

# USING CLUSTER ABUNDANCES AND PECULIAR VELOCITIES TO TEST THE GAUSSIANTY OF THE COSMOLOGICAL DENSITY FIELD

WEIHSUEH A. CHIU

Joseph Henry Laboratories, Department of Physics, Princeton University, Princeton, NJ 08544

AND

JEREMIAH P. OSTRIKER AND MICHAEL A. STRAUSS<sup>1</sup>

Princeton University Observatory, Princeton, NJ 08544

Received 1997 August 18; accepted 1997 October 1

## ABSTRACT

Comparing the frequency of typical events with that of unusual events allows one to test whether the cosmological density distribution function is consistent with the normally made assumption of Gaussianity. To this end, we compare the consistency of the tail-inferred (from clusters) and measured values (from large-scale flows) of the rms level of mass fluctuations for two distribution functions: a Gaussian, and a texture (positively skewed) probability distribution function (PDF). We find that if we average the recent large-scale flow measurements, observations of the rms and the tail at the  $10 h^{-1}$  Mpc scale disfavor a texture PDF at  $\sim 1.5 \sigma$  in all cases. If we take the most recent measurement of the rms, that from Willick et al., the comparison disfavors textures for low  $\Omega_0 = 0.3$ , and disfavors Gaussian models if  $\Omega_0 = 1$  (again at  $\sim 1.5 \sigma$ ). Predictions for evolution of high-temperature clusters can also be made for the models considered, and, as is known (e.g., Henry), strongly disfavor  $\Omega_0 = 1$  Gaussian models, while we find  $\Omega_0 = 1$  marginally disfavored in texture models. Taking the suite of tests as a whole, and using all of the quoted data, it appears that textures are strongly disfavored and only the low- $\Omega_0$  Gaussian models are consistent with all the data considered. But given evidence for the internal inconsistency of the observational data, had we only used the recent Willick et al. results, the strength of our conclusion would have been reduced to the  $\sim 1 \sigma$  level.

*Subject headings:* cosmology: theory — galaxies: abundances — galaxies: clusters: general — large-scale structure of universe — methods: analytical

## 1. INTRODUCTION

All current scenarios for the formation of structure in the universe assume that cosmic structures grow from initially small seed perturbations. A fundamental assumption underlying a large class of these theories is that the initial Fourier modes have random phases. This implies that the statistics of the initial density field  $\rho_i(x)$  are fully specified by the correlation function  $\xi_i(r) = \langle \rho_i(x) \rho_i(x+r) \rangle_\infty$  or equivalently its Fourier transform, the power spectrum  $P_i(k)$ . For any given smoothing length, then, the one-point probability distribution function (PDF) of the initial density field  $\rho_i$  is Gaussian. A measurement of the frequency of rare events (clusters of galaxies) coupled with a measurement of the level of rms fluctuations can tell us directly whether or not the distribution of density fluctuations satisfy the Gaussian hypothesis.

While theories based on the inflationary paradigm generally predict Gaussian initial conditions, there are many alternatives that do not. Global textures and other topological defects, for example, predict a strongly non-Gaussian distribution for the initial perturbations (e.g., Gooding, Spergel, & Turok 1991; Park, Spergel, & Turok 1991; Gooding et al. 1992; Pen, Spergel, & Turok 1994; Pen, Seljak, & Turok 1997). Global textures arise when a global non-Abelian symmetry, such as SU(2), is spontaneously and completely broken. When these defects collapse, energy gradients in the Higgs field perturb the metric and induce gravitational fluctuations in the matter fields. These texture knots, then, act as “causal seeds” of structure

formation. We take the texture model as our standard non-Gaussian model, since it is well studied and still viable as a theory of structure formation. In general, texture and Gaussian models give similar predictions for the *shape* of the power spectrum and the level of CMB fluctuations (Pen et al. 1994; Pen & Spergel 1995). In fact, it is often noted (e.g., Park et al. 1991) that texture and Gaussian cold dark matter (CDM) models are *most* different in their PDFs. If it could be demonstrated that the initial density field were Gaussian on some suitable scale, the texture model would be directly falsified.

The standard method for testing Gaussianity relies on measuring the  $N$ th order reduced moments, or cumulants  $\lambda_N$  of the present-day PDF, such as skewness or kurtosis, from redshift or peculiar velocity surveys (see Strauss & Willick 1995 for a review). The variance is  $\sigma^2 \equiv \lambda_2 \equiv \langle \delta^2 \rangle$ , the skewness is  $\lambda_3 \equiv \langle \delta^3 \rangle$ , and the kurtosis is  $\lambda_4 \equiv \langle \delta^4 \rangle - 3\sigma^4$ , where  $\delta$  is the density contrast  $\rho/\langle \rho \rangle - 1$ . For a Gaussian random field, all the cumulants of order  $N > 2$  are equal to zero. The growth of Gaussian initial perturbations under gravitational instability naturally produces non-Gaussianity in the distribution function of the density field. In particular, the cumulants follow a scaling relation  $\lambda_N = S_N \lambda_2^{N-1}$ , where the  $S_N$  are independent of scale in the mildly nonlinear regime for power-law power spectra (Fry 1984a, 1984b; Bernardeau 1992), and can be calculated from  $(N-1)$ th order perturbation theory (Bernardeau 1994). Significant departures from scale invariance are taken as signs of non-Gaussianity. Unfortunately, the higher order moments are very sensitive to finite-volume effects and non-Gaussian noise in the data (see the discussion in Kim & Strauss 1998 and references therein).

<sup>1</sup> Alfred P. Sloan Foundation Fellow, and Cottrell Scholar of Research Corporation.

Since many non-Gaussian distribution functions have long tails, it is better to *directly* compare the rms and the tail of the PDF, rather than to rely on the moments, which are integrals over the PDF. Given a distribution function, the level of tail fluctuations maps directly to a level of rms fluctuations; the tail-inferred rms, then, must be consistent with the directly measured rms. Furthermore, one should look for measures that are, as closely as possible, indicators of the initial, linear density field.

