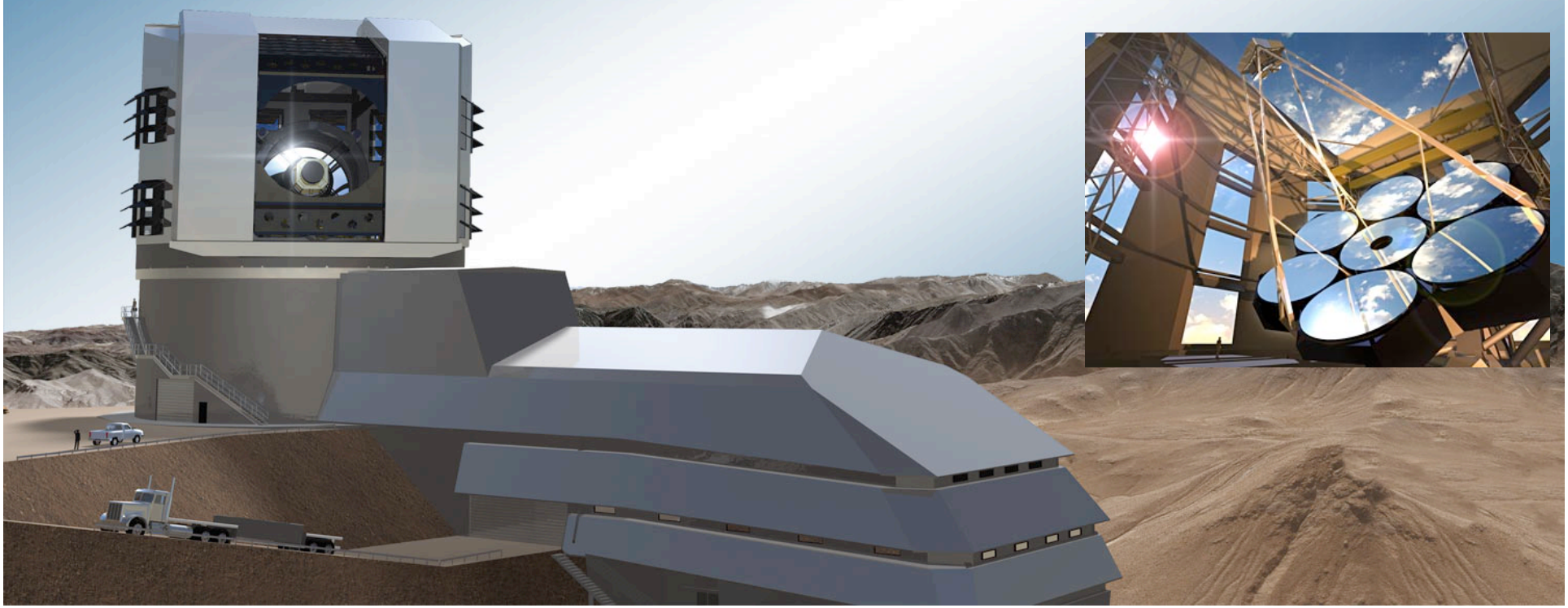
Photometric Redshift Calibration with Extremely Large Telescopes

Jeffrey Newman, U. Pittsburgh / PITT-PACC

Analysis Coordinator and acting Photo-z Working Group convener, LSST Dark Energy Science Collaboration

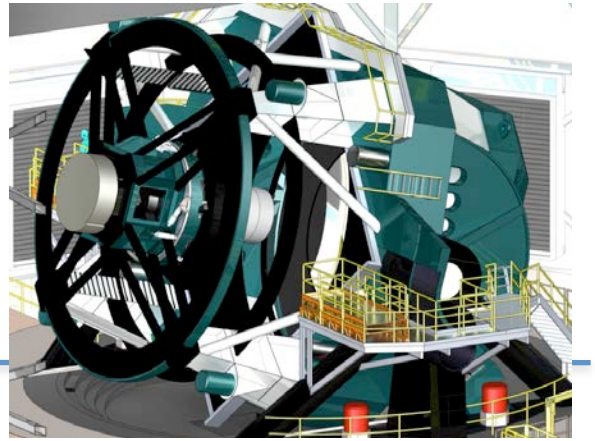


Outline

- The context in ~2020+: LSST
  - Euclid, WFIRST, etc. will have similar requirements
- The importance of photometric redshifts for cosmology
- The importance of GMT for photometric redshifts
  - GMT is likely the best ELT for photo-z calibration
  - The greatest impact of GMT on cosmology should come via photo-*z* calibration
  - A well-designed photo-*z* calibration survey would be a terrific dataset for studying galaxy evolution!
- What sort of training / calibration dataset could be achieved with GMT?

A brief review of LSST

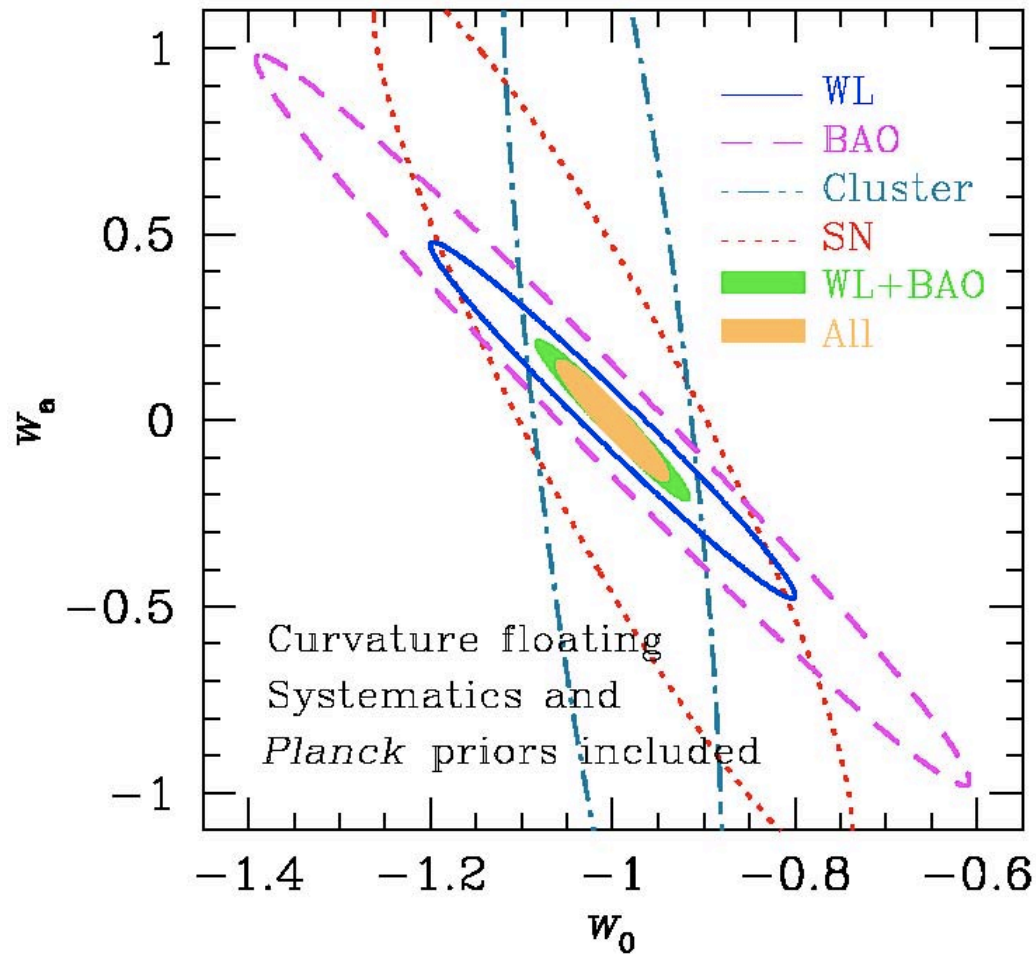
- 8m diameter survey telescope, deep imaging in 6 filters (*ugrizy*)
- 30 sec images of 10 sq. deg. at a time
  - 900 visits per night, cover visible sky every 3 nights
- 10-year total survey: combine for extremely deep imaging over 50% of sky
- Science enabled:
  - Cosmology (dark matter, dark energy, testing GR, etc.)
  - Mapping the Milky Way
  - Revealing the Transient Universe
  - Inventory of the Solar System



LSST constrains dark energy in many ways… all will rely on redshift information

- LSST will constrain dark energy via 4 major probes:
  - Weak gravitational lensing
  - Baryon Acoustic Oscillations
  - Type Ia supernovae
  - Cluster counts
  - (Plus strong lensing, etc.)

