Mon. Not. R. Astron. Soc. **000**, 1–4 (2015) Printed 3 April 2015 (MN LaTeX style file v2.2)

Galaxy-dark matter offsets in galaxy clusters and groups of the Illustris simulation

Karen Y. Ng,[1] Annalisa P. Pillepich,[2] William A. Dawson,[3] D. Wittman,[1]
Lars Hernquist,[2] etc. [order TBD]

arXiv

ABSTRACT
Being able to find the center of the dark matter component of a galaxy cluster or group enables correct stacking and mass estimation if any parametric halo profile is employed.

Key words: galaxy clusters, dark matter, something else

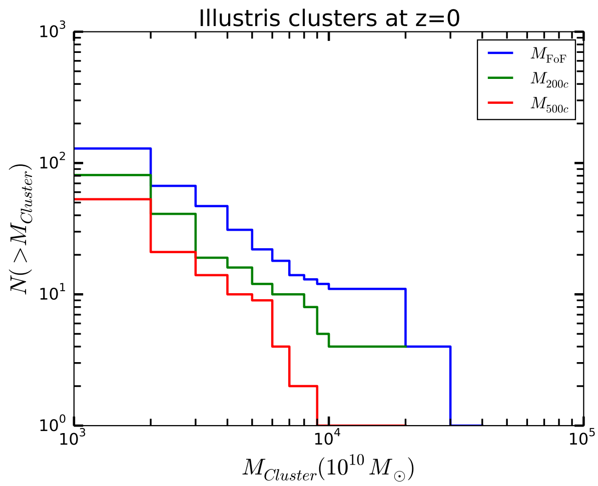

Figure 1.

1 INTRODUCTION

Galaxies that belongs to a galaxy cluster or group are stochastic samples of the underlying dark matter mass density.

We look for summary statistic that would have the least biased estimate of the dark matter density peak. For stacking the different sets of data, find a good proxy for the dark matter density peak is important for the peak not to be smoothed out.

As we try to utilize our knowledge of galaxy clusters for precision cosmology, we need to quantify and propagate our uncertainties carefully.

2 DATA

2.1 Data from the Illustris simulation

2.1.1 *Properties of the galaxy clusters / groups*

Most important properties of the galaxy clusters that we examine in this study include, colors, magnitude, richness, concentration and non-relaxedness. We define and examined these properties one by one. Fig 4. Relaxedness Fig 5.

Richness characterizes the number of galaxies in a given cluster. We provide several definitions of non-relaxedness to characterize whether the clusters underwent any recent merger activities. These definitions of non-relaxedness. The Shapiro-Wilk statistic can characterize the deviation from normality with the highest statistical power ().

2.1.2 *Volume Selection*

1. Rockstar halo finder 2. Light cone 3. spatial projections

2.1.3 *Galaxy Selection*

2.1.4 *Galaxy weights*

One of the most common weighting scheme employed for cluster galaxy data is to weight by the luminosity in a particular band.

2.1.5 *Data with and without noise*

1. Mass-richness diagram - with different cuts

3 METHODS

3.1 Galaxy Centers

3.1.1 Unweighted and weighted centroids

We follow the usual definition of spatial centroid as

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_i \vec{x}_i. \tag{1}$$

While the weighted centroids are just:

$$\bar{\mathbf{x}}_w = \frac{\sum_i w_i \vec{x}_i}{\sum_i w_i}, \tag{2}$$

for each spatial dimension and the weights w_i for the $i-th$ galaxy is described in section. Centroids and weighted centroids can be biased by merging activities yet do not provide explicit evidence for ongoing merger or accretion.

