IRAS VERSUS POTENT DENSITY FIELDS ON LARGE SCALES: BIASING PARAMETER AND $\Omega$

YAIR SIGAD, 1 AMIRAM ELDAR, 1 AVISHAI DEKEL, 1,2,3 MICHAEL A. STRAUSS, 4,5 AND AMOS YAHIL 6

Received 1997 August 7; accepted 1997 October 16

ABSTRACT

The galaxy density field as extracted from the IRAS 1.2 Jy redshift survey is compared to the mass-density field as reconstructed by the POTENT method from the Mark III catalog of peculiar velocities. The reconstruction is done with Gaussian smoothing of radius 12 $h^{-1}$ Mpc, and the comparison is carried out within volumes of effective radii 31–46 $h^{-1}$ Mpc, containing $\approx 10–26$ independent samples. Random and systematic errors are estimated from multiple realizations of mock catalogs drawn from a simulation that mimics the observed density field in the local universe. The relationship between the two density fields is found to be consistent with gravitational instability theory in the mildly nonlinear regime and a linear biasing relation between galaxies and mass. We measure $\beta_l \equiv \Omega^{0.6}/b_l = 0.89 \pm 0.12$ within a volume of effective radius 40 $h^{-1}$ Mpc, where $b_l$ is the IRAS galaxy biasing parameter at 12 $h^{-1}$ Mpc. This result is only weakly dependent on the comparison volume, suggesting that cosmic scatter is no greater than $\pm 0.1$. These data are thus consistent with $\Omega = 1$ and $b_l \approx 1$. If $b_l > 0.75$, as theoretical models of biasing indicate, then $\Omega > 0.33$ at 95% confidence. A comparison with other estimates of $\beta_l$, if not dominated by unaccounted-for systematic errors, may provide hints for scale dependence in the biasing relation for IRAS galaxies.

Subject headings: cosmology: theory — galaxies: clusters: general — galaxies: distances and redshifts — large-scale structure of universe

1. INTRODUCTION

A comparison of the galaxy density field derived from a redshift survey with the mass-density fluctuation field inferred from galaxy peculiar velocity data allows one to test both gravitational instability (GI) theory and models for the relation between galaxies and mass. If the data are consistent with the assumed model, one can then estimate the value of the cosmological density parameter $\Omega$. This can be illustrated, for simplicity, by the linear approximation to GI, for which the relation between the mass-density fluctuation field $\delta(x)$ and the peculiar velocity field $\mathbf{v}(x)$ is

$$V \cdot \mathbf{v} = -f(\Omega)\delta, \quad f(\Omega) \approx \Omega^{0.6}, \quad |\delta| \ll 1,$$

where $b$ refers to the galaxies in a specific redshift survey and at a fixed smoothing length. With this model for biasing, equation (1) can be rewritten as a relation between the observable quantities,

$$V \cdot \mathbf{v} = -\beta \delta, \quad \beta \equiv f(\Omega)/b.$$

Thus, in the context of linear GI and linear biasing, the comparison of peculiar velocities and the galaxy distribution enables a measurement of $\beta$. However, $\beta$ provides only an indirect estimate of $\Omega$, because it is a degenerate combination of $\Omega$ and $b$.

There is quite an extensive literature on the comparison of peculiar velocity and redshift survey data to measure $\beta$ (for reviews see Dekel 1994; Strauss & Willick 1995; Strauss 1997a, 1997b; for a general review of $\Omega$ measurements see Dekel, Burstein, & White 1997). The comparison is not straightforward: the peculiar velocity data are sparse, inhomogeneously distributed, limited to the radial component, and quite noisy, while the redshift data need to be corrected to real space and may trace the mass distribution in a nontrivial way. These difficulties give rise to a variety of statistical biases that depend on the details of the analysis carried out.

There are two approaches to this problem, depending on whether the quantities that are actually compared are velocities or densities. Integrating both sides of equation (3) yields a predicted velocity field given measurements of the galaxy density field (eq. [18] below). Comparison of these predictions to observed radial peculiar velocities allows a determination of $\beta$ (Strauss 1989; Kaiser et al. 1991; Hudson 1994). One can make the comparison more sophisticated by smoothing the two velocity fields before comparing them (Davis, Nusser, & Willick 1996), or by using the predicted velocity field to minimize the scatter in the distance indicator relation from which the peculiar velocities are measured in the first place (Strauss 1989; Roth 1994; Schlegel 1995; Shaya, Peebles, & Tully 1995; Willick et al. 1997b). For IRAS galaxies, these velocity comparisons yield values of $\beta_l$ ranging from 0.49 $\pm$ 0.07 (Willick et al. 1997b) to 0.86 $\pm$ 0.14 (Kaiser et al. 1991), depending on the details

1 Racah Institute of Physics, Hebrew University, Jerusalem 91904, Israel; sigad@astro.huji.ac.il, eldara@astro.huji.ac.il, dekel@astro.huji.ac.il.
2 Center for Particle Astrophysics, University of California, Berkeley, CA 94720.
3 Department of Physics/Lick Observatory, University of California, Santa Cruz, CA 95064.
4 Princeton University Observatory, Princeton University, Princeton, NJ 08544; strauss@astro.princeton.edu.
5 Alfred P. Sloan Foundation Fellow and Cottrell Scholar of Research Corporation.
6 Department of Physics and Astronomy, State University of New York, Stony Brook, NY 11794-3800; amos.yahil@sunysb.edu.
of the analysis, the smoothing scale, and the data used. Davis et al. (1996) have claimed inconsistencies between the peculiar velocity and redshift survey data in the context of GI and linear biasing. Alternatively, one can use the POTENT method (§ 2) to recover the density fluctuation field from the peculiar velocity data and use equation (3), or its nonlinear extension (see the discussion in § 2.4), to compare to the galaxy density field. Dekel et al. (1993, hereafter P1993) carried out such an analysis, using the IRAS 1.936 Jy redshift survey (Strauss et al. 1992a), and the Mark II compilation of peculiar velocities (Burstein 1989). The advantages of the differential form include the facts that the direct comparison of densities is local (whereas the velocity field is sensitive to the mass distribution in a large volume, perhaps even outside the sampled volume), it is independent of reference frame, and it allows direct control over the smoothing of the fields. Monte Carlo tests showed that the POTENT density field of P1993 was biased in a variety of ways, forcing the use of an elaborate maximum likelihood technique to quantify the consistency between data and mode and to measure \( \beta_1 \). P1993 did find consistency, and concluded that \( \beta_1 = 1.28^{+0.75}_{-0.53} \) at 95% confidence.

A similar comparison of POTENT densities with optically selected galaxies by Hudson et al. (1995) yielded an acceptable fit, with \( \beta_\text{opt} = 0.74 \pm 0.13 (1 \sigma) \).

The current paper, like P1993, follows the general approach of a density comparison, but is a significant step forward because of a number of improvements in the quantity and quality of the data and the methods of analysis. In particular:

1. The current analysis is based on the Mark III catalog of peculiar velocities (Willick et al. 1995, 1996, 1997a). With \( \sim 3400 \) galaxies, it is more than 3 times as big as the Mark II data set, and has better space coverage. The data sets composing the Mark III catalog have been treated with more care, especially in self-consistently calibrating the Tully-Fisher (TF) relations, grouping, and correcting for inhomogeneous Malmquist bias (§ 2).
2. The POTENT method for deriving the density fluctuation field \( \delta_\text{p} \) from peculiar velocity data has been much improved since P1993 (§ 2; Dekel et al. 1998).
3. The current analysis uses the IRAS 1.2 Jy redshift survey (Fisher et al. 1995), containing twice as many galaxies as in the 1.936 Jy survey.
4. The method for deriving a uniform galaxy density field \( \delta_1 \) from the redshift data has been improved in several ways since P1993 (§ 3).
5. The availability of much more realistic simulations of both the IRAS and Mark III data sets (§ 4.1; Kolatt et al. 1996) allows much better error analysis both in the peculiar velocity and density fields and in the comparison.

We use these simulations to assess biases in our determination of \( \beta_1 \). In contrast to the results of P1993, these biases turn out to be negligible, allowing us to sidestep the rather elaborate likelihood analysis of that paper. Indeed, we will use the simulations themselves as a guide to whether our data are statistically consistent with the null hypothesis of GI and linear biasing. We also use them to quantify the statistical errors in our final results.

The current analysis compares the density fields smoothed with a Gaussian window of radius 12 h\(^{-1}\) Mpc, where the fluctuations are of order unity and therefore require a mildly nonlinear treatment. The POTENT analysis computes the density field \( \delta_1 \), a generalization of \(-j(\Omega)^{-1}V \cdot \mathbf{v} \) that is a nonlinear function of \( \Omega \) and the spatial partial derivatives of the observed \( d(x) \) (eq. [15]). The IRAS reconstruction, in turn, yields a mildly nonlinear galaxy density field \( \delta_1 \), that is a weak function of \( \Omega \) and \( \beta_1 \) only via the corrections from redshift space to real space (eqs. [16] and [17]). Equation (3) is thus replaced by

\[
\delta_1 = b_1 \delta_\text{p} .
\]

The \( \Omega \) dependence of equation (3) is already included in \( \delta_\text{p} \). The density fields are recovered from the data for assumed values of \( \Omega \) and \( b_1 \). Then \( b_1 \) is determined by equation (4), and \( \beta_1 \) is quoted. The analysis is carried out for several initial values of \( \Omega \) and \( b_1 \) to confirm the robustness of the estimate of \( \beta_1 \).

The present work is less ambitious than P1993 in one respect. An attempt was made in P1993 to use the nonlinear effects to break the degeneracy between \( b_1 \) and \( \Omega \), but with only limited success. The resulting constraints on each parameter separately were rather weak, indicating that the nonlinear effects associated with these data are not sufficient for this purpose. Even with the new data, the nonlinear effects are appreciably smaller than the errors that accompany the reconstructions. Furthermore, if we are to consider nonlinear gravity, we should also allow nonlinear extensions to the biasing relation, equation (2), but in practice it is not clear how to distinguish nonlinear gravity effects from nonlinear biasing effects. We therefore limit ourselves in this paper to determining the degenerate combination \( \beta_1 \).

