Inferring redshift distributions from clustering measurements

Brice Ménard
Johns Hopkins University
& Kavli IPMU

in collaboration with

Mubdi Rahman
Johns Hopkins University

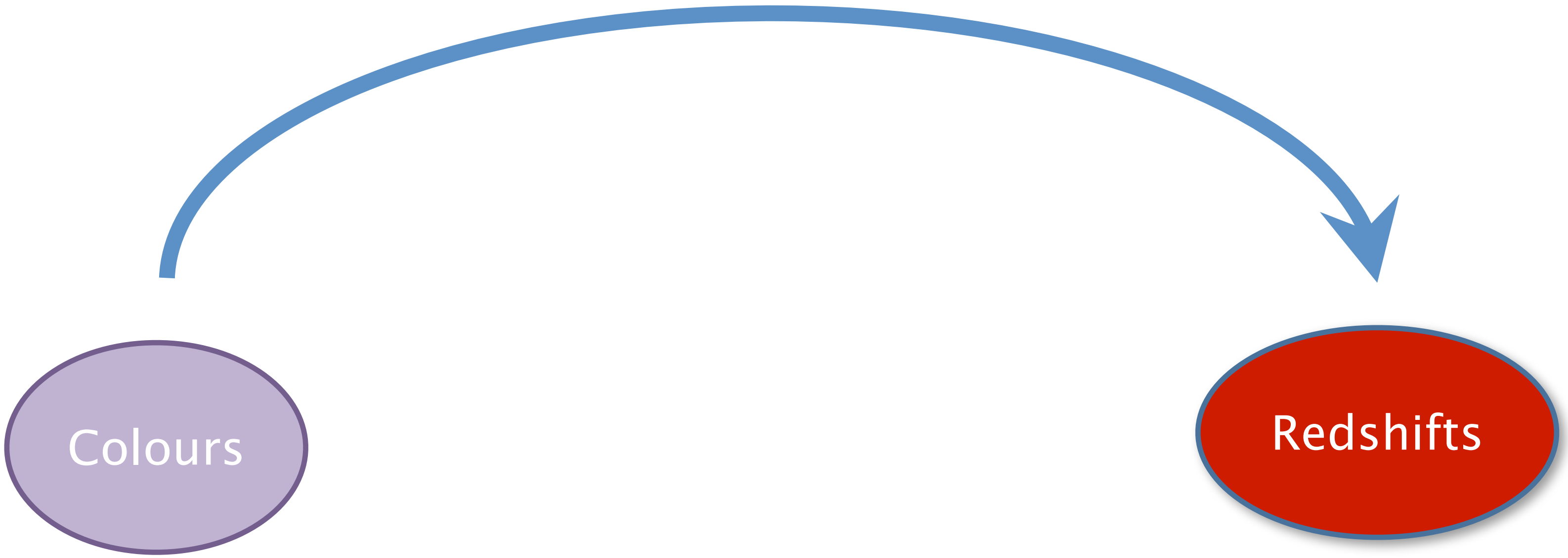
Ryan Scranton, Sam Schmidt, Chris Morrison (UC Davis)
Tamas Budavari (JHU)

spectroscopic redshifts

- They are well defined as long as spectral features are present and unambiguously identified
- The source needs to be bright enough
- They are expensive

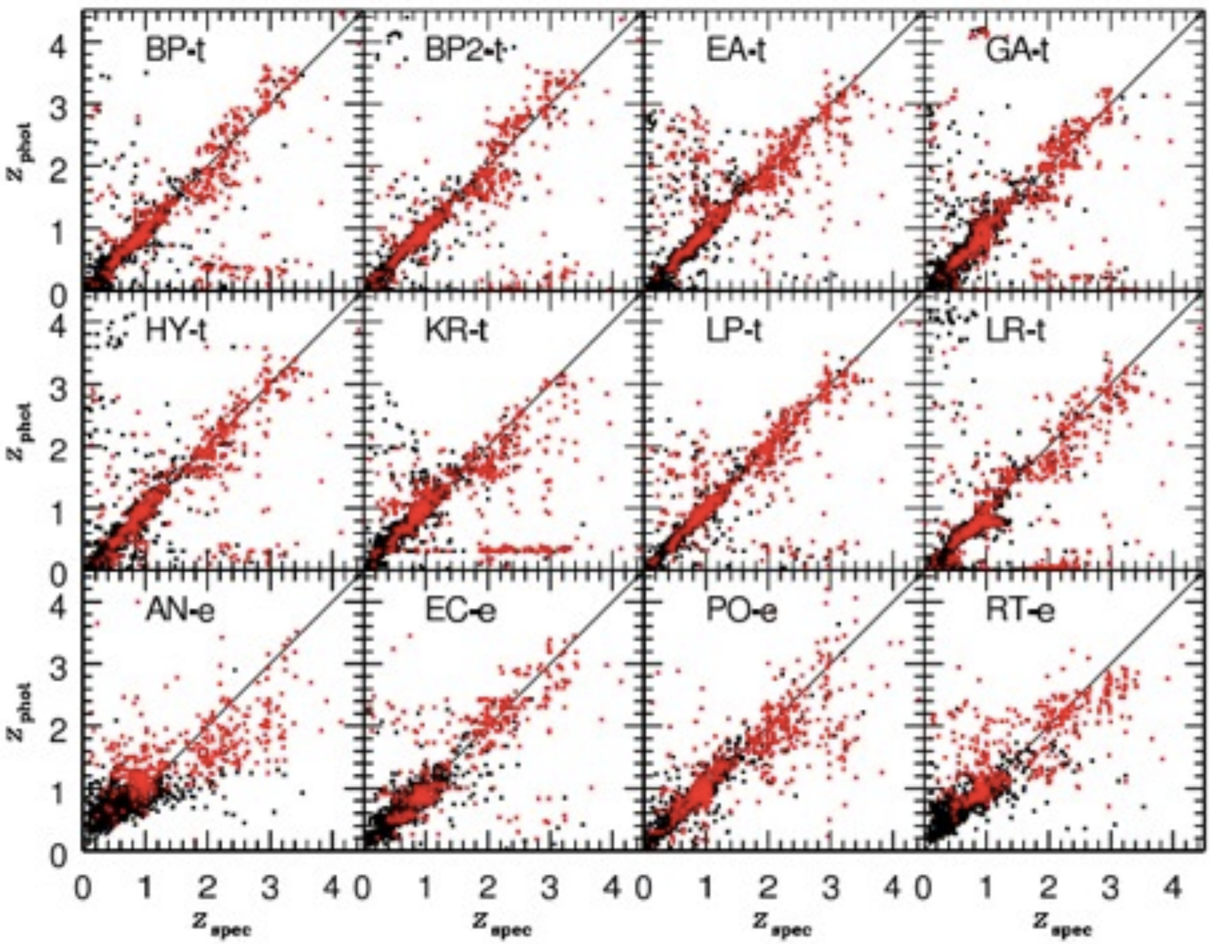
As time goes on the fraction of *known* galaxies for which we have a spectroscopic redshift decreases.

Photometric Redshifts
SEDs or Training Sets



Photometric redshifts

- They rely on templates (theoretical or observed)
- They require training sets. The answer is not unique.

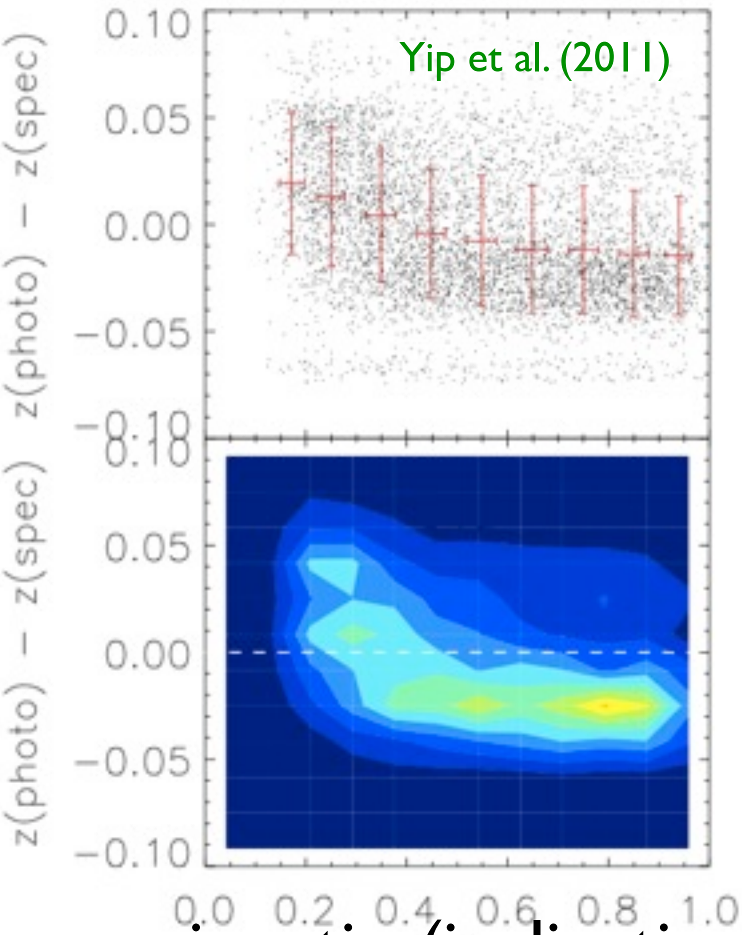


PHAT: PHoto-z Accuracy Testing ★

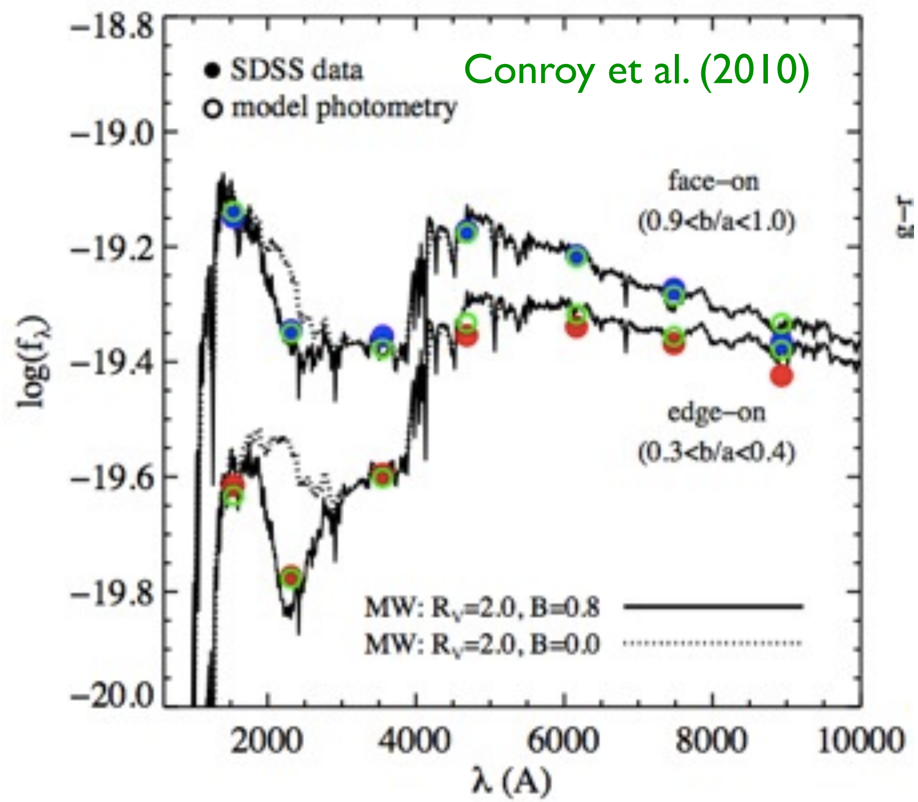
Hildebrandt et al.

Photometric redshifts

- They rely on templates (theoretical or observed)
- They require training sets. The answer is not unique.
- They are affected by dust extinction/reddening effects



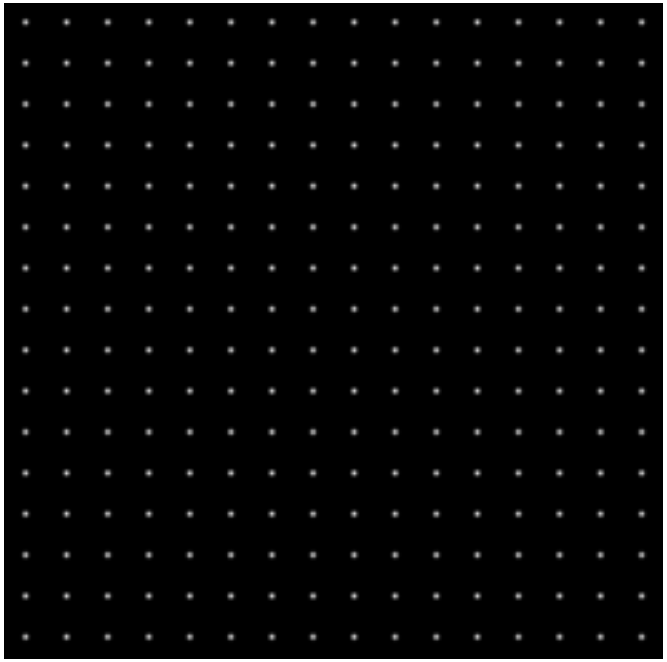
axis ratio (inclination)

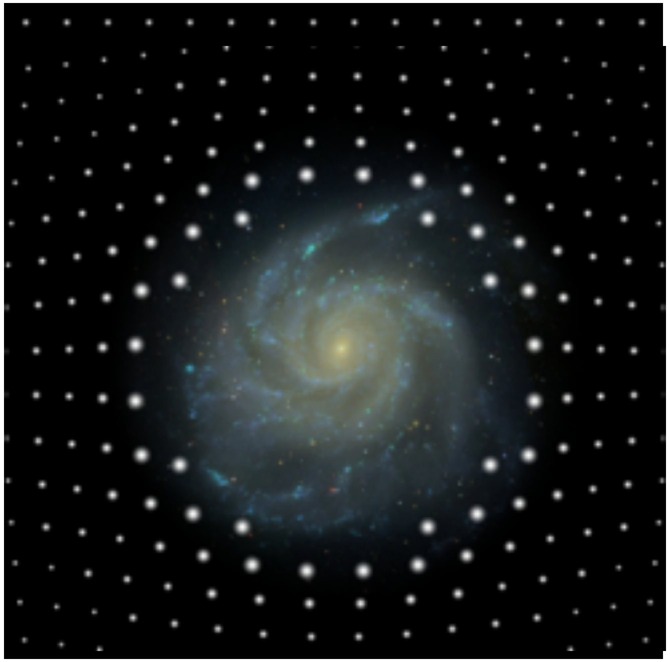


Photometric redshifts

- They rely on templates (theoretical or observed)
- They require training sets. The answer is not unique.
- They are affected by dust extinction/reddening effects
- They suffer from catastrophic failures

=> serious problem to select clean samples of foreground & background objects, needed for gravitational lensing.





