

# Exploring the relationship between black hole accretion and star formation with blind mid-/far-infrared spectroscopic surveys

M. Bonato,<sup>1,2★</sup> M. Negrello,<sup>2</sup> Z.-Y. Cai,<sup>3,4</sup> G. De Zotti,<sup>2,4</sup> A. Bressan,<sup>4</sup> A. Lapi,<sup>4,5</sup>  
F. Pozzi,<sup>6</sup> C. Gruppioni<sup>7</sup> and L. Danese<sup>4</sup>

<sup>1</sup>Dipartimento di Fisica e Astronomia ‘G. Galilei’, Università degli Studi di Padova, Vicolo Osservatorio 3, I-35122 Padova, Italy

<sup>2</sup>INAF, Osservatorio Astronomico di Padova, Vicolo Osservatorio 5, I-35122 Padova, Italy

<sup>3</sup>Center for Astrophysics, University of Science and Technology of China, Hefei 230026, China

<sup>4</sup>SISSA, Via Bonomea 265, I-34136 Trieste, Italy

<sup>5</sup>Dipartimento di Fisica, Università ‘Tor Vergata’, Via della Ricerca Scientifica 1, I-00133 Roma, Italy

<sup>6</sup>Dipartimento di Fisica e Astronomia, Università di Bologna, via Ranzani 1, I-40127 Bologna, Italy

<sup>7</sup>INAF, Osservatorio Astronomico di Bologna, Via Ranzani 1, I-40127 Bologna, Italy

Accepted 2014 August 12. Received 2014 August 11; in original form 2014 June 4

## ABSTRACT

We present new estimates of redshift-dependent luminosity functions of IR lines detectable by SPICA/SAFARI (SPace InfraRed telescope for Cosmology and Astrophysics/SpicA FAR infrared Instrument) and excited both by star formation and by AGN activity. The new estimates improve over previous work by using updated evolutionary models and dealing in a self-consistent way with emission of galaxies as a whole, including both the starburst and the AGN component. New relationships between line and AGN bolometric luminosity have been derived and those between line and IR luminosities of the starburst component have been updated. These ingredients were used to work out predictions for the source counts in 11 mid-/far-IR emission lines partially or entirely excited by AGN activity. We find that the statistics of the emission line detection of galaxies as a whole is mainly determined by the star formation rate, because of the rarity of bright AGNs. We also find that the slope of the line integral number counts is flatter than two implying that the number of detections at fixed observing time increases more by extending the survey area than by going deeper. We thus propose a wide spectroscopic survey of 1 h integration per field of view over an area of 5 deg<sup>2</sup> to detect (at 5 $\sigma$ )  $\sim$ 760 AGNs in [O IV] $\lambda$ 25.89  $\mu$ m – the brightest AGN mid-infrared line – out to  $z \sim 2$ . Pointed observations of strongly lensed or hyperluminous galaxies previously detected by large area surveys such as those by *Herschel* and by the South Pole Telescope can provide key information on the galaxy–AGN co-evolution out to higher redshifts.

**Key words:** galaxies: active – galaxies: evolution – galaxies: starburst – infrared: galaxies.

## 1 INTRODUCTION

One of the main scientific goals of the SPace InfraRed telescope for Cosmology and Astrophysics (SPICA)<sup>1</sup> with its SpicA FAR infrared Instrument (SAFARI; Roelfsema et al. 2012) is to provide insight into the interplay between galaxy evolution and the growth of active nuclei (AGNs) at their centres during the dust-enshrouded

active star formation and mass accretion phases. The imaging spectrometer SAFARI is designed to fully exploit the extremely low far-infrared (far-IR) background environment provided by the SPICA observatory, whose telescope will be actively cooled at 6 K. In each integration, it will take complete 34–210  $\mu$ m spectra split in three bands (34–60  $\mu$ m, 60–110  $\mu$ m, 110–210  $\mu$ m), spatially resolving the full 2 arcmin  $\times$  2 arcmin field of view (FoV). These bands are weakly, if at all, affected by dust extinction and thus allowing direct measurements of the intrinsic properties of the sources.

The mid-IR (MIR) lines that will be detected can come either from star-forming regions or from nuclear activity or from both.

★E-mail: [matteo.bonato@oapd.inaf.it](mailto:matteo.bonato@oapd.inaf.it)

<sup>1</sup>[http://www.ir.isas.jaxa.jp/SPICA/SPICA\\_HP/index-en.html](http://www.ir.isas.jaxa.jp/SPICA/SPICA_HP/index-en.html)

Distinguishing the two contributions is not easy. However, a key diagnostic relies on the fact that AGNs produce harder radiation and therefore excite metals to higher ionization states than starbursts. Thus, lines with high-ionization potential are a powerful tool for identifying AGN activity (Sturm et al. 2002; Meléndez et al. 2008; Satyapal et al. 2008, 2009; Dudik, Satyapal & Marcu 2009; Goulding & Alexander 2009; Weaver et al. 2010) that is easily missed by optical spectroscopic observations in the case of dusty objects, as galaxies with intense star formation are.

So far, these studies have been limited to local objects. Substantial progress in this area will be made possible by spectroscopy with SAFARI that will exploit for the first time the rich suite of mid- and far-IR diagnostic lines to trace the star formation and the accretion on to the supermassive black holes up to high redshifts through both blind spectroscopic surveys and pointed observations.

An exploratory study of the SPICA/SAFARI potential in detecting galaxies and AGNs in the main IR lines in several redshift ranges has been presented by Spinoglio et al. (2012). Their calculations were based on the phenomenological evolutionary models by Franceschini et al. (2010), Gruppioni et al. (2011) and Valiante et al. (2009). The Franceschini et al. (2010) and the Gruppioni et al. (2011) models, like most classical backward evolution models, evolve the star-forming galaxies and the AGNs independently. This class of models have been quite successful in reproducing far-IR to sub-millimetre source counts but are clearly unable to predict the composite line spectra of galaxies in which star formation and AGN activity co-exist.

Valiante et al. (2009) have carried out a first attempt to deal in a coherent way with the cosmic evolution of IR emissions of both the starburst and the AGN components. This was done by first estimating the distribution of the ratios between the  $6\ \mu\text{m}$  AGN continuum luminosity,  $\nu L_{6\ \mu\text{m}}$ , and the total IR luminosity,  $L_{\text{IR}} = L_{8-1000\ \mu\text{m}}$ , for a sample of local galaxies in five IR luminosity bins. The luminosity-dependent distributions of  $\nu L_{6\ \mu\text{m}}/L_{\text{IR}}$  were derived subtracting the estimated starburst contribution from the *Spitzer* Infrared Spectrograph (IRS) MIR data for a complete sample of *IRAS* galaxies. A check on the evolution with redshift of the mean  $\nu L_{6\ \mu\text{m}}$  versus  $L_{\text{IR}}$  relation was made using IRS measurements for a sample of galaxies at  $0.37 < z < 3.35$ . The mean ratio  $\nu L_{6\ \mu\text{m}}/L_{\text{IR}}$  was found to increase with luminosity as  $\nu L_{6\ \mu\text{m}}/L_{\text{IR}} \propto L_{\text{IR}}^{1.4 \pm 0.6}$ . The increase was assumed to stop at  $L_{\text{IR},f} = 10^{12.8} L_{\odot}$  for  $z < 0.5$  and at  $L_{\text{IR},f} = 10^{12} L_{\odot}$  for  $z \geq 0.5$ . Then an evolutionary law for  $L_{\text{IR}}$  was chosen and a value of  $\nu L_{6\ \mu\text{m}}/L_{\text{IR}}$  drawn at random from the appropriate distribution was assigned to each source.

The strong increase of the average AGN luminosity fraction with increasing  $L_{\text{IR}}$  up to  $L_{\text{IR},f}$  derived by Valiante et al. (2009) is however not supported by many other analyses, although the relationship between these two quantities is still matter of debate with studies showing either strong or marginal correlations between AGN and starburst luminosities (Lutz et al. 2008; Serjeant & Hatziminaoglou 2009; Alexander & Hickox 2012; Diamond-Stanic & Rieke 2012; Mullaney et al. 2012a,b; Rosario et al. 2012; Chen et al. 2013; Chen & Hickox 2014; Hickox et al. 2014). The apparently contradictory results of different analyses may be understood taking into account on one side that the characteristic time-scale of black hole accretion is very different from that of star formation and on the other side that any relation between star formation rate (SFR) and accretion rate must break down as the AGN luminosity is at the Eddington limit.

The different time-scales imply that accretion and star formation are not necessarily on at the same time. Hence, the relation among the two quantities may have a large dispersion when individual

objects are considered, still being strong when averaging over the whole population of star-forming galaxies (Chen et al. 2013). A coherent scenario for the interpretation of the various pieces of evidence on the connection between star formation and AGN activity has been elaborated by Lapi et al. (2014).

In this paper, we present a new approach to the problem taking advantage of the physical model for the co-evolution of supermassive black holes and massive protospheroidal galaxies at high redshift ( $z \geq 1.5$ ) worked out by Cai et al. (2013) who presented the evolving luminosity functions of these objects as a whole (starburst plus AGN), taking into account in a self-consistent way the variation with galactic age of the global spectral energy distribution (SED).

The physical model was complemented by a phenomenological model for lower  $z$  AGNs and star-forming galaxies, mostly late-type since massive spheroidal galaxies are observed to be in essentially passive evolution at  $z \lesssim 1-1.5$ . In this case, galaxies and AGNs were evolved separately. We combined the two components exploiting the average accretion rate as a function of the SFR derived by Chen et al. (2013), taking into account its dispersion. Bright optically selected QSOs, that do not obey such correlation, were taken into account adopting the best-fitting evolutionary model by Croom et al. (2009) up to  $z = 2$ .

The plan of the paper is the following. In Section 2, we present the adopted model for the evolution with cosmic time of the IR ( $8-1000\ \mu\text{m}$ ) luminosity function. In Section 3, we discuss the relations between line and continuum luminosity for the main mid-/far-IR lines. In Section 4, we work out our predictions for line luminosity functions, number counts and redshift distributions in the SPICA/SAFARI bands. In Section 5, we compare our results with previous estimates. In Section 6, we discuss possible SPICA/SAFARI observation strategies, considering different integration times per FoV. Section 7 contains a summary of our main conclusions.

We adopt a flat  $\Lambda$  cold dark matter cosmology with matter density  $\Omega_m = 0.32$ , dark energy density  $\Omega_\Lambda = 0.68$  and Hubble constant  $h = H_0/100\ \text{km s}^{-1}\ \text{Mpc}^{-1} = 0.67$  (Planck Collaboration XVI 2013).

## 2 EVOLUTION OF THE IR LUMINOSITY FUNCTIONS

As mentioned in Section 1, our reference model for the redshift-dependent IR luminosity functions is the one worked out by Cai et al. (2013) based on a comprehensive ‘hybrid’ approach combining a physical model for the progenitors of early-type galaxies with a phenomenological one for late-types. The evolution of progenitors of early-type galaxies and of massive bulges of Sa’s is described by an updated version of the physical model by Granato et al. (2004, see also Lapi et al. 2006, Lapi et al. 2011 and Mao et al. 2007). In the local Universe, these objects are composed of relatively old stellar populations with mass-weighted ages  $\gtrsim 8-9$  Gyr, corresponding to formation redshifts  $z \gtrsim 1-1.5$ , while the disc components of spirals and the irregular galaxies are characterized by significantly younger stellar populations (cf. Bernardi et al. 2010, their fig. 10). Thus, the progenitors of early-type galaxies, referred to as protospheroidal galaxies or ‘protospheroids’, are the dominant star-forming population at  $z \gtrsim 1.5$ , while IR galaxies at  $z \lesssim 1.5$  are mostly late-type galaxies.

