

GALAXY GROUP DYNAMICS: STATISTICAL ANALYSIS AND
COMPARISON OF GROUP PROPERTIES

GALAXY GROUP DYNAMICS: STATISTICAL ANALYSIS AND COMPARISON OF GROUP PROPERTIES

By

ANNIE HOU, B.Sc.

A Thesis

Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Master of Science

McMaster University

©Copyright by Annie Hou, 2009.

MASTER OF SCIENCE (2009)

McMaster University
Hamilton, Ontario

TITLE: Galaxy Group Dynamics: Statistical Analysis and Comparison of Group Properties

AUTHOR: Annie Hou, B.Sc.(University of Toronto)

SUPERVISORS: Dr. L. Parker, Dr. W. Harris

NUMBER OF PAGES: viii, 107

Abstract

The dynamical state of galaxy groups at intermediate redshifts can provide information about the growth of structure in the universe. We examine three goodness-of-fit tests, the Anderson–Darling (A–D), Kolmogorov and χ^2 tests, in order to determine which statistical tool is best able to distinguish between groups that are relaxed and those that are dynamically complex. We perform Monte Carlo simulations of these three tests and show that the χ^2 test is profoundly unreliable for groups with fewer than 30 members. Power studies of the Kolmogorov and A–D tests are conducted to test their robustness for various sample sizes. We then apply these tests to a sample of the second Canadian Network for Observational Cosmology Redshift Survey (CNOC2) galaxy groups and find that the A–D test is more reliable and powerful at detecting real departures from an underlying Gaussian distribution than the more commonly used χ^2 and Kolmogorov tests. We use this statistic to classify a sample of the CNOC2 groups and find that 34 of 106 groups are inconsistent with an underlying Gaussian velocity distribution, and thus do not appear relaxed. In addition, we compute velocity dispersion profiles (VDPs) for all groups with more than 20 members and compare the overall features of the Gaussian and non-Gaussian groups, finding that the VDPs of the non-Gaussian groups are distinct from those classified as Gaussian. We also compare group properties of both rich individual groups and stacked groups to determine if any there are any trends amongst the classified Gaussian and non-Gaussian groups.

Acknowledgements

Firstly, I would like to thank my supervisors, Dr. Laura C. Parker and Dr. William E. Harris. I am very grateful to them for suggesting this research topic and introducing me to the wondrous field of galaxy groups. Their guidance and enthusiasm over the past two years has greatly motivated me. I am also thankful for their constant support and encouragement, which has allowed me to explore many avenues of my research project. I would also like to thank the members of the CNOC2 collaboration for many useful discussions and ideas, as well as the other member of my committee Dr. Alison Sills.

These past two years at McMaster would not have been the same without several of the graduate students. Thanks for all the coffee breaks, Phoenix meals, Casbah shows, discussions, both science and non-science related, bets, and for generally making living in Hamilton fun.

I would like to thank Andrew Fotos for being a constant source of love and support. You have kept me sane and happy through it all. I am truly grateful to you for being so understanding and knowing that sometime school/science just has to happen. Lastly, I would like to thank my parents, who have always given me an unwavering supply of love and support. I would not be where or who I am without your guidance and encouragement throughout my life. Most importantly, thank you for constantly nurturing my interests in science. Without all of the astronomy books, the telescope, comet watching, etc., I would have undoubtedly ended with a much more boring life!

Contents

1	Introduction	1
1.1	Galaxy Groups	1
1.2	Group Dynamics	5
2	Statistical Tools	9
2.1	Goodness-of-Fits Tests	9
2.1.1	The Pearson's χ^2 Test	9
2.1.2	The Kolmogorov Test	11
2.1.3	The Anderson–Darling Test	12
2.2	Monte Carlo Simulations	14
2.3	Power Studies of the Kolmogorov and Anderson–Darling Tests	20
3	Application of Tests to Data	23
3.1	The Data	23
3.2	Estimation of Distribution Parameters	27
3.3	Comparison of Tests	30
3.3.1	Statistical Analysis	30
3.3.2	Velocity Distributions	31
4	Classification of the CNOC2 Groups	35

4.1	Maximum Radius Cut	35
4.2	Computation of the μ and σ values for the CNOC2 Groups	39
4.3	Application of the Anderson–Darling Test	50
5	Comparison of Group Properties	57
5.1	Individual Group Properties	57
5.1.1	Velocity Dispersion Profiles	57
5.1.2	Colour-Radius	61
5.1.3	Stellar Mass	65
5.2	Stacked Group Properties	68
5.2.1	Velocity Dispersion Profiles	68
5.2.2	Colour - Magnitude	75
5.2.3	Colour-Radius	79
5.2.4	Star Formation Rate: O[II] Equivalent Widths	83
5.2.5	Stellar Mass	87
6	Conclusions	91
	Appendices	105
	Complementary Analysis	105

List of Figures

2.1	Monte Carlo Simulations of the χ^2 Test	15
2.2	Monte Carlo Simulation of Kolmogorov Test	17
2.3	Monte Carlo Simulation of Anderson–Darling Test	19
3.1	Luminosity Weighted Position Plots of CNOC2 groups	26
3.2	Monte Carlo simulations of the Gapper Estimator and rms dispersion	29
3.3	Velocity Distribution Histograms for various CNOC2 groups	33
4.1	Position Plots and Velocity Distributions of sample CNOC2 groups .	37
4.2	Group Membership Distribution	38
4.3	Comparison of Intrinsic Velocity Dispersions with Wilman et al. (2005b)- 14 hr Field	45
4.4	Comparison of Intrinsic Velocity Dispersions with Wilman et al. (2005b)- 21 hr Field	46
4.5	Comparison of Intrinsic Velocity Dispersions with Wilman et al. (2005b)- 2 hr Field	47
4.6	Comparison of Intrinsic Velocity Dispersions with Wilman et al. (2005b)- 9 hr Field	48
4.7	Histogram comparing σ_{int} and σ_{int} Wilman	49

5.1	Velocity Dispersion Profiles of the CNOC2 Groups.	60
5.2	Colour-Radius Plots for Gaussian Groups with $n \geq 20$	63
5.3	Colour-Radius Plots for Non-Gaussian Groups with $n \geq 20$	64
5.4	Stellar Mass versus Radius Plots for Gaussian Groups with $n \geq 20$	66
5.5	Stellar Mass versus Radius Plots for Non-Gaussian Groups with $n \geq 20$	67
5.6	Velocity Dispersion Profiles all Stacked Gaussian/non-Gaussian Groups	70
5.7	Velocity Dispersion Profiles for the Stacked $n < 20$ Gaussian/non-Gaussian Groups	72
5.8	Velocity Dispersion Profile for the stacked CNOC2 groups	74
5.9	Colour-Magnitude diagram of the stacked Gaussian Groups	76
5.10	$g - r$ histogram for the stacked Gaussian groups	76
5.11	Colour-Magnitude diagram of the stacked non-Gaussian Groups	78
5.12	$g - r$ histogram for the stacked non-Gaussian groups	78
5.13	Colour-Radius diagram of the stacked Gaussian Groups	81
5.14	Binned Colour-Radius diagram of the stacked Gaussian Groups	81
5.15	Colour-Radius diagram of the stacked non-Gaussian Groups	82
5.16	Binned Colour-Radius diagram of the stacked non-Gaussian Groups	82
5.17	OII versus Radius Plot of Stacked Gaussian Groups	85
5.18	OII versus Radius Plot of Stacked non-Gaussian Groups	86
5.19	Plot of Stellar Mass versus Radius of the stacked Gaussian Groups	88
5.20	Plot of Stellar Mass versus Radius of the stacked non-Gaussian Groups	89
1	Velocity Dispersion Profiles of the 9 X-ray Selected Local Groups	107

Introduction

1.1 Galaxy Groups

The group environment represents an intermediate size and density scale between individual galaxies and rich galaxy clusters. With roughly half of the present day galaxy population in groups (Geller & Huchra, 1983; Eke et al., 2005), this environment plays an important role in the formation and evolution of galaxies. Despite the fact that a large fraction of galaxies reside in groups, relatively little is known about the group environment, especially in comparison to field galaxies and rich clusters. Only recently have there been large scale surveys of galaxy groups, such as the second Canadian Network for Observational Cosmology (CNOC2) (Carlberg et al., 2001), the Two-degree Field Galaxy Redshift Survey (2dFGRS) (Eke et al., 2004), the Sloan Digital Sky Survey (SDSS DR4) (Yang et al., 2007) and the high-redshift second Deep Extragalactic Evolutionary Probe (DEEP2) (Gerke et al., 2005) group catalogs. The lack of observational data results from the extreme difficulty in finding the groups themselves. Unlike rich galaxy clusters, groups have relatively little contrast with respect to the background, making them difficult to detect optically, thus increasing the probability of interlopers and false detections (Mamon, 2007). Galaxy groups also tend to have significantly lower hot gas density, in comparison to clusters,

(Li & Yee, 2008), making detection of extended X-ray emission from groups difficult (see §2 of this chapter for further discussion) .

With no clear distinction between a “rich” group and a “poor” cluster, the most challenging aspect of studying groups is the relatively arbitrary way in which they can be defined. Even within the field there is large debate as to what exactly constitutes a group. Theoretical astronomers tend to view groups as galaxies in virialized systems sharing one dark matter halo, while optical astronomers have a looser definition, viewing groups as over-densities of galaxies in space. Groups not only suffer definitional problems at the “rich” end, but also at the lower limit of members, with optical astronomers requiring a minimum of 3 member galaxies and X-ray astronomers viewing groups as having tens of galaxies.

In order to identify group membership, both photometric and spectroscopic data is needed to ensure proximity in both position- and redshift-space. There are several group finding algorithms, and one of the more commonly used methods determines group membership using a friends-of-friends (FOF) search algorithm, first developed by Huchra & Geller (1982). The FOF algorithm searches through a catalog of galaxy positions, redshifts and magnitudes and identifies the first galaxy in the catalog not assigned to a group. It then searches for companion galaxies, or friends, within a projected separation in position-, D_{12} (Equation 1.1), and redshift-space, V_{12} (Equation 1.2), from the initial galaxy, defined as:

$$D_{12} = \frac{2 \sin(\theta/2) V}{H_0} \leq D_L(V_1, V_2, m_1, m_2) \quad (1.1)$$

where $V = (V_1 + V_2)/2$, V_1 and V_2 are the redshifts of the initial galaxy and its friend, m_1 and m_2 are their magnitudes, θ is their angular separations, D_L is the spatial linking length parameter and,

$$V_{12} = |V_1 - V_2| \leq V_L(V_1, V_2, m_1, m_2), \quad (1.2)$$

where V_L is the redshift linking length parameter. This process is then carried out for each companion galaxy added to the list of group members, and is repeated until no more companions are found.

Although this method appears relatively simple, the definitions of the linking lengths, D_L and V_L , are arbitrary, with no strict or standard values, and changing them can drastically alter group membership. If the linking lengths are too restrictive then the FOF algorithm will find only compact groups that are likely already virialized, omitting more dynamically complex groups that may include interacting galaxies. Conversely, if the linking lengths are too relaxed then the defined groups tend to have very large group-centric radii, sometimes on the order of several Mpc, resulting in systems that may not be real. However, one can optimize the FOF search parameters by first testing out the algorithm on simulated galaxies. Nolthenius & White (1987) used cold dark matter (CDM) N-body simulations to find optimal linking lengths to find groups in the Center for Astrophysics (CfA) redshift survey. In general, with the appropriate choice of linking lengths and high quality spectroscopic data, the FOF algorithm generally picks out groups that are true associations of galaxies, while minimizing the contamination of interloping galaxies due to chance alignments along the line of sight.

The majority of the available galaxy group catalogs have been generated using the FOF algorithm, but the need for observationally expensive spectroscopic data limits these catalogs to either the local Universe or small patches of the sky. Recently Li & Yee (2008) have developed a new group finding method, the probability friends-of-friends (pFOF) algorithm, which uses photometric-, rather than spectroscopic-redshifts, allowing group surveys to be extended to both larger areas of the sky and higher redshifts. This method uses the FOF algorithm in the RA/DEC position space and photometric-redshift probability densities in the direction along the line of sight. Although this method allows for wider/deeper surveys, the uncertainties in photometric-redshifts can be 50-100 larger than spectroscopically determined values,

resulting in higher false detection rates (Li & Yee, 2008). Also, the pFOF algorithm can only reliably detect groups with more than eight members (Li & Yee, 2008), while current group catalogs include groups with as few as three galaxies. Improvements in photometric-redshift techniques will undoubtedly reduce the current errors associated with the pFOF group finding algorithm.

Understanding the group environment provides not only information about the evolution of galaxies and their properties, but can also be used to probe the surrounding dark matter haloes. Yang et al. (2005) developed a halo-based group finding algorithm that defines groups as galaxies which share a common dark matter halo. The first step of this algorithm involves defining a potential group, and group center, using the FOF algorithm. From this potential group, Yang et al. (2005) then compute the total luminosity of the group and the mass-to-light ratio, which in turn allows for estimates of the halo mass and radius, the virial radius and velocity and finally the velocity dispersion. Group membership is then assigned based on the computed properties of the associated dark matter halo. The group center and halo properties are re-computed as new members are added and the process is repeated until no new members are assigned to the group. Yang et al. (2005) applied their halo-based method on mock catalogues and concluded that their method was more successful at assigning galaxies to common dark matter haloes than the traditional FOF algorithm.

The recent advances in both observational techniques and group-finding algorithms have significantly increased our ability to study galaxy groups, and the important role they play with respect to galaxy evolution, as well as their use as probes of dark matter haloes on smaller scales. These surveys allow for detailed studies of the group environment and comparison with both the cluster and field environment. In particular, one can compare galaxy evolution in groups and in clusters and investigate which star formation truncation mechanism dominates in each environment.

1.2 Group Dynamics

Although the majority of galaxies live in a group environment only a small fraction of galaxy groups are actually relaxed or virialized systems (Mamon, 2007). Since the velocity dispersions of groups are relatively low, the group environment allows for a higher rate of galaxy interactions and mergers (Brough et al., 2006) than clusters. Major and repeated minor mergers can greatly alter the morphologies of galaxies in groups (Osmond & Ponman, 2004), changing the overall group properties. Another consequence of the frequent galaxy-galaxy interactions in galaxy groups is its effects on the star formation rates (SFR) of the member galaxies (Zabludoff & Mulchaey, 1998; Hashimoto et al., 1998; Tran et al., 2001). In fact, it has been suggested that a large portion of the evolution of galaxy properties occurs in the group environment (Balogh et al., 2004). Presumably the dynamical state of groups is correlated with star formation, mergers and other indicators of galaxy evolution.

The dynamical states of galaxy clusters are generally investigated using X-ray observations, i.e., X-ray luminosity (L_X) and X-ray temperature (T_X), which can provide information on the group centroid and depth of the group potential well (Osmond & Ponman, 2004). Although this method is well-established for clusters, it is not always applicable to galaxy groups, as some groups are not sufficiently massive to produce detectable diffuse X-ray emission. In fact, X-ray emission is generally only found in groups with a high fraction of early-type galaxies, which contain a centrally dominant galaxy or brightest group galaxy (BGG) (Mulchaey & Zabludoff, 1998; Mulchaey et al., 2003; Osmond & Ponman, 2004). These X-ray bright groups are systems that are believed to have already reached dynamical equilibrium (Jeltema et al., 2006), and since we wish to study both relaxed and dynamically complex systems, X-ray temperature observations would not provide enough information, restricting us to galaxy groups in a small range of dynamical states. Fortunately, T_X is not the only way to obtain dynamical information. The virial theorem states that the line of sight dispersion of the *velocities* of the member galaxies is also a measure of the grav-

itational potential. Thus, in order to probe both virialized and dynamically complex systems, we use optical observations of the *velocity distributions* to investigate the dynamical state of galaxy groups.

The usual assumption is that the underlying velocity distribution of galaxy clusters and groups is Gaussian in nature, but this is strictly true only for systems in dynamical equilibrium. On the other hand, groups with non-Gaussian velocity distributions could mark systems in the process of a merger or those that are in the early stages of evolution. To study group (or cluster) dynamics we need a reliable method of distinguishing between relaxed systems with Gaussian velocity distributions and more complex systems with non-Gaussian dynamics.

Previous analyses of cluster velocity distributions have resulted in differing views on the dynamics of these systems. Yahil & Vidal (1977) used the a-test, u-test and Shapiro-Wilks W-test for non-normality to show that the observed radial-velocity distributions of clusters of galaxies, with as few as 10 and as many as 122 members, are always consistent with an underlying Gaussian distribution. More recently, evidence of substructure has been found in clusters (Dressler & Shectman, 1988; Bird, 1994; Burns, 1998), indicating that cluster dynamics may be more complicated than initially assumed.

Beers et al. (1990) emphasize the difficulty in determining that a given velocity distribution differs significantly from Gaussian, stating that the goodness-of-fit tests used by Yahil & Vidal (1977) are sensitive to different departures from a true Gaussian distribution. That is to say one test may be sensitive to skewness, rejecting systems with asymmetrical shapes, but may not be able to detect non-Gaussian features in the wings, or edges, of the velocity distribution. Thus, the same system may be classified as either Gaussian or non-Gaussian depending on the statistic used. These difficulties are more severe when studying smaller systems, as in the case with galaxy groups. Since group membership can range from 3 to 50 or more galaxies, we need a statistical test that is reliable even for extremely small sample sizes. We also require

a test that is robust, or unaffected by small departures from normality, to ensure that the rejections are a result of real deviations from a Gaussian distribution and not sensitivities inherent to the test. The goodness-of-fit tests used to analyze rich clusters are generally not applicable to groups, where the challenges of small number statistics become relevant.

In this thesis we test three goodness-of-fit tests in order to determine which one in particular is best able to distinguish between Gaussian and non-Gaussian velocity distributions, especially for small sample sizes. In Chapter 2, we discuss the statistical tools we use to determine departures from Gaussianity and present the results of our Monte Carlo simulations of the χ^2 , Kolmogorov and Anderson–Darling tests, as well as our power studies of the Kolmogorov and Anderson–Darling tests. In Chapter 3, we apply the goodness-of-fit tests to the CNOC2 galaxy group data and compare the results. In Chapter 4, we apply the Anderson–Darling test to the CNOC2 group catalog and classify the dynamical state of the groups. In Chapter 5, we compare the group properties of both individual and stacked Gaussian and non-Gaussian groups and discuss the implications of our results and in Chapter 6 we summarize our findings.

Throughout this thesis we use Λ CDM cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

Chapter 2

Statistical Tools

2.1 Goodness-of-Fits Tests

Since groups can contain anywhere between 3 to 50 (or more) member galaxies, proper statistical analysis of groups requires a goodness-of-fit test that is reliable and robust for even extremely small sample sizes. The Pearson's χ^2 test (Equation 2.1) is arguably the most commonly used goodness-of-fit test. However, this statistic was developed as a large sample theory and its reliability begins to break down as one approaches small sample sizes (n). For small n , D'Agostino & Stephens (1986) (hereafter DA86) suggest the use of goodness-of-fits tests based on Empirical Distribution Functions (EDFs), such as the Anderson–Darling (A–D) and the Kolmogorov tests.

2.1.1 The Pearson's χ^2 Test

The Pearson's χ^2 Test is defined as:

$$\chi^2 = \sum_{i=1}^n \frac{(\text{observed}_i - \text{expected}_i)^2}{\text{expected}_i} \quad (2.1)$$

The critical points of the χ^2 test were developed for large samples and it is difficult to determine exactly where, at small n , this theory begins to break down (DA86). Although there is no standard method to apply the χ^2 test, it is often suggested that the bin widths be constructed such that they have the same number of data points in each bin. For small sample sizes Vessereau (1958) claims that it is not the number of data points per bin that is important but rather the *total* number of data points. Vessereau (1958) also finds that as long as $n \geq 10$ and one uses the 1% or 5% critical values, the χ^2 test should not produce a significant amount of error (with respect to false positives or negatives). Koehler & Larntz (1980) claim that for $k-1$ degrees of freedom, the Pearson's χ^2 test is 'reasonably adequate' when $k \geq 3$ and $n \geq 10$ and Roscoe & Byars (1971) also find that when the degrees of freedom are > 1 , the χ^2 test does not produce significant false positives, remaining robust against these type of errors.

The χ^2 test is often performed on binned data, testing the variance between histograms and a continuous Gaussian distribution. Unfortunately, binning can add additional errors to the test, a problem which becomes worse at small n . The choice of bin width can alter the results of the χ^2 test, even causing the same data to be classified as both Gaussian and non-Gaussian, depending on the bin width. Several authors (e.g., Heald (1984); Scott (1979)) have attempted to reduce this effect by minimizing the sampling fluctuations. With testing, Heald's optimal bin width was selected, where the bin width is defined as:

$$\delta x = \sigma \left(\frac{20}{n} \right)^{1/5} \quad (2.2)$$

where σ is the standard deviation.

The degrees of freedom in the system are given by $DOF = N - 1 - k$, where N is the number of bins and k is the number of free parameters in the assumed distribution. Thus, for a Gaussian distribution there are two free parameters ($k = 2$),

μ and σ , and the minimum number of bins is $N = 4$.

2.1.2 The Kolmogorov Test

The popular Kolmogorov test (DA86 and references therein) is a goodness-of-fit test based on supremum statistics, which makes use of the existence of a unique least upper bound or maximum value in the data set. For the Kolmogorov test, the supremum values is the measured maximal vertical difference between the empirical distribution function (EDF), $F_n(x)$, of the *ordered* data x_i and the cumulative distribution function (CDF), $F(x)$, of a given model ¹. The EDF statistic computed for the Kolmogorov test is the D value, which is derived from the D^+ and D^- values, and is defined as:

$$D = \max(D^+, D^-) \quad (2.3)$$

$$D^+ = \supremum \left| \frac{i}{n} - F(x) \right|, \quad (2.4)$$

$$D^- = \supremum \left| F(x) - \frac{(i-1)}{n} \right|, \quad (2.5)$$

where $F_n(x) = \frac{i}{n}$ for D^+ , $F_n(x) = \frac{(i-1)}{n}$ for D^- and $1 \leq i \leq n$.

Stephens (1974) has simplified the Kolmogorov test with the modification of the D -values, called the D^* -value (Equation 2.6), which allows for comparison with a single critical value table, rather than computing critical values for specific sample sizes and significance levels (Massey, 1951).

$$D^* = D \left(\sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}} \right) \quad (2.6)$$

¹It should be noted that the Kolmogorov-Smirnov test is technically the difference between two EDFs, while the Kolmogorov Test measures the difference between a model CDF and EDF. Despite this distinction, the Kolmogorov test is often referred to as the "Kolmogorov-Smirnov test", here we will refer to it by its proper name.

The use of the Kolmogorov test in place of the χ^2 test for small samples is suggested by Lilliefors (1967). In a comparison of the Kolmogorov and χ^2 tests, Massey (1951) concludes that the Kolmogorov test is generally more reliable than the χ^2 test, especially for small n , where the effects of binning, required by the χ^2 test but not the Kolmogorov test, can result in a large loss of information.

2.1.3 The Anderson–Darling Test

Like the Kolmogorov test, the Anderson–Darling (A–D) test is also based on EDF statistics and does not require binning or graphical analysis. Despite these advantages, the A–D test is not commonly used in astronomy. The A–D statistic involves the calculation of the A^2 and A^{2*} values, starting from the *ordered* data $\{x_i\}$:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) (\ln \Phi(x_i) + \ln(1 - \Phi(x_{n+1-i}))), \quad (2.7)$$

$$A^{2*} = A^2 \left(1 + \frac{0.75}{n} + \frac{2.25}{n^2} \right) \quad (2.8)$$

where $x_i \leq x < x_{i+1}$, $\Phi(x_i)$ is the CDF of the hypothetical underlying distribution and the A^{2*} values are the modified statistics for cases where the distribution parameters are not known *a priori*. From Equation 2.8, it is clear that A^{2*} approaches A^2 for large n .