- For all of these, we want to measure some observable as a function of redshift

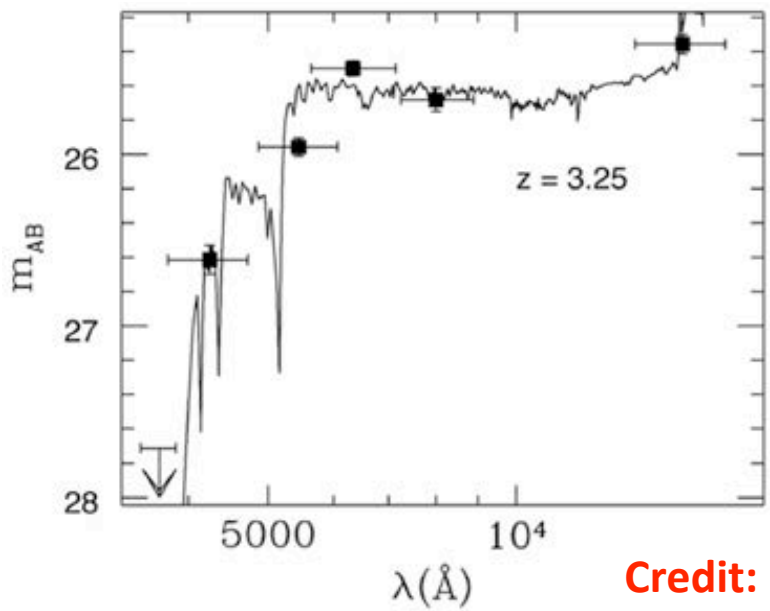


LSST Science Book

Spectroscopy provides ideal redshift measurements – but is infeasible for large samples

- At LSST "gold sample" (*i*<25.3) depths, ~180 hours on a 10m telescope to determine a redshift (70% of time) spectroscopically
- With a next-generation, 5000 fiber spectrograph on a 10m telescope, still >50,000 telescope-years to measure redshifts for LSST "gold" weak lensing sample (4 billion galaxies)!

- Alternative: use broad spectral features to determine *z* : a photometric redshift
- Advantage: high multiplexing
- Disadvantages: lower precision, calibration uncertainties



Credit: ESO

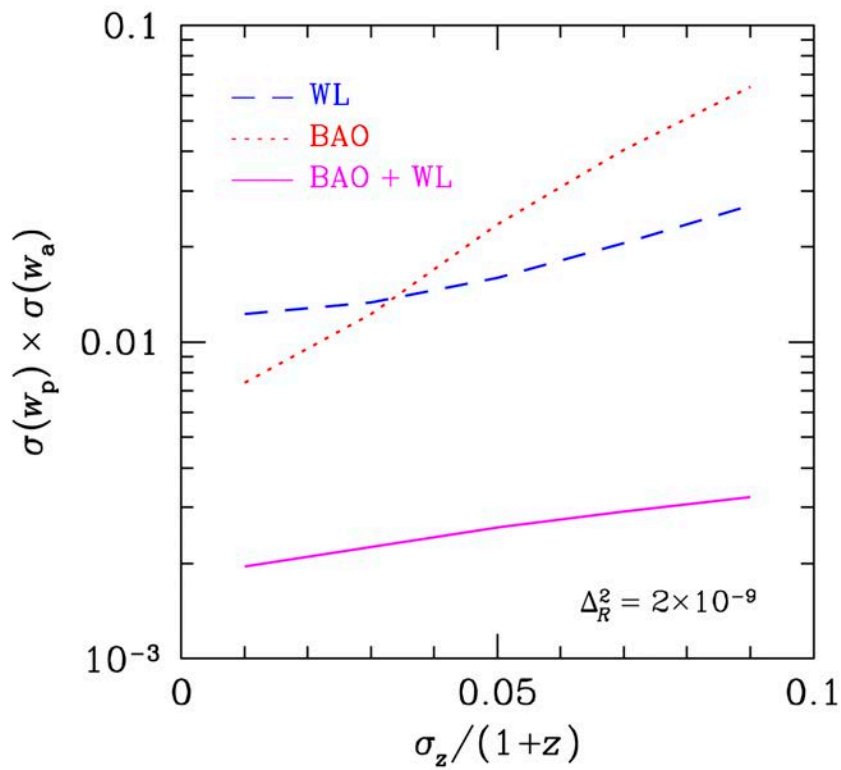
Two basic Photo-z methods: Template fitting and training-based

- **Template fitting**: use galaxies with known *z* to calibrate set of underlying galaxy spectral energy distributions (SEDs) and photometric band-passes
  - Determine posterior probability distribution for *z* | *ugrizy*
  - Needs spectra of galaxies spanning full range of possible properties to tune templates, establish priors, etc.

- **Training-based:** Use galaxies with known redshift **and** uniform/well-understood sampling to determine a relationship between *z* and colors
- Training set **MUST** span **full** range of properties & z of galaxies
  - **Pro:** Can take advantage of progress in machine learning & stats
  - **Con:** Sensitive to systematic incompleteness in training sets; extrapolate poorly

