

Euclid: Improving redshift distribution reconstruction using a deep-to-wide transfer function[★]

Y. Kang^{1,★★}, S. Paltani¹, W. G. Hartley¹, M. Bolzonella², A. H. Wright³, F. Dubath¹, F. J. Castander^{4,5}, D. C. Masters⁶, W. d'Assignies⁷, H. Hildebrandt³, O. Ilbert⁸, M. Manera^{9,7}, W. Roster¹⁰, S. A. Stanford¹¹, N. Aghanim¹², B. Altieri¹³, S. Andreon¹⁴, N. Auricchio², H. Aussel¹⁵, C. Baccigalupi^{16,17,18,19}, M. Baldi^{20,2,21}, S. Bardelli², P. Battaglia², A. Biviano^{17,16}, E. Branchini^{22,23,14}, M. Brescia^{24,25}, J. Brinchmann^{26,27,28}, S. Camera^{29,30,31}, G. Cañas-Herrera^{32,33}, V. Capobianco³¹, C. Carbone³⁴, V. F. Cardone^{35,36}, J. Carretero^{37,38}, S. Casas^{39,40}, M. Castellano³⁵, G. Castignani², S. Cavuoti^{25,41}, K. C. Chambers⁴², A. Cimatti⁴³, C. Colodro-Conde⁴⁴, G. Congedo³², L. Conversi^{45,13}, Y. Copin⁴⁶, A. Costille⁸, F. Courbin^{47,48,49}, H. M. Courtois⁵⁰, M. Cropper⁵¹, H. Degaudenzi¹, G. De Lucia¹⁷, H. Dole¹², C. A. J. Duncan³², X. Dupac¹³, S. Dusini⁵², A. Ealet⁴⁶, S. Escoffier⁵³, M. Farina⁵⁴, R. Farinelli², S. Farrens¹⁵, F. Faustini^{35,55}, S. Ferriol⁴⁶, F. Finelli^{2,56}, N. Fourmanoit⁵³, M. Frailis¹⁷, E. Franceschi², M. Fumana³⁴, S. Galeotta¹⁷, K. George⁵⁷, B. Gillis³², C. Giocoli^{2,21}, J. Gracia-Carpio¹⁰, A. Grazian⁵⁸, F. Grupp^{10,59}, S. V. H. Haugan⁶⁰, H. Hoekstra³³, W. Holmes⁶¹, F. Hormuth⁶², A. Hornstrup^{63,64}, P. Hudelot⁶⁵, K. Jahnke⁶⁶, M. Jhabvala⁶⁷, B. Joachimi⁶⁸, E. Keihänen⁶⁹, S. Kermiche⁵³, A. Kiessling⁶¹, B. Kubik⁴⁶, M. Kümmel⁵⁹, M. Kunz⁷⁰, H. Kurki-Suonio^{71,72}, R. Laureijs⁷³, A. M. C. Le Brun⁷⁴, S. Ligi³¹, P. B. Lilje⁶⁰, V. Lindholm^{71,72}, I. Lloro⁷⁵, G. Mainetti⁷⁶, D. Maino^{77,34,78}, E. Maiorano², O. Mansutti¹⁷, S. Marcin⁷⁹, O. Marggraf⁸⁰, M. Martinelli^{35,36}, N. Martinet⁸, F. Marulli^{81,2,21}, R. J. Massey⁸², E. Medinaceli², S. Mei^{83,84}, Y. Mellier^{85,65,†}, M. Meneghetti^{2,21}, E. Merlin³⁵, G. Meylan⁸⁶, A. Mora⁸⁷, M. Moresco^{81,2}, L. Moscardini^{81,2,21}, R. Nakajima⁸⁰, C. Neissner^{7,38}, S.-M. Niemi⁸⁸, C. Padilla⁷, F. Pasian¹⁷, K. Pedersen⁸⁹, V. Pettorino⁸⁸, S. Pires¹⁵, G. Polenta⁵⁵, M. Poncet⁹⁰, L. A. Popa⁹¹, L. Pozzetti², F. Raison¹⁰, A. Renzi^{92,52}, J. Rhodes⁶¹, G. Riccio²⁵, E. Romelli¹⁷, M. Roncarelli², R. Saglia^{59,10}, Z. Sakr^{93,94,95}, A. G. Sánchez¹⁰, D. Sapone⁹⁶, B. Sartoris^{59,17}, P. Schneider⁸⁰, T. Schrabback⁹⁷, A. Secroun⁵³, G. Seidel⁶⁶, S. Serrano^{5,98,4}, P. Simon⁸⁰, C. Sirignano^{92,52}, G. Sirri²¹, L. Stanco⁵², J. Steinwagner¹⁰, P. Tallada-Crespi^{37,38}, A. N. Taylor³², I. Tereno^{99,100}, N. Tessore⁵¹, S. Toft^{101,102}, R. Toledo-Moreo¹⁰³, F. Torradeflot^{38,37}, I. Tutusaus^{4,5,94}, J. Valiviita^{71,72}, T. Vassallo¹⁷, A. Veropalumbo^{14,23,22}, Y. Wang¹⁰⁴, J. Weller^{59,10}, G. Zamorani², F. M. Zerbi¹⁴, I. A. Zinchenko¹⁰⁵, E. Zucca², J. García-Bellido¹⁰⁶, J. Martín-Fleitas¹⁰⁷, V. Scottez^{85,108}, M. Viel^{16,17,19,18,109}, and R. Teysier¹¹⁰

(Affiliations can be found after the references)

Received 5 January 2026 / Accepted 3 March 2026

ABSTRACT

The *Euclid* mission of the European Space Agency seeks to understand the Universe's expansion history and the nature of dark energy, through measurements of cosmic shear. This requires a very accurate estimate of the true redshift distribution of the galaxies, with the systematic error in the mean redshift satisfying $\sigma_{\langle z \rangle} < 0.002(1+z)$ per tomographic bin. Achieving this accuracy relies on reference samples with spectroscopic redshifts, together with a procedure to match them to survey sources for which only photometric redshifts are available. One important source of systematic uncertainty is the mismatch in photometric properties between galaxies in the *Euclid* survey and the reference objects. We develop a method to degrade the photometry of objects with deep photometry to match the properties of any shallower survey in the multi-band photometric space, preserving all the correlations between the fluxes and their uncertainties. We compare our transfer method with more demanding image-based methods, such as Balrog from the Dark Energy Survey Collaboration. According to our metrics, our method outperforms Balrog. We implement our method in the redshift distribution reconstruction, based on the self-organising map approach, and test it using a realistic sample from the *Euclid* Flagship Mock Galaxy Simulation. We find that the key ingredient is to ensure that the reference objects are distributed in the colour space the same way as the wide-survey objects, which can be efficiently achieved with our transfer method. In our best implementation, the mean redshift biases are consistently reduced across the tomographic bins, bringing a significant fraction of them within the *Euclid* accuracy requirements in all tomographic bins. Equally importantly, the tests allow us to pinpoint which step in the calibration pipeline has the strongest impact on achieving the required accuracy. Our approach also reproduces the overall redshift distributions, which are crucial for applications such as angular clustering. The agreement between the reconstructed and true distributions demonstrates both the feasibility and robustness of the approach. This implementation is sufficient for *Euclid* Data Release 1 and provides a solid foundation for subsequent data releases.

Key words. methods: data analysis – methods: statistical – galaxies: statistics – cosmology: observations

[★] This paper is published on behalf of the Euclid Consortium.

^{★★} Corresponding author: yuzheng.kang@uniige.ch

[†] Deceased.

1. Introduction

Many cosmological experiments have demonstrated that the Universe is undergoing accelerated expansion, with an equation of state of dark energy consistent with the cosmological constant (Planck Collaboration VI 2020; Abbott et al. 2023; Ghirardini et al. 2024). The confirmation of any evidence for a deviation from this value, such as the recent studies from the Dark Energy Spectroscopic Instrument (DESI) Collaboration (Adame et al. 2025), would have a significant impact on our understanding of the Universe. Measuring this equation of state and its evolution is therefore one of the major goals of modern cosmology. To achieve this goal and deepen our understanding of the dark Universe, modern surveys employ a variety of complementary probes, including weak gravitational lensing tomography, baryon acoustic oscillations, galaxy clustering, and galaxy clusters. Among these, weak lensing tomography (Hu 1999), which uses the shear of galaxy images induced by the passage of light through the cosmic web, has emerged as a powerful method to trace the evolution of the power spectrum of matter. However, it imposes some very stringent requirements on the measurement of cosmic shear and on the accuracy of redshift estimates (Amara & Réfrégier 2008).

Euclid is a medium-class mission led by the European Space Agency, specifically designed to advance this goal. It will survey approximately $14\,000\text{ deg}^2$ of the extragalactic sky, delivering high-resolution galaxy shapes in the optical I_E band with the VIS imaging instrument (Euclid Collaboration: Cropper et al. 2025) and near-infrared spectroscopy and photometry in the Y_E , J_E , and H_E bands with the NISP near-infrared instrument (Euclid Collaboration: Jahnke et al. 2025). The Euclid Wide Survey (EWS; Euclid Collaboration: Scaramella et al. 2022) nominally reaches a depth of $I_E = 24.5$ at $S/N = 10$ for extended sources (Euclid Collaboration: Scaramella et al. 2022; Euclid Collaboration: McCracken et al. 2026). These data are expected to constrain the quasi-linear approximation of the w parameter of the equation of state $p = w(z)\rho c^2$, with $w(z) = w_0 + w_a z/(1+z)$ and z the redshift, to a 1σ precision of 0.02 for w_0 and 0.1 for w_a . To meet these scientific objectives, the mission imposes stringent requirements on knowledge of the redshift distribution $n(z)$ in at least ten tomographic bins, with a systematic error in the mean redshift in each tomographic bin $\sigma_{\langle z \rangle} < 0.002(1+z)$ (Euclid Collaboration: Mellier et al. 2025). Given the vast sky coverage of the EWS and its depth, photometric redshifts are the only feasible approach for assigning redshifts to the billions of galaxies in the EWS.

Photometric redshifts are derived by the *Euclid* scientific data processing pipeline, which is designed to support both cosmology and non-cosmology science, such as galaxy evolution. This pipeline receives multi-band photometric catalogues produced upstream (Euclid Collaboration: Romelli et al. 2026) and provides for each object an estimate of the redshift, the source classification, the reconstructed spectral energy distribution (SED), and estimates of the source’s physical parameters. It also constructs the $n(z)$ redshift distributions in each bin required for weak lensing and, more generally, for 3×2 pt cosmological inference. The Quick Release 1 version of the pipeline is described in Euclid Collaboration: Tucci et al. (2026), while significant changes will be implemented for Data Release 1, in particular to support cosmological science.

