

This is the final peer-reviewed accepted manuscript of:

Bisigello, L., Gruppioni, C., Calura, F., Feltre, A., Pozzi, F., Vignali, C., . . . Wang, L. (2021). Simulating infrared spectro-photometric surveys with a Spritz. Publications of the Astronomical Society of Australia, 38, E064.

The final published version is available online at:
<http://dx.doi.org/10.1017/pasa.2021.57>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

Simulating infrared spectro-photometric surveys with a SPRITZ

Bisigello, L.^{1*}, Gruppioni, C.¹, Calura, F.¹, Feltre, A.¹, Pozzi, F.^{1,2}, Vignali, C.^{2,1}, Barchiesi L.^{1,2}, Rodighiero, G.^{3,4}, Negrello, M.⁵, Carrera, F.J.⁶, Dasyra, K.M.^{7,8}, Fernández-Ontiveros, J.A.^{9,10}, Giard, M.¹¹, Hatziminaoglou, E.¹², Kaneda, H.¹³, Lusso, E.^{14,15}, Pereira-Santaella, M.¹⁶, Pérez González, P.G.¹⁶, Ricci, C.^{17,18}, Schaerer, D.¹⁹, Spinoglio, L.⁹ and Wang, L.²⁰

¹INAF Osservatorio di Astrofisica e Scienza dello Spazio, via Gobetti 93/3, I-40129 Bologna, Italy

²Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, I-40129 Bologna, Italy

³Dipartimento di Fisica e Astronomia, Università di Padova, Vicolo dell'Osservatorio, 3, I-35122, Padova, Italy

⁴INAF Osservatorio Astronomico di Padova, vicolo dell'Osservatorio 5, I-35122 Padova, Italy

⁵School of Physics and Astronomy, Cardiff University, The Parade, Cardiff CF24 3AA, UK

⁶Instituto de Física de Cantabria (CSIC-U. Cantabria), Avenida de los Castros, 39005 Santander, Spain

⁷Dep. of Astrophysics, Astronomy & Mechanics, Faculty of Physics, National and Kapodistrian University of Athens, Panepistimiopolis, Zografou 15784, Greece

⁸National Observatory of Athens, Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, Penteli 15236, Athens, Greece

⁹Istituto di Astrofisica e Planetologia Spaziali (INAF-IAPS), Via Fosso del Cavaliere 100, Roma I-00133, Italy

¹⁰Centro de Estudios de Física del Cosmos de Aragón, Unidad Asociada al CSIC, Plaza San Juan 1, E-44001 Teruel, Spain

¹¹Institut de Recherche en Astrophysique et Planétologie, Toulouse (CNRS-INSU), France

¹²European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching bei München, Germany

¹³Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan

¹⁴Dipartimento di Fisica e Astronomia, Università di Firenze, via G. Sansone 1, 50019 Sesto Fiorentino, Firenze, Italy

¹⁵ INAF – Osservatorio Astrofisico di Arcetri, 50125 Florence, Italy

¹⁶Centro de Astrobiología (CSIC-INTA), Ctra. de Ajalvir, Km 4, 28850, Torrejón de Ardoz, Madrid, Spain

¹⁷Núcleo de Astronomía de la Facultad de Ingeniería, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago 22, Chile

¹⁸Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People's Republic of China

¹⁹Observatoire de Genève, Département d'Astronomie, Université de Genève, 51 Chemin Pegasi, 1290 Versoix, Switzerland

²⁰SRON Netherlands Institute for Space Research, Landleven 12, 9747 AD, Groningen, The Netherlands

Abstract

Mid- and far-infrared (IR) photometric and spectroscopic observations are fundamental to a full understanding of the dust-obscured Universe and the evolution of both star formation and black hole accretion in galaxies. In this work, using the specifications of the SPace Infrared telescope for Cosmology and Astrophysics (SPICA) as a baseline, we investigate the capability to study the dust-obscured Universe of mid- and far-IR photometry at 34 and 70 μm and low-resolution spectroscopy at 17-36 μm using the state-of-the-art Spectro-Photometric Realisations of Infrared-selected Targets at all- z (SPRITZ) simulation. This investigation is also compared to the expected performance of the Origins Space Telescope and the Galaxy Evolution Probe. The photometric view of the Universe of a SPICA-like mission could cover not only bright objects (e.g. $L_{IR} > 10^{12} L_{\odot}$) up to $z=10$, but also normal galaxies ($L_{IR} < 10^{11} L_{\odot}$) up to $z\sim 4$. At the same time, the spectroscopic observations of such mission could also allow us to estimate the redshifts and study the physical properties for thousands of star-forming galaxies and active galactic nuclei by observing the polycyclic aromatic hydrocarbons and a large set of IR nebular emission lines. In this way, a cold, 2.5-m size space telescope with spectro-photometric capability analogous to SPICA, could provide us with a complete three-dimensional (i.e. images and integrated spectra) view of the dust-obscured Universe and the physics governing galaxy evolution up to $z\sim 4$.

Keywords: galaxies: active – galaxies: evolution – galaxies: star formation – infrared: galaxies – techniques: spectroscopic – techniques: photometric

Preface

The articles of this special issue focus on some of the major scientific questions that a future IR observatory will be able to address. We adopt the SPace Infrared telescope for Cosmology and Astrophysics (SPICA) design as a baseline to demonstrate how to achieve the major scientific goals in the fields of galaxy evolution, Galactic star formation and protoplanetary disks formation and evolution. The studies developed for the SPICA mission serve as a reference for future work in the field, even though the mission proposal has been cancelled by ESA from its M5 competition. The mission concept of SPICA employs a 2.5 m telescope, actively cooled to below ~ 8 K, and a suite of mid- to far-IR spectrometers and photometric cameras, equipped with state-of-the-art detectors (Roelfsema et al., 2018). In particular, the SPICA Far-Infrared Instrument (SAFARI) is a grating spectrograph with low ($R \sim 200\text{--}300$) and medium-resolution ($R \sim 3\,000\text{--}11\,000$) observing modes instantaneously covering the $35\text{--}210\ \mu\text{m}$ wavelength range. The SPICA Mid-Infrared Instrument (SMI) has three operating modes: a large field of view ($10' \times 12'$) low-resolution $17\text{--}36\ \mu\text{m}$ imaging spectroscopic ($R \sim 50\text{--}120$) mode and photometric camera at $34\ \mu\text{m}$ (SMI-LR), a medium-resolution ($R \sim 1\,300\text{--}2\,300$) grating spectrometer covering wavelengths of $18\text{--}36\ \mu\text{m}$ (SMI-MR), and a high-resolution echelle module ($R \sim 29\,000$) for the $10\text{--}18\ \mu\text{m}$ domain (SMI-HR). Finally, BBOP, a large field of view ($2'.6 \times 2'.6$), three-channel (70, 200 and $350\ \mu\text{m}$) polarimetric camera complements the science payload.

1 INTRODUCTION

As new stars form in dusty environments, infrared (IR) observations are key to studying obscured star formation and, in turn, achieving a better understanding of galaxy formation and evolution. This is particularly important around the very active period known as the Cosmic Noon (i.e. $z \sim 2$; Madau & Dickinson, 2014), where dust absorbs and re-radiates in the IR, on average, 80% of the ultra-violet (UV) and optical radiation (Reddy et al., 2012).

At $z > 3$ our knowledge of the on-going star formation is mainly based on UV rest-frame observations (e.g. Bouwens et al., 2015; Livermore et al., 2017, 2018; Oesch et al., 2018), given the large availability of optical and near-IR observatories such as the Hubble Space Telescope (*HST*). UV observations have shown that dust corrections at $z > 3$ may be relatively small (e.g. Bouwens et al., 2016; Dunlop et al., 2017). Conversely, studies based on emission lines have shown that UV-based star-formation rate may be underestimated for galaxies with high level of star formation, due to an increased impor-

tance of the dusty stellar populations in these galaxies (e.g. Reddy et al., 2015). Similarly, recent observations of galaxies at $z > 6$ have shown the presence of a large quantity of dust (e.g. $\sim 10^6 M_{\odot}$; Tamura et al., 2019), which may be more concentrated and warm than clouds in local galaxies, increasing the UV attenuation (e.g. Sommovigo et al., 2020)

In addition, when studying the cosmic star-formation-rate density (SFRD) of the Universe, there are some tensions between the results obtained from UV observations (e.g. Schenker et al., 2013; Bouwens et al., 2015) and results based on IR or radio data (e.g. Magnelli et al., 2013; Novak et al., 2017; Gruppioni et al., 2020), with the former indicating a steeper redshift evolution (i.e. lower SFRD at high- z) than the latter. This tension is at least partially due to UV observations not being representative of the whole galaxy population (Bouwens et al., 2020). Indeed, recent observations have shown the non-negligible presence of a dusty and optically-dark galaxy population (e.g. Franco et al., 2018; Wang et al., 2019; Talia et al., 2021), which contributes to 17% of the star-formation rate density at $z \sim 5$, as estimated by Gruppioni et al. (2020) from a sample of blindly detected far-IR galaxies in the ALMA Large Program to INvestigate survey (ALPINE; Le Fèvre et al., 2020).

Due to the capabilities of past and current IR and sub-millimetre observatories, e.g. *Herschel*, *Spitzer* and the Atacama Large Millimetre/submillimetre Array (ALMA), current observations have been limited to the brightest IR galaxies or to very small fields. A new IR telescope with both high sensitivity and large survey capability, such as a SPICA-like mission, the Origins Space Telescope¹ (OST; Leisawitz et al., 2019) or the Galaxy Evolution Probe (Glenn et al., 2018, GEP), is therefore essential to improving our understanding of the dust-obscured Universe at different epochs.

One of the primary science goals of SPICA was the detection and analysis of the dust-obscured phases of galaxy evolution, including activity from both star formation and accretion onto super massive black holes. Combined low-resolution (LR) spectroscopy and photometry at IR wavelengths could enable the detection of polycyclic aromatic hydrocarbons (PAHs) features, fine-structure lines as well as the dust continuum emission at mid-IR wavelengths. At these wavelengths the light is not heavily affected by dust absorption, as happens instead at UV and optical wavelengths, and, therefore, it allows the analysis of the evolutionary phases which are difficult to study at shorter wavelengths. In addition, the advantage of the spectro-photometric observations at mid- to far-IR wavelengths is the possibility to analyse the physics behind the processes responsible for the observed galaxy evolution and the environment in which they take place.

*laura.bisigello@inaf.it

¹<http://origins.ipac.caltech.edu/>

In this work we make use of the state-of-the-art Spectro-Photometric Realisations of Infrared-selected Targets at all- z simulation (SPRITZ; Bisigello et al., 2021, hereafter B21) that is particularly suited for predictions of IR surveys and includes both star-forming galaxies and active galactic nuclei (AGN). Using this simulation and following the works by Kaneda et al. (2017), Gruppioni et al. (2017) and Spinoglio et al. (2021), we aim at exploring the capability of a SPICA-like cold 2.5 m telescope to drive significant progress in our knowledge on galaxy evolution. Given the dismissal of the SPICA mission concept from the M5 competition, we expanded this work to include a comparison with the OST and the GEP, two NASA concepts for cryogenic IR observatory with 5.9m and 2.0m primary mirrors, respectively.

This paper is organised as follows. Section 2 introduces the SPRITZ simulation, the SPICA spectro-photometric surveys used as references and the OST and GEP surveys. Section 3 contains the different predictions for SPICA complemented, when possible, with OST and GEP expectations. We summarise our main findings in Section 4. Throughout the paper, we consider a Λ CDM cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$.

2 THE SPRITZ SIMULATION AND SPICA MOCK CATALOGUES

2.1 The SPRITZ simulation

To predict the galaxy population that could be observed by future IR missions, we created mock observations using the newly developed SPRITZ simulations (B21). All the mock catalogues considered in this paper are made publicly available².

Briefly, in SPRITZ all simulated galaxies are extracted from a set of observed luminosity functions (LFs) or galaxy stellar mass functions (GSMF):

- The *Herschel* LFs measured by Gruppioni et al. (2013) for different galaxy populations such as normal star-forming galaxies (hereafter spirals), starburst (SB) and two composite systems with an AGN contribution that is only visible in the mid-IR and, even if with some obscuration, in the X-rays (SF-AGN and SB-AGN). The first composite AGN population, i.e. SF-AGN, is common at $z=1-2$ and includes low-luminosity AGN, while the second one, i.e. SB-AGN, is common at $z>2$, hosts a bright, but heavily obscured AGN and has a high SFR (i.e. $\log_{10}(sSFR/yr) \sim 8$ in the simulation). All these LFs are described by a modified Schechter function (Saunders et al., 1990), whose parameter values are summarised in Table 1 in B21. At $z>3$ the LFs

are extrapolated assuming a constant characteristic luminosity and a decreasing number density at the knee as $\propto (1+z)^{k_\Phi}$, with $k_\Phi = -4$ to -1 . These values have been chosen to explore a large range of possible density evolution and are consistent with recent observations of the total IR LF observed at $z>3$ (Gruppioni et al., 2020).

- The obscured and un-obscured AGN (AGN2 and AGN1, respectively) LF, as derived by B21 using *Herschel* observations along with a set of far-ultraviolet observations up to $z=5$. The LF is described by a modified Schechter function and the evolution of the parameters describing the function at $z<5$ is extrapolated at higher redshift (see Table 1 in B21).
- The K-band LF of elliptical galaxies (Ell), taken from the average LF of Cirasuolo et al. (2007), Arnouts et al. (2007) and Beare et al. (2019). The LF is extrapolated for galaxies at $z>2$ by keeping the characteristic luminosity constant and with a number density at the knee decreasing as $\propto (1+z)^{-1}$.
- The GSMF of dwarf irregular galaxies (Irr) by Huertas-Company et al. (2016). For consistency with the *Herschel* LFs, at $z>3$ we extrapolated the GSMF, which is not well studied at these high- z , by keeping the characteristic mass constant and decreasing the number density at the knee as $\propto (1+z)^{k_\Phi}$, with k_Φ ranging from -4 to -1.

The different galaxy populations and the references for their LF or GSMF are reported in Table 1.

Mock fluxes for different current and future facilities, including SPICA, are derived by convolving the throughput of each filter with an empirical galaxy template associated to each galaxy population. We considered 35 empirical templates taken from Polletta et al. (2007); Rieke et al. (2009); Gruppioni et al. (2010) and Bianchi et al. (2018). We included these templates as they are the same used by Gruppioni et al. (2013) to derive the *Herschel* IR LF, except for dwarf galaxies which were not observed in large numbers by *Herschel*.

The main physical properties for each simulated galaxy, such as stellar mass, AGN fraction, gas metallicity and intrinsic luminosity at different wavelengths, are derived from the associated template or from a broad set of empirical relations (see B21, and references therein). The simulation also includes mock SMI and OST spectra, obtained considering the same galaxy templates used for deriving the simulated photometric data.

We added optical and IR nebular line emission to these templates, considering a set of empirical and theoretical relations (see B21, and references therein). All lines include star formation and, if present, the AGN contribution. The AGN component is limited to the contribution of the narrow-line gas-emitting regions. Therefore, as

²<http://spritz.oas.inaf.it/>

the contribution of the broad-line regions is missing, the fluxes of the permitted lines for the simulated unobscured AGN (AGN1) should be considered as lower limits. Similarly, the numbers of AGN1 with a detection of permitted lines should be considered as lower limits.

