

The PAU Survey: an improved photo-*z* sample in the COSMOS field

Alex Alarcon¹,^{1,2,3}★ Enrique Gaztanaga¹,^{2,3} Martin Eriksen,^{4†} Carlton M. Baugh,^{5,6} Laura Cabayol,⁴ Ricard Casas,^{2,3} Jorge Carretero,^{4†} Francisco J. Castander,^{2,3} Juan De Vicente,⁷ Enrique Fernandez,⁴ Juan Garcia-Bellido¹,⁸ Hendrik Hildebrandt,⁹ Henk Hoekstra¹,¹⁰ Benjamin Joachimi,¹¹ Giorgio Manzoni¹,^{5,6,12} Ramon Miquel,^{4,13} Peder Norberg,^{5,6,12} Cristobal Padilla,⁴ Pablo Renard¹,^{2,3} Eusebio Sanchez,⁷ Santiago Serrano,^{2,3} Ignacio Sevilla-Noarbe,⁷ Malgorzata Siudek¹,⁴ and Pau Tallada-Crespi^{7†}

¹HEP Division, Argonne National Laboratory, Lemont, IL 60439, USA

²Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain

³Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain

⁴Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain

⁵Institute for Computational Cosmology (ICC), Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

⁶Institute for Data Science (IDAS), Durham University, South Road, Durham DH1 3LE, UK

⁷CIEMAT, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Avda. Complutenses 40, E-28040 Madrid, Spain

⁸Instituto de Física Teórica (IFT-UAM/CSIC), Universidad Autónoma de Madrid, E-28049 Madrid, Spain

⁹Ruhr-University Bochum, Astronomical Institute, German Centre for Cosmological Lensing, Universitätsstr. 150, D-44801 Bochum, Germany

¹⁰Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands

¹¹Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

¹²Centre for Extragalactic Astronomy (CEA), Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

¹³Institució Catalana de Recerca i Estudis Avançats (ICREA), E-08010 Barcelona, Spain

Accepted 2020 November 17. Received 2020 November 17; in original form 2020 August 13

ABSTRACT

We present – and make publicly available – accurate and precise photometric redshifts in the ACS footprint from the COSMOS field for objects with $i_{AB} \leq 23$. The redshifts are computed using a combination of narrow-band photometry from PAUS, a survey with 40 narrow bands spaced at 100 Å intervals covering the range from 4500 to 8500 Å, and 26 broad, intermediate, and narrow bands covering the UV, visible and near-infrared spectrum from the COSMOS2015 catalogue. We introduce a new method that models the spectral energy distributions as a linear combination of continuum and emission-line templates and computes its Bayes evidence, integrating over the linear combinations. The correlation between the UV luminosity and the O II line is measured using the 66 available bands with the zCOSMOS spectroscopic sample, and used as a prior which constrains the relative flux between continuum and emission-line templates. The flux ratios between the O II line and H α , H β and O III are similarly measured and used to generate the emission-line templates. Comparing to public spectroscopic surveys via the quantity $\Delta_z \equiv (z_{\text{photo}} - z_{\text{spec}})/(1 + z_{\text{spec}})$, we find the photometric redshifts to be more precise than previous estimates, with $\sigma_{68}(\Delta_z) \approx (0.003, 0.009)$ for galaxies at magnitude $i_{AB} \sim 18$ and $i_{AB} \sim 23$, respectively, which is three times and 1.66 times tighter than COSMOS2015. Additionally, we find the redshifts to be very accurate on average, yielding a median of the Δ_z distribution compatible with $|\text{median}(\Delta_z)| \leq 0.001$ at all redshifts and magnitudes considered. Both the added PAUS data and new methodology contribute significantly to the improved results. The catalogue produced with the technique presented here is expected to provide a robust redshift calibration for current and future lensing surveys, and allows one to probe galaxy formation physics in an unexplored luminosity-redshift regime, thanks to its combination of depth, completeness, and excellent redshift precision and accuracy.

Key words: galaxies: distances and redshifts – galaxies: evolution – galaxies: photometry.

1 INTRODUCTION

Redshift galaxy surveys can be broadly divided into two categories: spectroscopic surveys and imaging surveys. The former obtains high-

resolution spectra of the object within some wavelength coverage, which is used to identify sharp features like emission and absorption lines to nail the redshift of the object with very high precision. However, these are expensive to obtain: they require knowing the position of the object beforehand and a large exposure time, which makes it observationally inefficient to observe faint objects over a large area. Such surveys also suffer from incompleteness both because not all objects in the field are always targeted and also since

* E-mail: aalarcongonzalez@anl.gov

† Also at Port d'Informació Científica (PIC), Campus UAB, C. Albareda s/n, 08193 Bellaterra (Cerdanyola del Vallès), Spain.

a fraction of the measured spectra fail to provide an accurate redshift, for example for the lack of obvious emission or absorption features in low signal-to-noise spectra, when only one line is observed, or when there is a line confusion. In contrast, imaging surveys are able to obtain measurements of every object in the field of view at the same time from a set of bandpass filtered images, which allows to cover large areas faster and to a greater depth. This happens at the expense of getting flux measurements with very poor spectral resolution since the width of typical broad-band filters is larger than 100 nm, which makes the photometric redshift (i.e. photo- z) determination much less precise and sometimes inaccurate.

The precision of photometric redshift surveys can be improved with narrower bands and with a broader wavelength range coverage (e.g. see fig. 3 from Salvato, Ilbert & Hoyle 2019), and in recent years a new generation of multiband photometric imaging surveys have emerged, for example, spanning from the ultraviolet (UV) to infrared (IR) (COSMOS, Ilbert et al. 2009), using sets of intermediate bands (SHARDS or ALHAMBRA, Pérez-González et al. 2013; Molino et al. 2014), or a number of narrow-band (NB) filters (PAUS, Eriksen et al. 2019). From the more precise photometric redshifts one can extract better measurements of galaxy properties (luminosity, stellar mass, star formation rate) to probe and understand galaxy formation and galaxy evolution physics with a denser sample, at higher redshifts, and with little selection effects. One can measure galaxy clustering in thin redshift shells as a function of these intrinsic galaxy properties, measure luminosity functions and star formation histories, or use galaxy-galaxy lensing to constrain the galaxy to halo connection in a less explored luminosity-redshift regime. In particular, the Physics of the Accelerating Universe Survey (PAUS) is an ongoing NB imaging survey that intends to cover 100 deg² using 40 NB filters, increasing the number of objects with sub percent redshift precision by two orders of magnitude (Eriksen et al. 2019). The role of environment in structure formation is limited by the poor redshift precision in broad-band surveys and by the tiny area or low density in spectroscopic surveys. The unique combination of area, depth and redshift resolution from PAUS allows to sample with high density several galaxy populations, with which studies targeting non-linear galaxy bias, intrinsic alignments, magnification or density field reconstruction can be developed.

On the other hand, imaging weak lensing galaxy surveys have entered the era of precision cosmology and have become one of the most powerful probes for the Lambda cold dark matter cosmological model by measuring the shape and position of hundreds of millions of galaxies. Current and future surveys such as the Dark Energy Survey (DES, Troxel et al. 2018; Abbott et al. 2018a,b), the Kilo-Degree Survey (KiDS, Hildebrandt et al. 2017, 2020; Wright et al. 2019b, 2020), Hyper Suprime-Cam survey (HSC, Aihara et al. 2018; Hikage et al. 2019), the Legacy Survey of Space and Time (LSST, LSST Dark Energy Science Collaboration 2012), or the *Euclid* mission (Laureijs et al. 2011) are reaching or will reach a point where systematic uncertainties limit the full exploitation of their statistical power. Among these, one of the most challenging systematic uncertainties is the characterization of the redshift distribution of the weak lensing tomographic samples, which contain millions of faint galaxies with a few colours measured using broad-band filters. A correct description of such redshift distributions is crucial to avoid introducing a bias in the cosmological inference (Huterer et al. 2006; Cunha et al. 2012; Hildebrandt et al. 2012; Benjamin et al. 2013; Huterer, Cunha & Fang 2013; Bonnett et al. 2016; Hildebrandt et al. 2017; Joudaki et al. 2017; Hoyle et al. 2018; Joudaki et al. 2019) and to allow a robust comparison between cosmological parameters from weak lensing analysis and from the cosmic microwave background (CMB,

Planck Collaboration VI 2018), especially when a number of recent studies suggest a mild tension between the values of cosmological parameters inferred for the early and late-time universe (Asgari et al. 2019; Joudaki et al. 2019; Wright et al. 2020).

The really large area and depth covered by these surveys makes it unfeasible to measure spectroscopic redshifts for each galaxy of interest, which is why a number of alternative techniques have been developed over the years to estimate their redshift distributions. These can be broadly grouped as those which use angular cross-correlations with an overlapping tracer sample with well-characterized redshifts (*clustering redshifts*, see Newman 2008; Ménard et al. 2013; Schmidt et al. 2013; Davis et al. 2017; Hildebrandt et al. 2017; Gatti et al. 2018), those that model the galaxy spectral energy distribution (SED) of each galaxy to connect their observed colours to redshift (e.g. Benítez 2000; Arnouts & Ilbert 2011; Tanaka 2015; Hoyle et al. 2018) and those that model the colour redshift relation empirically using calibration samples (e.g. Cunha et al. 2012; Bonnett et al. 2016; De Vicente, Sánchez & Sevilla-Noarbe 2016; Buchs et al. 2019; Wright et al. 2019a). Each of these methods present different intrinsic systematics and potential biases, and are typically best used in combination (e.g. Hildebrandt et al. 2017; Hoyle et al. 2018).

Direct or empirical calibration methods rely on galaxy samples where abundant redshift information is available, either through spectroscopy or many band photometric redshifts. The former can be very accurate, but estimates are only available for a subset of the sample, and the targeting strategy, quality selection and incompleteness can introduce a statistical redshift bias with respect to the redshift of the full sample (Bonnett et al. 2016; Gruen & Brimiouille 2017; Speagle et al. 2019; Hartley et al. 2020; Wright et al. 2020). On the other hand, multiband photometric surveys provide a complete redshift sample at the expense of degrading the redshift precision, and can be biased if the galaxy SED modelling is incorrect. The COSMOS field (Scoville et al. 2007) contains the most widely used multiband redshift calibration survey, which provides a unique combination of deep photometric observations ranging from the UV to the IR over an area of ~ 2 deg², and several photometric redshift catalogues have been produced over the years (Ilbert et al. 2009, 2013; Laigle et al. 2016).

Here, we combine the multiband photometry from Laigle et al. (2016), hereafter COSMOS2015, with 40 NB filters from the PAUS survey (Padilla et al. 2019) which span the wavelength range from 4500 to 8500 Å. The unique PAUS photometric set is able to determine very precise photometric redshifts thanks to its exquisite wavelength sampling, which specifies precisely the location of the very sharp features in the galaxy SED (Eriksen et al. 2019, 2020). To estimate the photometric redshifts we develop an algorithm that models the galaxy SED as a linear combination of continuum and emission-line templates, and marginalises over different combinations computing a Bayesian integral. Furthermore, we calibrate priors between the continuum and emission-line templates using a subsample with spectroscopic redshifts and the multiband photometry. In Eriksen et al. (2019), a similar model (BCNZ2) was used, where the best-fitting linear combination of templates was calculated for each galaxy and model instead of the Bayesian integral. There, BCNZ2 was used to measure redshifts using the 40 NBs from PAUS and a subset of six broad bands from COSMOS2015, for objects with $i_{AB} \leq 22.5$ until redshift $z_{max} = 1.2$. Here, we use a total of 66 bands which include 40 narrow bands from PAUS and a combination of 26 narrow, intermediate, and broad bands from COSMOS2015, and we extend the magnitude and redshift limits to $i_{AB} \leq 23$ and $z_{max} = 3$. We focus on objects that were identified as galaxies in COSMOS2015, removing AGN X-ray detected sources, for which

several photo-*z* studies exist (Salvato et al. 2009, 2011; Simm et al. 2015; Marchesi et al. 2016). We make the redshift catalogue publicly available, including the redshift distribution of each object.

This paper is organized as follows. In Section 2, we describe the photometric data catalogues and the spectroscopic redshift catalogue used in this work. Section 3 describes the methodology used to describe the galaxy SED and infer the photometric redshift. Section 4 presents the primary photometric redshift results of this work. In Section 5, we discuss more details of the analysis and possible extensions for future work. We conclude and summarize in Section 6. There are four appendices, which contain details of how to download the catalogue (Appendix A), details of how we combine heterogeneous data (Appendix B), the calibration of zero-point offsets (Appendix C), and the population prior on the models (Appendix D). Details of the photo-*z* algorithms can be found online.

2 DATA

In this section, we describe the data we are going to use throughout. We will use NB photometry from the PAU Survey, a combination of narrow, intermediate, and broad bands coming from various instruments publicly released by the COSMOS Survey, and a spectroscopic redshift catalogue including measurements from several public redshift surveys.

2.1 PAUCam NB photometry

The PAUS is an ongoing imaging survey using a unique instrument, PAUCam (Padilla et al. 2019), mounted at the William Herschel Telescope (WHT) and located in the Observatorio del Roque de los Muchachos (La Palma, Canary Islands, Spain). PAUCam carries a set of 40 NB filters with 12.5 nm full width at half-maximum (FWHM) that span the wavelength range from 450 to 850 nm, in steps of 100 nm. PAUS has been collecting data since 2015 during several observing runs imaging five different fields: the COSMOS field (Scoville et al. 2007) and the CFHT W1, W2, W3, and W4 fields.¹ The PAU/CFHT fields are larger and represent the main survey, which is intended to observe up to 100 deg², while the COSMOS field (2 deg²) has been targeted as a calibration field since many photometric observations already exist, ranging from UV all the way to far-IR, as well as spectroscopic surveys with relatively high completeness and depth. In this work, we use the data collected in the COSMOS field from campaigns between 2015 and 2017.

2.1.1 Data reduction overview

At the end of each observing night, the data taken at WHT is sent to Port d'Informació Científica (PIC) for its storage and processing (Tonello et al. 2019). The data reduction process starts with initial de-trending, where a number of signatures from the instrument are removed from the images, using the NIGHTLY pipeline (see Serrano in preparation, Castander in preparation for details). This includes removing electronic bias with an overscan subtraction, correcting the gain from the different amplifiers and compensation from readout patterns using bias frames. A master flat is created from exposures of the dome with homogeneous illumination that are taken every afternoon before the observation. It is used to correct the vignetting of the telescope corrector, among other effects such as dead and hot pixels. Each individual NB filter is only covering a single CCD,

instead of a unique broad-band that covers all the focal plane, and the visible edges of the filters in the supporting grid of the filter tray produced scattered light in the image edges. An adjustment to the camera in 2016 significantly reduced this effect, which is partly mitigated by the pipeline by using a low pass filter with sigma clipping. Cosmic rays are identified using a Laplacian edge detection (van Dokkum 2001) and masked from the image.

An astrometric solution is added to align the different exposures using SCAMP (Bertin 2011) by comparing to *GAIA* DR1 (Gaia Collaboration et al. 2016). To estimate the point spread function (PSF) and perform the photometric NB calibration, we select stars with $i_{AB} < 21$ from the Sloan Digital Sky Survey (SDSS) (Smith et al. 2002). The PSF is modelled using PSFEX (Bertin 2011) in these stars. We calibrate the NB photometry (Castander in preparation) using the SDSS $u, g, r, i,$ and z *psfMag* magnitudes of these stars, performing the following procedure. We fit the Pickles stellar library (Pickles 1998) to the SDSS photometry of the star assuming the Milky Way extinction of the star is given by the value from the Planck 2015 thermal dust map Planck Collaboration X (2016). Here we approximate all stars to go through all the dust of the MW in that direction, since the majority of our calibration stars are fainter than magnitude 17–18 in the r band. Then, we generate synthetic NB observations without the Milky Way extinction for each SED and combine them with a weight equal to their probability from the SDSS fit. This combined synthetic narrow band is compared to observations to obtain the NB zero-points for the star. Finally, all star zero-points in the same image are combined to obtain one zero-point per image and narrow band. Note that this procedure also corrects for Milky Way extinction as the synthetic NB fluxes are generated excluding the Milky Way extinction.

