Large Synoptic Survey Telescope

www.lsst.org

Photometric Redshift Performance of LSST

Samuel Schmidt[1], J. A. Newman[2], A. J. Connolly[3], Z. Ivezic[3], J. A. Tyson[1], D. J. Matthews[2], LSST Collaboration

[1]Univ. of California, Davis, [2]Univ. of Pittsburgh, [3]Univ. of Washington

LSST will provide a view of galaxies over a significant fraction of the age of the universe. In order to take full advantage of this vast dataset the LSST cosmological science requires photometric redshifts to meet stringent goals. We describe the expected photometric redshift performance of LSST based on empirical studies and simulations, and ongoing work to characterize and minimize redshift uncertainties. We discuss the accuracy of photometric redshifts and our ability to photometrically calibrate the data. Bayesian priors incorporate additional knowledge such as surface brightness in the redshift determination. We discuss training sets needed to span the full range of galaxy properties, and methods to deal with incompleteness. Cross-correlation enables measurement of accurate photometric redshift error distributions, constraining redshift uncertainties for faint galaxies even with spectra of only a bright subsample.

The LSST *ugrizy* filter system provides leverage for redshift estimation from 0<z<6. While photo-z's will be computed for galaxies to *r*~27.5, the majority of science cases will use a subset with *i*<25.3: the "Gold" sample. Science requirements for this S/N>20 "Gold" sample are:

- RMS scatter in uncertainties, $\sigma_z/(1+z)<0.05$
- Fraction of 3σ outliers below 10%
- Bias, $e_z=(z_{phot}-z_{spec})/(1+z_{spec}) < 0.003$
- Uncertainties on σ_z must be known to better than 1%.

Figure 1 shows the expected performance of the LSST Gold sample at full 10 year survey depth. These results are derived from simulated photometry that reproduces colors and luminosities as a function of redshift of the COSMOS survey. Magnitude and surface brightness priors have been applied, and sources with broad or multiply peaked redshift probability distributions have been excluded. For the LSST filters the uncertainty in photometric redshift scales approximately linearly with S/N, with a floor of σ_z~0.02. The tremendous increase in photometric S/N for LSST accounts for much of the improvement over previous photo-z surveys. The number of catastrophic outliers also depends on S/N (Mobasher et al. 2007). With the techniques described below we believe that LSST will yield significantly better photometric redshifts than previous optical surveys, sufficient to meet design goals.

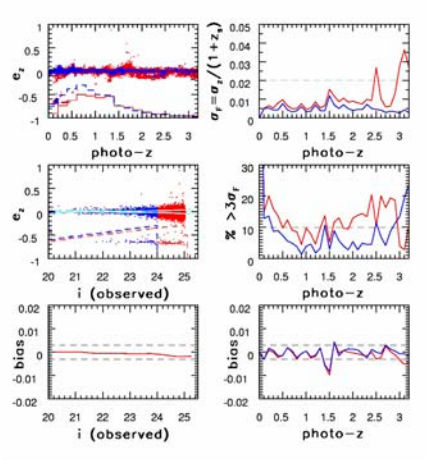

Figure 1: Illustration of the photometric redshift performance as a function of apparent magnitude and redshift for a simulation of the LSST *ugrizy* filter set. *Red* points and curves correspond to the "gold" sample defined by *i*<25, and *blue* points and curves to a subsample with *i*<24. The photometric redshift error is defined as $e_z=(z_{phot}-z_{spec})/(1+z_{spec})$.

Top Left: e_z vs.photometric redshift. The two histograms show redshift distributions of the galaxy simulations. Top right: RMS scatter of e_z as a function of photometric redshift. The horizontal dashed line shows science driven design goal. *Middle left*: e_z vs. observed *i* band magnitude. The two histograms show logarithmic differential counts. The horizontal lines show the 3σ envelope around the median e_z (where sigma is the RMS from the top right panel). *Middle right*: the fraction of 3σ outliers as a function of redshift (horizontal line shows the LSST design goal). *Bottom left*: Median value of e_z (bias) as a function of apparent magnitude. The two dashed lines show design goal for the bias. *Bottom right*: The median value of e_z vs. redshift. The two dashed lines show the design goal.

Priors And Outlier Rejection

To mitigate catastrophic redshift failures Bayesian methods have been developed (Benitez 2000), where prior information about the redshift distribution is invoked in estimating the redshift. A magnitude-type prior is most commonly used. Surface brightness, with its stronger redshift evolution, should provide stronger constraints (Stabenau et al. 2008), particularly for reducing the number of catastrophic outliers (**Figure 2**).

The mapping of broadband colors to redshift employed by photo-z algorithms is a classic inverse problem. As such, there are regions of parameter space where the mapping is not unique, causing redshift degeneracies and catastrophic failures. By identifying and prefiltering these regions, we can eliminate large percentage of the catastrophic outliers while retaining galaxies with well constrained redshifts. We further reject outliers with broad or bimodal redshift probability distributions. We will need extensive spectroscopic training sets to identify these regions and construct our spanning template set.

We estimate that we will need ~50,000 galaxies to *i*<25 spread over several calibration fields in order to characterize template SEDs, as well as to measure the bias and uncertainty for this spectroscopic subsample as a check on our algorithms. A portion of these redshift samples will also be used in the cross-correlation analysis described below. See the *LSST Science Book* for details.