The large-scale structure is included in the simulation using the algorithm by Soneira & Peebles (1978) to extract positions based on a two-point correlation function. In particular, we assumed the power-law approximation by Limber (1953) for the angular two-point correlation function $w(\theta) = A_w \theta^{1-\gamma}$, with $\gamma=1.7$, as suggested by observations (e.g. Wang et al., 2013), and A_w evolving with the stellar mass (Wake et al., 2011; Hatfield et al., 2016) but not with redshift (B  thermin et al., 2015; Schreiber et al., 2015). We highlight that the assumptions made on the clustering have no impact on the results presented in this work, as this information is used, at the moment of writing, only to create mock images and we do not analyse source blending.

As detailed in B21, SPRITZ agrees well with a broad set of observations (LFs and number counts) at different wavelengths. In particular, the simulation is in agreement, within the simulation uncertainties, with the mid- and far-IR number counts ($\sim 24\text{--}600 \mu\text{m}$) and with the recent IR LF observed by Gruppioni et al. (2020) up to $z\sim 6$, showing the reliability of the extrapolation at $z>3$, at least at IR wavelengths. In addition, Gruppioni et al. (2020) found a general agreement between observed galaxies at $z\sim 6$ and the SED templates used to build SPRITZ, confirming the validity of the templates used, even at high- z .

Looking at other wavelengths, the part of the UV LF dominated by galaxies (i.e. AB magnitude $M_{1600\text{ }} > -23$) in SPRITZ is underestimated at $z>2$, indicating that a population of dust-poor galaxy may be missing in the simulation. The absence of such population should however not impact the results presented in this paper focused on future IR missions.

The bright-end of the X-ray LF of the simulation is in agreement with previous observational results up to $z=6$ (Aird et al., 2010, 2015; Miyaji et al., 2015; Vito et al., 2018; Wolf et al., 2021), above which no observations are at the moment available. The faint-end slope, on the contrary, is slightly overestimated and the difference between our simulation and observations increases with redshift, from a factor of four at $z\sim 0.5$ to a factor of ~ 15 at $z\sim 5$, but still within the uncertainties. This will be further investigated in a future work and it is expected to have an impact on the comparison between IR and X-ray fluxes presented in Sec. 3.1.3.

We can overall conclude that the SPRITZ simulation can reliably be used to perform SPICA, GEP and OST predictions. We refer to B21 for more information about SPRITZ.

Table 1 The different galaxy populations included in SPRITZ and the reference of the LF or GSMF from which their number densities are derived.

Population	Reference
Spiral	Gruppioni et al. (2013)
SB	Gruppioni et al. (2013)
SF-AGN	Gruppioni et al. (2013)
SB-AGN	Gruppioni et al. (2013)
AGN1	Gruppioni et al. (2013), B21
AGN2	Gruppioni et al. (2013), B21
Elliptical	Cirasuolo et al. (2007), Arnouts et al. (2007) and Beare et al. (2019)
Dwarf	Huertas-Company et al. (2016)

2.2 Simulated catalogues

In the next sections we describe in detail the simulated catalogues we created resembling six different surveys of SPICA, OST and GEP. To show the improvement of such surveys with respect to previous ones, we show in Figure 1 the comparison between the depth and area of the considered SPICA, OST and GEP surveys with the depth and area of past surveys observing around 30 and 70 μm (i.e. Lonsdale et al., 2003; Sanders et al., 2007; Magnelli et al., 2013).

In the same figure, we also include a comparison with two approved James Webb Space Telescope (JWST; Gardner et al., 2009) surveys: the Mid Infrared Instrument (MIRI Kendrew et al., 2015; Rieke et al., 2015; Wright et al., 2015) observations at 21.8 μm which is part of the Cosmic Evolution Early Release Science (CEERS³) survey and the MIRI 7.7 μm observations of the Medium JWST Advanced Deep Extragalactic Survey (JADES⁴). JWST and the three far-IR missions analysed in this work are complementary in wavelengths, as JWST covers $\lambda < 28.8 \mu\text{m}$ while SPICA, OST and GEP cover $\lambda > 20, 25$ and $10 \mu\text{m}$, respectively. At the same time, JWST is particularly suited for performing very deep mid-IR observations, as in imaging it can reach a 10σ depth of 7.09 μJy at 21.0 μm in 10^4 s (Glasse et al., 2015), but they are limited to small area in the sky given the small MIRI field-of-view ($74'' \times 113''$; Bouchet et al., 2015).

2.2.1 SPICA

In this work, we simulated two different spectrophotometric surveys, following two planned SPICA surveys (Table 2):

- an Ultra-Deep Survey (UDS) covering 1 deg² in 600 h, reaching down to 3 μJy (5σ) with SMI at 34 μm and 60 μJy (5σ) with B-BOP at 70 μm . In the same amount of time SMI-LR would simultaneously

³<https://ceers.github.io>

⁴<https://www.cosmos.esa.int/web/jwst-nirspec-gto/jades>

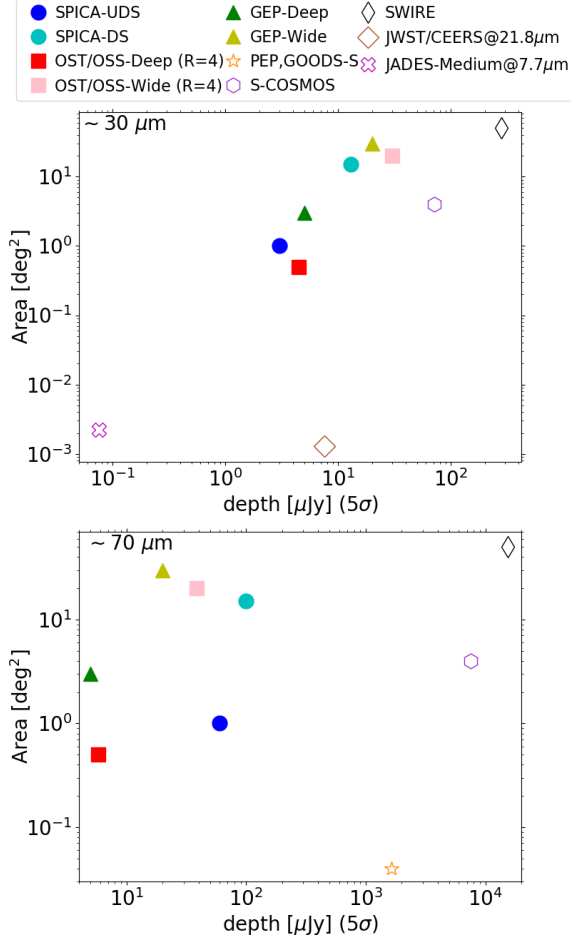


Figure 1. Comparison between area and depth covered by different surveys, observing around 30 (*top*) and 70 μm (*bottom*). Coloured points indicate the surveys considered in this work (see legend). Empty points indicate some past surveys at 24 and 70 μm . For comparison, we also include two future JWST surveys: the JWST/CEERS survey at 21.8 μm (*brown diamonds*) and the JWST/JADES medium survey at 7.7 μm (*magenta crosses*).

obtain spectra down to 39 μJy at 20 μm and 65 μJy at 30 μm (5σ).

- a deep survey (DS) covering 15 deg^2 in 600 h, reaching down to 13 μJy (5σ) with SMI at 34 μm and 100 μJy (5σ) with the B-BOP at 70 μm . In the same amount of time SMI-LR would simultaneously obtain spectra down to 150 μJy at 20 μm and 250 μJy at 30 μm (5σ).

Both spectroscopic limits are derived considering a low background level (i.e., 19 MJy/sr at 27 μm). Photometric confusion noise is not included in SPRITZ and corresponds at 5σ to 9 μJy and 0.25 mJy at 34 μm and 70 μm , respectively. The confusion noise may be broken with the help of the available spectroscopic data (Raymond et al., 2010), but it is necessary to consider with caution sources with fluxes similar or below the mentioned confusion noise, as they may need additional de-blending analysis.

Galaxies in the two mock catalogues are selected to have a signal-to-noise (S/N) larger than 3 either in the SMI or B-BOP photometric filters, similarly to what is done with real data to limit the number of false detection. In Figure 2 we report as an example a $10' \times 12'$ simulated image with SMI at 34 μm (*left panel*) and B-BOP at 70 μm (*central panel*). The chosen area corresponds to a single SMI pointing, whereas a B-BOP pointing covers a smaller area, i.e. $2'.6 \times 2'.6$. In the right panel we show a zoom-in of the B-BOP pointing with, superimposed, the SMI image. In this figure it is possible to appreciate the different size of the point-spread function (PSF), which has been approximated by a two-dimensional Gaussian. The PSF has a full-width-half-maximum of $3.5''$ at 34 μm and of $9''$ at 70 μm .

Table 2 5σ depths at four wavelengths corresponding to the two SPICA photometric filters and LR spectrograph at two reference wavelengths, as planned for the two spectro-photometric surveys considered in this work.

Instrument	λ [μm]	5σ depth [μJy]	
		UDS (1 deg^2)	DS (15 deg^2)
SMI	34	3	13
B-BOP	70	60	100
SMI-LR	20	39	150
	30	65	250

2.2.2 OST

The Origins Survey Spectrometer (OSS; Bradford et al., 2018) planned for OST consists of six channels covering between 25 and 589 μm with a resolution of $R=300$. In this work, we compared the two aforementioned SPICA surveys with the Deep and Wide surveys planned for OST (hereafter OST-Deep and OST-Wide), covering, respectively, 0.5 and 20 deg^2 with a 5σ depth of 39.4 and 262.5 μJy at 25-44 μm , at a resolution of $R=300$.

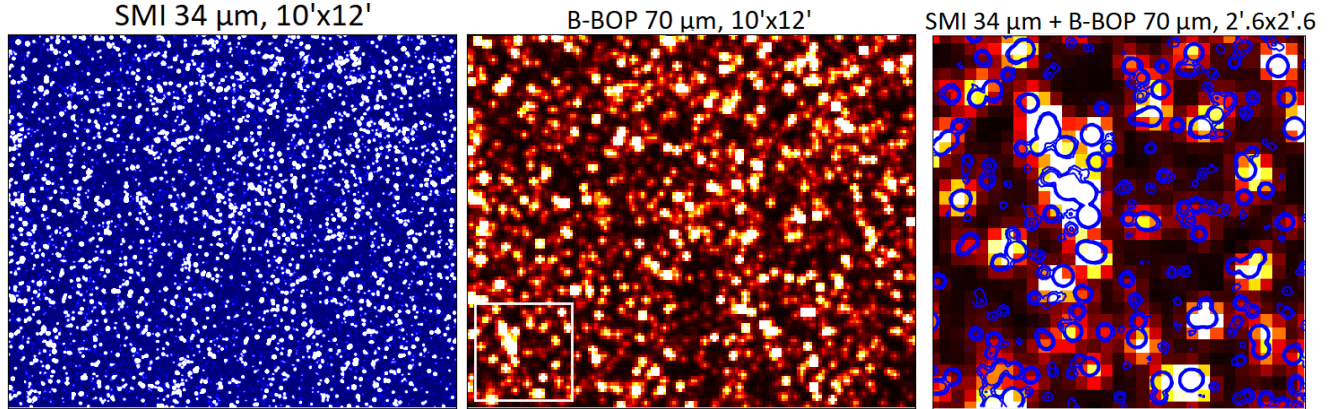


Figure 2. *Left:* Simulated SMI image at $34\ \mu\text{m}$ of a field of $10' \times 12'$, which corresponds to a single SMI pointing. *Centre:* Simulated B-BOP image at $70\ \mu\text{m}$ of the same field. The B-BOP field-of-view ($2'.6 \times 2'.6$) is shown for comparison in the bottom left (*white square*). *Right:* Zoom-in of the B-BOP image in the B-BOP field-of-view. Blue contours are the superimposed SMI image (5, 10 and 20σ). Each point corresponds to one galaxy or blends of several galaxies simulated with SPRITZ, depending on source blending, down to the noise level.

For comparison with the two SPICA photometric surveys, we also considered the photometric points derived considering each OST channel as a photometer, i.e. at a resolution of $R=4$ instead of $R=300$. This corresponds to a 5σ depth of 4.5 and $30.3\ \mu\text{Jy}$ at $25\text{-}44\ \mu\text{m}$ in the OST-Deep and OST-Wide survey, respectively. Taking into account the central wavelengths of the different filters, in the part of this work focused on results based on photometric data, we compared results of the OST/OSS Ch1 with the SPICA/SMI filter at $34\ \mu\text{m}$ and of OST/OSS Ch2 with the SPICA/B-BOP filter $70\ \mu\text{m}$. We instead considered all six OST/OSS channels when comparing the spectroscopic capability of the two instruments.

Confusion noise should not be a major problem in OST, given the large primary mirror, and it corresponds to $120\ \text{nJy}$ and $1.1\ \text{mJy}$ at 50 and $250\ \mu\text{m}$. As for SPICA, the available spectra can be used to de-blend galaxies and break the confusion noise (Raymond et al., 2010). The full list of OST/OSS channels and their expected observational depths are listed in Table 3.

Table 3 5σ depths at six wavelengths corresponding to the OST/OSS spectroscopic channels, as planned for the two spectroscopic surveys (i.e. OST-Deep and OST-Wide) considered in this work.

Channel	$\lambda\ [\mu\text{m}]$	5σ depth [μJy]			
		Deep ($0.5\ \text{deg}^2$)		Wide ($20\ \text{deg}^2$)	
		$R=4$	$R=300$	$R=4$	$R=300$
Ch1	25-44	4.5	39.4	30.3	262.5
Ch2	42-74	5.8	50.4	38.8	336.1
Ch3	71-124	8.4	72.4	55.8	483.1
Ch4	119-208	21.8	188.9	145.5	1260
Ch5	200-350	21.8	188.9	145.5	1260
Ch6	336-589	70.9	614.1	472.9	4100

2.2.3 GEP

GEP (Glenn et al., 2018) is a NASA concept of a 2 m cold (4 K) telescope with photometric and spectroscopic capability. The GEP Imager (GEP-I) is planned to have 23 bands covering $10\text{-}400\ \mu\text{m}$, with spectral resolution $R=8$ at $10\text{-}95\ \mu\text{m}$ and $R=3.5$ at $95\text{-}400\ \mu\text{m}$. Considering the central wavelengths of the different filters, in this work we compared results for the SPICA/SMI filter at $34\ \mu\text{m}$ and SPICA/B-BOP at $70\ \mu\text{m}$ with the GEP-I filters centred at ~ 33 (Band 10) and $73\ \mu\text{m}$ (Band 16).

At the moment of writing, there are four surveys planned:

- An all sky survey with a 5σ depth of $\sim 1\ \text{mJy}$.
- A survey covering $300\ \text{deg}^2$ with a 5σ depth of $\sim 50\ \mu\text{Jy}$.
- A survey covering $30\ \text{deg}^2$ with a 5σ depth of $\sim 20\ \mu\text{Jy}$ (hereafter GEP-Wide).
- A survey covering $3\ \text{deg}^2$ with a 5σ depth of $\sim 5\ \mu\text{Jy}$ (hereafter GEP-Deep).

