



# The many flavours of photometric redshifts

Mara Salvato, Olivier Ilbert, Ben Hoyle

## ► To cite this version:

Mara Salvato, Olivier Ilbert, Ben Hoyle. The many flavours of photometric redshifts. *Nature Astronomy*, 2019, 3, pp.212-222. 10.1038/s41550-018-0478-0 . insu-03666742

**HAL Id: insu-03666742**

**<https://insu.hal.science/insu-03666742>**

Submitted on 28 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The hundred flavours of photometric redshifts

Mara Salvato,<sup>\*</sup> Olivier Ilbert,<sup>†</sup> and Ben Hoyle<sup>‡</sup>

(Dated: June 6, 2018)

For more than seventy years, the measurements of fluxes of galaxies at different wavelengths and derived colours have been used to estimate their corresponding cosmological distances. From the fields of galaxy and AGN evolution to precision cosmology, the number of scientific projects relying on such distance measurements, called photometric redshifts, have exploded. The benefits of photometric redshifts is that all sources detected in photometric images can have some distance estimates relatively cheaply. The major drawback is that these cheap estimates have a low precision when compared with the resource-expensive spectroscopy. The methodology to estimate redshifts has been through several major revolutions throughout the last decades, triggered by increasingly more stringent requirements on the photometric redshift accuracy. Here, we review the various techniques to obtain photometric redshifts, from template-fitting to machine learning and hybrid systems. We also describe the state-of-the-art results on current extra-galactic samples and explain how survey strategy choices impact redshift accuracy. We close the review with a description of the photometric redshifts efforts planned for upcoming wide field surveys, which will collect data on billions of galaxies, aiming to solve the most exciting cosmological and astrophysical questions of today.

## I. INTRODUCTION

The distance of an extra-galactic source needs to be calculated before any more meaningful physical quantities can be inferred. The primary observable to measure a distance of an object is its electromagnetic spectral energy distribution (hereafter SED), which is composed of a continuum and emission/absorption lines. The expansion of the Universe stretches the SED toward longer wavelengths by a factor  $1 + z$  with  $z$  being the redshift. The observational difficulty in distance estimation consists of identifying a pair of characteristic features in the SED and measure the amount they have been stretched. Then, the measured redshift is related to a proper distance, assuming a cosmological model. Emission and absorption lines are sharp features which can be easily identified in the SED. In addition, two well known features shape the SED continuum (see Fig. 1): 1) the *Balmer break* below 4000Å, which is explained by the absorption of photons more energetic than the Balmer limit at 3646Å and the combination of numerous absorption lines by ionised metals in stellar atmospheres; 2) the *Lyman break* below 1216Å, explained by the absorption of light below the Lyman limit at 912Å and the absorption by the intergalactic medium along the line of sight.

When SEDs of sufficient wavelength resolution are available (i.e. spectra), the emission/absorption lines can be identified and the redshift precision can be measured to better than  $10^{-3}$  for a resolving power  $R = \lambda/\Delta\lambda > 200$  [e.g., 1]. Despite the efficiency of the current generation of multi-object spectrographs (MOS; e.g., VLT/VIMOS, KECK/DEIMOS, SUBARU/FMOS) we

can only obtain meaningful spectra for a few percent of the sources detected in deep imaging surveys even on 8-m class telescopes [e.g., 1]. A minimum of two well identified spectral features are required to obtain a robust redshift measurement. Given the MOS limited spectral coverage and the limited signal-to-noise in spectra for faint objects, the success rate of measuring spectroscopic redshifts (hereafter spec- $z$ ) can be lower than 50-70% in deep spectroscopic surveys [e.g., 2].

Alternatively, by measuring the flux of an object in broader filters, we can obtain a sparse sampling of the SED sufficient to constrain the continuum shape, define the extra-galactic nature of the sources and estimate the redshift based on broad features like the Lyman and Balmer breaks, or strong emission and absorption lines. Such a low resolution distance estimate is called a "photometric redshift" (hereafter photo- $z$ ). This principle was first applied by [3] to measure the photo- $z$  of elliptical galaxies at  $z \sim 0.4$ . A modern version of the photo- $z$  method has been published by [4] who was the first to apply a template-fitting procedure to radio galaxies.

The main advantage of using photo- $z$  is that we can derive a distance measurement for all sources identified in an imaging survey. The price paid is the lower redshift precision which is typically a factor 10-100 times worse than that obtained with a low resolution spectrograph [e.g., 5]. Any photo- $z$ -based study needs an accurate assessment of the photo- $z$  performance, which depends on the image properties (e.g. depth, wavelength coverage) and whether the sources are galaxies, active galaxy nuclei (AGN) or stars. *Assessing photo- $z$  performance requires deep and representative spectroscopic sample, which highlights the complementarity of photometric and spectroscopic redshift surveys.*

The photo- $z$  method has been adopted as a common tool to estimate galaxy distances, and a simple ADS/NASA search shows that the fraction of refereed

<sup>\*</sup> MPE; mara@mpe.mpg.de

<sup>†</sup> LAM; olivier.ilbert@lam.fr

<sup>‡</sup> USM,MPE; benhoyle1212@gmail.com

publications including the term "photometric redshift" has increased by a factor ten in the last two decades. This can be partly explained by the increasing number of multi-wavelength surveys. In addition, the confidence on the method has grown thanks to reassuring good comparison with spec-z's for an always increasing number of galaxies as faint as the limit of the imaging surveys. Finally, several mature and well-tested photo-z codes are both publicly available and well documented, making the technique readily accessible.

The quality of the photo-z defines the range of their possible scientific applications. Photo-z are used to study the formation and evolution of galaxies, allowing a statistical analysis of larger samples than those allowed by spec-z, also in those regions of the colour-magnitude space difficult to populate with spectroscopy. Some of the recent applications heavily relying on accurate photo-z are: the cosmic time evolution of galaxy properties [e.g., 6] or the search of primordial galaxies [e.g., 7]; the study of the relation between galaxy properties and their dark matter halos [e.g., 8] or cluster identifications [e.g., 9]. Galaxy environment or galaxy pairs evolution can also be investigated with photo-z [e.g., 10–12], although with a limited precision [e.g., 13, 14]. Similarly, photo-z are used for studying the evolution of galaxies hosting an Active Galactic Nucleus (AGN) [e.g., 15] or more peculiar objects [e.g., Blazars: 16].

Photo-z are also becoming a major tool to study properties of dark energy. The weak lensing tomography approach [19] has become one of the main cosmological probes in current and future cosmological experiments such as the Dark Energy Survey (DES), High Suprime Camera survey (HSC), Euclid, and the Large synoptic survey telescope (LSST) (see VI). For these applications, the required photo-z precision is not as stringent as for studying galaxy evolution, but several other challenges need to be faced. Among those are the amount of information to be processed; the need for homogeneous photo-z performances over thousands of  $\text{deg}^2$ ; and the need to characterise the redshift distribution of a population with extreme precision.

In this review we provide an overview on the ground principles of the photo-z technique, with details on the most common methods and with an eye on current and future surveys and scientific application challenging the current state-of-art.

## II. METHODS FOR ESTIMATING PHOTO-Z

We use Fig.1 to introduce how the photo-z problem is tractable. In panel a), we show key features such as the Lyman or Balmer breaks in galaxy SEDs can be diverse. The photo-z technique relies on the capacity to isolate the wavelength position of these red-shifted features. The breaks correspond to a rapid increase of the flux continuum from the blue to the red part of the SED. The photo-z technique aims to detect gradients between

observed fluxes in adjacent filters, which would reveal the presence of a break. Panel b) of Fig.1 illustrates this key principle. The  $i - z$  color varies with redshift and reaches a maximum at  $z = 1.1$  when the Balmer break falls between the  $i$  and  $z$  bands. *Thus, the first rule when designing a photometric survey is that the filter set should be chosen to encompass key features at the redshift range of interest.*

However, the same  $i - z$  color could correspond to multiple redshifts, implying degeneracies in the redshift solution. Such degeneracies can be broken by combining several colours. *A second important rule is that the multi-wavelength coverage should be as broad as possible to limit the risk of photo-z degeneracies* [e.g. 20].

The flow-chart of Fig. 2 shows the fundamental steps and ingredients used in all photo-z techniques. At the core there is always a model of the mapping between various colours (or fluxes) and redshift. The redshift solution and its associated Probability Distribution Function (PDF) are obtained by comparing this mapping and the observed fluxes of the studied source. In the case of template-fitting methods, the redshift-colours mapping is based on physical knowledge built by scientists over time. With Machine Learning methods, the mapping is obtained every time anew, using a representative 'training sample' of galaxies with both photometry and known redshifts. For both methods the results can be improved by using additional priors (see further in this section for more details).

While the basic principle is simple, the mapping and its application to the data can be established in many different ways. In this section, we highlight the most commonly used methods.

### A. Physically motivated methods

In this branch of photo-z techniques, the mapping between flux and redshift is predicted for extragalactic sources taking into account physical processes regulating the observed light emission.

Thus, the most basic ingredient is the definition of a set of SED templates, which could be either obtained from theory or observations. Theoretical templates are generated using stellar population synthesis models [e.g., 21–24] which progressed significantly in the last decades. Such theoretical templates rely on numerous assumptions, for instance on the galaxy star formation histories. Empirical templates are extracted from observed spectra [e.g., 25, 26], extended over the entire wavelength range with models [e.g., 27, 28]. Not only the type of templates determines the quality of photo-z, but also their optimal coverage of the colour-redshift space [e.g., 29].

