Accuracy of photometric redshifts for future weak lensing surveys from space

F. Bellagamba\textsuperscript{1}, M. Meneghetti\textsuperscript{2,3}, L. Moscardini\textsuperscript{1,2,3}, M. Bolzonella\textsuperscript{2}

\textsuperscript{1}Dipartimento di Astronomia, Universit`a di Bologna, Via Ranzani 1, 40127, Bologna, Italy
\textsuperscript{2}INAF-Osservatorio Astronomico di Bologna, Via Ranzani 1, 40127, Bologna, Italy
\textsuperscript{3}INFN-National Institute for Nuclear Physics, Sezione di Bologna, Viale Berti Pichat 6/2, 40127, Bologna, Italy

ABSTRACT

Photometric redshifts are a key tool to extract as much information as possible from planned cosmic shear experiments. In this work we aim to test the performances that can be achieved with observations in the near-infrared from space and in the optical from the ground. This is done by performing realistic simulations of multi-band observations of a patch of the sky, and submitting these mock images to software usually applied to real images to extract the photometry and then a redshift estimate for each galaxy. In this way we mimic the most relevant sources of uncertainty present in real data analysis, including blending and light pollution between galaxies. As an example we adopt the infrared setup of the ESA-proposed Euclid mission, while we simulate different observations in the optical, modifying filters, exposure times and seeing values. Finally, we consider directly some future ground-based experiments, such as LSST, Pan-Starrs and DES. The results highlight the importance of $u$-band observations, especially to discriminate between low ($z \lesssim 0.5$) and high ($z \sim 3$) redshifts, and the need for good observing sites, with seeing FWHM $< 1$. arcsec. The former of these indications clearly favours the LSST experiment as a counterpart for space observations, while for the other experiments we need to exclude at least 15% of the galaxies to reach a precision in the phot-z\textsuperscript{s} equal to $\langle \sigma_{\text{phot-z}} \rangle < 0.05$.

Key words: cosmology: dark energy – cosmology: observations – gravitational lensing: weak – galaxies: photometry – galaxies: distances and redshifts

1 INTRODUCTION

Cosmic shear experiments observe the large-scale structure of the Universe through the distortions in the shape of galaxies induced by the matter distribution along the line of sight. The statistical correlation between the observed ellipticities of the galaxies depends both on the geometry of the Universe and on the growth of structures. These are determined by the density of the main components of the Universe, such as dark matter and dark energy, and by the physics that governs their cosmological behaviour. Thus, weak lensing observations can give constraints over the fundamental physics that stands behind the evolution of the Universe. This is the idea that drives many proposed experiments, involving both ground-based telescopes [e.g. Kilo-Degree Survey, Pan-STARRS, Dark Energy Survey (DES), Large Synoptic Survey Telescope (LSST)] and space-based missions [Euclid, Joint Dark Energy Mission (JDEM)].

To get significant weak lensing measurements, we need first of all to overcome the noise induced by the intrinsic shape distribution of galaxies. This is done increasing the number density of target galaxies, observing with long exposures times and possibly with a broader filter than usual. The shape of these galaxies must then be measured in a precise and robust way, correcting for the blurring and the distortion produced by the instrumental point-spread function, and many efforts are being spent on this task (see e.g. the Great10 challenge, Kitching et al. 2011). Once a shear estimate is obtained for every galaxy, if we are able to locate it in the $z$-direction, the information included in the shape measurement becomes much more powerful. In this case, in fact, one can slice the galaxy catalogue in redshift bins and thus measure the variation of the matter power spectrum as a function of cosmic time. For this reason, we need a good redshift estimate for the observed galaxies, and they can only be obtained via photometric techniques for the vast majority of the sample observed by any of these experiments, because spectroscopical observations for such big and deep samples are not realistically achievable.

In this paper we will study the possibility of obtaining such precise redshift estimates for the galaxies observed by a space-based mission targeted to measure cosmic shear.
Most proposed space-based experiments include two main observing channels: optical imaging to measure the galaxy shapes, taking advantage to the stability of the PSF and the absence of atmospheric turbulence, and near-infrared photometry to characterise their spectral-energy distribution in a wavelength range where observations from the ground are very difficult due to the high background. In particular, as an example of a future weak lensing survey from space, we choose to simulate a Euclid-like setup for the space observations (Laureijs 2009). We will assess what is the achievable precision in the redshift estimation using these space data in connection with multi-band optical ground-based observations. This kind of study has been performed up to now by creating multi-band mock galaxy catalogues with MonteCarlo techniques, considering the nominal depth of the survey, and thus the expected error on the photometric measurements (Abdalla et al. 2008; Bordoloi et al. 2010). With this work we study in deeper detail this process, by using complete simulations, from the images of the galaxies in different observing bands to the photometric redshift estimation. In this way, the photometry of each galaxy has a more realistic uncertainty and we can take into account more sources of error in the measurements, such as those due to proximity of the galaxies. The simulations of observations are performed with the SkyLens simulator (Meneghetti et al. 2008, 2010), that provides images of galaxies under different observing conditions. To these images we apply up-to-date software to extract multi-band photometry for every detected galaxy. Then we submit the multi-band galaxy catalogue to the Bayesian algorithm BPZ (Benitez 2000) to obtain a redshift measurement. We first analyse how different parameters of the ground-based surveys (depth, wave-length coverage, seeing) can influence the photometric redshift quality. Then we make specific simulations to estimate the precision that can be obtained with some of the proposed ground-based surveys to be used in connection with space-based observations.

