Measuring galaxy environments with group finders: Methods & Consequences

Euge DÍAZ-GIMÉNEZ
IATE, Cordoba, Argentina

Manuel DUARTE
former IAP Doc student

Marina TREVISAN
Postdoc, IAP

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

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 $\Omega_m$?
**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, ...
  - \(= f(\text{global environment, local envt, large-scale envt, redshift})\)

- **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?

\[ \Delta = \text{overdensity} / \text{critical density} \]

\[ \frac{cz}{H_0D + v_{\text{pec}}} = 11 - 18 \]

GM, Biviano & Murante 10

\[ \frac{r_{\text{max}}}{r_v} = \kappa \sqrt{\frac{\Delta}{2}} \left( \frac{\sigma_v}{v_v} \right) \approx 11 - 18 \]
Group finders
for spectroscopic galaxy samples

- **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)

survey edge

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

Duarte & Mamon 14

**mean transverse link**

\[
\frac{\delta n}{n} = \frac{3}{4\pi b^3_\perp} - 1
\]

\[
b_\perp = \left( \frac{3/(4\pi)}{\Delta/\Omega_m + 1} \right)^{1/3}
\]

= 0.07

**max (95% c.l.) transverse link**

\[
b_\perp = \frac{\text{Max}(S_\perp)}{n^{-1/3}} = \left( \frac{3/(4\pi)}{\Delta/\Omega_m + 1} \right)^{1/3} \frac{\text{Max}(S_\perp)}{r_{\text{vir}}} N_{\text{vir}}^{1/3} \approx 0.09 N^{0.08}
\]

\[
b_\perp = 0.10 \text{ for } N = 4 \text{ and } b_\perp = 0.12 \text{ for } N = 40
\]

**mean line-of-sight link**

\[
\frac{b_\parallel}{b_\perp} = \left( \frac{v_{\text{max}}}{\sigma_v} \right) \left( \frac{\sigma_v}{v_{\text{vir}}} \right) \sqrt{\frac{\Delta}{2}} \approx 11
\]

\[\Rightarrow b_\parallel = 1.1\]

for \(v_{\text{max}}/\sigma_v = 1.65\) (95%)
Background regions:
expect $a = \text{Area/} \langle \text{Area} \rangle$ follows $f(a) \propto a^3 e^{-4a}$ Kiang+66

$\rightarrow$ Area threshold

$\rightarrow$ groups/clusters = connected regions w Area below threshold

$\leftrightarrow$ FoF on low area cells!

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

Ramella+01
Dendrograms

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

Materne 78

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 $v_{\text{LOS}}$)
  - or magnitudes (LF prior) or
  - or photo-zs (Gaussian prior)
Yang et al.’s Halo-based Group Finder

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

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)

- 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

weaknesses
probabilistic

\[ P(R, v_z) = \frac{g_{\text{halo}}(R, v_z)}{g_{\text{halo}}(R, v_z) + g_{\text{ilop}}(R, v_z)} \]

more realistic \( g_{\text{halo}} \) from \( \Lambda \)CDM 3D model with anisotropic velocities

\[ g_h(R, v_z) = \Sigma_{\text{sph}}^{\text{NFW}}(R) \langle h(v_z|R, r) \rangle_{\text{LOS-sph}} \]

\[ = 2 \int_{R}^{r_{200}} \nu(r) h(v_z|R, r) \frac{r \, dr}{\sqrt{r^2 - R^2}} \]

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

\[ \sigma_r^2(R, r) = \left(1 - \beta(r) \frac{R^2}{r^2}\right) \sigma_i^2(r) \]

\( \sigma_i(r) \) from solving Jeans equation \( \beta(r) \) from cosmo simulations
\[ P(R, v_z) = \frac{g_{\text{halo}}(R, v_z)}{g_{\text{halo}}(R, v_z) + g_{\text{ilop}}(R, v_z)} \]

\[ g_i(R, v_z) = \frac{N_{200}}{r_{200}v_{200}} \hat{g}_i \left( \frac{R}{r_{200}}, \frac{v_z}{v_{200}} \right) \]

\[ \hat{g}_i(X, u) = A(X) \exp \left[ -\frac{1}{2} \frac{u^2}{\sigma_i^2(X)} \right] + B \]

**DM particles in HD simulation:** Borgani+04

**SAM:** Guo+11
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**

\[ N_{est} \geq 3 \; \& \; N_{true} \geq 3 \; \& \; \text{unflagged} \]

