Clustering-redshifts

Brice Ménard
Johns Hopkins University
& Kavli IPMU Tokyo University

Mubdi Rahman
Johns Hopkins University

Ryan Scranton, Sam Schmidt & Chris Morrison
(UC Davis)

- Spectroscopic redshifts

- Photometric redshifts

- Clustering redshifts
  - the method
  - validation
  - exploration

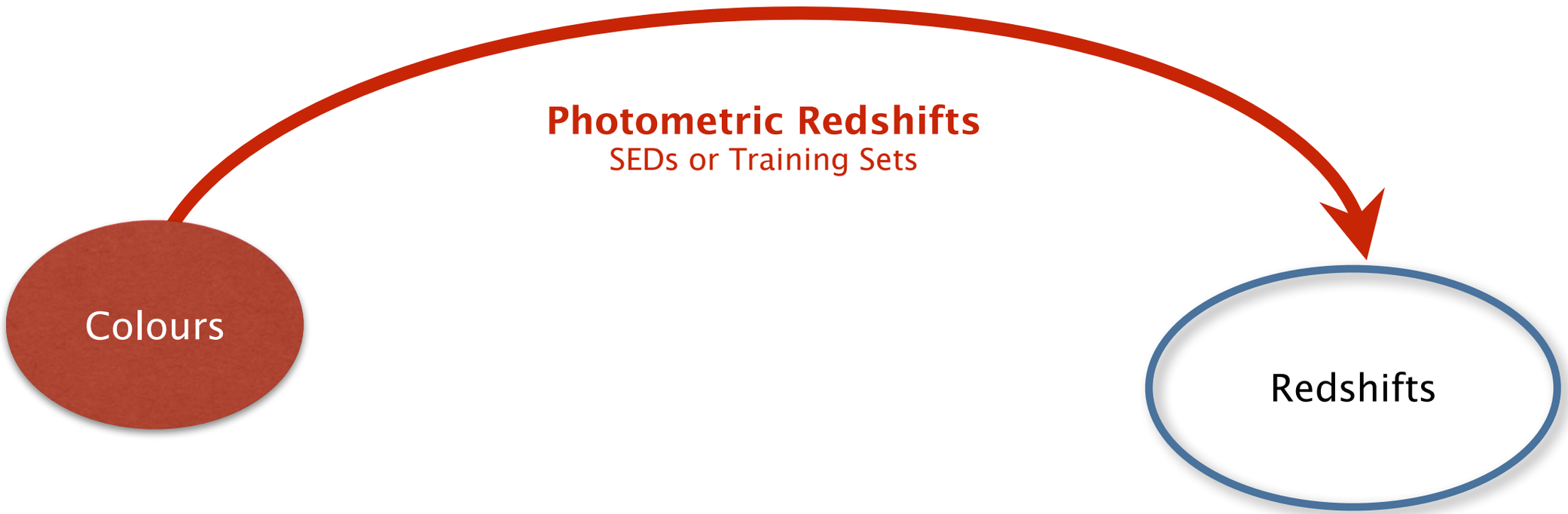
spectroscopic redshifts

- They are well defined as long as spectral features are present and unambiguously identified

- The source needs to be bright enough

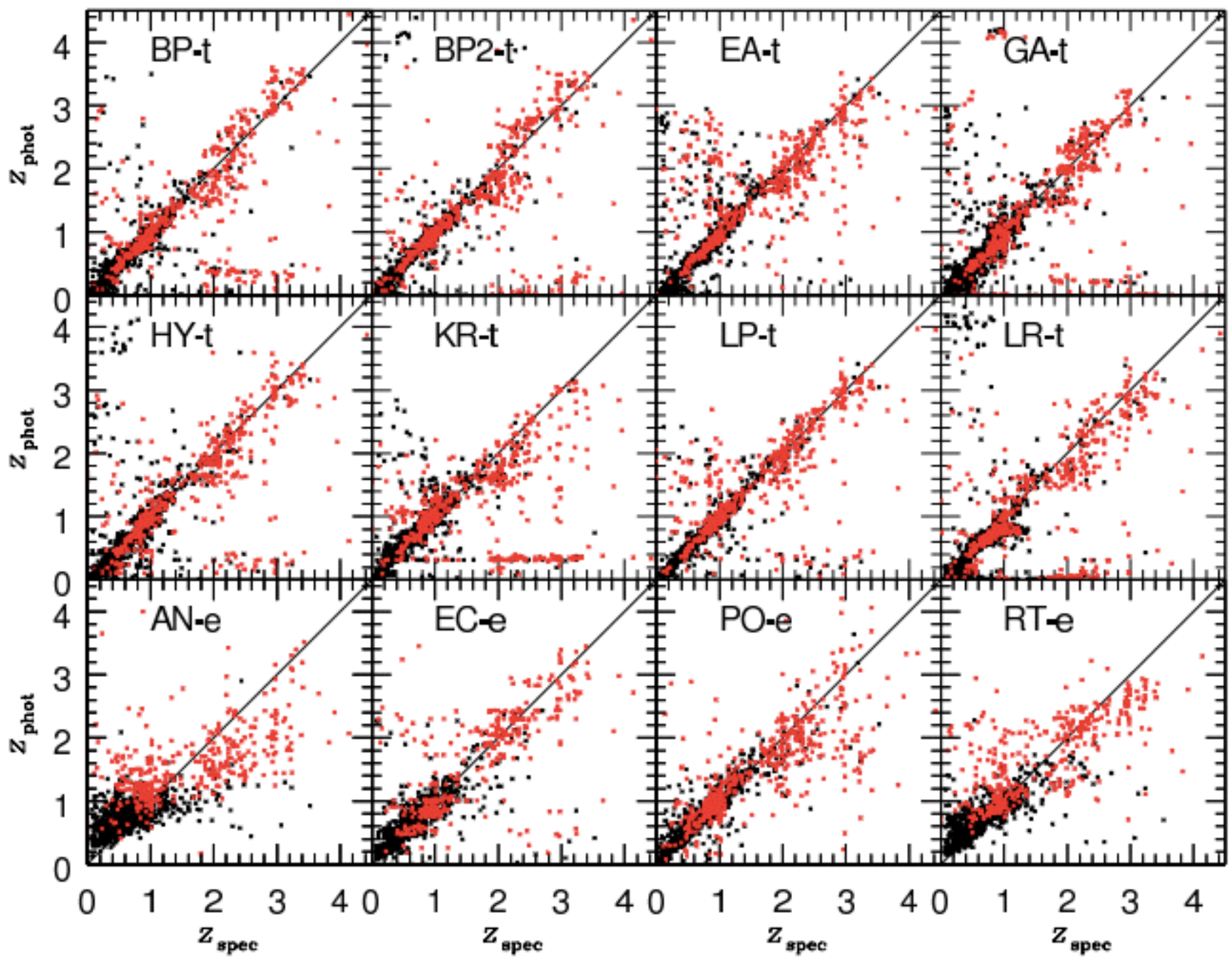
- They are expensive

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


Photometric Redshifts
SEDs or Training Sets
Colours
Redshifts

Photometric redshifts

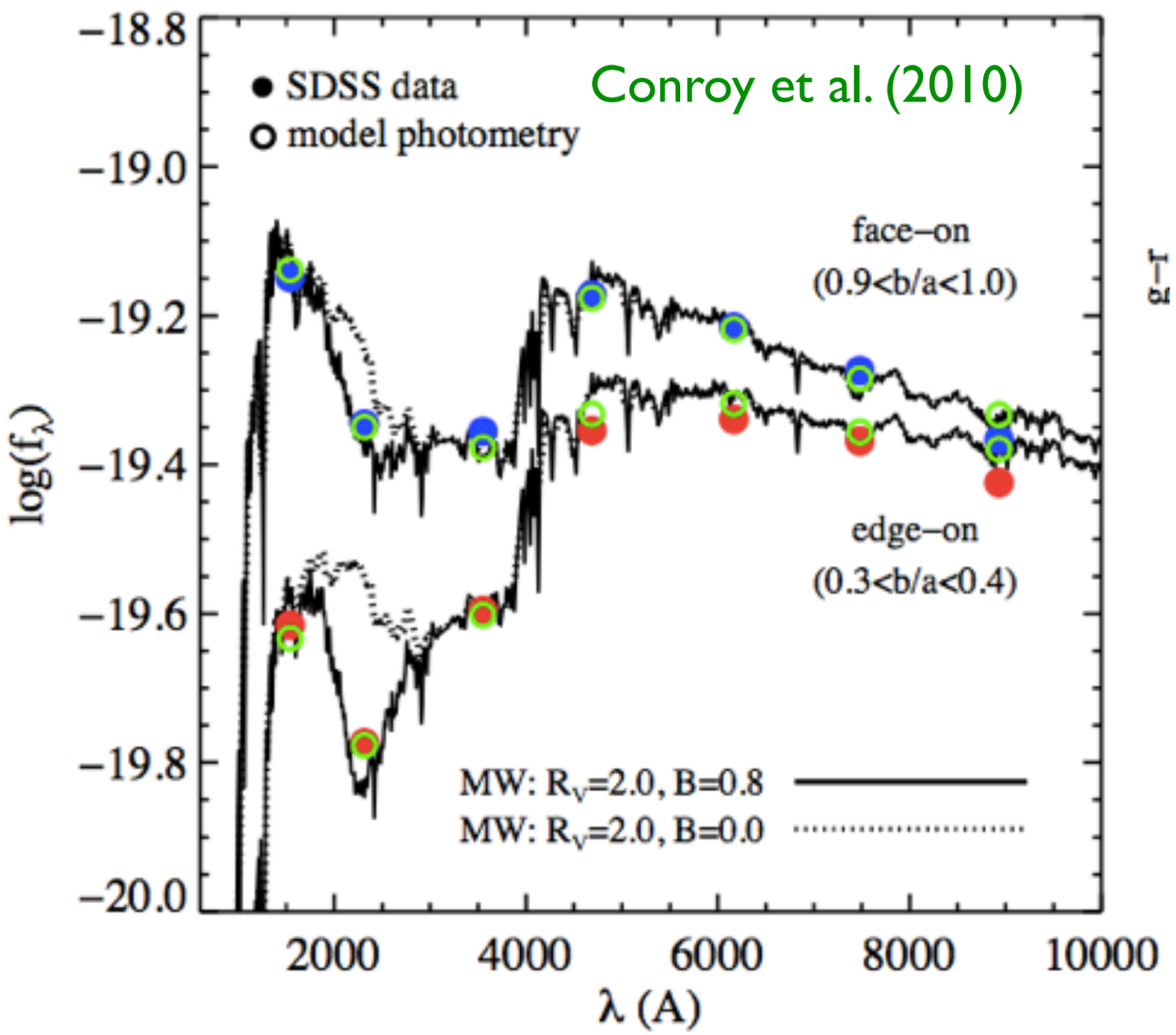
- They rely on templates (theoretical or observed)
- They require training sets. The answer is not unique.

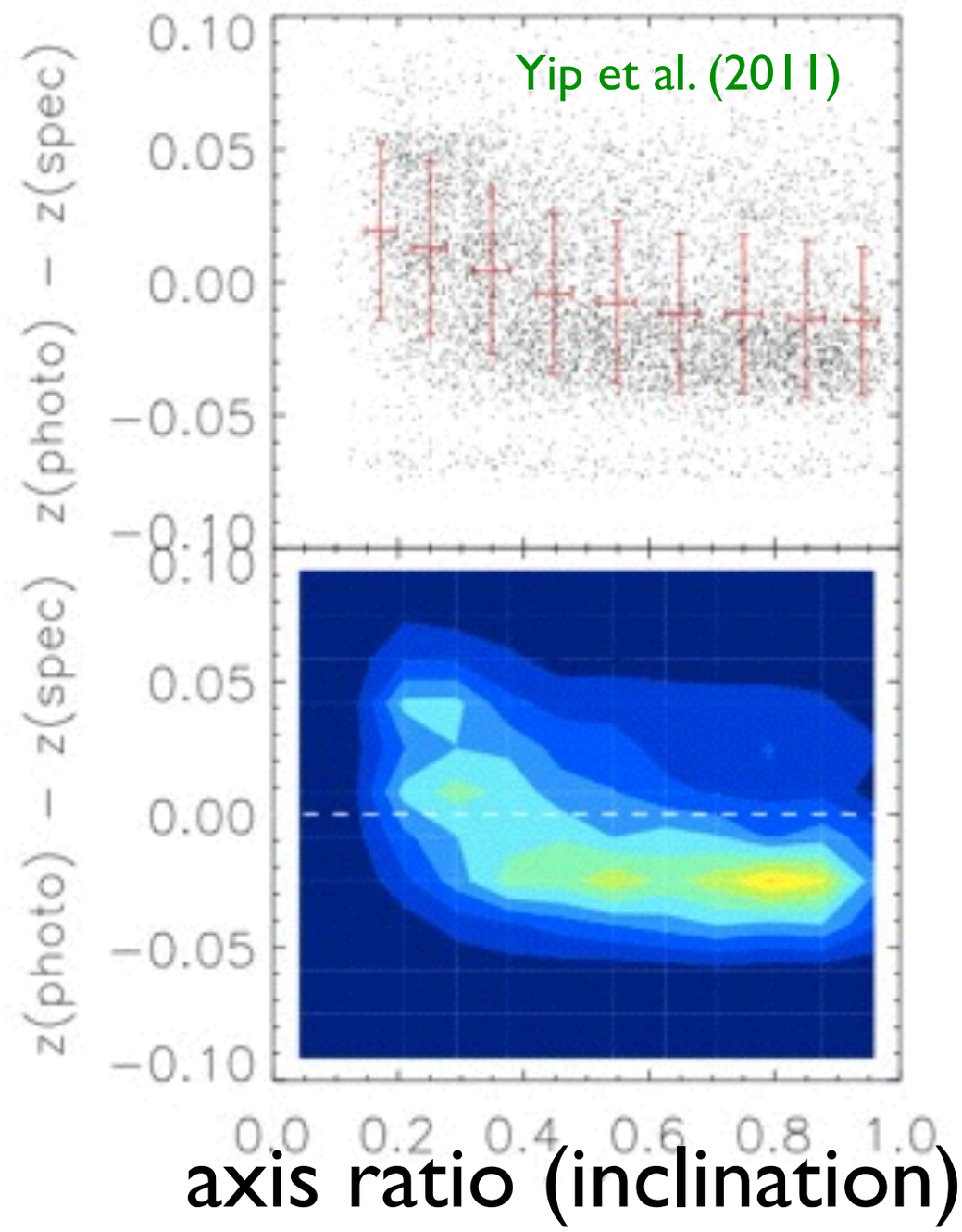


PHAT: PHoto-z Accuracy Testing ★ Hildebrandt et al.

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

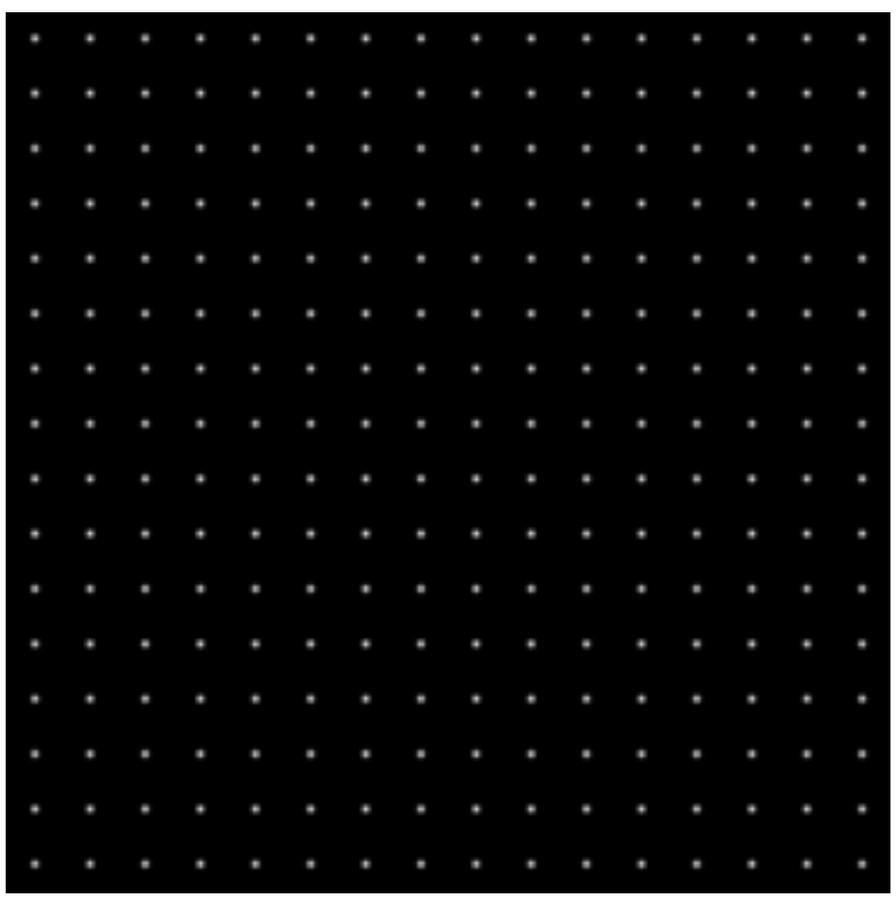


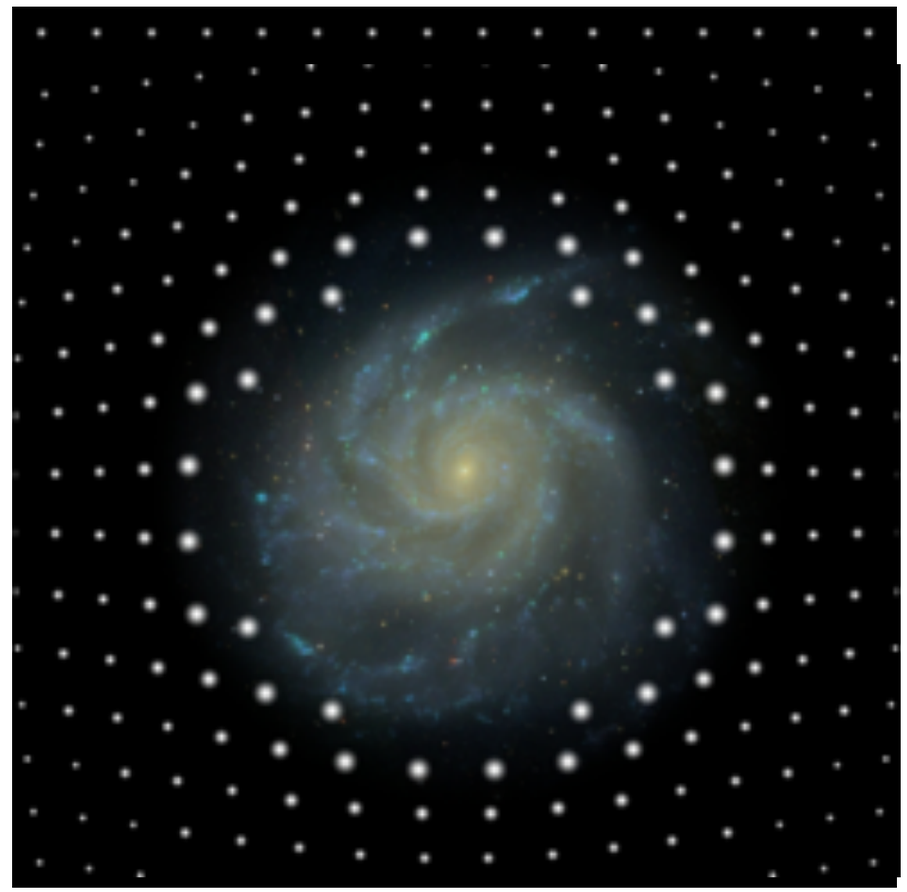


Photometric redshifts

- 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.



