{sSFR}_{\text{ms}}$ ), we split the sample at  $1.5 < z < 2.5$  only in bins of  $M_{\star}$ , and used a constant bolometric correction of 22.4 (from Vasudevan & Fabian 2007) to derive the average BHAR. We found that  $\langle\dot{M}_{\text{bh}}\rangle \propto M_{\star}^{0.86 \pm 0.08}$ , consistently with what inferred by M12, who derived the relation  $\langle\dot{M}_{\text{bh}}\rangle \propto M_{\star}^{1.05 \text{ pm}0.36}$ . However, this nearly linear relation vanishes as we split the sample as a function of both SFR and  $M_{\star}$ : keeping the same M12-like sample and changing the grid configuration only, the resulting best-fit relation with  $M_{\star}$  is  $\langle\dot{M}_{\text{bh}}\rangle \propto M_{\star}^{0.45 \pm 0.25}$ .<sup>13</sup>

To study the simultaneous dependence of  $\langle\dot{M}_{\text{bh}}\rangle$  on both SFR and  $M_{\star}$ , we performed a multiple-correlation fitting in the log space, as a function of  $\langle\dot{M}_{\text{bh}}\rangle$ , SFR and  $M_{\star}$  at the same time. We used a M12-like sample, but splitting as a function of both SFR and  $M_{\star}$ . We found that  $\langle\dot{M}_{\text{bh}}\rangle \propto \text{SFR}^{1.04 \pm 0.29} \cdot M_{\star}^{0.13 \pm 0.18}$ . These slopes may suggest a “pure” SFR-driven relation. According to our analysis, the most relevant factor that would reconcile our analysis with that presented by M12 is the grid configuration. Covering the SFR– $M_{\star}$  plane in both dimensions allows us to separate the two variables which are degenerate in an MS-only analysis as in M12 (see Section 5.5.2).

We made also a comparison with results from C13, who measured a nearly linear trend between  $\langle\dot{M}_{\text{bh}}\rangle$  and SFR at  $0.25 < z < 0.80$ , splitting their *Herschel* sample in bins of SFR only. C13 found  $\log \text{BHAR} = (-3.72 \pm 0.52) + (1.05 \pm 0.33) \log \text{SFR}$  (both in  $M_{\odot} \text{ yr}^{-1}$ ). By adjusting our grid configuration in bins of SFR to match their analysis, our fitting routine gives consistent coefficients, with  $\text{BHAR} = (-3.65 \pm 0.12) + (1.18 \pm 0.11) \log \text{SFR}$ . We note that a similar trend is found even in case we split the sample in bins of SFR and  $M_{\star}$ .

<sup>13</sup>The latter trend is similar to what would be obtained in Fig. 5.5 (highest redshift bin) by removing the two bins corresponding to off-sequence galaxies (black squares).



**Figure 5.6:** *Top panel:* distribution of bins including *Herschel* galaxies above the  $>80\%$  completeness limit, as a function of AGN accretion rate, SFR and  $M_\star$  at all redshifts. The  $x$ ,  $y$  axes represent the SFR- $M_\star$  plane, while the  $z$ -axis sets the colour coded AGN accretion rate. Note that the three-dimensional space is logarithmic. *Middle panel:* previous plot projected on the SFR- $\langle \dot{M}_{\text{bh}} \rangle$  plane, with colour coded redshift. *Bottom panel:* projection on the  $M_\star$ - $\langle \dot{M}_{\text{bh}} \rangle$  plane, with colour coded redshift.

## 5.6 Discussion

In Section 5.5.2, we described the relationship between AGN accretion rate and SFR,  $M_\star$  and offset from the main-sequence, over a wide range of galaxy properties and cosmic epochs, from  $z\sim 0$  up to  $z=2.5$ . In this Section we discuss and interpret our findings in the context of current AGN/galaxy evolutionary scenarios.

Mapping the average AGN accretion along both SFR and  $M_\star$  allows us to separate the degeneracy between the two parameters, shown by the presence of the MS relation since  $z\sim 2$ . In Section 5.5.2 we found that the average BHAR depends on both SFR and  $M_\star$  with similar significance levels. A more clear correlation with SFR is observed at  $z>1$  and becomes more significant (at  $\sim 2.2\sigma$ ) at  $z\sim 2$ . Consistent trends are obtained if the  $>80\%$  complete *Herschel* sample is used. This finding may support the SFR as the original driver of the correlation with average AGN accretion. In particular, the correlation between  $\langle \dot{M}_{\text{bh}} \rangle$  and  $M_\star$  inferred in the highest  $z$ -bin shows a much flatter trend compared to the previous redshift slices, while the correlation with SFR is still significant (see Table 5.1).

We relate the trends of BHAR with SFR,  $M_\star$  and redshift to gas content and the redshift evolution of the gas fraction:

$$f_{\text{gas}} = \frac{M_{\text{gas}}}{M_{\text{gas}} + M_\star} \quad (5.10)$$

which represents the ratio between molecular gas mass and total baryonic mass of a galaxy. Several studies have investigated the evolution of  $f_{\text{gas}}$  in main-sequence galaxies from the local Universe to  $z\sim 1$  (Leroy et al. 2008; Geach et al. 2011; Saintonge et al. 2011a, 2011b, 2012), at  $z\sim 1.5\text{--}2$  (Daddi et al. 2010a; Tacconi et al. 2010, 2013), finding a strong ( $f_{\text{gas}} \propto (1+z)^2$ ) evolution up to  $z\sim 2$  and a plateau at higher redshift ( $z\sim 3$ , Magdis et al. 2013). Direct CO observations presented by Daddi et al. (2010a) and Tacconi et al. (2010, 2013) provided first empirical evidence for the existence of very large gas fractions in  $z\sim 1\text{--}2$  main-sequence galaxies, with mass in gas even larger than the mass in stars (i.e.  $f_{\text{gas}} > 0.5$ ). These gas rich systems have also higher SFRs (Daddi et al. 2010b; Genzel et al. 2010), as expected from the Schmidt–Kennicutt relation (Schmidt 1959; Kennicutt 1998). This gas-dominated phase in  $z\sim 1\text{--}2$  main-sequence galaxies is also reflected in a difference in morphology. Indeed, while local main-sequence galaxies are preferentially regular disks, their  $z\sim 2$  analogs show a larger fraction of irregular morphologies and/or clumpy disks (e.g. Elmegreen et al. 2007; Förster Schreiber et al. 2009; Kocevski et al. 2012), which make them potentially more efficient in funnelling the cold gas inward onto the SMBH.

In the light of the above-mentioned literature, it is plausible to expect a relationship between average AGN accretion rate and SFR, especially with increasing redshift, where the fraction of actively star forming gas becomes larger. For the same reason, given that  $z\sim 1\text{--}2$  galaxies are truly gas rich, especially those at lower  $M_\star$  (e.g. Santini et al. 2014), the build-up of their final  $M_\star$  is at a relatively early stage. As a consequence, a weaker correlation between AGN accretion and  $M_\star$ , as suggested by our data, may be reasonably expected.

In line with these arguments, Vito et al. (2014) found that AGN hosts are significantly more gas rich than inactive galaxies, at a given  $M_\star$  and redshift, suggesting that the

probability that a SMBH is active is strongly connected to the amount of cold gas supply. This supports  $M_{\text{gas}}$  as the key ingredient to explain the mutual evolution of star formation and AGN accretion activity (e.g. Santini et al. 2012; Mullaney et al. 2012a; Rosario et al. 2013; local AGN review by Heckman & Best 2014). However, the link between BHAR and SFR that we observe induces a more specific meaning to the notion that AGN accretion is more closely related to the amount of cold *star forming* gas ( $M_{\text{gas,SF}}$ ) rather than  $M_{\text{gas}}$ . Though in this work we are not able to distinguish between nuclear and global star formation, we note that if AGN accretion were tracing global cold gas mass, the known strong differences in star forming efficiency (or inverse “depletion time”) between normal galaxies and luminous starburst (e.g. Solomon et al. 1997; Saintonge et al. 2012) would predict a trend in BHAR/SFR with MS offset that we do not observe (Section 5.5.2). A plausible hypothesis that would reconcile the dependence of AGN accretion on either  $M_{\text{gas}}$  or  $M_{\text{gas,SF}}$ , is that a significant part of the galaxy star formation is on circumnuclear (<1kpc) scales, as suggested by Diamond-Stanic & Rieke 2012 in nearby Seyferts and also supported by signatures of increasing star-formation “compactness” found in non-local starbursts (Elbaz et al. 2011; Gilli et al. 2014).

As argued in Section 5.5.2, our analysis may support the SFR as the original driver of the correlation with BHAR, though not well discernible from  $M_\star$  within individual redshift slices. By combining all redshift bins together we are able to increase the statistics and infer more general trends between AGN accretion and galaxy properties. Fig. 5.6 shows the sub-sample of our *Herschel* sources with FIR flux larger than the 80% completeness threshold (see Section 5.5.2). The middle and bottom panels in Fig. 5.6 show that the overall correlation of BHAR with SFR is much narrower and significant than with  $M_\star$ . A linear regression fit returns a positive slope of  $1.6 \pm 0.1$ , suggesting that the observed BHAR vs SFR relation is steeper than linear, and that their ratio intrinsically evolves with redshift. As discussed in Madau & Dickinson 2014 (2014, see their Fig. 15), both the global SFR density (SFRD) and black hole accretion rate density (BHAD) peak around  $z \sim 2$ , declining towards the local Universe. Despite the differences seen in recent derivations of accretion histories (e.g. Shankar et al. 2009; Aird et al. 2010; Delvecchio et al. 2014; Ueda et al. 2014), all of them systematically predict a slightly faster decay from  $z \sim 2$  down to  $z \sim 0$ , compared to the cosmic evolution of SFH. By fitting the BHAD/SFRD ratio in the log–log space at  $0 < z < 2$ , assuming the SFH from Madau & Dickinson 2014 and an average BHAD evolution from the afore-mentioned literature, we obtain  $\text{BHAD} \propto \text{SFRD}^{1.4 \pm 0.2}$ . This relation is consistent with the slope evolution that we found between black hole and SFRs in our IR-selected sample:  $\text{BHAR} \propto \text{SFR}^{1.6 \pm 0.1}$  (see Fig. 5.6, middle panel), supporting a scenario where the SMBH growth follows a faster fading compared to their host galaxies, from  $z \sim 2$  to the present epoch. This redshift evolution is parametrized as  $\text{BHAR}/\text{SFR} \propto (1+z)^{0.9 \pm 0.3}$  that we found through a linear regression fit in the log–log space. By integrating our estimates of average BHAR and SFR (see middle panel of Fig. 5.6) over cosmic time yields to  $\text{BHAR}/\text{SFR} \approx 1/(3000 \pm 1500)$ , marginally consistent with the prediction from the local  $M_{\text{bh}}-M_{\text{bulge}}$  relation (around 1/1000, e.g. Magorrian et al. 1998). This slight difference has been found and extensively argued in various studies (e.g. Jahnke et al. 2009; Rafferty et al. 2011; M12; C13). One possible factor

is that our SFRs for main-sequence galaxies refer to the entire galaxy, including a large disk contribution, while only stars ultimately ending in the bulge should be counted in comparisons to the local  $M_{\text{bh}}-M_{\text{bulge}}$  relation. Second, since in the local Universe the black hole accretion is less radiatively efficient compared to non-local AGN (Merloni & Heinz 2008a), and the most powerful AGN are found in massive and passively evolving systems, a sample of radiatively efficient AGN might miss a not negligible fraction of the local BH accretion rate density, leading to a smaller BHAR/SFR ratio.

As described in Section 5.5.2, we did not find any evidence for a correlation between  $\langle \dot{M}_{\text{bh}} \rangle$  and MS offset, at any redshift. Since this non-trend remains valid also if selecting only sources with FIR flux above the corresponding  $>80\%$  completeness level (see Section 5.5.2), we argue that this result is not affected by selection biases. We interpret this lack of correlation as the fact the offset from the MS is not a good predictor of AGN accretion. It is worth mentioning that almost all the SFR/ $M_{\star}$  bins used in our analysis lie within a factor of 4 from the MS relation, so that galaxies on and above the MS relation cover different ranges in our *Herschel* sample. This might wash out a possible dependence of  $\langle \dot{M}_{\text{bh}} \rangle$  on the MS offset between MS and starburst galaxies.

Recent studies based on deep *Hubble Space Telescope* (HST) images, such as the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS), allowed thorough morphological analyses of star forming galaxies at various redshifts. No difference in morphology between AGN hosts and “inactive” galaxies has been found at  $0.7 < z < 3$  (Cisternas et al. 2011; Schawinski et al. 2011; Kocevski et al. 2011, 2012; Villforth et al. 2014), which trims down the cosmic relevance of major mergers and supports a less violent picture, where secular processes (e.g clumpy and/or unstable disks, Bournaud et al. 2012) play a major role in triggering both AGN and star formation activity. Independent observations carried out with *Herschel* (e.g. Santini et al. 2012; Mullaney et al. 2012b; Rosario et al. 2013) have shown that the FIR-based sSFRs of moderately luminous ( $L_{\text{x}} < 10^{44}$  erg s $^{-1}$ ) X-ray selected AGN hosts up to  $z \sim 3$  are almost indistinguishable from those of inactive galaxies. According to these studies, it is plausible to expect a non-trend between MS offset and AGN activity in our *Herschel*-selected sample. We would have expected to observe a more clear correlation in local luminous AGN hosts, which are mostly missed in our study due to the relatively small comoving volumes covered by COSMOS at low redshift.