The outline of this paper is as follows: In § 2 we discuss the POTENT reconstruction of the mass-density field from peculiar velocities, and in § 3 the reconstruction of the galaxy density field from the IRAS redshift survey. In § 4 we use mock catalogs to evaluate the random and systematic errors in the reconstructed density fields, and to minimize them if possible. In § 5 we describe our method of comparison of the two fields and estimate the systematic and random errors in the measurement of \( \beta_1 \). In § 6 we perform the comparison of the real data, evaluate goodness of fit, and determine the value of \( \beta_1 \). In § 7 we conclude with the implications of our results and compare them with other recent determinations of \( \beta_1 \).

2. POTENT RECONSTRUCTION FROM PECULIAR VELOCITIES

The POTENT procedure recovers the underlying mass-density fluctuation field from a whole-sky sample of observed radial peculiar velocities. The steps involved are (a) preparing the data for POTENT analysis, including grouping and correcting for Malmquist bias, (b) smoothing the peculiar velocities into a uniformly smoothed radial velocity field with minimum bias, (c) applying the Ansatz of gravitating potential flow to recover the potential and three-dimensional velocity field, and (d) deriving the underlying density field \( \delta_\text{p} \) by an approximation to GI in the mildly nonlinear regime. The revised POTENT method, which grew out of the original method of Dekel, Bertschinger, & Faber (1990, hereafter DBF), is described in detail in Dekel et al. (1998) and is reviewed in the context of other methods by Dekel (1997a, 1997b). We emphasize below the important improvements since P1993.
2.1. Mark III Data

We use the most comprehensive catalog of peculiar velocity data available today, the Mark III catalog (Willick et al. 1995, 1996, 1997a), which is a careful compilation of several data sets of spiral and elliptical galaxies under the assumption that all galaxies trace the same underlying velocity field. The nontrivial procedure of merging the data sets accounts for differences in the selection criteria, the quantities measured, the method of measurement, and the TF calibration techniques. The data per galaxy consist of a redshift z, a “forward” TF (or $D_c(r)$) inferred distance $d$, and an error. The radial peculiar velocity is then $u = cz - d$.

The Mark II data used in P193 consisted of about 1000 galaxies (mostly ellipticals by Lynden-Bell et al. 1988 and spirals by Aaronson et al. 1982). The extended Mark III catalog consists of ~3400 galaxies, including the Mark II data (with improvements to the uniformity of Aaronson et al. 1982 by Tormen & Burstein 1995) as well as newer spiral data sets by Courteau (1992), Willick (1991), Mathewson, Ford, & Buchhorn (1992), and Han & Mould (1992 and references therein). This sample enables a reasonable recovery of the dynamical fields with 12 h$^{-1}$ Mpc smoothing in a sphere of radius ~60 h$^{-1}$ Mpc about the Local Group (LG).

2.2. Correcting Malmquist Bias

Even after the selection bias in the TF calibration is properly corrected, the inferred distance $d$ for a given galaxy, and therefore its peculiar velocity, suffer from a Malmquist bias due to the cross-talk between the random distance errors and the distribution of galaxies about $d$. We correct this bias in a statistical way before using the velocities as input to POTENT.

If the scatter around the TF relation is Gaussian (Willick et al. 1997a), i.e., the absolute magnitude $M$ is distributed normally for a given log width $\eta$ with standard deviation $\sigma_m$, then the forward inferred distance of a galaxy at a true distance $r$ is distributed log normally about $r$, with relative error $\Delta \approx 0.46\sigma_m$. Given $d$, the expectation value of $r$ is

$$E(r | d) = \left[ \frac{\int_0^\infty rP(r | d)dr}{\int_0^\infty P(r | d)dr} \right] = \left[ \frac{\int_0^\infty r^2n(r) \exp \left\{ -\ln[r/d]\right\}/2\Delta^2}{}\right]dr,$$

where $n(r)$ is the number density in the underlying distribution from which the galaxies were selected. The deviation of $E(r | d)$ from $d$ is the bias. The homogeneous part ($n = \text{constant}$), which arises from the three-dimensionality of space, reduces to $E(r | d) = de^{1.5\Delta^2}$ and thus is easy to correct.

Fluctuations in $n(r)$ are responsible for the inhomogeneous Malmquist (IM) bias, which, if uncorrected, will systematically enhance the inferred density perturbations $\delta_p$, and thus the deduced value of $\beta$. The galaxies are first grouped in z-space (Willick et al. 1996), reducing the distance error of each group and thus weakening the bias. The inferred distance of each object is then replaced by $E(r | d)$, with an assumed $n(r)$ properly corrected for grouping and for redshift cutoffs in the data. We use high-resolution density fields of IRAS and optical galaxies (Hudson 1993) for the required input $n(r)$ for spirals and ellipticals, respectively. Tests with mock data demonstrated that the IM bias can be reduced to the level of a few percent in $\delta_p$ (Dekel et al. 1998; Eldar, Dekel, & Willick 1998).

The resultant correction to $\delta_p$ is less than 20% even at the highest peaks, and is much smaller throughout most of space. Thus, the use of IRAS data for the IM correction of the Mark III data introduces only a weak coupling to the POTENT output density field. For the purpose of this paper, the POTENT and IRAS density fields can thus be regarded as essentially independent data sets.

Indeed, a different IM correction procedure in the framework of an “inverse” TF analysis, which does not rely on external input for $n(r)$, yields a density field that is consistent with the result of the forward analysis applied here (Dekel et al. 1998; Eldar et al. 1998; following Landy & Szalay 1992).

2.3. Smoothing the Radial Velocities

The goal of the POTENT analysis is to use as input the set of peculiar velocities $u_i$ at positions $\mathbf{d}_i$, to determine the underlying continuous fields $v(x)$ and $\delta(x)$, both smoothed with a Gaussian of radius $R_s$. We denote hereafter a three-dimensional Gaussian smoothing window of radius $R$ by GR, so that 12 h$^{-1}$ Mpc smoothing is indicated by G12, and so on. The first, most difficult step is the smoothing, or interpolation, into a continuous radial velocity field $u(x)$. The goal is minimum bias compared to the radial component of the true field, had it been sampled noiselessly, densely, and uniformly, and perfectly smoothed in three dimensions. The radial velocity field at $x_i$ is taken to be the value of an appropriate local velocity model $v(x_i, x - x_i)$ at $x = x_i$. The model parameters $\sigma_k$ are obtained by minimizing the weighted sum of residuals,

$$\sum_i W_i [(u_i - \mathbf{x}_i \cdot v(x_i, x_i))]^2,$$

within an appropriate local window $W_i \equiv W(x_i, x_i)$. The window is a Gaussian, modified in such a way that it minimizes the combined effect of the following three types of errors:

1. Tensor window bias.—Because the radial peculiar velocity vectors are not parallel, the $u_i$ cannot be averaged as scalars, and $u(x_i)$ requires a fit of a local three-dimensional velocity model. The POTENT version of DBF used the simplest local model with three parameters,

$$v(x) = B,$$

for which the solution can be expressed explicitly in terms of a tensor window function. However, a bias occurs because the tensorial correction to the spherical window has conical symmetry, and because variations of $v$ within the window, such as a spherical infall pattern, may be interpreted erroneously as a local bulk velocity. This bias amounts to ~300 km s$^{-1}$ at the Great Attractor (see below). The window bias can be reduced by introducing shear into the model,

$$v(x) = B + \mathbf{\tilde{A}} \cdot (x - x_i),$$

where $\mathbf{\tilde{A}}$ is a symmetric tensor that automatically ensures local irrotationality. The additional six terms tend to “absorb” most of the bias, leaving $v(x_i) = B$ less biased. Unfortunately, a high-order model tends to pick undesired small-scale noise. The optimal compromise for the Mark III data was found to be equation (8) out to $r = 40$ h$^{-1}$ Mpc, smoothly changing to equation (7) beyond 60 h$^{-1}$ Mpc (Dekel et al. 1998).
2. Sampling-gradient bias.—Gradients in the true velocity field within the effective window, coupled with the nonuniform sampling, will introduce a bias if the smoothing is not equal-volume weighted. We thus weight each object in proportion to the local volume that it “occupies,” estimated by \( W_i \propto R_i^3 \), where \( R_i \) is the distance to the fourth-nearest object. Simulations show that this reduces the bias to negligible levels for \( r < 60 h^{-1} \) Mpc, except for the very sparsely sampled Galactic zone of avoidance. We will use the \( R_i(x) \) field as a flag for poorly sampled regions, to be excluded from the comparison with IRAS.

3. Random errors.—The ideal weighting for reducing the effect of Gaussian noise would be \( W_i \propto \sigma_i^{-2} \), where \( \sigma_i \) are the distance errors. Unfortunately, this weighting spoils the volume weighting. As a compromise we weight by both, i.e.,

\[
W(x_i, x_j) \propto V_i \sigma_i^{-2} \exp \left[ - (x_i - x_j)^2 / 2R_i^2 \right].
\]  

(9)

2.4. From Radial Velocity to Density Field

Under GI, the large-scale velocity field is expected to be irrotational, \( V \times v = 0 \). This remains a good approximation in the mildly nonlinear regime as long as the field is properly smoothed (Bertschinger & Dekel 1989; DBF). This implies that the velocity field can be derived from a scalar potential, \( \Phi(x) = -V \Phi(x) \), and thus the potential can be computed by integration along the lines of sight,

\[
\Phi(x) = -\int_0^\infty u(r', \theta, \phi) dr'.
\]  

(10)

The two missing transverse velocity components are then recovered by differentiation.

