# Solving photometric redshift challenges with MSE

**Jeff Newman, U. Pittsburgh/PITT PACC**

# Two spectroscopic needs for photo-z work: training and calibration

- **Better training / optimization of algorithms using sets of objects with spectroscopic redshift measurements shrinks photo-z errors for individual objects, providing more detailed maps of large-scale structure and improved cosmology constraints, especially for BAO + cluster studies**

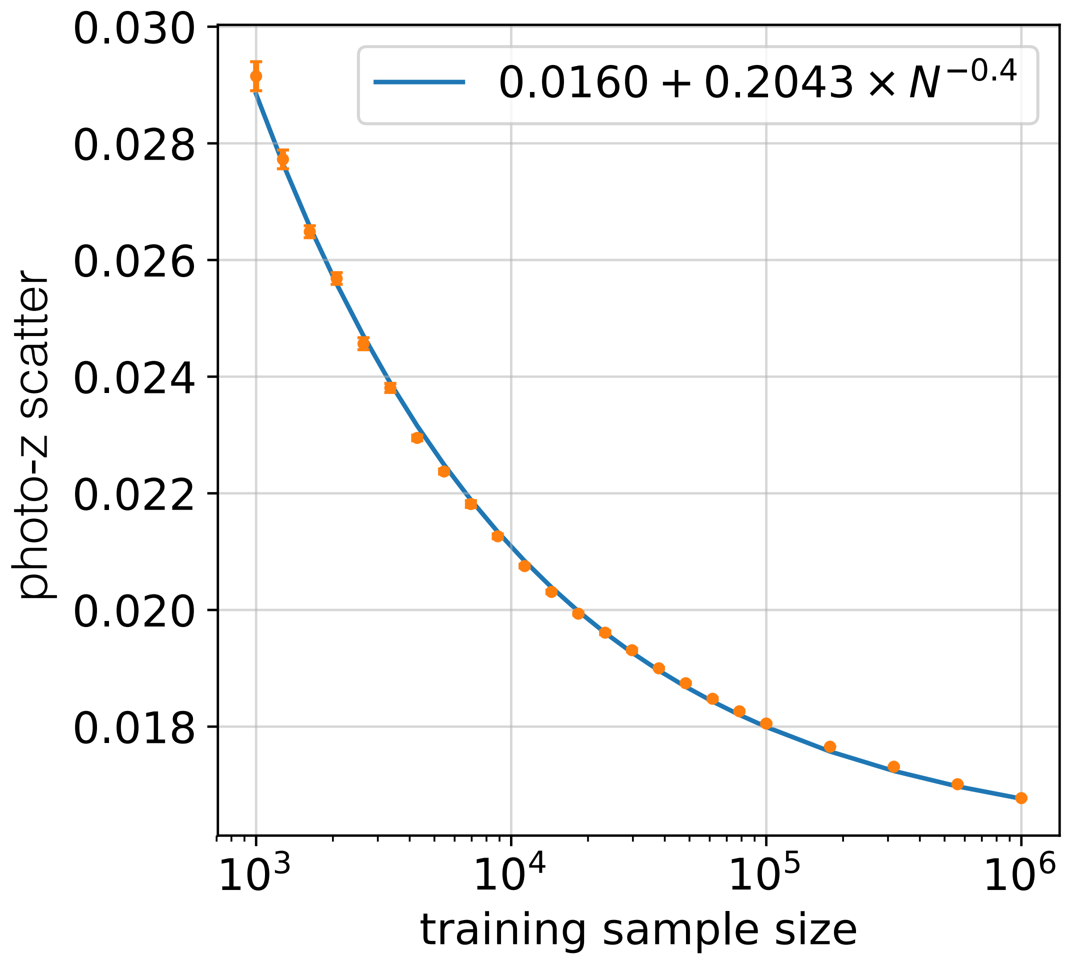
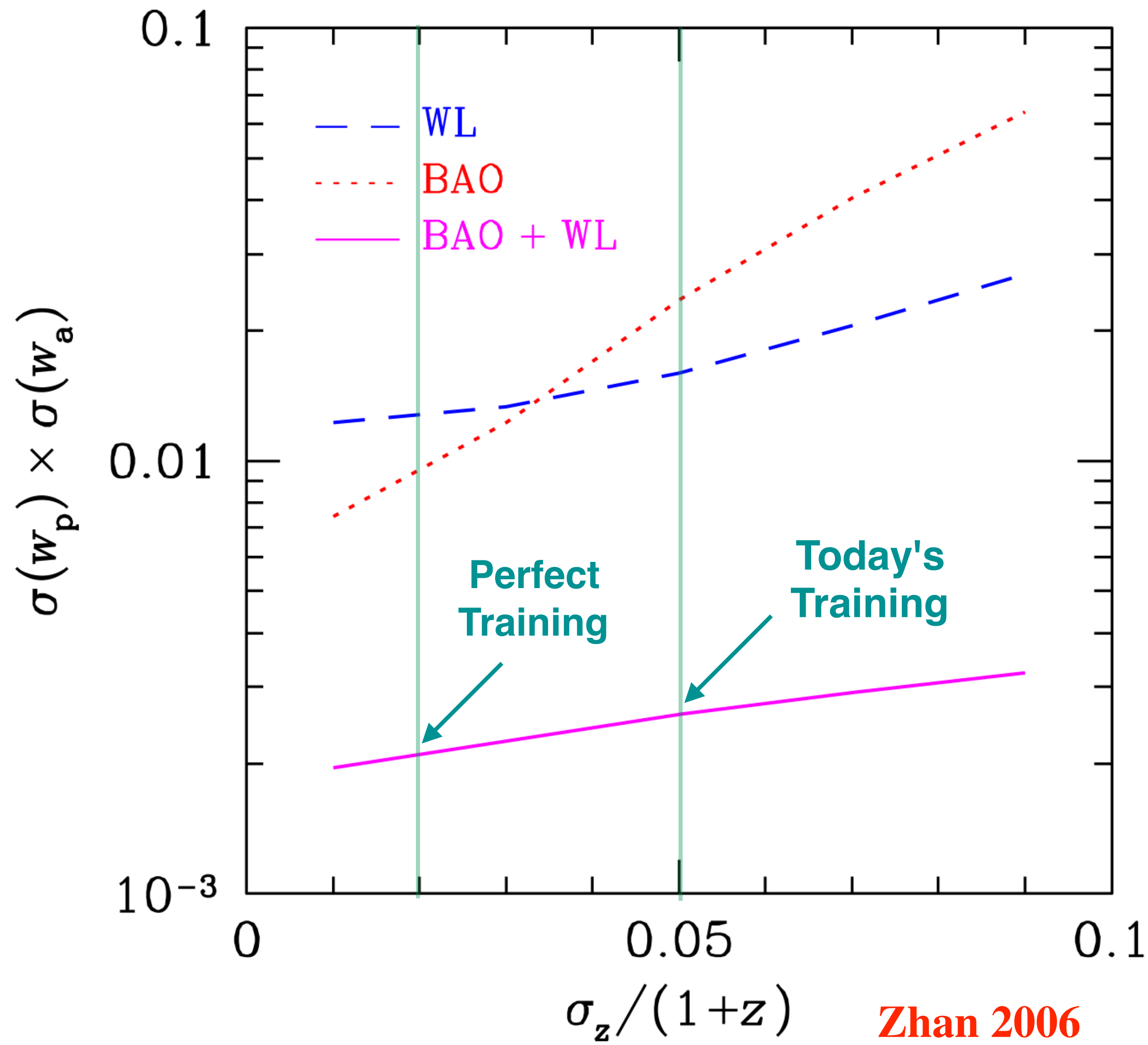