In this paper, we use the observed temperature distribution of clusters of galaxies as a measure of the tail of the distribution of initial density fluctuations, and galaxy peculiar velocities and redshift-space distortions as measures of the rms. Clusters of galaxies are the most massive virialized objects in the universe, with densities a hundred or more times that of the mean density of the universe, containing only a few percent of all galaxies. They are thus extreme excursions from the mean, and hence are measures of the high-potential tail of the distribution of fluctuations. We compare this to the rms level of peculiar velocity fluctuations. In Fourier space, peculiar velocities are more weighted toward lower  $k$  than density ( $v_k \sim \delta_k/k$ ), and hence are less affected by nonlinear growth. Further, the rms is not as sensitive to nonlinear growth as higher order statistics are (A. Dekel 1997, private communication). Both cluster and peculiar velocity observations are at roughly the same spatial scale,  $\sim 10 h^{-1}$  Mpc. Furthermore, they both measure gravitational potential fluctuations, and therefore have a similar dependence on  $\Omega_0$ , which can be approximated by defining

$$\eta_R \equiv \sigma_R \Omega_0^{0.6}, \quad (1)$$

where  $\sigma_R$  is the rms mass fluctuation on *top-hat* filter scale  $R$ . The remaining  $\Omega_0$  dependence is  $\leq 5\%$  for the peculiar velocity and redshift distortion measurements, and  $\leq 25\%$  in the range  $0.3 \leq \Omega_0 \leq 1.0$  for the cluster measurements. This is the quantity we will attempt to determine throughout our analysis, comparing estimates of it found by various methods.

In § 2 we examine various recent derivations of  $\eta_{10}$  from the Mk III peculiar velocity catalog (Willick et al. 1995, 1996, 1997a), and from *IRAS* galaxy redshift-space distortions. In § 3 we infer  $\eta_{10}$  from the redshift-zero X-ray cluster abundances, assuming a Gaussian and texture PDF, using a modified Press-Schechter approach. We calculate the  $\eta_{10}$  separately for  $\Omega_0 = 1.0$ , cosmologically flat ( $\Lambda$ -dominated)  $\Omega_0 = 0.3$ , and open  $\Omega_0 = 0.3$ . In § 4 we discuss the consistency of the rms and the tail-inferred measures of  $\eta$  by performing a likelihood analysis of the Gaussian and texture hypotheses. We also discuss further constraints from observations of clusters at moderate redshift. We summarize in § 5.

## 2. ROOT MEAN SQUARE MASS FLUCTUATIONS FROM PECULIAR VELOCITIES AND REDSHIFT DISTORTIONS

It has been noted for more than a decade that the rms optical galaxy counts in  $8 h^{-1}$  Mpc spheres is  $\langle (\delta N/N)_8^2 \rangle^{1/2} \equiv \sigma_{8,\text{Optical}} \approx 1$  (Davis & Peebles 1983). But given that galaxy formation is still not very well understood, it has become customary to relate this measured variance to the more fundamental rms fluctuations in the matter density  $\sigma_8$  by a bias parameter  $b$ . For optical galaxies,  $b_{\text{Optical}}$  is defined by  $\sigma_{8,\text{Optical}} \equiv b_{\text{Optical}} \sigma_8$ . The bias, however, is clearly dependent on the galaxy type (for example, as we will see below  $\sigma_{8,\text{IRAS}} = 0.69 \pm 0.04$ ; Fisher

et al. 1994a), and may be dependent on scale. We wish to estimate the rms value of fluctuations independent of bias. To do this, we look at recent analyses of peculiar velocity and redshift survey data.

### 2.1. Recent Measurements of Root Mean Square Fluctuations

The peculiar velocity  $\mathbf{v}$  is the deviation of the proper velocity of a galaxy from the local Hubble flow. Assuming that this measured velocity follows from gravitational potential fluctuations, we obtain from linear perturbation theory (Peebles 1993):

$$\nabla \cdot \mathbf{v} = -H_0 f(\Omega_0) \delta \approx -H_0 \Omega_0^{0.6} \delta. \quad (2)$$

Here  $\delta$  is the mass density contrast  $\delta(x) \equiv \rho_{\text{lin}}(x)/\bar{\rho} - 1$ , and the approximation is good to 5%. Now consider filtering the velocity and density fields with a top hat of radius  $R$ , yielding the filtered fields  $\mathbf{v}_R$  and  $\delta_R$ . Then  $\nabla \cdot \mathbf{v}_R$ , the divergence of the filtered velocity field, is proportional to the filtered mass density field  $\delta_R$ . Taking the rms values of  $\nabla \cdot \mathbf{v}_R$  and  $\delta_R$  gives

$$H_0^{-1} \langle (\nabla \cdot \mathbf{v}_R)^2 \rangle^{1/2} \approx \Omega_0^{0.6} \langle \delta_R^2 \rangle^{1/2} \equiv \Omega_0^{0.6} \sigma_R \equiv \eta_R. \quad (3)$$

The peculiar velocity field thus provides a direct measurement of mass fluctuations, appearing with the cosmological density parameter  $\Omega_0$ . Note that in linear theory, where individual Fourier components evolve independently, it makes no difference in which order one filters and takes the divergence of the velocity field.

Seljok & Bertschinger (1994) made an early attempt to measure  $\eta_R$  directly using the POTENT method (Bertschinger et al. 1990) to reconstruct the mass density and peculiar velocity fields from the Mk II catalog (D. Burstein 1989, private communication). They derived a value of  $\eta_8 = 1.3 \pm 0.3$ . More recently, several workers have used the Mk III catalog of peculiar velocities (Willick et al. 1995, 1996, 1997a) to derive values of  $\eta$ .

Zaroubi et al. (1997) perform a likelihood analysis of the peculiar velocities given families of CDM-type models. Because we are making a comparison only with clusters of galaxies, which are measures of fluctuations on scales of roughly  $10 h^{-1}$  Mpc, we only consider their *COBE*-independent results. Zaroubi et al. (1997) work with the measured peculiar velocities of individual galaxies, without any smoothing, so it is unclear to which scale they are most sensitive. We simply integrate their power spectra with  $10 h^{-1}$  Mpc top-hat smoothing to obtain the rms fluctuation on that scale. Their “maximum likelihood” model has  $\eta_{10} = 0.70$ , and the models on their “65% confidence” level contour span the range  $\eta_{10} = 0.61\text{--}0.78$ . We therefore take their measurement to be  $\eta_{10} = 0.70 \pm 0.09$ .

Kolatt & Dekel (1997) use the POTENT technique (Dekel, Bertschinger, & Faber 1990) to determine the quantity  $\nabla \cdot \mathbf{v}_R$  from the Mk III peculiar velocities, to compute the mass power spectrum directly, subtracting out a model for the noise power spectrum derived from mock catalogs. The power spectrum they give is hence linear and unsmoothed. Since POTENT applies a Gaussian filter of filter length  $12 h^{-1}$  Mpc, equivalent to a  $\sim 23 h^{-1}$  Mpc top-hat filter (see Appendix), we “refilter” their observed power spectrum, to derive a value of  $\eta_{23} = 0.32 \pm 0.11$ .

