STANFORD UNIVERSITY

KIPAC

KAVLI INSTITUTE FOR PARTICLE ASTROPHYSICS AND COSMOLOGY

Angular and redshift issues for large galaxy surveys

Carlos E. Cunha

Stanford University

TAUP 2013

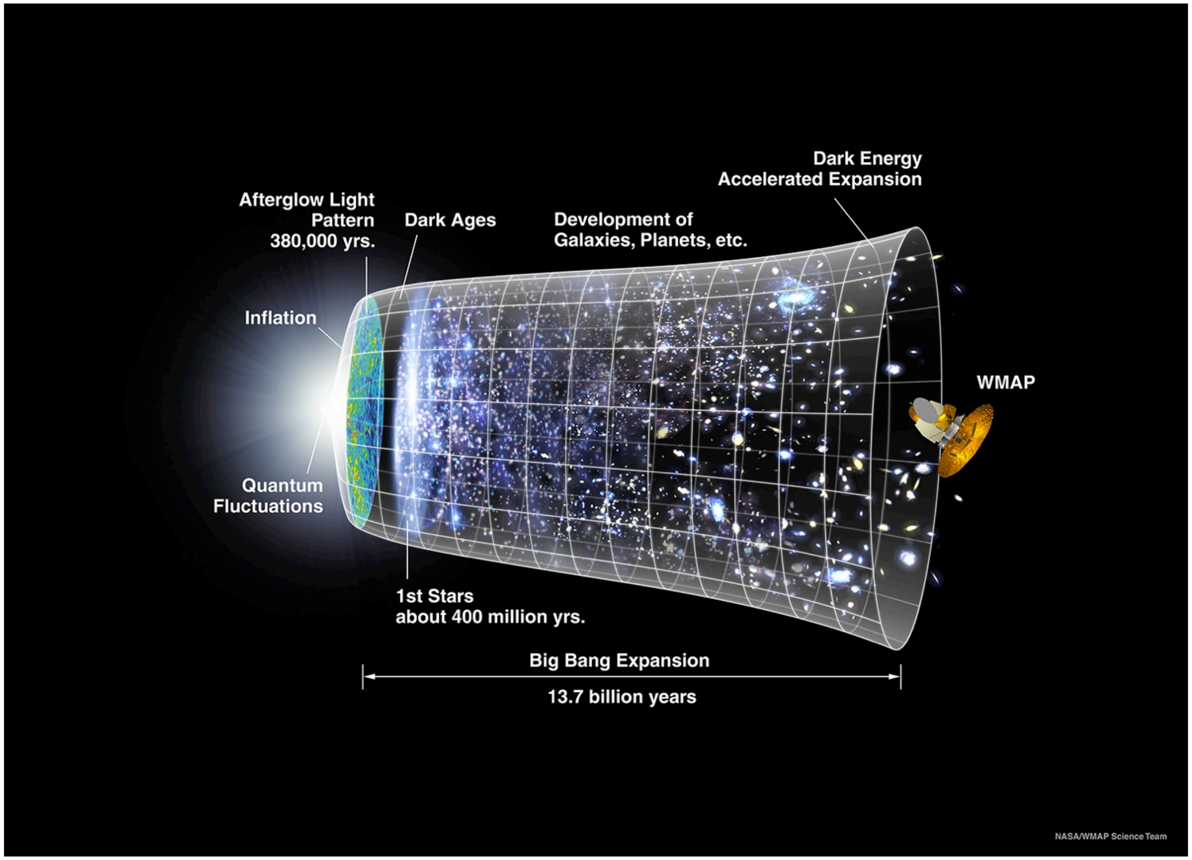
Asilomar, CA – September 10, 2013

Cosmological information

Cosmological information comes from two sources:

Geometry: tracing the expansion history

Growth of large-scale density fluctuations



Cosmological probes from large galaxy surveys

- Supernovae
- Gravitational Lensing
- Large-scale structure (BAO)
- Galaxy Clusters

Ongoing and upcoming optical surveys

Photometric:

- **DES**: Dark Energy Survey
- **PanStarrs**: Panoramic Telescope and Rapid Response System
- **KIDS**: Kilo-Degree Survey
- **HSC**: HyperSuprime Cam
- **LSST**: Large Scale Synoptic Telescope

Spectroscopic:

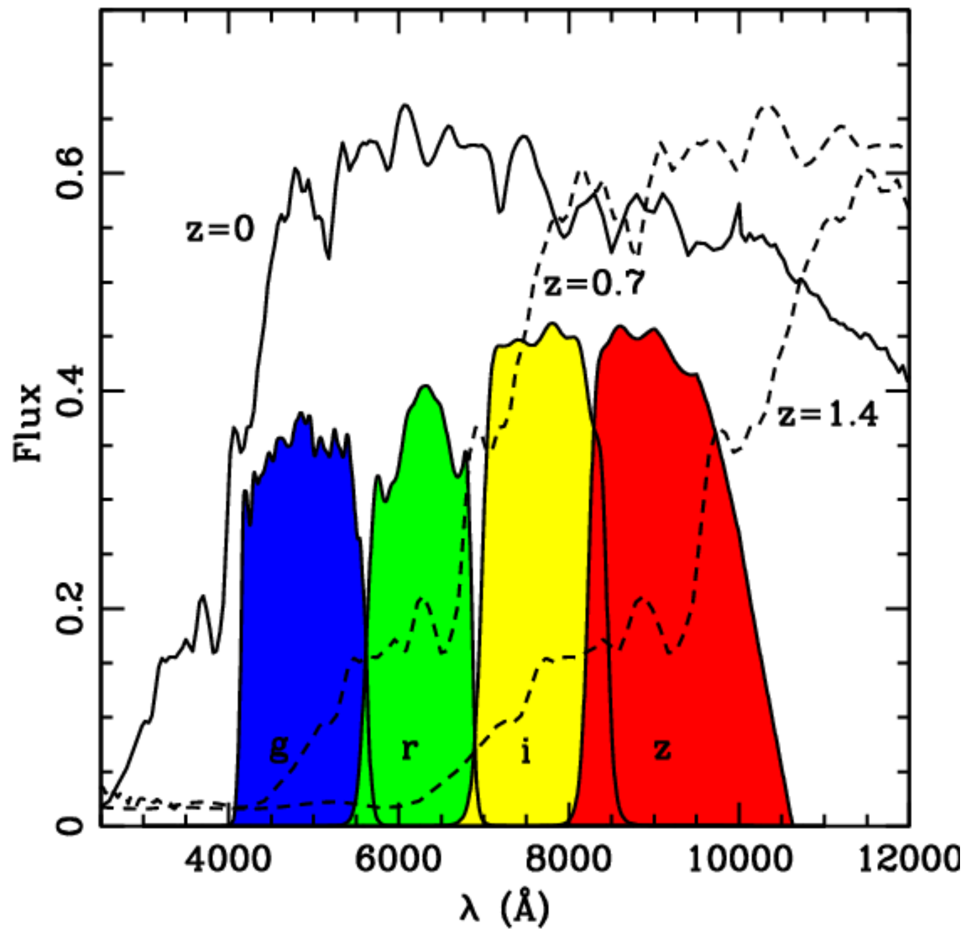
- **BOSS:** Baryon Oscillation Spectroscopic Survey
- **MS-DESI:** Mid-Scale Dark Energy Spectroscopic Instrument

Basics of photometric surveys

Basics of photometric surveys

- Collect light from galaxies in several broad-band filters in optical and near-IR.
- E.g. grizY (DES) + JK (Vista)
- Use flux in each filter to determine:
  - type:star/gal./QSO
  - gal. type: spiral, elliptical, …
  - (photometric) redshift

- Also have angular and shape information



Terminology:

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

Need redshifts

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

Accurate but expensive

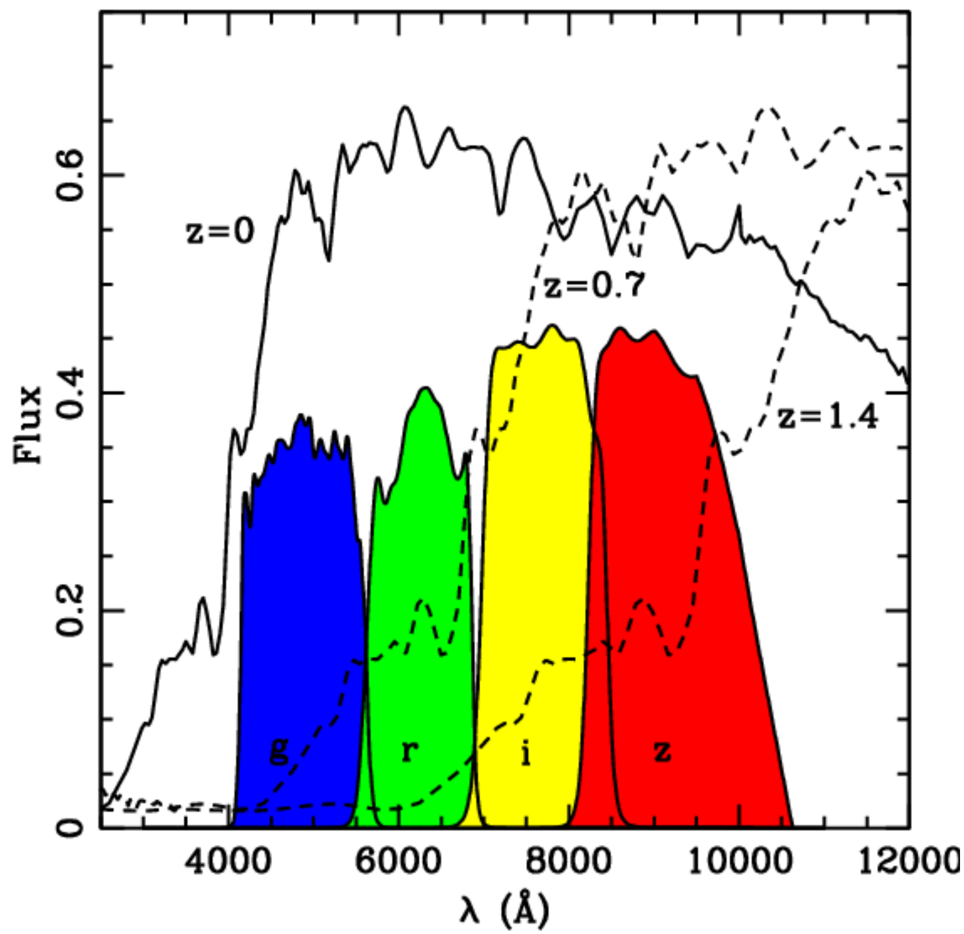
Innaccurate but cheap

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

Redshift issues

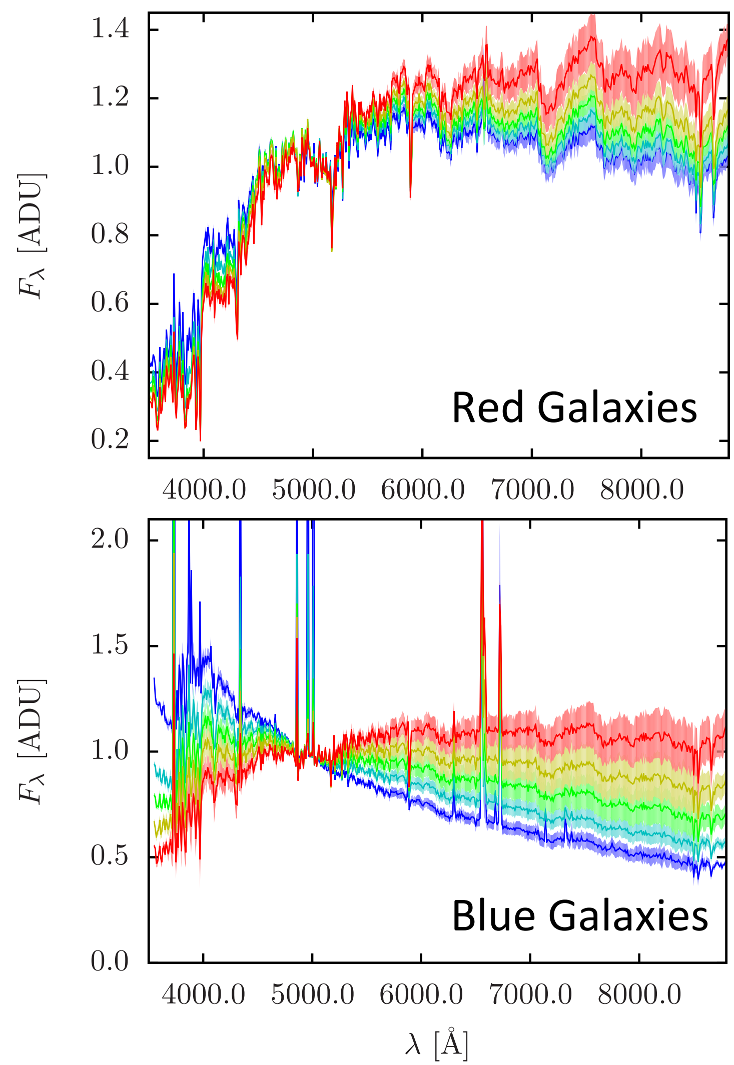
Basics of photo-zs

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



Basics of photo-zs

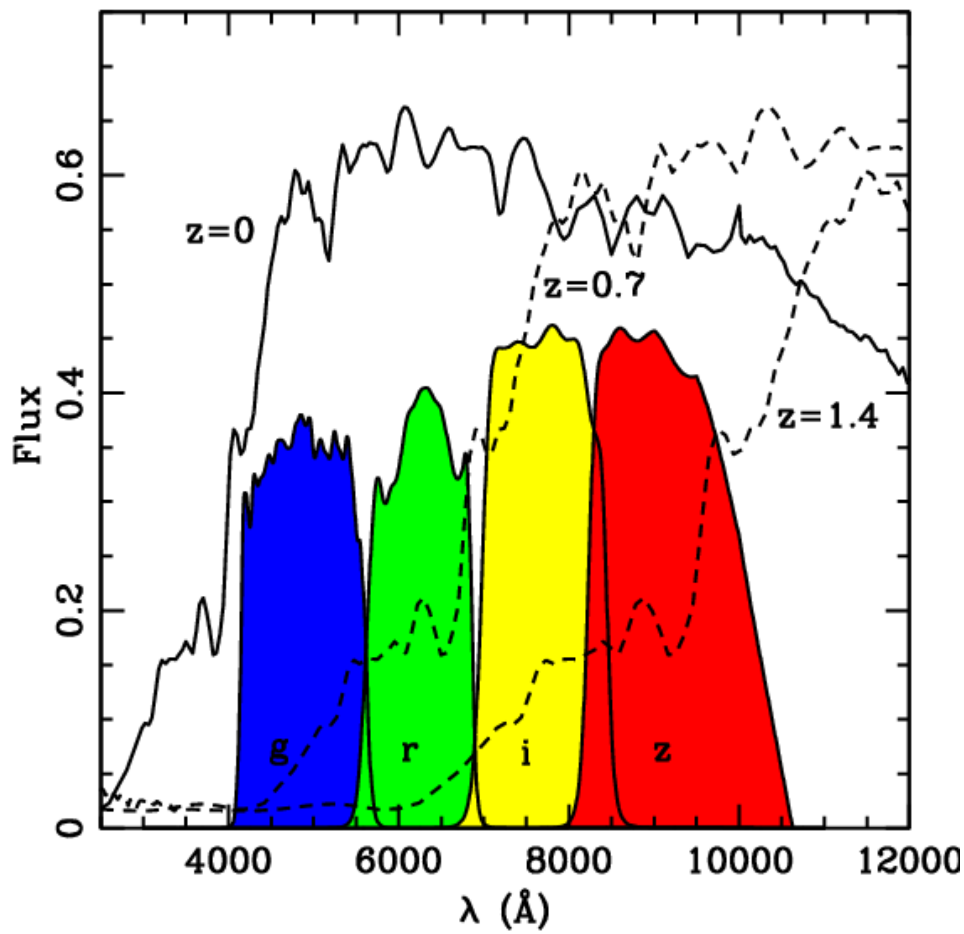
- A difficulty is that there is a distribution of spectral types.
- Galaxy spectra = sum of stellar spectra
- Stellar spectrum ≈ blackbody spectrum + absorption and emission



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.

