Measuring galaxy environments with group finders: Methods & Consequences

Diego STALDER
Doc student
INPE, São Jose dos Campos, Brazil
& IAP

Marina TREVISAN
Postdoc, IAP

Manuel DUARTE
former IAP Doc student

Euge DÍAZ-GIMÉNEZ
IATE, Cordoba, Argentina

Outline

- Motivations of Group Finders
- Review of Group Finders
  - FoF, matched-filter, Voronoi, Yang, MAGGIE …
- Do ≠ Group Finders give ≠ results?
  - surface density & LOS velocity dispersion profiles
  - environmental trends
- Are Group Finders so bad that they blur or bias our knowledge of environmental effects?
- Do group properties strongly depend on Ω_m?

Why are group finders useful?

- Study individual groups

- Statistics of environmental effects on galaxies
  - ★ Galaxy morphology, structure, kinematics, gas & dust content, luminosity & stellar mass functions, fertility, chemistry, …
    = *f* (global environment, local envt, large-scale envt, redshift)

- Cosmological tools
  - ★ evolution of group/cluster mass function
  - ★ velocity fields around groups

Why use Optical group finders?

- X-rays suffer least from projection effects

  X-rays are expensive!
  Difficult to blindly detect low-mass groups

  $L_X \propto T^3 \propto M^2$

- SZ low sensitivity

  $Y \propto M\,T \propto M^{5/3}$

- Lensing least affected by systematics

  Lensing is ~ cheap!
  Difficult to blindly detect low-mass groups

Optical group finders = cheapest way to blindly detect groups!

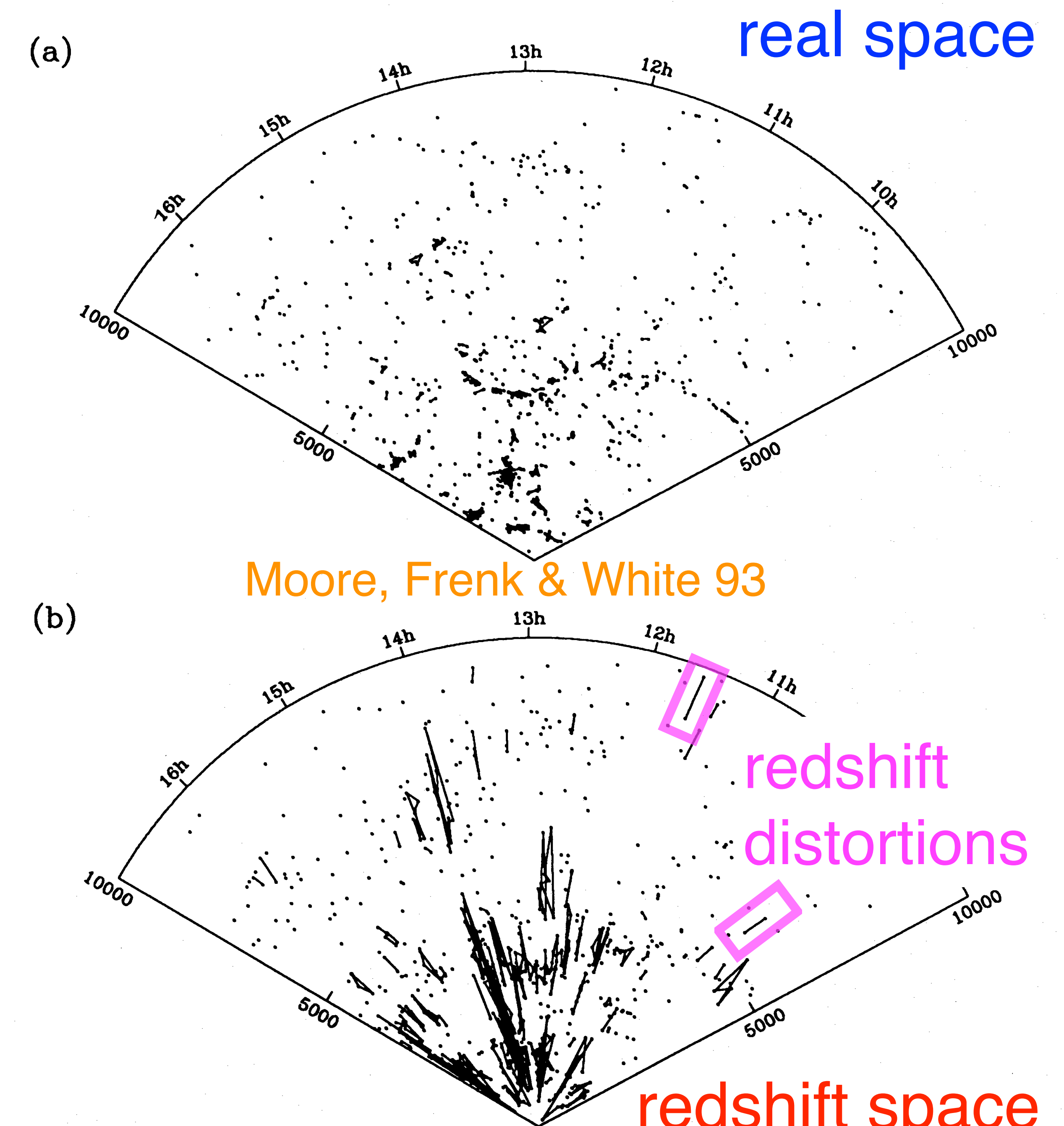
What should group finders provide?

- Positions (centers)
- Mean redshifts
- Group luminosities & stellar masses
- Group total masses (Global Environment)
- Galaxy positions and line-of-sight velocities in group (Local Environment)
- Galaxy membership (Probabilistic?)

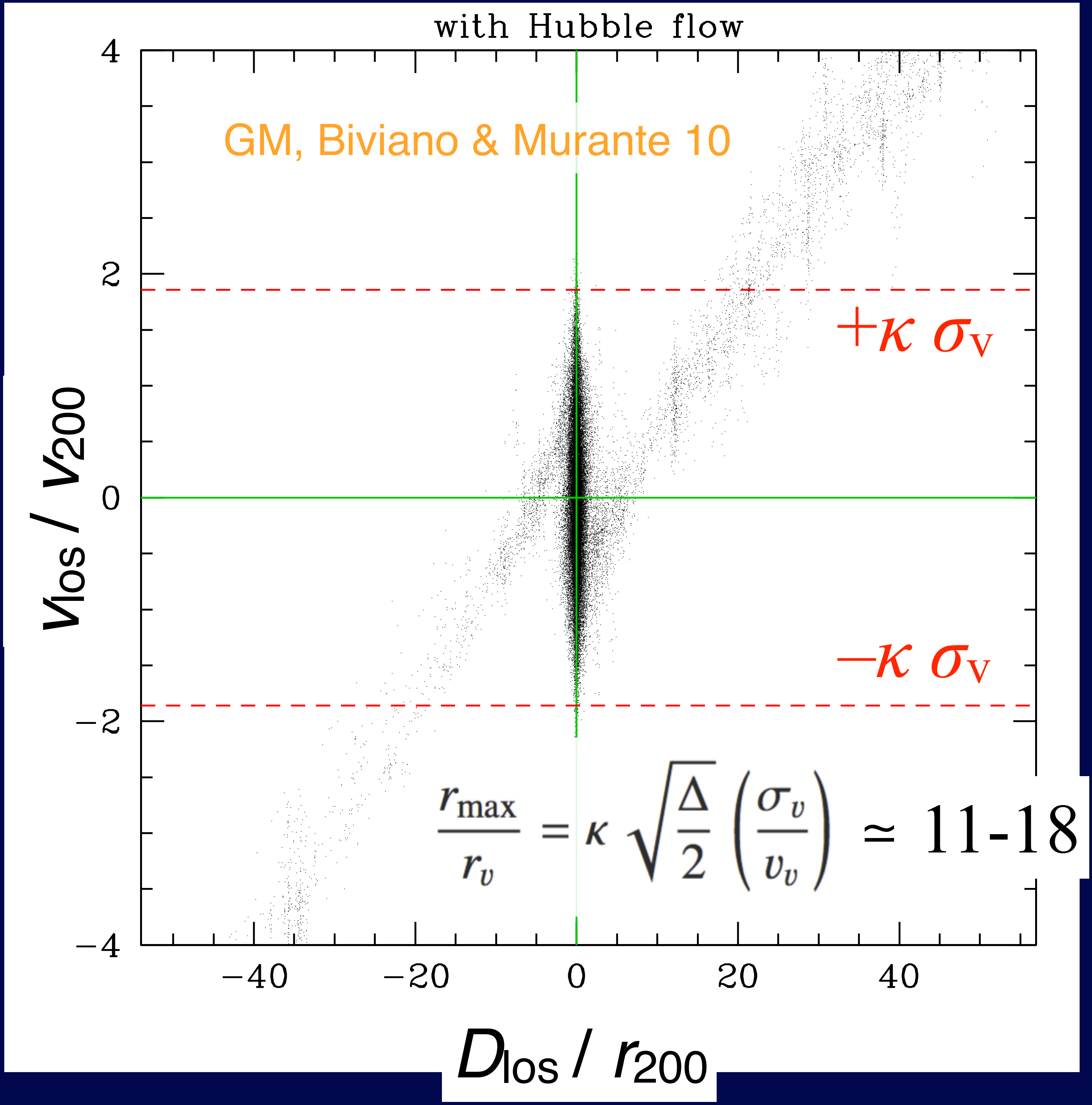
Review of Group Finders

this talk: ~ limited to spectroscopic surveys!

How to extract real-space groups from redshift-space data?



$cz = H_0 D + v_{\rm pec}$



$$\frac{r_{\rm max}}{r_v} = \kappa \sqrt{\frac{\Delta}{2}} \left(\frac{\sigma_v}{v_v}\right) \simeq 11\text{-}18$$

Δ = overdensity / critical density

Group finders for spectroscopic galaxy samples

incomplete list!

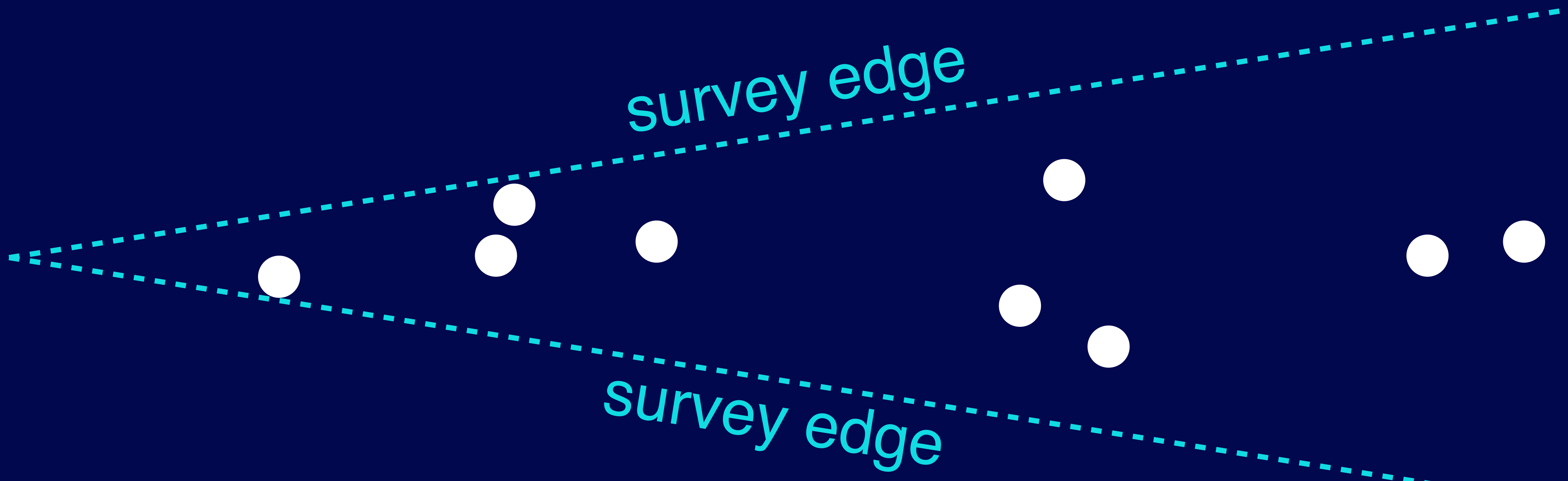
- *Frequentist*

Friends-of-Friends	Huchra & Geller 82
Voronoi Tessellation	Marinoni+02
Dendrograms	Tully 87