If we assume linear biasing,  $\delta_{\text{galaxies}} = b \delta_{\text{dark matter}}$ , comparison of peculiar velocity and galaxy density field data via

equation (2) allows one to determine the quantity  $\beta \equiv \Omega_0^{0.6}/b$ , the proportionality constant between the peculiar velocity and galaxy density fields for a particular set of galaxies. Note that if bias is independent of scale,  $\beta$  should be as well. Following equation (3) and the definition of  $b$ , we find that for a comparison with *IRAS* galaxies

$$\eta_R = \beta_{IRAS} \sigma_{R,IRAS}. \quad (4)$$

For  $\sigma_{R,IRAS}$ , we use the power-law analytic approximation of Fisher et al. (1994a):

$$\sigma_R = \left[ \frac{72(r_0/R)^\gamma}{2^\gamma(3-\gamma)(4-\gamma)(6-\gamma)} \right]^{1/2};$$

$$r_0 = 3.76^{+0.20}_{-0.23} h^{-1} \text{ Mpc}; \quad \gamma = 1.66^{+0.12}_{-0.09}. \quad (5)$$

Several groups have determined  $\beta$  from comparisons of *IRAS* density and peculiar velocity fields. Willick et al. (1997b) perform a likelihood analysis of Tully-Fisher observables of the Mk III data given a prior velocity model, and find  $\beta_{IRAS} = 0.49 \pm 0.07$ . Their analysis uses an effective Gaussian smoothing of  $\sim 4 h^{-1}$  Mpc and so corresponds roughly to an  $8 h^{-1}$  Mpc top hat. Using  $\sigma_{8,IRAS} = 0.69 \pm 0.04$ , we obtain  $\eta_8 = 0.34 \pm 0.05$ . In another approach, Sigad et al. (1997) compare the POTENT-derived peculiar velocity divergence field to the *IRAS* galaxy density field (with effective top-hat smoothing of  $\sim 23 h^{-1}$  Mpc due to the POTENT method) and derive  $\beta = 0.89 \pm 0.12$ . Multiplying by  $\sigma_{23,IRAS} = 0.29 \pm 0.02$  leads to  $\eta_{23} = 0.26 \pm 0.04$ .

Finally, as first pointed out by Kaiser (1987), one can use the anisotropy of the redshift-space correlation function of a redshift survey to infer the value of  $\beta$ . Fisher et al. (1994b) measured the anisotropy of the *IRAS* 1.2 Jy sample at an effective scale of  $10\text{--}15 h^{-1}$  Mpc, obtaining  $\beta = 0.45^{+0.27}_{-0.18}$ . Using the same procedure as above to convert  $\beta$  to  $\eta$ , we obtain  $\eta_{12.5} = 0.21 \pm 0.11$ . Cole, Fisher, & Weinberg (1994, 1995) use instead the power spectrum anisotropy, at an effective scale of  $\sim 35 h^{-1}$  Mpc, to obtain  $\beta = 0.52 \pm 0.13$ , which results in  $\eta_{35} = 0.11 \pm 0.03$ . Yet another method involves decomposing the redshift-space distribution into spherical harmonics. Fisher, Scharf, & Lahav (1994) (see Heavens & Taylor 1995) show that the inferred value of  $\beta$  depends sensitively on the assumed power spectrum. Scaling their results for each of the power spectra they use, we find that the results for  $\eta$  roughly converge at  $30 h^{-1}$  Mpc, with a value  $\eta_{30} = 0.20 \pm 0.03$ .

Table 1 summarizes the values of  $\eta_R$  cited above, all using rms fluctuations, while Figure 1 plots all the recent measurements, with their respective errors, at their effective top-hat filter scales. Most current theories of structure for-

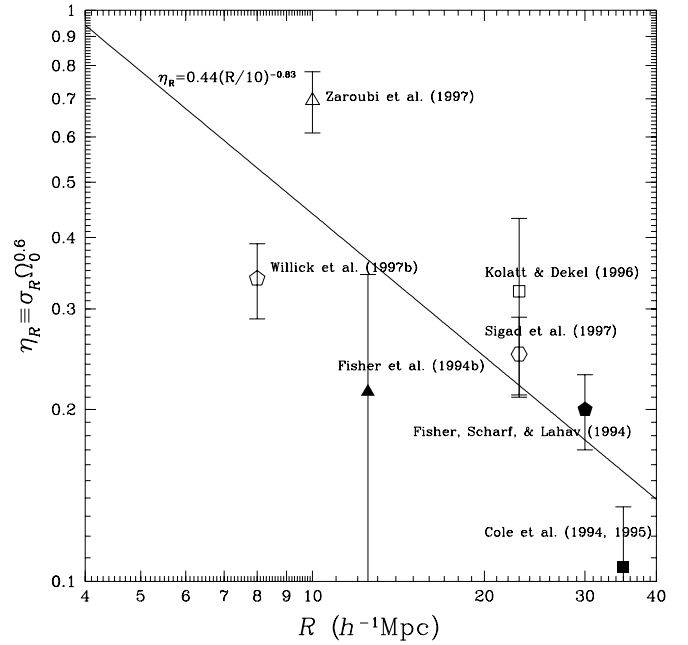


FIG. 1.—Root mean square mass fluctuations as inferred from peculiar velocity and redshift distortion analyses, as described in § 2, and listed in Table 1. The solid line is  $\eta_R = 0.44(R/10.0 h^{-1} \text{ Mpc})^{-0.83}$ , where  $\alpha$  is that for *IRAS* galaxies and  $\eta_{10}$  is taken from the maximum likelihood average in Table 1.

mation predict that  $\sigma_R$  is approximately a power law in the range  $8 < R < 35 h^{-1}$  Mpc. We thus approximate  $\eta_R$  as a power law  $\eta_R \propto R^{-\alpha}$ , to allow us to extrapolate the results of Table 1 and Figure 1 to a common scale. In linear theory, a standard CDM model with  $n = 1$  and  $\Omega_0 = 1$  has a logarithmic slope at  $10 h^{-1}$  Mpc of  $\alpha \approx 1.0$ , while a tilted  $n = 0.7$ , low-density  $\Omega_0 = 0.3$  CDM model has  $\alpha \approx 0.6$ . An  $\Omega_0 = 1$  texture model (using the power spectra of Pen et al. 1994 and Pen & Spergel 1995) gives  $\alpha \approx 1.2$  at  $10 h^{-1}$  Mpc. If galaxy bias does not change appreciably in this range, then  $\alpha$  is the same as that for galaxies. Equation (5) implies an exponent  $\alpha = 0.83 \pm 0.12$  for galaxy fluctuations. If the true value of  $\alpha$  for mass fluctuations is much less (greater) than 0.83, then bias must decrease (increase) with increasing scale. When extrapolating the above results to different scales, then, we take three different power laws to reflect this uncertainty in the slope of the fluctuation spectrum,  $\eta_R \propto R^{-\alpha}$ :  $\alpha = 0.83, 0.60$ , and  $1.10$ . Table 1 contains extrapolations of  $\eta_R$  to  $10 h^{-1}$  Mpc for each of these assumed values for  $\alpha$ . It is striking that the different methods give answers that differ from one another by significantly more than the quoted errors, a sure sign that systematic errors of

