Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **421,** 2355–2367 (2012) doi:10.1111/j.1365-2966.2012.20468.x

CFHTLenS: improving the quality of photometric redshifts with precision photometry⋆

H. Hildebrandt,[1,2]† T. Erben,[2] K. Kuijken,[3] L. van Waerbeke,[1] C. Heymans,[4] J. Coupon,[5] J. Benjamin,[1] C. Bonnett,[6] L. Fu,[7] H. Hoekstra,[3] T. D. Kitching,[4] Y. Mellier,[8,9] L. Miller,[10] M. Velander,[3,10] M. J. Hudson,[11,12] B. T. P. Rowe,[13] T. Schrabback,[3,14] E. Semboloni[3] and N. Benítez[15]

[1]*Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC, V6T 1Z1, Canada*
[2]*Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany*
[3]*Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands*
[4]*Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*
[5]*Astronomical Institute, Graduate School of Science, Tohoku University, Sendai 980-8578, Japan*
[6]*Institut de Ciències de l'Espai, CSIC/IEEC, F. de Ciències, Torre C5 par-2, Barcelona 08193, Spain*
[7]*Key Lab for Astrophysics, Shanghai Normal University, 100 Guilin Road, 200234, Shanghai, China*
[8]*Institut d'Astrophysique de Paris, Université Pierre et Marie Curie – Paris 6, 98 bis Boulevard Arago, F-75014 Paris, France*
[9]*Institut d' Astrophysique de Paris, CNRS, UMR 7095, 98 bis Boulevard Arago, F-75014 Paris, France*
[10]*Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH*
[11]*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1, Canada*
[12]*Perimeter Institute for Theoretical Physics, 31 Caroline St N, Waterloo, ON, N2L 2Y5, Canada*
[13]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT*
[14]*Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305-4060, USA*
[15]*Instituto de Astrofísica de Andalucía, CSIC, C/Camino Bajo de Huétor 24, Granada 18008, Spain*

Accepted 2011 December 30. Received 2011 December 29; in original form 2011 November 16

ABSTRACT

Here we present the results of various approaches to measure accurate colours and photometric redshifts (photo-z) from wide-field imaging data. We use data from the Canada–France–Hawaii Telescope Legacy Survey which have been re-processed by the Canada–France–Hawaii Telescope Lensing Survey (CFHTLenS) team in order to carry out a number of weak gravitational lensing studies. An emphasis is put on the correction of systematic effects in the photo-z arising from the different point spread functions (PSFs) in the five optical bands. Different ways of correcting these effects are discussed and the resulting photo-z accuracies are quantified by comparing the photo-z to large spectroscopic redshift (spec-z) data sets. Careful homogenization of the PSF between bands leads to increased overall accuracy of photo-z. The gain is particularly pronounced at fainter magnitudes where galaxies are smaller and flux measurements are affected more by PSF effects. We discuss ways of defining more secure subsamples of galaxies as well as a shape- and colour-based star–galaxy separation method, and we present redshift distributions for different magnitude limits. We also study possible re-calibrations of the photometric zero-points (ZPs) with the help of galaxies with known spec-z. We find that if PSF effects are properly taken into account, a re-calibration of the ZPs becomes much less important suggesting that previous such re-calibrations described in the literature could in fact be mostly corrections for PSF effects rather than corrections for real inaccuracies in the ZPs. The implications of this finding for future surveys like the Kilo Degree Survey (KiDS), Dark Energy Survey (DES), Large Synoptic Survey Telescope or Euclid are mixed. On the one hand, ZP re-calibrations with spec-z values might not be as

⋆Based on observations obtained with MegaPrime/MEGACAM, a joint project of CFHT and CEA/DAPNIA, at the Canada–France–Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada–France–Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS.
†E-mail: hendrik@phas.ubc.ca

© 2012 The Authors
Monthly Notices of the Royal Astronomical Society © 2012 RAS

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

accurate as previously thought. On the other hand, careful PSF homogenization might provide a way out and yield accurate, homogeneous photometry without the need for full spectroscopic coverage. This is the first paper in a series describing the technical aspects of CFHTLenS.

Key words: galaxies: abundances – galaxies: high-redshift – galaxies: photometry.

1 INTRODUCTION

Estimating distances of celestial objects has always been one of the major technical aspects in observational astronomy. Whenever approximate redshifts of a very large number of faint extragalactic objects are needed the estimation of redshifts (and hence distances) from colours, also termed photometric redshifts (photo-z; see e.g. Baum 1962; Puschell, Owen & Laing 1982; Koo 1985, 1999; Loh & Spillar 1986; Connolly et al. 1995; Benítez 2000; Bolzonella, Miralles & Pelló 2000; Wolf, Meisenheimer & Röser 2001; Csabai et al. 2003; Collister & Lahav 2004; Ilbert et al. 2006; Hildebrandt, Wolf & Benítez 2008; Coupon et al. 2009; Hildebrandt et al. 2010), represents the only practical solution. Over the last few decades this technique has become increasingly important in extragalactic studies. Cosmological observations, inherently statistical in nature, particularly benefit from the availability of redshifts for millions of objects over large cosmological volumes.

An example is weak gravitational lensing (WL; for reviews see Bartelmann & Schneider 2001; Hoekstra & Jain 2008; Munshi et al. 2008), which has been established as an important tool to study the dark sector of the Universe. The first WL detection of a cluster of galaxies was made by Tyson, Wenk & Valdes (1990). Thanks to the steady progress in analysis tools and better knowledge of the source redshift distribution WL studies have become an important tool to calibrate the masses of individual galaxy clusters (Hoekstra 2007; Heymans et al. 2008; Okabe et al. 2010; Hoekstra et al. 2011; Jee et al. 2011). By stacking the signals of many lenses, the average properties of clusters (Johnston et al. 2007; Hildebrandt et al. 2011) and groups (Hoekstra et al. 2001; Parker et al. 2005; Leauthaud et al. 2010) or even galaxies (e.g. Brainerd, Blandford & Smail 1996; Hudson et al. 1998; Hoekstra, Yee & Gladders 2004; Heymans et al. 2006; Mandelbaum et al. 2006; Parker et al. 2007; Leauthaud et al. 2011; van Uitert et al. 2011) can be studied. Furthermore, the WL effect of the large-scale structure of the Universe, called cosmic shear (Bacon, Refregier & Ellis 2000; Kaiser, Wilson & Luppino 2000; van Waerbeke et al. 2000; Wittman et al. 2000; Hoekstra et al. 2006; Semboloni et al. 2006; Hetterscheidt et al. 2007; Massey et al. 2007; Schrabback et al. 2007, 2010; Fu et al. 2008), has been identified as one of the most promising probes of the effects of dark energy (Albrecht et al. 2006; Peacock et al. 2006).

Since the lensing signals observed in these cases depend directly on the distances of the lenses and sources, it is important to have an accurate knowledge of the lens–source geometry through knowing the redshifts of the objects. Modern WL surveys are designed in such a way that simultaneous measurements of the WL observables as well as photo-z values are possible. In particular, large, future imaging surveys that will cover a fair fraction of the extragalactic sky – like KiDS, DES, LSST and Euclid – require extremely accurate photo-z to reach a systematic accuracy in the WL measurements that does not compromise the survey's statistical power. Different ways of achieving this goal have been discussed in the literature.

In the foreseeable future all of these surveys will rely on ground-based multicolour photometry, which means that atmospheric effects have to be corrected for. In this paper we present advanced techniques to arrive at homogeneous photo-z from inherently inhomogeneous, ground-based survey data. For this we use the most powerful WL survey to date, the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS). Being a ground-based survey, the CFHTLS involves some unavoidable inhomogeneities, e.g., in terms of seeing, atmospheric extinction, etc. The CFHT Lensing Survey (CFHTLenS)[1] team was formed to provide a reduction and analysis of the CFHTLS data optimized for WL science and addressing these challenges. The higher level requirements for the data to measure accurate shapes and redshifts of tens of millions of galaxies made this 'lensing-quality' reduction necessary.

In this first paper of a series we present the multicolour photometry and the resulting photo-z upon which the future CFHTLenS science projects will be based. In Section 2 the CFHTLS data set and the CFHTLenS data reduction are presented and compared to the public data available from the TERAPIX team. Section 3 deals with the crucially important correction for atmospheric effects needed for accurate multicolour photometry. In Section 4 our strategy to extract catalogues from the images in five different photometric bands is described. The photo-z estimation is then presented in Section 5 along with the results of the comparisons between photo-z and spectroscopic redshifts (spec-z). Conclusions are presented in Section 6.

2 DATA SET

2.1 The CFHTLS-Wide

The Wide component of the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS-Wide) commenced in mid-2003 and completed observations in early 2009. In more than 2300 h of dark and grey time over these five and a half years, the CFHTLS-Wide imaged 172 one square degree MEGACAM fields in five filters u, g, r, i, z to a 5σ point source limiting magnitude of $i_{\rm AB} \approx 25.5$. The data span four contiguous fields: W1 ($\sim$63.8 square degrees), W2 ($\sim$22.6 square degrees), W3 ($\sim$44.2 square degrees) and W4 ($\sim$23.3 square degrees) totalling 154 square degrees once the overlap regions are accounted for. W1, W2 and W4 are equatorial fields with W1 and W4 containing the VIMOS VLT Deep Survey (VVDS) and VIMOS Public Extragalactic Redshift Survey (VIPERS) spectroscopic surveys. W3 is a northern field containing the extended Groth Strip DEEP2 spectroscopic survey. A detailed report of the full CFHTLS Deep and Wide surveys can be found in the TERAPIX CFHTLS T0006 release document (Goranova et al. 2009).

The CFHTLS-Wide was optimized for the study of WL for which the crucial observables are of the shape of resolved galaxies as well as their redshifts. The observing strategy was therefore to reserve the best-seeing conditions with $\theta < 0.8$ arcsec for the lensing i-band filter and follow-up with the other bands in the poorer seeing

[1] http://www.cfhtlens.org

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

conditions. That is also the reason why the i-band is our primary object detection band (see Section 4).

