PFOF Group Finding in Pan-STARRS1 MD Survey Fields

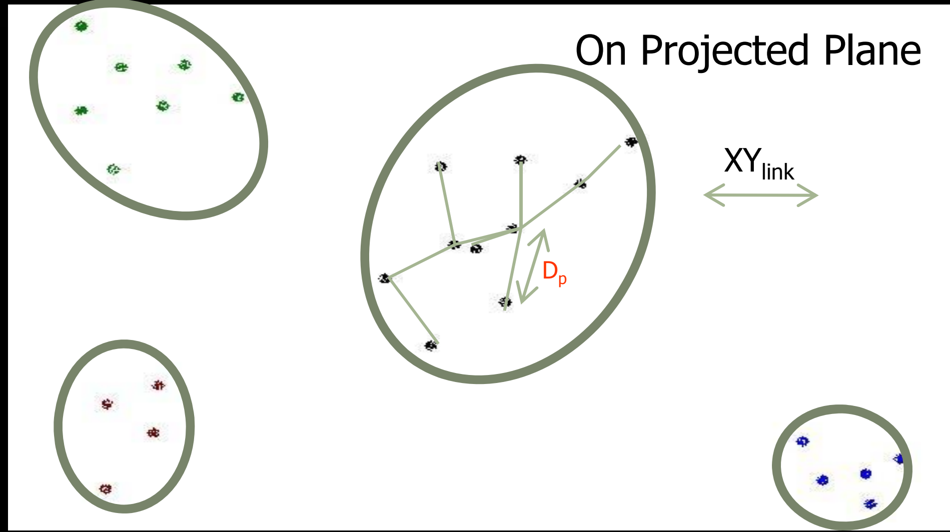
Hung-Yu Jian (NTU)

Collaborators: Lihwai Lin (ASIAA), Chin-Wei Chen (ASIAA), Sebastien Foucaud (NTNU), Tzihong Chiueh (NTU), Hau-Yu Liu (ASIAA), and Durham Group.

PFOF Algorithm (Liu et al. 2008)

Three input parameters:

xy_{link}, z_{link}, P_{th}



1) Criterion 1 – on projected plane

$$D_p \leq xy_{link} ,$$

where xy_{link} is the linking length in the projected plane

2) Criterion 2 – in redshift (line-of-sight) direction

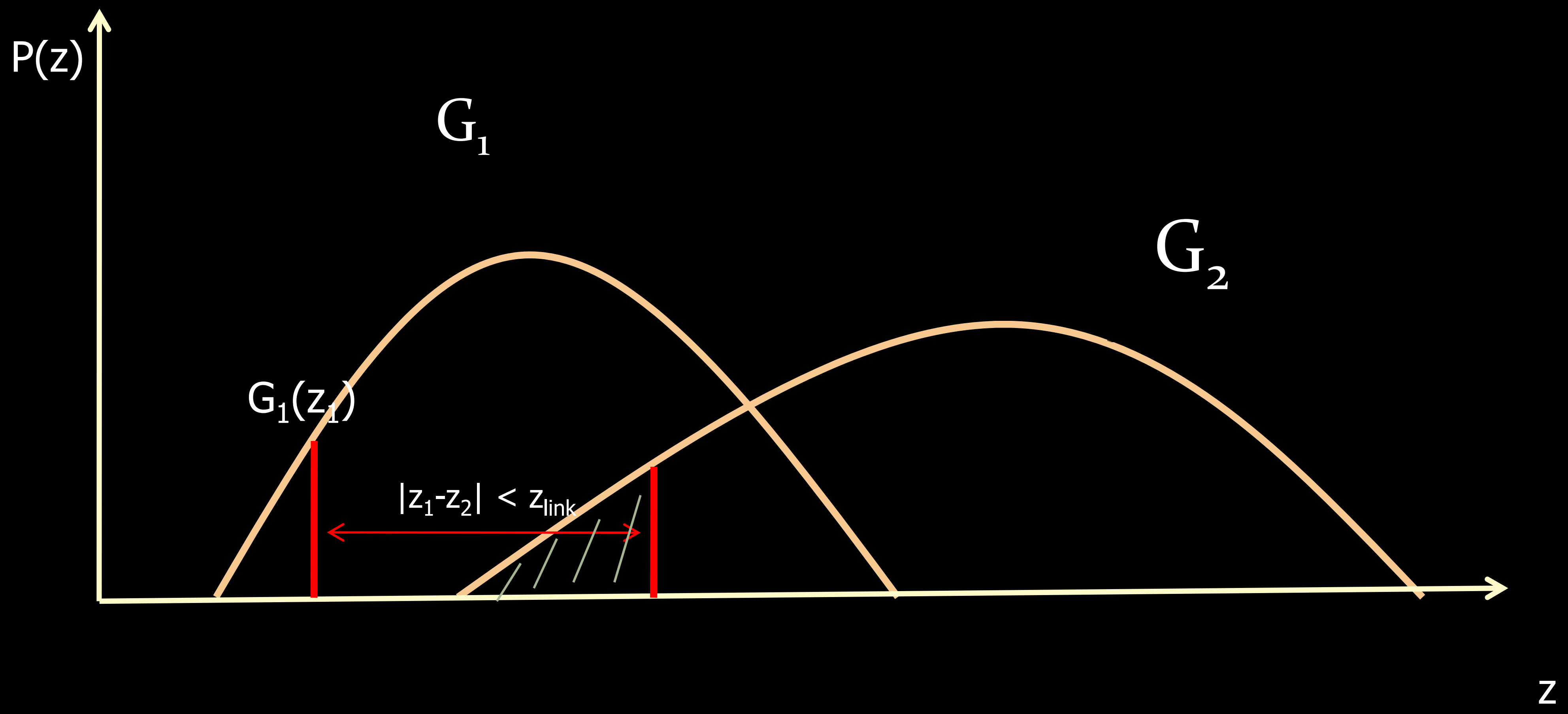
Calculate the linking probability $P(z_{link})$ of distance being less than z_{link}

$$P(z_{link}) \geq P_{th}$$

Two galaxies are physically linked if two criteria are both satisfied.

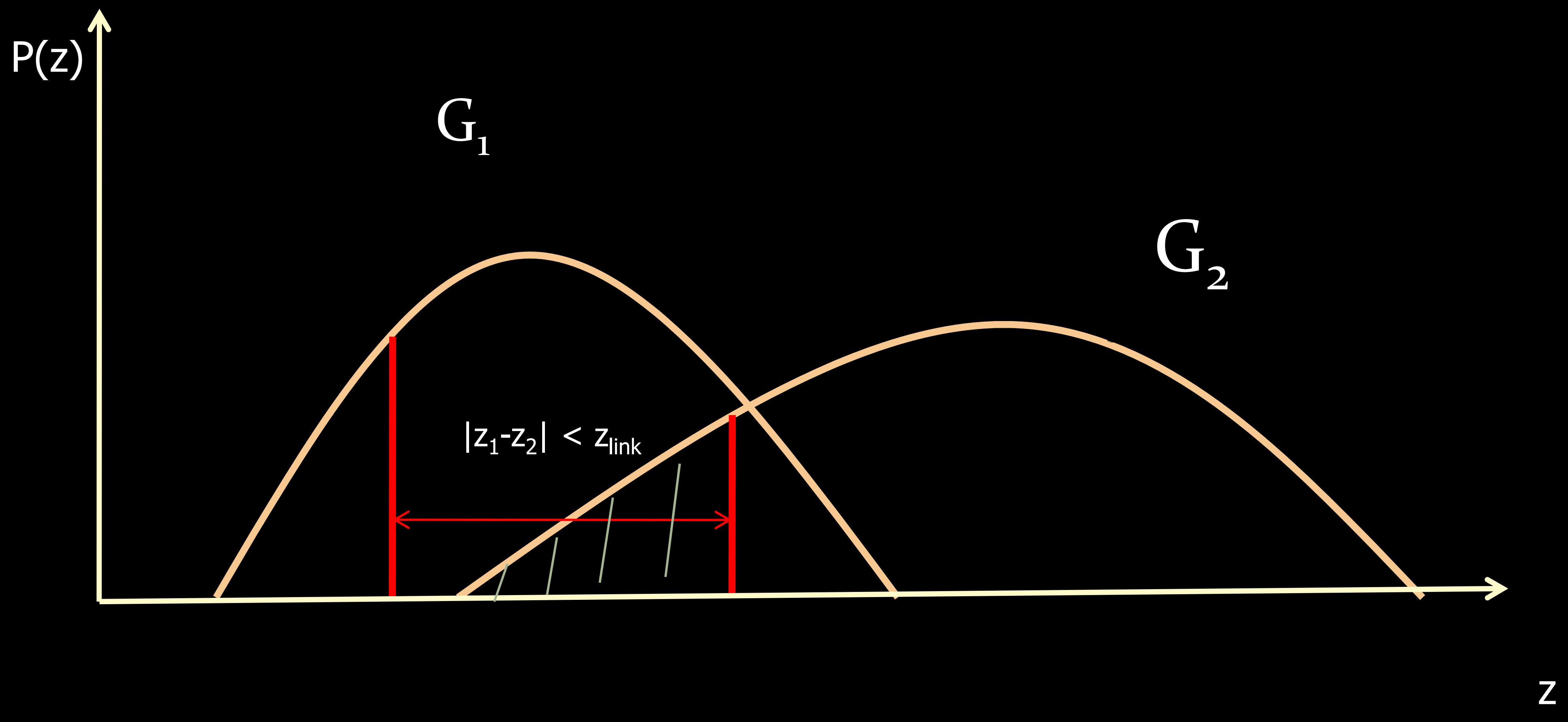
$$P(|z_1 - z_2| < z_{link}) \equiv \int_0^{\infty} dz G_1(z) \int_{z-z_{link}}^{z+z_{link}} dz' G_2(z') > p_{th},$$