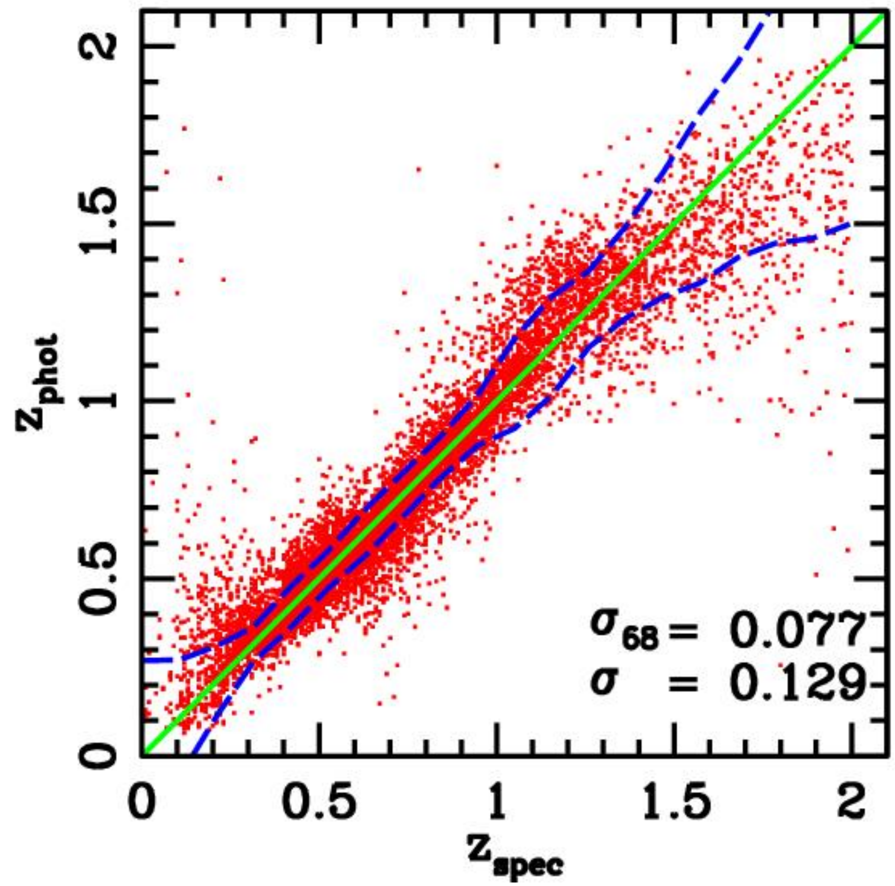


Basics of photo-z's

Photo-zs are often not very good.

<u>Three steps before getting to the cosmology:</u>

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

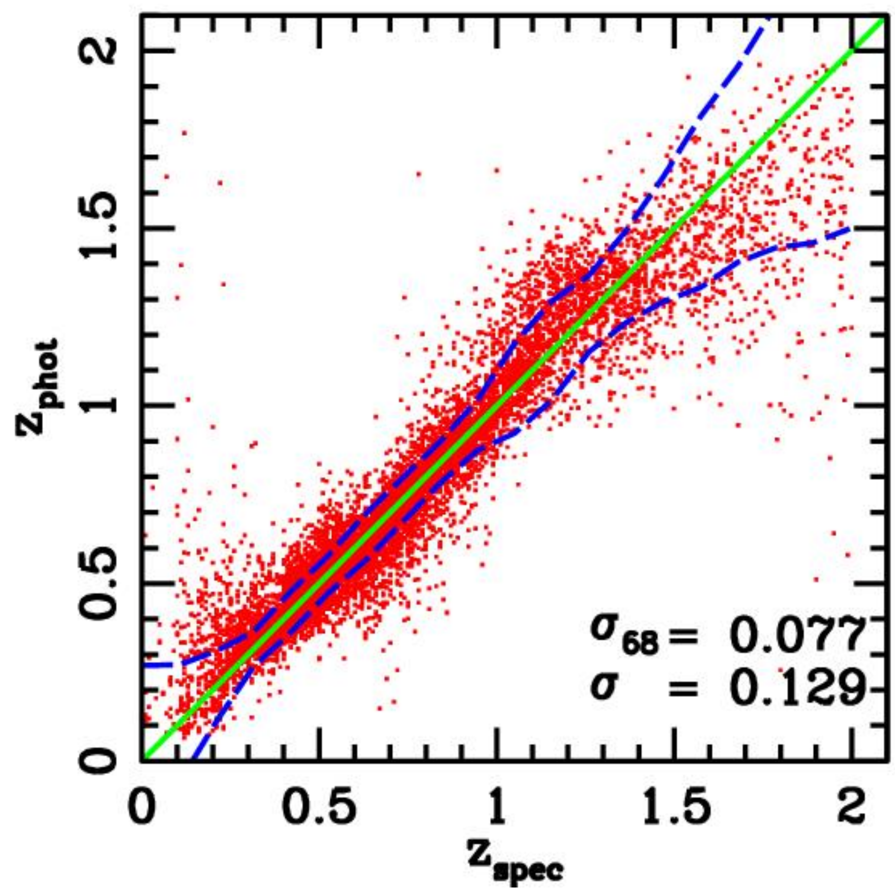


Basics of photo-z's

Photo-zs are often not very good.

<u>Three steps before getting to the cosmology:</u>

- **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.** (Cunha et al 2009)
- **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

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.** (Cunha et al 2009)
- **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

$N(z_{spec})$

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

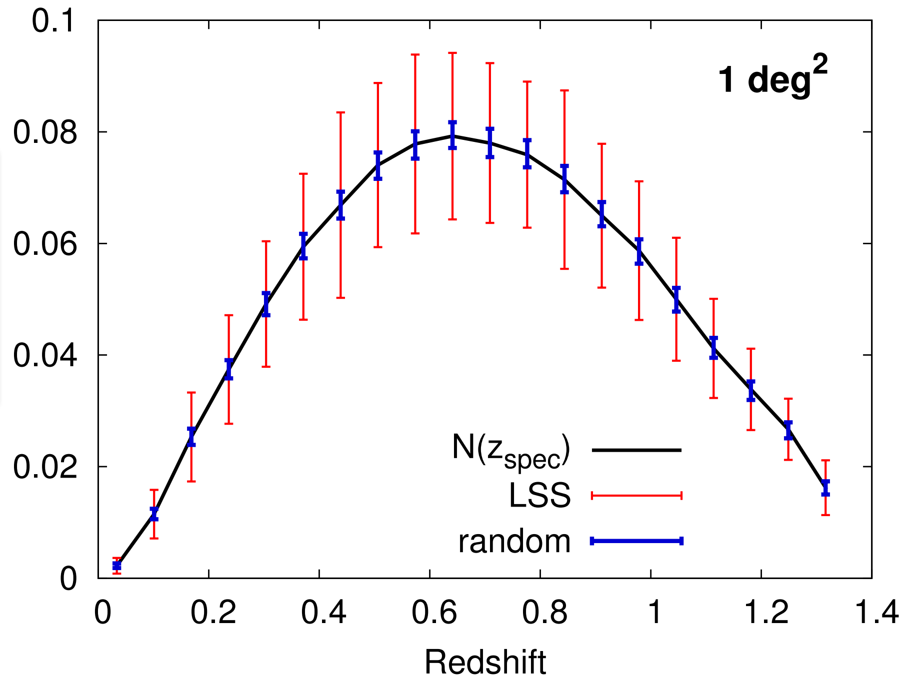


Figure 1. Normalized spectroscopic redshift distribution for the full data. The red (light gray) error bars show the 1-σ variability in the redshift distribution for contiguous 1 deg^2 angular patches. The blue (dark gray) error bars show the variability in the redshift distribution assuming random samples 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 et al, 2012a