According to Cai et al. (2013), the redshift-dependent luminosity functions of protospheroids as a whole,  $\Phi(L_{\text{tot}}, z)$ , can be

written as

$$\begin{aligned} \Phi(L_{\text{tot}}, z) &= \int_{M_{\text{min}}}^{M_{\text{max}}} dM_{\text{vir}} \int_z^{z_{\text{max}}} dz_{\text{vir}} \left| \frac{dt_{\text{vir}}}{dz_{\text{vir}}} \right| \\ &\times \frac{dN_{\text{ST}}}{dt_{\text{vir}}}(M_{\text{vir}}, z_{\text{vir}}) \int_{-\infty}^{\log L_{\text{tot}}} \frac{dx}{2\pi\sigma_*\sigma_*} \\ &\times \frac{L_{\text{tot}}}{L_{\text{tot}} - 10^x} \exp[-(x - \log \bar{L}_*)^2/2\sigma_*^2] \\ &\times \exp[-[\log(L_{\text{tot}} - 10^x) - \bar{L}_*]^2/2\sigma_*^2], \end{aligned} \quad (1)$$

where  $\bar{L}_* = \bar{L}_*(z|M_{\text{vir}}, z_{\text{vir}})$  and  $\bar{L}_* = \bar{L}_*(z|M_{\text{vir}}, z_{\text{vir}})$  are the redshift-dependent mean starburst and AGN luminosities at given halo mass,  $M_{\text{vir}}$ , and virialization redshift,  $z_{\text{vir}}$ , respectively, and  $\sigma_*$  and  $\sigma_*$  are the corresponding dispersions.  $dN_{\text{ST}}/dt$  is the dark matter halo mass function (Sheth & Tormen 1999).

The conditional luminosity function of the starburst component,  $\Phi_*(L_*|L_{\text{tot}}, z)$ , i.e. the number density of sources at redshift  $z$  with starburst luminosity  $L_*$  given that the total luminosity is  $L_{\text{tot}}$ , can be written as

$$\begin{aligned} \Phi_*(L_*|L_{\text{tot}}, z) d \log L_* &= \int_{M_{\text{min}}}^{M_{\text{max}}} dM_{\text{vir}} \\ &\times \int_z^{z_{\text{max}}} dz_{\text{vir}} \left| \frac{dt_{\text{vir}}}{dz_{\text{vir}}} \right| \frac{dN_{\text{ST}}}{dt_{\text{vir}}}(M_{\text{vir}}, z_{\text{vir}}) \frac{d \log L_*}{2\pi\sigma_*\sigma_*} \\ &\times \frac{L_{\text{tot}}}{L_{\text{tot}} - L_*} \exp[-(\log L_* - \log \bar{L}_*)^2/2\sigma_*^2] \\ &\times \exp[-[\log(L_{\text{tot}} - L_*) - \log \bar{L}_*]^2/2\sigma_*^2], \end{aligned} \quad (2)$$

while the conditional luminosity function of the AGN component can be written as

$$\begin{aligned} \Phi_*(L_*|L_{\text{tot}}, z) d \log L_* &= \int_{M_{\text{min}}}^{M_{\text{max}}} dM_{\text{vir}} \\ &\times \int_z^{z_{\text{max}}} dz_{\text{vir}} \left| \frac{dt_{\text{vir}}}{dz_{\text{vir}}} \right| \frac{dN_{\text{ST}}}{dt_{\text{vir}}}(M_{\text{vir}}, z_{\text{vir}}) \frac{d \log L_*}{2\pi\sigma_*\sigma_*} \\ &\times \frac{L_{\text{tot}}}{L_{\text{tot}} - L_*} \exp[-[\log(L_{\text{tot}} - L_*) - \log \bar{L}_*]^2/2\sigma_*^2] \\ &\times \exp[-[\log(L_* - \log \bar{L}_*)]^2/2\sigma_*^2]. \end{aligned} \quad (3)$$

The probability that an object at redshift  $z$  has a starburst luminosity  $L_*$  or an AGN luminosity  $L_*$  given the total luminosity of  $L_{\text{tot}} = L_* + L_*$  is then

$$P_*(L_*|L_{\text{tot}}, z) d \log L_* \equiv \frac{\Phi_*(L_*|L_{\text{tot}}, z) d \log L_*}{\Phi(L_{\text{tot}}, z)}, \quad (4)$$

or

$$P_*(L_*|L_{\text{tot}}, z) d \log L_* \equiv \frac{\Phi_*(L_*|L_{\text{tot}}, z) d \log L_*}{\Phi(L_{\text{tot}}, z)}. \quad (5)$$

The two probabilities have unit integral over the luminosity range  $0 - L_{\text{tot}}$  and are linked by the condition that, at fixed  $L_{\text{tot}}$ ,  $P_*(L_*|L_{\text{tot}}) = -P_*(L_*|L_{\text{tot}}) (d \log L_*/d \log L_*)$  with  $d \log L_*/d \log L_* = d \log(L_{\text{tot}} - L_*)/d \log L_* = -L_*/(L_{\text{tot}} - L_*)$ . Then

$$P_*(L_*|L_{\text{tot}}) = \frac{L_*}{L_{\text{tot}} - L_*} P_*(L_{\text{tot}} - L_*|L_{\text{tot}}) \quad (6)$$

or

$$P_*(L_*|L_{\text{tot}}) = \frac{L_*}{L_{\text{tot}} - L_*} P_*(L_{\text{tot}} - L_*|L_{\text{tot}}). \quad (7)$$

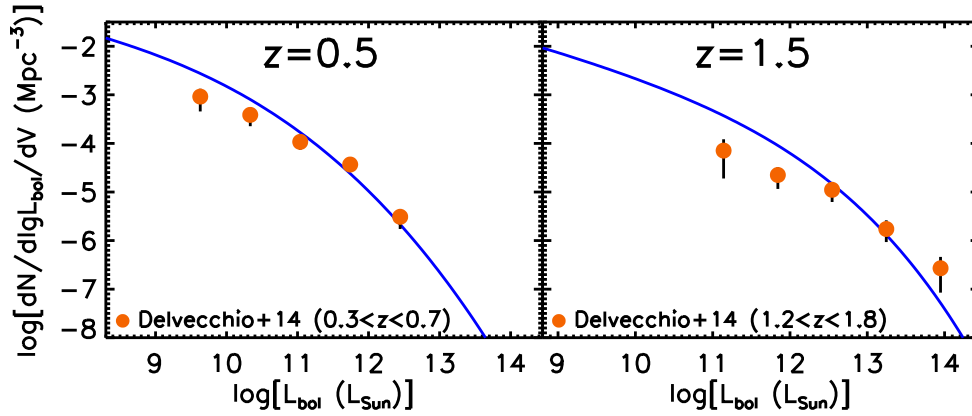
Each protospheroid of luminosity  $L_{\text{tot}}$  was assigned bolometric luminosities of the starburst and of the AGN components drawn

at random from the above probability distributions, with the parameter values given by Cai et al. (2013). Note that for the starburst component the bolometric luminosity is very close to the IR (8–1000  $\mu\text{m}$ ) luminosity, but this is not the case for AGNs, especially for type-1s. For example, adopting the AGN SEDs used by Cai et al. (2013) we get  $L_{\text{IR}}/L_{\text{bol}} = 0.189$  and  $0.776$  for type-1s, type-2s, respectively. This should be taken into account when comparing our results with those by Spinoglio et al. (2012) which rely on correlations between line and IR luminosity rather than between line and bolometric luminosity, as done in this paper.

The model makes detailed predictions also for strongly gravitationally lensed galaxies. These objects are of special interest because strong lensing allows us to measure their total mass distribution up to very large distances and to gain information on sources too faint to be detected with current instrument sensitivities, thus testing models for galaxy formation and dark matter. To deal with them, the probability distributions have been applied to the de-lensed luminosities. The de-lensing was made attributing to each lensed source an amplification factor randomly extracted from the amplification probability distributions by Negrello et al. (2007) and Lapi et al. (2012). After having done that, the luminosities of both components were re-amplified by the same factor.

As mentioned above, the evolution of late-type galaxies, the dominant star-forming population at  $z \lesssim 1.5$ , is described by a phenomenological, parametric model, distinguishing two sub-populations, ‘cold’ (normal) and ‘warm’ (starburst) galaxies. The AGNs associated with these populations are evolved independently, again with a phenomenological recipe. To combine galaxies and AGNs into sources including both components, we have exploited the correlation between SFR and average black hole accretion rate derived by Chen et al. (2013) on the basis of a sample of 1767 far-IR selected galaxies in the redshift range  $0.25 < z < 0.8$ . Using their equation 5, their bolometric correction for the AGN X-ray emission  $L_{\text{bol}}/L_X = 22.4$  and their relationship between the IR luminosity of the starburst,  $L_{\text{IR}}$ , and the SFR we got  $\langle L_{\text{bol}}/L_{\text{IR}} \rangle = 0.054(L_{\text{IR}}/10^{12} L_{\odot})^{0.05}$ .

It is also necessary to properly take into account the large (see the discussion in Section 1) dispersion around the mean relation. Such dispersion is difficult to evaluate because of the large fraction of galaxies in the Chen et al. (2013) sample for which only upper limits to the nuclear emission are available. We have estimated it by trial and error, looking for a value that yields redshift-dependent AGN luminosity functions consistent with the observational estimates by Delvecchio et al. (2014). In practice, we have carried out simulations associating galaxies having a given  $L_{\text{IR}}$  with an AGN with bolometric luminosity drawn at random from a Gaussian distribution with mean given by the average  $L_{\text{bol}}/L_{\text{IR}}$  relation and several values of the dispersion. The contribution of optically selected AGNs (see below) was then added to the derived luminosity functions. We get consistency with the data (see a couple of examples in Fig. 1) with a dispersion of 0.69 dex, in agreement with the distribution of data points in figs 3 and 4 of Chen et al. (2013). The fact that the model appears to overestimate the low-luminosity tail of the AGN bolometric luminosity function is not worrisome for the present purposes since these sources are anyways too weak to be detected by SPICA/SAFARI. It should also be noticed that the observational errors are purely statistical, i.e. do not include uncertainties on the decomposition of the observed SED to disentangle the AGN contribution from that of the host galaxy, on the bolometric correction and on the correction for incompleteness. These systematic errors increase with decreasing AGN luminosity. The



**Figure 1.** Comparison of the AGN bolometric luminosity functions at  $z = 0.5$  and  $z = 1.5$  yielded by our approach (solid lines) with the observational estimates by Delvecchio et al. (2014). The dotted red curve in the right-hand panel shows the contribution of strongly lensed (amplification  $\mu \geq 2$ ) sources; at  $z = 0.5$  such contribution is negligible. The vertical lines correspond to the minimum bolometric luminosity yielding a detection by SAFARI of the strongest AGN line ([O IV] 25.89  $\mu\text{m}$ ) with exposures of 1 and 10 h/FoV (solid and dotted, respectively), for the mean line to bolometric luminosity ratios.

evolution of sources as a whole is then driven by that of the SFR, as modelled by Cai et al. (2013).