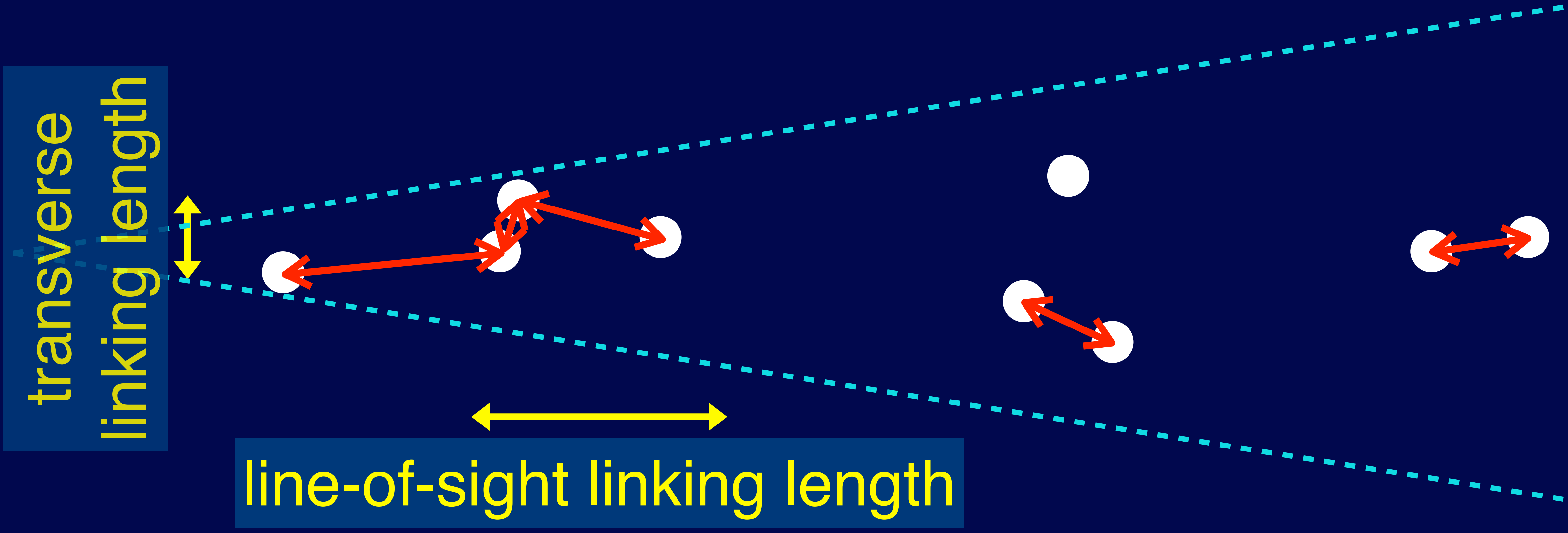
- *Prior-based*

Matched Filter	Kepner+99
Yang	Yang, Mo, van den Bosch +05, 07
MAGGIE	Duarte & Mamon 15

Friends of Friends (FoF)



Friends of Friends (FoF)



transverse linking length

line-of-sight linking length

Dimensionless linking lengths in terms of mean nearest neighbor separation: $b = \mathrm{LL}/\langle n(z)\rangle^{-1/3}$

Optimal FoF linking lengths

Duarte & Mamon 14

for Δ=200 & Ω_m=0.25

mean transverse link

$$\frac{\delta n}{n} = \frac{3}{4\pi b_\perp^3} - 1$$

$$b_\perp = \left(\frac{3/(4\pi)}{\Delta/\Omega_\mathrm{m} + 1}\right)^{1/3} = 0.07$$

max (95% c.l.) transverse link

$$b_\perp = \frac{\mathrm{Max}(S_\perp)}{n^{-1/3}} = \left(\frac{3/(4\pi)}{\Delta/\Omega_\mathrm{m} + 1}\right)^{1/3} \frac{\mathrm{Max}(S_\perp)}{r_\mathrm{vir}} N_\mathrm{vir}^{1/3} \simeq 0.09\, N^{0.08}$$

$b_\perp = 0.10$ for $N = 4$ and $b_\perp = 0.12$ for $N = 40$

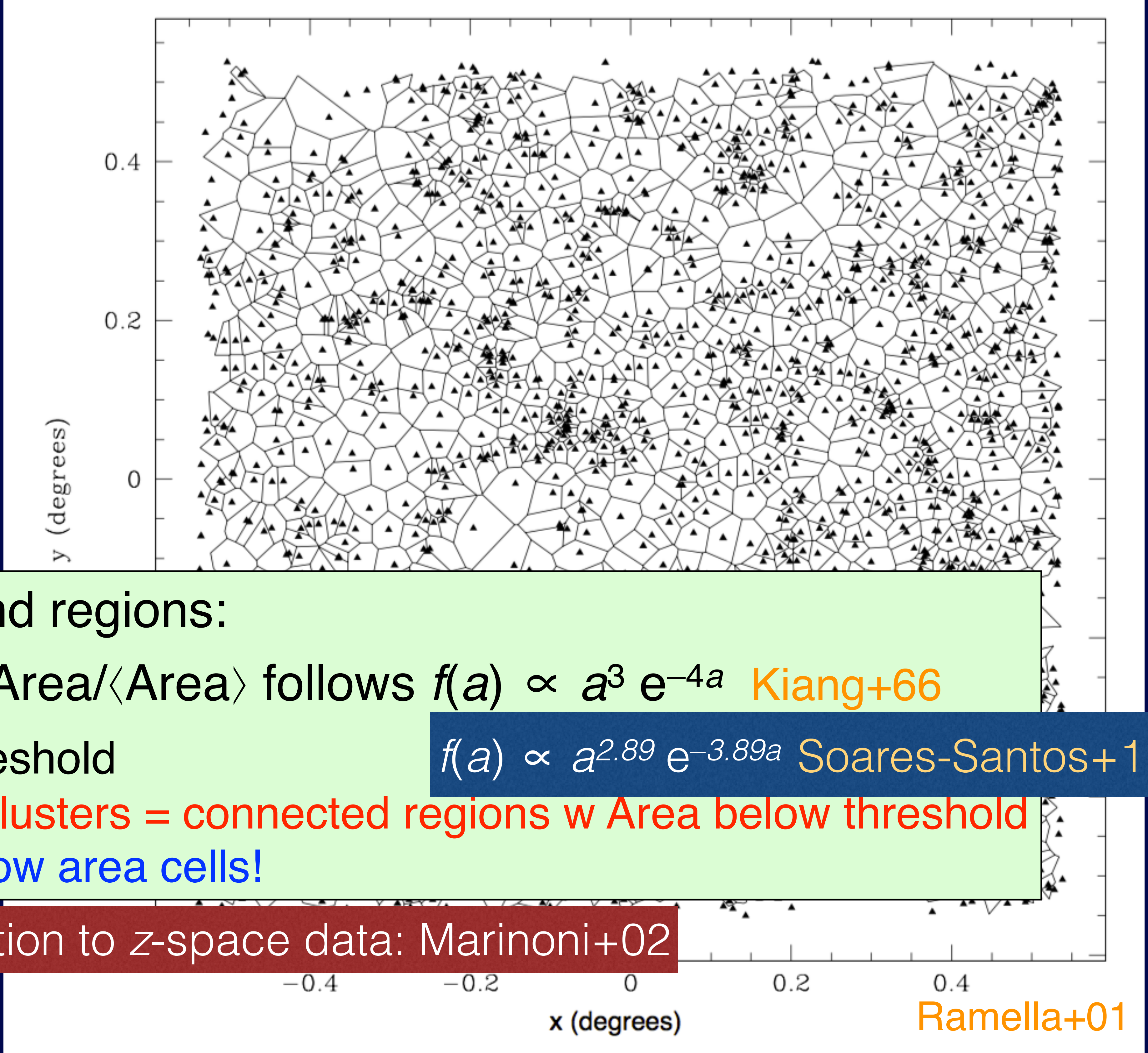
mean line-of-sight link

$$\frac{b_\parallel}{b_\perp} = \left(\frac{v_\mathrm{max}}{\sigma_\mathrm{v}}\right) \left(\frac{\sigma_\mathrm{v}}{v_\mathrm{vir}}\right) \sqrt{\frac{\Delta}{2}} \simeq 11$$

$$\Rightarrow b_\parallel = 1.1$$

for v_{max}/σ_v=1.65 (95%)

Voronoi tessellation



Ramella+01

Background regions:

expect a=Area/⟨Area⟩ follows $f(a) \propto a^3 e^{-4a}$ Kiang+66

→ Area threshold

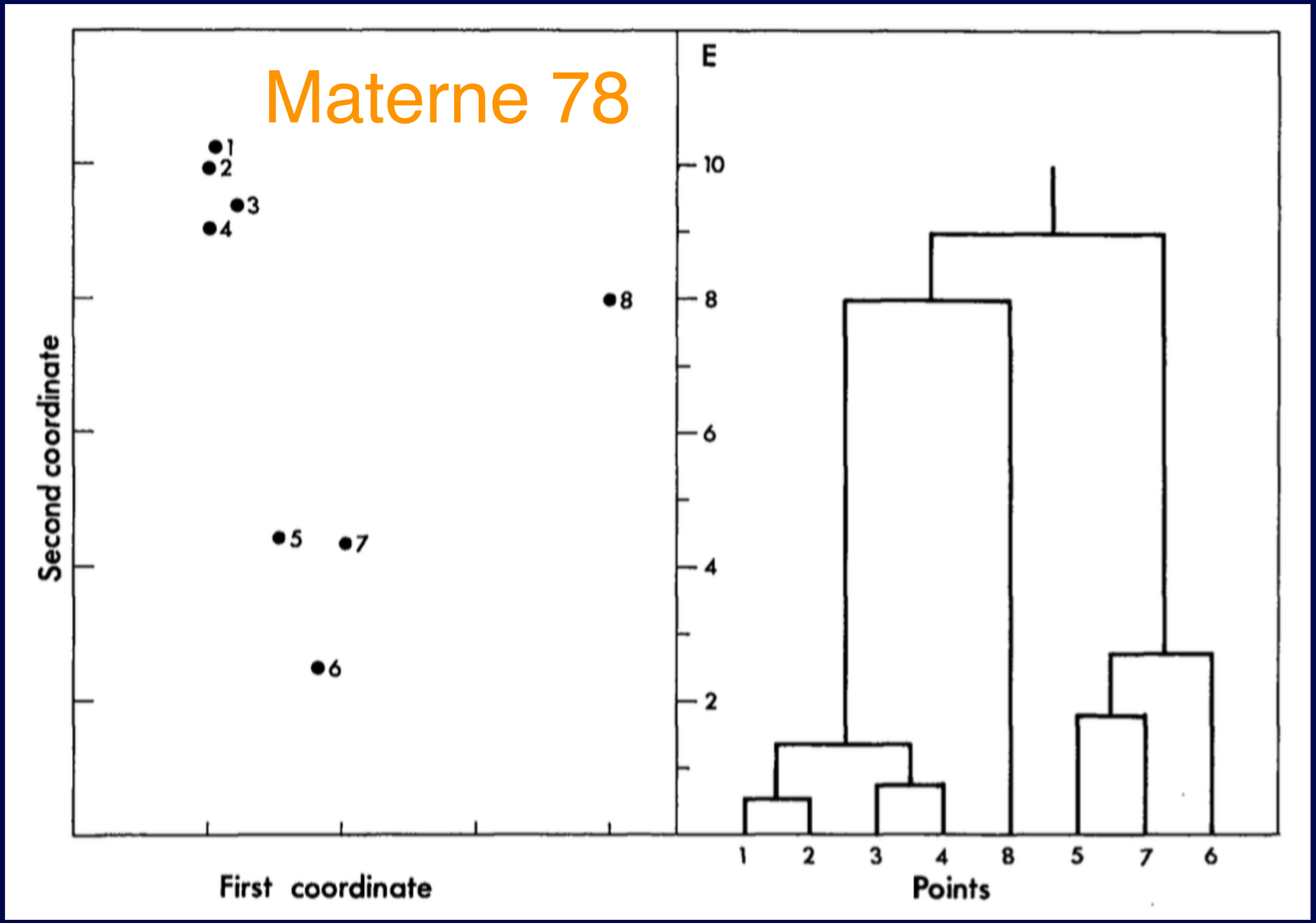
$f(a) \propto a^{2.89} e^{-3.89a}$ Soares-Santos+11

→ groups/clusters = connected regions w Area below threshold

⇔ FoF on low area cells!

