Evil preparation

XXXI. The effect of the variations in photometric passbands on photometric-redshift accuracy


(Affiliations can be found after the references)

Received 24 May 2023 / Accepted 26 September 2023

ABSTRACT

The technique of photometric redshifts has become essential for the exploitation of multi-band extragalactic surveys. While the requirements on photometric redshifts for the study of galaxy evolution mostly pertain to the precision and to the fraction of outliers, the most stringent requirement in their use in cosmology is on the accuracy, with a level of bias at the sub-percent level for the **Evil** cosmology mission. A separate, and challenging, calibration process is needed to control the bias at this level of accuracy. The bias in photometric redshifts has several distinct origins that may not always be easily overcome. We identify here one source of bias linked to the spatial or time variability of the passbands used to determine the photometric colours of galaxies. We first quantified the effect as observed on several well-known photometric cameras, and found in particular that, due to the properties of optical filters, the redshifts of off-axis sources are usually overestimated. We show using simple simulations that the detailed and complex changes in the shape can be mostly ignored and that it is sufficient to know the mean wavelength of the passbands of each photometric observation to correct almost exactly for this bias; the key point is that this mean wavelength is independent of the spectral energy distribution of the source. We use this property to propose a correction that can be computationally efficiently implemented in some photometric-redshift algorithms, in particular template-fitting. We verified that our algorithm, implemented in the new photometric-redshift code Phosphoros, can effectively reduce the bias in photometric redshifts using the CFHTLS T007 survey, with an average measured bias Δz over the redshift range 0 ≤ z ≤ 0.7 decreasing by about 0.02, specifically from Δz ≃ 0.04 to Δz ≃ 0.02 around z = 0.5. Our algorithm is also able to produce corrected photometry for other applications.

**Key words.** galaxies: distances and redshifts — cosmology: observations — surveys — techniques: photometric — techniques: miscellaneous

1. Introduction

Multi-megapixel cameras with large fields of view have revolutionised extragalactic astrophysics and observational cosmology by enabling photometric surveys of large sky areas in several optical and near-infrared bands. The Dark Energy Survey (DES; The Dark Energy Survey Collaboration 2005), the Kilo-Degree Survey (KIDS; de Jong et al. 2013), and the Hyper Suprime-Cam Strategic Survey Program (HSC-SSP; Aihara et al. 2018) are recent examples of photometric surveys with areas exceeding 1000 deg². These surveys enable the measurement of the cosmic shear, which is the distortion of the images of distant galaxies.
objects caused by the propagation of light rays through inhomogeneous matter (Blandford et al. 1991). The cosmic shear allows the reconstruction of dark-matter maps at different redshifts, from which the distribution of matter and its evolution can be inferred. Modern cosmological surveys have established cosmic shear as one of the main modern cosmological probes and have already provided important cosmological constraints (e.g. Abbott et al. 2018 for DES; Asgari et al. 2021 for KiDS; and Hamana et al. 2020 for HSC-SSP).

Euclid (Laureijs et al. 2011) is a mission of the European Space Agency that will perform a survey over 15,000 deg² of extragalactic sky (Euclid Collaboration 2022a) with optical and near-infrared imaging, as well as with slitless multi-object spectroscopy in the near-infrared. The main scientific probes of Euclid are the cosmic shear and galaxy clustering. They are supported by two instruments: The VIS optical camera (Cropper et al. 2014) will provide us with high-resolution images of galaxies for the determination of the cosmic shear; and the Near Infrared Spectrometer and Photometer near-infrared instrument (NISP; Maciaźek et al. 2016) will perform near-infrared photometry in three bands to support cosmic shear determination, as well as near-infrared spectroscopy for the study of three-dimensional galaxy clustering. Compared to current ground-based surveys, the determination of the cosmic shear with Euclid will greatly benefit from high-resolution imaging and near-infrared photometry from space, but also from the significantly larger sky area (three times the area of DES) and depth.

The measurement of redshifts for a large fraction of the surveyed galaxies is an indispensable step in the cosmological study of the cosmic shear. While redshift measurements can be performed through spectroscopy, spectroscopy is more challenging than photometry at very faint fluxes, which limits the number of measurable redshifts, even when efficient multi-object spectrographs are available. In the case of Euclid, spectroscopic redshifts will be determined for some 30 million galaxies, while the total number of galaxies for which sufficiently precise shapes can be measured will exceed one billion (Laureijs et al. 2011). The technique of photometric redshifts, that is, using photometric observations only, is currently the only way to determine redshifts on such a huge scale. The limited precision of photometric redshifts is not an issue for cosmic shear because the efficiency of lensing is a slowly varying function of the redshift. Photometric redshifts have thus become an essential tool of modern observational cosmology.

Photometric-redshift determination requires photometric observations of galaxies in several wavelength bands, which defines a multi-dimensional flux or colour space. The redshift is then obtained from the construction of a mapping between the position of an object in this colour space and the redshift. We can in principle determine this mapping through different approaches, either based on real objects with known redshifts or on simulated objects. The determination of redshifts of galaxies using a small number of broad-band photometric measurements was pioneered six decades ago by Baum (1962). The template-fitting (TF) technique, which involves the comparison of the magnitudes or fluxes of a galaxy with those of simulated objects, was first developed by Loh & Spiller (1986) and subsequently exploited in several codes that are still in use today (e.g. Arnouts et al. 1999; Bolzonella et al. 2000). A more recent approach using machine-learning (ML), whose first applications to photometric-redshift calculations were made by Firth et al. (2003) and Tagliaferri et al. (2003) using artificial neural networks, is now becoming more and more popular, and virtually every ML approach has been investigated, for instance support vector machine (Wadadekar 2005), decision trees (Carrasco Kind & Brunner 2013), and Gaussian processes (Almosallam et al. 2016). A review of the challenges of photometric-redshift determination in the context of cosmological surveys has been presented in Newman & Gruen (2022).

The usefulness of any quantity for a scientific application is bound to meet the requirements on the quality of its determination. In the case of photometric redshifts, this quality is usually expressed with three parameters. When a photometric-redshift determination is successful, the measurement is distributed around the true value; the dispersion $\sigma_z$ of this distribution is a measure of the precision of the determination. However, it sometimes happens that the photometric-redshift determination fails completely and the predicted redshift is found to lie very far from the true value; the probability of such a failure is called the “outlier fraction”. Finally, the bias $\Delta z$ is the peak of the distribution of the differences between predictions and true redshift values and determines the accuracy of the predictions.

The Euclid requirements on the precision of photometric redshifts (Amara & Réfrégier 2007) are not extremely demanding [$\sigma_z = 0.05 (1+z)$ and an outlier fraction of 10% at a magnitude 24.5] compared with what can be achieved on small fields; for instance, in the COSMOS field, Weaver et al. (2022) obtained, at similar depths, $\sigma_z$ better than 0.02 $(1+z)$ and an outlier fraction better than 5%. Meeting the Euclid requirements nevertheless remains challenging over the full Euclid wide survey because of the difficulty to obtain deep-enough photometry over a field almost 10,000 times larger (Euclid Collaboration 2020a). However, the biggest difficulty in using photometric redshifts for cosmology is the accuracy in the photometric-redshift determination that is necessary for the cosmic-shear probe; in Euclid, the requirement is expressed as the bias on the mean redshift in each of the approximately ten tomographic redshift bins at redshift $z$, which needs to be less than $\Delta z / (1+z) = 0.002$. Such a stringent requirement demands that some bias correction takes place after the photometric-redshift determination, for instance by calibrating the bias directly in the colour space occupied by galaxies (Masters et al. 2015). This study was the main motivation behind the C3R2 project to gather a gold-standard set of spectroscopic redshifts over the full colour space of galaxies (Masters et al. 2017, 2019; Euclid Collaboration 2020b, 2022b; Stanford et al. 2021). It is however essential to remove, as much as possible, any bias before the calibration step if we want to meet this requirement. We focus here on one specific source of bias that is inherent to the photometric measurement and that is linked to the time and spatial variations of the photometric passbands. Spatial variations are dependent on the positions of the sources in the field of view and thus induce spatial variations unrelated to the source in the photometric redshifts that, in turn, introduce a spurious signal into the correlation function of the cosmic shear. Such variations may have an effect on the dispersion, outlier fraction, and bias; we focus here on the bias since this is the most stringent requirement on photometric redshifts for Euclid.

