Research Paper

Simulating infrared spectro-photometric surveys with a SPRITZ

L. Bisigello1,2, C. Grupioni1, F. Calura3, A. Feltre1, F. Pozzi1,2, C. Vignali1,2, L. Barchiesi1,2, M. Negrello5, F. J. Carrera6, K. M. Dasyra7,8, J. A. Fernández-Ontiveros9,10, M. Giard1, E. Hatziminaoglou12, H. Kaneda13, E. Lusso14,15, M. Pereira-Santaella16, P. G. Pérez González16, C. Ricci17,18, D. Schaerer19, L. Spinoglio9 and L. Wang20

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, 7Department 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–IAAPS), 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-INSA), France, 12European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany, 13Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan, 14Dipartimento di Fisica e Astronomia, Università di Firenze, Vicolo dell’Osservatorio 5, I-35122 Padova, Italy, 15INAF – 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 and 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. LIR > 10^{12} L_{⊙}) up to z = 10, but also normal galaxies (LIR < 10^{11} L_{⊙}) up to z ∼ 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 ∼ 4.

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

(Received 17 August 2021; revised 16 November 2021; accepted 22 November 2021)

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 ∼ 200–300) and medium-resolution (R ∼ 3 000–11 000) observing modes instantaneously covering the 35–210 μm wavelength range. The SPICA Mid-Infrared Instrument (SMI) has three operating modes: a large field of view (10′ × 12′) low-resolution 17–36μm imaging spectroscopic (R ∼ 50–120) mode and photometric camera at 34 μm (SMI-IR), a medium-resolution

Author for correspondence: Laura Bisigello, e-mail: laura.bisigello@inaf.it

Cite this article: Bisigello L, Grupioni C, Calura F, Feltre A, Pozzi F, Vignali C, Barchiesi L, Rodighiero G, Negrello M, Carrera FJ, Dasyra KM, Fernández-Ontiveros JA, Giard M, Hatziminaoglou E, Kaneda H, Lusso E, Pereira-Santaella M, Pérez González PG, Ricci C, Schaerer D, Spinoglio L and Wang L (2021) Simulating infrared spectro-photometric surveys with a SPRITZ. Publications of the Astronomical Society of Australia 38, e064, 1–18. https://doi.org/10.1017/pasa.2021.57

© The Author(s), 2021. Published by Cambridge University Press on behalf of the Astronomical Society of Australia.
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 ∼ 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, Finkelstein, & Lotz 2017; Livermore et al. 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 importance 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. ∼10^6 M_☉; 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 ∼ 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, for example 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 Telescope (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 supermassive 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.

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 5.9 and 2.0 m 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 ΛCDM cosmology with H₀ = 70 km s⁻¹Mpc⁻¹, Ω_m = 0.27 and Ω_Λ = 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 (Bisigello et al. 2021). All the mock catalogues considered in this paper are made publicly available.a

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, that is SF-AGN, is common at z = 1–2 and includes low-luminosity AGN, while the second one, that is SB-AGN, is common at z > 2, hosts a bright, but heavily obscured AGN and has a high SFR (i.e. log_{10}(SFR/yr) ∼ 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 Bisigello et al. (2021). At $z > 3$ the LFs are extrapolated assuming a constant characteristic luminosity and a decreasing number density at the knee as $\propto (1 + z)^{4.5}$, 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 Bisigello et al. (2021) 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 Bisigello et al. 2021).

- The K-band LF of elliptical galaxies (EII), 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 $\propto (1 + z)^{4.5}$, with $k_\Phi$ 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 Bisigello et al. 2021, 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 Bisigello et al. 2021, 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 the contribution of the broad-line regions is missing, the fluxes of the permitted lines for the simulated un-obscured 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, Brunner, & Dolence 2013), and $A_w$ evolving with the stellar mass (Wake et al. 2011; Hatfield et al. 2016) but not with redshift (Béthermin 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 Bisigello et al. (2021), 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 ($\sim 24-600 \mu m$) and with the recent IR LF observed by Gruppioni et al. (2020) up to $z = 6$, 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 built 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 $M_{1600A} > -23$) 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; Aird et al. 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 $\sim 15$ 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 Section 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 Bisigello et al. (2021) for more information about SPRITZ.