Photometric redshifts offer much higher completeness, especially for faint sources, than spectroscopic redshifts. However, they are inherently subject to biases arising from different sources. In particular, the SED templates, in the case of template-

fitting algorithms, or the training sample, in the case of machine-learning algorithms, are never fully representative of the true population. These biases can propagate into cosmological analyses, where accurate redshift distributions are critical. As a result, meeting the stringent accuracy requirements using photometric redshifts alone remains a considerable challenge (Bordoloi et al. 2010). This motivates the need for additional procedures that can mitigate the systematic offsets and improve the reliability of the reconstructed redshift distribution used in cosmological inference, a process often referred to as ‘calibration’.

Modern cosmological surveys have employed a range of methods for photometric redshift calibration. One common method is direct calibration with spectroscopic redshifts originally proposed by Lima et al. (2008). Its basic principle is that the $n(z)$ of two galaxy samples should be identical if the two samples have the same magnitude and colour distributions. The method uses a spectroscopic sample in which objects are weighted according to their prevalence in the second sample, whose $n(z)$ we want to measure. However, the matching of the samples in a high-dimensional colour space becomes computationally and conceptually more and more complex as the number of photometric bands increases.

To address this issue, Masters et al. (2015) introduced a novel approach using self-organising maps (SOM; Kohonen 1990). The SOM is an algorithm in the category of manifold learning, which provides a 2D view of the complex, multi-dimensional manifold of galaxy colours, onto which spectroscopic and photometric samples can be projected. Within this framework, the redshift distribution in each cell is estimated using the number-weighted spectroscopic sample, which serves as a proxy for the true $n(z)$. This method improves upon the original direct calibration method by providing a visual and interpretable framework for the matching in colour space. Each cell is naturally occupied with galaxies with similar SEDs, and empty cells can be easily identified and either discarded or set up for follow-up observations. This was successfully demonstrated as part of the Complete Calibration of the Color-Redshift Relation (C3R2) programme (Masters et al. 2017).

The SOM calibration implicitly assumes that the photometric depth, luminosity distribution, and data quality of the spectroscopic sample match those of the sample to be calibrated. The KiDS survey (de Jong et al. 2013) benefits from extensive spectroscopic coverage that overlaps directly with the survey. As a result, KiDS constrained redshift biases to $\sigma_{\langle z \rangle} \leq 0.006$ across all tomographic bins, supported by spectroscopic data covering 99% of the full colour space of the photometric sample (Wright et al. 2020). In other cases, such as the Dark Energy Survey (DES; Dark Energy Survey Collaboration 2016), the spectroscopic sample originates from a different field with deeper observations and possibly additional passbands. Thus, its photometric properties cannot match those of the Wide survey, in particular because of the larger S/N. The SOMPZ method (Myles et al. 2021) developed for the DES addresses this issue by constructing separate SOMs for the Wide and Deep samples, connecting them through a transfer function that predicts how a deep-field source would appear when observed in the Wide survey. This information is then propagated to the wide-field SOM and used to calibrate the redshift distributions of wide-field galaxies. This approach enabled DES to achieve an uncertainty on the mean redshift in each tomographic bin $\sigma_{\langle z \rangle} \sim 0.01$ (Myles et al. 2021). The transfer function from the Deep to the Wide SOMs is implemented using the Balrog method (Everett et al. 2022), which involves injecting synthetic objects from the Deep sample into raw Wide imaging data

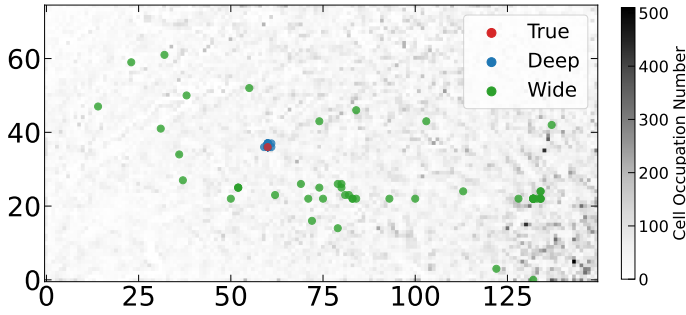


Fig. 1. Projection of sources onto the self-organising map (SOM) trained using 8-band photometric data (DES g, r, i, z bands and *Euclid* $I_E, Y_E, J_E,$ and H_E bands). Each SOM cell consists of objects with similar spectral energy distributions (SEDs). The background grayscale indicates the number of objects mapped to each cell. The red dot marks the true flux of a selected object projected onto the SOM, while the blue and green markers show 50 independent realisations of the same object with Deep and Wide photometric noise, respectively.

that are then processed in the same way as any other real source.

In the framework of the *Euclid* project, several studies on redshift calibration are explored in parallel, including using the SOM (Roster et al. 2026) and clustering redshift (d’Assignies et al. 2025). Roster et al. (2026) focuses on the issue of the definition of tomographic bins. Here, we focus on the mismatch between the photometric properties of the sources with spectroscopic redshifts and those used to determine the shear. The EWS is, indeed, in a situation similar to DES, where the photometric redshift calibration pipeline relies on the existence of specific sky areas with deep multi-band photometry and numerous spectroscopic-redshift measurements obtained from complementary surveys, the Euclid Auxiliary Fields (EAFs; Euclid Collaboration: Mellier et al. 2025), including AEGIS (Davis et al. 2007), CDFS (Giacconi et al. 2001), COSMOS (Scoville et al. 2007), the *Euclid* self-calibration field, ultra-deep field, GOODS-North (Giavalisco et al. 2004), SXDS (Furusawa et al. 2008), and VVDS (Le Fèvre et al. 2005). The EWS is approximately two magnitudes shallower than the EAFs. Thus, there are important differences in S/N. As a result, due to the different scatter in colours due to uncertainties, identical sources observed in the EWS and the EAFs cover different regions of the SOM, as illustrated in Fig. 1. As in the case of DES, it is necessary to transform the deep photometry to mimic wide-like observations before calibration. Implementing image-level simulations, as done in DES using Balrog, would, in principle, represent the most comprehensive approach to perform the transfer. However, such methods are computationally expensive and require substantial modifications to the very complex *Euclid* pipeline. We propose here a catalogue-level transformation that offers a fast and effective solution to determine the transfer function without significant additional processing or substantial modifications to the pipeline. It is important to note that our goal is not merely to reproduce wide-like measurements starting from a deep survey. Instead, we aim to study the impact of the transfer function on the calibration of $n(z)$.

This paper is structured as follows. In Sect. 2, we introduce the multi-passband transformation method and verify it using DES data. In Sect. 3, we apply this method to test its impact on the calibration of the redshift distribution in the context of *Euclid*. Section 4 presents a detailed discussion of the method, its

performance, and its broader applicability to cosmological analyses. Finally, we summarise our main conclusions in Sect. 5.

2. Transferring a deep-sample object into a wide-like one

As previously explained, because of the photometric uncertainties, a given source has a probability distribution of falling in a given SOM cell that depends on the depth of the observations, with deeper observations resulting in a more concentrated probability distribution around the true SOM cell, that is, the cell in which the source would fall if its true fluxes were known. Assuming that the tomographic bins are constructed from a set of SOM cells, the same source observed with both Deep and Wide photometry may end up in different bins (see Fig. 1), inducing differences in the $n(z)$ distributions of the cells. Thus, when constructing the SOM with sources that have deep photometry, their fluxes need to be degraded so that they are scattered across the SOM in the same way as if the sources were observed at Wide depths. For this to happen, it is crucial that the uncertainty distributions match, but also that all correlations between fluxes, uncertainties, and fluxes and uncertainties are preserved. We develop here an empirical method to match each deep catalogue source to multiple wide-like counterpart realisations given deep and wide multi-band catalogues from either simulations or real data.

2.1. Single-passband transfer

In many applications, a straightforward method is often employed to transform synthetic or deep data into wide-like observations using a fixed Gaussian error on the flux in each passband. We first add some realism by drawing the error from its distribution in the Wide sample. We call this method the single-passband transfer function (SPT). For a given object observed in the Deep sample with observed flux and uncertainty (d_X, σ_X) in passband X , the likelihood \mathcal{L} that this object is identical to a source in the Wide sample with observed flux and flux error (w_X, τ_X) is given by

$$\mathcal{L} = \frac{1}{\sqrt{2\pi} \sqrt{\sigma_X^2 + \tau_X^2}} \exp\left(-\frac{(d_X - w_X)^2}{2(\sigma_X^2 + \tau_X^2)}\right). \quad (1)$$

We compute the likelihood \mathcal{L} for all objects in the Wide sample and draw one wide-sample object (w_X, τ_X) at random following these likelihoods. The wide-like flux of the deep-sample source is given by $w'_X \sim \mathcal{N}(d_X, \tau_X^2)$, where $\mathcal{N}(a, b)$ is the normal distribution with mean a and variance b . We iterate on all passbands to degrade the deep-sample object into a wide-like one, each time potentially obtaining the wide-sample uncertainty from a different object. Applying this to all objects in the Deep sample, we obtain a degraded sample with 1D error distributions that match those of the Wide survey, provided that the assumption that σ_X is negligible compared to τ_X is true.

2.2. Multi-passband transfer

The SPT approach cannot capture the photometric correlations across the different quantities in the different bands. Poisson noise introduces a correlation between fluxes and flux errors, while flux errors across different passbands may be correlated due to the observing strategy or the local background. A realistic

method to transfer a Deep object into a wide-like object needs to preserve these correlations.

For this purpose, we propose a multi-passband transfer (MPT) method¹. We consider a deep-sample object, with a set of fluxes and flux errors ($d_{X,i}$, $\sigma_{X,i}$) in passband X_i , $i = 1, \dots, N$. For each wide-sample object with flux and flux error ($w_{X,i}$, $\tau_{X,i}$) in passband X_i , we compute the likelihood \mathcal{L} that the object is identical to the deep-sample one

$$\mathcal{L} = \prod_{i=1}^N \frac{1}{\sqrt{2\pi} \sqrt{\sigma_{X,i}^2 + \tau_{X,i}^2}} \exp\left(-\frac{(d_{X,i} - w_{X,i})^2}{2(\sigma_{X,i}^2 + \tau_{X,i}^2)}\right). \quad (2)$$

We then draw a single wide-sample object at random following these likelihoods. The procedure can be repeated to produce as many realisations of the transfer as needed. The flux errors $\tau_{X,i}$ from the selected neighbour(s) are then applied to the deep-sample object, and its new wide-like fluxes are obtained from the normal distributions

$$w'_{X,i} \sim \mathcal{N}(d_{X,i}, \tau_{X,i}^2). \quad (3)$$

This method ensures that the applied flux errors in the multiple passbands contain all the correlations present in the Wide sample.