In this analysis, we take $\Phi(x_i)$ to be the CDF of a Gaussian distribution, given as:

$$\Phi(x_i) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x_i - \mu}{\sqrt{2}\sigma} \right) \right) \quad (2.9)$$

where x_i are the radial velocities of the galaxy group members, arranged from lowest

to highest, μ is the computed mean velocity of the group and σ is the computed velocity dispersion. The $\Phi(x_i)$ values are then used in the A–D computing formulas (Equations 2.7 and 2.8) to obtain the A^{2*} values. These values can either be compared to known critical or limiting value tables, or used to compute the significance level α , which gives the probability of the null hypothesis (i.e., the underlying distribution is Gaussian) being true. A more detailed discussion of critical values and significance levels is given in Chapter 4. DA86 recommend the A–D test as the ‘omnibus’ test for EDF statistics when the underlying distribution is believed to be Gaussian. Furthermore, they claim that the A^2 and A^{2*} values can be reliably computed down to $n = 5$.

A method of quantifying the robustness of statistical tests involves conducting power studies, which investigates the percentage of false positives a given test may produce when the underlying distribution is distorted (i.e., skewed, shifted, wings, etc.). For the Gaussian distribution, this involves applying the tests to a variety of non-Gaussian samples and determining not only how often a specified test will reject the distribution as Gaussian, but also the specific types of departures from non-normality that affect the rejection rate (DA86 and references therein). In a comparison of the Kolmogorov, Cramér-von Mises, Kuiper, Watson and A–D tests, Stephens (1974) conducted power studies using a variety of non-Gaussian distributions and found the A–D test to be most powerful of the EDF statistics for detecting departures from Gaussianity, while the Kolmogorov test proved to be least powerful. As for the χ^2 test, DA86 claim that it is in general not a powerful test for Gaussian distributions and do not recommend its use. We discuss the results of our power studies in §2.3.

2.2 Monte Carlo Simulations

We test the claim of Stephens (1974) and DA86 that the A–D test is the most reliable test for Gaussianity for small sample sizes, by performing Monte Carlo simulations of the χ^2 , Kolmogorov and A–D tests. We perform 30 000 iterations for a variety of sample sizes, with $5 \leq n \leq 50$ and drawing from a random Gaussian distribution with input values $\mu = 0$, $\sigma = 1.0$, to determine the reliability of the tests and how accurately each test can reproduce published critical values.

The results of the χ^2 Monte Carlo simulations are shown in Figure 2.1, where we have plotted histograms of χ^2/DOF for our various sample sizes. Ideally a peak should occur at $\chi^2/DOF = 1.0$ for the χ^2 test, but we see that there is significant scatter in the histograms for $n < 30$ (Figure 2.1). Most notable is the $n = 5$ histogram in Figure 2.1, which has two peaks at $\chi^2/DOF > 10$, much higher than the $\alpha = 0.10$ critical value (i.e., the χ^2 value above which 10% of values fall) of 2.41 (DOF = 1). Thus for small n , the χ^2 test tends to overestimate the number of failed/non-Gaussian samples. It is only for $n \geq 30$ that we see the expected peak of 1.0. Unfortunately, the majority of the CNOC2 groups have $n_{\text{members}} < 30$, so use of the χ^2 test to classify galaxy group dynamics could result in the false identification of many groups with non-Gaussian velocity distributions.

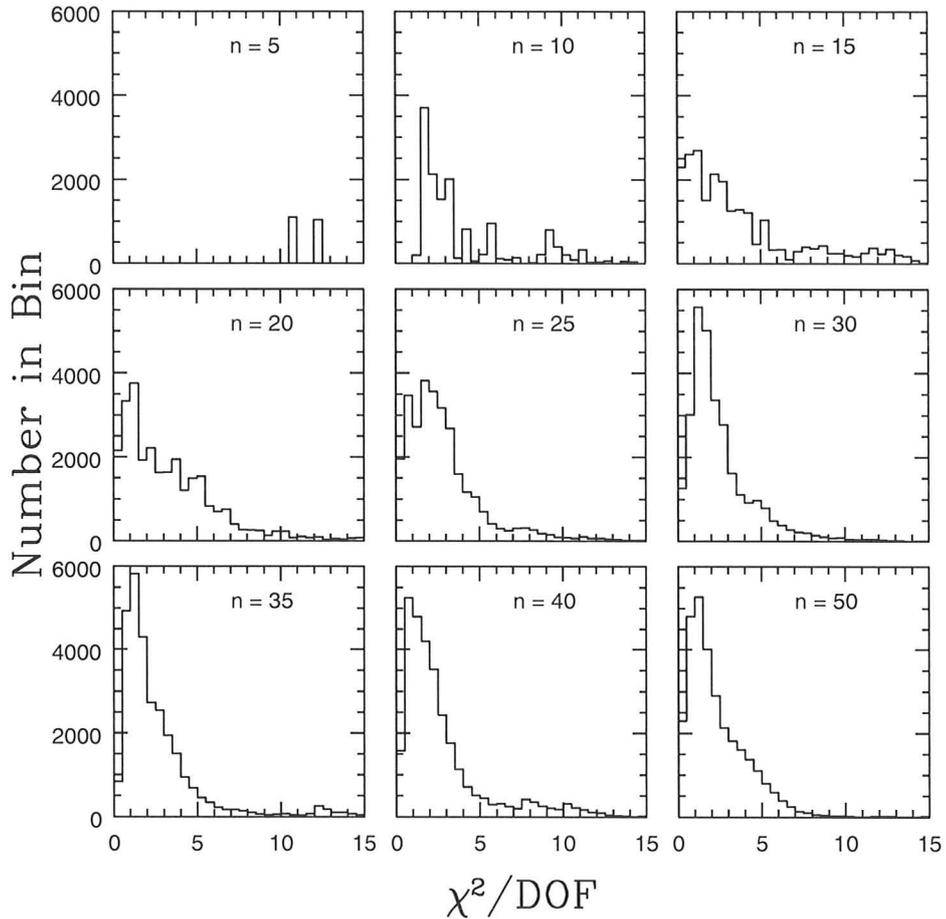


Figure 2.1: Monte Carlo simulations of the χ^2 test for different sample sizes, using a Gaussian random number generator with input values of $\mu = 0.0$ and $\sigma_{\text{int}} = 1.0$, and 30 000 iterations. The histograms are generated with a bin width of 0.5. For $n < 30$, the χ^2 test completely fails to recover the input distribution. This is most obvious for the $n = 5$ case, where the histogram shows two peaks at $\chi^2/\text{DoF} = 10, 12$, instead of the expected peak value of $\chi^2/\text{DoF} = 1.0$. Only for $n > 30$ do the simulations recover the expected peak value. It should be noted that for the $n = 5$ case, the remainder of the values fall at $\chi^2/\text{DOF} > 15$.

The results of the Kolmogorov Monte Carlo simulations are shown in Figure 2.2, where we show histograms of the computed D^* values and indicate the $\alpha = 0.10$ critical value of 1.224 (DA86) with a dotted vertical line. Unlike the histograms for the χ^2 test, we see no scatter in the D^* values, even at $n = 5$. The histograms of the D^* values remain remarkably stable over the sample size range, indicating the the Kolmogorov test is reliable for small n . To determine if our simulations produce the expected D^* critical values, we compute these values from our histograms (Figure 2.2) and compare them to published values cited in DA86. For each sample size, we are able to reproduce all of the given critical values for the case where the input distribution parameters, μ and σ , are known.

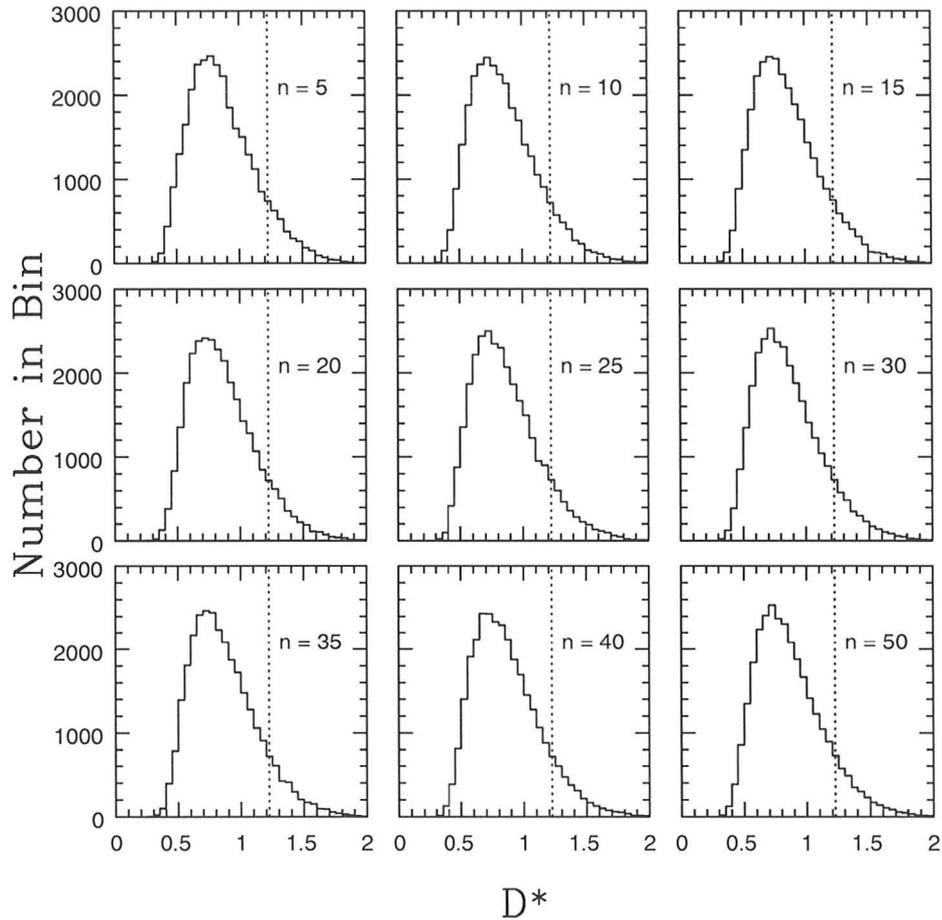


Figure 2.2: Monte Carlo Simulation of Kolmogorov test for different sample sizes, using a Gaussian random number generator with input values of $\mu = 0.0$ and $\sigma_{\text{int}} = 1.0$, and 30 000 iterations. The histograms of the D^* values are plotted using a bin width of 0.05. The results of the simulations show that even at $n = 5$, the Kolmogorov test is able to recover the input distribution. The histograms for $n = 5$ to 50, consistently reproduce the expected peak values, indicating that the test is reliable over a wide sample range. The vertical dotted line represents the D^* value above which 10% of the values lie, our computed values are in complete agreement with the known $\alpha = 0.10$ critical value of 1.224 (DA86).

In §2.1, we discussed the use of the A^{2*} values for the A–D test, but this is a modification for the case where the distribution parameters are unknown. When the input parameters are known *a priori*, as in the case with Monte Carlo simulations, DA86 state that no modification for the A–D test is needed and one should use the A^2 values when comparing to critical value tables. The results of the A–D test simulations are shown in Figure 2.3, where we plot histograms of the computed A^2 values and indicate the $\alpha = 0.10$ critical value of 1.933 (DA86) with a dotted vertical line. The histograms are similar to those of the Kolmogorov test, showing no scatter in the A^2 values over the entire sample size range. The stability of the A–D test, even at $n = 5$, supports the claim of DA86 that the statistic is reliable for all $n \geq 5$. We compute the $\alpha = 0.10$ critical values for the A–D test from the histograms in Figure 2.3 and find that our values are in agreement with those found in DA86.

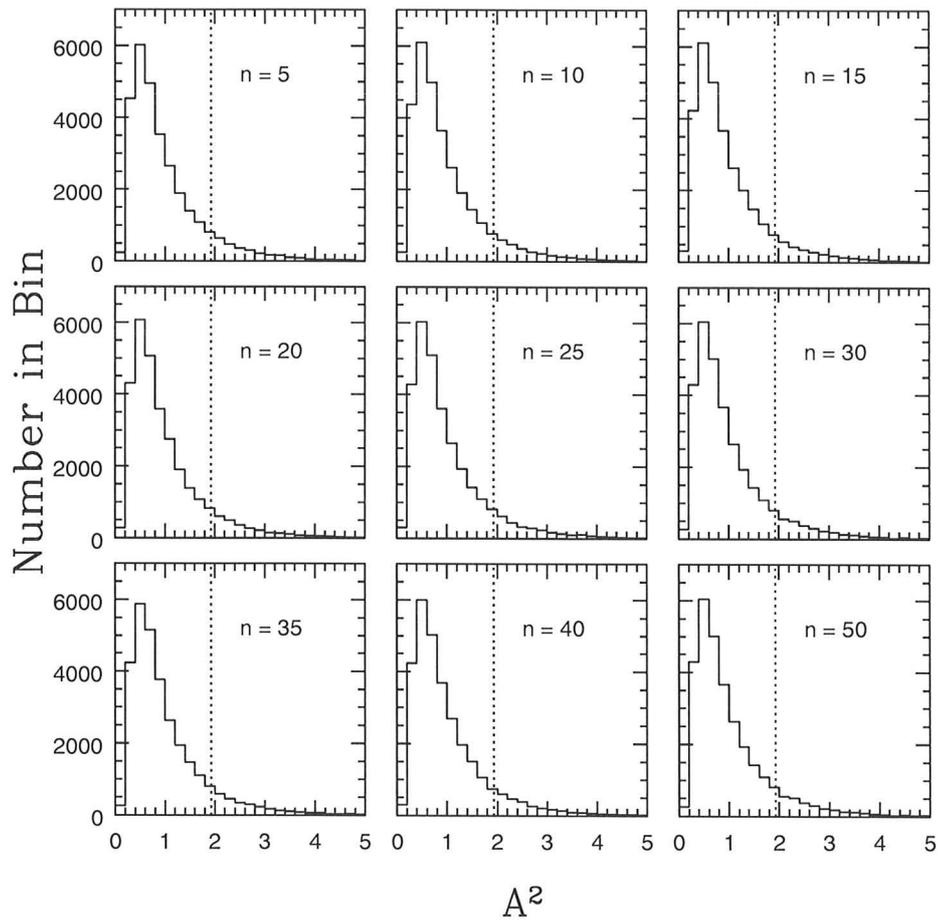


Figure 2.3: Monte Carlo Simulation of Anderson–Darling test for different sample sizes, using a Gaussian random number generator with input values of $\mu = 0.0$ and $\sigma_{\text{int}} = 1.0$, and 30 000 iterations. The histograms of the A^2 values are plotted using a bin width of 0.2. The results of the simulations show that even at $n = 5$, the A–D test is able to recover the input distribution. The histograms for $n = 5$ to 50, consistently reproduce the expected peak values, indicating that the test is reliable over a wide sample range. The vertical dotted line represents the A^2 value above which 10% of the values lie. Our computed values are in complete agreement with the known $\alpha = 0.10$ critical value of 1.933 (DA86).

2.3 Power Studies of the Kolmogorov and Anderson–Darling Tests

Our Monte Carlo simulations indicate that the χ^2 test is indeed an unreliable statistic for testing Gaussianity in small sample sizes, as suggested by DA86. Thus, we can eliminate the use of this test for classification of the galaxy group dynamics. However, the results of our simulation also show that both the Kolmogorov and A–D statistics are reliable down to $n = 5$. In order to determine which test is best able to accurately distinguish between Gaussian and non-Gaussian systems, we perform power studies of the Kolmogorov and A–D tests. Monte Carlo simulations of both tests are performed, using a skewed Gaussian distribution, $F(x_i)$ (Azzalini & Capitanio, 1999), given by:

$$F(x_i) = 2\phi(x_i)\Phi(\alpha_S x_i) \quad (2.10)$$

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (2.11)$$

where $\phi(x)$ is the probability density function of a Gaussian distribution, with $\mu = 0.0$ and $\sigma = 1.0$, $\Phi(x_i)$ is the CDF of a Gaussian distribution (Equation 2.9) and α_S is known as the shape or slant parameter. The value of α_S is proportional to the skewness of the Gaussian distribution, with higher values of α_S producing more skewed distributions and an $\alpha_S = 0.0$ producing a non-skewed Gaussian.

We draw various sample sizes, $5 \leq n \leq 100$, from a Gaussian random distribution and apply varying levels of skewness, $0 \leq \alpha_S \leq 5$, to determine the rejection rate of both the Kolmogorov and A–D tests. The results of our power studies are given in Table 2.1, where Column 1 indicates the test used, Column 2 indicates the α_S value, and Columns 3 - 9 indicate the percentage of rejection given a specific sample size. The rejection rates are determined using the 10% critical values given in DA86. In

order for a test to be considered powerful, the simulations with high α_S should have higher rejection rates, since the underlying distribution are increasingly less Gaussian. From Table 2.1, it is clear that for all levels of skewness and all sample sizes, the A–D statistic rejects more objects than the Kolmogorov test. For the $\alpha_S = 0.25$ and $n \leq 30$ cases the two tests are comparable, but as one increases α_S , the percentage of rejections for the Kolmogorov test are significantly lower than those of the A–D test. Looking at the rejection rates for the $\alpha_S = 1.0$, which is a heavily skewed Gaussian, it is clear that the Kolmogorov test underestimates the amount of non-Gaussian systems. This is especially evident when one focuses on the $n = 30$ case for the $\alpha_S = 1.0$ distribution, which has a 100% rejection rate for the A–D test, but only an 85% failure rate for the Kolmogorov test.

The results of our power studies indicate that the Kolmogorov test is much less powerful than the A–D statistic, which is in agreement with the findings of Stephens (1974). The strongest evidence for this claim is shown in the $\alpha_S = 5.0$, a completely non-Gaussian distribution, and $n = 5$ simulation, where the A–D test rejects 100% of the sample while the Kolmogorov test only rejects 75% of the sample.

Based on our Monte Carlo simulations and power studies, we conclude that the A–D test is the most reliable and power statistics, and thus is best able to distinguish between dynamically relaxed (Gaussian) and complex (non-Gaussian) galaxy groups.

Application of Tests to Data

3.1 The Data

The second Canadian Network for Observational Cosmology Redshift Survey (CNOC2) was conducted in the redshift range of $0.1 < z < 0.6$ and obtained photometry for $\sim 4 \times 10^4$ galaxies covering four patches, 1.5 deg^2 in area, in the $UBVR_CI_C$ bands down to a limiting magnitude of $R_C = 23.0$ (Yee et al., 2000). Spectra of more than 6000 galaxies were taken with the MOS spectrograph on the Canada-France-Hawaii Telescope (CFHT), with 48% completeness at $R_C = 21.5$ (Yee et al., 2000). Over 200 galaxy groups were identified using a FOF percolation algorithm in the CNOC2 survey (Carlberg et al., 2001).

Wilman et al. (2005b) (hereafter W05) obtained deeper spectroscopy of 20 selected CNOC2 groups and 6 serendipitous groups in the redshift range of $0.3 \leq z \leq 0.55$, with the Multi-Object Spectrograph Low Dispersion Survey Spectrograph (LDSS2) on the 6.5 m Magellan (Baade) telescope at Las Campanas Observatory in Chile. W05 redefined group membership with more relaxed algorithm parameters than those used by Carlberg et al. (2001). The original search parameters were optimized so that the group-finding algorithm would identify dense, virialized groups, while the W05 sample included looser group populations. We would like to quantify how many of these

groups have more complex velocity distributions, potentially identifying merger products or systems in the early stages of virialization.

The centroid of the CNOC2 groups is defined using a luminosity-weighted center (W05). From the viewpoint of hierarchical structure formation, one would expect that at the center of rich clusters lies a brightest cluster galaxy (BCG). Brough et al. (2006) suggests that these BCGs form in the group environment and that groups with extended X-ray emission tend to have a galaxies similar to BCGs close to the center of their X-ray emission. Unfortunately, not all groups are massive enough to have detectable X-ray emissions and do not always have obvious brightest group galaxies (BGG). In some cases, even when a BGG does exist, it does not necessarily lie at the center of the potential well (Osmond & Ponman, 2004). Brough et al. (2006) find that only groups in complete dynamical equilibrium have a BGG at the group centroid. Thus, the luminosity-weighted centroid does not necessarily lie within a particular galaxy, causing some debate as to how to properly define the group center (see Chapters 3.2 and 4.2 for further discussion). The somewhat arbitrary definition of the group centroid can lead to problems in defining galaxy member velocity offsets. In theory, one would expect the mean group velocity offset, $c(z_{\text{member}} - z_{\text{centroid}})$, and the redshift of the group centroid to have similar values, as in the case of Group 208, whose group centroid is located at a redshift of 0.269 and has a mean $c(z_{\text{member}} - z_{\text{centroid}})$ of $\approx 0 \text{ km s}^{-1}$. However, for a large fraction of the CNOC2 groups these values can be drastically different. For example, Group 237 is located at a redshift of 0.399, while the mean group velocity offset is $\sim 433 \text{ km s}^{-1}$. Assuming the equivalency of these two values can drastically alter the results of statistical tests that require accurate information about the mean, μ , of the group (see Chapter 3.2 for further discussion).

In Figure 3.1 we plot R-band luminosity weighted position plots for 4 CNOC2 groups, with member galaxies indicated by open circles and the group centroid (W05) indicated by a red cross, to illustrate the effects of defining a luminosity weighted

group centroid. Group 4 (top-left in Figure 3.1) contains 20 member galaxies and it is clear that the computed group centroid lies at the center of the brightest galaxy. Group 238 (top-right in Figure 3.1) has 21 members and although the group centroid does not lie exactly at the center of the brightest galaxy, it is relatively close to the centers of the three brightest galaxies in this system. Groups 117 (bottom-left in Figure 3.1) highlights the fact that even if there is a BGG, the luminosity weighted centroid does not necessarily have to correspond with the center of the BGG. The high fraction of galaxies below the BGG in Group 117 causes the centroid to lie completely outside of the BGG and closer to large grouping of less luminous galaxies. Group 320 (bottom-right in Figure 3.1) also contains 20 member galaxies and it is an example of a system with no obvious BGG, as there are 4 member galaxies with relatively similar R-band luminosities. The computed group centroid lies within a close grouping of 14 galaxies and is somewhat close to the center of one of the brightest galaxies. This group also illustrates the problem with assuming that the mean group velocity offset is equivalent to the redshift of the group centroid, with $\bar{z} = 0.245$ and a mean $c(z_{\text{member}} - z_{\text{centroid}})$ of $\approx 56 \text{ km s}^{-1}$.

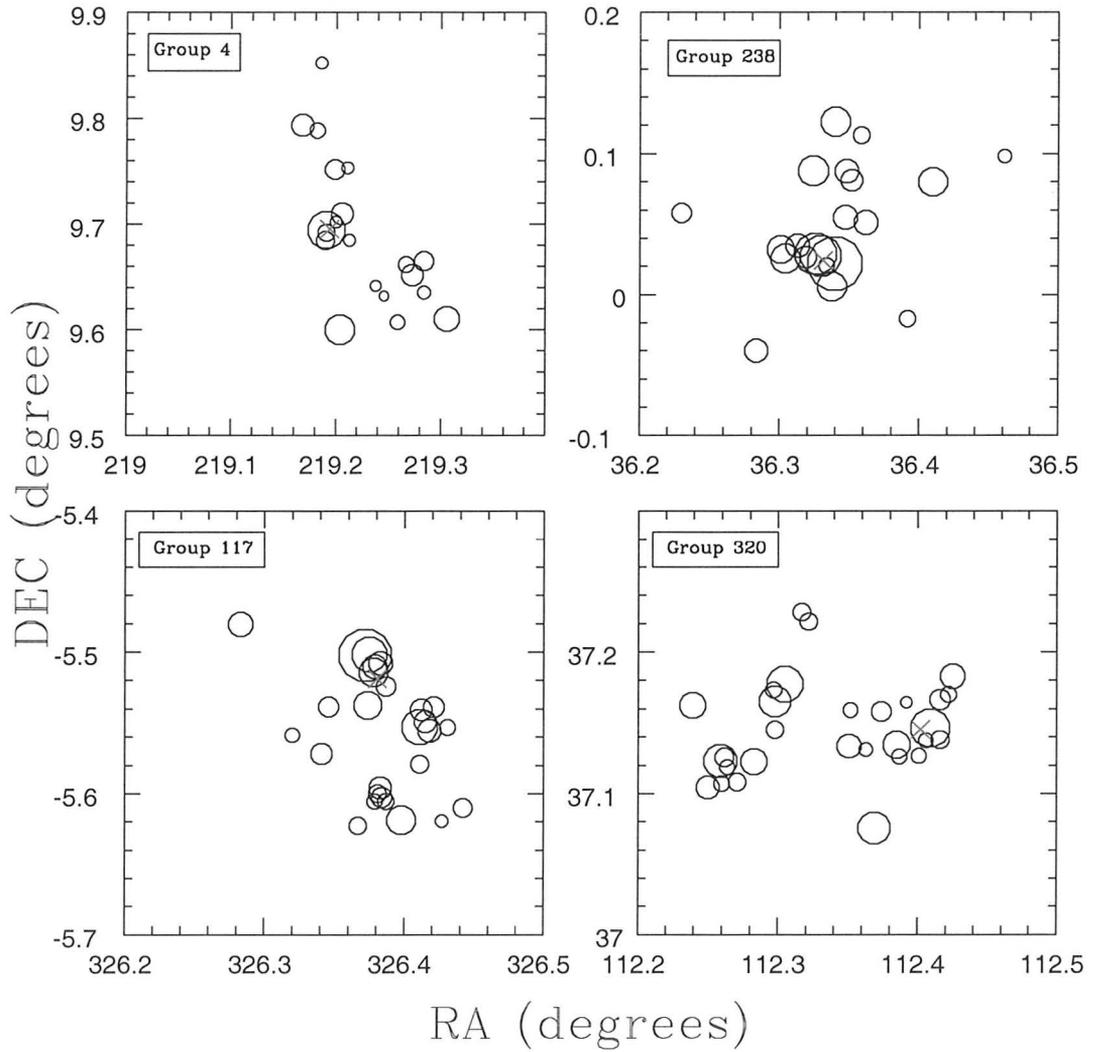


Figure 3.1: R-band luminosity weighted position plots of 4 CNOC2 groups, where larger symbols correspond to more luminous galaxies and the red crosses indicate the luminosity weighted group centroid.

