THE ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES, 104:1–36, 1996 May

EVALUATION OF STATISTICAL TESTS FOR SUBSTRUCTURE IN CLUSTERS OF GALAXIES

JASON PINKNEY, KURT ROETTIGER, AND JACK O. BURNS
Department of Astronomy, Box 30001, Department 4500, New Mexico State University, Las Cruces, NM 88003; jpinkney,kroettig,jburns@nmsu.edu
AND
CHRISTINA M. BIRD
Department of Physics and Astronomy, University of Kansas, Lawrence, KS 66045; tbird@kula.phsx.ukans.edu
Received 1995 August 10; accepted 1995 November 9

ABSTRACT

We use *N*-body simulations of galaxy cluster mergers to evaluate new and published statistical tests of substructure. These tests include 22 one-dimensional (normality) tests, four two-dimensional (spatial) tests, and five three-dimensional (velocity-spatial) tests. These tests are statistical in that they provide a significance level for the presence of substructure. All the tests are applied to the same data files so that their relative sensitivity can be compared. The data files contain positions and velocities of dark matter particles drawn randomly from the simulations.

Three noncosmological simulations are run in which the subclusters begin as King spheres: a single isothermal cluster, a merger of a $\frac{1}{6}$ mass subcluster, and a merger of a $\frac{1}{3}$ mass subcluster. In this way, we examine the dependence of the tests on the subcluster's relative mass. We examine also the dependence on the total sample size, the epoch of merger (pre- and post–core-crossing), and the projection angle of the merger axis.

The results allow a quantitative comparison of the effectiveness of each estimator under different observational scenarios. In general, the higher the dimensionality of the test, the more sensitive it is to substructure. The sensitivity of individual diagnostics depends on the line of sight relative to the merger axis. The three-dimensional tests are least sensitive to lines of sight perpendicular to the merger axis and are most sensitive to lines of sight 45°–60° from perpendicular. The two-dimensional tests are most sensitive to lines of sight perpendicular to the merger axis. The one-dimensional tests are the most sensitive to lines of sight parallel to the merger axis. No single substructure test is the most sensitive in all situations. Therefore, we recommend that a battery of tests be applied to each cluster. We provide a score for each test reflecting its relative sensitivity. We find that clusters with no merging components but a velocity dispersion gradient have an increased likelihood for a "false-positive" response from some three-dimensional tests.

We examine also the signatures of merger and the detectability of mergers in redshift surveys. We find that cluster masses are overestimated by up to a factor of 2 for clusters undergoing mergers. We plot the dependence of mass estimators on projection angle and epoch of merger. We find that the detectability of postmerger states is hampered by small search radii (e.g., $<2.0\ h_{75}^{-1}$ Mpc) in redshift surveys.

Subject headings: galaxies: clusters: general — galaxies: distances and redshifts — galaxies: interactions — methods: numerical

1. INTRODUCTION

The demand for statistical tests for substructure in clusters of galaxies has grown over the last decade. Substructure is the presence of two or more clumps of galaxies and/or gas within a cluster of galaxies. The most obvious observational signature is a multimodality in the spatial or velocity distributions of galaxies or gas. Substructure is a clear sign of incomplete relaxation in a cluster. Consequently, it complicates the use of clusters as cosmological probes. For example, dynamical estimators of cluster mass can change their values significantly after substructure is objectively removed from a cluster (Bird 1995). Whether clusters are used to study galaxy evolution, or to derive initial conditions through their shapes, radii, or masses, substructure is likely to complicate matters. The percentage of clusters with significant substructure is in itself a useful number for determining Ω (Richstone, Loeb, & Turner 1992; Kauffman & White 1993; Lacey & Cole 1993). Yet there is no consensus on the definition of "significant substructure."

A standardized, calibrated set of tests is clearly desirable in order to diagnose substructure in clusters. Additional motivation comes from the impending explosive expansion of the cluster redshift database. We first saw a great increase of redshifts during the 1980s with the establishment of multiobject spectroscopy (MOS). Several studies of large samples of clusters resulted (Colless & Hewett 1987; Dressler & Shectman 1988; Teague, Carter, & Grey 1990; Zabludoff, Huchra, & Geller 1990). Some of the large surveys are now reaching completion (ENACS, Katgert et al. 1995; APM, Maddox et al. 1990; CfA, Huchra, Geller, & Corwin 1995; cD clusters, Hill & Oegerle 1993; dense clusters, Zabludoff et al. 1993). With the Sloan Digital Sky Survey, the redshift database will expand by orders of magnitude: $\sim 10^6$ redshifts are expected from the northern hemisphere (Gunn & Weinberg 1995).

These substructure diagnostics should be robust and applicable for small data sets for several reasons. First, attention is turning to poor clusters and groups of galaxies. It has been suggested that poor clusters are presently in the process of forma-

tion, and therefore they are expected to be incompletely virialized (Beers et al. 1995; Burns et al. 1994; Venkatesan et al. 1994). They are also more common than rich clusters, thereby providing better statistics in volume-limited samples. Finding substructure in poor clusters will demand sensitive, accurately calibrated substructure tests because of the limited sample sizes. Second, programs studying high-redshift clusters (e.g., gravitational lenses) are likely to deal with small redshift sample sizes because of the flux limits of MOS systems. Third, X-ray observations from *Einstein* and *ROSAT* revealed significant subclumps, some of which coincide with optical regions of low galaxy concentration (e.g., the "dark" objects in A1367 and A2256; Bechtold et al. 1983, Henry & Briel 1991). The small number of galaxies in these systems may be difficult to distinguish dynamically. Finally, the bending of wide-angle tailed (WAT) radio sources may be caused by cluster-subcluster mergers rather than the motion of the WAT galaxy (Pinkney et al. 1993, 1995; Pinkney, Burns, & Hill 1994). The substructure in this case may be subtle because the subclusters must have already crossed and begun mixing in order to generate the tail-bending ram pressure.

In this paper, we will explore the limits of detectability of substructure. We will examine substructure tests in the literature, as well as new variations on published diagnostics. We are also interested in the dynamics of cluster-subcluster mergers. Two primary questions will be addressed in this paper: (1) What are the relative strengths of the substructure tests? (2) How detectable are two-body, cluster-subcluster mergers in redshift surveys?

We use N-body simulations to answer both of these questions. We sidestep the issues of cosmological initial conditions, identification of clusters, and identification of galaxies in simulations. Instead, we have contrived an initial condition of two clusters, modeled by lowered isothermal king spheres (Binney & Tremaine 1987, p. 232), initially separated by several megaparsecs. We randomly select subsamples of dark matter particles to represent observed galaxies. The subsamples are treated as cluster data sets, and standard diagnostics are performed on them. The simulations are ideal for goal (1) (above) because we *know* that substructure is present, and we know *which* particles originate in the subcluster. Also, they are uncluttered by extra subclusters, filaments, and late infall which appear in more sophisticated simulations. Our simulations are less ideal for goal (2) because they may lack realism (see § 4.4). This is an area for continued future work.

We will compare the performance of the substructure tests over a wide range of parameter space. The likelihood that any particular test will detect substructure depends on the relative size of the subcluster, the sample size, the projected spatial separation of the two components, and the radial velocity separation of the two cluster components. These factors can be adjusted by the choice of the fraction of particles representing the subcluster, the total number of particles drawn from the simulation, the merger epoch, and the projection angle of the merger axis.

This paper will proceed as follows. The substructure tests will be defined in § 2, and the N-body code will be discussed in § 3. In § 4 we describe the procedure of applying the tests to the simulations. We also discuss how comparable the N-body data sets are to real data sets. In § 5 we present the results, summarizing each test individually and discussing its relative performance when applied to the simulations. We then discuss the overall detectability of substructure in a two-body merger scenario. Next, we quantify the effect of substructure on mass estimators. Finally, we quantify the effect of a finite search radius in redshift surveys on the detectability of substructure. In § 6 we summarize the results of this study.

2. THE TESTS

For comparability, we have collected only substructure tests which can provide a significance level. The significance level of our tests will be given as a probability that the distribution (spatial or velocity) could be drawn randomly from a distribution free of substructure. Such "hypothesis tests" generally take qualities such as unimodality, Gaussianity, spatially uniform velocity dispersion, or symmetry as the definition of "substructure free." The majority, but not all, of the published hypothesis tests are included. The basic information about our tests is tabulated in Tables 1–3. We will use the short names in column (2) of Tables 2 and 3 for referring to the tests within figures and tables.

The tests are grouped based on their dimensionality. "One-dimensional" (Table 1) refers exclusively to the velocity distribution, never to a spatial dimension. Similarly, "two-dimensional" refers only to the two dimensions in the plane of the sky (e.g., right ascension and declination). A "three-dimensional" test utilizes the two spatial coordinates and the radial velocity coordinate.

TABLE 1

ONE-DIMENSIONAL SUBSTRUCTURE DIAGNOSTICS

Test Name (1)	Sensitivity (2)	Probabilities Available? (3)
A	Kurtosis	No
U	Kurtosis	Yes
W	Kurtosis	Yes
B_1	Skewness	Yes
B_2	Kurtosis	Yes
B_1B_2	Kurtosis/skewness	Yes
I	Kurtosis/skewness	Yes
KS	Kurtosis	Yes
V	Kurtosis	Yes
W^2	Kurtosis	Yes
U^2	Kurtosis	Yes
A^2	Kurtosis	Yes
DIP	Unimodality	Yes
AI	Skewness	No
TI	Kurtosis	No
Secondary Statistics		
V_{pec}	Peculiar velocity	No
SS1	Long on left or right	No
SS2	Tails long or short	No
SS3	Skewed left or right	No
av10	Kurtosis/skewness	Yes
skewav	Skewness	Yes
kurtav	Kurtosis	Yes

NOTES.—Col. (1) gives the name as used in the text. Col. (2) gives the substructure signature to which the test is sensitive. When "yes" occurs in col. (3), the centroid and scale of the test's significance level are tabulated in Tables 7–9.

TABLE 2

TWO-DIMENSIONAL SUBSTRUCTURE DIAGNOSTICS

Test Name (1)	Short Name (2)	Sensitivity (3)	Reference (4)
Angular separation test	AST	Noncentral clumping	West et al. 1988
Fourier elongation test	FE	Elongation	This paper
Lee statistic	Lee 2D	Spatial bimodality	Fitchett & Webster 1987
Symmetry test	β	Asymmetry	West et al. 1988

NOTES.—Col. (2) gives the name of the two-dimensional substructure test used in figures, tables, and most of the text.

3. THE *N*-BODY SIMULATIONS

The *N*-body code used for our three simulations is TREECODE (Hernquist 1987), modified to use a variable time step (Roettiger, Loken, & Burns 1995b). Our first simulation is a single, isolated, virialized cluster. The cluster is represented by 15,000 equal-mass "dark matter" particles in an isothermal King distribution. The cluster's velocity dispersion was chosen to be 800 km s^{-1}, and its core radius was 250 kpc. The softening parameter of the particles was $\sim$50 kpc.

The second scenario is the merger of a single cluster (identical to that in the first scenario) with a subcluster of $\frac{1}{3}$ the mass. The "subcluster" is also modeled by an isothermal King sphere containing 5000 particles and a core radius of 173 kpc. The initial core separation is 6 Mpc. The two components fall together from rest under the influence of gravity. The system is evolved for 250 time steps, corresponding to 11.12 Gyr. The head-on merger occurs at 6.8 Gyr (when the centers of mass cross).

The third scenario features the merger of a smaller, $\frac{1}{6}$ mass, subcluster with the same primary cluster. This subcluster was represented by 2500 particles, with a core radius of 138 kpc. This simulation was run for 12.51 Gyr (281 time steps), with the core crossing occurring at 7.5 Gyr. This allowed about the same amount of time for analysis before and after the core crossing as the 3:1 mass merger. About 40 data dumps, equally spaced in time, were created for both merger scenarios.

4. APPLICATION OF TESTS TO *N*-BODY SIMULATIONS

4.1. *Producing "Pseudocluster" Data Files*

We created pseudocluster data files containing random draws of 30–200 dark matter particles from the output of our simulations. The number of particles drawn from the two subcomponents was required to be proportional to their mass. We projected the six available coordinates into R.A., declination, and radial velocity. The units were changed to hours, degrees, and km s^{-1} to match real data files. Our projection angle is defined as the angle between the line of sight to the cluster and the axis of merger. For each time/projection angle/sample size combination, between five and 20 data files were created. Fifteen data files were used for most sample sizes, but 20 were used for $N = 30$ and $N = 45$, and five to 10 were used for $N = 140$ and $N = 200$, which required excessive CPU time for the analysis.

We selected six of 40 available epochs to make data files out of the merger simulations. These times were strategically chosen, using plots such as Figure 1, to provide a variety of spatial and velocity separations of the subcluster. For the single-cluster simulation, only one time from the middle of the simulation was used. For the 3:1 and 6:1 simulations, the first two epochs chosen were -2.0 and -0.63 Gyr (measured relative to the moment of core crossing). These feature a large spatial separation of the components and a small velocity separation. The subcluster is still intact and well defined. The third and fourth epochs are -0.19 and 0.19 Gyr. They were chosen only 0.4 Gyr apart because the system evolves rapidly during core crossing. Both feature a small spatial separation ($\sim$600 kpc) and large velocity separation ($\sim$2500 km s^{-1}) of the components. The postmerger, fifth and sixth epochs were chosen differently for the two mass ratios. For the 3:1 mass ratio, the fifth is during the second velocity separation maximum at 1.68 Gyr, when the majority of the subcluster is falling back through the primary. The sixth is 3.76 Gyr; the system consists of an escaped, dispersed population, and a relaxed system. For the 6:1 mass ratio, the sixth epoch is taken to be the secondary maximum of relative velocity, and the fifth occurs when the subcluster center of mass is still speeding away from the impact point.

TABLE 3

THREE-DIMENSIONAL SUBSTRUCTURE DIAGNOSTICS

Test Name (1)	Short Name (2)	Sensitivity (3)	Reference (4)
Delta test	Δ	Change of μ and σ with position	Dressler & Shectman 1988
Epsilon test	ϵ	Change of σ and ρ with position	Bird 1993
Alpha test	α	Centroid shift with velocity, σ weighted	West & Bothun 1990
Alpha variation	α var	Centroid shift with velocity, ρ weighted	This paper
Three-dimensional Lee statistic	Lee 3D	Bimodality	This paper

NOTES.—Col. (2) gives the name of the three-dimensional substructure test used in figures, tables, and most of the text.

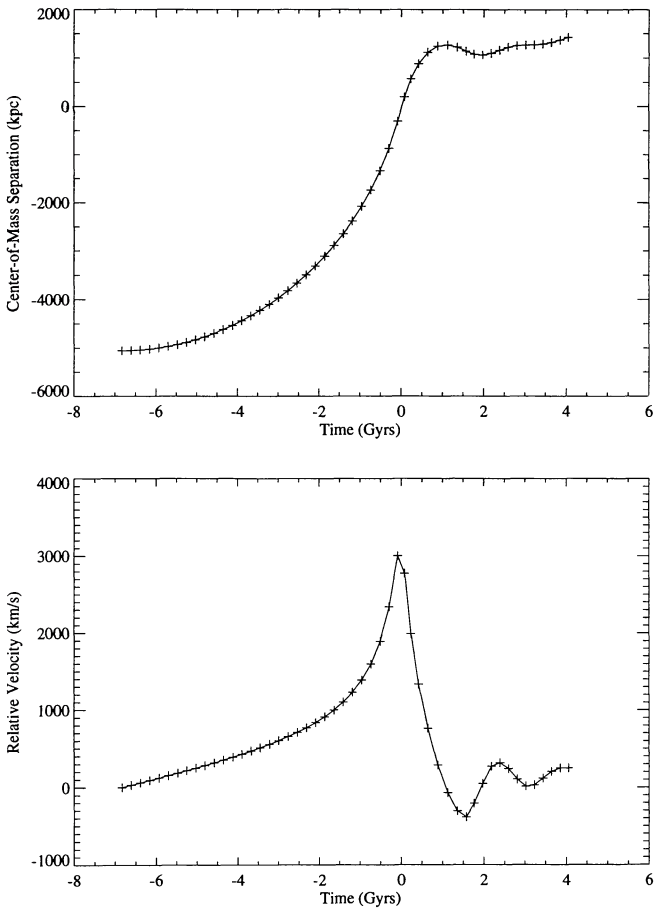

FIG. 1.—Spatial (*top*) and velocity (*bottom*) separation of the cluster and subcluster vs. time for the merger simulation containing the $\frac{1}{3}$ mass subcluster. Distances and velocities are measured along the axis of merger. Data dumps were available at the crosses. Data dumps were converted into data files at −2.0, −0.63, −0.19, 0.19, 1.68, and 3.76 Gyr.

Table 4 shows the parameters which were investigated for this paper. There are a total of 36 parameter combinations. (Note that redundant cases occur.) Figure 1 was an aid in selecting which epochs to study. The third epoch, $T = -0.19$ Gyr, was chosen for a detailed study of both projection angle and sample size. The default projection angle for the study of substructure versus time was 60° (30° was also used for the 3:1 mass ratio), and the default sample size was 100.

The appearance of the 3:1 cluster-subcluster system as a function of time is depicted in Figures 2 and 3. The 60° viewing angle causes a small (0.86 times) foreshortening along the merger axis while allowing some velocity separation of the components. The two clusters begin as spherical and compact systems. The subcluster becomes elongated in the epoch just prior to core crossing ($T = -0.19$ Gyr) and remains nonspherical thereafter. A dispersed population of both subcluster (mostly) and primary cluster particles is noticeable in the last two epochs.

4.2. *The Analysis*

Each data file is analyzed in the following manner. First, kinematical quantities like spatial centroid, velocity mean, and dispersion are calculated. These allow a "3σ" clipping, commonly applied to real data (Yahil & Vidal 1977). On average, less than 1% of the points are rejected from each data file. Second, these same quantities are determined for the subcluster and cluster individually. This allows us to find the projected velocity and spatial separation of the components, the "peculiar velocity" of a galaxy at rest with respect to the primary, and the extent to which the subcluster alters the centroid and scale of the cluster (see § 5.2). Third, substructure tests are applied to the data file. Finally, the significance level for each estimator is calculated by comparison with 300 Monte Carlo shuffles, or randomizations of the data file. We also recorded the dispersions of the significance levels which allow error bars to be plotted in § 5. We use the biweight estimators, S_{BI} and C_{BI}, for dispersion and mean, respectively (see Beers, Flynn, & Gebhardt 1990).

4.3. *The Normalization of Statistical Estimators*

A substructure test statistic is generally meaningless without normalization. The result for the input data file must be compared to the results for data files consistent with a *null hypothesis*. For a one-dimensional (normality) test, the null hypothesis is a Gaussian because we expect an isothermal distribution of self-gravitating bodies to have a Gaussian one-dimensional velocity distribution. For a two-dimensional (spatial) test, the null hypothesis is an azimuthally symmetric, smooth distribution of points in a plane, with the surface density decreasing with radius. For a three-dimensional (spatial-velocity) test, the null hypothesis is no correlation between position and velocity. This means that the velocity mean and dispersion should be the same locally as globally, within counting statistics, where "locally" refers to any subregion projected on the cluster.

The null hypothesis was produced in different ways for different tests. For three-dimensional tests, we ran Monte Carlo simulations of the input file in which the velocities were shuffled randomly with respect to the positions. The positions stayed fixed, except in the case of the three-dimensional Lee statistic, which required azimuthal randomizations of positions as well. Azimuthal randomization was also our means of creating null hypothesis files for two-dimensional tests. The distance of each galaxy from the cluster centroid is conserved, while the azimuth is assigned randomly. This is a good technique because it duplicates exactly the radial profile of the input data rather than assuming one (West, Oemler, & Dekel 1988). The one-dimensional tests did not require Monte Carlo simulations. Instead, the results were compared to look-up tables. We will discuss how our methods of normalization differ from the original methods for some tests in §§ 5.3 and 5.4.

We determine a significance level using Monte Carlo simulations in the following way. First, 300 simulation data files were created for each input data file. (Note that since 15 data files were used for each parameter set, 4500 total simulations are used.) Second, we calculate the number of Monte Carlo simulations which indicate greater substructure than the original statistic. Third, we divide this number by 300. For example, if only five Monte Carlo files have a statistic more extreme than the input file, the significance level is 0.017. We would claim that the null hypothesis was rejected at the 1.7% significance level, a marginal result.

Other methods have been used for determining significance

TABLE 4

PARAMETERS EXPLORED FOR THE DETECTABILITY OF SUBSTRUCTURE

			1:0	
SIMULATION PARAMETERS	3:1	6:1	Isothermal	Gaussian
Time:				
Epochs (Gyr)	−1.99, −0.63, −0.19, 0.19, 1.68, 3.76	−2.0, −0.64, −0.2, 0.18, 0.57, 2.18	...	...
Sample sizes	100	100	...	...
Projection angles	30, 60°	60°	...	...
Sample Size:				
Sample sizes	30, 45, 60, 80, 100, 140, 200	30, 45, 60, 80, 100, 140, 200	...	...
Projection angles	60°	60°	...	...
Epochs (Gyr)	−0.19	−0.20	...	...
Projection Angle:				
Projection angles	0°, 10°, 30°, 45°, 60°, 80°, 90°	0°, 10°, 30°, 45°, 60°, 80°, 90°	...	...
Sample sizes	100	100	...	...
Epochs (Gyr)	−0.19	−0.20	...	...
False Positive Rate:				
Sample sizes	...	...	30, 45, 60, 100, 140, 200	30, 100, 140
Projection angles	...	...	60°	...
Epochs (Gyr)	...	...	2.25	...

NOTES.—The first column lists parameter names, while the body of the table contains parameter values for each simulation. The top row gives the ratio of mass of the primary cluster to the secondary cluster. The 1:0 Isothermal mass ratio refers to the subcluster-free simulation. The 1:0 Gaussian column refers to the data files generated randomly from Gaussians. The epochs are measured in Gyr relative to the moment of core crossing except for the "False Positive Rate" epoch, which is measured since the start of the simulation, whereas they are given relative to the moment of core crossing for the other three parameters.

levels which are not rigorously correct. For example, West & Bothun (1990) determined the significance of their α test by modeling the distribution of Monte Carlo simulations as a Gaussian and calculating how many standard deviations from the mean the true α lies. Figures 4 and 5 show that the distributions of two- and three-dimensional statistics are not Gaussian; they are asymmetric with long upper tails. A Gaussian model will have consistently shorter upper tails than the actual distribution, resulting in an overestimate of the significance. Another normalization method is to use the Monte Carlo randomizations of one file for others. Dressler & Shectman (1988) stress that Monte Carlo models must be run for *each* cluster for the Δ test. We confirm that the centroid of the distribution of Monte Carlo statistics depends on the input data file. For some of our tests, the centroid shifts from one input data file to the next are on the order of the scale of the distribution. Thus, using the Monte Carlo distribution of one file for another may result in 1 σ errors in significance levels.

4.4. *How Realistic Are the N-Body Data Sets?*

Our numerical simulations contain accurate modeling of the behavior of dark matter particles in a complex, changing gravitational potential. The resulting data sets are therefore more physically realistic than if the x, y, and V_z coordinates were drawn from Gaussians.

Nevertheless, the data sets are less complicated than real data for several reasons. The simulation data contain no contamination by background and foreground galaxies. Real data sets are prone to observational errors in positions and velocities. Velocities derived from the cross-correlation procedure may be based on the wrong velocity peak for low-signal spectra. Also, the simulated cluster samples are not subject to unfair sampling, as are real redshift surveys. For example, a survey may favor the innermost galaxies, where the probability of cluster membership is highest.

Another difference is that no spatial cutoff was applied in drawing the data sets from the simulation. Redshift surveys often limit observations to 1.0–2.0 h_{75}^{-1} Mpc from a cluster center. This may exclude gravitationally bound substructures. (Recall that little evidence for substructure was found by West & Bothun 1990 when applying their tests to data which had a 1.3 h_{75}^{-1} Mpc radial cutoff.) It is clear in Figures 2 and 3 that our particles are distributed out to 2.5 h_{75}^{-1} Mpc from the cluster centers. We will quantify the effect of a radial cutoff on substructure in § 5.6.

