Mon. Not. R. Astron. Soc. **000**, 1–11 (2015) Printed 12 June 2016 (MN LATEX style file v2.2)

New physics or wrong statistics? The 'offsets' between dark matter and the galaxies in the galaxy clusters in the Illustris simulation

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

arXiv

ABSTRACT

Galaxy clusters, which mainly compose of dark matter (DM), can be rare test beds for the particle properties of DM. However, the continuous merger and accretion events of clusters also complicate the modeling of galaxy clusters. With uncertainties coming from various modeling choices and observational constraints, we need to carefully account for the uncertainties for us to give meaningful quantitative constraints from the studies of galaxy clusters. In this paper, we test various summary statistics of the member galaxy components of galaxy clusters by computing them from the data from a cosmological simulation, the Illustris simulation. We examine the uncertainties of the different summary statistics, and see if the galaxy population have statistics consistent with those of the DM population. TODO: result summary. We found that the uncertainty of the offset resulting from projection effects are non-negligible and vary in unpredictable ways.

Key words: galaxy clusters, dark matter, statistics

1 INTRODUCTION

During the latest stage of structure formation, the universe gave birth to non-linear, hierarchical structures known as galaxy clusters. These clusters, made up of dark matter, galaxies and hot gas, are constantly accreting, merging and evolving with their environments. Bright galaxies that belong to a galaxy cluster or group, in particular, highlight the overdensities of the underlying dark matter (DM) distribution.

In these dense regions of the clusters, the rates of particle interactions can be enhanced, including the long-suspected self-interaction of DM particles (hereafter, SIDM). Many papers have used the offsets between the summary statistics of the DM density and the galaxy density to give constraints on the self-interaction cross section, i.e. $\sigma_{\rm SIDM}$, of dark matter (Markevitch et al. 2004, Bradač et al. 2006, Mertel, Dawson, Jee etc.). However, it is unclear how large an offset there is between the two components without any effects of SIDM. A lot of simulations that galaxy-dark matter offset use parametric prescriptions of the spatial distribution galaxies (Kahlhoefer et al. 2013, Markevitch et al. 2004, Robertson et al. 2016), and may not have realistic substructures. These idealistic simulation settings may underestimate the contribution of the intrinsic scatter to the galaxy-dark matter offsets.

We investigate the Illustris dataset from a statistical perspective to see how to best compute the instrinsic scatter of the offsets without SIDM. This is the first study to investigate if it is a feasible idea to use the galaxy-DM offsets given the main constraints from observations, namely, the projections and the dynamical states of the galaxy clusters. These two latent variables are confounding and can increase the variance of population summary statistics of galaxy clusters.

Popular choice for computing the offsets involves first inferring the summary statistic of each of the DM and the galaxy population of a cluster before taking a difference. While there are well established procedure driven by lensing physics for inferring the DM spatial distribution, there is no standard procedure for mapping the sparse member galaxy distribution. We quantify the bias and uncertainty associated with the one-point summary statistic for summarizing the physical state of a galaxy cluster.

In this paper, we 1) extract realistic observables from the Illustris simulation for comparison with observations, 2) explore the pros and cons of the different statistic for summarizing *the member galaxy population* of a galaxy cluster, 3) give estimates for the offsets between the summary statistics of the galaxy population and the DM population under ΛCDM cosmology, which we call

$$\Delta s \equiv \boldsymbol{s}_{\rm gal} - \boldsymbol{s}_{\rm DM}. \quad (1)$$

where $\mathbf{s_{gal}}$ and $\mathbf{s_{DM}}$ are the two-dimensional (2D) spatial locations of the summary statistic of the galaxy population, and the density peak of DM respectively. This gives an estimate of the baseline scatter of offsets without any SIDM. And finally we 4) examine the properties of the clusters that give outliers in the offset distribution and 5) investigate the correlations between the 3-dimensional properties of a cluster and the projected observables such as Δs.

The organization of this paper is as follows: In section 2, we will describe the physical properties of the data of the Illustris simulation (Vogelsberger et al. 2014a, Genel et al. 2014), and the selection criteria that we have employed to ensure that the quantities that we examine resemble observables but without noise and

