LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE

MSU Is An Affirmative Action/Equal Opportunity Institution
c:\circ\datedue.pm3-p.1

Studies of Substructure in Clusters of Galaxies: A Two-Dimensional Analysis

By

Jeffrey R. Kriessler

A DISSERTATION

submitted to
Michigan State University
in partial fulfillment of the requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Department of Physics and Astronomy

1997

ABSTRACT

STUDIES OF SUBSTRUCTURE IN GALAXY CLUSTERS: A TWO DIMENSIONAL ANALYSIS

By

Jeffrey R. Kriessler

ABSTRACT

In this thesis I explore a procedure for the detection and quantification of substructure in the projected positions of galaxies in clusters. The method is first tested by application to the 56 well-studied galaxy clusters that make up the morphological sample of Dressler (1980). This method is then applied to a much larger, volume-limited sample of 119 Abell clusters originally identified of Hoessel, Gunn, & Thuan (1980). This sample includes all Abell clusters with distance class ≤ 4 and richness class ≥ 0 with $|b| > 30$. Two tests for substructure, one parametric and one nonparametric, are applied to the galaxy positions and the results are compared. The KMM algorithm partitions the data into Gaussian sub-populations and estimates their statistical significance via a hypothesis test. The DEDICA algorithm is a nonparametric technique that identifies peaks in the projected galaxy density and determines their significance with respect to the background. After a K-S test is employed on the magnitude distributions to remove background/foreground groups, 64% $\pm$ 15% of the large cluster sample is found to contain significant substructure.

Nonparametric methods of density estimation are explored and applied to the construction of contour plots and the calculation of radial number-density profiles for

each of the sample clusters. An average core radius of 150 ± 100 kpc ($H_0 = 75$ km s^{-1}) is obtained. This is however, likely to be an upper limit due to mis-specification of the cluster centers. Inside of 1 Mpc, the space density is found to vary as $\rho \propto r^{-1.9\pm0.3}$ after a correction is made for background galaxies.

The large fraction of clusters with presently-detectable substructure, as well as the shallow space-density profiles, are used to argue that rich clusters of galaxies are still in the process of formation during the present epoch and are not well described by equilibrium models. If clusters are currently accreting large amounts of material, this implies a high-density Universe, with $\Omega \gtrsim 0.4$.

To my family.

ACKNOWLEDGEMENTS

I would like to acknowledge all those who have contributed to the successful completion of this project. In particular the guidance of my advisors Tim Beers and Suzanne Hawley was very helpful. I would also like to thank the members of my committee Jim Linnemann, Ed Loh, Gerald Pollack, and Horace Smith.

I further wish to acknowledge the friends I have made here at Michigan State and all the fond memories I will take with me where ever I go. Thanks to Bill Abbott, Dave Bercik, Daniel Casavant, Jen Discenna, Normand Mousseau Dennis Kuhl, Rod Lambert, Vickie Plano, Jeff Schubert, and Steve Snyder, the years I have spent here passed quickly.

Table of Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

The human spirit is characterized, among other things, by an intense desire to explain the world around us. Of all the physical sciences, cosmology is perhaps the most ambitious, for it seeks to explain how and why the Universe we see today came into existence. From the fertile minds of scientists have sprung forth an incredible array of theories, from the Earth-centered, non-evolving universe of Ptolemy to the Inflationary Big Bang cosmologies which are popular today. The role of observational cosmology is to constrain these theories and try to find the model which best explains the real Universe. As a science, cosmology is still in its infancy. Only within this century have the theories been taken from the realm of pure conjecture about how the Universe *ought* to be, to testable theories which describe the Universe as it, at present, appears. The reasons for this transition are: 1) The existence of the mathematical frame work of General Relativity which describes the interaction of matter with space and time, 2) advances in high energy physics which have extended our understanding of the forces which act between particles at very high temperature and pressure, and 3) extension of the astronomical observations to include nearly the full range of the electromagnetic spectrum.

1.1 The Standard Model

Recent progress in observational cosmology has led to the general acceptance of the Hot Big Bang theory or, borrowing a term from high energy physics, the "standard model." The success of the standard model of the Universe rests on its ability to explain, in a simple manner, three important observations. These are: 1) the blackbody spectrum of the microwave background radiation, 2) the abundance of the light elements (D, He, and Li) in the Universe, and 3) the expansion of the Universe. Though there remain a number of problems with the standard model, the current evidence indicates that it is at least a useful approximation to the real Universe.

The standard model indicates that the Universe did not exist in its present state forever, but was created a finite time ago in an event commonly referred to as the Big Bang. At this time the Universe existed in a singular state of unimaginable temperature and pressure where current theories of physics can not be applied. From this hot, radiation-dominated state, the Universe started to expand and cool. This process continues today as observed in the redshifts of galaxies and the, now relatively cool, 2.7 K cosmic background radiation. While there are many interesting transitions which took place as the Universe cooled, this thesis is primarily concerned with events which occurred after the decoupling of the radiation and the matter, about 3×10^5 years after the Big Bang. This is commonly referred to as the "era of recombination", because only then is it possible for electrons to bind with protons to form atoms. This era is of interest because it signals the earliest possible formation time of structures such as galaxies and clusters of galaxies. Although it is believed that the density fluctuations which collapsed into these structures must have already existed before the era of recombination, they could not have *begun* to collapse due to the radiation pressure that dominated the Universe at these times. For a density fluctuation to

grow due to gravitational instability it must have a mass greater than the Jeans mass M_J. At the time of recombination the Universe becomes transparent to light and particles of matter can slow to non-relativistic speeds. This transition causes the Jeans mass to change abruptly from $M_J \approx 10^{16} M_\odot$ (about ten times the mass of the Coma cluster) to $M_J \approx 10^6 M_\odot$. The time scales for collapse of cluster-sized, Gaussian perturbations will be discussed in detail in Chapter 5.

Despite the successes of the standard model, there are a number of fundamental questions that remain unanswered. These questions include: What is the mass density of the Universe, Ω_0? Will the Universe expand forever, or collapse sometime in the future? What is the present value of the Hubble constant, H_0? How old is the Universe? What is the value of the cosmological constant, Λ? In a rapidly expanding, uniform Universe, how can structures such as galaxies, clusters of galaxies and superclusters of galaxies form? Does structure form on large scales and then fragment into smaller units, or do smaller units form first and join together to create larger objects? What is the nature of dark matter and how does it affect the evolution of the Universe?

Clusters of galaxies play an important role in addressing these questions. In a low-density Universe clusters at the present epoch are expected to be in free expansion. Therefore, they should not be accreting new material. On the other hand, in a high-Ω Universe clusters can continue to grow in the present epoch. The inflow of material into clusters should be observed as substructure (Richstone *et al.* 1992, Kauffmann & White 1993) and flattened density profiles (Crone *et al.* 1994, Jing *et al.* 1995, Crone *et al.* 1997).

1.2 Observational Properties of Clusters

One of the biggest challenges faced by the researcher in the field of galaxy clusters is defining just what constitutes a cluster, with different researchers adopting different criteria. The problem is to identify galaxies which are gravitationally bound to one another, often through the use of only their projected positions on the sky and their apparent magnitudes, parameters easily estimated from photographic surveys. Abell's (1958) solution was to visually examine the red plates of the Palomar Optical Sky Survey (POSS) and define a cluster as a region where there existed at least 30 galaxies within two magnitudes of the third-brightest galaxy within a projected radius of $1.5h^{-1}$ Mpc (one Abell radius). (Throughout this thesis $H_0 = 100h$ km s^{-1} Mpc^{-1} with h=0.75.) The centers of the clusters were determined by eye and the distance to the cluster was estimated using the apparent magnitude of the tenth-ranked galaxy. The resulting catalog of clusters (after extending it to the southern sky [Abell, Corwin, & Olowin 1989]) contains 4076 systems. Abell divided the sample into "richness classes" and "distance classes," as defined in Table 1.1. Column (1) provides the richness class R. Column (2) lists the number N of galaxies within two magnitudes of the third-ranked galaxy for each richness class. The distance class D is given in column (3). The magnitude range of the tenth-ranked galaxy m_{10} in the V band for each distance class is listed in column (4).

The Abell catalog has received a number of criticisms over the years. First, strict use of a radius within which to look for cluster members tends to favor inclusion of only those clusters which are concentrated and roughly circular, often referred to as "regular." Large, spread-out clusters or elongated clusters having a large fraction of their members outside the Abell radius would be missed. Second, since the clusters were chosen only on the basis of concentrations in the projected galaxy distribution,

Table 1.1. Abell Richness and Distance Classes

R (1)	N (2)	D (3)	m_{10} (4)
0	30-49	0	< 13.3
1	50-79	1	13.3-14.0
2	80-129	2	14.1-14.8
3	130-199	3	14.9-15.6
4	200-299	4	15.7-16.4
5	> 300	5	16.5-17.2

the possibility arises that a large number of the clusters may simply be due to the superposition of background and foreground groups. In fact a number of numerical simulations indicate that identifying clusters using Abell's method may lead to a catalog within which as much as 30% of the richness class 1 clusters are due simply to projection effects and that a similar percentage of real clusters have been missed (van Haarlem 1996). On the other hand, a study by Briel & Henry (1993) of Abell clusters with detectable X-ray emission by *ROSAT* (indicating the presence of a real potential well) found that only 10% of the Abell richness class 1 clusters are likely to have been mis-identified due to foreground/background projection.

In the past decade, the somewhat subjective nature of the Abell catalog has been addressed by the development of machine-generated catalogs (Dalton *et al.* 1994), but these do not as yet exist for the entire sky. Furthermore, advances in X-ray astronomy have led to the hope that a cluster catalog can be produced using X-ray-derived temperatures. If the hot intercluster gas is assumed to be in hydrostatic equilibrium, then the temperature of the gas will be a direct measure of the depth of the potential well, eliminating all possibility of mis-identified clusters. Although catalogs of clusters have been made from the *ROSAT* All Sky Survey (Giacconi &

Burg 1993, Ebeling *et al.* 1996), at present they are still incomplete. Furthermore, a higher-resolution survey (with 5-10 arcsec resolution) is required for easy separation of point and extended sources. Thus, despite the potential problems, the Abell catalog is still the most complete catalog of rich clusters available for the entire sky.

A typical line-of-sight velocity dispersion for a rich Abell cluster is $\sigma_r \approx 10^3$ km s^{-1}. If clusters of galaxies are assumed to be bound, the viral theorem can be used to determine the mass of the cluster. Typically this mass for rich Abell clusters is on the order of a few $\times 10^{15} M_\odot$.

One important dynamical time scale is the crossing time. The crossing time is the time it takes the average galaxy to get from one end of the cluster to the other. In convenient units it is given by:

$$t_{cross} \approx 10^9 yr \left(\frac{R}{\mathrm{Mpc}}\right)\left(\frac{10^3 \mathrm{km\ s^{-1}}}{\sigma_r}\right), \tag{1.1}$$

where R is the radius of the cluster and σ_r is the line-of-sight velocity dispersion. This provides a lower limit for the time it takes substructure to be erased. For rich clusters this is about a billion years, or one-tenth the age of the Universe. Analytical work suggests that the smallest groups on radial orbits will be disrupted by tidal forces in a single crossing time, while a merger between two equal-sized clusters may take as long as four crossing times to be erased (González-Casado *et al.* 1994). On the other hand, numerical simulations span the full spectrum of possible relaxation times, from a single crossing time to several Hubble times, depending on the initial conditions assumed (West, Oemler, & Dekel 1988; Cavaliere *et al.* 1992; Nakamura *et al.* 1995). Ultimately, it appears as though observations of clusters will be necessary to constrain these initial conditions.

1.3 Previous Studies

The modern study of clusters of galaxies was initiated with Zwicky's (1933) study of the Coma cluster (A1656). Using positions projected on the plane of the sky and line of sight velocities obtained from redshifts of spectral lines in the galaxies, he concluded that the amount of matter needed to keep the Coma cluster from flying apart on a time scale of a billion years was many times larger than the matter visible in the galaxies. This was the first indication of the existence of large amounts of dark matter in the Universe. In a follow-up study, Zwicky (1937) concluded that the distribution of bright galaxies was very similar to the distribution of mass density in an isothermal gas sphere. For the following five decades the Coma Cluster has been considered the prototype of a relaxed, rich galaxy cluster.

With advances in computer speed and availability over the past 20 years has come enormous strides in statistical techniques which can be used to analyze data in new ways. With this and the advent of X-ray astronomy, the idea that clusters of galaxies could be described as relaxed systems in isothermal equilibrium has been challenged. The evidence used to argue against equilibrium cluster models includes: a) "clumpy" distributions of galaxies seen in projection on the sky, b) apparent structure in the distribution of radial velocities for cluster members, and c) multiple centers of X-ray emission in the cluster, and other complexities in the X-ray-derived temperature profiles, suggestive of ongoing subcluster collisions.

Studies using solely the projected positions of galaxies in clusters have concluded that between 20% and 80% of clusters have statistically-significant substructure. Geller & Beers (1982, hereafter GB) made contour maps of the projected galaxy density for 65 clusters with data from Dressler (1976) and Dressler (1980). These authors concluded that 40% of the clusters in the combined Dressler samples have

substructure based on multiple peaks in the contour maps. Baier (1983), on the basis of secondary peaks in radial number-density distributions for some 100 clusters, concluded that as much as 80% of the clusters in his sample had substructure. More recent investigations are discussed in Baier *et al.* (1996). The image analysis techniques of West & Bothun (1990) led them to conclude that some 30% of the Dressler sample has substructure. Rhee, van Haarlem & Katgert (1991) applied six tests for substructure to the projected positions of galaxies in 104 Abell clusters obtained from digital scans of copies of the Palomar Sky Survey plates, and found that 26% had significant substructure. Salvador-Solé, Sanromà & González-Casdado (1993, hereafter SSG) looked for deviations in the density profiles of 15 clusters (after applying a redshift filter to remove obvious foreground/background galaxies), and found that 50% of their sample showed evidence of substructure.

Dressler & Shectman (1988) obtained the first sample of galaxies in clusters with a sufficient number of measured redshifts to include velocity information in the search for substructure, and concluded that 30% to 40% of their sample (of 15 clusters) exhibited deviations in the local *vs.* global kinematic properties, consistent with the existence of dynamically-significant substructure. Bird (1993) applied a number of statistical tests using both spatial and velocity data to demonstrate that between 30% and 80% of clusters could have substructure, depending on the test employed. Escalera *et al.* (1994) applied the wavelet analysis technique to projected galaxy positions and velocities in 16 clusters and found that only three clusters could be classified as unimodal. The simple kinematic test of Dressler & Shectman (the Δ-test) was applied to a sample of 73 clusters in the ESO Nearby Abell Cluster Survey (ENACS) by den Hartog (1995), who found that 50% showed evidence of substructure.

Jones & Forman (1992) found that of the 208 clusters observed by the EINSTEIN satellite with X-ray emission bright enough to classify, some 22% showed clear sub-

structure. Mohr, Frabricant & Geller (1992) examined X-ray surface brightness moments of 40 EINSTEIN cluster observations, and found that 68% showed evidence of substructure. Buote & Tsai (1996) used a power-ratio technique (essentially a ratio of higher-order moments of a two-dimensional potential to the monopole moment) to examine 59 clusters with X-ray maps available from *ROSAT* which had substructure obvious to the eye, in order to specify the dynamical states of the clusters in their sample. These authors conclude that most clusters have some level of substructure and that the evolutionary state of the cluster can be specified by its influence on the gravitational potential.

1.4 Goals of the Thesis

Most of the above studies have used relatively small numbers of clusters which suffer more or less from selection effects. Although Rhee *et al.* applied a battery of substructure tests to their sample clusters, they concluded that only 26% of them had evidence for substructure. Furthermore, no single test resulted in more than 10% of the sample being classified as containing significant substructure. Their conclusion is clearly at odds with the growing evidence that suggests most clusters do indeed have substructure. It may well be that the statistical tests applied by Rhee were simply not sensitive enough to the substructure which they were designed to detect. On the other hand, Jones & Forman studied a large sample of clusters with EINSTEIN pointed observations with which they went as far as classifying the morphologies of the clusters based on the appearance of the images. While such a catalog is potentially very useful, especially for comparison to optical maps such as the ones presented in this thesis, this catalog is not readily available. Also, in his thesis, Beers (1983) argued that estimates of the true fraction of X-ray clusters which exhibit substructure obtained with the EINSTEIN survey are only lower limits due to selection biases

in the observations. In particular, because the unvignetted field of view is only 40 arcminutes, subclusters with separations greater than about 20 arcminutes from the cluster center will be missed.

Therefore, armed with new and potentially very powerful techniques for the detection of substructure in clusters, the time is ripe to re-examine the question of substructure in the projected galaxy distributions of a statistically-complete sample of nearby Abell clusters. The goal of the thesis is to identify a subset of Abell clusters which are likely to contain dynamically-significant substructure. The fraction of clusters with substructure and the radial density profiles of clusters are used to place constrains on the cosmological density parameter, Ω_0.

1.5 Chapter Overview

In Chapter 2 the motivation and selection criteria for the cluster sample is explained. The use of the Automated Plate Scanner (APS) catalogs of star and galaxy positions is discussed. In particular, the accuracy and completeness of the galaxy catalogs is addressed. The background contamination within an Abell radius of each cluster is estimated to be between 10% and 30%.

Chapter 3 presents the basic concepts behind nonparametric density estimation. The motivation behind the use of the adaptive-kernel technique is explained. These concepts are applied in the construction of contour maps for each of the sample clusters. These maps can be used to identify peaks in the projected galaxy positions for detailed comparison with X-ray surface brightness maps.

The tests employed for the detection of substructure are presented in Chapter 4. Two tests are applied to the sample clusters and the results discussed and compared. Comparisons are made with other techniques and the advantages and disadvantages

are explored.

In Chapter 5, the theory behind the estimation of Ω_0 from the fraction of clusters with detectable substructure is reviewed. The results of Chapter 4 are used to argue that if the density perturbations are Gaussian at the time of recombination and if substructure is erased on the order of four crossing times, Ω_0 is likely to be greater than 0.4. This is about twice the amount of matter currently inferred from the dynamics of galaxy clusters.

Chapter 6 presents non-parametric density profiles for the sample clusters. The question of the existence of constant-density cores in clusters is addressed. The steepness of the radial-density profiles is compared to numerical simulations which suggest a high-Ω Universe.

Conclusions and suggestions for future work are presented in Chapter 7.

Chapter 2

DATASET

2.1 The Cluster Samples

In this thesis two samples of clusters will be examined for the presence of substructure in the projected galaxy positions. These are the 56 clusters included in Dressler's morphological study (1980) and the 119 clusters in the sample identified by Hoessel, Gunn, & Thuan (1980, hereafter HGT). There is an overlap of 25 clusters between the two samples which will be used to compare the APS (Automated Plate Scanner project at Minnesota) data with that obtained by Dressler. The combined samples contain a total of 150 clusters.

2.1.1 Dressler's Data

The sample of clusters selected by Dressler includes clusters with $z \leq 0.06$ ($cz \leq 18000$ km s^{-1}), and $N \geq 50$ with magnitude $m_V \leq 16.5$ contained in an area of few square degrees on the sky. Unlike Abell's definition of a cluster, the area definition was left purposefully vague in order to avoid selecting only circularly-symmetric clusters. This nevertheless includes 38 Abell clusters from the northern catalog, though some (*e.g.* A14) have richness class 0 due to the different area definitions. In addition, 18 southern clusters (those with prefix DC and Centaurus) satisfying these criteria

were identified in the southern sky using the plate copies of the ESO Quick Blue Sky Survey. Though many of the these clusters have since been given Abell numbers in the expanded ACO (Abell, Corwin, & Olowin 1989) catalog, the older designation will be retained here for easy comparison with previous work. The cluster redshift was obtained by Dressler from the literature when available. When not available he obtained redshifts for at least two cluster members.

Dressler (1980) lists positions, estimated magnitudes rounded to the nearest magnitude, bulge sizes, ellipticities, and morphological type for each galaxy in the survey. The fact that this information was published and that these are among the closest clusters has made members of the Dressler sample some of the most well-studied clusters in the sky, hence a good test bed for the evaluation of new statistical techniques.

The background in the Dressler sample was estimated by taking an additional 15 plates at random areas of the sky and repeating the same procedure of galaxy identification carried out for the program plates. A median value of 8 galaxies $\deg^{-2}$ is quoted by Dressler or 0.0022 galaxies arcmin^{-2}. Follow-up studies which included the gathering of redshift data confirmed that in most cases this is a good estimate (Dressler & Shectman 1988).

2.1.2 HGT Sample

In addition to a re-examination of the question of substructure in the 56 clusters of Dressler's morphological sample, this study also includes the 119 Abell clusters in the sample of HGT (1980). This sample was an attempt to be a volume-limited sample, in that it consists of all northern clusters in the Abell catalog with distance class less than or equal to 4 and richness class greater than 0, with galactic latitude $|b| \geq 30$ (107 clusters). In addition it contains 12 clusters with richness class 0 and distance class 3 or less at high galactic latitude.

Although the APS project offers the unique possibility of testing *all* 2714 northern Abell clusters, the reason for choosing to study these 119 clusters first is that these clusters are the richest and closest clusters to us. As such, they have attracted the most attention from the astronomical community and are likely to continue to do so in the coming years. They are among the most likely targets for new redshift surveys. At least 54 of these clusters have detectable X-ray emission indicating that they are real systems and not simply due to the projection of physically different foreground and background groups. (At the time of this writing 36 have pointed *ROSAT* observations available from the public archive.) Each cluster has a measured redshift, so that their distances do not need to be approximated. Lastly, avoiding clusters close to the plane of the Galaxy helps to minimize the effect of obscuring dust which would need to be estimated and corrected for in the calculation of a limiting magnitude, as well as helping to keep down the number of misclassified stars in the sample.

Table 2.1 is a listing of cluster parameters for the HGT sample. Column (1) lists the cluster name. Columns (2) and (3) list the center of the cluster as specified by Abell in 1950 coordinates. The galactic latitude is given in column (4). Distance classes and richness classes are given in columns (7) and (8), respectively. Column (7) lists the Bautz-Morgan (BM) type (Bautz & Morgan 1970). The revised Rood-Sastry (RS) type given by Struble & Rood (1982) is listed in column (8). Column (9) lists the cluster redshift from Struble & Rood (1991).

TABLE 2.1. HGT Cluster Parameters

Cluster (1)	RA (1950) (2)	DEC (1950) (3)	b (4)	D (5)	R (6)	BM (7)	RS (8)	z (9)
A 21	00 07.9	+28 22	−33.74	4	1	I	B	0.0948
A 76	00 07.2	+06 30	−55.97	3	0	II–III	L	0.0377
A 85	00 09.1	−09 38	−72.08	4	1	I	cD	0.0556
A 88	00 00.4	−26 20	−87.80	3	1	III		0.1086
A 104	00 47.1	+24 15	−38.35	4	1	II–III	F	0.0822
A 119	00 53.8	−01 32	−64.11	3	1	II–III	C	0.0446
A 121	00 55.0	−07 17	−69.83	4	1	III	I	0.1048
A 147	01 05.6	+01 55	−60.42	3	0	III	I	0.0441
A 151	01 06.4	−15 41	−77.62	3	1	II	cD	0.0526
A 154	01 08.3	+17 24	−44.95	3	1	II	B	0.0612
A 166	01 12.1	−16 33	−77.90	4	1	III	F	0.1156
A 168	01 12.6	−00 02	−62.05	3	2	II–III	I	0.0457
A 189	01 21.1	+01 24	−60.19	4	1	III	I	0.0349
A 193	01 22.5	+08 27	−53.25	4	1	II	cD	0.0478
A 194	01 23.0	−01 46	−63.10	1	0	II	L	0.0178
A 225	01 36.2	+18 38	−42.56	4	1	II–III	I	0.0692
A 246	01 42.1	+05 34	−54.62	4	1	II–III	F	0.0753
A 274	01 52.2	−06 32	−64.29	4	3	III	I	0.1289
A 277	01 53.3	−07 38	−65.04	3	1	III	I	0.0947
A 389	02 49.1	−25 07	−63.04	4	2	II	F	0.1160
A 399	02 55.2	+12 49	−39.47	3	1	I–II	cD	0.0725
A 400	02 55.0	+05 50	−44.93	1	1	II–III	I	0.0231
A 401	02 56.2	+13 23	−38.87	3	2	I	cD	0.0752
A 415	03 04.4	−12 15	−54.89	4	1	II	cD	0.0788
A 496	04 31.3	−13 22	−36.49	3	1	I	cD	0.0326
A 500	04 36.8	−22 12	−38.49	4	1	III	I	0.0666
A 514	04 45.5	−20 32	−36.02	3	1	II–III	F	0.0697
A 634	08 00.5	+58 12	+33.64	3	0	III	F	0.0266
A 671	08 25.4	+30 36	+33.11	3	0	II–III	C	0.0497
A 779	09 16.8	+33 59	+44.41	1	0	I–II	cD	0.0201
A 787	09 23.5	+74 38	+36.20	4	2	II	F	0.1355
A 957	10 11.4	−00 40	+42.88	4	1	I–II	L	0.0437
A 978	10 18.0	−06 17	+40.35	3	1	II	F	0.0527
A 993	10 19.4	−04 43	+41.69	3	0	III	I	0.0530
A 1020	10 25.2	+10 40	+52.33	4	1	II–III	I	0.0650
A 1035	10 29.2	+40 29	+58.46	3	2	II–III	F	0.0799
A 1126	10 51.3	+17 08	+60.98	4	1	I–II	B	0.0828
A 1139	10 55.5	+01 47	+52.66	3	0	III	I	0.0376
A 1185	11 08.1	+28 57	+67.76	2	1	II	C	0.0349
A 1187	11 08.9	+39 51	+65.85	3	1	III	I	0.0791
A 1213	11 13.8	+29 33	+69.01	2	1	III	C	0.0484
A 1216	11 15.2	−04 12	+51.14	4	1	III	F	0.0524
A 1228	11 18.8	+34 37	+69.44	1	1	II–III	F	0.0344
A 1238	11 20.4	+01 23	+56.42	4	1	II	C	0.0716
A 1254	11 23.8	+71 22	+44.46	3	1	III	I	0.0628
A 1257	11 23.4	+35 37	+70.05	3	0	III	F	0.0339
A 1291	11 29.3	+56 19	+57.77	3	1	III	F	0.0586

TABLE 2.1. (continued)

Cluster (1)	RA (1950) (2)	DEC (1950) (3)	b (4)	D (5)	R (6)	BM (7)	RS (8)	z (9)
A 1318	11 33.7	+55 15	+59.00	3	1	II	C	0.0189
A 1364	11 41.1	−01 30	+56.80	4	1	III	C	0.1070
A 1365	11 41.8	+31 11	+74.88	4	1	III	F	0.0763
A 1367	11 41.9	+20 07	+73.04	1	2	II–III	F	0.0205
A 1377	11 44.3	+56 01	+59.11	3	1	III	B	0.0509
A 1382	11 45.6	+71 43	+44.82	4	1	II	cD	0.1046
A 1383	11 45.5	+54 54	+60.17	4	1	III	I	0.0598
A 1399	11 48.6	−02 50	+56.45	4	2	III	I	0.0913
A 1412	11 53.1	+73 45	+43.07	4	2	III	C	0.0839
A 1436	11 57.9	+56 32	+59.47	3	1	III	I	0.0646
A 1468	12 03.1	+51 42	+64.20	4	1	I	C	0.0853
A 1474	12 05.4	+15 14	+74.17	4	1	III	I	0.0778
A 1496	12 10.9	+59 33	+57.18	4	1	III	I	0.0961
A 1541	12 24.9	+09 07	+70.86	4	1	I–II	B	0.0892
A 1644	12 54.6	−17 06	+45.48	4	1	II	cD	0.0456
A 1651	12 56.8	−03 56	+58.61	4	1	I–II	cD	0.0842
A 1656	12 57.4	+28 15	+87.96	1	2	II	B	0.0230
A 1691	13 09.1	+39 29	+77.22	3	1	II	cD	0.0722
A 1749	13 27.3	+37 53	+76.79	4	1	II	cD	0.0562
A 1767	13 34.2	+59 29	+56.99	4	1	II	cD	0.0712
A 1773	13 39.6	+02 30	+62.31	3	1	III	F	0.0776
A 1775	13 39.6	+26 37	+78.70	4	2	I	B	0.0718
A 1793	13 46.1	+32 32	+76.63	4	1	III	I	0.0849
A 1795	13 46.7	+26 51	+77.16	4	2	I	cD	0.0631
A 1809	13 50.8	+05 25	+63.55	4	1	II	cD	0.0788
A 1831	13 56.9	+28 14	+74.97	3	1	III	F	0.0749
A 1837	13 59.1	−10 56	+48.08	4	1	I–II	cD	0.0376
A 1904	14 20.3	+48 48	+62.29	3	2	II–III	C	0.0719
A 1913	14 24.5	+16 54	+65.59	4	1	III	I	0.0533
A 1927	14 28.8	+25 54	+67.68	4	1	I–II	cD	0.0740
A 1983	14 50.4	+16 57	+60.11	3	1	III	F	0.0458
A 1991	14 52.2	+18 51	+60.51	3	1	I	F	0.0589
A 1999	14 52.6	+54 32	+54.77	4	1	II–III	I	0.1032
A 2005	14 56.6	+28 01	+61.84	4	2	III	B	0.1251
A 2022	15 02.2	+28 38	+60.66	3	1	III	F	0.0565
A 2028	15 07.1	+07 43	+51.88	4	1	II–III	I	0.0772
A 2029	15 08.5	+05 57	+50.55	4	2	I	cD	0.0777
A 2040	15 10.3	+07 37	+51.18	4	1	III	C	0.0456
A 2048	15 12.8	+04 35	+48.86	4	1	III	C	0.0945
A 2052	15 14.3	+07 12	+50.12	3	0	I–II	cD	0.0351
A 2061	15 19.2	+30 50	+57.17	4	1	III	L	0.0782
A 2063	15 20.6	+08 49	+49.72	3	1	II	cD	0.0337
A 2065	15 20.6	+27 54	+56.56	3	2	III	C	0.0722
A 2067	15 21.2	+31 06	+56.76	4	1	III	cD	0.0726
A 2079	15 26.0	+29 03	+55.53	3	1	II–III	cD	0.0657
A 2089	15 30.6	+28 12	+54.43	4	1	iI	cD	0.0743
A 2092	15 31.3	+31 20	+54.61	4	1	II–III	I	0.0669

TABLE 2.1. (continued)

Cluster (1)	RA (1950) (2)	DEC (1950) (3)	b (4)	D (5)	R (6)	BM (7)	RS (8)	z (9)
A 2107	15 37.6	+21 56	+51.50	4	1	I	cD	0.0421
A 2124	15 43.1	+36 14	+52.31	3	1	I	cD	0.0671
A 2142	15 56.2	+27 22	+48.70	4	2	II	B	0.0911
A 2147	16 00.0	+16 03	+44.49	1	1	III	F	0.0377
A 2151	16 03.0	+17 53	+44.53	1	2	III	F	0.0360
A 2152	16 03.1	+16 35	+44.02	1	1	II	F	0.0444
A 2162	16 10.5	+29 40	+46.04	1	0	II–III	I	0.0318
A 2175	16 18.4	+30 02	+44.42	4	1	II	cD	0.0978
A 2197	16 26.5	+41 01	+43.81	1	1	III	L	0.0303
A 2199	16 26.9	+39 38	+43.71	1	2	I	cD	0.0312
A 2255	17 12.2	+64 09	+34.95	3	2	II–III	C	0.0747
A 2256	17 06.6	+78 47	+31.74	3	2	II–III	B	0.0550
A 2328	20 45.4	−18 00	−33.56	4	2	I	I	0.1470
A 2347	21 26.7	−22 26	−44.17	4	1	III	I	0.1196
A 2382	21 49.3	−15 53	−46.94	4	1	II–III	L	0.0648
A 2384	21 49.5	−19 47	−48.40	4	1	II–III	F	0.0943
A 2399	21 54.9	−08 02	−44.57	3	1	III	I	0.0587
A 2410	21 59.4	−10 09	−46.58	4	1	III	I	0.0806
A 2457	22 33.3	+01 13	−46.60	4	1	I–II	C	0.0597
A 2634	23 35.8	+26 46	−33.06	1	1	II	cD	0.0315
A 2657	23 42.3	+08 53	−50.29	3	1	II	F	0.0414
A 2666	23 48.4	+26 53	−33.80	1	0	I	cD	0.0273
A 2670	23 51.6	−10 41	−68.52	4	3	I–II	cD	0.0774
A 2675	23 53.0	+11 10	−49.12	4	1	II	F	0.0726
A 2700	00 01.3	+01 48	−58.63	4	1	II	cD	0.0978

2.2 Digital Sky Surveys

In an effort to make the large photographic surveys (such as POSS) conducted over the last half century or so more easily accessible and therefore more useful, a number of projects have been carried out, or are currently underway, to transfer them to an electronic form. These include the APM at Cambridge, APS at Minnesota, COSMOS at Edinburgh, PDS at STScI and PPM at the U.S. Naval Observatory (see Lasker 1995 for a review). The existence of these databases will greatly facilitate computerized procedures for catalog construction, and the making of finding charts. It also offers the possibility of doing science such as the large-scale distribution of stars in the Milky Way Galaxy (Larson 1996) or, as in this case, the distribution of galaxies within a large number of galaxy clusters.

While there are a number of such projects, the only one suitable for the present work is the APS catalog of POSS I at Minnesota. The APS survey offers positions and magnitudes as well as a classification of objects as stars or galaxies, unlike the PDS scans at STScI which only provide images. And, although only partially completed, the APS survey is available to the public, unlike the APM survey which is only available for collaborative use.

2.3 The Minnesota Automated Plate Scanner

The Automated Plate Scanner (APS) project at the University of Minnesota (Pennington *et al* 1993) uses a "flying spot" laser scanner. The light from a helium-neon laser is sent through a rapidly rotating, eight-sided prism. A lens and beam-splitting prism are used to form three 12 micron spots simultaneously on the E (red) and O (blue) plates of POSS I as well as a reticle for positional reference of the other two spots. The spots are detected using silicon photodetectors.

Of the three available modes of operation, the POSS I plates have been scanned in the threshold densitometry mode. That is, pixel information is saved only when the density exceeds 65% above the median background value, which corresponds roughly to 23.5 B mag arcsec^{-2}. The position of the ingress (when the density threshold has been met), the position of egress (when the density threshold is no longer met) and the pixel data in-between are recorded. Positions are measured at a resolution of 0.366 microns, which corresponds to 0.0245 arcsecs on the sky, with a repeatability error of 0.6 arcsecs.

The magnitudes on each plate are estimated from the image diameter size. The integrated isophotal magnitudes of galaxies are listed as being accurate to 0.5 magnitudes. However, this error is just the reproducibility error in the measurement on any given plate. Due to variation in the emulsion of various plates the actual error in magnitudes is higher. Comparison of galaxies in the plate overlap regions indicate the the plate to plate variation is in some cases as high as 1.0 magnitude on the blue plates. This indicates the need for better magnitude calibration if studies involving luminosity functions or comparisons between objects on different plates is to be done. This does not however, affect the current study, except that the sampling depth in each cluster is not really a constant but may vary slightly from cluster to cluster. In those clusters which include data from more than one plate, the magnitudes were calibrated using the mean value of all galaxies in the overlap region.

2.4 Neural Network Star/Galaxy Classification

In general, a single POSS I plate will produce on the order of 250,000 detected images. In order to classify this many objects, a fast and fully-automated procedure is needed. The solution of Odewahn *et al.* (1992) was to employ a neural network.

Neural networks are a family of artificial intelligence algorithms which are capable of performing pattern recognition. Typically, a neural network is trained using a sample of pre-classified objects. To perform the star/galaxy separation 14 parameters of the image are input into the neural network. These are: diameter, ellipticity, average transmission, central transmission, ratio of ellipse area to area from the pixel count, the logarithm of the area from pixel count, first moment of the image, rms error of ellipse fit to transit endpoints, the Y centroid error, and the five image gradients defined as:

$$G_{ij} = \frac{T_j - T_i}{r_i - r_j}, \tag{2.1}$$

where T_i is the median transmission value in an elliptical annulus and semimajor axis r_i.

Although the performance of a neural network can be judged by viewing the output, it is not generally clear by what criteria the classification is being made. For instance, the original training set at APS contained stars and galaxies, but did not contain double stars. As a result, any suitably-elongated image was classified as a galaxy with high probability. Although this problem has be identified and remedied with a training set of double stars and the plates are being re-processed, some of the data used in this thesis were classified before such improvements were made.

2.4.1 Contamination from Stars

The catalog of galaxy positions and magnitudes used in this thesis includes objects classified as galaxies by the neural net with a probability of 0.85 or higher. Each galaxy assigned a magnitude brighter than 19.0 on the blue plate was examined using the Digitized Sky Survey (DSS) done with the PDS machines at STScI. (Although there are plans to place the APS images online, at the time of writing these are still

only available for a small fraction of the online catalogs.) Objects which were actually stars or binary stars were deleted from the catalogs. From this, it was noticed that several plates had far more contamination than the expected 10-20%. A2666, on plate mlp 779, had the most misclassified stars with a contamination rate of 42%. The reason for such a large contamination is probably due to the fact that A2666 is a nearby cluster and therefore the 1.5 Mpc region covers a large area of sky, which includes relatively more optical binary and bright stars. Furthermore, A2666, at a galactic latitude of -33, is relatively close to the plane of the Galaxy. This was a major reason to limit this study to clusters with $|b| \geq 30$. In this and other clusters that showed a contamination rate greater than 10%, every galaxy was examined. In all 16,000 galaxies were examined, or about half of the catalog used in this thesis. Above 19.0 magnitude, an overall contamination rate of 12% was observed. The breakdown with magnitude is as follows: $m \leq 16.5$, 23%; $16.5 \leq m < 18.0$, 19%; $19.0 \leq m < 18.0$, 13%; and $m > 19.0$, 8%. It is somewhat surprising that the greatest contamination level is for galaxies brighter than 16.5 magnitude. In general this appears to be due to bright, saturated stars being classified as galaxies. Also, there is a sharp drop off in the number of stars deleted at magnitudes fainter than 19.0. This is likely to be a result of the greater difficulty encountered in the determination of which objects are stars and which are galaxies as the plate limit is approached, and should not be used to estimate contamination levels at these magnitudes. However, if it is assumed that the major source of misclassified galaxies (those which are actually double stars) remains constant with magnitude and that the overall contamination level is on the order of 15%, then the contamination left in the sample after deleting the probable stars should be near the 8% level. This is considered acceptable since misclassified stars should appear randomly (with a constant density) over the cluster and as such are very unlikely to be identified as coherent substructure.

There are also a number of objects in the APS catalog which are classified as galaxies, usually with with very high probability, with a magnitude of 8.00. In these cases the neural network has become confused by bright, and therefore large, stars or galaxies. When such an object was encountered and determined to be a galaxy from examination of the DSS image, the Third Reference Catalog of Bright Galaxies (de Vaucouleur *et al.* 1991, hereafter 3RC) was searched for a nearby galaxy. The photographic magnitude listed there (m_B) was transformed to the m_O magnitude of the APS using the average offset calculated from other galaxies on the same plate which also had m_B listed in 3RC. An average 1.0 magnitude needed to be added to m_B to obtain m_O. In 5 cases, an entry was not found in 3RC and these galaxies were assigned a magnitude of 16.5, the limit to which 3RC attempted to be complete plus one magnitude.

Lastly, one cluster, A2079, had to be deleted from the sample due to a bright star which caused a hole in the galaxy catalog. Its map is still given for completeness but it is not used in any calculations.

2.4.2 Completeness

Because the goal was to make the analysis of each cluster as consistent as possible, the magnitude limit for each cluster was set to an absolute magnitude $M_O = -16.2 + 5 \log h$, as opposed to simply a fixed apparent magnitude. The value of this limit was obtained by finding the absolute magnitude of the faintest galaxies on the plate which contained the furthest cluster. That is, the cluster A2328 at $z = 0.147$ has a magnitude limit of $m_O = 22.2$. With the magnitude limit so defined, each cluster is sampled to the same depth in the luminosity function, or about three magnitudes below the knee in the Schechter luminosity function.

The luminosity function for galaxies is usually fit to an analytic function due

to Schechter (1976). The number of galaxies with luminosity in a range of $L + dL$ according to the Schechter function is

$$n(L)dL = N^*(L/L^*)^{-\alpha}\exp(-L/L^*)d(L/L^*) \tag{2.2}$$

where L^* is a characteristic luminosity and is often referred to as the "knee" in the Luminosity function. Likewise the integrated luminosity function is

$$N(L) = n^*\Gamma(1-\alpha, L/L^*) \tag{2.3}$$

where $\Gamma(a,x)$ is the incomplete gamma function.

Two effects can cause cluster-to-cluster variation. The first is that the sensitivity of the emulsion on the photographic plates may vary either across a single plate or between plates and thereby cause variations in the measured magnitudes of the galaxies. As already mentioned, this could be as much as one magnitude, as measured by the difference in magnitudes of galaxies in the plate overlap regions. The second is variations in sample completeness, which is a function of the apparent magnitude cutoff used. In general there will be proportionately fewer galaxies detected at a magnitude of 22 than at a magnitude of 19. There are two reasons for this effect. First, galaxies near the plate limit may fall below the detection limit due to random fluctuations in the background or emulsion sensitivity. Second, small compact galaxies, such as dwarf ellipticals, are more likely to be misclassified as stars at fainter magnitudes.

Odewahn *et al.* (1993) have examined the completeness of the APS data by comparing the APS output for an area centered on the north galactic pole region with that of other studies of this region. From this it was concluded the APS data is 95% complete at an $m_O = 19.5$, 90% complete for $19.5 < m_O < 20.0$, and 80% $20 < m_O < 21$. Thus more distant clusters with a magnitude cutoff of 21 can have the galaxy counts depleted by as much as 20% as compared to a nearby cluster. One way

to avoid this completeness problem would be to use a brighter absolute magnitude for the cutoff. However, if this is done nearby clusters will have their magnitude cutoff raised as well. For some sparsely populated clusters, such as A194 or A634, this would result in too few member galaxies to carry out the substructure tests with any reliability. With this cutoff the cluster with the smallest number of galaxies is A634 with 71 galaxies. The Monte Carlo tests discussed in chapter 4 indicate that, with this number of galaxies, only very wide separations between the groups are likely to returned as significant most of the time. In retrospect, clusters with redshifts greater than $z = 0.1$ should probably not have been included.

2.5 Estimation of Background Contamination

Not all of the galaxies which appear in the cluster maps will be gravitationally bound to the clusters. The presence of some galaxies will be due to the projection of background or foreground galaxies onto the plane of the sky. (For convenience both background and foreground galaxies will be lumped together under the term background.) An estimate of an assumed constant-density background can be obtained for each cluster from the adaptive-kernel procedure discussed in Chapter 3. The background density can be taken as the density at the point with the largest bandwidth factor λ_i (see section 3.5). Defined as such, this density corresponds to the lowest density region (but not necessarily the lowest density) in each map. Although other definitions of the background density are possible, this one has the advantage of being based on a density measurement for each cluster. This is quite different from estimates that count the number of galaxies in random areas of the sky and then assume that the background rate is constant for all clusters. It is possible that this procedure overestimates the background since we would expect that even at the lowest densities in the $1.5h^{-1}$ Mpc region, some of the density there will be due to cluster members.

On the other hand, there is really no reason to assume that the background density is actually constant. Maps of the clusters may contain background groups and other clusters, providing a clumpy background. Some examples include A85, which contains the more distant cluster A89 (as well as the cluster A87 which is at the same distance as A85 [den Hartog 1995]), and A1999, which contains A2000. In such cases, contamination could be greater. Unidentified foreground clusters are not expected to be as big of a potential problem because they will appear larger and contribute a nearly constant density across the cluster maps, which should be well approximated by this method of background determination. It is this background estimate that the significance of the subclusters is estimated against in the program DEDICA discussed in Chapter 4.

The background density estimates are listed in Table 2.2 for the Dressler clusters and Table 2.3 for the HGT sample clusters. The cluster is listed in column (1). Column (2) provides the total number of galaxies in each cluster. The number of expected background galaxies is given in column (3). Column (4) is the percentage of the total number. Column (5) is the density of the estimated background in galaxies arcmin^{-2}. For the Dressler clusters, the estimated background varies from half that estimated by Dressler to nearly five times as much for A2256.

TABLE 2.2. Estimated Background – Dressler sample

Cluster (1)	N_{tot} (2)	N_{back} (3)	$\%N_{tot}$ (4)	σ_{back} (5)
A 0014	79	17	21	0.0035
A 0076	72	10	14	0.0028
A 0119	116	15	13	0.0040
A 0151	105	15	14	0.0019
A 0154	79	17	21	0.0047
A 0168	106	6	6	0.0016
A 0194	75	18	24	0.0022
A 0376	119	25	21	0.0080
A 0400	92	11	12	0.0011
A 0496	81	15	18	0.0041
A 0539	99	17	17	0.0021
A 0548	234	24	10	0.0030
A 0592	61	12	19	0.0014
A 0754	150	26	18	0.0033
A 0838	62	38	61	0.0046
A 0957	82	16	19	0.0020
A 0978	62	12	20	0.0015
A 0979	86	18	21	0.0022
A 0993	91	27	30	0.0034
A 1069	47	8	18	0.0010
A 1139	63	10	16	0.0013
A 1142	59	8	13	0.0009
A 1185	44	15	33	0.0030
A 1377	52	12	24	0.0029
A 1631	90	23	25	0.0028
A 1644	145	19	13	0.0024
A 1656	245	22	9	0.0028
A 1736	166	18	11	0.0022
A 1913	86	26	30	0.0035
A 1983	123	20	16	0.0025
A 1991	53	9	18	0.0013
A 2040	108	20	19	0.0028
A 2063	110	13	11	0.0017
A 2151	152	13	8	0.0017
A 2256	83	25	30	0.0100
A 2589	72	20	27	0.0055
A 2634	132	27	21	0.0064
A 2657	82	17	21	0.0048
DC 0003-50	79	11	14	0.0014
DC 0103-47	53	12	22	0.0014
DC 0107-46	55	10	18	0.0013
DC 0247-31	48	15	32	0.0019
DC 0317-54	65	17	26	0.0021
DC 0326-53	161	37	23	0.0045
DC 0329-52	190	12	6	0.0014

TABLE 2.2. (continued)

Cluster (1)	N_{tot} (2)	N_{back} (3)	$\%N_{tot}$ (4)	σ_{back} (5)
DC 0410-62	64	24	37	0.0029
DC 0428-53	131	21	16	0.0025
DC 0559-40	112	10	9	0.0013
DC 0608-33	122	6	5	0.0008
DC 0622-64	98	12	13	0.0015
DC 1842-63	55	15	27	0.0018
DC 2048-52	216	42	19	0.0052
DC 2103-39	108	12	11	0.0015
DC 2345-28	95	30	32	0.0037
DC 2349-28	68	24	35	0.0030
Centaurus	73	18	24	0.0022

TABLE 2.3. Estimated Background – HGT sample

Cluster (1)	N_{tot} (2)	N_{back} (3)	$\%N_{tot}$ (4)	σ_{back} (5)
A 21	291	40	14	0.0306
A 76	195	31	16	0.0045
A 85	323	64	20	0.0146
A 88	72	24	33	0.0237
A 104	151	27	18	0.0153
A 119	268	22	8	0.0036
A 121	145	34	23	0.0312
A 147	155	24	15	0.0038
A 151	342	43	13	0.0102
A 154	272	19	7	0.0069
A 166	157	19	12	0.0219
A 168	235	22	9	0.0038
A 189	157	26	16	0.0024
A 193	264	55	21	0.0108
A 194	129	8	7	0.0002
A 225	202	20	10	0.0082
A 246	103	33	32	0.0137
A 274	174	18	10	0.0246
A 277	230	43	19	0.0324
A 389	173	10	6	0.0116
A 399	254	16	6	0.0067
A 400	190	31	16	0.0014
A 401	288	23	8	0.0108
A 415	243	56	23	0.0295
A 496	226	20	9	0.0017
A 500	225	30	14	0.0114
A 514	282	24	8	0.0108
A 634	71	23	32	0.0014
A 671	293	52	18	0.0108
A 779	115	14	12	0.0006
A 787	154	29	19	0.0448
A 957	288	40	14	0.0066
A 978	295	29	10	0.0069
A 993	272	56	21	0.0135
A 1020	265	49	18	0.0175
A 1035	283	31	11	0.0169
A 1126	248	57	23	0.0349
A 1139	168	38	23	0.0047
A 1185	335	48	14	0.0037
A 1187	227	28	13	0.0150
A 1213	261	13	5	0.0025
A 1216	102	24	24	0.0057
A 1228	278	37	13	0.0038
A 1238	180	23	13	0.0101
A 1254	263	26	10	0.0087

TABLE 2.3. (continued)

Cluster (1)	N_{tot} (2)	N_{back} (3)	$\%N_{tot}$ (4)	σ_{back} (5)
A 1257	212	47	22	0.0046
A 1291	395	29	7	0.0069
A 1318	286	20	7	0.0055
A 1364	226	37	16	0.0358
A 1365	163	19	12	0.0094
A 1367	200	39	19	0.0015
A 1377	402	88	22	0.0198
A 1382	183	43	23	0.0398
A 1383	284	30	11	0.0092
A 1399	284	33	11	0.0229
A 1412	244	32	13	0.0189
A 1436	358	46	13	0.0161
A 1468	166	29	18	0.0175
A 1474	187	34	18	0.0182
A 1496	355	59	16	0.0437
A 1541	205	7	4	0.0048
A 1644	297	36	12	0.0061
A 1651	205	14	7	0.0083
A 1656	424	24	6	0.0011
A 1691	247	46	19	0.0203
A 1749	219	44	20	0.0130
A 1767	308	35	11	0.0146
A 1773	282	16	6	0.0080
A 1775	268	52	19	0.0213
A 1793	248	44	18	0.0269
A 1795	288	44	15	0.0142
A 1809	308	49	16	0.0259
A 1831	308	45	15	0.0205
A 1837	268	32	12	0.0038
A 1904	386	26	7	0.0111
A 1913	276	33	12	0.0080
A 1927	245	24	10	0.0111
A 1983	439	75	17	0.0123
A 1991	368	74	20	0.0215
A 1999	187	18	10	0.0164
A 2005	139	19	14	0.0250
A 2022	322	55	17	0.0148
A 2028	231	33	14	0.0167
A 2029	437	62	14	0.0309
A 2040	278	69	25	0.0121
A 2048	314	37	12	0.0280
A 2052	270	37	14	0.0038
A 2061	285	47	16	0.0233
A 2063	211	12	6	0.0012
A 2065	422	43	10	0.0188
A 2067	283	34	12	0.0153
A 2079	318	67	21	0.0247

TABLE 2.3. (continued)

Cluster (1)	N_{tot} (2)	N_{back} (3)	$\%N_{tot}$ (4)	σ_{back} (5)
A 2089	158	11	7	0.0050
A 2092	267	24	9	0.0089
A 2107	275	45	16	0.0067
A 2124	298	47	16	0.0169
A 2142	311	29	9	0.0197
A 2147	465	35	8	0.0037
A 2151	388	61	16	0.0071
A 2152	471	81	17	0.0095
A 2162	124	17	14	0.0015
A 2175	448	35	8	0.0284
A 2197	313	35	11	0.0027
A 2199	389	46	12	0.0035
A 2255	417	22	5	0.0117
A 2256	451	38	8	0.0115
A 2328	131	14	11	0.0258
A 2347	90	15	16	0.0176
A 2382	197	13	6	0.0044
A 2384	127	10	8	0.0075
A 2399	256	26	10	0.0075
A 2410	235	29	12	0.0157
A 2457	248	16	7	0.0049
A 2634	411	76	18	0.0062
A 2657	171	17	10	0.0025
A 2666	171	31	18	0.0018
A 2670	255	26	10	0.0121
A 2675	182	33	18	0.0149
A 2700	129	38	30	0.0308

Chapter 3

PROBABILITY DENSITY ESTIMATION

3.1 Introduction

Probability density estimation has a wide field of application. As such, it has received a great deal of interest from the statistical community. The question which density estimation attempts to answer is the following: given a sample of n independent observations, $X_i \ldots X_n$, what is the probability that the next observation will be at any given position x. Or, what is $f(x)$, the probability density function (PDF), such that

$$P(a < X < b) = \int_a^b f(x)\,dx. \tag{3.1}$$

This problem can be approached parametrically or nonparametrically. In the parametric approach, the form of the PDF is assumed and various parameters measured. The most commonly applied PDF is the normal or Gaussian distribution, where the average μ and the standard deviation σ of the observations are estimated from the data. In fact, it has become so widely used that many researchers continue to use μ and σ even for distributions which are not Gaussain, and for which robust estimators for the location and scale of the data are required. If, as is often the

case in astronomy, there is no *a priori* reason to assume a particular form of $f(x)$ a nonparametric approach that makes as few assumptions as possible about the density being estimated is desirable.

3.2 The Histogram

The oldest and most widely used form of nonparametric density estimate is the histogram, which dates back at least to the work of Graunt in 1662. The density estimate of a histogram, $\hat{f}(x)$, is defined as:

$$\hat{f}(x) = \frac{1}{Nh}(\text{no. of } X_i \text{ in same bin as } x), \tag{3.2}$$

where h is the width of the bins and N is the total number of observations. In addition to specifying the bin width h which controls the smoothness of the histogram, it is also necessary to specify an origin for the bins. While the choice of origin may seem to be a trivial matter, it can have quite an effect on the shape of the histogram constructed, especially with small to moderate-sized data batches. As an example, Figures 3.1 and 3.2 show histograms constructed from 88 measured redshifts of Abell 400. The data are taken from Beers *et al.* (1992) and have errors on the order of 50 km s^{-1}. Both histograms are constructed using the same data and the same bin width, 300 km s^{-1}. The only difference between the plots is the choice of bin origin. In Figure 3.1 the origin of the bins has been set at 5000 km s^{-1} while that of Figure 3.2 has been shifted by 200 km s^{-1} and set at 5200 km s^{-1}. Although statistically the two histograms show the same thing, one gives the impression of bimodality while the other does not. In this case, it turns out that most choices of bin origin lead to a bimodal histogram, as correctly identified by Beers *et al.*, and the first choice of origin was simply unfortunate.

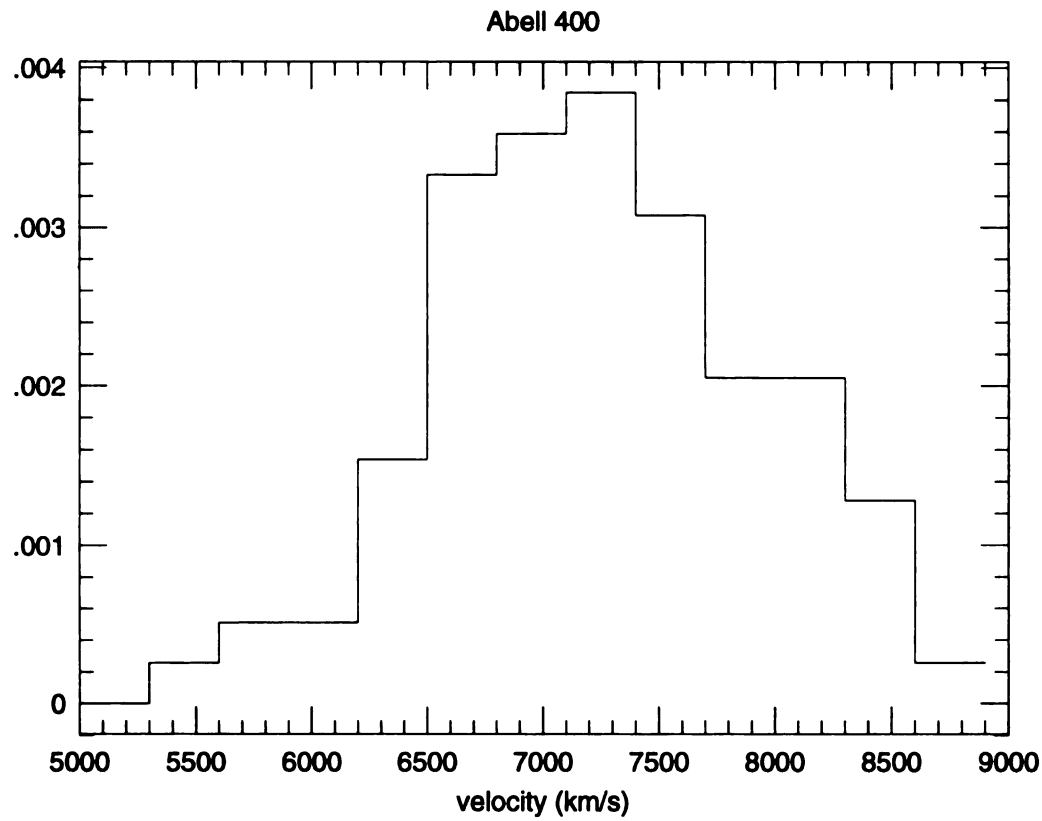

Fig. 3.1.— Histogram of A400 velocity measurements from Beers *et al.* (1992). The bin width is 300 km s^{-1} and the bin origin is at 5000 km s^{-1}.

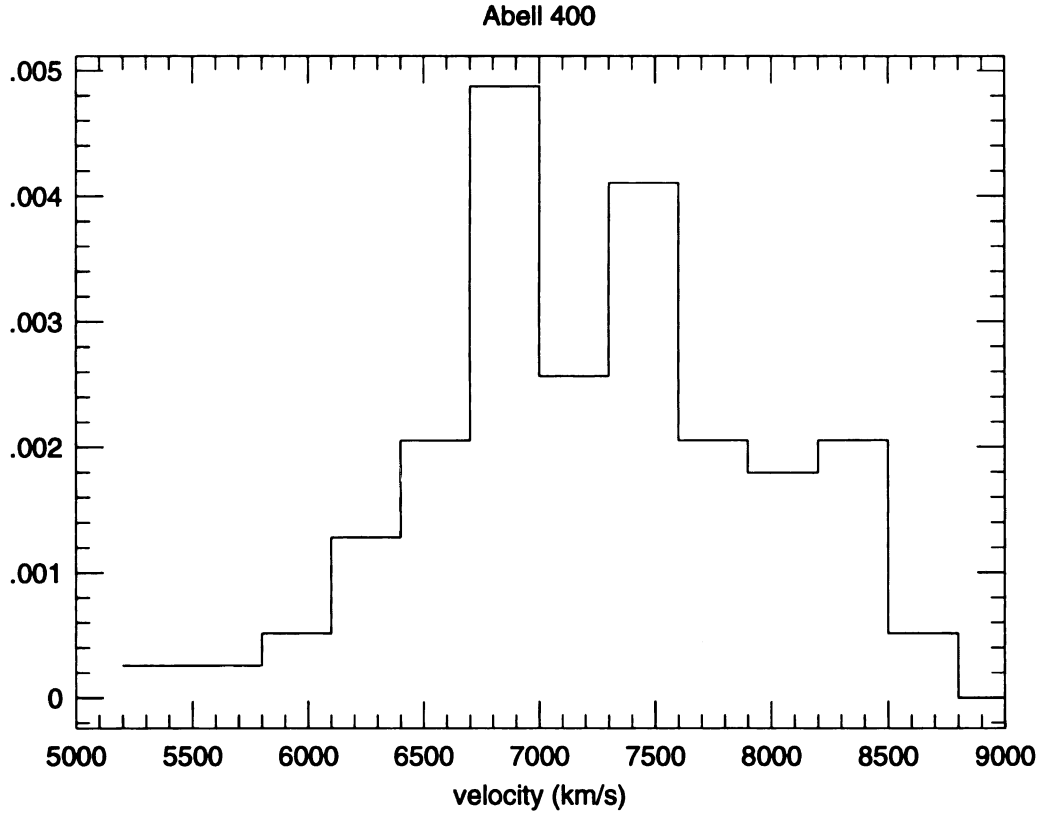

Fig. 3.2.— Histogram of A400 velocity measurements from Beers *et al.* (1992). The bin width is 300 km s^{-1} and the bin origin is at 5200 km $^{-1}$.

3.3 Generalizations of the Histogram

There are two simple ways to free the histogram from dependence on bin origin. Perhaps the most obvious solution is to construct m histograms, each one with the origin shifted by $\delta = h/m$, and average them together as discussed by Chamayou (1980). The result is referred to as an Average Shifted Histogram (ASH). Thus:

$$\hat{f}(x) = \frac{1}{m} \sum_{i=1}^{m} \hat{f}_i(x), \tag{3.3}$$

where $\hat{f}_i$ is the histogram estimator given in equation (3.2) with bin origins of 0, h/m, $2h/m$, ..., $(m-1)h/m$, respectively. It is possible to rewrite this in a more computationally convenient form by defining a new bin width $\delta = h/m$ (see Scott 1992). Then,

$$\hat{f}(x, m) = \frac{1}{nh} \sum_{i=1-m}^{m-1} w_m(i) \, [\text{no. of } X_i \text{ in same bin as } x], \tag{3.4}$$

where the bin is now the smaller bin and $w_m(i)$ is a weight function given by:

$$w_m(i) = 1 - \frac{|i|}{m}. \tag{3.5}$$

Note that unlike the histogram, with its box-shaped weight function, the weight function for the ASH is an isosceles triangle. An example of the ASH is given for the Abell 400 data in Figure 3.3 with $m = 32$ and bin width of 300 km s^{-1}. Here the bimodal nature of the PDF is clear, thus justifying the above statement that the choice of bin origin in Figure 3.1 was merely unfortunate. There is also evidence of the possible third peak at higher velocity causing the density estimate to flatten off.

An alternative solution is referred to by Silvermann as the "naive estimator." Instead of rigidly fixing the bins to some arbitrary origin on the coordinate axis and counting the number of observations which lie in each bin, as in the classical

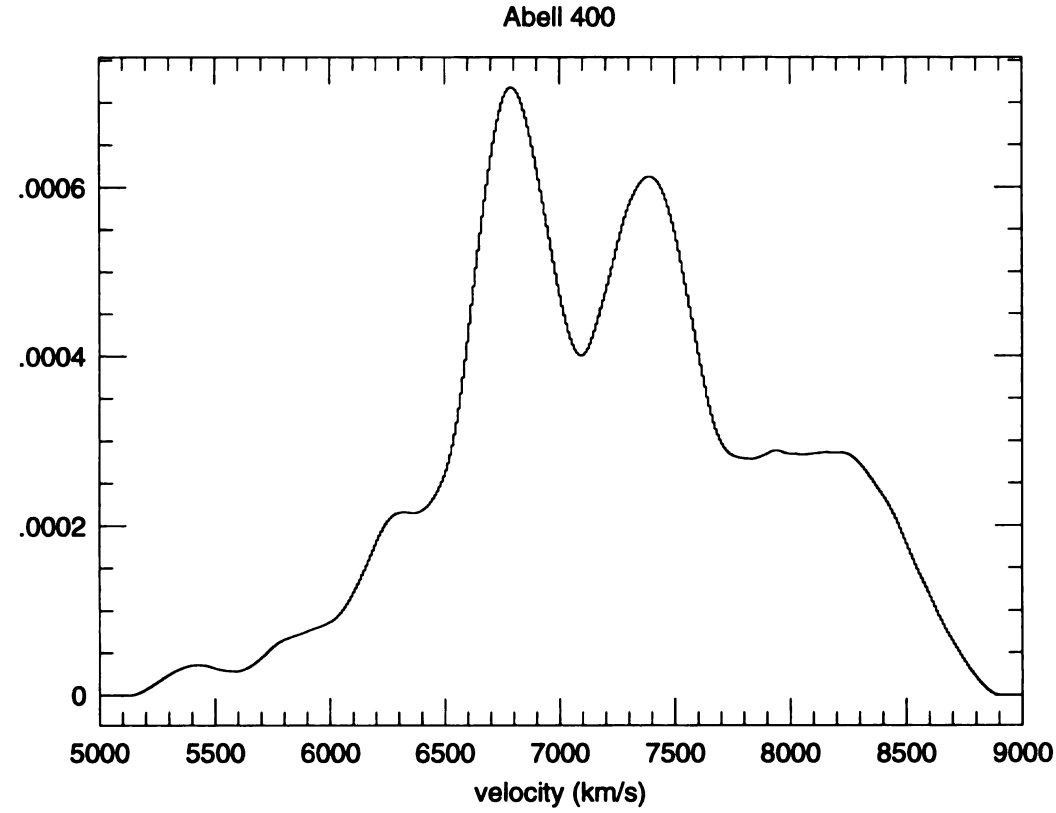

Fig. 3.3.— Average Shifted Histogram of A400 velocity measurements. The bin width is 300 km s^{-1} and m=32.

histogram, in the naive estimator the bin origins are allowed to float based on the position of each data point. This is accomplished by placing a box of width $2h$ and height $1/(2hn)$ centered on each data point and summing the box heights at each coordinate position x. Or:

$$\hat{f}(x) = \frac{1}{2nh} \sum_{i=1}^{n} w\left(\frac{x - X_i}{h}\right). \tag{3.6}$$

The function $w(x)$ is again a weight function and in this case is:

$$w(x) = \begin{cases} \frac{1}{2} & \text{if } |x| < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

Figure 3.4 shows a density estimate constructed in this manner, again for the A400 data. Notice that this density estimate leads to very sharp peaks which can be aesthetically unpleasant at best and misleading at worst. Also, like the histogram, its derivative is zero everywhere except at those points where it is discontinuous, in this case at each $X_i \pm h$. Nevertheless, it clearly shows the two peaks as well as the lower density plateau at high velocity; in short, all the information shown in the ASH.