While one should in principle compute the likelihood in Eq. (2) for all wide-sample objects, this is in general extremely inefficient, since most wide-sample sources have very small likelihoods to match the current deep-sample object. To speed up the processing, we restrict the computation of the likelihood, and hence the drawing of the flux errors, to wide-sample objects that are a priori likely to have large likelihoods. These objects are the nearest neighbours of the deep-sample source in flux space, neglecting flux uncertainties. They can be identified very efficiently using a k-d tree (Bentley 1975; Virtanen et al. 2020). The choice of the number of nearest neighbours for which the likelihoods are computed is driven by the computational efficiency on the one side and the accuracy of the transfer on the other side. We have found (see Sect. 2.4) that using the 50 nearest neighbours is sufficient to obtain an accurate transfer.

2.3. Selection function

Both the SPT and MPT methods convert all deep-sample objects into their wide-like counterparts. However, the deep-sample objects fainter than the wide-survey detection limit would not appear if they were observed at wide depth. To take this into account, after applying the deep-to-wide transformation, we introduce an empirical selection function derived from the wide-field data.

We use the distribution of the wide-field flux in the detection band X , dN_X/df_X , to construct an empirical selection function. We fit the high-flux part of dN_X/df_X with a power law with a fixed slope of -2.5 . We then divide dN_X/df_X by this power law model to obtain an empirical selection function of the wide field, randomly accepting objects with a probability given by this function. For simplicity, we perform the selection in a single band, although some surveys perform the detection in a combination of bands. For instance, Dark Energy Survey Collaboration (2016) use a combination of the $r+i+z$ bands for detection, which we approximate in the following with a simple i -band selection.

¹ The code is available at <https://github.com/yuzheng-cosmos/multi-passband-transfer>

2.4. Wide-like sample compared with DES Balrog

We tested the multi-passband transfer method using the DES Y3 Gold Catalogue (Sevilla-Noarbe et al. 2021) for the target sample and Balrog catalogue (Everett et al. 2022) as a comparison, which contains synthetic galaxies injected into single-epoch DES images and processed through the same pipeline as real wide-survey data. Out of 26 million objects, 11 million have been injected into co-add images, detected, and subsequently catalogued. Synthetic sources are created based on DES Deep Field sources (Hartley et al. 2022), providing both high S/N Deep photometry and realistic wide-like photometry from the reprocessed images. However, some injected sources originate from defective regions, for example, at the detector edges in the Deep field, or are re-injected into problematic areas in the Wide survey, leading to unrealistic photometric properties in either regime. To mitigate these effects, we applied quality cuts to the DES Deep catalogue, selecting only sources with `kNN_class=1` (galaxy classification), `badpix_frac < 0.75` (low fraction of bad pixels), and `log10(bdf_T) < 0.1 log10(bdf_flux_i)^2 + 0.2` (shear sample selection), where `bdf_T` is the model area and `bdf_flux_i` is the i -band flux from the Bulge+Disk model. These cleaned Deep sources were then cross-matched with the injected sample in the Balrog catalogue to eliminate objects that might affect the subsequent transformation.

To remove Deep sources that were injected into problematic Wide-field regions, we applied additional selection cuts on the Balrog catalogue. Specifically, we required `meas_FLAGS_GOLD_SOF_ONLY=0` (no processing issues), `meas_cm_flags=0` (successful composite model fit), all elements of `meas_psf_flags` to be zero (reliable PSF fitting), and a compact size cut on the object given by `log10(meas_cm_T) < 0.1 log10(flux_i)^2 + 0.2`, where `meas_cm_T` is the deconvolved area and `flux_i` is the measured i -band flux. We then selected 500 000 sources to form the Balrog Deep catalogue and 2 000 000 sources for Balrog Wide catalogue. For the Balrog Deep catalogue, we used `true_bdf_flux` and `true_bdf_flux_err` as the fluxes and associated errors. For the Balrog Wide catalogue, we adopted `meas_cm_flux` and extracted errors from the diagonal elements of `meas_cm_flux_cov`. We applied the same selection to the DES Y3 Gold product and randomly selected 2 000 000 sources to form the DES Wide catalogue; an additional, independent DES Wide catalogue of the same size was prepared for validation purposes. Two wide-like samples were generated from the Balrog Deep catalogue using the SPT (Sect. 2.1) and MPT (Sect. 2.2) methods. Ten realisations of each transfer have been performed.

We present the flux and flux error distributions across the four passbands in Fig. 2 for the different methods compared to those of the DES Wide data set. To evaluate the performance of the transfer methods, we adopt the Wasserstein distance (Kantorovich 1960), which is a metric that quantifies the similarity between two multi-dimensional probability distributions, with smaller values of the metric indicating a closer match. In this analysis, we compare the Balrog Deep sample (before transformation), the Balrog Wide sample (generated via image simulation), and the samples transformed using both the SPT and MPT methods against the DES Wide sample. We also used the independently drawn DES Wide catalogue, which should match the original DES Wide sample by construction, to establish a benchmark for the Wasserstein distance. The Wasserstein distances are computed across three feature spaces: all flux bands,

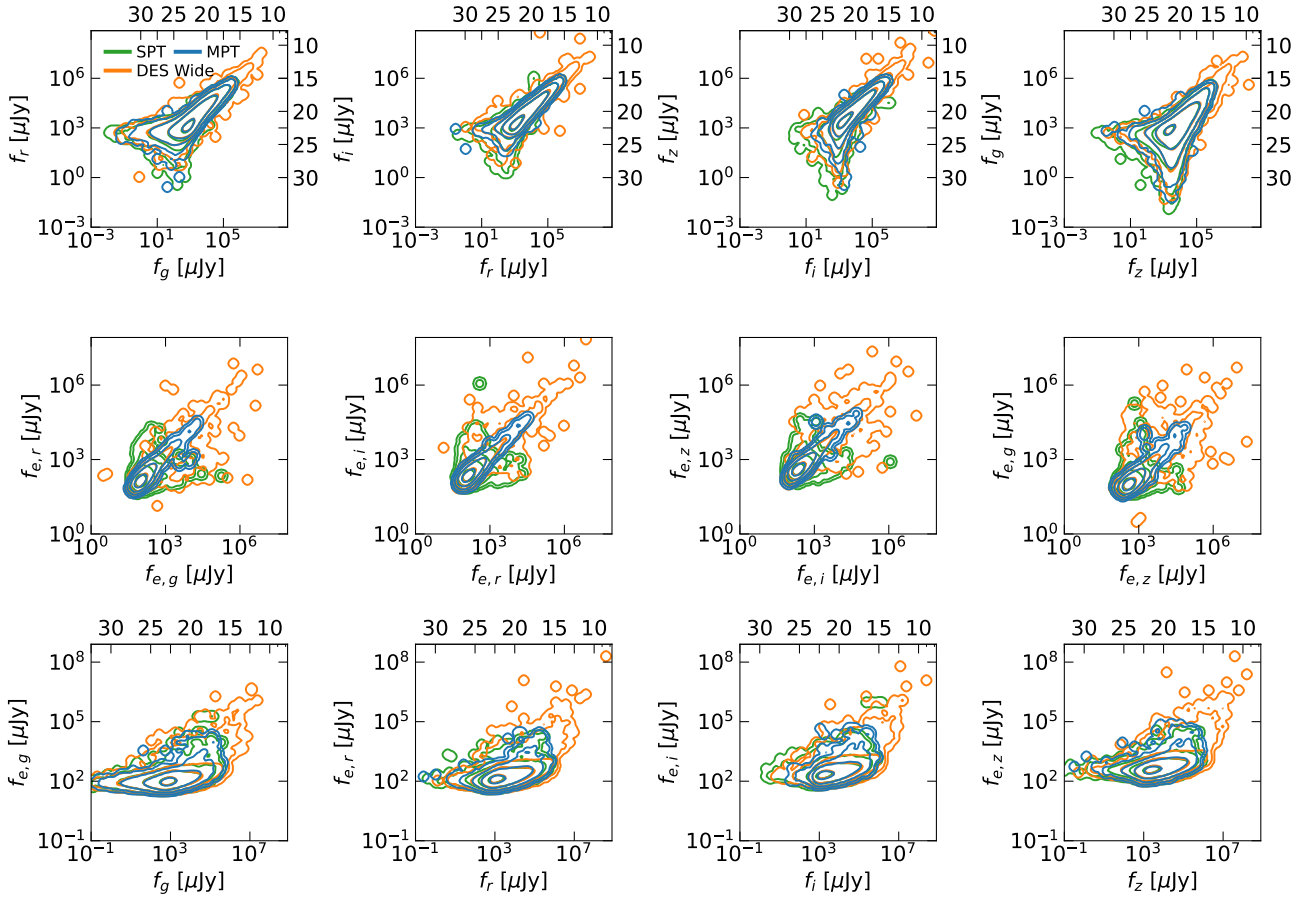


Fig. 2. Comparison between the DES Wide data set and our transformations in flux and colour space. The figure shows the distribution of the DES Wide data set (orange), compared with the multi-passband transfer data set introduced in this paper (blue), and the wide-like data set generated using the single-passband transfer (green), where f_x and $f_{e,x}$ indicate flux and flux error in passband X . The number on the top or right axis of the sub-figures indicates the corresponding magnitude.

Table 1. Wasserstein distance measured between the DES Wide reference subset and five comparison samples.

Sample	Fluxes	Errors	Fluxes + Errors
Balrog Deep	0.37	3.58	3.68
Balrog Wide	0.34	1.00	1.14
SPT	0.35	1.12	1.35
MPT	0.34	0.53	0.78
DES Wide'	0.31	0.37	0.61

Notes. Wasserstein distance across three feature spaces: fluxes, flux errors, and the combined flux and error space. ‘Balrog Deep’ and ‘Balrog Wide’ correspond to Balrog simulations with deep-like and wide-like noise levels, respectively. ‘SPT’ and ‘MPT’ refer to the transfer methods introduced in Sect. 2. ‘DES Wide’ denotes a second, independent subset of the DES Wide sample, used as a benchmark.

all flux error bands, and the full space combining both fluxes and flux errors. The results of this comparison are presented in Table 1.

While the ‘DES Wide’ to DES Wide comparison shows, as expected, the lowest Wasserstein distance for the three feature spaces, we find that all methods perform similarly in terms of reconstructing the flux-flux correlations. This is understandable, as the correlation is mostly due to the nature of the sources and, in particular, to their SEDs. However, when considering

errors, we find that all transfer methods can considerably reduce the Wasserstein distance compared to the Balrog Deep case (no transformation). Nevertheless, although Balrog Wide and the SPT methods are very similar, the MPT method significantly improves over the two other methods. This is achieved in spite of the huge computational cost of the Balrog Wide transfer method.