The physics of real clusters is more complex than our simulations in ways which may create differences in the detectability of substructure. In particular, we do not model subtle effects caused by the range of galaxy masses within each cluster. Real cluster galaxies have a large range in masses (as judged by their luminosities). Large galaxies are slowed preferentially through dynamical friction, resulting in mass segregation. The most massive galaxies, in turn, have been arguably found to have "bound populations" of smaller galaxies (Bothun & Schombert 1988, 1990; Gebhardt & Beers 1991; Tonry 1985). Note that such observable signatures make the assessment of dynamical state easier with real data. Our simulation contains no gas or hydrodynamical effects. In general, the distribution of the mass into gas, dark matter, and galaxies is not well known in clusters (Merritt 1987). Our simulated clusters were initialized with an isothermal King sphere. Simulations by many have found that the dark matter and galaxies segregate in subclusters, so that the galaxy velocities can underestimate the total mass by a factor of 10 (West & Richstone 1988; Carlberg & Couchman 1989; Serna, Alimi, & Scholl 1994). Such an effect should influence the detectability of substructure, since subcluster densities would be time variable, and because the sur-

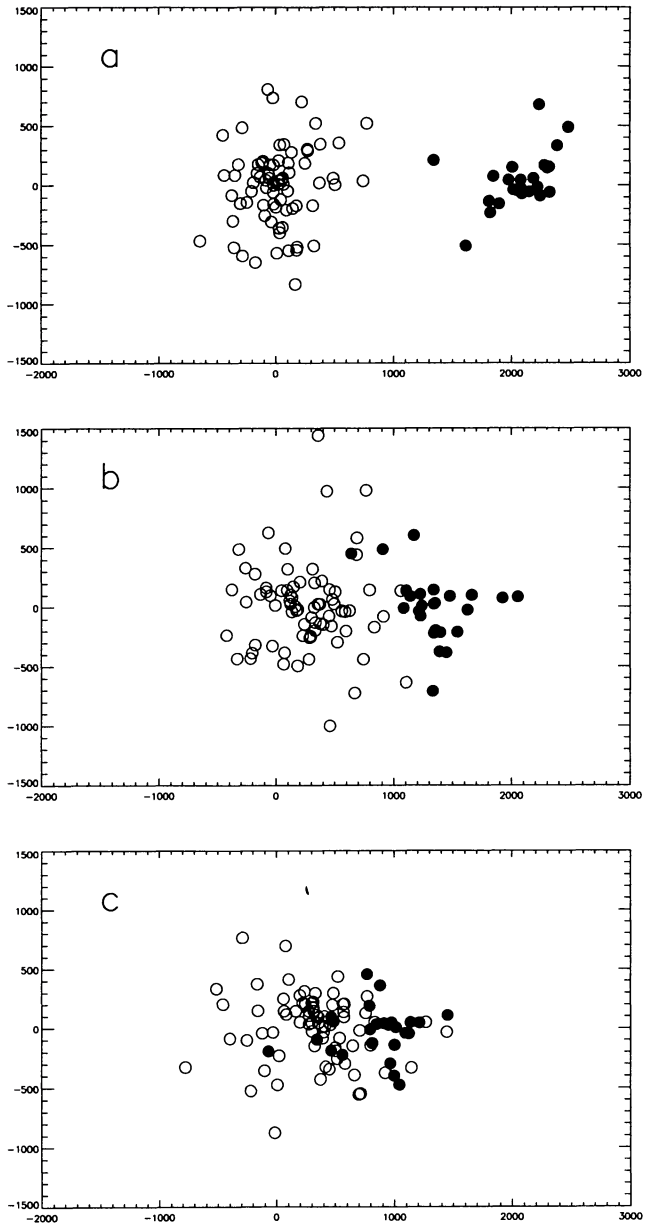

FIG. 2.—Equal area projection of the 3:1 mass ratio merger simulation viewed at (*a*) $T = -2.0$, (*b*) $T = -0.63$, and (*c*) $T = -0.19$ Gyr. The subcluster galaxies are depicted by filled circles. The projection angle (§ 4.1) is 60°. The vertical side of the box is 4.1 h_{75}^{-1} Mpc (the units are arcseconds assuming $z = 0.08$). $N = 100$ size samples are drawn randomly for each epoch.

vivability of subclusters should be altered (see Gonzáles-Casado, Mamon, Salvador-Solé 1994).

Some aspects of our initial conditions may also be rare in the universe. Observations reveal most clusters to have some sort of irregularity, be it asphericity, infalling populations, or substructure. Thus, the chance of finding two spherical, relaxed clusters within 6 Mpc is small. The peak velocity of the components may also be questionable; after falling from rest, the impact velocities of our subclusters reach $\approx$3000 km s^{-1}. Observations of clusters with radial velocity separations over 1500 km s^{-1} have previously been interpreted as unbound systems using simple two-body models (e.g., Beers et al. 1991). Exceptions are A2256 and A754, which have both been modeled as two recently collided subclusters with velocity differences of $\sim$2100 km s^{-1} (Briel et al. 1991; Roettiger, Burns, & Pinkney 1995; Zabludoff & Zaritsky 1995). A radial velocity difference of 2500 km s^{-1} between the two dominant ellipticals in A2255 is also suggestive of an energetic merger (Burns et al. 1995). Moreover, recent cosmological simulations suggest that these impact velocities should be occurring (Crone & Geller 1995). The lack of large observed velocity separations should be partly caused by the scarcity of line-of-sight mergers.

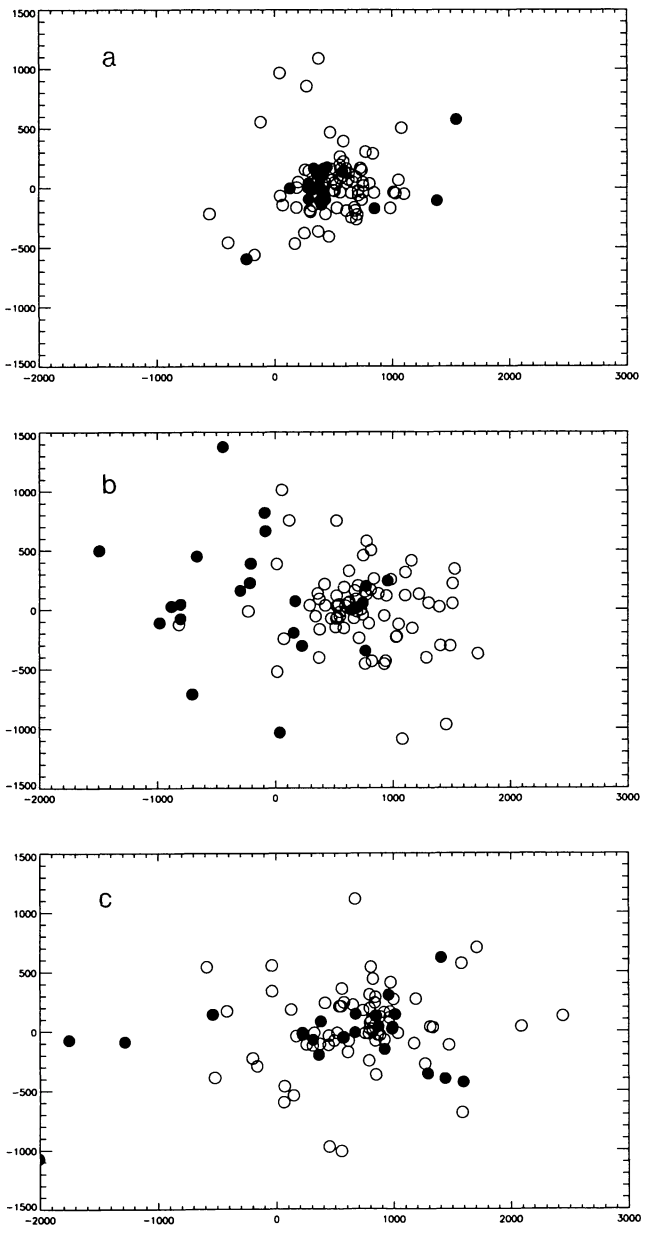

FIG. 3.—Equal area projection of the 3:1 mass ratio merger simulation viewed at (*a*) $T = 0.19$ (*b*) 1.68 and (*c*) 3.76 Gyr. Particles are coded the same as in Fig. 2.

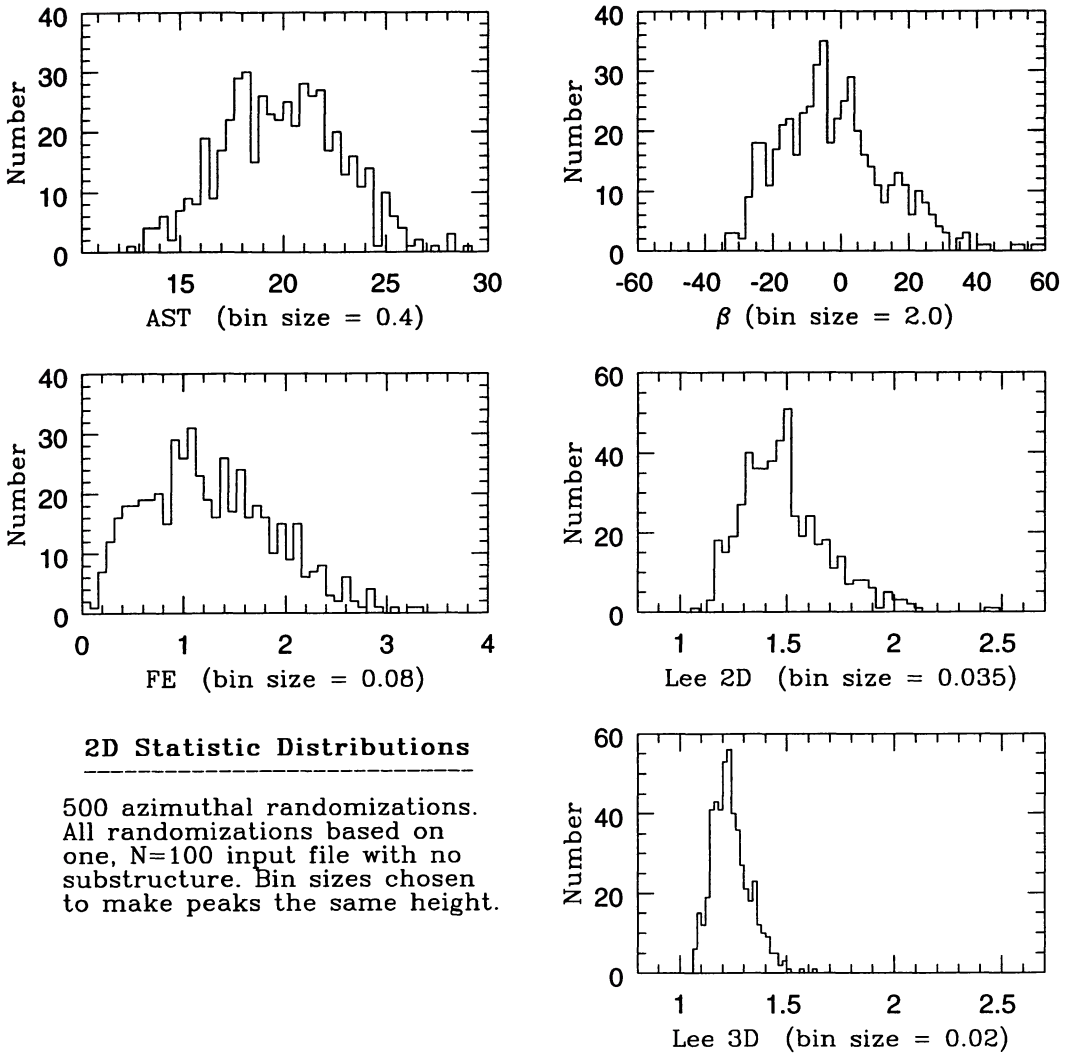

FIG. 4.—Distributions of the values of two-dimensional substructure tests when applied to 500 azimuthal randomizations of one input data file. The input data file came from the *N*-body simulation of a single, isothermal cluster. The peaks of these distributions are located at values typical of the substructure test. (β is multiplied by 1000 here.) The location of the peak depends on the sample size and the distribution of x, y, and v within the input data file.

5. RESULTS

In this section, we will present the dependence of substructure tests on sample size, projection angle, time in merger, and relative size of the substructure. We will discuss individual tests in §§ 5.2, 5.3, and 5.4 and the overall detectability of substructure in mergers in § 5.5. The algorithms for computing most of our substructure tests have been published previously. We will present brief descriptions here for completeness. The dependence of mass estimators on substructure will be discussed in § 5.6, and the effects of a finite search field size are discussed in § 5.7.

As an important check of our methods, we apply our tests first to numerous null hypothesis distributions.

5.1. *Testing for False Positives*

In order to verify that our normalization procedures (§ 4.3) are correct, we apply them to substructure-free data files. We actually did *two* experiments to estimate the *false positive* rates of our tests, i.e., the rate at which a test will claim significance for substructure-free samples. The first experiment was to run our tests on a simulated single, isothermal cluster. The second experiment was to run our tests on randomly generated clusters, with each of their x, y, and V distributions drawn from a Gaussian. The latter experiment ensured that the null hypothesis was met for all tests.

The analysis of substructure-free pseudoclusters was done differently than that of the simulated mergers (§ 4.2). Our program was modified to loop through 100 files instead of 15, and to add up the false positives at three significance levels: 1%, 5%, and 10%. The number of Monte Carlo simulations was reduced from 300 to 100. Since the appearance of the cluster at different projection angles and epochs is identical, the only parameter that required exploration was sample size.

If our simulated clusters contain some characteristic which differs from the null hypothesis, it will appear as an increase in false positives (hereafter, FPs) with sample size. If we produce the null hypothesis correctly in our data files, we expect a test

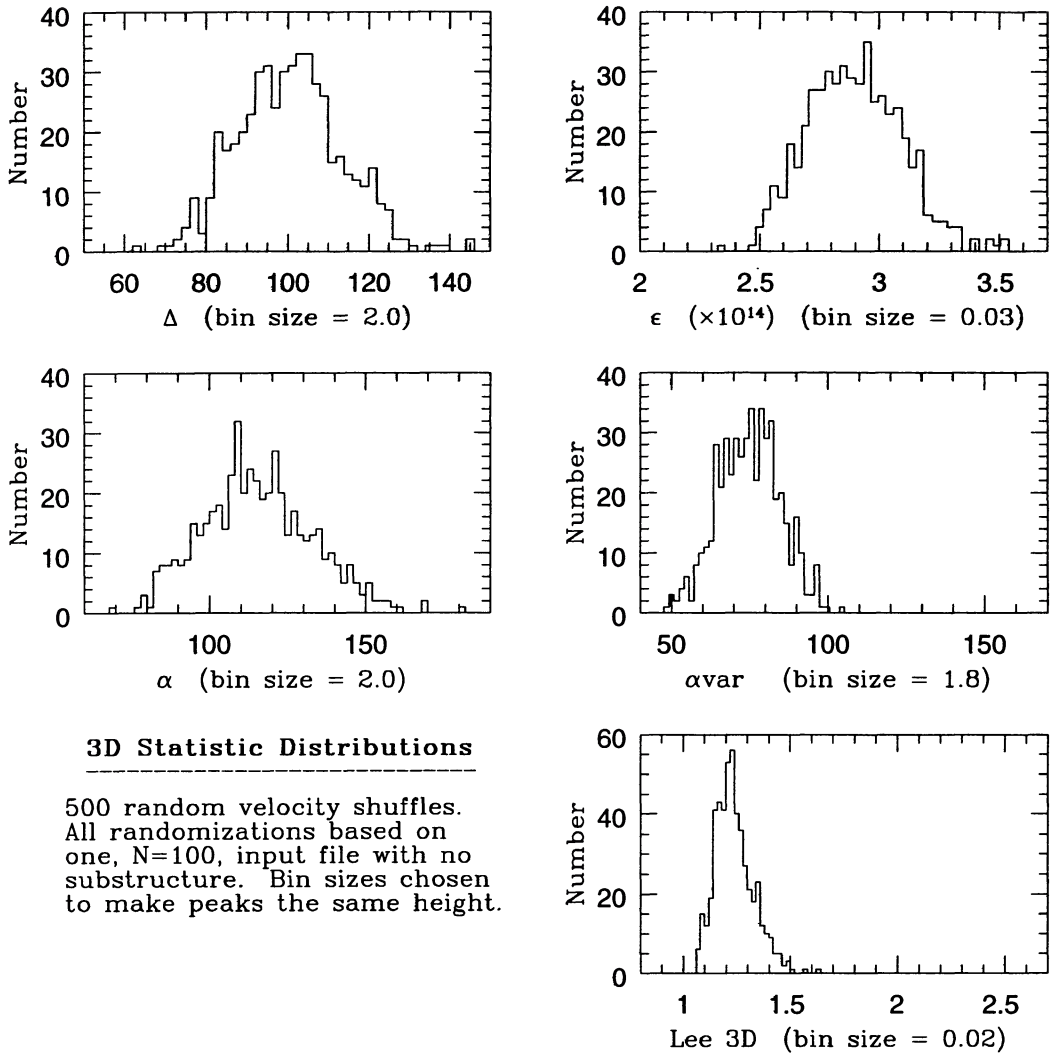

FIG. 5.—Distributions of the values of three-dimensional substructure tests when applied to 500 Monte Carlo simulations of one input data file. (See Fig. 4.)

to find one file out of 100 to be be significant at the 1% level (or better), five files out of 100 to be significant at the 5% level, and 10 files to be significant at the 10% level. Statistical scatter about these numbers is expected because of our finite sample sizes (100 files for each sample size, 100 Monte Carlo simulations of each file). This scatter can be quantified using χ^2 statistics. Variations of 1.65 σ (90% significance) for the counts in the 1%, 5%, and 10% bins are 1.64, 3.68, and 5.2, respectively, for $N = 100$.

5.1.1. *Isothermal Cluster Samples*

The isothermal clusters drawn from our N-body simulation did *not* fulfill the null hypothesis for the three-dimensional estimators. They exhibit a decrease in radial velocity dispersion with radius (Fig. 6). This is expected for a cluster with its mass centrally concentrated because the galaxy escape velocity will decrease with radius. Some clusters which have been well sampled out to large radii show this decrease in dispersion (A2670; Sharples, Ellis, & Gray 1988). The response of three-dimensional substructure tests to a gradient in velocity dispersion has not been quantified until now.

The FP rates for the isothermal cluster samples (Table 5) show more discrepant cases than would be expected from the χ^2 scatter alone (see above). The Δ, α, and ϵ tests give systematically high FP rates. Also, the α test and ϵ test show an increase in false positives with sample size. These deviations from the expected FPs reflect either a problem in the normalization procedure, or the fact that the null hypothesis is not being fulfilled.

A better indicator that the null hypotheses are being produced correctly is the *centroid* of the distribution of significances. The centroid is a more accurate indicator than FPs because all 100 data points are used to calculate it. Centroids of probabilities are plotted in Figures 7 and 8. This is the probability that a sample drawn randomly from the null hypothesis will give a substructure signal greater than or equal to the actual sample. A horizontal line at 0.5 probability is expected for the null hypothesis. This is exhibited, within errors, for all the two-dimensional tests. In contrast, only two of the three-di-

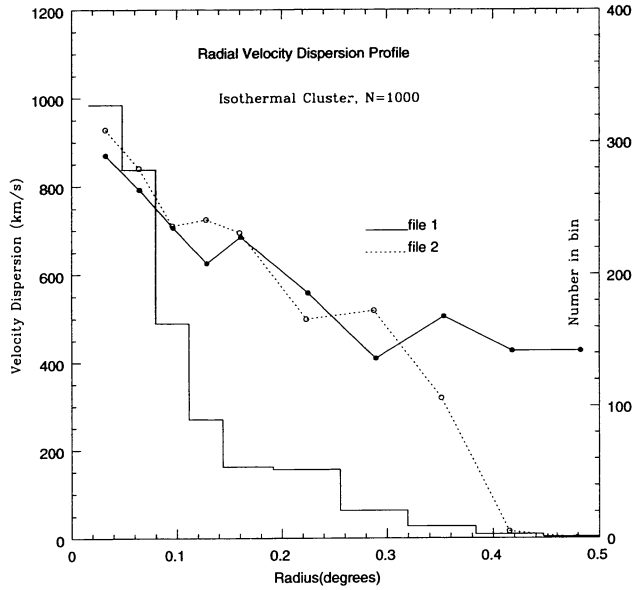

FIG. 6.—Velocity dispersion profiles of our simulated isothermal cluster. Two random samples of 1000 N-body particles are plotted. The number of points used to calculate each dispersion is shown for one file in the histogram (*right scale*). Here 0.5 corresponds to 2.45 h_{75}^{-1} Mpc.

mensional tests, Lee 3D and α var, behave as expected. The $S_{BI}/N^{1/2}$ error bars are plotted only for the Δ test, but the errors for the other tests are comparable. The ϵ test and, to a lesser extent, the α test, show a distinct increase in significance with sample size. This indicates that these tests are responding to structure in the isothermal data files.

5.1.2. *Gaussian Cluster Samples*

The false positive rates for the samples drawn from a Gaussian are shown in Table 6. Three sample sizes are sufficient to explain the unexpected results obtained using the isothermal clusters. We find that *all* the tests have the expected 1%, 5%, and 10% false positive rates, within errors. Moreover, Figures 9 and 10 exhibit the expected probability centroids of 0.5. This suggests that our normalization techniques are sound.

These results confirm that the three-dimensional tests respond to velocity dispersion gradients in clusters. The ϵ, α, and Δ tests are the most sensitive. Taking the α test as an example, about twice as many relaxed clusters with a velocity dispersion gradient will show significance at the 1%, 5%, or 10% level than without ($N = 100$). The ϵ test was the most sensitive to these gradients. Its probability centroid for large samples ($N = 200$) is about 10%. We do not know exactly what effect the gradient has for clusters which also have substructure.

We caution that the 1% level should be adhered to as the limit of statistical significance in blindly applying three-dimensional diagnostics for substructure. Also, 1%–10% significance levels should be considered marginal for the ϵ, α, and Δ tests. These values can be caused by velocity dispersion gradients in addition to real substructure and random fluctuations. However, if the data are of high quality, and are inspected for veloc-

TABLE 5

FALSE POSITIVE RATES FOR A SIMULATED ISOTHERMAL CLUSTER

		SAMPLE SIZE (N)				
TEST (1)	PROBABILITY (2)	30 (3)	45 (4)	100 (5)	140 (6)	200 (7)
FE	≤10%	7.0	10.0	10.0	10.0	6.0
	≤05%	3.0	8.0	5.0	4.0	4.0
	≤01%	0.0	1.0	3.0	1.0	2.0
β	≤10%	13.0	8.0	12.0	13.0	10.0
	≤05%	10.0	5.0	9.0	9.0	6.0
	≤01%	5.0	2.0	4.0	3.0	0.0
AST	≤10%	6.0	9.0	11.0	8.0	16.0
	≤05%	3.0	4.0	3.0	5.0	4.0
	≤01%	0.0	1.0	0.0	1.0	2.0
Lee 2D	≤10%	10.0	10.0	16.0	9.0	6.0
	≤05%	4.0	6.0	8.0	2.0	2.0
	≤01%	1.0	0.0	0.0	0.0	2.0
Lee 3D	≤10%	11.0	12.0	17.0	8.0	8.0
	≤05%	5.0	5.0	6.0	3.0	2.0
	≤01%	2.0	0.0	0.0	0.0	2.0
Δ	≤10%	14.0	12.0	21.0	13.0	14.0
	≤05%	8.0	9.0	12.0	12.0	6.0
	≤01%	1.0	2.0	1.0	1.0	2.0
ϵ	≤10%	3.0	7.0	12.0	25.0	50.0
	≤05%	2.0	4.0	8.0	15.0	30.0
	≤01%	0.0	0.0	1.0	1.0	6.0
α	≤10%	17.0	17.0	23.0	23.0	40.0
	≤05%	9.0	8.0	10.0	11.0	24.0
	≤01%	0.0	3.0	2.0	5.0	2.0
α var	≤10%	8.0	16.0	9.0	8.0	12.0
	≤05%	6.0	7.0	1.0	5.0	4.0
	≤01%	2.0	1.0	0.0	2.0	0.0

NOTES.—Cols. (3)–(7) contain false positive (FP) Counts out of 100 data files. The data files of 30–200 dark matter particles were drawn randomly from the simulation of a subcluster-free cluster. The tests in col. (1) were applied to these data files. Col. (2) contains the substructure significance level; the FP count is the number of files with this probability or a lower one. The numbers represent the number of files out of 100 that had a significance level given in col. (2).

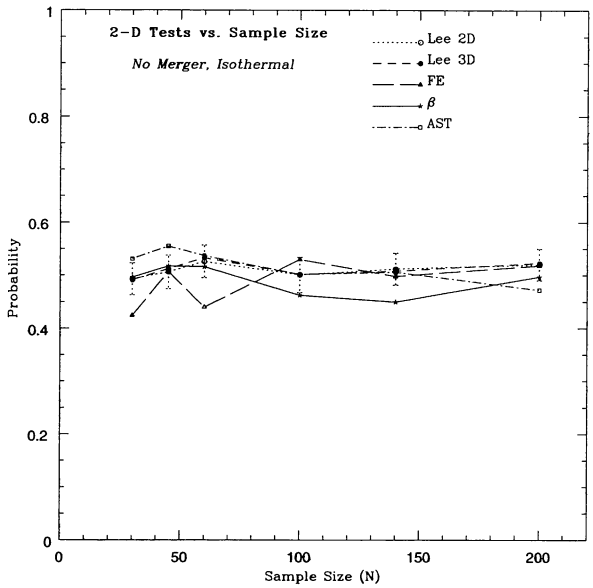

FIG. 7.—Significance levels of two-dimensional substructure estimators applied to the isothermal, single-cluster simulation.

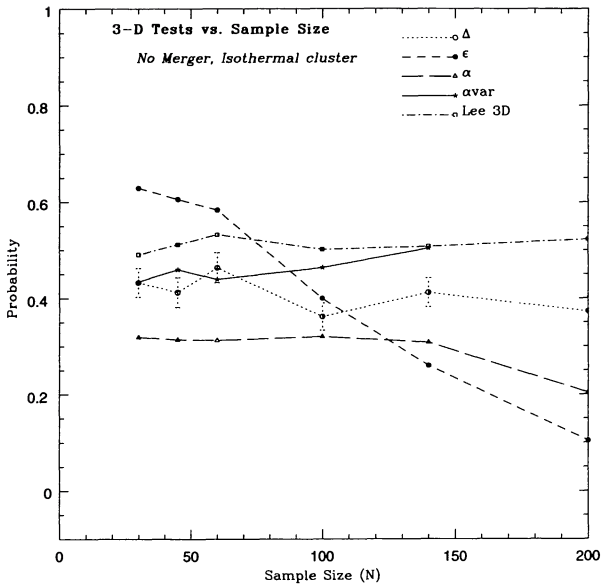

FIG. 8.—Significance levels of three-dimensional substructure estimators applied to the isothermal, single-cluster simulation.

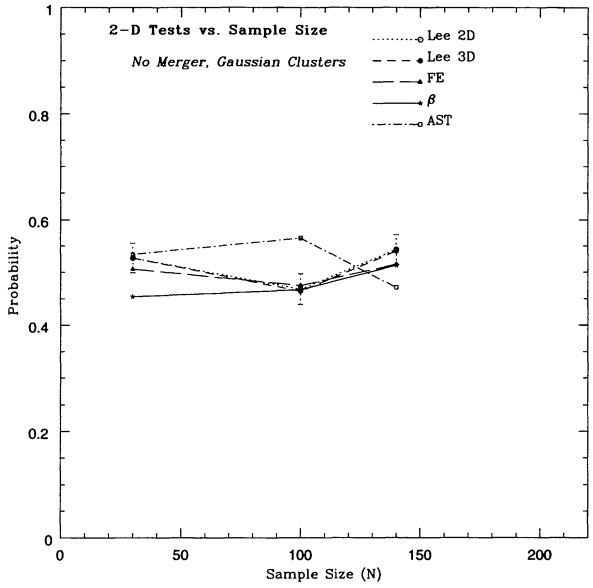

FIG. 9.—Significance levels of two-dimensional substructure estimators applied to the "toy clusters" drawn randomly from Gaussians.

