Refining Photometric Redshift Distributions with Cross-Correlations

Alexia Schulz
Institute for Advanced Study

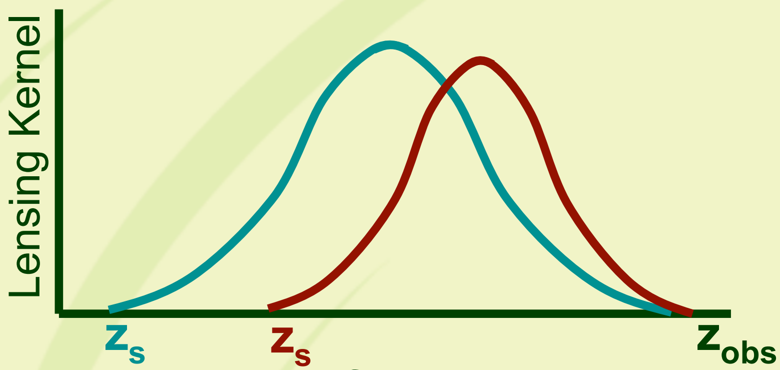
Collaborators: Martin White

Talk Overview

- Introduction
  - Weak lensing tomography can improve cosmological constraints
  - Accurate source galaxy redshift distribution is essential
  - Calibration with cross-correlations may be a better way to calibrate it
- Method
  - The idea and how it works
  - N-body simulations used to test it
- Results
  - Spectroscopic sub-sample
  - Other tracers of the dark matter
  - Relationship to galaxy bias
- Conclusions
  - Advantages
  - Outstanding theoretical questions

Why is the source galaxy distribution important for weak lensing?

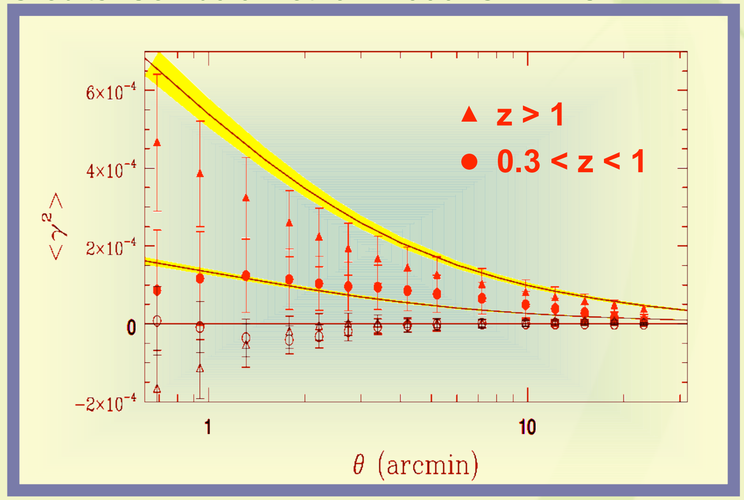
- Sources at different redshifts are lensed differently



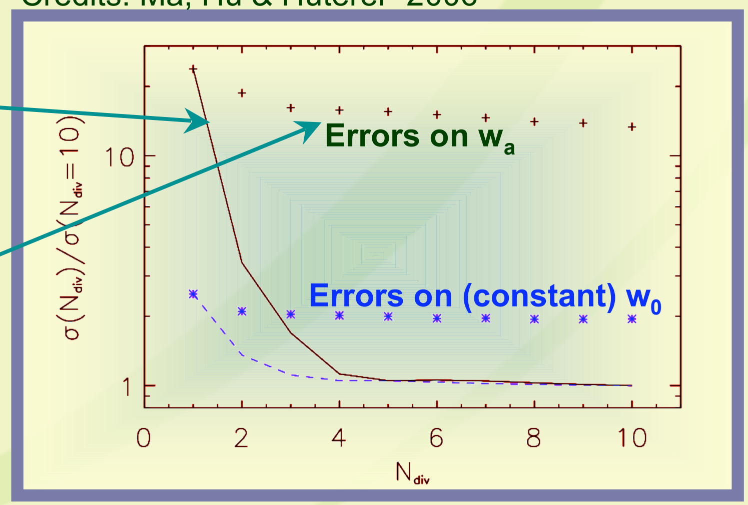
- Using just a few source redshift bins can drastically improve constraints on parameters (solid lines)

- Improvement is close to nothing if source distribution is unconstrained (points)

Credits: Semboloni et. al. 2006 CFHTLS

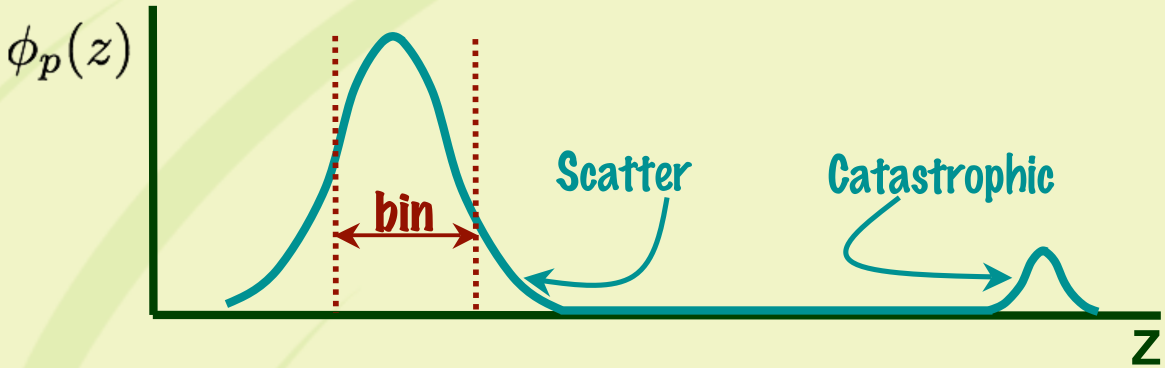


Credits: Ma, Hu & Huterer 2006



What if there are unknown errors in the photometry?

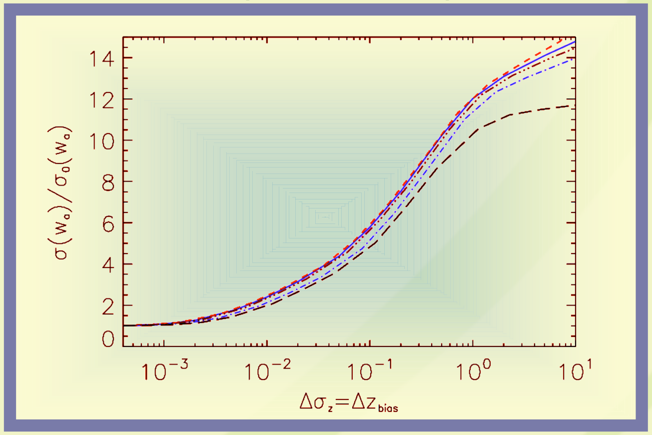
- In a photometric bin, the true redshift distribution has some scatter, and at some redshifts catastrophic errors



- Weak lensing tomography is very sensitive to (unknown) errors in the source galaxy distribution

Fractional increase in the constraint on w_a versus prior knowledge of scatter and offset in mean redshift

Credits: Ma, Hu & Huterer 2006



- To use tomography effectively, the true distribution must be well understood

Why use the cross-correlation technique?