3. Photometric redshift distribution calibration

To assess the performance of the MPT method to improve the calibration of $n(z)$ with the SOM, we implemented it in the calibration process and applied it to the *Euclid* Flagship Simulation (Euclid Collaboration: Castander et al. 2025). This state-of-the-art galaxy mock catalogue was developed to support *Euclid*’s scientific exploitation and to test data processing and calibration algorithms designed to meet the mission’s scientific goals. The Flagship Simulation is based on an N -body simulation with four trillion particles (Potter et al. 2017), generating a light cone populated with galaxies through halo occupation distribution (Berlind & Weinberg 2002; Zheng et al. 2005) and abundance matching techniques (Vale & Ostriker 2004; Conroy et al. 2006). The final data set contains 3.4 billion galaxies up to a magnitude of $H_E < 26$, covering one octant of the sky up to redshift $z = 3$, with photometric measurements across multiple bands, observed (cosmological and peculiar) redshifts, and true (cosmological only) redshift values. In this test, we do not use

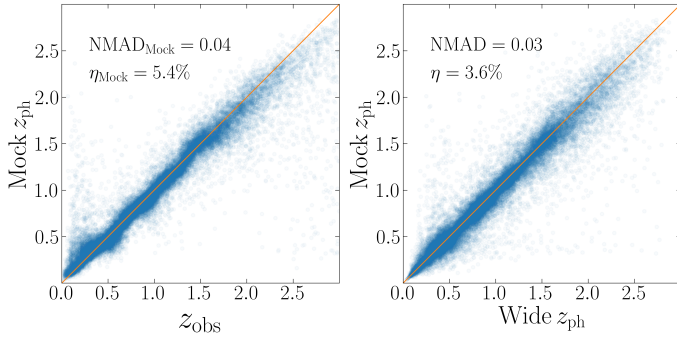


Fig. 3. Left: Comparison between RF photo- z estimates and spectroscopic redshifts for objects in the MPT-Mock sample. A magnitude cut of $I_E < 25$ and an $S/N \geq 10$ cut on the I_E band photometry are applied. Right: Comparison of RF photo- z estimates between matched objects in the Wide and MPT-Mock catalogues. The NMAD of the residuals and the outlier fractions are indicated in the figures, sources with $|z_{ph} - z_{obs}| > 0.15(1 + z_{obs})$ being defined as outliers.

real observational data because ground-truth values are required to validate the performance. Therefore, using Flagship Simulation provides us with a controlled environment in which the true bias can be measured.

By introducing the MPT method into the Flagship photometry, we can evaluate its effect on the accuracy of the recovered redshift distribution. Since Flagship Simulation realistically incorporates galaxy clustering and survey selection effects, it provides a controlled and yet representative environment to test the robustness of our approach.

3.1. Data preparation

We selected galaxies from a large sky patch defined by $220^\circ < RA < 230^\circ$ and $0^\circ < Dec < 10^\circ$, comprising 20 million objects from the Flagship Simulation version 2.1 Wide catalogue gathered from CosmoHub (Carretero et al. 2017; Tallada et al. 2020). Photometric measurements are available in the following bands: *Euclid* I_E , Y_E , J_E , and H_E bands, and DES g , r , i , z bands. The *Euclid* Flagship Simulation provides the true fluxes for each galaxy, which form the basis of our test. However, to obtain realistic observed fluxes in both Deep and Wide configurations, it is necessary to add photometric noise.

3.1.1. Preparation of the test catalogues

In the Flagship Simulation, photometric noise is added independently to each band, which does not reproduce the inter-band noise correlations present in real observations. While this approach is suitable for many validation tasks, it is not adequate for testing our method, which is designed to preserve correlations between all combinations of fluxes and flux errors across bands. To address this, we added to the true fluxes photometric noise resulting from the Poisson process and background fluctuations. Poisson noise depends on the count-to-flux conversion factor in each band and increases as the square root of the true fluxes. For background noise, we defined a correlation matrix of the uncertainties between the different passbands, with diagonal elements equal to 1 and off-diagonal elements arbitrarily set equal to 0.37 to introduce mild correlations in the amplitudes of the uncertainties. The value of 0.37 is motivated by empirical correlations measured in the *Euclid* Quick Data Release (*Euclid* Collaboration: Aussel et al. 2026). Each uncertainty is

then applied individually to their respective fluxes. We generated both Wide and Deep sample objects using the original 20 million sources, with the Wide and Deep samples assigned different Poisson and background noise levels. The reconstructed flux and flux error distributions of the new Wide sample catalogue match those of the Flagship Simulation 2.1 Wide catalogue’s observed fluxes and flux errors, except that the latter do not include error correlations. The Deep catalogue is created with a depth equivalent to 40 times the exposure time of the Wide survey. For validation purposes, we used the observed redshift (z_{obs}) provided in the Flagship Simulation as reference values for the redshift, which corresponds to the true galaxy redshift, including the contribution from the peculiar velocity. We followed the same procedure as described in Sect. 2.2 to transfer all deep-sample objects to wide-like objects. In the end, we constructed three catalogues: the Deep catalogue, the Wide catalogue, and the wide-like catalogue obtained from the Deep catalogue, which we refer to hereafter as the ‘MPT-Mock catalogue’.

3.1.2. Photometric redshift computation

The calibration of the redshift distribution requires a point estimate of the redshift for tomographic-bin assignment. Since our catalogues are generated by creating new realisations of the observed fluxes from the true fluxes, the original photo- z values from the Flagship Simulation cannot be used. As our goal is to evaluate the performance of the $n(z)$ calibration enabled by the MPT method, we aim to minimise the impact of redshift estimation on this validation. To keep the test focused, we adopted a simple machine-learning approach. We used a Random Forest (RF) regression (Breiman 2001), implemented using `RandomForestRegressor` from the `scikit-learn` Python package (Pedregosa et al. 2011) with `max_depth` set to 180 and all other parameters set to their default values, to compute the point estimates of the photo- z s. The regressor is trained using the seven photometric bands $grizY_EJ_EH_E$ from the Deep catalogue as input features and the corresponding spectroscopic redshifts as labels. The Deep catalogue is randomly split, with 70% of the objects used for training and the remaining 30% for validation. The resulting training and validation samples are obviously not consistent, because their selection functions and photometric properties differ. An estimator trained on the deep catalogue, therefore, cannot perform as accurately when applied to the wide data.

We assessed photo- z performance by comparing the predicted photo- z values with the observed redshifts z_{obs} . Figure 3 shows the comparison between the estimated photo- z and z_{obs} for the MPT-Mock-sample objects, with a normalised median absolute deviation (NMAD) of 0.04 and an outlier fraction of 5.4%, for samples that satisfy $I_E < 25$ and a $S/N \geq 10$ in the I_E band. However, this performance reflects both the quality of the MPT-Mock catalogue produced via MPT and the predictive power of the regression model. To disentangle the impact of the transfer method itself, we compared photo- z estimates of the same sources in the MPT-Mock and Wide catalogues. The MPT-Mock catalogue contains wide-like objects generated by degrading deep photometry using our method, while the Wide catalogue comprises the same objects as observed in the wide survey. By cross-matching the two sets and comparing their respective RF photo- z estimates, we find that they match very well, without evidence of additional bias, although, as expected due to the randomisation of the fluxes, there is some scatter, with NMAD = 0.03 and an outlier fraction of 3.6%. For wide-sample objects, the NMAD with respect to z_{obs} is 0.03 and the outlier

fraction is 3.4%. The corresponding figures for the MPT-Mock sample are $\text{NMAD} = 0.04$ and the outlier fraction is 5.4%. We find that the photo- z performance obtained using the MPT-Mock sample shows a slight degradation compared to that of the Wide sample, when evaluated against the observed redshift z_{obs} as the ground truth. The photo- z performance also reflects the intrinsic uncertainty of the RF model in the Deep sample. The photo- z of the Deep sample with respect to z_{obs} have an NMAD of 0.02 and an outlier fraction of 0.8%. Because transferring from photometry to photo- z is a highly non-linear process, there is no general formalism for the combination of uncertainties. To perform a rough error propagation, we added them quadratically. Regarding outliers, the maximum outlier fraction corresponds to the sum of the individual fractions. We find that the NMAD of the MPT-Mock sample matches well the quadratic propagation of errors due to the deep photometry, while the outlier fraction is a bit higher than the sum of the outlier fractions in the Deep and Wide samples.

3.2. Calibration using self-organising maps

To introduce MPT for $n(z)$ calibration, we adopted the calibration method proposed by Masters et al. (2015) as our baseline. This method has been further developed in subsequent applications of the SOM for the $n(z)$ calibration (Wright et al. 2020; Myles et al. 2021; Roster et al. 2026). We explore below several modifications making use of our MPT-Mock catalogue to identify which ones have the most significant impact on the resulting $n(z)$ reconstruction. We also used realistic numbers of reference objects and spectroscopic redshifts to avoid overoptimistic performance.

In our implementation of the original SOM-based calibration of Masters et al. (2015), which we refer to as Scenario A, we randomly selected 500 000 galaxies from the Deep catalogue to form the Deep subset, of which 15 000 galaxies are randomly selected and used as the Deep calibration subset with known z_{obs} ; these numbers have been chosen to reflect realistic numbers for *Euclid* (Masters et al. 2017, 2019; Euclid Collaboration: Stanford et al. 2021). These subsets are used to calibrate the $n(z)$ distribution of the Wide sample, from which we randomly selected 2 000 000 galaxies from the Wide catalogue to form the Wide subset. In all these samples, as well as those defined below, we applied a selection function to keep objects with the I_E $S/N \geq 10$, which is the threshold originally set to obtain a good quality shape measurement for the weak lensing analysis. We created a SOM with dimensions 75×150 and trained it with the Deep subset (see Appendix A for technical details of our SOM). A modification that we introduce in all our scenarios is that we use the ratios of the g , r , i , z , Y_E , J_E , and H_E fluxes to the I_E flux, instead of magnitude-based colours. Using these flux ratios allows the values used to train the SOM to remain roughly Gaussian distributed. We then projected all objects with z_{obs} from the Deep calibration subset onto the trained SOM. Each SOM cell containing at least one z_{obs} is assigned its average value in the cell, forming what we refer to as the z_{obs} map, and empty cells are discarded, as they are uncalibratable. We defined ten tomographic bins with redshift boundaries at 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25, and 2.5. The tomographic-bin map is then constructed by assigning each SOM cell to a bin based on its mean z_{obs} . For example, a cell with a mean z_{obs} of 0.1 is assigned to the first tomographic bin. We projected objects from both the Deep calibration subset and the Wide subset onto the trained SOM and assigned each object to a bin based on its position in the tomographic-bin map. The

weight of each cell is set by the number of wide-sample objects it contains, while the z_{obs} values of the deep-sample objects in that cell define a normalised redshift distribution in cell i , which we write $P_i(z)$. For each tomographic bin, we constructed a calibrated $n(z)$ of the source in the Wide sample as

$$n(z) = \sum_{i=1}^m N_{\text{wide},i} P_i(z), \quad (4)$$

where m is the total number of cells in the tomographic bin and $N_{\text{wide},i}$ is the number of wide-like samples in cell i . In this test setup, the Wide sample is constructed from the Flagship Simulation, so that the ground truth redshift distributions are known, which can be compared with our reconstructed $n(z)$.