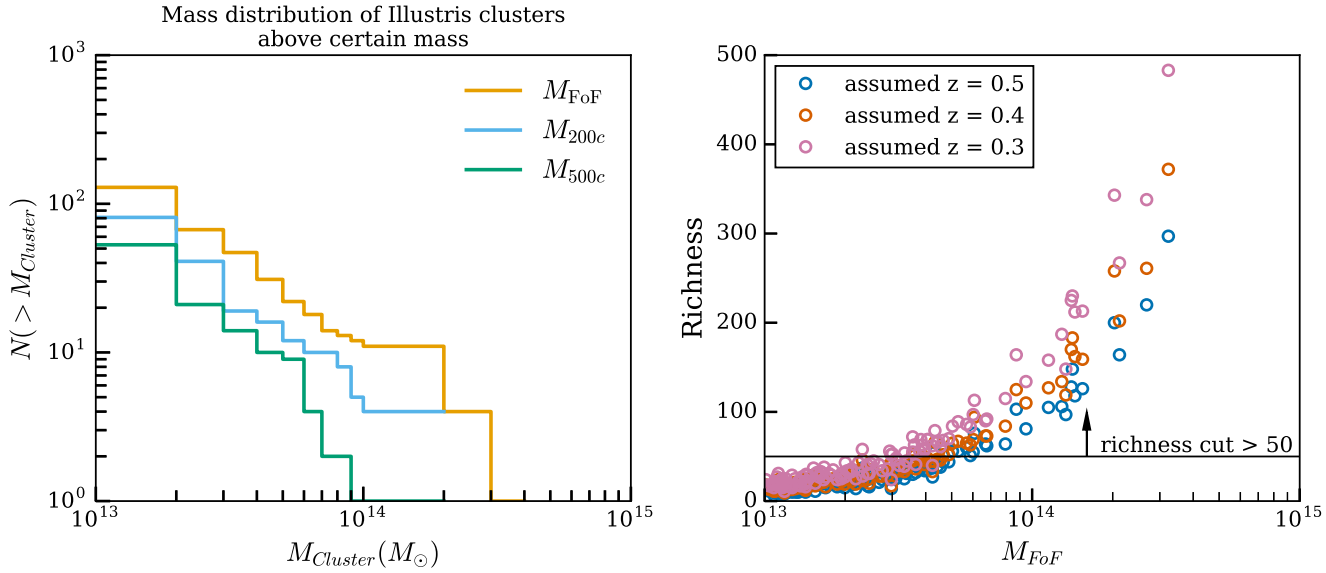

Figure 1. Left figure: Mass distribution of the group / cluster sized DM halos for different halo selection schemes. Mass estimates obtained by the FoF algorithm are labeled as $M_{\rm FoF}$. Masses centered on the most bound particle within a radius those the average density is 200 or 500 times the critical density of the universe are labeled as M_{200c} and M_{500c} respectively. **Right figure:** Mass-richness relationship of galaxy clusters and groups with $M_{\rm FoF} > 10^{13} M_\odot$ assuming different cosmological redshifts of the observed clusters.

systematics from observations. Then in section 3, we explain the methods for computing various one-point statistics of the spatial distribution of galaxies how we prepare our dark matter spatial data to resemble convergence maps. We show the statistical performance of the different summary statistics before we show the main results in section 4. In the discussion in section 5, we list the implications of our results and compare it to other simulations and observations. We also show how one may make use of the population offset statistical distribution from the Illustris data to construct a two-tail p-value test with a null hypothesis of $\sigma_{\rm SIDM} = 0$.

Our analysis makes use of the same flat Lambda Cold Dark Matter (ΛCDM) cosmology as the Illustris simulation. The relevant cosmological parameters are $\Omega_\Lambda = 0.7274, \Omega_m = 0.2726$, and $H_0 = 70.4$ km s^{-1} Mpc^{-1}.

2 THE ILLUSTRIS SIMULATION DATA

The Illustris simulation contains some of the most realistic, simulated galaxies to date, making it especially suitable for verifying the properties of galaxy clusters. We obtained our data from snapshot number 135 (cosmological $z = 0$) of the Illustris-1 simulation. The Illustris-1 simulation has the highest particle resolution and has incorporated the most comprehensive baryonic physics among the different Illustris simulation suites. The sophisticated galaxy formation model in Illustris-1 includes star formation rate, and stellar evolution due to environmental effects of the intracluster medium, such as ram pressure stripping and strangulation and feedback from Active Galactic Nuclei (AGN) etc. (Genel et al. 2014). The physics of stellar evolution were solved using a moving mesh code **AREPO** (Springel 2010). The observable properties of galaxies were statistically consistent with the Sloan Digital Sky Survey (SDSS) data (Vogelsberger et al. 2014a).

As the stellar population in Illustris were evolved from the initial condition, these makes the spatial distribution of galaxies in Illustris data more realistic than galaxies that are prescribed onto DM-only cosmological simulation data such as those used in Harvey et al. (2014). Gravitational effects in Illustris-1 have provided realistic dynamics and spatial distribution of subhalos. The simulated effects include tidal stripping, dynamical friction and merging. Since the profile of the galaxies clusters were not provided in symmetrical, parametric forms, we can study how asymmetry in the cluster profile affects the estimate of our summary statistic. This data allows us to examine cluster galaxies in a realistic, yet noise-free way. The softening length of the DM particles is 1.4 kpc and those of the stellar particles is 0.7 kpc, both in constant comoving units (Genel et al. 2014).

The two sets of data catalogs in use are obtained through two types of halo finders. The catalog that maps particles to the halo of a certain cluster was created by the **SUBFIND** algorithm. The friends-of-friends (FoF) finder (Davis et al. 1985) was further used to identify the affinity of galaxy-sized halos to a galaxy-cluster. These galaxy-size halos are referred to as *subhalos* and they are the dark matter hosts of what we refer to as galaxies in Illustris-1. Vogelsberger et al. (2014b) also extracted the absolute magnitude of each subhalo in the SDSS bands of g, r, i, z as part of the **SUBFIND** catalog using stellar population synthesis models.

For our analyses, we make use of galaxy clusters / groups with at least 50 member galaxies that are above a reasonable observation limit, i.e. apparent $i \leq 24.4$ when we assume a cosmological redshift of $z = 0.3$ in the i band. This is because of the relatively large statistical uncertainty if we try to analyze clusters with less than 50 member galaxies. As indicated by the right-hand panel of Fig. 1, a total of 43 clusters have survived this magnitude cut.

2.1 Cluster properties

2.1.1 *Relaxedness of the galaxy clusters*

Clusters undergo merger activities of a large range of physical scales and in the time scale of million of years. The dynamical history, or what we call "relaxedness" is hard to retrieve from simulations

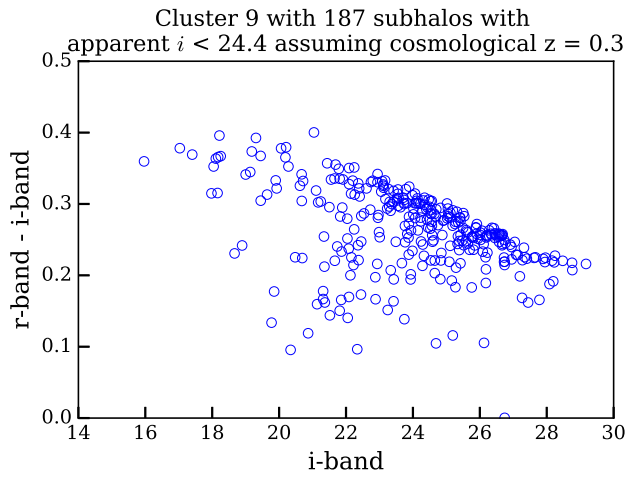

Figure 2. Color-magnitude diagram of one of the galaxy clusters that is selected for analysis. This cluster is the 9th most massive. The apparent magnitude is calculated assuming that the cosmological redshift (distance) is $z = 0.3$. We can see a clear overdense region that corresponds to a red-sequence. The color-magnitude diagrams of the other clusters can be found in the Jupyter notebook at https://goo.gl/TJmI6s.

across different saved states and is missing from observations. We quantify the state of the cluster by providing several quantitative definitions of relaxedness and see how they correlate with Δs. Some definitions of relaxedness referred by the simulation community include:

- the ratio of mass outside the dominant dark matter halo over the total mass of the galaxy cluster
- the distance between the most bound particle from the center of mass as a function of R_{200c}.

which are computable from the Illustris data. To relate these simulation quantities, we compute more observation oriented quantities in the method section 3.0.2.