- Large upcoming photometric galaxy surveys (DES, DUNE, Hyper SupremeCam, LSST, PanStarrs, PAU, SNAP, Vista and others) will increase the current number of survey galaxies by 2 orders of magnitude

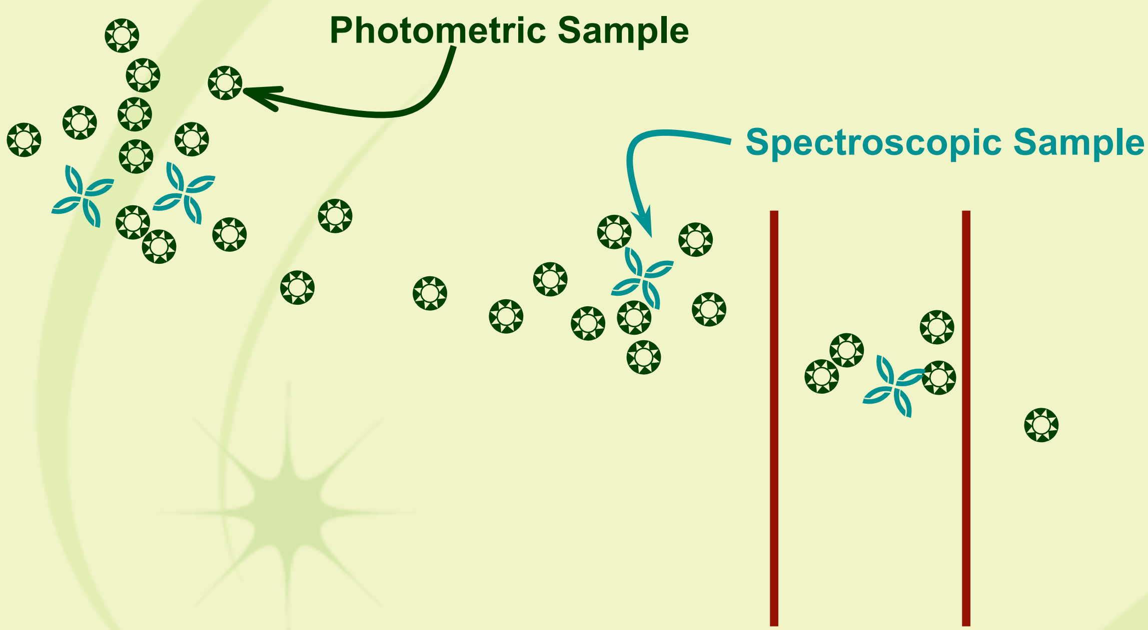
- Some of the surveys will be very faint, and very deep

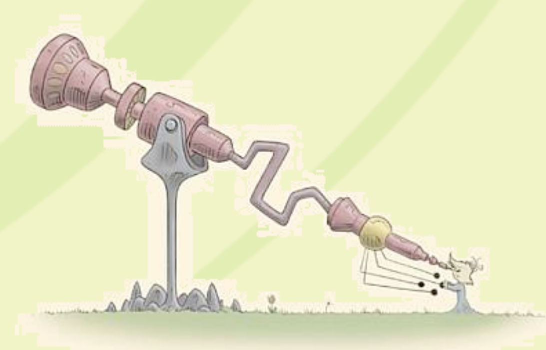
- Spectroscopic follow-up of a calibration sample could be very time consuming and expensive -- not very practical

- The cross-correlation technique proposed most recently by J. Newman (http://astron.berkeley.edu/%7Ejnewman/xcorr/xcorr.pdf) may not require such extensive spectroscopic follow up observations
  - Any other tracer (LRGs, quasars, etc.) as calibrating sample
  - May be able to use legacy data (SDSS)
  - Requirement: spatially overlapping

How does it work? (Picture people)

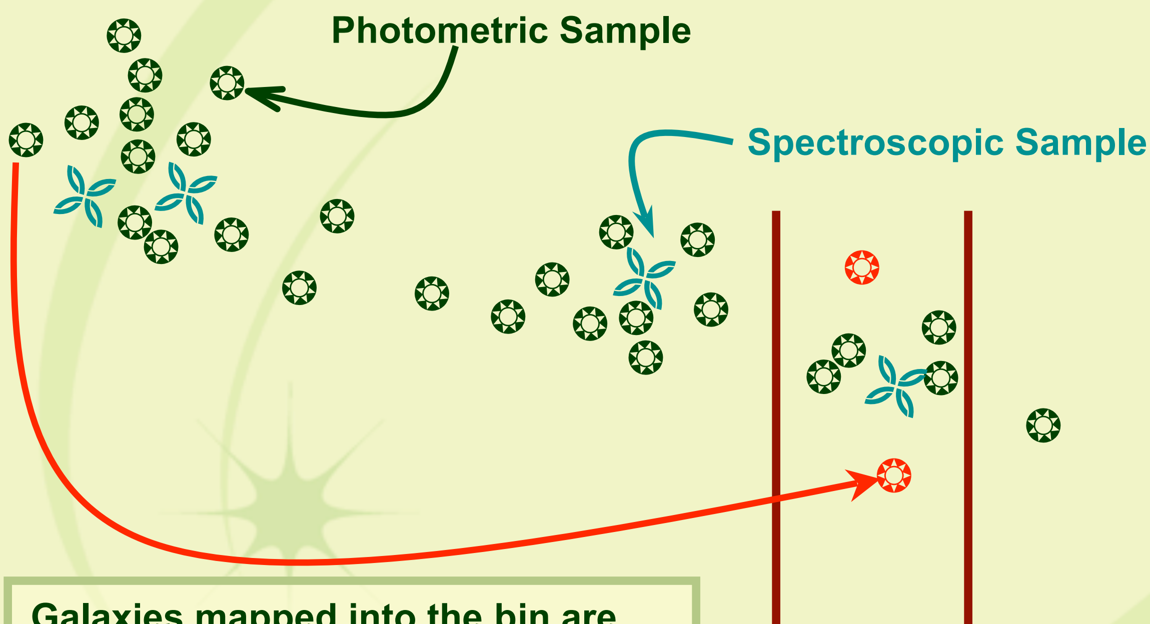
- Objects in the photometric and spectroscopic samples trace the same underlying dark matter structures
- Positions of photometric objects will be correlated with those of spectroscopic objects at the same redshift (physically close)