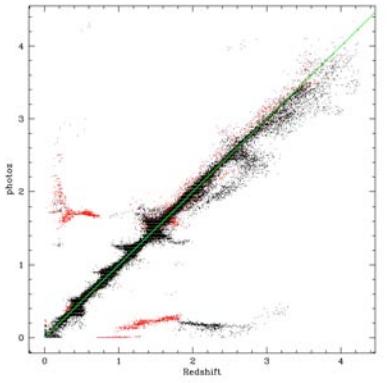

Figure 2: Photometric versus spectroscopic redshift for a mock galaxy simulation of the LSST "Gold" sample to i=25th magnitude. Half of the templates used to generate the mocks are removed from the photo-z algorithm to simulate template incompleteness. Redshift dependent scatter and catastrophic outliers are clearly visible. A type dependent magnitude prior has been applied. Red points indicate galaxies identified as outliers either by a naïve surface brightness/type prior or occupying degenerate regions of color space. Even this very simple formulation is able to eliminate a large fraction of the catastrophic failures.

Cross Correlation Techniques:

One method that avoids incompleteness issues employs cross-correlation information (Newman 2008). Past redshift surveys suggest that we will not be successful in obtaining redshifts for all targeted galaxies, and that the failure rate is a strong function of both type and redshift. Even with very high completeness redshift distributions may be biased beyond tolerances of LSST dark energy goals (Banerji et al. 2008). Because galaxies cluster spatially only on local scales, any observed clustering between a photo-z sample and galaxies at some fixed redshift z_s can only arise from galaxies near redshift z_s. By measuring the angular cross correlation between a photometric sample and a spectroscopic sample as a function of redshift we can recover the redshift distribution of the photometric sample, even if the spec-z sample is not representative. A cartoon description of the method is shown in **Figure 3**. **Figure 4** shows the redshift distribution reconstruction of a Monte Carlo simulation of galaxies derived from the Millennium simulation (Croton et al. 2006). Detecting non-Gaussianities (e.g. tails) in these distributions is straightforward.

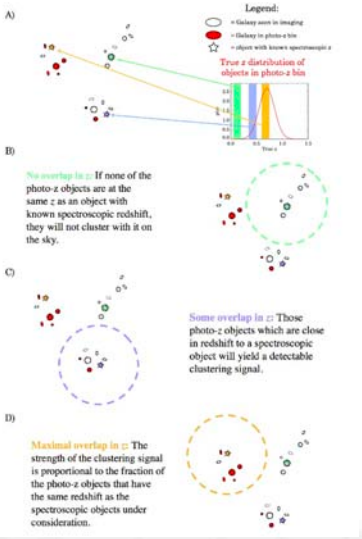

Figure 3: Cartoon of cross-correlation photometric redshift calibration (Newman 2008). **A)** shows the basic situation: we have imaging for many galaxies (circles and ellipses), some of which fall in the photometric redshift bin of interest (red). Galaxies that are near to each other in three dimensions cluster together on the sky. We know the spectroscopic redshifts of a smaller sample of objects (starred symbols). The true redshift distribution for the objects in the photo-z bin is assumed to be a Gaussian with mean z=0.7 (see inset plot); The stars are color-coded according to redshift range that the galaxy lies within, with corresponding colors shown on the inset. **B)** For spec-z objects that do not overlap in redshift with the photo-z bin there will be no excess of neighbors in the photo-z sample. **C)** If there is some overlap, there will be some excess of neighbors around the spectroscopic object from the photo-z bin **D)** The strength of this clustering signal will be stronger the greater the fraction of the photometric redshift sample lies at the same z as the spectroscopic object in question. Thus we can reconstruct the true redshift distribution of the photo-z sample by measuring the clustering with objects of known redshift as a function of spectroscopic redshift.

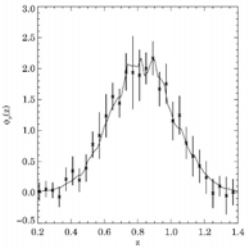

Figure 4: Recovery of redshift distribution with cross-correlation techniques. The solid line is the true redshift distribution of a subset of those galaxies with M_B-5log *h* < -17 in 24 0.5x2 degree light-cone mock catalogs constructed from the Millennium Simulation semi-analytic models of Croton et al. (2006), with a probability of being included in the set given by a Gaussian in redshift, centered at z=0.8 and with dispersion σ_z = 0.2. Points and error bars show the median and RMS variation in the cross-correlation reconstruction of this true distribution using a spectroscopic sample consisting of 60% of all galaxies down to R_{AB}=24.1 in only four of the 24 fields.

The true distribution may be reconstructed to the accuracy required by LSST using spectroscopic samples of realistic size.

References:

Banerji, M., et al. 2008, MNRAS, 386, 1219
Benitez, N. 2000, ApJ, 536, 571
Croton, D., et al. 2006, MNRAS, 365, 11
Mobasher, B., et al. 2007, ApJS, 172, 117
Newman, J. 2008, ApJ, 684, 88
Stabenau, H., et al. 2008, MNRAS, 387, 1215

January 2010

LSST is a public-private partnership. Design and development activity is supported in part by the National Science Foundation. Additional funding comes from private gifts, grants to universities, and in-kind support at Department of Energy laboratories and other LSSTC Institutional Members.