2.2 The CFHTLenS data reduction

The data reduction was conducted with the THELI pipeline (Schirmer et al. 2003; Erben et al. 2005) following the procedures outlined in Erben et al. (2009). We briefly summarize the most important differences in data processing between the CARS project, detailed in Erben et al. (2009) and Hildebrandt et al. (2009), and the current CFHTLenS data set. A more detailed description will be given in another publication of this series.

The CFHTLenS project makes use of the complete CFHTLS-Wide data set. This includes five-colour coverage of 172 square degrees of high-quality data subdivided into four patches W1–W4 (see Section 2.1). In addition we make use of the CFHTLS Pre-Survey which densely covers the complete survey area with shortly exposed r-band images. This Pre-Survey was acquired to optimize the astrometric calibration for the main science data. Similarly, to improve the photometric calibration of the CFHTLS, the survey area was (re-)observed in a sparse grid under photometric conditions during the year 2008 (CFHT program RunIDs 08AL99 and 08BL99).

All CFHT MEGACAM images are initially processed using the ELIXIR software at the Canadian Astronomical Data Centre (CADC) and it is this archived data that we use in the CFHTLenS project. The current work and all other CFHTLenS publications use the CFHTLS-Wide images, the astrometric Pre-Survey data and additional photometric data that were available at CADC on 2009 January 15. In total, the set contains 7997 ELIXIR processed CFHT MEGACAM images.

While we processed the data on a per-pointing level in the CARS project, we performed all calibrations on a per-patch level for CFHTLenS. The inclusion of all available data, especially the astrometric Pre-Survey and the photometric (re-)calibration, significantly improved the homogeneity of our data. Most important for this work is our improvement in photometric calibration. In Erben et al. (2009) we quoted the RMS uncertainty of our relative photometric calibration between fields as $\sigma_{\mathrm{abs},g'r'i'} \approx 0.01$–$0.04$ mag, $\sigma_{\mathrm{abs},z'} \approx 0.03$–$0.05$ mag and $\sigma_{\mathrm{abs},u^*} \approx 0.15$ mag. In CFHTLenS we now reach $\sigma \approx 0.01$–0.03 mag in all passbands. Fig. 1 compares magnitudes in the CFHTLenS field W1p1m1 with the magnitudes taken from the fifth data release of the Sloan Digital Sky Survey (SDSS-DR5; Adelman-McCarthy et al. 2007). For comparison the plot also shows this comparison for the corresponding CARS field (see also fig. A.7 of Erben et al. 2009).

2.3 The TERAPIX T0006 catalogue

The first reduction of the CFHTLS-Wide data was performed by TERAPIX.[2] The data were distributed in several public releases. The most recent release, T0006 (2010 November 15; Goranova et al. 2009), includes photo-z estimated in the same way as described in Coupon et al. (2009) for the T0004 release (see also Coupon et al. 2011, for more details on the T0006 photo-z) with the LE PHARE photo-z code (Arnouts et al. 2002; Ilbert et al. 2006). However, the photometric calibration improved significantly in the latest release and now covers the whole CFHTLS-Wide area. We will compare

[2] http://terapix.iap.fr/

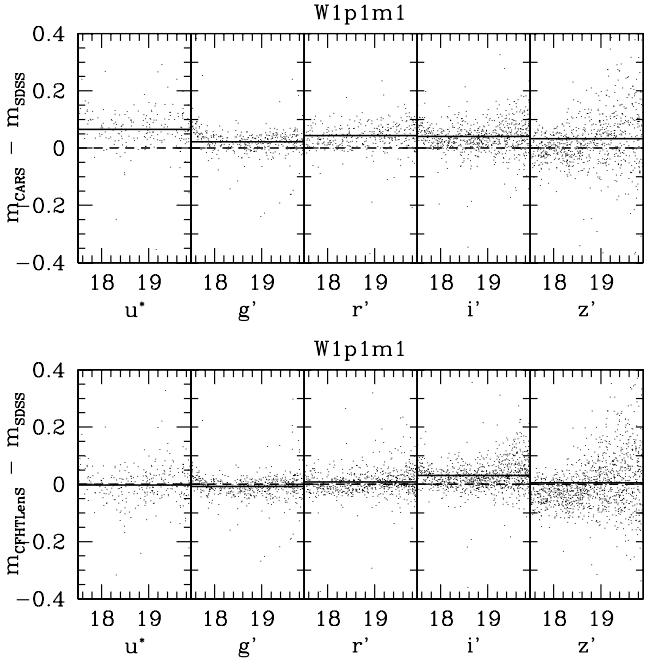


Figure 1. Photometric comparison between SDSS-DR5 and CARS (top; see Erben et al. 2009) and SDSS-DR5 and CFHTLenS (bottom) in the five different bands in one field of ≈1 square degree. The solid horizontal lines represent the average difference over this magnitude interval.

photo-z from the T0006 catalogues to CFHTLenS in the following sections.

The most notable differences in the multicolour photometry compared to CFHTLenS are as follows.

(i) The T0006 images are median stacks whereas the CFHTLenS images are mean stacks (Erben et al. 2009). While median stacks allow for easier rejection of image defects (cosmic rays, reflections, asteroids, etc.) mean stacks are less noisy and hence lead to deeper images and more precise photometry.[3]

(ii) T0006 does not implement point spread function (PSF) homogenization (a procedure correcting for the different seeings in the different bands as well as at different positions on the image described in Section 3) before catalogue extraction.

(iii) T0006 measures magnitudes in elliptical apertures (Kron 1980) instead of isophotal apertures, which we use for CFHTLenS (see Section 4).

(iv) The photometric zero-points were re-calibrated with the help of a large sample of spec-z values. See Section 5.3 for a detailed discussion of the benefits and dangers of such a re-calibration.

3 PSF HOMOGENIZATION

For photo-z we are particularly interested in measuring accurate colours of objects. For point sources this is a clearly defined problem. For extended objects, however, there is no unique definition of a colour and it depends on circumstances as to which definition is the

[3] A lot of effort was invested to automatically mask image defects in the CFHTLenS data reduction prior to the stacking. This leads to very clean images which can be averaged directly superseding the more noisy median (or also trimmed-mean) procedures. The standard error of the median is ≈25 per cent larger than the standard error of the mean, if the noise distribution is Gaussian, leading to a loss of ≈0.24 mag of depth.

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

most useful. For our purpose, to use the colours to estimate photo-z (see Section 5) we need colours that best match those modelled from spectral energy distribution (SED) templates. As we explain below this requires identical physical apertures[4] on the sky in the different bands. An extreme choice would be to define colour as the difference in total magnitude of an object in different bands, but since objects can have very different physical extent in different bands, this approach could lead to vastly different colours compared to the matched apertures, and is not optimal in terms of a signal-to-noise ratio (S/N).

When estimating photo-z we compare the observed colours of an object to colours modelled from the convolution of SED templates with instrument response curves. These SED templates are either based on observations (empirical templates) or on synthetic stellar evolution models (synthetic templates). Both approaches yield model colours for a particular stellar population, and it is therefore important that our observations yield colours that correspond to the same set of stars in each band, i.e. that they represent the same physical aperture on the galaxy.

Matching the physical apertures is complicated by the fact that different images taken through different filter bands will invariably show different degrees of blurring, expressed as the PSF and measurable as the observed shape of the stellar images in the field. These PSF differences can arise from different ambient conditions during the observations, or from chromatic effects in the atmosphere and optics of the telescope. Regardless of their origin they have the effect that identical apertures in different images of the same part of the sky do not represent the same physical part of a source. Not only different PSF sizes (seeing), but also the more subtle effects of PSF anisotropy, especially with a prime-focus camera like MEGACAM, will affect the photometry. Compensating for the PSF differences between bands can be done in two ways: either by adapting the apertures to compensate for the different PSFs or by manipulating the images so as to arrive at images with the same PSFs. Here we will follow the latter approach, because it is independent of the object size and leads to more homogeneous photometry across the survey. This involves constructing a suitable pixel convolution kernel for each image.

3.1 Constant Gaussian convolution kernel: *global*

In this method we assume for simplicity that the PSF can be described by a single Gaussian with width $\sigma_{\rm PSF}$. Under this assumption one can convolve an image in band X with a two-dimensional Gaussian filter function of width

$$\sigma_{{\rm filter},X} = \sqrt{\sigma^2_{\rm PSF,worst} - \sigma^2_{{\rm PSF},X}}\,, \quad (1)$$

to arrive at an image with a PSF size which matches the PSF size of the image with the worst seeing in a given set, $\sigma_{\rm PSF,worst}$. This method implies that the PSF size does not change with position on an image (or at least that the variation with position is the same in all bands),[5] an assumption that is not necessarily true for contemporary wide-field imaging cameras like MEGACAM on the CFHT. In the following we will call this approach *global* PSF homogenization. It is the same approach as used in Hildebrandt et al. (2006, 2007, 2009) and Erben et al. (2009) and is computationally straightforward.

3.2 Gaussianization of the PSF with a spatially varying kernel: *local*

In the *local* approach we drop the assumptions that the initial PSF is Gaussian and position independent, and construct a convolution kernel that is designed to make the PSF Gaussian everywhere, with the same width. Obviously this requires a non-Gaussian convolution kernel that changes with position on the image. Here we model the PSF and convolution kernels using the shapelet formalism (Refregier 2003; Refregier & Bacon 2003; Kuijken 2006) – essentially, each source is described as a sum of two-dimensional Gauss–Hermite functions. This method is closely related to the GAAP photometric package described in Kuijken (2008).

First, a catalogue is extracted from the i-band image with the SEXTRACTOR software (Bertin & Arnouts 1996). Stars are selected in a magnitude versus size diagram from the i-band data. In each of the five bands, the images of these stars are then modelled by sums of shapelets up to tenth order, through least-squares fitting of the pixel values. The size of the PSF, which is needed to scale the shapelets, is also taken from the initial SEXTRACTOR catalogue.

Next two-dimensional, fifth-order polynomials are fitted to the coefficients of the shapelet expansion to describe the variation of the PSF over the MEGACAM field in each band. From these high-order analytic descriptions of the PSF variations in each of the five bands, the convolution kernel is then created. The goal is to produce five images which have exactly the same Gaussian PSF over the whole field. To avoid deconvolution and the associated noise amplification, the target size for the final Gaussian PSF is chosen to match the largest PSF size found in any position in any of the five bands of a field. Within the shapelet formalism, it is easy to calculate a convolution kernel that transforms the modelled PSF into the Gaussian target PSF at each position in each band, and to perform the convolution in Fourier space. More details can be found in Appendix A.