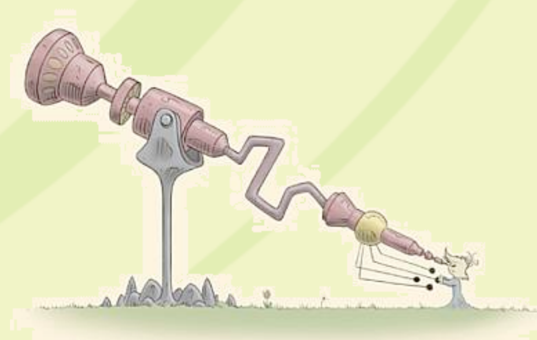


How does it work? (Picture people)

- Objects in the photometric and spectroscopic samples trace the same underlying dark matter structures
- Positions of photometric objects will be correlated with those of spectroscopic objects at the same redshift (physically close)

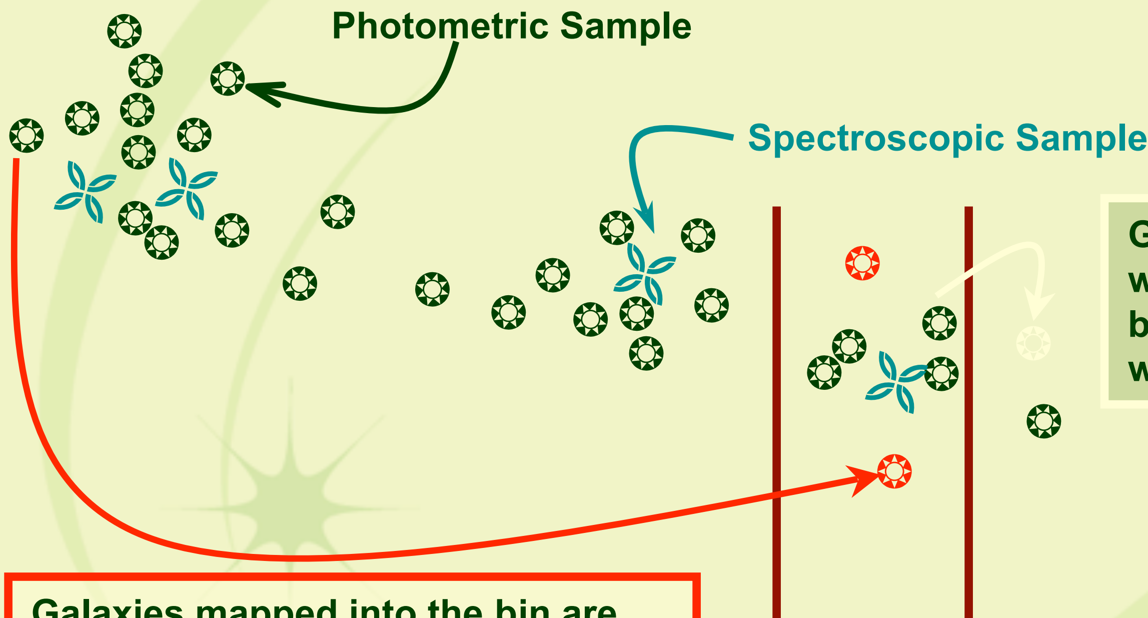


Galaxies mapped into the bin are uncorrelated with positions of spectroscopic objects ➔ they add to the noise but not the signal



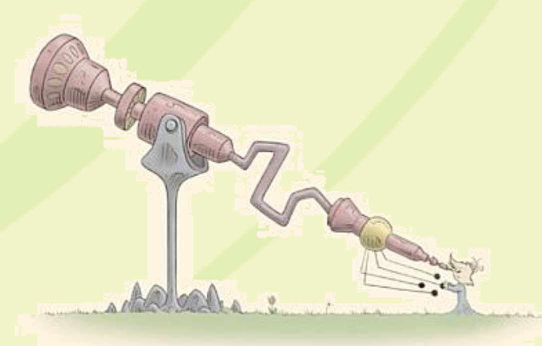
How does it work? (Picture people)

- Objects in the photometric and spectroscopic samples trace the same underlying dark matter structures
- Positions of photometric objects will be correlated with those of spectroscopic objects at the same redshift (physically close)



Galaxies mapped out of the bin will contribute to the signal in the bin because they are correlated with spectroscopic positions

Galaxies mapped into the bin are uncorrelated with positions of spectroscopic objects ➔ they add to the noise but not the signal



How does it work? (Equation people)

- The fundamental assumption relates the 3D cross-correlation function to the observable 3D autocorrelation function of the spectroscopic sample

2D Angular cross-correlation

3D cross-correlation function

Photometric selection function

$$\omega_{ps}(\theta, z_s) = \int dz\, \xi_{ps}(z, z_s, \theta)\, \phi_p(z)$$

At large (linear) scales assume: $\xi_{ps}(z, z_s, \theta) \propto \xi_{ss}(z, z_s, \theta)$

$$\omega_{ps}(\theta, z_s) \propto \int dz\ \xi_{ss}(z, z_s, \theta)\ \ \phi_p(z)$$

How does it work? (Equation people)

- The fundamental assumption relates the 3D cross-correlation function to the observable 3D autocorrelation function of the spectroscopic sample

2D Angular cross-correlation — 3D cross-correlation function — Photometric selection function

$$\omega_{ps}(\theta, z_s) = \int dz\, \xi_{ps}(z, z_s, \theta)\, \phi_p(z)$$

At large (linear) scales assume: $\xi_{ps}(z, z_s, \theta) \propto \xi_{ss}(z, z_s, \theta)$

$$\omega_{ps}(\theta, z_s) \propto \int dz\, \xi_{ss}(z, z_s, \theta)\, \phi_p(z)$$