2.2 Selection of the field-of-view

We make use of the **SUBFIND** member particle for the DM and the **FoF** subhalo identification as our default volume selection scheme for each cluster / group. We understand that this choice of volume selection can be more ideal than observational conditions. We make use of this volume selection scheme for baseline comparisons.

Assuming a conservative line-of-sight (los) distance , i.e. cosmological redshift, with [TODO] $z = 0.4$, the projected extent for most of the Illustris galaxy clusters and groups, fits inside the field of view of telescopes, such as the Subaru Suprime Camera, which covers a physical area of [TODO] ~ 9 Mpc $\times 7$ Mpc. (See https://goo.gl/ClZNvM for a Jupyter notebook showing the extent of the Dark Matter distribution of the most massive 129 clusters)

2.2.1 *Spatial Projections*

The summary statistics are computed all based on 2D projections of the spatial location. For computing the summary statistic, we note that the order of projecting the data and taking the summary statistic is non-commutative. In order to represent the projection uncertainty, we compute even angular orientation as our line-of-sight by using HealPy, which is a Python wrapper for HEALPix[1] (Gorski et al. 2005). The number of projections that we employed is 384 for each cluster. Details of the implementation of the projection is available in Appendix A.

2.3 Properties of the galaxies in Illustris clusters

2.3.1 *Galaxy weights*

Different galaxies have different masses, so they should not be considered with equal importance for peak identification, which requires summing the mass proxies of different galaxies. One of the most common weighting schemes employed for galaxy data is to weight by the luminosity in a particular band. We make use of the $i-$band magnitude associated with each subhalo as the weight. Since the $i-$band is one of the redder bands, the mass-to-light ratio is not skewed as much due to star formation activities. We further examined if the colors distribution of galaxies in Illustris-1 are similar to the observed color-magnitude diagrams for clusters. The Illustris cluster galaxies are realistic enough that it is easy to identify an overdense region of galaxies known as the red-sequence in the color-magnitude diagram such as Fig. 2). The red-sequence is prominent even if we use other colors formed by different combinations of r, i, z bands.

3 METHODS

A common and the most precise way of summarizing the DM distribution in a galaxy cluster is by finding the lensing peaks (Medezinski et al. 2013, Markevitch et al. 2004, Zitrin et al. 2013). Additionally, the peak region is physically interesting due to the higher particle density and interaction rates. The most direct analogous statistic for summarizing the member galaxy population in a cluster is therefore, also the peak. Comparing the DM peak with the summary statistics of the galaxy population that are not the peak therefore can have an *offset* purely due to the difference in the choice of the statistic for summarizing the two sets of identically distributed data.

We compare four common point statistic or location for summarizing the member galaxy population in a galaxy cluster:

(i) Weighted centroids
(ii) Weighted density peak via density estimate
(iii) Shrinking aperture estimate
(iv) Brightest cluster galaxy (BCG)

We avoid any manual methods for comparison purposes, scalability and reproducibility. Since all the methods listed in this paper are automated with the source code openly available, it is possible for future studies to reuse our code for comparisons. Furthermore, a major advantage for automation is that it allows us to scale up our analysis by applying the same methods across the different snapshots of the (Illustris) simulations to examine the variability of Δs across time.

3.0.1 *Computing the weighted centroid*

We follow the usual definition of spatial centroid as

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_i \boldsymbol{x}_i. \quad (2)$$

[1] HEALPix is currently hosted at http://healpix.sourceforge.net

Table 1. Selection criteria for stellar subhalos (member galaxies) for each cluster / group

Data	Selection strategy	Sensitivity	Relevant section
Field of view (FOV)	FoF halo finder	comparable to FOV of the Subaru Suprime camera	2.2
Observed filter	i-band	consistent among the redder r, i, z bands	
Cluster richness	$i \leq 24.4$ and $z = 0.3$	sensitive to the assumed cosmological redshift of cluster and the assumed limiting magnitude of telescope	2
Two-dimensional projections	even HEALPix samples over half a sphere	discussed as results	2.2.1

Table 2. Comparison between various methods for estimating the one-point statistics of the galaxies of a cluster

Method	One-point statistic	Sensitivity to biases	Uncertainty	Relevant section	Comment
Centroid	2D spatial averages	High	Low		
Shrinking aperture	proxy for density peak	Higher sensitivity to substructures	Medium		
Peak finding from KDE	density peak	Lower sensitivity to substructures	Higher		
Brightest cluster galaxy		Sensitive to foreground contaminants			
Most bound particle	bottom of gravitational potential well				

While the weighted centroid is just:

$$\bar{\mathbf{x}}_w = \frac{\sum_i w_i \boldsymbol{x}_i}{\sum_i w_i}, \tag{3}$$

with $\boldsymbol{x}_i$ being the positional vector of each subhalo and we use the i-band luminosity as the weight w_i for the i-th galaxy. Centroids can be biased 1) by subcomponents from merging activities yet the centroid estimate do not provide explicit evidence for ongoing merger or accretion. These estimates are also sensitive to odd boundaries of the field of view.

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

Finding the exact peak of a sets of data points involves computing the density estimate of the data points and sorting through the density estimates. A specific version of this density estimation process is known as histogramming. During the making of histogram, each data point is given some weight using a tophat kernel and the weights are summed up at specific data locations (e.g. $\mathbf{x_i}$). Whereas, for the density estimate discussed here, the form of the kernel is a Gaussian. The exact choice of the functional form of the smoothing kernel does not dominate the density estimate as long as the chosen kernel is smooth (Feigelson & Babu 2014).