Another argument discussed in Section 5.5.2 concerns how the BHAR/SFR ratio relates to major mergers. The overall picture is that both morphology and total star formation of AGN hosts do not show evidence for a dominant role of major galaxy mergers in AGN triggering, of the type assumed by classical merger scenarios for the triggering of (luminous) AGN (Sanders et al. 1988; Farrah et al. 2001; Springel et al. 2005; Hopkins et al. 2006; Sijacki et al. 2007; Di Matteo et al. 2008). Still, it is relevant to ask whether above MS starbursts stand out in the BH accretion as discussed in this work. Section 5.5.2 has argued that the MS offset is not a good predictor of absolute BHAR. We have also found that, at fixed  $M_{\star}$ , the enhanced SFRs of above MS galaxies go along with enhanced BHAR, in agreement with a BHAR/SFR ratio that does not change with position with respect to the main-sequence.

In other words, our results suggest that accretion onto the black hole and star forma-

tion broadly and on average trace each other, irrespective of whether the galaxy is evolving steadily on the MS or bursting. This picture supports a causal connection between (radiatively efficient) AGN activity and amount of cold star-forming gas.

## Concluding remarks and future prospects

In this Thesis we have investigated the connection between the physical properties of AGN (e.g. bolometric luminosities) and those of their host galaxies (e.g. SFR,  $M_*$ ). In Chapter 1, the physical and observational properties of AGN have been discussed from a panchromatic perspective, recalling also the AGN/galaxy evolutionary scenarios before the advent of *Herschel*. The complementary perspective of *Herschel* observations in the context of multi-wavelength studies of AGN host galaxies has been emphasized in Chapter 2, where we introduce and motivate the need for *Herschel* data to constrain emission processes related to AGN accretion and star formation activity in galaxy SEDs.

We exploited *Herschel* imaging data in the far-IR/sub-mm regime obtained from PACS and SPIRE observations, complemented with broad band ancillary photometry down to optical/UV wavelengths. This panchromatic approach was essential in characterising the global SEDs of *Herschel*-selected galaxies in the GOODS-S, GOODS-N and COSMOS fields. We performed a SED-fitting decomposition analysis to decouple the AGN emission from the host galaxy light, as presented in Chapter 3. This technique has been applied with the final goal of studying the evolution of the IR-selected AGN population over the cosmic time (Chapter 4), as well as to investigate how AGN accretion is related to the host galaxy properties (Chapter 5).

In the following Sections, we present our conclusions, highlighting the main results and the future projects that will enable to extend the work accomplished in this Thesis in the context of AGN and galaxy evolution.

### 6.1 Decomposition of broad-band SEDs

In Chapter 3 we have described our analysis to decouple the AGN emission from that related to the host galaxy. One of the main advantages of SED-fitting decomposition is the applicability of this technique to a large variety of galaxies, from passive systems to powerful ULIRGs at different redshifts. The energy balance technique incorporated in the MAGPHYS code (da Cunha et al. 2008) enables to fit the observed datapoints with meaningful and self-consistent galaxy SEDs from the optical/UV to sub-mm wavelengths. We also used a modified version of MAGPHYS, where an additional AGN emission component is considered

in the fit (Berta et al. 2013b). This code allows to estimate the AGN bolometric luminosity in active galaxies, but also to infer more precise measurements of the host galaxy properties, such as SFR and  $M_*$ , correcting for the possible contamination due to the AGN. We also distinguish our sources between “active” and “inactive” galaxies, on the basis of a Fisher-test that measures the statistical significance of the additional AGN component with respect to the fit without it.

Here below we list our main results.

*(i)* The fraction of star-forming galaxies in the redshift range  $0.1 \leq z \leq 3.8$  harbouring an AGN (at a  $>99\%$  confidence level) is  $\approx 37\%$ . We study the dependence of this fraction (or AGN detection rate) as a function of redshift and total IR (rest 1-1000  $\mu\text{m}$ ) luminosity  $L_{1-1000}$ . At  $z > 0.3$  we find that the AGN detection rate rises as a function of the total IR luminosity, suggesting that the AGN activity at these redshifts is likely to take place in stronger star-forming systems ( $L_{1-1000} \sim 10^{12}L_{\odot}$ ), rather than in “normal” ones ( $L_{1-1000} \sim 10^{11}L_{\odot}$ ). Furthermore, we find an additional dependence of the AGN detection rate on redshift, since the probability to pick up AGN in galaxies with a given luminosity increases with redshift, at least up to  $z \sim 2$ . This positive evolution as a function of both  $L_{1-1000}$  and redshift is in agreement with independent estimates from the literature (e.g. Imanishi et al. 2010; Alonso-Herrero et al. 2012).

*(ii)* We compare our classification based on SED decomposition with the mid-IR colour-colour selection criteria presented by Lacy et al. (2007) and Donley et al. (2012). The latter is more conservative and therefore more reliable than the former, though less complete at low luminosities. We find that  $\sim 12\%$  of *Herschel* galaxies classified as “AGN” fulfill the Donley et al. 2012 criterion and are characterised by AGN dominated systems in the mid-IR, with extreme ( $> 10^{12}L_{\odot}$ ) star formation and AGN accretion luminosities. On the other hand,  $\sim 70\%$  of our AGN sample lie inside the Lacy et al. 2007 wedge, showing more modest IR luminosities  $L_{1-1000} \sim 10^{11-11.5}L_{\odot}$ . Interestingly, we found our samples of AGN and galaxies to occupy different regions across the mid-IR colour-colour diagram, suggesting that our SED classification effectively labels as “AGN” sources with near/mid-IR SEDs systematically different from those of purely star forming galaxies.

*(iii)* The average (1-1000  $\mu\text{m}$ ) bolometric correction derived from each best-fit AGN template is around 3–5, therefore significantly lower than what derived from X-rays ( $\sim 20$ –30). Estimates of the AGN bolometric luminosity  $L_{\text{bol,AGN}}$  have been obtained through a Bayesian approach. The robustness of our measurements depends on the reliability of the SED-fitting decomposition itself. The comparison with a small (i.e. 14 sources) sample of highly obscured, Compton-Thick AGN candidates selected through X-ray spectral analysis (Lanzuisi et al. 2014), suggests that our estimates of  $L_{\text{bol,AGN}}$  inferred from SED-fitting decomposition agree with those derived from the X-rays, even for these extreme AGN.

Our analysis demonstrates that the combined *Herschel* and *Spitzer* capabilities of measuring the mid-to-far IR part of the spectrum are essential to constrain the possible presence of an AGN. Of course, this technique would largely benefit from *Spitzer*-IRS spectra, which



are currently available only for  $\sim$  a hundred of sources in the non-local Universe (e.g. Fadda et al. 2010; Pozzi et al. 2012).

Following on this reasoning, we stress that SED-fitting decomposition does not provide 100% complete and/or reliable AGN samples. As already stated in Section 3.6, the comparison with independent AGN diagnostics suggests that the reliability of SED-fitting decomposition in identifying AGN is dependent on the AGN-to-host luminosity ratio.

On the other hand, this method is almost unbiased towards nuclear and galaxy attenuation, and therefore enables to identify also heavily obscured AGN, missed even by deep X-ray surveys. Moreover, the large fraction of heavily obscured AGN detected by *Herschel* (60% instead of  $\sim 20\%$  found for the full X-ray sample) suggests that broad band SED-fitting including *Herschel* data offers a complementary perspective to studies conducted at other wavelengths hunting for this elusive AGN population.

## 6.2 AGN accretion history since $z \sim 3$

The predicting power of SED-fitting decomposition in evaluating the AGN bolometric luminosity  $L_{\text{bol,AGN}}$  enables to infer the cosmic evolution of the comoving luminosity density of AGN. We have considered a sample of PACS-160  $\mu\text{m}$  selected galaxies in the COSMOS and GOODS-S fields. The main purpose was to derive the AGN bolometric luminosity function since  $z \sim 3$  and compare our derivation with previous findings from the literature.

Here we report the main results achieved in Chapter 4.

(i) The best-fit AGN bolometric LF derived from our *Herschel* sample is in good agreement with that obtained by Hopkins et al. (2007) up to  $z \sim 3$ , despite the different and independent AGN selection criteria (Hopkins et al. 2007 exploited X-ray, optical and mid-IR selected samples of AGN). We stress that a correction for incompleteness has been applied to account for the fraction of AGN missed by *Herschel*, following the approach described in detail in Appendix A.

(ii) We analysed the comoving number density of *Herschel* sources showing a significant AGN component in their best-fit (at  $>99\%$  confidence level, on the basis of an F-test), as a function of redshift and AGN bolometric luminosity. We found the observed trends to resemble the “downsizing” behaviour. Indeed, Fig. 4.4 suggests that the bulk of the SMBH growth takes place in more powerful objects at earlier times. The similarity between our findings and those previously suggested by X-ray based works (e.g. Hasinger et al. 2005; Silverman et al. 2008) may imply that X-ray selected and *Herschel*-selected AGN hosts have likely experienced similar evolutionary paths.

(iii) By integrating the AGN bolometric luminosity function, we estimate the black hole accretion density (BHAD) at different cosmic epochs. Applying the Soltan’s argument to match the local BH mass density allows us to infer an estimate of the BH radiative efficiency  $\epsilon_{\text{rad}}$ . We found that both the redshift evolution of the BHAD and the resulting  $\epsilon_{\text{rad}}$  ( $\approx 0.07$ ) are in agreement with those obtained from X-ray based results (e.g. Merloni & Heinz 2008b; Ueda et al. 2014). This overall consistency suggests that the *Herschel*-based

selection, together with a proper SED decomposition, enables to trace consistent integrated luminosity densities with respect to X-ray selected AGN samples. This similarity does not necessarily imply that both selection criteria trace the evolution of the same AGN population, given the small overlap between IR and X-ray selected samples (i.e. 10–20% from both selections). The respective corrections for incompleteness account for the AGN population that is missed by the selection, which is necessary to reproduce the “complete” AGN accretion history.

### 6.3 The connection between AGN accretion and galaxy properties

We analysed the *average* AGN accretion rate (BHAR) in the SFR– $M_{\star}$  plane at  $0 < z \leq 2.5$  and investigated for the first time the mutual relations between AGN accretion rate and SFR,  $M_{\star}$  and specific SFR in about 8600 *Herschel*-selected galaxies taken from the GOODS-S, GOODS-N and COSMOS fields. The large statistics, together with the wealth of multi-wavelength data available in these fields, allowed us to perform broad-band SED-fitting decomposition and to measure SFR and  $M_{\star}$ , both corrected for a possible contamination due to the AGN.

Average AGN accretion rates have been derived from *Chandra* X-ray observations, considering all the *Herschel* galaxies within each bin of SFR,  $M_{\star}$  and redshift (both X-ray detected and undetected), and applying reasonable assumptions and widely used scaling factors.

The analysis of the relationships between AGN accretion rate and integrated galaxy properties has provided the following results.

(i) The average AGN accretion rate inferred in *Herschel*-selected galaxies shows positive evolution as a function of both SFR and  $M_{\star}$  at  $z < 1$ , while at higher redshift ( $1.5 < z \leq 2.5$ ) our data establish with  $>2\sigma$  significance the SFR as the best tracer of AGN accretion. This may suggest that the galaxy SFR is the original driver of the correlation with accretion onto the central black hole. A possible (but not unique) interpretation is that the observed trends reflect the evolution with redshift of the molecular gas fraction in star forming galaxies. It is now well established (e.g. Tacconi et al. 2013) that higher redshift galaxies are significantly richer in gas, compared to their local analogs at the same  $M_{\star}$ . This implies that at higher redshift, an increasing fraction of the total cold gas mass has not yet been converted into stars. Since AGN accretion occurs preferentially in gas rich galaxies (e.g. Vito et al. 2014), which are also more star forming (assuming a Schmidt-Kennicutt relation), this naturally explains at the same time the weaker correlation of AGN accretion with  $M_{\star}$  at  $z \sim 2$ , and the stronger correlation with SFR found at this epoch. Despite the lack of physical constraints to quantitatively test this hypothesis, we suggest that the key ingredient operating behind the scenes is the cold molecular gas mass.

(ii) Given a link between BHAR and SFR, we found that the BHAR/SFR ratio slightly evolves with redshift, suggesting a faster decline of the cosmic black hole accretion history

with respect to the star formation history from  $z \sim 2$  to recent epochs, in agreement with other derivations (e.g. Fig. 15 of Madau & Dickinson 2014). Integrating these evolutionary trends down to the local Universe leads to a mildly smaller BHAR/SFR ratio compared to the predictions of the local Magorrian et al. (1998) relation. Different interpretations could explain this apparent discrepancy (Rafferty et al. 2011): for instance, AGN that are not included in our sample (e.g. radiatively inefficient) would produce a lower BHAR/SFR ratio. A lower value of the BHAR/SFR might come out also if a significant fraction of the total SFR of *Herschel* galaxies occurs outside the galaxy bulge.

(iii) We compared the observed correlations between AGN accretion and key galaxy properties with those presented by Mullaney et al. (2012a) and Chen et al. (2013). Our analysis suggests that the correlation between AGN accretion rate and  $M_\star$  claimed at  $z \sim 2$  by Mullaney et al. 2012a is likely a consequence of the trend with SFR and of the main-sequence relation that holds between the two.