3.2 Estimation of Distribution Parameters

The χ^2 , Kolmogorov and A–D statistics were developed under the assumption that all of the parameters of the underlying distribution were completely specified. Modifications to the statistics, with the use of Monte Carlo simulations, have been carried out to allow these tests to be applied to cases where the distribution parameters are not completely known a priori, but may be estimated from the data (DA86).

The parameters required to define a Gaussian distribution are the mean, μ , and the dispersion, σ . In the analysis described in §3 of this chapter, μ is calculated using the standard mean, and the velocity dispersions are estimated with the Gapper Algorithm, given by:

$$\sigma_{\text{Gapper}} = \frac{\sqrt{\pi}}{n(n-1)} \sum_{i=1}^{n-1} w_i g_i \quad (3.1)$$

where $w_i = i(n-i)$, $g_i = x_{i+1} - x_i$, here the *ordered* x_i values are given by the observed radial velocities of the group members. For small number statistics, Beers et al. (1990) recommend the Gapper Estimator over the canonical rms standard deviation, as this algorithm is insensitive to outliers and thus more accurately reproduces the true dispersion of the system.

To ensure that the Gapper Estimator is truly a more reliable method of computing the dispersion of a system, we perform Monte Carlo simulations of both the Gapper algorithm and the canonical rms standard deviation. We draw various sample sizes ($n = 5, 15, 20, 50$) from a Gaussian random number generator (Galassi, 2006) with the inputs $\mu = 0.0$ and $\sigma_{\text{intrinsic}} = 100$ and then compute the dispersion using the two aforementioned methods.

The results are shown in Figure 3.2, where we have plotted histograms of the output dispersions for each sample size. Figure 3.2 that for the $n = 5$ case the canonical rms standard deviation underestimates the true dispersion by 25% and

the distribution is heavily skewed to lower values. Although the distribution of the Gapper Estimator is also skewed, the peak of the distribution occurs at the true dispersion value of 100, indicating that this method is indeed insensitive to outliers. As we increase the sample size, $n = 15$ and 20 , we can see that the rms dispersion continues to underestimate the true velocity dispersion, but also that the two methods begin to converge. The histogram for the $n = 50$ case shows that the rms dispersion and the Gapper Estimator both correctly identify the true dispersion. Thus, based on these results we choose to use the Gapper Estimator to compute the group velocity dispersion.

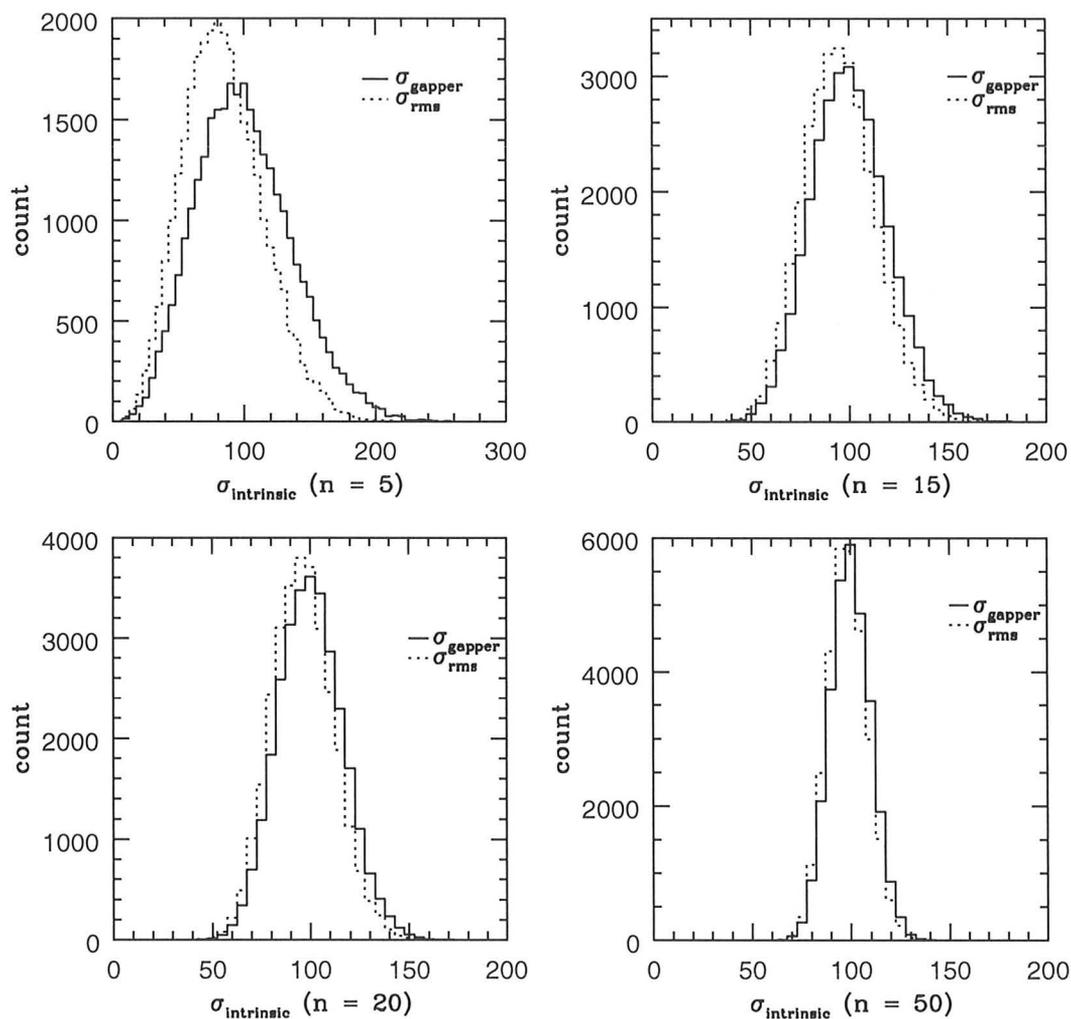


Figure 3.2: Monte Carlo simulations of the Gapper Estimator and the rms dispersion methods for $n = 5$ (top left), $n = 15$ (top right), $n = 20$ (bottom left) and $n = 50$ (bottom right). For each histogram, we compute the Gapper and rms dispersions using velocities generated from a random Gaussian distribution with input values of $\mu = 0.0$ and $\sigma_{\text{int}} = 100$. This process was done with 30 000 iterations. The dotted lines indicate dispersion values computed using the canonical rms standard deviation and the solid lines are values calculated using the Gapper algorithm.

3.3 Comparison of Tests

3.3.1 Statistical Analysis

The χ^2 , Kolmogorov and A–D tests are applied to a subset of 62 CNOC2 groups containing at least 10 members per group. We chose this membership cut based on the result of our $n = 5$ χ^2 Monte Carlo simulation, shown in Figure 2.1, which indicated that the test was completely unreliable at small n . This cut also helps to minimize the effect of binning. The potential complications introduced by the choice of bin width are also reduced by using the optimal bin width formula (Equation 2.2).

Despite our attempts to minimize the various uncertainties due to binning, we still encounter groups that do not meet the minimum number of bins requirement, $N = 4$, for our application of the χ^2 test, with 5 out of the 62 groups in the sub-sample having $N < 4$. Thus, we apply the χ^2 test to 57 groups ($n \geq 10$ and $N \geq 4$).

The modification to the D -statistic (Equation 2.6) given in §2.1 is used for the case when the input distribution parameters are known *a priori* (i.e., Monte Carlo simulations in §2.2). When applying the Kolmogorov test to real data sets, one must estimate the distribution parameters, μ and σ , and thus the modification for the D -value becomes (DA86):

$$D^* = D \left(\sqrt{(n)} - 0.01 + \frac{0.85}{\sqrt{(n)}} \right) \quad (3.2)$$

where D is given by equation 2.3.

Similarly, in §2.2 we used the A^2 values, but for real data sets one must use the A^{2*} values (Equation 2.8) in order to properly apply the A–D statistic. Since, the A–D and Kolmogorov tests do not require binned data, we are able to apply both tests to all 62 groups ($n \geq 10$).

The results of our analysis are presented in Table 3.1, where Column 2 indicates the number of groups used to perform the specified test and Column 3 indicates the

number of groups that failed at the 0.10 significance level. Our initial findings show that these three tests differed in the number of rejected (non-Gaussian) groups, with a 21% rejection rate for the χ^2 test, a $\sim 11\%$ rejection rate for the Kolmogorov test and a $\sim 16\%$ rejection rate for the A–D test.

Table 3.1: Results of the χ^2 , Kolmogorov and A–D tests applied to a sample of CNOC2 groups with $n \geq 10$.

Test	Number of Groups	Number of Failed Groups	Percentage of Failed Groups	Significance Level
Pearson's χ^2	57	12	$\sim 21\%$	0.10
Kolmogorov	62	7	$\sim 11\%$	0.10
A–D	62	10	$\sim 16\%$	0.10

The D -statistic of the Kolmogorov test is often used to test for goodness-of-fit, but both Stephens (1974) and DA86 do not recommend its use for testing Gaussian distributions, based on its lack of power. The results of our own power studies, in §2.3, also indicate that the Kolmogorov test lacks power and is unable to detect real departures from Gaussianity. The relatively low non-Gaussian detection rate in the CNOC2 group sub-sample, with only 7 out of the 62 CNOC2 groups failing at the 0.10 significance level, further supports the notion that the Kolmogorov test suffers from under-rejection.

3.3.2 Velocity Distributions

The reliability of the A–D test over the χ^2 and Kolmogorov tests is further demonstrated when one looks at the velocity distributions of specific groups. In Figure 3.3, we show the velocity distributions of four CNOC2 groups that have either failed the χ^2 , Kolmogorov and A–D tests (non-Gaussian groups), passed all three tests

(Gaussian groups), failed the χ^2 test but passed the Kolmogorov and A–D tests, or have passed the χ^2 and Kolmogorov tests but failed the A–D test. The histograms are made using Heald’s optimal bin width (Equation 2.2) and are over-plotted with a Gaussian distribution generated using the estimated mean and dispersion of the group. Although the Kolmogorov and A–D tests do not use binned data, we can look at the velocity distributions of groups classified as Gaussian or non-Gaussian by the various tests to see if there are obvious visual departures from normality.

Group 110 is classified as Gaussian by the χ^2 , Kolmogorov and A–D tests and from Figure 3.3 we can see that the shape and mean of the velocity distribution agrees well with the underlying Gaussian distribution. Group 208 (Figure 3.3) is a group that has been classified as non-Gaussian by all three tests, and it is evident that the velocity distribution is non-Gaussian, as the histogram shows a double peak.

While Groups 110 and 208 are examples of systems with obvious Gaussian or non-Gaussian features in their velocity distributions, this distinction is not so clear for several of the CNOC2 groups. Group 366 represents groups that have failed the χ^2 test, but have passed the Kolmogorov and A–D tests. The histogram for Group 366 (Figure 3.3), shows no obvious departures from the Gaussian distribution and despite the use of Heald’s optimal bin width, the relatively low group membership ($n = 15$) results in binning issues and causes the group to be rejected by the χ^2 test. The Kolmogorov and A–D tests use *ordered* rather than binned data, and thus do suffer from the issues introduced when binning small sample sizes.

Group 239 represents groups that have passed the χ^2 and Kolmogorov tests, but not the A–D test. This histogram for this group (Figure 3.3) appears skewed and has a mean that is slightly shifted from that of the Gaussian distribution. Despite these non-Gaussian features, only the A–D test is able to detect these departures and rejects Group 239 as having an underlying Gaussian distribution. This group highlights the fact that the Kolmogorov test lacks power, as discussed in §2.3, and is unable to detect slight departures from Gaussianity.

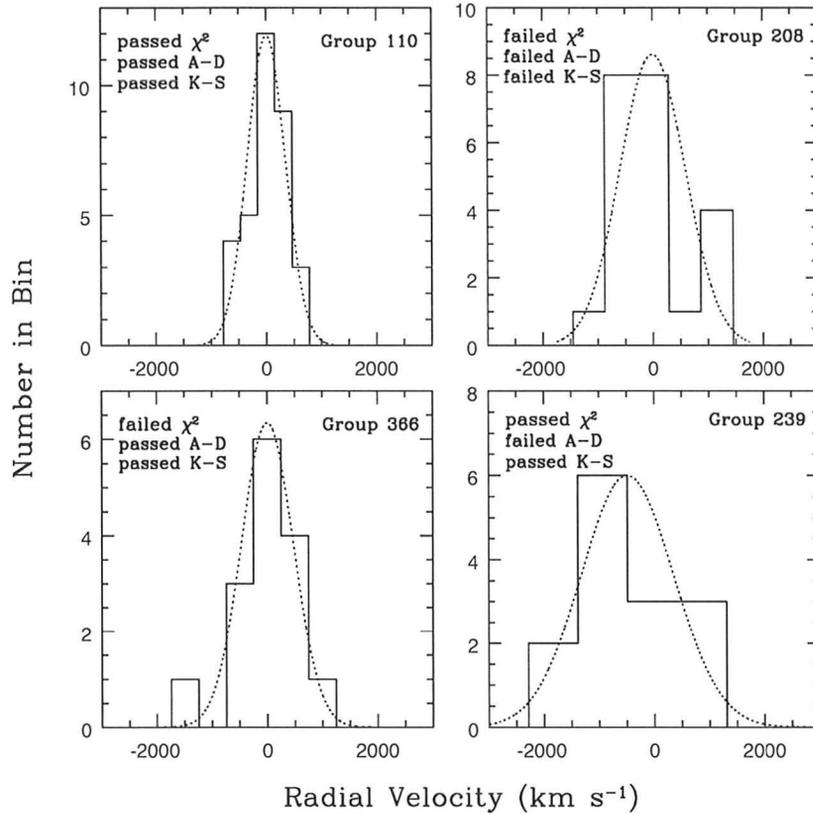


Figure 3.3: Velocity Distribution histograms for various CNOC2 groups. The dotted lines show Gaussian distributions computed using the estimated μ and σ of the group. It should be noted that the histograms are generated using Heald's optimal bin width (Equation 2.2), thus the bin widths vary from group to group. It should be noted that the Kolmogorov and A-D tests do not require binned data and are instead computed using ordered data. Also, note that the y-axes differ between plots.

Classification of the CNOC2 Groups

Having identified the Anderson–Darling test as the best statistical tool for the analysis of galaxy groups, we can now proceed to classify the CNOC2 groups as being either relaxed (Gaussian) or dynamically complex (non-Gaussian) systems. Before the test can be applied, we have to ensure that all galaxies identified by the FOF algorithm are truly members of the group. Once our group sample is defined, we can then compute the distribution parameters, apply the A–D test and finally classify our groups as either having an underlying Gaussian or non-Gaussian velocity distribution.

4.1 Maximum Radius Cut

In our analysis in §3, we apply only a minimum membership cut, but proper classification requires more detailed treatment of the group catalog. The relaxed linking length parameters used by W05 sometimes resulted in groups with large radii, with some members as far as 5.0 Mpc from the group center. This was done to ensure that FOF algorithm did not identify only virialized or relaxed systems, but also included more dynamically complex groups. Mamon (2007) suggests that the typical maximum r_{200} scale (i.e., the virial radius) for groups is 1.0 Mpc and the typical r_{200} for the Carlberg et al. (2001) CNOC2 groups is \ll 1.0 Mpc. We therefore apply a

1.0 Mpc radius cut to each of these re-identified CNOC2 groups to try and limit our sample to true groups rather than very loose associations of galaxies.

In Figure 4.1 we show both the position plots and velocity distributions of Group 208 and 226 to illustrate the effects of applying our maximum radius cut. The velocity distribution of Group 208 indicates that including all of the identified group members results in a group with a radius of ~ 3 Mpc, much higher than the suggested maximum radius of 1.0 Mpc. The position plot for Group 208 shows a compact grouping of 10 galaxies near the luminosity weighted group centroid with 12 galaxies scattered around the central clump. It is clear that once we apply our 1 Mpc radius cut on Group 208, indicated with a red circle centered at the group centroid in Figure 4.1, that the group now only includes the close arrangement of member galaxies at the group center.

Group 208 is a good example of how our maximum radius cut eliminates galaxies that are not truly members of the identified groups. Unfortunately, the distinction between ‘real’ group members and false identifications is not as clear for all the W05 identified CNOC2 groups, as seen with Group 226 in Figure 4.1. From the velocity distribution, we can see that inclusion of all 86 FOF identified group members results in a galaxy group over 4 Mpc in radius. The position plot for Group 226 does not show an obvious grouping of member galaxies, but rather several possible sub-groups. In spite of the possibility of substructure within Group 226, we apply the 1 Mpc radius cut around the group centroid, indicated by the red circle in Figure 4.1, and take the galaxies within this maximum radius as members of Group 226. This is done to ensure that we apply the same definition of membership to *all* CNOC2 groups.

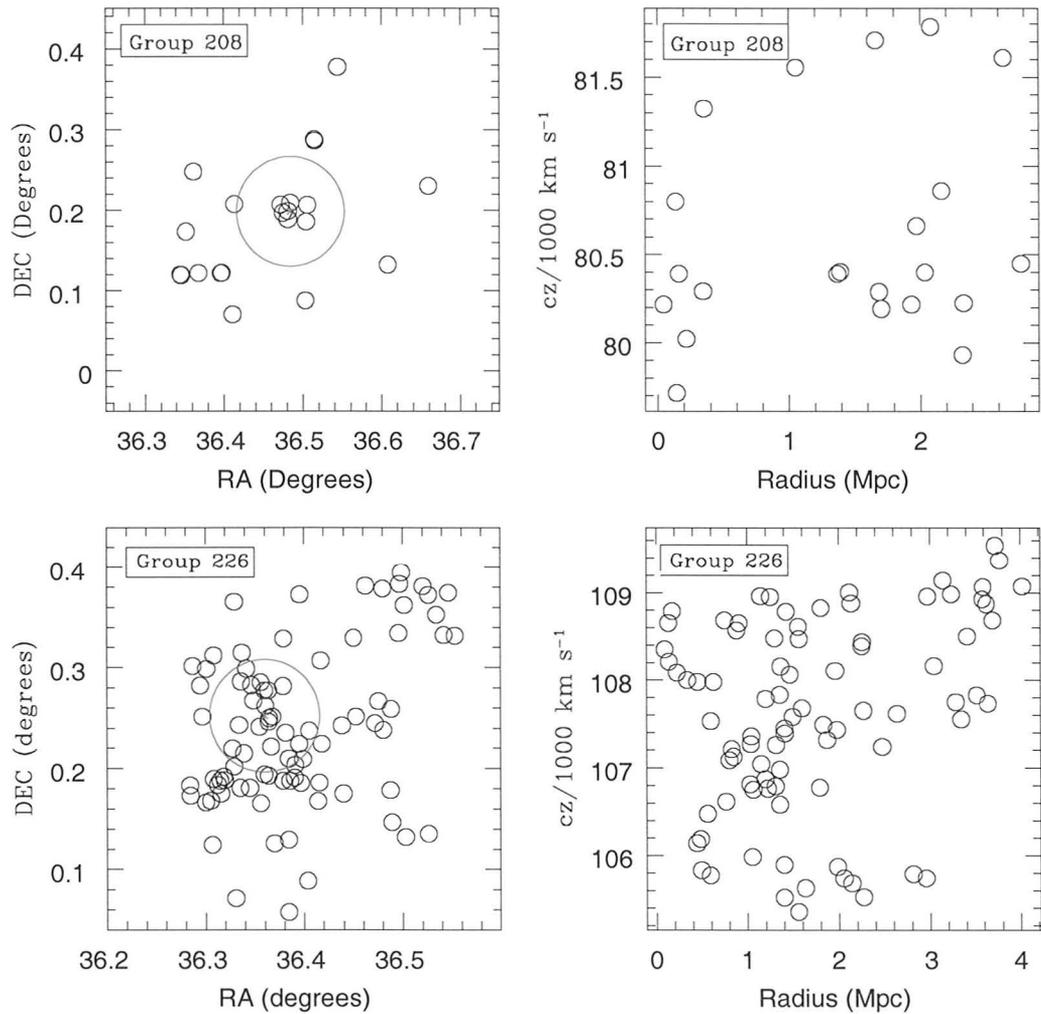


Figure 4.1: Left: Position plots of Groups 208 (top) and 226 (bottom), overplotted with a 1 Mpc radius circle around the group centroid. Right: Velocity distributions of Groups 208 (top) and 226 (bottom.)

Figure 4.2 shows a histogram of the number of group members both before (dotted red line) and after (solid black line) our 1 Mpc radius cut. It should be noted that Groups 138 ($n = 53$) and 226 ($n = 86$) are not included in the histogram. From the histograms we can see that after our radius cut the majority of groups, $\sim 76\%$, have 5 to 10 members, compared to only $\sim 55\%$ before our radius cut. Also, 16 of the original 122 groups, fall below the minimum group membership of $n = 5$ after our radius cut.

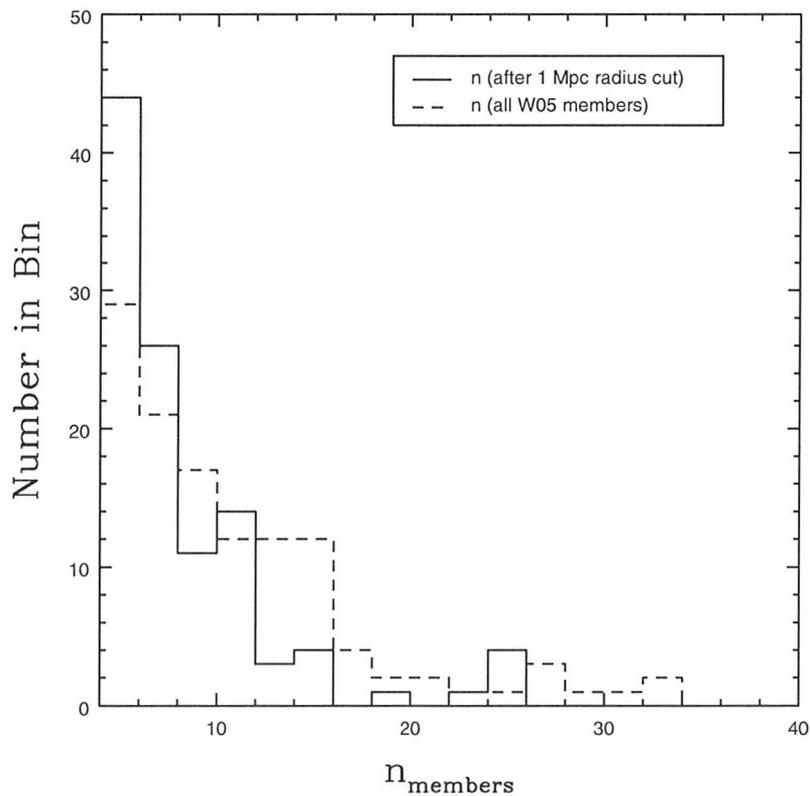


Figure 4.2: Histogram of the number of members per CNOC2 galaxy group in our sample, with the solid black line indicating the number of members after a 1 Mpc radius cut and considering only groups with $n \geq 5$ and the dashed black line indicating the number of members initially identified by W05 .

