Reconstructing Redshift Distributions with Cross-Correlations

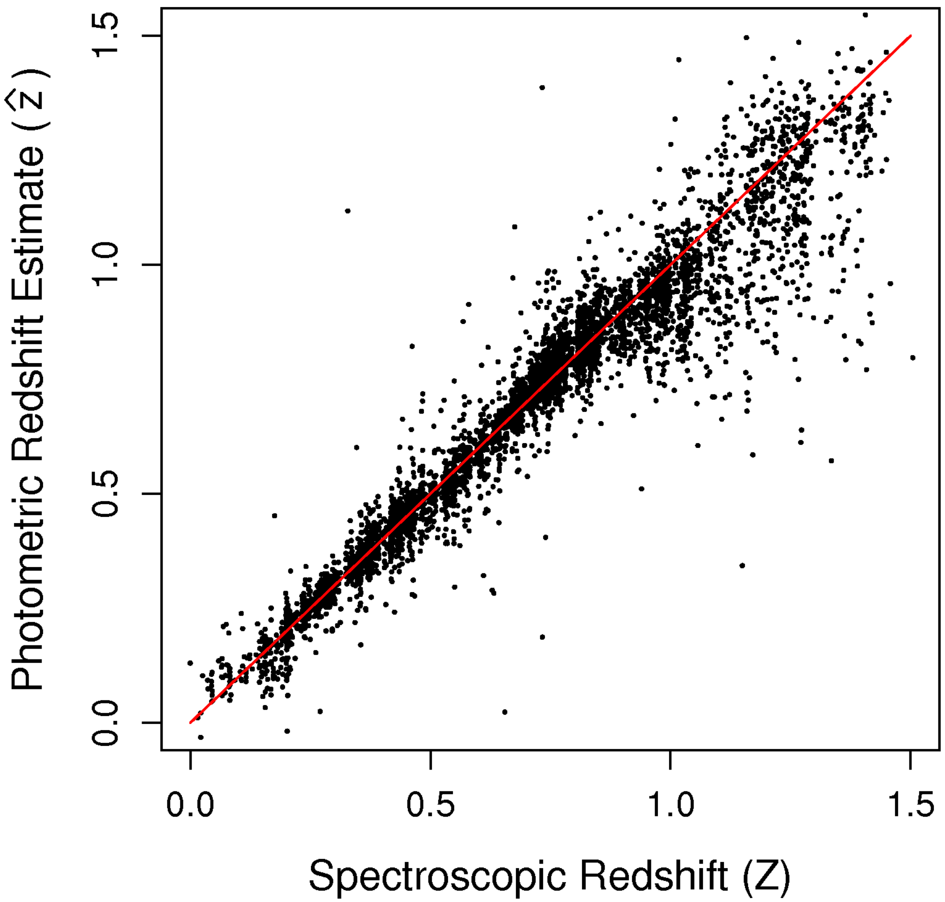
Dan Matthews
University of Pittsburgh
June 16th, 2010

Photometric Redshifts

- Redshifts measured from photometric data
- Advantages:
  - Can measure the redshifts of many more objects than with spectroscopy
  - Photometric surveys go to fainter magnitudes: find more galaxies
- Disadvantages:
  - Less accurate than spectroscopic redshifts
  - Catastrophic outliers
  - Spectroscopic surveys used for calibration are highly incomplete

Photometric Redshifts

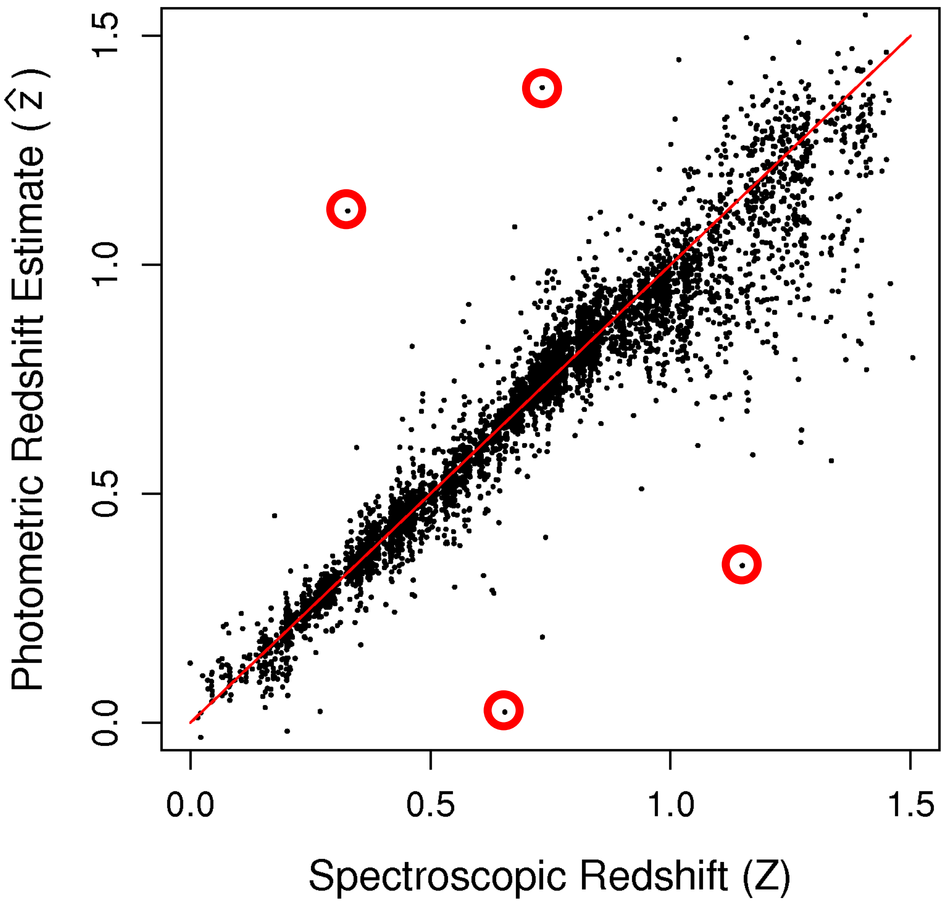
- Redshifts measured from photometric data
- Advantages:
  - Can measure the redshifts of many more objects than with spectroscopy
  - Photometric surveys go to fainter magnitudes: find more galaxies
- Disadvantages:
  - Less accurate than spectroscopic redshifts
  - Catastrophic outliers
  - Spectroscopic surveys used for calibration are highly incomplete



