Mon. Not. R. Astron. Soc. **000**, 1–19 (2015) Printed 4 July 2016 (MN LATEX style file v2.2)

Did we see self-interacting dark matter or statistical noise? The galaxy-dark matter offsets of 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

Assuming that dark matter has a zero interacting cross section, how likely is it for us to see the offset values between dark matter and galaxies from real data? This paper formulates a test using cluster data in the cosmological simulation, the Illustris simulation, to examine that hypothesis. We examine the uncertainties of the different summary statistics, and see if the spatial distribution of galaxy closely follows those of the Dark Matter. 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 mass, 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. A lot of observational studies focus on using merging galaxy clusters as they assume the high collisional velocity should further increase the chance of detecting the effects of SIDM. By assuming galaxies being relatively collisionless $\sigma_{\rm gal} \approx 0$, any offset of the DM population from the galaxy provides $\sigma_{\rm SIDM}$ relative to $\sigma_{\rm gal} \approx 0{\rm cm}^2$. These observational studies include Markevitch et al. (2004) and Bradač et al. (2006) reporting an offset of 25 kpc for the Bullet Cluster; Dawson (2013) reporting an offset of 129 kpc and 47 kpc for the southern and the northern subcluster respectively; Jee et al. (2015) reporting an offset of 190 kpc for MACSJ1752, and others that we list in detail in table [TODO]. However, other studies using 129 X-ray selected relaxed galaxy clusters, such as George et al. (2012) also report offsets of the same order of magnitude, between $50 - 150$ kpc.

On the other hand, there are many staged simulations of mergers of galaxy clusters that focused on detecting the signal from SIDM. These staged simulation usually have parametric prescriptions of the spatial distribution of galaxies (Randall et al. 2008, Kahlhoefer et al. 2014, Robertson et al. 2016), such as an NFW profile, and do not have realistic galaxy features, nor dynamical frictions. They try to show the magnitude of offsets solely due to SIDM (Kahlhoefer et al. 2014) as by initializing the galaxy-DM offset to be zero at the beginning of their simulations. Furthermore, these staged simulations commonly use a much higher number of galaxy particles than the realistic observable number of galaxies. Randall et al. (2008) found an offset of only 1.8 kpc in the staged merger simulation with $\sigma_{\rm SIDM} = 0{\rm cm}^2/$ g using 10^5 galaxy particles When assumed with zero impact parameter for mergers, Kim and Peter et al. (2016), using 5.7k or 57K galaxy tracer particles also show null galaxy-DM offset during most periods of their control staged simulation with $\sigma_{\rm SIDM} = 0{\rm cm}^2/$ g. While we provide a more in-depth comparison with these staged simulation in the discussion, we argue these staged simulations do not probe how statistical and observational uncertainties realistically contribute to the galaxy-DM offsets. As such, any offsets from aforementioned staged simulations when they increased $\sigma_{\rm SIDM}$, they can guarantee the offsets are maximally due to SIDM. When these simulations set the $\sigma_{\rm SIDM}$ to observationally motivated levels of < 3 cm^2/g, different authors have consistently reported offset signals on par with uncertainties estimated from individual observations. These simulations have raised questions about how strongly the galaxy-DM offsets can constrain the effects of SIDM. When Kahlhoefer et al. (2014) simulated SIDM with both low-momentum-transfer self-interaction and rare self-interactions of DM with high momentum transfer, they found maximum offsets that are < 30 kpc for $\sigma_{\rm SIDM}$ as high as 1.6 cm^2 / g. The reported offset from Randall et al. (2008) for $\sigma_{\rm SIDM} = 1.24{\rm cm}^2/{\rm g}$ is only 53.9 kpc. While [TODO] Kim and Peter et al. (2016) found a maximum offset < 50 kpc for $\sigma_{\rm SIDM} = 3{\rm cm}^2/{\rm g}$, and Robertson et al. (2016) also found a max-

imum offset $\lesssim 50$ kpc from a simulation suite of a Bullet Cluster analog with $\sigma_{\rm SIDM} = 1\ {\rm cm}^2/\ {\rm g}$.

An alternative explanation for the observed galaxy-DM offsets is due to statistical and observational uncertainties. Galaxies are very sparse samples of the underlying DM overdensities, it is possible that the summary statistic of the sparse sample be different from those of the underlying distribution. It is not clear if there is any physical cause of the galaxy-DM offset in a CDM universe, but any statistical noise leading to an offset can influence this method of the inference of $\sigma_{\rm SIDM}$. Since the Illustris simulation assumes no SIDM, this study is complementary to staged simulation for understanding what can contribute to the offsets. Simply put, we perform a hypothesis test with the galaxy-DM offsets in the Illustris simulation directly corresponding to our null hypothesis $\mathcal{H}_0$, with:

$$\begin{cases} \text{the null hypothesis } \mathcal{H}_0 : \text{Cold Dark Matter (CDM)} \\ \text{the alternative hypothesis } \mathcal{H}_1 : \text{Self-interacting Dark Matter (SIDM)} \end{cases} \tag{1}$$

This exercise is further complicated by the fact that there is no theoretical foundation showing which observable would be the most sensitive to each possible type of SIDM. In fact, Kahlhoefer et al. (2014) have argued that SIDM does not cause significant offsets between the galaxy and DM peaks, and only cause an offset between the corresponding centroids for a brief period of time after a merger. Popular choice for computing the offsets involves first inferring the summary statistic of the DM population and those of the galaxy population of a cluster independently before taking a difference. While there are well established procedures 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 member galaxy population.

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

$$\Delta \boldsymbol{s} \equiv \boldsymbol{s}_{\rm gal} - \boldsymbol{s}_{\rm DM}. \tag{2}$$

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 physical 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 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 test with a null hypothesis of $\sigma_{\rm SIDM} = 0$ and discuss the caveats.

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 $H_0 = 70.4\ {\rm km\ s^{-1}\ Mpc^{-1}}$.

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 within a reasonable observational limit, i.e. apparent $i \leq 24.4$ which is the limiting magnitude of the DEIMOS spectrometer on the Keck telescope, when we assume a cosmological redshift of $z = 0.3$ in the i band. The limiting magnitude of 24.2 in the F814W filter of the Hubble Space Telescope, and the limiting magnitude of the Canadian-Hawaii French Telescope of 24.5, are also close to our chosen limiting magnitude. 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. These simulated galaxy clusters (or groups) have masses ranging from $10^{13}\ {\rm M}_\odot$ to $10^{14}\ {\rm M}_\odot$.

