Velocity Dispersion Effects of Real and Mock Galaxies

Joe Hollowed[1,2]

[1]DePaul University
[2]Argonne National Laboratory
joehollowed@live.com

March 17, 2017

Abstract

The velocity distribution of galaxies within clusters has long been an object of interest in cosmology, since it has some promise of acting as a proxy for cluster mass. In addition, velocity distributions may encode information about how galaxies uniquely behave within clusters, as compared to hot intracluster gas (the ICM), or dark matter. At the same time, it is of utmost concern to those working on developing cosmological simulations to find reliable criteria by which to spatially distribute mock galaxies within gravity-only ("dark matter") simulations. Given these last two sentences, a potentially enlightening path can be seen; once mock galaxies are placed into a simulation via some model, we should be able to evaluate the quality of that model by looking for consistencies in the velocity distribution effects of both these mock galaxies and real spectroscopic data. This is what I have spent most of my time doing at Argonne National Laboratory (ANL) throughout DePaul University's Winter Quarter, and I will be presenting my work over the following pages. Specifically, I compare spectroscopic SPT and ACT galaxy data to simulated data output from "core-tacking", a method for mock galaxy placement and tracking produced at ANL, using our in-house simulation suites.

I. Introduction

Clusters of galaxies are the largest and most recently gravitationally relaxed objects to form thus far in the history of the cosmos. These clusters are the result of billions of years of evolution of initial small-scale density perturbations in an otherwise perfectly homogeneous environment - the early universe. In the midst of the expansion of space - the Hubble Flow - these density perturbations attracted matter and grew, until they were sufficiently massive as to decouple from the expansion, gravitationally collapse, and eventually (relatively recently) virialize.

Given this, it is reasonable to make the claim that the evolution of galaxy clusters is a reflection of the nature of large-scale structure formation in the universe. Thus, a study of clusters as a function of time (observing clusters at different redshifts) should enable us to constrain cosmological parameters. In short, this motivates cosmologists to put significant effort toward deducing the mass of galaxy clusters, since the universe is dominated by gravitational effects on these scales.

There are many observables discussed in the literature that can serve as proxies for cluster mass. For example, the optical richness - a measure of luminosity - of a galaxy cluster can give us an idea of its mass. Or, one can use the lensing of background galaxies by a cluster to estimate the depth of its gravitational potential.

There are also many other mass-relating observables in wavebands other than the optical. However, the observable that we will be concerning ourselves with for the rest of this paper is the velocity distribution and dispersion of a cluster's member galaxies.

The justification for using such a quantity to probe for the mass of a cluster is given by the *virial theorem*. The virial theorem simply states that for a self-gravitating, stable, spherical distribution of equal mass objects, the potential energy of the system is within a factor of two of the system's total kinetic energy. With most galaxy clusters, all of these assumptions are approximately true, which we can see, in the event that we have data for many galaxies within a single cluster.

There are indeed clusters that are sufficiently far from being stable (or "virialized") that they are better suited for other measures of mass. Such a cluster is usually in the stages of a large merger. Though, there have been studies whose results favor the view that the majority of clusters are dynamically relaxed systems (Evrard et al., 2008).

Eventually, the goal would be to try and find a reliable fitting function to relate a cluster's velocity dispersion to its mass as accurately as possible, ideally in a cosmology-independent way. The quality of such a relation could be checked by calibrating to other mass-observable relations, but also by comparing observational results to simulation results. That comparison would go something like this: we can measure the velocity dispersion of some simulated clusters, ideally using samples that match the nature of the observational sample as well as possible. Since we *know* the actual masses of the simulated clusters, we can then investigate the quality of our conclusions drawn from the real data.

The utility of this method, however, hinges on being able to confidently assume that our simulations are an accurate representation of reality. Herein lies where the inspiration for much of my work at Argonne comes from. I want to find out - given some mock galaxy catalogs output from simulation, do they behave how real galaxies behave? And what kind of analyses can we do to answer that question?

Over the next few pages, I will be discussing my attempts to do just that, with what we call *cores* taking on the role of mock galaxies in our simulation data.

As has likely been made evident in this introduction, in this paper I assume some working knowledge of cosmology by the reader, for the sake of brevity. Now, I shall proceed as follows: § 2 will be an overview of where I have sourced my data for this project, and also an introduction to the concept of core-tracking. In § 3, I address the statistical techniques and estimators I use to process data, and to explicitly define velocity dispersion. § 4 then discusses the methods I employ to make comparisons between real and mock galaxies, with respect to velocity dispersion information. Finally, § 5 provides a summary, a statement on future work, and some conclusive remarks.

II. Observational Catalogs and Simulation Data

2.1 Spectroscopy and Observational Catalogs

There are several large scale collaborations currently in the business of creating observational catalogs of galaxy clusters, most notably the South Pole Telescope (SPT) project and the Atacama Cosmology Telescope (ACT) project.

The South Pole Telescope is a 10-meter diameter telescope located in Antarctica at the Amundsen-Scott South Pole Station, designed for observations in the millimeter and submillimeter wavebands. The SPT-SZ survey has been my primary data source for observational galaxy clusters, all of which were discovered with SPT via the Sunyaev-Zel'dovich effect (Bleem et al., 2015).