In the following we will call this approach *local* PSF Gaussianization.

3.3 General remarks

It is clear that the benefits of PSF homogenization are most pronounced if the changes in PSF size (or also PSF shape) are large between bands. We will compare both the *global* and the *local* approaches to the case where no PSF homogenization is performed. This latter case will be referred to with the label *none*. See Table 1 for a summary of the different schemes to homogenize the PSF. The photo-z in the T0006 catalogues (see Section 2.3), which were estimated from multicolour photometry on median stacks that had not been corrected for PSF effects, will be referred to by the label *T0006*.

4 CATALOGUE EXTRACTION

Multicolour catalogues are extracted from a set of PSF homogenized images in the $ugriz$ filters using SEXTRACTOR in dual-image mode. The procedure is identical to the one presented in Erben et al. (2009) involving six SEXTRACTOR runs. The unconvolved (i.e. non-PSF-homogenized) i-band image is used as the detection image in all six runs. Five runs are performed with the convolved

[4] Here the term 'physical' refers to apertures which cover the same physical parts of an object in different bands, i.e. the same area on the sky before blurring by the atmosphere.

[5] However, it should be noted that PSF effects can have a collective effect changing the average properties of galaxies as a function of position if they are not corrected for. We defer an analysis of these PSF-photo-z correlations to a forthcoming paper.

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

Table 1. Different schemes to homogenize the PSFs and their properties. Note that the PSF size at a random position in the *global* scheme is only approximately constant between bands because the intrinsic PSF size changes from the centre to the edge of the MEGACAM mosaic. Using a constant convolution kernel for PSF homogenization between bands will only correct the PSF in one position (here chosen to be the centre of the image) and leave residuals at other positions. However, as we show in Section 5.2, these residuals are small enough to make the photo-z based on the *global* photometry superior to the ones based on the *none* or *T0006* photometry (but not as accurate as the ones based on the *local* photometry).

Scheme	PSF size constant between all five images in image centres	PSF size constant between all five images at any other position	PSF size & shape constant independent of position in all five images	PSF shape Gaussian in all five images
none/T0006	X	X	X	X
global	√	approximately	X	X
local	√	√	√	√

(i.e. PSF-homogenized) $ugriz$ images. One additional run with the unconvolved i-band image as detection as well as measurement image is performed to estimate total i-band magnitudes. This latter run is necessary because the isophotal apertures, which are used for flux measurements and hence colour estimation, are defined on the detection image by SEXTRACTOR. Using a measurement image with a larger PSF leads to flux leaking outside the aperture and therefore underestimated fluxes (overestimated magnitudes). Those biased fluxes are optimal for colour measurements because they correspond to the same physical parts of an object (see Section 3 above) but cannot readily be used to estimate total magnitudes. However, with the total i-band magnitudes estimated reliably in the sixth SEXTRACTOR run and the accurate colour measurements it is still possible to arrive at total magnitude estimates in all bands, e.g., for a band X:

$$X_{\rm tot} = i_{\rm tot} + (X - i)\,, \tag{2}$$

where $(X - i)$ is the corresponding colour index.

The multicolour photometry is complemented with the following quantities (for details see Erben et al. 2009):

(i) limiting magnitudes estimated from the local sky-background around an object,

(ii) extinction values extracted from the Schlegel, Finkbeiner & Davis (1998) maps at the object's position and

(iii) masks based on algorithms to reject regions of low S/N, haloes and diffraction spikes of bright stars and asteroids. These automated masks are then inspected by eye and further modified if necessary by people in the survey team.

4.1 Photometric quality control and creation of the mosaic catalogue

Before photometric redshifts are estimated (see Section 5) several tests are performed with the multicolour photometry alone to ensure the integrity of the data. The surface densities of objects in a fixed magnitude range are analysed and the object magnitude number counts are inspected on a field-by-field basis. The sky distributions of the same objects are plotted and eye-balled and distributions of quantities like the half light radius or the position angle of galaxies for each field are plotted as well as colour–colour diagrams of stars (selected by size and magnitude).

In order to arrive at a homogeneous mosaic catalogue we define hard cuts in right ascension and declination for each field. As the boundary between two fields we choose the mean of the extremal positions of objects taken from one field and from a neighbouring field. This method ensures that no celestial object appears more than once in our mosaic catalogue.[6]

5 PHOTOMETRIC REDSHIFTS

Photo-z is estimated with the BPZ code (Benítez 2000; Coe et al. 2006), one of the most widely used photo-z codes. Hildebrandt et al. (2010) tested several codes against simulated and real data showing that BPZ is amongst the most accurate codes when combined with the best available SED templates.

In the photo-z estimation process we properly take account of objects that are not detected in one or more of the $ugrz$ bands (identified by magnitude estimates fainter than the limiting magnitudes, which can happen when SEXTRACTOR is run in dual-image mode). As in Erben et al. (2009) and Hildebrandt et al. (2009) we use the recalibrated template set of Capak (2004).[7] Compared to Erben et al. (2009) we implemented some changes to improve the low-redshift behaviour of photo-z explained in the following.

5.1 Modifications of the prior

The Bayesian approach used by BPZ encompasses the calculation of the redshift likelihood and its subsequent multiplication with a prior to yield the posterior probability of an object having a certain redshift given the data. For noisy data the peak in the likelihood always has a finite width. For a $z = 0$ object the following happens: the likelihood function, which extends to redshifts $z > 0$ because of its finite width, is multiplied with a steep prior that behaves like $P(z) \approx z^{\alpha_t}$ for $z \ll 1$, with $0.9 \lesssim \alpha_t \lesssim 2.5$. In particular, $P(z = 0) = 0$ so that the peak of the posterior probability distribution is always at $z > 0$. Since BPZ picks the peak of the posterior as the photo-z estimate, this leads to a systematic overestimation of photo-z values at low redshift, whenever there is an appreciable number of low-z objects with limited S/N (and hence broad likelihoods) to make this effect visible. Similar biases of different severity can be seen in the Bayesian photo-z values of e.g. Csabai et al. (2003; fig. 7), Benítez (2000; fig. 7) and Coupon et al. (2009; figs 3 and 5). This behaviour is a general problem of template-based Bayesian redshifts. Here we present an ad hoc solution, but emphasize that further research is needed on this.

[6] Note that there could still be objects which appear in two neighbouring fields because of astrometric errors. But their number is so low that we do not account for this here.

[7] This template set is very similar to the one used for the T0006 catalogues.

© 2012 The Authors, MNRAS **421,** 2355–2367
Monthly Notices of the Royal Astronomical Society © 2012 RAS

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

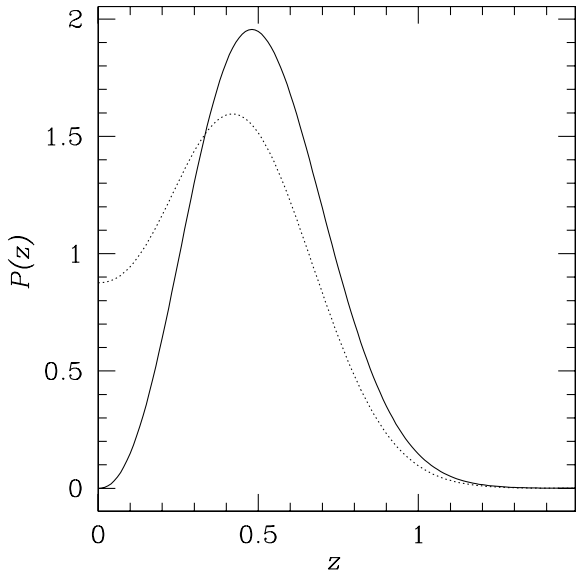


Figure 2. Redshift prior before (solid) and after (dotted) modification, here shown for an elliptical galaxy with $i = 20$.

To attenuate low-redshift bias we run BPZ twice, first with the original prior from Benítez (2000),

$$P(z) \propto z^{\alpha_t} \exp\left[-\left(\frac{z}{z_{mt}(m_0)}\right)^{\alpha_t}\right] \quad (3)$$

and again with a modified prior,

$$P_{\rm mod}(z) \propto (z^{\alpha_t} + 0.05) \exp\left[-\left(\frac{z}{z_{mt}(m_0)}\right)^{\alpha_t}\right]. \quad (4)$$

This modified prior no longer vanishes for $z = 0$, but levels off. Fig. 2 illustrates this modification. We keep the result of the first run unless the most probable redshift is below $z = 0.1$ and the ODDS parameter (i.e. the fraction of the integrated probability included in the primary peak of the posterior probability distribution; see Benítez 2000) associated with this solution is smaller than 0.8 indicating a possibly biased redshift. We found that this ad hoc modification yields an improved low-redshift performance when comparing photo-z to spec-z (see the next section). The choice of adding 0.05 in the first term of equation (4) yields the best results for this particular data set. But data with different noise properties might require different modifications. As mentioned above these prior modifications are a subject of ongoing research. It will be important to find robust, self-consistent ways of finding the optimal modification, possibly based on a large, spectroscopic, low-z training set.

5.2 Photo-z accuracy

We estimate the accuracy of the CFHTLenS photo-z by comparing them to spec-z from the VVDS (Le Fèvre et al. 2005; Garilli et al. 2008) and the DEEP2 galaxy redshift survey (Davis et al. 2007) which overlap with 20 fields of the CFHTLS-Wide. The deepest spectroscopic fields contain objects down to $i = 24$, however with an increasing incompleteness for fainter magnitudes. We also add spec-z from the SDSS-DR7 (Abazajian et al. 2009) if available in those 20 fields. We do not add the SDSS spec-z in the other $\sim$90 CFHTLS-Wide fields that overlap with SDSS because they do not contain significant additional information due to their low redshift. In Fig. 3 the photo-z for the three different PSF homogenization methods shown against the spec-z is. Visibly the photo-z accuracy improves with increasing sophistication of the PSF homogenization.