TABLE 6

FALSE POSITIVE RATES FOR THE NULL HYPOTHESIS

TEST (1)	PROBABILITY (2)	SAMPLE SIZE (N) 30 (3)	100 (4)	140 (5)
FE	≤10%	16.0	12.0	9.0
	≤05%	9.0	6.0	3.0
	≤01%	3.0	2.0	0.0
β	≤10%	15.0	13.0	10.0
	≤05%	6.0	7.0	6.0
	≤01%	1.0	2.0	2.0
AST	≤10%	11.0	9.0	9.0
	≤05%	7.0	5.0	3.0
	≤01%	1.0	2.0	0.0
Lee 2D	≤10%	4.0	10.0	6.0
	≤05%	3.0	6.0	3.0
	≤01%	0.0	0.0	0.0
Lee 3D	≤10%	6.0	11.0	7.0
	≤05%	2.0	5.0	3.0
	≤01%	0.0	0.0	0.0
Δ	≤10%	9.0	9.0	9.0
	≤05%	4.0	0.0	5.0
	≤01%	0.0	0.0	2.0
ε	≤10%	7.0	12.0	9.0
	≤05%	5.0	6.0	6.0
	≤01%	1.0	2.0	1.0
α	≤10%	11.0	14.0	14.0
	≤05%	5.0	5.0	8.0
	≤01%	3.0	1.0	3.0
α var	≤10%	10.0	5.0	9.0
	≤05%	4.0	2.0	6.0
	≤01%	1.0	0.0	1.0

NOTES.—Cols. (3)–(5) contain the number of files out of 100 that had a significance level given in the second column. (See Table 5.)

ity dispersion gradients, 1%–5% levels can be interpreted as strong evidence for substructure.

5.2. *Individual Tests: One-dimensional Estimators*

The majority of the one-dimensional estimators are normality tests, i.e., they test the null hypothesis that the velocity distribution is Gaussian. This is important for clusters because models for cluster evolution predict that a system of gravitationally bound particles will "relax" into a Maxwellian velocity distribution, and therefore, a Gaussian distribution is expected for line-of-sight galaxy velocities (e.g., Ueda, Itoh, & Suto

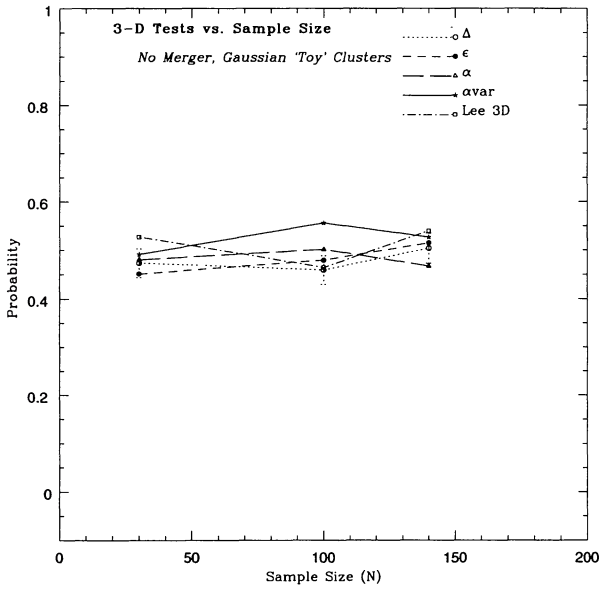

FIG. 10.—Significance levels of three-dimensional substructure estimators applied to the "toy clusters" drawn randomly from Gaussians.

1993). There is, however, the possibility that the equilibrium distribution of clusters is not exactly Gaussian (Fitchett 1988a; Zabludoff et al. 1990).

We adapted the routines in "ROSTAT," by Beers et al. (1990), to run in our program. Not all the routines provided significances with the statistic. We added significance calculations for the DIP test and U test.

Twelve of the ROSTAT tests are already described in Pinkney et al. (1993). Null values and interpretations are given in this paper. The tests utilized in our program are tabulated in Table 1. All these tests, except the A test, are tabulated in Tables 7, 8, and 9. When a "yes" occurs in column (3) of Table 1, the centroid and scale of the test's significance level are tabulated in these tables. The A test is only used in making "SS2." Because of the large number of normality tests, we will not describe them as thoroughly as the two- and three-dimensional tests.

It is evident from viewing Table 1 that most normality tests are sensitive to kurtosis and/or skewness. Kurtosis is a difference in the length, or heaviness, of the tails of the distribution compared to a Gaussian. Skewness is an asymmetry in the distribution. The canonical estimator of skewness, here called "B_1," is the third moment of the distribution,

$$B_1 = \frac{\sum_{i=1}^{N} (x_i - \mu)^3}{N\sigma^3}. \quad (1)$$

To estimate kurtosis, we use B_2, the fourth moment of the distribution,

$$B_2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^4}{N\sigma^4}. \quad (2)$$

The canonical estimator of kurtosis is $B_2 - 3$, where the 3 allows the kurtosis to be 0.0 in the case of a Gaussian distribution. The asymmetry index, AI (also "Finch" statistic), and tail index, TI, were recently tested as alternatives to B_1 and B_2 (Bird & Beers 1993). These authors find AI and TI to be more conservative tests, but also more efficient in most cases. We use the scaled versions so that a Gaussian distribution will give TI = 1.0, and AI = 0.0 (Table 8).

Some tests in Table 1 require special mention. The DIP statistic, for example, tests for consistency with a unimodal hypothesis (Hartigan & Hartigan 1985). Its statistic is the maximum difference between the empirical distribution function (a step function from 0 at the lowest datum to 1 at the highest datum) and the best-fitting, unimodal distribution function. The results for the DIP statistic are given in Table 8. $V_{\rm peculiar}$ is the difference in the velocity centroid ($C_{\rm BI}$) of the entire cluster (subcluster plus primary) and the primary cluster. It is akin to the "peculiar velocity" one would measure for a cD galaxy at rest with respect to the primary cluster. Values over 200 km s^{-1} can be detectable with significance in data sets larger than $\sim$100, depending on the cluster's dispersion and the error in the cD velocity (see Pinkney et al. 1993).

Other secondary statistics were created from the original statistics. The W^2, U^2, and A^2 kurtosis tests can distinguish between long tails on the left (low) or right (high) side of the mode. These were combined to make a statistic "SS1" which ranges from -3 to $+3$, $+3$ meaning a long tail to the right of the mode. The A and B_2 tests can distinguish between a long-tailed and a short-tailed distribution. These made a second statistic "SS2" which ranged from -2 to $+2$. Notice in equation (1) that B_1 can take a positive or negative sign depending on how many points are less than or greater than the mean. This was exploited for a third statistic, "SS3." SS3 is -1 when there are more counts to the left of the mean and $+1$ when there are more to the right. If these secondary statistics have the same value for all 10–20 realizations of a parameter set, it indicates that the distribution is strongly one-sided, or long- or short-tailed, as the case may be.

Finally, since there are so many normality tests, a logical shortcut to judging normality is to average the tests. We formed three such averages. First, "av10," averages the significance of 10 tests, mixing tests sensitive to kurtosis and skewness. Second, "skewav," averages the significance of B_1, the I test, and the B_1B_2 omni tests (all sensitive to skewness). Third, "kurtav," combines the six best estimators of kurtosis.

Our results for the one-dimensional tests are given in Tables 7–9. A subset of these results are plotted in Figures 11–14. Some results may appear suspect at first glance. The KS and V tests have such coarse percentile points that any significance greater than 0.25 will appear as 0.25. Likewise, the I test only shows a 0.15 significance level until the significance drops below 0.1. None of the simulated datasets were extreme enough to cause this. Consequently, scales ($S_{\rm BI}$) for I, KS, and V are 0.0 because no variations in significance occur over 15 data files. The program ROSTAT does not calculate the significance levels for B_1 or B_1 and B_2 correctly if $N \geq 200$. We manually recalculated av10, and skewav for $N = 200$, with these tests omitted. Another caution is that some of these significance levels are folded. Some statistics are two-sided, so that significance levels less than 0.1 and greater than 0.9 *both* imply nonnormality. In order to keep all nonnormal significances less than 0.5 (the convention), initial levels greater than 0.5 are folded (i.e., subtracted from 1.0) to obtain the final significance. This is done for B_1 and B_2 and influences av10 and skewav. One side effect is that the Gaussian distributions have significance centroids of $\sim$0.3 instead of 0.5. Significance centroids less than 0.1 for these tests are still accurate because very few values near 0.5 are included in their calculation.

The dependence on sample size was only inspected for the case of $\phi = 60°$, $T = -0.2$ Gyr. Seven sample sizes were inspected. The velocity separation was only $\sim$1200 km s^{-1} (1.5 σ_v), and so very few probabilities less than 0.1 occurred. Only when N reached 200 was the B_1 test able to detect the skewness.

The radial velocity difference between the cluster and subcluster appears to determine whether the velocity distributions are non-normal. Tables 9–11 show that, when the velocity difference is greater than about 2000 km s^{-1} (2.5 σ_v), the significance levels of most normality tests fall below 0.1 (similar results are reported by Ashman, Bird, & Zepf 1994). Epochs near the time of core crossing and projection angles near 0° will give significant detections of nonnormality for $N \geq 100$. This is shown in Figure 13 where projection angles are $\leq$30°. The change from insignificance to significance is also abrupt in Figures 11 and 12. Note, however, that the scatter in probability ($S_{\rm BI}$ in Tables 7 and 8) is large; parameter sets with a large

TABLE 7

ONE-DIMENSIONAL STANDARD NORMALITY TESTS: PROBABILITY CENTROIDS AND SCALES[a]

PARAMETERS		B_1	B_1B_2	B_2	U	W	KS	V	W^2	U^2	A^2
		c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}
-0.2Gyr,60°	N=30	0.273 0.166	0.487 0.292	0.299 0.219	0.519 0.314	0.449 0.323	0.250 0.000	0.250 0.000	0.501 0.327	0.489 0.321	0.488 0.328
3:1 ratio	45	0.233 0.143	0.469 0.262	0.248 0.152	0.712 0.342	0.499 0.267	0.250 0.000	0.250 0.000	0.558 0.297	0.560 0.300	0.556 0.283
	60	0.390 0.118	0.604 0.326	0.273 0.211	0.491 0.343	0.536 0.283	0.250 0.000	0.250 0.000	0.645 0.278	0.623 0.282	0.606 0.286
	80	0.272 0.122	0.499 0.245	0.212 0.182	0.348 0.266	0.389 0.356	0.250 0.000	0.250 0.000	0.549 0.287	0.534 0.294	0.553 0.268
	100	0.360 0.155	0.496 0.286	0.201 0.158	0.491 0.310	0.348 0.275	0.250 0.000	0.250 0.000	0.462 0.414	0.456 0.398	0.471 0.406
	140	0.380 0.075	0.593 0.336	0.244 0.162	0.508 0.446	0.469 0.345	0.250 0.000	0.250 0.000	0.494 0.219	0.466 0.201	0.562 0.251
	200	... ...	... ...	0.170 0.191	0.545 0.241	0.175 0.113	0.250 0.000	0.250 0.000	0.390 0.502	0.365 0.505	0.381 0.438
-0.2Gyr,60°	N=30	0.306 0.118	0.537 0.316	0.226 0.168	0.503 0.363	0.412 0.269	0.250 0.000	0.250 0.000	0.457 0.335	0.438 0.345	0.443 0.322
6:1 ratio	45	0.336 0.116	0.516 0.317	0.204 0.179	0.456 0.337	0.460 0.323	0.250 0.000	0.250 0.000	0.457 0.324	0.437 0.321	0.467 0.328
	60	0.271 0.143	0.471 0.287	0.252 0.188	0.362 0.349	0.169 0.240	0.250 0.000	0.250 0.000	0.357 0.339	0.325 0.338	0.334 0.364
	80	0.277 0.119	0.542 0.261	0.212 0.186	0.402 0.350	0.404 0.354	0.250 0.000	0.250 0.000	0.635 0.292	0.671 0.289	0.591 0.292
	100	0.253 0.144	0.474 0.249	0.167 0.115	0.474 0.376	0.367 0.258	0.250 0.000	0.250 0.000	0.498 0.338	0.493 0.343	0.475 0.327
	140	0.267 0.169	0.491 0.360	0.285 0.209	0.422 0.355	0.403 0.365	0.250 0.000	0.250 0.000	0.518 0.326	0.521 0.321	0.512 0.328
	200	... ...	... ...	0.072 0.116	0.256 0.351	0.120 0.084	0.194 0.096	0.176 0.122	0.153 0.369	0.144 0.341	0.268 0.417
N=100,30°	T=-2.0 Gyr	0.183 0.131	0.524 0.266	0.316 0.169	0.509 0.257	0.379 0.206	0.250 0.000	0.250 0.000	0.549 0.357	0.560 0.356	0.551 0.333
3:1 ratio	-0.6 Gyr	0.310 0.122	0.265 0.309	0.056 0.112	0.386 0.362	0.284 0.280	0.250 0.000	0.250 0.000	0.167 0.176	0.152 0.159	0.185 0.232
	-0.2 Gyr	0.127 0.111	0.058 0.029	0.070 0.076	0.000 0.000	0.018 0.021	0.056 0.106	0.015 0.024	0.016 0.035	0.015 0.032	0.015 0.040
	0.2 Gyr	0.025 0.024	0.058 0.058	0.251 0.159	0.247 0.334	0.002 0.003	0.015 0.031	0.015 0.031	0.004 0.010	0.007 0.014	0.001 0.005
	1.7 Gyr	0.228 0.152	0.568 0.277	0.383 0.169	0.587 0.253	0.478 0.294	0.250 0.000	0.250 0.000	0.442 0.316	0.458 0.310	0.459 0.317
	3.8 Gyr	0.172 0.133	0.467 0.310	0.278 0.199	0.607 0.294	0.575 0.325	0.250 0.000	0.250 0.000	0.415 0.335	0.418 0.330	0.418 0.366
N=100,60°	T=-2.0 Gyr	0.191 0.110	0.462 0.283	0.290 0.190	0.632 0.356	0.398 0.281	0.250 0.000	0.250 0.000	0.418 0.288	0.426 0.301	0.429 0.280
3:1 ratio	-0.6 Gyr	0.299 0.172	0.436 0.352	0.206 0.129	0.473 0.333	0.310 0.326	0.250 0.000	0.250 0.000	0.405 0.316	0.391 0.297	0.433 0.320
	-0.2 Gyr	0.360 0.155	0.496 0.286	0.201 0.158	0.491 0.310	0.348 0.275	0.250 0.000	0.250 0.000	0.462 0.414	0.456 0.398	0.471 0.406
	0.2 Gyr	0.209 0.125	0.502 0.242	0.265 0.154	0.483 0.328	0.439 0.282	0.250 0.000	0.250 0.000	0.512 0.312	0.521 0.318	0.513 0.296
	1.7 Gyr	0.248 0.141	0.574 0.265	0.326 0.169	0.466 0.324	0.506 0.260	0.250 0.000	0.250 0.000	0.584 0.314	0.584 0.310	0.584 0.313
	3.8 Gyr	0.263 0.131	0.653 0.315	0.307 0.132	0.609 0.275	0.408 0.262	0.250 0.000	0.250 0.000	0.450 0.340	0.444 0.339	0.421 0.344
N=100,60°	T=-2.0 Gyr	0.245 0.124	0.549 0.283	0.299 0.203	0.410 0.358	0.264 0.222	0.250 0.000	0.250 0.000	0.438 0.325	0.440 0.329	0.426 0.303
6:1 ratio	-0.6 Gyr	0.232 0.146	0.525 0.292	0.283 0.179	0.611 0.341	0.461 0.289	0.250 0.000	0.250 0.000	0.524 0.360	0.523 0.362	0.512 0.323
	-0.2 Gyr	0.253 0.144	0.474 0.249	0.167 0.115	0.474 0.376	0.367 0.258	0.250 0.000	0.250 0.000	0.498 0.338	0.493 0.343	0.475 0.327
	0.2 Gyr	0.229 0.176	0.495 0.302	0.311 0.173	0.602 0.254	0.402 0.283	0.250 0.000	0.250 0.000	0.467 0.325	0.477 0.318	0.453 0.315
	0.6 Gyr	0.297 0.127	0.786 0.211	0.370 0.121	0.746 0.252	0.594 0.235	0.250 0.000	0.250 0.000	0.561 0.269	0.543 0.259	0.563 0.255
	2.2 Gyr	0.205 0.153	0.447 0.287	0.269 0.176	0.530 0.384	0.303 0.277	0.250 0.000	0.250 0.000	0.249 0.186	0.259 0.167	0.296 0.206
N=100,-0.2Gyr	PA=0°	0.011 0.020	0.070 0.100	0.333 0.159	0.296 0.315	0.001 0.001	0.010 0.000	0.010 0.000	0.001 0.003	0.002 0.003	0.000 0.001
6:1 ratio	10°	0.010 0.012	0.058 0.052	0.389 0.136	0.368 0.347	0.003 0.004	0.010 0.000	0.010 0.000	0.002 0.002	0.002 0.004	0.001 0.001
	30°	0.021 0.026	0.089 0.107	0.402 0.150	0.305 0.347	0.009 0.012	0.026 0.023	0.012 0.035	0.014 0.018	0.019 0.027	0.011 0.016
	45°	0.219 0.133	0.484 0.326	0.272 0.216	0.439 0.362	0.356 0.280	0.250 0.000	0.250 0.000	0.320 0.316	0.291 0.318	0.354 0.319
	60°	0.253 0.144	0.474 0.249	0.167 0.115	0.474 0.376	0.367 0.258	0.250 0.000	0.250 0.000	0.498 0.338	0.493 0.343	0.475 0.327
	80°	0.248 0.191	0.453 0.327	0.252 0.182	0.520 0.376	0.405 0.302	0.250 0.000	0.250 0.000	0.504 0.321	0.512 0.302	0.524 0.323
	90°	0.245 0.159	0.658 0.255	0.373 0.148	0.871 0.132	0.487 0.257	0.250 0.000	0.250 0.000	0.494 0.299	0.489 0.280	0.482 0.283
Gaussian	N=30	0.288 0.151	0.594 0.294	0.313 0.203	0.491 0.327	0.540 0.292	0.250 0.000	0.250 0.000	0.517 0.278	0.507 0.276	0.519 0.275
	100	0.250 0.132	0.548 0.272	0.293 0.173	0.582 0.300	0.493 0.293	0.250 0.000	0.250 0.000	0.487 0.301	0.480 0.295	0.497 0.300
	140	0.258 0.148	0.571 0.283	0.301 0.156	0.557 0.296	0.415 0.289	0.250 0.000	0.250 0.000	0.482 0.315	0.478 0.311	0.492 0.320

[a] Probability centroids and scales for standard normality tests (one-dimensional substructure tests). These are the results of substructure tests applied to to output from *N*-body simulations (§ 5). The majority of entries come in pairs. These are the biweight centroid and scale (C_{BI} and S_{BI}) of the distribution of values over all datafiles. Usually 15 data files are produced for each parameter set described in the leftmost columns. The C_{BI} for 15 files gives a more accurate estimate of the value than a single data file, and S_{BI} characterizes the range of values possible. Probability centroids less than about 0.05 imply statistically significant substructure, and centroids between 0.05 and 0.1 imply marginally significant substructure.

TABLE 8

ONE-DIMENSIONAL STANDARD NORMALITY TESTS AND SECONDARY STATISTICS[a]

PARAMETERS		I-Prob	DIP-Prob	*av10*	*skewav*	*kurtav*	SS2	SS1	SS3	*TI*	*AI*
		c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	-2=short	-3=left	-1=left	1.0=Gaus.	0.0=Gaus.
-0.2Gyr,60°	N=30	0.150 0.000	0.315 0.278	0.395 0.186	0.303 0.130	0.455 0.243	-1.200 1.295	0.000 2.582	-0.400 0.894	0.929 0.170	-0.126 0.642
3:1 ratio	45	0.150 0.000	0.361 0.287	0.431 0.157	0.281 0.110	0.515 0.214	-0.933 1.560	-0.733 2.726	0.200 0.949	0.959 0.155	-0.072 0.575
	60	0.150 0.000	0.160 0.175	0.477 0.148	0.369 0.133	0.527 0.212	-1.067 1.390	-1.000 2.739	0.067 0.966	0.944 0.097	0.131 0.443
	80	0.150 0.000	0.186 0.190	0.403 0.148	0.307 0.099	0.452 0.206	-0.800 1.378	0.067 2.817	-0.333 0.913	0.958 0.086	-0.154 0.503
	100	0.150 0.000	0.301 0.313	0.372 0.195	0.330 0.126	0.404 0.269	-1.467 0.856	0.200 2.898	-0.467 0.856	0.965 0.112	-0.151 0.504
	140	0.150 0.000	0.275 0.286	0.460 0.150	0.352 0.148	0.553 0.166	-1.200 1.265	-0.600 2.530	-0.200 0.934	0.970 0.051	-0.173 0.557
	200	0.150 0.000	0.112 0.118	0.281 0.160	0.075 0.000	0.370 0.266	-2.000 0.000	3.000 0.000	-1.000 0.000	0.964 0.048	-0.254 0.626
-0.2Gyr,60°	N=30	0.150 0.000	0.226 0.206	0.373 0.175	0.323 0.118	0.413 0.240	-1.000 1.447	0.400 2.697	0.400 0.894	0.919 0.111	0.266 0.482
6:1 ratio	45	0.150 0.000	0.345 0.325	0.375 0.186	0.333 0.128	0.413 0.245	-1.500 1.215	0.500 2.820	-0.100 0.971	0.952 0.102	-0.152 0.492
	60	0.150 0.000	0.238 0.175	0.284 0.215	0.290 0.126	0.261 0.283	-1.333 1.155	1.000 2.739	0.333 0.913	1.009 0.156	0.169 0.571
	80	0.150 0.000	0.150 0.232	0.414 0.175	0.323 0.112	0.470 0.234	-1.600 1.049	-1.133 2.497	0.067 0.966	0.962 0.079	-0.048 0.429
	100	0.150 0.000	0.101 0.158	0.380 0.166	0.294 0.122	0.427 0.213	-1.467 1.111	0.333 2.799	-0.200 0.949	0.987 0.076	-0.089 0.636
	140	0.150 0.000	0.156 0.323	0.390 0.207	0.296 0.160	0.447 0.262	-0.800 1.265	-0.600 2.803	-0.200 0.934	0.999 0.080	0.022 0.572
	200	0.150 0.000	0.046 0.033	0.265 0.283	0.575 0.500	0.168 0.201	-2.000 0.000	3.000 0.000	1.000 0.000	0.917 0.066	0.129 0.115
N=100,30°	-2.0 Gyr	0.150 0.000	0.051 0.064	0.413 0.143	0.275 0.123	0.489 0.180	-1.200 1.183	-0.600 2.846	-0.867 0.483	0.981 0.092	-0.495 0.430
3:1 ratio	-0.6 Gyr	0.150 0.000	0.361 0.267	0.252 0.171	0.248 0.099	0.262 0.217	-1.600 1.049	1.800 2.324	-0.067 0.966	0.904 0.101	0.033 0.388
	-0.2 Gyr	0.150 0.000	0.073 0.193	0.092 0.050	0.126 0.068	0.071 0.052	-1.867 0.483	3.000 0.000	1.000 0.000	0.941 0.110	0.511 0.555
	0.2 Gyr	0.150 0.000	0.084 0.187	0.076 0.097	0.075 0.022	0.110 0.120	-1.200 0.949	3.000 0.000	1.000 0.000	1.037 0.107	0.737 0.416
	1.7 Gyr	0.150 0.000	0.121 0.163	0.406 0.149	0.316 0.140	0.455 0.197	-0.667 1.528	0.600 2.665	-0.467 0.856	0.990 0.101	-0.272 0.500
	3.8 Gyr	0.150 0.000	0.179 0.198	0.383 0.211	0.267 0.146	0.449 0.266	0.933 1.390	0.600 2.846	-0.200 0.949	1.079 0.111	-0.092 0.293
N=100,60°	-2.0 Gyr	0.150 0.000	0.153 0.121	0.373 0.181	0.268 0.120	0.437 0.250	-0.533 1.653	0.600 2.470	-0.600 0.775	0.978 0.097	-0.570 0.279
3:1 ratio	-0.6 Gyr	0.150 0.000	0.207 0.217	0.340 0.185	0.293 0.146	0.372 0.226	-0.667 1.683	0.333 2.708	-0.600 0.775	0.982 0.090	-0.247 0.402
	-0.2 Gyr	0.150 0.000	0.301 0.313	0.372 0.195	0.330 0.126	0.404 0.269	-1.467 0.856	0.200 2.898	-0.467 0.856	0.965 0.112	-0.151 0.504
	0.2 Gyr	0.150 0.000	0.070 0.116	0.387 0.157	0.290 0.112	0.435 0.217	-1.333 1.155	0.067 2.817	0.733 0.658	1.035 0.112	0.313 0.436
	1.7 Gyr	0.150 0.000	0.065 0.093	0.432 0.158	0.316 0.117	0.501 0.223	-0.400 1.449	-0.600 2.846	0.067 0.966	1.022 0.121	0.166 0.150
	3.8 Gyr	0.150 0.000	0.058 0.137	0.392 0.178	0.345 0.144	0.440 0.222	1.067 1.197	0.333 2.799	-0.067 0.966	1.071 0.125	0.141 0.489
N=100,60°	-2.0 Gyr	0.150 0.000	0.190 0.247	0.368 0.154	0.314 0.124	0.410 0.207	-0.667 1.528	0.467 2.781	-0.867 0.483	1.012 0.091	-0.128 0.431
6:1 ratio	-0.6 Gyr	0.150 0.000	0.272 0.257	0.421 0.188	0.299 0.131	0.499 0.259	-0.533 1.494	-0.333 2.708	-0.733 0.658	0.971 0.111	-0.325 0.336
	-0.2 Gyr	0.150 0.000	0.101 0.158	0.380 0.166	0.294 0.122	0.427 0.213	-1.467 1.111	0.333 2.799	-0.200 0.949	0.987 0.076	-0.089 0.636
	0.2 Gyr	0.150 0.000	0.076 0.176	0.386 0.169	0.291 0.149	0.445 0.221	-0.400 1.449	0.067 2.726	0.067 0.966	0.986 0.074	0.251 0.607
	0.6 Gyr	0.150 0.000	0.178 0.213	0.479 0.118	0.421 0.110	0.546 0.168	-0.533 1.653	-0.867 2.689	-0.333 0.913	1.023 0.132	-0.064 0.246
	2.2 Gyr	0.150 0.000	0.194 0.185	0.314 0.143	0.273 0.128	0.335 0.177	0.267 1.713	1.800 2.324	-0.067 0.966	1.091 0.143	0.194 0.546
N=100,-0.2Gyr	PA=0°	0.150 0.000	0.300 0.349	0.090 0.038	0.076 0.060	0.099 0.032	0.267 1.390	3.000 0.000	1.000 0.000	1.203 0.162	0.572 0.444
6:1 ratio	10°	0.150 0.000	0.036 0.072	0.093 0.043	0.068 0.037	0.136 0.066	0.667 1.354	3.000 0.000	1.000 0.000	1.229 0.178	0.814 0.396
	30°	0.150 0.000	0.024 0.046	0.102 0.046	0.097 0.064	0.128 0.065	0.133 1.317	2.600 1.449	1.000 0.000	1.109 0.112	0.801 0.561
	45°	0.150 0.000	0.031 0.053	0.329 0.182	0.285 0.143	0.362 0.235	-1.067 1.197	1.133 2.497	0.867 0.483	0.983 0.092	0.199 0.313
	60°	0.150 0.000	0.101 0.158	0.380 0.166	0.294 0.122	0.427 0.213	-1.467 1.111	0.333 2.799	-0.200 0.949	0.987 0.076	-0.089 0.636
	80°	0.150 0.000	0.264 0.237	0.385 0.183	0.284 0.166	0.447 0.220	-0.267 1.713	-0.200 2.720	0.200 0.949	0.986 0.099	0.144 0.503
	90°	0.150 0.000	0.196 0.175	0.445 0.150	0.351 0.136	0.510 0.185	-0.267 1.390	-0.067 2.817	-0.200 0.949	1.030 0.110	-0.241 0.660
Gaussian	N=30	0.150 0.000	0.281 0.289	0.385 0.168	0.327 0.120	0.424 0.230	-0.740 1.583	0.340 2.843	0.080 0.992	0.974 0.197	0.029 0.506
	100	0.150 0.000	0.163 0.258	0.415 0.169	0.343 0.133	0.467 0.227	-0.660 1.593	-0.020 2.849	-0.040 0.994	1.007 0.109	0.045 0.461
	140	0.150 0.000	0.122 0.152	0.394 0.158	0.313 0.130	0.446 0.207	-0.920 1.500	0.040 2.904	0.060 0.993	0.999 0.080	0.037 0.538

[a] Probability centroids and scales for standard normality tests (one-dimensional substructure tests). (See Table 7.)