Freeman et al. 2009

Photometric Redshifts

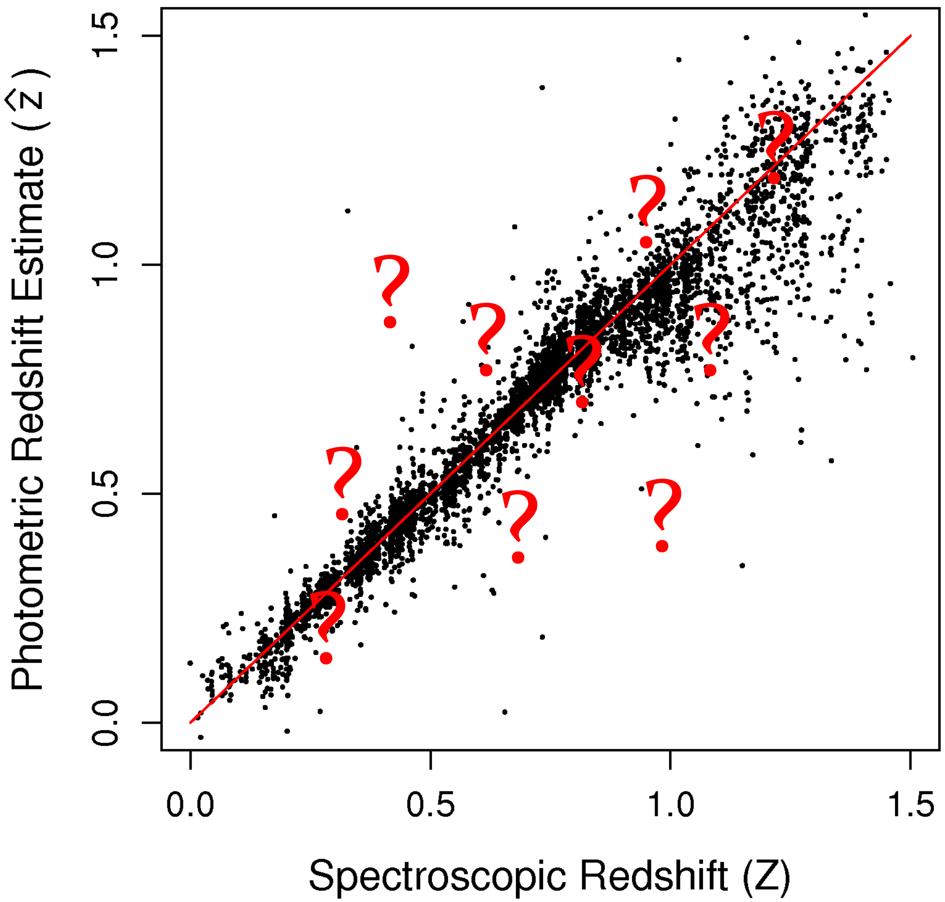
- Redshifts measured from photometric data
- Advantages:
  - Can measure the redshifts of many more objects than with spectroscopy
  - Photometric surveys go to fainter magnitudes: find more galaxies
- Disadvantages:
  - Less accurate than spectroscopic redshifts
  - Catastrophic outliers
  - Spectroscopic surveys used for calibration are highly incomplete



Freeman et al. 2009

Photometric Redshifts

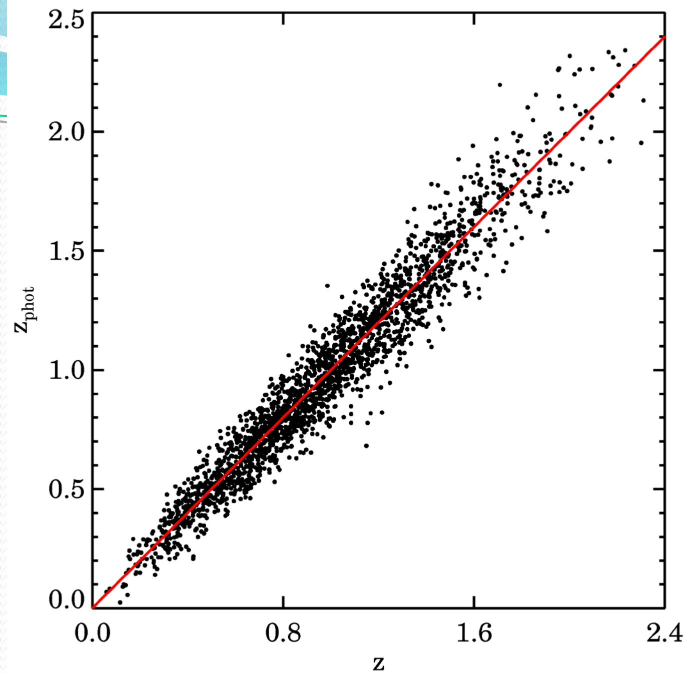
- Redshifts measured from photometric data
- Advantages:
  - Can measure the redshifts of many more objects than with spectroscopy
  - Photometric surveys go to fainter magnitudes: find more galaxies
- Disadvantages:
  - Less accurate than spectroscopic redshifts
  - Catastrophic outliers
  - Spectroscopic surveys used for calibration are highly incomplete



Freeman et al. 2009

Photo-z bins

- Dark energy experiments are unlikely to treat redshifts of individual objects as known
- Objects will often be divided into photo-z bins
- Because of photo-z uncertainties, *true* redshift distribution will differ from photo-z distribution
- Cross-correlation method recovers this true distribution



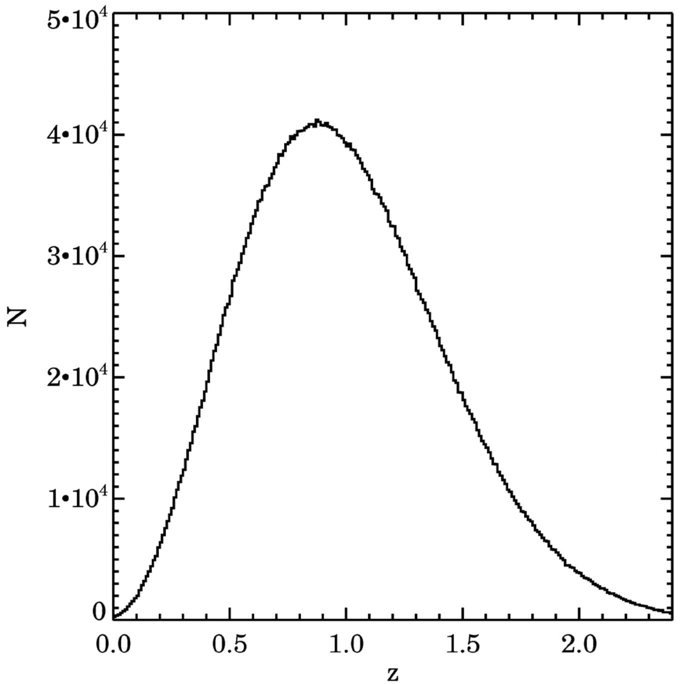
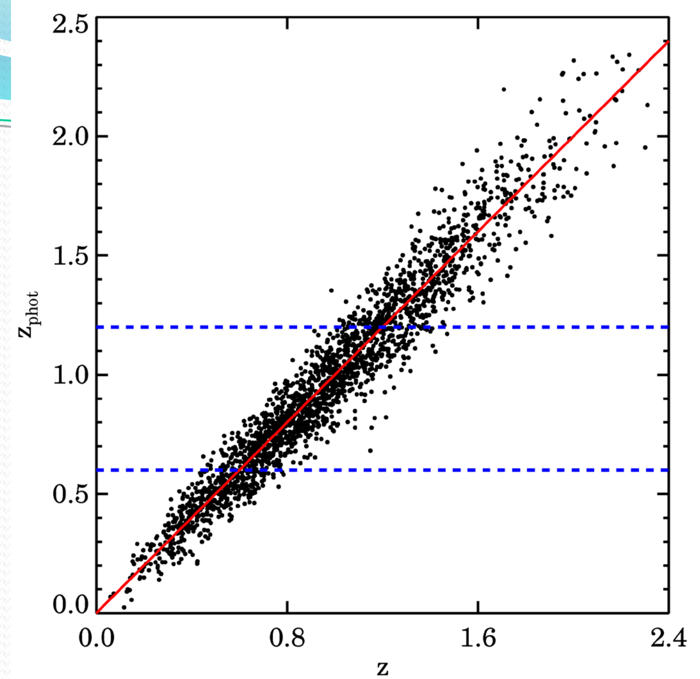