In this paper, we first review the current knowledge about passband variations for surveys relevant to Euclid. We then discuss the problems caused by these variations and demonstrate the effect quantitatively using idealised simulations. We finally propose an efficient implementation that can remove most of the
bias resulting from this issue, and validate it using a real photometric survey. While we focus on the spatial dependence in this paper, the correction also applies to time variations of the passbands.

2. Photometry

Photometric measurements are performed over ranges of wavelengths called "passbands" that are defined by a transmission curve. This curve, denoted $T(\lambda)$, is a function of the wavelength $\lambda$ that describes the fraction of photons (or of the energy) entering the telescope that is ultimately recorded on the detector. The main component affecting the transmission curve is an optical filter that restricts the range of accessible wavelengths; other components include the detector quantum efficiency and the transmission through the other optical elements. Usually $T(\lambda)$ transistions are designed to approximate more or less a top-hat filter. One defines the AB flux $F_T$ of a source with rest-frame spectral energy distribution (SED) $L(\lambda_0)$, with $\lambda_0 = \lambda/(1+z)$ where $z$ is the redshift, through the photometric transmission curve $T(\lambda)$ as (Oke & Gunn 1983)

$$F_T = \frac{\int_0^\infty s(\lambda) T(\lambda) \lambda \, d\lambda}{c \int_0^\infty T(\lambda) \frac{d\lambda}{\lambda}},$$

(1)

where $c$ is the speed of light and

$$s(\lambda) = \frac{L(\lambda_0) 10^{-0.4 E_{B-V}} k(\lambda)}{4\pi (1+z) D_L^2}$$

(2)

is the observed flux as a function of wavelength, with $D_L$ the luminosity distance, $E_{B-V}$ the Galactic reddening and $k(\lambda)$ the extinction law. The above expression assumes that the detector is counting photons, which is the case for the vast majority of modern detectors in the optical and near-infrared range. The AB convention provides the flux of a source with flat SED (when expressed per unit of frequency) that would leave the same number of counts on the detector as the source under study when passing through the transmission curve $T(\lambda)$. We point out that, with the definition of Eq. (1), the normalisation of $T(\lambda)$ is arbitrary. We assume here that $T(\lambda)$ is normalised so that $\int_0^\infty T(\lambda) \, d\lambda = 1$.

Photometric extragalactic surveys use a (small) set of $N$ passbands with transmission curves $T_i(\lambda)$, $i = 1, \ldots, N$, covering distinct, and usually very slightly overlapping, wavelength ranges. The $ugriz$ set of passbands of the Sloan Digital Sky Survey Photometric Camera (Gunn et al. 1998) is now the most common system used for extragalactic surveys, with slight differences in the filters depending on the cameras and telescopes. The Euclid photometric survey consists of the very broad $I_E$ passband of the VIS optical instrument, which covers roughly the $riz$ bands and which we will not consider further here, and the three passbands in the near-infrared of the NISP instrument, the $Y_E$, $J_E$, and $H_E$ passbands (see Sect. 3.4). The full set of passbands used in this paper, which we denote $T$, are shown in Fig. 1.

3. Spatial variation of photometric passbands

A consistent photometric system implicitly assumes that the transmissions $T_i(\lambda)$, $i = 1, \ldots, N$ are the same for all objects. However, small variations from object to object are possible. In the presence of passband variations, the colours of the different sources occupy different colour spaces, and cannot be compared any more. Passband variations can occur due to several effects as a function of time or of position on the detector. Time dependence can be introduced by atmospheric effects and by the evolution or degradation of the properties of the optical elements, filters, or detectors in a way that depends on the wavelength. Spatial dependence of the passbands is another issue that affects differently, and systematically, the sources in the field of view. Such effects can be introduced by non-uniformities of the filters or detectors, or by the fact that the optical beams of off-axis sources hit the filter at angles that are different from those in the case of on-axis sources. The latter effect can in principle be predicted. In order to obtain photometric measurements with the highest possible accuracy, several teams have measured passband variations across the field of view of their camera. We briefly review studies of the spatial variations of the passbands below, in order to demonstrate that the issue is general, and not limited to a particular survey.

Passband variations have been measured for several photometric systems. Denoting $E(x)$ the average of any function $x(\lambda)$

Fig. 1. Set of transmission curves $T = \{ugrizY_EJ_EH_E\}$ used for the Euclid mission (from left to right). The $ugriz$ passbands are only fiducial, since different sets will be used by Euclid; those represented here are from SDSS. The $Y_EJ_EQ_E$ passbands are from NISP on board Euclid. Only the filter transmissions are shown, without atmospheric, telescope, and detector quantum efficiency effects.
over the transmission $T(\lambda)$, that is,
\[
E(x) = \frac{\int_0^\infty x(\lambda) T(\lambda) \, d\lambda}{\int_0^\infty T(\lambda) \, d\lambda},
\]
we characterise the measured variations using the changes in the first four moments of the normalised transmission curve $T(\lambda)$ in different locations in the field of view: the mean $\mu = E(\lambda)$; the variance $\sigma^2 = E((\lambda - \mu)^2)$; the skewness $\chi = E((\lambda - \mu)^3)$; and the kurtosis $K = E((\lambda - \mu)^4)$. We point out that the skewness and the kurtosis are the standardised third and fourth moments, respectively. We will also use the dispersion $\sigma$, instead of the second moment $\sigma^2$, as it is more intuitively understandable. We stress that these moments do not depend on the SED of the source.

3.1. Sloan Digital Sky Survey Photometric Camera

The response of the Sloan Digital Sky Survey Photometric Camera (Gunn et al. 1998), which is mounted on the Sloan Foundation Telescope at Apache Point Observatory, has been characterised in extensive detail in Doi et al. (2010). The camera is composed of 30 CCDs in a matrix of 5 rows and 6 columns. The passband variations are determined empirically using a series of monochromatic 1 nm-wide dome flats (Rheault et al. 2012). They measured in particular the response for the six columns of the imager for the five ugriz bands, in order to estimate the column-to-column passband variation. Figure 2 (top) shows the six r passbands obtained in one campaign. Stronger variations due to temperature and ageing are however reported by Doi et al. (2010); they quote variations in the passband effective wavelengths from about 3 nm in the best case ($g$ and $i$ passbands) to 12 nm in the worst case ($z$ passband, which is strongly affected by the CCD quantum efficiency), and 3.5 nm in the $r$ passband. Unfortunately, the measurements obtained at other periods do not seem to be available.

Figure 2 (bottom) shows the first four moments of the $r$ passbands. One detector column (3) shows particularly large variations of mean wavelength and dispersion. The overall shape remains however very similar, with very minor change in skewness (<0.01) and kurtosis (0.007).