Nebular emission lines emitted by HII regions should also be considered in the templates [e.g. 30–32]. [5] have clearly shown how considering the contribution of the nebular emission improve on the accuracy of photo-z by a factor of 2.5. The importance of this component is eas-

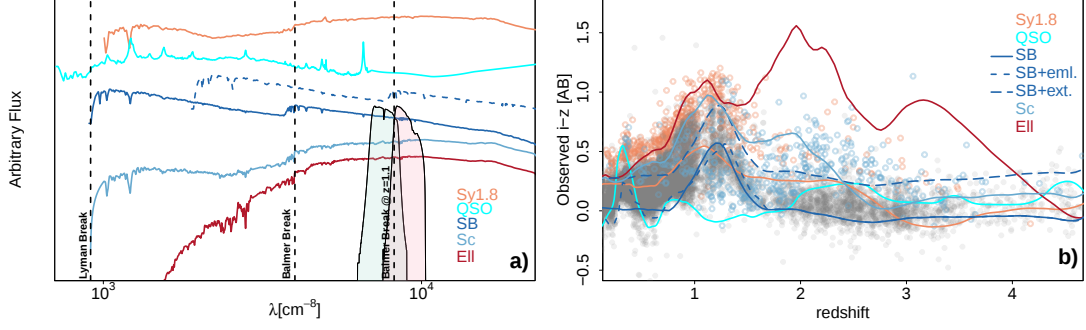


FIG. 1: **The colour-redshift mapping principle.** **a):** Examples of SEDs for various type of galaxies (elliptical:Ell; Starburst:SB; spiral with small bulge: Sc) and AGN (luminous quasar: QSO; low luminosity, obscured AGN: Sy1.8 (Seyfert 1.8)). The Lyman and Balmer breaks, among the key features in determining the redshifts, at rest-frame are indicated by vertical dashed lines. One template and the Balmer break are also plotted at redshift 1.1. For clarity, the transmission curves of  $i$  and  $z$  filters, covering the wavelength range between 700 and 1100 nm are also indicated. **b):**  $(i - z)$  color as a function of redshift. Galaxies with reliable spectroscopic redshifts are represented with grey dots while AGN obscured/un-obscured by dust are represented with red/blue small dots. The solid lines represent the expected redshift evolution of the  $(i - z)$  color for the templates presented in the left panel, without any extinction. The star-burst galaxy (in blue) is also shown i) considering extinction (long-dashed line) and ii) considering the contribution from emission lines (short-dashed line).

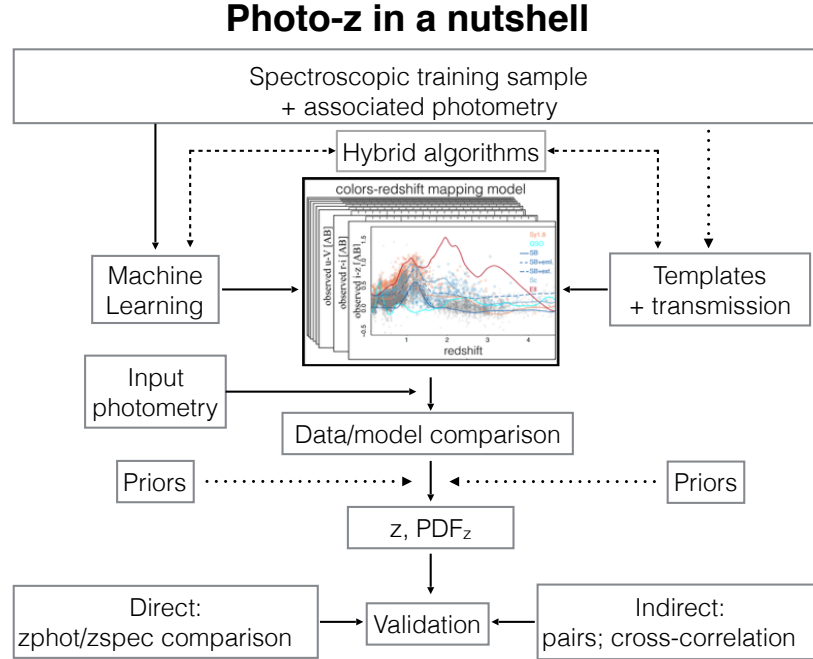


FIG. 2: **Photometric redshift technique in a nutshell.** Flowchart representing the various steps involved in computing photometric redshift. The core of ML, template-fitting and hybrid methods is the multi-dimensional colour-redshift relation. The spectroscopic sample, essential for ML, is optional (although desirable for some tuning) for template-fitting techniques. With the accumulated knowledge over the past decade we are able to define reliable priors that are improving dramatically our results. After the computation of photo-z and PDF, a huge effort goes into the validation of the results.

ily understood looking at the panel b) of Fig. 1. There, the SED of a starburst galaxy is plotted with and without a contribution from emission lines, with the former better matching the colour-redshift map of spectroscopically confirmed galaxies.

To obtain accurate photo- $z$  at  $z > 1$  it is particularly important to model the effect of dust attenuation, because optical data will sample the rest-frame UV part of the SED which is most affected. The dust grains present in the interstellar medium absorb and scatter the stellar light, and reddens the SED continuum. Depending on the amount of dust and its grain properties, the impact can be severe (see panel b of Fig. 1, comparing long-dashed and continuous lines). When computing photo- $z$  with template-fitting codes the dust attenuation is typically modelled as a free parameter, using one of several dust attenuation laws [e.g., 33, 34].

Then, the templates are corrected for the absorption of the galaxy light that get absorbed when crossing the neutral gas present in the inter-galactic medium [e.g. 35, 36].

The light is also attenuated by the dust from our Milky Way along the optical path between the source and the observer. Normally, catalogues are corrected for these effects using the dust extinction maps of [37], but a better treatment is proposed in [38] which accounts for variations in the intrinsic SED of galaxies.

Finally, the redshifted templates should be integrated through the filter transmission curves to produce the modelled fluxes. All other factors that could modify the light distribution should be considered, such as the impact of Earth atmosphere, the optics of the telescope, the shape and efficiency of the filter curves, the Charged Coupled Device (CCD) transmission. These various components are usually integrated into one single transmission curve per filter and stored within the code libraries.

As a final note, while this section discussed the modelling of extra-galactic sources, template-fitting techniques also need to consider stellar templates, because we can not rule out the Galactic nature of the studied sources. Typically, the fit to the stellar templates is performed independently and the extragalactic/Galactic nature of the source is decided a posteriori.

## B. Data driven methods

Rather than relying upon any physical models one may estimate photo- $z$  using data driven approaches. The most common method is Machine Learning (hereafter ML). ML algorithms can be categorised into unsupervised and supervised learning, the difference being that unsupervised learning requires only photometry, while supervised methods require both photometry and reliable spec- $z$  for a data sample during the training of the algorithm. In the photo- $z$  framework, these training samples allow the algorithm to *learn* the mapping between colour and redshift (see e.g, Fig.1).

In general all supervised ML methods perform function approximation, and for photo- $z$  this entails finding a function which maps between the multi-dimensional photometry space (e.g., fluxes or colours) and the redshift values of the training sample. Indeed the function is assumed to be “well behaved” between the multi-dimensional locations of the training data, and thus supervised ML methods can be viewed as performing high dimensional interpolation. Once this function has been learned, the photometry of a source with an unknown redshift is localised in the multi-dimensional photometry space and paired with a corresponding redshift value of probability distribution. Due to the interpolative nature of supervised ML approaches, one must ensure that the space of photometric properties of the sample for which predictions will be made, is well sampled by the training examples [see, e.g., 39–41]. Failing this, the algorithms may perform extrapolation and lose accuracy. Furthermore, the complexity of the (unknown) manifold to be approximated using ML provides an estimate of the number density of training data in each region of color space that are required.

Two popular supervised ML systems are often used for photo- $z$  computation, Random Forests [42] and Neural Networks [43] which we briefly describe below.

Random forests [e.g., 44, 45] are aptly named because they are constructed from collections of ‘decision trees’ and each tree is grown using a random subset of the training data. Decision trees are a set of logical if-statements that group the properties of the training sample into cells, which are defined to minimise the spec- $z$  dispersion of the data in each cell. Successively, each cell is assigned the average value of the spec- $z$ . Once all the trees have been grown, predictions are made on the redshift of the sources in a test sample by passing their photometric properties through each tree and averaging over all of the obtained redshifts. The number of trees or the maximum number of cells are some of the variables of the algorithm.

Neural Networks are a class of popular ML algorithms that were inspired by how neurons in the animal brain process information and learn. The NN algorithms [e.g., 40, 46, 47, 49, 50] make increasingly complex non-linear matrix transformations of the input properties. The structure consists of different layers, and each layer can be viewed as a new transformation of those which come before it. The output of the NN is defined by the user, and represents the target value of interest, such as the redshift value. The NN algorithm tunes the matrix transformations using the training data, such that the output value minimises the residuals between true redshift and predicted redshift.

Finally, Deep Machine Learning (DML), the current state of the art in computer science, has been also applied to problem of photo- $z$  estimation in [e.g. 51, 52]. DML is based on normal neural networks, but with many thousands of neurons in each hidden layer. In particular DML may also work directly from galaxy images. On many square degrees surveys this means avoiding the

alternative of computing photo- $z$  gathering photometry from catalogs of different surveys, often using different measures.

Given that supervised ML methods require training data with both photometry and spectroscopic redshifts to approximate the mapping between colours and redshift, they are of course limited by the nature of the training data, often limited to low redshift and bright luminosity (see Fig. 3). Given the dearth of both high redshift (e.g.,  $z > 1$ ) and very faint galaxies, machine learning methods should be used with caution in these regimes.