(iv) We investigated the observed correlations between AGN accretion rate and offset from the star-forming main-sequence. We found that, at fixed  $M_\star$ , the enhanced SFR of above MS galaxies goes along with enhanced BHAR, in agreement with a BHAR/SFR ratio that does not change with position with respect to the main-sequence. Our interpretation is that accretion onto the central black hole and star formation broadly trace each other, irrespective of whether the galaxy is evolving steadily on the main-sequence or bursting.

## 6.4 Future prospects

The SED-fitting decomposition technique can provide several crucial information for statistical samples of AGN and galaxies. To constrain for individual galaxies the major physical processes that regulate in the IR domain galaxy star formation and AGN activity, IR spectroscopy is mandatory. First attempts to improve SED-fitting decomposition have been done using *Spitzer*-IRS spectra in the mid-IR regime (e.g. Feltre et al. 2013), though these are available only in  $\sim$  a hundred of non-local galaxies.

The IR space telescope SPICA (Swinyard et al. 2009; Nakagawa et al. 2014), a mission that will be submitted in response to the incoming ESA M5 call, should enable to conduct systematic spectroscopic studies in the mid-IR and far-IR regime for thousands of galaxies up to  $z \gtrsim 3$ . In particular, the *SpicA FAR-infrared Instrument* (SAFARI), the far-IR (30–210  $\mu\text{m}$ ) imaging spectrometer onboard SPICA, would be crucial to detect fine-structure lines (e.g. [O I]63  $\mu\text{m}$ , [O IV]26  $\mu\text{m}$ , [C II]158  $\mu\text{m}$ ), which are powerful tracers of the physical conditions of the obscuring medium, therefore essential to characterise the ionizing radiation field ascribed to star formation processes and to a possible AGN.

The next generation X-ray observatory ATHENA+ (Nandra et al. 2013) will provide new constraints on the formation and evolution of SMBHs. The broad overlap in terms of science topics, and the complementary perspectives that Athena+ and SPICA will enable to achieve, will greatly improve our current understanding of AGN and galaxy evolution, paving the way to take a full census of the AGN accretion in the Universe.

Constraining the library of AGN templates, such as the assumptions on the torus size and geometry (i.e. smooth or clumpy), would significantly improve of our SED-fitting decomposition. Interferometric studies at mid-IR wavelengths tend to prefer a clumpy, compact structure (e.g. Tristram & Schartmann 2011, 2014), though not well discernible from currently available *Spitzer* or *Herschel* data (e.g. Feltre et al. 2012). An interesting prediction from clumpy models is that the dusty torus should show the same morphology across the far-IR and sub-mm regime, while smooth configurations predict a monotonic decrease of the dust temperature with distance from the SMBH. Only sub-mm interferometric studies with ALMA will allow us to infer precise measurements of the size and geometry of the dusty torus in type-1 and type-2 AGN (Maiolino 2008).

In the next future, we will significantly extend our investigation of the AGN accretion vs galaxy properties, by combining our sample with state-of-the-art radio sky surveys (e.g. JVLA-COSMOS, PI: V. Smolcic). This will allow us to constrain the stellar and SMBH growth over the cosmic time, exploring the role of AGN-driven feedback in galaxy evolution and the impact to the cosmic star formation history. These studies will also pave the way for the next generation (>2024) radio interferometer Square Kilometer Array (SKA), that will reach flux limits orders of magnitudes deeper than current radio instruments, thus probing the AGN and galaxy evolution taking advantage of a deep synergy with above-mentioned future space missions at other wavelengths.

# Bibliography

- Ahn C. P. et al., 2014, *ApJS*, 211, 17
- Aird J. et al., 2012, *ApJ*, 746, 90
- Aird J. et al., 2010, *MNRAS*, 401, 2531
- Akylas A., Georgantopoulos I., Georgakakis A., Kitsionas S., Hatziminaoglou E., 2006, *A&A*, 459, 693
- Alexander D. M. et al., 2003, *AJ*, 126, 539
- Alexander D. M., Bauer F. E., Chapman S. C., Smail I., Blain A. W., Brandt W. N., Ivison R. J., 2005, *ApJ*, 632, 736
- Alexander D. M., Hickox R. C., 2012, *New A Rev.*, 56, 93
- Allamandola L. J., Tielens A. G. G. M., Barker J. R., 1989, *ApJS*, 71, 733
- Alonso-Herrero A., Pereira-Santaella M., Rieke G. H., Rigopoulou D., 2012, *ApJ*, 744, 2
- Alonso-Herrero A., Pérez-González P. G., Rieke G. H., Alexander D. M., Rigby J. R., Papovich C., Donley J. L., Rigopoulou D., 2008, *ApJ*, 677, 127
- Antonucci R. R. J., 1984, *ApJ*, 278, 499
- Antonucci R. R. J., Miller J. S., 1985, *ApJ*, 297, 621
- Arnaud K. A., 1996, in *Astronomical Society of the Pacific Conference Series*, Vol. 101, *Astronomical Data Analysis Software and Systems V*, Jacoby G. H., Barnes J., eds., p. 17
- Aussel H., Elbaz D., Cesarsky C. J., Starck J. L., 1999, in *ESA Special Publication*, Vol. 427, *The Universe as Seen by ISO*, Cox P., Kessler M., eds., p. 1023
- Avni Y., Bahcall J. N., 1980, *ApJ*, 235, 694
- Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
- Balestra I. et al., 2010, *A&A*, 512, A12

- Ball W. H., Tout C. A., Żytkow A. N., Eldridge J. J., 2011, MNRAS, 414, 2751
- 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., Wang W.-H., 2008, ApJ, 689, 687
- Barthel P. D., 1989, ApJ, 336, 606
- Barvainis R., 1987, ApJ, 320, 537
- Bauer F. E., Alexander D. M., Brandt W. N., Schneider D. P., Treister E., Hornschemeier A. E., Garmire G. P., 2004, AJ, 128, 2048
- Bauer F. E., Yan L., Sajina A., Alexander D. M., 2010, ApJ, 710, 212
- Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
- Bell E. F. et al., 2005, ApJ, 625, 23
- Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38
- Berta S. et al., 2013a, A&A, 555, L8
- Berta S. et al., 2013b, A&A, 551, A100
- Berta S. et al., 2010, A&A, 518, L30
- Berta S. et al., 2011, A&A, 532, A49
- Bevington P. R., Robinson D. K., 2003, Data reduction and error analysis for the physical sciences
- Bolatto A. D. et al., 2013, ArXiv e-prints
- Bongiorno A. et al., 2012, MNRAS, 427, 3103
- Bournaud F. et al., 2012, ApJ, 757, 81
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
- Boyle B. J., Terlevich R. J., 1998, MNRAS, 293, L49
- Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
- Brandt W. N., Hasinger G., 2005a, ARA&A, 43, 827
- Brandt W. N., Hasinger G., 2005b, ARA&A, 43, 827
- Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

- Brusa M. et al., 2009, *A&A*, 507, 1277
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Bundy K. et al., 2006, *ApJ*, 651, 120
- Bundy K. et al., 2008, *ApJ*, 681, 931
- Bunker A. J., Stanway E. R., Ellis R. S., McMahon R. G., McCarthy P. J., 2003, *MNRAS*, 342, L47
- Burgarella D. et al., 2013, *A&A*, 554, A70
- Burlon D., Ajello M., Greiner J., Comastri A., Merloni A., Gehrels N., 2011, *ApJ*, 728, 58
- Burtscher L. et al., 2013, *A&A*, 558, A149
- Capak P. et al., 2007, *ApJS*, 172, 99
- Cappelluti N. et al., 2009, *A&A*, 497, 635
- Caputi K. I. et al., 2007, *ApJ*, 660, 97
- Cardamone C. N. et al., 2008, *ApJ*, 680, 130
- Carilli C. L., Walter F., 2013, *ARA&A*, 51, 105
- Casey C. M. et al., 2012, *ApJ*, 761, 139
- Casey C. M., Chapman S. C., Smail I., Alaghband-Zadeh S., Bothwell M. S., Swinbank A. M., 2011, *MNRAS*, 411, 2739
- Casey C. M., Narayanan D., Cooray A., 2014, *Phys. Rep.*, 541, 45
- Cesarsky D., Lequeux J., Abergel A., Perault M., Palazzi E., Madden S., Tran D., 1996, *A&A*, 315, L309
- Chabrier G., 2003, *ApJ*, 586, L133
- Charlot S., Fall S. M., 2000, *ApJ*, 539, 718
- Chary R., Elbaz D., 2001, *ApJ*, 556, 562
- Chen C.-T. J. et al., 2013, *ApJ*, 773, 3
- Cicone C., Feruglio C., Maiolino R., Fiore F., Piconcelli E., Menci N., Aussel H., Sturm E., 2012, *A&A*, 543, A99
- Cicone C. et al., 2014a, *ArXiv e-prints*
- Cicone C. et al., 2014b, *A&A*, 562, A21
- Ciliegi P., Gruppioni C., McMahon R., Rowan-Robinson M., 2001, *Ap&SS*, 276, 957

- Ciotti L., Ostriker J. P., 2007, *ApJ*, 665, 1038
- Ciotti L., Ostriker J. P., Proga D., 2010, *ApJ*, 717, 708
- Cisternas M. et al., 2011, *ApJ*, 726, 57
- Civano F. et al., 2012, *ApJS*, 201, 30
- Coil A. L. et al., 2011, *ApJ*, 741, 8
- Coil A. L. et al., 2009, *ApJ*, 701, 1484
- Cooper M. C. et al., 2012, *MNRAS*, 425, 2116
- Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, *AJ*, 112, 839
- Cristiani S. et al., 2000, *A&A*, 359, 489
- Croom S. M., Smith R. J., Boyle B. J., Shanks T., Loaring N. S., Miller L., Lewis I. J., 2001, *MNRAS*, 322, L29
- Croton D. J. et al., 2006, *MNRAS*, 365, 11
- Cucciati O. et al., 2012, *A&A*, 539, A31
- Cusumano G., 2009, in *American Institute of Physics Conference Series*, Vol. 1126, American Institute of Physics Conference Series, Rodriguez J., Ferrando P., eds., pp. 104–107
- da Cunha E., Charlot S., Elbaz D., 2008, *MNRAS*, 388, 1595
- Daddi E. et al., 2010a, *ApJ*, 713, 686
- Daddi E., Cimatti A., Renzini A., Fontana A., Mignoli M., Pozzetti L., Tozzi P., Zamorani G., 2004, *ApJ*, 617, 746
- Daddi E. et al., 2007, *ApJ*, 670, 156
- Daddi E. et al., 2010b, *ApJ*, 714, L118
- Dale D. A., Helou G., 2002, *ApJ*, 576, 159
- de Graauw T. et al., 2010, *A&A*, 518, L6
- Decarli R., Falomo R., Treves A., Labita M., Kotilainen J. K., Scarpa R., 2010, *MNRAS*, 402, 2453
- Dekel A. et al., 2009, *Nature*, 457, 451
- Del Moro A. et al., 2013, *A&A*, 549, A59
- Della Ceca R. et al., 2008, *A&A*, 487, 119
- Delvecchio I. et al., 2014, *MNRAS*, 439, 2736



- Devecchi B., Volonteri M., 2009, *ApJ*, 694, 302
- Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, *ApJ*, 676, 33
- Di Matteo T., Springel V., Hernquist L., 2005, *Nature*, 433, 604
- Diamond-Stanic A. M., Rieke G. H., 2012, *ApJ*, 746, 168
- Dickinson M. et al., 2004, *ApJ*, 600, L99
- Diolaiti E., Bendinelli O., Bonaccini D., Close L., Currie D., Parmeggiani G., 2000, in *Astronomical Society of the Pacific Conference Series*, Vol. 216, *Astronomical Data Analysis Software and Systems IX*, Manset N., Veillet C., Crabtree D., eds., p. 623
- Doherty M., Bunker A. J., Ellis R. S., McCarthy P. J., 2005, *MNRAS*, 361, 525
- Domínguez Sánchez H. et al., 2014, *MNRAS*, 441, 2
- Donley J. L. et al., 2012, *ApJ*, 748, 142
- Draine B. T., Allaf-Akbari K., 2006, *ApJ*, 652, 1318
- Dunlop J. S., McLure R. J., Kukulka M. J., Baum S. A., O’Dea C. P., Hughes D. H., 2003, *MNRAS*, 340, 1095
- Efstathiou A., Rowan-Robinson M., 1995, *MNRAS*, 273, 649
- Elbaz D. et al., 2007, *A&A*, 468, 33
- Elbaz D. et al., 2011, *A&A*, 533, A119
- Elbaz D. et al., 2010, *A&A*, 518, L29
- Elitzur M., Shlosman I., 2006, *ApJ*, 648, L101
- Elmegreen D. M., Elmegreen B. G., Ravindranath S., Coe D. A., 2007, *ApJ*, 658, 763
- Elvis M. et al., 2009, *ApJS*, 184, 158
- Elvis M., Maccacaro T., Wilson A. S., Ward M. J., Penston M. V., Fosbury R. A. E., Perola G. C., 1978, *MNRAS*, 183, 129
- Elvis M. et al., 1994, *ApJS*, 95, 1
- Engel H. et al., 2010, *ApJ*, 724, 233
- Fabian A. C., Iwasawa K., 1999, *MNRAS*, 303, L34
- Fadda D. et al., 2010, *ApJ*, 719, 425
- Fanaroff B. L., Riley J. M., 1974, *MNRAS*, 167, 31P
- Fanidakis N. et al., 2012, *MNRAS*, 419, 2797