3.2. Dark Energy Camera

The Dark Energy Survey (The Dark Energy Survey Collaboration 2005) is a griz photometric survey of about 5000 deg$^2$ with the 2.2 deg-diameter Dark Energy Camera (Flaugher et al. 2015) located on the Blanco telescope at Cerro Tololo Observatory. Passband variations have been studied in detail by Li et al. (2016). They measured the change of the passband as a function of the distance to the centre of the detector up to the edge of the field-of-view around 1.1 deg. The $i$ filter was found to show the largest variation, with a wavelength shift of the cut-on wavelength of about 6 nm. Figure 3 (top) shows the variation of the $r$ passband. As in the case of SDSS, there is little variation in the shape of the passband.

Figure 3 (bottom) shows the first four moments of the $r$ passbands; larger position indices indicate larger off-axis distances. The mean wavelength is very stable, with an amplitude of variation of about 1 nm compared to the passband in the centre of the field of view, which we refer to in the following as the “central passband”, without clear dependence on the off-axis angle; the dispersion variation is very comparable to that of SDSS. Again very small changes in skewness and kurtosis are observed.

3.3. MegaCam

The MegaCam instrument (Boulade et al. 2003) is a 1-deg$^2$ imaging camera located on the prime focus of the Canada-France-Hawaii Telescope in Hawaii. The detailed calibration of the camera has been performed in the framework of the Supernova Legacy Survey project (Guy et al. 2010) by Betoule et al. (2013). They found the location of the source as a function of the distance to the centre of the detector plane to be a major driver of passband variations, with transmissions moving towards shorter wavelengths. This is expected for optical interference filters (see discussion in, e.g., Sect. 3.2 of Euclid Collaboration 2022c, and references therein). Figure 4 (top) shows the variation of the $r$ passband, where the position index is correlated with the off-axis distance, with every shift in the position index corresponding to about an additional 5 arcmin distance from the centre of the field (see also Sect. 6).

Figure 4 (bottom) shows the first four moments of the $r$ passbands. The changes in the mean wavelength are much more significant, with a peak-to-peak amplitude of about 8 nm, although the...
the change in the dispersion is not significantly larger than in the case of SDSS or DES. Similar changes in mean wavelengths do occur with other SDSS or DES passbands. The amplitude of skewness and (especially) kurtosis variations are several times larger than for SDSS, indicating the presence of more significant changes in the shape of the passband.

3.4. Euclid NISP

Euclid NISP (Maciaszek et al. 2016) is equipped with 3 photometric filters covering the Euclid $Y_E$, $I_E$, $H_E$ bands, shown in Fig. 1. The NISP photometric system is described in detail in Euclid Collaboration (2022c). The passbands were computed as a function of position in the NISP field of view, based on local filter passband measurements and full ray tracing of the Euclid NISP optical system to account for angle-of-incidence variations on the filter surface, reaching an accuracy of 0.8 nm. The study focuses on the determination of cut-on and cut-off wavelengths, but does not addresses other variations of the transmission. Polynomial expressions have been provided to compute the cut-on and cut-off wavelengths at any position (see Fig. 8 in Euclid Collaboration 2022c). These measurements showed the existence of a blue shift that depends on the off-axis distance, which can be explained by the different incident angles of the incoming light. Blue shifts between 2.5 nm and 6.1 nm have been observed for the cut-on or cut-off wavelengths of the three filters, with the $Y_E$ and $H_E$ filters being the least and the most affected, respectively (see Fig. 9 in Euclid Collaboration 2022c).

4. Consequences of passband variations

4.1. Effect on the photometry

Photometric observations require a calibration in order to derive, for each observation frame, the so-called zero-point, i.e. the relation between the physical flux (or magnitude) and the count rate on the detector. This is generally achieved by using reference stars located in the field of view. The AB definition from Eq. (1) provides an exact relation only for sources that have the same SED as the calibration stars. In order to cope with galaxies with different SEDs, the calibration is sometimes refined with the addition of a colour term (e.g., Padmanabhan et al. 2008; de Jong et al. 2015), which is a correction based on the ratio of the counts in adjacent photometric bands, providing a coarse approximation for the true shape of $\sigma(\lambda)$. This method calibrates...
all bands simultaneously, which can be done only for the bands that are observed with the same telescope. It is therefore not well adapted to the *Euclid* survey, which will use a combination of telescopes, and thus colour-term correction is not used.

With regard to passband variations, the fact that a unique zero-point correction is computed for each frame implies that any spatial variations of the passband across the field-of-view are ignored. The calibration is therefore formally valid only for some average passband. In the case of *Euclid*, a second step is performed which corrects for any spatially dependent systematic deviations of the reconstructed fluxes. This step makes the fluxes independent of possible changes in the normalisation of the passband, provided they only depend on the location in the field of view. In the case of passband variations, this correction would remove any effect linked to a change in the effective area of the transmission. Atmospheric absorption also results in passband variations; however, they can be considered as spatially uniform over a single observation. Hence the calibration process will absorb this effect in the zero point, although in reality the correction depends on the colour of the object. For extended objects, the passband could in principle change across the object, which would affect photometric extraction in a complicated way; however, in cosmological applications useful galaxies are very small compared to the scale on which passband variations are measured, so we can safely ignore this effect.

As a result of the *Euclid* photometric calibration, variations that are not limited to a change in normalisation do impact the photometric measurements. In their very detailed analysis of the photometric stability of the SDSS camera, Doi et al. (2010) found that column-to-column variations of the passbands induce errors on the $g$, $r$, $i$, and $z$ fluxes of up to $1\%$ (0.01 mag). Based on the “Scientific Challenge 8” simulations of the *Euclid* performance, *Euclid Collaboration (2022c)* found that the effect is of the order of a few millimags in the *Euclid* near-infrared bands. Such bias would impact the performance of photometric-redshift determination. In order to remove this bias, it would be necessary to include the correct transmission curve $T(\lambda)$ in Eq. (1); however, the bias would also depend on the a priori unknown SED of the object, so that the correction cannot be performed on isolated frames.

### 4.2. Effects on photometric-redshift determination

#### 4.2.1. First-order effect

We build here a toy model to allow us to estimate the amplitude of the effect of passband variations on the photometric redshift using a simplistic SED consisting of a step function (when expressed per wavelength) at the wavelength of the Balmer break, which we set here to be exactly 400 nm. We consider a system of three top-hat transmission curves $UGR$, with $U(\lambda) = 1$ if $300 \text{ nm} < \lambda < 400 \text{ nm}$, $G(\lambda) = 1$ if $400 \text{ nm} < \lambda < 500 \text{ nm}$, and $R(\lambda) = 1$ if $500 \text{ nm} < \lambda < 600 \text{ nm}$, all transmission curves being 0 outside of these ranges. From Eq. (1), we find that the fluxes $f_U$ and $f_R$ are constant if we consider only redshifts $z < 0.25$, since the Balmer break remains within the $G$ band. As the redshifted Balmer breaks moves across the $G$ passband, $f_G$ changes as a function of $z$ according to Eq. (1) (see Fig. 5), so that there is a direct relationship between $f_G$ and the redshift. Because of the $\lambda$ term in the numerator of Eq. (1), the relationship is not exactly linear.

We consider a varying $G$ passband where the only possible variation is a shift in the mean wavelength of the passband by an amount of $\delta$, i.e. $G(\lambda) = 1$ if $400 \text{ nm} < \lambda - \delta < 500 \text{ nm}$, and $G(\lambda) = 0$ otherwise. Obviously, the mean $\mu$ of $G$ is shifted by the same amount $\Delta \mu = \delta$. As discussed in Sect. 3, typical values of $\delta$ can be of the order of a few nanometres. Figure 6 shows the bias that results from a shift of $\Delta \mu$ of the passband for a true redshift $z = 0.15$.