We limit the analysis to galaxies, as identified by the spec-z surveys. The stellar spectra will be used in Section 5.5 to develop criteria for star–galaxy separation. It should be noted that the spec-z catalogues used here are incomplete for $i \gtrsim 22$ and might paint too positive a picture of the photo-z accuracy (see Hildebrandt et al. 2008, 2010, for extensive discussions of these effects). In Section 6 we discuss strategies to acquire a complete picture of the photo-z accuracy that will be presented in Benjamin et al. (in preparation).

For each object with a reliable spec-z measurement (quality flags 3, 4, 23 and 24 for VVDS; quality flags 3 and 4 for DEEP2; quality flag 3 for SDSS) we calculate the quantity $\Delta z = \frac{z_{\rm phot} - z_{\rm spec}}{1+z_{\rm spec}}$, where $z_{\rm phot}$ is the peak of the posterior probability distribution. Objects with $|\Delta z| > 0.15$ are regarded as outliers.[8] For a given sample we then calculate the mean of Δz and the standard deviation around this mean, which corresponds to the photo-z scatter. This is done after outliers have been excluded. Furthermore we report the total bias of the sample, i.e. the mean of Δz including the outliers. We would like to stress that these three numbers (outlier fraction, scatter, bias) are not independent of each other and cannot reflect the full error distribution, which is highly non-Gaussian.

These statistics are calculated for different narrow i-band magnitude bins as well as for different narrow redshift bins (with a pre-selection of $19 < i < 24.5$) for each of the three different PSF homogenization approaches, *none*, *global* and *local*, as well as for the *T0006* catalogues. Errors are calculated assuming Poissonian shot noise. It should be noted that there is non-negligible correlation between the errors in neighbouring magnitude/redshift bins.

Fig. 4 shows the photo-z accuracy as a function of the i-band magnitude and redshift for the different methods. The effects of the PSF homogenization can be clearly seen. While the performance at bright magnitudes is similar in all methods the photo-z scatter and outlier rates at fainter magnitudes for the two methods without PSF homogenization (*none* and *T0006*) are larger than for the methods with PSF homogenization (*global* and *local*). Fainter objects are also smaller on average and hence their shape is more strongly dominated by the PSF. Not correcting for PSF effects biases their colours and leads to less accurate photo-z values.

Looking at the accuracy as a function of redshift shows that PSF homogenization leads to greater accuracy over the whole redshift range. The effect is more pronounced at higher redshifts, but since there are also many faint low-redshift galaxies the low-z statistics for *global* and *local* are generally better than for *none* and *T0006* as well.

The *global* and *local* schemes show similar redshift accuracy. At redshifts $z \gtrsim 0.9$, *local* shows somewhat reduced photo-z scatter, but the differences are small. This finding supports the hypothesis from Section 3.3 that although the PSF still varies from centre to edge in the *global* images (unlike the *local* images) these variations are similar in all bands and do not lead to strong colour biases.

The photo-z bias is non-negligible for most magnitudes and redshifts regardless of the method. With typical template-based photo-z methods it is very hard to suppress this bias without

[8] This choice is arbitrary and is mainly used for historical reasons since many photo-z studies in the past have adopted the same definition of an outlier. With a typical photo-z scatter of $\sigma \sim 0.04$ it ensures that only $\approx 4\sigma$ outliers are counted.

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

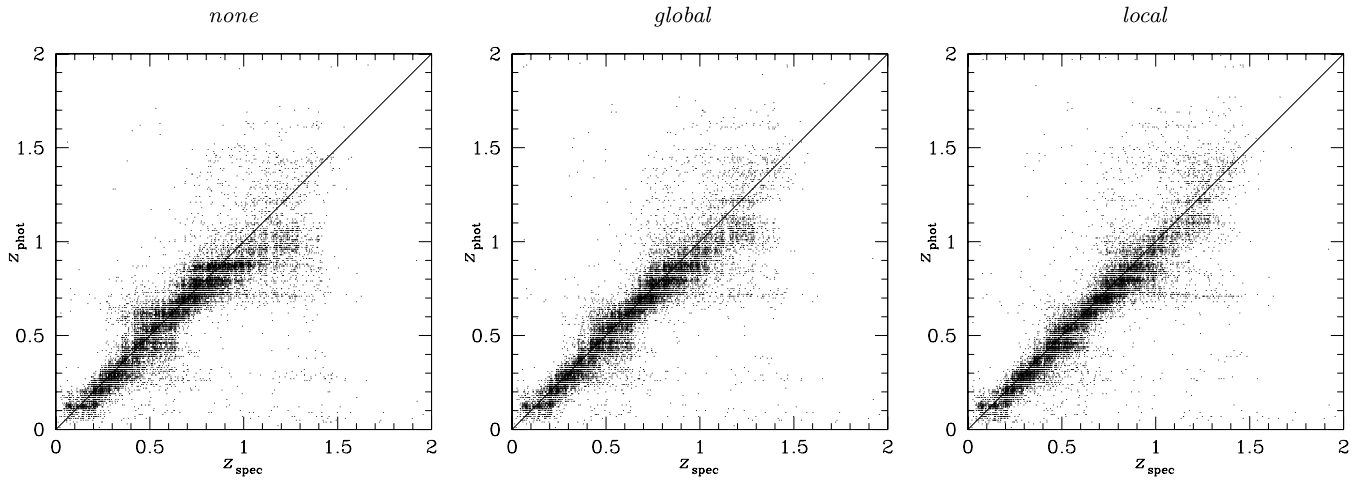


Figure 3. Photo-z versus spec-z for the three different ways of homogenizing the PSF. Shown are all objects with secure spec-z values (from VVDS, DEEP2 and SDSS) of the 20 fields with VVDS or DEEP2 overlap. No magnitude or ODDS cut is applied.

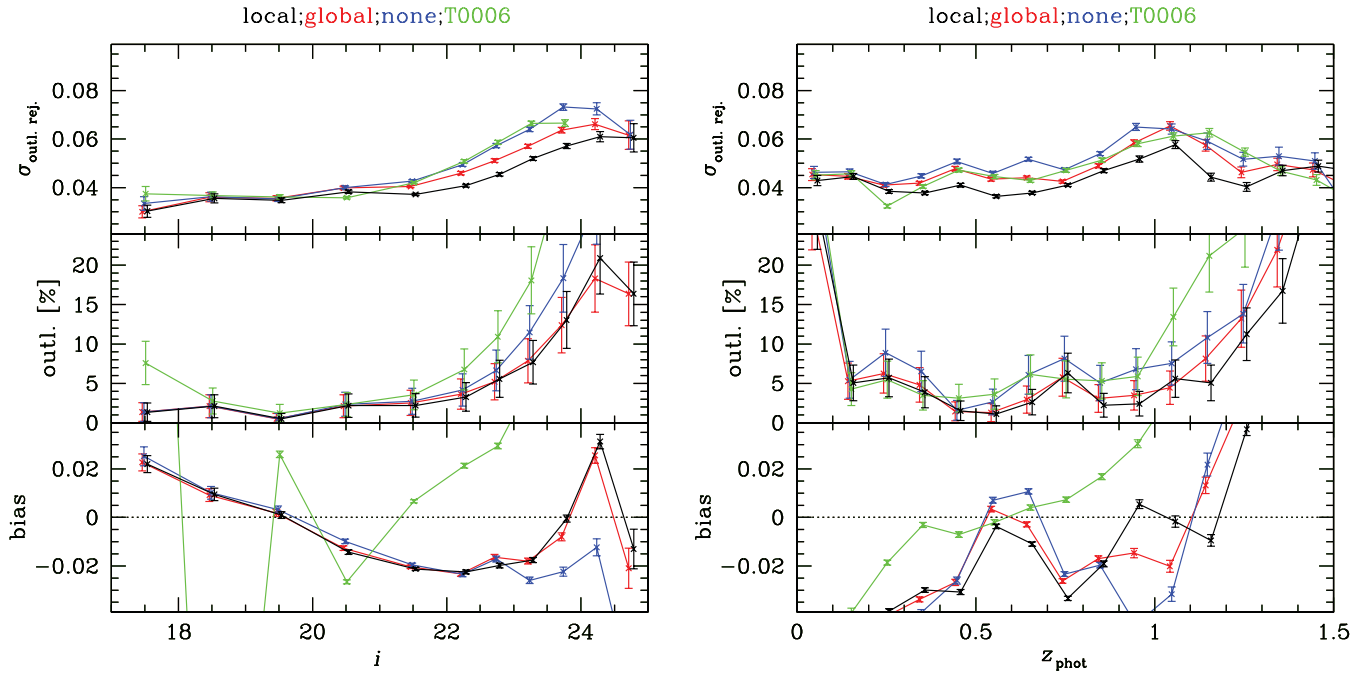


Figure 4. Photo-z statistics as a function of magnitude (left) and redshift (right) showing the effects of different ways of homogenizing the PSF. While *none* and *T0006* correspond to no PSF homogenization, *global* corresponds to a constant Gaussian kernel used for the convolution of the image in one band and *local* corresponds to a non-Gaussian, spatially varying kernel that leads to the same Gaussian PSF over the whole field. The top panel shows the photo-z scatter after outliers were rejected, the middle panel shows the outlier rate and the bottom panel shows the bias (outliers included; positive means photo-z values overestimate the spec-z values). Errors are purely Poissonian. Note that the errors between magnitude/redshift bins are correlated.

introducing larger scatter or more outliers. The strategy for scientific studies using such photo-z values must be to properly calibrate and account for this bias in the analysis. We do not correct for the bias at this stage.

In Fig. 5 we present the photo-z accuracy for the different SED templates as determined by the photo-z code. Similar to Ilbert et al. (2006) we find that the accuracy for elliptical galaxies and spiral galaxies is very similar. Only actively star-forming galaxies show a degraded photo-z accuracy. It is obvious that the accuracy for ellipticals/spirals suffers considerably once the 4000 Å-/Balmer-break starts to leave the filter set at $z \approx 1.1/1.3$.

