Euclid view of the dusty star-forming galaxies at $z \gtrsim 1.5$ detected in wide area submillimetre surveys

Dipanjan Mitra,[1]★ Mattia Negrello [®],[1] Gianfranco De Zotti [®][2] and Zhen-Yi Cai[3]

[1]*School of Physics and Astronomy, Cardiff University, The Parade, Cardiff CF24 3AA, UK*
[2]*INAF, Osservatorio Astronomico di Padova, Vicolo Osservatorio 5, I-35122 Padova, Italy*
[3]*Department of Astronomy, University of Science and Technology of China, Hefei 230026, China*

ABSTRACT

We investigate the constraints provided by the *Euclid* space observatory on the physical properties of dusty star-forming galaxies (DSFGs) at $z \gtrsim 1.5$ detected in wide area submillimetre surveys with *Herschel*. We adopt a physical model for the high-z progenitors of spheroidal galaxies, which form the bulk of DSFGs at $z \gtrsim 1.5$. We improve the model by combining the output of the equations of the model with a formalism for the spectral energy distribution (SED). After optimizing the SED parameters to reproduce the measured infrared luminosity function and number counts of DFSGs, we simulated a sample of DSFGs over $100\,\mathrm{deg}^2$ and then applied a $5\,\sigma$ detection limit of 37 mJy at $250\,\mu\mathrm{m}$. We estimated the redshifts from the *Euclid* data and then fitted the *Euclid+Herschel* photometry with the code CIGALE to extract the physical parameters. We found that 100 per cent of the *Herschel* galaxies are detected in all 4 *Euclid* bands above $3\,\sigma$. For 87 per cent of these sources the accuracy on $1 + z$ is better than 15 per cent. The sample comprises mostly massive, i.e. $\log(M_\star/\mathrm{M}_\odot) \sim 10.5$–12.9, highly star forming, i.e. $\log(\mathrm{SFR}/\mathrm{M}_\odot\mathrm{yr}^{-1}) \sim 1.5$–4, dusty, i.e. $\log(M_{\mathrm{dust}}/\mathrm{M}_\odot) \sim 7.5$–9.9, galaxies. The measured stellar masses have a dispersion of 0.19 dex around the true value, thus showing that *Euclid* will provide reliable stellar mass estimates for the majority of the bright DSFGs at $z \gtrsim 1.5$ detected by *Herschel*. We also explored the effect of complementing the *Euclid* photometry with that from the *Vera C. Rubin Observatory/LSST*.

Key words: galaxies: evolution – galaxies: formation – galaxies: general – galaxies: high-redshift – galaxies: photometry – galaxies: star formation.

1 INTRODUCTION

In the 1990s, surveys using the Submillimeter Common User Bolometer Array (SCUBA) at $850\,\mu\mathrm{m}$ on the 15 m *James Clerk Maxwell Telescope* (*JCMT*; Smail, Ivison & Blain 1997) for the first time detected a population of high-redshift ($z \gtrsim 1.5$) galaxies that are extremely luminous at far-infrared (FIR) and submillimetre (sub-mm) wavelengths and are highly dust-obscured at optical wavelengths. These are now commonly referred to as dusty star-forming galaxies (DSFGs) or submillimetre galaxies (SMGs) and are the dominant star-forming galaxies at $z \gtrsim 1.5$ (Casey et al. 2012; Bothwell et al. 2013; Cai et al. 2013) and most likely to be in the early evolutionary phase of present-day elliptical galaxies (Fu et al. 2013; Ivison et al. 2013; Dokkum et al. 2015; Simpson et al. 2017). These galaxies undergo intense star formation with star formation rates (SFRs) from a few hundreds to thousands solar masses per year (Blain et al. 2002; Magnelli et al. 2012; Riechers et al. 2013; Swinbank et al. 2014; MacKenzie et al. 2017; Michałowski et al. 2017; Gullberg et al. 2019; Castillo et al. 2023). During the cosmic noon (at $z \sim 2$) (Lilly et al. 1996; Madau et al. 1996) when the star formation rate density (SFRD) of the Universe peaked, these galaxies contributed heavily

to the overall galaxy population (Magnelli et al. 2011; Burgarella et al. 2013; Rowan-Robinson et al. 2016). Therefore, the discovery of these galaxies caused a big revolution in the field of extragalactic astronomy making their study vital for a detailed understanding of the formation and evolution of galaxies, which led to the development of more sensitive instruments like the SCUBA-2 (Holland et al. 2013), the Large Apex BOlometer CAmera (Siringo et al. 2009), the AZtronomical Thermal Emission Camera (AzTEC; Wilson et al. 2008), and the *Herschel Space Observatory* (*Herschel*; Pilbratt et al. 2010).

Our understanding of the DSFGs has increased drastically thanks to *Herschel* which, during its 4 yr of operation, from 2009 to 2013, has mapped ~ 1300 square degrees of the sky at wavelengths in the range of 100–$500\,\mu\mathrm{m}$ leading to the detection of more than a million FIR and sub-mm bright galaxies. In particular, the *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al. 2010) and the *Herschel* Multitiered Extragalactic Survey (Oliver et al. 2012) together surveyed over ~ 930 square degrees of the sky at 250, 350, and $500\,\mu\mathrm{m}$. Despite such surveys, our understanding of the underlying physical processes responsible for the evolution of galaxies is far from complete, but it is speculated that there will be a clearer understanding of these processes in the next decade with big projects like the *Euclid* space observatory (Laureijs et al. 2011), which was launched on 2023 July 1, and the *Vera*
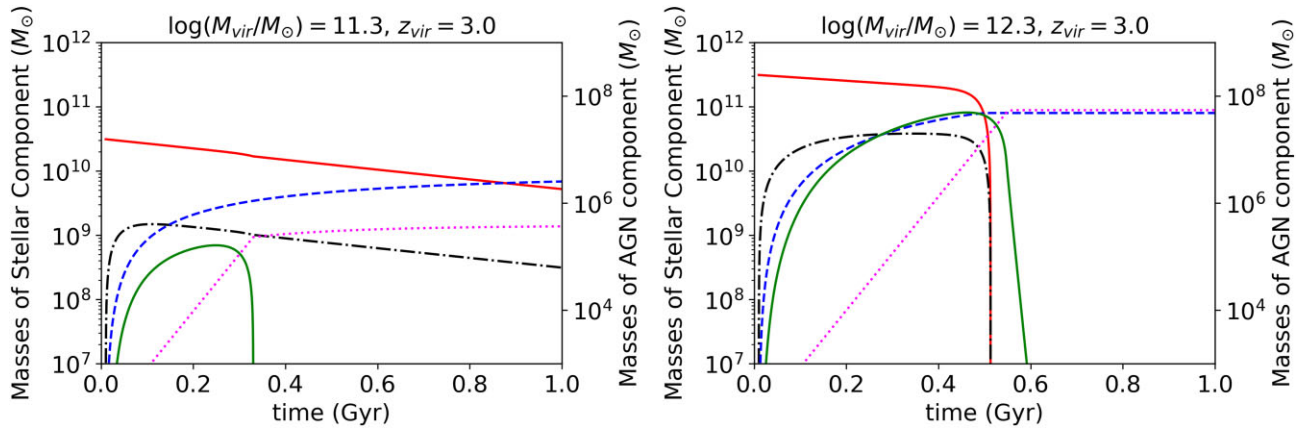
★ E-mail: mitraD@cardiff.ac.uk

Figure 1. Evolution of the properties of the stellar and AGN components of proto-spheroidal galaxies as a function of galactic age for a halo of mass $M_{vir} = 10^{11.3}$ M$_\odot$ (left) and a halo of mass $M_{vir} = 10^{12.3}$ M$_\odot$ (right). Both have virialization redshift $z_{vir} = 3$. In each plot, the left y-axis refers to the mass evolution of the stellar component: infalling gas (solid red line), cold gas (dot–dashed black line), and stellar mass (blue dashed line); while the right y-axis refers the mass evolution of AGN component: gas falling in the reservoir (solid green line) and supermassive black hole (SMBH) mass (dotted magenta line). We observe that the stellar mass increases as more and more cold gas is condensed into stars, thus leading to a corresponding steady decline in the mass of the gas components. At the same time, the central BH mass increases as a fraction of the cold gas loses its angular momentum via the stellar radiation drag and settles down into a reservoir around the central SMBH. The BH accretes matter from the reservoir via viscous dissipation, which powers the nuclear activity. With the evolution of the galactic age, the AGN feedback sets in. Its effect can be observed in the most massive DM haloes where the star formation activity and, consequently, the stellar mass growth, are stopped within less than 1 Gyr from the galaxy's birth.

C. Rubin Observatory/LSST (Ivezić et al. 2019). The *Euclid* Wide Survey (EWS; Euclid Collaboration 2022b) will observe 15 000 square degrees of the sky at both optical (Cropper et al. 2018) and near-infrared (NIR; Euclid Collaboration 2022a) wavelengths, thus providing crucial constraints on the stellar mass of galaxies. On the other hand, *Rubin's* LSST (Legacy Survey of Space and Time) Survey will observe 18 000 square degrees of the sky in 6 bands – u, g, r, i, z, and y – having a wavelength coverage of 0.36–1.01 µm.

Many sophisticated phenomenological models have been developed to study the cosmological evolution of galaxies (Matteucci et al. 2009; Franceschini et al. 2010; Popesso et al. 2011; Béthermin et al. 2012). These models adopt simple functional forms to describe the luminosity function of multiple galaxy populations, which are assigned different spectral energy distributions (SEDs) and varying evolutionary properties. Active galactic nuclei (AGNs) are also considered in some of the models. In this paper, we adopt instead the physically motivated model by Cai et al. (2013, C13 hereafter) for the formation and evolution of $z \gtrsim 1.5$ DSFGs, which links the star formation activity and the growth of a central supermassive black hole in a self-consistent way. The C13 model has proved successful in reproducing data on luminosity functions, number counts, and redshift distributions over a broad wavelength range, from the mid-IR to the millimetre. We use the model to investigate how *Euclid* can detect and study the $z \gtrsim 1.5$ proto-spheroidal galaxies discovered by the H-ATLAS survey. For this purpose, we improve on the C13 original model by implementing a more sophisticated formalism for the SED, which allows us to predict the optical/near-IR properties of these galaxies.

The paper is organized in the following manner. In Section 2, we provide a brief description of the model and theory used in this paper. In Section 3, we present a detailed methodology of catalogue creation, deriving photometric flux densities and estimating photometric redshifts and other physical properties using SED fitting. In Section 4, we discuss the surveys for which we make forecasts. The results obtained are analysed and discussed in Section 5 and the final conclusions are summarized in Section 6. Throughout the paper, we adopt a flat Λ-CDM cosmology with present-day matter and

baryon density (in units of the critical density), $\Omega_{m,0} = 0.3153$ and $\Omega_{b,0} = 0.0493$. We set the value of the Hubble–Lemaître constant to $h = H_0/100 = 0.6736$, the slope of the spectrum of primordial density perturbations to $n = 0.9649$ and the normalization of the density fluctuations on a scale of $8h^{-1}$ Mpc to $\sigma_8 = 0.8111$ (Planck Collaboration VI, 2020).

2 ADOPTED MODEL FOR THE PROTO-SPHEROIDAL GALAXIES

In this section, we provide a brief description of the C13 model for the proto-spheroidal galaxies and illustrate the formalism used for predicting the SED of these galaxies.

2.1 Model outline

The model by C13 is based on the fact that spheroids in the local universe (i.e. ellipticals and bulges of disc galaxies) are mostly comprised of old stellar populations that formed at redshifts $z \gtrsim 1.5$. On the other hand, disc galaxies consist of comparatively younger stellar populations with luminosity-weighted age $\lesssim 7$ Gyr, indicating a formation redshift $z \lesssim 1$. Therefore, the progenitors of local spheroids (also called proto-spheroidal galaxies or proto-spheroids) are the dominant star-forming galaxies at $z \gtrsim 1.5$, whereas star formation in discs takes place mostly at $z \lesssim 1.5$. The model provides a physically motivated co-evolution of the SFR of the proto-spheroids along with the supermassive black hole at their centres. The star formation history (SFH) of these proto-spheroids is calculated using a set of equations considering the evolution of gas phases and of the AGN, including cooling, condensation to form stars, accretion into the supermassive black hole, and feedback from both the stellar component and the AGN (see Fig. 1). The model depicts two ways of dark matter (DM) halo formation, i.e. a fast collapse of the halo triggering star formation, followed by a slow growth of the halo outskirts, which have little effect on the inner region of the halo. The first phase is the one driving the build-up of most of the stellar mass in these objects. Both the star formation and the growth of

the supermassive black hole at the centre are controlled by feedback mechanisms from the supernovae and the AGN. The AGN feedback is effective in quenching the star formation in the most massive proto-spheroids, while SFR is mainly affected by SN feedback in less massive ones (Granato et al. 2004; Scannapieco et al. 2008; Rosas-Guevara et al. 2016; Rosito et al. 2021). A detailed explanation of the equations governing the evolution of both the stellar and the AGN components is given in Appendix A.

The bolometric luminosity function of the proto-spheroidal galaxies, i.e. the volume density of galaxies per unit interval of luminosity, is calculated by convolving the halo formation rate function with the galaxy luminosity distribution predicated by the equations of the model, i.e.

$$\Phi(\log L, z) = \int_{M_{\mathrm{vir}}^{\mathrm{min}}}^{M_{\mathrm{vir}}^{\mathrm{max}}} \int_{z}^{z_{\mathrm{vir}}^{\mathrm{max}}} \left| \frac{\mathrm{d}t}{\mathrm{d}z_{\mathrm{vir}}} \right| \frac{\mathrm{d}N_{\mathrm{ST}}}{\mathrm{d}t}$$
$$P(\log L, z; M_{\mathrm{vir}}, z_{\mathrm{vir}}) \mathrm{d}z_{\mathrm{vir}} \mathrm{d}M_{\mathrm{vir}}, \tag{1}$$

where N_{ST} is the halo mass function, for which the model adopts the Sheth & Tormen (1999) approximation, while $P(\log L, z; M_{\mathrm{vir}}, z_{\mathrm{vir}})$ is the luminosity distribution function of galaxies at redshift z with virialization mass and redshift given by M_{vir} and z_{vir}, respectively. The distribution function is a lognormal distribution given by

$$P[\log L | \log \bar{L}] \mathrm{d} \log L = \frac{\mathrm{e}^{-\frac{\log^2(L/\bar{L})}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \mathrm{d} \log L, \tag{2}$$

where σ is the dispersion around the mean luminosity, $\bar{L}$, which is computed using the formalism illustrated in Section 2.2 for the stellar component. The AGN luminosity is computed using equation (A17). The total luminosity of a galaxy is the sum of the stellar luminosity and the AGN luminosity. The dispersion is considered to reflect the uncertainties in the adopted values of the main parameters that are involved in the equations of the model.

2.2 SED modelling

The SED of a galaxy is an observable that reflects the properties of the stars, dust, and gas in the galaxy. As such, it can be used to extract important information. In fact, while the UV/optical/near-IR part of the SED depends on the SFH, metallicity, and dust extinction, the far-IR/sub-mm/mm portion of the SED reveals the star formation activity occurring in the dust-enshrouded birth clouds and can be used to constrain the properties of the dust itself.

C13 adopted a single representative SED for the whole population of high-z proto-spheroidal galaxies, i.e. the SED of the well-studied $z = 2.3$ lensed galaxy SMM J2135-0102. That choice was dictated by the need of speeding up the computation of the relevant statistical properties on the galaxies, such as number counts and redshift distributions. Also, C13 were mainly focused on observables at far-IR/sub-mm/mm wavelengths where the SED depends on a very limited number of free parameters, e.g. the dust temperature and the dust emissivity index. In fact, with the SED of SMM J2135-0102, C13 did achieve a good qualitative fit to the multiwavelength (from the mid-IR to millimetre waves) data on number counts, both global and per redshift slices (e.g. Béthermin et al. 2012). Moreover, the model made a successful prediction of the sub-mm (Negrello et al. 2010, 2017) and mm (Vieira et al. 2010; Cai, Zotti & Bonato 2020; Cai, Negrello & Zotti 2022) number counts of strongly lensed galaxies.

However, the use of a single SED does limit the applicability of the model to the investigation and the forecast of the statistical properties of dusty proto-spheroidal galaxies at optical/near-IR

wavelengths, where the SED is much more complex than at longer wavelengths. Therefore, we have decided to improve the C13 model by implementing an SED modelling that reflects such complexity. The SED formalism is based on that put forward by da Cunha, Charlot & Elbaz (2008), with a few modifications, and it is described in this section.

2.2.1 Stellar component and dust attenuation

The stellar emission spectrum of the proto-spheroidal galaxies is calculated by coupling the SFH, as obtained from the equations of the model (equation A7) for any given value of M_{vir} and z_{vir}, with some single stellar population (SSP) synthesis templates. There are several SSP templates available in literature (Buzzoni 1993; Jimenez, Flynn & Kotoneva 1998; Bruzual & Charlot 2003). Here, we use the SSP templates by Bruzual & Charlot (2003), which are provided at wavelengths from 91 Å to 160 μm and between ages 1×10^5 and 2×10^{10} yr, for metallicities $Z = 0.008$, 0.02, and 0.05 and three IMFs, i.e. Salpeter, Korupa and Chabrier. We adopt the SSP template corresponding to $Z_{\odot}$ and a Chabrier IMF. Dust attenuation is modelled using the attenuation law by Charlot & Fall (2000). It takes into account attenuation by both birth clouds (BC) and the interstellar medium (ISM), represented by different power laws. This also helps in differentiating between the emission from the young stars still embedded in their dust-enshrouded birth clouds and the old stars located in the ISM, respectively, which is important for IR galaxies (Buat et al. 2018; Małek et al. 2018).

The luminosity emitted by a galaxy per unit interval of wavelength at time t is given by Charlot & Fall (2000)

$$L_\lambda(t) = \int_0^t \psi(t - t') S_\lambda(t') e^{-\hat{\tau}_\lambda(t')} dt' \tag{3}$$

where $\psi(t - t')$ is the SFR at $(t - t')$, and $S_\lambda(t')$ is the luminosity per unit interval of wavelength and per unit mass of an SSP at t'. The function $\tau_\lambda(t')$ represents the absorption optical depth of dust at the time t' and is defined as

$$\hat{\tau}_\lambda(t') = \begin{cases} \hat{\tau}_\lambda{}^{\mathrm{BC}} + \hat{\tau}_\lambda{}^{\mathrm{ISM}} & \text{for } t' \leq t_0, \\ \hat{\tau}_\lambda{}^{\mathrm{ISM}} & \text{for } t' > t_0 \end{cases} \tag{4}$$

t_0 being the time-scale of BC dissipation, while $\hat{\tau}_\lambda{}^{\mathrm{BC}}$ and $\hat{\tau}_\lambda{}^{\mathrm{ISM}}$ are the absorption optical depths of the dust in the BCs and in the ambient ISM, respectively, modelled as

$$\hat{\tau}_\lambda{}^{\mathrm{BC}} = (1 - \mu)\hat{\tau}_V \left(\frac{\lambda}{5500\,\text{Å}} \right)^{-1.3} \tag{5}$$

and

$$\hat{\tau}_\lambda{}^{\mathrm{ISM}} = \mu\hat{\tau}_V \left(\frac{\lambda}{5500\,\text{Å}} \right)^{-0.7}. \tag{6}$$

Here, $\hat{\tau}_V$ is the total effective V-band absorption optical depth of the dust in the BCs and $\mu = \frac{\hat{\tau}_V{}^{\mathrm{ISM}}}{\hat{\tau}_V{}^{\mathrm{BC}} + \hat{\tau}_V{}^{\mathrm{ISM}}}$ is the fraction contributed by the dust in the ISM.