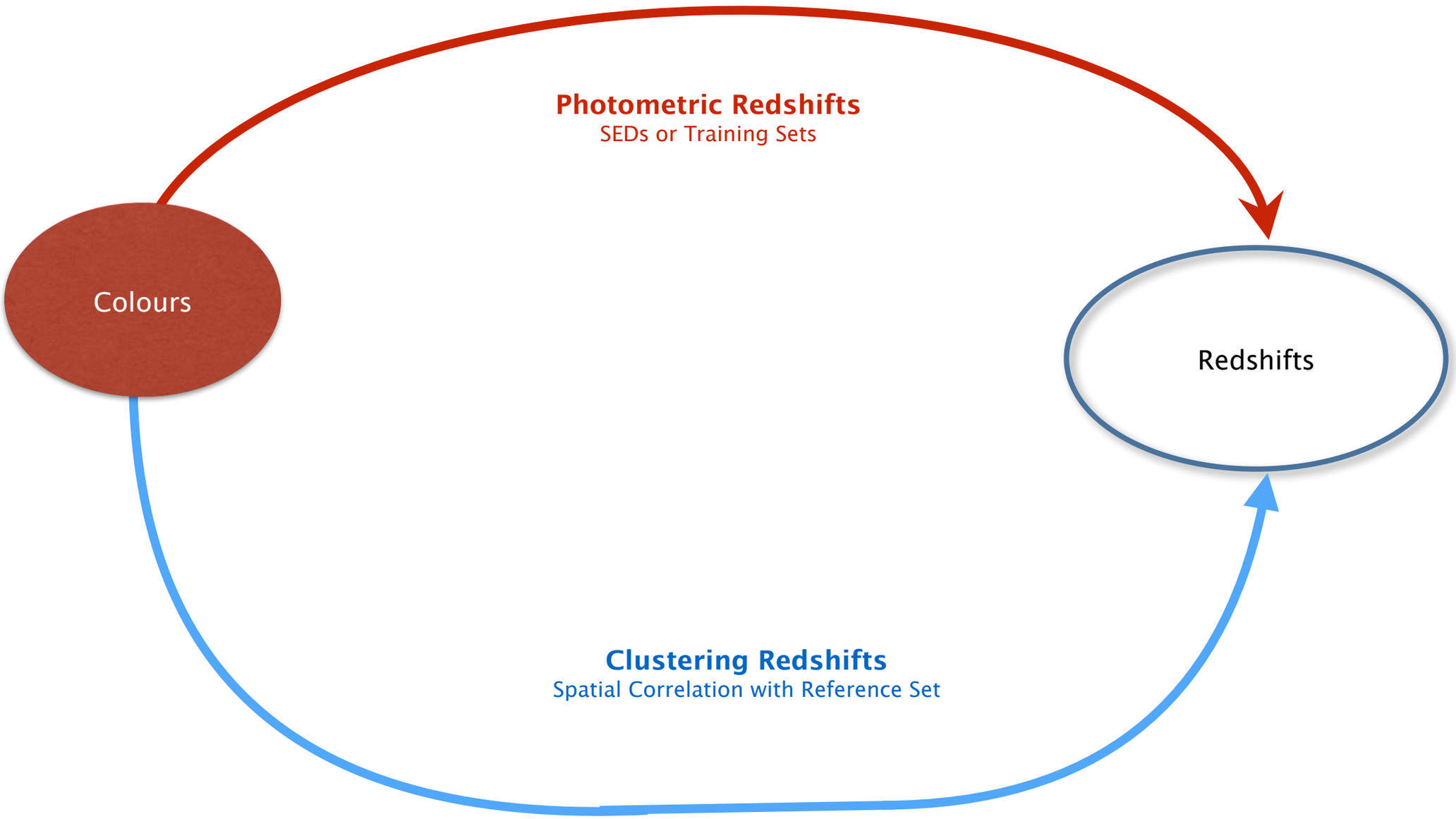


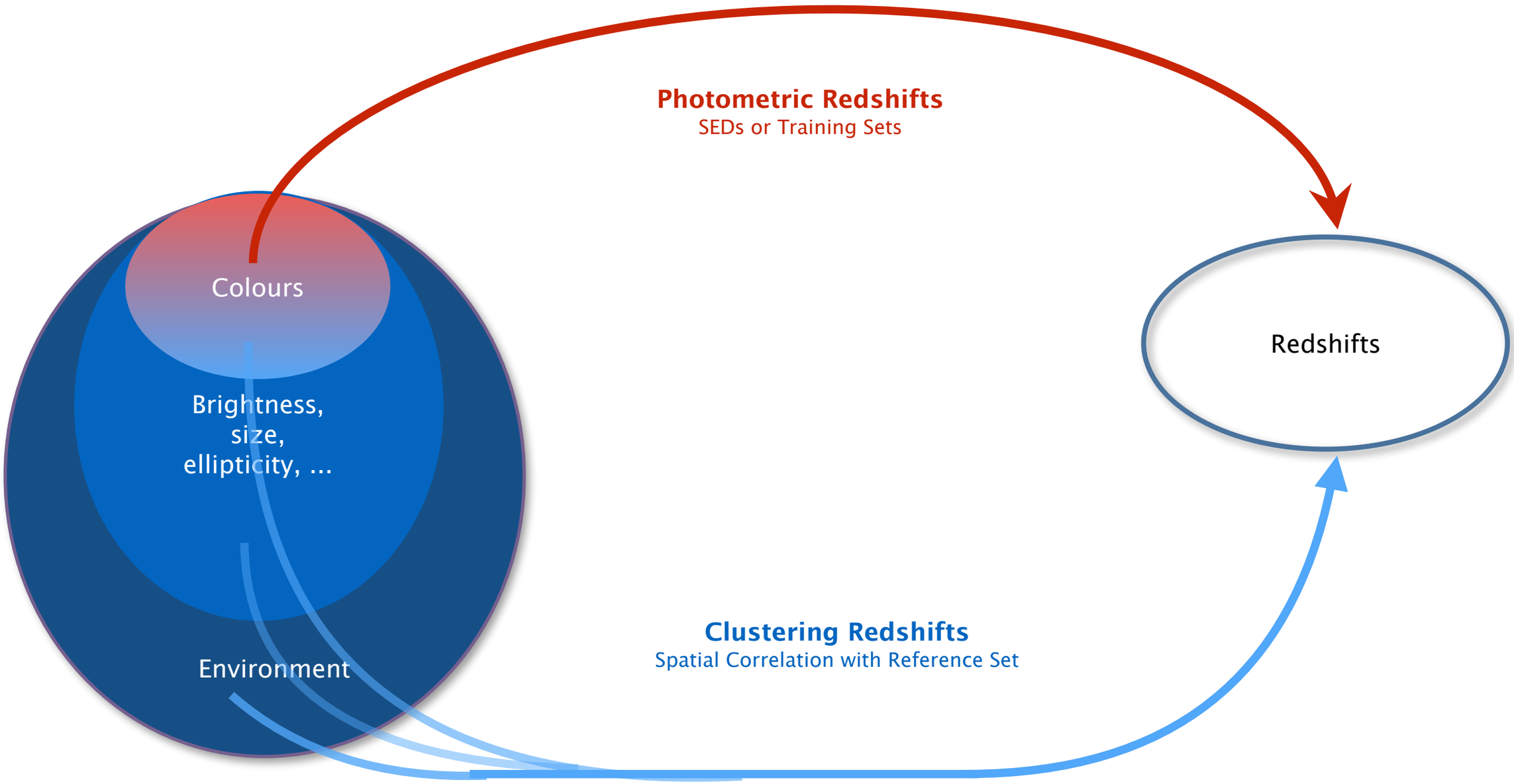
Photometric redshifts

- 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.

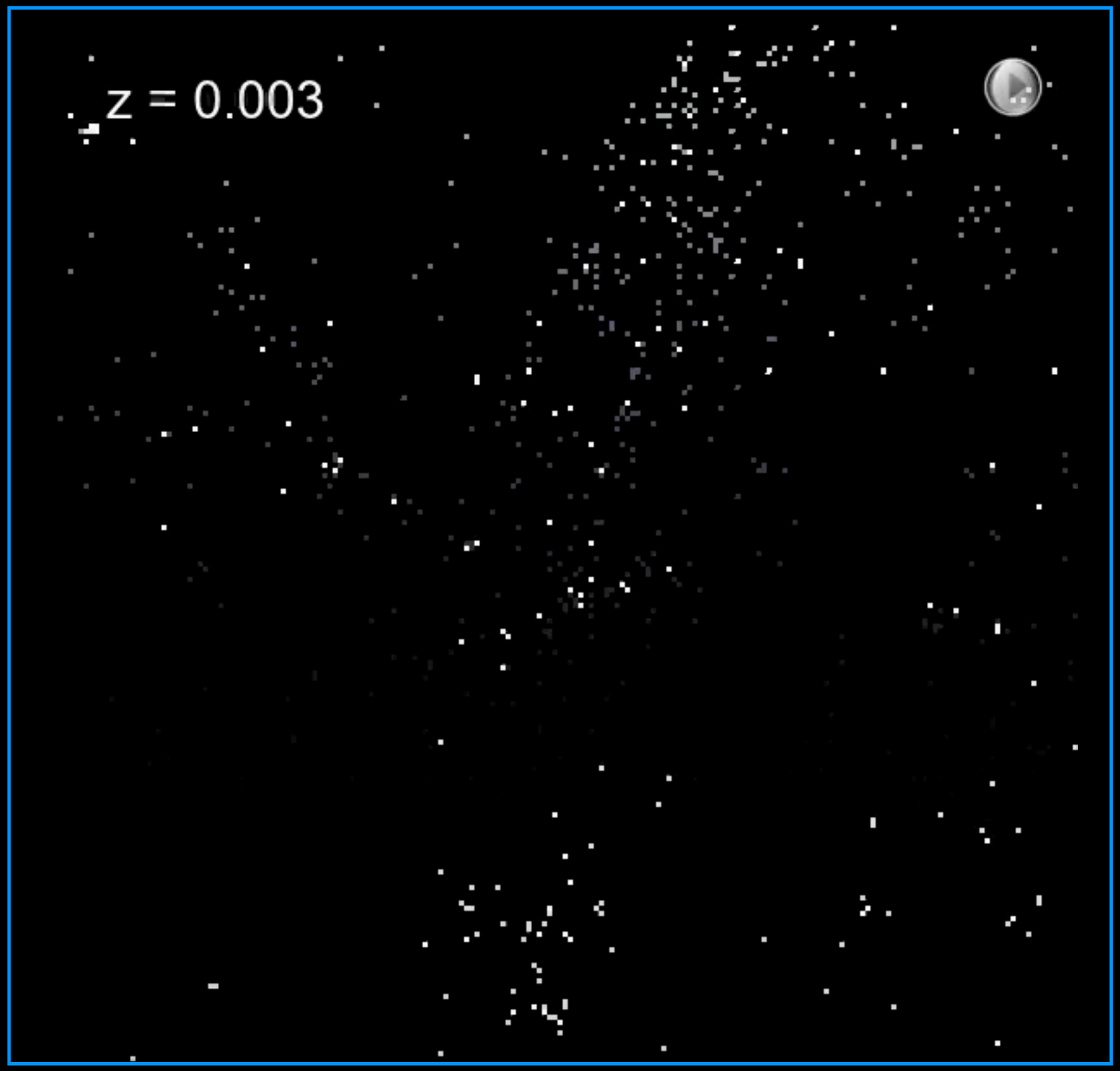
We could completely miss an entire population of galaxies.


Photometric Redshifts
SEDs or Training Sets
Colours
Redshifts
Clustering Redshifts
Spatial Correlation with Reference Set



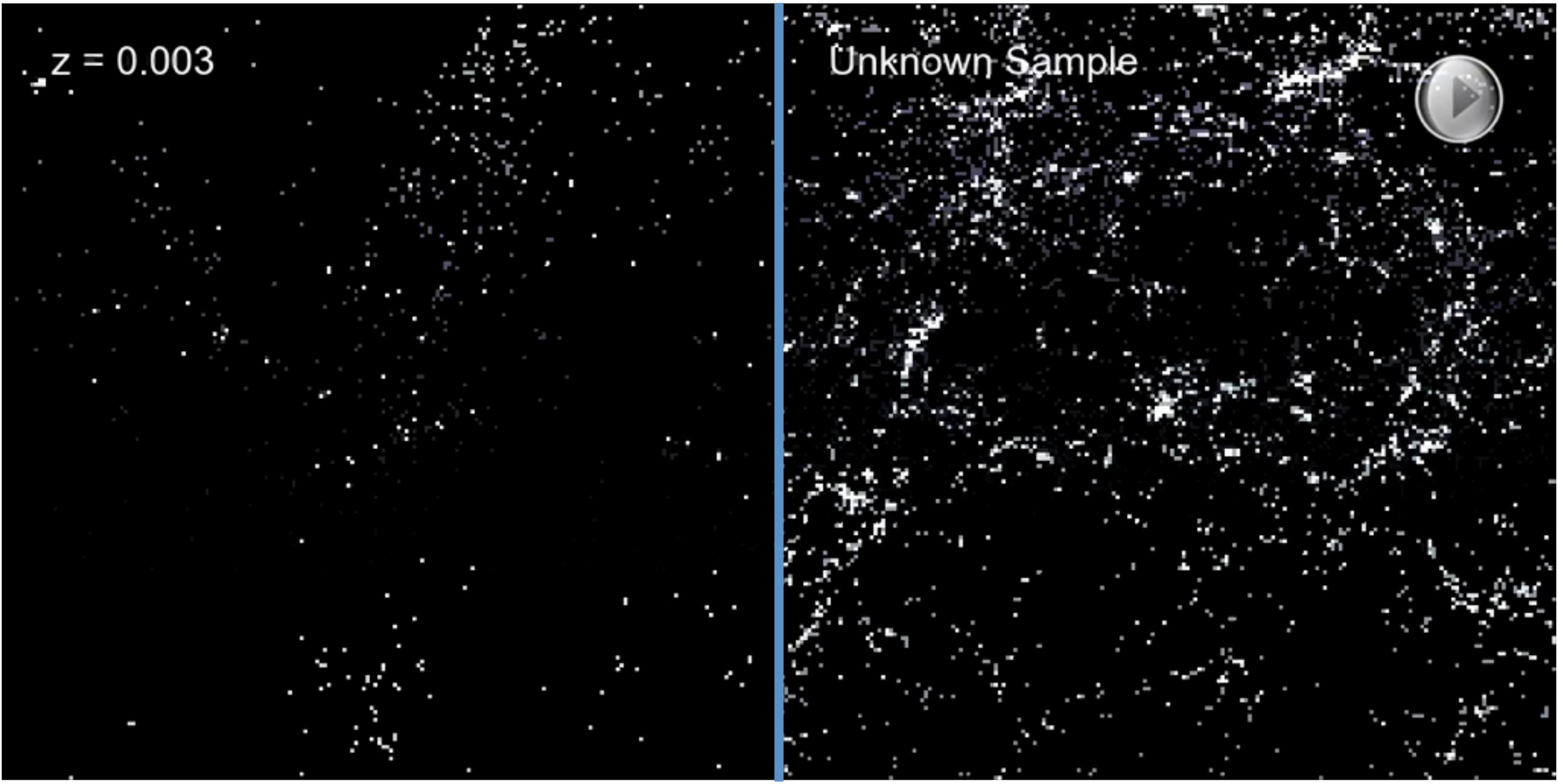
This can be used to infer the redshift pdf of one galaxy

SDSS Spectroscopic Galaxies

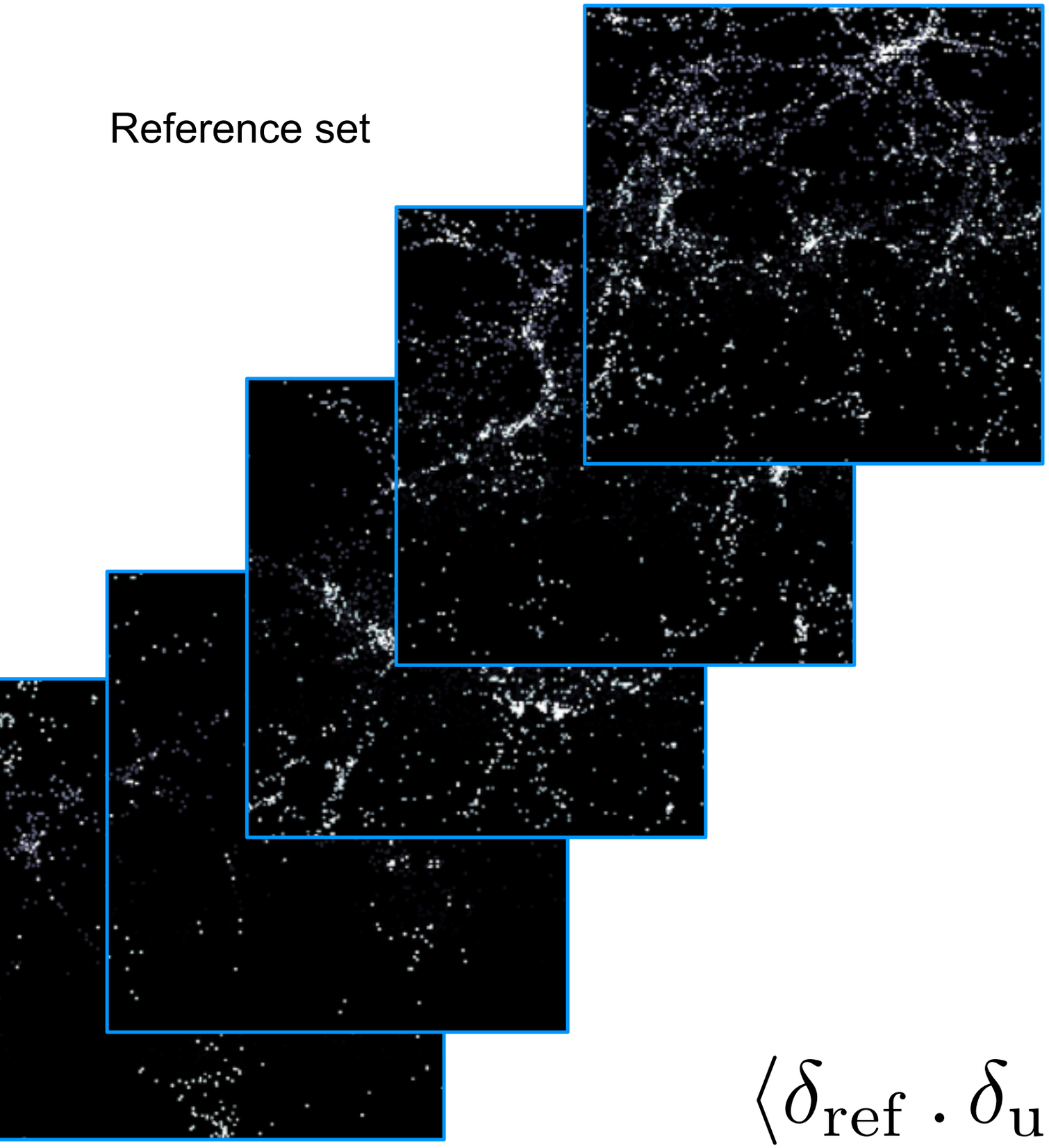


60 degrees

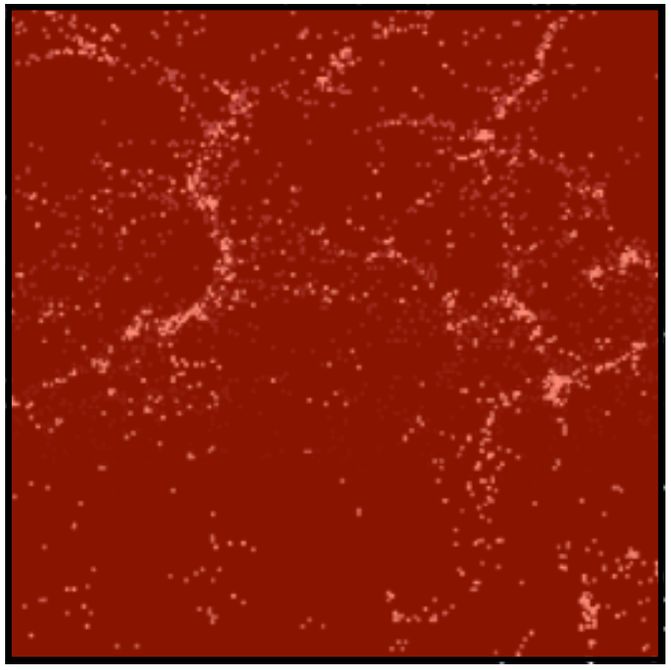
Using Spatial Correlations to Measure Redshift



Reference set



Sample at unknown redshift



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

Metric: 2-point correlation function

The clustering redshift technique

expected data: one "unknown" sample $\{N_u, dN_u/dz, b_u(z)\}$
one reference sample $\{N_r, dN_r/dz, b_r(z)\}$

1. Ideal case: $$\frac{\mathrm{dN_u}}{\mathrm{d}z} = \mathrm{N_u}\,\delta_D(z - z_0)$$

location in redshift space: $$\mathrm{dN_u}/\mathrm{d}z \propto w_{ur}(z_i)$$

normalization: $$\int \mathrm{d}z\,\frac{\mathrm{dN_u}}{\mathrm{d}z} = \mathrm{N_u}$$

The clustering redshift technique

$$\mathrm{dN_u/dz} \propto w_{ur}(z_i) \qquad w_{ur}(\theta, z_i) = \frac{\langle n_u(\theta, z_i)\rangle_r}{n_u} - 1$$