- Farrah D. et al., 2001, MNRAS, 326, 1333
- Farrah D. et al., 2012, ApJ, 745, 178
- Feltre A., Hatziminaoglou E., Fritz J., Franceschini A., 2012, MNRAS, 426, 120
- Feltre A. et al., 2013, MNRAS, 434, 2426
- Ferrarese L., 2002, ApJ, 578, 90
- Feruglio C., Maiolino R., Piconcelli E., Menci N., Aussel H., Lamastra A., Fiore F., 2010, A&A, 518, L155
- Fiore F. et al., 2003, A&A, 409, 79
- Fiore F. et al., 2008, ApJ, 672, 94
- Fiore F. et al., 2009, ApJ, 693, 447
- Flores H. et al., 1998, ArXiv Astrophysics e-prints
- Fontana A. et al., 2004, A&A, 424, 23
- Förster Schreiber N. M. et al., 2009, ApJ, 706, 1364
- Förster Schreiber N. M. et al., 2014, ApJ, 787, 38
- Förster Schreiber N. M., Shapley A. E., Erb D. K., Genzel R., Steidel C. C., Bouché N., Cresci G., Davies R., 2011, ApJ, 731, 65
- Franceschini A., Aussel H., Cesarsky C. J., Elbaz D., Fadda D., 2001, A&A, 378, 1
- Fritz J., Franceschini A., Hatziminaoglou E., 2006, MNRAS, 366, 767
- Gabor J. M., Davé R., Oppenheimer B. D., Finlator K., 2011, MNRAS, 417, 2676
- Gabor J. M. et al., 2009, ApJ, 691, 705
- Gaibler V., Khochfar S., Krause M., 2011, MNRAS, 411, 155
- Gao Y., Solomon P. M., 2004, ApJ, 606, 271
- García-Burillo S., Combes F., Schinnerer E., Boone F., Hunt L. K., 2005, A&A, 441, 1011
- Geach J. E., Smail I., Moran S. M., MacArthur L. A., Lagos C. d. P., Edge A. C., 2011, ApJ, 730, L19
- Gebhardt K. et al., 2000, ApJ, 539, L13
- Gehrels N., 1986, ApJ, 303, 336
- Genzel R., Cesarsky C. J., 2000, ARA&A, 38, 761
- Genzel R., Eckart A., Ott T., Eisenhauer F., 1997, MNRAS, 291, 219

- Genzel R. et al., 2010, MNRAS, 407, 2091
- Georgakakis A. et al., 2009, MNRAS, 397, 623
- Georgakakis A. et al., 2007, ApJ, 660, L15
- Georgakakis A. et al., 2008, MNRAS, 385, 2049
- Gerke B. F. et al., 2007, MNRAS, 376, 1425
- Gilli R., Comastri A., Hasinger G., 2007, A&A, 463, 79
- Gilli R. et al., 2014, A&A, 562, A67
- Gilli R., Vignali C., Mignoli M., Iwasawa K., Comastri A., Zamorani G., 2010, A&A, 519, A92
- Glikman E., Helfand D. J., White R. L., 2006, ApJ, 640, 579
- Granato G. L., Danese L., 1994, MNRAS, 268, 235
- Grazian A. et al., 2006, A&A, 449, 951
- Griffin M. J. et al., 2010, A&A, 518, L3
- Grogin N. A. et al., 2005, ApJ, 627, L97
- Gruppioni C., Lari C., Pozzi F., Zamorani G., Franceschini A., Oliver S., Rowan-Robinson M., Serjeant S., 2002, MNRAS, 335, 831
- Gruppioni C. et al., 2013, MNRAS, 432, 23
- Gültekin K., Cackett E. M., Miller J. M., Di Matteo T., Markoff S., Richstone D. O., 2009, ApJ, 706, 404
- Haiman Z., Loeb A., 2001, ApJ, 552, 459
- Harrison C. M. et al., 2012, ApJ, 760, L15
- Hasinger G., 2008, A&A, 490, 905
- Hasinger G. et al., 2007, ApJS, 172, 29
- Hasinger G., Miyaji T., Schmidt M., 2005, A&A, 441, 417
- Hatziminaoglou E. et al., 2008, MNRAS, 386, 1252
- Hatziminaoglou E., Fritz J., Jarrett T. H., 2009, MNRAS, 399, 1206
- Hatziminaoglou E. et al., 2010, A&A, 518, L33
- Heckman T., Best P., 2014, ArXiv e-prints

- Herbert P. D., Jarvis M. J., Willott C. J., McLure R. J., Mitchell E., Rawlings S., Hill G. J., Dunlop J. S., 2011, *MNRAS*, 410, 1360
- Hickox R. C. et al., 2009, *ApJ*, 696, 891
- Hickox R. C., Mullaney J. R., Alexander D. M., Chen C.-T. J., Civano F. M., Goulding A. D., Hainline K. N., 2014, *ApJ*, 782, 9
- Hine R. G., Longair M. S., 1979, *MNRAS*, 188, 111
- Hiroi K., Ueda Y., Akiyama M., Watson M. G., 2012, *ApJ*, 758, 49
- Hönig S. F., Kishimoto M., Gandhi P., Smette A., Asmus D., Duschl W., Polletta M., Weigelt G., 2010, *A&A*, 515, A23
- Hopkins P. F., Hernquist L., 2006, *ApJS*, 166, 1
- Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, *ApJS*, 163, 1
- Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, *ApJS*, 175, 356
- Hopkins P. F., Hickox R., Quataert E., Hernquist L., 2009, *MNRAS*, 398, 333
- Hopkins P. F., Kereš D., Murray N., Hernquist L., Narayanan D., Hayward C. C., 2013, *MNRAS*, 433, 78
- Hopkins P. F., Quataert E., 2010a, *MNRAS*, 407, 1529
- Hopkins P. F., Quataert E., 2010b, *MNRAS*, 407, 1529
- Hopkins P. F., Quataert E., 2011, *MNRAS*, 415, 1027
- Hopkins P. F., Richards G. T., Hernquist L., 2007, *ApJ*, 654, 731
- Hubeny I., Agol E., Blaes O., Krolik J. H., 2000, *ApJ*, 533, 710
- Ilbert O. et al., 2009, *ApJ*, 690, 1236
- Ilbert O. et al., 2010, *ApJ*, 709, 644
- Imanishi M., Nakagawa T., Shirahata M., Ohyama Y., Onaka T., 2010, *ApJ*, 721, 1233
- Iverson R. J. et al., 2010, *MNRAS*, 402, 245
- Iwasawa K. et al., 2012, *A&A*, 546, A84
- Jaffe W. et al., 2004, *Nature*, 429, 47
- Jahnke K. et al., 2009, *ApJ*, 706, L215
- Jahnke K., Macciò A. V., 2011, *ApJ*, 734, 92

- Juneau S. et al., 2013, *ApJ*, 764, 176
- Juneau S. et al., 2005, *ApJ*, 619, L135
- Kartalpe J. S. et al., 2010, *ApJ*, 721, 98
- Kellermann K. I., Sramek R., Schmidt M., Shaffer D. B., Green R., 1989, *AJ*, 98, 1195
- Kennicutt R. C., Evans N. J., 2012, *ARA&A*, 50, 531
- Kennicutt, Jr. R. C., 1998, *ApJ*, 498, 541
- Kessler M. F. et al., 1996, *A&A*, 315, L27
- Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35
- Kewley L. J., Geller M. J., Jansen R. A., 2004, *AJ*, 127, 2002
- Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, *MNRAS*, 372, 961
- Khachikian E. Y., Weedman D. W., 1974, *ApJ*, 192, 581
- Kishimoto M., Hönig S. F., Beckert T., Weigelt G., 2007, *A&A*, 476, 713
- Kocevski D. D. et al., 2012, *ApJ*, 744, 148
- Kocevski D. D. et al., 2011, *ApJ*, 736, 38
- Kormendy J., Kennicutt J. R. C., 2004, *ARA&A*, 42, 603
- Kurk J. et al., 2013, *A&A*, 549, A63
- La Franca F. et al., 2005, *ApJ*, 635, 864
- Lacy M., Petric A. O., Sajina A., Canalizo G., Storrie-Lombardi L. J., Armus L., Fadda D., Marleau F. R., 2007, *AJ*, 133, 186
- Lacy M. et al., 2004, *ApJS*, 154, 166
- Laing R. A., Jenkins C. R., Wall J. V., Unger S. W., 1994, in *Astronomical Society of the Pacific Conference Series*, Vol. 54, *The Physics of Active Galaxies*, Bicknell G. V., Dopita M. A., Quinn P. J., eds., p. 201
- Laird E. S., Nandra K., Pope A., Scott D., 2010, *MNRAS*, 401, 2763
- Lanzuisi G. et al., 2013, *MNRAS*, 431, 978
- Lanzuisi G. et al., 2014, *ArXiv e-prints*
- Lauer T. R. et al., 2007, *ApJ*, 662, 808
- Lawrence A., 1991, *MNRAS*, 252, 586
- Le Fèvre O. et al., 2005, *A&A*, 439, 845

- Le Flo'ch E. et al., 2009, ApJ, 703, 222
- Le Flo'ch E. et al., 2005, ApJ, 632, 169
- Lee N. et al., 2013, ApJ, 778, 131
- Leger A., Puget J. L., 1984, A&A, 137, L5
- Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782
- Lilly S. J. et al., 2009, ApJS, 184, 218
- Lilly S. J., Le Fèvre O., Hammer F., Crampton D., 1996, ApJ, 460, L1
- Lilly S. J. et al., 2007, ApJS, 172, 70
- Lodato G., Natarajan P., 2007, MNRAS, 377, L64
- Luo B. et al., 2008, ApJS, 179, 19
- Lusso E. et al., 2012, MNRAS, 425, 623
- Lusso E. et al., 2013, ApJ, 777, 86
- Lutz D., 2014, ARA&A, 52, 373
- Lutz D. et al., 2010, ApJ, 712, 1287
- Lutz D. et al., 2011, A&A, 532, A90
- Lutz D., Spoon H. W. W., Rigopoulou D., Moorwood A. F. M., Genzel R., 1998, ApJ, 505, L103
- Lutz D. et al., 2008, ApJ, 684, 853
- Lynden-Bell D., 1969, Nature, 223, 690
- Madau P., Dickinson M., 2014, ArXiv e-prints
- Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388
- Magdis G. E. et al., 2011, ApJ, 740, L15
- Magdis G. E. et al., 2013, A&A, 558, A136
- Magdis G. E., Rigopoulou D., Huang J.-S., Fazio G. G., 2010, MNRAS, 401, 1521
- Magliocchetti M. et al., 2014, MNRAS, 442, 682
- Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2009, A&A, 496, 57

- Magnelli B. et al., 2010, *A&A*, 518, L28
- Magnelli B. et al., 2014, *A&A*, 561, A86
- Magnelli B. et al., 2013, *A&A*, 553, A132
- Magorrian J. et al., 1998, *AJ*, 115, 2285
- Mainieri V. et al., 2007, *ApJS*, 172, 368
- Maiolino R., 2008, *New A Rev.*, 52, 339
- Malkan M. A., Sargent W. L. W., 1982, *ApJ*, 254, 22
- Mancini C. et al., 2011, *ApJ*, 743, 86
- Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, *MNRAS*, 351, 169
- Marshall H. L., 1985, *ApJ*, 299, 109
- Matt G. et al., 1999, *A&A*, 341, L39
- McLure R. J., Willott C. J., Jarvis M. J., Rawlings S., Hill G. J., Mitchell E., Dunlop J. S., Wold M., 2004, *MNRAS*, 351, 347
- Meisenheimer K. et al., 2007, *A&A*, 471, 453
- Menci N., Fiore F., Puccetti S., Cavaliere A., 2008, *ApJ*, 686, 219
- Merloni A. et al., 2010, *ApJ*, 708, 137
- Merloni A. et al., 2014, *MNRAS*, 437, 3550
- Merloni A., Heinz S., 2008a, *MNRAS*, 388, 1011
- Merloni A., Heinz S., 2008b, *MNRAS*, 388, 1011
- Merloni A., Heinz S., 2013, *Evolution of Active Galactic Nuclei*, Oswald T. D., Keel W. C., eds., p. 503
- Merloni A., Rudnick G., Di Matteo T., 2004, *MNRAS*, 354, L37
- Meurer G. R., Heckman T. M., Calzetti D., 1999, *ApJ*, 521, 64
- Mignoli M. et al., 2005, *A&A*, 437, 883
- Mignoli M. et al., 2013, *A&A*, 556, A29
- Mineo S., Gilfanov M., Lehmer B. D., Morrison G. E., Sunyaev R., 2014, *MNRAS*, 437, 1698
- Mineo S., Gilfanov M., Sunyaev R., 2012a, *MNRAS*, 419, 2095