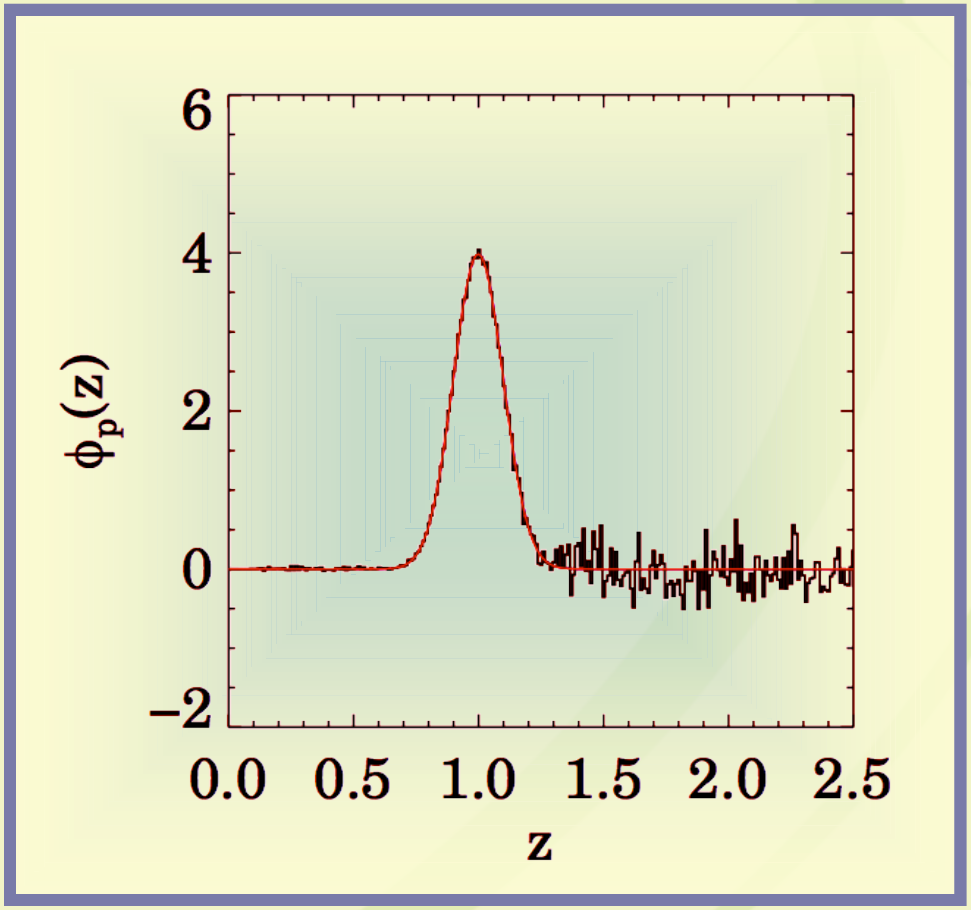
Observables: $\omega_{ps}(\theta, z_s)$ and $\xi_{ss}(z, z_s, \theta)$

Invert integral to solve for the selection function: $\phi_p(z)$

Tests with MCMC

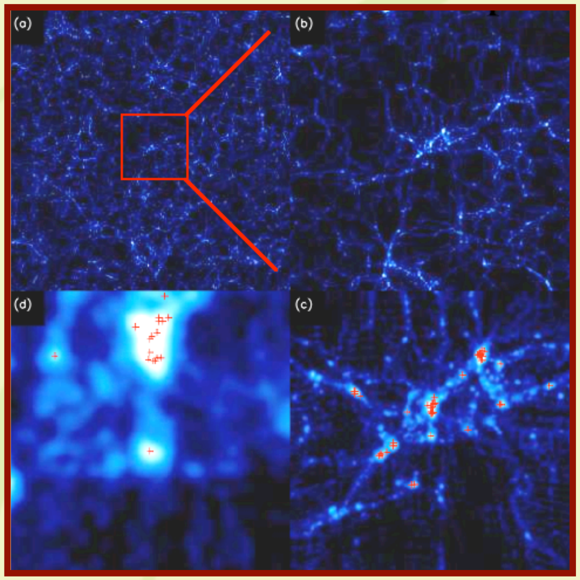
- 10,000 Monte Carlo realizations of the photometric distribution have been used to test the reconstruction technique.
- Each realization split into redshift bins
- Error has been added to the true $\phi_p(z)$ for each bin
- The technique seems promising

http://astron.berkeley.edu/%7Ejnewman/xcorr/xcorr.pdf



Test with N-body simulations

- Perform cosmological dark matter simulations



Volume: 1 $(\mathrm{Gpc}/h)^3$

Dark matter halos: 7.5 Million

Halo masses: $M > 5 \times 10^{11}\ M_\odot/h$

Central galaxy at gravitational potential minimum

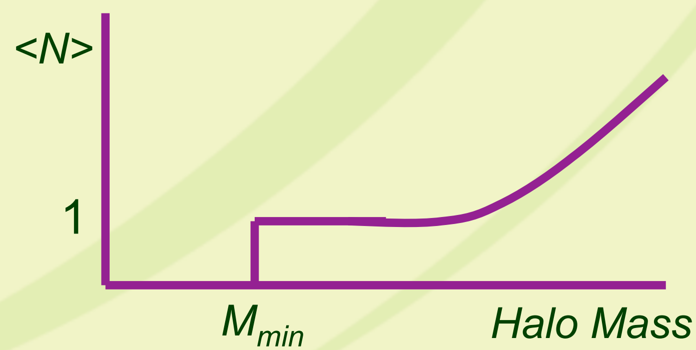
Satellite galaxies trace dark matter

Satellite galaxies Poisson distributed about $\langle N_s \rangle$

- Use halo model to populate with "galaxy" populations of differing biases

One parameter family for $\langle N_{gal} \rangle$

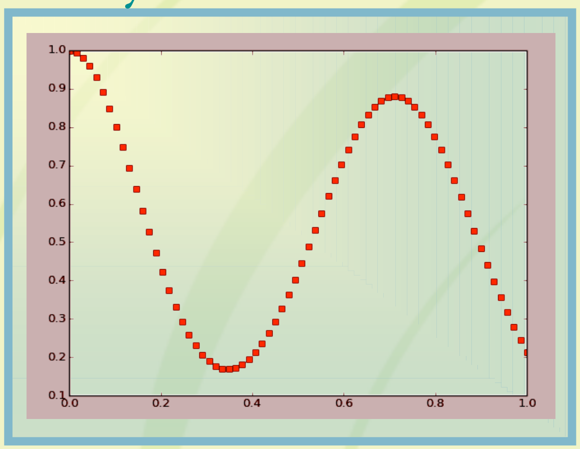
$$\langle N_{\rm gal}(M)\rangle = \Theta(M - M_{\rm min})\left(1 + \frac{M - M_{\rm min}}{10\, M_{\rm min}}\right)$$