Not all AGNs obey the accretion rate – SFR correlation. In particular, hosts of bright optically selected QSOs, with high accretion rates, are found to have SFRs ranging from very low to moderate, particularly at relatively low redshifts (Schweitzer et al. 2006; Netzer et al. 2007). These objects were taken into account adopting the best-fitting evolutionary model by Croom et al. (2009) up to  $z = 2$  since at higher redshifts the abundance of AGNs associated with late-type galaxies is negligible compared to that of AGNs associated with protospheroids whose evolution up to the optically bright phase is described by the Cai et al. (2013) physical model.

The consistency between the model and the observed bolometric luminosity functions, illustrated by Fig. 1 for  $z \leq 1.5$  and demonstrated by Cai et al. (2013) at higher  $z$ , is crucial for the reliability of the derived line luminosity functions because they are calculated coupling the bolometric luminosity functions with the relations between line and bolometric luminosities (see Section 3).

In summary, the different IR galaxy populations considered in our paper are

- (i) protospheroidal galaxies (unlensed and lensed), the progenitors of early-type galaxies and the dominant star-forming population at  $z \gtrsim 1.5$ , whose evolution is described by the physical model by Cai et al. (2013);
- (ii) AGNs associated with protospheroids whose co-evolution with the host galaxies is self-consistently built in the Cai et al. (2013) physical model;
- (iii) late-type galaxies, the dominant star-forming population at  $z \lesssim 1.5$ , divided into two sub-populations (spiral and starburst galaxies), whose evolution is described by the Cai et al. (2013) phenomenological model;
- (iv) type-1 and type-2 AGNs associated with late-type galaxies, whose evolution is linked to that of host galaxies via the Chen et al. (2013) accretion rate – SFR correlation;
- (v) optical AGNs, whose evolution is described by the Croom et al. (2009) best-fitting model up to  $z = 2$  and included in the Cai et al. (2013) physical model at higher  $z$ .

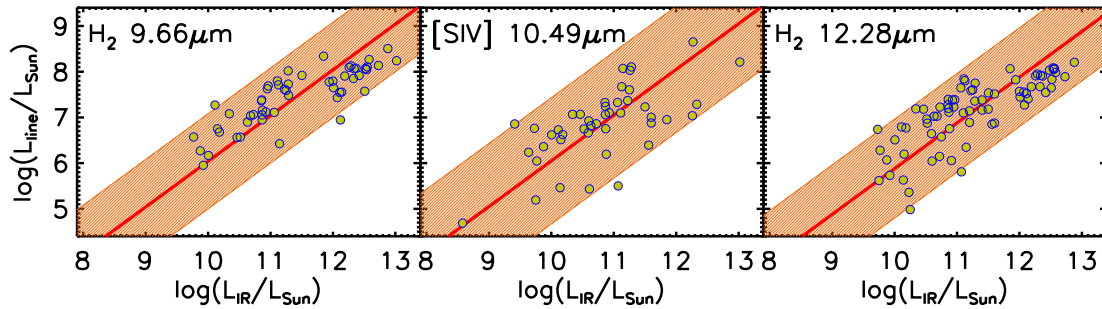
While referring to the Cai et al. (2013) paper for full details, we summarize here, for the sake of clarity, a few points. Protospheroidal galaxies and massive bulges of disc galaxies are assumed to *virialize* at  $z \geq 1.5$  but their evolution is followed down to  $z = 0$ . Their

lifetimes in the star formation phase increase with decreasing halo mass, consistent with the observed ‘downsizing’, from  $\simeq 0.7$  Gyr for the most massive ones to a few Gyr for the least massive. Hence, they are contributing to the IR luminosity functions also at  $z < 1.5$ , but their contribution decreases with decreasing  $z$ . At  $z < 1.5$ , late-type galaxies take over them as the dominant contributors. The phenomenological model describing their evolution includes a smooth decline of their space density above  $z \simeq 1.5$  but their contribution to the global IR luminosity function is still significant up to  $z \simeq 2$ .

The evolution of supermassive black holes associated with protospheroidal galaxies is followed by the physical model *only* during their active star formation phase when black holes acquire most of their mass. In other words, the present implementation of the physical model is unable to follow the later AGN evolution when spheroids are in essentially passive evolution and nuclei can be re-activated by, e.g., interactions, mergers, or dynamical instabilities, especially if they acquire a disc, i.e. become the bulges of later-type galaxies. These later phases are dealt with by the phenomenological model, as described above.

### 3 LINE VERSUS IR LUMINOSITY

IR lines excited by star formation activity have been dealt with by Bonato et al. (2014), with an analysis of the relations between line and starburst IR luminosities that considered differences among source populations and that was supported by extensive simulations taking into account dust obscuration. We complement their study considering the brightest lines that are excited by, or also by, AGNs. These include three typical AGN lines ([Ne V] 14.32  $\mu\text{m}$ , [Ne V] 24.31  $\mu\text{m}$  and [O IV] 25.89  $\mu\text{m}$ ) and eight lines that can also be produced in star formation regions (the five H II region lines: [S IV] 10.49  $\mu\text{m}$ , [Ne II] 12.81  $\mu\text{m}$ , [Ne III] 15.55  $\mu\text{m}$ , [S III] 18.71  $\mu\text{m}$  and [S III] 33.48  $\mu\text{m}$ ; three molecular hydrogen lines, H<sub>2</sub> 9.66  $\mu\text{m}$ , H<sub>2</sub> 12.28  $\mu\text{m}$  and H<sub>2</sub> 17.03  $\mu\text{m}$ ). The other IR lines considered by Bonato et al. (2014) have negligible AGN contributions so that we adopt directly the luminosity functions derived in that paper. In addition, we present results for the 6.2, 7.7 and 8.6  $\mu\text{m}$  polycyclic aromatic hydrocarbons (PAH) lines, overlooked by Bonato et al. (2014) but of considerable interest for SPICA/SAFARI surveys. Of course, to detect the broad PAH lines the SAFARI ( $R = 2000$ ) spectral resolution needs to be degraded without loss of signal, using



**Figure 2.** Line luminosity as a function of the IR luminosity of the starburst component for the (H<sub>2</sub>)9.66 μm, [S IV]10.49 μm and (H<sub>2</sub>)12.28 μm lines. Luminosities are in solar units. Data from Roussel et al. (2006), Farrah et al. (2007), Bernard-Salas et al. (2009), Hao et al. (2009), O’Dowd et al. (2009, 2011), Pereira-Santaella et al. (2010) and Cormier et al. (2012), excluding all objects for which there is evidence of a substantial AGN contribution. The orange bands show the  $2\sigma$  spread around the mean linear relations (red lines) in the log–log plane. The mean ratios  $\langle \log(L_{\ell}/L_{\text{IR}}) \rangle$  and the dispersions around them are given in Table 1.

appropriate algorithms. All the lines have been taken into account to compute the number of sources detected in more than one line.

As in Bonato et al. (2014), the line luminosities are derived from the continuum ones exploiting the relations between line and IR (or bolometric in the AGN case) luminosities, taking into account the dispersions around the mean values. For five of the lines listed above, namely [Ne II]12.81 μm, [Ne III]15.55 μm, H<sub>2</sub> 17.03 μm, [S III]18.71 μm and [S III]33.48 μm, these relationships were already derived, for the starburst component only, by Bonato et al. (2014). Again for the starburst component, we have collected from the literature data on three additional lines (H<sub>2</sub> 9.66 μm, [S IV]10.49 μm and H<sub>2</sub> 12.28 μm), excluding objects for which there is evidence of a substantial AGN contribution. These data are shown in Fig. 2 together with the best-fitting linear relations. We have neglected the starburst contribution to the three typical AGN lines.

The mean ratios  $\langle \log(L_{\ell}/L_{\text{IR}}) \rangle$  for star-forming galaxies and the dispersions around them are given in Table 1. The values for the 6.2,

7.7 and 8.6 μm PAH lines (also given in Table 1) were inferred from that for the PAH 11.25 μm line, given by Bonato et al. (2014), using the average ratios  $L_{\text{PAH}6.2}/L_{\text{PAH}11.25} = 1.22$ ,  $L_{\text{PAH}7.7}/L_{\text{PAH}11.25} = 4.5$  and  $L_{\text{PAH}8.6}/L_{\text{PAH}11.25} = 1.36$  given by Fiolet et al. (2010).

The relationships between line and bolometric luminosities of AGNs were determined using the data on the QSO samples by Veilleux et al. (2009), on the local Seyfert galaxies whose MIR emission was found to be 100 per cent AGN dominated by Sturm et al. (2002) and, only for the three pure AGN lines ([Ne V]14.32 μm, [Ne V]24.31 μm and [O IV]25.89 μm), on the local Seyfert galaxies by Tommasin et al. (2008, 2010) coupled with the continuum nuclear emission measurements by Gorjian et al. (2004), Gandhi et al. (2009) and Hönig et al. (2010).

The IR luminosities given in Veilleux et al. (2009) and Sturm et al. (2002) are defined in Sanders & Mirabel (1996), but this definition applies to starburst not to AGN SEDs. Therefore, to convert them to AGN bolometric luminosities we cannot use the correction factors given in Section 2. The AGN bolometric luminosities were then estimated using the appropriate Cai et al. (2013) SEDs (i.e. for QSOs and Seyfert 1 galaxies, the mean QSO SED of Richards et al. 2006 extended to millimetre wavelengths as described in Cai et al. 2013; for Seyfert 2 galaxies, the SED of the local AGN-dominated ULIRG Mrk 231 taken from the SWIRE library) normalized to the observed 12 μm luminosities which are expected to be modestly affected by contamination from the host galaxy. Normalizing instead to the AGN 6 μm fluxes of Sani et al. (2010), we obtain, for a sub-sample containing 65 per cent of our objects, an average decrease of the estimated AGN bolometric luminosity by about 17 per cent. On the other hand, we caution that the Mrk 231 SED might have a non-negligible host galaxy contamination. In fact its far-IR emission peaks at  $\sim 100$  μm. At such relatively long wavelength the contribution from dust heated by young stellar populations can be substantial or even dominant (Mullaney et al. 2011). Therefore, our approach likely overestimates the bolometric luminosities of Seyfert 2 galaxies. This however has little impact on our conclusions since these AGNs are minor contributors to line counts at the SPICA/SAFARI detection limits.

At variance with star-forming galaxies for which data and extensive simulations are consistent with a constant line-to-continuum IR luminosity ratio (Bonato et al. 2014), for AGNs observational data indicate luminosity-dependent ratios (Hill et al. 2014). We adopt a relation of the form  $\log(L_{\ell}) = a \log(L_{\text{bol}}) + b$ . The best-fitting coefficients for the 11 lines of our sample and the dispersions,  $\sigma$ , are listed in Table 2. There are no significant differences among

**Table 1.** Mean values of the log of line to IR (8–1000 μm) continuum luminosities for star-forming galaxies,  $\langle \log(L_{\ell}/L_{\text{IR}}) \rangle$  and associated dispersions  $\sigma$ . For the H<sub>2</sub> 17.03 μm line, the tabulated values, taken from Bonato et al. (2014), have been computed excluding local ULIRGs, for which the luminosity in this line was found to be uncorrelated with  $L_{\text{IR}}$ . The mean value of  $\log(L_{\ell})$  in this line for local ULIRGs was found to be  $\log(L_{\ell}/L_{\odot}) = 8.07$  with a dispersion of 0.34.