For the actual galaxy data (I obtain galaxy velocities via spectroscopic redshifts), I have been using the SPT-GMOS survey, consisting of spectroscopic follow-up for 62 of the galaxy clusters in the SPT-SZ survey, totaling 2243 galaxies (Bayliss et al., 2016). In addition, I have received galaxy spectroscopic data from M. Bayliss, as presented in Bayliss et al. (2017), which adds 1431 more galaxies my dataset, spanning 27 clusters as presented in the literature of other projects, mostly ACT (Sifón et al., 2016).

2.2 Simulations and Core Tracking

Argonne's cosmology group is one of the world's leaders in developing state of the art cosmological simulations. Our group runs very large dark matter (gravity only) simulations of the universe, from the initial density perturbations of the universe, up to the present time at redshift $z = 0$, boasting unprecedented mass resolutions in a cosmological volume.

Specifically, for my analysis here, I use one of our "mini" simulations, dubbed *Alpha Quadrant* (AlphaQ). AlphaQ simulates billions of "dark matter" particles, each representing $10^9 M_\odot$, as they evolve through time according Newtonian gravitation. This all happens within a box of 360^3Mpc in volume, with periodic boundary conditions, and a WMAP-7 cosmology (Komatsu et al., 2011).

Again, these are gravity only simulations. There are no baryonic physics involved, preventing any true simulated galaxies from populating the *halos* that form in the later stages of the simulations. To address this problem, we use a technique called *core-tracking* (as far as our concerns here go, a "halo" is the simulated analog of a galaxy cluster - a large gravitationally bound clump of mass).

In simple terms, core-tracking works like this: when clumps of matter in the simulation grow above some mass threshold that we define, we will assign a *core* to them. What this means is that we collect data on the 20 most bound particles in that dense clump, and save them as a group under a unique identifier, a *core tag*. As the cores move through the simulation, we track each of their 20 constituent particles, and monitor how compact they remain. If these 20 particles become too disperse, then the core is considered *disrupted*, and we discard it.

We populate the simulation with cores wherever we find massive, dense regions, which should be in agreement with where galaxies form in reality. The creation of a core, then, represents the spawning of a galaxy, and disruption is interpreted as the diffusion of a galaxy into the ICM due to tidal forces. *Fig.1* shows the a single halo output from running core-tracking on AlphaQ, where each red point is a core, each blue point is a disrupted core, and the grey shading represents the dark matter density.

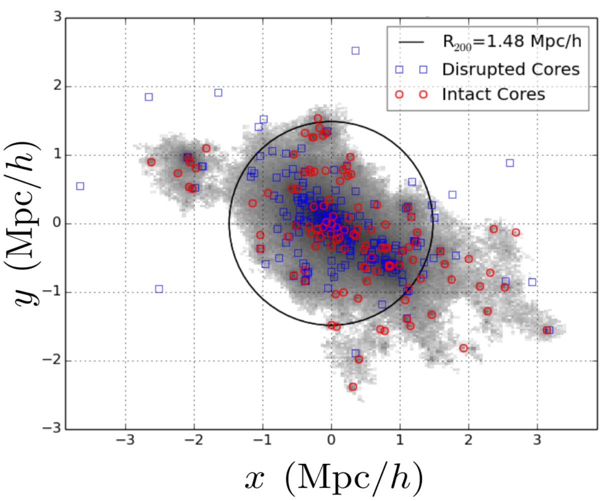

Figure 1: *A single dark matter halo from AlphaQ, showing the spatial distribution of intact cores (red), disrupted cores (blue), and the dark matter density (grey). The black curve represents the virial radius of the halo, denoted R_{200}. We generally find the highest density of cores in the center of the virial radius, often surrounded by a cloud of disrupted cores.*

This halo, populated with cores, should be interpreted as the simulated analog to a galaxy cluster populated with galaxies. As a reminder, the question we seek to answer in this work is: do cores behave as real galaxies do with respect to velocity dispersion?

III. Statistical Estimators

3.1 Dispersion Estimators

As noted in § 1, clusters can, to a significant extent, be considered a relaxed and stable population. If the assumptions stated in our discussion of the virial theorem are true - that a cluster is spherically symmetric and stable - then the velocity distribution of its member galaxies should be a typical Gaussian. However, clusters do not perfectly match these assumptions, and in most cases, we can only say that the velocity distribution is *nearly* Gaussian (see Becker et al. (2007) for a technical discussion).

Combine this with the fact that we often have rather small sample sizes for observational clusters (~1 to ~100), and it is clear that we have to be careful with our choice of definition of "dispersion" and "average", as to not yield biased results. What we want are *robust* and *resistant* estimators (robustness is characterized by some degree of insensitivity to the assumed distribution from which a sample of data is drawn; resistance implies insensitivity to locally misbehaved data). To this end, I use the *biweight dispersion* to measure velocity dispersions, and the *biweight average* (Beers et al., 1990) to find the center of clusters in both redshift and velocity space. The biweight dispersion, σ_{BI}, and the biweight average, C_{BI}, are given as

$$\sigma_{BI} = n^{1/2} \frac{\left[\sum_{|u_i|<1}(x_i - M)^2(1-u_i^2)^4\right]^{1/2}}{\left|\sum_{|u_i|<1}(1-u_i^2)(1-5u_i^2)\right|} \quad (1)$$

$$C_{BI} = M + \frac{\sum_{|u_i|<1}(x_i - M)(1-u_i^2)^2}{\sum_{|u_i|<1}(1-u_i^2)^2} \quad (2)$$

where x_i are our data points, M is the sample median, n is the sample size, and the weights u_i are given by

$$u_i = \frac{(x_i - M)}{c\mathrm{MAD}} \quad (3)$$

in which c is a tuning constant, and MAD is the median absolute deviation from the sample median. As can be seen, C_{BI} requires an auxiliary estimate for the location of the data, which we take to be M. In my analysis, I compute this C_{BI} iteratively, replacing M with the previous measure of C_{BI}, until convergence, and likewise for σ_{BI}.