For a better comparison with the SPICA surveys, in this work we considered only the last two surveys. Given the size of the primary mirror, GEP is expected to reach the confusion noise level in imaging at approximately $70\ \mu\text{m}$ in the two considered surveys (Glenn et al., 2019).

In addition, the GEP Spectrometer (GEP-S) concept consists of four gratings covering from 24 to $193\ \mu\text{m}$ with $R=200$. However, current plans for GEP-S include only pointed observations (Glenn et al., 2019). Therefore, we decided to limit the comparison of GEP and SPICA only to their photometric capability.

3 PREDICTIONS

3.1 Galaxy populations expected from photometric surveys

In this section, we give an overview of the physical properties of the galaxy populations that are expected to be detected by the considered photometric surveys, as derived from SPRITZ. We concentrate on the nature of such objects, but we highlight that, for their analysis, it may be necessary to rely on additional spectroscopic data (see Sec. 3.2) or a larger set of photometric data including also other wavelengths.

3.1.1 IR luminosities

We start by investigating the IR luminosity of the galaxies that we expect to observe with a SPICA-like mission (Figure 3), compared to predictions for OST (Figure 4) and GEP (Figure 5).

In both the UDS and DS, SMI would be able to detect galaxies with IR luminosities at the knee of the luminosity function, as derived by Gruppioni et al. (2013), up to $z=6$. This would be possible, at least for some objects, even up to $z=10$, while it would be limited to $z<4$ for B-BOP observations in both surveys.

The depth and area of the DS are ideal to statistically investigate the population of Ultra-luminous IR (ULIRG) and Hyper-luminous IR galaxies (HyLIRG), i.e. respectively $L_{IR} > 10^{12} L_{\odot}$ and $L_{IR} > 10^{13} L_{\odot}$. We indeed would expect to detect around $4\text{--}6 \times 10^4$ ULIRGs up to redshift $z=10$, showing a mean redshift in the range $z=2.8\text{--}3.5$ and $\sim 700\text{--}900$ HyLIRGs with a mean redshift $z=2.6\text{--}3.8$. For comparison, in the UDS, which covers only 1 deg^2 , we would expect to detect only a few tens (30-50) of HyLIRGs. For an area of 100 deg^2 and considering the same observational depth of the DS, which was not possible with SPICA but may be feasible with future IR missions, we would expect $\sim 5000\text{--}6000$ HyLIRGs with mean redshift $z=3.6\text{--}3.8$ and $3.5\text{--}4.7 \times 10^5$ ULIRGs with mean redshift $z=3.0\text{--}3.4$.

The uncertainties on the numbers of ULIRG and HyLIRG are due to the extrapolation of the IR LF at $z>3$, which impacts also the mean redshift of the simulated sample. In fact, if we consider $k_{\Phi} = -4$ instead of $k_{\Phi} = -1$, the average redshift decreases as there are fewer galaxies at $z>3$.

The predicted HyLIRGs number (i.e. $30\text{--}50 \text{ deg}^{-2}$) corresponds to a larger surface density than the one observed by Wang et al. (2021), which ranges between 5 and 18 HyLIRGs per deg^2 . However, these surface densities are consistent with each other, once we take into account the uncertainties on the bright-end of the IR *Herschel* LF by Gruppioni et al. (2013), which varies between 0.07 and 0.43 dex for $L_{IR} > 10^{13} L_{\odot}$, depending on the considered redshift. Even considering the more optimistic estimation of the HyLIRGs surface density, we need an area larger than $\sim 70 \text{ arcmin}^2$ to detect at

least one of them. This consideration shows the complementarity of missions such as SPICA, OST or GEP with respect to JWST, for which it is difficult to observe large areas, particularly in the mid-IR.

The galaxy population observed by a SPICA-like mission would not be limited to extreme IR galaxies. Indeed, with SMI at $34 \mu\text{m}$, we would expect to observe below the luminosity of the so-called Luminous IR galaxies (LIRGs, $L_{IR} = 10^{11} L_{\odot}$) up to $z=4.0$ and $z=2.5$ in the UDS and DS, respectively. Even if these observations would be mainly limited to $z<2.0$ at $70 \mu\text{m}$, they would lead to breakthrough results such as the unprecedented, direct measure of the faint-end slope of the luminosity function at these wavelengths.

The OST/OSS Ch1 depth in the OST-Deep survey is similar to the SMI depth at $34 \mu\text{m}$ in the SPICA UDS and, indeed, the IR luminosity of observed galaxies is expected to be similar. In the OST-Wide Survey the OST/OSS Ch1 is instead planned to be shallower than SMI at $34 \mu\text{m}$, i.e. $30.3 \mu\text{Jy}$ compared to $13 \mu\text{Jy}$. This limits the number of the detectable faint galaxies with respect to what was expected for SPICA. However, the planned observational depths for the OST/OSS Ch2 are deeper than the ones planned for SPICA/B-BOP at $70 \mu\text{m}$. For this reason, in the OST/OSS Ch2 we expect to observe galaxies below the ULIRG limit up to $z\sim 7.5$ in the OST-Deep Survey and up to $z\sim 5.5$ in the OST-Wide Survey.

GEP-I observational depths are expected to be of ~ 5 and $20 \mu\text{Jy}$, in the 3 and 30 deg^2 surveys, respectively. Such surveys are therefore deeper than both SPICA surveys at $70 \mu\text{m}$ and the SPICA DS at $34 \mu\text{m}$. For this reason, we expect to see galaxies with IR luminosity below the ULIRGs regime up to $z=5.5$ (5.0) and 8.5 (4.0) at 33 and $73 \mu\text{m}$ in the deepest (widest) considered GEP survey. However, we note that the number and the maximum observable redshift of galaxies with $L_{IR} < 10^{12} L_{\odot}$ detected by GEP-I may be underestimated, as we consider only two of the 23 filters available. A similar case does not hold for OST, as the most sensitive filters are those at the shortest wavelengths and they are the filters we considered in this work.

We also compared the 80% limits of the L_{IR} , derived using SPRITZ, of two samples selected in two JWST/MIRI filters: [F2100W] <21.7 (i.e. $7.5 \mu\text{Jy}$) and [F770W] <26.7 (i.e. $0.07 \mu\text{Jy}$). The two selections correspond to the magnitude limits of the CEERS and JADES surveys, respectively, in their two longest wavelength filters (see Sec. 2.2.1). For this comparison, photometric errors are not included in the JWST fluxes.

The first flux cut corresponds to a L_{IR} limit similar to the UDS and DS limits at $70 \mu\text{m}$, to the OST-Wide limit at $34 \mu\text{m}$ and to the GEP-Wide limit at $33 \mu\text{m}$, at least up to $z=3$, while it probes galaxies with brighter L_{IR} than the other SPICA, OST and GEP surveys. The second JWST flux threshold corresponds to a L_{IR} limit

much fainter than all the considered SPICA, OST and GEP surveys, showing the power of JWST to observe very faint galaxies.

The expected number of galaxies observed by JWST will be smaller than the number observed by the considered SPICA, OST and GEP surveys, given the small sky area covered by the two considered JWST surveys (i.e. ~ 4.6 and 8 arcmin² for the CEERS F2100W and for the JADES F770W observations, respectively). These small areas would also limit or preclude observations of very bright galaxies. In addition, JWST does not allow one to fully sample the far-IR dust bump at any redshift, but this is instead possible with OST and GEP, enabling the characterisation of the properties of the dust in high-redshift galaxies. Moreover, the same would have been possible with SPICA, thanks to the capabilities of the B-BOP instrument.

3.1.2 Redshift distribution

In Figure 3 we show the expected redshift distribution per unit redshift interval of the galaxies that would be detected in the SPICA UDS and DS with SMI at $34 \mu\text{m}$ and B-BOP at $70 \mu\text{m}$. In Table 4 we list the predicted number of galaxies that would be detected in the considered filters in the UDS and DS in different redshift intervals. At $z < 2$, we would expect a few times 10^4 objects per redshift bin to be detected at both wavelengths in the UDS, while this number increases to a few times 10^5 in the DS. These are predominantly normal star-forming galaxies, i.e. labelled as spiral, irregular or SF-AGN in the figures, with an increment of the starburst fraction with increasing redshift. Such observations would have allowed for outstanding improvements of our knowledge of the LF, not only at its knee, thanks to the large number of observed galaxies, but also at the faint end, considering the possible observations of galaxies below the LIRG limit.

Up to $z=2$, we would also expect to observe a sample of elliptical galaxies, whose analysis in the IR has been limited up to now mainly to nearby objects (e.g. Smith et al., 2012). Their numbers would vary between 10^2 and 10^3 objects in total, depending on the instrument considered and the area covered.

At increasing redshift, the number of observed galaxies would decrease, but we would expect to detect galaxies up to extremely high- z at $34 \mu\text{m}$ and at least up to $z=7-8$ at $70 \mu\text{m}$, thanks to the combination of depth and area coverage. At $34 \mu\text{m}$, we would expect $\sim 100-900$ objects at $z > 6$ detected in the UDS and slightly less, i.e. $\sim 300-500$, in the shallower DS. The uncertainties are due to the high- z extrapolation. The $70 \mu\text{m}$ filter would have been less sensitive than the $34 \mu\text{m}$ filter, resulting on fewer observed objects. In particular, we predict at maximum a few tens of detections in the UDS and around one hundred in the DS.

Keeping in mind the limitations of the simulation at

$z > 6$, we would expect these high- z galaxies to be SBs or composite galaxies with an AGN component (i.e., SB-AGN and SF-AGN), with some AGN-dominated system only in the DS. In the UDS (DS) these galaxies would have an average stellar mass of $\log_{10}(M^*/M_{\odot}) = 10.7$ (11.1), with values ranging from 8.5 (9.1) to 11.8 (12.2). The average SFR, derived combining UV and IR estimations, would be $\log_{10}(\text{SFR}_{\text{UV+IR}}/M_{\odot} \text{ yr}^{-1}) = 2.1$ in the UDS and 2.3 in the DS, but it could even reach up to $\log_{10}(\text{SFR}_{\text{UV+IR}}/M_{\odot} \text{ yr}^{-1}) = 3.8$. However, these SFRs need to be considered as lower limits, as comparing the simulation with observations at $z > 6$, we derived that a population of highly star-forming systems (i.e. $\log_{10}(\text{sSFR}/\text{yr}) \gtrsim -7.5$) may be missing (see B21). The gas-phase metallicity, derived considering the mass-metallicity relation by Wuyts et al. (2014), would be on average $12 + \log_{10}(\text{O}/\text{H}) = 8.3-8.4$, depending on the considered survey, but it could go as down as 7.5. All these values are derived considering the high- z extrapolation with $k_{\Phi} = -1$.

The improvement with respect to previous surveys is outstanding, given that *Herschel*, for example, has been mainly limited to galaxies below $z \leq 3.5$ (e.g. Gruppioni et al., 2013). This shows the need for similar surveys when investigating the evolution of both star-formation and black hole accretion in galaxies over a large redshift range.

In Figure 4 we show the results for OST/OSS. On one hand, with the OST/OSS Ch1 filter, we expect to be limited to $z < 8$ in both the OST-Deep and OST-Wide surveys, with the possibility to observe few tens (300) of galaxies $z > 8$ ($z > 6$) in the OST-Deep survey if we consider the flattest redshift evolution of the LF (i.e. $k_{\Phi} = -1$). On the other hand, with the OST/OSS Ch2 filter we expect to observe $0.8-2.1 \times 10^3$ galaxies with $z > 6$ in the OST-Wide survey of 20 deg^2 , depending on the considered extrapolation of the LF at high redshift.

In the OST-Deep (OST-Wide) survey such high- z ($z > 6$) galaxies are expected to have, on average, $\log_{10}(M^*/M_{\odot}) = 10.5$ (11.0) and $\log_{10}(\text{SFR}_{\text{UV+IR}}/M_{\odot} \text{ yr}^{-1}) = 2.0$ (2.4), considering the $k_{\Phi} = -1$ high- z extrapolation. These galaxies are, on average, less massive and more star-forming than the ones that would be detected by SPICA, because OST/OSS Ch2 observations are planned to be deeper than B-BOP observations. We remind the reader that for a fair comparison with SPICA capabilities we have considered only the first two OST/OSS channels. However, because of the positive k -correction (i.e. rest-frame wavelengths moving closer to the peak of dust emission at increasing redshift), the number of $z > 6$ galaxies detected by OST is expected to increase when considering all six channels.

Figure 5 shows the results for GEP. In the survey of 3 deg^2 , we instead expect to observe $1-7 \times 10^3$ galaxies at $z > 6$ with the narrow filter centred at $73 \mu\text{m}$ (Band

