Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **519,** 1792–1808 (2023) https://doi.org/10.1093/mnras/stac3585
Advance Access publication 2022 December 8

Mapping variations of redshift distributions with probability integral transforms

J. Myles [1,2,3,4★] D. Gruen,[4] A. Amon [5,6] A. Alarcon,[7] J. DeRose,[8] S. Everett,[9] S. Dodelson,[10,11] G. M. Bernstein [12] A. Campos,[10] I. Harrison,[13] N. MacCrann,[14] J. McCullough,[2] M. Raveri,[12] C. Sánchez [12] M. A. Troxel,[15] B. Yin,[10] T. M. C. Abbott,[16] S. Allam,[17] O. Alves,[18] F. Andrade-Oliveira,[18] E. Bertin,[19,20] D. Brooks,[21] D. L. Burke,[2,3] A. Carnero Rosell,[22,23,24] M. Carrasco Kind,[25,26] J. Carretero,[27] R. Cawthon,[28] M. Costanzi,[29,30,31] L. N. da Costa,[23] M. E. S. Pereira,[32] S. Desai,[33] P. Doel,[21] I. Ferrero,[34] B. Flaugher,[17] J. Frieman,[17,35] J. García-Bellido,[36] M. Gatti,[12] D. W. Gerdes,[37,18] R. A. Gruendl,[25,26] J. Gschwend,[23,38] G. Gutierrez,[17] W. G. Hartley,[39] S. R. Hinton,[40] D. L. Hollowood,[41] K. Honscheid,[42,43] D. J. James,[44] K. Kuehn,[45,46] O. Lahav,[21] P. Melchior,[47] J. Mena-Fernández,[48] F. Menanteau,[25,26] R. Miquel,[49,27] J. J. Mohr,[50,4] A. Palmese,[51] F. Paz-Chinchón,[25,5] A. Pieres,[23,38] A. A. Plazas Malagón,[47] J. Prat,[52,35] M. Rodriguez-Monroy,[48] E. Sanchez,[48] V. Scarpine,[17] I. Sevilla-Noarbe,[48] M. Smith,[53] E. Suchyta,[54] M. E. C. Swanson, G. Tarle,[18] D. L. Tucker,[17] M. Vincenzi,[55,53] and N. Weaverdyck[18,8] (DES Collaboration)

Affiliations are listed at the end of the paper

Accepted 2022 December 2. Received 2022 November 21; in original form 2022 August 24

ABSTRACT

We present a method for mapping variations between probability distribution functions and apply this method within the context of measuring galaxy redshift distributions from imaging survey data. This method, which we name PITPZ for the probability integral transformations it relies on, uses a difference in curves between distribution functions in an ensemble as a transformation to apply to another distribution function, thus transferring the variation in the ensemble to the latter distribution function. This procedure is broadly applicable to the problem of uncertainty propagation. In the context of redshift distributions, for example, the uncertainty contribution due to certain effects can be studied effectively only in simulations, thus necessitating a transfer of variation measured in simulations to the redshift distributions measured from data. We illustrate the use of PITPZ by using the method to propagate photometric calibration uncertainty to redshift distributions of the Dark Energy Survey Year 3 weak lensing source galaxies. For this test case, we find that PITPZ yields a lensing amplitude uncertainty estimate due to photometric calibration error within 1 per cent of the truth, compared to as much as a 30 per cent underestimate when using traditional methods.

Key words: gravitational lensing: weak – methods: numerical – galaxies: distances and redshifts.

1 INTRODUCTION

The matter density field of the Universe and its evolution over time relate directly to the cosmological model of the Universe. Galaxy surveys provide observable proxies of the matter density field, and thus can be used to place competitive constraints on parameters of cosmological models. Specifically, experiments such as the Dark Energy Survey (DES), Kilo-Degree Survey (KiDS), and the Hyper Suprime-Cam Survey (HSC) as well as the future Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST), *Euclid*, and *Roman* Space Telescope missions measure statistics such as correlation functions of galaxy positions and shapes to probe the underlying matter density field (Green et al. 2011; Laureijs et al. 2011; LSST Dark Energy Science Collaboration 2012; Hildebrandt et al. 2017; Abbott et al. 2018; Hikage et al. 2019; Hildebrandt et al. 2020; Heymans et al. 2021; Abbott et al. 2022). In these analyses, determining the impact of weak gravitational lensing on the observed galaxy images provide crucial information to relate observations to the underlying matter density field that galaxies live in. Among the data products needed for these experiments, *redshift distributions*, which encode the relative contribution of galaxies at different redshifts to the gravitational lensing signal observed loom large due to their key role in enabling interpretation of the effect of weak lensing on the apparent shapes and sizes of galaxies (For a review, see e.g. Newman & Gruen 2022. See also Huterer et al. 2006; Lima et al. 2008; Cunha et al. 2012; Hildebrandt et al. 2012; Benjamin et al. 2013; Bsemi Huterer, Cunha & Fang 2013; Bonnett et al. 2016; Samuroff et al. 2017; BsemiHoyle et al. 2018; BsemiWright et al. 2020a,b; Euclid Collaboration et al. 2020; Joudaki et al. 2020; Tessore & Harrison 2020; BsemiMyles

★ E-mail: jmyles@stanford.edu

et al. 2021; Hildebrandt et al. 2021; BsemiCabayol et al. 2022; Gatti et al. 2022; Sánchez et al. 2022).

In lensing survey nomenclature, the term 'redshift distribution' refers to a function describing the relative probability of a galaxy in a sample to have come from a particular narrow redshift histogram bin. A typical lensing survey will divide its data set into a few tomographic bins, each with its own redshift distribution. We highlight that a redshift distribution is distinct from the photometric redshift for any individual galaxy, and the uncertainty requirements of redshift distributions are likewise distinct from uncertainty requirements of individual galaxy photometric redshifts. As reducing systematic uncertainties in redshift distributions is necessary to meet uncertainty goals on estimated cosmological parameters, greater attention is being drawn to the importance of modelling redshift distribution uncertainty with sufficient complexity (see e.g. Malz et al. 2018; Hadzhiyska et al. 2020; Malz 2021; Myles et al. 2021; Stölzner et al. 2021; Cordero et al. 2022; Malz & Hogg 2022; Zhang et al. 2022). Redshift distributions have been historically described as a single probability density function together with, for example, a shift parameter describing uncertainty on the mean redshift value (e.g. Hoyle et al. 2018). More recently, redshift distributions have been described as *joint* probability distribution function (PDF) for redshift histogram bin heights, meaning each bin in a redshift histogram has a full associated PDF (see e.g. Leistedt et al. 2019; Sánchez & Bernstein 2019; Alarcon et al. 2020), or alternatively as an *ensemble* of slightly varying PDFs that collectively describe the full uncertainty in knowledge of galaxy redshift (see e.g. Hildebrandt et al. 2017; Myles et al. 2021). In this work, we present a method for characterizing such an ensemble of PDFs that collectively represent the knowledge of the redshift distribution for a galaxy sample.

Measuring and quantifying the uncertainty of redshift distributions often involves detailed studies of simulated galaxy catalogues, where particular sources of error can be tightly controlled. For example, simulation codes easily facilitate changes in the number and spatial extent of galaxies used, biases in the assumed distribution of true galaxy redshifts, and the level of photometric noise in the survey. In this work, we present a methodology for mapping the variation present in an ensemble of redshift distributions measured in simulations to redshift distributions measured from the data, and vice versa. Our methodology relies on probability integral transformations to transfer the variation in an ensemble of distributions to another fiducial distribution. We call this method PITPZ for the probability integral transformations (PITs) that characterize and enable it and for the redshift 'z' distributions that it is designed to help to estimate. Although this method is designed and discussed in the context of relating effects measured in cosmological simulations to analogous measurements on data, its potential for application is notably broader than this.

This paper is organized as follows: in Section 2, we describe the PITPZ method and its differences compared to related existing methods, in Section 3, we discuss how we implement our method as software, in Section 4, we derive quantities conserved by the transformations of the method, in Section 5, we show an example use of this method for propagating photometric calibration uncertainty to redshift distributions of galaxies in the Dark Energy Survey, in Section 6, we show results of the experiment outlined in Section 5, and in Section 7 we conclude.

A flat ΛCDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_\mathrm{m} = 0.3$ is assumed throughout this work. Other cosmological parameters are taken to be consistent with *Planck* 2018 ΛCDM cosmology Planck Collaboration VI (2020).

2 METHOD

This section describes the PITPZ method for transferring the variation measured in one ensemble of distributions to another distribution. We provide a visual illustration of the method in Fig. 1 to accompany the text of this section.

In our description of the PITPZ method, we use notation $p(z)$ to denote the probability distribution function of a random variable z of interest. In this work, the variable of interest is galaxy redshift for a weak lensing sample of galaxies, but we refer only to abstract general probability distributions in Sections 2, 3, and 4, because our method is broadly applicable to any problem with an ensemble of probability distribution functions describing some uncertainty. We thus defer specific redshift discussion until the analyses discussed in the sections thereafter.

PITPZ requires two inputs and produces one output. Namely, the two inputs are:

I. A fiducial $p(z)$ measurement or ensemble of measurements. We denote this ensemble with $p^{\mathrm{fid.}}(z)$. While only one such measurement is needed for the purposes of this algorithm, the algorithm accommodates having an ensemble of fiducial $p(z)$ measurements to, for example, sequentially propagate multiple independent sources of uncertainty.

II. An ensemble of redshift distributions whose variation we want to map to $p^{\mathrm{fid.}}(z)$. We call this ensemble the input ensemble and denote it with $p_i^{\mathrm{in.}}(z)$, where i is an index for each realization in the ensemble.

The sole output is:

III. An ensemble of $p(z)$ whose variation is related to the variation between realizations of the input ensemble but which is mapped onto $p^{\mathrm{fid.}}(z)$. We call this ensemble the output ensemble and denote it with $p^{\mathrm{out.}}(z)$. We describe quantitatively the relationship between the variation of the input ensemble and the variation of the output ensemble in Section 4.

We begin by computing the inverse cumulative distribution function (inverse CDF, also called the quantile function) F_i^{-1} for each realization $p_i(z)$ in the input ensemble. This can be written as

$$F_i^{-1}(p) = \{z : F_i(z) = p\}, \tag{1}$$

where the CDF is defined as

$$F_i(z) = \int_{-\infty}^{\infty} p_i(z')\,\mathrm{d}z' = \int_{0}^{z_{\mathrm{max.}}} p_i(z')\,\mathrm{d}z'. \tag{2}$$

The integral transforming $p(z)$ to $F(z)$ is called a probability integral transformation (Dodge et al. 2006). Our method relies on these transformations to generate the cumulative distribution functions necessary to subsequently produce a transformation that transfers variation from the input ensemble onto $p^{\mathrm{fid.}}(z)$.

We note that our method, while making use of PITs, differs from past uses of PITs for galaxy redshift estimation. Such past work includes the use of PITs to assess redshift biases by taking advantage of the fact that the PIT of a proper PDF is uniformly distributed, so deviations from uniform distributions in PITs computed from redshift PDFs indicate the presence of biases in these underlying PDFs (see e.g. Bordoloi, Lilly & Amara 2010; Polsterer, D'Isanto & Gieseke 2016; Freeman, Izbicki & Lee 2017; Tanaka et al. 2018; Schmidt et al. 2020; Shuntov et al. 2020; Hasan et al. 2022; Zhang et al. 2022). Our method, by contrast, uses PITs to construct another transformation entirely, which is used to alter $p(z)$ to make them more like some other $p(z)$, as to be described in greater detail in the following text.