3.4 The Kernel Estimator

Although both the ASH and the naive estimator discussed above are independent of the the choice of origin, they still retain the discontinuous nature of the histogram. This prevents them from being useful when derivatives of the density estimate are sought, as in the peak identification procedure employed in Chapter 4. The discontinuities in both the naive estimate and the ASH arise from the discontinuity of the weight functions: the box shape in the naive estimator and the histogram or the triangle shape in the ASH. This problem can be overcome by generalizing the weight functions to different shapes which are themselves continuous. Thus we can gener-

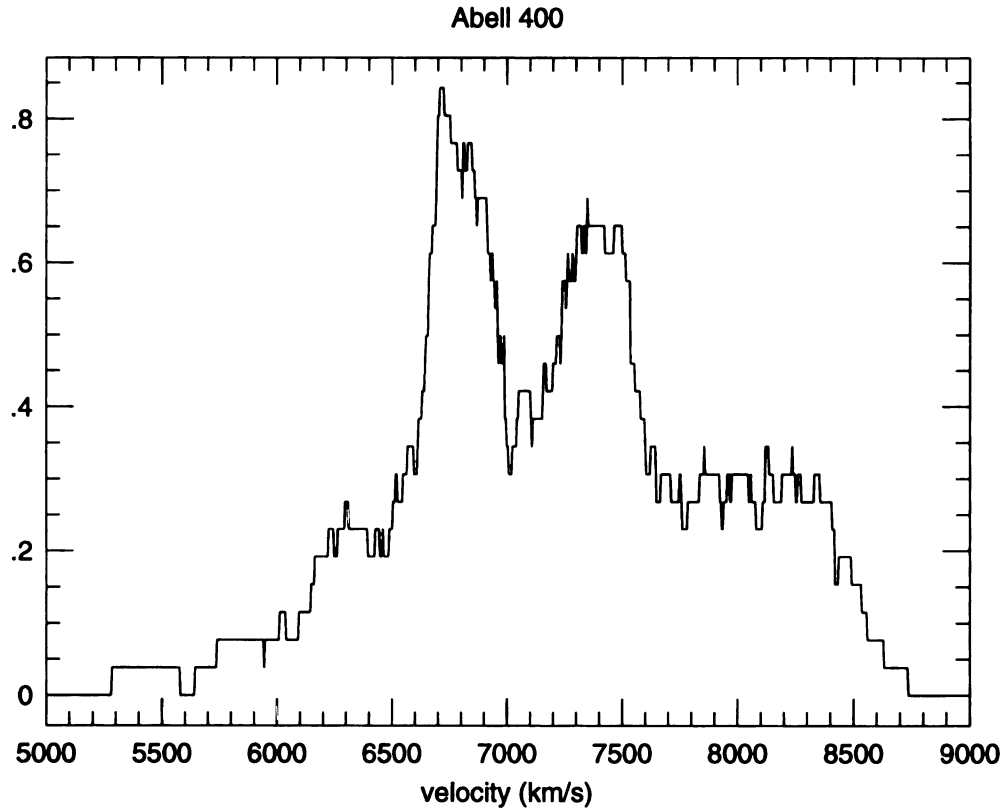

Fig. 3.4.— "Naive estimator" for the A400 velocity data. Again the bin width is set to 300 km s^{-1}.

alize equation (3.6) by replacing the weight function $w(x)$ with a continuous kernel function $K(x)$ where:

$$\int_{-\infty}^{\infty} K(x)dx = 1. \tag{3.8}$$

Instead of placing a box over each data point and summing the boxes, a bump with shape controlled by the kernel function and width specified by the smoothing parameter h is used. In the ASH, the weights must sum to m so that $w_m(i)$ could be replaced by

$$w_m(i) = \frac{mK(i/m)}{\sum_{j=1-m}^{m-1} K(j/m)} \quad i = 1-m, \ldots, m-1. \tag{3.9}$$

Some commonly used kernel functions are:

Epanechnikov:

$$K(x) = \begin{cases} \frac{3}{4}(1-x^2) & \text{for } x^2 < 1 \\ 0 & \text{otherwise} \end{cases}, \tag{3.10}$$

Biweight:

$$K(x) = \begin{cases} \frac{15}{16}(1-x^2)^2 & \text{for } |x| < 1 \\ 0 & \text{otherwise} \end{cases}, \tag{3.11}$$

and Normal:

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{3.12}$$

A density estimate constructed by summing a series of $K(X_i)$ will inherit all of the continuous and differential properties of $K(x)$.

It can be shown that the Epanechnikov kernel function (Epanechnikov 1969) minimizes the mean integrated square error (MISE) between the density estimate and the true density, provided the optimal value of h is used. The choice of the form of

the kernel function is not critical since the efficiencies of the other kernels differ from that of the Epanechnikov kernel only slightly (Silvermann 1986). Thus, the choice can be made on the basis of the desired differentiability of the estimate or speed of calculation. For instance, the Epanechnikov and the biweight kernels both have discontinuous derivatives at $x = \pm 1$. On the other hand, the normal kernel has a continuous derivative everywhere, but suffers from infinite tails.

Figure 3.5 shows the density function constructed with the A400 data employing a normal kernel. As with the ASH the two peaks are seen clearly in the kernel estimator as is the possible third, lower-density peak at higher velocity. This similarity between the ASH and the kernel density estimates is not an accident. It can be shown (Scott 1992) that in the limit as $m \to \infty$ the ASH estimate approaches that of the kernel estimate and the two techniques are equivalent.

Although the previous examples used only one-dimensional data, the same arguments apply in two. In fact, the problem of bin origin becomes even worse since not only is the two dimensional histogram affected by shifts in the x and y position of the origin, but also by rotations of the coordinate axis. Since this thesis is primarily concerned with density estimation in two dimensions, in the following discussion the kernel estimator is generalized appropriately.

3.5 Adaptive Smoothing Methods

The kernel estimator provides a smooth density estimate which is independent of origin. However, use of a fixed value of h will yield a density estimate which is *over*-smoothed in high-density regions, tending to hide real structure, and *under*-smoothed in low-density regions, which are subject to Poisson noise. One solution to this difficulty is to vary the kernel width based on the local density.

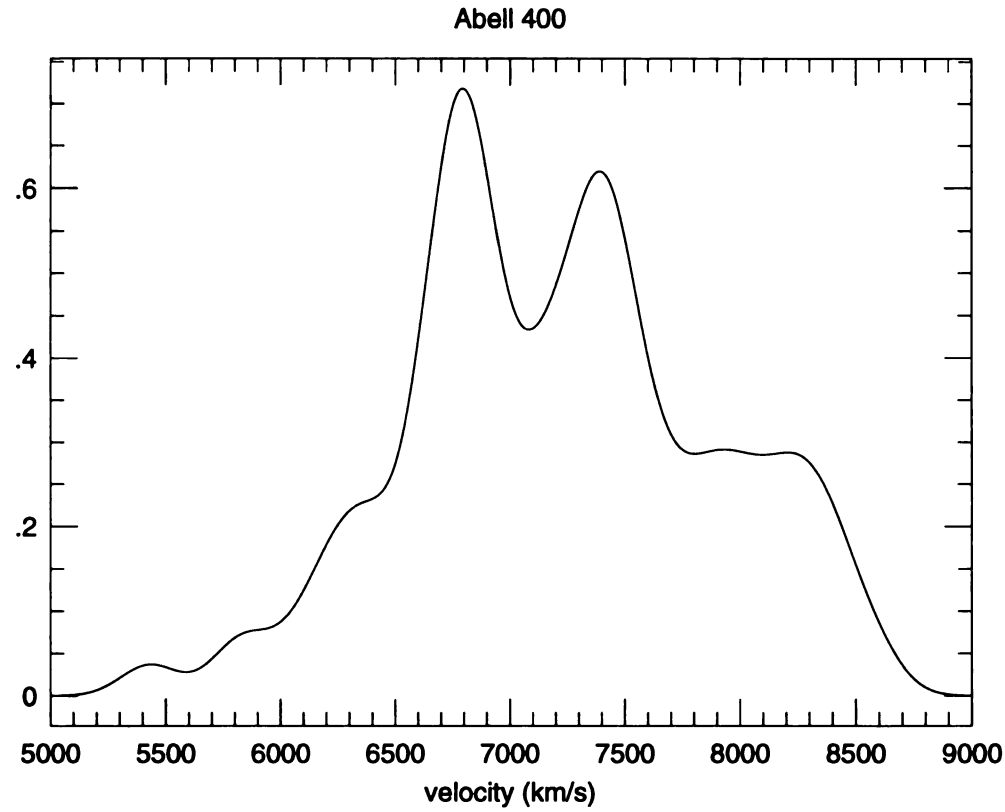

Fig. 3.5.— Adaptive-kernel density estimate of A400 velocity data. The smoothing parameter $h = 150$ km s^{-1}.

Perhaps the simplest nonparametric adaptive smoothing method is the nearest neighbor density (hereafter NND.) The NND estimate in two dimensions is defined by:

$$\hat{f}(x) = \frac{(k)}{\pi n r_k(x)^2}, \tag{3.13}$$

where $r_k(x)$ is the distance to the kth nearest neighbor and n is the number of observations. In the NND, it is the value of k which controls the smoothing. Astronomers generally set $k = 10$, though $k \propto n^{4/(d+4)}$ gives the best estimate in d dimensions (Silvermann 1986) with the "constant" of proportionality depending on the position x.

There are, however, a number of problems with the NND estimate. The tails of the estimate will fall off very slowly, $\propto r^{-1}$, regardless of the true distribution. Thus the NND estimate *over*-smooths in low density regions and generally performs worse than the fixed-kernel estimator. (This is true in one and two dimensions, but for $d \geq 5$ the NND performs better than the fixed kernel estimator at least in the case of a normal density distribution.) Furthermore, the NND is not a smooth function since the derivative is discontinuous at the points $1/2(X_{(j)} - X_{(j+k)})$ where the $X_{(j)}$ are the order statistics of the sample (*i.e.* $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(N)}$.)

An alternative to the NND, first proposed by Fix & Hodges (1951) and discussed in detail in the monograph by Silvermann (1986), is to vary the kernel width based on an estimate of the local density. Construction of the adaptive-kernel density estimate proceeds as follows:

Obtain a "pilot estimate" $\tilde{f}(x)$ which satisfies $\tilde{f}(X_i) > 0$ for all i.
Define "local bandwidth factors" λ_i by:

$$\lambda_i = \left\{ \frac{\tilde{f}(X_i)}{g} \right\}^{\alpha}, \tag{3.14}$$

where g is the geometric mean of the $\tilde{f}(X_i)$:

$$\log g = n^{-1} \sum_{i=1}^{n} \log \tilde{f}(X_i) \tag{3.15}$$

and α is a sensitivity parameter satisfying $0 \leq \alpha \leq 1$.

Define the "adaptive-kernel estimate" $\hat{f}(x)$ by:

$$\hat{f}(x) = n^{-1} \sum_{i=1}^{n} h^{-d} \lambda_i^{-d} K\{h^{-1} \lambda_i^{-1} (x - X_i)\}, \tag{3.16}$$

where d is the dimensionality of the data (in this case $d = 2$).

Note that the adaptive-kernel estimator defined above is the same as the non-adaptive (fixed) kernel estimator, except that the window width h is now replaced by $h\lambda_i$, the local bandwidth indicators. The result is that the window width in the adaptive-kernel estimator is *decreased* in high-density regions and *increased* in low-density regions. The amount by which the window width is decreased or increased can be altered by changing the sensitivity parameter α. With $\alpha = 0$, for instance, all the bandwidth factors become 1 and the pilot density estimate is returned. As α approaches 1, the density estimator is similar to a nearest-neighbor method, which is prone to local noise and has heavy tails. It can be shown (Abramson 1982) that a value of $\alpha = 1/2$ provides a smaller bias in the density estimate than that obtained using a fixed kernel width in both one and two dimensions (this is not necessarily true for other choices of α.) Thus the value of $\alpha = 1/2$ is adopted throughout this thesis.

It needs to be pointed out that current statistical research indicates that the adaptive-kernel technique outlined above is not well behaved asymptotically (as the number of observations approaches infinity.) For data sets larger than 20,000, it can be shown that the adaptive procedure performs significantly worse than a fixed kernel estimator (Scott 1992). While this is cause for concern the adaptive-kernel

has, at present, the best track record for a wide variety of PDF's with $N \lessapprox 200$ and is therefore employed in this thesis.

So far in the discussion, little attention has been paid to the subject of the smoothing parameter size. Since the choice of smoothing parameter has the greatest effect on the accuracy of the density estimate, a great deal of effort has be expended by statisticians searching for the best possible value of h. Unfortunately, the optimal smoothing size depends on the unknown density for which an estimate is sought and no value of h (or even a prescription for finding h) will give the best estimate for all density distributions. Because of this, many statisticians recommend choosing the smoothing parameter subjectively. That is, vary h until the desired level of smoothness is achieved. In fact, Scott (1992) has expressed the opinion that there is no wrong choice of h, as information is gained by all values. The drawback of this method is that different researchers will no doubt have different opinions of when an estimate is "smooth enough." There is a great temptation to oversmooth since this leads to neater looking plots. This temptation should be avoided since further smoothing can be done by eye, but an oversmoothed estimate can not be un-smoothed, and real effects in the data can be lost. To avoid this, an automatic selection is clearly desirable.

Two different methods for choosing h will be applied in this thesis. The first involves choosing the optimal smoothing parameter based on minimizing the MISE for a given kernel function with respect to a particular distribution or family of distributions. For density estimation of a bivariate-normal distribution this is:

$$h_n = A(K)\{1/2(\sigma_x^2 + \sigma_y^2)\}^{1/2} N^{-1/6}, \tag{3.17}$$

where $A(K)$, a constant that depends on the kernel function, is 0.96 and 2.04 for the normal and biweight kernels, respectively. This prescription is often referred to

as the "normal rule." Being based on the normal distribution, density estimates constructed with h_n will oversmooth multimodel densities. Based on experience, Silvermann suggests using $h = 0.85h_n$ as a good compromise, as it works well with bimodal as well as skewed distributions. This method, which is quick and simple to calculate, will be used in constructing a consistent set of contour maps for the galaxy clusters. Because most of the clusters show evidence of multimodality, a value of $h = 0.75h_n$ is used. It should be noted that the factor of 0.75 is based on experience using the adaptive-kernel technique with clusters of galaxies and is somewhat arbitrary. Clearly, there is a trade off. A unimodal-normal cluster will be undersmoothed to some extent, while a multi-modal cluster will tend to be oversmoothed. Use of an adaptive-kernel technique however, partially compensates for this effect, and the final density estimate is relatively insensitive to variations in kernel width within 15% to 20% of the optimal value.

The other method, least squares cross validation (LSCV), involves finding the value for $h = h_{CV}$ which minimizes the cross validation term in the expression for the integrated square error (ISE). The ISE is given by:

$$ISE(\hat{f}) = \int f^2(x)dx + \int \hat{f}^2(x)dx - 2\int f(x)\hat{f}(x)dx. \tag{3.18}$$

Because the first term depends only on the actual density it is constant with respect to changes in h. Therefore minimizing the ISE is equivalent to minimizing the term:

$$M_0(h) = \int \hat{f}(x)^2dx - 2\int f(x)\hat{f}(x)dx. \tag{3.19}$$

Unfortunately, this still depends on the unknown density $f(x)$. To get around this, it can be shown that:

$$E\left\{2\int f(x)\hat{f}(x)dx\right\} = E\left\{2N^{-1}\sum_{i=1}^{N}\hat{f}_{-i}(X_i)\right\}, \tag{3.20}$$

where E is the expectation operator and $\hat{f}_{i-1}$ is the density estimate obtained with the ith observation deleted from the calculation. If

$$M_0(h) = \int \hat{f}(x) - 2N^{-1} \sum_{i=1}^{N} \hat{f}_{-i}(X_i), \tag{3.21}$$

then under the mild assumption that the minimizer of $E\left\{M_0(\hat{f})\right\}$ is near the minimizer of M_0, h_{CV} will minimize the ISE. Furthermore, due to a result of Stone (1984), in the limit as $N \to \infty$, h_{CV} will be the best possible choice of smoothing parameter (in the sense that a density estimate constructed using it will have the minimum ISE) and could not be improved upon even if $f(x)$ were known exactly.

Despite the superior asymptotic performance of the LSCV method, it nevertheless can run into problems when applied to real data sets. In particular, Silvermann (1986) shows that for small values of h, $M_0(h)$ can become extremely sensitive to small scale effects (such as the rounding of real numbers) in the data. He therefore recommends searching of h_{CV} only in the range of $0.25h_n < h < 1.5h_n$, where h_n is given by the normal rule.

3.6 Application: Galaxy Number-Density Plots

To find the pilot estimate of density in the cluster contour maps, the kernel estimator is employed on a 100×100 grid with a fixed window width. The window width is set automatically based on the total number of galaxies in each cluster and scaled by their average marginal variance, For this application, the kernel function is taken to be the biweight kernel:

$$K_2(\mathbf{x}) = \begin{cases} 3\pi^{-1}(1 - \mathbf{x}^{\mathbf{T}}\mathbf{x})^2 & \text{for } \mathbf{x}^{\mathbf{T}}\mathbf{x} < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.22}$$

The biweight kernel function is employed because it gives a smoother appearing contour plot than the Epanechnikov calculated on the same number of grid points.

The contour plots in Figure 3.6 are constructed from the positions of galaxies previously published by Dressler (1980). The bar in each plot indicates the initial smoothing scale ($h = 0.75h_n$), the size of which varies from 0.16 to 1.11 Mpc. The positions of the galaxies classified as D or cD by Dressler are indicated by filled circles. The crosses and plus marks indicate the median positions of the groups that are returned as significant by the KMM and DEDICA algorithms, respectively, as discussed in Chapter 4. The maps are centered on the median position of all the galaxies in each cluster.

Table 3.1 presents the parameters of each map. Column (1) gives the cluster name. The number of galaxies in each cluster is listed in column (2). The RA and DEC (1950 coordinates) of the median galaxy position for each cluster are in columns (3) and (4). Column (5) lists the surface density of the highest contour in galaxies per square arcmin. Column (6) lists the surface density of the lowest contour, which is set to one galaxy per resolution element. Listed in column (7) is the contour spacing. Columns (8) and (9) are the initial smoothing scale in arcmin and Mpc, respectively.

TABLE 3.1. Adaptive Kernel Map Parameters – Dressler Sample

Cluster	N	RA (1950)	DEC (1950)	σ_{max} (gal/arcmin^2)	σ_{min} (gal/arcmin^2)	$\delta\sigma$ (gal/arcmin^2)	$0.75h_n$ arcmin	$0.75h_n$ Mpc
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 14	79	00 09 37.4	−24 39 41	0.2262	0.0127	0.0214	8.7	0.59
A 76	72	00 35 02.8	+06 04 23	0.1152	0.0100	0.0105	9.8	0.41
A 119	116	00 51 40.1	−02 01 20	0.3054	0.0112	0.0294	9.2	0.42
A 151	105	01 03 23.3	−16 21 25	0.2771	0.0052	0.0272	13.5	0.77
A 154	79	01 05 32.3	+16 55 06	0.4196	0.0105	0.0409	9.5	0.66
A 168	106	01 09 09.2	−00 49 21	0.3241	0.0132	0.0311	8.5	0.42
A 194	75	01 20 30.2	−02 19 25	0.0947	0.0038	0.0091	15.8	0.32
A 376	119	02 39 58.1	+36 13 06	0.4659	0.0163	0.0450	7.7	0.39
A 400	92	02 52 23.6	+05 04 46	0.1370	0.0049	0.0132	13.9	0.36
A 496	81	04 29 02.9	−13 51 31	0.1225	0.0090	0.0114	10.3	0.41
A 539	99	05 11 02.6	+05 39 35	0.2757	0.0055	0.0270	13.2	0.41
A 548	234	05 44 06.6	−26 04 47	0.2257	0.0052	0.0221	13.5	0.59
A 592	61	07 37 10.2	+08 45 35	0.0731	0.0034	0.0070	16.8	1.11
A 754	150	09 04 09.4	−10 14 41	0.2026	0.0061	0.0197	12.5	0.73
A 838	62	09 30 10.4	−05 18 09	0.0415	0.0027	0.0039	18.9	1.04
A 957	82	10 08 06.9	−01 24 20	0.1666	0.0038	0.0163	15.9	0.76
A 978	62	10 14 51.7	−06 51 26	0.0793	0.0058	0.0074	12.8	0.73
A 979	86	10 14 41.7	−08 22 41	0.0900	0.0035	0.0087	16.5	0.98
A 993	91	10 15 37.8	−05 26 30	0.0718	0.0038	0.0068	15.8	0.90
A 1069	47	10 35 44.6	−09 17 20	0.0568	0.0038	0.0053	15.8	1.06
A 1139	63	10 53 34.2	+01 03 35	0.0918	0.0048	0.0087	14.1	0.59
A 1142	59	10 54 50.1	+10 15 56	0.0540	0.0047	0.0049	14.3	0.57
A 1185	44	11 05 18.7	+28 19 48	0.1299	0.0074	0.0123	11.4	0.44
A 1377	52	11 39 15.9	+55 16 09	0.2303	0.0078	0.0223	11.1	0.61
A 1631	90	12 47 41.3	−15 55 20	0.1671	0.0074	0.0160	11.4	0.65
A 1644	145	12 51 38.8	−17 49 03	0.1114	0.0054	0.0106	13.4	0.71
A 1656	245	12 54 32.6	+27 30 35	0.3618	0.0065	0.0355	12.1	0.31
A 1736	166	13 21 14.0	−27 35 14	0.1262	0.0053	0.0121	13.4	0.70
A 1913	86	14 21 08.9	+16 12 40	0.1450	0.0040	0.0141	15.5	0.88

TABLE 3.1. (continued)

Cluster	N	RA (1950)	DEC (1950)	σ_{max} (gal/arcmin^2)	σ_{min} (gal/arcmin^2)	$\delta\sigma$ (gal/arcmin^2)	$0.75h_n$ arcmin	$0.75h_n$ Mpc
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 1983	123	14 47 28.8	+16 09 24	0.2173	0.0051	0.0212	13.6	0.68
A 1991	53	14 49 27.0	+18 09 07	0.0609	0.0055	0.0055	13.1	0.83
A 2040	108	15 07 07.9	+06 50 19	0.1604	0.0047	0.0156	14.3	0.70
A 2063	110	15 17 42.2	+08 05 57	0.2228	0.0059	0.0217	12.7	0.48
A 2151	152	16 00 33.7	+16 58 02	0.1684	0.0060	0.0162	12.7	0.50
A 2256	83	17 05 18.5	+78 25 34	0.5759	0.0216	0.0554	6.6	0.43
A 2589	72	23 18 31.9	+16 00 24	0.3355	0.0151	0.0320	7.9	0.37
A 2634	132	23 34 03.4	+25 59 44	0.2752	0.0119	0.0263	8.9	0.31
A 2657	82	23 39 18.2	+08 04 55	0.2883	0.0127	0.0276	8.7	0.39
DC 0003-50	79	00 00 04.2	−51 37 47	0.0727	0.0036	0.0069	16.4	0.63
DC 0103-47	53	01 02 22.2	−47 46 31	0.0462	0.0038	0.0042	15.8	0.41
DC 0107-46	55	01 04 37.7	−46 58 25	0.0627	0.0040	0.0059	15.4	0.40
DC 0247-31	48	02 44 27.1	−32 08 52	0.1661	0.0041	0.0162	15.3	0.36
DC 0317-54	65	03 14 44.7	−54 48 03	0.0988	0.0043	0.0094	15.0	0.88
DC 0326-53	161	03 23 14.6	−54 15 08	0.1084	0.0044	0.0104	14.7	0.91
DC 0329-52	190	03 26 00.5	−53 24 09	0.3172	0.0076	0.0310	11.2	0.68
DC 0410-62	64	04 03 25.1	−63 45 41	0.0344	0.0061	0.0031	18.0	0.35
DC 0428-53	131	04 25 04.7	−54 41 03	0.2334	0.0068	0.0227	11.9	0.53
DC 0559-40	112	05 56 33.8	−40 51 52	0.1859	0.0058	0.0180	12.8	0.68
DC 0608-33	122	06 05 22.3	−34 26 52	0.1488	0.0048	0.0144	14.1	0.54
DC 0622-64	98	06 19 09.8	−65 39 41	0.1288	0.0053	0.0123	13.4	0.40
DC 1842-63	55	18 40 34.0	−63 58 26	0.2182	0.0111	0.0207	9.3	0.16
DC 2048-52	216	20 45 33.4	−53 34 58	0.2864	0.0059	0.0281	12.8	0.64
DC 2103-39	108	21 00 49.2	−40 01 47	0.1457	0.0058	0.0140	12.9	0.72
DC 2345-28	95	23 43 14.7	−28 54 48	0.2834	0.0093	0.0274	10.2	0.31
DC 2349-28	68	23 46 31.4	−29 03 37	0.0832	0.0069	0.0076	11.7	0.37
Centaurus	73	12 42 10.5	−42 05 36	0.0562	0.0036	0.0053	16.3	0.20

Fig. 3.6.— Adaptive kernel maps of the Dressler sample clusters. Map paramters are listed in Table 3.1.

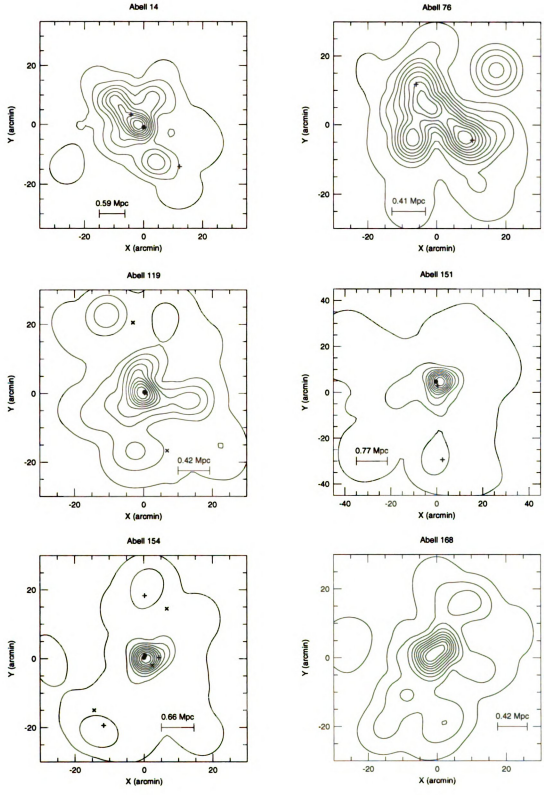

Figure 3.6. Adaptive-Kernel Maps for the Dressler Clusters

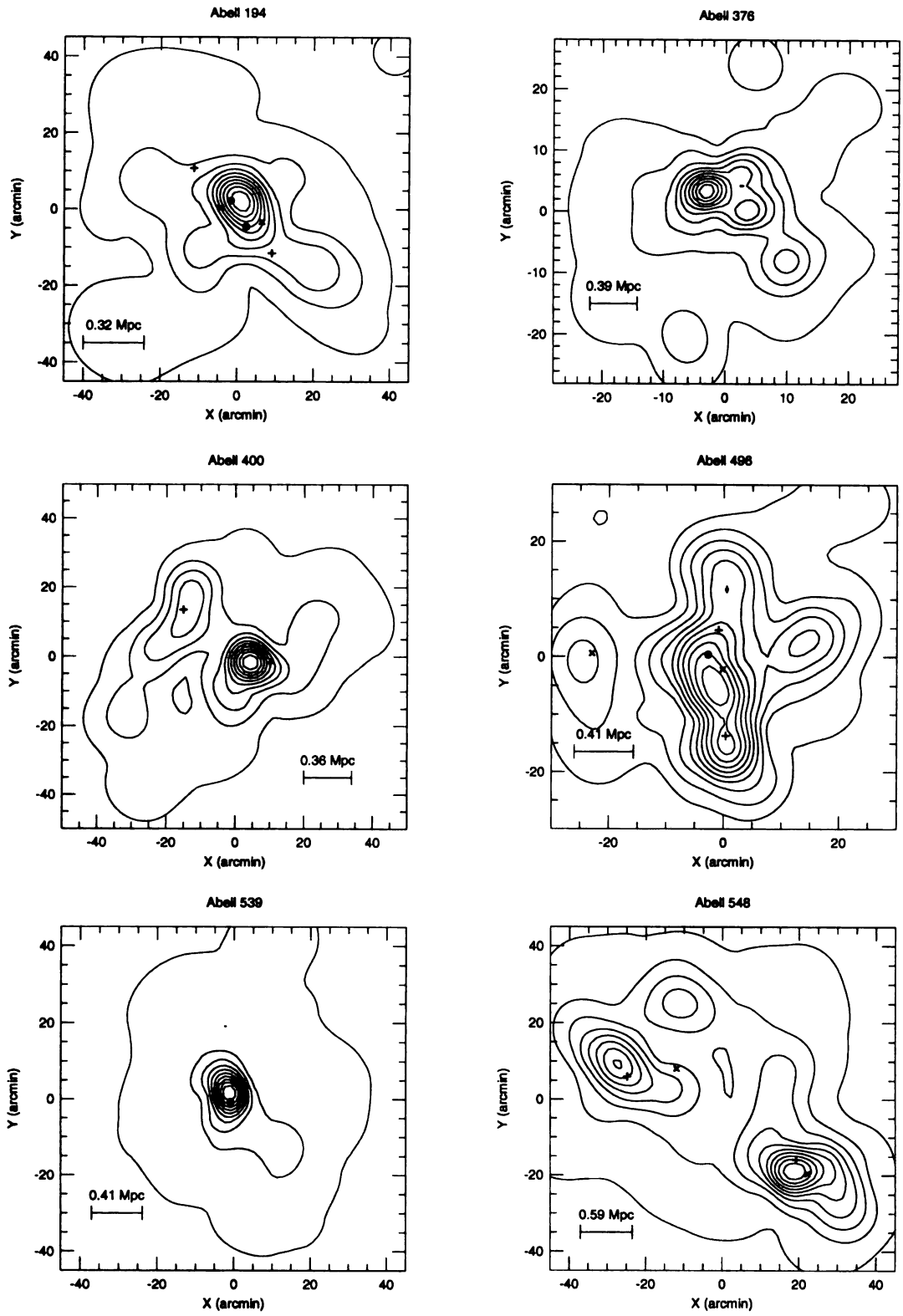

Figure 3.6 (cont'd).

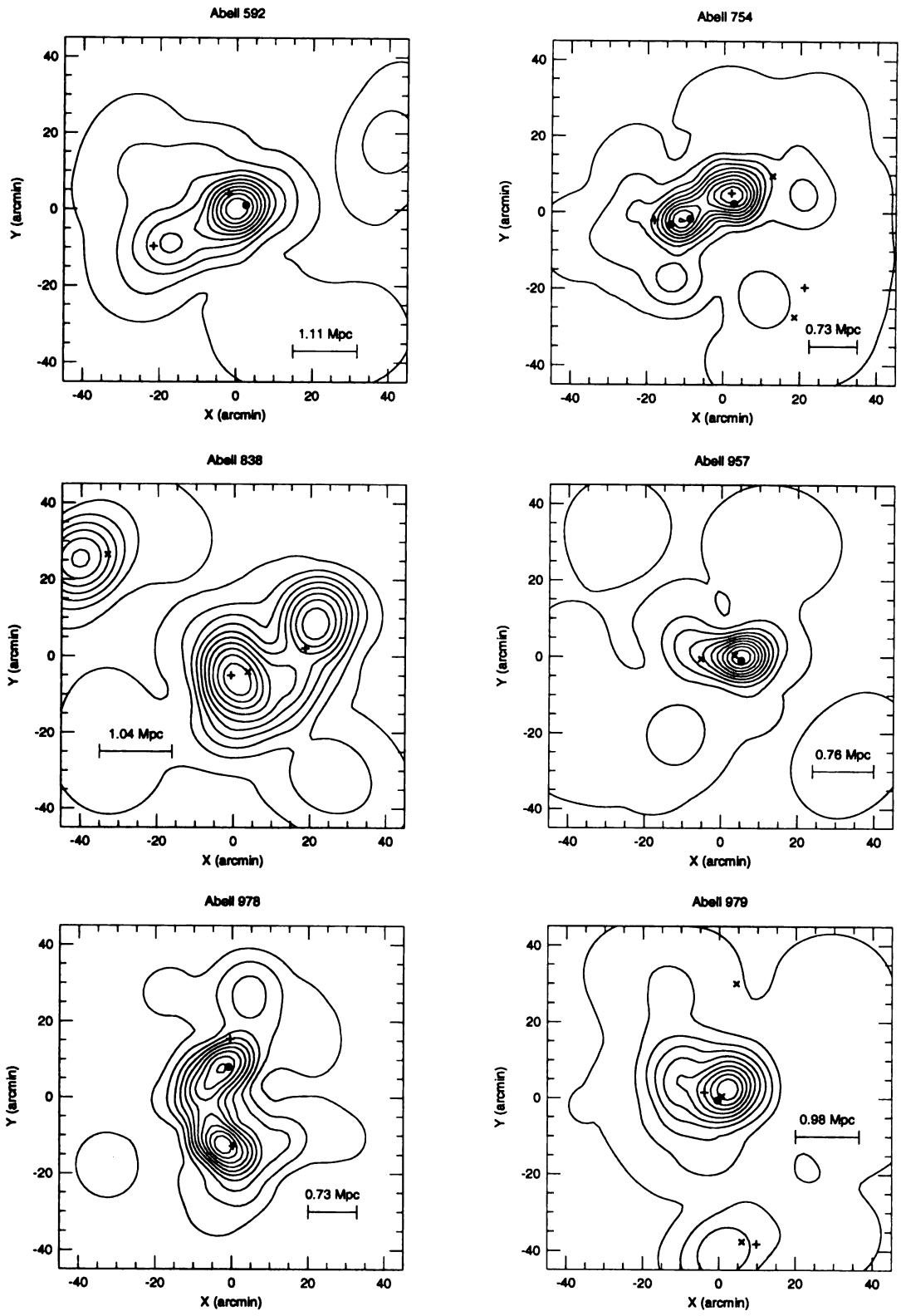

Figure 3.6 (cont'd).

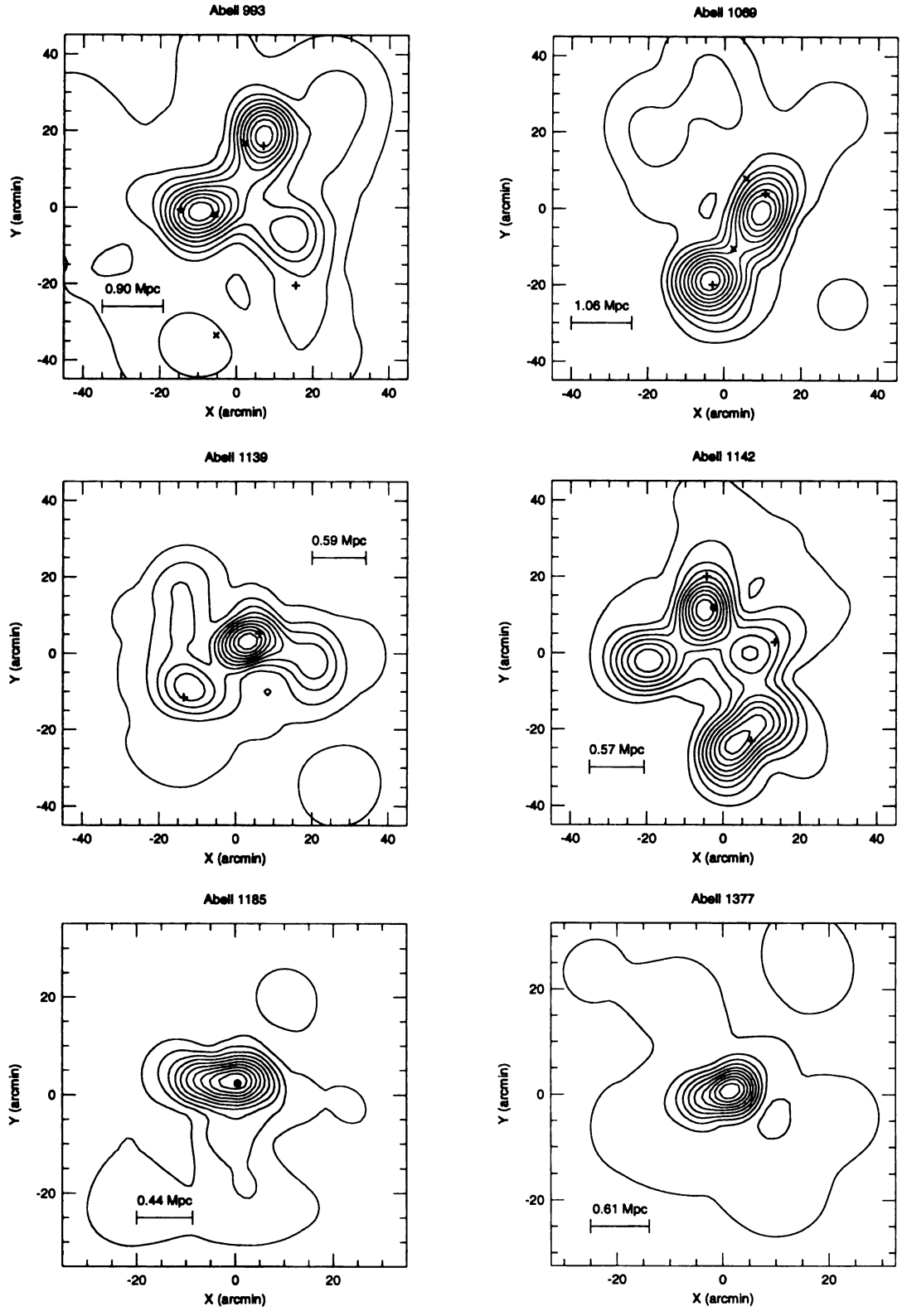

Figure 3.6 (cont'd).

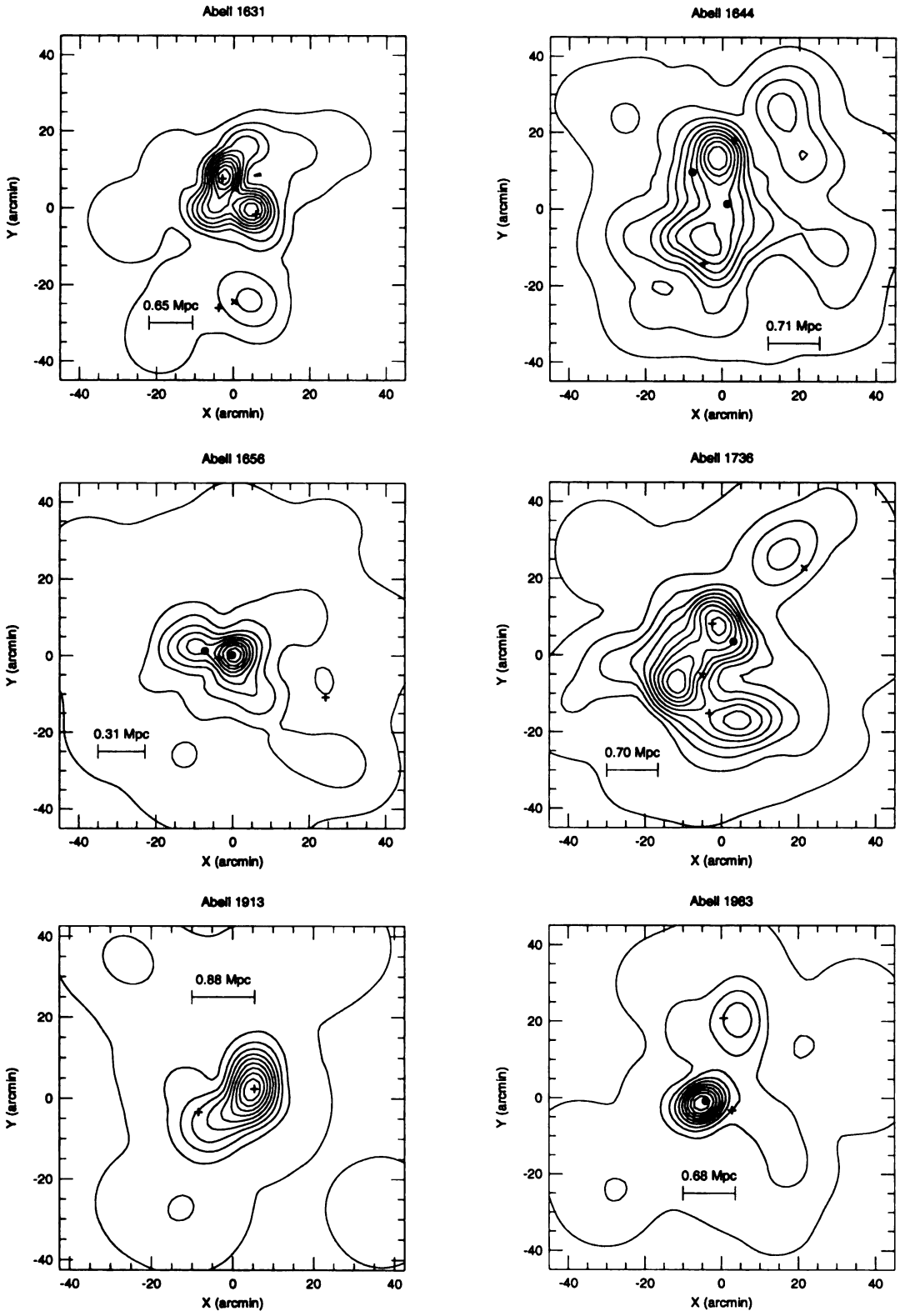

Figure 3.6 (cont'd).

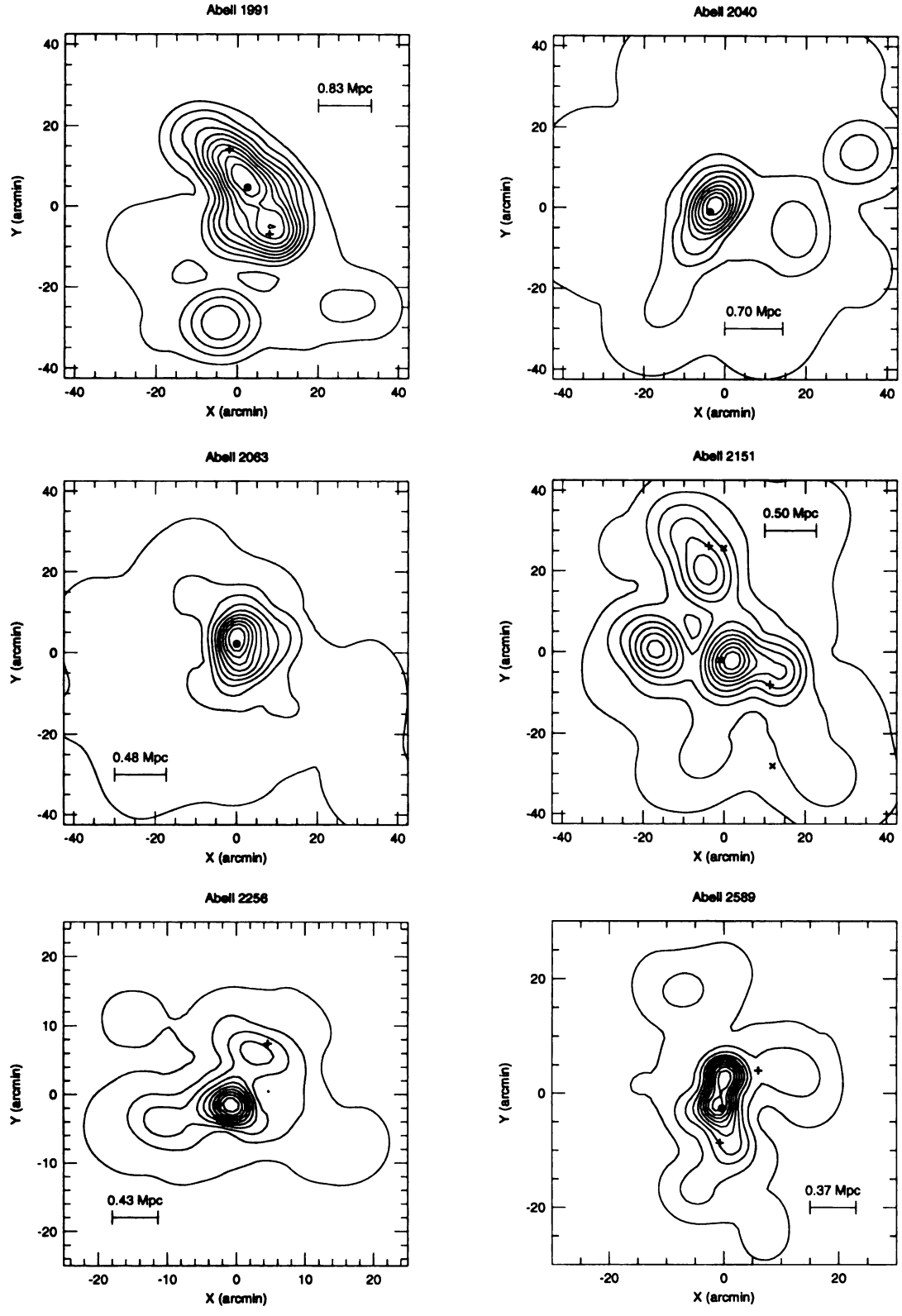

Figure 3.6 (cont'd).

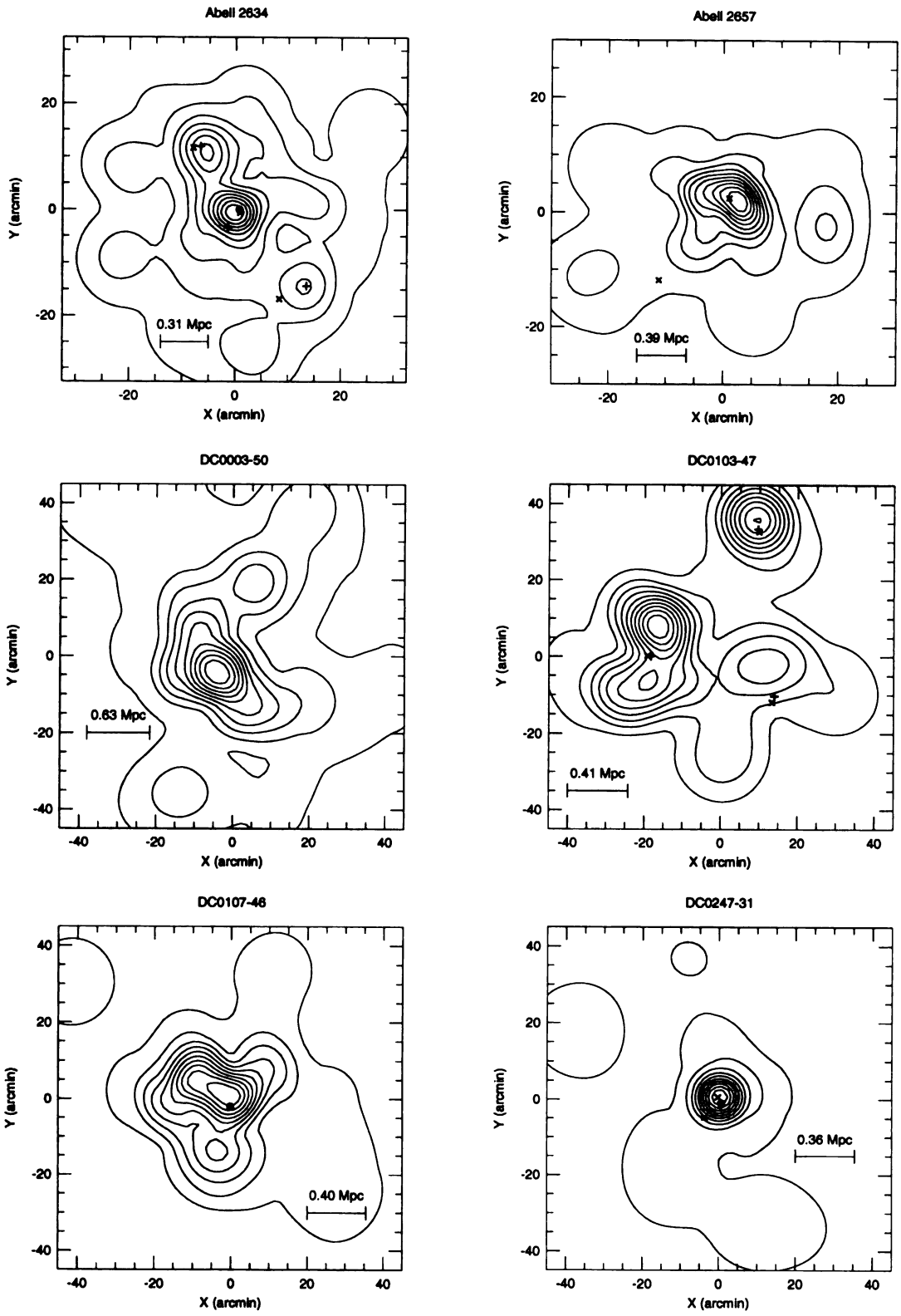

Figure 3.6 (cont'd).

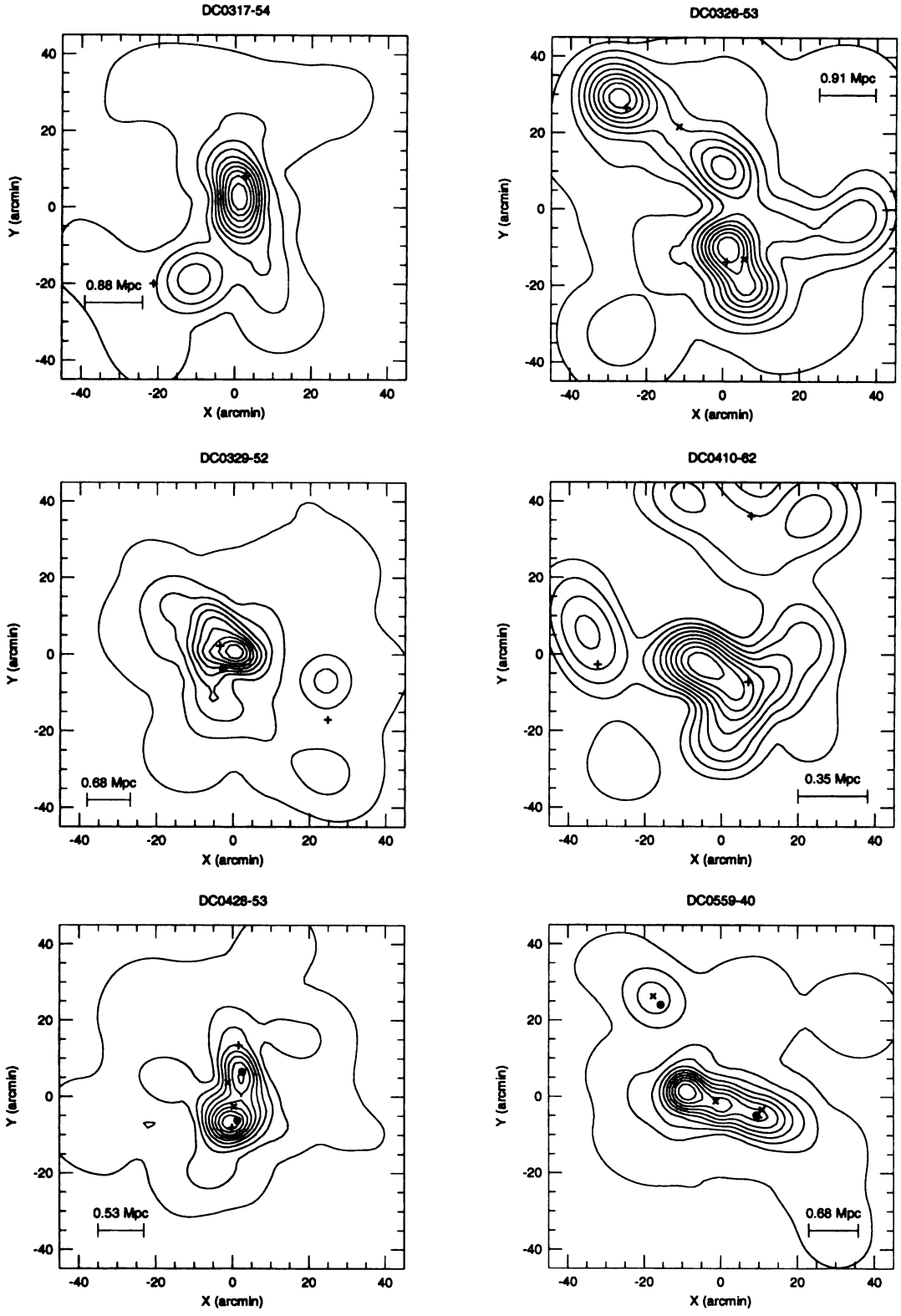

Figure 3.6 (cont'd).

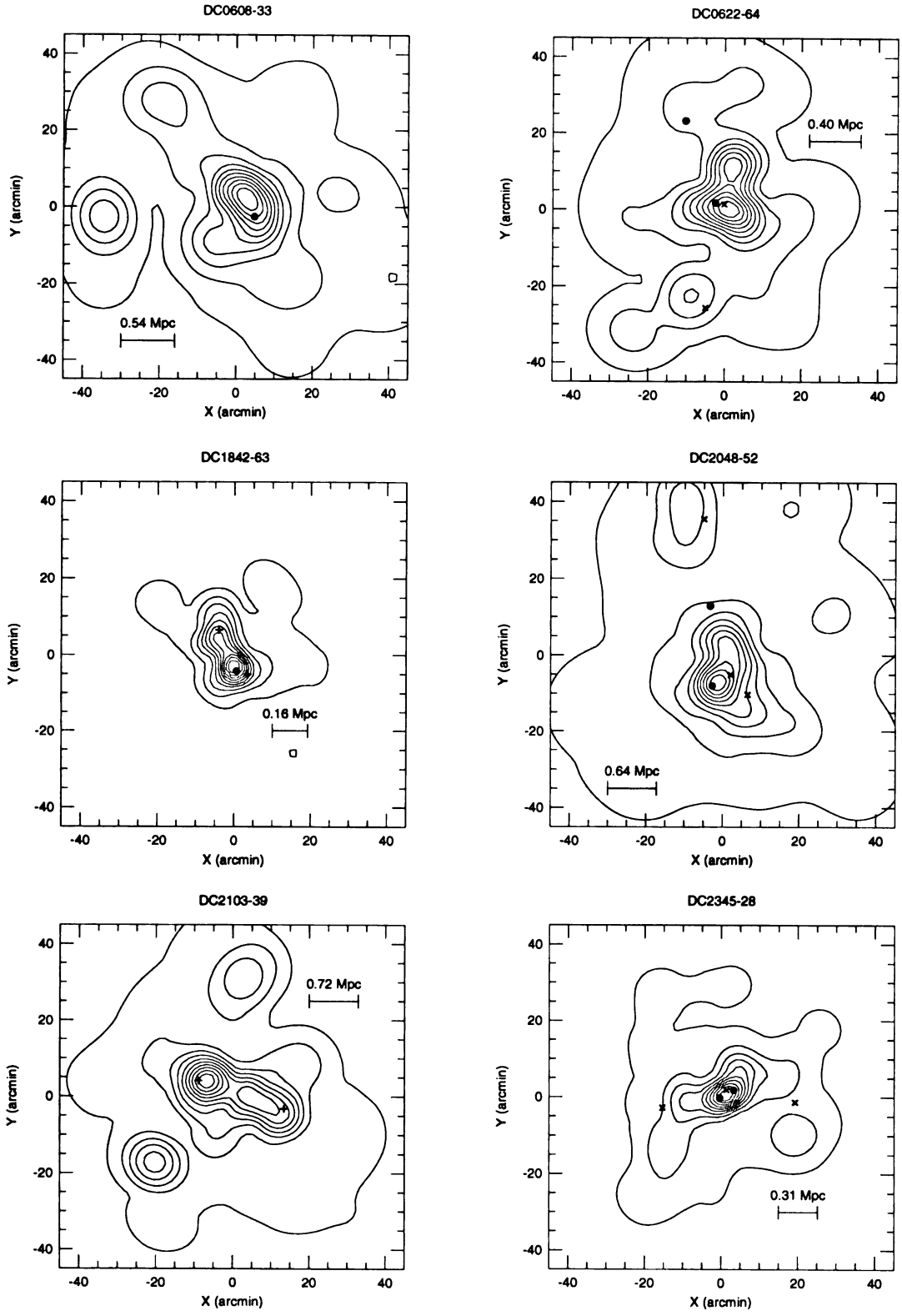

Figure 3.6 (cont'd).

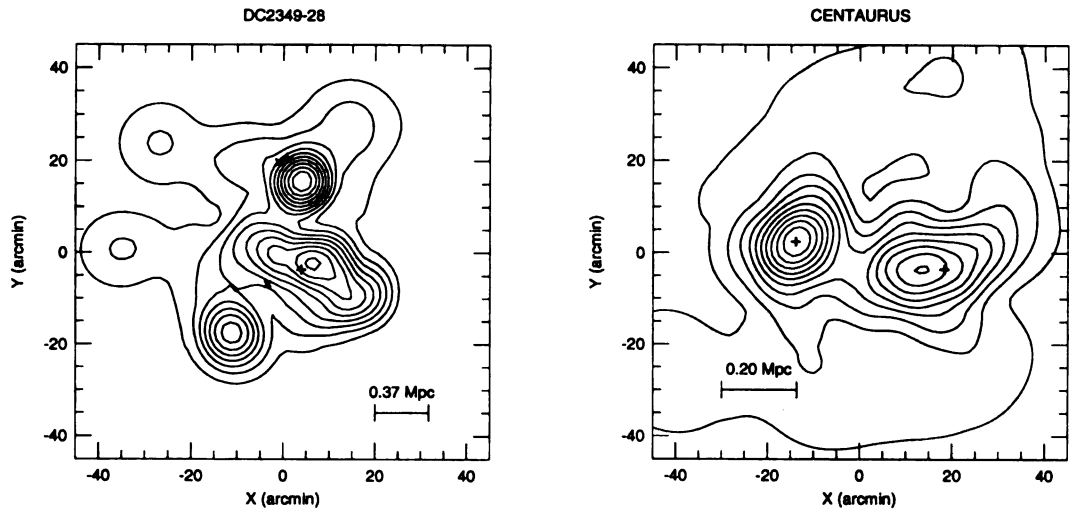

Figure 3.6 (cont'd).

For comparison, Figures 3.7 and 3.8 show four examples of the improvement of the adaptive-kernel maps over those of GB. The GB map is shown on the left, while the corresponding adaptive-kernel map (produced with the same data) is on the right. The GB maps were constructed using 50% overlapping boxes (in essence a two-dimensional ASH with m=2) with fixed window width of 0.24, 0.48 or $0.72h^{-1}$ Mpc (the scale bar at the upper left of the GB maps indicates the width used). The plots in Figure 3.7 were chosen to illustrate how details in high-density regions could easily be oversmoothed, hiding structure in the core of the clusters. In both Abell 496 and Abell 754 the GB maps were smoothed using their smallest window width of $0.24h^{-1}$ Mpc. Although Abell 754 is obviously elongated, the structure in the core is not resolved. This structure is clearly resolved in the adaptive-kernel maps, even though the initial smoothing window is larger than that used in the corresponding GB map. Figure 3.8 illustrates how undersmoothing of low-density regions can lead to "noise". Most commonly, this noise arises from small numbers of galaxies located in isolated regions on the outskirts of clusters, but with separations smaller than the size of the fixed window width. While it is possible that some of these density fluctuations in the outskirts of the clusters may be due to galaxies that are gravitationally bound to the cluster and not just Poisson noise of the background, they are very unlikely to be dynamically significant to the evolution of the cluster. Essentially all of this noise is eliminated in the adaptive-kernel maps. For comparison purposes the positions of the galaxies are plotted in the adaptive-kernel maps.

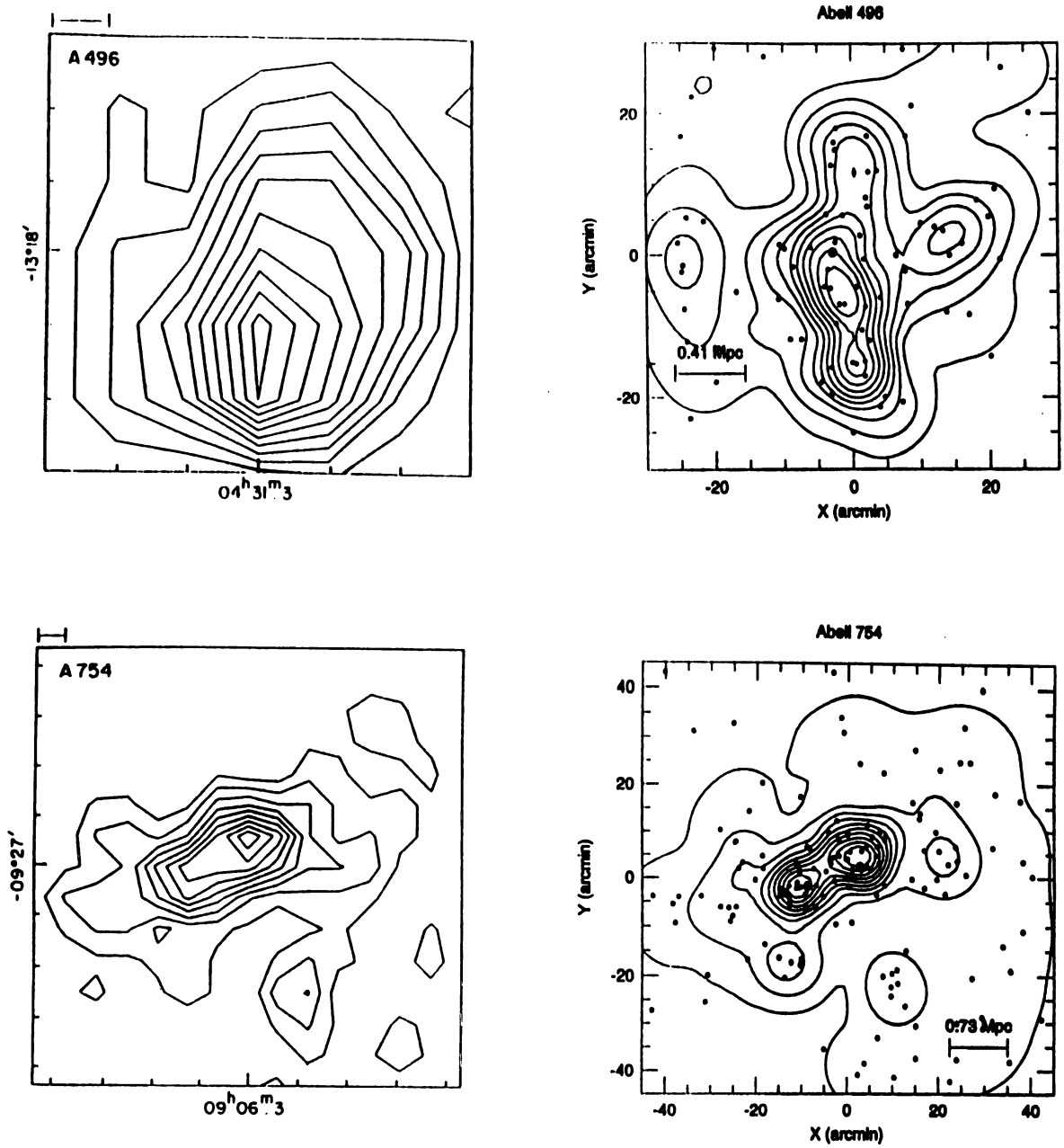

Fig. 3.7.— Comparison of adaptive-kernel maps with those of GB. In both A496 and A754 detail in the high density regions, which is oversmoothed in the GB maps, is resolved in the adaptive-kernel map.