Photo-z bins

- Dark energy experiments are unlikely to treat redshifts of individual objects as known
- Objects will often be divided into photo-z bins
- Because of photo-z uncertainties, *true* redshift distribution will differ from photo-z distribution
- Cross-correlation method recovers this true distribution



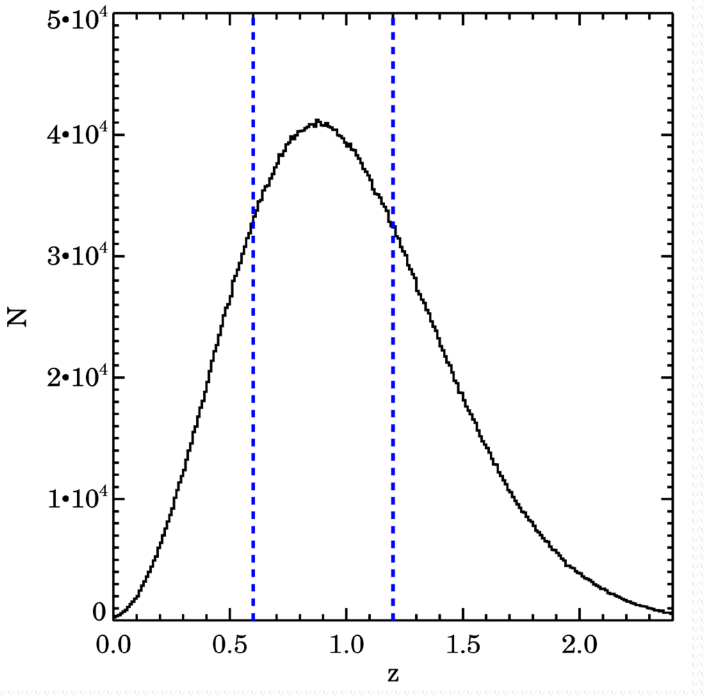
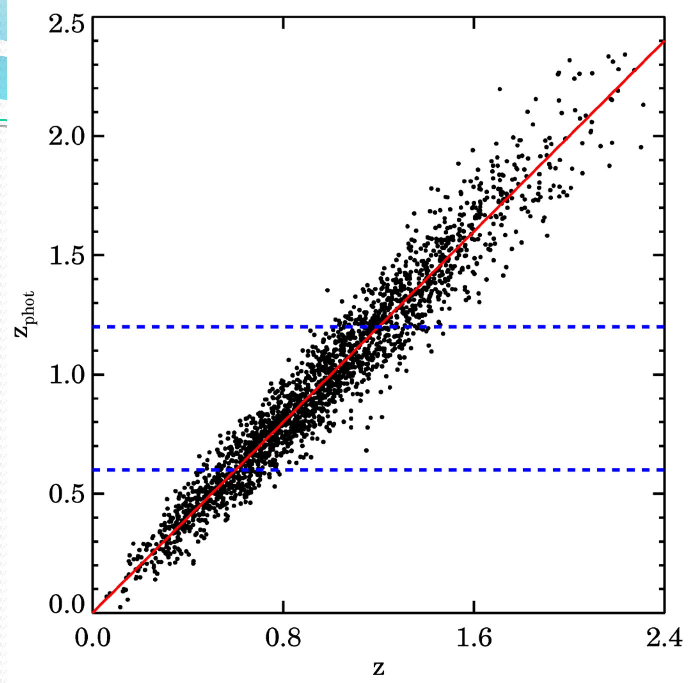
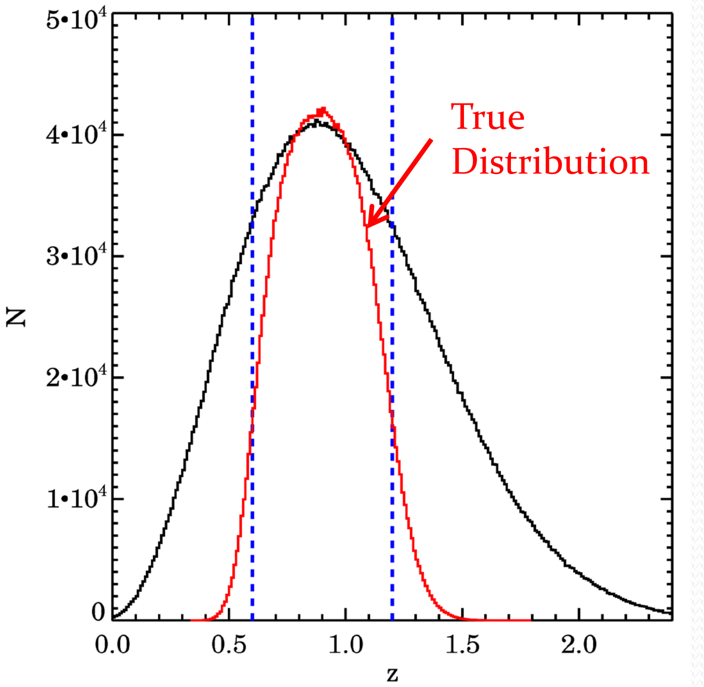