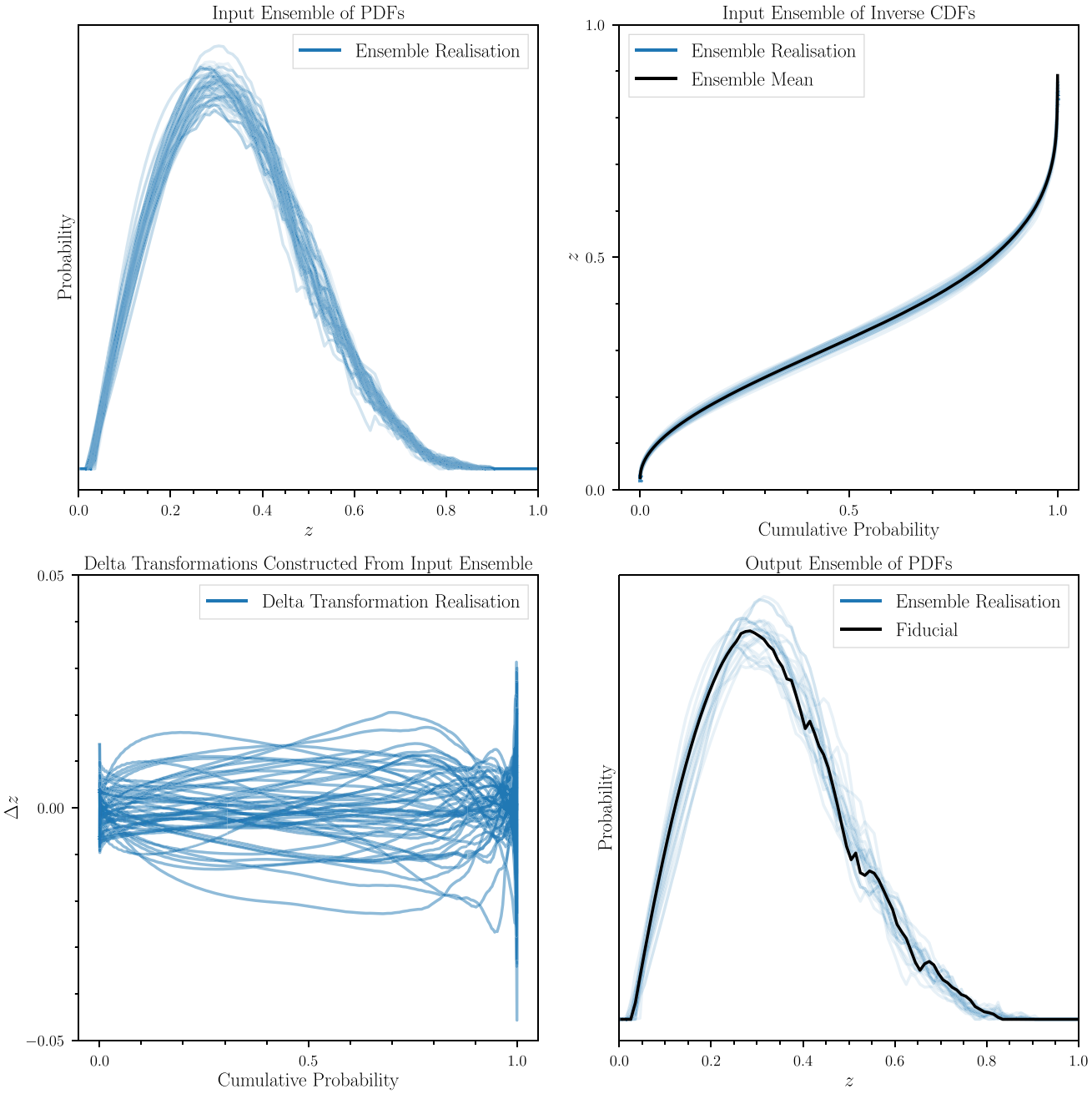

Figure 1. PITPZ method used to propagate the uncertainty associated with the mock ensemble shown in the top left-hand panel onto the mock fiducial curve of the bottom right-hand panel. *Top left*: Input ensemble of PDFs. The variation between these curves is the information we want to transfer. *Top right*: Input ensemble of CDFs. *Bottom left*: Delta transformations constructed from the input ensemble by taking the difference of inverse CDFs with respect to the mean inverse CDF. *Bottom right*: Output ensemble of PDFs constructed by applying delta transformations to the inverse CDF of the fiducial $p(z)$, then converting the result to a PDF.

We define a new transformation which we call a delta transformation (denoted here as T) as the difference between the inverse CDF F_i^{-1} of a given realization in the input ensemble and the *average* inverse CDF of the input ensemble:

$$T_i = F_i^{-1,\text{in.}} - \langle F^{-1,\text{in.}} \rangle. \tag{3}$$

Given this definition, each delta transformation encodes the difference between a given realization of the input ensemble and the mean of the realizations of said input ensemble. We apply these transformations by adding each delta transformation to the inverse CDF $F_{\text{fid.}}^{-1}$ of the fiducial data $n(z)$:

$$F_i^{-1,\text{out.}} = F_i^{-1,\text{fid.}} + T_i. \tag{4}$$

Given this ensemble of transformed inverse CDFs of $p^{\text{fid.}}(z)$, we construct the output ensemble by taking the inverse of these inverse CDFs to yield CDFs, then differentiating to yield PDFs:

$$p_i^{\text{out.}}(z) = \frac{\text{d}}{\text{d}z}\left(F_i^{\text{out.}}\right). \tag{5}$$

3 IMPLEMENTATION

The conceptual algorithm described in Section 2 for generating an ensemble of $p(z)$ involves manipulating smooth probability density and quantile functions. We circumvent implementation complications that arise from operating on smooth functions by *evenly sampling* each PDF to generate an ordered list of n samples $\{z_0, \ldots, z_n\}$ from each $p(z)$ and manipulating these samples, rather than the quantile functions directly as follows. In practice, the probability density functions used are often stored digitally as histograms, in which case our sampling procedure avoids complications related to differing normalizations and bin size and range.

In brief, applying a delta transformation (as in equation 4) amounts to generating an ordered list of samples z from each $p(z)$, adjusting the values of those samples with the delta transformation, and computing the distribution of the adjusted samples for a specified histogram binning. We first determine the number of samples to be apportioned to each histogram bin, then use those samples to compute and apply each delta transformation, and finally compute the new $p(z)$ from each array of ordered, adjusted samples.

We use the largest remainder method to apportion the discrete samples among histogram bins as closely to the bins' relative probability as is possible (Tannenbaum 2010). This method consists of dividing the total number n of samples to be apportioned by the histogram value $p(z)$ of each histogram bin. Each bin is apportioned a number of samples equal to the integral part of its respective quotient. The histogram bins are then ranked by the size of their remainders, and each bin is assigned an additional sample until the remaining samples have been fully allocated. This procedure is done for the fiducial distribution $p(z)$, and for each realization $p_i^{\text{in.}}(z)$ constituting the input ensemble. After using this method to compute the appropriate number of samples apportioned to each bin, we distribute those samples evenly across the width of the bin. This yields the following sets of ordered redshift values:

I. 1 (or more) set $\{z_0, z_1, \ldots, z_n\}^{\text{fid.}}$
II. N sets $\{z_0, z_1, \ldots, z_n\}_i^{\text{in.}}$

Here the j^{th} value $z_{i,j}$ of the i^{th} set of ordered redshift samples $\{z\}_i$ represents the redshift corresponding to the $\frac{j}{n}^{\text{th}}$ quantile of the distribution. In other words, these samples constitute the quantile function for $p(z)$.

We then compute the delta transformations by taking the difference of each ordered sample of a realization in the input ensemble and the corresponding ordered sample for the mean of these realizations:

$$\begin{aligned} T_i &= \{z_0, z_1, \ldots, z_n\}_i^{\text{in.}} - \{z_0, z_1, \ldots, z_n\}^{\langle \text{in.} \rangle} \\ &= \{\Delta_0, \Delta_1, \ldots, \Delta_n\}_i. \end{aligned} \tag{6}$$

Applying these delta transformations amounts to adding each of these Δz values to the value of its corresponding quantile in the list of ordered samples of $p^{\text{fid.}}(z)$. For a single delta transformation $T = \{\Delta_0, \Delta_1, \ldots, \Delta_n\}$, the implementation of equation (4) is then:

$$\left\{z_0^{\text{out.}}, z_1^{\text{out.}}, \ldots, z_n^{\text{out.}}\right\} = \left\{z_0^{\text{fid.}} + \Delta_0, z_1^{\text{fid.}} + \Delta_1, \ldots, z_n^{\text{fid.}} + \Delta_n\right\}. \tag{7}$$

We note that as a result of the delta transformation some samples can be shifted outside of the range of acceptable values, e.g. below zero in the case of cosmological redshift. In the case of redshift distributions, we discard these samples and increase the value of the remaining samples such that the mean redshift of the distribution is not changed. Once we have the perturbed samples described by equation (7), constructing the final modified $p(z)$ is done by binning the samples with any given histogram bin edges, which is done in lieu of equation (5).

4 CONSERVATION RULES OF DELTA TRANSFORMATIONS

Recall the goal of the PITPZ method: we aim to propagate uncertainties to measured redshift distributions. Past analyses have used coherent shifts of measured redshift distributions to lower and higher values, with the shifts drawn from a Gaussian distribution whose standard deviation encapsulates mean redshift uncertainty (see e.g. Hoyle et al. 2018). This approach produces an output ensemble of PDFs that only varies in mean redshift, but in reality many sources of uncertainty produce more complicated variations than simple mean shifts. The goal of PITPZ is to preserve the full correlation structure across an input ensemble in a constructed output ensemble. This section is dedicated to illustrating how this information is conserved by the PITPZ method.

Recall that the starting point for applying the PITPZ method is two inputs: a fiducial measured $p^{\text{fid.}}(z)$ (or an ensemble of such fiducial measurements) and an input ensemble $p_i^{\text{in.}}(z)$ of redshift distributions whose variation encodes uncertainty due to some relevant effect(s). Our algorithm produces an output ensemble $p_i^{\text{out.}}(z)$ which has mapped the variation in the input ensemble onto the fiducial measurement $p^{\text{fid.}}(z)$. Posing the question of information conservation in the broadest possible sense, we want to relate each central moment of each realization in $p_i^{\text{in.}}(z)$ to the corresponding central moment of its counterpart realization in $p_i^{\text{out.}}(z)$. We proceed by deriving the conservation rules for the mean, variance, and skewness of a realization of the output ensemble in terms of the corresponding moments of the fiducial $p(z)$, the realization of the input ensemble used, and the mean of the realizations of the input ensemble. Fig. 2 shows the performance of our software implementation of PITPZ to conserve the rules derived for mean and variance. Inspection of this figure illustrates that PITPZ produces an output $p(z)$ realization whose mean differs from the fiducial in proportion to how the mean of the corresponding realization of the input ensemble differs from the mean of the input ensemble. By contrast, mean shifts maintain this relationship only when sufficiently far from the edges of the allowed parameter limits. The fact that the observed numerical noise lies within the LSST uncertainty region illustrates that the deviation from conservation of the mean value is negligible for near-term weak lensing redshift calibration applications. PITPZ preserves a similar relationship for the variance, but mean shifts do not transfer the relative change in width of realizations in the input ensemble to the constructed output ensemble. Although for the source of uncertainty propagated for this figure (see Section 5), the changes in $p(z)$ width introduced by the mean shift method are within the LSST year 10 target uncertainty, it is the combined value for all sources of uncertainty that should be ultimately compared to the target error budget. In practice, using PITPZ may be necessary to meet the LSST year 10 target uncertainties.

In this section, we introduce the following notation convention: Overlines represent averages over the redshift value samples, which are indexed with j. For example, the mean redshift z of $p(z)$ is represented by $\bar{z}$. Brackets represent averages over the redshift distribution realizations of an ensemble, which are indexed by i. For example, the mean $p(z)$ of the input ensemble, $p_i^{\text{in.}}(z)$, is represented by $p^{\langle \text{in.} \rangle}(z)$.

4.1 Mean of redshift distributions

Measuring the mean redshift of each constituent realization of the input ensemble yields a distribution of mean redshifts $\bar{z}_i^{\text{in.}}$ where $\bar{z}_i^{\text{in.}} = \int z' p_i^{\text{in.}}(z')\, \mathrm{d}z'$. We aim to derive the relation between

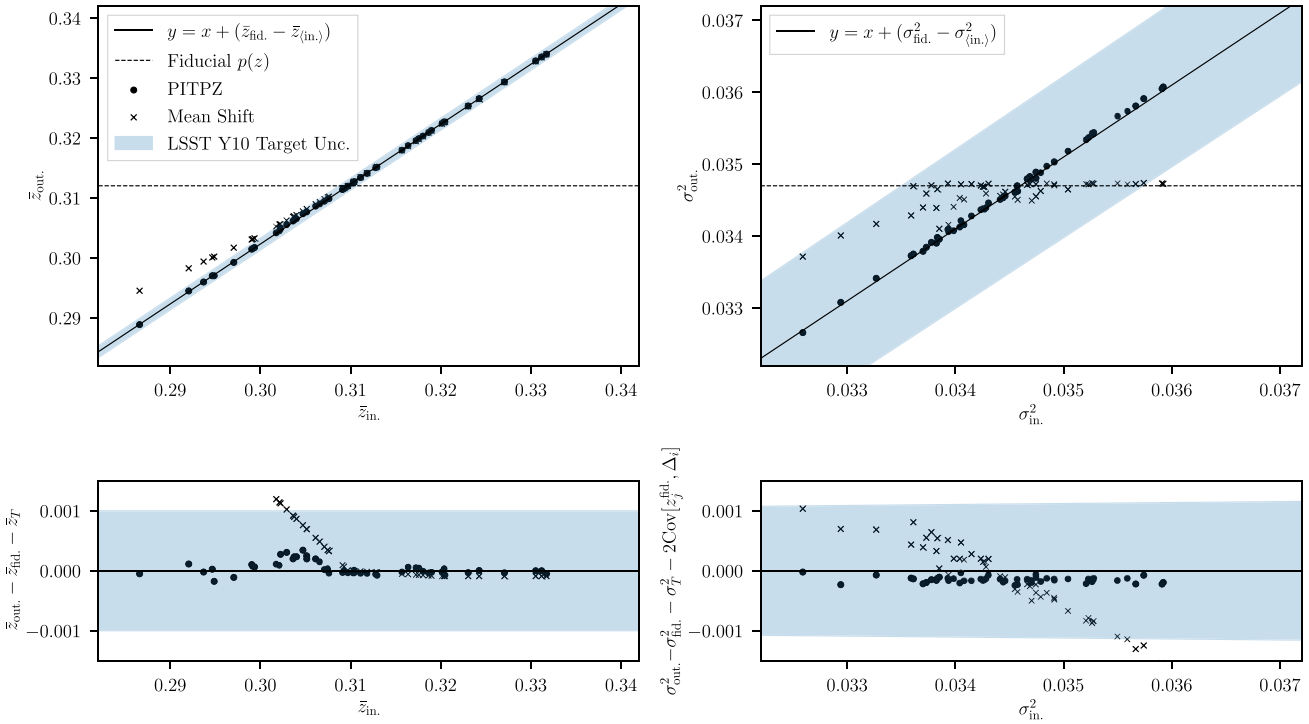

Figure 2. Illustration of mean and variance conservation by the PITPZ method and of improved behaviour compared to mean shifts. Shown here are results for the first tomographic bin of the experiment described in Section 5. *Top*: Relationship in redshift distribution moments between the input ensemble and output ensemble realizations. *Bottom*: Deviations from the conservation rules derived in Section 4 due to numerical noise in our software implementation of the formalism described. The blue uncertainty region corresponds to the LSST Y10 WL analysis uncertainty requirements of $0.001\,(1+z)$ on the mean and $0.003\,(1+z)$ on the standard deviation (here scaled to variance) of redshift at $z = 0$ (The LSST Dark Energy Science Collaboration et al. 2018).

each mean redshift in this ensemble and the mean redshift of the corresponding output in the output ensemble produced by the PITPZ algorithm, $\bar{z}^{\rm out.}_i$.