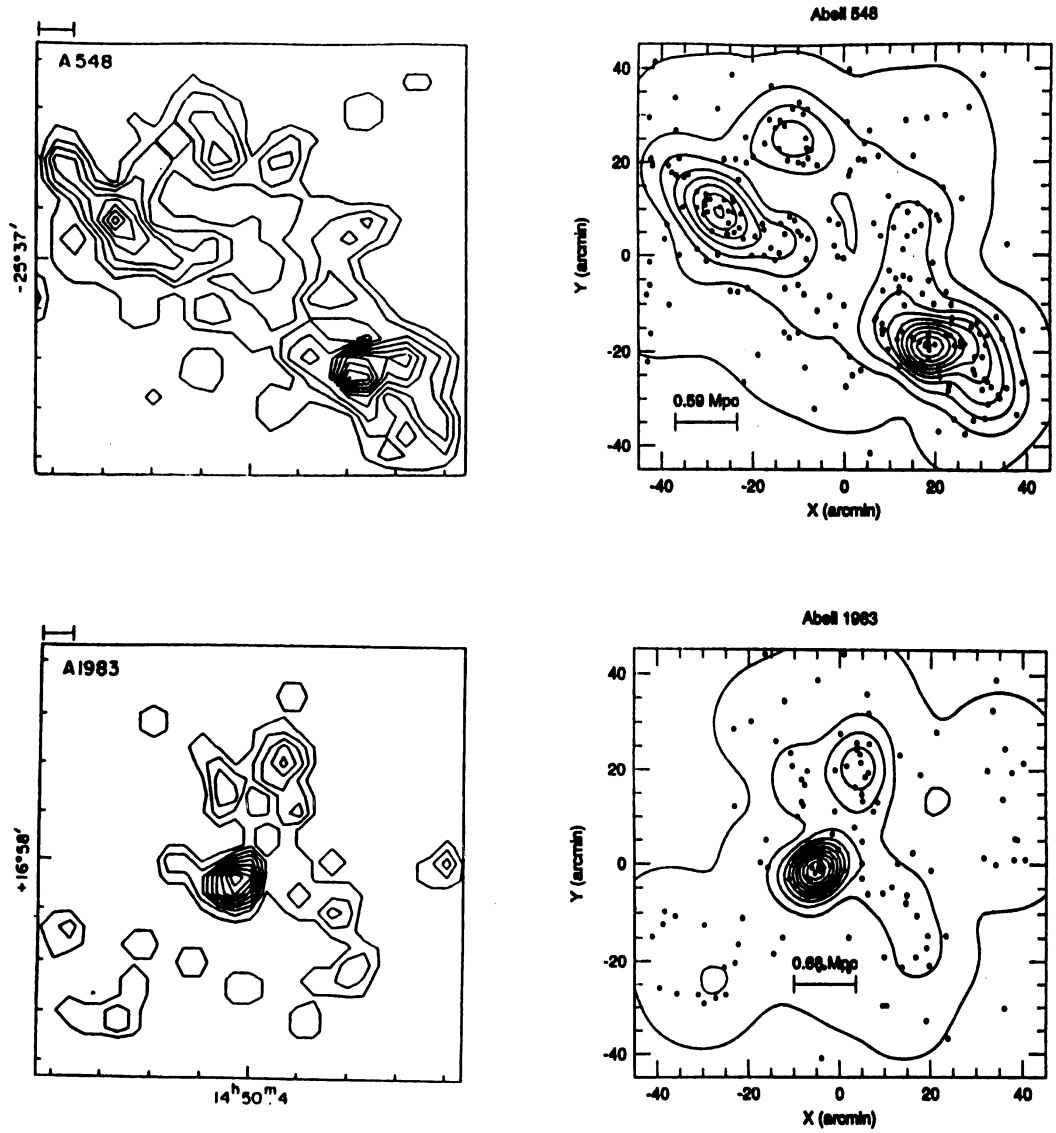

Fig. 3.8.— Comparison of adaptive-kernel maps with those of GB. In both A548 and A1983 noise in the low-density regions of the GB maps is smoothed away in the adaptive-kernel map.

Adaptive-kernel maps of the HGT sample clusters are given in Figure 3.9. Again, the crosses and plus marks indicated the positions of groups found significant by either KMM or DEDICA respectively. The cross and plus marks with squares and circles around them indicate possible foreground or background groups based on the magnitude distribution (see Chapter 4). The positions of the filled circles in this case indicate the galaxies identified as cD by Struble & Rood (1982). The scale bar indicates the size of the initial smoothing window in Mpc.

Table 3.2 gives the same quantities for the HGT sample as Table 3.1 did for the Dressler sample, except that column (5) now lists the apparent O magnitude cutoff used for each cluster. For the 25 clusters which are in common between the two samples comparison of APS data with that of Dressler shows that on average the maps made from the APS data have 33% more galaxies than in the Dressler maps. This is due to the inclusion of fainter magnitude galaxies in the maps made with the APS data. Because Dressler was interested in studying the morphology of the galaxies in these clusters, he was forced to use a brighter limiting magnitude to ensure accurate classification.

TABLE 3.2. Adaptive Kernel Map Parameters – HGT Clusters

Cluster	N	RA (1950)	DEC (1950)	m_{lim}	σ_{max} (gal/arcmin^2)	σ_{min} (gal/arcmin^2)	$\delta\sigma$ (gal/arcmin^2)	$0.75h_n$ arcmin	Mpc
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A 21	291	00 17 54	+28 22	21.3	1.5684	0.0327	0.1536	5.4	0.52
A 76	195	00 37 12	+06 30	19.5	0.1303	0.0179	0.0125	13.3	0.61
A 85	323	00 39 06	−09 38	20.0	0.3758	0.0089	0.0367	10.3	0.58
A 88	72	00 40 24	−26 20	21.6	0.4315	0.0247	0.0407	6.2	0.67
A 104	151	00 47 06	+24 15	21.0	0.7542	0.0198	0.0734	6.9	0.59
A 119	268	00 53 48	−01 32	19.6	0.3068	0.0074	0.0299	11.4	0.54
A 121	145	00 55 00	−07 17	21.5	0.4239	0.0655	0.0398	6.1	0.64
A 147	155	01 05 36	+01 55	19.6	0.1804	0.0058	0.0175	12.8	0.61
A 151	342	01 06 24	−15 41	20.0	0.5542	0.0103	0.0544	9.6	0.55
A 154	272	01 08 18	+17 24	20.5	0.7189	0.0137	0.0705	8.4	0.58
A 166	157	01 12 06	−16 33	21.7	0.9070	0.0327	0.0874	5.4	0.62
A 168	235	01 12 36	−00 02	19.7	0.3498	0.0064	0.0343	12.2	0.60
A 193	264	01 22 30	+08 27	19.8	0.4063	0.0067	0.0400	11.9	0.62
A 194	129	01 23 00	−01 46	17.6	0.0249	0.0008	0.0024	35.2	0.71
A 225	202	01 36 12	+18 38	20.6	0.4014	0.0150	0.0386	8.0	0.58
A 246	103	01 42 06	+05 34	20.6	0.1060	0.0092	0.0097	10.2	0.75
A 274	174	01 52 12	−06 32	21.9	0.7451	0.1131	0.0702	4.7	0.59
A 277	230	01 53 18	−07 38	21.3	1.0577	0.0288	0.1029	5.8	0.55
A 389	173	02 49 06	−25 07	21.7	1.0089	0.1360	0.0970	4.9	0.57
A 399	254	02 55 12	+12 49	20.7	0.4224	0.0157	0.0407	7.8	0.58
A 400	190	02 55 00	+05 50	18.2	0.0637	0.0020	0.0062	22.1	0.57
A 401	288	02 56 12	+13 23	20.8	0.7803	0.0210	0.0759	6.7	0.53
A 415	243	03 04 24	−12 15	20.9	0.8260	0.0222	0.0804	6.6	0.54
A 496	226	04 31 18	−13 22	18.9	0.1422	0.0037	0.0138	16.1	0.57
A 500	225	04 36 48	−22 12	20.5	0.8030	0.0165	0.0787	7.6	0.53
A 514	282	04 45 30	−20 32	20.7	0.6142	0.0171	0.0597	7.5	0.57
A 634	71	08 10 30	+58 12	18.5	0.0506	0.0015	0.0049	25.1	0.75
A 671	293	08 25 24	+30 36	19.9	0.6408	0.0084	0.0632	10.6	0.57
A 779	115	09 16 48	+33 59	18.2	0.0238	0.0034	0.0023	29.0	0.74
A 787	154	09 23 30	+74 38	22.0	1.2019	0.0474	0.1155	4.5	0.59

TABLE 3.2. (continued)

Cluster	N	RA (1950)	DEC (1950)	m_{lim}	σ_{max} (gal/arcmin^2)	σ_{min} (gal/arcmin^2)	$\delta\sigma$ (gal/arcmin^2)	$0.75h_n$ arcmin	$0.75h_n$ Mpc
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A 957	288	10 11 24	−00 40	19.6	0.3575	0.0059	0.0352	12.7	0.61
A 978	295	10 18 00	−06 17	20.0	0.5138	0.0104	0.0503	9.6	0.54
A 993	272	10 19 24	−04 43	20.0	0.2974	0.0076	0.0290	11.2	0.64
A 1020	265	10 25 12	+10 40	20.4	0.6698	0.0133	0.0656	8.5	0.58
A 1035	283	10 29 12	+40 29	20.9	0.9445	0.0231	0.0921	6.4	0.53
A 1126	248	10 51 18	+17 08	21.0	1.0821	0.0231	0.1059	6.4	0.56
A 1139	168	10 55 30	+01 47	19.3	0.1089	0.0040	0.0105	15.4	0.65
A 1185	335	11 08 06	+28 57	18.8	0.3511	0.0037	0.0347	16.1	0.54
A 1187	227	11 08 54	+39 51	20.9	0.4888	0.0197	0.0469	7.0	0.57
A 1213	261	11 13 48	+29 33	19.7	0.3573	0.0072	0.0350	11.6	0.59
A 1216	102	11 15 12	−04 12	20.0	0.0708	0.0059	0.0065	12.7	0.72
A 1228	278	11 18 48	+34 37	19.1	0.1235	0.0037	0.0120	16.0	0.62
A 1238	180	11 20 24	+01 23	20.7	0.5417	0.0120	0.0530	8.9	0.67
A 1254	263	11 23 48	+71 22	20.4	0.5248	0.0125	0.0512	8.7	0.58
A 1257	212	11 23 24	+35 37	19.0	0.1002	0.0028	0.0097	18.6	0.70
A 1291	395	11 29 18	+56 19	20.0	0.6750	0.0113	0.0664	9.2	0.52
A 1318	286	11 33 42	+55 15	20.1	0.2214	0.0730	0.0212	10.1	0.61
A 1364	226	11 41 06	−01 30	21.5	1.0809	0.0306	0.1050	5.6	0.60
A 1365	163	11 41 48	+31 11	20.8	0.7088	0.0166	0.0692	7.6	0.60
A 1367	200	11 41 54	+20 07	18.0	0.0796	0.0024	0.0077	19.9	0.48
A 1377	402	11 44 18	+56 01	19.9	0.3694	0.0086	0.0361	10.5	0.59
A 1382	183	11 45 36	+71 43	21.5	0.7013	0.1629	0.0673	5.8	0.61
A 1383	284	11 45 30	+54 54	20.3	0.4082	0.0909	0.0397	9.1	0.58
A 1412	244	11 53 06	+73 45	21.0	1.0375	0.0212	0.1016	6.7	0.58
A 1436	358	11 57 54	+56 32	20.4	0.8415	0.0162	0.0825	7.7	0.52
A 1468	166	12 03 06	+51 42	21.0	0.3701	0.0884	0.0352	7.3	0.64
A 1474	187	12 05 24	+15 14	20.9	0.3826	0.0175	0.0365	7.4	0.61
A 1496	355	12 10 54	+59 33	21.3	1.1066	0.0275	0.1079	5.9	0.56
A 1541	205	12 24 54	+09 07	21.1	0.9976	0.0277	0.0970	5.9	0.54
A 1644	297	12 54 36	−17 06	19.6	0.3113	0.0077	0.0304	11.1	0.54

TABLE 3.2. (continued)

Cluster	N	RA (1950)	DEC (1950)	m_{lim}	σ_{max} (gal/arcmin^2)	σ_{min} (gal/arcmin^2)	$\delta\sigma$ (gal/arcmin^2)	$0.75h_n$ arcmin	$0.75h_n$ Mpc
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A 1651	205	12 56 48	−03 56	21.0	0.7909	0.0200	0.0771	6.9	0.59
A 1656	424	12 57 24	+28 15	18.2	0.1638	0.0023	0.0162	20.4	0.53
A 1691	247	13 09 06	+39 29	20.7	0.4949	0.0151	0.0480	7.9	0.60
A 1749	219	13 27 18	+37 53	20.2	0.4798	0.0105	0.0469	9.5	0.60
A 1767	308	13 34 12	+59 29	20.6	0.6229	0.0178	0.0605	7.3	0.54
A 1773	282	13 39 36	+02 30	20.8	1.3750	0.0238	0.1351	6.3	0.51
A 1775	268	13 39 36	+26 37	20.6	0.4372	0.0164	0.0421	7.6	0.56
A 1793	248	13 46 06	+32 32	21.0	1.3789	0.0234	0.1356	6.4	0.56
A 1795	288	13 46 42	+26 51	20.3	0.9357	0.0133	0.0922	8.5	0.55
A 1809	308	13 50 48	+05 25	20.9	1.1872	0.0229	0.1164	6.5	0.53
A 1831	308	13 56 54	+28 14	20.7	0.6616	0.0178	0.0644	7.3	0.56
A 1837	268	13 59 06	−10 56	19.3	0.1546	0.0047	0.0150	14.2	0.59
A 1904	386	14 20 18	+48 48	20.6	0.7130	0.0173	0.0696	7.4	0.55
A 1913	276	14 24 30	+16 54	20.0	0.4321	0.0101	0.0422	9.7	0.56
A 1927	245	14 28 48	+25 54	20.7	0.5220	0.0173	0.0505	7.4	0.58
A 1983	439	14 50 24	+16 57	19.6	0.4065	0.0071	0.0399	11.6	0.56
A 1991	368	14 52 12	+18 51	20.2	0.3843	0.0108	0.0373	9.4	0.59
A 1999	187	14 52 36	+54 32	21.5	1.1878	0.0346	0.1153	5.3	0.54
A 2005	139	14 56 36	+28 01	21.9	0.8724	0.0396	0.0833	4.9	0.60
A 2022	322	15 02 12	+28 38	20.1	0.8781	0.0122	0.0866	8.9	0.53
A 2028	231	15 07 06	+07 43	20.8	0.8320	0.0168	0.0815	7.5	0.60
A 2029	437	15 08 30	+05 57	20.8	0.9588	0.0211	0.0938	6.7	0.54
A 2040	278	15 10 18	+07 37	19.7	0.2760	0.0063	0.0270	12.3	0.61
A 2048	314	15 12 48	+04 35	21.3	1.4985	0.0301	0.1468	5.6	0.54
A 2052	270	15 14 18	+07 12	19.1	0.1235	0.0035	0.0120	16.6	0.64
A 2052	270	15 14 18	+07 12	19.1	0.1235	0.0035	0.0120	16.6	0.64
A 2061	285	15 19 12	+30 50	20.8	0.8732	0.0215	0.0852	6.7	0.53
A 2063	211	15 20 36	+08 49	19.0	0.1869	0.0033	0.0184	17.1	0.64
A 2065	422	15 20 36	+27 54	20.7	1.4274	0.0191	0.1408	7.1	0.54
A 2067	283	15 21 12	+31 06	20.7	0.5411	0.0146	0.0526	8.1	0.62

TABLE 3.2. (continued)

Cluster	N	RA (1950)	DEC (1950)	m_{lim}	σ_{max} (gal/arcmin^2)	σ_{min} (gal/arcmin^2)	$\delta\sigma$ (gal/arcmin^2)	$0.75h_n$ arcmin	$0.75h_n$ Mpc
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A 2079	318	15 26 00	+29 03	20.5	0.4389	0.0127	0.0426	8.7	0.61
A 2089	158	15 30 36	+28 12	20.7	0.3427	0.0155	0.0327	7.9	0.61
A 2092	267	15 31 18	+31 20	20.5	0.4496	0.0132	0.0436	8.5	0.60
A 2107	275	15 37 36	+21 56	19.5	0.3421	0.0052	0.0337	13.6	0.62
A 2124	298	15 43 06	+36 14	20.5	0.7221	0.0133	0.0709	8.5	0.59
A 2142	311	15 56 12	+27 22	21.2	0.7166	0.1634	0.0691	6.2	0.57
A 2147	465	16 00 00	+16 03	19.1	0.2334	0.0504	0.0229	14.4	0.56
A 2151	388	16 02 60	+17 53	19.2	0.2504	0.0050	0.0245	13.9	0.57
A 2152	471	16 03 06	+16 35	19.2	0.1613	0.0362	0.0156	13.9	0.57
A 2162	124	16 10 30	+29 40	18.9	0.1091	0.0030	0.0106	17.8	0.63
A 2175	448	16 18 24	+30 02	21.3	2.1481	0.0366	0.2112	5.1	0.50
A 2197	313	16 26 30	+41 01	18.8	0.0960	0.0218	0.0093	17.3	0.58
A 2199	389	16 26 54	+39 38	18.8	0.1990	0.0036	0.0195	16.3	0.55
A 2255	417	17 12 12	+64 09	20.9	1.2430	0.0263	0.1217	6.0	0.50
A 2256	451	17 06 36	+78 47	20.3	1.2692	0.0176	0.1252	7.4	0.47
A 2328	131	20 45 24	−18 00	22.2	1.0706	0.0557	0.1015	4.1	0.58
A 2347	90	21 26 42	−22 26	21.8	0.4459	0.1138	0.0415	5.6	0.66
A 2382	197	21 49 18	−15 53	20.4	0.4651	0.0151	0.0450	8.0	0.55
A 2384	127	21 49 30	−19 47	21.3	0.9371	0.0261	0.0911	6.1	0.58
A 2399	256	21 54 54	−08 02	20.2	0.3516	0.0791	0.0341	9.3	0.58
A 2410	235	21 59 24	−10 09	20.9	0.5376	0.0198	0.0518	6.9	0.58
A 2457	248	22 33 18	+01 13	20.3	0.5639	0.0126	0.0551	8.7	0.55
A 2634	411	23 35 48	+26 46	18.9	0.1323	0.0031	0.0129	17.5	0.61
A 2657	171	23 42 18	+08 53	19.5	0.2029	0.0047	0.0198	14.3	0.65
A 2666	171	23 48 24	+26 53	18.5	0.0271	0.0067	0.0025	24.5	0.73
A 2670	255	23 51 36	−10 41	20.7	0.9085	0.0194	0.0889	7.0	0.55
A 2675	182	23 52 60	+11 10	20.7	0.4681	0.0163	0.0452	7.6	0.58
A 2700	129	00 01 18	+01 48	21.3	0.4560	0.0662	0.0433	6.5	0.64

Fig. 3.9.— Adaptive kernel maps of the HGT Sample. Map parameters are listed in Table 3.2

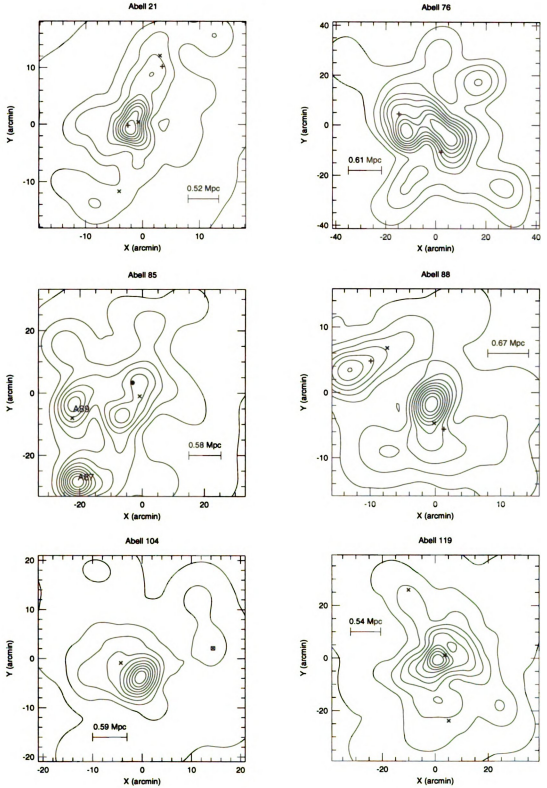

Figure 3.9. Adaptive-Kernel Maps for the HGT Clusters

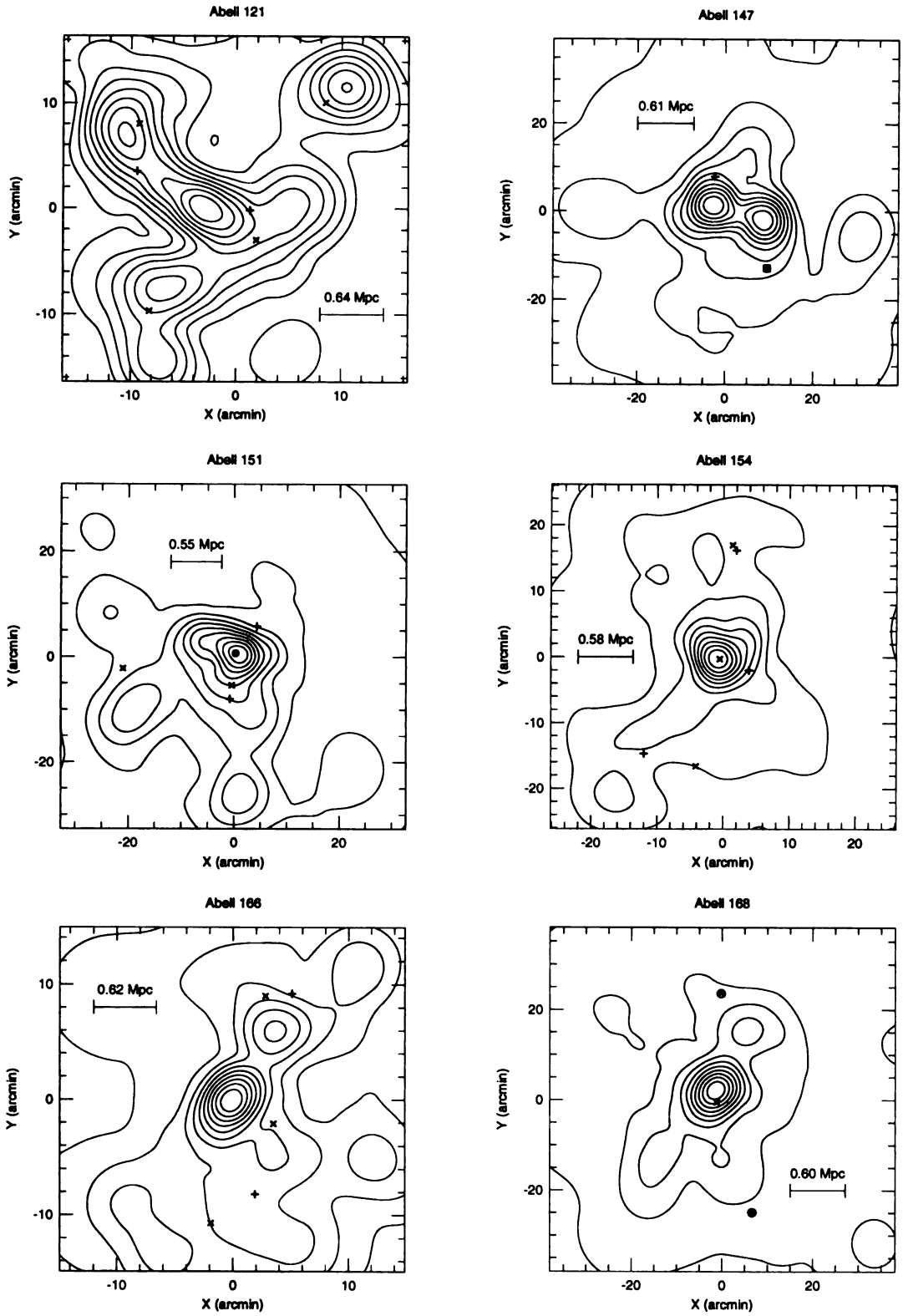

Figure 3.9 (cont'd).

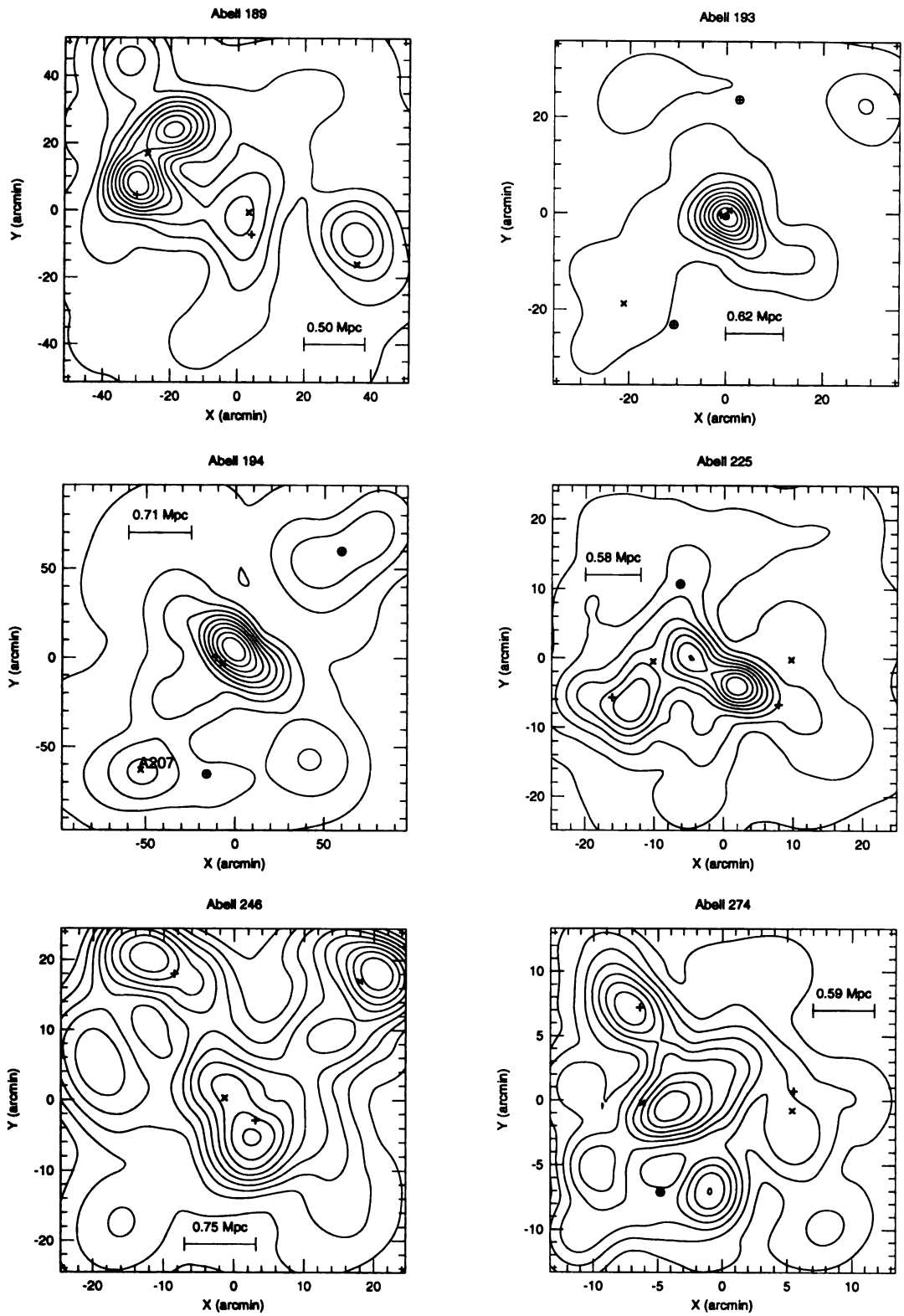

Figure 3.9 (cont'd).

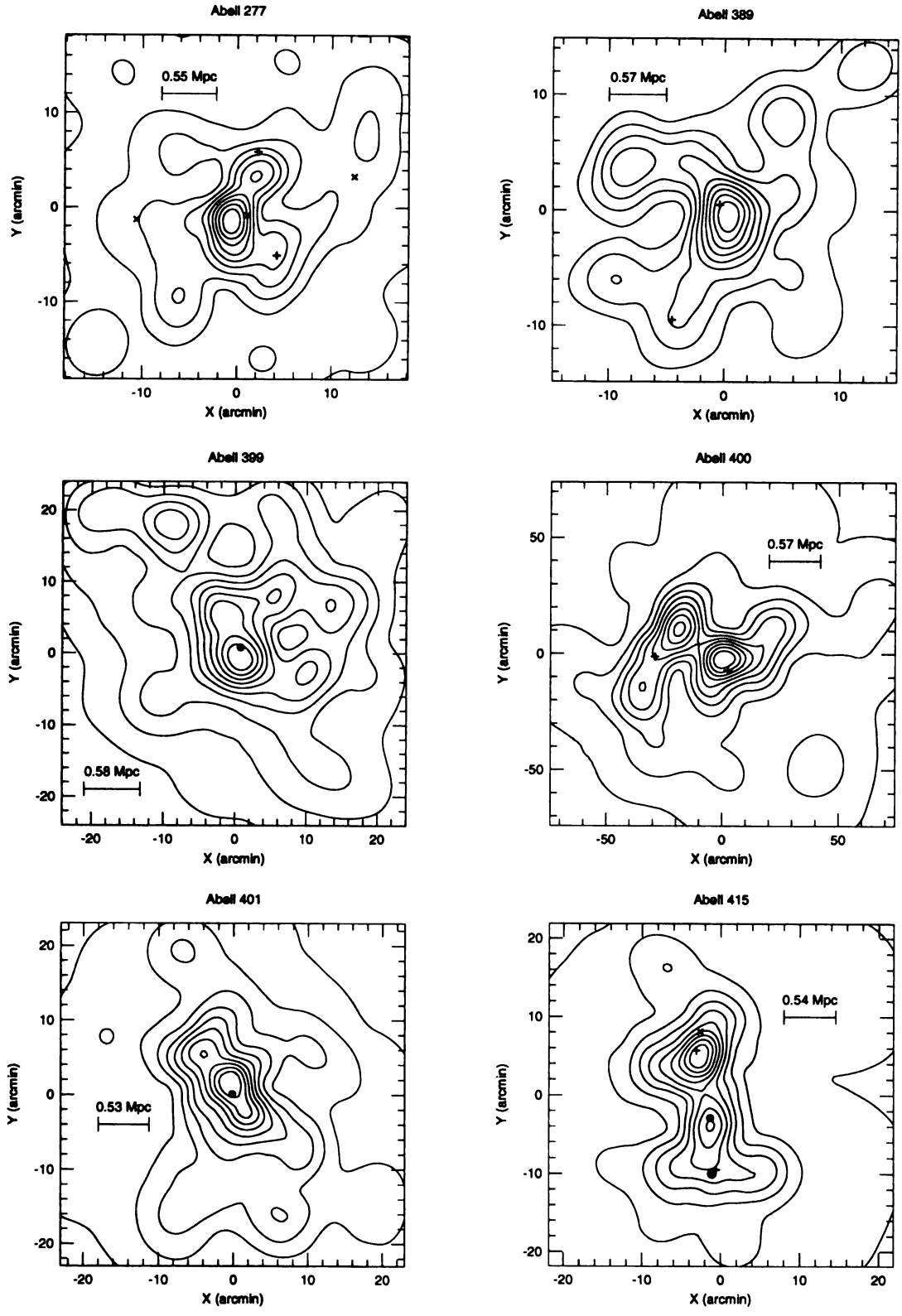

Figure 3.9 (cont'd).

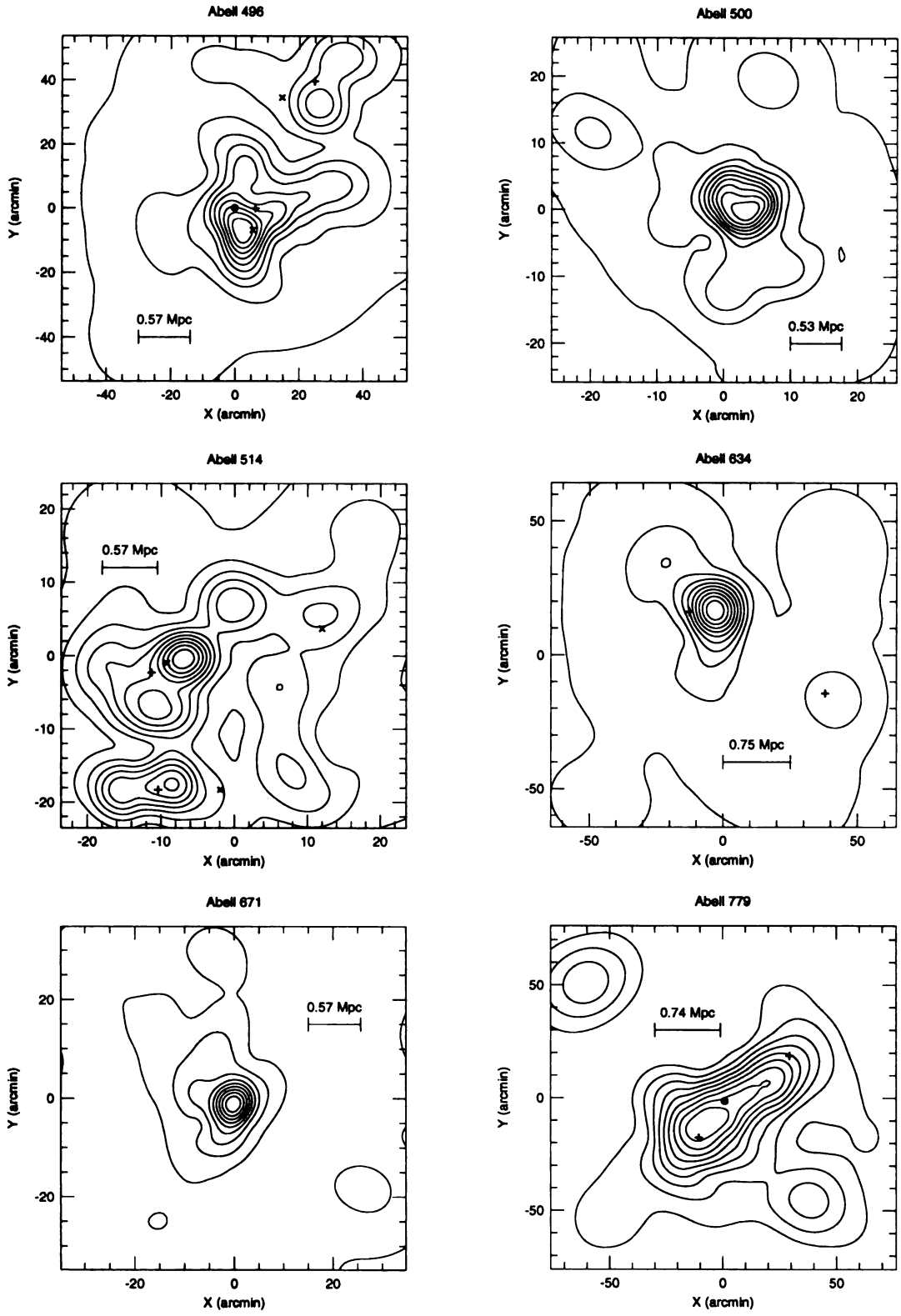

Figure 3.9 (cont'd).

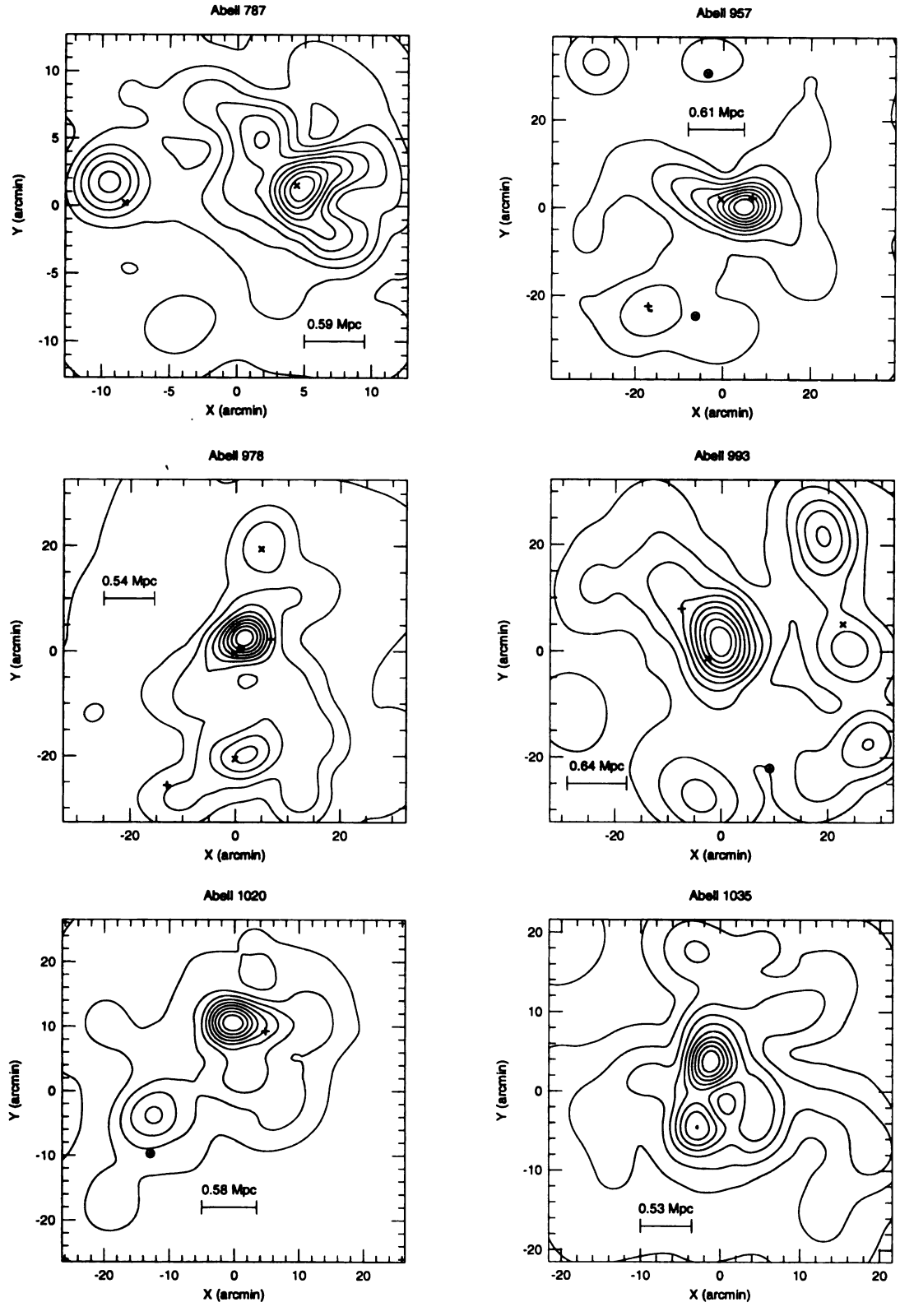

Figure 3.9 (cont'd).

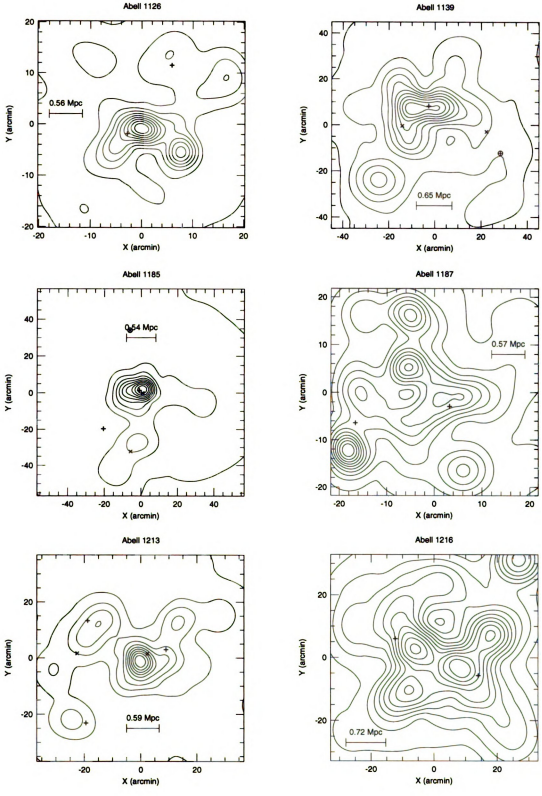

Figure 3.9 (cont'd).

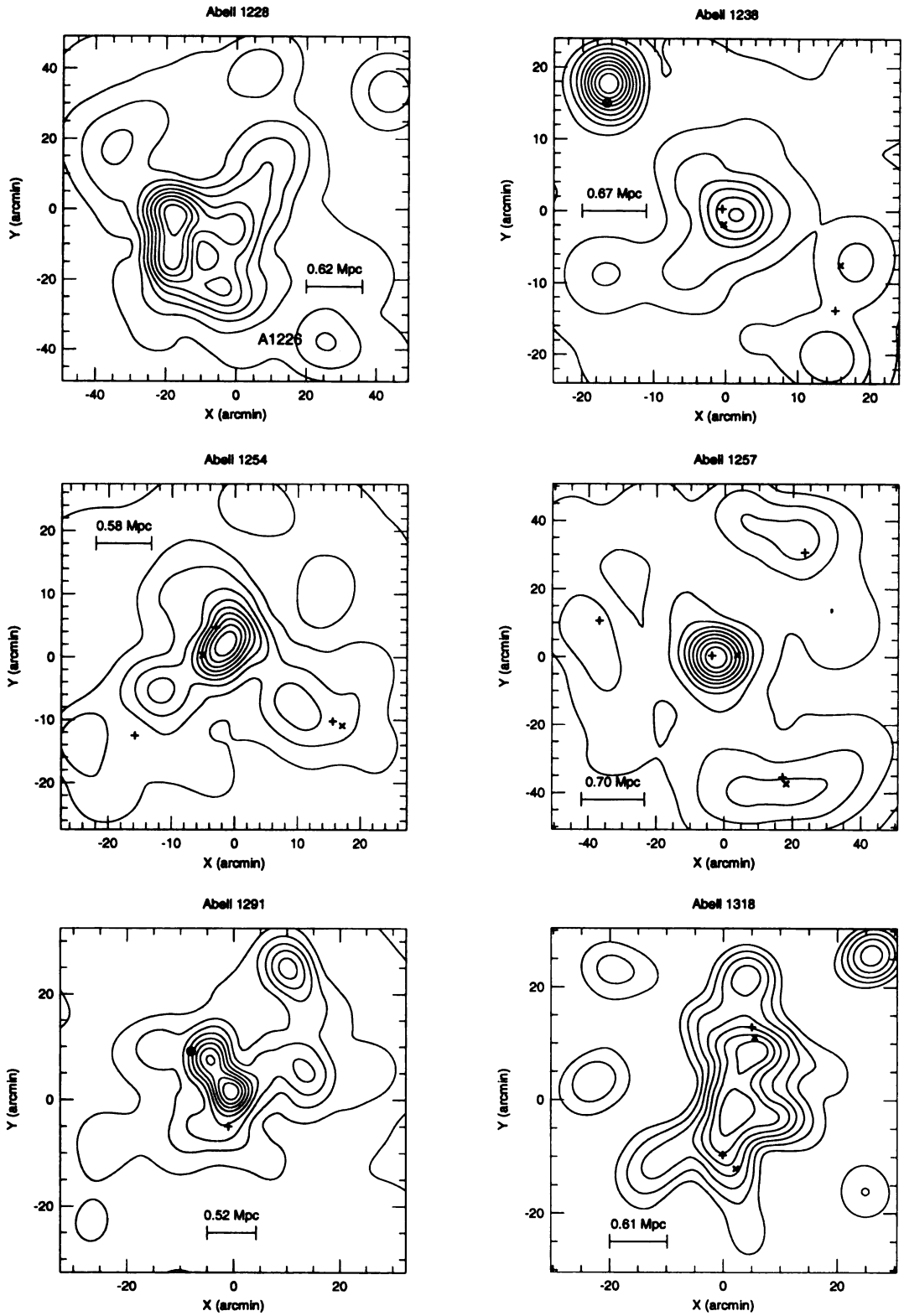

Figure 3.9 (cont'd).

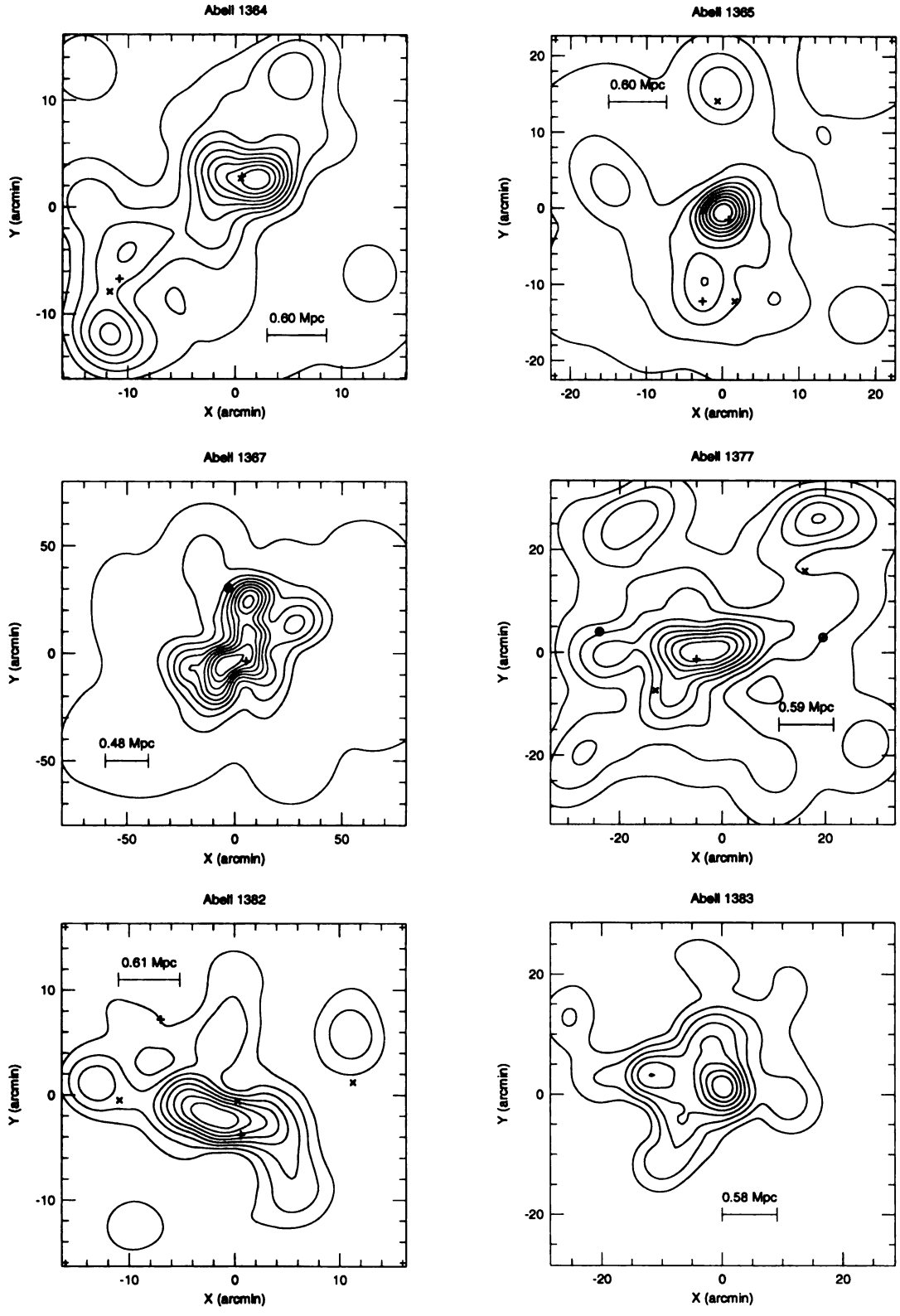

Figure 3.9 (cont'd).

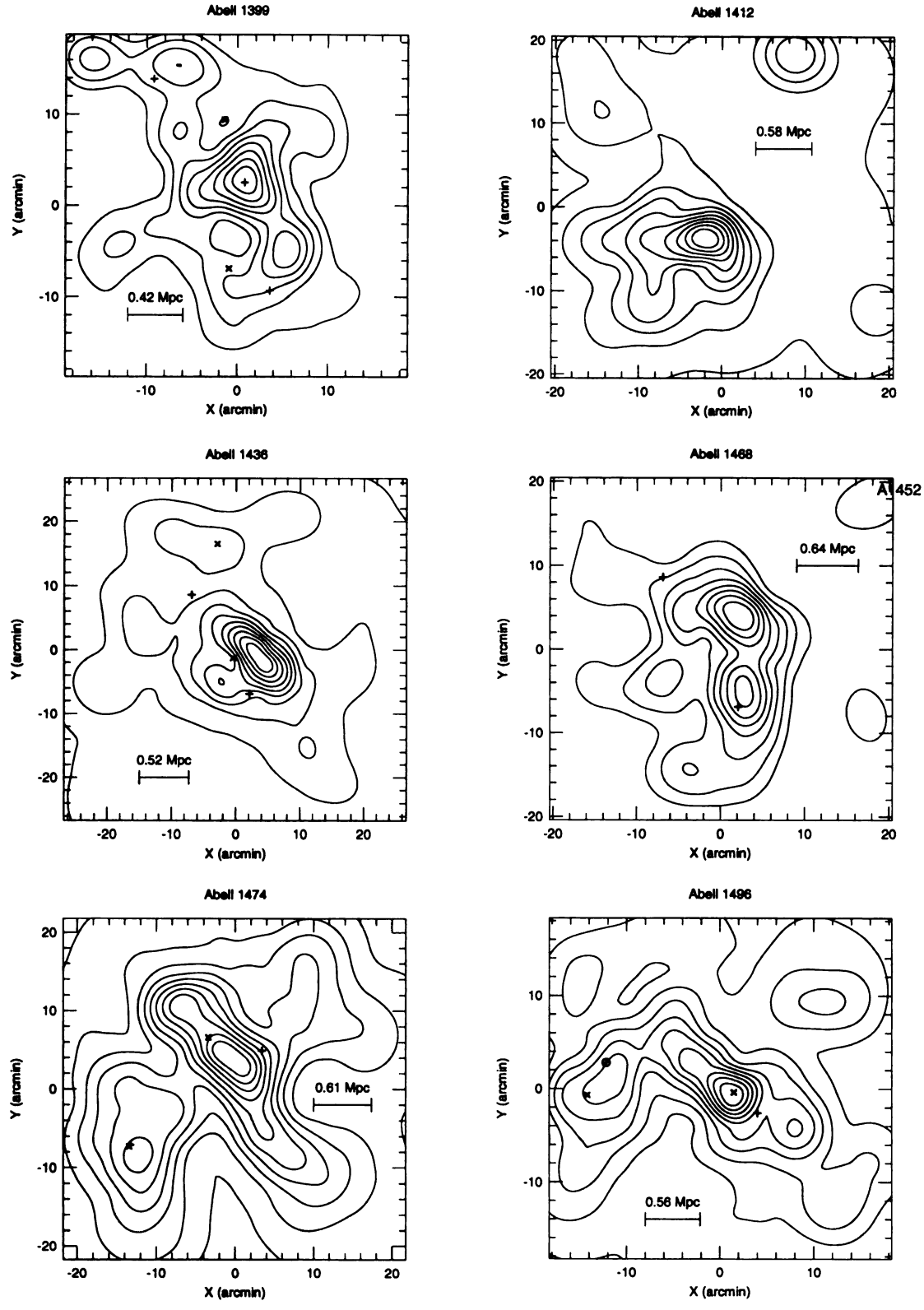

Figure 3.9 (cont'd).

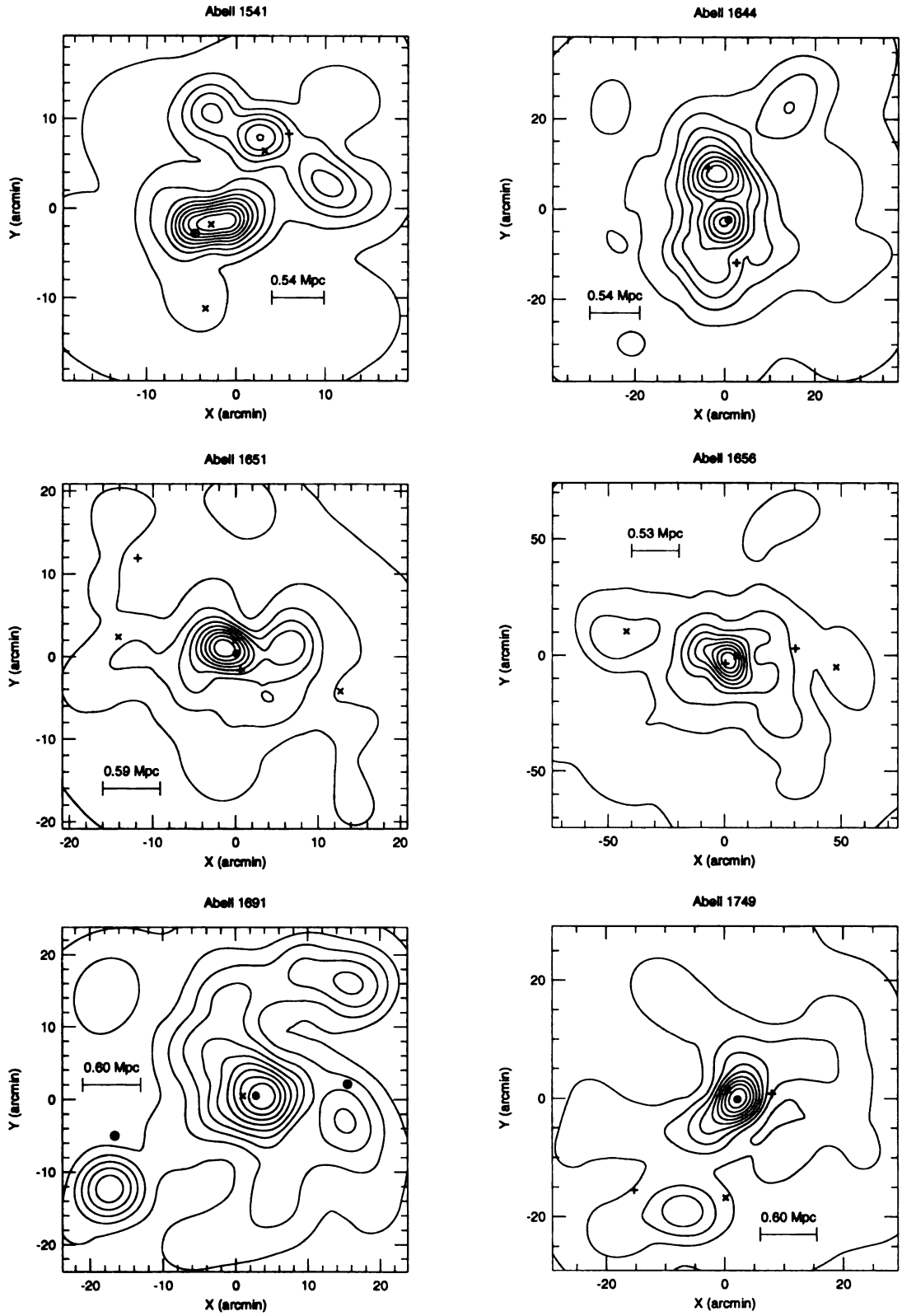

Figure 3.9 (cont'd).

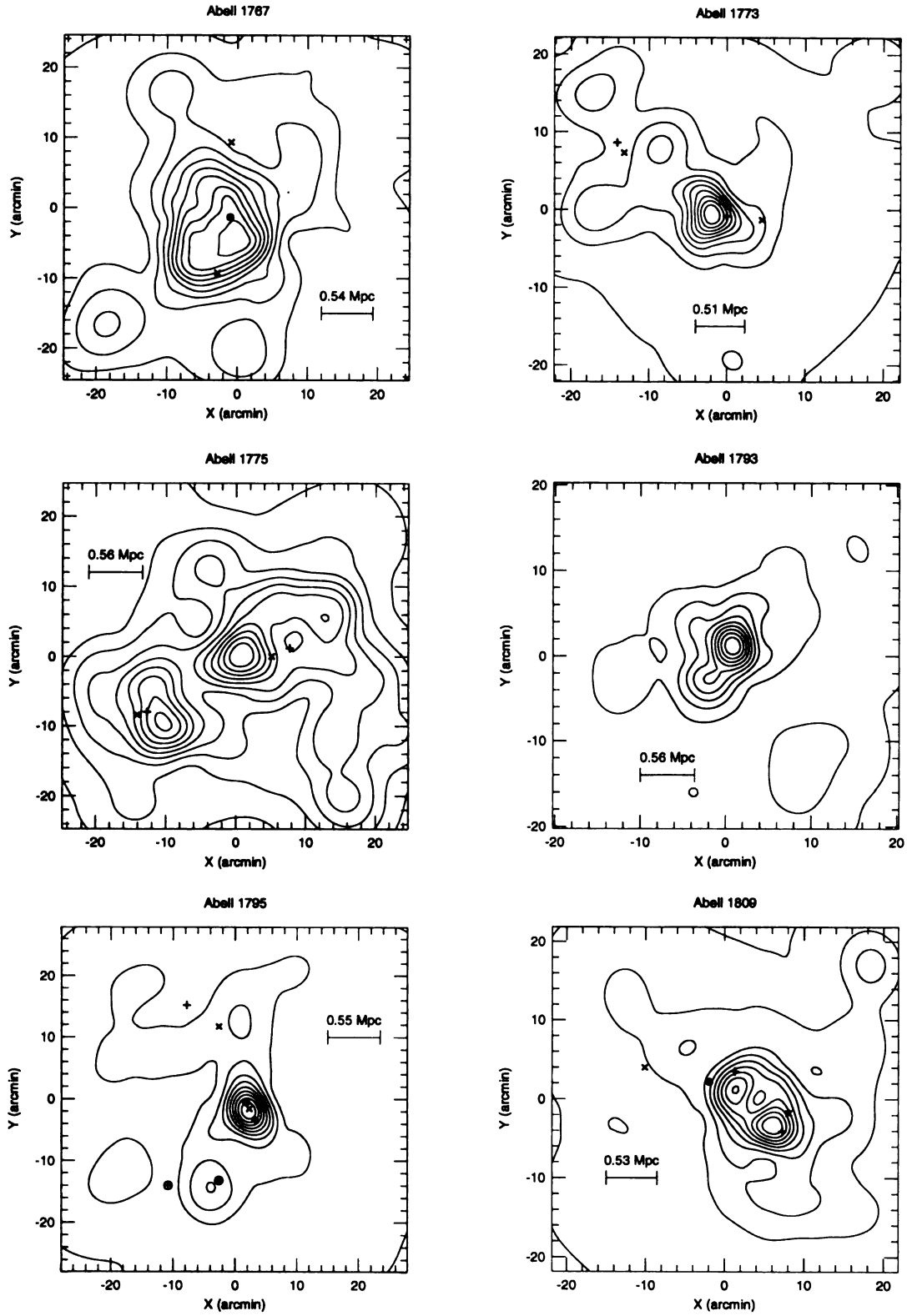

Figure 3.9 (cont'd).

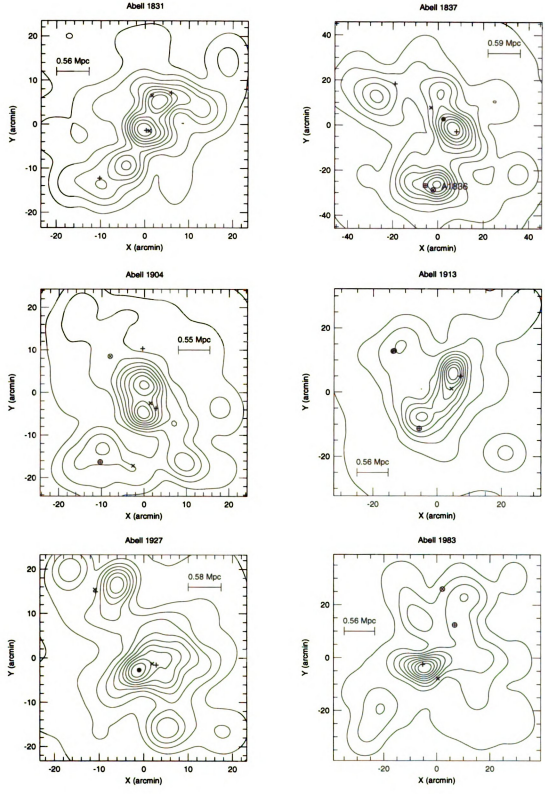

Figure 3.9 (cont'd).