5.3 Re-calibration of the photometric zero-points

It has been suggested in the literature (e.g. Coe et al. 2006; Ilbert et al. 2006; Coupon et al. 2009) that a re-calibration of the photometric zero-points of the images with the help of spec-z can lead to an enhanced accuracy of photo-z. The procedure involves running the photo-z code on a subsample of objects with reliable spec-z values and just fitting for the template while fixing the redshift to the spectroscopic value. The averaged magnitude differences in a band between the best-fitting templates and the observed photometry can then be applied as corrections to the

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

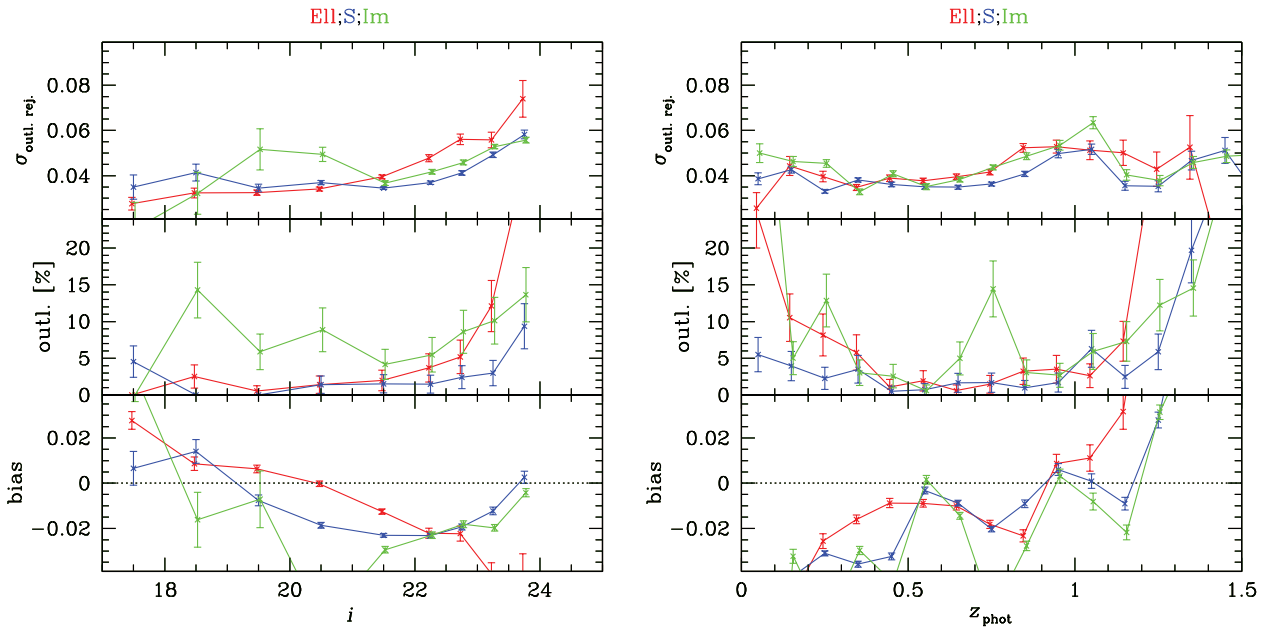


Figure 5. Same as Fig. 4 but showing the accuracy for the different SED templates using the *local* photometry.

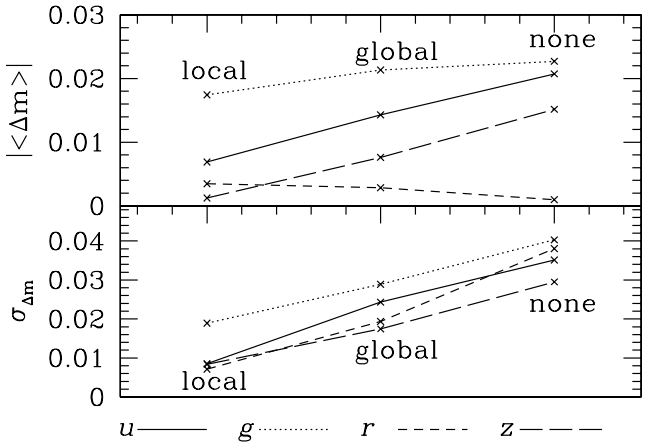


Figure 6. Absolute values of the means (top) and standard deviations (bottom) of the 20 zero-point offsets for the different PSF homogenization methods and different bands.

zero-point in that band. Usually this is done iteratively until convergence is reached.

We performed such a zero-point re-calibration for the *none*, *global* and *local* methods in the 20 fields with VVDS/DEEP2 coverage. This yields 20 zero-point corrections for four of the five bands[9] and for each method. The absolute values of the mean and the standard deviation of the 20 zero-point offsets are shown in Fig. 6.

From the figure it is clear that both the mean and the width of the distributions become smaller going from *none* to *global* to *local*. In particular, the corrections for the *local* method mostly vanish, i.e. most *local* offsets are similar to or smaller than the error of a single correction (≈0.02 mag).

[9] The re-calibration procedure is only sensitive to colours. So we decide to fix the offset in the i-band – the detection band – to $\Delta i = 0$.

The zero-points used for the catalogue extraction were identical in all three methods. Thus, our results for the zero-point re-calibrations strongly suggest that mostly PSF effects are corrected by such a procedure.[10] Furthermore, it means that proper PSF homogenization, in combination with an accurate absolute photometric calibration supersedes zero-point re-calibration procedures like the ones presented in Ilbert et al. (2006) and Coupon et al. (2009).

This is confirmed by looking at the photo-z accuracy before and after re-calibration. In Fig. 7 the photo-z statistics are shown again as a function of magnitude and redshift, for the *none* method and a re-calibrated *none* method. The *local* method is also plotted as a benchmark. The improvement is striking although the accuracy of the 'none-recalib' method does not reach the accuracy of the *local* method at the faintest magnitudes. Interestingly the lines for a re-calibrated *local* method are nearly indistinguishable from the basic *local* method.

We would like to stress that these findings do not tell the whole story and the situation might be even worse for re-calibrated photometry. Re-calibrations of the photometric zero-points might well fool the user into believing that the photo-z accuracy is better than it is in reality. If most of the corrections are due to PSF effects as our results suggest, then these corrections depend on the average angular size of the objects. But the average size of the photometric galaxy sample used for the science projects often differs from the average size of the spectroscopic calibration sample which is not only used for the zero-point re-calibration but also for the following

[10] Certainly, there are other effects that play a role here. For example, absolute photometric calibrations are often done with standard stars that were observed in a slightly different photometric system than the instrumental one. Conversions between the standard system and the instrumental system depend on the colour term of the object. Thus, a correction that is correct on average for stars is not correct on average anymore for galaxies, which have very different SEDs. Often also the filter curves used for template-based photo-z have some uncertainty, especially in the ultraviolet where the differential atmospheric transparency has a large influence on the effective throughput.

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

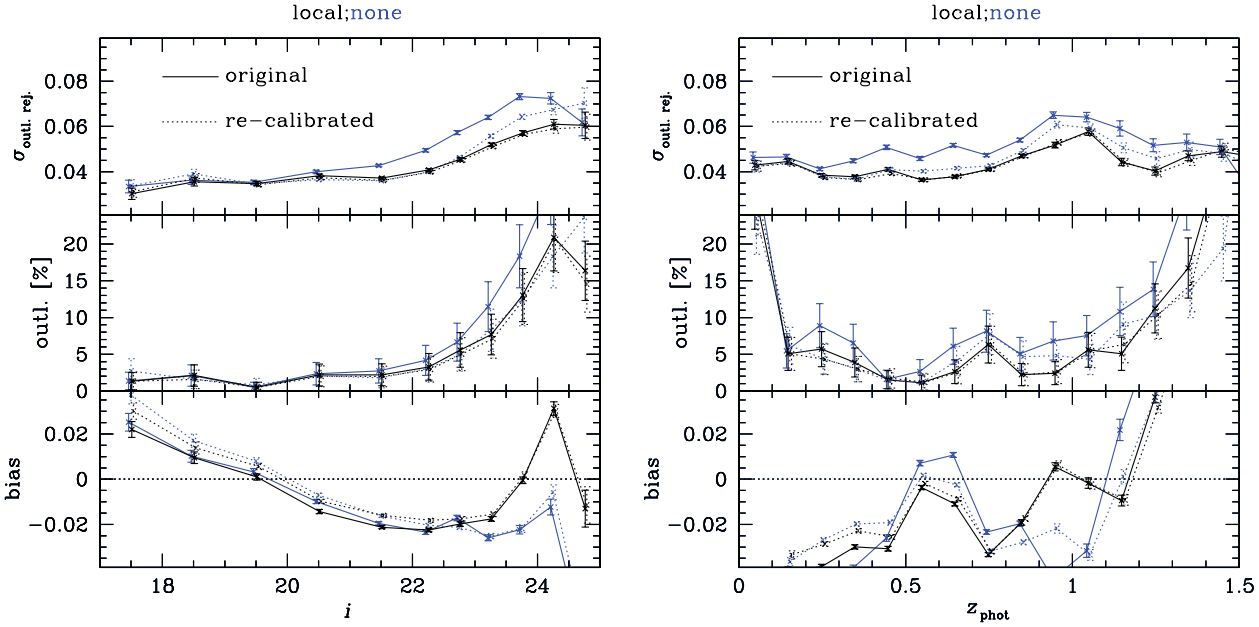


Figure 7. Same as Fig. 4 but showing the effects of zero-point re-calibration on the *none* method (i.e. without PSF homogenization) in comparison to the default *local* method (i.e. local PSF Gaussianization). The solid lines correspond to the original methods whereas the dotted lines correspond to the re-calibrated methods. It is clearly visible that the re-calibrated *none* method performs similar to the original (i.e. not re-calibrated) *local* method, suggesting strongly that zero-point re-calibrations are mostly correcting for PSF effects and not for real biases in the photometric zero-points.

assessment of the photo-z accuracy. This circular use is dangerous if the systematic effects that are corrected for depend on the nature of the objects (e.g. their size) and are not identical for all objects (like e.g. real photometric zero-points). Furthermore, if the zero-point corrections depend sensitively on seeing, it is not advisable to apply the correction found on one particular field to another field.

Based on these findings we decide to use the *local* method for all scientific projects in the CFHTLenS. It offers the best photo-z accuracy combined with the most stable photometry and does not require re-calibration of the zero-points to achieve this. The photo-z values for this method are well understood in the redshift range $0.1 < z_{\rm phot} < 1.3$ for $i < 24.5$ – two magnitudes fainter than the analysis presented in Coupon et al. (2009) – with photo-z scatter values in the range $0.03 < \sigma < 0.06$ and outlier rates smaller than 10 per cent.