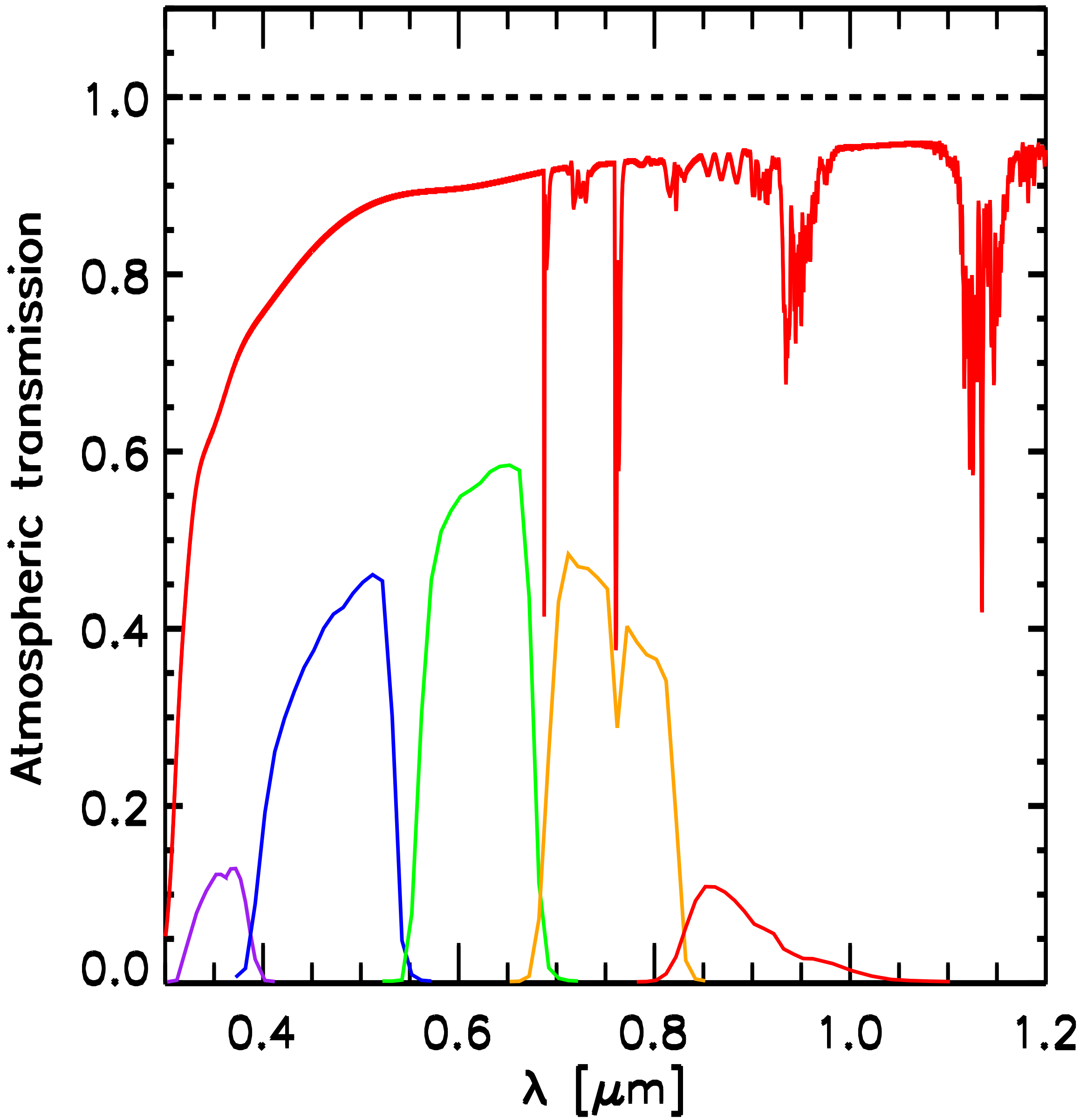
Photometric redshifts

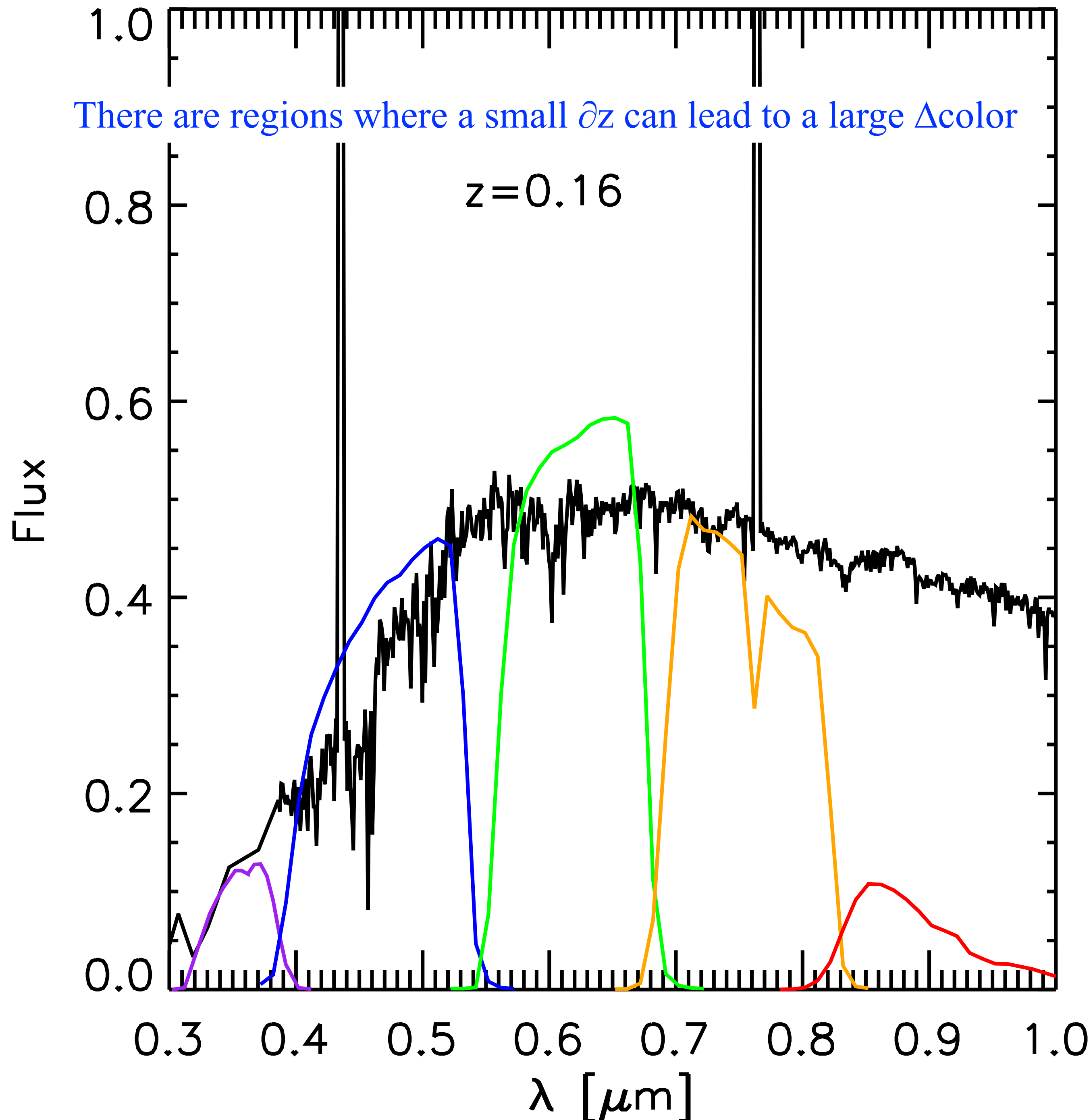
- They rely on templates (theoretical or observed)
- They require training sets. The answer is not unique.
- They are affected by dust extinction/reddening effects
- They suffer from catastrophic failures

- Most importantly they rely on our a priori knowledge of the sources. When exploring the unknown, they may no longer be reliable.

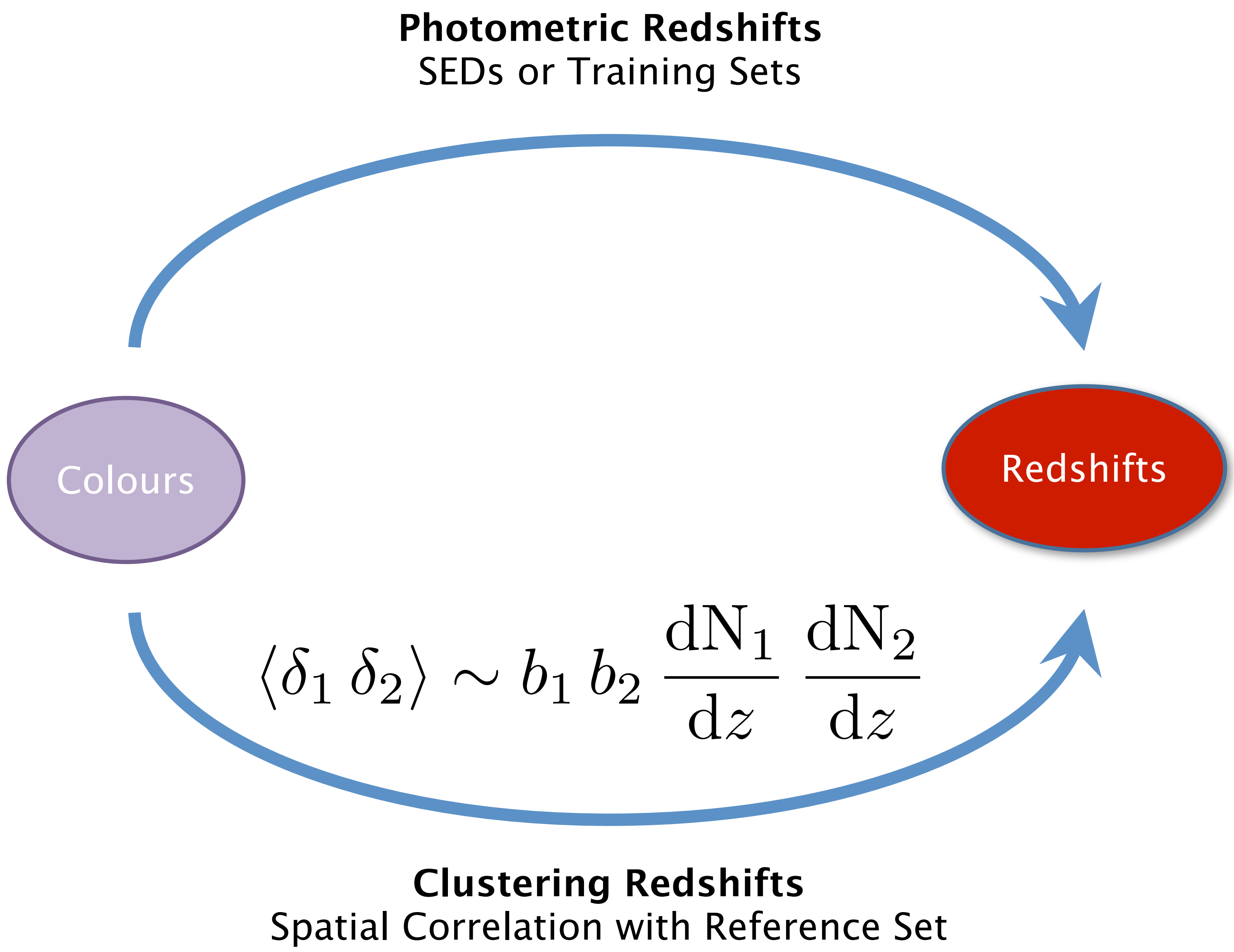
We could completely miss an entire population of galaxies.

Atmospheric transmission





There are regions where a small ∂z can lead to a large Δcolor
z=0.16
Flux
λ [μm]
1.0
0.8
0.6
0.4
0.2
0.0
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0


Photometric Redshifts
SEDs or Training Sets
Colours
Redshifts
$\langle \delta_1 \, \delta_2 \rangle \sim b_1 \, b_2 \, \frac{\mathrm{dN}_1}{\mathrm{d}z} \, \frac{\mathrm{dN}_2}{\mathrm{d}z}$
Clustering Redshifts
Spatial Correlation with Reference Set

Estimating redshifts

$$\langle \delta_1 \, \delta_2 \rangle \sim b_1 \, b_2 \, \frac{\mathrm{d}N_1}{\mathrm{d}z} \, \frac{\mathrm{d}N_2}{\mathrm{d}z}$$

Seldner & Peebles (1979), Roberts & Odell (1979)
Landy, Szalay, Koo (1996) —— [local sampling]
Newman (2008), Matthews & Newman (2010, 2012)
Schultz (2010), McQuinn & White (2013), de Putter et al. (2013)

Schmidt et al. (2012), Ménard et al. (2013), —— [local sampling]
Rahman et al. (2013)

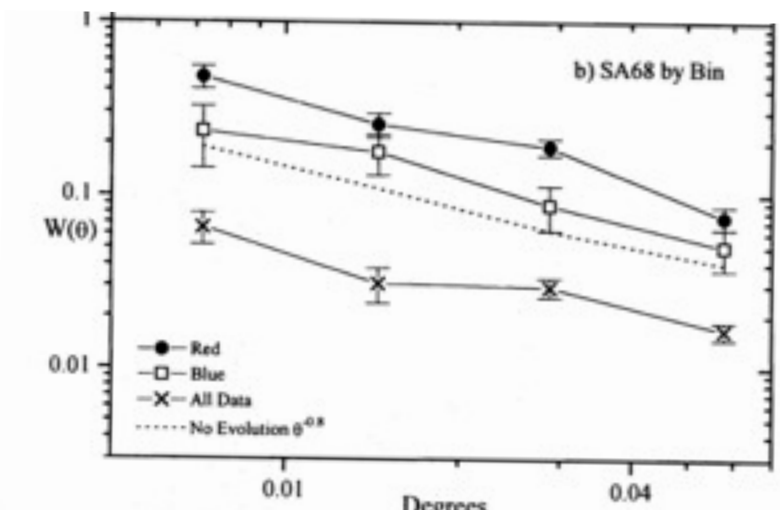
STRONG ANGULAR CLUSTERING OF VERY BLUE GALAXIES: EVIDENCE OF A LOW-REDSHIFT POPULATION

STEPHEN D. LANDY,[1] ALEXANDER S. SZALAY,[2,3] AND DAVID C. KOO[4]

Received 1995 February 17; accepted 1995 September 22

ABSTRACT

We have studied galaxy two-point angular correlations as a function of color using 4 m plate photometry in two independent fields. Each field consists of over 2900 galaxies with magnitudes $20 < B_J < 23.5$ in an area of ~ 750 arcmin^2. We find that the autocorrelation amplitude of the bluest 15% of galaxies is surprisingly strong, with a relative increase in clustering amplitude of a factor of 6 over that of the complete data set, while exhibiting a power-law slope consistent with the canonical value of -0.8. These very blue galaxies are also found to be weakly correlated with galaxies of median color and marginally anticorrelated with the reddest subset. These correlation properties are incompatible with existing simple models of the galaxy distribution; they suggest that a significant fraction, more than 50%, of these very blue galaxies are a faint population that lies at nearby redshifts, $z < 0.3$.

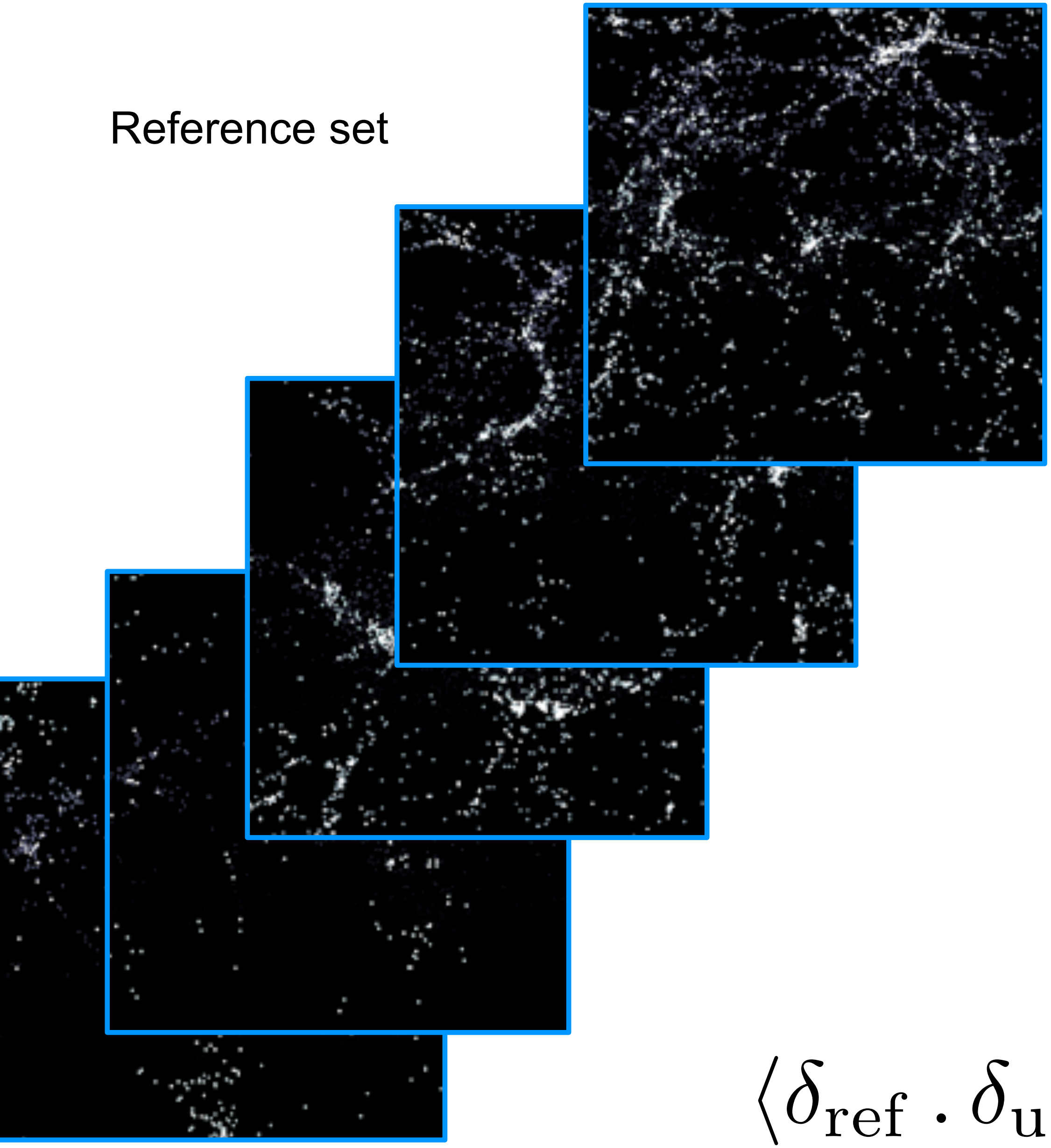


AUTO- AND CROSS-CORRELATIONS BY COLOR BAND

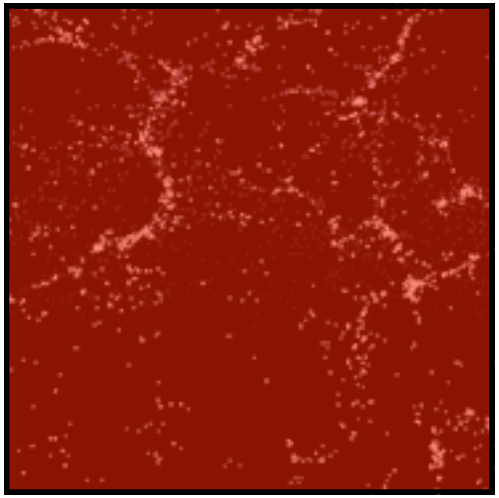
	Blue (20%)	Mid (60%)	Red (20%)
Blue......	68 ± 20	9 ± 7	−18 ± 10
Mid	...	14 ± 1	20 ± 9
Red	...	...	130 ± 30

NOTES.—Cross- and autocorrelations between the 20% most extreme blue and red galaxies and the mid 60%, averaged over the two fields. The mean A_w for all galaxies in the two fields SA 57 and SA 68 was 18.0 ± 1.8. Given the enhanced amplitude of the autocorrelations for the red and blue subsets, together with the weak negative cross-correlation between them, we conclude that these subsets are disparate populations.