1st application to z-space data: Marinoni+02

Dendrograms



link pairs by values of max(L_i,L_j)/R_{ij}^3 Tully 87

Matched filter

Postman+96 (2D)   Kepner+99 (2D,2+½D,3D)

Convolve data with filter using

- position (prior on surface density profile)
- redshift (Gaussian prior on distribution of $v_{\rm LOS}$)
  - or magnitudes (LF prior) or
  - or photo-zs (Gaussian prior)

Yang et al.'s Halo-based Group Finder

$$g(R, v_z) = \Sigma_{\rm NFW}(R)\,\exp\left(-\frac{v_z^2}{2\,\sigma_{\rm LOS}^2}\right) > 10\,\frac{c\,\rho_{\rm Univ}}{H_0}$$

Yang, Mo & van den Bosch 04; Yang+07
Domínguez Romero, García Lambas & Muriel 12

group masses (hence virial radii) from:
- FoF group luminosities (1st pass: M=300L)
- *Halo Abundance Matching* (next passes)

Accurate group masses (global environment), BCG at center (local environment)

weaknesses

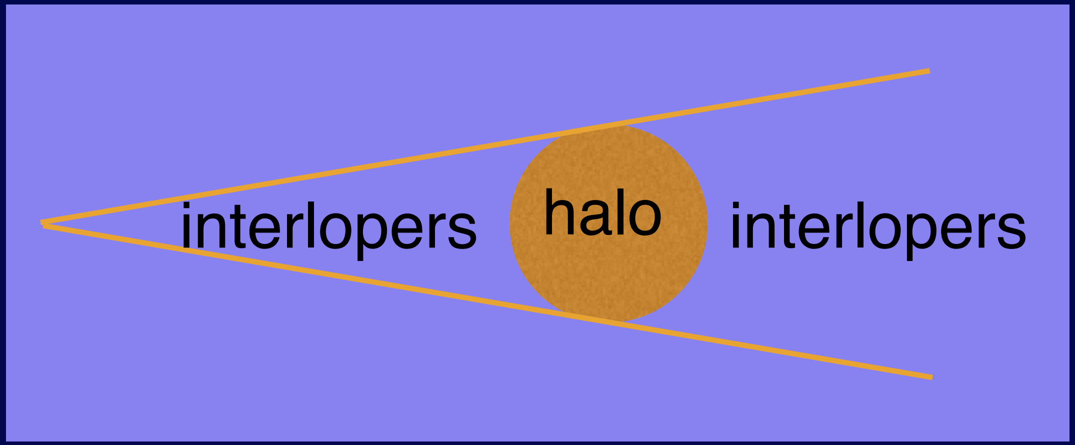
- LOS velocity dispersion profile should be convex in log-log (not cst)
- LOS velocity distributions not Maxwellian (outer radial vel. anisotropy)
- ad hoc threshold for membership (10)
- imprecise correction for lum. incompleteness (for SDSS flux-limited sample)
- hard group assignment is unstable

MAGGIE: *Models & Algorithms for Galaxy Groups, Interlopers & Environment*

Duarte & Mamon 15

probabilistic

$$P(R, v_z) = \frac{g_{\rm halo}(R, v_z)}{g_{\rm halo}(R, v_z) + g_{\rm ilop}(R, v_z)}$$



more realistic $g_{\rm halo}$ from ΛCDM 3D model with anisotropic velocities

$$g_{\rm h}(R, v_z) = \Sigma^{\rm NFW}_{\rm sph}(R)\ \langle h(v_z|R, r)\rangle_{\rm LOS-sph}$$

$$= 2 \int_R^{r_{200}} \nu(r)\, h(v_z|R, r)\, \frac{r\,{\rm d}r}{\sqrt{r^2 - R^2}}$$

$$h(v_z|R, r) = \frac{1}{\sqrt{2\pi\sigma_z^2(R, r)}} \exp\left[-\frac{v_z^2}{2\,\sigma_z^2(R, r)}\right]$$

$$\sigma_z^2(R, r) = \left(1 - \beta(r)\frac{R^2}{r^2}\right)\ \sigma_{\rm r}^2(r)$$

$\sigma_r(r)$ from solving Jeans equation

$\beta(r)$ from cosmo simulations

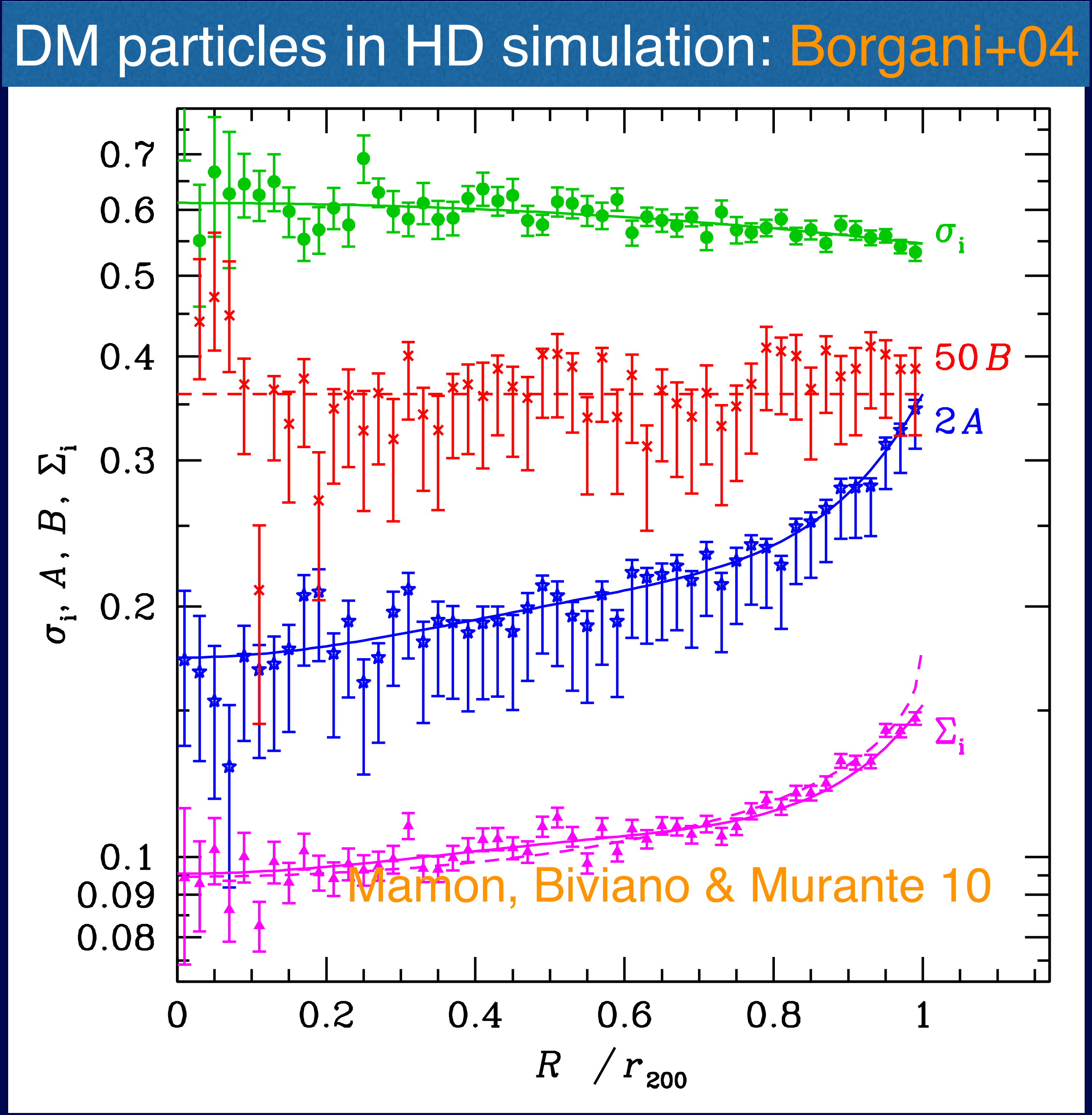
MAGGIE: interlopers

Duarte & Mamon 15

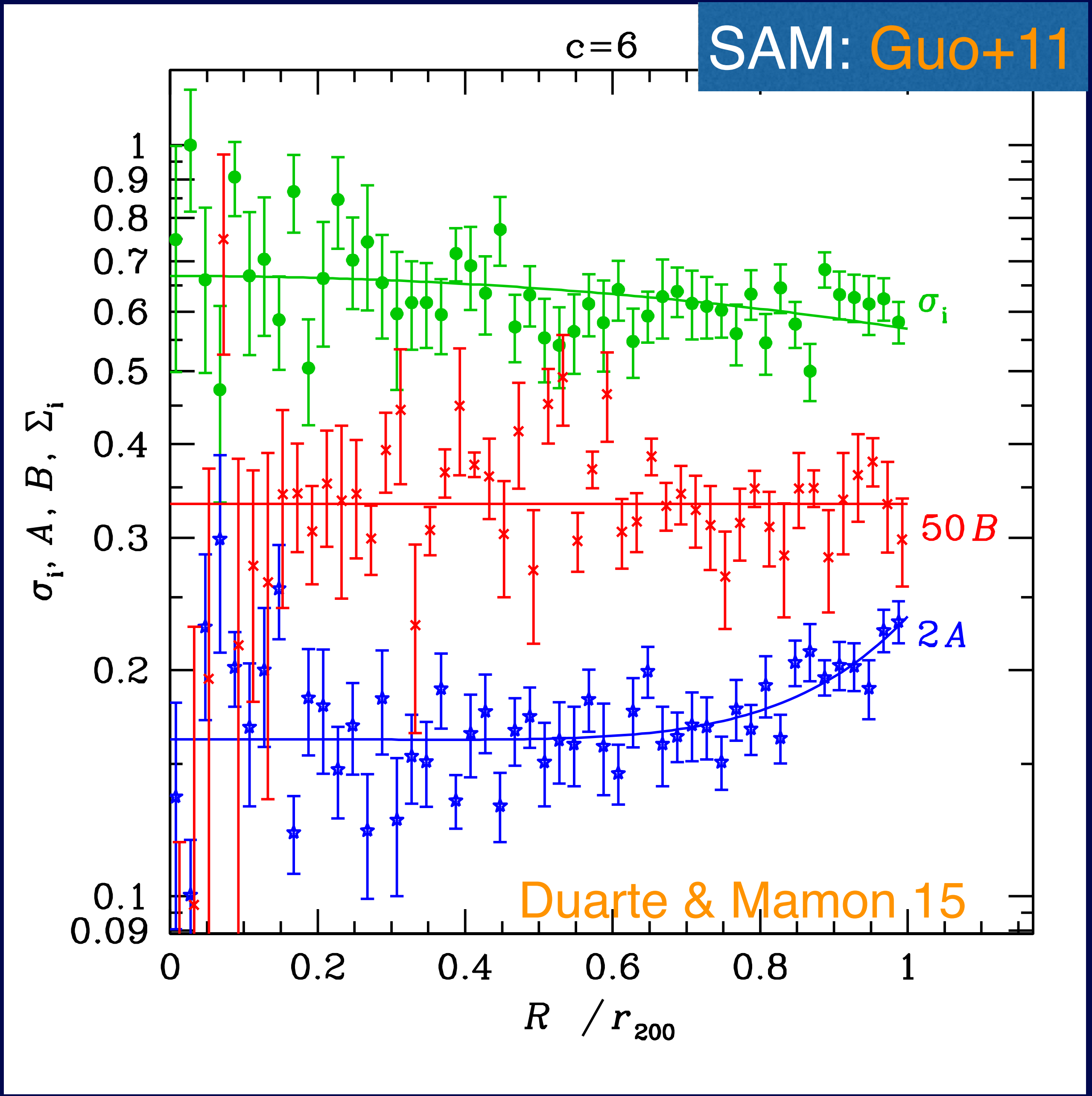
$$P(R,v_z) = \frac{g_{\rm halo}(R,v_z)}{g_{\rm halo}(R,v_z) + g_{\rm ilop}(R,v_z)}$$

$$g_{\rm i}(R,\,v_z) = \frac{N_{200}}{r_{200}^2 v_{200}}\,\widehat{g}_{\rm i}\left(\frac{R}{r_{200}},\,\frac{v_z}{v_{200}}\right)$$

$$\widehat{g}_{\rm i}(X,u) = A(X)\,\exp\left[-\frac{1}{2}\,\frac{u^2}{\widehat{\sigma}_{\rm i}^2(X)}\right] + B$$


DM particles in HD simulation: Borgani+04

Mamon, Biviano & Murante 10


SAM: Guo+11

Duarte & Mamon 15

Mamon & Duarte 15

MAGGIE: Models & Algorithms for Galaxy Groups, Interlopers & Environment

- group masses by Halo Abundance Matching
    - on central galaxy luminosity or stellar mass (1st pass)
    - on total group luminosity or stellar mass (next passes)
- groups extracted from *D*- & *L*-complete subsamples
- group properties = sums weighted by probabilities

Testing Group Finders

How can group finders go wrong?