4.2 Computation of the μ and σ values for the CNOC2 Groups

For the distribution parameters, we estimate μ , the mean group velocity offset, using the standard mean, and σ , the group velocity dispersion, by computing the intrinsic velocity dispersion (W05), obtained by first computing the observed dispersion, given by;

$$\sigma_{obs} = 1.135 * \sigma_{Gapper} \tag{4.1}$$

where σ_{Gapper} is given by equation 3.1. The factor of 1.135 is a correction applied by W05 for their redshift space cut of $2\frac{\sigma_{obs}}{c}$. The next step is to compute the rest-frame velocity dispersion:

$$\sigma_{rest} = \frac{\sigma_{obs}}{1+z} \tag{4.2}$$

where z is the redshift of the group centroid. The final step involves removing the measurement uncertainty of each galaxy, $\langle\delta(v)\rangle$, from the rest-frame dispersion, as;

$$\sigma_{int}^2 = \sigma_{rest}^2 - \langle\delta(v)\rangle^2 \tag{4.3}$$

where $\langle\delta(v)\rangle = 142 \text{ km s}^{-1}$ for the LDSS2 data and $\langle\delta(v)\rangle = 103 \text{ km s}^{-1}$ for the original CNOC2 data (W05).

Using the method outline above we compute final intrinsic velocity dispersions (Tables 4.1, 4.2, 4.3 and 4.4) for all CNOC2 groups with $n \geq 5$ after our 1 Mpc radius cut.

Table 4.1: Computed velocity dispersions for the CNOC2 groups with $n \geq 5$ after a 1 Mpc radius cut in the 14 hr field.

Group	n	μ km s ⁻¹	σ_{obs} km s ⁻¹	σ_{rest} km s ⁻¹	σ_{int} km s ⁻¹
1	8	1.12	274	235	211
2	6	-51.5	492	411	398
4	8	121	446	371	357
6	8	-41.6	345	280	261
8	8	-12.4	326	265	238
9	8	1.12	277	219	194
11	12	0.500	269	211	185
13	11	135	552	434	421
14	5	0	398	310	292
16	5	-9.00	358	274	254
17	6	233	508	388	373
19	6	161	516	390	376
20	6	20.5	461	347	332
21	6	57.5	268	199	170
23	8	35.2	709	524	510
24	10	0.900	151	111	41.6
25	16	29.6	627	461	445
28	6	0	279	203	161
29	9	16.0	465	338	322
30	11	-1.09	428	307	289
32	8	-61.8	755	542	532
33	6	0	235	167	126
34	6	-1.50	254	173	134
37	10	-37.2	371	252	225
38	16	141	1.21×10^3	800	793
39	13	-42.2	733	477	462

Table 4.2: Computed velocity dispersions for the CNOC2 groups with $n \geq 5$ after a 1 Mpc radius cut in the 21 hr field.

Group	n	μ km s ⁻¹	σ_{obs} km s ⁻¹	σ_{rest} km s ⁻¹	σ_{int} km s ⁻¹
103	10	-0.300	179	156	118
104	19	14.5	471	411	398
107	9	-76.3	286	249	227
108	7	105	405	353	337
110	26	-15.3	422	365	350
111	9	51.6	452	384	370
113	7	37.7	614	512	501
114	5	-70.1	382	314	297
117	15	-29.2	328	269	248
120	6	-38.0	248	200	171
121	13	-0.922	348	280	260
122	8	-22.9	599	481	470
123	12	-5.25	292	231	207
124	9	-52.6	349	271	251
126	5	0.600	230	177	144
127	6	1.00	180	137	90.7
128	5	130	586	445	433
129	5	-118	371	282	262
131	5	101	593	437	424
132	8	25.5	542	399	382
134	10	-23.4	471	338	317
135	7	1.28	379	271	251
137	8	-0.749	477	334	315
138	23	-226	1.06×10^3	740	731
139	10	0.899	363	252	226
140	5	0	219	149	100

Table 4.3: Computed velocity dispersions for the CNOC2 groups with $n \geq 5$ after a 1 Mpc radius cut in the 2 hr field.

Group	n	μ km s ⁻¹	σ_{obs} km s ⁻¹	σ_{rest} km s ⁻¹	σ_{int} km s ⁻¹
201	5	29.4	353	312	290
202	5	63.6	423	356	341
204	6	0	187	153	113
206	11	90.8	344	280	261
208	8	-70.5	762	601	592
211	8	-0.375	234	184	153
212	11	1.36	333	256	235
213	7	58.7	581	446	434
216	6	-48.0	471	361	346
217	5	-661	920	704	696
218	5	252	443	338	322
220	5	-0.600	353	262	241
221	6	0.500	171	126	73.1
226	25	-174	1.16×10^3	853	847
227	7	-1.25	491	360	342
228	8	0	475	343	327
230	10	0	212	153	113
231	5	82.1	438	316	299
232	11	-76.3	443	317	300
233	7	-61.7	795	568	559
234	7	0.857	422	302	284
237	7	534	720	515	505
238	11	-219	899	638	630
239	6	-169	869	617	608
240	6	111	536	379	365
241	6	-0.999	316	222	197
244	15	-0.999	356	242	211

Table 4.4: Computed velocity dispersions for the CNOC2 groups with $n \geq 5$ after a 1 Mpc radius cut in the 9 hr field.

Group	n	μ km s ⁻¹	σ_{obs} km s ⁻¹	σ_{rest} km s ⁻¹	σ_{int} km s ⁻¹
307	6	220	643	532	522
308	25	2.52	639	522	512
312	8	-217	424	344	328
315	12	11.7	360	290	271
317	11	-30.5	382	307	289
320	14	-50.3	575	462	450
322	6	-135	545	437	425
323	7	-88.2	610	490	479
324	11	112	418	336	319
333	6	-212	351	265	245
334	11	128	590	446	434
336	5	-62.4	965	708	700
337	6	-232	797	581	571
338	9	-13.7	503	367	352
344	7	8.99	368	268	248
346	26	-80.4	613	446	434
349	6	-110	766	555	545
350	5	226	386	280	260
351	6	0.500	243	176	42.6
352	5	118	581	422	409
355	5	-161	506	364	349
356	5	-153	540	388	374
357	7	-0.428	234	168	133
358	11	-19.6	409	294	275
360	7	295	878	631	622
363	5	-119	620	424	411
366	11	71.1	468	317	300

We compare our intrinsic velocity dispersion values with those computed by W05. It should be noted that although we follow the general method outlined by W05, there are several differences in our computations of σ_{int} . W05 defines a maximum radius of 0.5 Mpc, as opposed to our 1.0 Mpc cut. They also compute σ_{obs} using an iterative process, making several smaller radius cuts (1.0, 0.75 and 0.5 Mpc), redefining the group center and recalculating σ_{Gapper} after each cut. Once W05 obtains a stable σ_{obs} , they calculate σ_{rest} and σ_{int} using equations 4.2 and 4.3. Also, we do not compute dispersion values for groups with $n < 5$ after our radius cut and therefore have fewer groups with computed dispersion values than W05. Despite these computational differences, we find that for the majority of groups, our intrinsic velocity dispersions agree with those of W05. We plot both values of σ_{int} in Figures 4.3, 4.4, 4.5 and 4.6, where our dispersion values are plotted in red crosses and W05 values are plotted in open circles with errors, computed using a jackknife method. This method of computing errors involves re-calculating the dispersion on a subset of the data (i.e., $n - 1$ data points) and is computed n times, eliminating a different data point for each dispersion. The jackknife error is then obtained by determining the maximum difference between the dispersions of the subsets and that of the entire data set.

In Figure 4.3 shows the dispersion values for the CNOC2 14 hour field, where we see that for 13 of the 26 groups, our values are exactly the same as those of W05. For the remainder 18 groups, our values agree within error to W05 σ_{int} values. We plot both intrinsic dispersion values for the CNOC2 21 hour field in Figure 4.4, and find that 17 of the 26 CNOC2 groups have the same σ_{int} , with all but one of the remaining values falling within the W05 errorbars. The CNOC2 2 hour field (Figure 4.5) has a high percentage of matches between ours and W05 σ_{int} values, with 25 out of 26 groups having the same intrinsic dispersion. Similarly, the CNOC2 9 hour field has a very high fraction of corresponding intrinsic dispersion values (Figure 4.6), with 26 of the 27 σ_{int} values in complete agreement. In Figure 4.7, we plot a histogram of the difference between our intrinsic velocity dispersions and those computed by

W05, where we see a sharp peak in the histogram at $\sigma_{\text{int}} - \sigma_{\text{int Wilman}} \approx 0 \text{ km s}^{-1}$, further indicating that the two values are generally in good agreement.

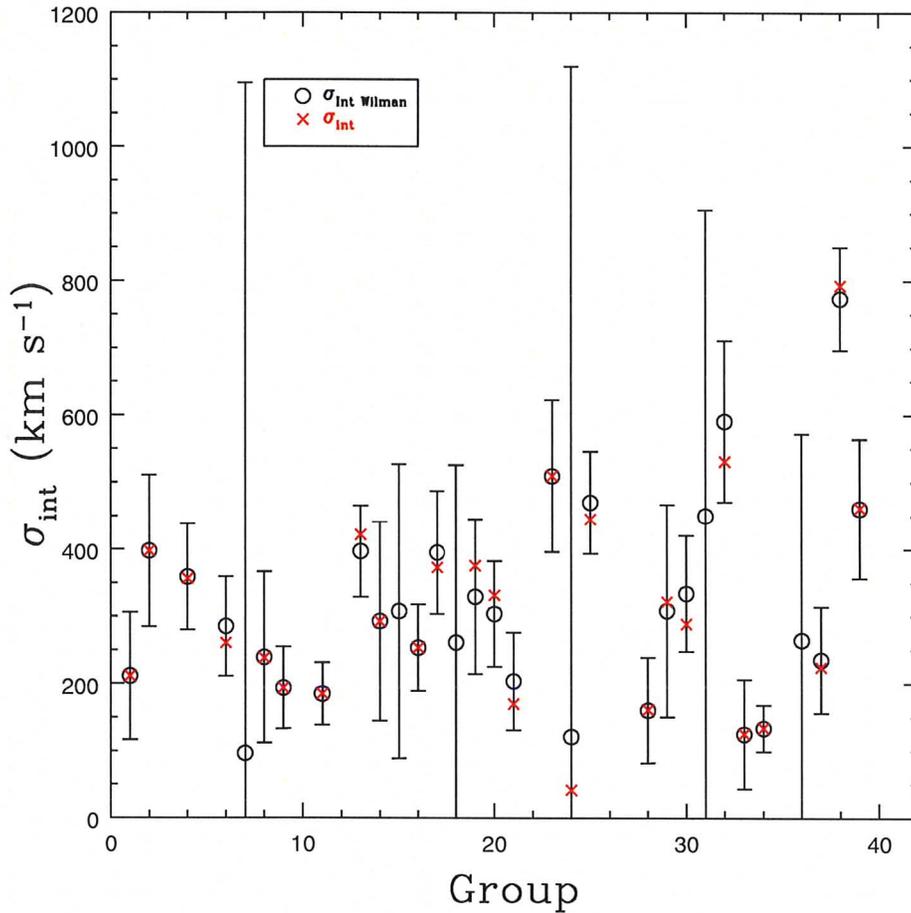


Figure 4.3: Comparison of our computed intrinsic velocity dispersion with those of Wilman et al. (2005b) in the 14 hr field. Our dispersions are plotted in red crosses and Wilman et al. (2005b) dispersions are plotted in open circles with error bars computed using a jackknife method.

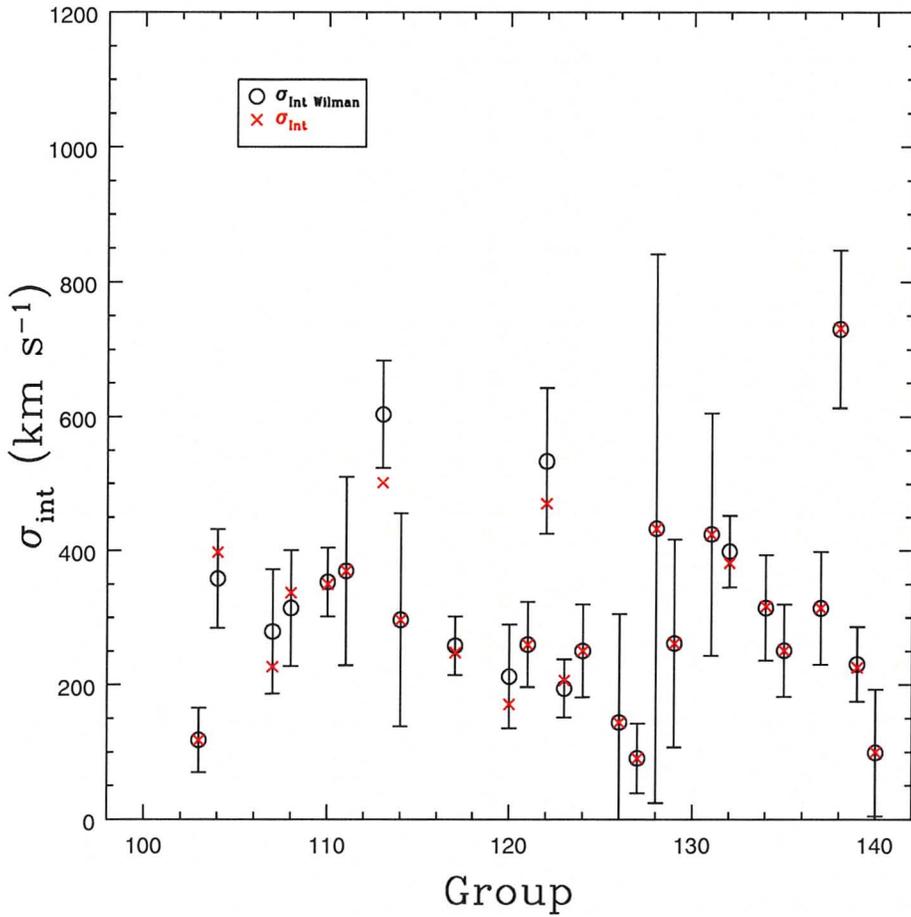


Figure 4.4: Comparison of our computed intrinsic velocity dispersion with those of Wilman et al. (2005b) in the 21 hr field. Our dispersions are plotted in red crosses and Wilman et al. (2005b) dispersions are plotted in open circles with error bars computed using a jackknife method.

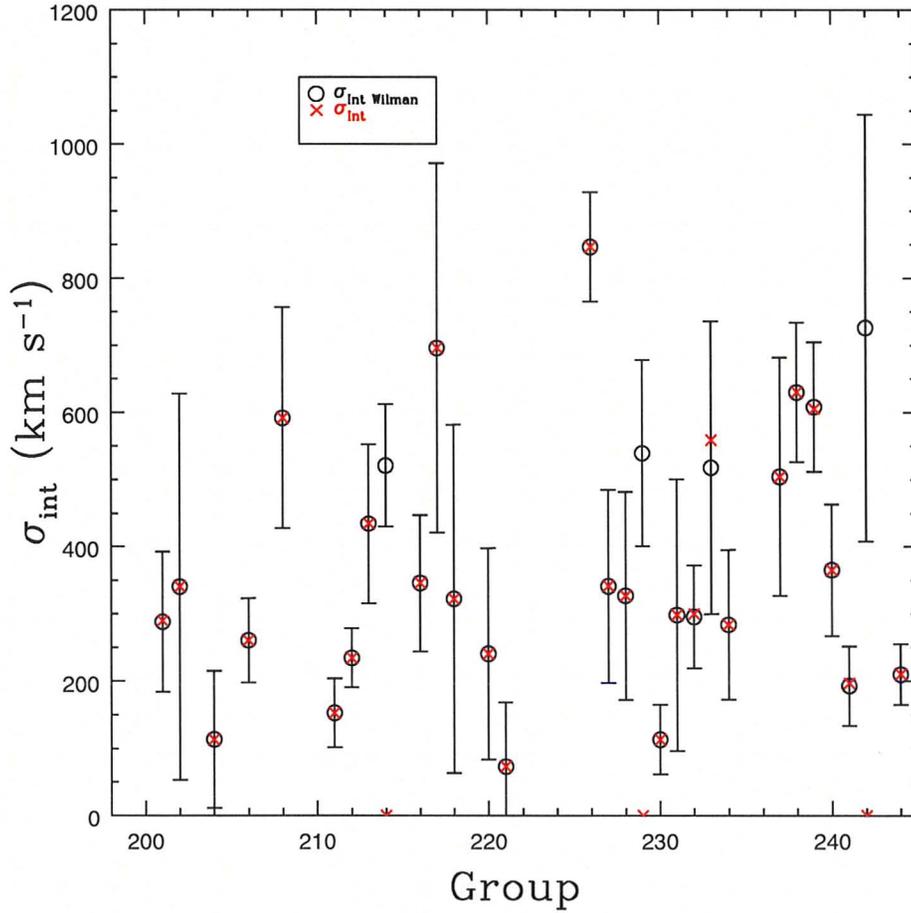


Figure 4.5: Comparison of our computed intrinsic velocity dispersion with those of Wilman et al. (2005b) in the 2 hr field. Our dispersions are plotted in red crosses and Wilman et al. (2005b) dispersions are plotted in open circles with error bars computed using a jackknife method.

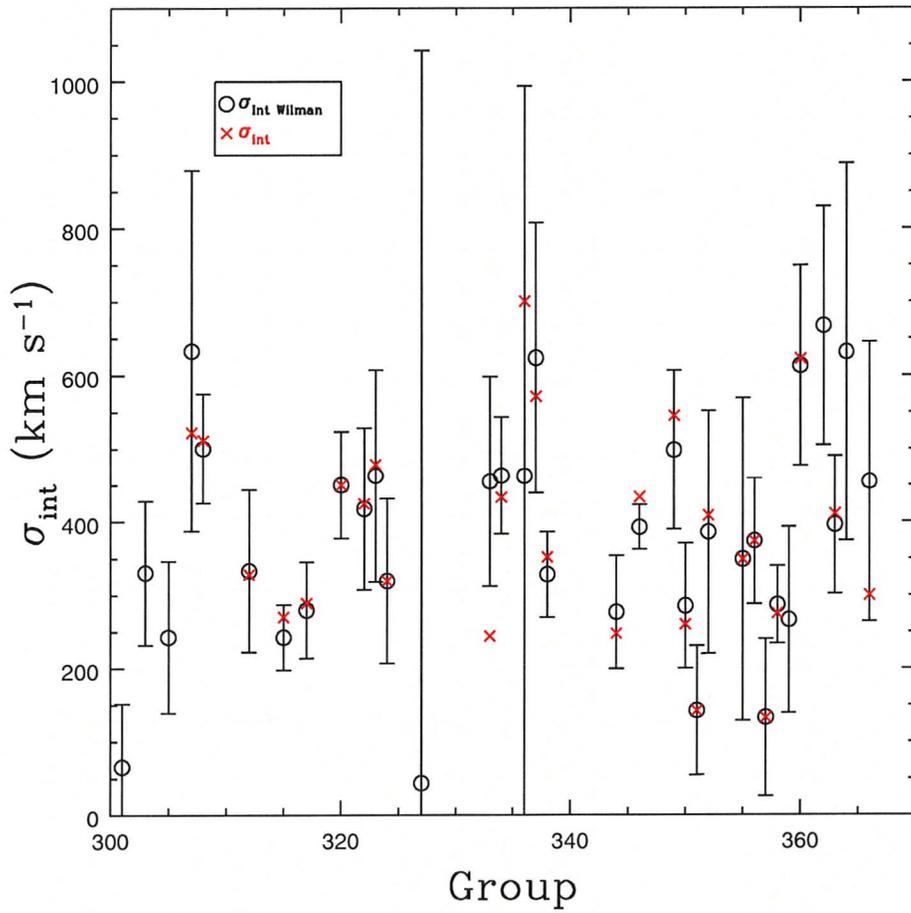


Figure 4.6: Comparison of our computed intrinsic velocity dispersion with those of Wilman et al. (2005b) in the 9 hr field. Our dispersions are plotted in red crosses and Wilman et al. (2005b) dispersions are plotted in open circles with error bars computed using a jackknife method.

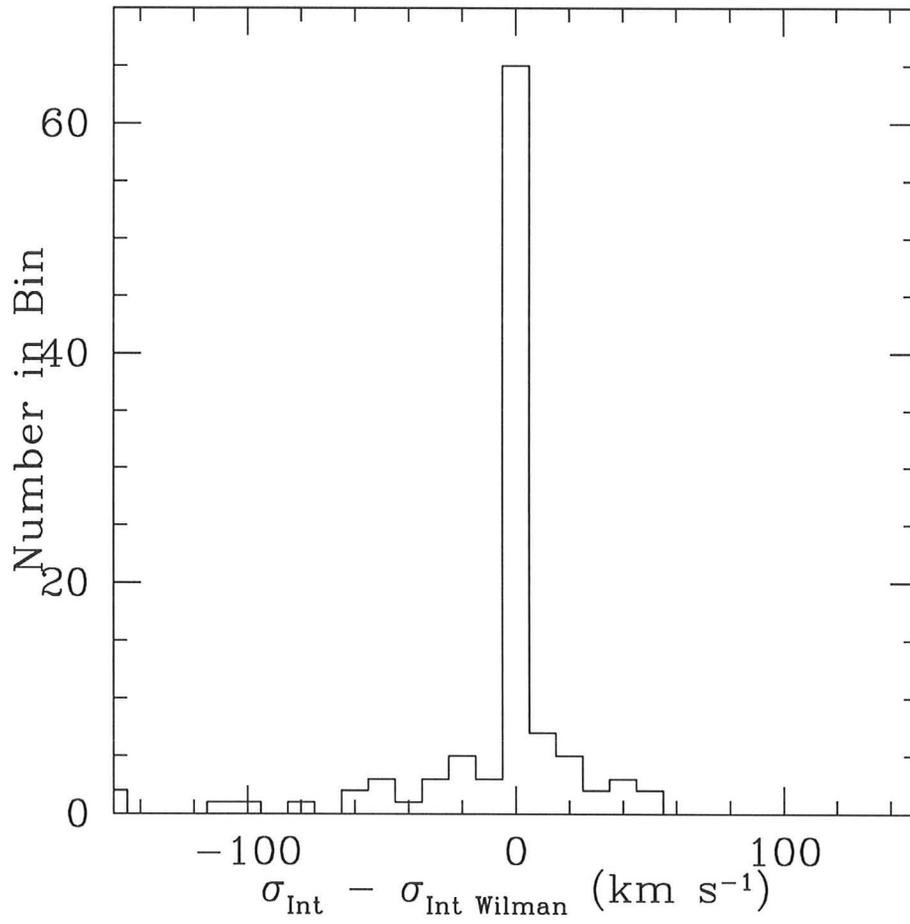


Figure 4.7: Histogram of $\sigma_{\text{int}} - \sigma_{\text{int Wilman}}$ for all 4 fields.