From Fig. 6 we found that the bias $\Delta z$ is almost a linear function of $\mu$, with $\Delta z = -0.0033 \Delta \mu / 1 \text{ nm}$. This is close, but not identical, to the expected bias if one is able to locate the Balmer break with an error of $\Delta \mu$. In such a case, we would have: $\Delta z = \Delta \mu / 400 \text{ nm} = -0.0025 \Delta \mu / 1 \text{ nm}$. The larger slope is due to the stronger weight of long wavelengths in Eq. (1). This very simplistic analysis shows nevertheless that a shift of 1 nm leads to a bias that is of similar amplitude to *Euclid*’s requirement on the bias, $\Delta z / (1+z) = 0.002$, which shows that changes in passband variations need to be taken into account in the computation of photometric redshifts. In Fig. 6, we also showed the range of $\Delta \mu$ observed in Sect. 3. We found that, because of the tendency of off-axis transmissions to move towards the blue, photometric redshifts are expected to be biased positively if the central passbands are used.
4.2.2. Consequences for photometric-redshift algorithms

Passband variations add a major complexity in the process, because potentially each source is observed with a (slightly) different set of passbands, which means that each source lives in a different colour space (which we assume is known through the measurements of its actual passband). Therefore each source requires a different mapping from colour space to redshift. The two main approaches to photometric-redshift determination face significant, but distinct, issues when dealing with multiple colour spaces.

Template-fitting. The TF approach to determine photometric redshifts (e.g., Arnouts et al. 1999; Bolzonella et al. 2000; Paltani et al., in prep.) involves the knowledge of galaxy SEDs, the so-called templates, which are assumed to be known at all relevant rest-frame wavelengths. The source fluxes are compared with reference fluxes that are computed by integrating the templates through the passbands. Hence, it is straightforward to compute the reference fluxes in the colour space of the source of interest. However, the calculation of the reference fluxes implies integrations through passbands and becomes computationally expensive for any reasonably large survey, and impossible for catalogues of billions of sources such as the future Euclid Wide survey. In the case where all sources occupy the same colour space, this is solved by first calculating a single grid of fluxes for all models at all redshifts. If each passband has n variations, the model grid becomes n times larger, which could become quite cumbersome.

Machine-learning approaches. ML is a vast class of algorithms that share the basic principle to infer the desired relation (in this case between the source’s fluxes and the redshift) in a purely data-driven manner. ML approaches almost always involve a training phase, where the algorithms build up internally (learn) the colour-redshift relation. After the training phase, photometric redshifts can be very efficiently computed through this newly determined relation. However, the training phase of ML algorithms can be a quite computationally intensive process, depending on the algorithm. This is not an issue in the case where all the sources and reference objects occupy the same colour space, since it needs to be performed only once. However, in the presence of passband variations, all sources can in principle occupy distinct colour spaces, so that the training phase would need to be performed many times, which may become computationally difficult. Another difficulty lies in finding enough training objects in each colour space, so that the colour-redshift relation can be accurately learned. As a matter of fact, finding a reference sample that covers entirely a single colour space of galaxies is already extremely difficult (Masters et al. 2015).

5. Implementation in template-fitting algorithms

The TF algorithm is based on the assumption that the SEDs of the real objects are drawn from a known set of SEDs, so that we are able to compute the predicted model colours in any colour space. However, associating to every source the full passband information, including the variations specific to this source, is quite demanding in terms of data management. Furthermore, this results in a much larger model grid size. We thus propose here a simplified correction that only takes into account the shifts in the mean wavelengths of the passbands, and we then validate our approach with simulations.

5.1. Template-fitting likelihood

TF algorithms compare the source fluxes with those obtained from simulated objects with known parameters α. These parameters are typically the set of reference SEDs used to match the observed SED of the source, the redshift of the source, the internal reddening law (e.g., Prevot et al. 1984; Calzetti et al. 2000, etc.), and the value of internal reddening E(B-V). The likelihood of the match as a function of α is given by exp(−χ^2_α/2), with

\[ \chi^2_α = \sum_{T ∈ ℋ} \frac{(f_T^α - α f_T^0)^2}{σ_T^α}, \]

where the sum runs over all passbands T ∈ ℋ, f_T^α is the source flux through passband T, r_T is the reference flux of the simulated object with parameters α obtained from Eq. (1), and σ_T^α are the uncertainties of the source fluxes in passband T. Finally, α is a scale factor that is left free to minimise χ^2_α, as in Eq. (4); alternatively, α can be included in α, so that it can be marginalised upon or its posterior can be obtained (this is especially useful for the determination of physical parameters; see, e.g., Phosphoros; Paltani et al., in prep.).

5.2. Correction factor for passband variations

The likelihood in Eq. (4) does not take into account the possibility of filter variations. If the source flux is measured through a variation T’ of passband T, and the simulated flux with parameters α is computed using passband T, we introduce a correction factor C_{T→T’} to be applied to r_T’ in order to obtain r_T in Eq. (4) as

\[ r_T' = C_{T→T'} r_T, \]

so that we get a new equation for χ^2_α:

\[ χ^2_α = \sum_{T ∈ ℋ} \frac{(f_T' - α C_{T→T'} f_T^0)^2}{σ_T^{α}}, \]

Since we know the SED s(λ) and the passbands T and T’, we can use Eq. (1) to determine the correction factor C_{T→T’}:

\[ C_{T→T'} = \frac{r_T' r_{T'}^0}{\int_0^\infty s(λ) T(λ) dλ \int_0^\infty T'(λ) \frac{dT'}{dλ} \int_0^\infty s(λ) T(λ) \frac{dT}{dλ}}, \]

The assumption that the main parameter affecting the bias is the mean wavelength of the passband allows us to propose a very important simplification. With Δλ being the difference of mean wavelength between T and T’,

\[ Δλ = \int_λ [T'(λ) - T(λ)] dλ, \]

the correction factor can be expressed as a function of Δλ, that is, C_{T→T’} = C_{T’}(Δλ). We can thus compute C_{T’}(Δλ) for different wavelength shifts of the passband T using Eq. (7). We note that Δλ is simply the difference between the mean wavelengths of the two passbands, and, crucially, does not depend on the SED. In the case of multiple exposures with different variations of the passband, the resulting Δλ is defined as the exposure time-weighted average of the individual Δλ_i of the different exposures. We point

3 Available using Anaconda; see https://anaconda.org/astrorama/phosphoros
out that this approach is very similar to the correction of the Galactic extinction using the full knowledge of the SED as developed by Galementz et al. (2017; see our Appendix B.1 for more details), which is implemented in Phosphoros (Paltani et al., in prep.).

Appendix A describes in detail how $C^r_Ω(\Delta \lambda)$ can be approximated with an analytical function of $\Delta \lambda$. We found that we get excellent approximations of $C^r_Ω(\Delta \lambda)$ for all sets of parameters $\alpha$, all redshifts, and all four CWW templates using a second-order polynomial with the constant term fixed to 1:

$$C^r_Ω(\Delta \lambda) \approx a^r_Ω \Delta \lambda^2 + b^r_Ω \Delta \lambda + 1.$$ 

(9)

Since $a^r_Ω$ and $b^r_Ω$ depend only on the passband $T$ and on the set of parameters $\alpha$, we can precompute grids of these parameters.

