Photometric redshifts: The long road ahead

Carlos Cunha
Stanford University

Berkeley Cosmology Seminar, February 25, 2014

W_o=-0.95±0.01

W_a=-0.5±0.1

New physics!

Do you believe the result? How about the error bar?

From the 2nd-ever photo-z-only meeting (Taipei, Sept. 2013)

Need redshifts

- **Spectroscopic** or **photometric** redshifts (photo-zs).

Accurate but expensive

Innaccurate but cheap

- For large surveys such as DES, PanSTARRS and LSST, photo-zs are the only option (besides cross-correlation techniques).

Conclusion

- Mistrust of photo-zs implies enormous costs to verify that photo-zs are okay.

Conclusion

- Mistrust of photo-zs implies enormous costs to verify that photo-zs are okay.

- We cannot trust photo-zs because we don’t trust our understanding of galaxy populations and distributions.

Conclusion

- Mistrust of photo-zs implies enormous costs to verify that photo-zs are okay.

- We cannot trust photo-zs because we don't trust our understanding of galaxy populations and distributions.

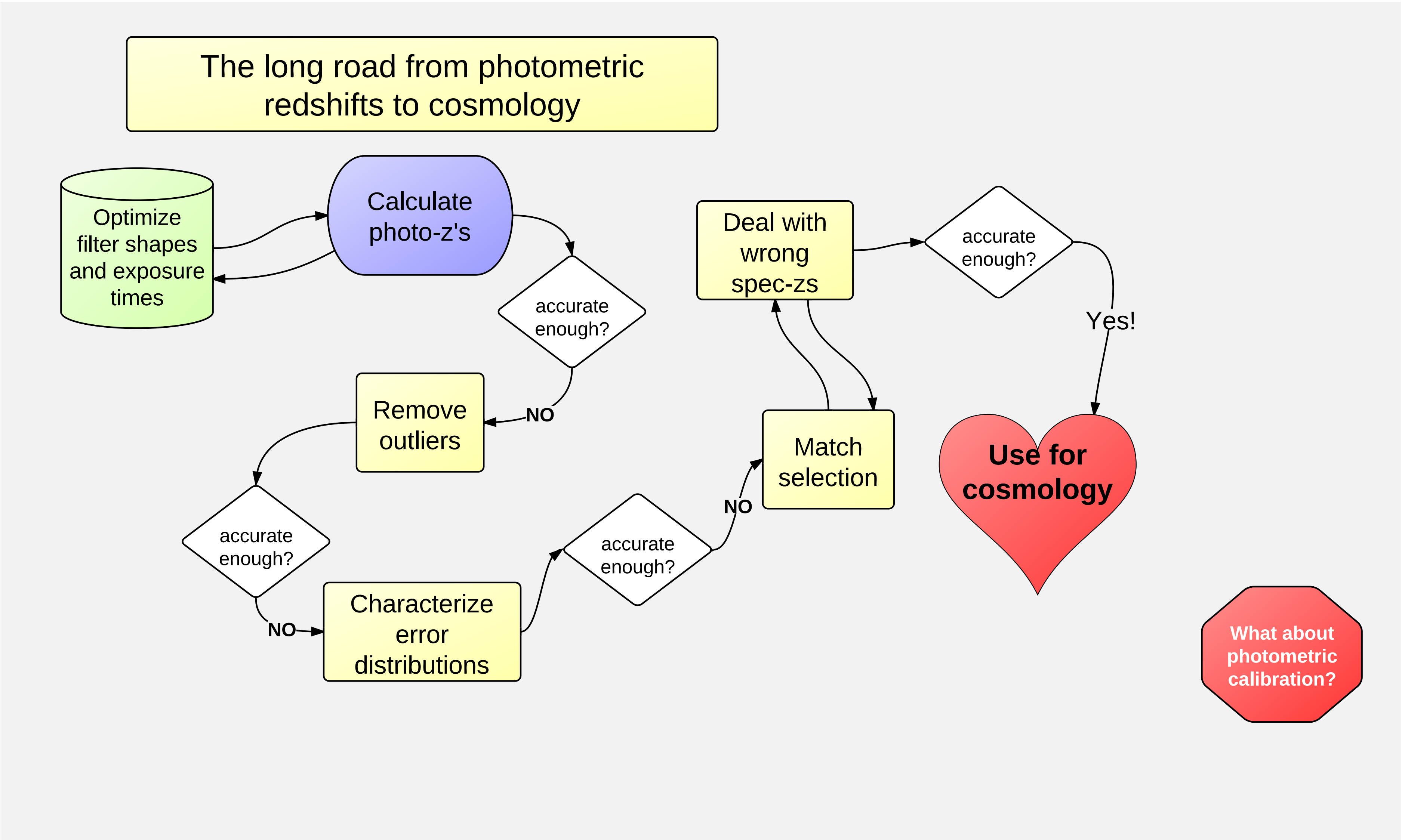
- But galaxy formation studies require redshifts.

Conclusion

- Mistrust of photo-zs implies enormous costs to verify that photo-zs are okay.
- We cannot trust photo-zs because we don't trust our understanding of galaxy populations and distributions.
- But galaxy formation studies require redshifts.
- Both have to be done simultaneously.

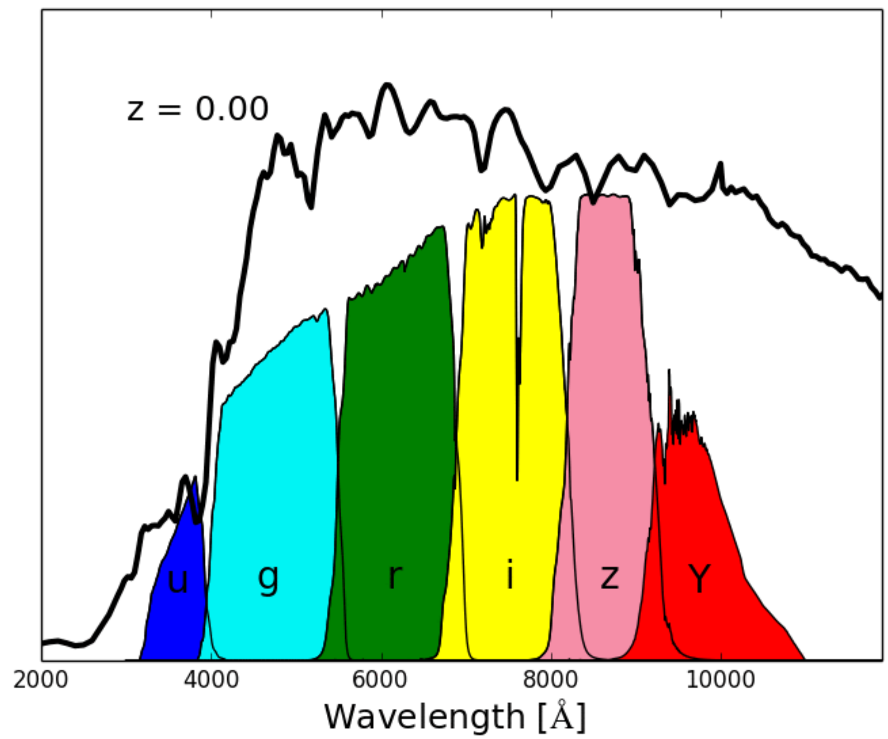
Conclusion

- Mistrust of photo-zs implies enormous costs to verify that photo-zs are okay.
- We cannot trust photo-zs because we don't trust our understanding of galaxy populations and distributions.
- But galaxy formation studies require redshifts.
- Both have to be done simultaneously.
- Simulations are the best framework with which to assess our state of knowledge (and I'll only trust cosmological results from LSST when we can produce a photometric simulation that closely resembles observations).


The long road from photometric redshifts to cosmology
Optimize filter shapes and exposure times
Calculate photo-z's
accurate enough?
NO
Remove outliers
accurate enough?
NO
Characterize error distributions
accurate enough?
NO
Match selection
Deal with wrong spec-zs
accurate enough?
Yes!
Use for cosmology
What about photometric calibration?

Basics of photo-zs

- Probe strong spectral features (4000 Å break)
- Flux in each filter depends on galaxy's type and redshift.



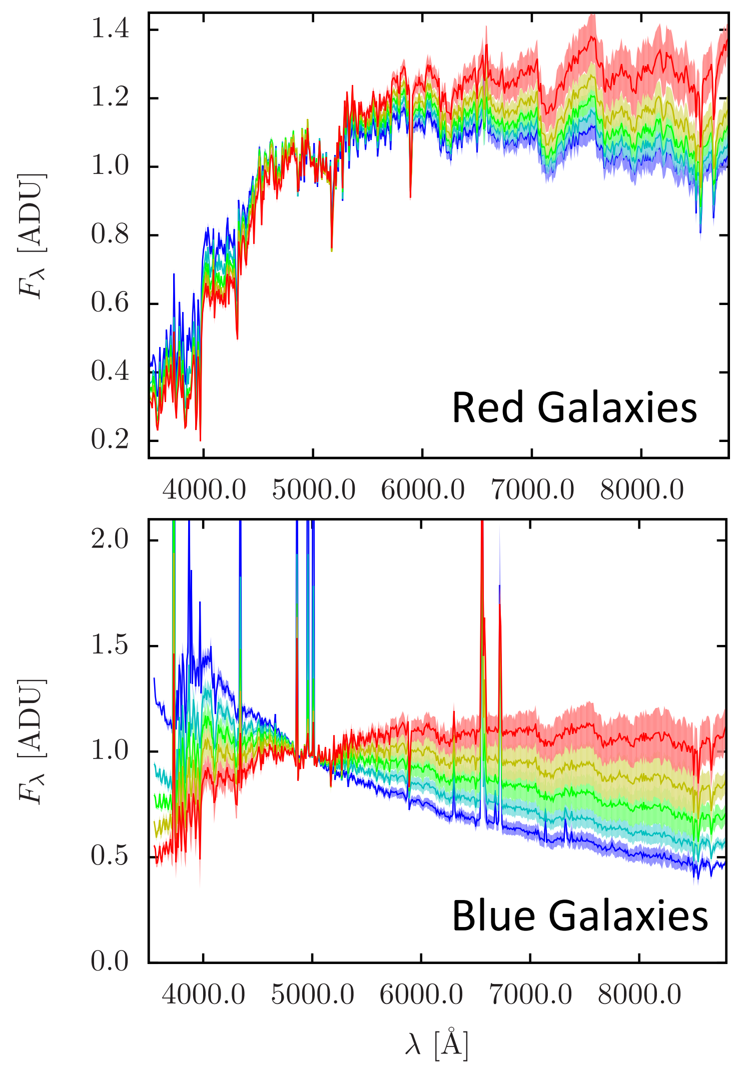
Terminology:

magnitude = A – log(flux)
color = magnitude - magnitude

Basics of photo-zs

- A difficulty is that there is a distribution of spectral types.