TABLE 9

RADIAL VELOCITY AND PROJECTED SPATIAL SEPARATION OF SUBCLUSTER

PARAMETERS		V_{pec} km s^{-1}		Δs_{BI} (Tot-Prim) km s^{-1}		ΔV_r km s^{-1}		Δs_{BI} (Sub-Prim) km s^{-1}		Δx kpc	
-0.2Gyr,60°	N=30	334	58	101	90	1137	252	-319	150	779	130
3:1 ratio	45	297	52	122	57	1142	204	-212	143	763	53
	60	276	53	114	85	1099	286	-204	156	779	113
	80	293	29	145	34	1210	107	-161	102	767	75
	100	292	26	104	41	1149	127	-204	45	772	45
	140	307	23	120	40	1159	117	-211	117	751	53
	200	295	38	130	54	1159	179	-202	84	769	74
-0.2Gyr,60°	N=30	168	86	88	81	1163	342	168	202	560	100
6:1 ratio	45	181	71	68	62	1176	229	18	212	560	100
	60	194	46	102	51	1437	140	122	163	560	100
	80	175	18	97	38	1301	194	-271	90	562	108
	100	173	22	75	44	1183	219	-361	124	573	106
	140	181	21	80	38	1290	209	-256	99	601	107
	200	182	17	107	20	1272	128	-265	52	547	28
N=100,30°	T=-2.0 Gyr	211	42	7	46	748	168	-289	187	1650	74
3:1 ratio	-0.6 Gyr	346	49	161	51	1385	161	-275	123	914	85
	-0.2 Gyr	449	31	393	54	2014	109	-221	93	432	63
	0.2 Gyr	491	78	482	95	2425	320	4	120	152	43
	1.7 Gyr	-95	84	97	51	-422	433	266	140	321	254
	3.8 Gyr	64	55	27	42	259	235	102	187	184	197
N=100,60°	T=-2.0 Gyr	120	51	-30	28	439	176	-195	131	2856	57
3:1 ratio	-0.6 Gyr	214	35	16	45	729	141	-286	95	1518	69
	-0.2 Gyr	292	26	104	41	1149	127	-204	45	772	45
	0.2 Gyr	320	55	190	59	1434	237	6	137	240	46
	1.7 Gyr	-52	45	-5	37	-179	185	-48	134	886	353
	3.8 Gyr	46	49	-25	41	204	188	-109	144	389	355
N=100,60°	T=-2.0 Gyr	85	18	-27	18	490	121	-367	145	2734	78
6:1 ratio	-0.6 Gyr	126	22	9	31	731	171	-365	86	1364	105
	-0.2 Gyr	173	22	75	44	1183	219	-361	124	573	106
	0.2 Gyr	145	31	73	46	1144	307	-71	193	434	48
	0.6 Gyr	92	34	1	23	599	220	-205	159	965	104
	2.2 Gyr	-18	29	-15	25	-108	204	-72	167	891	499
N=100,-0.2Gyr	PA=0°	182	61	318	24	2437	158	-405	129	63	33
6:1 ratio	10°	203	47	331	22	2465	164	-349	73	120	50
	30°	196	29	262	35	2054	141	-286	192	318	82
	45°	207	19	178	14	1705	107	-301	108	480	87
	60°	173	22	75	44	1183	219	-361	124	573	106
	80°	69	28	-23	21	406	180	-328	129	675	43
	90°	-7	42	-34	14	-15	113	-251	133	690	71

NOTES.—Cols. (1) and (2) define the parameter space. Cols. (3)–(7) each contain the biweight centroid and scale of the quantity over 15 data files. Col. (3), "V_{pec}," is the radial velocity centroid of the entire system minus the radial velocity centroid of the primary cluster alone. Col. (4) is the velocity dispersion of the entire system minus the dispersion of the primary cluster alone. Col. (5) is the radial velocity of the subcluster minus the radial velocity of the primary. Col. (6) is the velocity dispersions of the subcluster minus the velocity dispersion of the primary. Col. (7) contains the projected spatial separation of the centers of mass.

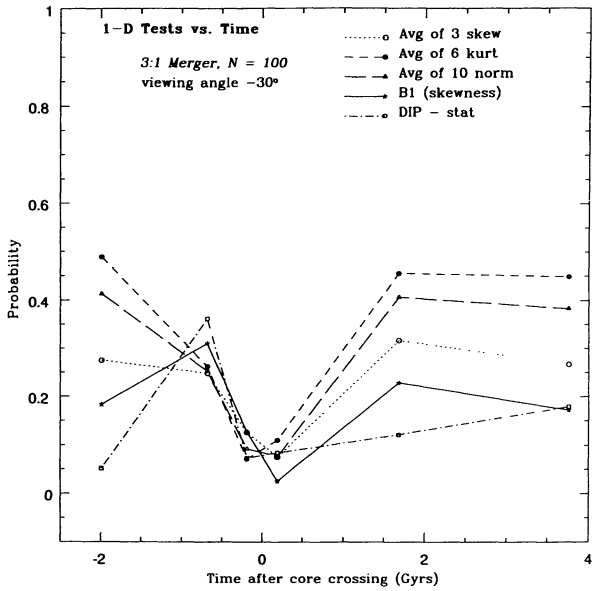

FIG. 11.—Significance levels of one-dimensional (velocity) substructure estimators applied to the 3:1 merger simulation at six epochs. The projection angle is 30°.

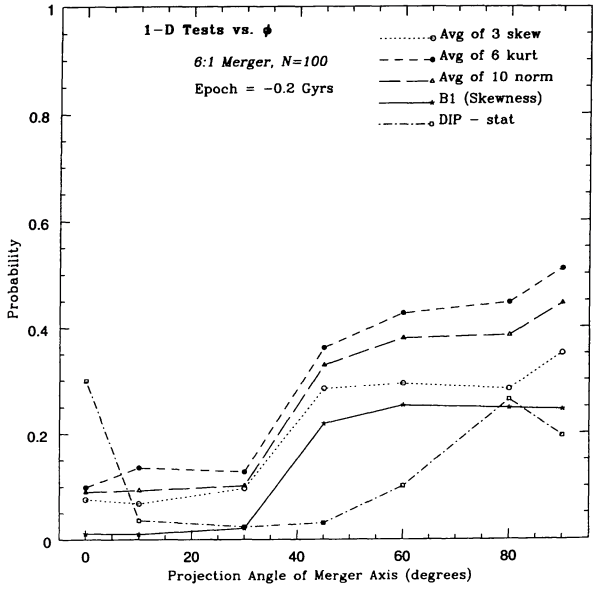

FIG. 13.—Significance levels of one-dimensional substructure estimators applied to the 6:1 merger simulation at seven projection angles. The epoch is 0.2 Gyr before the merger.

probability centroid can produce small individual probabilities.

Significance levels do not correlate as well with the relative size of the subcluster as with velocity separation. In Table 8, comparing the row for $N = 100$, 30°, 3:1 ratio, $T = -0.2$, with that for $N = 100$, $T = -0.2$, 6:1 ratio, P.A. = 30°, one sees that increasing the size of the subcluster does not necessarily improve the significance level. The B_1, W, KS, V, W^2, and A^2 tests all show a decrease in significance with subcluster size. A subcluster size increase seems to help the B_2 and B_1 and B_2 tests immensely. The dependence on sample size can be explained as follows. When a small subcluster (e.g., $\frac{1}{8}$ of the primary in number) is placed in the wings of a Gaussian (e.g., $2\sigma_v$ from the centroid), a distribution will be observed to be long-tailed and asymmetric. As the subcluster grows relative to the primary, the best-fit Gaussian will become broader, and the data will be interpreted as platykurtic. Eventually, when the subcluster equals the primary in size, the distribution will no

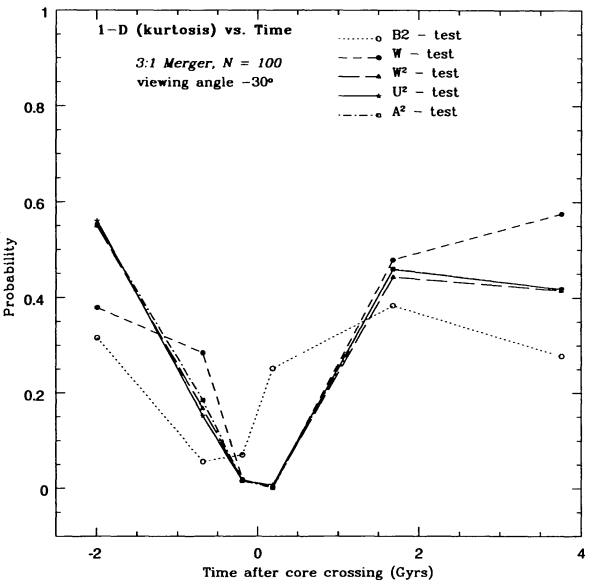

FIG. 12.—Significance levels of one-dimensional kurtosis estimators applied to the 3:1 merger simulation at six epochs. The projection angle is 30°.

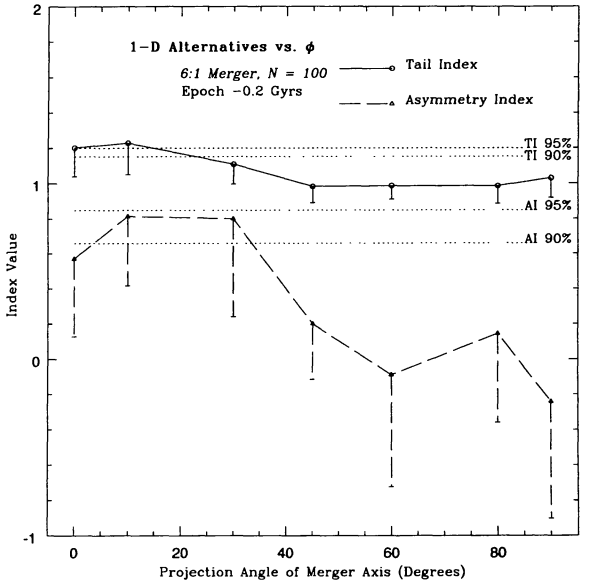

FIG. 14.—Alternative estimators of skewness and kurtosis, AI and TI, applied to the 6:1 merger simulation at seven projection angles. The epoch is 0.2 Gyr before the merger.

TABLE 10

TWO-DIMENSIONAL SUBSTRUCTURE TESTS: CENTROIDS AND SCALES OF SUBSTRUCTURE PROBABILITIES

PARAMETERS		Lee 2D	Lee 3D	F.E.	β	AST	Δ V	Δ x
		c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	km s^{-1}	kpc
-0.2Gyr,60°	N=30	0.372 0.305	0.361 0.332	0.044 0.098	0.322 0.281	0.330 0.210	1137	779
3:1 ratio	45	0.313 0.325	0.198 0.249	0.019 0.034	0.029 0.088	0.327 0.239	1142	763
	60	0.295 0.271	0.187 0.257	0.002 0.005	0.287 0.317	0.208 0.233	1099	779
	80	0.317 0.298	0.215 0.215	0.000 0.000	0.018 0.038	0.274 0.221	1210	767
	100	0.297 0.318	0.179 0.188	0.000 0.000	0.006 0.014	0.227 0.223	1149	772
	140	0.235 0.274	0.055 0.057	0.000 0.000	0.000 0.002	0.098 0.111	1159	751
	200	0.024 0.029	0.000 0.000	0.000 0.000	0.000 0.000	0.245 0.316	1159	769
-0.2Gyr,60°	N=30	0.526 0.289	0.539 0.305	0.497 0.296	0.317 0.251	0.513 0.308	1163	560
6:1 ratio	45	0.399 0.362	0.407 0.351	0.268 0.273	0.435 0.342	0.375 0.272	1176	560
	60	0.411 0.319	0.387 0.380	0.018 0.056	0.320 0.246	0.503 0.299	1437	560
	80	0.418 0.361	0.420 0.344	0.057 0.113	0.331 0.336	0.476 0.256	1301	562
	100	0.444 0.316	0.351 0.283	0.170 0.262	0.323 0.366	0.281 0.270	1183	573
	140	0.354 0.305	0.317 0.277	0.043 0.062	0.362 0.220	0.460 0.212	1290	601
	200	0.127 0.115	0.073 0.090	0.069 0.086	0.095 0.158	0.406 0.404	1272	547
N=100,30°	-2.0 Gyr	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.031 0.042	748	1650
3:1 ratio	-0.6 Gyr	0.014 0.025	0.004 0.008	0.000 0.000	0.017 0.034	0.194 0.149	1385	914
	-0.2 Gyr	0.396 0.321	0.255 0.301	0.186 0.160	0.396 0.250	0.420 0.341	2014	432
	0.2 Gyr	0.452 0.316	0.425 0.313	0.406 0.261	0.301 0.384	0.553 0.287	2425	152
	1.7 Gyr	0.518 0.272	0.511 0.283	0.262 0.254	0.441 0.259	0.463 0.376	-422	321
	3.8 Gyr	0.365 0.278	0.349 0.262	0.430 0.342	0.337 0.352	0.443 0.339	259	184
N=100,60°	-2.0 Gyr	0.000 0.000	0.000 0.000	0.002 0.004	0.000 0.000	0.003 0.005	439	2856
3:1 ratio	-0.6 Gyr	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.022 0.054	729	1518
	-0.2 Gyr	0.297 0.318	0.179 0.188	0.000 0.000	0.006 0.014	0.227 0.223	1149	772
	0.2 Gyr	0.392 0.326	0.421 0.336	0.013 0.033	0.502 0.272	0.362 0.303	1434	240
	1.7 Gyr	0.142 0.434	0.175 0.436	0.008 0.020	0.043 0.095	0.506 0.272	-179	886
	3.8 Gyr	0.080 0.099	0.071 0.096	0.048 0.044	0.018 0.030	0.158 0.283	204	389
N=100,60°	-2.0 Gyr	0.000 0.000	0.000 0.000	0.119 0.155	0.000 0.000	0.179 0.116	490	2734
6:1 ratio	-0.6 Gyr	0.000 0.000	0.000 0.000	0.134 0.216	0.000 0.000	0.310 0.232	731	1364
	-0.2 Gyr	0.444 0.316	0.351 0.283	0.170 0.262	0.323 0.366	0.281 0.270	1183	573
	0.2 Gyr	0.090 0.129	0.101 0.104	0.039 0.065	0.355 0.268	0.354 0.296	1144	434
	0.6 Gyr	0.359 0.317	0.337 0.325	0.069 0.145	0.016 0.044	0.452 0.249	599	965
	2.2 Gyr	0.237 0.307	0.247 0.308	0.192 0.202	0.207 0.164	0.274 0.339	-108	891
N=100,-0.2Gyr	PA=0°	0.387 0.325	0.388 0.340	0.473 0.320	0.558 0.366	0.546 0.267	2437	63
6:1 ratio	10°	0.529 0.295	0.514 0.297	0.494 0.337	0.542 0.331	0.447 0.260	2465	120
	30°	0.525 0.320	0.511 0.313	0.148 0.298	0.303 0.304	0.480 0.295	2054	318
	45°	0.434 0.346	0.351 0.343	0.095 0.157	0.452 0.321	0.456 0.330	1705	480
	60°	0.444 0.316	0.351 0.283	0.170 0.262	0.323 0.366	0.281 0.270	1183	573
	80°	0.401 0.329	0.389 0.306	0.038 0.083	0.161 0.163	0.348 0.242	406	675
	90°	0.362 0.290	0.357 0.289	0.063 0.077	0.137 0.378	0.292 0.285	-15	690
Gaussian	N=30	0.527 0.280	0.527 0.289	0.506 0.356	0.454 0.311	0.534 0.321	...	...
	100	0.468 0.293	0.464 0.299	0.475 0.303	0.467 0.307	0.565 0.320	...	...
	140	0.543 0.278	0.540 0.284	0.515 0.308	0.513 0.287	0.471 0.303	...	...

NOTE.—The last two columns are the projected velocity and position difference between the subcluster and cluster. (See Table 7.)

TABLE 11

THREE-DIMENSIONAL SUBSTRUCTURE TESTS: CENTROIDS AND SCALES OF SUBSTRUCTURE PROBABILITIES

PARAMETERS		Δ	ϵ	α	α var	Lee 3D	Δ V	Δ x
		c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	c_{BI} s_{BI}	km s^{-1}	kpc
-0.2Gyr,60°	N=30	0.066 0.124	0.377 0.300	0.247 0.238	0.234 0.251	0.361 0.332	1137	779
3:1 ratio	45	0.031 0.047	0.260 0.315	0.075 0.097	0.046 0.122	0.198 0.249	1142	763
	60	0.011 0.046	0.415 0.335	0.042 0.061	0.025 0.066	0.187 0.257	1099	779
	80	0.000 0.000	0.059 0.089	0.029 0.033	0.014 0.029	0.215 0.215	1210	767
	100	0.000 0.000	0.004 0.017	0.004 0.011	0.011 0.029	0.179 0.188	1149	772
	140	0.000 0.000	0.005 0.010	0.004 0.006	0.001 0.003	0.055 0.057	1159	751
	200	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	1159	769
-0.2Gyr,60°	N=30	0.288 0.207	0.578 0.268	0.344 0.259	0.347 0.325	0.539 0.305	1163	560
6:1 ratio	45	0.323 0.346	0.710 0.318	0.323 0.250	0.375 0.290	0.407 0.351	1176	560
	60	0.056 0.115	0.253 0.320	0.143 0.209	0.125 0.184	0.387 0.380	1437	560
	80	0.042 0.054	0.254 0.250	0.089 0.109	0.195 0.155	0.420 0.344	1301	562
	100	0.025 0.050	0.183 0.205	0.230 0.229	0.319 0.300	0.351 0.283	1183	573
	140	0.001 0.004	0.008 0.030	0.000 0.000	0.006 0.021	0.317 0.277	1290	601
	200	0.000 0.000	0.002 0.004	0.004 0.009	0.033 0.036	0.073 0.090	1272	547
N=100,30°	-2.0 Gyr	0.000 0.000	0.019 0.023	0.000 0.000	0.007 0.009	0.000 0.000	748	1650
3:1 ratio	-0.6 Gyr	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.004 0.008	1385	914
	-0.2 Gyr	0.000 0.000	0.003 0.013	0.000 0.000	0.003 0.006	0.255 0.301	2014	432
	0.2 Gyr	0.087 0.100	0.111 0.283	0.091 0.163	0.181 0.135	0.425 0.313	2425	152
	1.7 Gyr	0.138 0.120	0.066 0.073	0.127 0.179	0.412 0.340	0.511 0.283	-422	321
	3.8 Gyr	0.127 0.113	0.028 0.031	0.080 0.116	0.534 0.262	0.349 0.262	259	184
N=100,60°	-2.0 Gyr	0.015 0.022	0.431 0.382	0.021 0.042	0.086 0.098	0.000 0.000	439	2856
3:1 ratio	-0.6 Gyr	0.000 0.000	0.070 0.075	0.002 0.007	0.007 0.014	0.000 0.000	729	1518
	-0.2 Gyr	0.000 0.000	0.004 0.017	0.004 0.011	0.011 0.029	0.179 0.188	1149	772
	0.2 Gyr	0.005 0.024	0.092 0.151	0.124 0.187	0.279 0.269	0.421 0.336	1434	240
	1.7 Gyr	0.225 0.319	0.134 0.197	0.193 0.225	0.409 0.306	0.175 0.436	-179	886
	3.8 Gyr	0.159 0.178	0.014 0.029	0.051 0.054	0.421 0.333	0.071 0.096	204	389
N=100,60°	-2.0 Gyr	0.009 0.028	0.078 0.079	0.003 0.007	0.034 0.066	0.000 0.000	490	2734
6:1 ratio	-0.6 Gyr	0.013 0.022	0.079 0.103	0.000 0.000	0.005 0.018	0.000 0.000	731	1364
	-0.2 Gyr	0.025 0.050	0.183 0.205	0.230 0.229	0.319 0.300	0.351 0.283	1183	573
	0.2 Gyr	0.084 0.084	0.258 0.182	0.238 0.245	0.267 0.264	0.101 0.104	1144	434
	0.6 Gyr	0.165 0.122	0.265 0.278	0.081 0.171	0.332 0.303	0.337 0.325	599	965
	2.2 Gyr	0.166 0.126	0.044 0.129	0.063 0.078	0.504 0.296	0.247 0.308	-108	891
N=100,-0.2Gyr	PA=0°	0.075 0.170	0.045 0.077	0.244 0.299	0.386 0.268	0.388 0.340	2437	63
6:1 ratio	10°	0.017 0.028	0.058 0.068	0.047 0.122	0.329 0.250	0.514 0.297	2465	120
	30°	0.005 0.008	0.148 0.147	0.058 0.075	0.214 0.263	0.511 0.313	2054	318
	45°	0.002 0.006	0.072 0.159	0.053 0.065	0.075 0.066	0.351 0.343	1705	480
	60°	0.025 0.050	0.183 0.205	0.230 0.229	0.319 0.300	0.351 0.283	1183	573
	80°	0.078 0.113	0.291 0.284	0.363 0.256	0.532 0.366	0.389 0.306	406	675
	90°	0.293 0.236	0.394 0.098	0.355 0.284	0.471 0.357	0.357 0.289	-15	690
Gaussian	N=30	0.473 0.295	0.451 0.282	0.480 0.299	0.491 0.286	0.527 0.289	...	...
	100	0.459 0.289	0.479 0.313	0.501 0.321	0.556 0.291	0.464 0.299	...	...
	140	0.504 0.313	0.515 0.312	0.467 0.340	0.527 0.317	0.540 0.284	...	...

NOTE.—The last two columns are the projected velocity and position difference between the subcluster and cluster, given in units of km s^{-1} and kpc, respectively. (See Table 7.)

longer be measurably asymmetric. These changes in "interpretation" cause many normality tests to be insensitive to velocity substructure.

The estimators of skewness and kurtosis, B_1 and B_2, seem fairly complementary in detecting a merger (Table 8). The alternative estimators, AI and TI, can best be compared using Figure 14. The value of each index is plotted with respect to the 5% and 10% significant levels. The tail index, TI, detects kurtosis with over 95% certainty for the two smallest angles. It appears *more* sensitive than B_2 for these cases. The asymmetry index, AI, exceeds only 90% certainty in detecting skewness and is comparable to B_1 in sensitivity.

The secondary statistics SS1, SS2, and SS3 reflect the side on which the velocity distribution's tail is long, whether the tails are longer or shorter than a Gaussian, and on which side the distribution is skewed, respectively. These statistics can only take integer values over a small range determined by the number of normality tests used to create them. Table 8 shows the average of these statistics over 10–20 files for each parameter set. We see by the size of the S_{BI} that there is a large scatter in these statistics for a given parameter set. The exceptions are cases with large sample size ($N = 200$) or extreme velocity separation ($T = -0.2$ Gyr, $\phi = 30°$). Here the statistics are the same for each file. The C_{BI} column shows that SS1, SS2, and SS3 also take their extreme values of -3 or 3, -2 or 2, and -1 or 1, respectively, for these extreme parameter sets. When applying SS1, SS2, and SS3 to a single data set, extreme values for all three reveals significant nonnormality. SS1 has a value of 3.0 in the extreme cases of our simulations because the velocities of our subcluster galaxies are higher than those of the primary, and the tails are longer on the side of the subcluster. Notice that SS2 indicates a short-tailed or boxy distribution when the velocity substructure is pronounced. Thus, the presence of a subcluster separated by $\leq 1.5\sigma_v$ will result in *shorter* tails than the best Gaussian fit. As the velocity separation increases, SS2 becomes positive, but only for the 6:1 simulations. The subcluster must be small and well separated to be interpreted as a long tail. Note that this would be difficult to distinguish from foreground or background contamination.