Test with N-body simulations

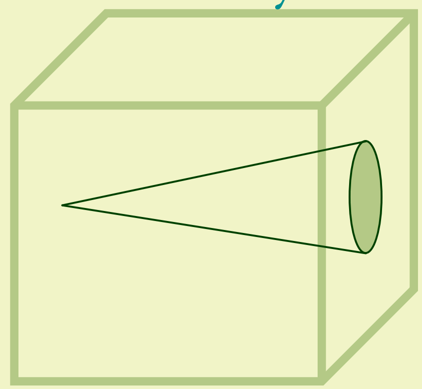
- Impose a selection function and survey geometry on the mock photometric sample to be calibrated (these lead to a $\phi_p(z)$)

Wacky selection function



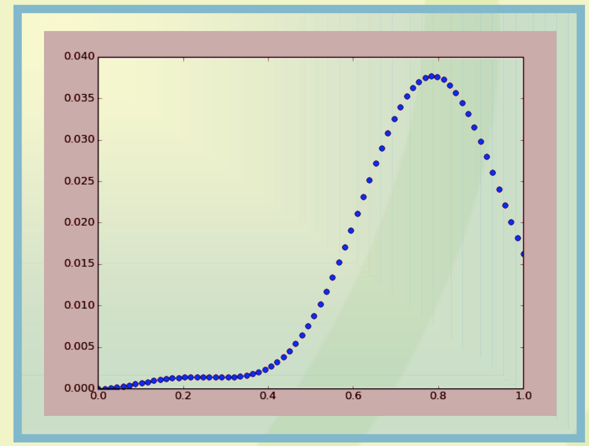
+

Geometry



→

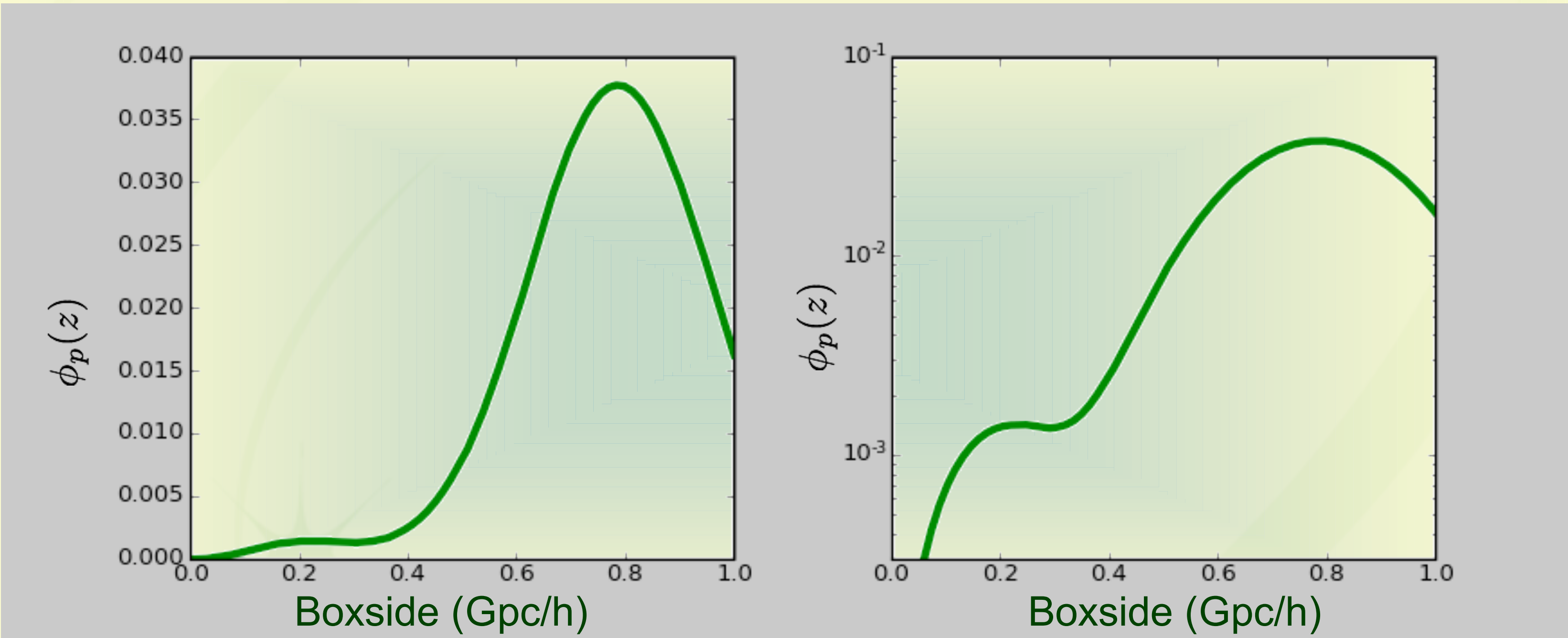
Photometric distribution



- Compute $\omega_{ps}(\theta, z_s)$ and $\xi_{ss}(\theta, z, z_s)$, and invert the relation to obtain the reconstructed photometric distribution $\phi_p(z)$
- Compare to the input photometric redshift distribution
- Test 1: A fair sub-sample of the photometric galaxies $(\xi_{ss} = \xi_{ps})$
- Test 2: A completely unrelated class of tracer with a different bias
- Test 3: Tests 1 and 2 in a light cone with evolving ξ_{ss} and ξ_{ps}