Beers et al. (1990) shows these estimators to be quite robust for reasonable sample sizes ($n \geq 10$), in that they retain high efficiency in the case of non-Gaussian populations. As a meaningful comparison, a typical sample mean has 100% efficiency in a pure Gaussian distribution, but drops to zero almost immediately even for a slight deviations from normality.

These estimators have a history of being used in the literature for velocity dispersion applications, with some degree of reliability (Bayliss et al., 2016, 2017; Becker et al., 2007; Ruel et al., 2014; White et al., 2010).

3.2 Membership Selection

We must take great care in preforming any kind of statistical analysis with spectroscopic follow-up of cluster surveys, since we can never be sure about true cluster membership. This means that, for a given cluster, our sample of galaxies is almost certainly contaminated with foreground and background galaxies that appear to lie in the cluster from our line of sight, but are actually entirely external to the system. These non-member galaxies are known as *interlopers*.

In this subsection, I will outline my procedure that aims to clean up these contaminated clusters. To serve as visual reference during this discussion, *Fig.2* shows the velocity distribution of 29 galaxies with spectroscopic data in a certain SPT cluster.

To begin interloper removal, I make a hard velocity cut of ± 5000 km/s, relative to the bulk cluster velocity. I find this cluster velocity by evaluating Eq.(2) for C_{BI} on the all of the galaxy redshifts, resulting in an average value of z for the cluster, which can be converted to a line of sight velocity.

Then, I begin a simple procedure known as *3σ clipping*, in which I iteratively remove all galaxies further than $3\sigma_{BI}$ from the average cluster velocity (recalculating C_{BI} on each iteration), until convergence. The motivation here is that galaxies that are far outside the velocity distribution of the cluster have a very low chance of being member galaxies.

The velocity distribution in *Fig.2* shows some obvious interlopers. The red dashed line is the initial ± 5000 km/s cut, and the blue lines are the final iteration of the 3σ clipping. In this instance, all galaxies determined to be interlopers were removed during the first step.

Again, there are many examples of 3σ interloper removal in the literature on velocity dispersions; for more information see Bayliss et al. (2016, 2017); Becker et al. (2007); Ruel et al. (2014). There are also examples of much more sophisticated removal, notably White et al. (2010).

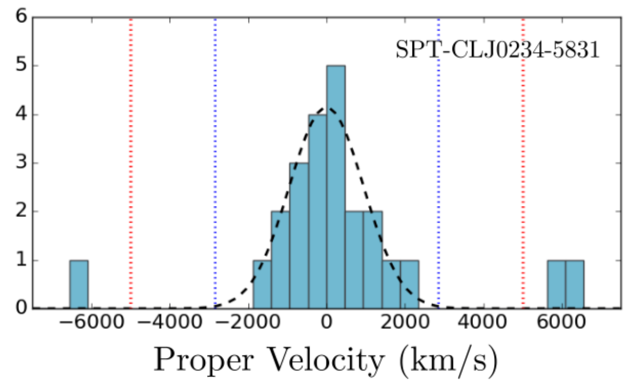

Figure 2: *The velocity distribution of a single cluster from the SPT-GMOS survey, which had 29 galaxies before interloper removal. The Gaussian fit to the histogram is defined to have a center of zero, and a width equal to σ_{BI} of the 21 member galaxies remaining after the initial velocity cut (red dotted lines), and subsequent 3σ clipping (blue dashed lines).*

IV. Velocity Dispersion Effects

4.1 Velocity Bias

Now, we discuss the ways in which velocity dispersions can be used as a comparison between observed and simulated galaxies. We could plot both the SPT clusters and the AlphaQ halos on the same plot, with mass on the x-axis, and velocity dispersion on the y-axis. But this would be folding in complications due to mass measurements, since we know the mass perfectly for simulated halos, and not for real clusters.

A better approach would be to look for some statistical effects present in the observational data, and see if we can pull those effects out of the core-tracking data as well (which we would expect to, if the cores are behaving as galaxies).

In this subsection, I shall address one such effect - bias between velocity dispersion-mass relations of galaxies and (simulated) dark matter. Many studies have been done in the past to try to compare observational velocity dispersions to those output from simulations, with the intent to try and detect any possible velocity bias, b_v, between galaxies and the rest of the continuous mass within a cluster (dark matter). Values have been reported anywhere within the wide range $b_v \simeq (0.95 - 1.3)$ (Bayliss et al., 2017; Saro et al., 2013; White et al., 2010).

Evidently, this bias is quite poorly constrained as of the present time, and is sometimes left as a free parameter where assumptions are required (Becker et al., 2007). However, we are not so much interested in the specific value of b_v, as we are interested in how b_v may vary as a function of cluster sample size.

Shown in *Fig.3* is a velocity dispersion vs. mass plot for 83 SPT clusters, along with a line representing the scaling relationship between velocity dispersion and cluster mass from dark-matter simulation data (Saro et al., 2013). The 52 grey points are clusters having members counts $n < 30$.