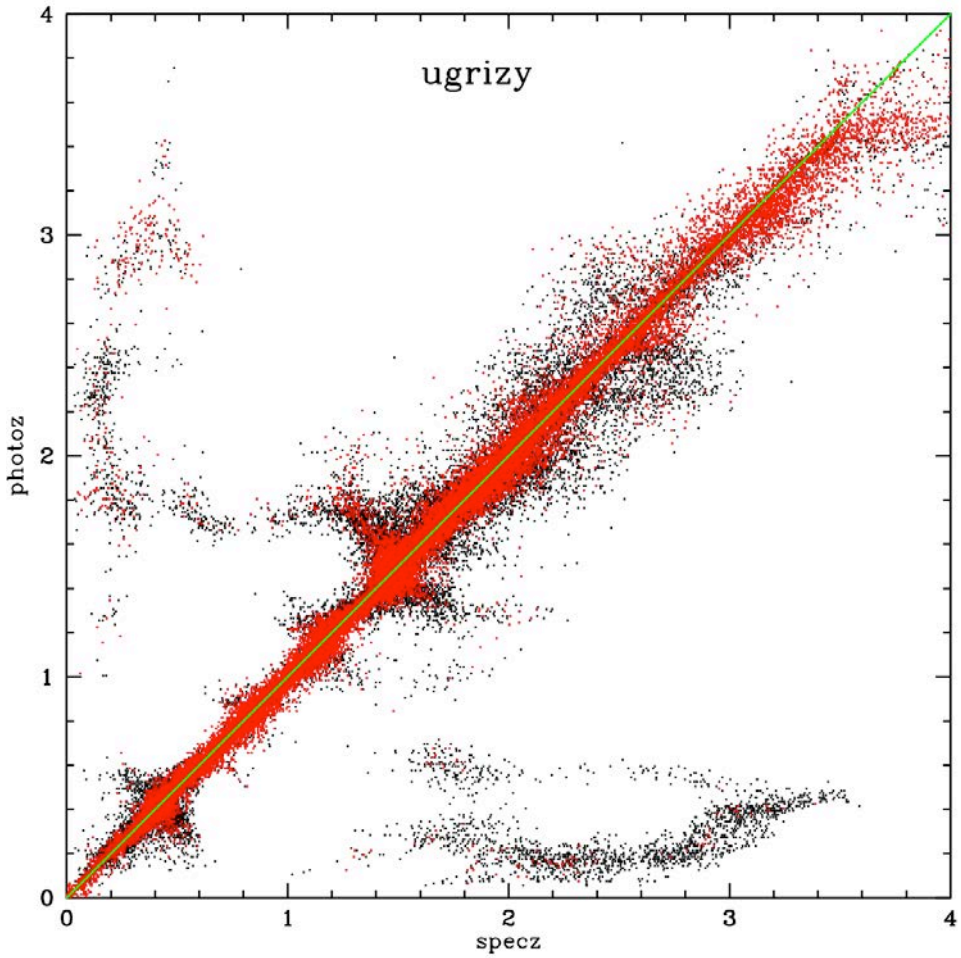
Two spectroscopic needs for photo-z's: calibration and training

- For LSST weak lensing and supernovae, individual-object photo-*z*'s do not need high precision, but the calibration must be accurate
  - bias ($\delta_z = \langle z_p - z_s \rangle$), and error in scatter, $\sigma(\sigma_z) = \sigma(\mathrm{RMS}(z_p - z_s))$, must both be $<\sim 0.002(1+z)$)
- Science reach, especially for BAO and clusters, improves as photo-z errors get smaller: better training sets allow LSST to perform better than requirements
  - Training sets will contribute to calibration, including via cross-correlations. Perfect training sets can solve calibration needs.

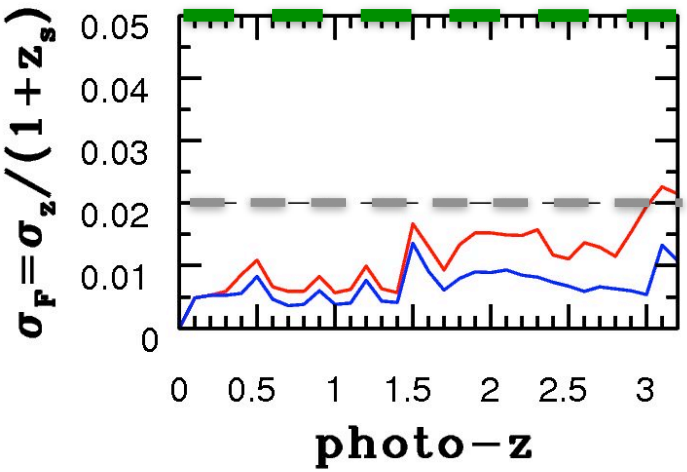


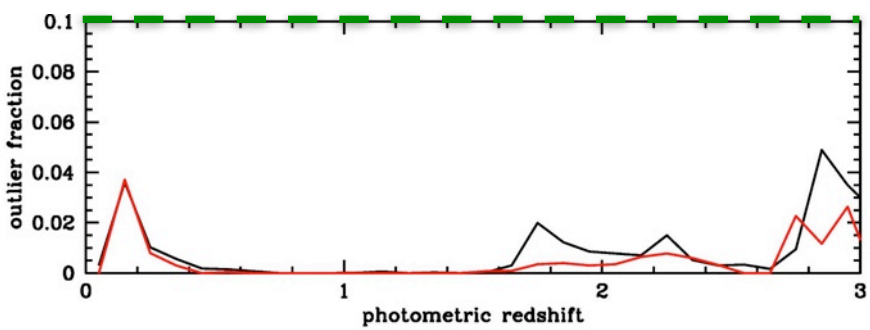
Zhan 2006, JCAP

Expected photo-z performance for LSST *ugrizy*



Green: Requirements on actual performance; grey: requirements on performance with perfect template knowlege (as in these sims)





Spectroscopic training set requirements

- Goal: make δ_z and $\sigma(\sigma_z)$ so small that systematics are subdominant
- Many estimates of training set requirements (Ma et al. 2006, Bernstein & Huterer 2009, Hearin et al. 2010, LSST Science Book, etc.)
- General consensus that roughly 20k-30k extremely faint galaxy spectra are required to characterize:
  - Typical z_{spec}-z_{phot} error distribution
  - Accurate catastrophic failure rates for all objects with $z_{phot} < 2.5$
  - Characterize all outlier islands in z_{spec}-z_{phot} plane via targeted campaign (core errors easier to determine)
- Those numbers of redshifts are achievable with GMT, if multiplexing is high enough

What qualities do we desire in our training sets?

- **Sensitive spectroscopy of faint objects (to *i*=25.3)**

  - Need a combination of large aperture and long exposure times; >20 Keck-nights (=4 GMT-nights) equivalent per target, minimum

- **High multiplexing**

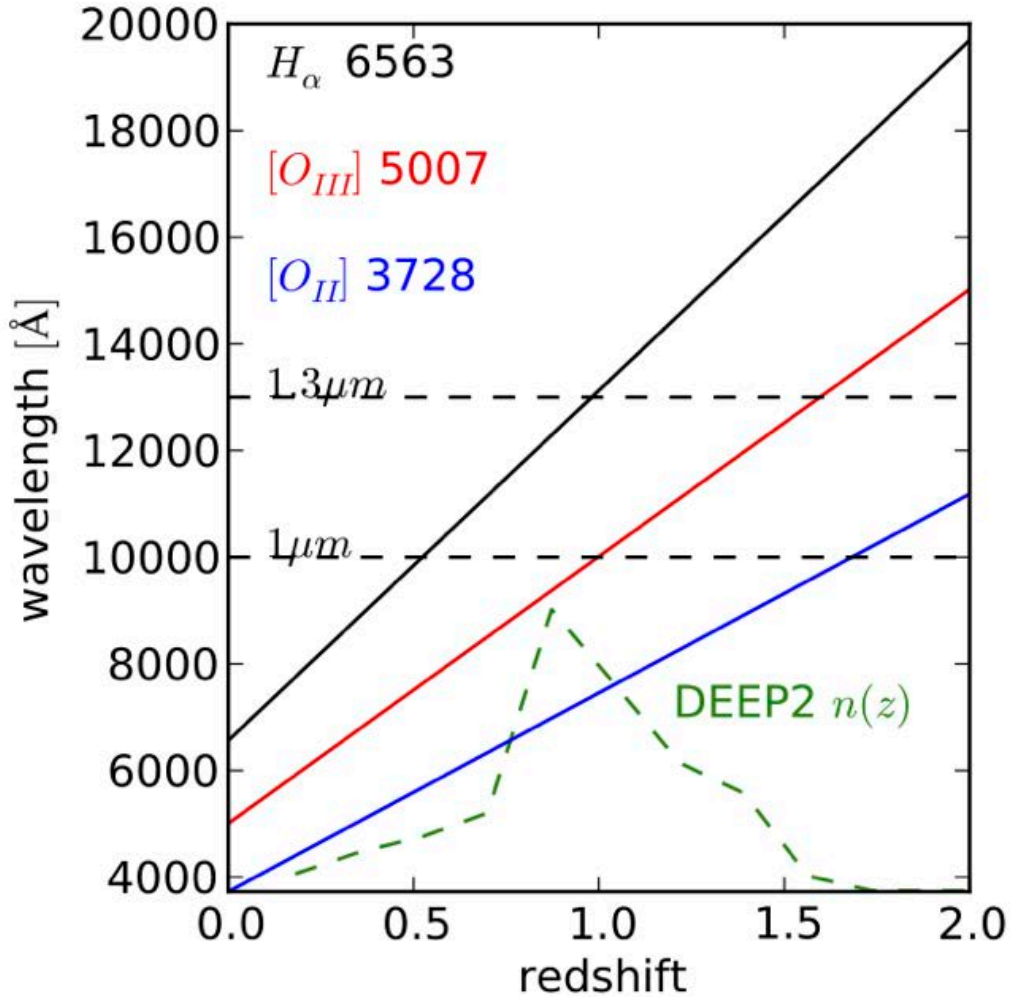
  - Obtaining large numbers of spectra is infeasible without it

What qualities do we desire in our training sets?