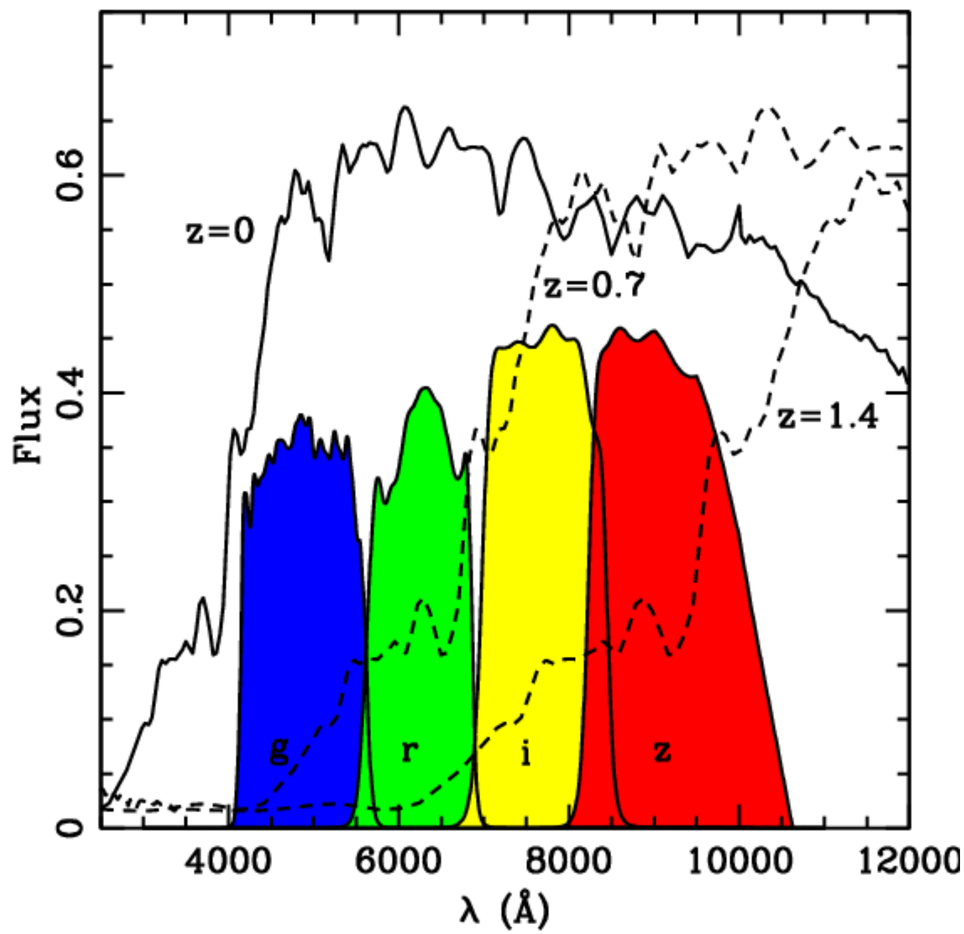
- And you have to separate galaxies from stars and QSOs (and from other galaxies – deblending).



Basics of photo-z's

Two classes of methods:

- **Template-fitting:** compare observed fluxes with predicted fluxes from library of galaxy spectra.
- **Training set:** use subsample with known redshifts to "train" flux-redshift relation.



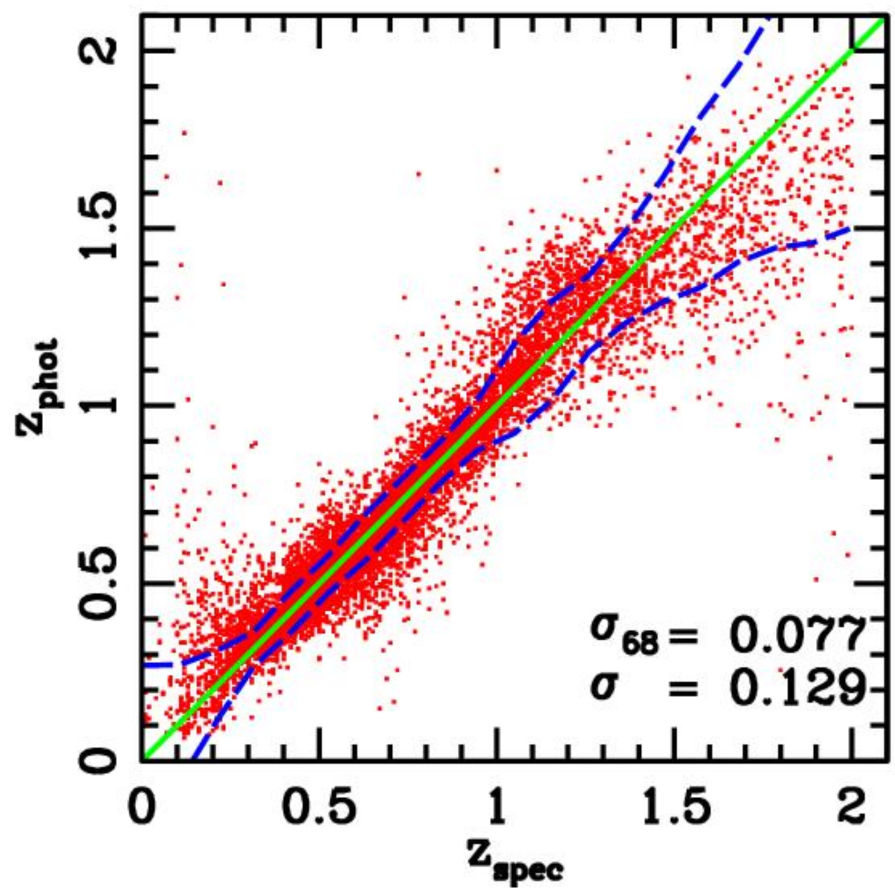
Courtesy M. Lima

Basics of photo-z's

Photo-zs are often not very good.

Three steps before getting to the cosmology:

- **Get photo-zs;**
- **Estimate photo-z errors and cull outliers;**
- **Calibrate error distribution, e.g. $P(z_s|z_p)$.**

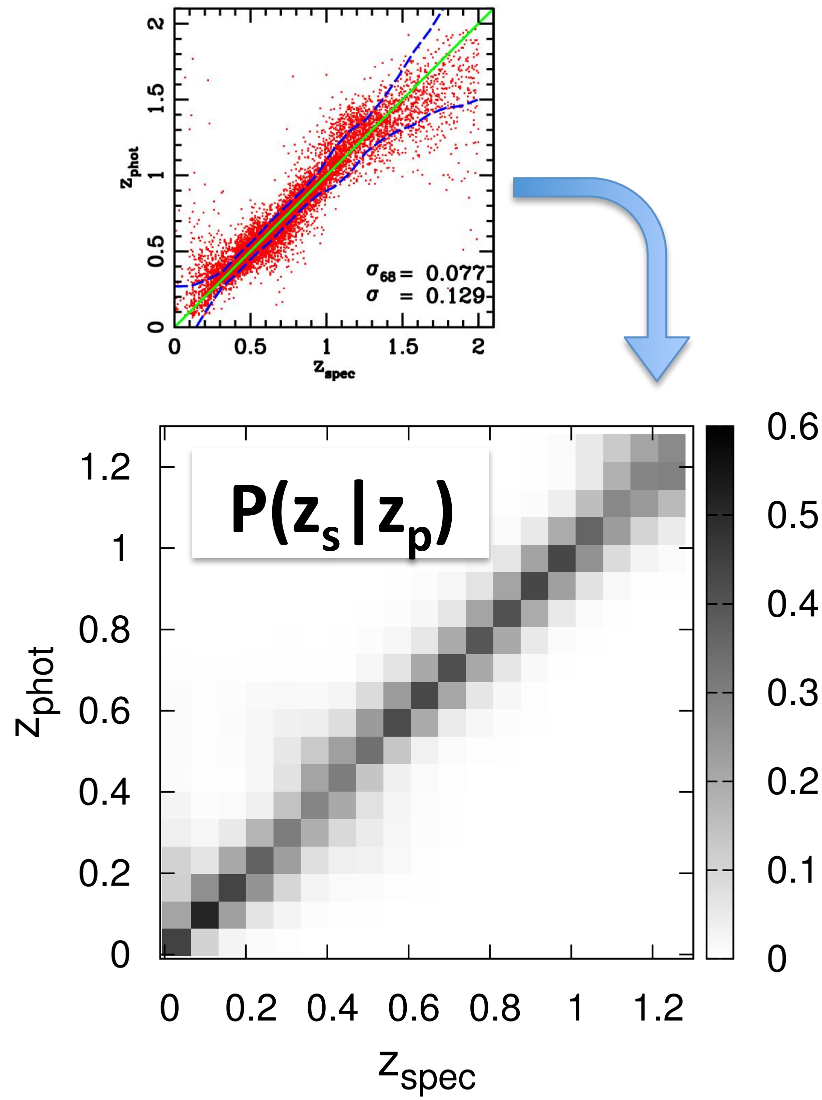


Basics of photo-z's

Photo-zs are often not very good.

Three steps before getting to the cosmology:

- **Get photo-zs; spectra recommended**
- **Estimate photo-z errors and cull outliers; spectra recommended**
- **Calibrate error distribution, e.g. $P(z_s|z_p)$. spectra required**



Need spectra, so what?

Good spectroscopic samples are hard to come by. **Issues**

- **Selection in observables:** typically have many more bright samples than faint samples.
- **Selection in non-observables:** sample selected for a different purpose with different bands (e.g. DEEP2 survey).
- **Shot-noise:** samples are small.
- **Sample variance:** surveys are pencil-beam.
- **Spectroscopic failures:**
  - Can't get spectra for certain galaxies.
  - Wrong spectroscopic redshifts.

Need spectra, so what?

Good spectroscopic samples are hard to come by. **Solutions**

- **Selection in observables: e.g. Weights** (Lima, Cunha et al 2008)
- **Selection in non-observables: Don't do it.**
- **Shot-noise: need many galaxies**
- **Sample variance: need lots of area.**

  } Cunha et al. 2012a

- **Spectroscopic failures:**
  - **Can't get spectra for certain galaxies.**
  - **Wrong spectroscopic redshifts.**

  } Cunha et al. 2012b

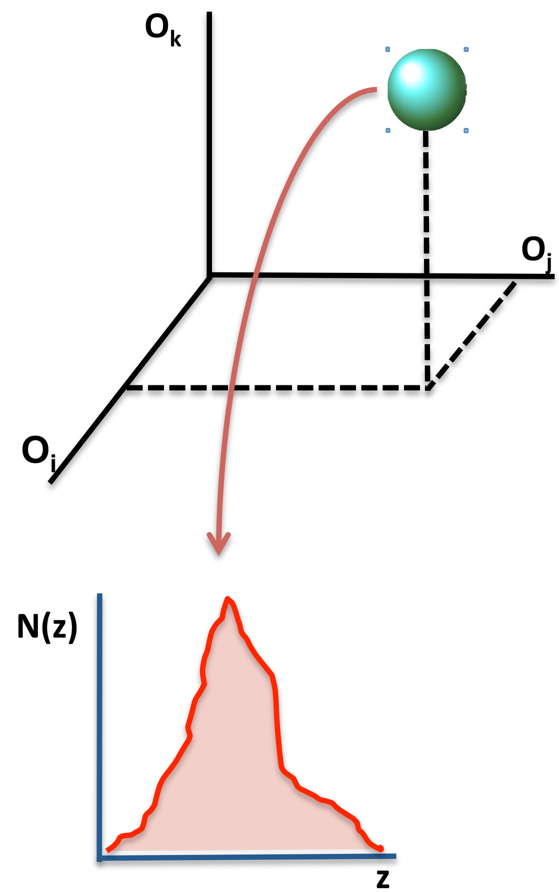
Weights

Match distributions of observables in **training** (spectroscopic or simulated) sample and **photometric** sample by assigning **weights** to training set galaxies.

$$Weight \propto \frac{\rho_{photo}}{\rho_{train}} \quad \text{where} \quad \rho_i = \frac{N_i}{V}$$

N_i : number of galaxies within ball of volume V.

The radius of the ball is determined by the distance to 100th nearest neighbor in the training set in space of observables (colors and magnitudes).



Assumption: Training set is **locally** representative of photometric set.

Is that true? Yes, **if** differences in selection are only in observable space.

Need spectra, so what?