Taking into consideration the AGN feedback which quenches star formation (see Fig. 2), we modify and re-define $\tau_\lambda(t')$ as

$$\hat{\tau}_\lambda(t') = \begin{cases} \hat{\tau}_\lambda{}^{\mathrm{BC}} + \hat{\tau}_\lambda{}^{\mathrm{ISM}} & \text{for } t' \leq \min(t_0, t_{\mathrm{AGN}}), \\ \hat{\tau}_\lambda{}^{\mathrm{ISM}} & \text{for } t' > \min(t_0, t_{\mathrm{AGN}}), \\ 0 & \text{for } t' > t_{\mathrm{AGN}} \end{cases} \tag{7}$$

where t_{AGN} is the time at which the AGN feedback has completely stopped the star formation. In practice, when this happens, all the gas and dust in the galaxy have been removed by the AGN feedback so that $\hat{\tau}_\lambda(t') = 0$.
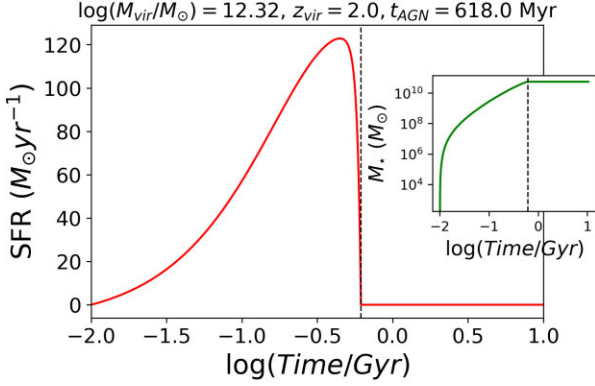
Figure 2. SFR as a function of the galaxy age in an object forming in a halo of mass $M_{\rm vir} = 10^{12.32}$ M$_\odot$ virialized at redshift $z_{\rm vir} = 2$. The inset plot shows the evolution of the stellar mass with galaxy age. The black dashed line marks the time ($t_{\rm AGN}$) when AGN feedback quenches SF in this galaxy and the SFR drops to zero. For this galaxy, $t_{\rm AGN} = 618.0$ Myr. As there is no more SF going on, the stellar mass remains constant afterwards as illustrated by the stellar mass versus time plot.

The amount of starlight absorbed by dust present in the BCs and in the ISM is reprocessed and re-radiated at infrared wavelengths. The dust luminosity is thus given by

$$L_{\rm total}^{\rm dust}(t) = L_{\rm BC}^{\rm dust}(t) + L_{\rm ISM}^{\rm dust}(t) \tag{8}$$

where

$$L_{\rm BC}^{\rm dust}(t) = \int_0^\infty (1 - e^{-\hat{\tau}_\lambda^{\rm BC}}) {\rm d}\lambda \int_0^{t_0} \psi(t - t') S_\lambda(t') {\rm d}t' \tag{9}$$

and

$$L_{\rm ISM}^{\rm dust}(t) = \int_0^\infty (1 - e^{-\hat{\tau}_\lambda^{\rm ISM}}) {\rm d}\lambda \int_{t_0}^{t} \psi(t - t') S_\lambda(t') {\rm d}t' \tag{10}$$

denote the amount of starlight absorbed and re-radiated by dust in the BCs and in the ISM, respectively. For later use in the analysis, we also define the fraction of IR luminosity coming from the dust in the ISM as

$$f_\mu = \frac{L_{\rm ISM}^{\rm dust}(t)}{L_{\rm total}^{\rm dust}(t)}, \tag{11}$$

Clearly, this fraction depends on $\hat{\tau}_V$, μ, and the SFH.

2.2.2 Infrared emission by dust

Here, we briefly describe the way in which we compute the infrared part of the SED, which is associated with the dust emission. We assume three main dust components: polycyclic aromatic hydrocarbons (PAHs), which produce strong emission in the NIR at wavelengths of about 3–20 μm, small dust grains (size < 0.01 μm), which are responsible for the hot mid-infrared (MIR) emission, and large dust grains (size $\sim 0.01 - 0.25$ μm) which are in thermal equilibrium with the radiation field and produce the colder FIR emission.

(i) *PAH emission:* in star-forming galaxies, strong emission features at 3.3, 6.2, 7.7, 8.6, 11.3, and 12.7 μm are observed. They are attributed to emission from PAH molecules. In the model, the PAH emission is represented by the spectrum of the photo-dissociation region (PDR) in the star-forming region M17 SW from ISOCAM observation of Cesarsky et al. (1996) and extracted by Madden et al. (2006). In order to include the 3.3 μm PAH emission feature in the template spectra we extend it bluewards beyond 5 μm using a

Lorentzian profile by Verstraete et al. (2001). The SED of the PAHs in the model is calculated as

$$l_{\lambda,{\rm PAH}} = L_{\lambda,M17} \left(\int_0^\infty L_{\lambda,M17} {\rm d}\lambda \right)^{-1}, \tag{12}$$

where $L_{\lambda,M17}$ is the template spectra used here.

(ii) *MIR emission:* Besides the emission from PAH molecules, the spectra of star-forming galaxies consist of a MIR continuum emission due to the heating of small dust grains to very high temperatures. This part of the spectra is computed using a 'greybody' function as

$$l_{\lambda,T_{\rm dust}} = \kappa_\lambda B_\lambda(T_{\rm dust}) \left(\int_0^\infty \kappa_\lambda B_\lambda(T_{\rm dust}) {\rm d}\lambda \right)^{-1}, \tag{13}$$

where $B_\lambda(T_{\rm dust})$ is the Planck function at temperature $T_{\rm dust}$ and κ_λ is the dust mass absorption coefficient of the form

$$\kappa_\lambda \propto \lambda^{-\beta}, \tag{14}$$

where β is the dust emissivity index. Studies show that $\beta \approx 1$ for carbonaceous grains (radiating energy in the MIR), while $\beta \approx 1.5$–2 for silicate grains (Hildebrand 1983; Draine & Lee 1984), radiating most of the energy in the far-IR/sub-mm wavelengths. In the model, the MIR continuum emission is characterized by the sum of two greybody functions at temperatures 130 and 250 K respectively, having equal contribution to the IR luminosity. Mathematically,

$$l_{\lambda,{\rm MIR}} = (l_{\lambda,130K} + l_{\lambda,250K}) \left(\int_0^\infty (l_{\lambda,130K} + l_{\lambda,250K}) {\rm d}\lambda \right)^{-1} \tag{15}$$

where $l_{\lambda,130K}$ and $l_{\lambda,250K}$ are calculated using equation (13). The value of β for the MIR emission is set to 1.

(iii) *Dust grains in equilibrium:* The FIR portion of the SED comprises emission from bigger dust grains which are in thermal equilibrium and have comparatively low temperatures. The model takes into account two types of grains: warm grains which can be present in BCs and in the ISM, with temperatures $T_{w,{\rm BC}}$ and $T_{w,ISM}$, respectively; and cold grains that are found in the ISM at temperatures $T_{c,ISM}$. The emissions are computed using greybody functions (equation 13) with $\beta = 1.5$ and 2.0, respectively. The total IR emission from BCs writes

$$L_{\lambda,{\rm BC}}^{\rm dust} = (\xi_{\rm BC}^{\rm PAH} l_{\lambda,{\rm PAH}} + \xi_{\rm BC}^{\rm MIR} l_{\lambda,{\rm MIR}} + \xi_W^{\rm BC} l_{\lambda,T_{w,{\rm BC}}})$$
$$\times (1 - f_\mu) L_{\rm total}^{\rm dust}(t) \tag{16}$$

satisfying the condition

$$\xi_{\rm BC}^{\rm PAH} + \xi_{\rm BC}^{\rm MIR} + \xi_W^{\rm BC} = 1 \tag{17}$$

Likewise, the ISM emission can be computed as

$$L_{\lambda,{\rm ISM}}^{\rm dust} = (\xi_{\rm ISM}^{\rm PAH} l_{\lambda,{\rm PAH}} + \xi_{\rm ISM}^{\rm MIR} l_{\lambda,{\rm MIR}} + \xi_W^{\rm ISM} l_{\lambda,T_{w,{\rm ISM}}}$$
$$+ \xi_C^{\rm ISM} l_{\lambda,T_{c,{\rm ISM}}}) \times f_\mu L_{\rm total}^{\rm dust}(t) \tag{18}$$

satisfying the condition

$$\xi_{\rm ISM}^{\rm PAH} + \xi_{\rm ISM}^{\rm MIR} + \xi_W^{\rm ISM} + \xi_C^{\rm ISM} = 1 \tag{19}$$

The adopted values of the fractions mentioned in equations (17) and (19) are discussed in Section 3.3. Fig. 3 shows an example of the different components that make up the modelled SED of a protospheroidal galaxy.

2.2.3 Emission associated with the AGN

The SED of the AGN component is modelled using the smooth torus model introduced by Fritz, Franceschini & Hatziminaoglou (2006,
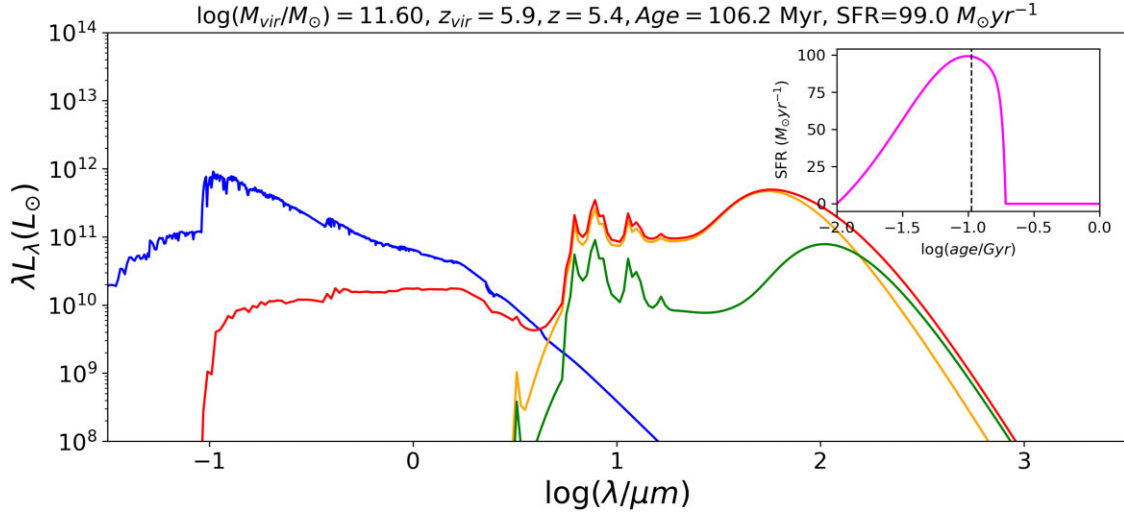
Figure 3. Example of the SED of a proto-spheroidal galaxy (red curve) with halo mass of $M_{\rm vir} = 10^{11.60}\,{\rm M}_\odot$ and formation redshift $z_{\rm vir} = 5.9$. The SED accounts for the effect of the suppression of the stellar emission in the UV/optical/near-IR due to dust absorption and for the reprocessing at longer wavelengths of the starlight that has been absorbed by dust. The galaxy is observed at $z = 5.4$ at an age of 106.2 Myr when the SFR is $\sim 99\,{\rm M}_\odot\,{\rm yr}^{-1}$. The evolution of the SFR with time is shown in the inset plot, where the dashed black line marks the age of the galaxy at the time of observation. The blue line shows the unattenuated stellar spectrum. The IR emission from the BCs is shown in orange while that from the ISM is shown in green.

hereafter F06). We considered a smooth torus model with the ratio of maximum to minimum radii $R_{\rm max}/R_{\rm min} = 10$–$100$. The torus amplitude angle, Θ, has values of $40°$, $100°$, and $140°$. The variation of the gas density inside the torus along the radial and angular direction is given by the equation $\rho(r,\theta) = \alpha r^\beta e^{-\gamma|\cos(\theta)|}$, where $\beta = 0$ and $\gamma = 0$ mean constant density. The model assumes that the luminosity of the central power source of the torus is described by a broken power law as $\lambda L_\lambda(\lambda) = L_0\lambda^\xi$ (erg s^{-1}), where $\xi = 1.2$ if $0.001\,\mu{\rm m} < \lambda < 0.03\,\mu{\rm m}$, $\xi = 0$ if $0.03\,\mu{\rm m} < \lambda < 0.125\,\mu{\rm m}$ and $\xi = -0.5$ if $0.125\mu{\rm m} < \lambda < 20\mu{\rm m}$. The equatorial optical depth at $9.7\,\mu{\rm m}$, $\tau_{\rm eq}(9.7)$, ranges from 0.1 to 10. The model considers 10 viewing angles ranging from $0°$ (type 1 AGN) to $90°$ (type 2 AGN). For more details about the model, we refer the reader to F06.

In our simulations, as explained in Section 3, we associate to each simulated galaxy an AGN SED by randomly sampling the available range of values of the SED parameters. We then normalize the SED so that the resulting bolometric AGN luminosity equals the one obtained from equation (A17).

Some of the light emitted from the AGN gets absorbed by the dust in the ambient ISM. So, the AGN luminosity emitted per unit wavelength at time t is given by

$$L_\lambda^{\rm AGN}(t) = \int_0^t L_\lambda e^{-\hat\tau_{\lambda,{\rm AGN}}(t')}{\rm d}t', \tag{20}$$

where L_λ is the luminosity per unit wavelength interval from the template. We define

$$\hat\tau_{\lambda,{\rm AGN}}(t') = \begin{cases} \hat\tau_\lambda^{\rm ISM} & {\rm for\ } t' < t_{\rm AGN}, \\ 0 & {\rm for\ } t' > t_{\rm AGN}. \end{cases} \tag{21}$$

The total AGN luminosity absorbed by dust in the ISM is then given by

$$L_{\rm ISM}^{\rm dust,AGN}(t) = \int_0^\infty \int_0^t (1 - e^{-\hat\tau_{\lambda,{\rm AGN}}(t')}) L_\lambda\,{\rm d}t'\,{\rm d}\lambda. \tag{22}$$

Therefore, the total dust luminosity of the galaxy is redefined as

$$L_{\rm total}^{\rm dust}(t) = L_{\rm BC}^{\rm dust}(t) + L_{\rm ISM}^{\rm dust}(t) + L_{\rm ISM}^{\rm dust,AGN}(t), \tag{23}$$

and the fraction of the dust luminosity coming from the ISM is now

$$f_\mu = \frac{L_{\rm ISM}^{\rm dust}(t) + L_{\rm ISM}^{\rm dust,AGN}(t)}{L_{\rm total}^{\rm dust}(t)}. \tag{24}$$

The absorbed AGN luminosity is reprocessed by dust and emitted at IR wavelengths. Likewise, the total IR emission from the BC and the ISM is now re-defined as

$$L_{\lambda,{\rm BC}}^{\rm dust} = (\xi_{\rm BC}^{\rm PAH}l_{\lambda,{\rm PAH}} + \xi_{\rm BC}^{\rm MIR}l_{\lambda,{\rm MIR}} + \xi_W^{\rm BC}l_{\lambda,T_{w,{\rm BC}}})$$
$$\times(1 - f_\mu)(L_{\rm BC}^{\rm dust}(t) + L_{\rm ISM}^{\rm dust}(t) + L_{\rm ISM}^{\rm dust,AGN}(t)), \tag{25}$$

and

$$L_{\lambda,{\rm ISM}}^{\rm dust} = (\xi_{\rm ISM}^{\rm PAH}l_{\lambda,{\rm PAH}} + \xi_{\rm ISM}^{\rm MIR}l_{\lambda,{\rm MIR}} + \xi_W^{\rm ISM}l_{\lambda,T_{w,{\rm ISM}}}$$
$$+ \xi_C^{\rm ISM}l_{\lambda,T_{c,{\rm ISM}}}) \times f_\mu(L_{\rm BC}^{\rm dust}(t) + L_{\rm ISM}^{\rm dust}(t) + L_{\rm ISM}^{\rm dust,AGN}(t)). \tag{26}$$

respectively.

Fig. 4 shows the F06 AGN template SED as a function of wavelength for different viewing angles (from $0°$ to $90°$) and a constant density profile of the gas in the torus. An example of the SED of a proto-spheroidal galaxy with a visible AGN component is shown in Fig. 5.

Apart from the dusty torus and the dust in the ISM, 'polar dust' can also contribute to the emission from the AGN. This dust extends in the polar direction of the dusty torus over parsec scales (Hönig et al. 2013; Yang et al. 2020). Its emission peaks in the mid-IR (Hönig et al. 2013, see fig. 8), and it is negligible at the wavelengths of interest here, i.e. UV/optical/near-IR wavelengths. Hence, we do not include the effect of polar dust in our model.

2.2.4 Dust mass estimation

To estimate the dust mass in galaxies, we follow the prescriptions of Cunha et al. (2008), which are based on the formula by Hildebrand (1983). For dust grains in thermal equilibrium at the temperature T_d the dust mass (M_d) is given by

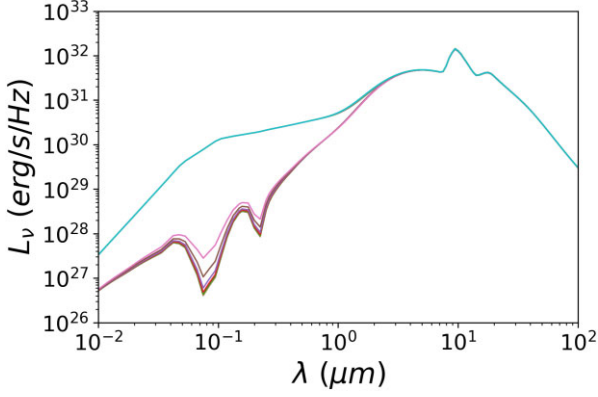$$L_\lambda^{T_d} = 4\pi M_d \kappa_\lambda B_\lambda(T_d), \tag{27}$$

is estimated as

$$M_{\text{dust}}^{\text{total}} \approx 1.1 \left(M_{\text{dust,W}}^{\text{BC}} + M_{\text{dust,W}}^{\text{ISM}} + M_{\text{dust,C}}^{\text{ISM}} \right). \tag{31}$$

3 BUILDING A SIMULATED SAMPLE OF PROTO-SPHEROIDAL GALAXIES

Here, we describe the way in which we implement the formalism presented in the previous section to generate a sample of simulated proto-spheroidal galaxies. The sample will then be used for forecast studies.

3.1 Step 1: sampling the halo formation rate function

We start by randomly sampling the halo formation rate function (equation A25) in both halo virialization mass and virialization redshift, within the intervals $z_{\text{vir}} \in [1.5, 12]$ and $\log(M_{\text{vir}}/\text{M}_\odot) \in [11.3, 13.3]$, respectively. The total number of randomly sampled pairs of M_{vir} and z_{vir} is dictated by the adopted survey area, which, in turn, determines the sampled volume at any given virialization redshift. For each simulated pair of values of M_{vir} and z_{vir}, we then solve the C13 equations to compute the SFH as a function of the galaxy age, with the birth of the galaxy set at the corresponding virialization redshift.