- **Coverage of full ground-based window**

  - Ideally, from below 4000 Å to ~1.5µm

  - Require multiple features for secure redshift

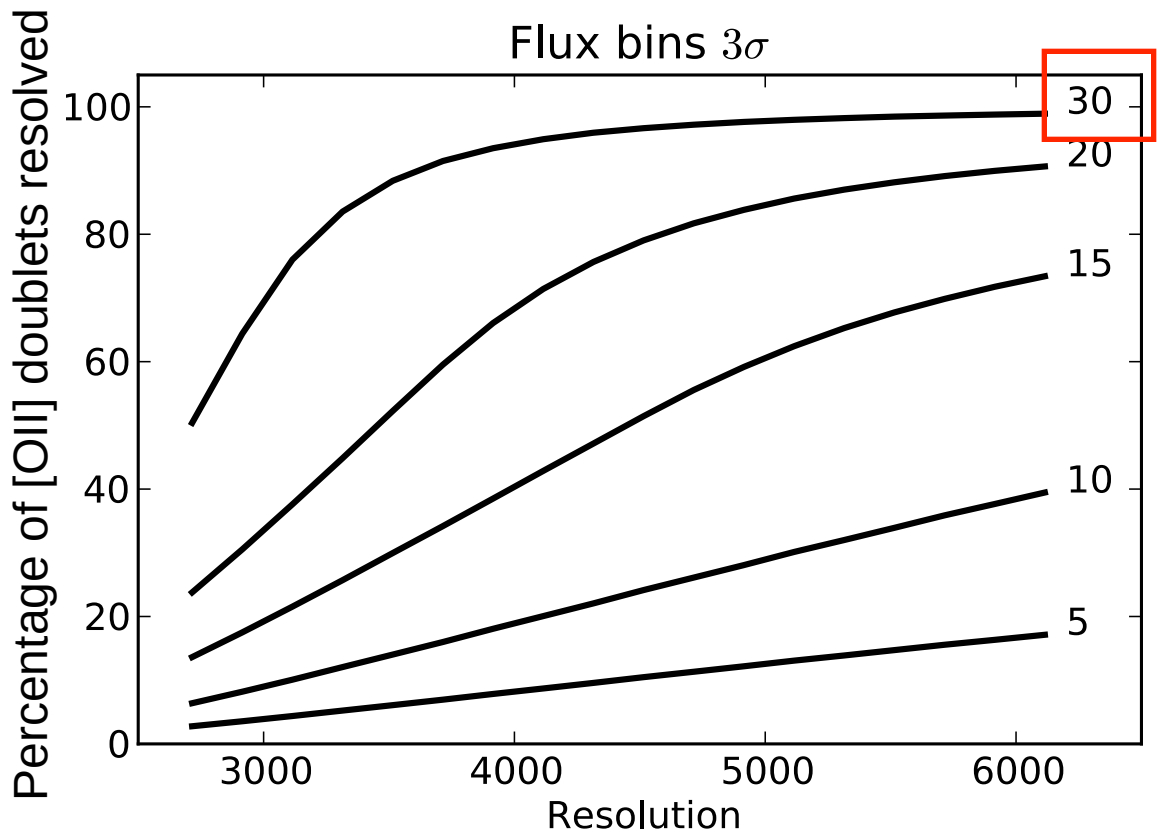


Comparat et al. 2013, submitted

What qualities do we desire in our training sets?

- **Significant resolution (R>~4000) at red end**

  - Allows redshifts from [OII] 3727 Å doublet alone, key at *z*>1



Comparat et al. 2013, submitted

What qualities do we desire in our training sets?

- **Field diameters > ~20 arcmin**

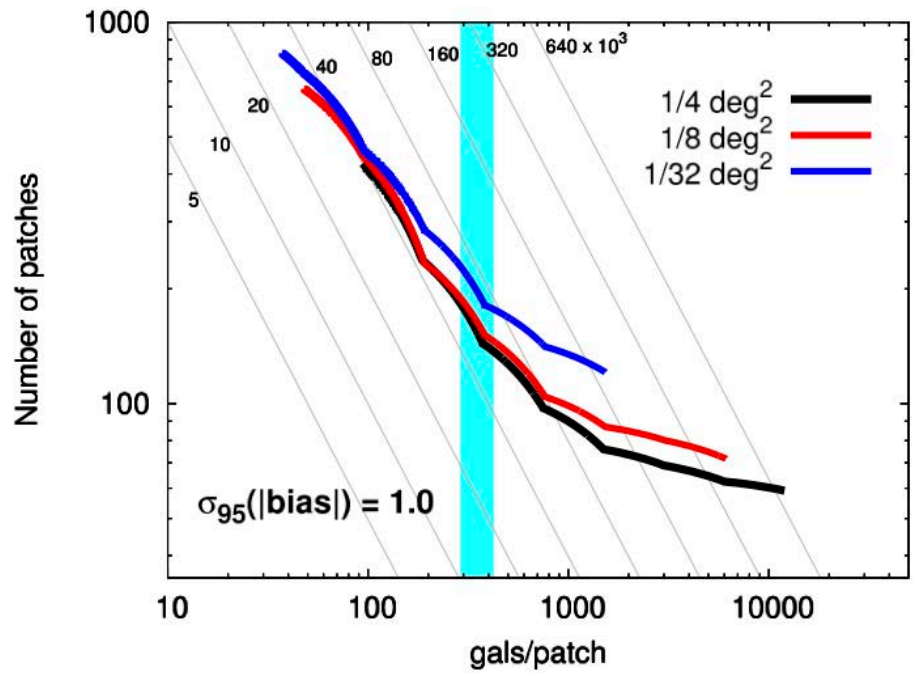
  - Need to span several correlation lengths for accurate clustering measurements (key for galaxy evolution science and cross-correlation techniques)

  - r_0 ~ 5 h^{-1} Mpc comoving corresponds to ~7.5 arcmin at z=1, 13 arcmin at z=0.5

- **Many fields**

  - Minimizes impact of sample/ cosmic variance.