Good spectroscopic samples are hard to come by. **Solutions**

- **Selection in observables: e.g. Weights** (Lima, Cunha et al 2008)
- **Selection in non-observables: Don't do it.**
- **Shot-noise: need many galaxies**
- **Sample variance: need lots of area.**

Cunha et al. 2012a

- **Spectroscopic failures:**
  - **Can't get spectra for certain galaxies.**
  - **Wrong spectroscopic redshifts.**

Cunha et al. 2012b

The Dark Energy Survey

- Study Dark Energy using 4 complementary techniques:
  I. Cluster Counts
  II. Weak Lensing
  III. Baryon Acoustic Oscillations
  IV. Supernovae

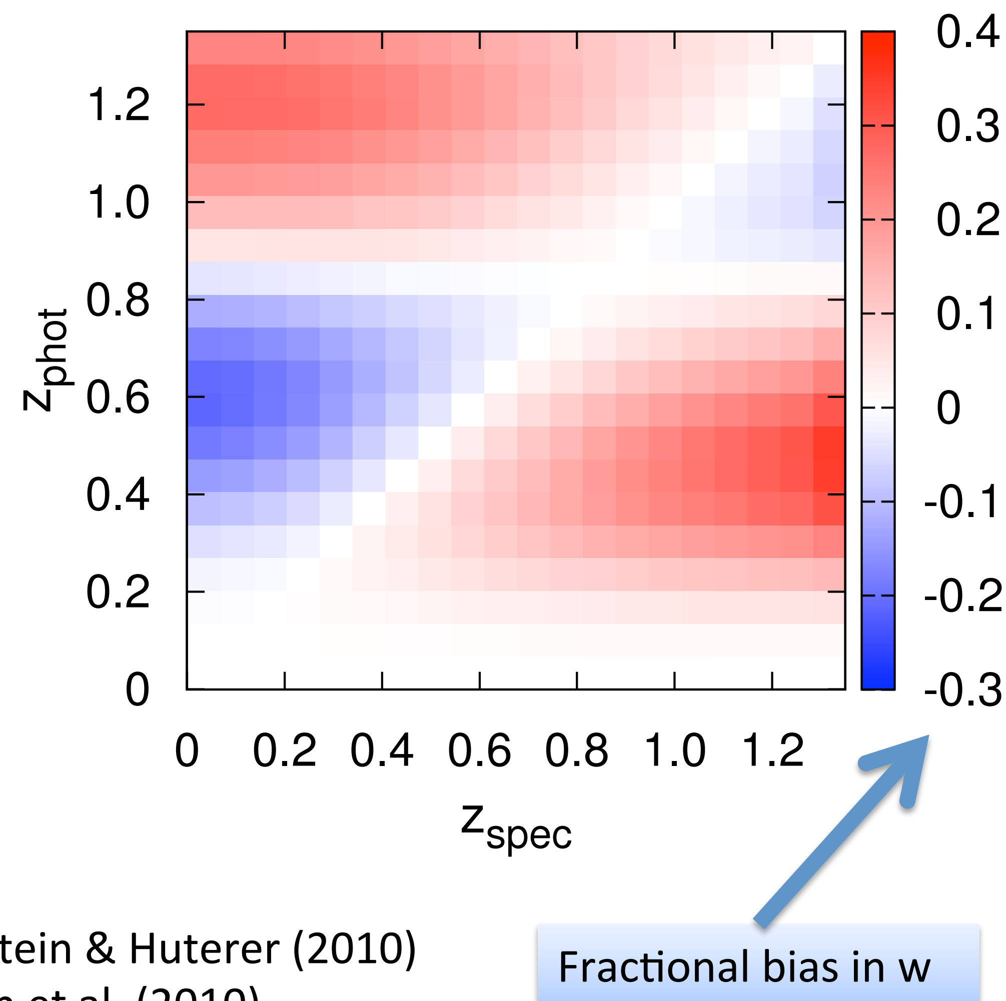
- Two multiband surveys:
  Main: 5000 deg^2 ≈ 5 (h^{-1}Gpc)^3
  300 million galaxies
  g, r, i, z, Y to 24th mag
  SNe: 15 deg^2 repeat

- Build new 3 deg^2 FoV camera and Data management sytem in Blanco 4-m telescope
  Survey 2012-2017 (525 nights)
  Camera available for community use the rest of the time (70%)

www.darkenergysurvey.org

Biases in w from error in $P(z_s|z_p)$ estimation

- Fixed **0.01** error in $\mathbf{P(z_s|z_p)}$ estimation, i.e, **$\Delta P(z_s|z_p)$ =0.01** at a single bin.
- For DES shear-shear analysis.



Bernstein & Huterer (2010)
Hearin et al. (2010)
Cunha et al. 2012a

Biases due to sample variance

Sample variance

For typical existing spectroscopic samples, sample variance is significantly larger than shot noise.

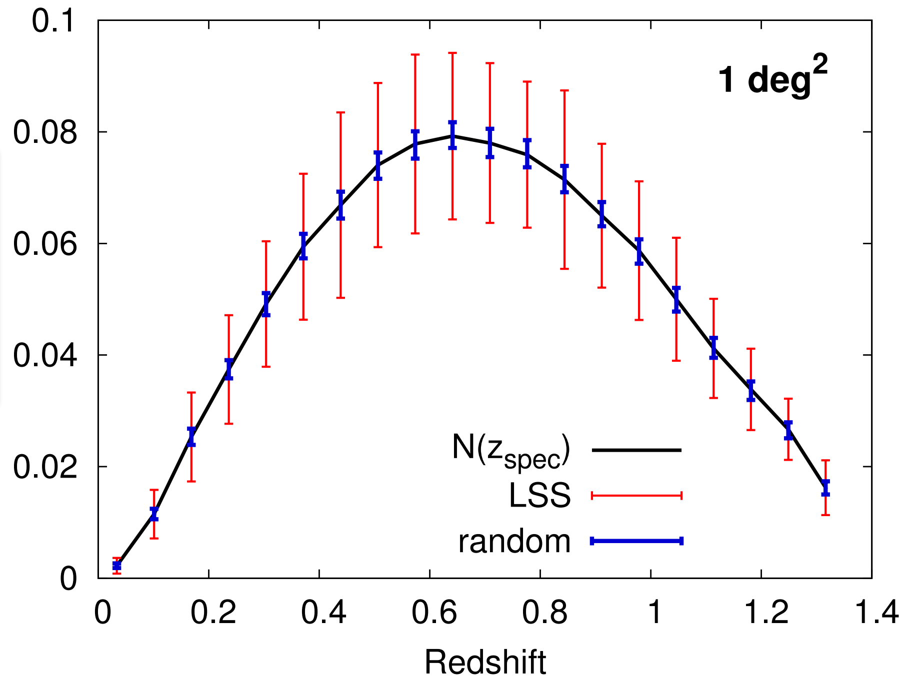


Figure 1. Normalized spectroscopic redshift distribution for the full data. The red (light gray) error bars show the 1-σ variability in the redshift distribution for contiguous 1 deg^2 angular patches. The blue (dark gray) error bars show the variability in the redshift distribution assuming random samples of with the same mean number of objects as the 1 deg^2 patches. We assume that only a 25% random subsample of each patch is targeted for spectroscopy, yielding about 1.2×10^4 galaxies per patch on average.

Cunha, Huterer, Busha & Wechsler
arXiv: 1109:5691

Sample variance in photo-zs and zspecs

Example:

Distribution of galaxies in photometric sample:

z_{phot} (rows) vs. z_{spec} (columns)

1	1	2
1	6	1
2	1	1

Distribution of galaxies in calibration sample:

z_{phot} (rows) vs. z_{spec} (columns)

1	1	4
1	6	2
2	1	2

LSS fluctuation!!!

Sample variance in photo-zs and zspecs

photometric sample:

1	1	2
1	6	1
2	1	1

calibration sample:

1	1	4
1	6	2
2	1	2

$P(z_p | z_s)$

0.25	.125	0.50
0.25	0.75	0.25
0.50	.125	0.25

=

0.25	.125	0.50
0.25	0.75	0.25
0.50	.125	0.25

$P(z_s | z_p)$

0.25	0.25	0.50
.125	0.75	.125
0.50	0.25	0.25

≠

.167	.167	.667
.111	.667	.222
0.4	0.2	0.4

Colums:

z_{phot}

Rows:

z_{spec}

Sample variance in photo-zs and zspecs

photometric sample:

1	1	2
1	6	1

1	1	4
1	6	2

calibration sample:

Conclusion:

$P(z_s|z_p)$ is sensitive to z_{spec} fluctuations, but **$P(z_p|z_s)$** is not. Conversely, only **$P(z_p|z_s)$** is sensitive to z_{phot} fluctuations.

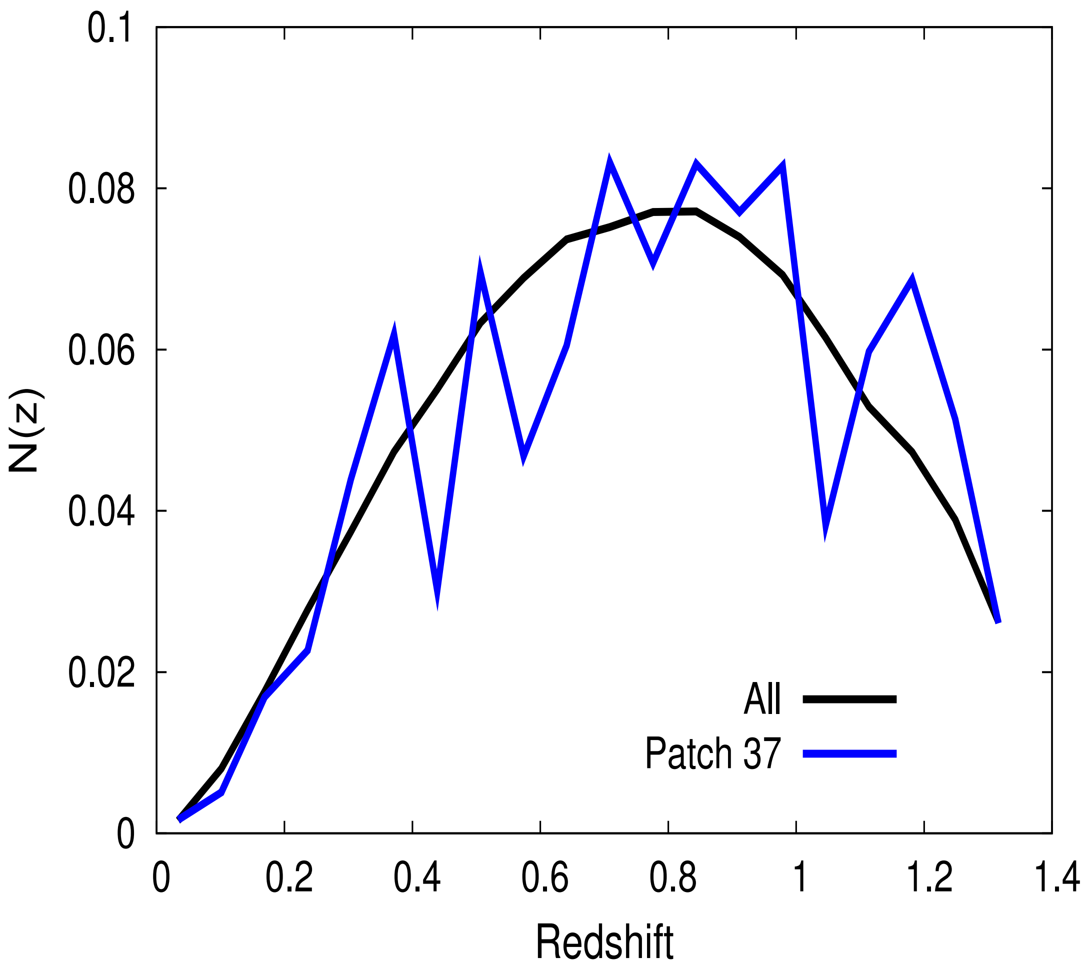
$P(z_s|z_p)$

.125	0.75	.125
0.50	0.25	0.25

$\neq$

.111	.667	.222
0.4	0.2	0.4

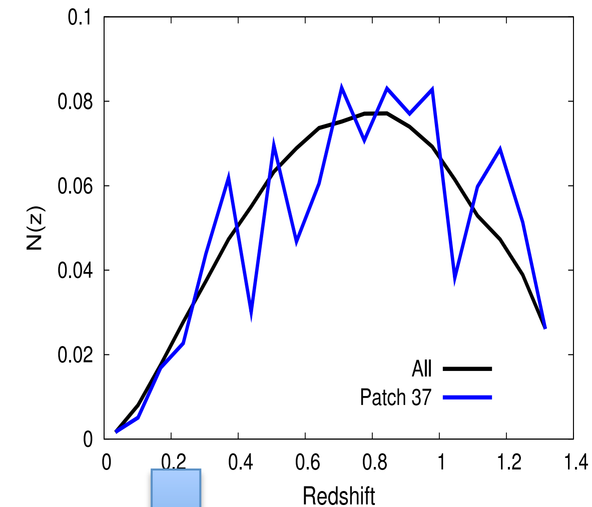
Example: Patch 37



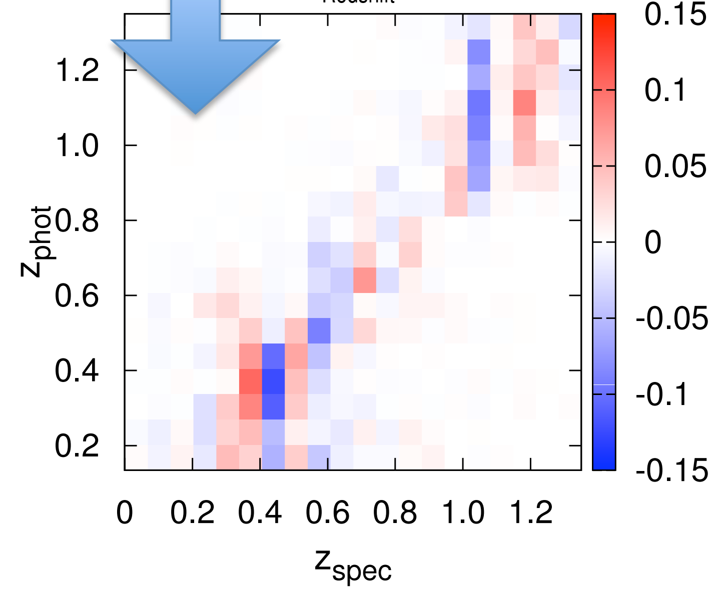
Example: Patch 37

An example:

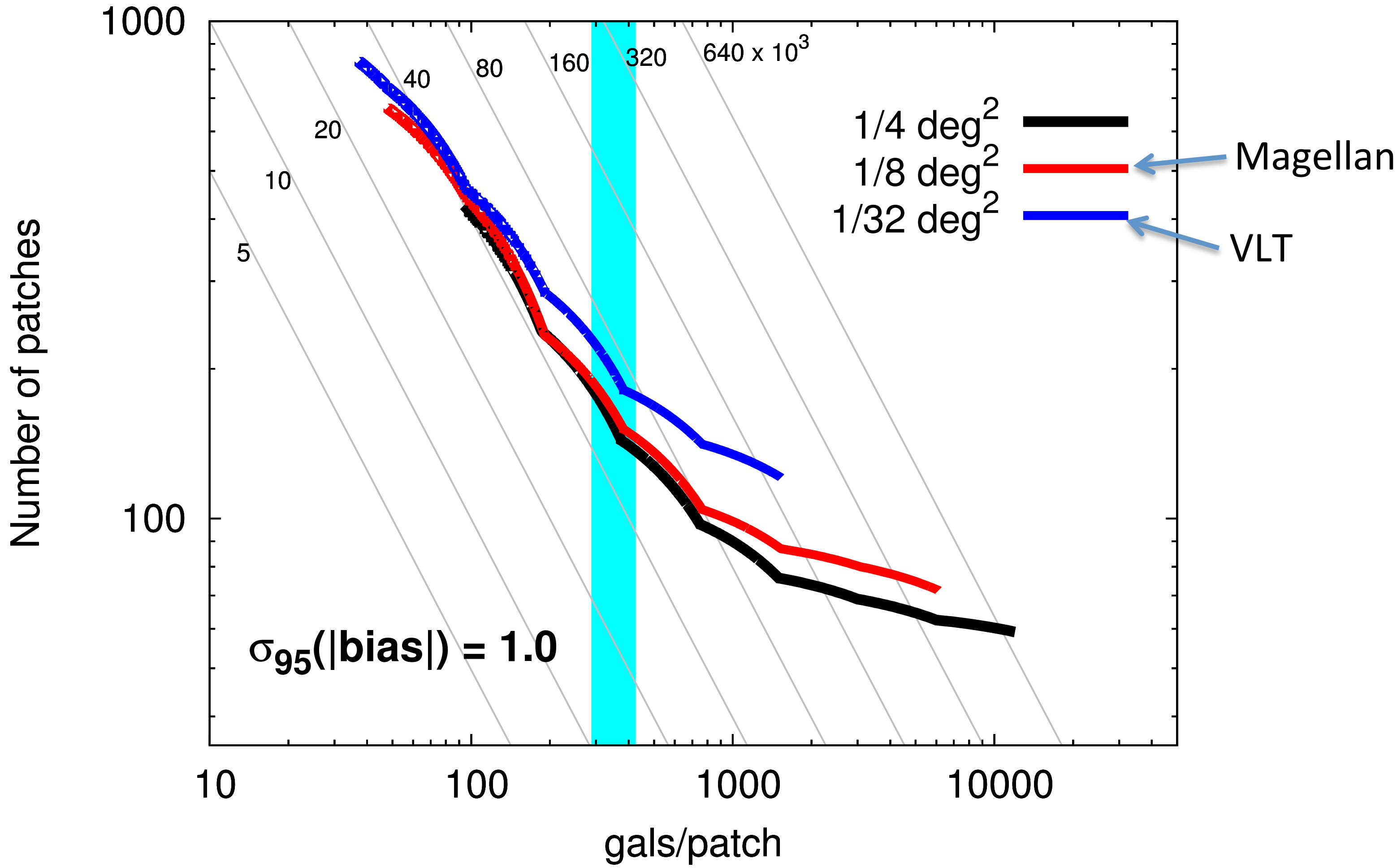
- Errors in $N(z_{spec})$ translate into errors in the error matrix estimation.



$$\Delta P(z_s|z_p) = P(z_s|z_p)_{phot} - P(z_s|z_p)_{train}$$



Survey Calculator

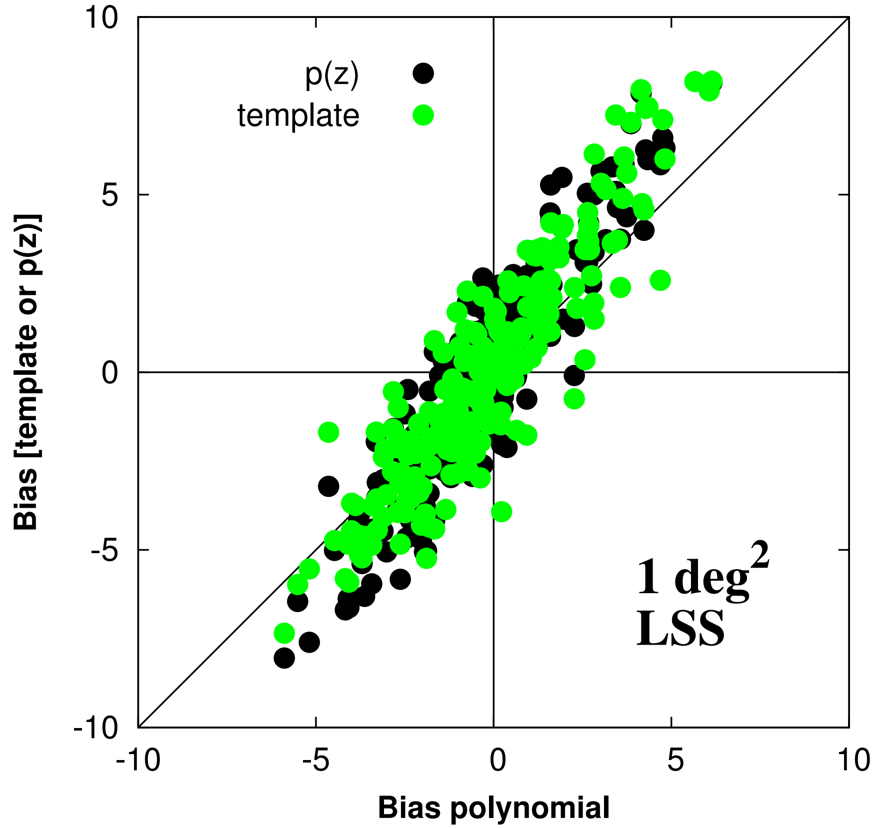


Assuming fiducial σ(w)=0.035, and perfect spectroscopic selection.

Cunha, Huterer, Busha & Wechsler (2012a)

Do we really need all those patches?

w-biases for each patch



Cunha, Huterer, Busha & Wechsler (2012)

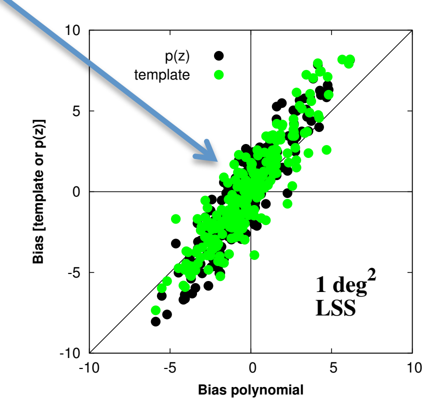
Getting lucky with cosmic variance

- What if the patches were not chosen randomly?

Getting lucky with cosmic variance

- What if the patches were not chosen randomly?

- If we are **lucky**, a single patch may result in very small bias.



Getting lucky with cosmic variance

- What if the patches were not chosen randomly?
- If we are **lucky**, a single patch may result in very small bias.
- **Question:** how do we find that patch?

Getting lucky with cosmic variance

- **Idea:** The patches that look **most similar** to the **survey average**, will produce the smallest biases if used for calibration.

- **Question:** What do you mean by **most similar**?

Choosing the most similar patches

Several options:

- 1-point statistics (e.g. Redshift distribution)
  - Rms (χ^2)
  - Kolmogorov-Smirnov (KS) – more sensitive to biases.
- 2-point statistics (e.g. correlation function)

Based on:

- Photometric properties
- Spectroscopic properties (perhaps of a brighter sample)

The simulations

- 8000 sq. degrees
- DES depth (griz bands), cut at i<23.5
- BPz photo-zs (for plots shown).

- Photo-z stats (pretty awful, at present – more on this later):
  - σ= 0.2
  - σ_{68}= 0.13

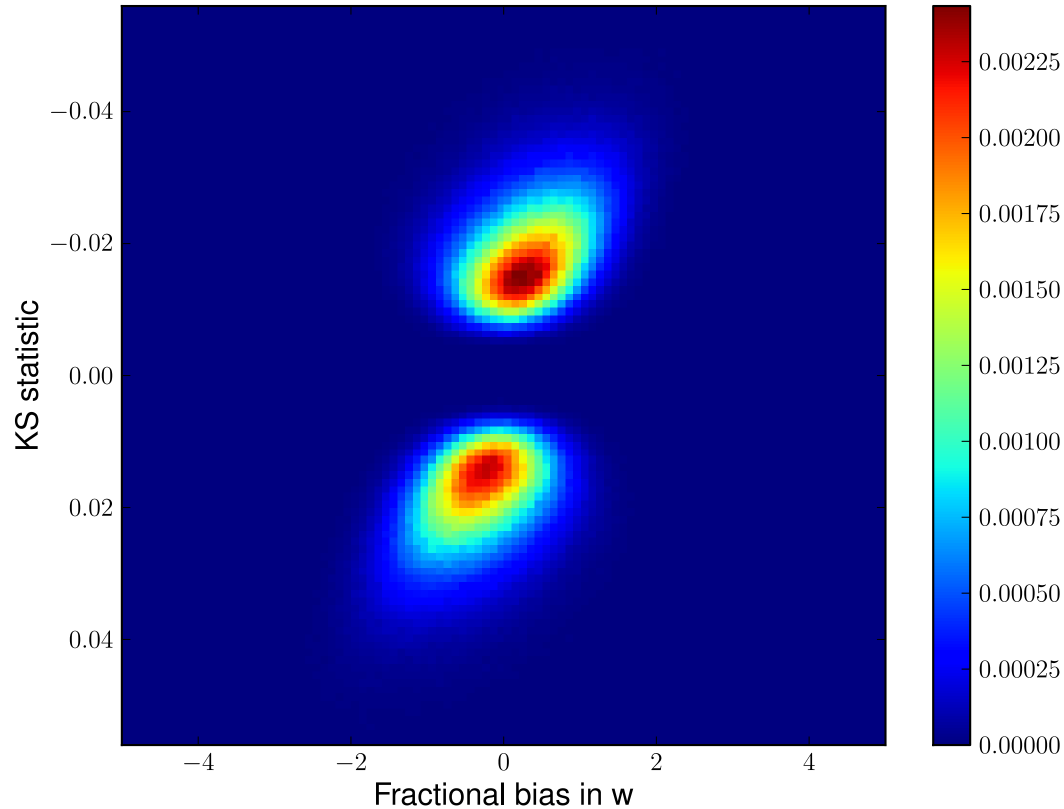
Procedure

- Split simulation **into thousands of patches** of area 1/8 sq. deg – comparable to aperture of Magellan.
- To improve statistics, generate **millions of sets of 20 patches**, randomly picked.
- Look **at fractional biases in w** from using each of the sets of patches.
- Look for correlations between biases in w and how well N(z) of patches reproduces the simulation mean.