- Mineo S., Gilfanov M., Sunyaev R., 2012b, MNRAS, 426, 1870
- Moorwood A. F. M., 1999, in ESA Special Publication, Vol. 427, The Universe as Seen by ISO, Cox P., Kessler M., eds., p. 825
- Mor R., Netzer H., 2012, MNRAS, 420, 526
- Morganti R., Holt J., Saripalli L., Oosterloo T. A., Tadhunter C. N., 2007, A&A, 476, 735
- Mortlock D. J. et al., 2011, Nature, 474, 616
- Mullaney J. R., Alexander D. M., Goulding A. D., Hickox R. C., 2011, MNRAS, 414, 1082
- Mullaney J. R. et al., 2012a, ApJ, 753, L30
- Mullaney J. R. et al., 2012b, MNRAS, 419, 95
- Nakagawa T., Shibai H., Onaka T., Matsuhara H., Kaneda H., Kawakatsu Y., Roelfsema P., 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9143, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 1
- Nandra K. et al., 2013, ArXiv e-prints
- Nandra K. et al., 2007, ApJ, 660, L11
- Nardini E., Risaliti G., Watabe Y., Salvati M., Sani E., 2010, MNRAS, 405, 2505
- Natarajan P., 2011, arxiv:1105.4902
- Nenkova M., Ivezić Ž., Elitzur M., 2002, ApJ, 570, L9
- Nenkova M., Sirocky M. M., Nikutta R., Ivezić Ž., Elitzur M., 2008, ApJ, 685, 160
- Netzer H., 2009, MNRAS, 399, 1907
- Netzer H., Lira P., Trakhtenbrot B., Shemmer O., Cury I., 2007, ApJ, 671, 1256
- Neugebauer G. et al., 1984, ApJ, 278, L1
- Nguyen H. T. et al., 2010, A&A, 518, L5
- NICASTRO F., Martocchia A., Matt G., 2003, ApJ, 589, L13
- Noeske K. G. et al., 2007, ApJ, 660, L43
- Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, A&A, 507, 1793
- Nordon R. et al., 2012, ApJ, 745, 182
- Nordon R. et al., 2010, A&A, 518, L24



- Novikov I. D., Thorne K. S., 1973, in *Black Holes (Les Astres Occlus)*, Dewitt C., Dewitt B. S., eds., pp. 343–450
- Oliver S. J. et al., 2012, *MNRAS*, 424, 1614
- Olsen K. P., Rasmussen J., Toft S., Zirm A. W., 2013, *ApJ*, 764, 4
- Page M. J. et al., 2012, *Nature*, 485, 213
- Papovich C. et al., 2007, *ApJ*, 668, 45
- Pérez-González P. G. et al., 2005, *ApJ*, 630, 82
- Peterson B. M., 1997, *An Introduction to Active Galactic Nuclei*
- Peterson B. M. et al., 2004, *ApJ*, 613, 682
- Piconcelli E., Jimenez-Bailón E., Guainazzi M., Schartel N., Rodríguez-Pascual P. M., Santos-Lleó M., 2005, *A&A*, 432, 15
- Pier E. A., Krolik J. H., 1992, *ApJ*, 401, 99
- Pierce C. M. et al., 2007, *ApJ*, 660, L19
- Pilbratt G. L. et al., 2010, *A&A*, 518, L1
- Poglitsch A. et al., 2010, *A&A*, 518, L2
- Polletta M. et al., 2007, *ApJ*, 663, 81
- Popesso P. et al., 2009, *A&A*, 494, 443
- Pozzi F. et al., 2007, *A&A*, 468, 603
- Pozzi F. et al., 2012, *MNRAS*, 423, 1909
- Puget J. L., Leger A., 1989, *ARA&A*, 27, 161
- Raban D., Jaffe W., Röttgering H., Meisenheimer K., Tristram K. R. W., 2009, *MNRAS*, 394, 1325
- Rafferty D. A., Brandt W. N., Alexander D. M., Xue Y. Q., Bauer F. E., Lehmer B. D., Luo B., Papovich C., 2011, *ApJ*, 742, 3
- Ranalli P., Comastri A., Setti G., 2003, *A&A*, 399, 39
- Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, *ApJS*, 175, 48
- Rees M. J., 1978, in *IAU Symposium, Vol. 77, Structure and Properties of Nearby Galaxies*, Berkhuijsen E. M., Wielebinski R., eds., pp. 237–242
- Rees M. J., 1984, *ARA&A*, 22, 471

- Richards G. T. et al., 2006, *ApJS*, 166, 470
- Richards G. T. et al., 2009, *ApJS*, 180, 67
- Risaliti G., Elvis M., 2004, in *Astrophysics and Space Science Library*, Vol. 308, Supermassive Black Holes in the Distant Universe, Barger A. J., ed., p. 187
- Rodighiero G. et al., 2011, *ApJ*, 739, L40
- Rola C. S., Terlevich E., Terlevich R. J., 1997, *MNRAS*, 289, 419
- Rosario D. J. et al., 2013, *ApJ*, 763, 59
- Rosario D. J. et al., 2012, *A&A*, 545, A45
- Roseboom I. G. et al., 2012, *MNRAS*, 419, 2758
- Roseboom I. G. et al., 2010, *MNRAS*, 409, 48
- Rovilos E. et al., 2012, *A&A*, 546, A58
- Rowan-Robinson M., Crawford J., 1989, *MNRAS*, 238, 523
- Saintonge A. et al., 2011a, *MNRAS*, 415, 32
- Saintonge A. et al., 2011b, *MNRAS*, 415, 61
- Saintonge A. et al., 2012, *ApJ*, 758, 73
- Sajina A., Spoon H., Yan L., Imanishi M., Fadda D., Elitzur M., 2009, *ApJ*, 703, 270
- Sajina A., Yan L., Fadda D., Dasyra K., Huynh M., 2012, *ApJ*, 757, 13
- Sakamoto K., Aalto S., Combes F., Evans A., Peck A., 2014, *ArXiv e-prints*
- Sales D. A., Pastoriza M. G., Riffel R., Winge C., Rodríguez-Ardila A., Carciofi A. C., 2011, *ApJ*, 738, 109
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Salpeter E. E., 1964, *ApJ*, 140, 796
- Salvato M. et al., 2011, *ApJ*, 742, 61
- 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, *ARA&A*, 34, 749
- Sanders D. B., Scoville N. Z., Young J. S., Soifer B. T., Schloerb F. P., Rice W. L., Danielson G. E., 1986, *ApJ*, 305, L45
- Sanders D. B., Soifer B. T., Elias J. H., Madore B. F., Matthews K., Neugebauer G., Scoville N. Z., 1988, *ApJ*, 325, 74

- Santini P. et al., 2009, *A&A*, 504, 751
- Santini P. et al., 2014, *A&A*, 562, A30
- Santini P. et al., 2012, *A&A*, 540, A109
- Santos-Sanz P. et al., 2012, *A&A*, 541, A92
- Sargent M. T., Béthermin M., Daddi E., Elbaz D., 2012, *ApJ*, 747, L31
- Saunders W., Rowan-Robinson M., Lawrence A., Efstathiou G., Kaiser N., Ellis R. S., Frenk C. S., 1990, *MNRAS*, 242, 318
- Sazonov S., Revnivtsev M., Krivonos R., Churazov E., Sunyaev R., 2007, *A&A*, 462, 57
- Schartmann M., Meisenheimer K., Camenzind M., Wolf S., Henning T., 2005, *A&A*, 437, 861
- Schartmann M., Meisenheimer K., Camenzind M., Wolf S., Tristram K. R. W., Henning T., 2008, *A&A*, 482, 67
- Schawinski K. et al., 2006, *Nature*, 442, 888
- Schawinski K. et al., 2009, *MNRAS*, 396, 818
- Schawinski K., Simmons B. D., Urry C. M., Treister E., Glikman E., 2012, *MNRAS*, 425, L61
- Schawinski K., Treister E., Urry C. M., Cardamone C. N., Simmons B., Yi S. K., 2011, *ApJ*, 727, L31
- Schinnerer E. et al., 2007, *ApJS*, 172, 46
- Schmidt M., 1959, *ApJ*, 129, 243
- Schmidt M., 1968, *ApJ*, 151, 393
- Schreiber C. et al., 2014, *ArXiv e-prints*
- Sellgren K., 1984, *ApJ*, 277, 623
- Seyfert C. K., 1943, *ApJ*, 97, 28
- Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337
- Shankar F., Salucci P., Granato G. L., De Zotti G., Danese L., 2004, *MNRAS*, 354, 1020
- Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, *ApJ*, 690, 20
- Shao L. et al., 2010, *A&A*, 518, L26
- Shen Y., Kelly B. C., 2010, *ApJ*, 713, 41

- Shi Y. et al., 2006, *ApJ*, 653, 127
- Shields G. A., 1978, *Nature*, 272, 706
- Siebenmorgen R., Haas M., Krügel E., Schulz B., 2005, *A&A*, 436, L5
- Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, *MNRAS*, 380, 877
- Silk J., 2005, *MNRAS*, 364, 1337
- Silk J., Di Cintio A., Dvorkin I., 2013, *ArXiv e-prints*
- Silk J., Norman C., 2009, *ApJ*, 700, 262
- Silva L., Granato G. L., Bressan A., Danese L., 1998, *ApJ*, 509, 103
- Silverman J. D. et al., 2008, *ApJ*, 679, 118
- Silverman J. D. et al., 2009, *ApJ*, 696, 396
- Simões Lopes R. D., Storchi-Bergmann T., de Fátima Saraiva M., Martini P., 2007, *ApJ*, 655, 718
- Simmons B. D., Urry C. M., Schawinski K., Cardamone C., Glikman E., 2012, *ApJ*, 761, 75
- Simpson C., 2005, *MNRAS*, 360, 565
- Smith D. J. B. et al., 2012, *MNRAS*, 427, 703
- Smolčić V. et al., 2009, *ApJ*, 696, 24
- Soifer B. T., Neugebauer G., Houck J. R., 1987, *ARA&A*, 25, 187
- Solomon P. M., Downes D., Radford S. J. E., Barrett J. W., 1997, *ApJ*, 478, 144
- Soltan A., 1982, *MNRAS*, 200, 115
- Springel V. et al., 2005, *Nature*, 435, 629
- Stanway E. R., Bunker A. J., McMahon R. G., Ellis R. S., Treu T., McCarthy P. J., 2004a, *ApJ*, 607, 704
- Stanway E. R. et al., 2004b, *ApJ*, 604, L13
- Steffen A. T., Barger A. J., Cowie L. L., Mushotzky R. F., Yang Y., 2003, *ApJ*, 596, L23
- Steffen A. T., Strateva I., Brandt W. N., Alexander D. M., Koekemoer A. M., Lehmer B. D., Schneider D. P., Vignali C., 2006, *AJ*, 131, 2826
- Stern D. et al., 2005, *ApJ*, 631, 163
- Storchi-Bergmann T., 2010, *ArXiv e-prints*

- Strateva I. et al., 2001, *AJ*, 122, 1861
- Strolger L.-G. et al., 2004, *ApJ*, 613, 200
- Sturm E. et al., 2011, *ApJ*, 733, L16
- Sutherland W., Saunders W., 1992, *MNRAS*, 259, 413
- Swain M. et al., 2003, *ApJ*, 596, L163
- Swinyard B. et al., 2009, *Experimental Astronomy*, 23, 193
- Symeonidis M. et al., 2014, *ArXiv e-prints*
- Symeonidis M. et al., 2013, *MNRAS*, 431, 2317
- Szokoly G. P. et al., 2004, *ApJS*, 155, 271
- Tacconi L. J. et al., 2010, *Nature*, 463, 781
- Tacconi L. J. et al., 2008, *ApJ*, 680, 246
- Tacconi L. J. et al., 2013, *ApJ*, 768, 74
- Tadhunter C., Morganti R., Rose M., Oonk J. B. R., Oosterloo T., 2014, *Nature*, 511, 440
- Tozzi P. et al., 2006, *A&A*, 451, 457
- Treister E., Krolik J. H., Dullemond C., 2008, *ApJ*, 679, 140
- Treister E., Schawinski K., Urry C. M., Simmons B. D., 2012, *ApJ*, 758, L39
- Treister E., Urry C. M., 2006, *ApJ*, 652, L79
- Treister E., Urry C. M., Virani S., 2009, *ApJ*, 696, 110
- Tristram K. R. W., Burtscher L., Jaffe W., Meisenheimer K., Hönig S. F., Kishimoto M., Schartmann M., Weigelt G., 2014, *A&A*, 563, A82
- Tristram K. R. W. et al., 2007, *A&A*, 474, 837
- Tristram K. R. W. et al., 2009, *A&A*, 502, 67
- Tristram K. R. W., Schartmann M., 2011, *A&A*, 531, A99
- Trump J. R. et al., 2009, *ApJ*, 696, 1195
- Trump J. R. et al., 2007, *ApJS*, 172, 383
- Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, *ApJ*, 786, 104
- Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, *ApJ*, 598, 886
- Urry C. M., Padovani P., 1995, *PASP*, 107, 803