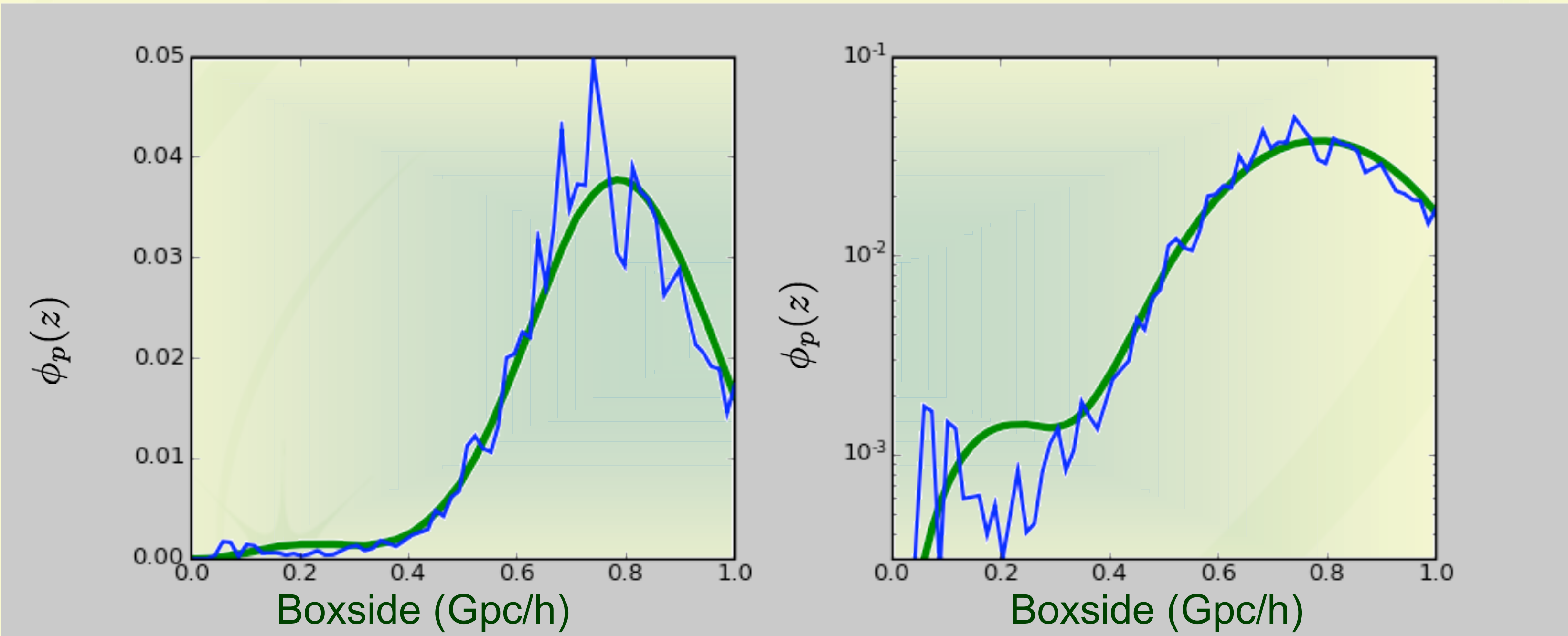
Test 1: Spectroscopic sub-sample

Input redshift distribution



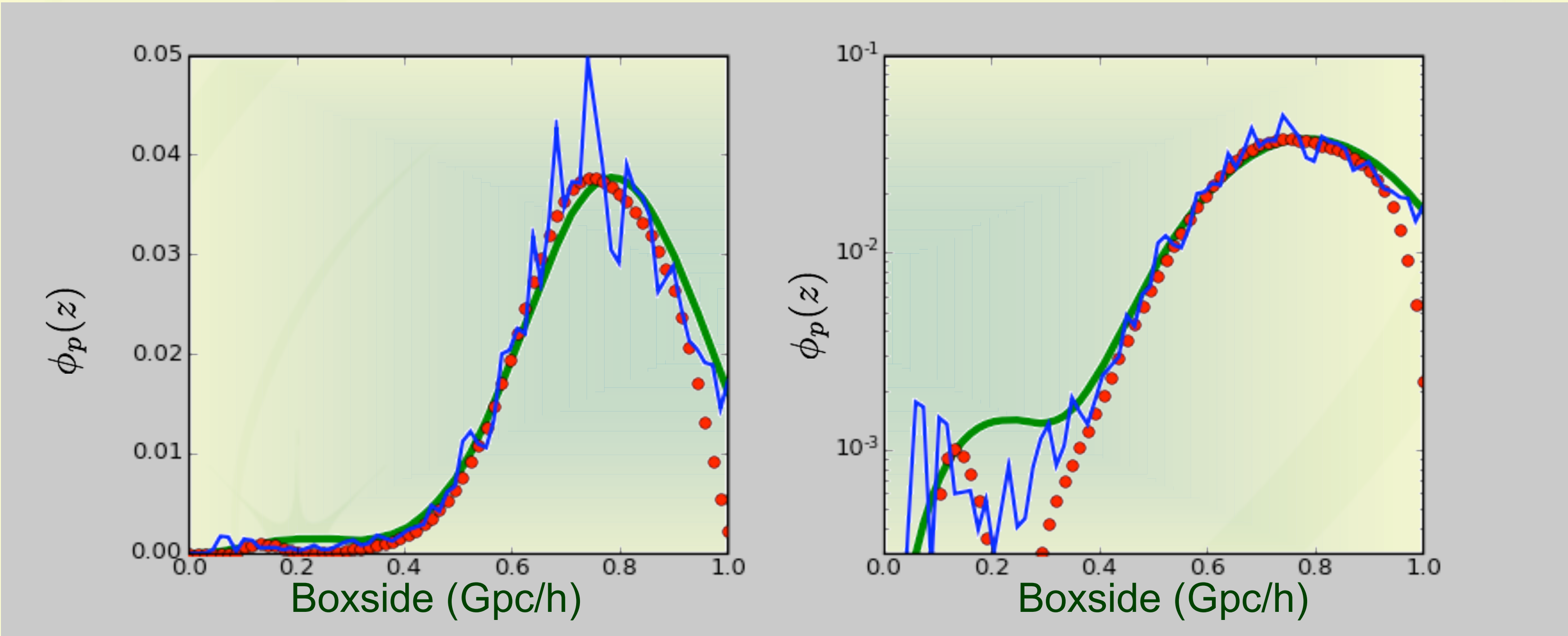
Test 1: Spectroscopic sub-sample

One realization of galaxies in halos (M_{min}=5e11 $M_{\odot}$/h)