The paper is organised as follows. In Section 2 we present the way in which simulations of observations are performed, and how multi-band photometry and then a redshift estimate for each galaxy are extracted from them. Finally, we describe the way in which the galaxy catalogue can be cleaned to exclude objects that are likely to have a bad redshift estimate. In Section 3 we analyse the impact of different characteristics of ground-based surveys on the goodness of photometric redshifts. In Section 4 we show more specific results obtained simulating observations of future or ongoing ground-based surveys. In Section 5 we compare our results with those obtained by other authors, while a summary and concluding remarks are provided in Section 6.

2 METHODS

2.1 Simulations of observations

SkyLens (Meneghetti et al. 2008, 2010) is a software that creates realistic simulations of galaxy observations, considering characteristics of the telescope and of the detector and the effect of the atmosphere. It starts from a photometric and morphological database of ∼ 10000 real galaxies observed in the Hubble-Ultra-Deep-Field (HUDF) (Beckwith et al. 2006). A spectral type and a photometric redshift have been assigned to each galaxy (Coe et al. 2006), and this allows to re-observe it in different observing bands. To create a realistic observation, the program first builds a random distribution of galaxies taken from the HUDF catalogue. A flux in the desired band is assigned to each galaxy considering its observed magnitude, its spectral-energy distribution and its redshift. The shape of the galaxy is reconstructed using its decomposition in shapelets basis functions (Melchior et al. 2007), and rotated by a random amount. Then, from the surface brightness distribution of the galaxy, the counts in every pixel are calculated considering the quantum efficiency of the CCD, the transmission curve of the lenses, the mirror reflectivity and the atmosphere extinction. Poisson noise is added to the image, mimicking photon noise, read-out-noise and flat-field finite precision.

For the purpose of this work, we simulate the observation of the same patch of the sky with different instruments. As an example of future space-based survey, we choose to consider the Euclid experiment (Laureijs 2009), that consists of an optical imaging channel and a near-infrared photometry channel. In particular, we start from a 800′′ × 800′′ image of a simulated sky observed with a large optical imaging band analogue to the RIZ of the Euclid experiment (see Fig. 1). Simulations made with SkyLens of Euclid RIZ images are currently being used to predict the galaxy number density available for shear measurements from the observations (Meneghetti et al., in prep.). We underline that this broad band is proposed by the Euclid experiment to maximize the galaxy density used in cosmic shear measurements, but it is practically useless in photometric redshift determination, so the results of this work do not depend strongly on this choice. We then simulate the near-infrared coverage of the same sky, in particular considering three bands (Y J H). Characteristics of the filters and exposure times are taken from the Euclid proposal (Laureijs 2009). Finally, we simulate the observation of the same sky with different ground-based instruments and different observing conditions. In the following, we explain how we analyse these images to extract the redshift estimation. In Fig. 2 we show as an example the appearance of the same galaxy in different observing bands and conditions.

2.2 Multiband photometry

Our intent is to submit the images to software pipelines usually adopted for the analysis of real observations, to thoroughly simulate the extraction of a redshift estimate from images taken with different instruments. A fundamental step to get a good photometric redshift estimate is to correctly measure the colours from multi-band observations of the same galaxy. As the images have different PSFs and different levels of noise, measuring naively the magnitude in each band in a matched aperture is not appropriate. Instead, we use the Colorpro software, presented by Coe et al. (2006), that measures all the magnitudes in the same aperture, and then corrects them for the different PSFs. This software has been used in many experiments, including some that combine space-based and ground-based data, as in our case (e.g. Umetsu et al. 2010; Rafelski et al. 2009).

In particular, we use the image of the galaxy in the RIZ band as the detection image, where the galaxy is detected...
Figure 1. Simulation of an observation with Euclid in the RIZ band. The picture has a side of 400 arcsec. The exposure time is 1800 sec.

Figure 2. A galaxy observed with Euclid in the RIZ band (left), with Euclid in the H band (centre), and with Subaru in the i band (right).

and its aperture defined. To catch possible optical dropouts, we detect galaxies also in the H-band image, and we merge the two catalogues. Then the images in all the bands are remapped to the higher resolution of the RIZ image, and SExtractor MAG_ISO magnitudes are measured from these images, considering the apertures previously defined. We underline that all the apertures we use for the photometry are drawn from the space-based RIZ image, to take advantage of the better PSF, and are typically small. In fact, the distribution of apertures as a function of number of pixels has 95 as median, that corresponds to a radius \( \sim 0.5'' \) in case of a circular shape. To estimate the amount of flux lost in each band due to the worse seeing, the RIZ image is degraded to the PSF of each band, and its new MAG_ISO measured. The difference between this magnitude and the MAG_AUTO measured on the original RIZ frame is the desired correction. This procedure is rigorously correct only if the surface brightness distribution is the same in all bands. This is true by construction in our simulations, because the object morphology is not dependent on wavelength. All magnitudes stated in this paper are AB magnitudes.

