Calibrating photometric redshifts with intensity mapping observations

All of us[1]
[1] *University of Wherever*
(Dated: January 8, 2017)

TODO: Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

I. INTRODUCTION

TODO: Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

II. FORMALISM

A. Clustering-based photo-z calibration

Consider two galaxy samples with redshift distributions $\phi_i(z)$ ($i = \{1,2\}$), and let $a^i_{\ell m}$ be the harmonic coefficients of their projected overdensity of counts on the sky. Their cross-correlation is given by:

$$\langle a^i_{\ell m}(a^j_{\ell m})^*\rangle = N^{ij}_\ell + S^{ij}_\ell \tag{1}$$

$$S^{ij}_\ell = \frac{2}{\pi}\int dz \int dz'\,\phi_i(z)\phi_j(z')\times \tag{2}$$

$$\times \int dk\,k^2\,b_i(z)b_j(z')P_m(k,z,z')\,j_\ell(k\chi(z))\,j_\ell(k\chi(z')),$$

where P_m is the matter power spectrum, $j_\ell(x)$ is a spherical Bessel function, N^{ij}_ℓ is the cross-noise power spectrum between samples i and j, b_i is the linear bias of the i-th sample and we have neglected redshift-space distortions and all other sub-dominant contributions to the observed power spectrum. In the Limber approximation ($j_\ell(x) \to \sqrt{\pi/(2\ell+1)}\delta^{\mathcal{D}}(\ell+1/2-x)$), this simplifies to:

$$S^{ij}_\ell = \int dk\,P_m(k,z_\ell)\,\frac{H^2(z_\ell)b^i(z_\ell)b^j(z_\ell)}{\ell+1/2}\phi_i(z_\ell)\phi_j(z_\ell), \tag{3}$$

where $\chi(z_\ell) \equiv (\ell+1/2)/k$.

For the purposes of this discussion, the most important feature of Equation 3 is the fact that the amplitude of the cross-correlation is proportional to the overlap between the redshift distributions of those samples. This is especially relevant if one of the samples has good radial resolution, in which case it can be split into narrow bins of redshift. The cross-correlations of all narrow bins with the other sample will therefore trace the amplitude of its redshift distribution, and can effectively be used to constrain it. This is illustrated in Fig. **TODO:** make fig, which shows the cross-power spectrum between a Gaussian photo-z bin of width $\sigma = 0.05$ and a set of narrow redshift bins ($\delta z = 0.002$).

Different recipes have been formulated to carry out this kind of analysis, such as the optimal quadratic estimator method of **TODO:** cite. The forecasts presented here will interpret the redshift distribution (in a parametric or non-parametric form) as a set of extra nuisance parameters, on which we will carry out the Fisher matrix analysis described in Section II D. Thus, even though our results will be optimistic in as much as the Fisher matrix saturates the Rao-Cramer bound, they will account for all correlations between redshift distribution parameters and with the cosmological parameters, as well as the presence of redshift-space distortions and magnification bias (effects that have been overseen in previous works).

For the purposes of estimating the ability of future surveys to calibrate photometric redshift distributions through cross-correlations, we will always consider an individual redshift bin for a photometric sample with unknown distribution together with a set of overlapping narrow redshift bins of spectroscopic galaxies or intensity mapping observations. Let $N^p(z)$ be the overall true redshift distribution of the photometric sample, and let $p(z_{\rm ph}|z)$ be the conditional distribution for a photo-z $z_{\rm ph}$ given the true redshift z. Then, the redshift distribution in a photo-z redshift bin b with bounds $z^i_b < z_{\rm ph} < z^f_b$ is given by

$$\phi_b(z) \propto N^p(z)\int_{z^i_b}^{z^f_b} dz_{\rm ph}\,p(z_{\rm ph}|z). \tag{4}$$

In what follows we will consider 3 degrees of complexity in terms of describing the unknown redshift distribution:

1. We will assume Gaussianly-distributed photo-zs

with a given variance (σ_z^2) and bias Δz:

$$p(z_{\rm ph}|z) \equiv \mathcal{N}(z_{\rm ph} - \Delta z; z, \sigma_z) \equiv \frac{\exp\left[-\frac{1}{2}\frac{(z_{\rm ph}-z-\Delta z)^2}{\sigma_z^2}\right]}{\sqrt{2\pi\sigma_z}}, \quad (5)$$

and we will assume that the uncertainty in the redshift distribution is fully described by Δz and σ_z.

2. We will introduce hard tails in the photo-z distribution (possibly caused by catastrophic outliers) and parametrize it by a pseudo-Voigt profile, combining a Gaussian and Cauchy distributions:

$$p(z_{\rm ph}|z) \equiv f(\sigma_z,\gamma)\,\mathcal{N}(z_{\rm ph} - \Delta z; z, \sigma_z) + (1 - f(\sigma_z,\gamma))\frac{\gamma/\pi}{\gamma^2 + (z-\Delta z)^2}, \quad (6)$$

where f is given as in **TODO:** cite:

$$f(\sigma_z,\gamma) = a. \quad (7)$$

We thus add γ to Δz and σ_z in describing the redshift distribution.