Spectral line	$\left\langle \log \left( \frac{L_{\ell}}{L_{\text{IR}}} \right) \right\rangle (\sigma)$
PAH6.2 μm	−2.20 (0.36)
PAH7.7 μm	−1.64 (0.36)
PAH8.6 μm	−2.16 (0.36)
H <sub>2</sub> 9.66 μm	−3.96 (0.52)
[S IV]10.49 μm	−3.95 (0.69)
H <sub>2</sub> 12.28 μm	−4.12 (0.54)
[Ne II]12.81 μm	−3.11 (0.45) <sup>a</sup>
[Ne III]15.55 μm	−3.69 (0.47) <sup>a</sup>
H <sub>2</sub> 17.03 μm	−4.04 (0.46) <sup>a</sup>
[S III]18.71 μm	−3.49 (0.48) <sup>a</sup>
[S III]33.48 μm	−3.05 (0.31) <sup>a</sup>

<sup>a</sup>Taken from Bonato et al. (2014).

**Table 2.** Coefficients of the best-fitting linear relations between line and AGN bolometric luminosities,  $\log(L_\ell) = a \log(L_{\text{bol}}) + b$ , and  $1\sigma$  dispersions associated with the relations.

Spectral line	$a$	$b$	Disp. ( $1\sigma$ )
H <sub>2</sub> 9.66 $\mu\text{m}$	1.07	− 5.32	0.34
[S IV] 10.49 $\mu\text{m}$	0.90	− 2.96	0.24
H <sub>2</sub> 12.28 $\mu\text{m}$	0.94	− 3.88	0.24
[Ne II] 12.81 $\mu\text{m}$	0.98	− 4.06	0.37
[Ne V] 14.32 $\mu\text{m}$	0.78	− 1.61	0.39
[Ne III] 15.55 $\mu\text{m}$	0.78	− 1.44	0.31
H <sub>2</sub> 17.03 $\mu\text{m}$	1.05	− 5.10	0.42
[S III] 18.71 $\mu\text{m}$	0.96	− 3.75	0.31
[Ne V] 24.31 $\mu\text{m}$	0.69	− 0.50	0.39
[O IV] 25.89 $\mu\text{m}$	0.70	− 0.04	0.42
[S III] 33.48 $\mu\text{m}$	0.62	0.35	0.30

the line–bolometric luminosity correlations for the different AGN types (see Fig. 3).

#### 4 LINE LUMINOSITY FUNCTIONS AND NUMBER COUNTS

To compute the line luminosity functions and the number counts, we adopted a Monte Carlo approach, starting from the redshift-dependent luminosity functions of the source populations described in Section 2, including both the starburst and the AGN component. Each source was assigned luminosities of each component drawn at random from the appropriate probability distributions (see Section 2). Then, we associated each component line luminosity drawn at random from Gaussian distributions with the mean values and the dispersions given in Tables 1 and 2. This procedure gives the line luminosity functions of the objects as a whole, as well as of

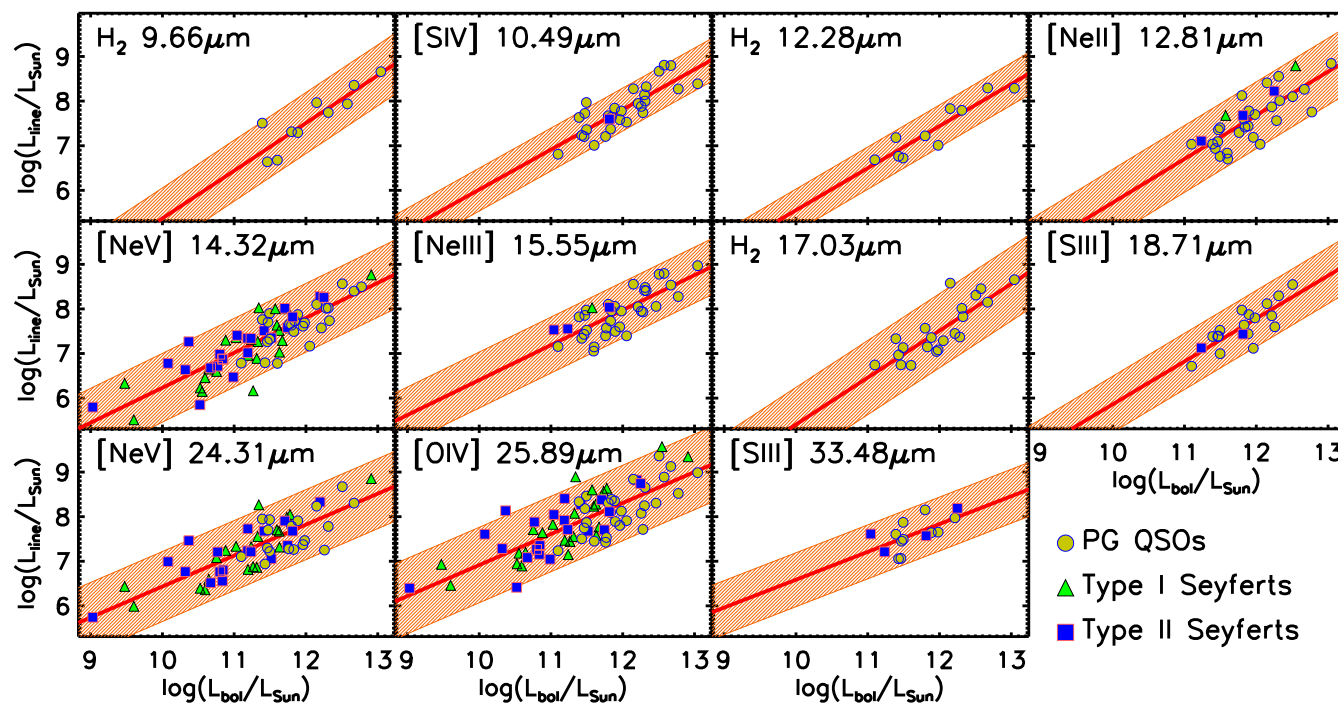
their starburst and AGN components. Examples are shown in Figs 4 and 5.

The counts are then straightforwardly computed. The wavelength range covered by each SPICA/SAFARI band,  $[\lambda_{\text{min}}, \lambda_{\text{max}}]$ , and the line rest-frame wavelength,  $\lambda_\ell$ , define the minimum and the maximum redshift within which the line is detectable in that band,  $z_{\ell, \text{min/max}} = \lambda_{\text{min/max}}/\lambda_\ell - 1$ . The integral counts predicted by our model in the SPICA/SAFARI wavelength range for the 14 IR lines of our sample are shown in Fig. 6.

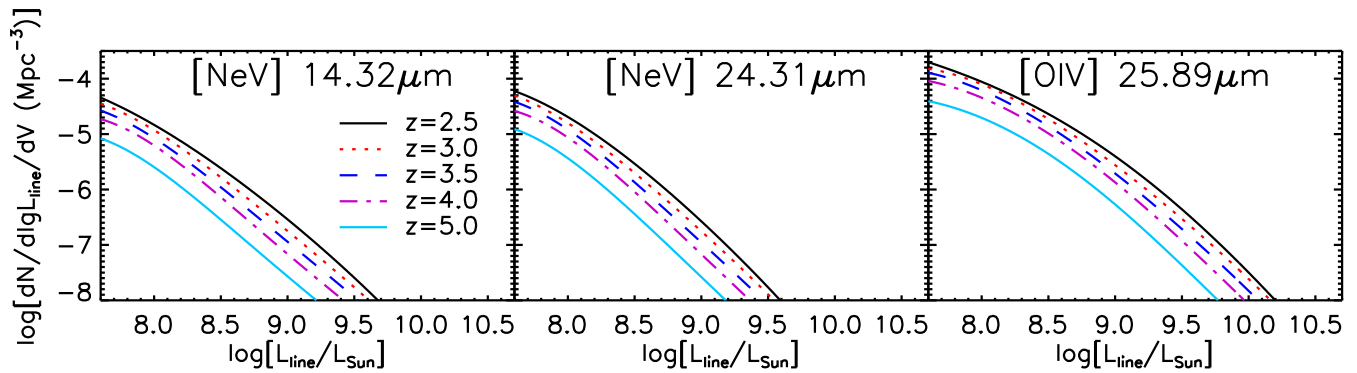
The SPICA/SAFARI reference  $5\sigma$  detection limits for an integration of 1 h per FoV are (Sibthorpe, private communication)  $3.7 \times 10^{-19} \text{ W m}^{-2}$  for the first band (34–60  $\mu\text{m}$ ),  $3.4 \times 10^{-19} \text{ W m}^{-2}$  for the second band (60–110  $\mu\text{m}$ ) and  $2.9 \times 10^{-19} \text{ W m}^{-2}$  for the third band (110–210  $\mu\text{m}$ ). We present predictions for a survey of  $5 \text{ deg}^2$  with 1 h per FoV. The survey area is 10 times larger compared to the reference survey considered in Bonato et al. (2014). In fact, as discussed later on in this section and in Section 6, AGN lines are more difficult to detect than starburst lines and thus a wider/deeper survey is required to achieve good statistics. The predicted numbers of sources detected in each line by a SPICA/SAFARI survey covering  $5 \text{ deg}^2$  with a 1 h integration/FoV are given in the last column of Table 3. The other columns detail the redshift distributions of detected sources. This SPICA/SAFARI survey will be able to detect many hundreds of AGNs (762, according to our calculations) in the strongest line ([O IV] 25.89  $\mu\text{m}$ ).

The redshift distributions of sources detected in the three AGN lines, [Ne V] 14.32  $\mu\text{m}$ , [Ne V] 24.31  $\mu\text{m}$  and [O IV] 25.89  $\mu\text{m}$ , are displayed in Fig. 7, where the redshift ranges covered by the different bands are identified by different colours. Fig. 8 shows the contributions to the counts of AGNs of the various kinds.