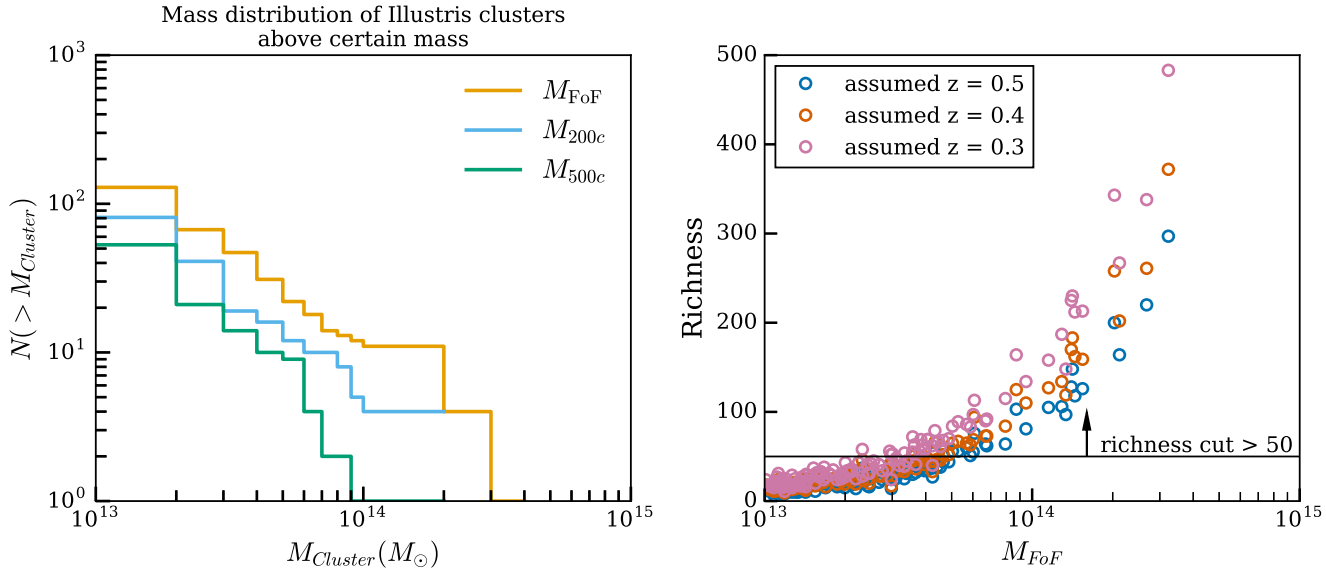

Figure 1. Left figure: Mass distribution of the group / cluster sized DM halos for different halo selection schemes. Mass estimates obtained by the FoF algorithm are labeled as $M_{\rm FoF}$. Masses centered on the most bound particle within a radius those the average density is 200 or 500 times the 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.

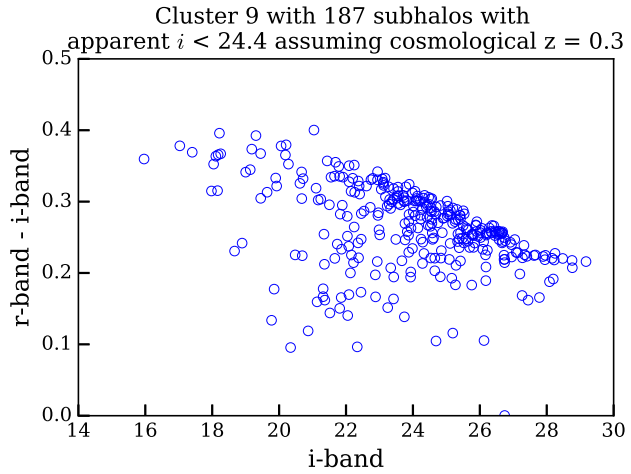

Figure 2. Color-magnitude diagram of one of the galaxy clusters that is selected for analysis. This cluster is the 9th most massive. The apparent magnitude is calculated assuming that the cosmological redshift (distance) is $z = 0.3$. We can see a clear overdense region that corresponds to a red-sequence.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

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 "non-relaxedness" is hard to retrieve from simulations across different saved states and is missing from observations. We quantify the state of the cluster by providing several quantitative definitions of non-relaxedness and see how they correlate with Δs. Some possible definition of non-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 lower the ratio, the less substructures there are in the cluster.
- the distance between the most bound particle from the center of mass as a function of R_{200c}. The smaller the distance, there are less asymmetric substructures.

TODO velocity dispersion from selected galaxies those selection criteria will be explained in

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.

The Pearson product-moment correlation coefficient (aka Pearson's r) of the first two relaxedness criteria for the 43 selected clusters is as high as 0.82.

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

Unlike in staged simulation, picking out a particular projection for a cluster does not always make physical sense. For highly symmetrical clusters, most projections are similar. However, for mergers

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	2.3
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

or asymmetrical clusters, there is no obvious choice for the plane of projection that can allow us to understand the cluster. There is no simplistic merger axis that should be projected along the plane of the field of view due to possible 3D spiraling trajectories of the merging components.

We therefore compute observables based on even sampling of angular orientation as our line-of-sight. As the order of projecting the data and estimating the summary statistic is non-commutative, we first project the data before estimating any projected observable. The computation of even angular orientation is done by using HealPy, which is a Python wrapper for **`HEALPix`** [1] (Gorski et al. 2005). Each line-of-sight centers on a **`HEALPix`** pixel. The number of projections that we employed is 768 for each cluster. Even though there are at least 2 identical projections for each cluster due to one possible line-of-sight from the front and one from the back, it does not affect any summary statistic. We do not remove the duplication as it breaks the rotational symmetry in the 2D plane when we try to compute the 2D population distribution of offsets. Details of the implementation of the projection is available in Appendix **??**.

2.3 Properties of the galaxies in Illustris clusters

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. A comparison of the projected DM map, the luminosity map and the number density map of 129 clusters can be found at https://goo.gl/kZUWrg, https://goo.gl/R7VNi9 and https://goo.gl/lmQUPd respectively. It is clear from the comparison that the luminosity maps computed from the KDE trace the DM distribution more closely than the number density maps.

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 the 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 apply 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 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. $\mathbf{x_i}$). Histogram is not good for peak estimate for *sparse* data for two reasons: 1) the choice of laying down the bin boundaries affects the count in each bin, 2) the choice of bin width also affects the count in the bin. Only when the available number of data points for binning is large, the estimates of histograms and smoothed density estimates are approximately the same. The number of member galaxies (< 500) is sparse enough for the uncertainty introduced by histogramming to bias our peak estimate. For the density estimate of galaxy luminosity, we adopt

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

a Gaussian kernel. 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 bandwidth of the smoothing kernel, which takes the form of a matrix in the 2D case. 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 too many peaks due to noise. 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 $\hat{f}$:

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

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 and w_i is again the $i-$band luminosity weights for each galaxy. 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 parameter.

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 magnitude 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 ($\lesssim 2$ second per projection per cluster).

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 and compute the offset. The exact procedure is discussed in section 3.3.

For each projection of each cluster, we normalize the density of all significant luminosity peaks to those of the brightest peak. Then we sum the density of all the peaks for a cluster and call this value ν. When the value of ν gets bigger than 1, it indicates the presence of projected substructures for that particular projection.

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 peak of the luminosity map. 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 extent of the aperture, or to partition the data with another (statistical) algorithm. Furthermore, the convergence of results of this method is unstable. 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 the pseudo-code can be find in Appendix A.

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 of the cluster galaxy population by several orders of magnitude. However, star formation can cause less massive galaxies to be brighter in the bluer photometric bands. To avoid star formation from biasing our algorithm for identifying the BCG, we find the brightest galaxies in redder bands i.e. the r, i, z bands and found that they give consistent results for all selected clusters. We used the $i-$band to pick the BCG for the plots and the final results.

3.1 Comparison of the methods from Gaussian mixture 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 source galaxies that are being lensed, 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. (2016) found a resolution of 11 arcseconds can best fit

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.