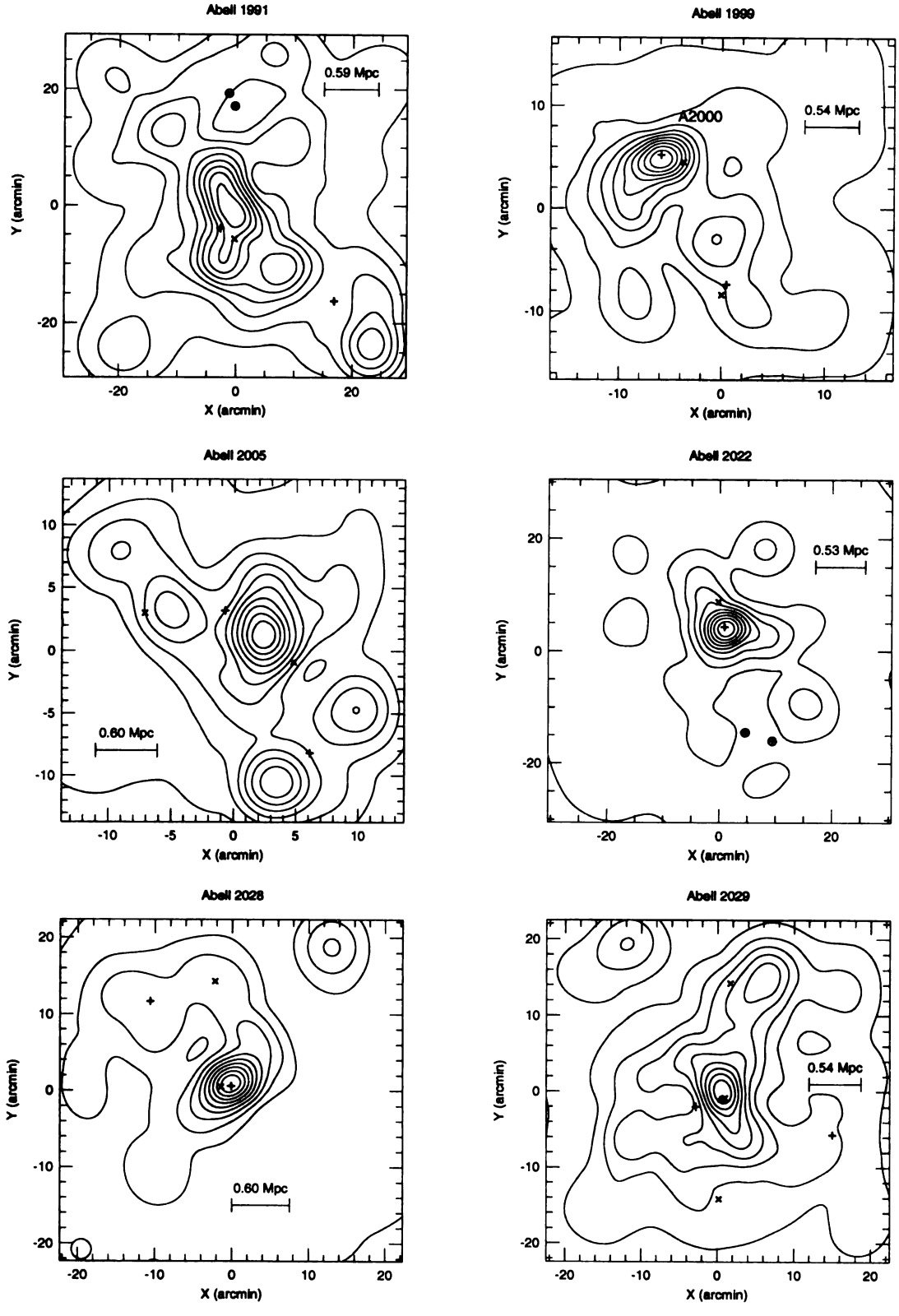

Figure 3.9 (cont'd).

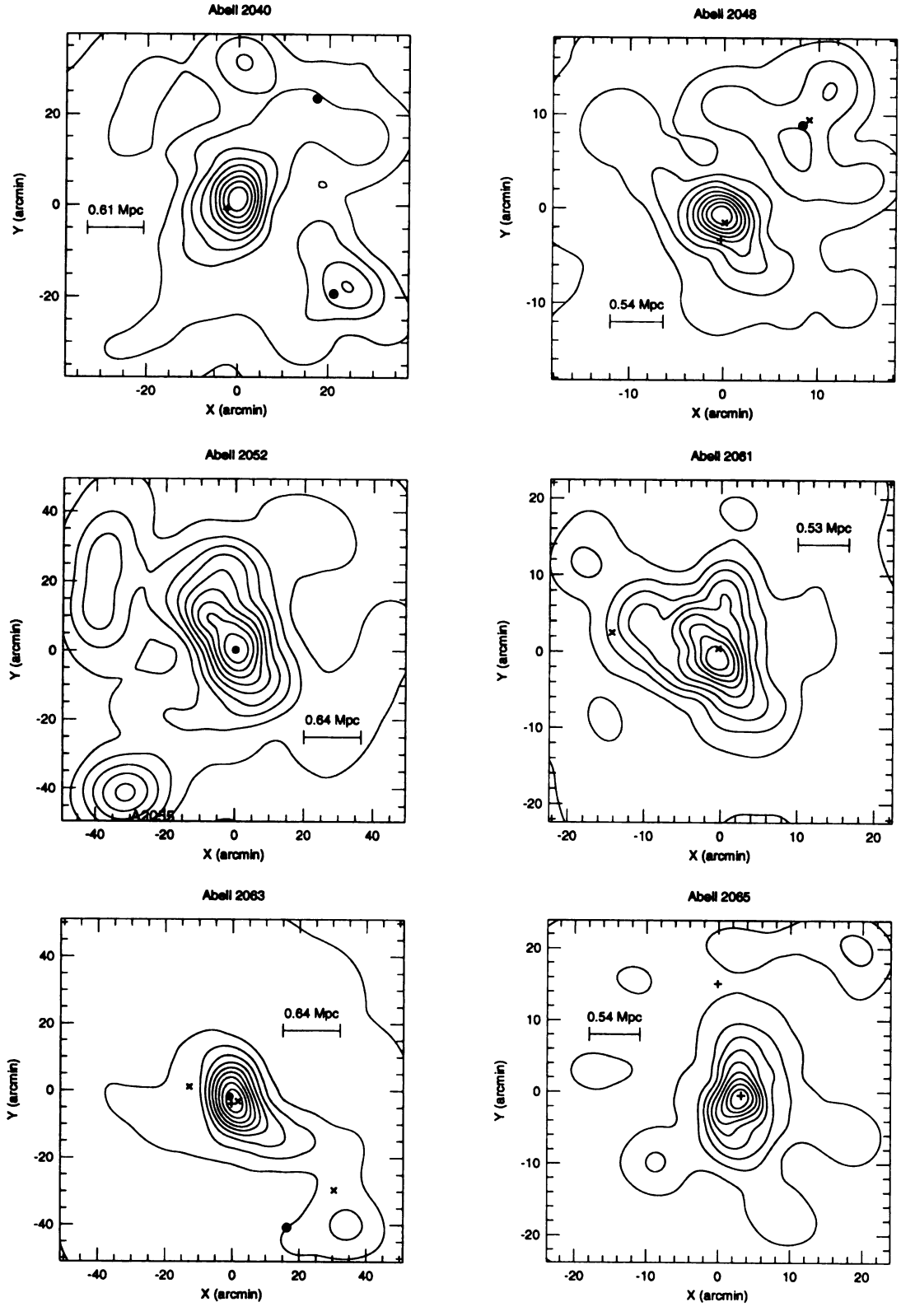

Figure 3.9 (cont'd).

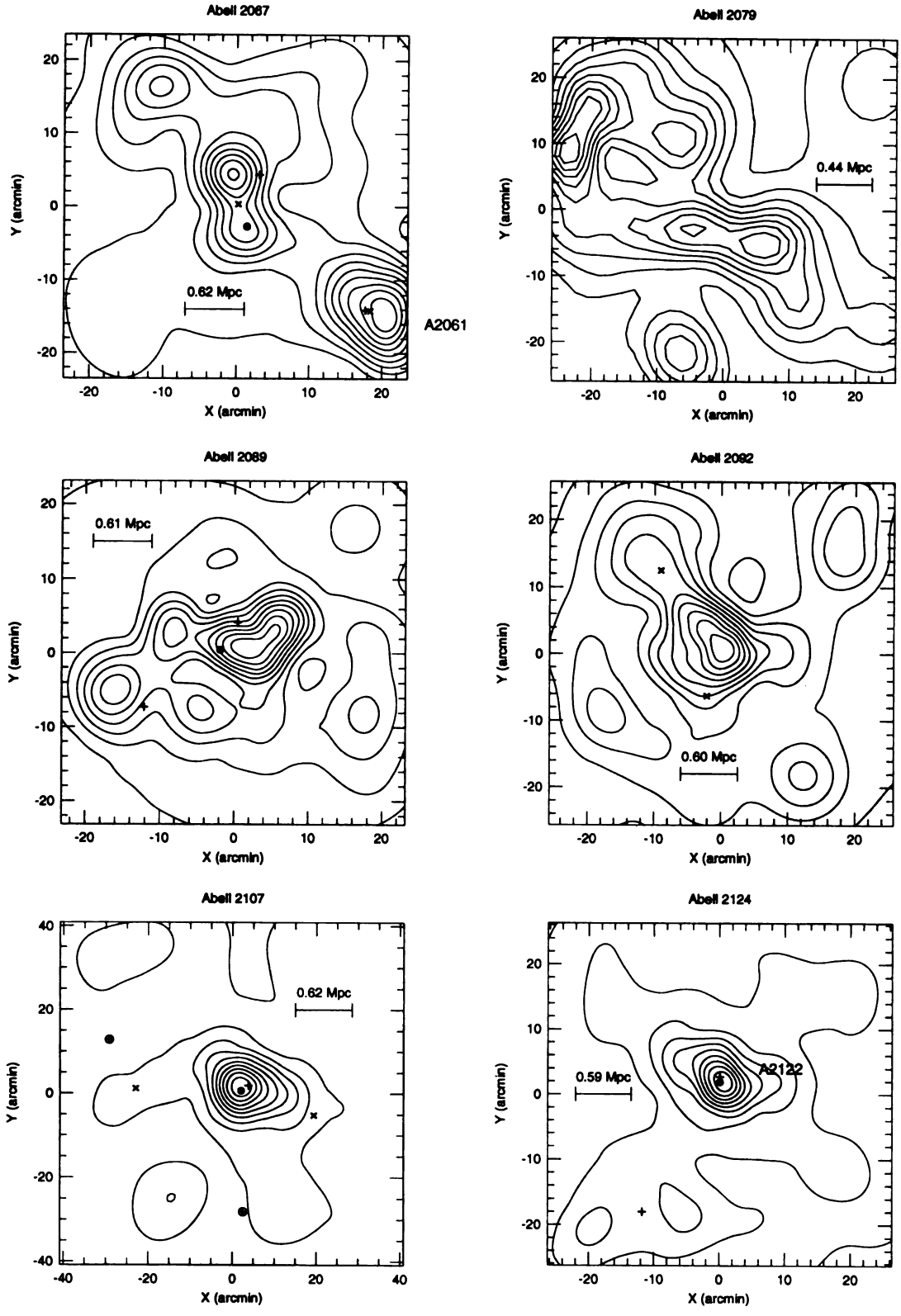

Figure 3.9 (cont'd).

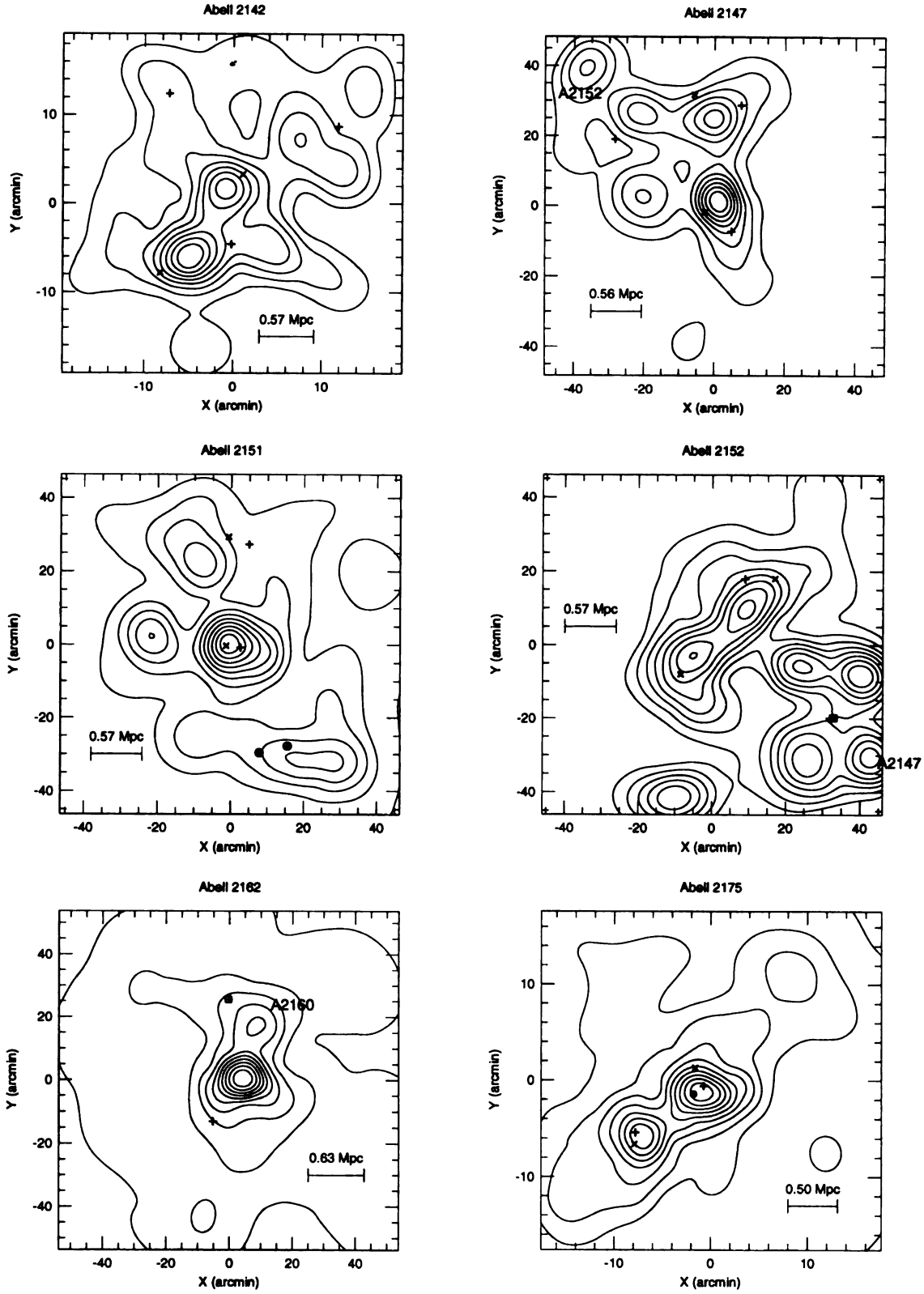

Figure 3.9 (cont'd).

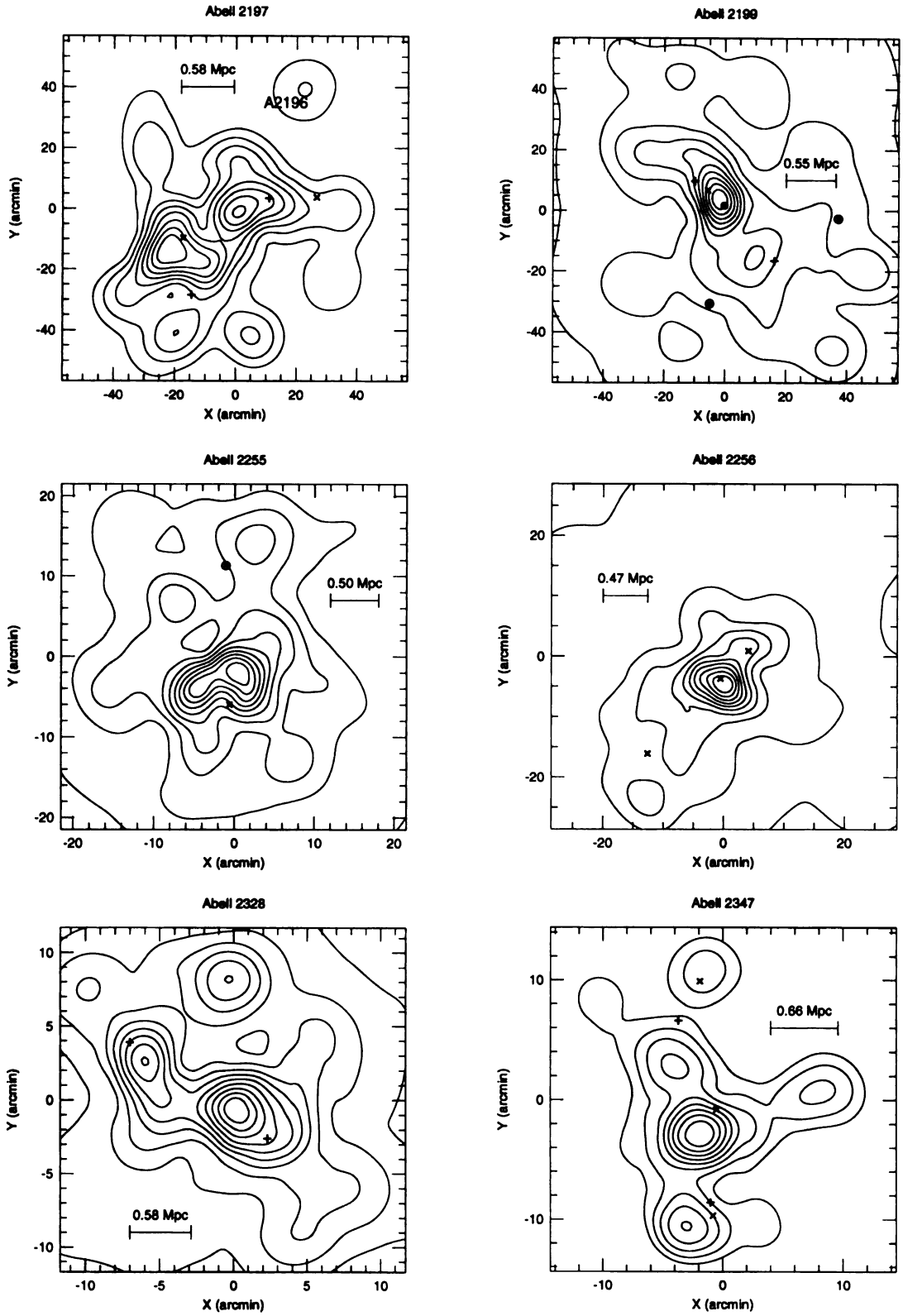

Figure 3.9 (cont'd).

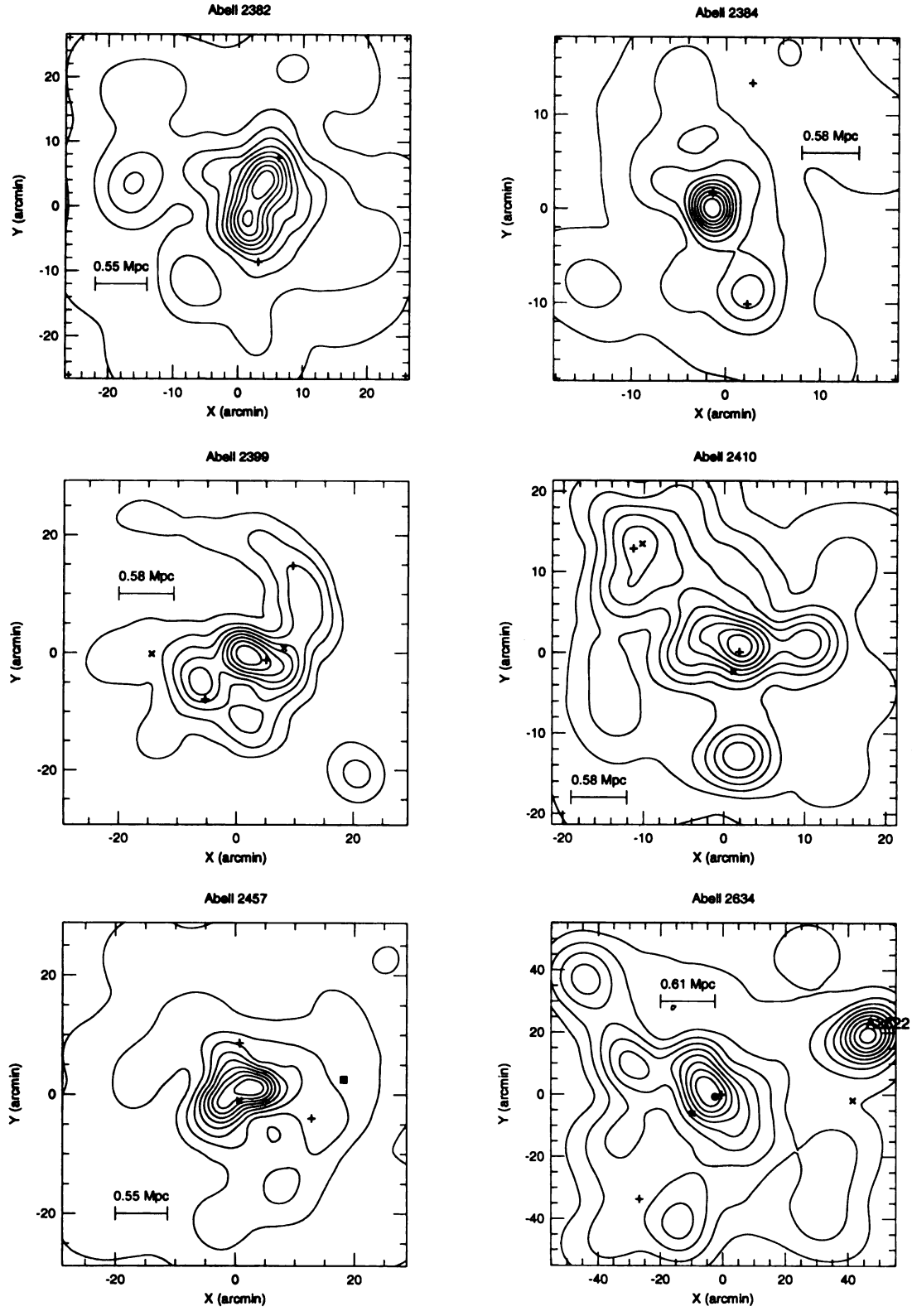

Figure 3.9 (cont'd).

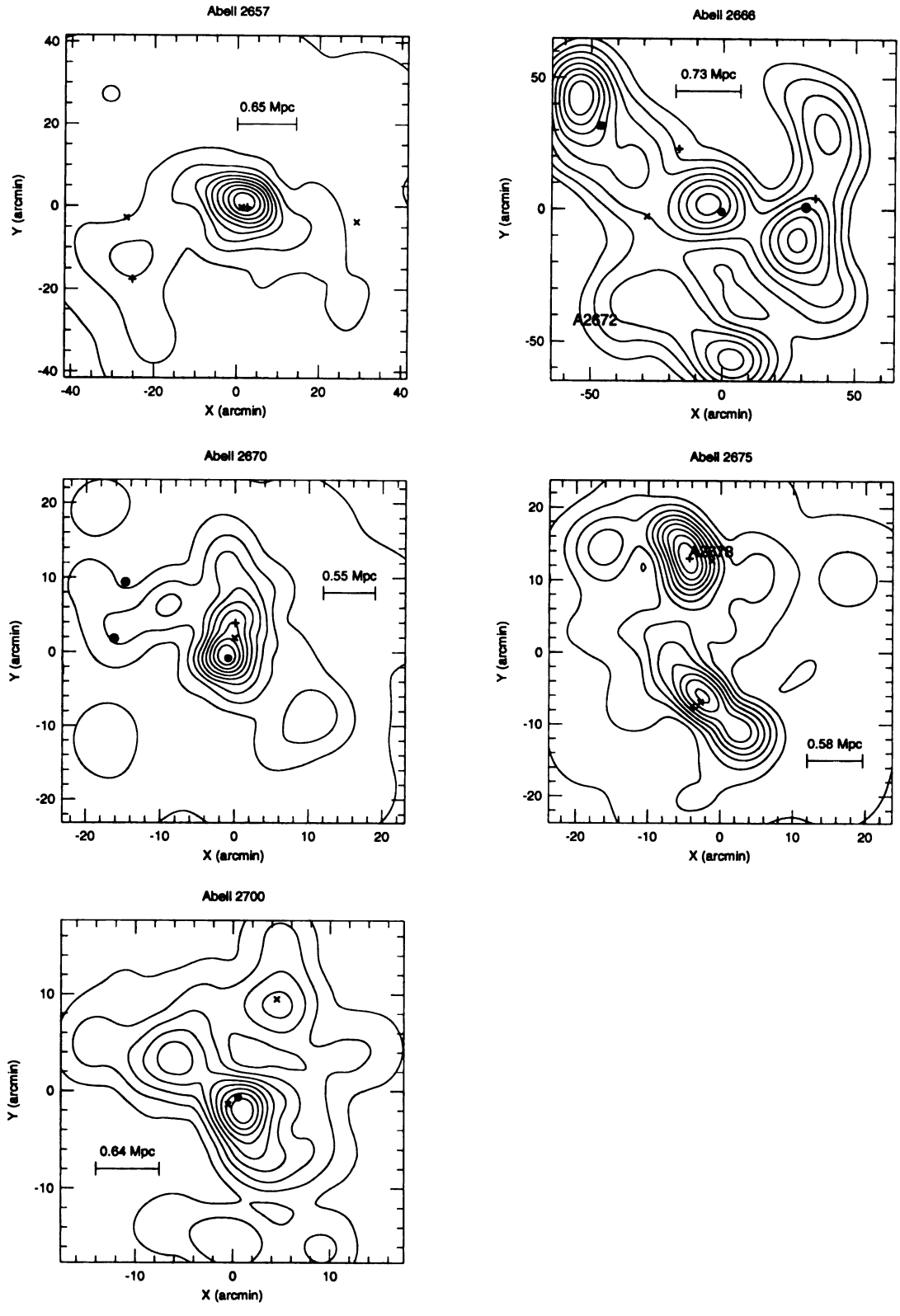

Figure 3.9 (cont'd).

For the clusters which are common to both samples it is of interest to compare the maps made with the different data sets. Although this comparison is made somewhat difficult by the different sizes of the maps, it is clear that most of the features visible in the Dressler maps are still visible in the maps made with the APS data. There are a number of special cases which require further attention.

The peaks to the north and south of A119 are washed out in the larger APS data set, so much so that the DEDICA program no longer finds them significant, although the KMM partition remains. This could mean that the fainter magnitude cutoff used in the second map ($m_V \approx 18.6$ for the APS data as opposed to $m_V = 16.5$ in the Dressler data) has increased the background to the point where the structure is being washed out. However, comparison of the estimates for background listed in tables 2.2 and 2.3 indicate the opposite is true; the Dressler map has a *higher* background. This is likely a consequence of the slightly larger area plotted in the APS map, which follows A119 into a lower density region near the edge of the map. The elongated peak to the northeast of A400 in the map made from the Dressler data is beginning to become resolved into two components in the APS map. Lastly, the elongation seen in the core of A1656 in the Dressler map is completely obscured in the APS map. This is not due to the larger data set employed but to the increased area plotted in the map. In Figure 3.10 the core region of the cluster is plotted using the APS data. Here the bimodal nature of the density is evident.

In fact, a large fraction of the clusters in the HGT sample show a similar effect. In the case of the Coma cluster the reality of the structure in the core can be confirmed by corresponding peaks in the X-ray surface brightness map of the cluster (Davis & Mushotzky 1993). A400, shown in Figure 3.11, is another example of a cluster with apparent core substructure. In this case the EINSTEIN X-ray map is elongated in a direction similar to the small scale structure in the core of the cluster (Beers

et al. 1992). Further evidence of the reality of this substructure is given by the bimodal (and possibly even trimodal) appearance of the velocity distribution plotted in the examples of density estimation given above. While a more detailed study is clearly called for, this raises the intriguing possibility that most clusters could have substructure in their cores at scales of approximately $250h^{-1}$ kpc, or the canonical size of the core radius of clusters. By choosing to search for substructure in the $1.5h^{-1}$ region, this study misses identification of this small-scale core structure.

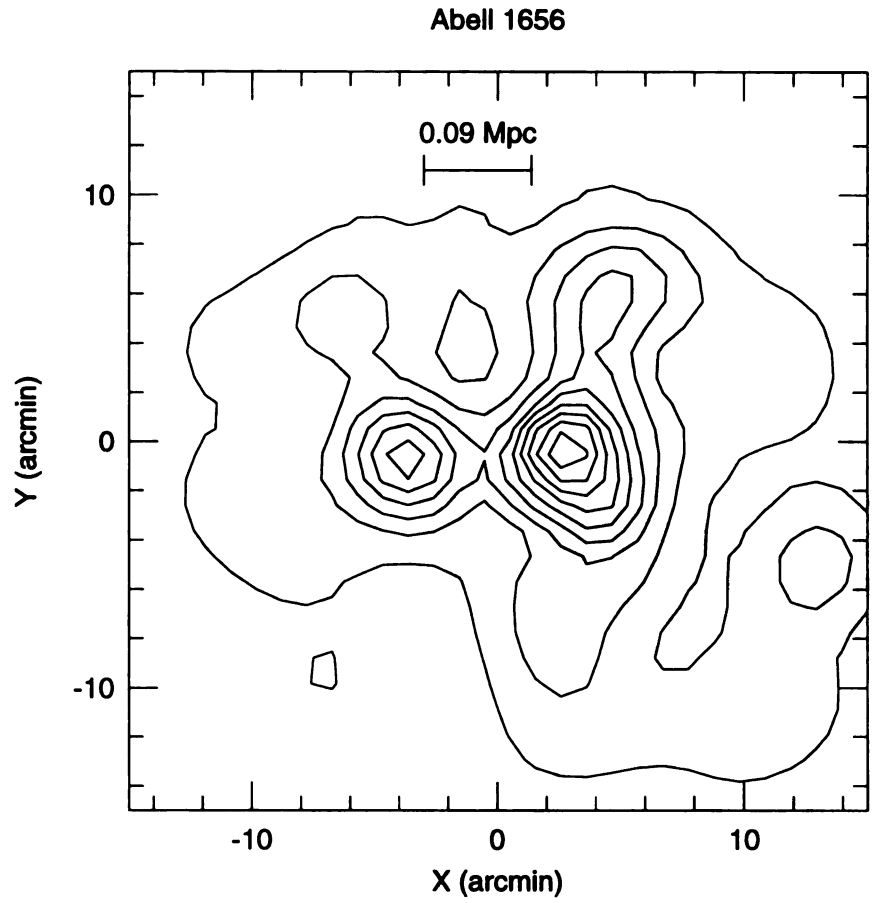

Fig. 3.10.— Core region of A1656 showing that substructure can exist even in the cores of very massive clusters.

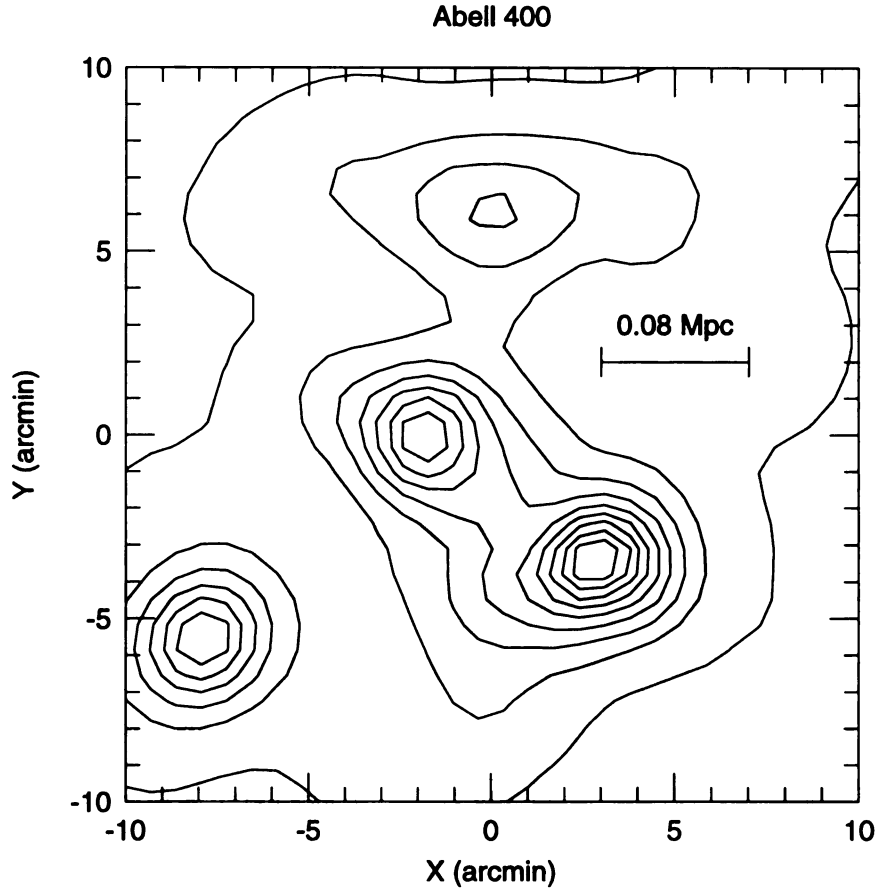

Fig. 3.11.— Adaptive-kernel map of the core region of Abell 400 indicating the possible existence of core substructure.

Chapter 4

TESTS FOR SUBSTRUCTURE

4.1 Introduction

Although the adaptive-kernel maps discussed in Chapter 3 can give an indication of the existence of substructure within galaxy clusters, the presence of peaks in the maps alone cannot prove the statistical significance of those groups. For this a test for substructure must be employed. Proving a particular feature, such as a second density peak, exists in a particular data set can be a difficult problem. In general it is a simpler task to show that alterative descriptions of the data are *not* true. This is referred to as hypothesis testing.

In this chapter two tests for substructure will be discussed and applied to the projected galaxy positions of the sample clusters, both of which are hypothesis tests. The first test, KMM (McLachlan & Basford 1988, hereafter MB), is a parametric approach which fits the data to bivariate Gaussian distributions and tests whether a single Gaussian or multiple Gaussians provide the best fit to the data. Significant substructure is claimed for those clusters where a likelihood-ratio test rejects the hypothesis that the data are drawn from a single Gaussian.

Although fitting the projected distribution of galaxies in a self-gravitating system to Gaussians cannot be justified from theoretical arguments, the Gaussian neverthe-

less remains the best available choice. Given the freedom to choose any mean and standard deviation, the two-dimensional Gaussian is capable of approximating a wide variety of shapes. More importantly, from the perspective of hypothesis testing as applied here, it is the most well-studied statistical distribution. Other theoretically-derived PDFs for clusters such as the Michie-King models or the isothermal sphere are mathematically more complex and have not been well-studied by statisticians. Furthermore, the observational evidence for such distributions in galaxy clusters is still open to debate (see discussion in Chapter 6.)

The second test, DEDICA, is the nonparametric technique due to Pisani (1993, 1996). In this approach no assumptions are made about the shapes which the groups might have, only that each group is identifiable by a single peak in the PDF. Each detected peak, in turn, is assumed to belong to the background and the likelihood of the fit evaluated. Significant substructure is claimed for for those groups where a likelihood-ratio test rejects the hypothesis that they belong to the background.

Although in this thesis these tests will be applied only in two dimensions, the x and y positions of the galaxies projected onto the plane of the sky, both algorithms are capable of using redshifts to test for clustering in three dimensions. With the large redshift surveys currently planned, such as SLOAN (see Bahcall 1995) and ENACS (see den Hartog 1995), it is important that these potentially very powerful techniques be well studied and tested.

4.2 The KMM Algorithm

KMM implements the Expectation Maximization (EM) algorithm of Dempster *et al.* (1977) as described by McLachlan & Basford (1988, hereafter MB). The program is also discussed for use in the detection of bimodality in univariate data by Ashman,

Bird & Zepf (1994, hereafter ABZ), and has been applied extensively in the analysis of substructure in clusters by Bird (Bird 1994a, Bird 1994b, Bird 1995, Bird, Davis & Beers 1995). The KMM program provides a maximum-likelihood fit of a data set to a mixture of Gaussian distributions, and can be used for hypothesis testing by evaluation of a likelihood-ratio test. The number of Gaussians to be fit, g, as well as an initial g-group partition of the data is specified by the user. Alternatively, the user can specify a first guess at the parameters of the individual Gaussians to be fit (locations and covariance matrices) along with an estimate of the mixing proportions. For this application specification of an initial partition is our preferred choice, rather than specification of the unknown positions and sizes of the groups. Furthermore, several objective partitioning algorithms exist that can be employed to specify the initial partition (see Kaufmann & Rousseeuw 1990 and references therein).

If it is assumed that the data points, $\mathbf{x}_1, \ldots \mathbf{x}_N$ (the x and y positions of the galaxies) are independently drawn from g Gaussian probability density functions (PDF), $G_1, \ldots G_g$, then the PDF for the superpopulation G can be represented as:

$$f(\mathbf{x}; \phi) = \sum_{i=1}^{g} \pi_i f_i(\mathbf{x}; \theta), \tag{4.1}$$

where $f_i(\mathbf{x}; \theta)$ is the PDF of G_i and π_i is the fraction of the superpopulation belonging to G_i. Here θ contains the elements of the mean vectors μ_i and the covariance matrices $\mathbf{\Sigma}_i$ for each group, and the vector

$$\phi = (\pi', \theta')' \tag{4.2}$$

is the vector transpose of all unknown model parameters. The log-likelihood of the complete data can then be defined as:

$$L_C(\phi) = \sum_{i=1}^{g} \sum_{j=1}^{n} z_{ij} [\log \pi_i + \log f_i(\mathbf{x}_i; \theta)], \tag{4.3}$$

where z_{ij} is an indicator variable:

$$z_{ij} = \begin{cases} 1 & \text{if } \mathbf{x}_j \in G_i \\ 0 & \text{if } \mathbf{x}_j \ni G_i. \end{cases}$$

Once an initial partition has been specified by the user, KMM calculates the log-likelihood of the fit using equation (4.3). The program then proceeds to find the value of ϕ, say $\phi^{(1)}$, which maximizes the expectation of the log-likelihood conditional on the observed data and the initial fit. Using $\phi^{(1)}$, posterior probabilities for group membership can be estimated by:

$$\tau_i(\mathbf{x}_j; \phi^{(1)}) = \frac{\pi_i f_i(\mathbf{x}_j; \theta)}{\sum_{t=1}^{g} \pi_t f_t(\mathbf{x}_j, \theta)}, \tag{4.4}$$

for $i = 1 \ldots g$. Here $\tau_i(\mathbf{x}_j; \phi)$ is the probability that the object with observation $\mathbf{x}_j$ is a member of group G_i. The expectation of the log-likelihood is then re-calculated using equation (4.3) with the z_{ij} replaced by the $\tau_i(\mathbf{x}_j, \phi^{(1)})$, the posterior probabilities. The program then searches for the value of ϕ, say $\phi^{(2)}$, which maximizes the expectation of the log-likelihood. These two steps, E (expectation) and M (maximization), are repeated iteratively until $L_C(\phi)$ has converged to a local maximum, provided a maximum exists (for a discussion of the convergence properties see Wu 1983). Objects are then assigned to the group for which their posterior probability of membership is the highest.

The final value of $L_C(\phi)$ is used to evaluate the improvement of the g-group fit over the null hypothesis that the galaxies are drawn from a distribution of g_0 Gaussians by calculating the log-likelihood ratio λ:

$$\lambda = \frac{L_C(\phi)}{L_C(\phi)^{(g_0)}}, \tag{4.5}$$

where $L_C(\phi)^{(g_0)}$ is the log-likelihood of the g_0 group fit. The greater the value of λ, the greater the improvement in the fit.

In order to quantify the improvement in the fit obtained by the addition of another Gaussian, the percentile (p-value) of the log-likelihood ratio can be estimated using a bootstrap procedure, as follows. Random data samples are generated under the null hypothesis that the data are drawn from a mixture of g_0 Gaussians with means, covariance matrices, and mixing proportions specified by the likelihood estimates from the g_0-group fit to the original data. For each bootstrap sample, λ is calculated after fitting mixture models for both g_0 and g groups. The value of λ from the actual data can then be compared to the null distribution of λ values calculated from the bootstrap samples to find the significance. In this paper $g_0 = g - 1$, thus requiring that any g-group fit be a significant improvement over the $(g-1)$-group fit.

It is important to note that the EM algorithm discussed above is not the only way to maximize the likelihood equation. Various other algorithms have been proposed and applied. The most well known of these are a group of algorithms based on Newton's method (Press *et al.* 1986). There are also algorithms due to Fletcher & Reeves (1964) and Nelder & Mead (1965.) Six methods are compared by Everitt (1984) for the case of a mixture of two univariate normal densities. In general, convergence was fastest using Newton's method with exact expressions of the first and second derivatives of the likelihood equation. This advantage over the EM algorithm disappears when finite-difference approximations for the derivatives are used. Both the Fletcher-Reeves and the Nelder-Mead algorithms showed a tendency to find points from which no improvement could be made even though not at a local maximum. The main advantage of the EM algorithm is that each iteration is guaranteed to improve the fit. This is not always true for Newton's method (McLachlan & Basford 1988). In general less than 100 iterations are required for convergence, which on a Sun Sparc 2 takes less than 30 seconds. Thus speed of convergence was not a big issue.

4.2.1 Monte Carlo Simulations

In order to assess the strengths and weaknesses of the KMM algorithm three questions need to be addressed. First, how often is KMM likely to classify a given data set as having significant substructure when such substructure does *not* exist? Second, for cases where real substructure exists, under what circumstances is KMM likely to fail to recover it? The former is traditionally referred to as an error of type 1, while the later is referred to as an error of type 2. Finally, what is the effect of *random* background/foreground contamination? In order to answer these questions a number of Monte Carlo experiments was conducted, following ABZ.

Data points were drawn randomly from two-dimensional Gaussian distributions with a covariance matrix of $\sigma_{xx} = \sigma_{yy} = 1.0$ and $\sigma_{xy} = 0.0$, which remained fixed for all data sets. For each case 250 data sets were generated, with the number of points set to 50, 100, 250, and 500, with an assumed constant-density background of 0%, 10% and 20% of the total number (numbers most relevant to the data sets, see Chapter 2). KMM was run on each data set for both the homoscedastic (common covariance) and the heteroscedastic (independent covariance) situations. In the homoscedastic tests, the Gaussians are forced to share a common shape, while they are allowed to have independent shapes in the heteroscedastic case. The significance of the resulting partitions was evaluated using the bootstrap procedure (1000 resamples) described in the previous section, with the modification that the analytical means and covariance matrices of the null hypothesis were used instead of the likelihood estimates. This avoided the need to bootstrap each realization of a data set, which was impractical due to the CPU time required.

The experiments can be divided into two broad categories, corresponding to the different error types. Category I contains those data sets generated under the null

hypothesis in order to test KMM's propensity to identify substructure which is not real. Category II contains data sets generated with hypothesized substructure in order to test the ability of KMM to correctly recover the input (real) substructure.

In category I, data points were generated from a single Gaussian with mean (0,0) and covariance described above, and KMM was requested to find two groups. Futhermore, random data sets were generated from two equally-populated Gaussians with the mean of the first group set to (0,0) and the mean of the second group varied between (1.50,0.0) and (4.00,0.0) in steps of $\delta x = 0.25$. Again, the covariance matrix of each group remained the same as above. In this instance, KMM was asked to identify three groups. In category II, two equally-populated Gaussians were generated as described above, and KMM was requested to find two groups.

For category I experiments KMM was started with an objective partitioning of the data supplied by the program PAM (Partitioning Around Medoids). As described by Kaufmann & Rousseeuw (1990), PAM searches for a user-specified number of representative objects (the medoids). The medoids are chosen such that the dissimilarity (or distance) between the groups is maximized while at the same time the dissimilarity within each group is minimized. A final partition is effected by simply assigning each object to the closest medoid. In category II the initial partition of the data was obtained by assigning each object to the closest Gaussian (note that this is the same as running PAM with the medoids forced to be the centers of the Gaussians, without the CPU time required by actually running PAM).

In order to compare the results of the experiments with the fits obtained using real data, a generalization of the dimensionless parameter $\Delta\mu$ defined by ABZ is employed:

$$\Delta\mu_{ij} = \frac{d_{ij}}{\sqrt{\sigma_i \sigma_j}}\,, \tag{4.6}$$

where d_{ij} is the distance between the averages of groups i and j, and σ_i is the standard deviation of group i along the vector joining the average positions of groups i and j. Note that with the averages and covariance matrices of the Gaussians described above, $\Delta\mu$ is simply equal to the average x position of the second group.

For all category I experiments using homoscedastic fits, the rate of false positives (type 1 errors) at the 90% significance level remained below 10%. When groups with less the 20% of the total number of galaxies in each cluster were rejected, as done in this thesis, the rate of false detection falls to 5%. The corresponding numbers for the 95% and 99% significance levels are 3% and 1%, respectively. These error rates changed little by the rejection of small groups or the distance between the groups. In general, the effect of adding a constant density background lowered the rate of substructure detection.

The results for the heteroscedastic runs show the opposite behavior. The highest error rate at the 90% significance level reached 85%. These error rates depend less on the significance level then on the separation of the groups and the background level added, with wider separations and higher background leading to higher error rates. However, the error rate is most sensitive to the small-group cutoff level. By rejecting groups with less than 20% of the total number, the error rate, in the case of 500 galaxies with a 20% background, is cut from 84% to 30%, with these values remaining constant for the 90%, 95% and 99% significance levels. Applying a 20% cutoff for small groups, the error rate only reaches the 10% level for groups with 250 or more members. It can be concluded that in the heteroscedastic case KMM has a propensity to return highly-significant groups with few members. Therefore, in large data sets a 20% small-group cutoff needs be employed because increasing the significance level does not lower the error rate. When dealing with large data sets, $N \gtrapprox 200$, the following procedure has been found to give good results. KMM

is run first for the homoscedastic case. If a g-group partition is found to be a good improvement (at the 90% level or better) over the $(g-1)$-group partition, then KMM can be run for the heteroscedastic case using the same initial partition to see if an improvement can be made over the homoscedastic fit, as judged by the Anderson-Darling statistic (see McLachlan & Basford 1988 and section 4.2.3 of this thesis). In this way the freedom to attain a better fit by allowing the groups to have different covariance matrices is retained without losing the robustness of the homoscedastic case. Furthermore, the Hawkins test described below will often give a good indication of whether a heteroscedastic fit is required.

The results of the two-group, category II, homoscedastic fits with no background are in good agreement with the univariate experiments of ABZ where the p-value was obtained by assuming that the null distribution of the LRTS was distributed as χ^2, as opposed to the bootstrap procedure employed here. In Figures 4.1 and 4.2 the results are plotted for the homoscedastic and heteroscedastic mixture models, respectively, using a 20% small-group cutoff, with a 10% background. With $N = 50$, the rate of detection at the 99% significance level does not achieve 90% until $\Delta\mu = 4.00$, a rather large separation. The corresponding numbers for $N = 100$, 250 and 500 are $\Delta\mu$=3.25, 2.75, and 2.50, respectively. In the heteroscedastic experiments, it can be seen that the necessary separation needs to be larger for a given rate of detection. Again, the addition of a constant-density background, at least to the 20% level, lowers the significance of the fits. This result underscores the need for deep catalogs of galaxies which sample well into the cluster luminosity function, without over-sampling background galaxies.

Although these Monte Carlo experiments can provide a useful guide to situations in which KMM is likely to succeed or fail to detect substructure, too much emphasis should not be placed on them because the cases tested are quite specific. Questions

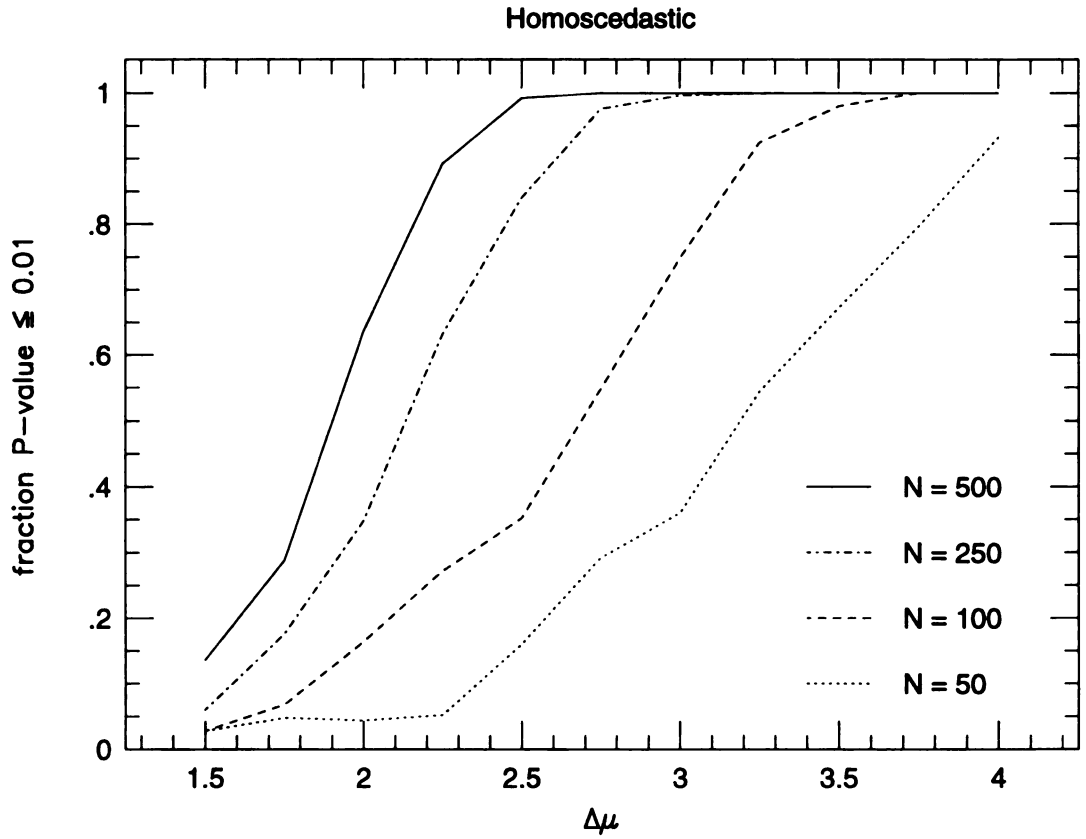

Fig. 4.1.— KMM success rate *vs.* group separation – homoscedastic case.

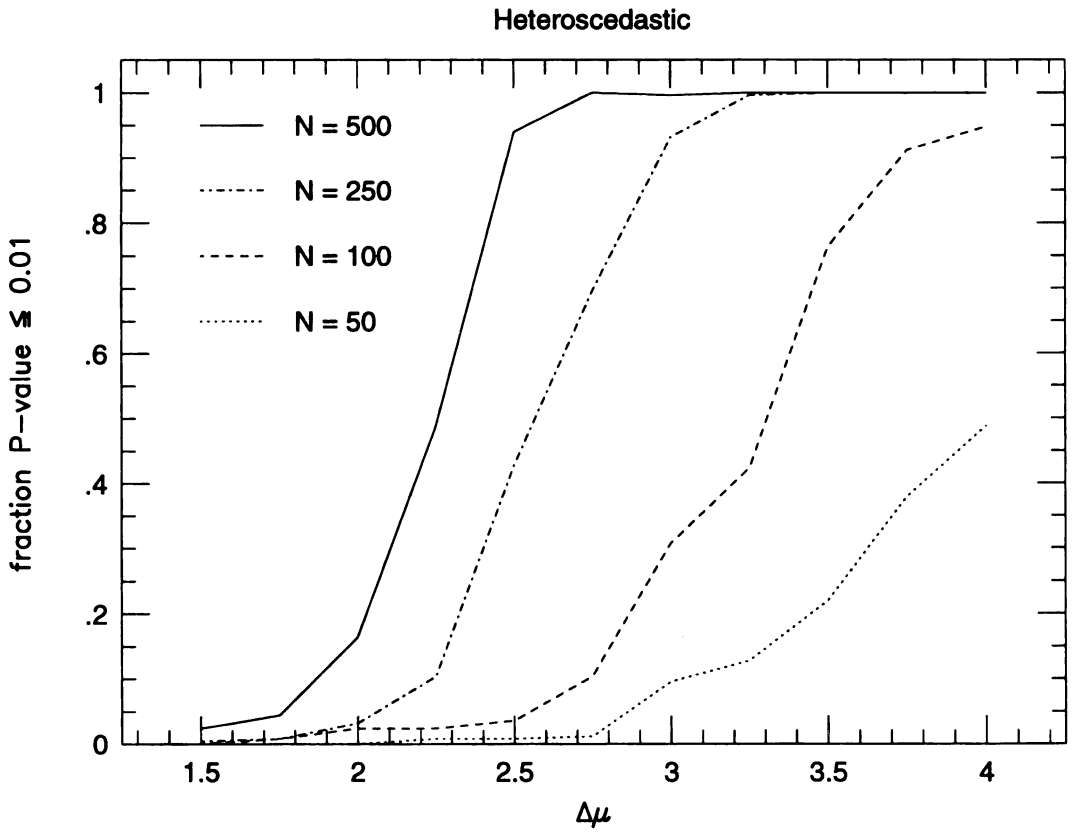

Fig. 4.2.— KMM success rates – heteroscedastic case.

not addressed by these experiments include how often in KMM likely to identify a single Gaussian as three or even four Gaussians, and how good is KMM at detecting substructure which is not Gaussian. Testing the vast array of possible parameter space where KMM might be applied is beyond the scope of this study. However, limited tests indicate that if a group is peakier than Gaussian, as clusters of galaxies are expected to be (Beers & Tonry 1986), KMM will perform slightly better than the above results indicate in the sense that closer groups can be identified. On the other hand, highly skewed distributions may be identified as significant substructure, especially with a large background. This can be guarded against by employing the Hawkins test as discussed in section 4.2.3 of this thesis.

4.2.2 Application of KMM to the Dressler Sample

KMM was run on each of the Dressler clusters for $g = 2$, $g = 3$, and $g = 4$ groups. The initial partition of the data was effected by assigning each galaxy to the closest user-chosen medoid. The medoids were generally chosen to correspond to the density peaks observed in the adaptive-kernel maps. However, multiple medoids were used for each cluster to ensure that KMM converged to a global maximum. In the cases where different solutions were found, the mixture model with the highest log-likelihood was chosen. The log-likelihood ratio was calculated with $g_0 = g - 1$ and the bootstrap carried out with 1000 iterations to estimate the significance of the fit. A p-value ≤ 0.05 was considered to be significant. From the Monte Carlo simulations, it can be estimated that this choice corresponds to an error rate of approximately 8%. Although a smaller error is easily achievable by applying more strict criteria, it was decided that without redshift data such refinements would not be meaningful.

The clusters for which KMM rejects the null hypothesis at the 95% level are listed in Table 4.1. Although a number of clusters have more than one acceptable partition,

only the one with the best Anderson-Darling statistic is listed. Furthermore, only those clusters with at least two groups containing more than 20% of the total are listed. (A151 is listed in Table 4.1, even though the second group does not meet the 20% criteria, for comparison with the DEDICA results discussed below.) Column (1) lists the cluster name. Column (2) gives the number of galaxies in each group. The percentage of the total number of galaxies for each group in given in column (3). In columns (4) and (5) the x and y positions along with their respective one-σ errors of the groups are listed in arcminutes. Column (6) is the significance of the partition.

It is interesting to note that several of the clusters which show multiple condensations in their adaptive-kernel maps are returned by KMM as not having significant substructure. For instance, there are four clusters, A978, A1991, DC1842-63 and Centaurus, which appear to have two similar-density condensations near their centers. These groups are fit by KMM and have $\Delta\mu$ values of 2.1, 1.3, 0.5, and 1.7 respectively. From the Monte Carlo experiments it can estimated that the probability (assuming these structures are real) of KMM returning significant p-values to be 0.09, 0.06, 0.02, and 0.10, respectively. These numbers would of course improve with a larger number of galaxies. Therefore, the absence of a given cluster from Table 3.1 should not be interpreted to mean that the cluster does not have substructure, but that a more detailed analysis (or deeper catalog of galaxies) might be required to detect it. Centaurus for instance, is known to have substructure in its velocity distribution (Lacey, Currie & Dickens 1986), a result that might have been predicted from the adaptive-kernel map.

TABLE 4.1. Mixture Model Parameters – Dressler Sample

Cluster	N	% of total	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	S
(1)	(2)	(3)	(4)	(5)	(6)
A 119	65	56	0.7 ± 11.1	0.0 ± 5.3	1.000
	28	24	−3.1 ± 11.1	20.5 ± 5.3	
	23	20	6.8 ± 11.1	−16.7 ± 5.3	
A 151	88	84	−2.9 ± 18.3	3.0 ± 15.1	0.972
	17	16	−2.6 ± 20.0	−32.5 ± 4.5	
A 154	37	47	2.5 ± 7.6	−2.0 ± 8.8	0.999
	21	27	6.7 ± 7.6	14.5 ± 8.8	
	21	27	−14.4 ± 7.6	−15.0 ± 8.8	
A 194	45	60	−4.3 ± 21.1	0.3 ± 21.3	0.980
	30	40	6.5 ± 16.8	−3.3 ± 17.7	
A 496	51	63	−0.2 ± 5.0	−2.2 ± 12.8	0.999
	17	21	−22.9 ± 5.0	0.6 ± 12.8	
	13	16	17.2 ± 5.0	3.7 ± 12.8	
A 548	157	67	−11.9 ± 19.1	8.1 ± 16.0	1.000
	77	33	22.2 ± 8.5	−19.6 ± 7.8	
A 754	84	56	−12.1 ± 13.9	−0.2 ± 12.6	0.999
	40	27	13.1 ± 13.9	9.7 ± 12.6	
	26	17	18.6 ± 13.9	−27.4 ± 12.6	
A 838	38	61	3.7 ± 15.7	−4.0 ± 12.8	0.999
	16	26	−33.1 ± 15.7	26.5 ± 12.8	
	8	13	27.5 ± 15.7	−31.2 ± 12.8	
A 957	46	56	−7.9 ± 20.2	−0.5 ± 22.9	1.000
	36	44	5.4 ± 19.2	0.1 ± 17.9	
A 979	50	58	0.8 ± 19.3	0.5 ± 7.8	0.998
	19	22	6.0 ± 19.3	−37.6 ± 7.8	
	17	20	4.5 ± 19.3	30.2 ± 7.8	
A 993	45	49	−6.4 ± 20.8	4.9 ± 11.8	0.982
	25	27	13.2 ± 6.6	−5.5 ± 23.0	
	21	23	−9.6 ± 24.1	−38.1 ± 7.4	
A 1069	26	55	2.4 ± 12.8	−10.5 ± 15.1	0.995
	21	45	5.7 ± 20.5	8.0 ± 21.0	
A 1631	69	77	0.3 ± 16.7	4.9 ± 8.6	1.000
	21	23	0.3 ± 16.7	−24.5 ± 8.6	
A 1736	133	80	−5.1 ± 17.0	−5.4 ± 15.6	0.984
	33	20	21.4 ± 17.0	22.8 ± 15.6	
A 2151	74	49	−1.0 ± 16.7	−2.0 ± 7.9	1.000
	47	31	−0.1 ± 16.7	25.7 ± 7.9	
	31	20	12.0 ± 16.7	−28.0 ± 7.9	

TABLE 4.1. (continued)

Cluster	N	% of total	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	S
(1)	(2)	(3)	(4)	(5)	(6)
A 2634	66	50	0.9 ± 11.4	−0.6 ± 7.5	0.953
	34	26	−8.0 ± 11.4	11.7 ± 7.5	
	32	24	8.4 ± 11.4	−16.8 ± 7.5	
A 2657	61	74	4.4 ± 8.4	−0.8 ± 8.9	0.960
	21	26	−19.4 ± 8.4	−2.3 ± 8.9	
DC 0103-47	25	47	−19.2 ± 10.8	0.2 ± 9.6	0.995
	16	30	13.4 ± 10.8	−11.9 ± 9.6	
	12	23	9.9 ± 10.8	32.7 ± 9.6	
DC 0247-31	34	71	−3.7 ± 19.8	−4.8 ± 24.0	1.000
	14	29	−0.2 ± 2.1	0.6 ± 1.8	
DC 0326-53	97	60	5.5 ± 20.1	−13.1 ± 13.9	1.000
	64	40	−11.5 ± 20.1	21.6 ± 13.9	
DC 0428-53	89	68	−1.2 ± 20.7	3.8 ± 19.2	0.999
	42	32	0.4 ± 3.7	−2.6 ± 8.5	
DC 0559-40	47	42	−9.5 ± 12.6	6.4 ± 16.6	0.999
	45	40	2.4 ± 11.4	−2.3 ± 4.0	
	20	18	32.9 ± 10.2	−7.8 ± 26.1	
DC 0622-64	55	67	−0.2 ± 18.9	9.3 ± 18.9	0.992
	24	24	−5.0 ± 18.9	−25.7 ± 9.3	
	19	19	−1.8 ± 18.9	30.0 ± 9.3	
DC 2048-52	96	44	2.1 ± 19.3	−5.1 ± 14.4	1.000
	77	36	6.5 ± 15.2	−10.4 ± 12.8	
	43	20	−5.0 ± 14.7	35.5 ± 6.4	
DC 2103-39	94	87	1.9 ± 19.7	−1.3 ± 16.8	
	14	13	−7.9 ± 11.3	3.5 ± 23.6	
DC 2345-28	41	43	1.4 ± 5.0	1.8 ± 13.7	0.998
	31	33	−15.3 ± 5.0	−2.8 ± 13.7	
	23	24	19.3 ± 5.0	−1.4 ± 13.7	
DC 2349-28	42	62	−3.4 ± 14.6	−6.9 ± 8.0	0.973
	26	38	−1.0 ± 14.6	19.7 ± 8.0	

Two other clusters – A400 and Coma (A1656) – are also known to have substructure from detailed kinematic and X-ray surface-brightness studies, but are absent from Table 3.1. Although the adaptive-kernel map of A1656 shows an elongated plateau to the east (suggesting unresolved substructure), the adaptive-kernel map of A400 shows no evidence of substructure in its core (Beers *et al.* 1992 suggest that the low-density peak to the northeast is likely to be a background group). In both of these clusters, the substructure *is* resolved in the adaptive-kernel maps made with a deeper galaxy survey (Kriessler, Beers & Odewahn 1995).

From table 4.1 it can be seen that 26 out of 56 clusters (46%) in the Dressler sample are better described by a multi-modal Gaussian fit than by a unimodal Gaussian, at the 95% confidence level (again, A151 is not counted.) The Gaussian sub-groups have a median separation from the global cluster centers of $0.6h^{-1}$ Mpc. These results are in concert with those of Geller & Beers (1982, hereafter GB), although there are disagreements for individual clusters. There are six clusters – A1142, A1983, A1991, DC 0317-54, DC 0326-53, and DC 0410-62 – for which GB claim substructure which is not confirmed by KMM. A1142 has a three-group partition that is significant at the 90% level, and therefore did not make the 95% cut. A1991 has two clear peaks in the central region of the cluster, one of which contains the D/cD galaxy. As discussed above, the groups are simply be too close together in this cluster for KMM to find a significant two-group partition.

4.2.3 Application of KMM to the HGT Sample

In the same manner as above, KMM was applied to the APS data for the HGT sample of clusters. Because the limiting magnitude used for each cluster was fainter than the $m_V = 16.5$ used in the Dressler sample, the HGT sample has more galaxies available and possibly larger field contamination. In an attempt to keep the error

rate low, the cutoff for significant structure was raised from 95% to the 99% level for this study. Furthermore, in the previous section no attention was paid as to whether or not a Gaussian was a good fit to the individual groups. The small sizes of the groups found in the Dressler sample made rejecting the the Gaussian hypothesis for the individual groups difficult and unreliable in many cases. Futhermore, the Monte Carlo experiments suggested little need for such precautions. (In any case, all of the partitions listed in Table 4.1 for the Dressler clusters meet the criteria outlined below.)

The situation is different using the larger data sets offered by the APS. With some clusters having 500 members and the possibility of 30% field contamination, the Monte Carlo simulations discussed previously suggest that the errors could in these cases be much larger than desired. A simple experiment illustrates the dangers. If KMM is run on a data set which consists of 250 points drawn randomly from a bivariate Gaussian distribution and 100 points drawn from a uniform distribution, there is a high probability of a three-group partition being a significant improvement over that of a two-group partition. In these cases the heavy tails on either side of the single-peaked Gaussian have been fit as two separate Gaussians.

It should be noted that this type of error is not simply a problem which pertains specifically to the KMM algorithm, but to all hypothesis tests. If the null hypothesis is fundamentally wrong, it is possible to get positive results even if the hypothesis being tested for does not pertain.