The subsequent derivation of the underlying mass-density fluctuation field requires a solution to the equations of GI in the mildly nonlinear regime. The linear approximation is limited to the small dynamical range between a few tens of megaparsecs and the \( \sim 100 h^{-1} \) Mpc extent of the current samples. Our current peculiar velocity samples enable us to perform reliable dynamical analyses with a smoothing radius as small as \( \sim 10 h^{-1} \) Mpc, where \( |V \cdot v| \) reaches values larger than unity and therefore nonlinear effects play a role. For example, the deviation of the linear approximation, \(-V \cdot v\), from the true density field at \( \delta \sim 1 \) is typically \(-0.4\) (Ganon et al. 1998). We appeal to the Zeldovich (1970) approximation, which is known to be a successful tool in the mildly nonlinear regime. Substituting an Eulerian version of the Zeldovich approximation in the continuity equation yields (Nusser et al. 1991)

\[
\delta(x) = \| I - f^{-1} \hat{\nabla} \hat{\nabla} x \| - 1,
\]  

(11)

where the bars denote the Jacobian determinant and \( I \) is the unit matrix. This was the approximation used in PI93. Tests with \( N \)-body simulations show that it does an excellent job for \( \delta \geq 0 \), but it tends to be an overestimate for \( \delta < 0 \) (Mancinelli et al. 1994; Ganon et al. 1998).

The Zeldovich displacement is first order in \( f^{-1} \) and \( v \), and therefore the determinant in \( \delta \) includes second- and third-order terms as well, involving sums of double and triple products of partial derivatives:

\[
\delta = -f^{-1} V \cdot v + f^{-2} \Delta_2 + f^{-3} \Delta_3,
\]  

(12)

where

\[
\Delta_2(x) = \sum_{i,j} \left\{ \left( \frac{\partial v_i}{\partial x_j} \right)^2 - \frac{\partial v_i}{\partial x_j} \frac{\partial v_i}{\partial x_j} \right\},
\]  

(13)

and

\[
\Delta_3(x) = \sum_{i,j,k} \left( \frac{\partial v_i}{\partial x_j} \frac{\partial v_j}{\partial x_k} \frac{\partial v_k}{\partial x_i} \right).
\]  

(14)

The sum is over the three cyclic permutations of \( (i, j, k) = (1, 2, 3) \). The approximation can be improved by slight adjustments to the coefficients of the three terms in equation (12),

\[
\delta = -(1 + \epsilon_1) f^{-1} V \cdot v + (1 + \epsilon_2) f^{-2} \Delta_2 + (1 + \epsilon_3) f^{-3} \Delta_3.
\]  

(15)

These coefficients were empirically tuned to best fit a family of \( CDMS \) simulations of \( 12 h^{-1} \) Mpc, smoothing over the whole range of \( \delta \)-values, with \( \epsilon_1 = 0.06, \epsilon_2 = -0.13, \) and \( \epsilon_3 = -0.3 \). This approximation is found to be robust to unknown quantities such as the value of \( \Omega \), the shape of the power spectrum, and the degree of nonlinearity as determined by the fluctuation amplitude and the smoothing scale (Ganon et al. 1998). We adopt \( \delta = \delta_{\epsilon^+} \), equation (15), in this paper.

3. THE IRAS RECONSTRUCTION

The redshifts of galaxies in the IRAS sample differ from the true distances by the same peculiar velocities that one is attempting to measure in the Mark III data set,

\[
cz = r + \hat{x} \cdot v(x),
\]  

(16)

where \( cz \) and \( v \) are measured in the same frame of reference.

Because of peculiar velocities, the galaxy density field measured in redshift space, \( \delta_{\epsilon}(x) \), differs systematically from that in real space, \( \delta(x) \) (Kaiser 1987; see reviews in Dekel 1994; Strauss & Willick 1995; Strauss 1999b). Gravitational instability theory enables us to correct for the effects of these velocities. Given the divergence field \( V \cdot v \) and appropriate boundary conditions, the velocity is

\[
v(x) = \frac{1}{4\pi} \int \left[ \frac{V \cdot v(x')}{|x' - x|^3} \right] d^3x'.
\]  

(17)

In the linear regime, equation (3) yields

\[
v(x) = \frac{\beta}{4\pi} \int \delta(x') \left( \frac{x - x'}{|x' - x|^3} \right) d^3x'.
\]  

(18)

Of course, the input density field is defined in real space, which requires that we find a simultaneous solution for the real space density and velocity fields. Yahil et al. (1991, hereafter YSDH) and Strauss et al. (1992b) describe an iterative technique which implements equations (16) and (18) to calculate the peculiar velocity field and convert the redshifts to distances for an assumed value of \( \beta \). Our current implementation of the iteration procedure has three new features: the treatment of triple-valued zones, the filtering of the density field, and the nonlinear corrections.

3.1. Modeling Triple-valued Zones

Given a model for the velocity field from equations (3) and (17), one can compute the redshift-distance relation along a line of sight to a given galaxy from equation (16). This can be compared with the observed redshift of the galaxy to solve for the distance \( r \). However, in the vicinity of prominent overdensities, the redshift as a function of distance can become nonmonotonic, in that there are three different distances corresponding to a given redshift. YSDH
describe a distance-averaging procedure that recognizes the existence of triple-valued zones.

The more general approach applied here is inspired by the VELMOD maximum likelihood analysis of Willick et al. (1997b). Along a given line of sight, we ask for the joint probability distribution for observing a galaxy with redshift $cz$, flux density $F$, and (unknown) distance $r$:

$$P(cz, F, r) = P(cz | r)P(F | r)P(r) .$$  \hfill (19)

The first term is provided by our velocity field model. We assume that the contributions to the velocity field of shot noise, the nonlinearities, and the gravitational influence of material within one smoothing length are incoherent, and thus we model the scatter around the assumed redshift-distance relation as a Gaussian, with dispersion $\sigma_v$:

$$P(cz | r) = \frac{1}{\sqrt{2\pi} \sigma_v} \exp \left(-\frac{[cz - \bar{c} \cdot \nu(x)]^2}{2\sigma_v^2} \right).$$  \hfill (20)

For what follows, we set $\sigma_v = 150$ km s$^{-1}$, independent of position (see Willick et al. 1997b). Like YSDH, we collapse clusters of galaxies to a common redshift.

The second term of equation (19) is given by the luminosity function of galaxies, $\Phi(L)$,

$$P(F | r) = \Phi(L = 4\pi^2 F)dL/d\Phi \propto r^2 \Phi(L) ,$$  \hfill (21)

where the derivative is needed because the probability density is defined in terms of $F$, not $L$.

The third term in equation (19) is given by the galaxy density distribution along the line of sight:

$$P(r) \propto n(x)r^2 \propto [1 + \delta_p(x)]r^2 .$$  \hfill (22)

Finally, given the joint probability distribution, equation (19), we estimate the distance of a given galaxy to be the expectation value,

$$\langle r \rangle = \frac{\int r P(cz, F, r)dr}{\int P(cz, F, r)dr} .$$  \hfill (23)

3.2. Power-preserving Filtering

In a flux-limited sample, the mean number density of objects is a monotonically decreasing function of distance, and therefore the shot noise in the density field and predicted peculiar velocity field both increase with distance. Several approaches to this problem have been taken in the literature, but they all involve variable smoothing, which is not desirable for the purpose of comparing with the uniformly smoothed POTENT output. YSDH set a top-hat smoothing length equal to the mean interparticle spacing, which of course increases with distance from the origin. Nusser & Davis (1994) expand the density field in spherical harmonics and radial Bessel functions. Given a finite number of angular modes (independent of distance), the effective smoothing length again increases with scale. Fisher et al. (1995) perform a similar expansion of the density field, but in addition they apply a Wiener filter to the expansion coefficients.

The Wiener filter (Press et al. 1992) provides the minimum variance reconstruction of the smoothed density field, given a priori knowledge of the underlying power spectrum and the noise (Lahav et al. 1994; Zaroubi et al. 1995). Let the observed smoothed density field be $\delta_p(x)$ and the true underlying density field be $\delta_p(x)$, so that $\delta_p(x) = \delta_p(x) + \epsilon(x)$, where $\epsilon$ is the local contribution from shot noise. We wish to minimize the squared error in the density field by filtering the observed density field. Filtering is best done in Fourier space, $\delta_p(k) = F(k)\delta_d$, where the tilde indicates a Fourier transform. Parseval’s theorem allows us to write the squared error of the filtered density field as

$$\langle [\delta_p(x) - \delta_p(x)]^2 \rangle = \langle [\delta_p^2(k) - F(k)\delta_d(k)]^2 \rangle .$$  \hfill (24)

Minimizing this expression with respect to the filter $F$ gives the Wiener filter,

$$F W = \frac{\langle \delta_p^2 \rangle}{\langle \delta_p^2 \rangle + \langle \epsilon^2 \rangle} .$$  \hfill (25)

If the underlying density field and the noise are both Gaussian-distributed, the Wiener filter provides the most probable solution for the mean underlying field.

However, the Wiener filter has a serious drawback for the present application. The expectation value of the square of the filtered field is

$$\langle F W \delta_p \rangle = \left(\frac{\langle \delta_p^2 \rangle}{\langle \delta_p^2 \rangle + \langle \epsilon^2 \rangle} \right)^2 \langle \delta_p^2 \rangle + \langle \epsilon^2 \rangle = F W \langle \delta_p^2 \rangle .$$  \hfill (26)

Since $F W < 1$, the Wiener-filtered field has a variance that is always smaller than that of the true underlying field (contrast this with the case of the observed field, whose variance is systematically greater than that of the true field, by an amount $\langle \epsilon^2 \rangle$). Even worse, because $\epsilon$ is in general an increasing function of distance in a flux-limited sample, the variance in the Wiener-filtered field is a decreasing function of distance. In fact, when the noise dominates over the signal, the most likely value of $\delta$ tends to the mean value of the prior model, i.e., $\delta \rightarrow 0$. This is unacceptable for the current application, because any such systematic effect in $\delta$ will translate directly into a bias in $\beta$, (see eq. [3]).