Unlike supervised ML, unsupervised ML algorithms does not use spec- $z$  in the training phase, but rather performs clustering in the input data space to identify groups of similar objects. The most popular example in the photo- $z$  literature is the Self Organising Map (SOM). A nice introduction to SOMs can be found in [e.g., 53, 54].

All ML algorithms used for photo- $z$  are able to generate PDFs of varying sophistication. The simplest method is to randomly sampling from values and errors of each input parameter many times, and associate the normalised distribution of predictions as a PDF [e.g., 45, 55].

Other exploratory techniques from ML have also been ported to photo- $z$  estimation, such as anomaly detection, data augmentation, and feature importance [56–58]. In turn, these techniques allow ML methods to identify problematic training data, provide some extrapolation ability outside of the original training sample, and motivate the choice of input parameters that provide the most photo- $z$  predictive power. As in the case of template-fitting methods, ML can also improve on the results by considering additional information as a prior, often in the form of re-weighting the training data [59].

### C. Using additional data as prior information.

Many photo- $z$  codes, often produce a redshift  $z$  which is the maximum of the likelihood  $\mathcal{L}(z|T, D)$  given the data  $D$  and galaxy type  $T$  or choice of SED. This redshift inference can be improved by adding other prior information based on the empirical knowledge that we have acquired over time of the redshift evolution of various galaxy properties.

When the prior is considered in a Bayesian context, the posterior redshift prediction is the product of the multi-dimensional likelihood and the prior which is then marginalised over, to produce a one dimensional posterior in redshift. Such a Bayesian approach has been established in BPZ [60] by introducing a prior on the redshift distribution per galaxy spectral type and magnitude bin. Other choices of prior have also been recently made, e.g. as luminosity function for the GOODZ code [61] and mass-SFR relation in [62]. One of the most recent public code is BEAGLE [29], which is among the few that adopt a fully Bayesian approach and thus a prior can be assigned to each parameter. Prior should be used with care given the risk of introducing an inaccurate information

(e.g. the redshift distribution beyond the spectroscopic limits is hardly known). Still, in the case of sparse wavelength coverage which produce several redshift peaks in the PDFs, [60] showed that using the most probable solution based on prior knowledge helps in reducing the fraction of catastrophic failures for the average population, at a price of associating the wrong photo- $z$  to specific populations (as high- $z$  sources, or AGN). Independently on the type of prior, we believe it is a responsibility of the authors to describe in detail the priors that are adopted and to discuss the implication that the prior may have on the photo- $z$  results.

### D. Introducing spatial information

The correlation of source positions can also be used as a redshift estimation technique. This method uses the fact that galaxies are not distributed randomly in the Universe but reside in large scale structures. By using the spec- $z$  of a reference sample, the redshift distribution of an unknown sample can be found by maximising the spatial cross-correlation signal between the unknown and the reference samples. As early as [63], galaxy clustering was used to extract information on QSO distances. This method has been mainly revived in the last decade to estimate the mean redshift of photo- $z$  selected samples for weak lensing applications [e.g. 64]. While considering usually only the large-scale clustering, [65] generalise the method using also clustering on smaller scales. However, such approach requires the unknown sample to be preselected in a narrow redshift range, in photo- $z$  or in colours [e.g. 66]. By considering extremely narrow sample of preselected galaxies, it becomes even possible to measure individual redshifts [67].

[68] propose a new estimate of the photo- $z$  using the cosmic web. Such method requires a dense spectroscopic coverage to characterise accurately void regions, filaments, walls and clusters. The redshift is obtained from the product between the photo- $z$  derived from the colours, the cosmic web and the density field. One of the limitation of the method is the need of using accurate photo- $z$  (better than 1%) from colours in order to avoid a catastrophic association between the photo- $z$  and the wrong cosmic web structure. For 3% of the SDSS multicolor catalogue, [68] were able to apply this method, and show that an accuracy as good as  $10^{-5} - 10^{-4}$  was possible.

## III. WHICH METHOD FOR WHICH SURVEY

But how to decide between different methods? Is template-fitting better than ML? Or vice versa? Or, is there among the SED-fitting and ML codes a favoured one? In general, a direct and fair comparison between performances of different template-fitting and ML codes is possible only within tests specifically designed for it

[e.g., 39, 41, 69–71]. Here we list the most used and/or public codes, together with their key features, pointing to the papers that attempted a more quantitative comparison of the various algorithms. More on this is also discussed in Sec. VI.

### A. Template-fitting codes

Numerous template-fitting codes are publicly available, with Hyperz [72] being a precursor. While the ultimate common goal of all the codes is to compute photo- $z$ , each code developed their own specificity. For example, template libraries change from code to code. EAZY [73] combines basic templates and creates new ones on-the-fly, while ZEBRA [74] trains the templates using a spectroscopic sample (a risk, when the spectroscopic sample is not representative of the entire population [see Fig. 11 and related text in 75]). Some codes include dust attenuation in their templates [e.g. BPZ, 60] while others include dust as a free parameter (e.g. Hyperz). No papers have yet to clearly highlight that one template set is superior to another. Still, EAZY include an error function associated to the templates which improve the quality of the PDF.

Depending on the dataset and the scientific objective, some features are crucial to improve the photo- $z$  accuracy. It is now widely accepted that galaxy photo- $z$  performances improves, especially at high redshift, when including emission lines [31] in the templates, like in Le Phare [30, 76] or EAZY, especially when the photometry from narrow and intermediate bands are used like in COSMOS [5, 78], MUSYC [79], SHARDS [80] and ALHAMBRA [81].

Another example of adapting the code configuration to the photometric data available is the use of prior. [62] showed that for a faint survey at  $i < 25$  and with only *griz* bands, he was able to decrease the fraction of outliers from 45% to 25%. But if the survey provides an extensive multi-wavelength coverage including NIR, adding a prior is usually not considered, meaning that an equal probability is associated to each value of the parameter. Another important aspect is the sensitivity of the photo- $z$  bias to the absolute flux calibration of a survey: [30] showed that only a few percent uncertainty in the absolute calibration of the photometry could create a bias at the same order as the photo- $z$  precision for even bright sources. This calibration issue can be corrected using a spec- $z$  training sample, with a overall improvement in accuracy by a factor of 2. The zero-point calibration was also pointed out in [70], as a way to reduce the fraction of outliers and to increase the accuracy.

### B. ML codes

The choice about which ML algorithm will provide the best performance for a particular dataset or problem is

nontrivial. Results from computer science show that if the training sample is large enough, and the training time of each algorithm is long enough, then the performance from different algorithms converges [see e.g., 41, 57, for a photo- $z$  setting]. If the training samples are small, or the training times are limited, then the performance of different codes depends on the complexity of the underlying surface to be learned. Given the plethora of ML algorithms on the market, each group typically considers only a few algorithms for any given problem, often chosen from those of which they have obtained good results in the past. Furthermore those algorithms which are finally implemented in large scale astronomical surveys, such as the TPZ [45], Skynet [50] and aNNz2 [83] methods for DES, are often those in which the code authors were part of the collaboration, and had a vested interest in seeing their codes advanced. A good starting point for deciding which ML algorithms should be applied to new datasets, is to use those which are very fast to implement so that any subsequent performances can be easily bench-marked. For this purpose the authors would suggest starting with decision tree based algorithms, such as Random Forests.

### C. SED-fitting or ML?

Generally, the choice between template-fitting and ML methods strongly depends on the scientific application and on the spec- $z$  sample available for the training. For instance, ML applied to SDSS data outperform template-fitting for low-redshift sources. This is because the relatively few optical bands are still sufficient to accurately correlate with redshift, and the spectroscopic sample used for the training is rich and complete. But in deep pencil-beam surveys covering a large redshift range, the reverse case is shown in [39]. In other words, the use of ML is limited to those surveys with a sufficient training set, while if the scientific objective is to study galaxy population with limited spectroscopic coverage, template-fitting code should be favoured.

Template-fitting codes could also model uncertainties in the absolute calibration, but ML methods are also insensitive to photometry biases depending on magnitude or colours. Therefore if the goal is to limit biases over a large field well covered with spectroscopy, ML would normally expected to be the favoured approach. In term of speed, ML clearly outperform the template-fitting techniques [47]. One of the largest benefits of ML algorithms come from their optimisation to work smoothly on massive data sets. They have been specifically developed to process massive volumes of data, and can easily accommodate current and future astronomical sized data sets [see e.g., 47].

## IV. VALIDATION AND STATE OF ART

### 1. Photo-z performance and validation

An extensive characterisation of the photo-z performance should be made for every catalog and sub-sample of interest. It is typical to compare a photo-z point prediction (such as the mean or mode of the PDF,  $z_{\text{phot}}$ ) and the spec-z ( $z_{\text{spec}}$ ) as a way to assess performance. Note that any spec-z which are used for ML training, or prior construction, should generally be excluded from the validation sample.

The following three measures are often adopted to assess photo-z performances:

- Bias: defined as  $\langle z_{\text{phot}} - z_{\text{spec}} \rangle$ , it characterises the average separation between the predicted and the true redshift [e.g., 84].
- Outliers (or catastrophic failures) fraction: usually defined as the fraction of sources for which  $|z_{\text{phot}} - z_{\text{spec}}| > N\sigma$  [e.g., 70], or  $|z_{\text{phot}} - z_{\text{spec}}|/(1 + z_{\text{spec}}) > 0.15$  [e.g., 39]. It highlights the fraction of sources with unexpectedly large errors.
- Precision ( $\sigma_{z_p}$ ): is often defined as the standard deviation of  $(z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$  or as  $1.48 \times \text{median}(|z_{\text{phot}} - z_{\text{spec}}|/(1 + z_{\text{spec}}))$  [30], the latter being less sensitive to outliers. The precision describes expected scatter between predictions and truths.