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.** (Cunha et al 2009)
- **Shot-noise: need many galaxies**
- **Sample variance: need lots of area.**

Cunha et al. 2012a

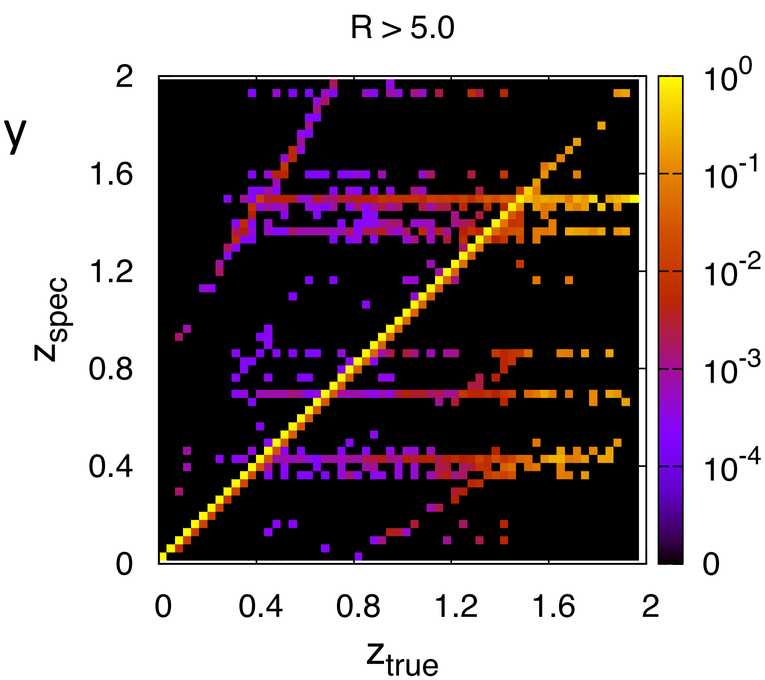
- **Spectroscopic failures:**
  - **Can't get spectra for certain galaxies.**
  - **Wrong spectroscopic redshifts.**

Cunha et al. 2012b

Spectroscopic failures (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.



Sample used in the plot has **98.6%** correct redshifts and constitutes **60%** of total sample (bias in w: O(100%)).

R: cross-correlation parameter (measures redshift confidence)

Case study: Simulations of DES photometry + VVDS-like spec-z's

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

Angular Issues

Growth of structure

- Want:

$$\delta = \frac{\rho - \overline{\rho}}{\overline{\rho}}$$

Growth of structure

- Want: $$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}}$$

- (Think you) get: $$\delta = \frac{N - \bar{N}}{\bar{N}}$$

N: number of galaxies, clusters, …

Growth of structure

- Want: $$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}}$$
- (Think you) get: $$\delta = \frac{N - \bar{N}}{\bar{N}}$$ N: number of galaxies, clusters, …
- Actually get: $$\delta_{obs} = \frac{N_{obs} - \bar{N}_{obs}}{\bar{N}_{obs}}$$

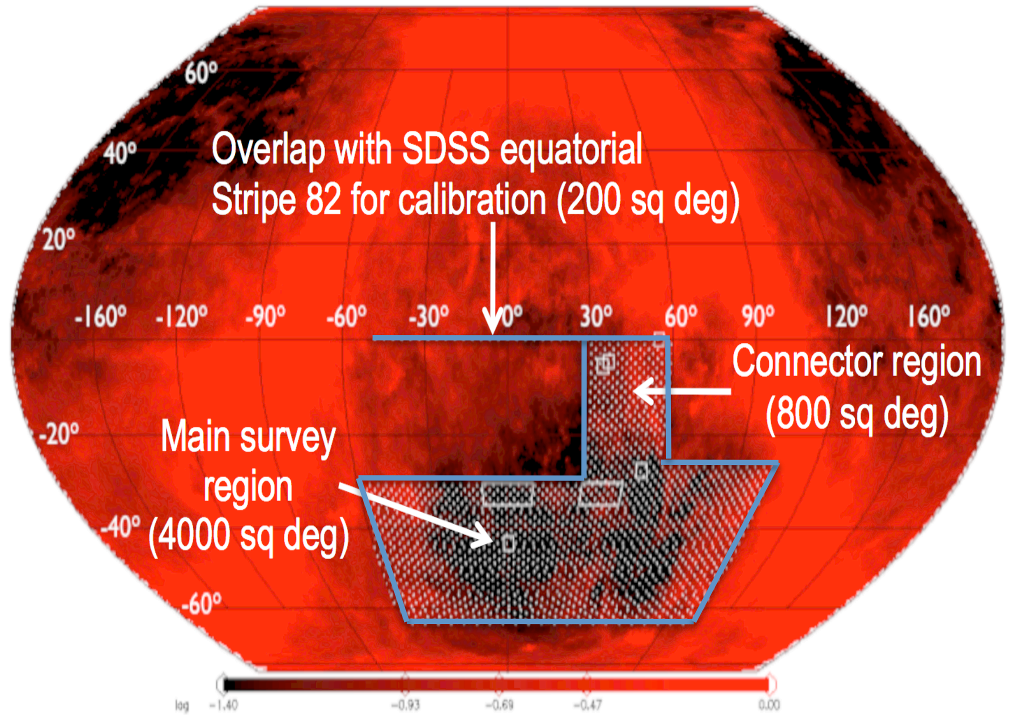
where $$N_{obs}(x) \equiv (1 + c(x))N(x)$$

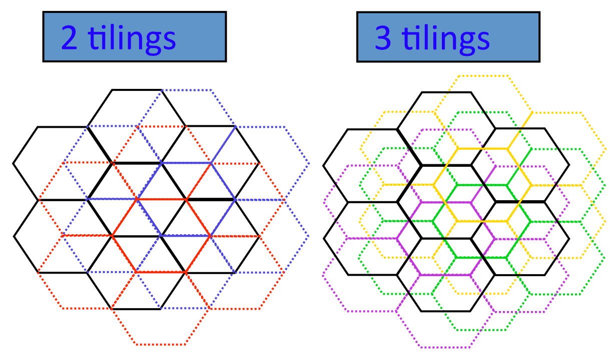
Selection or calibration errors

Sources of calibration error

Photometric calibration is complicated

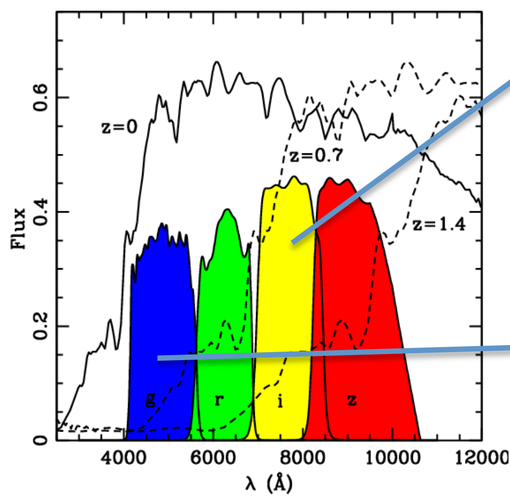
- Deal with: telescope/camera, atmosphere, seasons, Moon, Milky Way – over several years.
- Multiple overlapping tilings with varying orientations + standard stars + …