5.4 Selection of subsamples with higher photo-z accuracy

The ODDS parameter introduced by Benítez (2000) and described in Section 5.1 can be used to select subsamples of galaxies with a higher photo-z accuracy, with the trade-off of a decreased completeness and an implicit colour selection. In Fig. 8 the photo-z statistics for different cuts on the ODDS parameter are shown. Also the completeness of the sample is reported when such cuts are applied.

The main effect of a cut on ODDS is that the outlier rates at faint magnitudes are reduced. However, by looking at the redshift dependence (right-hand panel of Fig. 8) it becomes clear that these problematic objects are mostly assigned very high redshifts of $z_{\rm phot} > 1.5$ (i.e. they do not appear in the right-hand panel). One exception is a feature at $z_{\rm phot} \sim 0.75$ where the outlier rate can be effectively suppressed by cutting on ODDS. But overall ODDS has a negligible impact in the well-understood redshift range of $0.1 < z_{\rm phot} < 1.3$. For most applications it is probably not meaningful to apply a global cut on ODDS, most importantly because such a selection always entails an implicit colour selection. A redshift-dependent cut on ODDS could make sense for applications that could tolerate such a selection (e.g. selection of background sources for WL).

5.5 Star–galaxy separation

We separate stars from galaxies using a combination of size, magnitude and colour information. Although a pure stellar sample can be obtained by isolating the stellar branch in the size–magnitude plane, the mixing of faint and small galaxies with stars considerably complicates the separation of both classes of objects in this regime. For science cases requiring a pure and complete galaxy sample, having a robust star–galaxy estimator becomes a key issue.

A great advantage of template-fitting methods to estimate photo-z is the ability to use different template sets. Therefore, in addition to galaxy templates, one is able to test for stellar templates as well. We run BPZ again fixing the redshift to $z = 0$ and using the stellar spectral library from Pickles (1998). For each object, BPZ outputs a best-fitting estimate from the stellar library with an associated $\chi^2_{\rm star}$ which can be compared to $\chi^2_{\rm gal}$ associated with the best-fitting galaxy template and redshift. Then ideally a star would satisfy the relation $\chi^2_{\rm star} < \chi^2_{\rm gal}$. However, one should keep in mind that both estimators were computed from independent template libraries and the comparison is therefore not straightforward because of different degrees of freedom. To overcome this difficulty, the estimator can be tested and calibrated on spectroscopic data. The method was applied previously on the CFHTLS-Wide (Coupon et al. 2009) and tested using spectroscopic data from VVDS F02 (Le Fèvre et al. 2005) and VVDS F22 (Garilli et al. 2008) surveys. The following criteria, combining size, magnitude and colour information, have been found to give the best compromise between a pure and a complete galaxy sample when tested on the VVDS:

(i) for $i < 21$: all objects with $r_{\rm h} < r_{\rm h,limit}$ are flagged as stars,
(ii) in the range $21 < i < 23$, objects with $r_{\rm h} < r_{\rm h,limit}$ *and* $\chi^2_{\rm star} < 2 \times \chi^2_{\rm gal}$ are flagged stars and
(iii) for $i > 23$ all objects are flagged as galaxies,

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

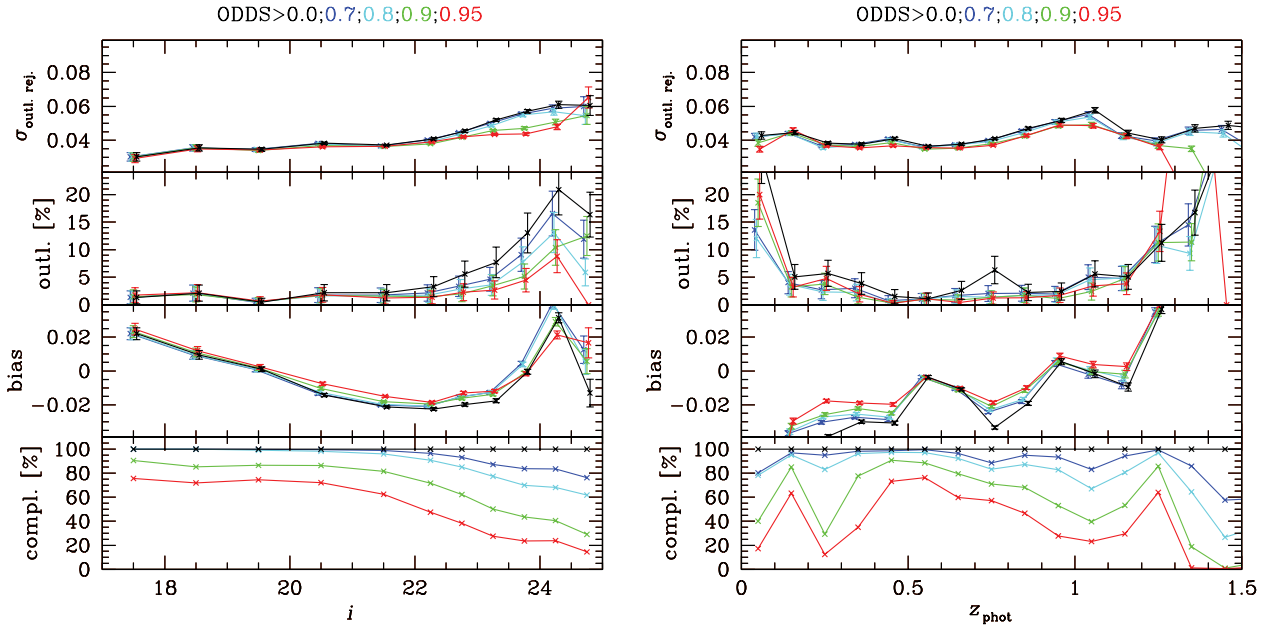


Figure 8. Same as Fig. 4 but showing the effect of cutting on ODDS for the default *local* method (i.e. local PSF Gaussianization; no zero-point re-calibration). The additional bottom panel shows the completeness of the galaxy sample after the ODDS cut is applied.

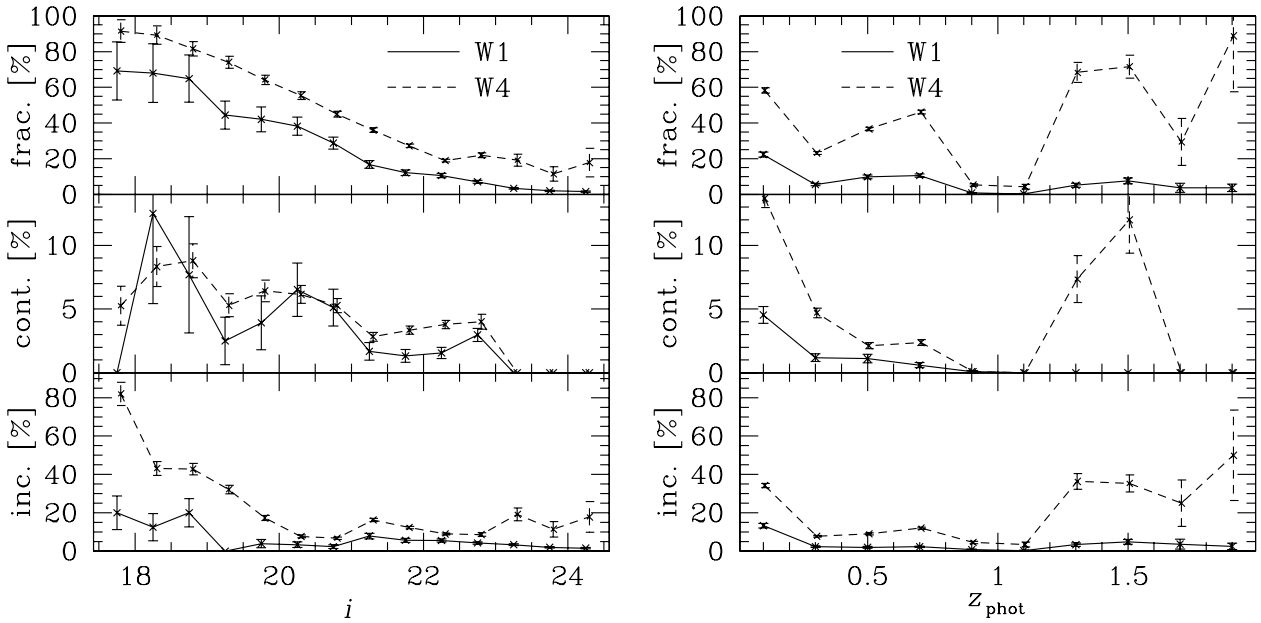


Figure 9. Star–galaxy separation efficiency as a function of the i-band magnitude (left) and photo-z (right). Shown are the fraction of stars (frac.; top), contamination of the galaxy sample (cont.; middle) and the incompleteness of the galaxy sample (inc.; bottom). Our estimator is tested and calibrated using the VVDS spectroscopic samples in W1 (solid line) and W4 (dashed line). Incompleteness represents the percentage of galaxies lost after selection, and contamination the percentage of stars misidentified as galaxies.

where $r_{\rm h}$ is the average half light radius in a field computed in the i-band image and $r_{\rm h,limit}$ the 3σ upper limit of the $r_{\rm h}$ distribution in a single image. The $r_{\rm h,limit}$ values for each field are determined by manually inspecting the size–magnitude diagrams.[11]

We present the method efficiency as a function of magnitude and redshift in Fig 9. True galaxies and true stars are given by spectroscopic information. The incompleteness is defined as the percentage of galaxies lost after selection compared to the total number of true galaxies, and the contamination as the number of true stars misidentified as galaxies compared to the total number of true galaxies. A robust estimator should lead to vanishing numbers for both estimators. In the range $21 < i < 23$, where size, magnitude and colour information are used, the estimator performed the best, keeping the contamination below 5 per cent and the incompleteness below 15 per cent. Extending the estimator to brighter magnitude in W4, where the star concentration is very high, increases the galaxy incompleteness. The star–galaxy separation strongly depends on the

[11] It should be noted that the r_h estimation becomes less and less reliable at faint magnitudes. But for $i < 23$, where we use it, it is still largely unbiased.