Photo-z bins

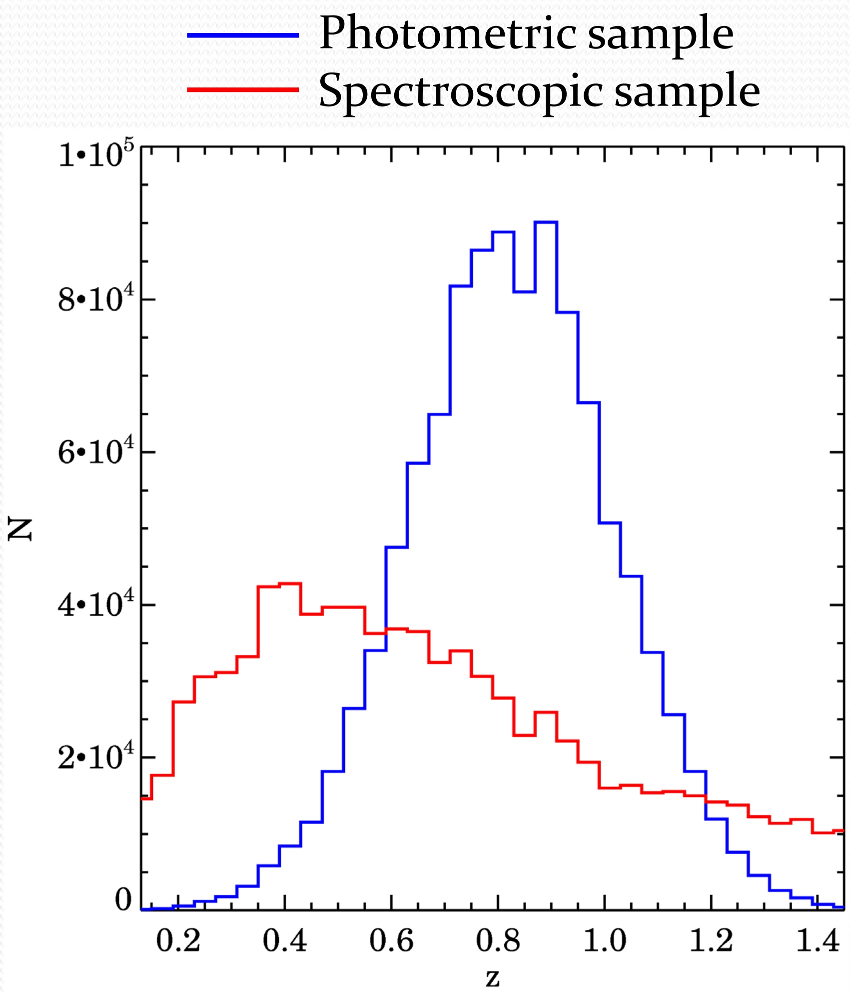
- Dark energy experiments are unlikely to treat redshifts of individual objects as known
- Objects will often be divided into photo-z bins
- Because of photo-z uncertainties, *true* redshift distribution will differ from photo-z distribution
- Cross-correlation method recovers this true distribution





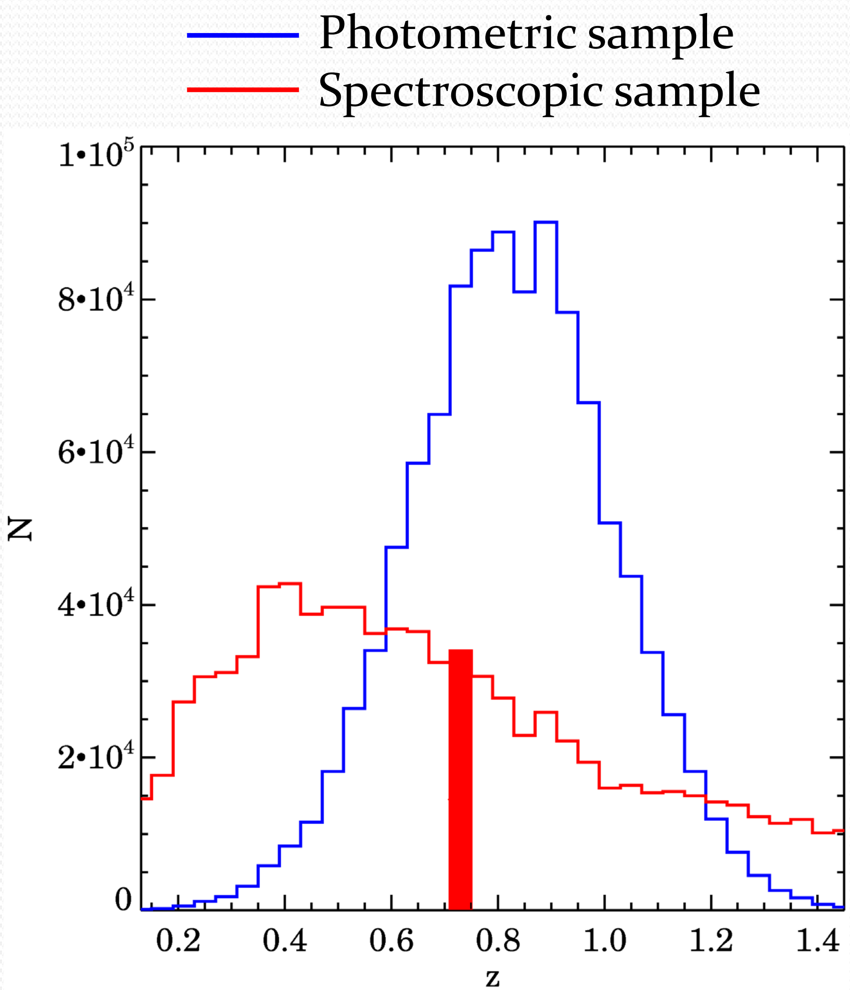
Cross-Correlation Method

- Two samples in the same region of sky: one with only photometric information, the other with known spectroscopic redshifts
- Measure the angular cross-correlation between objects in the two samples as a function of spectroscopic z



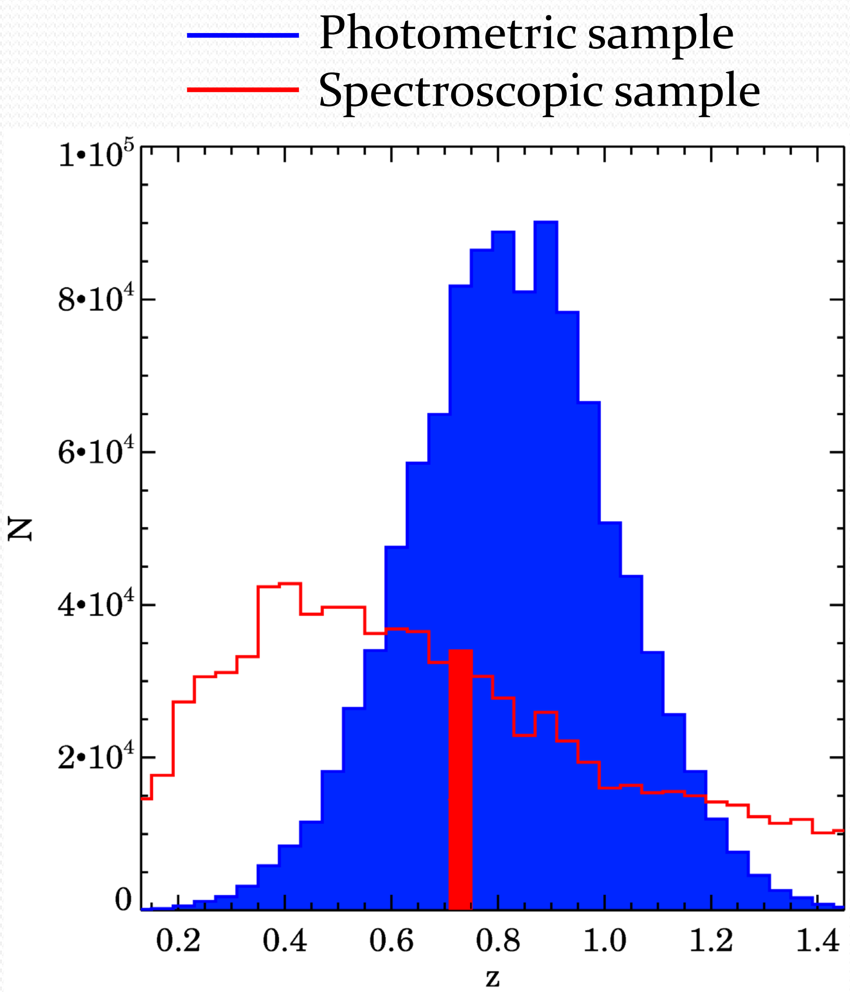
Cross-Correlation Method

- Two samples in the same region of sky: one with only photometric information, the other with known spectroscopic redshifts
- Measure the angular cross-correlation between objects in the two samples as a function of spectroscopic z