4.3 Application of the Anderson–Darling Test

We compute the intrinsic velocity dispersion for all W05 re-identified CNOC2 groups with $n_{\text{members}} \geq 5$, then classify the groups based on the computed A^{2*} values. A crucial step in any statistical analysis is to identify the appropriate significance level or critical value used for classification. The critical points of the A–D test, as with all EDF statistics, change depending on the accuracy with which one knows the input distribution parameters (DA86). For a Gaussian distribution, there are four cases to be considered:

- case 0: both μ and σ^2 are known a priori, i.e., a fully specified distribution;
- case 1: σ^2 is known and μ must be estimated;
- case 2: μ is known and σ^2 must be estimated, and;
- case 3: both μ and σ^2 must be estimated

Each of these cases result in different critical values, which can greatly alter the number of rejections. For example, the 5% critical value is 2.492 for Case 0, 1.087 for Case 1, 2.308 for Case 2 and 0.752 for Case 3 (DA86). When testing for Gaussianity in data sets, Stephens (1974) suggests that Case 3 is the most practical choice, as distribution parameters are in general estimated and not known *a priori*. In this situation, there are two approaches one can take in distinguishing between Gaussian and non-Gaussian systems; the first involves comparing the computed A^{2*} values with critical value tables and the other uses the A^{2*} values to directly compute the significance level α , which gives the probability that the system has an underlying Gaussian distribution. We chose to follow the latter method and compute α using the formula;

$$\alpha = a \exp(-A^{2*}/b) \tag{4.4}$$

where $a = 3.6789468$ and $b = 0.1749916$, and both factors are determined via Monte Carlo methods (Nelson, 1998).

Using Equation 4.4, we were able to determine the probability of whether or not each CNOC2 group had a Gaussian velocity distribution, classifying all groups with $\alpha < 5\%$ as being non-Gaussian. The results of the A-D test are given in Table 4.5, where we see that $\sim 32\%$ of the CNOC2 groups are classified as non-Gaussian at the $\alpha = 0.05$ significance level and also that the \bar{n} , \bar{z} and $\bar{\sigma}$ are similar for the Gaussian/non-Gaussian groups. Using this classification scheme, we can now investigate specific group properties to determine if there are any obvious trends or differences between the Gaussian and non-Gaussian groups. The properties of the CNOC2 groups that are classified as dynamically complex (non-Gaussian) are given in Table 4.8.

Table 4.5: Anderson-Darling Classification of the CNOC2 groups with $n \geq 5$ after the 1 Mpc cut.

Classification	Number of Groups	Percentage of All Groups	\bar{n}	\bar{z}	$\bar{\sigma}_{\text{int}}$ (km s ⁻¹)
Gaussian	72	$\sim 68\%$	9	0.30	347
non-Gaussian	34	$\sim 32\%$	9	0.37	327

In Tables 4.6 and 4.7 we list several properties of the groups classified as Gaussian by the A-D test, where Column 1 indicates the Group, Column 2 gives n after our 1 Mpc radius cut, Column 3 gives σ_{int} , Column 4 gives the A^{2*} values and Column 5 gives the values of α (Equation 4.4). Looking at the number of member galaxies, it is evident that the Gaussian groups are not restricted to either low or high membership groups, with $5 \leq n \leq 26$. Similarly, the intrinsic velocity dispersion of the groups show no preference for any given dispersion range, with the values ranging from ~ 110 -700 km s⁻¹.

Table 4.6: Properties of Gaussian CNOC2 Groups in the 14 and 21 hour Fields.

Group	n	σ_{int} (km s ⁻¹)	A^{2*}	α
1	8	211	0.293	0.691
2	6	398	0.240	0.933
4	8	357	0.287	0.717
6	8	261	0.392	0.391
8	8	238	0.437	0.303
9	8	194	0.270	0.785
11	12	185	0.379	0.423
13	11	422	0.451	0.279
14	5	293	0.283	0.729
16	5	254	0.607	0.115
17	6	373	0.720	0.060
19	6	376	0.472	0.248
20	6	332	0.396	0.383
21	6	170	0.390	0.395
23	8	510	0.453	0.277
25	16	445	0.608	0.114
28	6	161	0.646	0.092
37	10	225	0.717	0.06
39	13	462	0.685	0.073
103	10	118	0.623	0.105
104	19	398	0.277	0.754
107	9	227	0.271	0.781
108	7	337	0.251	0.876
110	26	350	0.200	1.17
111	9	370	0.353	0.489
113	7	501	0.440	0.298
114	5	297	0.243	0.916
117	15	248	0.682	0.075
121	13	260	0.459	0.267
122	8	470	0.210	1.11
123	12	207	0.656	0.086
124	9	251	0.339	0.529
126	5	144	0.742	0.0531
131	5	424	0.314	0.613
134	10	317	0.487	0.228
137	8	315	0.407	0.360

Table 4.7: Properties of Gaussian CNOC2 Groups in the 2 and 9 hour Fields.

Group	n	σ_{int} (km s ⁻¹)	A^{2*}	α
201	5	290	0.204	1.15
204	6	113	0.498	0.214
206	11	261	0.360	0.469
208	8	592	0.385	0.407
212	11	235	0.747	0.0514
216	6	346	0.595	0.123
217	5	696	0.353	0.489
220	5	241	0.455	0.274
227	7	342	0.255	0.859
228	8	327	0.608	0.114
231	5	299	0.691	0.071
232	11	300	0.335	0.543
234	7	284	0.297	0.674
237	7	505	0.651	0.089
238	11	630	0.629	0.101
239	6	608	0.733	0.056
240	6	365	0.649	0.090
307	6	522	0.389	0.399
308	25	512	0.590	0.126
315	12	271	0.476	0.242
317	11	289	0.370	0.443
320	14	450	0.360	0.470
322	6	425	0.240	0.931
323	7	479	0.358	0.476
324	11	319	0.406	0.361
333	6	245	0.625	0.103
334	11	434	0.432	0.312
337	6	571	0.208	1.12
344	7	248	0.530	0.178
349	6	545	0.428	0.319
350	5	260	0.656	0.086
352	5	409	0.430	0.315
355	5	349	0.541	0.167
356	5	374	0.450	0.282
363	5	411	0.604	0.117
366	11	300	0.742	0.053

The properties of the groups classified as non-Gaussian by the A–D test are given in Table 4.8. Like the Gaussian groups, the intrinsic velocity dispersions of the non-Gaussian groups show no obvious trend, as the values range from $\sim 40 - 800 \text{ km s}^{-1}$, so the more dynamically complex systems are not restricted to low or high velocity dispersions. Although, it should be noted that the dispersion range for the non-Gaussian groups is slightly larger than that of the Gaussian groups (Tables 4.6 and 4.7). Similarly, there is a wide range in the number of members for failed groups, having anywhere between 5 and 26 members, suggesting that the A–D test is not biased towards a specific sample size.

Table 4.8: Properties of non-Gaussian CNOC2 Groups in all 4 Fields.

Group	n	σ_{int} (km s ⁻¹)	A^{2*}	α
24	10	42	13.4	2.69×10^{33}
29	9	322	0.780	0.0426
30	11	289	1.18	0.00424
32	8	532	0.945	0.0166
33	6	126	1.61	0.000375
34	6	134	1.67	0.000271
38	16	793	1.66	0.00940
120	6	171	0.752	0.0499
127	6	91	1.18	0.00427
128	5	433	1.22	0.00342
129	5	262	0.813	0.0353
132	8	382	1.31	0.00202
135	7	251	1.03	0.0105
138	23	731	0.973	0.0142
139	10	226	1.39	0.00128
140	5	100	1.18	0.00432
202	5	341	0.836	0.0309
211	8	153	0.892	0.0225
213	7	434	0.763	0.0469
218	5	322	0.880	0.241
221	6	73.1	2.37	4.86×10^6
226	25	847	1.04	0.00940
230	10	113	1.25	0.00298
233	7	559	1.57	0.000463
241	6	197	0.757	0.0487
244	15	211	1.35	0.00168
312	8	328	1.14	0.00545
336	5	700	0.947	0.0164
338	9	352	0.822	0.0336
346	26	434	1.26	0.00277
351	6	143	0.814	0.0352
357	7	133	1.80	0.000128
358	11	275	0.776	0.0436
360	7	622	0.831	0.0319

Comparison of Group Properties

Although the majority of our group sample have $n \leq 10$, we do have 5 CNOC2 groups with $n \geq 20$ after our radius cut. As previously mentioned, there is no clear distinction between rich groups and poor clusters, thus one might expect groups with a large number of members and underlying Gaussian velocity distributions to exhibit trends similar to those observed in galaxy clusters. We therefore look at our 5 rich groups to determine if any trends exist or if there are any differences between groups classified as Gaussian or non-Gaussian by the Anderson–Darling test.

5.1 Individual Group Properties

5.1.1 Velocity Dispersion Profiles

One way to investigate the differences between relaxed and dynamically complex systems is to study the velocity dispersion profiles (hereafter VDP) of galaxy groups. If a classified non-Gaussian group is in fact dynamically more complex than a Gaussian one, then the corresponding VDP may exhibit different features from those of relaxed systems. In their analysis of merging clusters, Menci & Fusco-Femiano (1996) find that radially increasing VDPs indicate significant galaxy merging in the cluster core. Girardi et al. (1996) also suggest that the presence of neighbouring clusters resulted

in a VDP with an initially flat profile, which then increased strongly at larger radii.

We follow the method outlined in Bergond et al. (2006) to generate VDPs for the CNOC2 groups in which the radial velocities are binned with an exponentially weighted moving window. The window function is given by:

$$w_i(R) = \frac{1}{\sigma_R} \exp \left[-\frac{(R - R_i)^2}{2\sigma_R^2} \right] \quad (5.1)$$

where σ_R is the width of the window, which can be constant or a function of radius R , and the R_i 's are the radial positions of the members of the system. The projected velocity dispersions are then defined as;

$$\sigma_p(R) = \sqrt{\frac{\sum_i w_i(R)(x_i - \bar{x})^2}{\sum_i w_i(R)}} \quad (5.2)$$

where the x_i 's are the radial velocities and \bar{x} is the mean velocity of the system.

This ‘‘moving window’’ prescription for computing VDPs takes into account the contribution of every radial velocity measurement at each value of R . It also removes the restriction of computing binned projected velocity dispersion. Instead, a smoothed profile can be generated, since the projected dispersions can be computed at any radius, not just at the radii corresponding to the observed velocities.

In order to use this method to probe the dynamics of a system, one must be careful with the choice of window width, σ_R , as a too-large window can wash out real features, and a too-small window, or very small n , tends to add spurious features in the profile. Since the projected velocity dispersions are computed using weighted values of every velocity measurement in the data, any large individual deviations from the mean can alter the overall shape of the profile. This effect is significantly more pronounced in small samples, as there are not enough data points to counteract or outweigh the effects of an outlier. With testing we choose a window width of 0.35 Mpc, approximately one-third the value of the maximum radius. We also enforce a minimum group membership of 20 members, after our 1 Mpc radius cut, to ensure

that any visible trend in the VDPs is not a result of outliers.

There are five CNOC2 groups that meet our minimum group membership criteria, two of which are classified as Gaussian, Groups 110 ($n = 26$) and 308 ($n = 25$) and three as non-Gaussian systems, Groups 138 ($n = 23$), 226 ($n = 25$) and 346 ($n = 26$). The VDPs for these groups are shown in Figure 5.1, where the Gaussian groups are shown with filled symbols and the non-Gaussian groups with open symbols. Both Gaussian groups have decreasing profiles, while two of the three non-Gaussian groups (138 and 346) have increasing profiles. The VDP for Group 226 (Figure 5.1) does not exhibit the same overall trend as the other non-Gaussian groups, as the profile increases initially but turns over at roughly 0.4 Mpc. Although the profile for Group 226 does not continually increase, it does show distinct features from the profiles of the Gaussian groups. It is impossible to make general statements on the overall shape of all Gaussian or non-Gaussian groups based on these five groups alone, but the results do support our claim that galaxy groups classified as non-Gaussian by the A–D test are more dynamically complex systems. In their investigation of the relation between velocity dispersion and X-ray temperatures in galaxy clusters, Girardi et al. (1996) found that clusters with rising profiles, such as Abel 3391 and 3395, had a neighbouring cluster or group, resulting in velocity anisotropies and sharp increases in their VDPs. Although we are unable to determine by the VDPs alone whether these non-Gaussian groups are indeed undergoing a merger, the clear differences between Gaussian and non-Gaussian group profiles suggest that these two types of systems are dynamically different.

A closer inspection of the individual profiles of the Gaussian groups suggests another interesting result. The profile for Group 110 shows an initial decrease with an eventual flattening of the profile towards the outer radius, a general trend that was observed in clusters by Girardi et al. (1996). The VDP of 308 does not exhibit the same trend of a flattened profile, continually decreasing towards the edge of the group.

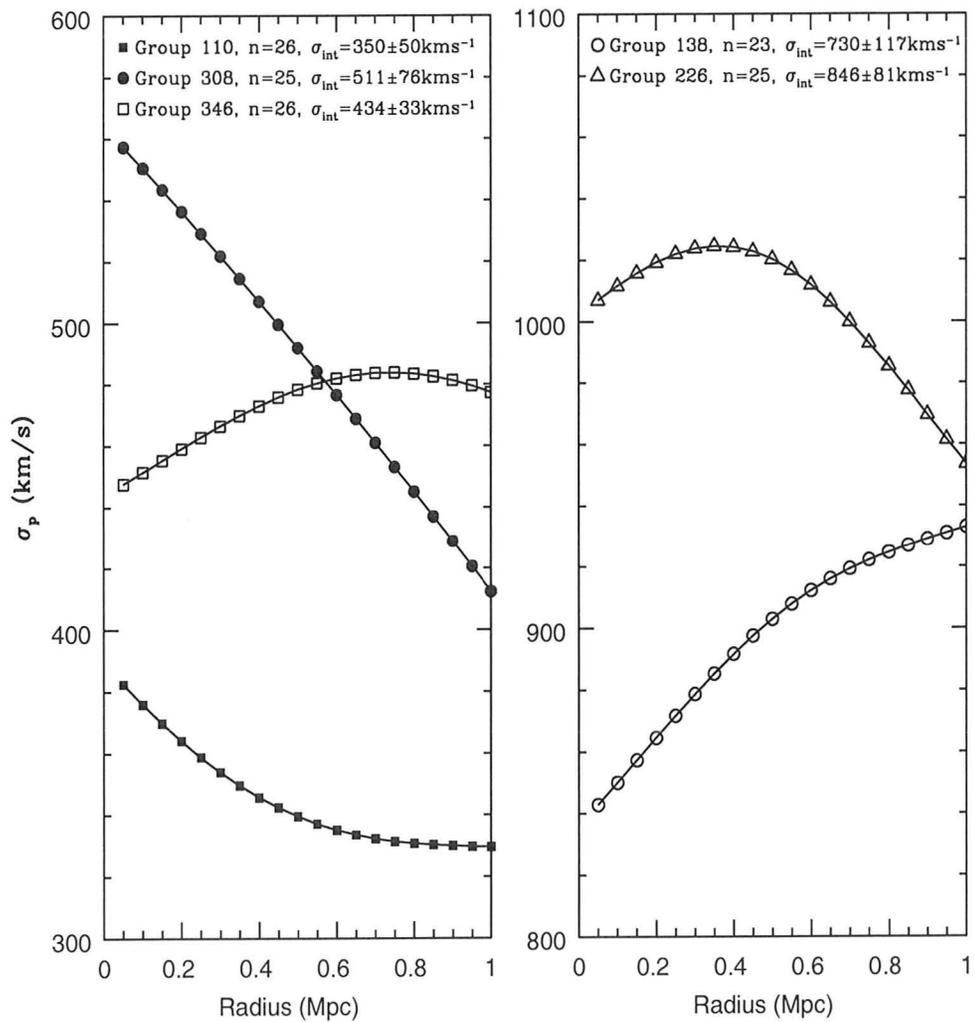


Figure 5.1: Velocity dispersion profiles for the CNOC2 groups with $n > 20$ after a 1Mpc radius cut and using a constant window width of 0.35 Mpc. The open symbols indicate groups classified as non-Gaussian and the closed symbols indicate those classified as Gaussian.

5.1.2 Colour-Radius

Along with the $UBVR_CI_C$ photometry there is also a set of CFHT $ugriz$ photometry for the CNOC2 groups, which have been k -corrected to obtain rest-frame colours. Since colour is an indicator of star formation history, we look at colour versus radius plots for the five CNOC2 groups, with $n \geq 20$ after our 1 Mpc radius cut, to investigate possible differences between the classified Gaussian and non-Gaussian groups. We also aim to determine if the trend of an increasing fraction of blue galaxies with radius observed in galaxy clusters (Butcher & Oemler, 1984), is also present in the group environment.

In Figure 5.2, we plot $g - r$ versus radius for Groups 110 and 308, which are classified as having underlying Gaussian velocity distributions by the A-D test. Since Groups 110 and 308 are our two largest Gaussian groups, we might expect them to exhibit trends similar to those observed in galaxy clusters. The colour-clustercentric radius relation observed in galaxy clusters out to $z \sim 0.3$ (Goto et al., 2004) is seen in the colour versus radius plots of Groups 110 and 308 (Figure 5.2), which shows a clear increase of blue galaxies with radius. In Group 110, the $\overline{g - r}$ value for the galaxies within 0.4 Mpc of the group center is 0.728 ± 0.033 , while $\overline{g - r} = 0.549 \pm 0.044$ for galaxies with radii beyond 0.4 Mpc. Group 308 shows a similar decreasing trend with $\overline{g - r} = 0.558 \pm 0.054$ for galaxies within 0.4 Mpc and $\overline{g - r} = 0.504 \pm 0.054$ for galaxies outside this radius.

The $g - r$ versus radius plots for our three classified non-Gaussian groups (138, 226 and 346) are shown in Figure 5.3. Unlike the colour-radius plots for the Gaussian groups (Figure 5.2), the colour-radius relation is not observed in any of these three groups. In fact, each of the plots for Groups 138, 226 and 346 (Figure 5.3) show some interesting and unexpected features. Group 138 has a rising VDP (Figure 5.1), a possible indication of a minor or major interaction and one might expect this group to have a high fraction of blue galaxies, especially at the edge of the group where tidal effects would be the strongest. From Figure 5.3, we see that Group 138 has only

1 member galaxy with $g - r < 0.5$, which lies relatively close to the group centroid at 0.3 Mpc. In contrast to the Gaussian groups, the mean colour of the galaxies in Group 138 appear to remain flat with radius, with $\overline{g - r} = 0.593 \pm 0.047$ for members within 0.4 Mpc of the group centroid and $\overline{g - r} = 0.612 \pm 0.012$ for galaxies between 0.4 - 1.0 Mpc of the center. Group 226 also shows no significant increase of blue galaxies with radius (Figure 5.3), with the $\overline{g - r} = 0.569 \pm 0.065$ for galaxies within 0.4 Mpc of the group center and $\overline{g - r} = 0.510 \pm 0.035$ for galaxies beyond this radius. Similarly to Groups 138 and 226, the colour-radius relation is not evident in Group 346 (Figure 5.3), where the mean values are $\overline{g - r} = 0.623 \pm 0.057$ for galaxies within 0.4 Mpc of the group centroid, and $\overline{g - r} = 0.630 \pm 0.033$ for galaxies between 0.4 - 1.0 of the center.

Based on these results, it would appear that Gaussian groups obviously follow the colour-radius relation seen in clusters, with visible negative gradients in their $g - r$ versus radius plots. The non-Gaussian groups show no visible correlation between colour and radius, and the mean $g - r$ values remain relatively flat for galaxies within 0.4 Mpc of the group center and those between 0.4 - 1.0 Mpc of the center.

We also look at the fraction of blue galaxies (hereafter f_b) of each group, defined as:

$$f_b = \frac{n(g - r < 0.5)}{n(\text{all members})} \quad (5.3)$$

We choose a $g - r$ value of 0.5 to distinguish between the red sequence and blue cloud based on the $g - r$ histograms for the stacked CNOC2 groups (Figures 5.10 and 5.12) discussed in detail in Chapter 5.2.2.

The f_b values differ between the two Gaussian groups, with Group 110 having 30% blue galaxies and Group 308 having 44% blue galaxies. The range in f_b values is even more significant when we look at the non-Gaussian groups. Group 138 has a very low fraction of blue galaxies with only 4%, Group 226 has the highest f_b value of 44% and Group 346 has an f_b value of 27%. Contrary to what was expected, we do

not find any trends or differences in the fraction of blue galaxies in the five, $n \geq 20$, CNOC2 groups. This result is quite interesting, as the VDPs of the non-Gaussian groups show a rising trend with radius (Figure 5.1), a possible indicator of galaxy interactions or merger activity, which is expected to trigger star formation and thus increase the blue fraction.

We emphasize that although we do not observe correlations in these five groups, we do not exclude the possibility that in general a trend may exist in the f_b values of Gaussian and non-Gaussian groups (see Chapter 5.2.2 and 5.2.3 for discussion on the colour-magnitude and -radius plots of stacked groups).

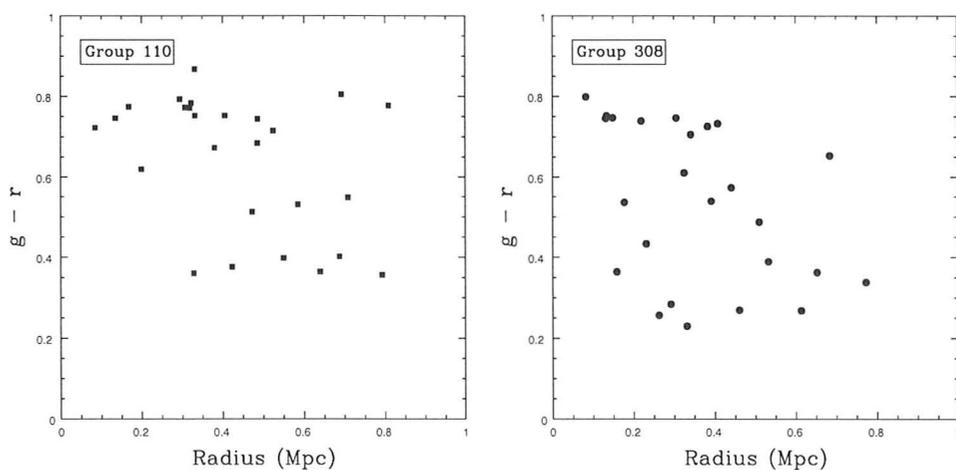


Figure 5.2: Left: Colour-radius plot for Group 110. Right: Colour-radius plot for Group 308. The closed symbols indicate that these groups have been classified as Gaussian by the A–D test.

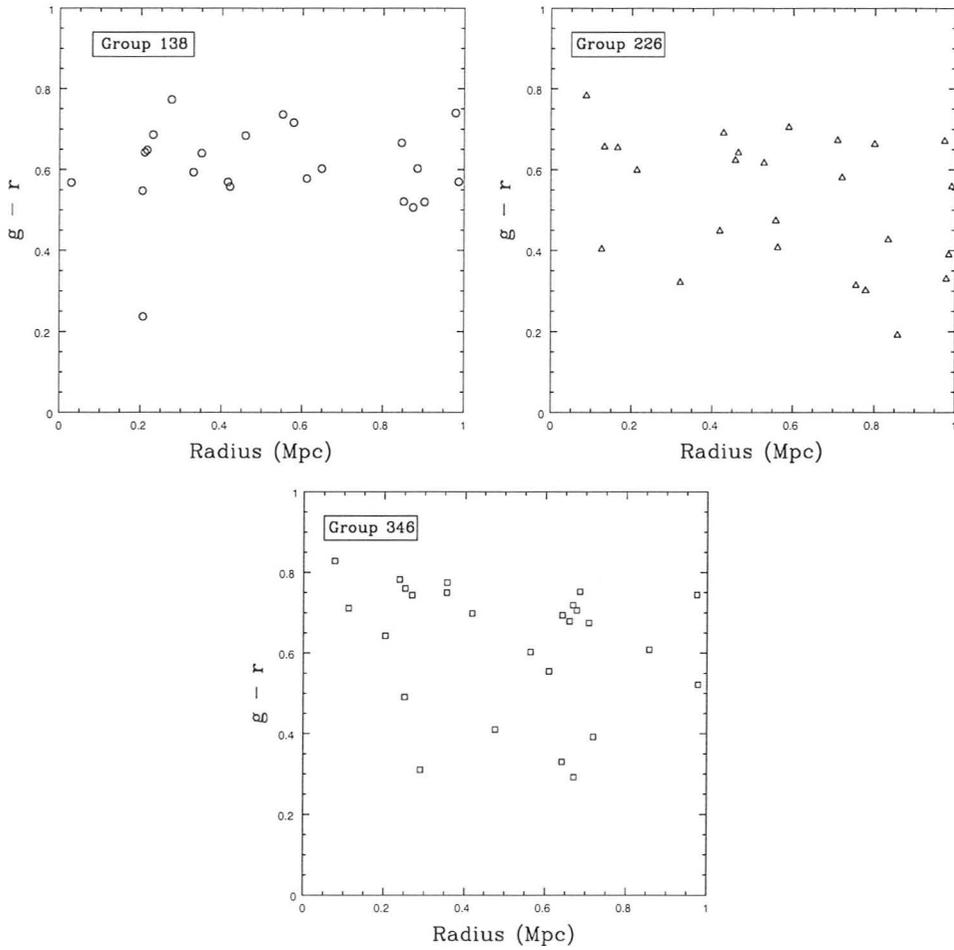


Figure 5.3: Top-left: Colour-radius plot for Group 138. Top-right: Colour-radius plot for Group 226. Bottom: Colour-radius plot for Group 346. The open symbols indicate that these groups have failed the A–D test and are classified as non-Gaussian.