Reference set

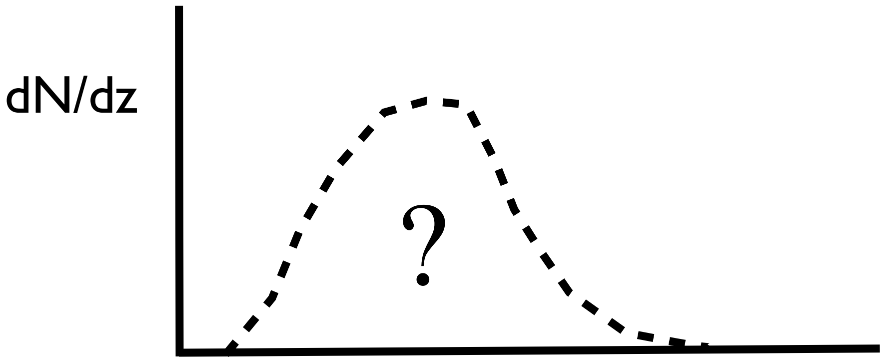


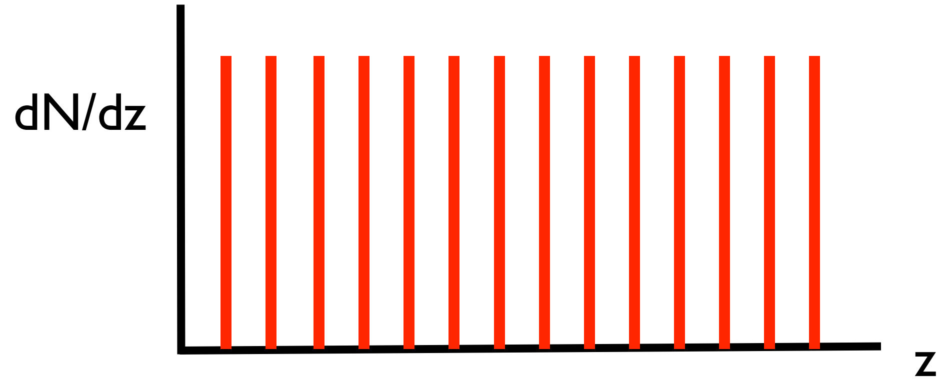
Sample at unknown redshift



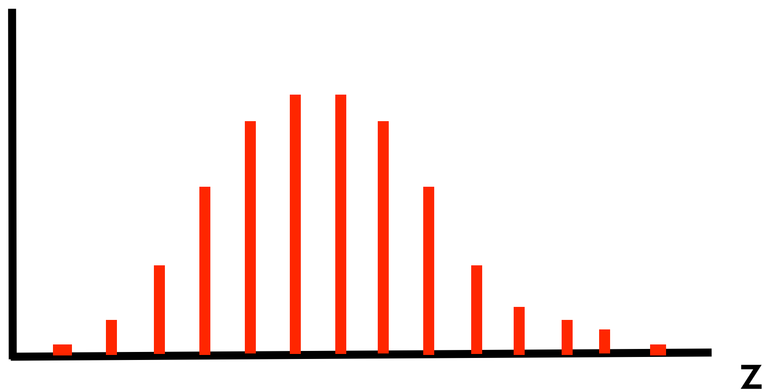
$$\langle \delta_{\mathrm{ref}} \cdot \delta_{\mathrm{unknown}} \rangle$$

Metric: 2-point correlation function



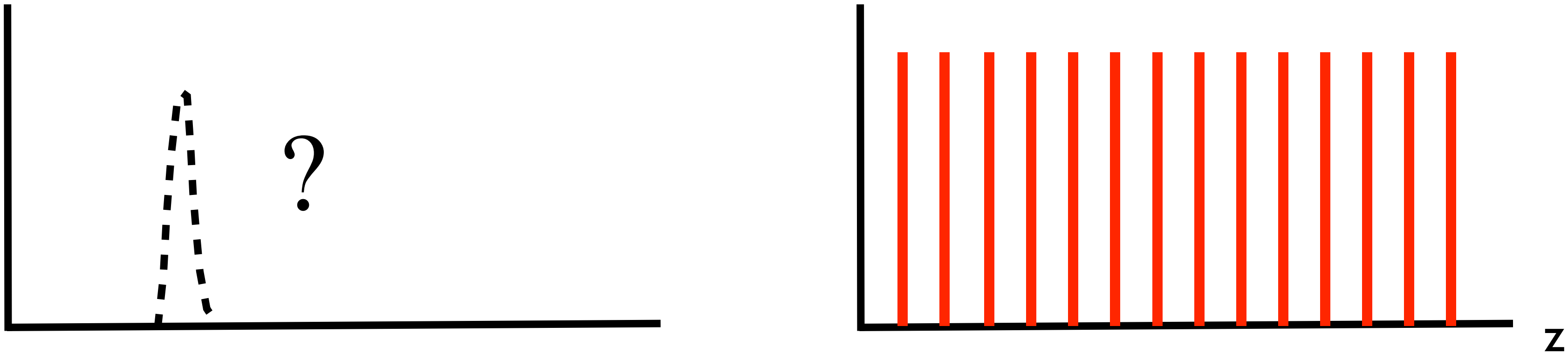


$< \partial_{\mathrm{unknown}} \; . \; \partial_{\mathrm{reference}} >$

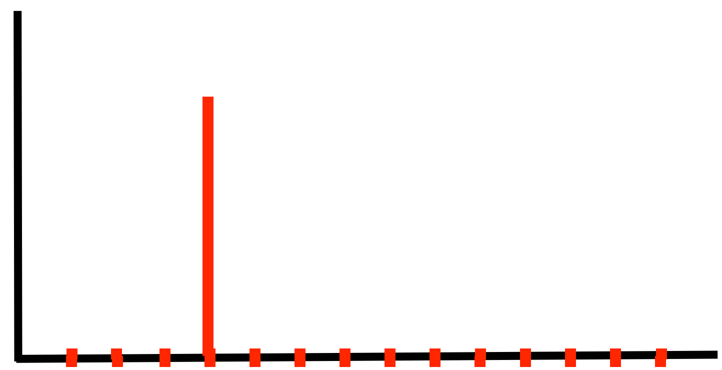


$$\bar{w}_{ur}(z) \sim \int \mathrm{d}z' \, \frac{\mathrm{dN_u}}{\mathrm{d}z'} \frac{\mathrm{dN_r}}{\mathrm{d}z'} \, \bar{b}_u(z')\bar{b}_r(z') \, \bar{w}(z')$$

$< \partial_{\mathrm{unknown}} \, . \, \partial_{\mathrm{reference}} >$



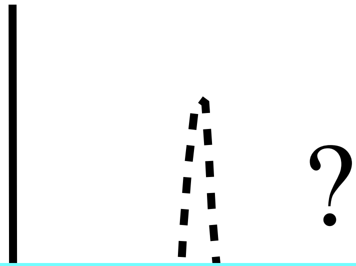
if $$\frac{\mathrm{d}\log \mathrm{dN}/\mathrm{d}z}{\mathrm{d}z} \gg \frac{\mathrm{d}\log \mathrm{b_u}(z)}{\mathrm{d}z}$$

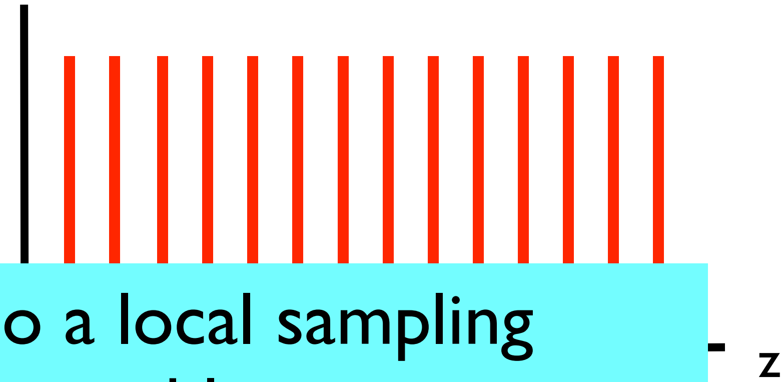


then $$\mathrm{dN}/\mathrm{d}z \propto \bar{w}_{ur}(z) \left(\frac{1}{\bar{b}_r(z)\, \bar{w}(z)} \right)$$

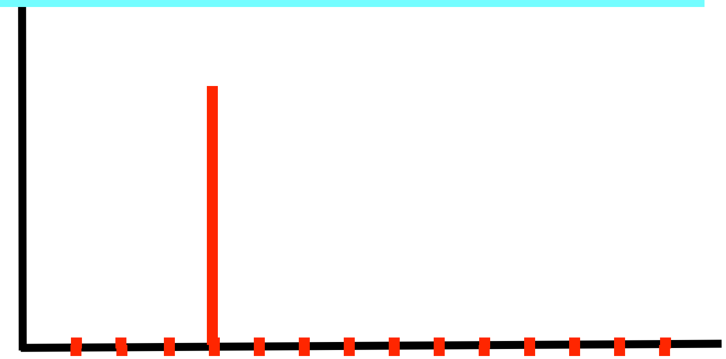
and the redshift distribution is simply normalized by $$\int \mathrm{d}z \, \mathrm{dN}/\mathrm{d}z = N_{\mathrm{tot}}$$

$< \partial_{\rm unknown} \, . \, \partial_{\rm reference} >$





The key point here is to do a local sampling in the space of observables



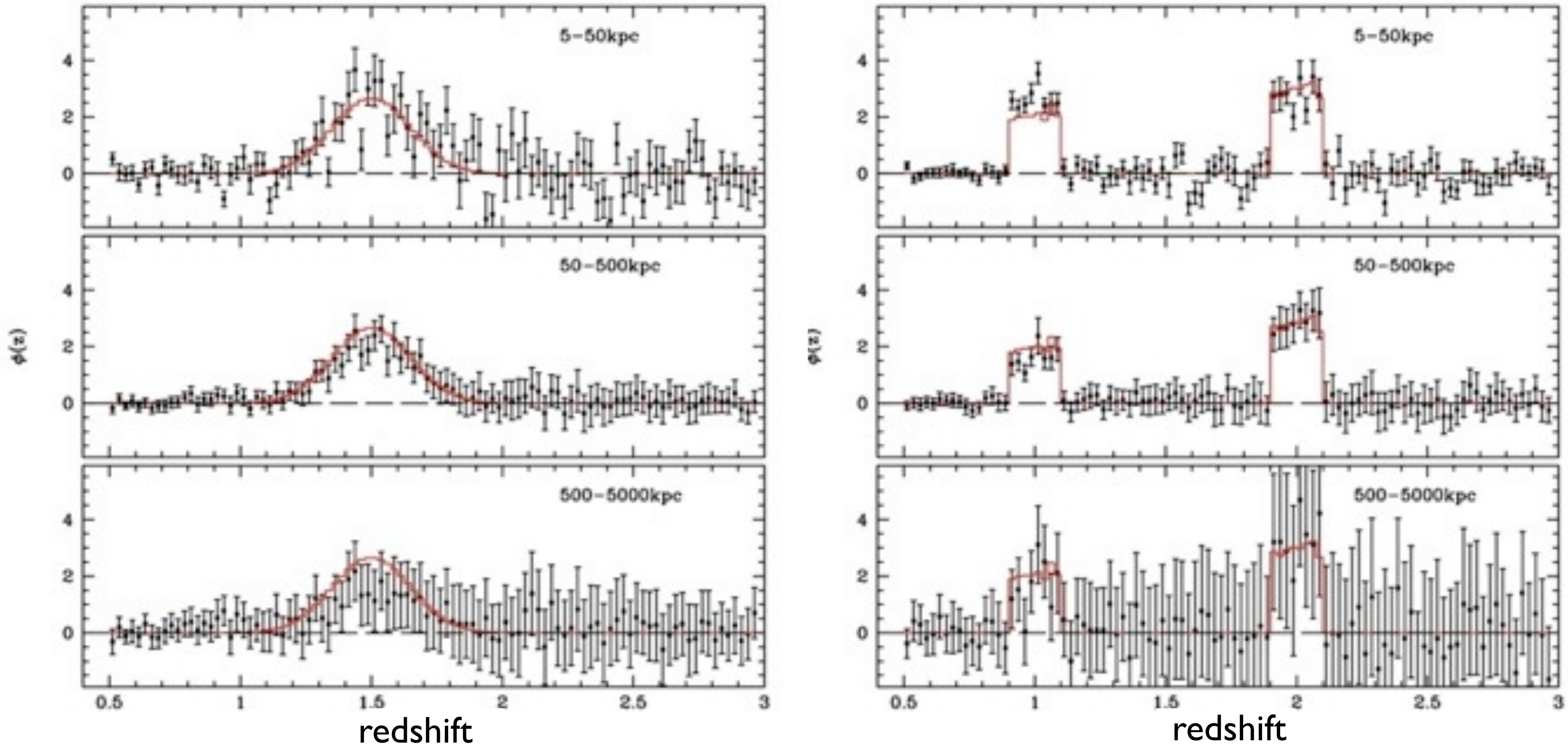
if $\frac{\rm d \log dN/d\mathit{z}}{{\rm d}z} \gg \frac{\rm d \log b_u(\mathit{z})}{{\rm d}z}$

then $$\mathrm{dN/d}z \propto \bar{w}_{ur}(z) \left(\frac{1}{\bar{\bar{b}}_r(z)\,\bar{w}(z)} \right)$$

and the redshift distribution is simply normalized by $$\int \mathrm{d}z\, \mathrm{dN/d}z = N_{\rm tot}$$