where p stands for probability, G1 and G2 are the redshift probability distribution functions of two galaxies, and z_{link} is the linking length in the redshift direction



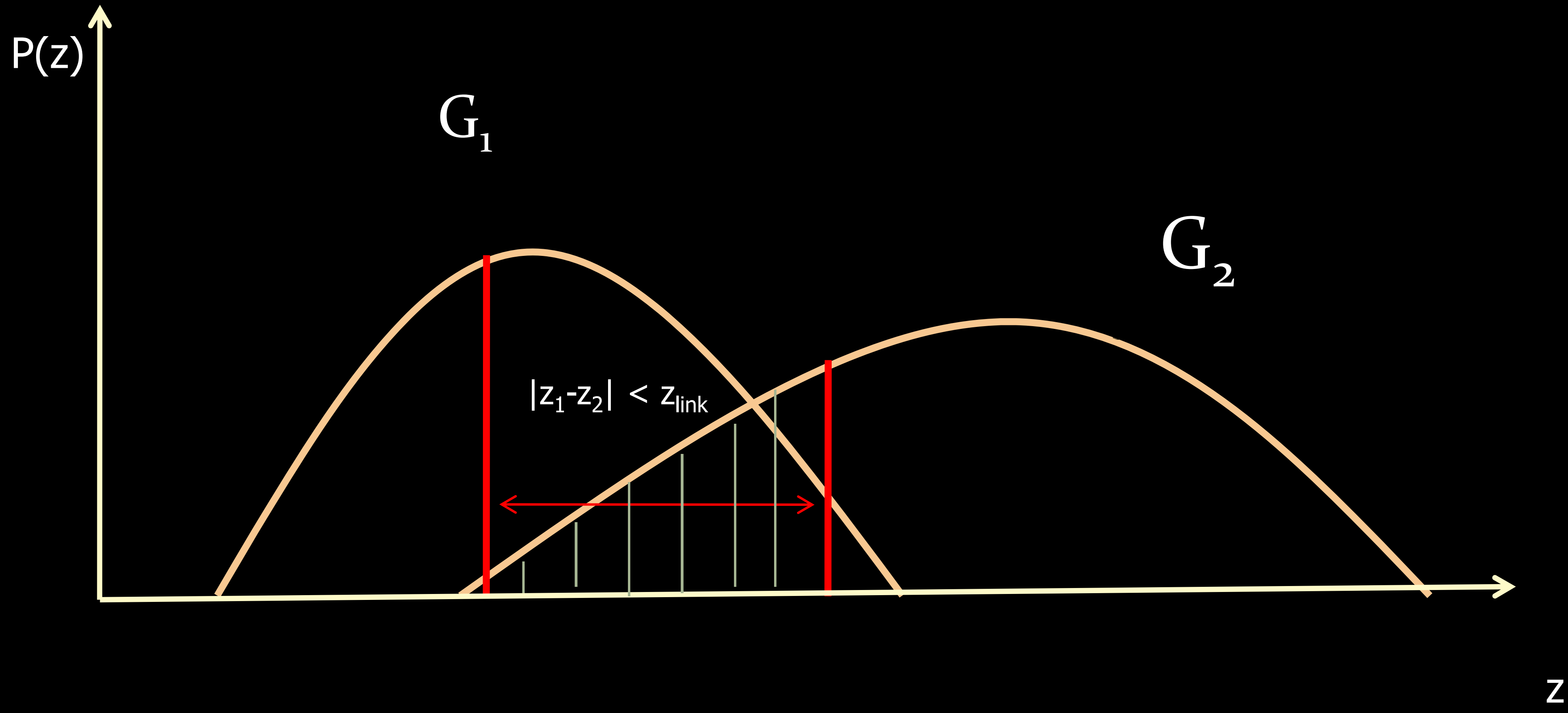
$$P(|z_1 - z_2| < z_{link}) \equiv \int_0^{\infty} dz G_1(z) \int_{z - z_{link}}^{z + z_{link}} dz' G_2(z') > p_{th},$$

where p stands for probability, G1 and G2 are the redshift probability distribution functions of two galaxies, and z_{link} is the linking length in the redshift direction



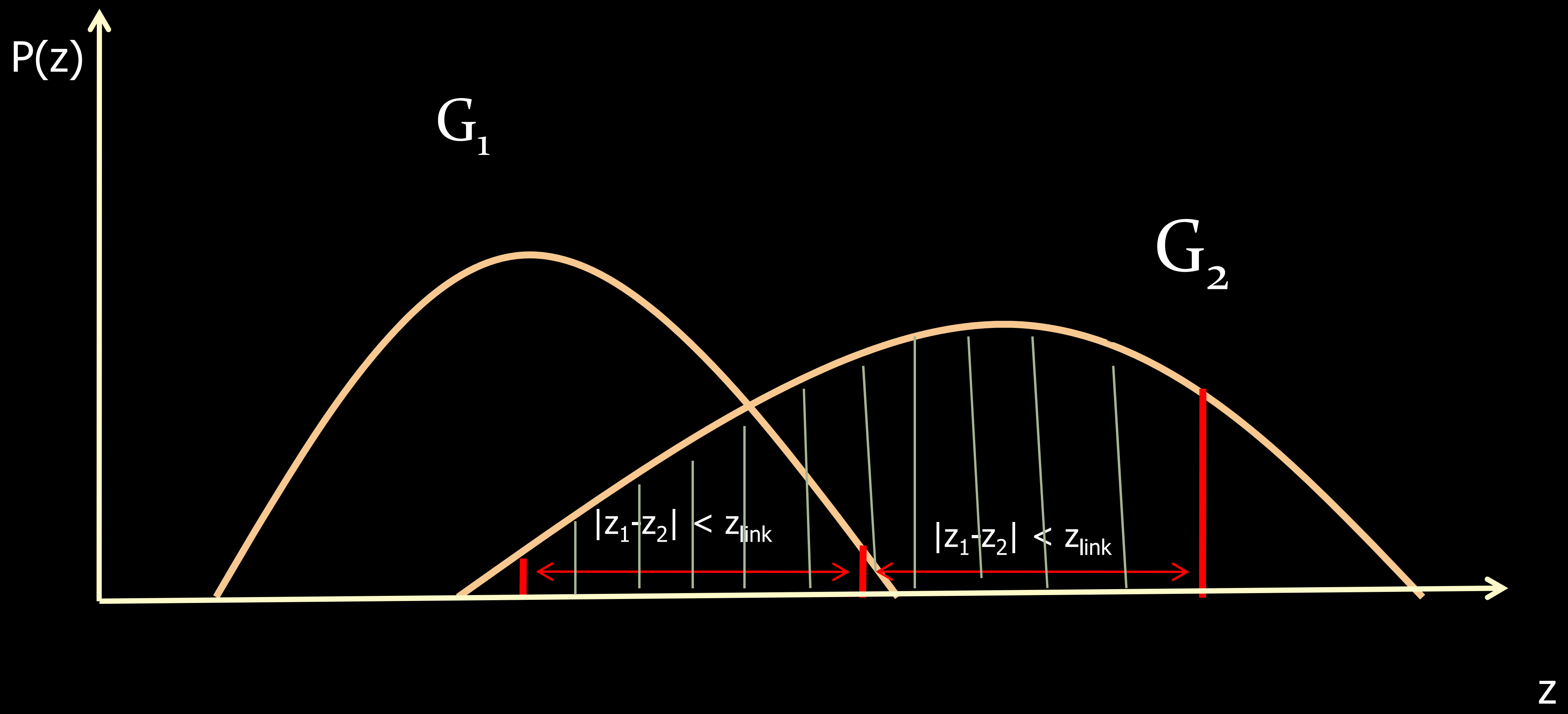
$$P(|z_1 - z_2| < z_{link}) \equiv \int_0^{\infty} dz G_1(z) \int_{z-z_{link}}^{z+z_{link}} dz' G_2(z') > p_{th},$$