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

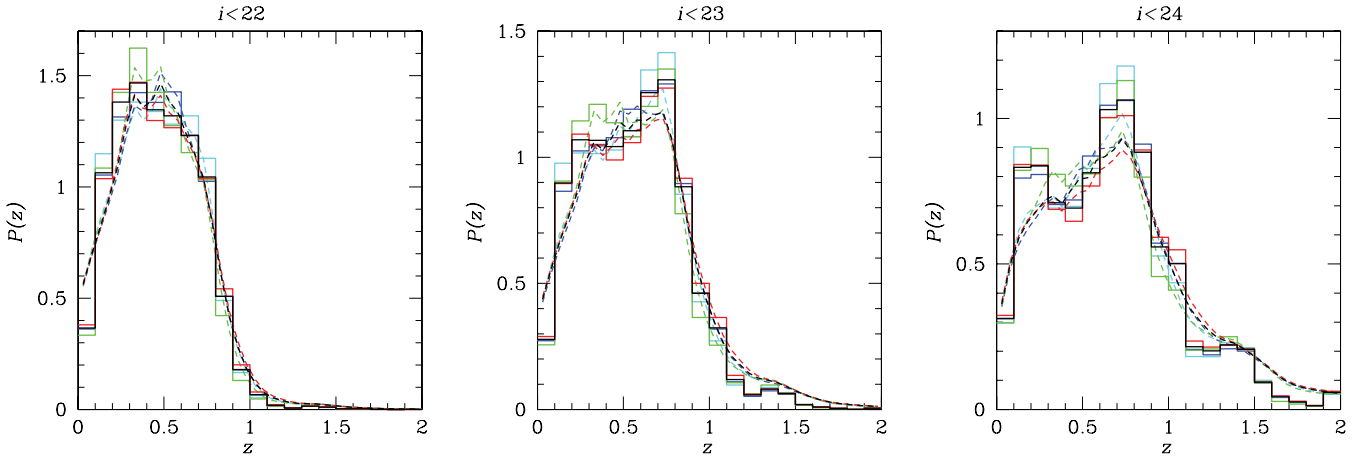


Figure 10. Redshift distributions for $i < 22$ (left), $i < 23$ (middle) and $i < 24$ (right). The dashed lines represent the stacked PDFs and the solid histograms represent the distributions of the most probable photo-z. The whole Wide survey is shown in black, W1 in red, W2 in green, W3 in blue and W4 in cyan.

redshift estimate of the objects. As seen in Section 5.2, the redshift estimator is the most robust in the range $0.1 < z < 1.3$, where both incompleteness and contamination remain below 5 per cent and 10 per cent, respectively. It seems however that a conservative redshift cut $z < 1.1$ should be observed if a very complete galaxy sample is needed in star-crowded fields, or alternatively a hard size cut of $r_{\rm h} > r_{\rm h,limit}$ should be applied so that no star will pass. It should be noted that the high-galactic-latitude fields of the CFHTLS contain only very few stars so that star–galaxy separation is not a major issue for most WL science projects planned by the CFHTLenS team.

5.6 Redshift distributions

The redshift distributions of all objects classified as galaxies are shown in Fig. 10 for three different magnitude limits. We show the distributions of the most probable photo-z values as well as the stacked posterior probabilities output by the photo-z code. While there is good agreement between the two for $i < 23$, the double-peaked histogram for the faintest magnitude cut ($i < 24$) is not reproduced in the stacked posterior probabilities. This redshift-focusing effect occurs when the prior dominates the posterior probability distribution for wide, flat likelihoods (plateaus) in the low S/N case. A large number of objects are then assigned the peak value of the prior which leads to artificial peaks in the redshift histograms. This is the reason why we recommend to use the full probability distributions, after a proper deconvolution taking the photometric uncertainties into account was performed, instead of the most probable redshifts (i.e. the peak of the posterior) in science analyses.[12] Similar double-peaked photo-z distributions for $i < 24$ in the CFHTLS can be seen in Ilbert et al. (2006) and Coupon et al. (2009) so we suspect the filter set plays a crucial role here. But also the prior can lead to multipeaked distributions. A final answer to this question requires a highly complete spec-z catalogue all the way down to $i = 24$.

There are several methods discussed in the literature to correct these redshift probability densities (e.g. Newman 2008; Benjamin et al. 2010) to make them more realistic using angular cross-correlation functions between different photo-z bins or between photo-z and spec-z samples. We defer such an analysis carried out with the CFHTLenS redshifts to a forthcoming paper (Benjamin et al. in preparation).

We intentionally do not present functional fits to these distributions. Contemporary WL surveys (like the CFHTLS) have reached such a precision that redshift distributions taken from external surveys would dominate the total error budget on most cosmological WL measurements (van Waerbeke et al. 2006). It is one great advantage of the CFHTLS over previous surveys to have a photo-z estimate for each galaxy used in the WL analysis.

6 CONCLUSIONS

The CFHTLS represents the largest and therefore most powerful WL survey to date. Here we present new, improved photometric catalogues and photo-z which will be used for the work of the CFHTLenS team. We show that the correction for PSF effects in data taken in different bands at different times is crucially important. Not only does the accuracy of the resulting photo-z increase significantly if PSF effects are thoroughly corrected for, but also the overall photometric homogeneity of a survey can be improved considerably by employing such corrections. We show that this can be done up to the point where re-calibrations of the photometric zero-points with subsamples of galaxies with spec-z become largely unnecessary.

This implies that the re-calibration of photometric zero-points described in the literature are mostly corrections for PSF effects. It is hence dangerous to assume that the photo-z accuracy measured on the same spec-z sample that was also used for the re-calibration will be matched by the photometric sample. On the positive side, these findings suggest that future surveys like KiDS, DES, LSST or Euclid will benefit tremendously from careful PSF homogenization. These projects will not necessarily need a spectroscopic coverage over the whole area to achieve their absolute photometric calibration goals, if PSF effects are corrected for with high precision. For example, Euclid will calibrate its redshift distributions directly from very complete spectroscopic catalogues that will not cover the whole area. Relating the results from these calibration fields to the rest of the survey requires exquisite photometric homogeneity.

The CFHTLenS team will use these photo-z catalogues for a wide variety of WL-related science projects ranging from galaxy–galaxy lensing and cluster lensing to cosmic shear tomography using the shear as well as the magnification effect of WL. Especially the cosmological measurements will benefit from this very homogeneous

[12] It should be noted that the stacked posterior probabilities are certainly affected by the ad hoc modification of the prior described in Section 5.1.

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

photo-z catalogue. Measuring tiny correlations over large angular distances, as it is done in cosmic shear or cosmic magnification studies, requires an exquisite level of control of systematic effects. The removal of PSF effects from the photometry and the resulting accurate photo-z represent an important step to reach this goal.

The shear measurement technique of CFHTLenS will be presented in another technical paper (Miller et al. in preparation) and the resulting shear catalogue will be carefully inspected for systematic effects (Heymans et al. in preparation). Together with the photo-z presented in this technical paper this will set the basis for CFHTLenS science analyses.

ACKNOWLEDGMENTS

We are grateful to the CFHTLS survey team for conducting the observations and the TERAPIX team for developing software used in this study. We acknowledge use of the Canadian Astronomy Data Centre operated by the Dominion Astrophysical Observatory for the National Research Council of Canada's Herzberg Institute of Astrophysics.

We would like to thank Mischa Schirmer for his work on the THELI pipeline as well as Stephane Arnouts, Jean-Charles Cuillandre, Yuliana Goranova, Patrick Hudelot, Olivier Ilbert and Henry J. McCracken for their work on the TERAPIX T0006 catalogues.

The computational infrastructure for this project was supported by an NSERC RTI grant as well as the Marie Curie IOF 252760.

H. Hildebrandt is supported by the Marie Curie IOF 252760 and by a CITA National Fellowship. TE is supported by the BMBF through project 'GAVO III' and by the DFG through project ER 327/3-1 and the TR 33. LVW is supported by NSERC and CIfAR. CH acknowledges support from the European Research Council under the EC FP7 grant number 240185. JC is supported by the Japanese Society for the Promotion of Science. CB is supported by the Spanish Science MinistryAYA2009-13936, Consolider-Ingenio CSD2007-00060, project2009SGR1398 from Generalitat de Catalunya and by the the European Commission's Marie Curie Initial Training Network CosmoComp (PITN-GA-2009-238356). LF is supported by the NSFC grant numbers 11103012 & 10878003, Innovation Programme 12ZZ134 and Chen Guang project 10CG46 of Shanghai Municipal Education Commission and Science and Technology Commission of Shanghai Municipal Grant No. 11290706600. TDK was supported by an RAS 2010 Fellowship. MV acknowledges support from the Netherlands Organization for Scientific Research (NWO) and from the Beecroft Institute for Particle Astrophysics and Cosmology. H. Hoekstra and ES acknowledge support from an NWO Vidi grant and a Marie Curie IRG. BR acknowledges support from the European Research Council in the form of a Starting Grant with number 240672. TS acknowledges support from NSF through grant AST-0444059-001, the Smithsonian Astrophysics Observatory through grant GO0-11147A, and NWO.

AUTHOR CONTRIBUTIONS

H. Hildebrandt led this paper and the photometry working group, created the catalogues, estimated photo-z and conducted the main analysis.

TE reduced the complete data set and developed code for different stages of the photometry and photo-z estimation.

KK developed the shapelet-based code for the Gaussianization of the PSF.

LVW and CH managed and supervised the CFHTLenS collaboration and the data flow. LVW also manually selected stars from size–magnitude diagrams of >6000 chips of the MEGACAM mosaic.

JC led the photometric star–galaxy separation and contributed to the T0006 photo-z.

JB, CB, LF, LM and MV ran different tests for systematic effects in the data (see Section 4.1).

H. Hoekstra, TK, YM, MH, BR, TS and ES contributed to this paper and the photo-z catalogue by organizing the CFHTLS survey operations, using the photo-z, providing valuable feedback, and engaging in numerous discussion on how to improve the catalogue.