With this in mind, Yahil (1994) developed a power-preserving variant of the Wiener filter, defined by

$$F_T = F W^{1/2} .$$  \hfill (27)

A calculation analogous to that of equation (26) shows that

$$\langle F_T^2 \delta_p^2 \rangle = \langle \delta_p^2 \rangle ;$$

that is, the filter preserves the power in the underlying true density field. It is not optimal in the sense of minimum variance, but it is straightforward to show that the ratio of the mean square difference between true and filtered fields for the Yahil and Wiener filters can be expressed in terms of the signal-to-noise ratio as

$$\frac{\langle (\delta_T - F_T \delta_p) \rangle^2}{\langle (\delta_T - F W \delta_p) \rangle^2} = 1 + [(1 + y)^{1/2} - y^{1/2}]^2 , \quad y \equiv \frac{\langle \delta_p^2 \rangle}{\langle \epsilon^2 \rangle} .$$  \hfill (28)

The square root of this ratio approaches unity when the signal dominates, $y \gg 1$. It rises to only 1.08 when $y = 1$, and is still only 1.41 even when the noise dominates, $y \ll 1$. That is, we pay only a mild price in density field errors when replacing the Wiener filter by the Yahil filter, while gaining an unbiased estimate of the variance of the density field everywhere.

Applying any of these filters requires that we first define a continuous density field. We assign the galaxies in the sample (weighted by the inverse of the selection function) via cloud-in-cell to a 128$^3$ Cartesian grid of spacing 2 $h^{-1}$ Mpc, and smooth the density field further with a small Gaussian window of G5 (or G4; see § 4.3). We then apply the Yahil filter to the Fourier transform of the density field,
where $\langle \delta_2^2 \rangle$ and $\langle e^2 \rangle$ are calculated as appropriate for the smoothing window, using as a priori the power spectrum of IRAS galaxies as derived by Fisher et al. (1993).

Of course, the appropriate Yahil filter to apply at each point in real space is a function of the distance from the origin, since the noise term $\langle e^2 \rangle$ is an increasing function of distance. We thus calculate a series of Yahil-filtered density fields, each assuming the shot noise appropriate for one of 13 different distances from the origin: one distance equal to 5, and 20 going to 240 $h^{-1}$ Mpc in steps of 20 $h^{-1}$ Mpc. The value of the Yahil-filtered density field at any other point in space is then spline-interpolated from those 13 fields.

### 3.3. Nonlinear Correction

Because we are calculating a continuous density field, we can take advantage of nonlinear extensions to equation (3) before plugging $V \cdot v$ into equation (17) and solving it in Fourier space. Nusser et al. (1991) provided a simple functional fit to the inverse of equation (11), which does a good job of fitting nonlinear effects for smoothing scales of G5 and larger (see also Mancinelli & Yahil 1995). The Nusser et al. (1991) expression has a mean differing from zero, which gives rise to an erroneous monopole term in the derived density field galaxies. The galaxies are assigned log line widths $\delta_0$, and absolute magnitudes $M$ are randomly scattered about the TF value, $M_{TF}(\eta)$, following a Gaussian distribution of width appropriate to the corresponding subsample of the Mark III catalog. Field galaxies are selected in the angular regions corresponding to each of the subsamples, with the appropriate magnitude limits and redshift cutoffs.

Next, “galaxies” are identified in the simulation and assigned the relevant physical properties. They are then “observed” to make mock catalogs that include all the relevant errors and selection effects in both the peculiar velocity and redshift survey data.

For the mock Mark III catalogs, each of the N-body particles is considered a galaxy candidate, and is identified as elliptical or spiral, depending on the local neighborhood density of particles (following Dressler 1980). Rich clusters are identified, mimicking the cluster samples in the real data, and the remaining particles are left as candidates for field galaxies. The galaxies are assigned log line widths $\eta$ drawn at random from the observed $\eta$ distribution function (corresponding to the observed galaxy luminosity function). A Tully-Fisher (or $D_n-\sigma$) relation is assumed, and absolute magnitudes $M$ are randomly scattered about the TF value, $M_{TF}(\eta)$, following a Gaussian distribution of width appropriate to the corresponding subsample of the Mark III catalog. Field galaxies are selected in the angular regions corresponding to each of the subsamples, with the appropriate magnitude limits and redshift cutoffs. The only feature of the observational procedure that is not simulated is the calibration of the TF relations for each sample in the Mark III catalog, and the matching of them in their overlap; the TF relations are assumed to be known perfectly a priori.

These data are used to infer TF distances to all the galaxies in the mock catalog. The “observed” redshifts are taken to be the true velocities of the particles in the simulation. Finally, the galaxies selected are grouped using the same code used for the real data, and are corrected for Malmquist bias as in § 2.2, using the galaxy number density profile $n(r)$ as derived from a randomly selected mock IRAS catalog (see § 4.3).

### 4. EVALUATING THE RECONSTRUCTION METHODS

The POTENT and IRAS reconstruction methods laid out in the previous sections are evaluated using mock catalogs based on N-body simulations. The systematic errors are investigated and corrected as much as possible, and the random statistical errors are quantified. We first describe the mock catalogs (§ 4.1), then summarize the testing of the POTENT and the IRAS methods themselves (§§ 4.2, 4.3), and eventually describe the calibration of the comparison (§ 5.2).

#### 4.1. The Mock Catalogs

The mock catalogs and the underlying N-body simulations are described in detail in Kolatt et al. (1996); we present only a brief outline here.

A special effort was made to generate simulations that mimic the actual large-scale structure in the real universe, in order to take into account any possible dependence of the errors on the signal. The present-day density field, G5 smoothed, is taken to be that of IRAS galaxies as reconstructed by the method described in § 3, with $b_7 = \Omega = 1$. The field is traced back in time to remove nonlinear effects by integrating the Zeldovich-Bernoulli equation (Nusser & Dekel 1992). Non-Gaussian features are removed, and structure on scales smaller than the smoothing length is added using the method of constrained realizations (Hoffman & Ribak 1991), with the power spectrum of the IRAS 1.2 Jy survey (Fisher et al. 1993) as a prior. The resulting density field is fed as initial conditions to a particle mesh (PM) N-body code (Bertschinger & Gelb 1991), which then follows the nonlinear evolution under gravity, with $\Omega = 1$. The present epoch is defined by an rms density fluctuation of $\sigma_8 = 0.7$ at a top-hat smoothing of radius $8 h^{-1}$ Mpc, as is observed for IRAS galaxies (Fisher et al. 1994a). The periodic box of side 256 $h^{-1}$ Mpc is simulated with a $128^3$ grid and $128^3$ particles.

The mock Mark III catalogs, each of the N-body particles is considered a galaxy candidate, and is identified as elliptical or spiral, depending on the local neighborhood density of particles (following Dressler 1980). Rich clusters are identified, mimicking the cluster samples in the real data, and the remaining particles are left as candidates for field galaxies. The galaxies are assigned log line widths $\eta$ drawn at random from the observed $\eta$ distribution function (corresponding to the observed galaxy luminosity function). A Tully-Fisher (or $D_n-\sigma$) relation is assumed, and absolute magnitudes $M$ are randomly scattered about the TF value, $M_{TF}(\eta)$, following a Gaussian distribution of width appropriate to the corresponding subsample of the Mark III catalog. Field galaxies are selected in the angular regions corresponding to each of the subsamples, with the appropriate magnitude limits and redshift cutoffs. The only feature of the observational procedure that is not simulated is the calibration of the TF relations for each sample in the Mark III catalog, and the matching of them in their overlap; the TF relations are assumed to be known perfectly a priori.

4.2. Errors in the POTENT Reconstruction

The left-hand panel of Figure 1 demonstrates how well POTENT can do with ideal data of dense and uniform sampling and no distance errors. The reconstructed density field, from input that consisted of the exact, G12-smoothed radial velocities is compared with the true G12 density field of the simulation. The comparison is done at grid points of spacing $5 h^{-1}$ Mpc inside our “standard” comparison volume of effective radius $40 h^{-1}$ Mpc (see below). We see that no bias is introduced by the POTENT procedure itself. The small rms scatter of 2.5% reflects the cumulative effects
of small deviations from potential flow, scatter in the nonlinear approximation (eq. [15]), and numerical errors (compare to the discussion of the nonlinear effects on this smoothing scale in P193).

We execute the POTENT algorithm on each of 10 noisy mock realizations of the Mark III catalog, recovering 10 corresponding density fields. The error in the POTENT density field at each point in space, $\sigma_p$, is taken to be the rms difference over the realizations, between $\delta_{p,T}$ and $\delta_T$, the true G12-smoothed density field of the mass in the simulation. This scatter includes both systematic and random errors.

We evaluate the density field and the errors on a Cartesian grid with $5 h^{-1}$ Mpc spacing. In the well-sampled regions, which extend in Mark III out to $40-60 h^{-1}$ Mpc, the errors are $\sigma_p \approx 0.1-0.3$, but they are much larger in certain regions at large distances. We will use $\sigma_p$ below to exclude noisy regions in the POTENT-IRAS comparison (§ 5.1). In particular, our standard comparison volume is defined by $\sigma_p < 0.3$ and $R_s < 9.2 h^{-1}$ Mpc. The resulting volume $V$ has an effective radius $R_s = 40 h^{-1}$ Mpc, defined by $V = (4\pi/3)R_s^3$. The rms of $\sigma_p$ in the standard volume is 0.19.

Some part of these errors is systematic. We can quantify this by comparing $\delta_{p,T}$ to the recovered mean density field, averaged over the mock catalogs at each point. The right-hand panel of Figure 1 makes this comparison at the points of a uniform grid inside the standard volume.

The residuals in this scatter plot ($\langle \delta_p \rangle$ vs. $\delta_T$) are the local systematic errors. Their rms value over the standard volume (and over the realizations) is 0.06. The corresponding rms of the random errors ($\delta_p$ vs. $\langle \delta_p \rangle$) is 0.18. The systematic and random errors add in quadrature to the total error ($\delta_p$ vs. $\delta_T$), whose rms over the realizations at each point, $\sigma_p$, is used in the analysis below.