where p stands for probability, G1 and G2 are the redshift probability distribution functions of two galaxies, and z_{link} is the linking length in the redshift direction



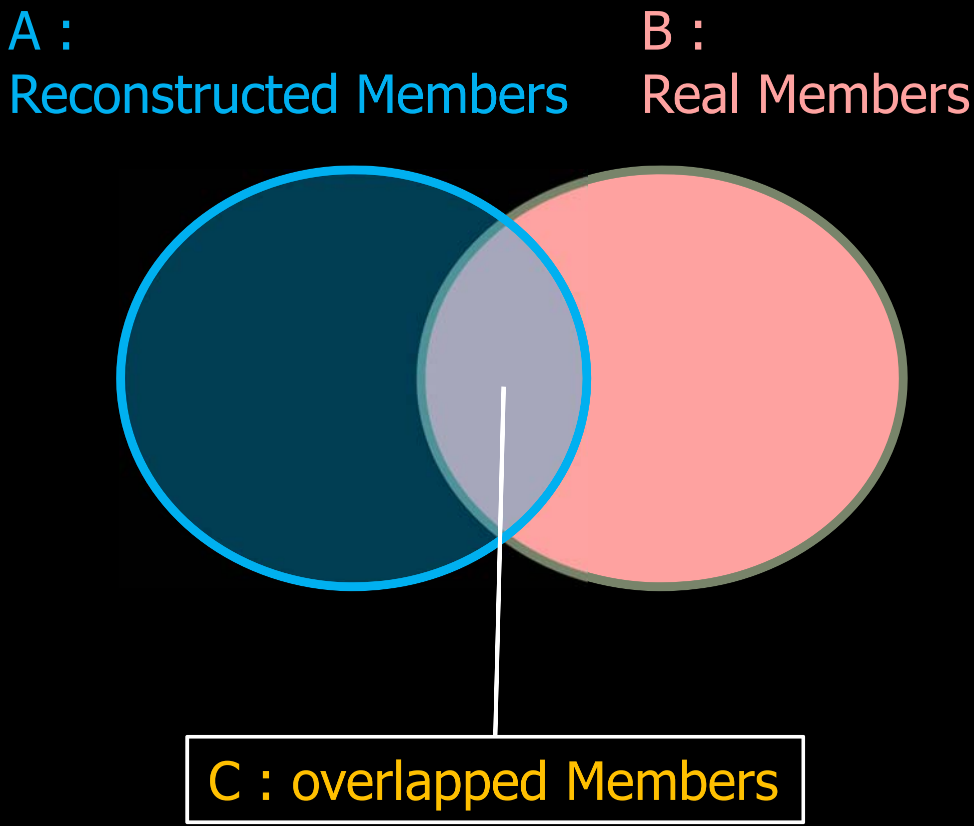
$$P(|z_1 - z_2| < z_{link}) \equiv \int_0^\infty dz G_1(z) \int_{z-z_{link}}^{z+z_{link}} dz' G_2(z') > p_{th},$$

where p stands for probability, G1 and G2 are the redshift probability distribution functions of two galaxies, and z_{link} is the linking length in the redshift direction



(1) Definition: One-way Purity and Completeness

(Gerke et al., 2005)



p1 (Purity)

= Number of Pure Groups /
Number of Reconstructed Groups

A Pure Group:
C / A > 0.5

c1 (Completeness)

= Number of Complete Groups /
Number of True Groups

A Complete Group:
C / B > 0.5

(2) Definition: Two-way Purity and Completeness mainly for optimization

A Pure-Complete Group: C / A > 0.5 and C / B > 0.5

p2 = Number of Pure-Complete Groups / Number of Reconstructed Groups

c2 = Number of Pure-Complete Groups / Number of True Groups

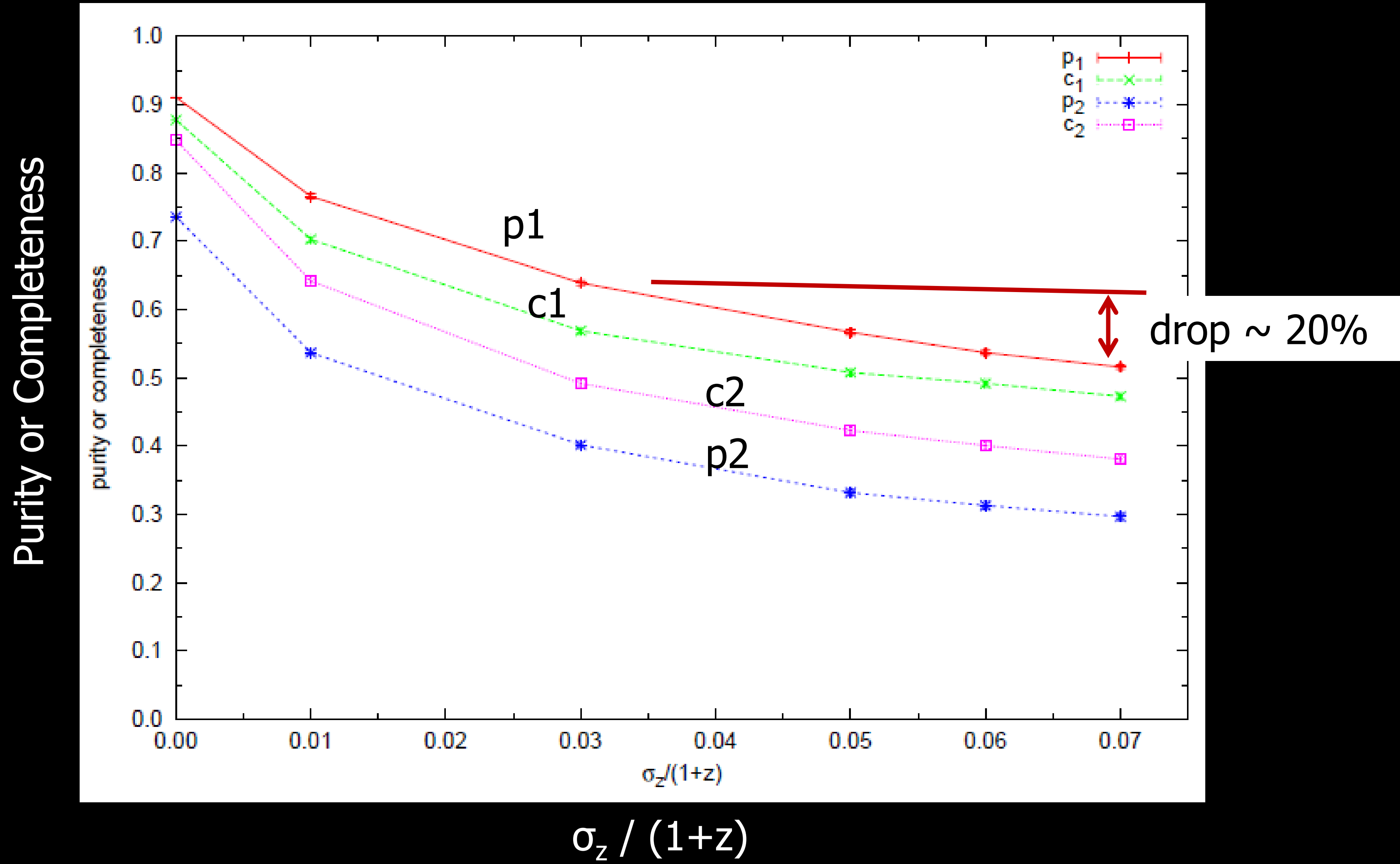
the optimization measure (Knobel et al., 2012)