5.1.3 Stellar Mass

Balogh et al. (2007) obtained near-infrared (NIR) imaging for 58 CNOC2 groups, from the Isaac Newton Group Red Imaging Device (INGRID) on the William Herschel Telescope and from the Infrared Array Camera (IRAC) on the Spitzer Space Telescope. Of the 58 observed groups, 29(38) of the groups have good INGRID(IRAC) coverage, which are listed in Table 1 of Balogh et al. (2007). Using either combined NIR and optical or optical only imaging, estimates of the stellar mass per galaxy in the CNOC2 groups is obtained via spectral energy distribution (SED) fitting. This method involves broadband photometry and finding the best-fit templates to the SEDs. From these fits, one is able to determine properties such as star formation rates (SFRs), metallicities and stellar masses. For the CNOC2 groups, Balogh et al. (2007) use the Bruzual & Charlot (2003) stellar population model and the two-component dust model of Charlot & Fall (2000) to fit the SEDs.

We plot stellar mass per galaxy versus radius for the two Gaussian groups in Figure 5.4, where we see that Groups 110 and 308 show a similar trend of decreasing stellar mass content with radius. Both Gaussian groups have a galaxy with a maximum stellar mass of $\sim 2.2 \times 10^{11} M_{\odot}$ close to the group center. Past a radius of 0.4 Mpc, we do not observe any galaxies with stellar masses $> 3 \times 10^{10} M_{\odot}$.

The stellar mass versus radius plots for the three non-Gaussian groups (Figure 5.5) show much more scatter than the Gaussian groups. For Groups 138 and 226, the galaxy with the maximum stellar mass does not lie close to the group center, as with the Gaussian groups, but beyond a radius of 0.4 Mpc. Also with the Gaussian groups, we find that at larger radii the stellar mass values of the galaxies decreased significantly from those closer to the group center. This is a trend that we do not observe in Groups 138, 226 and 346 (non-Gaussian), which all have galaxies with stellar masses $> 5 \times 10^{10} M_{\odot}$ at the very edge of the group.

It should be noted that estimates of stellar mass are currently being improved with the addition of GALEX and MIPS 24 μm photometry, which will allow for a

more accurate comparison between the Gaussian and non-Gaussian groups.

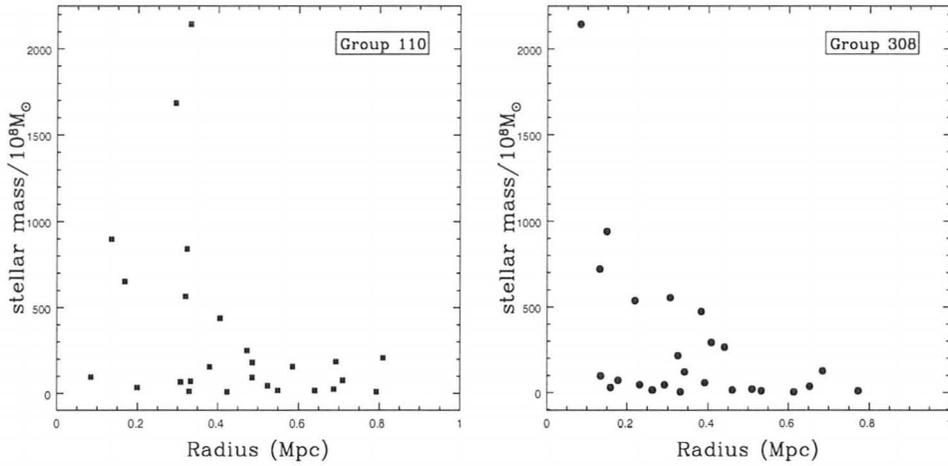


Figure 5.4: Left: Stellar mass vs radius radius plot for Group 110. Right: Stellar mass vs radius plot for Group 308. The closed symbols indicate that these groups have been classified as Gaussian by the A–D test.

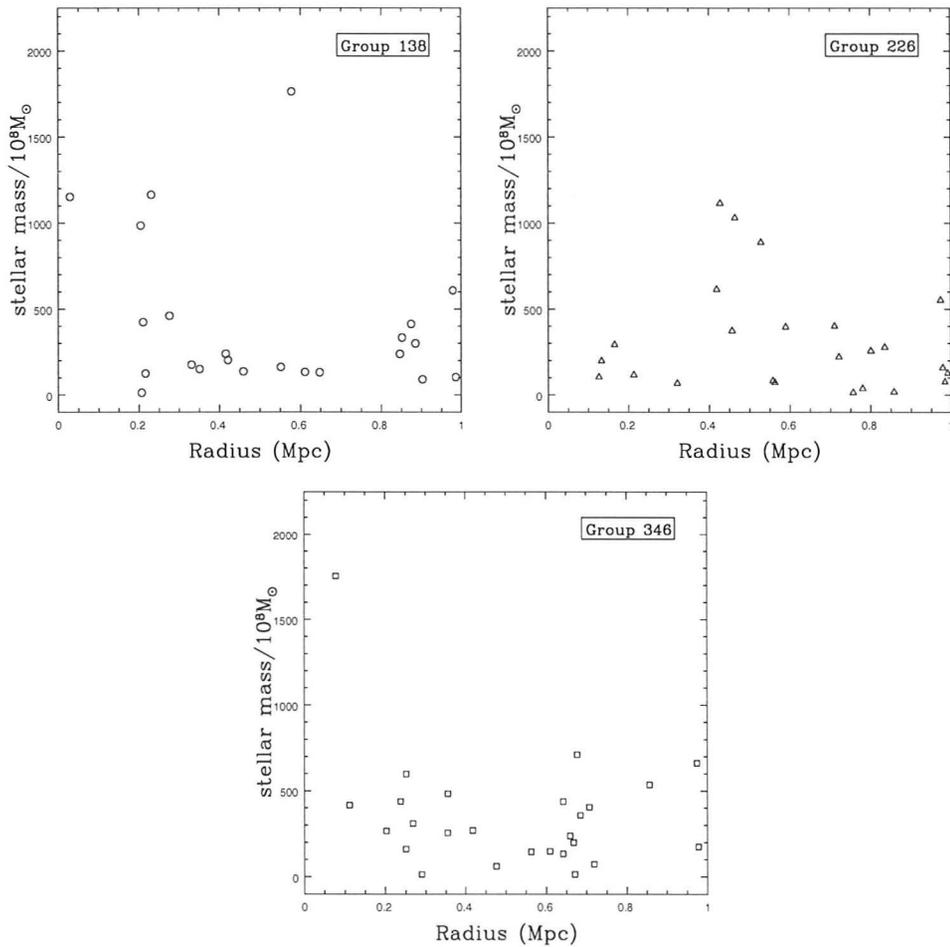


Figure 5.5: Top-left: Stellar mass vs radius plot for Group 138. Top-right: Stellar mass vs radius plot for Group 226. Bottom: Stellar mass vs radius plot for Group 346. The open symbols indicate that these groups have failed the A–D test and are classified as non-Gaussian.

5.2 Stacked Group Properties

One of the goals of this project was to compare group properties of the Gaussian and non-Gaussian groups to see if any obvious trends would emerge. This task is extremely challenging due to the small sizes of the majority of the CNOC2 groups (Figure 4.2) and it is impossible to draw conclusions about radial trends, for example, for a group with 5 members. Analysis of our largest galaxy groups, examined in §1 of Chapter 5, showed some possible trends, but since our sample contained only five groups it is difficult to claim that any of the observed correlations were true for *all* Gaussian or non-Gaussian groups. One way to overcome this issue is to stack the groups classified as either Gaussian or non-Gaussian by the A–D test and study the stacked group properties. This method increases our sample size and thus allows us to more adequately search for trends. The stacked Gaussian group is a composite of 598 member galaxies that lie within 1 Mpc of the group center and the stacked non-Gaussian groups contains a total of 314 member galaxies which lie within our radius cut.

5.2.1 Velocity Dispersion Profiles

In §1.1 of this chapter, we presented VDPs for five individual CNOC2 groups, with $n \geq 20$ (Figure 5.1), and concluded that the profiles of the rich groups classified as non-Gaussian by the A–D test showed distinctive features, i.e., rising profiles, from the classified Gaussian groups. Although, this result suggests that the non-Gaussian groups, 138, 226 and 346, are dynamically more complex than Groups 110 and 308 (Gaussian), we could not make any statements about the profiles of *all* groups classified as non-Gaussian (or Gaussian) based on such a small sample. Therefore, in order to search for general trends we compute VDPs for the stacked Gaussian and non-Gaussian groups, following the method outlined in §1.1 of this chapter. The profiles of both stacked Gaussian (open circles) and non-Gaussian (closed triangles)

groups are shown in Figure 5.6. From these profiles it would appear that the combined non-Gaussian groups have a higher σ_p than the combined Gaussian groups, with a peak value of $\sim 630 \text{ km s}^{-1}$, significantly more than the peak Gaussian group σ_p of $\sim 440 \text{ km s}^{-1}$. However, when we compute the velocity dispersion of the stacked groups, using the Gapper algorithm (Equation 3.1), we find that $\sigma_{\text{Gapper}} = 462 \pm 111 \text{ km s}^{-1}$ for the stacked Gaussian groups and $\sigma_{\text{Gapper}} = 611 \pm 173 \text{ km s}^{-1}$ for the stacked non-Gaussian groups, indicating that velocity dispersions of the stacked Gaussian and non-Gaussian groups are in fact in agreement within error.

The profile of the stacked non-Gaussian groups shows an overall rising trend, spanning roughly 150 km s^{-1} , which is larger than the mean intrinsic velocity dispersion error of 125 km s^{-1} computed by W05¹. The rising trend seen in the non-Gaussian groups differs from the relatively flat trend seen in Gaussian groups, with a projected dispersion range of only $\sim 40 \text{ km s}^{-1}$.

¹We can qualitatively use the errors computed by W05 since our final intrinsic velocity dispersion values are comparable (see Chapter 4.2)

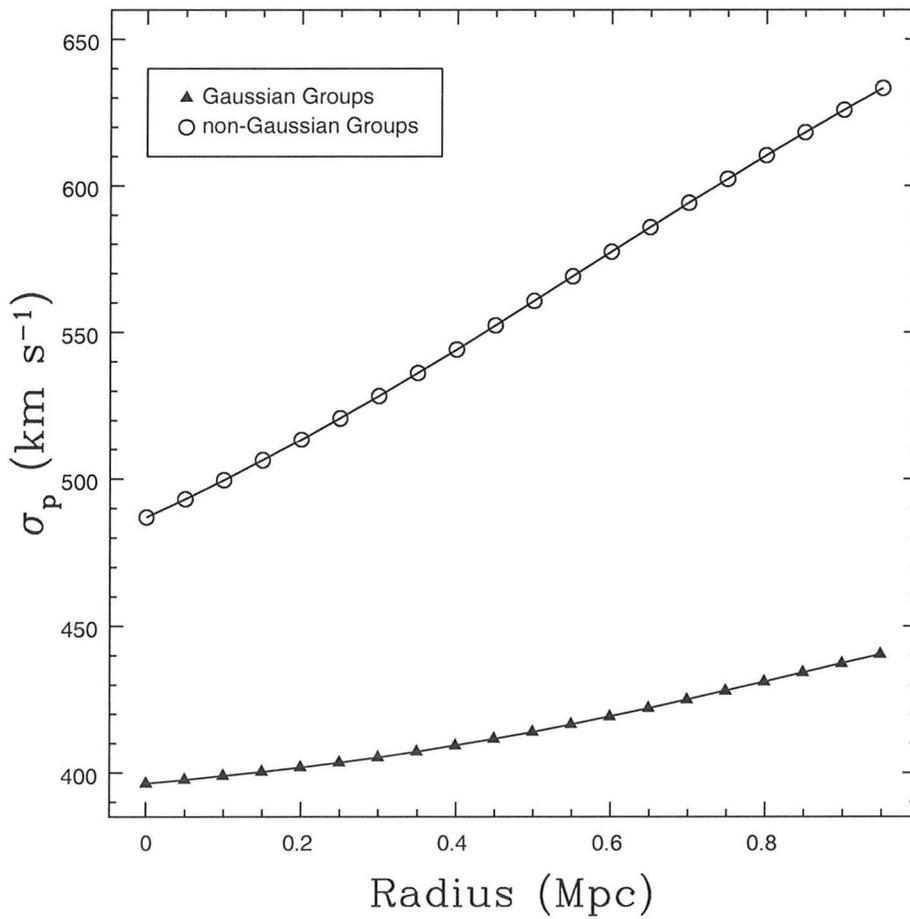


Figure 5.6: Velocity dispersion profiles for all the stacked Gaussian (triangle) and non-Gaussian (circles) groups.

To ensure that the features shown in Figure 5.6 are indeed real, and not driven by the individual rich groups, we compute VDPs of the stacked groups excluding the rich, $n \geq 20$, groups. The stacked Gaussian groups now contain a total of 547 member galaxies, while the stacked non-Gaussian groups are now a composite of 240 members. The results of our stacked ($n \leq 20$) VDPs are plotted in Figure 5.7, where we see that the higher σ_p values of the combined non-Gaussian groups over the Gaussian groups almost disappears. The computed velocity dispersions now become $\sigma_{\text{Gapper}} = 489 \pm 148 \text{ km s}^{-1}$ for the stacked non-Gaussian groups and $\sigma_{\text{Gapper}} = 456 \pm 112 \text{ km s}^{-1}$ for the stacked Gaussian groups. Similarly to the dispersion values when all groups are included, the velocity dispersions of the stacked $n < 20$ Gaussian and non-Gaussian groups are also in good agreement, suggesting that there is no significant difference in the velocity dispersions of the relaxed and dynamically complex groups.

The rising trend seen with *all* non-Gaussian groups combined, is arguably still observed when we exclude the $n \geq 20$ groups, but the slope is less steep, spanning a smaller σ_p range of $\sim 110 \text{ km s}^{-1}$, which is equivalent to the mean intrinsic velocity dispersion error (W05). The overall shape of the stacked Gaussian groups, minus the two $n \geq 20$ groups (110 and 308), does not change significantly from the profile that includes *all* groups, maintaining a generally flat trend and covering the same σ_p range.

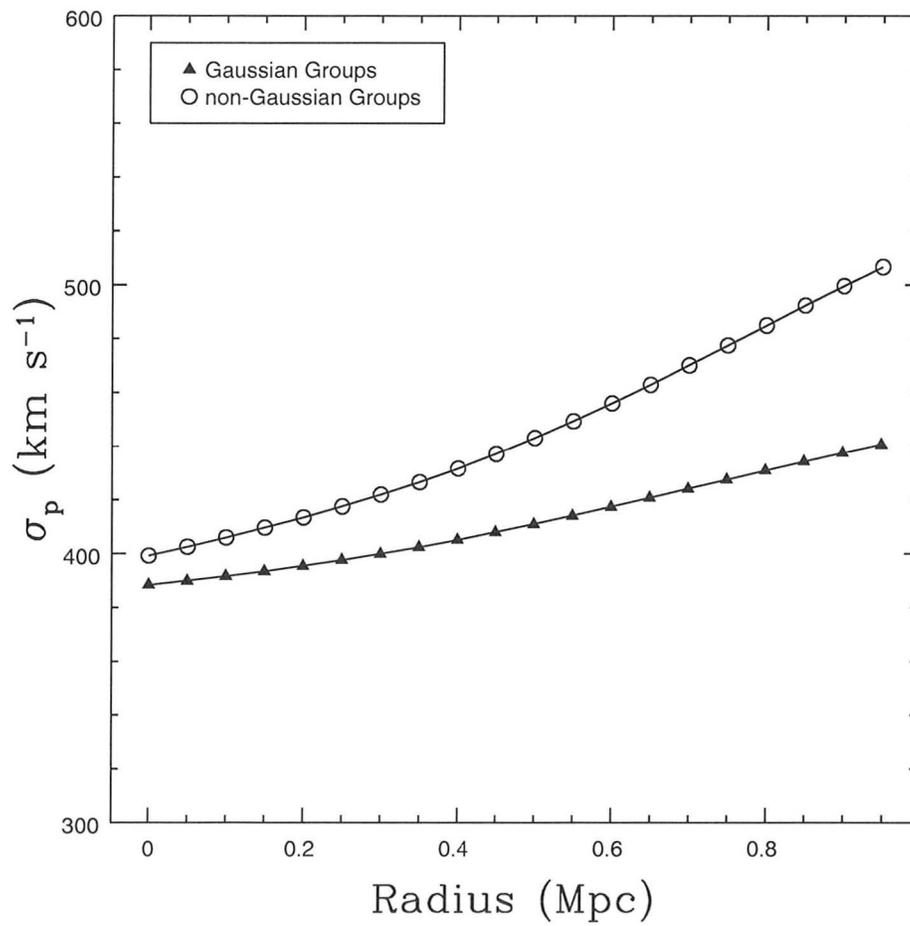


Figure 5.7: Velocity dispersion profiles of the stacked Gaussian (triangle) and non-Gaussian (circles) groups, excluding the $n \geq 20$ groups.

In general, when we stack all 106 CNOC2 groups in our sample, we find that the VDP shown in Figure 5.8 seems to follow a general flat trend, with the projected velocity dispersion values covering a range of 100 km s^{-1} , which is less than the mean intrinsic velocity dispersion error of 125 km s^{-1} . Although, it should be noted that the stacked VDP could also be interpreted as increasing with radius. Since the range in dispersion values is only slightly less than the mean intrinsic velocity dispersion error, we cannot definitively rule out the possibility that the observed rising profile is not real feature of the groups.

Other methods of computing projected VDPs for the CNOC2 groups have resulted in conflicting conclusions. Carlberg et al. (2001) derived projected velocity dispersions of the CNOC2 groups by modelling the projection of a three-dimensional velocity dispersion. They found that the projected VDP rose slowly with radius. Conversely, Parker et al. (2005) obtained weak-lensing signals for a sample of 116 CNOC2 galaxy groups and found that the projected velocity dispersion decreased with radius, which is similar to the profiles of observed galaxy clusters (Girardi et al., 2002; Lokas & Mamon, 2003). In studies of other group catalogs, Zabludoff & Mulchaey (1998), Girardi et al. (2002) and Brough et al. (2006) also found that the velocity dispersion profiles fell with radius.

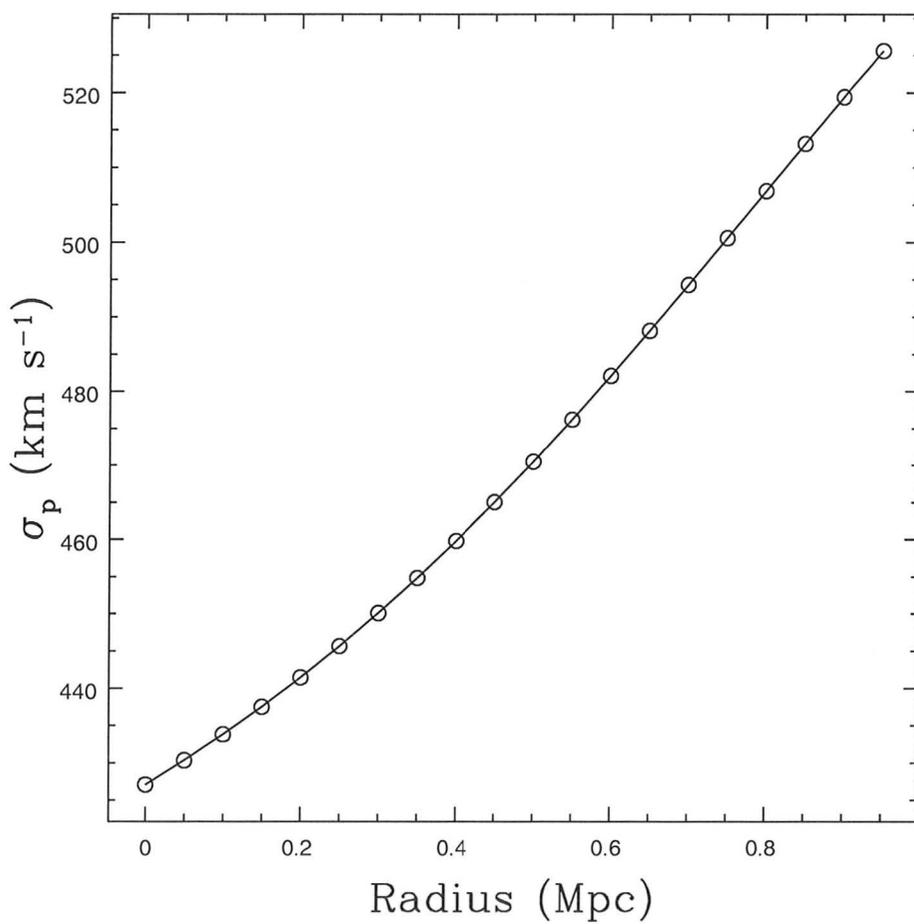


Figure 5.8: Velocity dispersion profile of the stacked 106 CNOC2 groups with $n \geq 5$ after a 1 Mpc radius cut.

5.2.2 Colour - Magnitude

The bimodal nature of the galaxy rest-frame colour-magnitude diagrams (hereafter CMD) is now well established, with the majority of galaxies falling in either the red sequence or blue cloud and it is believed that this trend was already in place by $z \sim 1$ (Bell et al., 2004). Hogg et al. (2002) show with the SDSS catalog that the red sequence mainly consists of early-type galaxies, with little or no star formation, while the blue cloud is dominated by late-type (spiral) galaxies, where star formation is on-going. In Chapter 1, we discussed how the increased galaxy interaction or merger rates in the group environment (Brough et al., 2006) could trigger star formation in groups where these processes occur. This would suggest that the dynamically complex groups should have more galaxies on the blue cloud than groups that are relaxed and virialized. Therefore, if the groups classified as non-Gaussian by the A–D test are indeed dynamically more complex, we would expect them to have more members which are morphologically late-type galaxies.

In Figure 5.9 we plot colour versus magnitude ($g - r$ vs M_r) of the stacked Gaussian groups, where we clearly observe a distinct red sequence and blue cloud. The bimodality of the stacked Gaussian group CMD is also seen in the $g - r$ histogram shown in Figure 5.10, where we see two clear peaks. From this histogram, it appears that $g - r = 0.5$ defines the line between galaxies which are on the red sequence and those that are in the blue cloud. Of the 598 galaxies, only 113 fall below $g - r < 0.5$, indicating that over 80% of the Gaussian group members lie in the red sequence. This result is in agreement with our claim that groups with underlying Gaussian velocity distributions are indeed relaxed and virialized systems, similar to observations of the know well-established heavily populated red sequence observed in clusters (Gladders & Yee, 2000).

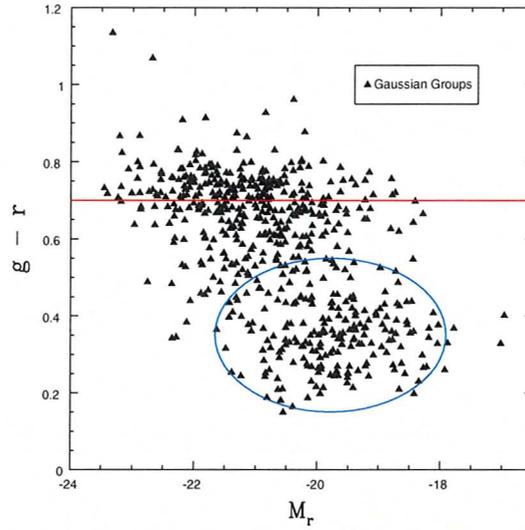


Figure 5.9: Colour-Magnitude diagram of all stacked Gaussian Groups. The red line indicates the peak $g-r$ value of the red sequence, determined from Figure 5.10 and the blue ellipse indicates the approximate area of the blue cloud.