The signal can be obtained by measuring clustering on any scale.
The best strategy is therefore to integrate over all available scales

$$\bar{w}_{ur}(z) = \int_{\theta_{\mathrm{min}}}^{\theta_{\mathrm{max}}} \mathrm{d}\theta\, W(\theta)\, w_{ur}(\theta, z)$$

The clustering redshift technique

2. Departure from the ideal case: $$\frac{\mathrm{dN_u}}{\mathrm{d}z} = \mathrm{N_u}\, G(z_0, \sigma_z)$$

The observable $w_{ur}(z)$ is now a function of:

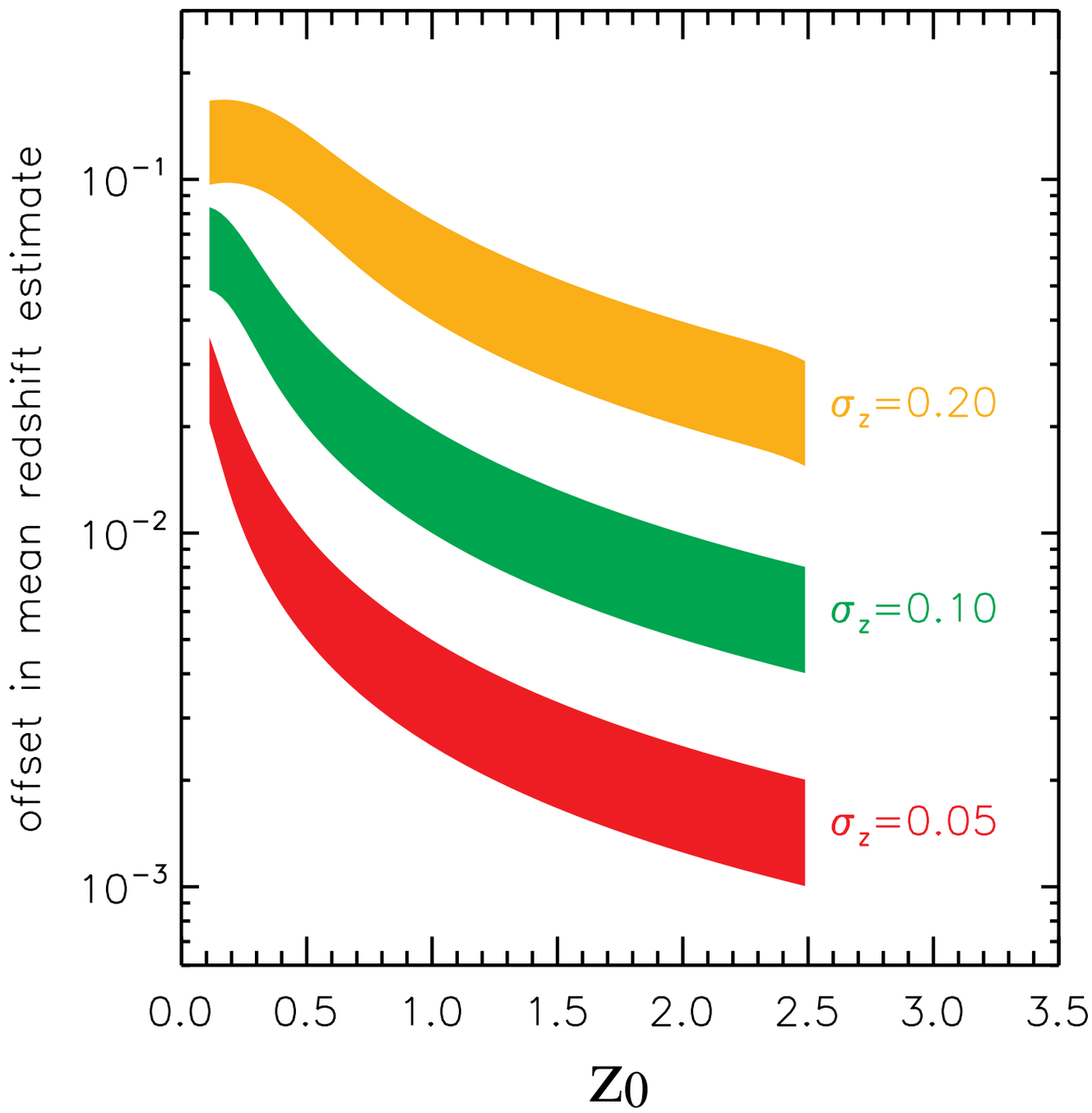
$$\bar{w}_{ur}(z_i) \propto \frac{\mathrm{dN_u}}{\mathrm{d}z}(z_i)\, \bar{b}_u(z_i)\, \bar{b}_r(z_i)\, \bar{w}_{\mathrm{DM}}(z_i)$$

But if the relative change in b(z) is much smaller than in dN/dz

$$\frac{\mathrm{d}\log \mathrm{dN_u}/\mathrm{d}z}{\mathrm{d}z} \gg \frac{\mathrm{d}\log \bar{b}_u}{\mathrm{d}z}$$ then we are in a quasi ideal case

What error are we making if $b_u(z)$ changes?

example: $\bar{b}_u(z) \propto z^{\alpha}$ but we assume $b_u(z)$ = constant



$$\frac{dN_u}{dz} = N_u\, G(z_0, \sigma_z)$$

The clustering redshift technique

Overall strategy:

- Divide the sample into small cells in the space of observables (mag, color, size, etc.)

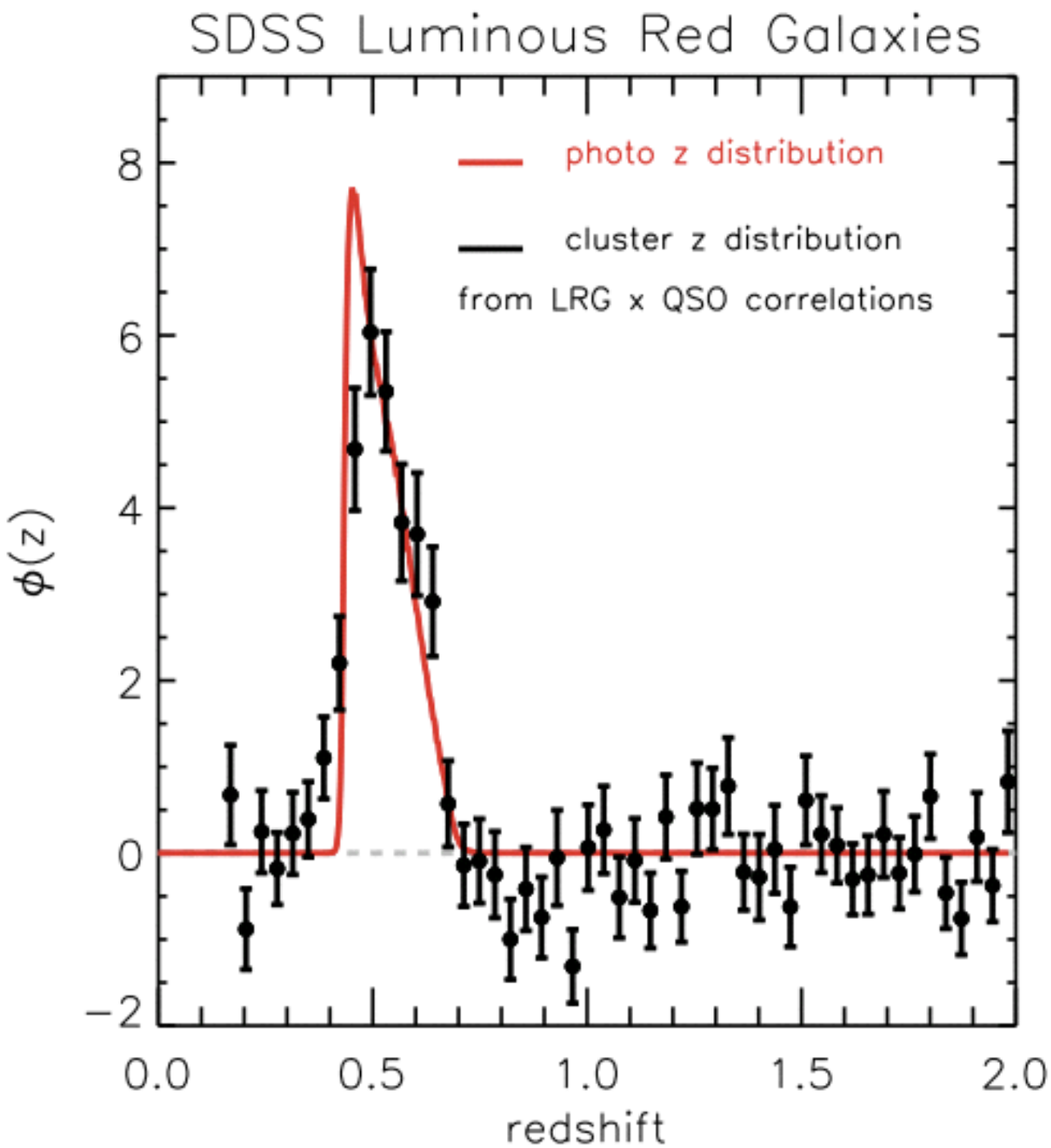
- For each cell, estimate the clustering redshift distribution. The compactness of the distribution can be used to estimate the likely error on dN/dz

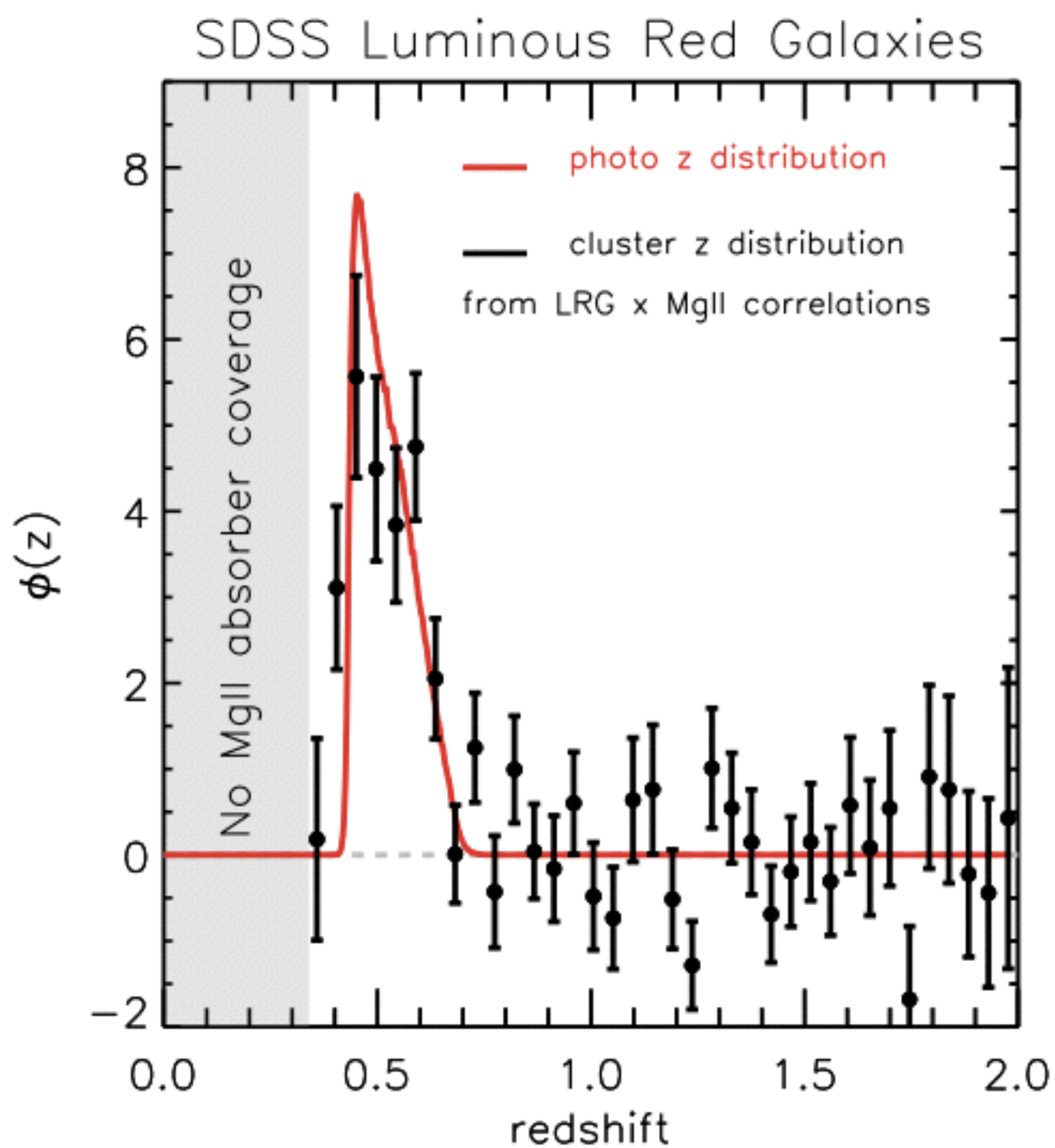
- Combine all (good) cells $$\frac{\mathrm{dN_u}}{\mathrm{d}z} = \sum_j \frac{\mathrm{dN_u}}{\mathrm{d}z}(\mathrm{voxel}\, j)$$

Pilot study with real data

Ménard et al. (2013)

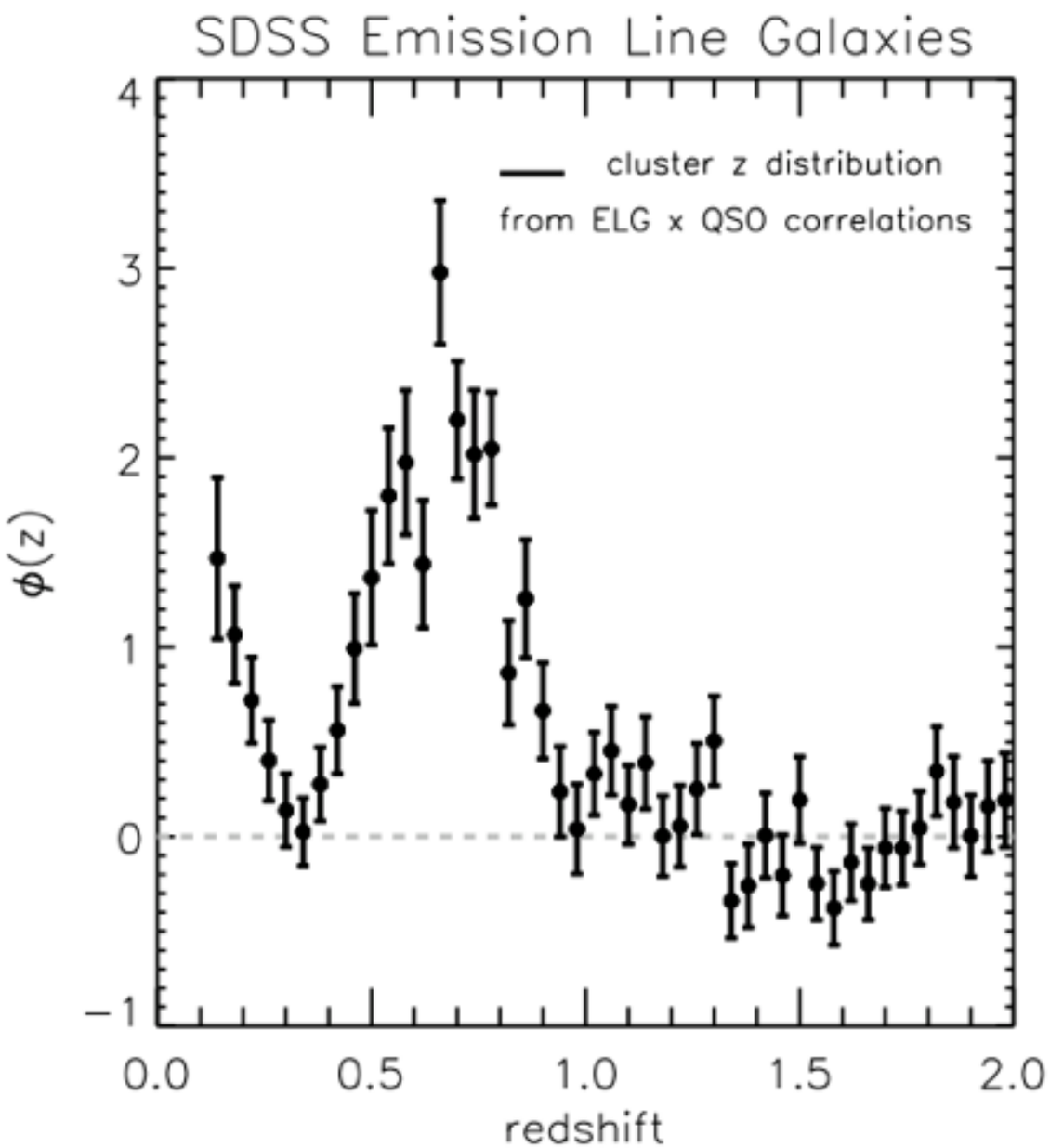
clustering-based redshifts

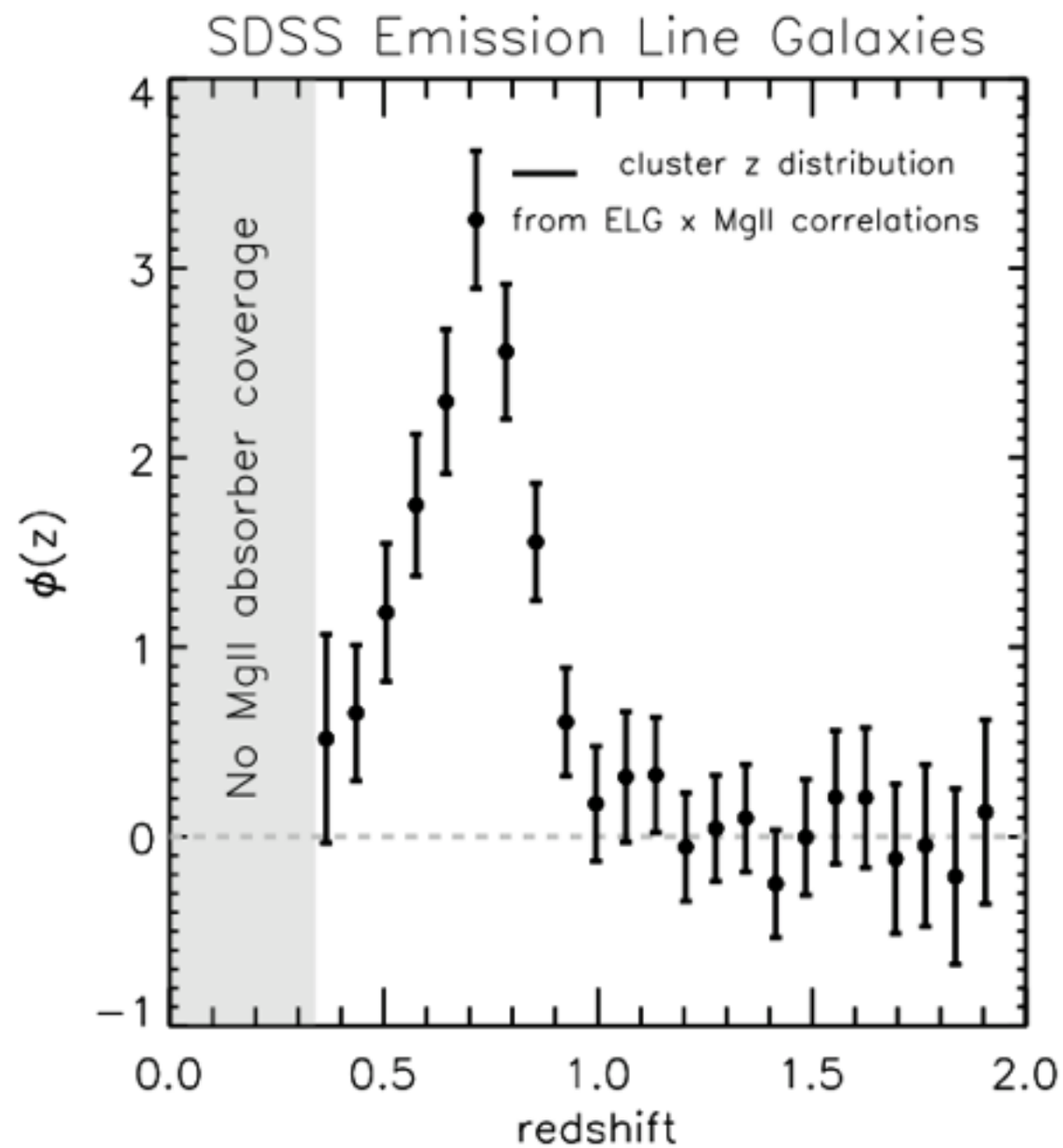




Ménard et al. (2013)