- group fragmentation

→ secondary fragments bring down group purity

→ reduced galaxy completeness

- group merging

→ reduced group completeness

→ reduced purity of galaxy membership

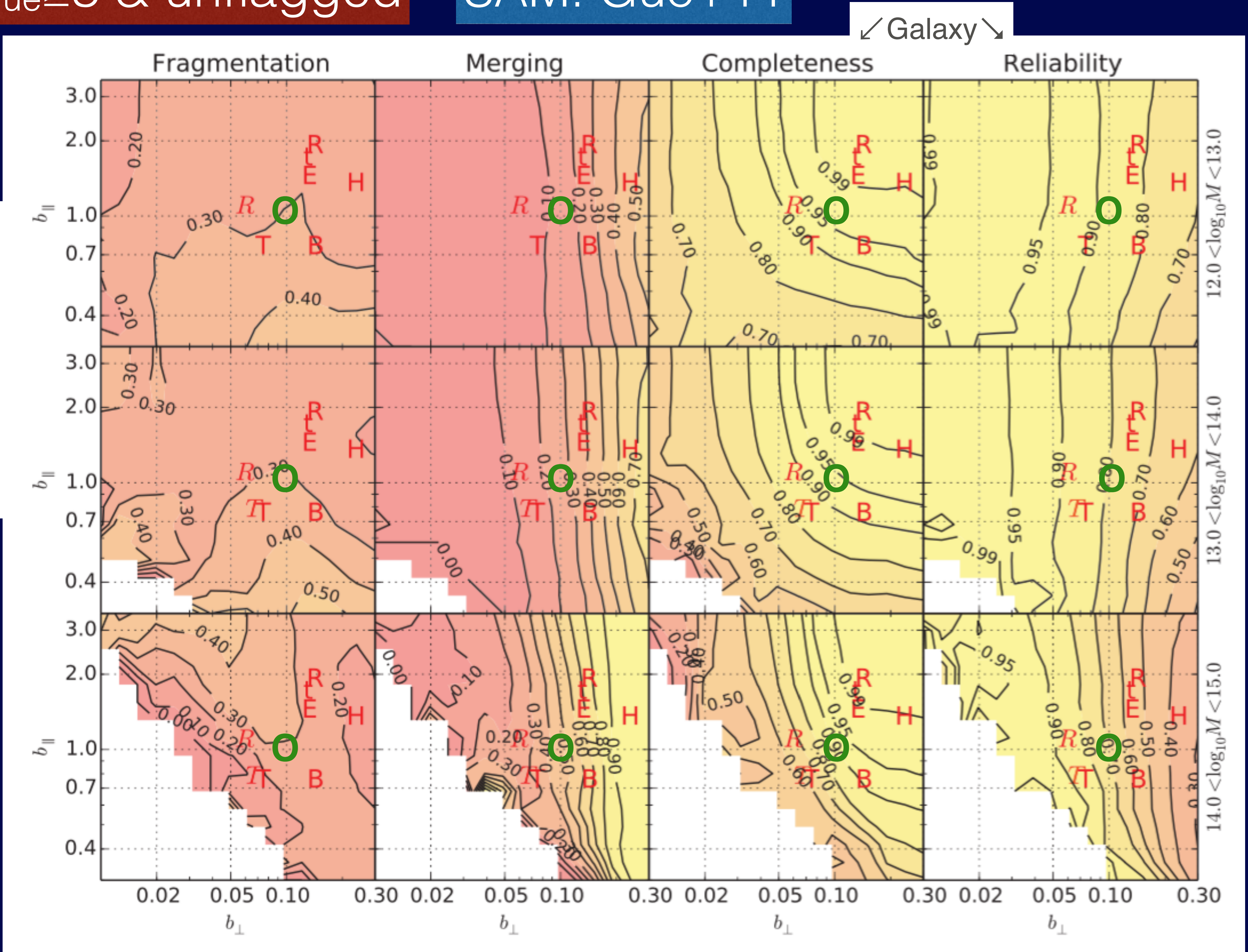
Friends-of-Friends optimization

Duarte & Mamon 14

$N_{\mathrm{est}} \geq 3$ & $N_{\mathrm{true}} \geq 3$ & unflagged

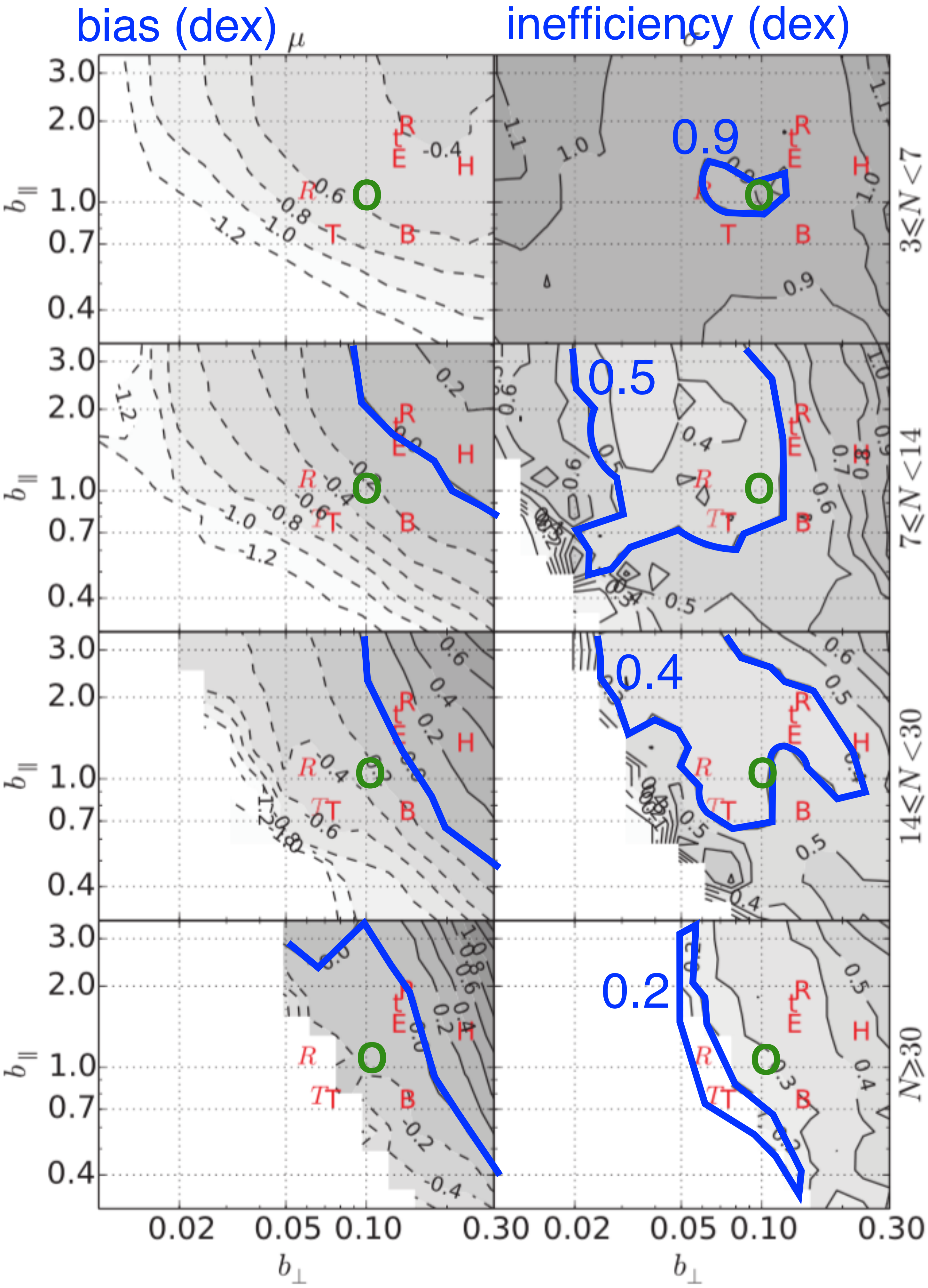
SAM: Guo+11

- H: Huchra & Geller 82
- R: Ramella_89
- t: Trasarti-Battistoni 98
- E: Eke+04
- B: Berlind+06
- T: Tago+10
- *R*: Robotham+11
- *T*: Tempel+14
- **O**: optimal (theoretically)



FoF optimization: mass accuracy

Duarte & Mamon 14



H: Huchra & Geller 82
R: Ramella_89
t: Trasarti-Battistoni 98
E: Eke+04
B: Berlind+06
T: Tago+10
R: Robotham+11
T: Tempel+14
O: optimal (theoretically)

Best compromise is:
→ $b_\perp = 0.07$ & $b_{//} = 1.1$
≃ theoretical for mean separation
≃ Robotham+11

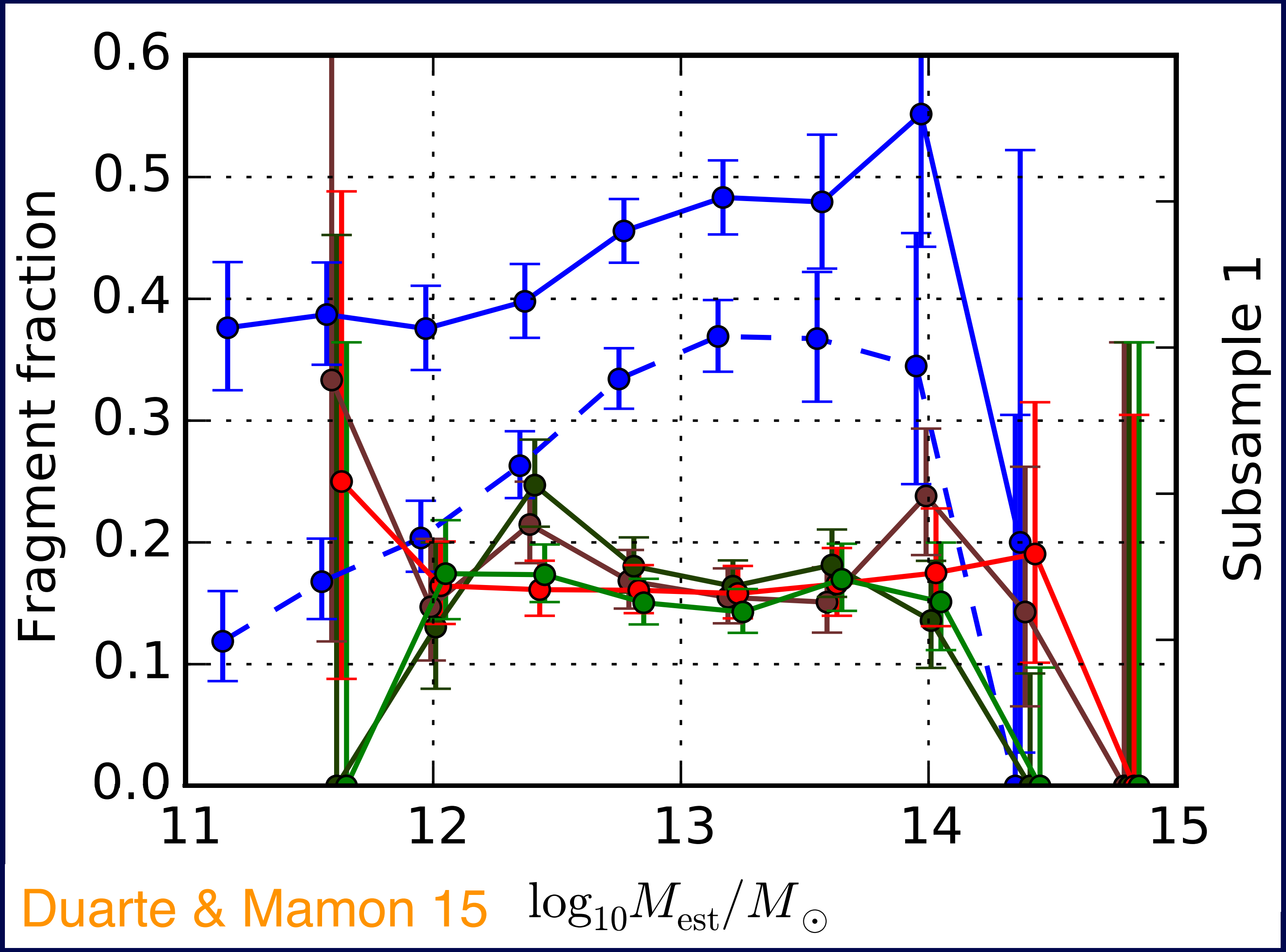
Tests: Group Fragmentation

mocks SDSS galaxy catalog with errors on luminosities (0.08 dex) & stellar masses (0.2 dex)