NB photometry is obtained using the MEMBA pipeline (Multi-Epoch and Multi-Band Analysis, Serrano in preparation, Gaztanaga in preparation). In general, we rely on deep overlapping observations from lensing surveys to provide a detection catalogue with high-quality shape measurements to perform forced aperture photometry. In the COSMOS field, positions and shape measurements from ACS are used. The half-light radius, r_{50} , is used along with ellipticity measurements from Sargent et al. (2007) and the PAUS PSF FWHM to determine the aperture size and shape to target 62.5 per cent of the light, set to optimize the signal-to-noise ratio. A flux measurement is obtained for each individual exposure using this aperture measurement and a background subtraction estimated from a fixed annulus of 30 to 45 pixels around the source, where sources falling in the annulus get sigma clipped (for more details of the annulus background subtraction see Cabayol et al. 2019). Fluxes measured in different exposures are corrected with the estimated image zero-points and get combined with a weighted average to produce a narrow band coadded flux measurement.

The data reduction pipeline propagates flags for each individual exposure and object, and flagged measurements (indicating problems in the photometry) are not included in the weighted average. We remove objects with fewer than 30 NB measurements.

2.2 COSMOS survey photometry

Along with the NB data described in the previous section, we include photometry from several filters from the released COSMOS2015 catalogue² (Laigle et al. 2016), with filters from UV to near-infrared (NIR). Here, we list the bands, we use in this work,

¹<http://www.cfht.hawaii.edu/Science/CFHLS/cfhtlsdeepwidefields.html>

²ftp://ftp.iap.fr/pub/from_users/hjmcc/COSMOS2015/

along with the original instrument/survey: NUV data from *GALEX*; u^* from the Canada–France Hawaii Telescope (MegaCam); (B , V , r , i^+ , z^{++}) broad bands, ($IA427$, $IA464$, $IA484$, $IA505$, $IA527$, $IA574$, $IA624$, $IA679$, $IA709$, $IA738$, $IA767$, $IA827$) intermediate bands and ($NB711$, $NB816$) narrow bands from Suprime-Cam/*Subaru*; Y broad-band from HSC/*Subaru*; (Y , J , H , K_s) from VIRCAM/VISTA (UltraVISTA-DR2); and (H , K_s) data from WIRCAM/CFHT.

We point the reader to (Laigle et al. 2016) and references therein for a detailed overview of these observations and the data reduction. We use the 3 arcsec diameter PSF homogenized coadded flux measurements available in COSMOS2015 and apply several corrections as described and provided in the catalogue release. In particular, we correct for Milky Way dust extinction using the $E(B - V)$ value available for each galaxy and an effective factor F_x for each filter x (see table 3 and equation 10 in Laigle et al. 2016) using

$$\text{mag}_{\text{corrected}} = \text{mag}_{\text{uncorrected}} - E(B - V) \times F_x, \quad (1)$$

where the factors F_x are derived from the filter response function and integrated against the galactic extinction curve (for a sense of scale, $F_{\text{NUV}} \approx 8.6$ and $F_{z^{++}} \approx 1.5$). The fluxes have also been corrected from aperture to total flux, using the column OFFSET. In addition, we remove masked objects by using the flag parameter FLAG_PETER = 0, we remove both objects identified as stars or AGN by selecting TYPE = 0, and we select objects with $0 < \text{PHOTOZ} < 9$.

2.3 Combined photometric catalogue

We combine the NB catalogue from PAUS and the photometry from COSMOS2015 by matching objects by (RA, Dec.) position, keeping objects within 1 arcsec radius. We restrict the analysis to objects with an AUTO i -band magnitude brighter than $i_{\text{AB}} \leq 23$, which we obtain from Ilbert et al. (2009) (i_{AB} is also the reference magnitude in the PAUS data reduction). For objects fainter than this magnitude the PAUS NB photometry has a typical signal-to-noise ratio well below five in all bands (see fig. 2 in Eriksen et al. 2019), but emission lines can still be significantly detected in the NB filters. We defer to future work how including PAUS NBs can improve the photo- z performance in this fainter regime. The final catalogue after masking contains 40672 galaxies within the ACS footprint (see fig. 1 from Laigle et al. 2016, for different COSMOS footprints).

The combination of heterogeneous photometry from different instruments can be complicated and ultimately degrade the photometric redshift performance of the catalogue if it is not performed consistently. While in the case of PAUS the flux measurement is obtained from a variable aperture that targets 62.5 per cent of the total light of each galaxy, adapting to the galaxy’s apparent size and taking into account the PSF of each individual image, the COSMOS2015 photometry measures flux with a fixed aperture of 3 arcsec on PSF homogenized images, which is then corrected from aperture flux to total flux. Therefore, each survey measures a different fraction of the light for each galaxy, which depends on its apparent size. To deal with this effect, we have developed a self-calibration algorithm that benefits from the overlap between the *Subaru* r band and the PAUS narrow bands. We introduce a synthetic *Subaru* r -band flux, defined as a linear combination of NB flux measurements, and find a rescaling factor for each object that we apply to its NB photometry. In this way, we homogenize the photometry across surveys. We describe this method in Appendix B. Finally, we note that the colours measured by each survey could also be slightly different, but we will consider this is a negligible effect.

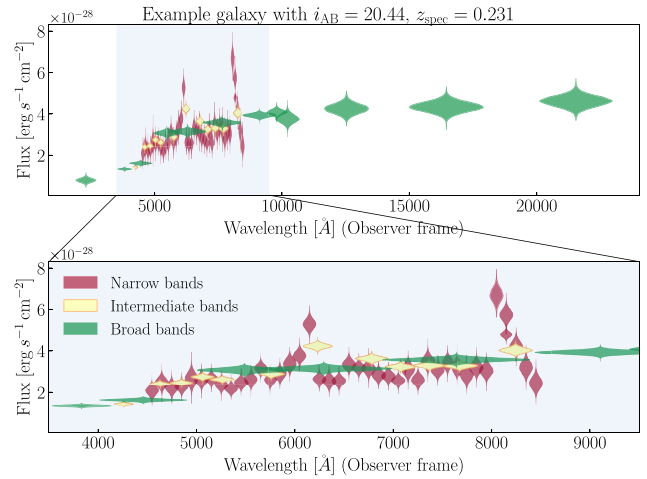


Figure 1. The 66 photometric bands used in this work for an example galaxy from the catalogue which has magnitude $i_{\text{AB}} = 20.44$ and spectroscopic redshift $z_{\text{spec}} = 0.231$. Bands with qualitatively similar FWHM have the same colour: broad bands in green, intermediate bands in yellow, and narrow bands in red. The horizontal width of the violin plots shows the FWHM of each filter, while the violin plot is a Gaussian distribution centred at the measured flux and with variance equal to the measured flux variance plus an extra systematic variance. Two blocks of emission lines are clearly visible in the narrow bands, one containing the redshifted H_δ line around $\lambda \sim 8100 \text{ \AA}$, and another containing the O III doublet and H_β around $\lambda \sim 6100 \text{ \AA}$.

In addition, we add a systematic error to each band for all objects, which accounts for any residual error that remains either from inaccuracies in the galaxy modelling, errors in the data reduction or from joining inconsistent heterogeneous photometry from different surveys. We add a 2 per cent error to every band except for the UltraVista broad bands Y , J , H , K_s (5 per cent error) and for the *GALEX* NUV band (10 per cent error), similar to those used by COSMOS2015 (private communication).

Fig. 1 shows an example galaxy with the 66 photometric bands we use in this work, coloured according to their width (broad bands in green, intermediate bands in yellow and narrow bands in red). The horizontal width of the violins shows the FWHM of each filter, while the violin is showing a Gaussian distribution centred at the measured flux and with variance equal to the measured flux variance plus the previously described extra systematic variance. Regarding the photometric band completeness, after flagging, 68 per cent of the catalogue has measurements available for all 66 bands, while 92 per cent of objects have at least 65 bands measured. All objects have measurements for the 12 intermediate bands and the four Y , J , H , K_s UltraVista NIR bands. All bands have measurements available in at least 98 per cent of the catalogue, except for the *GALEX* NUV band which is missing in 23 per cent of the objects.

We make the photometric redshift catalogue publicly available along with the redshift distribution of each object (for details of the catalogue and how to download it see Appendix A).

2.4 Spectroscopic data

To measure the precision and accuracy of the photometric redshifts we compare to spectroscopic redshifts. We use a compilation of public spectroscopic surveys (courtesy of Mara Salvato, private communication) and we apply a quality flag to keep only objects with a very reliable measurement. This compilation includes redshifts from the following instruments or surveys: zCOSMOS DR3 (Lilly

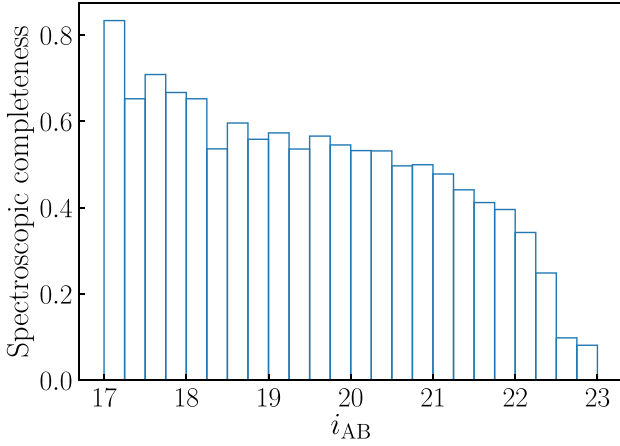


Figure 2. The spectroscopic completeness of the spectroscopic catalogue used in this work as a function of i_{AB} magnitude. The completeness is defined as the fraction of objects in the catalogue, which have a spectroscopic redshift measured. The i_{AB} magnitude is the reference magnitude in the PAUS data reduction, and is the AUTO i -band magnitude from the Ilbert et al. (2009) catalogue.

et al. 2009), C3R2 DR1&DR2 Masters et al. (2017), Masters et al. (2019), 2dF (Colless 1999), DEIMOS (Casey et al. 2017; Hasinger et al. 2018; Masters et al. 2019), FMOS (Kashino et al. 2019), LRIS (Lee et al. 2018), MOSFIRE (Kriek et al. 2015), MUSE (Rosani et al. 2019), Magellan (Calabrò et al. 2018), and VIS3COS (Paulino-Afonso et al. 2018). Whenever more than one redshift measurement is available for the same object we take the mean of all observations, and if the multiple observations disagree by more than 0.002 in redshift, we do not assign any spectroscopic redshift to that object (which removes 95 objects in the spectroscopic catalogue). The photometric catalogue from Section 2.3 has a total of 12112 objects with an overlapping spectroscopic match. Fig. 2 shows a summary of the spectroscopic completeness of the catalogue as a function of i_{AB} magnitude, which stays above 40 per cent for bright magnitudes and decreases below 10 per cent for the fainter objects considered, at $i_{AB} > 22.5$. Since the spectroscopic redshifts in this catalogue come from high-resolution instruments, we expect their precision error to be negligible compared to the photo- z precision. However, there could be a sample of outlier spectroscopic measurements that we cannot flag, since the amount of duplicate spectroscopic measurements is very small.

3 METHODOLOGY

In this section, we describe the methodology used to model the SED of a galaxy and obtain its redshift distribution.

3.1 Redshift distribution

We want to obtain the redshift probability distribution of a galaxy $p(z|f)$ from some flux observations f . To model the relation between redshift and f , we introduce a set of models denoted by $\{M\}$ that can predict the fluxes f as a function of redshift for different type of galaxies. Therefore, we write

$$p(z|f) = \sum_{\{M\}} p(z, M|f) \propto \sum_{\{M\}} p(f|z, M)p(z, M). \quad (2)$$

Each model M is defined as a linear combination with parameters $\{\alpha_j\}$ of a particular set of SEDs $\{t_j\}$ (SEDs),

$$M(z) = \sum_j \alpha_j(z) t_j(z). \quad (3)$$

The SED templates $\{t_j\}$ can either be continuum templates of different galaxy populations or the flux from emission lines. In this work we will use two continuum templates and one emission-line template, so $\alpha = \{\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, \alpha^{\text{EL}}\}$. For details on the SED templates see the next subsections. We predict the colours at different redshifts by redshifting the rest-frame SEDs and convolving with each filter accordingly. The amplitudes α are the free parameters of the model at each redshift. The probability $p(f|z, M)$ is given by

$$p(f|z, M) = \int p(f, \alpha|z, M) d\alpha, \\ = \int p(f|\alpha, z, M) p(\alpha|z, M) d\alpha, \quad (4)$$

which is commonly referred as the Bayes evidence. We assume the likelihood $p(f|\alpha, z, M)$ is given by a normal multivariate distribution,

$$p(f|\alpha, z, M) = \frac{1}{\sqrt{(2\pi)^d \prod_k \sigma(f_k)}} \times \\ \exp \left[-\frac{1}{2} \sum_k \left(\frac{f_k}{\sigma(f_k)} - \sum_{j=1}^n \alpha_j \frac{t_{jk}}{\sigma(f_k)} \right)^2 \right], \quad (5)$$

where k runs over the bands, j runs over the templates in the model M , t_{jk} is the flux for band k and template t_j in the model, and $\sigma(f_k)$ is the measured flux error for band k .

The integral in equation (4) requires that we specify a prior on the model parameters $p(\alpha|z, M)$. We distinguish between continuum- and emission-line SEDs in the prior,

$$p(\alpha|z, M) = p(\alpha_0^{\text{Cont}}|z, M) \times \\ p(\alpha_1^{\text{Cont}}|z, M) \times \\ p(\alpha^{\text{EL}}|\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, z, M), \quad (6)$$

where we choose $p(\alpha_0^{\text{Cont}}, z, M)$ and $p(\alpha_1^{\text{Cont}}, z, M)$ to be top hat functions,

$$p(\alpha_0^{\text{Cont}}|z, M) \propto \Theta(\alpha_0^{\text{Cont}}) \Theta(\Delta_0 - \alpha_0^{\text{Cont}}), \\ p(\alpha_1^{\text{Cont}}|z, M) \propto \Theta(\alpha_1^{\text{Cont}}) \Theta(\Delta_1 - \alpha_1^{\text{Cont}}), \quad (7)$$

with Θ the Heaviside step function. Ideally, the prior for the continuum amplitudes (equation 7) should contain information about the luminosity function. We use a top hat prior between 0 and a maximum flux of (Δ_0, Δ_1) , which regards values outside these bounds as unphysical. The values (Δ_0, Δ_1) are set such that the maximum flux does not exceed a given maximum luminosity threshold in the *Subaru i* band, and the thresholds are calculated from the *Subaru i*-band absolute magnitude which is derived from the data and the best model (see Section 3.3).

The prior probability term $p(\alpha^{\text{EL}}|\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, z, M)$ is one of the key ingredients of our modelling, which constrains the colours that arise from emission lines with respect to continuum flux. We estimate this prior using data, modelling a relation between the luminosity in the UV and the luminosity of the OII line, which we describe in detail in Section 3.3.

We note that the Bayesian integral of equation (4) is a generalization of previous work (see EAZY or BCNZ2 codes Brammer, van Dokkum & Coppi 2008; Eriksen et al. 2019), which approximate the

integral with the maximum likelihood

$$p(f|z, M) \approx p(f|\alpha_{\max}, z, M), \quad (8)$$

where α_{\max} are the maximum likelihood values of the parameters within the positive orthant. We will use this approximation only when calibrating the prior for the emission lines and in the zero-point calibration step (see Sections 3.3 and 3.4). To find the maximum likelihood values α_{\max} , we will use the minimization algorithm from BCNZ2 (Eriksen et al. 2019).

To efficiently compute the integral from equation (4), we have implemented a code based on a Gaussian integral algorithm from Genz (1992). Details of the algorithms can be found online.