### 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 \(\mu m\) (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 \(\mu m\) which is part of the Cosmic Evolution Early Release Science (CEERS\(^4\)) survey and the MIRI 7.7 \(\mu m\) observations of the Medium JWST Advanced Deep Extragalactic Survey (JADES\(^5\)). JWST and the three far-IR missions analysed in this work are complementary in wavelengths, as JWST covers \(\lambda < 28.8 \mu m\) while SPICA, OST and GEP cover \(\lambda > 20, 25\) and \(10 \mu m\), respectively.

At the same time, JWST is particularly suited for performing very deep mid-IR observations, as in imaging it can reach a 10\(\sigma\) depth of 7.09 \(\mu Jy\) at 21.0 \(\mu m\) in \(10^4\) s (Glaeser 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 spectro-photometric surveys, following two planned SPICA surveys (Table 2):

- an Ultra-Deep Survey (UDS) covering 1 \(\text{deg}^2\) in 600 h, reaching down to 3 \(\mu Jy\) (5\(\sigma\)) with SMI at 34 \(\mu m\) and 60 \(\mu Jy\) (5\(\sigma\)) with B-BOP at 70 \(\mu m\). In the same amount of time SMI-LR would simultaneously obtain spectra down to 39 \(\mu Jy\) at 20 \(\mu m\) and 65 \(\mu Jy\) at 30 \(\mu m\) (5\(\sigma\)).
- a deep survey (DS) covering 15 \(\text{deg}^2\) in 600 h, reaching down to 13 \(\mu Jy\) (5\(\sigma\)) with SMI at 34 \(\mu m\) and 100 \(\mu Jy\) (5\(\sigma\)) with the B-BOP at 70 \(\mu m\). In the same amount of time SMI-LR would simultaneously obtain spectra down to 150 \(\mu Jy\) at 20 \(\mu m\) and 250 \(\mu Jy\) at 30 \(\mu m\) (5\(\sigma\)).

Both spectroscopic limits are derived considering a low background level (i.e. 19 \(\text{M} \text{Jy sr}^{-1}\) at 27 \(\mu m\)). Photometric confusion noise is not included in SPRITZ and corresponds at 5\(\sigma\) to 9 \(\mu Jy\) and 0.25 \(\mu Jy\) at 34 \(\mu m\) and 70 \(\mu m\), 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 \(\mu m\) (left panel) and B-BOP at 70 \(\mu m\) (central panel). The chosen area corresponds to a single SMI pointing, whereas a B-BOP pointing covers a smaller area, that is 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 \(\mu m\) and of 9’’ at 70 \(\mu m\).

2.2.2. OST

The Origins Survey Spectrometer (OSS; Bradford et al. 2018) planned for OST consists of six channels covering between 25 and

---

4https://ceers.github.io.
5https://www.cosmos.esa.int/web/jwst-nirspec-gto/jades.
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 $R = 300$.

For comparison with the two SPICA photometric surveys, we also considered the photometric points derived considering each OST channel as a photometer, that is at a resolution of $R = 4$ instead of $R = 300$. This corresponds to a 5σ depth of 4.5 and 30.3 μJy at 25–44 μm 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 μm and of OST/OSS Ch2 with the SPICA/B-BOP filter 70 μm. 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 nJy and 1.1 mJy at 50 and 250 μm. 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 | $\lambda$ (μm) | Deep (0.5 deg$^2$) | Wide (20 deg$^2$) |
|---------|----------------|--------------------|------------------|
| 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 10–400 μm, with spectral resolution $R = 8$ at 10–95 μm and $R = 3.5$ at 95–400 μm. Considering the central wavelengths of the different filters, in this work we compared results for the SPICA/SMI filter at 34 μm and SPICA/B-BOP at 70 μm with the GEP-I filters centred at ~33 (Band 10) and 73 μm (Band 16).