### 5.3. Simulations

In order to validate our approach, we performed idealistic, noise-less simulations to estimate the bias resulting from passband variations. We used the four CWW templates (Elliptical, Sab, Sbc, Irregular; Coleman et al. 1980) in order to estimate the bias over a range of galaxy types. We simulated MegaCam $ugriz$ and Euclid $Y_J H_K$ photometry of objects modelled with the CWW templates ignoring internal reddening and photometric uncertainty over the redshift range 0–3. We chose MegaCam because of its rather large passband variations; in addition, MegaCam $u$ and $r$ passbands will be used for the northern part of the Euclid wide survey (Euclid Collaboration 2022a). Objects were simulated in each band with randomly chosen instances of its ten possible MegaCam variants (see Sect. 3.3). For the Euclid NISP passbands, we created ten arbitrary passbands by shifting the nominal passband by 0 to 8 nm. We then determined the redshift using TF using only the central passbands, ignoring passband variations. In absence of passband variations, this setup would produce perfect photometric redshifts, without any uncertainty, nor bias, so that any uncertainty or bias is entirely due to the mismatch between the passbands used to determine the source and reference fluxes, respectively. We performed three different tests of increasing complexity: firstly, only the $r$ passband can vary; secondly all passbands can vary independently; and finally fluxes are measured using a stack of four exposures, each of them having random sets of passband variations.

Figure 7 shows the resulting bias in the three configurations. When the central passbands are used to determine the photometric redshifts, some bias is clearly present at a level that is of the same order as the accuracy requirement for Euclid in all three configurations. When only the $r$ passband is randomised, the bias is completely concentrated in a specific redshift range, which matches quite well that where the Balmer break, which is present to different extents in all four CWW templates, falls into the $r$ passband. When all passbands vary, the bias is present at all redshifts and reaches about 0.007$\,(1+ z)$ in the worst case. Considering an average passband shift of 5 nm, in our toy model the bias reaches about 0.011$(1+ z)$ at $z = 0.15$, which is about 50% larger than the bias we found in the simulations. This difference was expected because the features in the CWW templates are not as sharp as the step function we used in our toy model; the fact that we are at redshift $\sim 0.5$ instead of 0.15 further smooths the transition. When multiple exposures were stacked, the bias was practically identical to that in the single-exposure case because using four exposures makes the effective transmission less variable, without changing its average.

Figure 8 shows plots similar to Fig. 7, but for the dispersion. We see again that using wrong passbands has an effect on the quality of the photometric-redshift predictions. This effect, which reaches at most 0.005$(1+ z)$ is however quite small compared to the Euclid requirement on the dispersion $[\sigma_z = 0.05(1+ z)]$ in all configurations. We note that, in the case of four exposures, the dispersion is a factor 2 lower, as expected from the averaging of four exposures.

When we applied the $C^r_Ω(\Delta \lambda)$ correction factors, we found that the bias was significantly reduced, such that it always remained within the requirements (see Fig. 7). It is even in general below 0.0005$(1+ z)$, except in the case of the “Irregular” SED, where a peak at about 0.0015$(1+ z)$ remains. This is probably due to the presence of sharp features, such as strong emission lines, in this SED. In Fig. 8 we see that using only the shifts in mean wavelengths was also able to reduce the dispersion to a large extent, except in the case of the “Irregular” SED, where some residual dispersion remains. This is again probably an effect of the presence of sharp features, such as strong emission lines, in this SED. As a conclusion, for realistically varying passbands, the knowledge of the full passbands for each objects...
is not necessary; it is sufficient to know the mean wavelengths of the passbands, which is a quantity that is much easier to handle by photometric-redshift algorithms, and in particular TF. The bias and dispersion found when using the second-order polynomial approximation of $C^0_0(\Delta \lambda)$ were extremely close to those involving the full $C^0_0(\Delta \lambda)$, demonstrating that the second-order polynomial approximation provides a very good representation of the correction factors.

6. Application to real data

We verified the capability of the method developed here to reduce the bias in the photometric-redshift determination by applying it to real data. We used the seventh (final) data release of the Legacy Survey performed at the Canada-France-Hawaii Telescope (CFHTLS\textsuperscript{4}). We used only the W1 wide-field, which we matched with the VIPERS Public Data Release 2 (VIPERS-PDR2; Scoditti et al. 2018) spectroscopic-redshift catalogue obtained with the VIMOS multi-object spectrogaph on the ESO-VLT. VIPERS is colour-selected to include mostly sources at redshifts 0.5 < z < 1.2 (Garilli et al. 2014). The CFHTLS catalogue contains the MegaCam versions of the usual ugriz passbands, with some objects being observed with a different $i$ passband denoted $g'$\textsuperscript{4}. Since we were mostly interested in the bias, we selected sources brighter than $r = 21.5$, in order to remove as much as possible the statistical uncertainties from the photometric-redshift determinations. From the VIPERS-PDR2 catalogue, we only kept very secure objects with flags either 3.5 or 4.5, excluding stars (identified with a spectroscopic redshift of 0). The match between the two catalogues resulted in 4915 objects. With the cut at bright magnitudes, we found that most of the sources are found in the redshift range 0.4 < z < 0.7.

The calibrated data we were using do not contain any information regarding the atmospheric effects on the passband; consequently the only effect that we could take into account are the spatial variations of the passbands. As discussed in Sect. 3.3, the transmission curves of the MegaCam instrument have been measured at ten different off-axis distances (Betoule et al. 2013). Using the pixel scale and pixel size of MegaCam, we could convert these physical distances into off-axis angles. The pointing strategy of CFHTLS is such that a given object is observed with the same off-axis angles in all passbands, which maximises the effect of passband variations. The shifts in the mean wavelengths of the passbands as a function of off-axis angle are shown in Fig. 9. The off-axis dependence is present in all passbands, but with quite different amplitudes. We note however that the relations are more complex than expected from purely geometric considerations, which means that the shift is spatially dependent in a more complex way that we can model here.

We first computed the photometric redshifts of the 4915 sources with a fully standard TF approach using Phosphoros. We used the 31 COSMOS templates used in Ilbert et al. (2009) and applied internal reddening up to $E_{B-V} = 0.5$ on the templates of spiral and starburst galaxies only using either the SMC extinction curve from Prevot et al. (1984), or the Calzetti et al. (2000) extinction law for star-forming galaxies, with the addition of a bump at 2175 Å introduced by Massarotti et al. (2001). Standard emission lines in Phosphoros have been added to the COSMOS templates based on the Kennicutt relation and the line flux ratios observed in sources in the SDSS-III/Baryonic Oscillation Spectroscopic Survey (Kennicutt 1998; Thomas et al. 2013; see Paltani et al., in prep., for details). We do not apply any refinement in the algorithm, such as a luminosity prior, brightness prior, or zero-point correction, because we focussed on the determination of the bias, and not on the production of the best possible catalogue. Figure 10 shows the overall quality of the photometric redshifts. While the details of the performance are not important, we obtained very good predictions even with this limited analysis, with a dispersion (measured with the normalised median absolute deviation, NMAD) of 0.038 and an outlier fraction of 2.9%. The bias appears significant, with the region with the highest density of sources lying $\Delta z \sim 0.04$ above the 1:1 relation.

Using the off-axis angles of each source, we obtained wavelength shifts for each passband by interpolating the relations shown in Fig. 9. We used then Phosphoros in the exact same configuration as above, but this time taking into account the correction for wavelength shifts presented in Sect. 5.2. We obtained

---

\textsuperscript{4} \url{http://terapix.calet.org/terapix.iap.fr/rubrique5c64.html?id_rubrique=268}

\textsuperscript{5} This $g'$ band mostly overlaps the $i$ band, and should not be confused with the usual $g$ or Euclid $Y_B$ bands, which lie at the limit between optical and near-infrared.
Fig. 9. Mean wavelength shifts of the six MegaCam transmissions as a function of the off-axis angle compared to the on-axis transmissions. The $y$ transmission has been measured at slightly different off-axis distances from the other transmissions.

a normalised median absolute deviation NMAD of 0.039 and an outlier fraction of 3.2%. Both values are very close to, but slightly worse than, those obtained without the application of the correction for the mean wavelength shifts. The additional noise could result from the too simplistic assumption we made here that the wavelength shifts only depend on the off-axis angle. Figure 11 compares the densest part of the photometric-redshift prediction plots with and without mean wavelength shifts. The bulk of the sources very clearly moved towards the 1:1 relationship when the shifts are applied, indicating a reduction in the bias, although some significant bias remains.