Systematic errors may also be correlated. We quantify this by performing a regression of $\delta_p$ on $\delta_T$: $\delta_p = m\delta_T + n$, for each realization, and then measuring the scatter around this best-fit line:

$$S' = \frac{1}{N_{\text{grid}}} \sum_{i} \left[ \frac{\delta_{p,i} - (m\delta_{T,i} + n)}{\sigma_{p,i}} \right]^2,$$

which we compute for each realization within the standard volume. This statistic is not necessarily distributed like $\chi^2$ because the local systematic errors may be correlated, but it should still average to $\langle S' \rangle = 1$ as long as the errors of $\delta_p$ about $\delta_T$ average to zero. The average of $S'$ over the 10 realizations is $\langle S' \rangle = 0.94$, with a standard deviation of 0.14 and therefore an error of 0.05 in the mean. The proximity of this average to unity indicates that, indeed, the local errors roughly average to zero.

The main point of Figure 1 is that the global systematic errors in $\delta_p$ as a function of $\delta_T$ are small. The figure shows no systematic deviations from the $y = x$ line. The average of the slopes $m$ of the regression lines over the 10 realizations is $\langle S' \rangle = 0.99$, with a standard deviation of 0.14 and therefore an error of 0.05 in the mean. The proximity of this average to unity indicates that, indeed, the local errors roughly average to zero.

The errors in $\delta_p$ due to POTENT errors should still average to zero.

The main point of Figure 1 is that the global systematic errors in $\delta_p$ as a function of $\delta_T$ are small. The figure shows no systematic deviations from the $y = x$ line. The average of the slopes $m$ of the regression lines over the 10 realizations is $\langle S' \rangle = 0.99$, with a standard deviation of 0.14 and therefore an error of 0.05 in the mean. The proximity of this average to unity indicates that, indeed, the local errors roughly average to zero.

4.3. Errors in the IRAS Reconstruction

Mock IRAS redshift surveys are drawn from the simulation following the observed luminosity function and redshift distribution in the IRAS 1.2 Jy catalog. We allow each particle to be chosen as a galaxy with equal probability, so $b_\Omega = 1$ for these samples. Those regions of the sky not surveyed in the IRAS sample (Strauss et al. 1990) are excluded.

Each of nine mock IRAS redshift catalogs is now put through the reconstruction code described in § 3, assuming $\Omega = b_\Omega = 1$ and G5 smoothing (but see below). The resulting density fields are smoothed further with a Gaussian of $10.9 h^{-1}$ Mpc to yield a nominal total effective smoothing of G12, in order to match the smoothing used in
the POTENT reconstruction. Finally, the density field recovered from each of the nine realizations is compared with the true G12 density field of the simulation.

The average recovered density field is compared with the true field in Figure 2 (left-hand panel). The comparison is done within the standard comparison volume (defined by POTENT errors as before). The average slope of the regression lines, shown in this figure, is 0.92 ± 0.08, where the error quoted is the standard deviation over the nine realizations, i.e., the mean slope is about 3σ away from unity. This bias reflects an effective oversmoothing of the IRAS density field, which is partly due to the cloud-in-cell algorithm used to place galaxies on the grid. We find that this bias is significantly reduced by artificially replacing the desired G5 smoothing with a G4 smoothing. After applying the same additional G10.9 smoothing, the average slope of the regression becomes 0.98 ± 0.08, which we regard as satisfactorily unbiased for the purpose of comparison with POTENT. The average density, with G4 + G10.9 smoothing, is compared with the true G12 density in the right-hand panel of Figure 2.

The total error in the IRAS G12 density field at each grid point, σp, is taken to be the rms value of δp over the realizations. Its rms within the standard volume is σp = 0.09. The random and systematic contributions are estimated to be 0.08 and 0.04, respectively.

As we did with POTENT, we do regressions of each realization of δp on δp, and quantify the scatter S, following equation (30). We find ⟨S^2⟩ = 0.60 with a standard deviation of ±0.18. This is significantly smaller than unity, indicating that there are systematic errors in the determination of δp that do not average to zero. This is not of major concern; when we later evaluate goodness of fit of the real data using an S'-like statistic, we do not assume that it is distributed like χ^2, but rather determine its distribution directly from the mock realizations.

5. METHOD OF COMPARISON

5.1. Measuring β1

Since the mock IRAS and mock Mark III catalogs are both drawn from the same underlying density field of an Ω = 1 simulation with b = 1, a perfect method of comparison should yield β = 1.

As mentioned in §4.2, we limit the comparison to the regions where the recovery is reliable in both data sets. In fact, the POTENT errors are typically twice as large as the IRAS errors, so we simply restrict ourselves to regions with small σp. In addition, in order to minimize the effects of sampling gradient bias, we also avoid regions of sparse sampling as indicated by a large Rs.

We carry out the comparison using three alternative sets of cuts on σp and Rs, defining different volumes of space about the Local Group, in order to test our sensitivity to these limits and the volume surveyed. These three cuts are given in the left-hand section of Table 1, which lists the Rs and σp limits, the number of grid points (of spacing 5 h⁻¹ Mpc) that satisfy this cut, the effective radius R* and the corresponding effective number of independent smoothing volumes N_eff, which we estimate by N_eff = N_grid ∑N.setLayoutManager(“line”){
  strokeLineCap(“round”)
  strokeLineJoin(“round”)
  strokeLineWidth(2.0)
}

\[ N_{\text{eff}} = \frac{1}{12} \sum_{i=1}^{N_{\text{grid}}} \sum_{j=1}^{N_{\text{real}}} \exp \left( -\frac{r_{ij}^2}{2R^2} \right), \]

where \( R = 12 \text{ h}^{-1} \text{ Mpc} \) is the smoothing radius. This range of comparison volumes represents a compromise between our need to avoid noisy and sparsely sampled regions and our wish to include as large a volume as possible in order to reduce cosmic scatter and come closer to a fair sample. We have already introduced

\[ N_{\text{eff}} \] is not simply the ratio of the comparison volume to the effective volume of the smoothing window; it also takes into account the shape of the comparison volume, in the sense that \( N_{\text{eff}} \) becomes appropriately larger as the window deviates from a sphere. In contrast to Hudson et al. (1995), our calculations of errors do not use \( N_{\text{eff}} \); it is provided only as a reference.
the volume of $R_e = 40 h^{-1}$ Mpc as our standard comparison volume.

In our model the variables $\delta_t$ and $\delta_p$ are linear functions of the (unknown) true density field $\delta_T$:

$$ \delta_t = b_t \delta_T + \epsilon_t , $$  

(31)

and

$$ \delta_p = \delta_T + c + \epsilon_p , $$  

(32)

where $\epsilon_t$ and $\epsilon_p$ are independent random variables with dispersions $\sigma_t$ and $\sigma_p$, respectively. In this case, we can estimate $b_t$ and $c$ by minimizing the $\chi^2$-like quantity (Lawley & Maxwell 1971):

$$ \chi^2 = \sum \frac{(\delta_{pi} - b_i^{-1} \delta_{Tii} - c)^2}{\sigma_{pi}^2 + b_i^{-2} \sigma_{Tii}^2} , $$  

(33)

with respect to $b_t$ and $c$, taking into account the errors in both fields. The allowed offset $c$ reflects the zero-point freedom in the TF data and the uncertainty in the mean density in the IRAS field. Remember that $\delta_p$ already includes the factor $f^{-1}(\Omega)$ (eq. [15]), which is why $b_t$ does not appear in this equation.

We sample the density fields every $5 h^{-1}$ Mpc while our smoothing length is $12 h^{-1}$ Mpc; thus the fields are greatly oversampled, $N_{\text{grid}} \gg N_{\text{eff}}$. It is meaningful to define a statistic,

$$ S = \chi^2 / N_{\text{grid}} , $$  

(34)

for estimating goodness of fit, as long as we remember that its probability distribution function is not that of a reduced $\chi^2$; the expectation value may be $\langle \chi^2 \rangle = 1$, but the dispersion reflects $N_{\text{eff}}$ and not $N_{\text{grid}}$. We thus do not use it to determine the statistical error in $b_t^{-1}$ (as was done, for example, in Hudson et al. 1995). We instead determine the distribution of this statistic from the mock catalogs that satisfy our null hypotheses of GI and linear biasing (§ 5.2), and then compare the value of $S$ obtained from the real data with this distribution (§ 6.2).

### 5.2. Testing the Comparison with Mock Catalogs

We carry out the comparison using 90 pairs of mock catalogs, pairing the 10 mock POTENT with the nine mock IRAS realizations. For each of the three comparison volumes, we calculate the mean and standard deviation over the pairs of the quantities $S$ and $b_t^{-1}$, and list them in the middle section of Table 1.

The most important conclusion from the comparison of the mock data is that our method provides an almost bias-free estimate of $b_t^{-1}$. This is consistent with our previous tests showing that the POTENT and IRAS reconstructions themselves are both almost unbiased (§§ 4.2, 4.3). When we combine the two, and fit for $b_t^{-1}$ using equation (33), we find that the values of $\langle b_t^{-1} \rangle$ deviate by only 2%–4% from the correct value of unity.

How significant is this small bias? The 90 pairs of realizations are not all independent because the IRAS errors are smaller than the POTENT errors. We crudely estimate the effective number of independent pairs to be $\sim 20$, so the error in $\langle b_t^{-1} \rangle$ is roughly $\sigma_{b_t^{-1}}/(20)^{1/2} \sim 0.03$, and thus the apparent 3% bias is not significant.

Another important result is that $b_t^{-1}$ is fairly robust to the comparison volume in the simulations. A factor of 3 increase in the comparison volume leads to only a 2% change in the value of $\langle b_t^{-1} \rangle$. One conclusion is that cosmic scatter in the simulations does not play a very important role in the determination of $b_t^{-1}$ by this method. Our formal error in $b_t^{-1}$, the derived scatter in $b_t^{-1}$ between realizations, does not include the cosmic scatter, and thus underestimates the true error, but not by a large amount. This cosmic scatter will be estimated again similarly in the real universe in § 6 by comparing results from different volumes.

The linear fit of equation (33) is affected by the systematic errors between $\delta_t$ and $\delta_T$, and between $\delta_p$ and $\delta_T$, that we estimated using equation (30). Given that $\langle \delta_T \rangle$ was substantially less than unity for the IRAS reconstruction, we expect $\sigma_t$ to be somewhat of an overestimate, and therefore the $S$ of the $\delta_p-\delta_t$ comparison may be somewhat less than unity. As Table 1 shows, this is indeed what we find for the mock catalogs.