The effects of velocity "contamination" by a subcluster are tabulated in Table 9. "$V_{peculiar}$" is the difference in the global value of C_{BI} and C_{BI} for the primary cluster. If the kinematic centers of relaxed clusters were reliably marked by, say, a cD galaxy, then a nonzero systematic velocity of the cD would indicate the presence of substructure (Beers et al. 1991; Pinkney et al. 1993; Zabludoff et al. 1993; Bird 1994). Our results show that such a cD peculiar motion would be at least marginally significant in many of our cases. For example, if C_{BI} for the cluster has an error of 100 km s^{-1}, and the velocity of the cD has an error of 50 km s^{-1}, a velocity difference of 224 km s^{-1} (2σ) is marginally significant (using the formalism of Teague et al. 1990). This difference is not quite attainable for the 6:1 merger. The 3:1 merger, however, produces V_{pec} values as large as 490 km s^{-1} and gives marginal or better results even for angles of 60°. Consequently, if large-scale mergers are frequent, apparent cD peculiar velocities should be common.

Histograms of velocity distributions for many of the interesting parameter sets discussed above are shown in Figure 15. The test probabilities are given for the distribution shown, and the averages over ~15 data files drawn from the same parameter set are given in parentheses. One can judge by eye whether the addition of the subcluster (cross-hatched) has made the total distribution (open) non-Gaussian. The distributions in Figures 15*a* and 15*b* are drawn from simulations which have identical parameters except for the relative size of the subcluster. This pair demonstrates the difference in behavior of the kurtosis tests B_2, W, A^2, and skewav to changing subcluster size. Figures 15*c* and 15*d* demonstrate the same effect, but for a different velocity separation of the two components. Figures 15*e* and 15*f* are for time epochs before and after the core crossing. These demonstrate a difference in the dispersion in the substructure velocities before and after phase mixing has occurred.

In summary, the one-dimensional tests are not very effective at detecting substructure in our simulation data files. Relatively few detections (i.e., significance levels less than 0.1) appeared. The parameters (especially sample size) were chosen to probe marginal significance levels of the more sensitive two- and three-dimensional tests. Nevertheless, when the velocity separation between the cluster and subcluster exceeded ~2000 km s^{-1}, the nonnormality becomes significant for redshift survey sizes of $N \sim 100$. (Compare to Zabludoff et al. 1990, who do not find that the KS test can distinguish $N \sim 100$–200 data sets at 5% significance.) We find that averaging the result from many tests is not useful. Many normality tests will show a decrease in significance with an increase in subcluster size. V_{pec} complements this behavior nicely by becoming more significant with an increase in subcluster size. It appears more likely for mergers to manifest themselves as boxy velocity distributions than long-tailed ones. Measures of skewness are especially sensitive to our two-body mergers.

5.3. *Individual Tests: Two-dimensional Estimators*

Selected results are plotted for the two-dimensional tests in Figures 16–21. Again, we show the behavior of the tests as a function of time, projection angle, sample size, and mass ratio. In § 5.1, 100 data files were used for each probability centroid. Here 15 data files are used in the case of $N = 60$, 80, or 100 sample sizes; 20 data files for $N = 30$ and 45; 10 data files for $N = 140$; and six data files for $N = 200$.

5.3.1. *The Angular Separation Test*

The angular separation test (hereafter AST) was introduced by West et al. (1988) for the analysis of their *N*-body simulations. The algorithm is based on the premise that substructure should create an excess of small angular separations compared to a Poisson distribution. The "angular separation" between galaxies *i* and *j* is the angle between the rays extending from the cluster centroid to the two galaxies. The harmonic mean of the angular separations is calculated to be

$$\theta_{hm} = \left[2/N(N-1) \sum_{i>j} \theta_{ij}^{-1} \right]^{-1}, \quad (3)$$

where the sum is over all $N(N-1)$ pairs, and the angular separation between galaxies i and j, θ_{ij}, is measured consistently in the same sense (e.g., counterclockwise), so that the angles can range between 0 and 2π. The test statistic is formed by

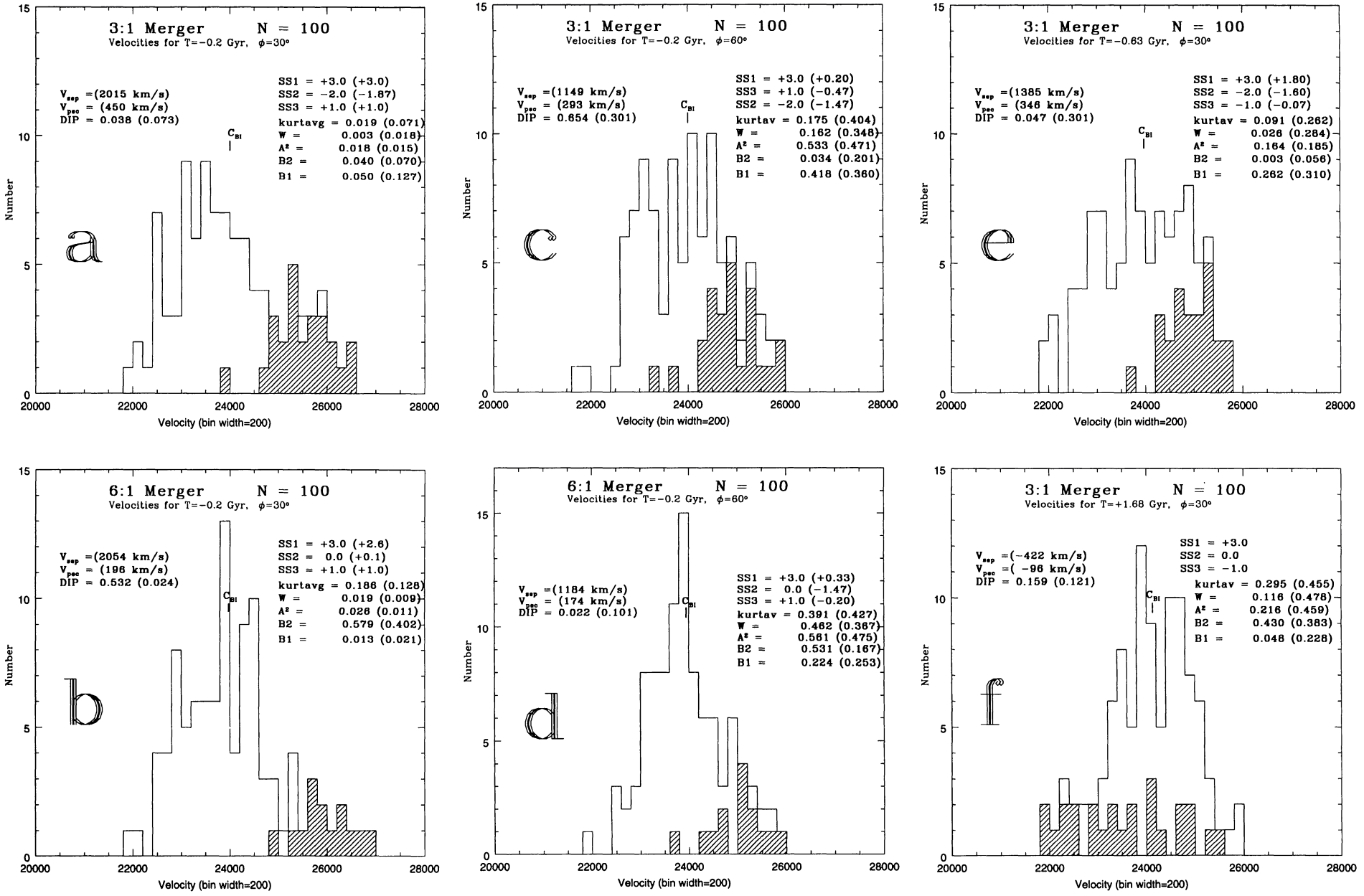

FIG. 15.—Sample radial velocity histograms taken from the simulation data files. The cross-hatched histogram is the subcluster alone, while the open histogram is the combined sample. (See text.)

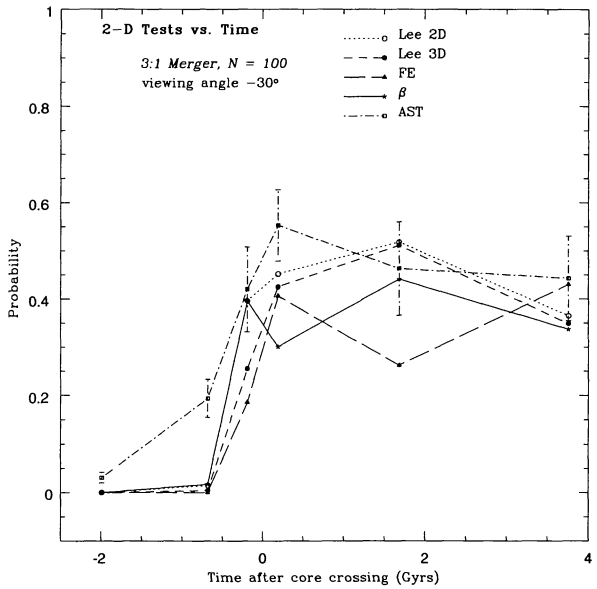

FIG. 16.—Significance levels of two-dimensional substructure estimators applied to the 3:1 merger simulation at six epochs. The projection angle is 30°.

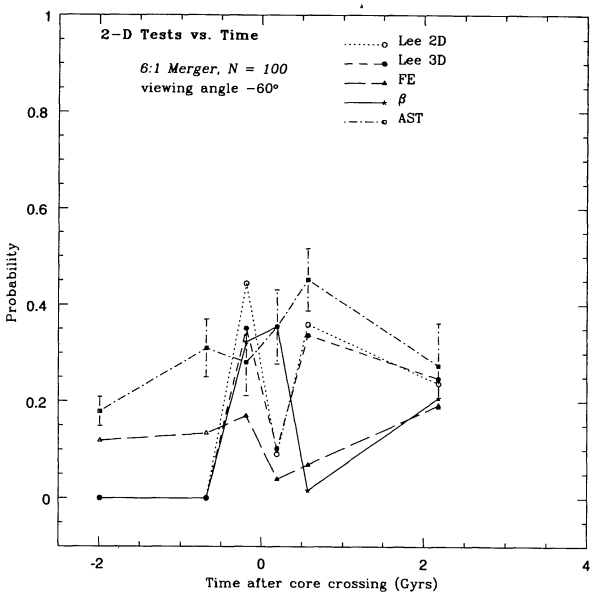

FIG. 18.—Significance levels of two-dimensional substructure estimators applied to the 6:1 merger simulation at six epochs. The projection angle is 60°.

dividing θ_{hm} for the actual distribution by the angular harmonic mean for a Poisson distribution with the same N, i.e., AST = θ_{hm}/θ_P. This ratio will be about 1.0 for substructure-free clusters, and AST < 1.0 for clumpy distributions. The θ_P were calculated by averaging 1000 Monte Carlo simulations at sample sizes between 20 and 500, and by interpolating between these values. The Monte Carlo data sets were created using the azimuthal randomization technique discussed in § 4.3.

West et al. made two refinements to this test. First, they examine AST as a function of radius by grouping the data in two radial bins, one for points less than the half-mass radius and the other for points greater than the half-mass radius. Second, they rejected all θ_{ij} that were less than 1% of the mean interparticle separation $[0.01(2\pi/N)]$. This improved the stability of the test because very small angles produced large fluctuations in AST.

There were some minor differences between our implementation of AST and that of West et al. We did not incorporate

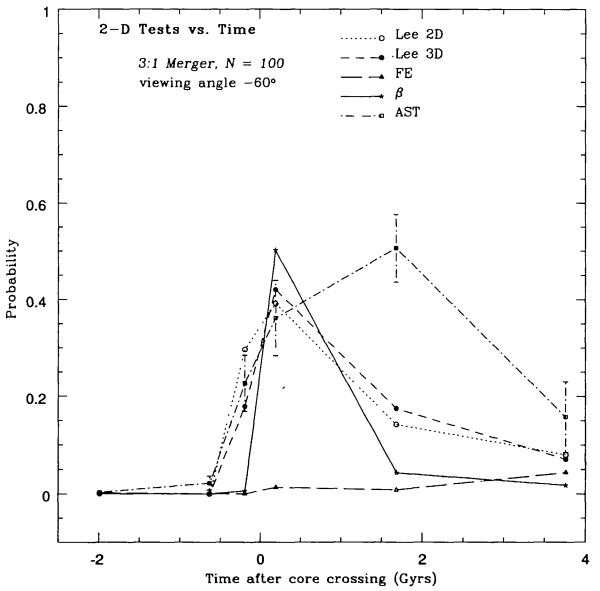

FIG. 17.—Significance levels of two-dimensional substructure estimators applied to the 3:1 merger simulation at six epochs. The projection angle is 60°.

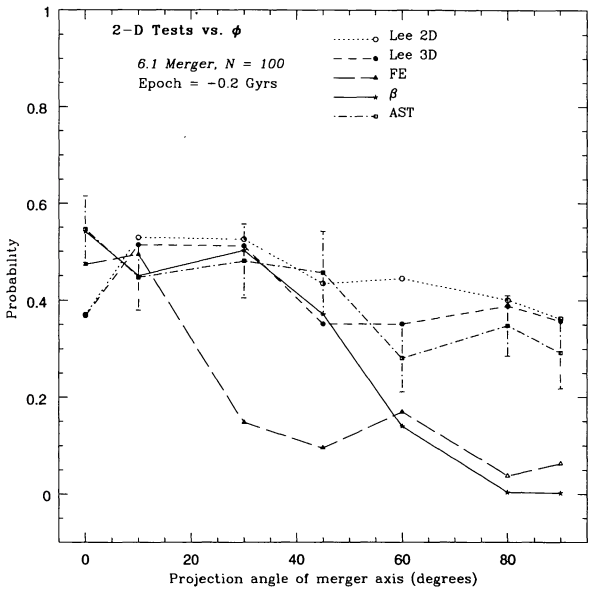

FIG. 19.—Significance levels of two-dimensional substructure estimators applied to the 6:1 merger simulation for several projection angles of the merger axis. The epoch is −0.19 Gyr.

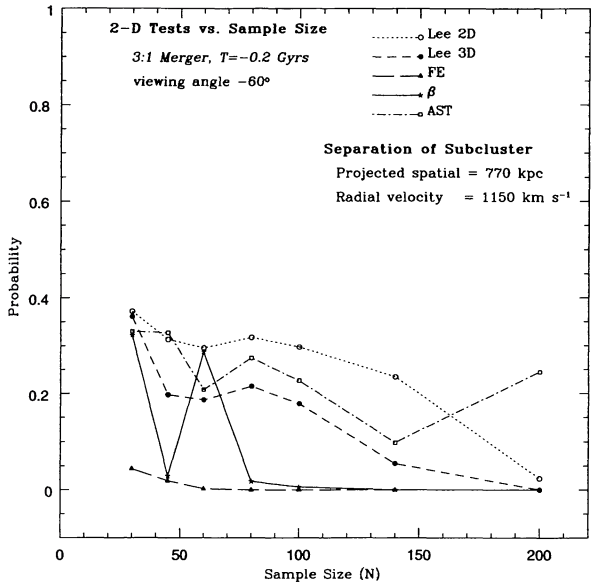

FIG. 20.—Significance levels of two-dimensional substructure estimators applied to the 3:1 merger simulation for various sample sizes. The projection angle is 60°.

the radial binning of data in our algorithm. However, we did use the small angle filter. Another minor difference between our code and the algorithm above is that we defined our angle to be the smaller of the two possible for each pair of points. Thus, θ_{ij} ranged between 0° and 180°. Finally, we used azimuthal normalization to create 300 Monte Carlo simulations for *each* data file. This way, the sample sizes are matched and interpolation is unnecessary. We show the distribution of θ_{ij} (in degrees) for 500 Monte Carlo simulations in Figure 4.

West et al. tested AST on various substructure scenarios. One would expect the test to respond to substructure-free distributions which are simply elongated. They found that θ_{hm} did not give values significantly lower than 1.0 until ellipticities were greater than 0.5. They also generated distributions containing from one to five small (e.g., 10% of the total mass), dense subclusters and found the test to be sensitive to them. West et al. point out that the test may be relatively insensitive because there is a loss of information going from two spatial dimensions to one angular dimension.

Our results for AST are given in Figures 16 to 21 and Table 9. Figures 7 and 9 can be inspected to see how the test responds in substructure-free conditions. We see that AST has the expected mean probability of 0.5 and the expected false positive rates. For the merger cases, AST proves to be very insensitive to our substructure. Only one case rejects the Poisson hypothesis at greater than the 99% level; for $T = -2.0$ Gyr (the earliest epoch inspected). A subcluster of a given surface density will create more small angles if located at a large radius than at a small radius. This explains why AST responded to the epochs in which the subcluster has a projected spatial separation of over 1200 kpc. It produced marginally significant results for three additional cases. All four cases occurred for the 3:1 ratio; the $\frac{1}{6}$ size subcluster did not produce even a marginally significant detection.

AST would be a more sensitive diagnostic in other scenarios. First, if the subclusters were more compact, they could be detected at smaller radii. Second, the system could contain more than one subclump. Since many of our tests rely on bimodality for a strong detection (e.g., the Lee statistic), AST may actually excel as a sensitive diagnostic for clumping in multiple systems.

5.3.2. *The β Test*

The β, or *symmetry,* test was introduced by West et al. (1988). It tests for significant deviations from mirror symmetry about the cluster center. The first step of the algorithm is to find the mean distance from each galaxy, i, to its five nearest neighbors. Next, the mean distance from a point diametrically opposite each galaxy to its five nearest neighbors is found. The asymmetry for galaxy i is given by

$$\beta = \log_{10}\left(\frac{d_o}{d_i}\right), \tag{4}$$

where d_o is the local density diametrically opposite galaxy i. The average value of β over all galaxies, $\langle\beta\rangle$, is the actual β statistic. Finally, West et al. calculated β in the same two radial bins used for the angular separation test. For symmetric distributions, β is ≈ 0. Typical values for random distributions are $\pm 10^{-2}$.

Azimuthal randomization of the input data set was used to create files for normalization. This was recommended by West et al. over random draws from an $r^{1/4}$ or Hubble profile. They generated 60 spatial distributions for comparison with each. We made a few changes to the original algorithm. We calculated β for the entire sample, not in two radial bins. We defined our cluster centroid as the mean of x and y. We used 300 azi-

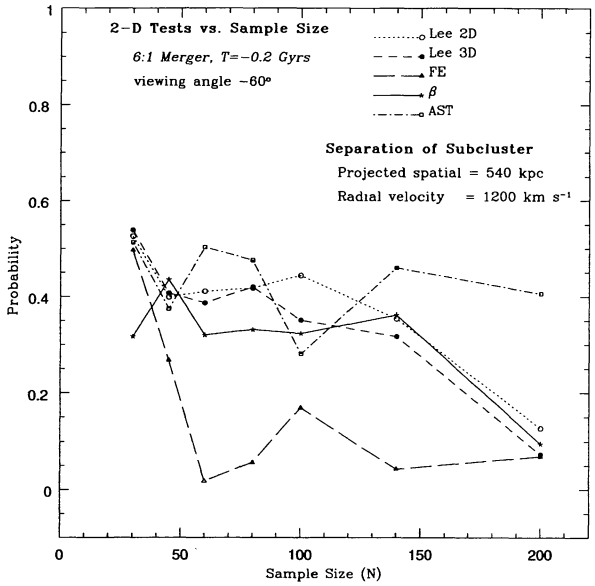

FIG. 21.—Significance levels of two-dimensional substructure estimators applied to the 6:1 merger simulation for various sample sizes. The projection angle is 60°.

muthally randomized Monte Carlo simulations instead of 60. Instead of using five nearest neighbors for all sample sizes, we used $N^{1/2}$. Also, for convenience, we multiplied β by 1000 (see Fig. 4).

The β test behaved as expected from the algorithm. β is sensitive to deviations from *mirror* symmetry, but not necessarily *circular* symmetry. Consequently, this test will not mistake an elongated but smooth cluster for one containing substructure (i.e., clumpiness). In Figures 16–18 we see that β finds significant asymmetries (probabilities $\lesssim 1\%$) for the first two epochs of merger. During these epochs, the subcluster is well separated and distinct, especially for the 60° projection angle. For epochs near core crossing, the spatial distribution of galaxies shows elongation but only slight asymmetry, so the significance of β is diminished. Nevertheless, β responds to the slight asymmetry at $T = -0.2$ Gyr (Figs. 19 and 20); the probability decreases as ϕ approaches 90° and as N increases. After core crossing, there is asymmetry due to the dispersed population on the right side of the cluster (see Fig. 3). β reveals the asymmetry to be significant at 60° but not at 30°. The 3:1 merger gives a stronger signal than the 6:1; however, β will become ineffective as the mass ratio approaches 1:1 because the centroid will be midway between the two components. The figures show that β is more sensitive to substructure than all other two-dimensional tests, except the FE test.

5.3.3. *Fourier Elongation Test*

The elongation of a cluster is not a definitive sign of substructure; however, it is a signature. Some authors interpret deviations from circular symmetry in the X-ray surface brightness of clusters as substructure (Mohr, Fabricant, & Geller 1993). Jones & Forman (1992) include elliptical morphologies in their estimate of the occurrence rate of substructure in X-ray clusters. Clusters with substructure will, in most cases, have an elongated distribution of galaxies. Therefore, it is desirable to say how significantly elongated a cluster is compared to a circularly symmetric distribution. It should be noted, however, that the galaxy distribution may be more elongated than the cluster's gravitational potential (Buote & Canizares 1992).

In this spirit, we have taken one of several available estimators of elongation and transformed it into a statistical test. We use the method of Fourier analysis which was applied to clusters by Rhee, van Haarlem, & Katgert (1991). These authors define the cluster centers as the brightest galaxy within 0.5 Mpc of a nominal cluster center or, if this was problematic, the density peak. Using this center, they assume that the azimuthal galaxy distribution, $N(\phi)$, resembles the model

$$N(\phi) = \left(\frac{N_0}{2\pi}\right)[1 + (N_1/N_0)\cos(2\phi - 2\phi_0)], \quad (5)$$

where N_0 is the total number of galaxies in the cluster, N_1 is the elongation amplitude, and ϕ_0 is the position angle of the cluster. We use the elongation strength as our statistic, defined as

$$FE = \frac{N_1}{\sqrt{2N_0}} = \frac{2(S^2 + C^2)^{1/2}}{\sqrt{2N_0}}, \quad (6)$$

where

$$S = \sum N(\phi)\sin(2\phi), \quad C = \sum N(\phi)\cos(2\phi). \quad (7)$$

We use the $C_{\rm BI}$ of the spatial distributions to define our centroid. Finally, we normalize FE by applying it to 300 azimuthal randomizations of the original data set (as with AST and β).

Figures 16–21 and Table 9 show how FE responds to substructure more readily than all other two-dimensional tests. FE excels especially at detecting the deformation of the cluster just before and after core crossing ($T = \pm 0.19$ Gyr). It is so sensitive that a 15% difference in the x and y widths of our Gaussian toy clusters doubled the FP counts for $N = 100$ and tripled them for $N = 140$. The FE test does not give probabilities less than 0.1 in all cases; no elongations were significant for projection angles $\leq 30°$.

Since elongation is a signature of other formation scenarios besides hierarchical mergers, further criteria should be used with FE before it is used as a substructure diagnostic. We recommend these criteria: FE must reject the circular hypothesis at the 99% level ($P \leq 0.01$), *and* FE must be greater than 2.5. Applying this criterion, the number of "detections" in our simulations goes from 24 down to 15 cases out of the 36 in which we know a subcluster is present.

5.3.4. *The Lee Statistic*

The Lee (1979) statistic was introduced as a means of testing for bimodality in distributions. Its astronomical applications were investigated by Fitchett (1988b), and it was used to demonstrate the presence of substructure in the Coma Cluster by Fitchett & Webster (1987).

The Lee statistic uses the maximum likelihood technique to separate a data set of two or higher dimensions into two parts. For a set of N points distributed in the plane, the algorithm begins by projecting the points onto a line. The line makes the angle ϕ with respect to a second line, which we take to have constant declination. The first line is rotated by small increments for $0 \leq \phi \leq 180°$. For each orientation, the points are projected onto the line so that each point now has a new coordinate, x_j, along that line. A search for the best partition into a "left" and "right" clump is then made. For all $N - 1$ possible partitions, the quantities σ_l, σ_r, and σ_T are calculated for the left, right, and total samples, respectively. The σ values are given by $\sigma = \sum (x_j - \mu)^2$ (not a standard deviation), where $\mu = \sum x_j/n$. Next, the quantity L is defined to be

$$L = \max_{\rm partitions}\left(\frac{\sigma_T}{\sigma_l + \sigma_r} - 1\right). \quad (8)$$

L will vary with ϕ such that it is large when the projection axis connects two, distinct clumps in the dataset. Two useful statistics can be formed from L: $L_{\rm max}$ and $L_{\rm rat}$. $L_{\rm max}$ is simply the maximum L over all ϕ, while $L_{\rm rat}$ is the ratio of $L_{\rm max}$ to $L_{\rm min}$. Fitchett (1988b) has shown that $L_{\rm rat}$ is a more powerful statistic than $L_{\rm max}$ for small N. They perform similarly for large N. Consequently, we have chosen to use only $L_{\rm rat}$ in this study.

We experimented with the choice of increments for ϕ. One

hundred subdivisions of 180° have typically been used (e.g., Bird 1993). This provides a smooth plot of L versus ϕ. However, we found that 30 subdivisions (6° increments) will find the same optimum orientation angle, and will provide a statistic $L_{\rm rat}$ which is negligibly different. Computing the Lee statistic with 100 subdivisions is more time consuming than any other two- or three-dimensional statistic. Thus, we adopted 30 subdivisions to save time.