  - e.g., Cunha et al. (2012) estimate that 40-150 ~0.1 deg^2 fields are needed for DES for sample variance not to impact errors (unless we get clever)



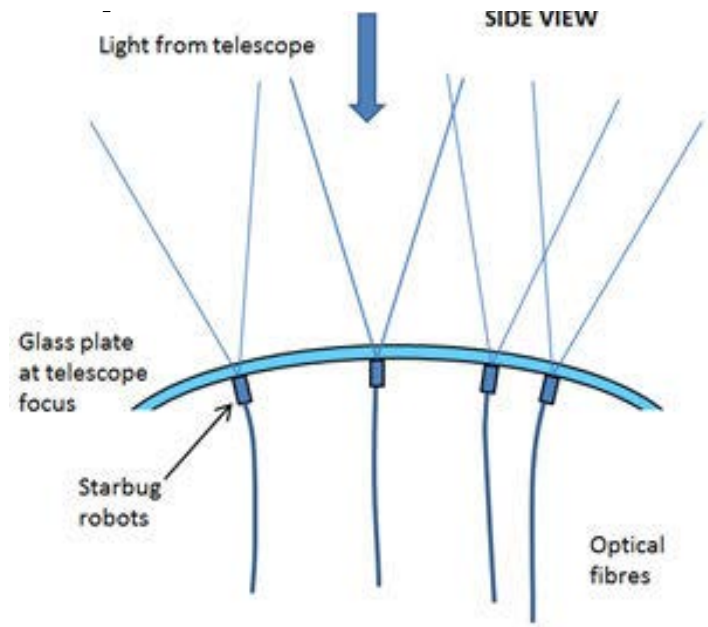
Cunha et al. 2012

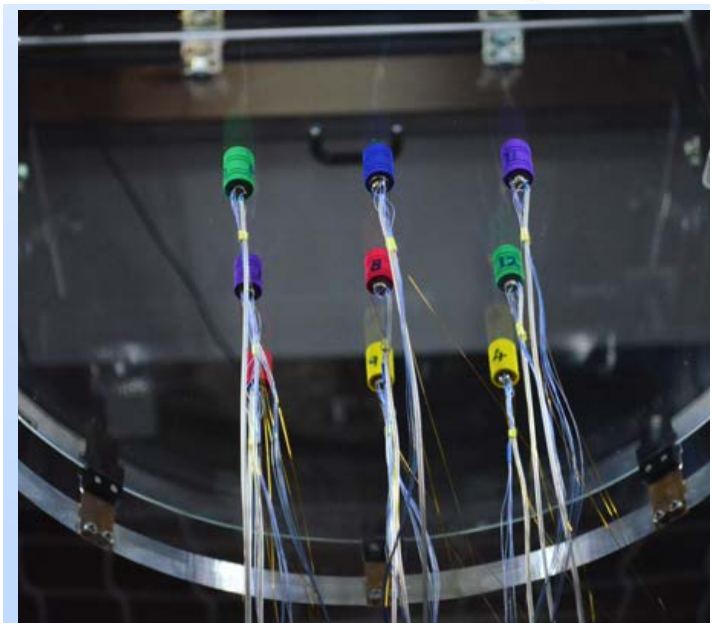
GMT can be particularly well-suited to this work

- For photo-z calibration, field of view and multiplexing are key: doubling the number of slits/fibers is as useful as doubling the collecting area

- ~40 LSST "gold sample" lensing objects per square arcmin: to get ~2500x multiplexing, need at least 60 arcmin^2 FOV, but more is MUCH better as can optimize targets

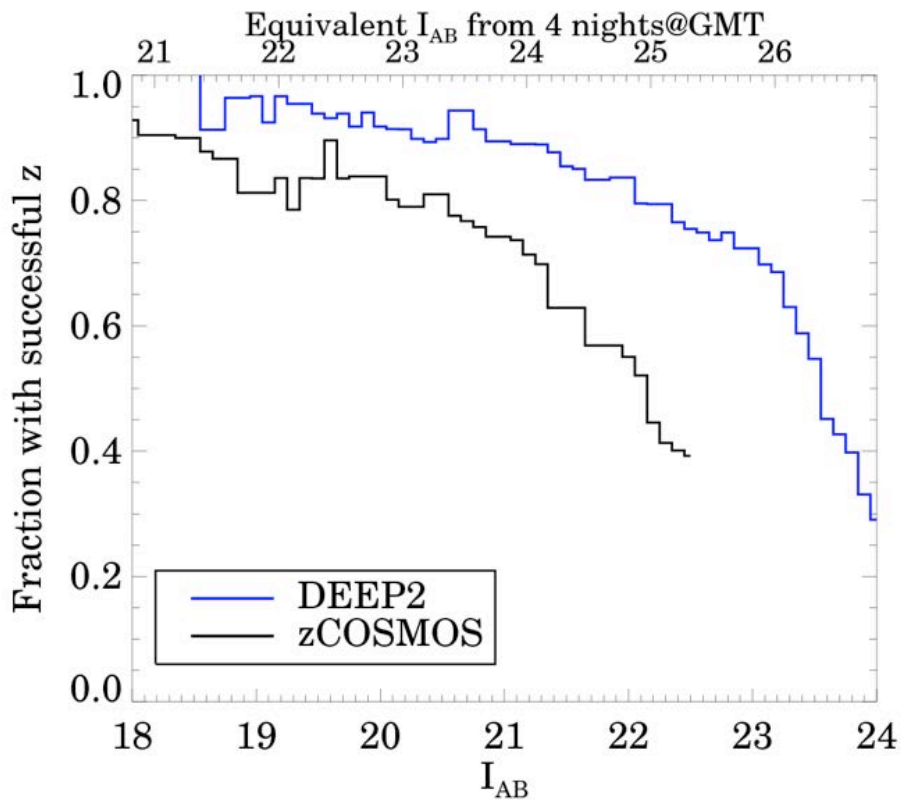
- MANIFEST could be VERY nice for this work; GMACS is close to being the spectrograph we'd like





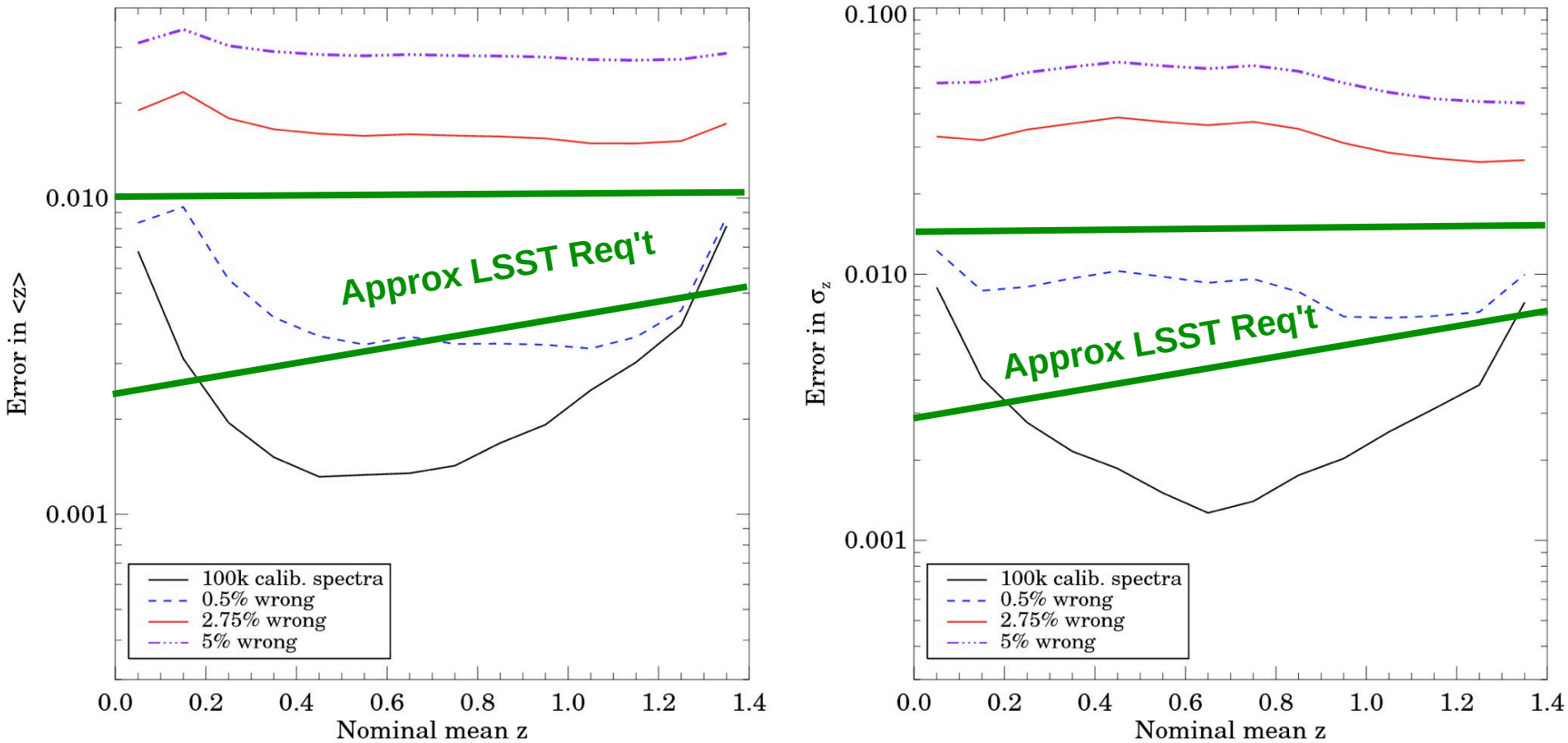
Biggest concern: incompleteness in training sets