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

- An all sky survey with a 5σ depth of ~1 mJy.
- A survey covering 300 deg$^2$ with a 5σ depth of ~50 μJy.
- A survey covering 30 deg$^2$ with a 5σ depth of ~20 μJy (hereafter GEP-Wide).
- A survey covering 3 deg$^2$ with a 5σ depth of ~5 μJy (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 μm 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 μm 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 SPITZ. 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 Section 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 redshift $z = 6$. 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), that is, respectively $L_{IR} > 10^{12} L_\odot$ and $L_{IR} > 10^{13} L_\odot$. We indeed would expect to detect around $4\times6 \times 10^4$ ULIRGs up to redshift $z = 10$, showing a mean redshift in the range $z = 2.8$--3.5 and $\sim700$--900 HyLIRGs with a mean redshift $z = 2.6$--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 $\sim5\times6\times10^4$ HyLIRGs with mean redshift $z = 3.6$--3.8 and $3.5$--$4.7 \times 10^3$ ULIRGs with mean redshift $z = 3.0$--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 $k_b = -4$ instead of $k_b = -1$, the average redshift decreases as there are fewer galaxies at $z > 3$.

The predicted HyLIRGs number (i.e. $30$--$50$ 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 extrapolation with $k_b = -1$, the average redshift decreases.

The galaxy population observed by a SPICA-like mission would not be limited to extreme IR galaxies. Indeed, with SMI at 34 $\mu$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 $\mu$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.
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 $z = 5$.

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 $\mu$m, of some galaxies below the ULIRG luminosity at $z > 5$. 
The OST/OSS Ch1 depth in the OST-Deep survey is similar to the SMI depth at 34 $\mu$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 $\mu$m, that is 30.3 $\mu$Jy compared to 13 $\mu$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 $\mu$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 $\sim$5 and 20 $\mu$m, in the 3 and 30 deg$^2$ surveys, respectively. Such surveys are therefore deeper than both SPICA surveys at 70 $\mu$m and the SPICA DS at 34 $\mu$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 $\mu$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 $\mu$Jy) and [F770W] $< 26.7$ (i.e. 0.07 $\mu$Jy). The two selections correspond to the magnitude limits of the CEERS and JADES surveys, respectively, in their two longest wavelength filters (see Section 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 $\mu$m, to the OST-Wide limit at 34 $\mu$m and to the GEP-Wide limit at 33 $\mu$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. $\sim$4.6 and 8arcmin$^2$ 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 $\mu$m and B-BOP at 70 $\mu$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, that is 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 $\mu$m and at least up to $z = 7$–8 at 70 $\mu$m, thanks to the combination of depth and area coverage. At 34 $\mu$m, we would expect $\sim$100–900 objects at $z > 6$ detected in the UDS and slightly less, that is $\sim$300–500, in the shallower DS. The uncertainties are due to the high-$z$ extrapolation. The 70 $\mu$m filter would have been less sensitive than the 34 $\mu$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.

### 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. $k_B = -4$).

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

Downloaded from https://www.cambridge.org/core. University of Groningen, on 27 Dec 2021 at 10:58:02, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2021.57
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_☉) = 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}(SFR_{UV+IR}/M_☉ yr^{-1}) = 2.1 in the UDS and 2.3 in the DS, but it could even reach up to log_{10}(SFR_{UV+IR}/M_☉ 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}(SFR/yr) ≥ −7.5) may be missing (see Bisigello et al. 2021). The gas-phase metallicity, derived considering the mass-metallicity relation by Wuyts et al. (2014), would be on average 12 + log_{10}(O/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_β = −1.