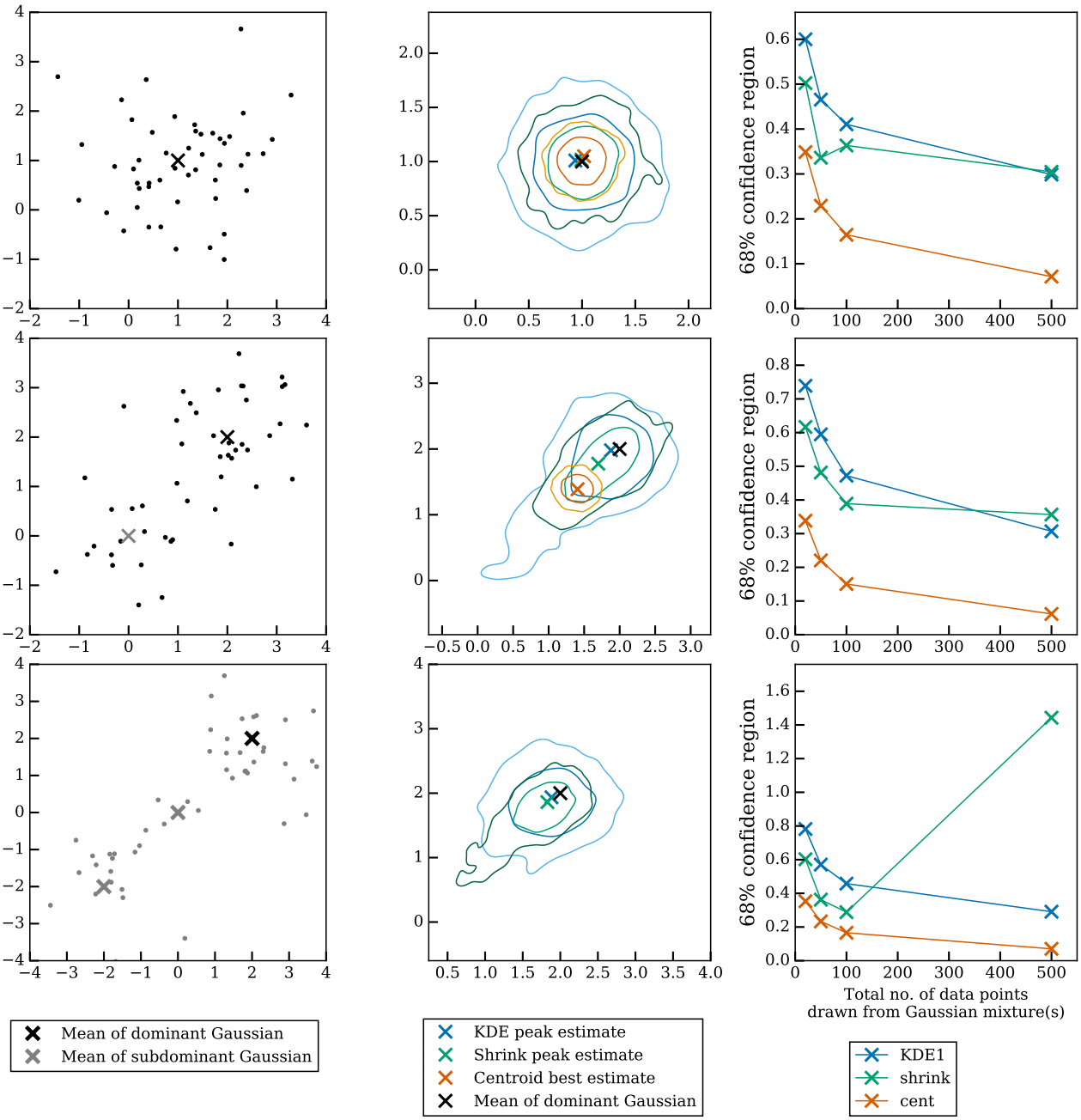

Figure 4. Comparison of peak finding performances of different methods by drawing data points (i.e. 20, 50, 100, 500) from known number of Gaussian mixtures. Panels from the top row contain data drawn from a single Gaussian mixture. The panels from the middle row contain data from two Gaussian mixtures 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% Monte Carlo 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. [TODO]

the data. This kernel translates to a physical size of 50 kpc assuming a cosmological redshift of $z \approx 0.3$. To compute a DM spatial distribution, we first make histogram with 2 kpc $\times$ 2 kpc bin size which is larger than the DM softening length of 1.4 kpc. After that, we use a Gaussian smoothing kernel of the DM histogram Illustris DM particle data. We do not perform a cross-validated KDE that has $O(n^2)$ runtime on the DM data because the number of DM particles for each cluster is of [TODO double check] the order of millions. The DM resolution is high enough for the histograms to be accurate. Physically, the histograms of the dark matter of each cluster is analogous to a convergence map from a lensing analysis.

3.3 Finding the offsets

It is possible to have several peak estimates from the KDE of the member galaxy population of a cluster. From the density estimate at each peak, we can sort the peaks according to their densities. We only match make use of luminosity peaks that are at least 20% as dense as the brightest galaxy-luminosity peak to avoid computing the offsets of spurious substructures, such as the peaks due to small number of galaxies that are located far away from the main concentration of mass.

In general, there are many more DM peaks because there are many more dark subhalos than galaxies for each cluster and the resolution of the DM data is much higher. To find the nearest match to the significant galaxy peaks, we construct a k-dimensional tree (KD-Tree; in our case, k = 2) using the densest $n_{\rm DM}$ number of DM peaks:

$$n_{\rm DM} = \begin{cases} 3 \times (n_{gal} + 1) & \text{if } n_{gal} < 3 \\ 3 \times n_{gal} & \text{if } n_{gal} \geq 3. \end{cases} \quad (5)$$

where n_{gal} is the number of significant galaxy peak, and $n_{\rm DM}$ is the number of peaks that went into the construction of the KD-tree. When there are more than one dense galaxy peaks located far away from one another, the top few densest DM peaks (subhalos) can be located around the same galaxy peak. i.e. there is no one-to-one matching between the luminosity of galaxies and the density of detected DM peaks. Matching purely based on density and luminosity leads to larger offsets. From inspection, using eq. (5) works well to match the appropriate peaks. After identifying the DM peaks, we also compute the offsets between the DM peaks, and the following spatial estimates, including 1) the most bound particle 2) the shrinking aperture peaks, 3) the number density peaks, 4) the BCGs, and 5) the luminosity weighted centroid.

Since there can be more than one number density peak from the corresponding KDE map, we also use a KDE tree to location the closest peak to the identified DM peak.

[TODO] Figure out how the Illustris simulation figure out the most bound particle. The most bound particle is the location with minimum gravitational potential of the **`Subfind`** identified cluster. Due to substructures, it is possible for there to be several minima of similar gravitational potential level.

3.4 Constructing the non-parametric hypothesis test

After matching the peaks, we use the offsets as the basis of our non-parametric hypothesis test. We compute the p-value as the highest density interval of simulated offsets that are below observed values of offsets in the literature. This gives us a rough estimation of the probability of seeing the offset from real observations under the null hypothesis of CDM being true. We also provide the biweight location and the density interval characterizing the distribution of offsets computed from each of the listed methods. For instance, the 95% interval is computed as the density mass that encompasses 95% of total density (i.e. 2.5% of mass at each end of the tail are excluded.) centered on the location estimate for the symmetrical distribution of Δy.

The different representations of Δs have different statistical power for the hypothesis test, i.e. the spread of the offset distribution in different representation differ so it affects the significance of an observation. The most faithful representation of the offsets without any information loss is:

$$\Delta\mathbf{s} = (\mathbf{x}_{\rm gal} - \mathbf{x}_{\rm DM}, \mathbf{y}_{\rm gal} - \mathbf{y}_{\rm DM}). \quad (6)$$

The PDF in 6 peaks at (0, 0) when there is no real offset. However, when one takes the magnitude of $\Delta\mathbf{s}$, i.e.:

$$|\Delta\mathbf{s}| = \sqrt{(\mathbf{x}_{\rm gal} - \mathbf{x}_{\rm DM})^2 + (\mathbf{y}_{\rm gal} - \mathbf{y}_{\rm DM})^2}, \quad (7)$$

the resulting 1D distribution of $|\Delta\mathbf{s}|$, those support being $[0, \infty)$, will not peak at zero even if the original distribution of $\Delta\mathbf{s}$ peaks at (0, 0). On the other hand, the 1D distributions along a particular spatial axis, e.g. $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$, each with a support of $\mathbb{R}$, will not exhibit a discontinuity when the offset is zero. Also, since we have enough samples for there to be rotational symmetry for the distribution of $(\Delta x, \Delta y)$, it does not matter if we picked Δx or Δy for the 1D representation.

Due to the asymmetry of representing the offset as $|\Delta\mathbf{s}|$ that can mislead readers to think that there is an offset when there is none , we compute the hypothesis test significance level with the offset $\Delta\mathbf{x}$ along one of the spatial axes. We also provide a table of statistic of different ways of representing the offsets.