Figure: Rongpu Zhou

  - **Training datasets will contribute to calibration of photo-z's. ~Perfect training sets can solve calibration needs.**

# Two spectroscopic needs for photo-z work: training and calibration

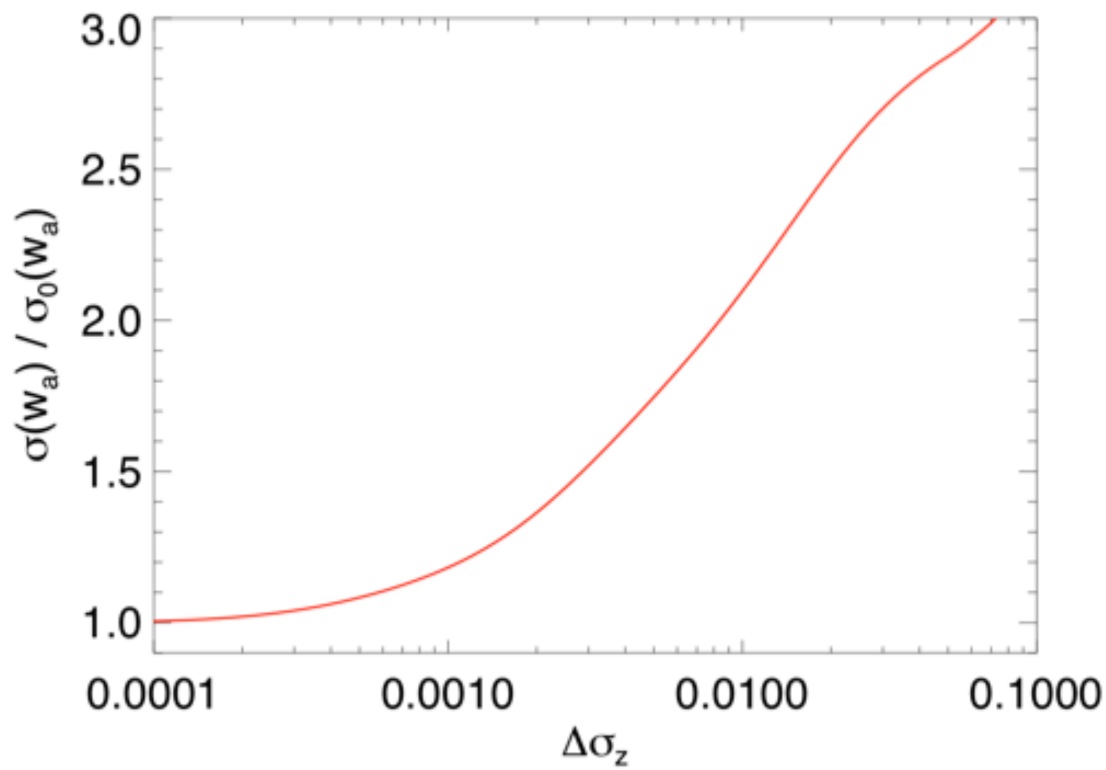
- Better training / optimization of algorithms using sets of objects with spectroscopic redshift measurements shrinks photo-z errors for individual objects, providing more detailed maps of large-scale structure and improved cosmology constraints, especially for BAO + cluster studies




  - Training datasets will contribute to calibration of photo-z's. ~Perfect training sets can solve calibration needs.

# Two spectroscopic needs for photo-z work: training and calibration

- For weak lensing and supernovae, individual-object photo-*z*'s do not need high precision, but the calibration must be accurate  - i.e., bias and errors need to be *extremely* well-understood or dark energy constraints will be off




Newman et al. 2015

  - *uncertainty in* bias, $\sigma(\delta_z) = \sigma(\langle z_p - z_s \rangle)$, and in scatter, $\sigma(\sigma_z) = \sigma(\mathrm{RMS}(z_p - z_s))$, must both be <~0.002(1+*z*) in each bin for Stage IV surveys.  Calibration may be done via cross-correlation methods using DESI/4MOST redshifts (Newman 2008)

## Estimated requirements for LSST photo-z training survey are well-matched to MSE (cf. Newman et al. 2015)

- Sensitive spectroscopy of ~20-30,000 faint objects (to $i$=25.3 for LSST)
  - Needs a combination of large aperture and long exposure times
- High multiplexing
  - Required to get large numbers of spectra
- Coverage of full ground-based spectral window
  - Minimum: 0.37-1 micron, 0.35-1.3 microns preferred
- Significant resolution ($R=\lambda/\Delta\lambda>\sim4000$) at red end
  - Allows secure redshifts from [OII] 3727 Å line at z>1
- Field diameters > ~20 arcmin
  - Need to span several correlation lengths for accurate clustering
- Many fields, >~15
  - To mitigate sample/cosmic variance
- Ideally, a Southern hemisphere site
  - To enable sampling across the LSST or WFIRST footprint

# MSE is extremely competitive for this work: almost as fast as the ELTs, but much cheaper to run

| Instrument / Telescope | Collecting Area (sq. m) | Field area (sq. arcmin) | Multiplex | Total time, Photometric Redshift Training (y) | Approx. cost per year |
|---|---|---|---|---|---|
| 4MOST | 10.75 | 14,400 | 1,400 | 5.4 | $3,900,000 |
| Mayall 4m / DESI | 11.40 | 25,500 | 5,000 | 5.1 | $4,200,000 |
| WHT / WEAVE | 13.00 | 11,300 | 1,000 | 6.0 | $4,700,000 |
| Subaru / PFS | 53.00 | 4,500 | 2,400 | 1.1 | $19,000,000 |
| VLT / MOONS | 58.20 | 500 | 500 | 2.7 | $21,000,000 |
| Keck / DEIMOS | 76.00 | 54 | 150 | 6.8 | $28,000,000 |
| Keck / FOBOS | 76.00 | 314 | 1,800 | 0.8 | $28,000,000 |
| ESO SpecTel | 87.89 | 17,676 | 3,333 | 0.7 | $32,000,000 |
| MSE | 97.59 | 6,359 | 3,249 | 0.6 | $36,000,000 |
| GMT/MANIFEST + GMACS | 368.00 | 314 | 420 | 0.5 | $130,000,000 |
| TMT / WFOS | 655.00 | 25 | 100 | 1.2 | $130,000,000 |
| E-ELT / Mosaic Optical | 978.00 | 39 | 200 | 0.5 | $240,000,000 |
| E-ELT / MOSAIC NIR | 978.00 | 46.00 | 100 | 0.8 | $240,000,000 |