These metrics are not unique [e.g., see discussion in Sec.4 in 70] and they could be science-dependent. However, because they are the most used, they allow for a easy comparison between surveys.

We note that photo-z performance measured using a direct comparison between photo-z and spec-z is less meaningful if the spectroscopic coverage is not fully representative of the entire photo-z sample, e.g., if the spec-z sample is biased toward bright sources [see e.g. 41, 85].

When this happens, a statistical mapping between the photometric parameter space covered by the two samples should be defined [e.g. 54]. For regions of photometric space without spectroscopic coverage, alternative methods have been developed to establish the photo-z performances. For example, the galaxy closed pairs technique [78, 86], uses the fact that close pairs have a significant probability of being associated and that they therefore should have similar redshifts. The comparison of the spatial cross-correlation between two redshift bins with that expected from a model may also be used to measure redshift precision and contamination [87].

Characterising the performance of the full photo-z PDF is slowly becoming popular, as it can highlight the lack of templates, training sample, or sub-optimal redshift prediction routines. One usual problem affecting PDFs is caused by under-estimated photometric uncertainties, creating a PDF peak which is too narrow [e.g.

70]. In numerous works authors validate the PDFs by insuring that 68% of the spec-z falls within the  $1\sigma$  uncertainties derived from the PDF. In [84] the authors introduce the statistical method of the Continuous Rank Probability Score (CRPS) to the photo-z community as a means of testing PDF performance. The CRPS asserts that the value of the cumulative PDF at  $z_{\text{spec}-z}$  must be uniform for an ensemble of galaxies.

### 2. State of the art

The wavelength coverage (see panel a) of Fig. 4) and the quality of the input data determines the accuracy of the photo-z, independently of the merits of each redshift code. One of the main advantages of photo-z is the ability to estimate distance information for faint sources, granted that the photometric errors associated to the measures are sufficiently small to constrain the redshift by limiting the degeneracy in the solutions. Hence, because the photometric errors are larger when the sources in an image are faint, the only way to keep small the photometric errors (and thus keep high the accuracy of photo-z) is to obtain images as deep as possible (see Fig.3). The photo-z estimates in CANDELS represent one of the deepest photo-z samples available today with  $\sigma_{z_p}$  increasing from 0.040 (8%) to 0.055 (28%) between  $H < 24$  and  $26 < H < 28$ , respectively. Over large fields of several  $\text{deg}^2$ ,  $\sigma_{z_p} \sim 0.05$  is routinely reached at  $i < 26$  [e.g. 88].

With broad band photometric data alone, the best  $\sigma_{z_p}$  is limited to approximately 0.025 (see Fig.3 for results on pencil-beam and wide area surveys) and this does not seem to improve even above large signal-to-noise values [e.g.,  $\text{SN} > 40$ ; 89]. In the last decade, medium band data with filter widths around  $400\text{\AA}$  have enabled a breakthrough in precision by improving the SED resolution. COMBO-17 [77] was the first survey to use medium bands imaging to produce a photo-z catalogue. There,  $\sigma_{z_p}$  of bright sources ( $i < 22$ ) reached  $\sim 0.02$ , thanks also to the ability to precisely locate the Balmer break. The COSMOS, SHARDS and ALHAMBRA surveys improved upon this precision to  $\sim 0.01$  out to a redshift  $\sim 1.5$  with deeper medium-band photometry and by introducing emission lines into the templates.

The increase in sensitivity of NIR detectors has allowed astronomers to extend their study beyond the *redshift desert* by following both the Balmer break between  $1.6 < z < 4$  [e.g. 90], and the Lyman break above  $z > 8$  [91]. In this way high precision photo-z ( $\sigma_{z_p} \sim 0.02$ ) has been reached between  $1.6 < z < 4$  for massive galaxies in the NEWFIRM survey [92].

Photo-z enable the identification of galaxy samples across thousands of  $\text{deg}^2$ . For example the DES survey is reaching depths of  $i \sim 24$  over  $5000 \text{ deg}^2$ , with a photo-z accuracy expected to be below  $\sim 0.08$ . The major challenge for wide-field surveys is the calibration of photometric noise precision over large areas while be-

ing observed over time scales of many years with varying sky conditions, and a potential degradation in quality of the instruments. Also, the spectroscopic training sample should not be limited to a small area of the sky, but rather homogeneously distributed.

Finally, beside these other difficulties, the procedure to extract the photometry from images is also crucial. Source extraction is commonly performed with SExtractor [94] but the region within which the galaxy flux is measured is important because too large an area would compromise the signal-to-noise of faint sources [95, 96]. Moreover the light should be produced by the same regions of the studied sources despite the Point Spread Function (PFS) variation from one band to another. In cases of PFS variation between the various bands, PFS homogenisation is necessary [e.g. 5, 88, 95, 97], and efficient tools are being developed to produce fluxes measured in a consistent way using the high-resolution images as reference for example see PYGFIT : [98]; Synmag: [99]; TFIT: [100]; T-PHOT: [101].

## V. PHOTO-Z OF EXOTIC SOURCES

The discovery that virtually every galaxy hosts a super massive black hole [e.g. 102], suddenly increased the interest of the scientific community for those galaxies hosting an AGN, now seen as a key ingredient in galaxy evolution models. Therefore, computing their redshifts became crucial. However, while the number density of AGNs is sufficiently high that spec-z follow-up via single slit spectroscopy would be too time consuming, they are so sparse that a multi-object spectroscopic campaign would be totally inefficient. Thus, the photo-z technique is favoured, although these sources are particularly challenging. In fact, their SED is characterised by the sum of two unknown relative contributions, host and AGN, and in any photometric band this depends on the type of host and the type and strength of the AGN [e.g., Figure 5 in 103, and related text]. When these constraints are ignored, and for example only templates of galaxies are used (i.e. assuming a dormant black hole), the photo-z obtained with SED-fitting will be characterised by a high number of degenerated photo-z solutions, or, worse, catastrophic failures. This is clearly demonstrate in the right panel of Figure 4. There, we show the difference between photo-z computed using either AGN [104] or galaxy [88] templates, for a sample of 1672 X-ray selected AGN with a secure spec-z from the Legacy *Chandra* COSMOS field [105]. In particular, the discrepancy between the photo-z solutions increases often with the strength of the AGN [104].

The choice of templates to include in a library is dictated by the type of AGN that are treated [e.g., depending on their hardness ratio 106], the selection technique (e.g. excess flux, variability) used for the identification and by the selection band (e.g., X-ray, radio, optical, infrared) and its depth. This is clearly demonstrated

in Figure 14 of [78], where the X-ray flux distribution of the AGN detected in various X-ray surveys is shown. Clearly, the population of sources that are represented in the wide but shallow XMM-COSMOS [e.g. 107] survey, is very underrepresented in the deep but narrow *Chandra* observations of the Chandra Deep Field South field [e.g., 75]. Reciprocally, the bulk of the sources in the latter field is not represented at all in the former. Similar results are shown recently in [71] for Radio selected AGN. This is particular relevant considering that wide or all-sky surveys such as eROSITA [108] in X-ray and EMU [109] in the Radio are expected to detect million of AGN that will be used as cosmological probes and for evolutionary studies.

The problems is alleviated in rich photometric datasets which are inclusive of both intermediate and narrow bands filters because they help to identify emission lines [78, 79, 110], which are the only feature visible in the power-law continuum of a QSO (bright AGN outshining their host) for which colours are independent of redshift (see for example the cyan track in panel b) of 1).

Priors are a crucial ingredient when computing photo-z for AGN via template-fitting. Specifically, given the typical absolute magnitude of these sources (typically brighter than -22 in blue optical bands), a prior based on the apparent size allows to limit the range in redshift solution and thus the degeneracy [e.g., 104, 110, 111].

ML techniques have also estimated photo-z for AGN but the best results are mostly reliable for QSO at high redshift, where the host contribution can be neglected. Examples of some successful applications are presented in [112, 113] and [114]. The mentioned works are all focused on the SDSS footprints, where the plethora of spectroscopic follow-up data are well suited to ML techniques. Very recently [115] applied ML to a sample of bright X-ray selected AGN, providing reliable results also at low-redshift.

An additional problem when computing photo-z for AGN is that variability across the spectrum is an intrinsic property of the sources, with different cadence and intensity. This makes it difficult to reliably estimate photo-z when multi-wavelength data are collected over time scales of years [116].