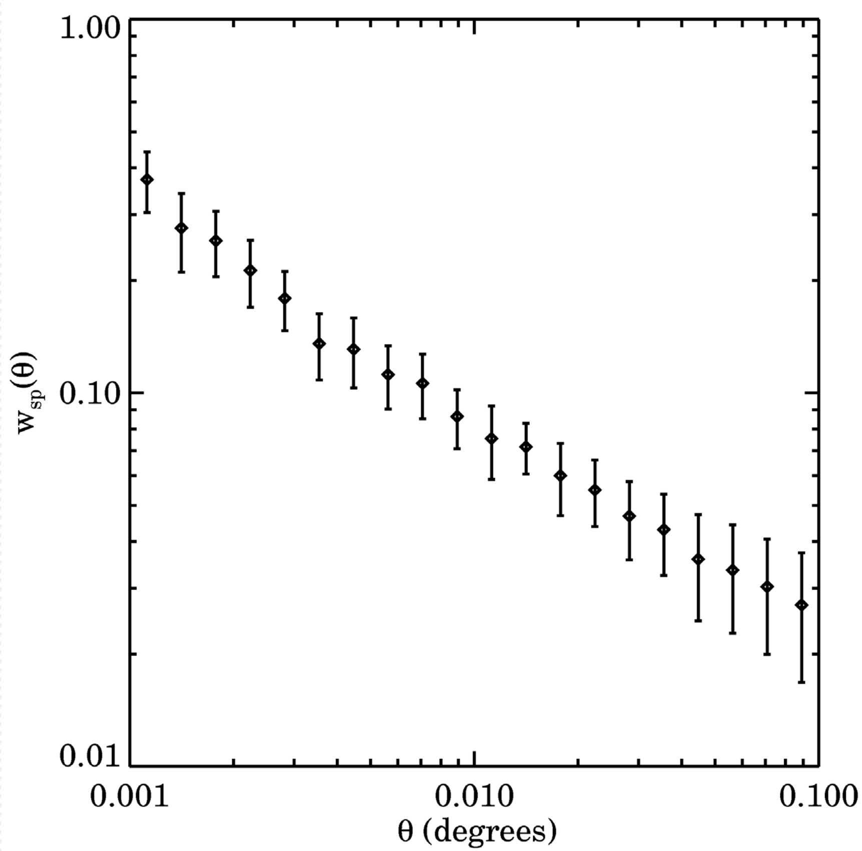
Cross-Correlation Method

- Two samples in the same region of sky: one with only photometric information, the other with known spectroscopic redshifts
- Measure the angular cross-correlation between objects in the two samples as a function of spectroscopic z



Cross-Correlation Method

- Two samples in the same region of sky: one with only photometric information, the other with known spectroscopic redshifts
- Measure the angular cross-correlation between objects in the two samples as a function of spectroscopic z



Cross-Correlation Method

- Basic idea:
  - Objects at similar redshifts tend to cluster with each other
  - The more overlap in z between the two samples, the larger the correlation amplitude is
- Model the real-space correlation function as a power law

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma}$$

$$w_{sp}(\theta, z) \propto \underbrace{\phi_p(z)}_{\text{Redshift Distribution}} \underbrace{r_{0,sp}^{\gamma_{sp}}}_{\text{Intrinsic Clustering}} \theta^{1-\gamma_{sp}}$$

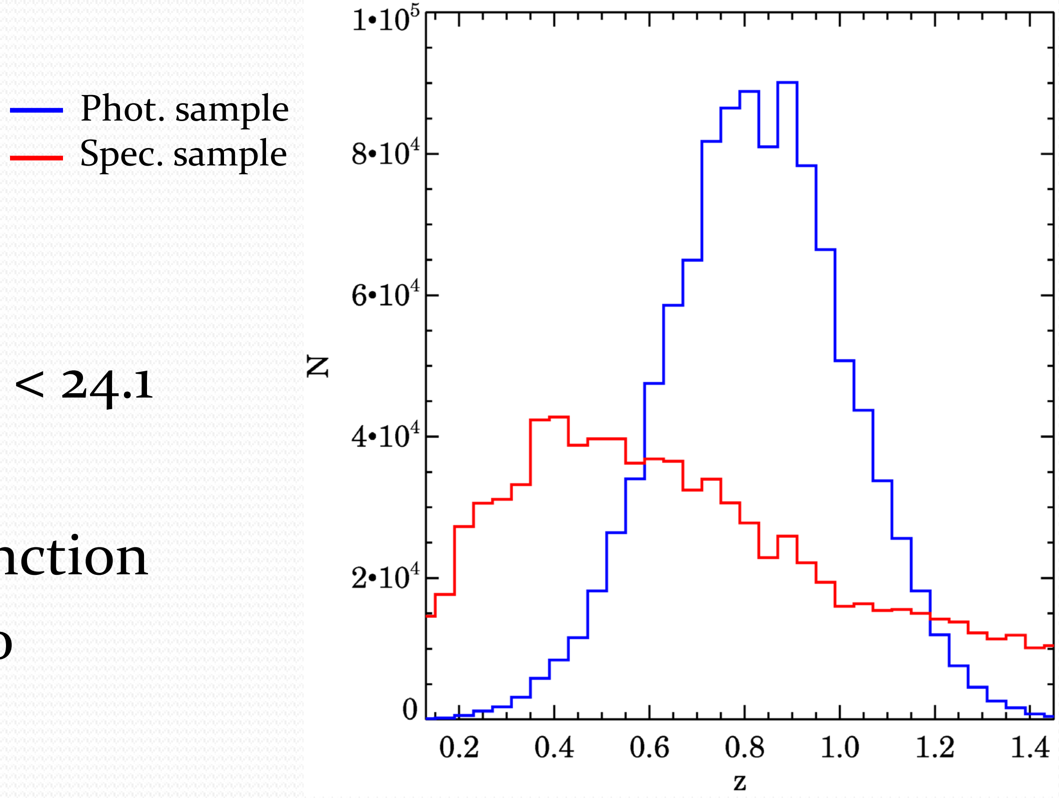
- Autocorrelation measurements for each sample give information about their intrinsic clustering