3.2 Step 2: generating the UV/optical/near-IR part of the SED

At this point, we apply the SED formalism illustrated in Section 2.2.1 to compute the SED of the dust-attenuated stellar component (equation 3) as well as the SED of the dust-attenuated AGN (equation 20) at any given age of the galaxy. The main parameters affecting the stellar component are the age of the BCs, t_0, and the total effective V-band absorption optical depth of the dust in the BCs, $\hat{\tau}_V$, and in the ISM, $\hat{\tau}_V^{\text{ISM}}$. For them, we adopt the values measured by Rowlands et al. (2014, R14 hereafter), who performed SED fitting on a sample of 29 250 μm-selected high-redshift DSFGs from H-ATLAS, using MAGPHYS (da Cunha & Charlot 2011). They obtained a good

Figure 4. AGN SEDs as a function of wavelength for 10 different viewing angles from 0° (cyan, topmost curve) to 90° (blue, lowest curve) with $\tau_{\text{eq}}(9.7) = 0.1$ and $\Theta = 140°$.

where L_λ^d is the infrared luminosity of the dust grains, while $B_\lambda(T_d)$ and κ_λ are defined in equations (13) and (14), respectively. After normalizing κ_λ using $\kappa_{850\,\mu m} = 0.77\ \text{g}^{-1}\ \text{cm}^2$, the mass contribution of warm dust in BC and ISM is calculated as

$$M_{\text{dust,W}}^{\text{BC}} = \xi_W^{\text{BC}}(1 - f_\mu)L_{\text{total}}^{\text{dust}} \left(4\pi \int_0^\infty \kappa_\lambda B_\lambda(T_{w,\text{BC}}) d\lambda \right)^{-1} \tag{28}$$

and

$$M_{\text{dust,W}}^{\text{ISM}} = \xi_W^{\text{ISM}} f_\mu L_{\text{total}}^{\text{dust}} \left(4\pi \int_0^\infty \kappa_\lambda B_\lambda(T_{w,\text{ISM}}) d\lambda \right)^{-1}, \tag{29}$$

respectively, and that from the cold dust in the ISM is calculated as

$$M_{\text{dust,C}}^{\text{ISM}} = \xi_C^{\text{ISM}} f_\mu L_{\text{total}}^{\text{dust}} \left(4\pi \int_0^\infty \kappa_\lambda B_\lambda(T_{c,\text{ISM}}) d\lambda \right)^{-1}. \tag{30}$$

Taking into account the contribution from stochastically heated dust grains and considering only a small contribution of a few per cent from PAH grains (Draine & Li 2007), the total dust mass of a galaxy
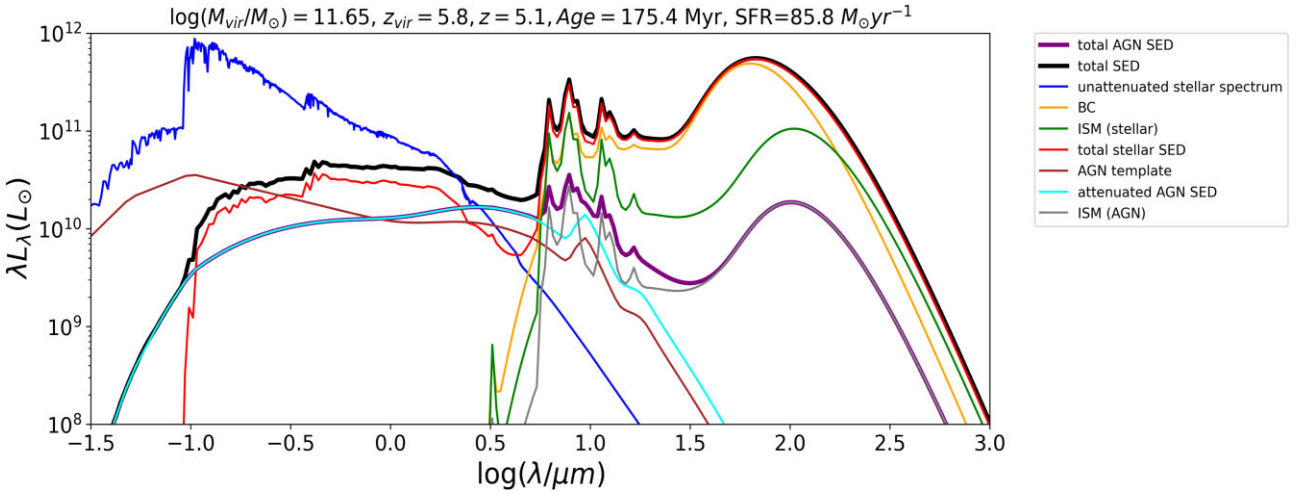

Figure 5. Example of the SED (stellar + AGN components) of a proto-spheroidal galaxy (black curve) with halo mass of $M_{\text{vir}} = 10^{11.65}\ \text{M}_\odot$ and formation redshift $z_{\text{vir}} = 5.8$. The galaxy is observed at $z = 5.1$ at an age of 175.4 Myr when the SFR is $\sim 86\ \text{M}_\odot\ \text{yr}^{-1}$. The blue curve is the unattenuated stellar spectrum, while the red curve shows the starlight attenuated by dust in the UV/optical/near-IR and reprocessed by the same dust at longer wavelengths. The IR emission from the BCs is shown in orange while that from the ISM is shown in green. The unattenuated SED of the AGN component is shown in brown while the attenuated AGN SED is shown in cyan. The grey curve depicts the IR emission from the ISM due to the absorption of AGN luminosity by ISM dust. The purple curve shows the total AGN SED. The AGN SED is typical of a Type 2 AGN ($\theta = 90°$) with density profile $\rho(r) = \alpha r^{-0.5}$, $\tau_{\text{eq}}(9.7) = 10$, $R_{\text{max}}/R_{\text{min}} = 60$, and $\Theta = 100°$.

fit by allowing t_0 to vary uniformly in logarithmic space between 10^7 and 10^8 yr, while for $\hat{\tau}_V$ and $\hat{\tau}_V^{ISM}$ they derived mean values of 5.1 ± 0.6 and 1.0 ± 0.1, respectively. In practice, we take Gaussian priors for $\tau_V^{\hat{B}C}$ and $\tau_V^{I\hat{S}M}$ with mean 5.1 and 1.0, respectively. The standard deviation is taken as 0.6 and 0.1, respectively, for all the simulated objects, and the value of $\log(t_0\,\mathrm{yr}^{-1})$ is randomly assigned to each object within the range of 7–8.

Before moving on, we checked that the infrared luminosity function (LF) of the simulated sample is consistent with current measurements. For a given redshift, we first computed the age of each object at that redshift and then used equation (23) to compute the corresponding dust luminosity, L_d. Because the latter represents the amount of energy re-emitted in the far-IR by dust, we took L_d to be the total infrared luminosity, L_{IR}. We then binned the objects in L_{IR} to generate a simulated infrared luminosity function to be compared with the data.

The results are illustrated in Fig. 6 for $z = 2$, 3, and 4. The simulated LF (red curve) shows a very good agreement with the C13 model (black curve) at moderate to bright infrared luminosities, i.e. above $L_{IR} \gtrsim 10^{11.5}\,L_\odot$. However, it lies significantly below the prediction from the C13 model at low luminosities. This is a consequence of the way C13 computed the infrared luminosity. Indeed, they calculated L_{IR} from the SFR of the objects using the relation by Kennicutt (1998, K98 hereafter). However, that relation only applies during the early dust-obscured phase of the evolution of galaxies, when the far-IR emission is mainly associated with BCs where the dust is heated by young massive stars (R14). It does not hold for more evolved, low-luminosity, galaxies where the ISM contributes a significant fraction of the dust emission.

This is illustrated in Fig. 7 where the SFR of our simulated galaxies is shown as a function of their infrared luminosity, computed using equation (8), and colour-coded according to the value of f_μ. A high value of f_μ corresponds to a dominant contribution of the ISM to the dust luminosity. Below $\log(L_{IR}/L_\odot) \simeq 11$, where $f_\mu > 0.5$, there is a clear deviation from the K98 relation, with the latter significantly underestimating the infrared luminosity for any given SFR. As a result, more objects with lower L_{IR} are expected when using the K98 relation. This is the reason why our simulated LF curve lies below the C13 one at lower luminosities.

The shape of the LFs is highly dependent on the values of the parameters that determine L_d via equation (8), especially of t_0, τ_V, and τ_V^{ISM}. Fig. 8 shows how the IR LF at $z = 2$ is affected by changes in the values of the above parameters. The parameters are changed one at a time while keeping the others fixed at their prescribed values. It is evident that a deviation from the R14 values produces LFs that lie below the reference LF by C13 (black curve).

By increasing the dissipation timescale of the BC, stars are allowed to reside inside the BC for a longer time. One would expect that a prolonged dust absorption of UV/optical light from the stars will lead to a higher IR luminosity, hence to an increased infrared LF. But, in reality, that is not the case. The massive stars, which contribute the bulk of the IR luminosity, have significantly evolved by the time the BC dissipates and hence, have reduced luminosity. So, there are fewer UV/optical photons being emitted from them to be absorbed and re-emitted at IR wavelengths. Moreover, by increasing the value of t_0 and by getting it closer to t_{AGN}, the BCs are eventually destroyed by the AGN feedback. This effect also leads to a lower IR luminosity.

Similarly, by increasing τ_V^{ISM} while keeping τ_V constant (which is equivalent to decreasing τ_V^{BC}), a lower fraction of the starlight is re-radiated in the far-IR, leading to a decrease in the number of galaxies for any given L_{IR} compared to the reference case.
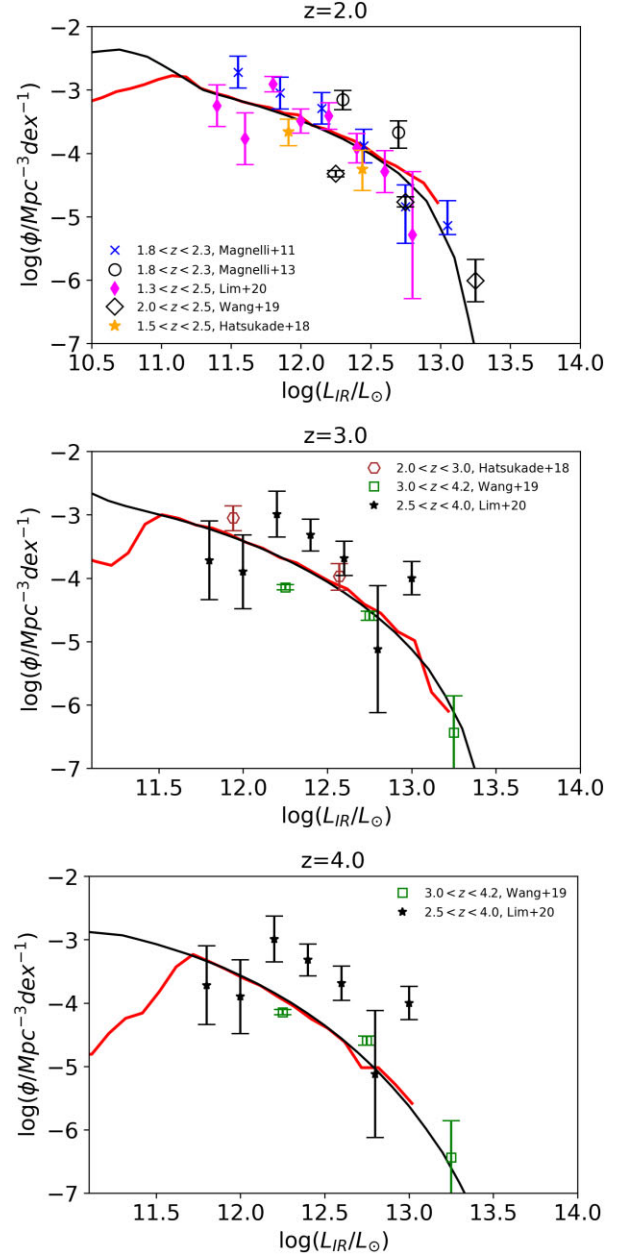
Figure 6. Comparison between the theoretical LF (black), the simulated curves (red) and observational data at different redshifts. The data points are taken from Magnelli et al. (2011), Magnelli et al. (2013), Hatsukade et al. (2018), Wang et al. (2019), and Lim et al. (2020).

3.3 Step 3: generating the far-IR/sub-mm/mm part of the SED

For each simulated object, we compute the far-IR/sub-mm/mm part of the SED by implementing the formalism illustrated in Section 2.2.2, and setting the values of the relevant parameters to those proposed by da Cunha et al. (2015). The indices β for warm and cold dust grains were fixed at 1.5 and 2, respectively. The temperature of the warm dust grains in the BCs, $T_{w,BC}$, is allowed to vary uniformly in the interval 30–80 K, while that of the ISM, $T_{w,ISM}$, is fixed at 45 K. The temperature of the cold grains in the ISM, $T_{c,ISM}$, is taken from a uniform distribution between 20 and 40 K. In the BCs, ξ_{BC}^{PAH} and ξ_{BC}^{MIR} are allowed to vary uniformly in the ranges 0–0.1 and 0.1–0.2, respectively. For each choice of both ξ_{BC}^{PAH} and ξ_{BC}^{MIR}, ξ_W^{BC}
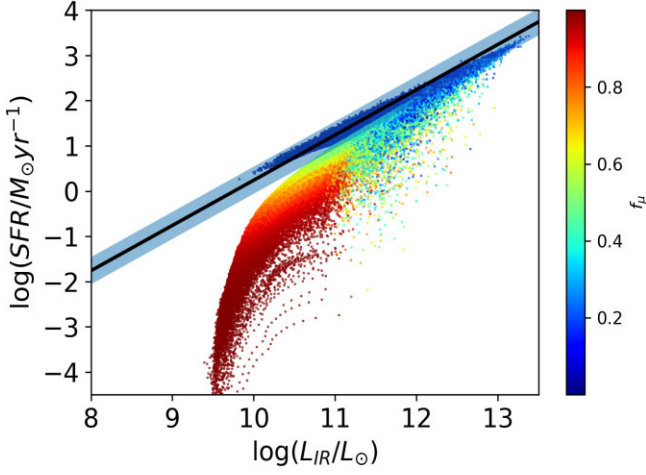
Figure 7. Correlation between SFR and total IR luminosity, L_{IR}, for a sample of simulated galaxies at $z = 2$. L_{IR} is calculated using equation (8). Data points are colour coded in accordance with the value of f_μ. The black line shows the relation between SFR and total IR luminosity (integrated from 8 to 1000 μm) from K98 with a 30 per cent dispersion shown by the shaded blue area. Galaxies with a low value of f_μ lie on the K98 relation. However, for galaxies where the ISM contributes significantly to the dust luminosity (i.e. those with high values of f_μ) the K98 relation leads to an underestimate of L_{IR}.

follows from equation (17). For the ISM, we extract ξ_C^{ISM} from a uniform distribution between 0.5 and 1. The other parameters are set to $\xi_{ISM}^{PAH} = 0.550(1 - \xi_C^{ISM})$, $\xi_{ISM}^{MIR} = 0.275(1 - \xi_C^{ISM})$, and $\xi_W^{ISM} = 0.175(1 - \xi_C^{ISM})$ so that equation (19) is satisfied.

With this parameter set, we generated differential number counts at 850, 870, and 1100 μm. We did that by first computing the objects' flux density at several redshifts from $z = 1$ to $z = 8$ in steps of 0.1. Each object was randomly assigned a redshift within the corresponding bin assuming a uniform distribution. We then computed the number of objects per unit interval of redshift and per unit interval of flux density at each redshift step. Finally, we integrated the redshift-dependent number counts to produce the differential number counts. The result is shown by the red curve in Fig. 9, where it is compared with the C13 number counts model curve (black line) of the proto-spheroids and with existing data (see also Table 1).

We notice that the simulated number counts significantly underestimate the observed number counts and the C13 curve at low flux densities, i.e. $\lesssim 6$ mJy, while producing, at the same time, an excess of bright objects. This result indicates that the assumed values for the parameters that describe the far-IR/sub-mm/mm SED of the proto-spheroids need to be revised. In principle, the best-fitting values of the SED parameters could be found using a minimum χ^2 approach on the number counts. However, that method, in this specific case, is very time consuming. Therefore, we opted instead, for a trial and error approach, which led to the following changes in the value of some of the SED parameters. We assigned a higher value to the dust emissivity index of the warm dust grains, raising it from $\beta = 1.5$ to $\beta = 2$, while we kept the one of the cold dust grains to its original value of $\beta = 2$. The temperatures $T_{w,BC}$ and $T_{c,ISM}$ were allowed to vary uniformly within the intervals 30–45 and 20–30 K, respectively. The value of $T_{w,ISM}$ was left unchanged. By using these updated values for the SED parameters we achieved a very good agreement with the observed number counts at the faint end, as illustrated by
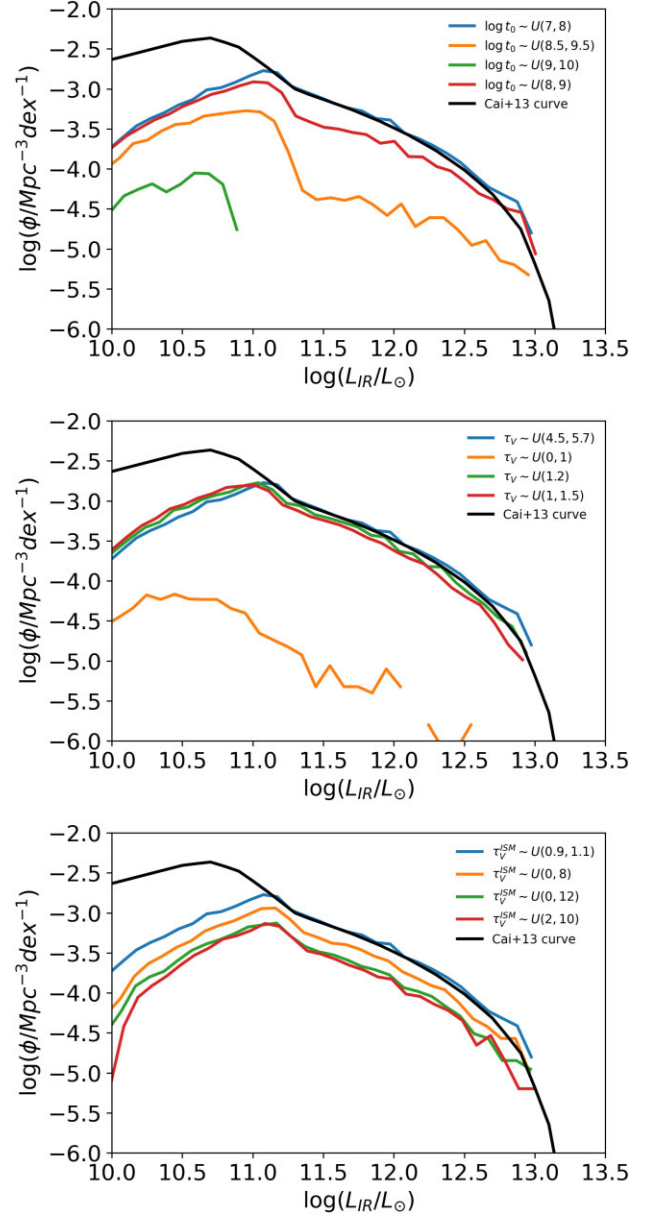

Figure 8. Effect of changing the values of t_0, $\hat\tau_V$, and $\tau_V^{I\hat SM}$ on the shape of the total IR LF at $z = 2$. In all panels, the black curve represents the theoretical curve obtained by C13 and the blue curve represents the simulated LF obtained by using the optimal values of the parameters: $\log t_0 \in U(7, 8)$, $\tau_V \in U(4.5, 5.7)$, and $\tau_V^{ISM} \in U(0.9, 1.1)$. *Top panel*: Varying t_0 uniformly in logarithmic scale in the ranges 8–9 (red), 8.5–9.5 (orange), and 9–10 (green), respectively. *Middle panel*: Varying τ_V uniformly in the ranges 0–1 (orange), 1–2 (green), and 1–1.5 (red), respectively. *Bottom panel*: Varying τ_V^{ISM} uniformly in the ranges 0–8 (orange), 0–12 (green), and 2–10 (red), respectively.

the solid cyan curve in Fig. 9. However, we were still left with an excess of objects at bright flux densities.