Fitchett & Webster (1987) used a problematic method for calculating the significance of $L_{\rm rat}$. As usual, pseudocluster data sets were produced randomly for comparison. However, the spatial distributions were randomly drawn from one of two radial surface density profiles:

$$\text{case A: } \sigma = \sigma_0 r^{-1} \quad \text{(power law)}, \tag{9}$$

and

$$\text{case B: } \sigma = \sigma_0 (r^2 + r_c^2)^{-1/2}. \tag{10}$$

They found that the data (Coma Cluster) appeared twice as significant when compared to case A than case B, presumably because the constant density core of case B can produce a chance clumping more readily than case A. Indeed, any profile steeper than the actual data will produce an overestimate of the significance.

We resorted to azimuthal randomization to produce null cases. This way, each pseudocluster will have the same surface density profile as the original data set. This method produced the expected false positive rates (§ 5.1).

We refer to our adaptation of the Lee statistic as "Lee 2D" in the figures and tables. Lee 2D is not very sensitive to the substructure in our merger simulations, despite the fact that they contained two components. In particular, a projected spatial separation of over 900 kpc was needed for bimodality to be marginally detected in an $N = 100$ sample. There is one counterexample in the 6:1 simulation at $T = 0.2$ Gyr. For $T = -0.2$ Gyr, when the subcluster is only separated by $\sim$760 km s^{-1}, only the 3:1 merger with $N = 200$ and $\phi = 60°$ showed significant bimodality. Typically, the significant detections by Lee 2D were for the early epochs in which the separation between components is obvious. As with AST, a more compact subcluster should be more detectable. Lee 2D is a conservative test; it will not see elongations as substructure, and it will lose sensitivity when more than two subclusters are present or when the size of the two clumps becomes dissimilar.

5.4. *Individual Tests: Three-dimensional Estimators*

In this section, we discuss the results for substructure tests which analyze all three dimensions available to observers: right ascension, declination, and radial velocity. All five of the tests discussed here have the null hypothesis of constant mean velocity and velocity dispersion as a function of position. As discussed in § 5.1, a cluster with a velocity dispersion gradient does *not* satisfy the null hypothesis. The two components in our merger simulation were modeled as isothermal King spheres with decreasing velocity dispersion with radius. This may lead to an enhancement of significance in these tests. The results in the plots to follow should be compared with Figures 8 and 10.

Each of the three-dimensional tests (except Lee 3D) calculate a statistic both *locally* and *globally,* and then compare the two to quantify substructure. "Local" refers to one galaxy and its $N_{\rm nn}$ nearest neighbors in the plane of the sky, or in velocity space. For the Δ and α tests, $N_{\rm nn}$ was originally chosen to be 10, irregardless of sample size, N. This can be problematic in cases of small N. In developing the ϵ test, Bird (1993) chose $N_{\rm nn}$ to be the greatest integer not exceeding $N^{1/2}$. This was based on attempts by Silverman (1986) to measure local densities in data sets of comparable size to cluster data sets. He found that $N^{1/2}$ nearest neighbors is optimum to avoid smoothing over substructure and for robustness against nonsignificant random fluctuations. We decided to adopt $N_{\rm nn} = N^{1/2}$ for the other three-dimensional tests as well.

5.4.1. *The Lee Three-dimensional Statistic*

Fitchett & Webster (1987) discuss the extension of the two-dimensional Lee statistic (§ 5.3) to three-dimensional (x, y, V) data. They recommend treating the velocity as a third spatial dimension and then rotating the projection axis over two angles, θ and ϕ. In practice, the units of x, y, and V should be rescaled so that they have comparable ranges; however, the statistics are invarient over linear transformations. For normalization, they recommend Monte Carlo simulations which draw random velocities from a Gaussian and random positions from case A and B (above). Fitchett (1988b) finds that $L_{\rm max}$ is the optimal statistic for this test.

We were reluctant to implement this algorithm for several reasons. First, if 30 increments in angle were used for the loops over θ and ϕ, this test would be about 30 times more CPU intensive than Lee 2D. Second, it would require a special means of normalization, not used for the other three-dimensional tests. Third, as discussed in § 5.3, the normalization data sets should not be drawn from distributions which are different from the original.

We have developed an alternative algorithm which utilizes velocities. It begins with the Lee 2D algorithm but uses velocity dispersion weighting to calculate L. The idea is that the velocity dispersion of the "left" and "right" subsamples should be minimized for the cleanest separation into two subclusters. The projection axis is only rotated in ϕ space, as for Lee 2D. For each ϕ, the data is projected on the axis. The number L is now defined as (see eq. [8])

$$L = \max_{\rm partitions}\left(\frac{2\sigma_T}{\sigma_l + \sigma_r + \sigma_{vl+vr}} - 1\right), \tag{11}$$

where

$$\sigma_{vl+vr} = \frac{\sigma_T}{4\sigma_v}(\sigma_{vl} + \sigma_{vr}), \tag{12}$$

the σ values are still defined as $\sum (x_j - \mu)^2$, and the v subscripts refer to velocities. The term in front of equation (12) scales the velocity "variances" to the spatial variances. The factor of.4 is

needed because, on average, the spatial variance of the left and right subsamples is about $\frac{1}{4}$ of the total spatial variance (the sample is split in two), while the velocity variance of the left and right subsamples is about the same as the total velocity variance. Finally, the Lee 3D statistic, analogous to Lee 2D, is $L_{rat} = L_{max}/L_{min}$, where the maximum and minimum are found over the 30 projection angles.

Figure 4 shows that the distributions of Lee 2D and Lee 3D (for the same 500 simulated data files) are different. Both the Lee 2D and Lee 3D statistics are bounded by 1.0 on the low end, but Lee 3D has a smaller centroid and scale in all cases.

The sensitivity of Lee 3D to bimodality in our simulations is a little higher than Lee 2D in nearly all cases. The probability curves for Lee 3D and Lee 2D are nearly parallel in plots such as Figure 18, indicating the same sensitivity. It would be expected that scenarios with large velocity separations between the two components would give Lee 3D the greatest improvement in sensitivity. In fact, it seems that an increase in the mass ratio from 6:1 to 3:1 causes a greater improvement (compare Figs. 20 and 21). Figure 19 shows that Lee 3D does not improve much over Lee 2D as ϕ approaches 0°, at least not for the 6:1 mass ratio. This behavior is probably caused by the inevitable overlap of the primary and subcluster particles for $\phi <$ 30°. The best partition along the axis of projection cannot prevent high-velocity particles from the subcluster from mixing with those of the primary, and vice versa. The number of marginal or better detections out of 36 total was only one better than Lee 2D. Compared to the other three-dimensional tests, Lee 3D is *insensitive*. However, Lee 3D (like Lee 2D) has the advantages of insensitivity to nonsubstructure configurations which appear as substructure to other tests (e.g., elongation, velocity dispersion gradient). The code for both Lee tests can also output the position angle of the axis joining likely subsystems and allocate the data into two groups objectively (Fitchett 1988b).

5.4.2. *The Δ Test*

The Δ test (also "DS test") was introduced by Dressler & Shectman (1988, hereafter DS) to find substructure in their sample of 15 clusters.

The algorithm begins by calculating the velocity mean and standard deviation for each galaxy and its N_{nn} nearest neighbors. Next, these local means and σ values are compared with the global mean and σ (using all galaxies with velocities). A deviation from the global values can then be defined for galaxy i as

$$\delta_i^2 = \left(\frac{N_{nn}+1}{\sigma^2}\right)[(\bar{v}_{local}-\bar{v})^2 + (\sigma_{local}-\sigma)^2]. \quad (13)$$

Finally, the cumulative deviation Δ is defined as $\sum \delta_i$ and serves as the statistic for quantifying the substructure. The value of Δ is $\sim N$ for samples with no substructure (see Fig. 5). The test was originally normalized with 500–1000 Monte Carlo simulations where the velocities are shuffled among the positions (see § 4.3). It is also typically presented with a "bubble skyplot" (see DS and Fig. 22). These skyplots represent galaxies by circles whose radii are proportional to e^{δ_i}. The purpose of this plot is to reveal the likely location of the subcluster. An example is given in Figure 22 for a draw of 100 particles from the 6:1 mass ratio simulation at $T = -0.2$ Gyr, $\phi = 60°$. The clump of the largest circles coincides nicely with the subcluster. The only differences in our adaptation of the Δ test are the use of $N_{nn} = N^{1/2}$, rather than $N_{nn} = 10$, and the use of 300 simulations.

Two cases of insensitivity can be deduced by looking at equation (13). The first was mentioned by DS: the Δ test is insensitive when the subclusters are superimposed. The Δ test is also insensitive when the subclusters have the same dispersion and velocity mean, e.g., if they have equal mass and their merger axis is in the plane of the sky. The occurrence of both of these situations requires projection angles which are rare.

DS also explored the sensitivity of the test to (1) uncertainties in $\bar{v}$ and σ, and (2) presence of velocity outliers. The $\bar{v}$ and σ used in the real data sets and the Monte Carlo shuffles were perturbed by numbers of order the error in these quantities (e.g., 8%–20% for σ, 75–100 km s^{-1} for $\bar{v}$). The resulting Δ remained significant for the strongest cases, but it could become both more and less significant for clusters that were marginal to begin with. To test sensitivity to outliers, DS removed the two extreme velocities from each end of the distribution and reran the test. This typically changed the global dispersion by 11%. Again, borderline cases would change their significance rating, but extreme cases remained significant. DS concluded that the Δ test was resistant to outliers and robust to parametric uncertainties.

Our results for the Δ test and other three-dimensional tests are found in Figures 23–28, and Table 11. We find Δ to be the most sensitive of *all* the tests to the substructure in our simulations (Table 15). Moreover, the probability curves (Figs. 23–28) remained less than 0.40, indicating that Δ was responding to deviations from the null hypothesis in all cases with substructure (see Fig. 10).

The Δ test rejected the null hypothesis with ≥99% significance in the early epochs of both merger simulations (Figs. 23 and 24). After core crossing, there is a distinct rise in probability centroid due to the loss of spatial separation. The post-merger probability centroids are ~0.1, meaning that a random draw may have a marginally significant result.

The sample size dependence makes the detection of substructure in small samples look hopeful. Figure 27 shows marginal detections down to the $N = 30$ limit for the 3:1 merger (velocity separation ~ 1100 km s^{-1}, spatial separation ~ 780 kpc). But $N \geq 60$ was required for a marginal detection of the 6:1 merger (similar subcluster separation).

Also of interest is the dependence of Δ on the projection angle of the merger axis (Fig. 26). In this epoch (−0.2 Gyr), there is a large true velocity separation of about 2500 km s^{-1} between the two components and a fairly small true spatial separation of 700 kpc. Although projection angles near 0° and 90° are both expected to be problematic, only 90° seriously reduced the sensitivity of Δ. At 0°, corresponding to perfect superposition, the Δ test gives marginal significance. This must be caused by the smaller core radius of the subcluster. It may also be caused partly by the physics of the simulation: the subcluster's particles are *channeled* into the potential of the primary cluster, thereby decreasing its apparent radius when viewed at 0°.

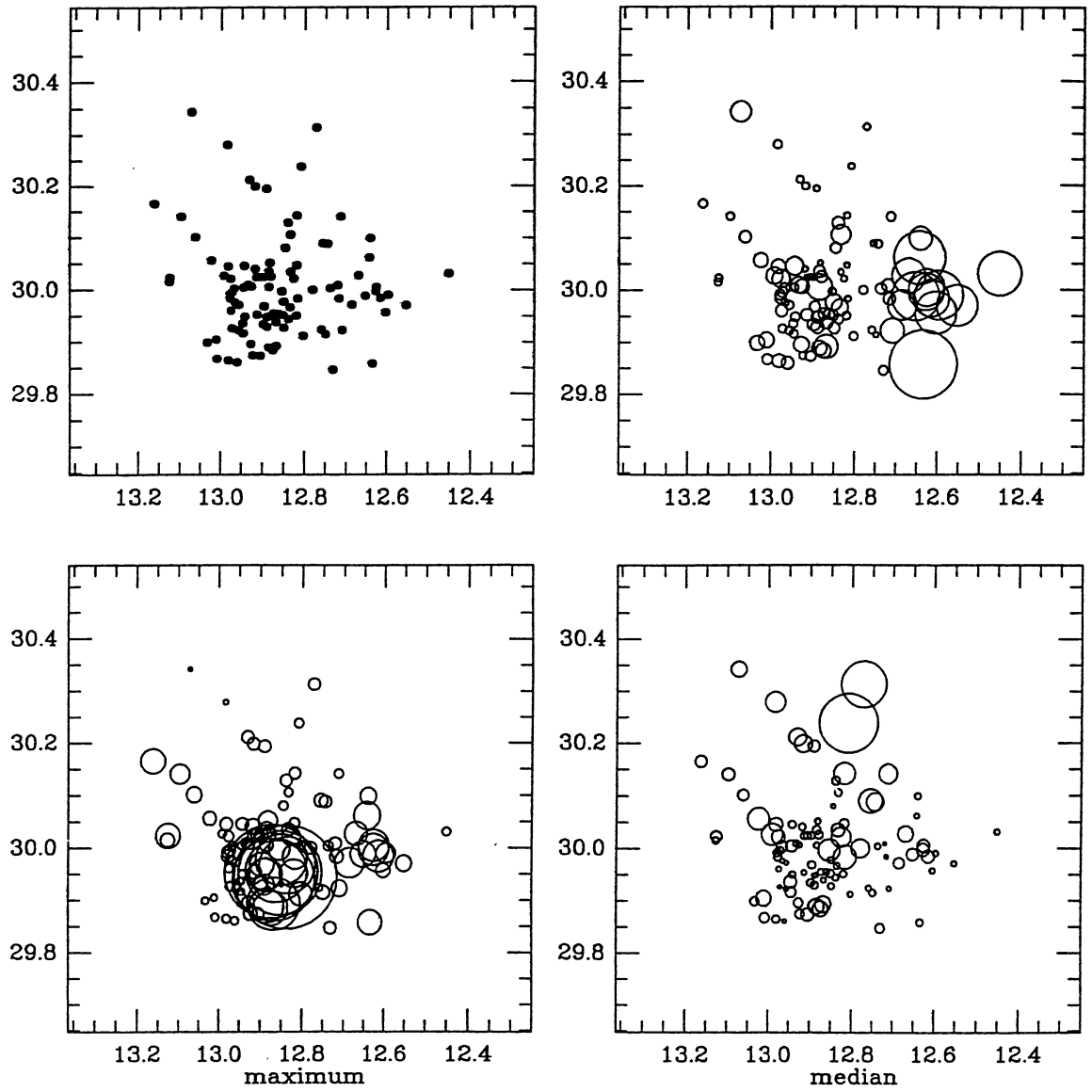

FIG. 22.—Bubble plot for the Δ test. The upper left panel shows the spatial distribution only. The upper right panel represents "galaxies" by circles whose radii are proportional to e^{δ}, (see text). A circle is large when the velocity mean and/or standard deviation of the galaxies near it have a pronounced difference from the global values. The lower panels are the same, but for Monte Carlo simulations of input data file. The median Δ (*right*) and maximum Δ (*left*) cases are shown out of 500 simulations. This sample contains substructure significant at the 2% level.

5.4.3. *The ϵ Test*

The ϵ test was created by Bird (1993). It is distinguished as the first nearest neighbor group (NNG) test in which the number of nearest neighbors is a function of sample size. It quantifies correlations between position and the projected mass estimator, defined as

$$M_{\rm PME} = \xi\left(\frac{24}{\pi GN}\right)\sum_{j=1}^{N_{\rm nn}} v_{zj}^2 r_j \qquad (14)$$

(see Heisler, Tremaine, & Bahcall 1985). For the ϵ test, v_{zj} is a radial peculiar velocity with respect to the NNG average velocity ($C_{\rm BI}$), r_j is its projected distance from the NNG center, and ξ is a constant which takes the value of 4/3 for isotropic orbits. Since the center of the NNG is defined as galaxy i, r_i is 0, and the central galaxy does not contribute to $M_{\rm PME}$. The substructure statistic is then defined as

$$\epsilon = \frac{1}{N_{\rm gal}}\sum_{i=1}^{N} M_{\rm PME}\,. \qquad (15)$$

Thus, ϵ is "the average mass per NNG." Its units are solar masses; however, it does not provide an accurate estimate of the cluster's total mass. Since the galaxies in NNGs have small

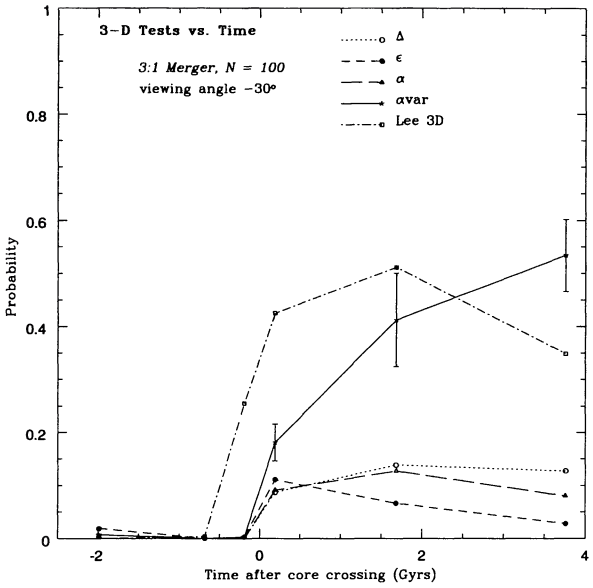

FIG. 23.—Significance levels of three-dimensional substructure estimators applied to the 3:1 merger simulation at six epochs. The projection angle is 30°.

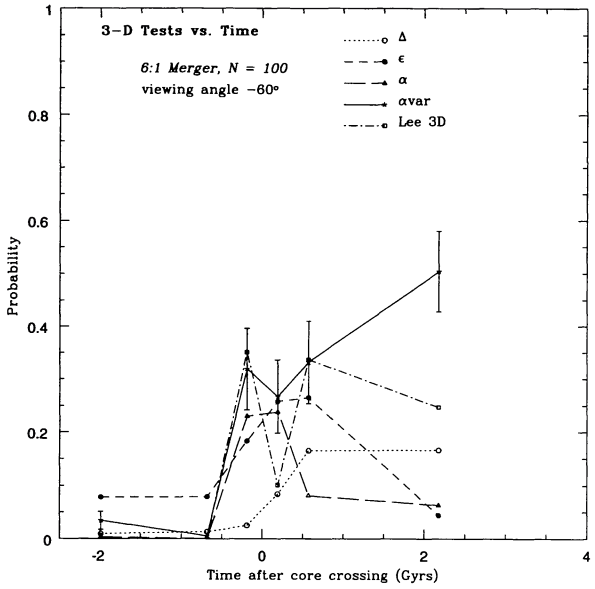

FIG. 25.—Significance levels of three-dimensional substructure estimators applied to the 6:1 merger simulation at six epochs. The projection angle is 60°.

projected separations, ϵ is generally smaller than global mass estimates. The test is normalized in the same way as Δ; the velocities are shuffled randomly among the positions. This erases any strong correlations between dispersion and position. Here ϵ will be lower for a cluster with substructure than the Monte Carlo simulations of that cluster.

Bird (1993) describes the ϵ test as sensitive to mergers without redshift separation, but not as sensitive as the Δ test. Our results support these assertions. Recall that ϵ reacted most strongly to dispersion gradients in our isothermal clusters (Fig. 8). The ϵ test finds marginal or better significance in 20 out of 36 cases studied. This is not as sensitive as the Δ or α tests, but it is more sensitive than the α var and two-dimensional tests. Figures 23 and 24 show the conditions under which this test actually dominates over Δ; the last two epochs of both the 3:1 and 6:1 mergers. In these epochs, the subcluster has passed through the primary and appears as a dispersed spray of particles on the other side of the cluster. Some velocity and spatial

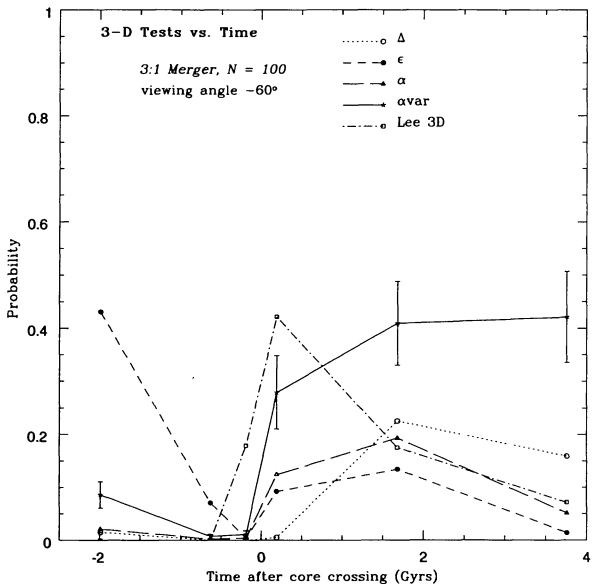

FIG. 24.—Significance levels of three-dimensional substructure estimators applied to the 3:1 merger simulation at six epochs. The projection angle is 60°.

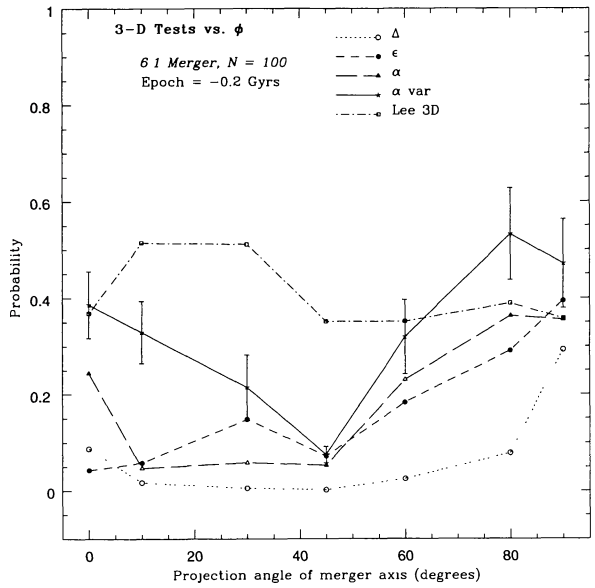

FIG. 26.—Significance levels of three-dimensional substructure estimators applied to the 6:1 merger simulation for several projection angles of the merger axis. The epoch is −0.19 Gyr.

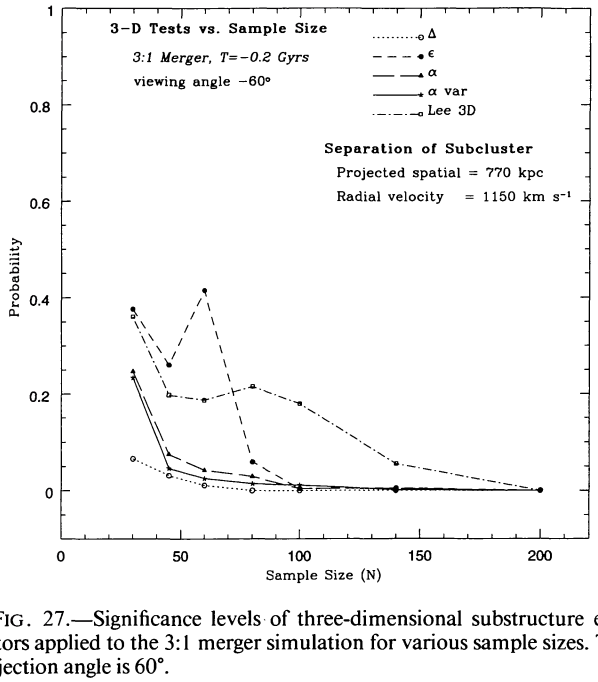

FIG. 27.—Significance levels of three-dimensional substructure estimators applied to the 3:1 merger simulation for various sample sizes. The projection angle is 60°.

separation exists in these 30° and 60° views, but the cause of small $M_{\rm PME}$ estimates must be the large projected separations, r_j, between galaxies in NNGs. $M_{\rm MPE}$ increases with r_j; however, the $M_{\rm MPE}$ for the Monte Carlo simulations (which have the same spatial distribution) will be much greater with interspersed velocities from the primary. This explanation is supported by Figure 26, showing the dependence of probability on projection angle. It is apparent that ϵ is not sensitive for $\phi > 45°$ in the 6:1 merger during an epoch in which the subcluster is not yet dispersed (-0.2 Gyr, $N = 100$). Thus, a velocity separation of the subcluster is important for ϵ. Note that ϵ would not be as sensitive in the late epochs of merger if there were a spatial cutoff in the sample (§ 4.4).

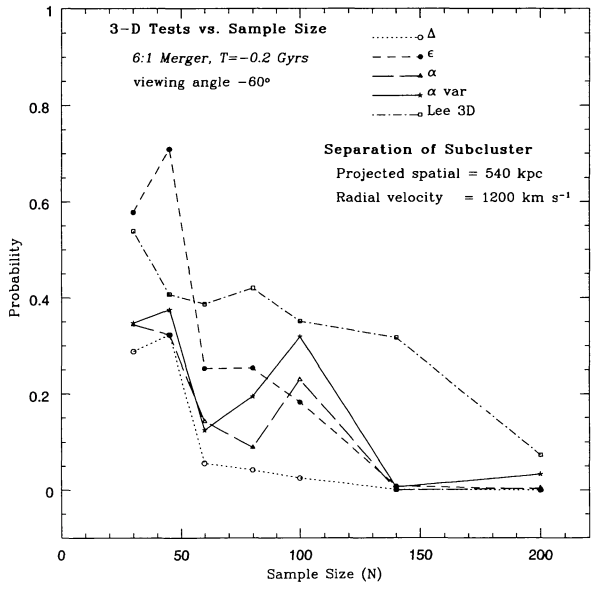

FIG. 28.—Significance levels of three-dimensional substructure estimators applied to the 6:1 merger simulation for various sample sizes. The projection angle is 60°.