To explore the goodness of the Gaussian assumption for the groups a test due to Hawkins (1981) is employed. Although more complicated than a χ-square test, the Hawkins' test does not require binning of the data and can be used to test for homoscedasticity at the same time. This test is briefly outlined below.

To apply the Hawkins test it is necessary to assume that the mixture model returned by KMM is the true underlying density distribution. Then the Mahalanobis squared distance is calculated between each observation and the average of the group (calculated excluding x_{ij}) to which it belongs. Or, if x_{ij} is the jth observation in the ith group and $\bar{x}_i$ is the average of the i group, the Mahalanobis squared distance is defined as:

$$D(x_i, \bar{x}_i, ; S) = (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)' S^{-1} (x_{ij} - \bar{x}_i). \tag{4.7}$$

Here, the matrix S is given by

$$S = \sum_{i=1}^{g} (n_i - 1) S_i / (N - g) \tag{4.8}$$

with S_i the covariance matrix of the ith group

$$S_i = \frac{\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)'}{n_i - 1}. \tag{4.9}$$

As with the mean, S_i is calculated with x_{ij} deleted in case it contaminates the estimates of the mean and covariance matrix. It can be shown that the quantity:

$$\frac{(n_i - 1)\nu}{(n_i p)(\nu + p - 1)} D(x_{ij}, \bar{x}_i; S) \tag{4.10}$$

is an F distribution with d and $\nu = n - g - d$ degrees of freedom (d is the dimensionality of the data, in this case $d = 2$.) If $a_{(ij)}$ denotes the tail area under $F_{d,\nu}$ to the right of the calculated value of equation (4.10), then under the hypothesis that the ith group is normal the a_{ij} will be distributed approximately uniformly over the interval [0,1]. A summary of this information can be given by the Anderson-Darling statistic, defined as:

$$W_i = -n_i - \sum_{j=1}^{n_i} (2j - 1) \left[\log a_{i(j)} + \log(1 - a_{i(m_i - j + 1))} \right] / n_i, \tag{4.11}$$

where for each $i = 1, \ldots, g$, $a_{i(1)} \leq a_{i(2)} \leq \ldots \leq a_{i(n_i)}$. According to the Monte Carlo simulations carried out by Hawkins (1981), if W_i is greater than 2.5 the normality of the ith group can be rejected at the 95% level. However, in applying the Hawkins test MB recommend *not* using a hard cutoff since equation (4.11) only holds exactly as $N \rightarrow \infty$. This is because the a_{ij} can not be exactly uniform on the interval [0,1] unless there is an infinite number of them. This advice has been followed here. Groups with an Anderson-Darling statistic much greater than 3 were rejected as not significant. However, groups in the range of 2.5 to 4 were kept if DEDICA also returned the group as significant. The intention was to avoid rejecting significant groups simply because they are not well fit by a Gaussian, yet at the same time avoiding the identification of skewness or heavy tails, which usually have $W_i \approx 7 - 12$, as additional groups.

The results are given in Table 4.2. Again, only the best-fitting partition is listed. In this case, 83 clusters out of 118 (70%) have a significant multimodal Gaussian fit. Of the 25 clusters that are in both the Dressler and the HGT samples, if substructure was identified in the Dressler sample then in general it was identified in the HGT sample. Again this comparison is made more difficult because of the different sizes used. An illustrative case is that of A194. Since a similar magnitude cutoff is employed for both data sets (m_O=17.6 is approximately m_V=16.5), the larger number of galaxies in the APS data is due almost entirely from the larger area used. In the Dressler data, KMM splits the core region of the cluster into two groups. Even though the core region is more clearly elongated in the map made with the APS data, it is not partitioned. Instead, KMM is drawn to the groups which lie to the southeast (A207) and to the northwest. This same situation applies to the clusters A496, A957, and A2634. This indicates that substructure in the core regions of the clusters is being missed by the analysis done here and that a follow-up study which includes only the galaxies within 1/2 of an Abell radius show be conducted.

TABLE 4.2. Mixture Model Parameters for HGT Sample

Cluster	N_g	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 21	146	50	52	-0.8 ± 7.7	0.4 ± 4.0	19.9	18.6	1.000
	82	28	28	3.0 ± 7.7	12.1 ± 4.0	20.1	18.9	
	63	22	21	-4.0 ± 7.7	-11.7 ± 4.0	20.2	19.1	
A 85	204	63	63	-0.6 ± 13.7	2.5 ± 16.5	18.8	17.5	1.000
	69	21	19	-22.8 ± 3.8	-2.9 ± 16.6	19.0	18.1	
	50	15	18	-9.4 ± 14.4	-29.8 ± 1.7	18.9	18.2	
A 88	43	60	68	-0.2 ± 7.3	-4.7 ± 4.9	17.8	17.0	0.996
	22	31	26	-7.4 ± 7.1	6.8 ± 3.6	18.0	18.5	
	7	10	6	-1.2 ± 10.7	0.0 ± 11.4	18.4	0.0	
A 104	115	76	68	-4.3 ± 7.2	-0.9 ± 10.0	19.9	18.8	0.992
	36	24	32	14.4 ± 3.5	2.1 ± 9.7	19.5	19.0	
A 119	142	53	56	3.7 ± 15.9	1.1 ± 9.4	18.6	17.0	0.999
	67	25	22	5.1 ± 15.9	-23.8 ± 9.4	18.4	18.0	
	59	22	22	-10.1 ± 15.9	25.9 ± 9.4	18.5	17.8	
A 121	44	30	28	-9.2 ± 4.2	8.0 ± 4.9	19.9	19.3	1.000
	42	29	38	1.9 ± 4.2	-3.0 ± 4.9	20.3	19.6	
	30	21	17	-8.2 ± 4.2	-9.7 ± 4.9	20.0	20.0	
	29	20	17	8.5 ± 4.2	10.1 ± 4.9	20.4	20.4	
A 151	184	54	59	-0.4 ± 6.5	-5.4 ± 16.1	19.0	17.6	1.000
	93	27	26	-21.0 ± 6.5	-2.2 ± 16.1	19.0	18.1	
	65	19	15	20.9 ± 6.5	-5.1 ± 16.1	19.2	18.4	
A 154	111	41	35	-0.5 ± 12.4	-0.3 ± 5.8	19.5	18.4	0.996
	91	33	42	-4.1 ± 12.4	-16.6 ± 5.8	19.3	17.9	
	70	26	23	1.4 ± 12.4	17.0 ± 5.8	19.3	18.6	
A 166	60	38	38	3.5 ± 6.8	-2.1 ± 3.2	20.1	19.2	0.994
	52	33	34	2.8 ± 6.8	9.0 ± 3.2	20.4	19.8	
	45	29	27	-1.9 ± 6.8	-10.7 ± 3.2	20.2	19.6	
A 168	114	49	48	-1.2 ± 18.8	-0.5 ± 8.3	18.6	17.4	0.995
	74	31	25	-0.2 ± 18.8	23.6 ± 8.3	18.8	18.1	
	47	20	28	6.6 ± 18.8	-24.9 ± 8.3	19.0	18.4	
A 189	66	42	54	-26.7 ± 10.3	16.9 ± 19.1	18.2	17.5	0.996
	61	39	30	3.4 ± 17.4	-0.6 ± 28.0	18.5	17.6	
	30	19	17	35.8 ± 9.6	-16.1 ± 18.7	18.3	18.3	
A 193	168	64	58	0.6 ± 13.8	0.6 ± 15.0	18.8	17.4	0.992
	56	21	15	-21.1 ± 13.8	-18.8 ± 15.0	19.0	18.4	
	40	15	27	21.3 ± 13.8	19.8 ± 15.0	19.2	18.9	

TABLE 4.2. (continued)

Cluster	N_g	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 194	66	51	57	-3.7 ± 34.0	0.5 ± 40.4	16.9	16.2	0.990
	34	26	26	-47.0 ± 28.9	-53.6 ± 23.5	17.0	16.8	
	29	22	17	58.8 ± 22.0	59.4 ± 22.9	17.1	17.1	
A 225	142	70	69	-10.2 ± 8.0	-0.5 ± 11.9	19.6	18.4	0.997
	60	30	31	9.8 ± 8.0	-0.2 ± 11.9	19.5	18.8	
A 246	82	80	80	-1.3 ± 12.8	0.3 ± 13.1	19.8	19.3	0.999
	21	20	20	18.0 ± 4.3	17.0 ± 4.7	20.0	0.0	
A 274	122	70	68	-6.2 ± 4.4	-0.2 ± 7.0	19.7	18.5	0.999
	52	30	32	5.4 ± 4.4	-0.8 ± 7.0	20.0	19.1	
A 277	115	50	50	1.1 ± 3.9	-0.8 ± 8.7	19.7	18.2	0.996
	70	30	34	-10.6 ± 3.9	-1.3 ± 8.7	20.0	18.9	
	45	20	16	12.4 ± 3.9	3.3 ± 8.7	19.5	19.0	
A 415	121	50	61	-2.6 ± 9.6	8.1 ± 6.4	20.0	18.7	1.000
	122	50	39	-1.1 ± 9.6	-10.0 ± 6.4	19.9	18.9	
A 496	143	63	61	5.5 ± 23.9	-6.9 ± 16.0	18.2	17.1	0.992
	83	37	39	13.3 ± 23.9	34.5 ± 16.0	18.1	17.1	
A 514	130	46	40	-9.3 ± 7.1	-1.0 ± 9.2	19.6	18.5	0.994
	78	28	35	-1.9 ± 12.6	-18.3 ± 2.7	19.5	18.5	
	74	26	25	11.9 ± 6.1	3.7 ± 10.4	19.5	18.7	
A 787	111	72	59	4.4 ± 4.5	1.5 ± 5.6	20.4	18.9	0.995
	43	28	41	-8.3 ± 2.8	0.2 ± 7.0	20.2	19.5	
A 957	143	50	54	-0.5 ± 19.5	2.0 ± 7.6	18.6	17.5	1.000
	88	31	31	-6.2 ± 19.5	-24.6 ± 7.6	18.9	17.7	
	57	20	15	-3.5 ± 19.5	30.8 ± 7.6	18.9	18.4	
A 978	114	39	39	-0.3 ± 14.3	-0.4 ± 7.0	19.0	17.9	0.996
	124	42	40	-0.1 ± 14.3	-20.6 ± 7.0	19.2	17.7	
	57	19	21	4.9 ± 14.3	19.3 ± 7.0	19.0	18.2	
A 993	215	79	77	-2.5 ± 16.4	-1.3 ± 17.8	19.1	17.4	1.000
	57	21	23	22.8 ± 4.9	5.1 ± 17.9	19.0	18.1	
A 1139	106	63	73	-14.0 ± 13.7	-0.3 ± 22.8	18.4	16.7	0.990
	62	37	27	22.6 ± 13.7	-3.2 ± 22.8	18.4	17.7	
A 1185	155	46	52	0.9 ± 23.8	-0.1 ± 7.3	17.8	16.3	1.000
	105	31	31	-5.7 ± 26.6	-32.3 ± 11.7	17.8	16.7	
	75	22	17	-6.1 ± 30.6	34.1 ± 13.5	18.1	17.2	
A 1213	118	45	57	2.3 ± 8.1	1.5 ± 17.6	18.6	17.2	0.994
	94	36	28	-22.5 ± 8.1	1.6 ± 17.6	18.9	17.9	
	49	19	15	24.5 ± 8.1	-2.4 ± 17.6	18.8	18.4	

TABLE 4.2. (continued)

Cluster	N_g	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 1238	77	43	49	-0.3 ± 4.8	-1.9 ± 12.0	19.8	18.9	1.000
	57	32	27	-16.4 ± 4.8	7.7 ± 12.0	20.0	19.6	
	46	26	24	15.9 ± 4.8	-7.5 ± 12.0	19.9	19.6	
A 1254	211	80	70	-5.1 ± 12.2	0.3 ± 13.4	19.8	17.9	0.990
	52	20	30	17.1 ± 6.4	-10.9 ± 9.3	19.6	19.0	
A 1257	89	42	44	2.9 ± 21.8	2.9 ± 18.4	18.2	16.6	0.997
	47	22	16	18.9 ± 18.7	-36.5 ± 8.8	18.2	17.6	
	39	18	15	5.5 ± 25.1	40.5 ± 5.9	18.4	17.9	
	37	17	25	-41.8 ± 5.5	6.0 ± 26.6	18.0	17.6	
A 1318	114	40	70	2.2 ± 14.5	-12.1 ± 9.6	19.3	17.7	0.995
	104	36	20	5.5 ± 10.5	11.1 ± 10.0	18.9	17.6	
	49	17	8	-23.0 ± 4.6	10.7 ± 13.2	18.9	18.3	
	19	7	2	25.5 ± 2.8	25.1 ± 2.3	19.3	0.0	
A 1364	178	79	80	0.5 ± 7.9	2.7 ± 8.0	19.8	18.1	1.000
	48	21	20	-11.7 ± 2.5	-7.9 ± 5.3	19.6	19.1	
A 1365	58	36	43	-2.6 ± 8.6	-0.4 ± 3.3	19.3	18.5	1.000
	67	41	39	1.6 ± 11.5	-12.3 ± 4.3	19.7	18.5	
	38	23	18	-0.6 ± 13.2	14.2 ± 5.1	20.0	19.6	
A 1377	176	44	56	-12.1 ± 10.5	-6.8 ± 10.3	18.9	17.3	1.000
	91	23	22	16.0 ± 10.5	15.3 ± 10.3	19.1	18.0	
	68	17	12	17.5 ± 10.5	-18.5 ± 10.3	19.2	18.2	
	67	17	10	-18.1 ± 10.5	19.6 ± 10.3	18.9	18.2	
A 1382	84	46	56	0.3 ± 3.5	-0.6 ± 8.4	20.0	19.0	0.995
	55	30	23	-10.9 ± 3.5	-0.5 ± 8.4	20.4	19.8	
	44	24	20	11.3 ± 3.5	1.2 ± 8.4	20.3	19.9	
A 1399	164	58	72	-1.2 ± 8.8	3.5 ± 7.2	19.7	18.2	1.000
	83	29	21	3.1 ± 8.3	-10.3 ± 4.8	20.1	18.9	
	37	13	7	-10.5 ± 5.5	15.8 ± 1.6	20.3	20.1	
A 1436	197	55	57	-0.4 ± 12.6	-1.3 ± 5.5	19.4	17.7	1.000
	95	27	31	-2.9 ± 12.6	16.5 ± 5.5	19.2	18.2	
	66	18	12	2.4 ± 12.6	-18.1 ± 5.5	19.5	19.0	
A 1474	76	41	39	-3.4 ± 5.6	6.6 ± 7.1	19.8	18.8	1.000
	50	27	29	-13.4 ± 5.6	-7.2 ± 7.1	20.0	19.1	
	32	17	14	11.5 ± 5.6	11.4 ± 7.1	19.9	19.8	
	29	16	18	10.3 ± 5.6	-10.8 ± 7.1	19.3	19.4	
A 1496	233	66	71	1.5 ± 8.3	-0.4 ± 9.1	19.8	17.8	0.997
	82	23	22	-14.2 ± 2.4	-0.7 ± 8.1	20.0	18.5	
	40	11	7	12.1 ± 3.6	12.1 ± 3.9	20.0	19.7	

TABLE 4.2. (continued)

Cluster	N_g	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 1541	37	18	9	-2.8 ± 3.0	-1.8 ± 1.3	19.6	19.4	0.998
	123	60	81	3.2 ± 8.6	6.4 ± 5.9	19.8	18.2	
	45	22	10	-3.4 ± 10.4	-11.2 ± 4.7	19.9	19.6	
A 1651	103	50	49	-0.3 ± 4.3	0.7 ± 10.2	19.9	18.9	0.991
	54	26	29	-14.1 ± 4.3	2.0 ± 10.2	19.7	19.3	
	48	23	23	12.7 ± 4.3	-4.3 ± 10.2	20.1	19.3	
A 1656	226	53	50	5.3 ± 16.6	0.0 ± 34.2	17.1	15.3	1.000
	108	25	20	-42.9 ± 16.6	10.2 ± 34.2	17.2	16.0	
	90	21	30	47.1 ± 16.6	-7.9 ± 34.2	17.2	15.9	
A 1691	119	48	60	1.0 ± 4.9	0.5 ± 12.1	19.1	17.8	0.997
	72	29	24	15.5 ± 4.9	2.1 ± 12.1	19.6	18.6	
	56	23	17	-16.6 ± 4.9	-5.0 ± 12.1	19.6	19.0	
A 1749	116	53	56	0.7 ± 14.6	1.3 ± 6.8	19.1	18.0	0.999
	65	30	31	0.2 ± 14.6	-16.8 ± 6.8	19.2	18.3	
	38	17	13	-9.3 ± 14.6	19.9 ± 6.8	19.7	19.3	
A 1767	186	60	62	-2.8 ± 11.5	-9.5 ± 8.3	19.4	17.9	0.999
	122	40	38	-0.9 ± 11.5	9.2 ± 8.3	19.3	18.3	
A 1773	119	42	40	4.5 ± 9.4	-1.3 ± 11.2	19.8	18.7	1.000
	111	39	38	-13.2 ± 5.2	7.3 ± 7.8	19.8	18.6	
	52	18	22	-2.2 ± 2.3	-0.6 ± 2.6	19.4	18.8	
A 1775	202	75	76	5.1 ± 11.4	0.0 ± 11.5	19.6	17.8	0.998
	66	25	24	-14.1 ± 5.6	-8.4 ± 7.4	19.3	18.6	
A 1795	36	12	12	2.3 ± 1.4	-1.6 ± 2.4	19.3	19.2	1.000
	138	48	51	-2.6 ± 14.1	11.7 ± 9.3	19.2	18.1	
	114	40	37	-2.6 ± 13.7	-13.2 ± 7.1	19.5	18.3	
A 1809	244	79	67	8.1 ± 7.2	-1.8 ± 10.4	19.8	17.8	1.000
	64	21	33	-10.1 ± 7.2	4.0 ± 10.4	19.7	18.7	
A 1831	132	43	44	1.6 ± 11.6	6.5 ± 9.8	19.6	18.2	0.995
	122	40	43	1.0 ± 11.5	-1.5 ± 10.9	19.4	18.1	
	54	18	13	-3.4 ± 13.0	-12.3 ± 6.5	19.7	19.1	
A 1837	182	68	70	-3.1 ± 23.1	7.9 ± 18.0	18.6	17.1	0.995
	86	32	30	-2.0 ± 19.7	-28.7 ± 9.0	18.7	17.9	
A 1904	135	35	29	-8.0 ± 9.1	8.5 ± 9.8	19.7	18.4	
	107	28	28	-2.6 ± 11.9	-17.2 ± 4.1	19.5	18.1	
	88	23	30	1.5 ± 4.5	-2.5 ± 4.9	19.2	18.1	
	56	15	13	16.7 ± 4.6	1.1 ± 12.1	19.6	18.8	

TABLE 4.2. (continued)

Cluster	N_g	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 1913	193	70	67	4.4 ± 12.9	1.1 ± 14.8	19.2	17.6	0.994
	60	22	26	−13.6 ± 10.0	12.9 ± 10.5	19.2	18.4	
	23	8	7	23.7 ± 5.0	−18.0 ± 8.3	19.3	19.5	
A 1927	154	63	53	1.9 ± 10.2	−1.3 ± 8.5	19.8	18.4	0.999
	53	22	25	−11.0 ± 6.3	15.4 ± 4.7	19.8	19.2	
	38	16	22	11.2 ± 7.5	−16.0 ± 3.3	19.8	19.6	
A 1983	328	75	78	0.4 ± 19.8	−7.7 ± 15.9	18.7	16.6	1.000
	111	25	22	2.1 ± 19.1	26.0 ± 7.8	18.9	17.5	
A 1991	235	64	68	−0.3 ± 13.3	−5.7 ± 11.9	19.1	17.5	1.000
	96	26	24	−1.3 ± 16.4	19.3 ± 6.6	19.4	18.2	
	37	10	8	23.5 ± 2.9	−20.7 ± 5.6	19.1	18.9	
A 1999	116	62	59	−3.7 ± 7.6	4.5 ± 4.9	20.1	18.7	1.000
	71	38	41	0.0 ± 7.6	−8.4 ± 4.9	19.8	18.7	
A 2005	95	68	73	4.8 ± 4.3	−0.9 ± 6.8	20.3	18.9	0.998
	44	32	27	−7.0 ± 4.3	3.0 ± 6.8	20.4	20.0	
A 2022	214	66	74	−0.2 ± 14.2	8.7 ± 9.9	19.2	17.4	0.995
	108	34	26	4.6 ± 14.2	−14.5 ± 9.9	19.4	18.5	
A 2028	164	71	68	−2.0 ± 11.6	−1.1 ± 10.3	19.8	18.3	0.997
	67	29	32	−0.5 ± 12.2	15.7 ± 4.0	19.7	18.9	
A 2048	200	64	63	−0.2 ± 8.3	−1.4 ± 8.1	20.1	18.4	
	100	32	34	9.2 ± 5.5	9.4 ± 5.3	20.0	19.0	
	14	4	3	−11.7 ± 4.6	−16.0 ± 1.2	20.3	0.0	
A 2063	117	55	53	−12.8 ± 23.2	1.1 ± 28.6	18.3	17.3	1.000
	48	23	22	30.4 ± 10.9	−29.6 ± 15.2	18.4	18.1	
	46	22	25	1.7 ± 4.5	−3.2 ± 7.2	18.0	17.6	
A 2067	181	64	59	0.2 ± 10.7	0.3 ± 11.8	19.9	18.8	1.000
	67	24	25	18.4 ± 3.8	−14.2 ± 4.1	19.6	19.1	
	35	12	16	−11.1 ± 5.2	15.6 ± 4.0	19.7	19.4	
A 2079	154	48	53	−14.7 ± 7.4	8.0 ± 9.6	19.5	18.2	0.995
	109	34	33	8.5 ± 9.3	−8.5 ± 6.9	19.5	18.7	
	29	9	8	−6.4 ± 10.8	−20.9 ± 2.9	19.7	19.7	
	26	8	6	21.3 ± 2.6	13.5 ± 9.2	19.7	19.9	
A 2092	167	63	71	−2.1 ± 13.1	−6.7 ± 10.0	19.3	18.1	1.000
	63	24	21	−8.6 ± 6.7	11.9 ± 8.0	19.2	18.7	
	37	14	9	17.8 ± 4.5	17.9 ± 5.3	19.7	19.5	
A 2142	198	64	66	1.1 ± 8.8	3.3 ± 9.1	20.2	18.8	0.994
	86	28	28	−8.2 ± 6.5	−7.8 ± 6.8	20.2	19.2	
	27	9	6	14.6 ± 2.6	11.2 ± 5.0	20.4	20.4	

TABLE 4.2. (continued)

Cluster	N_g	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 2147	220	47	49	-2.8 ± 20.4	-1.9 ± 17.2	18.4	16.6	1.000
	168	36	37	-5.7 ± 24.1	31.7 ± 9.6	18.3	16.8	
	43	9	7	-0.8 ± 27.4	-43.0 ± 2.9	18.7	18.1	
	34	7	6	-44.5 ± 1.7	0.7 ± 24.1	18.4	18.3	
A 2151	155	40	46	-1.3 ± 22.2	-0.3 ± 9.1	18.3	16.7	1.000
	123	32	30	-0.7 ± 22.2	29.3 ± 9.1	18.3	16.9	
	110	28	24	7.8 ± 22.2	-29.6 ± 9.1	18.6	17.2	
A 2152	180	38	38	-8.5 ± 13.7	-8.0 ± 25.1	18.5	17.0	0.995
	152	32	37	32.9 ± 9.6	-19.8 ± 14.4	18.5	16.8	
	112	24	21	17.0 ± 13.5	18.0 ± 16.4	18.5	17.4	
	27	6	4	-40.4 ± 2.8	-1.4 ± 26.4	18.8	18.8	
A 2175	222	50	46	-1.6 ± 8.3	1.3 ± 8.3	20.2	18.9	1.000
	161	36	38	-7.9 ± 5.7	-6.6 ± 5.3	20.2	18.9	
	65	15	15	9.3 ± 4.3	10.5 ± 3.9	20.4	19.6	
A 2197	202	65	62	-17.3 ± 17.1	-9.7 ± 28.5	17.9	15.9	1.000
	111	35	38	26.6 ± 17.1	3.8 ± 28.5	17.9	16.3	
A 2199	155	40	45	-5.7 ± 18.7	6.8 ± 16.5	17.8	16.3	0.996
	93	24	19	-5.0 ± 24.5	-30.7 ± 14.0	18.1	16.9	
	93	24	23	37.1 ± 10.2	-2.6 ± 27.8	17.9	16.7	
	48	12	12	-9.4 ± 20.1	45.7 ± 6.5	17.7	17.2	
A 2255	276	66	71	-0.6 ± 9.8	-6.0 ± 7.0	19.6	17.9	1.000
	141	34	29	-1.1 ± 9.8	11.3 ± 7.0	19.8	18.5	
A 2256	237	53	54	4.1 ± 13.7	0.9 ± 13.5	19.0	17.3	0.995
	121	27	21	-12.6 ± 7.9	-16.1 ± 7.5	19.3	18.0	
	93	21	26	-0.5 ± 4.2	-3.7 ± 3.8	18.8	18.0	
A 2347	44	49	39	-0.5 ± 7.0	-0.8 ± 2.7	20.6	19.8	1.000
	25	28	40	-1.9 ± 7.0	9.9 ± 2.7	20.9	21.0	
	21	23	21	-0.8 ± 7.0	-9.7 ± 2.7	20.6	21.0	
A 2399	149	58	57	8.4 ± 9.4	0.7 ± 14.9	19.0	17.7	0.999
	107	42	43	-14.3 ± 9.4	-0.3 ± 14.9	19.2	17.9	
A 2410	183	78	84	1.2 ± 10.7	-2.3 ± 9.3	19.6	17.6	0.999
	52	22	16	-10.2 ± 5.4	13.5 ± 5.0	19.5	18.9	
A 2457	148	60	63	2.1 ± 13.5	1.2 ± 5.9	19.1	17.8	1.000
	58	23	19	2.7 ± 13.5	-18.4 ± 5.9	19.5	18.8	
	42	17	17	7.9 ± 13.5	19.9 ± 5.9	19.3	18.9	
A 2634	280	68	73	-9.8 ± 22.0	-6.1 ± 26.9	18.0	16.4	0.999
	97	24	20	41.4 ± 9.0	-1.9 ± 29.0	18.2	17.1	
	34	8	8	-43.2 ± 7.1	36.7 ± 7.3	18.0	17.7	

TABLE 4.2. (continued)

Cluster	N_g	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 2657	77	45	46	2.0 ± 8.5	−1.4 ± 20.2	18.5	17.8	0.999
	56	33	32	−26.5 ± 8.5	−4.9 ± 20.2	18.5	18.0	
	38	22	22	28.9 ± 8.5	−6.6 ± 20.2	18.5	18.2	
A 2666	103	60	70	−28.7 ± 22.2	−2.7 ± 37.1	17.8	16.5	1.000
	68	40	30	31.4 ± 16.7	0.9 ± 35.7	18.0	17.5	
A 2670	132	52	57	0.0 ± 4.4	1.9 ± 10.8	19.0	17.9	1.000
	77	30	25	−16.2 ± 4.4	1.8 ± 10.8	19.4	18.6	
	46	18	18	13.1 ± 4.4	−3.0 ± 10.8	19.1	18.7	
A 2675	100	55	50	−2.7 ± 10.4	−6.9 ± 8.4	19.5	18.4	0.999
	82	45	50	−1.2 ± 11.4	12.7 ± 5.3	19.5	18.2	
A 2700	66	51	51	−0.8 ± 8.6	−1.4 ± 5.6	19.6	18.9	1.000
	43	33	35	4.3 ± 7.9	9.0 ± 5.2	19.8	19.2	
	20	16	14	−1.5 ± 8.0	−15.3 ± 1.4	19.5	0.0	

Likewise, A151, with a larger area in the Dressler data, clearly has a low density elongation to the east which is partly resolved in the smaller and deeper map of the APS data. It is this elongation that KMM identifies as significant and not the peak to the south as in the Dressler data. The peak to the south may still be significant except that a number of its members did not make the one Abell radius cutoff. Other clusters similar to A151 include A978, A1139, A1913, A1983, and A1991.

In six cases – A168, A1185, A1377, A1656, A2063, and A2256 – KMM indicates the existence of substructure that is not present in the Dressler data. While it is possible that three, A168, A1185, and A1656, are due simply to the increased size of the APS map, the others are of similar size and the complexity in the maps is most likely due the the inclusion of fainter galaxies. Whether or not these structures are real, or simply due to the background, can only be resolved with redshift measurements. However, using the procedure discussed below, it appears likely that the substructure in A168, A1377, and A2063 is due to background contamination.

4.3 The DEDICA Algorithm

Many of the disadvantages associated with the use of KMM to detect substructure in galaxy clusters can be eliminated by using a method which evaluates the significance of the peaks in the probability density distribution. The easiest way to do this is a procedure similar to that adopted by GB: count a peak as significant if its density is 3σ above the background density. The main drawback to such a procedure is that the significance of the group depends critically on the assumed background density. Furthermore, as in KMM, it is advantageous to be able to assign individual galaxies to a specific group and to ascribe membership probabilities, this time without making any assumptions about the form of the underlying PDF for the groups. The details of

such a clustering procedure have been worked out and implemented by Pisani (1996) in the program DEDICA, which is summarized below.

First, the PDF of the cluster needs to be calculated. This is accomplished using the adaptive-kernel method used in Chapter 3 for construction of the contour maps. In this case it is advantageous to use the normal kernel despite its infinite support and lower efficiency since the next step requires the calculation of the derivative of $\hat{f}(x)$. Furthermore, more attention needs to be given to the choice of initial smoothing parameter h, since the number of groups and their significance will be dependent on the density estimate.

Pisani (1996) recommends choosing h by the method of minimizing the cross validation term in the IMSE (least squares cross validation, LSCV). To accomplish this DEDICA sets $h = 4h_n$ (h_n being the smoothing window as specified by the normal rule) and reduces it by a factor of 2 with each iteration till the value which minimizes the CV term is found. However, because LSCV can be sensity to small-scale effects in the data for small smoothing parameters (Silverman 1986), in this thesis h is given a lower bound of 100 kpc. This is about the size of a large galaxy. If LSCV returns of value of h smaller than this size, it is deemed to have failed and a value of $h = 0.85h_n$ is adopted. Without taking this precaution, h can, at times, be reduced to scales of 20 kpc or so and even the high-density peaks are assigned less than 5% of the galaxies within an Abell radius. Thus the trend appears to be that larger values of h will generally lead to larger subcluster sizes. This is due to the fact that smaller smoothing parameters lead to density estimates with larger gradients, and therefore a smaller spatial extent for the peaks found in the next step. In most of the clusters $h_{CV} \lessapprox 0.85h_{opt}$, and is deemed to have failed in about 40% of the cases. This poor success rate, taken with the fact that a majority of the rest of the clusters have $h_{CV} \simeq h_{opt}$, makes it difficult to justify the extra expense in CPU time

necessary to calculate h_{CV}. Furthermore, it is unclear whether the goal of minimizing the ISE of the density estimate is appropriate when calculation of the derivative of the density estimate is sought.

The next step is to identify the peaks in the density estimate, and therefore possible subclusters. This is done by an iterative scheme due to Fukunaga & Hostetler (1975.) The local maxima of $f(x)$ can be found by the limit of the sequence:

$$r_{m+1} = r_m + a_2 \frac{\nabla f(r_m)}{f(r_m)}, \tag{4.12}$$

where r_0, the position vector, is set to the position vector of each data point in turn. The factor a_2 controls the rate of convergence of the sequence which is optimized with:

$$a_2 = \frac{2}{\sum_{i=1}^{N} [\nabla f(r_i)/f(r_i)]^2}. \tag{4.13}$$

The iterative procedure is stopped when $|r_{m+1} - r_m|/r_m \leq 10^{-8}$.

A cluster is then defined as the group of points with positions vectors r_i for which equation (4.12) converges to the same value of r. Clusters with only a single member are considered to be isolated points.

Once the set of ν groups, C_μ, and n_0 isolated points has been identified, $f(r)$, the PDF of the entire sample, can be defined as:

$$f(r) = \sum_{\mu=0}^{\nu} f_\mu(r), \tag{4.14}$$

where $f_\mu(r)$ is the PDF for each of the ν groups:

$$f_\mu = \frac{1}{N} \sum_{j \in C_\mu} K(r_j, \sigma_j; r), \tag{4.15}$$

and $f_0(r)$ is the PDF of the background. The statistical significance of the μth group can be evaluated using a likelihood-ratio test:

$$\chi^2 = -2 \ln \frac{L(\mu)}{L_N}. \tag{4.16}$$

Thus defined, the likelihood ratio is distributed as chi-square with one degree of freedom (Materne 1979). The quantity L_N is the sample likelihood or:

$$L_N = \prod_{i=1}^{N} \left[\sum_{\mu=0}^{\nu} f_\mu(r_i) \right], \tag{4.17}$$

and $L(\mu)$ is the value that L_N would have if the μth cluster were described by $f_0(r)$ and thus actually belonged to the background. Or:

$$L_\mu = \prod_{i=1}^{N} \left[f(r_i) - f_\mu(r_i) + \frac{1}{N} \sum_{j \in C_\mu} K(r_j, \sigma_0; r_i) \right]. \tag{4.18}$$

As discussed in Chapter 2, the background density adopted here is the density at the point with the largest bandwidth factor $h\lambda_0$ in the adaptive-kernel density estimate. Another choice of the background density $f_0(r)$ could be the density due to all points defined as isolated. In many cases though, all points are assigned to one group or another with no points listed as isolated. Finally, the probability that the ith galaxy belongs to the μth cluster can be defined as:

$$P(i \in \mu) = \frac{f_\mu(r_i)}{f(x_i)}. \tag{4.19}$$

One added degree of freedom available with the DEDICA algorithm which needs to be mentioned is the ability to merge nearby groups into a single group. The merging is accomplished by a minimum spanning tree technique. Thus the user can specify, to some extent, on what length scales DEDICA is to search for substructure by specifying a distance d_{crit}. Density peaks which are further apart then d_{crit} will be considered distinct, while those closer will be merged. With d_{crit} set very large,

all galaxies will be in one large group; with d_{crit} very small most galaxies will be in their own group and isolated. This is different from the behavior of KMM, where the scale of the substructure is set by the scale of the data, *i.e.*, by one Abell radius in this case.

In the application of DEDICA to the clusters d_{crit} is set such that all galaxies are assigned to one large group. It is then lowered and the results analyzed as each successive group is split off from the main cluster. The process is halted before the significant groups found in previous steps are split into their generally non-significant component parts.

4.3.1 Application of DEDICA to the Dressler Sample

The results of the DEDICA runs are given in Table 4.3. Only groups significant at the 99% level are included. Column (1) lists the cluster. Column (2) gives the number of galaxies assigned to each group while column (3) gives the percentage of the total number in the cluster. The median x and y positions of the galaxies in each group is listed in columns (5) and (6) along with their one sigma errors. Column (7) gives the significance of each group evaluated against the background which is listed in Table 2.1.

TABLE 4.3. DEDICA Cluster Parameters for Dressler Sample

Cluster	N	% of total	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	S
(1)	(2)	(3)	(4)	(5)	(6)
A 14	61	77	−4.3 ± 10.1	3.3 ± 8.7	1.000
	18	23	12.2 ± 7.7	−14.1 ± 7.7	1.000
A 76	35	49	10.3 ± 8.4	−4.4 ± 12.0	1.000
	25	35	−6.0 ± 7.3	11.8 ± 7.9	1.000
A 119	85	73	3.6 ± 11.6	0.6 ± 10.1	1.000
	16	14	−13.1 ± 7.0	21.7 ± 5.2	1.000
	15	13	0.2 ± 6.2	−17.9 ± 4.0	1.000
A 151	72	69	0.5 ± 15.4	2.9 ± 10.4	1.000
	21	20	2.5 ± 10.6	−29.4 ± 7.2	1.000
A 154	41	52	4.4 ± 8.3	0.3 ± 6.7	1.000
	18	23	−11.7 ± 7.4	−19.4 ± 5.9	1.000
	16	20	0.2 ± 5.1	18.3 ± 4.8	1.000
A 194	37	49	9.1 ± 16.5	−11.0 ± 11.5	1.000
	29	39	−11.3 ± 13.6	10.8 ± 13.9	1.000
A 400	51	55	10.1 ± 13.9	−1.8 ± 13.8	1.000
	29	32	−15.1 ± 10.4	13.5 ± 13.8	1.000
A 496	60	74	−0.9 ± 16.0	4.6 ± 12.5	1.000
	21	26	0.4 ± 4.4	−13.7 ± 5.9	1.000
A 548	112	48	19.3 ± 10.3	−15.9 ± 14.2	1.000
	89	38	−24.8 ± 11.9	6.0 ± 12.2	1.000
	33	14	−7.5 ± 10.0	25.6 ± 6.5	1.000
A 592	29	48	−2.0 ± 14.0	4.0 ± 10.6	1.000
	14	23	−21.6 ± 7.6	−9.7 ± 8.6	0.999
	9	15	39.6 ± 0.8	18.1 ± 17.0	0.994
A 754	46	31	−18.1 ± 11.3	−2.1 ± 6.9	1.000
	36	24	21.3 ± 12.8	−19.6 ± 24.9	1.000
	33	22	2.1 ± 4.3	5.0 ± 4.2	1.000
A 838	17	27	18.7 ± 7.1	2.1 ± 10.4	1.000
	14	23	−0.7 ± 3.1	−5.0 ± 9.5	0.999
	8	13	−40.8 ± 3.5	24.2 ± 3.1	0.998
A978	32	52	−0.5 ± 14.2	15.4 ± 12.6	1.000
	30	48	0.3 ± 15.1	−12.8 ± 7.7	1.000
A 979	47	55	−3.8 ± 11.5	1.5 ± 10.3	1.000
	17	20	9.8 ± 15.8	−38.2 ± 4.9	1.000

TABLE 4.3. (continued)

Cluster	N	% of total	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	S
(1)	(2)	(3)	(4)	(5)	(6)
A 993	28	31	−14.6 ± 14.5	−0.5 ± 7.7	1.000
	21	23	7.0 ± 7.6	16.0 ± 6.0	1.000
	19	21	15.5 ± 6.0	−20.5 ± 15.8	1.000
A 1069	15	32	10.7 ± 6.3	4.0 ± 11.8	1.000
	15	32	−3.0 ± 6.8	−20.0 ± 7.3	1.000
A 1139	39	62	6.2 ± 18.7	5.2 ± 10.7	1.000
	20	32	−13.5 ± 10.8	−11.6 ± 11.0	0.999
A1142	19	32	−4.4 ± 3.2	20.0 ± 14.2	1.000
	17	29	7.3 ± 8.4	−23.1 ± 6.3	1.000
	12	20	13.4 ± 11.7	2.8 ± 9.8	0.997
	11	19	−21.5 ± 5.6	−2.4 ± 4.8	1.000
A 1631	50	56	−2.9 ± 17.6	7.6 ± 8.5	1.000
	20	22	5.9 ± 5.5	−1.6 ± 5.8	1.000
	20	22	−3.8 ± 14.4	−26.1 ± 7.9	1.000
A 1644	67	46	3.3 ± 19.7	18.1 ± 11.0	1.000
	63	43	−4.7 ± 17.4	−14.0 ± 11.2	1.000
A 1656	132	54	−3.7 ± 11.3	−0.5 ± 10.8	1.000
	64	26	24.3 ± 9.2	−10.8 ± 19.7	1.000
A 1736	76	46	−3.2 ± 22.0	−15.2 ± 10.5	1.000
	56	34	−2.4 ± 13.3	8.2 ± 9.8	1.000
A 1913	29	34	5.4 ± 4.2	2.4 ± 8.0	1.000
	19	22	−8.4 ± 6.6	−3.4 ± 10.8	0.997
A 1983	54	44	2.8 ± 12.7	−3.2 ± 10.0	1.000
	35	28	0.6 ± 11.0	20.8 ± 8.0	1.000
A 1991	26	49	8.0 ± 14.9	−6.9 ± 9.7	1.000
	18	34	−2.0 ± 12.9	14.2 ± 8.6	1.000
A 2151	73	48	11.3 ± 12.8	−8.1 ± 13.6	1.000
	46	30	−3.8 ± 12.6	26.2 ± 8.8	1.000
	25	16	−18.5 ± 5.2	0.3 ± 7.4	1.000
A 2256	53	64	−2.7 ± 9.8	−3.4 ± 3.9	1.000
	26	31	4.6 ± 6.5	7.4 ± 4.7	1.000
A 2589	31	43	−0.7 ± 5.7	−8.7 ± 7.8	1.000
	30	42	6.0 ± 10.0	4.0 ± 4.7	1.000
A 2634	66	50	−1.4 ± 11.8	−3.4 ± 7.4	1.000
	38	29	−6.6 ± 9.6	11.9 ± 5.3	1.000
	16	12	13.5 ± 5.8	−14.4 ± 4.8	0.999

TABLE 4.3. (continued)

Cluster	N	% of total	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	S
(1)	(2)	(3)	(4)	(5)	(6)
A 2657	58	71	−0.5 ± 8.0	1.4 ± 9.4	1.000
	13	16	17.8 ± 2.6	−0.1 ± 7.3	1.000
	11	13	−20.7 ± 4.5	−11.1 ± 4.7	0.999
DC 0103-47	26	49	−18.3 ± 10.5	0.0 ± 10.9	1.000
	15	28	14.0 ± 13.7	−10.3 ± 11.4	0.999
	12	23	9.6 ± 6.6	33.4 ± 7.5	1.000
DC 0317-54	47	72	2.8 ± 10.7	8.2 ± 19.5	1.000
	17	26	−21.2 ± 12.8	−20.0 ± 9.3	1.000
DC 0326-53	49	30	1.0 ± 11.4	−13.8 ± 10.0	1.000
	33	20	−25.4 ± 8.1	26.5 ± 7.1	1.000
DC 0329-52	146	77	−3.5 ± 11.9	2.2 ± 14.3	1.000
	39	21	24.8 ± 7.1	−17.1 ± 14.5	1.000
DC 0410-62	31	48	7.0 ± 14.5	−7.2 ± 13.2	1.000
	18	28	7.6 ± 18.1	36.1 ± 8.2	0.996
	15	23	−32.4 ± 7.9	−2.7 ± 21.7	0.995
DC 0428-53	55	42	−0.1 ± 9.7	−8.1 ± 8.6	1.000
	31	24	1.6 ± 5.4	13.4 ± 9.2	1.000
DC 0559-40	53	47	9.8 ± 12.4	−5.4 ± 9.2	1.000
	34	30	−12.1 ± 11.9	3.6 ± 11.4	1.000
DC 0622-64	71	72	3.4 ± 15.9	1.7 ± 14.2	1.000
	16	16	−13.0 ± 10.8	−26.4 ± 5.8	1.000
	11	11	−10.5 ± 8.3	32.8 ± 8.2	0.998
DC 1842-63	25	45	−4.0 ± 6.8	6.5 ± 5.9	1.000
	23	42	3.5 ± 10.7	−5.1 ± 6.7	1.000
DC 2048-52	144	67	1.4 ± 13.1	−8.8 ± 13.0	1.000
	35	16	−11.0 ± 8.1	34.4 ± 7.7	1.000
	28	13	25.3 ± 9.1	21.1 ± 16.2	1.000
DC 2103-39	39	36	13.4 ± 10.3	−3.2 ± 12.4	1.000
	36	33	−8.9 ± 9.9	4.1 ± 8.1	1.000
	17	16	0.0 ± 10.5	32.6 ± 7.3	1.000
	16	15	−18.5 ± 5.4	−16.5 ± 3.8	1.000
DC 2345-28	57	60	1.8 ± 11.2	2.8 ± 6.4	1.000
	17	18	−15.9 ± 6.9	−17.9 ± 7.6	1.000
	11	12	18.9 ± 5.1	−9.4 ± 3.5	1.000
DC 2349-28	31	46	3.9 ± 10.8	−3.7 ± 6.7	1.000
	24	35	0.8 ± 15.8	20.6 ± 6.6	1.000
	10	15	−11.8 ± 4.0	−17.8 ± 4.7	1.000
Centaurus	32	44	18.4 ± 10.7	−3.6 ± 14.4	1.000
	30	41	−13.9 ± 8.9	2.4 ± 19.0	1.000

There are 39 clusters (70%) in the Dressler sample that have significant groups at the 99% level. Many of the peaks visible in the adaptive-kernel maps which KMM was unable to return as significant, because of the small μ values of the groups, have been identified by DEDICA. These include A76, A400, A978, A1139, A1983, A2657, A1991, DC 0317-54, DC 0329-52, DC0410-62, DC 1842-63, and Centaurus. The partitions of A14 and A1142 are returned by KMM with a significance of 90% and therefore just missed the 95% cutoff. Note that all of the clusters identified by GB as having substructure have been confirmed by DEDICA (although four clusters – A119, A2657 DC 0622-64, and DC 2048-52 – are not counted in this study since in these cases the number of galaxies in the groups is less than 20% of the total.) This is not surprising given the similarity of the two approaches. Likewise, there are 22 clusters for which DEDICA gives positive results which did not meet the GB criteria. In these cases the difference is caused by the use of the adaptive-kernel density instead of the fixed, box-car smoothed density of GB. With a smaller kernel width used in the high-density regions, the peaks can have a higher density than if the galaxies in that peak had been spread out over the larger area of the bin width used by GB.

On the other hand, there are six clusters for which substructure found by KMM is not confirmed by DEDICA. These clusters are A957, A2657, DC 0247-31, DC 0608-33, DC 2048-52, and DC 2345-28. Where these groups are identified by DEDICA, they are listed in table 4.3. In most cases the groups simply did not have enough members to make the 20% cutoff. This applies to a number of other groups identified by KMM as well, such as the group to the south in A119 and the groups to the east and west in A496. The exception to this is DC 2048-52 which does not have any similar partition given by DEDICA. Although the Monte Carlo experiments indicate that the probability of a false positive in a cluster with only 48 members is small, the fact that there is no confirming homoscedastic fit or DEDICA partition for this case

leads to the suspicion that the structure in this cluster is not real.

4.3.2 Application of DEDICA to the HGT Sample

In a similar fashion, DEDICA was applied to the HGT clusters with the results listed in Table 4.4. Column (1) lists the cluster name. Column (2) gives the number of galaxies in each group. The percentage of the total number of galaxies for each group in given in column (3), while the percent of the total luminosity contained in each group is given in column (4). In columns (5) and (6) the x and y positions, along with their respective 1σ errors of the groups are listed in arcminutes. The median apparent magnitude m_{med} for each group is listed in column (7). The average apparent magnitude of the 10th-to-20th ranked galaxy is given in column (8). Lastly, column (9) lists the significance of each group.

In all, 96 (81%) of the clusters in the HGT sample are returned by DEDICA as having significant substructure. A comparison between the Dressler and APS samples shows that the partitions found significant by DEDICA are much more stable than KMM to variations in the chosen field size of the clusters. Despited the changes in size and magnitude limit between the two samples, 12 clusters are returned with essentially the same partition. In five cases – A957, A1185, A1377, A2040, and A2063 – substructure was detected by DEDICA in the deeper survey that was not detected in the Dressler data. However, these are all likely to be due to background contamination. Other differences can be explained as groups leaving (as in A151) or entering (as in A194 and A496) the field of view. In the case of A2256, however, it is clear that the larger background population in the APS map has washed out the structure so that it can no longer be identified. On the other hand, this larger number of galaxies has enabled KMM to identify the group, which it could not in the smaller Dressler sample.

TABLE 4.4. DEDICA Cluster Parameters for HGT Sample

Cluster	N	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 21	141	48	46	-2.6 ± 5.5	-0.2 ± 5.4	20.0	18.8	1.000
	69	24	25	3.4 ± 5.8	10.2 ± 4.6	20.0	19.0	1.000
	39	13	17	-10.6 ± 4.8	-12.8 ± 3.3	19.7	19.2	1.000
A 76	90	46	50	2.1 ± 13.9	-10.7 ± 16.1	18.5	17.3	1.000
	47	24	22	-14.8 ± 7.1	4.5 ± 10.2	18.4	18.0	1.000
A 85	112	35	34	0.1 ± 10.8	-1.6 ± 8.7	18.8	17.9	1.000
	59	18	24	-18.1 ± 6.5	-27.9 ± 4.1	18.7	17.9	1.000
	51	16	15	-20.4 ± 4.5	-2.6 ± 7.3	18.8	18.2	1.000
A 88	43	60	59	1.3 ± 7.3	-5.6 ± 5.3	18.0	17.3	1.000
	24	33	35	-9.9 ± 5.0	4.8 ± 3.8	17.7	18.1	1.000
A 104	112	74	72	-1.6 ± 8.8	-3.1 ± 6.9	19.7	18.6	1.000
	21	14	23	13.4 ± 3.5	8.2 ± 6.3	19.1	19.8	1.000
A 119	226	84	85	4.5 ± 15.5	-4.3 ± 17.2	18.5	16.9	1.000
	42	16	15	-18.0 ± 12.1	27.3 ± 8.1	18.5	18.2	1.000
A 121	59	41	35	-9.4 ± 4.2	3.5 ± 8.8	20.0	19.3	1.000
	40	28	35	1.3 ± 4.7	-0.2 ± 3.0	20.0	19.5	1.000
	20	14	12	10.8 ± 2.9	10.4 ± 2.8	20.2	20.6	1.000
	13	9	6	-7.2 ± 3.1	-14.9 ± 1.7	20.3	0.0	0.998
A 147	56	36	33	-2.3 ± 9.5	7.9 ± 12.6	18.7	18.1	1.000
	39	25	38	9.5 ± 5.9	-12.8 ± 12.8	18.1	17.8	1.000
A 151	161	47	53	-0.8 ± 6.6	-8.0 ± 12.9	19.0	17.7	1.000
	130	38	30	4.3 ± 22.3	5.7 ± 20.3	19.0	18.0	1.000
	46	13	14	-18.8 ± 4.7	-9.9 ± 5.4	19.0	18.6	1.000
A 154	96	35	31	3.9 ± 8.7	-2.0 ± 7.9	19.3	18.3	1.000
	83	31	39	-12.0 ± 8.8	-14.7 ± 7.2	19.2	18.0	1.000
	80	29	25	2.0 ± 13.5	16.2 ± 6.2	19.3	18.5	1.000
A 166	78	50	45	1.9 ± 8.3	-8.2 ± 4.5	20.1	19.2	1.000
	44	28	32	5.1 ± 5.3	9.2 ± 3.3	20.2	19.8	1.000
	27	17	20	-0.2 ± 2.4	0.1 ± 1.9	20.0	20.1	1.000
A 168	88	37	37	-0.4 ± 7.8	-1.2 ± 9.8	18.6	17.6	1.000
	44	19	19	7.0 ± 9.9	18.3 ± 8.1	18.7	18.3	1.000
	33	14	11	-18.6 ± 7.4	-15.6 ± 10.1	19.0	18.9	1.000
A 189	55	35	27	4.2 ± 18.0	-7.2 ± 31.4	18.5	17.9	1.000
	37	24	40	-29.9 ± 10.4	4.5 ± 10.3	18.1	17.8	1.000
	29	18	14	36.8 ± 8.9	-10.6 ± 16.9	18.3	18.3	1.000
	28	18	15	-16.2 ± 11.4	27.0 ± 6.7	18.1	18.3	1.000

TABLE 4.4. (continued)

Cluster	N	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 193	93	35	38.	−1.1 ± 7.9	−0.3 ± 7.8	18.6	17.7	1.000
	77	29	20	−10.7 ± 23.7	−23.2 ± 9.4	19.0	18.2	1.000
	74	28	38	2.7 ± 23.1	23.6 ± 7.7	19.1	18.2	1.000
A 194	51	40	50	−10.9 ± 23.6	0.4 ± 20.7	16.7	16.2	1.000
	33	26	17	−15.9 ± 47.0	−65.3 ± 11.9	17.2	17.1	1.000
	27	21	18	59.6 ± 20.4	59.9 ± 22.8	17.0	17.1	1.000
A 225	79	39	31	−6.3 ± 8.8	10.7 ± 10.9	19.8	19.0	1.000
	67	33	40	7.9 ± 8.2	−6.7 ± 7.6	19.3	18.7	1.000
	56	28	29	−16.1 ± 5.0	−5.7 ± 7.5	19.5	18.9	1.000
A 246	34	33	35	3.1 ± 8.6	−2.9 ± 6.5	19.8	19.8	0.998
	25	24	24	16.2 ± 5.6	15.6 ± 6.2	19.8	19.9	1.000
	16	16	14	−8.5 ± 6.5	18.0 ± 3.9	19.9	0.0	1.000
	12	12	14	−18.9 ± 3.1	4.6 ± 5.8	19.4	0.0	0.995
A 274	60	34	29	−4.8 ± 5.3	−7.1 ± 3.6	20.1	19.4	1.000
	56	32	36	5.5 ± 4.5	0.7 ± 6.1	19.6	19.0	1.000
	44	25	26	−6.4 ± 3.8	7.2 ± 3.6	19.7	19.2	1.000
A 277	69	30	34	4.2 ± 6.6	−5.1 ± 4.5	19.6	18.6	1.000
	50	22	14	2.2 ± 3.6	5.9 ± 5.5	19.9	19.4	1.000
A 389	104	60	62	−0.4 ± 6.6	0.5 ± 4.6	20.2	19.2	1.000
	38	22	25	−4.5 ± 7.4	−9.5 ± 3.2	20.2	20.0	1.000
	31	18	13	7.4 ± 4.8	10.4 ± 2.8	20.4	20.4	1.000
A 400	52	27	28	−29.0 ± 17.6	−1.3 ± 16.1	17.3	16.9	1.000
	38	20	20	2.5 ± 10.4	−7.1 ± 9.2	17.3	17.1	1.000
A 415	81	33	52	−3.1 ± 6.0	5.7 ± 4.4	19.8	18.8	1.000
	69	28	20	−3.4 ± 4.4	−9.0 ± 5.4	20.0	19.4	1.000
A 496	110	49	51	6.3 ± 14.0	−0.1 ± 13.6	18.0	17.1	1.000
	69	31	33	24.8 ± 20.5	39.6 ± 10.8	18.0	17.2	1.000
A 514	101	36	32	−11.3 ± 5.7	−2.3 ± 6.3	19.6	18.6	1.000
	58	21	27	−10.4 ± 7.0	−18.3 ± 2.9	19.4	18.6	1.000
	27	10	9	14.2 ± 4.4	5.1 ± 4.7	19.5	19.7	1.000
A634	44	62	65	−12.6 ± 19.0	16.1 ± 20.8	17.9	17.6	1.000
	23	32	31	38.0 ± 19.2	−14.4 ± 39.7	17.7	18.0	1.000
A 779	45	39	39	−10.6 ± 16.9	−17.6 ± 18.8	17.5	17.1	1.000
	28	24	14	29.4 ± 16.8	18.7 ± 22.7	17.5	17.6	1.000
A 787	113	73	56	4.5 ± 4.7	2.2 ± 5.6	20.3	19.0	1.000
	27	18	32	−9.1 ± 2.4	2.5 ± 3.9	20.0	20.4	1.000
A 957	151	52	56	6.7 ± 16.9	2.3 ± 11.8	18.6	17.5	1.000
	86	30	30	−17.0 ± 15.0	−22.4 ± 10.5	18.7	17.8	1.000
	51	18	14	−6.7 ± 22.0	33.2 ± 4.3	18.9	18.5	1.000

TABLE 4.4. (continued)

Cluster	N	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 978	119	40	42	1.2 ± 13.9	2.3 ± 8.0	18.9	17.9	1.000
	107	36	36	6.9 ± 10.0	−19.6 ± 7.1	19.0	17.8	1.000
	39	13	14	1.7 ± 11.9	22.3 ± 5.5	18.8	18.6	1.000
A 993	117	43	48	−7.5 ± 11.4	8.0 ± 11.3	18.9	17.7	1.000
	78	29	20	9.1 ± 16.3	−22.1 ± 6.7	19.2	18.3	1.000
	40	15	18	17.9 ± 6.0	21.5 ± 6.0	18.6	18.1	1.000
A 1020	163	62	66	4.8 ± 9.9	9.3 ± 9.4	19.5	17.9	1.000
	102	38	34	−12.9 ± 12.5	−9.7 ± 12.1	19.7	18.7	1.000
A 1126	98	40	26	−2.8 ± 6.5	−1.9 ± 4.1	19.8	18.7	1.000
	89	36	29	5.9 ± 11.8	11.5 ± 4.5	19.8	18.5	1.000
	41	17	40	8.8 ± 4.4	−6.4 ± 4.3	20.1	19.7	1.000
A 1139	84	50	58	−2.7 ± 12.4	8.2 ± 13.1	18.0	17.0	1.000
	44	26	18	28.3 ± 11.2	−12.2 ± 26.8	18.4	18.1	1.000
	20	12	12	−24.4 ± 7.7	−22.5 ± 6.5	18.7	19.0	1.000
A 1185	152	45	42	−20.5 ± 23.5	−19.7 ± 39.2	17.8	16.6	1.000
	126	38	41	−0.9 ± 11.1	1.3 ± 8.9	17.7	16.5	1.000
A 1187	78	34	36	3.2 ± 9.0	−2.9 ± 7.1	19.7	18.7	1.000
	56	25	21	−16.6 ± 2.9	−6.4 ± 7.0	19.6	18.9	1.000
	36	16	16	−7.1 ± 6.8	16.2 ± 2.6	19.6	19.3	1.000
	31	14	14	−5.5 ± 3.9	5.5 ± 2.8	19.4	19.4	1.000
A 1213	139	53	63	9.0 ± 12.4	2.9 ± 13.5	18.6	17.1	1.000
	63	24	18	−18.8 ± 7.7	13.3 ± 10.4	18.8	18.2	1.000
	59	23	18	−19.4 ± 21.8	−23.1 ± 10.8	18.7	18.0	1.000
A 1216	47	46	45	14.0 ± 9.4	−5.7 ± 13.1	19.3	18.9	1.000
	47	46	48	−12.3 ± 10.5	6.0 ± 17.0	19.1	18.7	1.000
A 1238	82	46	52	−0.5 ± 6.5	0.3 ± 9.3	19.7	18.8	1.000
	42	23	21	15.2 ± 5.6	−13.8 ± 8.4	19.9	19.7	1.000
	35	19	16	−16.8 ± 2.8	17.5 ± 3.3	20.1	20.0	1.000
A 1254	88	33	24	−3.0 ± 5.7	4.7 ± 7.7	19.8	19.0	1.000
	73	28	25	−15.8 ± 7.7	−12.5 ± 8.3	19.6	18.8	1.000
	66	25	37	15.6 ± 7.2	−10.2 ± 8.7	19.6	18.7	1.000
A 1257	57	27	35	−36.8 ± 10.3	10.6 ± 28.8	17.9	17.2	1.000
	53	25	18	23.4 ± 16.2	30.7 ± 14.5	18.6	17.7	1.000
	54	25	20	17.1 ± 20.1	−35.4 ± 9.9	18.2	17.5	1.000
	48	23	26	−3.6 ± 9.9	0.4 ± 11.9	17.9	17.2	1.000
A 1291	103	26	31	−1.0 ± 7.5	−5.0 ± 8.6	18.8	17.5	1.000
	94	24	20	−8.0 ± 7.5	9.1 ± 6.3	18.9	18.0	1.000
A 1318	116	41	63	−0.1 ± 12.1	−9.7 ± 9.4	19.1	17.5	1.000
	74	26	20	5.0 ± 8.5	12.9 ± 7.1	19.0	18.1	1.000

TABLE 4.4. (continued)

Cluster	N	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 1364	98	43	49	0.6 ± 4.6	2.9 ± 5.4	19.7	18.5	1.000
	71	31	28	−10.8 ± 3.4	−6.7 ± 6.0	19.8	18.9	1.000
A 1365	49	30	37	0.8 ± 4.5	−1.5 ± 3.5	19.3	18.6	1.000
	48	29	28	−2.6 ± 8.9	−12.2 ± 3.9	19.4	18.8	1.000
	25	15	16	−15.7 ± 3.6	5.3 ± 5.6	19.4	19.7	1.000
	16	10	8	18.8 ± 2.6	−13.7 ± 5.7	19.7	0.0	1.000
	15	9	7	−0.6 ± 2.1	15.8 ± 2.8	19.7	0.0	1.000
A 1367	130	65	72	5.3 ± 25.6	−3.6 ± 21.6	17.0	15.8	1.000
	42	21	19	−3.0 ± 22.7	30.6 ± 11.3	17.5	17.3	1.000
A 1377	165	41	57	−5.0 ± 9.0	−1.3 ± 10.0	18.7	17.1	1.000
	134	33	23	19.5 ± 8.7	2.9 ± 22.4	18.9	17.7	1.000
	103	26	20	−23.9 ± 7.1	4.0 ± 21.4	19.1	18.3	1.000
A 1382	78	43	50	0.6 ± 5.3	−3.8 ± 5.0	20.0	19.2	1.000
	56	31	29	−7.0 ± 6.9	7.2 ± 5.4	20.4	19.7	1.000
	25	14	11	11.6 ± 2.7	7.5 ± 4.7	20.2	20.3	0.999
A 1399	83	29	25	0.8 ± 4.9	2.5 ± 4.9	19.6	18.6	1.000
	79	28	20	4.8 ± 6.5	−9.6 ± 4.8	19.9	18.9	1.000
	61	21	12	−8.7 ± 5.9	13.8 ± 3.8	20.1	19.5	1.000
A 1436	170	47	45	1.8 ± 7.4	−3.0 ± 7.0	19.2	17.9	1.000
	145	41	46	−6.8 ± 15.6	11.1 ± 10.3	19.3	17.8	1.000
A 1468	95	57	49	2.1 ± 10.3	−6.9 ± 6.8	19.8	18.7	1.000
	53	32	40	−6.9 ± 7.8	8.6 ± 5.6	19.7	19.0	1.000
A 1474	134	72	69	3.4 ± 9.5	4.9 ± 10.9	19.6	18.3	1.000
	53	28	31	−13.3 ± 4.6	−7.1 ± 8.0	19.7	19.0	1.000
A 1496	173	49	59	4.0 ± 7.1	−2.6 ± 7.1	19.5	17.8	1.000
	125	35	30	−12.2 ± 4.6	2.8 ± 9.9	19.9	18.3	1.000
A 1541	98	48	76	5.9 ± 8.3	8.3 ± 5.3	19.7	18.4	1.000
	95	46	22	−4.6 ± 5.9	−2.8 ± 5.9	20.0	19.0	1.000
A 1644	156	53	52	2.6 ± 19.5	−11.9 ± 11.4	18.7	17.3	1.000
	88	30	31	−3.7 ± 10.3	9.1 ± 7.2	18.7	17.6	1.000
A1651	142	69	69	0.6 ± 9.0	−1.7 ± 7.6	19.8	18.6	1.000
	48	23	23	−11.8 ± 9.4	11.9 ± 10.5	19.9	19.5	1.000
A 1656	196	46	45	0.4 ± 20.6	−3.6 ± 20.1	17.0	15.3	1.000
	165	39	43	30.5 ± 32.4	3.0 ± 51.5	17.1	15.4	1.000
	57	13	10	−46.9 ± 12.1	14.0 ± 14.1	17.1	16.4	1.000
A 1691	163	66	75	4.2 ± 9.1	−1.8 ± 10.5	19.2	17.7	1.000
	34	14	10	−17.6 ± 4.1	−13.8 ± 6.3	19.6	19.4	1.000
	32	13	9	13.9 ± 5.3	16.4 ± 3.1	19.5	19.4	1.000

TABLE 4.4. (continued)