Scale dependence of the results

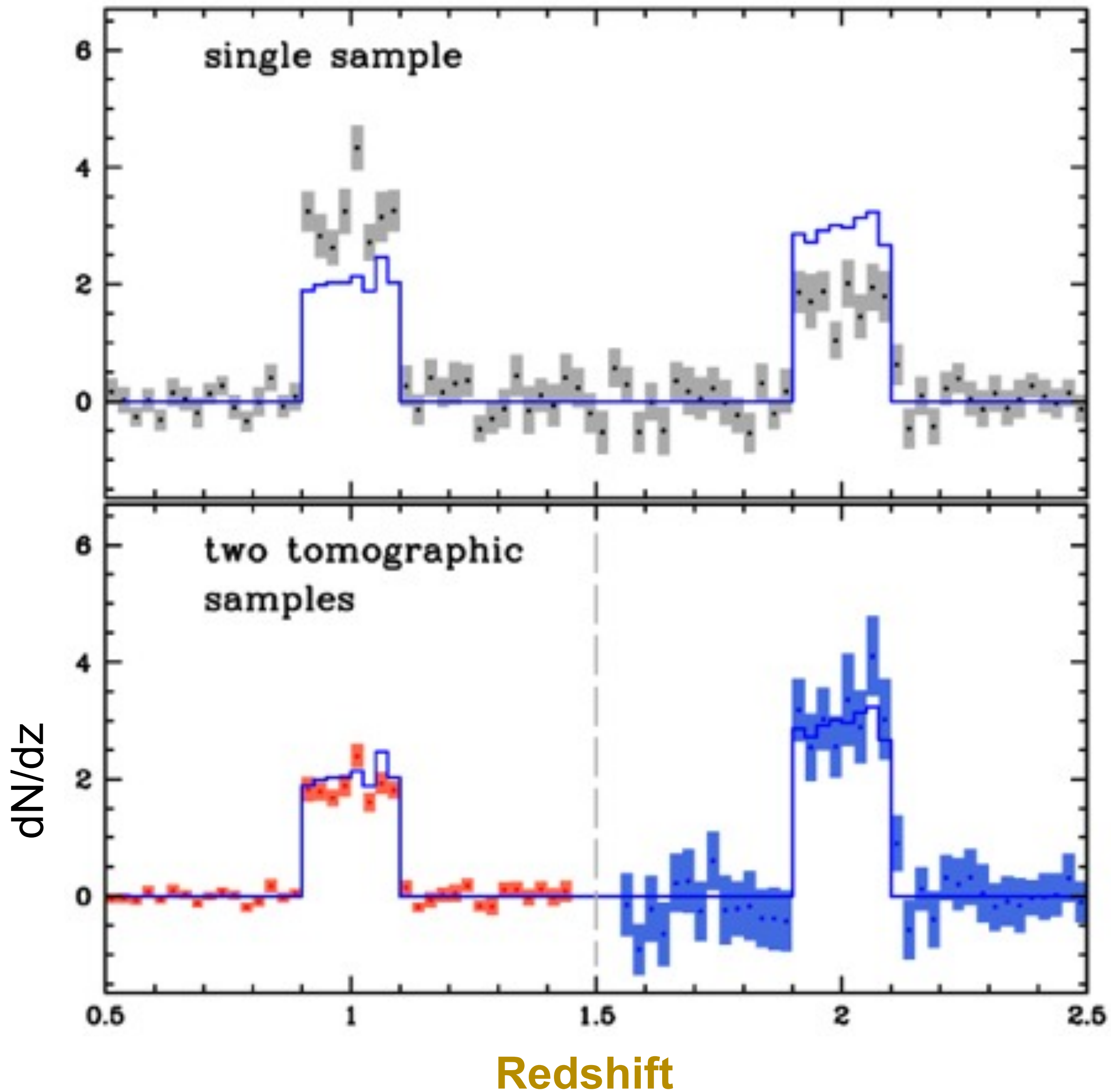
- Clustering exists on all scales but with different amplitudes. The dependence on scale can be explored with simulations



- The results depend only weakly on scale
- There is plenty of signal to be extracted down to small scales

Schmidt et al. (2013)

Bimodal redshift distributions

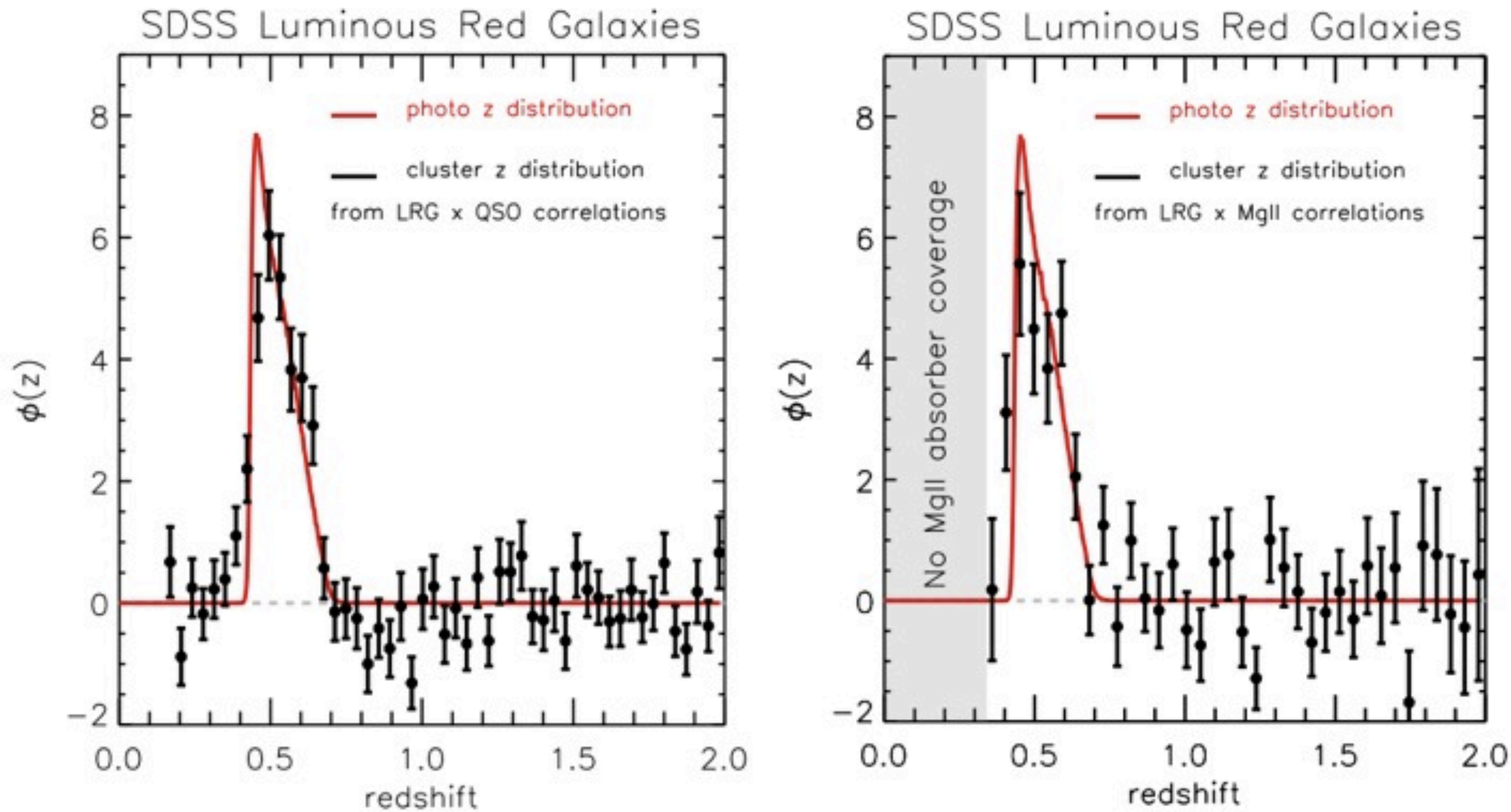


Schmidt et al. (2013)

Application to real datasets:

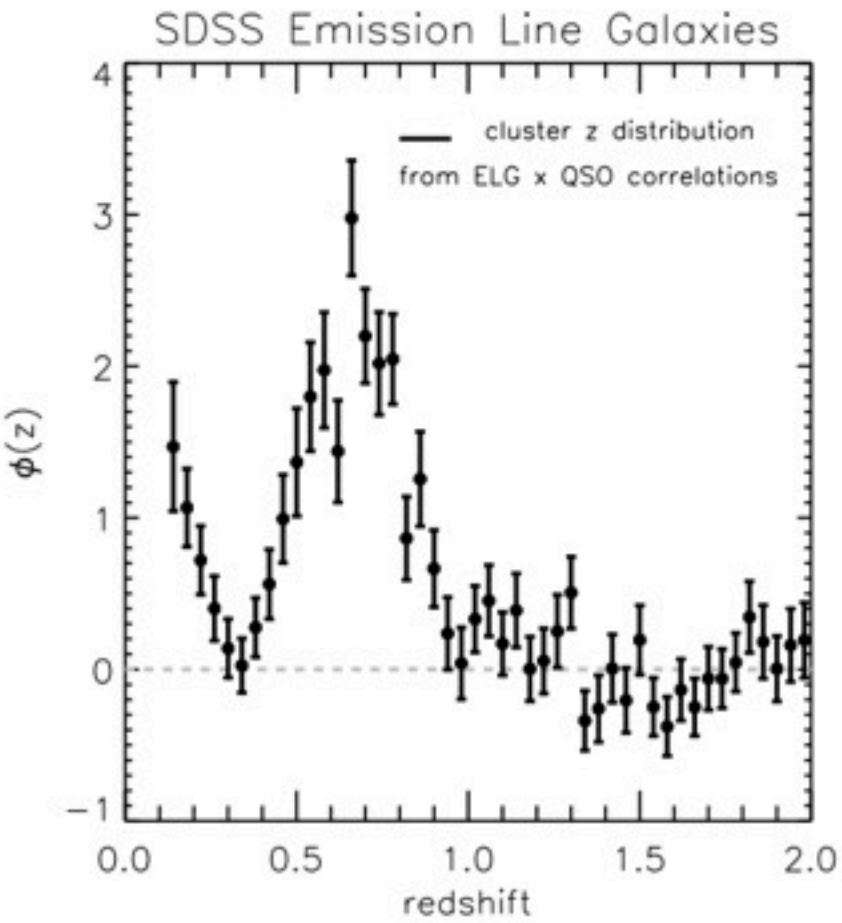
reference sample: spectroscopic
'unknown' sample: photometric

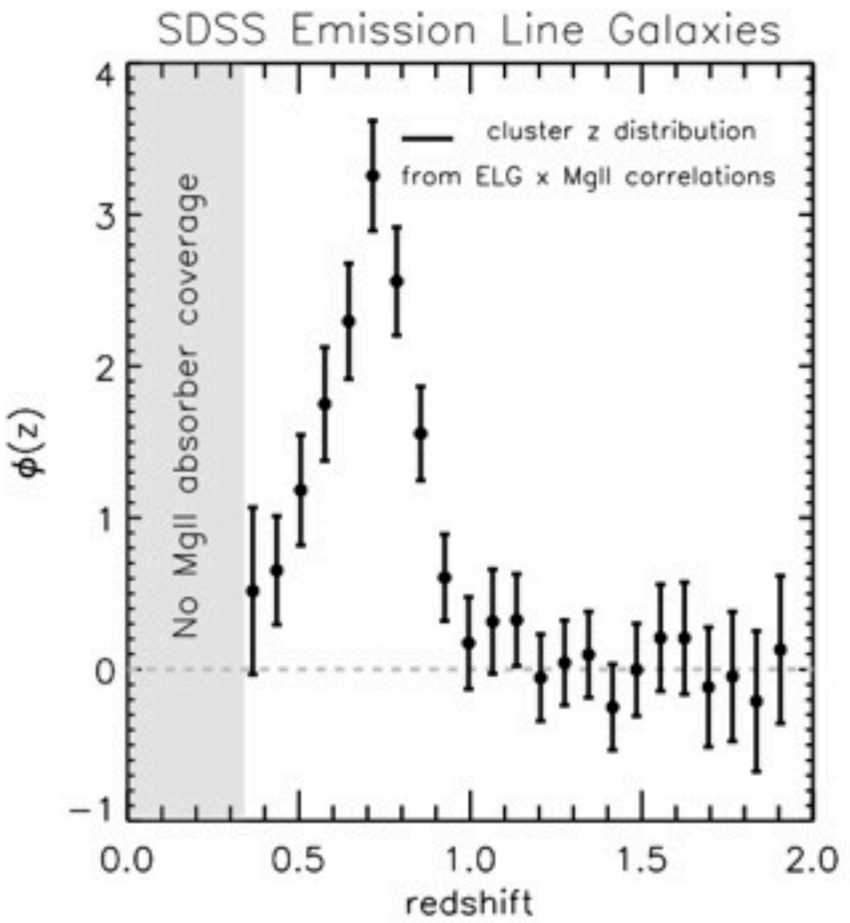
clustering-based redshifts



Ménard et al. (2013)

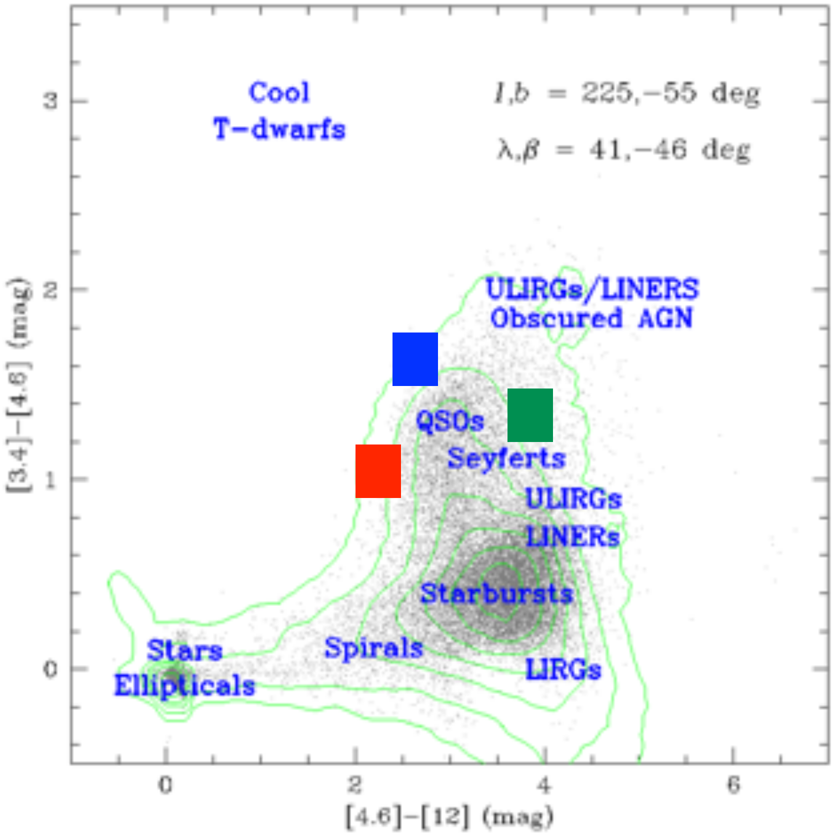
clustering-based redshifts

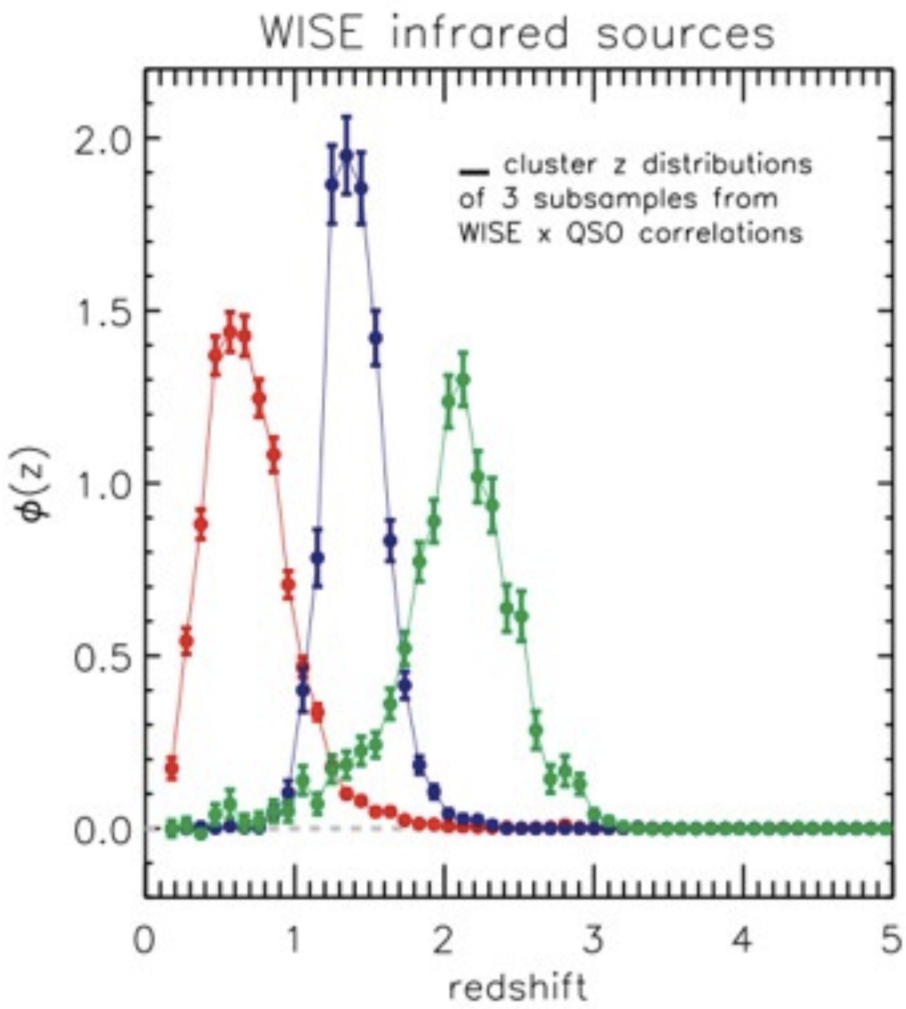




Ménard et al. (2013)