3.2 Galaxy SED

3.2.1 Continuum templates

Here, we describe the continuum galaxy SED templates used in this work. We will use a library of synthetic SED templates generated using recipes from Bruzual & Charlot (2003) and Polletta et al. (2007). This library, which is similar to the ones used in Arnouts & Ilbert (2011), Ilbert et al. (2009), Ilbert et al. (2013), Laigle et al. (2016), Eriksen et al. (2019), contains a set of Elliptical, Spiral and Starburst synthetic templates. The templates in this library do not account for the light attenuation due to the internal dust present in each galaxy. Generally, this effect varies in each galaxy. Following the aforementioned references, we model extinction using an extinction law $k(\lambda)$ and a colour excess $E(B - V)$ that adds this effect to the template as

$$F_{\text{observed}}(\lambda) = F_{\text{no dust}}(\lambda) \times 10^{-0.4 E(B-V)k(\lambda)}. \quad (9)$$

We include this effect by modifying our default templates and generating new ones with different amounts of dust attenuation, since we cannot parametrize it linearly. For the starburst galaxies, we will use the extinction law from Calzetti et al. (2000), while for spirals we will model dust attenuation using the Prevot et al. (1984) law. We will not add extinction to the reddest galaxy templates, like ellipticals templates. We include two modifications of the Calzetti law with an additional bump around 2175 Å (see Section 3.4 of Ilbert et al. 2009) which was not prominent in the original calibration of the law, but which has later been found in starburst galaxies (Stecher & Donn 1965; Xiang, Li & Zhong 2011; Fitzpatrick & Massa 1986, 2007). We use two different amplitudes of the bump following Laigle et al. (2016), Eriksen et al. (2019). We generate a grid of templates with different $E(B - V)$ values ranging from 0.05 to 0.5 in steps of 0.05. We implement the IGM absorption using the analytical correction from Madau (1995).

As mentioned in the previous section, each model M used in this work contains two continuum templates. The complete list of groups of two continuum templates is shown in Table 1. Some of the groups exist with different extinction laws created with a range of different $E(B - V)$ values, as indicated in the table. We have tested some combinations of three continuum templates but found it to have little impact on the probability. The Bayes evidence has preference for simpler models (an effect often named as Bayesian Occam's Razor, e.g. Ghahramani 2012) so that models with more templates than needed to describe the data naturally get a lower probability. A group of at least two continuum templates guarantees a more continuous coverage in colour space. We leave an exploration of other combinations of templates, and the addition of different synthetic templates to future work.

Table 1. List of continuum synthetic templates considered in different models M . We use continuum templates used in Laigle et al. (2016). The elliptical (Ell) and spiral (Sc) templates were generated by Polletta et al. (2007), while the starburst models (SB) were generated by Bruzual & Charlot (2003). The additional BC03 templates (BC) have their metallicity (Z) and age (Gyr) specified in parentheses, and were introduced in Ilbert et al. (2013). We apply reddening to the continuum SEDs using extinction laws for spiral and starburst templates, using a grid with 10 different $E(B - V)$ values spaced by 0.05 from 0.05 to 0.5 (see the text for further details). In total, there are 97 combinations of continuum templates with different extinction laws and extinction values.

Groups of continuum templates	Extinction laws
(0) BC(0.008,10), BC(0.008,13)	None
(1) BC(0.008,8), BC(0.008,10)	None
(2) BC(0.008,6), BC(0.008,8)	None
(3) BC(0.008,4.25), BC(0.008,6)	None
(4) BC(0.008,2.6), BC(0.008,4.25)	None
(5) BC(0.02,10), BC(0.02,13)	None
(6) BC(0.02,8), BC(0.02,10)	None
(7) BC(0.02,6), BC(0.02,8)	None
(8) BC(0.02,4.25), BC(0.02,6)	None
(9) BC(0.02,2.6), BC(0.02,4.25)	None
(10) Ell1, Ell4	None
(11) Ell4, Ell7	None
(12) Ell7, Sc	None
(13) Sc, SB0	None, Prevot
(14) SB0, SB4	None, Prevot
(15) SB4, SB8	None, Calzetti, Calz. + Bump1, Calz. + Bump2
(16) SB8, SB11	None, Calzetti, Calz. + Bump1, Calz. + Bump2

Table 2. List of emission lines included in the SED modelling. The lines are modelled with a Gaussian distribution with a width of 10 Å centred around the air wavelengths shown in the second column.

Line	λ (Å)
Ly α	1215.7
O II	3726.8
O II ₁	4959
O III ₂	5007
H α	6562.8
H β	4861
N II ₁	6548
N II ₂	6583
S II ₁	6716.4
S II ₂	6730.8

3.2.2 Emission lines

All the models M include a third template which models the flux from emission lines. Table 2 shows a list of the lines we include and the wavelength on which they are centred. We model each line with a Gaussian distribution of 10-Å width (similar to LEPHARE, Arnouts & Ilbert 2011, which accounts for some Doppler broadening due to the galaxy's rotational velocity). Therefore, our model will integrate over different combinations of continuum flux and emission-line flux, with a greater ability to describe the observed flux of each object than if we fixed these quantities.

We build the emission-line template t_{EL} as

$$t_{\text{EL}} \equiv \psi^{\text{O II}} + \beta^{\text{O III}} \psi^{\text{O III}} + \beta^{\text{H}\alpha} \psi^{\text{H}\alpha} + \beta^{\text{H}\beta} \psi^{\text{H}\beta}, \quad (10)$$

where

$$\begin{aligned}\psi^{\text{OII}} &\equiv \text{O II} + 2\text{Ly}_\alpha, \\ \psi^{\text{OIII}} &\equiv \frac{1}{3}\text{O III}_1 + \text{O III}_2, \\ \psi^{\text{H}_\alpha} &\equiv \text{H}_\alpha + 0.35\left(\frac{1}{3}\text{N II}_1 + \text{N II}_2 + \text{S II}_1 + \text{S II}_2\right), \\ \psi^{\text{H}_\beta} &\equiv \text{H}_\beta,\end{aligned}\quad (11)$$

The notation O II, for example, in the above equation means a Gaussian distribution centred at $\lambda = 3726.8$ and 10-\AA width whose flux integrates to $10^{-17}\text{erg s}^{-1}\text{cm}^{-2}$, and similar for the other lines using the wavelength values from Table 2. The parameters β^{OIII} , β^{H_α} , and β^{H_β} indicate the relative amount of flux between $(\psi^{\text{OIII}}, \psi^{\text{H}_\alpha}, \psi^{\text{H}_\beta})$ and ψ^{OII} and are determined from data (see Section 3.3).

The term ψ^{OII} contains the O II and Ly_α lines. We include the Ly_α line following previous photo-z analysis in the COSMOS field (Ilbert et al. 2009; Laigle et al. 2016) which used a fixed ratio of 2 between Ly_α and O II. The Ly_α line has a small impact since it only changes the flux of the *GALEX* NUV filter for most of the galaxies in this analysis, which has the smallest signal-to-noise ratio, and it only enters the CFHT *u* band above redshift 2. The term ψ^{OIII} contains the O III doublet, where the factor $\frac{1}{3}$ comes from atomic physics (e.g. Storey & Zeippen 2000). The term ψ^{H_α} contains the H_α line and the N II and S II doublets, where the factor $\frac{1}{3}$ also comes from atomic physics (Storey & Zeippen 2000). Note that the H_α and N II lines are essentially blended together in our filter set, and we assign a fixed ratio of 0.35 to the N II lines with respect to H_α , although this ratio is smaller for lower stellar mass galaxies (e.g. Faisst et al. 2018). H_α and S II doublet are distinguishable in different narrow bands until $z \sim 0.3$, where they get redshifted outside of the PAUS NB coverage and become blended, so we model them together to have an homogeneous modelling at all redshifts. The H_β line is modelled separately in ψ^{H_β} . The emission-line modelling presented here can be improved further using known relations: the BPT diagram (Baldwin, Phillips & Terlevich 1981) which establishes relations between the O III, N II, H_α and H_β lines; the intrinsic case B recombination Balmer decrement $(\text{H}_\alpha/\text{H}_\beta) = 2.86$ (Storey & Hummer 1995; Moustakas, Kennicutt & Tremonti 2006); or modelling S II and H_α differently below and above $z \sim 0.3$. We defer a thorough exploration of these possibilities for future work.

Finally, we apply the same extinction law for the continuum templates and the emission-line template, and modify the ratio values according to the attenuation from extinction. The amount of dust attenuation for emission lines in nebular regions can be different than that from stellar continuum (see Calzetti, Kinney & Storchi-Bergmann 1994; Calzetti et al. 2000; Puglisi et al. 2016; Saito et al. 2020). Measuring the ratios β^{OIII} , β^{H_α} , β^{H_β} and the prior of the emission-line template $p(\alpha^{\text{EL}}|\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, z, M)$ from data assuming the same amount of dust attenuation in the model yields a more consistent flux model overall (see Section 3.3).

3.3 Prior on template parameters

In this section, we describe how we estimate several parameters in the model using spectroscopic data. To recap, we need to estimate the parameters (Δ_0, Δ_1) introduced in the continuum templates prior (equation 7), describe the prior on the emission-line template $p(\alpha^{\text{EL}}|\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, z, M)$ (equation 6) and set the parameters $(\beta^{\text{OIII}}, \beta^{\text{H}_\alpha}, \beta^{\text{H}_\beta})$ from equation (11).

For each galaxy with a confident³ spectroscopic redshift in the zCOSMOS-Bright spectroscopic catalogue (see Section 2.4), we carry out the following steps:

- (i) Find the model M with the largest Bayes evidence at the galaxy's spectroscopic redshift (equation 4).
- (ii) For the most probable model M^{max} , we find the maximum likelihood parameters α_{max} using the minimization algorithm (equation 8).
- (iii) Subtract from the data the estimated continuum using the continuum α_{max} parameters.
- (iv) Produce Gaussian realizations centred at the subtracted flux using the measured flux error. For each realization, find the best fit values for that galaxy: $(\alpha_g^{\text{EL}}, \beta_g^{\text{OIII}}, \beta_g^{\text{H}_\alpha}, \beta_g^{\text{H}_\beta})$. Note that since we separately model additional internal dust reddening, these values are extinction free by definition.
- (v) Estimate the mean and standard deviation of the parameters $(\alpha_g^{\text{EL}}, \beta_g^{\text{OIII}}, \beta_g^{\text{H}_\alpha}, \beta_g^{\text{H}_\beta})$ from the previous step.

From step (ii), we can also obtain absolute magnitudes for every galaxy, which we can correct for internal dust extinction using the best model (and the best extinction parameters). Of interest for us are the absolute magnitudes M_{NUV} , M_I , which stand for the *GALEX* NUV and *Subaru i* bands. In summary, this algorithm provides the parameters $(\langle\alpha_g^{\text{EL}}\rangle, \langle\beta_g^{\text{OIII}}\rangle, \langle\beta_g^{\text{H}_\alpha}\rangle, \langle\beta_g^{\text{H}_\beta}\rangle, M_{\text{NUV}}, M_I)$ for each of the aforementioned galaxies.

The distribution of M_I peaks around magnitude -22 in our sample, and we find no galaxies brighter than -26 . We assume this value to be the brightest magnitude a galaxy could be in this band, and set the upper limit of the top hat prior of the continuum templates (Δ_0, Δ_1) (equation 7) to fulfil this condition at all redshifts.

Fig. 3 shows a density plot of the estimated mean values $(\beta_g^{\text{OIII}}, \beta_g^{\text{H}_\alpha}, \beta_g^{\text{H}_\beta})$. For many of the galaxies, the measured signal-to-noise ratio on these parameters is low. Therefore, the density plot is a convolution of some underlying distribution convolved with this noise. We do not attempt to recover the true distribution in this work, but instead just measure the median and σ_{68} of the marginal noisy distribution, finding:

$$\begin{aligned}\log_{10}(\beta_g^{\text{OIII}}) &= -0.50 \pm 0.35, \\ \log_{10}(\beta_g^{\text{H}_\alpha}) &= -0.56 \pm 0.34, \\ \log_{10}(\beta_g^{\text{H}_\beta}) &= -0.08 \pm 0.24.\end{aligned}\quad (12)$$

For each pair of continuum SEDs, we add one emission-line model with parameters β^{OIII} , β^{H_α} , and β^{H_β} equal to the median values from equation (12). In order to account for the breadth of the distribution, we include six additional models per continuum group, each with one of the β parameters set at a value two times the σ_{68} with respect to its median.⁴ In total, we have 679 different models M , each with different continuum or emission-line templates.

It is worth noting that step (i) in this section requires an initial assumption about the values of $(\beta^{\text{OIII}}, \beta^{\text{H}_\alpha}, \beta^{\text{H}_\beta})$ and the prior $p(\alpha^{\text{EL}}|\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, z, M)$. We initially fix these values to the line flux ratio values from Ilbert et al. (2009). After the initial run, we

³For this exercise we use only galaxies in the zCOSMOS-Bright DR3 release with a confidence flag c in the set: [3.x, 4.x, 2.4, 2.5, 1.5, 9.3, 9.4, 9.5, 18.3, 18.5]

⁴These six additional models can be expressed as $([2,0,0],[0,0,2],[0,2,0],[-2,0,0],[0,0,-2],[0,-2,0])$, where for example $[2,0,0]$ would mean $(\beta_g^{\text{H}_\alpha}, \beta_g^{\text{H}_\beta})$ are set at their median value, while β_g^{OIII} is set at its median value plus two times the measured σ_{68} .

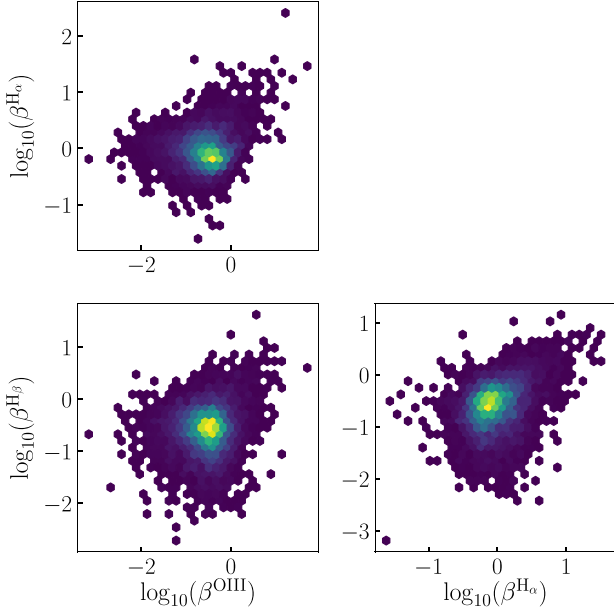


Figure 3. This figure shows the line flux ratio between (mainly) the O III, H α , and H β lines with respect to the O II line, as the density plot of the logarithmic measured values of β^{OIII} , $\beta^{\text{H}\alpha}$, and $\beta^{\text{H}\beta}$ (equation 10) which give the relative amount of flux between different emission lines, as defined by equation (11). The combined photometry of COSMOS and PAUS is used to make the measurement for each object. We fit models of the emission lines to a continuum subtracted measured flux for a subset of objects with spectroscopic redshift, correcting for internal dust extinction (see Section 3.3).

repeat the process to measure all of these parameters and calibrate the prior (which is described later in this section). We repeat this process a couple of times, after which we find the values do not change. All the values reported in this section are the final values, which are used in the remainder of this work.

3.3.1 Prior $p(\alpha^{\text{EL}}|\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, z, M)$

Our model requires a prior on the emission-line template free amplitude α^{EL} . We will use the known correlation between the UV luminosity of a galaxy and its O II emission-line flux, which has been used before in photometric redshift estimation (Kennicutt 1998; Ilbert et al. 2009), to build a model between α^{EL} and the absolute magnitude M_{NUV} . We assume the following relation:

$$p(\eta|M_{\text{NUV}}) \sim \mathcal{N}(\mu = aM_{\text{NUV}} + b, \sigma = c),$$

$$\eta \equiv -2.5 \log_{10}(\alpha^{\text{EL}}) - \text{DM}, \quad (13)$$

where DM is the distance modulus, \mathcal{N} is a Gaussian distribution, and a, b, c are parameters to be determined from data. The parameter η can be interpreted as an emission-line absolute magnitude. Therefore, we assume there is a linear relation between η and M_{NUV} with an intrinsic scatter perpendicular to the correlation given by a normal distribution. In the model, the absolute magnitude M_{NUV} is a function of the three free amplitudes, $M_{\text{NUV}} = M_{\text{NUV}}(\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, \alpha^{\text{EL}})$. Therefore, equation (13) describes the probability of α^{EL} given $(\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, z)$, which is our probability $p(\alpha^{\text{EL}}|\alpha_0^{\text{Cont}}, \alpha_1^{\text{Cont}}, z, M)$ for a given model M .