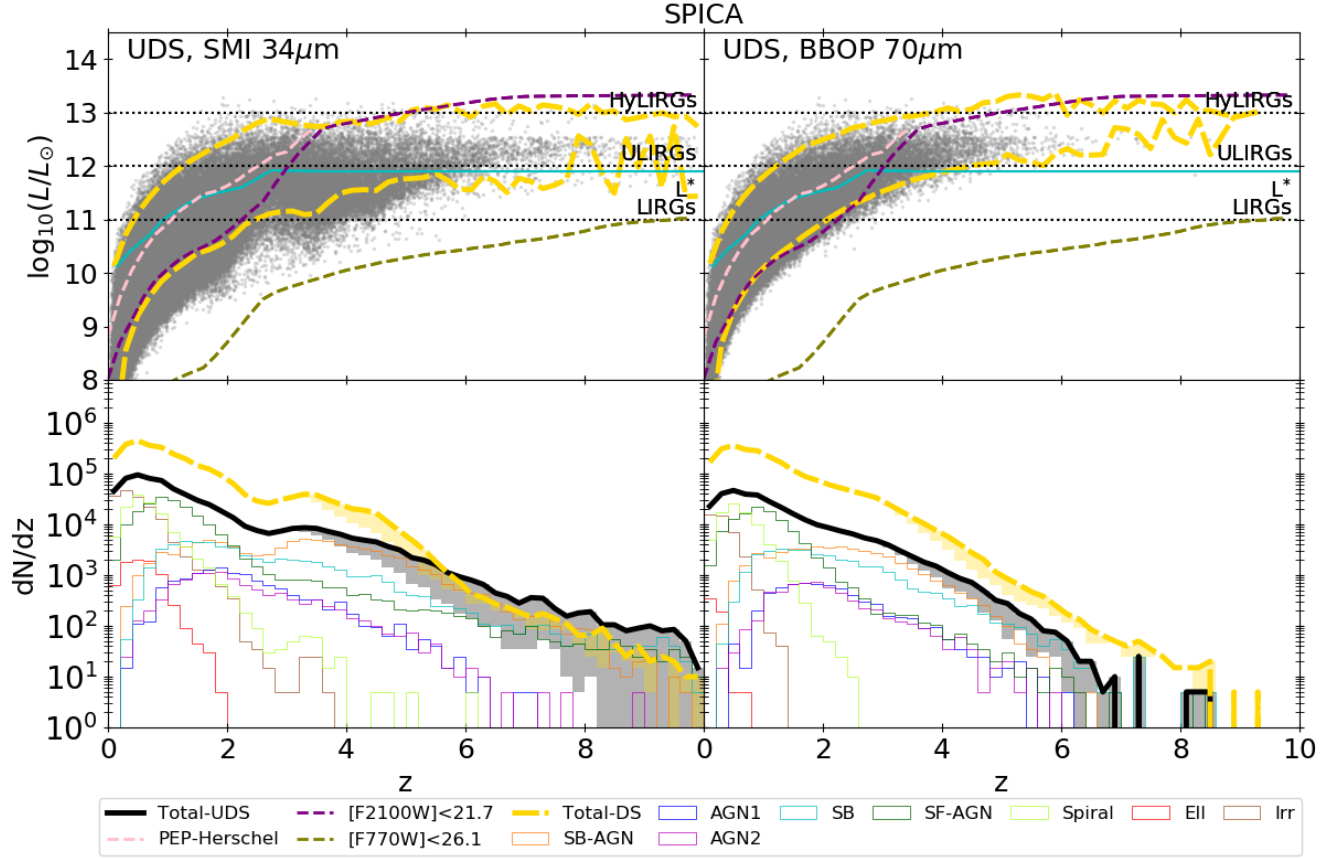


Figure 3. *Top:* L_{IR} vs. redshift for simulated galaxies (grey points) detected at $34\ \mu\text{m}$ (left) and $70\ \mu\text{m}$ (right) in the SPICA UDS. The dashed tick yellow lines refer to the L_{IR} area occupied by galaxies detected in SPICA DS. In each panel, the cyan solid line indicates the luminosity of the knee of the LF, as derived by Gruppioni et al. (2013), while the pink dotted line show the 80% completeness of the *Herschel*-PEP survey at $100\ \mu\text{m}$ in the GOODS-S (Berta et al., 2013). We also report the 80% completeness of the two samples with JWST/MIRI $[F2100]<21.7$ and $[F770W]<26.7$. The horizontal dotted black lines indicate the IR luminosity limit of HyLIRGs, ULIRGs and LIRGs. Results are obtained considering the high- z extrapolation with $k_{\Phi} = -1$. Both surveys would have allowed to observe not only ULIRGs, as done with *Herschel*, but also less luminous galaxies. *Bottom:* Redshift distribution per unit redshift interval of sources detected at $34\ \mu\text{m}$ with SMI (left) and at $70\ \mu\text{m}$ with B-BOP, as expected in the UDS (tick black lines) and in the DS (tick yellow dashed lines). For the UDS, we report also the distributions of the different sub-populations (coloured lines), all derived considering $k_{\Phi} = -1$. Results obtained with other $z>3$ extrapolations, i.e. from $k_{\Phi} = -4$ to -2 , are included in the grey and yellow areas. Overall, observations at $34\ \mu\text{m}$ could reach galaxies up to $z\sim 8$.

16) and 200-3500 with the filter centred at $33\ \mu\text{m}$ (Band 10). In the shallower survey of $30\ \text{deg}^2$ we instead expect ~ 500 -800 galaxies detected in Band 10 and between 2700 and 1.3×10^4 in Band 16.

In the GEP-Deep (GEP-Wide) survey such galaxies are expected to have, on average, $\log_{10}(M^*/M_{\odot}) = 10.5$ (10.9) and $\log_{10}(\text{SFR}_{UV+IR}/M_{\odot}\ \text{yr}^{-1}) = 1.9$ (2.2), considering the $k_{\Phi} = -1$ high- z extrapolation. In the GEP-Deep survey we also expect some AGN-dominated galaxies (i.e. AGN1 and AGN2) too rare for the OST-Deep survey or the UDS.

Differences between SPICA, OST and GEP are driven

by the different observational depths and different survey sizes, but also, in particular in the case of GEP, by the different filter widths.

3.1.3 X-ray fluxes

If we focus on the AGN activity, we can investigate what are the expected soft (0.5-2 keV, Figure 6) and hard (2-10 keV, Figure 7) X-ray fluxes for each galaxy detected by a SPICA-like mission, OST or GEP. This is important because, for galaxies with both IR and X-ray observations, it will be possible to analyse the global energetics of the AGN. In addition, IR observations are

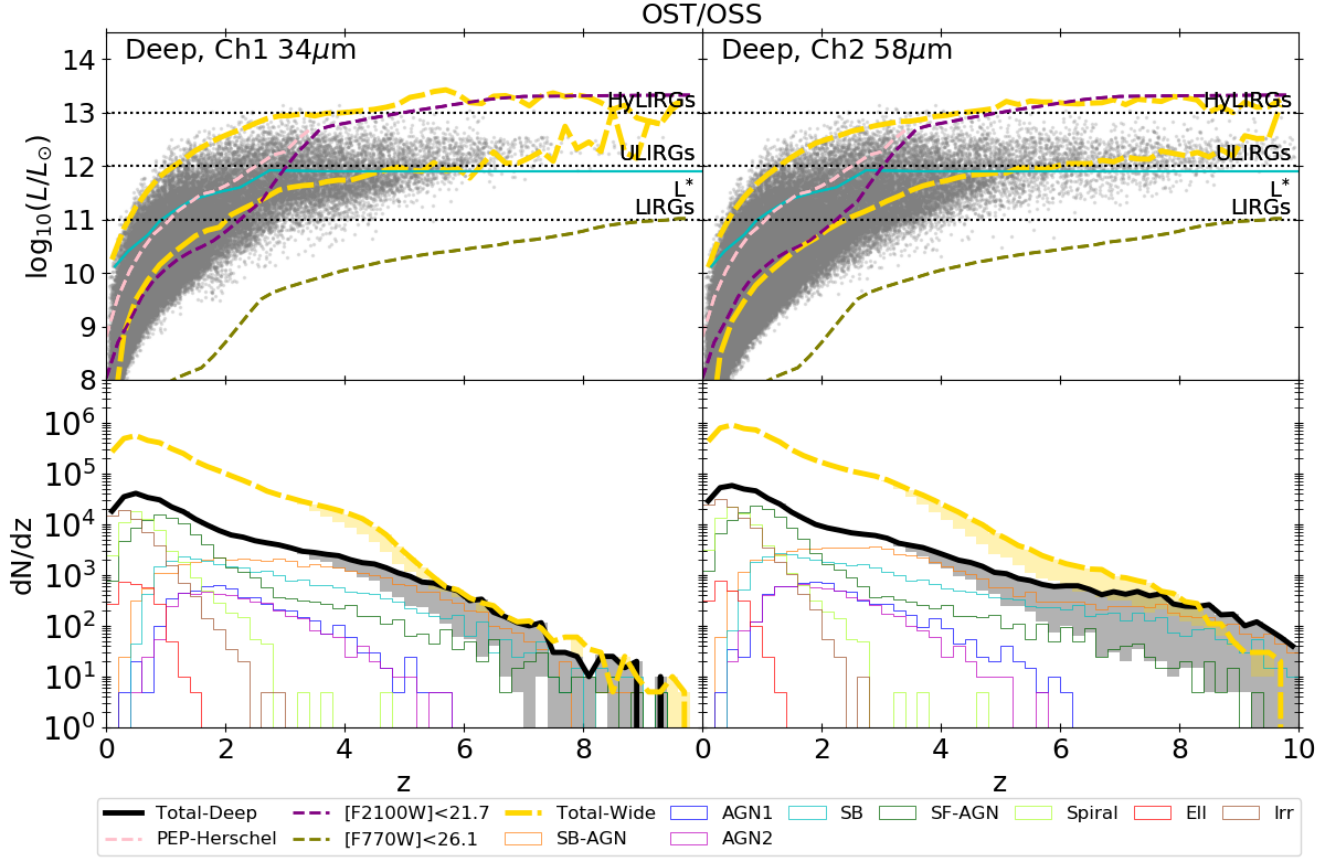


Figure 4. Same as Figure 3, but for galaxies detected with OST in channel 1 and 2 in the OST-Deep survey. Tick yellow dashed lines show the results for the OST-Wide survey. Both surveys will allow the detection of galaxies up to $z=7$, thus extending the existing observations well below the ULIRG typical luminosity, at least up to $z=5$.

Survey Filter	UDS		DS	
	SMI	B-BOP	SMI	B-BOP
$0 \leq z < 1$	7.5×10^4	3.7×10^4	3.4×10^5	2.8×10^5
$1 \leq z < 2$	3.2×10^4	1.8×10^4	1.6×10^5	1.3×10^5
$2 \leq z < 3$	8.8×10^3	6.5×10^3	3.7×10^4	4.6×10^4
$3 \leq z < 4$	7.7×10^3 (5.8×10^3)	2.5×10^3 (1.9×10^3)	3.3×10^4 (2.5×10^4)	1.5×10^4 (1.2×10^4)
$4 \leq z < 5$	4.3×10^3 (1.9×10^3)	700 (340)	1.5×10^4 (6.6×10^3)	3.1×10^3 (1.6×10^3)
$5 \leq z < 6$	1.6×10^3 (400)	150 (40)	2.4×10^3 (950)	520 (270)
$6 \leq z < 8$	760 (110)	30 (2)	420 (290)	130 (90)
$8 \leq z < 10$	180 (2)	3 (0)	70 (30)	12 (10)

Table 4 Approximated number of galaxies in different redshift intervals detected by SMI and B-BOP, as expected for the UDS and DS. Numbers in brackets are for the most conservative extrapolation at $z > 3$ considered in this work (i.e. $k_{\Phi} = -4$).

complementary to X-ray ones (Barchiesi et al., 2021), as they allow observations of the most dust-obscured AGN that are difficult to detect in the X-rays. As mentioned in Sec. 2.1, the observed faint end of the X-ray LF is overestimated by the predictions obtained with SPRITZ.

Therefore, the fractions of IR-detected AGN above certain X-ray fluxes have to be considered as upper limits, at least for the deepest X-ray threshold considered.

We compared the expected soft (hard) X-ray fluxes of IR detected AGN with the depth of cur-

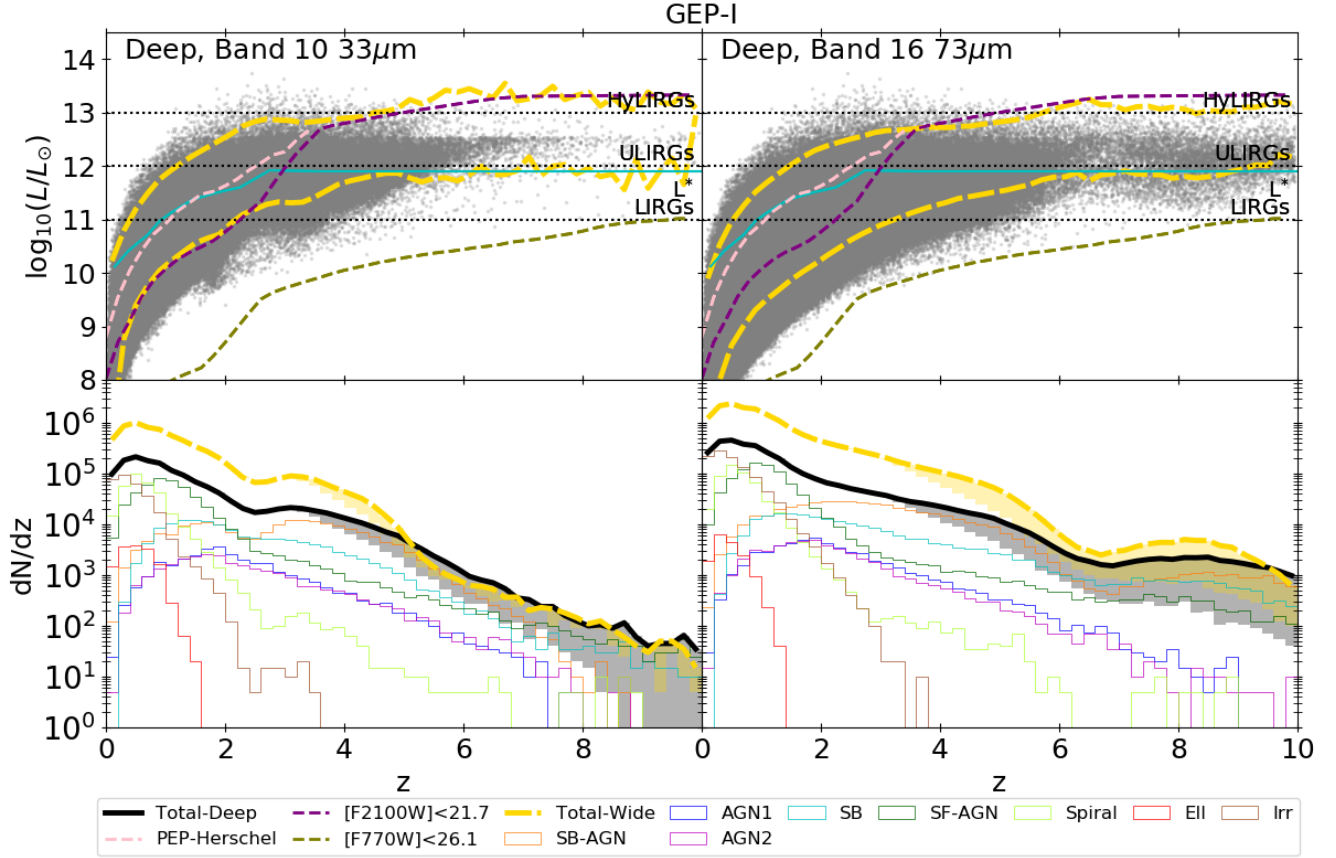


Figure 5. Same as Figure 3, but for galaxies detected with GEP-I in channel 10 and 16 in the GEP-Deep survey. *Tick yellow dashed lines* show the results for the GEP-Wide survey. At both wavelengths, these surveys will allow the detection of galaxies up to $z=9$ and, at $73 \mu\text{m}$, of some galaxies below the ULIRG luminosity at $z>5$.

rent and future X-ray surveys: $7 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ ($3.3 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$) like the XMM-Newton Deep survey in COSMOS (Hasinger et al., 2007), $2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ ($1.5 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$) like the Chandra Deep survey in COSMOS-Legacy (Civano et al., 2016) and $4 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ ($2.0 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$) like the planned *Athena* (Nandra et al., 2013) Deep survey.

In the SPICA UDS, we expect, on average, that 26%, 5% and 1% of the detected AGN would have soft X-ray fluxes above 4×10^{-17} , 2×10^{-16} and $7 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$, respectively. This fraction generally would have a peak around $z \sim 1.3$ and then it would decrease with increasing redshift, going from $\sim 65\%$ to less than 30% of galaxies having $f_{0.5-2\text{keV}} > 4 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ at $z > 2.5$.

In the DS, the fractions of detected AGN with soft X-ray fluxes above the considered limits would be larger than in the UDS, i.e. 37%, 22% and 9%, on average, of

detected AGN with soft X-ray fluxes above 4×10^{-17} , 2×10^{-16} and $7 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$, respectively. The fraction of galaxies above each X-ray flux cut would be large (~ 15 , 25 and 70%) even at low redshift, reflecting the reduced number of faint AGN in the DS survey. At $z > 4-5$, the fraction decreases with increasing redshift reaching fractions below 30%, considering the lowest flux cut $f_{0.5-2\text{keV}} > 4 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$.