NB developed the BPZ photo-z code, which was used extensively in this study.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Adelman-McCarthy J. K. et al., 2007, ApJS, 172, 634
Albrecht A. et al., 2006, preprint (arXiv:astro-ph/0609591)
Arnouts S. et al., 2002, MNRAS, 329, 355
Bacon D. J., Refregier A. R., Ellis R. S., 2000, MNRAS, 318, 625
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Baum W. A., 1962, in McVittie G. C., ed., Proc. IAU Symp. 15, Problems of Extra-Galactic Research. Macmillan Press, New York, p. 390
Benítez N., 2000, ApJ, 536, 571
Benjamin J., van Waerbeke L., Ménard B., Kilbinger M., 2010, MNRAS, 408, 1168
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bolzonella M., Miralles J.-M., Pelló R., 2000, A&A, 363, 476
Brainerd T. G., Blandford R. D., Smail I., 1996, ApJ, 466, 623
Capak P. L., 2004, PhD thesis, Univ. Hawaii
Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, AJ, 132, 926
Collister A. A., Lahav O., 2004, PASP, 116, 345
Connolly A. J., Csabai I., Szalay A. S., Koo D. C., Kron R. G., Munn J. A., 1995, AJ, 110, 2655
Coupon J. et al., 2009, A&A, 500, 981
Coupon J. et al., 2011, preprint (arXiv:1107.0616)
Csabai I. et al., 2003, AJ, 125, 580
Davis M. et al., 2007, ApJ, 660, L1
Erben T. et al., 2005, Astron. Nachrichten, 326, 432
Erben T. et al., 2009, A&A, 493, 1197
Fu L. et al., 2008, A&A, 479, 9
Garilli B. et al., 2008, A&A, 486, 683
Goranova Y. et al., 2009, http://terapix.iap.fr/cplt/T0006-doc.pdf
Hetterscheidt M., Simon P., Schirmer M., Hildebrandt H., Schrabback T., Erben T., Schneider P., 2007, A&A, 468, 859
Heymans C. et al., 2006, MNRAS, 371, L60
Heymans C. et al., 2008, MNRAS, 385, 1431
Hildebrandt H. et al., 2006, A&A, 452, 1121
Hildebrandt H., Pielorz J., Erben T., Schneider P., Eifler T., Simon P., Dietrich J. P., 2007, A&A, 462, 865
Hildebrandt H., Wolf C., Benítez N., 2008, A&A, 480, 703
Hildebrandt H., Pielorz J., Erben T., van Waerbeke L., Simon P., Capak P., 2009, A&A, 498, 725
Hildebrandt H. et al., 2010, A&A, 523, A31
Hildebrandt H. et al., 2011, ApJ, 733, L30
Hoekstra H., 2007, MNRAS, 379, 317
Hoekstra H., Jain B., 2008, Annu. Rev. Nuclear Part. Sci., 58, 99
Hoekstra H. et al., 2001, ApJ, 548, L5
Hoekstra H., Yee H. K. C., Gladders M. D., 2004, ApJ, 606, 67
Hoekstra H. et al., 2006, ApJ, 647, 116
Hoekstra H., Donahue M., Conselice C. J., McNamara B. R., Voit G. M., 2011, ApJ, 726, 48
Hudson M. J., Gwyn S. D. J., Dahle H., Kaiser N., 1998, ApJ, 503, 531
Ilbert O. et al., 2006, A&A, 457, 841
Jee M. J. et al., 2011, ApJ, 737, 59
Johnston D. E. et al., 2007, preprint (arXiv e-prints)

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022

Kaiser N., Wilson G., Luppino G. A., 2000, preprint (arXiv:astro-ph/0003338).
Koo D. C., 1985, AJ, 90, 418
Koo D. C., 1999, in Weymann R., Storrie-Lombardi L., Sawicki M., Brunner R., eds, ASP Conf. Ser. Vol. 191, Photometric Redshifts and the Detection of High Redshift Galaxies. Astron. Soc. Pac., San Francisco, p. 3
Kron R. G., 1980, ApJS, 43, 305
Kuijken K., 2006, A&A, 456, 827
Kuijken K., 2008, A&A, 482, 1053
Le Fèvre O. et al., 2005, A&A, 439, 845
Leauthaud A. et al., 2010, ApJ, 709, 97
Leauthaud A. et al., 2012, ApJ, 744, 159
Loh E. D., Spillar E. J., 1986, ApJ, 303, 154
Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, MNRAS, 368, 715
Massey R. et al., 2007, ApJS, 172, 239
Munshi D., Valageas P., van Waerbeke L., Heavens A., 2008, Phys. Rep., 462, 67
Newman J. A., 2008, ApJ, 684, 88
Okabe N., Takada M., Umetsu K., Futamase T., Smith G. P., 2010, PASJ, 62, 811
Parker L. C., Hudson M. J., Carlberg R. G., Hoekstra H., 2005, ApJ, 634, 806
Parker L. C., Hoekstra H., Hudson M. J., van Waerbeke L., Mellier Y., 2007, ApJ, 669, 21
Peacock J. A., Schneider P., Efstathiou G., Ellis J. R., Leibundgut B., Lilly S. J., Mellier Y., 2006, Technical report, ESA–ESO Working Group on 'Fundamental Cosmology'
Pickles A. J., 1998, PASP, 110, 863
Puschell J. J., Owen F. N., Laing R. A., 1982, ApJ, 257, L57
Refregier A., 2003, MNRAS, 338, 35
Refregier A., Bacon D., 2003, MNRAS, 338, 48
Schirmer M., Erben T., Schneider P., Pietrzynski G., Gieren W., Carpano S., Micol A., Pierfederici F., 2003, A&A, 407, 869
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Schrabback T. et al., 2007, A&A, 468, 823
Schrabback T. et al., 2010, A&A, 516, A63
Semboloni E. et al., 2006, A&A, 452, 51
Tyson J. A., Wenk R. A., Valdes F., 1990, ApJ, 349, L1
van Uitert E., Hoekstra H., Velander M., Gilbank D. G., Gladders M. D., Yee H. K. C., 2011, A&A, 534, A14
van Waerbeke L. et al., 2000, A&A, 358, 30
van Waerbeke L., White M., Hoekstra H., Heymans C., 2006, Astropart. Phys., 26, 91
Wittman D. M., Tyson J. A., Kirkman D., Dell'Antonio I., Bernstein G., 2000, Nat, 405, 143
Wolf C., Meisenheimer K., Röser H.-J., 2001, A&A, 365, 660

APPENDIX A: DETAILS OF THE PSF HOMOGENIZATION

A key part of our analysis is the homogenization of the PSF between images taken through different filters, so that proper colours, representing the same part of each source, can be measured. Our approach is to convolve the images with a kernel that renders the PSF close to Gaussian, with a width that is set by the worst-seeing image.

The PSF Gaussianization is performed by first constructing a suitable, spatially variable, kernel and then convolving the images with it. Both steps take advantages of some of the mathematical properties of the shapelet formalism (Refregier 2003).

Shapelets are two-dimensional Gauss–Hermite functions, of the form

$$S_{ab}(x, y) = N_{ab} H_a(x/\beta) H_b(y/\beta) \mathrm{e}^{-r^2/2\beta^2}. \quad \text{(A1)}$$

N_{ab} is a normalization constant and β the scale radius. x and y are Cartesian sky coordinates with respect to a suitably chosen centre; r is the corresponding polar coordinate. H_n is a Hermite polynomial.

Any source with intensity $I(x, y)$ can be written as a superposition of shapelets:

$$I(x, y) = \sum_{ab} s_{ab} S_{ab}(x, y), \quad \text{(A2)}$$

where the s_{ab} are the amplitudes of the different shapelets.

Shapelets have the useful property that a convolution of any two of them can be written as a new (generally infinite) shapelet series:

$$S_{ab} \otimes S_{cd} = \sum_{ef} C_{ace} C_{bdf} S_{ef}, \quad \text{(A3)}$$

and expressions for the matrix elements C_{lmn} are given in Refregier & Bacon (2003). (Note that there is no requirement for the scale radii of the shapelets to be the same.) If we express the PSF and the kernel as shapelet series, with coefficients p_{ab} and k_{ab}, respectively, then this allows us to write the result of the convolution $P \otimes K$ as a new shapelet series with coefficients

$$t_{ef} = \sum_{ab} \sum_{cd} C_{ace} C_{bdf} p_{ab} k_{cd} \equiv \sum_{cd} M_{cd,ef} k_{cd}, \quad \text{(A4)}$$

with

$$M_{cd,ef} = \sum_{ab} C_{ace} C_{bdf} p_{ab} \quad \text{(A5)}$$

encoding the effect of PSF convolution in shapelet space. If the PSF is known, M can be computed. Our Gaussianization technique involves constructing a shapelet kernel by inverting equation (A4). As the target PSF on the left-hand side we stipulate a Gaussian; all its elements are zero except the t_{00} component.

Before calculating the kernel, we model the PSF variation across the full CFHTLS image. This is done by identifying all stars above a certain S/N, choosing a suitable scale radius for the PSF, making a flux-normalized shapelet expansion for each star and then separately fitting (with a two-dimensional polynomial) the variation of each coefficient s_{ab} across the image, with outlier rejection.

This PSF map is then sampled on a regular grid across the image, the convolution kernel is calculated at each of those points and a polynomial model fitted to the variation of its coefficients across the image. This kernel is then convolved with the original CFHT image. Again here we can use a nice mathematical property of shapelets: each is its own Fourier transform, making the convolution efficient.

In practice, we truncate the shapelet series: we have only pixellated information on the PSF and so cannot sample it arbitrarily finely. For robustness we truncate the kernel at a lower order than the PSF, to prevent overfitting. The inversion of equation (A4) is therefore recast as a least-squares problem, in which we determine the k_{cd} that best approximate the target Gaussian PSF t_{ab}.

For the CFHT data we find that the following parameters work well: shapelet order (maximum $a + b$ in the expansion equation A1) equal to 10 for the PSF, 8 for the kernel; polynomial order for fitting spatial variation across the image 5; shapelet scale radius of the PSF map 1.3 times the median Gaussian radius fitted to the stars and target Gaussian radius 0.8 times the largest scale radius of the images to be compared.

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

Downloaded from https://academic.oup.com/mnras/article/421/3/2355/1078193 by guest on 05 June 2022