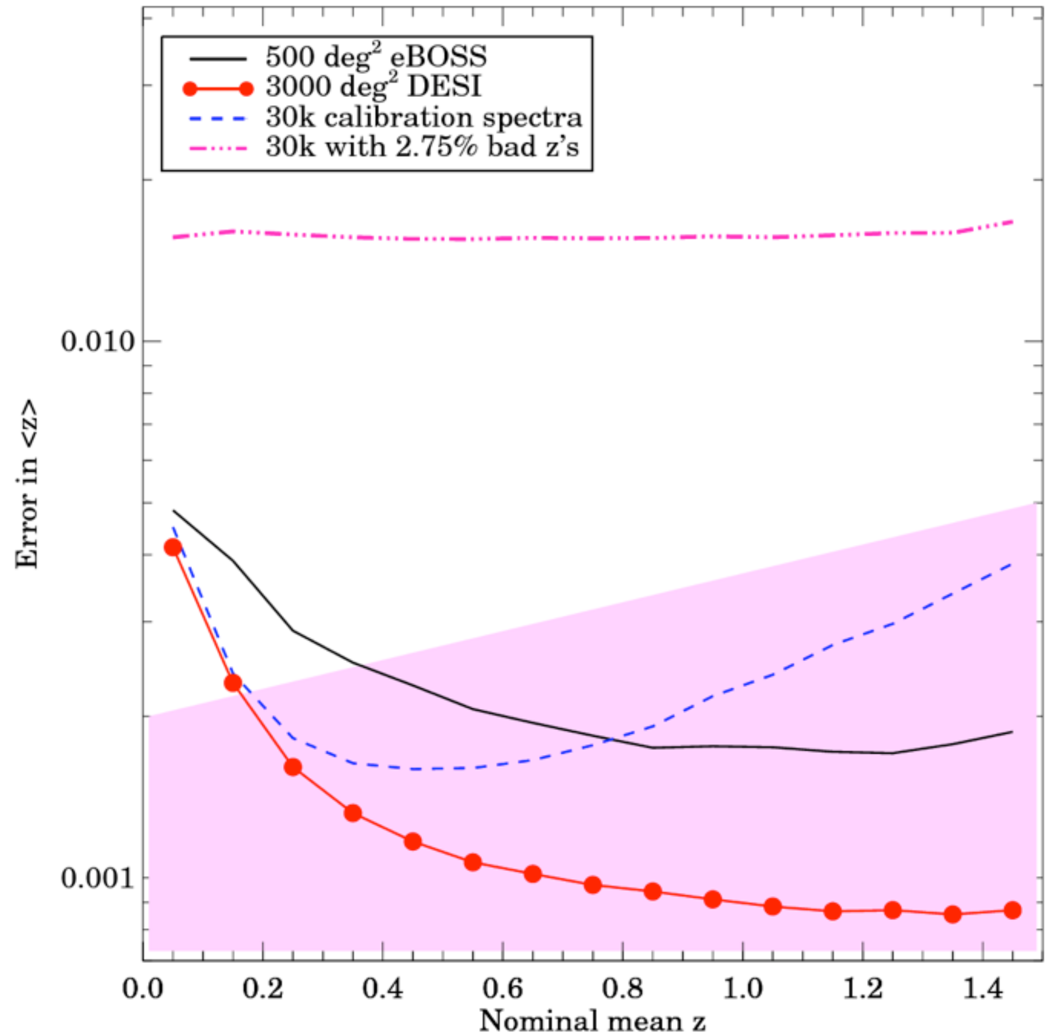
**Updated from Newman et al. 2015, *Spectroscopic Needs for Imaging Dark Energy Experiments***

# MSE for photo-*z* training

- ~0.6 dark-years on MSE for fiducial LSST training survey (>75% success rates, 15 fields, 30k spectra; includes weather losses)
- MSE would exceed requirements on area per field and sample size by a fair margin: better sample/cosmic variance than assumed in Newman et al.
  - Unless the 15 paintings are not widely-separated...
  - Almost certainly could get away with ~10 paintings given MSE FoV, cutting survey time by 1/3
- Could trade out fibers as secure redshifts are achieved, so actual sample could include extra bright objects (or have better S/N for galaxy evolution studies)
- Highly synergistic with surveys intended to study galaxy evolution: the photo-z training survey would determine the range of galaxy SEDs as a function of redshift, and how SED relates to local environment / overdensity

# MSE for photo-*z* calibration

- If spectroscopic samples remain incomplete or redshifts are not highly secure, best hope for calibration is cross-correlation methods (Newman 2008); a.k.a. "clustering redshifts"
- For LSST, easiest to lower cross-correlation errors by using dilute samples (e.g., DESI QSOs) over wide areas. However, DESI is very sparse at $1.6 < z < 2.2$, where QSOs are the only tracer available
- For WFIRST, optimal cross-correlation survey would cover as many z=0-4 galaxies as possible over full survey footprint.
- The proposed MSE cosmology survey could be very useful for this cross-correlation work




Newman et al. 2015

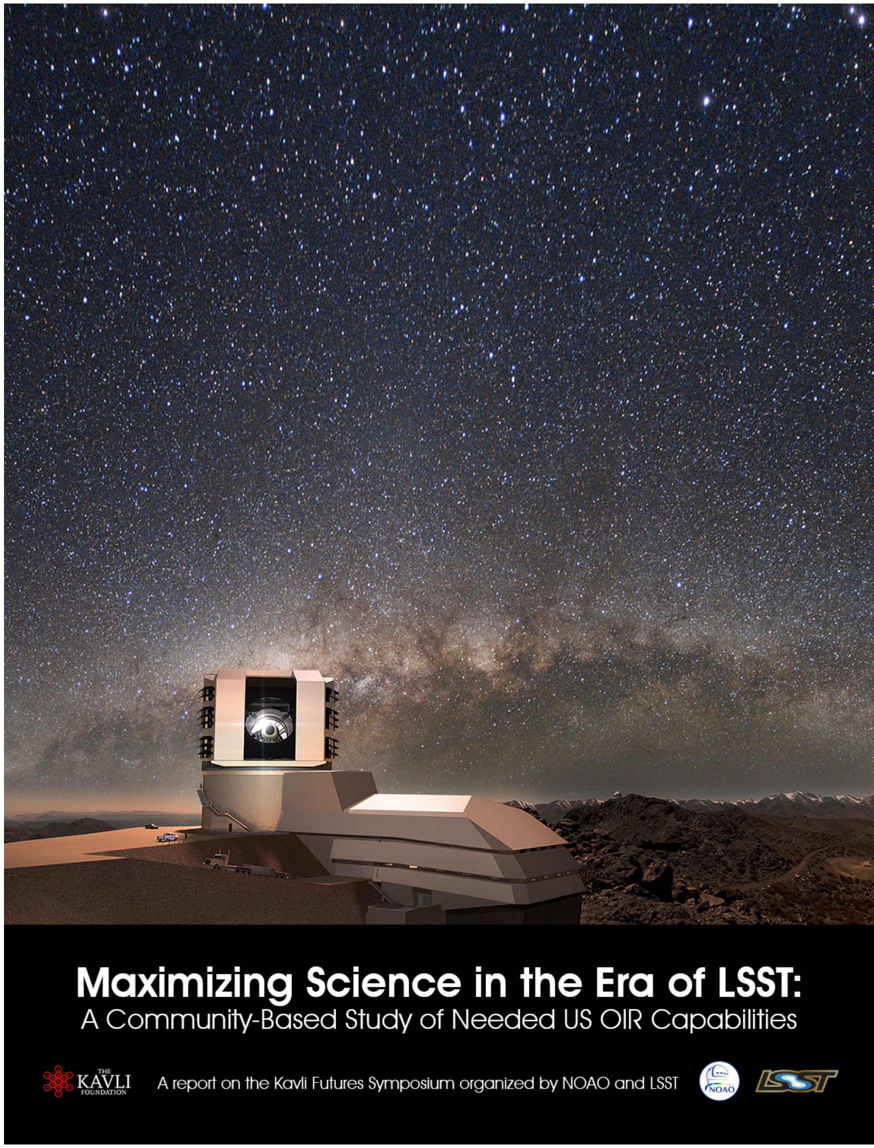
# Conclusions

- MSE can make major contributions to cosmology through photo-z training survey
- Baseline photo-z training survey, >~75% complete:
  - 15 (or 10?) widely-separated pointings, DEEP2 S/N at $i$=25.3

    = 150-220 dark nights for >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. Synergistic with proposed MSE galaxy evolution studies, if include enough objects in between the target redshift regimes.
- MSE is an extremely efficient option for this work.
- See the Newman et al. *Spectroscopic Needs* white paper for more: http://adsabs.harvard.edu/abs/2015APh....63...81N