An example of the precision in the colour measurements obtained applying Colorpro to our simulations is shown in Fig. 3 where we plot the mean offset of measured colours with respect to the input values. We considered combinations between the four space-based bands (RIZ Y J H) and the ground-based bands that will be used in the photometric redshift estimation (griz). For the purpose of this plot, we simulated the observation of a small field around one galaxy at a time, to avoid blending and pollution and thus check the robustness of the photometry estimation process in an ideal case. We see that in this conditions the mean bias is negligible and the standard deviation is below 0.1 magnitudes. We then focus on a single colour (g-J) and plot in Fig. 4 the error in the colour measurement as a function of the photometric uncertainty for each object in the sample. We see that most of the objects have very small values of bias and that the statistical standard deviation is in good accordance with the uncertainty quoted by the program. This means that the implemented method is able to compensate for the different PSFs and pixel sizes of the images.
2.3 Redshift estimation

With the procedure described above, we have obtained realistic multi-band catalogues, with luminosity measurements in the 4 space-based bands (RIZYJH) and in the proper ground-based optical bands. Now we use these catalogues in the proper estimation of photometric redshifts for the galaxies. As before, we perform this task applying a method that is developed for and usually applied to real data. Different techniques have been proposed to extract the redshift estimate of a galaxy from its multi-band photometry, such as template fitting (e.g. Bolzonella et al. 2000; Benitez 2000; Feldmann et al. 2006) and training-set methods (e.g. Collister & Lahav 2004; Way & Srivastava 2006). They are probed to obtain similar results (Abdalla et al. 2011), with the former being more flexible, because they do not need to be trained with a complete spectroscopic subsample. For this work we use the BPZ software (Benitez 2000), the implementation of a Bayesian template-fitting method, that has been applied to real data in many works (e.g. Mei et al. 2006; Menanteau et al. 2006). It requires as input a set of galaxy template SEDs, from which the program calculates the colours when observed through the considered filters at different redshifts. The probability of a galaxy with colours $C$ of being at redshift $z$ is then evaluated applying

$$p(z|C, m_0) \propto \sum_T p(z, T|m_0)p(C|z, T)$$

where $m_0$ is the magnitude used to compute the prior and $T$ is the index that runs over the templates. In particular, $p(z, T|m_0)$ is the prior given by the $i$ band distribution, as observed in the Hubble Deep Field North, while $p(C|z, T)$ is the likelihood of the colours given the template and the redshift, calculated via a $\chi^2$ approach,

$$-\log p(C|z, T) \propto \chi^2(z, T, \alpha) = \sum_\alpha \frac{(f_{\alpha} - a f_{T\alpha})^2}{\sigma_{f\alpha}^2},$$

where $f_{T\alpha}$, $f_{\alpha}$ and $\sigma_{f\alpha}$ are the theoretical flux of the template, the observed flux and the noise in the band $\alpha$, respectively, and $a$ is the normalisation of the template. The BPZ software has been proven to perform competitively with other methods in Abdalla et al. (2011). Further details and a little modification of the original algorithm can be found in the Appendix.

In our case, we know the SEDs of our simulated galaxies, so we give the ‘right’ template SEDs to the program. This set is formed by the six templates used by Benitez (2000), plus the two added by Coe et al. (2006) to fit the bluest galaxies of the HUDF sample. This is an ideal case, as in photo-z template-fitting algorithms the correct choice of the template set is one of the critical steps (Yee 1998; Rowan-Robinson et al. 2008). The input magnitude that BPZ uses to compute the prior is the one measured in the $i$ band by the appropriate ground-based instrument.

2.4 Catalogue cleaning

Photometric redshift estimation suffers from colours degeneracy for galaxies at distinct redshifts. Although the Bayesian prior implemented in Eq. [1] reduces this problem, e.g. avoiding that very brilliant galaxies are given high redshift estimates, it does not solve it completely. When two templates at different redshifts have very similar theoretical colours, photon noise will sometimes move the observed magnitudes of a galaxy towards the wrong template, and the redshift extracted from the algorithm will be catastrophically wrong. Because of this, to match the strong requirements on the precision in the photometric redshifts for cosmic shear surveys, we need to clean our catalogue from galaxies that are likely to have a wrong redshift estimate. A way to do so is looking at the full redshift probability distribution of the galaxy, instead of considering only the most probable value. Galaxies that have colours in regions of degeneracy, i.e. that lie near the theoretical colours of templates at different redshifts, will have a multi-peaked probability distribution. Galaxies that have a controversial photometry, for example because are polluted by brilliant neighbours, will not show a single strong peak, but a diffuse distribution with much of the probability outside the vicinity of the most probable redshift. Thus, measuring the integral of the redshift probability distribution of the galaxy in the proximity of the most prominent peak will give an indication of how certain is the redshift estimation. This can be done via the ODDS parameter calculated by BPZ. It is defined as

$$\text{ODDS} = \int_{z' - 2\Delta_z}^{z' + 2\Delta_z} p(z),$$

where $z'$ is the most probable redshift value, and $\Delta_z$ is a free parameter, set by default in BPZ to 0.067(1 + $z'$). ODDS is thus a measure of how much the distribution is concentrated around the principal peak. This is indeed a powerful way to assess the goodness of the estimate, as can be seen in Fig. [5] where we plot for each galaxy the error in the redshift estimate against the ODDS parameter calculated from its $p(z)$. The fraction of galaxies with a catastrophic redshift estimate decreases rapidly as ODDS increases.

If ODDS > 0.95, it means that the distribution is more peaked than a Gaussian with a width equal to $\Delta_z$. This may seem a reasonable cut for the catalogue, and has been used
in some works (Coe et al. 2006, Hildebrandt et al. 2008), while Erben et al. (2009) preferred a looser limit of 0.90. In the example of Fig. 3, we see that catastrophic failures become rarer for galaxies with an increasing ODDS value, and disappear for galaxies with ODDS > 0.80. In general, we need to trade off the amount of galaxies to be kept in the catalogue with the mean accuracy of the redshift estimation. In the following, we will try to optimise the selection, by looking at how the standard deviation of the redshift measurements varies as a function of the ODDS value used as a minimum threshold. An optimised selection is crucial in cosmic shear observations, as it is important to consider as many galaxies as possible to reduce the noise induced by the intrinsic shape of the galaxies.