Fig. 10 shows that, as expected, the bright tail of the number counts is made up by the intrinsically brightest objects, i.e. those with total infrared luminosity $L_{IR} \gtrsim 10^{13} L_\odot$. Therefore, we assigned a separate set of SED parameter values to these very bright objects. Because the dust temperature increases with increasing infrared luminosity (Hwang et al. 2010), we let $T_{w,BC}$ and $T_{c,ISM}$ to vary uniformly in the range 30–70 K and 20–40 K, respectively. At the
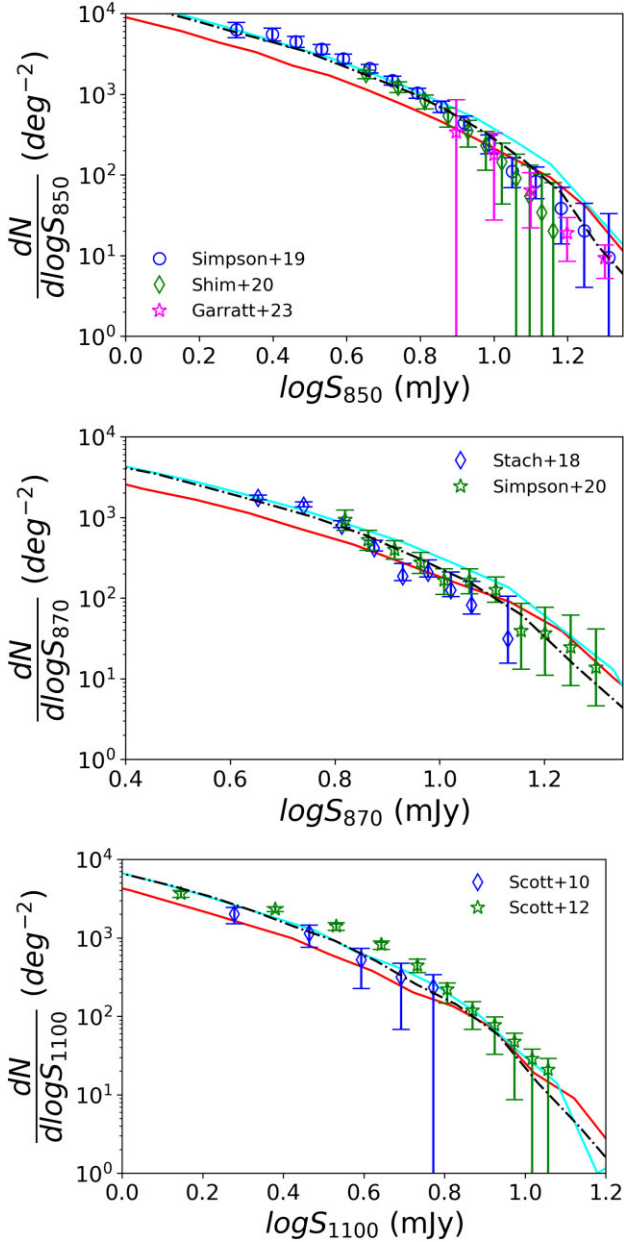
Figure 9. Comparison between the simulated differential number counts (red curve) with the C13 model (dashed black dots) and the observed data at 850, 870, and 1100 μm. The cyan curve shows the number counts obtained by modifying the parameter values by trial and error. The new curve fits the faint end but produces an excess of objects at the brightest luminosities. The data points are taken from Simpson et al. (2019), Shim et al. (2020), Garratt et al. (2023), Stach et al. (2018), Simpson et al. (2020), Scott et al. (2010), and Scott et al. (2012).

same time, we also decreased the dust emissivity index of the warm dust grains to $\beta = 1.8$. With these modifications, we obtained the number counts shown by the red solid curve in Fig. 11. The agreement with the measured number counts is now satisfactory across the whole range of flux densities probed by the data.

3.4 Step 4: generating the final sample

With the adopted values of the SED parameters, which reproduce both the infrared luminosity function and number counts of proto-

Table 1. Information about the measured differential number counts at 850, 870, and 1100 μm shown in Figs 9–11.

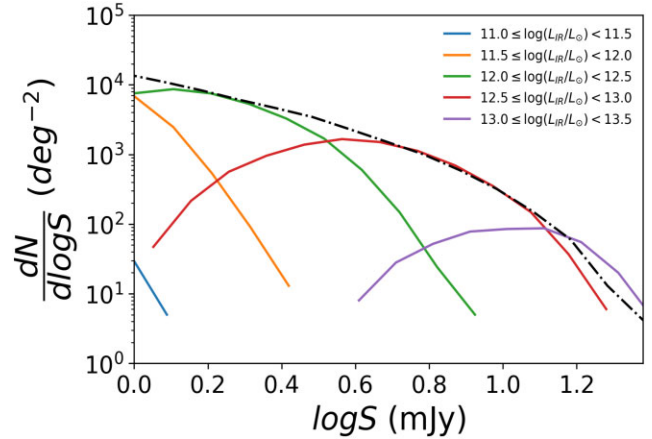| Wavelength (μm)	Instrument	Field	Reference
850	SCUBA-2	S2COSMOS	Simpson et al. (2019)
850	SCUBA-2	NEPSC2	Shim et al. (2020)
850	SCUBA-2	S2LXS XMM-LSS	Garratt et al. (2023)
870	SCUBA-2	S2CLS UDS	Stach et al. (2018)
870	SCUBA-2	S2COSMOS	Simpson et al. (2020)
1100	ASTE/AzTEC	GOODS-S	Scott et al. (2010)
1100	*JCMT*/AzTEC	GOODS-N	Scott et al. (2012)

Figure 10. Simulated differential number counts at 850 μm, with contributions from intervals of total infrared luminosity. The dot-dashed black curve represents the model number counts by C13.

spheroidal galaxies, we generate the final catalogue of simulated objects used for forecast studies. We simulate proto-spheroidal galaxies over an area of 100 deg^2, which is of the same order of the wide area surveys conducted with *Herschel*.

4 FORECASTS

We investigate the detectability of *Herschel*-selected sources in the *Euclid* wide area survey. To increase the wavelength coverage, we also test how many of these *Euclid* sources are detected by the *Rubin* observatory. We first provide a brief description of the *Herschel*, *Euclid*, and *Rubin's* wide-area surveys and then present the results of the analysis.

4.1 Surveys

4.1.1 The Herschel-ATLAS

The *Herschel*-ATLAS (H-ATLAS; Eales et al. 2010) is a survey of 600 deg^2 conducted in five photometric bands (100, 160, 250, 350, and 500 μm), using the Photoconductor Array Camera and Spectrometer (PACS) and the Spectral and Photometric Imaging Receiver (SPIRE) cameras. The H-ATLAS observed fields in the Northern and Southern hemispheres and on the celestial equator. The fields were chosen to minimize the dust emission from the Milky Way. The observed fields are as follows:

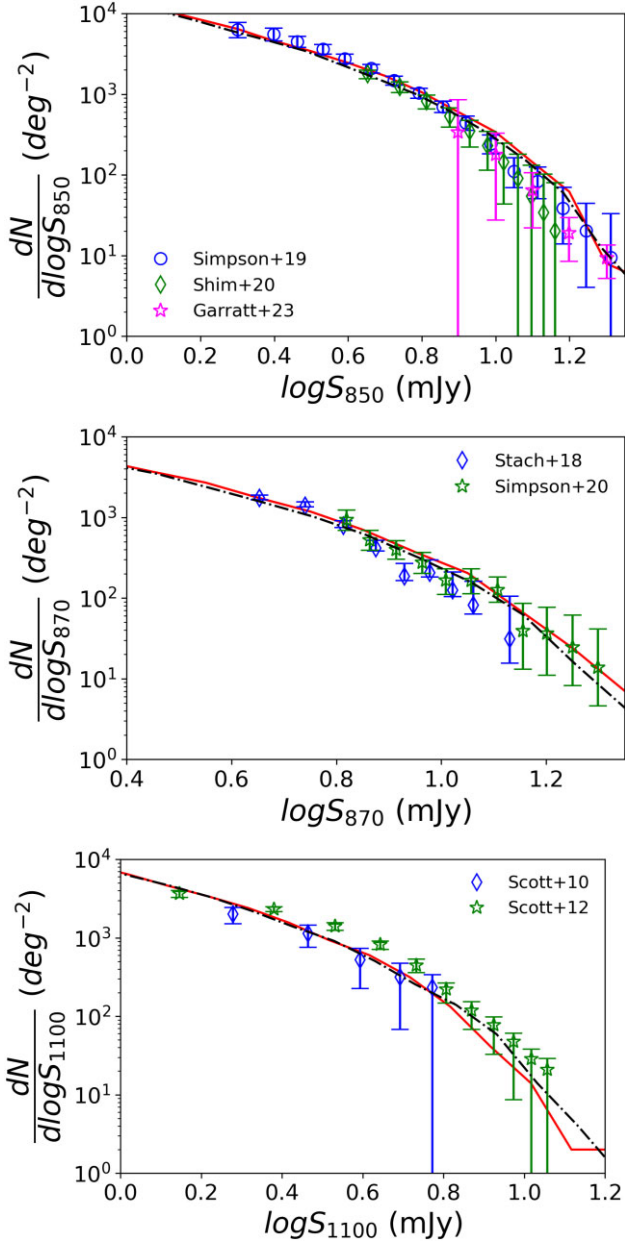(i) A field with an area of 150 deg^2 close to the North Galactic Pole.

Figure 11. Final version of the simulated differential number counts (solid red curve) compared to the C13 model prediction (dot–dashed black curve) and to observed data at 850, 870, and 1100 μm. The data points are taken from Simpson et al. (2019, 2020), Shim et al. (2020), Garratt et al. (2023), Stach et al. (2018), and Scott et al. (2010, 2012).

(ii) Three equatorial fields each of approximate area of 56 deg² coinciding with the previously surveyed fields in the Galactic and Mass Assembly spectroscopic survey.

(iii) Two fields having a total area of 250 deg², close to the South Galactic Pole.

Because of the higher sensitivity to high-z sources of SPIRE compared to PACS, hereafter we will only focus on the SPIRE wavelengths. The 1σ noise for source detection at 250, 350, and 500 μm is 7.4 , 9.4, and 10.2 mJy, respectively.

Table 2. Euclid filters along with their central wavelengths and bandwidths.

Filter	Central wavelength (μm)	Bandwidth (μm)
I_E/VIS/*Euclid*	0.731	0.378
Y_E/NISP/*Euclid*	1.081	0.262
J_E/NISP/*Euclid*	1.342	0.398
H_E/NISP/*Euclid*	1.772	0.498
u/Rubin	0.358	0.068
g/Rubin	0.478	0.133
r/Rubin	0.623	0.127
i/Rubin	0.755	0.119
z/Rubin	0.871	0.101
y/Rubin	1.013	0.162

4.1.2 Euclid wide area survey

The EWS (Euclid Collaboration 2022b) will map 15 000 deg² of the sky from the Sun-Earth Lagrange point L2, using all four of its filters – visible (I_E) and NIR (Y_E, J_E, and H_E). The survey will have a 5σ depth of 26.2, 24.3, 24.5, and 24.4 mag for I_E, Y_E, J_E, and H_E, respectively.

4.1.3 Rubin's Legacy Survey of Space and Time

The *Vera C. Rubin Obervatory* (previously known as Large Synoptic Survey Telescope; Ivezić et al. 2019), which is under construction in Cerro Pachón in Northern Chile, is planned to conduct a survey named *Legacy Survey of Space and Time (LSST)* covering approximately 18 000 deg² of the sky in the optical in six bands – u, g, r, i, z, and y. The survey will have a 5σ depth of 23.8, 24.5, 24.03, 23.41, 22.74, and 22.96 mag, respectively.

A list of the *Euclid* and *Rubin* filters along with their central wavelengths and bandwidths is provided in Table 2.

4.2 Simulation set-up

A simulated source is said to be detected by *Herschel* if it has a flux density above the 5σ limit of 37 mJy at 250 μm, and to be detected in one of the *Euclid* bands if it has a flux density higher than the 3σ limit at that band, i.e. 26.75, 24.85, 25.05, 24.95 mag for I_E, Y_E, J_E, and H_E, respectively. Moreover, we say that an *Euclid* source is detected by *LSST* if it has a flux density $> 3\sigma$ in at least four out of the six bands.

To speed up the calculation, we simulated objects with a minimum virialization halo mass of $\log(M_{\rm vir}/M_\odot) = 12.5$, which corresponds to an infrared luminosity of $\log(L_{\rm IR}/L_\odot) = 12$, as illustrated in Fig. 12. In fact, simulated sources with $S_{250} = 37$ mJy have a minimum halo mass of $10^{12.5}\,M_\odot$ and a corresponding minimum IR luminosity of $10^{12}\,L_\odot$.

The simulated catalogue comprised 458 994 galaxies prior to the application of any selection criterion. Approximately 33 per cent of them (i.e. 151 720 objects) have $S_{250} > 37$ mJy. All of these *Herschel* detected galaxies are also detected in all the four *Euclid* bands above 3σ. Among these sources, 140 524 have a 350 μm flux density above 3σ and 30 337 are detected above 3σ at both 350 and 500 μm. Approximately, 75 per cent of the *Euclid* detected *Herschel* galaxies are also detected by *Rubin/LSST* above 3σ in at least four out of the six bands.
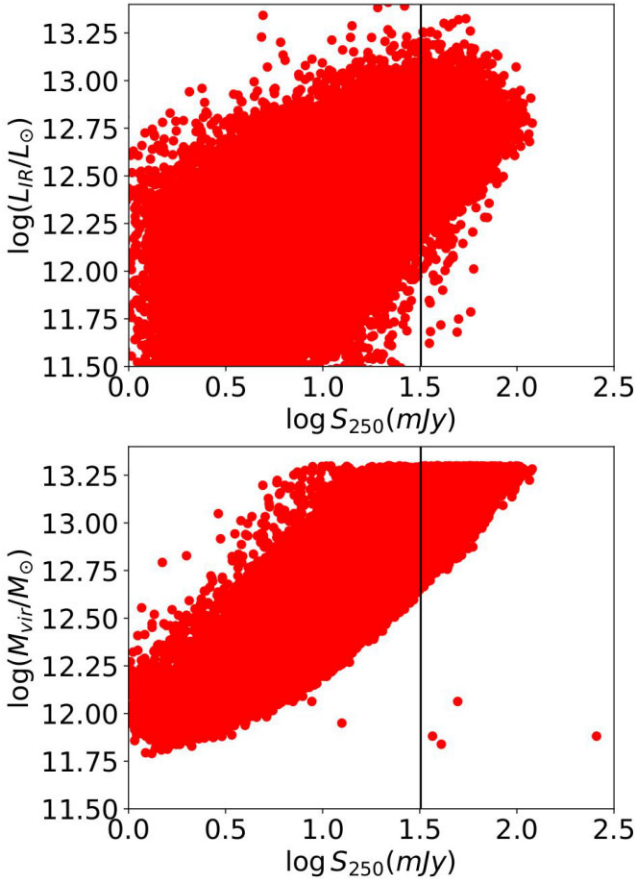
Figure 12. Distribution of infrared luminosity, L_{IR}, and virial mass, M_{vir}, of the simulated proto-spheroidal galaxies as a function of their 250 μm flux density, S_{250}. The vertical black line denotes the *Herschel*-ATLAS 5σ flux density limit of 37 mJy at 250 μm.

4.3 Photometric redshifts

The first thing needed to extract the intrinsic properties of galaxies is an estimate of the redshift. When spectra are not available, an estimate of the redshift can be obtained from the photometry. Because we are dealing with *Herschel* sources detected by *Euclid*, we can use either the optical/near-IR photometry from *Euclid* or the sub-mm *Herschel* photometry, or a combination of both to estimate the redshift.

4.3.1 Redshifts from the optical/near-IR Euclid photometry

To estimate the photo-z from the *Euclid* data, we used the photometric redshift estimation code EAZY (Brammer, van Dokkum & Coppi 2008), which is designed to work with UV/optical/near-IR data. EAZY fits the observed SED by linearly combining a number of template SEDs. It thus calculates the best-fitting coefficients of the linear combination together with the photometric redshift. We used the EAZY version 'v1.3' that comprises 9 (5+1+1+1+1) templates, shown in Fig. 13. The original 'v1.0' version of EAZY comprises five 'principal component' templates obtained from Grazian et al. (2006) using the Blanton & Roweis (2007) algorithm. One dusty starburst template ($t = 50$ Myr, $A_v = 2.75$) was added to account for the extremely dusty galaxies. The new adopted version has these six templates modified to include line emissions. It also includes one dust template, a template taken from Erb et al. (2010) and the evolved SSP by Maraston (2005) to account for massive old galaxies
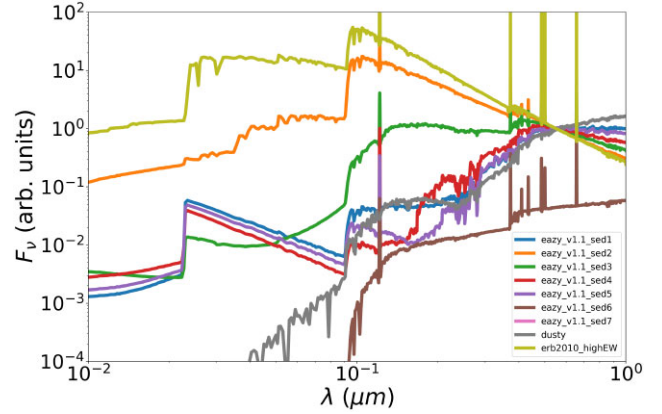

Figure 13. The set of EAZY SED templates used in this work.

at $z < 1$. Dust absorption from the intergalactic medium (IGM), in accordance with Madau (1995), is also included by EAZY during its calculations. We selected a redshift range from $z = 1$ to 8 in steps of $\Delta z = 0.01$.

4.3.2 Redshifts from the sub-mm photometry

Another redshift estimate was derived from the *Herschel*/SPIRE photometry at 250, 350, and 500 μm. We fitted it with the empirical SED of Pearson et al. (2013), which is the sum of two modified blackbody spectra:

$$F_\nu = N[\nu^\beta B_\nu(T_1) + \xi \nu^\beta B_\nu(T_2)] \quad (32)$$

where F_ν is the flux density at the rest-frame frequency ν, N is the normalization factor, B_ν is the Planck function, β is the dust emissivity index, T_1 is the hot dust temperature, T_2 is the cold dust temperature, and ξ is the ratio of the mass of the cold dust to that of the hot dust. We adopt $T_1 = 46.9$ K, $T_2 = 23.9$ K, $\beta = 2$, and $\xi = 30.1$ (Pearson et al. 2013; Negrello et al. 2017). The template SED was redshifted between $0 \leq z \leq 8$ in steps of 0.01 and a χ^2 minimization was performed to estimate the photometric redshift.