The most important parameter of computing the density estimate is the width of the smoothing kernel. The kernel width can greatly affect both the density estimates and the peak location(s). We illustrate the choice of kernel width with Fig. 3. When the kernel width is too large (bottom left panel), the data is over-smoothed, resulting in a bias of the peak estimate. On the other hand, when the kernel width is too small, it results in high variances of the estimate and result in multiple sub-peaks. The decision of having to balance between creating high bias or high variance estimates is also known as the bias-variance tradeoff.

A well-known way to minimize the fitting error from the density estimate is through a data-based approach called cross-validation to obtain the optimal 2D smoothing bandwidth matrix (H) of the 2D Gaussian kernel for the density estimate $\hat{f}$:

$$\hat{f}(\chi; H) = \frac{1}{n} \frac{1}{(2\pi)^{d/2}|H|^{1/2}} \sum_{i=1}^{n} \exp((\chi - \mathbf{x}_i)^T H^{-1} (\chi - \mathbf{x}_i)), \tag{4}$$

where the dimensionality is $d = 2$ for our projected quantities, χ represents the uniform grid points for evaluation, and $\mathbf{x_i}$ contains the spatial coordinates for each of the identified member galaxies that survived our brightness cut. The idea behind cross-validation is to leave a small fraction of data point out as the test set, and use the rest of the data points as the training set for computing the estimated density. Then it is possible to minimize the asymptotic mean-integrated squared error (AMISE) by searching for the best set of bandwidth matrix values, eliminating any free parameters. Specifically, we made use of the smoothed-cross validation (Hall et al. 1992) bandwidth selector in the statistical package **ks** (Duong 2007) in the **R** statistical computing environment (R Core Team 2014). Among all the different **R** packages, **ks** is the only package capable of handling the different weights of the data points while inferring the density estimates (Deng & Wickham 2011). Although the particular implementation of KDE has a computational runtime of $O(n^2)$, the number of cluster galaxies is small enough for this method to finish quickly.

After obtaining the KDE estimate, we employed both a first and second-order finite differencing algorithm to find the local maxima. The local maxima were then sorted according to the KDE density in a descending fashion before we perform peak matching. The density of the peaks are normalized as a fraction of the densest peak for each projection of each cluster in our sample.

3.0.3 *Shrinking aperture estimates*

Another popular method among astronomers for finding the peak of a spatial distribution is what we call the shrinking aperture method. While we do not endorse this method, we test if the shrinking aperture method is able to reliably recover the densest peak. This method is dependent on the initial diameter and the initial center location of the aperture. This method does not evaluate if the cluster is made up of several components. The estimate using the shrinking aperture algorithm can be biased by substructures. The only way to inform the algorithm about substructures would be to introduce another parameter to restrict the center of the aperture, or to partition the data with another (statistical) algorithm. Furthermore, the convergence rate for this iterative algorithm is not analytical and is dependent on both the data and the parameters. We use a convergence criteria of having the aperture distance not change more than 2% between successive iterations as a reference. The actual implementation in Python can be found at https://goo.gl/nqxJl8 while a pseudo-code can be find in the Appendix B.

Figure 3. This figure is adapted from VanderPlas et al. 2012 from http://www.astroml.org/book_figures/chapter6/fig_hist_to_kernel.html under the fair use of the BSD license.

3.0.4 Brightest Cluster Galaxies (BCG)

The BCGs are formed by the merger of many smaller galaxies. The galaxy-cannibalism makes BCGs typically brighter than the rest the cluster galaxy population by several orders of magnitude. (CITE?) However, star formation can result in galaxies brighter in the bluer photometric bands. To avoid star formation from biasing our algorithm for identifying the BCG, we find at the brightest galaxies in redder bands i.e. the r, i, z bands and found that they give consistent results for all selected clusters. The band that we used to pick the BCG for presentation and the final results is the i-band.

3.1 Comparison of the methods from test data

In order to examine the statistical properties 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. (See Fig. 4). The main factors that affect the performance of the methods are sensitive to the statistical fluctuations of the drawn data, e.g. the spatial distribution of the data, including 1) the density profile and 2) the location(s) of subdominant mixtures, , and 3) the number of data points that we draw. It is also not enough to just compare the performance by applying each method for one realization of the data. We provide the 68% and the 95% confidence regions by applying the each method for many Monte Carlo realizations. In general, the peaks identified from the KDE density is closer to the peak of the dominant mixture (more accurate) than both the weighted centroid method and the shrinking aperture method. For example, in the bottom middle panel, it is clear that the green contours that represents the confidence region for the shrinking aperture peak is biased due to the substructure, whereas the confidence region for the centroid is so biased that it is outside the field of view of that panel. For the bottom right plot, there is also a catastrophic outlier for the shrinking aperture method for 500 data points. The outlier shows how the shrinking aperture method can have radical behavior when there are subclusters in the data.

3.2 Modeling the DM map in Illustris-1 and the lensing kernel

The most well established method of inferring the projected dark matter spatial distribution from observations is through gravitational lensing. It works by detecting subtle image distortions of background galaxies due to the foreground dark matter. The resolution of the inferred map therefore depends on the properties of the galaxies, such as the projected number density, intrinsic ellipticities and morphology etc. To achieve a sufficient signal-to-noise ratio for lensing, Hoag et al. 2016 has performed simulation for inferring the optimal size for a Gaussian smoothing kernel for the cluster MACSJ0416. In the strong lensing regime, Hoag et al. found a resolution of 11 arcseconds can best maximize the fit. This kernel translates to a physical size of 50 kpc assuming a cosmological redshift of $z \approx 0.3$. With the lack of better physically motivated kernel size, we adopt 50 kpc as the smoothing kernel of the DM histogram (with 2 kpc bin size) for the Illustris DM particle data. Physically, the histograms of the dark matter of each cluster is analogous to a convergence map from a lensing analysis.

[TODO] cite papers that find DM peaks!