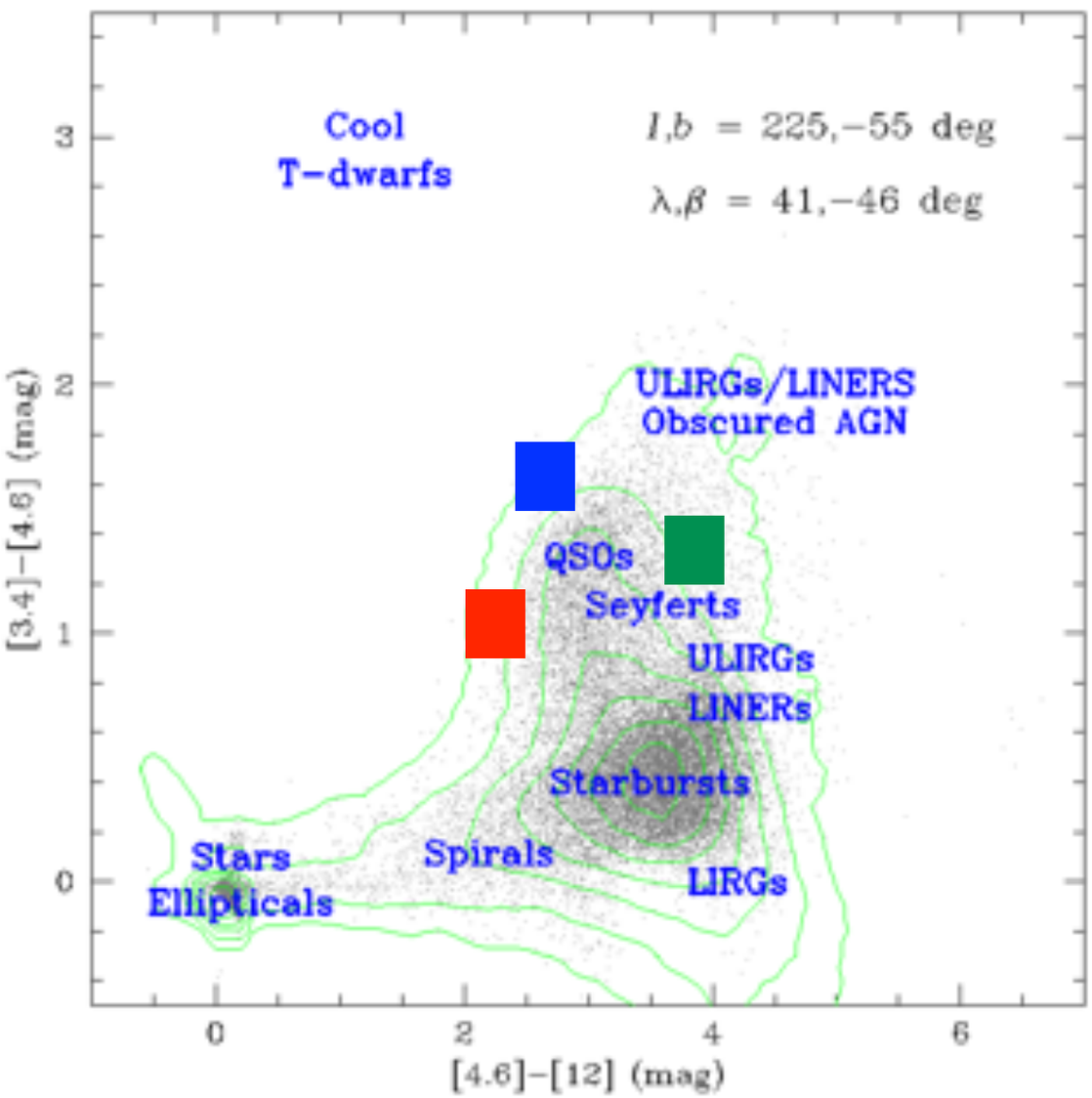
clustering-based redshifts

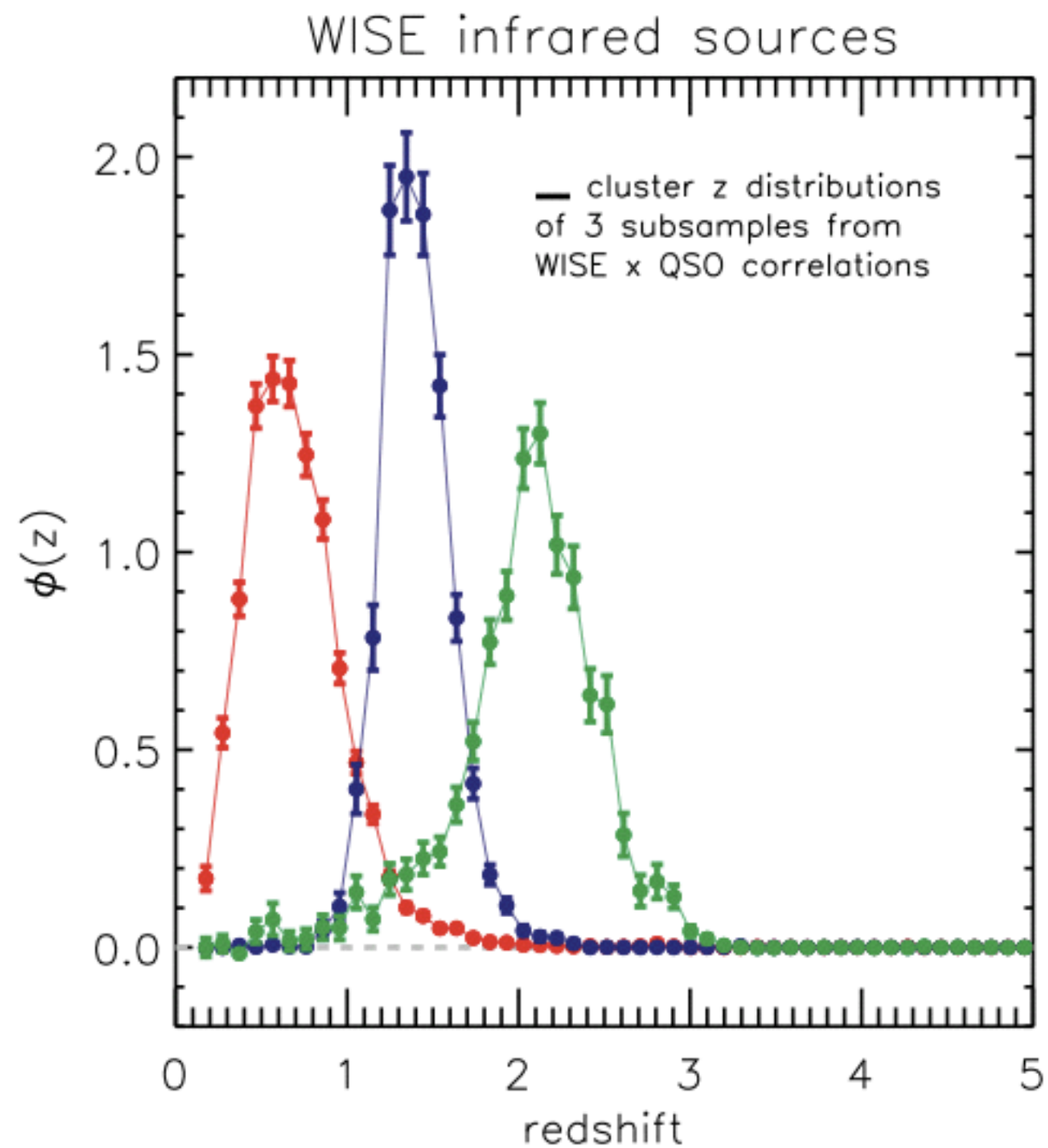




Ménard et al. (2013)

clustering-based redshifts

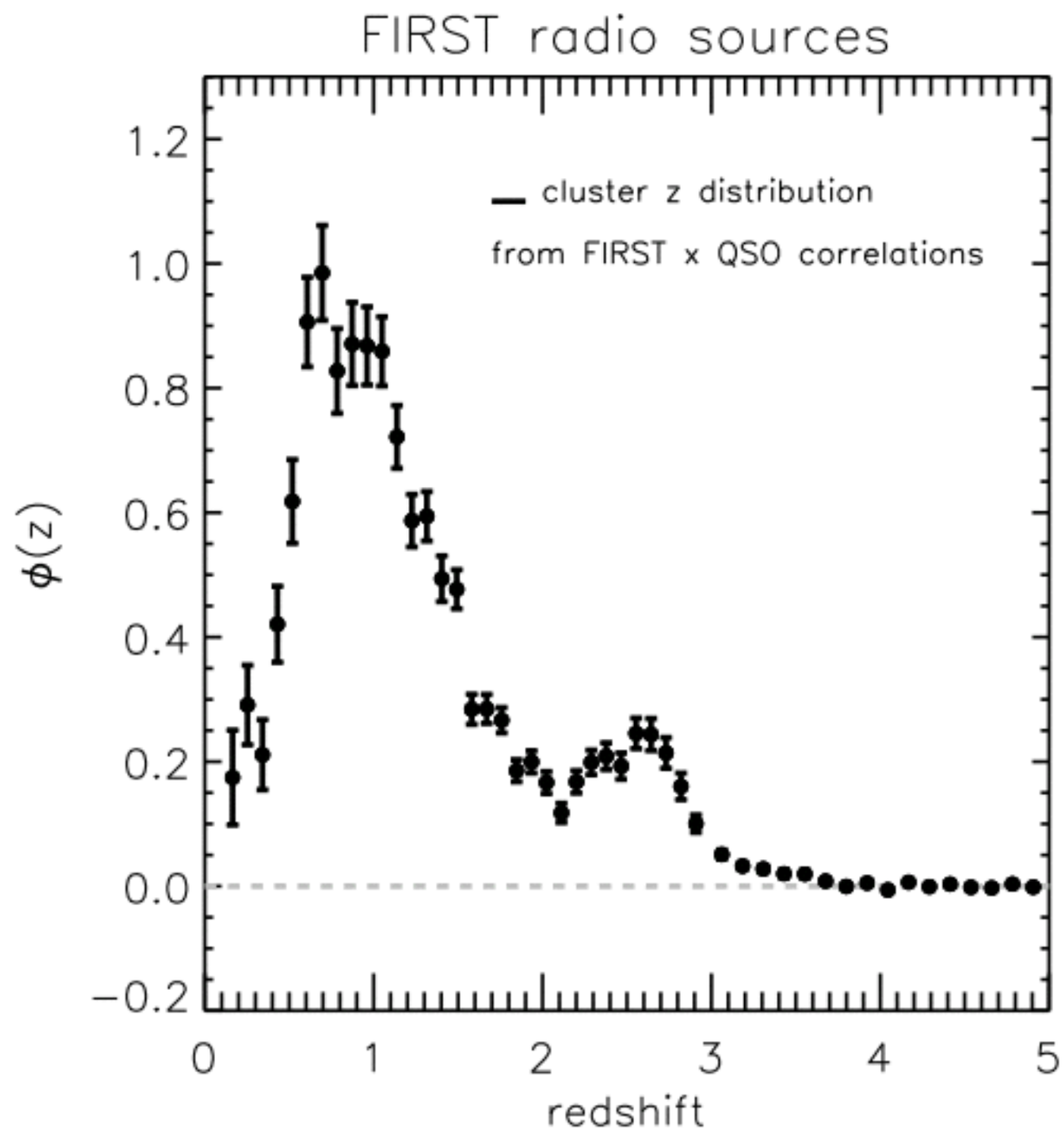




Ménard et al. (2013)

clustering-based redshifts

The FIRST VLA



Ménard et al. (2013)

Photometric redshifts

- rely on SED templates
- require training sets.
- answer not unique.
- affected by dust
- subject to catastrophic failures

- Require an a priori knowledge of the sources.

Clustering redshifts

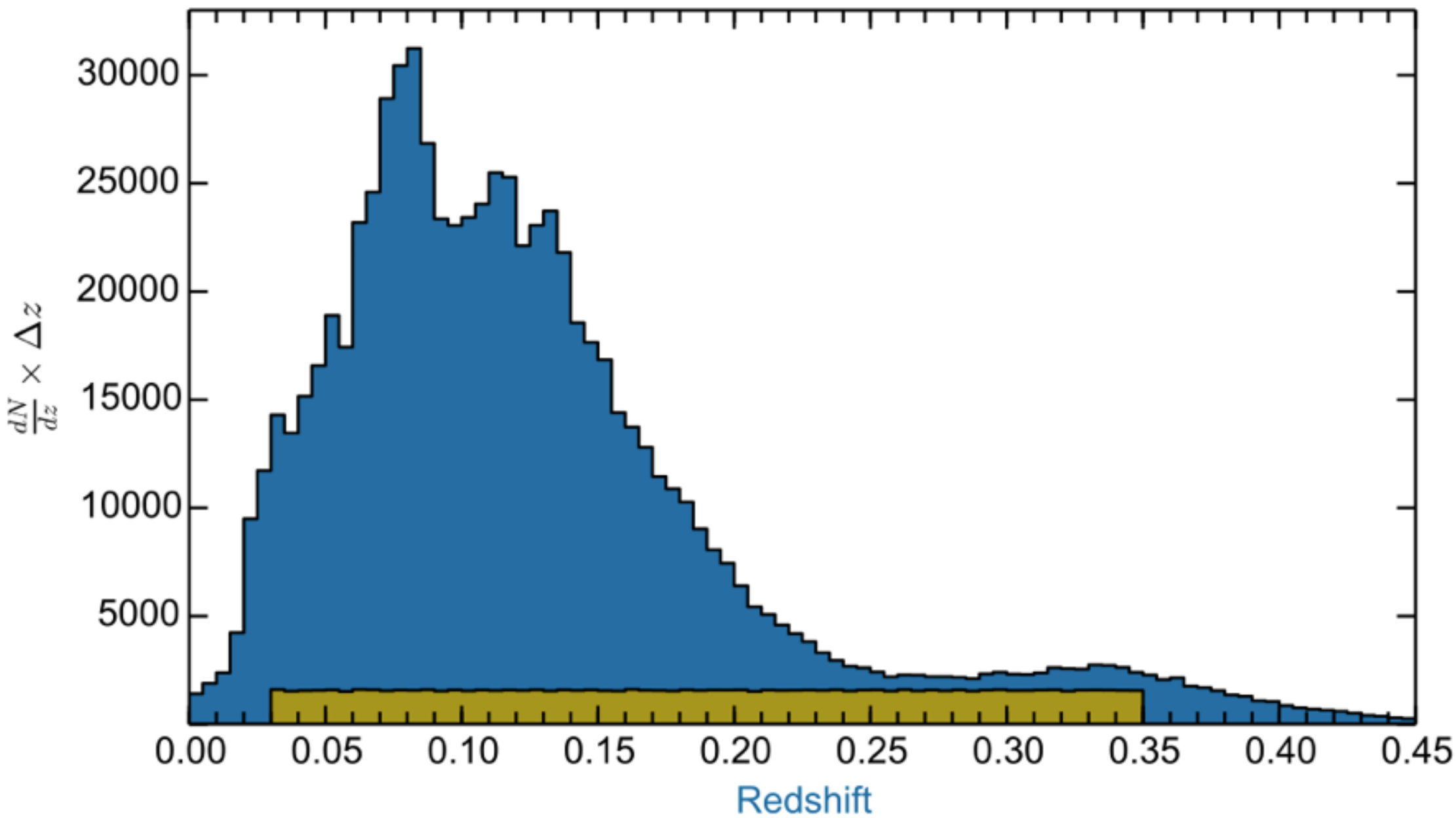
- no SED templates
- no training sets.
- unique answer.
- not affected by dust
- no catastrophic failures

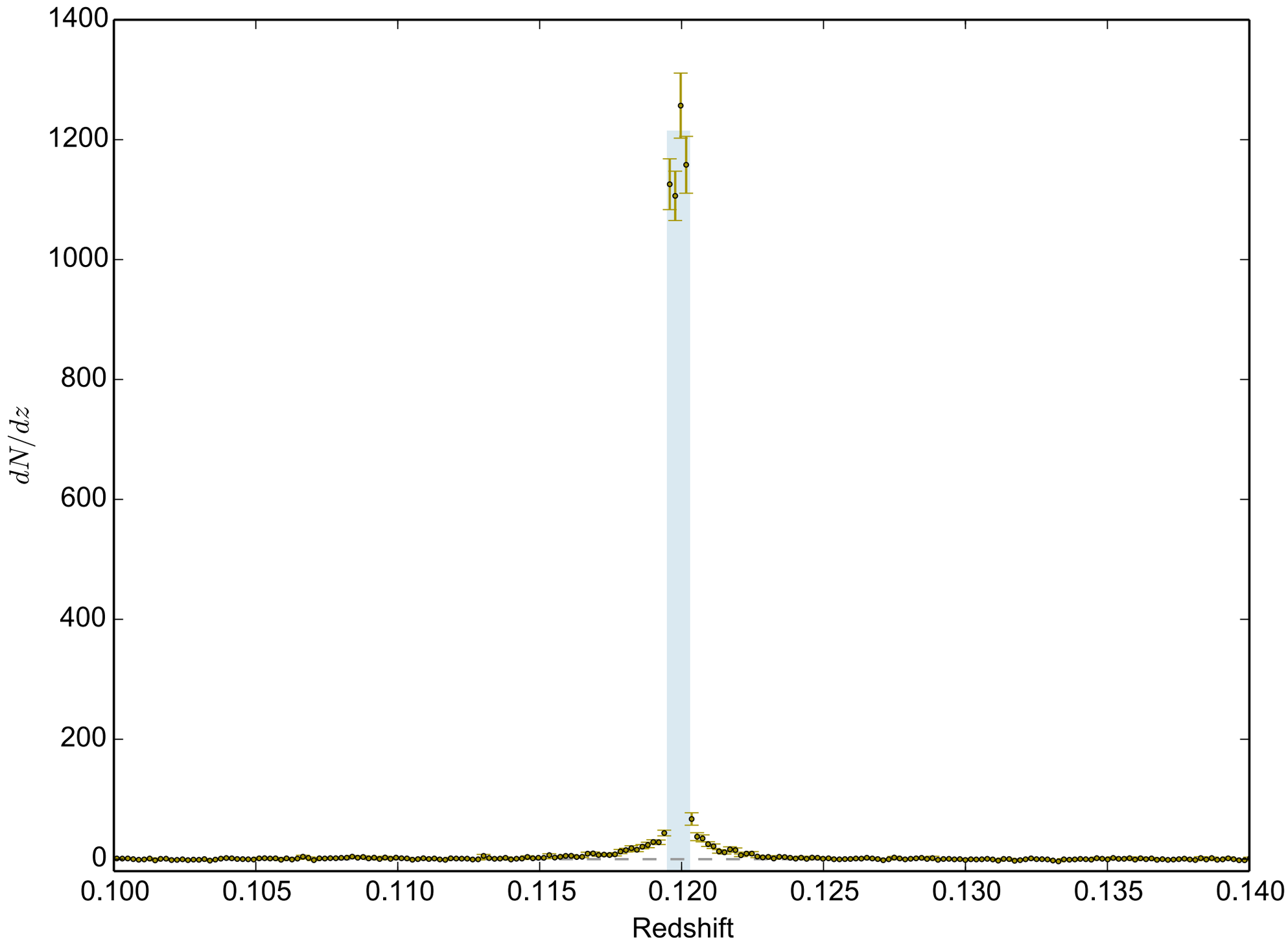
- Can be used blindly without any a priori knowledge of the sources