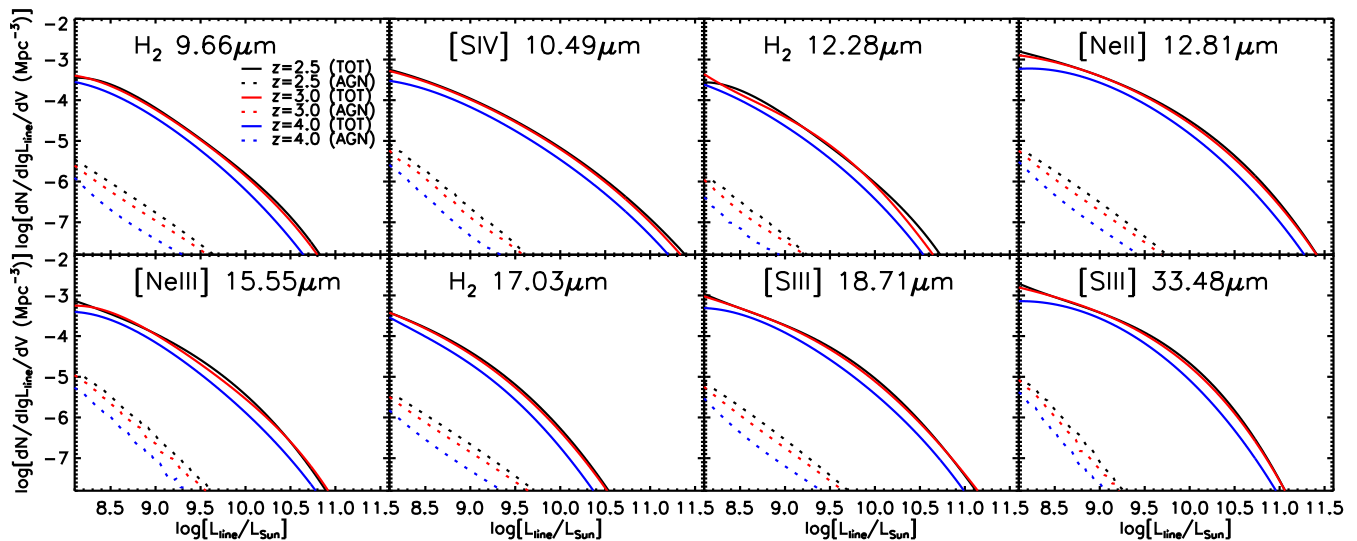
Fig. 9 illustrates the luminosity and redshift distributions of sources (starburst plus AGN components) for which the SPICA/SAFARI  $5 \text{ deg}^2$  survey will detect at least two, three or



**Figure 3.** Line versus bolometric luminosity, in solar units, for AGNs. PG QSO data (yellow circles) are taken from Veilleux et al. (2009), Seyfert 1 (green triangles) and Seyfert 2 (blue squares) data from Sturm et al. (2002) and, only for the [Ne V] 14.32  $\mu\text{m}$ , [Ne V] 24.31  $\mu\text{m}$  and [O IV] 25.89  $\mu\text{m}$  lines, also from Tommasin et al. (2008, 2010). The orange bands show the  $2\sigma$  spread around the linear relations  $\log(L_\ell) = a \log(L_{\text{bol}}) + b$  (red line).



**Figure 4.** Predicted luminosity functions of the three typical AGN lines of our sample at different redshifts at which these lines can be detected by SPICA/SAFARI.



**Figure 5.** Predicted luminosity functions for the total (starburst plus AGN) emission (solid lines) and for the AGN component only (dashed lines) at different redshifts at which these lines can be detected by SPICA/SAFARI.

four lines, taking into account both the spectral lines considered in this paper and the pure starburst lines considered in Bonato et al. (2014). We expect that 11 4903, 50 817, 29 492, 18 412, 11 555 and 6930 sources will be detected in at least one, two, three, four, five and six lines, respectively. The numbers of protospheroidal galaxies detectable in at least one, two, three, four, five and six lines are 44 335, 13 854, 6149, 3151, 1749 and 907, respectively. Sources detected in at least one line include  $\sim 724$  strongly lensed galaxies at  $z > 1$ ; about 300 of them will be detected in at least two lines.

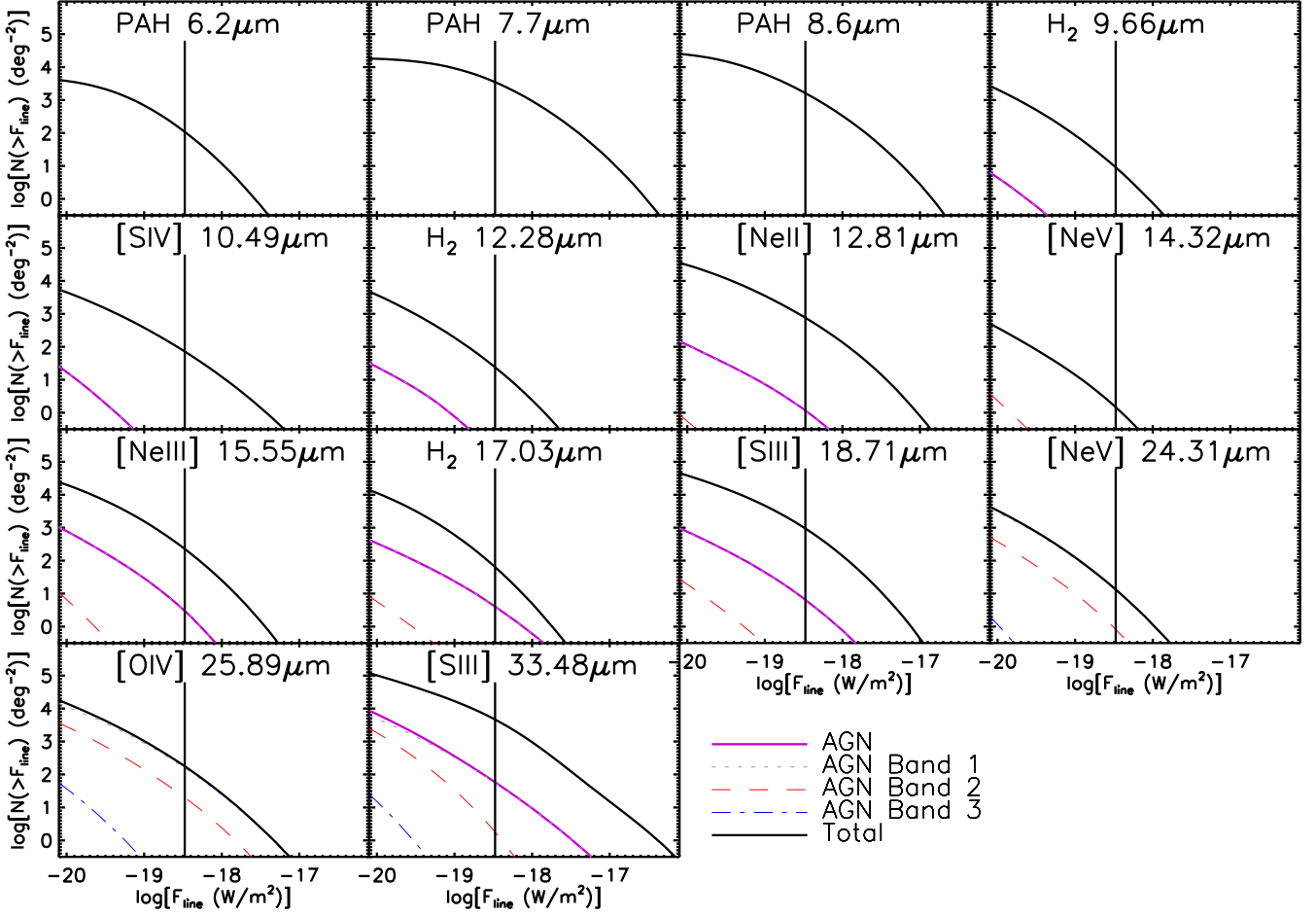
It is easily seen that these values, once rescaled to the same survey area, a factor of 10 smaller, are similar to those calculated in Bonato et al. (2014) considering only star-forming galaxies: the AGN component adds a minor contribution. The reason for that is illustrated by Fig. 10 where the luminosity functions of the brightest AGN line ([O IV]25.89  $\mu\text{m}$ ) are compared with those of two bright starburst lines at two redshifts. Two factors concur to yield a much higher space density of detectable starbursts compared to AGNs. First, AGNs are far less numerous than starburst galaxies, due to the much shorter lifetime of their bright phase. Secondly, many starburst lines have ratios to bolometric luminosity substantially higher than even the brightest AGN line (cf. the luminosity ratios for starbursts in table 1 of Bonato et al. 2014 with those for AGNs in Table 2 of this paper).

## 5 COMPARISON WITH PREVIOUS ESTIMATES

Fig. 11 compares our predictions for the redshift distributions of sources detected by SPICA/SAFARI with those from the three models used by Spinoglio et al. (2012). The total number of detections, for each line, is also summarized in Table 4.

The different results reflect the different approaches adopted by the authors (only very marginally the different evolutionary models used). In particular, Spinoglio et al. (2012)<sup>2</sup> based their estimate on empirical relations between the line luminosities and the total IR luminosity for the Seyfert galaxies of the 12  $\mu\text{m}$  galaxy sample (Tommasin et al. 2008, 2010), and for the pure starburst galaxies of the sample of Bernard-Salas et al. (2009). These luminosity ratios were then used to convert the total IR luminosity functions from different evolutionary models (i.e. Valiante et al. 2009; Franceschini et al. 2010; Gruppioni et al. 2011) into line luminosity functions and to derive the numbers of objects

<sup>2</sup> In the calculation of the IR luminosities used to calibrate the line to continuum relations in the Spinoglio et al. (2012) paper a factor of 1.8 was inadvertently omitted, resulting in an overestimate of the line counts (an erratum on that has been submitted to ApJ).



**Figure 6.** Integral counts in several lines of galaxies as a whole (starburst plus AGN components; solid black lines) and of AGNs only (dash-dotted violet lines) over the full SPICA-SAFARI wavelength range, and, for the AGNs, in each of its three bands. The vertical lines correspond to the average detection limit in the three bands for 1 h integration/FoV.

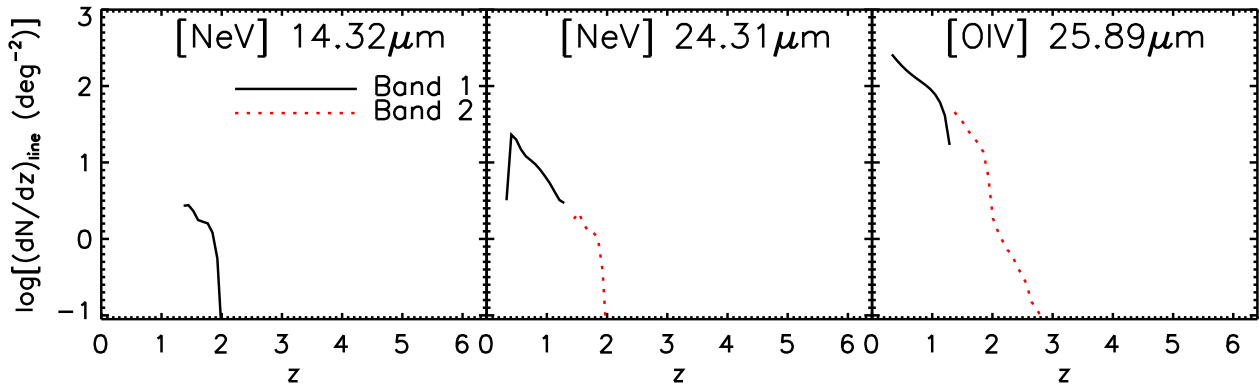
**Table 3.** Predicted redshift distributions of AGNs and, in parenthesis, of galaxies as a whole (starburst plus AGN components) detectable by a SPICA/SAFARI survey covering an area of  $5 \text{ deg}^2$  in 1 h integration per FoV.