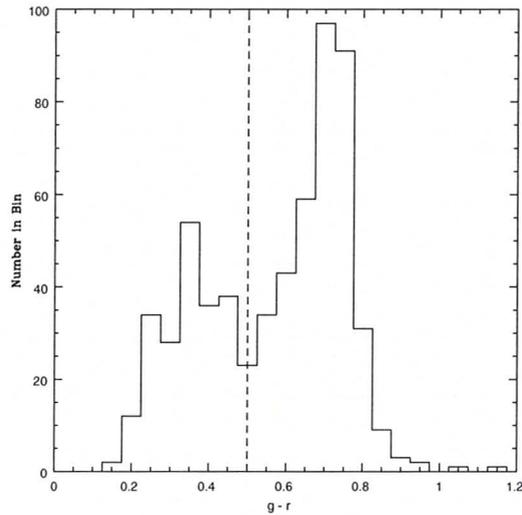


Figure 5.10: $g-r$ histogram for the stacked Gaussian groups using a bin width of 0.05. The dashed vertical line indicates the $g-r$ value that distinguishes between galaxies in the red sequence and blue cloud.

In Figure 5.11 we plot $g - r$ vs M_r for the stacked non-Gaussian groups. Here, we also see a distinct red sequence and blue cloud, which is also shown in Figure 5.12 with a double peak in the $g - r$ histogram. Similarly to the colour histogram for the stacked Gaussian groups (Figure 5.10), the distinction between red sequence and blue cloud appears to lie at $g - r \approx 0.5$. Qualitatively, there do not appear to be a higher fraction of galaxies in the blue cloud of the non-Gaussian groups (Figure 5.11), but when we compute the percentage of members with $g - r < 0.5$, the f_b , we find that $\sim 42\%$ of galaxies lie in the blue cloud. Thus, the stacked non-Gaussian groups have $\geq 20\%$ more blue galaxies than the combined Gaussian groups. Since galaxies in the blue cloud are generally late-type star forming galaxies, the results of our stacked colour-magnitude diagrams support our claim that classified non-Gaussian groups are dynamically more complex systems and have increased SFRs.

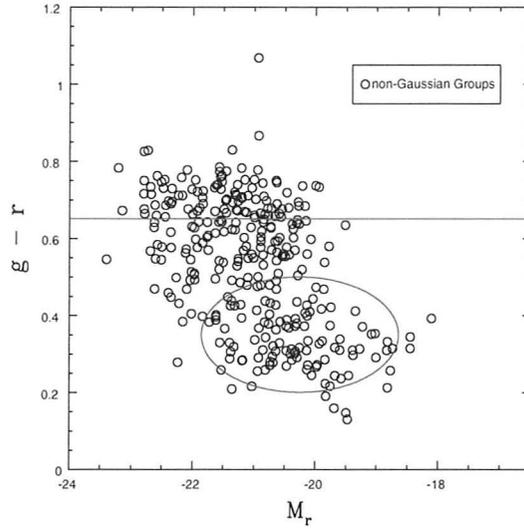


Figure 5.11: Colour-Magnitude diagram of all stacked non-Gaussian Groups. The red line indicates the peak $g-r$ value of the red sequence, determined from Figure 5.12 and the blue ellipse indicates the approximate area of the blue cloud.

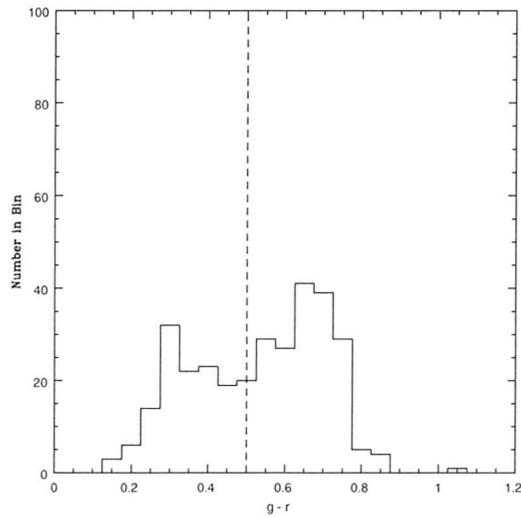


Figure 5.12: $g-r$ histogram for the stacked non-Gaussian groups using a bin width of 0.05. The dashed vertical line indicates the $g-r$ value that distinguishes between galaxies in the red sequence and blue cloud.

5.2.3 Colour-Radius

In §1.2 of this chapter we compared colour versus radius plots of the five Gaussian and non-Gaussian CNOC2 groups with $n \geq 20$ and found that the two Gaussian groups (110 and 308) followed the colour-radius relation observed in clusters, while the non-Gaussian groups (138,226 and 346) did not show any obvious correlations. We now look at the colour versus radius plots of the stacked Gaussian (Figure 5.13) and non-Gaussian (Figure 5.15) groups to see if the trends observed in the individual groups are still present.

When we combine all of the galaxies in the groups classified as Gaussian by the A-D test, we find that the colour-radius relation observed in Groups 110 and 308 does not appear to exist in the colour versus radius plot (Figure 5.13), with galaxies populating the entire colour range of $0.1 \leq g - r \leq 0.9$ from the center of the group out to 1.0 Mpc and showing no visible gradient. However, when the mean colours for the galaxies within 0.4 Mpc of the group center are computed, we find that $\overline{g - r} = 0.605 \pm 0.009$, while $\overline{g - r} = 0.508 \pm 0.012$ for galaxies in the group-centric radius range of 0.4 - 1.0 Mpc. These mean values suggest that the colour-radius relation observed in the rich individual galaxy groups (Figure 5.2) is still present when one stacks all of the classified Gaussian groups. In Figure 5.14, we compute the $\overline{g - r}$ values in bins of 0.1 Mpc and find that the cluster population gradient, an increase of the blue fraction with radius (Ellingson et al., 2001), is also observed in the stacked Gaussian groups. We performed a weighted least square fit on Figure 5.14 and found the slope to be consistent with a non-zero negative value.

Similarly, the colour versus radius plot for the stacked non-Gaussian groups (Figure 5.15) show similar features as the rich individual non-Gaussian groups (Figure 5.3), showing no correlation between colour and radius. Computing the $\overline{g - r}$ value for galaxies in the inner and outer regions of the group further indicates that there is no correlation between colour and radius for the non-Gaussian groups, where we find that $\overline{g - r} = 0.534 \pm 0.014$ for galaxies within 0.4 Mpc of the group centroid and

$\overline{g-r} = 0.526 \pm 0.13$ for galaxies beyond 0.4 Mpc. The lack of a population gradient is better observed when we look at the binned $\overline{g-r}$ values, shown in Figure 5.16, where we observe a relatively flat trend, as opposed to the clear negative slope observed in the stacked Gaussian groups (Figure 5.14). We carried out a weighted least square fit for Figure 5.16 and found the slope to be consistent with zero.

We reiterate that although stacking groups is needed, we cannot claim that the colour-radius relation does not exist in the classified non-Gaussian groups, as the wide range in redshift and velocity dispersion in our group sample may affect our results. We would require a large sample of groups at the same redshift and in a smaller dispersion range in order to effectively rule out any correlations or alternatively applying a more sophisticated method of stacking our groups, with a proper normalization scheme which would allow us to better compare groups at different redshifts and dispersions.

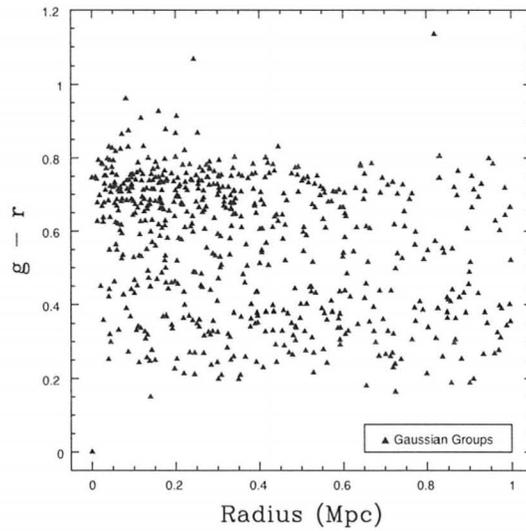


Figure 5.13: Colour-Radius diagram of all stacked Gaussian Groups.

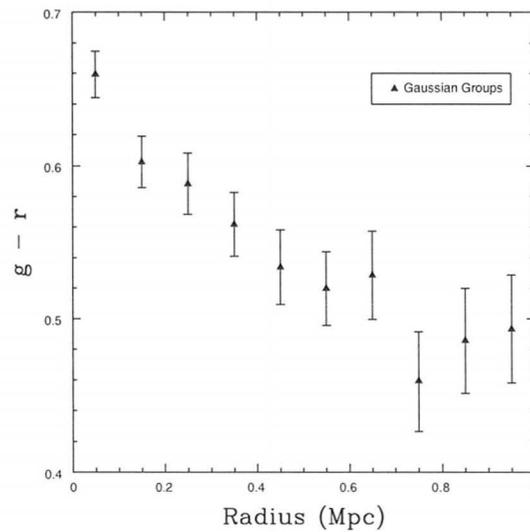


Figure 5.14: Binned Colour-Radius diagram of all stacked Gaussian Groups. Mean $g - r$ values are computed in bin widths of 0.1 Mpc. We applied a weighted least square fit and found the slope to be consistent with a non-zero negative slope.

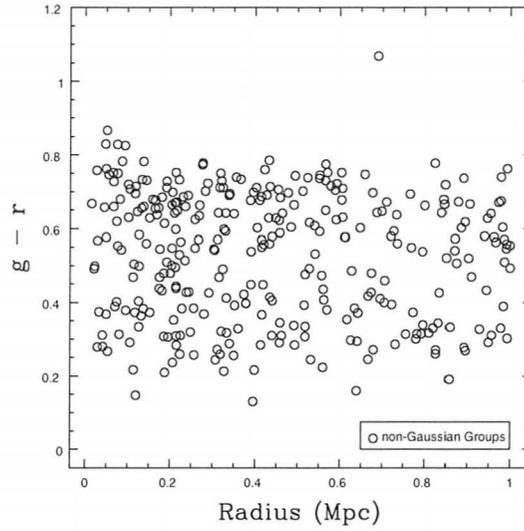


Figure 5.15: Colour-Radius diagram of all stacked non-Gaussian Groups.

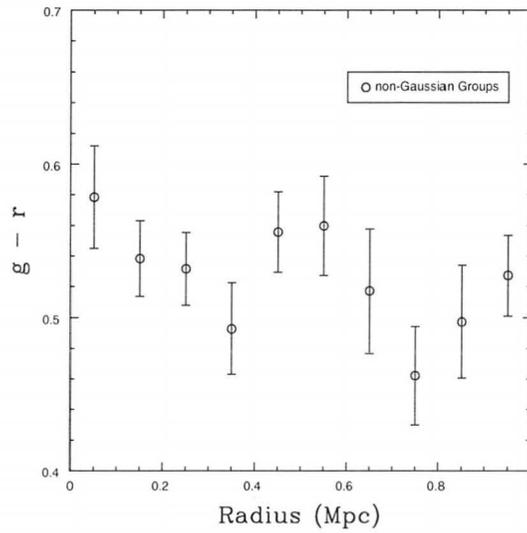


Figure 5.16: Binned Colour-Radius diagram of all stacked non-Gaussian Groups. Mean $g - r$ values are computed in bin widths of 0.1 Mpc. We applied a weighted least square fit and found the slope to be consistent with zero.

5.2.4 Star Formation Rate: O[II] Equivalent Widths

In addition to colour, there are other frequently used indicators of star formation. In particular, indicators of specific star formation rates (hereafter SSFR), defined as the star formation rate per unit stellar mass, such as $H\alpha$ and O[II] $\lambda 3727$ emission line equivalent widths (hereafter EW[$H\alpha$] and EW[OII]) and K-band luminosities, allow for the study of how star formation contributes to the growth of galaxies (Bauer et al., 2005). The bimodal nature observed in galaxy populations is also present in measures of SSFRs. Balogh et al. (2004) used EW[$H\alpha$] as an indicator of SSFR to show that galaxies that reside in the group environment form two distinct populations, star-forming and quiescent galaxies. This result is similar to the red sequence and blue cloud galaxy populations observed in the CMDs of galaxy clusters.

Although EW[$H\alpha$] is often used to determine SSFR, the $H\alpha$ emission line is not visible with the LDSS2 spectrograph beyond $z \sim 0.21$, thus in the CNOC2 redshift range of $0.3 \leq z \leq 0.55$, W05 measures SSFR using EW[OII], targeting 20 CNOC2 groups and 6 serendipitous groups.

The results of our stacked CMDs (see Chapter 5.2.2) indicated that the non-Gaussian groups had a higher fraction of galaxies in the blue cloud. Based on this result one would expect measures of SSFR to also be higher for the stacked non-Gaussian groups, since blue galaxies are generally star-forming. Following Wilman et al. (2005a) we employ a division at $EW[OII] = 5 \text{ \AA}$ to distinguish between passive ($EW[OII] < 5 \text{ \AA}$) and star-forming ($EW[OII] \geq 5 \text{ \AA}$) galaxies. Based on this definition, we find no significant difference in the fraction of star-forming galaxies between the Gaussian (Figure 5.17) and non-Gaussian (Figure 5.18) groups. For the non-Gaussian groups, there are 144 galaxies with measured EW[OII], 52% of which have equivalent widths $> 5 \text{ \AA}$. There are 146 galaxies with measured EW[OII] for the Gaussian groups and 56% have values $> 5 \text{ \AA}$. These results indicate that the members of the non-Gaussian groups do not have more star-forming galaxies, which seems to contradict the results of CMDs (Figure 5.9 and 5.11) which show that the non-Gaussian groups

have 20% more galaxies in the blue cloud.

Although we do not observe any significant differences in the EW[OII] values between the Gaussian and non-Gaussian groups, there does appear to be a trend, for both categories, with radius. From Figures 5.17 and 5.18, it is clear that there is a large concentration of galaxies with $\text{EW}[\text{OII}] < 5 \text{ \AA}$ close to the group center. As you move out towards the edge of the group, the number of star-forming ($\text{EW}[\text{OII}] \geq 5 \text{ \AA}$) increases.

It should be noted that although the O[II] is used as an indicator of star formation at intermediate redshifts, Jansen et al. (2001) claim that it is in fact not the best measure of SFR due to the effects of reddening and metallicity-dependent excitations in the interstellar medium. Additionally, these authors find that the O[II]/H α ratio is higher, by a factor of ~ 7 , for lower luminosity galaxies. Currently better indicators of SFRs, using GALEX and MIPS 24 μm observations, are being obtained and will be available for comparison.

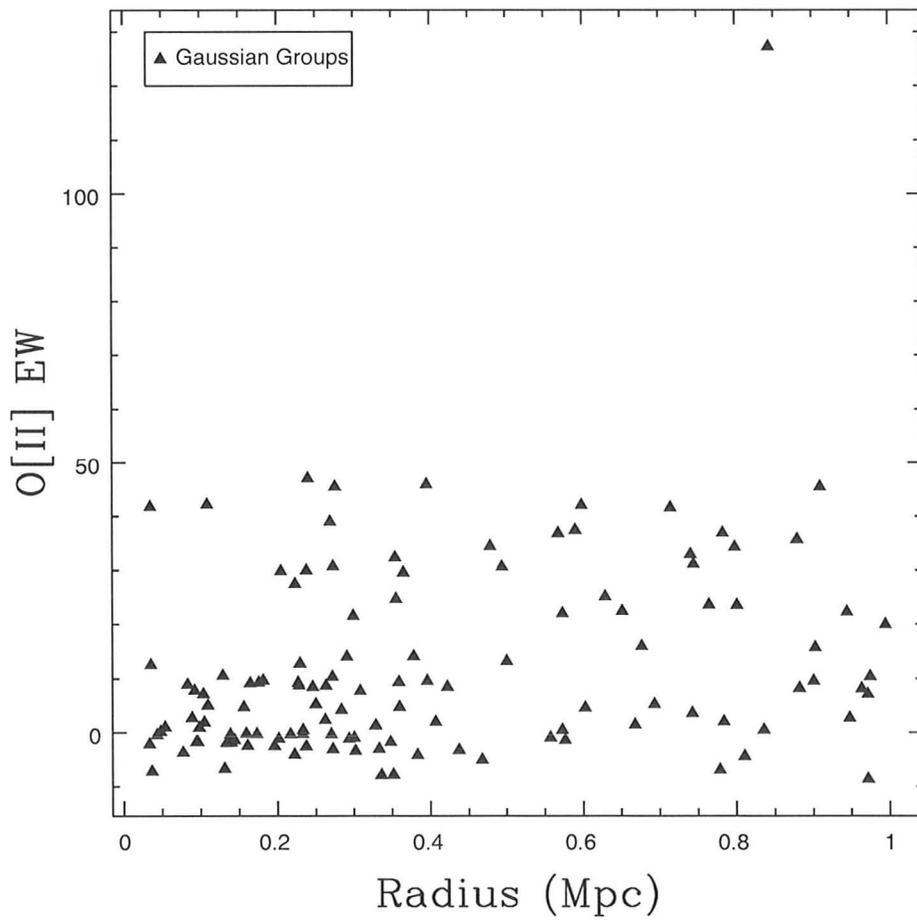


Figure 5.17: O[II] equivalent width vs radius for the stacked Gaussian groups.

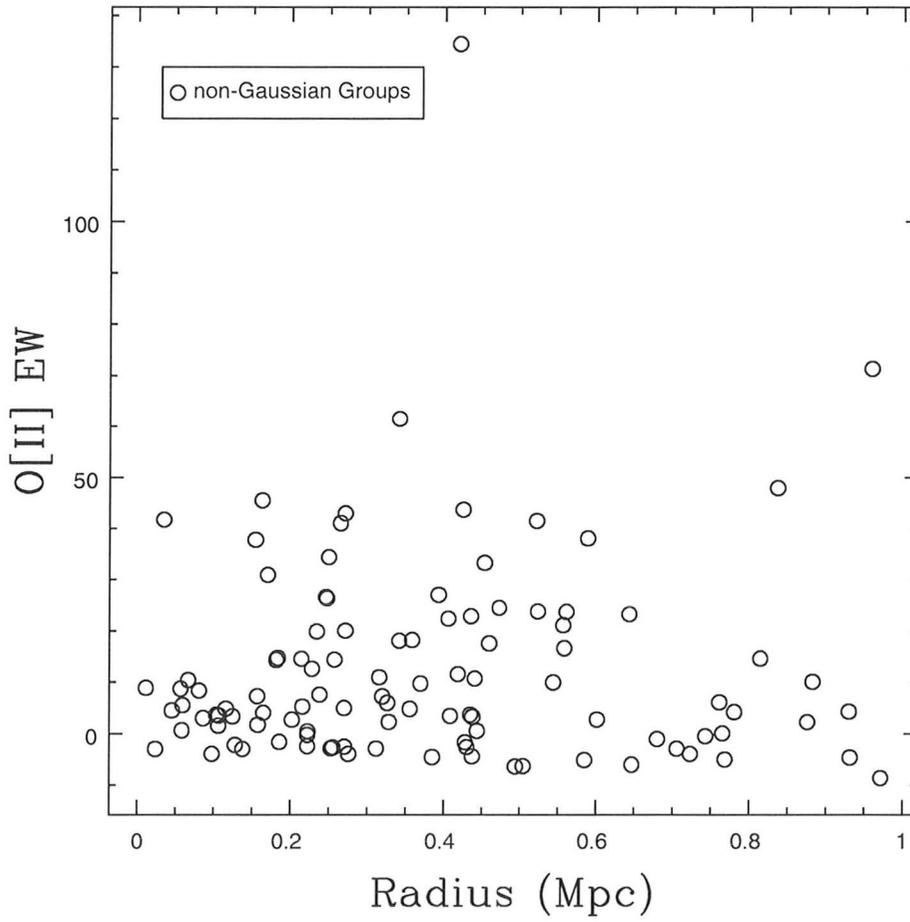


Figure 5.18: O[II] equivalent width vs radius for the stacked non-Gaussian groups.

5.2.5 Stellar Mass

In §1.4 of this chapter, we compared the stellar mass per galaxy of individual CNOC2 groups with $n \geq 20$. Here we compare the stellar mass versus radius plots of the stacked Gaussian (Figure 5.19) and non-Gaussian (Figure 5.20) groups. Both the stacked Gaussian and non-Gaussian groups show a similar trend with radius. With the exception of a few galaxies, the stellar mass content of the galaxies decreases with increasing radius, a trend observed in Groups 110 and 308 (Figure 5.4). Conversely, the scatter seen in the individual non-Gaussian groups (Figure 5.5) seems to disappear when we stack all of the galaxies in the classified non-Gaussian groups together. The stellar masses in Figure 5.20 now appear to follow a stellar-mass versus radius trend, although the gradient is stronger in the stacked Gaussian group. The galaxies with the highest stellar mass per galaxy generally lie close to the group centers for the stacked non-Gaussian groups, where as in Groups 138 and 226, the high stellar mass galaxies fell closer to the edge of the group.

Another feature observed in the individual Gaussian groups was that at radii > 0.4 Mpc, there were no galaxies with stellar masses $> 3 \times 10^{10} M_{\odot}$. This is not true of the stacked Gaussian groups (Figure 5.19), where there are several galaxies with stellar masses of $\sim 1 \times 10^{11} M_{\odot}$ out to 1 Mpc. This feature is observed for both the individual and stacked non-Gaussian groups.

A qualitative comparison of the stellar mass versus radius plots for the stacked Gaussian and non-Gaussian groups appears to show no significant difference between the two categories. Both Figures 5.19 and 5.20 exhibit stellar mass versus radius trends, and have galaxies with similar maximum stellar mass content ($\sim 3 \times 10^{11} M_{\odot}$). In the future we intend to perform a quantitative comparison of these two distributions using a two dimensional Kolmogorov-Smirnov test to accurately determine whether or not these two distributions are similar.

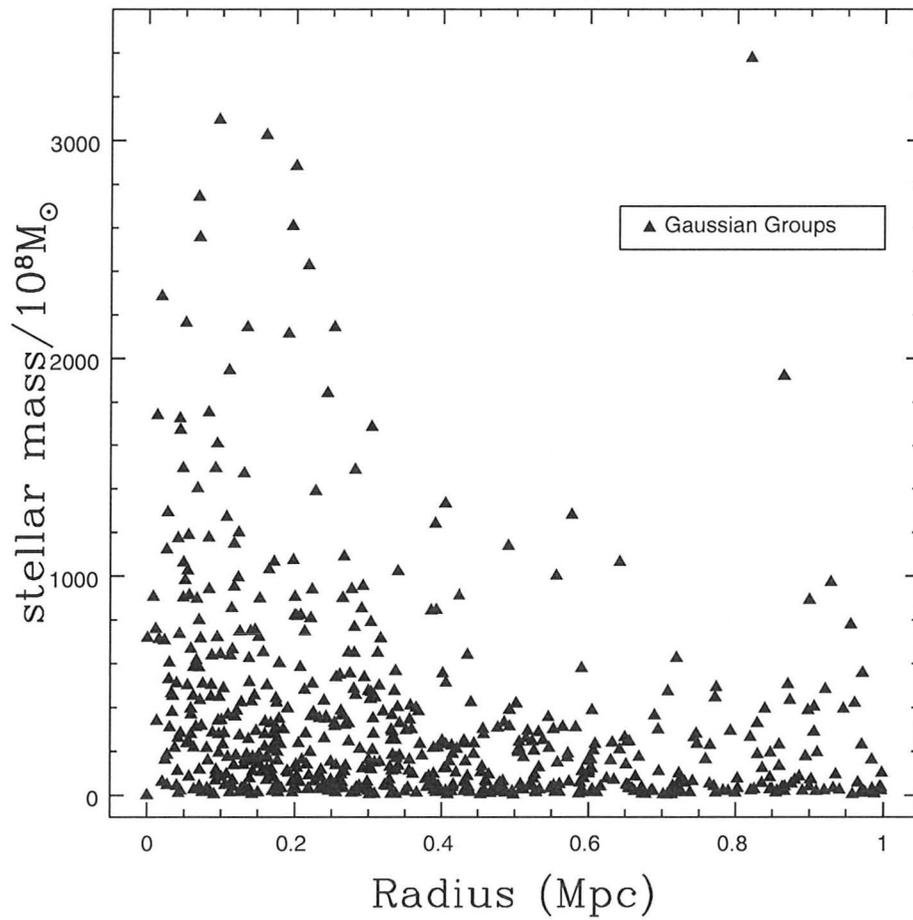


Figure 5.19: Stellar Mass/ $10^8 M_{\odot}$ versus Radius (Mpc) diagram of all stacked Gaussian groups.