As noted in Roster et al. (2026), using z_{obs} to define tomographic bins can affect the calibration performance due to the double use of the z_{obs} information – first when constructing the tomographic-bin map, and then again when constructing the $n(z)$ distribution of the weighted deep-sample objects. As a first modification to the original method of Masters et al. (2015), we defined Scenario B, where the tomographic binning is based on the mean of the photo- z s of the 500 000 deep-sample objects in each cell, the photo- z values being computed using the method presented in Sect. 3.1.2.

We then implemented four additional configurations of the calibration procedure based on Scenario B, where the MPT-Mock sample is used at different stages. The configurations for the following scenarios are also listed in Appendix B.

- Scenario C: The SOM is constructed and the tomographic-bin map is defined, as in Scenario B. The Deep calibration subset is then transformed fifty times using MPT to form the MPT-Mock calibration subset, which then consists of $50 \times 15\,000$ objects. The MPT-Mock calibration subset is then projected onto the SOM, and the weights of the SOM cells are calculated according to the MPT-Mock sample’s number density. The corresponding z_{obs} measurements from the MPT-Mock calibration subset are used to calculate the average z_{obs} in each cell.
- Scenario D: Building from Scenario B, the tomographic-bin map is defined by the average of the photo- z s of the MPT-Mock subset objects.
- Scenario E: The SOM is trained using the MPT-Mock subset instead of on deep-sample objects from the Deep subset.
- Scenario F: All the steps added in Scenarios C, D, and E are implemented: the Deep sample is fully replaced by the MPT-Mock sample throughout the pipeline. The SOM is trained on the MPT-Mock-sample objects from the MPT-Mock subset; the photo- z map and the tomographic bins are defined using photo- z values from the MPT-Mock subset, and the weight of each cell is determined from the number of MPT-Mock-sample objects from the MPT-Mock calibration subset.
- Scenario G: Similar to Scenario F, except that the objects are assigned to the tomographic bins based on their individual wide-sample photo- z , following Roster et al. (2026).

Figure 4 illustrates the redshift and occupation maps in different scenarios: a Deep SOM trained with the Deep-sample objects populated with the Deep calibration subsets shows a large fraction (46.5%) of uncalibratable cells. In Scenario D, the same SOM is populated by MPT-Mock-sample objects for the definition of the tomographic bins, showing that the distribution of redshifts on the SOM matches that of z_{obs} . The occupation of the MPT-Mock calibration subset in Scenario C leaves a much

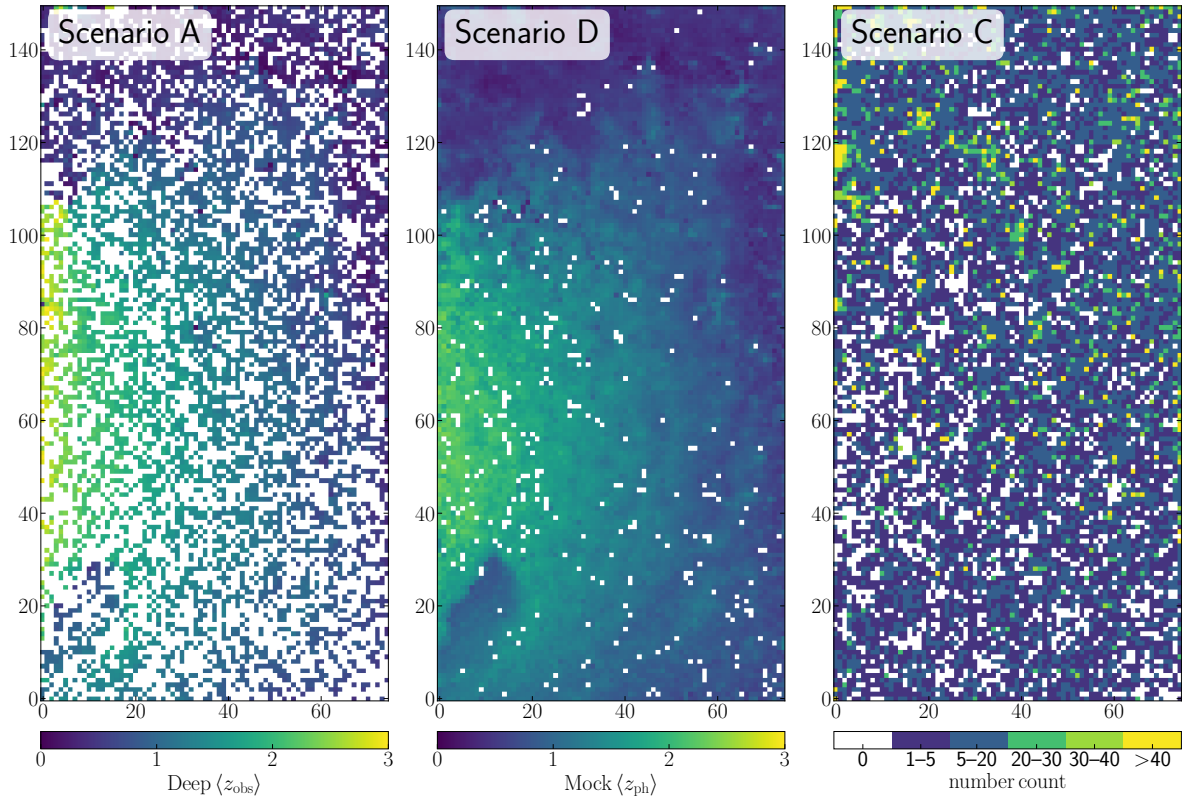


Fig. 4. SOM constructed from the Deep sample populated by redshift and objects number count. The SOM is constructed using 500 000 deep-sample objects with dimensions 75×150 . Left: SOM populated with the 15 000 deep-sample objects that have z_{obs} information, where each cell shows the mean z_{obs} of the samples it contains. This shows the original Masters et al. (2015) method. Middle: SOM populated using the 500 000 MPT-Mock-sample objects, which are the wide-like counterparts of the deep-sample objects used to construct the SOM. Each cell shows the mean photo- z of the samples it contains (Scenario D, photo- z map). Right: SOM populated with the MPT-Mock-sample objects corresponding to the 15 000 deep-sample galaxies with z_{obs} information, each replicated through 50 independent realisations, in total 750 000 objects. The colour scale indicates the number of objects per cell (Scenario C, MPT-Mock sample occupation for the z_{obs} -sample projection). White colour indicates the empty cells. The middle and right plots reflect the MPT-Mock-sample object onto a deep-sample constructed SOM.

smaller fraction of the SOM cells (6.7%) not covered with z_{obs} , due to multiple realisations of the transfer.

Figure 5 shows the bias in the mean redshift of each tomographic bin relative to the ground truth distribution for the seven calibration scenarios discussed above. To estimate the uncertainty in each measurement, we repeated the entire data selection and calibration process 500 times by drawing sources at random from the parent catalogues and computed the standard deviation and distributions of the resulting mean redshifts. Table 2 shows the average scaled biases (i.e. divided by $1 + z$) in each bin. For distributions that overlap with the *Euclid* requirements, we computed the corresponding probability as a percentage. We find that only Scenarios C, F, and G have a non-negligible probability to meet the requirements in most tomographic bins, Scenario G being the only one to have a non-negligible probability in all bins (larger than 30%).

Figure 6 shows one of the resulting $n(z)$ distributions for the tomographic bin in the redshift range $1.0 < z \leq 1.25$. The true wide-sample $n(z)$ distribution is well reproduced with the calibration method from Scenario C, while the calibration method from Masters et al. (2015) with photo- z binning (Scenario B) underestimates the width of the distribution. Quantitatively, we measure a variance of 0.098 for the true $n(z)$, 0.113 with Scenario B, and 0.099 with Scenario C. Thus, an error of 15% is obtained with Scenario B, which reaches only 1% with Scenario C. Scenario F uses a different SOM and a different binning scheme, so its true $n(z)$ distributions differ from those in Scenarios B and C.

We find that the variance in the same bin in Scenario F is 0.089, while the true $n(z)$ has a variance of 0.088, again providing an error of about 1% only.

4. Discussion

4.1. Performance of the multi-passband transfer

Because of the differences in the survey strategies, there is a mismatch in photometric properties between deep- and wide-sample objects. A common approach to mitigate this mismatch is to degrade deep-sample objects by adding random scatter to approximate the noise properties of Wide observations. However, this simple noise-adding method disrupts the correlation structure imprinted in the measurements, particularly the correlations between flux errors across different passbands.

In this work, we introduce and test successfully a multi-passband transfer method designed to simulate wide-like photometry from Deep observations. This catalogue-level transfer serves as a practical alternative to image-based simulations such as Balrog for matching deep- and wide-sample objects, offering a computationally efficient approach capable of processing one million sources in eight bands in five minutes with an Intel(R) Xeon(R) Gold 5218 CPU.