clustering-based redshifts

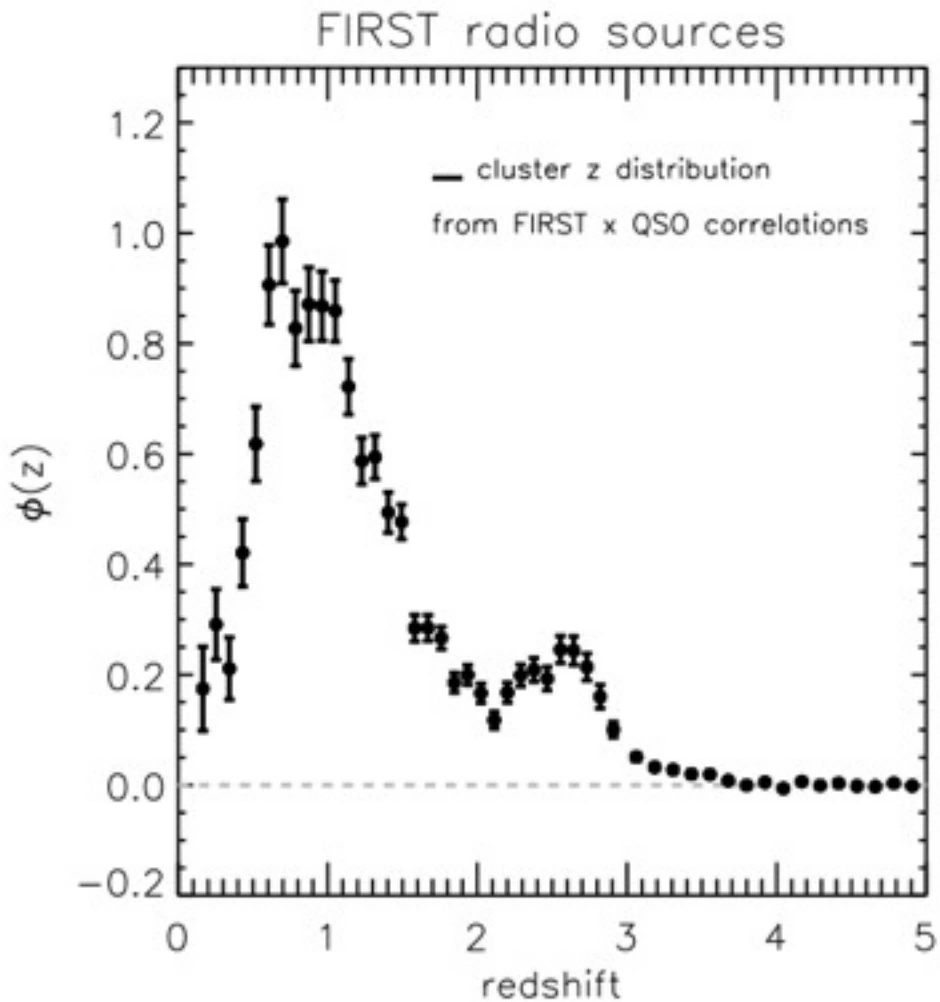




Ménard et al. (2013)

clustering-based redshifts

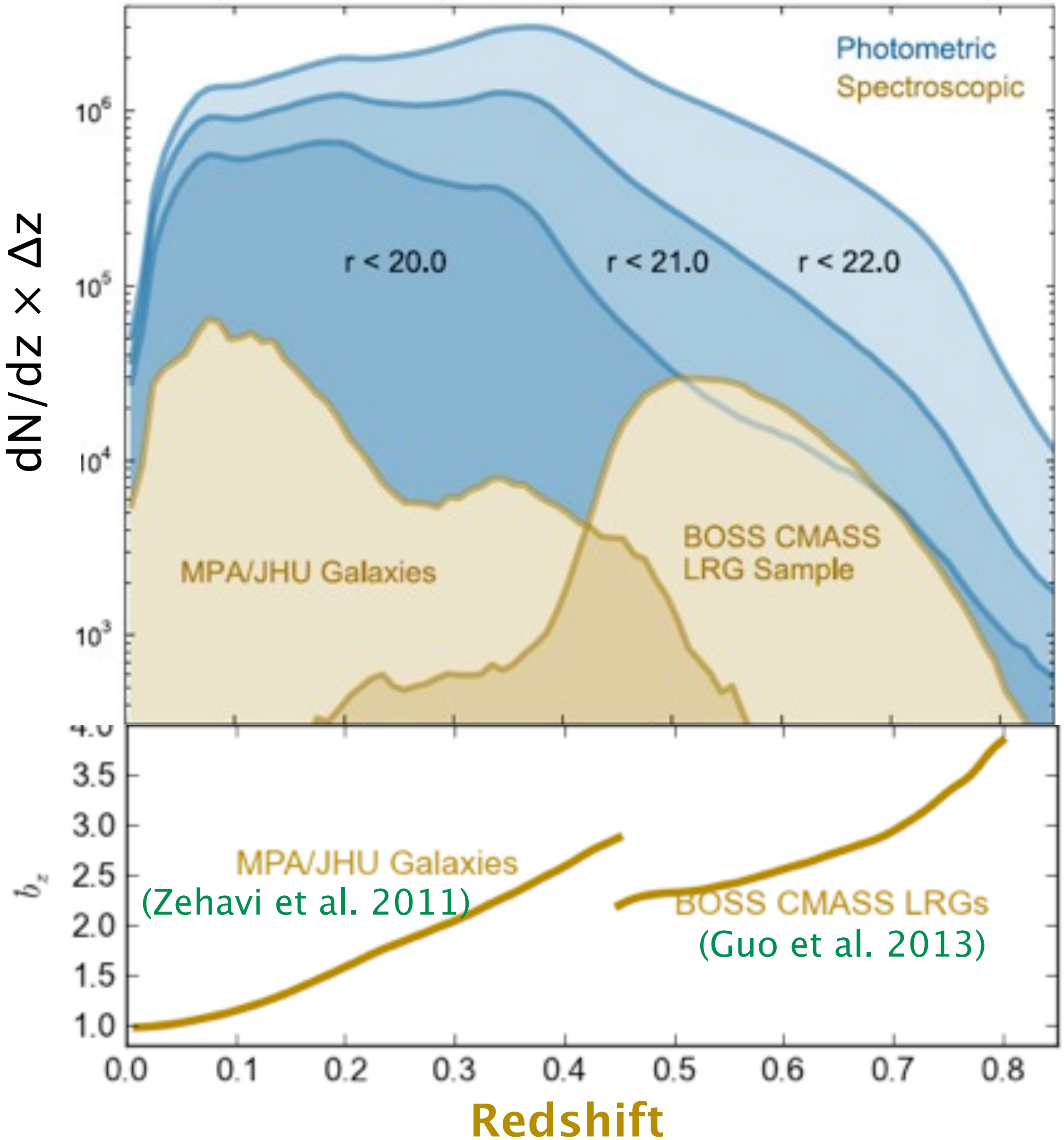
The FIRST VLA



Ménard et al. (2013)

Exploration of the SDSS photometric galaxies

Characterizing photometric galaxies



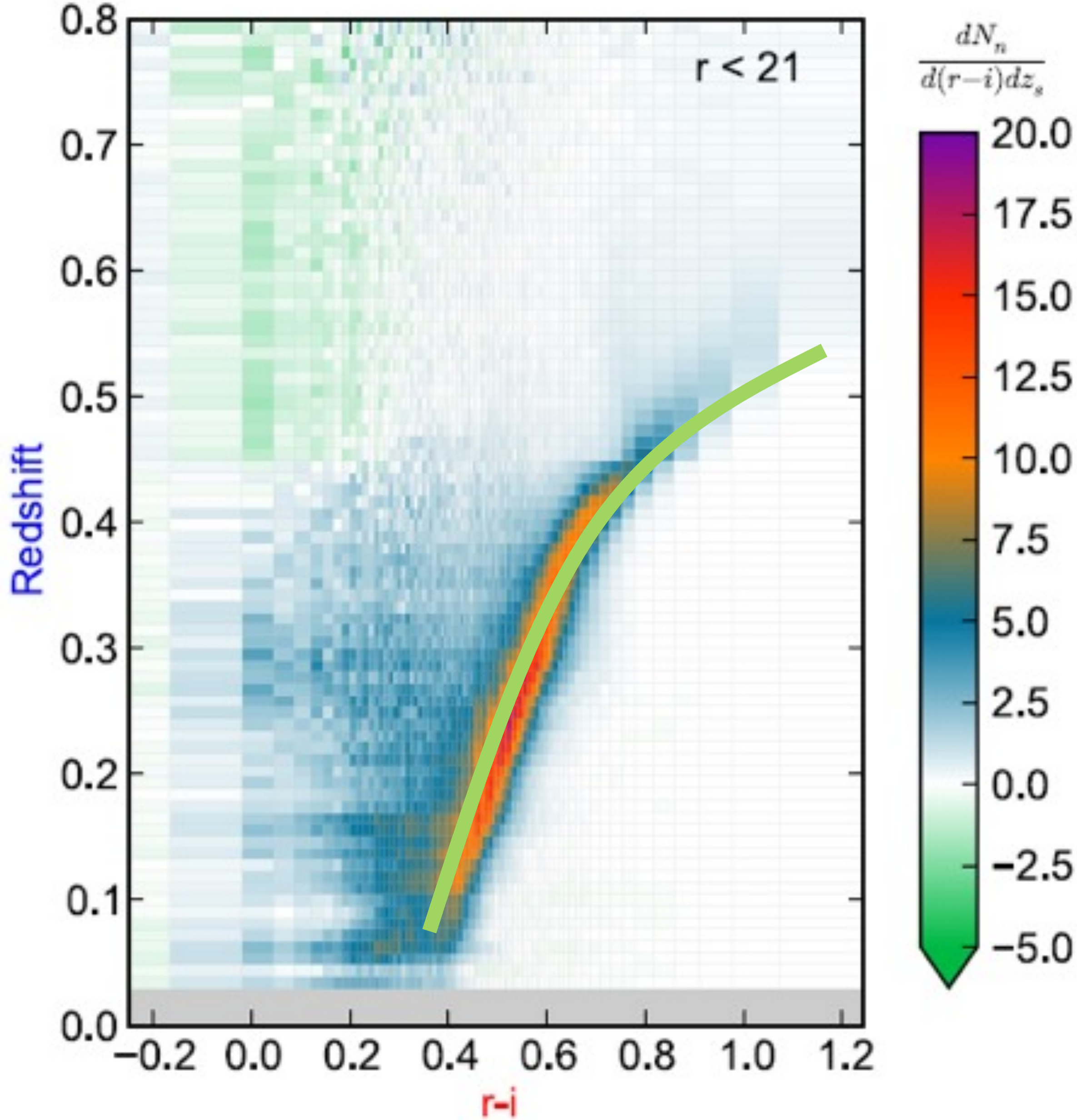
The (spectroscopic) reference sample is much smaller and does not need to be representative of the unknown galaxies

Cluster-z distribution of a color-selected sample

- Assembling clustering redshift distribution of all r–i selected samples

- 80 slices in r–i
- 80 slices in z ($\Delta z \sim 10^{-2}$)

- 6400 cross-correlation measurements

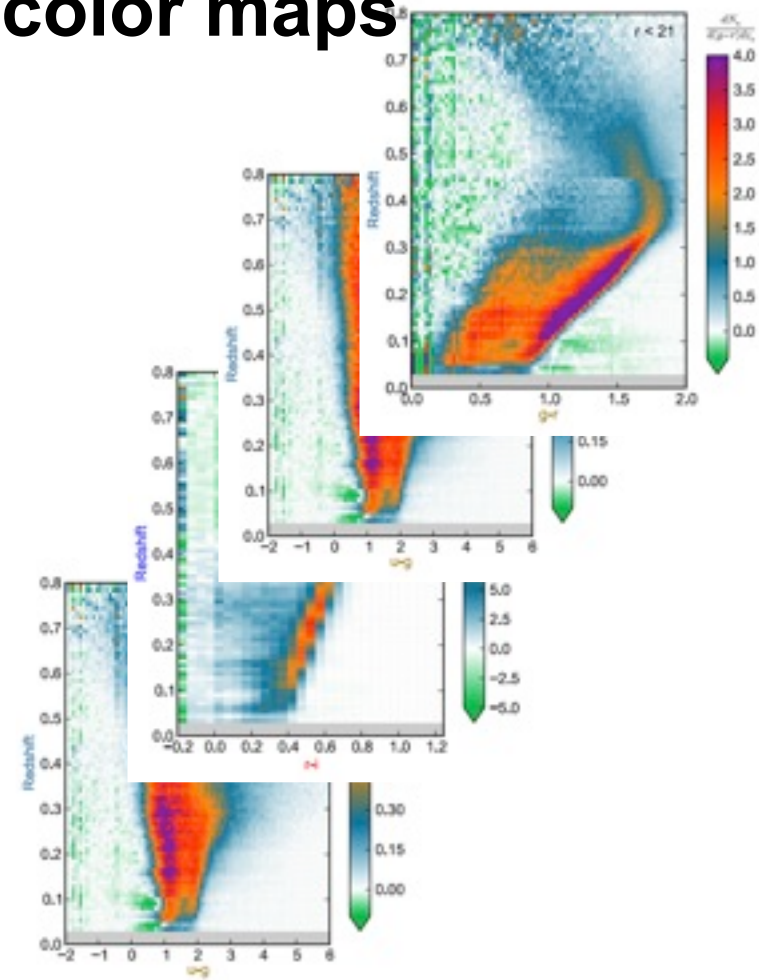


Rahman et al. (in prep)

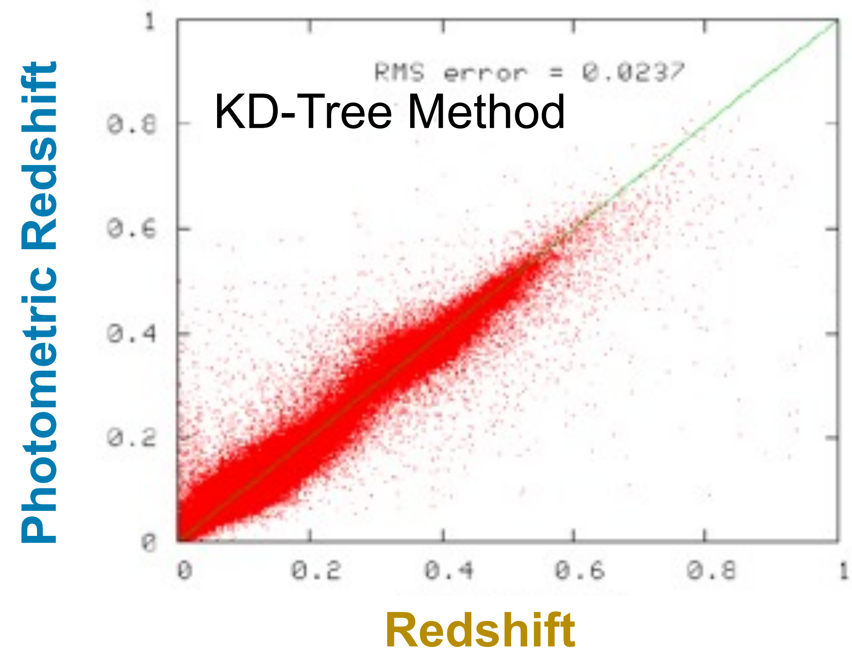
Reducing the dimensionality of the problem