Cluster	N	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A1749	128	58	66	8.0 ± 10.9	0.8 ± 10.4	19.0	17.7	1.000
	63	29	26	−15.3 ± 9.8	−15.5 ± 9.9	19.2	18.5	1.000
A 1767	277	90	92	−0.6 ± 10.7	−0.3 ± 12.4	19.3	17.8	1.000
	31	10	8	−19.1 ± 3.6	−16.3 ± 5.0	19.4	19.4	1.000
A 1773	133	47	51	0.0 ± 6.7	−0.8 ± 5.8	19.6	18.4	1.000
	100	35	34	−14.1 ± 5.3	8.6 ± 8.1	19.7	18.7	1.000
A 1775	181	68	68	7.7 ± 10.7	1.2 ± 12.1	19.6	17.9	1.000
	87	32	32	−12.7 ± 6.6	−8.0 ± 7.9	19.4	18.4	1.000
A 1793	135	54	54	0.6 ± 5.8	2.3 ± 6.2	19.6	18.5	1.000
	47	19	27	5.7 ± 8.0	−14.0 ± 4.1	20.1	19.8	1.000
	46	19	12	−15.5 ± 3.2	1.7 ± 10.2	20.2	19.9	1.000
A 1795	108	38	41	−7.8 ± 13.3	15.2 ± 8.2	19.2	18.2	1.000
	89	31	34	3.3 ± 6.2	−3.5 ± 6.8	19.2	18.3	1.000
	65	23	17	−10.8 ± 9.0	−14.0 ± 5.8	19.7	19.0	1.000
A 1809	108	35	28	7.3 ± 6.3	−4.1 ± 4.3	19.7	18.6	1.000
	61	20	19	1.3 ± 3.9	3.5 ± 4.9	19.5	18.7	1.000
A 1831	105	34	33	6.0 ± 8.4	7.1 ± 7.3	19.4	18.3	1.000
	85	28	38	0.3 ± 7.7	−1.4 ± 6.2	19.3	18.2	1.000
	72	23	21	−10.2 ± 6.7	−12.3 ± 5.1	19.5	18.8	1.000
A 1837	89	33	33	−19.0 ± 17.5	18.3 ± 12.6	18.5	17.6	1.000
	73	27	29	−5.6 ± 10.2	−26.8 ± 9.9	18.6	17.9	1.000
	53	20	18	8.2 ± 7.7	−2.8 ± 7.0	18.7	18.3	1.000
A 1904	161	42	52	2.8 ± 6.5	−3.7 ± 8.8	19.2	17.7	1.000
	149	39	27	−0.4 ± 16.6	10.3 ± 11.1	19.7	18.4	1.000
	76	20	22	−10.3 ± 6.7	−16.3 ± 5.2	19.3	18.3	1.000
A 1913	101	37	38	7.2 ± 7.2	5.0 ± 8.6	19.0	17.9	1.000
	66	24	17	−5.6 ± 8.4	−11.3 ± 7.6	19.3	18.7	1.000
	55	20	21	−13.8 ± 7.9	12.8 ± 8.0	19.2	18.7	1.000
A 1927	148	60	56	2.8 ± 10.4	−1.6 ± 7.9	19.7	18.4	1.000
	59	24	28	−10.8 ± 6.5	14.9 ± 5.4	19.8	19.1	1.000
	38	16	15	10.5 ± 7.5	−16.0 ± 3.1	19.8	19.6	1.000
A 1983	211	48	47	6.7 ± 16.7	12.4 ± 17.8	18.7	16.9	1.000
	103	23	22	−5.3 ± 11.1	−2.5 ± 5.4	18.6	17.4	1.000
	68	15	19	−24.8 ± 8.4	−22.4 ± 9.4	18.5	17.8	1.000
A 1991	135	37	42	−2.7 ± 7.5	−3.9 ± 8.9	18.9	17.8	1.000
	128	35	31	−0.3 ± 18.3	17.1 ± 8.0	19.3	18.1	1.000
	85	23	23	16.9 ± 7.8	−16.3 ± 7.6	19.1	18.2	1.000
	20	5	5	−20.1 ± 5.4	−24.3 ± 3.1	19.2	19.8	1.000

TABLE 4.4. (continued)

Cluster	N	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 1999	102	55	53	−5.9 ± 7.4	5.2 ± 5.0	20.0	18.9	1.000
	85	45	47	0.5 ± 7.6	−7.4 ± 5.1	19.8	18.5	1.000
A 2005	106	76	82	−0.7 ± 7.4	3.2 ± 5.8	20.2	18.7	1.000
	33	24	18	6.1 ± 4.1	−8.2 ± 3.4	20.2	20.1	1.000
A 2022	137	43	60	0.9 ± 8.8	4.3 ± 6.7	19.0	17.4	1.000
	75	23	17	9.4 ± 9.8	−16.0 ± 7.4	19.5	18.8	1.000
A 2028	76	33	32	−0.1 ± 5.0	0.6 ± 5.5	19.9	18.8	1.000
	75	32	40	−10.6 ± 7.6	11.7 ± 6.3	19.6	18.7	1.000
A 2029	186	43	41	−2.8 ± 6.9	−2.0 ± 6.2	19.6	18.4	1.000
	119	27	25	15.0 ± 6.4	−5.7 ± 13.1	19.7	18.6	1.000
	67	15	13	5.2 ± 4.2	14.3 ± 3.9	19.7	19.1	1.000
A 2040	108	39	42	−2.4 ± 8.6	−0.9 ± 12.5	18.7	17.8	1.000
	70	25	26	17.3 ± 14.4	23.5 ± 10.3	18.9	18.3	1.000
	59	21	18	21.2 ± 9.0	−19.3 ± 9.7	19.1	18.7	1.000
A 2048	127	40	44	−0.3 ± 5.0	−3.3 ± 5.5	19.8	18.6	1.000
	121	39	38	8.3 ± 6.2	8.9 ± 5.6	20.0	19.0	1.000
A 2063	134	64	67	−0.3 ± 21.9	−3.9 ± 14.6	18.1	17.3	1.000
	51	24	18	16.3 ± 28.4	−40.8 ± 7.9	18.4	18.1	1.000
	26	12	14	−31.3 ± 15.3	38.4 ± 10.4	18.1	18.4	1.000
A 2065	181	43	44	3.2 ± 4.5	−0.5 ± 6.6	19.6	18.6	1.000
	136	32	30	−0.1 ± 16.2	15.1 ± 7.7	19.7	18.7	1.000
A 2067	136	48	45	3.1 ± 7.5	4.3 ± 9.8	19.8	18.9	1.000
	75	27	28	17.7 ± 4.9	−14.1 ± 4.7	19.7	19.1	1.000
	42	15	19	−12.3 ± 4.9	15.1 ± 5.2	19.6	19.2	1.000
	30	11	8	−16.3 ± 4.5	−14.3 ± 6.3	20.0	20.0	1.000
A 2079	142	45	49	−15.7 ± 7.7	9.9 ± 9.5	19.5	18.2	1.000
	121	38	35	5.4 ± 9.9	−6.2 ± 6.5	19.6	18.7	1.000
A 2089	86	54	59	0.5 ± 8.9	4.2 ± 8.4	19.5	18.7	1.000
	37	23	22	−12.1 ± 6.8	−7.3 ± 6.2	19.6	19.4	1.000
A 2092	95	36	37	−0.3 ± 7.6	0.1 ± 5.2	19.2	18.3	1.000
	50	19	17	−11.6 ± 6.3	16.3 ± 5.1	19.4	18.8	1.000
A 2107	106	39	45	3.8 ± 10.9	1.9 ± 7.6	18.4	17.5	1.000
	69	25	25	2.5 ± 19.5	−28.1 ± 7.5	18.8	18.1	1.000
	69	25	20	−29.2 ± 8.5	12.8 ± 23.6	18.7	18.0	1.000
A 2124	138	46	48	0.1 ± 8.2	2.7 ± 6.8	19.3	18.2	1.000
	75	25	26	−11.8 ± 9.5	−18.0 ± 6.2	19.5	18.7	1.000
A 2142	134	43	42	−0.2 ± 7.7	−4.6 ± 5.5	20.1	19.0	1.000
	68	22	18	11.9 ± 4.3	8.7 ± 5.6	20.3	19.7	1.000
	63	20	24	−7.2 ± 6.4	12.4 ± 5.0	20.1	19.4	1.000

TABLE 4.4. (continued)

Cluster	N	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 2147	156	34	33	−28.3 ± 11.0	19.0 ± 18.4	18.2	17.0	1.000
	111	24	24	7.6 ± 15.6	28.7 ± 9.7	18.3	17.2	1.000
	105	23	26	4.9 ± 9.0	−7.3 ± 14.1	18.3	17.0	1.000
A 2151	118	30	29	4.9 ± 24.2	27.3 ± 10.1	18.3	17.0	1.000
	116	30	24	15.5 ± 18.5	−27.8 ± 10.6	18.5	17.2	1.000
	82	21	30	2.6 ± 7.9	−0.8 ± 6.9	18.2	17.1	1.000
	50	13	12	−24.7 ± 8.4	0.0 ± 9.5	18.1	17.6	1.000
A 2152	162	34	39	31.9 ± 10.4	−19.9 ± 14.4	18.4	16.8	1.000
	112	24	21	8.9 ± 15.2	17.9 ± 10.4	18.4	17.5	1.000
	84	18	18	−7.3 ± 9.6	−7.7 ± 9.2	18.3	17.3	1.000
A 2162	51	41	44	−0.3 ± 25.0	25.6 ± 12.7	18.0	17.6	1.000
	32	26	33	−5.2 ± 16.4	−13.1 ± 12.5	18.5	18.4	1.000
A 2175	121	27	27	−0.7 ± 3.1	−0.6 ± 3.3	20.1	18.3	1.000
	102	23	23	−7.8 ± 2.8	−5.4 ± 2.8	20.3	18.4	1.000
	79	18	18	8.4 ± 4.4	9.4 ± 4.2	20.3	18.9	1.000
A 2197	152	49	51	−14.5 ± 25.3	−28.5 ± 16.3	17.7	16.0	1.000
	91	29	32	10.9 ± 18.2	3.5 ± 11.5	17.6	16.6	1.000
	36	12	7	27.3 ± 16.6	42.3 ± 9.7	18.3	18.2	1.000
A 2199	130	33	38	−9.9 ± 15.2	9.7 ± 12.7	17.8	16.5	1.000
	93	24	23	16.2 ± 12.9	−16.6 ± 17.2	17.8	16.8	1.000
A 2255	315	76	78	−3.8 ± 8.5	−2.4 ± 9.8	19.6	17.9	1.000
	50	12	14	2.6 ± 4.3	14.6 ± 3.3	19.4	18.9	1.000
A 2256	304	67	69	−0.4 ± 10.7	−2.8 ± 8.0	19.0	17.2	1.000
	71	16	10	−11.3 ± 7.6	−21.8 ± 3.9	19.4	18.6	1.000
A 2328	82	63	63	2.3 ± 5.6	−2.6 ± 4.8	20.2	19.1	1.000
	30	23	12	−7.0 ± 2.5	3.9 ± 3.7	20.3	20.3	1.000
	19	15	25	0.0 ± 2.6	8.6 ± 2.0	20.2	0.0	1.000
A 2347	41	46	58	−3.7 ± 7.3	6.6 ± 5.3	20.5	20.1	1.000
	29	32	27	−1.0 ± 5.1	−8.6 ± 3.4	20.5	20.6	1.000
	15	17	11	7.8 ± 3.8	−0.3 ± 2.5	21.0	0.0	0.997
A 2382	95	48	51	3.1 ± 11.3	−8.6 ± 7.7	19.3	18.5	1.000
	54	27	26	6.2 ± 6.1	7.5 ± 7.4	19.4	18.9	1.000
	16	8	9	−14.8 ± 4.8	19.3 ± 3.6	19.2	0.0	1.000
A 2384	54	43	42	−1.4 ± 4.7	1.7 ± 4.2	19.9	19.4	1.000
	32	25	30	2.3 ± 5.8	−10.1 ± 4.1	19.9	19.9	1.000
	31	24	24	2.8 ± 9.6	13.4 ± 4.1	20.1	20.0	1.000
A 2399	57	22	23	−5.3 ± 6.7	−7.9 ± 8.7	19.2	18.5	1.000
	55	21	24	9.6 ± 8.4	14.8 ± 8.0	18.9	18.1	1.000
	50	20	24	5.1 ± 6.0	−1.2 ± 4.3	18.6	18.0	1.000

TABLE 4.4. (continued)

Cluster	N	$\%N_{tot}$	$\%L_{tot}$	$x \pm \sigma_x$ (arcmin)	$y \pm \sigma_y$ (arcmin)	m_{med} (O)	m_{jm} (O)	S
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 2410	134	57	71	1.9 ± 11.3	0.1 ± 7.2	19.3	17.6	1.000
	60	26	17	−11.3 ± 5.2	12.9 ± 5.7	19.5	18.9	1.000
	38	16	11	4.2 ± 7.8	−14.4 ± 3.5	19.7	19.5	1.000
A 2457	122	49	51	0.7 ± 8.6	2.9 ± 7.6	19.2	17.8	1.000
	91	37	37	12.8 ± 10.4	−4.0 ± 19.3	19.2	18.0	1.000
	30	12	11	−19.5 ± 5.2	−4.2 ± 11.2	19.1	19.1	1.000
A 2634	119	29	33	−0.4 ± 10.8	−0.1 ± 16.0	17.9	16.9	1.000
	90	22	23	−26.8 ± 15.6	−33.7 ± 14.4	17.9	17.0	1.000
	56	14	10	44.1 ± 8.6	18.3 ± 9.2	18.3	17.9	1.000
A 2657	78	46	47	2.5 ± 11.6	−0.3 ± 9.7	18.4	17.8	1.000
	38	22	20	−25.2 ± 9.6	−17.6 ± 13.3	18.5	18.3	1.000
A 2666	62	36	28	35.0 ± 15.7	4.3 ± 36.9	18.0	17.5	1.000
	44	26	42	−46.4 ± 14.1	31.7 ± 17.7	17.3	16.6	1.000
	43	25	17	−16.6 ± 23.0	−45.1 ± 14.9	18.0	17.8	1.000
	22	13	13	−7.7 ± 11.1	1.7 ± 10.0	17.7	17.9	0.999
A 2670	115	45	50	0.1 ± 4.4	3.9 ± 7.7	19.0	18.1	1.000
	58	23	22	−14.7 ± 5.1	9.4 ± 7.3	19.3	18.7	1.000
	46	18	18	9.5 ± 6.0	−11.9 ± 5.9	19.2	18.8	1.000
A 2675	78	46	47	2.5 ± 11.6	−0.3 ± 9.7	18.4	17.8	1.000
	38	22	20	−25.2 ± 9.6	−17.6 ± 13.3	18.5	18.3	1.000
A 2700	36	28	32	3.4 ± 3.8	−3.2 ± 3.6	19.6	19.3	1.000
	19	15	15	−6.3 ± 2.3	3.0 ± 3.6	19.6	0.0	0.995
	14	11	11	3.2 ± 3.5	8.9 ± 1.8	19.6	0.0	0.993

4.4 Background/Foreground Cluster Identification

With the larger groups and unbinned magnitudes of the APS data, identification of possible background/foreground groups can be carried out by applying a Kolmogorov-Smirnov (K-S) test to the magnitude distributions for the galaxies assigned to the different groups. In theory, if a group is sufficiently far away from the main cluster, along the line of sight, its galaxies will on average be fainter than that of a closer group of galaxies. Use of the K-S test in this fashion depends on the assumption that all clusters and subclusters share a common luminosity function. If galaxies in some clusters or subclusters are intrinsically fainter than others, the results of the K-S test on the magnitudes will be misleading. In general it appears this assumption holds, although there are a number of exceptions (Schechter 1976). Recently, Jones & Mazure (1996) have used the ESO Nearby Abell Cluster Survey (ENACS) to examine this assumption in detail. They conclude that galaxy clusters do *not* have a universal luminosity function, with significant variations occurring at both the bright and faint ends of the distribution. However, the middle of the distribution, from the 10th-ranked galaxy to the 20th-ranked galaxy appears to be the most stable region. Thus, they recommend using the average of the 10th to 20th ranked galaxies, or:

$$m_{jm} = \frac{1}{11} \sum_{i=10}^{20} m_i, \tag{4.20}$$

where the m_i are the sorted magnitudes, to determine a redshift-independent distance estimate to clusters. They again warn that there are exceptions and that in some clusters the galaxies are simply fainter at all magnitudes. Despite these shortcomings, without a massive redshift survey the magnitude distribution is the most reliable way to determine distances to the clusters. Furthermore, subclusters of galaxies that are intrinsically brighter or fainter would, in and of themselves, be interesting in the clues they may hold for galaxy formation.

Thus, with the above caution, the K-S test has been applied to each of the groups identified by either KMM or DEDICA. Any groups which could be rejected as being drawn from the same distribution at greater than the 90% level were considered to be not physically associated with the main cluster. To assign the groups as either background or foreground, the median magnitude and m_{jm} in each group were examined. In general these two numbers agreed. However, as examination of Tables 4.2 and 4.3 reveals, there were a number of cases where they gave opposite results. In these cases, the ratio of the number fraction and the luminosity fractions for the groups was examined. Ratios greater than 1 were considered background, while ratios less than one were classified as foreground. The groups classified as background are plotted in Figure 3.9 with an open circle and those that are foreground with an open square.

A number of clusters have other Abell clusters within an Abell radius and are so labeled in Figure 3.9. In the cases where redshifts were quoted in the literature it is possible to check the results of the K-S test. A85 has two clusters which appear nearby, A87 and A89. Only A87 has a redshift quoted at z=0.055 which is quite close the the value of A85 at z=0.0518. A89 however, is likely to be a background group even though the K-S test fails to reject it. This may result from contamination of the group with galaxies that actually belong to A85, and contamination of A85 with galaxies that actually belong to A89. The case of A1837 and A1836 should provide a warning. Although both the DEDICA and KMM partitions reject the hypothesis that the magnitudes of the two clusters are drawn from the same population at greater than the 95% level, the KMM partition has both m_{med} and m_{jm} greater for A1836 indicating it as a background object, while the opposite is true for for the DEDICA partition. In actual fact A1836 with z=0.0362 is at roughly the same redshift as A1837A at z = 0.03722 but is foreground to a second component to A1837 with z=0.0718. The only other cluster with a redshift that are unambiguously identified

by the substructure tests is the binary cluster A2675 and A2678 which have nearly the same redshift and are likely to be gravitationally bound. The K-S test makes no distinction between the two magnitude distributions.

4.5 Comparison of Results

With background and foreground groups removed, the best estimate for the fraction of clusters in the HGT sample with significant substructure is 64±15%. The estimate of error is the internal error between KMM and DEDICA. This error is only a lower limit because it does not include the errors associated with the K-S determination of background groups. Here a group was considered background/foreground if either the KMM group or a nearby DEDICA group failed to pass the K-S test criteria. According to the K-S test, 20% of the clusters in the sample are contaminated with background groups within an Abell radius. This is midway between the numerical results of van Haarlem (1996), which suggest contamination at the 30% level, and the X-ray results of Briel (1993), which found contamination at the 10% level. Therefore, the results of the K-S test are not wildly off from what is expected.

4.6 Comparison to other Studies

In order to compare the present results with those of other studies the characteristics of the statistic used need to be taken into account. For instance, the study of Rhee *et al.* (1991) found that of their six tests for substructure the test with the highest rate of detection was the Lee test (described by Fitchett 1988), with 10% of the sample clusters having substructure. With all the tests included, 26% of the sample showed some evidence of substructure. However, there are a number of important differences between that study and the work in this thesis. First, in an attempt to keep

background contamination low, Rhee *et al.* only considered the 100 brightest galaxies in each cluster. Second, substructure needed to be significant at the 99% level to be considered "real" by Rhee *et al.* As seen with the Monte Carlo experiments above, with only 100 galaxies in a cluster the 99% significance level is probably too restrictive for KMM, and thus may also be for a number of the tests employed by Rhee *et al.* More importantly, some of the tests used were only sensitive to very specific kinds of substructure. The Lee test is only sensitive to bimodal structures; multi-modal structures tend to lower the significance of the statistic (Fitchett & Webster 1987). Other tests employed, such as the percolation test and the angular separation test may not be sensitive enough to the structures they were designed to detect. Thus, the higher percentage of substructure detected here is due the increased power of the tests employed, especially the ability of KMM and DEDICA to fit clusters with more than two subclusters, as well as the more complete sampling of the luminosity function. In fact, of the 61 clusters common to both studies, 19 were identified by Rhee *et al.* as containing substructure. All of these 19 clusters have been identified by either KMM or DEDICA as containing substructure, though four are probably due to background contamination.

Other recent studies have tended to find more substructure than Rhee *et al.*, in better agreement with the present results. The study by Salvador-Solé *et al.* (1993) found that 50% of the 15 Dressler clusters they looked at had substructure at the 95% significance level. If Abell 1736 is removed from consideration (since it was analyzed as two separate clusters by SSG), the KMM and DEDICA results differ from those of SSG for only three clusters. Substructure is found in A1644 with a significant four-group partition from KMM (with three of the groups having less than 20% of the total number) and a two-group partition from DEDICA. Other tests which use velocity information, such as the Δ-statistic and the ϵ-statistic (Bird 1993), and X-ray

data (Davis 1994), confirm the existence of substructure in the cluster. Although a two-group partition is found by KMM for DC 0247-31, it is likely that this structure is not real. Finally, DEDICA finds a significant second peak in the PDF of A1656.

4.7 Conclusions

There are several conclusions to be drawn. First, the ability of DEDICA to separate and test the significance of close groups of galaxies is clearly superior to that of KMM. Second, DEDICA is not as sensitive as KMM to the choice of boundary for the clusters. On the other hand KMM has more power than DEDICA to detect substructure in the presence of background contamination, as seen in the case of A2256.

One potential problem with the current version of DEDICA is the selection of the smoothing parameter. The derivative of the density, as well as the density, is important in the peak identification procedure. Scott (1992) shows that accurate calculation of the derivative requires a larger smoothing window and more data points than accurate calculation of the density. Furthermore, as will be discussed in the next chapter, Merritt & Trembly (1994) find that when calculating the derivative of the kernel density, the rules for choosing the smoothing parameter do not work very well. Thus it is likely that the current implementation of the LSCV technique for finding the smoothing parameter is undersmoothing and thus obtaining a steeper gradient in the density than the true gradient. In general, this leads to a higher significance of the groups, and smaller group sizes.

It is important to stress that the strength of a two-dimensional analysis lies not in the ability of the statistics to establish, once and for all, whether a given cluster does or does not contain substructure. A complete analysis needs to take advantage

of all available data, including redshifts and X-ray surface-brightness maps. The purpose here is to take the fullest possible advantage of the readily-available galaxy position data offered by digitized sky surveys, in order to provide a guide to clusters which might or might not harbor substructure. Once this type of analysis has been carried out the researcher can more efficiently select clusters for a large redshift survey, identify which galaxies within a cluster should have redshifts measured, predict how the X-ray map for a given cluster is likely to appear, and have a framework within which to discuss possible deviations. This same type of analysis could be done on data from numerical simulations (such as those described by Pinkney *et al.* 1996 or van Haarlem 1996), where problems of interpretation are similar to those encountered in the study of real clusters.

These algorithms have several advantages over alternative techniques for the detection of substructure in projected galaxy positions. They can fit any number of subgroups, unlike the Lee test, which is only sensitive to bimodal structures. Secondly, the KMM algorithm is very robust. Although small numbers of outlying galaxies can perturb the parameters of the fit (the estimated means and covariance matrices of the groups) since all galaxies are assigned to at least one of the groups, KMM very rarely returns such groups as significant. Because KMM fits the groups to two-dimensional Gaussian distributions, a wide variety of shapes can be fit, from spherical to rather elongated structures. DEDICA has even more flexibility in this respect since it does not need the Gaussian assumption. Finally, unlike the method of SSG, these methods are very visual. The positions, shapes, and sizes of the identified groups can be seen in the adaptive-kernel maps and compared to X-ray maps for the clusters, unlike the centroid shift methods which can only give a positive or negative result.

The disadvantages of this, or any two-dimensional analysis, is potential contamination from foreground or background galaxies. While the Monte Carlo experiments

indicate that a constant-density background of less then 15% lowers the significance of substructure, foreground/background clusters pose a more serious problem. With the larger catalogs of the APS data, very distant clusters can be detected by applying a K-S test to the magnitude distributions for each group. Ultimately however, X-ray and/or redshift data will need to be considered to confirm the results. Furthermore, merger events occurring along the line of sight will not be detected. Lastly, KMM fits the groups to two-dimensional Gaussians. Departures of the actual density profiles from Gaussian will reduce the usefulness of the partitions obtained. While this can be guarded against by employing the Hawkins test when the individual groups have a large number of galaxies, for small groups the results are likely to be misleading.

These results indicate that a large fraction of the clusters in the sample of galaxy clusters exhibits evidence of substructure in their projected galaxy distributions. This substructure is very often seen in the core of the clusters, even if not identified in this study. As a result of this, and with the possibility of line of sight mergers, the 64% fraction of clusters with significant substructure is likely to be a lower limit. However, a great deal of redshift information will need to be gathered in the coming years to confirm or deny these results.

Chapter 5

ESTIMATION OF THE COSMIC DENSITY PARAMETER Ω_0

5.1 Introduction

One of the reasons for studying substructure in clusters of galaxies is to place constraints on the curvature of the universe. In this chapter the possibility of using the fraction of clusters with presently-detectable substructure to estimate Ω_0 is explored. First the theory is described, then the cluster catalog in this thesis is used to obtain an estimate of Ω_0. Other possible explanations for large amounts of substructure are also discussed. The argument given below is based on the work of Gunn & Gott (1972) and Richstone, Loeb & Turner (1992, hereafter RLT).

5.2 The Theory

From General Relativity, the equation of motion in a Freidmann universe with the Robertson-Walker metric is given by:

$$\frac{d^2r}{dt^2} = -\frac{4\pi G\rho_i r_i^3}{3} r^{-2} + \frac{\Lambda}{3} r, \tag{5.1}$$

where r_i and ρ_i are the separation of two fundamental observers (*i.e.*, observers that are expanding with the Universe) and the density at any given time t_i. Λ is the

cosmological constant. The analysis of RLT has shown that Λ has little effect on the results for flat cosmologies, *i.e.* $\Omega+\Lambda = 1$. Therefore, for convenience the cosmological constant is assumed to be zero in what follows. Integrating, this equation becomes:

$$\dot{r}^2 = 2E + \frac{8\pi G\rho_i r_i}{3} r^{-1}, \tag{5.2}$$

where E is a constant of integration and has units of specific energy. Equation (5.2) holds for any spherically-symmetric, homogeneous matter distribution with no pressure. It is standard to define:

$$H(t) = \frac{\dot{r}}{r} \text{ and } \Omega(t) = \frac{8\pi G\bar{\rho}}{3H^2}, \tag{5.3}$$

where $\bar{\rho}$ is the mean background density. In terms of the redshift $z = r_0/r - 1$:

$$H(z) = H_0(1+z)\sqrt{1+\Omega_0 z}, \tag{5.4}$$

and

$$\Omega(z) = \left[1 + \frac{1-\Omega_0}{\Omega_0(1+z)}\right]^{-1}. \tag{5.5}$$

Equations (5.4) and (5.5) hold only for a matter-dominated universe; *i.e.* after recombination, or $z \lessapprox 10^3$. For $z \gg \Omega_0^{-1}$ and defining the small quantity ϵ as:

$$\epsilon(z) = \frac{1-\Omega_0}{\Omega_0(1+z)} \ll 1, \tag{5.6}$$

equation (5.4) can be approximated by:

$$\Omega(z) \approx 1 - \epsilon(z). \tag{5.7}$$

It is convenient to characterize the perturbations as the fractional overdensity δ defined as:

$$\delta = \frac{\rho}{\bar{\rho}} - 1. \tag{5.8}$$

Then for the perturbations, equation (5.2) can be rewritten as:

$$\dot{r} = \sqrt{2E + \frac{\Omega_i r_i^3}{r}(1+\delta)}. \tag{5.9}$$

Upon substitution of equation (5.7) this becomes:

$$\dot{u} = H_i \sqrt{\frac{2E}{r_i^2 H_i^2} + \frac{1 - \epsilon_i + \delta_i}{u}}, \tag{5.10}$$

where $u = r/r_i$.

At this point in the calculation RLT set the constant of integration so that $\dot{u}_i = H_i$. In other words the perturbation is initially expanding with the Hubble flow of the background Universe. This choice has been criticized in a study by Bartelmann *et al.* (1993, BES). These authors argue from the Zel'dovich approximation (Zel'dovich 1970, Buchert 1989, 1992) that in fact the existence of a density perturbation at time t_i implies the existence of a potential perturbation, the gradient of which gives rise to a velocity perturbation. They therefore conclude that the time scale for collapse in the RLT analysis is too large. BES find:

$$\dot{u} = H_i u^{-1/2}[(1 - \epsilon_i + \delta_i) + (\epsilon_i - c\delta_i)u]^{1/2}, \tag{5.11}$$

where the constant c is 5/3 in the analysis of BES and is 1 if the perturbation is assumed to be expanding with the Hubble flow.

Inverting equation (5.11) and integrating, the time scale for collapse is given by twice the time needed to reach maximum expansion at u_{max}, or:

$$\tau = 2H_i^{-1} \int_0^{u_{max}} \frac{u^{1/2}}{[(1 - \epsilon_i + \delta_i) + (\epsilon_i - c\delta_i)u]^{1/2}}, \tag{5.12}$$

where the initial time t_i is small compared to the present age of the Universe and has therefore been set to 0 for simplicity. The maximum expansion is found by setting

(5.11) equal to 0. Thus:

$$u_{max} = \frac{1 - \epsilon_i + \delta_i}{c\delta_i - \epsilon_i}, \tag{5.13}$$

from which it can be seen that perturbations with $\delta \leq (\epsilon_i/c)$ will never collapse. Integrating (5.12) and keeping only the leading terms in ϵ_i and δ_i:

$$\tau \approx \frac{\pi}{H_i(c\delta_i - \epsilon_i)^{3/2}}. \tag{5.14}$$

The time τ to collapse can be written in terms of the present age of the Universe t_0 as:

$$t' = \tau/t_0 = \frac{\pi}{T(c\delta - \epsilon_i)^{3/2}}, \tag{5.15}$$

where $T = t_0 H_i$.

Further progress cannot be made until adopting some distribution for the density perturbations δ_i. The choice of a Gaussian distribution is again one of convenience. Although there exists the possibility of someday testing this assumption with a higher-resolution microwave satellite, the present COBE results cannot resolve perturbations on the scale of galaxy clusters. Nevertheless, the Gaussian assumption should be able to give a good first-order estimate. With a Gaussian probability distribution:

$$dP(\delta_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\frac{\delta_i^2}{\sigma^2}\right] d\delta_i, \tag{5.16}$$

the probability of finding a perturbation $\delta_i \geq \delta'$, where δ' at any arbitrary threshold, is:

$$P(\delta_i \geq \delta\prime) = \frac{1}{2}\text{erfc}\left[\frac{\delta'}{\sqrt{2}(\sigma)^2}\right]. \tag{5.17}$$

Solving equation (5.15) for δ_i as a function of t' gives the minimum perturbation size necessary to collapse before t':

$$\delta_i(t') = \frac{1}{c}\left[\left(\frac{\pi}{Tt'}\right)^{2/3} + \epsilon_i\right]. \tag{5.18}$$

Substituting this for δ' in equation (5.17) yields the probability for density perturbations to collapse before t':

$$P(t') = \frac{1}{2}\mathrm{erfc}\left\{\frac{1}{\sqrt{2}c(\sigma)}\left[\left(\frac{\pi}{Tt'}\right)^{2/3} + \epsilon_i\right]\right\}. \tag{5.19}$$

And finally, the fraction of present day clusters which have collapsed within the last time interval t' is given by:

$$\delta F(t') = \frac{P(1) - P(1-t')}{P(1)} \tag{5.20}$$

Equation (5.20) can be evaluated numerically for any Ω and time interval t' once a value of σ, the standard deviation of the distribution of density perturbations, is known. RLT choose σ such that $P(1)$ gives the correct fraction of the Universe currently in virialized clusters of mass $\approx 1 \times 10^{15} h^{-1} M_\odot$. That faction is given by:

$$f = \frac{\langle n \rangle M}{\rho_c \Omega_0}, \tag{5.21}$$

where $\langle n \rangle$ is the number density of rich clusters of mass M and $\rho_c\Omega_0$ is the mean density of the universe. With a number density of $6 \times 10^{-6} h^3$ Mpc^{-3} from Bahcall (1988), and a mean cluster velocity dispersion of 750 km s^{-1} giving a mass of $10^{15} h^{-1} M_\odot$ and a critical density $\rho_c = 1.9 \times 10^{-29} h^2$ g cm^{-3}, RLT find $f = 0.021\Omega_0^{-1}$.

This choice is criticized by BES who argue that the number density of currently-collapsed clusters is poorly known. Indeed the number density quoted from Bahcall is calculated for Abell clusters with richness class greater than 0 and $z \leq 0.08$. From the discussion of the completeness of the Abell catalog given in Chapter 1, this could

be in error by 10 to 30%. Furthermore, a similar measurement given by Postman *et al.* (1992) finds $\langle n \rangle = 1.2 \times 10^{-}5h^3$ Mpc^{-3}, or about twice that of Bahcall. Because of these uncertainties, BES argue that it is better to calculate σ from the assumed power spectrum of the primordial density fluctuations. BES find that the argument of RLT is strengthened by their analysis.

If the more conservative method of RLT is adopted, it can be seen that the use of $c = 5/3$ merely has the effect of changing σ to $(3/5)\sigma$ and has no effect on the probability of collapse given by equation (5.20). Thus, a lack of understanding in the collapse time scales of clusters is normalized out, at least to some extent, in the final calculation. This is not true for errors in the estimate of the current number density of rich clusters of galaxies discussed above, nor for the times scales for relaxation in clusters discussed below.

Along with the errors in the normalization discussed above, the major source of error in determining Ω via the percentage of clusters with substructure is the estimate of the time for such structures to be eliminated by dynamical processes. RLT adopt a value of $t/t_0 = 0.1$ or about the crossing time of a rich cluster. As discussed in Chapter 1 this is only a lower limit on the relaxation time; substructure can not be eliminated on time scales shorter than the crossing time. Numerical simulations suggest that this time is likely to be much higher, in the range of 4 to 10 crossing times, depending on the density profiles adopted for the model clusters (Nakamura *et al.* 1995). The shortest time scales for erasure of substructure are for those clusters with small core radii and steep density profiles. Given the results obtained from clusters acting as gravitational lenses which indicate very small core radii, a value of $t/t_0 = 0.4$ is adopted here.

5.3 Discussion

The results are shown in Figure 5.1, assuming that 64% ± 15% of the clusters have presently-detectable substructure. The solid line indicates the 64% fraction of clusters with substructure, while the dotted lines are the estimated error. The solid curve is calculated using the Bahcall (1988) normalization while the normalization used in the dashed cureve is due to Postman *et al.* (1990). The results indicate $\Omega_0 \gtrapprox 0.4 - 0.6$, though $\Omega = 2.5$ is not ruled out. However, this result needs to be viewed with a great deal of caution. First, the fraction of clusters with substructure presented here is likely to be an underestimate since line of sight mergers will be missed. Second, the relaxation times for clusters is very poorly known. Although a number of numerical simulations have been performed, many of the results can be called into question because of lack of resolution or arbitrary initial conditions. Also, most of these numerical simulations have been conducted for head-on collisions only: substructure resulting from collisions with a non-zero impact parameter is likely to last longer. Furthermore, the model of gravitational collapse of a Gaussian perturbation is very specific. Although BES have shown that the argument of RLT is affected little by the generalization to the collapse of ellipsoidal perturbations, the effects of small scale substructure on the collapse times of larger objects is not yet fully understood. In a CDM dominated universe, structures are expected to grow hierarchically, from the merging of smaller objects into larger ones. A necessary by-product of this is the creation of small scale structures. As shown by Peebles (1990) gravitational collapse of larger objects in the presence of these smaller clumps can be delayed.

Given the above uncertainties in the relaxation times, it may be interested to turn the question around. If a value of Ω_0 is assumed, these results can be used to place limits on the relaxation time scales of the projected galaxy positions in clusters.

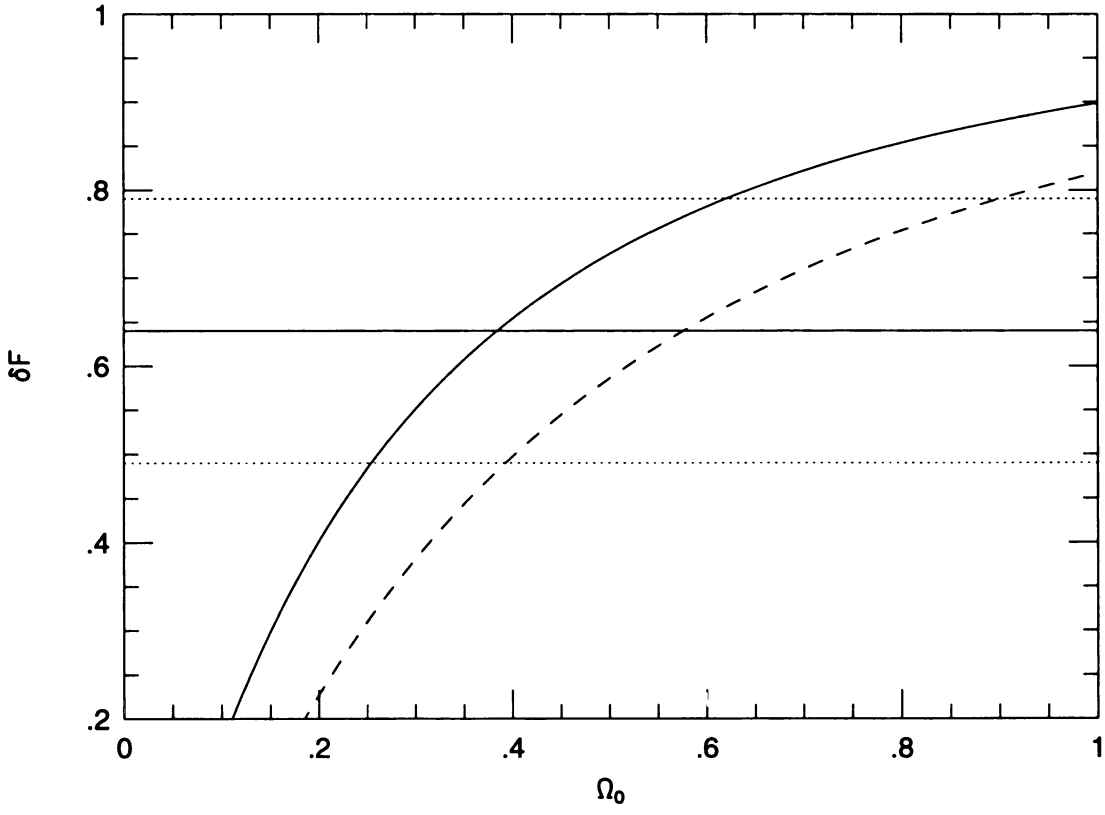

Fig. 5.1.— Fraction of clusters with substructure *vs.* Ω_0. The solid line is for the normalization of Bahcall while the dashed line is for the normalization due to Postman *et al.* The dotted lines show the estimated error in the fraction of clusters with substructure.

As measured by cluster dynamics, $\Omega_0 \approx 0.20$. To reproduce the current fraction of clusters with substructure, the relaxation time of rich clusters needs to be on the order of 5.8-7.4 crossing times (depending on the normalization used) or about $6.6 \times 10^9 h^{-1}$ years.

Chapter 6

RADIAL NUMBER-DENSITY PROFILES

6.1 Introduction

Estimation of the radial density profiles of clusters is important for several reasons. First, numerical simulations which seek to explore dynamical evolution and mergers of clusters generally assume a mass distribution function for the model clusters. Typical forms adopted are ones which involve the density approaching a constant in the core of the cluster. Some models which have been chosen in the past include the modified Hubble law, the Michie-King models, and the non-singular isothermal sphere. A number of the results, such as the survivability of substructure, depend critically on the adopted size of the core radius. In general, smaller core radii lead to shorter relaxation times because groups passing near these cores will be more efficiently disrupted by tidal forces. Furthermore, several clusters are now known to contain arcs of background galaxies. These arcs are formed by the foreground cluster acting as a gravitational lens. The size and curvature of these arcs depend sensitively on the form of the potential well of the lensing object. Since the light reacts to all gravitating material, whether it is due to dark or luminous matter, for the first time it is possible to compare the distribution of dark matter to the distribution of galaxies in clusters.

Furthermore, clusters which have a density cusp at their centers or very small core radii, are much more likely to act as lensing sources. Because the background, lensed objects, are often times very faint they require long exposures with large telescopes to detect. A great deal of observing time could be saved if a way existed to identify, from a large sample of clusters, which ones were likely to be detectable lensing objects. Lastly, formation theories of cD galaxies depend sensitively on the form of the potential well near the cores of clusters. The existence of a density cusp in the core of a cluster makes merger events between galaxies much less likely and mass accretion by tidal stripping more important.

In all of these applications it is crucial to know, not the projected profile which is measured, but the true space-density profile. Because only projected positions are available, it is necessary to make assumptions about the three-dimensional geometry. Statisticians refer to this type of problem as "ill-conditioned." Typically, spherical symmetry is assumed and the de-projection is carried out using Abel's equation:

$$\nu(r) = -\frac{1}{\pi}\int_r^\infty \frac{d\Sigma}{dR}\frac{dR}{\sqrt{R^2 - r^2}}, \tag{6.1}$$

where ν is the space density, r is the three dimensional radius and Σ is the projected distribution with R the projected radius. Note that this depends on the derivative of Σ and not directly on Σ.

It is customary to use a parametric approach when measuring the projected density profiles of clusters. The form is chosen (usually one of the above mentioned forms) and various parameters measured. However, there is a great deal of danger in such a procedure. Even fits which are statistically "good" in a χ^2 sense can have significant deviations from the parametric model. The statistical literature is full of examples where use of a parametric model masks information in the data which is inconsistent with the model (for example see Gasset *et al.* 1984). In this case the

problem is compounded because the Abel inversion requires the derivative be calculated and small errors in Σ can become large errors in ν (Anderssen & Jakeman 1975, Wahba 1990).

These problems were addressed by Merritt & Tremblay (1994). They recommend using a completely nonparametric approach based on a maximum likelihood density estimate. Use of this method allows the smoothing to be carried out directly on the estimate of ν without the necessity of calculating Σ first. What follows is a brief description of the use of maximum likelihood in density estimation and details of the approach used by Merritt & Tremblay.

6.2 Maximum Penalized Likelihood Estimator

Given the success of the maximum likelihood approach used to estimate the PDF of the Gaussian decomposition of clusters in Chapter 4, the question arises: can the maximum-likelihood technique be employed to obtain a nonparametric density estimate? The answer is yes, with a modification.

Given a set of n independent observations, $X_i \ldots X_n$, the likelihood that a curve g represents the underlying density is:

$$L(g) = \prod_{i=1}^{n} g(X_i). \tag{6.2}$$

Unfortunately, this likelihood has no finite maximum. A little thought reveals, for instance, that the likelihood will approach infinity as $h \to 0$ in any of the kernel density estimates. It is easy to see with a box-shaped kernel:

$$\hat{f}(X_i) \geq \frac{1}{nh}, \tag{6.3}$$

and

$$\prod \hat{f}(X_i) \geq \left(\frac{1}{nh}\right)^n. \tag{6.4}$$

Therefore, the maximum-likelihood density estimate is a sum of Dirac delta functions placed at the positions of the observations. While this preserves all of the information in the data, it is not a useful probability density function.

The solution, first applied to density estimation by Good and Gaskins (1971), is to penalize any density estimate which is not smooth to obtain the Maximum Penalized Likelihood (MPL) estimate. Thus, the maximum is sought for the function:

$$\log L_\lambda(f) = \sum_{i=1}^{n} \log f(X_i) - \lambda P(f), \tag{6.5}$$

where $P(f)$ (the penalty function) is some function which quantifies the roughness of the curve f and λ is a smoothing parameter. Typically $P(f)$ is chosen to depend on the squared derivatives of f. Good & Gaskins (1971) suggested using the penalty function:

$$P(f) = \int_{-\infty}^{\infty} \left[\left(\frac{d}{dx}\right)^2 \sqrt{f}\right]^2 dx, \tag{6.6}$$

which will penalize density estimates with large curvature. It is also advantageous to have $P(f)$ depend on the logarithm of f so that the density estimate is forced to be positive. Furthermore, the fluctuations in the estimate are penalized via their *relative* size and not their *absolute* size.

Like the choice of kernel function in the kernel density estimate discussed in Chapter 3, any reasonable choice of the penalty function can be used to provide good estimates of the density as long as the proper smoothing parameter is used. However, as the smoothing parameter λ is increased, the shape of the density estimate will

tend toward the shape of the penalty function, yielding in effect a parametric density estimate. For example, the penalty function defined by:

$$P(f) = \int_{-\infty}^{\infty} \left[\left(\frac{d}{dx} \right)^3 \log f, \right]^2 dx \tag{6.7}$$

is zero if and only if f is a normal distribution. Thus as λ approaches infinity, the MPL estimate will be a normal distribution with the mean and variance of the data.

The behavior of the limiting estimate for large smoothing parameter differs from that of the kernel-based density estimators which approach a constant value as h is increased. In the construction of the maps of Chapter 3 or the identification of peaks in the DEDICA algorithm of Chapter 4 this effect made little difference. It becomes important in constructing radial density profiles when the primary interest is in the behavior near the core. The kernel-based estimator will in general return an estimate which approaches a constant density near the core unless the smoothing parameter is made very small, which leads to a rather noisy estimate.

This problem can be overcome by using MPL with a penalty function of the form employed by MT:

$$P(\nu) = \int \left[d^2 \log \nu / d \log r^2 \right]^2 d(\log r), \tag{6.8}$$

which is zero when $\nu = ar^{-b}$. Thus an oversmoothed estimate will be the best-fit power law approximation. Note also that it is the space density ν that is being penalized and therefore estimated. By calculating the maximum likelihood estimate for the space density directly from the observations, the effect of compounding the errors when calculating the derivative of the estimate of Σ can be avoided. An estimate of Σ, can then be obtained by integrating the estimate of ν, which is a well-conditioned problem. Note that this does not change the nature of the problem

in the sense that it is still ill-conditioned. It is merely that the smoothing is now being performed directly on the derivative of Σ.

As with the kernel estimators, the quality of the density estimate will depend mostly on the choice of smoothing parameter. In this case there are really two density estimates which are sought: the estimate of Σ and the estimate of ν, or in actual fact the derivative of Σ. Although they are related the smoothing parameter that provides the best estimate for the one will not necessarily provide the best estimate for the other. Scott (1992) shows that the derivative of the density requires a larger smoothing parameter and larger number of points to achieve the same MISE as compared to the density estimate. Furthermore, the Monte Carlo simulations performed by MT show that none of the prescription for choosing the smoothing parameter work well in this application. They recommend constructing a number of profiles and only accepting as real those features that appear over a range of smoothing parameters.

Fig. 6.1.— Nonparametric number-density profiles - HGT clusters

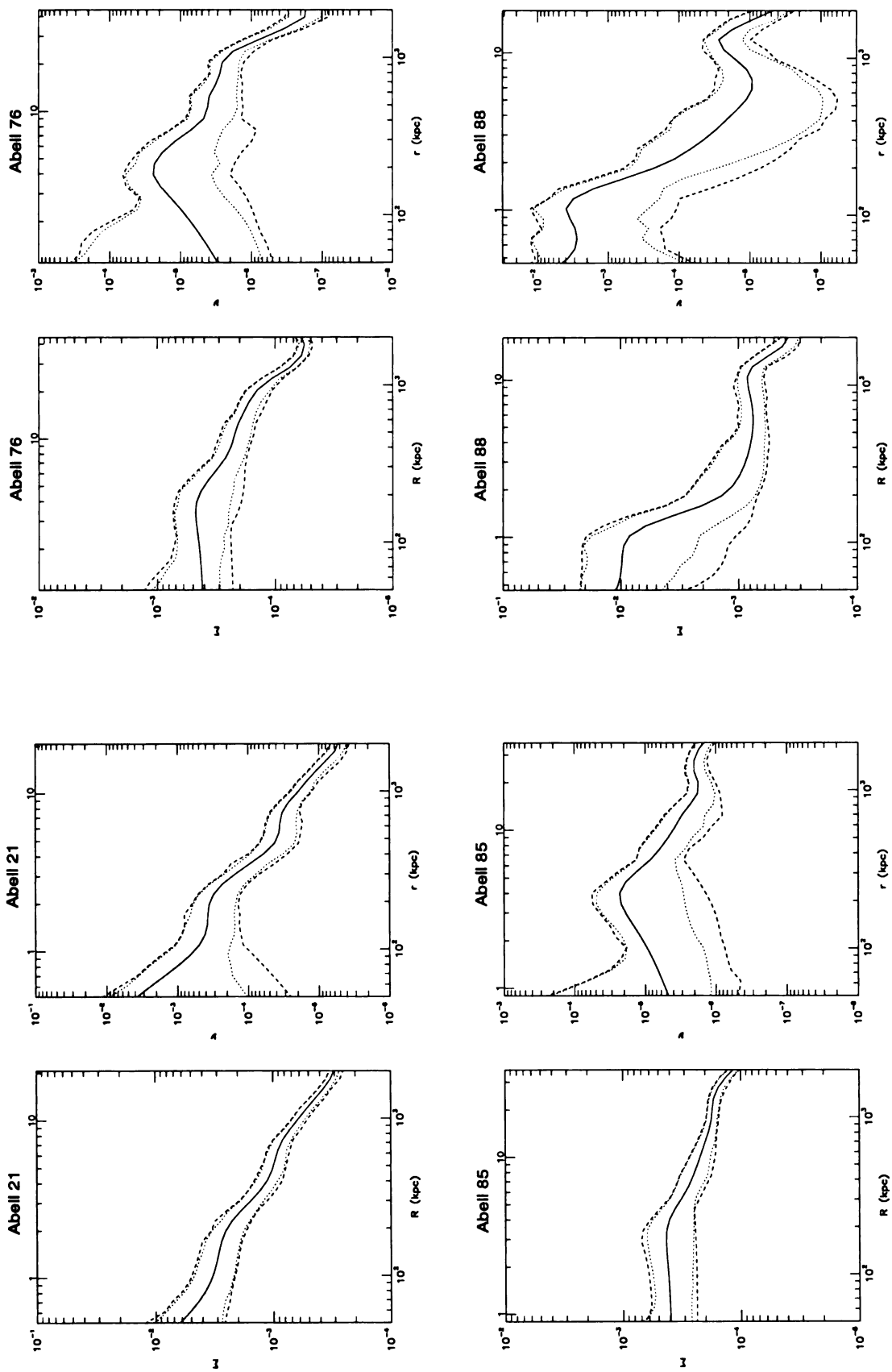

Figure 6.1. Nonparametric density profiles

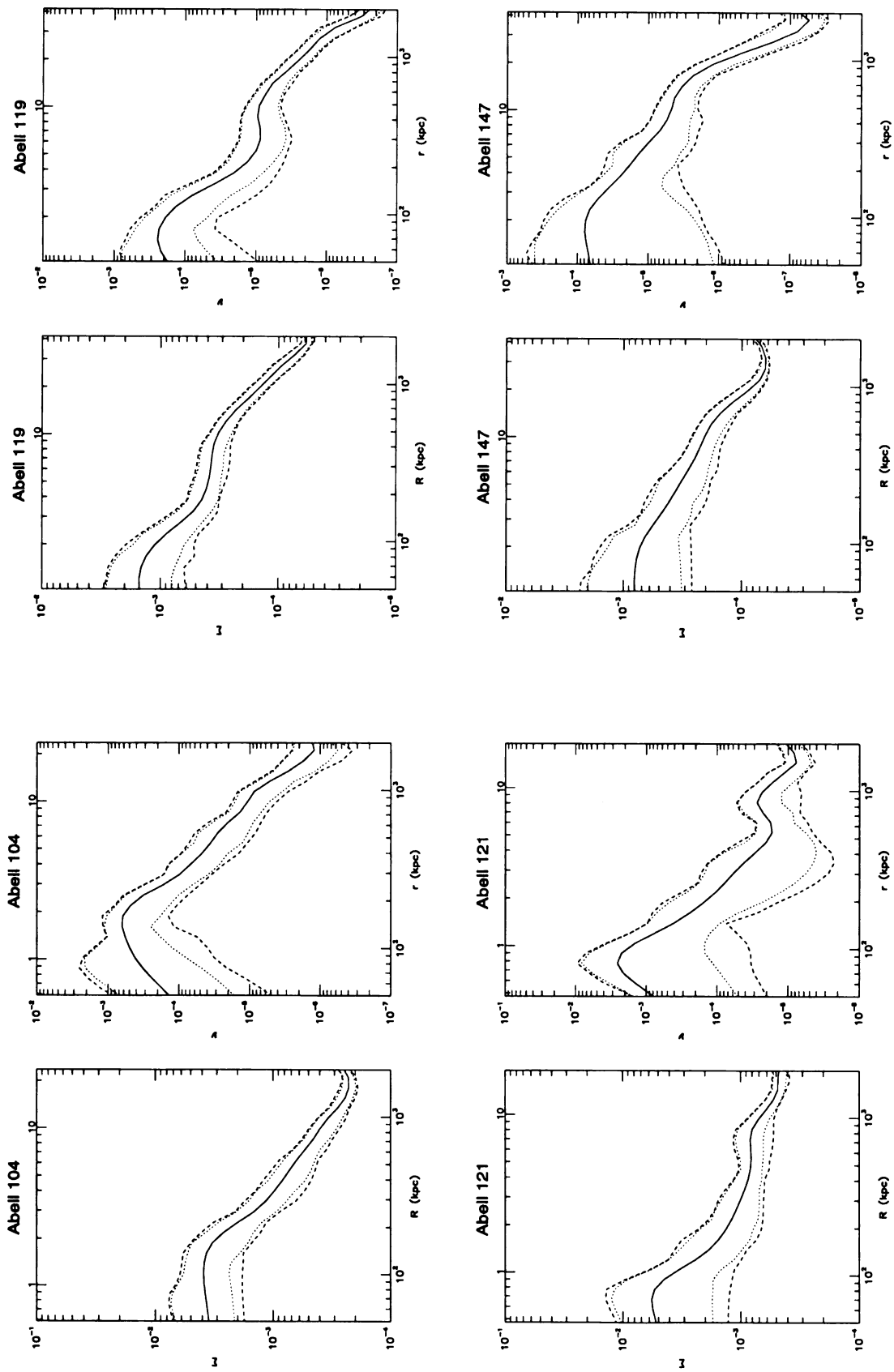

Figure 6.1. (cont'd)

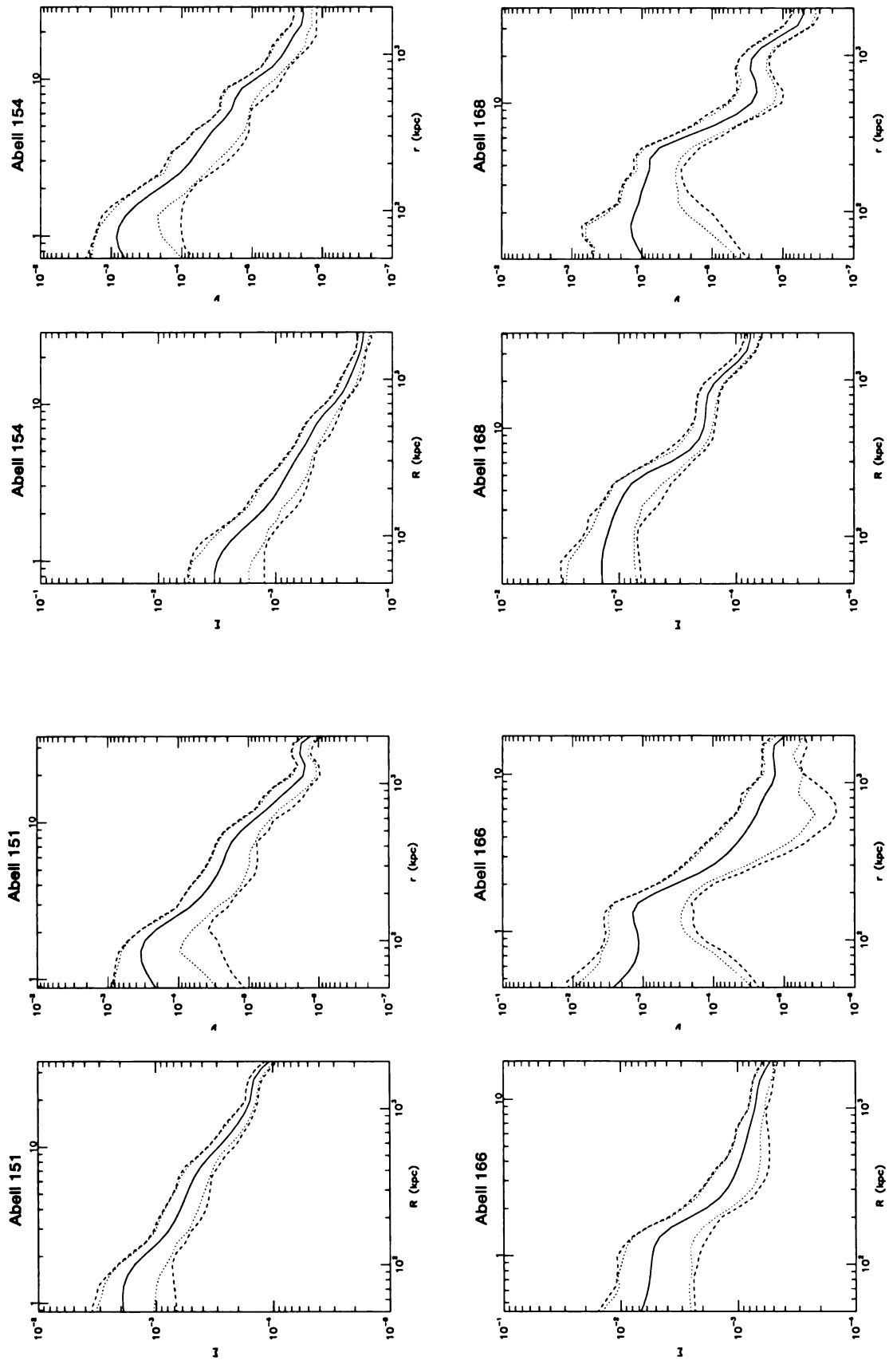

Figure 6.1. (cont'd)

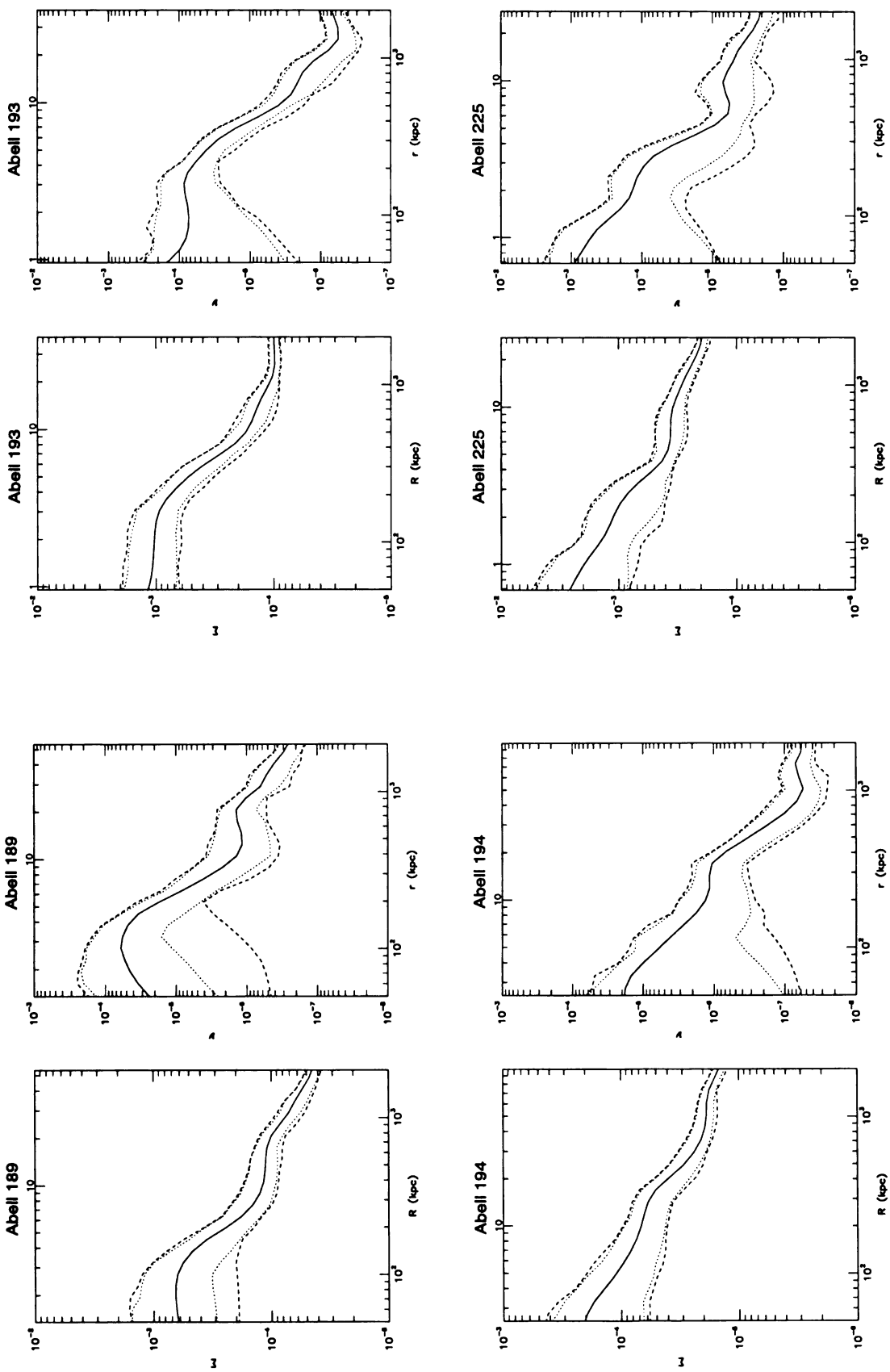

Figure 6.1. (cont'd)