Variability does not only effect QSOs, but also objects as Bl-Lac, Gamma-ray Bursts (GRBs) and Supernovae (SN). Bl-Lac belong to the family of AGN, in particular Blazars, i.e. sources that are observed through the jet launched from the center and for this reason are characterised by a featureless power-law continuum, unusable for assessing redshift via spectroscopy. Here, simultaneous observations in multiple bands are required in order to measure the absorption bluer than  $\text{Ly}\alpha$  by the intervening material [117]. Gamma-Ray-Burst (GRB) and Super Novae (SN) are sudden and quickly occurring explosive events that allow us to measure the distance of galaxies that would be otherwise too faint to be detected. For GRBs, we rely on simultaneous observations in optical+NIR and SED fitting templates [118], while the most

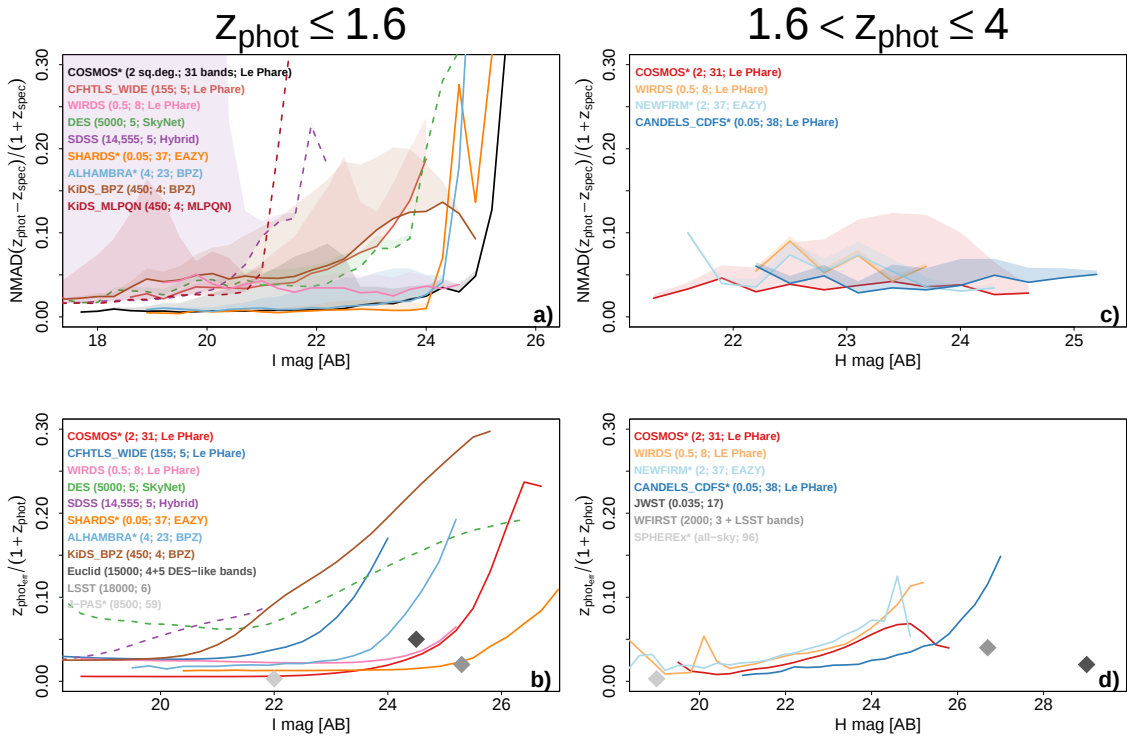


FIG. 3: **Accuracy of various photo- $z$  surveys as a function of their depth and number of bands.** The low redshift range is shown in the panels a) and b) where the accuracy is measured as function of the  $i$  band. For the high redshift range (panels c) and d)), the accuracy is measured as a function of the near-infrared depth in the  $H$  band. In all panels only sources with a photometric error smaller than 0.3 are considered. The accuracy is measured with NMAD, with the solid lines indicating the surveys where photo- $z$  were computed via template-fitting, while the dashed lines indicate the surveys with photo- $z$  computed via ML or Hybrid methods. Surveys with intermediate or narrow band photometry (COSMOS: [88]; SHARDS: [80]; ALHAMBRA: [81]; CANDELS/CDFS: [78]; NEWFIRM: [92]) are highlighted with an asterisk. The coloured areas are proportional to the number of reliable spectroscopic redshifts used for the training, at that magnitude, while the numbers in brackets beside each survey indicate their size in square degrees, the number of bands and the code used. In general, the less accurate photo- $z$  are at the faint end of each survey. A measure of the accuracy beyond the availability of spectroscopic redshift is provided in panels b) and d), where the solid lines show the median photo- $z$  error divided by  $(1+z_{\text{phot}})$ , as a function of the depth of the survey. It is clear how the accuracy in all the panels is better in the region well populated by the spectroscopy, where the sources are brighter. Surveys with intermediate and narrow band photometry perform better than those having only broad band photometry (WIRDS: [132]; CFHTLS.WIDE: [89]; KiDS: [133]; SDSS:[82]; DES: [85]). Finally, with grey symbols we add the expected accuracy for the future cosmological surveys.

reliable photo- $z$  for SN rely on the use of priors [e.g. 119].

## VI. THE FUTURE

### A. Reaching the deep Universe

The photo- $z$  technique gained momentum when the first images from the Hubble Deep Field became available in 1995. [120] published a first photo- $z$  catalogue to study 1683 galaxies out to  $z=6$ , which was an incredible leap forward in studying the high redshift Universe. Such analysis was followed by numerous attempts to make the photo- $z$  estimation technique more robust [e.g. 72, 76], and to exploit the following generations of HST deep surveys [e.g. 70]. The template-fitting technique has been central since only sparsely spectroscopic coverage is possible in these deep fields. The study of

primordial galaxies will trigger a new burst of activity in the next years with the launch of the James Web Space Telescope [121, JWST]. With the efficient near-infrared NIRCAM camera, surveys could be conducted at depths of  $mag_{AB} \sim 30 - 31$  also by adopting a strategy similar to the CANDELS survey. [122] showed that NIRCAM is a powerful camera to produce photo- $z$  above  $z > 5$  out to  $z \sim 20$ .

Based on the HST experience, template-fitting will probably remain the main photo- $z$  technique in these extremely faint fields. However, new difficulties will arise because of the increasing contribution of the emission lines with redshift [123] and an evanescent Balmer break feature. While emission lines create gradients in the colour- $z$  relation that could be used to measure a redshift [e.g. 124], it makes also the estimate more challenging because more prone to degenerated solutions [122]. The knowledge of the galaxies and space at redshift be-

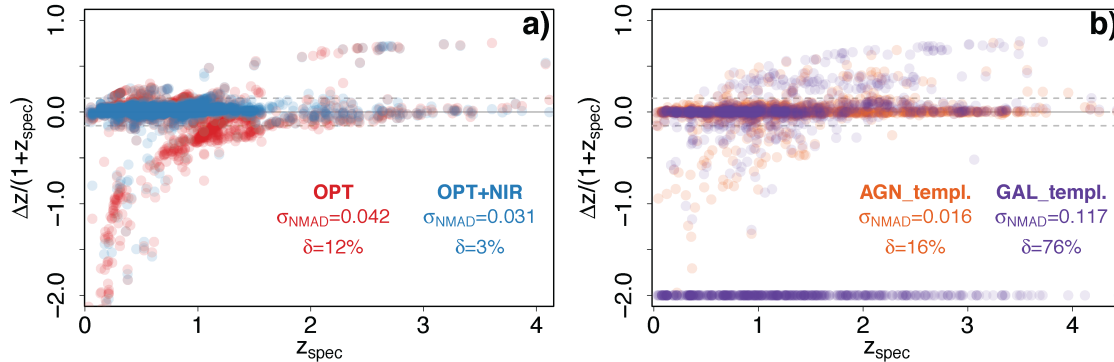


FIG. 4: **Impact of multi-lambda coverage and choice of templates.**

**a):** Comparison of photo- $z$  for galaxies computed using the optical bands (ugriz; red) or the optical and near infrared (NIR) (ugrizJHK; blue). The addition of NIR infrared photometry improves dramatically the precision at  $z > 1$ . Sources external to the dashed lines  $\pm 0.15(1+z_{\text{spec}})$  are considered outliers. **b):** Difference between photo- $z$  obtained using either galaxy (violet) or AGN (orange) templates for sources that are X-ray detected in *Chandra* Legacy COSMOS and with a reliable spec- $z$ . The 33% of sources for which no template of normal galaxy could compute the photo- $z$  are artificially set to -2.

yond  $z > 10$  will be soon revealed thanks to the deep (28 AB magnitude in Near-infrared for a signal/noise=10 in  $10^4$  s[? ]) for spectroscopy that JWST will perform in optical and near-infrared. Template-fitting methods will then have to incorporate the new knowledge in form of new templates and new physics.

## B. Building blocks to unveil new cosmology

The next decade will see many exciting imaging surveys dedicated to cosmology. The goal of these surveys is to constrain the nature of the dark energy through the combination of various probes. Several probes rely on photo- $z$ , such as galaxy cluster counts, weak lensing tomography, and Baryon Acoustic Oscillations (BAO). These probes require large statistical samples over large areas to extract their cosmological signals, which has triggered this period of gigantism for the next generation of cosmological surveys. While the ongoing DES survey will cover  $5000 \text{ deg}^2$  in 5 bands and gathering information for 300 millions of sources, the next generation of surveys will increase the number of galaxies by one order of magnitude. For instance, LSST [18] will begin operations in 2021 and cover  $18000 \text{ deg}^2$  of the sky, gathering information on 4 billions sources down to  $r_{AB} = 27.5$  after 10 years of operation. The Euclid mission [17] will be launched in 2021 and remain in operation for 6 years and use its VIS camera to measure the galaxy shapes of 1.5 billion of galaxies over  $15000 \text{ deg}^2$ .

Computation of photo- $z$  and storage of the PDF will be as challenging as to maintain the required photometry quality across such a large area. For instance, one requirement of the LSST survey is to get band-to-band

calibration errors not larger than 0.005 mag and no more than 0.01 mag variation across the sky. Being successful in maintaining such relative calibration is necessary to insure homogeneous performances.