3.3 Finding the offsets and its population distributions

We computed the projected offsets between the galaxy population and the DM populations of each cluster . Furthermore, to avoid undesirable effects from setting hard boundary at offset = 0 when plotting the PDF of the offsets, we randomly assign the sign of an offset before smoothing the PDF. Finally, after smoothing the PDF, we take the absolute magnitude of the offset as a comparison of how taking absolute magnitude affects our estimates. We provide the best estimate of the offset using the biweight statistic [CITE].

3.3.1 Peak matching

3.3.2 Possible populations of galaxy clusters

4 RESULTS

We summarize the main results here and leave the detailed tables of results in Appendix C.

4.1 Galaxy-DM Offset in Illustris

4.1.1 Projected offsets

From Fig. 6, we can tell that the population estimates.

The variance of the offsets due to projection is [TODO]. In particular for the most massive galaxy clusters with $10^{14} M_{\odot}$

- those between BCG, the most bound particle and the other masses.
- explain the variation of the offsets for the same cluster under different projections

While there is a tight distribution for Δs_{BCG} that peaks around zero, there are a non-negligible population of data points that lie outside this peak.

4.2 Correlations between different variables and the offsets

We subset the data and visually inspected the samples with the largest galaxy-DM offsets $\Delta s_{\rm KDE}$. We found that ...

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

- relaxedness
- mass
- richness

5 DISCUSSION

It is not easy to compare the result of this study to other study due to the differences in the multi-step method for inferring the "peak", or the "centroid".

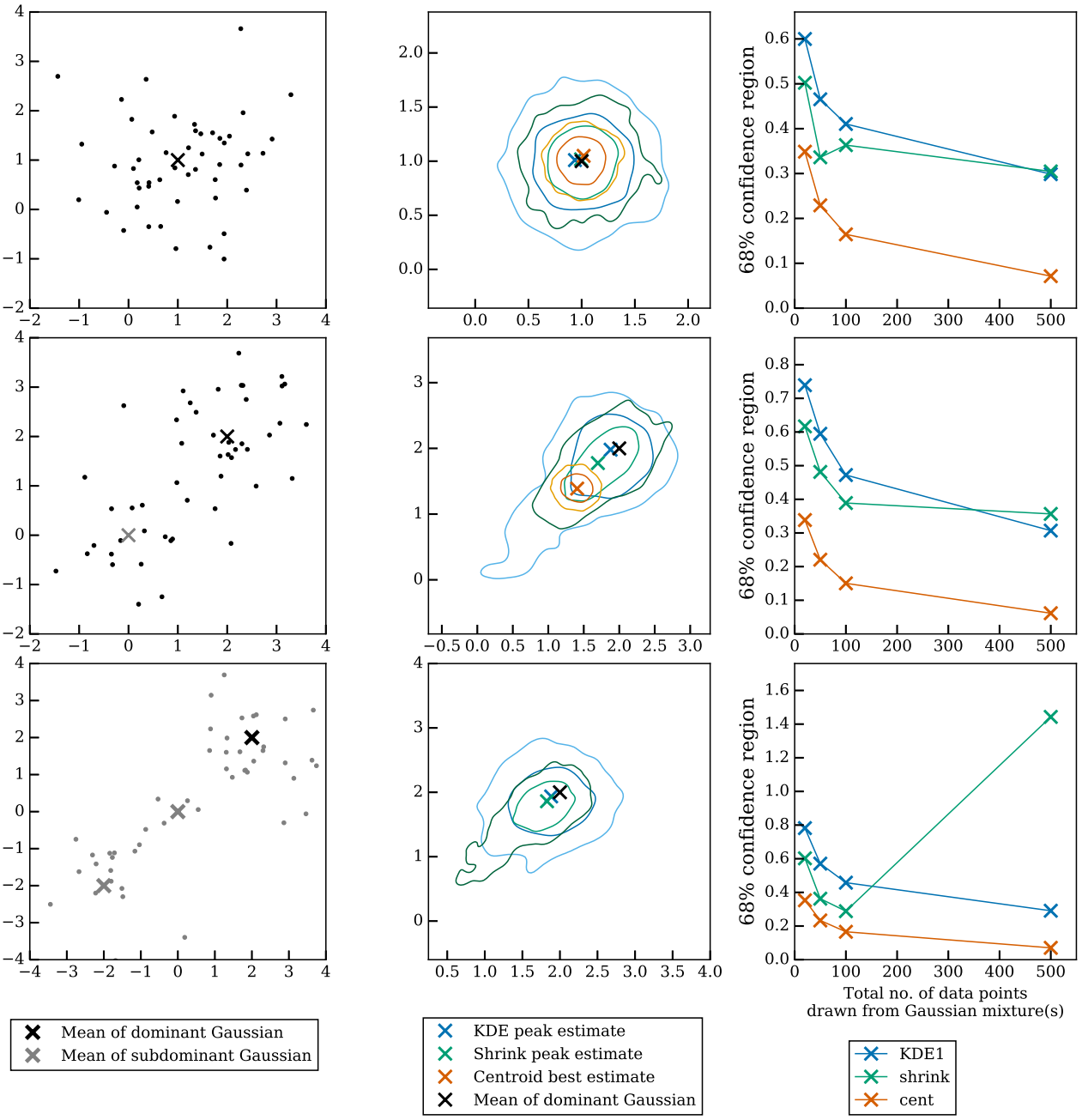

Figure 4. Comparison of peak finding performances of different methods by drawing data points (i.e. 20, 50, 100, 500) from known number of Gaussian mixtures. Panels from the top row contain data drawn from a single Gaussian mixture. The panels from the middle row contain data from two Gaussian mixtures with weight ratio = 7:3. The panels from the bottom row contain data drawn from three Gaussian mixtures with weight ratio = 55:35:10. The left column shows how 50 data points drawn from the fixed number of Gaussian mixtures look like. Due to the statistical nature of this exercise, we sampled the data and performed the analyses [TODO: state how many times] many times to create the (68% and 95%) confidence contours of the estimates in the zoomed-in view of the data in the middle column. The rightmost column shows how the size (median contour radius) of the confidence regions vary as a function of the number of drawn data points from the Gaussian mixtures. From the middle and the rightmost column, we can tell that the KDE peak estimate is the most accurate but less precise for estimating the sampled data from each set of data.