In particular we will measure the precision of the redshift estimates calculating the normalised standard deviation

\[
\langle \sigma_{z} \rangle = \frac{1}{N} \left[ \sum_{i=1}^{N} \left( \frac{z_{i,\text{real}} - z_{i,\text{phot}}}{1 + z_{i,\text{phot}}} \right)^{2} \right]^{1/2},
\]

where the index \(i\) runs over the \(N\) galaxies of the sample in the redshift range \(0.3 < z < 3\). We fix the requirement on the precision in the redshift measurement to \(\langle \sigma_{z} \rangle < 0.05\), that is considered valid to extract cosmological information from cosmic shear experiments, such as Euclid (Laureijs 2009). We will measure how many galaxies we must eliminate from the sample via a selection based on the ODDS parameter to fulfil this condition. As an example, in Fig. 6 we plot the value of \(\langle \sigma_{z} \rangle\) as a function of the ODDS value used as a lower threshold, for one of the simulated observations. In practice, we continue to throw away galaxies from the sample according to their ODDS value until the normalised standard deviation of photo-z estimates goes below 0.05. In this procedure, we will unavoidably eliminate some of the galaxies with a good photo-z estimate, and we will as well keep in the catalogue some outliers, as the indication given by ODDS is obviously only valid in a statistical sense. Nonetheless, the mean standard deviation drops almost monotonically increasing the ODDS threshold, so in this way we are in the end able to extract a robust galaxy sample from the data.

For every survey setup our main result will be the ratio of the number of galaxies that are kept in the catalogue after the cut over the total number of galaxies in the initial sample. We will take into account only galaxies with a measured RIZ magnitude below the declared limit of the Euclid mission (24.5). Although with a Euclid-like setup we are able to go slightly deeper than this limit in the optical imaging observation, some tests we did including dimmer galaxies proved not to improve our results, probably because they suffer from poor infrared photometry, so we decided to concentrate on a RIZ-magnitude limited sample. To increase the statistic, we create 1000 random realisations of each catalogue via bootstrap resampling, and report the mean of the measurements.

### 3 Dependence on survey parameters

In order to obtain general results that do not depend on the specific instrument or survey, we first simulate some optical ground-based multi-band observations taken with a fictional ideal instrument. This instrument has a very large collecting area (diameter = 8.2 m), a pixel angular size of 0.2 arcsec and does not suffer from any losses due to mirrors and optics. The considered filters are the standard SDSS ugriz. The performances of this unrealistic setup are compared with the ones of real telescopes considered in this paper in Tab. 1. Given the exposure time in a certain band with the ideal telescope, it is easy to calculate what is the time needed to reach the same depth with another telescope, considering the efficiency and the area of the instruments.

We underline that all the results presented in the following are obtained using data from the 4 space-based bands (RIZ Y J H), plus the indicated ground-based data.

#### 3.1 Exposure time

The quality of the photometry, and thus of the photometric redshifts, depends primarily on the depth of the survey. In this view, we simulate a set of observations with a constant exposure time in each griz band. (The importance of the u band will be studied separately in the next Section.) The results of the simulation are shown in Fig. 7, where we show...
the standard deviation of the total sample (i.e. formed by all the galaxies with RIZ magnitude < 24.5, without any selection based on ODDS), and the fraction of galaxies that are kept in the catalogue after a selection analogue to Fig. 5 is performed. We see that the standard deviation of the whole sample does not improve significantly as the exposure time increases, but there is a great enhancement of the number of available galaxies after the cut is performed. This happens because the standard deviation of the total sample depends mainly on catastrophic failures due to redshift-colours degeneracies. These are only partially avoidable improving the quality of the photometry, because they are an intrinsic drawback of the photometric redshift estimation. Above this, we have a number of galaxies with poor photometry due to bright companions or to blending with close dim neighbours. In this case, the observed flux is the sum of two different objects, and thus the observed SED will likely not correspond to any of the two physical states. Luckily, the selection via the ODDS parameter allows us to reject these galaxies. The improvement in the fraction of available galaxies after the selection for longer exposures is then given by the more precise photometry that we obtain. Reducing the errors on the measured magnitudes, the non-pathological galaxies (i.e. not exactly on a redshift-space degeneracy and without close companions) will be assigned a redshift with a smaller uncertainty. Thus they will have a more peaked redshift probability distribution and a bigger ODDS parameter. This will avoid them from being confused with catastrophic failures when the selection is performed. We note that with the deepest exposure, 1000 sec, just the 14% of the total sample of galaxies brighter than 24.5 in the RIZ band must be excluded to meet the requirements in the redshift precision.

3.2 Observing bands

It is sometimes possible to extend the observed wavelength range using the u band. This band is more time consuming then the others due to the increased absorption of the atmosphere with respect to the redder wavelengths and to the lower quantum efficiency of the CCD detectors. It is then usually necessary to trade off the availability of data at shorter wavelength with the loss of depth in the other bands. To test the importance of u-band data, we run the BPZ algorithm on the same images of the previous section, but with an additional u-band observation of the same exposure time. The results are shown in Fig. 8.