- In current deep redshift surveys (to i~22.5/R~24), 30-60% of targets fail to yield secure (>95% confidence) redshifts
- Losses are worst at the faint end
- Redshift success rate varies with galaxy color, redshift, etc.
- In DEEP2, best parts of *BRI* color space have ~90% redshift success
- 4 night GMT depth would yield ~70% completeness; achieving 90% would require ~25 nights/ pointing



Data from DEEP2 (Newman et al. 2013) and zCOSMOS (Lilly et al. 2009)

Note: even for 100% complete samples, current false-z rates would compromise Dark Energy inference



- **Only the highest-confidence redshifts may be useful for calibration / training: lowers spectroscopic completeness further**

Based on actual redshift distributions for ANNz-defined DES bins in mock catalog from Huan Lin, UCL & U Chicago, provided by Jim Annis

Conclusions

- GMT is likely the best ELT for photo-z training & calibration
- The greatest impact of GMT on cosmology is via photo-*z*'s
- Minimum photo-z training survey, ~70% complete:
  - 15 widely-separated pointings, 4 nights/pointing

    = 60 clear nights, 30,000 spectra to i = 25.3
  - Sample objects over full range of galaxy SEDs, $0 < z < 3.7$
  - This would be a VERY interesting galaxy-evolution survey
  - LSST DESC Clusters WG is interested in cluster sight-line surveys
- Our dream survey: 50 widely-separated pointings, 25 nights/pointing = 3.4 clear years, 100,000 spectra to i = 25.3
  - This would be ~19 clear years with Subaru/PFS. . .
- Even the dream survey should be systematically incomplete. Cross-correlation methods (Newman 2008) can solve this problem.

3 Ways to address spectroscopic incompleteness – all appear feasible (but we need to explore this!)

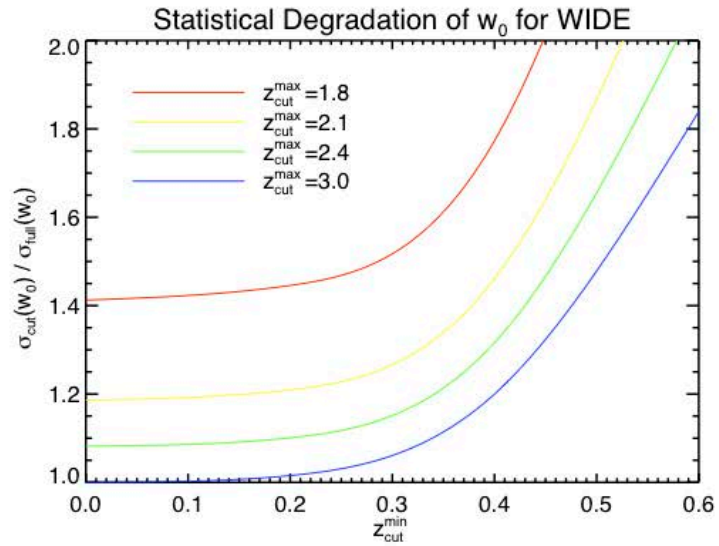
I. Throw out objects lacking secure photo-*z* calibration
   - ID regions in *ugrizy* space where redshift failures occurred
   - Eliminating a fraction of sample has modest effect on FoM

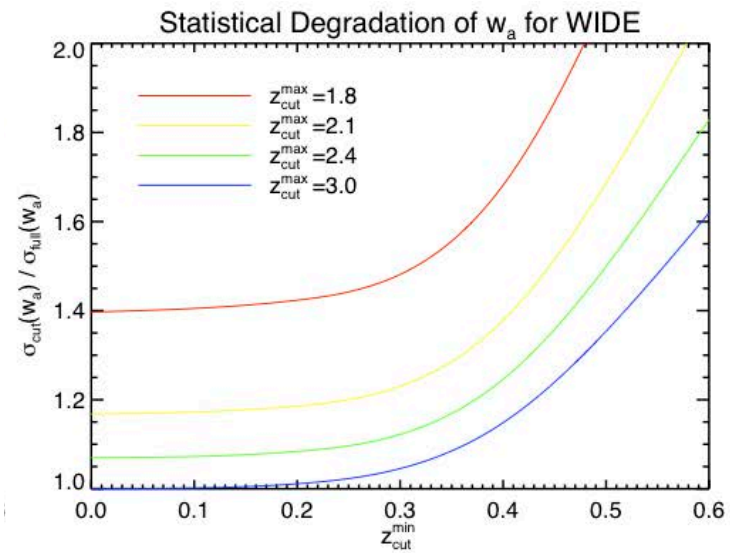
II. Incorporate additional information
   - Longer exposure/wider wavelength range spectroscopy (JWST, etc.) for objects that fail to give redshifts in first try
   - Deep, medium-band/multiwavelength imaging in limited regions to get sample with secure, accurate photo-z's as training set
   - Develop comprehensive model of galaxy spectral evolution constrained by redshifts obtained

III. Cross-correlation techniques

Effect of rejecting objects with particularly low or particularly high photo-z's

- Can mitigate catastrophic outlier impact by throwing out objects with photo-z's in problematic ranges
- Plots at right: weak lensing error degradation (vs. random errors only) as change minimum redshift (x axis) and maximum redshift (different-colored curves)