4 RESULTS

4.1 Offset between galaxy summary statistic and the most bound particle

As a sanity check, we computed the offsets between different galaxy summary statistic and the most gravitationally bound particle (aka most bound particle). The ranking in terms of increasing distance to the most bound particle computed by different method is as follows:

- BCG
- densest peak of the luminosity map created by weighted the KDE
- shrinking aperture center from the luminosity weighted galaxy data
- densest peak of the number density map created by the unweighted KDE
- centroid (center of mass) estimate using luminosity weights

Table B2 contains the different percentile and robust estimates of the distribution of offsets from the most bound particle. In fact, most of the BCG offsets are very small except for two clusters with ID 13 and 33. Both clusters have all the values of $\nu > 1.5$ over different projections and from the projected visualization, we confirm that both clusters have significant substructures. It is therefore possible for the most bound particle to have similar gravitational potential level as another substructure where the BCG is located.

4.2 Relaxedness of the clusters

Out of the $43 \times 768 = 33024$ projections, $\sim 45\%$ of the projections have only one dominant luminosity peak, with the total peak density

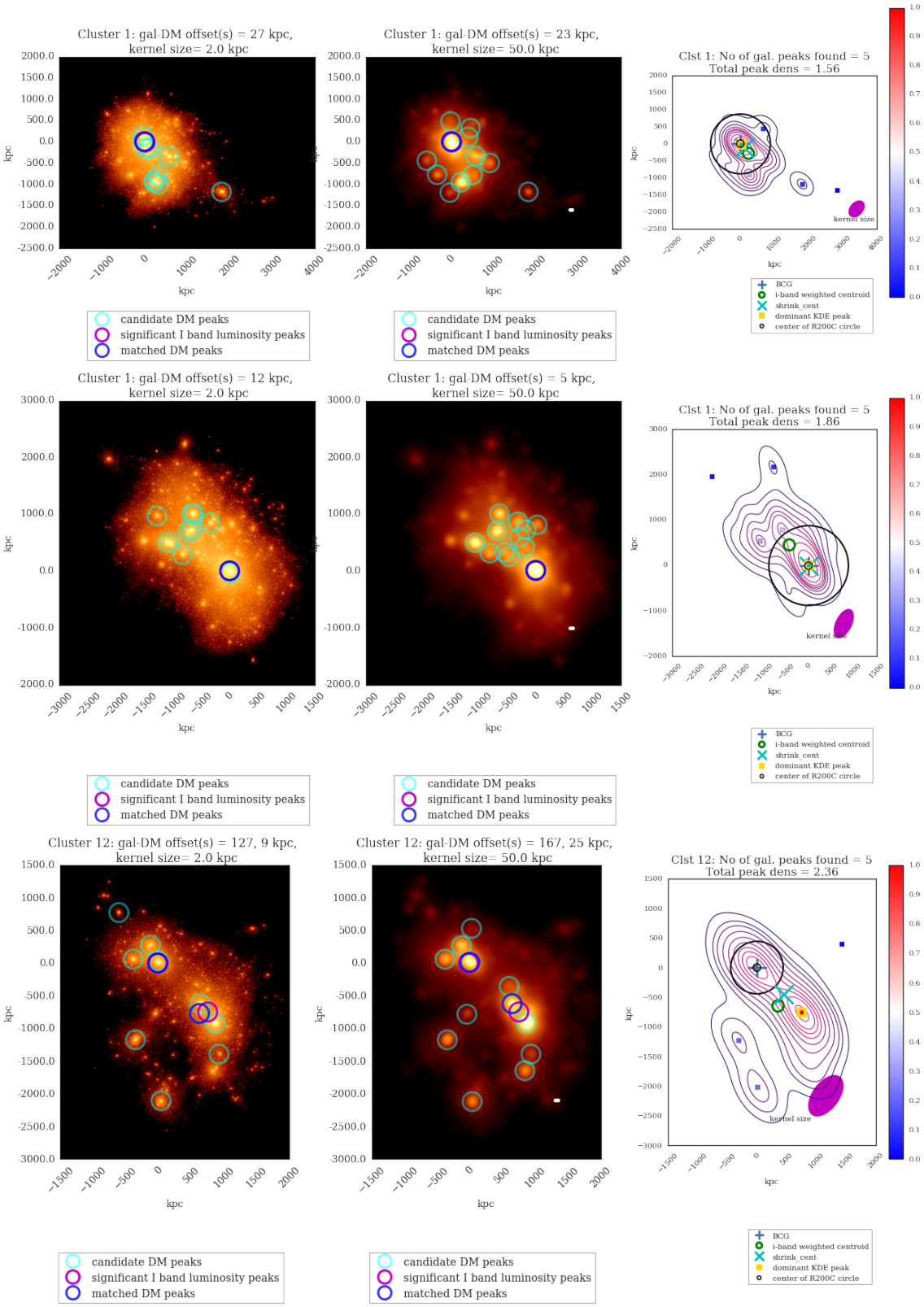

Figure 5. [TODO merge margin between left and middle panels] Visualization of clusters (each row is for the same projection of the same cluster). **Left column:** Projected density distribution of DM particle data (density overlay). The identified density peaks are indicated by colored circles. **Middle 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 colored 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). The big black circle is centered on the most bound particle as identified by **`SUBFIND`** and the radius of the circle indicates the three-dimensional region in which the average density is 200 times the critical density of the universe (a.k.a. R_{200C}). See http://goo.gl/WiDijQ and http://goo.gl/89edcM for the visualization of the selected clusters inside two Jupyter notebooks.

Table 2. Robust estimates and the distribution of offsets along the y-axis (This is different from the magnitude which has discontinuity at zero).

sample	offset (kpc)	location	lower 68%	lower 95%	lower 99%	upper 68%	upper 95%	upper 99%
all ν	$\Delta y_{\rm BCG}$	0	-3	-22	-496	3	456	1449
all ν	$\Delta y'_{\rm KDE}$	0	-25	-79	-127	25	79	126
all ν	$\Delta y_{\rm num.dens}$	0	-84	-303	-693	84	302	691
all ν	$\Delta y'_{\rm shrink}$	0	-65	-295	-652	65	295	655
$\nu < 1.2$	$\Delta y_{\rm BCG}$	0	-3	-10	-19	2	9	19
$\nu < 1.2$	$\Delta y'_{\rm KDE}$	0	-18	-48	-82	18	48	83
$\nu < 1.2$	$\Delta y_{\rm centroid}$	0	-108	-255	-395	108	254	394
$\nu < 1.2$	$\Delta y_{\rm num.dens}$	0	-73	-195	-303	73	195	302
$\nu < 1.2$	$\Delta y'_{\rm shrink}$	0	-51	-187	-285	51	187	285
$1.2 < \nu < 2.2$	$\Delta y_{\rm BCG}$	0	-3	-160	-684	4	807	1570
$1.2 < \nu < 2.2$	$\Delta y'_{\rm KDE}$	0	-32	-89	-125	32	89	124
$1.2 < \nu < 2.2$	$\Delta y_{\rm centroid}$	0	-262	-663	-905	262	663	904
$1.2 < \nu < 2.2$	$\Delta y_{\rm num.dens}$	0	-87	-299	-739	87	298	738
$1.2 < \nu < 2.2$	$\Delta y'_{\rm shrink}$	0	-85	-386	-777	85	386	779

Table 3. Observed offsets from bimodal clusters with confirmed major mergers.