Fig. 4 shows the relation between η and M_{NUV} as a density plot of the measured values for each galaxy, with α^{EL} expressed in units of $10^{-17} \text{erg s}^{-1} \text{cm}^{-2}$. Similar to Fig. 3, the plot does not show

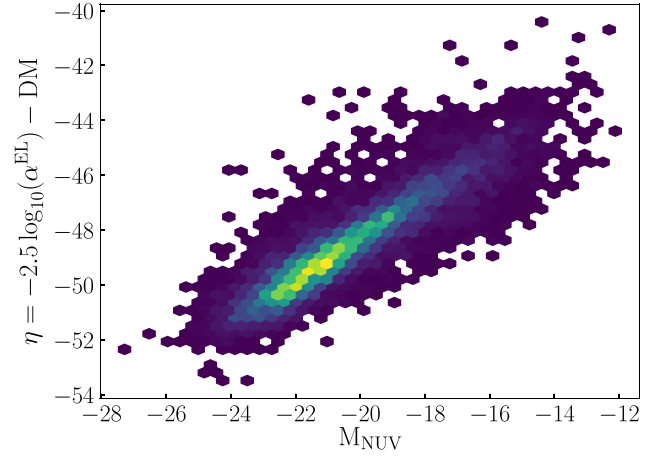


Figure 4. This figure shows the correlation between the O II line flux and the UV light as a density plot of the measured values of $\eta \equiv -2.5 \log_{10}(\alpha^{\text{EL}}) - \text{DM}$ and M_{NUV} (equation 13) for a subset of objects with spectroscopic redshift, where DM is the distance modulus. We expect these variables to be correlated since the O II line flux is correlated with the UV luminosity of a galaxy. We model this as a linear relation with an intrinsic Gaussian scatter, and find the most likely parameters of this model to calibrate the prior between continuum- and emission-line templates, which is one of the key ingredients in our redshift estimation model (see Section 3.3). The measurements have been corrected for internal extinction.

the relation directly, since it is convolved with the measurement noise of both variables. The inference of the parameters (a, b, c) in equation (13) has to be done carefully to avoid introducing a bias since the data is noisy in both axes (for more details see Kelly 2007; Hogg, Bovy & Lang 2010).

We will infer the parameters $\theta \equiv (a, b, c)$ by writing the likelihood of the observations given the parameters, following a model similar to Kelly (2007). Let η_g, M_g be the true values of the variables η, M_{NUV} for galaxy g , and $\hat{\eta}_g, \hat{M}_g$ be their noisy observational counterparts. We will assume M_g follows from a probability distribution $p(M_g|\xi)$, where ξ are the parameters of the distribution. The joint distribution of η_g and M_g is then $p(\eta_g, M_g|\theta, \xi) = p(\eta_g|M_g, \theta)p(M_g|\xi)$. We will assume a Gaussian and independent measurement error in $\hat{\eta}_g$ and \hat{M}_g , so that $p(\hat{\eta}_g, \hat{M}_g|\eta_g, M_g) = p(\hat{\eta}_g|\eta_g)p(\hat{M}_g|M_g)$ are two Gaussian distributions with means (η_g, M_g) and variances $(\sigma^2(\hat{\eta}), \sigma^2(\hat{M}_g))$. Therefore, we can hierarchically express the model as

$$M_g \sim p(M_g|\xi),$$

$$\eta_g|M_g \sim \mathcal{N}(aM_g + b, c),$$

$$\hat{\eta}_g, \hat{M}_g|\eta_g, M_g \sim \mathcal{N}(\eta_g, \sigma(\hat{\eta})) \times \mathcal{N}(M_g, \sigma(\hat{M}_g)). \quad (14)$$

The likelihood function of the measured data $p(\hat{\eta}_g, \hat{M}_g|\theta, \xi)$ can be obtained by integrating the complete data likelihood over the missing data η_g, M_g

$$p(\hat{\eta}_g, \hat{M}_g|\theta, \xi) = \int \int p(\hat{\eta}_g, \hat{M}_g, \eta_g, M_g|\theta, \xi) d\eta_g dM_g,$$

$$= \int \int p(\hat{\eta}_g, \hat{M}_g|\eta_g, M_g)p(\eta_g|M_g, \theta) \times p(M_g|\xi) d\eta_g dM_g. \quad (15)$$

In this work, we will model $p(M_g|\xi)$ with a mixture of K Gaussian distributions,

$$p(M_g|\xi) = \sum_{k=1}^K \frac{\pi_k}{\sqrt{2\pi\tau_k^2}} \exp \left\{ -\frac{1}{2} \frac{(M_g - \mu_k)^2}{\tau_k^2} \right\}, \quad (16)$$

where $\sum_k \pi_k = 1$. Defining $\pi \equiv (\pi_1, \dots, \pi_K)$, $\mu \equiv (\mu_1, \dots, \mu_K)$ and $\tau \equiv (\tau_1, \dots, \tau_K)$, note that we have $\xi = (\pi, \mu, \tau)$. This mixture model is flexible enough to describe a wide variety of distributions, and it is also convenient since it simplifies the mathematics for writing the likelihood of the measured data (see Kelly 2007). Assuming the data for different galaxies is statistically independent, the full data likelihood is the product of the measurement likelihood of each galaxy

$$p(\{\hat{\eta}_g\}, \{\hat{M}_g\} | \theta, \xi) = \prod_{g=1}^n \prod_{k=1}^K \frac{\pi_k}{2\pi |V_{g,k}|^{1/2}} \times \exp \left\{ -\frac{1}{2} (z_g - \zeta_k) V_{g,k}^{-1} (z_g - \zeta_k)^T \right\}, \quad (17)$$

with

$$\begin{aligned} z_g &= (\hat{M}_g, \hat{\eta}_g), \\ \zeta_k &= (a\mu_k + b, \mu_k), \\ V_{g,k} &= \begin{pmatrix} a^2\tau_k^2 + c^2 + \sigma^2(\hat{\eta}_g) & a\tau_k^2 \\ a\tau_k^2 & \tau_k^2 + \sigma^2(\hat{M}_g) \end{pmatrix}. \end{aligned} \quad (18)$$

We fix the number of mixture Gaussians to $K = 2$ (although we have verified that the results for θ do not change if $K = \{2, 3, 4, 5\}$). We maximize the likelihood in equation (17) and find the most likely parameters θ^{ML} to be

$$a^{\text{ML}} = 0.750 \quad b^{\text{ML}} = -33.38 \quad c^{\text{ML}} = 0.327 \quad (19)$$

We will use these values for all the results in this work. Finally, we numerically integrate the prior, equation (6), to compute the prior normalization, which is needed for the Bayes evidence, using a Metropolis–Hastings integration algorithm.

3.4 Systematic zero points offsets

A common approach in the literature is to find systematic relative (not global) zero-points between different bands before running the photo-z algorithm (e.g. Benítez 2000; Coe et al. 2006; Hildebrandt et al. 2012; Molino et al. 2014; Laigle et al. 2016; Eriksen et al. 2019). This attempts to optimize the colours predicted by the model in comparison to the observed colours in the data to improve the photo-z (e.g. Dahlen et al. 2013). A zero-point offset does not need to come from the zero-point estimation itself, but can also be due to an incorrect PSF modelling (Hildebrandt et al. 2012), and from incorrect or missing templates.

We calibrate the systematic offsets with the same spectroscopic catalogue described in Section 3.3. We also use a similar algorithm. For every galaxy, we find the model M with the largest Bayes evidence M^{max} and its most likely parameters α_{max} . We build the most likely fluxes T_i according to M^{max} and α_{max} .

We assume the measured and predicted fluxes to be statistically independent for every galaxy and band, and find the offsets $\{\kappa_j\}$ that maximize the likelihood

$$\begin{aligned} p(\{f_g\}, \{\sigma_g\}, \{T_{g,k}\} | \{\kappa_j\}) &\approx \prod_k \prod_g p(f_g, \sigma_g, T_{g,k} | \kappa_j), \\ &= \prod_k \prod_g \mathcal{N}(f_g \kappa_j - T_{g,k}, \kappa_j \sigma_g). \end{aligned} \quad (20)$$

We apply the offsets (or factors for fluxes) to the data and run again, repeating the process until convergence. We exclude galaxies with a very bad fit ($\chi^2 > 120$, with ~ 63 degrees of freedom). The values of

the offsets can be found in Appendix C, in Table C1 and Fig. C1. It is worth noting that the calibration of the prior described in Section 3.3 and the offset calibration described in this section depend on each other. We hierarchically run each part of the calibration, using the prior parameters and the zero-point offsets from the previous step. We perform this a couple of times.

3.5 Population prior and redshift posterior

To obtain the redshift posterior $p(z|f)$ for each galaxy, we need to calculate equation (2), which requires us to know the population's distribution over different redshifts and models, $p(z, M)$. This quantity is unknown a priori, and previous template codes and analysis have made different assumptions, such as assuming it is uniform, or introducing analytical functions with hyperparameters calibrated with a spectroscopic population. Once a target population has been identified, the posterior on the redshift and model of each galaxy in the population can be jointly and hierarchically inferred along with the population's distribution over different redshifts and models (e.g. Leistedt, Mortlock & Peiris 2016). This can be further extended to include a dependence of galaxy density on the line-of-sight position due to galaxy clustering (see Alarcon et al. 2019; Sánchez & Bernstein 2019).

The redshift posterior of a galaxy is not unique since it depends on the population to which it belongs, or in other words different galaxy sample selections will yield different $p(z, M)$, and thus a different posterior $p(z|f)$ for each galaxy. One proposed application of this redshift sample is to empirically calibrate the redshift distribution of galaxy samples from weak lensing surveys. Various techniques exist (Alarcon et al. 2019; Buchs et al. 2019; Sánchez & Bernstein 2019; Wright et al. 2019a; Sánchez et al. 2020) which write a probability relation between the weak lensing galaxies and the galaxies from the calibration samples. The most correct output for such studies would be to produce and release the full likelihood $p(z, M|f)$ for each galaxy, so that the redshift posterior and population distribution can be inferred correctly for any galaxy sample. This is impractical since the likelihood contains over a million values for each galaxy. Instead, we will compute and release a pseudo probability $\tilde{p}(z|f)$ defined as

$$\tilde{p}(z|f) = \sum_{\{M\}} p(z, M|f) \approx \sum_{\{M\}} p(f|z, M) p(M), \quad (21)$$

where we marginalize over the 679 models M , explicitly assuming a uniform prior in redshift $p(z|M)$. This probability can be interpreted as an effective likelihood of a unique *pseudo* model, since it is a weighted likelihood over different models and has no explicit redshift prior. This effective *likelihood* \tilde{p} can be used to infer the redshift distribution of a given population, and it is a good approximation when all the $p(f|z, M)$ which contribute significantly are similar, and given that $p(M)$ is close to the real distribution of the target population.

We calibrate $p(M)$ using a subset of the broad-band colours we have available in data and the same colours predicted by each model. The details are given in Appendix D.

3.6 Comparison to previous models

The flux model developed here shares several elements with those implemented in BCNZ2 (used in Eriksen et al. 2019, Er19) and LEPHARE (used in Laigle et al. 2016, COSMOS2015), and it is worth highlighting some differences between them. Regarding the SED

templates, the galaxy continuum templates and dust extinction laws used are the same in all models, but the emission-line templates implementation differs. Er19 and COSMOS2015 create emission-line templates using fixed-line flux ratios between several lines and the O II line, as measured by different spectroscopic surveys, and originally collected in Ilbert et al. (2009). Here, we use a few different values for the line flux ratios of line templates that contain the O II, H α , H β , and O III emission lines (equations 10–12), measured directly from the 66 photometric bands and the best model. In COSMOS2015 the emission-line template was combined with the continuum template at three fixed amplitudes with respect to the continuum template, which were given by the correlation between the UV luminosity and the O II line from Kennicutt (1998). Er19 left the amplitudes of the templates free (with a nonnegativity constraint), and allowed for setups where the O III doublet was an extra template, separated from the other lines. Here, each model contains one emission-line template, and its amplitude is marginalized with a prior that also accounts for the distribution between UV luminosity and the O II line, which we measure from the 66 photometric bands (equation 19).

Finally, in COSMOS2015 one single template which combined continuum and emission-line flux was fitted to the data, and the best-fitting amplitude was used to infer the likelihood at each redshift; in Er19 the combination of several continuum templates (6 to 10) and several emission-line templates (0 to 2) were maximized, and the best-fitting combination was used for the redshift inference; here, we marginalize over the amplitudes of two continuum templates and one emission-line template with priors and compute the Bayesian integral for the redshift inference.

4 RESULTS

In this section, we present the photometric redshift measurements obtained using the model described in Section 3, with the photometry from PAUS and COSMOS described in Section 2. We will compare the redshift estimates with the spectroscopic catalogue described in Section 2.4.

4.1 Photometric redshift precision

We define our photo- z point estimate z_{phot} as the mode of the redshift distribution of each galaxy $\hat{p}(z|f)$ (equation 21). When comparing to the COSMOS2015 photo- z we will use the column PHOTOZ from their public catalogue.

To assess the accuracy and precision of the photo- z point estimates with respect to the spectroscopic point estimates, we consider the distribution of the following quantity: $\Delta_z \equiv (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$. We define two metrics to assess the photo- z precision. One is the central dispersion of the Δ_z distribution, σ_{68} , defined as

$$\sigma_{68} \equiv \frac{P[84] - P[16]}{2}, \quad (22)$$

where $P[x]$ is the value of the distribution Δ_z for the percentile x , which is more robust to outliers than the standard deviation of the distribution. The second metric addressing the photo- z precision is the outlier rate, which is the percentage of outlier galaxies, defined as galaxies that fulfil

$$\text{Outlier} \equiv \frac{|z_{\text{phot}} - z_{\text{spec}}|}{1 + z_{\text{spec}}} > 0.1, \quad (23)$$

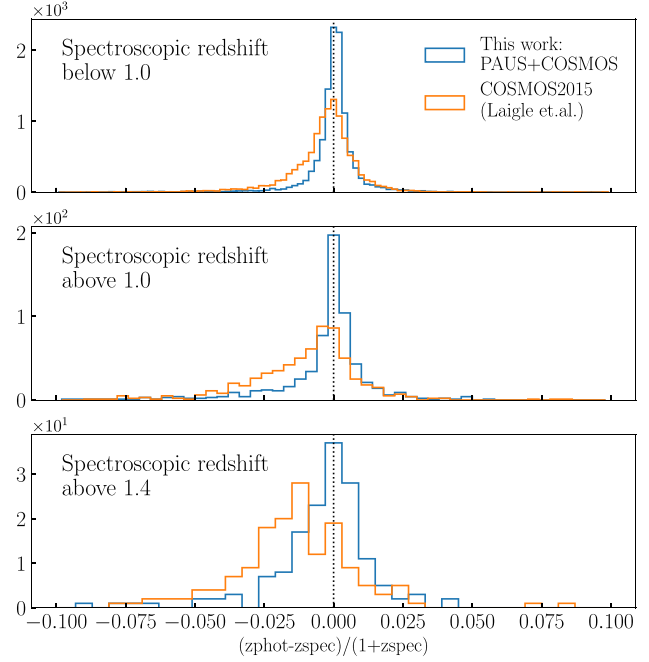


Figure 5. This figure shows the $\Delta_z \equiv (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$ histogram for spectroscopic objects below redshift 1 (top panel), for spectroscopic objects above redshift 1 (middle panel) and spectroscopic objects above redshift 1.4 (bottom panel). The photo- z from this work are used in the orange histograms, while the blue histograms use the photo- z from the public COSMOS2015 catalogue. The histograms clearly show the differences in the photo- z precision, and also hint at a difference in the accuracy, i.e. how well the distributions are centred on 0, which is shown by a vertical dashed line.

which is similar in magnitude to what a galaxy would be considered as an outlier in typical lensing surveys (Hildebrandt et al. 2017; Hoyle et al. 2018).