TABLE 1  
ROOT MEAN SQUARE MEASUREMENTS OF  $\eta_{10}$

REFERENCE	R	$\eta_R$	$\eta_{10}$		
			$\alpha = 0.60$	$\alpha = 0.83$	$\alpha = 1.10$
Zaroubi et al. 1997.....	10	$0.70 \pm 0.09$	$0.70 \pm 0.09$	<b><math>0.70 \pm 0.09</math></b>	$0.70 \pm 0.09$
Kolatt & Dekel 1996 .....	23	$0.32 \pm 0.11$	$0.53 \pm 0.18$	<b><math>0.64 \pm 0.22</math></b>	$0.80 \pm 0.27$
Willick et al. 1997b .....	8	$0.34 \pm 0.05$	$0.30 \pm 0.04$	<b><math>0.28 \pm 0.04</math></b>	$0.27 \pm 0.04$
Sigad et al. 1997.....	23	$0.25 \pm 0.04$	$0.41 \pm 0.07$	<b><math>0.50 \pm 0.08</math></b>	$0.62 \pm 0.10$
Fisher et al. 1994b .....	12.5	$0.21 \pm 0.13$	$0.25 \pm 0.15$	<b><math>0.26 \pm 0.16</math></b>	$0.27 \pm 0.17$
Cole et al. 1994, 1995 .....	35	$0.11 \pm 0.03$	$0.22 \pm 0.06$	<b><math>0.30 \pm 0.08</math></b>	$0.42 \pm 0.12$
Fisher, Scharf, & Lahav 1994.....	30	$0.20 \pm 0.03$	$0.39 \pm 0.06$	<b><math>0.50 \pm 0.07</math></b>	$0.67 \pm 0.10$
Maximum likelihood average .....	10	...	$0.39 \pm 0.12$	<b><math>0.44 \pm 0.12</math></b>	$0.51 \pm 0.15$

unknown origin are important. We now present a method for combining these measurements.

## 2.2. An Error Model for the Different Measurements of $\eta$

Let us assume that each of the seven measurements (labeled  $i = 1, \dots, 7$ ) listed in Table 1 has an unknown systematic error  $\mu_i$  from the true value of  $\eta$ . That is, each one *actually* measures  $\eta + \mu_i$ , rather than  $\eta$ , where the stated statistical error  $\epsilon_i$  is the error in the measurement of  $\eta + \mu_i$ . We assume that the systematic error is drawn from a Gaussian distribution with zero mean and unknown variance  $\theta^2$ , and that the statistical errors  $\epsilon_i$  are also Gaussian distributed. The probability of a measurement  $x_i$ , given unknown  $\eta$  and  $\theta$ , is given by

$$\mathcal{P}(x_i | \eta, \theta) \equiv \int d\mu_i \mathcal{P}(x_i | \mu_i, \eta) \mathcal{P}(\mu_i | \theta) \quad (6)$$

$$= \frac{1}{\sqrt{2\pi(\theta^2 + \epsilon_i^2)}} \exp \left[ -\frac{(x_i - \eta)^2}{2(\theta^2 + \epsilon_i^2)} \right]. \quad (7)$$

The systematic error simply adds an error  $\theta$  in quadrature to each measurement error  $\epsilon_i$ .

If we define the likelihood function of the unknowns  $\eta$  and  $\theta$  in the usual way, we obtain

$$\ln \mathcal{L}(\eta, \theta) \equiv \ln \left[ \prod_i \mathcal{P}(x_i | \eta, \theta) \right] \quad (8)$$

$$= -\frac{1}{2} \sum_i \left\{ \frac{(x_i - \eta)^2}{\theta^2 + \epsilon_i^2} + \ln [2\pi(\theta^2 + \epsilon_i^2)] \right\}. \quad (9)$$

The maximum likelihood value for  $\eta$  is obtained by setting the derivative of equation (9) to zero:

$$\hat{\eta} = \left( \sum_i \frac{x_i}{\theta^2 + \epsilon_i^2} \right) \times \left( \sum_i \frac{1}{\theta^2 + \epsilon_i^2} \right)^{-1}. \quad (10)$$

Note that equation (10) gives the traditional weighted average for  $\theta \ll \epsilon_i$ , and an unweighted average for  $\theta \gg \epsilon_i$ . Since we are primarily concerned with  $\eta$ , we numerically integrate the likelihood  $\mathcal{L}(\eta, \theta)$  over  $\theta$  to obtain the marginal likelihood over  $\eta$

$$\mathcal{L}_{\text{rms}}(\eta) = \int_0^\infty \mathcal{L}(\eta, \theta) d\theta. \quad (11)$$

In Figure 2, we show the likelihood  $\mathcal{L}_{\text{rms}}(\eta)$ , each normalized to  $\int d\eta \mathcal{L} = 1$ , in the case of  $\alpha = 0.60, 0.83$ , and  $1.10$ , along with a Gaussian distribution of the same 68% confidence level in the case of  $\alpha = 0.83$ . Note the significant tail to the distribution relative to a Gaussian, which reflects the fact that the measurements are not consistent with their individual normal errors. The values of  $\eta_{10}$  corresponding to the maximum of  $\mathcal{L}_{\text{rms}}(\eta)$ , equation (11), along with  $1\sigma$  (68% confidence) errors, are given in the last row of Table 1, for various  $\alpha$ . For  $\alpha = 0.83$ , the rms measurements of the velocity field give  $\eta_{10} = 0.44 \pm 0.08$  ( $1\sigma$ ).