We quantified more precisely the reduction in the bias by computing the median difference in redshift between the photometric-redshift predictions and the true redshifts, in bins of redshift. Figure 12 compares the bias in the predictions from the original catalogue and from our computations involving the mean wavelength shifts over the redshift range $0.45 < z < 0.65$, where the bias can be reliably estimated. From Fig. 12, we found that the bias, which is positive everywhere, was reduced over this redshift interval by 0.018 in average. As expected from the field-of-view dependence of the transmissions, passband variations induce a positive bias on the photometric redshifts. This reduction matches very well the expected bias due to passband variations for an average shift of the order of 5 nm, which is the value expected in the case of the MegaCam passbands. However, at redshift $z = 0.48$, where the reduction reaches a maximum, the residual bias is about 0.016 and still exceeds the Euclid requirement of $\Delta z/(1 + z) < 0.002$ by a factor about 8 (Laureijs et al. 2011).

7. Discussion

7.1. Origin of the residual bias in photometric-redshift algorithms

The method we proposed here is able to remove a significant fraction of the bias due to the variations of the passbands. We designed the method to cope with spatial variations of the passbands, although in principle the same approach could be used for, for instance, passband variations due to atmospheric effects. In practice, this might be made difficult by the calibration process, which performs a colour-independent correction. In Appendix B we describe other processes that can lead to (at least apparent) passband variations. However, the residual bias is still larger than the requirements by a large factor. This is due to the fact that there are other sources of bias that are not due to the changes in the passbands. This could be due to other issues in the photometry. Any defect in the calibration process (for instance a wrong zero-point) will lead to biased determinations of the photometric redshift. In the latter case, a zero-point correction (Ilbert et al. 2006) can be introduced in the TF algorithms, and has been proven to be quite effective to reduce the bias. We did not use this correction here on purpose, since we wanted to focus on the effect of passband shifts. Other, in particular non-linear, issues in the photometry may not be easily removed. One example is the imperfect PSF homogenisation that needs to be applied to the different photometric bands.

Another source of bias is the mismatch between the SED models and reality. For instance, SEDs are too few, or they do not match accurately enough the SEDs of true galaxies. To
alleviate these issues, some photometric-redshift codes include linear combination of templates (EAZY; Brammer et al. 2008), or SED templates can be adapted to match better with the observed colours (Coupon et al. 2009). We point out that the latter method may also be able to remove some bias inherent to the photometry. Other model errors that cannot be corrected by these methods may contribute to the bias, such as, for instance, imprecise estimate of the Galactic reddening, incorrect Galactic reddening attenuation curve, inaccurate intrinsic reddening laws, or wrong assumptions on the IGM absorption.

Even if the models were correct, the priors could be incorrect, leading to biased estimates. In algorithms implementing Bayesian statistics the choice (or not) to apply a specific prior (e.g., luminosity function or similar, or the colour-space coverage of the SED models) leads to biases. Algorithms using maximum-likelihood are even more subject to bias, since the maximum-likelihood solution of an optimisation problem is biased when the error distribution is not symmetric. Even if one does not consciously use Bayesian statistics, some choices act as priors and affect the determination of the photometric redshifts, such as, for example, the choice of a maximum value for the internal reddening or for a galaxy luminosity.

ML algorithms have in principle significant advantages over TF algorithms with respect to the bias, as they can learn the defects in the photometry, although we argued in Sect. 4.2.2 that passband shifts are difficult to take into account. They also do not rely on our imperfect knowledge of the Universe. However, the quality of the ML model depends a lot, and in a very complicated way, on how well the training sample matches the target data set. In particular the requirement to get a spectroscopic redshift biases the training samples towards bright emission-line galaxies (see Hartley et al. 2020, for an in-depth discussion). Thus ML algorithms are also affected by model imperfections and incorrect priors, which are hard to determine objectively.

Ultimately, a bias correction remains necessary, since not all sources of bias can be identified precisely enough to be corrected. Different approaches have been proposed to this end. In the case of Euclid, a direct calibration in colour-space based on the approach developed in Masters et al. (2015) will be implemented. Such calibration at the level of the Euclid requirements remains nevertheless extremely challenging (Wright et al. 2020), so that any well-understood source of bias, such as the passband variation, should be removed beforehand as far as possible.

7.2. Comparison with the colour-term approach

Using an approach similar to the colour-term calibration (see Sect. 4.1), Betoule et al. (2013) implemented a correction for the effect of passband variations based on a colour term. Using their notation (see their Eq. (6)), the corrected magnitude $m_{i\lambda}$ of a source with magnitude $m_i$ is given by

$$m_{i\lambda} \approx m_i - \delta k(x) (c_i - c^0(x)),$$

(10)

where $x_0$ refers to the centre of the field of view, $x$ is the location of the source in the field of view, $\delta k(x)$ is the position-dependent colour term, $c_i$ is the colour of the object, and $c^0(x)$ is the reference colour for the affine transformation; we ignore here the flat-field and zero-point corrections, as they are not relevant for our discussion. Betoule et al. (2013) chose $c_i = (g - i)_0$, which is a good temperature indicator for stars. In the limit of small passband variations, $\delta k(x)$ is small, which alleviates the need to obtain accurate colours; consequently, Betoule et al. (2013) ignore the effect of passband variations in the determination of $(g - i)_0$.

We can compare the above equation with the correction in Eq. (5). Both approaches use an estimate of the SED of the source and some position-dependent factor to determine corrected flux or magnitude. A noticeable difference is that our correction is expressed as a function of a property, the mean wavelength shift, that is independent of the intrinsic properties of the object, contrarily to the observed colour. The colour-term correction is also based on a very crude approximation of the SED of the source based on a single colour, which may be unsuitable for the estimation of the SED for passbands that are far from the $g$ and $i$ passbands, such as the Euclid $Y_i,J_F,H_J$ near-infrared passbands. By contrast, in our method the SED is determined over all passbands based on all the available photometry simultaneously using empirically or physically motivated templates. Eq. (10) also introduces an unwanted correlation with the $g$ and $i$ bands.

One additional advantage of our approach is that the SED is not needed during the photometric extraction and calibration stages, which makes it easier to combine photometric catalogues from different surveys. In the case of the Euclid survey of the southern hemisphere, the Euclid near-infrared photometry will be complemented with the DES survey (and with the Rubin LSST later on; see Guy et al. 2022). The northern sky will be even more complicated, with the optical survey consisting of observations from several telescopes (Euclid Collaboration 2022a); the CFHT (Canada-France Imaging Survey; Ibata et al. 2017) for the $u$ and $r$ bands; Subaru with the Hyper Suprime-Cam (Miyazaki et al. 2018) for the $g$ and $z$ band; and PanSTARRS (Chambers et al. 2016) for the $i$ band. In order to constrain colour terms, the photometric-calibration approach would require the simultaneous analysis and calibration of all these photometric data, which can become impractical. In addition, it
7.3. Photometry corrected for passband variations

The $C_{\alpha}^T(\Delta \lambda)$ correction factors can be computed only with TF algorithms, because these are the only photometric-redshift algorithms for which the SEDs of the reference objects are known in general. ML algorithms require only a spectroscopic redshift. However, TF algorithms can be used to create corrected photometry that can later be used for any purpose, including the computation of photometric redshifts using any other algorithms, simply by applying the $C_{\alpha}^T(\Delta \lambda)$ correction factor to the photometric flux and uncertainty through passband $T$ for the best-fit $\alpha$ parameters and the measured $\Delta \lambda$ wavelength shift. The possibility to provide corrected photometric measurements has been implemented in Phosphoros.