The fraction of IR AGN detected in the X-rays would be larger in the DS than in the UDS because the DS covers a larger area than the UDS and, therefore, it would contain on average brighter AGN than the UDS. This is visible in Figure 8, where we show the overall distribution of the bolometric AGN luminosity obtained in the two surveys.

A second peak is present in the redshift evolution of some fractions mentioned above, particularly in the DS. This second peak is due to the $9.7 \mu\text{m}$ absorption feature that biases IR photometric observations to detect the

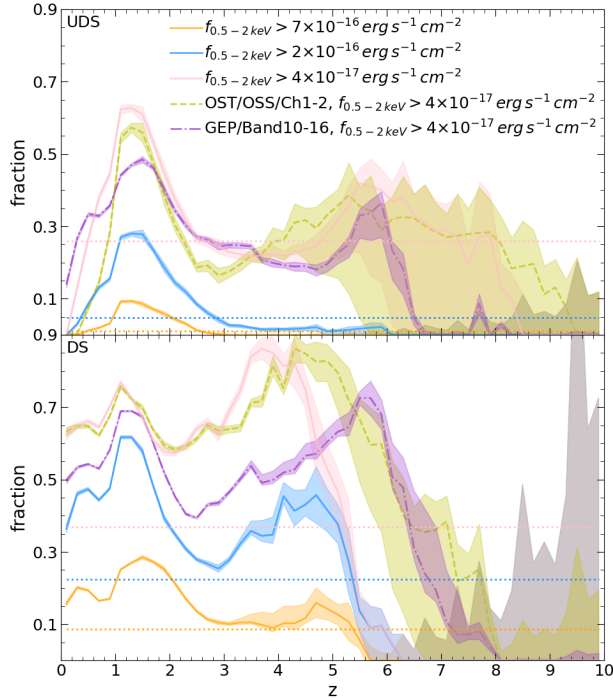


Figure 6. Fraction of simulated AGN detected in the IR with X-ray flux at 0.5-2 keV above 4×10^{-17} (pink lines), 2×10^{-16} (blue lines) and $7 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ (orange lines), in the SPICA UDS (top) and DS (bottom). We also report the fraction of simulated AGN with X-ray flux at 0.5-2 keV above $4 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ detected with OST/OSS (green dashed line) in the OST-Deep (top) and OST-Wide survey (bottom). The same fraction is shown for AGN detected with GEP (purple dot-dashed line) in the survey of 3 (top) and 30 deg² (bottom). Shaded areas show the uncertainties, including Poisson errors and the extrapolation at $z > 3$, while horizontal dotted lines indicate the average fraction for SPICA across the entire redshift range. IR observations are key to putting together a complete picture of AGN activity, as X-ray observations may miss a large fraction of them because of dust absorption.

brightest objects. It is however necessary to consider that the predicted depth of this feature is subject of various uncertainties. First, different torus models predict different $9.7 \mu\text{m}$ absorption strength, due to differences in the dust distribution and the silicates composition (e.g. Feltre et al., 2012; Hatziminaoglou et al., 2015). Second, while some low-metallicity galaxies with a clear $9.7 \mu\text{m}$ absorption feature have been observed (e.g. Thuan et al., 1999; Houck et al., 2004), some other studies (e.g. Kulkarni et al., 2011) have pointed out a possible reduction of the $9.7 \mu\text{m}$ absorption strength with decreasing metallicity. If this is the case, we would expect high- z galaxies to have a less pronounced $9.7 \mu\text{m}$ absorption feature compared to low- z galaxies, used as template in our simulation. Therefore, SPICA observations around $z=4$ would not be biased towards bright IR galaxies, thus reducing the fraction of IR AGN above each X-ray

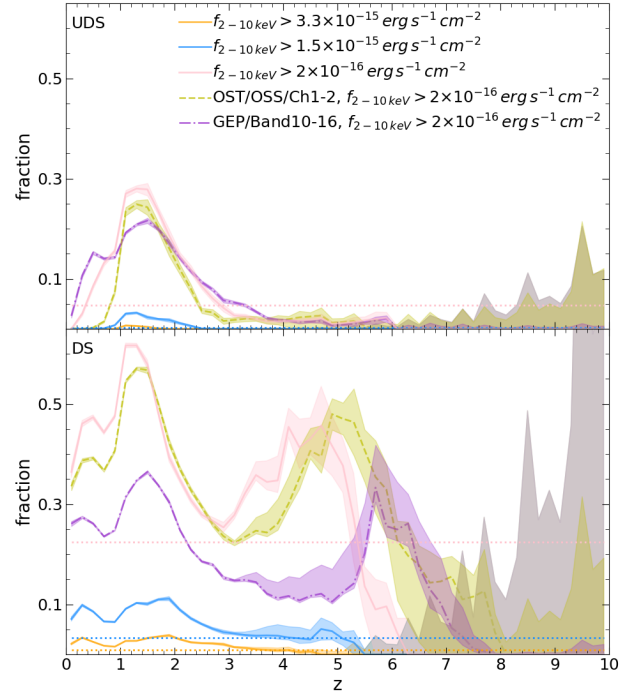


Figure 7. Same as Figure 6, but for AGN with X-ray flux at 2-10 keV above 3.3×10^{-15} (pink lines), 2.5×10^{-15} (blue lines) and $2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ (orange lines).

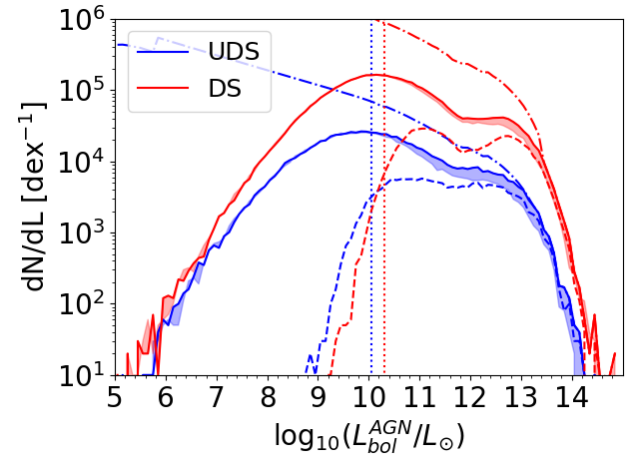


Figure 8. Distribution of the bolometric luminosity of the AGN detected in the SPICA UDS (blue solid line) and DS (red solid line). The dashed histograms indicate the distribution of the AGN at $z > 3$. Shaded areas show the uncertainties due to the $z > 3$ extrapolation. Vertical dotted lines are the median bolometric luminosities of the two surveys across the entire redshift range. The dashed-dotted lines indicate the intrinsic distribution of the bolometric luminosity, before applying any flux selection but normalised to the observed area.

threshold, i.e. smoothing out the secondary peaks visible in Fig. 6.

The fractions of IR-detected AGN above the consid-

ered hard X-ray fluxes would have a redshift evolution similar to the fractions of AGN detected in soft X-ray of the same X-ray surveys, although these fractions would be generally lower. These low fractions are due to the hard X-ray depths, which are shallower than the soft X-ray flux depths. In addition, as redshift increases, observed fluxes corresponds to higher rest-frame energies, where the emission falls like a power law with an exponential cutoff. This effect is larger at 2-10 keV than for 0.5-2 keV.

We would expect on average less and 1% of AGN having hard X-ray fluxes above $f_{2-10\text{keV}} > 1$ and $3 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$ and only 5% above the hard X-ray flux of $2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$, with a maximum of 30% around $z \sim 1.3$. In the DS, the average fraction would increase to 1, 3 and 22% above 3.3×10^{-15} , 1.5×10^{-15} and $2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$. The fraction of IR-detected AGN above the deepest hard X-ray flux cut, i.e. $2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$, would show a double peak around $z \sim 1.3$ and $z \sim 4.5$ with almost no AGN above the hard X-ray flux cut at $z > 6$, similarly to the soft X-ray AGN fraction in the DS. The detection in both soft and hard X-ray bands will allow for analysing in more detail the spectral type of the AGN, directly fitting the X-ray spectra or through the analysis of the hardness ratio.

In Figures 6 and 7 we also report the expected fraction of AGN with $f_{0.5-2\text{keV}} > 4 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ and $f_{2-10\text{keV}} > 2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ detected with OST/OSS, in Channel 1 or 2, and with GEP-I, in band 10 or 16. For OST, the average fraction is 25% (3%) in the OST-Deep survey of 0.5 deg^2 and 45% (25%) in the OST-Wide Survey of 20 deg^2 in the soft (hard) X-rays. The similar area and depth of the SPICA UDS and OST-Deep survey are reflected in similar fractions in both soft and hard X-rays. The average fractions for GEP-I are slightly smaller, 18% (5%) and 38% (15%) for the 3 and 30 deg^2 survey in soft (hard) X-ray, respectively. Even in this case, the fraction of AGN detected in both IR and X-rays generally decreases with increasing redshift.

It is important to highlight that, as mentioned before, the fraction of IR-detected AGN with faint X-ray fluxes is not limited to low-power AGN, but contains also Compton-thick AGN (i.e. obscured AGN with hydrogen column density $\log_{10}(N_H/\text{cm}^{-2}) \geq 24$) that are challenging to detect at X-rays, but they become bright at IR-wavelengths. For a more detailed analysis on the possible synergies between SPICA or OST and future X-ray telescopes, such as *Athena*, we refer to the dedicated work of Barchiesi et al. (2021).

3.2 Detection of IR spectral features

SMI-LR spectroscopy could have been performed simultaneously to the photometric SMI observations for both surveys, as the exposure times were long enough to allow for a complete scan of the sky observed in one pointing.

Even taking into account that the spectroscopic limits were significantly shallower than the photometric ones, a fraction of galaxies observed in the photometric UDS and DS would have had low-resolution SMI spectra as well. We predicted the capability of SMI-LR to detect some of the main IR spectral features produced by star formation or AGN activity, considering the expected observational depth of the parallel spectroscopic surveys (see Table 2).

The detection of multiple lines enables robust redshift estimates and then, depending on the line, the study of different physical properties of the galaxies, e.g. the nature of the ionising sources, gas densities, gas metal abundances. For a detailed analysis of the spectral capability of the other SPICA spectrometer, i.e. SAFARI, we refer to the work by Spinoglio et al. (2021). The results presented in the next paragraphs are consistent with results from Spinoglio et al. (2021) for the detection of PAH and fine-structure lines, taking into account that, in the aforementioned paper, there is no attempt at separating the AGN and star-forming galaxy populations and the resulting number counts are, consequently, always larger than the ones presented here.

As mentioned previously, the current plan for the GEP spectrometer is to perform only pointed observations, therefore we decided to not include any comparison with SPICA SMI-LR. The comparison between SPICA-LR and OST/OSS is instead possible by considering the six OST/OSS channels with their full resolution, i.e. $R=300$, and not with a reduced resolution of $R=4$, as done for the comparison between the photometric surveys.

3.2.1 PAH features

We start our analysis from the PAH features, ranging from $3.3 \mu\text{m}$ to $17.0 \mu\text{m}$, derived in SPRITZ considering the relations by Gruppioni et al. (2016), Bonato et al. (2019) and Mordini et al. (2021). The latter includes different relations for low-metallicity galaxies (i.e. $12 + \log(\text{O}/\text{H}) < 8.2$), as some observations point out that these galaxies may have suppressed PAH features (Shipley et al., 2016). In Figure 9 we report the redshift range in which each PAH feature is inside the wavelength coverage of SPICA/SMI and OST/OSS, for a better comparison between the wavelength coverage of the two instruments.

In Figure 10 we report, for both the SPICA UDS and DS, the expected redshift distribution per unit redshift interval of galaxies with detections in each PAH feature, where a detection means a $S/N > 3$ as derived by integrating the feature over its entire width. In the bottom panels of the same figure we report similar histograms, where we separate galaxies by the number of detected PAHs. This analysis allows us to verify for which and how many galaxies the redshift estimation would be solid, as multiple strong lines would be detectable inside the spectra.

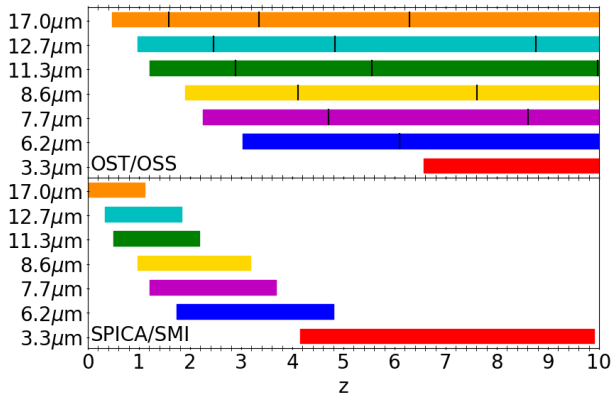


Figure 9. Redshift range in which each PAH feature is inside the wavelength coverage of OST/OSS (*top*) and SPICA/SMI (*bottom*). Vertical black lines show the separation between the different OST/OSS channels. The wide wavelength coverage of OST allows for the simultaneous detection of multiple PAH lines.

In the SPICA UDS, we would expect more than $\sim 17\%$ of galaxies at $z=1-4.5$ to have at least two detected PAH features, with a maximum of $\sim 60\%$ at $z\sim 2.5$. These would be mainly star-forming galaxies without an AGN component ($\sim 40\%$) or composite systems ($\sim 54\%$). In particular, PAH detections would be extremely rare among dwarf irregular galaxies, i.e. less than 2% of all dwarf irregular galaxies detected in photometry at $z=1-4.5$. The limited number of multiple PAH detections is largely limited by the wavelength coverage of the SPICA/SMI instrument, as seen in Figure 9.

In the SPICA DS, the fraction varies from 7 to 40% in the same redshift range, with a maximum at $z=2.6$. At $z>5$, only the PAH at $3.3\ \mu\text{m}$ would be covered by the SMI spectral range, though it would be too faint to be detected for a large fraction (i.e. $>99\%$) of the sample. Therefore, for these galaxies the redshift estimation would have to rely on other bright IR emission lines (see Sec. 3.2.2-3.2.3) or on the additional information provided by the simultaneous photometric surveys.

PAH features are not only important for deriving precise redshifts, but also for identifying potential AGN activity. It has indeed been shown that the equivalent widths of the 6.2 and $11.3\ \mu\text{m}$ PAH features anticorrelate with the dominance of AGN activity (Spoon et al., 2007; Tommasin et al., 2008; Hernán-Caballero & Hatziminaoglou, 2011; Feltre et al., 2013). This technique to identify potential AGN would be feasible between redshift 0.5 and 5 for 10^2-10^3 galaxies per redshift bin (dN/dz) in the UDS and DS.

The OST spectroscopic wavelength coverage is larger than the SPICA/SMI one, allowing the simultaneous detection of a significant number of PAH features and over a larger redshift range. This is visible in Figure 11, where we show the redshift distribution per unit redshift interval of galaxies with different PAH features detected

in at least one of the six OST/OSS channels.

In the OST-Deep survey, it will be possible to detect six PAH lines, using the spectroscopic capability of all OST/OSS channels, at $z=3$ to 8 for 4% of galaxies on average. The $3.3\ \mu\text{m}$ PAH feature is expected to be too faint to be detected with OST/OSS, but the application of some wavelength binning can boost the signal, reducing the wavelength resolution, but allowing for the detection of this feature similarly to what is predicted for SPICA/SMI.