We see that even the shorter exposure that includes the u band (50 sec, corresponding to a 10-σ depth equal to 23.5) gives a total redshift standard deviation lower than with the deepest exposure with the griz bands only. This result is not surprising as the power of the u band in photometric redshift estimation is widely recognised in the literature. In particular, it helps in identifying galaxies at low redshift (z < 0.5) thanks to the ultraviolet continuum, and the ones at z ∼ 3 when the Lyman break enters the blue side of the filter. The realisation of this can be seen in Fig. 9, where we plot the output redshift against the input one, both with and without using the u-band data. The reduction of catastrophic failures especially for galaxies at high redshift is very evident.

On the other hand, when considering the number of galaxies after the selection, we see that the results are closer to the ones with the griz bands only, thanks to the fact that many catastrophic failures are in that case pointed out and eliminated. The difference between the two setups is around 10% of the total number of galaxies for the whole range of exposure times tested. With the longest exposures, the ugriz configuration permits to use around 94% of the sample.

3.3 Seeing

A problem that influences the quality of the photometry in ground-based observations is the blurring of the images that occurs as the light travels through the atmosphere. This is
often quantified by the seeing, i.e. the mean FWHM of the observed image of an intrinsic point-like source as a star. The method used to perform these simulations allows us to study the influence of this parameter on the precision of the photometric redshifts. In particular, we simulated five different observations of 100 sec per filter with the ideal telescope, with a seeing varying from 0.4 to 1.2. Results are shown in Fig. 10.

As in the case of the observing time, the change in the precision of the whole sample is not significant, while much more evident is the improvement in the number of available galaxies after the cut is performed. In fact, the number of galaxies with a certain redshift estimate can change by a factor of two depending on the quality of the atmosphere of the telescope site. This of course depends on the leakage of light between different galaxies that occurs if the seeing is greater than the angular distance between them. This creates an observed SED without a corresponding galaxy template and thus with a poor redshift estimate.

In reality, many astronomical images are the sum of exposures taken in different moments and with different observing conditions, i.e. without a single seeing value that can be attributed to the whole image, as in our idealised work. Nonetheless, our results are a strong indication that good observing sites are required to perform the high-precision photometry needed for redshift estimation of deep and vast galaxy catalogues.

4 PLANNED SURVEYS

Different projects are considered as possible ground-based counterparts of space-based weak lensing observations. We underline that any space-based mission that will observe both the northern and southern sky will need collaboration from different projects covering different regions of the sky. In some sense, some of the projects we study will not be competitors against each other in practice, because any space-based project will need a patchwork of collaborations around the globe.

In this work we consider three ground-based surveys: the Dark Energy Survey (DES), Pan-Starrs and the Large Synoptic Survey Telescope (LSST). The first two are officially supporting the Euclid mission, while the third is a case we chose in order to explore the potential of observations that consider the u band, following the results of the previous Section. In Table 2 we show the exposure times and the depths of the three surveys, calculated with our simulator.

- **DES** is a project that will begin in late 2012 with the Blanco 4-meter telescope at Cerro Tololo Inter-American Observatory, on which a new camera named Decam will be mounted at the scope (Abbott et al. 2005). It will observe 5000 square degrees in the South Galactic Gap for 525 nights in 5 years.
- **Pan-Starrs** is a new telescope, consisting of four mirrors, one of which is already working at Mount Haleakaka in Hawaii (Kaiser et al. 2010). It will observe the 30000 square degrees visible from the site. The exposure times of Table 2 are calculated from the ones declared for the ongoing 3-years 3r survey (Magnier et al. 2007), multiplied by 10/3, to make them realistic for a 10-year project.
- **LSST** is a telescope that will start observing in 2017 from Cerro Pachon in Chile (Abell et al. 2009). It will observe the southern half of the sky. The observing times are the ones declared for the 10-year total observing project.

For each project, we perform the simulations as described in Section 2, keeping the seeing fixed at 0.8 arcsec. This is a reasonable value, that can be obtained in all three sites (Els et al. 2009; Schoeck et al. 2009; Tokovinin & Travouillon 2006). The results are given in Fig. 11.

© 2010 RAS, MNRAS 000.
As expected from the simulations with the ideal telescope, there is a big difference between the two griz instruments and LSST, when concerning the standard deviation in the redshift measurement of the whole sample. On the other hand, the number of galaxies available after the selection differs a lot also between the two griz instruments. The longer exposure times of Pan-Starrs make its results almost the best achievable for a griz survey, with more than 80% of the galaxies retained after the selection. The LSST-like survey is the only one that goes beyond 90%, as could be guessed by the simulations with the ideal telescope. A detailed plot of the redshift standard deviation as a function of the fraction of galaxies kept in the sample can be seen in Fig. 12. From this figure it is possible to evaluate a possible trade-off between redshift precision and galaxy fraction for the different instruments. We see that, even with the most severe selection, the redshift standard deviation remains above 0.03 in all three cases. This is due to very unfortunate overlapping of galaxies, that were detected as single objects by our routines. The redshift deduced by the combined photometry of the two galaxies is then compared to the one of the most brilliant object in the RIZ band, that is sometimes totally different. This kind of problem depends very mildly on the number of bands and on the quality of the photometry and has been only partially solved running a detection on the H-band image, in addition to the one on the RIZ image (See Sect. 2.2).