4.4 Estimating the galaxies properties

Once the photo-z is determined, we estimate the physical properties of the simulated galaxies using the SED fitting code **CIGALE** (**C**ode **I**nvestigating **GAL**axy **E**mission; Burgarella, Buat & Iglesias-Páramo 2005; Noll et al. 2009; Boquien et al. 2019). **CIGALE** is a code written in *Python* which is used to efficiently model the multi-wavelength (UV to radio) spectrum of galaxies and to estimate the physical properties of the galaxies. It is based on the principle of 'energy balance', where the stellar energy absorbed by dust in the UV-optical is entirely re-emitted in the far-IR/sub-mm/mm.

CIGALE creates model SEDs based on the modules and parameter values given as input by the user. Once an SFH is chosen, CIGALE couples it with the stellar population synthesis models to create the stellar spectrum of the galaxies. Several SFH functions are available depending on the physics of star formation and on the morphology of the galaxies. In selecting the modules of CIGALE, we took help from Traina et al. (2023), who studied 1620 sub-mm galaxies from $0.5 \leq z \leq 6$ by using the data from the largest available Atacama Large Millimeter Array (ALMA)[1] survey in the COSMOS field

[1] https://www.almaobservatory.org/en/home/

Table 3. Parameter values given as input to CIGALE for SED fitting.

SFH: sfhdelayed – delayed SFH with optional exponential burst	
e-folding time of the main stellar population (Myr)	1000, 2000, 5000, 6000, 7000
e-folding time of the late starburst population (Myr)	5000, 10 000
Mass fraction of the late starburst population	0.0, 0.1, 0.15, 0.30
Age of the main stellar population (Myr)	500, 1000, 2000, 3000, 5000, 6000, 7000, 8000, 9000, 10 000
Age of the late starburst (Myr)	1, 10, 30, 50, 70, 100, 150, 300
SSP: bc03 (Bruzual & Charlot 2003)	
Initial mass function (IMF)	Chabrier (Chabrier 2003)
Metallicity	0.02 ($Z_\odot$)
Dust attenuation: dustatt_modified_CF00 (Charlot & Fall 2000)	
V-band attenuation in ISM (A_V^{ISM})	0.3, 1.7, 2.8
μ	0.44, 0.5
Power-law slope of attenuation in the ISM	-0.7
Power-law slope of attenuation in the BCs	-0.7
Dust emission: dl2014 (Draine et al. 2014)	
PAH mass fraction (q_{PAH})	2.5, 3.9, 4.58, 5.26
Minimum radiation field (U_{min}) (Habing)	5, 10, 15, 25, 40
Dust emission power-law slope (α)	2
Fraction illuminated from U_{min} to U_{max} (γ)	0.02
AGN: fritz2006 (Fritz et al. 2006)	
Ratio of the maximum to minimum radii of the dusty torus (r_{ratio})	60, 150
Equatorial optical depth at 9.7 μm (τ)	0.6, 1.0, 3.0, 6.0,10.0
Radial dust distribution within the torus (β)	-0.5, 0.0
Angular dust distribution within the torus (γ)	0, 6
Full opening angle of the dusty torus (Opening angle)	100, 140
Angle between the equatorial axis and line of sight (Ψ)	0.001, 30.1, 60.1, 89.990
AGN fraction (f_{AGN})	0.0, 0.1, 0.15, 0.25, 0.5

(Scoville et al. 2007), known as A^3 COSMOS (Liu et al. 2019). We chose a delayed SFH with an additional late burst of star formation. For modelling the stellar spectrum, we adopted the Chabrier IMF (Chabrier 2003) and chose the SSP models by BC03. The metallicity was set at the solar value ($Z_\odot = 0.02$). The dust attenuation was modelled using the attenuation law by Charlot & Fall (2000). To model the dust emission, we adopted the templates from Draine et al. (2014). Finally, the AGN component was modelled using the templates from Fritz et al. (2006).

The chosen modules and parameter values, which are summarized in Table 3, produced a total of 39 029 760 SED templates. CIGALE considered all these templates to perform SED fitting by implementing a χ^2 minimization process (Noll et al. 2009).

5 RESULTS

This section presents the results obtained for the estimation of photometric redshift, stellar mass and SFR for the simulated catalogue for *Euclid*, using the SED fitting technique.

5.1 Derived photometric redshift

In Fig. 14, we show the photometric redshift estimated with EAZY versus the input redshift for the detected galaxies. There is a good level of agreement between the photometric redshift and the input redshifts for $z_{input} \leq 3$. The fraction of outliers, i.e. of objects with $|\Delta z|/(1 + z_{input}) > 0.15$ where $|\Delta z| = |z_{input} - z_{phot}^{EAZY}|$, as commonly defined in the literature (e.g. Laigle et al. 2016), is $f_{outlier} = 0.128$. Most of the disastrous estimates of the photometric redshift are associated with galaxies at $z_{input} \gtrsim 3.5$, which are wrongly assigned a redshift in the range $z_{phot}^{EAZY} \sim 1$–1.5. Also,
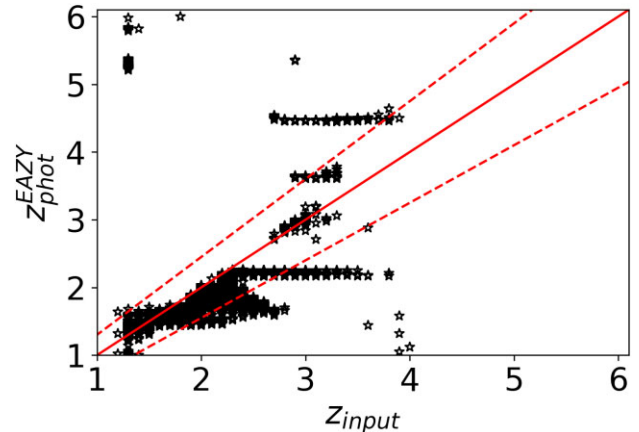

Figure 14. Derived photometric redshift using EAZY versus the input redshift of the galaxies detected by *Herschel* which are also detected by all the four bands of *Euclid* at $\geq 3\sigma$. The solid red line denotes $z_{input} = z_{phot}^{EAZY}$, while the dashed red lines define the region where $|\Delta z| \leq 0.15(1 + z_{input})$.

there are some galaxies with input redshift $z_{input} \sim 1$–2 which are estimated to be at much higher redshift, i.e. $z_{EAZY} \sim 5$–6. Fig. 15 shows some of the best-fitting SEDs produced by EAZY along with the corresponding estimated redshift. In order to achieve a good photometric redshift, EAZY mainly exploits the 4000 Å break. At low redshifts (i.e. $z_{input} \lesssim 3$), such feature falls within the wavelength range sampled by *Euclid*, thus leading to photometric redshifts that agree with the input values. However, at higher redshifts, the 4000 Å break is missed by *Euclid* thus causing catastrophic redshift estimates. Moreover, due to the limited wavelength coverage, the
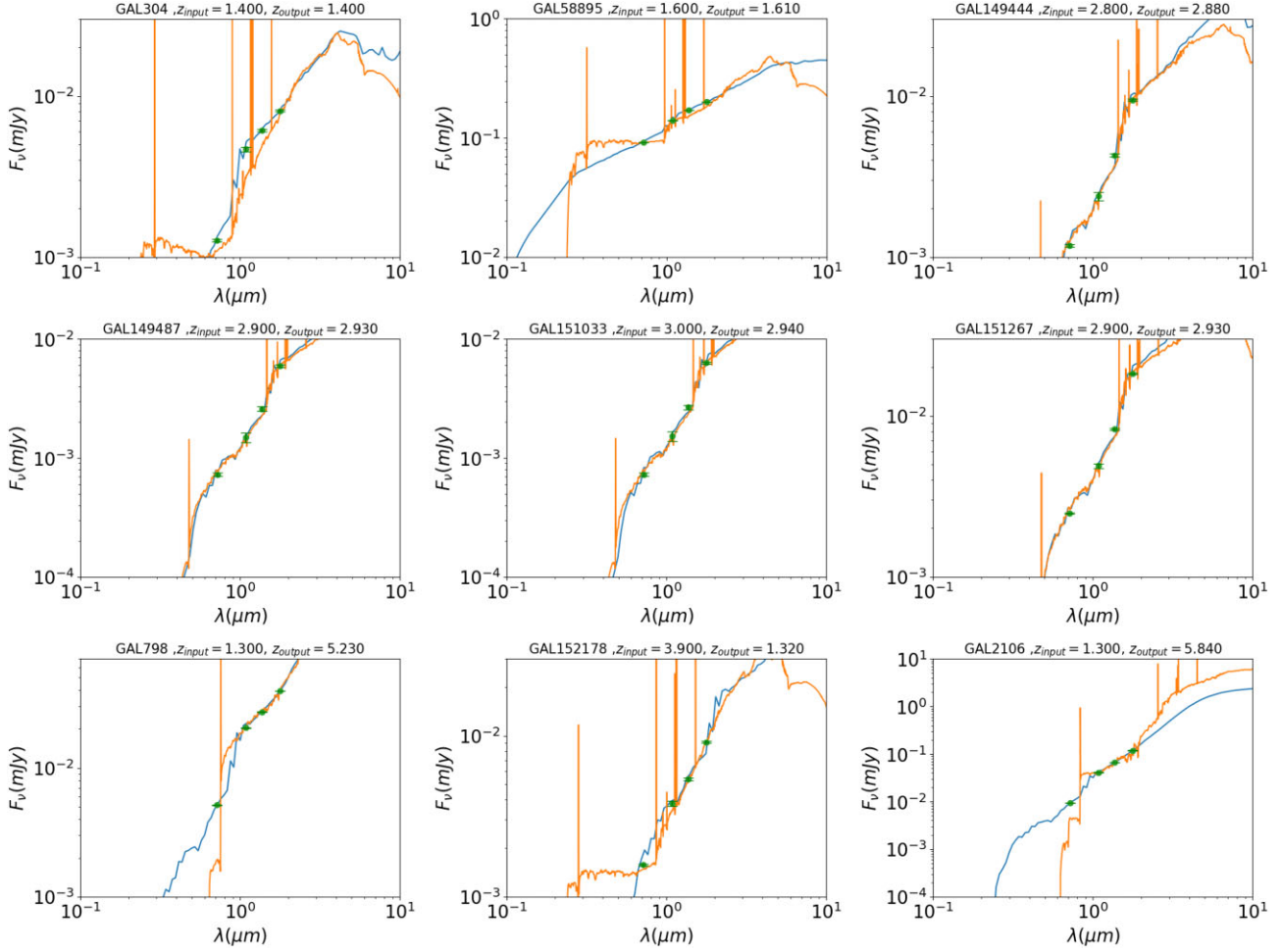
Figure 15. Examples of simulated near-IR SEDs fitted with EAZY, illustrating the cases of both good and catastrophic redshift estimates. The blue curve denotes the simulated SED while the orange curve is the best-fitting SED produced by EAZY. The green points denote the 'observed' flux density values at the four *Euclid* bands.

4000 Å break is wrongly identified as the Ly α break at 912 Å, which also leads to a wrong redshift estimate. The addition of filters bluewards of *Euclid* can help to identify the Ly α break, thus leading to a more accurate redshift estimate. This is observed when the *Euclid* photometry is complemented with the *LSST* data. Fig. 16 shows a plot of the input redshift versus the photometric redshift obtained by EAZY from the *Euclid* +*LSST* photometry. Out of the 113 897 *Euclid* sources that are also detected by *LSST* in at least four bands, 112 831 (approximately 99 per cent) have $|\Delta z|/(1 + z_{\mathrm{input}}) \leq 0.15$. The outlier fraction f_{outlier} is thus down to just 1 per cent.

Fig. 17 shows the contour plot of the derived photometric redshift of the same galaxies using the *Herschel*/SPIRE photometry alone. The fraction of outlier is $f_{\mathrm{outlier}} = 0.37$, more than twice the one obtained from the *Euclid* photometry. Moreover, there are 47 618 sources for which the Pearson template gives $|\Delta z|/(1 + z_{\mathrm{input}}) > 0.15$ while EAZY produces $|\Delta z|/(1 + z_{\mathrm{input}}) \leq 0.15$. At the same time, there are 10 623 sources with $|\Delta z|/(1 + z_{\mathrm{input}}) \leq 0.15$ when using the Pearson template while EAZY produces $|\Delta z|/(1 + z_{\mathrm{input}}) > 0.15$. Therefore, in general, the near-IR data seem to be more successful at estimating the redshift of the simulated galaxies compared to the *Herschel* photometry. However, it is worth pointing out that for the objects which are also detected above 3σ at both 350 and 500 μm (only 20 per cent of the sample), the sub-mm photometry
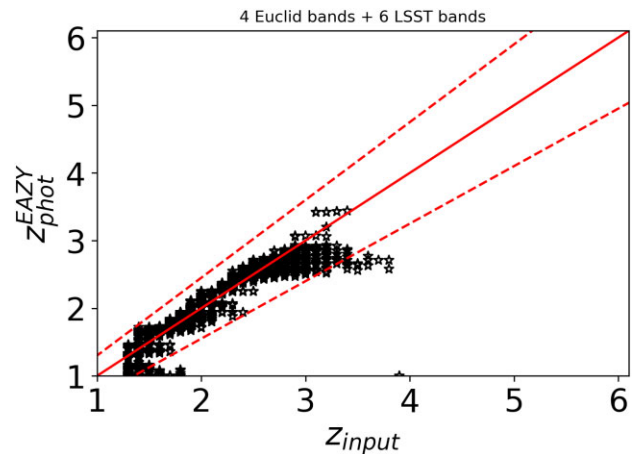

Figure 16. Derived photometric redshift using EAZY versus the input redshift of the galaxies detected by *Herschel* which are also detected by all the four bands of *Euclid* and any four out of six bands of *LSST* at $\geq 3\sigma$. The solid red line denotes $z_{\mathrm{input}} = z_{\mathrm{phot}}^{\mathrm{EAZY}}$, while the dashed red lines define the region where $|\Delta z| \leq 0.15(1 + z_{\mathrm{input}})$.
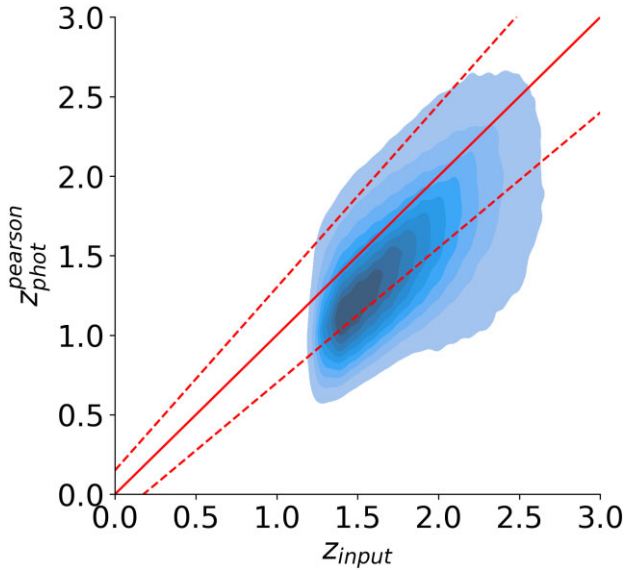
5.2 Derived physical properties

In order to extract the physical properties of the simulated galaxies, such as the stellar mass, $M_\star$, and the star formation rate, SFR, we fitted the SED of the simulated *Herschel* galaxies detected with *Euclid* by using CIGALE and adopting the photometric redshift produced by EAZY. In the following sections, we discuss the results obtained using the *Euclid* photometry alone and the *Euclid* plus *Herschel*/SPIRE photometry combined. We also analyse the improvement of the above when including the *LSST* photometry.

5.2.1 Stellar mass

Fig. 19 shows the comparison between the derived stellar masses, $M_\star^{\mathrm{CIGALE}}$, and the input ones, $M_\star^{\mathrm{input}}$. Also shown in the figure is the histogram of the difference between the logarithm of the stellar masses defined as $\Delta \log(M_\star) \equiv \log(M_\star^{\mathrm{input}}) - \log(M_\star^{\mathrm{CIGALE}})$. Overall, the retrieved stellar mass is in good agreement with the input values. The median value of the stellar mass is found to be $\log(M_\star/\mathrm{M}_\odot) = 11.29 \pm 0.181$. The 1σ dispersion in $\Delta \log(M_\star)$ are 0.16 and 0.19 from the *Euclid* and *Euclid*+*Herschel* photometry, respectively. These dispersions are mainly a reflection of the uncertainties in the derived photometric redshifts. As expected, because the evolved stellar populations – which contribute most of the stellar mass in a galaxy – dominate the rest-frame optical/near-IR part of the spectrum, the stellar mass estimates produced by CIGALE do not improve when the SED fitting also includes the sub-mm *Herschel* data. However, the addition of the *LSST* data does improve the stellar mass estimate, by reducing the 1σ dispersion in $\Delta \log(M_\star)$ to 0.14.

We observe that, on average, CIGALE overestimates the stellar mass, the effect being particularly relevant for the higher masses. This offset can be attributed to the choice of the SFH and its parametrization. Here, we choose a multicomponent SFH, which, in general, has a tendency to yield a higher mass-to-light ratio; hence, higher stellar mass estimates, as pointed out by Michałowski et al. (2012). The discrepancy in stellar mass estimates could also be explained by the difference in the value of the dust attenuation law in birth clouds used in CIGALE versus the one adopted here, i.e. -0.7 versus -1.3, respectively, and therefore to the extension of the extinction curve to the near-IR, where it directly impacts the stellar mass.