# Assumptions made for time calculations

- This is extracted from an attempt to take the largest surveys proposed in the Kavli/NOAO/LSST report (Najita, Willman et al. 2016) and work out how long would be needed to do them
- Common set of assumptions: one-third loss to instrumental effects, weather and overheads; 4m = Mayall/DESI; 8m = Subaru/PFS; all instrumental efficiencies identical; equivalent # of photons will yield equal noise; ignoring differences in seeing/ image quality and fiber/slitlet size. Only medium-resolution fibers included. Assuming full spectral range can be covered simultaneously (likely not true for E-ELT).
- See report (available at http://arxiv.org/abs/1610.01661 ) for details of these surveys
- Estimating time in years on each platform; note that this is generally dark time (very faint targets!)
- Costs based on TSIP + inflation: $1k/m²/night



# Two spectroscopic needs for photo-z work: training and calibration

- **Better training / optimization of algorithms using sets of objects with spectroscopic redshift measurements shrinks photo-z errors for individual objects, providing more detailed maps of large-scale structure and improved cosmology constraints, especially for BAO + cluster studies**

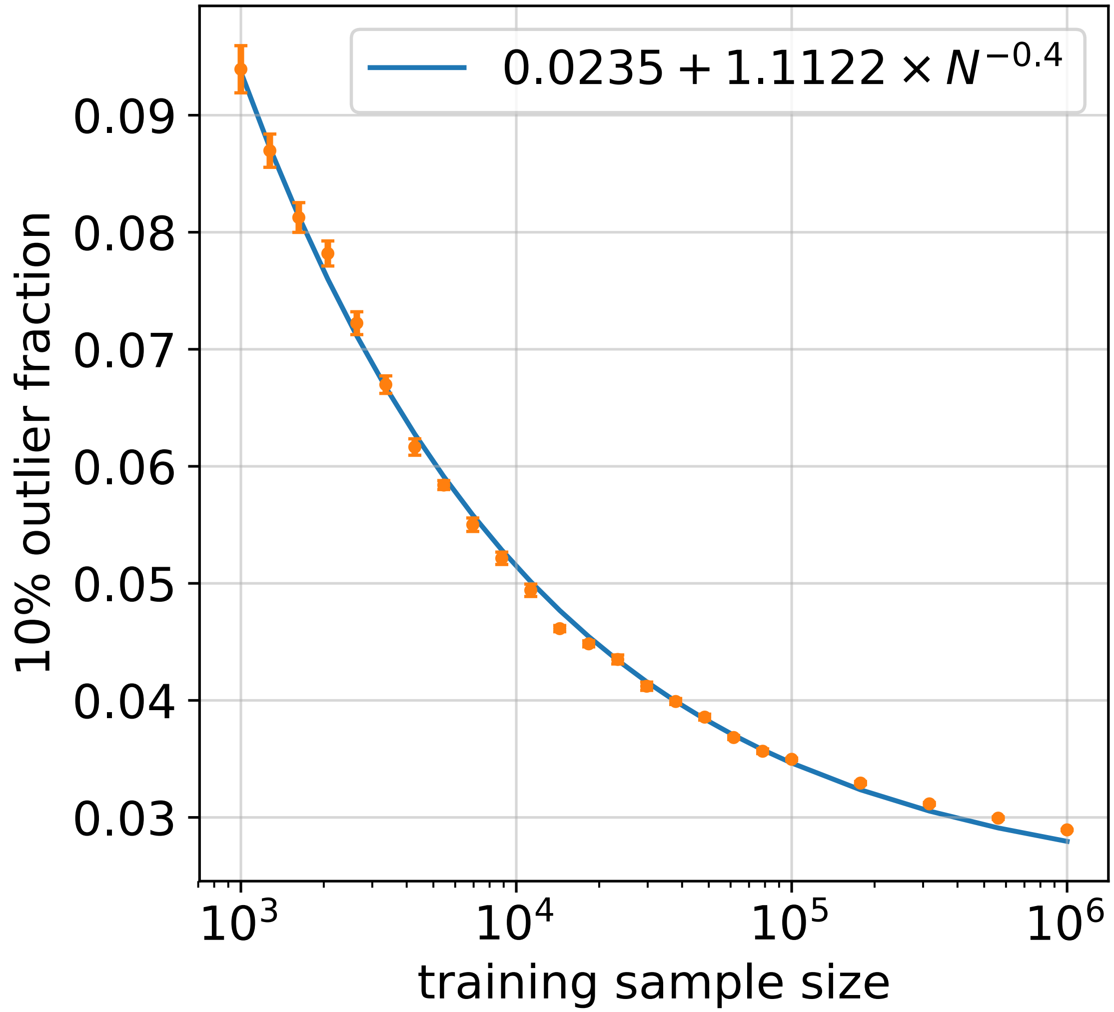
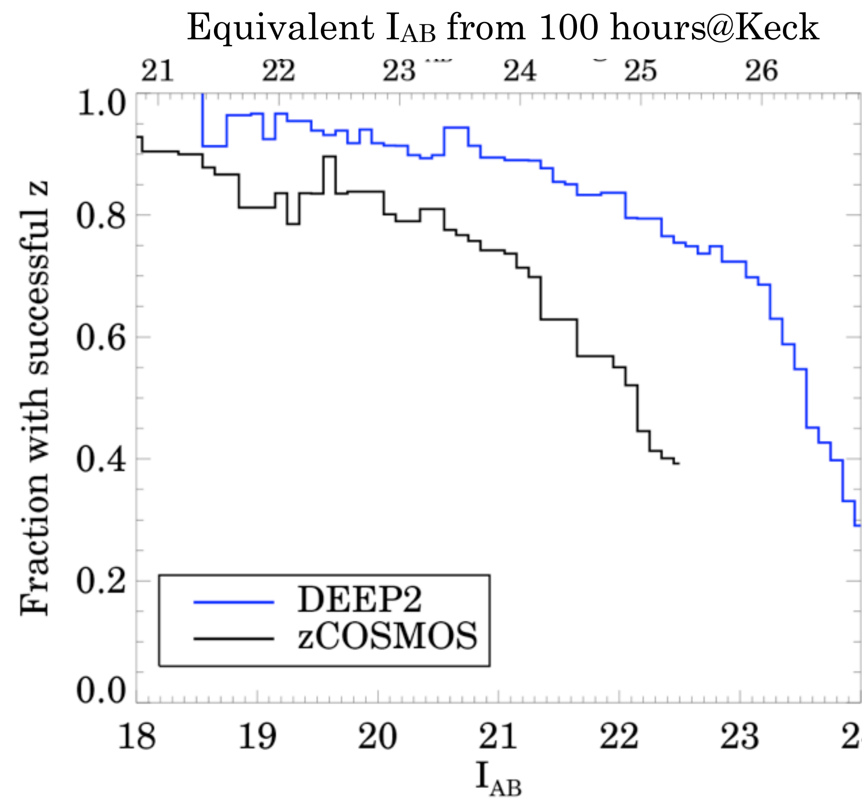



Figure: Rongpu Zhou

  - **Training datasets will contribute to calibration of photo-z's. ~Perfect training sets can solve calibration needs.**