One thing of note is that (not surprisingly), the points of highest scatter are those clusters with low member counts. Despite the robustness of the biweight estimators, if the data is lacking, then the result shall be poor.

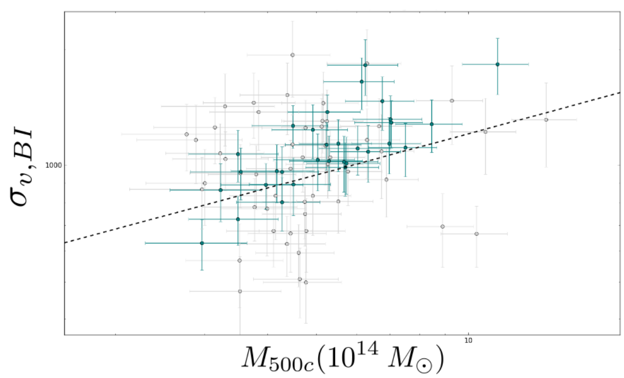

Figure 3: *The biweight velocity dispersion versus SZ-based SPT cluster masses. Shaded points are those clusters that had at least 30 members, while grey points are the opposite. The dashed line is the dispersion-mass relation as found from analyses of dark matter simulations (Saro et al., 2013).*

What is interesting about this is that observers conducting similar analyses notice that clusters with low member counts generally seem to prefer scattering low, rather than high (Though I can not speak to confirm or reject this claim via *Fig.3*). I have even been informed that J. Ruel did revisit a similar plot of his, gathered more spectroscopic data for low-biased clusters, and noticed the dispersions for those clusters rose (Bleem, private communication).

The thought, between a few colleagues and I, is that this effect could be due to substructure - if you can only manage to get spectroscopic redshifts for 10 galaxies in a cluster, and most of them happen to reside is some kind of recently infallen clump, then you will pick up the velocity dispersion of a smaller system than the cluster you intended to observe.

Even if one is hesitant to accept this reasoning, it is at least true that an asymmetric scatter may indeed be due to real physical effects. And if that is the case, then we should be able to find a similar effect in the core-tracking data from AlphaQ.

Fig.4 shows an attempt to look for that effect. Plotted is nearly the same as *Fig.3*, using cores in place of galaxies, and a different mass definition. Each halo has two points - one dispersion measured via the dark matter particle velocities, and one dispersion measured via the core velocities. Also included are fitting function to these two datasets, and a simulation-derived relation due to Evrard et al. (2008).

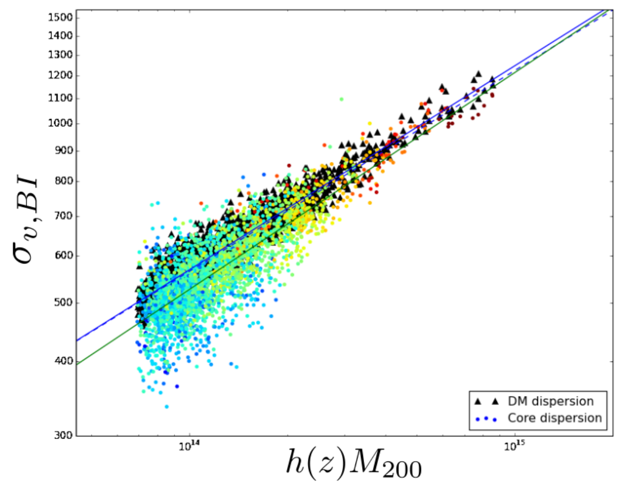

Figure 4: *The biweight velocity dispersion versus halo masses. Each halo has two points: the black triangles represent velocity dispersions measured via dark matter particle velocities, while the colored points measured dispersion via core velocities. These latter points are color coded according to the number of cores in that halo (which range from ~10-~200). The solid blue line is a least squares fit to the dark matter dispersions, while the green line is fitted to the core dispersion points. The dashed line is a dispersion vs. mass form by Evrard et al. (2008).*

We see that the points with the highest member counts (red) are in good agreement with the dark matter points, while points with problematically few members (blue) seem to heavily favor scattering low rather than high.

As of right now, this is interesting, and shows that the cores are indeed tracing some kind of physical effect not seen by the dark matter particles alone. If more suggestive evidence can be found for a similar effect in observational data, then more work can be done to quantify the potential correlation here.

4.2 Velocity Segregation

Another effect that can be pulled out of velocity distribution information is something known as *velocity segregation*, as per Bayliss et al. (2017). The idea here is straightforward; we can take all the galaxies in a particular cluster, and bin ("segregate") them by some property (such as color, magnitude, richness, etc.). Then, we can investigate whether the velocity dispersion of one sub-population differs appreciably from another. Again, if we expect cores to behave as do galaxies, then we should expect to be able to recover those "segregation effects" in simulation.

There is one significant complication that first must be addressed, however; we don't have many clusters in the SPT sample that have enough spectroscopic members to bin the distribution in such a way without degrading its statistical ability. To explain how we work around this problem, I now introduce the idea of the *stacked cluster*.