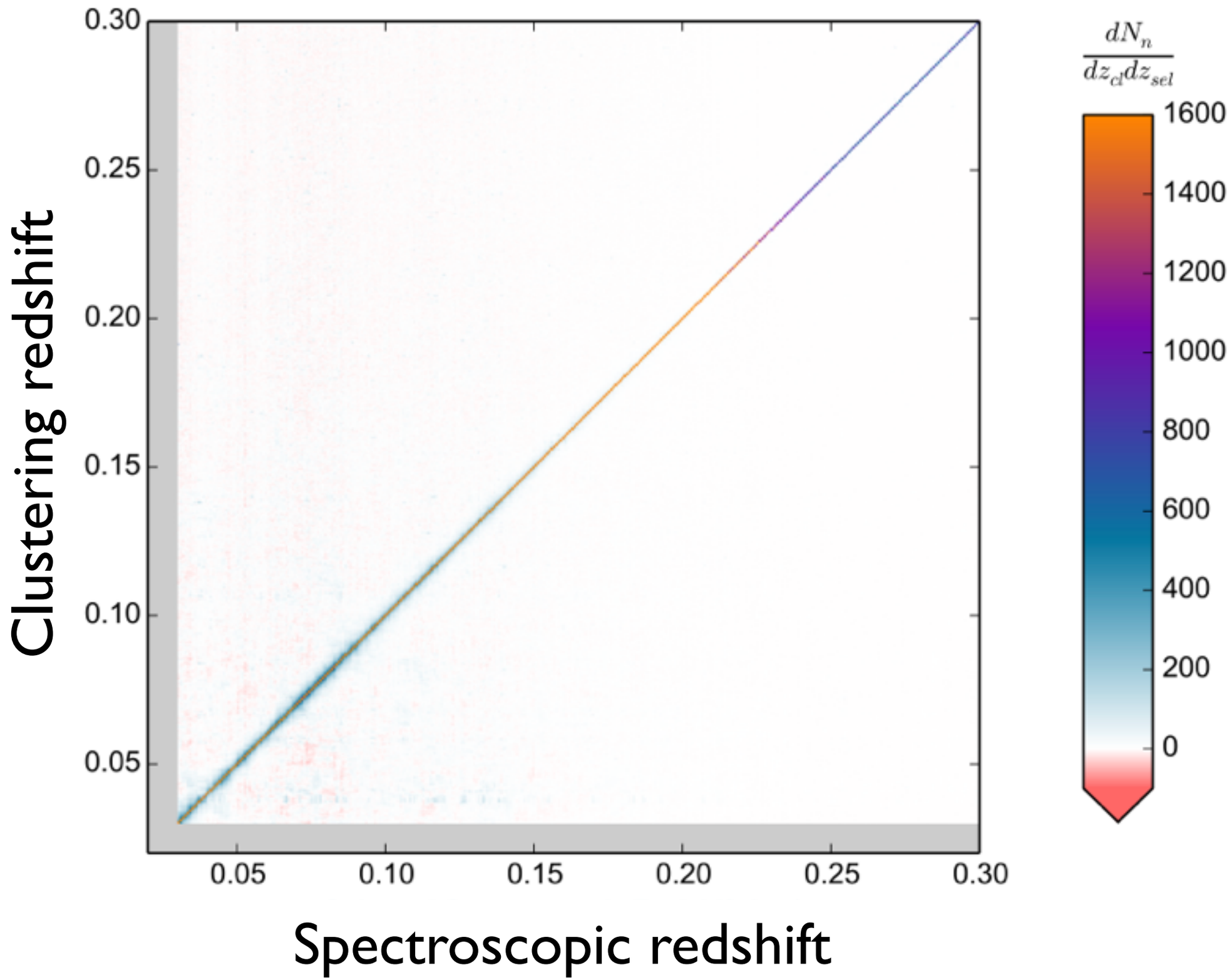
Testing the accuracy of clustering redshifts

Application to
SDSS spectroscopic galaxies

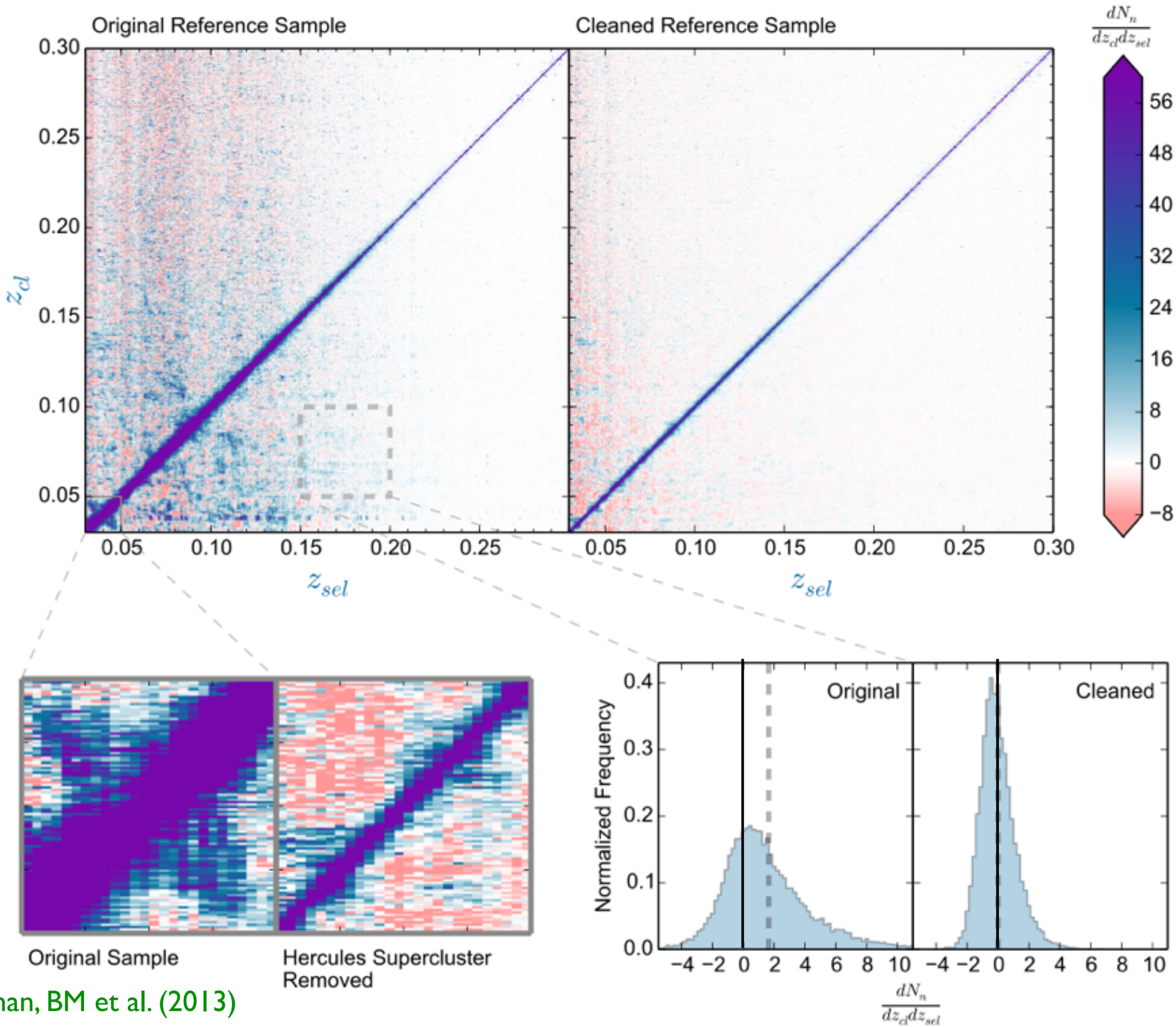
1 million spectroscopic galaxies




1400
1200
1000
800
600
400
200
0
dN/dz
0.100
0.105
0.110
0.115
0.120
0.125
0.130
0.135
0.140
Redshift



Rahman, BM et al. (2013)



Rahman, BM et al. (2013)