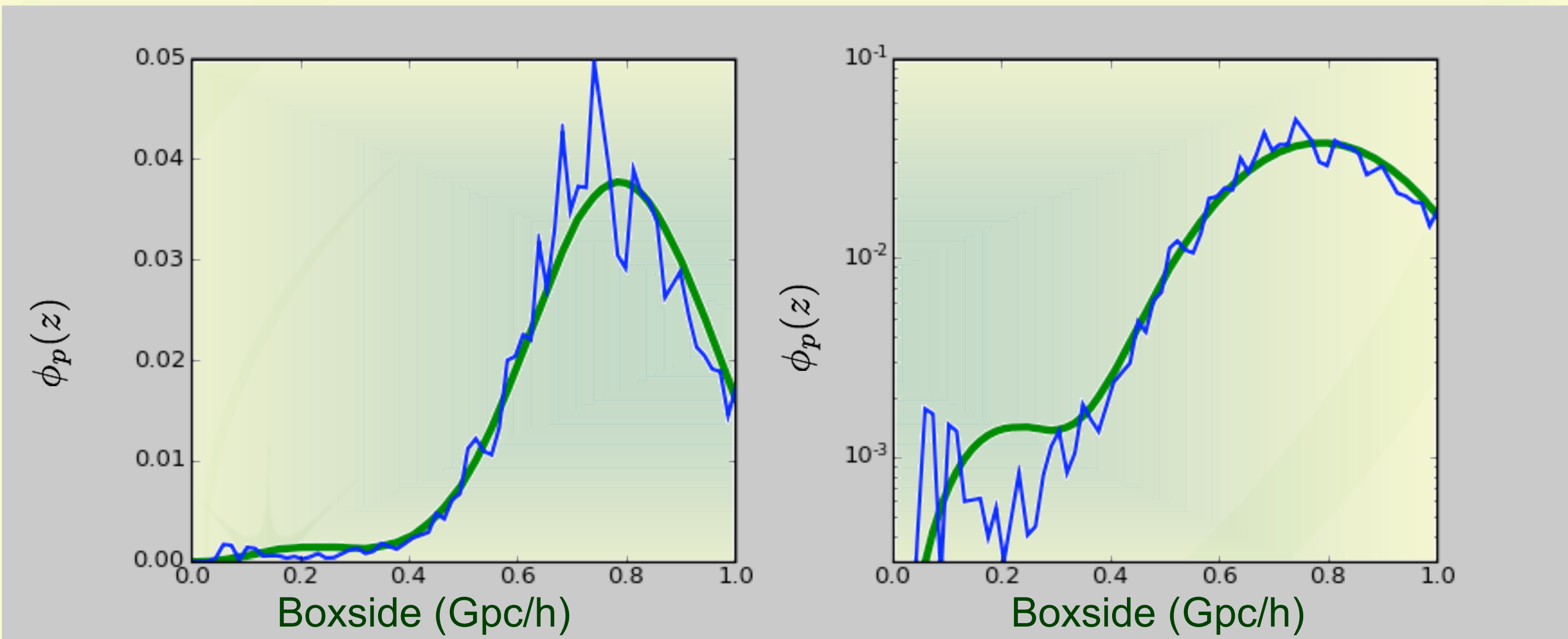
Test 1: Spectroscopic sub-sample

Our reconstruction using 1/10 randomly selected subsample



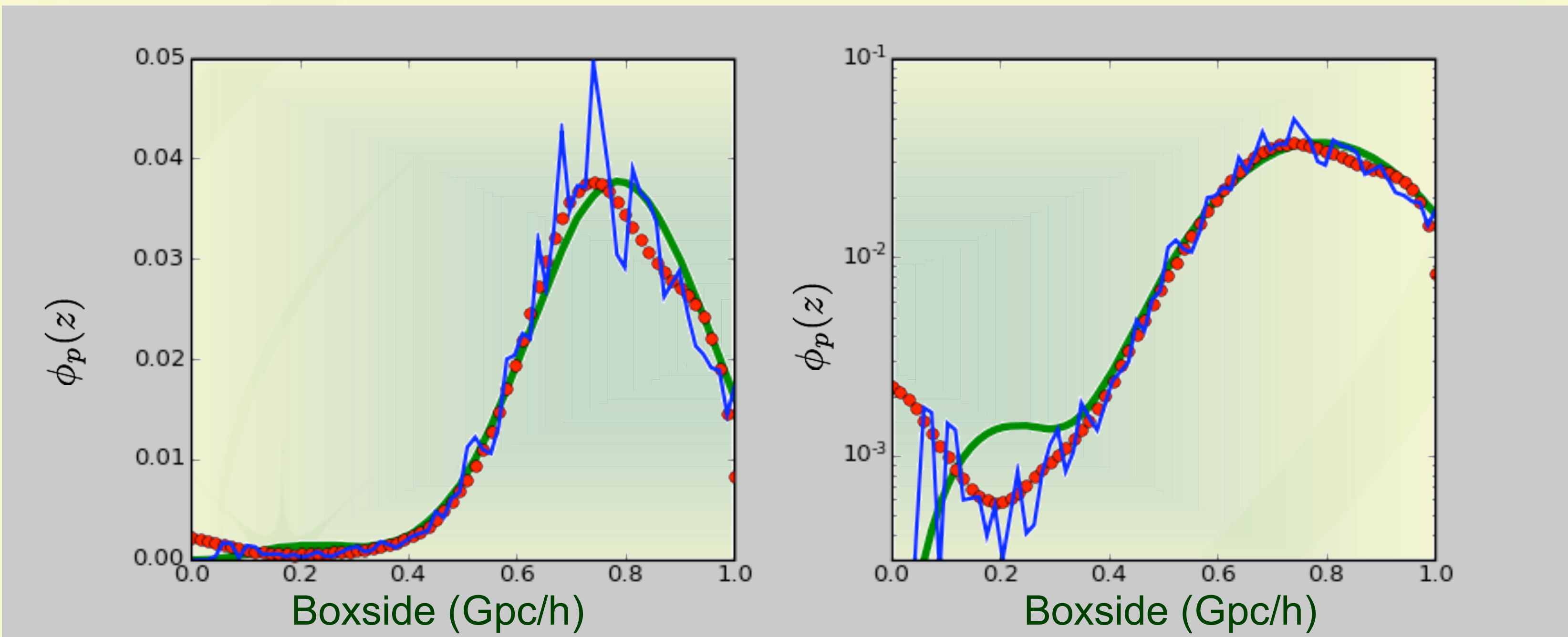
Test 2: Other tracers as calibrators

Same setup



Test 2: Other tracers as calibrators

Reconstruction using more biased tracers (M_{min}=7e12 $M_\odot$/h)



Test 3: Light cones with evolving biases

- DNE (yet!)
- It's crucial to investigate light cone evolution due to a potential degeneracy between $\phi_p(z)$ and $b_p(z)$

$$\xi_{ps}(r,z) = \frac{b_p(z)}{b_s(z)}\,\xi_{ss}(r,z) \quad \Rightarrow \quad (b_p(z)\phi_p(z))$$

In terms of observables

- $\omega_{pp}(\theta)$ does not break this degeneracy:

$$\xi_{pp}(r,z) = \left(\frac{b_p(z)}{b_s(z)}\right)^2 \xi_{ss}(r,z)$$

$$\omega_{pp} = \int dz_1 \int dz_2\, \phi_p(z_1)\, \phi_p(z_2)\, \xi_{pp}(\theta, z_1, z_2)$$