On average, around 37% of the galaxies above $z=1$ are expected to have at least two PAH features detected. This fraction changes to 11% in the OST-Wide survey, where we expect to be able to detect at maximum four PAH features at the same time. This happens, on average, for $\sim 7\%$ of the galaxies at $z>3$. The difference is mainly due to the difficulties, given the planned OST-Wide survey observational depths, on detecting the 12.7 and $11.3\ \mu\text{m}$ features.

The fraction of galaxies with a detected $11.3\ \mu\text{m}$ PAH feature is highly uncertain in the OST-Wide survey, depending on the relation considered to derive the flux of this feature, and its detection may be possible at $z>1.5$. Even the detection of the $6.2\ \mu\text{m}$ feature is very uncertain because, when considering the uncertainties in our prediction, it may not be possible to detect this feature at all above $z>3.5$. As mentioned for the $3.3\ \mu\text{m}$ feature, wavelength binning could help on boosting the signal and increase the detection rates for this PAH feature.

3.2.2 IR hydrogen recombination lines

We now focus on the detection of the main hydrogen nebular emission lines in the IR. In Figures 12 and 13 we plot the redshift distribution per unit redshift interval of galaxies with expected detection ($S/N>3$) of different hydrogen recombination lines from the Paschen, Brackett, Pfund and Humphreys series. We do not consider hydrogen lines that are outside the SMI or OST wavelength coverage in the considered redshift range. Hydrogen lines include the contribution from both star formation and AGN activity. As the broad-line contribution has not been included in SPRITZ, the AGN nebular emission, which accounts only for the emission from the gas in the narrow-line regions, has to be considered as a lower limit to the total (broad+narrow) nebular emission.

All the aforementioned hydrogen lines would be too faint to be detected with SPICA/SMI for the large majority of galaxies. In the UDS, only a few percent of galaxies would have detectable Pa_α , Br_α , Br_β or Pf_α , while all the other lines would be detected for less than 1% of the total amount of observed galaxies at any redshift. In the DS, only Pa_α would be detected for 4% of the galaxies, while the other hydrogen lines would be detected in less than 1% of the observed galaxies at any redshifts.

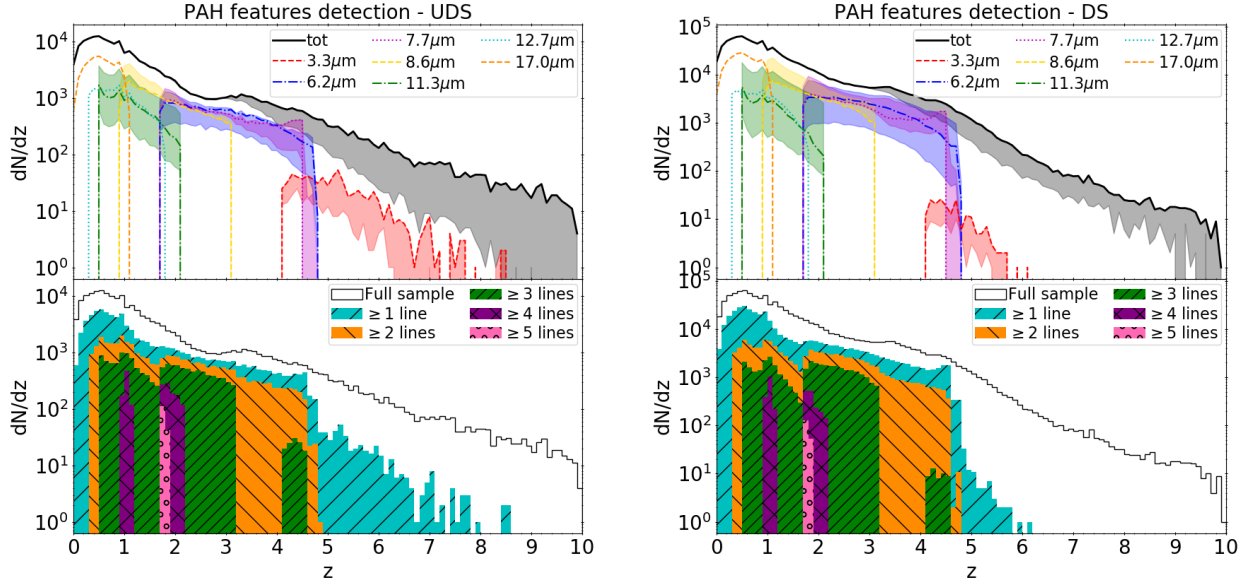


Figure 10. *Top:* Redshift distribution per unit redshift interval of galaxies in the SPICA UDS (*left*) and DS (*right*) with detected PAH lines (*coloured lines*), i.e. integrated $S/N > 3$. We also report the complete sample of galaxies detected with SMI or B-BOP in photometry (*black solid line*). The shaded areas show the uncertainties due to the different empirical relation $L_{IR}-L_{line}$, when present, and the extrapolation at $z > 3$. *Bottom:* Redshift distribution per unit redshift interval of galaxies in the UDS (*left*) and DS (*right*). The sample is divided by the number of detected PAH features. The black solid line is the complete photometric sample. At $z < 5$, multiple PAH lines would be observed by SPICA in both surveys, for a sizeable amount of galaxies.

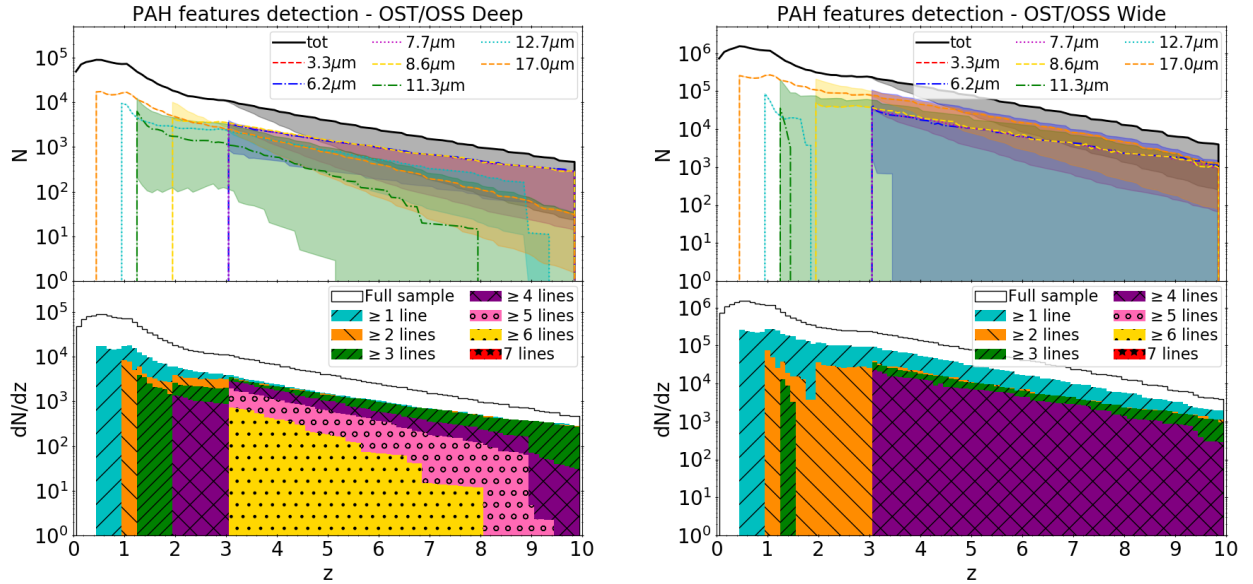


Figure 11. Same as Figure 10, but for OST. The complete sample (*black solid line*) consists on galaxies detected in at least one OST/OSS channel considering $R=4$. PAH detections are instead derived considering $R=300$. At $z > 1$, multiple PAH lines are expected to be observed by OST in both surveys, for a sizeable amount of galaxies.

In the range between $z=4.9$ and $z=8.0$, where only the $3.3 \mu\text{m}$ PAH feature would be detectable (see Sec. 3.2.1), the few detections of the Br_α and Br_β lines could be used for the redshift estimation, for which only the

$3.3 \mu\text{m}$ PAH feature is otherwise observable. At even higher redshifts ($z > 8$), in the UDS only a few tens of galaxies would have a detection of the $\text{Pa}_{\alpha, \beta}$, making the redshift estimation difficult for the rest of the sample.

This task would be even more difficult in the DS, where the Pa_α line would never be detected.

The OST-Deep survey is expected to have a similar depth to the SPICA/SMI UDS survey, but it covers a quarter of the area with a larger wavelength range. The shortest available wavelength is however $25 \mu\text{m}$, compared to $17 \mu\text{m}$ of the SPICA/SMI, moving Pa_α outside the wavelength coverage at all analysed redshifts. These differences result in the expected detection of few lines, mainly Br_α , Pf_α and Hu_α , for at maximum 2% of the galaxies detected in photometry (i.e. OST/OSS with $R=4$), but around 1% on average.

On the OST-Wide survey, the situation is slightly better, given that brighter objects are observed in the OST-Wide survey than in the OST-Deep one. The maximum is reached at $z \sim 5$, where 4% of all galaxies observed in photometry have detected Br_α . In addition, in the OST-Wide survey few galaxies have also detected Hu_β , Hu_γ and Pf_β , which are expected never to be observed in the Deep survey.

Overall, both for both a SPICA-like mission and OST the detection of hydrogen lines is expected to be very challenging and limited to the brightest galaxies. This needs to be taken into account when planning to derive the gas metallicity using the hydrogen lines, although metallicity determinations based only on the IR fine-structure lines have proven to be robust (e.g. Figure 7 in Fernández-Ontiveros et al., 2021).

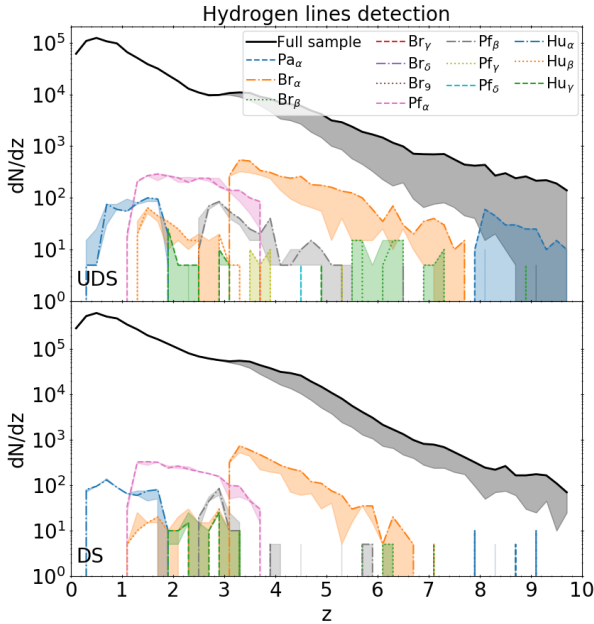


Figure 12. Redshift distribution per unit redshift interval of galaxies with detected ($S/N > 3$) hydrogen lines in the SPICA UDS (top) and DS (bottom). The list of considered lines is reported in the legend. The shaded areas show the uncertainties due to the extrapolation at $z > 3$.

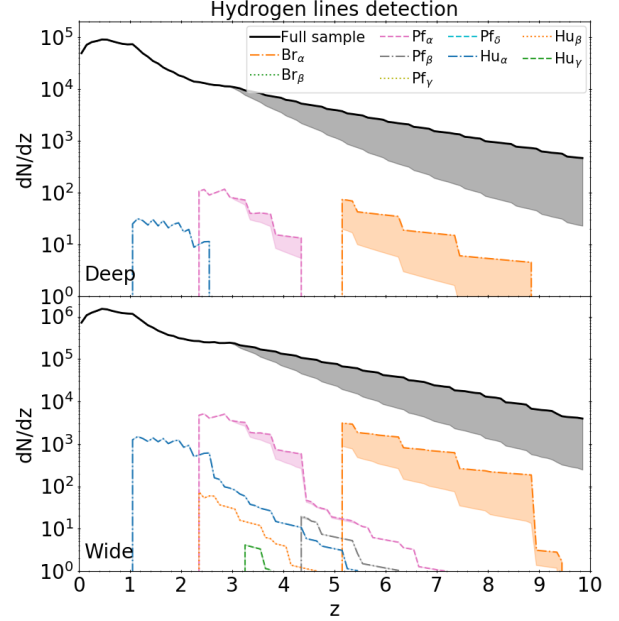


Figure 13. Same as Figure 12, but for galaxies detected in at least one OST/OSS channel in the OST-Deep (top) and OST-Wide survey (bottom). Only few galaxies are expected to have a detection in an Hydrogen line.

3.2.3 IR fine-structure lines

In this last section we report the predictions for the IR fine-structure lines as included in SPRITZ, due to star formation or AGN activity. These fine-structure lines can be used to analyse different physical properties. For example, the use of two lines of the same ionic species, like the two $[\text{S III}]$ lines at 18 and $33 \mu\text{m}$, can be used to investigate the gas density. The use of two lines of the same element, but of two different ionisation levels, like $[\text{Ne III}]$ at $15.55 \mu\text{m}$ and $[\text{Ne II}]$ at $12.81 \mu\text{m}$ (Thornley et al., 2000), can instead be used to investigate the nature of the ionising source. In addition, the detection of specific fine-structure lines that need high ionising energies, like the $[\text{Ne V}]$ at 14.32 and $24.31 \mu\text{m}$, is an indication of AGN activity (e.g. Spinoglio & Malkan, 1992; Genzel & Cesarsky, 2000; Meléndez et al., 2008; Feltre et al., 2016). More details on the use of mid-IR lines to study the nature of ionising sources, the gas densities and the gas metal abundances can be found in Spinoglio et al. (2015) and Fernández-Ontiveros et al. (2016, 2017).

The redshift at which we would expect each fine-structure line to be observed is shown in Figures 14 and 15, for the SPICA UDS and DS, respectively. The detection of these lines with SMI-LR would be limited at $z < 4.5$, because of the SMI wavelength coverage. At higher redshift, it would be necessary to search for emission lines in the rest-frame near-IR, precisely in the range $1.5 \mu\text{m} < \lambda < 6.5 \mu\text{m}$ that are not currently included in SPRITZ.

There would be no significant difference in the relative numbers of detections between the UDS and DS. The precise percentage of detections depends on the considered lines and redshift, ranging from less than 1% to almost all the galaxies. This gives an idea of the great power of such IR spectroscopy, offering a plethora of nebular emission lines not only to derive a precise redshift estimation, but also to study both star formation and AGN activity.

In Figures 16 and 17, we report results for the OST/OSS, considering all six available channels. We consider also additional fine-structure lines at wavelengths longer than what would be observed by SPICA/SMI, i.e. $\lambda > 37 \mu\text{m}$, as OST/OSS covers up to $589 \mu\text{m}$. The number of galaxies with fine-structure-line detections varies with redshift, line and survey.