5.2.2 SFR

The comparison between the input values of the SFR and the estimated ones is shown in Fig. 20. In this case, the SED fit has been performed using both the *Euclid* and the *Herschel* photometry. In fact, for dusty objects, the NIR wavelengths are not good tracers of the star formation (Pforr, Maraston & Tonini 2012, 2013; Collaboration et al. 2023) and the use of sub-mm data becomes crucial to reliably infer the SFR. The difference in logarithm between the input and the output values of the SFR has a dispersion of 0.26. As the LSST filters cover the UV/optical wavelengths, adding the LSST data to the NIR and FIR photometry can help to trace the SFR more accurately as can be seen in Fig. 20 (right panels). In fact, when using the *Euclid*+*Herschel*+*LSST* photometry, the 1σ dispersion in $\Delta \log(\dot{M}_\star)$ drops to 0.18. However, we observe a systematic overestimation of the SFR by CIGALE when adding the *LSST* photometric data. As for the stellar mass estimates, the discrepancy in the SFR measurements can also be explained in terms of differences in the adopted values of the slope of the dust attenuation law between CIGALE and our formalism. Overall, estimating the SFR seems to
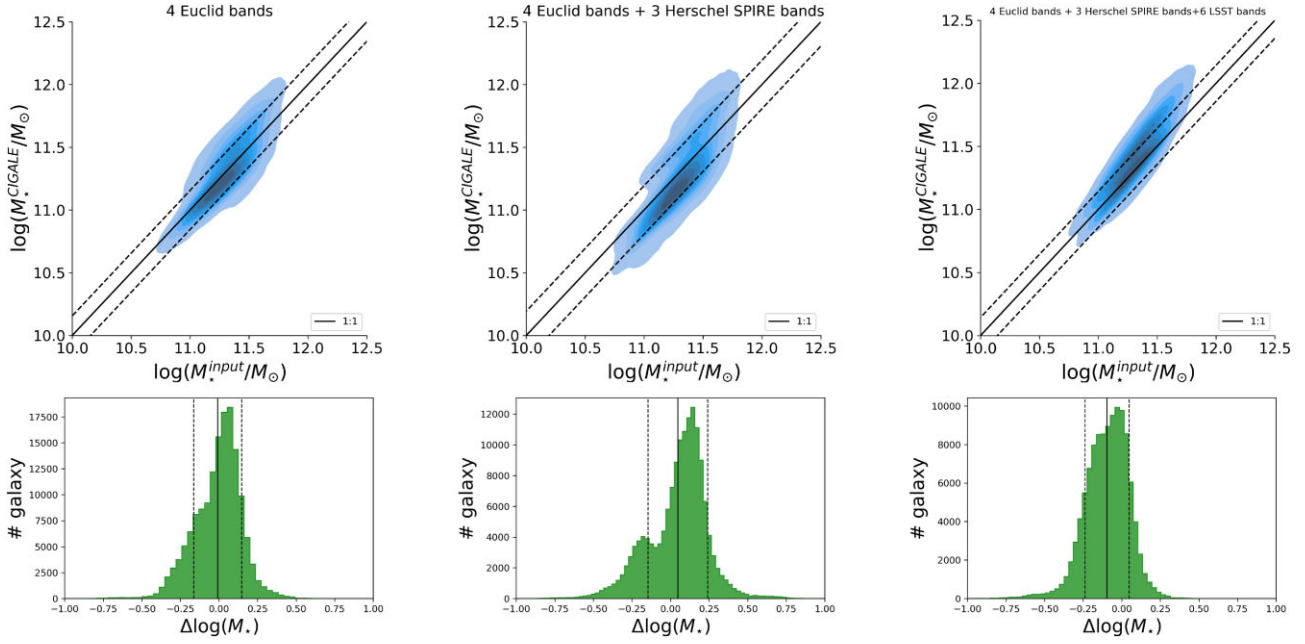

Figure 17. Contour plot of the derived photometric redshifts using Pearson templates on *Herschel* photometry versus the input redshift of the *Euclid* detected *Herschel* galaxies. The red solid line denotes $z_{\mathrm{input}} = z_{\mathrm{phot}}^{\mathrm{pearson}}$, while the red dashed lines bound the region where $|\Delta z| \leq 0.15(1 + z_{\mathrm{input}})$.
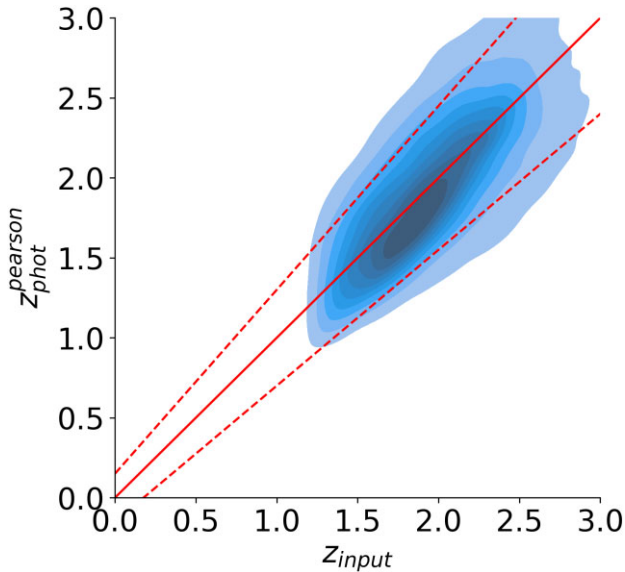
Figure 18. Contour plot of the derived photometric redshifts using Pearson templates on *Herschel* photometry versus the input redshift of the *Herschel*/*Euclid* galaxies detected above 3σ at both 350 and 500 μm. The solid red line denotes $z_{\mathrm{input}} = z_{\mathrm{phot}}^{\mathrm{pearson}}$, while the dashed red lines bound the region where $|\Delta z| \leq 0.15(1 + z_{\mathrm{input}})$.

provides a better estimate of the photometric redshift as illustrated in Fig. 18. Therefore, the inaccuracy of the sub-mm photometry in estimating the redshift for the bulk of the objects can be attributed to the absence of robust flux density measurements at 350 and 500 μm.

In the following analysis, we will rely on the photometric redshift estimated by EAZY from the *Euclid* photometry.

Figure 19. Comparison between the input and the estimated values of the stellar masses obtained with CIGALE from *Euclid* photometry alone (left panels), *Euclid + Herschel* photometry (middle panels) and *Euclid+Herschel+ LSST* photometry (right panels). The top panels show the contour of the estimated versus the true stellar masses for both cases. The black straight line marks the 1:1 relation. The black dashed line marks the boundaries of the 1σ dispersion around the mean. The bottom panels show the distribution of differences between the logarithms of the input stellar mass and of the estimated stellar mass. The black solid line denotes the mean value while the dashed lines mark the 1σ dispersion around the mean.

be more challenging than inferring the stellar mass. This may be due to the more sensitive dependence of the latter on the SFH compared to the former. In fact, the SFR is an *instantaneous* quantity, while $M_\star$ is an integrated quantity. To test this, we compared the values of the SFR averaged over a defined time-scale. We introduce two quantities, $\dot{M}_{\star,10}$ and $\dot{M}_{\star,100}$, which represent the SFR averaged over the 10 Myr and the 100 Myr prior to the time of observation, respectively, i.e.

$$\dot{M}_{\star,10} = \frac{1}{10\,\mathrm{Myr}} \int_{t_{obs}-10\,\mathrm{Myr}}^{t_{obs}} \dot{M}_\star(t)\mathrm{d}t \qquad (33)$$

and

$$\dot{M}_{\star,100} = \frac{1}{100\,\mathrm{Myr}} \int_{t_{obs}-100\,\mathrm{Myr}}^{t_{obs}} \dot{M}_\star(t)\mathrm{d}t. \qquad (34)$$

Fig. 21 shows the distribution of the input values versus the measured values of $\dot{M}_{\star,10}$ and $\dot{M}_{\star,100}$ along with the difference in logarithm between the true and the estimated values of the SFR from the *Euclid+Herschel* and the *Euclid+Herschel+LSST* photometry. Both $\Delta \log(\dot{M}_{\star,10})$ and $\log(\Delta \dot{M}_{\star,100})$ have a 1σ dispersion of 0.26 when estimated from the *Euclid+Herschel* photometry. Upon adding the LSST data to the former, the 1σ dispersion reduces to 0.18. Even though $\dot{M}_{\star,10}$ and $\dot{M}_{\star,100}$ are integrated properties, we get similar dispersion to what is derived for the instantaneous SFR. Therefore, we can conclude that the use of an instantaneous SFR cannot explain the larger dispersion in the recovered values of SFR compared to what was observed for the stellar mass. However, the galaxies we are dealing with are dust-obscured; therefore, the accuracy in the estimate of the SFR is also dependent on the availability of data at far-IR/sub-mm/mm wavelengths. Only 20 per cent of the simulated DSFGs detected by Herschel above 5σ at 250 μm are also detected above 3σ at 350 μm and at 500 μm. Indeed, if we restrict the calculation of the SFR to that subsample we find a

1σ dispersion of 0.23 and 0.14 for *Euclid+Herschel* photometry and *Euclid+Herschel+LSST* photometry, respectively, which clearly shows an improvement in the estimation of the SFR. The median SFR of the sources in our sample is $\log(\dot{M}_\star/\mathrm{M}_\odot\mathrm{yr}^{-1}) = 2.77 \pm 0.162$.

5.2.3 SFR–$M_\star$ relation

Studies of large galaxy samples have shown that most of the star-forming galaxies in the Universe follow an SFR–$M_\star$ correlation, called the galaxy main sequence (MS), while a fraction of galaxies lie above that correlation and are regarded as *starburst* (SB) galaxies (Speagle et al. 2014; Mancuso et al. 2016; Popesso et al. 2022). Fig. 22 shows the SFR–$M_\star$ plot of our simulated galaxies for several bins of redshifts, along with the redshift-dependent galaxy–MS relation derived by Speagle et al. (2014). Our sample comprises both MS and starburst galaxies. The bulk of the galaxies, approximately 57 per cent, with $z \sim 1$–2, are above the MS relation and they are in their starburst stage, forming stars at an average rate of ~ 300–$3000\,\mathrm{M}_\odot$ per year. There are very few cases (11) where the estimated SFR is $> 10^4\,\mathrm{M}_\odot\mathrm{yr}^{-1}$. This is due to the wrong redshift estimation by EAZY, leading to the wrong estimation of SFR by CIGALE. Around 15 per cent of the galaxies at $z \sim 2$–3 are above the MS relation and in their starburst stage. On the other hand, at high redshifts $z \sim 3.5$–5.0, we find that the bulk of the sources are MS galaxies.

5.2.4 Dust luminosity and dust mass

The comparison between the measured values of the dust luminosity and the input values is shown in Fig. 23. We considered again the results obtained by using the *Euclid+Herschel* photometry (left-hand panels) and the *Euclid+Herschel+LSST* photometry (right-
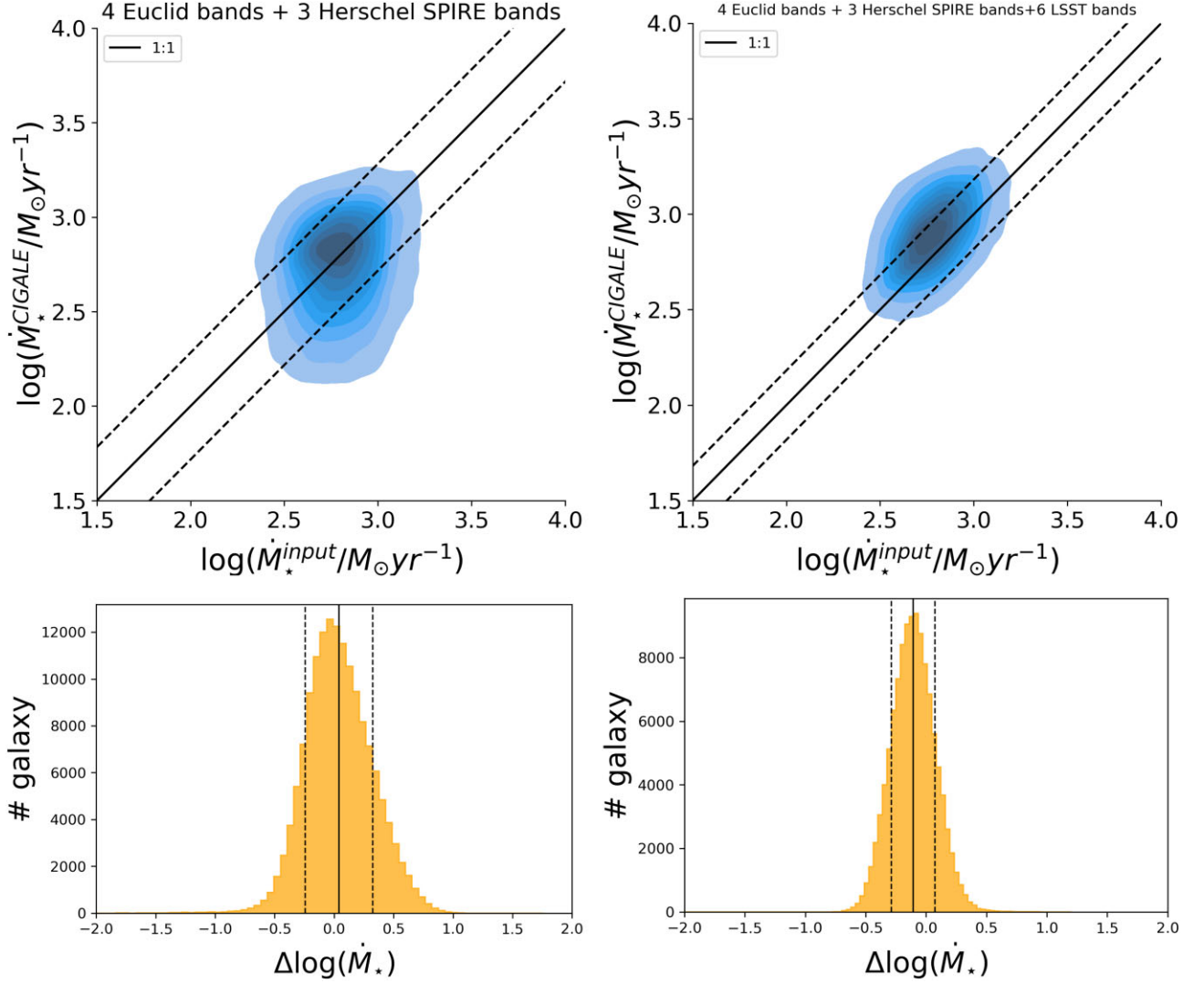
Figure 20. Same as Fig. 19, but for the SFR, with the SED fit being performed on *Euclid+Herschel*/SPIRE photometry (left) and *Euclid+Herschel*/SPIRE+*LSST* photometry (right).

hand panels). The histogram of the differences between the input values and the measured values is shown in the bottom panels. We found that, when fitting the *Euclid* and the *Herschel* data together, we could estimate the dust luminosity (L_{dust}) accurately, with a 1σ dispersion in $\Delta \log(L_{dust})$ of about 0.22. On adding the *LSST* photometry to the fit, the constraint on L_{dust} becomes tighter and the 1σ dispersion in $\Delta \log(L_{dust})$ reduces to 0.17. In fact, by tracing the UV/optical wavelengths, the LSST data help to better constrain the extinction of starlight by dust, thus leading to a more precise estimate of L_{dust}. It is worth noticing, though, that in this case CIGALE systematically overestimates the dust luminosity by a factor of ~1.07. This is due to a higher standard value of the slope of the dust-attenuation law adopted by CIGALE. Overall, with a median dust luminosity of $\log(L_{dust}/L_\odot) = 12.8 \pm 0.12$, these galaxies sample the bright tail of the population of ultra-luminous infrared galaxies (ULIRGs).

In Fig. 24, we show the results for the dust mass estimates. The dust mass has a median value of $\log(M_{dust}/M_\odot) = 8.9 \pm 0.10$. This is consistent with the values of M_{dust} of other ULIRGs reported by Dudzevičiūtė et al. (2020) and Pantoni et al. (2021), and shows that these DSFGs have higher dust content than local ($z < 1$) star-

forming galaxies selected with Herschel (e.g. R14). The dust mass is affected by larger uncertainties compared to the dust luminosity and the use of the LSST photometry does not help to reduce that uncertainty. We find, indeed, that the 1σ dispersion in $\Delta \log(M_{dust})$ is 0.35 and 0.32 for the *Euclid+Herschel* photometry and the *Euclid+Herschel*+LSST photometry, respectively. We also observe that CIGALE systematically overestimates M_{dust} by a factor of 1.2. Interestingly, Liao et al. (2024) fit the SED of a sample of 18 bright ($S_{870} = 12.4 - 19.4\,\text{mJy}$) 870 μm-selected DSFGs from AS2COSMOS (Simpson et al. 2020) using both CIGALE and MAGPHYS, the latter being based on the same Cunha et al. (2008) formalism we have used in this work. They found that CIGALE produced values of M_{dust} higher by a factor 1.5 than those estimated with MAGPHYS. An overestimation factor of 1.3 is also quoted by Birkin et al. (2021) when comparing dust mass estimates by CIGALE and MAGPHYS for a sample 61 sources selected from ALMA-identified 870 μm-selected DSFGs from AS2COSMOS, AS2UDS (Stach et al. 2019), and ALESS (Hodge et al. 2013) surveys. These discrepancies in M_{dust} can be mainly explained by the difference in the standard values of κ_0 adopted by the two SED-fitting codes (Liao et al. 2024).
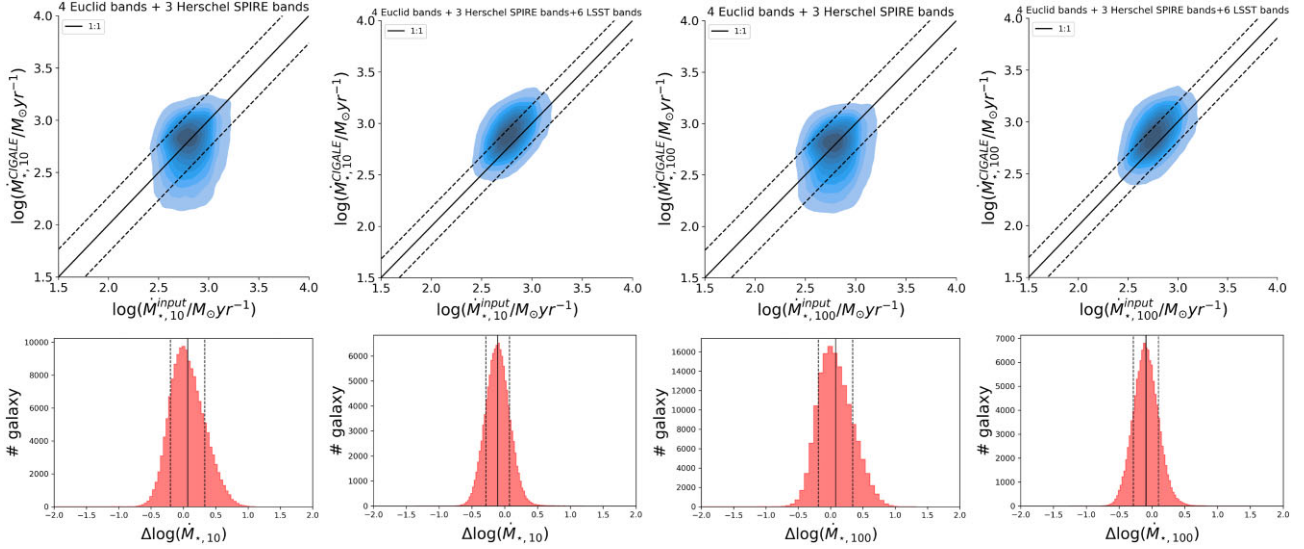
Figure 21. Same as Fig. 19, but for the SFR averaged over 10 Myr (first two top and bottom figures) and 100 Myr (last two top and bottom figures) respectively, with the SED fit being performed on *Euclid+Herschel*/SPIRE photometry and *Euclid+Herschel*/SPIRE+*LSST* photometry.

6 SUMMARY AND CONCLUSIONS

We have presented predictions for how well the recently launched space telescope *Euclid* can study the $z \gtrsim 1.5$ DSFGs detected by the *Herschel*-ATLAS survey. To this end, we have built a simulated catalogue of galaxies exploiting the C13 model for the evolution of protospheroidal galaxies. We combined the output of model equations with the SED formalism of D08, which is based on the principle of energy balance in the UV-optical and IR.

By setting the values of the main parameters of the model to those derived by R14 from the study of a sample of DSFGs we obtained a good agreement between the predicted infrared luminosity function of DSFGs and the measured one in the redshift range of $z = 2 - 4$. However, compared with the prescriptions of R14, we fine-tuned the values of both the dust emissivity index and the temperature of the warm dust in the BCs and of the cold dust in the ISM to reproduce the observed number counts of at 850 μm, 870 μm, and 1.1 mm.

After these preliminary checks, we simulated a catalogue of 458 994 galaxies with $L_{IR} \gtrsim 10^{12} L_\odot$ within a survey area of 100 deg^2, which is of the same order of magnitude of the total area surveyed by H-ATLAS, and applied a 5σ detection condition at 250 μm for *Herschel* and a 3σ detection condition for all the 4 *Euclid* bands. All the 151 720 sources detected by *Herschel* at 250 μm are also detected in all the *Euclid* bands. We used EAZY on the *Euclid* photometry to estimate the photometric redshifts of the above sources. We then exploited CIGALE to extract the physical properties of the galaxies from the SED fitting, assuming the previously derived photometric redshift.