$$\widetilde{g_1} = \sqrt[2]{[(1-\mathrm{p2})^2 + (1-c2)^2]/2}$$

Looking for the minimum of $\widetilde{g_1}$ ⟹ obtain optimized p1 and c1

(For example, if p2 =1 and c2 =1, $\widetilde{g_1} = 0$)

Redshift Uncertainty Test using Durham Mock :

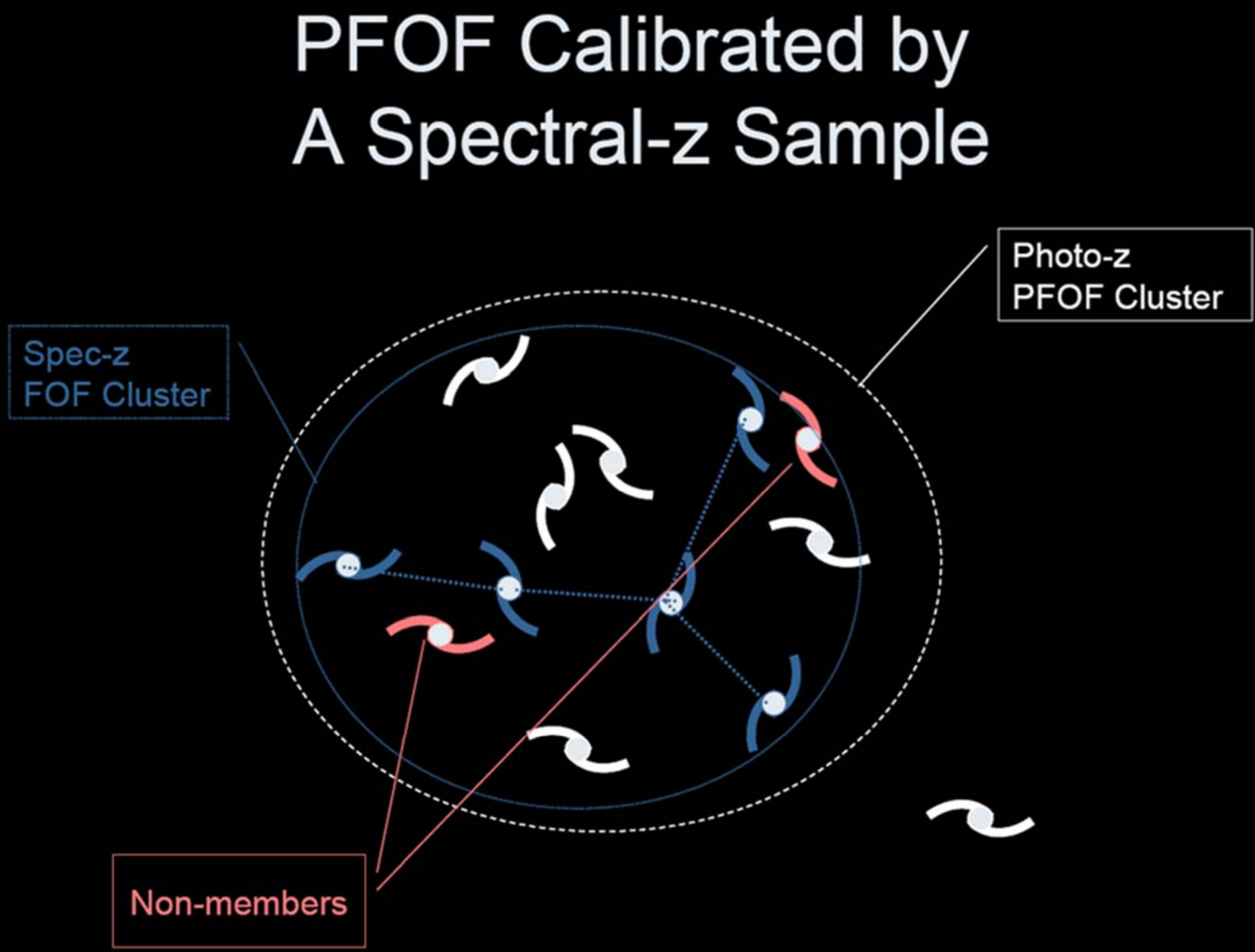


(1) From $\sigma_z/(1+z)$ = 0.03 to 0.07, p1 and c1 drop < 20%.

(2) PFOF Performance does not change significantly with redshift uncertainty between 0.03 and 0.07.

How to find the optimized linking lengths?

1. Using a mock catalog.
2. Using spectral-z groups in the field as a training set.
   ---For example, the ZCOSMOS groups in MD04.



“Subset Optimization” steps

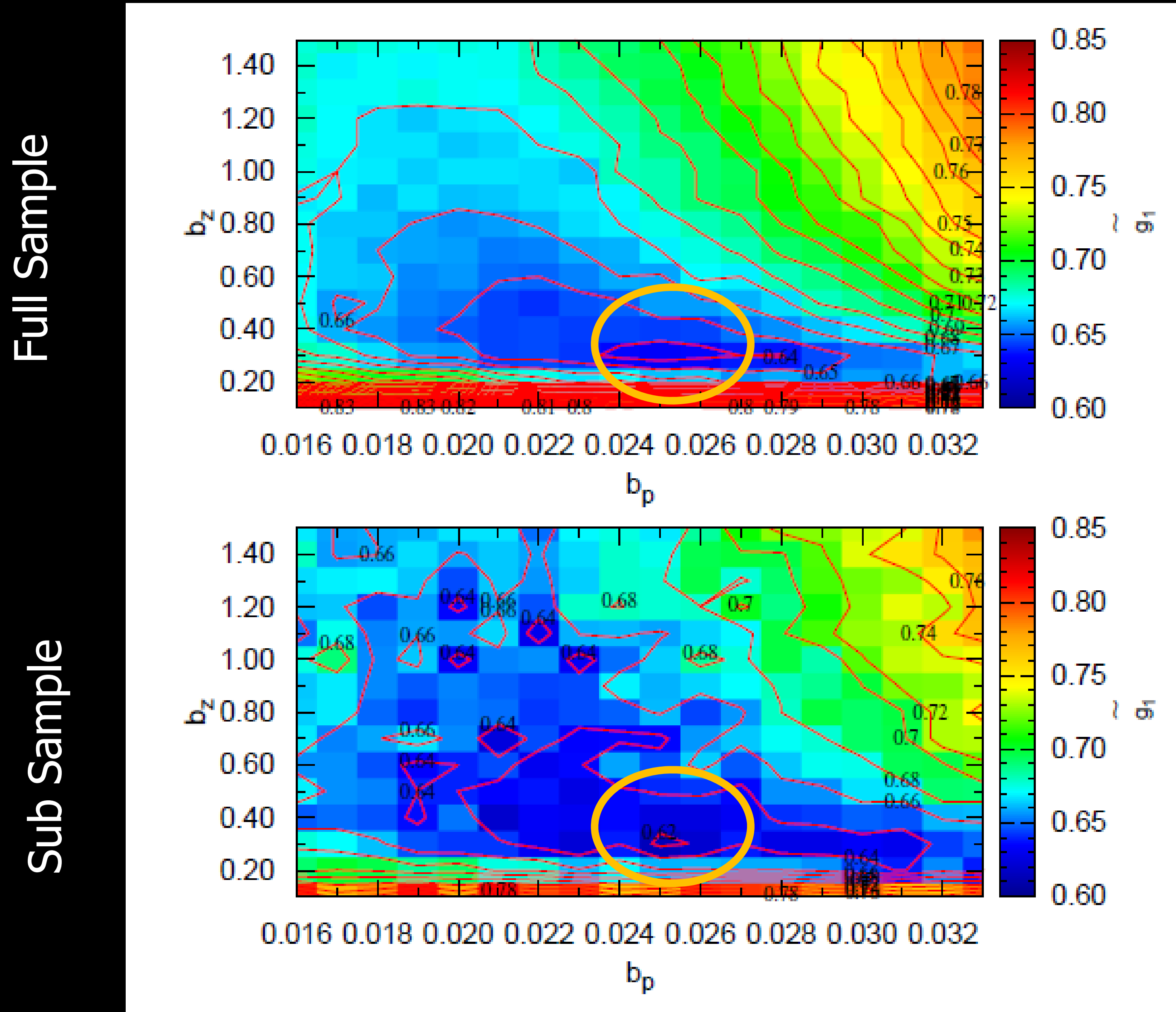
- Select a set of input parameters, bp, bz, p_{th}, for PFOF grouping on the full sample (photometric sample).
- Identify PFOF group members containing spectral-zs.
- Compute Purity and Completeness for these spectral-z members referenced to spectral-z groups.
- Re-select input parameters until the minimum optimization measure $\widetilde{g}_1$ is found