As introduced in Section 3, we can represent a given realization of the input ensemble $p^{\rm in.}(z)$, a given delta transformation T, and the resulting realization of the output ensemble $p^{\rm out.}(z)$ as a set of ordered samples:

$$\begin{aligned} p^{\rm in.}(z) &\Leftrightarrow \left\{z^{\rm in.}_0, z^{\rm in.}_1, \ldots, z^{\rm in.}_n\right\} \\ T &\Leftrightarrow \left\{\Delta_0, \Delta_1, \ldots, \Delta_n\right\} \\ p^{\rm out.}(z) &\Leftrightarrow \left\{z^{\rm out.}_0, z^{\rm out.}_1, \ldots, z^{\rm out.}_n\right\} \\ &= \left\{z^{\rm fid.}_0 + \Delta_0, z^{\rm fid.}_1 + \Delta_1, \ldots, z^{\rm fid.}_n + \Delta_n\right\}. \end{aligned} \tag{8}$$

It is straightforward to prove that the mean redshift of each realization of the output ensemble is the sum of the mean redshift of the fiducial $p(z)$ and the mean value of the shifts comprising the delta transformation. In the following we use our customary labels of 'in.' and 'out.' to represent single realizations of the input and output ensembles, respectively, and the letter T to likewise represent a single delta transformation. With this convention, each input-output pair follows the following conservation rule:

$$\begin{aligned} \bar{z}^{\rm out.} &= \frac{1}{n}\sum_j^n z^{\rm out.}_j \\ &= \frac{1}{n}\sum_j^n \left(z^{\rm fid.}_j + \Delta_j\right) \\ &= \frac{1}{n}\sum_j^n z^{\rm fid.}_j + \frac{1}{n}\sum_j^n \Delta_j \\ &= \bar{z}^{\rm fid.} + \bar{\Delta} \\ &= \bar{z}^{\rm fid.} + \bar{z}^{\rm in.} - \bar{z}^{\langle \rm in. \rangle}. \end{aligned} \tag{9}$$

4.2 Higher order moments of redshift distributions

We present results for the variance and skewness here, deferring the full derivation to Appendix A.

Our expression for the variance of a realization in the output ensemble is

$$\begin{aligned} \sigma^2_{\rm out.} &= \sigma^2_{\rm fid.} + \sigma^2_T + 2\,{\rm Cov}\left[z^{\rm fid.}_j, \Delta_j\right] \\ &= \sigma^2_{\rm fid.} + \sigma^2_{\rm in.} + \sigma^2_{\langle \rm in. \rangle} - 2\,{\rm Cov}\left[z^{\rm in.}_j, z^{\langle \rm in. \rangle}_j\right] + 2\,{\rm Cov}\left[z^{\rm fid.}_j, \Delta_j\right]. \end{aligned} \tag{10}$$

Our expression for the skewness of a realization in the output ensemble is:

$$\begin{aligned}\sigma^3_{\rm out.}\widetilde{\mu}^{\rm out.}_3 &= \sigma^3_{\rm fid.}\widetilde{\mu}^{\rm fid.}_3 + \sigma^3_{\rm in.}\widetilde{\mu}^{\rm in.}_3 + \sigma^3_{\langle{\rm in.}\rangle}\widetilde{\mu}^{\langle{\rm in.}\rangle}_3 \\ &+ 3\sigma^2_{\rm fid.}\sigma_{\rm in.}S\left(z^{\rm fid.}_j, z^{\rm fid.}_j, z^{\rm in.}_j\right) \\ &- 3\sigma^2_{\rm fid.}\sigma_{\langle{\rm in.}\rangle}S\left(z^{\rm fid.}_j, z^{\rm fid.}_j, z^{\langle{\rm in.}\rangle}_j\right) \\ &+ 3\sigma_{\rm fid.}\sigma^2_{\rm in.}S\left(z^{\rm fid.}_j, z^{\rm in.}_j, z^{\rm in.}_j\right) \\ &- 6\sigma_{\rm fid.}\sigma_{\rm in.}\sigma_{\langle{\rm in.}\rangle}S\left(z^{\rm fid.}_j, z^{\rm in.}_j, z^{\langle{\rm in.}\rangle}_j\right) \\ &+ 3\sigma_{\rm fid.}\sigma^2_{\langle{\rm in.}\rangle}S\left(z^{\rm fid.}_j, z^{\langle{\rm in.}\rangle}_j, z^{\langle{\rm in.}\rangle}_j\right) \\ &- 3\sigma^2_{\rm in.}\sigma_{\langle{\rm in.}\rangle}S\left(z^{\rm in.}_j, z^{\rm in.}_j, z^{\langle{\rm in.}\rangle}_j\right) \\ &+ 3\sigma_{\rm in.}\sigma^2_{\langle{\rm in.}\rangle}S\left(z^{\rm in.}_j, z^{\langle{\rm in.}\rangle}_j, z^{\langle{\rm in.}\rangle}_j\right), \end{aligned} \tag{11}$$

where the S denotes the coskewness of three random variables X, Y, and Z:

$$S(X, Y, Z) = \frac{\mathrm{E}[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))(Z - \mathrm{E}(Z))]}{\sigma_X \sigma_Y \sigma_Z}. \tag{12}$$

5 COSMOLOGICAL IMPACT ANALYSIS

Having defined PITPZ as a statistical method and illustrated the rules by which it conserves and transfers information from one distribution of PDFs to another, we now turn to understand how this can affect scientific conclusions in the context of weak lensing cosmology experiments. For the remainder of this work, we choose to denote our probability distribution function of interest as $n(z)$ to remain consistent with the redshift calibration literature, in which $n(z)$ represents a weighted number density of galaxies at redshift z, where each galaxy's may be weighted according to its contribution to the associated shear catalogue (for more information about weight choices see e.g. Gatti et al. 2021). We note that $n(z)$ has a different normalization than the probability density function of a galaxy in the survey having a specific redshift and emphasize that $n(z)$ is not the probability distribution function for the redshift of an individual galaxy.

Weak gravitational lensing refers to the accumulated deflections to the path of light from a distant source galaxy as it travels through the large-scale structure of the Universe toward an observer. In order to interpret the coherent distortions in the shapes of large samples of observed galaxies due to this effect, we must have a constraint on the redshift of the source galaxies and the intervening distribution of lensing matter. In this context, the salient question is how using PITPZ to generate $n(z)$ realizations whose variation encodes uncertainties in the redshift distributions of the selected galaxy sample will affect the uncertainty on parameters of the cosmological model being tested with weak lensing analyses. In practice, the relationship between variations of $n(z)$ realizations and cosmology uncertainty is that evaluating the cosmology likelihood function given weak lensing data should sample over an ensemble of $n(z)$ realizations. For the purpose of our work, the question of how $n(z)$ uncertainty and cosmology are related can be reduced to assessing the impact that using PITPZ to construct redshift distributions has on the resulting distribution of lensing signal amplitude (for a given lens redshift). To this end, we first briefly summarize the way galaxy photometry is used in the redshift calibration scheme applied in this work, deferring to Myles et al. (2021) for a full description.

5.1 DES Year 3 redshift methodology

The DES Y3 redshift calibration relies on a method called SOMPZ developed to take advantage of the DES deep-drilling fields, where longer exposure times and spatial overlap with near-infrared surveys provides more information to use for redshift inference (Buchs et al. 2019; Myles et al. 2021; Hartley et al. 2022). In this method, the deep-field galaxies serve as an intermediary between galaxies with secure (e.g. spectroscopic) redshifts and the overall wide-field sample; the deep-field galaxies play the crucial role of enabling secure redshifts to be used for subsamples of galaxies while avoiding selection bias between the secure redshift sample and galaxies in the overall wide-field survey sample (for more information on such selection bias, see Gruen & Brimioulle 2017). Within this scheme, redshift distributions are computed in small regions of deep-field colour-magnitude space. The wide-field galaxy density is determined in small regions of wide-field colour-magnitude space. The ultimate calibrated redshift distributions of the wide-field sample are the weighted sum of redshift distributions in deep-field colour-magnitude space, where weights are the likelihood of given deep galaxies being detected and selected in the wide-field sample as determined using the BALROG image simulation package (Everett et al. 2022). SOMPZ is additionally combined with independent information from galaxy clustering and shear ratios (Myles et al. 2021; Gatti et al. 2022; Sánchez et al. 2022). The final product of this kind of redshift calibration is not a single $n(z)$, but rather an ensemble of $n(z)$ whose variations encode the uncertainty. This ensemble can be used in cosmology analyses by sampling the ensemble for each evaluation of the cosmological likelihood function. PITPZ is designed as a method for generating such an ensemble to be sampled in cosmology analyses.

5.2 Experimental design

Among the several sources of uncertainty inherent to the DES Year 3 redshift methodology, the photometric calibration of the deep-field galaxies stands out due to the novel use of these galaxies to improve our calibration. This uncertainty is best understood by taking advantage of realistic simulations in which photometric calibration error can be easily scaled at will. We therefore choose this source of uncertainty to illustrate the characteristics of our PITPZ method for propagating uncertainty.

Our experimental design to illustrate the impact of PITPZ consists of the procedure described in the following test and illustrated in Fig. 3.

We begin with an ensemble of 100 $n(z)$ produced using the Buzzard simulations (DeRose et al. 2019), where each realization has zero-point offsets according to the photometric calibration uncertainty measured by Hartley et al. (2022) are introduced to the deep-field photometry. The variation between the $n(z)$ realizations in this ensemble reflects the uncertainty in $n(z)$ due to deep-field photometric zero-point uncertainty.

We split this ensemble into two halves of 50 realizations each. The first half is used to construct delta transformations relative to the mean. Because it is used in this way, the first half serves the role of the input ensemble as defined in §2, so it is labelled $n^{\rm in.}(z)$. The second half is to construct the fiducial $n^{\rm fid.}(z)$: $n^{\rm fid.}(z)$ is simply the mean of the $n(z)$ comprising the second half.