In the OST-Deep survey this number is generally below 10^4 galaxies per redshift bin, while in the OST-Wide surveys some lines, like the [O I] at $63 \mu\text{m}$, are expected to be detected in 5×10^5 galaxies per redshift bin, at least at very low- z (i.e. $z < 0.8$). The redshift range in which the different lines can be detected is very large, compared to what would have been possible with SPICA, thanks to the OST/OSS large wavelength range. Therefore, we expect that the analysis of many fine-structure lines, such as the [Ne V] at 14.32 and $24.3 \mu\text{m}$ and [O IV] at $25.9 \mu\text{m}$, tracers of AGN activity, can be extended even at $z > 4.5$ with future OST/OSS observations.

4 CONCLUSIONS

In this paper we presented a set of predictions for the spectro-photometric surveys planned for the SPICA mission concept, derived with the state-of-the-art simulation SPRITZ. Results are valid also for similar 2.5 m actively cooled space missions. We also present detailed predictions for NASA's OST and GEP mission concepts, obtained with the same simulation.

In particular, we simulated the SPICA-like DS and UDS considering galaxies with detection in at least one of the broad-band filters centred at 34 (SMI) and $70 \mu\text{m}$ (B-BOP) and observations with the simultaneous SMI-LR spectrometer ranging from 17 to $36 \mu\text{m}$. We considered a 5σ depth of 3 and $13 \mu\text{Jy}$ at $34 \mu\text{m}$ and 60 and $100 \mu\text{Jy}$ at $70 \mu\text{m}$, in the UDS and DS respectively. We also simulated two surveys for OST, a Deep one of 0.5 deg^2 and a Wide one of 20 deg^2 , and other two surveys of 3 and 30 deg^2 for GEP.

Our main findings are:

- The most luminous galaxies would be observed up to $z=10$, in the SPICA DS of 15 deg^2 , and, even considering the most conservative results, at least up to $z=8$ in the SPICA UDS. Similar conclusions apply also to OST and GEP.

- LIRGs would be observed at $34 \mu\text{m}$ up to $z=4.0$ in the UDS and up to $z=2.5$ in the DS, whereas ULIRGs would be detected at all redshifts in the both the UDS and DS. Observations of LIRGs would be mainly limited at $z < 2$ at $70 \mu\text{m}$ in both surveys. The observational depth planned for the OST/OSS Chanel 2 in the OST-Deep survey should instead allow for detecting at least few LIRGS up to $z=4$.
- On average, 26% (37%) of AGN detected at 34 or $70 \mu\text{m}$ in the SPICA UDS (DS) would have soft $0.5\text{-}2 \text{ keV}$ X-ray fluxes above $4 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$, i.e. corresponding to the observational depth of the *Athena* Deep Survey. The fractions are lower if we consider detections in the hard X-rays. This shows the complementary nature of X-ray and IR observations in the study of AGN in general, and obscured AGN in particular.
- The redshift estimation could be derived for SPICA/SMI using at least two PAH features from $z=1.0$ to $z=5$, for 29% (19%) of the sample in the UDS (DS) on average. At $z \sim 5\text{-}8$, the redshift estimation could be based on the Br_α and Br_β lines at these redshifts, at least for the few galaxies for which these lines are detected. At higher redshifts, only few tens of galaxies could have a detection in an hydrogen recombination line.
- Considering all OST/OSS channels and galaxies detected in photometry (i.e. $R=4$) in the OST-Deep survey, we expect to detect six different PAH features for 4% of the galaxies at $z=3\text{-}8$ and at least two PAH features for 37%, on average, of galaxies at $z > 1$. However, OST wavelength range, starting at wavelengths longer than SPICA/SMI, limits the number of observable bright hydrogen recombination lines.
- The AGN and host galaxy analysis could be performed with SPICA/SMI at $z < 4.5$ and it will be possible at even higher redshifts with OST/OSS. This is possible using a large set of IR nebular emission lines originated from star formation or AGN activity, such as different neon lines like the [Ne V] at $14.32 \mu\text{m}$ and at $24.31 \mu\text{m}$ and the [Ne II] at $12.81 \mu\text{m}$.

These results show the capability of SPICA, OST and GEP to detect galaxies up to $z=10$ without being affected by dust obscuration and also to study the physics of AGN and star formation for a large fraction of such objects.

5 ACKNOWLEDGEMENTS

This paper is dedicated to the memory of Bruce Swinyard, who initiated the SPICA project in Europe, but unfortunately died on 22 May 2015 at the age of 52. He was ISO-LWS calibration scientist, *Herschel*-SPIRE instrument sci-

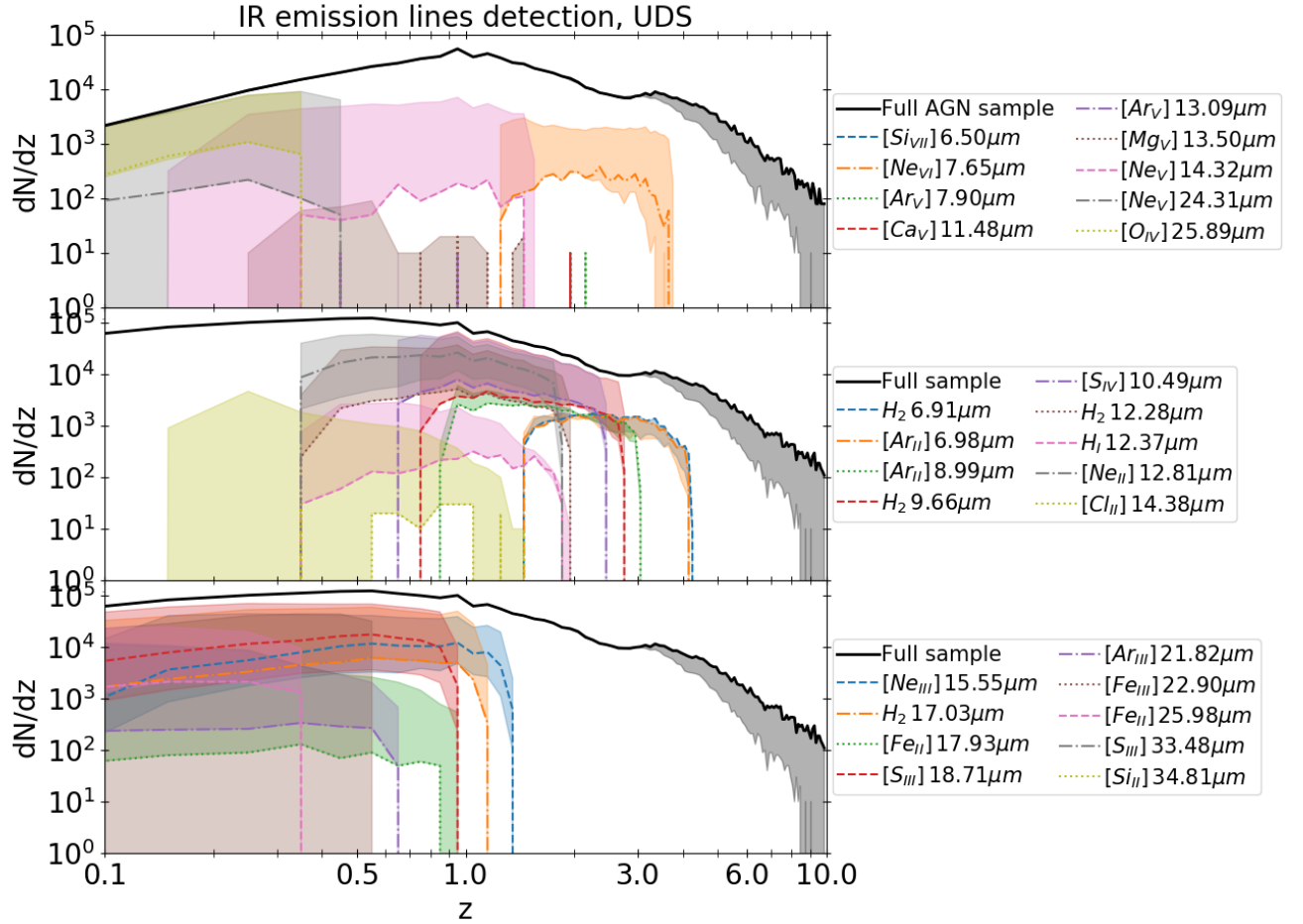


Figure 14. Redshift distribution per unit redshift interval of galaxies in the SPICA UDS with IR fine-structure lines with $S/N > 3$. The list of lines is reported on the right of each panel. The first panel show AGN lines while the other two panels show lines originated both by star formation and AGN activity. The complete sample of galaxies detected with SMI or B-BOP in photometry is shown in black, in the top panel we limit the sample to AGN only. We consider as reference the line luminosity derived from the L_{IR} considering the relation by [Bonato et al. \(2019\)](#). The shaded areas show the uncertainties due to different empirical relations $L_{IR}-L_{line}$ (i.e. [Grupponi et al., 2016](#); [Mordini et al., 2021](#)), when present, and the extrapolation at $z > 3$.

entist, first European PI of SPICA and first design lead of SAFARI. We acknowledge the whole SPICA Collaboration Team. LB, CG, LS and JAFO acknowledge financial support by the Agenzia Spaziale Italiana (ASI) under the research contract 2018-31-HH.0. LB would like to thank F. La Franca and M. Malkan for providing constructive comments. AF and FC acknowledges the support from grant PRIN MIUR2017-20173ML3WW_001. MPS acknowledges support from the Comunidad de Madrid through the Atracción de Talento Investigador Grant 2018-T1/TIC-11035 and PID2019-105423GA-I00 (MCIU/AEI/FEDER,UE). FJC acknowledges financial support from the Spanish Ministry MCIU under project RTI2018-096686-B-C21 (MCIU/AEI/FEDER/UE), cofunded by FEDER funds and from the Agencia Estatal de Investigación, Unidad de Excelencia María de Maeztu, ref. MDM-2017-0765.

REFERENCES

- Aird J., et al., 2010, *MNRAS*, **401**, 2531
- Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, *MNRAS*, **451**, 1892
- Arnouts S., et al., 2007, *A&A*, **476**, 137
- Barchiesi L., et al., 2021, *PASA*, **38**, e033
- Beare R., Brown M. J. I., Pimbblet K., Taylor E. N., 2019, *ApJ*, **873**, 78
- Berta S., et al., 2013, *A&A*, **551**, A100
- Béthermin M., et al., 2015, *A&A*, **573**, A113
- Bianchi S., et al., 2018, *A&A*, **620**, A112
- Bisigello L., et al., 2021, *A&A*, **651**, A52
- Bonato M., et al., 2019, *PASA*, **36**, e017
- Bouchet P., et al., 2015, *Publications of the Astronomical Society of the Pacific*, **127**, 612

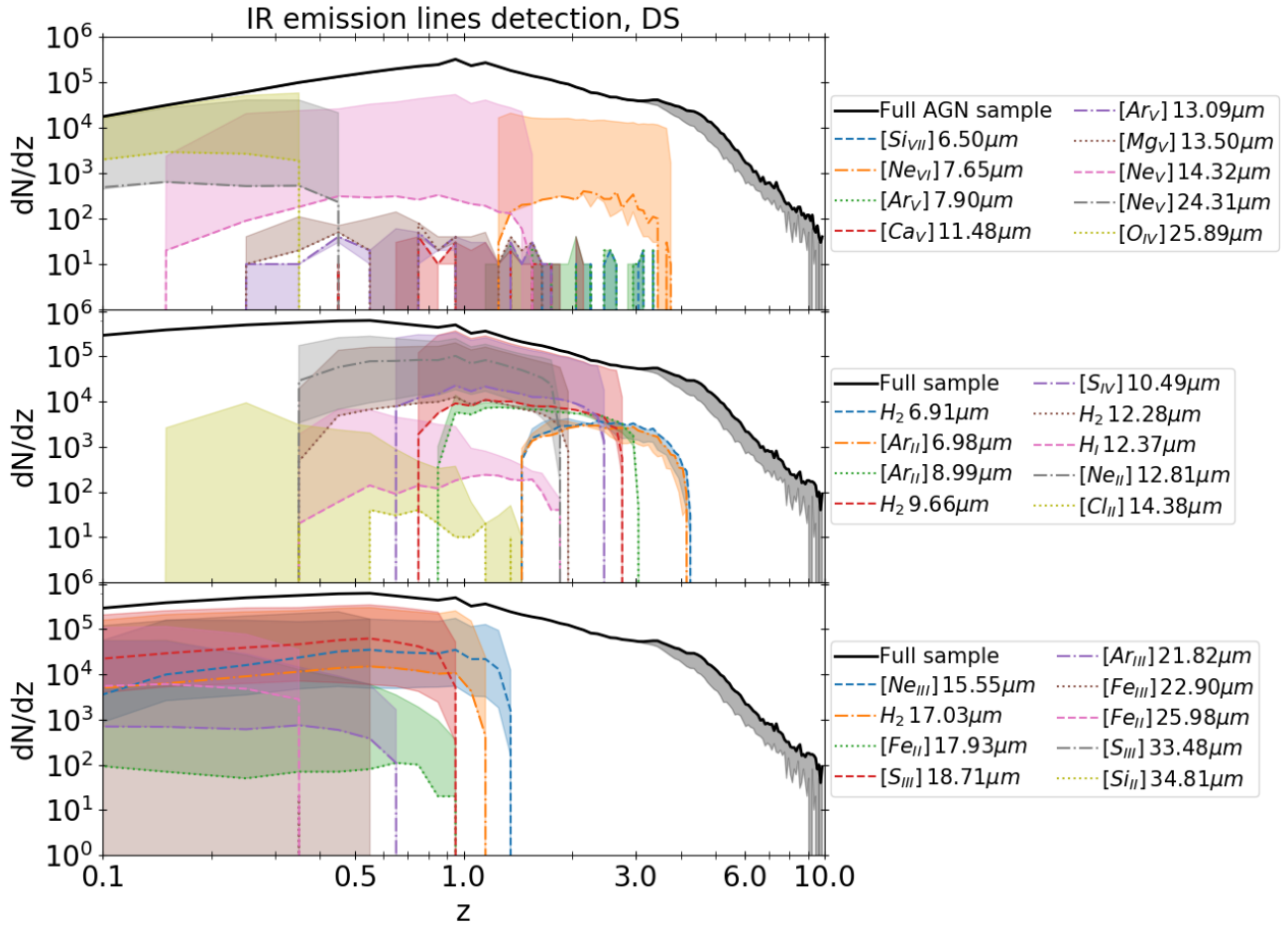


Figure 15. Same as of Figure 14, but for the SPICA DS.