Optimized linking lengths from sub-sample

= Optimized linking lengths from full-sample ?

Mock Test on Subset Optimization

Randomly select 50% members from a sample as the sub-sample.



$\sigma z/(1+z) = 0.06$

At fixed Pth= 0.001

Linking parameters found from sub-sample with the minimum optimization measure are the same as the ones found from full-sample.

Real Data Application:
(I) COSMOS Field & PS1 MD04

(1) COSMOS Catalog

for input Data

- 0 < z < 1.4 and i AB< 24.2
- Field of View : ~ 2 deg^2
- # of galaxies: 115,844
- Photometric samples
- 30 bands: $\sigma z/(1+z) \sim 0.012$ (from Ilbert et al. 2009)

(2) PS1 MD04 Galaxy Catalog

for input data

- 0 < z < 1.4 and i < 24.1
- Field of View : ~ 7 deg^2
- # of galaxies: 278,860 PS1_I_CLASS_STAR =0.4)
- Photometric samples
- 6 bands: u, g, r, i, z, y bands: $\sigma z/(1+z) \sim 0.046$

(3) 20k ZCOSMOS Group Catalog for training set

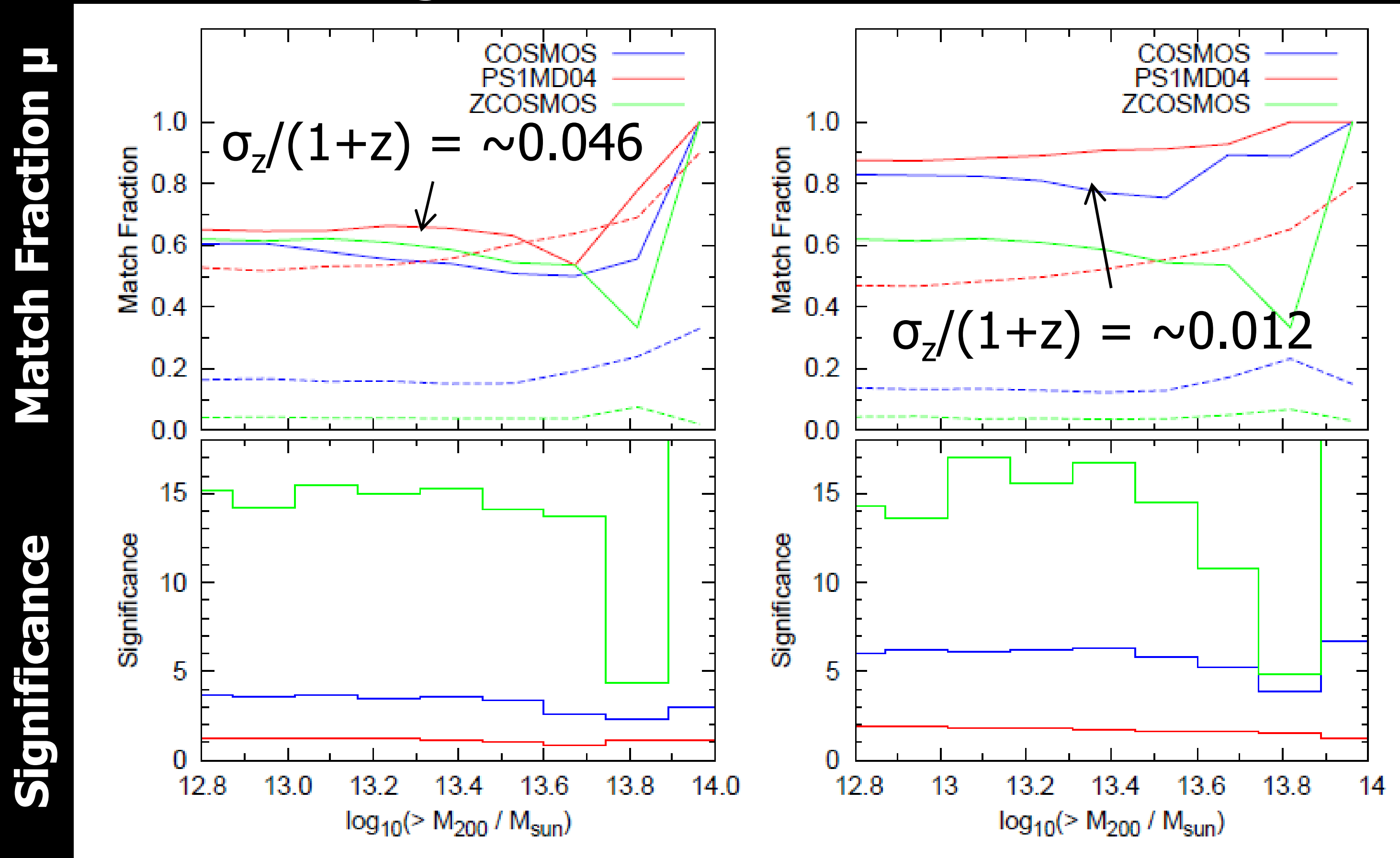
- 0 < z < 1.2 and i < 22.5
- Field of View: ~ 1.7 deg^2
- # of galaxy: ~ 16,500
- # of groups: ~ 1,284
- Spectroscopic samples
- the group catalog: Knobel et al., 2012, ApJ, 753, 112.

(4) XMM-NEWTON X-RAY Group Catalog for performance test

- # of Groups: 129
- z < 1.
- Field of View: ~ 2 deg2
- X-ray group catalog: Finoguenov et al. in prep.

Cross Matching with the XMM-NEWTON X-RAY Group Catalog

(i) zCOSMOS groups as the training set | **(ii) x-ray groups as the training set**



Match criterion: Group center has to be inside r200 of a x-ray group and their redshift difference is within 1.5 $\sigma_z/(1+z)$.

Significance= Solid line / dotted line

(1) When using zCOSMOS groups as the training set, COSMOS and PS1MD0 groups both have a match fraction ~60% similar to the match rate of zCOSMOS, but there is huge difference in significance; i.e. it is ~ 15 for zCOSMOS, ~ 4 for COSMOS, and ~1 fro PS1MD04.