Cluster	Δs (kpc)	galaxy peak	DM peak	p-value	subcluster	mass (10^{14} $M_\odot$)	reference
MACS J0025.4-1222					southeast	2.5	Bradač et al. 2008
MACS J0025.4-1222					northwest	2.6	Bradač et al. 2008
DLSCL J0916.2+2951	129	number density	weak-lensing		southern	3.1	
DLSCL J0916.2+2951	47	number density	weak-lensing		northern	1.7	
Abell 520							
ACT-CL J0102-4915	570	number density	weak-lensing		southeast	11	Jee et al. 2014
CIZA J2242.8+5301	∼200	number and luminosity	weak-lensing		both subclusters		Jee et al. 2015

of the projection being $\nu \leq 1.2$, while $\sim 50\%$ of the projections have approximately two dominant luminosity peaks with $1.2 < \nu < 2.2$. From B1, most clusters show some kind of secondary peak, only 7 clusters (with ID 15, 16, 17, 22, 31, 35, 51) out of 43 clusters have $\nu \lesssim 1.2$ for most of the projections. There is also very a strong correlation of ~ 0.8 between each of the two non-relaxedness quantities and the median total density of the luminosity peaks ν of each cluster.

4.3 Galaxy-DM Offset in Illustris

4.3.1 Two-dimensional (2D) offsets

The 2D distribution of Δs from most methods peak at around zero ($\lesssim 4$ kpc) with rotational symmetry, except the luminosity weighted centroid method with slight asymmetry along x-axis (See table B3). The spread of Δs computed by each method differ enough that one needs to know which offset method an astronomical study used for a fair comparison (See Fig. 6 for the distribution along the y-axis). The offset $\Delta s_{\rm BCG}$ peaks sharply near zero but contains outliers. Having outlier is possible because the DM peak is chosen as the closest DM peak to match the brightest luminosity peak in a particular projection. When there are distantly separated subclusters of similar masses, the brightest projected luminosity peak may shift from one subcluster to another subcluster between different projections, while the BCG identification is unchanged between projections.

For the full sample in table The offsets computed by the peak from the luminosity weighted KDE has the second smallest variance. The 68% percentile of $\Delta y'_{\rm KDE}$ is at around ± 25 kpc. Using shrinking aperture to estimate the peak location from the luminosity map increased the 68-th percentile of the offset to more than double those of $\Delta y'_{\rm KDE}$ at ± 65 kpc.

Our results from the samples with $\nu < 1.2$ also show that even for samples with one dominant component, substructures can still cause non-negligible biases of the peak estimate from the shrinking aperture method.

The peak estimate from the number density map has even larger variance (with its 68-th percentile at ± 84 kpc). The ranking of the variance for different methods stay the same even if we only consider projections with $\nu < 1.2$ and that we only consider the offsets from the dominant DM peak. The spread of the offsets inferred by each method will affect their ability for constraining $\sigma_{\rm SIDM}$. We will further elaborate on this point when we compare our results with staged simulations in section .

4.4 Offset projection uncertainty of each cluster

When we gather the offsets $\Delta s'_{\rm KDE}$ of the 768 projections for each cluster, we can find the offset uncertainty due to projection effects. The distribution of $\Delta y'_{\rm KDE}$ of each cluster is illustrated in the box plot of Fig. 7. We use a robust statistic, the biweight mid-variance to characterize the spread of the distribution instead of the standard deviation, which can be more susceptible to outliers. The biweight mid-variance of the $\Delta s_{\rm KDE}$ of half of the clusters is < 23 kpc. Only 4 clusters have mid-variance > 40 kpc.

4.5 Correlations between different variables and the offsets

Physical quantities that have positive correlation ~ 0.7 with the maximum $\Delta s_{\rm KDE}$ of each cluster include the two non-relaxedness criteria mentioned in section 2.1.1, The offset $\Delta s_{\rm KDE}$ also show a high correlation of 0.77 with ν. The FoF mass of clusters shows slight correlation of 0.28 with the median total density of luminosity peaks ν of each cluster but no significant correlation (0.14) with $\Delta s_{\rm KDE}$.

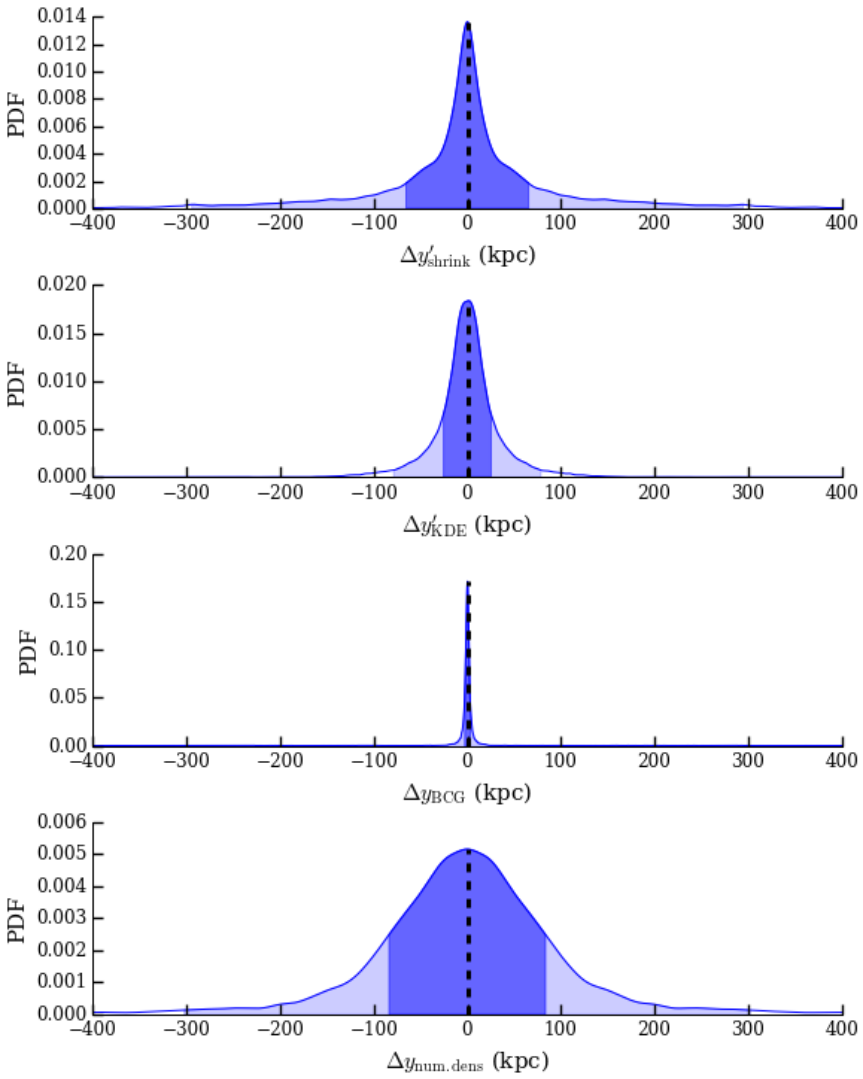

Figure 6. The distribution of different offsets of 43 clusters with all 768 projections. For estimates where several peaks of galaxy data are possible, only the densest peak is matched to the DM peak for computing the offsets in this figure. The dark blue area indicates the 68% density interval while the light blue area shows the 95% density interval. The table summarizing the statistic of each distribution is available in table

There is slightly negative correlation (~ -0.2) between the different mass measured within a certain density threshold, i.e. M_{200C} and M_{500C}, with $\Delta s_{\rm KDE}$ and ν. A lot of the mentioned correlations are consistent with the measurements of substructures in the cluster. Surprisingly, there is no significant correlation between the richness of the clusters and $\Delta s_{\rm KDE}$ and ν. This may be due to the fact that the peak estimate is only affected strongly by a few bright galaxies near the peak.

5 DISCUSSION

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

We inspected both the luminosity maps and the number density maps of the member galaxy populations. With the same selection of bright galaxies of apparent i-band< 24.4 at $z = 0.3$, the luminosity maps in general resemble the DM maps more closely than the number density maps. In real observations, missing selection of member galaxies, or foreground objects can both affect the inference of the galaxy spatial distribution. The number density map can be less susceptible to bias from bright foreground objects. It is less clear about the effect of missing member galaxies for computing the luminosity map because there is a selection bias favoring bright member galaxies.