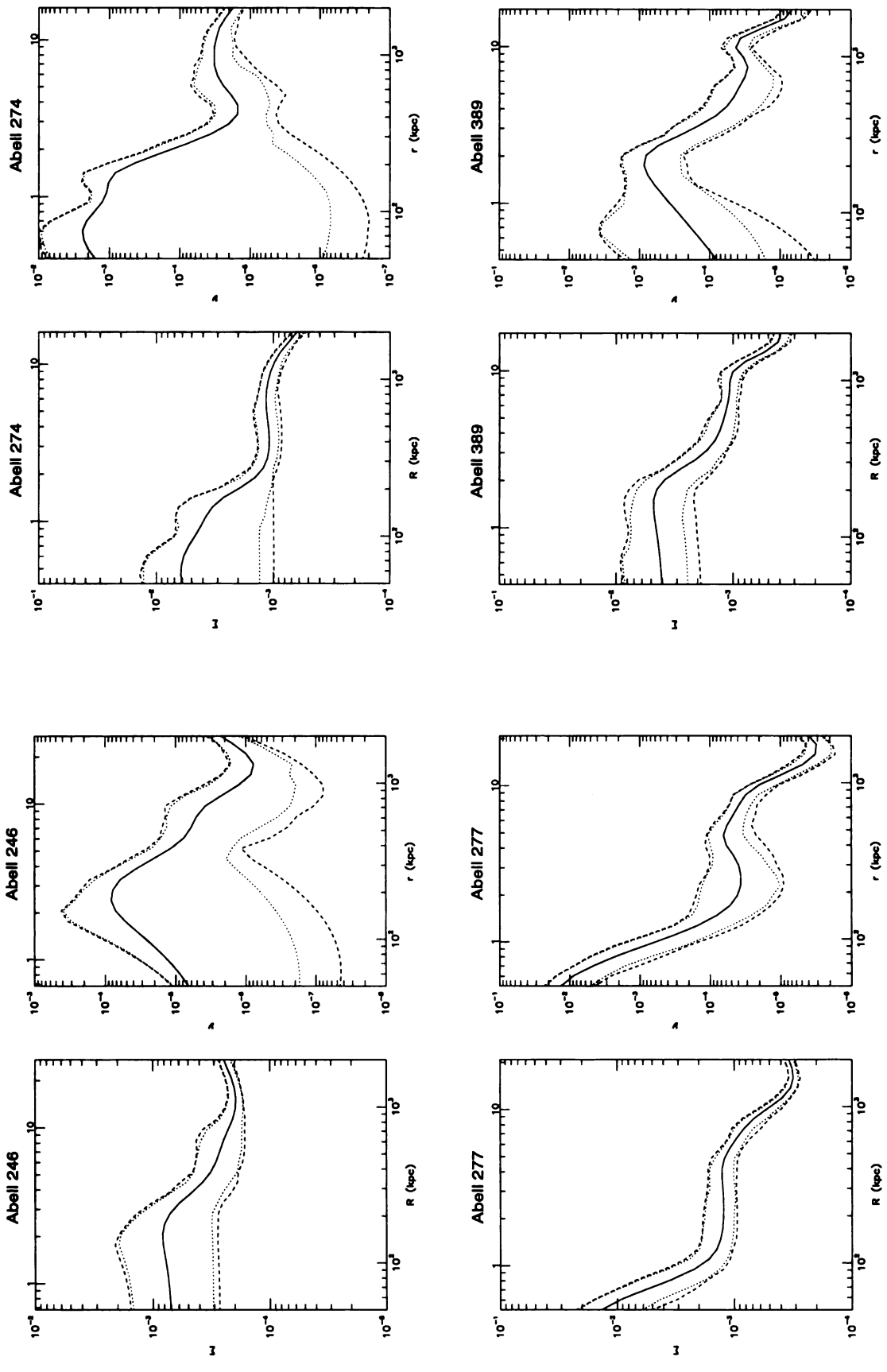

Figure 6.1. (cont'd)

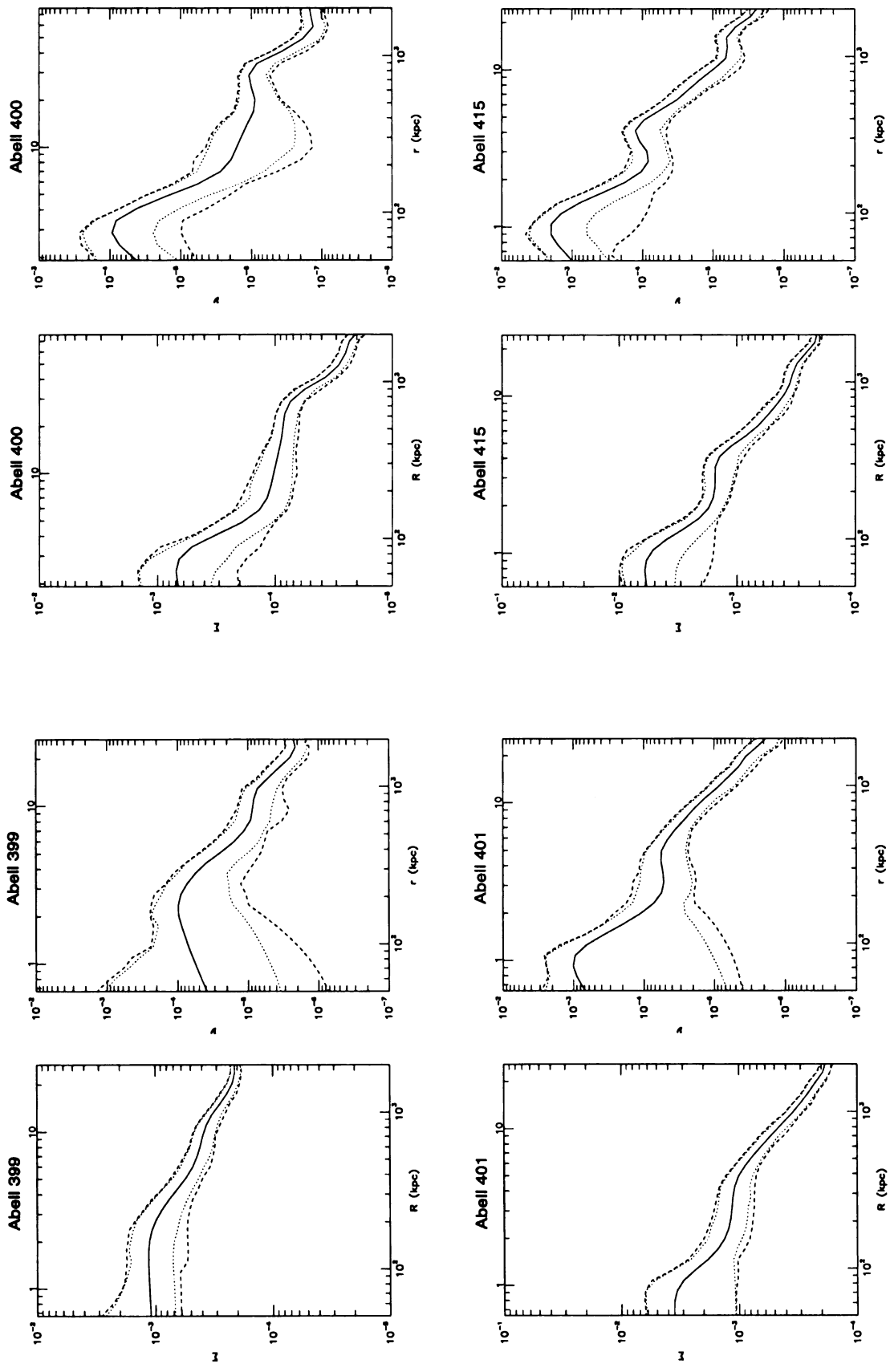

Figure 6.1. (cont'd)

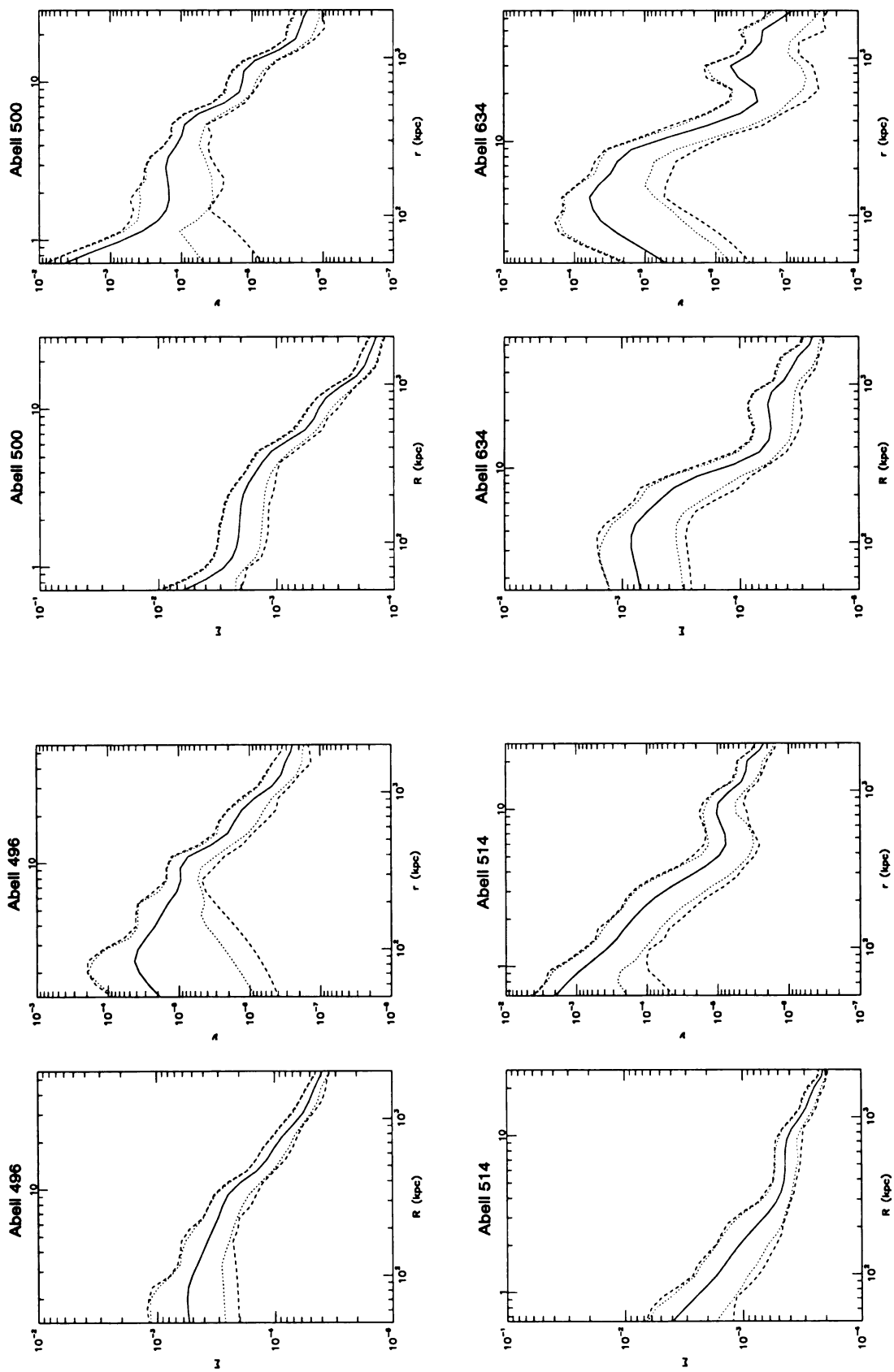

Figure 6.1. (cont'd)

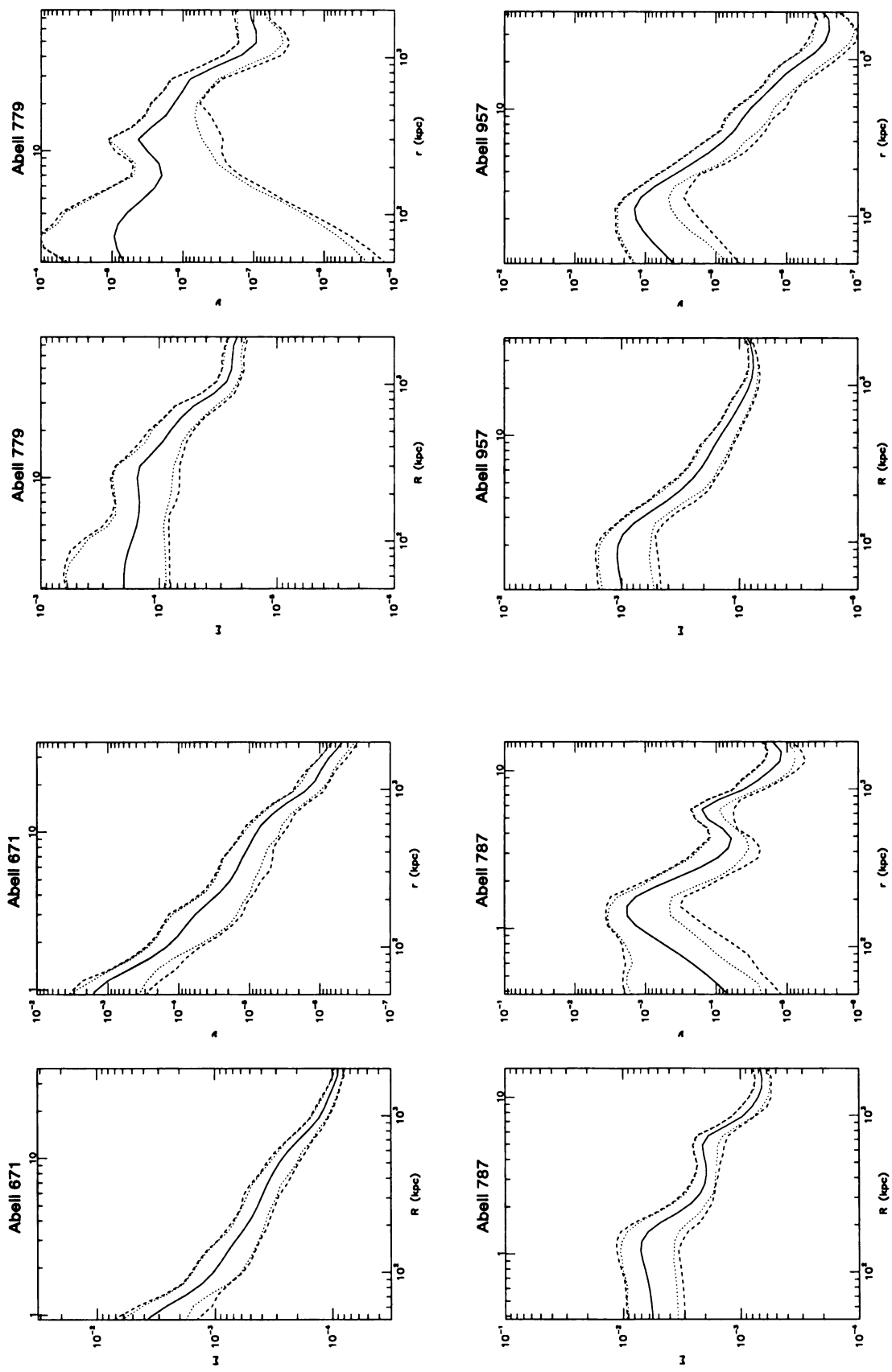

Figure 6.1. (cont'd)

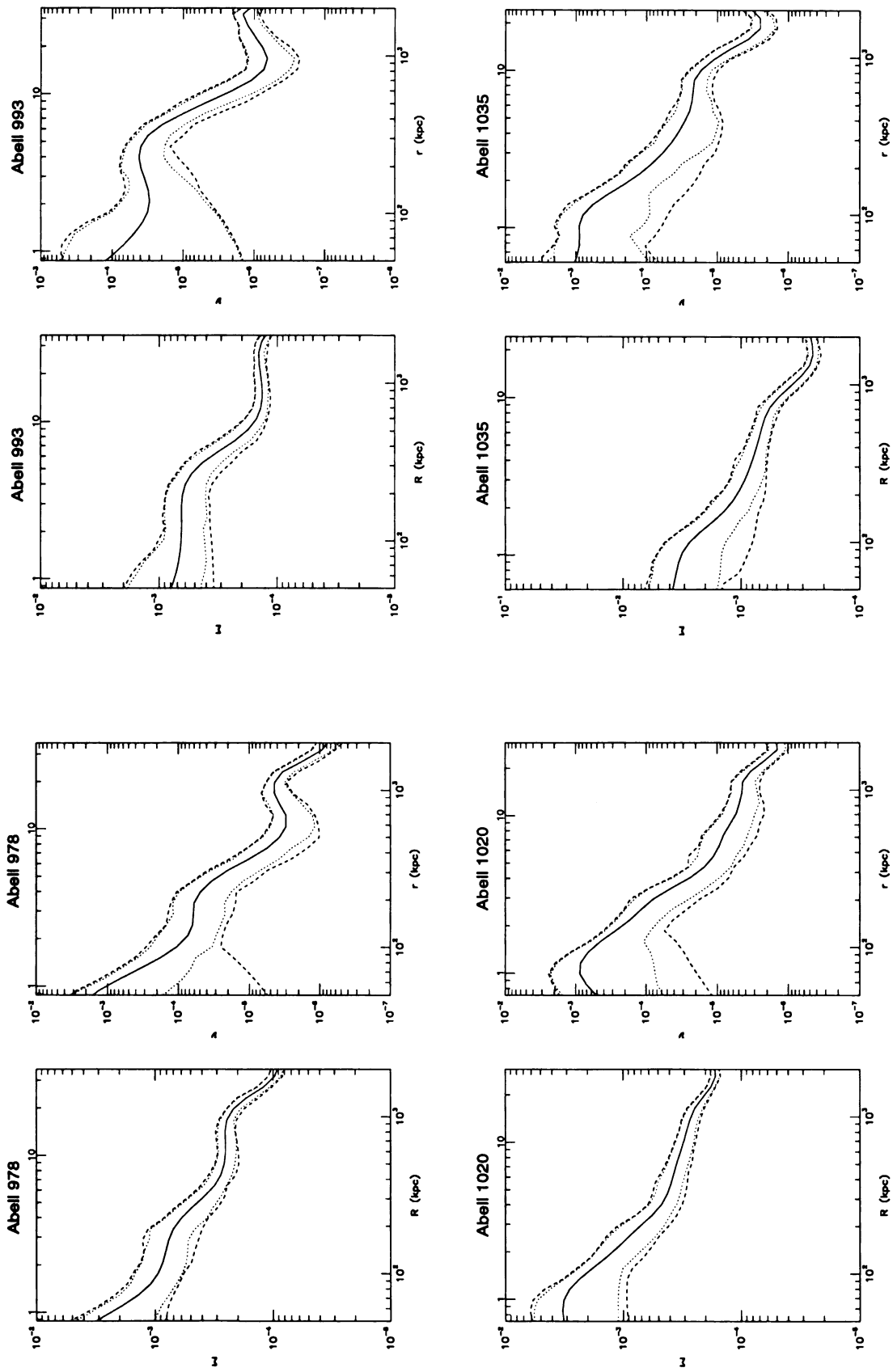

Figure 6.1. (cont'd)

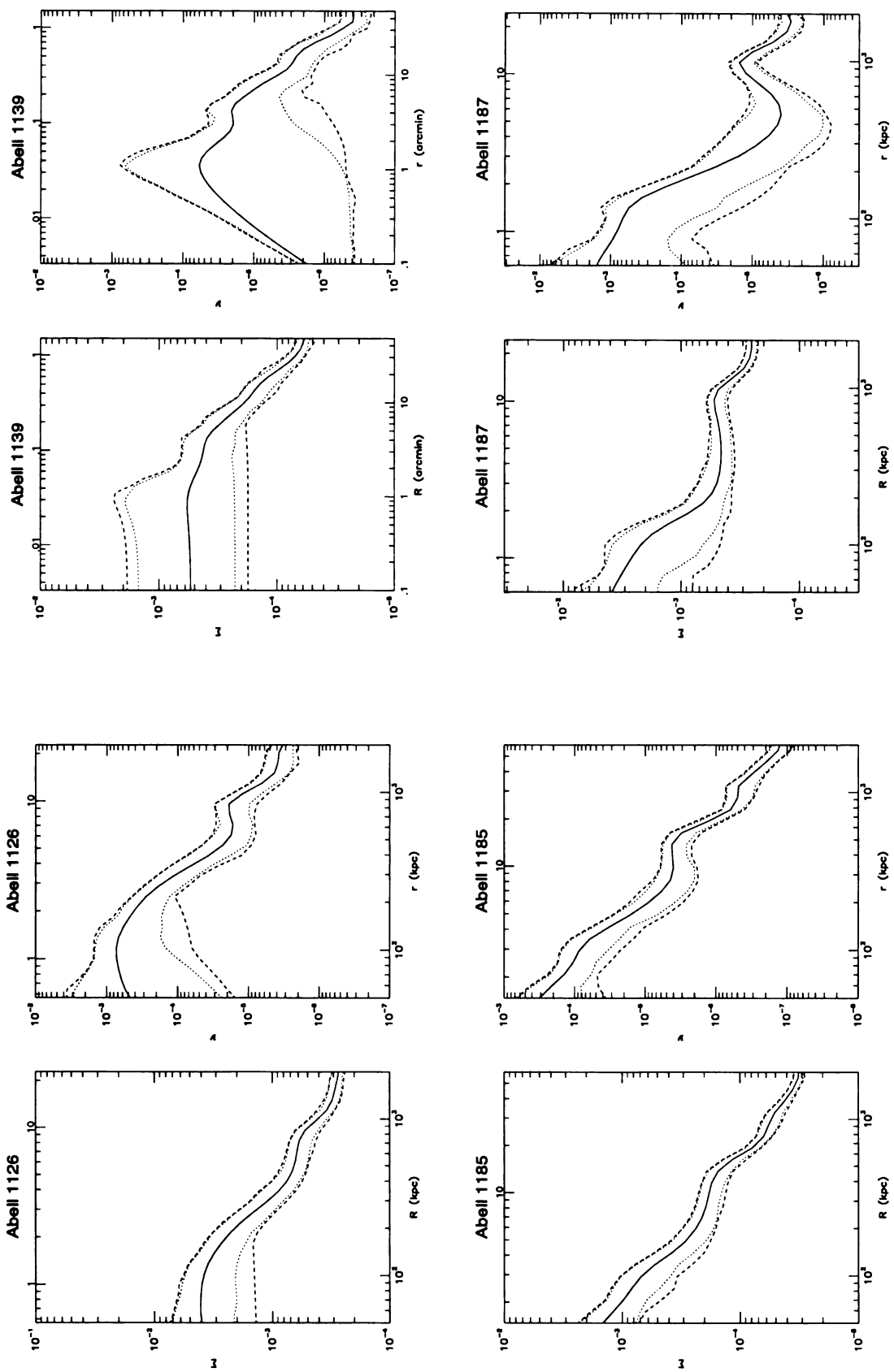

Figure 6.1. (cont'd)

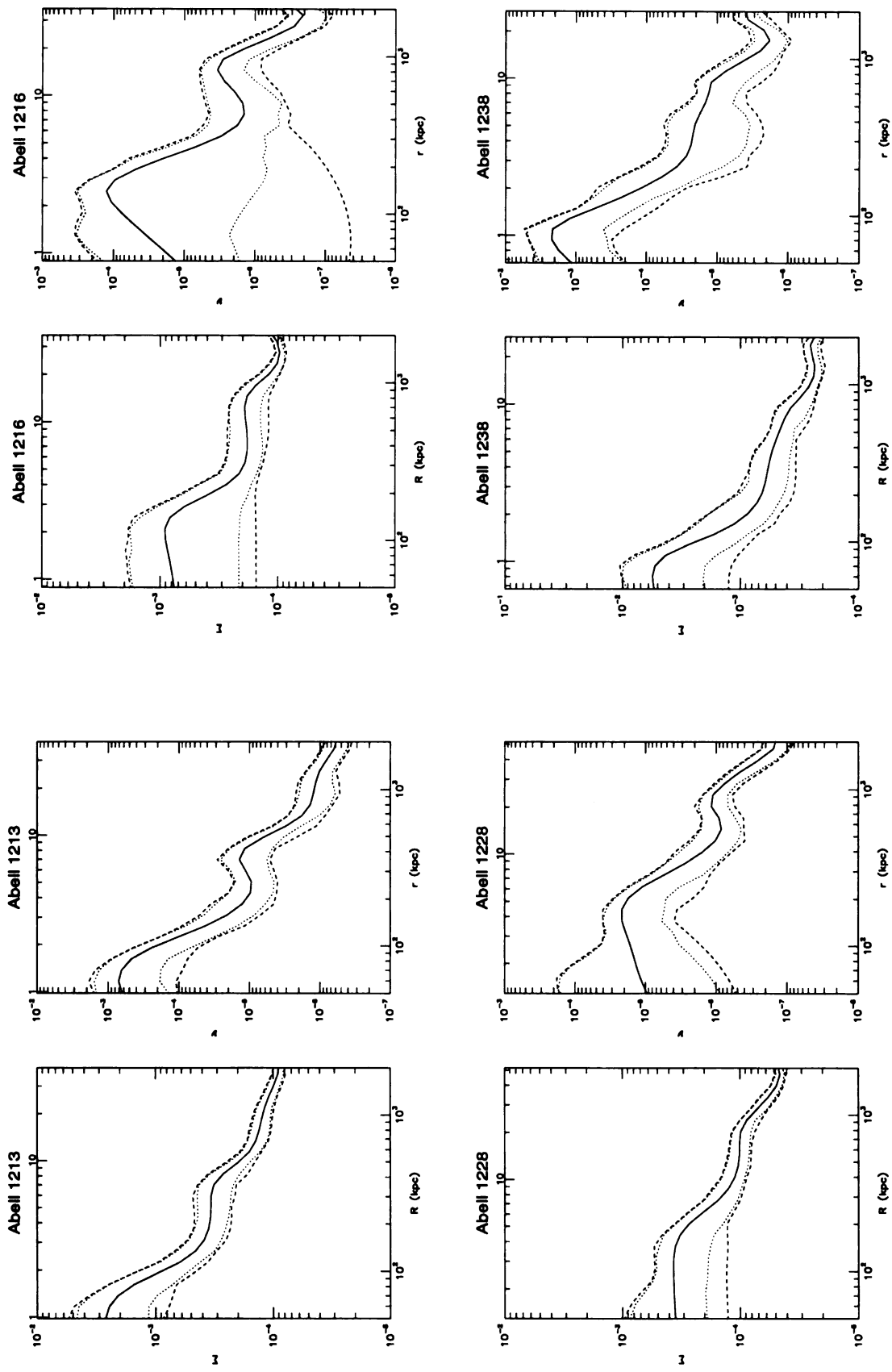

Figure 6.1. (cont'd)

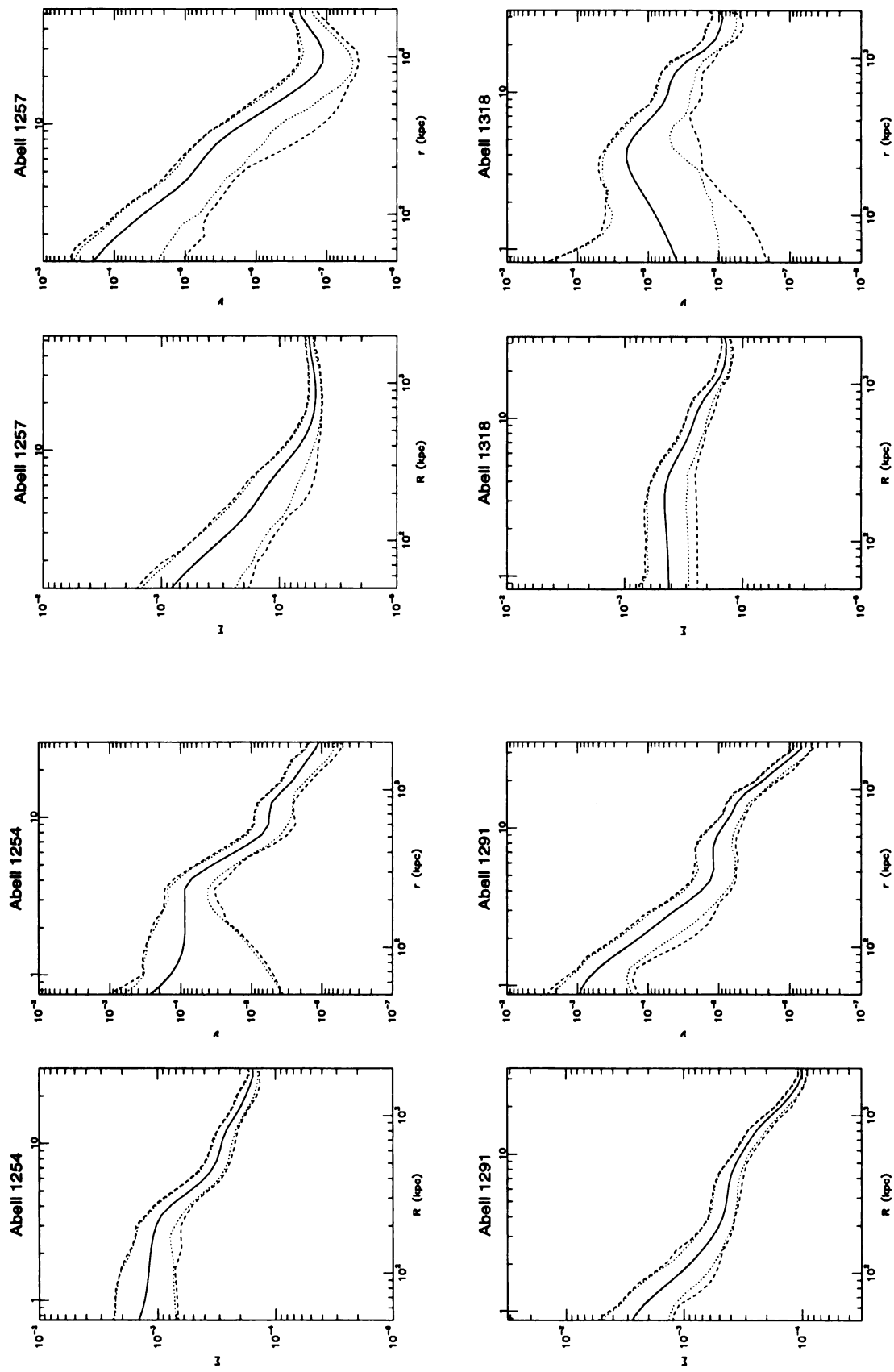

Figure 6.1. (cont'd)

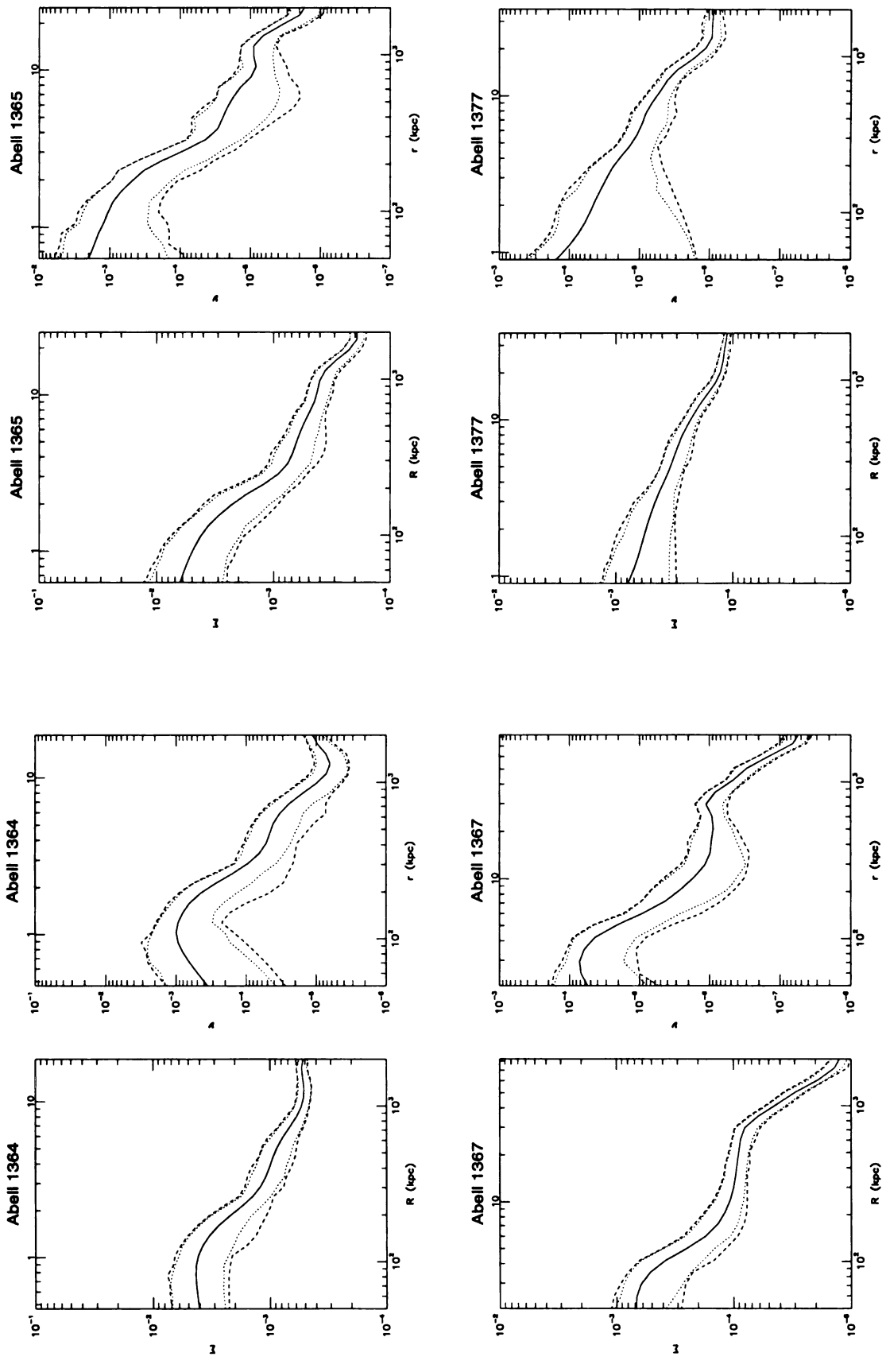

Figure 6.1. (cont'd)

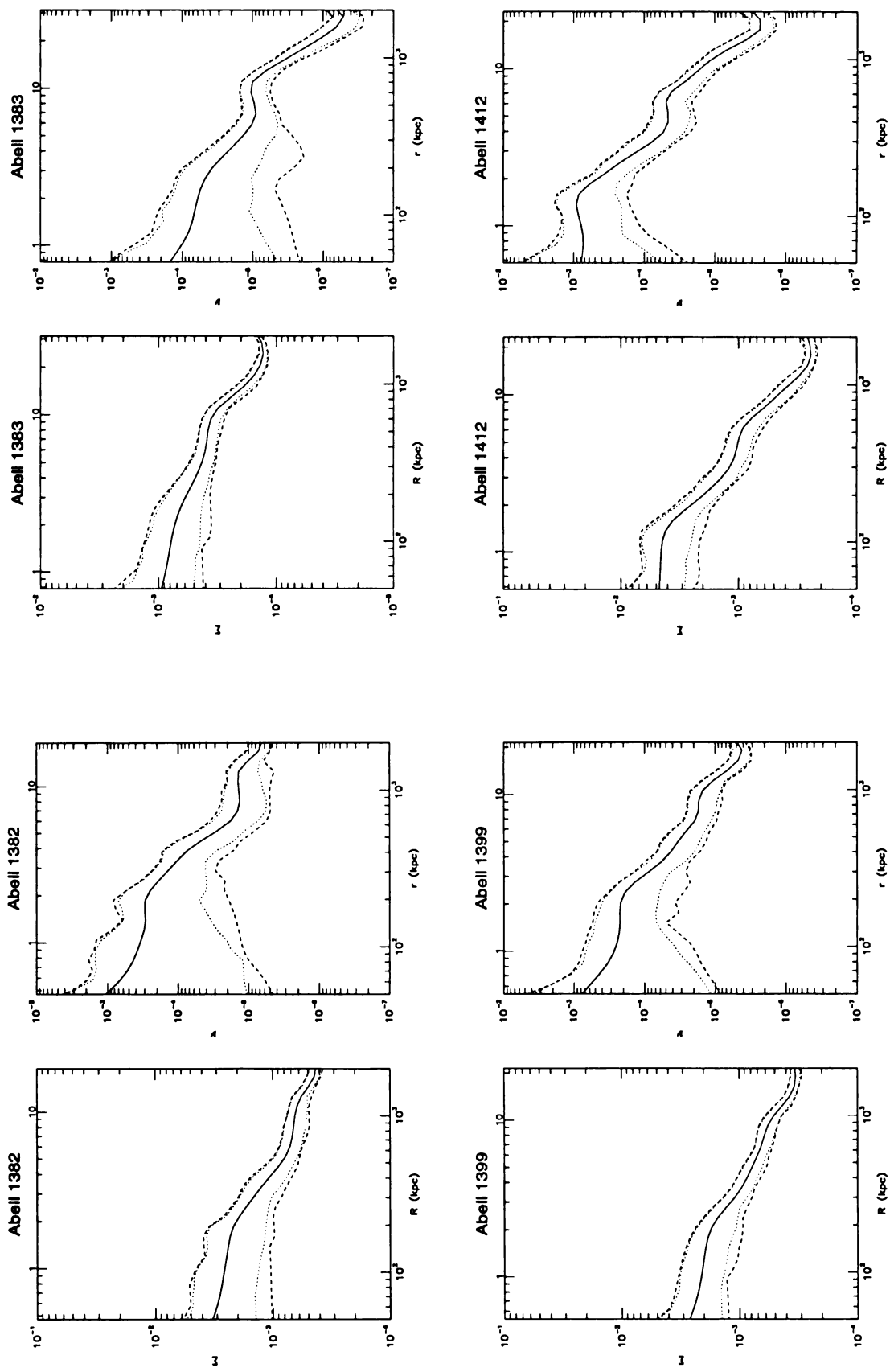

Figure 6.1. (cont'd)

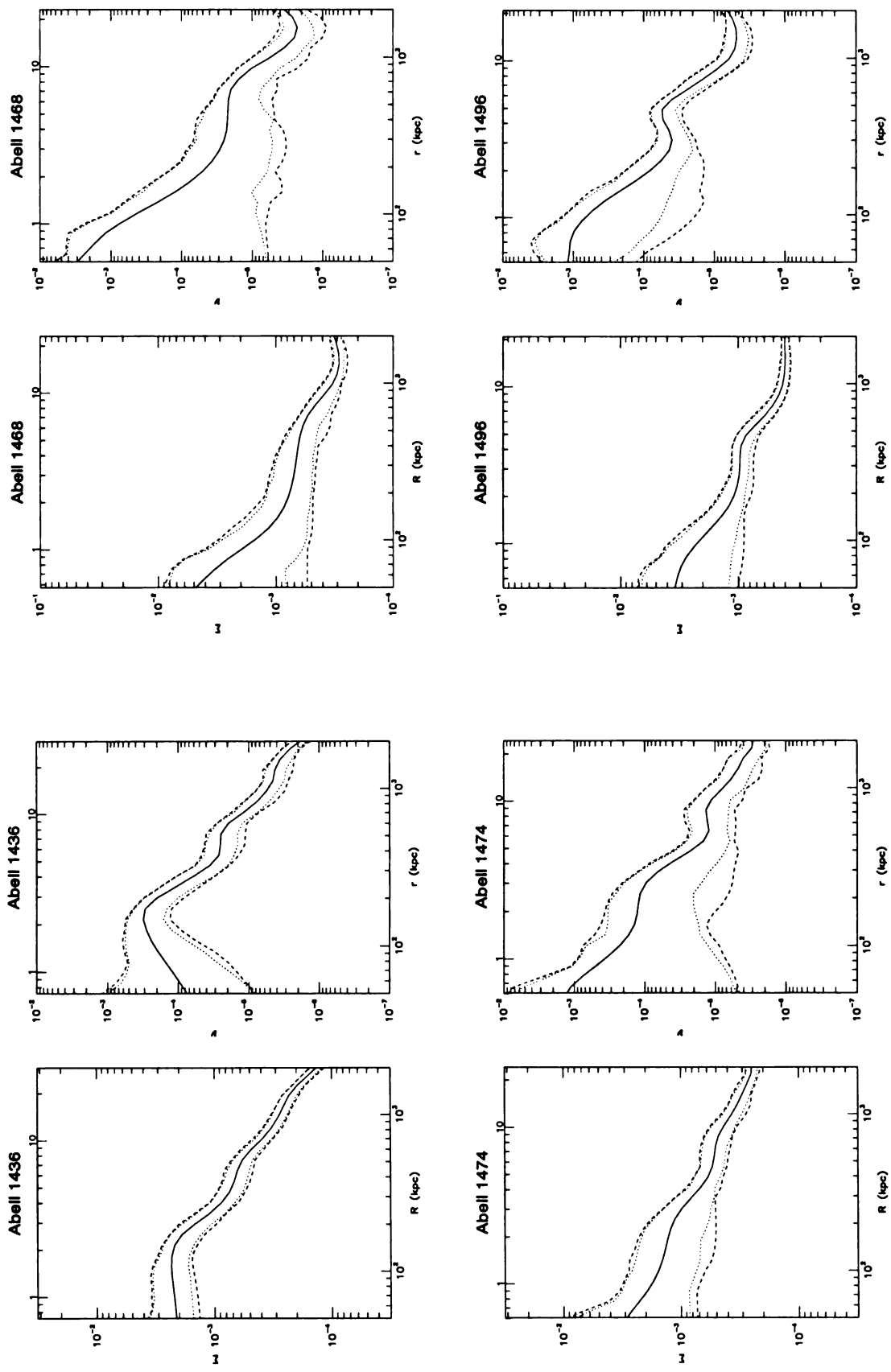

Figure 6.1. (cont'd)

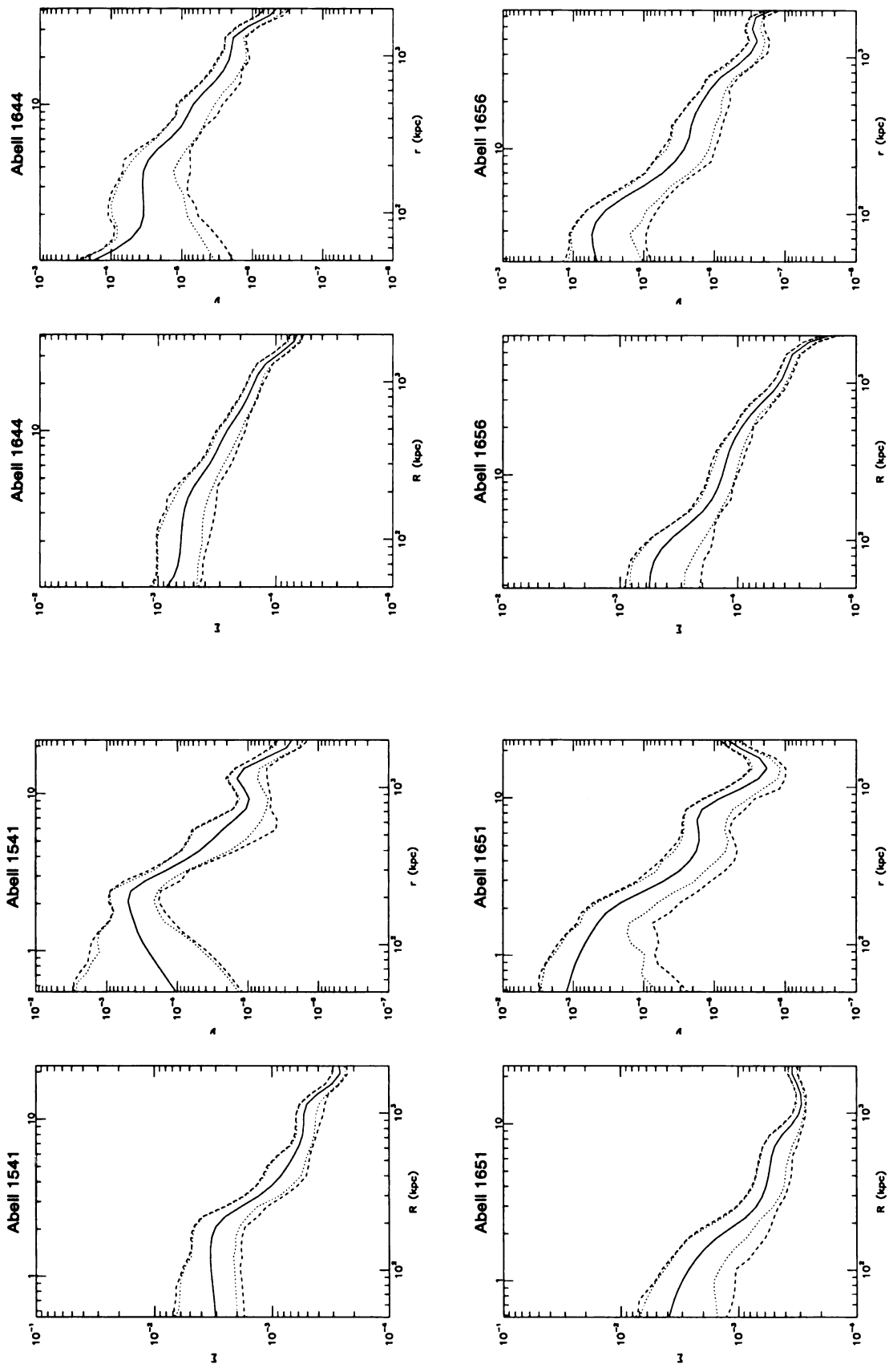

Figure 6.1. (cont'd)

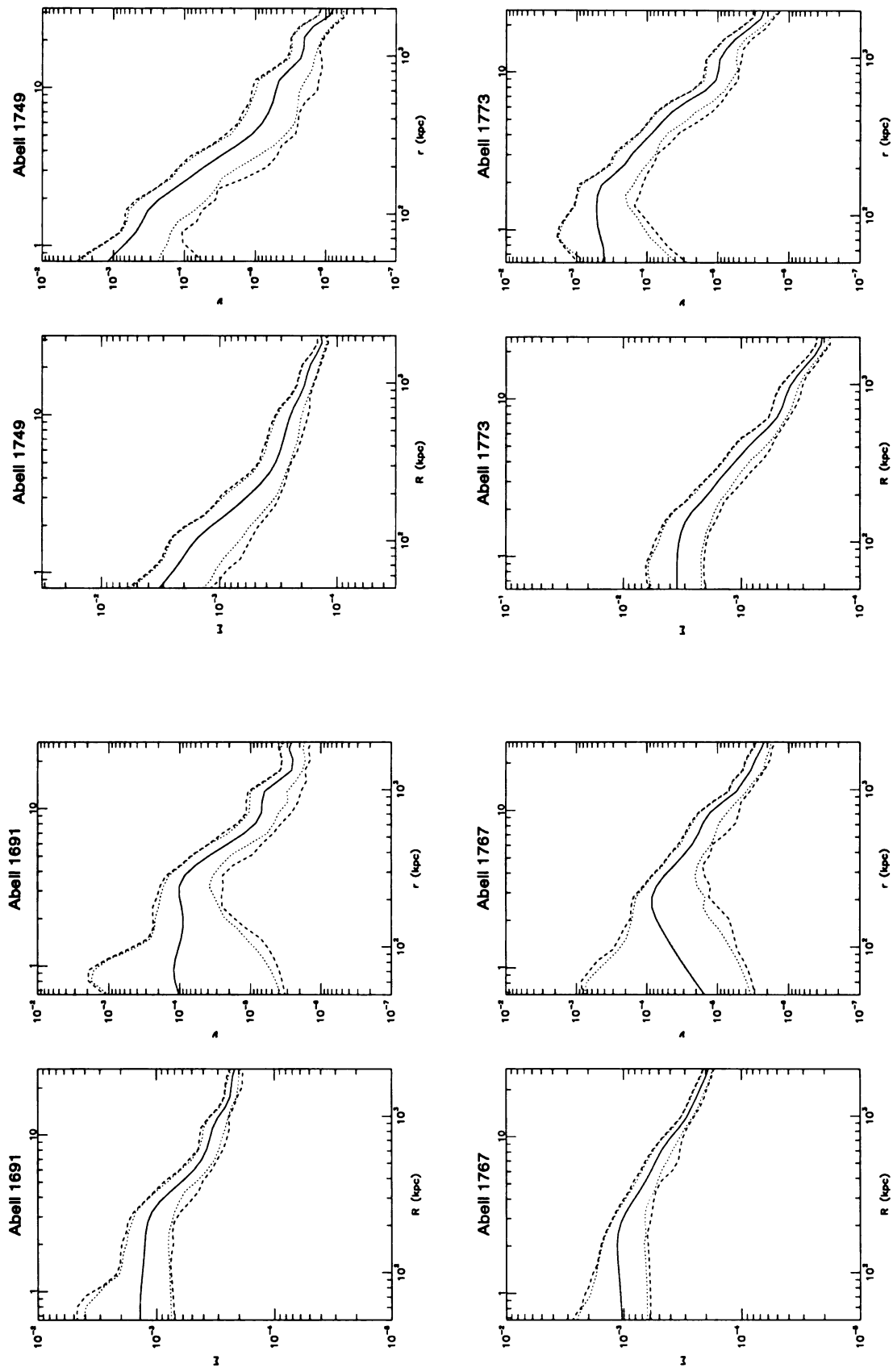

Figure 6.1. (cont'd)

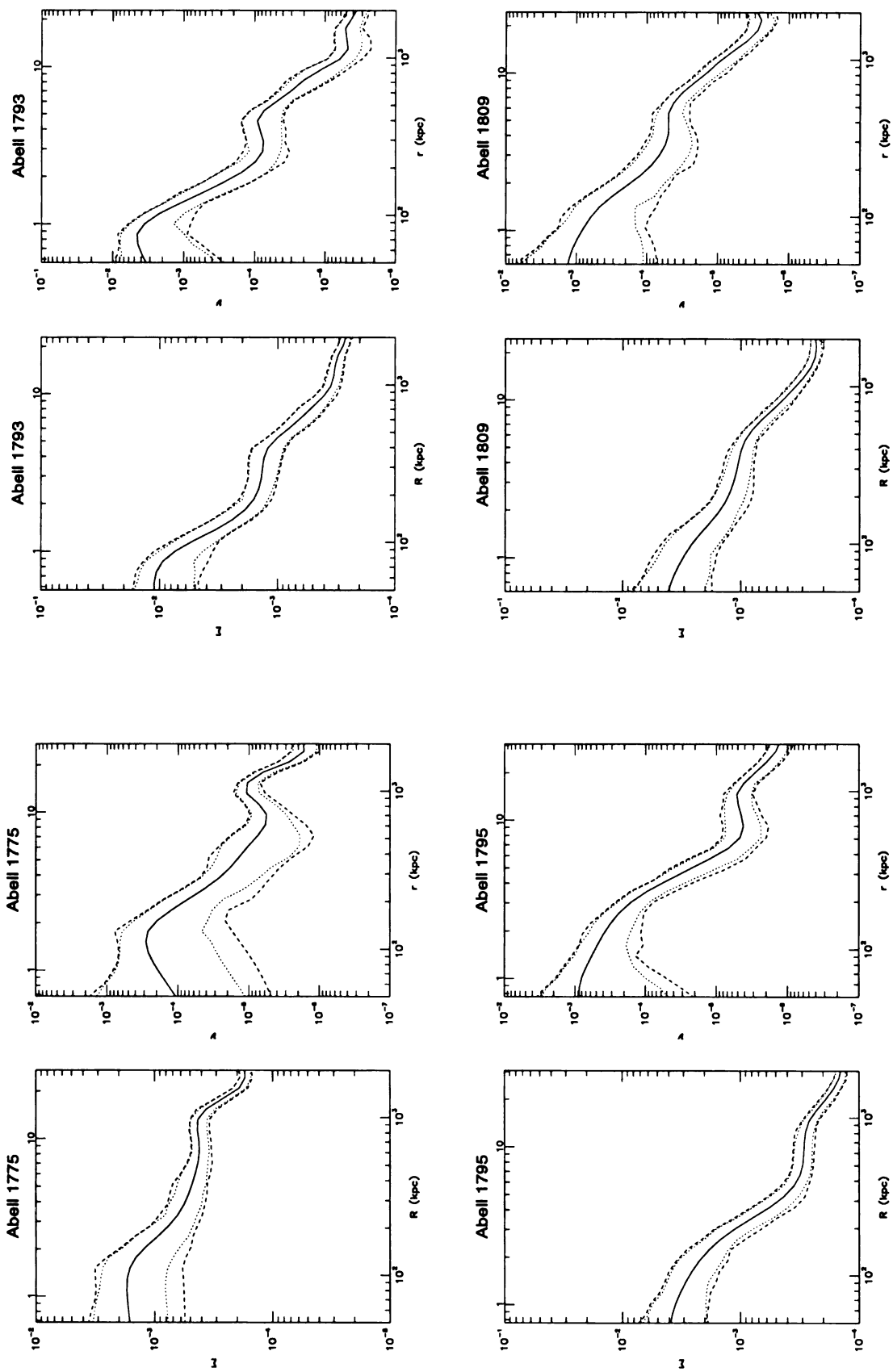

Figure 6.1. (cont'd)

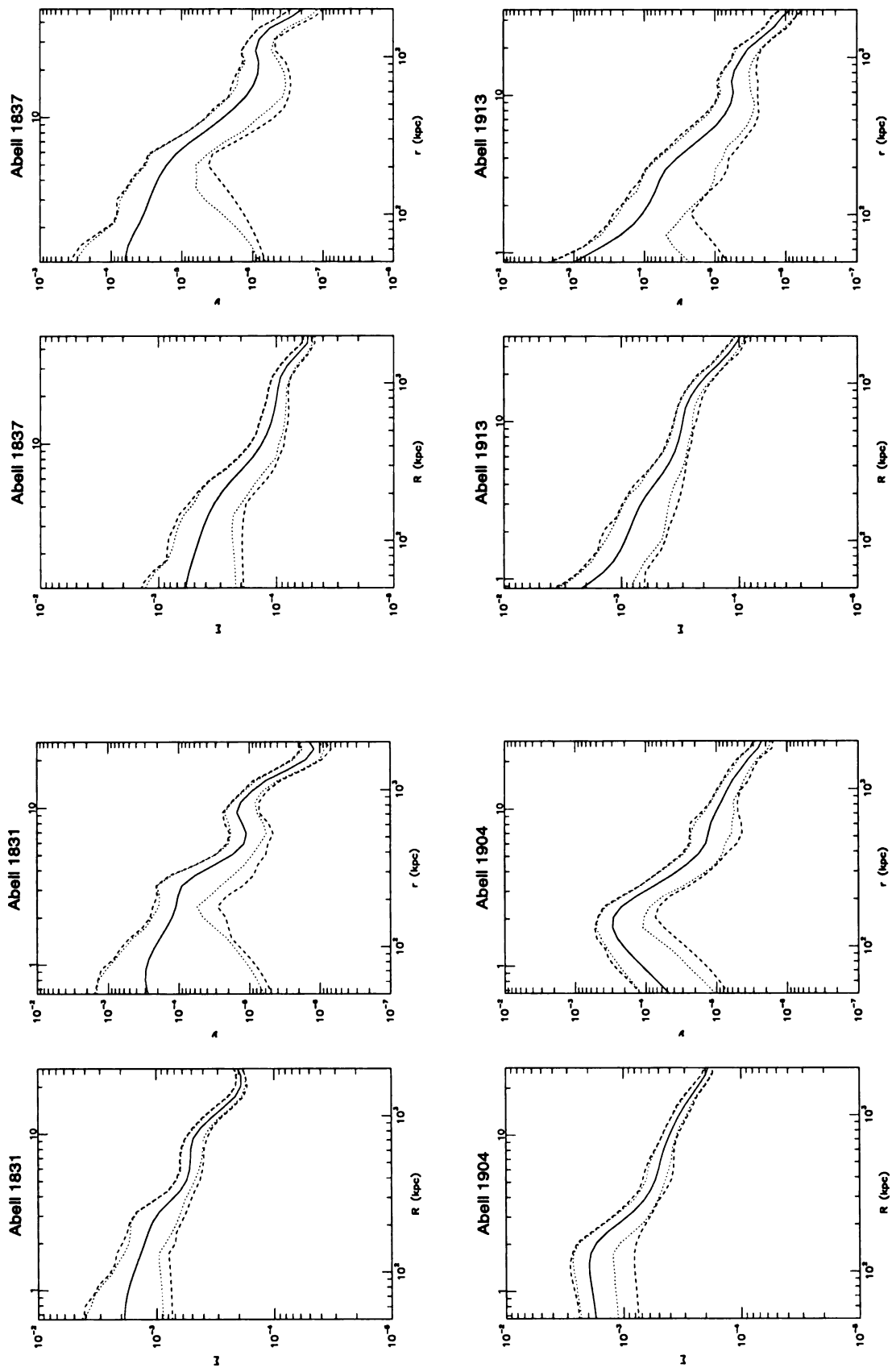

Figure 6.1. (cont'd)

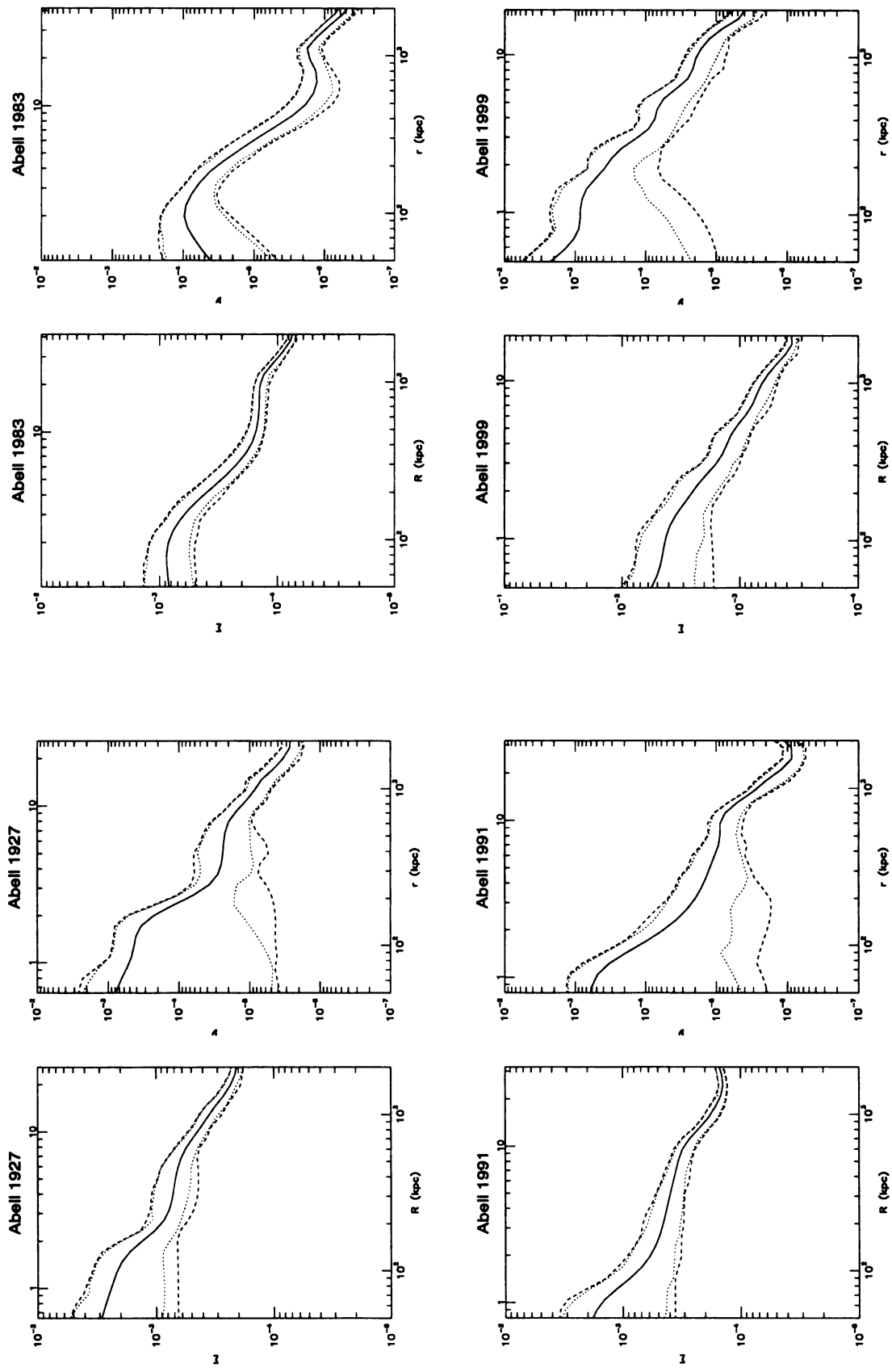

Figure 6.1. (cont'd)

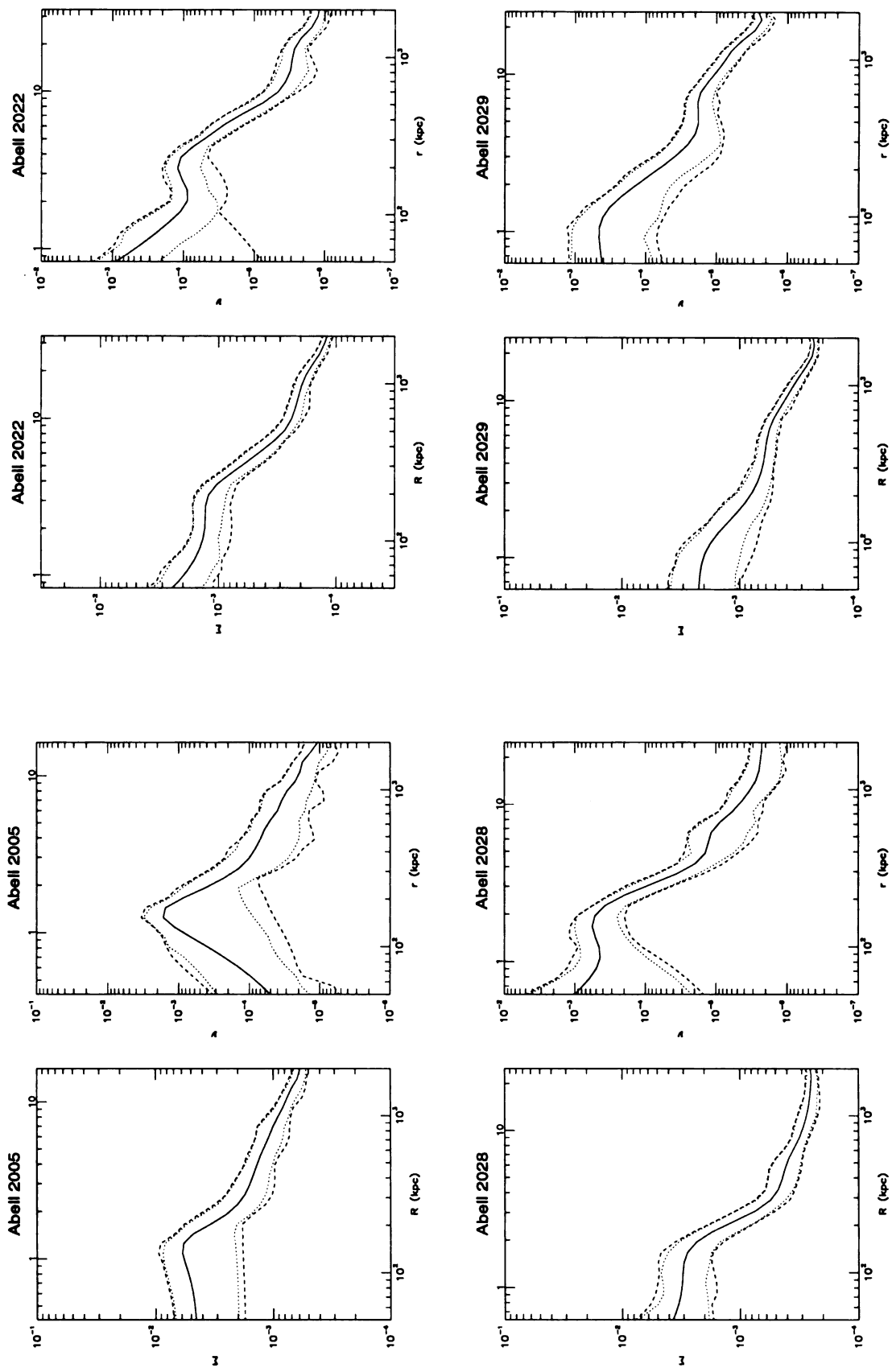

Figure 6.1. (cont'd)