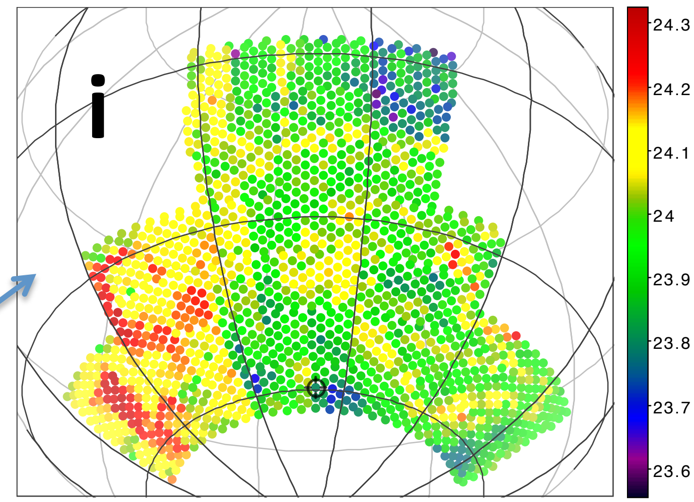


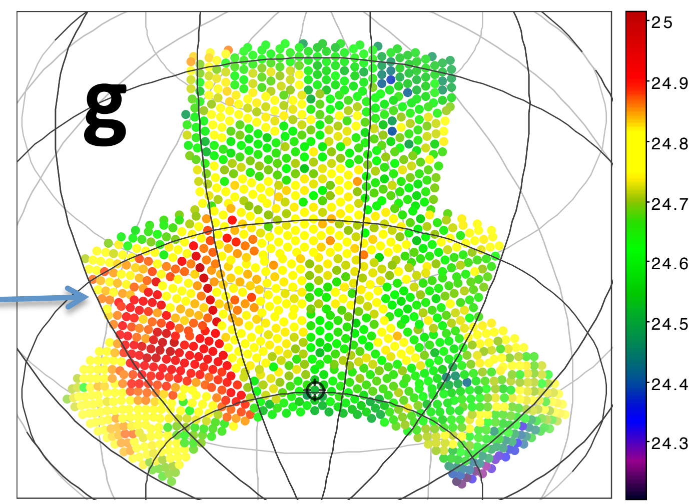
Photometric calibration is complicated

- Deal with: telescope/camera, atmosphere, seasons, Moon, Milky Way – over several years.
- Multiple overlapping tilings with varying orientations + standard stars + …



DES 5yr mag – limits



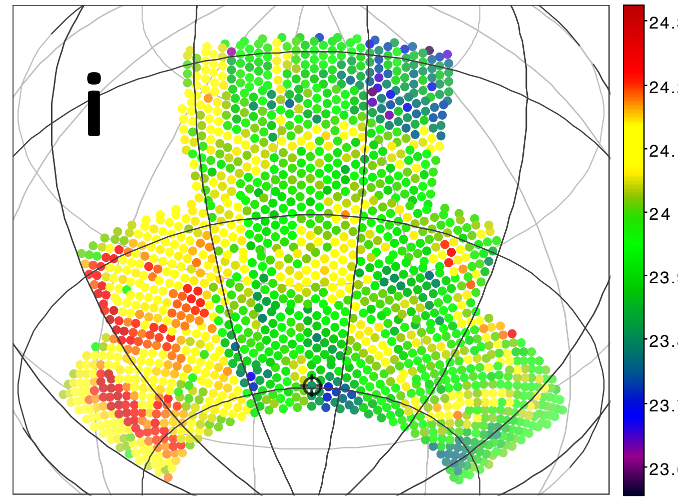


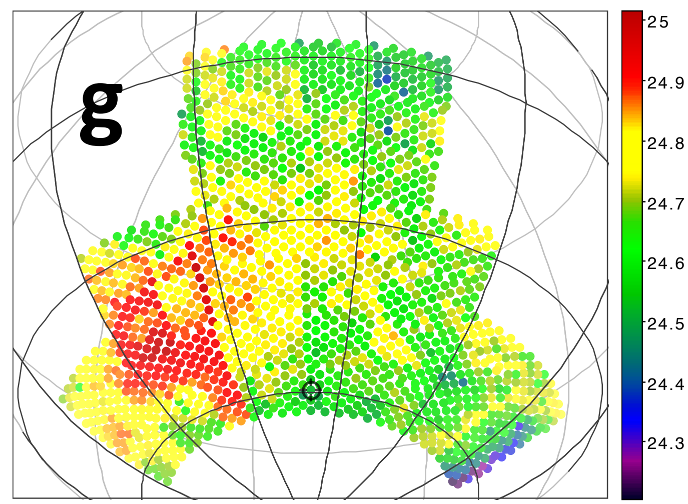
Simulations courtesy of Jim Annis

Life is complicated (even for perfect calibration)

- Mag. Limits affect N(θ)
- Mag. limits affect redshift distribution **-> coupling between angular and radial effects** (problem is more complicated if using photo-zs).

DES 5yr mag – limits



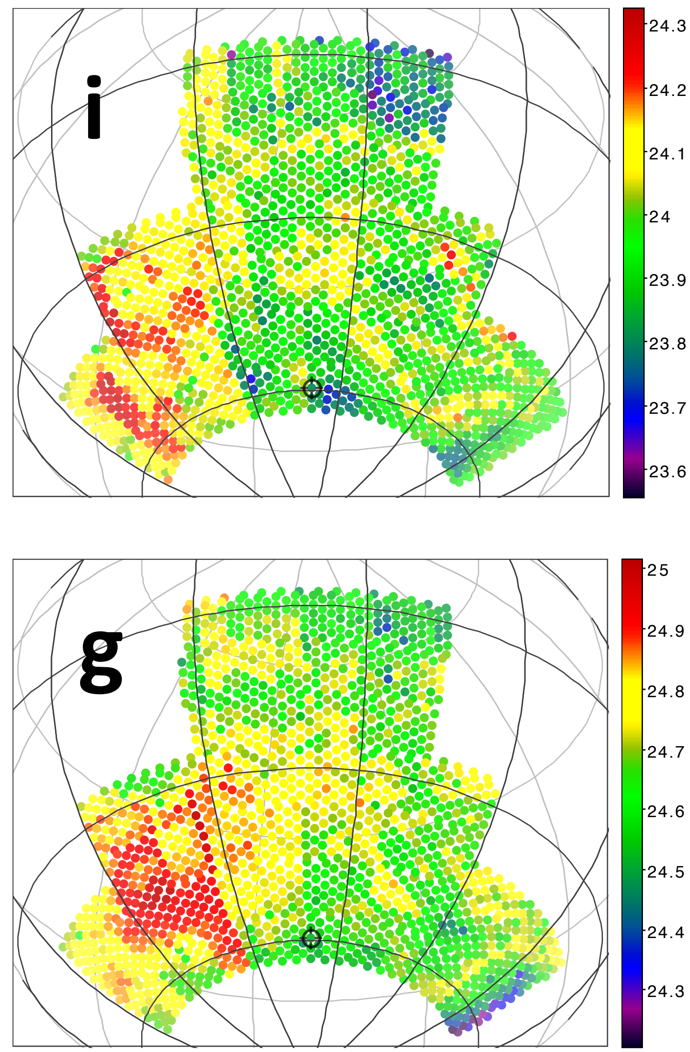


Simulations courtesy of Jim Annis

Life is complicated (even for perfect calibration)

- Mag. Limits affect N(θ)
- Mag. limits affect redshift distribution -> **coupling between angular and radial effects** (problem is more complicated if using photo-zs).
- Varying colors, affect galaxy types being selected.
  - Different types have different HODs, with different biases.
  - variation in color -> **scale-dependent** halo bias
- Need to couple radial-angular mask
- Uncertainty in calibration will still be a problem.