Fig. 5 visually shows the Δ_z distribution for the range $-0.1 < \Delta_z < 0.1$, highlighting that the Δ_z values from this new catalogue are not only tighter, but also clearly less biased than COSMOS2015, especially for higher redshifts. Fig. 6 shows the σ_{68} (top panel) and outlier rate (bottom panel) for this work’s photo- z from the combination of PAUS and COSMOS2015 photometry (orange lines) as a function of the *Subaru* i_{AB} (each point is a different magnitude bin). We find a $\sigma_{68} = 0.0026$ at bright magnitudes of $i_{\text{AB}} \sim 18$ which increases up to $\sigma_{68} \sim 0.009$ at $i_{\text{AB}} \sim 23$. There is a jump at $i_{\text{AB}} \sim 22.5$, $\sigma_{68} \sim 0.0057$, where the spectroscopic sample completeness changes significantly since the zCOSMOS survey stops at $i_{\text{AB}} = 22.5$.

The quasi spectroscopic precision at bright magnitudes is a common feature of analysis containing PAUS photometry (Eriksen et al. 2019, 2020), since sharp features present in the galaxy spectrum can be precisely identified with the narrow bands, which have an FWHM of 100 Å. For each magnitude bin, we compute the σ_{68} using the photo- z from the COSMOS2015 catalogue (blue lines) for the same galaxies. We find a similar trend with magnitude, with $\sigma_{68} \sim 0.007$ at the bright end, and $\sigma_{68} \sim 0.015$ at the faint end. In comparison, we find this new catalogue yields a redshift precision which is 1.66 times tighter than COSMOS2015 at $i_{\text{AB}} \sim 23$, and up to three times at bright magnitudes. We also find the same jump at $i_{\text{AB}} = 22.5$ with the COSMOS2015 photo- z , which means it is unrelated to the NB photometry from PAUS, and likely a feature of the spectroscopic sample. If the case, one explanation could be

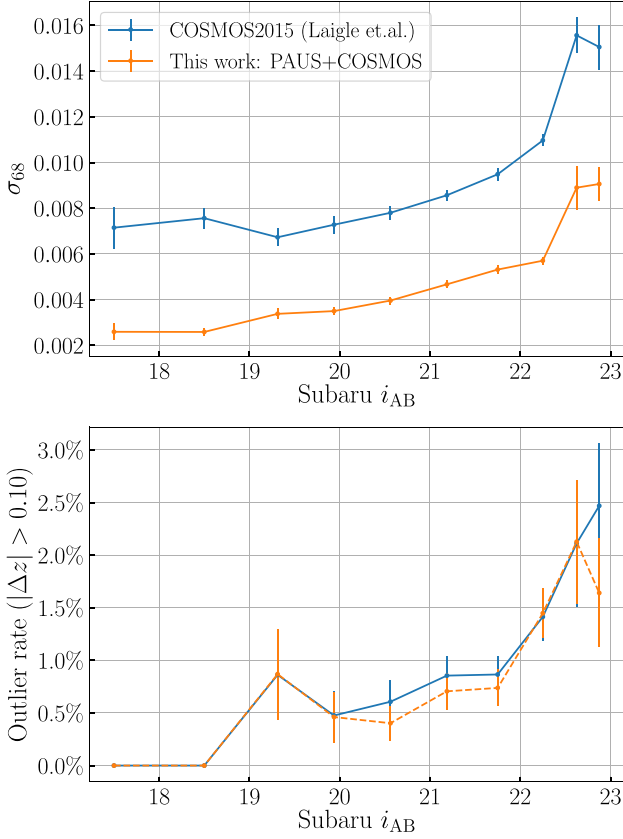


Figure 6. The precision of the photo- z point estimates from this work (orange lines) with respect to the spectroscopic redshift catalog. The top panel shows the σ_{68} (equation 22) of the $\Delta_z \equiv (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$ distribution as a function of i_{AB} magnitude, while the bottom panel shows the percentage of galaxies classified as photo- z outliers, defined as objects that fulfill $|\Delta_z| > 0.1$, also as a function of i_{AB} magnitude. We compute the same statistics using the photo- z estimate from the COSMOS2015 public catalogue (blue lines). We find a significant improvement in the redshift precision (lower σ_{68}) at all magnitudes considered in this work. The error bars are found by computing the dispersion of each metric when bootstrapping the objects in each magnitude bin.

that the spectra is covering the galaxy population differently in the last two magnitude bins, since the completeness is lower (Fig. 2), or the spectroscopic redshifts could be noisier and have a worse performance.

We find a very similar outlier rate for this analysis and COSMOS2015 (bottom panel), which is ~ 1 per cent at $i_{AB} \leq 22.5$, and increases to ~ 2 per cent at $i_{AB} \sim 23$, with slightly smaller values for our work. Fig. 7 shows the scatter plot between spectra and PAUS + COSMOS photo- z (this work, top panel) and the COSMOS2015 photo- z (bottom panel). Visually, both show a very tight concentration along the diagonal, with a small number of outliers seen in both catalogues. We note that a group of outliers confusing the O III and H α lines at $\Delta_z \approx -0.24$ present in the COSMOS2015 catalogue get assigned the correct redshift in our new catalogue. Catastrophic outliers can occur for a number of different reasons, including failure of the model, missing or wrong SED templates, outliers in the photometry, and also outliers in the spectroscopic redshifts measurement. Both this work and the COSMOS2015 analysis share some of the above aspects, which explains why some outliers are present in both catalogues.

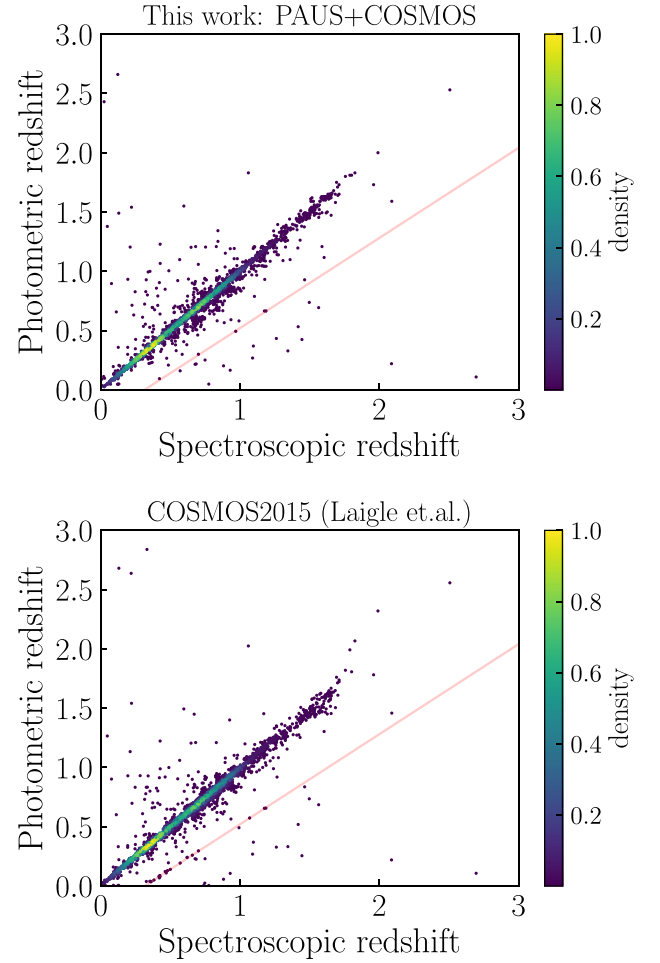


Figure 7. The scatter plot of the spectroscopic redshifts and the photo- z point estimates of this work (top panel) and the photo- z point estimates from COSMOS2015 (bottom panel). Points are coloured according to the proximity (or density) of other nearby objects in this space. We find fewer outliers at $\Delta_z \approx -0.24$ (highlighted with a faint red line in both panels), which are consistent with a confusion between the O III and H α lines, where $\Delta_z \equiv (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$.

4.2 Photometric redshift accuracy

The previous section focused on the width of the Δ_z distribution to assess the average precision of the photo- z point estimates in this catalogue. It is equally important to assess if the photo- z estimates are also statistically unbiased, especially if they have to be used to calibrate the mean redshift of another sample very accurately.

Fig. 8 presents the mean, median and σ_{68} of the Δ_z distribution as a function of Subaru i_{AB} magnitude. Fig. 9 presents the same variables but as a function of the spectroscopic redshift. While the mean is sensitive to outliers, the median should be more robust to them, which is why the bootstrap error bars are smaller in the middle panels in both figures. We find the median of the Δ_z distribution to be consistent within $|\text{median}(\Delta_z)| \leq 0.001$ (which is shown as a grey band) at all redshifts and magnitudes considered in this analysis (orange lines, middle panels), a threshold that we consider good enough to call these distributions statistically unbiased because it is much smaller than the central 68 percent dispersion. In comparison, we find COSMOS2015 (blue lines) to be biased towards lower redshifts with respect to the spectroscopic sample, with a larger bias at higher

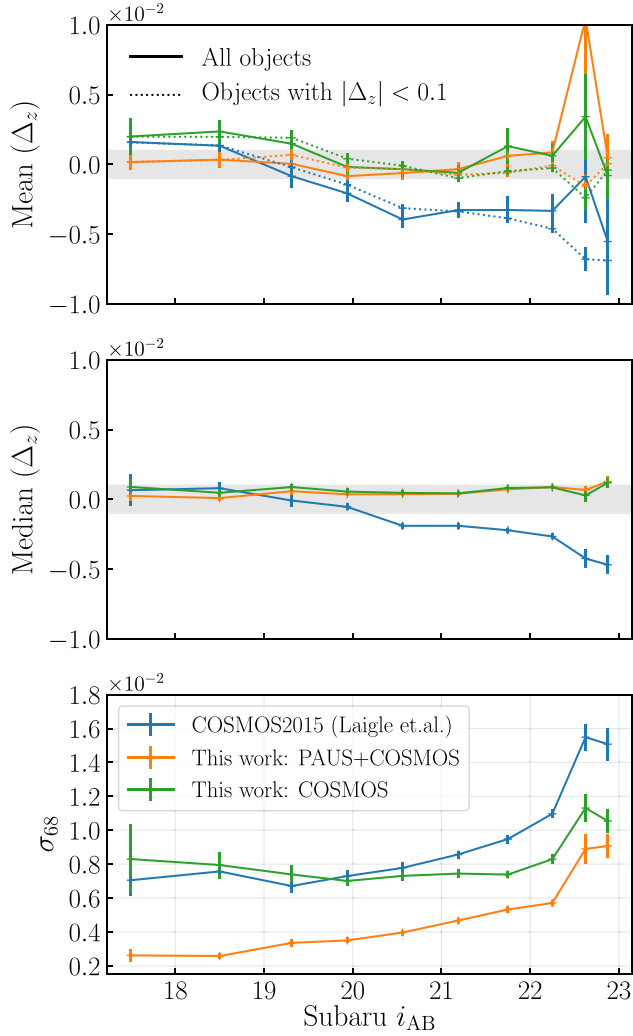


Figure 8. The mean (top panel), median (middle panel) and σ_{68} (bottom panel) of the Δ_z distribution as a function of i_{AB} magnitude. The orange line shows the results from the fiducial calculation in this work, which includes photometry from PAUS and COSMOS surveys, while the green line is an additional calculation using only photometry from the COSMOS survey. For comparison, the blue lines show results using the photo- z from COSMOS2015. We find the mean and median from this work to be statistically unbiased as a function of i_{AB} magnitude, as we find them consistent with $|\text{mean}(\Delta_z)| \leq 0.001$ and $|\text{median}(\Delta_z)| \leq 0.001$, which are shown as a shaded area in the two top panels. The bottom panel is equivalent to the top panel from Fig. 6, but adding the run which drops the PAUS photometry as a green line. The green lines across the panels show that most of the improvement in accuracy and a significant fraction of the improvement in precision at the faint end comes from the new methodology presented in this work, and not from the addition of the PAUS NB data. Most or all of the improvement in precision at $i_{AB} \lesssim 21$ is achieved by including PAUS photometry. In the top panel, the dashed lines show the mean(Δ_z) only for objects with $|\Delta_z| < 0.1$, which shows the impact of the extreme outliers on this metric.

redshift and fainter magnitudes. We find the mean from our work to also be unbiased at different magnitude bins (except for one bin), but we find a biased mean as a function of spectroscopic redshift. We find the catastrophic outlier fraction, objects with $|\Delta_z| > 0.1$, to be mainly responsible for these biases in the mean, as shown by the dashed lines in the top panels of Figs 8 and 9, which show the mean

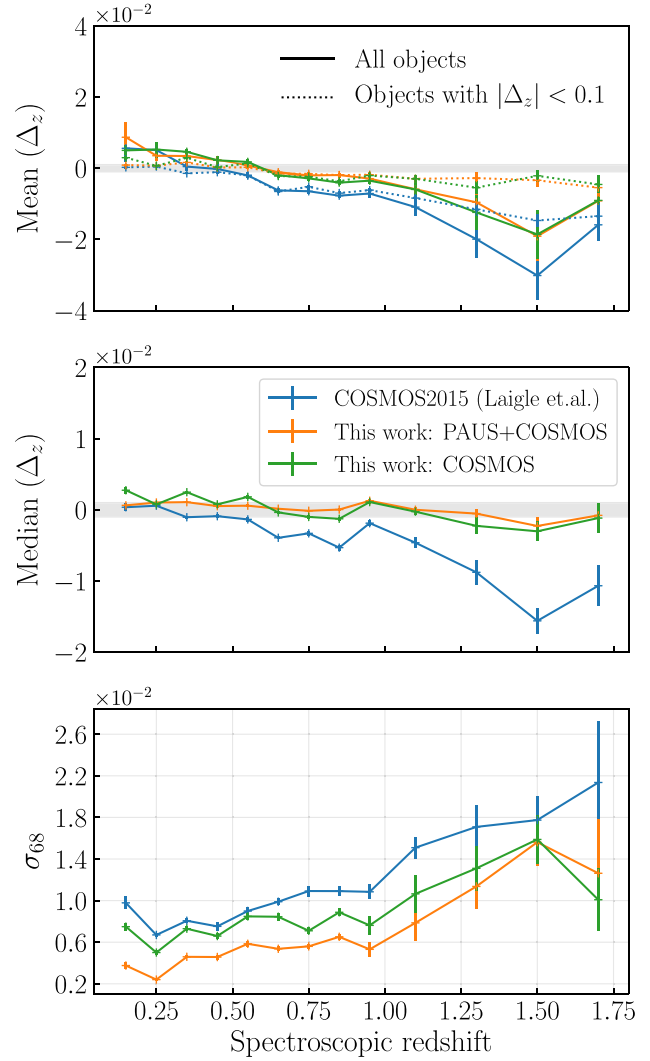


Figure 9. This figure is analogous to Fig. 8, showing the mean, median and σ_{68} as a function of spectroscopic redshift. We find the median for the fiducial run in this work (orange line) to be consistent with $|\text{median}(\Delta_z)| \leq 0.001$ at all redshifts, showing that the centre of the distribution is also unbiased as a function of redshift. The top panel, which shows the mean, shows significantly biased values, towards higher redshifts at low redshift, and towards lower redshift at higher redshifts. The dashed lines show the mean when removing extremely large outliers ($|\Delta_z| > 0.1$), in which case the mean are pulled significantly closer to the 0.

when removing these catastrophic outliers. Fig. 10 shows the mean, median, and σ_{68} as a function of photometric redshift. Similar to Figs 8 and 9 the median is unbiased, consistent with $|\text{median}(\Delta_z)| \leq 0.001$, as a function of photo- z . We find the mean to be more unbiased as a function of photo- z than as a function spec- z , especially when removing the extreme outliers $|\Delta_z| > 0.1$.