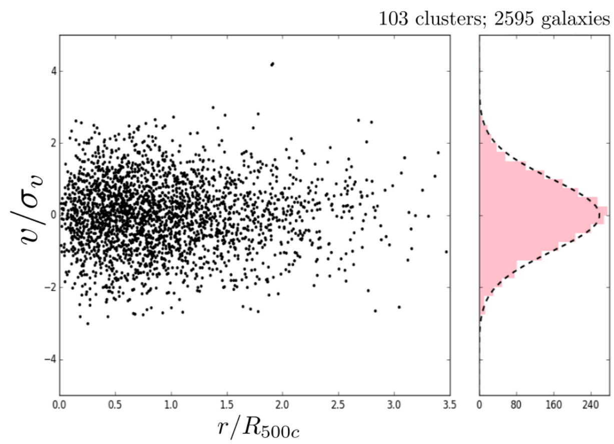

Figure 5: *A stacked cluster, including 2595 spectroscopic galaxies with normalized velocities and radial distances, with respect to their host clusters.*

Fig.5 shows one such of these stacked clusters, in which I have "stacked" all of the SPT cluster galaxies (for all clusters with at least 15 spectroscopic members via SPT-GMOS). I have normalized the distance of each galaxy from its host cluster center by the cluster radius, and normalized each galaxy velocity by the dispersion of its host cluster. With this done, I can put all of the galaxies, from all clusters, into one distribution.

As is evident from *Fig.5*, this essentially gives us one cluster in which we have a high sample size, giving us much more statistical flexibility. Specifically, we end up with a stack consisting of 103 clusters and 2595 galaxies. We can then preform velocity segregation analysis on this stack and learn about the dynamics of different sub-populations of galaxies.

As an example, I show the results of preforming this velocity segregation with respect to galaxy spectral type in *Fig.7* on the next page. The three sub-populations present in this figure are passive galaxies (abbreviated PA; generally older, redder galaxies, that have been in the cluster longer and are hence more virialized), star-forming galaxies (SF; generally young, bluer galaxies, found at larger radii since they are more recently infallen), and post-starburst galaxies (PS; mid-range galaxies, in a transition stage between star-forming and passive.)

By taking the velocity dispersion of these new sub-populations, and normalizing it by the dispersion of the full stack (*Fig.5*), I find the following results:

- $\sigma_{v,\mathrm{PA}}/\sigma_{v,\mathrm{all}} = 0.98 \pm 0.02$
- $\sigma_{v,\mathrm{PS}}/\sigma_{v,\mathrm{all}} = 1.12 \pm 0.10$
- $\sigma_{v,\mathrm{SF}}/\sigma_{v,\mathrm{all}} = 1.28 \pm 0.18$

We interpret this as being reflective of the fact that bluer star-forming galaxies have, in general, been in the cluster for much less time than the passives, and have not yet had time to experience much dynamical friction and virialize.

In *Fig.6*, I repeat this same segregation analysis with a dataset containing more non-SPT clusters (via Bayliss et al. (2017)). This time, rather than plotting the individual distributions, I just plot the ratio $\sigma_v/\sigma_{v,\mathrm{all}}$, as I vary the percentage of star-forming galaxies that I include in the stack (see figure caption for more info). This gives the same conclusion - star-forming galaxies as a sub-population have a higher velocity dispersion than the full sample.

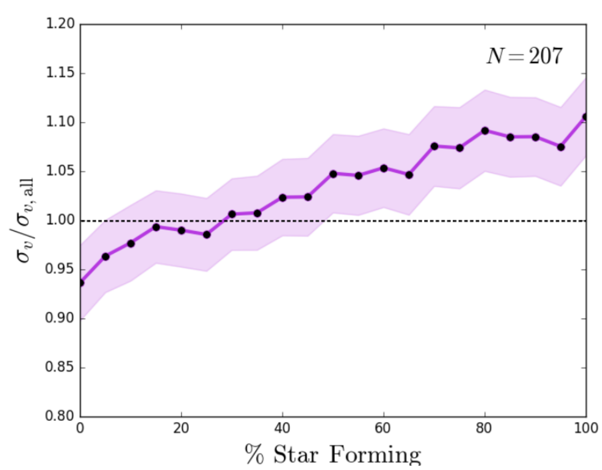

Figure 6: *Velocity dispersions with respect to that of the full stack as we vary the percentage of star-forming galaxies. The leftmost point is 0% SF galaxies, so it represents a stack containing only PA and PS galaxies. Every subsequent point is a new stack, including more SF and less PA/PS galaxies. The right-most point is a stack entirely composed of SF galaxies. The shaded region is 1σ certainty.*