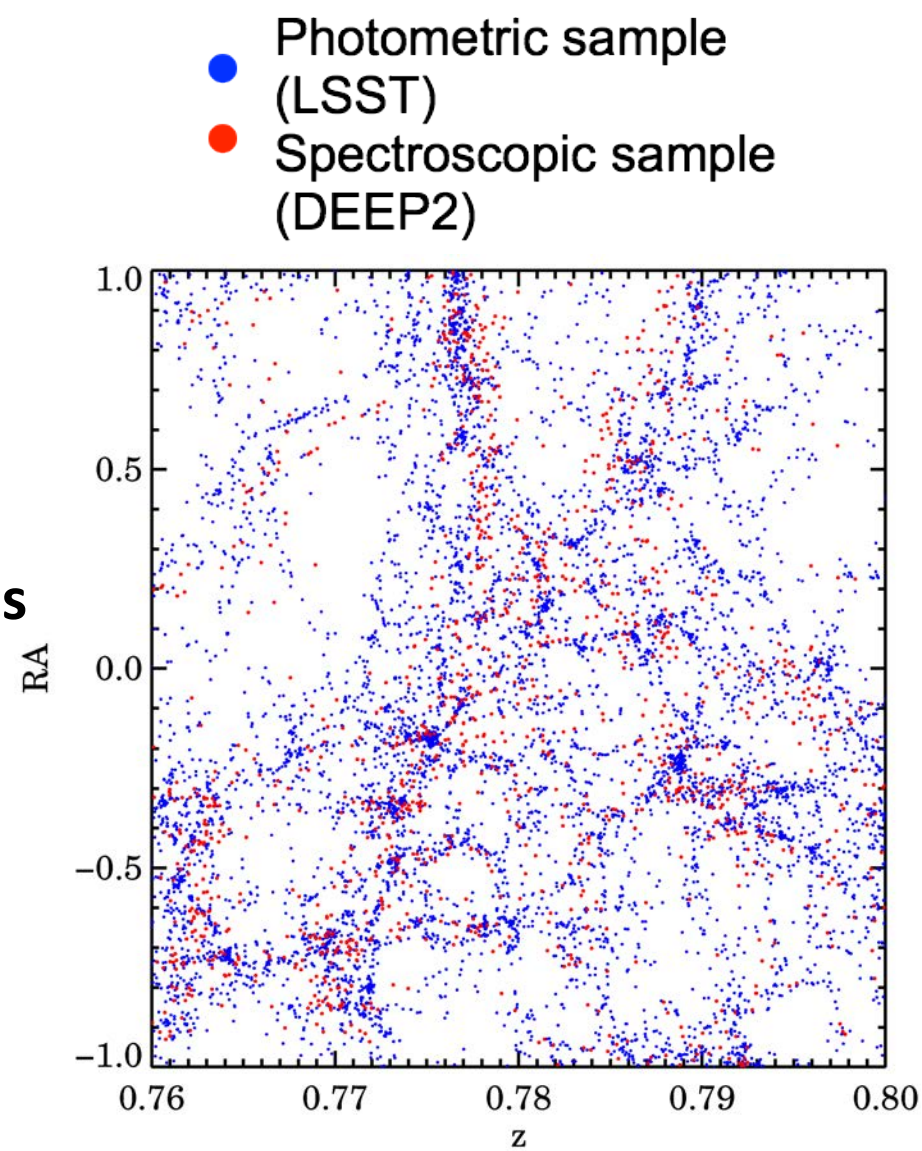




Hearin et al. 2010

Cross-correlation methods: exploiting redshift information from galaxy clustering

- Galaxies of all types cluster together: trace same dark matter distribution
- Galaxies at significantly different redshifts do not cluster together
- From observed clustering of objects in one sample with another (as well as information from their clustering on their own), can determine fraction of objects in overlapping redshift range



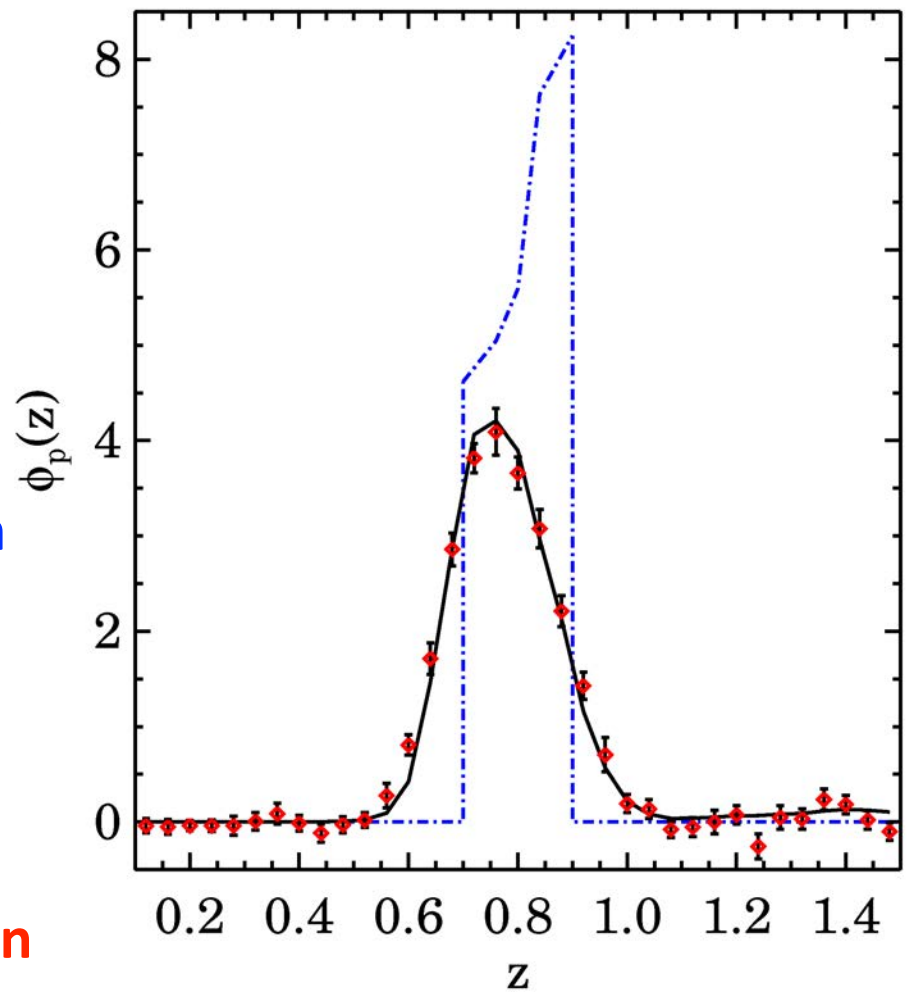
Higher-resolution information can be obtained by cross-correlating with spectroscopic samples

- Key advantage: spectroscopic sample can be systematically incomplete and include only bright galaxies
- See: Newman 2008, Ho et al. 2008, Matthews & Newman 2010, 2011

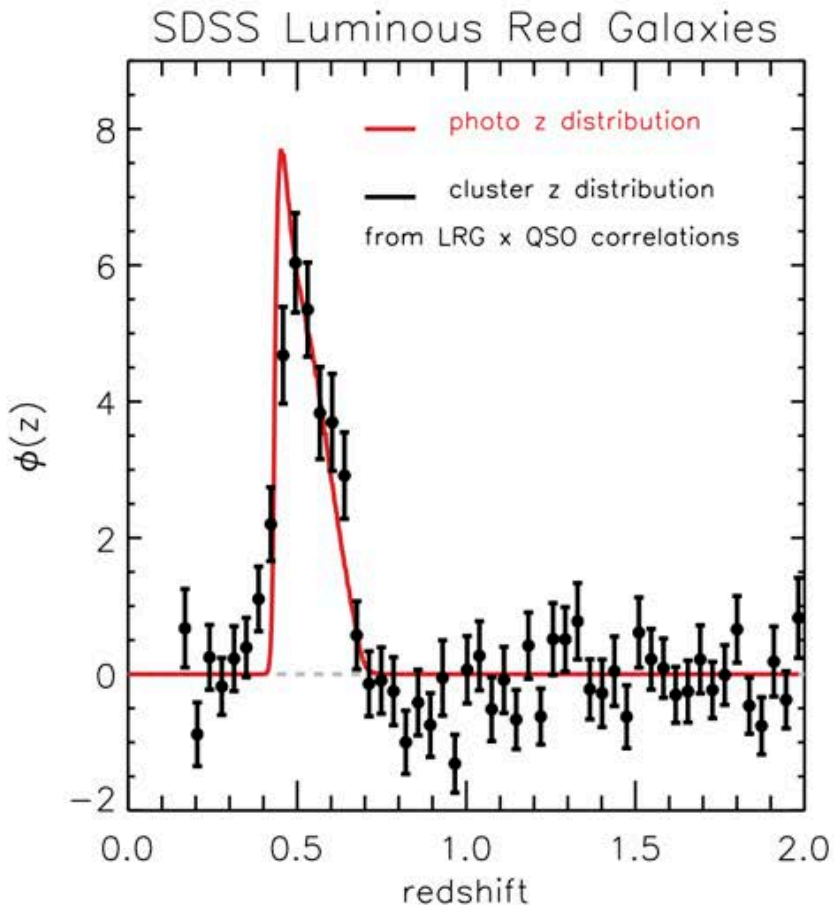
Blue: z_{phot} distribtion of objects with $0.7 < z_{phot} < 0.9$