matching extracted & true groups by most luminous (L) or massive in stars (M) member

only unflagged groups $N_{\mathrm{true}} \geqq 3$ & $N_{\mathrm{est}} \geqq 3$

fraction of extracted groups = secondary fragments of true groups



Duarte & Mamon 15

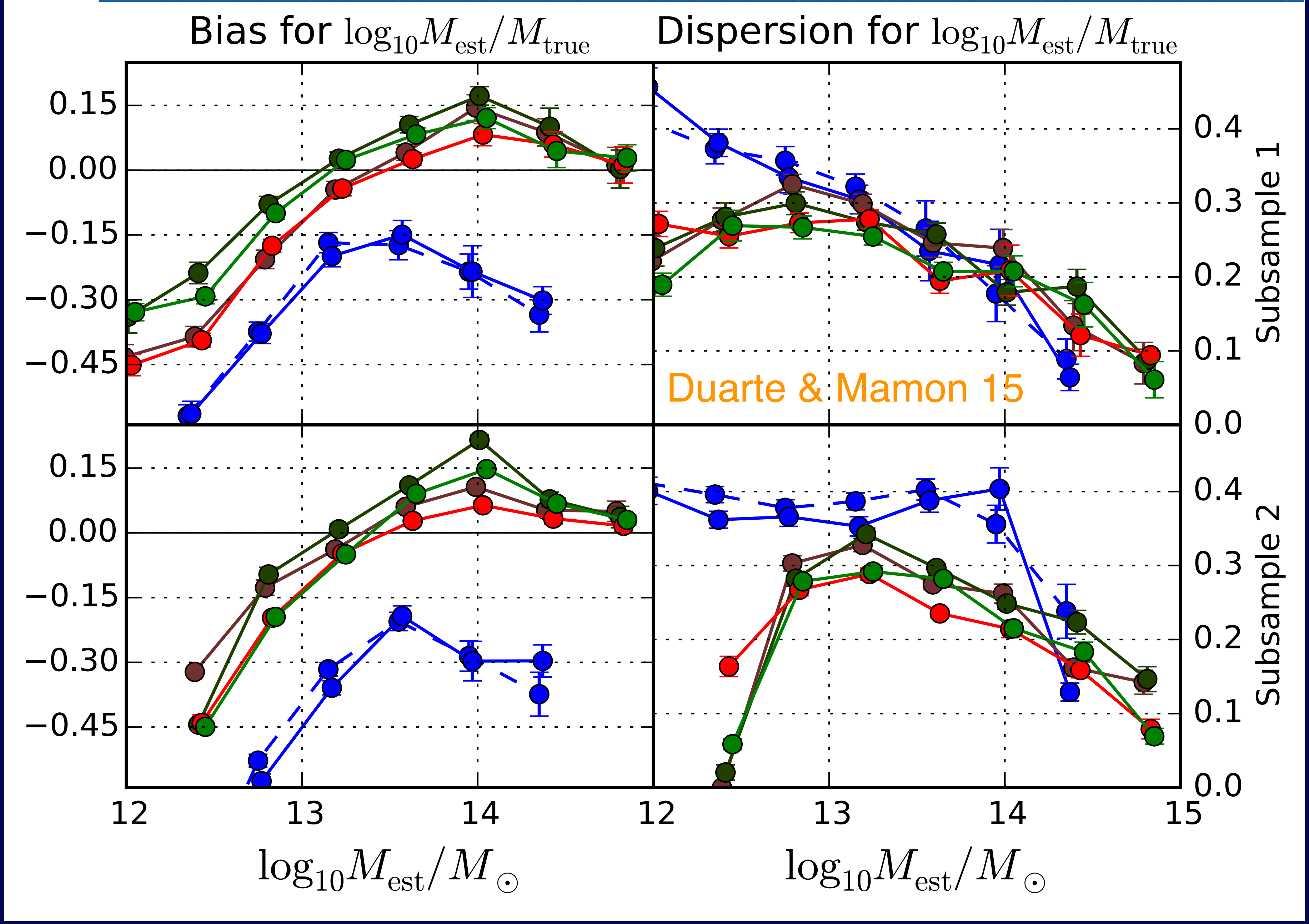
FoF-M (solid)
FoF-L (dashed)
Yang-M
Yang-L
MAGGIE-M
MAGGIE-L

FoF clusters: high probability of being secondary fragment!

M-based: more physical than L-based, but higher (systematic) errors

Group total mass accuracy

only unflagged primary-fragment groups $N_{\rm true} \geqq 3$ & $N_{\rm est} \geqq 3$



FoF-M (solid)
FoF-L (dashed)
Yang-M
Yang-L
MAGGIE-M
MAGGIE-L

FoF masses biased low by 0.15 to 0.5 dex, 0.3 dex at hi mass

mass accuracy (dex)

@log M = 13: 0.35 (FoF), 0.32 (Yang), 0.28 (MAGGIE)

@log M = 14: 0.2-0.4 (FoF), 0.23 (Yang), 0.20 (MAGGIE)

Euclid Cluster Finders

Euclid:

- deep
- mainly based on photo-zs

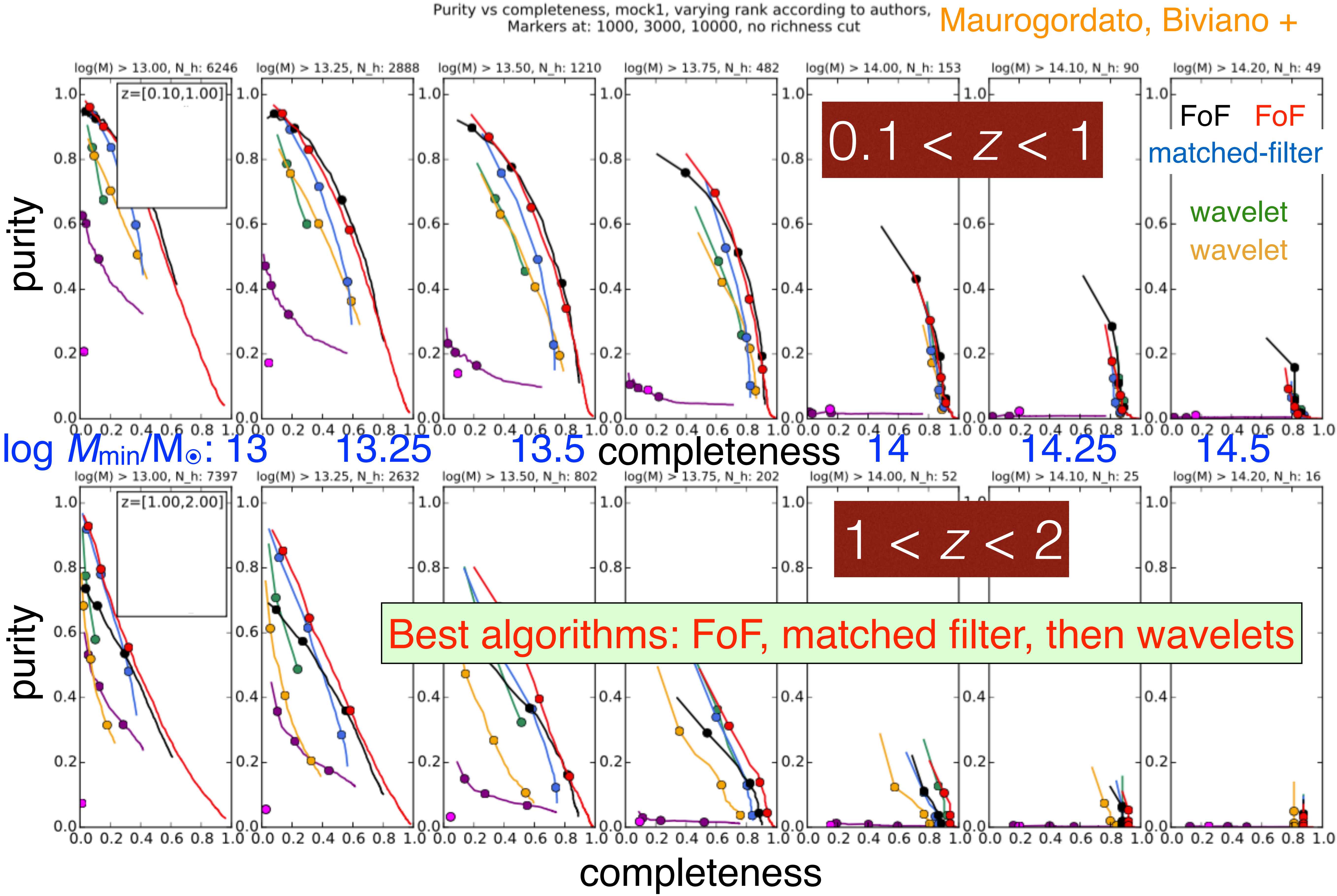
Euclid Cluster Finder Challenge (4 versions)

Maurogordato & Biviano

8 algorithms on mock galaxy catalogs (SAM & HOD)
with photo-*z* errs few galaxies will have spec-*z*s