5.2 Comparison to surveys of relaxed galaxy clusters

Using X-ray selected relaxed galaxy clusters, (George et al. 2012) reported Δs the order of 50 kpc to 150 kpc.

huge offset

galaxy centroid

5.3 Comparing the offset distributions with merging cluster observations

Our estimated distribution present in table 2 represents a lower bound of the spread of offsets. This is due to the high purity and

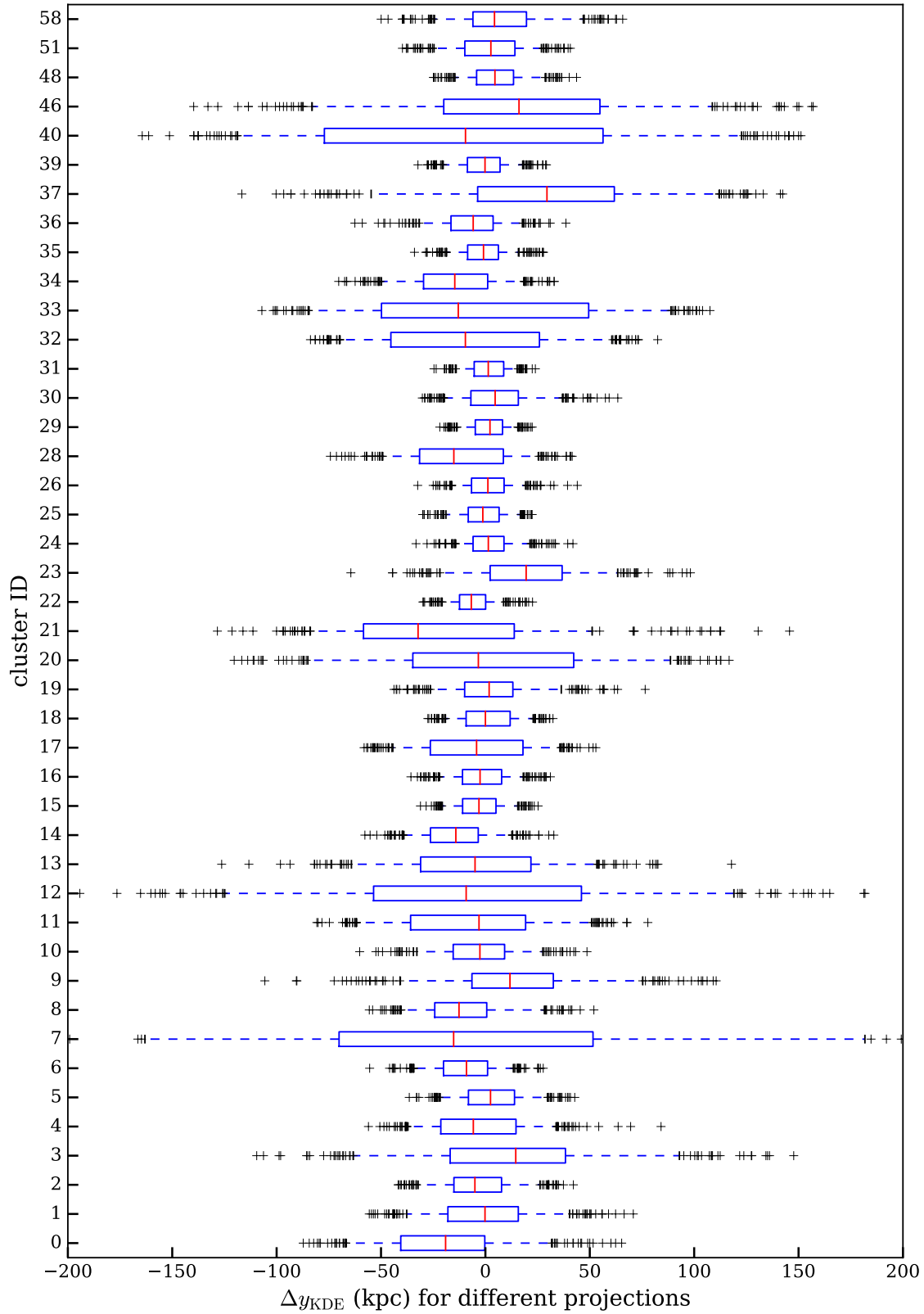

Figure 7. A box plot showing the distribution of projected offsets for each cluster based on 768 projections. The red line shows the median of the projections, the box encompasses the 25% and 75% percentile of the distribution while the whiskers mark the 5% and the 95% percentile. The other black crosses are data points with extreme values beyond the 5% and 95% percentile. The offsets were computed between the closest DM peak to the brightest luminosity peak of each cluster.

completeness of the Illustris data. Most observations of massive galaxy clusters with total mass of $M_\odot \approx 10^{14}$ have richness $\lesssim 300$, while the most massive Illustris cluster ($M_{FoF} = 3.23 \times 10^{14} M_\odot$) has a richness of 483.

We do not further divide the cluster sample based on the dynamical state because they are realizations of the possible dynamical states of clusters. Most of them are relaxed while a small number (~ 3).

If we restrict to comparison of clusters total mass $\sim 10^{14} M_\odot$, there is no clusters with offsets that go beyond our reported 95-th percentile interval.

The largest observed $\Delta s_{\rm num.dens} \approx 570$ kpc to date is from the southeastern (SE) subcluster of El Gordo (ACT-CL J0102-4915; Jee et al. 2014). And the number density peak in (Jee et al. 2014) is determined via a shrinking aperture algorithm (via private communication). The SE luminosity peak of El Gordo, however, is found to be much closer to the DM peak location of the lensing convergence map. El Gordo is also more much more massive ($10^{15} M_\odot$) than the Illustris clusters. Other large Δs from mergers are reported at the level of 200 kpc using the number density peak of CIZA J2242.8+5301 (Jee et al. 2015) . These offset level are within the 95% percentile level when we look at $\Delta s_{\rm num.dens}$. The reported luminosity peak offset $\Delta s \approx 200$ kpc, however, goes beyond the 99% level of our reported intervals. CIZA J2242.8+5301 is located close to the plane of the Milky Way with high chance of star contamination. Again, CIZA J2242.8+5301 is more massive (total mass $\sim 10^{15} M_\odot$) than the Illustris clusters.

5.4 Comparison to other simulations

5.4.1 Comparison to other cosmological simulations

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

5.4.2 Comparison to other staged simulations

5.5 Comparison to other observational studies

Caution when comparing to observational studies.

With high completeness and purity of member galaxy data, luminosity peaks

5.6 Offset between the BCG and the DM peak

Zitrin et al. (2012a) Zitrin et al. (2012b) Mohammed et al. (2014) miscentering lensing peak for stacking weak lensed signal from clusters.

Centering is not a well posed problem when there are multiple dense subhalos in the densest region of a cluster.

The problem is easier when the dense subhalos are spaced below observation resolution. However, when the dense subhalos are spaced within 50 kpc of each other in comoving units.

Unfortunate reliance of outliers of offset distribution to detect signal scrutiny of their parameter choices for the inference spatial distribution

Statistical aspects that one should take into account when constructing such a test include:

- whether the population PDF of offsets have converged

Characterizing the state of the cluster is not trivial.

6 STATISTICALLY INFERRING SIDM FROM A POPULATION OF GALAXY CLUSTERS

The main implications of our results for inferring $\sigma_{\rm SIDM}$ from the offset are the following.

There is no deterministic one-to-one mapping between Δs and $\sigma_{\rm SIDM}$. This result does not mean that $\sigma_{\rm SIDM}$ is always overestimated from past literature. On the contrary, it means that individual SIDM offset estimate may either experience a random increase or decrease due to the statistical noise of the data and different offset inference methods have different sensitivity to the noise. If the offset signal from SIDM is small enough, the offset distribution may be degenerate between different σ_{SIDM}.

plagued by expensive observations and small number of cluster samples.

not a statistically sound way of rejecting $\sigma_{\rm SIDM} = 0$ warn against p-value 'hacking' using this Illustris data

sound statistical analyses to claim discovery

Need expensive cosmological simulations with SIDM to gather sufficient statistics

Since SIDM is not the only source of contribution to the galaxy-DM offset in galaxy clusters, any method for small number of clusters do not account for the intrinsic scatter will have unknown bias $\sigma_{\rm SIDM}$. This is because the intrinsic Δs may or may not align with the SIDM offset contribution.