5.1 Other findings from the visual inspection of the simulated galaxy clusters

From the high resolution visualization of the DM maps (two-dimensional histograms with 2 kpc bins) in Fig. 5, we can tell that some clusters clearly possess multiple subclusters with visible separations between the subclusters, e.g. cluster 12 visualized on the panels on row 3 of the plot. However, there are also many clusters that contain only one main component, or several closely packed components as indicated by several significant peaks. This illustrates why finding a 'center' or peak of a cluster is an ill-defined notion. These statistics are only unambiguous for symmetric, relaxed clusters.

Furthermore, we show a lower-resolution visualization of the

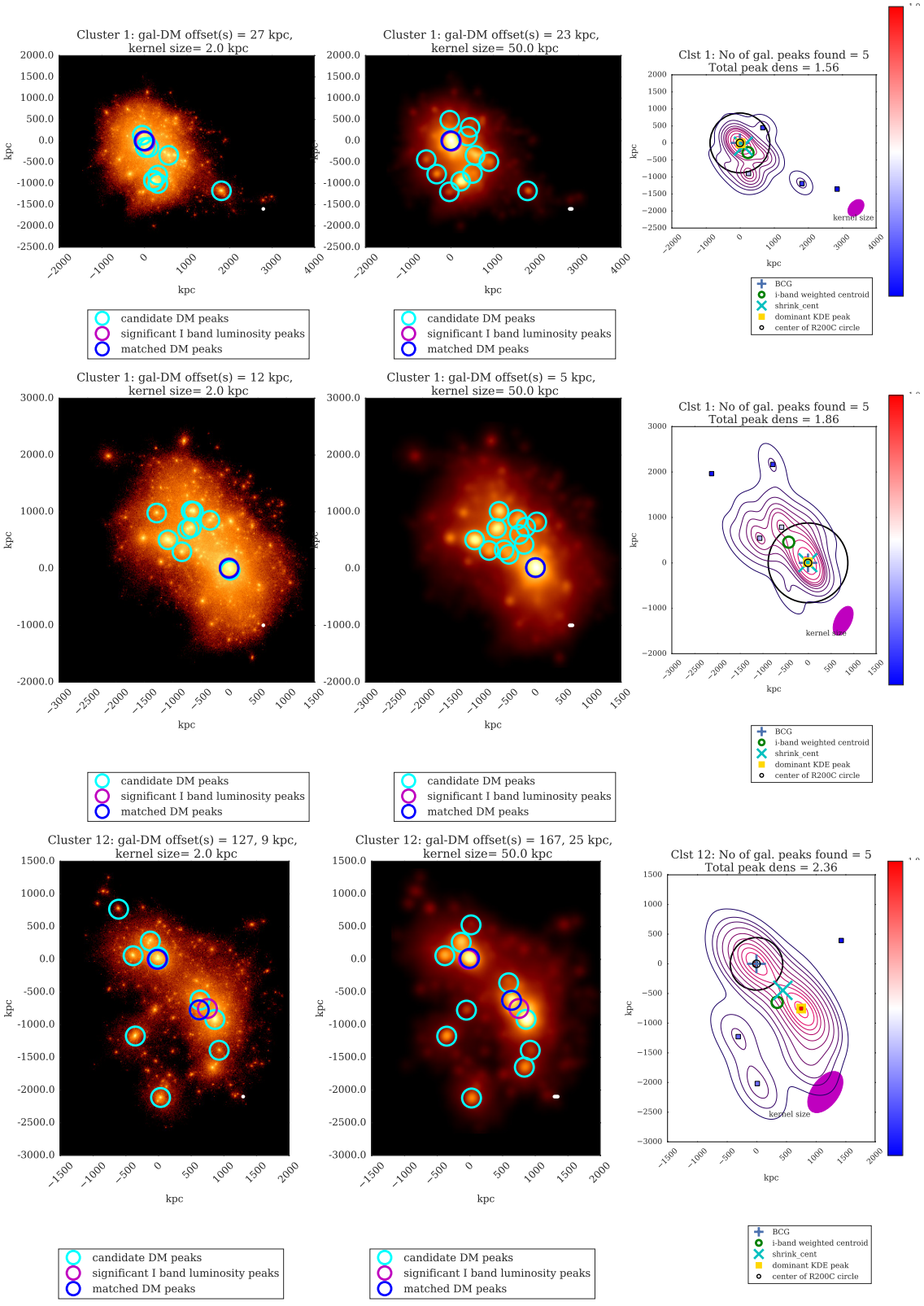

Figure 5. Visualization of clusters (each row is for the same projection of the same cluster). **Left column:** Projected density distribution of DM particle data (density overlay). The identified density peaks are indicated by colored circles. **Middle column:** The same DM projection but with treated with a 50 kpc smoothing kernel (kernel size indicated by white dot on lower right of the figure. Note that the thickness of the dot may be larger than 2 kpc for the plots on left hand column). **Right column:** Projected galaxy kernel density estimates (KDE) of the i-band luminosity map for the member galaxies of the same clusters. Each contour denotes a 10% drop in density mass starting from the highest level in red. Each of the magenta ellipse on the bottom right corner of each plot show the Gaussian kernel matrix H from eq. (4) and the orientation. See http://goo.gl/WiDijQ and http://goo.gl/89edcM for the visualization of the selected clusters inside two Jupyter notebooks.