Application to real data

We select 3 galaxy samples from the SDSS

magnitude limited: $r<17.7$

sample 1: $0.5 < g-r < 0.6$

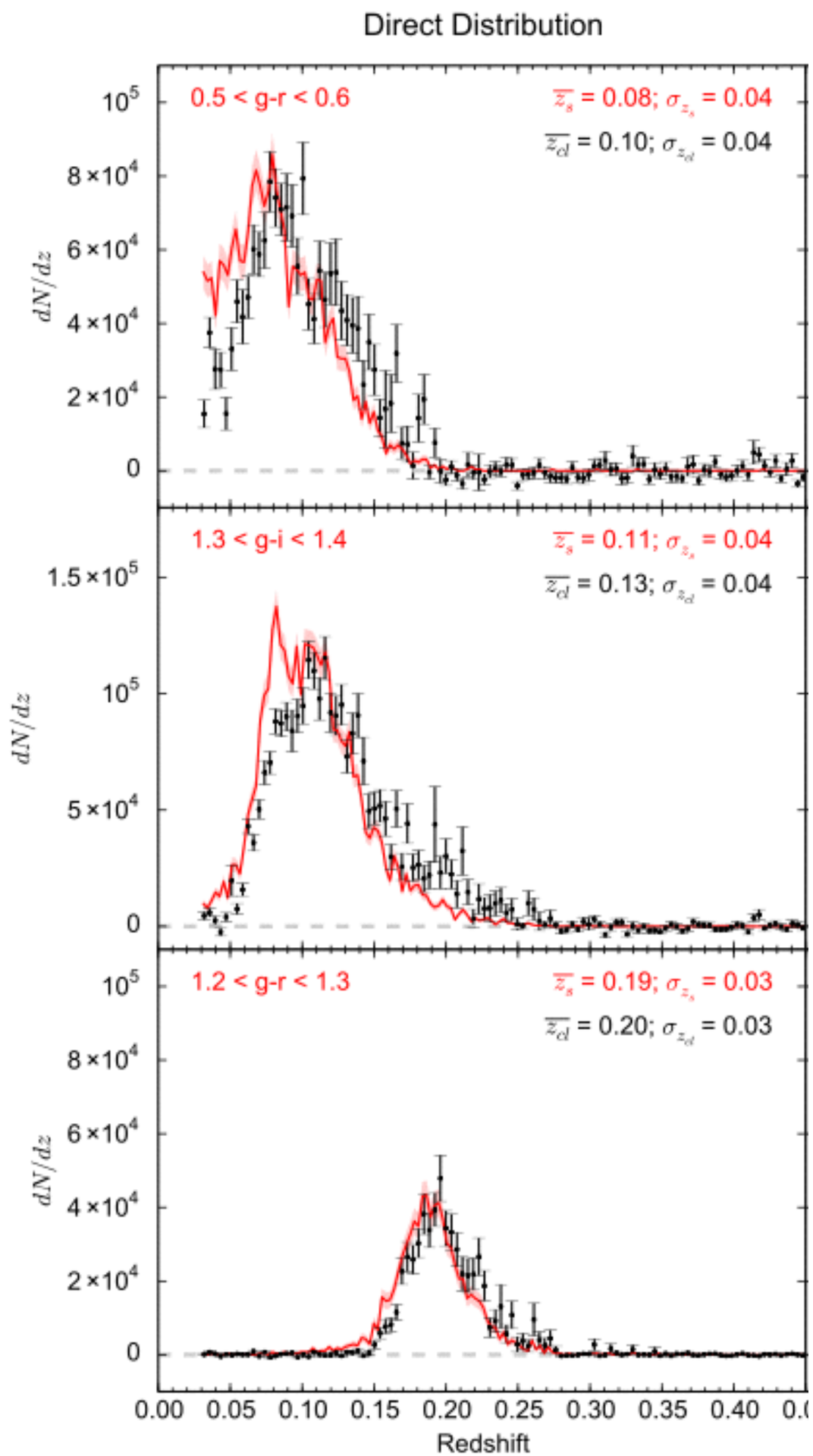
~ 6,300 galaxies

sample 2: $1.3 < g-i < 1.4$

~ 10,000 galaxies

sample 3: $1.2 < g-r < 1.3$

~ 2,500 galaxies



sample 1: 0.5 < g-r < 0.6

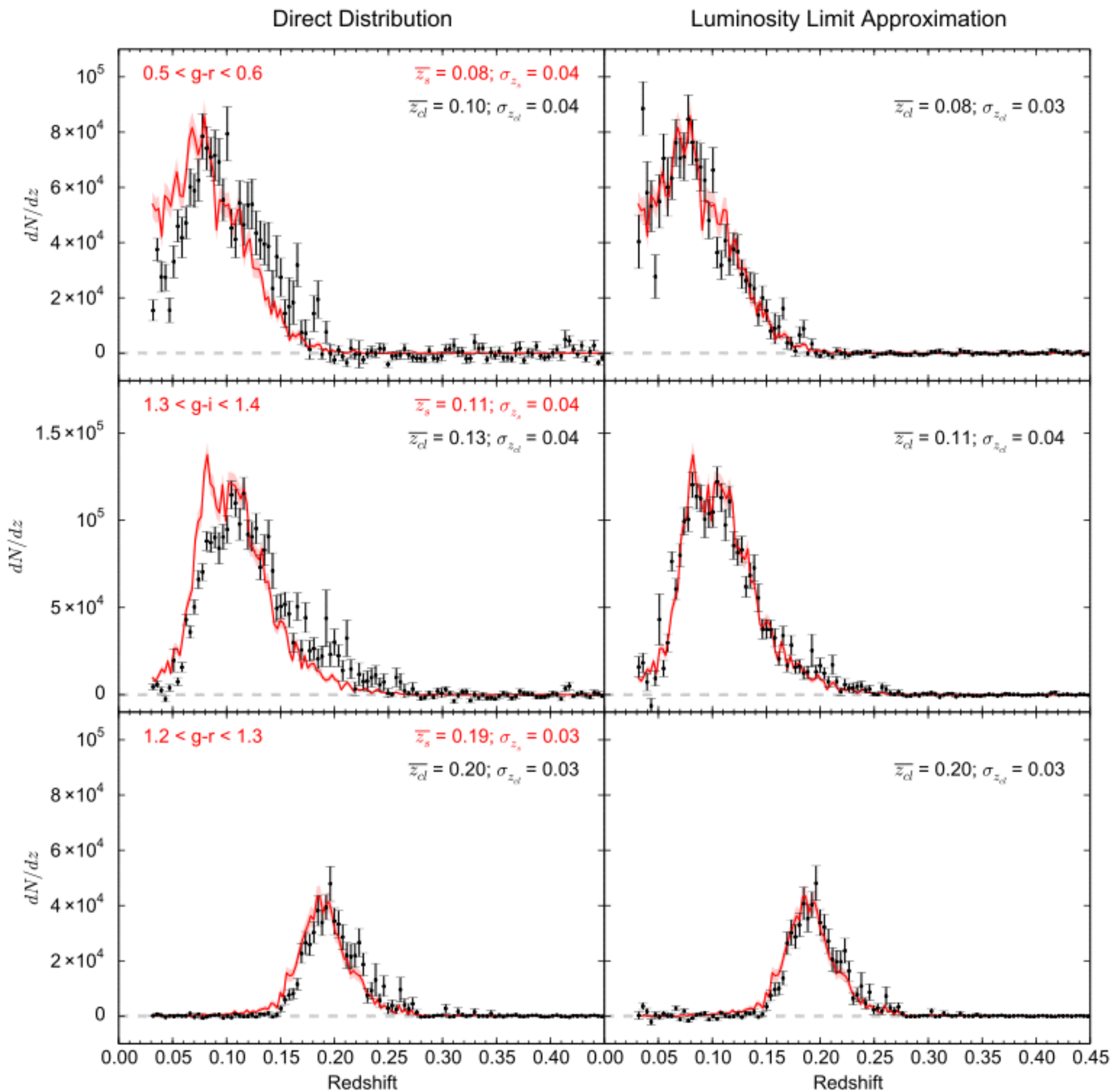
~ 6,300 galaxies

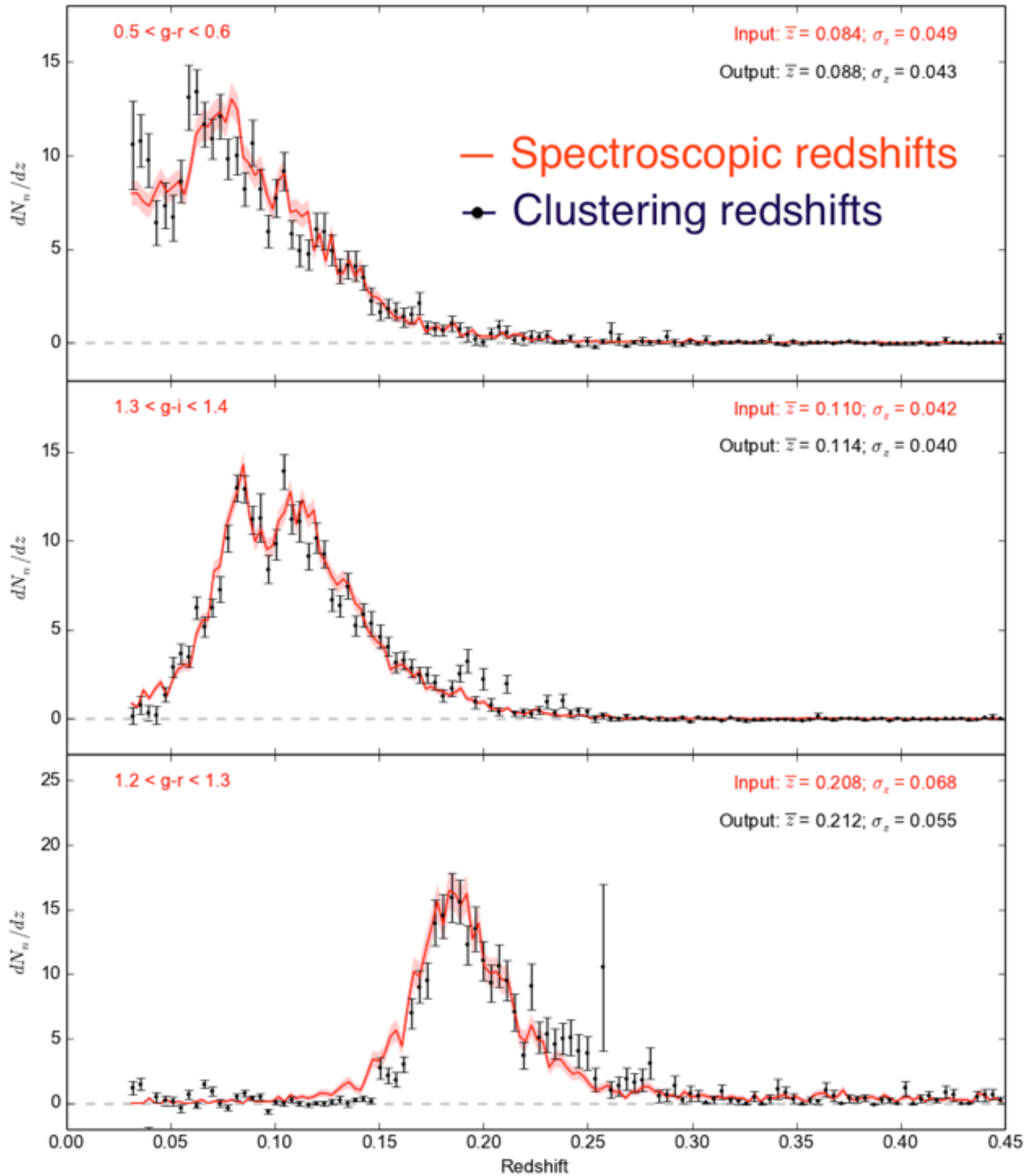
sample 2: 1.3 < g-i < 1.4

~ 10,000 galaxies

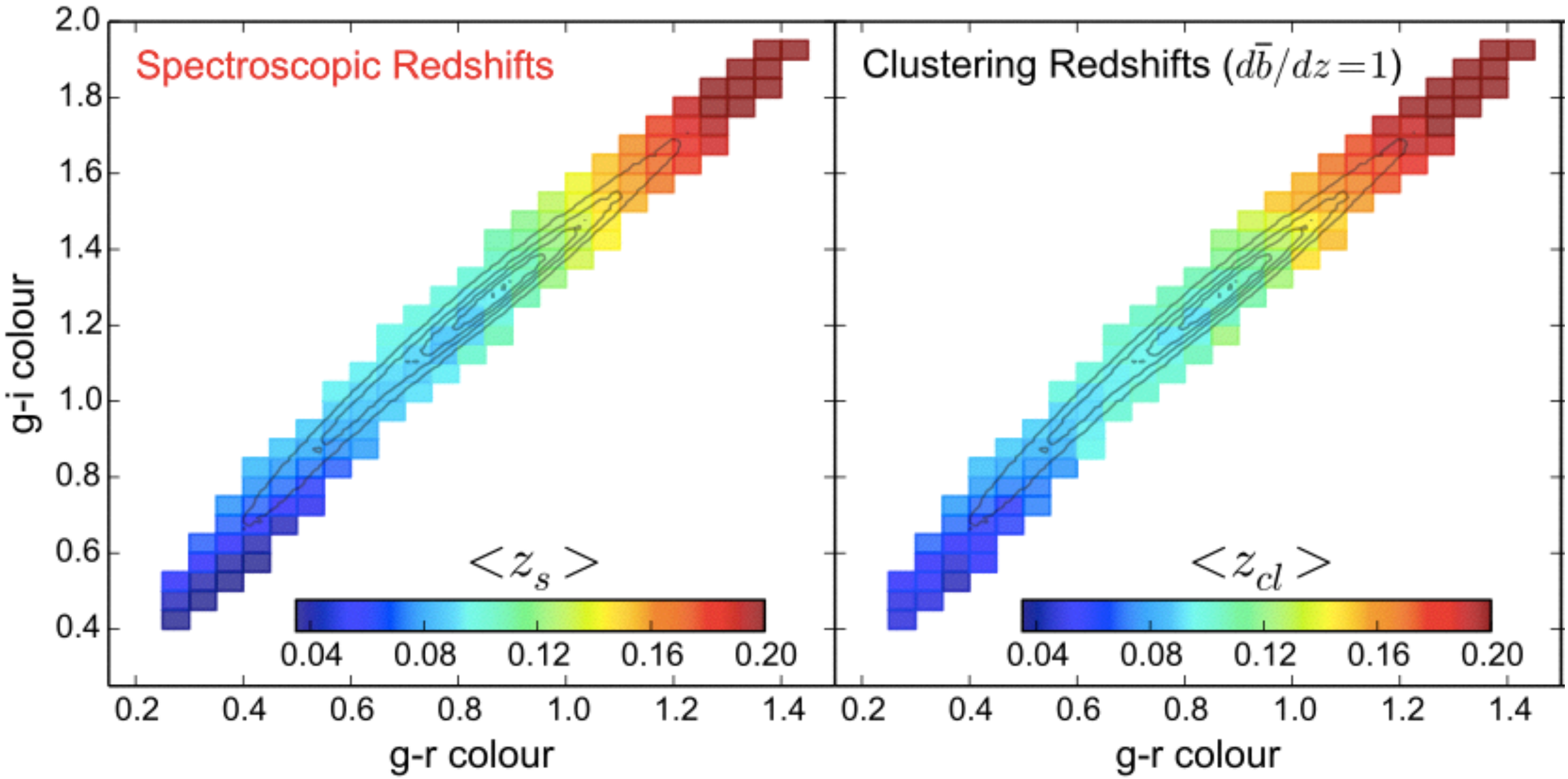
sample 3: 1.2 < g-r < 1.3

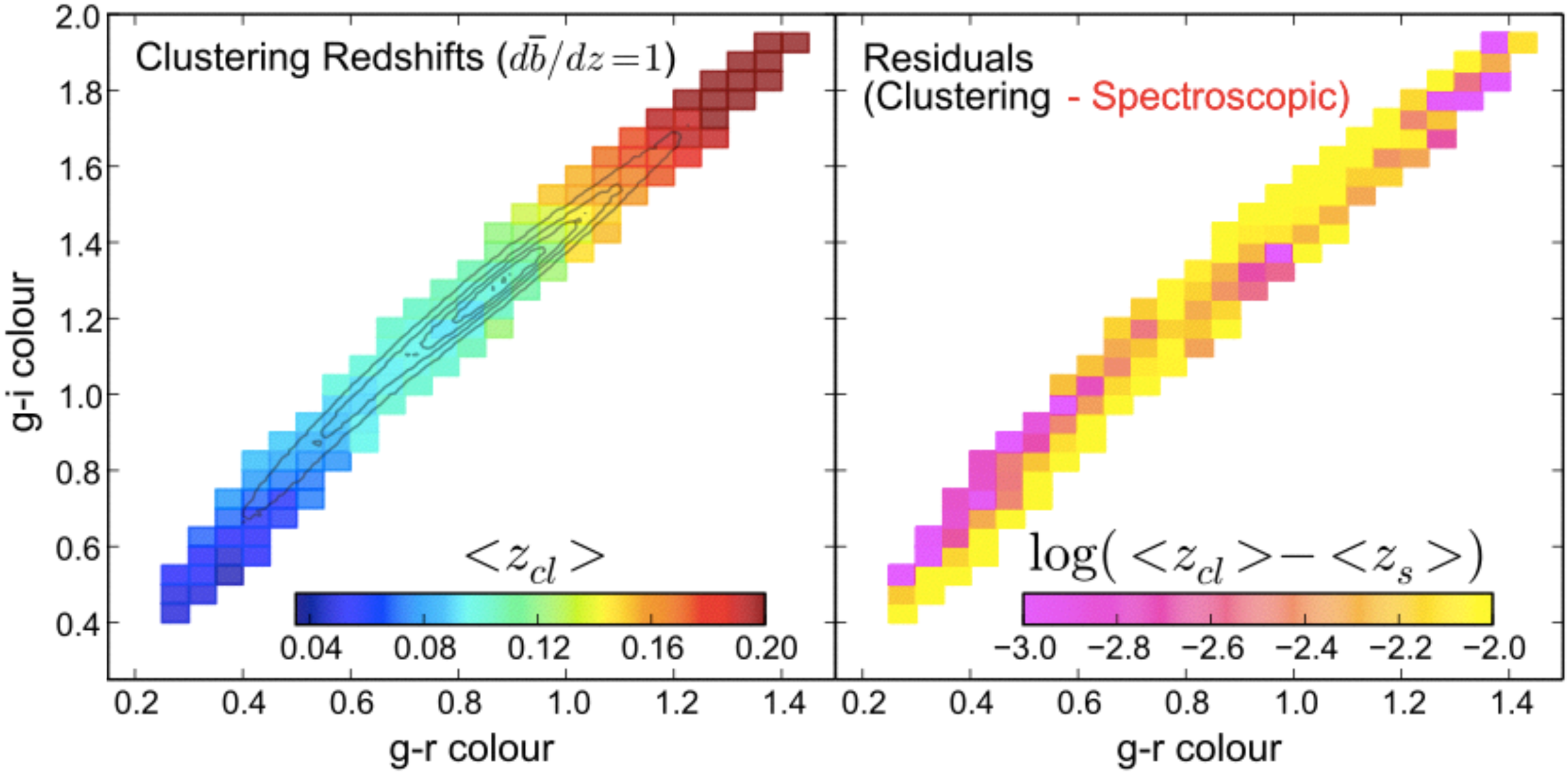
~ 2,500 galaxies


Direct Distribution
Luminosity Limit Approximation
0.5 < g-r < 0.6
$\overline{z_s}$ = 0.08; σ_{z_s} = 0.04
$\overline{z_d}$ = 0.10; σ_{z_d} = 0.04
$\overline{z_d}$ = 0.08; σ_{z_d} = 0.03
1.3 < g-i < 1.4
$\overline{z_s}$ = 0.11; σ_{z_s} = 0.04
$\overline{z_d}$ = 0.13; σ_{z_d} = 0.04
$\overline{z_d}$ = 0.11; σ_{z_d} = 0.04
1.2 < g-r < 1.3
$\overline{z_s}$ = 0.19; σ_{z_s} = 0.03
$\overline{z_d}$ = 0.20; σ_{z_d} = 0.03
$\overline{z_d}$ = 0.20; σ_{z_d} = 0.03
dN/dz
Redshift
Redshift

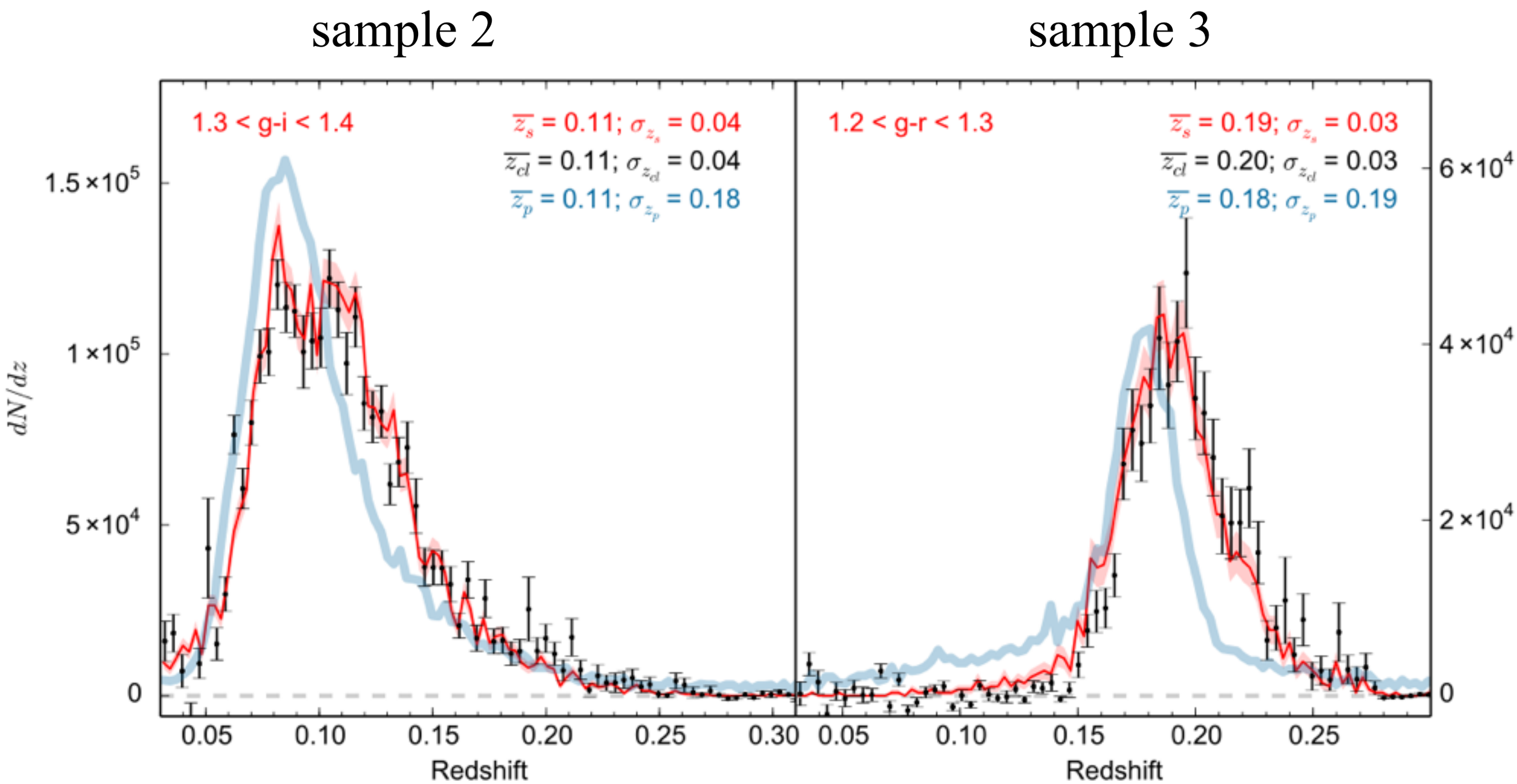


Rahman et al. 2014


Spectroscopic Redshifts
Clustering Redshifts ($d\bar{b}/dz = 1$)
g-i colour
g-r colour
$< z_s >$
$< z_{cl} >$
0.04
0.08
0.12
0.16
0.20
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0


Clustering Redshifts ($d\bar{b}/dz=1$)
Residuals
(Clustering - Spectroscopic)
g-i colour
g-r colour
g-r colour
$<z_{cl}>$
0.04 0.08 0.12 0.16 0.20
$\log(<z_{cl}>-<z_s>)$
−3.0 −2.8 −2.6 −2.4 −2.2 −2.0
0.2 0.4 0.6 0.8 1.0 1.2 1.4
0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0

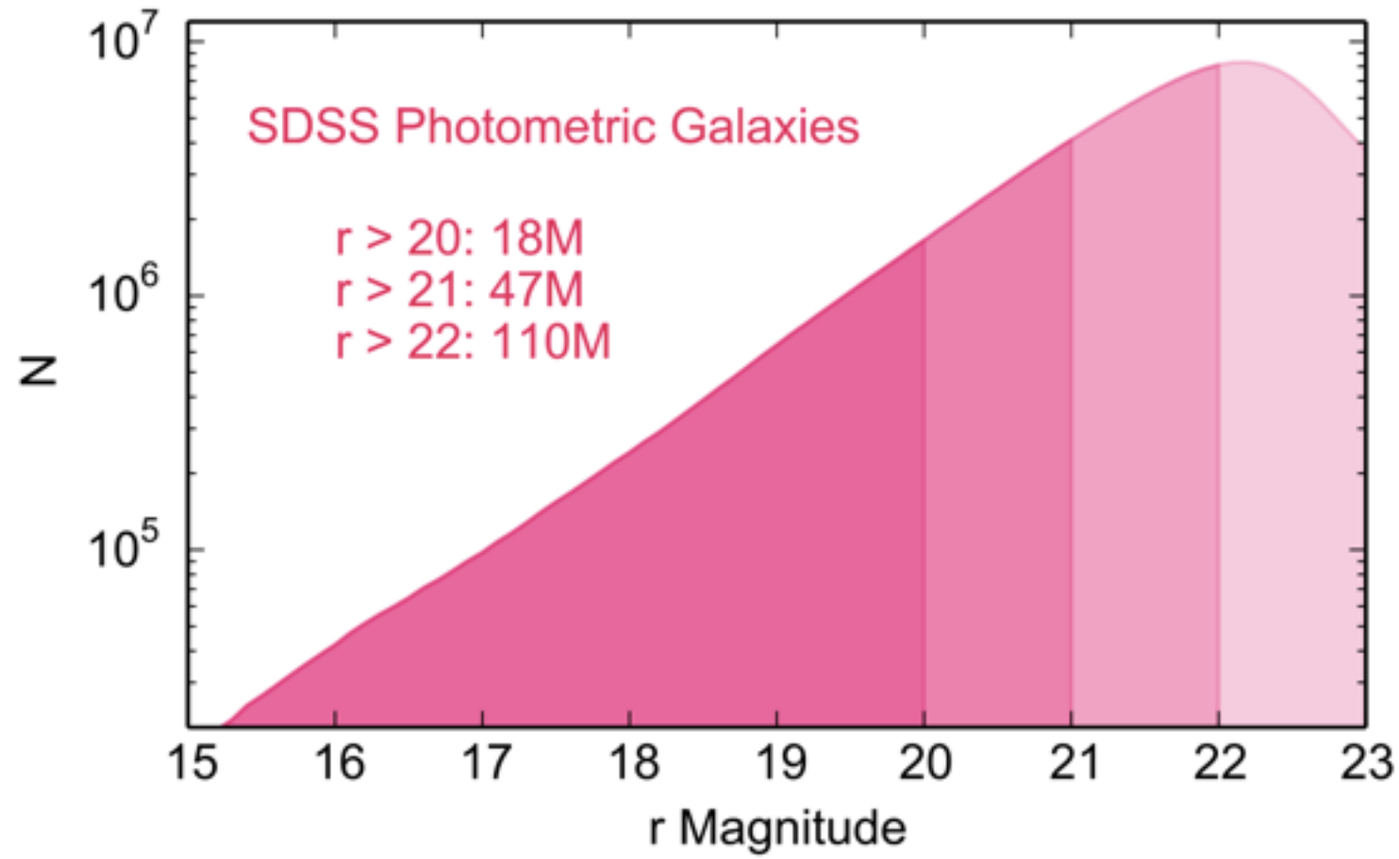
Comparison to photometric redshifts



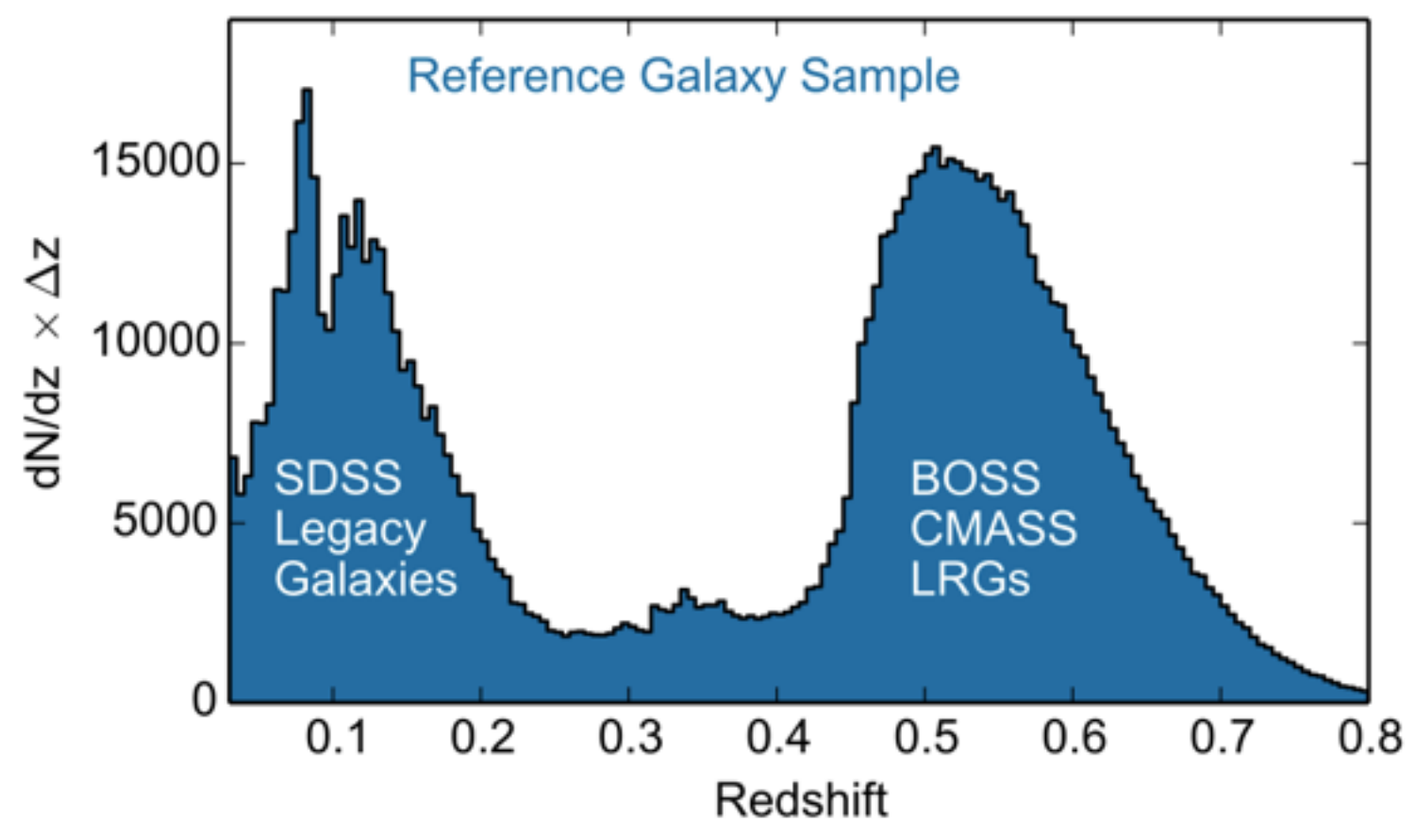
SDSS KD-tree photometric redshifts

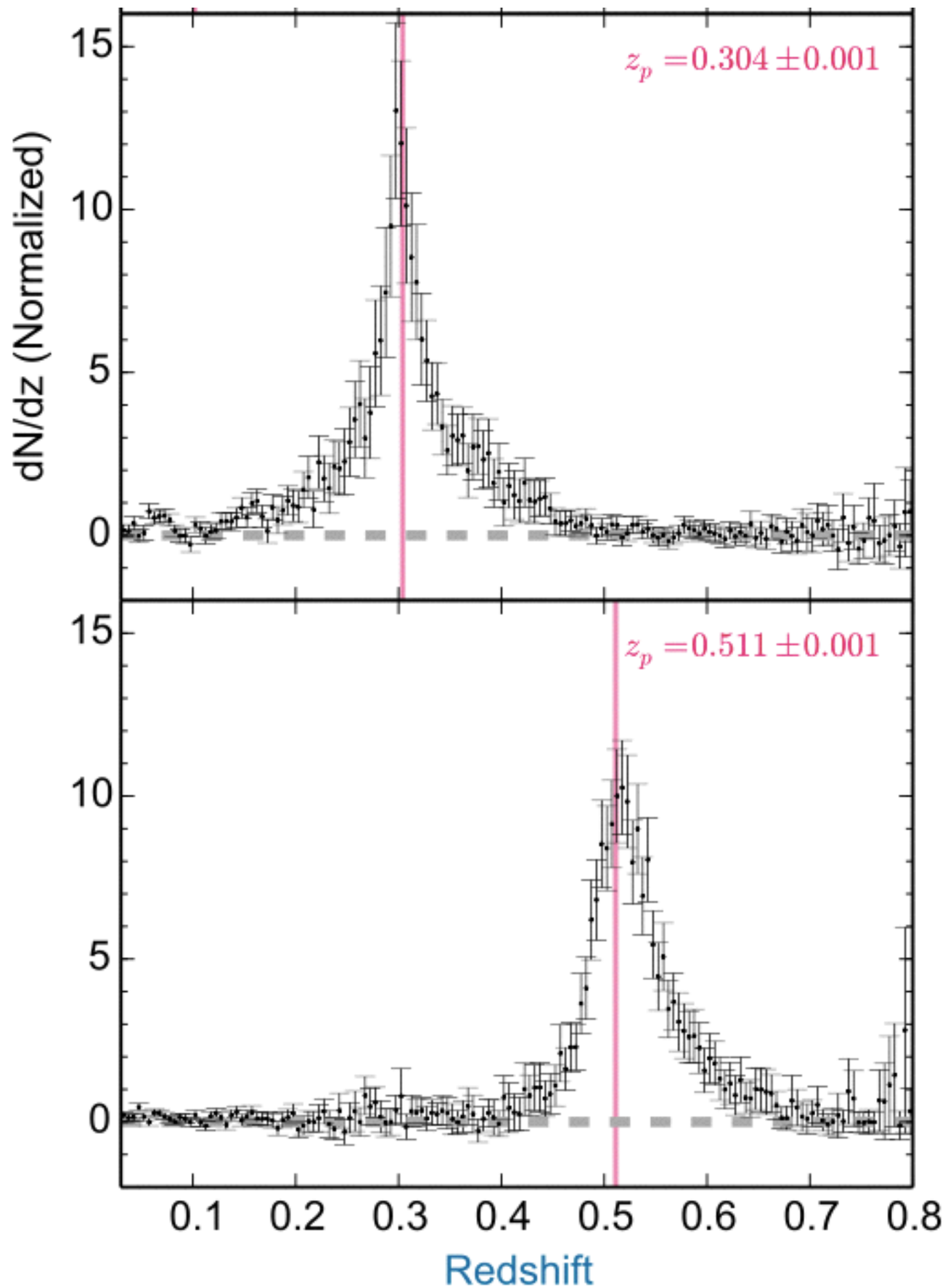
Application to SDSS photometric galaxies