Recovering $\phi_p(z)$

- Cross-correlation measurements in multiple spec-z bins, along with autocorrelations, allow us to recover the true redshift distribution
- Advantages:
  - The spectroscopic sample does not have to be complete at any given redshift
  - Possible to use only the brightest objects from which it is easier to obtain secure spec. redshift measurements
  - Effective even when the two samples do not have similar properties (e.g. differing bias)

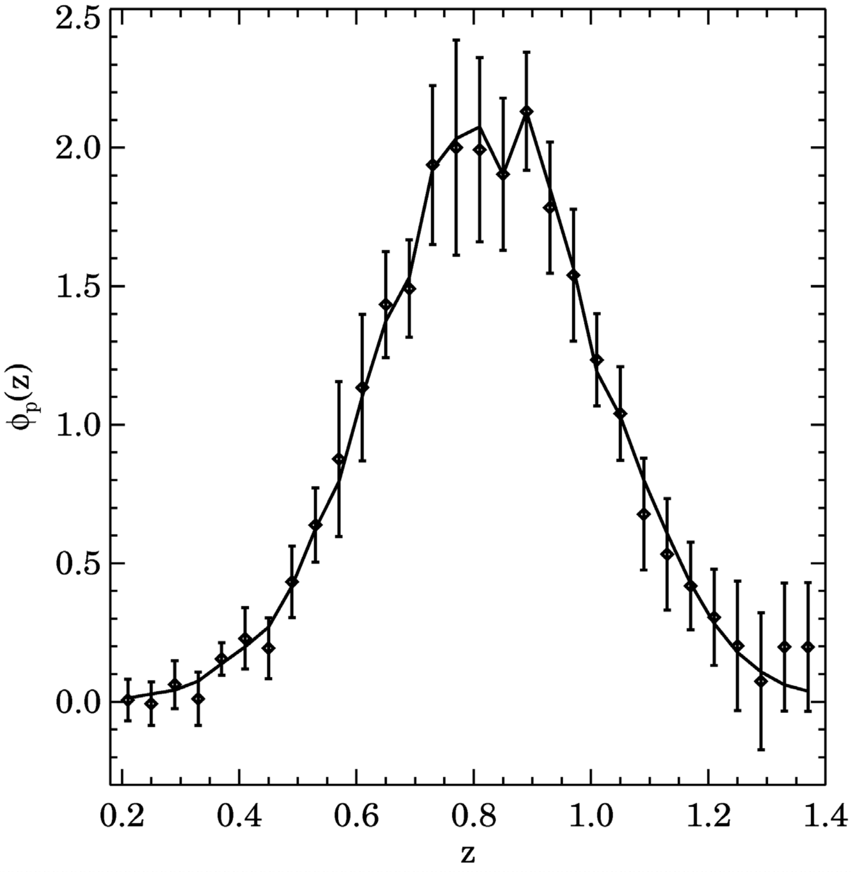
Testing on mock catalogs

- Darren Croton's mock DEEP2 Redshift Survey light-cones
  - 24 light-cones constructed using lines-of-sight through the Millennium Simulation plus a semi-analytic model
    - 0.5 x 2.0 deg. region of sky and redshift range, $0.1 < z < 1.5$
  - Spectroscopic sample
    - 60% of objects with $R < 24.1$
  - Photometric sample
    - Gaussian selection function
      - $\langle z \rangle = 0.75 \quad \sigma_z = 0.20$



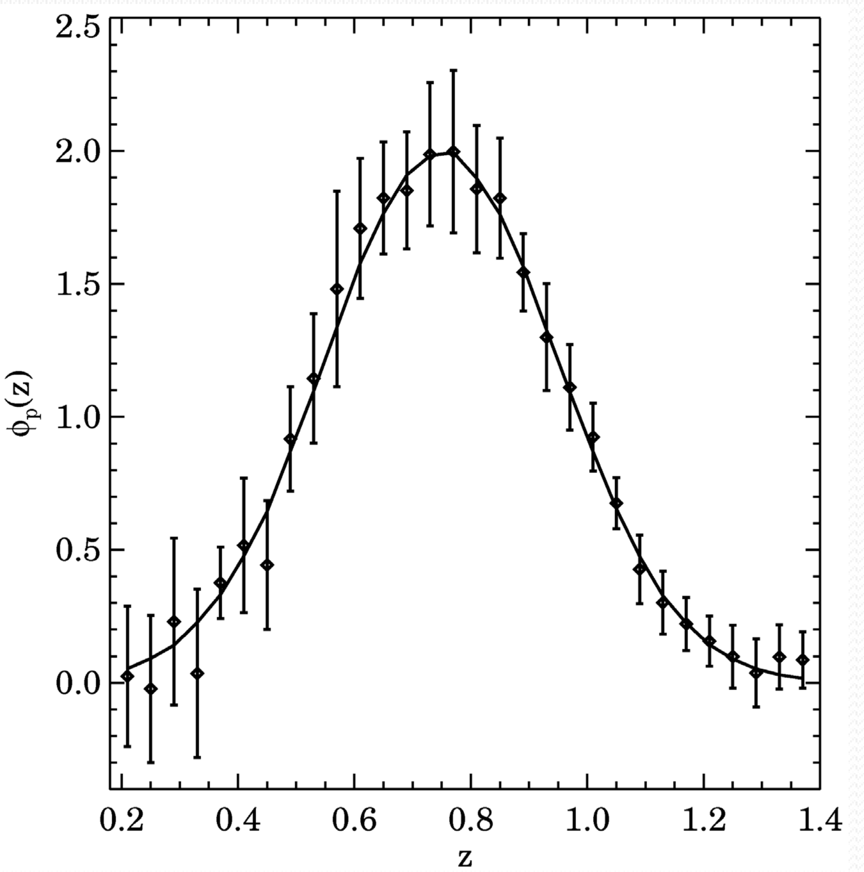
Recovery of $\phi_p(z)$

- Points show the median of 10^4 measurements, error bars are their standard deviation
  - Each measurement: averaging 4 fields picked at random from the 24 total fields
- Errors are larger than expected using weak clustering formalism
  - Extra variance terms in the correlation measurements (Bernstein 1994)



Error Analysis

- Divide out parent sample redshift distribution and apply empirical correction for cosmic variance to recover Gaussian selection function
  - $\sigma[<z>] = 7.4 \times 10^{-3}$
  - $\sigma[\sigma_z] = 8.5 \times 10^{-3}$
- Estimated requirements for LSST at z=0.75
  - $\sigma[<z>] \sim 3.5 \times 10^{-3}$
  - $\sigma[\sigma_z] \sim 5.3 \times 10^{-3}$



Conclusions

- We have applied the cross-correlation method to realistic mock catalogs and get a good overall recovery
  - Includes the effects of bias evolution and cosmic variance
- Errors in recovered $\langle z \rangle$ and σ_z are larger than projected requirements
  - This is solvable with:
    - Larger samples
    - Wider sky coverage at constant sample size
    - Improved correlation estimators (Padmanabhan 2007)