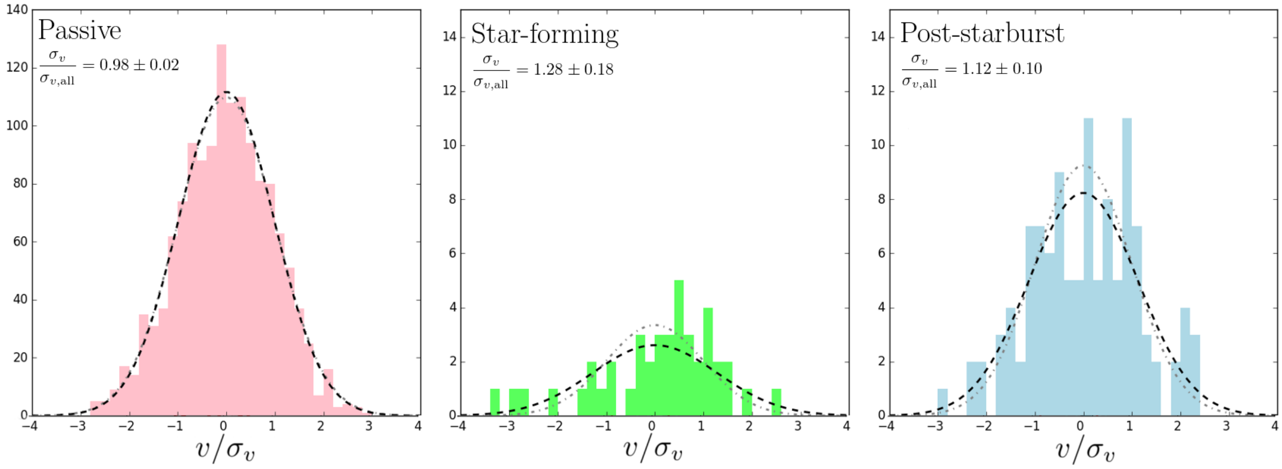

Figure 7: *The stacked cluster from Fig.5, binned into sub-populations according to galaxy spectral type. We find that star-forming galaxies have the highest dispersion as compared to the full stack. The black dashed Gaussians are a normal distribution having a center of zero, a width of σ_v for that spectral type, and scaled so the area under the curve matches the area of the histogram. The lighter dash-dot lines are the same thing, but using $\sigma_{v,all}$ as the width.*

To be able to compare such an analysis directly to core-tracking data, we would need some way to assign spectral types (or, at least, red/blue color assignment) to cores, which we do not have. However, we *can* repeat this segregation strategy on the cores using any other observable that we believe to correlate in an unknown, but real, way with galaxy spectral type/color.

Two such observables are the core *radius*, and the core *infall step*. Recall that a core is nothing more than a collection of 20 particles that are gravitationally bound. The core radius is an expression of how compact or "poofy" those 20 particles are. So, cores with a higher radius are closer to being disrupted than cores with a small radius.

This could potentially correlate with galaxy spectral type in the following way: recall in § 2, we noted that dark matter halos often end up with a cloud of disrupted cores at their centers. In general, core tracking seems to suggest that cores that are disrupted or nearly disrupted (having larger radii) have been in the cluster longer than compact cores, and thus are found near the potential minimum. This makes sense, given that we would expect galaxies that have been in a cluster longer to have experienced more tidal forces and merger events.

The converse also seems reasonable - cores that have recently infallen into a halo could have come from a much less dense and exciting environment, and would never have had cause to loose its compactness. For these reasons, I plot the velocity segregation now as a function of core radius in *Fig.8*.

The results are indeed just as we expect, with compact cores having a high velocity dispersion relative to the full stack. Likewise, "poofy" cores with higher radii are biased low in velocity dispersion. My exact findings are

- $\sigma_{v,\text{compact}}/\sigma_{v,\text{all}} = 1.038 \pm 0.006$
- $\sigma_{v,\text{poofy}}/\sigma_{v,\text{all}} = 0.931 \pm 0.005$

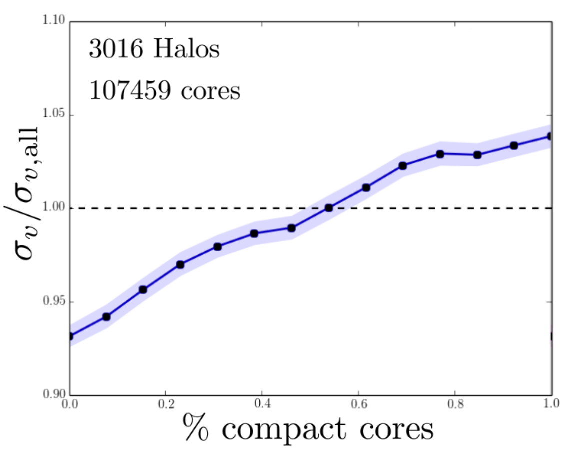

Figure 8: *Velocity dispersions for a stacked halo containing varying fractions of compact cores (cores having low radii). It is clear that compact cores, as a sub-population, have a dispersion biased high as compared to the full stack of all cores. This calculation involves halos from all redshifts, and the shaded contour represents 1σ certainty.*

(I should note, whether a core is placed into the "poofy" or "compact" sub-sample is just determined by which side of the median of all core radii that particular core finds itself. This criteria applies to the rest of the segregation analyses in this paper).

The other mentioned observable that may correlate with galaxy spectral type is the core infall step. The infall step is a property of a core that specifies how long ago (at what simulation time step) the core fell into the halo.

Once again, it is the assertion that passive red galaxies have been in a cluster long enough to be tidally stripped, ending active star-formation, and virialize. As *Fig.6* implies, this old, red population has a dispersion biased low. The inverse of this statement is true for blue star forming galaxies. Given this, we should recover the exact same trend if we preform velocity segregation analysis as a function of core infall step.

The result of such an analysis is shown in *Fig.9*. This too agrees well with our treatment of the real galaxy data. Recently infallen cores are seen to have a dispersion biased high with respect to the full stack, while cores that have been in the halo for some time are biased quite low. Specifically, I find:

- $\sigma_{v,\text{old}}/\sigma_{v,\text{all}} = 0.916 \pm 0.006$
- $\sigma_{v,\text{young}}/\sigma_{v,\text{all}} = 1.02 \pm 0.006$