The trade-off between redshift precision and galaxy fraction is also strongly connected to the effectiveness of the procedure of removing galaxies that are likely to have a wrong redshift. We used a selection via ODDS, while other authors preferred using the 1-σ error quoted by the photometric redshift code (Abdalla et al. 2008), or investigating the presence of secondary peaks (Bordoloi et al. 2010). We tried these other methods, but the results were in general poorer that the ones we obtained with ODDS. If we were able to make an ideal selection, knowing perfectly what are the most wrong redshifts in the sample, we would need to remove ∼1% of the galaxies with the LSST data, and ∼5% with the DES and PanStarrs ones, to get to the required precision.

5 COMPARISON WITH PREVIOUS STUDIES

Our simulations are of course a simplified version of what is done to measure redshifts from multi-band real images, nonetheless they are more realistic than the simple creation of dirty catalogues. Thus, is not surprising that our results are in general worse than what is obtained by previous works done with that method (Abdalla et al. 2008; Bordoloi et al. 2010). For example, Bordoloi et al. (2010) find, for a DES-like survey, a standard deviation of the selected sample equal to 0.064 (so in fact higher than our threshold) after removing 12% of the galaxies. As can be seen in Fig. 12 we instead would need to eliminate ∼15% of the sample to get to the same precision. To better understand this difference, we made some tests using simulated catalogues created without going through image simulation. We calculated magnitudes in the necessary bands from the theoretical SEDs of the galaxies in our sample, and we cut the resulting catalogue at RIZ < 24.5. Then, we added Gaussian noise according to the observing depth for the Euclid-like survey and for the

![Figure 11](image1.png)

Figure 11. Quality of the photometric redshifts obtained with the different ground-based surveys. The solid line represents \( \langle \sigma_z/(1+z) \rangle \), while the dashed line represents the fraction of galaxies with observed RIZ magnitude < 24.5 that are available after a cut is performed to reduce the redshift standard deviation below 0.05.

![Figure 12](image2.png)

Figure 12. Redshift standard deviation as a function of the fraction of galaxies kept in the sample via a ODDS selection for the three simulated surveys.

As expected from the simulations with the ideal telescope, there is a big difference between the two griz instruments and LSST, when concerning the standard deviation in the redshift measurement of the whole sample. On the other hand, the number of galaxies available after the selection differs a lot also between the two griz instruments. The longer exposure times of Pan-Starrs make its results almost the best achievable for a griz survey, with more than 80% of the galaxies retained after the selection. The LSST-like survey is the only one that goes beyond 90%, as could be guessed by the simulations with the ideal telescope. A detailed plot of the redshift standard deviation as a function of the fraction of galaxies kept in the sample can be seen in Fig. 12. From this figure it is possible to evaluate a possible trade-off between redshift precision and galaxy fraction for the different instruments. We see that, even with the most severe selection, the redshift standard deviation remains above 0.03 in all three cases. This is due to very unfortunate overlapping of galaxies, that were detected as single objects by our routines. The redshift deduced by the combined photometry of the two galaxies is then compared to the one of the most brilliant object in the RIZ band, that is sometimes totally different. This kind of problem depends very mildly on the number of bands and on the quality of the photometry and has been only partially solved running a detection on the H-band image, in addition to the one on the RIZ image (See Sect. 2.2).
three realistic ground-based surveys. The resulting multi-band catalogues have been submitted to BPZ and analysed in the same way as those obtained from Colorpro on the simulated images. In Fig. 13 we show the results, namely the redshift standard deviation as a function of the fraction of galaxies kept in the catalogue. The results we obtain in this way are significantly better than the ones obtained with the complete simulation (Fig. 12), and are in close agreement with Bordoloi et al. (2010).

As can be seen from Figs. 12 and 13, the improvement we obtain with the simpler set of simulations is relevant especially in the cleaning procedure, that becomes more effective. This may depend on the fact that in this case the errors quoted in the photometry catalogue are by construction the real uncertainties on the magnitudes, while the procedure of multiband photometry when dealing with images can bring hidden errors, such as those due to the matching of images with different quality. As the errors in the photometry translate in the uncertainty in the redshift determination, the value of the ODDS parameter is certainly more trustable in the simpler case and cleaning the catalogue from possible catastrophic outliers is thus more straightforward.

The work of Abdalla et al. (2008) differs in that they use the neural-network code ANNz to compute the redshifts. This may explain some differences on the importance of different bands between this work and theirs. In fact, they quote a $\sigma_z$ (note: without dividing by $1+z$) of 0.307 for a DES survey, 0.274 for Pan-Starrs, 0.197 for LSST. We get 0.398 (0.359 with the ‘simple’ simulations) in the DES case, 0.400 (0.311) for Pan-Starrs, 0.146 (0.131) for LSST. Thus, they perform slightly better in the gri $z$ cases, while we do much better when the $u$-band is available. This reflects also in the fact that they are able to get a $\sigma_z = 0.113$, below the LSST case, for an ‘ideal’ optical survey, while we do not obtain such a result, as shown in Section 3. This is reasonably due to the completely different techniques adopted. Although in general they get comparable results, it is possible that in this case we are showing strengths and weaknesses on both sides. Of course a firm statement on this topic can be done only submitting the same photometric catalogue to differentphoto-z algorithms.