3.1.2 Cross-validated Kernel Density Estimation (KDE) and the peak finder

We employed a KDE algorithm to infer a smooth density distribution of the galaxies while using smoothed cross-validation to obtain the optimal smoothing bandwidth matrices (H). Specifically, we made use of the KDE function in the statistical package ks (Duong) in the R statistical computing environment (R Core Team 2014). Cross validation eliminates free parameters in the KDE and minimizes the asymptotic mean-integrated squared error (AMISE) for a best fit to the data. After obtaining the KDE estimate, we employed a finite differencing algorithm to find the local maxima. We sorted the local maxima according to the KDE density at the maxima locations and identified the dominant peak.

Spatial location and the density of the subdominant peaks are also stored. The subdominant

3.1.3 Shrinking aperture

Another popular method among astronomers for finding the peak of a spatial distribution include what we call a shrinking aperture method. We test if the shrinking aperture method is able to recover the highest peak reliably. This method is dependent on the initial diameter and the center location of the aperture. Furthermore, the convergence criteria is not set objectively. We present the convergence criteria for reference. We note that different implementation may result in different performances.

Data: subhalo that satisfy cuts as a galaxy

```
initial center = mean(data_array)
dist_array = euclidean dist(initial center, data array)
apert = get_90th_percentile(dist_array)
while (newCenterDist - oldCenterDist) /
oldCenterDist ≥ 2e-2 do
    new data array = old data array within apert
    newCenter = mean value of new data along each
    spatial dimension
end
```

Algorithm 1: Shrinking aperture algorithm

3.1.4 Brightest Cluster Galaxies (BCG)

3.1.5 Comparison of the methods from test data

In order to examine the performance of commonly used point-estimates of the distribution of the galaxy data, we test them on data drawn from Gaussian mixtures with known mean and variance.
Fig 1. one Normal mixture
Fig 2. one big normal mixture and one smaller normal mixture
Fig 3. three bridged normal mixtures
We compare the properties and performance of each of the methods for finding the peaks of the galaxy and dark matter, except the BCG since it does not rely on the cluster member population. The main factors that affect the performance of the methods depend heavily on statistical fluctuations of the drawn data. Namely, the performance of each method depends on: 1) the number of Gaussian mixture used, 2) the number of data points in each mixture

Due to the statistical nature of this exercise, it is not enough to just compare the performance from one realization of each method. We also provide the 68% and the 95% confidence regions from

The details and implementation can be found in our Bitbucket git repository.

3.2 DM Centers

3.3 Finding offsets

We computed the projected offsets between the galaxy density peaks inferred from the cross-validated KDE and the dark matter density peak. (TBD) The viewing angles of the projections are defined by an elevation angle ξ and an azimuthal angle ϕ. We sample at 5 evenly spaced values of ξ where $0 \leq \xi \leq \pi/2$ and 10 evenly spaced values of ϕ where $0 \leq \phi \leq \pi$, to give 50 samples for each cluster with mass $> 10^{14} M_{\odot}$.

This method gives us samples of the joint distribution:

$$P(\Delta\eta, \phi, \xi, h_{mer}, \psi_g, \psi_D, I_g, I_D | \theta, \sigma_{CDM}). \tag{3}$$

Binning the samples based on $\Delta\eta$ and h_{mer} only gives us the marginal distribution:

$$P(\Delta\eta, h_{mer} | \theta, \sigma_{CDM}), \tag{4}$$

that we have far too few samples of ...

4 RESULTS

4.1 Galaxy-DM Offset in Illustris

4.1.1 Projected offsets

4.1.2 Correlations between the offsets and properties of the cluster / groups

5 DISCUSSION

5.1 Comparison to other simulations

5.2 Comparison to other observational studies

5.3 Galaxy-DM Offset in Merging Galaxy Clusters

5.4 Summary

We showed that

- the peak finding method To-be-finalized for the density of cluster galaxies is the least biased due to substructures from our test data.
- all existing peak finding methods have non-negligible uncertainty due to the small number of data points. When dealing with small number of cluster samples, the uncertainties of the peak locations should not be ignored.

6 ACKNOWLEDGEMENTS

APPENDIX A: KDE

This paper has been typeset from a TEX/ LATEX file prepared by the author.