Correlation Function

- 3-D two-point real-space correlation function

$$dP = n[1 + \xi(r)]dV$$

- Angular two-point correlation function

$$dP = \Sigma[1 + w(\theta)]d\Omega$$

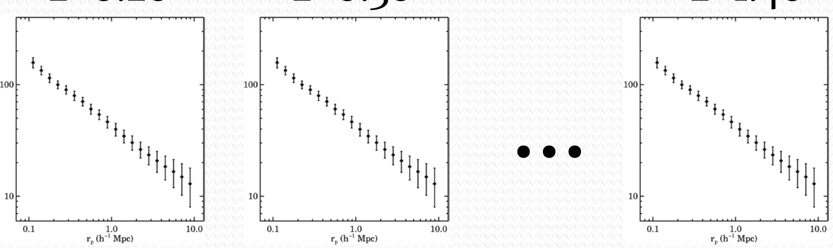
- Cross-correlation: excess probability of finding a galaxy from one sample at a given separation from a galaxy in another sample

Autocorrelation

Spectroscopic Sample

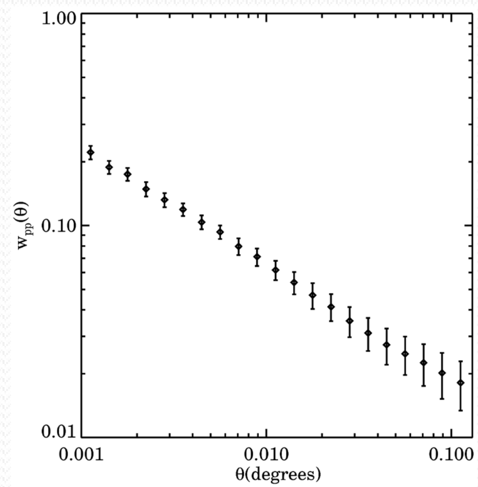
- Real-space correlation function
  - Projected correlation function
    - Integral along the line-of-sight reduces impact of redshift space distortions
  - Measured in multiple spec-z bins

z=0.20 z=0.30 z=1.40



Photometric Sample

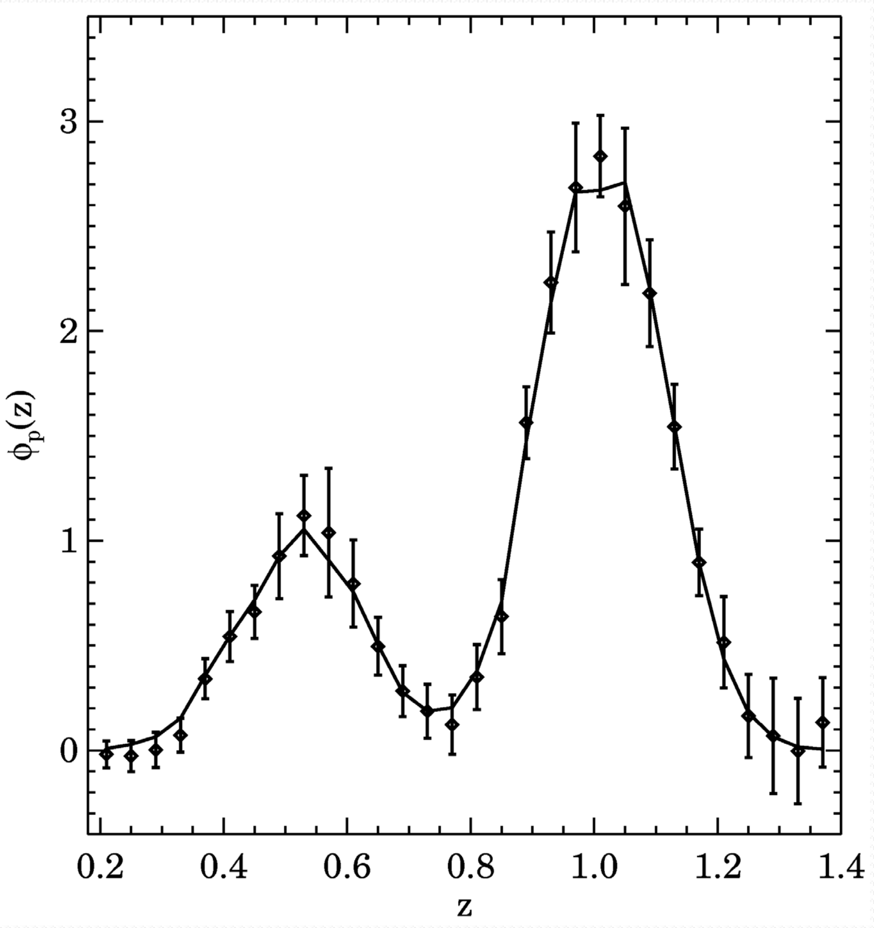
- Angular correlation function
  - Measured for the full sample



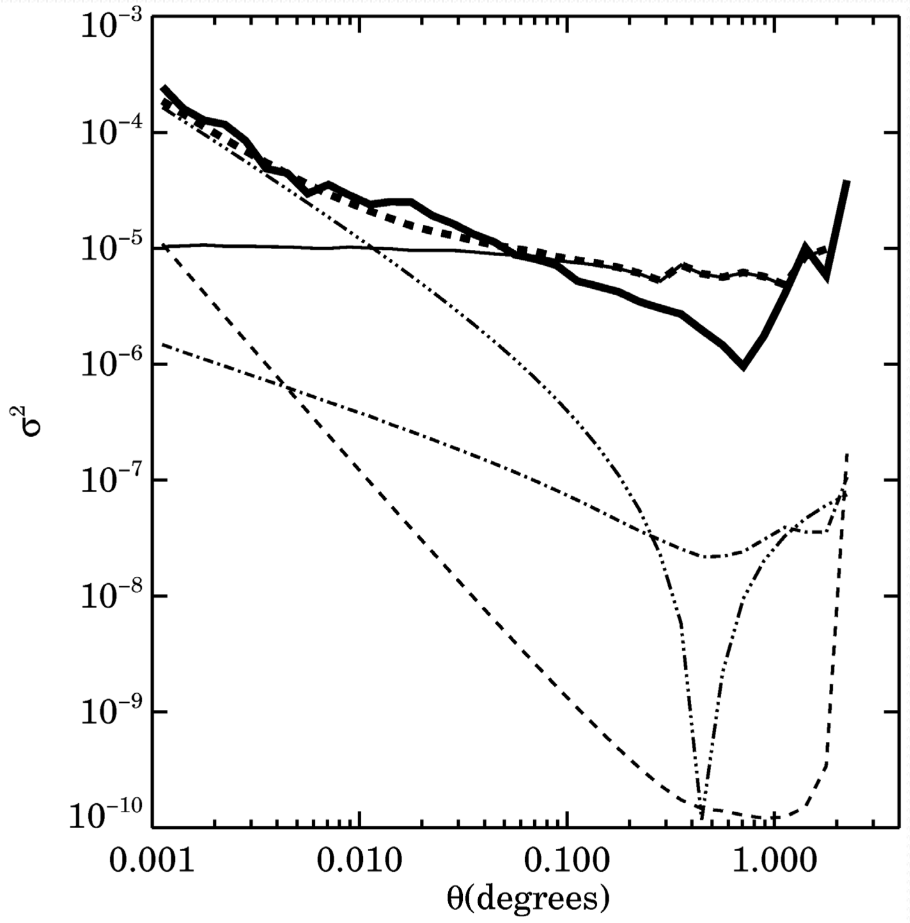
Intrinsic clustering of the the samples with each other