As mentioned above, we will not use $S$ and $\chi^2$ statistics to determine the error in $b_t^{-1}$. Instead, we use the scatter of $b_t^{-1}$ over the pairs of mock catalogs as our estimate for the error in the value of $b_t^{-1}$ as derived from the real data. We find that this error is $0.01\pm0.13$ (middle section of Table 1).

For the offset density we find from the mock catalogs small values of $\langle c \rangle = 0.04, 0.03$, and 0.02 for the three comparison volumes from small to large, with a scatter $\sigma_c$ of 0.07, 0.06, and 0.07, respectively. The displacement is thus consistent with zero (although a small displacement is expected, especially at small volumes, due to the fact that the mean density was determined separately in the reconstruction from each IRAS mock catalog).

The left-hand panel of Figure 3 shows a scatter plot of the average of the 10 POTENT realizations versus the average of the nine IRAS realizations for the standard volume. The line of slope 1.03, which is the average of the slopes over the individual pairs of realizations, is shown for reference (it is not necessarily identical to the best-fit line of $\langle \delta_T \rangle$ vs. $\langle \delta_T \rangle$). This figure provides a visual impression of the correlation, in the spirit of Figures 1 and 2. The fact that the slope is close to unity reflects the lack of global bias, and the scatter is the local bias at the grid points.

### Table 1

**Comparison Parameters for Different Volumes**

| Comparison Volume | Mock Data | Real Data |
|-------------------|-----------|-----------|
| $\sigma_p$ | $R_p$ | $N_{\text{grid}}$ | $N_{\text{eff}}$ | $R_e$ | $\langle \delta_T \rangle$ | $\sigma_b$ | $\langle b_t^{-1} \rangle$ | $\sigma_{b_t^{-1}}$ | $S$ | $b_t^{-1}$ |
| 0.20 | 8.0 | 1005 | 10 | 31 | 0.80 | 0.14 | 1.04 | 0.13 | 1.03 | 0.83 |
| 0.30 | 9.2 | 2081 | 18 | 40 | 0.89 | 0.16 | 1.03 | 0.12 | 1.06 | 0.89 |
| 0.40 | 10.0 | 3342 | 26 | 46 | 0.93 | 0.15 | 1.02 | 0.11 | 1.16 | 0.93 |

* In units of $h^{-1}$ Mpc.
For a visualization of the actual scatter that one may expect to see in the real universe, we show in the right-hand panel of Figure 3 the comparison of one arbitrary pair of realizations. The scatter is naturally larger than in the left panel, but the correlation is still very strong.

Figure 4 displays contour maps from the mock catalogs; the contouring is described in the figure caption. The upper and middle rows show G12 density fluctuation fields in the supergalactic plane. The upper row is the average over the realizations of POTENT and IRAS in comparison with the true G12 field, and the middle row shows one random realization of each and the difference between them, in units of $\sigma_p$, defined in equation (35) below. It is interesting to compare these maps to Figure 6, which shows the reconstructions of the real fields. Note that the IRAS errors are smaller than the POTENT errors, so the mock IRAS reconstruction resembles the true one more closely than the POTENT reconstruction. The difference between the IRAS and POTENT fields can be very large, especially at large distances, where we know the POTENT reconstruction is less reliable. For most of the volume, the difference is of low statistical significance.

The bottom row of Figure 4 shows the boundaries of the small ($R_c = 31$ h$^{-1}$ Mpc) and large ($R_c = 46$ h$^{-1}$ Mpc) comparison volumes in the supergalactic plane ($Z = 0$) and two parallel planes ($Z = \pm 20$ h$^{-1}$ Mpc), superposed on the true field. The narrowing of the comparison volume in the zone of avoidance (roughly corresponding to $Y = 0$), due to the large $R_4$, is clearly visible.

6. RESULTS

With the method calibrated using the mock catalogs, we are in a position to compare the POTENT and IRAS reconstructions of the real data. We carry out the comparison in the three comparison volumes of Table 1, using the errors $\sigma_p$ and $\sigma_f$ as estimated from the mock catalogs in the $\chi^2$ expression of equation (33). We start with a visual comparison using maps (§ 6.1), proceed to a more quantitative analysis of the residuals in order to evaluate the goodness of fit (§ 6.2), and conclude with a measurement of $\beta_f$ (§ 6.3).

6.1. Visual Comparison

Figure 5 shows the G12 density fields $\delta_{\Omega} (\Omega = 1)$ and $\delta_f$ in the supergalactic plane. The general correlation is evident: the Great Attractor (GA), Perseus-Pisces (PP), and the void in between all show up both as dynamical entities and as structures in the galaxy distribution.

More quantitative visual comparisons of the two density fields are provided in Figure 6, which shows in three rows three different planes in supergalactic coordinates: the supergalactic plane $Z = 0$ and the planes $Z = -20$ h$^{-1}$ Mpc and $Z = +20$ h$^{-1}$ Mpc. The first and second columns show the G12 density fields of POTENT ($\Omega = 1$) and IRAS, and the third column is the difference between the two, $\Delta \delta \equiv \delta_{\Omega} - (b_f^{-1} \delta_f + c)$, divided by the effective error

$$\sigma_e \equiv (\sigma_p^2 + b_f^{-2} \sigma_f^2)^{1/2},$$

where $b_f$ and $c$ are the best-fit values for the standard volume (§ 6.3; Table 1). The contouring is described in the figure caption.

Although the IRAS map is noticeably less noisy than the POTENT map, they both reveal the same main features. The prominent density peak near $(X, Y, Z) = (-30, 15, 0)$ h$^{-1}$ Mpc in the IRAS field is centered on the Hydracentaurus supercluster, which we associate with the Great Attractor. It is elongated along the $Y$ direction because, under G12 smoothing, the Virgo cluster [located roughly at $(15, 0, 0)$] is an extension of the GA. The GA is clearly seen in the POTENT field roughly in the same region of space, except that it peaks at a different location ($-45, 0, 0$) than the IRAS peak. The GA is also clearly visible in the $Z = -20$ h$^{-1}$ Mpc plane, where the peak appears in a similar location ($-30, 15$) in both panels. There is an apparently statistically significant difference between the two, centered at $(-45, 0, 0)$; see also Figure 5. However, it lies in the zone of avoidance, where the sampling is very
Fig. 4.—Density fluctuation fields of POTENT mass vs. IRAS galaxies from the mock catalogs. The smoothing is G12. Contour spacing is 0.2 in $\delta$; the heavier contour is $\delta = 0$; solid contours refer to $\delta > 0$ and dashed contours to $\delta < 0$. The density is also indicated by shading. Top row: Maps in the supergalactic plane of the average fields over the mock realizations, in comparison with the true G12 density field of the simulation. Middle row: Density fields reconstructed from one pair of realizations, and the difference field $\Delta\delta$ in units of the error $\sigma_{\delta}$, with contour spacing of unity. Bottom row: Boundaries of the $R_e = 31$ and $46\ h^{-1}\ Mpc$ comparison volumes in the supergalactic plane ($Z = 0$) and in two parallel planes ($Z = \pm 20\ h^{-1}\ Mpc$).

Fig. 5.—POTENT mass ($\Omega = 1$) vs. IRAS galaxy density fields in the supergalactic plane, both smoothed G12. The height of the surface plot is proportional to $\delta$. The LG is at the center, the GA on the left, PP on the right, and the Sculptor void in between. The region shown is $160\ h^{-1}\ Mpc$ on a side.
sparse and the sampling gradient bias has the strongest effect. This is why the comparison volume does not extend to this region (Fig. 6, middle row). The agreement between the fields within the comparison volume is seen in Figure 7. Note in particular that very little of the standard comparison volume is discrepant by more than 2σ (see Fig. 8 below). In the $Z = +20 \ h^{-1} \text{Mpc}$ plane, the GA lies outside the comparison volume.

The extended density enhancement on the right of the $Z = 0$ maps is the Perseus-Pisces supercluster, which has two peaks, both outside the comparison volume in either map. The PP and GA superclusters are separated by the large Sculptor void, which is similar in the two maps, both at $Z = 0$ and at $Z = -20 \ h^{-1} \text{Mpc}$. Above the supergalactic plane ($Z = +20 \ h^{-1} \text{Mpc}$), the density field is dominated by a large void, prominent in both POTENT and IRAS. Another density peak, far outside our comparison volume, is the Coma supercluster, centered at (5, 75, 0) in the IRAS map, and at (20, 50, 0) in the POTENT reconstruction. This difference need not worry us, as the POTENT reconstruction is known to be extremely noisy there.

6.2. Goodness of Fit

Only after we convince ourselves that the data are consistent with the assumed model of GI and linear biasing can we estimate the parameter $b^{-1}$ and interpret it properly. We therefore make an effort to assess goodness of fit by more quantitative investigations of the residuals $\Delta \delta$ between the POTENT and IRAS fields. We do it in three ways; by
If we have estimated our errors accurately, and if they are Gaussian-distributed, then we expect the scatter \( \Delta \delta / \sigma_e \) to be Gaussian-distributed with a mean of zero and a variance of unity. The observed differential and cumulative distribution function of this quantity are shown in Figure 8 in comparison with the expected Gaussian. The Gaussian is a reasonable match to the observed distribution; the observed standard deviation of 1.03 (i.e., the square root of the reduced \( \chi^2 \) statistic \( S \) [Table 1, right section]) is only slightly different from unity. Note in particular the reasonable fit in the tails of the distribution. Our conclusion is that the distribution of residuals is fairly consistent with a Gaussian.

Figure 9 shows the observed value of \( S \) within the standard volume in comparison with the distribution of \( S \) over the 90 pairs of mock realizations of the data, representing the null hypothesis of GI and linear biasing. Note that this distribution has no extended tails, and would be well fitted by a Gaussian. The observed value of \( S \) falls 1 \( \sigma \) from the center of the distribution of \( S \) in the mock realizations (Table 1), which is a confirmation for the goodness of fit.