5.4.4. *The α Test*

The centroid shift or α test was introduced by West & Bothun (1990, hereafter WB). It is a measure of how much the centroid of the galaxy distribution shifts as a result of correlations between local kinematics and the projected galaxy distribution. First, one must find the centroid of the two-dimensional galaxy distribution;

$$x_c = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad y_c = \frac{1}{N}\sum_{i=1}^{N} y_i . \tag{16}$$

Second, a weight is assigned to each galaxy i, equal to $w_i = 1/\sigma_i$, where σ_i is the line-of-sight velocity dispersion for galaxy i and its $N_{\rm nn}$ nearest neighbors in projection (originally taken to be 10). Third, for each galaxy i and its $N_{\rm nn}$ nearest neighbors in *velocity* space, calculate a spatial centroid

$$x'_c = \frac{\sum_{i=1}^{N_{\rm nn}+1} x_i w_i}{\sum_{i=1}^{N_{\rm nn}+1} w_i}, \quad y'_c = \frac{\sum_{i=1}^{N_{\rm nn}+1} y_i w_i}{\sum_{i=1}^{N_{\rm nn}+1} w_i} . \tag{17}$$

The weights ensure that galaxies in regions with small velocity dispersions have more influence on the centroid. Fourth, the difference in centroid of these nearest neighbor velocity groups and a global, unweighted centroid is calculated to be

$$\gamma_i = [(x_c - x'_c)^2 + (y_c - y'_c)^2]^{1/2} . \tag{18}$$

Finally, the statistic α quantifies the overall substructure as an average of the γ values,

$$\alpha = \frac{1}{N}\sum_{i=1}^{N} \gamma_i . \tag{19}$$

The statistic α will have the same units as the x_i and y_i values.

The test is normalized by comparison with the value α takes for Monte Carlo distributions created by shuffling the velocities randomly. WB used 1000 such simulations, but we used 300. To determine the significance, WB calculated the mean and standard deviation, σ, for the distribution of Monte Carlo α values and then calculated how many σ the real α was from the mean. As discussed in § 4.3, this distribution is *not* Gaussian, and assuming as much will produce different results than using the p-value of the statistic. We used the p-value significance instead. This should make the test slightly more *conservative* because of the long, upper tail of the distribution. On the other hand, WB used the 5% level as a cutoff for marginal significance, whereas we use 10% below. A final difference in our application of the test was the use of $N_{\rm nn} = N^{1/2}$ instead of $N_{\rm nn} = 10$.

WB point out that the presence of a subcluster will influence the centroid for the velocity group in which it falls, provided that it is separated from the primary both spatially and in velocity. One can qualitatively predict the performance of the α

test from a plot of the galaxy positions with different velocity intervals represented by different symbols. If any symbol is grouped together on one side of the cluster, that will increase α. WB also mention that, when no substructure is present, α is affected by Poisson fluctuations and by the intrinsic shape of the velocity and spatial distributions. This agrees with our warning in § 4.3 not to apply the Monte Carlo simulations of one distribution to another.

One might expect the α test to have problems with some cases of superposition. If a subcluster with large velocity separation is superposed on many primary cluster galaxies, the local standard deviations will be large for its galaxies. Therefore, the velocity dispersion weighting will *reduce* the centroid shift for galaxies in the velocity range of the subcluster. On the other hand, nearly all the galaxies in the velocity range of the subcluster will be members of the subcluster. It is not intuitively obvious which effect dominates.

Our results show that the α test is a sensitive diagnostic of substructure (see Figs. 23–28 and Table 11); α had probability centroids less than 0.1 (marginal significance) in 24 out of 36 cases, second only to the Δ test. We do not find superposition to be a major obstacle to α because spatial separations of only 100–500 kpc gave significant detections of the $\frac{1}{6}$ size subcluster for velocity separations of $\sim$2000 km s^{-1} ($N = 100$). In fact, spatial/velocity separations of 550 kpc/1100 km s^{-1} gave significant detections for $N \geq 140$ (6:1).

The velocity dispersion weights, w_i, used in this test give it a behavior very similar to ϵ. This includes dominance over the Δ test for late stages of merger (Figs. 23 and 24), and preference to small projection angles (Fig. 26). Since α measures a centroid shift, one expects sensitivity to be lost when the centroid of the subcluster is superimposed on the primary cluster. This is verified by the steep drop in α between 0° and 10° in Figure 26.

5.4.5. *The α Variation*

None of the three-dimensional tests drawn from the literature have a strong dependence on the surface density of the subcluster (only ϵ to a small extent). This inspired us to develop a variation on the α test which uses distance dispersions as weights instead of velocity dispersions. The algorithm is the same for α (above) except for the assignment of weights to each galaxy, i. Here we use $w_i = 1/\sigma_{ri}$, where

$$\sigma_{ri} = \left[\frac{\sum_{i=1}^{N_{nn}} \sum_{j=i}^{N_{nn}} (\mathrm{sep}_{i,j} - \overline{\mathrm{sep}})^2}{N_{nn}(N_{nn} - 1)} \right]^{1/2}, \qquad (20)$$

where $\mathrm{sep}_{i,j}$ is the projected separation between galaxies i and j, and $\overline{\mathrm{sep}}$ is the mean galaxy separation for all $N_{nn}(N_{nn} - 1)$ galaxy pairs. This is akin to a positional dispersion, whereby more compact groups will have larger weights, σ_{ri}. The remaining steps incorporate this "density" weighting just as the α test used velocity weighting.

The α var test did not excel as a diagnostic. It detects slightly more cases than ϵ at probabilities ≤ 0.05 (13), but for total marginal cases, it detects fewer than the three other tests. It is interesting that large differences between α var and α or ϵ occur for cases in which the subcluster is dispersed. The α and ϵ test continue giving significant detections for $T \geq 0.2$ Gyr; however, α var switches immediately to null value probabilities after the subcluster is dispersed (Figs. 23 and 24). Thus, α var serves as a diagnostic for "dispersed-ness" when combined with ϵ and/or α. Surprisingly, α var excels over ϵ but *not* α at early epochs when the subcluster is distinct. Also, α var seems to give smaller probabilities than α for $N \leq 80$ (3:1, velocity separation $\sim$ 1100 km s^{-1}, spatial separation $\sim$ 780 kpc). The difference is small, but it occurs four times. The α var test has the advantage that it is not sensitive to velocity dispersion gradients (§ 5.1).

5.5. *Substructure in the Simulations*

In this section, we will discuss what remnants of a merger are noticeable in our simulations when all the information about the N-body particles is available, and then we extrapolate these results to the observation regime, limited to three dimensions (x, y, V) and smaller sample sizes.

With access to all the output of our N-body code, we know that evidence of a merger exists throughout all epochs of the simulation. The output includes the three dimensions of position and three dimensions of velocity for all particles in our simulations (17,500 for the 6:1 merger and 20,000 for the 3:1 merger). In the first two sampled epochs, the spatial separation of the two components is obvious. The first two panels of Figure 2 demonstrate this. The two components are also accelerating toward each other. This allows for increasing distinguishability with time in the x velocity dimension, as seen in Figure 1. During the middle two epochs, −0.2 and +0.2 Gyr relative to core crossing, the subcluster is difficult to identify spatially. Nevertheless, one can "see" the subcluster as an asymmetry when viewed perpendicular to the merger. (Notice that Fig. 2 is viewed 30° off of perpendicular.) Even in the brief ($\sim$0.2 Gyr) time span in which the cores are overlapping, the subcluster is discernible in velocity space. Figure 1 shows that the velocity separation is 3000 km s^{-1}, and 3 times the velocity dispersion of the primary is 2400 km s^{-1}. With this large a separation, the particles can be identified as subcluster members without knowing their history. A more subtle effect observed during core crossing is the compression of the subcluster along the y and z dimensions. The dark matter particles of the subcluster are channeled through the potential well of the primary in a focused flow, much like the gas component described in Roettiger, Burns, & Loken (1993) and Roettiger et al. (1995b).

Phase mixing also occurs between the subcluster and primary during core crossing. This is the principal cause of the asymmetry with respect to the time of core crossing in Figure 1. The subcluster center of mass does not leave the merger with the same speed that it arrived, and the subcluster is expanding. Some of the kinetic energy of the subcluster has gone into the particles of the primary. The members of both components are seen in a "spray" of particles on the side of departure. The velocity dispersion of the entire system measured *perpendicular* to the merger axis rises *after* the core crossing. These facts demonstrate the transfer of momentum into dimensions perpendicular to the merger axis. Some of the subcluster particles depart at high velocities, and the simulation does not run long enough to show them turn around and fall back. This explains the fact that the center-of-mass separation (Fig. 1) never

reaches 0 kpc again after core crossing, despite the fact that a sizable fraction of subcluster galaxies merge with the primary by 2 Gyr after the core crossing.

The phase mixing essentially makes it impossible to assign particles to the subcluster or primary with 100% confidence. Nevertheless, we find that the combined distribution appears unrelaxed to the end of the simulation (4 Gyr). Spatially, the dispersed population is asymmetric, appearing mostly on the side of departure. Also, slight asymmetric elongations exist in the central regions as the subcluster oscillates preferentially along the x-axis. The velocities remain very anisotropic in the 4 Gyr after core crossing, as seen in Figure 29. In particular, the velocity dispersion of the system as a whole is larger along the x-axis than along orthogonal axes.

Now we will view our merger simulation from the perspective of an observer. We are limited to no more than two spatial dimensions and one velocity dimension. Our substructure tests can tell us how significant the clumping is for sample sizes $\leq$ 200 (see Tables 7, 8, 10, and 11).

In the earliest epochs examined, 2.0 and 0.6 Gyr before merger, the blatant spatial separation allows significant detections by two- and three-dimensional tests. Most of the two-dimensional diagnostics reject the null hypothesis at over the 99% level for both 3:1 and 6:1 mergers, $N = 100$, viewed at $\geq 60°$. (The separation is obvious to the eye.) Lee 2D does particularly well because it is sensitive to bimodal spatial distributions. (Note that if three or more subclusters were present, the Lee statistic would be hampered in its sensitivity, while the β and angular separation tests could excel.) Similarly, the three-dimensional diagnostics detect velocity-spatial correlations easily, but their maximum sensitivity is for viewing angles closer to 45°. The velocity separations of the subcluster along the merger axis are about 800 km s^{-1} ($1\sigma_{\rm primary}$) and 1500 km s^{-1} ($2\sigma_{\rm primary}$), for -2.0 and -0.6 Gyr, respectively. This is not extreme enough to measure significant nonnormality in the velocity distributions (for 3:1 or 6:1, $N = 100$). It is doubtful that *any* test would find significant structure for views nearly parallel to the merger axis ($\phi \sim 0°$) at $T \leq -0.6$ Gyr. We did not cover these particular parameter sets in our testing. However, the parameter set (0°, $T = -0.2$ Gyr, $N = 100$) was studied, and three tests, α, ϵ, and Δ, produced marginal significance

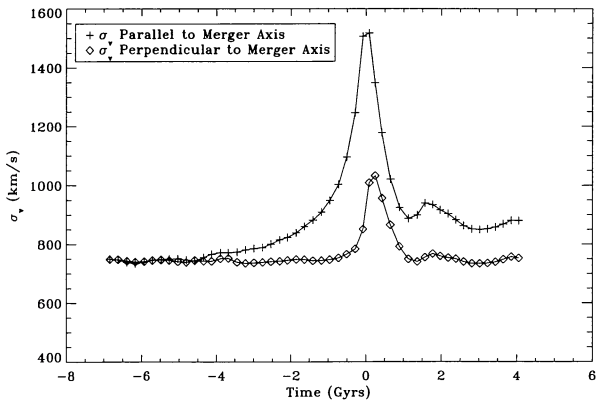

FIG. 29.—Velocity dispersion vs. time for the 3:1 mass ratio merger simulation.

levels. Since the relative velocity of the subcluster is much smaller at $T \leq -0.6$ than $T = -0.2$, it is doubtful that these tests would detect the merger.

During the epochs -0.2 and 0.2 Gyr, which straddle the moment of core crossing, the relative importance of the tests change. These are the only epochs in which the one-dimensional (normality) tests are quite sensitive. However, they only give low probabilities for $\phi \leq 30°$ ($N = 100$, 3:1 or 6:1). The subcluster spatially overlaps the primary for all projection angles, so the two-dimensional tests are ineffective. This is especially true for the 6:1 mass ratio, where the β test requires $N = 200$ to detect asymmetry at $\phi = 60°$. In contrast, for the 3:1 mass ratio, β can detect marginal asymmetry at $\phi = 60°$ with $N = 80$.

During core crossing, the subcluster is spatially mixed and weakly detectable from any viewing angle. The velocity distribution is a key indicator of the ongoing merger. This is where one-dimensional tests prevail. Ironically, if these mergers happened to occur along the line of sight, the velocity separation of 3000 km s^{-1} would make it observationally questionable whether the subcluster was physically associated or projected.

Despite the small spatial separation, the three-dimensional tests were sensitive to substructure during the core crossing epochs. Surprisingly, even at 0°, two tests (Δ and ϵ) found marginally significant spatial-velocity correlations (Fig. 26). The subcluster's core radius is smaller than that of the primary ($0.55\times$ for the 6:1 merger), so the explanation may lie in the fact that the two components do not overlap completely. The observed compression of the subcluster during its passage through the primary may enhance this effect. Projection angles near 90° at the time of core crossing reduce the probability of detection because only the FE and β two-dimensional test can detect the elongation or asymmetry, and there is no velocity separation. However, this epoch is so short lived, and the viewing angle so improbable, that such occurrences are rare.

The detectability wanes after core crossing because of the phase mixing (described above). However, the ϵ and α tests do quite well at finding a significant spatial-velocity correlation in the dispersed population. A high-velocity population still exists and can trigger one- and three-dimensional tests for views nearly parallel to the merger axis.

The merger signatures in X-rays were presented for the hybrid hydro/N-body code by Roettiger, Burns, & Loken (1995a), which used the same TREECODE plus an Eulerian hydrodynamics code, ZEUS-3D. They examined variations in the X-ray morphology as a function of relative cluster mass, finding examples of multiple peaked X-ray distributions, isophotal twisting, and centroid shifting throughout the merger evolution. It is expected that lines of sight along the merger axis will confound the interpretation of X-ray data (see Buote & Tsai 1995). Here the use of traditional redshift surveys can provide a clearer view of the situation. Ideally, a combination of X-ray and optical data can be used to demonstrate the presence of substructure.

5.6. *Mass Estimators and Substructure*

It has long been realized that cluster mass estimates based on the virial theorem are prone to systematic errors (Bahcall & Tremaine 1981). These errors include contamination by non-

member galaxies, the unknown radial distribution of dark matter, anisotropic galaxy orbits (e.g., caused by anisotropic collapse), and the presence of merging subclusters.

Here we examine the systematic error induced by the presence of merging subclusters. We have applied previously published dynamical mass estimators to the same files to which we applied substructure diagnostics. This allows a quantitative look at the errors incurred in applying estimators (which are intended for a relaxed system of self-gravitating particles) to systems with substructure. Furthermore, the errors can be studied as a function of sample size, projection angle, subcluster size, and stage of merger.

We use two dynamical mass estimators introduced in Heisler et al. (1985) and frequently found in the literature. These are the virial theorem (hereafter $M_{\rm VT}$) and the projected mass estimator (hereafter, $M_{\rm PME}$). The virial theorem mass can be expressed in terms of observables as follows:

$$M_{\rm VT} = \frac{3\pi N}{2G} \frac{\sum_i V_{zi}^2}{\sum_{i<j} 1/R_{\perp,ij}}, \qquad (21)$$

where N is the galaxy sample size, V_{zi} is the radial velocity with respect to the velocity centroid of galaxy i (cosmologically corrected), and $R_{\perp,ij}$ is the projected distance between galaxies i and j. Similarly, the projected mass estimator is

$$M_{\rm PME} = \frac{\xi}{GN} \sum_i V_{zi}^2 R_{\perp i}, \qquad (22)$$

where $R_{\perp i}$ is now the projected distance between galaxy i and the cluster centroid, and ξ is $32/\pi$ assuming isotropic orbits (see eq. [14]). A factor of H_0^{-1} is introduced by the $R_\perp$ terms. In addition, we use the robust virial theorem, $M_{\rm RVT}$, where the canonical mean and standard deviation are replaced by the biweight estimators of centroid and scale, as recommended by Beers et al. (1990).

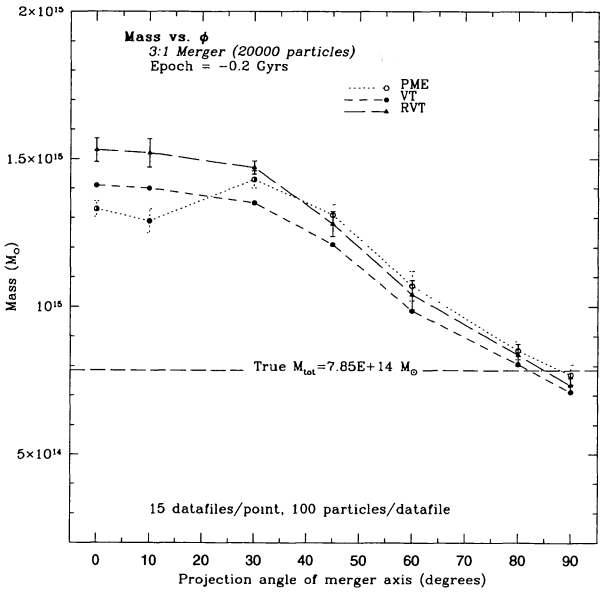

FIG. 30.—Dynamical mass estimators vs. projection angle for the 3:1 mass ratio merger simulation. The horizontal line marks the true total mass.

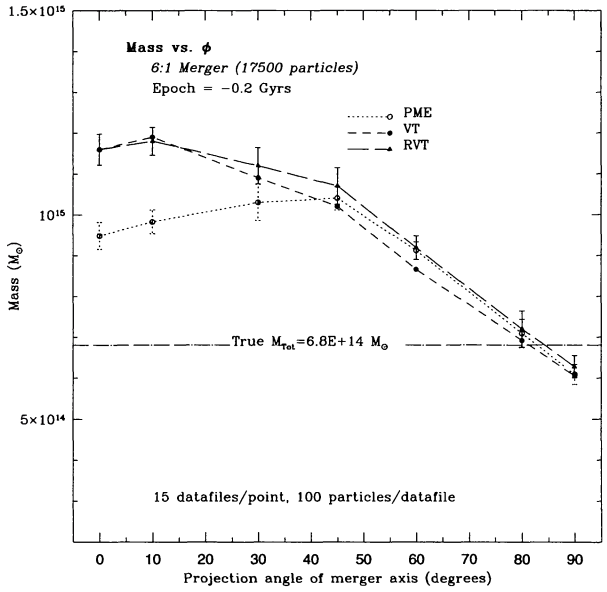

FIG. 31.—Dynamical mass estimators vs. projection angle for the 6:1 mass ratio merger simulation.

The results of the behavior of mass estimators are shown in Figures 30–34. On each figure, the actual mass of the cluster plus subcluster is shown. Fifteen data files of 100 points were created for each point (except Fig. 34). Error bars of 68% are shown; however, the errors are $15^{1/2}$ times greater when applied to a single, $N = 100$ data file.

The dependence of mass estimates on projection angle are shown in Figures 30 and 31. The epoch 0.19 Gyr prior to core crossing was examined in both the 3:1 and 6:1 mergers. At this

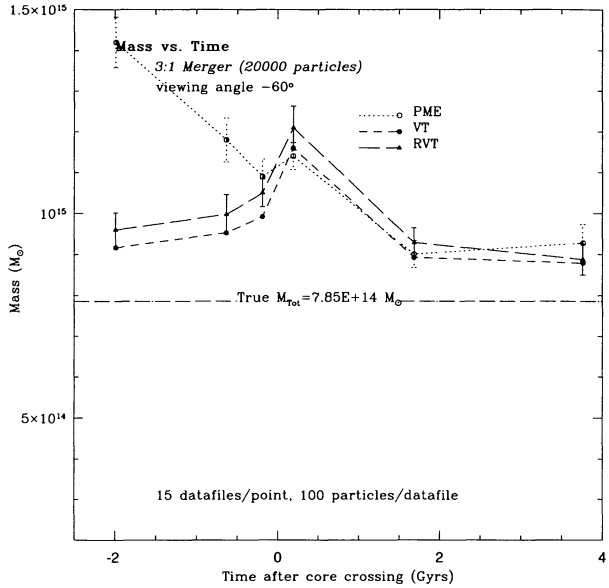

FIG. 32.—Dynamical mass estimators vs. time for the 3:1 mass ratio merger simulation.

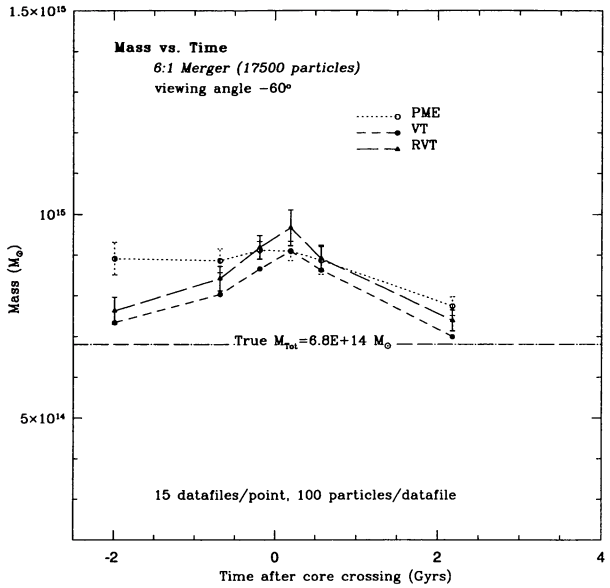

FIG. 33.—Dynamical mass estimators vs. time for the 6:1 mass ratio merger simulation.

time, the subcluster and cluster were separated by 700 kpc and approaching each other at 2450 km s^{-1}. The projected separations for each angle are given in Table 9. We see that the masses are overestimated by as much as a factor of 2 when the merger occurs within 30° of the line of sight. The discrepancy is not as great for $M_{\rm PME}$ as $M_{\rm VT}$ and $M_{\rm RVT}$. The factor decreases smoothly to less than 1 at 90°. This is to be expected because the velocity dispersion of each component should be less than the dispersion of a single, relaxed system with the same mass as the cluster plus subcluster.

The dependence of the estimators on time is depicted for the 3:1 and 6:1 mergers in Figures 32 and 33, respectively. We keep the projection angle fixed at 60°, where the apparent velocity difference of the components is $\frac{1}{2}$ of the true value. The mass is overestimated at all epochs; however, the error is only 15%–20% at most times. Comparison with Figures 30 and 31 suggests that masses would be overestimated by 30%–40% for viewing angles under 30° and epochs over ~1 Gyr from core crossing. The outstanding feature is the peak in mass which occurs about 0.2 Gyr after the core crossing. This time delay probably increases with viewing angle; it takes time for phase mixing to convert the large-scale motion of the subcluster into random motions perpendicular to the merger axis.

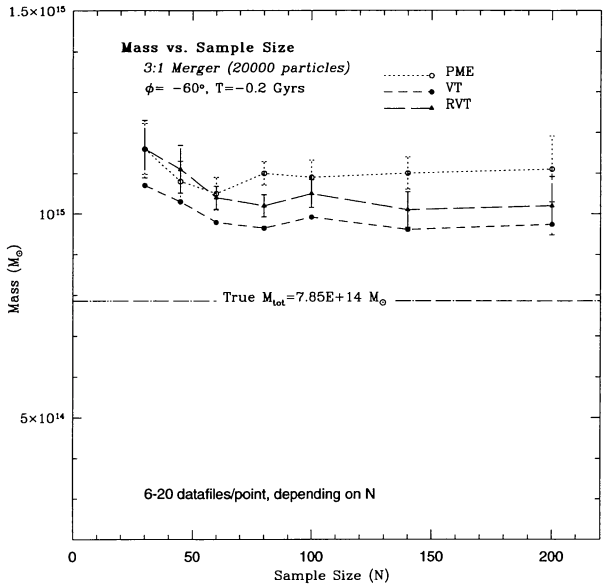

FIG. 34.—Dynamical mass estimators vs. sample size for the 6:1 mass ratio merger simulation.

Some significant differences between $M_{\rm PME}$ and the other global estimators appear in the figures. These all arise because the radius $R_{\perp i}$ is measured with respect to the centroid of all particles, whereas $R_{\perp i,j}$ is a harmonic mean separation (smaller distances get greater weights). During premerger cases, the centroid will be *in between* the two clusters, resulting in a larger $R_{\perp i}$ and a proportionately larger mass for $M_{\rm PME}$. This accounts for the large differences in Figures 32 and 33. In Figures 30 and 31, $M_{\rm PME}$ is *less* than than the others for angles less than 30° because $R_{\perp i}$ is smaller than $R_{\perp i,j}$ when the two clusters are superimposed on each other.

We see a consistently greater $M_{\rm RVT}$ than $M_{\rm VT}$. For any given point, the difference is not significant, but the trend exists for nearly all points. There are two places in the calculation of these estimators at which the robust estimators replace the canonical mean and standard deviation. The first is the dispersion calculation, where $S_{\rm BI}$ is substituted. The second is in calculating $R_{\perp i,j}$: the conversion from arcseconds to megaparsecs depends on the centroid of the velocity distribution, and here we substitute $C_{\rm BI}$ for μ. The small difference in $R_{\perp i,j}$ cannot account for the ~5% mass differences. The larger size of $S_{\rm BI}$, however, is sufficient to account for the difference since $M \propto \sigma^2$. $S_{\rm BI}$ was advertised as an estimate of scale which is more resistant to outliers than the standard deviation; however, the subcluster is not small enough or fast enough to be "seen" as an outlier by the estimator. Instead, $S_{\rm BI}$ sees a broader distribution. Notice that $M_{\rm VT}$ and $M_{\rm RVT}$ become the same when there is no radial velocity difference between the two components (Figs. 30 and 31).

The "resistant" nature of the biweight estimators in RVT was not important here. $M_{\rm RVT}$ would have an advantage in cases in which there is velocity contamination by nonmembers (i.e., the observational situation), but nonmembers were not included in our simulation.