(2) When using X-ray groups as the training set, the match fraction is ~ 81% for COSMOS and ~ 84% for PS1MD04, and the significance is ~ 6 for COSMOS and ~ 2 for PS1MD04.

Grouping Results:

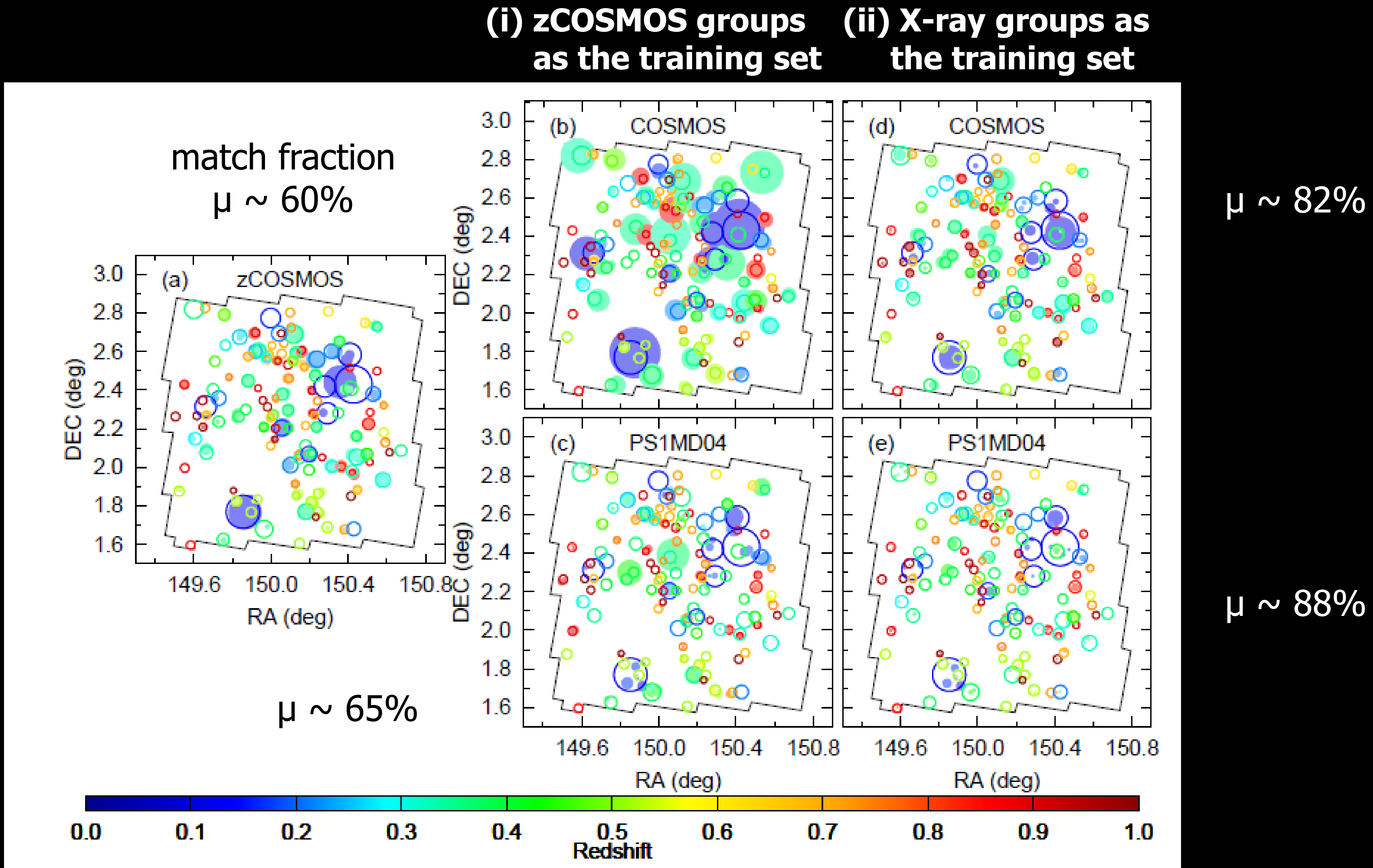
(1) Optimized with 20 k ZCOSMOS groups

	COSMOS	PS1 MD04
p1	0.94	0.86
c1	0.72	0.49
X-ray match fraction	0.60	0.65
BCG match fraction	0.63	0.37
Number of Groups	5,100	5,093

(2) Optimized with x-ray groups

	COSMOS	PS1 MD04
p1	0.98	0.91
c1	0.45	0.31
X-ray match fraction	0.82	0.88
BCG match fraction	0.72	0.44
Number of Groups	3,724	4,741