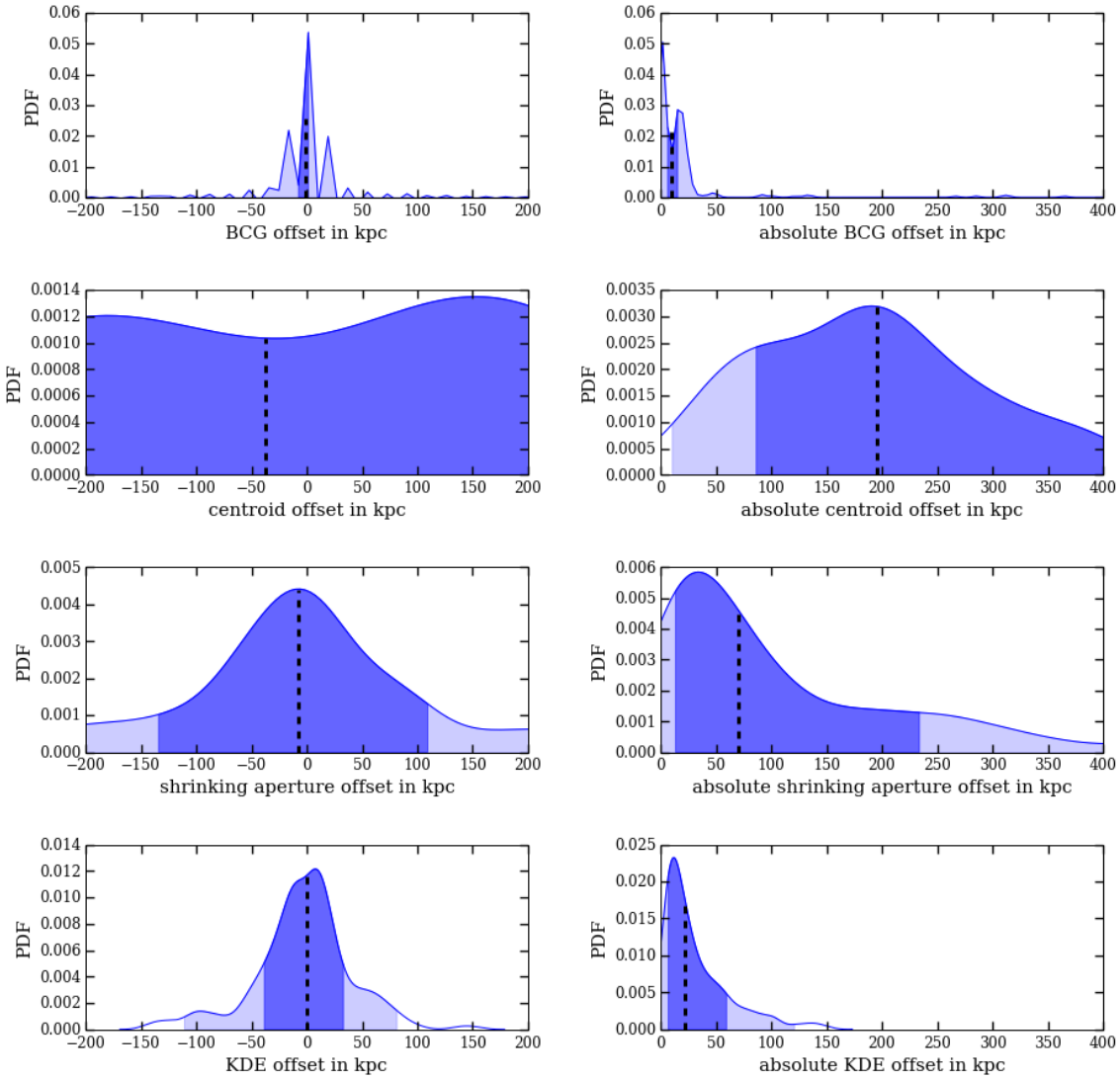

Figure 6. The distribution of different offsets of [TODO] clusters with [TODO] projections. The dark blue area indicates the 68% confidence interval while the light blue area shows the 95% confidence interval. We provide two ways of summarizing the offsets, the **left column** shows the offsets when we randomly denote the sign of the offset. The direction of the offset in the Illustris simulation without SIDM has no physical meaning. The estimates of the offsets on the left are all consistent with 0 within the 68% confidence interval. On the **right column**, we plot *the same data* after taking the absolute magnitude. The estimates from the absolute magnitude of the offsets are pushed towards larger values due to the cutoff at zero from taking the absolute values. None of the estimates from the absolute offsets is consistent with 0 within the 68% interval.

DM surface map in Fig. [TODO] compared to the higher resolution map, we can see a clear shift in the peak location. This illustrates that peak / center finding is also subject to noise from the data.

5.2 Comparison to other simulations

5.2.1 Comparison to other cosmological simulations

Assume dynamical equillibrium

Cui et al. (2016) X-ray center, BCG

kernel smoothing approach use minimum potential position as a proxy for the maximum SPH density

Voronoi Tessellation Density (VTD)

gives no confidence estimate for comparison. width of histograms can affect the distribution taking the absolute magnitude of the offset can bias the data

Harvey et al. (2014) consistent with zero offsets galaxies prescribed to the DM simulation and not simulated to see how the surrounding environment

[TODO] address the point about how many substructures is needed to average out the offset

They have only verified that their model of SIDM work for $\sigma_{\rm SIDM} = 0$ but not any model with $\sigma_{\rm SIDM} > 0$

5.2.2 Comparison to other staged simulations

Robertson et al. (2016)

5.2.3 *Estimation of the SIDM cross section from the Illustris offsets*

Since the Illustris simulation assumes CDM, it is interesting to see if we can infer any non-zero $\sigma_{\rm SIDM}$ from the offset estimates based on different methods.

5.3 Comparison to other observational studies

Weak lensing observations infer mass distribution of both baryonic and dark matter.

Central galaxy paradigm (CGP)

- Ford et al. (2014) paper about miscentering in CFHT and the difference between taking the absolute magnitude of an offset or not

There are many aspects of the analysis that is not covered by this study that are performed for analyzing observational data, such as

- galaxy membership identification along the line of sight
- removal of foreground galaxies

that are important for calculating the $\sigma_{\rm SIDM}$ with using a galaxy-DM offset.

5.4 Galaxy-DM Offset in observations of Merging Galaxy Clusters

how the offsets will translate to a $\sigma_{\rm SIDM}$.