We apply the delta transformations made from the first half (i.e. from the input ensemble) to this fiducial $n^{\rm fid.}(z)$. As an alternative to applying the delta transformations, we also apply to the fiducial $n^{\rm fid.}(z)$, the mean shifts corresponding to the difference in mean redshift between each realization of the input ensemble and the mean of the realizations of the aforementioned input ensemble; this is a simpler alternative to PITPZ, which has been employed for past redshift calibration analyses (e.g. Jee et al. 2013; Bonnett et al. 2016; Hoyle et al. 2018). As a result, we have produced two

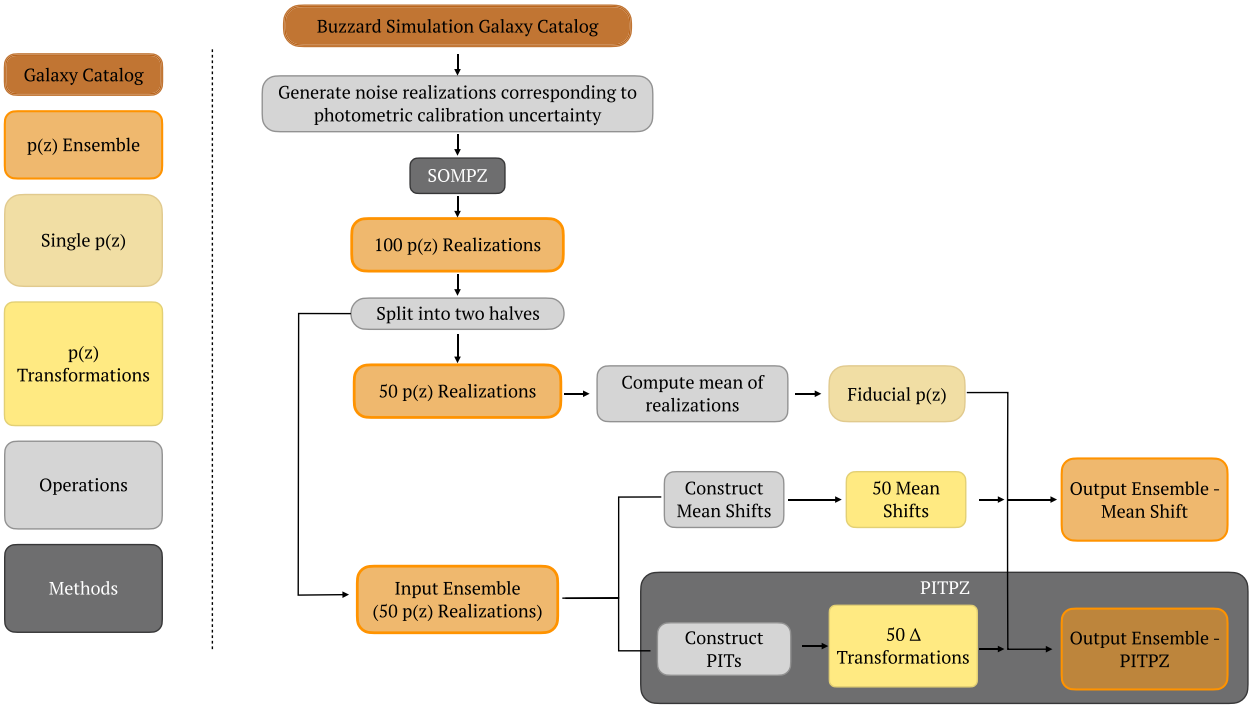

Figure 3. Illustration of the experimental design of the cosmological impact analysis in this work. The input ensemble is produced by running SOMPZ 50 times with varying deep-field photometric zero-points. The fiducial $n(z)$ is produced by taking the mean of an ensemble produced by running SOMPZ 50 times again with varying deep-field photometric zero-points. 'Output Ensemble – Mean Shift' is constructed by shifting the fiducial $n(z)$ by the mean value of each PIT; 'Output Ensemble – PITPZ' is constructed with the PITPZ method, i.e. by applying the full-shape delta transformations constructed from the input Ensemble to alter the fiducial $n(z)$.

versions of the output ensemble: one with PITPZ and one with mean shifts. The mean shift ensemble transfers only changes in the mean redshift between realizations in $n^{\text{in.}}(z)$ by contrast PITPZ transfers, the information for higher than mean-order moments according to the conservation rules shown in 4. In short, PITPZ transfers the full correlation structure of the realizations generated by the simulations. These two versions of the output ensemble should have transferred a different aspect or 'amount' of information from $n^{\text{in.}}(z)$ to $n^{\text{fid.}}(z)$. The difference between these two versions of the output ensemble will demonstrate the benefits of using PITPZ rather than mean shifts. To summarize, the three $n(z)$ ensembles discussed are:

I. (Input ensemble): First determine random zero-point offsets due to the uncertainty of the photometric calibration error by drawing from a Gaussian centred on zero with standard deviation set to the uncertainty of the deep-field photometric calibration in each band. Shift all deep-field magnitudes according to the result of this draw in each respective band for each deep field. Use these altered deep-field magnitudes as input to a run of the SOMPZ method on the Buzzard simulated galaxy catalogues. Select the first 50 realizations and construct delta transformations from them.

II. (Output ensemble – mean shift): $n(z)$ constructed by applying mean shifts (rather than full-shape delta transformations) to the fiducial $n(z)$.

III. (Output ensemble – PITPZ): $n(z)$ constructed by applying full-shape delta transformations to the fiducial $n(z)$. Following the notation of §2, this ensemble is labelled $n^{\text{out.}}(z)$.

These $n(z)$ are shown in Fig. 4. With these mock redshift distribution ensembles produced, we turn to assessing the difference between them for cosmology analysis. Our analysis consists in computing the uncertainty on the lensing amplitude associated with each ensemble, which relates closely to uncertainty on cosmological parameters.

We are interested in the following comparisons of the lensing amplitude distribution results yielded from these analyses:

(1) The difference between the lensing amplitude distributions associated with II and III illustrates the residual effect on redshift distributions of zero-point uncertainties beyond the first-order shift of the mean redshift. This is equivalent to illustrating the importance of using PITPZ, rather than simpler mean shifts, to incorporate this systematic uncertainty into redshift distributions.

(2) Because the input ensemble serves as a ground truth for the degree of variation due to photometric calibration uncertainty present in the simulations, any difference between the lensing amplitude distributions associated with I and III illustrates the residual effect on redshift distribution of zero-point uncertainties beyond what is corrected for with delta transformations produced with Buzzard. This is equivalent to illustrating the impact of higher than first-order moments due to the effect of photometric calibration uncertainty beyond what can be accounted for with the PITPZ method. In summary, any difference here illustrates shortcomings of the PITPZ method.

While the primary goal of this work is the illustration of the importance of using the delta transformation to preserve higher-order information than lower n-th order statistics in generating ensembles of probability distributions (i.e. comparison 1), this experimental design facilitates a secondary goal of illustrating the impact of our chosen source of uncertainty – photometric calibration error – on cosmology constraints. This secondary goal can play a role in informing future observing strategy decisions to collect the data necessary to reduce this uncertainty.

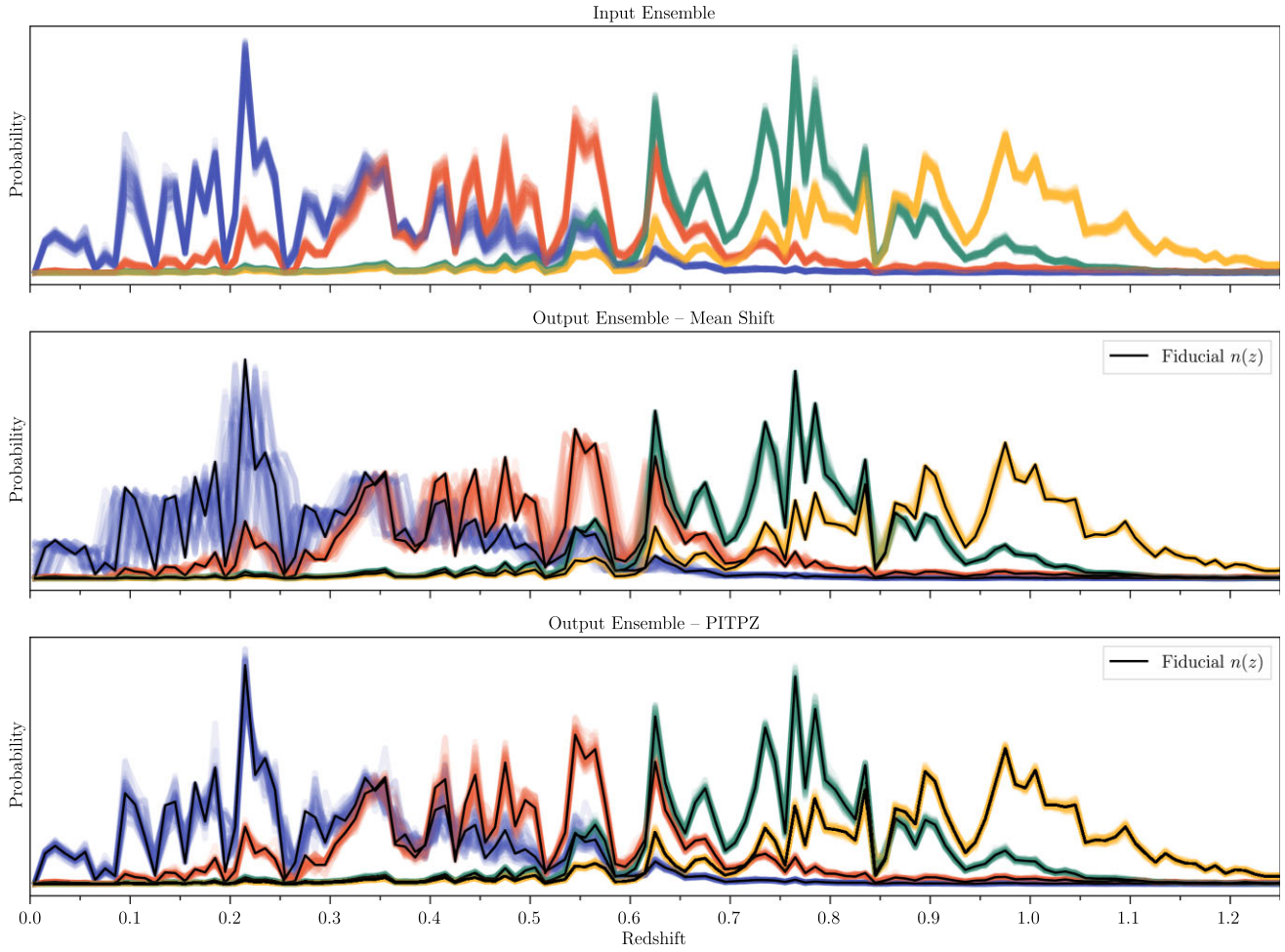

Figure 4. Illustration of the $n(z)$ distributions used in the simulated likelihood analysis in this work. The input ensemble is produced by running SOMPZ 50 times with varying deep-field photometric zero-points. 'Output ensemble – mean shift' is constructed by shifting the fiducial $n(z)$ by the mean value of each PIT; 'Output Ensemble – PITPZ' is constructed with the PITPZ method, i.e. by applying the full-shape delta transformations constructed from the input ensemble to alter the fiducial $n(z)$.

It remains to describe the relevant statistic that relates redshift distributions to constraints on the parameters of a given cosmological model. In practice, weak gravitational lensing involves inferring the matter distribution from coherent distortions in the measured shapes of galaxies. The presence of tangential alignment in galaxy shapes measured on the sky corresponds to the presence of a matter overdensity along the line of sight. The observed mean tangential shear γ_t associated with a separation angle θ on the sky can be expressed in terms of the lensing *convergence* that describes the amount of lensing

$$\langle\gamma_t\rangle(\theta) = \overline{\kappa}(<\theta) - \langle\kappa\rangle(\theta). \tag{13}$$

Convergence, in turn, can be written in terms of the total projected mass density Σ along a line-of-sight $\vec{\theta}$, and a *critical surface density* parameter which characterizes the lensing system

$$\kappa(\vec{\theta}) \equiv \frac{\Sigma(\vec{\theta})}{\Sigma_{\rm crit.}}. \tag{14}$$

This critical surface density due to lensing of a source at distance $D_{\rm s}$ from the observer by a lens (i.e. deflector) at distance $D_{\rm d}$ from the observer in a universe, where the distance between the source and the lens is $D_{\rm ds}$ is defined as follows under the assumption that the distances between source, lens, and observer are all much greater than the spatial extent of the lens (see e.g. Bartelmann & Schneider 2001)

$$\Sigma_{\rm crit.}^{-1} \equiv \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm d}D_{\rm ds}}. \tag{15}$$

This definition illustrates that uncertainty on galaxy distance corresponds directly to uncertainty on critical surface density, which in turn directly limits the degree to which projected mass density, and therefore cosmology can be constrained. For this reason, we choose critical surface density to test the impact of PITPZ on cosmology.

The shear $\gamma(\vec{\theta}, z_{\rm s})$ to which a particular source galaxy image is subject is a function of source galaxy redshift, so the mean shear observed along a line of sight $\vec{\theta}$ must be expressed with respect to the source galaxy redshift distributions (MacCrann et al. 2022; Amon et al. 2022)

$$\gamma(\vec{\boldsymbol{\theta}}) = \int {\rm d}z_{\rm s} n(z_{\rm s}) \gamma(\vec{\boldsymbol{\theta}}, z_{\rm s}). \tag{16}$$

Similarly, the total averaged lensing signal amplitude can be expressed in terms of the critical surface density integrated in the

Table 1. Summary statistics for each ensemble in the cosmological impact analysis of this work. We show the mean value and standard deviation for each of two statistics – mean redshift ($\bar{z}$) and lensing amplitude ($\overline{\Sigma^{-1}}_{\rm crit.}$) of an $n(z)$, for each of four tomographic bins; we also show the relative value of the lensing amplitude mean and standard deviation compared to the input ensemble to directly highlight the difference between PITPZ and mean shifts. We find that using our PITPZ method recovers, the uncertainty in $\bar{z}$ and $\overline{\Sigma^{-1}}_{\rm crit.}$ of the input ensemble (the ground truth in our experiment), using simpler mean shifts recovers only a portion of the total uncertainty in these parameters. The extent to which mean shifts underestimate uncertainty depends on the context of which underlying physical effect is being considered. In our case of photometric calibration uncertainty, we find that using mean shifts underestimates the uncertainty in lensing amplitude by as much as approximately 30 per cent in each of the bins. We choose $z = 0.25$ as the lens redshift for the lensing amplitudes shown in this table.