# A few notes

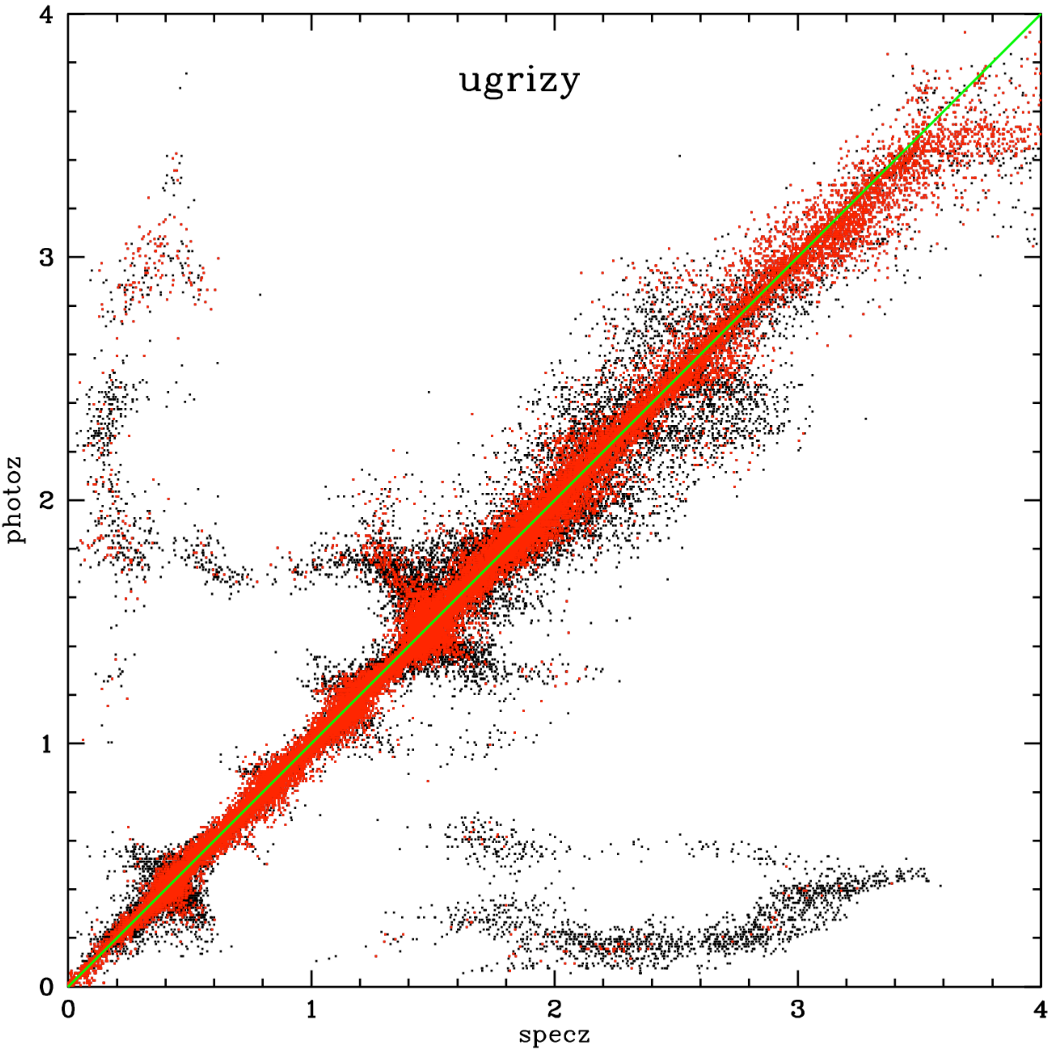
- Basic assumption for exposure time calculations: with comparable resolution and greater wavelength coverage than DEEP2, redshift success at the same continuum S/N should be no lower than DEEP2's (partially because [OII] EW distribution shows little evolution, so [OII] S/N $\propto$ continuum S/N)
- Difficult to make guarantees about how high success rates will be: many failures are "unknown unknowns", especially for IR-selected samples (as WFIRST). DEEP2-based predictions are for success vs. *i* magnitude (and hence optical S/N).



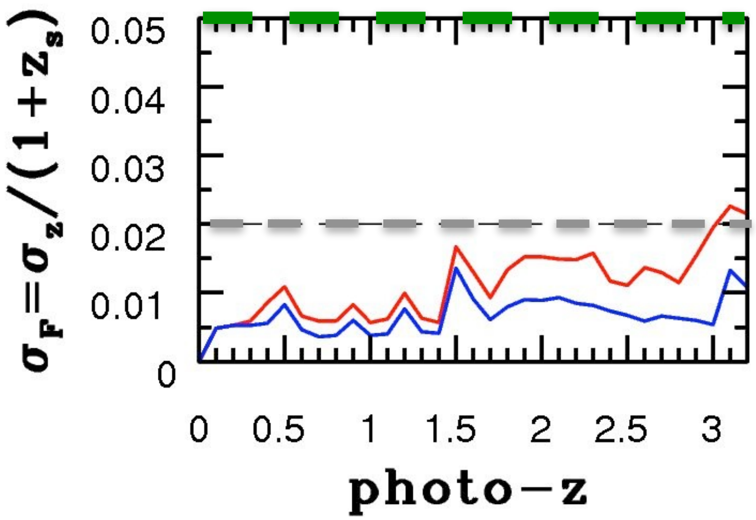

Newman et al. 2015

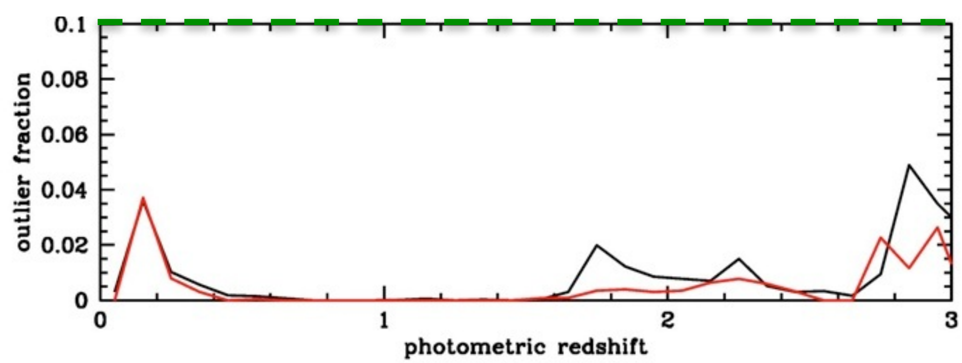
- Based on theory papers, need >99% completeness over the full color/magnitude range used for analyses to keep biases subdominant for Stage IV (it takes a large number of independent spectra to demonstrate that is reached!). If that <u>is</u> achieved, training samples provide calibration too.

# Potential photo-z performance for LSST *ugrizy*




**Green**: Requirements on actual performance; **grey**: requirements on performance **<u>with perfect template knowledge</u>** (as in these sims)



## Spectroscopic training set requirements

- Goal: make $\delta_z$ 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 ELTs, 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 equivalent per target, minimum**