Need to sample enough galaxy clusters with the same selection criteria, both from simulations and from observations.

Asymmetry of DM map may not be statistically strong signal. Different choice of the size of smoothing kernel and projection effects can both lead to spurious features such as tails in the convergence map.

In simulation show low statistical noise floor so that the selected observable of SIDM can discriminate between CDM and SIDM.

Hard to formulate a multivariate probability distribution to account for all the uncertainties from time-since-collision, impact parameter, concentration of clusters before infall

How to characterize subclusters - use different statistical methods for clustering the DM and galaxy population. using NFW halos for DM halos and using Gaussian mixture model for galaxy clustering Not physical

Quantitative constraints carefully consider the contribution of each component

7 SIGNIFICANCE OF THE INTRINSIC OFFSET TO INFERENCE OF $\sigma_{\rm SIDM}$

Modeling a galaxy cluster is a high dimensional problem with missing variables such as line-of-sight information and merging history which can all lead to large uncertainties.

It is hard to define an analog for observed galaxy clusters due to missing information such as projection, merger history which may not be described well by parametric representation.

Given the uncertainties purely due to offset estimation methods and intrinsic scatter of Δs, any estimates of $\sigma_{\rm SIDM}$ without account for these uncertainties will overestimate.

no analytical form of how the effects of SIDM directly map as observables

model fitting that takes observational uncertainty into account.

show unambiguous signal of SIDM that is detectable despite the noise

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

8 SUMMARY

We showed that

• the contribution of statistical uncertainty to the galaxy-DM offsets for ΛCDM clusters is not negligible when compared to the reported levels of offset from staged simulations (~ 50 kpc). Any observational study that uses galaxy-DM offsets to constrain $\sigma_{\rm SIDM}$ has to account for this contribution or else $\sigma_{\rm SIDM}$ will be biased.

• while the location estimates of the 2D spatial distribution of offsets and the 1D marginal distributions in the Illustris simulation is approximately zero. The location estimate of the magnitude (or the root-mean-square) of the offsets is ~ 30 kpc from the Illustris sample. If one uses the magnitude of the offset to constrain $\sigma_{\rm SIDM}$, $\sigma_{\rm SIDM} = 0$ does not map to zero magnitude of offset.

• with high completeness and purity of member galaxy data, the luminosity map resembles the DM spatial distribution more closely than the number density map of member galaxies.

• the BCG has the closest offset to the dominant DM peak; and the KDE peak of the luminosity map gives the second tightest offsets from the DM peak.

• some of the observed offsets from various merging galaxy clusters have a p-value of [TODO] when compared to the offsets from a ΛCDM simulation. However, due to the idealistic data selection in the Illustris simulation, we have not accounted for more observational constraints such as incompleteness of member galaxy selection, line-of-sight substructures, or the lower resolution of DM data due to the lack of strong lensing DM peak estimates. The uncertainties from missing data is especially hard to quantify.

• the DM peak locations are consistent with the BCG.

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.

9 ACKNOWLEDGEMENTS

KN would like to thank Professor Thomas Lee for the helpful discussion of the construction of the p-value hypothesis test.

Part of the work before the conception of this paper was discussed during the AstroHack week 2014. Part of this work was performed under HST grant (TODO ask Dave for grant number).

REFERENCES

Bradač M., et al., 2006, ApJ, 652, 937
Bradač M., Allen S. W., Treu T., Ebeling H., Massey R., Morris R. G., von der Linden A., Applegate D., 2008, ApJ, 687, 959
Cui W., et al., 2016, MNRAS, 456, 2566
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Dawson W. A., 2013, PhD thesis, University of California, Davis
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
Genel S., et al., 2014, MNRAS, 445, 175
George M. R., et al., 2012, ApJ, 757, 2
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
Jee M. J., Hughes J. P., Menanteau F., Sifón C., Mandelbaum R., Barrientos L. F., Infante L., Ng K. Y., 2014, ApJ, 785, 20
Jee M. J., et al., 2015, ApJ, 802, 46
Kahlhoefer F., Schmidt-Hoberg K., Frandsen M. T., Sarkar S., 2014, 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
Mohammed I., Liesenborgs J., Saha P., Williams L. L. R., 2014, MNRAS, 439, 2651
R Core Team 2014, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/
Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradač M., Bradac M., 2008, ApJ, 679, 1173
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
Zitrin A., Bartelmann M., Umetsu K., Oguri M., Broadhurst T., 2012a, MNRAS, 426, 2944
Zitrin A., Bartelmann M., Umetsu K., Oguri M., Broadhurst T., 2012b, MNRAS, 426, 2944
Zitrin A., Menanteau F., Hughes J. P., Coe D., Barrientos L. F., Infante L., Mandelbaum R., 2013, ApJ, 770, L15

APPENDIX A: ALGORITHM OF THE SHRINKING APERTURE ESTIMATES

Data: subhalo that satisfy cuts as a galaxy

initial aperture centroid = weighted 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 weighted distance array
while *(newCenterDist - oldCenterDist) / oldCenterDist* $\geq$ *2e-2* **do**
    new data array = old data array within aperture
    newCenter = weighted mean value of new data along each spatial dimension
end

Algorithm 1: Shrinking aperture algorithm with luminosity weights

APPENDIX B: TABLE OF RESULTS

Table B1. Properties of the clusters used in the analysis. Richness is computed based on $i-$band < 24.4 assuming $z = 0.3$.

ID	richness	$M_{200C}(10^{14}M_{\odot})$	$M_{500C}(10^{14}M_{\odot})$	$M_{FoF}(10^{14}M_{\odot})$	non-relaxedness$_0$	non-relaxedness$_1$	midvar(Δy_{KDE}) (kpc)	max(Δy_{KDE}) (kpc)	med(ν)
0	483	1.64	1.09	3.23	29	33	31	65	1.43
1	338	1.57	0.62	2.68	20	16	25	71	1.59
2	267	1.53	0.87	2.12	17	3	18	42	1.30
3	343	0.82	0.56	2.03	37	59	44	148	2.01
4	213	1.19	0.66	1.54	21	4	24	84	1.58
5	212	0.90	0.56	1.44	20	27	16	43	1.19
6	225	0.96	0.60	1.40	18	7	15	28	1.16
7	230	0.31	0.17	1.41	54	280	101	379	2.83
8	148	0.83	0.54	1.34	24	26	20	52	1.32
9	187	0.79	0.50	1.29	23	12	33	111	1.45
10	158	0.73	0.53	1.15	19	8	19	49	1.35
11	134	0.57	0.33	0.95	20	9	36	78	1.12
12	164	0.20	0.09	0.87	64	142	77	218	1.99
13	115	0.22	0.14	0.79	63	143	38	118	2.26
14	90	0.45	0.29	0.67	15	8	17	33	1.08
15	92	0.51	0.35	0.68	11	3	11	25	1.00
16	113	0.40	0.23	0.61	19	4	13	31	1.06
17	97	0.42	0.18	0.60	21	8	27	53	1.09
18	83	0.45	0.31	0.59	15	8	14	32	1.20
19	86	0.26	0.19	0.57	30	68	18	77	1.52
20	84	0.15	0.11	0.50	60	122	54	117	1.65
21	89	0.26	0.12	0.53	23	8	47	146	1.24
22	70	0.42	0.30	0.49	14	7	10	23	1.06
23	68	0.25	0.17	0.47	30	25	26	98	1.03
24	66	0.33	0.26	0.44	14	14	11	42	1.17
25	79	0.23	0.15	0.43	23	25	11	22	1.25
26	61	0.26	0.18	0.45	28	40	11	44	1.30
28	69	0.30	0.16	0.41	22	12	26	42	1.01
29	62	0.30	0.20	0.42	16	14	9	22	1.20
30	59	0.18	0.14	0.40	42	78	17	63	1.35
31	57	0.29	0.21	0.40	14	15	10	24	1.06
32	56	0.18	0.13	0.38	35	23	43	83	1.54
33	69	0.19	0.10	0.38	49	54	60	108	2.11
34	63	0.21	0.14	0.39	23	20	22	33	1.07
35	69	0.29	0.22	0.41	12	3	11	28	1.01
36	72	0.24	0.16	0.36	21	22	16	39	1.18
37	63	0.21	0.16	0.36	25	23	51	142	1.11
39	55	0.27	0.18	0.36	11	3	12	29	1.00
40	54	0.18	0.10	0.33	44	69	81	151	1.39
46	52	0.08	0.06	0.30	57	73	59	157	1.65
48	53	0.12	0.08	0.30	40	104	13	44	1.82
51	56	0.19	0.13	0.29	12	5	17	41	1.00
58	58	0.14	0.09	0.23	29	10	21	66	1.00