Spectroscopic case

Start with naïve scenario:
- Use **N(z_{spec})** to choose lucky patches.

Distribution of patch-sets in **KS – w-bias** space.

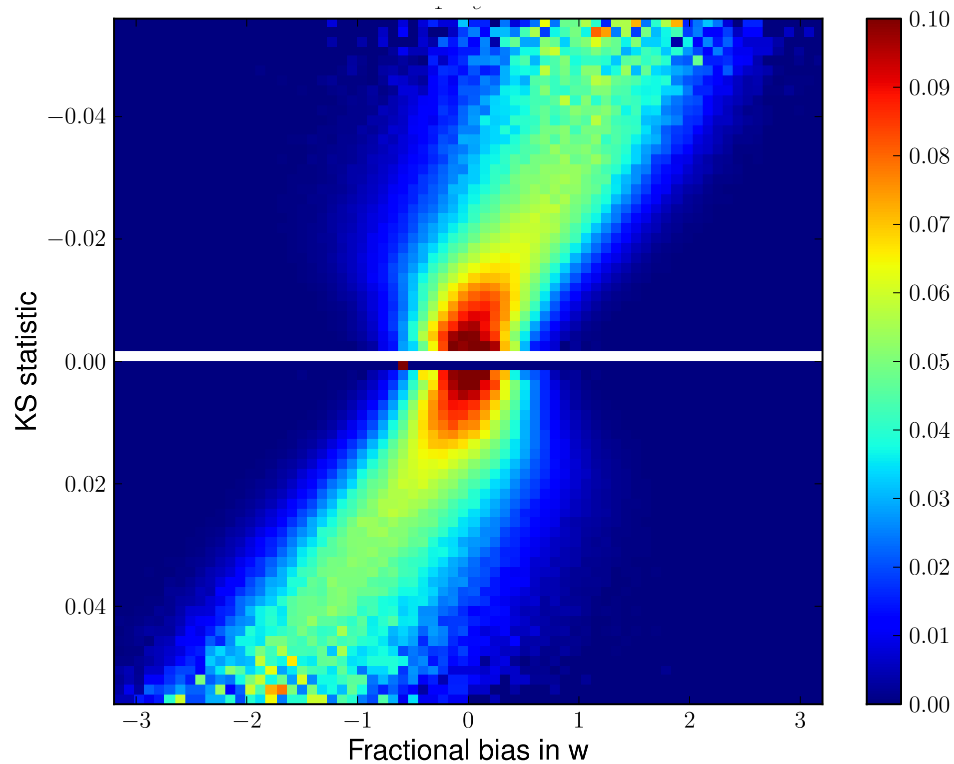


Spectroscopic case

Start with naïve scenario:
- Use **N(z_{spec})** to choose lucky patches.

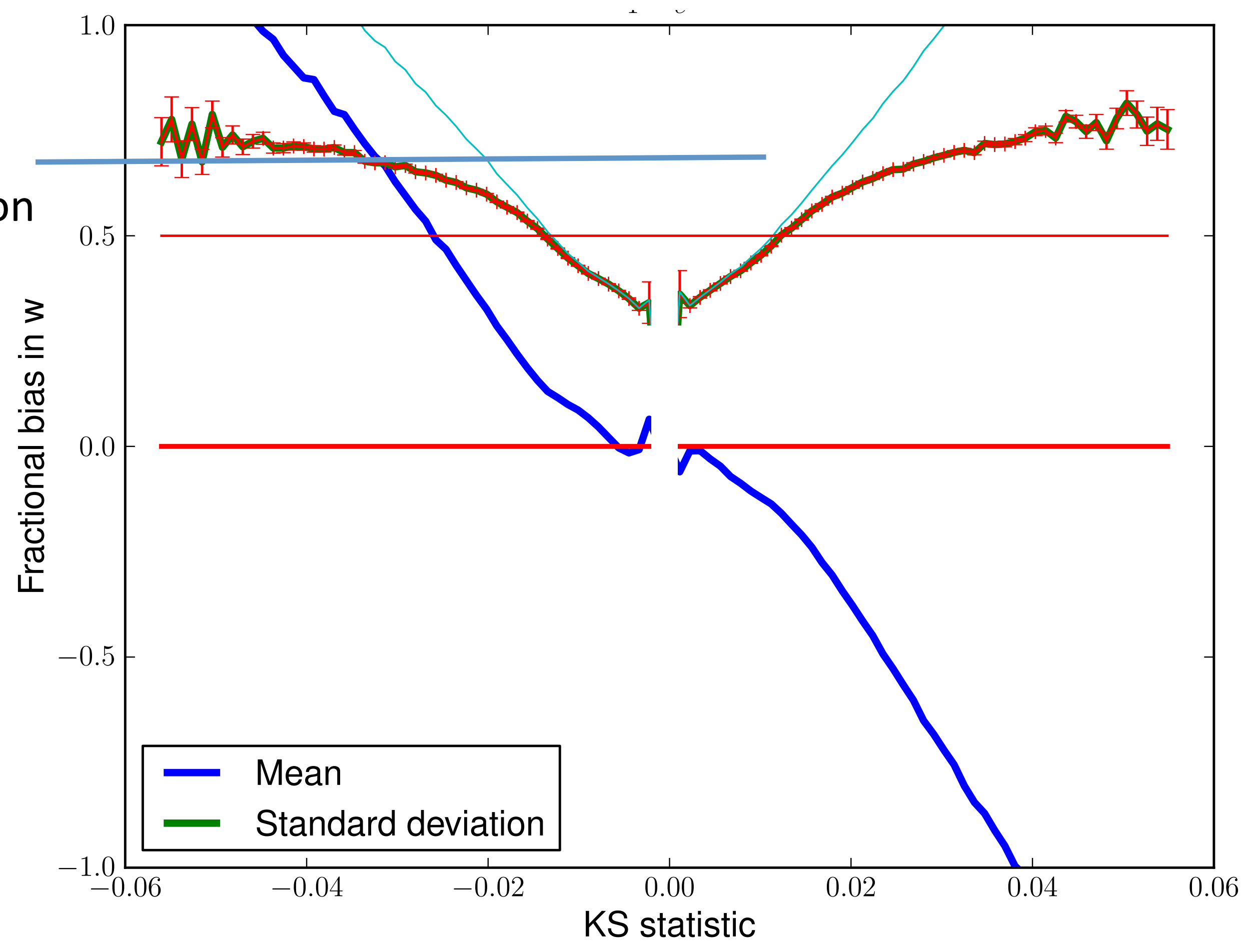
Probability of w-bias given KS-statistic:

P(w-bias|KS)