- **High multiplexing**

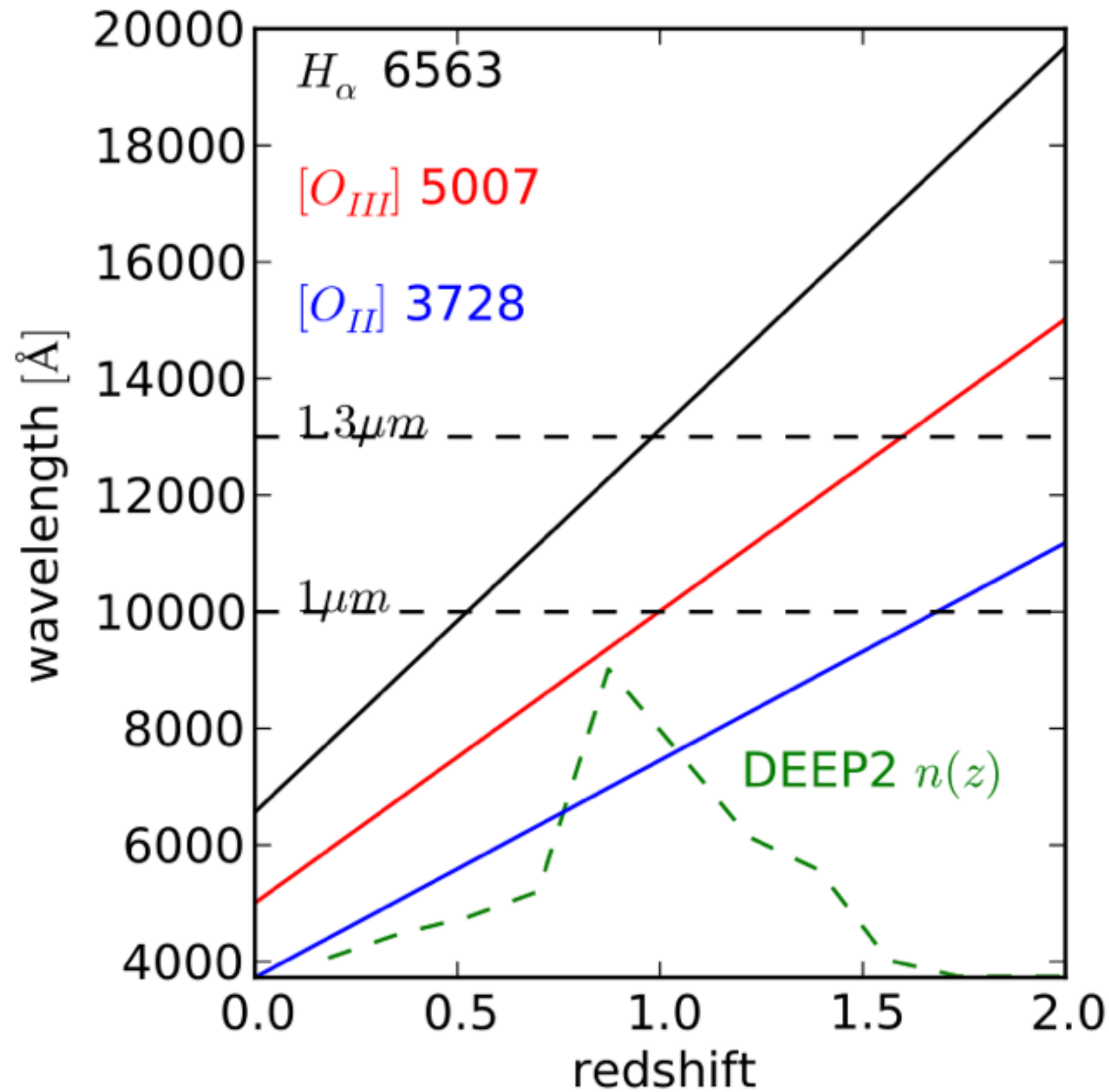
**- Obtaining large numbers of spectra is infeasible without it**