This model is only a rough attempt at quantifying the degree of systematic errors present in these measurement (see Press 1996 for an alternate method). The assumptions that each systematic error  $\mu_i$  is independent and normally distributed are certainly not completely correct. All of the analyses use the *IRAS* galaxy redshift catalog, and many of them use the same Mk III catalog; therefore, their results must be somewhat correlated. Furthermore, systematic

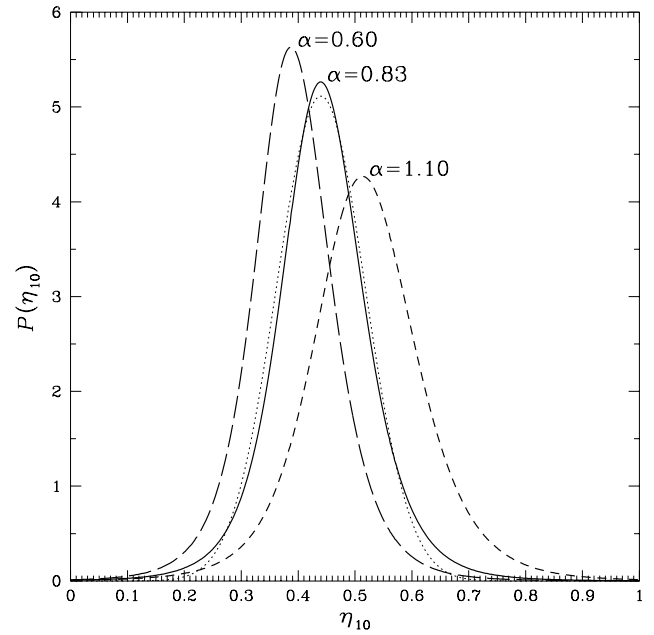


FIG. 2.—Probability density for the error analysis in § 2 of the rms measurements of  $\eta_{10}$ , for power-law indices  $\alpha = 0.60, 0.83$ , and  $1.10$ . The thin dotted line is a Gaussian distribution with the same mean and 68% confidence interval ( $1\sigma$  error) as the probability density for  $\alpha = 0.83$ .

errors are rarely normally distributed; however, the sensitivity to this assumption is somewhat alleviated by integrating over all possible standard deviations  $\theta$ . Perhaps one of the measurements is correct and all the others are erroneous. In comparing rms and tail-inferred values of  $\eta$ , we therefore also use the Willick et al. (1997b) result alone, which is the most recent result that explicitly filters at a scale close to the  $\sim 10 h^{-1}$  Mpc scale from which clusters form.

## 3. TAIL FLUCTUATIONS: ABUNDANCES OF X-RAY CLUSTERS

Galaxy statistics and peculiar velocities measure the rms fluctuations of the density field, i.e., fluctuations within  $\sim 1\sigma$  from the mean. Most probability distributions look similar to Gaussians at this fluctuation level; they are peaked and convex near the mean. The more dramatic differences between Gaussian and non-Gaussian distributions come in their tails. Rich clusters of galaxies, the most massive virialized objects in the universe, offer the best measure of the tail of the probability distribution of the density field on scales of  $\sim 10 h^{-1}$  Mpc. Approximately 5% of the  $L_*$  galaxies are within one Abell radius of a rich cluster (Bahcall 1996), and about 10% of the baryons reside in clusters (taking a big bang nucleosynthesis (BBN) value of  $\Omega_b = 0.0125 h^{-2}$  from Walker et al. 1991). Thus in the Gaussian scenario, they constitute  $\gtrsim 1.5\sigma$  perturbations. They are well observed by a number of techniques; in particular, their abundance as a function of X-ray temperature is well measured (Henry & Arnaud 1991). Thus they provide an accurate estimate of the integral over the tail of the distribution function at the  $\gtrsim 1.5\sigma$  level.

The temperature function of rich clusters can be predicted from the statistics of peaks in a density field; it depends on the rms level of fluctuations  $\sigma_R$  on the scale on which clusters form, as well as  $\Omega_0$ , through the relation between

comoving radius and mass. Comparing these predictions with observations constrains the quantity  $\sigma_R \Omega_0^v$ ,  $v = 0.4-0.6$  (see White, Efstathiou, & Frenk 1993; Eke, Cole, & Frenk 1996; Pen 1997, and references therein). The mass of a rich cluster within the  $1.5 h^{-1}$  Mpc Abell radius is  $10^{14}-2 \times 10^{15} h^{-1} M_\odot$ , which corresponds to a region with initial comoving radius  $R = (3M/4\pi\bar{\rho}_0)^{1/3} = (4.4-12) \Omega_0^{-1/3} h^{-1}$  Mpc. Numerical simulations have confirmed that the cosmic abundance of these objects is determined by the power spectrum at  $k^{-1} = 5-20 h^{-1}$  Mpc, so that their frequency provides a direct measure of  $\sigma_R$  in this range (e.g., Bahcall & Cen 1992; Evrard, Metzler, & Navarro 1996; Cen et al. 1998). Hence the cluster abundances constrain  $\eta_{10}$  with only a weak dependence on  $\Omega_0$ . In § 3.1 we derive the relation between the distribution of clusters as a function of temperature and the assumed density PDF. This requires a relation between the observed temperature of a cluster and the initial comoving radius from which it formed; this is given in § 3.2. This relation depends on the formation epoch of the cluster, as derived in § 3.3. We compare these results to the observed temperature distribution of clusters in § 3.4 to derive the quantity  $\eta_{10}$ .

### 3.1. Press-Schechter and the Temperature Function

We use the Press-Schechter (1974, hereafter PS) *Ansatz* to derive the number density of clusters. Many have done this in the Gaussian case (e.g., White et al. 1993), and here we make a straightforward generalization to non-Gaussian distributions. In PS, an object of mass  $\geq M$  forms when the linearly extrapolated density contrast filtered on a mass-scale  $M$ , corresponding to a top-hat comoving length-scale  $R$ , exceeds the threshold  $\delta_c \approx 1.69$ . The exact value of  $\delta_c$  depends on  $\Omega_0$  and  $\Lambda$ , and can be derived assuming uniform spherical collapse (see Eke et al. 1996), but varies by only a few percent for  $\Omega_0 > 0.2$ . The differential abundance from PS, given here as the comoving number density at the present of objects with initial comoving radius in the interval  $(R, R + dR)$ , can be written as

$$n_R dR = \frac{3}{4\pi R^3} \left| \frac{d}{dR} \int_{\delta_c/\sigma_R}^{\infty} fP(y) dy \right| dR, \quad (12)$$

where we assume that the mass distribution function  $P(y)$  only depends on  $y \equiv \delta/\sigma_R$ . This is true in Gaussian theories; in the texture model, the seeding of density perturbations is scale invariant, and thus the distribution function must also be scale invariant. The prefactor  $f \equiv [\int_0^\infty P(y) dy]^{-1}$  normalizes  $P$  so that all the mass in the universe is accounted for. Assuming, as before, a power-law dependence  $\sigma_R \propto R^{-\alpha}$  gives

$$n_R = \frac{3}{4\pi R^3} \frac{\alpha y}{R} fP(y), \quad y \equiv \frac{\delta_c}{\sigma_R}. \quad (13)$$

The initial comoving radius is of course not an observable, so we will convert to the abundance as a function of present-day temperature:

$$n_T \frac{dT}{dR} = \frac{3}{4\pi R^3} \frac{\alpha y}{R} fP(y). \quad (14)$$

The observed X-ray temperature function gives the  $n_T$  of equation (14) as a function of temperature. In order to find  $y$  (and hence  $\sigma$  and  $\eta$ ), we need the appropriate PDF for each model, and the relationship between the initial comoving radius and the temperature of the clusters (next section).