The improvement with respect to previous surveys is outstanding, given that Herschel, for example, has been mainly limited to galaxies below z ≤ 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_β = −1). On the other hand, with the OST/OSS Ch2 filter we expect to observe 0.8–2.1 × 10^4 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_☉) = 10.5 (11.0) and log_{10}(SFR_{UV+IR}/M_☉ yr^{-1}) = 2.0 (2.4), considering the k_β = −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 × 10^3 galaxies at z > 6 with the narrow filter centred at 73 μm (Band 16) and 200–3500 with the filter centred at 33 μm (Band 10). In the shallower survey of 30 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_☉) = 10.5 (10.9) and log_{10}(SFR_{UV+IR}/M_☉ yr^{-1}) = 1.9 (2.2), considering the k_β = −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 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 Section 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 current and future X-ray surveys: 7 × 10^{-16} erg s^{-1} cm^{-2} (3.3 × 10^{-15} erg s^{-1} cm^{-2}) like the XMM-Newton Deep survey in COSMOS (Hasinger et al. 2007), 2 × 10^{-16} erg s^{-1} cm^{-2} (1.5 × 10^{-15} erg s^{-1} cm^{-2}) like the Chandra Deep survey in COSMOS-Legacy (Civano et al. 2016) and 4 × 10^{-17} erg s^{-1} cm^{-2} (2.0 × 10^{-16} erg s^{-1} 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 $4 \times 10^{-17}$, $2 \times 10^{-16}$ and $7 \times 10^{-16}$ erg s$^{-1}$ 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 ~65% to less than 30% of galaxies having $f_{0.5-2\text{keV}} > 4 \times 10^{-17}$ erg s$^{-1}$ 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, that is 37%, 22% and 9%, on average, of detected AGN with soft X-ray fluxes above $4 \times 10^{-17}$, $2 \times 10^{-16}$ and $7 \times 10^{-16}$ erg s$^{-1}$ 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 $f_{0.5-2\text{keV}} > 4 \times 10^{-17}$ erg s$^{-1}$ 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 μm absorption feature that biases IR photometric observations to detect the 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 μm 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 μm absorption feature have been observed (e.g. Thuan, Sauvage, & Madden 1999; Houck et al. 2004), some other studies (e.g. Kulkarni et al. 2011) have pointed out a possible reduction of the 9.7 μm absorption strength with decreasing metallicity. If this is the case, we would expect high-z galaxies to have a less pronounced 9.7 μm 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 threshold, that is smoothing out the secondary peaks visible in Figure 6.

The fractions of IR-detected AGN above the considered 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 than 1% of AGN having hard X-ray fluxes above $f_{2–10\text{keV}} > 1$ and $3 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ and only 5% above the hard X-ray flux of $2 \times 10^{-16}$ erg s$^{-1}$ cm$^{-2}$, with a maximum of 30% around $z \sim 1.3$. In the DS, the average fraction would increase to 13 and 22% above $3 \times 10^{-15}$, $1.5 \times 10^{-15}$ and $2 \times 10^{-16}$ erg s$^{-1}$ cm$^{-2}$. The fraction of IR-detected AGN above the deepest hard X-ray flux cut, that is $2 \times 10^{-16}$ erg s$^{-1}$ 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 trough the analysis of the hardness ratio.

In Figures 6 and 7, we also report the expected fraction of AGN with $f_{0.5-2\text{keV}} > 4 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$ and $f_{1-10\text{keV}} > 2 \times 10^{-16}$ erg s$^{-1}$ 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 deg² 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-1 are slightly smaller, 18% (5%) and 38% (15%) for the 3 and 30 deg² 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 $\log(N_{HI}/cm^{-2}) > 24$) 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, for example the nature of the ionising sources, gas densities, gas metal abundances. For a detailed analysis of the spectral capability of the other SPICA spectrometer, that is 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 SPCA-LR and OST/OSS is instead possible by considering the six OST/OSS channels with their full resolution, that is $R = 300$, 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 to 17.0 µm, 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. $12 + \log(O/H) < 8.2$), 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 SPCA/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.