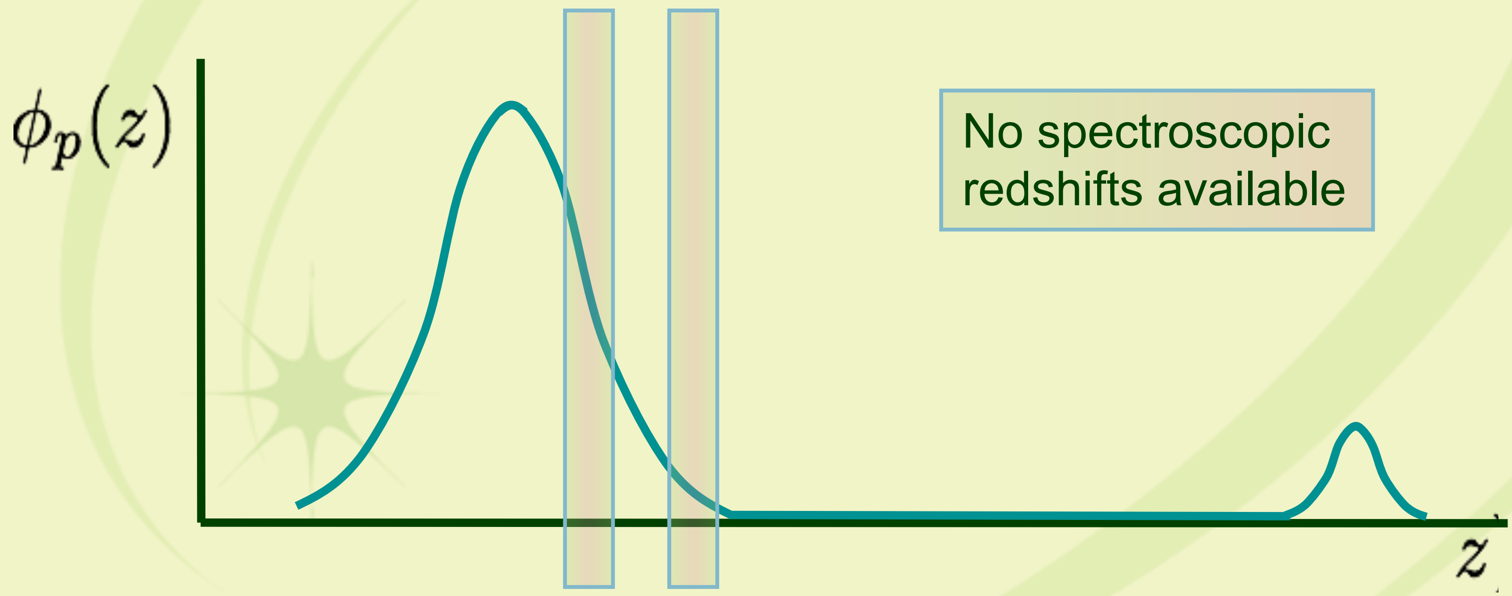
$$\Rightarrow \quad (b_p(z)\phi_p(z))^2$$

In terms of observables

- Could be difficult to put a reasonable prior on $b_p(z)$
- Fair sub-sample will avoid this problem (but is inconvenient)

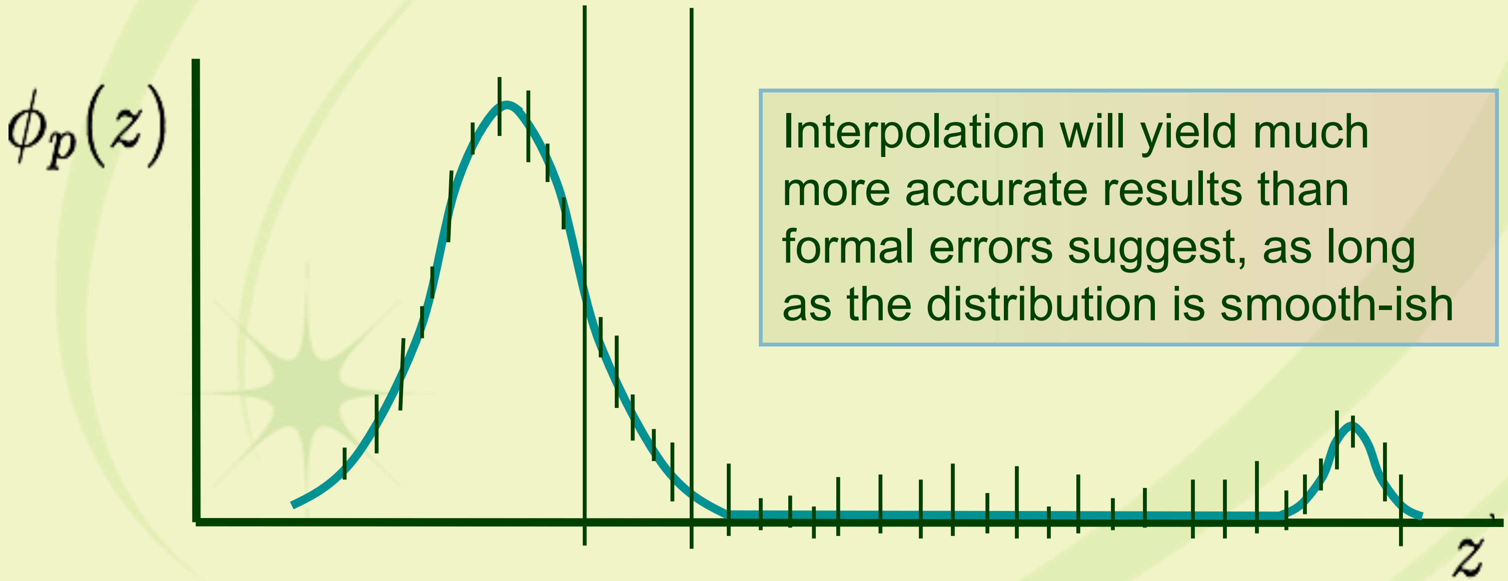
Cross-correlations may have some advantages over direct calibration

- If a fair sub-sample is needed, one might still use the cross-correlation method to help calibrate the photometry
- Contains some independent information: improve calibration
- Insensitive to catastrophic errors in photometry
- Not adversely affected by spectroscopic redshift deserts



Cross-correlations may have some advantages over direct calibration

- If a fair sub-sample is needed, one might still use the cross-correlation method to help calibrate the photometry
- Contains some independent information: improve calibration
- Insensitive to catastrophic errors in photometry
- Not adversely affected by spectroscopic redshift deserts



Interpolation will yield much more accurate results than formal errors suggest, as long as the distribution is smooth-ish

Conclusions

- Weak lensing constraints on cosmological parameters are highly improved by splitting source galaxies into redshift bins
- Tomographic lensing measurements are quite sensitive to uncertainties in source galaxy redshift distribution
- Upcoming surveys will yield an unprecedented volume of photometric data, making spectroscopic follow-up inconvenient
- Cross-correlations are a promising method for calibrating galaxy redshift distributions
- Potential degeneracy between the redshift distribution and evolution of the bias of the photometric sample may compel sub-sample follow-ups
- The cross-correlation method may have some advantages over conventional direct calibration methods, even with spectroscopic follow up observations