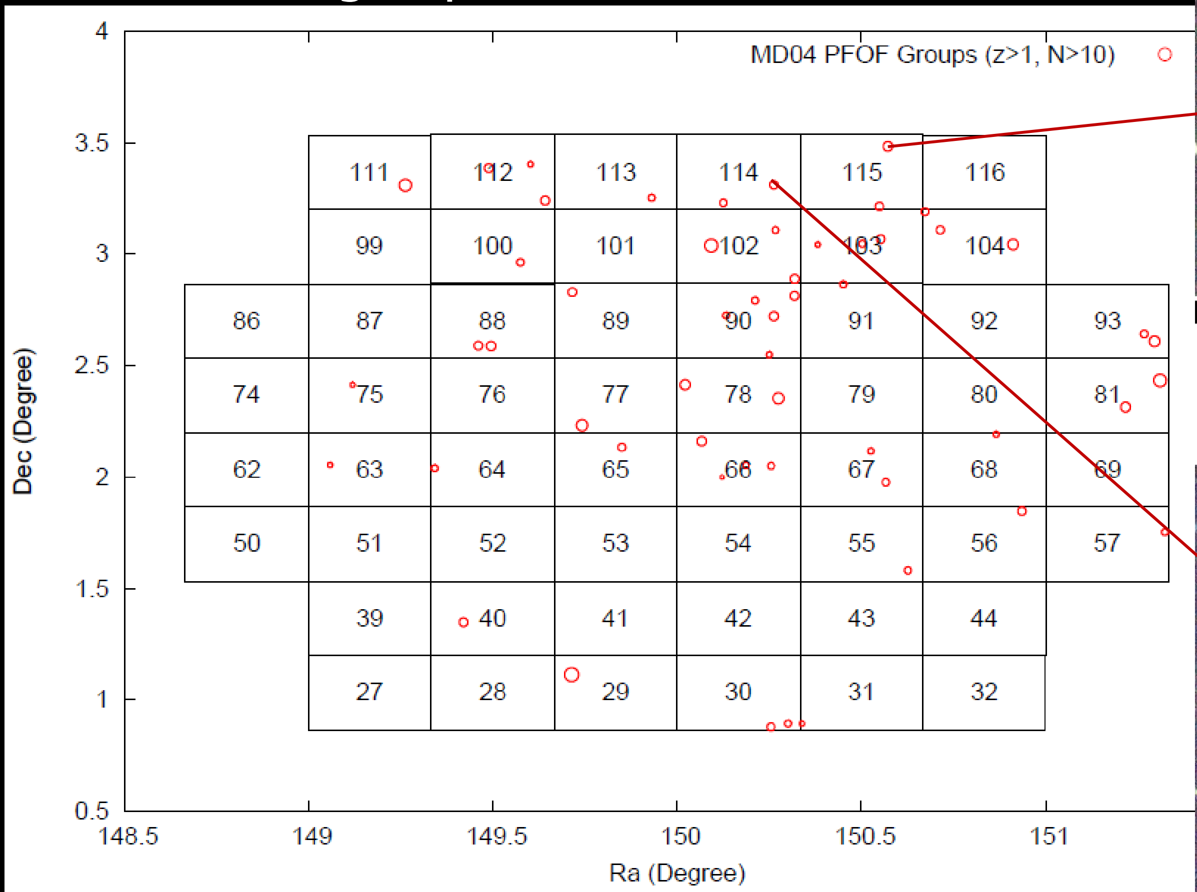
Map for X-ray and Match PFOF Groups



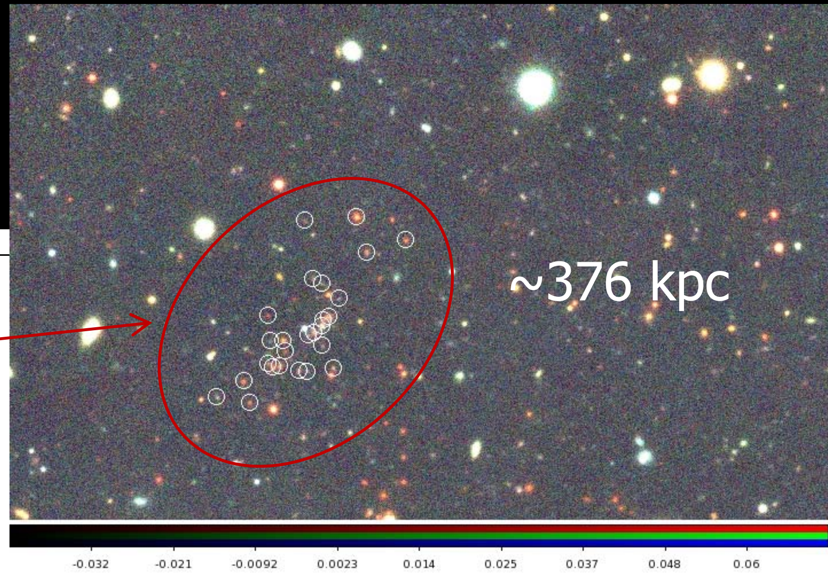
Two Examples of High-z PFOF Groups in MD04

rgb color :
Red -- z
Green -- r
Blue -- g

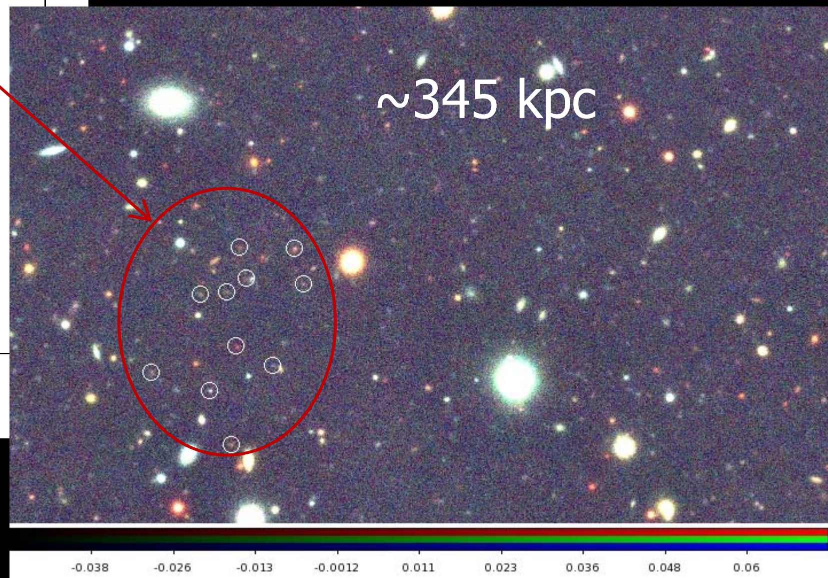
Find 54 groups with z > 1. & N > 10



z=1.06



z=1.19



Real Data Application:
(II) EGS Field & PS1 MD07

(1) EGS Photometric Catalog for input Data

- $0 < z < 1.4$ and $R < 24.1$
- Field of View : ~ 0.5 deg^2
- # of galaxies: 9,594
- Photometric samples
- bands: B, R, I, r $\sigma z/(1+z) \sim 0.03$

(2) PS1 MD04 Galaxy Catalog for input data

- $0 < z < 1.4$ and $r < 24.1$
- Field of View : ~ 7 deg^2
- # of galaxies: 328,119 (PS1_I_CLASS_STAR =0.4)
- Photometric samples
- 6 bands: u, g, r, i, z, y $\sigma z/(1+z) \sim 0.051$

(3) DEEP2 EGS Group Catalog for training set

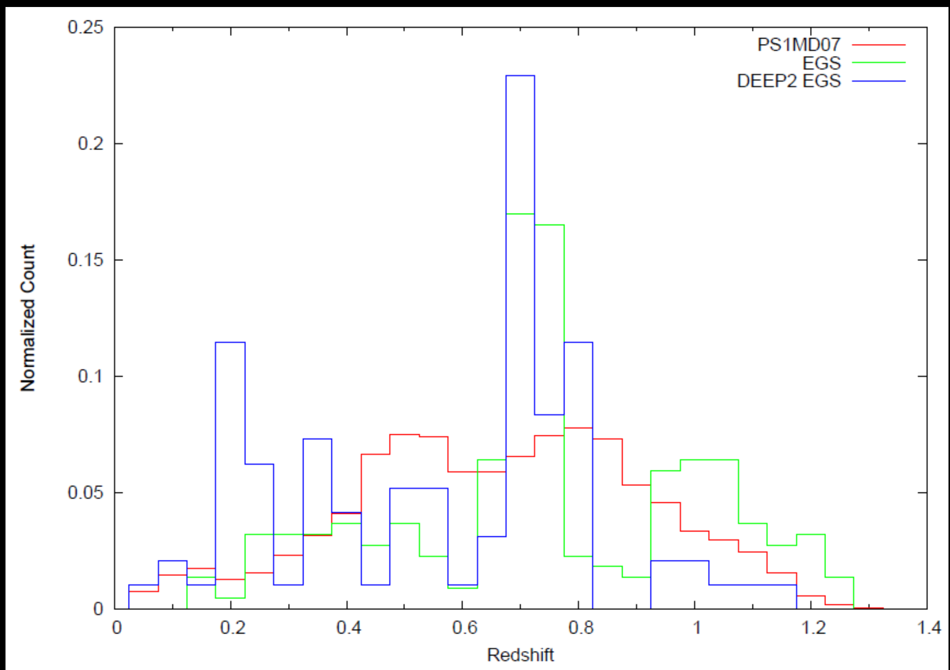
- $0 < z < 1.5$ and $R < 24.1$
- Field of View: ~ 1. deg^2
- # of groups: ~ 1165
- Spectroscopic samples
- the group catalog: Gerke et al., 2012.

Grouping Results:

Optimized with DEEP2 EGS groups

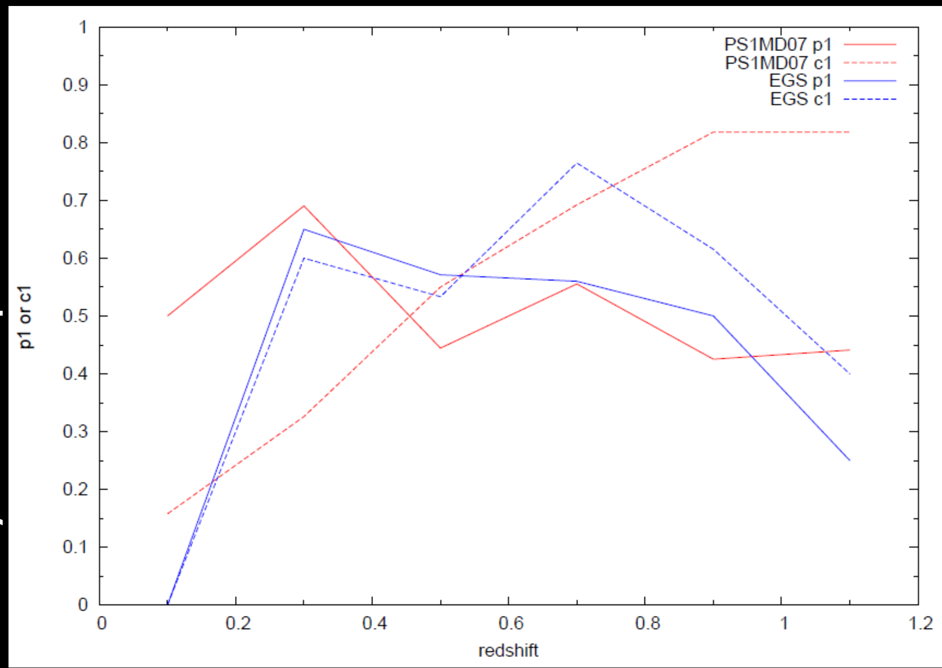
	EGS	PS1 MD07
p1	0.553	0.497
c1	0.615	0.504