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)

---

Gary Mamon (IAP), Measuring galaxy environments w group finders: methods & consequences, IAP, 13 Dec 2016, Physics of Groups and Galaxy Properties therein
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:
→ $b_\perp = 0.07$ & $b_\parallel = 1.1$
≈ 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 $N_{\text{true}} \geq 3$ & $N_{\text{est}} \geq 3$

fraction of extracted groups = secondary fragments of true groups

M-based: more physical than L-based, but higher (systematic) errors
**Group total mass accuracy**

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)

only unflagged primary-fragment groups $N_{\text{true}} \geq 3$ & $N_{\text{est}} \geq 3$
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-zs
Maurogordato, Biviano +

Best algorithms: FoF, matched filter, then wavelets
Group properties vs. group finder
Surface density profiles of SDSS groups

$N \geq 5$

$\log M/M_\odot > 13.1$

$M_r < -19$

FoF & Yang consistent with NFW for $0.04 (F) \text{ or } 0.08 (Y) < R/r_{200} < 1$
Surface density profiles of SDSS groups

\[ N \geq 5 \]
\[ \log \frac{M}{M_\odot} > 13.1 \]
\[ M_r < -19 \]

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
returned $\sigma_v/\nu_{200}$:

- Yang: too high
- FoF: too low
- MAGGIE: just right

At high richness: Yang & FoF get worse!
At high richness: Yang & FoF get worse!

Yang $r_{200}$ biased high by 0.6 dex?

0.2 dex bias in $r_{200}$ & $v_{200} \Rightarrow 0.6$ dex bias in Yang!

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

\[ f_{\text{blue}}(R) = f_{\infty} \frac{R}{R + a_{\text{blue}}} \]

\( \log M_{\text{group}}/M_\odot \downarrow \)

\( \log L/L_\odot \rightarrow \)

<table>
<thead>
<tr>
<th>Log L/L_\odot</th>
<th>9.96-10.07</th>
<th>10.07-10.18</th>
<th>10.18-10.31</th>
<th>10.31-10.48</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.19-13.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.86-14.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.06-14.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Log M_{\text{group}}</th>
<th>9.96-10.07</th>
<th>10.07-10.18</th>
<th>10.18-10.31</th>
<th>10.31-10.48</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.19-13.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.86-14.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.06-14.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Stalder+ in prep.

SDSS
Group Finders on mock-SAM

PRELIMINARY! check scaling w L & M

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

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

differences of $\sim 0.2$ dex in quenching projected radii

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

Gary Mamon (IAP), Measuring galaxy environments w group finders: methods & consequences, IAP, 13 Dec 2016, Physics of Groups and Galaxy Properties therein
# Group Finders on SDSS

**PRELIMINARY!**
check scaling w L & M

**FoF**
Yang
MAGGIE

\[
f_{\text{blue}}(R) = f_\infty \frac{R}{R + a_{\text{blue}}}
\]

\[
\log \left( \frac{a_{\text{blue}}}{r_{200}} \right) = a_1 + a_2 \log \left( \frac{L}{10^{10} L_\odot} \right) + a_3 \log \left( \frac{M_{\text{group}}}{10^{13} M_\odot} \right)
\]

\[
f_\infty = b_1 + b_2 \log \left( \frac{L}{10^{10} L_\odot} \right) + b_3 \log \left( \frac{M_{\text{group}}}{10^{13} M_\odot} \right)
\]

Stalder+ in prep.

quenching radii 10x lower than in SAM
Do group properties depend on $\Omega_m$?
### What fraction of mock CGs are *physically dense*?

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

<table>
<thead>
<tr>
<th>(DM) simulation</th>
<th>SAM</th>
<th>physically dense</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^3$ virialized SAM (!)</td>
<td>Mamon 87</td>
<td>40%</td>
<td>Mamon 86, 87</td>
</tr>
</tbody>
</table>
What fraction of mock CGs are *physically dense*?

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

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

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

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

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

Hernquist, Katz & Weinberg 95
Conclusions

• z-distortions $\Rightarrow$ no group finder can be perfect: fragmentation, etc.

• Prior-based group finders are much better for nearby spec-z surveys

• $\neq$ group finders lead to $\neq$ results
  ➡ LOS velocity dispersion profile
  ➡ quenching radii

• 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 $\Omega_m$:
  ➡ 2x less frequent
  ➡ 1.5x more contaminated by chance alignments (now > 50%!)