Spectral line	0.00–0.75	0.75–1.25	1.25–1.75	1.75–2.25	2.25–2.75	2.75–4.00	4.00–6.00	6.00–8.00	All $z$
PAH6.2 $\mu\text{m}$	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (432)	0 (19)	0 (451)
PAH7.7 $\mu\text{m}$	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (8750)	0 (6787)	0 (237)	0 (15 775)
PAH8.6 $\mu\text{m}$	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (5778)	0 (1280)	0 (26)	0 (7085)
H <sub>2</sub> 9.66 $\mu\text{m}$	0 (0)	0 (0)	0 (0)	0 (0)	0.2 (18)	0.2 (21)	0 (0.7)	0 (0)	0.4 (39)
[SIV]10.49 $\mu\text{m}$	0 (0)	0 (0)	0 (0)	0 (9)	0.1 (169)	0 (119)	0 (12)	0 (0.2)	0.1 (308)
H <sub>2</sub> 12.28 $\mu\text{m}$	0 (0)	0 (0)	0 (0)	0.3 (58)	0 (26)	0 (11)	0 (1)	0 (0)	0.3 (95)
[Ne II]12.81 $\mu\text{m}$	0 (0)	0 (0)	1 (490)	1 (1515)	0.4 (748)	0.3 (495)	0 (43)	0 (0.3)	3 (3293)
[Ne V]14.32 $\mu\text{m}$	0 (0)	0 (0)	9 (9)	2 (2)	0.4 (0.4)	0.2 (0.2)	0 (0)	0 (0)	11 (11)
[Ne III]15.55 $\mu\text{m}$	0 (0)	1 (82)	10 (493)	2 (220)	0.3 (92)	0.3 (52)	0 (1)	0 (0)	13 (941)
H <sub>2</sub> 17.03 $\mu\text{m}$	0 (0)	9 (84)	10 (100)	1 (45)	0.4 (20)	0.2 (10)	0 (0.3)	0 (0)	20 (260)
[S III]18.71 $\mu\text{m}$	0 (0)	13 (2040)	10 (1125)	1 (499)	0.4 (261)	0.1 (143)	0 (10)	0 (0.3)	25 (4078)
[Ne V]24.31 $\mu\text{m}$	28 (28)	19 (19)	8 (8)	1 (1)	0.3 (0.3)	0.2 (0.2)	0 (0)	0 (0)	55 (55)
[O IV]25.89 $\mu\text{m}$	426 (426)	227 (227)	89 (89)	20 (20)	1 (1)	0.3 (0.3)	0 (0)	0 (0)	762 (762)
[S III]33.48 $\mu\text{m}$	235 (9876)	10 (6760)	1 (2947)	1 (1289)	0.2 (771)	0.1 (438)	0 (21)	0 (0)	248 (22 104)

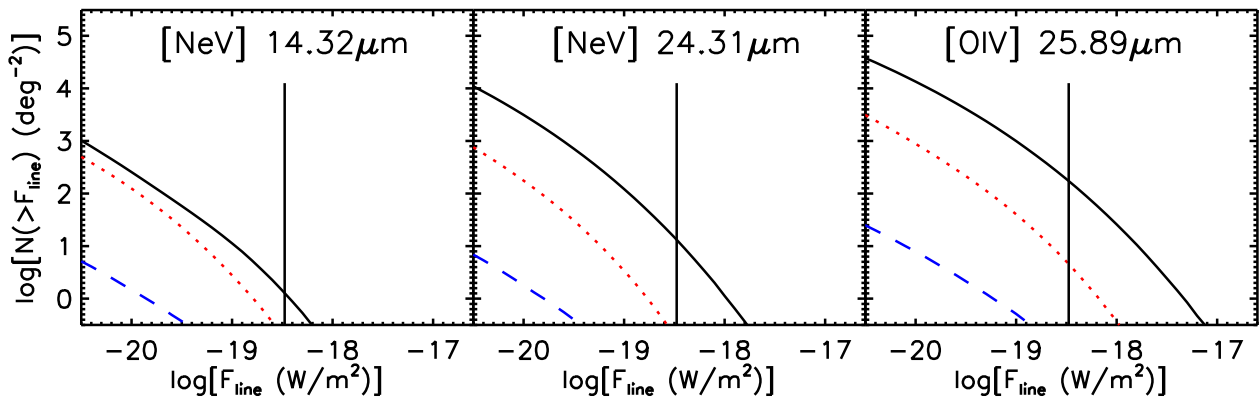
detectable by SPICA at different redshifts and luminosities. This approach – purely empirical – could be taken as an upper limit for the SPICA detections, since it applies the relations found for a sample of AGNs to all the AGN populations considered in the models, regardless of the AGN contribution to  $L_{\text{IR}}$ , while –

thanks to *Herschel* – we now know that  $L_{\text{IR}}$  in IR detected AGNs is typically composed by both AGN and starburst contributions (Hatziminaoglou et al. 2010) in different proportions (in large fractions of objects the AGN contributes to  $L_{\text{IR}}$  for <10 per cent; see Delvecchio et al. 2014).

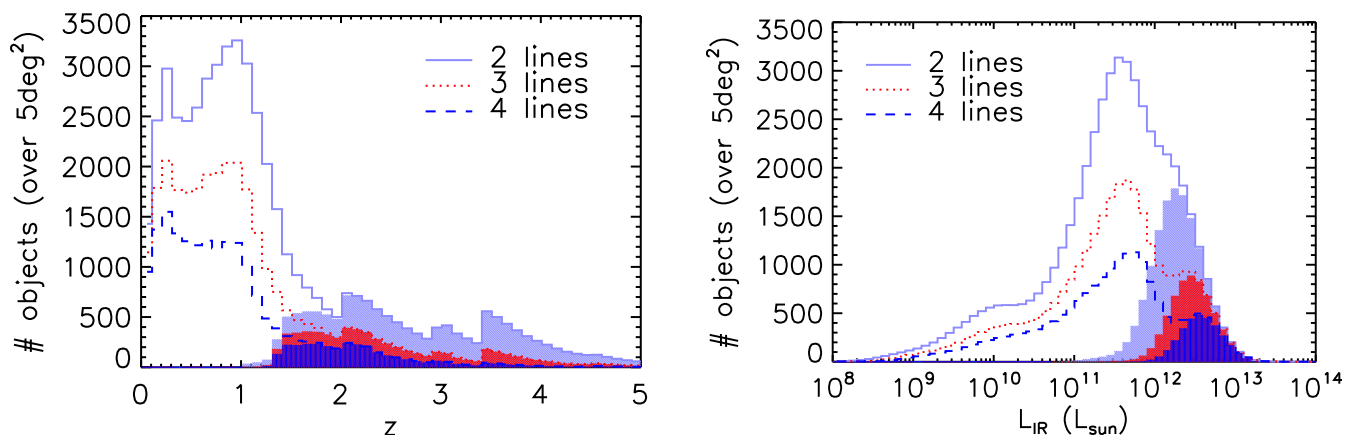




**Figure 7.** Examples of the predicted redshift distributions of AGNs detected by SPICA/SAFARI for a 1 h exposure per FoV. The different colours identify the spectral bands.



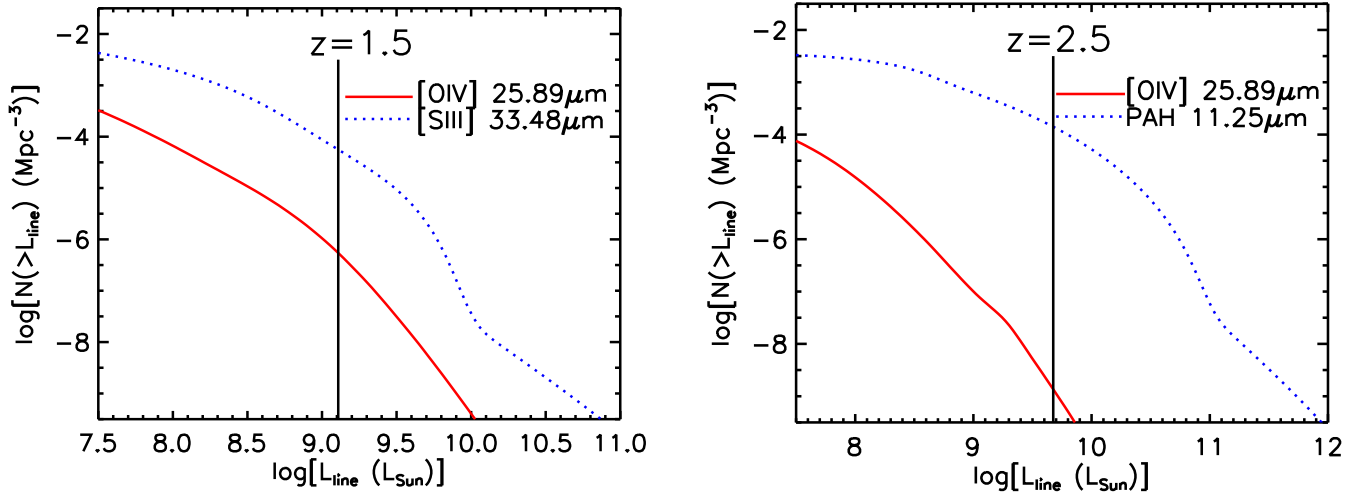
**Figure 8.** Contributions of different AGN populations to the integral counts in the [Ne v]14.32  $\mu\text{m}$ , [Ne v]24.31  $\mu\text{m}$  and [O IV]25.89  $\mu\text{m}$  lines. Solid black line: AGNs associated with late-type galaxies plus optically selected AGNs; dotted red line: AGN associated with unlensed protospheroids; dashed blue line: AGNs associated with strongly gravitationally lensed protospheroids. The vertical lines correspond to the average value of the detection limits ( $3.7 \times 10^{-19} \text{ W m}^{-2}$ ,  $3.4 \times 10^{-19} \text{ W m}^{-2}$ ,  $2.9 \times 10^{-19} \text{ W m}^{-2}$ ) for the three bands (34–60  $\mu\text{m}$ , 60–110  $\mu\text{m}$ , 110–210  $\mu\text{m}$ ) for 1 h exposure per FoV.