In Gaussian theories, the distribution function is

$$fP(y) = \frac{2}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right), \quad (15)$$

recalling that the factor  $f \equiv [\int_0^\infty P(y) dy]^{-1} = 2$ . For textures, we take the distribution function from the numerical simulations of Park et al. (1991), which are the only published simulations that are at scales appropriate for clusters of galaxies. We find that the high- $y$  tail of their distribution gives

$$fP(y) = 2.2 \exp(-1.45y), \quad y \geq 1.5. \quad (16)$$

Equation (16) is not valid for  $y < 1.5$ , but as we are only considering the tail of the distribution, it suffices for our purposes. We show the tails of the functions  $fP(y)$  for a Gaussian and texture PDF in Figure 3. Note that the texture PDF has a significantly wider tail for  $y \gtrsim 2$ .

### 3.2. The Radius-Temperature Relation

We next consider the relationship between the initial comoving radius of the perturbation and the present temperature of a cluster. We make the standard assumption that a cluster forms by undergoing a spherical collapse to a singular isothermal sphere with a virial radius equal to half its maximum expansion radius (e.g., White et al. 1993; Eke et al. 1996; Pen 1997). The X-ray temperature simply reflects the depth of the gravitational potential well of the cluster. We assume a hydrogen mass fraction of 0.76. For clusters collapsing at redshift  $z_f$ , we obtain

$$T = \frac{\mu m_p GM}{2r_{\text{vir}}} \quad (17)$$

$$= 7.8 \text{ keV} \left( \frac{M}{10^{15} h^{-1} M_\odot} \right)^{2/3} \left[ \frac{\Delta_c(z_f)}{18\pi^2} \right]^{1/3} \times \left[ \frac{\Omega_0}{\Omega(z_f)} \right]^{1/3} (1 + z_f) \quad (18)$$

$$= 8.6 \text{ keV } \Omega_0^{2/3} \left( \frac{R}{10 h^{-1} \text{ Mpc}} \right)^2 \left[ \frac{\Delta_c(0)}{18\pi^2} \right]^{1/3} \times \left\{ \left[ \frac{\Delta_c(z_f)}{\Delta_c(0)} \cdot \frac{\Omega_0}{\Omega(z_f)} \right]^{1/3} (1 + z_f) \right\} \quad (19)$$

$$\equiv T_0(R) \times \phi(z_f), \quad (20)$$

where the quantity in the braces is designated  $\phi$ , and only depends on the background cosmology and the formation redshift  $z_f$ . The function  $T_0(R)$  is the radius-temperature relation for collapse at  $z_f = 0$  for a given background cosmology; thus  $\phi(z_f = 0) = 1$ . Equation (18) was derived previously by Eke et al. (1996). The factor  $\Delta_c \equiv 3M/(4\pi r_{\text{vir}}^3 \rho_c)$  is the ratio of the virialized physical mass density to the *critical* (not the mean) cosmological density at the time of collapse, and  $r_{\text{vir}}$  is taken to be  $r_{\text{max}}/2$ . In an  $\Omega = 1$ , matter-dominated universe,  $\Delta_c = 18\pi^2 \approx 178$ ; for  $\Omega < 1$ ,  $\Delta_c$  can be derived from uniform spherical collapse (see Eke et al. 1996). The quantity  $\Delta_c$  depends only on the background cosmology. Because  $\Delta_c$  and  $\Omega$  asymptote to  $18\pi^2$  and 1, respectively, as  $z_f \rightarrow \infty$ , and hence change by a factor less than  $\Omega_0^{-1}$ , the  $z_f$  dependence of  $\phi$  is dominated by the factor  $1 + z_f$ .

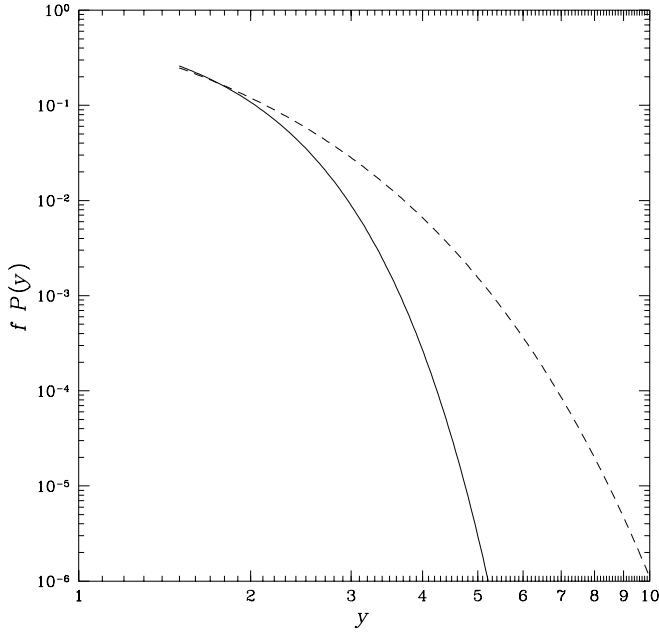


FIG. 3.—Tails of the effective distribution functions  $fP(y)$ , for a Gaussian (solid line) and texture (dashed line) PDF.