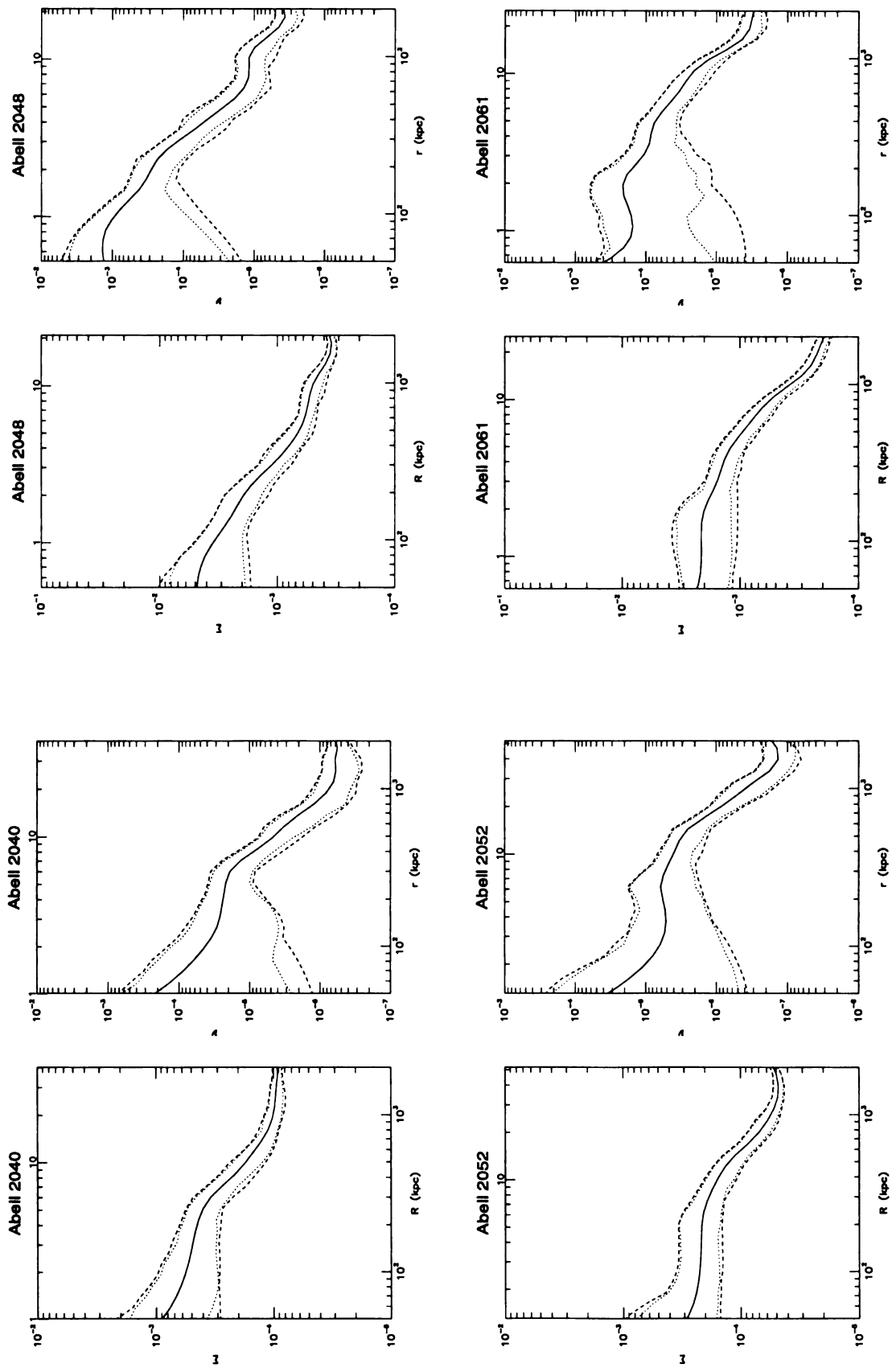

Figure 6.1. (cont'd)

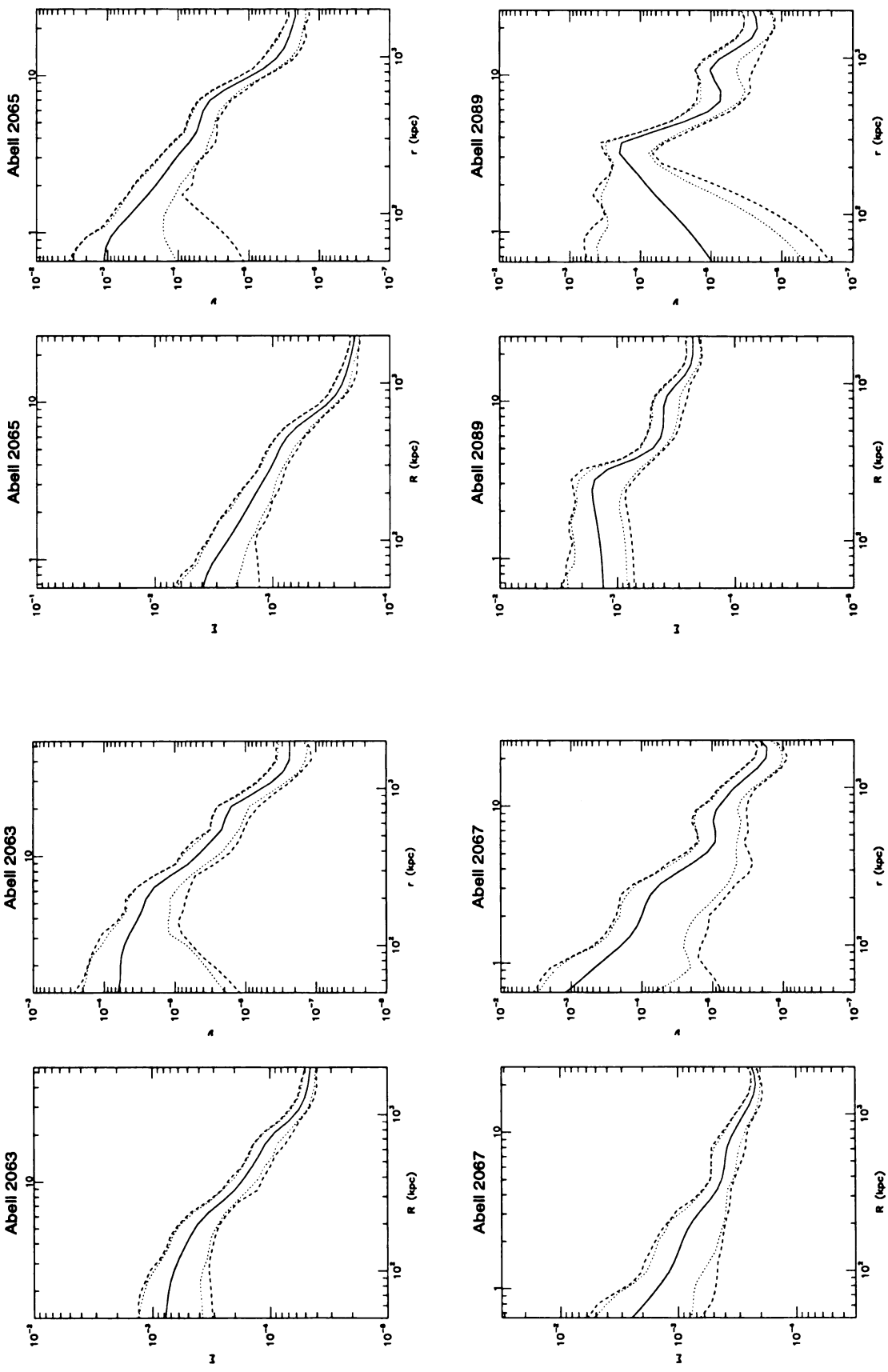

Figure 6.1. (cont'd)

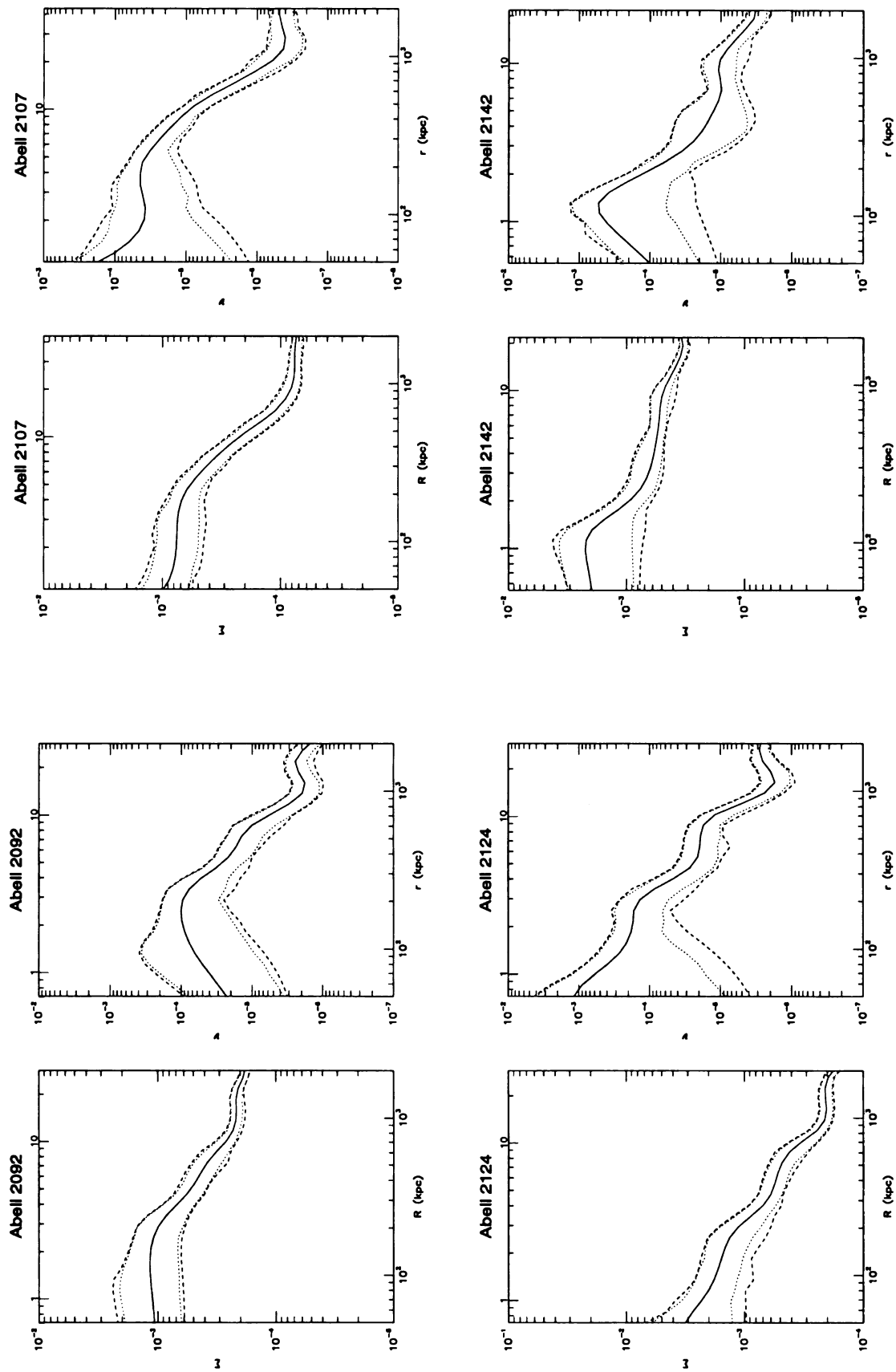

Figure 6.1. (cont'd)

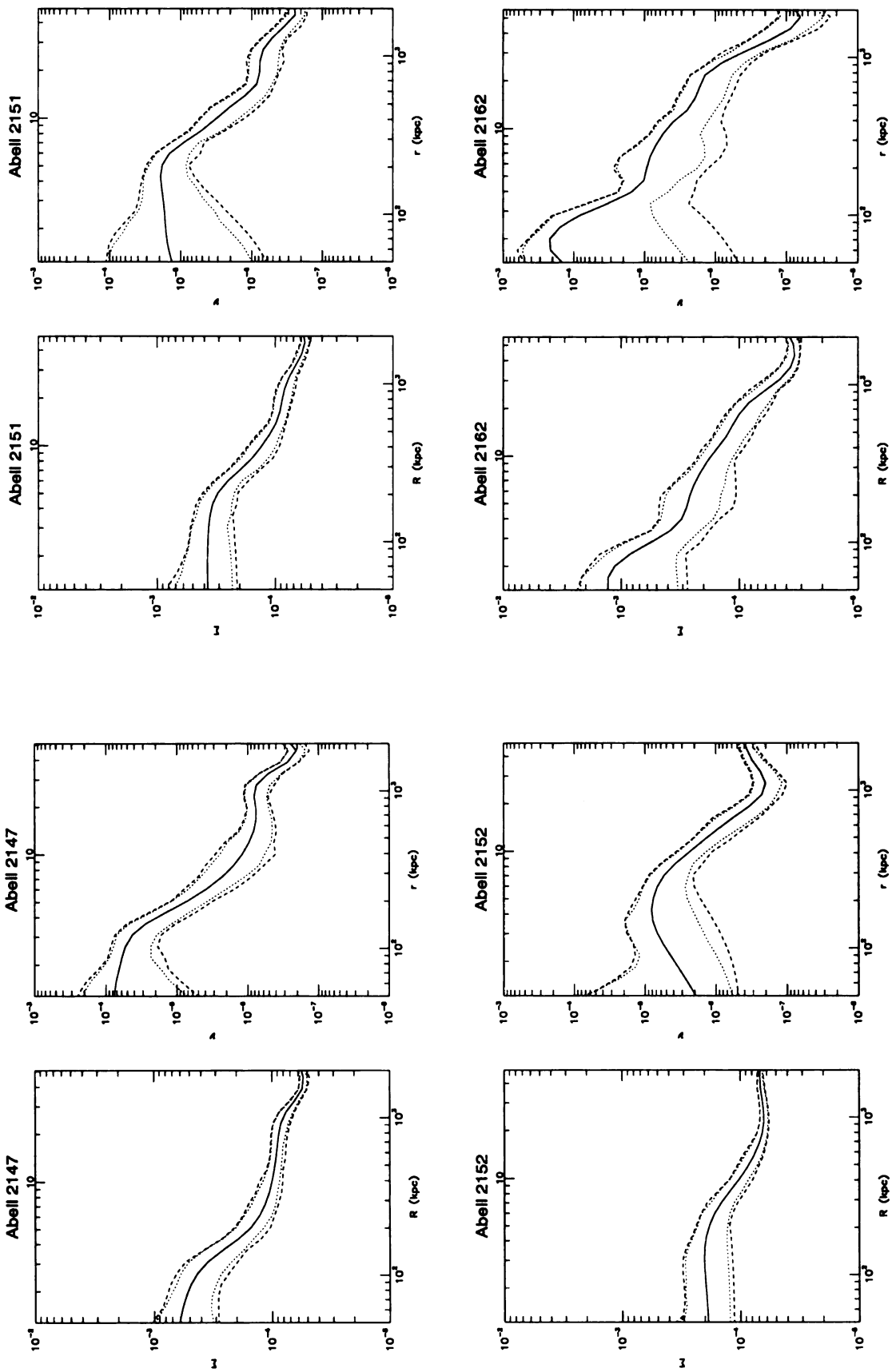

Figure 6.1. (cont'd)

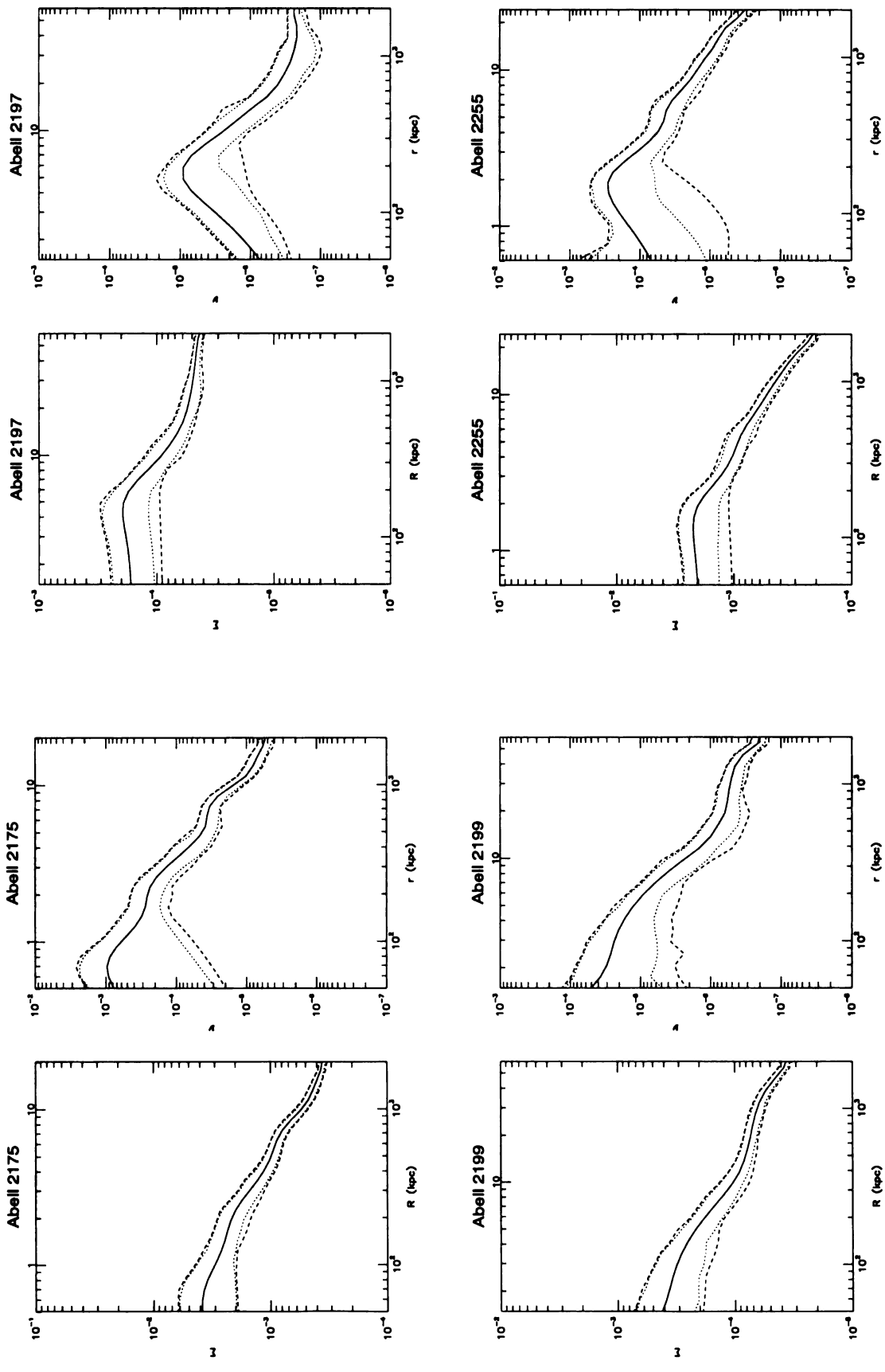

Figure 6.1. (cont'd)

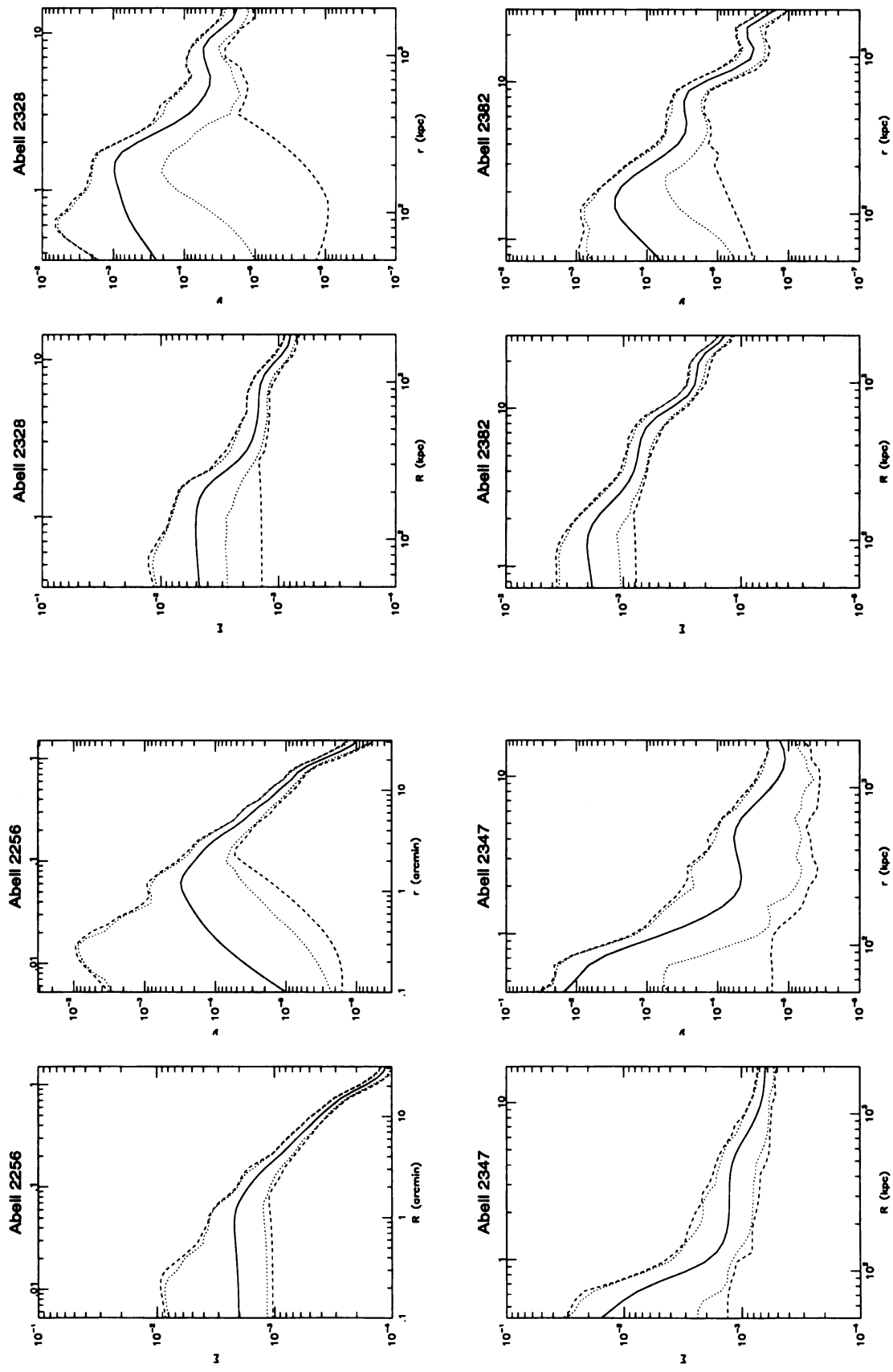

Figure 6.1. (cont'd)

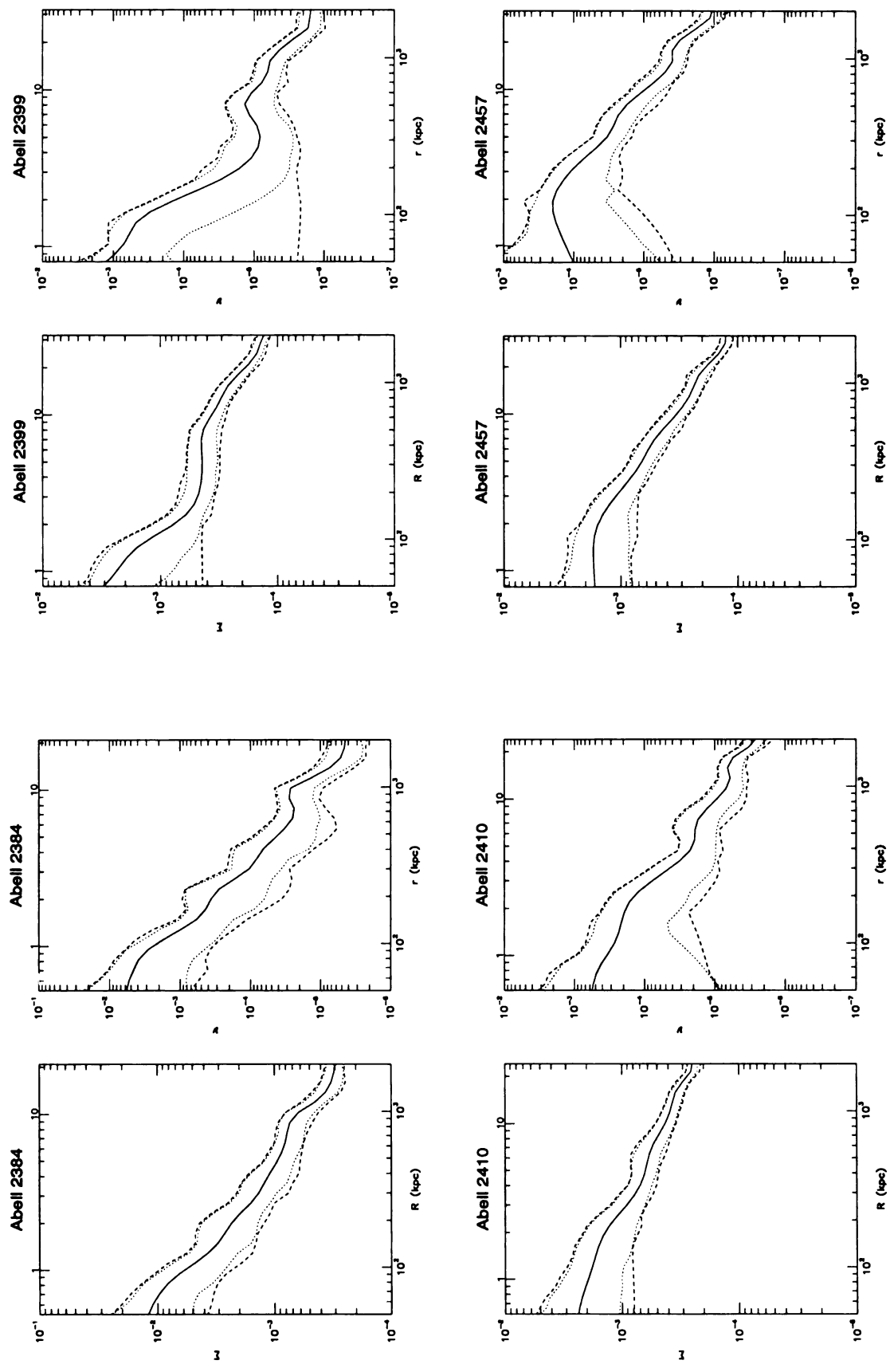

Figure 6.1. (cont'd)

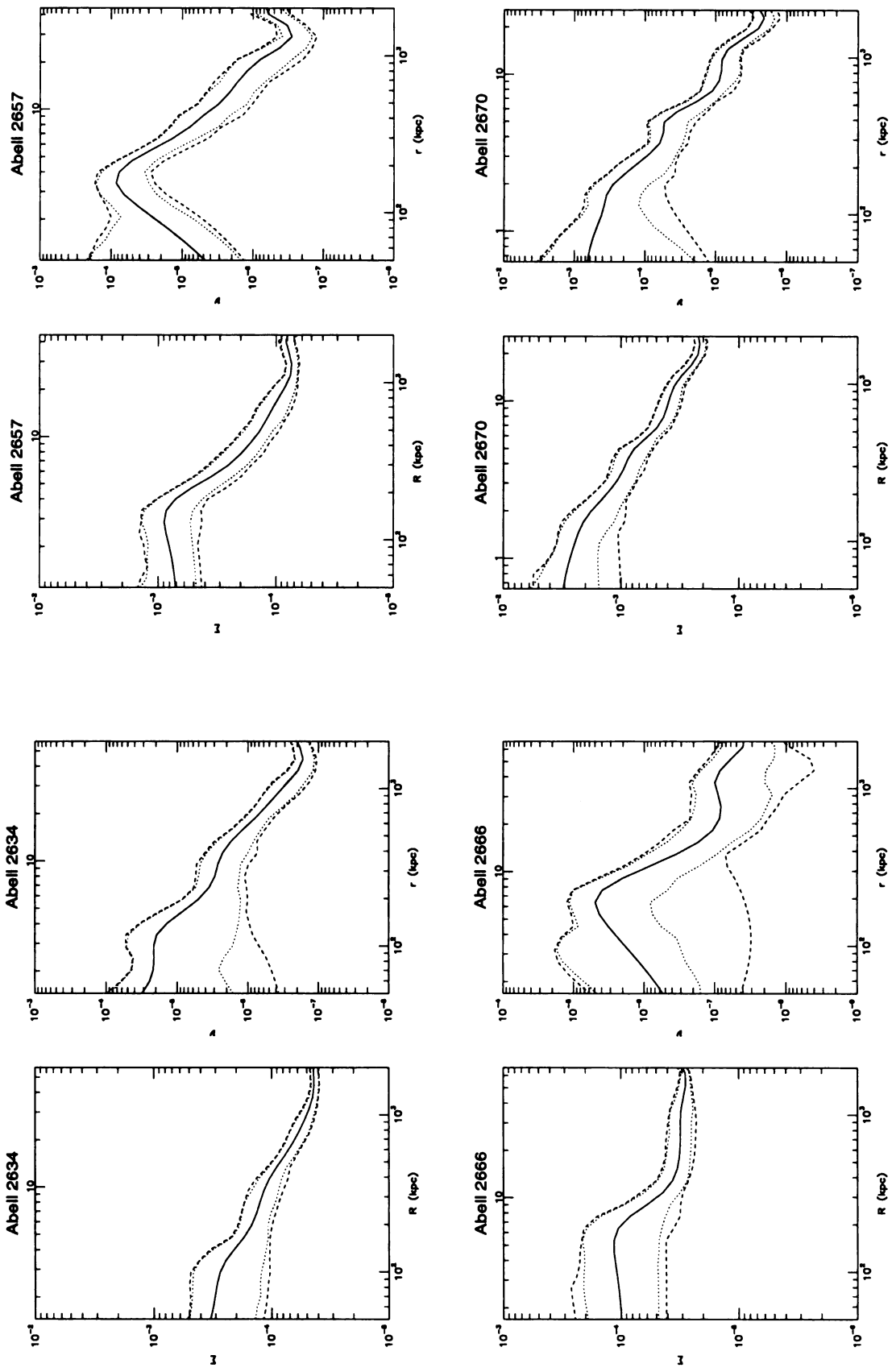

Figure 6.1. (cont'd)

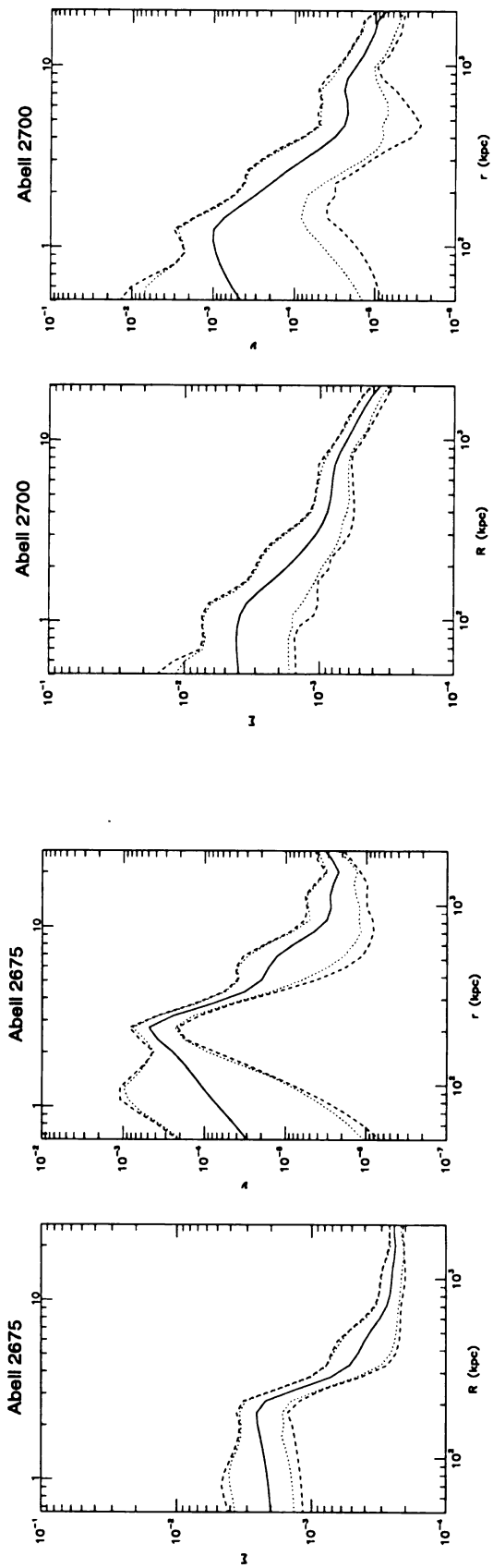

Figure 6.1. (cont'd)

6.3 Application to the Cluster Sample

The plots given in Figure 6.1 are constructed using the MPL method with the above penalty function (equation 6.8). Both the projected density and the space density are presented with ν and Σ plotted on a log-log scale with the bottom x axis labeled in kpc and the top labeled in arcminutes. The center of each cluster was chosen to be the highest peak in the adaptive-kernel map of the 2 Mpc region. The solid line is the estimate for either Σ or ν while the dotted and dashed lines are the 95% and 97.5%, respectively, bootstrap confidence intervals. Because of the desire to avoid spurious cores due to oversmoothing and the lack of a reliable, objective, data-based choice for the smoothing parameter, each of the the estimates was calculated with $\lambda = 1 \times 10^{-5}$. With this choice the space density profiles are deliberately *under*-smoothed with further smoother left to the viewer. As a result, the estimates for ν can be seen in many cases to "fall" near the centers of the clusters. As the smoothing parameter is increased, the density estimate in these cases will flatten out with a value close to that of the maximum. Further smoothing of course leads profile with a spurious, central cusp. Comparison with the numerical simulations of MT indicates that this type of behavior is a possible indication of a density that approaches a constant value in the core. In the case of a de Vaucouleur profile, which increases all the way to the center, the undersmoothed estimates generally oscillate about the true density.

6.3.1 Core Radius

It has been argued that a natural consequence of relaxation in a self-gravitating system is the development of a physical core in which the mass density approaches a constant value (King 1966). The core radius is usually defined as the radius at

which the projected density falls to one half the central value. Because the density estimate may decrease toward the center, due to the estimate being undersmoothed, the maximum value of Σ was used here instead of the value at the center. Unlike the plots shown in Figure 6.1, the profiles were constructed with a range of smoothing parameters, including those for which the density estimate was clearly undersmoothed and those which were merely power laws. In order to avoid creating spurious cores, the density profiles which decreased near the core and were closest to flat were used. For clusters that showed no sign of a flattened core in any of the calculated profiles, the smallest smoothing parameter, $\lambda = 1 \times 10^{-6}$, was used.

The results are listed in Table 6.1. The x and y positions in arcminutes of the center (chosen as the position of the maximum in the adaptive-kernel map) are listed in columns (1) and (2). Column (3) is the estimate of the core radius, with its 95% upper-bootstrap confidence interval listed in column (4). The median core radius for the clusters in the HGT sample is 150 ± 96 kpc, where the error is the one-sigma error. This result is approximately half that of the often quoted value of 250 h^{-1} kpc, or with $h = 0.75$, $r_{core} = 333$ kpc. On the other hand it is about three times the value obtained for some clusters with gravitational lens observations (Squires *et al* 1996).

TABLE 6.1. Core Radius of HGT clusters

Cluster	x (arcmin)	y (arcmin)	r_c (kpc)	95% error (kpc)	α	1σ error
(1)	(2)	(3)	(4)	(5)	(6)	(7)
A 21	−2.0	−0.5	130	254	−1.74	0.086
A 76	5.4	−5.4	289	632	−1.47	0.037
A 85	−6.3	−7.0	278	378	−1.67	0.182
A 88	−0.5	−1.7	121	130	−2.67	0.226
A 104	1.2	−0.4	168	223	−1.68	0.193
A 119	1.2	−0.4	92	144	−1.76	0.135
A 121	−2.8	−0.2	155	163	−1.94	0.104
A 147	−2.7	1.2	37	209	−1.63	0.097
A 151	1.0	0.3	149	170	−1.89	0.047
A 154	−0.8	−0.1	132	155	−2.01	0.048
A 166	−0.1	−0.1	37	218	−1.68	0.212
A 168	−1.1	1.9	201	237	−2.00	0.105
A 189	−29.3	7.7	170	192	−2.05	0.101
A 193	0.4	−1.1	38	272	−1.80	0.247
A 194	−1.0	4.8	94	332	−2.06	0.095
A 225	1.7	−3.7	70	177	−1.72	0.128
A 246	2.7	−5.2	317	718	−2.37	0.076
A 274	−3.9	−0.4	354	704	−1.12	0.106
A 277	−0.5	−1.3	82	91	−1.88	0.144
A 389	0.4	−0.4	295	311	−1.82	0.082
A 399	0.7	−0.7	252	501	−1.54	0.062
A 400	0.7	−2.2	91	100	−1.82	0.147
A 401	−0.2	1.6	85	584	−1.64	0.098
A 415	−2.8	4.6	136	146	−1.77	0.037
A 496	2.7	−7.0	119	374	−1.74	0.062
A 500	3.4	−0.3	238	282	−1.58	0.190
A 514	−6.8	−0.2	66	98	−2.06	0.165
A 634	−3.2	16.1	170	205	−2.52	0.123
A 671	−0.3	−1.0	62	70	−2.10	0.128
A 779	−6.8	−9.9	485	627	−0.83	0.186
A 787	4.7	1.1	137	291	−1.72	0.046
A 957	5.1	0.4	157	181	−2.44	0.058
A 978	−0.3	1.6	55	70	−2.16	0.215
A 993	−0.3	1.6	357	382	−3.11	0.120
A 1020	−0.3	10.3	112	168	−2.04	0.098
A 1035	−1.1	3.7	108	146	−1.63	0.124
A 1126	−0.2	−1.0	184	244	−1.95	0.089
A 1139	−9.4	6.7	172	515	−1.93	0.172
A 1185	0.6	1.7	78	89	−2.20	0.082
A 1187	−5.4	5.0	114	155	−1.96	0.357
A 1213	−0.4	−1.1	80	91	−2.12	0.092
A 1216	8.2	−3.6	31	175	−1.41	0.220
A 1228	−17.2	−2.5	286	334	−1.60	0.066
A 1238	1.2	−0.7	94	102	−2.10	0.088

TABLE 6.1. (continued)

Cluster	x (arcmin)	y (arcmin)	r_c (kpc)	95% error (kpc)	α	1σ error
(1)	(2)	(3)	(4)	(5)	(6)	(7)
A 1254	−1.4	1.9	41	335	−1.52	0.196
A 1257	−2.5	−0.5	42	83	−2.55	0.085
A 1291	−0.3	1.6	80	95	−1.84	0.147
A 1318	2.1	−2.1	678	1010	−1.62	0.060
A 1364	2.1	2.7	195	220	−2.16	0.074
A 1365	−0.2	−0.7	50	116	−2.11	0.145
A 1367	−2.4	−4.0	95	127	−1.61	0.103
A 1377	−2.3	0.3	97	654	−1.53	0.062
A 1382	−2.5	−1.8	324	446	−2.14	0.088
A 1383	0.3	1.4	34	700	−1.47	0.226
A 1399	0.9	2.8	360	394	−1.63	0.046
A 1412	−2.2	−3.5	176	200	−1.50	0.174
A 1436	4.0	3.9	221	273	−1.30	0.159
A 1468	2.2	3.9	64	551	−2.07	0.210
A 1474	−1.1	4.6	346	470	−1.68	0.032
A 1496	1.2	−0.5	66	107	−1.99	0.248
A 1541	−2.5	−1.7	187	279	−1.50	0.165
A 1644	−1.9	8.0	161	336	−1.65	0.076
A 1651	−1.5	1.0	104	164	−1.78	0.193
A 1656	2.2	−2.2	113	146	−1.78	0.081
A 1691	3.6	0.2	320	432	−1.53	0.152
A 1749	2.0	−0.3	44	81	−2.09	0.118
A 1767	−0.7	−4.7	351	722	−0.87	0.179
A 1773	−2.0	−0.7	35	243	−1.68	0.206
A 1775	0.7	0.2	186	275	−1.66	0.126
A 1793	1.0	1.0	44	115	−2.12	0.164
A 1795	2.0	−2.0	127	178	−2.09	0.146
A 1809	5.9	−3.3	67	112	−1.73	0.082
A 1831	−0.2	−1.2	213	381	−1.35	0.087
A 1837	6.9	−1.4	197	396	−1.63	0.115
A 1904	−0.2	−4.6	217	227	−1.64	0.057
A 1913	4.8	6.1	51	161	−1.57	0.180
A 1927	−1.2	−2.6	123	623	−1.47	0.137
A 1983	−4.3	−3.5	169	213	−1.83	0.112
A 1991	−0.9	−0.9	71	527	−1.57	0.108
A 1999	−5.8	4.8	101	275	−1.88	0.046
A 2005	2.3	1.2	262	419	−2.10	0.051
A 2022	0.9	4.0	63	95	−2.11	0.181
A 2028	−0.2	0.7	166	191	−1.83	0.242
A 2029	0.7	−0.2	50	321	−1.51	0.151
A 2040	−0.4	1.1	34	353	−1.74	0.172
A 2048	−0.2	−0.5	132	266	−1.88	0.078
A 2052	0.5	0.5	30	584	−1.22	0.312
A 2061	−0.7	−0.7	365	519	−1.55	0.072
A 2063	0.5	−3.6	194	267	−1.54	0.267

TABLE 6.1. (continued)

Cluster	x (arcmin)	y (arcmin)	r_c (kpc)	95% error (kpc)	α	1σ error
(1)	(2)	(3)	(4)	(5)	(6)	(7)
A 2065	3.1	−0.7	365	582	−1.55	0.072
A 2067	−0.7	4.5	90	200	−1.84	0.069
A 2089	1.6	0.7	323	350	−2.12	0.068
A 2092	0.3	0.8	314	425	−2.08	0.090
A 2107	2.0	1.2	35	313	−1.72	0.296
A 2124	0.8	1.8	102	232	−1.86	0.103
A 2142	−4.8	−5.9	204	737	−1.58	0.072
A 2147	1.4	1.4	95	153	−1.81	0.211
A 2151	−0.5	0.5	241	314	−1.67	0.113
A 2152	9.7	9.7	144	589	−1.49	0.060
A 2162	3.8	0.5	65	138	−2.19	0.170
A 2175	−0.9	−1.2	354	400	−1.68	0.061
A 2197	−21.0	−13.1	321	402	−1.70	0.099
A 2199	−1.7	4.0	81	218	−1.95	0.102
A 2255	0.6	−1.9	287	495	−1.72	0.043
A 2256	−0.3	−4.3	213	320	−1.33	0.164
A 2328	0.1	−0.6	284	379	−1.80	0.065
A 2347	−1.9	−2.7	61	119	−2.24	0.379
A 2382	−1.6	0.2	273	405	−1.62	0.074
A 2384	−1.6	0.2	90	177	−1.93	0.155
A 2399	1.5	−0.3	50	64	−1.91	0.292
A 2410	1.9	1.1	186	374	−1.50	0.089
A 2457	2.6	0.9	37	389	−1.55	0.140
A 2634	−5.0	1.7	160	450	−1.93	0.098
A 2657	2.1	1.2	36	237	−1.65	0.254
A 2666	−5.8	1.9	259	781	−1.85	0.123
A 2670	−1.2	−0.7	161	230	−1.44	0.141
A 2675	−2.6	−5.9	57	244	−2.07	0.166
A 2700	0.9	−1.9	166	225	−1.83	0.104

Examination of the data in Table 6.1 indicates the presence of two outliers: A1318 with $r_c = 678$ kpc, and A779 at $r_c = 485$ kpc. In both of these cases, the reason for such a large value is poor center specification. If, instead of using the maximum in the adaptive-kernel map, the Abell center is used, these values drop to $r_c = 36$ kpc, and $r_c = 57$ kpc, respectively. This variation in core radius with center specification is a well-known problem (Beers & Tonry 1986). In an attempt to minimize this variation, it has become customary to seek a center which can be specified independently of the galaxy distribution, such as the position of the X-ray peak or that of the cD galaxy. However, these alternative center specifications do not solve the problem, but merely provide other centers to try. For instance, in this study, if the position of the cD galaxies are used as the center, approximately half of the core radii for these clusters can be reduced, but several become significantly larger. The explanation for this sensitivity appears to be the presence of small-scale structure in the cores of clusters.

As previously discussed (see Chapter 3) many of the clusters in this survey show evidence of small-scale structure in their cores. In the case of A1656, this structure is believed to be real because of the matching peaks in the the X-ray surface brightness. A1656 has $r_c = 113$ kpc listed in Table 6.1, with the center at $(2'.2, -2'.2)$, obtained from the maximum density of the 2 Mpc region. Comparison with Figure 3.11 reveals that this center misses the density peak in the smaller map by about two arcminutes. Choosing the position of this peak as the center, the core radius is less than 50 kpc. A similar situation applies to A400. It has been argued in this thesis that the core structure seen in this cluster is also likely to be real. Choosing the position of the south-western peak as the center, the core radius falls from the 91 kpc quoted in Table 6.1 to 30 kpc. If on the other hand, the position of the dumb-bell galaxy, which lies between the peaks in Figure 3.12, is chosen instead, $r_c = 140$ kpc. Many more examples could be cited from the HGT sample clusters. In these cases however, the

reality of the structures is less certain.

The above discussion indicates two possible conclusions. If clusters really have constant density cores, the core substructure observed in the adaptive-kernel maps can not be real and must be due to contamination from background/foreground objects. Hence it is difficult to estimate the size of these cores with confidence, due to the sensitivity of the measurement to this contamination. Previous estimates of the core radii of clusters are unlikely to be accurate. The more likely conclusion is that the core substructure seen in the maps is, in many cases, real. When this substructure is taken into account, the core radii of clusters drops to values that are consistent with those measured for clusters containing arcs. The constant density cores observed for clusters in the past as well as those seen in the profiles of Figure 6.1 are caused by the failure to recognize the complexity of the cluster core.

The difference between the core radius measured from galaxies and that measured from either X-ray gas or the gravitational lens observations, has been used to support the claim that the dark matter in clusters is distributed differently than the galaxies. While it may indeed be the case that the two distributions differ, the core radii measured using the galaxy projected positions, either obtained by a parametric fit or by the nonparametric procedure above, should *not* be used as evidence for this. The reason is simply that the measurement errors are too large and have been consistently *under*-estimated. While the bootstrap error listed above is an improvement over the errors obtained using a parametric approach, it needs to be realized that this error is only an error in the density *estimate*, and not in the *actual* density. Perhaps a better estimate of the actual error would be to include in the bootstrap procedure variations in smoothing parameter as well as variations in the position of the center. Such refinements would be quite expensive computationally.

6.3.2 Power Law Profiles and Estimation of Ω_0

A number of studies, both numerical (Crone *et al.* 1994) and observational (Beers & Tonry 1986), have found that power law fits are a good approximation for the density profiles of galaxy clusters. Furthermore, the slope of the power law may have cosmological implications. Simple gravitational collapse can be shown to give rise to an r^{-4} profile (Gott 1975; Gunn 1977). Observations, on the other hand, indicate shallower profiles of r^{-2} (Beers & Tonry 1986). This discrepancy can be explained by secondary infall of surrounding material or accretion of smaller groups. Clusters which have recently accreted material will exhibit a flatter density profile than clusters which have not recently accreted material. As discussed in the previous chapter, in a high-density Universe clusters continue to accrete material in the present epoch. On the other hand, in a low-density Universe the mass should be too spread out at the present epoch for accretion of appreciable amounts of material.

The best fit power law is calculated for each of the clusters in the sample. A line was fit to the log-log plots of the space density profiles using a robust least-squares algorithm. Since the estimates for ν are not truly linear over the entire range, the slope was calculated from the core radius of each cluster out to a radius of one Mpc. Unlike the estimated values of core radius, changes in the smoothing parameter had little effect on the slopes. The values for each cluster are listed in column (6), with the one-sigma error in column (7) of Table 6.1. The median slope of -1.8 ± 0.3 was obtained. If an average constant-density background of 30% is assumed, then Monte-Carlo simulations indicate that the profiles will be too shallow by 0.15 ± 0.05. The background corrected slope is therefore likely to be -1.9 ± 0.3, in good agreement with previous results (Beers & Tonry 1986).

Comparison with numerical simulations (Efstathiou *et al.* 1988; Crone *et al.* 1994;

Jing *et al.* 1995) indicates that this is most consistent with an $\Omega = 1$ Universe with a power-law initial perturbation spectra $P(k) \propto k^{-2}$. Because of the degeneracy between the spectral-power index and Ω_0, a lower-density universe with a power index in the range of -3 to -4 cannot be ruled out. However, such a steep power law for the initial perturbation spectra is currently inconsistent with the observations (Henry & Arnaud 1991, Fisher *et al.* 1993, Peacock & Dodds 1994, Feldman *et al.* 1994). On the other hand, the flattened profiles seen here and in other studies using projected galaxy positions, may be due to incomplete background subtraction.

6.4 Conclusions

Despite the common conception of clusters containing constant-density cores in the galaxy distribution, there is *no* real evidence to support this. As the core of a cluster is approached, the number of galaxies available becomes smaller and the density estimate is poorly constrained, as is clearly demonstrated by the bootstrap confidence curves in Figure 6.1. It is invariably this poorly constrained region that is identified as a constant-density core using parametric techniques. Mis-specification of the cluster center merely leads to less galaxies near the center, a larger poorly-constrained region, and therefore a larger core radius. Thus, the core radii listed in Table 6.1 should be viewed as an upper-limit which is likely to be set more by the sample statistics than by a physical size of the core.

Chapter 7

CONCLUSIONS

7.1

This thesis calls into question the concept of clusters of galaxies being modeled as isothermal spheres, Michie-King distributions, or with any other equilibrium model. Clearly the idea that a galaxy cluster can be described by the two parameters of core radius and velocity dispersion is an extreme over-simplification, and probably not even a useful approximation to actual clusters. The growing body of evidence suggests that clusters are still in the process of formation and accretion of material, as evidenced by the large fraction of clusters with substructure and the flattened density profiles, continues at the present epoch. While both of these can be taken as evidence of a high cosmic density parameter Ω_0, possible alternative explanations exist.

The studies in this thesis have been based on the positions of galaxies in clusters. However, the mass in luminous galaxies is only expected to make up about 1% of the total cluster mass. Given this, it is quite possible that the galaxy positions have an entirely different distribution from that of the total mass. If the gravitational lens observations find steeper slopes are the norm, we would be forced to accept a low-density, open cosmology. Although the results are still somewhat uncertain, studies of several clusters with arcs indicate an isothermal profile away from the core, or

$\rho \propto r^{-2}$, is a good approximation (Bonnet *et al.* 1994, Smail *et al.* 1992, Tyson & Fisher 1995, Luppino & Kaiser 1997, Squires *et al* 1996), consistent with the results found here.

Another possibility is the poorly-known effects of small-scale structure on the collapse times of larger objects. If small-scale structure can delay the collapse of galaxy clusters till the present epoch, then the substructure seen in clusters, as well as the flattened density profiles, could be due to material that was accreted long ago and just now collapsing.

7.2 Future Work

The most obvious extension of the work contained in this thesis is to apply this type of cluster analysis to the entire Abell catalog. The existence of the APS project, as well as other similar projects such as the APM survey, beg for such an extension to be carried out. The importance of a consistent set of contour maps is hard to over estimate. Any researcher studying Abell 787, for instance, should be aware that the Abell center appears to be near of low-density trough with two high-density peaks to the east and west.

With the significant groups identified in the projected surface density, it remains for projection effects to be eliminated with the gathering of redshifts. Although a K-S test on the distribution of magnitudes can be used to estimate whether distributions of the groups are radically different, as in this thesis, evidence exists that there is not a universal luminosity function for galaxy clusters (Jones & Mazure 1996). It appears some clusters are simply fainter than other clusters. Therefore, the classification of groups as foreground/background in this thesis may be more misleading then than informative, as is demonstrated in the case of A1837 and A1836.

For the clusters with X-ray observations, detailed comparison could be made between the X-ray maps and the adaptive-kernel maps. This may help clear up some of the questions regarding background contamination, as well as indicate clusters that have undergone recent mergers. Furthermore, a detailed comparison needs to be done between the X-ray-derived mass profiles and the projected galaxy number-density profiles where *both* profiles are calculated using a fully nonparametric technique.

LIST OF REFERENCES

Abell, G. O. 1958, ApJS, 3, 211

Abell, G. O., Corwin, H. G., Olowin, R. P. 1989, ApJS, 70, 1

Anderssen, A. J. & Jakeman, A. J. 1975, J. Microscopy, 105, 135

Bautz, L. P. & Morgan, W. W. 1970, ApJ162, L149

Abramson, I. S. 1982, Ann. Statist. 10, 1217

Antonuccio-Delogu, V. 1996, in Mapping, Measuring, and Modelling the Universe, ASP Conference Series, 94, eds. P. Coles, V. J. Martinez & M.-J. Pons-Borderia (San Francisco: Astronomical Society of the Pacific), p. 63

Ashman, K. M., Bird, C. M., & Zepf, S.E. 1994, AJ, 108, 2348

Bahcall, N. A. 1988, ARA&A, 26, 631

Bahcall, N. A. 1995, PASP, 107, 790

Baier, F. W. 1983, Astr. Nach. Ap., 304, 211

Baier, F. W., Lima Neto, B.G., Wipper, H. & Braun, M. 1996, in Mapping, Measuring, and Modelling the Universe, ASP Conference Series, 94, eds. P. Coles, V. J. Martinez & M.-J. Pons-Borderia (San Francisco: Astronomical Society of the Pacific), p. 215

Struble, M. F., & Rood, H. J. 1982, AJ, 87, 7

Struble, M. F., & Rood, H. J. 1991, ApJS, 70, 1

Beers, T. C. 1992, in Statistical Challenges in Modern Astronomy, eds. E. Feigelson & G. J. Badu, (New York: Springer-Verlag Inc.), p. 111

Beers, T. C., Forman, W., Huchra, J. P., Jones, C., & Gebhardt, K. 1991, AJ, 102, 1581

Beers, T. C., Gebhardt, K., Huchra, J. P., Forman, W., Jones, C. & Bothun, G. D. 1992, ApJ, 400, 410

Beers, T. C. & Geller, M. J. 1983, ApJ, 274, 491

Beers, T. C. & Tonry, J. L. 1986, ApJ, 300, 557

Bird, C. M. 1993, PhD. Thesis, University of Minnesota

Bird, C. M. 1994a, ApJ, 422, 480

Bird, C. M. 1994b, AJ, 107, 1637

Bird, C. M. 1995, ApJ, 445, L81

Bird, C. M., Davis, D.S. & Beers, T. C. 1995, AJ, 109, 920

Bonnet, H., Mellier, Y., & Fort, B 1994, ApJ427, L83

Briel, U. G. & Henry, J. P. 1993, A&A, 278, 379

Buote, D. A. & Tsai, J. C. 1996, in Mapping, Measuring, and Modelling the Universe, ASP Conference Series, 94, eds. P. Coles, V. J. Martinez,& M.-J. Pons-Borderia (San Francisco: Astronomical Society of the Pacific), p. 189

Cavaliere, A., Colafrancesco, S., & Menci, N. 1992, ApJ, 392, 41

Chamayou, J. M. F. 1980, Computational Physics Communications, 21, 145

Crone, M. M., Evrard, A. E., & Richstone, D. O. 1994, ApJ, 434, 402

Crone, M. M., Governato, F., Stadel, J., & Quinn, T. 1997, ApJ, 477, L5 & Richstone, D. O. 1994, ApJ, 434, 402

Dalton, G. B., Efstathiou, G., Maddox, S. J., & Sutherland, W. J. 1994, MNRAS, 269, 151

Davis, D. S. 1994 PhD. Thesis, University of Maryland

Davis, D. S., & Mushotzky, R. F. 1993, AJ, 105, 409

Davis, D. S., Bird, C. M., Mushotzky, R. F. & Odewahn, S. C. 1995, ApJ, 440, 48

Dempster, A. P., Laird, N. M. & Rubin, D. B. 1977, J. R. Statist. Soc. B, 39, 1

de Vaucouleurs, A., Corwin, Jr., H. G., Buta, R. J., Paturel, G., & Fouque, P. 1991, Third Reference Catalogue of Bright Galaxies (Springer-Verlag New York Inc., New York)

den Hartog, R. 1995 PhD. Thesis, Leiden University

Dressler, A. 1976, PhD. Thesis, University of California, Santa Cruz

Dressler, A. 1980, ApJS, 42, 565

Dressler, A. Shectman, S. A. 1988, ApJ, 95, 985

Dutta, S. N. 1995, MNRAS, 276, 1109

Ebeling, H., Mendes de Oliveira, C. & White, D. A. 1995, MNRAS, 277, 1006

Ebeling, H., Voges, W., Bohringer, H., Edge, A. C., Huchra, J. P., & Briel, U. G. 1996 MNRAS, 281, 799

Epanechnikov, V. A. 1969, Theor. Probab. Appl., 14, 153

Efstathiou, G., Frenk, C. S., White, S. D. M., & Davis, M. 1988, MNRAS, 235, 715

Escalera, E., Biviano, A., Girardi, M., Giuricin, G., Mardirossian, F., Mazure, A. & Mezzetti, M. 1994, ApJ, 423, 539

Miralda-Escudé, J. 1995, ApJ, 438, 514

Everitt, B. S. 1984, Statistician, 33, 205

Feldman, H. A., Kaiser, N., Peacock, J. A. 1994, ApJ, 426, 23

Fisher, K.B., Davis, M., Strauss, M. A., Yahil, A., Huchra, J. P., 1993, ApJ, 402, 42

Fitchett, M. J. 1988, MNRAS 230, 161

Fitchett, M. J. & Webster, R. L. 1987, ApJ, 317, 653

Fletcher, R., & Reeves, C. M. 1964, Comp. J., 7, 149

Frenk, C. S., White, S. D. M., Efstathiou ,G. & Davis, M. 1990, ApJ, 351,10

Fix, E. & Hodges, J. L. 1951, Report No. 4, Project 21-49-004, USAF School of Aviation Medicine, Randolph Field, Texas

Fukunaga K., Hostetler, L. D. 1975, IEEE Tran. Inf. Theory, 21, 32

Gasser, Th., Müller, H. G., Köhler, W., Molinari, L. & Prader, A. 1984, Ann. Statist., 12, 210

Geller, M. J. & Beers, T. C., 1982, PASP, 92, 421 (GB)

Giacconi, R., & Burg, R. 1993, in Observational Cosmology APS Conference Series, 51, eds. G. Chincarini, A. Iovino, T. Maccacaro, & D. Maccagni (San Francisco: Astronomical Society of the Pacific), p. 342

Good, I. J. & Gaskins, R. A. 1971, Biometrika, 58, 255.

Gonzàlez-Casado, G., Mamon, G. A., & Salvador-Solé, E. 1994, ApJ, 433, L61

Gott, J. R. 1975, ApJ, 201, 296

Gunn, J. 1977 ApJ, 218, 592

Gunn, J., & Gott, J. R. 1972, ApJ, 209, 1

Hawkins, D. M. 1981, Technometrics, 23, 105

Henry, J. P., & Arnaud, K. A. 1991, ApJ, 372, 410

Hoessel, J. G., Gunn, J. E., & Thuan, T. X. 1980, ApJ, 241, 486

Jing, Y.P., Mo, H.J., Börner, G., & Fang, L. Z. 1995, MNRAS, 276, 417

Jones, C. & Forman, W. 1992, in Clusters and Superclusters of Galaxies, ed. A. C. Fabian (Kluwer, Dordrecht), p. 49

Jones, B. J. T., & Mazure, A. 1996, in Mapping, Measuring, and Modelling the Universe, ASP Conference Series, 94, eds. P. Coles, V. J. Martinez & M.-J. Pons-Borderia (San Francisco: Astronomical Society of the Pacific), p. 197

Kauffmann, G., & White, S. D. M. 1993, MNRAS, 261, 921

Kaufmann, L. & Rousseeuw, P.J. 1990, Finding Groups in Data: An Introduction to Cluster Analysis (Wiley, New York)

King, I. R. 1966, AJ, 71, 64

Kriessler, J. R., Beers, T. C., & Odewahn, S. C. 1995, BAAS, 186, 7.02

Lacey, C. G. & Cole, S. 1993, MNRAS, 262, 627

Lacey, C. G., Currie, M. J. & Dickens, R. J. 1986, MNRAS, 221, 453

Larson J. A. 1996, PhD. thesis, University of Minnesota

Lasker, B. M. 1995, PASP, 107, 763

Luppino, G. A., & Kaiser, N. 1997, ApJ, 475 L20

Materne, J. 1979, å, 74, 235

McLachlan, G. J. & Basford, K. E. 1988, Mixture Models: Inference and Applications to Clustering (Marcel Dekker, New York)

Merritt, D. & Tremblay, B. 1994, AJ, 108, 514

Mohr, J. J., Fabricant, D. G. & Geller, M. J. 1992, in Proc. 3rd Teton Summer School, The Evolution of Galaxies and Their Environment, ed. D. Hollenbach, H. Thronson, & J.M. Shull (NASA CP-3190), p. 285

Mohr, J. J., Fabricant, D. G. & Geller, M. J. 1993, ApJ, 413, 492

Nelder, J. A., & Mead, R. 1965, Comp. J., 7, 308

Nakamura, F. E., Hattori, M., & Mineshige, S. 1995, å, 302, 649

Peacock, J. A., & Dodds, S. J. 1994 MNRAS, 280, 19

Peebles, P. J. E., 1990, ApJ, 365, 27

Pinkney, J., Roettinger, K., Burns, J. O. & Bird, C. M. 1996, ApJS, 104, 1

Pisani, A. 1993, MNRAS, 265, 706

Pisani, A. 1996, MNRAS, 278, 697

Postman, M., Huchra, J. P. & Geller, M, J. ApJ, 348 404

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1986, Numerical Recipes (Cambridge: Cambribge University Press)

Rhee, G. F. R. N., van Haarlem, M. P. & Katgert, P. 1991, A&A, 246, 301

Odewahn, S. C., Stockwell, E. B., Pennington, R. L., Humphreys, R. M. & Zumach, W. A. 1992, AJ, 103, 318

Odewahn, S. C., Humphreys, R. M. Alserin, G., & Thurmes, P. 1993, PASP, 105, 1354

Pennington, R. L., Humphreys, R. M., Odewahn, S. C., Zumach, W. A. & Thurmes, P. M. 1993, PASP, 105, 521

Richstone, D., Loeb, A. & Turner, E. L. 1992, ApJ, 393, 477

Salvador-Solé, E., Sanromà, M. & González-Casado, G. 1993, ApJ, 402, 398 (SSG)

Schechter, P. 1976, ApJ, 203, 1976

Scott, D. W. 1992, Multivariate Density Estimation (New York: John Wiley & Sons)

Silvermann, B. W. 1986, Density Estimation for Statistics and Data Analysis, (Chapman & Hall, London)

Smail, I., Ellis, R. S., Fitchett, M. J., & Edge, A. 1994, MNRAS, 273, 277

Squires, G., Kaiser, N. Babul, A., Fahlman, G., Woods, D., Neuman, D. M. & Böhringer, H. 1996, ApJ, 461, 572

Stone, C. J. 1984, Ann. Statist. 12, 1285

Tyson, J. A., & Fischer, P. 1995, ApJ, 446, L55

Ulmer, M. P., Wirth, G. P. & Kowalski, M. P. 1992, ApJ, 397, 430

van Haarlem, M. P. 1996, in Mapping, Measuring, and Modelling the Universe, ASP Conference Series, 94, eds. P. Coles, V. J. Martinez,& M.-J. Pons-Borderia (San Francisco: Astronomical Society of the Pacific), p. 191

West, M. J, 1994, in Clusters of Galaxies (Proceedings of the XIVth Moriond Astrophysics Meeting), ed. F. Durret A. Mazure & J. Tran Tranh Van (Editions Frontiers, Gif sur-Yvette), p. 23

West, M. J. & Bothun, G.D. 1990, ApJ, 350, 36

West, M. J., Oemler, A. & Dekel, A. 1988, ApJ, 327, 1

Wahba, G. 1990, Spline Models for Observational Data (Philadelphia: SIAM)

Wu, C. F. J. 1983, Ann. Statist. 11, 95

Zabludoff, A. I. & Zaritsky, D. 1995, ApJ, 447, L21

Zwicky, F. 1933, Helv. Phys. Acta, 6, 110

Zwicky, F. 1937, Phys. Rev. 51, 290

MICHIGAN STATE UNIV. LIBRARIES
31293015951720