Our main findings are as follows:

(i) For 87 per cent of the sample, the *Euclid* photometry alone provides a good redshift estimate up to $z \sim 3$, with a discrepancy ≤ 15 per cent from the true value of $1 + z$. For input redshifts $z \gtrsim 3$, we observe some catastrophic outliers, where the estimated redshift is significantly lower than the input one, i.e. $z_{phot} \sim 1$–1.5. This discrepancy is attributed to the fact that the 4000 Å break, which is used as the main feature to constrain the redshift, is not sampled by *Euclid* above $z \sim 3$. Also, there are some galaxies at $z \sim 1 - 2$ which

are estimated to be at $z \sim 4 - 6$, due to the Balmer break (4000 Å) being mistaken as the Lyman break (912 Å). Adding the shorter wavelength data from LSST does help to break some degeneracies. In fact, the outlier fraction reduces to just 1 per cent for the 75 per cent of the *Euclid* detected galaxies that are also detected above 3σ in at least four LSST bands.

We also investigated the use of the sub-mm *Herschel* photometry to derive the photometric redshift by fitting those data with the empirical template of Pearson et al. (2013). However, we found a higher fraction of outliers compared to what was obtained with the *Euclid* photometry, i.e. $f_{outlier} = 0.35$ compared to $f_{outlier} = 0.13$. This is mainly due to the lack of sufficient sub-mm data for most of the sources in the sample. In fact, only 20 per cent of the *Herschel* detected galaxies (i.e. those above 5σ at 250 μm) are also detected at both 350 μm and 500 μm above 3σ. For this sub-sample, the outlier fraction is reduced to 13 per cent.

In summary, for the vast majority of the simulated DSFGs, we cannot gain any significant improvement on the photo-z from the sub-mm measurements and that is why we have adopted the photometric redshift estimates based on the *Euclid* data alone.

(ii) The distribution of the stellar masses, as recovered from the *Euclid* and *Euclid+Herschel* photometry, has a dispersion of 0.16 and 0.19 dex, respectively, around the true value. As expected, adding the *Herschel* photometry to the SED fit does not significantly improve the results, as the stellar mass is sensitive to the rest-frame optical/near-IR portion of the SED. The median value of the stellar mass is found to be $\log(M_\star/M_\odot) = 11.29 \pm 0.181$, while the range of values of $\log(M_\star/M_\odot)$ is $10.5 - 12.9$.

(iii) The recovered SFR shows a wider distribution around the true value compared to the stellar mass, with a dispersion 0.26 dex. Overall, the sources in our sample have a median SFR of $\log(\dot{M}_\star/M_\odot yr^{-1}) = 2.77 \pm 0.162$, while the probed range of values of $\log(\dot{M}_\star/M_\odot yr^{-1})$ is $1.51 - 3.99$. When the SFRs are averaged over 10 and 100 Myr, respectively, we obtain a similar 1σ dispersion of 0.26 in both cases. The median value is $\log(\dot{M}_{\star,10}/M_\odot yr^{-1}) = 2.77 \pm 0.16$ and $\log(\dot{M}_{\star,100}/M_\odot yr^{-1}) = 2.74 \pm 0.16$.
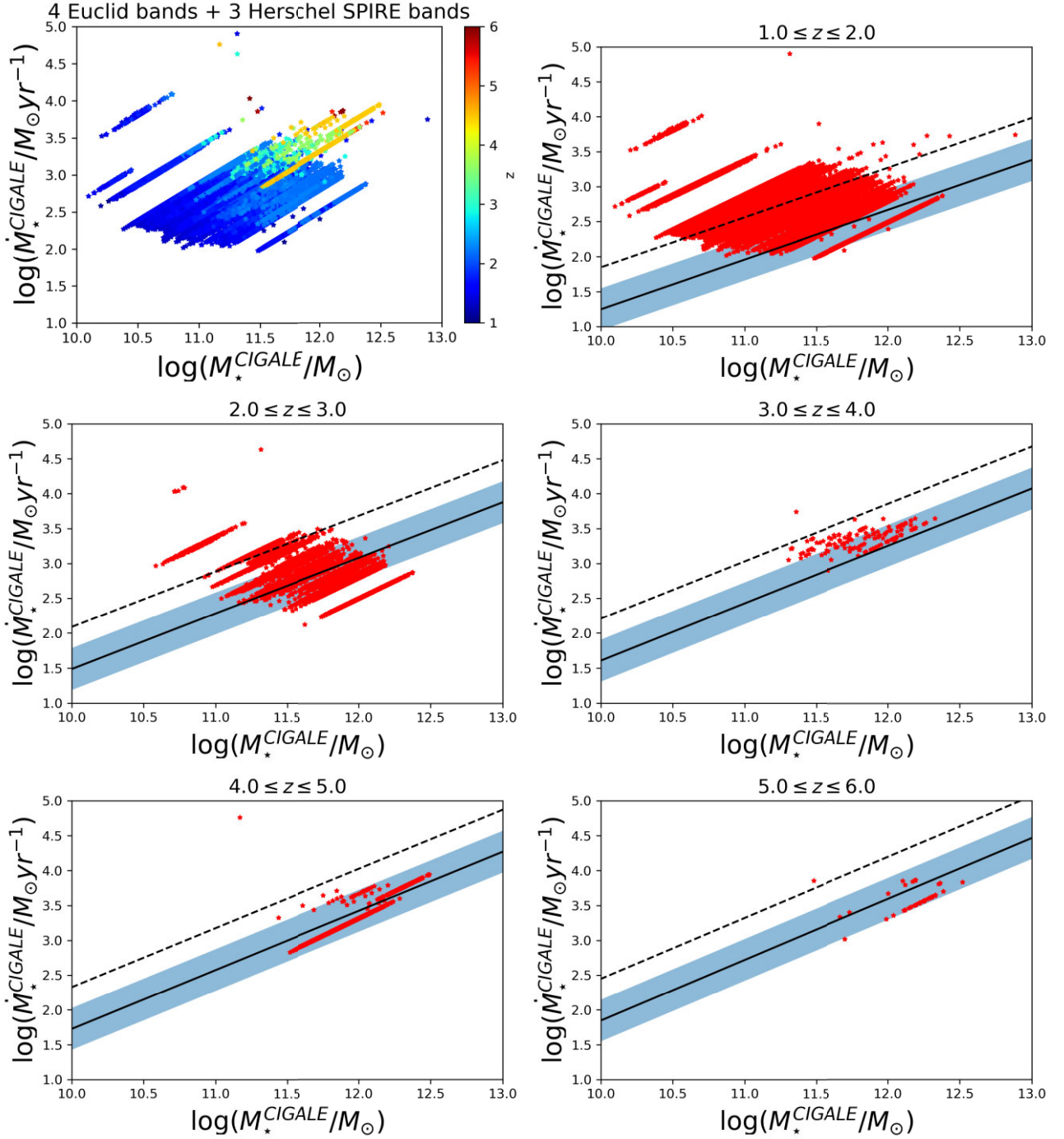
Figure 22. Star formation versus stellar mass for the simulated sample of galaxies detected by *Herschel* and *Euclid*. The top left panel shows the correlation for the entire catalogue of galaxies. The points are colour-coded according to their photometric redshift, z_{phot}^{EAZY}. The remaining plots show the SFR versus $M_\star$ for different redshift bins as follows: $1 \leq z \leq 2$ (top right), $2 < z \leq 3$ (middle left), $3 < z \leq 4$ (middle right), $4 < z \leq 5$ (bottom left), and $5 \leq z \leq 6$ (bottom right). The black line in each of the plots represents the galaxy–MS relation from Speagle et al. (2014) calculated at the central redshift of the interval. The blue-filled region denotes the 0.3 dex scatter around that relation. The dashed black line shows the MS relation scaled up by a factor of 4, the conventional boundary between MS and starburst galaxies.

(iv) From the SFR$-$M$_\odot$ relation, we found that our simulated catalogue comprised both MS galaxies and starburst galaxies. Approximately, 55 per cent of the galaxies in our sample lie above the MS relation and are starbursting, with a median SFR of $\log(\text{SFR}/\text{M}_\odot\,\text{yr}^{-1}) = 2.87$.

(v) The median dust luminosity obtained for our simulated DSFG sample is $\log(L_{\text{dust}}/L_\odot) = 12.8 \pm 0.12$ and the dust mass has a median value of $\log(M_{\text{dust}}/\text{M}_\odot) = 8.9 \pm 0.10$. These DSFGs have a dust luminosity in the range $\log(L_{\text{dust}}/L_\odot) \sim 12.2\text{--}13.3$ and dust mass in the range $\log(M_{\text{dust}}/\text{M}_\odot) \sim 7.5 - 9.9$.
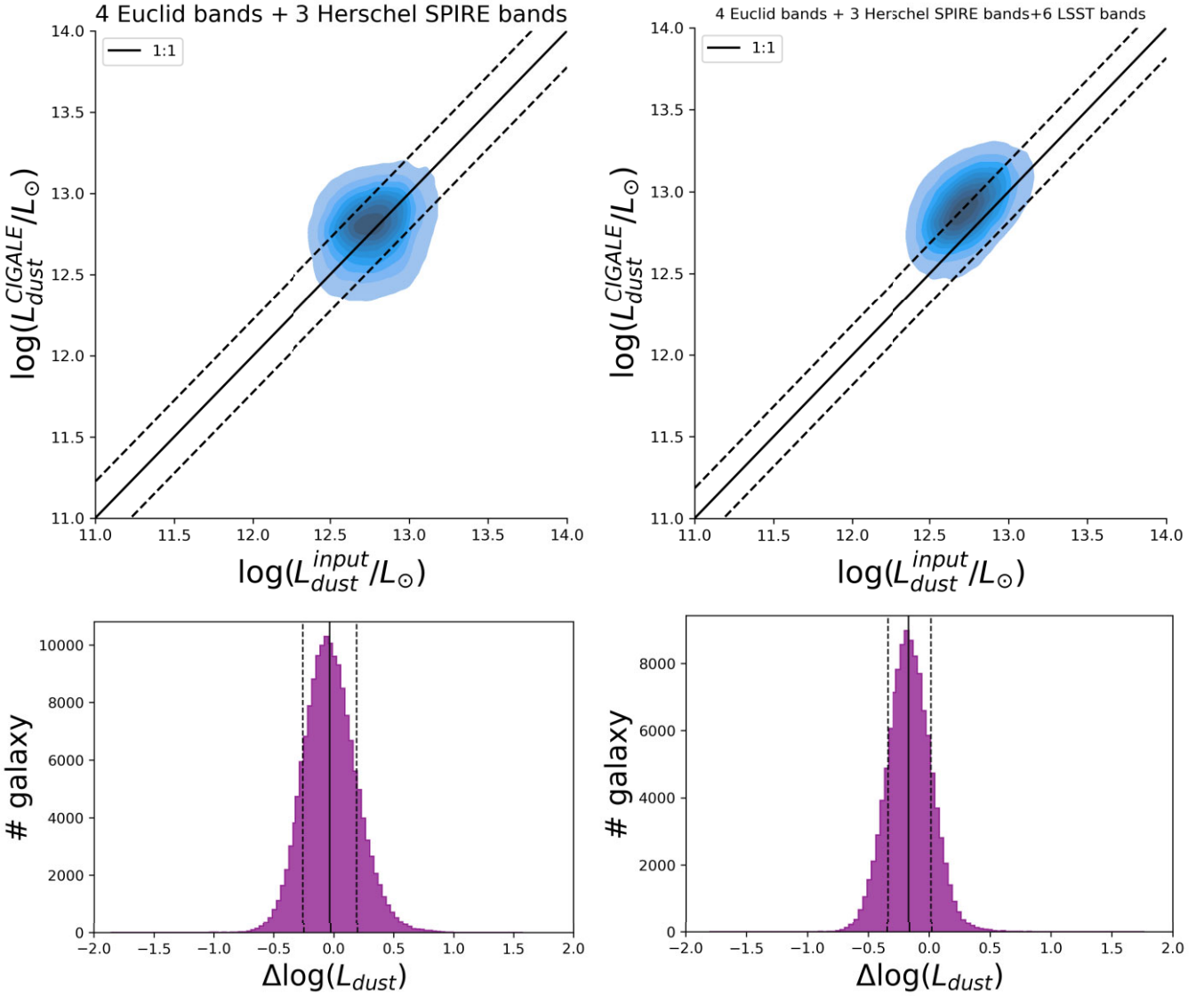
Figure 23. Same as Fig. 19, but for the dust luminosity (L_{dust}), with the SED fit being performed on *Euclid+Herschel*/SPIRE photometry (left-hand panels) and *Euclid+Herschel*/SPIRE+*LSST* photometry (right-hand panels).

(vi) The photometric redshift can be improved by complementing the *Euclid* data with the *LSST* observations. From the combined *Euclid+LSST* photometry, EAZY provides an accurate redshift, i.e. $\Delta z/(1 + z) \leq 0.15$, for 99 per cent of the 87 per cent *Euclid* objects that are also detected in at least four of the six *LSST* bands. For this subsample of galaxies that have both *Euclid* and LSST photometry,

the dispersion on the recovered stellar mass and SFR reduces to 0.14 and 0.18 dex, respectively. By covering the UV/optical wavelengths using *LSST* filters, we could constrain the extinction of starlight by dust in the UV, which reduced the dispersion in the estimation of L_{dust}. However, the dispersion in dust mass shows no significant improvement with the inclusion of *LSST* photometry.
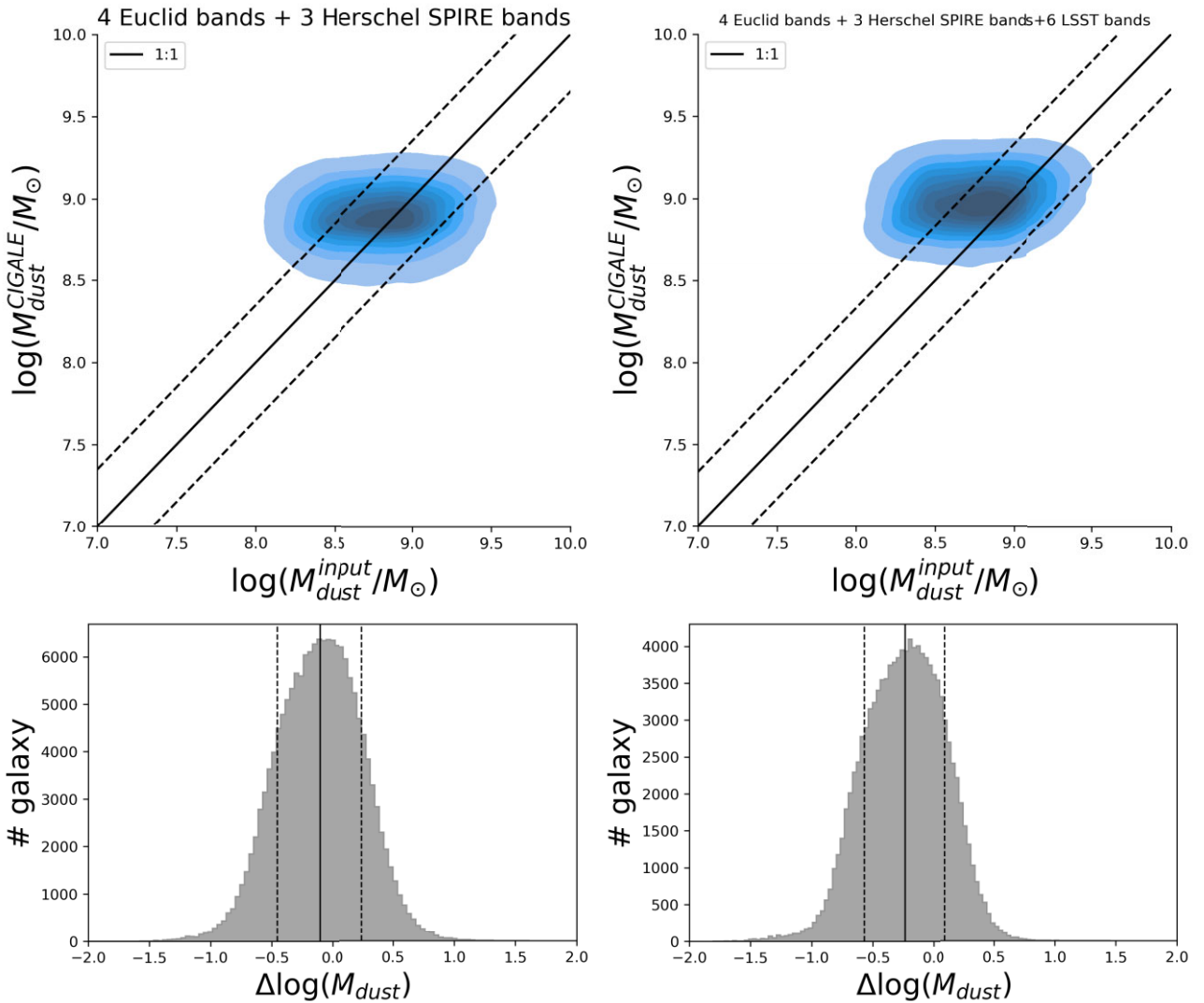
Figure 24. Same as Fig. 19, but for the dust mass (M_{dust}), with the SED fit being performed on *Euclid+Herschel*/SPIRE photometry (left-hand panels) and *Euclid+Herschel*/SPIRE+*LSST* photometry (right-hand panels).

[2] https://www.astropy.org/
[3] https://scipy.org/index.html
[4] https://numpy.org/
[5] https://joblib.readthedocs.io/en/stable/
[6] https://pypi.org/project/colossus/
[7] https://seaborn.pydata.org/index.html
[8] https://matplotlib.org/

DATA AVAILABILITY

Data in this paper will be made available upon request.