Spectroscopic case

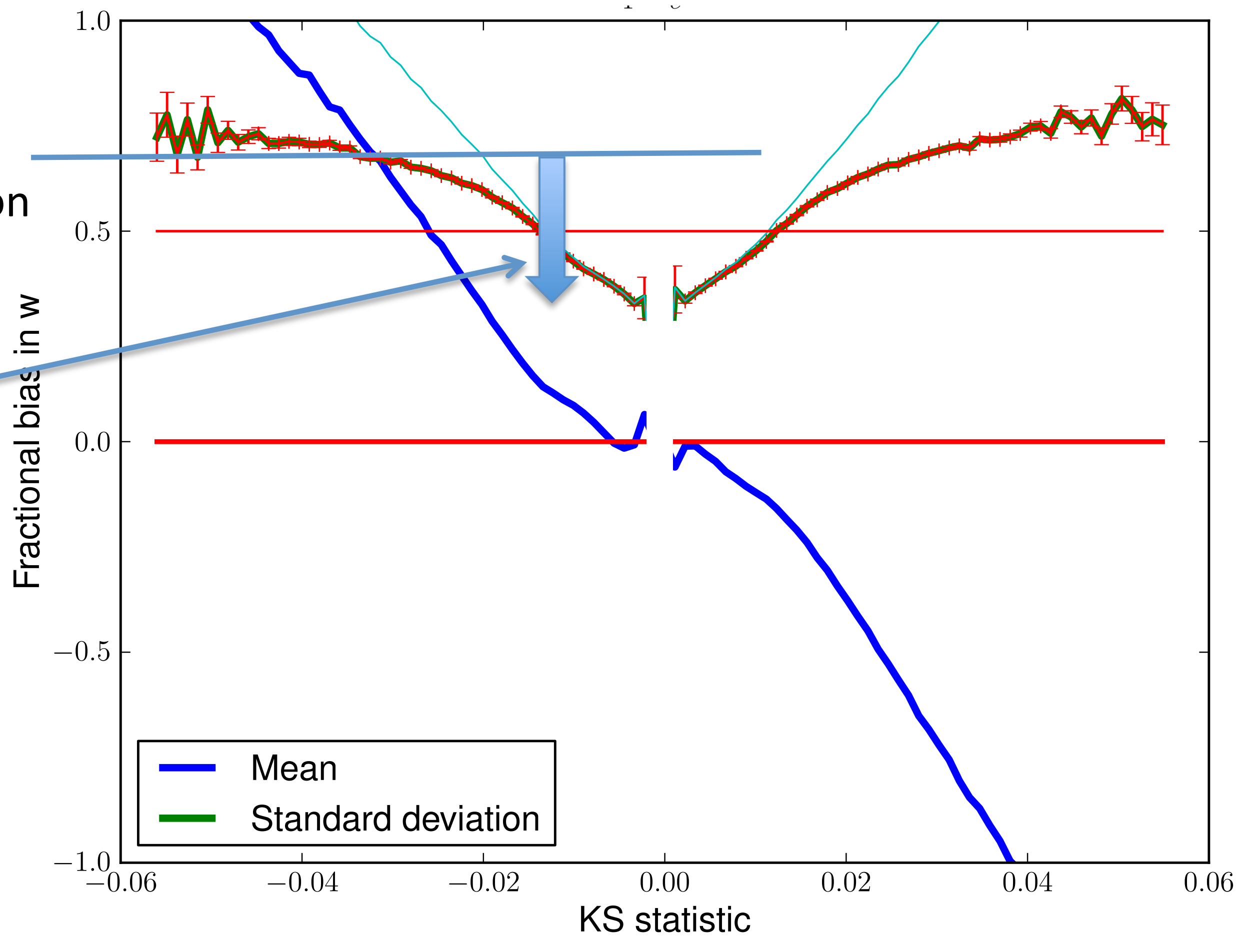
σ of total population



Spectroscopic case

σ of total population

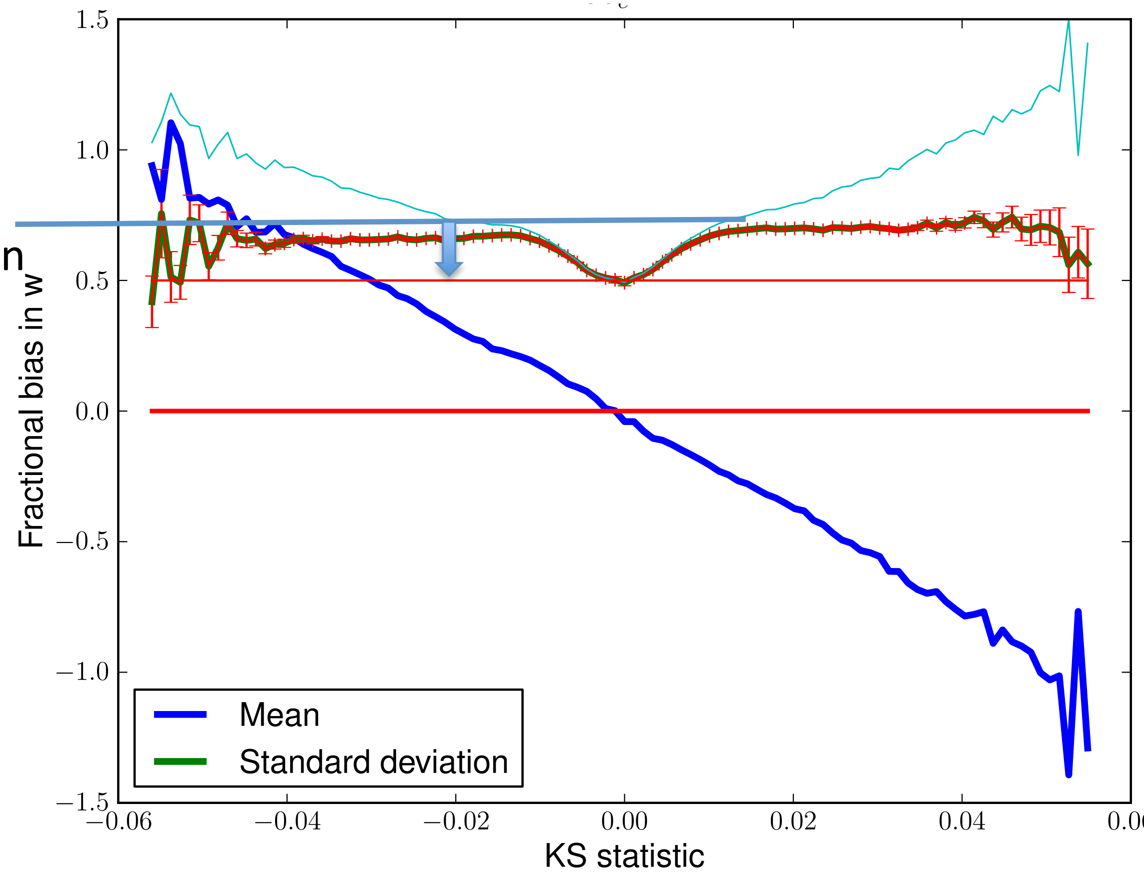
4x improvement in patches from using lucky patches (**2x in w-bias**)



Photometric case

Use N(z_{phot}) to choose lucky patches.

σ of total population



2x improvement in patches from using lucky patches (**1.4x in w-bias**)

Parenthesis: How do photo-z errors affect results?

- Effects on statistical constraints for fixed calibration:
  - Square-root of scatter (Zhaoming Ma, Fisher matrix, depends on priors)

- Effects on calibration for fixed statistical constraints:
  - Shot-noise dominated limit: **No dependence**, believe it or not.
  - Sample-variance dominated limit: **Square-root of scatter.**

- Impact on lucky-patch selection:
  - ??? – at least square-root of scatter.

Conclusions

- Lucky patch selection is free and reduces calibration requirements.
- Can this approach bias things in any way?
- What is the best 1-point statistic and the best tracer population?
- How much will 2-point statistics contribute?

Recap

Good spectroscopic samples are hard to come by. **Solutions**

- **Selection in observables: e.g. Weights** (Lima, Cunha et al 2008)
- **Selection in non-observables: Don't do it.**
- **Shot-noise: need many galaxies**
- **Sample variance: need lots of area.**

  } Cunha et al. 2012a

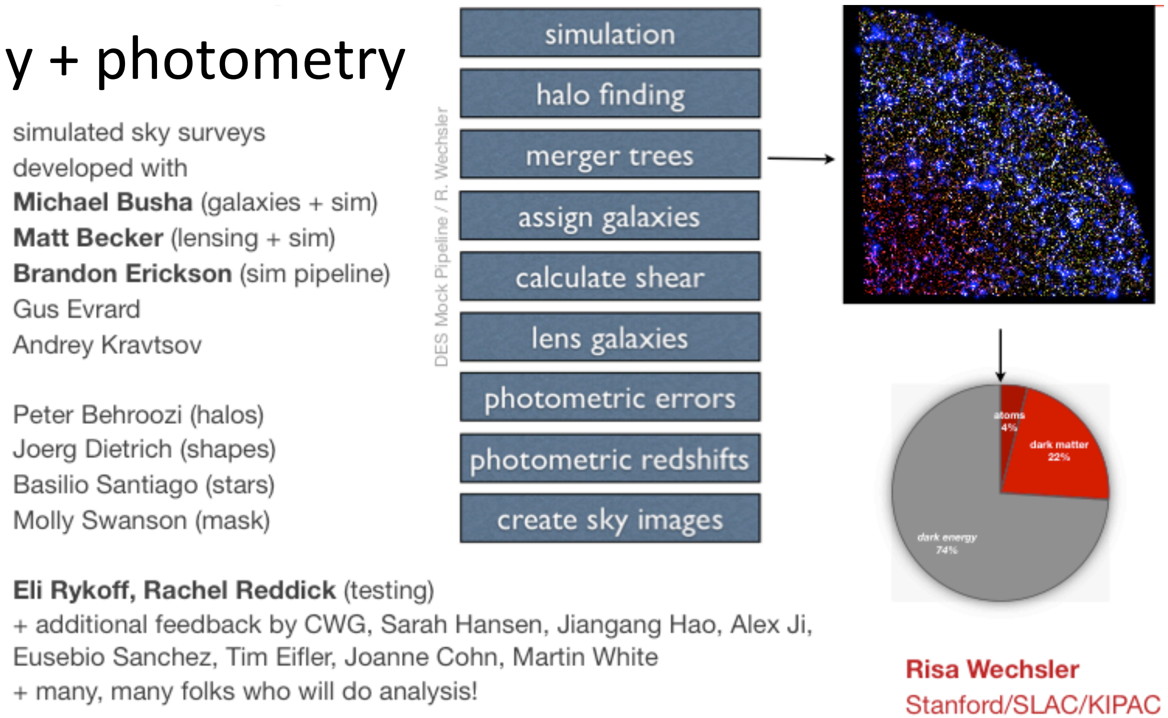
- **Spectroscopic failures:**
  - **Can't get spectra for certain galaxies.**
  - **Wrong spectroscopic redshifts.**

  } Cunha et al. 2012b

Spectroscopic failures

Spectroscopic simulations

- N-body + photometry



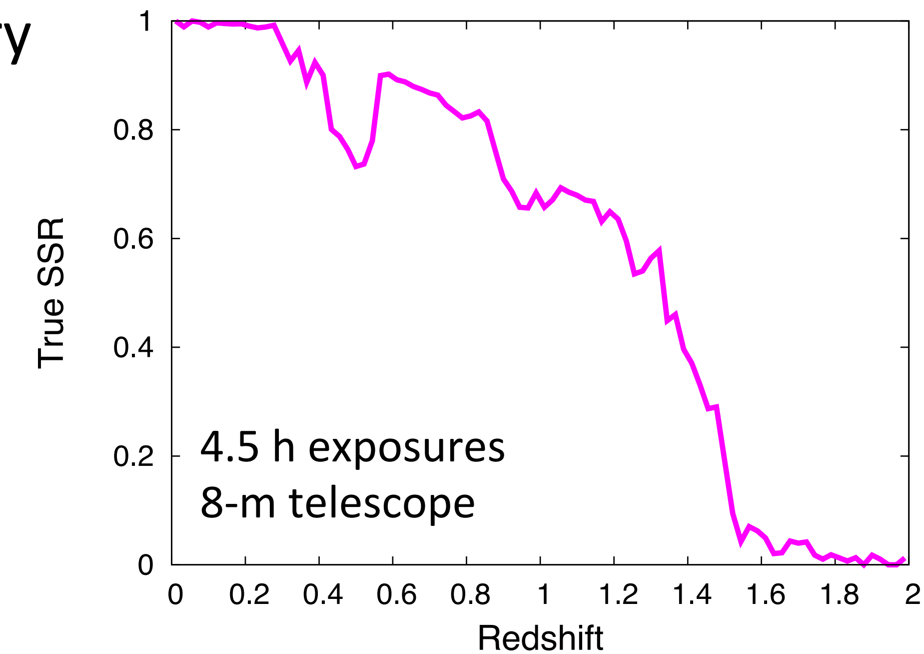
- Spectroscopy:
  - **Simulated spectra:** K-correct templates + resolution + noise
  - **Spectroscopic redshifts:** IRAF-rvsao on 1-D simulated spectra

Completeness issues

- Spectroscopic samples are very incomplete

Case study: Simulations of

DES photometry + VVDS-like spec-z's

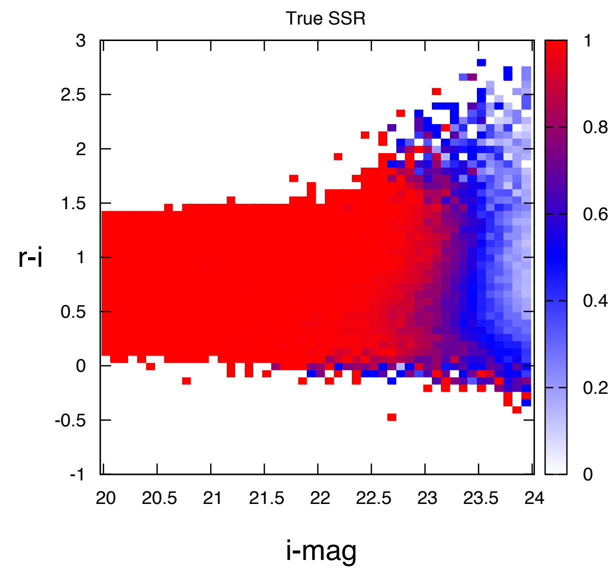


True SSR: fraction of galaxies with correct redshifts.

Completeness issues

SSR: Spectroscopic Success Rate

True SSR: fraction of galaxies with correct redshifts.



Completeness issues

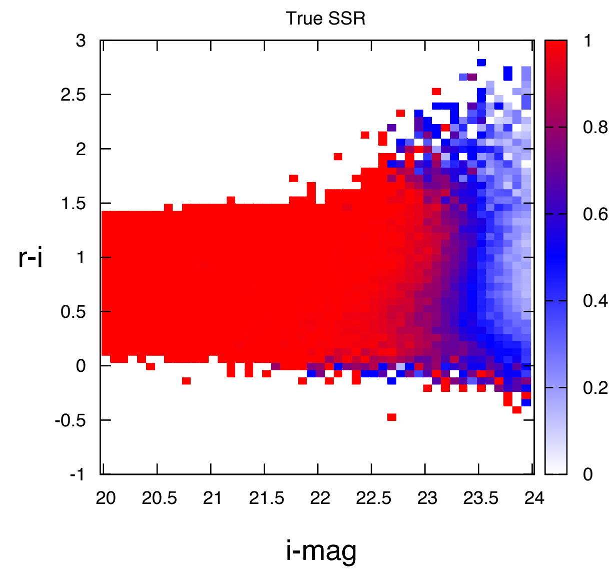
SSR: Spectroscopic Success Rate

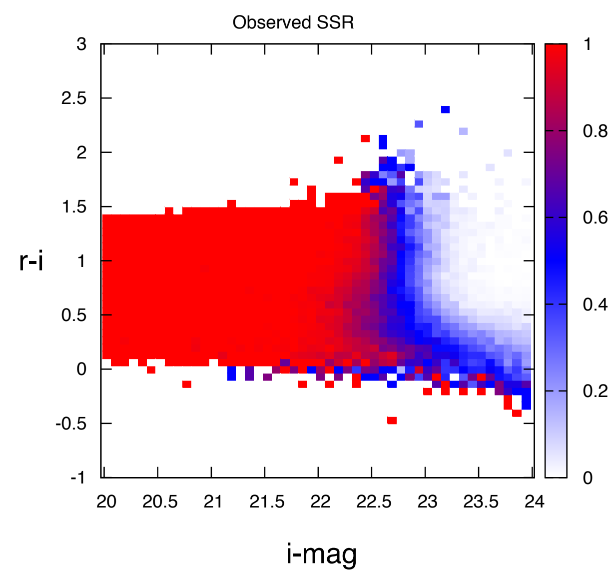
True SSR: fraction of galaxies with correct redshifts.