6 SUMMARY AND CONCLUSIONS

In this paper we investigated the quality of photometric redshifts that can be obtained for the galaxies observed by future space-based missions targeted to measure cosmic shear. In order to create a realistic setup for our work, we chose to simulate a Euclid-like mission, with imaging in an optical band and multi-band photometry in the near infrared (Laureijs 2009). Fulfilling the requirement on photometric redshift precision is a crucial step to reduce the error in cosmic shear measurement, and it is important as well in other scientific subjects that can be studied with the same data (e.g. galaxy clusters, galaxy evolution). To this end, the space-based survey will need the collaboration of ground-based multi-band observations to characterise the SED of the galaxies in the optical range. We studied this topic using complete simulations, from realistic images of a patch of the sky, obtained with Skylens (Meneghetti et al. 2008), through multi-band photometry, to the photometric redshift estimation itself, performed with a template-fitting method (Benitez 2000). This allowed us to consider many sources of uncertainty in the process of determining a galaxy photo-z. Our main aim was to determine what fraction of the galaxy sample observed by a Euclid-like survey must be discarded to achieve $<\sigma_z> < 0.05$, a precision considered suitable for cosmic shear precision measurements.

We first studied the impact of different parameters (depth, seeing, bands) on the quality of the redshift estimation. We found that:

• The number of good photo-zs increases as expected with the exposure time, as more and more galaxies acquire a good photometry that enables to assign to them a certain redshift. The improvement is not so relevant if we instead consider the mean precision of the whole sample, as some pathological galaxies (affected by light of neighbour objects, or by severe colour-redshift degeneracies) pollute the catalogue even with the deepest survey we simulated.

• Seeing is potentially as important as depth, as blurring of the images is a great source of noise in the photometry, especially for FWHM > 1 arcsec. Increasing the seeing from 0.8 to 1.2 arcsec, we need to remove twice the number of galaxies to reach the same precision in photo-z estimation.

• Having u-band observations seems very important to improve the goodness of photo-zs. Even a relatively shallow survey with a 10-$\sigma$ magnitude limit equal to 23.5 has a great impact on the redshift standard deviation of the total sample. Ultraviolet light is particularly relevant for galaxies at redshift below 0.5 and above 2.5 and resolves a lot of magnitude-colour degeneracies between these two ranges. The impact is not so strong when dealing with the fraction of available galaxies: in this case the improvement is around 10 % of the total sample in comparison with a corresponding gri $z$ survey.

Then, we moved to some more specific tests, aimed to study the results that would be obtained with some of the ground-based surveys that are currently considered as possible Euclid counterparts. We focused on three projects: DES, Pan-

---

Figure 13. As Fig. 12 but from mock catalogues obtained without going through image simulation. The points represent the results obtained by Bordoloi et al. (2010).
Starrs and LSST. DES and Pan-Starrs suffer the lack of u-band observations, and to clear the catalogue from the catastrophic failures and get to the required precision we need to eliminate respectively ~ 35 % and ~ 15 % of the galaxies of the sample. LSST performs much better than the other two, and seems to be the best, even if the farthest in time, counterpart to a space-based cosmic-shear mission.

With the aim of comparing our results with previous works, we made also some simpler simulations, creating catalogues with known photometric uncertainties without making mock observations. In this way, as expected, the results improve significantly, especially in the cleaning procedure. This means that the depth of the survey alone does not include all the information about the noise of the observation. Other uncertainties are added because of the proximity of galaxies, that pollute each other’s photometry, and to the necessity of matching images obtained with different observing conditions and telescopes. A possible solution to the problems caused by the proximity of galaxies would be to use some automatic algorithm to eliminate a priori objects that appear to be very close to each other or even completely blended (i.e., detected as a single object). This points towards the need of even more realistic simulations, where it will be possible to optimize the procedure of object detection and multi-band photometry extraction from the images, a topic we did not tackle in this work. For example [Kuijken et al. (2008)] proposed a more complex approach, that makes use of shapelets decomposition to model both the source and the point-spread function.

We want to stress here that we used a photometric redshift code that has performed well in other applications to real and simulated data, but we did not select it as the best in the treatment of the specific data set we applied it to, nor we optimised it in this view. Our sample cleaning method could maybe be improved too, e.g. tuning the integration limits in the ODDS calculation, or considering secondary peaks, as done by [Bordoloi et al. (2010)], or the 1-σ error quoted by the photometric redshift code, as done by [Abdalla et al. (2008)]. In general, our results should not be considered as optimal forecasts for a Euclid-like experiment, but more as a quantification of strengths and challanges for this kind of mission. In particular, the difference between results obtained with noisy catalogues (Fig. 13) and with image simulations (Fig. 12) hints at the importance of multi-band photometry extraction and a-priori object selection for this kind of measurement. Further testing and development of the algorithms is needed to understand what is the optimal way to extract photometric redshifts from this kind of observations and to optimise the methods for each single case.

Overall, with the cautions we underlined above, our work show that it is indeed possible to obtain precise (\(\frac{\sigma_{\Delta z}}{1+z} < 0.05\)) and robust photometric redshifts for more than 90 % of the galaxies with RIZ \(< 24.5\), in case u-band observations are available. In the case of griz ground-based observations, on the other hand, to get to the required precision we need to eliminate at least 15 % of the sample (See Fig. 12), depending on the depth of the survey.

**ACKNOWLEDGMENTS**

We acknowledge financial contributions from contracts ASI-INAF 1/023/05/0, ASI-INAF 1/088/06/0, ASI 1/016/07/0 COFIS, and ASI EUCLID-DUNE 1/064/08/0, ASI-Uni Bologna-Astronomy Dept. Euclid-NIS 1/039/10/0, and PRIN MIUR Dark energy and cosmology with large galaxy surveys. This work was sustained in part by the German Academic Exchange Service (DAAD) under the Vigoni programme. We thank G. Zamorani, H. Hoekstra, F. Abdalla, L. Guzzo, W. Percival and T. Kitching for the useful discussions.