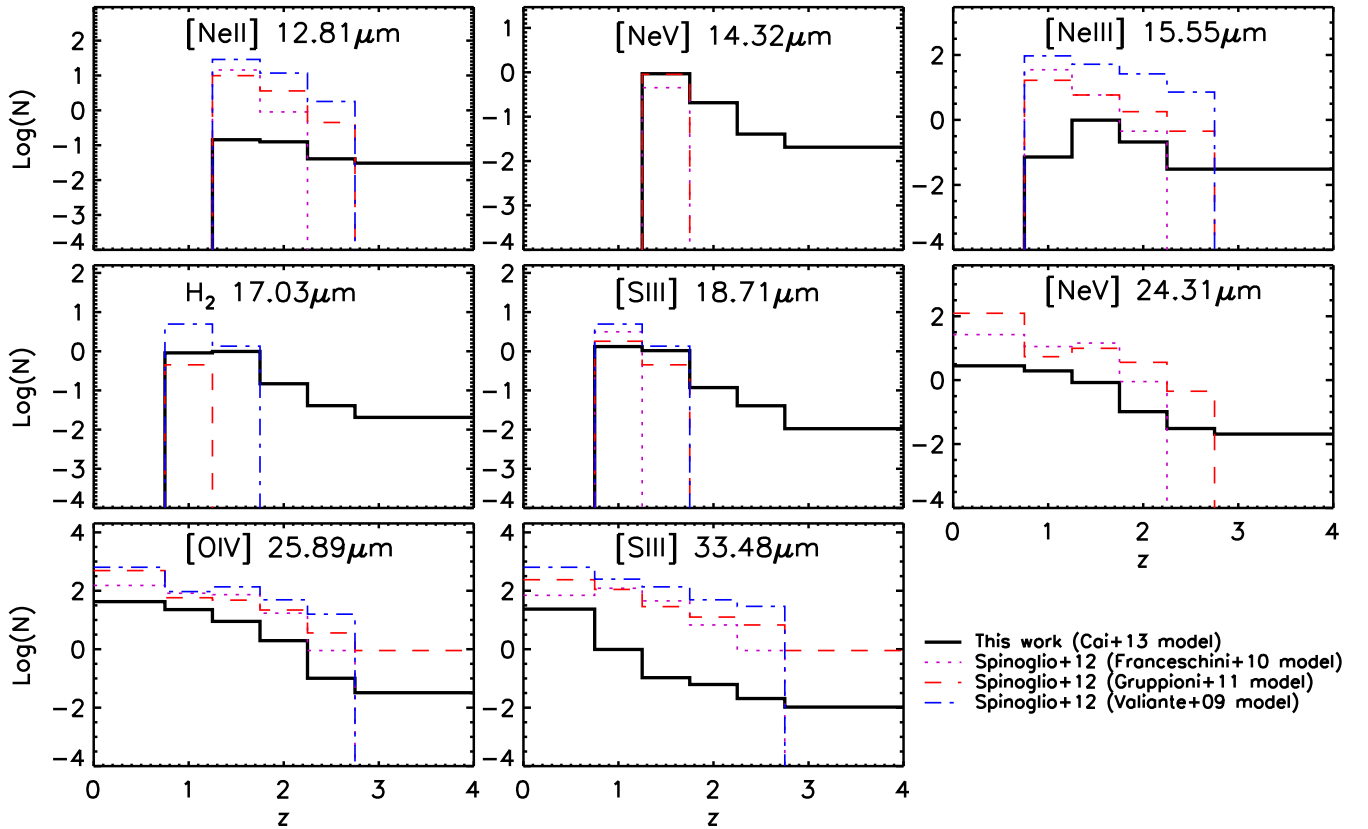
**Figure 9.** Predicted redshift (left) and total IR luminosity (right) distributions of galaxies (starburst plus AGN components) detectable in two (cyan histogram), three (red) and four (blue) spectral lines, by a SPICA/SAFARI survey covering 5  $\text{deg}^2$  in 1 h integration/FoV (taking into account both the spectral lines of this sample and those studied in Bonato et al. 2014). The shaded areas represent the contributions of protospheroids.

On the other hand, the approach adopted in this paper requires bolometric corrections, endowed with a substantial uncertainty, to derive the relationships between line and bolometric luminosities. We are now working on a semi-empirical approach, based on SED decomposition (see Berta et al. 2013; Delvecchio et al. 2014) of local (Tommasin et al. 2008, 2010) and high- $z$  samples of AGN (from

*Herschel* PACS Evolutionary Probe; Lutz et al. 2011), considering nuclear observational data in the MIR for estimating the  $L_{\text{line}}/L_{\text{bol}}$  relations and the AGN accretion luminosity function of Delvecchio et al. (2014) to derive expected detection numbers. The results of that work will be published in a forthcoming paper (Grupponi et al. in preparation).



**Figure 10.** Comparison of the cumulative luminosity functions at two redshifts of the brightest AGN line ([O IV]25.89  $\mu\text{m}$ ) with those of two bright starburst lines. The dotted curves show the contributions of strongly lensed sources to the total luminosity functions (solid curves). At each redshift the vertical line corresponds to the minimum luminosity detectable by a survey with 1 h integration/FoV.



**Figure 11.** Redshift distributions of AGNs detectable in each of the eight lines by a SPICA/SAFARI survey covering  $0.5 \text{ deg}^2$  with 1 h integration/FoV. The predictions of our model (black solid lines) are compared with those of the three models used by Spinoglio et al. (2012); see the legend on the bottom right.

## 6 OBSERVATION STRATEGY

Quantitative predictions on the number of detections of AGNs and of galaxies as a whole for a survey of  $0.5 \text{ deg}^2$  and different integration times per FoV are given in Table 5. As illustrated by Fig. 6, the integral counts have a slope flatter than two at and below the detection limit for 1 h integration/FoV. This means that the number of detections at fixed observing time increases more extending the

survey area than going deeper. This is particularly true for the three typical AGN lines (cf. Fig. 8), due to the fact that the AGN contribution to the bolometric luminosity is higher at higher luminosities. For a survey of  $0.5 \text{ deg}^2$  and 1 h integration/FoV the detection rate of AGNs is quite low and the predicted redshift distributions sink down at  $z \gtrsim 1$  (cf. Fig. 7). To investigate with sufficient statistics the galaxy–AGN co-evolution a deeper or a wider area survey is necessary. For example, a survey of 1 h/FoV over  $5 \text{ deg}^2$  would

**Table 4.** Numbers of  $5\sigma$  AGN detections by SPICA/SAFARI for a survey covering  $0.5 \text{ deg}^2$  in 1 h integration/FoV, as predicted by the present model (second column, in boldface) compared with those given (only for eight lines) by Spinoglio et al. (2012) for the three models used by them.

Spectral line	This work		Spinoglio et al. (2012)	
	Cai et al. (2013)	Franceschini et al. (2010)	Gruppioni et al. (2011)	Valiante et al. (2009)
[Ne II]12.81 $\mu\text{m}$	<b>0.3</b>	15	14	42
[Ne V]14.32 $\mu\text{m}$	<b>1</b>	0.4	1	–
[Ne III]15.55 $\mu\text{m}$	<b>1</b>	41	25	179
H <sub>2</sub> 17.03 $\mu\text{m}$	<b>2</b>	0	0.4	6
[S III]18.71 $\mu\text{m}$	<b>3</b>	3	2	6
[Ne V]24.31 $\mu\text{m}$	<b>5</b>	53	143	–
[O IV]25.89 $\mu\text{m}$	<b>76</b>	326	623	938
[S III]33.48 $\mu\text{m}$	<b>25</b>	244	400	1103

**Table 5.** AGN and total (in parenthesis) detections obtainable by SPICA-SAFARI surveys covering  $0.5 \text{ deg}^2$  in 1 h integration/FoV (2nd column) and 10 h integration/FoV (3rd column).

Spectral line	$t = 1\text{h}$	$t = 10\text{h}$
	$0.5 \text{ deg}^2$	$0.5 \text{ deg}^2$
PAH6.2 $\mu\text{m}$	0 (45)	0 (280)
PAH7.7 $\mu\text{m}$	0 (1578)	0 (4183)
PAH8.6 $\mu\text{m}$	0 (709)	0 (2661)
H <sub>2</sub> 9.66 $\mu\text{m}$	$4 \times 10^{-2}$ (4)	0.1 (33)
[S IV]10.49 $\mu\text{m}$	$1 \times 10^{-2}$ (31)	$3 \times 10^{-2}$ (156)
H <sub>2</sub> 12.28 $\mu\text{m}$	$3 \times 10^{-2}$ (10)	0.1 (74)
[Ne II]12.81 $\mu\text{m}$	0.3 (330)	3 (1490)
[Ne V]14.32 $\mu\text{m}$	1 (1)	5 (5)
[Ne III]15.55 $\mu\text{m}$	1 (94)	11 (627)
H <sub>2</sub> 17.03 $\mu\text{m}$	2 (26)	9 (236)
[S III]18.71 $\mu\text{m}$	3 (409)	18 (1979)
[Ne V]24.31 $\mu\text{m}$	5 (5)	48 (48)
[O IV]25.89 $\mu\text{m}$	76 (76)	427 (427)
[S III]33.48 $\mu\text{m}$	25 (2211)	147 (7826)

yield about 760 AGN detections in the [O IV]25.89  $\mu\text{m}$  line, with a total observing time of 4500 h. A deeper survey with the same observing time (exposure time of 10 h/FoV over  $0.5 \text{ deg}^2$ ) would give 427 AGN detections (see Table 5).

The blind spectroscopic survey may be complemented by follow-up observations of bright high- $z$  galaxies already discovered at (sub-)mm wavelengths over much larger areas. Fig. 12 shows, as an example, the SPICA/SAFARI exposure time per FoV necessary for a  $5\sigma$  detection of the three typical AGN lines at  $z = 1.5$  and  $2.5$  as a function of the AGN bolometric luminosity. When South Pole Telescope (SPT) and *Herschel* survey data will be fully available, we will have samples of many hundreds of galaxies with either intrinsic or apparent (i.e. boosted by strong gravitational lensing; Negrello et al. 2010, 2014; Vieira et al. 2013) IR luminosities larger than  $10^{13} L_{\odot}$ . These IR luminosities correspond to  $\text{SFR} > 10^3 M_{\odot} \text{ yr}^{-1}$  (Kennicutt & Evans 2012). Comparing the stellar with the halo mass function Lapi et al. (2011) estimated that the active star formation phase lasts for  $\sim 0.7 \text{ Gyr}$ . Thus, objects that bright are the likely progenitors of spheroidal galaxies with stellar masses  $M_{*} \gtrsim 7 \times 10^{11} / \mu M_{\odot}$ ,  $\mu$  being the gravitational magnification factor. As shown by Fig. 12, SPICA/SAFARI can detect, in 10 h, the [O IV]25.89  $\mu\text{m}$  line from an AGN with (real or apparent) bolometric luminosity of  $\simeq 3 \times 10^{12} L_{\odot}$  at  $z = 1.5$  and of  $\simeq 2 \times 10^{13} L_{\odot}$  at  $z = 2.5$ . For Eddington limited accretion, these luminosities correspond to black hole masses of  $\simeq 9 \times 10^7 / \mu M_{\odot}$  and  $\simeq 6 \times 10^8 / \mu M_{\odot}$ . For

comparison, from the black hole/stellar mass correlation (Kormendy & Ho 2013) we get that the final black hole masses associated with spheroidal galaxies with stellar masses of  $\simeq 10^{11}$  and  $\simeq 10^{12} M_{\odot}$  are  $\simeq 5 \times 10^8$  and  $\simeq 7 \times 10^9 M_{\odot}$ , respectively. This means that pointed SPICA/SAFARI observations can allow us to investigate early phases of the galaxy/AGN co-evolution, when the black hole mass was one or even two orders of magnitude lower than the final one.