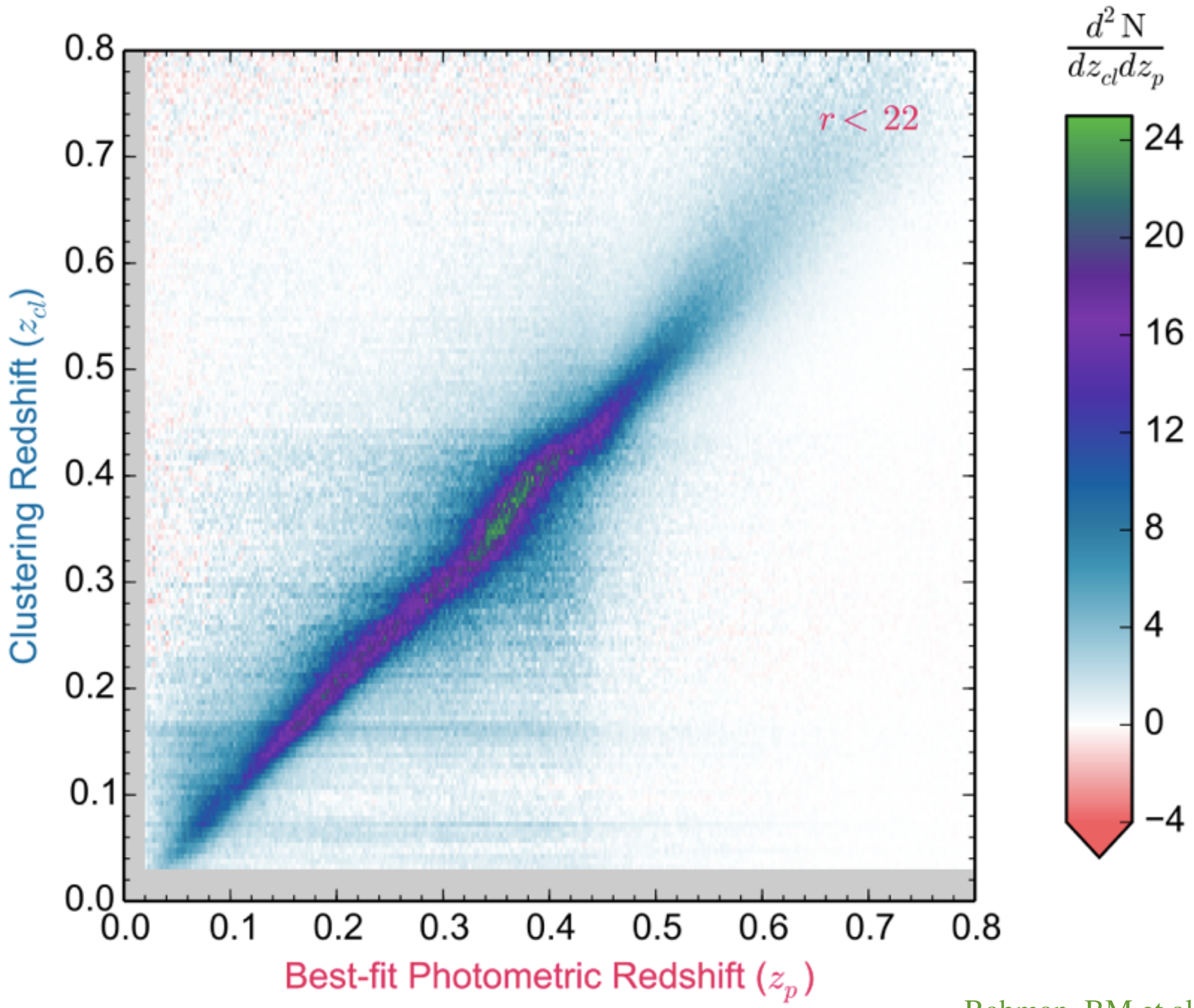
“unknown” sample



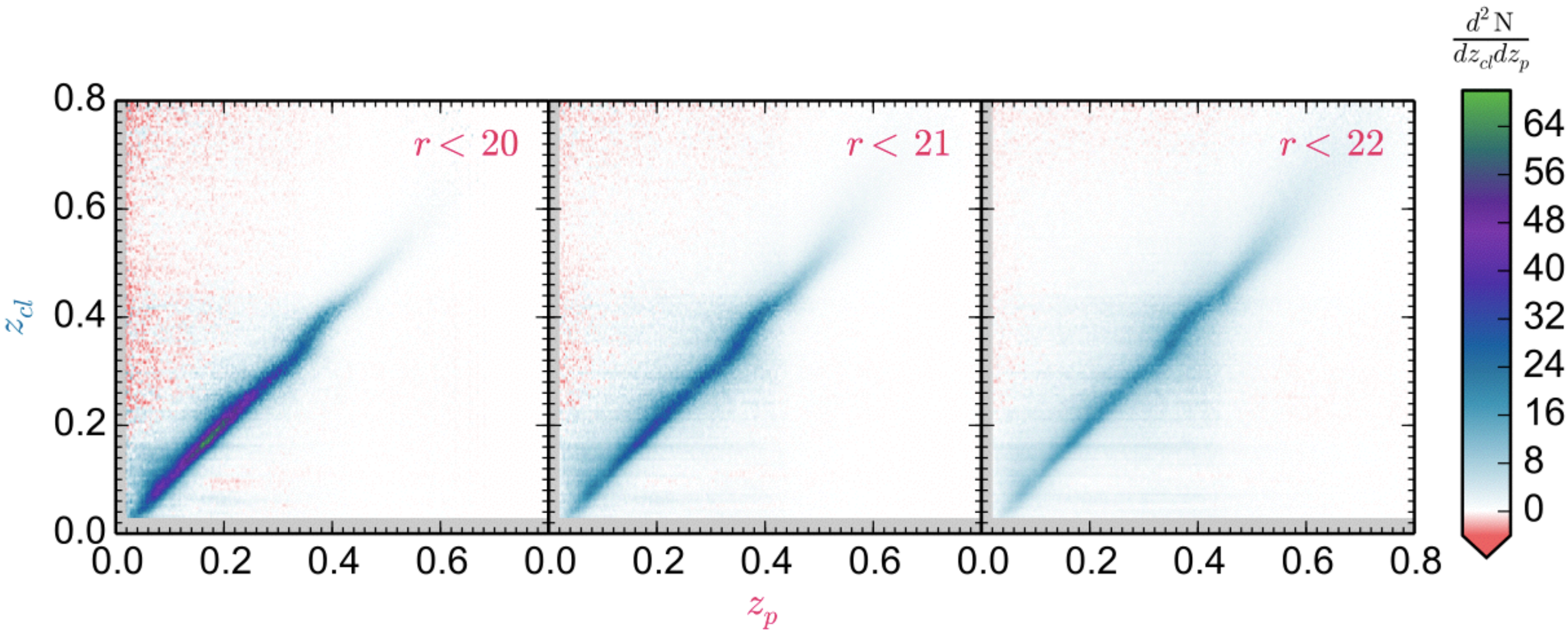
reference sample




$z_p = 0.304 \pm 0.001$
$z_p = 0.511 \pm 0.001$
dN/dz (Normalized)
Redshift
15
10
5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

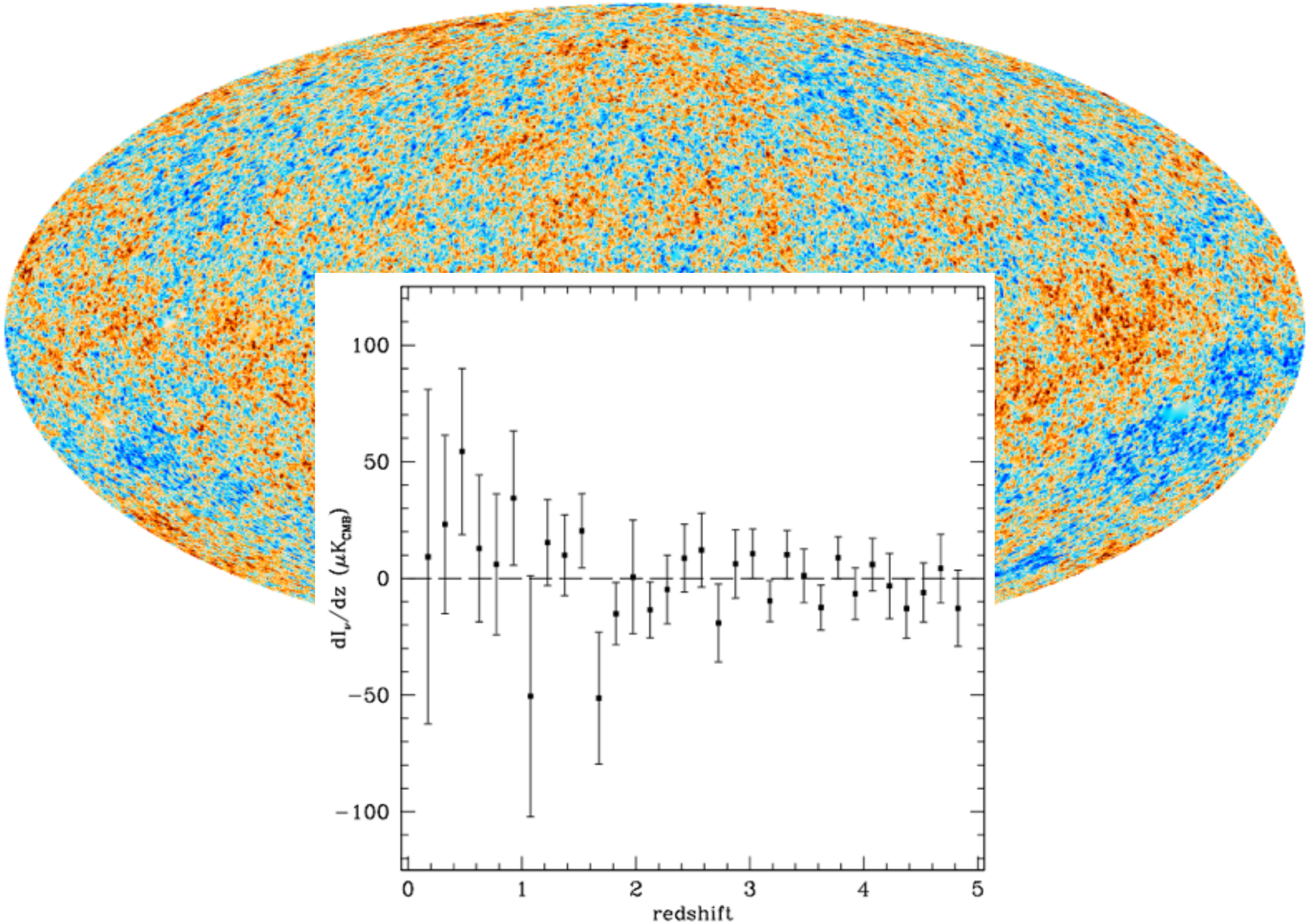


Rahman, BM et al., in prep



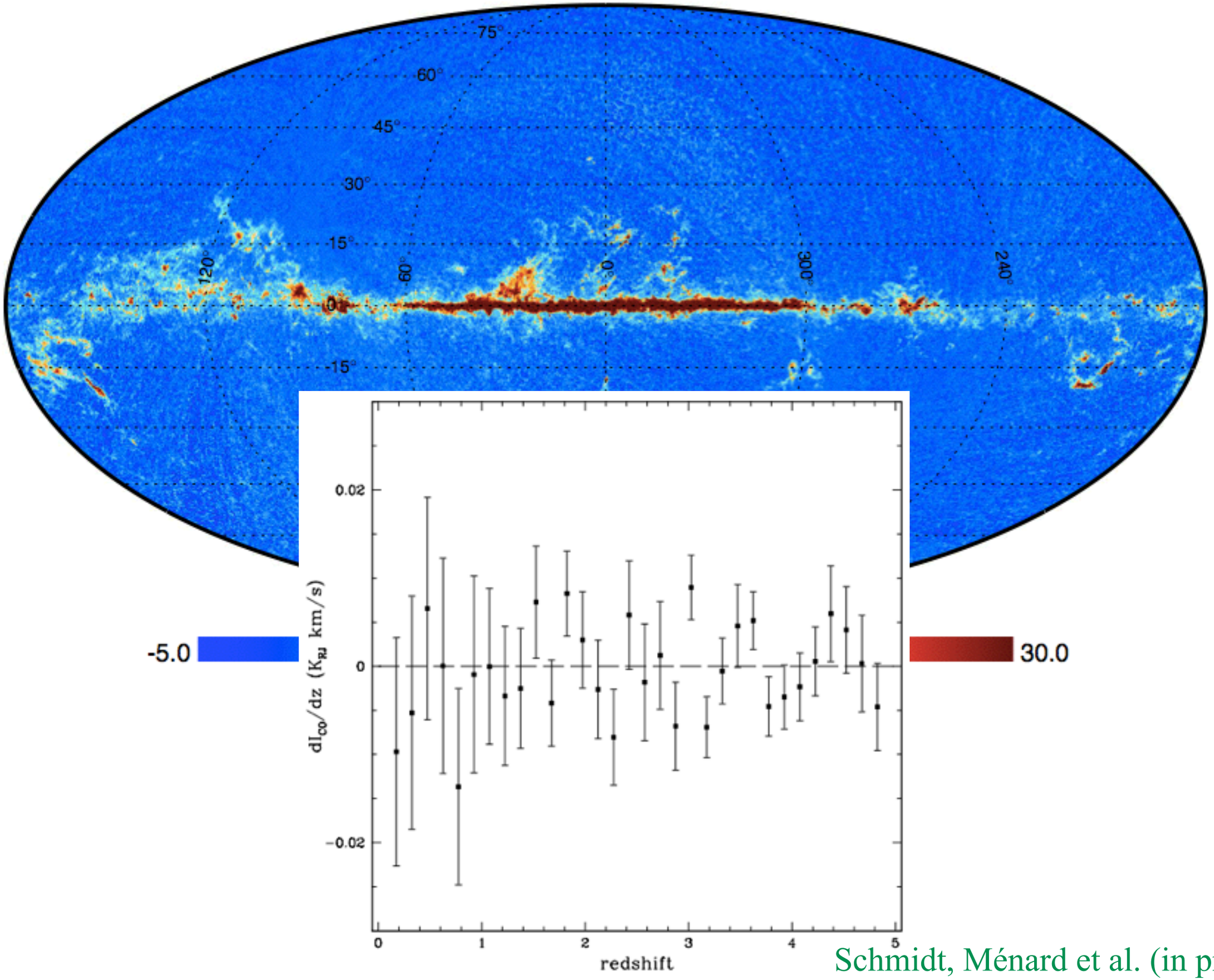
Rahman, BM et al., in prep

Planck CMB SMICA map



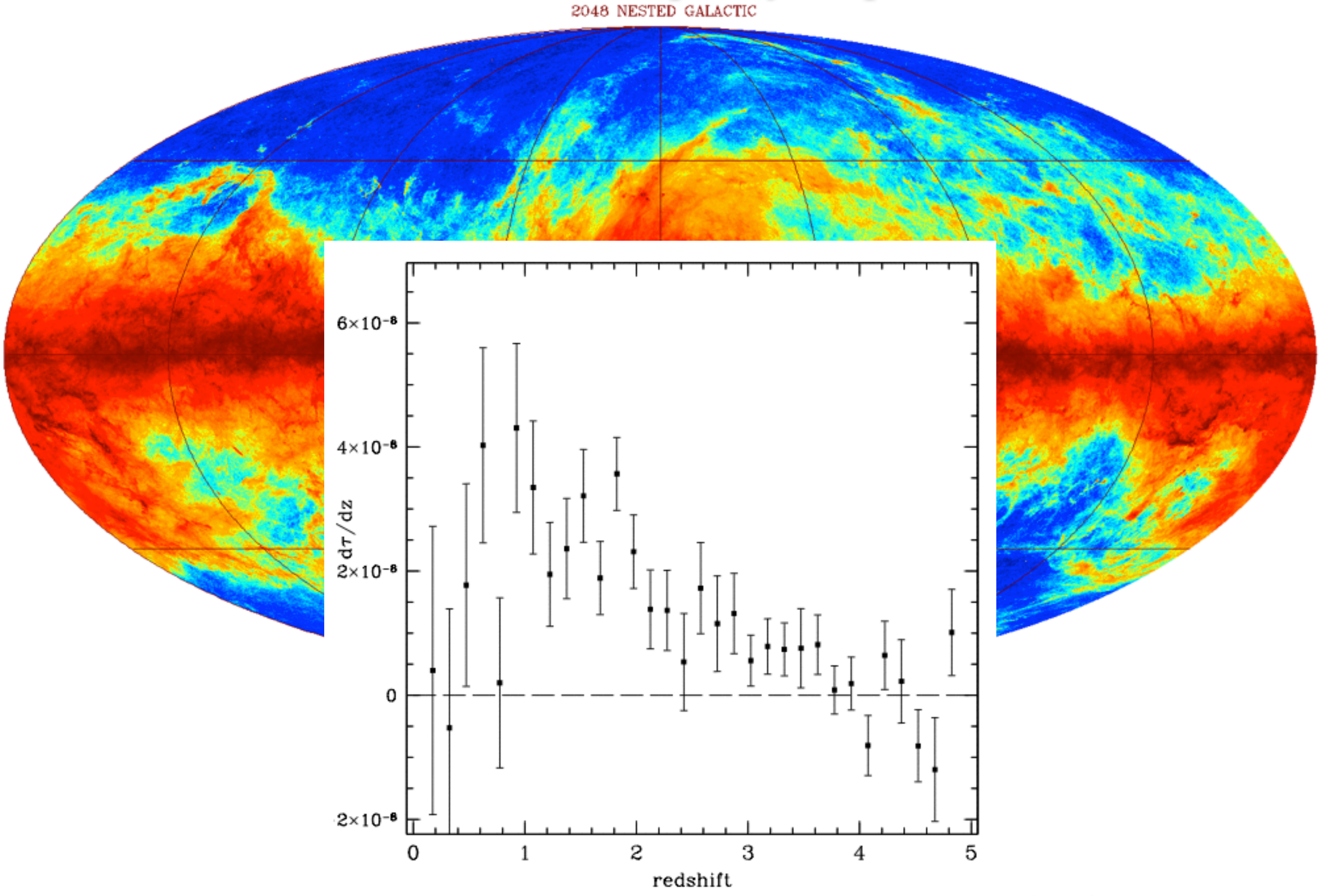
Schmidt, Ménard et al. (in prep)