$M_{\rm PME}$ was recently used in a study of the effects of substructure on real cluster mass estimates (Bird 1995). She used an objective partitioning scheme (KMM) to remove substructure from clusters. The remaining cluster was typically less massive by factors often greater than 2 (using $M_{\rm PME}$). In contrast, Biviano et al. (1993) and Escalera et al. (1994) claimed that substructure did not affect their optical mass estimates for real clusters. Our results agree with Bird's: if two-body, head-on mergers are occurring in clusters, cluster masses will be overestimated for nearly all projection angles.

5.7. *Effects of Spatial Limits on Cluster Samples*

A major difference between our simulated data sets and observations is the lack of spatial limits on the sampled galaxies.

TABLE 12

EFFECTS OF SPATIAL LIMITS

PARAMETER SET			PARTICLES LOST		
Simulation (1)	Epoch (Gyr) (2)	ϕ (3)	Outside 1.0 Mpc (4)	Outside 1.5 Mpc (5)	Outside 2.0 Mpc (6)
Isothermal	...	...	0.033	0.013	0.003
3:1 merger	−1.9	−60°	0.098	0.023	0.004
	−0.6	−60	0.073	0.015	0.003
	−0.2	−60	0.067	0.019	0.004
	+0.2	−60	0.080	0.020	0.003
	+1.7	−60	0.224	0.099	0.035
	+3.8	−60	0.258	0.165	0.124
	−0.6	−30	0.072	0.017	0.003
	+1.7	−30	0.222	0.096	0.038
	+3.8	−30	0.253	0.158	0.126
6:1 merger	+0.6	−60	0.050	0.012	0.001

NOTES.—Cols. (1)–(3) specify the simulation parameters. Cols. (4)–(6) show the fraction of particles (representing galaxies) which would be excluded by imposing a spatial limit (circles of 1, 1.5, and 2.0 Mpc radius, see text) on draws from the simulations. The fractions in cols. (4)–(6) are calculated with 100 random data files. A typical scatter in these fractions is ~0.002. Each data file contains 100 particles.

When cluster redshifts are measured, the region from which galaxies are chosen is generally minimized either for technical reasons or for time constraints. Our galaxies, on the other hand, are drawn randomly from the entire simulation box (6 h_{75}^{-1} Mpc across). Here we assess the importance of spatial limits on substructure detection.

We began by estimating the fraction of galaxies that would be lost in a spatially limited survey in the following way. First, a large sample was drawn randomly and projected onto the sky from the desired viewing angle, ϕ. Second, a centroid was determined for the two components separately. Third, we counted the number of particles which fell outside *two* circles, one centered on the subcluster and the other centered on the cluster. We tried radii typical of redshift surveys: 1, 1.5, and 2.0 h_{75}^{-1} Mpc. We used the same radius for both circles. This is a preferable method over using a single circular region because during the early epochs of merger the subcluster would be well separated and obvious to the observer.

Table 12 shows the fraction of particles in simulations that would be rejected as spatial outliers. The fractions for the isothermal case merely reflect the radial profile of a cluster with a core radius of 250 kpc. The merger cases show higher fractions than the isothermal case. We find that the late epochs are most likely to have ejected galaxies. For example, 3.8 Gyr after a 3:1 merger, as many as 25% of galaxies can be lost by a redshift survey within only 1 Mpc radius. Fortunately, 0.2 Gyr before core crossing for the 3:1 merger, only 2% of galaxies are lost with a 1.5 Mpc radial limit (this epoch was used for most of our parameter sets). A comparison of the 6:1 and 3:1 mergers at late epochs shows that the small-scale mergers will scatter fewer particles outside the radial limits. A comparison of $\phi = -60°$ and $-30°$ shows that the viewing angle does not make a significant difference on the fraction lost.

We conclude that a $\gtrsim$1.5 Mpc spatial cutoff on cluster surveys will not have a great effect on the galaxy count except for epochs following a large-scale merger. This is because no more than 3% of the galaxies will be lost from the search field until postmerger epochs, when the loss rate reaches 10% by about 2 Gyr (3:1 merger). This result depends somewhat on the true shape of clusters.

We have yet to determine how much a spatial cutoff influences the significance levels of substructure tests. We chose to run our analysis program on 1.5 Mpc spatially limited data files drawn from six epochs of the 3:1 simulation, viewed at −60°. The results are compared with data files without spatial limits in Table 13. Again, the average over 15 data files is shown for each entry. The S_{BI} for the datafiles with no spatial limit can be found in Tables 11 and 12. The 68% errors are $S_{BI}/15^{1/2}$.

We find that the two-dimensional substructure tests are most strongly affected by the spatial cutoff. For "FE," β, and AST, the cutoff lowers the detectability of substructure, espe-

TABLE 13

EFFECTS OF SPATIAL LIMITS ON SUBSTRUCTURE TESTS

EPOCH (Gyr) (1)	LIMIT? (2)	SUBSTRUCTURE TEST								
		FE (3)	β (4)	AST (5)	Lee2 (6)	Lee3 (7)	Δ (8)	ϵ (9)	α (10)	α var (11)
−1.9	No	0.000	0.000	0.000	0.000	0.000	0.026	0.423	0.031	0.106
	Yes	0.000	0.000	0.000	0.000	0.000	0.015	0.056	0.000	0.043
−0.6	No	0.000	0.000	0.041	0.000	0.000	0.000	0.070	0.002	0.009
	Yes	0.000	0.000	0.145	0.000	0.000	0.006	0.089	0.015	0.079
−0.2	No	0.000	0.006	0.227	0.297	0.179	0.000	0.004	0.004	0.011
	Yes	0.007	0.047	0.364	0.006	0.001	0.000	0.024	0.006	0.002
+0.2	No	0.013	0.502	0.426	0.392	0.421	0.005	0.092	0.124	0.230
	Yes	0.240	0.448	0.431	0.410	0.396	0.208	0.366	0.386	0.511
+1.7	No	0.008	0.043	0.506	0.142	0.175	0.225	0.134	0.193	0.409
	Yes	0.055	0.396	0.264	0.286	0.310	0.129	0.257	0.105	0.384
+3.8	No	0.048	0.018	0.158	0.080	0.071	0.159	0.014	0.051	0.421
	Yes	0.614	0.149	0.551	0.397	0.373	0.237	0.440	0.247	0.445

NOTES.—Col. (1) contains the epoch of merger. The other parameters are the same for all rows: mass ratio = 3:1, projection angle = 60°. Col. (2) contains a "y" if a 1.5 Mpc spatial limit was imposed on the selection of particles, and an "n" for no limit. Cols. (3)–(11) contain significance levels in the form of probabilities for each two- and three-dimensional test. The lower the probability, the more significant the substructure.

cially in the last two epochs. During premerger epochs, the cutoff actually improves detectability for Lee 2D and Lee 3D because it makes the two components more spatially distinct. After core crossing, the cutoff again decreases detectability. Three-dimensional tests also lose sensitivity when the spatial limit is used. This is most obvious for the epochs +0.2 and +3.8 Gyr. However, many exceptions occur, especially when neither of the two cases is significant.

6. SUMMARY

6.1. *Prescription of a Battery of Tests*

We find that the most information can be extracted from a cluster data set using not one but many tests. Here we review which tests are the most useful. We provide a quantitative evaluation of the tests based on our parameter search in Tables 14 and 15. Recommended tests are footnoted.

A comparison of Table 14 with Table 15 shows that the one-dimensional tests are generally much less sensitive to substructure than two- or three-dimensional tests. Substructure will most likely first appear significant in two or three dimensions as a redshift survey grows. However, the velocity distribution can reveal substructure hidden from two- and three-dimensional tests. This occurs when the subcluster is closely superimposed on the primary cluster and rapidly moving with respect to the primary (i.e., mergers along the line of sight). Moreover, substructure due to mergers is expected to produce a certain *kind* of nonnormality so that one-dimensional tests can be used as a check that the substructure signal in two or three dimensions is not caused by velocity outliers (more below).

Our results confirm that velocity distributions are skewed by a merger of two unequal-sized components. We recommend that the canonical skewness test B_1 should be used along with the more conservative AI. B_1 is also used to produce the statistic SS3 (more below).

The hybrid statistics av10, skewav, and kurtav are not as useful as we had hoped. For instance, av10 averaged skewness tests with kurtosis tests, each rejecting the Gaussian hypothesis for a different reason. Consequently, av10 will indicate that the distribution is consistent with a Gaussian unless it is *both* strongly skewed and long- or short-tailed. The statistics kurtav and skewav were similarly overconservative because they included conservative tests in their averages; skewav was particularly ineffective because it included the I test which has coarse percentile points.

We advise that several kurtosis tests should be applied separately. We would recommend the TI, W^2, W, and B_2 tests. Bird & Beers (1993) recently tabulated percentage points for TI and discuss its performance. W had the most detections of significance at the 1% level. U^2 and A^2 behave very similarly to W^2, but W^2 had lower probabilities. B_2 appears less sensitive to long tails than than W^2 and more sensitive to short tails. It has often been used in the literature as the canonical "kurtosis" test, and so it allows comparison to previous results. These all have precisely estimated probabilities and a variety of sensitivities to complement each other.

Many of the recommended kurtosis and skewness tests can provide useful information about the velocity distribution in addition to the significance level. This information was used to create secondary statistics, SS1, SS2, and SS3. Our results confirm previous findings that genuine substructure is more likely to appear as short tails, whereas long tails can be caused by erroneous velocities or foreground/background contamination (Ashman et al. 1994; Bird & Beers 1993). The statistics A, B_2 (used in SS2), and TI can distinguish between short- and long-tailed distributions. Thus, they should be combined with the kurtosis tests to allow the interpretation of substructure. Similarly, B_1 (SS3) or AI can indicate which side of the velocity distribution contains an excess of counts, and W^2, U^2, and A^2 (SS1) can indicate on which side of the distribution a long tail exists.

TABLE 14

ONE-DIMENSIONAL SUBSTRUCTURE TEST SUMMARY

	SUBSTRUCTURE DETECTIONS (OUT OF 36 CASES)			
TEST NAME (1)	$P = 1\%$ (2)	$P = 5\%$ (3)	$P = 10\%$ (4)	COMMENTS (5)
A	...	...	...	Indicates long or short tails (SS2); significance level not found in "ROSTAT"
B1[a]	0	4	4	Skewness; required sample size <200; used in SS3
B2[a]	0	0	3	Detects boxiness more easily than others; used in SS2; 3:1 merger easiest; commonly used
W[a]	4	4	5	Commonly used in literature; used in SS1, SS2
A^{2a}	3	5	5	Same strengths as W^2 and U^2 but more sensitive; used in "SS1"
U^2	3	5	5	Kurtosis
W^2	3	5	5	Kurtosis
DIP[a]	0	3	10	Can detect bimodality in 3:1 and 6:1; needs more testing
av10	0	0	4	Insensitive; mixes kurtosis and skewness
kurtav	0	0	2	Insensitive; 6 kurtosis tests
skewav	0	0	5	Insensitive, only 3 tests averaged

NOTES.—Cols. (2)–(4) give the number of substructure detections in 36 cases with less than 1%, 5%, and 10% significance, respectively. The 36 cases probe parameters such as sample size, time in merger, subcluster mass ratio, and projection angle. See Table 1 for more information on each test.

[a] Recommended for a battery of tests.

TABLE 15

TWO- AND THREE-DIMENSIONAL SUBSTRUCTURE TEST SUMMARY

TEST NAME (1)	SUBSTRUCTURE DETECTIONS (OUT OF 36 CASES) $P = 1\%$ (2)	$P = 5\%$ (3)	$P = 10\%$ (4)	COMMENTS (5)
Two-dimensional Tests				
FE[a]	11	19	24	Very sensitive to elongations, but should be used with a threshold on the FE statistic; provides a P.A.
β[a]	9	15	16	Most sensitive two-dimensional substructure test
AST	1	3	4	Very insensitive; requires compact, spatially distinct subcluster
Lee 2D[a]	5	7	9	Requires spatially distinct subcluster, Provides P.A. of axis connecting components
Three-dimensional Tests				
Lee 3D[a]	7	7	10	Slight improvement on 2D; P.A. should be the same as Lee 2D for real subcluster
Δ[a]	14	24	30	Most sensitive test! Graphic can show location of subcluster
ϵ[a]	7	13	21	Sensitive to dispersed subcluster; provides mass information
α[a]	12	16	25	Sensitive to dispersed subcluster
αvar[a]	8	15	17	Not sensitive to dispersed subcluster, should excel for compact subclusters

NOTES.—Columns are the same as in Table 9. See Tables 2 and 3 for more information on each test.

[a] Recommended for a battery of tests.

Three of our two-dimensional diagnostics, FE, Lee 2D, and β, are recommended for substructure analysis (Table 15). Some caveats must be mentioned. The Lee 2D test is sensitive only to genuine, two-dimensional substructure, while FE and, to a lesser extent, β can respond to structure in a cluster which is not necessarily a merging subcluster. Lee 2D is limited in that it is insensitive to more than two clumps, and to clumps of very unequal size. The FE test should not be considered significant unless the elongation strength is larger than 2.5. To avoid using these tests blindly, the data can be visualized using "skyplots" or an adaptive kernel contouring technique (Silverman 1986; Beers et al. 1992). Finally, these tests are intended for flux-limited samples of galaxies which are confirmed cluster members. However, they can provide accurate results if the sample is not photometrically complete but is spatially fair (i.e., with no biases toward certain regions of the field). If the galaxies do not have measured redshifts, there is the danger of foreground and background groups appearing as substructure.

Finally, we recommend Lee 3D, Δ, ϵ, α, and α var for clusters with radial velocities. Substructure can appear as a correlation between velocity mean and/or dispersion with position. Combining all three observable dimensions allows marginally significant substructure detections for velocity samples as small as 30 (the limit of our testing). Since the null hypothesis is spatially uniform velocity mean and dispersion, three-dimensional tests do not require fair spatial sampling as stringently as two-dimensional tests (Lee 3D is an exception). The Δ test was the most sensitive three-dimensional test. However, there are some situations in which Δ is not as sensitive and other tests should be used. These include the ϵ test and the α test, which both excel in postmerger epochs. The α var test is expected to excel in cases with high-density subclusters. Again, the three-dimensional results should be cross-checked with skyplots, bubble plots (§ 5.4.2), and velocity dispersion profiles to understand the origin of the velocity-spatial correlation.

6.2. *Detectability of Substructure in Mergers*

This work attempts to improve the connection between optical observations and numerical models of cluster-subcluster mergers. It complements the predictions for X-ray observations drawn from hydrodynamic simulations of the gas in mergers (Evrard 1990; Roettiger et al. 1993, 1995a).

We have shown that, in most circumstances, substructure should be detectable in the galaxies of clusters undergoing a merger using a battery of one-, two-, and three-dimensional diagnostics on typical cluster redshift surveys. The premerger and ongoing merger stages appear statistically significant in $N = 100$ data sets for views other than perpendicular to the merger axis. For perpendicular views at the time of core crossing, the merger will not be detectable. Fortunately, signatures in X-ray emission should be strong here. During post-merger epochs, the substructure remains detectable in three dimensions for over 3 Gyr, provided that the dispersed population is sampled in the redshift survey.

The main qualification of our results is that our simulations may differ from real clusters in ways which affect the detectability of subclusters. Observations are complicated by superpositions of foreground and background objects and velocity errors. Cluster redshift surveys are often limited to small radii so that substructure tests cannot react to asymmetries beyond $\sim 2\ h_{75}^{-1}$ Mpc. The survival of our subclusters is shortened by the head-on nature of our simulations (Gonzáles-Casado et al. 1994). The detectability of merging subclusters should also be altered by differences in the behavior of dark matter and galax-

ies. For example, the subcluster galaxies should be slowed by dynamical friction more readily than the subcluster dark matter. Our peak relative velocities are high (§ 4.4), but the paucity of large observed relative velocities may be explained by the short duration of the epochs in which the subcluster has the highest relative speed (see Fig. 1).

6.3. *Conclusions*

We have used N-body simulations of cluster-subcluster mergers to evaluate estimators of substructure in one-, two-, and three-dimensional data sets. In addition, we have determined the influence of two-body mergers on dynamical mass estimators, and the influence of spatial limits on detecting mergers in redshift surveys.

To encourage the future application of substructure tests to redshift surveys, we have collected numerous tests from the literature, described algorithms for their calculation, suggested improvements, and provided a fair means of intercomparison. Our false positive testing helps to ensure that our algorithms are properly coded and the normalization procedures are valid. We also point out two tempting but invalid normalization procedures. The significance levels for two- and three-dimensional diagnostics are found using numerous Monte Carlo simulations of the *original* data set. (We recommend 1000 simulations for real cluster data sets.) During this testing, we also found that a velocity dispersion gradient increases the significance of most three-dimensional substructure tests compared to clusters with flat dispersion profiles. Consequently, we recommend that a 1% significance level is adhered to when blindly applying such tests for diagnosing substructure. This applies especially to the ϵ, α, and Δ tests, which are the most dependent on the velocity dispersion profile.

Applying the tests to random draws from our cluster-subcluster merger simulations revealed their relative strengths. The tests are capable of detecting the substructure in redshift data sets as small as $N = 30$ (3:1 merger) and $N = 60$ (6:1 merger). As expected, the higher the dimensionality of the test, the greater its sensitivity to substructure. However, our parameterization of projection angle, time in merger, and the relative size of the subcluster demonstrates that the tests are very complementary. This is a natural consequence of the fact that the two subclusters increase their relative speed as they decrease their spatial separation. The two-dimensional tests are sensitive to large projected spatial separations but insensitive to large radial velocity separations. The one- and three-dimensional tests are most sensitive to radial velocity separation. The one-dimensional tests are most sensitive to mergers oriented along the line of sight, while the three-dimensional tests are most sensitive to merger axes 30°–45° from the line of sight. The most difficult parameter space for the tests to diagnose is during core crossing, with the line of sight perpendicular to the merger. Fortunately, this is where X-ray data can reveal shocks, high temperatures, and unusual structure in the gas density and temperature distributions (Roettiger et al. 1995a).

A battery of tests is prescribed to capitalize on this complementary behavior. We recommend the B_1, B_2, W, A^2, and DIP statistics for normality, the FE, Lee 2D, and β statistics for spatial substructure, and the Lee 3D, Δ, ϵ, α, and α var statistics for three-dimensional substructure. If these are combined with visualization of the data (e.g., velocity dispersion profiles, contour plots), one can rule out spurious substructure detections, and the 5% significance level can be taken as the lower limit of marginal significance.

By applying mass estimators to our simulated clusters, we have quantified the effects of a two-body merger on optical mass estimates. Each of the three estimators, $M_{\rm PME}$, $M_{\rm VT}$, and $M_{\rm RVT}$, overestimate the known total mass by 0%–90%, depending on the projection angle of the merger. Mergers oriented along the line of sight cause the largest overestimate. The projected mass estimator $M_{\rm PME}$ can depart significantly from the other two estimators in situations in which the subcluster is spatially separated from the primary cluster by large distances.

Finally, we have quantified the effect of spatial limits in redshift surveys on substructure diagnoses. In general, when substructure tests are applied to two samples of equal number, one with a radial limit and the other without, the tests are more sensitive to substructure in the unlimited sample. This difference was greater for the two-dimensional tests than the three-dimensional tests and becomes insignificant for radial limits above 2 Mpc.

We would like to thank M. West and A. Zabludoff for stimulating discussions. We thank George Rhee for discussions and for supplying code for Fourier spatial analysis. We also thank Lars Hernquist for his N-body TREECODE. J. P. thanks his Ph.D. committee (Jack Burns, Erica Ellingson, George Goedecke, Michael Ledlow, and René Walterbos) for their comments on the text. This work was supported by NSF grant (AST-9317596) and NASA Long Term Space Astrophysics grant (NAGW-3152) to J. O. B. The authors will make available standard FORTRAN source code for our substructure analysis program. Send requests to jpinkney@nmsu.edu or bird@kula.phsx.ukans.edu.

REFERENCES

Ashman, K. M., Bird, C. M., & Zepf, S. E. 1994, AJ, 108, 2348
Bahcall, J. N., & Tremaine, S. 1981, ApJ, 244, 805
Bechtold, J., Forman, W., Giacconi, R., Jones, C., Schwarz, J., Tucker, W., & Van Speybroeck, L. 1983, ApJ, 256, 26
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
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. 1992, ApJ, 400, 410
Beers, T. C., Kriessler, J. R., Bird, C. M., & Huchra, J. P. 1995, AJ, 109, 874
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Bird, C. M. 1993, Ph.D. thesis, Univ. Minnesota
———. 1994, ApJ, 422, 480
Bird, C. M., & Beers, T. C. 1993, AJ, 105, 1586
Bird, C. M. 1995, ApJ, 445, L81
Biviano, A., Girardi, M., Giuricin, G., Mardirossian, F., & Mezzetti, M. 1993, ApJ, 411, L13
Bothun, G. D., & Schombert, J. M. 1988, ApJ, 335, 617
———. 1990, ApJ, 360, 436
Briel, U. G., et al. 1991, A&A, 246, L10

Buote, D. A., & Canizares, C. R., 1992, ApJ, 400, 385
Buote, D. A., & Tsai, J. C. 1995, ApJ, 439, 29
Burns, J. O., Roettiger, K., Pinkney, J., Loken, C., Doe, S., Owen, F., Voges, W., & White, R. 1994, in Proc. *ROSAT* Science Symposium, ed. M. Schlegel & R. Petre (New York: AIP)
Burns, J. O., Roettiger, K., Pinkney, J., Perley, R. A., Owen, F. N., & Voges, W. 1995, ApJ, 446, 583
Carlberg, R. G., & Couchman, H. M. P. 1989, ApJ, 340, 47
Colless, M. A., & Hewett, P. C. 1987, MNRAS, 224, 453
Crone, M. M., & Geller, M. J. 1995, AJ, 110, 21
Dressler, A., & Shectman, S. A. 1988, AJ, 95, 985 (DS)
Escalera, E., Biviano, A., Girardi, M., Giuricin, G., Mardirossian, F., Mazure, A., & Mezzetti, M. 1994, ApJ, 423, 539
Evrard, A. E. 1990, ApJ, 363, 349
Fitchett, M. J., 1988a, in Minnesota Lectures on Clusters of Galaxies and Large Scale Structure, Vol. 5, ed. J. M. Dickey 143
———. 1988b, MNRAS, 230, 161
Fitchett, M. J., & Webster, R. 1987, ApJ, 317, 653
Gebhardt, K., & Beers, T. C. 1991, ApJ, 383, 72
Gonzáles-Casado, G., Mamon, G. A., & Salvador-Solé, E. 1994, ApJ, 433, L61
Gunn, J. E., & Weinberg, D. H. 1995, in Wide-Field Spectroscopy and the Distant Universe, Proc. 35th Herstmonceaux Workshop, ed. S. J. Maddox & A. Aragón-Salamanca (Singapore: World Scientific)
Hartigan, J. A., & Hartigan, P. M. 1985, Ann. Stat., 13, 70
Heisler, J., Tremaine, S., & Bahcall, J. N. 1985, ApJ, 298, 8
Henry, J. P., & Briel, U. G. 1991, A&A, 246, L14
Hernquist, L. 1987, ApJS, 64, 715
Hill, J. M., & Oegerle, W. R. 1993, AJ, 106, 831
Huchra, J. P., Geller, M. J., & Corwin, H. G., Jr. 1995, ApJS, 99, 391
Jones, C., & Forman, W. 1992, in Clusters and Superclusters of Galaxies, ed. A. C. Fabian (Dordrecht: Kluwer), 49
Katgert, P., et al. 1995, A&A, in press
Kauffman, G., & White, S. D. M. 1993, MNRAS, 261, 921
Lacey, C., & Cole, S. 1993, MNRAS, 262, 627
Lee, K. L. 1979, J. Am. Stat. Assoc., 74 (367), 708
Maddox, S. J., Efstathiou, G., Sutherland, W., & Loveday, J. 1990, MNRAS, 246, 433
Merritt, D. 1987, ApJ, 313, 121
Mohr, J. J., Fabricant, D. G., & Geller, M. J. 1993, ApJ, 413, 492
Pinkney, J., Burns, J. O., & Hill, J. M. 1994, AJ, 108, 2031
Pinkney, J., Burns, J. O., Ledlow, M. J., & Hill, J. M. 1995, in preparation
Pinkney, J., Rhee, G., Burns, J. O., Hill, J. M., Oegerle, W., Batuski, D., & Hintzen, P. 1993, ApJ, 416, 36
Rhee, G. F. R. N., van Haarlem, M. P., & Katgert, P. 1991, A&AS, 91, 513
Richstone, D., Loeb, A., & Turner, E. L. 1992, ApJ, 393, 477
Roettiger, K., Burns, J. O., & Loken, C. 1993, ApJ, 407, L53
———. 1995a, ApJ, submitted
Roettiger, K., Burns, J. O., & Pinkney, J. 1995, ApJ, 453, 634
Roettiger, K., Loken, C., & Burns, J. O. 1995b, ApJS, submitted
Serna, A., Alimi, J.-M., & Scholl, H. 1994, ApJ, 427, 574
Sharples, R. M., Ellis, R. S., & Gray, P. M. 1988, MNRAS, 231, 479
Silverman, B. 1986, Density Estimation for Statistics and Data Analysis (London: Chapman & Hall)
Teague, P. F., Carter, D., & Gray, P. M. 1990, ApJS, 72, 715
Tonry, J. L. 1985, AJ, 90, 2431
Ueda, H., Itoh, M., & Suto, Y. 1993, ApJ, 408, 3
Venkatesan, T. C. A., Batuski, D. J., Hanisch, R. J., & Burns, J. O. 1994, ApJ, 436, 67
West, M. J., & Bothun, G. D. 1990, ApJ, 350, 36 (WB)
West, M. J., Oemler, A., & Dekel, A. 1988, ApJ, 327, 1
West, M. J., & Richstone, D. O. 1988, ApJ, 335, 532
Yahil, A., & Vidal, N. V. 1977, ApJ, 214, 347
Zabludoff, A. I., Geller, M. J., Huchra, J. P., & Vogeley, M. S. 1993, AJ, 106, 1273
Zabludoff, A. I., Huchra, J. P., & Geller, M. J. 1990, ApJS, 74, 1
Zabludoff, A. I., & Zaritsky, D. 1995, ApJ, 447, L21