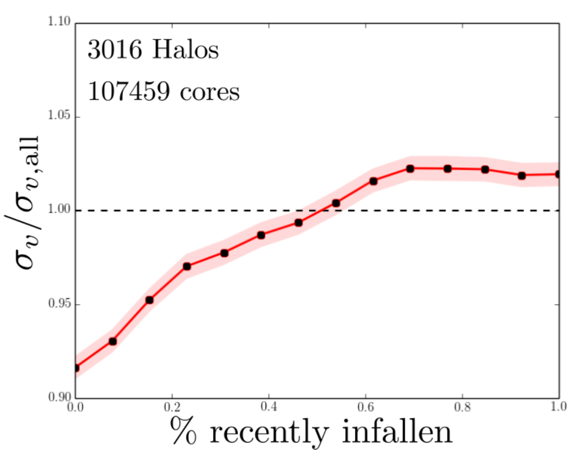

Figure 9: *Velocity dispersions for a stacked halo containing varying fractions of recently infallen cores (cores with a high infall step). It is clear that recently infallen ("young") cores, have a dispersion biased high as compared to the full stack of all cores. This calculation involves halos from all redshifts, and the shaded contour represents 1σ certainty.*

This agrees with both *Fig.6* and *Fig.8*, if we feel confident in our assumption that passive red galaxies can be represented by both more "poofy" cores, and cores that have been in the cluster for a longer time.

I note that similar analyses have been done using Semi-Analytic Models for galaxy placement in simulations. Such models *can* give you colors for mock galaxies. Gifford et al. (2013) does a velocity segregation analysis with these SAM galaxies with several other properties, which includes segregation as a function of color (The simulational analog to *Fig.6*). Their results also qualitatively agree with what is stated here; dispersions of simulated blue galaxies are biased high in their results by as much as $\approx 35\%$ (see also Bayliss et al. (2017) for a direct comparison of these results to observation).

Finally, I conclude this section by noting a curious feature I've found; preforming velocity segregation as a function of core *radial distance* seems to give unexpected results. One would expect that both older cores (less recently infallen) and poofier cores (larger radii) would be located close to the halo center. Also, red galaxies are typically found to be concentrated in the center of observational clusters. All three of these sub-populations, in my previous analyses, are biased low in velocity dispersion.

However, when I preform a velocity segregation scheme on all core's radial distance from the halo center, I yield the exact opposite answer, as shown in *Fig.10*. The biases for each population are

- $\sigma_{v,\text{distant}}/\sigma_{v,\text{all}} = 0.97 \pm 0.006$
- $\sigma_{v,\text{near}}/\sigma_{v,\text{all}} = 1.02 \pm 0.006$

Though the effect is rather small, I think it is worth some further investigation.

V. Summary & Conclusions

Throughout this paper, we have discussed the statistical methods used in the investigation of cluster velocity dispersions. We introduce the *biweight* estimators due to Beers et al. (1990), and apply them SPT-GMOS spectroscopic data and simulated core-tracking data, as introduced in § 2.2.

We have touched upon a simple method for cleaning potentially interloper contaminated observational datasets, and proceed to use the results to compare to core-tracking in velocity space. In § 1, we stressed the importance of doing such a comparison; simulations can serve as a powerful verification tool for cluster mass-observable conclusions, as long

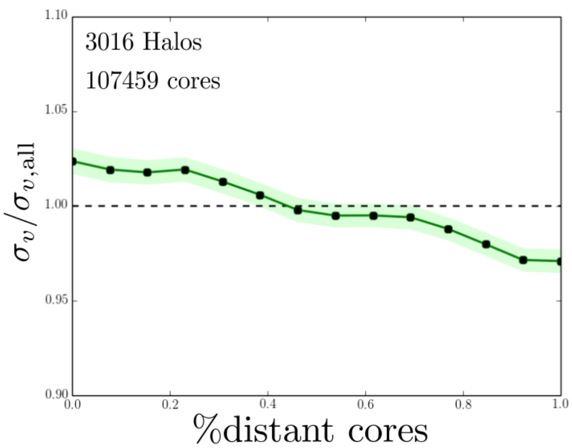

Figure 10: *Velocity dispersions for a stacked halo containing varying fractions of "distant" cores (those found close to the halo center). Here, we see the unexpected result that cores located further from the center are biased low in velocity dispersion, while cores nearer the center are biased high.*

as we are confident that our simulation reasonably represents reality. Checking for common velocity dispersion effects in both spectroscopic galaxy samples and simulated galaxy tracers can give us an idea of how well reality is being reflected in our synthetic datasets.

§ 4 mostly shows interesting effects in the velocity space of cores that *hints* at some commonalities shared with the dynamics of real galaxies, pending some further investigation. I show that dark matter halos with low core counts preferentially scatter low on a mass-vs-dispersion plot, suggesting the possibility of substructure detection.

Also, we review the idea of *velocity segregation*, and show how the biases in velocity dispersion of different sub-populations of a full stacked halo seem to agree across observational and simulated data. We find that red or passive galaxies in a sample of SPT and ACT galaxies have a velocity dispersion biased low by ~2%, and blue star-forming galaxies are biased high by ~28%.

Until we can make direct comparisons to cores with respect to galaxy color or spectral type, we can use core properties which we believe to be in some way correlated with color. Those properties are, as shown in § 4.2, core *radius*, *infall step*, and *radial distance* from its host halo's center.

I report *compact* cores to be biased high by ~4%, while *"poofy"* cores (with a large radius) are biased low by ~7%. Likewise, cores that have fell into the cluster long ago are biased low by ~8% in velocity dispersion, and those that fell in relatively recently are biased high by ~2%. Finally, the results for preforming velocity on core radial distance gave unexpected results, as we noted at the end of § 4.2. Cores far from the halo center were biased low by ~3%, and those near the halo center were biased high by ~2%.