A more advanced implementation would involve the creation of posteriors for the $C_{\alpha}^T(\Delta \lambda)$ correction factors. This would have the advantage of relying less on a specific best-fit solution, and is probably more robust. However, this makes the output more cumbersome to use, since the flux and uncertainty are replaced by a full posterior distribution. A convenient way to deal with these distributions, including the correlations between the correction factors for different passbands, is to provide a sampling of the $C_{\alpha}^T(\Delta \lambda)$ posteriors.

We note finally that relying on TF to determine the $C_{\alpha}^T(\Delta \lambda)$ does not lead to degeneracies. Indeed, even if two comparable solutions at very different redshifts exist, they would have by definition similar spectral shapes over the range of wavelengths covered by the passbands. The fact that the templates do not match exactly the SEDs of real objects is not a serious problem either, especially if the full posteriors of the correction factors are used, because what is most relevant is the range of colours provided by the templates.

7.4. Passband variations in ML algorithms

It is not straightforward to apply passband variation corrections to ML algorithms. The main advantage of these approaches is indeed that they do not rely on the knowledge of any SED, which
is necessary to compute the photometric corrections. It is possible to determine the SEDs of the objects in the training set with TF, making use of the fact that their true redshifts are known. Passband variations can be taken into account using a procedure where the SED and the photometric corrections are determined iteratively until the procedure converges on consistent SED and corrections.

Once the SEDs of the training set objects are obtained, it is possible to correct their photometric measurements to match the colour space of any source in the target sample. However, as anticipated in Sect. 4.2.2, most ML algorithms would require a potentially computationally expensive training phase to match every single colour space of the target sample. One notable exception is the $k$-nearest neighbours algorithm ($k$-NN; Cover & Hart 1967), which does not require any training. The algorithm finds the $k$ closest reference objects to a given source, and compute the value of interest using some (weighted or not) average over theirs. NNPS (Tanaka et al. 2018) is an example of the implementation of the $k$-NN algorithm for the determination of photometric redshifts. In such a case, it is straightforward to apply the $k$-NN by correcting first, for each target, the fluxes of the training set objects based on the correction factors of the target.

The $k$-NN algorithm relies on a definition of the distance between two objects. One possibility is to adopt a $\chi^2$ distance, with the likelihood given by $\exp(-\chi^2/2)$, where $\chi^2$ is given by Eq. (4), with the exception that $\sigma$ runs over the training set, instead of a grid of models. In such a case, the correction procedure and all equations from Sect. 5.2 can be used with minimal and straightforward modifications.

8. Conclusion

We studied the effects of variations of photometric passbands in the process of photometric-redshift determination. Biases due to passband variations are expected to be positive due to the usual shift of transmissions towards shorter wavelengths for off-axis sources. We found that, when taking into account the SED of the objects, it is sufficient to consider only the change in the mean wavelengths of the passbands in order to get a very accurate determination of the correction to be applied to the observed flux. Crucially, the mean wavelength can be determined irrespectively of the SED of the source. This simplification allowed us to propose an accurate and efficient correction that can be readily implemented in any TF algorithm, and has already been done in Phosphoros.

While the method we proposed here is able to remove effectively some bias in the photometric-redshift predictions, a significant bias remains. In the application to the CFHTLS photometric catalogue, we found a reduction by a factor 2 at $z = 0.5$, although it remained a factor 6 higher than the Euclid requirements. Remaining sources of bias were discussed. Bias may originate from the photometry itself. It could also be due to the inadequacy of the model: for instance, the templates might not represent accurately the SEDs of galaxies, or they may lack diversity. But the bias could also be due to the application of wrong priors, in particular hidden priors (Schmidt et al. 2020). These sources of bias could in principle be alleviated with deeper understanding of the properties of galaxies, but this remains a very challenging task. ML algorithms are immune to some photometric issues and to model-dependent effects, but are particularly sensitive to the hidden priors resulting from the construction of the training sample.

TF has different strategies to cope with biases inherent to the photometry itself, such as corrections for offset in the zero-points (Ilbert et al. 2006), or template adaptations (Coupon et al. 2009). Each of these methods should be used to remove the bias as much as possible. In many situations, ML algorithms should be largely insensitive to such biases. Passband variations, however, induce biases in the photometry that depend on the object, and taking them into account in ML algorithms would require an extremely complex training sample. Passband variations, if not corrected, can also imprint spurious spatial patterns that can bias cosmological parameter estimation; this is especially a concern for photometric galaxy clustering.

The method proposed here is able to remove most of the bias due to time and spatial variations of the passbands. While this is only one of the many biases affecting photometric-redshift determination, it is important to perform these corrections, because the cosmological requirements on the accuracy of photometric redshifts are so stringent for a mission such as Euclid that a post-processing calibration step is required; starting this step with the smallest possible bias is necessary if we want this calibration step to succeed. As a bonus, our method is able to provide corrected photometry, so that it can be used for other applications, for instance, the computation of new photometric-redshift predictions based on any other algorithm, in particular using ML.

Acknowledgements. The Euclid Consortium acknowledges the European Space Agency and a number of agencies and institutes that have supported the development of Euclid, in particular the Academy of Finland, the Agenzia Spaziale Italiana, the Belgian Science Policy, the Canadian Euclid Consortium, the French Centre National d’Études Spatiales, the Deutsches Zentrum für Luft- und Raumfahrt, the Danish Space Research Institute, the Fundação para a Ciência e a Tecnologia, the Ministerio de Ciencia e Innovación, the National Aeronautics and Space Administration, the National Astronomical Observatory of Japan, the Netherlands Organisation for Astronomical Research, the Norwegian Space Agency, the Romanian Space Agency, the State Secretariat for Education, Research and Innovation (SERI) at the Swiss Space Office (SSO), and the United Kingdom Space Agency. A complete and detailed list is available on the Euclid website (http://www.euclid-eoo.org). Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/IRFU, 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, 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 available at the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS.

References

Appendix A: Computation of the correction functions

We describe here how the correction factors \( C_f(T_\ell) \) can be computed for the transmission curve \( T(\lambda) \), in order to correct the fluxes in Eq. (6), with \( \alpha \) being an index that runs through all the model parameters. In the case of real reference objects with known SEDs, \( \alpha \) is just an integer sequence. In the case of TF, \( \alpha \) runs through all points in the computed grid, that is, \( \alpha \in \{ \alpha_i, S E D_j, E_{B-V_k} \} \), where \( i, j, k \) are integer sequences that enumerate all redshifts, all SEDs, and all internal reddening values, respectively, in the grid.