4 independent color maps

Reducing to one dimension

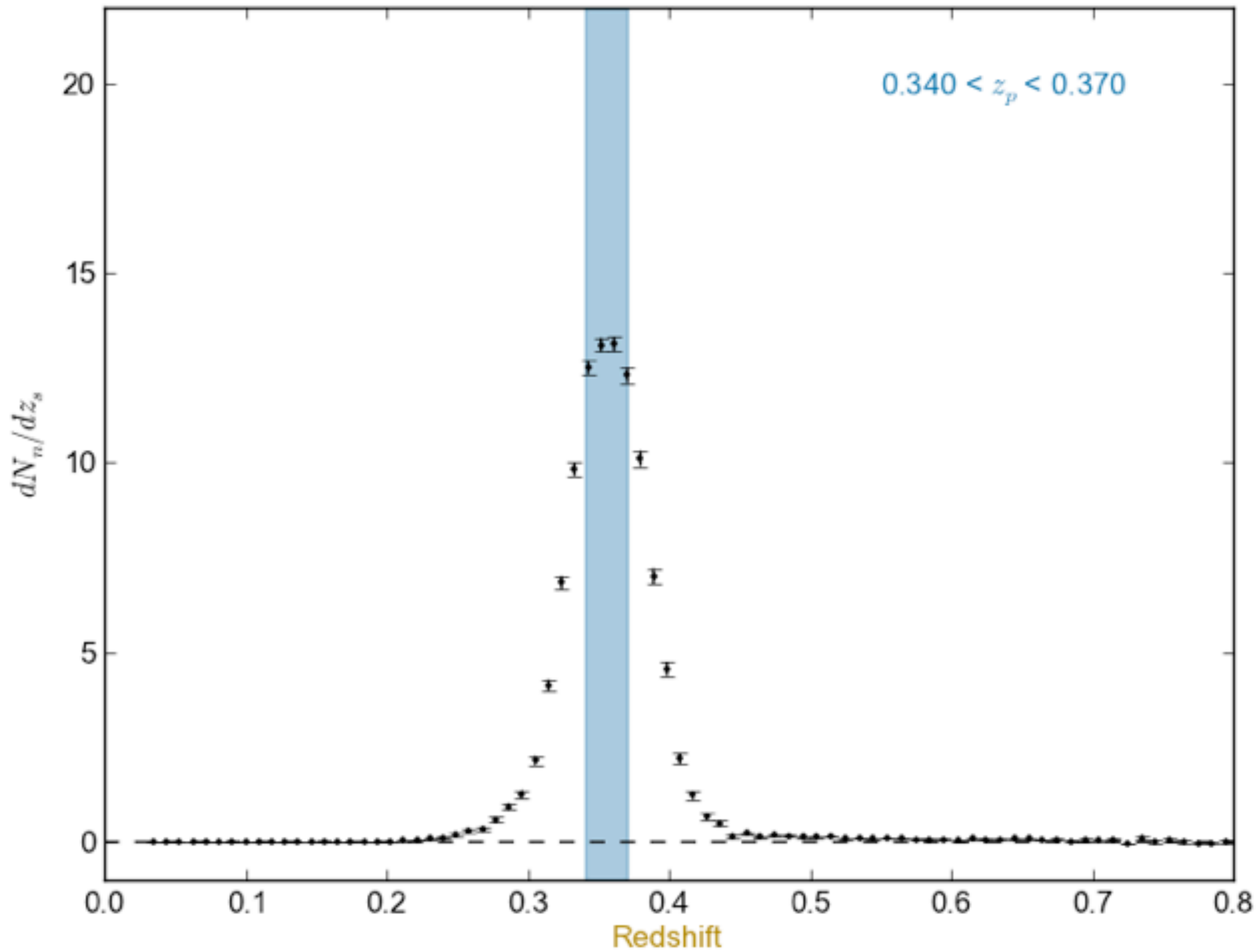


KD-Tree photometric redshifts



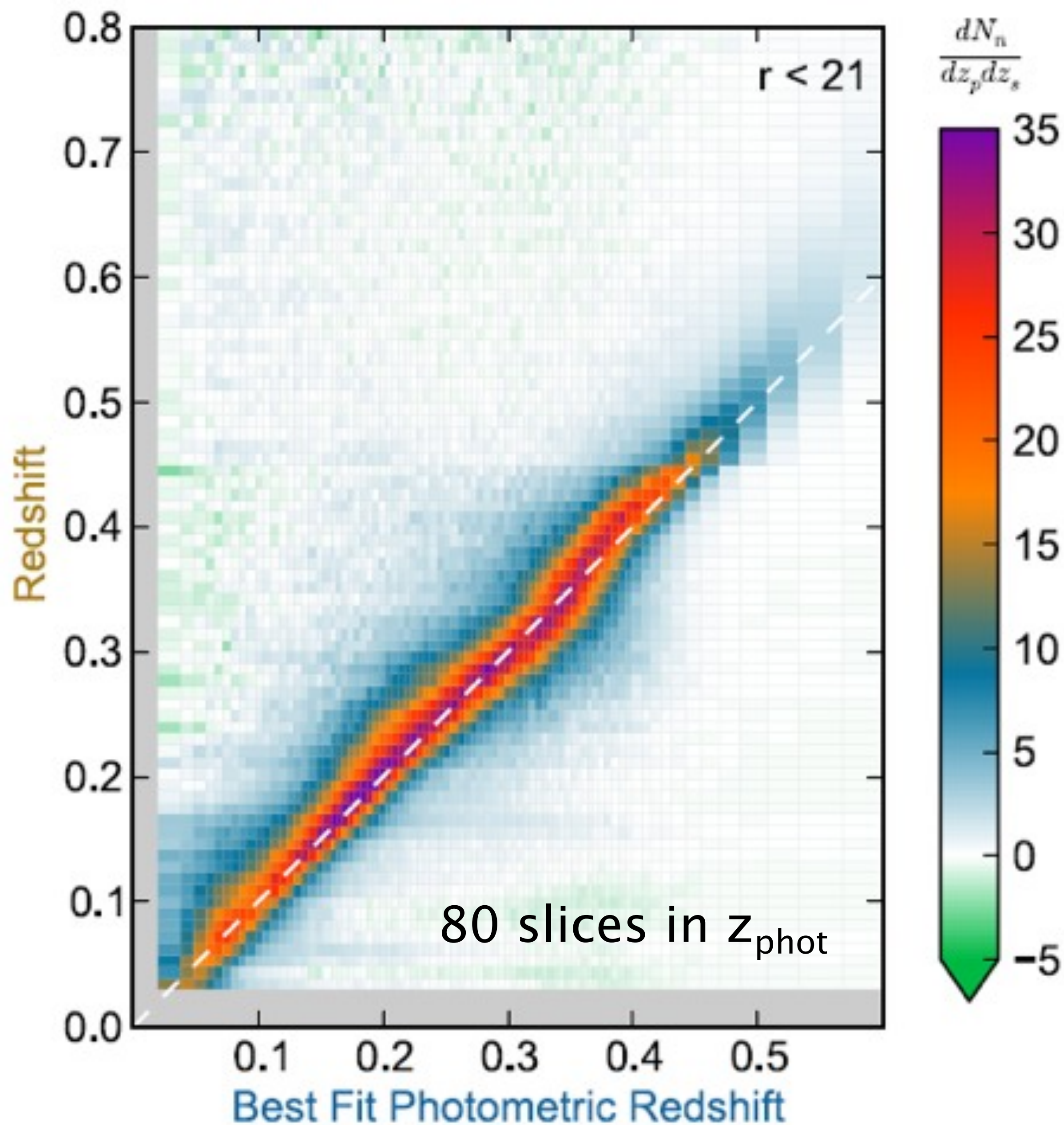
(Csabai et al. 2007)

Clustering redshift distribution



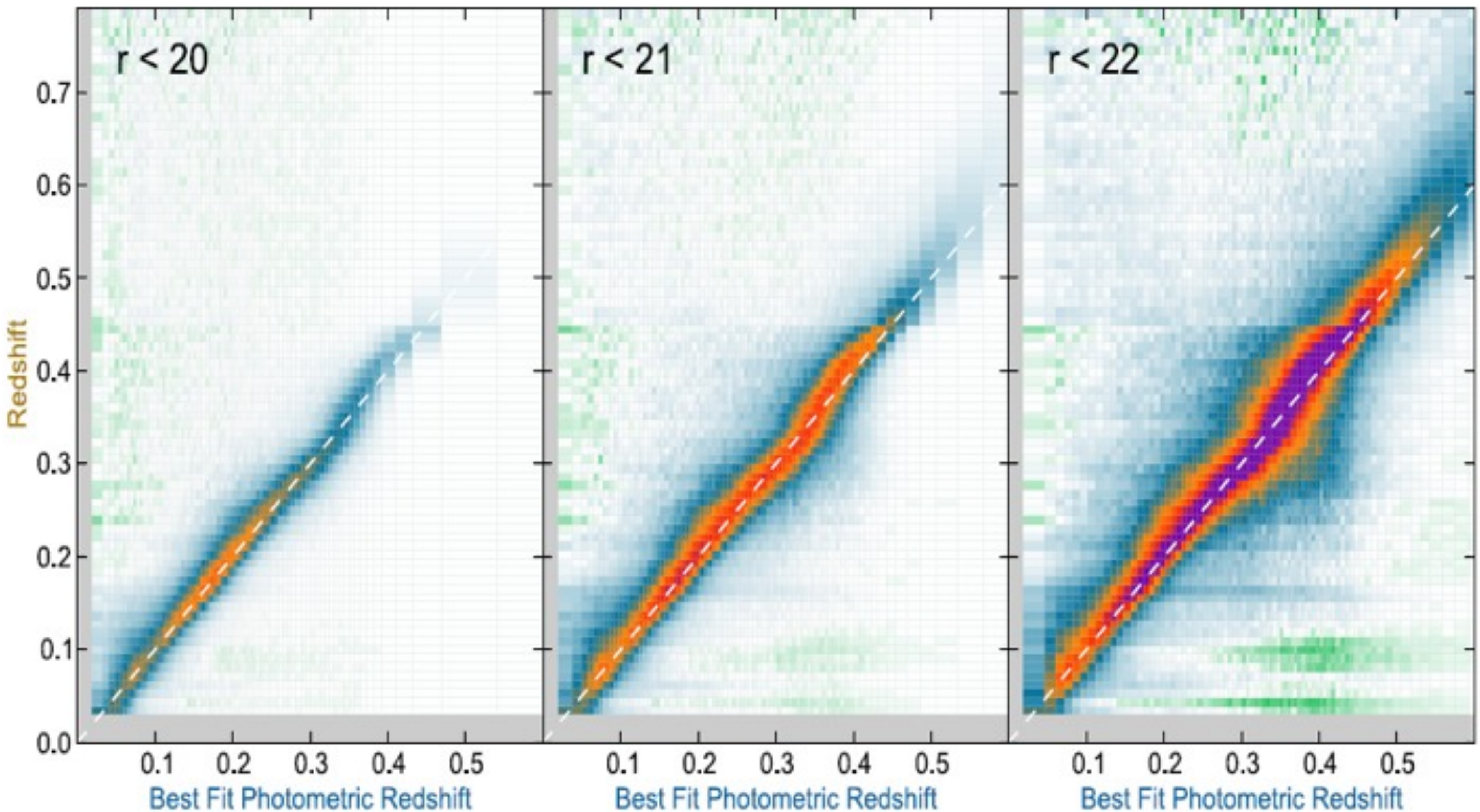
Density map of clustering-z distribution

80 slices in z
($\Delta z \sim 10^{-2}$)



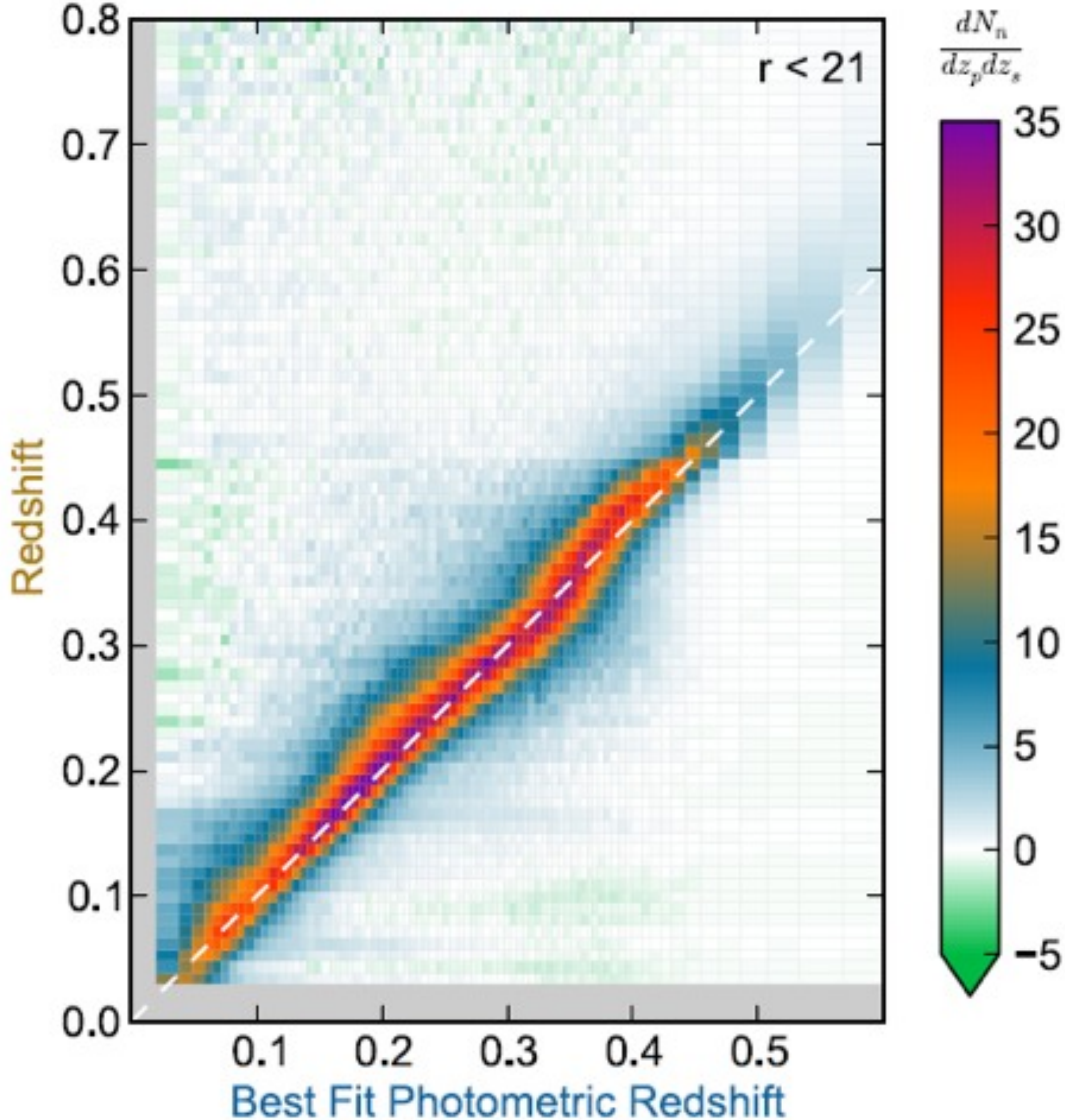
Rahman et al. (in prep)