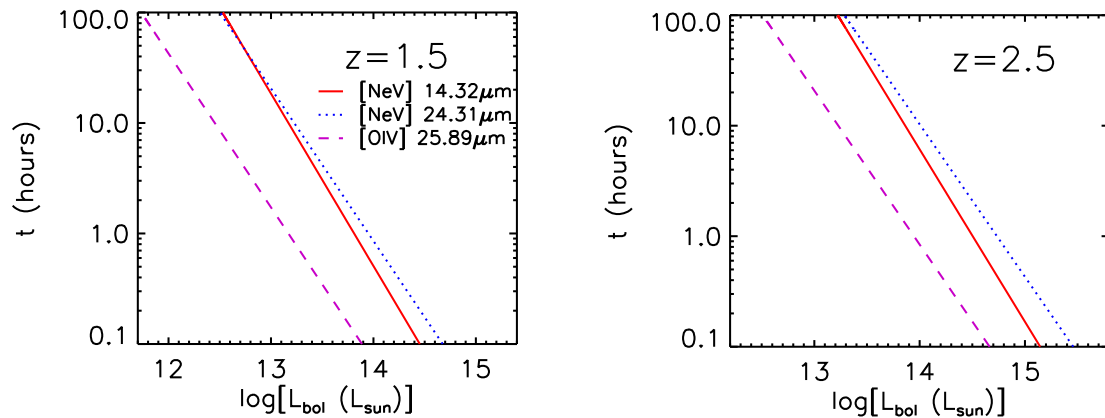
## 7 CONCLUSIONS

We have improved over earlier estimates of redshift-dependent luminosity functions of IR lines detectable by SPICA/SAFARI by building a model that deals in a self-consistent way with emission of galaxies as a whole, including both the starburst and the AGN component. For protospheroidal galaxies, which dominate the cosmic SFR at  $z \gtrsim 1.5$ , we have derived analytic formulae giving the probability that an object at redshift  $z$  has a starburst luminosity  $L_{*}$  or an AGN luminosity  $L_{\bullet}$  given the total luminosity of  $L_{\text{tot}} = L_{*} + L_{\bullet}$ . Each protospheroid of luminosity  $L_{\text{tot}}$  was assigned IR luminosities of the starburst and of the AGN component based on the above probability distributions. The association of the AGN component to late-type galaxies was made on the basis of the observed correlation between SFR and accretion rate.

The relationships between line and IR luminosities of the starburst component derived by Bonato et al. (2014) have been updated whenever new data have become available in the meantime and additional lines have been taken into account. New relationships between line and AGN bolometric luminosities have been derived.

These ingredients were used to work out predictions for the source counts in 11 mid-/far-IR emission lines partially or entirely excited by AGN activity, as well as in the 6.2, 7.7 and 8.6  $\mu\text{m}$  PAH lines, overlooked by Bonato et al. (2014). The expected outcome of a survey covering  $5 \text{ deg}^2$  with a 1 h integration/FoV is specifically discussed. Taking into account also the additional nine lines, with negligible AGN contribution, investigated by Bonato et al. (2014), i.e. a total of 23 IR lines, we have estimated the number of objects detectable in at least two, three, four, five and six lines. We estimate that such survey will yield about 760 AGN detections in the [O IV]25.89  $\mu\text{m}$  line. A deeper survey (10 h/FoV over  $0.5 \text{ deg}^2$ ) requiring the same observing time (4500 h) will yield about  $\sim 430$  detections in the same line.

The AGN contribution to the detectability of galaxies as a whole is minor. We show that this is due to the combination of two factors: the rarity of bright AGNs, due to their short lifetime, and the relatively low (compared to starbursts) line to bolometric luminosity ratios.



**Figure 12.** SPICA/SAFARI exposure time per FoV required for a  $5\sigma$  line detection of the three typical AGN lines as a function of the AGN bolometric luminosity for  $z = 1.5$  (left) and  $2.5$  (right).

We also recommend pointed observations of the brightest (either because they are endowed with extreme SFRs or because luminosities are strongly magnified by gravitational lensing) galaxies previously detected by large area surveys such as those by *Herschel* and by the SPT. Such observations will allow us to investigate early phases of the galaxy/AGN co-evolution when the black hole mass was much lower than the final one, given by the present-day correlation with the stellar mass.

## ACKNOWLEDGEMENTS

We are grateful to the anonymous referee for many constructive comments that helped us in improving this paper and to Professor Kotaro Kohno and Professor Takehiko Wada for pointing out the importance of the 6.2, 7.7 and 8.6  $\mu\text{m}$  PAH lines, overlooked by our previous study. We acknowledge financial support from ASI/INAF Agreement 2014-024-R.0 for the Planck LFI Activity of Phase E2 and from PRIN INAF 2012, project ‘Looking into the dust-obscured phase of galaxy formation through cosmic zoom lenses in the *Herschel* Astrophysical Large Area Survey’.

## REFERENCES

Alexander D. M., Hickox R. C., 2012, *New Astron. Rev.*, 56, 93  
 Bernard-Salas J. et al., 2009, *ApJS*, 184, 230  
 Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, *MNRAS*, 404, 2087  
 Berta S. et al., 2013, *A&A*, 551, A100  
 Bonato M. et al., 2014, *MNRAS*, 438, 2547  
 Cai Z.-Y. et al., 2013, *ApJ*, 768, 21  
 Chen C.-T. J., Hickox R. C., 2014, in Micaelian A. M., Sanders D. B., Proc. IAU Symp. 304, Multiwavelength AGN Surveys and Studies. Cambridge Univ. Press, Cambridge, p. 302  
 Chen C.-T. J. et al., 2013, *ApJ*, 773, 3  
 Cormier D. et al., 2012, *A&A*, 548, A20  
 Croom S. M. et al., 2009, *MNRAS*, 399, 1755  
 Delvecchio I. et al., 2014, *MNRAS*, 439, 2736  
 Diamond-Stanic A. M., Rieke G. H., 2012, *ApJ*, 746, 168  
 Dudik R. P., Satyapal S., Marcu D., 2009, *ApJ*, 691, 1501  
 Farrah D. et al., 2007, *ApJ*, 667, 149  
 Fiolet N. et al., 2010, *A&A*, 524, A33  
 Franceschini A., Rodighiero G., Vaccari M., Berta S., Marchetti L., Mainetti G., 2010, *A&A*, 517, A74  
 Gandhi P., Horst H., Smette A., Hönig S., Comastri A., Gilli R., Vignali C., Duschl W., 2009, *A&A*, 502, 457

Gorjian V., Werner M. W., Jarrett T. H., Cole D. M., Ressler M. E., 2004, *ApJ*, 605, 156  
 Goulding A. D., Alexander D. M., 2009, *MNRAS*, 398, 1165  
 Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2004, *ApJ*, 600, 580  
 Gruppioni C., Pozzi F., Zamorani G., Vignali C., 2011, *MNRAS*, 416, 70  
 Hao L., Wu Y., Charmandaris V., Spoon H. W. W., Bernard-Salas J., Devost D., Lebouteiller V., Houck J. R., 2009, *ApJ*, 704, 1159  
 Hatziminaoglou E. et al., 2010, *A&A*, 518, L33  
 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  
 Hill A. R., Gallagher S. C., Deo R. P., Peeters E., Richards G. T., 2014, *MNRAS*, 438, 2317  
 Hönig S. F., Kishimoto M., Gandhi P., Smette A., Asmus D., Duschl W., Polletta M., Weigelt G., 2010, *A&A*, 515, A23  
 Kennicutt R. C., Evans N. J., 2012, *ARA&A*, 50, 531  
 Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511  
 Lapi A., Shankar F., Mao J., Granato G. L., Silva L., De Zotti G., Danese L., 2006, *ApJ*, 650, 42  
 Lapi A. et al., 2011, *ApJ*, 742, 24  
 Lapi A., Negrello M., González-Nuevo J., Cai Z.-Y., De Zotti G., Danese L., 2012, *ApJ*, 755, 46  
 Lapi A., Raimundo S., Aversa R., Cai Z.-Y., Negrello M., Celotti A., De Zotti G., Danese L., 2014, *ApJ*, 782, 69  
 Lutz D. et al., 2008, *ApJ*, 684, 853  
 Lutz D. et al., 2011, *A&A*, 532, A90  
 Mao J., Lapi A., Granato G. L., de Zotti G., Danese L., 2007, *ApJ*, 667, 655  
 Meléndez M. et al., 2008, *ApJ*, 682, 94  
 Mullaney J. R., Alexander D. M., Goulding A. D., Hickox R. C., 2011, *MNRAS*, 414, 1082  
 Mullaney J. R. et al., 2012a, *MNRAS*, 419, 95  
 Mullaney J. R. et al., 2012b, *ApJ*, 753, L30  
 Negrello M., Perrotta F., González-Nuevo J., Silva L., de Zotti G., Granato G. L., Baccigalupi C., Danese L., 2007, *MNRAS*, 377, 1557  
 Negrello M. et al., 2010, *Science*, 330, 800  
 Negrello M. et al., 2014, *MNRAS*, 440, 1999  
 Netzer H. et al., 2007, *ApJ*, 666, 806  
 O’Dowd M. J. et al., 2009, *ApJ*, 705, 885  
 O’Dowd M. J. et al., 2011, *ApJ*, 741, 79  
 Pereira-Santaella M., Alonso-Herrero A., Rieke G. H., Colina L., Díaz-Santos T., Smith J.-D. T., Pérez-González P. G., Engelbracht C. W., 2010, *ApJS*, 188, 447  
 Planck Collaboration XVI, 2013, preprint ([arXiv:1303.5076](https://arxiv.org/abs/1303.5076))  
 Richards G. T. et al., 2006, *ApJS*, 166, 470  
 Roelfsema P. et al., 2012, *Proc. SPIE*, 8442, 84420R  
 Rosario D. J. et al., 2012, *A&A*, 545, A45  
 Roussel H. et al., 2006, *ApJ*, 646, 841

- Sanders D. B., Mirabel I. F., 1996, *ARA&A*, 34, 749  
Sani E., Lutz D., Risaliti G., Netzer H., Gallo L. C., Trakhtenbrot B., Sturm E., Boller T., 2010, *MNRAS*, 403, 1246  
Satyapal S., Vega D., Dudik R. P., Abel N. P., Heckman T., 2008, *ApJ*, 677, 926  
Satyapal S., Böker T., Mcalpine W., Gliozzi M., Abel N. P., Heckman T., 2009, *ApJ*, 704, 439  
Schweitzer M. et al., 2006, *ApJ*, 649, 79  
Serjeant S., Hatziminaoglou E., 2009, *MNRAS*, 397, 265  
Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119  
Spinoglio L., Dasyra K. M., Franceschini A., Gruppioni C., Valiante E., Isaak K., 2012, *ApJ*, 745, 171  
Sturm E., Lutz D., Verma A., Netzer H., Sternberg A., Moorwood A. F. M., Oliva E., Genzel R., 2002, *A&A*, 393, 821  
Tommasin S., Spinoglio L., Malkan M. A., Smith H., González-Alfonso E., Charmandaris V., 2008, *ApJ*, 676, 836  
Tommasin S., Spinoglio L., Malkan M. A., Fazio G., 2010, *ApJ*, 709, 1257  
Valiante E., Lutz D., Sturm E., Genzel R., Chapin E. L., 2009, *ApJ*, 701, 1814  
Veilleux S. et al., 2009, *ApJS*, 182, 628  
Vieira J. D. et al., 2013, *Nature*, 495, 344  
Weaver K. A. et al., 2010, *ApJ*, 716, 1151

This paper has been typeset from a  $\text{\TeX/L\TeX}$  file prepared by the author.