Observed SSR: Fraction of galaxies with redshift confidence above some threshold (e.g. Q>4).

Q: Strength of correlation between observed spectra and best-fitting spectrum in a template library.

Cannot use **spectroscopic** sample for calibration of photo-zs of **photometric** sample if selection is different.





Completeness issues

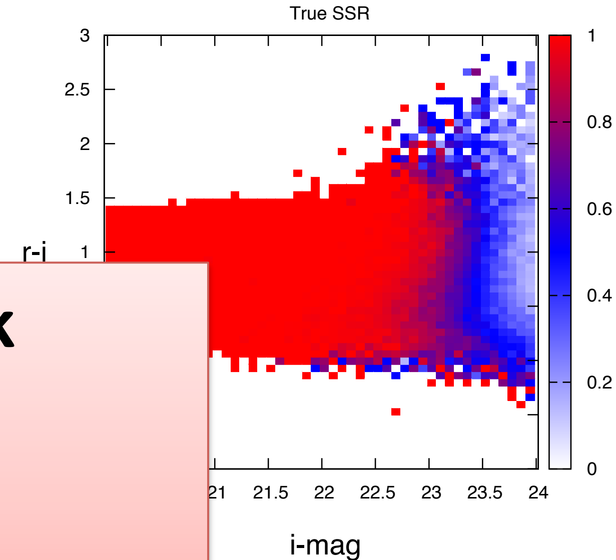
SSR: Spectroscopic Success Rate

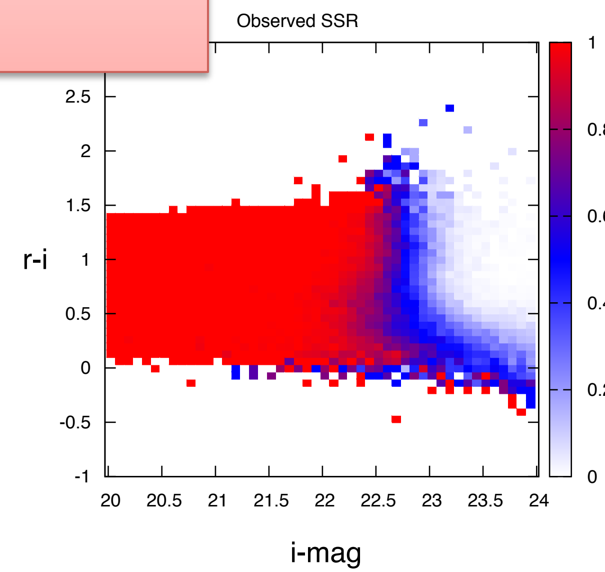
True SSR: fraction of galaxies with correct redshifts.

Observe
redshift
Q>4).

Q: Stren
spectra and best-fitting spectrum in a template library.

In simulations, **neural network** approach was able to **match** spectroscopic selection in photometric sample.





Cannot use **spectroscopic** sample for calibration of photo-zs of **photometric** sample if selection is different.

Selection matching with neural net

- Have a **redshift confidence (Q)** for galaxies in **spectroscopic sample**.

- Use neural net to find a relation between **Q** and photometric observables (magnitudes). This is $\mathbf{Q_{est}}$.

- $\mathbf{Q_{est}}$ can be calculated for **all galaxies** in the **spectroscopic** and **photometric** samples.

Q_{est}: redshift confidence estimated with neural net.

SSR_T: Percentage of correct redshifts in training sample.

z_{true}: bias due to selection matching with neural networks: **is negligible**

z_{spec}: bias due to selection matching + **wrong redshifts**: **is substantial**

Shear-Shear constraints on w

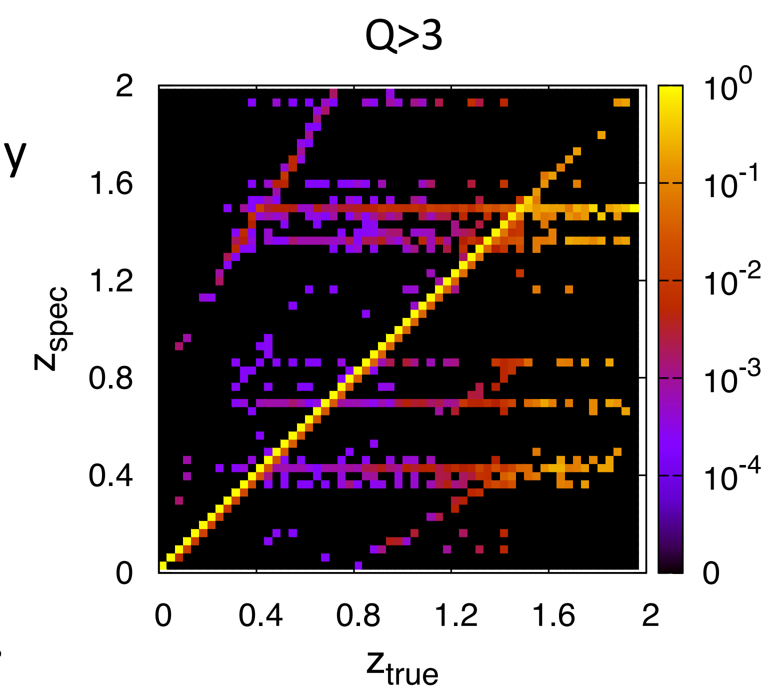
16200 secs				bias(w)	
Selection	Gal. Frac.	SSR_T (%)	$\sigma(w)$	z_{true}	z_{spec}
$Q_{est} > 1.5$	0.75	91.4	0.07	0.004	- 0.52
$Q_{est} > 2.5$	0.59	97.8	0.09	0.002	- 0.13
$Q_{est} > 3.5$	0.46	99.6	0.10	-0.001	- 0.02

Cunha, Huterer, Lin, Busha, Wechsler et al, 2012

Spectroscopic failures (wrong redshifts)

Issues:

- When spec-z's are wrong, they're really wrong.
- A small speck of wrong redshifts is enough to mess up cosmological constraints.



Q: cross-correlation parameter (measures redshift confidence)

Sample used in the plot has 98.6% correct redshifts and constitutes 60% of total sample.

Case study: Simulations of

DES photometry + VVDS-like spec-z's

Cunha, Huterer, Lin, Busha, Wechsler et al, 2012b

Q_{est}: redshift confidence estimated with neural net.

SSR_T: Percentage of correct redshifts in training sample.

z_{true}: bias due to selection matching with neural networks: **is negligible**

z_{spec}: bias due to selection matching + **wrong redshifts**: **is substantial**

Shear-Shear constraints on w

16200 secs				bias(w)	
Selection	Gal. Frac.	SSR_T (%)	$\sigma(w)$	z_{true}	z_{spec}
$Q_{est} > 1.5$	0.75	91.4	0.07	0.004	- 0.52
$Q_{est} > 2.5$	0.59	97.8	0.09	0.002	- 0.13
$Q_{est} > 3.5$	0.46	99.6	0.10	-0.001	- 0.02

Cunha, Huterer, Lin, Busha, Wechsler et al, 2012

Conclusions

- **Incompleteness:**
  - **Does not introduce cosmological biases** if selection matching is performed.
  - Statistical constraints suffer with reduction of sample size.

- **Wrong redshifts:**
  - Cause severe biases.
  - Need better than 99% correct redshifts.
  - If 99% accuracy not possible, need to calibrate spectroscopic error distribution $P(z_{true}|z_{spec})$ with deeper sample/better instrument.

- **Moral of the story:** Focus has to be on accuracy of derived redshifts.

How can we get so many spectra?

Existing instruments:

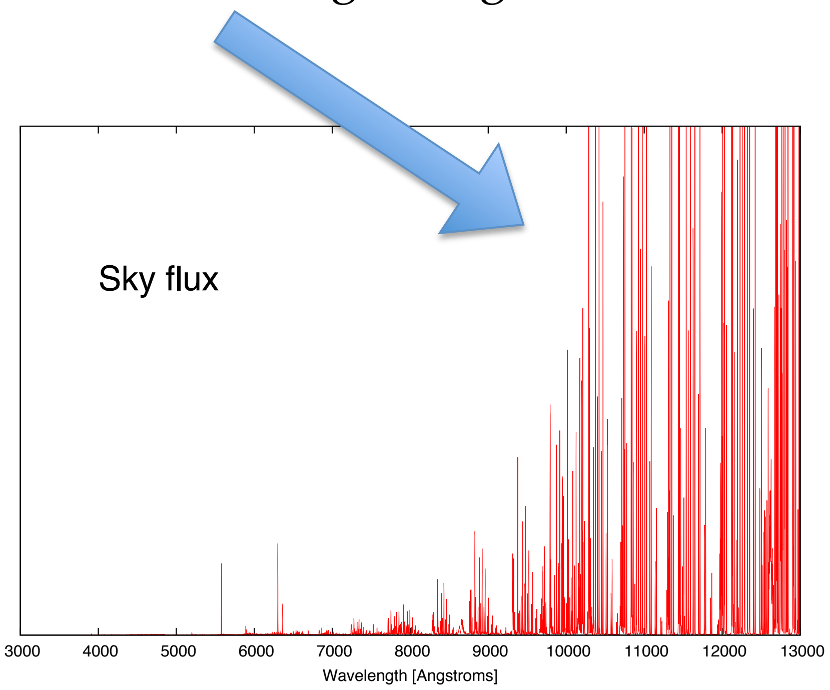
- VLT (8-m)
- Magellan (6.5 m)
- Gemini (8-m)
- Keck (10-m)

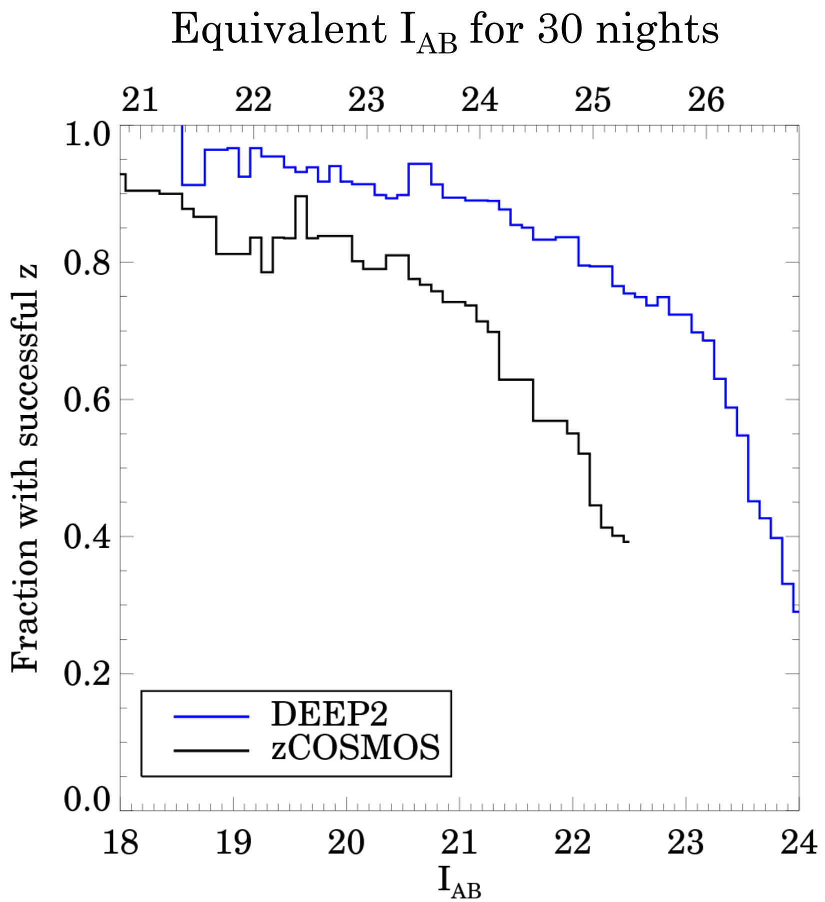
Planned:

- PFS on Subaru (8 m)
- ngCFHT (8 m)
- IFU on WFIRST (2.5 m)
- GMACS (24.5 m)

Calibrating photo-*z*'s at LSST depth is limited by incompleteness in redshift surveys

• Even with instruments now being built, this will be extremely **difficult from the ground** at z>2, degrading DE FoM

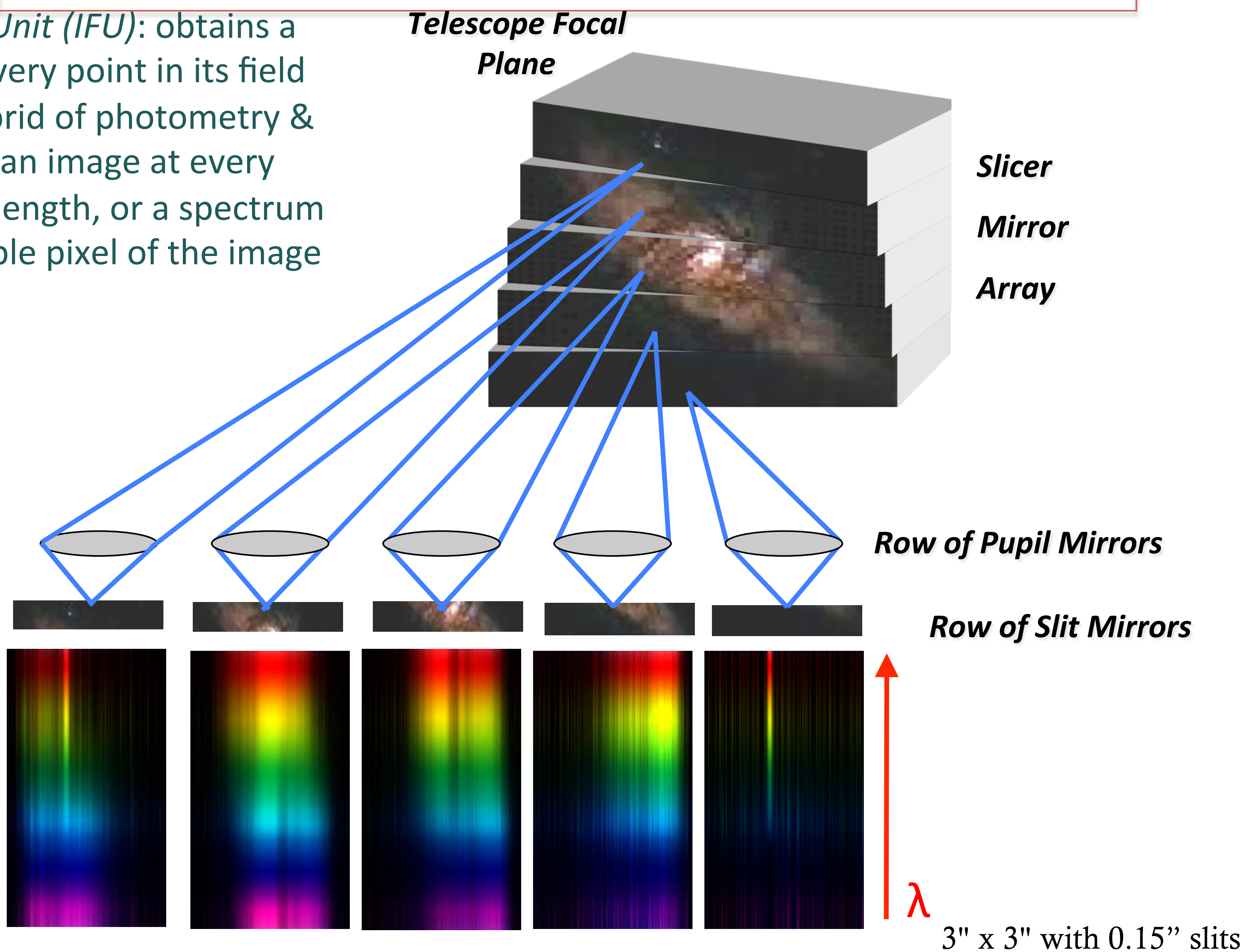




Redshift success rates from DEEP2 (Newman et al. 2012), zCOSMOS (Lilly et al. 2009)

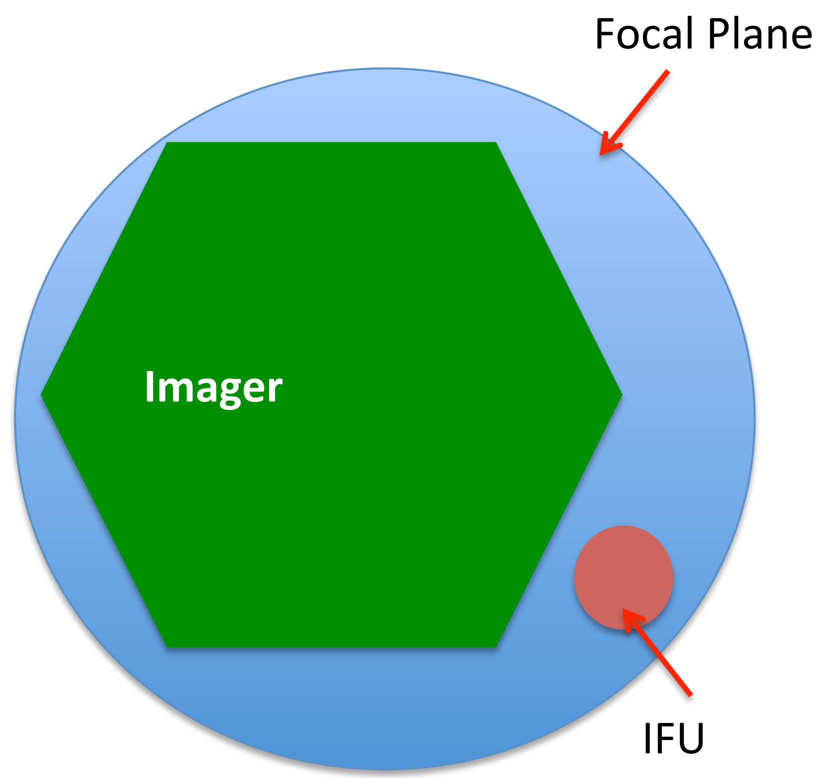
Integral Field Spectroscopy Concept

Integral Field Unit (IFU): obtains a spectrum at every point in its field of view. A hybrid of photometry & spectroscopy: an image at every possible wavelength, or a spectrum at every possible pixel of the image



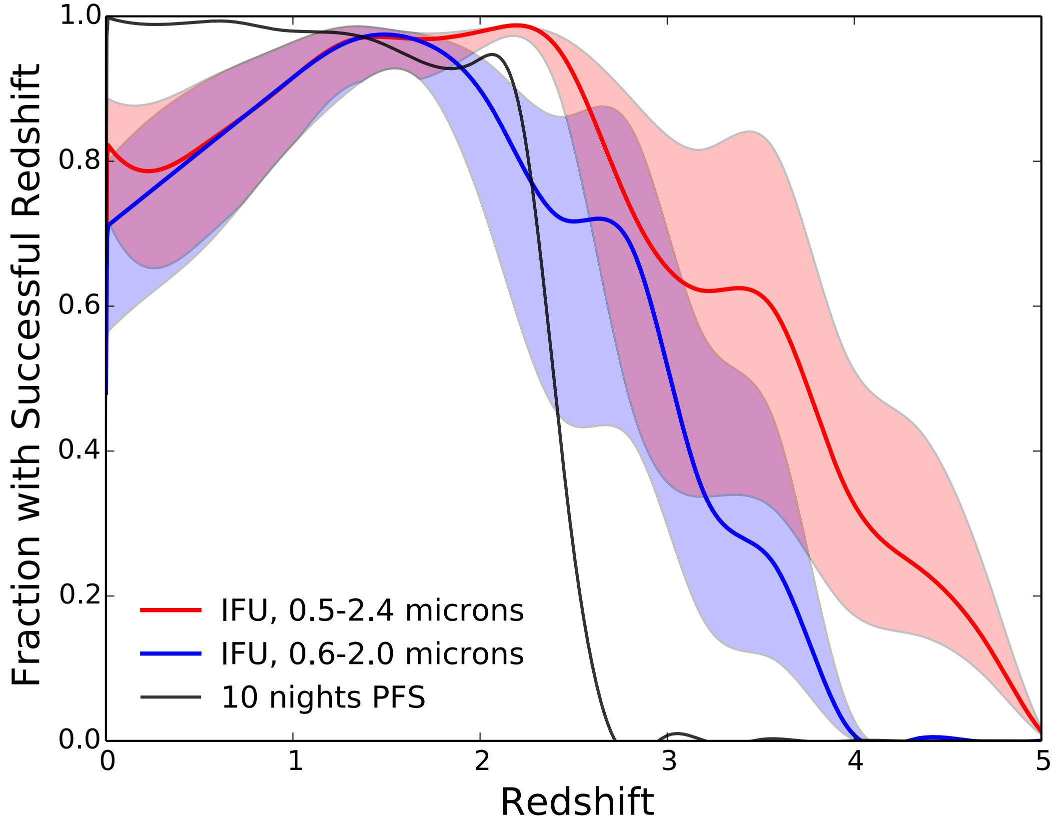
Why an IFU on WFIRST?

- Get supernovae spectra while performing the imaging survey.
- But SN only takes up one pixel in IFU
- **Use rest of IFU to get spectra!**



Why an IFU on WFIRST?

- About 3000 sq degrees.
- few ×10^4 low-res spectra
- Extending LSST calibration beyond z~2 can improve FoM by 30-40%.

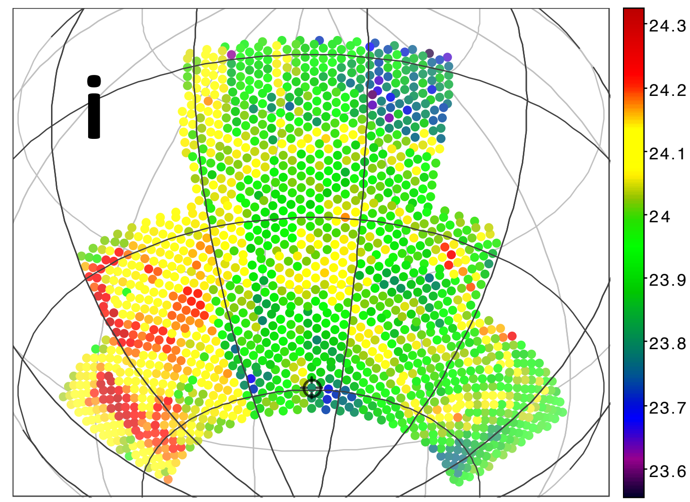


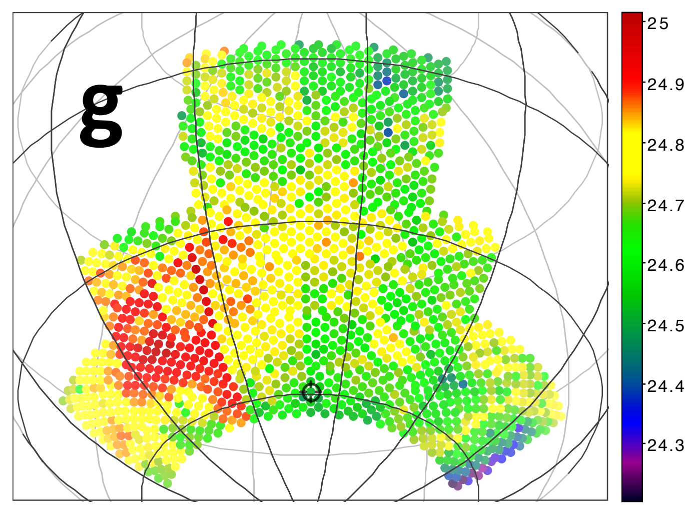
In collaboration with S. Perlmutter, J. Newman and C. Hirate

Part II - Angular selection issues

Strong angular dependence of photo-z scatter

DES 5yr
mag – limits





The End