As describe in section IV, the multi-lambda coverage is also crucial to define the redshift range of interest. In the new generation of imaging surveys, unless SPHEREx [125] with its 96 medium band NIR photometry is approved, Euclid will be the only survey to map the Universe in NIR using three filters between  $9200\text{\AA}$ - $20000\text{\AA}$ . The goal of these filters is to insure precise photometric redshifts at  $z > 1.3$  which is not possible without NIR (see panel a) of Fig. 4). However, Euclid will need to be complemented with ground-based data in optical wavelengths.

While most of the cosmological imaging surveys are performed with broad bands (HSC, DES, LSST, Euclid), two surveys are performed using medium bands: the PAU and the J-PAS surveys. For instance, J-PAS [126] will cover  $8500 \text{ deg}^2$  with 54 narrow band filters. They should observe 300 million galaxies with a photo- $z$  precision of 0.3%. With this precision one of the main objectives of this survey will be to the measure the BAO.

Finally, the variable Universe will become more easily accessible with LSST due to the 1000 repeated observations of each location of the sky over 10 years, starting in 2022, in up to six bands. In particular it will include the  $u$  band, assuring the capability of breaking the degeneracy between low and high redshift solutions. More in general, LSST will be extremely powerful to break the identify stars, AGN and exotic sources.

### C. Evolution in the photo-z technique and synergy with spectroscopic surveys

The evolution of the photo-z methods will depend on the spectroscopic data that will become available. For the next generation of cosmological surveys, the imaging and spectroscopic surveys are often conceived in parallel. For instance, the HSC imaging survey will be complemented with spectroscopy using the multi-object Prime Focus Spectrograph (PFS) instrument covering a wavelength range 3800Å-13000Å, and gathering redshift for millions of galaxies at  $0.8 < z < 6$  [127]. Similarly, the Euclid survey is complemented by the spectroscopy performed by the NISP instrument, able to perform NIR (11000Å-20000Å) slit-less spectroscopy to detect the  $H_\alpha$  emission for more than 50 millions of galaxies at  $0.7 < z < 2$ . With such a bright future for spectroscopic surveys, the ML techniques will become competitive in the high redshift Universe. It must also be stressed that because of the incumbent massive photometric surveys designed as cosmological probes, the algorithms and techniques for computing reliable photo-z are in continuous development. The most recent methods are hybrids that combines the best of the ML and SED fitting techniques [e.g., 47, 82, 128–130].

Specific needs will also trigger new techniques. Scientific application based on the weak lensing tomography require the characterisation of the true redshift distribu-

tion with an extreme precision in order to map the result of a weak lensing shear analysis onto cosmological parameters. In a very simplified setting, we show in Fig.5 how a biased photo-z distribution expected for one tomographic redshift bin for a Euclid-like survey, will impact the predicted auto-correlation function, and therefore will lead to incorrect best fitting cosmological parameters. The actual effect of such an offset depends strongly on the redshift of the galaxy sample, and the full shape of the redshift distribution function. In a weak lensing setting, we direct the reader to [Fig.1.3 in 2, 131] showing how a bias on the mean redshift translates into an error on the time evolution component of the dark energy equation of state in the LSST case. With current methods, such low bias can not be achieved using photo-z alone. The development of new techniques, able to combine the photo-z with the sources position (see §. IID) have the potential to revolutionise the field.

### ACKNOWLEDGMENTS

The Authors are grateful to the referees and the editor who contribute to the review with their comments. The authors are grateful to Li-Ting Hsu, Jonas Chavez-Montego, Pablo Gonzalez, Hendrik Hildebrandt, Massimo Brescia and Stefano Cavuoti for providing the CANDELS/CDFS, ALHAMBRA, SHARDS, KiDS.BPT, KiDS.MLPQN data that were used in Fig. 3.

- 
- [1] Le Fevre, O., *et al.* The VIMOS VLT deep survey. First epoch VVDS-deep survey: 11 564 spectra with  $17.5 \leq \text{IAB} \leq 24$ , and the redshift distribution over  $0 \leq z \leq 5$ . *A&A* **439**, 845 (2005)
  - [2] Newman, J.A., *et al.* Spectroscopic needs for imaging dark energy experiments. *Astroparticle Physics* **63**, 81 (2015)
  - [3] Baum W.A., Photoelectric determinations of redshifts beyond 0.2 c. *AJ* **62**, 6 (1957)
  - [4] Puschell, J.J., Owen, F.N., Laing, R.A. Near-infrared photometry of distant radio galaxies - Spectral flux distributions and redshift estimates. *ApJL* **257**, L57 (1982)
  - [5] Ilbert O., *et al.* Cosmos Photometric Redshifts with 30-Bands for 2-deg<sup>2</sup>. *ApJS* **189**, 270 (2009)
  - [6] Fontana A., *et al.* Photometric Redshifts and Selection of High-Redshift Galaxies in the NTT and Hubble Deep Fields. *AJ* **120**, 2206 (2000)
  - [7] Dunlop J. S., *et al.* A critical analysis of the ultraviolet continuum slopes ( $\beta$ ) of high-redshift galaxies: no evidence (yet) for extreme stellar populations at  $z > 6$ . *MNRAS* **420**, 901 (2012)
  - [8] Coupon J., *et al.* The galaxy-halo connection from a joint lensing, clustering and abundance analysis in the CFHTLenS/VIPERS field. *MNRAS* **449**, 1352 (2015)
  - [9] Finoguenov A., *et al.* The XMM-Newton Wide-Field Survey in the COSMOS Field: Statistical Properties of Clusters of Galaxies. *ApJS* **172**, 182 (2007)
  - [10] López-Sanjuan C., *et al.* The dominant role of mergers in the size evolution of massive early-type galaxies since  $z = 1$ . *A&A* **548**, A7 (2012)
  - [11] Man A. W. S., Zirm A. W., and Toft S., Resolving the Discrepancy of Galaxy Merger Fraction Measurements at  $z = 0-3$ . *ApJ* **830**, 89 (2016)
  - [12] Etherington J., *et al.* Environmental dependence of the galaxy stellar mass function in the Dark Energy Survey Science Verification Data *MNRAS* **466**, 228 (2017)
  - [13] Etherington J., Thomas D., Measuring galaxy environments in large-scale photometric surveys. *MNRAS* **451**, 660 (2015)
  - [14] Malavasi N., Pozzetti L., Cucciati O., Bardelli S., and Cimatti A., Reconstructing the galaxy density field with photometric redshifts. I. Methodology and validation on stellar mass functions *A&A* **585**, A116 (2016)
  - [15] Miyaji T., *et al.* Detailed Shape and Evolutionary Behavior of the X-Ray Luminosity Function of Active Galactic Nuclei. *ApJ* **804**, 104 (2015)
  - [16] Padovani P., Giommi P., and Rau A., The discovery of high-power high synchrotron peak blazar. *MNRAS* **422**, L48 (2012)
  - [17] Laureijs R., *et al.* Euclid Definition Study Report. *2011arXiv1110.3193L* (2011)
  - [18] Ivezić Z., *et al.* *2008arXiv0805.2366I* (2008)
  - [19] Hu W., Power Spectrum Tomography with Weak Lensing. *ApJL* **522**, L21 (1999)
  - [20] Benitez N., *et al.* Optimal Filter Systems for Photometric Redshift Estimation. *ApJL* **692**, 5 (2009)

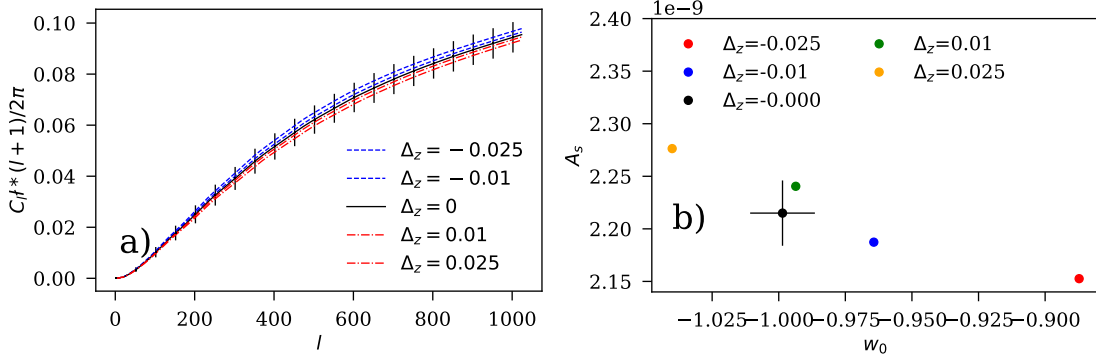


FIG. 5: **Cosmological parameter biases incurred by photo-z inaccuracy.** **a):** the predicted auto-correlation function ( $C_l$ ) is shown for a simulated Euclid-like redshift distribution (black line) and for a redshift distribution modified by a translation of up to  $\pm 0.025$  (red and blue lines). **b):** The subsequent biases induced in the cosmological parameters,  $w_0$  - the equation of state of dark energy, and  $A_s$  - the amplitude of the fluctuations on the surface of last scattering, from the redshift translations.