		$\bar{z}$		$\overline{\Sigma^{-1}}_{\rm crit.}\left[\frac{\rm M_\odot}{\rm Mpc}\right]^{-1}$		$\overline{\Sigma^{-1}}_{\rm crit.}\left[\frac{\rm M_\odot}{\rm Mpc}\right]^{-1}$	
Name	Symbol	μ	σ	μ	σ	$\mu/\mu_{\rm in.}$	$\sigma/\sigma_{\rm in.}$
				Bin 1			
Input ensemble	$p_{\rm in.}(z)$	0.310	1.044e-02	9.043e-17	5.856e-18	1.000	**1.000**
Output ensemble – mean shift	–	0.313	9.478e-03	9.181e-17	4.256e-18	1.015	**0.727**
Output ensemble – PITPZ	$p_{\rm out.}(z)$	0.312	1.041e-02	9.166e-17	5.919e-18	1.014	**1.011**
				Bin 2			
Input ensemble	$p_{\rm in.}(z)$	0.506	5.424e-03	2.007e-16	3.588e-18	1.000	**1.000**
Output ensemble – mean shift	–	0.506	5.385e-03	2.013e-16	2.592e-18	1.003	**0.722**
Output ensemble – PITPZ	$p_{\rm out.}(z)$	0.506	5.387e-03	2.013e-16	3.576e-18	1.003	**0.997**
				Bin 3			
Input ensemble	$p_{\rm in.}(z)$	0.745	1.998e-03	2.910e-16	6.246e-19	1.000	**1.000**
Output ensemble – mean shift	–	0.745	1.999e-03	2.909e-16	4.504e-19	1.000	**0.721**
Output ensemble – PITPZ	$p_{\rm out.}(z)$	0.745	1.964e-03	2.909e-16	6.174e-19	1.000	**0.989**
				Bin 4			
Input ensemble	$p_{\rm in.}(z)$	0.911	2.024e-03	3.185e-16	3.861e-19	1.000	**1.000**
Output ensemble – mean shift	–	0.912	2.023e-03	3.186e-16	3.169e-19	1.000	**0.821**
Output ensemble – PITPZ	$p_{\rm out.}(z)$	0.912	2.002e-03	3.186e-16	3.830e-19	1.000	**0.992**

same way as the total shear

$$\overline{\Sigma^{-1}}_{\rm crit.} = \frac{\int_{z_l}^{z_{\rm s,max.}} \frac{4\pi G}{c^2} \frac{D_{\rm d}(z_l) D_{\rm ds}(z_{\rm s}, z_l)}{D_{\rm s}(z_{\rm s})} n(z_{\rm s})\, {\rm d}z_{\rm s}}{\int_0^{z_{\rm s,max.}} n(z_{\rm s})\, {\rm d}z_{\rm s}}, \tag{17}$$

where the denominator is a normalization factor. Here $D_{\rm d}$, $D_{\rm s}$, and $D_{\rm ds}$ are determined by the lens and source redshifts $z_{\rm l}$ and $z_{\rm s}$. Equation (17) is a statistic to relate uncertainty on $n(z)$ to uncertainty on cosmology results. Note that this statistic is a weighted integral of $n(z)$, and effectively measures the probability density at redshift higher than the lens redshift $z_{\rm l}$, with higher redshift probability being weighted higher. As such, this statistic depends on higher than mean-order moments in $n(z)$. While mean redshift is the most important determining factor in the value of this statistic, at fixed mean redshift increasing variance, for example, will increase the probability at the highest redshifts. As a result, we expect this quantity to be more accurately evaluated from $n(z)$ constructed with PITPZ than from simpler mean shifts because PITPZ propagates uncertainty to higher-order moments (cf. Fig. 2).

We compute the distribution in $\overline{\Sigma^{-1}}_{\rm crit.}$ for each of our redshift distribution ensembles using the `lenstronomy` (Birrer & Amara 2018; Birrer et al. 2021) software, and report the resulting values in Table 1. Since the uncertainty on constraints on cosmology from a cosmic shear analysis such as that conducted with the Dark Energy Survey Year 3 data set (Amon et al. 2022; Secco et al. 2022) is proportional to the uncertainty on lensing amplitude, the distribution of possible lensing amplitudes functions as a proxy for the resulting uncertainty on cosmological parameters. In addition to the statistic defined in equation (17), we compute the cosmic shear two-point correlation function $\xi_{+/-}$ with each $n(z)$ in our input and output ensembles using the `CCL` package of Chisari et al. (2019) (for details on cosmic shear, see e.g. Amon et al. 2022; Secco et al. 2022). We integrate over this cosmic shear data vector ξ, and show results relating input and output values of this quantity in Fig. 6.

6 RESULTS

Our primary results are shown in Figs 5 and 6 and Table 1. Fig. 5 illustrates that PITPZ propagates the relative strength of the lensing signal amplitude, which depends on higher-order moments of $n(z)$ across all scales. By contrast, the loss of higher than mean-order moment information associated with mean shifts causes deviations from linearity in the relationship between lensing amplitude in the input ensemble and output ensemble realizations. As a result, the overall scatter in $\overline{\Sigma^{-1,\rm out.}}_{\rm crit.}$ is smaller in the case of using mean shifts. As shown in Table 1, the scatter in the output ensemble lensing amplitude using the full PITPZ method matches the true scatter from the input ensemble to within 1 per cent for all tomographic bins. By contrast, using mean shifts underestimate this scatter by 27, 28, 28, and 18 per cent in the four tomographic bins, respectively ($z_{\rm lens} = 0.25$). We can summarize the imperfections of the mean shift method relative to PITPZ in terms of two effects visually apparent in Fig. 5: first, the slope of the relationship between input and output lensing amplitude deviates from the value of unity, leading to the bulk of the loss of scatter in lensing amplitude. Second, however, the mean shift method introduces significant scatter about the linear relationship, which has an overall additive effect to the scatter in the lensing amplitude. In this sense, our estimate of the degree to which mean shifts underestimate the uncertainty in lensing amplitude are a lower bound because they include this additive effect. Our result that using mean shifts on $n(z)$ underestimates uncertainty applies not only to lensing signal amplitude, but to any quantity that is a weighted integral of $n(z)$, as any such quantity will depend on higher-order moments in $n(z)$. We finally highlight that since $n(z)$, unlike $\overline{\Sigma^{-1}}_{\rm crit.}$, is cosmology independent, our method does not depend on an assumed cosmology. By contrast, an attempt to propagate uncertainty by way of mean shifts on lensing signal amplitude itself would require an assumed cosmology to determine the $D_{\rm ds}$ factor present in the definition of $\overline{\Sigma^{-1}}_{\rm crit.}$. This is an additional advantage of operating directly on $n(z)$ with PITPZ. We emphasize that although the

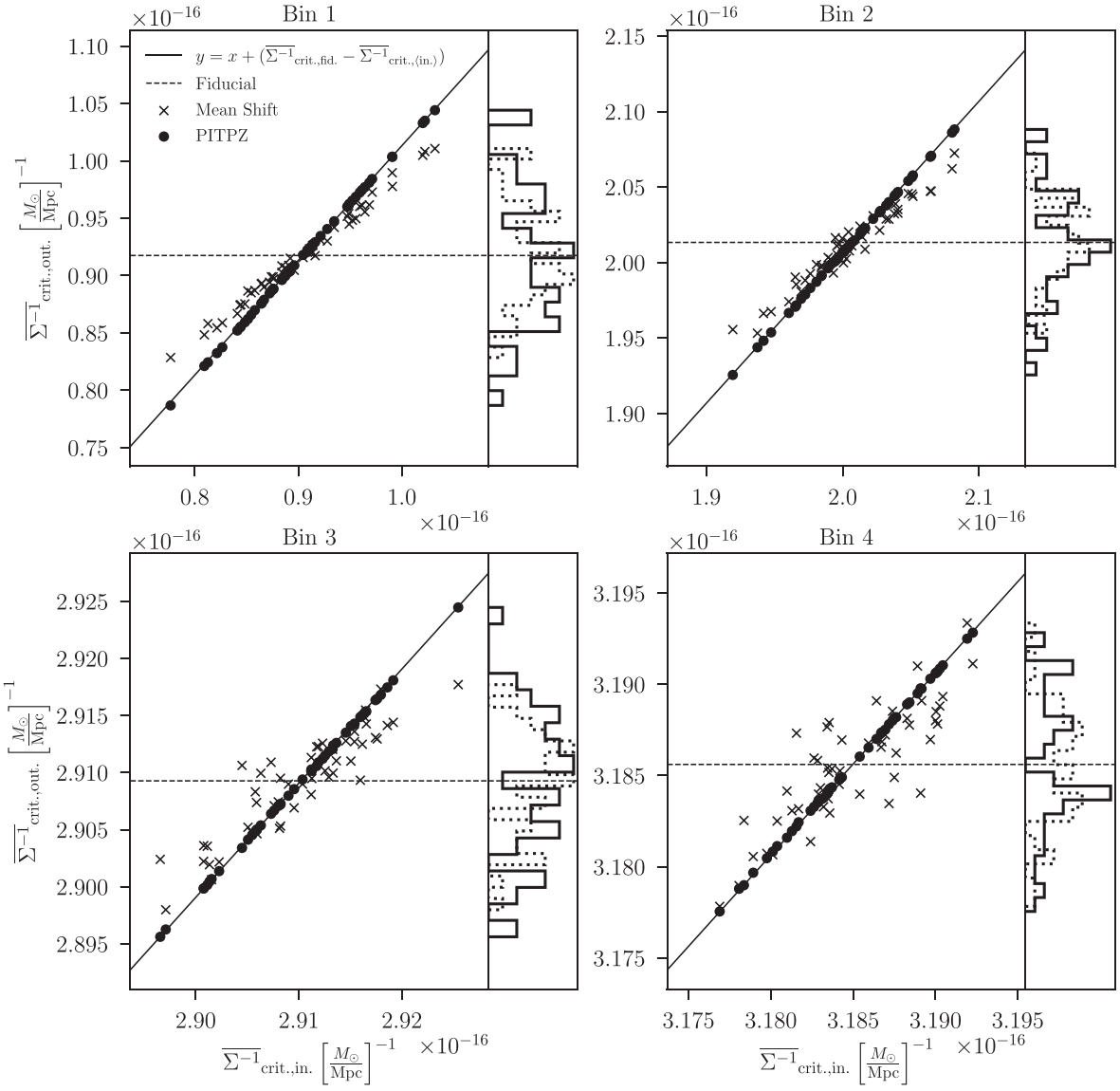

Figure 5. Relationship between lensing signal amplitude in the input ensemble and the output ensemble realizations using PITPZ or mean shifts for the experiment described in Section 5 with $z_{\rm lens} = 0.25$. We find that PITPZ more reliably transfers lensing amplitude information than mean shifts. This is explained by the fact that the lensing amplitude is a weighted integral of $n(z)$, so higher-order moments of $n(z)$, which are conserved by PITPZ but not conserved by mean shifts will cause the mean shift to underestimate the scatter in lensing amplitude. Histograms on the side panels illustrate the distribution of lensing signal amplitude for the output ensemble, where the solid line corresponds to the output ensemble produced with PITPZ, and the dotted line corresponds to that produced with mean shifts.

qualitative results shown are applicable in general, the quantitative difference between PITPZ and mean shifts is specific to the source of uncertainty under consideration, and the redshift distributions of the source and lens galaxy samples observed. Larger values of lens redshift eliminate the impact of differences between realizations in the input ensemble at redshift values less than $z_{\rm lens}$. As one scales lens redshift up from zero, the degree to which the effect shown in Fig. 5 varies depends on how much relative variation in $n(z)$ is below and above the value of the lens redshift. As a result, the degree to which these results change for a higher choice of lens redshift is again specific to the source of uncertainty and the redshift distribution of the galaxy survey in question. Fig. 6 shows the relationship between input and output values of the cosmic shear data vector ξ_+. In particular, for each $n(z)$ realization in the input ensemble, we compute the galaxy shape two-point correlation function $\xi_+(\theta)$ (given the assumed cosmology defined in Section 1) and the integral $\int {\rm d}\theta\, \theta \xi_+$. We likewise compute this value for each realization of the output ensembles produced by the mean shift and PITPZ methods, respectively. Fig. 6 shows that PITPZ again preserves a linear relationship between input and output realizations, whereas mean shifts do not.