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, that is 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 SPCA/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 µm would be covered by the SMI spectral range, thought 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 Sections 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 µm PAH features anti-correlate 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 SPCA/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.

---

![Figure 9](https://example.com/figure9.png) **Figure 9.** Redshift range in which each PAH feature is inside the wavelength coverage of OST/OSS (top) and SPC/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.
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), that is 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 of 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 OST-Deep survey, it will be possible to detect six PAH lines, using the spectroscopic capability of all OST/OSS channels, at $z = 3–8$ for 4% of galaxies on average. The 3.3 μm 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 μm features.
The fraction of galaxies with a detected 11.3 \( \mu \)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 \( \mu \)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 \( \mu \)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 Pas, Br\( \alpha \), Br\( \beta \) or Pf\( \alpha \), 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\( \alpha \) 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.

In the range between \( z = 4.9 \) and \( z = 8.0 \), where only the 3.3 \( \mu \)m PAH feature would be detectable (see Section 3.2.1), the few detections of the Br\( \alpha \) and Br\( \beta \) lines could be used for the redshift estimation, for which only the 3.3 \( \mu \)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 Pa\( \alpha \), making the redshift estimation difficult for the rest of the sample. This task would be even more difficult in the DS, where the Pa\( \alpha \) 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 \( \mu \)m, compared to 17 \( \mu \)m of the SPICA/SMI, moving Pa\( \alpha \) outside the wavelength coverage at all analysed redshifts. These differences result in the expected detection of few lines, mainly Br\( \alpha \), Pf\( \alpha \) and Hu\( \gamma \), 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\( \alpha \). In addition, in the OST-Wide survey few galaxies have also detected Hu\( \delta \), Hu\( \gamma \) and Pf\( \alpha \), 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).

### 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 [SIII] lines at 18 and 33 µm, can be used to investigate the gas density. The use of two lines of the same element, but of two different ionisation levels, like [NeIII] at 15.55 µm and [NeII] at 12.81 µm (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 [NeV] at 14.32 and 24.31 µm, is an indication of AGN activity (e.g. Spinoglio & Malkan 1992; Genzel & Cesarsky 2000; Meléndez et al. 2008; Feltre, Charlot, & Gutkin 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); Fernández-Ontiveros et al. (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 $1.5 \mu m < \lambda < 6.5 \mu m$ 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, that is $\lambda > 37 \mu m$, as OST/OSS covers up to $589 \mu m$. 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 µm, are expected to be detected in $5 \times 10^5$ 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 µm and [O IV] at 25.9 µm, 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 spectrophotometric 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
Figure 15. Same as of Figure 14, but for the SPICA DS.

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 $R = 300$. 
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 µm (B-BOP) and observations with the simultaneous SMI-LR spectrometer ranging from 17 to 36 µm. We considered a 5σ depth of 3 and 13 µJy at 34 µm and 60 and 100 µJy at 70 µm, in the UDS and DS respectively. We also simulated two surveys for OST, a Deep one of 0.5 deg² and a Wide one of 20 deg², and other two surveys of 3 and 30 deg² for GEP.

Our main findings are:

- The most luminous galaxies would be observed up to \(z = 10\), in the SPICA DS of 15 deg², 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 µm 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 µm 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 µm in the SPICA UDS (DS) would have soft 0.5–2 keV X-ray fluxes above \(4 \times 10^{-17}\) erg s⁻¹ cm⁻², that is 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 \(z \sim 5–8\), the redshift estimation could be based on the Brα and Brβ 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.
- On average, 26% (37%) of AGN detected at 34 or 70 µm in the SPICA UDS (DS) would have soft 0.5–2 keV X-ray fluxes above \(4 \times 10^{-17}\) erg s⁻¹ cm⁻², that is 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 \(z \sim 5–8\), the redshift estimation could be based on the Brα and Brβ 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 \(z = 3–8\) 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 µm and at 24.31 µm and the [Ne II] at 12.81 µm.
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.