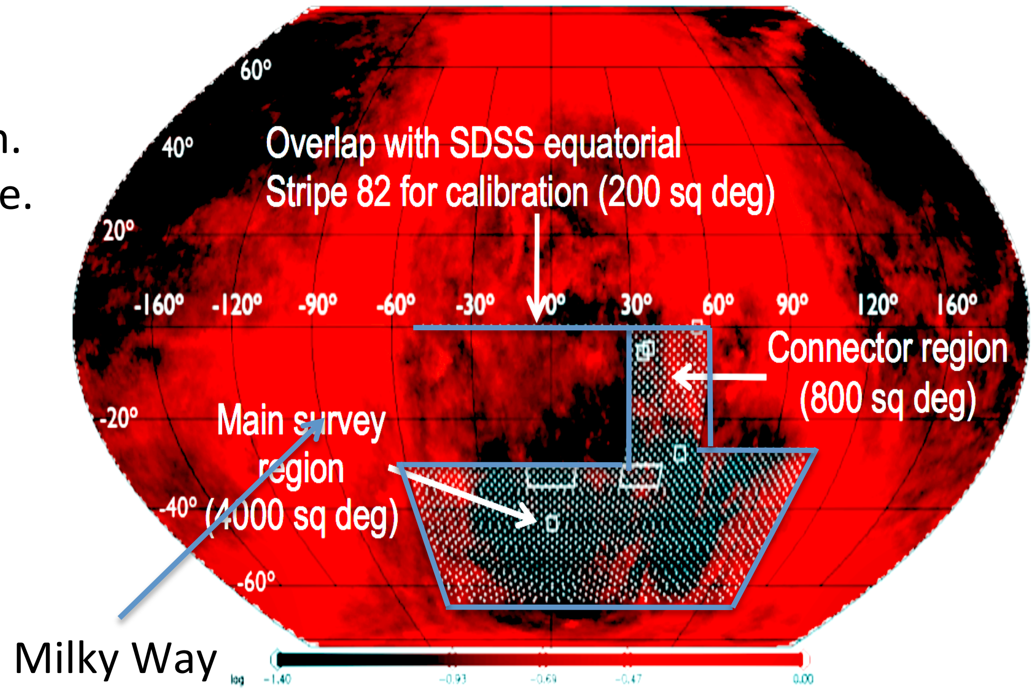
DES 5yr mag – limits



Simulations courtesy of Jim Annis

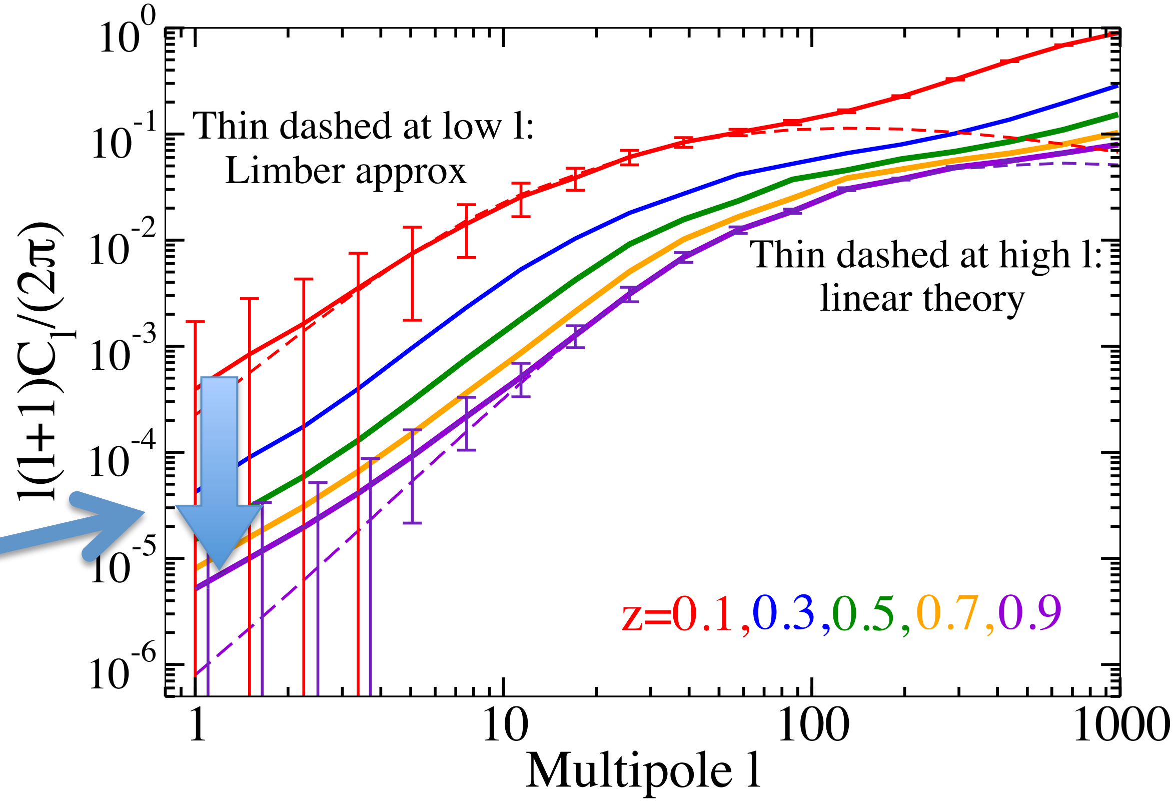
Star/Galaxy separation

- Distribution of stars is not random. Pronounced variation with latitude.
- Classification using colors (magnitudes)



Galaxy power-spectrum for DES

Power spectrum decreases with increasing redshift



Error bars: cosmic variance

Conclusions

- Calibration of wide surveys is a tough challenge.
- Requirements are more stringent for non-Gaussianity constraints.
- The larger scales are more sensitive to the calibration bias.
- Isotropy tests should provide important cross-check.

Modeling the calibration error

Modeling calibration issues

$$\delta(\hat{\mathbf{n}}) \equiv \frac{N(\hat{\mathbf{n}}) - \bar{N}(\hat{\mathbf{n}})}{\bar{N}(\hat{\mathbf{n}})} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}})$$

without calibration errors

$$\langle a_{\ell m} a^*_{\ell m} \rangle = \delta_{mm'} \delta_{\ell \ell'} C_\ell$$

Assuming statistical isotropy!

Modeling calibration issues

without calibration errors

$$\delta(\hat{\mathbf{n}}) \equiv \frac{N(\hat{n}) - \overline{N}}{\overline{N}} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}})$$

$$\langle a_{\ell m} a^*_{\ell m} \rangle = \delta_{mm'} \delta_{\ell\ell'} C_\ell$$

with calibration errors

$$\delta^{\mathrm{obs}}(\hat{\mathbf{n}}) \equiv \frac{N_{obs}(\hat{n}) - \overline{N}_{obs}}{\overline{N}_{obs}} = \sum_{\ell m} t_{\ell m} Y_{\ell m}(\hat{\mathbf{n}})$$

$$N_{\mathrm{obs}}(\hat{\mathbf{n}}) = [1 + c(\hat{\mathbf{n}})]\, N(\hat{\mathbf{n}}) \qquad c(\hat{\mathbf{n}}) = \sum_{\ell_1=0}^{\ell_{\mathrm{calib,max}}} \sum_{m_1=-\ell_1}^{\ell_1} c_{\ell_1 m_1} Y_{\ell_1 m_1}(\hat{\mathbf{n}})$$

Modeling calibration issues

without calibration errors

$$\delta(\hat{\mathbf{n}}) \equiv \frac{N(\hat{n}) - \overline{N}}{\overline{N}} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}})$$

$$\langle a_{\ell m} a^*_{\ell m} \rangle = \delta_{mm'} \delta_{\ell\ell'} C_\ell$$

with calibration errors

$$\delta^{\rm obs}(\hat{\mathbf{n}}) \equiv \frac{N_{obs}(\hat{n}) - \overline{N}_{obs}}{\overline{N}_{obs}} = \sum_{\ell m} t_{\ell m} Y_{\ell m}(\hat{\mathbf{n}})$$