# 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**

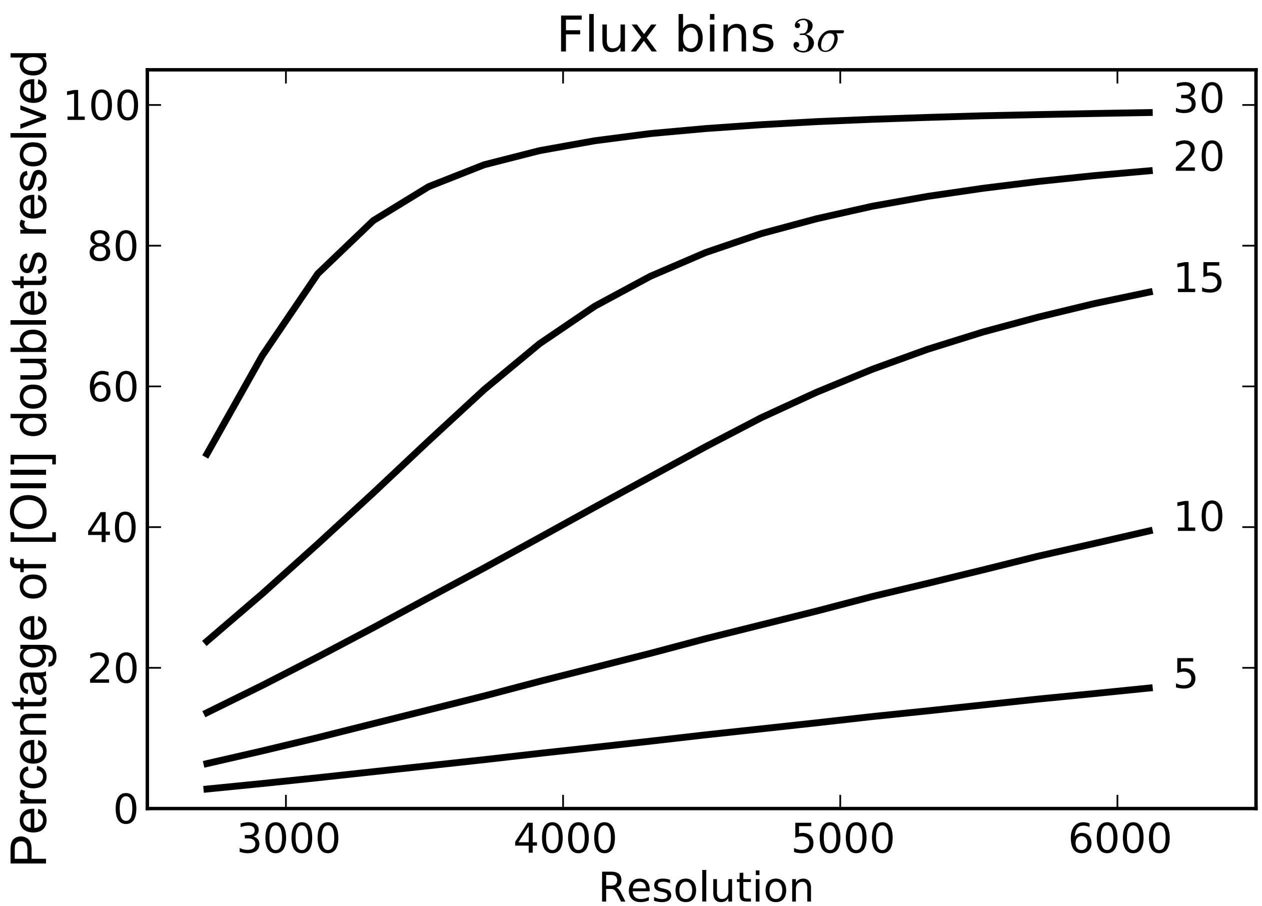



**Comparat et al. 2013, submitted**

# 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

# What qualities do we desire in our training sets?

- **Field diameters > ~20 arcmin**

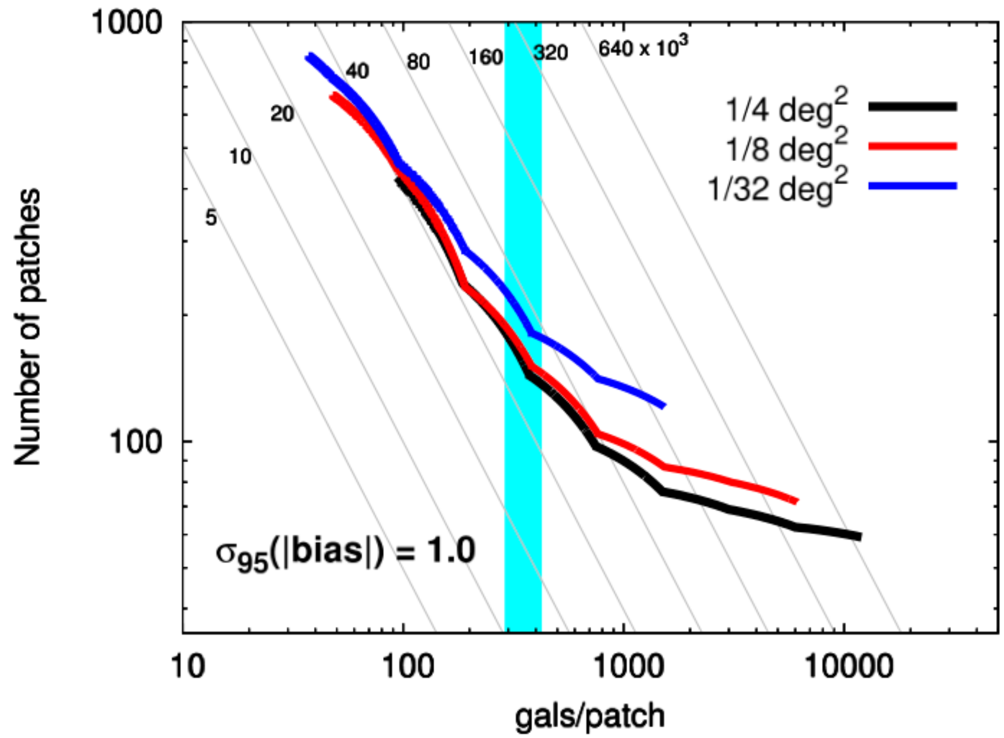
  - Need to span several correlation lengths for accurate clustering measurements (key for galaxy evolution science and cross-correlation techniques)

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

- **Many fields**

  - Minimizes impact of sample/ cosmic variance.

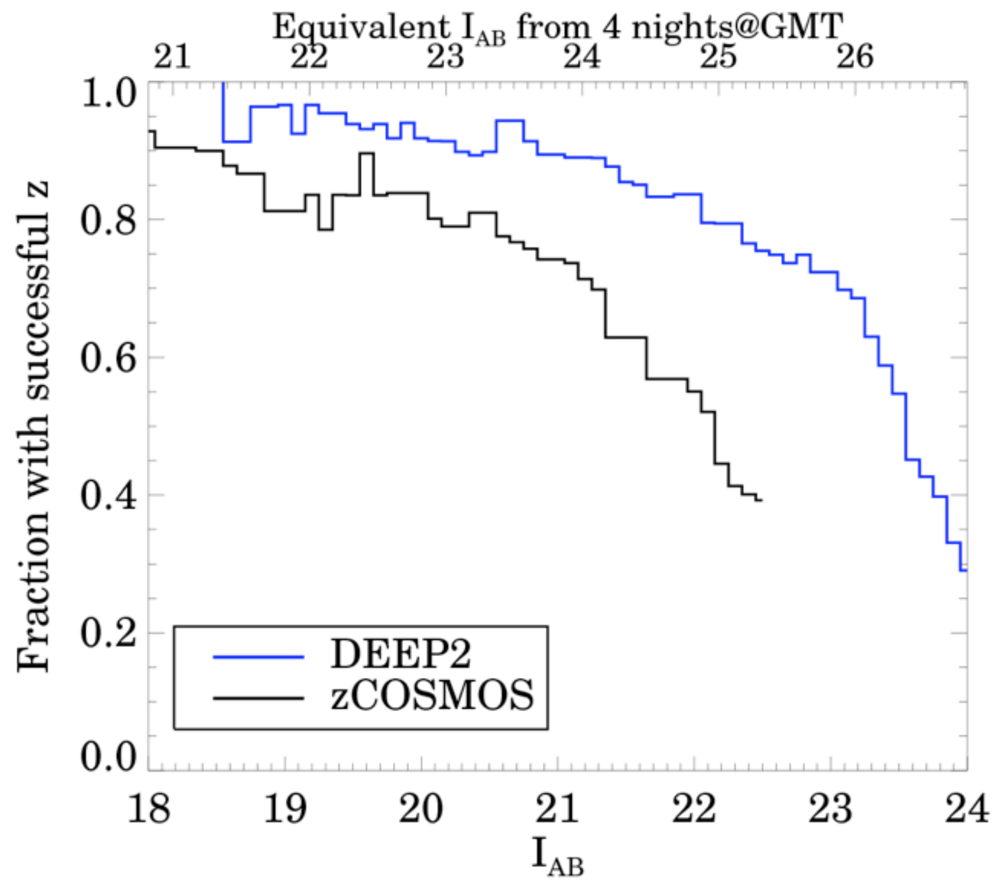
    - e.g., Cunha et al. (2012) estimate that 40-150 ~0.1 deg² fields are needed for DES for sample variance not to impact errors (unless we get clever)




Cunha et al. 2012

## Biggest obstacle: incompleteness in training sets

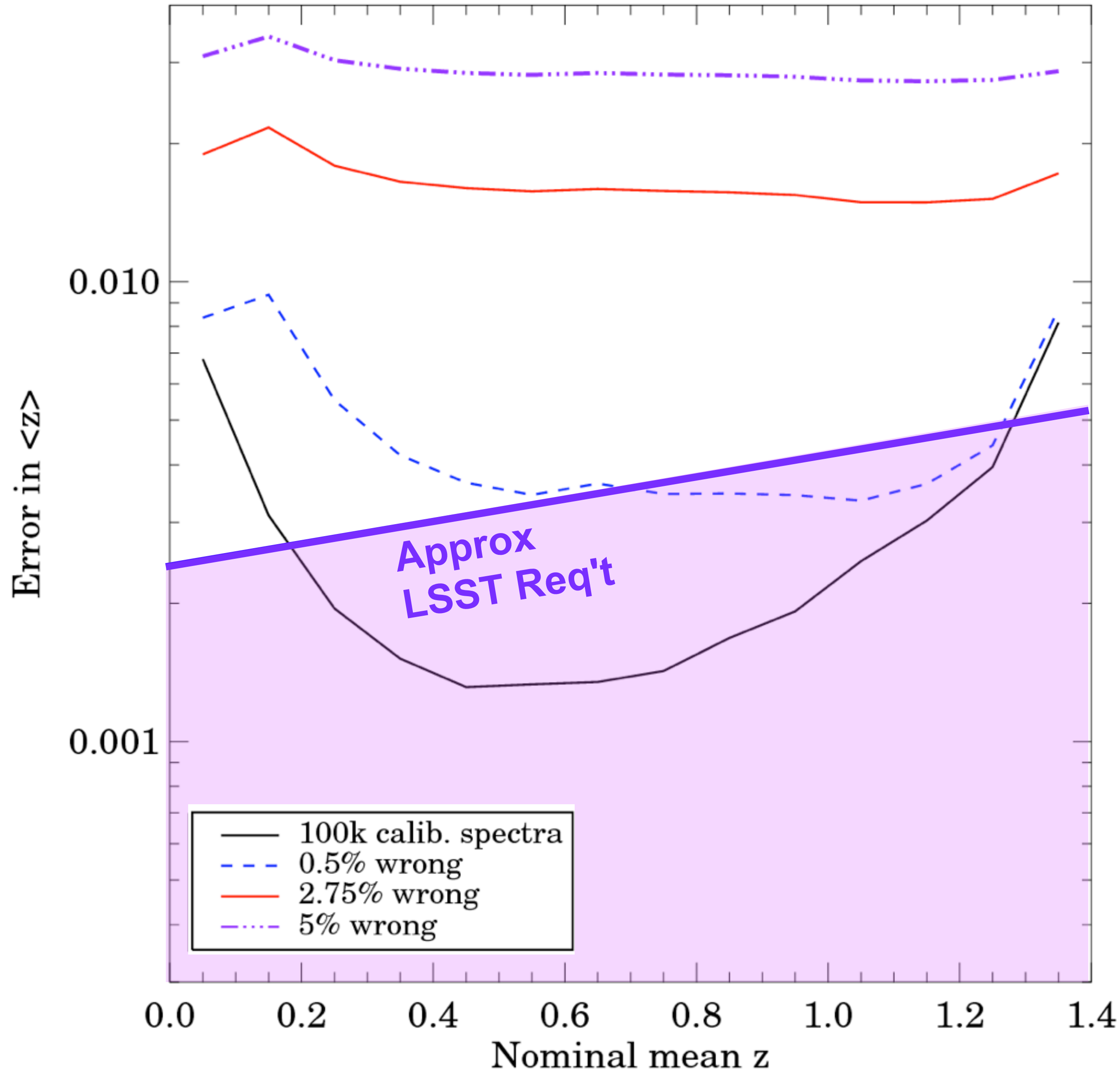
- In current deep redshift surveys (to i~22.5/R~24), 30-60% of targets fail to yield secure (>95% confidence) redshifts
- Losses are worst at the faint end
- Redshift success rate varies with galaxy color, redshift, etc.
- In DEEP2, best parts of *BRI* color space have ~90% redshift success
- 4 night GMT=15 night MSE depth would yield >~75% completeness; achieving >90% would require ~25 nights/pointing on GMT