5.5 How to use Δs to constrain $\sigma_{\rm SIDM}$

As a fast, preliminary step to show that σ_{SIDM} is plausible, one can compute the two-tail p-value for offsets computed with the *same methodology* based on the marginal distribution in table C1 from studies like this one.

However the p-value can only tell you given there is no SIDM, how unlikely it is to see the level of Δs that we see.

To properly infer σ_{SIDM} as a parameter estimation problem and account for the uncertainty, one should produce a simulation suite with a fixed $\sigma_{\rm SIDM}$ and compute the marginal likelihood of fitting the offset distributions with observations. Proper marginalization of nuisance parameters may include projections, mass and relaxedness etc. Extremely computationally intensive because nobody has shown that painted on galaxy population is realistic enough that the physical effects are sufficient.

Machine learning methods to paint galaxies to DM halos http://arxiv.org/abs/1510.07659

6 SUMMARY

We showed that

- the peak finding method from KDE for the density of cluster galaxies is the second 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.
- the resolution of the DM distribution can affect individual estimates Δs but do not show significant bias for the population estimate. However, the lower the resolution, the higher the variance of the population estimate.

Furthermore, we have provided a set of python functions for making accurate contour levels for spatial maps and inferring density peaks at https://goo.gl/MNrSQV.

While this paper does not provide a solution to the complete statistical model for galaxy clusters, it points out some of the aspects that good models should incorporate, especially when the utility of studying the cluster is to constrain $\sigma_{\rm SIDM}$.

Aspects that should be handle with care include:

- the mass profile, especially when the cluster is multimodal
- the evaluation of galaxy membership and the corresponding luminosity peak(s)

7 ACKNOWLEDGEMENTS

Part of the work before the conception of this paper was discussed during the AstroHack week 2014. KN would like to thank Phil Marshall and Jake Vanderplas for preliminary discussions for analyzing galaxy clusters. Part of this work was performed under HST grant (TODO ask Dave for grant number).

REFERENCES

Bradač M., et al., 2006, ApJ, 652, 937
Cui W., et al., 2016, MNRAS, 456, 2566
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Deng H., Wickham H., 2011, Technical report, Density estimation in R. had.co.nz
Duong T., 2007, J. Stat. Softw., 21, 1
Feigelson E. D., Babu G. J., 2014, Contemp. Phys., 55, 126
Ford J., et al., 2014, MNRAS, 447, 1304
Genel S., et al., 2014, MNRAS, 445, 175
Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Hall P., Marron J. S., Park B. U., 1992, Probab. Theory Relat. Fields, 92, 1
Harvey D., et al., 2014, MNRAS, 441, 404
Hoag A., et al., 2016, arXiv Prepr., p. 48
Kahlhoefer F., Schmidt-Hoberg K., Frandsen M. T., Sarkar S., 2013, MNRAS, 437, 2865
Markevitch M., Gonzalez a. H., Clowe D., Vikhlinin A., Forman W., Jones C., Murray S., Tucker W., 2004, ApJ, 606, 819
Medezinski E., et al., 2013, ApJ, 777, 43
R Core Team 2014, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/
Robertson A., Massey R., Eke V., 2016, arXiv Prepr., p. 20
Springel V., 2010, MNRAS, 401, 791
VanderPlas J., Connolly A. J., Ivezic Z., Gray A., 2012, in 2012 Conf. Intell. Data Underst.. IEEE, pp 47–54, doi:10.1109/CIDU.2012.6382200, http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6382200
Vogelsberger M., et al., 2014a, MNRAS, 444, 1518
Vogelsberger M., et al., 2014b, Nature, 509, 177

Table C1. Offsets for the full sample of 43 clusters at 2 kpc resolution.

Offset (kpc)	Location	68% CI†	95% CI†
BCG	-0.0	-7.486 , 10.35	-123.4 , 848.7
Weighted centroid	-37.19	-344.8 , 246.1	-806.0 , 567.9
Shrinking aperture	-7.191	-134.0 , 113.7	-368.4 , 343.8
KDE	-0.0	-38.38 , 33.79	-110.5 , 82.13

† CI stands for the credible interval centered on the biweight location estimate, with 68% of the probability density contained in the 68% credible interval.

Zitrin A., Menanteau F., Hughes J. P., Coe D., Barrientos L. F., Infante L., Mandelbaum R., 2013, ApJ, 770, L15

APPENDIX A: GETTING UNIQUE 2D PROJECTIONS OF THE CLUSTERS

In 3-dimensional space, rotation operations are non-commutative. We first actively rotate our clusters by the azimuthal angle ϕ before we rotate the particle according to the elevation angle ξ. Then we project to the transformed x-y plane. With this rotation scheme, two projections are identical if

$$|\phi_1 - \phi_2| = \pi \tag{A1}$$

It represents viewing the same cluster from opposite sides. To save disk space, we only compute the statistics from one of the two unique projections.

APPENDIX B: ALGORITHM OF THE SHRINKING APERTURE ESTIMATES

Data: subhalo that satisfy cuts as a galaxy

initial aperture centroid = mean galaxy location in each spatial dimension
distance array = euclidean distances between initial aperture center and each galaxy location
aperture radius = 90th percentile of the distance array
while *(newCenterDist - oldCenterDist) / oldCenterDist* $\geq$ *2e-2* **do**
    new data array = old data array within aperture
    newCenter = mean value of new data along each spatial dimension
end

Algorithm 1: Shrinking aperture algorithm

APPENDIX C: TABLE OF RESULTS

.

Table C2. Absolute offsets for the full sample of 43 clusters at 2 kpc resolution.

Offset (kpc)	Location	68% CI†	95% CI†
BCG	9.353	6.337 , 19.82	-2.656 , 1292.0
Weighted centroid	195.2	86.41 , 402.7	10.59 , 797.5
Shrinking aperture	70.07	13.44 , 236.4	-29.63 , 446.7
KDE	21.7	6.886 , 59.93	-3.722 , 121.3

† CI stands for the credible interval centered on the biweight location estimate, with 68% of the probability density contained in the 68% credible interval.

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