Planck CO map J=0-1

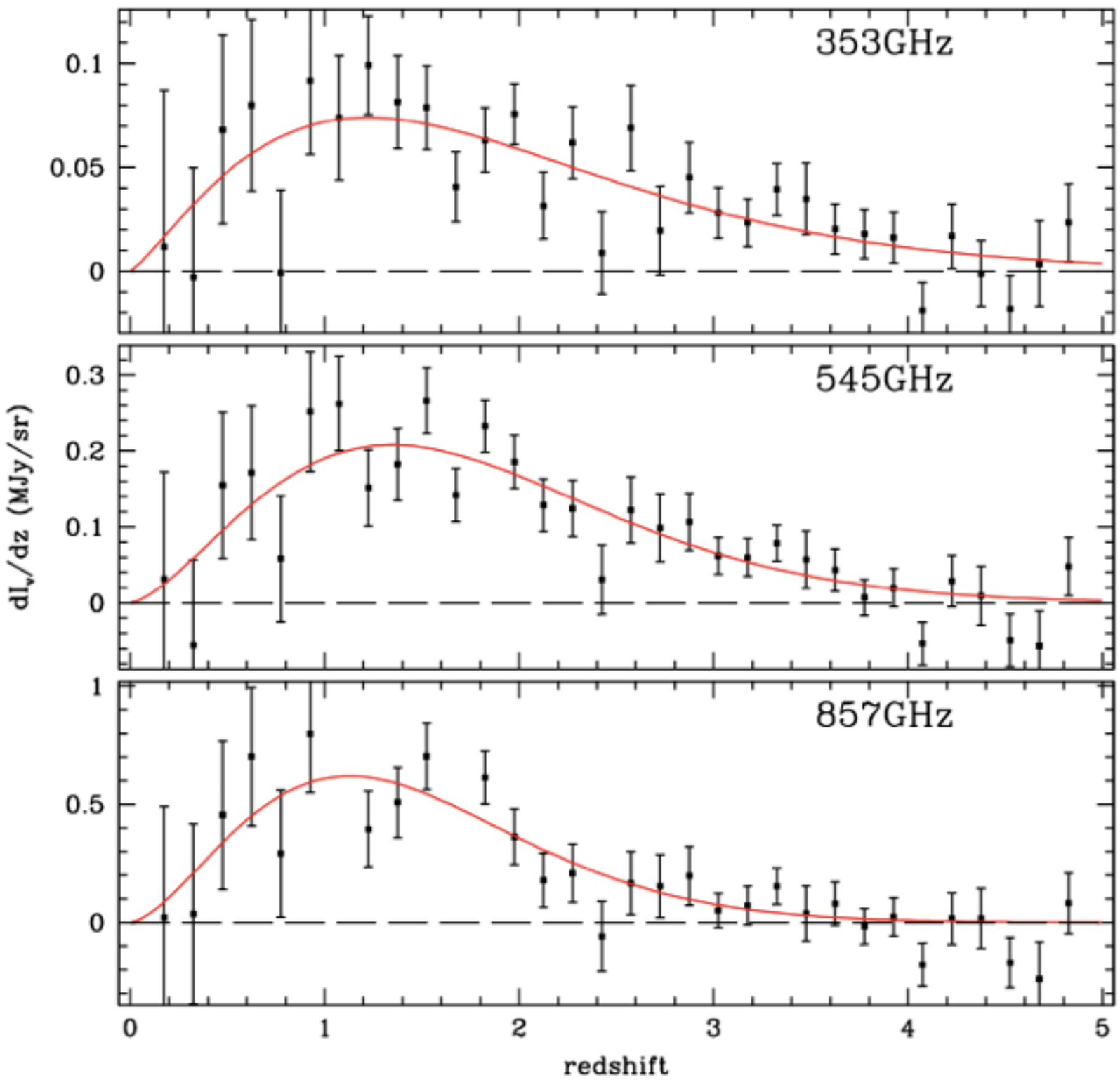


Schmidt, Ménard et al. (in prep)

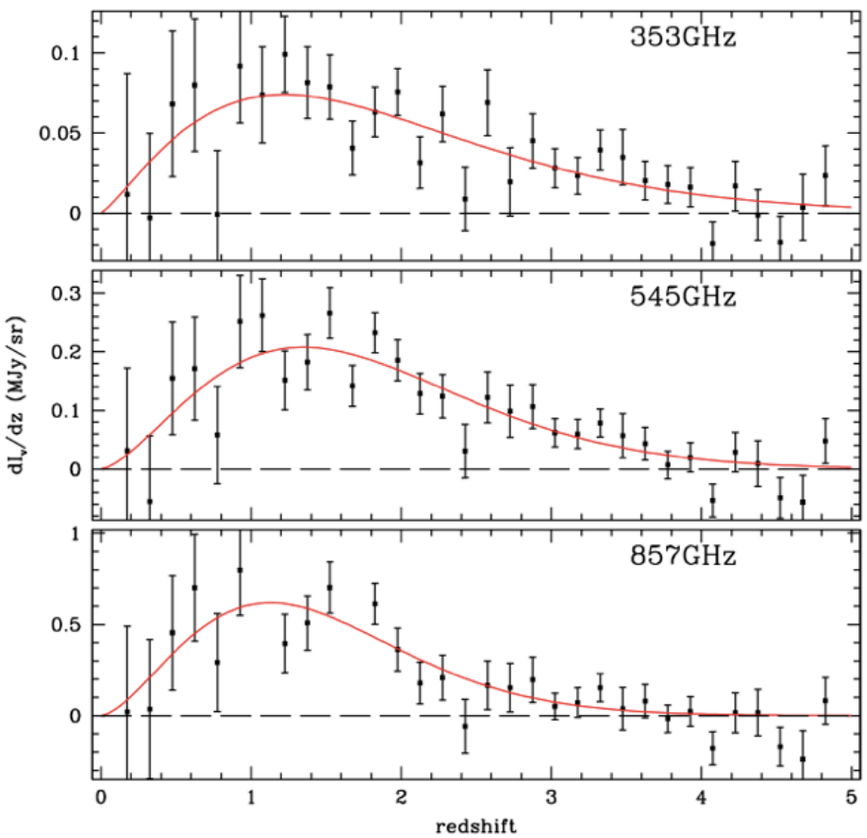
Planck dust opacity map



Schmidt, Ménard et al. (in prep)

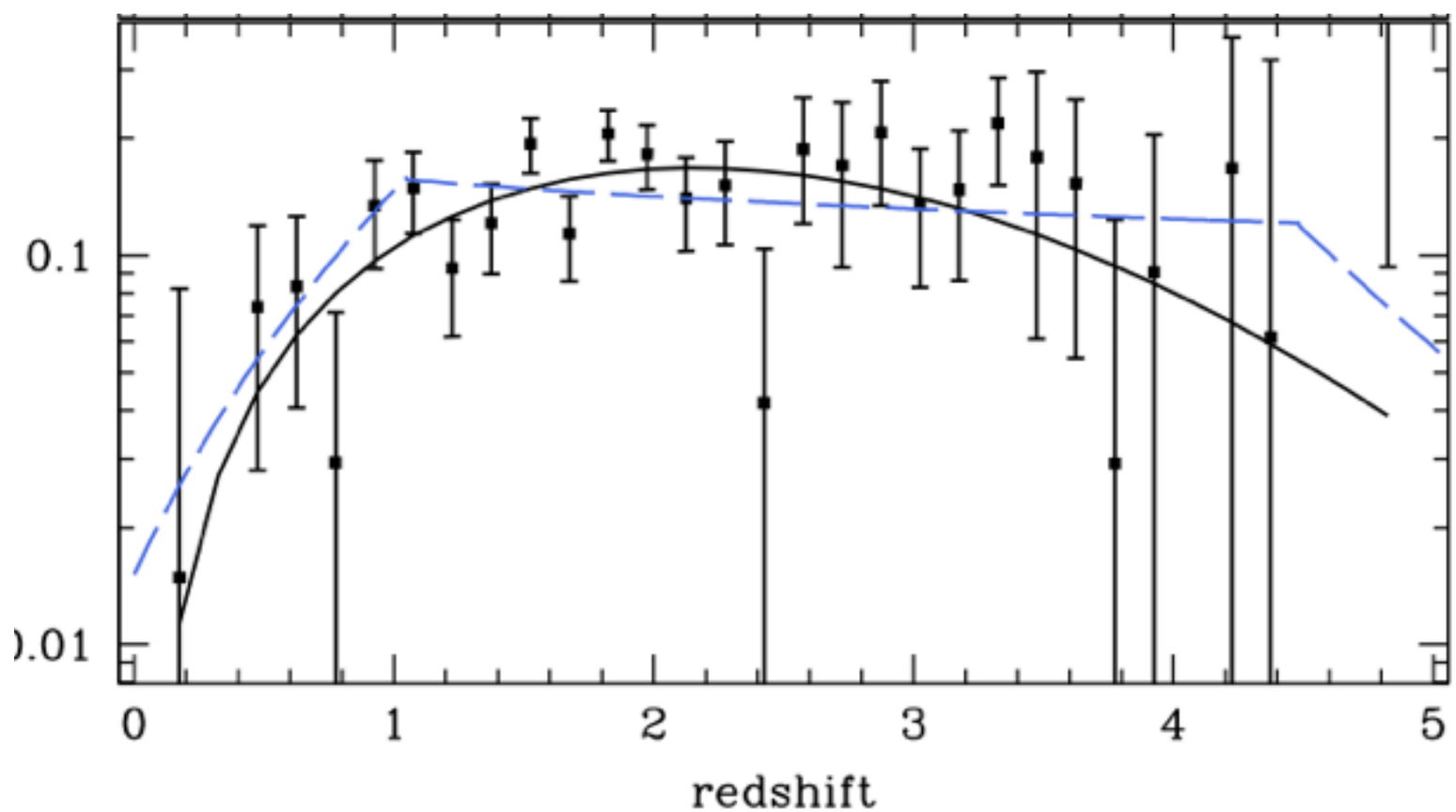


Schmidt, Ménard et al. (in prep)



Hopkins & Beacom 2006

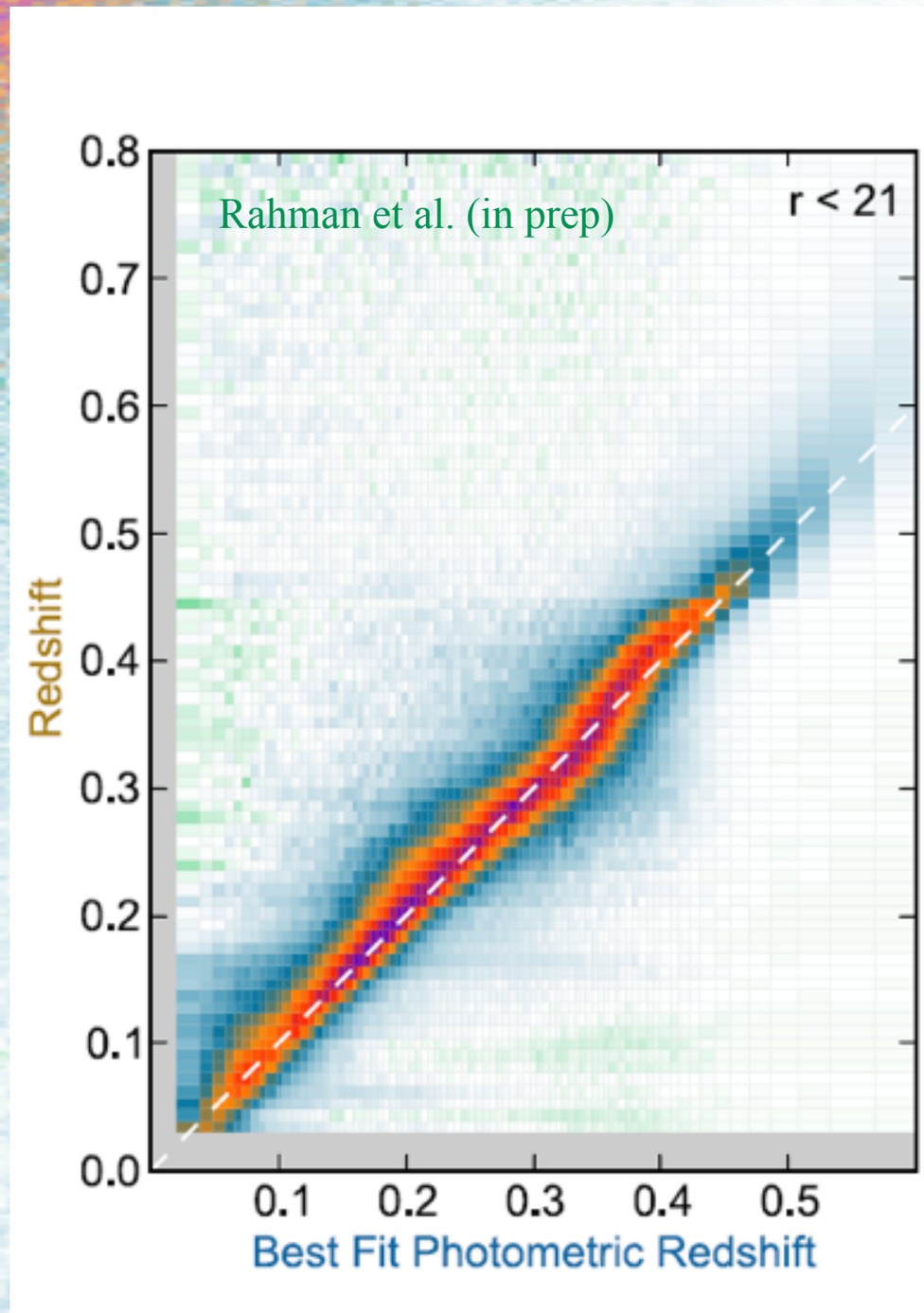
Star formation rate density [M_o/yr/Mpc^3]



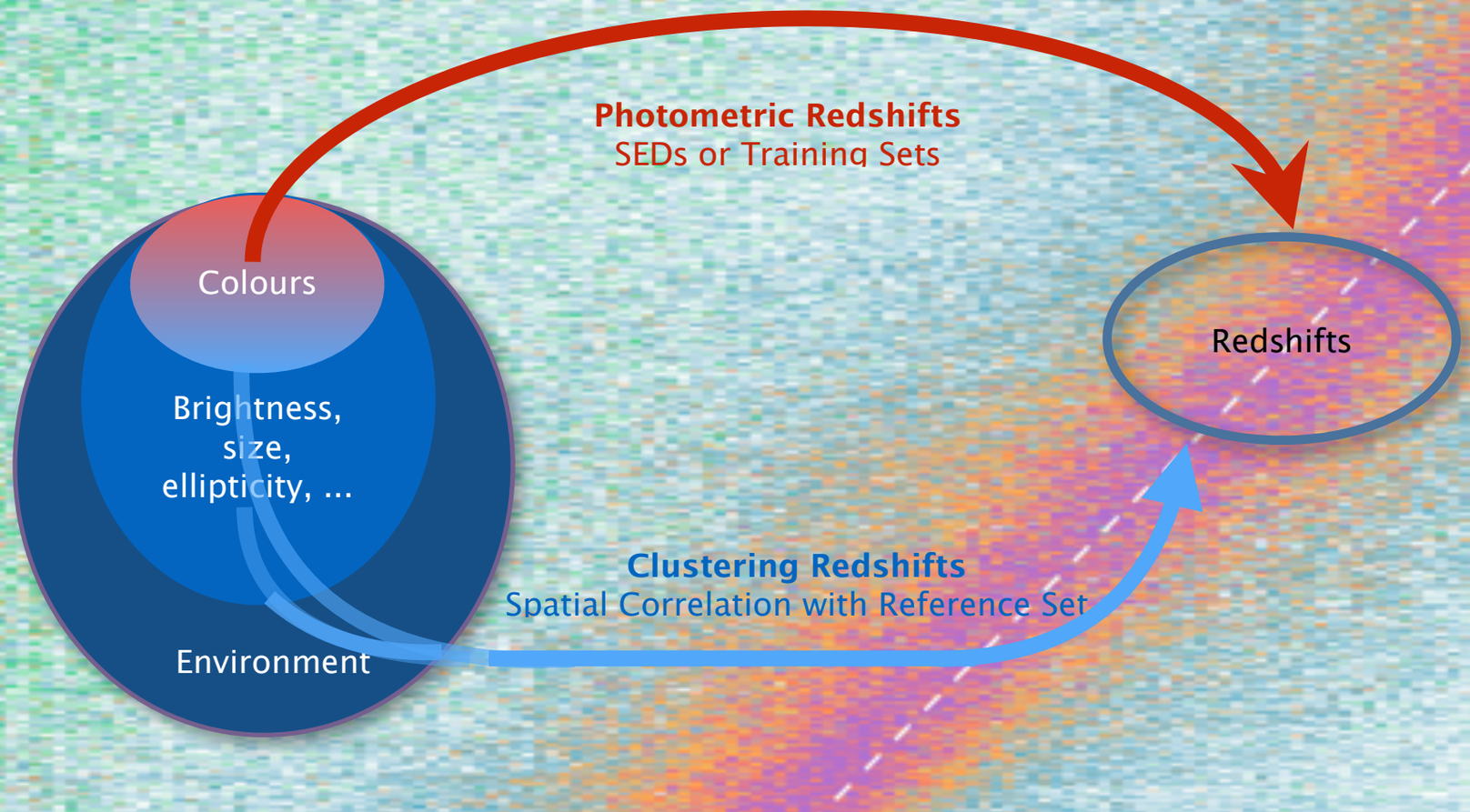
Schmidt, Ménard et al. (in prep)

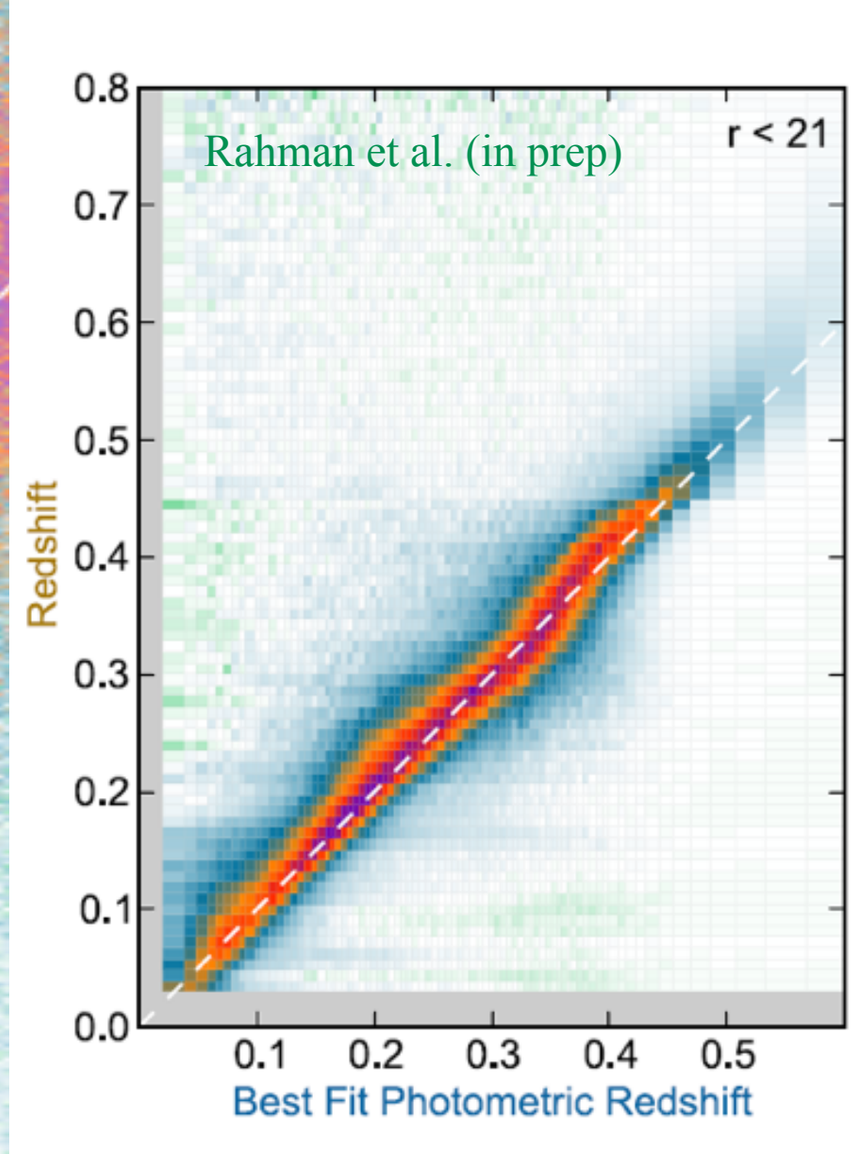
Summary

- We now have a new tool to explore the 3rd dimension of any sample
- The velocity accuracy is exquisite
- It alleviates many of the photo-z limitations
- A new tool to explore the Universe



Summary





- A new tool to explore the 3rd dimension of any sample
- Very high velocity resolution
- It alleviates many of the photo-z limitations