7 CONCLUSION

We have presented a method for transferring variations between realizations of PDFs in one ensemble onto another PDF (or ensemble of PDFs). Our method, dubbed PITPZ, may have general applications for propagating uncertainties on posterior probability functions. In

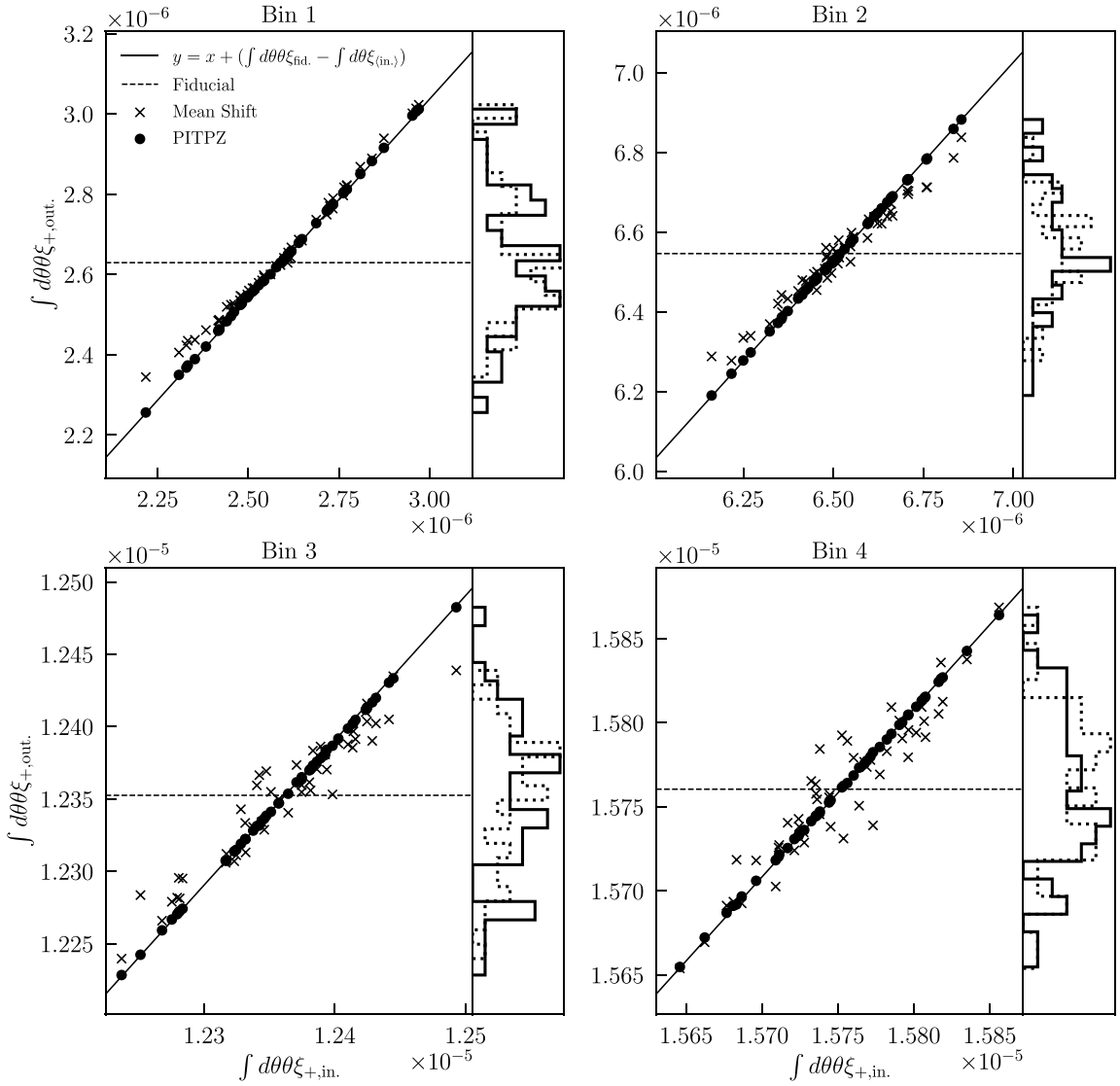

Figure 6. Relationship between the cosmic shear signal amplitude as inferred from input ensemble $n(z)$ realizations to the cosmic shear signal amplitude as inferred from output ensemble $n(z)$ realizations. The output ensembles are produced with PITPZ or mean shifts with the experiment described in Section 5. Axis values are integrals over the full cosmic shear data vector ξ_+. As in Fig. 5, we find that PITPZ more reliably transfers information than mean shifts. Histograms on the side panels illustrate the distribution of signal amplitude for the output ensemble, where the solid line corresponds to the output ensemble produced with PITPZ, and the dotted line corresponds to that produced with mean shifts.

addition to providing a treatment of the algorithm, we derive analytic estimates of the conservation rules for the first three moments (mean, variance, and skewness) of the PDFs used.

We illustrate the use of this method with an experiment in the context of the weak gravitational lensing survey redshift calibration problem, for which the redshifts for large numbers of galaxies are estimated. We find that our method is an improvement over simpler mean shifts of PDFs for transferring higher-order information. We show that this higher-order information is critically important in the context of redshift calibration by propagating redshift distributions to total gravitational lensing signal amplitude, which relates directly to the cosmological constraints of lensing surveys. In summary, we find for our fiducial test case involving photometric zero-point uncertainty for a DES Y3-like survey ($z_{\rm lens} = 0.25$) that our method recovers the true uncertainty on lensing amplitude to within 1 per cent, in contrast to an underestimate of as much as 30 per cent when using mean shifts. The difference between PITPZ and mean shift on lensing amplitude reflects the importance of this method for cosmology analyses requiring redshift distributions.

We confirm that the numerical errors associated with our software implementation of our method fall well-within the LSST DESC Year 10 uncertainty targets for redshift calibration. By contrast, using simple mean shifts exceeds this uncertainty target in the mean redshift in our test case. While in our test case, the error on the variance introduced by mean shifts is still so small as to fall within the LSST DESC Y10 uncertainty target in the scatter in redshift, it is the accumulated effect for all higher moments, and when also accounting for multiple independent sources of redshift uncertainty that propagates directly to uncertainty on cosmological parameters, which may justify the additional complexity of PITPZ relative to

mean shifts. Based on these results, we conclude that future galaxy lensing surveys should consider using PITPZ for propagating redshift uncertainties.

Development of the PITPZ method has been motivated by the significant and consequential challenges of the redshift calibration problem to accomplish the stated goals of upcoming galaxy imaging surveys like the Legacy Survey of Space and Time (LSST Science Collaboration et al. 2009; The LSST Dark Energy Science Collaboration et al. 2018; Ivezić et al. 2019). In this context, improvements in our ability to measure redshift distributions from photometric galaxy samples are a prerequisite to fulfill the promise of the next generation of weak lensing experiments and of the investments made to this end. As we have discussed, PITPZ will facilitate more accurate uncertainty characterization of these measurements by enabling a transfer of uncertainties from simulations where certain observational effects can be scaled at-will to the measurements on data. Similarly, uncertainties measured in data products can be likewise transferred to measurements in simulations, which will facilitate realistic end-to-end analyses in simulations for cosmology pipeline validation. Noting the characterization of the redshift calibration problem as being within a category for which 'promising ideas exist, but more exploration is needed to determine which will work and how exactly to use them at the level of precision needed for future surveys' (Mandelbaum 2018), we highlight that although this work has focused on weak lensing source galaxies, our method has important implications for lens redshift calibration. Given that lens redshift distributions appear as a quadratic term in the galaxy clustering signal by way of the radial selection function of lens galaxies for a given source galaxy tomographic bin (i.e. the 'galaxy clustering kernel'), the galaxy clustering signal is especially sensitive to the width of the lens $n(z)$ (see e.g. Pandey et al. (2021), Porredon et al. (2021), Rodríguez-Monroy et al. (2022)). PITPZ, as a first solution to propagating $n(z)$ uncertainty for the width of $n(z)$ (and other higher than mean-order moments) may prove an essential component to calibrating lens redshift distributions within uncertainty requirements for upcoming galaxy clustering analyses. Because PITPZ is part of an effort to express redshift distribution uncertainty with sufficient complexity to meet future uncertainty goals, a natural question to ask is whether the form of redshift distribution uncertainty relates to degeneracies between redshift distribution uncertainty, and other nuisance parameters in weak lensing cosmology analyses such as intrinsic alignment model parameters. We leave this question to future work.

PITPZ is a flexible solution with numerous potential applications in the context of weak lensing redshift calibration to address the clear needs for higher precision in scheduled next-generation galaxy surveys. More broadly, recognizing the trend within astrophysics and cosmology toward the use of Bayesian statistical methods that produce full posterior probability distributions for model parameters of interest, PITPZ can serve a useful role of sophisticated propagation of uncertainties in a wide variety of subfields of astronomy.

ACKNOWLEDGEMENTS

We thank Nickolas Kokron for helpful discussion. This work was supported by the Department of Energy, Laboratory Directed Research and Development programme at SLAC National Accelerator Laboratory, under contract DE-AC02-76SF00515. This work was supported by the Bavaria California Technology Center (BaCaTeC). JM acknowledges funding from the Diversifying Academia Recruiting Excellence (DARE) and the Robert H. Siemann Fellowship Programmes at Stanford University. JM thanks the LSSTC Data Science Fellowship Program, which is funded by LSSTC, NSF Cybertraining Grant #1829740, the Brinson Foundation, and the Moore Foundation; his participation in the program has benefited this work. IH acknowledges support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 849169).

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute for Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NFS's NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA programme of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia Midas (INCT) do e-Universo (CNPq grant 465376/2014-2).

DATA AVAILABILITY

The DES Buzzard simulated data redshift distributions used for this work are available upon request.

REFERENCES

Abbott T. M. C. et al., 2018, Phys. Rev. D, 98, 043526
Abbott T. M. C. et al., 2022, Phys. Rev. D, 105, 023520
Alarcon A., Sánchez C., Bernstein G. M., Gaztañaga E., 2020, MNRAS, 498, 2614
Amon A. et al., 2022, Phys. Rev. D, 105, 023514
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Benjamin J. et al., 2013, MNRAS, 431, 1547
Birrer S., Amara A., 2018, Phys. Dark Universe, 22, 189
Birrer S. et al., 2021, J. Open Source Softw., 6, 3283
Bonnett C. et al., 2016, Phys. Rev. D, 94, 042005
Bordoloi R., Lilly S. J., Amara A., 2010, MNRAS, 406, 881
Buchs R. et al., 2019, MNRAS, 489, 820
Cabayol L. et al., 2022, preprint (arXiv:2209.10161)
Chisari N. E. et al., 2019, ApJS, 242, 2
Cordero J. P. et al., 2022, MNRAS, 511, 2170
Cunha C. E., Huterer D., Busha M. T., Wechsler R. H., 2012, MNRAS, 423, 909
DeRose J. et al., 2019, preprint (arXiv:1901.02401)
Dodge Y., Cox D., Institute I. S., Commenges D., 2006, The Oxford Dictionary of Statistical Terms. Oxford Univ. Press, Oxford. Available at: https://books.google.com/books?id=OnjBgpuhWcC
Euclid Collaboration et al., 2020, A&A, 642, A192
Everett S. et al., 2022, ApJS, 258, 15
Freeman P. E., Izbicki R., Lee A. B., 2017, MNRAS, 468, 4556
Gatti M. et al., 2022, MNRAS, 510, 1223
Gatti M. et al., 2021, MNRAS, 504, 4312
Green J. et al., 2011, preprint (arXiv:1108.1374)
Gruen D., Brimioulle F., 2017, MNRAS, 468, 769
Hadzhiyska B., Alonso D., Nicola A., Slosar A., 2020, J. Cosmology Astropart. Phys., 2020, 056
Hartley W. G. et al., 2022, MNRAS, 509, 3547
Hasan I. S., Schmidt S. J., Schneider M. D., Tyson J. A., 2022, MNRAS, 511, 1029
Heymans C. et al., 2021, A&A, 646, A140
Hikage C. et al., 2019, PASJ, 71, 43
Hildebrandt H. et al., 2012, MNRAS, 421, 2355
Hildebrandt H. et al., 2020, A&A, 633, A69
Hildebrandt H. et al., 2021, A&A, 647, A124
Hildebrandt H. et al., 2017, MNRAS, 465, 1454
Hoyle B. et al., 2018, MNRAS, 478, 592
Huterer D., Cunha C. E., Fang W., 2013, MNRAS, 432, 2945
Huterer D., Takada M., Bernstein G., Jain B., 2006, MNRAS, 366, 101
Ivezić Ž. et al., 2019, ApJ, 873, 111
Jee M. J., Tyson J. A., Schneider M. D., Wittman D., Schmidt S., Hilbert S., 2013, ApJ, 765, 74
Joudaki S. et al., 2020, A&A, 638, L1
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Leistedt B., Hogg D. W., Wechsler R. H., DeRose J., 2019, ApJ, 881, 80
Lima M., Cunha C. E., Oyaizu H., Frieman J., Lin H., Sheldon E. S., 2008, MNRAS, 390, 118
LSST Dark Energy Science Collaboration, 2012, preprint (arXiv:1211.0310)
LSST Science Collaboration et al., 2009, preprint (arXiv:0912.0201)
MacCrann N. et al., 2022, MNRAS, 509, 3371
Malz A. I., 2021, PRD, 103, 083502
Malz A. I., Hogg D. W., 2022, ApJ, 928, 127
Malz A. I., Marshall P. J., DeRose J., Graham M. L., Schmidt S. J., Wechsler R., (LSST Dark Energy Science Collaboration, 2018, AJ, 156, 35
Mandelbaum R., 2018, ARA&A, 56, 393
Myles J. et al., 2021, MNRAS, 505, 4249
Newman J. A., Gruen D., 2022, ARA&A, 60, 363
Pandey S. et al., 2021, Phys. Rev. D, 106, 043520
Planck Collaboration VI, 2020, A&A, 641, A6
Polsterer K. L., D'Isanto A., Gieseke F., 2016, MNRAS, preprint (arXiv:1608.08016)
Porredon A. et al., 2021, Phys. Rev. D, 106, 103530
Rodríguez-Monroy M. et al., 2022, MNRAS, 511, 2665
Samuroff S., Troxel M. A., Bridle S. L., Zuntz J., MacCrann N., Krause E., Eifler T., Kirk D., 2017, MNRAS, 465, L20
Sánchez C., Bernstein G. M., 2019, MNRAS, 483, 2801
Sánchez C. et al., 2022, Phys. Rev. D, 105, 083529
Schmidt S. J. et al., 2020, MNRAS, 499, 1587
Secco L. F. et al., 2022, Phys. Rev. D, 105, 023515
Shuntov M. et al., 2020, A&A, 636, A90
Stölzner B., Joachimi B., Korn A., Hildebrandt H., Wright A. H., 2021, A&A, 650, A148
Tanaka M. et al., 2018, PASJ, 70, S9
Tannenbaum P., 2010, Excursions in Modern Mathematics, Pearson Custom Mathematics. Prentice Hall. Available at: https://books.google.com/books?id=jHAbOgAACAAJ
Tessore N., Harrison I., 2020, The Open J. Astrophys., 3, 6
The LSST Dark Energy Science Collaboration et al., 2018, preprint (arXiv:1809.01669)
Wright A. H., Hildebrandt H., van den Busch J. L., Heymans C., 2020a, A&A, 637, A100
Wright A. H., Hildebrandt H., van den Busch J. L., Heymans C., Joachimi B., Kannawadi A., Kuijken K., 2020b, A&A, 640, L14
Zhang T., Rau M. M., Mandelbaum R., Li X., Moews B., 2022, MNRAS, 518, 709