A systematic bias in the photo- z estimates can lead to a bias in the cosmological inference of lensing surveys (e.g. see fig. 5 from Salvato et al. 2019). In particular, Joudaki et al. (2019) estimate the DES-Y1 source redshift distributions to have a mean redshift about 0.01 \sim 0.05 lower when estimated using COSMOS2015 instead of spectra. Here, we also find biases towards lower redshift between COSMOS2015 and spectra, although we find them to be lower than $\lesssim 0.01$ in the magnitudes and redshift considered, and these get

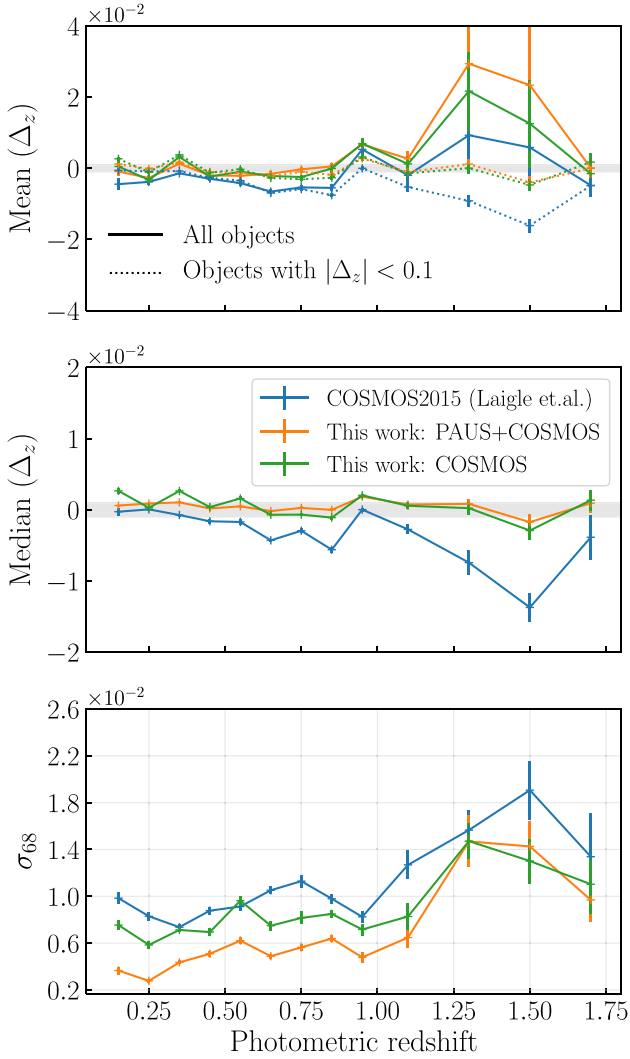


Figure 10. This figure is analogous to Figs 8 and 9, showing the mean, median and σ_{68} as a function of photometric redshift. We find the median for the fiducial run in this work (orange line) to be consistent with $|\text{median}(\Delta_z)| \leq 0.001$ at all redshifts, showing that the centre of the distribution is also unbiased as a function of photo- z . The top panel, which shows the mean, shows more unbiased results than when binning as a function of spectroscopic redshift (Fig. 9). The dashed lines show the mean when removing extremely large outliers ($|\Delta_z| > 0.1$), in which case the mean becomes very consistent with $|\text{mean}(\Delta_z)| \leq 0.001$.

significantly reduced in our new catalogue. Understanding the biases found in Joudaki et al. (2019) (importance of the faint end $i_{AB} > 23$, different spectroscopic samples, $N(z)$ methodology) is beyond the scope of this work.

To understand which improvement comes from adding the NB photometry from PAUS and which comes from the new methodology, we have run the code excluding the PAUS photometry, using only the photometry coming from COSMOS. We show the mean, median and σ_{68} for this study in the green lines in Figs 8–10. We find comparable values for the mean and median between this run and the fiducial run which includes the PAUS photometry, indicating that the new methodology is responsible for obtaining the largely unbiased photo- z estimates. Regarding the improvement in precision, the bottom panel of Fig. 8 shows that part of the improvement in σ_{68}

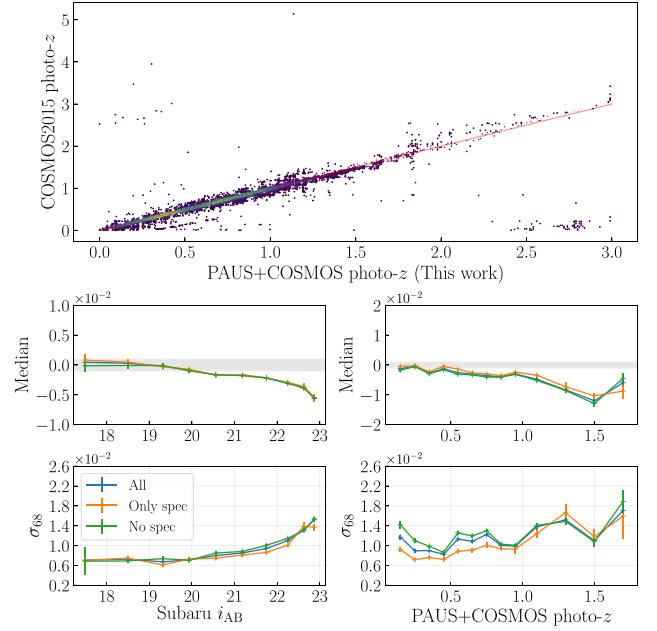


Figure 11. Performance of the complete sample. The top row shows the scatter plot between the photo- z from PAUS + COSMOS and the photo- z from (Laigle et al. 2016). The middle and bottom row show the median and σ_{68} of the quantity $\Delta_z \equiv (z_{\text{COSMOS2015}} - z_{\text{P+C}})/(1 + z_{\text{P+C}})$, where $z_{\text{COSMOS2015}}$ and $z_{\text{P+C}}$ are the photo- z from COSMOS2015 and PAUS + COSMOS, as a function of the *Subaru* i -band (left-hand panels) and the PAUS + COSMOS photo- z (right-hand panels). In the bottom panels, the blue lines show results for all objects in the catalogue, the orange lines for the subset that has spectra, and the green lines for the subset that does not have spectra.

at fainter magnitudes is explained by the new methodology, while most or all of the improvement in precision at $i_{AB} \gtrsim 21$ is achieved by including PAUS photometry.

4.3 Performance of the complete sample

Fig. 11 shows the performance of the full sample, including objects without spectroscopy. We do so via the quantity $\Delta_z \equiv (z_{\text{COSMOS2015}} - z_{\text{P+C}})/(1 + z_{\text{P+C}})$, that compares the photo- z from this work ($z_{\text{P+C}}$) and the photo- z from COSMOS2015. We compare the value of the median and σ_{68} of this quantity as a function of the *Subaru* i -band and the PAUS + COSMOS photo- z for all objects and for two subsets: the one with spectroscopic redshift and the rest. We find the accuracy of the subset without spectra to be very similarly biased towards lower redshift, and to have a similar σ_{68} as a function of magnitude (~ 15 per cent higher), than the subset with spectra. The larger σ_{68} as a function of photo- z is likely explained by the different magnitude distribution, since the subset with spectra has a brighter magnitude distribution (see Fig. 2).

In the top panel, we observe again the outlier stripe at $\Delta_z \approx -0.24$, which shows a discrepancy between both catalogues at deciding between the O III and H_α lines. We also observe an outlier cloud with objects with low photo- z according to COSMOS2015, that we find to be at high- z ($z > 2.3$) in our catalogue. Overlapping spectroscopic redshifts largely agree with our redshift assignment in the O III v H_α outlier stripe, but we do not find enough spectra to validate the existence of the low- z v high- z outlier cloud. However, note this outlier cloud follows the trend where the COSMOS2015 catalogue

systematically assigns redshifts to lower possible values, and a similar cloud was also observed when using LEPHARE in simulated data (Laigle et al. 2019).

5 DISCUSSION AND FUTURE WORK

Our results from Section 4 demonstrate excellent photometric redshift average precision and accuracy across the redshifts and magnitudes considered in this work, which represents an important step towards a deep and complete redshift sample which is at the same time highly accurate and precise in redshift. There are several lines one can take to improve upon these results. One of them is to investigate the origin of the catastrophic outliers ($|\Delta_z| > 0.1$). A rapid inspection by eye shows that a (small) fraction of these outliers could be explained by outlier photometry which has not been flagged during the data reduction process. Typically there are several exposures for the same filter and object, which are later *coadded* either at the image level or later by averaging the flux measured in each exposure. Running the photo- z algorithm directly on flux measurements from individual exposures could help reduce the impact of such photometric outliers. An alternative could be to develop a photometric outlier detection algorithm which finds and removes measurements which are likely to be spurious, but one would need to be careful to not remove real features (i.e. emission lines). Finally, changing the Gaussian likelihood (equation 5) to allow some measurements to be outliers could be an option, but it would require a significant transformation of the algorithm.

The zero-point re-calibration improves the photometric redshift estimation in template-fitting techniques, as has been extensively pointed out in the literature (see Section 3.4). The zero-point factors from this work are shown in Fig. C1, which are larger than the typical error of the measurements (statistical + systematic), and would degrade the photo- z performance if the corrections were ignored. We have found that the iterative algorithm performs well using simulated data, finding that the correct answer is reached after convergence. However, we have found that if the colours are on average biased with respect to the colours in the sample, the iterative zero-point re-calibration can introduce a very significant colour trend. In particular, in earlier stages of this work, when we set the emission-line ratios to values from the literature instead of taking the median measured values from Fig. 3, the iterative algorithm would find a very large colour trend, lifting the flux from shorter wavelengths, which produced a statistically biased photo- z estimation towards lower redshifts.

The zero-point re-calibration essentially compares the measured fluxes to the most likely predicted fluxes from the model, for those objects which have a spectroscopic redshift. Fig. 12 shows the density of residual differences between the measured flux and the most likely model for different broad bands. The differences are weighted by the error, which includes both measured error and the systematic error per band discussed in Section 2.3. Furthermore, we show the density of residual differences for three subsets of the sample: whether the most likely model contains elliptical (orange lines), spiral (green) or starburst (red) templates. We find some interesting trends: while starburst galaxies seem to have unbiased residuals (although the errors appear to be overestimated), elliptical galaxies show a significant offset in the NIR bands, where the models seem to lack flux. There are also some different trends for the spiral galaxies, which use the Prevot extinction law, as opposed to the Calzetti law in the starburst galaxies.

The trends from Fig. 12 could indicate that there are missing templates; that the elliptical templates are wrong in the NIR; that there

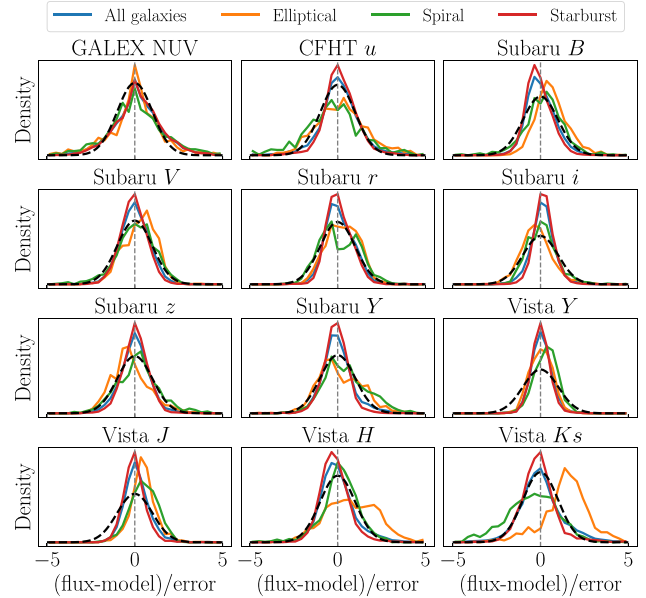


Figure 12. The density of residuals, defined as the difference between the measured flux and the most likely model, divided by the error (which includes both statistical and systematic error). Each panel shows one of the broad bands used in this work, and the residuals are shown for all galaxies and for three populations based on whether galaxies are classified to be elliptical, spiral or starburst according to which is the most likely model. The residuals show some significant offsets for elliptical galaxies and NIR bands, which could indicate a problem with the models at larger wavelengths.

are many more starburst galaxies, which dominate the zero-point recalibration and end up creating problems for elliptical galaxies; or that the Prevot law does not describe well on average the extinction of the spiral templates. We defer to future work the exploration of these possibilities, as well as studying combinations of elliptical templates with starburst templates (since the bulge and disc of a galaxy can have very different SEDs). An improvement of the models and templates would allow to decrease the systematic error per band and lead to a better exploitation of the statistical signal-to-noise.

Another interesting point of discussion is the redshift distribution of individual galaxies, which is required to be correct in many science applications (i.e. to correctly describe the uncertainties from the inference). In practice, this is very rarely achieved because inaccuracies at describing the data will yield the wrong PDF. A common metric in the photo- z literature for testing the quality of the PDF is the probability integral transform (PIT, e.g. Tanaka et al. 2018; Eriksen et al. 2019, 2020; Schmidt et al. 2020), which is defined as the cumulative distribution (CDF) evaluated at the spectroscopic redshift. If the PDFs were statistically correct an ensemble of PIT values should follow a uniform distribution. Fig. 13 shows the PIT values we obtain in this work. The top panel shows the PIT distribution for all galaxies (blue histogram), which shows an excess at both low and high PIT values. All catastrophic outliers ($|\Delta_z| > 0.1$) live at the extremes of the PIT distribution, but also a few more objects with a reasonable photo- z ($|\Delta_z| < 0.01$) have PIT values closer to 0.01 to either PIT=0 or PIT=1 than they should. The top panel also shows the PIT distributions for elliptical, spiral and starburst galaxies, revealing that although the distribution for all galaxies looks relatively flat, there are some clear trends for different populations: red galaxies seem to have a PDF biased low in redshift, while blue galaxies seem to be biased high, even if the photo- z (the

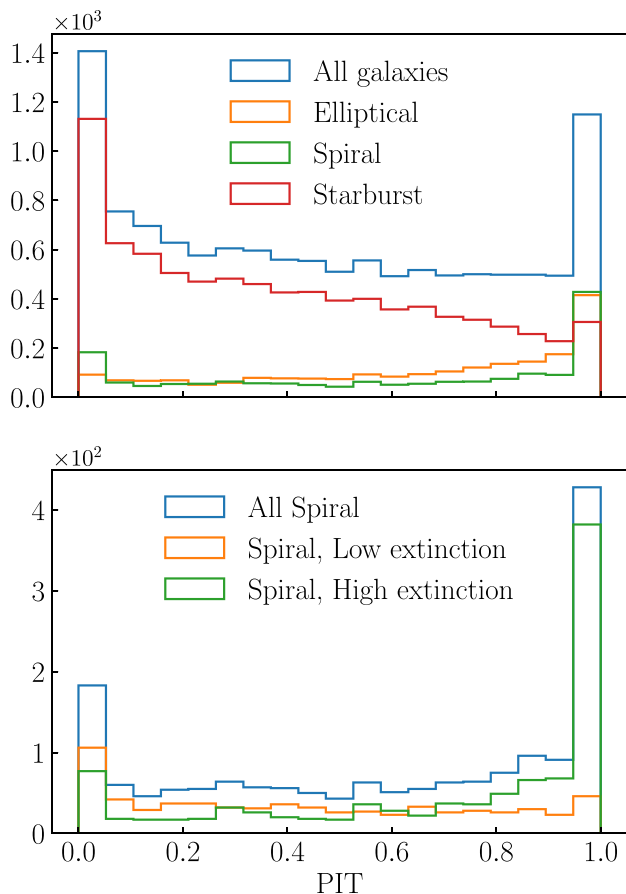


Figure 13. The distribution of the PIT values, defined as the cumulative distribution evaluated at the value of the spectroscopic redshift. The top panel shows the distribution for all objects with spectra, and for elliptical, spiral, and starburst galaxies. The bottom panel show the PIT for spiral galaxies, and for spiral galaxies with a lower or higher internal dust extinction value.

mode of the pdf) is still very close to the spectroscopic redshift for all of them. Therefore, the PDFs can be deemed unreliable from a statistical point of view, which is a common feature in many photo- z algorithms (e.g. see fig. 2 from Schmidt et al. 2020). However, it is important to point out that the redshift errors implied by the PIT distributions are smaller when the PDFs are narrow, as is the case for the PDFs we obtain. For example, we find $\text{median}(\Delta_z) = 0.0012$ for starburst galaxies and $\text{median}(\Delta_z) = -0.0016$ for spiral and elliptical galaxies, which means that the errors in the PDFs seen in the PIT distribution do not translate into large redshift inaccuracies, as the PDFs are indeed narrow in redshift.