Our concise conclusion is that core tracking does seem to be successful in mimicking the behavior of galaxies in velocity space, as far as our vague relations between galaxy age, radius, and color can tell us. More work certainly must be done, however. I feel that the most immediately important step to take is to find some way to assign colors to cores, or at least to find a valid relation between color and some our our core properties. This would facilitate a more direct comparison between SPT and ACT galaxies and AlphaQ cores.

Finally, as an undergrad finishing up my senior year at DePaul University, I have another, perhaps more important conclusion. Working at ANL over the past year has given me a good appreciation of how physics actually works. By this, I mean that I now better understand how one proceeds from theory, to experimentation and modeling, to results. Also, I now see how these things operate within the context of collaboration.

I have been exposed to a lot of one-on-one collaboration with my coworkers, and at least weekly time for discussion with the whole cosmology group at ANL. I've also also had the opportunity to discuss my work with visitors from other institutions, and attend talks on all sorts of topics within astronomy, cosmology, and physics in general.

Working in a stimulating environment on the front line of physics, at a world class laboratory, has certainly convinced me of my interests in contributing to the scientific endeavor.

Acknowledgments

I wish to thank everyone at ANL who has provided me with guidance and been so willing to ensure that I have a positive experience at the lab. I extend my gratitude to Lindsey Bleem, who has helped with with any question that ever arose, regularly offered to look over my work, and gave me a sense of direction. I also thank Salman Habib for accepting me into the cosmology group, and providing indispensable guidance. Dan Korytov has been one of my most frequent hands-on collaborators, and is the developer of the core-tracking algorithms; I thank him for constant assistance. I have received help on the simulation side of things from Katrin Heitmann and Steve Rangel on numerous occasions, and also received observational datasets and advice from Matt Bayliss.

Finally, I wish to thank DePaul University and my sponsoring professor Dr. Jesús Pando for the possibility of receiving Independent Study credits for my work at ANL, without which I would have been very limited in the amount of time and effort I could dedicate to this project.

References

Bayliss, M. B., Ruel, J., Stubbs, C. W., Allen, S. W., Applegate, D. E., Ashby, M. L. N., et al. (2016). SPT-GMOS: A Gemini/GMOS-South Spectroscopic Survey of Galaxy Clusters in the SPT-SZ Survey. *Astrophysical Journal, Supplement*, 227:3.

Bayliss, M. B., Zengo, K., Ruel, J., Benson, B. A., Bleem, L. E., Bocquet, S., et al. (2017). Velocity Segregation and Systematic Biases In Velocity Dispersion Estimates with the SPT-GMOS Spectroscopic Survey. *Astrophysical Journal*, 837:88.

Becker, M. R., McKay, T. A., Koester, B., Wechsler, R. H., Rozo, E., Evrard, A., et al. (2007). The Mean and Scatter of the Velocity Dispersion-Optical Richness Relation for maxBCG Galaxy Clusters. *Astrophysical Journal*, 669:905–928.

Beers, T. C., Flynn, K., and Gebhardt, K. (1990). Measures of location and scale for velocities in clusters of galaxies - A robust approach. *Astonomical Journal*, 100:32–46.

Bleem, L. E., Stalder, B., de Haan, T., Aird, K. A., Allen, S. W., Applegate, D. E., et al. (2015). Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the 2500-Square-Degree SPT-SZ Survey. *Astrophysical Journal, Supplement*, 216:27.

Evrard, A. E., Bialek, J., Busha, M., White, M., Habib, S., Heitmann, K., et al. (2008). Virial Scaling of Massive Dark Matter Halos: Why Clusters Prefer a High Normalization Cosmology. *Astrophysical Journal*, 672:122–137.

Gifford, D., Miller, C., and Kern, N. (2013). A Systematic Analysis of Caustic Methods for Galaxy Cluster Masses. *Astrophysical Journal*, 773:116.

Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Nolta, M. R., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. (2011). Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophysical Journal, Supplement*, 192:18.

Ruel, J., Bazin, G., Bayliss, M., Brodwin, M., Foley, R. J., Stalder, B., et al. (2014). Optical Spectroscopy and Velocity Dispersions of Galaxy Clusters from the SPT-SZ Survey. *Astrophysical Journal*, 792:45.

Saro, A., Mohr, J. J., Bazin, G., and Dolag, K. (2013). Toward Unbiased Galaxy Cluster Masses from Line-of-sight Velocity Dispersions. *Astrophysical Journal*, 772:47.

Sifón, C., Battaglia, N., Hasselfield, M., Menanteau, F., Barrientos, L. F., Bond, J. R., Crichton, D., Devlin, M. J., Dünner, R., Hilton, M., Hincks, A. D., Hlozek, R., Huffenberger, K. M., Hughes, J. P., Infante, L., Kosowsky, A., Marsden, D., Marriage, T. A., Moodley, K., Niemack, M. D., Page, L. A., Spergel, D. N., Staggs, S. T., Trac, H., and Wollack, E. J. (2016). The Atacama Cosmology Telescope: dynamical masses for 44 SZ-selected galaxy clusters over 755 square degrees. *Monthly Notices of the Royal Astronomical Society*, 461:248–270.

White, M., Cohn, J. D., and Smit, R. (2010). Cluster galaxy dynamics and the effects of large-scale environment. *Monthly Notices of the Royal Astronomical Society*, 408:1818–1834.