Cluster Finder Challenge 3 on Durham SAM mock

Maurogordato, Biviano +



0.1 < z < 1

FoF FoF

matched-filter

wavelet

wavelet

purity

log $M_{\rm min}/M_\odot$: 13 13.25 13.5 14 14.25 14.5

completeness

1 < z < 2

purity

Best algorithms: FoF, matched filter, then wavelets

completeness

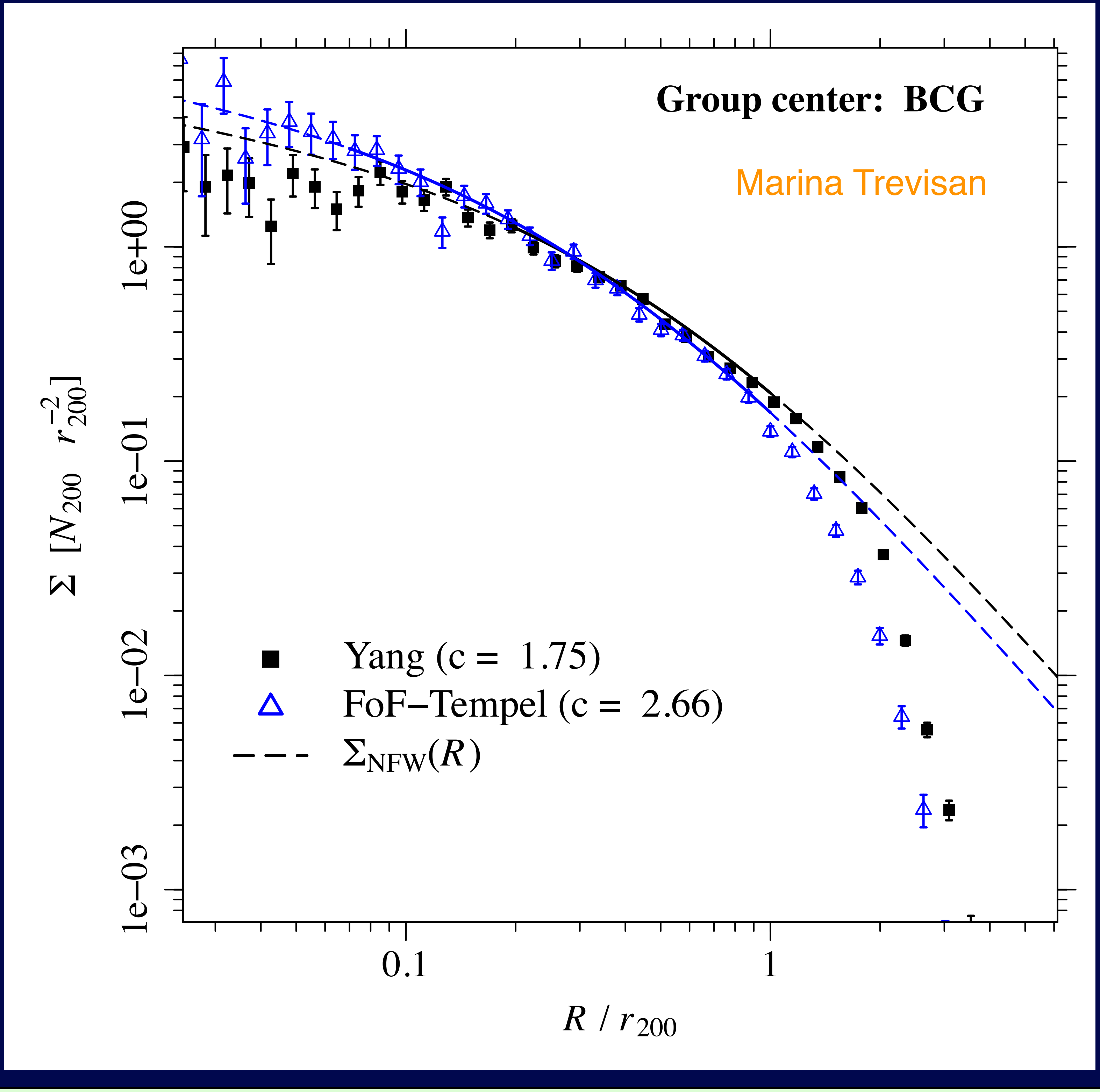
Group properties vs. group finder

Surface density profiles of SDSS groups

$N \geqq 5$

$\log M/\mathrm{M}_\odot > 13.1$

$M_r < -19$



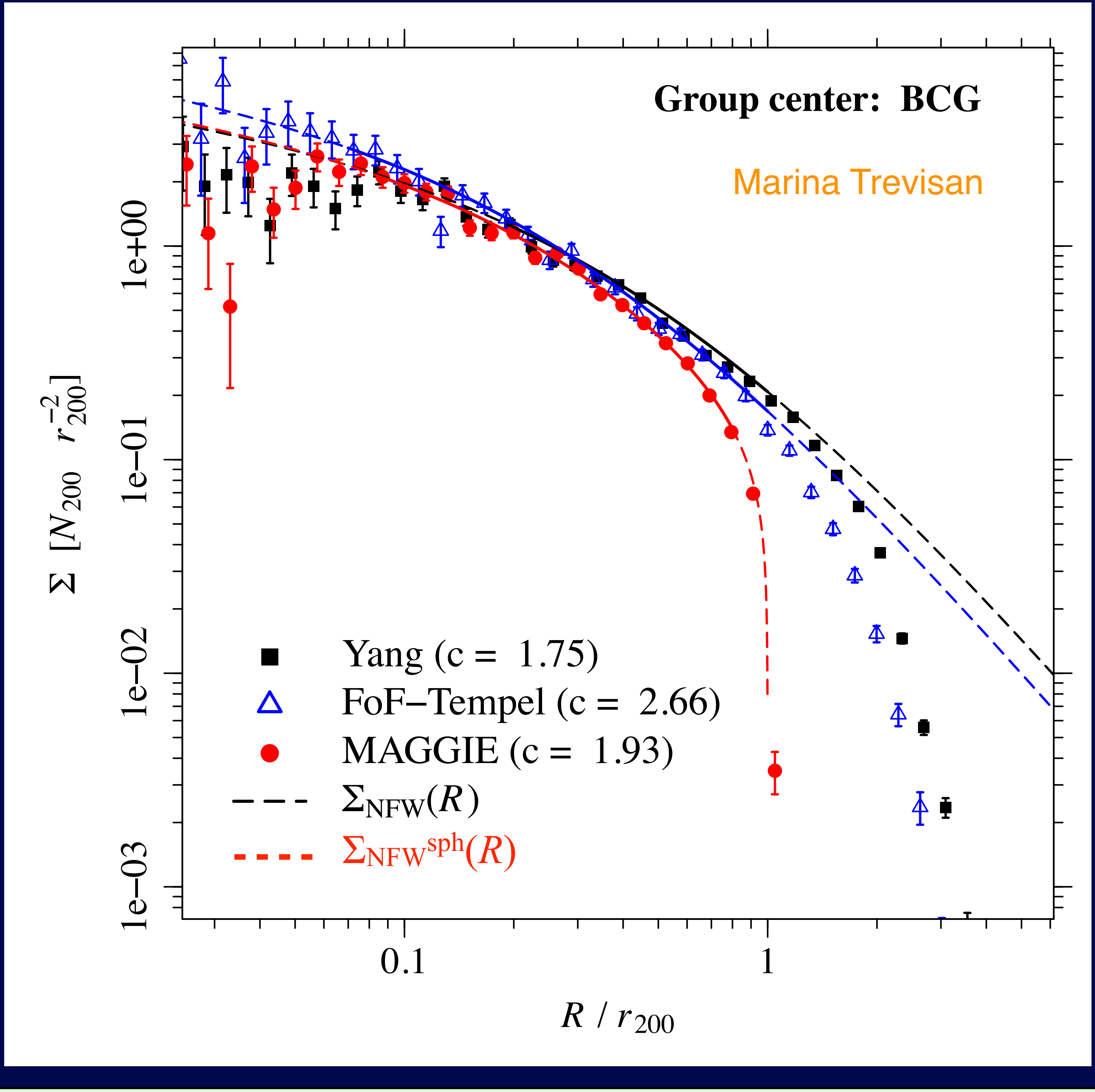
FoF & Yang consistent with NFW for 0.04 (F) or 0.08 (Y) < R/r_{200} < 1

Surface density profiles of SDSS groups

$N \geqq 5$

$\log M/M_\odot > 13.1$

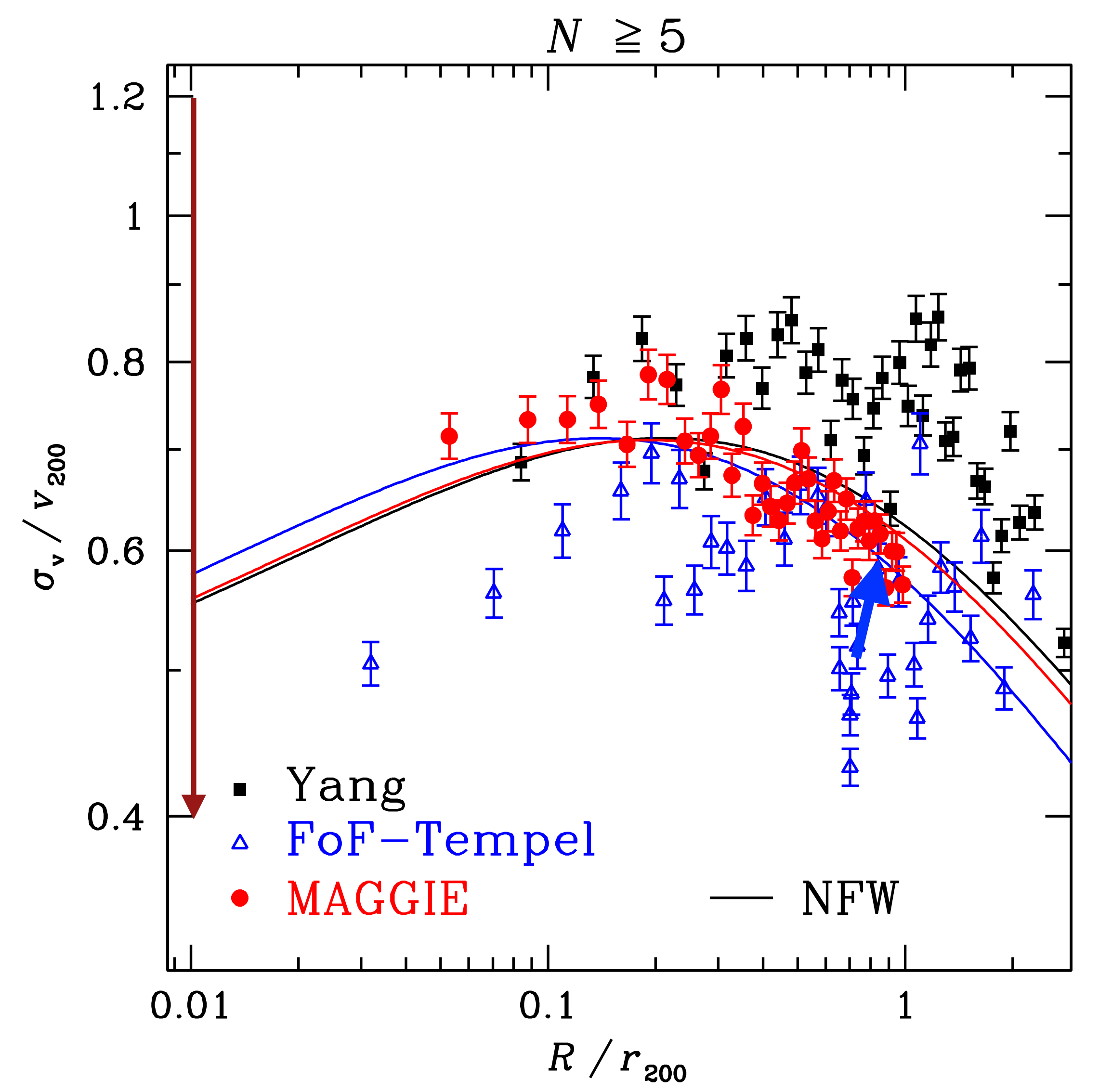
$M_r < -19$

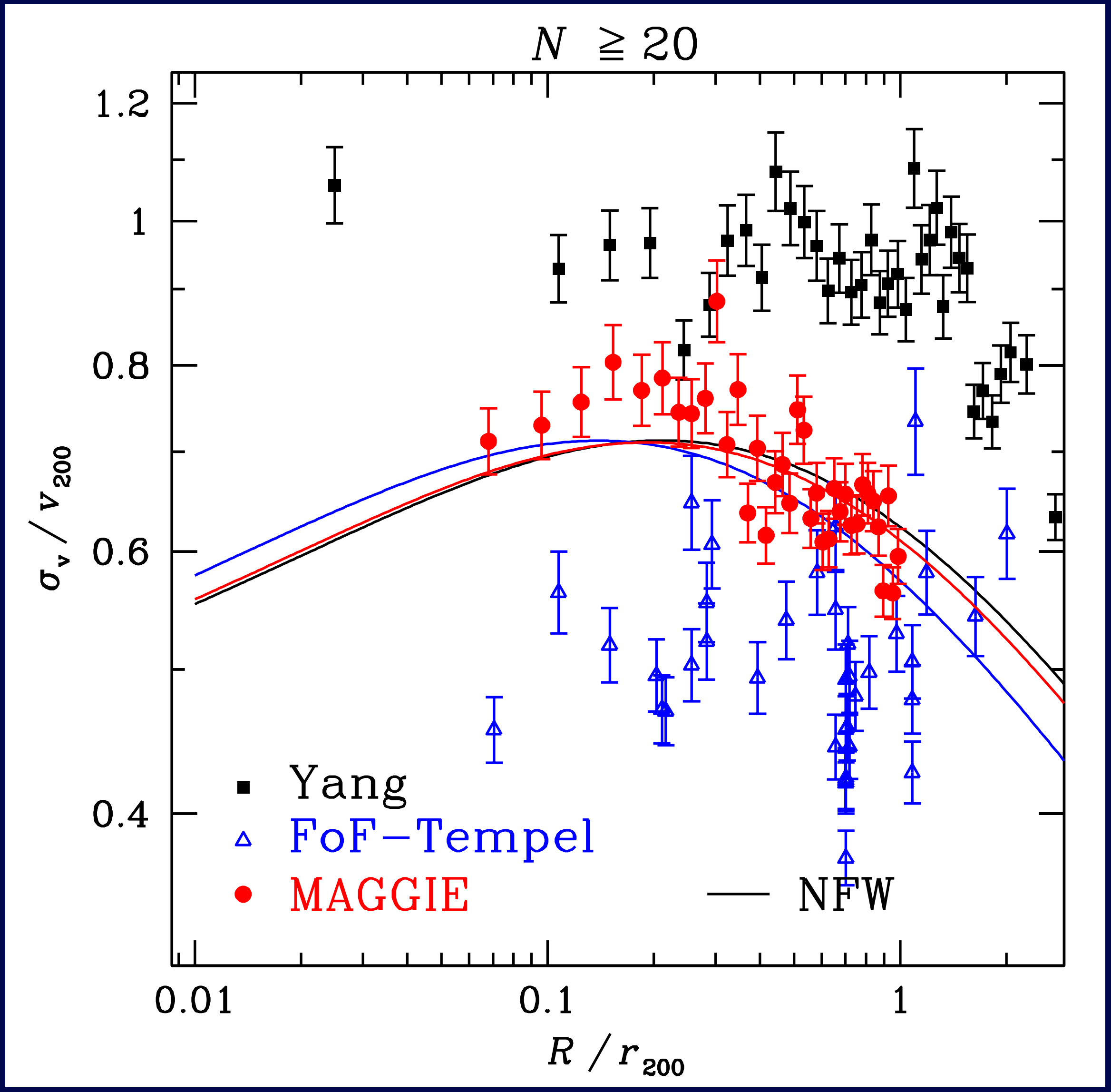


FoF & Yang consistent w NFW for 0.04 (F) or 0.08 (Y) < R/r_{200} < 1(F) or 1.2 (Y)