$$\xi_{sp}(r) = (\xi_{ss}\xi_{pp})^{1/2}$$

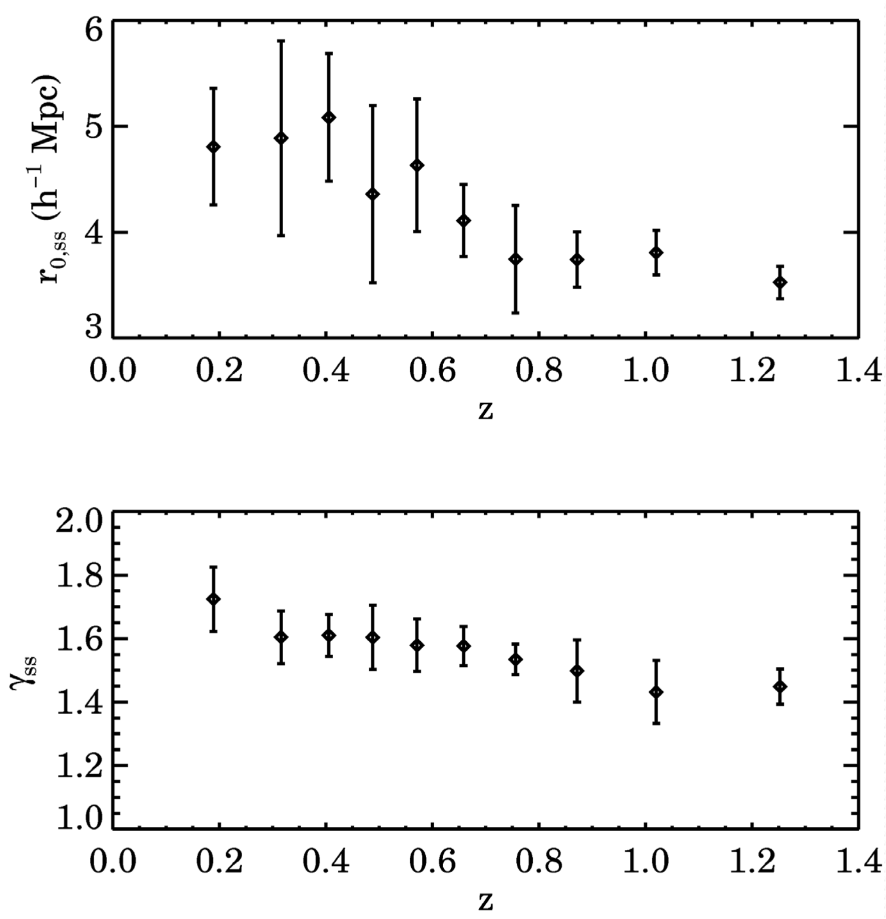
Double peak distribution



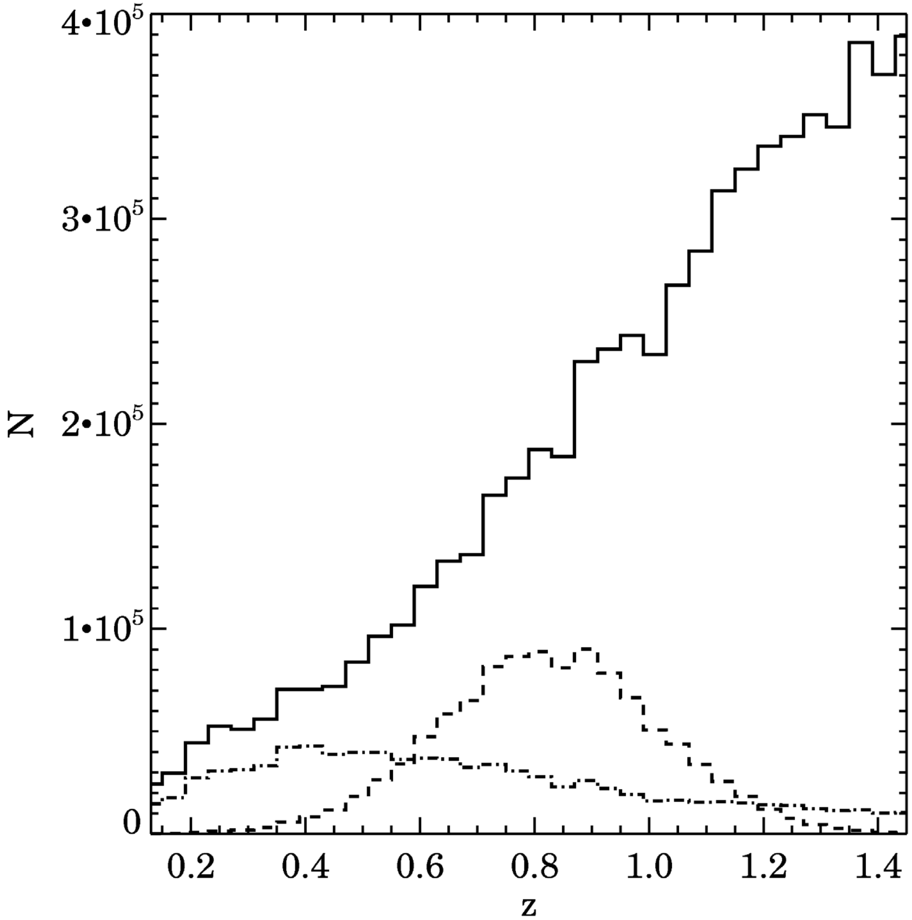
Bernstein variance terms



Spectroscopic sample parameters



Sample distributions



Equations

$$w_{sp}(\theta, z) = \frac{\phi_p(z) H(\gamma_{sp}) r_{0,sp}^{\gamma_{sp}} \theta^{1-\gamma_{sp}} D(z)^{1-\gamma_{sp}}}{dl/dz}$$

$$w_{pp}(\theta) = H(\gamma_{pp}) \theta^{1-\gamma_{pp}} \int_0^\infty \phi_p^2(z) r_{0,pp}^{\gamma_{pp}} \frac{D(z)^{1-\gamma_{pp}}}{dl/dz} dz$$

$$w_p(r_p) = 2 \int_0^\infty \xi[(r_p^2 + \pi^2)^{1/2}] d\pi$$

$$\xi_{sp}(r) = (\xi_{ss}\xi_{pp})^{1/2}$$