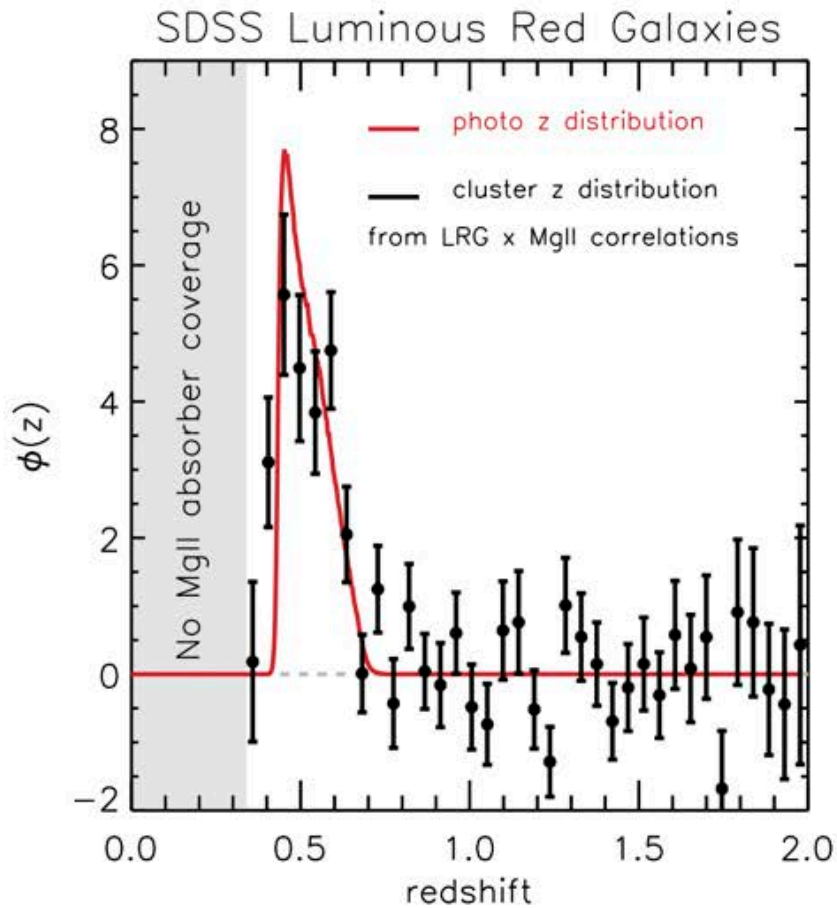
Black: True z distribution of sample, spanning 24 widely-separated fields

Red: Cross-correlation reconstruction with only a R<24, 4 deg^2 survey



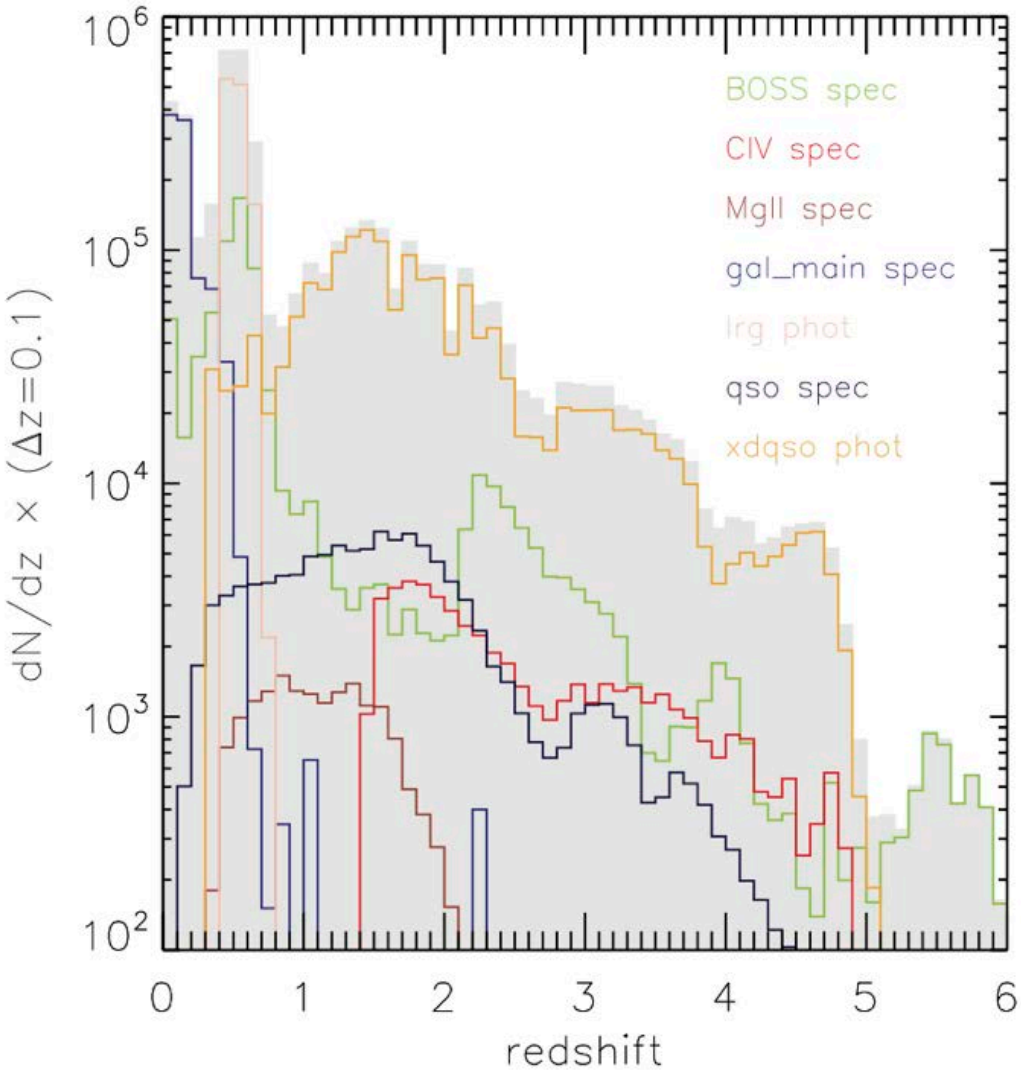
Absorption-line studies can contribute to cross-correlation samples





Menard et al. 2013

Absorption-line studies can contribute to cross-correlation samples



Menard et al. 2013