MAGGIE consistent with NFW-*in-sphere* for 0.05 < R/r_{200} < 1

Line-of-sight velocity dispersion profiles





returned σ_v/v_{200}:

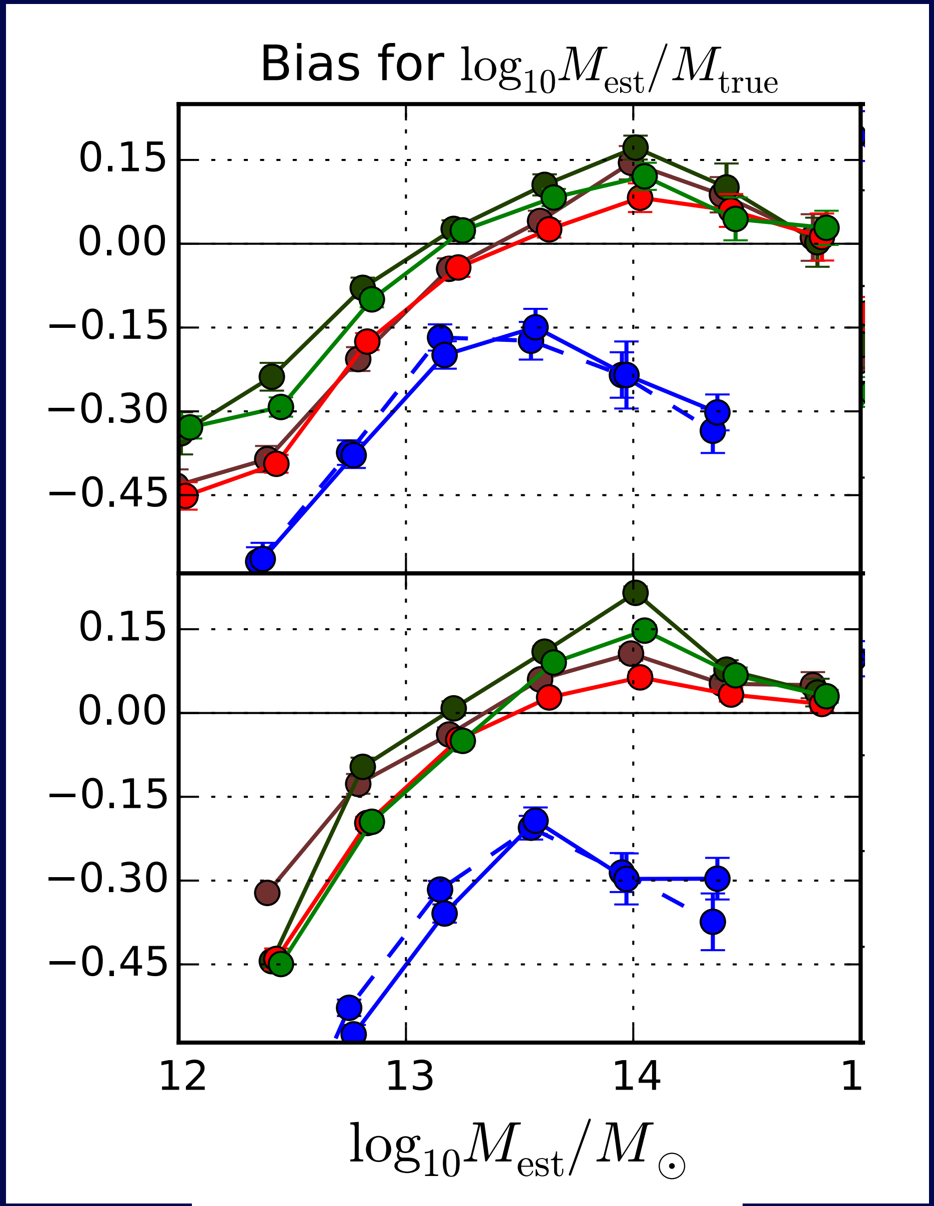
➡ Yang: too high

➡ FoF: too low

➡ MAGGIE: just right

At high richness: Yang & FoF get worse!

Line-of-sight velocity dispersion profiles



FoF-M (solid)

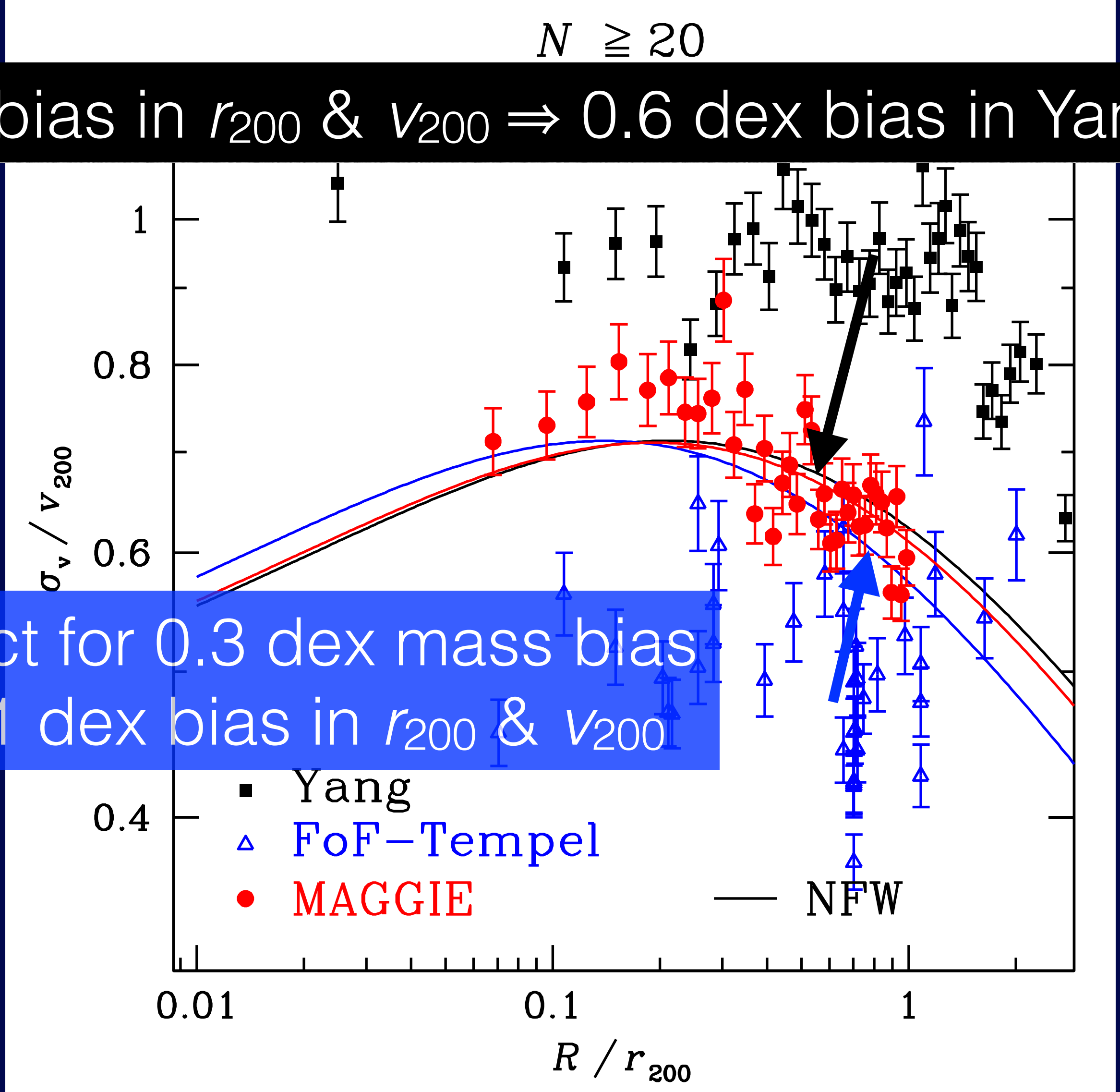
FoF-L (dashed)

Yang-M

Yang-L

MAGGIE-M

MAGGIE-L



0.2 dex bias in r_{200} & v_{200} ⇒ 0.6 dex bias in Yang!

correct for 0.3 dex mass bias
= 0.1 dex bias in r_{200} & v_{200}

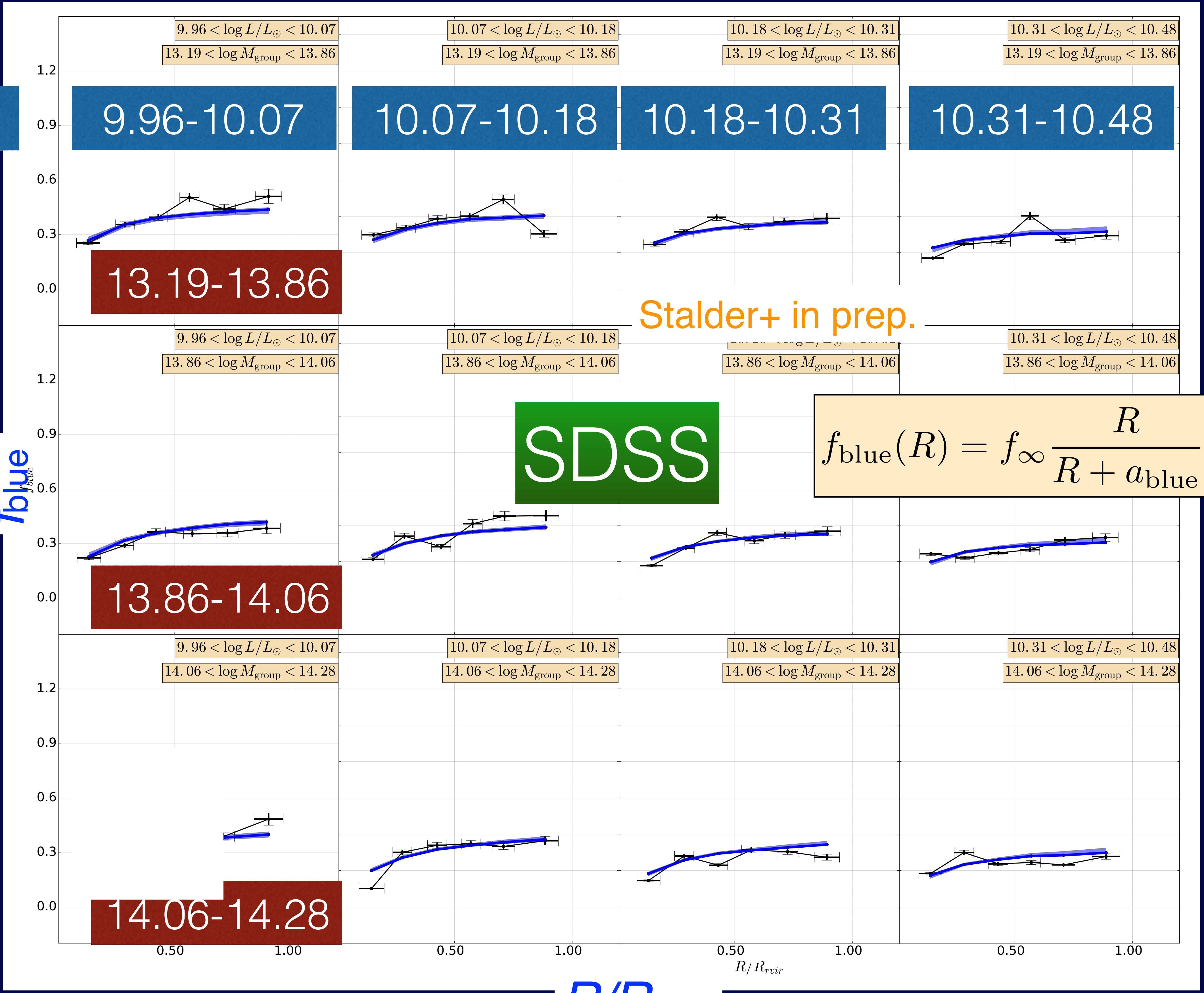
At high richness: Yang & FoF get worse!