APPENDIX A: HIGHER ORDER MOMENTS OF REDSHIFT DISTRIBUTION

In the following, we provide more complete algebra deriving the conserved quantities associated with the variance and skewness of the distributions used in our work. We use the same convention as Section 2, where 'in.' represents a $p(z)$ in the input ensemble that contains variation we wish to map to another $p(z)$, $\langle \text{in.} \rangle$ denotes the mean of these realizations, T represents a delta transformation, and 'out.' represents the output realization resulting from the PITPZ algorithm. The index j runs over the number of samples used to represent smooth $p(z)$, as described in Section 3.

A1 Variance

We now turn to representing the variance of the output ensemble in terms of the variances of the inputs.

$$\begin{aligned}\sigma^2_{\text{in.}} &= \frac{1}{n}\sum_j^n \left[\left(z_j^{\text{in.}} - \bar{z}^{\text{in.}}\right)^2\right] \\ &= \frac{1}{n}\sum_j^n \left[z_j^{\text{in.}^2} - 2z_j^{\text{in.}}\bar{z}^{\text{in.}} + \bar{z}^{\text{in.}^2}\right].\end{aligned} \tag{A1}$$

$$
\begin{aligned}
\sigma_T^2 &= \frac{1}{n}\sum_j^n \left[(\Delta_j - \bar{\Delta})^2\right] \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right) - \overline{\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)}\right]^2 \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right) - \overline{\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)}\right]\left[\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right) - \overline{\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)}\right] \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)^2 - 2\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)\overline{\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)} + \overline{\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)}^2\right] \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\text{in.}}\right)^2 - 2z_j^{\text{in.}} z_j^{\langle\text{in.}\rangle} + \left(z_j^{\langle\text{in.}\rangle}\right)^2 - 2z_j^{\text{in.}}\overline{\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)} + 2z_j^{\langle\text{in.}\rangle}\overline{\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)} + \overline{\left(z_j^{\text{in.}} - z_j^{\langle\text{in.}\rangle}\right)}^2\right] \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\text{in.}}\right)^2 - 2z_j^{\text{in.}} z_j^{\langle\text{in.}\rangle} + \left(z_j^{\langle\text{in.}\rangle}\right)^2 - 2z_j^{\text{in.}}(\bar{z}^{\text{in.}} - \bar{z}^{\langle\text{in.}\rangle}) + 2z_j^{\langle\text{in.}\rangle}(\bar{z}^{\text{in.}} - \bar{z}^{\langle\text{in.}\rangle}) + (\bar{z}^{\text{in.}} - \bar{z}^{\langle\text{in.}\rangle})^2\right] \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\text{in.}}\right)^2 - 2z_j^{\text{in.}} z_j^{\langle\text{in.}\rangle} + \left(z_j^{\langle\text{in.}\rangle}\right)^2 - 2z_j^{\text{in.}}\bar{z}^{\text{in.}} + 2z_j^{\text{in.}}\bar{z}^{\langle\text{in.}\rangle} + 2z_j^{\langle\text{in.}\rangle}\bar{z}^{\text{in.}} - 2z_j^{\langle\text{in.}\rangle}\bar{z}^{\langle\text{in.}\rangle} + (\bar{z}^{\text{in.}})^2 - 2\bar{z}^{\text{in.}}\bar{z}^{\langle\text{in.}\rangle} + (\bar{z}^{\langle\text{in.}\rangle})^2\right] \\
&= \frac{1}{n}\sum_j^n \left[\left(\left(z_j^{\text{in.}}\right)^2 - 2z_j^{\text{in.}}\bar{z}^{\text{in.}} + (\bar{z}^{\text{in.}})^2\right) + \left(\left(z_j^{\langle\text{in.}\rangle}\right)^2 - 2z_j^{\langle\text{in.}\rangle}\bar{z}^{\langle\text{in.}\rangle} + (\bar{z}^{\langle\text{in.}\rangle})^2\right) - 2z_j^{\text{in.}} z_j^{\langle\text{in.}\rangle} + 2z_j^{\text{in.}}\bar{z}^{\langle\text{in.}\rangle} + 2z_j^{\langle\text{in.}\rangle}\bar{z}^{\text{in.}} - 2\bar{z}^{\text{in.}}\bar{z}^{\langle\text{in.}\rangle}\right] \\
&= \sigma_{\text{in.}}^2 + \sigma_{\langle\text{in.}\rangle}^2 + \frac{1}{n}\sum_j^n \left[-2z_j^{\text{in.}} z_j^{\langle\text{in.}\rangle} + 2z_j^{\text{in.}}\bar{z}^{\langle\text{in.}\rangle} + 2z_j^{\langle\text{in.}\rangle}\bar{z}^{\text{in.}} - 2\bar{z}^{\text{in.}}\bar{z}^{\langle\text{in.}\rangle}\right] \\
&= \sigma_{\text{in.}}^2 + \sigma_{\langle\text{in.}\rangle}^2 - \frac{2}{n}\sum_j^n \left[z_j^{\text{in.}} z_j^{\langle\text{in.}\rangle} - z_j^{\text{in.}}\bar{z}^{\langle\text{in.}\rangle} - z_j^{\langle\text{in.}\rangle}\bar{z}^{\text{in.}} + \bar{z}^{\text{in.}}\bar{z}^{\langle\text{in.}\rangle}\right] \\
&= \sigma_{\text{in.}}^2 + \sigma_{\langle\text{in.}\rangle}^2 - \frac{2}{n}\sum_j^n \left[\left(z_j^{\text{in.}} - \bar{z}_j^{\text{in.}}\right)\left(z_j^{\langle\text{in.}\rangle} - \bar{z}^{\langle\text{in.}\rangle}\right)\right] \\
&= \sigma_{\text{in.}}^2 + \sigma_{\langle\text{in.}\rangle}^2 - 2\,\text{Cov}\left[z_j^{\text{in.}}, z_j^{\langle\text{in.}\rangle}\right].
\end{aligned}
\tag{A2}
$$

In summary, we find that the variance of a delta transformation can be written as the sum of the variance of the input $p(z)$ used in its construction, the variance of the mean of the realizations of the input $p(z)$ ensemble used in its construction, and the covariance between these

elements. This covariance is computed directly from the ordered, evenly-spaced samples of the relevant PDFs.

$$\begin{aligned}
\sigma^2_{\rm out.} &= \frac{1}{n}\sum_j^n \left[\left(z_j^{\rm out.} - \bar{z}_j^{\rm out.}\right)^2\right] \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\rm fid.} + \Delta_j\right) - \overline{\left(z_j^{\rm fid.} + \Delta_j\right)}\right]^2 \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\rm fid.} + \Delta_j\right) - \overline{\left(z_j^{\rm fid.} + \Delta_j\right)}\right]\left[\left(z_j^{\rm fid.} + \Delta_j\right) - \overline{\left(z_j^{\rm fid.} + \Delta_j\right)}\right] \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\rm fid.} + \Delta_j\right)^2 - 2\left(z_j^{\rm fid.} + \Delta_j\right)\overline{\left(z_j^{\rm fid.} + \Delta_j\right)} + \overline{\left(z_j^{\rm fid.} + \Delta_j\right)}^2\right] \\
&= \frac{1}{n}\sum_j^n \left[z_j^{\rm fid.^2} + \Delta_j^2 + 2z_j^{\rm fid.}\Delta_j - 2z_j^{\rm fid.}(\bar{z}^{\rm fid.} + \bar{\Delta}) - 2\Delta_j(\bar{z}^{\rm fid.} + \bar{\Delta}) + (\bar{z}^{\rm fid.} + \bar{\Delta})^2\right] \\
&= \frac{1}{n}\sum_j^n \left[z_j^{\rm fid.^2} + \Delta_j^2 + 2z_j^{\rm fid.}\Delta_j - 2\bar{z}^{\rm fid.}z_j^{\rm fid.} - 2\bar{\Delta}z_j^{\rm fid.} - 2\Delta_j\bar{z}^{\rm fid.} - 2\Delta_j\bar{\Delta} + (\bar{z}^{\rm fid.} + \bar{\Delta})^2\right] \\
&= \frac{1}{n}\sum_j^n \left[z_j^{\rm fid.^2} + \Delta_j^2 + 2z_j^{\rm fid.}\Delta_j - 2\bar{z}^{\rm fid.}z_j^{\rm fid.} - 2\bar{\Delta}z_j^{\rm fid.} - 2\Delta_j\bar{z}^{\rm fid.} - 2\Delta_j\bar{\Delta} + \bar{z}^{\rm fid.^2} + 2\bar{z}^{\rm fid.}\bar{\Delta} + \bar{\Delta}^2\right]. \\
&= \frac{1}{n}\sum_j^n \left[\left(z_j^{\rm fid.^2} - 2\bar{z}^{\rm fid.}z_j^{\rm fid.} + \bar{z}^{\rm fid.^2}\right) + \left(\Delta_j^2 - 2\Delta_j\bar{\Delta} + \bar{\Delta}^2\right) + 2z_j^{\rm fid.}\Delta_j - 2\bar{\Delta}z_j^{\rm fid.} - 2\Delta_j\bar{z}^{\rm fid.} + 2\bar{z}^{\rm fid.}\bar{\Delta}\right] \\
&= \sigma^2_{\rm fid.} + \sigma_T^2 + \frac{2}{n}\sum_j^n \left[z_j^{\rm fid.}\Delta_j - \bar{\Delta}z_j^{\rm fid.} - \Delta_j\bar{z}^{\rm fid.} + \bar{z}^{\rm fid.}\bar{\Delta}\right] \\
&= \sigma^2_{\rm fid.} + \sigma_T^2 + 2\,{\rm Cov}\left[z_j^{\rm fid.}, \Delta_j\right]. \qquad \text{(A3)}
\end{aligned}$$

Using equation (A2) to replace σ_T^2 with quantities from the input ensemble, this yields our final expression for the variance of a realization in the output ensemble

$$\sigma^2_{\rm out.} = \sigma^2_{\rm fid.} + \sigma^2_{\rm in.} + \sigma^2_{\langle{\rm in.}\rangle} - 2\,{\rm Cov}\left[z_j^{\rm in.}, z_j^{\langle{\rm in.}\rangle}\right] + 2\,{\rm Cov}\left[z_j^{\rm fid.}, \Delta_j\right]. \qquad \text{(A4)}$$