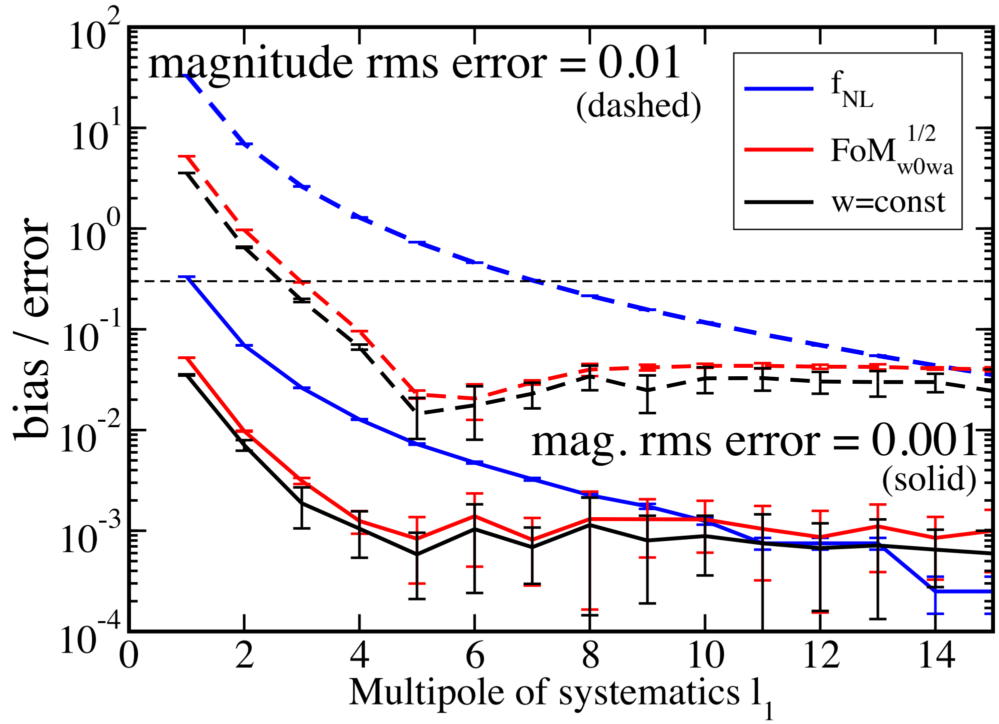
$$N_{\rm obs}(\hat{\mathbf{n}}) = [1 + c(\hat{\mathbf{n}})]\, N(\hat{\mathbf{n}}) \qquad c(\hat{\mathbf{n}}) = \sum_{\ell_1=0}^{\ell_{\rm calib,max}} \sum_{m_1=-\ell_1}^{\ell_1} c_{\ell_1 m_1} Y_{\ell_1 m_1}(\hat{\mathbf{n}})$$

$$\langle t_{\ell m} t^*_{\ell' m'} \rangle = \frac{1}{(1+\epsilon)^2} \left\{ \underbrace{\delta_{mm'}\delta_{\ell\ell'} C_\ell}_{\text{isotropic}} + \underbrace{\left[U^{\ell\ell'}_{mm'} C_{\ell'} + (U^{\ell\ell'}_{mm'})^* C_\ell \right] + \sum_{\ell_2 m_2} U^{\ell_2 \ell}_{m_2 m} (U^{\ell_2 \ell'}_{m_2 m'})^* C_{\ell_2} + c_{\ell m} c^*_{\ell' m'}}_{\text{breaks statistical isotropy}} \right\}$$

Bias from calibration error in single multipole

Error bars: variations from allocatting e_{calib} to different m.

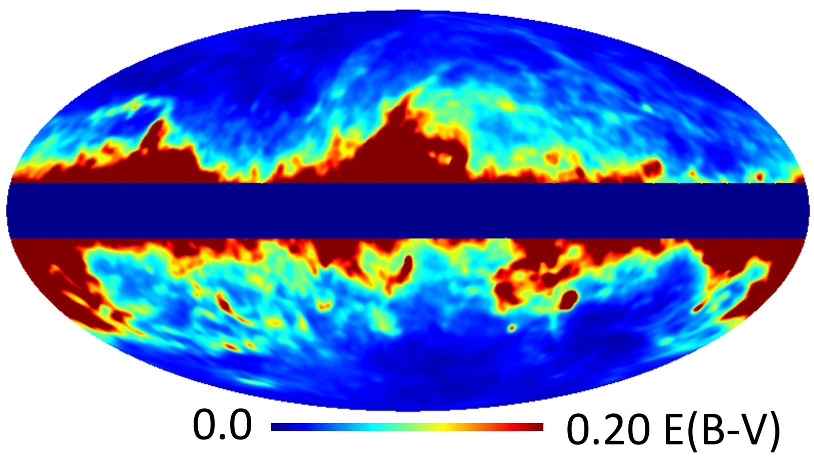
Changes in $m_{lim}(\theta)$ imply changes in N (θ,z)



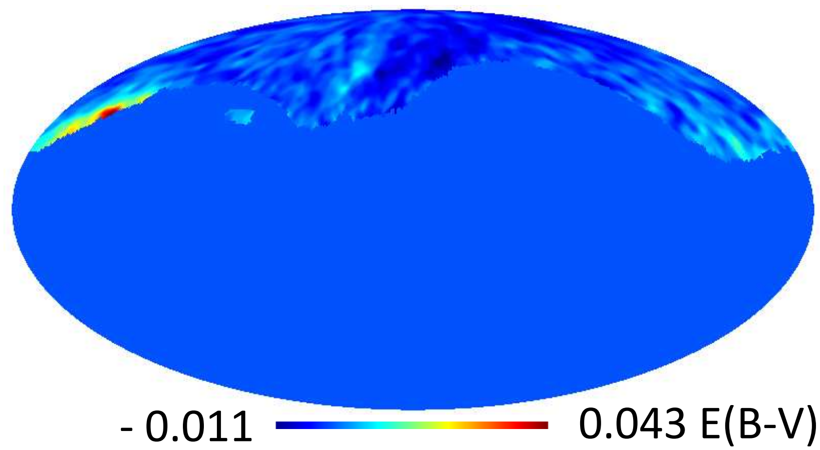
Huterer, Cunha & Fang, 2013, MNRAS.

Example: Dust extinction

Schlegel, Finkbeiner & Davis
(SFD) extinction maps

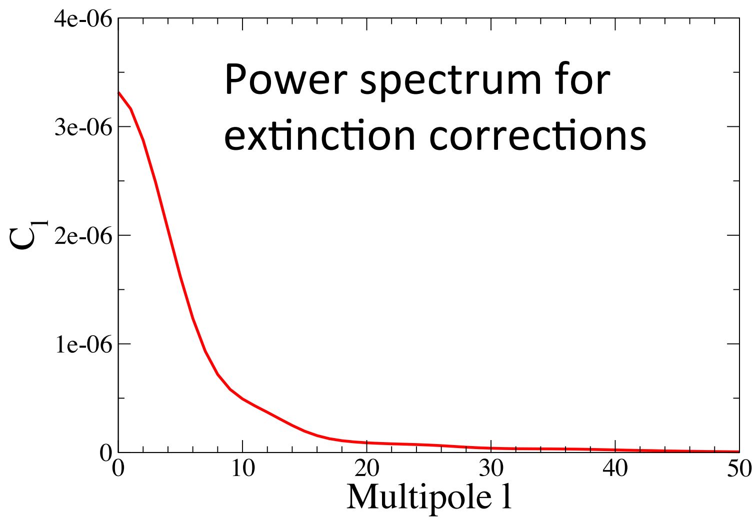


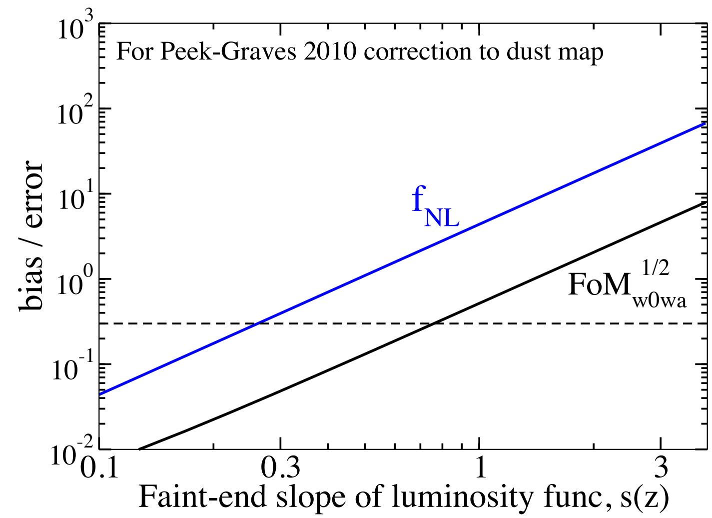
Peek & Graves corrections
to extinction maps