- van Bemmell I. M., Dullemond C. P., 2003, *A&A*, 404, 1
- van der Wel A., Franx M., van Dokkum P. G., Rix H.-W., 2004, *ApJ*, 601, L5
- Vanzella E. et al., 2008, *A&A*, 478, 83
- Vasudevan R. V., Fabian A. C., 2007, *MNRAS*, 381, 1235
- Vasudevan R. V., Fabian A. C., 2009, *MNRAS*, 392, 1124
- Vattakunnel S. et al., 2012, *MNRAS*, 420, 2190
- Veilleux S. et al., 2013, *ApJ*, 776, 27
- Vignali C., Alexander D. M., Gilli R., Pozzi F., 2010, *MNRAS*, 404, 48
- Villforth C. et al., 2014, *MNRAS*, 439, 3342
- Vito F. et al., 2014, *MNRAS*, 441, 1059
- Vito F. et al., 2013, *MNRAS*, 428, 354
- Volonteri M., Begelman M. C., 2010, *MNRAS*, 409, 1022
- Volonteri M., Silk J., 2014, *ArXiv e-prints*
- Volonteri M., Stark D. P., 2011, *MNRAS*, 417, 2085
- Wall J. V., Pope A., Scott D., 2008, *MNRAS*, 383, 435
- Weiß A., Ivison R. J., Downes D., Walter F., Cirasuolo M., Menten K. M., 2009, *ApJ*, 705, L45
- Werner M. W. et al., 2004, *ApJS*, 154, 1
- Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, *ApJ*, 754, L29
- Wilson A. S., Baldwin J. A., Ulvestad J. S., 1985, *ApJ*, 291, 627
- Wisnioski E. et al., 2014, *ArXiv e-prints*
- Wittkowski M., Kervella P., Arsenault R., Paresce F., Beckert T., Weigelt G., 2004, *A&A*, 418, L39
- Wuyts S. et al., 2011a, *ApJ*, 738, 106
- Wuyts S. et al., 2011b, *ApJ*, 742, 96
- Xue Y. Q. et al., 2010, *ApJ*, 720, 368
- Xue Y. Q. et al., 2011, *ApJS*, 195, 10
- Yu Q., Tremaine S., 2002, *MNRAS*, 335, 965

Yuan T.-T., Kewley L. J., Sanders D. B., 2010, *ApJ*, 709, 884

Zamorani G. et al., 1981, *ApJ*, 245, 357

Zheng W., Kriss G. A., Telfer R. C., Grimes J. P., Davidsen A. F., 1997, *ApJ*, 475, 469

Zubovas K., Nayakshin S., Sazonov S., Sunyaev R., 2013, *MNRAS*, 431, 793





# Appendices



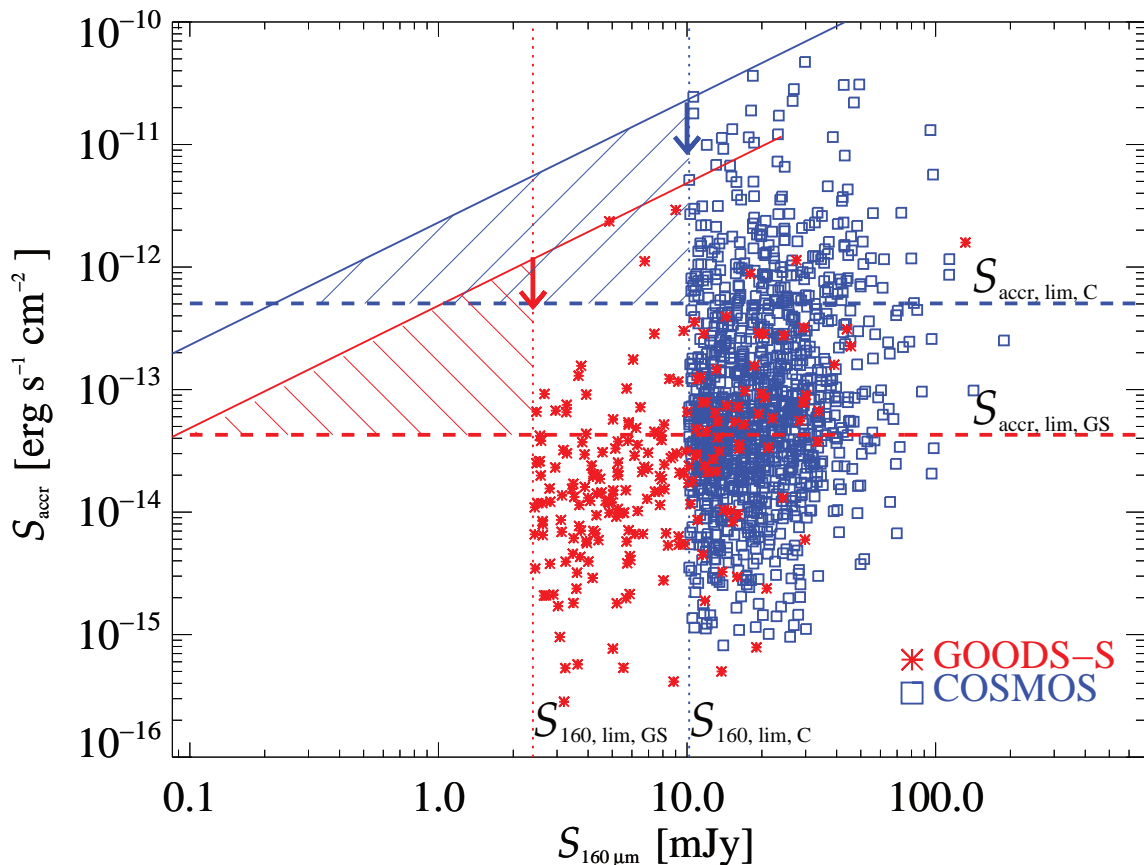
## Incompleteness in accretion luminosity

As mentioned in section 4.3.2, in this Appendix we describe in detail our approach to account for the incompleteness in accretion luminosity. We follow a similar approach to that developed by Fontana et al. (2004), who build the stellar mass function correcting for the incompleteness in stellar mass, starting from a  $K$ -band selected sample. The same argument could be in principle generalised, under some assumptions, to deal with the evaluation of the incompleteness affecting some unobservable quantity when selecting sources with any other observable quantity. We make use of a similar algorithm, but referred to the “accretion flux”  $S_{\text{accr}}$  as unobservable and to the  $160 \mu\text{m}$  flux as observable. We define  $S_{\text{accr}} = L_{\text{bol,AGN}}/4\pi D_L^2$  and compare it with the corresponding  $160 \mu\text{m}$  flux, simply rescaling the accretion power by its luminosity distance factor  $D_L$ .

As described in section 2.3, the reference sample was selected above a given detection threshold (2.4 mJy in the GOODS-S, 10.2 mJy in the COSMOS field). Nevertheless, we are potentially ruling out some objects that are below the observational limit, but likely having a relatively high accretion power (i.e. a larger  $L_{\text{bol,AGN}}$  value with respect to sources with FIR flux density above the detection limit). In order to quantify the missing fraction of AGN as a function of the observed  $160 \mu\text{m}$  flux density, the following hypotheses have been made.

1. We assume that the distribution of the flux ratio  $S_{\text{accr}}/S_{160}$  does not depend on redshift. Indeed, we are not able to characterise such distribution over sizeable samples at any redshift, especially at  $z > 2$  where the low statistics prevents us from deriving the ratio with high significance.
2. We rely on the basic assumption that the  $S_{160}$  completeness function derived by Berta et al. (2010, 2011) for  $160 \mu\text{m}$  sources in GOODS-S preserves the same trend also below the detection limit. Furthermore, the  $S_{160}$  flux completeness function derived for COSMOS sources has been extrapolated below 10.2 mJy on the existing one, by assuming the same trend for that derived in GOODS-S.

We fit the logarithmic distribution of flux ratio with a Gaussian function, for GOODS-S and COSMOS samples separately. In support to the assumption (ii), we observed a similar



**Figure A.1:** Distribution of GOODS-S (red asterisks) and COSMOS (blue squares) sources classified as AGN, as a function of  $S_{\text{accr}}$  and  $S_{160}$ . Dotted lines show the flux density limits at  $160 \mu\text{m}$  ( $S_{160,\text{lim}}$ ), whereas horizontal dashed lines set the thresholds in accretion flux that we found through this analysis. The red and blue solid lines trace the upper value of the distribution of flux ratio  $S_{\text{accr}} / S_{160}$ , as a function of the accretion flux. The red and blue shaded areas define the flux-to-flux regions that we have explored to calibrate our correction for incompleteness in  $S_{\text{accr}}$ . See text for details.

distribution of flux ratio among COSMOS and GOODS-South, which means that most likely the two fluxes scale in a similar way as a function of  $S_{160}$ . Such similar trend at lower  $S_{160}$  allows us to assume that our distribution is also preserved for lower  $S_{160}$  than the GOODS-S detection limit. As expected, the distributions of flux ratio are quite broad (about 0.6 dex), so they do not imply a tight correlation between star-formation and AGN accretion activities (see e.g. Rosario et al. 2012). However, the approach developed by Fontana et al. (2004) allows us to take into account the full distribution of flux ratios, even if quite broad.

In order to correct for the incompleteness in  $S_{\text{accr}}$ , it is necessary to evaluate the intrinsic number of AGN that are missed because  $S_{160}$  is lower than the detection limit  $S_{160,\text{lim}}$ , as a function of the accretion flux. In Fig. A.1 we plot the distribution of COSMOS (blue squares) and GOODS-S (red asterisks) sources as a function of both  $160 \mu\text{m}$  and accretion fluxes. The vertical dotted lines set the FIR detection limit of each PEP field. Each coloured solid line marks the upper boundary of the distribution of flux ratios sampled

by each value of the accretion flux. For sake of simplicity, hereafter we focus on the red asterisks only (representing GOODS-S). The following reasoning may be also extended to the blue squares (representing COSMOS).

Following Fontana et al. (2004), we set the starting point for our analysis in such a way that the highest flux ratio of the distribution samples the flux density limit  $S_{160,\text{lim}}$  value of the survey (see the down-arrow placed where the solid line crosses the dotted one in Fig. A.1). We may reasonably assume that above such threshold we are 100 per cent complete in accretion flux.

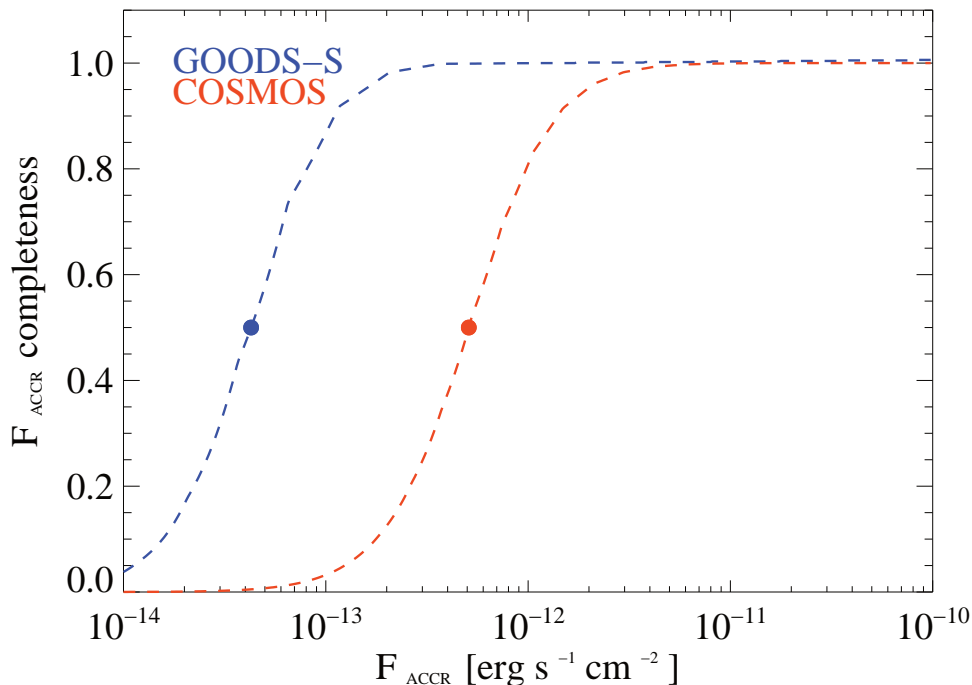
Since we are interested in quantifying the missing AGN fraction as a function of the accretion flux, we follow the down-arrow (see Fig. A.1), moving along the solid line down to lower and lower values of  $S_{\text{accr}}$ . According to the hypothesis (ii) we assume that also the distribution shifts down with the same trend. For each step in  $S_{\text{accr}}$  we need to evaluate the expected (i.e. intrinsic) number of sources  $N_{\text{exp}}$ , that is nothing more than the “classic”  $\log N_{\text{exp}} - \log S_{160}$ . This is achievable by extrapolating the existing  $S_{160}$  completeness curve down to lower and lower FIR flux densities.

As we are interested in deriving such information as a function of the accretion flux rather than FIR flux density, we convert the obtained  $\log N_{\text{exp}} - \log S_{160}$  into  $\log N_{\text{exp}} - \log S_{\text{accr}}$  through a convolution with the distribution of flux ratios. In other words, we use our distribution to derive the expected number of sources for each input accretion flux. From the comparison between the observed and expected number of sources, it is possible to obtain the detectable AGN fraction  $f_{\text{det}}$  as a function of  $S_{\text{accr}}$ :

$$f_{\text{det}}(S_{\text{accr}}) = \frac{\int_{S_{160,\text{lim}}}^{S_{160,\text{max}}} \frac{dN_{\text{exp}}}{dS_{160}} \cdot g(\bar{x}) dS_{160}}{\int_{S_{160,\text{min}}}^{S_{160,\text{max}}} \frac{dN_{\text{exp}}}{dS_{160}} \cdot g(\bar{x}) dS_{160}} \quad (\text{A.1})$$

where  $(\bar{x}) = S_{\text{accr}} / S_{160}$  and  $g(\bar{x})$  is the interpolated value of the distribution of flux ratios at  $[S_{\text{accr}}, S_{160}]$ . While  $S_{160,\text{lim}}$  is constant,  $S_{160,\text{min}}$  and  $S_{160,\text{max}}$  change step by step in accretion flux, since they represent the minimum and maximum FIR fluxes, respectively, sampled by the distribution of flux ratios for a given accretion flux  $S_{\text{accr}}$ .