REFERENCES

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Birkin J. E. et al., 2021, MNRAS, 501, 3926
Blain A. W., Smail I., Ivison R. J., Kneib J. P., Frayer D. T., 2002, Phys. Rep., 369, 111
Blanton M. R., Roweis S., 2007, AJ, 133, 734
Bothwell M. S. et al., 2013, MNRAS, 429, 3047
Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, A&A, 622, A103
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Buat V., Boquien M., Małek K., Corre D., Salas H., Roehlly Y., Shirley R., Efstathiou A., 2018, A&A, 619, A135
Burgarella D., Buat V., Iglesias-Páramo J., 2005, MNRAS, 360, 1413
Burgarella D. et al., 2013, A&A, 554, A70
Buzzoni A., 1993, A&A, 275, 433
Béthermin M. et al., 2012, ApJ, 757, L23

Cai Z.-Y. et al., 2013, ApJ, 768, 21(C13)
Cai Z.-Y., Zotti G. D., Bonato M., 2020, ApJ, 891, 74
Cai Z.-Y., Negrello M., Zotti G. D., 2022, ApJ, 932, 13
Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499
Casey C. M. et al., 2012, ApJ, 761, 139
Castillo M. F. et al., 2023, ApJ, 945, 128
Cesarsky D., Lequeux J., Abergel A., Perault M., Palazzi E., Madden S., Tran D., 1996, A&A, 315, L309
Chabrier G., 2003, PASP, 115, 763
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Euclid Collaboration 2023, MNRAS, 520, 3529
Cropper M. S. et al., 2018, in MacEwen H. A., Lystrup M., Fazio G. G., Batalha N., Tong E. C., Siegler N., eds, Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave. SPIE, Austin, p. 78
da Cunha E., Charlot S., 2011, Astrophysics Source Code Library, record ascl:1106.010
da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595
da Cunha E. et al., 2015, ApJ, 806, 110
Diemer B., 2018, ApJS, 239, 35
Dokkum P. G. v. et al., 2015, ApJ, 813, 23
Draine B. T., Lee H. M., 1984, ApJ, 285, 89
Draine B. T., Li A., 2007, ApJ, 657, 810
Draine B. T. et al., 2014, ApJ, 780, 172
Dudzevičiūtė U. et al., 2020, MNRAS, 494, 3828
Eales S. et al., 2010, PASP, 122, 499
Erb D. K., Pettini M., Shapley A. E., Steidel C. C., Law D. R., Reddy N. A., 2010, ApJ, 719, 1168
Euclid Collaboration, 2022a, A&A, 662, A92
Euclid Collaboration, 2022b, A&A, 662, A112
Franceschini A., Rodighiero G., Vaccari M., Berta S., Marchetti L., Mainetti G., 2010, A&A, 517, A74
Fritz J., Franceschini A., Hatziminaoglou E., 2006, MNRAS, 366, 767
Fu H. et al., 2013, Nature, 498, 338
Garratt T. K. et al., 2023, MNRAS, 520, 3669
Granato G. L., Zotti G. D., Silva L., Bressan A., Danese L., 2004, ApJ, 600, 580
Grazian A. et al., 2006, A&A, 449, 951
Gullberg B. et al., 2019, MNRAS, 490, 4956
Hatsukade B. et al., 2018, PASJ, 70, 105
Hildebrand R. H., 1983, QJRAS, 24, 267
Hodge J. A. et al., 2013, ApJ, 768, 91
Holland W. S. et al., 2013, MNRAS, 430, 2513
Hu W., Sugiyama N., 1995, ApJ, 444, 489
Hwang H. S. et al., 2010, MNRAS, 409, 75
Hönig S. F. et al., 2013, ApJ, 771, 87
Ivezić et al., 2019, ApJ, 873, 111
Ivison R. J. et al., 2013, ApJ, 772, 137
Jimenez R., Flynn C., Kotoneva E., 1998, MNRAS, 299, 515
Kennicutt R. C., 1998, ARA&A, 36, 189
Laigle C. et al., 2016, ApJS, 224, 24
Lapi A., Shankar F., Mao J., Granato G. L., Silva L., Zotti G. D., Danese L., 2006, ApJ, 650, 42
Lapi A. et al., 2011, ApJ, 742, 24
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Liao C.-L. et al., 2024, ApJ, 961, 226
Lilly S. J., Fèvre O. L., Hammer F., Crampton D., 1996, ApJ, 460, L1
Lim C.-F. et al., 2020, ApJ, 889, 80
Liu D. et al., 2019, ApJ, 887, 235
MacKenzie T. P. et al., 2017, MNRAS, 468, 4006

Madau P., 1995, ApJ, 441, 18
Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388
Madden S. C., Galliano F., Jones A. P., Sauvage M., 2006, A&A, 446, 877
Magnelli B., Elbaz D., Chary R. R., Dickinson M., Borgne D. L., Frayer D. T., Willmer C. N. A., 2011, A&A, 528, A35
Magnelli B. et al., 2012, A&A, 539, A155
Magnelli B. et al., 2013, A&A, 553, A132
Maraston C., 2005, MNRAS, 362, 799
Mancuso C., Lapi A., Shi J., Cai Z.-Y., Gonzalez-Nuevo J., Béthermin M., Danese L., 2016, ApJ, 833, 152
Matteucci F., Spitoni E., Recchi S., Valiante R., 2009, A&A, 501, 531
Małek K. et al., 2018, A&A, 620, A50
Michałowski M. J., Dunlop J. S., Cirasuolo M., Hjorth J., Hayward C. C., Watson D., 2012, A&A, 541, A85
Michałowski M. J. et al., 2017, MNRAS, 469, 492
Negrello M. et al., 2010, Science, 330, 800
Negrello M. et al., 2017, MNRAS, 465, 3558
Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, A&A, 507, 1793
Oliver S. J. et al., 2012, MNRAS, 424, 1614
Pantoni L. et al., 2021, MNRAS, 504, 928
Pearson E. A. et al., 2013, MNRAS, 435, 2753
Pforr J., Maraston C., Tonini C., 2012, MNRAS, 422, 3285
Pforr J., Maraston C., Tonini C., 2013, MNRAS, 435, 1389
Pilbratt G. L. et al., 2010, A&A, 518, L1
Planck Collaboration VI, 2020, A&A, 641, A6
Popesso P. et al., 2011, A&A, 532, A145
Popesso P. et al., 2022, MNRAS, 519, 1526
Riechers D. A. et al., 2013, Nature, 496, 329
Rosas-Guevara Y., Bower R. G., Schaye J., McAlpine S., Dalla Vecchia C., Frenk C. S., Schaller M., Theuns T., 2016, MNRAS, 462, 190
Rosito M. S. et al., 2021, A&A, 652, A44
Rowan-Robinson M. et al., 2016, MNRAS, 461, 1100
Rowlands K. et al., 2014, MNRAS, 441, 1017
Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2008, MNRAS, 389, 1137
Scott K. S. et al., 2010, MNRAS, 405, 2260
Scott K. S. et al., 2012, MNRAS, 423, 575
Scoville N. et al., 2007, ApJS, 172, 1
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Shim H. et al., 2020, MNRAS, 498, 5065
Simpson J. M. et al., 2017, ApJ, 839, 58
Simpson J. M. et al., 2019, ApJ, 880, 43
Simpson J. M. et al., 2020, MNRAS, 495, 3409
Siringo G. et al., 2009, A&A, 497, 945
Smail I., Ivison R. J., Blain A. W., 1997, ApJ, 490, L5
Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15
Stach S. M. et al., 2018, ApJ, 860, 161
Stach S. M. et al., 2019, MNRAS, 487, 4648
Swinbank A. M. et al., 2014, MNRAS, 438, 1267
Traina A. et al., 2024, A&A, 681, A118
Verstraete L. et al., 2001, A&A, 372, 981
Vieira J. D. et al., 2010, ApJ, 719, 763
Wang L., Pearson W. J., Cowley W., Trayford J. W., Béthermin M., Gruppioni C., Hurley P., Michałowski M. J., 2019, A&A, 624, A98
Wilson G. W. et al., 2008, MNRAS, 386, 807
Yang Q. et al., 2020, ApJ, 900, 58

APPENDIX A: EQUATIONS GOVERNING THE EVOLUTION OF PROTO-SPHEROIDAL GALAXIES

A1 The model

Massive proto-spheroidal galaxies are high-z sub-mm galaxies with high SFRs (Lapi et al. 2006, 2011). The model depicts two ways of DM halo formation as follows:

(i) fast collapse of the halo bulk triggering star formation, and

(ii) slow growth of the halo outskirts (accretion flows) having little effect on the inner portion.

The star formation takes place due to the cooling of gas in the region of radius

$$\approx 70 \left(\frac{M_{\text{vir}}}{10^{13} M_\odot} \right)^{\frac{1}{3}} \left[\frac{(1 + z_{\text{vir}})}{3} \right]^{-1} \text{kpc}, \qquad (A1)$$

where M_{vir} and z_{vir} are the halo mass and the virialization redshift, respectively. This region is approximately 30 per cent of the halo virial radius, and comprises ≈ 40 per cent of the total mass (C13). Both the star formation and the growth of BH at the centre are controlled by feedback mechanisms from the supernovae (SNe) and the AGN. AGN feedback is responsible for sweeping out the residual gas in the most massive proto-spheroids, while the SFR is mostly controlled by SN feedback in the less massive ones.

Consider a galactic halo mass M_{vir}. The initial gas mass fraction in the halo has the cosmological value $f_b = M_{\text{gas}}/M_{\text{vir}} = 0.165$. The gas is heated to the virial temperature at the virialization redshift z_{vir}. There are three gas phases:

(i) Hot dense medium infalling and/or cooling towards the centre, M_{inf};

(ii) Cold gas condensing into stars, M_{cold};

(iii) Gas stored in a reservoir around the SMBH, M_{res} (low angular momentum gas).

In addition, there are two condensed phases:

(i) Total mass of stars, $M_\star$;

(ii) Black hole mass, $M_\bullet$.

The equations governing the evolution of the gas phases are

$$\dot{M}_{\text{inf}} = -\dot{M}_{\text{cond}} - \dot{M}_{\text{inf}}^{\text{QSO}}, \qquad (A2)$$

$$\dot{M}_{\text{cold}} = \dot{M}_{\text{cond}} - [1 - R(t)]\dot{M}_\star - \dot{M}_{\text{cold}}^{SN} - \dot{M}_{\text{cold}}^{\text{QSO}}, \qquad (A3)$$

$$\dot{M}_{\text{res}} = \dot{M}_{\text{inflow}} - \dot{M}_{\text{BH}}. \qquad (A4)$$

The cooling of hot gas and its flow to the centre take place at a rate

$$\dot{M}_{\text{cond}} = \frac{\dot{M}_{\text{inf}}}{t_{\text{cond}}}, \qquad (A5)$$

where

$$t_{\text{cond}} \simeq 8 \times 10^8 \left(\frac{1 + z}{4} \right)^{-1.5} \left(\frac{M_{\text{vir}}}{10^{12} M_\odot} \right)^{0.2} \text{yr}, \qquad (A6)$$

and we take the initial value of M_{inf} as $M_{\text{inf}}^0 = f_b M_{\text{vir}}$. The SFR is (for the Chabrier IMF, $R = 0.54$)

$$\dot{M}_\star = \frac{\dot{M}_{\text{cold}}}{t_\star} = \frac{\dot{M}_{\text{cold}}}{t_{\text{cond}}/s} = \frac{5 \dot{M}_{\text{cold}}}{t_{\text{cond}}}. \qquad (A7)$$

The rate of gas loss due to SN feedback is

$$\dot{M}_{\text{cold}}^{SN} = \beta_{\text{SN}} \dot{M}_\star = \frac{N_{\text{SN}} \epsilon_{\text{SN}} E_{\text{SN}}}{E_{\text{bind}}} \dot{M}_\star \simeq 0.6 \left(\frac{N_{\text{SN}}}{8 \times 10^{-3}/M_\odot} \right) \left(\frac{\epsilon_{\text{SN}}}{0.05} \right)$$

$$\left(\frac{E_{\text{SN}}}{10^{51} \text{erg}} \right) \times \left(\frac{M_{\text{vir}}}{10^{12} M_\odot} \right)^{-\frac{2}{3}} \left(\frac{1 + z}{4} \right)^{-1} \dot{M}_\star. \qquad (A8)$$

The following values are adopted:

(i) Number of SNe /$M_\odot$, $N_{\text{SN}} \simeq 1.4 \times 10^{-2}/M_\odot$;

(ii) Fraction of energy released used to heat the gas, $\epsilon_{\text{SN}} = 0.05$;

(iii) Kinetic energy/SN, $E_{\text{SN}} \simeq 10^{51}$ erg;

(iv) Binding energy of the halo per unit stellar mass;

$$E_{\text{bind}} \simeq 3.2 \times 10^{14} \left(\frac{M_{\text{vir}}}{10^{12} M_\odot} \right)^{\frac{2}{3}} \left(\frac{1 + z}{4} \right) \quad \text{cm}^2 \text{ s}^{-2}.$$

The IR luminosity related to the dust-enveloped SF is given by

$$L_{\star, IR}(t) = k_{\star, IR} \times 10^{43} \left(\frac{\dot{M}_\star}{M_\odot \text{yr}^{-1}} \right) \text{erg s}^{-1}, \qquad (A9)$$

where $k_{\star, IR}$ is SED dependent.

The rate of cold gas inflow into the reservoir around the SMBH is

$$\dot{M}_{\text{inflow}} \simeq \frac{L_\star}{c^2} (1 - e^{-\tau_{RD}}) \simeq \alpha_{RD} \times 10^{-3} \dot{M}_\star (1 - e^{-\tau_{RD}}), \qquad (A10)$$

with

$$\tau_{RD} = \tau_{RD}^0 \left(\frac{Z_{\text{cold}}(t)}{Z_\odot} \right) \left(\frac{M_{\text{cold}}}{10^{12} M_\odot} \right) \left(\frac{M_{\text{vir}}}{10^{13} M_\odot} \right)^{-\frac{2}{3}}. \qquad (A11)$$

The cold gas metallicity evolves according to the equation

$$Z_{\text{cold}}(t) = Z_{\text{inf}}^0 + \frac{s}{s\gamma - 1} \varepsilon_Z(t) - \frac{st/t_{\text{cond}}}{e^{(s\gamma - 1)t/t_{\text{cond}}} - 1}$$
$$\left\{ \varepsilon_Z(t) + B_Z \sum_{i=2}^{\infty} \frac{1}{i.i!} \left[(s\gamma - 1) \frac{\min(t, t_Z)}{t_{\text{cond}}} \right]^{i-1} \right\}, \qquad (A12)$$

where $\gamma = 1 - R - \beta_{\text{SN}}$.

The mass fraction of metals formed in stars is given by

$$\varepsilon_Z(t) = A_z + B_z \left[\frac{\min(t, t_{\text{saturated}})}{t_z} \right]$$
$$= 0.03 + 0.02 \left[\frac{\min(t, t_{\text{saturated}})}{t_z} \right], \qquad (A13)$$

where $t_z = 20$ Myr, $t_{\text{saturated}} = 40$ Myr (for a Chabrier IMF) and $Z_\odot = 0.02$.

The above equation accounts for the fact that after star formation begins, contributions to the metal yield take place

(i) initially from massive stars with masses $\geq 20 M_\odot$ and lifetime ≤ 20 Myr;

(ii) afterwards from moderate/intermediate stars having masses $\sim 9 - 20 M_\odot$ and lifetime $\sim 20 - 40$ Myr;

(iii) finally, shifting to low mass stars ($M \leq 9 M_\odot$) with lifetime of approximately ≥ 40 Myr.

The black hole accretion rate is given by

$$\dot{M}_{\text{BH}} = \min(\dot{M}_{\text{BH}}^{\text{visc}}, \lambda_{\text{Edd}} \dot{M}_{\text{Edd}}). \qquad (A14)$$

The accretion rate into the reservoir around the SMBH of gas which has dissipated its angular momentum is

$$\dot{M}_{\text{BH}}^{\text{visc}} = \frac{M_{\text{res}}}{\tau_{\text{visc}}} = \kappa_{\text{acc}} 5 \times 10^3 \left(\frac{V_{\text{vir}}}{500 \text{ kms}^{-1}} \right)^3 \left(\frac{M_{\text{res}}}{M_\bullet} \right)^{\frac{3}{2}} \left(1 + \frac{M_\bullet}{M_{\text{res}}} \right)^{\frac{1}{2}}, \qquad (A15)$$

where

(i) $\kappa_{\mathrm{acc}} = 10^{-2}$;

(ii) $V_{\mathrm{vir}} = GM_{\mathrm{vir}}^{2/3}[(4/3)\pi \triangle_{\mathrm{vir}}(z)\bar{\rho}_m(z)]^{-1/3}$;

(iii) $\dot{M}_{\mathrm{Edd}} = M_{\bullet}/(\epsilon t_{\mathrm{Edd}})$; for a mass-to-light conversion efficiency $\epsilon = 0.1$ the Salpeter time is $\epsilon t_{\mathrm{Edd}} = 4.5 \times 10^7$ yr;

(iv) $\lambda_{\mathrm{Edd}} \simeq 0.1(z - 1.5)^2 + 1.0$ for $z \gtrsim 1.5$ ($\lambda_{\mathrm{max,Edd}} = 4$).

The BH grows according to the equation

$$\dot{M}_{\bullet}(t) = (1 - \epsilon)\dot{M}_{\mathrm{BH}}. \tag{A16}$$

The AGN luminosity is given by

$$L_{\bullet} = \epsilon \dot{M}_{\mathrm{BH}}c^2 = 5.67 \times 10^{45} \left(\frac{\epsilon}{0.1}\right)\left(\frac{\dot{M}_{\mathrm{BH}}}{\mathrm{M}_{\odot}\mathrm{yr}^{-1}}\right) \text{ erg s}^{-1}. \tag{A17}$$

The feedback from the AGN is given by

$$\dot{M}_{\mathrm{inf,cold}}^{\mathrm{QSO}} = \dot{M}_{\mathrm{wind}} \frac{M_{\mathrm{inf,cold}}}{M_{\mathrm{inf}} + M_{\mathrm{cold}}} = \frac{L_{\mathrm{ISM}}^{\mathrm{QSO}}}{E_{\mathrm{bind}}} \frac{M_{\mathrm{inf,cold}}}{M_{\mathrm{inf}} + M_{\mathrm{cold}}}, \tag{A18}$$

where

$$L_{\mathrm{ISM}}^{\mathrm{QSO}} = 2 \times 10^{44} \epsilon_{\mathrm{QSO}} \left(\frac{\dot{M}_{\mathrm{BH}}}{\mathrm{M}_{\odot}\mathrm{yr}^{-1}}\right)^{3/2} \text{ erg s}^{-1}. \tag{A19}$$

A2 Halo mass function and halo formation rate function

In cosmology, the halo mass function specifies the mass distribution of DM haloes. Technically speaking, it gives the number density of haloes per unit interval of mass. We adopt the analytical expression given by Sheth & Tormen (1999),

$$N_{\mathrm{ST}}(M_{\mathrm{vir}}, z)dM_{\mathrm{vir}} = \frac{\bar{\rho}_{m,0}}{M_{\mathrm{vir}}^2} f_{\mathrm{ST}}(\nu) \frac{d \ln \nu}{d \ln M_{\mathrm{vir}}} dM_{\mathrm{vir}}, \tag{A20}$$

where $\bar{\rho}_{m,0} = \Omega_{m,0}\rho_{c,0}$ is the present-day mean matter density of the Universe, $\nu \equiv (\delta_c(z)/\sigma(M_{\mathrm{vir}}))^2$ with $\delta_c(z) = \delta_0(z)(D(0)/D(z))$

where

$$\delta_0(z) \approx 1.6865[1 + 0.0123 \log \Omega_m(z)] \tag{A21}$$

and

$$D(z) = \frac{5\Omega_m(z)/[2(1 + z)]}{\frac{1}{70} + \frac{209}{140}\Omega_m(z) - \frac{1}{140}\Omega_m^2(z) + \Omega_m^{4/7}(z)} \tag{A22}$$

from Carroll, Press & Turner (1992).

The mass variance is given by Bardeen et al. (1986) and Hu & Sugiyama (1995)

$$\sigma(M_{\mathrm{vir}}) = \frac{0.8}{0.84}[14.110393 - 1.1605397x - 0.0022104939x^2$$
$$+0.0013317476x^3 - 2.1049631 \times 10^{-6}x^4], \tag{A23}$$

where $x = \log(M_{\mathrm{vir}}/\mathrm{M}_{\odot})$ and $10^6 < M_{\mathrm{vir}}/\mathrm{M}_{\odot} < 10^{16}$. Furthermore,

$$f_{\mathrm{ST}}(\nu) = A[1 + (a\nu)^{-p}]\left(\frac{a\nu}{2}\right)^{1/2} \frac{e^{-a\nu/2}}{\pi^{1/2}}, \tag{A24}$$

where $A = 0.322$, $p = 0.3$, and $a = 0.707$. The halo formation rate (HFR) is approximated by the positive cosmic time derivative of the halo mass function, N_{ST}, valid for $z \gtrsim 1.5$. It is given by

$$\frac{dN_{\mathrm{ST}}}{dt} = N_{\mathrm{ST}} \frac{d \ln f_{\mathrm{ST}}(\nu)}{dt}$$
$$\approx N_{\mathrm{ST}} \left[\frac{a\nu}{2} + \frac{p}{1 + (a\nu)^p}\right] \frac{d \ln \nu}{dz} \left|\frac{dz}{dt}\right|, \tag{A25}$$

where $dz/dt = -H_0(1 + z)E(z)$ with $E(z) = \sqrt{\Omega_{\Lambda,0} + \Omega_{m,0}(1 + z)^3}$.

This paper has been typeset from a TEX/LATEX file prepared by the author.