Yang r_{200} biased high by 0.6 dex?

Environmental trends

$\log M_{\rm group}/M_\odot$ ↓

$\log L/L_\odot$ →



9.96-10.07

10.07-10.18

10.18-10.31

10.31-10.48

13.19-13.86

13.86-14.06

14.06-14.28

Stalder+ in prep.

SDSS

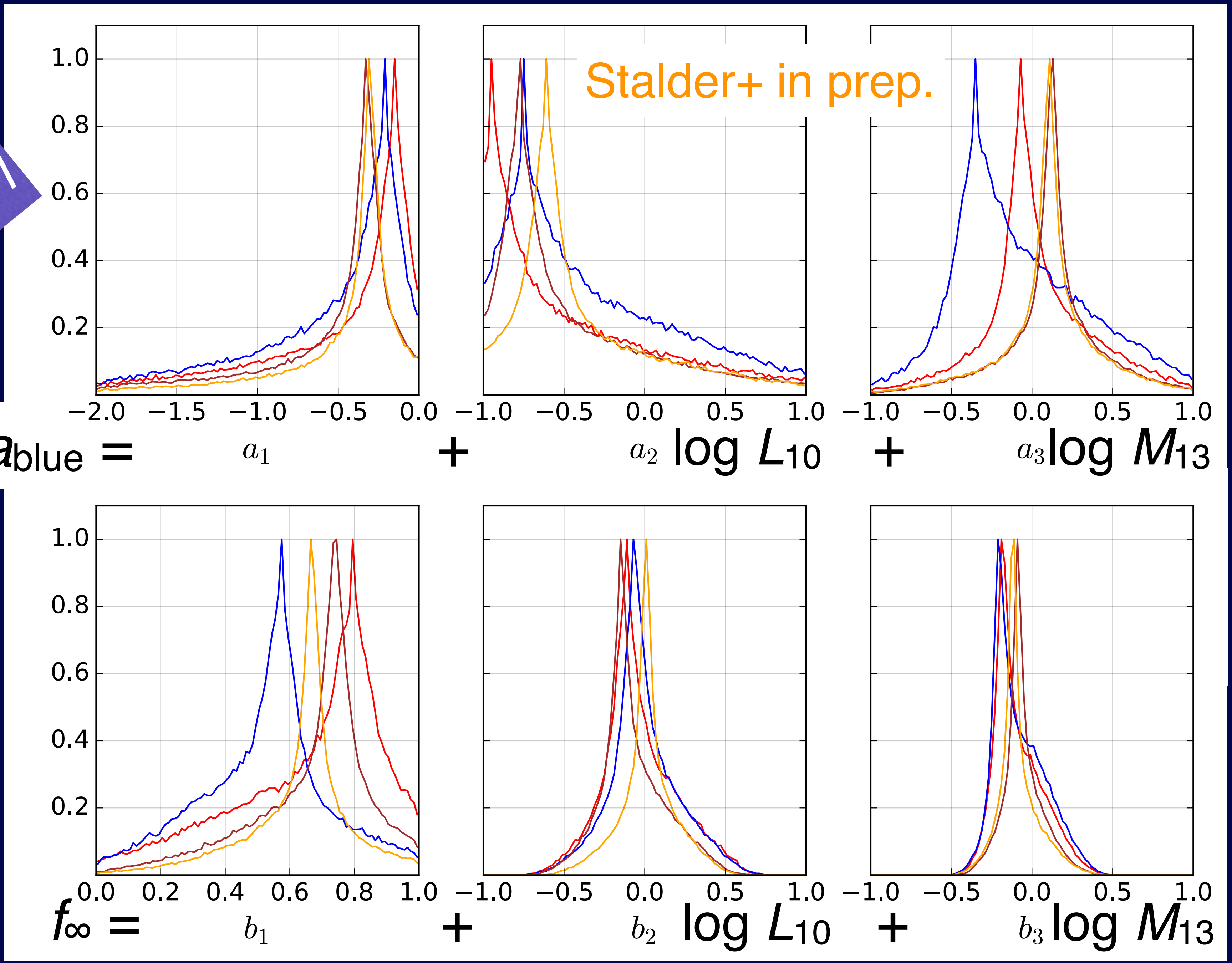
$$f_{\rm blue}(R) = f_\infty \frac{R}{R + a_{\rm blue}}$$

$f_{\rm blue}$

R/R_{200}

≠ Group Finders on mock-SAM

PRELIMINARY!
check scaling w L & M



Stalder+ in prep.

$$\log a_{\rm blue} = a_1 + a_2 \log L_{10} + a_3 \log M_{13}$$

$$f_\infty = b_1 + b_2 \log L_{10} + b_3 \log M_{13}$$

FoF
Yang
MAGGIE
Perfect

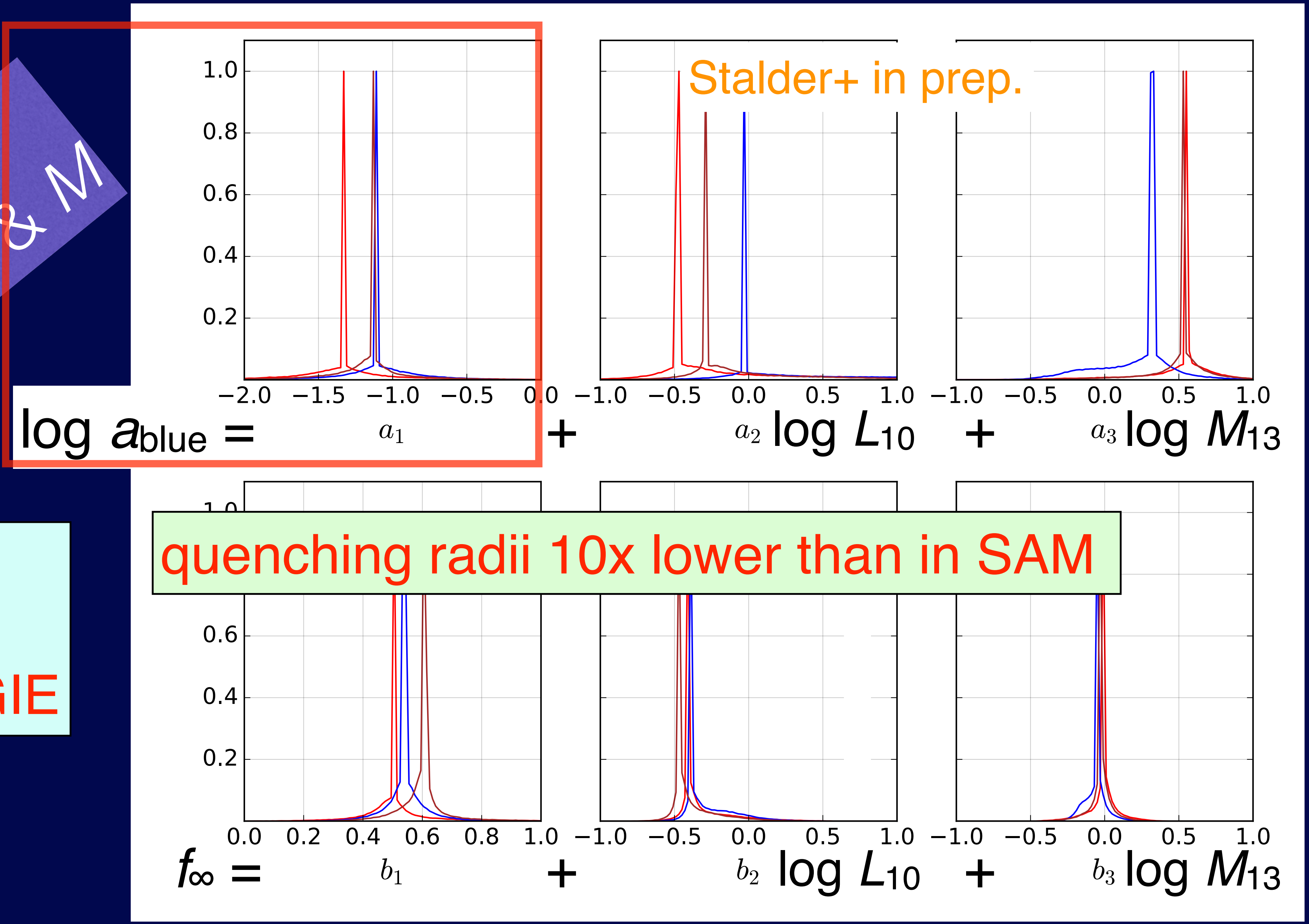
$$f_{\rm blue}(R) = f_\infty \frac{R}{R + a_{\rm blue}}$$

differences of ~ 0.2 dex in quenching projected radii

perfect mocks have lower quenching radii: i.e. less efficient quenching (!?)

≠ Group Finders on SDSS

PRELIMINARY!
check scaling w L & M



Stalder+ in prep.

quenching radii 10x lower than in SAM

FoF
Yang
MAGGIE

$$f_{\rm blue}(R) = f_\infty \frac{R}{R + a_{\rm blue}}$$

$$\log\left(\frac{a_{\rm blue}}{r_{200}}\right) = a_1 + a_2 \log\left(\frac{L}{10^{10}{\rm L}_\odot}\right) + a_3 \log\left(\frac{M_{\rm group}}{10^{13}{\rm M}_\odot}\right)$$

$$f_\infty = b_1 + b_2 \log\left(\frac{L}{10^{10}{\rm L}_\odot}\right) + b_3 \log\left(\frac{M_{\rm group}}{10^{13}{\rm M}_\odot}\right)$$

Do group properties depend on Ω_m?

What fraction of mock CGs are physically dense?

Díaz-Giménez & Mamon 10; Díaz-Giménez, GM+12

(DM) simulation	SAM	physically dense	Reference
2^3 virialized SAM (!)	Mamon 87	40%	Mamon 86, 87

What fraction of mock CGs are physically dense?

Díaz-Giménez & Mamon 10; Díaz-Giménez, GM+12

(DM) simulation	SAM	physically dense	Reference
2^3 virialized SAM (!)	Mamon 87	40%	Mamon 86, 87
2160^3 MS	Bower+06	77%	DíazG & GM 10
2160^3 MS	Croton+06	73%	DíazG & GM 10
2160^3 MS	De Lucia & Blaizot 07	58%	DíazG & GM 10
2160^3 MS-II	Guo+11	69%	DíazG+12
2160^3 MS	Guo+11	56%	DG+ in prep
2160^3 MS	Henriques+12	53%	DG+ in prep

1/2–2/3 CGs physically dense (90% within virialized groups)
1/3–1/2 chance alignments (80% within virialized groups)

What fraction of mock CGs are physically dense?

Díaz-Giménez & Mamon 10; Díaz-Giménez, GM+12

(DM) simulation	SAM	physically dense	Reference
2^3 virialized SAM (!)	Mamon 87	40%	Mamon 86, 87
2160^3 MS	Bower+06	77%	DíazG & GM 10
2160^3 MS	Croton+06	73%	DíazG & GM 10
2160^3 MS	De Lucia & Blaizot 07	58%	DíazG & GM 10
2160^3 MS-II	Guo+11	69%	DíazG+12
2160^3 MS	Guo+11	56%	DG+ in prep
2160^3 MS	Henriques+12	53%	DG+ in prep
MS → Planck	Henriques+15	31%	DG+ in prep

higher $\Omega_m \Rightarrow$ more CGs by chance alignments (now 70%!)

expect more chance alignments within filaments

Hernquist, Katz & Weinberg 95

Conclusions

- z-distortions ⇒ no group finder can be perfect: fragmentation, etc.
- Prior-based group finders are much better for nearby spec-z surveys
- ≠ group finders lead to ≠ results
  - ➡ LOS velocity dispersion profile
  - ➡ quenching radii

use ≠ Group Finders
e.g. w GGA (FoF, Yang, MAGGIE)
public release early ‘17

- Environmental effects NOT washed out by imperfect group finders(?)
- SDSS quenching radii 10x smaller than expected from Guo+11 SAM
- Compact groups: mocks with higher Ω_m:
  - ➡ 2x less frequent
  - ➡ 1.5x more contaminated by chance alignments (now > 50%!)