The quantity in equation (A.1) is by definition  $\leq 1$  and scales downward as long as  $S_{\text{accr}}$  decreases, since the FIR-based selection becomes progressively more incomplete in accretion flux. As done by Fontana et al. (2004), we choose to run the iterations until the detectable fraction matches the lost fraction (i.e.  $f_{\text{det}} = 0.5$ ). The latter value identifies the minimum accretion flux  $S_{\text{accr},\text{lim}}$  that a source should have to enter a  $>50$  per cent complete accretion-based sample, regardless of its FIR flux density. In Fig. A.1 the shaded areas delimit the flux-to-flux regions that we have explored to calibrate our correction for incompleteness in accretion flux. We show in Fig. A.2 the how the completeness fraction increases at higher values of  $S_{\text{accr}}$ . Filled circles identify the 50% threshold in COSMOS (blue) and GOODS-S (red) samples. These thresholds in accretion flux are equal to  $4 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$  and  $5 \times 10^{-13} \text{ erg s}^{-1} \text{ cm}^{-2}$  in the GOODS-S and the COSMOS field respectively, also marked with horizontal dashed lines in Fig. A.1.



**Figure A.2:** Completeness fraction  $y$ -axis as a function of the accretion flux  $S_{\text{accr}}$   $x$ -axis, calculated for our *Herschel* sample in the COSMOS (blue dashed line) and in the GOODS-S (red dashed line) fields. Filled circles mark the 50% completeness flux that is applied in our analysis.

Within our AGN sample, only those sources with accretion fluxes larger than the corresponding threshold have been considered for the AGN bolometric luminosity function. By applying the above mentioned corrections, the predicted number of sources in each  $L_{\text{bol,AGN}}$  bin may increase up to a factor of 2. Nevertheless, the  $V_{\text{max}}$  method (Eq. 4.1) was referred to the accessible comoving volume for a source to be detected at  $160 \mu\text{m}$ , but not necessarily to enter at the same time an accretion-based selection. As a consequence, we define an “effective” maximum comoving volume  $V_{\text{max}}^{\text{eff}}$  as:

$$V_{\text{max}}^{\text{eff}} = \min[V_{\text{max}}, V_{\text{max,accr}}] \quad (\text{A.2})$$

where  $V_{\text{max,accr}}$  is the equivalent of the classic  $V_{\text{max}}$  definition, except for the fact that  $V_{\text{max,accr}}$  is computed as a function of the accretion flux rather than of FIR flux. In other words, we calculate a  $V_{\text{max,accr}}$  value for each object by following Eqs. (4.1) and (4.2) and replacing the FIR flux completeness function  $f_c(z)$  with that based on the accretion flux  $f_{\text{det}}(S_{\text{accr}})$ .

Finally, the effective volume  $V_{\text{max}}^{\text{eff}}$  (Eq. A.2) is computed for sources having  $S_{\text{accr}} \geq S_{\text{accr,lim}}$  and  $S_{160} \geq S_{160,lim}$ , fitting both simultaneously with FIR and accretion-based selections. At the same time,  $V_{\text{max}}^{\text{eff}}$  allows us to correct for the incompleteness in accretion flux, that is to place our *Herschel*-selected AGN within the appropriate comoving volume where all AGN with  $S_{\text{accr}} \geq S_{\text{accr,lim}}$ , either detected or undetected by *Herschel*, would be observable.

## Details on X-ray stacking

The stacking analysis allows to increase the signal-to-noise of individually non-detected sources by grouping them together and piling-up their photon counts. This technique returns an average net (i.e. background subtracted) count rate for a given input list. As mentioned in section 5.4.1, we stacked individual *Herschel*-selected galaxies not detected in X-rays. We masked all point-like sources listed in public X-ray catalogues (Alexander et al. 2003 for GOODS-N, Xue et al. 2011 for GOODS-S and Elvis et al. 2009 for C-COSMOS). We have assumed a circular shape for each mask, which radius is 50% larger than the area enclosing 90% of photon counts from the source. The size of the Point Spread Function (PSF) follows a radial (and energy-dependent) profile in the *Chandra* Deep Fields, while in COSMOS it shows fluctuations in between overlapping pointings (see Elvis et al. 2009). We accounted for that and ensured that the mean background level, after masking all X-ray sources, was fully consistent in each field with that presented in the corresponding reference paper.

To optimize the final signal-to-noise ratio (SNR), we have taken a fixed source aperture (i.e. 2 arcsec radius) centered on each optical position. The number of photon counts has been corrected by accounting for off-axis angle and observed frequency<sup>1</sup>.

Each background region is an annulus centered on the corresponding optical position, with inner radius placed 5 pixels (1 pixel = 0.492 arcsec) apart from the PSF size enclosing 90% photon counts. The outer radius is 5 pixels apart from the inner one. This avoids contamination from the PSF of the stacked source.

We subtracted the background counts, rescaled to the source area, to the photon counts collected from the source region, to get the net number of source counts. The exposure time is taken from energy-dependent time-maps, which return for each (x,y) position in the sky the effective exposure time (i.e. corrected for vignetting, bad pixels, dithering and including also a spatially dependent quantum efficiency). Finally, the average count rate is computed by summing (over all stacked sources) all net source counts and dividing by the total effective exposure time.

---

<sup>1</sup>This correction has been made using the *psf* module (<http://cxc.harvard.edu/ciao/ahelp/psf.html>) from the Chandra Interactive Analysis of Observations (CIAO) package.





## Statistics in the SFR– $M_{\star}$ plane

In Table C.1 we show the values of AGN accretion rate and its uncertainties, as a function of SFR,  $M_{\star}$  and redshift, in order to quantify the colour coded scale shown in Fig. 5.2. In addition, we list the number of sources included in each bin (total numbers, X-ray detected and X-ray undetected). Note that bins of SFR change as a function of redshift, as they are arranged to follow the cosmic evolution of the main-sequence relation.

**Table C.1:** List of average AGN accretion rate, as a function of SFR,  $M_*$  and redshift, with related  $1\sigma$  uncertainties. Alongside of each redshift range, we report the median redshift value  $\langle z \rangle$  and the sSFR<sub>ms</sub> that follows the main-sequence at  $\langle z \rangle$ . In some bins only an upper limit (upp) or a lower limit (low) is available. Numbers between parentheses in each bin represent: total number of *Herschel* sources, X-ray detected and X-ray stacked, respectively.

SFR [ $M_\odot \text{yr}^{-1}$ ]	$0.01 \leq z < 0.25$ $\langle z \rangle = 0.18$ sSFR <sub>ms</sub> = $1.27e-10 \text{ yr}^{-1}$			
	9.00 – 9.50	9.50 – 10.00	10.00 – 10.50	10.50 – 11.00
0.13 – 0.40	$-5.00^{+0.26}_{-0.24}$ (25 4 21)			11.00 – 11.50
0.40 – 1.27	$-4.40^{+0.36}_{-0.35}$ (31 3 28)	$-4.56^{+0.27}_{-2.24}$ (51 4 47)	$-5.84$ (low) (32 3 29)	$-4.04^{+0.20}_{-0.16}$ (22 5 17)
1.27 – 4.00		$-4.12^{+0.13}_{-0.09}$ (51 3 48)	$-4.43^{+0.20}_{-0.14}$ (89 8 81)	$-2.77^{+0.78}_{-0.32}$ (34 7 27)
4.00 – 12.66			$-3.92^{+0.29}_{-0.15}$ (19 3 16)	$-3.85^{+0.29}_{-0.17}$ (33 6 27)
<hr/>				
SFR [ $M_\odot \text{yr}^{-1}$ ]	$0.25 \leq z < 0.50$ $\langle z \rangle = 0.35$ sSFR <sub>ms</sub> = $1.79e-10 \text{ yr}^{-1}$			
	9.00 – 9.50	9.50 – 10.00	10.00 – 10.50	10.50 – 11.00
0.57 – 1.79		$-4.21^{+0.48}_{-0.34}$ (36 2 34)	$-5.65$ (upp) (29 0 29)	11.00 – 11.50
1.79 – 5.67	$-3.89^{+0.63}_{-0.45}$ (24 2 22)	$-3.98^{+0.98}_{-0.48}$ (128 4 124)	$-4.26^{+0.40}_{-0.20}$ (287 9 278)	$-2.50^{+0.64}_{-0.33}$ (177 12 165)
5.67 – 17.94		$-4.10$ (upp) (15 2 13)	$-3.42^{+0.86}_{-0.42}$ (104 10 94)	$-2.62^{+0.38}_{-0.23}$ (59 10 49)
17.94 – 56.72			$-3.12^{+0.28}_{-0.25}$ (31 7 24)	$-1.91^{+1.50}_{-0.56}$ (16 4 12)

		$0.50 \leq z < 0.80$		$\langle z \rangle = 0.65$	$sSFR_{ms} = 3.13e-10 \text{ yr}^{-1}$		
SFR [ $M_{\odot}\text{yr}^{-1}$ ]		$\log(M_{\star}/M_{\odot})$					
		9.00 – 9.50	9.50 – 10.00	10.00 – 10.50	10.50 – 11.00	11.00 – 11.50	11.50 – 12.00
0.99 – 3.13			-4.59 <sup>+0.94</sup> -0.31 (15 1 14)				
3.13 – 9.91			-3.20 <sup>+0.83</sup> -0.43 (78 4 74)	-3.27 <sup>+0.47</sup> -0.05 (171 13 158)	-2.51 <sup>+0.30</sup> -0.20 (141 11 130)	-1.79 <sup>+1.37</sup> -0.45 (34 4 30)	
9.91 – 31.33			-2.16 <sup>+1.13</sup> -0.16 (80 8 72)	-3.29 <sup>+0.25</sup> -0.16 (300 16 284)	-2.85 <sup>+0.21</sup> -0.15 (522 26 496)	-1.98 <sup>+0.38</sup> -0.18 (168 28 140)	
31.33 – 99.08				-2.07 <sup>+0.94</sup> -0.49 (20 2 18)	-2.07 <sup>+0.35</sup> -0.17 ( 82 18 64)	-2.81 <sup>+0.64</sup> -0.37 (53 7 46)	
		$0.80 \leq z < 1.50$		$\langle z \rangle = 1.00$	$sSFR_{ms} = 5.54e-10 \text{ yr}^{-1}$		
SFR [ $M_{\odot}\text{yr}^{-1}$ ]		$\log(M_{\star}/M_{\odot})$					
		9.00 – 9.50	9.50 – 10.00	10.00 – 10.50	10.50 – 11.00	11.00 – 11.50	11.50 – 12.00
5.54 – 17.52			-3.28 <sup>+0.16</sup> -0.16 (56 3 53)	-2.77 <sup>+0.27</sup> -0.21 (154 7 147)	-2.15 <sup>+0.22</sup> -0.15 (202 15 187)	-1.91 <sup>+0.74</sup> -0.33 (69 8 61)	
17.52 – 55.40			-2.30 <sup>+1.01</sup> -0.53 (71 5 66)	-2.38 <sup>+0.23</sup> -0.15 (420 18 402)	-2.16 <sup>+0.13</sup> -0.11 (985 65 920)	-1.77 <sup>+0.14</sup> -0.11 (501 55 446)	
55.40 – 175.18			-1.40 <sup>+0.50</sup> -0.32 (91 11 80)	-1.38 <sup>+0.36</sup> -0.20 (339 37 302)	-1.36 <sup>+0.29</sup> -0.19 (296 42 254)	-1.72 <sup>+0.71</sup> -0.33 (19 1 18)	
175.18 – 553.98				-1.46 <sup>+0.38</sup> -0.34 (18 5 13)	-1.41 <sup>+0.95</sup> -0.35 (19 3 16)		

SFR [ $M_\odot \text{yr}^{-1}$ ]	$1.50 \leq z \leq 2.50$					$\langle z \rangle = 1.91$	sSFR $_{\text{ms}} = 1.79\text{e-}9 \text{ yr}^{-1}$
	9.00 – 9.50	9.50 – 10.00	$\log(M_\star/M_\odot)$ 10.00 – 10.50	10.50 – 11.00	11.00 – 11.50		
17.92 – 56.68	$-1.03^{+1.29}_{-0.20}$ (19 3 16)	$-2.44^{+0.61}_{-0.40}$ (47 2 45)	$-1.74^{+0.24}_{-0.16}$ (103 11 92)	$-1.38^{+0.95}_{-0.40}$ (35 3 32)			
56.68 – 179.24	$-0.53^{+0.36}_{-0.34}$ (16 4 12)	$-0.82^{+0.39}_{-0.27}$ (125 20 105)	$-1.50^{+0.29}_{-0.18}$ (230 19 211)	$-1.13^{+0.15}_{-0.12}$ (272 43 229)	$-0.57^{+0.25}_{-0.20}$ (38 12 26)		
179.24 – 566.80		$-0.49^{+0.81}_{-0.49}$ (22 6 16)	$-1.01^{+0.39}_{-0.27}$ (75 14 61)	$-0.63^{+0.49}_{-0.25}$ (128 24 104)	$-0.87^{+0.32}_{-0.21}$ (34 10 24)		

## Maximum Likelihood function

In this Appendix we describe the construction of a maximum likelihood (ML) function to estimate the most probable value of a given parameter, in case the sample that we want to analyse include both detected and undetected sources. This method provides an alternative approach compared to the weighted average presented in Eq. 5.1: differently from what is stated in Eq. 5.1, the ML function considers the single undetected source as an upper limit, rather than assuming the same stacked value for all undetected sources.