Alternatively we can expand Δ_j to yield

$$\sigma^2_{\rm out.} = \sigma^2_{\rm fid.} + \sigma^2_{\rm in.} + \sigma^2_{\langle{\rm in.}\rangle} - 2\,{\rm Cov}\left[z_j^{\rm in.}, z_j^{\langle{\rm in.}\rangle}\right] + 2\,{\rm Cov}\left[z_j^{\rm fid.}, z_j^{\rm in.}\right] - 2\,{\rm Cov}\left[z_j^{\rm fid.}, z_j^{\langle{\rm in.}\rangle}\right]. \qquad \text{(A5)}$$

A2 Skewness

We now turn to developing an expression for the skewness of a realization of the output ensemble in terms of moments of the input ensemble. We use the standardized moments, which are normalized to be scale invariant. For a random variable X with probability distribution P with mean μ, the standardized moment of degree k is defined as the ratio of the moment of degree k, and the standard deviation σ,

$$\widetilde{\mu}_k \equiv \frac{\mu_k}{\sigma^k} = \frac{{\rm E}[(X-\mu)^k]}{({\rm E}[(X-\mu)^2])^{k/2}}. \qquad \text{(A6)}$$

The standardized moment of degree k of a realization of the output ensemble can be written as follows. Using $\sigma_{\rm out.}$ to represent the standard deviation of a given realization (see equation 10),

$$\begin{aligned}
\widetilde{\mu}_k^{\rm out.} &= \frac{1}{\sigma^k_{\rm out.}}{\rm E}\left[\left(z_j^{\rm out.} - \overline{z}_j^{\rm out.}\right)^k\right] \\
&= \frac{1}{\sigma^k_{\rm out.}}\frac{1}{n}\sum_j^n \left[\left(z_j^{\rm out.} - \overline{z}_j^{\rm out.}\right)^k\right]. \qquad \text{(A7)}
\end{aligned}$$

Expanding $z_j^{\rm out.} = z_j^{\rm fid.} + \Delta_j = z_j^{\rm fid.} + z_j^{\rm in.} - z_j^{\langle{\rm in.}\rangle}$ yields

$$\begin{aligned}
\widetilde{\mu}_k^{\rm out.} &= \frac{1}{\sigma^k_{\rm out.}}\frac{1}{n}\sum_j^n \left[\left(z_j^{\rm fid.} + z_j^{\rm in.} - z_j^{\langle{\rm in.}\rangle}\right) - \overline{\left(z_j^{\rm fid.} + z_j^{\rm in.} - z_j^{\langle{\rm in.}\rangle}\right)}\right]^k \\
&= \frac{1}{\sigma^k_{\rm out.}}\frac{1}{n}\sum_j^n \left[\left(z_j^{\rm fid.} + z_j^{\rm in.} - z_j^{\langle{\rm in.}\rangle} - \overline{z}_j^{\rm fid.} - \overline{z}_j^{\rm in.} + \overline{z}_j^{\langle{\rm in.}\rangle}\right)^k\right] \\
&= \frac{1}{\sigma^k_{\rm out.}}\frac{1}{n}\sum_j^n \left[\left(\left(z_j^{\rm fid.} - \overline{z}_j^{\rm fid.}\right) + \left(z_j^{\rm in.} - \overline{z}_j^{\rm in.}\right) - \left(z_j^{\langle{\rm in.}\rangle} - \overline{z}_j^{\langle{\rm in.}\rangle}\right)\right)^k\right]. \qquad \text{(A8)}
\end{aligned}$$

The standardized skewness is thus

$$\begin{aligned}\widetilde{\mu}_3^{\text{out.}} = \frac{1}{\sigma_{\text{out.}}^3}\frac{1}{n}\sum_j^n \Big[&\left(z_j^{\text{fid.}} - \bar{z}_j^{\text{fid.}}\right)^3 + 3\left(z_j^{\text{fid.}} - \bar{z}_j^{\text{fid.}}\right)^2\left(z_j^{\text{in.}} - \bar{z}_j^{\text{in.}}\right) - 3\left(z_j^{\text{fid.}} - \bar{z}_j^{\text{fid.}}\right)^2\left(z_j^{\langle\text{in.}\rangle} - \bar{z}_j^{\langle\text{in.}\rangle}\right) + 3\left(z_j^{\text{fid.}} - \bar{z}_j^{\text{fid.}}\right)\left(z_j^{\text{in.}} - \bar{z}_j^{\text{in.}}\right)^2 \\ &- 6\left(z_j^{\text{fid.}} - \bar{z}_j^{\text{fid.}}\right)\left(z_j^{\text{in.}} - \bar{z}_j^{\text{in.}}\right)\left(z_j^{\langle\text{in.}\rangle} - \bar{z}_j^{\langle\text{in.}\rangle}\right) + 3\left(z_j^{\text{fid.}} - \bar{z}_j^{\text{fid.}}\right)\left(z_j^{\langle\text{in.}\rangle} - \bar{z}_j^{\langle\text{in.}\rangle}\right)^2 + \left(z_j^{\text{in.}} - \bar{z}_j^{\text{in.}}\right)^3 \\ &- 3\left(z_j^{\text{in.}} - \bar{z}_j^{\text{in.}}\right)^2\left(z_j^{\langle\text{in.}\rangle} - \bar{z}_j^{\langle\text{in.}\rangle}\right) + 3\left(z_j^{\text{in.}} - \bar{z}_j^{\text{in.}}\right)\left(z_j^{\langle\text{in.}\rangle} - \bar{z}_j^{\langle\text{in.}\rangle}\right)^2 - \left(z_j^{\langle\text{in.}\rangle} - \bar{z}_j^{\langle\text{in.}\rangle}\right)^3\Big].\end{aligned} \tag{A9}$$

Compare to the individual expressions for the k^{th} moment of each ingredient in the recipe for constructing each realization in the output ensemble,

$$\widetilde{\mu}_k^{\text{fid.}} = \frac{1}{\sigma_{\text{fid.}}^k}\frac{1}{n}\sum_j^n \left[\left(z_j^{\text{fid.}} - \bar{z}_j^{\text{fid.}}\right)^k\right], \tag{A10}$$

$$\widetilde{\mu}_k^{\text{in.}} = \frac{1}{\sigma_{\text{in.}}^k}\frac{1}{n}\sum_j^n \left[\left(z_j^{\text{in.}} - \bar{z}_j^{\text{in.}}\right)^k\right], \tag{A11}$$

$$\widetilde{\mu}_k^{\langle\text{in.}\rangle} = \frac{1}{\sigma_{\langle\text{in.}\rangle}^k}\frac{1}{n}\sum_j^n \left[\left(z_j^{\langle\text{in.}\rangle} - \bar{z}_j^{\langle\text{in.}\rangle}\right)^k\right]. \tag{A12}$$

We identify these terms in the expression to write the standardized skewness as

$$\begin{aligned}\sigma_{\text{out.}}^3\widetilde{\mu}_3^{\text{out.}} = {}&\sigma_{\text{fid.}}^3\widetilde{\mu}_3^{\text{fid.}} + \sigma_{\text{in.}}^3\widetilde{\mu}_3^{\text{in.}} + \sigma_{\langle\text{in.}\rangle}^3\widetilde{\mu}_3^{\langle\text{in.}\rangle} \\ &+ 3\sigma_{\text{fid.}}^2\sigma_{\text{in.}}S\left(z_j^{\text{fid.}}, z_j^{\text{fid.}}, z_j^{\text{in.}}\right) - 3\sigma_{\text{fid.}}^2\sigma_{\langle\text{in.}\rangle}S\left(z_j^{\text{fid.}}, z_j^{\text{fid.}}, z_j^{\langle\text{in.}\rangle}\right) \\ &+ 3\sigma_{\text{fid.}}\sigma_{\text{in.}}^2S\left(z_j^{\text{fid.}}, z_j^{\text{in.}}, z_j^{\text{in.}}\right) - 6\sigma_{\text{fid.}}\sigma_{\text{in.}}\sigma_{\langle\text{in.}\rangle}S\left(z_j^{\text{fid.}}, z_j^{\text{in.}}, z_j^{\langle\text{in.}\rangle}\right) \\ &+ 3\sigma_{\text{fid.}}\sigma_{\langle\text{in.}\rangle}^2S\left(z_j^{\text{fid.}}, z_j^{\langle\text{in.}\rangle}, z_j^{\langle\text{in.}\rangle}\right) - 3\sigma_{\text{in.}}^2\sigma_{\langle\text{in.}\rangle}S\left(z_j^{\text{in.}}, z_j^{\text{in.}}, z_j^{\langle\text{in.}\rangle}\right) \\ &+ 3\sigma_{\text{in.}}\sigma_{\langle\text{in.}\rangle}^2S\left(z_j^{\text{in.}}, z_j^{\langle\text{in.}\rangle}, z_j^{\langle\text{in.}\rangle}\right),\end{aligned} \tag{A13}$$

where the coskewness of three random variables X, Y, and Z is defined as

$$S(X, Y, Z) = \frac{\text{E}[(X - \text{E}(X))(Y - \text{E}(Y))(Z - \text{E}(Z))]}{\sigma_X\sigma_Y\sigma_Z}. \tag{A14}$$

APPENDIX B: NULLING

Here we introduce an additional optional procedure, which we call nulling that can reduce the error on the mean redshift caused by the PITPZ algorithm. Nulling enforces a requirement that the mean of the delta transformation values be zero for each sample index j, i.e. the mean of the delta transformations be zero for each percentile of the delta transformation distributions.

Recalling our definition of the delta transformation in Section 2, we can write the j^{th} sample of the i^{th} delta transformation as the following difference in redshift values between the i^{th} realization of the input ensemble ($p_i^{\text{in.}}(z)$) and the mean of the input ensemble, $p^{\langle\text{in.}\rangle}(z)$.

$$T_{ij} = z_{ij}^{\text{in.}} - z_j^{\langle\text{in.}\rangle}. \tag{B1}$$

The mean value of the j^{th} sample of each delta transformation over all realizations in the input ensemble is thus:

$$\begin{aligned}\langle T_{ij}\rangle &= \frac{1}{n_{\text{real.}}}\sum_i z_{ij}^{\text{in.}} - z_j^{\langle\text{in.}\rangle} \\ &= -z_j^{\langle\text{in.}\rangle} + \frac{1}{n_{\text{real.}}}\sum_i z_{ij}^{\text{in.}} \\ &= -z_j^{\langle\text{in.}\rangle} + \langle z_{ij}^{\text{in.}}\rangle.\end{aligned} \tag{B2}$$

This quantity does not vanish in general, in particular at the lowest and highest percentiles. These non-zero mean values at each percentile of the delta transformation sum to a non-zero mean value of the ensemble of the delta transformations.

We find empirically that without this procedure, the mean of the delta transformations is approximately 10^{-5}, which leads to an error on the mean redshift in the $p(z)$ of the output ensemble at the level of 10^{-5}. By contrast, applying this procedure decreases the mean of the delta transformations to the level of approximately 10^{-10} at the expense of a slightly more complicated method, and a slight deviation from the conservation rules in Section 4.

[1]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[2]*Kavli Institute for Particle Astrophysics & Cosmology (KIPAC), Stanford University,, PO Box 2450, Stanford, CA 94305, USA*
[3]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[4]*University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstraße 1, D-81679 Munich, Germany*
[5]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

[6]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[7]*Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA*
[8]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[9]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA*
[10]*Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15312, USA*
[11]*NSF AI Planning Institute for Physics of the Future, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[12]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[13]*School of Physics and Astronomy, Cardiff University, CF24 3AA, UK*
[14]*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK*
[15]*Department of Physics, Duke University Durham, NC 27708, USA*
[16]*Cerro Tololo Inter-American Observatory, NSF's National Optical-Infrared Astronomy Research Laboratory, Casilla 603, La Serena, Chile*
[17]*Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA*
[18]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[19]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[20]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[21]*Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK*
[22]*Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain*
[23]*Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ – 20921-400, Brazil*
[24]*Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[25]*Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*
[26]*Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA*
[27]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain*
[28]*Physics Department, William Jewell College, Liberty, MO, 64068, USA*
[29]*Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy*
[30]*INAF-Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy*
[31]*Institute for Fundamental Physics of the Universe, Via Beirut 2, I-34014 Trieste, Italy*
[32]*Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, D-21029 Hamburg, Germany*
[33]*Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India*
[34]*Institute of Theoretical Astrophysics, University of Oslo. PO Box 1029 Blindern, NO-0315 Oslo, Norway*
[35]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[36]*Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, E-28049 Madrid, Spain*
[37]*Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*
[38]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[39]*Department of Astronomy, University of Geneva, ch. d'Écogia 16, CH-1290 Versoix, Switzerland*
[40]*School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
[41]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
[42]*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
[43]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[44]*Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
[45]*Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia*
[46]*Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA*
[47]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[48]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Av. Complutense, 40, E-28040, Madrid, Spain*
[49]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
[50]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany*
[51]*Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, CA 94720, USA*
[52]*Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA*
[53]*School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK*
[54]*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[55]*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK*

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