Evrard et al. (1996) have compared numerical simulations for Gaussian  $\Omega_0 = 1.0$  and  $\Omega_0 = 0.2$  models, and found that  $T$  scales as  $M^{2/3}$ , as equation (18) implies. Fitting a power law to their results, and converting mass to comoving scale implies  $kT = (8.8 \pm 0.8 \text{ keV})(\Delta_c/18\pi^2)^{0.164}\Omega_0^{2/3}(R/8 \text{ h}^{-1} \text{ Mpc})^2$ . This equation is consistent with equation (19) for all three background cosmologies we consider, using the results of the next section to determine  $z_f$  and  $\phi(z_f)$ . The detailed simulations necessary to check the cluster temperatures for the texture scenario have not been done. Bartlett (1997) notes, however, that for any given density profile of a spherical collapse, the power law  $T \propto M^{2/3} \propto R^2$  should be the same; the only difference might be in the constant of proportionality (e.g., through a different  $\phi$  dependence).

Taking the derivative of equation (20) with respect to  $R$  gives

$$\frac{dT}{dR} = T_0(R) \frac{d\phi}{dR} + \phi[z_f] \frac{dT_0}{dR} = \frac{T}{R} \left( 2 + \frac{d \ln \phi}{d \ln R} \right), \quad (21)$$

where the redshift of formation  $z_f$  must be calculated as a function of  $R$ . Finally, inverting equation (20) gives  $R$  as a function of the temperature:

$$R = 10 \text{ h}^{-1} \text{ Mpc } \Omega_0^{-1/3} \left( \frac{T}{8.6 \text{ keV}} \right)^{1/2} \left[ \frac{18\pi^2}{\Delta_c(0)} \right]^{1/6} \times \left\{ \left[ \frac{\Delta_c(0)}{\Delta_c(z_f)} \cdot \frac{\Omega(z_f)}{\Omega_0} \right]^{1/6} \frac{1}{(1+z_f)^{1/2}} \right\} \quad (22)$$

$$\equiv R_0(T) \times \phi[z_f(y)]^{-1/2}, \quad (23)$$

where  $\phi$  is defined as above, the factor in front of the braces has been replaced by  $R_0(T)$ , the radius-temperature relation at  $z_f = 0$ . Note that equation (23) so defined is only valid when  $z_f$  is known.

### 3.3. Formation Epoch of Clusters in PS

We now have another parameter to consider, the redshift of formation  $z_f$ . The standard assumption is to assume that

the cluster has “just formed” at the redshift  $z_0$  at which we observe it, i.e., that  $z_f = z_0$ . In the  $\Omega_0 = 1$  Gaussian CDM scenario, this assumption is valid, since clusters form late and accretion continues to the present (e.g., Lacey & Cole 1993). If  $\Omega_0 < 1$ , however, linear growth of perturbations freezes out when the universe enters its phase of free expansion at  $1+z \sim 1/\Omega_0$  (see Peebles 1993). Therefore, a given structure must collapse at a higher redshift than in the  $\Omega_0 = 1$  case in order to form. If there is a sufficiently positively skewed non-Gaussian tail to the distribution function, as there is for textures, a larger fraction of the universe collapses nonlinearly early on (see Bartlett, Gooding, & Spergel 1993 and Pen et al. 1994). Furthermore, for textures and other “causal-seed” models, whose density perturbations are compensated, there is a limit to the final mass of a collapsing structure, defined by the size of the causal horizon at the time the perturbation is seeded. Because the integrated overdensity is zero on larger scales, no further accretion can occur (Bartlett et al. 1993).

We now derive the mean redshift of formation  $\langle z_f \rangle$ , and use that redshift in equation (20) to determine the typical temperature of clusters as a function of  $R$ . The net differential number density of clusters that form in redshift interval  $(z, z+dz)$  follows from equation (13)

$$\frac{dn_R}{dz} dz = \frac{3}{4\pi R^3} \frac{\alpha}{R} f \frac{d}{dz} \{y(z)P[y(z)]\} dz, \quad (24)$$

where

$$y(z) \equiv \frac{\delta_c(z)}{\sigma_R(z)} = \frac{\delta_c(0)}{\sigma_R(0)} \cdot \frac{\delta_c(z)}{\delta_c(0)D(z)} \quad (25)$$

$$\equiv y_0 \frac{\delta_c(z)}{\delta_c(0)D(z)}. \quad (26)$$

Here  $y_0$  refers to the value of  $y(z)$  at  $z = 0$ ,  $D(z)$  is the linear growth factor normalized to unity at redshift zero, and  $\delta_c(z)$  is the critical linear density of collapse at redshift  $z$ . For clusters observed at redshift  $z_0$ , the mean redshift of formation  $\langle z_f \rangle_{z_0}$  is obtained by averaging the redshift  $z$  weighted by equation (24):

$$\langle z_f \rangle_{z_0} = \frac{\int_{z_0}^{\infty} z (dn_R/dz) dz}{\int_{z_0}^{\infty} (dn_R/dz) dz} \quad (27)$$

$$= z_0 + \frac{1}{y(z_0)P[y(z_0)]} \int_{z_0}^{\infty} y(z)P[y(z)] dz, \quad (28)$$

where the second line comes from integrating the numerator by parts, integrating the denominator, and canceling terms. The mean redshift of formation of clusters at  $z_0$  is a function only of the background cosmology through  $\delta_c(z)$  and  $D(z)$ , and the current value of  $y_0$  through equation (26). We take the formation redshift required in equation (20) to be  $\langle z_f \rangle$ . When using equation (28) with the present epoch abundances below, we set  $z_0 = 0$ .

### 3.4. Determining $\eta$ from X-Ray Cluster Surveys

Using equations (20) and (21) in equation (14) gives

$$2n_T T \left( 1 + \frac{\alpha}{2} \frac{d \ln \phi}{d \ln y} \right) = \frac{3}{4\pi R^3} \alpha y f P(y), \quad (29)$$

where  $y$  is now understood to be its present-day value, and we have used  $(d \ln y / d \ln R) = -(d \ln n / d \ln R) = \alpha$ , because of our power-law assumption. Replacing  $R$  with equation (23) and separating those quantities depending on  $T$  from

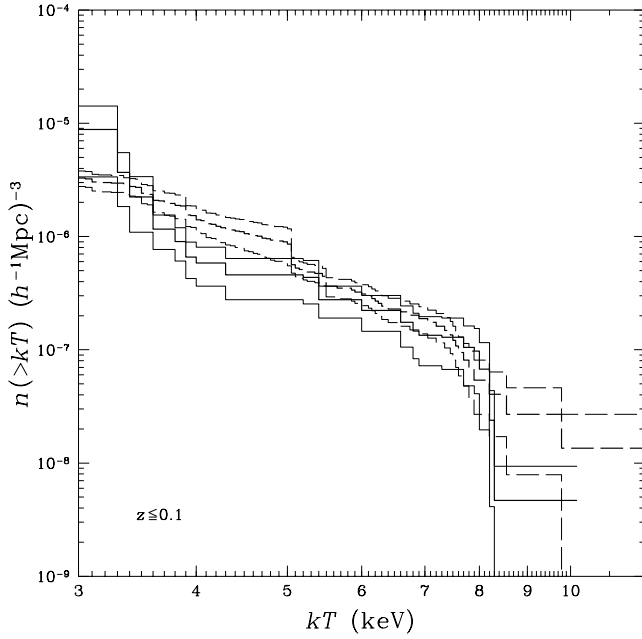


FIG. 4.—Cumulative temperature function from Henry & Arnaud (1991) (thick solid line), and from the XBACS survey (Ebeling et al. 1996) (thick dashed line), both with  $1\sigma$  Poisson errors from eq. (32) (thin histograms).