The ML method relies on two assumptions: first, both detected and stacked sources are supposed to belong to the same underlying population. This means that the distribution of the parameter that we are interested in (in this case the observed X-ray luminosity  $L_X$ ) is the same for all sources. This assumption can be tested afterwards. Second, one has to assume a given distribution of the parameter of interest: for simplicity we assume a Gaussian distribution in  $\log L_X$  for our sample. Under this assumption, the likelihood function depends only on the average  $L_X$  and on the dispersion of the  $L_X$  distribution. Adopting the same syntax used in Eq. 5.1, let's assume that our sample includes  $m$  detected and  $n$  undetected sources. We computed the likelihood function  $\mathcal{L}$  according to the formula:

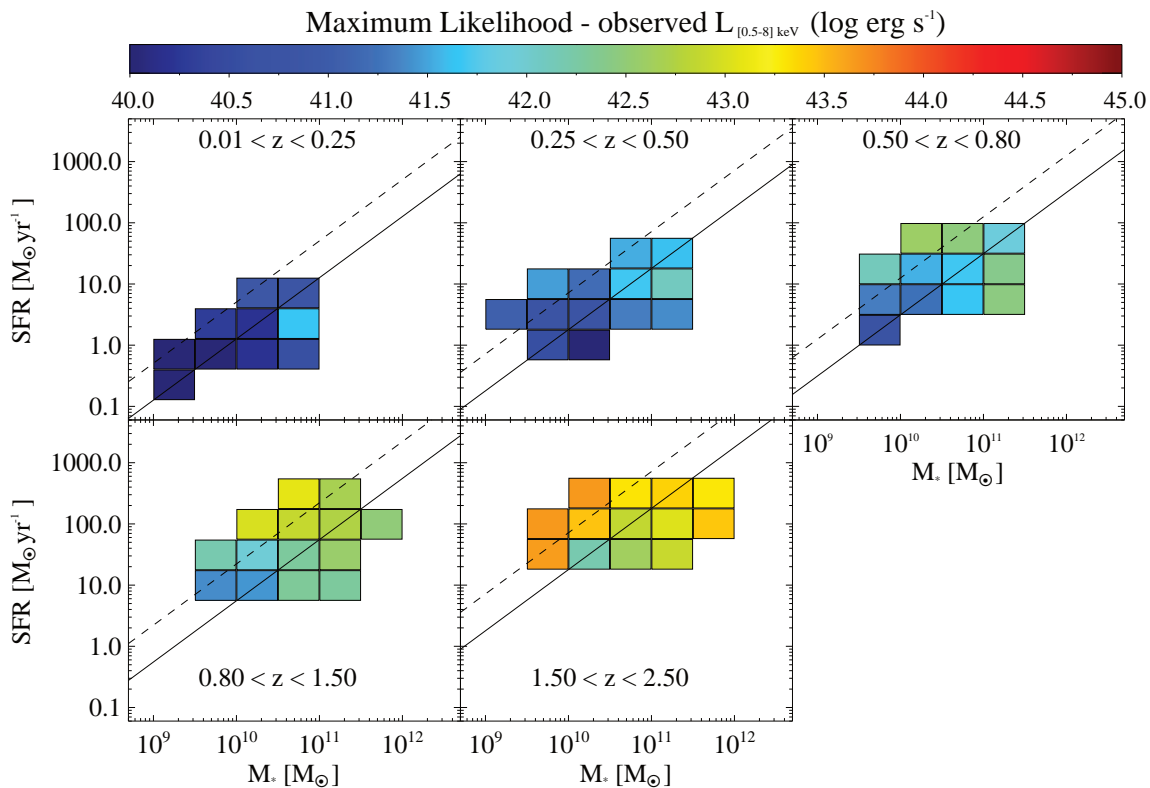
$$\mathcal{L} = \prod_{i=1}^m f(\log L_{X,i}) \cdot \prod_{j=1}^n \int_{-\infty}^{\log L_{X,\text{lim}}} f(\log L_{X,j}) d \log L_{X,j} \quad (\text{D.1})$$

where  $f$  is the Gaussian function in  $\log L_X$ :

$$f(\log L_X) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(\log L_X - \log L_{X,0})^2}{2\sigma^2}} \quad (\text{D.2})$$

The left factor of Eq. D.1 represents the likelihood function calculated for the  $m$  detected objects, while the right factor refers to the  $n$  undetected ones. The integral within the right factor is calculated for each source and measures the total probability of finding that source below the nominal detection limit  $L_{X,\text{lim}}$ .

The likelihood function  $\mathcal{L}$  enables to estimate the most probable value  $\log L_{X,0}$  and the dispersion  $\sigma$  of the underlying Gaussian distribution. The best estimate for  $(\log L_{X,0}, \sigma)$  is found by maximizing  $\mathcal{L}(\log L_{X,0}, \sigma)$ , while the uncertainties on  $(\log L_{X,0}, \sigma)$  are obtained



**Figure D.1:** Distribution of the observed X-ray luminosity (in colour coding) estimated through ML in the SFR– $M_*$  plane at  $0.01 \leq z \leq 2.5$ .

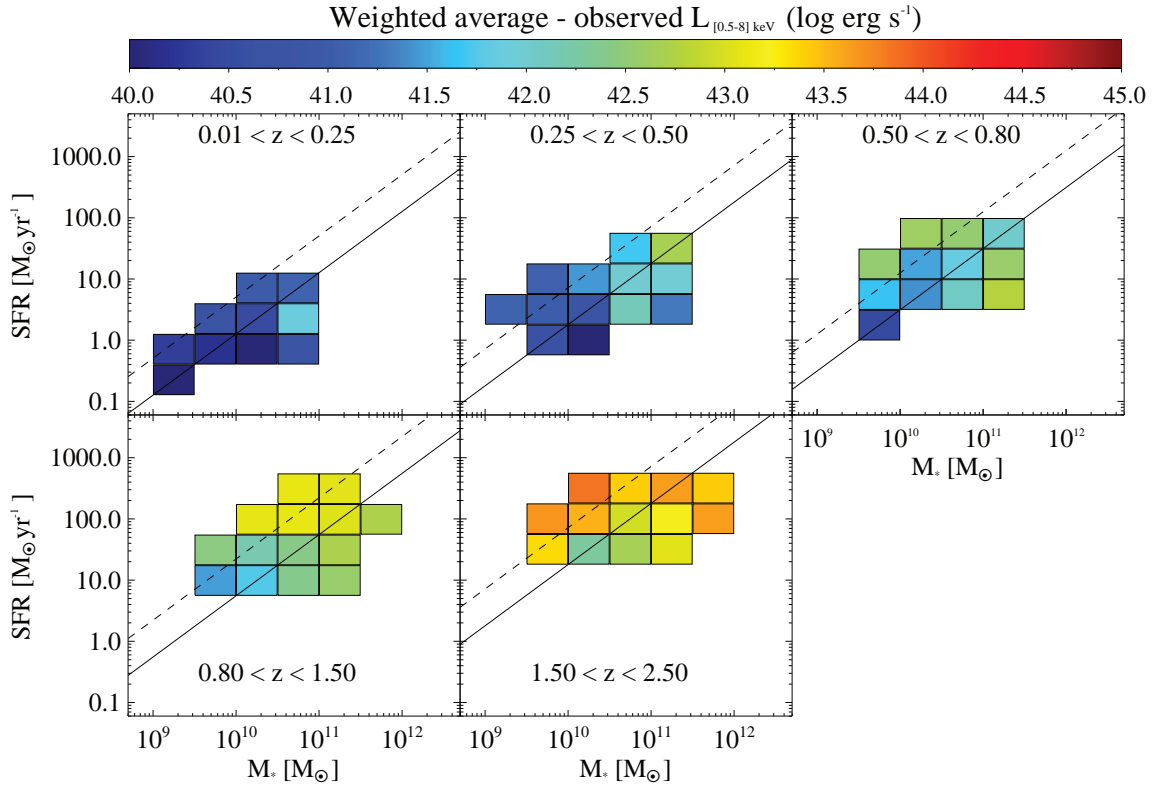
by finding all the  $(\log L_{X,0}, \sigma)$  pairs for which  $\mathcal{L} = e^{-1/2} \mathcal{L}_{\max}$ , where  $\mathcal{L}_{\max}$  is the maximum likelihood ML computed for the  $(n + m)$  sources.

We iterated the calculation of the ML function for each bin of SFR,  $M_*$  and redshift shown in Fig. 5.2. We estimated the X-ray luminosity limit  $L_{X,\text{lim}}$  (in the 0.5–8 keV band) for each undetected source directly from the X-ray sensitivity maps available for the GOODS-S, the GOODS-N and the COSMOS fields. The best estimate for  $\sigma$  is around 0.8, suggesting that the underlying Gaussian distribution is quite broad, as expected given that both X-ray detected and undetected sources are considered in the same sample. The most probable  $\log L_{X,0}$  obtained in each bin is calculated in logarithmic scale. The conversion from logarithmic to linear scale, under the assumption of a Gaussian distribution, depends only on the dispersion  $\sigma$  inferred from the ML:

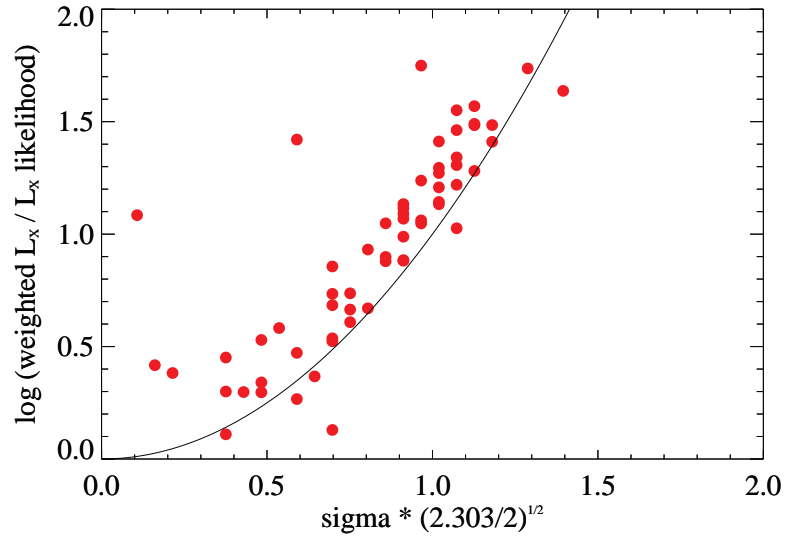
$$\log L_{X,\text{mean}} = \log L_{X,0} + \frac{2.303}{2} \sigma^2 \quad (\text{D.3})$$

where  $L_{X,\text{mean}}$  is the most probable X-ray luminosity estimated through the ML function in linear scale. This value can be directly compared to that derived from Eq. 5.1 to test the hypothesis that X-ray detected and undetected sources belong to the same underlying population.

We estimated this luminosity in each bin of SFR,  $M_*$  and redshift studied in Fig. 5.2. The distribution of the ML-based X-ray luminosities in the SFR– $M_*$  plane is shown in



**Figure D.2:** Distribution of the observed X-ray luminosity (in colour coding) estimated through weighted average (see Eq. 5.1) in the SFR– $M_*$  plane at  $0.01 \leq z \leq 2.5$ .



**Figure D.3:** Weighted X-ray luminosity over ML-based X-ray luminosity for each SFR– $M_*$ – $z$  bin (red circles), as a function of  $\sigma$ . The black solid line indicates the parabola fitting the expected  $\propto \sigma^2$  correlation.

Fig. D.1. For comparison, we show also the distribution of the observed X-ray luminosities derived from a weighted average in Fig. D.2. The colour coding X-ray luminosity

looks reasonably similar. To quantify possible differences, we study the logarithmic ratio between the weighted- $L_X$  and the ML-based  $L_{X,0}$ , as a function of  $\sigma$ . Fig. D.3 shows this comparison: the solid line marks the expected parabolic correlation (given Eq. D.3) in case of perfect match between X-ray luminosities. Red circles represent individual bins of SFR,  $M_*$  and redshift shown in Fig. D.1 and D.2.

The observed distribution of points shows an offset as high as 0.18 dex, with the dispersion of about 0.25 dex. This check suggests that the weighted average provides with reasonably consistent estimates of X-ray luminosity compared to what obtained from the ML function. As a consequence, this implies that the assumption that *Herschel* galaxies, both detected and undetected in X-rays, belong to the same underlying Gaussian population, is quite reasonable.