- [21] Fioc M. and Rocca-Volmernage B., Spectrophotometric Evolution from the UV to the NIR. *From Stars to Galaxies: the Impact of Stellar Physics on Galaxy Evolution*. *Astronomical Society of the Pacific Conference Series* **98**, 67 (1996)
- [22] Bruzual G., and Charlot S., Stellar population synthesis at the resolution of 2003. *MNRAS* **344** 1000 (2003)
- [23] Maraston C. Evolutionary population synthesis: models, analysis of the ingredients and application to high-z galaxies. *MNRAS* **362**, 799 (2005)
- [24] Conroy C. On the Birth Masses of the Ancient Globular Clusters. *apj* **758**, 21 (2012)
- [25] Coleman G. D., Wu C.-C., and Weedman D. W. Colors and magnitudes predicted for high redshift galaxies. *ApJS* **43**, 393 (1980)
- [26] Kinney A. L., *et al.* Template Ultraviolet to Near-Infrared Spectra of Star-forming Galaxies and Their Application to K-Corrections. *ApJ* **467**, 38 (1996)
- [27] Polletta M., *et al.* Spectral Energy Distributions of Hard X-Ray Selected Active Galactic Nuclei in the XMM-Newton Medium Deep Survey. *ApJ* **663**, 81 (2007)
- [28] Noll S., *et al.* The FORS Deep Field spectroscopic survey. *A&A* **418**, 885 (2004)
- [29] Chevallard J., and Charlot S., Erratum: Modelling and interpreting spectral energy distributions of galaxies with BEAGLE. *MNRAS* **464**, 2349 (2017)
- [30] Ilbert O., *et al.* Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey. *A&A* **457**, 841 (2006)
- [31] Schaerer D. The importance of nebular emission for SED modeling of distant star-forming galaxies. *The Spectral Energy Distribution of Galaxies - SED 2011, IAU Symposium* **284**, 20 (2012)
- [32] Pacifici C., Charlot S., Blaizot J., Brinchmann, J. Relative merits of different types of rest-frame optical observations to constrain galaxy physical parameters *MNRAS* **421**, 2002 (2012)
- [33] Calzetti D., *et al.* The Dust Content and Opacity of Actively Star-forming Galaxies. *ApJ* **533**, 682 (2000)
- [34] Prevot M. L., Lequeux J., Prevot L., Maurice E., and Rocca-Volmerange E., The typical interstellar extinction in the Small Magellanic Cloud. *A&A* **132**, 389 (2013)
- [35] Madau P., Radiative transfer in a clumpy universe: The colors of high-redshift galaxies. *ApJ* **441**, 18 (1995)
- [36] Draine B. T., Title *Physics of the Interstellar and Intergalactic Medium by Bruce T. Draine*. *Princeton University Press, 2011. ISBN: 978-0-691-12214-4*, (2011)
- [37] Schlegel D. J., Finkbeiner D. P., and Davis M., Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds. *ApJ* **500**, 525 (1998)
- [38] Galametz A., Saglia R., Paltani S., Apostolakis N., and Dubath P. SED-dependent galactic extinction prescription for Euclid and future cosmological surveys *A&A* **598**, A20 (2017)
- [39] Hildebrandt H., *et al.* PHAT: PHoto-z Accuracy Testing. *A&A* **523**, A31 (2010)
- [40] Cavauti S., Brescia M., Longo G., and Mercurio A., Photometric redshifts with the quasi Newton algorithm (MLPQNA) Results in the PHAT1 contest. *A&A* **546**, A13 (2012)
- [41] Beck R., *et al.* On the realistic validation of photometric redshift. *MNRAS* **468**, 4323 (2017)
- [42] Breiman L. Random Forests. *Machine Learning* **45**, 5 (2001)
- [43] Rosenblatt, F., Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms *Spartan Books* (2005)
- [44] Carliles S., Budavári T., Heinis S., Priebe c., Szalay A. S., Random Forests for Photometric Redshifts. *ApJ* **712**, 511 (2010)
- [45] Carrasco Kind, M., & Brunner, R. J. TPZ: photometric redshift PDFs and ancillary information by using prediction trees and random forests. *MNRAS* **432**, 1483 (2013)
- [46] Collister A. A., and Lahav O., ANNz: Estimating Photometric Redshifts Using Artificial Neural Networks. *PASP* **116**, 345 (2004)
- [47] Vanzella E., *et al.* Photometric redshifts with the Multilayer Perceptron Neural Network: Application to the HDF-S and SDSS. *A&A* **423**, 761 (2004)
- [48] Bonnett C. Using neural networks to estimate redshift distributions. An application to CFHTLenS. *MNRAS*

- 449**, 1043 (2015)
- [49] Brescia M., Cavaoti S., and Longo G., Automated physical classification in the SDSS DR10. A catalogue of candidate quasars. *mnras* **450**, 3893 (2015)
- [50] Bonnett, C., Using neural networks to estimate redshift distributions. An application to CFHTLenS. *MNRAS* **449**, 1043 (2015)
- [51] Hoyle, B., Measuring photometric redshifts using galaxy images and Deep Neural Networks *Astronomy and Computing* **16**, 34 (2016)
- [52] D’Isanto A., and Polsterer K. L., Photometric redshift estimation via deep learning. *A&A*, (2018)
- [53] Carrasco-kind M., and Brunner R. J., SOMz: photometric redshift PDFs with self-organizing maps and random atlas *MNRAS* **438**, 3409 (2014)
- [54] Masters D., *et al.* Mapping the Galaxy Color-Redshift Relation: Optimal Photometric Redshift Calibration Strategies for Cosmology Surveys. *ApJ* **813**, 53 (2015)
- [55] Cavaoti S., *et al.* METAPHOR: A machine-learning-based method for the probability density estimation of photometric redshifts. *MNRAS* **465**, 1959 (2017)
- [56] Hoyle B., *et al.* Anomaly detection for machine learning redshifts applied to SDSS galaxies. *MNRAS* **452**, 4183 (2015)
- [57] Hoyle B., Rau M. M., Seitz S., and Weller J., Data augmentation for machine learning redshifts applied to Sloan Digital Sky Survey galaxies. *MNRAS* **450**, 305 (2015)
- [58] Hoyle B., Rau M. M., Zitlau R., Seitz S., and Weller J., Feature importance for machine learning redshifts applied to SDSS galaxies. *MNRAS* **449**, 1275 (2015)
- [59] Lima, M., *et al.* Estimating the redshift distribution of photometric galaxy sample *MNRAS* **390**, 118 (2008)
- [60] Benítez N., Bayesian Photometric Redshift Estimation. *ApJ* **536**, 571 (2000)
- [61] Dahlen T., *et al.* A Detailed Study of Photometric Redshifts for GOODS-South Galaxies. *ApJ* **724**, 425 (2010)
- [62] Tanaka, M. Photometric Redshift with Bayesian Priors on Physical Properties of Galaxies. *ApJ* **801**, 20 (2015)
- [63] Seldner M., and Peebles P. J. E., Statistical analysis of catalogs of extragalactic objects. XI - Evidence of correlation of QSOs and Lick galaxy counts. *ApJ* **227**, 30 (1979)
- [64] Newman J. A., Calibrating Redshift Distributions beyond Spectroscopic Limits with Cross-Correlations. *ApJ* **684**, 88 (2008)
- [65] Ménard B., Clustering-based redshift estimation: method and application to data. *arXiv:1303.4722*, (2013)
- [66] Scottez V., *et al.* Clustering-based redshift estimation: application to VIPERS/CFHTLS. *MNRAS* **462**, 1683 (2016)
- [67] Rahman M., Ménard B., Scranton R., Schmidt C. B., and Morrison C. B., Clustering-based redshift estimation: comparison to spectroscopic redshifts *MNRAS* **447**, 3500 (2015)
- [68] Aragon-Calvo M. A., van de Weygaert R., Jones B. J. T., and Mobasher B., Submegaparsec individual photometric redshift estimation from cosmic web constraints. *MNRAS* **454**, 463 (2015)
- [69] Sánchez C., *et al.* Photometric redshift analysis in the Dark Energy Survey Science Verification data. *MNRAS* **445**, 1482 (2014)
- [70] Dahlen T., *et al.* A Critical Assessment of Photometric Redshift Methods: A CANDELS Investigation. *ApJ* **775**, 93 (2013)
- [71] Duncan K. J., *et al.* Photometric redshifts for the next generation of deep radio continuum surveys - I: Template fitting. *arxiv.org/abs/1709.09183*, (2017)
- [72] Bolzonella M., Miralles J.-M., and Pelló R., Photometric redshifts based on standard SED fitting procedures. *A&A* **363**, 476 (2000)
- [73] Brammer G. B., van Dokkum P. G., and Coppi P., EAZY: A Fast, Public Photometric Redshift Code. *ApJ* **686**, 1503 (2008)
- [74] Feldmann R., *et al.* The Zurich Extragalactic Bayesian Redshift Analyzer and its first application: COSMOS. *MNRAS* **372**, 565 (2006)
- [75] Luo B., *et al.* Identifications and Photometric Redshifts of the 2 Ms Chandra Deep Field-South Sources. *ApJS* **187**, 560 (2010)
- [76] Arnouts S., *et al.* Measuring and modelling the redshift evolution of clustering: the Hubble Deep Field North *MNRAS* **310**, 540 (1999)
- [77] Wolf C., *et al.* A catalogue of the Chandra Deep Field South with multi-colour classification and photometric redshifts from COMBO-17. *A&A* **421**, 913 (2004)
- [78] Hsu L.-T., *et al.* CANDELS/GOODS-S, CDFS, and ECDFS: Photometric Redshifts for Normal and X-Ray-Detected Galaxies. *ApJ* **796**, 60 (2004)
- [79] Cardamone C.N., *et al.* Dust-corrected Colors Reveal Bimodality in the Host-galaxy Colors of Active Galactic Nuclei at  $z \approx 1$ . *ApJS* **182**, 270 (2010)
- [80] Pérez-González P. G., SHARDS: An Optical Spectrophotometric Survey of Distant Galaxies. *ApJ* **762**, 46 (2013)
- [81] Molino A., *et al.* The ALHAMBRA Survey: Bayesian photometric redshifts with 23 bands for 3 deg. *MNRAS* **441**, 2891 (2014)
- [82] Beck R., Dobos L., Budavári T., Szalay A. S., and Csabai I., Photometric redshifts for the SDSS Data Release 12. *MNRAS* **460**, 1371 (2016)
- [83] Sadeh I., Abdalla F.B., Lahav O., ANNz2: Photometric Redshift and Probability Distribution Function Estimation using Machine Learning. *PASP* **128**, 104302 (2017)
- [84] Bordoloi R., Lilly S. J., Amara A., Photo-z performance for precision cosmology. *MNRAS* **406**, 881 (2010)
- [85] Bonnett C., *et al.* Redshift distributions of galaxies in the Dark Energy Survey Science Verification shear catalogue and implications for weak lensing. *MNRAS* **408**, 1168 (2016)
- [86] Quadri R. F., and Williams R. J., Quantifying Photometric Redshift Errors in the Absence of Spectroscopic Redshifts. *ApJ* **725**, 794 (2010)
- [87] Benjamin J., van Waerbeke L., Ménard B., and Kilbinger M., Photometric redshifts: estimating their contamination and distribution using clustering information *MNRAS* **408**, 1168 (2010)
- [88] Laigle C., *et al.* The COSMOS2015 Catalog: Exploring the  $1 < z < 6$  Universe with Half a Million Galaxies. *ApJS* **224**, 24 (2016)
- [89] Coupon J., *et al.* Photometric redshifts for the CFHTLS T0004 deep and wide fields. *A&A* **500**, 981 (2009)
- [90] Cimatti A., *et al.* The K20 survey. I. Disentangling old and dusty star-forming galaxies in the ERO population. *A&A* **381**, L68 (2002)
- [91] Bouwens R. J., *et al.* UV Luminosity Functions at Redshifts  $z = 4$  to  $z = 10$ : 10,000 Galaxies from HST Legacy