For the smaller and larger comparison volumes, the observed value of \( S \) is at 1.6\( \sigma \) and 1.5\( \sigma \), respectively, indicating a weaker but still acceptable goodness of fit.

It is also important to rule out possible spatial correlations between the residuals, which could be indicative of large-scale systematic errors in the POTENT or IRAS density fields. We define a correlation function of residuals by

\[
S(r) = \frac{\langle \Delta \delta(x)\Delta \delta(x + r) \rangle}{\sigma_e(x)\sigma_e(x + r)} x,
\]

where the average is over pairs of grid points in the comparison volume, with pair separation in the range \( (r, r + \Delta r) \), for \( \Delta r = 5 h^{-1} \text{ Mpc} \).

Figure 10 shows this statistic for the real data in comparison with the 90 pairs of mock catalogs. The correlation checking the Gaussianity of the residuals, by evaluating the mean amplitude of the residuals, and by searching for possible spatial correlations between them.

A visual impression can be gained from Figure 7, which shows a scatter plot of \( \delta_p \) versus \( \delta_l \) at the grid points within the standard volume. Recall that the \( \sim 2000 \) points are highly correlated because of the G12 smoothing, with only \( N_{\text{eff}} \sim 18 \) independent subvolumes. This gives rise to coherent clouds of points in the scatter diagram. Notice the qualitative similarity between this figure and the equivalent one for a single, arbitrary pair of mock realizations (Fig. 3).
would have been a direct measure of the comparison volumes, are POTENT density fields, at G12 smoothing and within included mildly nonlinear corrections in both the POTENT analysis, the simple relation $V \cdot v = -\beta_1 \delta_1$ would have been strictly valid, and the slope of the best-fit line quoted would have been a direct measure of $\beta_1$. Since we have included mildly nonlinear corrections in both the POTENT and the IRAS analyses, we may expect the actual relation between $\delta_p$ and $\delta_1$ to deviate slightly from the linear relation. What we have done, in fact, was to assume $\Omega = 1$ in the POTENT reconstruction (and, less important, $\Omega = b_1 = 1$ in the IRAS conversion from redshift to real space), and then determine $b_1$ [and thus $\beta_1 = f(\Omega)/b_1$] from the best-fit line of $\delta_p$ versus $\delta_1$. The fact that we ended up with a value of $\beta_1$ that is consistent at the $\sim 1 \sigma$ level with the assumed value of unity indicates that this is a self-consistent solution for $\Omega$ and $b_1$ near unity.

In order to validate our result for the more general case of $\beta_1 \sim 1$ where $\Omega$ and $b_1$ may differ from unity, one should redo the N-body simulation of the mock catalogs for values of $\Omega \neq 1$. In the present paper, we estimate the errors for $\Omega < 1$ by scaling the errors determined from the $\Omega = 1$ simulation. We first redo the comparison analysis in the standard volume with assumed values of $\Omega = 0.3$ and $b_1 = 0.5$, which also correspond to $\beta_1 = 1$. This makes a difference mostly in the POTENT reconstruction. We crudely scale the errors $\sigma_p$ in proportion to $f(\Omega)^{-1}$. For this case we obtain $\beta_1 = 1.00 \pm 0.18$ from our fit to the real data in our standard volume. The quoted error is now only a rough estimate based on the increase in $\sigma_p$. This result is consistent with the value $\beta_1 = 0.89 \pm 0.12$, obtained in our standard case of $\Omega = b_1 = 1$. This confirms our assessment that we are indeed measuring $\beta_1$, and are not very sensitive to the actual values of $\Omega$ and $b_1$ that give rise to that value of $\beta_1$.

Finally, we try two cases where the assumed $\beta_1$ is $\sim 0.5$ rather than unity. When we start with $\Omega = 1$ and $b_1 = 2$, we obtain $\beta_1 = 0.93 \pm 0.12$. When we start with $\Omega = 0.3$ and $b_1 = 1$, we obtain $\beta_1 = 0.90 \pm 0.18$ where the error is roughly scaled as before. In both cases, the resultant $\beta_1$ are significantly larger than the “assumed” value of $\beta_1 = 0.5$. This test demonstrates that the result is not determined by the assumed input and the method is discriminatory; we can rule out low values of $\beta_1 \leq 0.5$. Note, however, that the rejection of the case $b_1 = 2$ is stronger than the rejection of the case $\Omega = 0.3$, because the POTENT errors become larger when $\Omega$ is lower.

The above tests strengthen our confidence in the generality of our result from the standard case, $\beta_1 \approx 0.89$. However, the error estimate of $\pm 0.12$ is strictly valid only near $\Omega \sim 1$. If $\Omega$ is lower, the error is expected to be larger. The calibration of the method and the error estimate can be fine-tuned further using full POTENT and IRAS reconstructions from mock catalogs drawn from simulations with different values of $\Omega$ and $b_1$ (work in progress). However, since we have no reason to suspect any significant change in the result, we defer this further testing to a subsequent paper.

We have demonstrated, using the mock catalogs, that the nonlinear corrections (especially those to $\delta_p$, eq. [15]) provide an unbiased value of $\beta_1$. An analogous linear analysis at the same smoothing scale results in an underestimate of $\sim 10\%$ in $\beta_1$ in our simulations. The curvature induced by nonlinear effects in the POTENT-IRAS comparison is completely swamped by the scatter in Figure 7, and therefore cannot be measured to remove the degeneracy between $\Omega$ and $b_1$.

7. DISCUSSION AND CONCLUSIONS

We have compared the density of mass and light as reconstructed from the Mark III catalog of peculiar velocities and the IRAS 1.2 Jy redshift survey, with Gaussian...
smoothing of 12 $h^{-1}$ Mpc and within volumes of effective radii 31–46 $h^{-1}$ Mpc. Our two main conclusions are:

1. The data are consistent with gravitational instability theory in the mildly nonlinear regime and a linear biasing relation between galaxies and mass on these large scales.

2. The value of the corresponding $\beta$ parameter at this smoothing scale is $\beta_I = 0.89 \pm 0.12$. The relative robustness of this value to changes in the comparison volume suggests that the cosmic scatter is no greater than $\pm 0.1$.

This result is consistent with the simplest cosmological model of an Einstein–de Sitter universe ($\Omega = 1$, $\Lambda = 0$) without biasing $\gamma_{\text{IRAS}}$ galaxies, and it also permits somewhat lower values of $\Omega$ and $b_I$. However, values as low as $\Omega \approx 0.3$ would require large-scale antibiasing of $b_I < 0.75$ (with 95% confidence) for a phenomenon that is not easily reproduced in theoretical simulations (e.g., Cen & Ostriker 1992; Evrard, Summers, & Davis 1994).

Although our analysis uses an extension of the Zeldovich approximation, we have not tried to directly measure nonlinear effects in these data to break the degeneracy between $\Omega$ and $b_I$: the effects are unmeasurable given the amount of noise in the POTENT-$\gamma_{\text{IRAS}}$ comparison, and are of the same order as possible nonlinearities in the biasing relation. However, we show that ignoring nonlinear effects in the determination of $\delta_r$ introduces $\sim 10\%$ bias in the determination of $\beta_I$. An interesting area for future work is to see whether we can put an upper limit on the degree of nonlinearity; if we can, this puts a lower limit on $b_I$, and therefore a lower limit on $\Omega$.

As explained in § 1, the current analysis is a significantly improved version of the PI93 analysis done with the earlier Mark II and $\gamma_{\text{IRAS}}$-best 1.936 Jy data, and of the Hudson et al. (1995) comparison to optical galaxies. The result of PI93 was $\beta_I = 1.28^{+0.75}_{-0.55}$ at 95% confidence. The new result is lower, but only by about 1.3 $\sigma$.

The current analysis is superior in many respects. The improvements in the data include denser sampling of a larger volume, a careful procedure of selection, calibration, and merging of the TF data sets, and a better correction for Malmquist bias. The methods of reconstruction were considerably improved, using new techniques and with extensive testing and error analysis using realistic mock catalogs drawn from simulations. The optimization of the comparison method using the mock catalogs led to an unbiased estimate of $\beta$ with well-defined statistical errors.

The quantity $\beta$ has been measured using a variety of techniques from observational data sets (as discussed in the references in § 1); we here discuss other estimates of $\beta_I$ from the Mark III and $\gamma_{\text{IRAS}}$-best 1.2 Jy data.

Estimates of $\beta_I$ from redshift distortions in redshift surveys are almost all within 2 $\sigma$ of our current result (Peacock & Dodds 1994, $\beta_I = 1.0 \pm 0.2$; Fisher et al. 1994b, $\beta_I = 0.45^{+0.3}_{-0.2}$; Fisher, Sharf, & Lahav 1994c, $\beta_I = 1.0 \pm 0.3$; Cole, Fisher, & Weinberg 1995, $\beta_I = 0.5 \pm 0.15$; Hamilton 1995, $\beta_I = 0.7 \pm 0.2$; Heavens & Taylor 1995, $\beta_I = 1.1 \pm 0.3$; Fisher & Nusser 1996, $\beta_I = 0.6 \pm 0.2$).

The scatter between these results is mostly due to differences in the method of analysis, but cosmic scatter, nonlinear effects, and complications in the biasing scheme probably also play a role.

Our current analysis is a comparison at the density level (a $\delta$-$\delta$ comparison). Alternative analyses that performed the comparison at the velocity level ($v$-$v$ comparisons) typically yielded somewhat lower values for $\beta_I$, ranging from 0.5 to 0.9 (Strauss 1989, $\beta_I \approx 0.8$; Kaiser et al. 1991, $\beta_I = 0.86^{+0.15}_{-0.1}$; Roth 1994, $\beta_I = 0.6 \pm 0.2$; Nusser & Davis 1994, $\beta_I = 0.6 \pm 0.2$; Davis et al. 1996, $\beta_I = 0.6 \pm 0.2$; Willick et al. 1997b, $\beta_I = 0.49 \pm 0.07$). We focus here on the two most recent and sophisticated analyses of the $v$-$v$ type, termed ITF and VELMOD.