Simple test: Assume Peek & Graves corrections are the calibration error

Example: Dust extinction





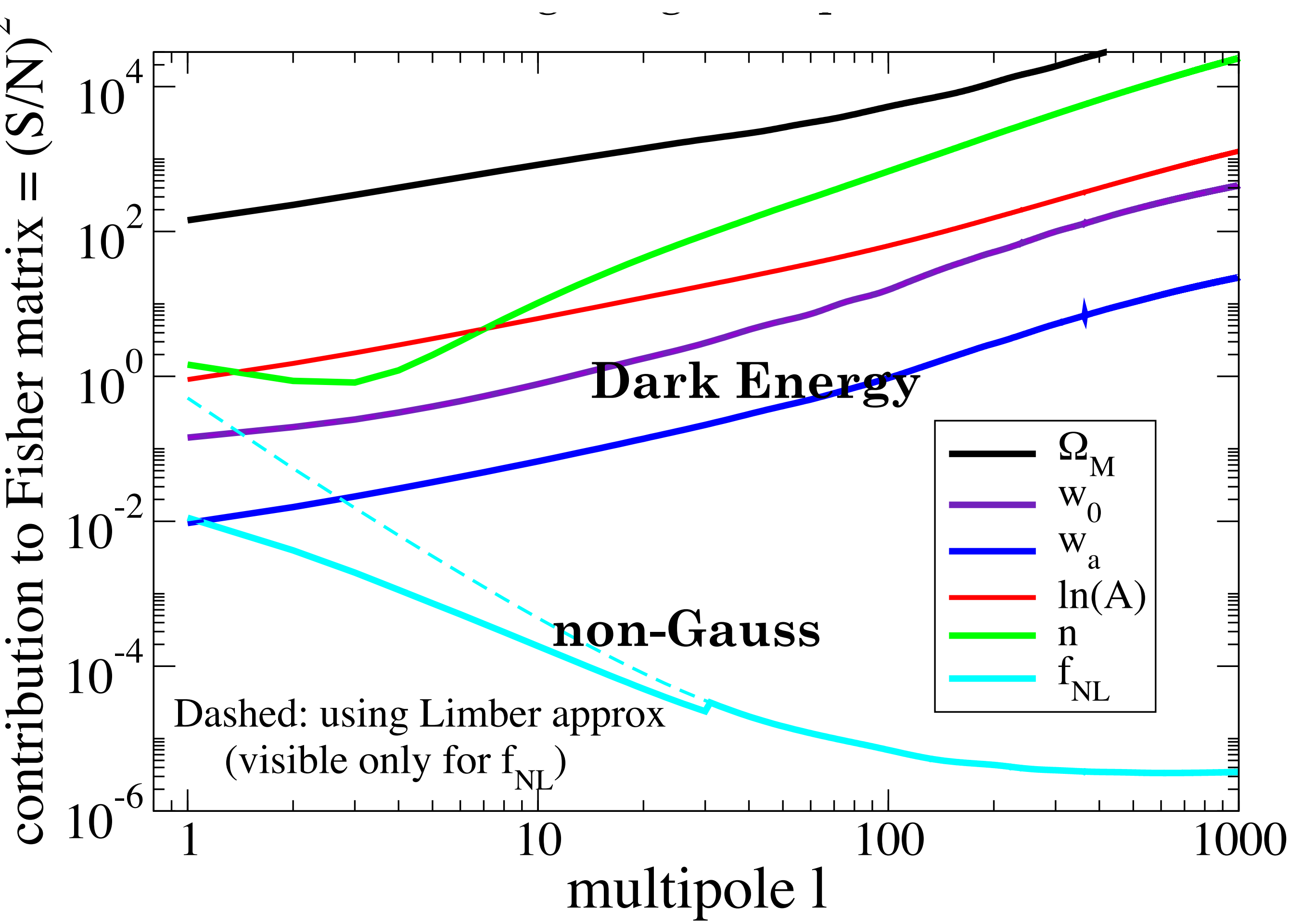
$$s(z) = \left.\frac{d\log_{10} N}{dm}\right|_{m_{\rm lim}}$$

From COSMOS sims,
at i=24:

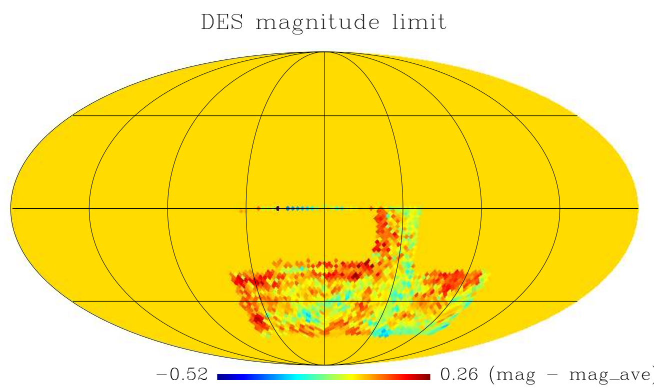
$$s(z) = 0.094 + 0.155z + 0.165z^2$$

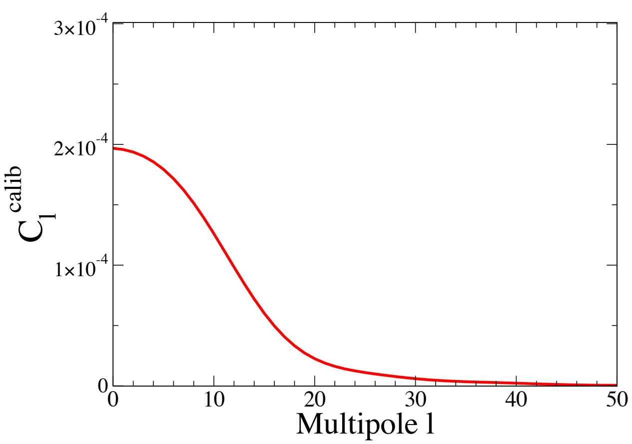
$$\delta p_i \propto \delta C_\ell \propto |c_{\ell m}|^2 \propto s(z)^2$$

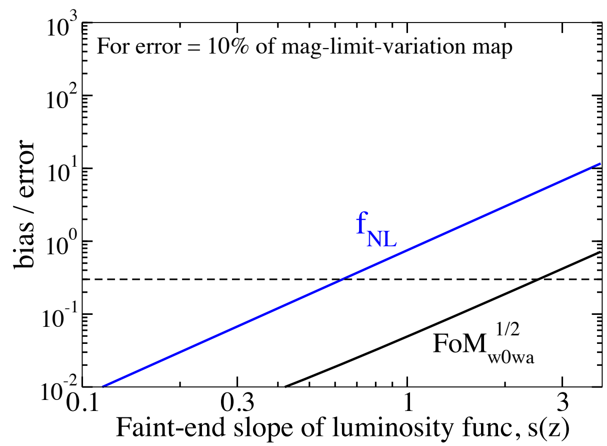
Where is the information and how much is there?



Example II: DES i-band limits







Observational issues for f_{nl} measurement

- Artificial correlations can mimic f_{nl}. For $f_{\rm NL}^{\rm local}$, separations >100 Mpc (several degrees) are crucial.

Because of $1/k^2$ scale dependence of bias

$$b(k) = b_G + f_{NL} \frac{const}{k^2}$$

- Artificial correlations can be due to:
  - photometric calibration
  - photometric redshifts
  - star/galaxy separation

  (photometric redshifts and star/galaxy separation:) More relevant for galaxies than clusters

- Clusters have own selection issues

The Dark Energy Survey

www.darkenergysurvey.org

- 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%)

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

$\neq$

.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

≠

.111	.667	.222
0.4	0.2	0.4