For each passband and each \( \alpha \), we compute the fluxes \( F_\lambda^T(\Delta \lambda) \) over a sequence of \( \Delta \lambda \) that encompasses (but should not largely exceed) the range of shifts in the mean wavelengths for the passband of interest. We use here 20 values between \(-10\) and \(+10\) nm. The correction is therefore:

\[
C_f^T(\Delta \lambda) = F_\lambda^T(\Delta \lambda)/F_\lambda^T(0).
\] (A.1)

We then approximate \( C_f^T(\Delta \lambda) \) with a polynomial expression. In general, we would like to minimise the number of parameters used to represent \( C_f^T \), because the number of models \( \alpha \) might be very large, and because high-degree polynomials might cause unwanted fluctuations. Figure A.1 shows a few examples of correction factors as a function of \( \Delta \lambda \) for different values of \( T \) and \( \alpha \). Investigating the \( C_f^T(\Delta \lambda) \) functions, we find that a second-degree polynomial always provides a very good approximation, requiring three parameters for each \((T, \alpha)\). However, we have the constraint \( C_f^T(0) = 1 \), so that there are effectively only two free parameters. Thus we define

\[
\hat{C}_f^T(\Delta \lambda) = \frac{C_f^T(\Delta \lambda) - 1}{\Delta \lambda}.
\] (A.2)

We point out that \( \hat{C}_f^T(\Delta \lambda) \) is well defined for \( \Delta \lambda \neq 0 \). We can then approximate \( \hat{C}_f^T(\Delta \lambda) \) with a linear function:

\[
\hat{C}_f^T(\Delta \lambda) \approx a_f^T \Delta \lambda + b_f^T.
\] (A.3)

The coefficients \( a_f^T \) and \( b_f^T \) can be obtained using a simple least-square minimisation. Finally, we have, using Eq. (A.2):

\[
C_f^T(\Delta \lambda) \approx a_f^T \Delta \lambda^2 + b_f^T \Delta \lambda + 1.
\] (A.4)

The second-order approximation from Eq. (A.4) is shown on Fig. A.1 for four arbitrary values of the passband \( T \) and of the model parameter \( \alpha \). Figure 7 shows that using the approximated correction Eq. (A.4) is practically indistinguishable from computing the flux using the full shifted passband.

Appendix B: Other effects leading to passband variations

Several other effects may affect the passband, and induce passband variations. However, depending on the details of the effects, they may be treated differently. We discuss three such effects below.

B.1. Galactic reddening

Dust in our Galaxy scatters the light emitted by extragalactic sources. The induced, extrinsic, so-called Galactic reddening depends on the direction in the sky, and its amplitude, parameterised by the value of the reddening \( E_{B-V} \), can be determined either from the observation of stars or extragalactic sources with well known intrinsic SEDs, such as quasars or passive galaxies (e.g., Schlegel et al. 1998; Schlafly et al. 2010; Schlafly & Finkbeiner 2011; Mörtsell 2013), or from direct measurement of the quantity of dust, for instance by combining Planck data with IRAS data.
100\,\mu m data (Planck Collaboration XI 2014), Galactic reddening is also strongly wavelength-dependent (hence the name), and affects the colours of the observed objects, so that its wavelength dependence can be determined from observations of well-characterised objects (Fitzpatrick 1999).

Photometry is often corrected for Galactic reddening by computing a colour term that depends only on the amount of dust in the line of sight. However, the (unknown) SED of the source needs to be taken into account in order to compute an accurate correction. Galactic reddening could be applied to all models in the grid, in order to derive, for all pass-bands, the shifts in the mean wavelength that are induced considering the model and the reddening $E_{B-V}$; however, when computing these shifts, the flux corrections are obtained at the same time. Therefore, in Galametz et al. (2017) a more direct approach has been adopted, where the flux corrections are derived directly, without computing the wavelength shifts. Thus, while Galactic reddening could be corrected by computing the wavelength shifts following the approaches we used here, in Phosphoros the two corrections are computed separately. We point out again that the work of Galametz et al. (2017) is very similar to the approach described here, and has been the inspiration for the wavelength-shift correction presented in this paper.

### B.2. Absorption in the intergalactic medium

The intergalactic medium (IGM) is filled with hydrogen that very efficiently absorbs the ultraviolet emission of objects at cosmological distances either through Lyman bound-bound transitions, or through the Lyman bound-free transition below 912 Å. Thus, while Galactic reddening could be corrected by computing a colour term that depends only on the amount of gas in the line of sight cannot be measured, and the SED cannot be recovered exactly. However, an average IGM transmission curve $G_i(\lambda_0)$, which depends strongly on the redshift $z$ and on the rest-frame wavelength $\lambda_0 = \lambda/(1+z)$, can be determined based on an analytic modelling of the distribution of gas clumps (Madau 1995; Inoue et al. 2014), or on cosmological numerical simulations (Meiksin 2006).

The IGM transmission curve $G_i(\lambda_0)$ can be seen as modifying the transmission curve $T_i(\lambda)$, such that the object is observed through an effective curve $T_i'(\lambda) = G_i(\lambda_0)/(1+z)$ $T_i(\lambda)$. However, for TF algorithms, the IGM transmission depends only on the parameters $\alpha$ of the model, so that it can be applied directly to the model SED when computing the predicted fluxes. Therefore, there is no need express the IGM absorption as an effect of pass-band variation.

### B.3. Effect of the motion of the Earth

Light from extragalactic sources is affected by Doppler effects due to the velocity $v$ of the observer. Several components contribute to $v$: the motion of the observatory due to the Earth’s rotation, or, in the case of Euclid, due to the motion of the spacecraft around the second Lagrangian point $L_2$; the revolution of the Earth, or of $L_2$, around the Solar system barycentre; and finally the peculiar motion of the Solar system barycentre with respect to the Universe’s comoving frame. Velocities due to Earth’s rotation and revolution can be very easily calculated (e.g., using the barycorrpy Python package\(^6\); Kanodia & Wright 2018). Earth’s rotation induces maximum velocities in the direction of the Earth’s equatorial plane below 1 km s\(^{-1}\), that is, $v/c \sim 10^{-6}$, which have negligible effect. Likewise, the motion of Euclid around $L_2$ has velocities below 1 km s\(^{-1}\). Earth’s revolution velocity is however of the order of 30 km s\(^{-1}\), that is, $v/c \sim 10^{-4}$, reaching a maximum along the ecliptic plane. Finally, the peculiar motion of the Solar system barycentre can be obtained from the dipole of the cosmic microwave background (CMB), and implies a peculiar velocity of 371 ± 1 km s\(^{-1}\) towards Galactic coordinates $(\ell, b) = (264\pm14, 48\pm15)$, or $v/c \sim 10^{-3}$ (Fixsen et al. 1996); this direction is also in the ecliptic plane. Velocities must be added as vectors, and only the resulting radial velocity matters (at these velocities, transverse Doppler effect can be ignored). Even though the Euclid survey will avoid the ecliptic plane because of the zodiacal light (Euclid Collaboration 2022a), the avoidance angle of 10° means that velocities up to 98% of the maximal velocities of the Earth’s revolution and Solar-system peculiar motion will occur.

The resulting radial velocities affect the photometric fluxes through Doppler boosting, which is the combination of three effects: relativistic aberration, time dilation, and blue- or red-shifting of the spectrum. To first order, the first two effects lead to an achromatic amplification $(1−3v/c)$, where a positive velocity indicates a motion away from the source. The amplification or attenuation can reach a level of a few $10^{-3}$, or a few millimag-nitudes, but it is removed by the calibration process, as it affects all sources in an exposure in the same way.

The third effect results in a shift of the observed wavelengths by a term $(1 + v/c)$ (with the same definition of the sign of $v$ as above). This induces approximately a translation of the passband of $(1 + v/c)$ times the mean wavelength of the passband. Considering the mean wavelength of the MegaCam r filter is 630 nm, the maximum shift due to these velocities is about 0.8 nm, which represents about 10% of the wavelength shift due to the off-axis dependence of the filter, and is therefore non-negligible. Since the radial velocity can be computed easily, the resulting wavelength shift can be applied to the photometric-redshift determinations using the algorithm presented here.

\(^6\) [https://pypi.org/project/barycorrpy/](https://pypi.org/project/barycorrpy/)