In addition, with the metrics we used here, we find that MPT recovers the statistical properties of the wide sample slightly better than Balrog. This arises possibly because Balrog

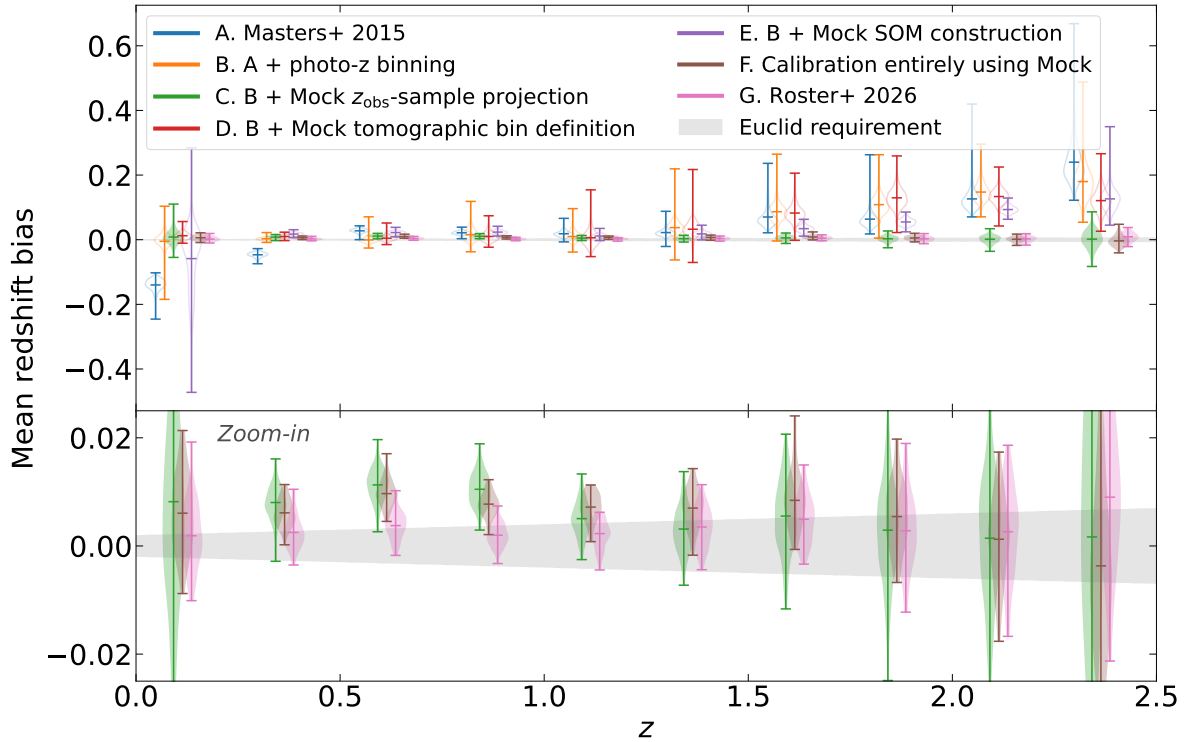


Fig. 5. Mean redshift bias for different configurations of the calibration pipeline, shown as a function of redshift for ten equal- z tomographic bins. The bias is computed relative to the true mean redshifts of the wide-sample $n(z)$ distributions. Violin points represent the distribution of 500 realisations, where violin points with face colour indicate calibration that used MPT-Mock-sample objects for projection and $n(z)$ reconstruction and points with no face colour indicate the projection and reconstruction are done by using deep-sample objects. The grey shaded region indicates the *Euclid* requirements for $n(z)$ accuracy in weak lensing cosmology. Blue data points represent Scenario A (original Masters et al. 2015 method); orange data points show Scenario B (photo- z s are used to define the tomographic binning); green ones Scenario C (z_{obs} projection replaced with MPT-Mock-sample objects); red data points show Scenario D (tomographic bin defined by MPT-Mock-sample objects); purple data points are Scenario E (SOM constructed by MPT-Mock-sample objects); brown data points show Scenario F where full calibration is based on MPT-Mock sample. The pink data points correspond to Scenario G, use the per-object photo- z binning introduced in Roster et al. (2026). The data points have been slightly shifted along the x -axis for clarity.

Table 2. Fraction of the tests that lie within the *Euclid* requirements.

	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1.0)	[1.0, 1.25)	[1.25, 1.5)	[1.5, 1.75)	[1.75, 2.0)	[2.0, 2.25)	[2.25, 2.5)
A	[0.124]	[0.034]	[0.017]	[0.011]	3%	2%	[0.027]	[0.022]	[0.040]	[0.071]
B	13%	44%	36%	16%	20%	2%	[0.033]	[0.038]	[0.047]	[0.053]
C	13%	2%	[0.007]	[0.006]	38%	66%	43%	48%	41%	21%
D	9%	3%	34%	23%	18%	5%	[0.031]	[0.045]	[0.043]	[0.036]
E	5%	[0.013]	[0.014]	[0.013]	8%	2%	[0.013]	[0.019]	[0.030]	[0.037]
F	18%	10%	3%	7%	20%	39%	48%	57%	51%	25%
G	36%	57%	41%	77%	76%	61%	51%	63%	57%	31%

Notes. When this fraction is 0, the table provides instead the average bias divided by $1 + z$, which can be directly compared to the *Euclid* requirements of 0.002; the value is then presented in square brackets.

injects sources using a synthetic model, and their photometry is extracted with a similar model, making the injected sources less realistic than actual wide-sample objects. Consequently, the flux errors measured for Balrog Wide objects are smaller than they would be if the same objects appeared in real DES Wide observations. This discrepancy can degrade the performance of Balrog Wide in recovering the wide-sample error distributions.

With the eight photometric bands of the Flagship Simulation, the transfer is performed in 16 dimensions. In principle, our probabilistic sampling should include the entire wide-sample data set, without pre-selection of the neighbours. However, since

wide surveys typically contain millions to billions of sources, this approach becomes computationally infeasible. We use a k -d tree to identify the nearest neighbours of each deep-sample object, and then perform a likelihood-based sampling within this local subset. The use of nearest neighbours as a pre-selection strategy significantly improves the efficiency with a moderate loss of precision. We find that setting $k = 50$ provides a good balance between transfer accuracy and computational performance.

Figure 2 shows that MPT successfully reproduces the flux and flux-error distributions of wide-sample objects by degrading deep-sample objects. While matching the distributions in flux

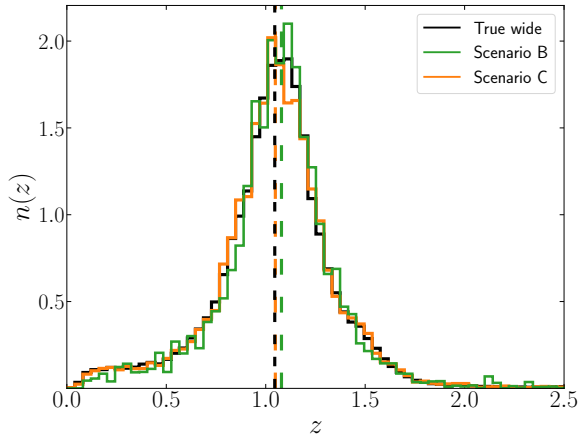


Fig. 6. Distributions $n(z)$ for the tomographic bin in the redshift range $1.0 < z \leq 1.25$. The black line shows the true $n(z)$ distribution of the wide-sample objects (mean $z = 1.045$, variance 0.098), with the dashed vertical line indicating its mean redshift. The orange line represents the distribution obtained when projecting z_{obs} from MPT-Mock-sample objects onto the SOM (Scenario C; mean $z = 1.048$, variance 0.099), where both the SOM and tomographic binning are based on the MPT-Mock sample. The green line shows the result from Scenario B (mean $z = 1.079$, variance 0.113).

and uncertainty space is a necessary first step, the true value of this method lies in its ability to support downstream scientific applications. One such application is photometric redshift estimation. The right panel of Fig. 3 illustrates the implementation of this transfer function. The photo- z values estimated from MPT-Mock-sample objects closely match those obtained from the same objects observed directly in the Wide data set. The NMAD (0.03) and the outlier fraction (3.4%) for the wide-sample objects indicate the upper limit that can be achieved by degrading the corresponding deep-sample objects. The MPT-Mock-sample object shows a slightly worse performance than with NMAD equal to 0.04 and an outlier fraction equal to 5.4%, but this can be reasonably well explained by the intrinsic scatter in the photo- z of the Deep sample objects. In addition, we do not find any evidence of bias in the photo- z introduced by the MPT. This shows that the transformation does not significantly degrade information compared to wide-sample objects and enables a reliable estimation of derived properties, such as redshift, highlighting its effectiveness beyond simple distribution matching.

4.2. Impact on the redshift distribution calibration

A basic requirement for an accurate photometric redshift calibration is the quite intuitive idea of Lima et al. (2008) that the colour distributions of test and reference objects should be identical. Since this is an extremely hard requirement, the SOM is used to partition the colour space into cells in which the occupation of both samples can be scaled. The properties of the photometric uncertainties are also significant in the matching of the colour spaces. We address this here using MPT, which degrades the photometry of deep-sample objects, in our case from the EAFs, to the performance of the wide survey, the EWS. This mimics the situation in KiDS (de Jong et al. 2013), where extensive spectroscopic coverage directly overlaps the survey area, which does not exist for *Euclid*. We test different ways to make use of this transfer method.

Figure 4 shows the Deep z_{obs} map exhibits a large fraction of unpopulated cells, with 46.5% of SOM cells lacking cover-

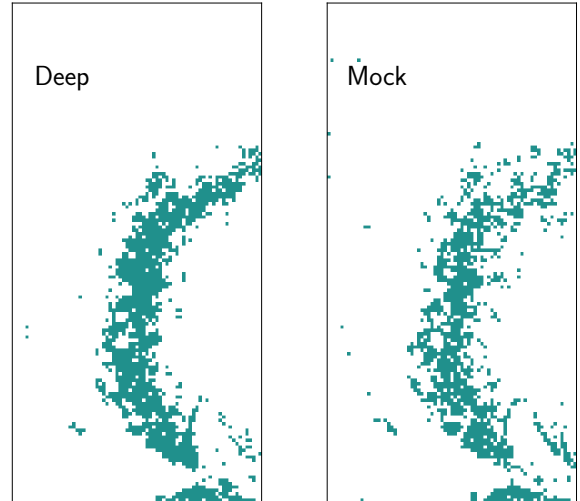


Fig. 7. Comparison of tomographic-bin areas defined using deep-sample objects and MPT-Mock-sample objects. The maps show the SOM regions where the mean photo- z of the cells falls within $1.25 < z \leq 1.5$, populated by deep-sample objects (left) and by MPT-Mock-sample objects (right). The MPT-Mock-sample set contains the same number of objects.

age and therefore having to be discarded. This shows that the number of z_{obs} is insufficient to cover the colour space of wide-survey galaxies adequately. Some of these unoccupied cells might be the result of the spread in colour space due to the wide-photometry uncertainties. In Scenario C, the colour space is populated more uniformly in the SOM through the use of multiple wide-like realisations, reducing the fraction of unpopulated cells to 3.4%. Scenarios F and G follow the same strategy and exhibit a similar improvement in SOM coverage.

Figure 5 clearly shows that MPT applied to Deep-sample objects with z_{obs} turns out to be the most important step towards improving the SOM-based calibration, as MPT improves the matching of both samples in colour space. With Scenario C, on average, 27.2% of the bins are well calibrated and 27.8% for Scenario F, while this value reaches 55% for Scenario G. As argued in Roster et al. (2026), assigning objects to tomographic bins based on their individual photo- z prevents the occurrence of situations where a SOM cell is assigned to a non-representative tomographic bin, because the taking of the mean of a bimodal distribution. Performance will improve with a larger sample of z_{obs} and better photo- z s.