The inverse TF (ITF) analysis (Davis et al. 1996) is a mode-by-mode comparison of velocity-field models, based on expansion in spherical harmonics and radial Bessel functions, that were fitted in redshift space to a processed version of the inverse TF data from the Mark III catalog and the $\gamma_{\text{IRAS}}$-best 1.2 Jy redshift survey. The ITF analysis finds a good fit between the velocity fields and the model, in the form of a distance-dependent dipole at large distances, in apparent contrast to the good fit obtained here. One possible explanation for this difference is an error in one of the large-scale velocity fields, or in both, that does not propagate to their local derivatives, the density fields. A differential offset in the zero point of the TF relation between data sets of the Mark III catalog, which generates an artificial coherent bulk motion, is one such possibility. An inaccurate definition of the Local Group frame in the $\gamma_{\text{IRAS}}$ reconstruction, or missing density data beyond the sampled volume, are other possible sources of error in the velocities. Another source of error is the somewhat arbitrary adjustments that had to be made to the Mark III data sets before combining them into one system for the ITF analysis (see the discussion in Davis et al. 1996).

If one ignores the poor fit, the ITF analysis yields the formal estimate $\beta_I = 0.6 \pm 0.2$, $\sim 1 \sigma$ lower than our current result.

The local nature of the density-density comparison, and the resulting better goodness of fit, argue that it may provide a more reliable estimate of $\beta_I$.

The VELMOD analysis (Willick et al. 1997b) is a high-resolution $v$-$v$ comparison at a small smoothing scale of 3 $h^{-1}$ Mpc, using about one-quarter of the Mark III data within a volume of radius less than 30 $h^{-1}$ Mpc, and applying a sophisticated likelihood analysis in the linear regime. The comparison revealed a good fit and an estimate of $\beta_I = 0.49 \pm 0.07$. We can see several possible reasons for the difference in the estimates of $\beta_I$.

It is possible that the difference originates from systematic errors that were somehow overlooked despite the successful testing using the same mock catalogs. For example, the VELMOD analysis assumes pure linear theory, and may thus suffer from nonlinear effects that were not taken into account. In fact, it was never fully understood how the linear analysis of VELMOD did well at G3 smoothing where nonlinear effects are expected to be important. One suspicion is that the mock simulation is too smooth on small scales.

Another source for the difference may be the partial data used in the VELMOD comparison, where the comparison is in fact dominated by data within $\sim 20 h^{-1}$ Mpc. The difference in $\beta_I$ can thus partly reflect cosmic scatter in $b_I$, or in the local effective $\Omega$. Table 1 shows that $\beta_I$ increases by an insignificant $1 \sigma$ from the smallest to the largest comparison volume; the VELMOD analysis also found no statistically significant growth of $\beta_I$ with scale. This allows us to put an upper limit on the cosmic scatter of order 0.1.

Finally, a possible explanation for the difference in $\beta_I$ is scale dependence of the biasing relation between Gaussian
smoothing of 3 and 12 $h^{-1}$ Mpc, which would be associated with nonlinear biasing.

Some support for such a trend is found by the SIMPOT analysis (Nusser & Dekel 1998), which fits a parametric model of velocity and density fields and $\beta$ simultaneously to the peculiar velocity and redshift data (a $v$-$\delta$ comparison). This analysis yields $\beta_1 = 1.0 \pm 0.15$ for G12 smoothing, and lower values of $\beta_2$ for smaller smoothing scales, in qualitative agreement with the comparison of the results of the current paper and of VELMOD.

Current theoretical simulations indicate possible scale dependence in the biasing relation between scales of one to a few megaparsecs (e.g., Kauffmann, Nusser, & Steinmetz 1997; cf. Mo & White 1996), but it remains to be seen whether such a trend can arise on scales of 6–12 $h^{-1}$ Mpc.

The biasing scheme, which could be nontrivial in several ways (see Dekel & Lahav 1998), is clearly a bottleneck in the effort to measure the cosmological parameter $\Omega$ via $\beta$; the biasing is inevitably involved whenever galaxy density data are used.

The largest source of error in this analysis lies in the peculiar velocity data, and thus improved peculiar velocity data sets, with better control of systematic errors, denser sampling, and more complete sky coverage, will be of great importance for this work (see the reviews of Strauss 1997a; Giovanelli 1997). It will be interesting to make the comparison of the POTENT maps with the optical redshift data of Santiago et al. (1995, 1996), although a proper calculation of $\beta_1$, $\beta_2$, and $\beta_3$ will yield values similar to our finding here. Finally, more work lies ahead in testing the robustness of our results to the assumed value of $\Omega$ (see the discussion in § 6.3), and carrying out the analysis at other smoothing lengths to look for nonlinear effects.

We acknowledge Tsafir Kolatt for his work on the mock catalogs, Galit Ganon for her work on the nonlinear approximations, and Idit Zehavi for thoughtful comments on the errors. We thank the rest of the Mark III team, David Burstein, Stéphane Courteau, Sandy Faber, and Jeffrey Willick. This paper benefited from earlier work with Edmund Bertschinger, Marc Davis, Michael Hudson, and Adi Nusser. This research was supported in part by the US-Israel Binational Science Foundation grant 95-00330, by the Israel Science Foundation grant 950/95, by NSF grants AST 96-16901, and by a NASA Theory grant at the University of California at Santa Cruz. M. A. S. gratefully acknowledges the support of an Alfred P. Sloan Foundation Fellowship.

REFERENCES

Aaronson, M., et al. 1982, ApJS, 50, 241
Bardeen, J., Bond, J. R., Kaiser, N., & Szalay, A. 1986, ApJ, 304, 15
Bertschinger, E., & Dekel, A. 1989, ApJ, 339, L5
Bertschinger, E., & Gelb, G. 1991, Comput. Phys., 5, 164
Burstein, D. 1989, privately circulated computer files (Mark II)
Cen, R., & Ostriker, J. P. 1992, ApJ, 402, 42
Cole, S., Fisher, K. B., & Weinberg, D. 1995, MNRAS, 275, 515
Courteau, S. 1992, Ph.D. thesis, Univ. California, Santa Cruz
da Costa, L. N., Freudling, W., Wegner, G., Giovanelli, R., Haynes, M. P., & Salzer, J. J. 1996, ApJ, 468, L5
nda Costa, L. N., Nusser, A., Freudling, W., Giovanelli, R., Haynes, M. P., Salzer, J. J., & Wegner, G. 1997, MNRAS, submitted (astro-ph/9707299)
Davies, M., Nusser, A., & Willick, J. A. 1996, ApJ, 463, 22
Dekel, A. 1994, MNRAS, 267, 271
Dekel, A. 1997, MNRAS, 286, 515
Kolatt, T., Dekel, A., Ganon, G., & Willick, J. A. 1996, ApJ, 458, 419
Lahav, O., Fisher, K. B., Hoffman, Y., Scharf, C. A., & Zaroubi, S. 1994, ApJ, 423, 193
Landy, S. D., & Szalay, A. S. 1992, ApJ, 391, 494
Lawley, D. N., & Maxwell, A. E. 1971, Factor Analysis as a Statistical Method (New York: Elsevier)
Lynden-Bell, D., Faber, S. M., Burstein, D., Davies, R. L., Dressler, A., Terlevich, R. J., & Wegner, G. 1988, ApJ, 326, 19
Mancinelli, P. J., Yahil, A., & Steinmetz, M. 1997, MNRAS, 286, 795
Ostriker, J. P., Yahil, A., Ganon, G., & Dekel, A. 1994, in Cosmic Veloc-
ity Fields, ed. F. Bouchet & M. Lachieze-Rey, (Gif-sur-Yvette: Editions Frontières), 215
Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. H. 1992, Numerical Recipes: The Art of Scientific Computing (2d ed.; Cambridge: Cambridge Univ. Press)
Roth, J. R. 1994, in Cosmic Velocity Fields, ed. F. Bouchet & M. Lachieze-Rey (Gif-sur-Yvette: Editions Frontières), 233
Shaya, E. J., Peebles, P. J. E., & Tully, R. B. 1996, ApJ, 454, 15
Strauss, M. A. 1989, Ph.D. thesis, Univ. California, Berkeley
Strauss, M. A. 1997a, in Critical Dialogues in Cosmology, ed. N. Turok (Singapore: World Scientific), in press (astro-ph/9610032)
Strauss, M. A. 1997b, in Formation of Structure in the Universe, ed. A. Dekel & J. P. Ostriker (Cambridge: Cambridge Univ. Press), in press (astro-ph/9610033)
Strauss, M. A., Davis, M., Yahil, A., & Huchra, J. P. 1990, ApJ, 361, 49
Strauss, M. A., Huchra, J. P., Davis, M., Yahil, A., Fisher, K. B., & Tonry, J. 1992a, ApJS, 83, 29
Strauss, M. A., & Willick, J. A. 1995, Phys. Rep., 261, 271
Strauss, M. A., Yahil, A., Davis, M., Huchra J. P., & Fisher, K. B. 1992b, ApJ, 397, 395
Tormen, G., & Burstein, D. 1995, ApJS, 96, 123
Willick, J. 1991, Ph.D. thesis, Univ. California, Berkeley

Willick, J. A., Courteau, S., Faber, S. M., Burstein, D., & Dekel, A. 1995, ApJ, 446, 12
Willick, J. A., Courteau, S., Faber, S. M., Burstein, D., Dekel, A., & Kolatt, T. 1996, ApJ, 457, 460
Willick, J. A., Courteau, S., Faber, S. M., Burstein, D., Dekel, A., & Strauss, M. A. 1997a, ApJS, 109, 333
Willick, J. A., Strauss, M. A., Dekel, A., & Kolatt, T. 1997b, ApJ, 486, 629
Yahil, A. 1994, in Evolution of the Universe and Its Observational Quest, ed. K. Sato (Tokyo: Universal Academy Press), 197
Yahil, A., Strauss, M. A., Davis, M., & Huchra J. P. 1991, ApJ, 372, 380 (YSDH)
Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O. 1995, ApJ, 449, 446
Zeldovich, Ya. B 1970, A&A, 5, 84