The PDFs could be improved by extending the template baseline used in this work, using synthetic stellar population synthesis models to generate more adequate templates (e.g. Conroy, Gunn & White 2009; Conroy & Gunn 2010; Chaves-Montero & Hearin 2020). We also find different PIT distribution for low or high extinction spiral galaxies (bottom panel), which could hint at a problem on how additional dust extinction is being modelled in spiral galaxies. One could also try to combine with AGN templates, or to add additional extinction to elliptical galaxies. A natural extension of the algorithm would be to allow more emission lines to be free, including adequate priors for different emission lines, instead of choosing some fixed values. The implicit redshift priors included in the priors for the continuum and emission-line templates can also yield the wrong PIT

distribution. We leave for a future study a detailed examination of these points.

Finally, PAUS has collected deeper data for a fraction of the COSMOS field, which will be a natural extension of this work towards improved redshift calibration samples for fainter objects.

6 SUMMARY AND CONCLUSIONS

We have presented a novel method to estimate photometric redshifts with multiband photometric data. We have applied the method to data from the COSMOS field and we have assessed its performance comparing to available public spectroscopic redshifts. We have used a combination of 40 NB photometric filters from the PAUS survey and 26 broad, intermediate and NB filters from the COSMOS survey (Laigle et al. 2016, COSMOS2015) to estimate the most precise and accurate photometric redshifts available in the COSMOS field for objects with $i_{AB} \leq 23$. We have developed a new algorithm that models the galaxy SED using a linear combination of continuum and emission-line templates and integrates over their possible different combinations using priors. The main primary results and conclusions are:

- (i) We find a redshift precision of $\sigma_{68}(\Delta_z) \sim 0.0057, 0.009$ at magnitude $i_{AB} \sim 22.5, 23$, respectively, which is over a factor $1.66 \times$ tighter than previous results from COSMOS2015, with $\Delta_z \equiv (z_{\text{photo}} - z_{\text{spec}})/(1 + z_{\text{spec}})$. The precision gets much better at bright magnitudes, where we find $\sigma_{68} \sim 0.0026$ at $i_{AB} \sim 18$ (see Fig. 6).
- (ii) We find the redshifts to be statistically unbiased, with the median of the Δ_z distribution consistent within $|\text{median}(\Delta_z)| \leq 0.001$ at all redshifts and magnitudes considered in this analysis (see Figs 5, 8, 9, 10).
- (iii) We measure different emission lines using the 66 photometric bands and a subsample with spectroscopic redshifts (see Figs 3, 4). We use these measurements to build the emission-line templates and calibrate a model to balance between the emission-line and the continuum templates. This yields a galaxy model with more accurate colours, with produces a statistically unbiased redshift inference.
- (iv) We make the redshift catalogue publicly available through the COSMOHUB platform (see Appendix A for details of the catalogue and how to download it).

The results from our work indicate that deeper PAUS data in the COSMOS field will significantly improve the performance and redshift precision of this photo- z sample and potentially yield a redshift sample with $\sigma_{68}(\Delta_z) < 0.01$ at $i_{AB} > 23$.

ACKNOWLEDGEMENTS

AA would like to thank Mara Salvato for providing a compilation of public spectroscopic redshifts in the COSMOS field. AA would like to thank Clotilde Laigle and Olivier Ilbert for providing information on SED templates, filter transmission curves and filter zero points at early stages of this work. AA thanks Jonas Chaves-Montero, Andrew Hearin and Nesar Ramachandra for useful discussions on this work. Argonne National Laboratory's work was supported by the US Department of Energy, Office of High Energy Physics. Argonne, a U.S. Department of Energy Office of Science Laboratory, is operated by UChicago Argonne LLC under contract no. DE-AC02-06CH11357. The PAU Survey is partially supported by MINECO under grants CSD2007-00060, AYA2015-71825, ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, MDM-2015-0509 and Juan de la Cierva fellowship and LACEGAL and EWC Marie Skłodowska-Curie grant No 734374

and no. 776247 with ERDF funds from the EU Horizon 2020 Programme, some of which include ERDF funds from the European Union. IIEC and IFAE are partially funded by the CERCA and Beatric de Pinos program of the Generalitat de Catalunya. Funding for PAUS has also been provided by Durham University (via the ERC StG DEGAS-259586), ETH Zurich, Leiden University (via ERC StG ADULT-279396 and Netherlands Organisation for Scientific Research (NWO) Vici grant 639.043.512), University College London and from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 776247 EWC. The PAU data centre is hosted by the Port d'Informació Científica (PIC), maintained through a collaboration of CIEMAT and IFAE, with additional support from Universitat Autònoma de Barcelona and ERDF. We acknowledge the PIC services department team for their support and fruitful discussions. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776247. H. Hildebrandt is supported by a Heisenberg grant of the Deutsche Forschungsgemeinschaft (Hi 1495/5-1) as well as an ERC Consolidator Grant (No. 770935). H. Hoekstra acknowledges support from Vici grant 639.043.512 from the Netherlands Organization for Scientific Research (NWO). GM acknowledges support from ST/P006744/1. CMB acknowledges support from ST/P000541/1 and ST/T000244/1. PN acknowledges support from ST/P000541/1 and ST/T000244/1. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Maria Skłodowska-Curie (grant agreement No 754510), the National Science Centre of Poland (grant UMO-2016/23/N/ST9/02963) and by the Spanish Ministry of Science and Innovation through Juan de la Cierva-formation program (reference FJC2018-038792-I).

DATA AVAILABILITY

The data underlying this paper available at

(i) Redshift catalog: Redshift information for 40672 galaxies is publicly available at <https://cosmohub.pic.es/home> under Catalogs labelled as PAUS + COSMOS photo-z catalog.

(ii) PAUS images used in this work are available from the ING Observatory and also from the authors on request. The PAUS photometry used in this work is preliminary and is being used in other applications. We expect to make it available in upcoming publications.

REFERENCES

- Abbott T. M. C. et al., 2018a, *Phys. Rev. D*, 98, 043526
 Abbott T. M. C. et al., 2018b, *ApJS*, 239, 18
 Aihara H. et al., 2018, *PASJ*, 70, S8
 Alarcon A., Sánchez C., Bernstein G. M., Gaztañaga E., 2019, *MNRAS*, 498, 2614
 Arnouts S., Ilbert O., 2011, Astrophysics Source Code Library, record ascl:1108.009
 Asgari M. et al., 2019, *A&A*, 624, A134
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
 Benítez N., 2000, *ApJ*, 536, 571
 Benjamin J. et al., 2013, *MNRAS*, 431, 1547
 Bertin E., 2011, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, ASP Conf. Ser., Vol. 422, Astronomical Data Analysis Software and Systems XX, Astron. Soc. Pac., San Francisco, p. 435
 Bonnett C. et al., 2016, *Phys. Rev. D*, 94, 042005
 Brammer G. B., van Dokkum P. G., Coppi P., 2008, *ApJ*, 686, 1503
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
 Buchs R. et al., 2019, *MNRAS*, 489, 820
 Cabayol L. et al., 2019, *MNRAS*, 483, 529
 Calabrò A. et al., 2018, *ApJ*, 862, L22
 Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, *ApJ*, 429, 582
 Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
 Carretero J. et al., 2017, Proceedings of the European Physical Society Conference on High Energy Physics., Venice, p. 488
 Casey C. M. et al., 2017, *ApJ*, 840, 101
 Chaves-Montero J., Hearin A., 2020, *MNRAS*, 495, 2088
 Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, *AJ*, 132, 926
 Colless M., 1999, *Phil. Trans. R. Soc. A*, 357, 105
 Conroy C., Gunn J. E., 2010, Astrophysics Source Code Library, record ascl:1010.043
 Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
 Cunha C. E., Huterer D., Busha M. T., Wechsler R. H., 2012, *MNRAS*, 423, 909
 Dahlen T. et al., 2013, *ApJ*, 775, 93
 Davis C. et al., 2017, preprint ([arXiv:1710.02517](https://arxiv.org/abs/1710.02517))
 De Vicente J., Sánchez E., Sevilla-Noarbe I., 2016, *MNRAS*, 459, 3078
 Eriksen M. et al., 2019, *MNRAS*, 484, 4200
 Eriksen M. et al., 2020, *MNRAS*, 497, 4565
 Faisst A. L., Masters D., Wang Y., Merson A., Capak P., Malhotra S., Rhoads J. E., 2018, *ApJ*, 855, 132
 Fitzpatrick E. L., Massa D., 1986, *ApJ*, 307, 286
 Fitzpatrick E. L., Massa D., 2007, *ApJ*, 663, 320
 Gaia Collaboration et al., 2016, *A&A*, 595, A2
 Gatti M. et al., 2018, *MNRAS*, 477, 1664
 Genz A., 1992, *J. Comput. Graph. Stat.*, 1, 141
 Ghahramani Z., 2012, *Phil. Trans. R. Soc. A*, 371, 20110553
 Gruen D., Brimiouille F., 2017, *MNRAS*, 468, 769
 Hartley W. G. et al., 2020, preprint ([arXiv:2003.10454](https://arxiv.org/abs/2003.10454))
 Hasinger G. et al., 2018, *ApJ*, 858, 77
 Hikage C. et al., 2019, *PASJ*, 71, 43
 Hildebrandt H. et al., 2012, *MNRAS*, 421, 2355
 Hildebrandt H. et al., 2017, *MNRAS*, 465, 1454
 Hildebrandt H. et al., 2020, *A&A*, 633, A69
 Hogg D. W., Bovy J., Lang D., 2010, *MNRAS*, 496, 4769
 Hoyle B. et al., 2018, *MNRAS*, 478, 592
 Huterer D., Takada M., Bernstein G., Jain B., 2006, *MNRAS*, 366, 101
 Huterer D., Cunha C. E., Fang W., 2013, *MNRAS*, 432, 2945
 Ilbert O. et al., 2009, *ApJ*, 690, 1236
 Ilbert O. et al., 2013, *A&A*, 556, A55
 Joudaki S. et al., 2017, *MNRAS*, 465, 2033
 Joudaki S. et al., 2019, *A&A*, 638, 8
 Kashino D. et al., 2019, *ApJS*, 241, 10
 Kelly B. C., 2007, *ApJ*, 665, 1489
 Kennicutt Robert C. J., 1998, *ARA&A*, 36, 189
 Kriek M. et al., 2015, *ApJS*, 218, 15
 Laigle C. et al., 2016, *ApJS*, 224, 24
 Laigle C. et al., 2019, *MNRAS*, 486, 5104
 Laureijs R. et al., 2011, preprint ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))
 Lee K.-G. et al., 2018, *ApJS*, 237, 31
 Leistedt B., Mortlock D. J., Peiris H. V., 2016, *MNRAS*, 460, 4258
 Lilly S. J. et al., 2009, *ApJS*, 184, 218
 LSST Dark Energy Science Collaboration, 2012, preprint ([arXiv:1211.0310](https://arxiv.org/abs/1211.0310))
 Madau P., 1995, *ApJ*, 441, 18
 Marchesi S. et al., 2016, *ApJ*, 830, 100
 Masters D. et al., 2015, *ApJ*, 813, 53
 Masters D. C., Stern D. K., Cohen J. G., Capak P. L., Rhodes J. D., Castander F. J., Paltani S., 2017, *ApJ*, 841, 111
 Masters D. C. et al., 2019, *ApJ*, 877, 81
 Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013, preprint ([arXiv:1303.4722](https://arxiv.org/abs/1303.4722))
 Molino A. et al., 2014, *MNRAS*, 441, 2891
 Moustakas J., Kennicutt Robert C. J., Tremonti C. A., 2006, *ApJ*, 642, 775
 Newman J. A., 2008, *ApJ*, 684, 88
 Padilla C. et al., 2019, *AJ*, 157, 246

- Paulino-Afonso A., Sobral D., Darvish B., Ribeiro B., Stroe A., Best P., Afonso J., Matsuda Y., 2018, *A&A*, 620, A186
- Pérez-González P. G. et al., 2013, *ApJ*, 762, 46
- Pickles A. J., 1998, *PASP*, 110, 863
- Planck Collaboration et al., 2016, *A&A*, 594, A10
- Planck Collaboration et al., 2018, *A&A*, 641, 67
- Polletta M. et al., 2007, *ApJ*, 663, 81
- Prevot M. L., Lequeux J., Maurice E., Prevot L., Rocca-Volmerange B., 1984, *A&A*, 132, 389
- Puglisi A. et al., 2016, *A&A*, 586, A83
- Rosani G., Caminha G. B., Caputi K. I., Deshmukh S., 2019, *A&A*, 633, A159
- Saito S., de la Torre S., Ilbert O., Dubois C., Yabe K., Coupon J., 2020, *MNRAS*, 494, 199
- Salvato M. et al., 2009, *ApJ*, 690, 1250
- Salvato M. et al., 2011, *ApJ*, 742, 61
- Salvato M., Ilbert O., Hoyle B., 2019, *Nature Astronomy*, 3, 212
- Sánchez C., Bernstein G. M., 2019, *MNRAS*, 483, 2801
- Sánchez C., Raveri M., Alarcon A., Bernstein G. M., 2020, *MNRAS*, 498, 2984
- Sargent M. T. et al., 2007, *ApJS*, 172, 434
- Schmidt S. J., Ménard B., Scranton R., Morrison C., McBride C. K., 2013, *MNRAS*, 431, 3307
- Schmidt S. J. et al., 2020, *MNRAS*, 621, 1587
- Scoville N. et al., 2007, *ApJS*, 172, 1
- Simm T. et al., 2015, *A&A*, 584, A106
- Smith J. A. et al., 2002, *AJ*, 123, 2121
- Speagle J. S. et al., 2019, *MNRAS*, 490, 5658
- Stecher T. P., Donn B., 1965, *ApJ*, 142, 1681
- Storey P. J., Hummer D. G., 1995, *MNRAS*, 272, 41
- Storey P. J., Zeppen C. J., 2000, *MNRAS*, 312, 813
- Tallada P. et al., 2020, *Astron. Comput.*, 32, 100391
- Tanaka M., 2015, *ApJ*, 801, 20
- Tanaka M. et al., 2018, *PASJ*, 70, S9
- Tonello N. et al., 2019, *Astron. Comput.*, 27, 171
- Troxel M. A. et al., 2018, *Phys. Rev. D*, 98, 043528
- van Dokkum P. G., 2001, *PASP*, 113, 1420
- Wright A. H., Hildebrandt H., van den Busch J. L., Heymans C., 2019a, *A&A*, 637, 20
- Wright A. H. et al., 2019b, *A&A*, 632, A34
- Wright A. H., Hildebrandt H., van den Busch J. L., Heymans C., Joachimi B., Kannawadi A., Kuijken K., 2020, *A&A*, 640, 10
- Xiang F. Y., Li A., Zhong J. X., 2011, *ApJ*, 733, 91

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: REDSHIFT CATALOGUE DETAILS

We make the catalogue publicly available through the COSMOHUB platform: <https://cosmohub.pic.es> (Carretero et al. 2017; Tallada et al. 2020). After registering, the catalogue is available under Catalogs labelled as PAUS + COSMOS photo-z catalog. We provide one catalogue with a list of summary properties for each galaxy, which include the best photometric redshift and best model information and a few absolute magnitudes and line flux measurements derived at the best photo-z and best model. Table A1 shows a summary of the columns provided in the catalogue, and a description of each column. A second file containing the redshift distributions (equation 21) for all galaxies can be directly downloaded from the Value Added Data section. The redshift distribution is evaluated between redshift 0 and 3: in steps of 0.001 between redshift 0 and 1, in steps of 0.002 between redshift 1 and 1.5, and in steps of 0.01 between redshift 1.5 and 3.0. Note the different redshift step sizes, especially when integrating over the redshift distribution.

Table A1. Description of the columns provided in the photo-z catalogue.