Scenario F is only marginally better than Scenario C. However, conceptually, Scenario F makes more sense, especially the construction of the SOM using the MPT-mocked sample. Indeed, because of the uncertainties, the Wide-sample colour space is larger than the Deep-sample one. When the SOM is constructed using deep photometry, some wide-sample objects may end up outside the colour space specified by the SOM, and then be projected on a random cell, thereby contaminating the $N_{\text{wide},i}$ occupations.

With Scenario D, we test whether defining the tomographic binning using MPT-Mock-sample objects leads to any improvement. Figure 7 shows an example of the tomographic bin comparison between using the deep subset (Scenario B) and the MPT-Mock subset (Scenario D) to define the bin areas. On large scales, the resulting distributions are broadly consistent with each other. As shown in Fig. 5 and Table 2, the performance in this case is comparable to Scenario B, indicating that using

MPT-Mock-sample objects for tomographic binning does not provide a huge improvement. This is because, while the MPT-Mock photo- z s are less precise, the averaging over many objects makes the cell assignments nearly identical to those obtained with deep-sample objects. This might not be true if the deep and wide photometries are affected by different biases. We also notice that Scenario D performs better in lower redshift bins. This is due to the fact that the variance of the probability distributions spans a wider range, even though the mean redshift biases further deviate from the zero offset, those data points within the *Euclid* requirement will take into account the probability.

To explore the impact of the type of data used to create the SOM, we test the SOM trained on deep-sample objects (Scenario B) and the SOM trained on MPT-Mock-sample objects (Scenario E), all other steps being as in Scenario B. The photometry of deep-sample objects is close to the true values, so the SOM approximately represents the real colour space. In contrast, MPT-Mock-sample objects occupy a broader colour space due to the photometric degradation process. This corresponds to Scenario E, which actually performs the worst in our tests. This can be explained by the fact that a given location in the true colour space gets spread over many cells, while the reference objects are not, unless the technique of Scenario C is used at the same time. Scenario E therefore creates a mismatch between the colour space of the SOM and of the reference objects. Nevertheless, a SOM constructed from deep photometry spans a smaller region of the colour-space manifold than one built from wide photometry due to the larger uncertainties. Projecting wide-sample objects that lie outside the deep colour-space manifold onto a deep-based SOM can introduce bias. Although we did not observe a significant reduction in $n(z)$ bias in our tests, employing a wide-like SOM may prove to be important in principle.

In summary, we confirm that matching the photometric properties of the Deep sample and the Wide sample is necessary. This is particularly true for the spectroscopic sample, as the only scenarios that produce reasonably calibrated bins are those where the Deep calibration subset is mocked using MPT and projected onto the SOM. Multiple realisations of the mocking are needed to mitigate the shot noise introduced by the small size of the sample. Even though the steps in Scenario D and E do not show very significant improvements, they can, in principle, improve the matching of the Deep and Wide samples.

4.3. Comparison with Roster et al. (2026)

In Roster et al. (2026), the authors tested several modifications to the original SOM $n(z)$ calibration of Masters et al. (2015) using Flagship Simulations. Namely, they used a SOM constructed using wide-like photometry, and the reference objects are projected onto this SOM after additional uncertainties typical of the wide survey are applied. Each object is assigned to a tomographic bin based on its photo- z . Scenario G implements the different steps of the method presented in Roster et al. (2026), with several differences. We also focused on building a test case of the method that is as representative as possible of the *Euclid* case, in terms of the number of objects in the Deep calibration sample, their distribution on the SOM, and the properties of the different photometric samples, in particular with the introduction of correlations between errors. In Roster et al. (2026), wide-like uncertainties of the Deep sample are known from the simulated catalogues, which requires that the same sources are observed at both deep and wide depths. Instead, MPT allows us to construct a wide-like sample from a deep survey by matching the properties of any wide sample, even if there is no overlap between

the deep and wide samples. We also evaluate each step of the process one by one to understand its importance in calibration.

In this paper, we have achieved performances similar to those obtained in Roster et al. (2026) under conditions that apply to the *Euclid* weak-lensing survey, while confirming that the matching of the photometric properties of the Deep and Wide samples is a key requirement for the SOM calibration to perform correctly. We confirm their key findings, in particular regarding the need to assign sources to tomographic bins based on their individual photo- z s. Roster et al. (2026) propagated the uncertainties and biases in the tomographic-bin redshift distributions to the cosmological-parameter inference. As our results are similar, the same conclusions apply, namely that the bias on the parameter should be limited to about 0.3σ under somewhat realistic conditions.

4.4. Implications beyond weak lensing tomography

For weak lensing tomography, an accurate estimate of the mean redshift of the galaxy distribution is crucial to minimise systematic biases in cosmological parameter estimation. While the mean of $n(z)$ captures the primary requirement, knowledge of the full distribution remains necessary (Ma & Bernstein 2008; Bordoloi et al. 2010; *Euclid* Collaboration: Mellier et al. 2025). As illustrated in Fig. 6, MPT not only reproduces the mean redshift of the wide-sample distribution with high accuracy but also preserves the overall shape of the $n(z)$ distribution; in particular, the second moment is recovered to within about 1% with Scenario C or Scenario F. This demonstrates that the calibration procedure with MPT faithfully retains the statistical structure of the redshift distribution beyond just its central value. Consequently, the method has potential applications beyond weak lensing and will in particular benefit other cosmological analyses that rely on accurate modelling of the full $n(z)$ distribution, such as angular clustering (Peebles 1980).

5. Conclusions

In this work, we present and evaluate a multi-passband transfer method that creates wide-like photometry from deep-field observations. The method is computationally efficient and preserves the intrinsic correlations between fluxes and errors across all passbands, offering a significant improvement over naive noising techniques and even, with the metrics we used here, over computationally heavy image-based source injection methods such as Balrog. Applied to the photometric redshift calibration method from Masters et al. (2015) in a somewhat realistic setup, especially in terms of the number of sources in the different samples and the presence of correlated errors, MPT is instrumental in reproducing with good performance not only the mean of the $n(z)$ redshift distribution, but also its overall shape. We find that the key role of MPT lies in the projection of the Deep z_{obs} sample, which allows the distribution of these objects on the SOM to match that of identical objects from the wide sample. When all other steps are kept based on deep-sample objects, this step results in the largest improvement in calibration performance. We also confirm the need to assign objects to tomographic bins based on their individual photo- z s. In addition, the MPT method integrates naturally into the *Euclid* calibration pipeline, where wide-field objects lack corresponding deep-field observations. Moreover, by generating multiple realisations, this method distributes the z_{obs} sources around their true locations, thereby reducing the number of sparsely populated cells in the SOM. Overall, this work provides a practical solution for improving

photometric redshift calibration, with broad relevance for any cosmological studies that rely on accurate measurements of the $n(z)$ distribution. Although improvements are still needed after *Euclid* DR1, if only by improving the quality of the reference sample, this approach offers a robust framework for future data releases.

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 Agenzia Spaziale Italiana, the Austrian Forschungsförderungsgesellschaft funded through BMIMI, the Belgian Science Policy, the Canadian Euclid Consortium, the Deutsches Zentrum für Luft- und Raumfahrt, the DTU Space and the Niels Bohr Institute in Denmark, the French Centre National d'Etudes Spatiales, the Fundação para a Ciência e a Tecnologia, the Hungarian Academy of Sciences, the Ministerio de Ciencia, Innovación y Universidades, the National Aeronautics and Space Administration, the National Astronomical Observatory of Japan, the Nederlandse Onderzoeksschool Voor Astronomie, the Norwegian Space Agency, the Research Council of Finland, the Romanian Space Agency, the Swiss Space Office (SSO) at the State Secretariat for Education, Research, and Innovation (SERI), and the United Kingdom Space Agency. A complete and detailed list is available on the *Euclid* web site (www.euclid-ec.org/consortium/community/). This work has made use of CosmoHub, developed by PIC (maintained by IFAE and CIEMAT) in collaboration with ICE-CSIC. CosmoHub received funding from the Spanish government (MCIN/AEI/10.13039/501100011033), the EU NextGeneration/PRTR (PRTR-C17.I1), and the Generalitat de Catalunya.

References

- Abbott, T. M. C., Aguena, M., Alarcon, A., et al. 2023, *Phys. Rev. D*, **107**, 083504
- Adame, A. G., Aguilar, J., Ahlen, S., et al. 2025, *JCAP*, **02**, 021
- Amara, A., & Réfrégier, A. 2008, *MNRAS*, **391**, 228
- Bentley, J. L. 1975, *Commun. ACM*, **18**, 509
- Berlind, A. A., & Weinberg, D. H. 2002, *ApJ*, **575**, 587
- Bordoloi, R., Lilly, S. J., & Amara, A. 2010, *MNRAS*, **406**, 881
- Breiman, L. 2001, *Mach. Learn.*, **45**, 5
- Carretero, J., Tallada, P., Casals, J., et al. 2017, in *Proceedings of the European Physical Society Conference on High Energy Physics. 5–12 July*, 488
- Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, *ApJ*, **647**, 201
- Dark Energy Survey Collaboration (Abbott, T., et al.) 2016, *MNRAS*, **460**, 1270
- d'Assignies, W., Manera, M., Padilla, C., et al. 2025, *A&A*, **702**, A155
- Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, *ApJ*, **660**, L1
- de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, *Exp. Astron.*, **35**, 25
- Euclid Collaboration (Stanford, S. A., et al.) 2021, *ApJS*, **256**, 9
- Euclid Collaboration (Scaramella, R., et al.) 2022, *A&A*, **662**, A112
- Euclid Collaboration (Castander, F., et al.) 2025, *A&A*, **697**, A5
- Euclid Collaboration (Cropper, M., et al.) 2025, *A&A*, **697**, A2
- Euclid Collaboration (Jahnke, K., et al.) 2025, *A&A*, **697**, A3
- Euclid Collaboration (Mellier, Y., et al.) 2025, *A&A*, **697**, A1
- Euclid Collaboration (Aussel, H., et al.) 2026, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202554610>
- Euclid Collaboration (McCracken, H. J., et al.) 2026, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202554594>
- Euclid Collaboration (Romelli, E., et al.) 2026, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202554586>
- Euclid Collaboration (Tucci, M., et al.) 2026, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202554588>
- Everett, S., Yanny, B., Kuropatkin, N., et al. 2022, *ApJS*, **258**, 15
- Furusawa, H., Kosugi, G., Akiyama, M., et al. 2008, *ApJS*, **176**, 1
- Ghirardini, V., Bulbul, E., Artis, E., et al. 2024, *A&A*, **689**, A298
- Giacconi, R., Rosati, P., Tozzi, P., et al. 2001, *ApJ*, **551**, 624
- Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJ*, **600**, L93
- Hartley, W. G., Choi, A., Amon, A., et al. 2022, *MNRAS*, **509**, 3547
- Hu, W. 1999, *ApJ*, **522**, L21
- Kantorovich, L. V. 1960, *Manage. Sci.*, **6**, 366
- Kohonen, T. 1990, *IEEE Proc.*, **78**, 1464
- Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, *A&A*, **439**, 845
- Lima, M., Cunha, C. E., Oyaizu, H., et al. 2008, *MNRAS*, **390**, 118
- Ma, Z., & Bernstein, G. 2008, *ApJ*, **682**, 39
- Masters, D., Capak, P., Stern, D., et al. 2015, *ApJ*, **813**, 53
- Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2017, *ApJ*, **841**, 111
- Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2019, *ApJ*, **877**, 81
- Myles, J., Alarcon, A., Amon, A., et al. 2021, *MNRAS*, **505**, 4249
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *J. Mach. Learn. Res.*, **12**, 2825
- Peebles, P. J. E. 1980, *The Large-Scale Structure of the Universe* (Princeton University Press)
- Planck Collaboration VI. 2020, *A&A*, **641**, A6
- Potter, D., Stadel, J., & Teyssier, R. 2017, *Comput. Astrophys. Cosmol.*, **4**, 2
- Roster, W., Wright, A. H., Hildebrandt, H., et al. 2026, *A&A*, **707**, A277
- Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, **172**, 1
- Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., et al. 2021, *ApJS*, **254**, 24
- Tallada, P., Carretero, J., Casals, J., et al. 2020, *Astron. Comput.*, **32**, 100391
- Vale, A., & Ostriker, J. P. 2004, *MNRAS*, **353**, 189
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nat. Methods*, **17**, 261
- Wright, A. H., Hildebrandt, H., van den Busch, J. L., & Heymans, C. 2020, *A&A*, **637**, A100
- Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, *ApJ*, **633**, 791