Normalized Count



redshift

Purity or Completeness

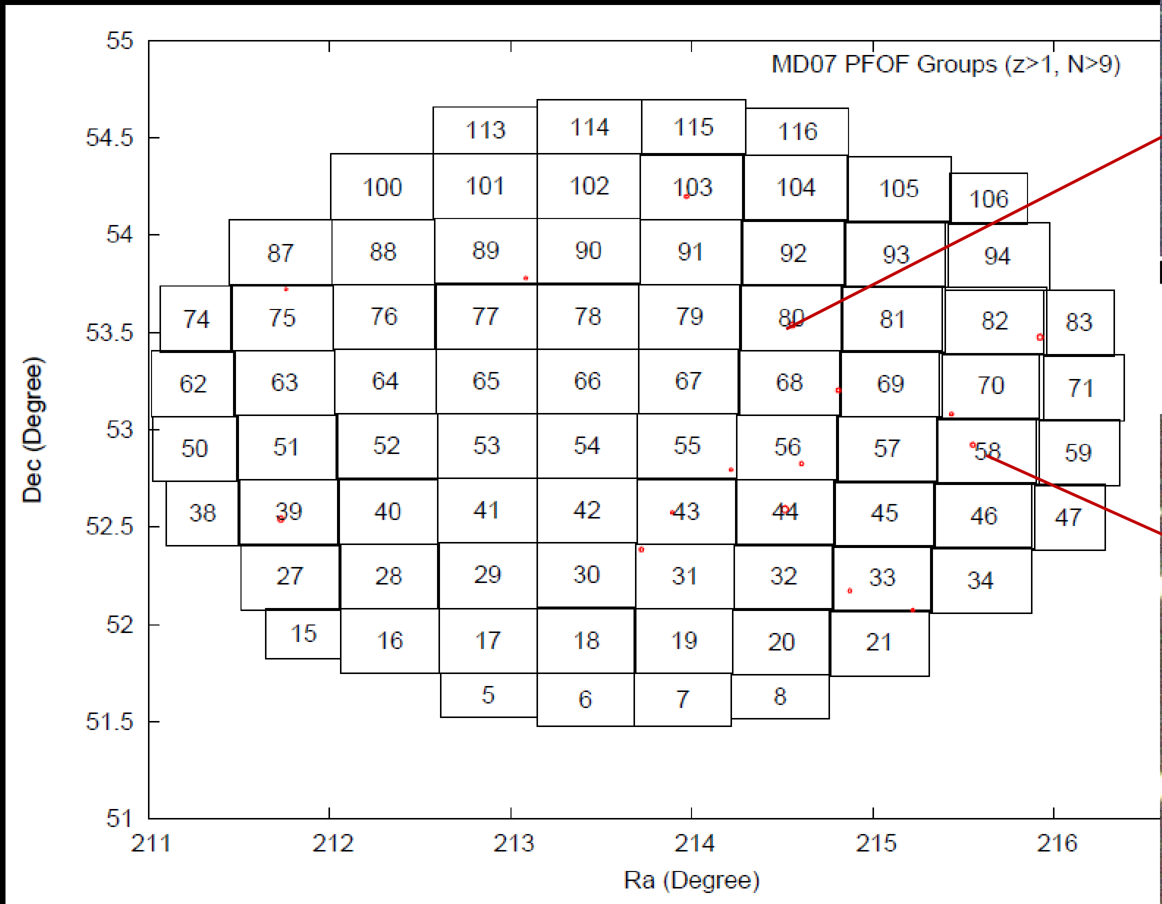


redshift

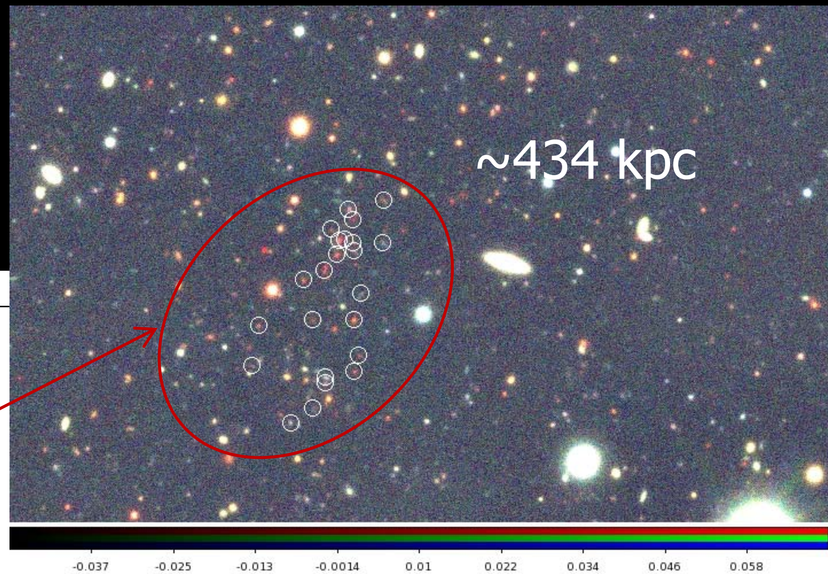
Two Examples of High-z PFOF Groups in MD07

rgb color :
Red -- z
Green -- r
Blue -- g

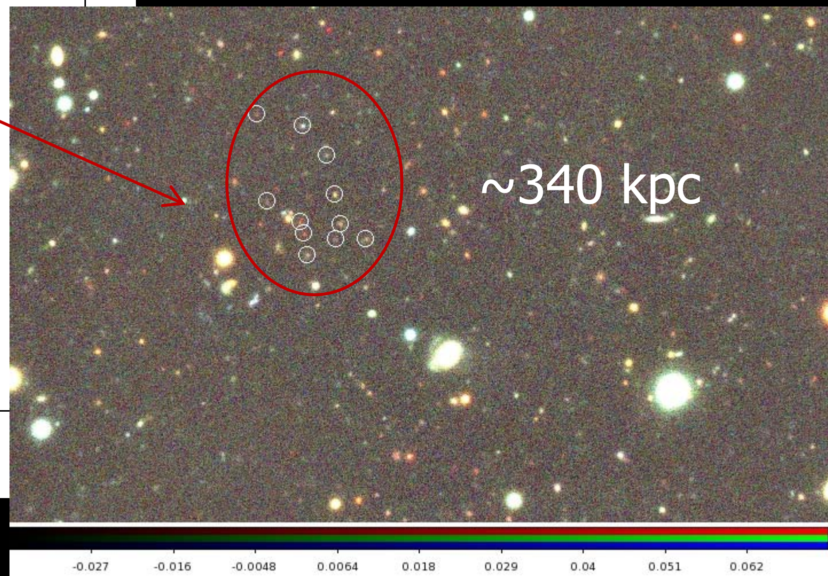
Find 16 groups with z > 1. & N > 9



z=1.14



z=1.12



Summary

- We found:

1. Purity p1 and Completeness c1 remain roughly the same if a catalog is with photo-z uncertainty in the range between 0.03 and 0.07.
2. From the mock test, "subset optimization" is feasible.
3. The x-ray group match fraction is ~60% for all M_h in both COSMOS and PS1MD04 groups when using zCOSMOS groups as the training set.
4. When using X-ray groups as the training set, the match fraction is ~82% in COSMOS and ~88% in PS1MD04.
5. Significance is affected by two factors, number of groups for match and redshift difference constrain.

Future

1. Make detail comparisons with ORCA to know difference between two group finding algorithms.
2. Produce ORCA-PFOF combined MDS-field group catalogs for scientific studies.

Thank You Very Much