**REFERENCES**

Abdalla F. B., Amara A., Capak P., Cypriano E. S., Lahav O., Rhodes J., 2008, MNRAS, 387, 969
Abdalla F. B., Banerji M., Lahav O., Rashkov V., 2011, MNRAS, 417, 1891
Abell P. A., Allison J., Anderson S. F., et al., 2009, arXiv preprint, astro-ph/0912.0201L

Beckwith S. V. W., Stiavelli M., Koekemoer A. M., et al., 2006, AJ, 132, 1729
Benitez N., 2000, ApJ, 536, 571
Bertin E., Arnouts S., 1996, A&A, 117, 393
Bordoloi R., Lilly S. J., Amara A., 2010, MNRAS, 406, 881

Bolzonella M., Miralles J.-M., Pello R., 2000, A&A, 363,476
Coe D., Benitez N., Sanchez S. F., Jee M., Bouwens R., Ford H., 2006, ApJ, 132, 926
Collister A. A., Lahav O., 2004, PASP, 116, 345
Dahlen T., Mobasher B., Dickinson M., et al., 2010, ApJ, 724, 425

Els S. G., Schoeck M., Bustos E., et al., 2009, PASP, 121, 922

Erben T., Hildebrandt H., Lercher M., 2009, A&A, 493, 1197
Feldmann R., Carollo C. M., Porciani C., 2006, MNRAS, 372, 565

Hildebrandt H., Wolf C., Benitez N., 2008, A&A, 480, 703
Kaiser N. et al., 2002, in Tyson J. A., Wolff S., eds, Proc. SPIE Vol. 4836, Survey and Other Telescope Technologies and Discoveries. SPIE, Bellingham, p. 154
Kitching T. et al., 2011, Annals Appl. Stat., 5, 2231
Kuijken K., 2008, A&A, 482, 1053
Laureijs R., 2009, arXiv preprint, astro-ph/0912.0914
Magnier E., 2007, in Sterken C., ed., ASP Conf. Ser. Vol. 364, The Future of Photometric, Spectrophotometric and Polarimetric Standardization. Astron. Soc. Pac., San Francisco, p. 153

Mei S., Holden B. P., Blakeslee J. P., et al., 2006, ApJ, 644, 759

Melchior P., Meneghetti M., Bartelmann M., 2007, A&A, 463, 1215
Menanteau F., Ford H. C., Motta V., Benitez N., Martel A. R., Blakeslee J. P., Infante L., 2006, AJ, 131, 208
Meneghetti M., Melchior P., Grazian A., et al., 2008, A&A, 482, 403
Meneghetti M., Rasia E., Merten J., Bellagamba F., Eltori S., Mazzotta P., Dolag K., Marri S., et al., 2010, A&A, 514, 93
APPENDIX A: CORRECTION TO BPZ ALGORITHM

In Section 3.2 of Benitez (2000), the author introduces the normalisation parameter $a$ of the template SED as a nuisance parameter in the Bayesian formalism. The probability of observing the colours $C$ given the template $T$ and the redshift $z$ is obtained by integrating the $\chi^2$ likelihood (see Eq. 2) over $a$, instead of considering only the best-fit value $a_{\text{min}}$ as in the simple maximum-likelihood approach. Thus, we have

$$p(C|z,T) \propto \int da \ p(a|m_0) \ p(C|z,T,a),$$

where $m_0$ is the magnitude value used as a prior. The prior of $a$ with respect of $m_0$, $p(a|m_0)$, is conservatively assumed flat. Evaluating the integral, he gets

$$p(C|z,T) \propto \frac{1}{2} \exp \left[ - \frac{\chi^2(z,T,a_{\text{min}})}{2} \right].$$

where

$$\chi^2(z,T,a_{\text{min}}) = \sum_{T \alpha} \frac{f_{T \alpha}^2}{\sigma_{T \alpha}^2} (\alpha - a_{\text{min}}).$$

Eq. A2 and A3 mean that the value of $p(C|z,T)$ depends over the ratios of the expected fluxes over the observed errors in the different bands and thus over the input normalisation of the templates (i.e., their value with $a = 1$). Without any indication on the way one should normalise $f_{T \alpha}$, this introduces an arbitrary factor in the calculation.

To overcome this problem, one can repeat the calculations assuming this time that the template is re-normalised at its best-fit parameter. We define our new templates $f'_{T \alpha} \equiv a_{\text{min}} f_{T \alpha}$ and we introduce the nuisance parameter $a' \equiv a/a_{\text{min}}$. In this way, we avoid any arbitrariness in the calculation, by defining the templates and the nuisance parameter in a physically motivated way. With the new definitions, Eq. A2 becomes

$$p(C|z,T) \propto \frac{1}{2} \exp \left[ - \frac{\chi^2(z,T,a'_{\text{min}} = 1)}{2} \right],$$

where $F'_{TT}$ is the analogue of $F_{TT}$ with the new template, and the value of $a'$ that maximizes the likelihood is unity by definition. We can now re-express Eq. A4 in terms of the original templates. We have

$$\sqrt{F'_{TT}} = \left( \sum_{\alpha} \frac{f_{T \alpha}^2}{\sigma_{T \alpha}^2} \right)^{1/2} = \left( \sum_{\alpha} \frac{a_{\text{min}}^2 f_{T \alpha}^2}{\sigma_{T \alpha}^2} \right)^{1/2}.$$