- ²⁹ Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy
- ³⁰ INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
- ³¹ INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy
- ³² Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
- ³³ Leiden Observatory, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands
- ³⁴ INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy
- ³⁵ INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy
- ³⁶ INFN-Sezione di Roma, Piazzale Aldo Moro, 2 – c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy
- ³⁷ Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain
- ³⁸ Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain
- ³⁹ Institute for Theoretical Particle Physics and Cosmology (TTK), RWTH Aachen University, 52056 Aachen, Germany
- ⁴⁰ Deutsches Zentrum für Luft- und Raumfahrt e. V. (DLR), Linder Höhe, 51147 Köln, Germany
- ⁴¹ INFN section of Naples, Via Cinthia 6, 80126 Napoli, Italy
- ⁴² Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
- ⁴³ Dipartimento di Fisica e Astronomia “Augusto Righi” – Alma Mater Studiorum Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
- ⁴⁴ Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain
- ⁴⁵ European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy
- ⁴⁶ Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, Villeurbanne F-69100, France
- ⁴⁷ Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
- ⁴⁸ Institutíó Catalana de Recerca i Estudis Avançats (ICREA), Passeig de Lluís Companys 23, 08010 Barcelona, Spain
- ⁴⁹ Institut de Ciències de l'Espai (IEEC-CSIC), Campus UAB, Carrer de Can Magrans, s/n Cerdanyola del Vallés, 08193 Barcelona, Spain
- ⁵⁰ UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, 4 rue Enrico Fermi, 69622 Villeurbanne, France
- ⁵¹ Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
- ⁵² INFN-Padova, Via Marzolo 8, 35131 Padova, Italy
- ⁵³ Aix-Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
- ⁵⁴ INAF-Istituto di Astrofisica e Planetologia Spaziali, via del Fosso del Cavaliere, 100, 00100 Roma, Italy
- ⁵⁵ Space Science Data Center, Italian Space Agency, via del Politecnico snc, 00133 Roma, Italy
- ⁵⁶ INFN-Bologna, Via Irnerio 46, 40126 Bologna, Italy
- ⁵⁷ University Observatory, LMU Faculty of Physics, Scheinerstr. 1, 81679 Munich, Germany
- ⁵⁸ INAF-Osservatorio Astronomico di Padova, Via dell'Osservatorio 5, 35122 Padova, Italy
- ⁵⁹ Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstr. 1, 81679 München, Germany
- ⁶⁰ Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, 0315 Oslo, Norway
- ⁶¹ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
- ⁶² Felix Hormuth Engineering, Goethestr. 17, 69181 Leimen, Germany
- ⁶³ Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark
- ⁶⁴ Cosmic Dawn Center (DAWN), Denmark
- ⁶⁵ Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France
- ⁶⁶ Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
- ⁶⁷ NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
- ⁶⁸ Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- ⁶⁹ Department of Physics and Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
- ⁷⁰ Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland
- ⁷¹ Department of Physics, P.O. Box 64, University of Helsinki, 00014 Helsinki, Finland
- ⁷² Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
- ⁷³ Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
- ⁷⁴ Laboratoire d'étude de l'Univers et des phénomènes eXtremes, Observatoire de Paris, Université PSL, Sorbonne Université, CNRS, 92190 Meudon, France
- ⁷⁵ SKAO, Jodrell Bank, Lower Withington, Macclesfield SK11 9FT, UK
- ⁷⁶ Centre de Calcul de l'IN2P3/CNRS, 21 avenue Pierre de Coubertin, 69627 Villeurbanne Cedex, France
- ⁷⁷ Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy
- ⁷⁸ INFN-Sezione di Milano, Via Celoria 16, 20133 Milano, Italy
- ⁷⁹ University of Applied Sciences and Arts of Northwestern Switzerland, School of Computer Science, 5210 Windisch, Switzerland
- ⁸⁰ Universität Bonn, Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
- ⁸¹ Dipartimento di Fisica e Astronomia “Augusto Righi” – Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, 40129 Bologna, Italy
- ⁸² Department of Physics, Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK
- ⁸³ Université Paris Cité, CNRS, Astroparticule et Cosmologie, 75013 Paris, France
- ⁸⁴ CNRS-UCB International Research Laboratory, Centre Pierre Binétruy, IRL2007, CPB-IN2P3, Berkeley, USA
- ⁸⁵ Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014 Paris, France
- ⁸⁶ Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
- ⁸⁷ Telespazio UK S.L. for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ⁸⁸ European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
- ⁸⁹ DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 155, 2200 Copenhagen, Denmark
- ⁹⁰ Centre National d'Etudes Spatiales – Centre spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
- ⁹¹ Institute of Space Science, Str. Atomistilor, nr. 409 Măgurele, Ilfov 077125, Romania
- ⁹² Dipartimento di Fisica e Astronomia “G. Galilei”, Università di Padova, Via Marzolo 8, 35131 Padova, Italy
- ⁹³ Institut für Theoretische Physik, University of Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany
- ⁹⁴ Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France
- ⁹⁵ Université St Joseph, Faculty of Sciences, Beirut, Lebanon
- ⁹⁶ Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile
- ⁹⁷ Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria

- ⁹⁸ Satlantis, University Science Park, Sede Bld, 48940 Leioa-Bilbao, Spain
- ⁹⁹ Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, PT1749-016 Lisboa, Portugal
- ¹⁰⁰ Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
- ¹⁰¹ Cosmic Dawn Center (DAWN)
- ¹⁰² Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark
- ¹⁰³ Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, Plaza del Hospital 1, 30202 Cartagena, Spain
- ¹⁰⁴ Caltech/IPAC, 1200 E. California Blvd., Pasadena, CA 91125, USA
- ¹⁰⁵ Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstr. 12-14, 69120 Heidelberg, Germany
- ¹⁰⁶ Instituto de Física Teórica UAM-CSIC, Campus de Cantoblanco, 28049 Madrid, Spain
- ¹⁰⁷ Aurora Technology for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ¹⁰⁸ ICL, Junia, Université Catholique de Lille, LITL, 59000 Lille, France
- ¹⁰⁹ ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing, Via Magnanelli 2, Bologna, Italy
- ¹¹⁰ Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA

Appendix A: Technical note on the construction of the SOM

Another important consideration is the computational performance and stability of the SOM algorithm, especially when projecting large numbers of sample objects onto the map. Some SOM implementations in interpreted languages are too slow to deal with the millions of objects found in modern surveys. Some packages produce unwanted features in the trained maps, such as patterns of empty cells at regular intervals, or cells that are populated with an unrealistically large fraction of the objects. In the SOM redshift map, there are sharp discontinuities between regions of high- and low-redshift cells, leaving valleys of empty cells at the transition boundaries. These issues can severely limit the applicability of such tools in large-scale survey contexts, where both efficiency and robustness are essential. Given that the available packages do not comply with all the needs of *Euclid* in terms of parallelisation, robustness and control of the input parameters, for this work, we adopted the SOM package developed by the Swiss *Euclid* Science Data Centre, available at https://github.com/astrorama/PHZ_SOMBiasCorrection/tree/0.18.1, which has been implemented from scratch in C++ and integrated into the official *Euclid* data processing pipeline. This implementation not only provides stable and reliable map construction, free from the aforementioned issues, but also offers fast projection speed, making it well-suited for large surveys.

During SOM construction, as in Wright et al. (2020) and Roster et al. (2026), we test various input configurations, including a combination of colours plus *Euclid* I_E magnitude, or more general combinations of colours plus a single passband magnitude. However, the additional dimension led to an increased number of empty cells. Moreover, due to the uncertainties, some objects contain negative flux values that cannot be converted into magnitudes or colours and have to be discarded from the SOM, which biases the colour space. Even when the flux is positive, fluxes close to the detection limit generate magnitudes that are strongly biased towards high values. However, such cuts can introduce selection biases and reduce the sample size. In contrast, using flux ratios at the SOM construction level while making sure that the flux in the denominator is well above the detection limit allows all objects, including those with negative fluxes in the other bands, to be retained. Taking into account all these considerations, we concluded that using flux ratios was more robust than using colours and magnitudes. We point out in addition that the cut on the *Euclid* I_E S/N is similar to that which is applied by the *Euclid* weak lensing pipeline.

Appendix B: Calibration configurations

Scenario	SOM construction	Binning assignment	Binning redshift	Projection sample
A	Deep	by SOM cell	Deep spec- z	Deep
B	Deep	by SOM cell	Deep photo- z	Deep
C	Deep	by SOM cell	Deep photo- z	MPT-Mock
D	Deep	by SOM cell	MPT-Mock photo- z	Deep
E	MPT-Mock	by SOM cell	Deep photo- z	Deep
F	MPT-Mock	by SOM cell	MPT-Mock photo- z	MPT-Mock
G	MPT-Mock	by object	photo- z	MPT-Mock