Acknowledgements. This paper is dedicated to the memory of Bruce Swinyard, who initiated the SPICA project in Europe, but unfortunately died on 2015 May 22 at the age of 52. He was ISO-LWS calibration scientist, Herschel-SPIRE instrument scientist, first European PI of SPICA and first design lead of SAFARI. We acknowledge the whole SPICA Collaboration Team. LB, CG, LS and JFaro 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 MIUR 2017-2019 MFWV_001. MPS acknowledges support from the Comunidad de Madrid through the Atención de Talentos Investigador Grant 2018-T1/TIC-11035 and PID2019-105423GA-I00 (MCIU/AEI/FEDER,UE). FJC acknowledges 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 Marfa de Maestu, 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
Bisigello, L., et al. 2021, A&A, 651, A52
Bouwens, R. J., et al. 2010, MNRAS, 401, 2531
Decataldo, D. 2020, MNRAS, 497, 956
Donald, J. S., et al. 2017, MNRAS, 466, 861
Ellis, R. S., & Frenk, C. S. 1990, MNRAS, 242, 318
Feltre, A., Charlot, S., & Gutkin, J. 2016, MNRAS, 456, 3354
Feltre, A., et al. 2013, MNRAS, 434, 2426
Fernández-Ontiveros, J. A., Tommasin, S., & Busquet, G. 2015, ApJ, 799, 21
Fernández-Ontiveros, J. A., Pereira-Santaella, M., Carniani, S., Gallerani, S., & Decataldo, D. 2020, MNRAS, 497, 956
Hasinger, G., & Malkan M. A. 1992, ApJ, 399, 504
Hasinger, G., et al. 2007, ApJS, 172, 29
Hatziminaoglou, E., Hernán-Caballero, A., Feltre, A., & Piñol Ferrer, N. 2015, PASA, 34, e055
Hernández-Sánchez, P., Elbaz, D., & BP, J. 2011, ApJ, 751, 19
Hoekstra, H., et al. 2007, ApJ, 655, 67
Huertas-Company, M., et al. 2016, MNRAS, 459, 2618
Kaneda, H., et al. 2017, PASA, 34, e059
Kendrew, S., et al. 2015, PASP, 127, 623
Kulkarni, V. P., Torres-Garcia, L. M., Som, D., York, D. G., Welty, D. E., & Vladilo, G. 2011, ApJ, 726, 14
Le Févre, O., et al. 2020, A&A, 643, A1
Le Fèvre, O., et al. 2019, in UV/Optical/IR Space Telescopes and Instruments: Innovative Technologies and Concepts IX, Vol. 11115, SPIE, ed. A. A. Barto, J. B. Breckinridge, & H. P. Stahl, 184, 10.1117/12.2530514
Limber, D. N. 1953, ApJ, 117, 134
Livermore, R. C., Finkelestein, 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, A&A, 482, 94
Miyaji, T., et al. 2015, ApJ, 804, 104
Mordoni, S., Spinoglio, L., & Hernández-Ontiveros, J. A. 2021, ApJ, 853, 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, PASP, 127, 584
Roeckesa, P. M., et al. 2018, PASA, 35, e030
Sanders, D. B., et al. 2007, ApJS, 172, 86
Schenker, M. A., et al. 2013, ApJ, 768, 196
Shipley, H. V., Papovich, C., Rieke, G. H., Brown, M. J. L., & Moustakas, J. 2016, ApJ, 818, 60
Smith, M. W. L., et al. 2012, ApJ, 748, 123
Sommovigo, L., Ferrara, A., Pallottini, A., Carniani, S., Galleroni, 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., Dausy, 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., Giuseppe, M., Zamorani, G., Bethermin, M., Faissart, A., Le Fèvre, O., & Smolčič, 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., Gonzalez-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, Nat, 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, PASP, 127, 595
Wuyts, E., et al. 2014, ApJ, 789, L40