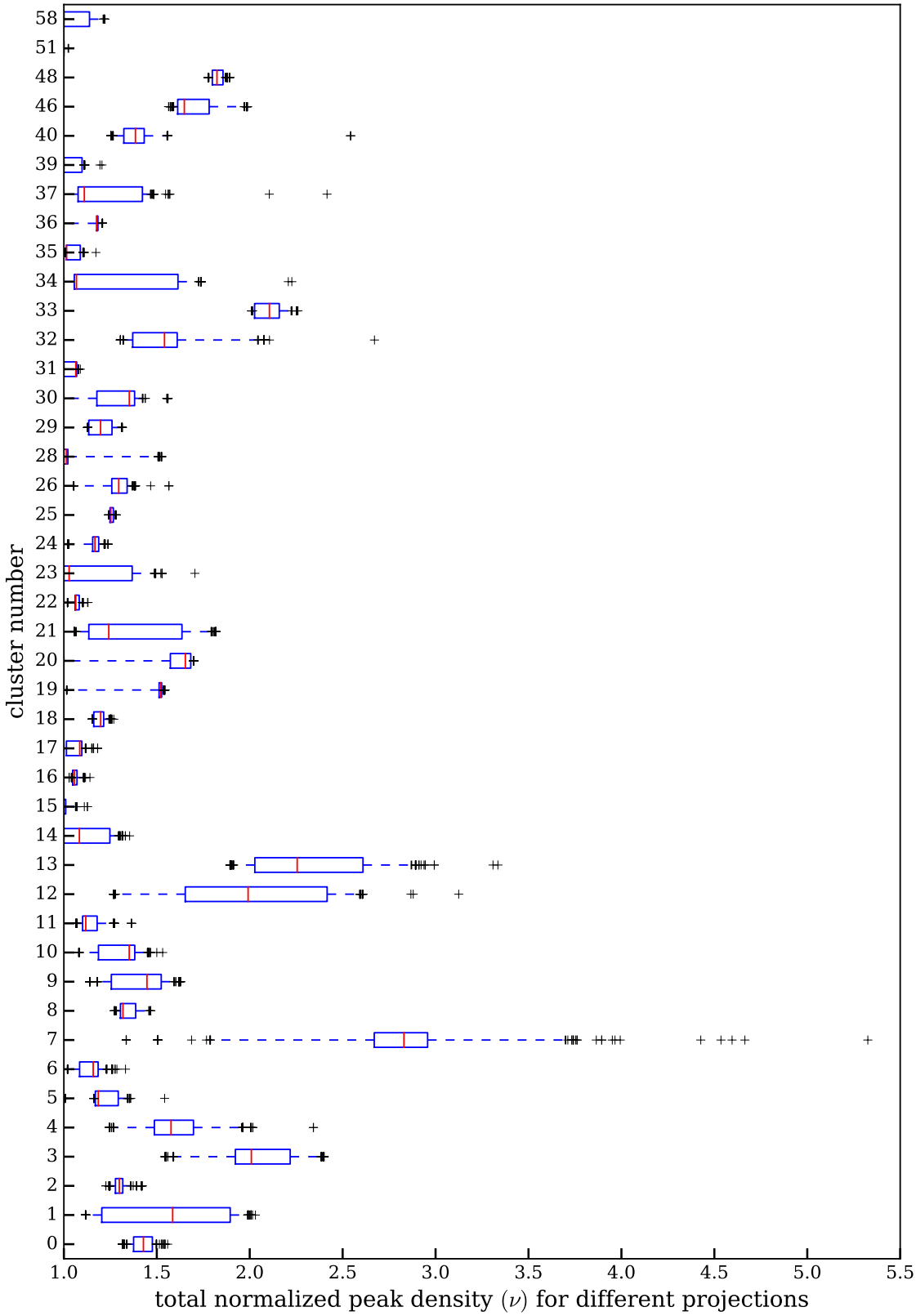

Figure B1. A box plot showing the distribution of the total normalized peak density for each cluster based on 768 projections. The red line shows the median of the projections, the box encompasses the 25% and 75% percentile of the distribution while the whiskers mark the 5% and the 95% percentile. The other black crosses are data points with extreme values beyond the 5% and 95% percentile.

Table B2. Summary statistic characterizing the offset distributions between the most bound particle and various summary statistics of the member galaxy population

	location	lower 68%	lower 95%	lower 99%	upper 68%	upper 95%	upper 99%
$\Delta y_{\rm BCG}$	0	-2	-2	-252	2	528	1107
$\Delta y'_{\rm centroid}$	0	-134	-491	-1176	134	491	1176
$\Delta y'_{\rm KDE}$	0	-19	-82	-1182	19	82	1182
$\Delta y_{\rm num.dens}$	0	-83	-302	-1114	83	302	1114
$\Delta y'_{\rm shrink}$	0	-50	-288	-1025	50	288	1025

The offsets represented with the prime $'$ symbols are estimated using the luminosity weighted galaxy data.

Table B3. Summary statistic characterizing the offset distributions for between the DM peak and the estimated galaxy location. All 43 clusters and all 768 projections are used in this table. The highest density values were used for the computation when there were more than one peak value estimated from the KDE.

kpc	mean	std	min	25%	50%	75%	max
$\lvert\Delta s_{\rm BCG}\rvert$	69	294	0	2	3	7	2335
$\Delta x_{\rm BCG}$	-14	226	-2331	-2	-0	1	2327
$\Delta y_{\rm BCG}$	23	197	-1980	-2	0	2	2332
$\lvert\Delta s'_{\rm centroid}\rvert$	261	209	2	114	202	317	1103
$\Delta x'_{\rm centroid}$	-42	224	-1022	-164	-37	66	1101
$\Delta y'_{\rm centroid}$	0	244	-1102	-111	-0	111	1100
$\lvert\Delta s'_{\rm shrink}\rvert$	118	156	0	21	60	165	1454
$\Delta x'_{\rm shrink}$	-7	131	-1089	-39	-3	23	969
$\Delta y'_{\rm shrink}$	0	145	-1091	-32	0	32	1109
$\lvert\Delta s'_{\rm KDE}\rvert$	37	35	0	14	26	49	498
$\Delta x'_{\rm KDE}$	-2	35	-330	-17	-2	12	386
$\Delta y'_{\rm KDE}$	-0	37	-439	-15	0	15	440
$\lvert\Delta s_{\rm num.KDE}\rvert$	136	161	1	56	92	147	2126
$\Delta x_{\rm num.KDE}$	-12	142	-1967	-55	-4	53	993
$\Delta y_{\rm num.KDE}$	-0	155	-1415	-54	-0	54	1417

The offsets represented with the prime $'$ symbols are estimated using the luminosity weighted galaxy data.

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