Column name	Description
ref_id	PAUdm reference ID.
I_auto	Auto <i>i</i> -band magnitude from Ilbert et al. (2009) cosmos photo-z catalogue.
photoz	Photo-z, defined as the mode of the $p(z)$ of each object obtained from the templates.
ra	Right ascension.
dec	Declination.
nbands	Number of bands used in the photo-z code.
id_laigle	COSMOS2015 (Laigle et al. 2016) reference ID.
zspec_mean	Average spectroscopic redshift from several public redshift surveys (when available).
zspec_std	Standard deviation spectroscopic redshift from several public redshift surveys (when available).
best_model	Best model in the photo-z code.
best_extlaw	Best extinction law in the photo-z code (0: No extinction; 1: Prevot; 2: Calzetti; 3: Calzetti+Bump1; 4: Calzetti + Bump2).
best_continuum	Best continuum template group ([0,16] see Table 1).
best_ebv	Best extinction $E(B - V)$ value (from 0 to 0.5, in steps of 0.05). See equation (9).
best_type	Best galaxy type, based on <code>best_continuum</code> (0: red ([0,12]), 1: green ([13,14]), 2: blue ([15,16])).
MUV	Absolute magnitude <i>GALEX</i> NUV band. Corrected for internal galaxy extinction.
MU	Absolute magnitude CFHT <i>u</i> band. Corrected for internal galaxy extinction.
MR	Absolute magnitude <i>Subaru</i> <i>r</i> band. Corrected for internal galaxy extinction.
MI	Absolute magnitude <i>Subaru</i> <i>i</i> band. Corrected for internal galaxy extinction.
flux_OII	Flux for the O II line template in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. Corrected for internal galaxy extinction.
flux_Hbeta	Flux for the $H\beta$ line template in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. Corrected for internal galaxy extinction.
flux_OIII	Flux for the O III line template in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. Corrected for internal galaxy extinction.
flux_Halpha	Flux for the $H\alpha$ line template in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. Corrected for internal galaxy extinction.
flux_err_OII	Flux error for the O II line template in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. Corrected for internal galaxy extinction.
flux_err_Hbeta	Flux error for the $H\beta$ line template in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. Corrected for internal galaxy extinction.
flux_err_OIII	Flux error for the O III line template in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. Corrected for internal galaxy extinction.
flux_err_Halpha	Flux error for the $H\alpha$ line template in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. Corrected for internal galaxy extinction.
line_coverage_OII	Maximum coverage of the O II flux line. Range is [0,1], with 0 being poorly covered and 1 great coverage. Measurements with poor coverage are unreliable and should be flagged.
line_coverage_Hbeta	Maximum coverage of the $H\beta$ flux line. Range is [0,1], with 0 being poorly covered and 1 great coverage. Measurements with poor coverage are unreliable and should be flagged.
line_coverage_OIII	Maximum coverage of the O III flux line. Range is [0,1], with 0 being poorly covered and 1 great coverage. Measurements with poor coverage are unreliable and should be flagged.
line_coverage_Halpha	Maximum coverage of the $H\alpha$ flux line. Range is [0,1], with 0 being poorly covered and 1 great coverage. Measurements with poor coverage are unreliable and should be flagged.

APPENDIX B: COMBINING HETEROGENEOUS PHOTOMETRY

Here, we describe the method to produce a synthetic broad-band flux from an overlapping set of NB filters. This synthetic flux can be used to estimate a factor that calibrates heterogeneous photometry (see Section 2.3).

Let $W_B(\lambda) \equiv \lambda^{-1} T_B(\lambda)$ be a broad-band filter that we want to express as a linear combination of 40 NB filters $W_N(i, \lambda) \equiv \lambda^{-1} T_N(\lambda)$, where $T(\lambda)$ are the filter responses (in units of photon) and $i = 1, \dots, 40$. One can find a solution by writing the coefficients $c(i)$ such that

$$W_B(\lambda) = \sum_{i=1}^{i=40} c(i) W_N(i, \lambda), \quad (\text{B1})$$

where λ is the wavelength which we will bin in integer values. We define

$$\begin{aligned} \langle W_B W_B \rangle &\equiv \int d\lambda W_B^2(\lambda) = \sum_{\lambda} W_B^2(\lambda) = 1, \\ \langle W_N W_N \rangle &\equiv \int d\lambda W_N^2(\lambda) = \sum_{\lambda} W_N^2(\lambda) = 1, \end{aligned} \quad (\text{B2})$$

which indicates the transmission curve's norm is normalized to unity. We can then define 40 elements of broad-narrow projection vector:

$$BN(j) \equiv \langle W_B W_N(j) \rangle = \int d\lambda W_B(\lambda) W_N(j, \lambda) \quad (\text{B3})$$

and a 40 x 40 narrow band overlap matrix

$$NN(i, j) \equiv \langle W_N(i) W_N(j) \rangle = \int d\lambda W_N(i, \lambda) W_N(j, \lambda). \quad (\text{B4})$$

If we multiply equation (B1) by $W_N(j, \lambda)$ and then integrate we get

$$BN(j) = \sum_i c(i) NN(i, j). \quad (\text{B5})$$

As $NN(i, j)$ is invertible, the unique solution is

$$\begin{aligned} c(i) &= \sum_j NN^{-1}(i, j) BN(j) \\ &= \sum_j \int d\lambda W_B(\lambda) W_N(j, \lambda) \left[\int d\lambda W_N(i, \lambda) W_N(j, \lambda) \right]^{-1} \end{aligned} \quad (\text{B6})$$

We can use the $c(i)$ values from equation (B6) to build a synthetic broad-band flux from 40 NBs fluxes as

$$f_{\text{syn}}^{BB} \equiv \sum_{i=1}^{i=40} c(i) f_i^{NB}. \quad (\text{B7})$$

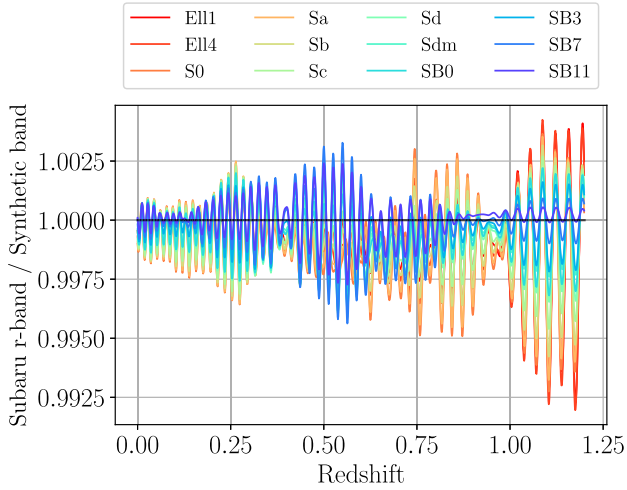


Figure B1. Ratio between the *Subaru* r -band true flux and the synthetic flux obtained from the NB fluxes for different SED models (colour coded) at different redshifts (x -axis). The deviations from unity can be predicted and included in the SED modelling.

Comparing the synthetic *Subaru* r band with the measured broad-band we estimate a factor that we can apply to bring two different photometric systems together.

Due to the particular shape and small overlap between different NB filters, the synthetic broad-band flux will slightly differ from the true one. This effect can be predicted and included in the modelling of the galaxy SED. Fig. B1 shows the ratio between the synthetic broad-band flux and the true broad-band flux for different continuum SEDs at different redshifts, showing some oscillations around 1. Additionally, when NB measurements are missing, we extrapolate from the remaining narrow bands.

APPENDIX C: SYSTEMATIC OFFSETS

Here, we show the systematic offsets described in Section 3.4. Table C1 and Fig. C1 show the value of the offsets we obtain for each of the bands we use in this work.

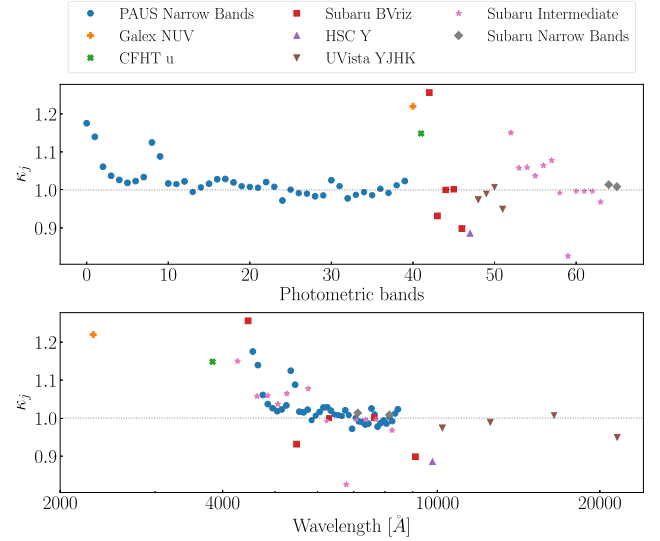


Figure C1. Systematic flux factors (values from Table C1) colour coded in blocks of bands. The bottom panel shows the zero-point of each band at the mean wavelength of the band.

Table C1. Systematic flux factors, κ_j in equation (20), (the flux version of systematic magnitude offsets) measured in this work to make the colours predicted by the models and the ones measured in data more similar. A subsample of objects with spectroscopic redshift available is used (see Section 3.4 for more details).

NB455 \rightarrow 1.175	NB465 \rightarrow 1.140	NB475 \rightarrow 1.061
NB485 \rightarrow 1.037	NB495 \rightarrow 1.026	NB505 \rightarrow 1.018
NB515 \rightarrow 1.023	NB525 \rightarrow 1.034	NB535 \rightarrow 1.125
NB545 \rightarrow 1.088	NB555 \rightarrow 1.017	NB565 \rightarrow 1.015
NB575 \rightarrow 1.022	NB585 \rightarrow 0.994	NB595 \rightarrow 1.007
NB605 \rightarrow 1.016	NB615 \rightarrow 1.028	NB625 \rightarrow 1.029
NB635 \rightarrow 1.020	NB645 \rightarrow 1.010	NB655 \rightarrow 1.008
NB665 \rightarrow 1.006	NB675 \rightarrow 1.021	NB685 \rightarrow 1.008
NB695 \rightarrow 0.972	NB705 \rightarrow 1.001	NB715 \rightarrow 0.991
NB725 \rightarrow 0.990	NB735 \rightarrow 0.983	NB745 \rightarrow 0.985
NB755 \rightarrow 1.025	NB765 \rightarrow 1.010	NB775 \rightarrow 0.978
NB785 \rightarrow 0.987	NB795 \rightarrow 0.994	NB805 \rightarrow 0.986
NB815 \rightarrow 1.003	NB825 \rightarrow 0.992	NB835 \rightarrow 1.012
NB845 \rightarrow 1.023	GALEX NUV \rightarrow 1.22	CFHT u \rightarrow 1.148
Subaru B \rightarrow 1.256	Subaru V \rightarrow 0.932	Subaru r \rightarrow 1.000
Subaru i \rightarrow 1.001	Subaru z \rightarrow 0.898	HSC y \rightarrow 0.886
UVista Y \rightarrow 0.974	UVista J \rightarrow 0.989	UVista H \rightarrow 1.007
UVista K \rightarrow 0.950	Sub IA427 \rightarrow 1.150	Sub IA464 \rightarrow 1.058
Sub IA484 \rightarrow 1.059	Sub IA505 \rightarrow 1.037	Sub IA527 \rightarrow 1.064
Sub IA574 \rightarrow 1.078	Sub IA624 \rightarrow 0.992	Sub IA679 \rightarrow 0.826
Sub IA709 \rightarrow 0.996	Sub IA738 \rightarrow 0.996	Sub IA767 \rightarrow 0.996
Sub IA827 \rightarrow 0.968	Sub NB711 \rightarrow 1.014	Sub NB711 \rightarrow 1.009

APPENDIX D: MODEL PRIOR

This section describes how we compute the population prior on the models $p(M)$ described in Section 3.5 and equation (21). We will use 10 broad bands: the CFHT u band, the *Subaru* B , V , r , i , z bands and the UltraVista Y , J , H , K bands. These 10 bands constitute a nine-dimensional colour space \mathbf{C} . We write

$$p(M) = \int p(M, z, \mathbf{C}) dz d\mathbf{C} = \int p(M, z|\mathbf{C})p(\mathbf{C}) dz d\mathbf{C}, \quad (\text{D1})$$

where $p(\mathbf{C})$ is the distribution of colours, which we estimate from the observed colour distribution in the data. Using Bayes theorem

$$p(M, z|\mathbf{C}) \propto p(\mathbf{C}|M, z)p(M, z). \quad (\text{D2})$$

Note $p(M, z)$ is unknown, so we assume here it is a uniform distribution, and we update this with the colour space information $p(\mathbf{C})$ that we observe. We obtain $p(\mathbf{C}|M, z)$ by integrating over the colours each model can produce at each redshift

$$\begin{aligned} p(\mathbf{C}|M, z) &= \int p(\mathbf{C}, \boldsymbol{\alpha}|M, z) d\boldsymbol{\alpha}, \\ &= \int p(\mathbf{C}|\boldsymbol{\alpha}, M, z)p(\boldsymbol{\alpha}|M, z) d\boldsymbol{\alpha}, \end{aligned} \quad (\text{D3})$$

where the colour $p(\mathbf{C}|\boldsymbol{\alpha}, M, z)$ is completely specified for each $(\boldsymbol{\alpha}, M, z)$, and $p(\boldsymbol{\alpha}|M, z)$ is the prior from equation (6).

In practice, we discretize the colour space \mathbf{C} using a self-organizing map (or SOM), which is an unsupervised machine-learning technique which projects data from a high-dimensional space into a lower dimensional grid. We refer to each unit of the grid as an SOM cell. The process preserves the topology of the higher dimensional data, which means that objects that were close in the original space will also be closer in the lower dimensional grid. We will use the same algorithm described in Masters et al. (2015), which we implement to produce a two-dimensional SOM.

We estimate $p(M, z|\mathbf{C})$ by randomly drawing values of $(\boldsymbol{\alpha}, M, z)$ from $p(\boldsymbol{\alpha}, M, z)$ and assigning them to an SOM cell. The distribution $p(\mathbf{C})$ is estimated from the number counts of observed colours in each SOM cell. Finally, we sum over cells and redshift using equation (D1). One can further refine this prior by allowing it to also depend on the observed *Subaru* i band magnitude m_i of each galaxy

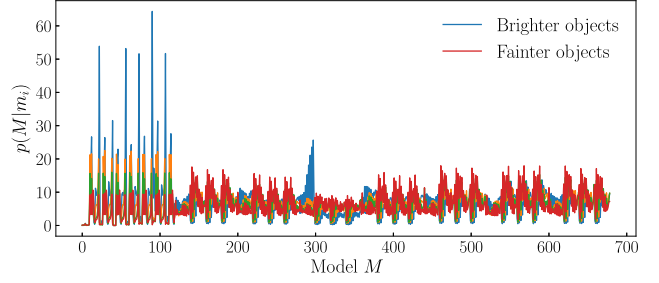


Figure D1. The prior for each model M used in this work for each magnitude bin, $p(M|m_i)$. There are four magnitude bins with equal number of galaxies for which $p(M|m_i)$ is computed, which appear colour coded from brighter to fainter as blue, orange, green, and red. The prior is computed using the observed colours of the data, and the colours spanned by each model M . See Appendix D and Section 3.5 for details.

as

$$\begin{aligned} p(M|m_i) &= \int p(M, z, \mathbf{C}|m_i) dz d\mathbf{C}, \\ &= \int p(M, z|\mathbf{C}, m_i)p(\mathbf{C}|m_i) dz d\mathbf{C}, \\ &\approx \int p(M, z|\mathbf{C})p(\mathbf{C}|m_i) dz d\mathbf{C}, \end{aligned} \quad (\text{D4})$$

In practice, the only change between equations (D1) and (D4) is $p(\mathbf{C}) \rightarrow p(\mathbf{C}|m_i)$. We compute $p(\mathbf{C}|m_i)$ by binning galaxies in four magnitude bins with equal numbers, and linearly interpolating the values of those four bins to any value of m_i .

Fig. D1 shows the model prior used in this work for the four magnitude bins, $p(M|m_i)$. There are a total of 679 models in this work, featuring a number of different continuum templates groups, extinction laws, and emission-line templates. The figure shows the models ordered following Table 1, putting models with the same continuum group but different emission-line models or extinction law together.

This paper has been typeset from a \LaTeX file prepared by the author.