- Fields. *A&A* **803**, 34 (2015)
- [92] Whitaker K. E., *et al.* The NEWFIRM Medium-band Survey: Photometric Catalogs, Redshifts, and the Bi-modal Color Distribution of Galaxies out to  $z \sim 3$ . *ApJ* **735**, 86 (2011)
- [93] Drlica-Wagner A., *et al.*, Dark Energy Survey Year 1 Results: Photometric Data Set for Cosmology.
- [94] Bertin E., and Arnouts S., SExtractor: Software for source extraction. *A&AS* **117**, 393 (1996)
- [95] Hildebrandt H., *et al.* CFHTLenS: improving the quality of photometric redshifts with precision photometry.. *MNRAS* **421**, 2355 (2012)
- [96] Moutard T., *et al.* The VIPERS Multi-Lambda Survey. I. UV and near-IR observations, multi-colour catalogues, and photometric redshifts. *A&A* **590**, A102 (2016)
- [97] Grazian A., *et al.* A comparison of LBGs, DRGs, and BzK galaxies: their contribution to the stellar mass density in the GOODS-MUSIC sample. *A&A* **465**, 393 (2007)
- [98] Mancone C. L., Gonzales A. H., Moustakas L. A., and Price A., PyGFit: A Tool for Extracting PSF Matched Photometry. *PASP* **125**, 1514 (2013)
- [99] Bundy K., *et al.* SYNMAJ Photometry: A Fast Tool for Catalog-level Matched Colors of Extended Sources. *AJ* **144**, 188 (2012)
- [100] Laidler V. G., *et al.* TFIT: A Photometry Package Using Prior Information for Mixed-Resolution Data Sets. *PASP* **119**, 1325 (2007)
- [101] Merlin E., *et al.* T-PHOT: A new code for PSF-matched, prior-based, multiwavelength extragalactic deconvolution photometry. *A&A* **582**, A15 (2015)
- [102] Magorrian J., *et al.* The Demography of Massive Dark Objects in Galaxy Centers. *AJ* **115**, 2285 (1998)
- [103] Bongiorno A., *et al.* Accreting supermassive black holes in the COSMOS field and the connection to their host galaxies. *MNRAS* **427**, 3103 (2012)
- [104] Salvato M., *et al.* Dissecting Photometric Redshift for Active Galactic Nucleus Using XMM- and Chandra-COSMOS Samples. *ApJ* **742**, 61 (2011)
- [105] Marchesi S., *et al.* The Chandra COSMOS Legacy survey: optical/IR identifications. *ApJ* **817**, 34 (2016)
- [106] Zheng W., *et al.* Photometric Redshift of X-Ray Sources in the Chandra Deep Field South. *ApJS* **155**, 73 (2004)
- [107] Brusa M., *et al.* The XMM-Newton Wide-field Survey in the Cosmos Field (XMM-COSMOS): Demography and Multiwavelength Properties of Obscured and Unobscured Luminous Active Galactic Nuclei. *ApJ* **716**, 348 (2010)
- [108] Merloni A., *et al.* eROSITA Science Book: Mapping the Structure of the Energetic Universe *2012arXiv1209.3114M*
- [109] Norris R., *et al.* EMU: Evolutionary Map of the Universe., *PASA* **28**, 215 (2011)
- [110] Salvato M., *et al.* Photometric Redshift and Classification for the XMM-COSMOS Sources. *ApJ* **690**, 1250 (2009)
- [111] Kitsionas S., Hatziminaoglou E., Georgakakis A., and Georgantopoulos I., On the use of photometric redshifts for X-ray selected AGNs. *A&A* **434**, 475 (2005)
- [112] Bovy J., *et al.* Photometric Redshifts and Quasar Probabilities from a Single, Data-driven Generative Model. *ApJ* **749**, 41 (2012)
- [113] Brescia M., Caviuoti S., D'Abrusco R., Long G., and Mercurio A., Photometric Redshifts for Quasars in Multi-band Surveys. *ApJ* **772**, 140(2013)
- [114] Budavari T., *et al.* Photometric Redshifts from Reconstructed Quasar Templates. *AJ* **122**, 1163 (2001)
- [115] Mountrichas E. A., Estimating Photometric Redshifts for X-ray sources in the X-ATLAS field, using machine-learning techniques. *arXiv:1710.01313*, (2017)
- [116] Simm T., *et al.* Pan-STARRS1 variability of XMM-COSMOS AGN. I. Impact on photometric redshifts. *A&A* **584**, A106 (2015)
- [117] Rau, A., *et al.* BL Lacertae objects beyond redshift 1.3 - UV-to-NIR photometry and photometric redshift for Fermi/LAT blazars. *A&A* **538**, A26 (2012)
- [118] Krühler T., *et al.* Photometric redshifts for gamma-ray burst afterglows from GROND and Swift/UVOT. *A&A* **526**, A153 (2011)
- [119] Palanque-Delabrouille N., *et al.* Photometric redshifts for type Ia supernovae in the supernova legacy survey. *A&A* **514**, A63 (2010)
- [120] Lanzetta K. M., Yahil A., and Fernández-Soto A., Star-forming galaxies at very high redshifts. *Nature* **381**, 759 (1996)
- [121] Finkelstein S., *et al.*, The Case for a James Webb Space Telescope Extragalactic Key Project *2015arXiv151204530F* (2015)
- [122] Bisigello, L., *et al.* The Impact of JWST Broadband Filter Choice on Photometric Redshift Estimation. *ApJS* **227**, 19 (2016)
- [123] Schaerer D., and de Barros S., On the physical properties of  $z$  Appl. Phys. 6-8 galaxies. *A&A* **515**, A73 (2010)
- [124] Labbe, I., *et al.* The Spectral Energy Distributions of  $z$  8 Galaxies from the IRAC Ultra Deep Fields: Emission Lines, Stellar Masses, and Specific Star Formation Rates at 650 Myr. *ApJL* **777**, L19 (2013)
- [125] Doré O., *et al.* Science Impacts of the SPHEREx All-Sky Optical to Near-Infrared Spectral Survey: Report of a Community Workshop Examining Extragalactic, Galactic, Stellar and Planetary Science, *2016arXiv160607039D*
- [126] Benitez N., *et al.* J-PAS: The Javalambre-Physics of the Accelerated Universe Astrophysical Survey. *2014arXiv1403.5237B* (2014)
- [127] Tamura N. *et al.* Prime Focus Spectrograph (PFS): A Very Wide-Field, Massively Multi-Object, Optical and Near-Infrared Fiber-Fed Spectrograph on the Subaru Telescope. *ASP Conference Series* **507**, 387 (2016)
- [128] Carrasco-Kind M., Brunner R. J. Exhausting the information: novel Bayesian combination of photometric redshift PDFs *MNRAS* **468**, 4323 (2014)
- [129] Leistedt B., and Hogg D. W., Data-driven, Interpretable Photometric Redshifts Trained on Heterogeneous and Unrepresentative Data. *ApJ* **838**, 5 (2017)
- [130] Speagle J., Eisenstein D.j., Deriving photometric redshifts using fuzzy archetypes and self-organizing maps - I. Methodology. *MNRAS* **469**, 1186 (2017)
- [131] Amendola L., *et al.* Cosmology and Fundamental Physics with the Euclid Satellite. *Living Reviews in Relativity* **16**, 6 (2013)
- [132] Bielby R., *et al.* The WIRCам Deep Survey. I. Counts, colours, and mass-functions derived from near-infrared imaging in the CFHTLS deep fields. *A&A* **545**, A23 (2012)
- [133] De Jong J. T. A., *et al.* The third data release of the Kilo-Degree Survey and associated data products., *A&A*

604, A134 (2017)