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

## Note: even for 100% complete samples, current false-z rates would be a problem

- Only the highest-confidence redshifts should be useful for precision calibration: lowers spectroscopic completeness further when restrict to only the best
- A major reason why splitting [OII] is important




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

# Amount of time required for each survey from the Kavli/NOAO/LSST report (sorted by telescope aperture; in dark-years). Leader for each column shown in bold.

| Instrument / Telescope | Total time, Photometric Redshift Training (y) | Milky Way halo survey (8000 sq. deg., y) | Local dwarfs and halo streams | Galaxy evolution | Supernova hosts | Total (8000 sq. deg. halo survey, y) |
|---|---|---|---|---|---|---|
| 4MOST | 5.4 | 12.6 | 10.1 | 4.2 | 0.05 | 32.4 |
| Mayall 4m / DESI | 5.1 | 6.7 | 9.5 | 1.1 | 0.03 | 22.5 |
| WHT / WEAVE | 6.0 | 13.3 | 8.3 | 4.9 | 0.06 | 32.5 |
| Subaru / PFS | 1.1 | 8.2 | 2.0 | 0.5 | 0.04 | 11.9 |
| VLT / MOONS | 2.7 | 67.0 | 1.9 | 2.2 | 0.29 | 74.0 |
| Keck / DEIMOS | 6.8 | 473.1 | 8.3 | 5.6 | 2.04 | 495.7 |
| Keck / FOBOS | 0.8 | 81.7 | 1.4 | 0.5 | 0.35 | 84.7 |
| ESO SpecTel | 0.7 | **1.3** | 1.2 | 0.2 | **0.01** | **3.4** |
| MSE | 0.6 | 3.1 | 1.1 | 0.2 | **0.01** | 5.1 |
| GMT/MANIFEST + GMACS | **0.5** | 16.9 | **0.3** | 0.4 | 0.07 | 18.2 |
| TMT / WFOS | 1.2 | 119.6 | 2.1 | 1.0 | 0.51 | 124.3 |
| E-ELT / Mosaic Optical | 0.5 | 51.8 | 0.9 | 0.3 | 0.22 | 53.7 |
| E-ELT / MOSAIC NIR | 0.8 | 43.4 | 0.8 | 0.6 | 0.19 | 45.7 |

## Total time required for <u>all</u> surveys from the Kavli/NOAO/LSST report (sorted by telescope aperture; in dark-years). Leader for each column shown in bold.

| Instrument / Telescope | Total (no halo survey, dark-years) | Total (8000 sq. deg. halo survey, dark-years) | Total (20k sq. deg. halo survey, dark-years) | Approx. cost per year |
|---|---|---|---|---|
| 4MOST | 19.8 | 32.4 | 51.3 | $3,900,000 |
| Mayall 4m / DESI | 15.8 | 22.5 | 32.5 | $4,200,000 |
| WHT / WEAVE | 19.2 | 32.5 | 52.4 | $4,700,000 |
| Subaru / PFS | 3.7 | 11.9 | 24.1 | $19,000,000 |
| VLT / MOONS | 7.0 | 74.0 | 174.6 | $21,000,000 |
| Keck / DEIMOS | 22.7 | 495.7 | 1205.4 | $28,000,000 |
| Keck / FOBOS | 3.0 | 84.7 | 207.3 | $28,000,000 |
| ESO SpecTel | 2.1 | **3.4** | **5.3** | $32,000,000 |
| MSE | 1.9 | 5.1 | 9.8 | $36,000,000 |
| GMT/MANIFEST + GMACS | **1.3** | 18.2 | 43.5 | $130,000,000 |
| TMT / WFOS | 4.8 | 124.3 | 303.7 | $130,000,000 |
| E-ELT / Mosaic Optical | 1.9 | 53.7 | 131.4 | $240,000,000 |
| E-ELT / MOSAIC NIR | 1.6 | 44.9 | 109.9 | $240,000,000 |

# Brief descriptions of the Kavli/NOAO/LSST surveys

- Photometric redshift training sample: Minimum of 30,000 galaxies total down to $i$=25.3 in 15 fields >20' diameter
  - 100 hours/pointing on 10m
  - To improve photo-z accuracy for LSST (and study galaxy SED evolution)
  - Highly-complete survey would require ~6x greater exposure time than used here

- Milky Way halo survey: ~125 $g$<23 luminous red giants deg$^{-2}$ over 8,000 (or preferably 20,000) square degrees of sky
  - 2.5 hours/pointing with 8m
  - Allows reconstruction of MW accretion history using stars to the outer limits of the stellar halo. Other objects could be targeted on remaining fibers.

# Brief descriptions of the Kavli/NOAO/LSST surveys

- Local dwarfs and halo streams: Local dwarfs were estimated to require 3200 hours on an 8m to measure velocity dispersions of LSST-discovered dwarfs within 300 kpc
  - Requires FoV ≥ 20 arcmin (1 deg preferred) and minimum slit/fiber spacing < 10 arcsec.
  - Characterizing ~10 halo streams to test for gravitational perturbations by low-mass dark matter halos was estimated to require ~25% as much time on similar instrumentation.

# Brief descriptions of the Kavli/NOAO/LSST surveys

- Galaxy evolution survey: Minimum of 130,000 galaxies total down to $M=10^{10}\ M_{Sun}$ at $0.5 < z < 2$ over a 4 sq. deg. field
  - 18 hours per pointing on 8m
  - To study relationship between galaxy properties and environment across cosmic time

- Supernova host survey: Annual spectroscopy of ~100 new galaxy hosts of supernovae $deg^{-2}$ with $r$<24 over the ~5 LSST deep drilling fields (10 sq. deg. each)
  - ~8 hours per pointing on 4m
  - Provides redshifts for most of the ~50,000 best-characterized LSST SN Ia (other transients/hosts could be observed on remaining fibers)