3. We will use a non-parametric form for $\phi_b(z)$, given as a piecewise function with a free amplitude for each spectroscopic redshift bin.

Our assumed fiducial value for Δz, σ_z and γ, as well as the binning scheme used are described in Section II C.

We finish this section by noting that the use of cross-correlations with spectroscopic surveys or intensity mapping observations for photo-z calibration is not limited to the measurement of the redshift distribution of a given galaxy sample, but that they can also be used to improve the precision of photometric redshift estimates for individual galaxies (e.g. **TODO:** CITE). Although we leave the discussion of this possibility for future work, we describe a Bayesian formalism for this task in Appendix A.

B. Intensity mapping

TODO:

- Briefly describe intensity mapping
- Interferometer vs single dish. Noise expressions.
- Foregrounds

C. Photometric redshift surveys

TODO:

- Describe models used for LSST.
- Describe binning scheme.

D. Forecasting formalism

Our formalism will distinguish between two types of tracers of the density field:

- Spectroscopic: tracers whose redshift distribution is well known. This would correspond to tracers with good radial resolution such as a narrow redshift bin of spectroscopic sources or an intensity map in a narrow frequency band, as well as other tracers with a well-known window function, such as a CMB lensing map.
- Photometric: tracers whose redshift distribution is unknown or uncertain. This would correspond to e.g. a photometric-redshift bin, a radio continuum survey or a map of the Cosmic Infrared Background.

Let us start by considering a set of sky maps corresponding to a number of tracers, and let a be the corresponding vector of maps expressed in a given basis. In the following sections we will often be in a situation where a are stored in terms of spherical harmonic coefficients and takes the form $\mathsf{a}_{\ell m} = (p_{\ell m}, s^1_{\ell m}, ...s^{N_s}_{\ell m})$, where $p_{\ell m}$ is a photometric tracer and $s^i_{\ell m}$ is a set of spectroscopic tracers. For the moment, however, we will keep the discussion general.

Assuming that a is Gaussianly distributed with zero mean and covariance $\hat{\mathsf{C}} \equiv \langle \mathsf{a}\mathsf{a}^\dagger \rangle$, its log-likelihood is given by:

$$\mathcal{L} \equiv -2\log p(\mathsf{a}) = \mathsf{a}^\dagger \hat{\mathsf{C}}^{-1}\mathsf{a} + \log(\det(2\pi\hat{\mathsf{C}})). \quad (8)$$

Now let q_i be a set of parameters modelling $\hat{\mathsf{C}}$, including (but not limited to) the parameters describing the photometric redshift distribution. A maximum-likelihood estimator for q_i can be defined by using an iterative Newton-Raphson method to minimize the likelihood in Eq. 8. This is described in **TODO:** CITES, and yields the iterative algorithm:

$$q_i^n = q_i^{n-1} + [\hat{\mathsf{F}}^{-1}]_{ij}\left[\mathsf{a}^\dagger\hat{\mathsf{C}}^{-1}\hat{\mathsf{C}}_{,j}\hat{\mathsf{C}}^{-1}\mathsf{a} - {\rm Tr}(\hat{\mathsf{C}}_{,j}\hat{\mathsf{C}}^{-1})\right], \quad (9)$$

$$\hat{\mathsf{F}}_{ij} \equiv \left\langle \frac{\partial^2\mathcal{L}}{\partial q_i \partial q_j}\right\rangle = {\rm Tr}\left(\hat{\mathsf{C}}^{-1}\hat{\mathsf{C}}_{,i}\hat{\mathsf{C}}^{-1}\hat{\mathsf{C}}_{,j}\right),$$

where, in Eq. 9 there is an implicit summation over j, the sub-index $_{,i}$ implies differentiation with respect to q_i, $\hat{\mathsf{F}}$ is the Fisher matrix, q_i^n is the $n-$th iteration of the solution for q_i and the previous iteration q_i^{n-1} is used to compute $\hat{\mathsf{C}}$ and $\hat{\mathsf{C}}_{,i}$ in the second term. Note that we have simplified a pure Newton-Raphson iteration by taking the ensemble average of the likelihood Hessian (i.e. the Fisher matrix). Furthermore, in the case where the likelihood is well-approximated by a Gaussian, $\hat{\mathsf{F}}^{-1}$ is the covariance matrix of the q_i. Eq. 9 is the basis of the method proposed in **TODO:** CITE (with a number of simplifications) and used in **TODO:** CITE to constrain the redshift distribution of galaxies in the KiDS survey.