Effect of limiting magnitude

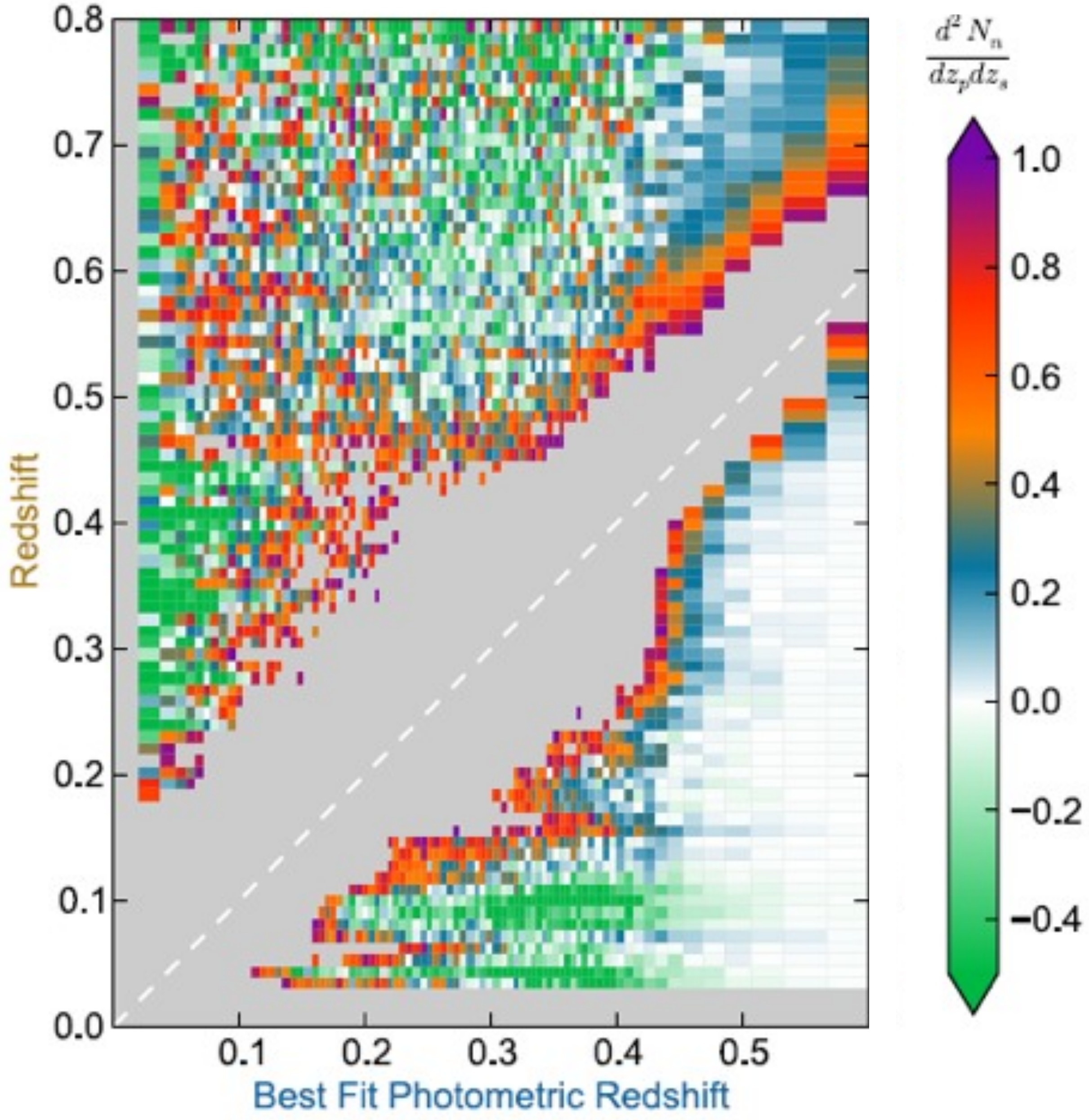


18 Million Objects 47 Million Objects 110 Million Objects

Density map of clustering-z distribution

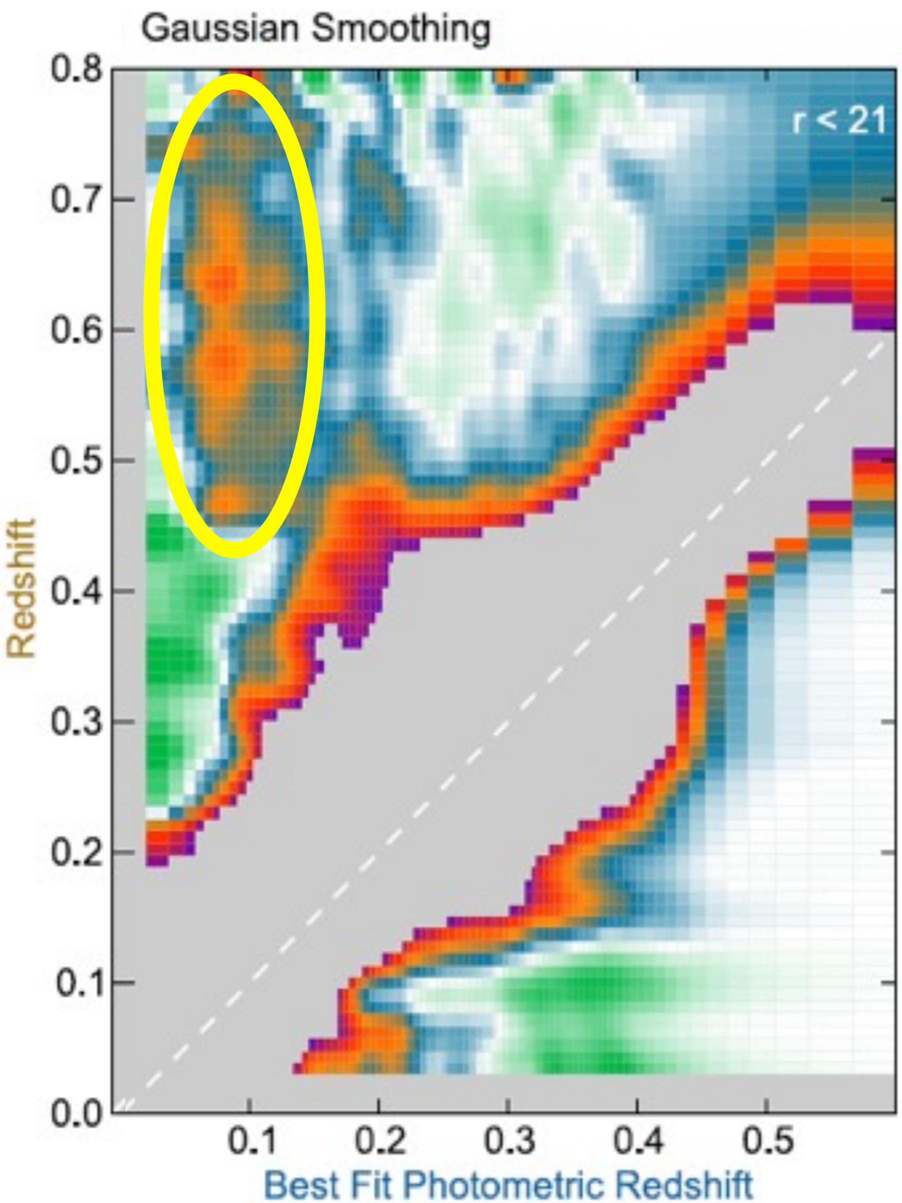


Density map of clustering-z distribution



Low-level Features

Locating the Emission Line Galaxies (ELGs)



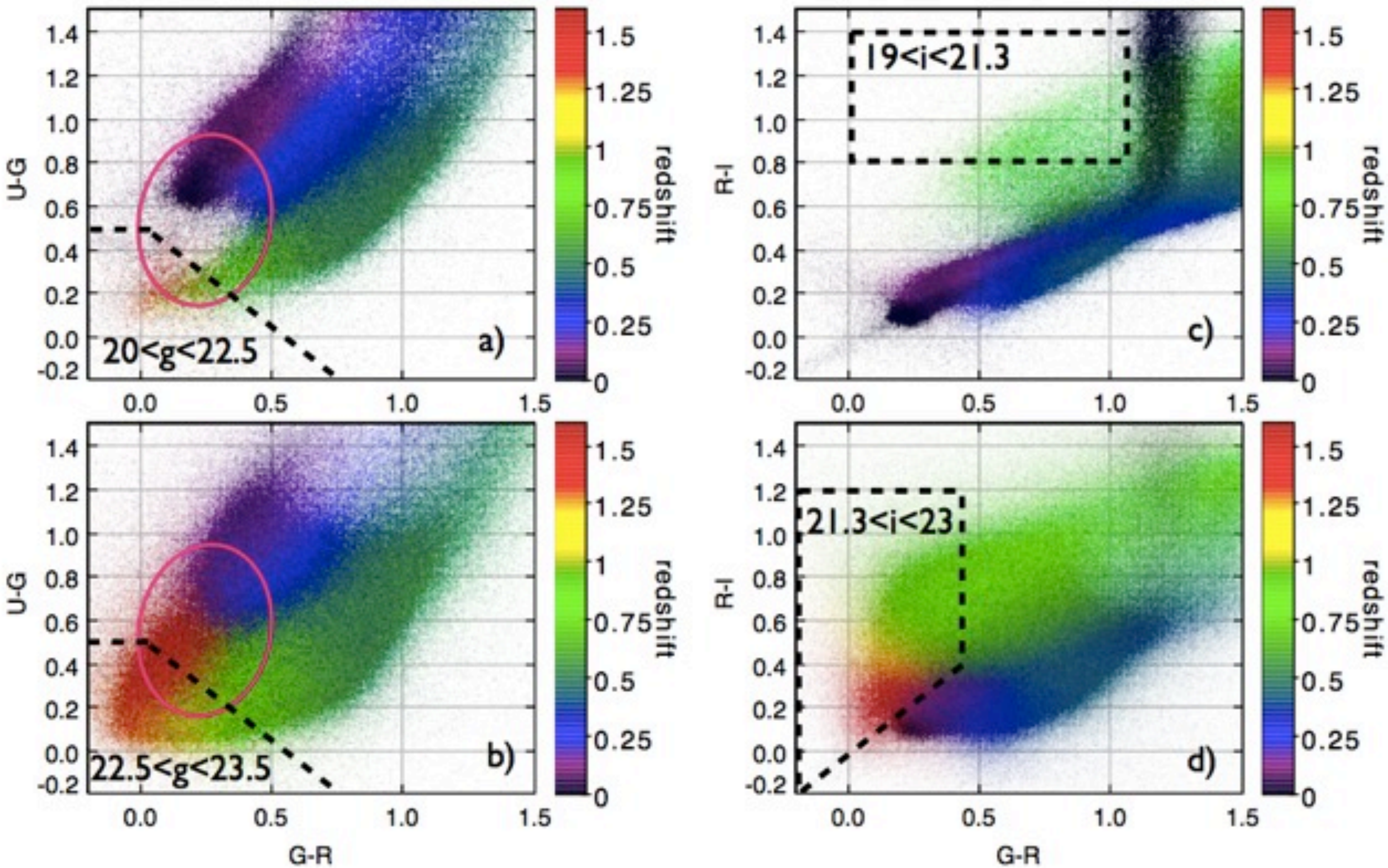
- Star forming (blue) galaxies at high z
- Colours dominated by emission lines: O III, O II and Hβ
- Problematic for photo-z estimates
- we could have discovered them a long time ago
- They have become key populations for upcoming surveys

Photometric selection of ELGs

Goal: selecting galaxies with $0.6 < z < 1.7$

the “ugr” selection

the “gri” selection



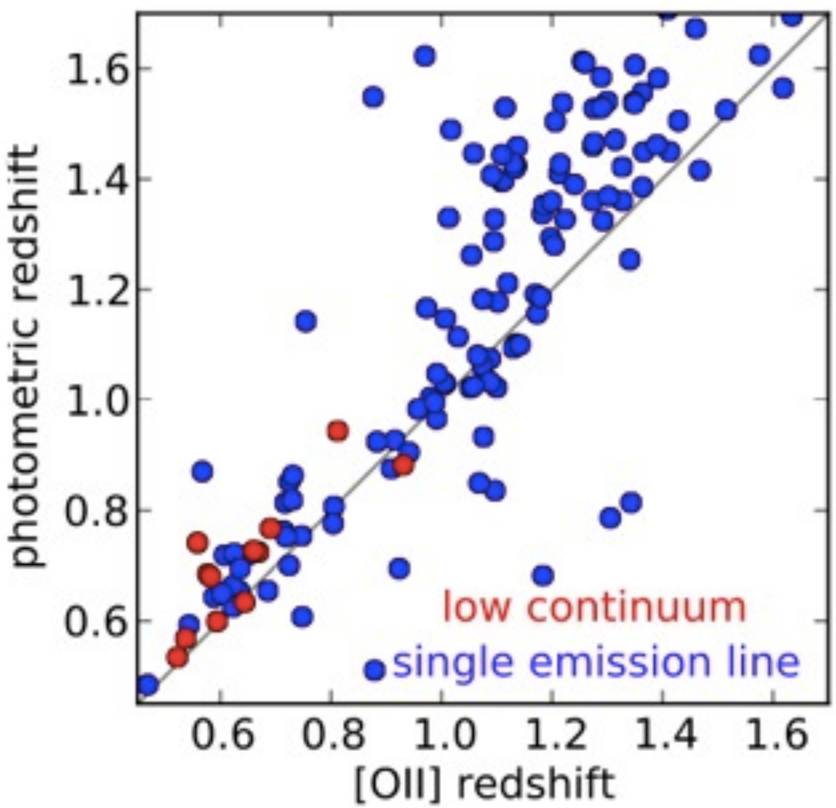
Comparat, Kneib et al. et al. (2012)

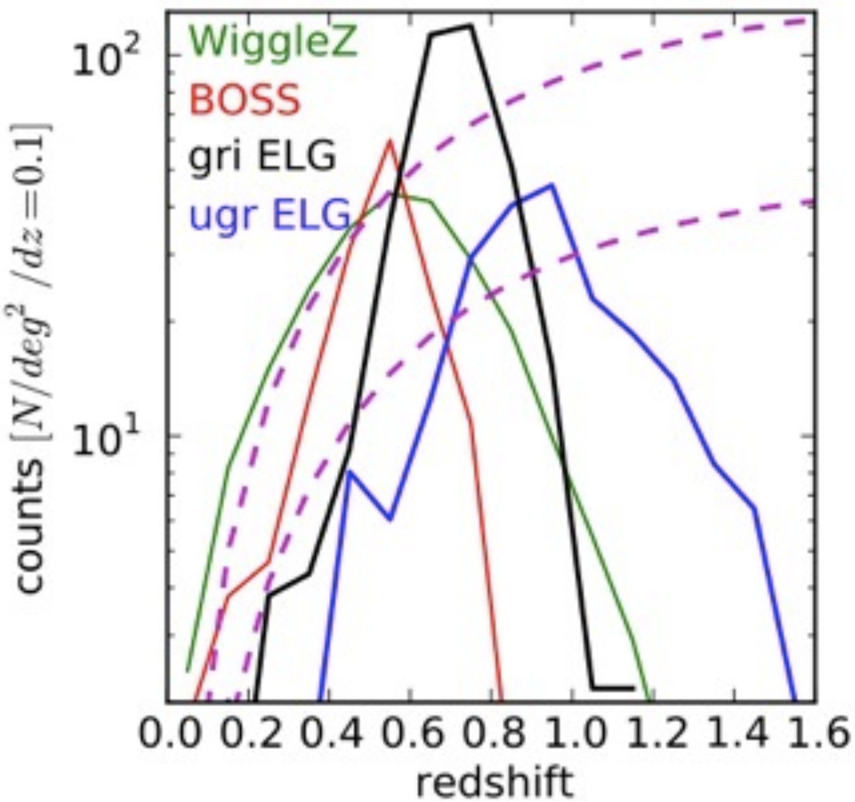
Photometric selection of ELGs

ELG SDSS-III/BOSS ancillary program, 2000 spectra:

Objects with robust redshifts: blue galaxies, red galaxies, QSOs,

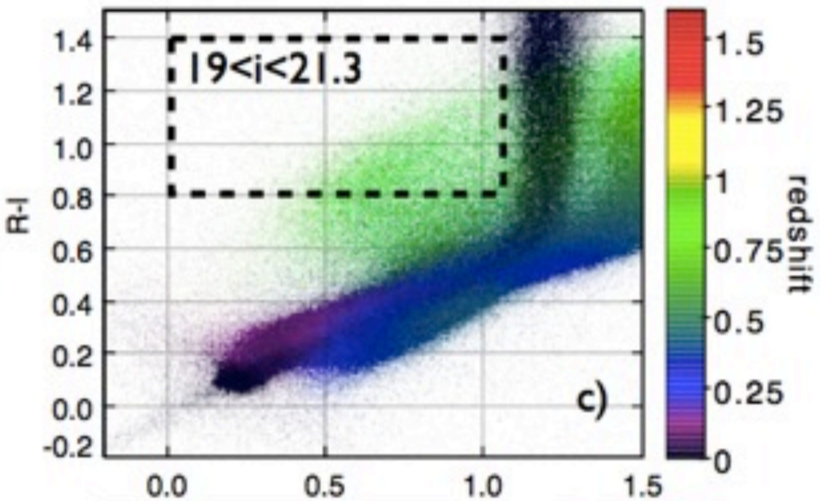
Objects with unreliable redshifts: single emission line, low continuum level, bad quality data



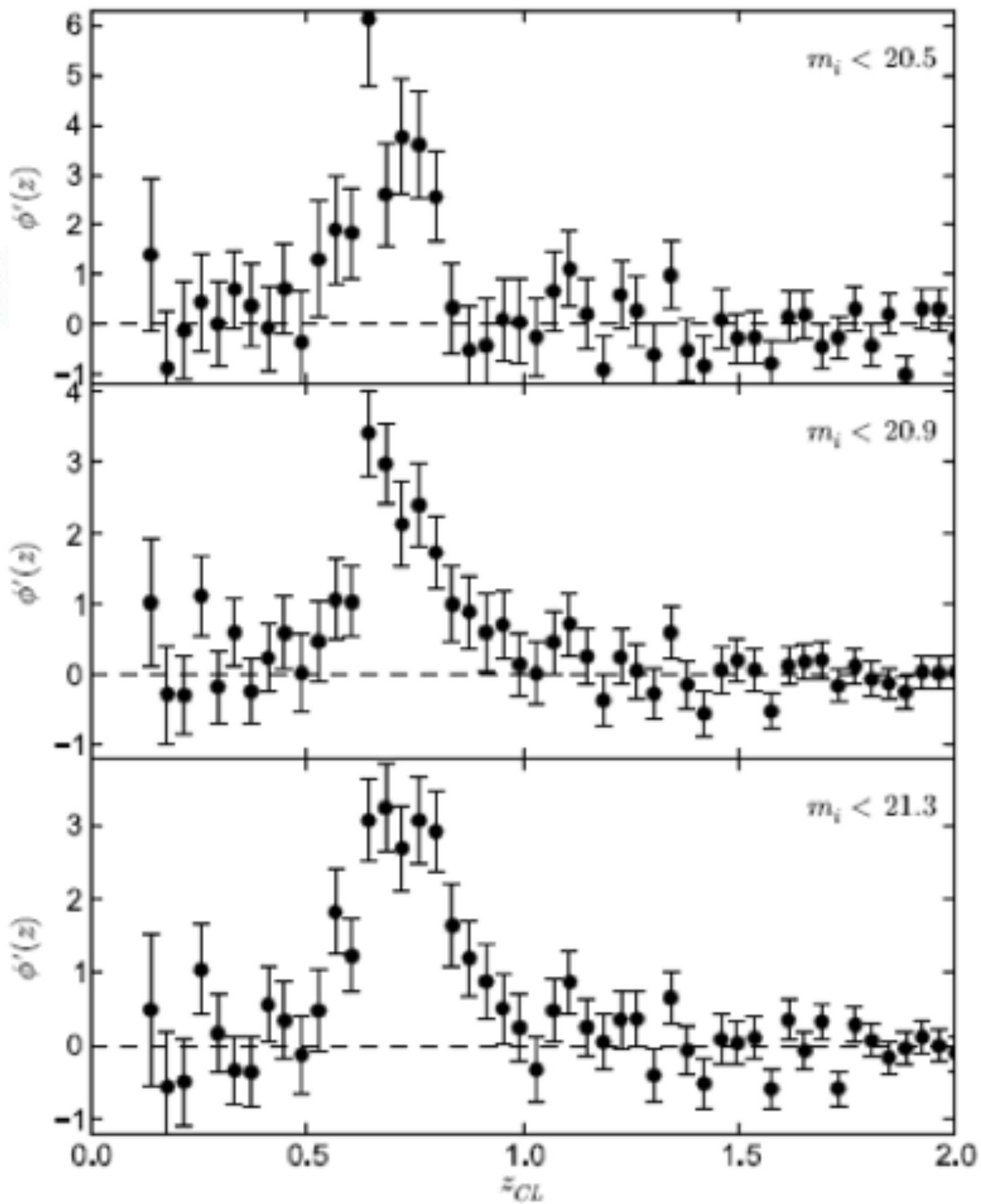


Comparat, Kneib et al. et al. (2012)