Bouwens R. J., et al., 2015, *ApJ*, **803**, 34

Bouwens R. J., et al., 2016, *ApJ*, **833**, 72

Bouwens R., et al., 2020, *ApJ*, **902**, 112

Bradford C. M., et al., 2018, in Lystrup M., MacEwen H. A., Fazio G. G., Batalha N., Siegler N., Tong E. C., eds, Vol. 10698, *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*. SPIE, pp 385 – 401, doi:10.1117/12.2314049

Cirasuolo M., et al., 2007, *MNRAS*, **380**, 585

Civano F., et al., 2016, *ApJ*, **819**, 62

Dunlop J. S., et al., 2017, *MNRAS*, **466**, 861

Feltre A., Hatziminaoglou E., Fritz J., Franceschini A., 2012, *MNRAS*, **426**, 120

Feltre A., et al., 2013, *MNRAS*, **434**, 2426

Feltre A., Charlot S., Gutkin J., 2016, *MNRAS*, **456**, 3354

Fernández-Ontiveros J. A., Spinoglio L., Pereira-Santaella M., Malkan M. A., Andreani P., Dasyra K. M., 2016, *ApJS*, **226**, 19

Fernández-Ontiveros J. A., et al., 2017, *PASA*, **34**, e053

Fernández-Ontiveros J. A., Pérez-Montero E., Vílchez J. M., Amorín R., Spinoglio L., 2021, *A&A*, **652**, A23

Franco M., et al., 2018, *A&A*, **620**, A152

Gardner J. P., et al., 2009, in Thronson H. A., Stiavelli M., Tielens A., eds, *Astrophysics in the Next Decade*. Springer Netherlands, Dordrecht, pp 1–29

Genzel R., Cesarsky C. J., 2000, *ARA&A*, **38**, 761

Glasse A., et al., 2015, *PASP*, **127**, 686

Glenn J., et al., 2018, in Lystrup M., MacEwen H. A., Fazio G. G., Batalha N., Siegler N., Tong E. C., eds, Vol. 10698, *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*. SPIE, pp 205 – 220, doi:10.1117/12.2314076

Glenn J., et al., 2019, in *Bulletin of the American Astronomical Society*. p. 112

Gruppioni C., et al., 2010, *A&A*, **518**, L27

Gruppioni C., et al., 2013, *MNRAS*, **432**, 23

Gruppioni C., et al., 2016, *MNRAS*, **458**, 4297

Gruppioni C., et al., 2017, *PASA*, **34**, e055

Gruppioni C., et al., 2020, *A&A*, **643**, A8

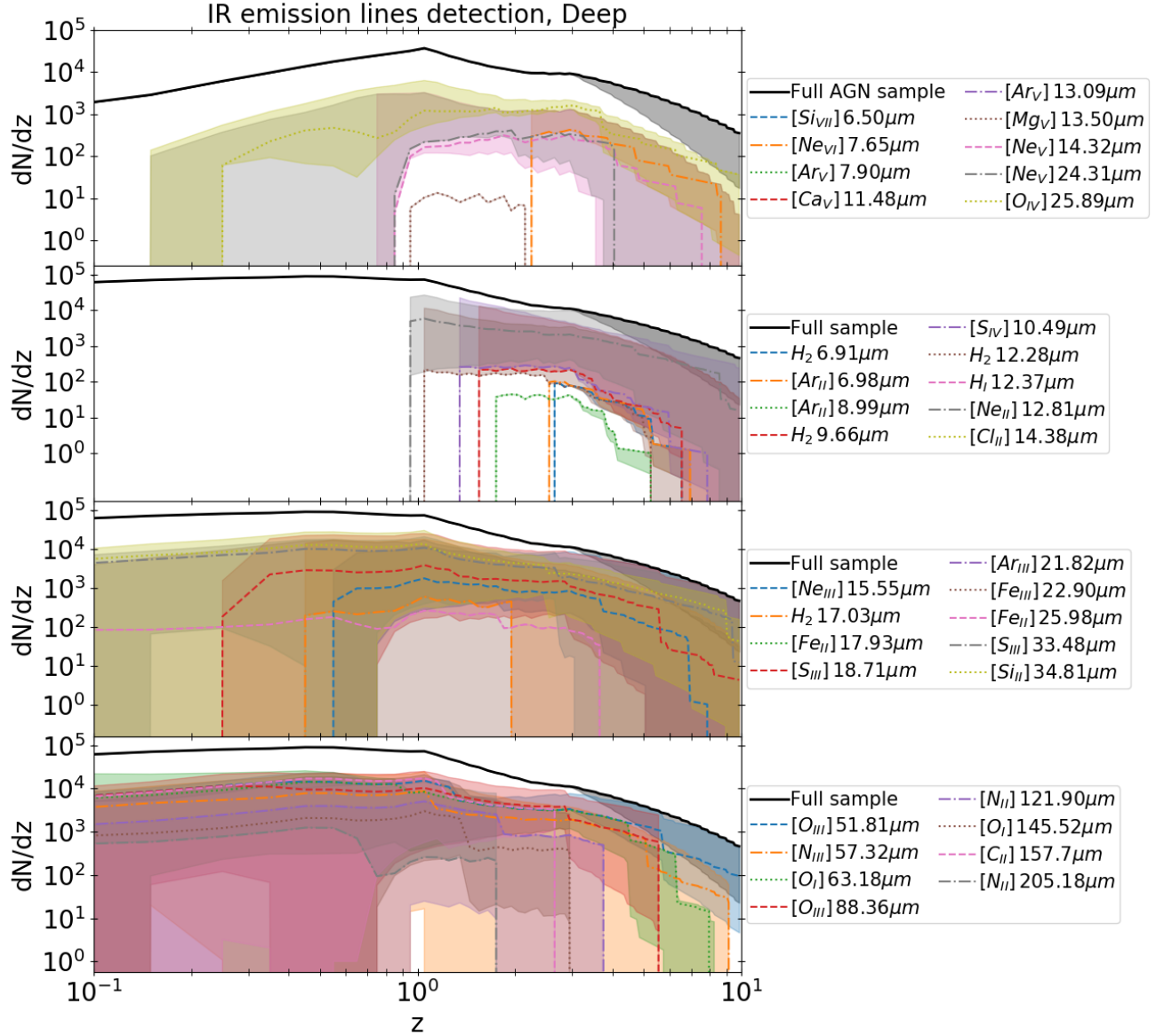


Figure 16. Same as of Figure 14, but for the OST-Deep survey. The complete sample refers to galaxies detected in at least one OST/OSS channel considering $R=4$, while line detections are instead derived considering $R=300$.

Hasinger G., et al., 2007, *ApJS*, **172**, 29

Hatfield P. W., Lindsay S. N., Jarvis M. J., Häußler B., Vaccari M., Verma A., 2016, *MNRAS*, **459**, 2618

Hatziminaoglou E., Hernán-Caballero A., Feltre A., Piñol Ferrer N., 2015, *ApJ*, **803**, 110

Hernán-Caballero A., Hatziminaoglou E., 2011, *MNRAS*, **414**, 500

Houck J. R., et al., 2004, *ApJS*, **154**, 211

Huertas-Company M., et al., 2016, *MNRAS*, **462**, 4495

Kaneda H., et al., 2017, *PASA*, **34**, e059

Kendrew S., et al., 2015, *Publications of the Astronomical Society of the Pacific*, **127**, 623

Kulkarni V. P., Torres-Garcia L. M., Som D., York D. G., Welty D. E., Vladilo G., 2011, *ApJ*, **726**, 14

Le Fèvre O., et al., 2020, *A&A*, **643**, A1

Leisawitz D., et al., 2019, in Barto A. A., Breckinridge J. B., Stahl H. P., eds, Vol. 11115, *UV/Optical/IR Space Telescopes and Instruments: Innovative Technologies and Concepts IX*. SPIE, pp 184 – 195, doi:10.1117/12.2530514

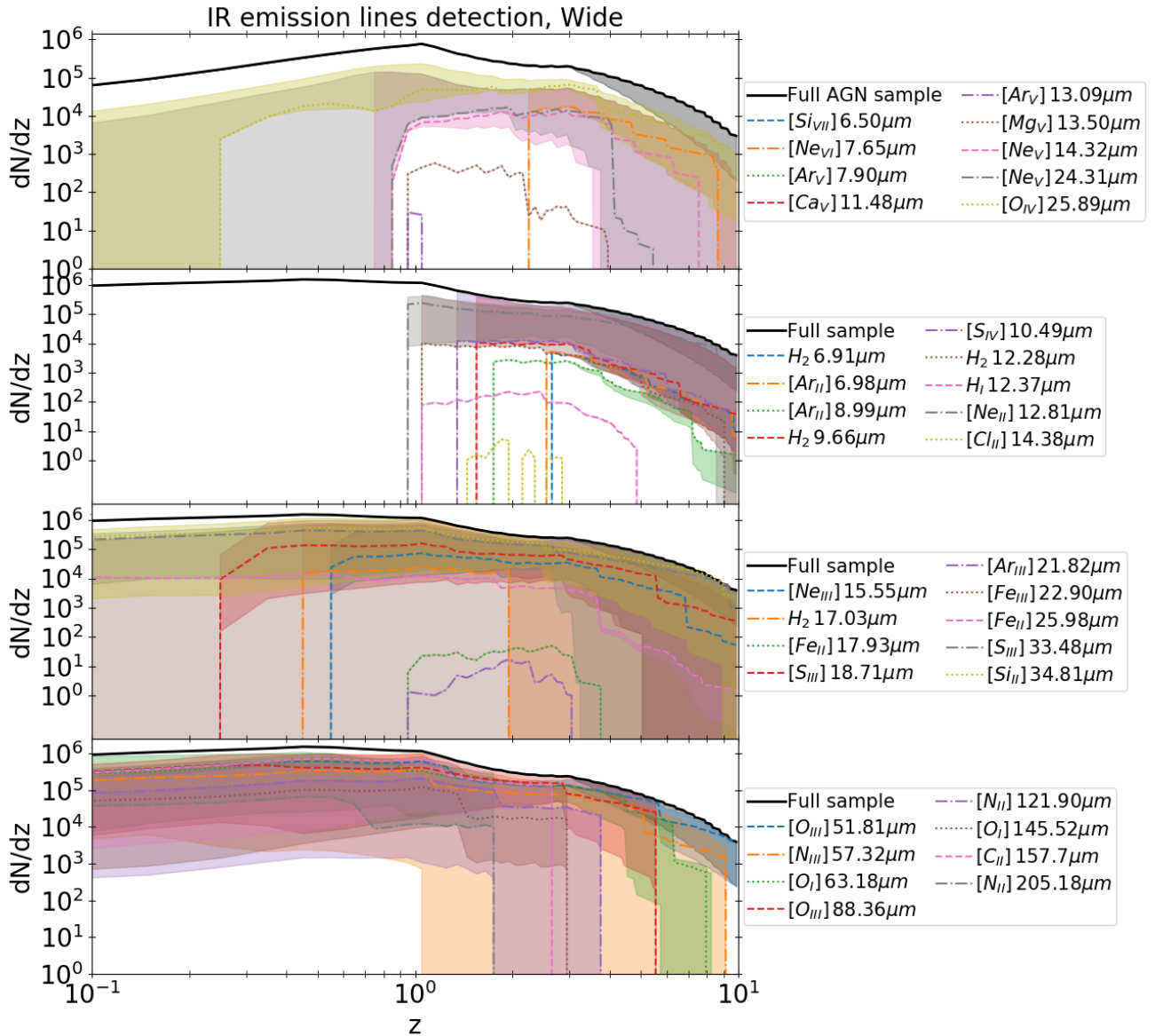


Figure 17. Same as Figure 14, but for the OST-Wide survey. The complete sample refers to galaxies detected in at least one OST/OSS channel considering $R=4$, while line detections are instead derived considering $R=300$.

Limber D. N., 1953, *ApJ*, **117**, 134

Livermore R. C., Finkelstein S. L., Lotz J. M., 2017, *ApJ*, **835**, 113

Livermore R. C., Trenti M., Bradley L. D., Bernard S. R., Holwerda B. W., Mason C. A., Treu T., 2018, *ApJ*, **861**, L17

Lonsdale C. J., et al., 2003, *PASP*, **115**, 897

Madau P., Dickinson M., 2014, *ARA&A*, **52**, 415

Magnelli B., et al., 2013, *A&A*, **553**, A132

Meléndez M., et al., 2008, *ApJ*, **682**, 94

Miyaji T., et al., 2015, *ApJ*, **804**, 104

Mordini S., Spinoglio L., Fernández-Ontiveros J. A., 2021, *A&A*, **653**, A36

Nandra K., et al., 2013, arXiv e-prints, p. arXiv:1306.2307

Novak M., et al., 2017, *A&A*, **602**, A5

Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, *ApJ*, **855**, 105

Polletta M., et al., 2007, *ApJ*, **663**, 81

Raymond G., Isaak K. G., Clements D., Rykala A.,

- Pearson C., 2010, *PASJ*, **62**, 697
- Reddy N., et al., 2012, *ApJ*, **744**, 154
- Reddy N. A., et al., 2015, *ApJ*, **806**, 259
- Rieke G. H., Alonso-Herrero A., Weiner B. J., Pérez-González P. G., Blaylock M., Donley J. L., Marcillac D., 2009, *ApJ*, **692**, 556
- Rieke G. H., et al., 2015, *Publications of the Astronomical Society of the Pacific*, **127**, 584
- Roelfsema P. R., et al., 2018, *PASA*, **35**, e030
- Sanders D. B., et al., 2007, *ApJS*, **172**, 86
- Saunders W., Rowan-Robinson M., Lawrence A., Efsthathiou G., Kaiser N., Ellis R. S., Frenk C. S., 1990, *MNRAS*, **242**, 318
- Schenker M. A., et al., 2013, *ApJ*, **768**, 196
- Schreiber C., et al., 2015, *A&A*, **575**, A74
- Shiple H. V., Papovich C., Rieke G. H., Brown M. J. I., Moustakas J., 2016, *ApJ*, **818**, 60
- Smith M. W. L., et al., 2012, *ApJ*, **748**, 123
- Sommovigo L., Ferrara A., Pallottini A., Carniani S., Gallerani S., Decataldo D., 2020, *MNRAS*, **497**, 956
- Soneira R. M., Peebles P. J. E., 1978, *AJ*, **83**, 845
- Spinoglio L., Malkan M. A., 1992, *ApJ*, **399**, 504
- Spinoglio L., Pereira-Santaella M., Dasyra K. M., Calzoletti L., Malkan M. A., Tommasin S., Busquet G., 2015, *ApJ*, **799**, 21
- Spinoglio L., et al., 2021, *PASA*, **38**, e021
- Spoon H. W. W., Marshall J. A., Houck J. R., Elitzur M., Hao L., Armus L., Brandl B. R., Charmandaris V., 2007, *ApJ*, **654**, L49
- Talia M., Cimatti A., Giulietti M., Zamorani G., Bethermin M., Faisst A., Le Fèvre O., Smolčić V., 2021, *ApJ*, **909**, 23
- Tamura Y., et al., 2019, *ApJ*, **874**, 27
- Thornley M. D., Förster Schreiber N. M., Lutz D., Genzel R., Spoon H. W. W., Kunze D., Sternberg A., 2000, *ApJ*, **539**, 641
- Thuan T. X., Sauvage M., Madden S., 1999, *ApJ*, **516**, 783
- Tommasin S., Spinoglio L., Malkan M. A., Smith H., González-Alfonso E., Charmandaris V., 2008, *ApJ*, **676**, 836
- Vito F., et al., 2018, *MNRAS*, **473**, 2378
- Wake D. A., et al., 2011, *ApJ*, **728**, 46
- Wang Y., Brunner R. J., Dolence J. C., 2013, *MNRAS*, **432**, 1961
- Wang T., et al., 2019, *Nature*, **572**, 211
- Wang L., et al., 2021, *A&A*, **648**, A8
- Wolf J., et al., 2021, *A&A*, **647**, A5
- Wright G. S., et al., 2015, *Publications of the Astronomical Society of the Pacific*, **127**, 595
- Wuyts E., et al., 2014, *ApJ*, **789**, L40