In our case, we mainly care about the uncertainty in the redshift distribution parameters included in the q_i, and therefore we will simply estimate the Fisher matrix $\hat{\mathsf{F}}$. In the case where a is a set of spherical harmonic coefficients with power spectrum $\langle \mathsf{a}_{\ell m} \mathsf{a}^\dagger_{\ell' m'} \rangle = \delta_{\ell\ell'}\delta_{mm'}\hat{\mathsf{C}}_\ell$, $\hat{\mathsf{F}}$ is given by

$$\hat{\mathsf{F}}_{ij} = \sum_{\ell=2}^{\ell_{\rm max}} f_{\rm sky}(\ell+1/2)\,{\rm Tr}\left(\hat{\mathsf{C}}_\ell^{-1}\hat{\mathsf{C}}_{\ell,i}\hat{\mathsf{C}}_\ell^{-1}\hat{\mathsf{C}}_{\ell,j}\right), \qquad (10)$$

where we have approximated the effects of a partial sky coverage by scaling the number of independent modes per ℓ by the sky fraction $f_{\rm sky}$. The form of the power spectra $\hat{\mathsf{C}}_\ell$ for the different tracers considered in this work is given in Appendix B.

As explicitly shown in Eq. 10, smaller-scale modes carry a higher statistical weight (proportional to $\sim \ell$), and would in principle dominate the redshift distribution constraints. The smallest scales are, however, dominated by theoretical uncertainties from non-linearities in the evolution of the density field and the galaxy-halo connection, and therefore a multipole cutoff $\ell_{\rm max}$ must be used to contain the constraining power of systematics-dominated modes. In this paper we use a redshift-dependent cutoff defined as follows. Let z be the mean redshift of a given redshift bin, and let $\sigma^2(k_*)$ be the variance of the linear density field at that redshift on modes with wavenumber $k < k_*$:

$$\sigma^2(k_{\rm max}) \equiv \frac{1}{2\pi^2}\int_0^{k_*} dk\, k^2\, P_m(k,z). \qquad (11)$$

We then define the cutoff scale as $\ell_{\rm max}(z) = \chi(z)\,k_{\rm max}(z)$, where $k_{\rm max}(z)$ satisfies $\sigma(k_{\rm max}, z) = \sigma_{\rm thr}$ for some choice of $\sigma_{\rm thr}$. In what follows we will use a fiducial threshold $\sigma_{\rm thr} = 0.75$, corresponding to $k_{\rm max}(z=0) \simeq 0.2\,h\,{\rm Mpc}^{-1}$, and we will study the dependence of our results on this choice. Besides this choice of $\ell_{\rm max}$, we will also impose a hard cutoff for all galaxy-survey and intensity-mapping tracers of $\ell < 2000$ (thus, in reality, $\ell_{\rm max} = \min(\chi k_{\rm max}, 2000)$).

III. RESULTS

A. Baseline forecasts

TODO:

- Describe results for SKA and HIRAX in comparison with DESI, Euclid, WFIRST.
- Show forecasts on parameters before and after calibration.

B. Dependence on experimental parameters

TODO:

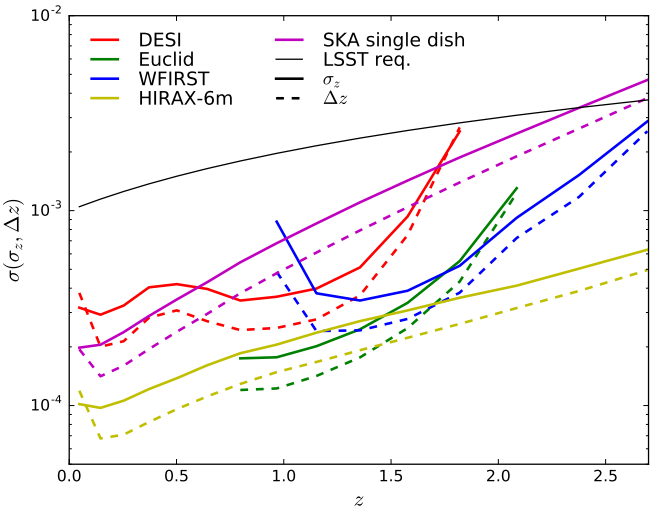

FIG. 1: **TODO:**

- Show results as a function of $f_{\rm sky}$, and noise level.
- Show results as a function of dish size for single-dish experiments
- Show results as a function of minimum baseline for interferometers

C. Foregrounds

TODO:

- Show results in the presence of foreground residuals.
- As a function of residual amplitude?
- Discuss correlated extragalactic foregrounds?

D. Generalized redshift distributions

TODO:

- Include hard-tail parameters (e.g. Cauchy contribution - Voigt profile).
- Non-parametric calibration. Show results for bins of dN/dz of fixed width as a function of experiment and redshift.

IV. DISCUSSION

TODO: Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy

sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo.

Acknowledgments

We thank Odin the almighty for useful comments and discussions.

APPENDIX A: INDIVIDUAL CLUSTERING REDSHIFTS

TODO: Possibly describe formalism to sharpen redshifts for individual redshifts.

APPENDIX B: ANGULAR POWER SPECTRA

TODO: Describe models used for the angular power spectra