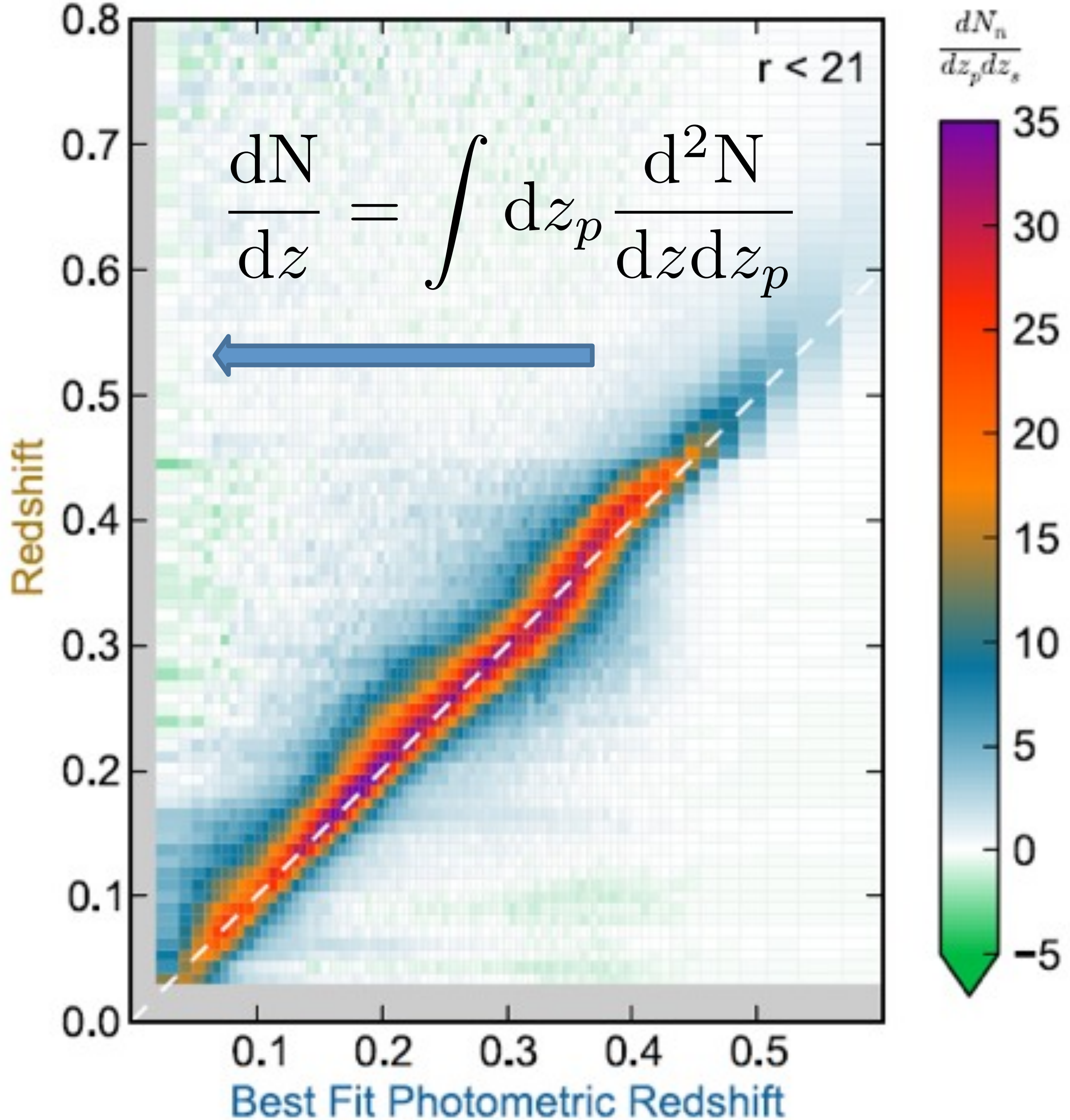
ELG redshift distributions



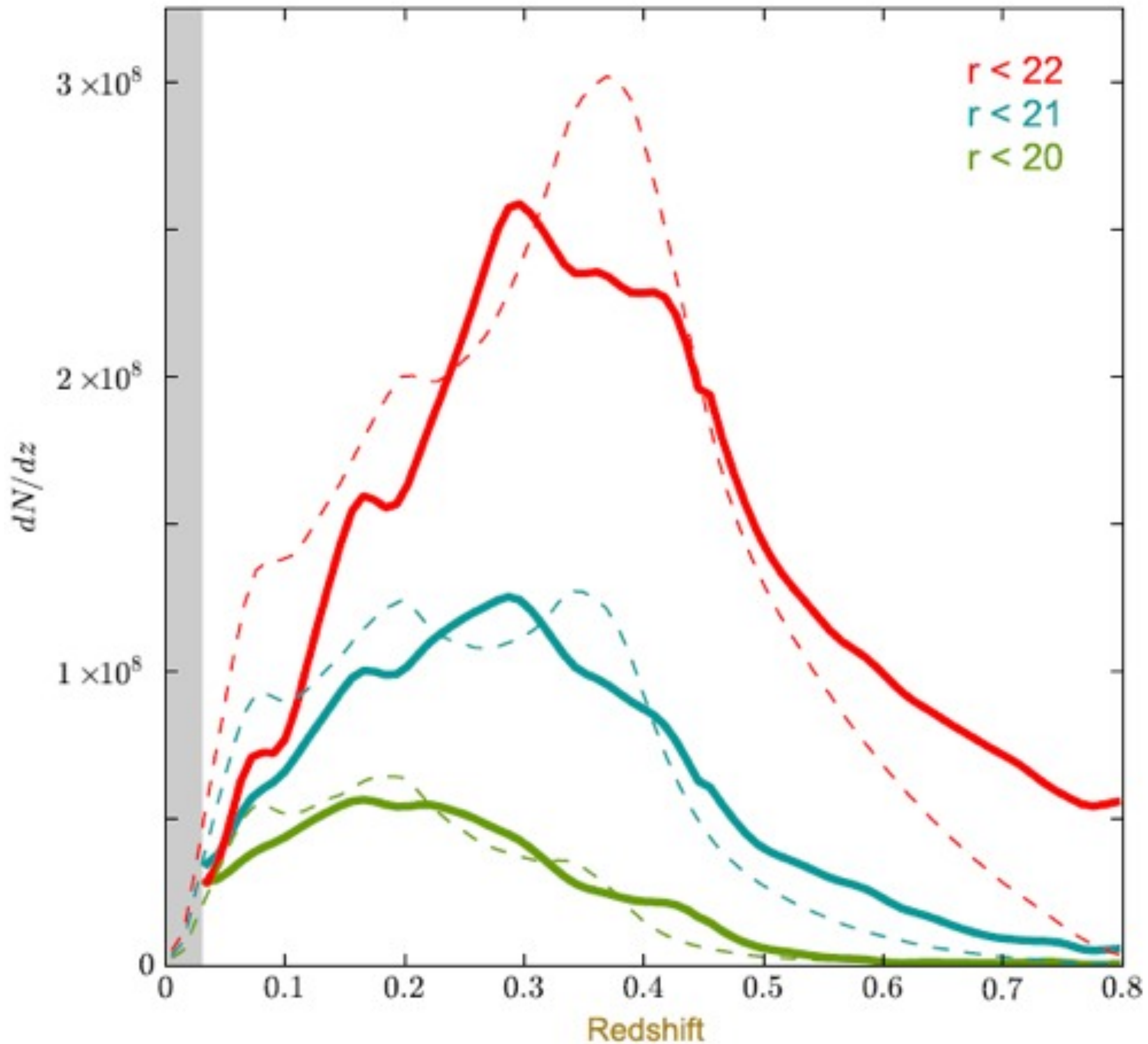
gri selection from
Comparat et al. (2013)



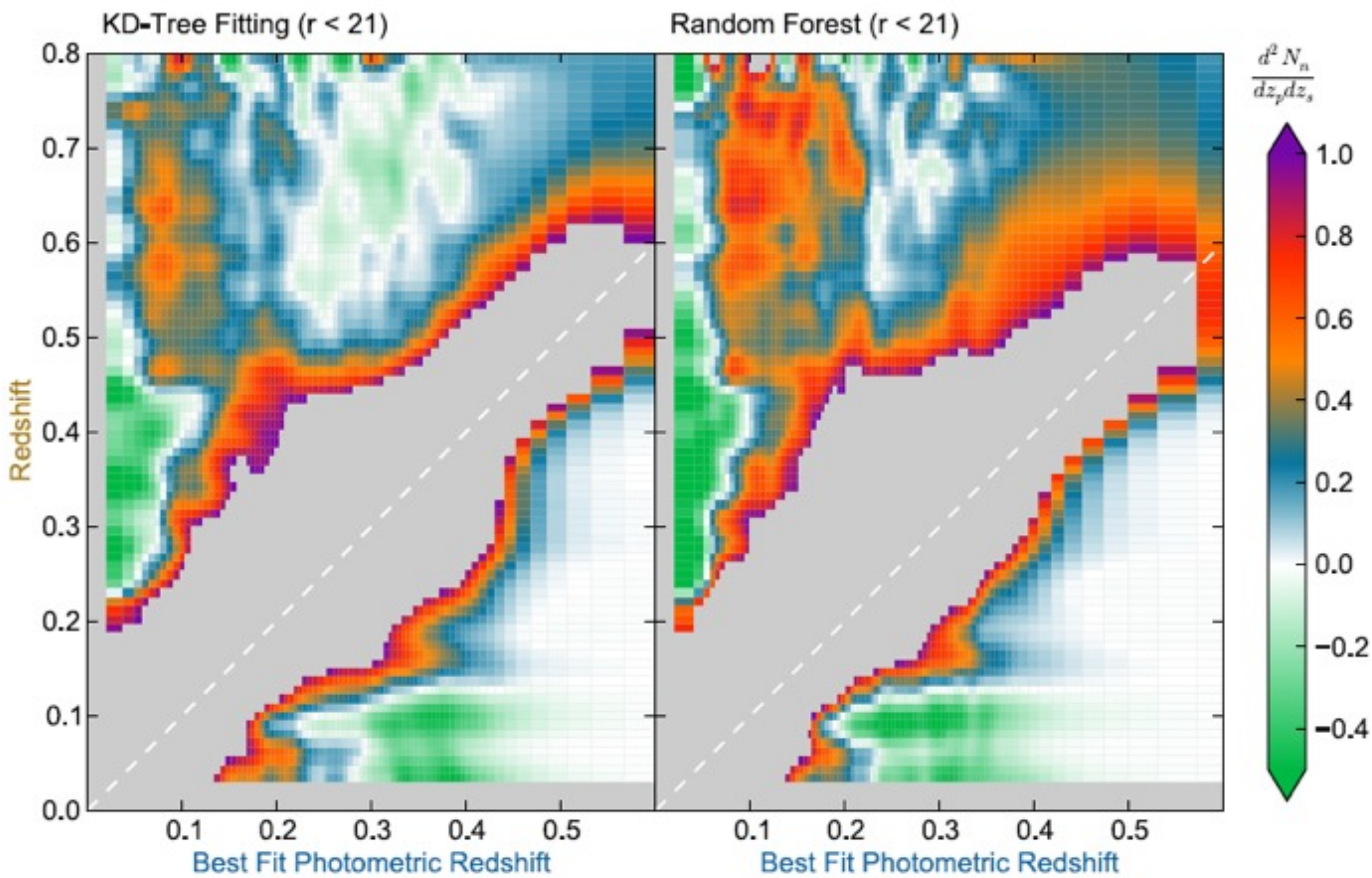
Global redshift distribution



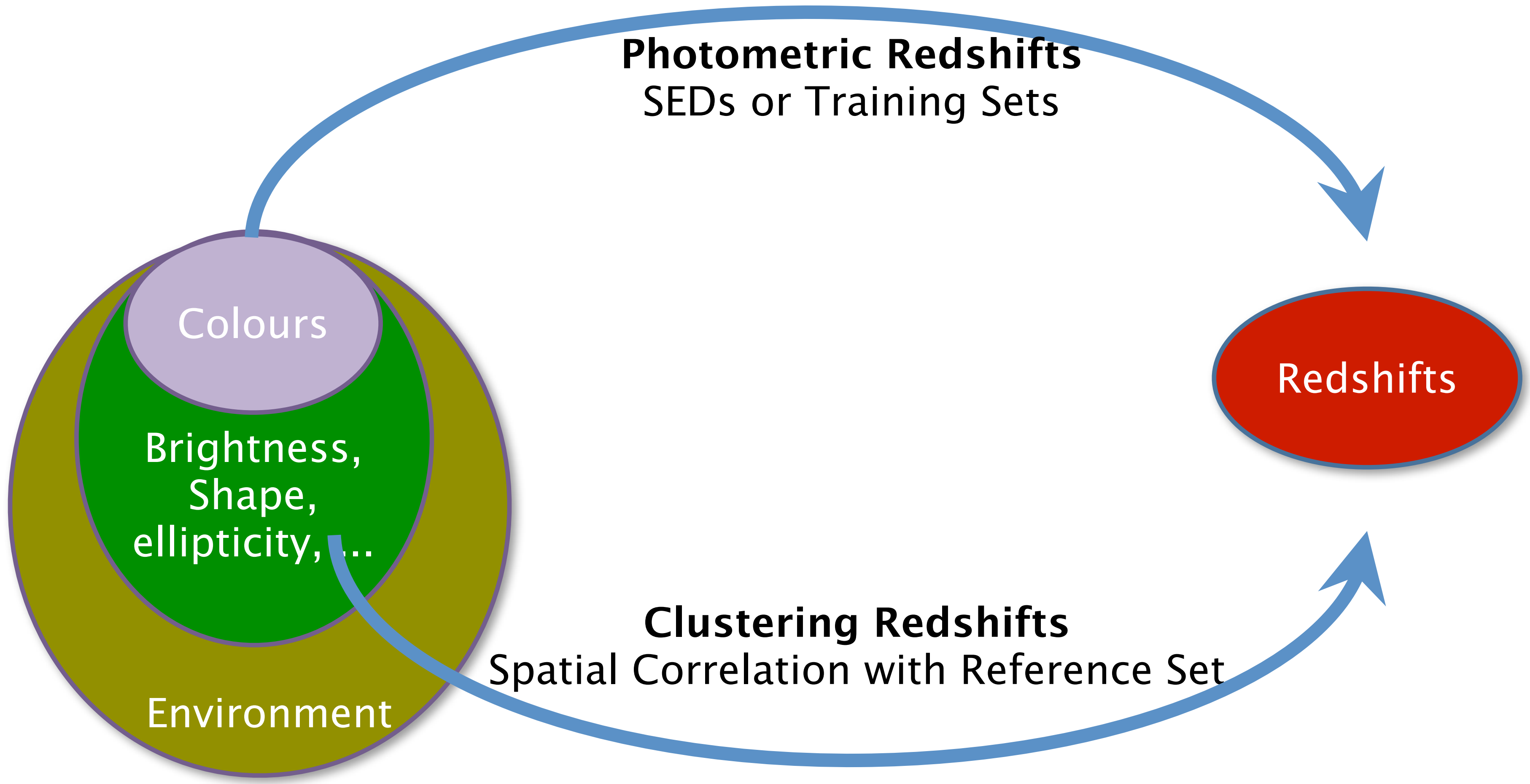
Global redshift distributions inferred from photo-zs and cluster-zs



Comparing different photo-z methods



What’s next?



- multidimensional sampling/selection

This will be done without any reference to photo-zs

This can be used to infer the redshift pdf of one galaxy

Summary

• Spatial correlations give us an estimate of b(z) . dN/dz

• This can be used as a tool to explore the 3rd dimension of the Universe.

• This method can process *any* dataset:

  ⋆ does not require any spectral feature
  ⋆ can even be done with one band
  ⋆ can be applied to diffuse signals

• Interesting preliminary results. More to come soon...