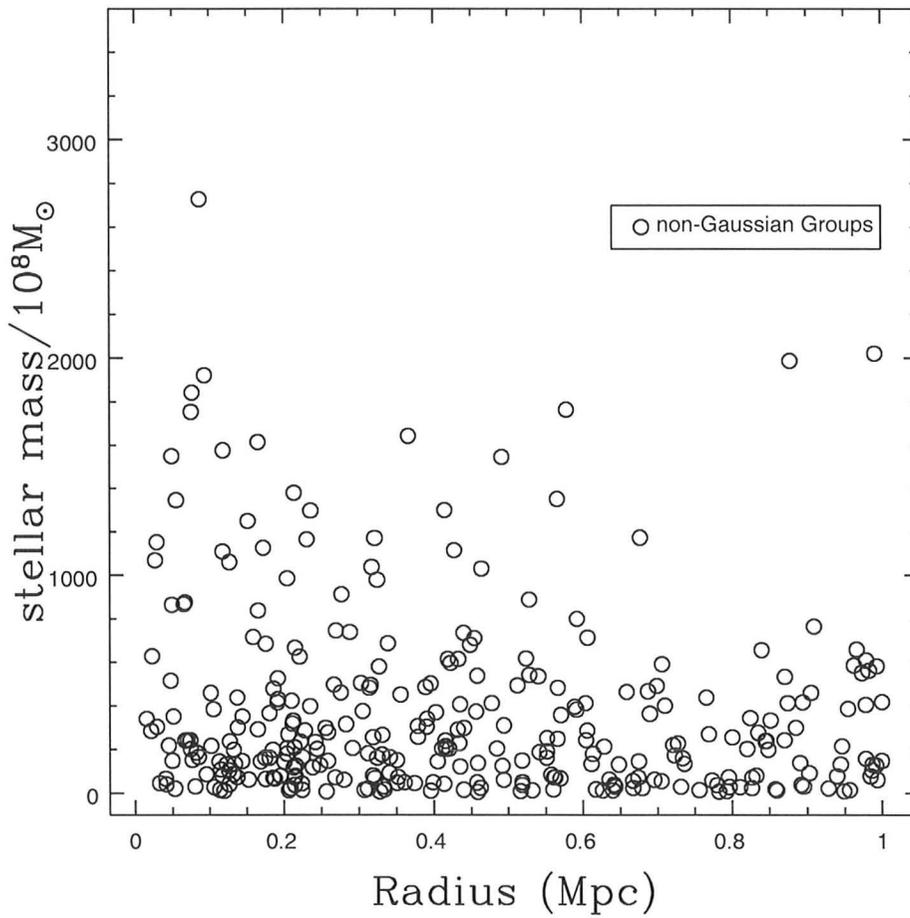


Figure 5.20: Stellar Mass/ $10^8 M_{\odot}$ versus radius (Mpc) diagram of all stacked non-Gaussian groups.

Chapter 6

Conclusions

Analysis of galaxy group dynamics requires the use of tools that are reliable and powerful for small sample sizes. We have applied three goodness-of-fit tests, the χ^2 , Kolmogorov and Anderson–Darling tests, to a subset of the CNOC2 groups in order to determine which test can best classify galaxy group dynamics. Based on our initial application of the aforementioned tests and on the results of our Monte Carlo simulations and power studies, we conclude that the Anderson–Darling test is the most reliable statistic to distinguish between relaxed (Gaussian) and dynamically complex (non-Gaussian) groups. The problems introduced by small sample sizes for the χ^2 test and the lack of power and sensitivity of the Kolmogorov test rule out these statistics as useful classification tools.

The results of our Monte Carlo simulations for the Gapper Estimator (Equation 3.1) and rms dispersion calculations indicate that for small sample size, $n < 50$, the Gapper algorithm is a more accurate estimate of the true velocity dispersion, which is in agreement with Beers et al. (1990).

We then apply the Anderson–Darling test to all CNOC2 groups with $n_{\text{members}} \geq 5$, after a 1 Mpc radius cut, using the mean velocity of the group members as the estimated μ and the intrinsic velocity dispersion (Equation 4.3) as the estimated σ . The groups are then classified as being in either a relaxed (Gaussian) or complex

(non-Gaussian) dynamical system, based on the $\alpha = 0.05$ critical value. The results of our analysis indicate that 34 of the 106, or $\sim 32\%$ of the sample of CNOC2 groups are non-Gaussian.

To investigate our claim that classified non-Gaussian groups are indeed dynamically more complex than Gaussian ones, we look at the velocity dispersion profiles of 5 CNOC2 groups with $n_{\text{members}} \geq 20$. Analysis of the resulting profiles indicates that;

1. The profiles of the 2 Gaussian groups (110 *and* 308) show a declining projected velocity dispersion with radius;
2. Two non-Gaussian groups (138 *and* 346) have rising profiles, a possible signature of merger activity (Menci & Fusco-Femiano, 1996), and;
3. The profile of Group 110 flattens towards larger radii, a trend observed by Girardi et al. (1996) in galaxy clusters.

We find that the VDPs of the Gaussian and non-Gaussian groups are distinct, supporting our claim that the classified non-Gaussian groups are dynamically different from the Gaussian systems.

In our comparison of the group properties of the 5 rich, $n \geq 20$, CNOC2 groups, classified as either Gaussian (Groups 110 and 308) or non-Gaussian (Groups 138, 226 and 346) by the Anderson–Darling test, we conclude that:

1. Groups 110 and 308 appear to follow the colour-radius relation observed in clusters, while Groups 138, 226 and 346 show no visible correlation between colour and radius, although there is an increase of blue galaxies with radius for Groups 226 and 346;
2. The average stellar mass per galaxy of Groups 110 and 308 decreases with radius, while Groups 138, 226 and 346 show no obvious stellar mass-radius trend;

We also look at the group properties for the stacked Gaussian and non-Gaussian groups and find that in general the trends observed in the individual rich groups tend to disappear when all the galaxies are stacked together. Comparison of the stacked group properties indicates;

1. The VDP of the stacked non-Gaussian groups, including *all* groups with $n \geq 5$, rises with radius, while the stacked Gaussian group profiles is consistent with being flat;
2. The velocity dispersions, σ_{Gapper} , of the stacked Gaussian and non-Gaussian groups are in agreement to within error;
3. The VDPs of the stacked groups, excluding the $n \geq 20$ groups, shows that the profile of the non-Gaussian groups still rises, but less steeply, while the VDP of the Gaussian groups remains flat, and also the computed velocity dispersion for both stacked groups is relatively the same;
4. Stacking the galaxies in all 106 CNOC2 groups produces a profile that is consistent with being flat with radius;
5. We compute the fraction of blue galaxies, f_b , from the stacked colour-magnitude diagrams and find that the non-Gaussian groups have 20% more galaxies in the blue cloud than the Gaussian groups;
6. The stacked Gaussian groups show an increase of blue galaxies with radius, while the stacked non-Gaussian groups show no correlation between colour and radius;
7. The EW[OII] values also show no significant difference between the Gaussian and non-Gaussian groups, although both groups do exhibit a slight trend with radius;

8. Both the Gaussian and non-Gaussian groups have decreasing average stellar mass per galaxy with radius;

We emphasize that although we may not observe some trends, we do not exclude the possibility that they may exist. Combining galaxies in groups from such a wide range of velocity dispersion and sizes could wash out real trends. Also, a more sophisticated method of stacking the groups, i.e., normalizing the groups to the virial radius or to a specific velocity dispersion or mass, would reduce these effects.

Recently, Finoguenov et al. (2009) have observed 25 X-ray selected groups in two of the CNOC2 fields, with 6 groups having CNOC2 optical counterparts, Groups 1, 11, 28, 104, 117, and 138. X-ray bright groups are of particular interest, as only groups in a relaxed dynamical state are expected to produce extended X-ray emissions. In a preliminary analysis of the 6 groups with both optical and X-ray observations, we find that Groups 1, 11, 28, 104 and 117 have all been classified as having underlying Gaussian velocity distributions by the Anderson–Darling test, as we had expected. The really interesting case is Group 138, which is one of the brightest X-ray sources in the field, but has been classified as non-Gaussian and also has a rising velocity dispersion profile. The colour-radius and SFR-radius trends discussed in §1 of Chapter 5 also indicate that Group 138 is a more complex environment than one would have imagined and further analysis and comparison of optically- and X-ray selected groups is required to better understand the dynamics of galaxy groups.

Based on our analysis of several statistical tools, we conclude that the Anderson–Darling goodness-of-fit test is the most reliable statistical tool for classification of galaxy group dynamics. Although we have used this test to determine departures from normality, its application is not restricted to Gaussian distributions and can be used with many continuous or discrete distributions. Not only is this test reliable and powerful, especially when dealing with small sample sizes, but its application is simple and has the potential to be useful in many other areas of astronomy.

Future work will include further comparison of group properties, which will be

updated with MIPS and GALEX observations, allowing us to better identify correlations between the dynamical state of the group and its properties. We also plan to study the dynamics of X-ray selected galaxy groups (Connelly, J. et al, in prep; Finoguenov et al. (2009)) in the CNOC2 fields, comparing optically versus X-ray selected groups. A detailed study of simulated galaxy groups will allow us to understand the relationship between velocity dispersion and mass for systems in different dynamical states. Simulations will also help us to quantify how projection and contamination by interloping galaxies affect our measured velocity distributions. Finally, we aim to improve our classification of relaxed versus dynamically complex groups by adding spatial information to the velocity distribution, allowing us to identify both substructure and non-Gaussian velocity distributions.

Bibliography

Azzalini, A. & Capitanio, A. 1999, *J. Roy. Statist. Soc.*, series B, 61, 579

Balogh, M., Eke, V., Miller, C., Lewis, I., Bower, R., Couch, W., Nichol, R., Bland-Hawthorn, J., Baldry, I. K., Baugh, C., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Cross, N., Dalton, G., de Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Gomez, P., Gray, A., Hawkins, E., Jackson, C., Lahav, O., Lumsden, S., Maddox, S., Madgwick, D., Norberg, P., Peacock, J. A., Percival, W., Peterson, B. A., Sutherland, W., & Taylor, K. 2004, *MNRAS*, 348, 1355

Balogh, M. L., Wilman, D., Henderson, R. D. E., Bower, R. G., Gilbank, D., Whitaker, R., Morris, S. L., Hau, G., Mulchaey, J. S., Oemler, A., & Carlberg, R. G. 2007, *MNRAS*, 374, 1169

Bauer, A. E., Drory, N., Hill, G. J., & Feulner, G. 2005, *ApJ*, 621, L89

Beers, T. C., Flynn, K., & Gebhardt, K. 1990, *AJ*, 100, 32

Bell, E. F., Wolf, C., Meisenheimer, K., Rix, H.-W., Borch, A., Dye, S., Kleinheinrich, M., Wisotzki, L., & McIntosh, D. H. 2004, *ApJ*, 608, 752

Bergond, G., Zepf, S. E., Romanowsky, A. J., Sharples, R. M., & Rhode, K. L. 2006, *A&A*, 448, 155

- Bird, C. M. 1994, *AJ*, 107, 1637
- Brough, S., Forbes, D. A., Kilborn, V. A., & Couch, W. 2006, *MNRAS*, 370, 1223
- Bruzual, G. & Charlot, S. 2003, *MNRAS*, 344, 1000
- Burns, J. O. 1998, *Science*, 280, 400
- Butcher, H. & Oemler, Jr., A. 1984, *ApJ*, 285, 426
- Carlberg, R. G., Yee, H. K. C., Morris, S. L., Lin, H., Hall, P. B., Patton, D. R., Sawicki, M., & Shepherd, C. W. 2001, *ApJ*, 552, 427
- Charlot, S. & Fall, S. M. 2000, *ApJ*, 539, 718
- D'Agostino, R. & Stephens, M. 1986, *Goodness-of-fit Techniques* (Marcel Dekker Inc.)
- Dressler, A. & Shectman, S. A. 1988, *AJ*, 95, 985
- Eke, V. R., Baugh, C. M., Cole, S., Frenk, C. S., King, H. M., & Peacock, J. A. 2005, *MNRAS*, 362, 1233
- Eke, V. R., Baugh, C. M., Cole, S., Frenk, C. S., Norberg, P., Peacock, J. A., Baldry, I. K., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Dalton, G., de Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Peterson, B. A., Sutherland, W., & Taylor, K. 2004, *MNRAS*, 348, 866
- Ellingson, E., Lin, H., Yee, H. K. C., & Carlberg, R. G. 2001, *ApJ*, 547, 609
- Finoguenov, A., Connelly, J. L., Parker, L. C., Wilman, D. J., Mulchaey, J. S., Saglia, R. P., Balogh, M. L., Bower, R. G., & McGee, S. L. 2009, *ApJ*, *subm*
- Galassi, M. e. a. 2006, *GNU Scientific Library Reference Manual*, 2nd edn. (Network Theory Ltd)

- Geller, M. J. & Huchra, J. P. 1983, *ApJS*, 52, 61
- Gerke, B. F., Newman, J. A., Davis, M., Marinoni, C., Yan, R., Coil, A. L., Conroy, C., Cooper, M. C., Faber, S. M., Finkbeiner, D. P., Guhathakurta, P., Kaiser, N., Koo, D. C., Phillips, A. C., Weiner, B. J., & Willmer, C. N. A. 2005, *ApJ*, 625, 6
- Girardi, M., Fadda, D., Giuricin, G., Mardirossian, F., Mezzetti, M., & Biviano, A. 1996, *ApJ*, 457, 61
- Girardi, M., Manzato, P., Mezzetti, M., Giuricin, G., & Limboz, F. 2002, *ApJ*, 569, 720
- Gladders, M. D. & Yee, H. K. C. 2000, *AJ*, 120, 2148
- Goto, T., Yagi, M., Tanaka, M., & Okamura, S. 2004, *MNRAS*, 348, 515
- Hashimoto, Y., Oemler, A. J., Lin, H., & Tucker, D. L. 1998, *ApJ*, 499, 589
- Heald, M. 1984, *Am. J. Phys.*, 52, 254
- Hogg, D. W., Blanton, M., Strateva, I., Bahcall, N. A., Brinkmann, J., Csabai, I., Doi, M., Fukugita, M., Hennessy, G., Ivezić, Ž., Knapp, G. R., Lamb, D. Q., Lupton, R., Munn, J. A., Nichol, R., Schlegel, D. J., Schneider, D. P., & York, D. G. 2002, *AJ*, 124, 646
- Huchra, J. P. & Geller, M. J. 1982, *ApJ*, 257, 423
- Jansen, R. A., Franx, M., & Fabricant, D. 2001, *ApJ*, 551, 825
- Jeltema, T. E., Mulchaey, J. S., Lubin, L. M., Rosati, P., & Böhringer, H. 2006, *ApJ*, 649, 649
- Koehler, K. & Larntz, K. 1980, *J. Amer. Statist. Assoc.*, 75, 336
- Li, I. H. & Yee, H. K. C. 2008, *AJ*, 135, 809

- Lilliefors, H. 1967, *J. Amer. Statist. Assoc.*, 62, 399
- Lokas, E. L. & Mamon, G. A. 2003, *MNRAS*, 343, 401
- Mamon, G. A. 2007, in *Groups of Galaxies in the Nearby Universe*, ed. I. Saviane, V. D. Ivanov, & J. Borissova, 203–+
- Massey, F. 1951, *J. Amer. Statist. Assoc.*, 46, 68
- Menci, N. & Fusco-Femiano, R. 1996, *ApJ*, 472, 46
- Mulchaey, J. S., Davis, D. S., Mushotzky, R. F., & Burstein, D. 2003, *ApJS*, 145, 39
- Mulchaey, J. S. & Zabludoff, A. I. 1998, *ApJ*, 496, 73
- Nelson, L. 1998, *Journal of Quality Technology*, 30, 298
- Nolthenius, R. & White, S. D. M. 1987, *MNRAS*, 225, 505
- Osmond, J. P. F. & Ponman, T. J. 2004, *MNRAS*, 350, 1511
- Parker, L. C., Hudson, M. J., Carlberg, R. G., & Hoekstra, H. 2005, *ApJ*, 634, 806
- Roscoe, J. & Byars, J. 1971, *J. Roy. Statist. Soc.*, 66, 755
- Scott, D. 1979, *Biometrika*, 66, 605
- Stephens, M. 1974, *J. Amer. Statist. Assoc.*, 69, 730
- Tran, K.-V. H., Simard, L., Zabludoff, A. I., & Mulchaey, J. S. 2001, *ApJ*, 549, 172
- Vessereau, A. 1958, *Bulletin de l'Institut Internationale de Statistique*, 36, 87
- Wilman, D. J., Balogh, M. L., Bower, R. G., Mulchaey, J. S., Oemler, A., Carlberg, R. G., Eke, V. R., Lewis, I., Morris, S. L., & Whitaker, R. J. 2005a, *MNRAS*, 358, 88

- Wilman, D. J., Balogh, M. L., Bower, R. G., Mulchaey, J. S., Oemler, A., Carlberg, R. G., Morris, S. L., & Whitaker, R. J. 2005b, MNRAS, 358, 71
- Yahil, A. & Vidal, N. V. 1977, ApJ, 214, 347
- Yang, X., Mo, H. J., van den Bosch, F. C., & Jing, Y. P. 2005, MNRAS, 356, 1293
- Yang, X., Mo, H. J., van den Bosch, F. C., Pasquali, A., Li, C., & Barden, M. 2007, ApJ, 671, 153
- Yee, H. K. C., Morris, S. L., Lin, H., Carlberg, R. G., Hall, P. B., Sawicki, M., Patton, D. R., Wirth, G. D., Ellingson, E., & Shepherd, C. W. 2000, ApJS, 129, 475
- Zabludoff, A. I. & Mulchaey, J. S. 1998, ApJ, 496, 39

Appendices

Complementary Analysis

In private communication with John Mulchaey ¹, we obtained a sample of 9 X-ray selected galaxy groups in the local Universe ($\bar{z} = 0.0268$), observed at the *Chandra X-Ray Observatory*. The group properties are given in Table 1.

Table 1: Group Properties of the local X-ray selected Groups obtained from Mulchaey.

Chandra ID	Alt. name	RA	DEC	z
RXCJ1204.4+0154	MKW 4	181.1065	1.9010	0.0203
RXCJ1223.1+1037	NGC 4325	185.7772	10.6240	0.0255
RXCJ1324.1+1358	NGC 5129	201.0497	13.9792	0.0233
RXCJ1440.6+0328	MKW 8	220.1592	3.4765	0.0269
RXCJ1604.9+2355	AWM 4	241.2377	23.9206	0.0321
RXCJ1617.5+3458	NGC 6107	244.3635	34.9367	0.0310
RXCJ1627.6+4055	A2197	246.9173	40.9197	0.0307
RXCJ1658.0+2751	AWM 5	254.5032	27.8544	0.0347
	NGC 2563	125.1017	21.0961	0.0163

Since these are rich ($n > 20$) X-ray bright, local groups, one would expect that they are probably virialized and relaxed systems. Thus, we apply the Anderson–Darling test on these groups to determine if their velocity distributions are consistent with an underlying Gaussian distribution. The results of our analysis are given in Table 2.

¹Observatories of the Carnegie Institution, 813 Santa Barbara Street, Pasadena, California, USA

Table 2: Results of Statistical Analysis of the local X-ray selected Groups obtained from Mulchaey.

Group	n	σ_{gapper}	A^{2*}
RXCJ1204	52	626	0.387
RXCJ1223	23	362	0.115
RXCJ1324	23	346	0.442
RXCJ1440	64	459	0.436
RXCJ1604	43	463	0.296
RXCJ1617	57	595	0.161
RXCJ1627	67	605	0.683
RXCJ1658	45	564	0.164
NGC2563	63	370	0.195

At the 5% confidence level, 8 of the 9 groups are classified as having underlying Gaussian velocity distributions, with RXCJ1627 failing at the same level. This interesting result requires further analysis to determine if this group is truly dynamically complex or whether this is a false negative.

We also compute VDPs for these groups following the method outline in §1.1 of Chapter 5, using a width that is 1/3 the maximum group-centric radius, which are shown in Figure A.1. It is difficult to comment on any profile trends without knowledge of the velocity errors. The profiles for RXCJ1204, 1617, 1627 and NGC 2563 appear to be decreasing with radius, with their projected velocity dispersions covering a range $> 60 \text{ km s}^{-1}$. While RXCJ 1223 and 1440 may have VDPs consistent with flat profiles, since their projected velocity dispersions cover a much smaller range of 10-15 km s^{-1} . The groups RXCJ1224, 1605 and 1658 appear to have increasing profiles with radius, but we emphasize that this feature may not be real if the velocity errors are larger the range in dispersion covered.

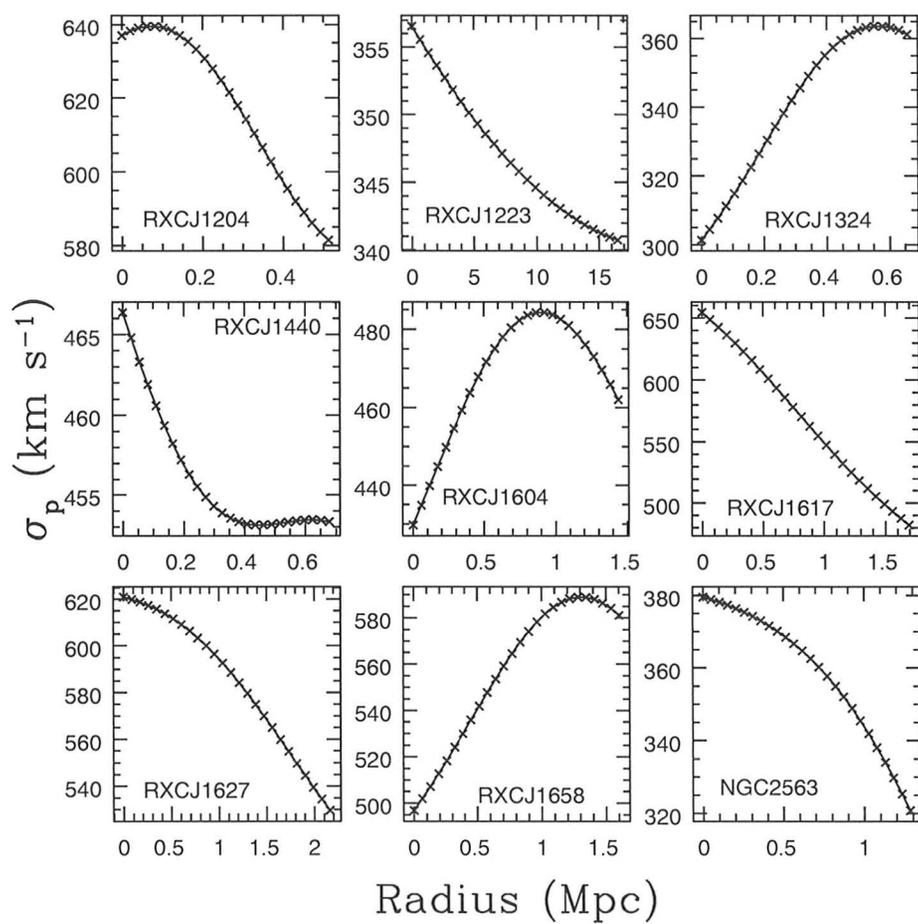


Figure 1: Velocity dispersion profiles of the 9 X-ray selected local groups observed by John Mulchaey. It should be noted that the y-axis ranges differ from plot to plot.