those depending on  $y$  gives

$$n_T T \frac{8\pi R_0^3(T)}{3} = \frac{\phi^{3/2}}{1 + (\alpha/2)(d \ln \phi / d \ln y)} \alpha y f P(y). \quad (30)$$

We need now only to relate observations of clusters to the left-hand side of equation (30).

There are two virtually complete X-ray flux-limited surveys of rich galaxy clusters at low redshift. The first was compiled by Henry & Arnaud (1991), and the second is the X-Ray Brightest Abell Cluster Survey (XBACS) from the *ROSAT* all-sky survey (Ebeling et al. 1996). The Henry & Arnaud (1991) survey is essentially contained within the XBACS survey. An unbiased estimator of the number density of clusters in the temperature interval  $(T, T + \Delta T)$  for a flux-limited survey is given by

$$n_T \Delta T = \sum_{T < T_i < T + \Delta T} V_{\max,i}^{-1}, \quad (31)$$

where  $V_{\max,i}$  is the maximum volume in which the cluster of temperature  $T_i$  (in keV) could be detected given the flux limit and geometric boundaries of the survey. In this case, we require a geometrical boundary at galactic latitude  $|b| > 20^\circ$ , and redshift  $z \leq 0.1$ . The full XBACS sample takes the maximum redshift to be  $z_{\max} = 0.2$ . The Poisson

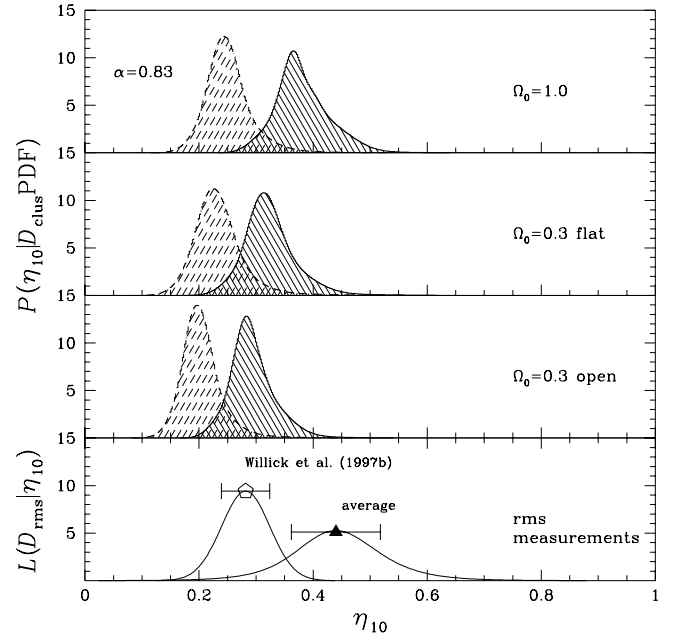


FIG. 5.—Cluster-inferred  $\eta_{10}$  compared to peculiar velocity/redshift distortion-inferred  $\eta_{10}$  at  $10 h^{-1} \text{ Mpc}$  for  $\alpha = 0.83$ , for various assumed background cosmologies. The solid-filled distribution functions are for a Gaussian PDF, and the dashed-filled distribution functions are for a texture PDF. The filled triangle is the maximum likelihood average of the measurements in Table 1. The open pentagon is the Willick et al. (1997b) result.

variance, following Pen (1997), is given by

$$\sigma^2(n_T \Delta T) = \sum_{T < T_i < T + \Delta T} V_{\max,i}^{-2}. \quad (32)$$

The cumulative temperature functions  $n(>T)$  for the two samples are shown in Figure 4. The two temperature functions are quite consistent in the overlapping temperature range greater than 3 keV, so we use this part of the (10 times) larger XBACS survey for our analysis.

Substituting equation (31) gives, finally,

$$\left( \frac{T}{\Delta T} \sum_{T < T_i < T + \Delta T} V_{\max,i}^{-1} \right) \frac{8\pi R_0^3(T)}{3} = \frac{\phi^{3/2} \alpha y f P(y)}{1 + (\alpha/2)(d \ln \phi / d \ln y)}. \quad (33)$$

We define the temperature bins  $(T, T + \Delta T)$  so that there are at least two clusters in each. For a fixed background cosmology and power-law index  $\alpha$ , the left-hand side of equation (33) contains only observed quantities, and the right-hand side is a function of  $y$  and the PDF. For each PDF, we therefore can solve for  $y$  implicitly as a function of  $T$ . We then convert  $T$  to  $R$  using equation (23) and our

TABLE 2  
XBACS CLUSTER MEASUREMENTS OF  $\eta_{10}$

CLUSTER SCENARIO	$\eta_{10}$		
	$\alpha = 0.60$	$\alpha = 0.83$	$\alpha = 1.10$
Gaussian $\Omega_0 = 1$ .....	$0.423 \pm 0.036$	<b><math>0.381 \pm 0.041</math></b>	$0.342 \pm 0.049$
Texture $\Omega_0 = 1$ .....	$0.287 \pm 0.033$	<b><math>0.252 \pm 0.031</math></b>	$0.220 \pm 0.032$
Gaussian $\Omega_0 = 0.3$ flat .....	$0.326 \pm 0.035$	<b><math>0.325 \pm 0.038</math></b>	$0.329 \pm 0.047$
Texture $\Omega_0 = 0.3$ flat .....	$0.245 \pm 0.037$	<b><math>0.235 \pm 0.034</math></b>	$0.231 \pm 0.036$
Gaussian $\Omega_0 = 0.3$ open .....	$0.298 \pm 0.029$	<b><math>0.293 \pm 0.033</math></b>	$0.293 \pm 0.042$
Texture $\Omega_0 = 0.3$ open .....	$0.215 \pm 0.028$	<b><math>0.204 \pm 0.027</math></b>	$0.198 \pm 0.029$

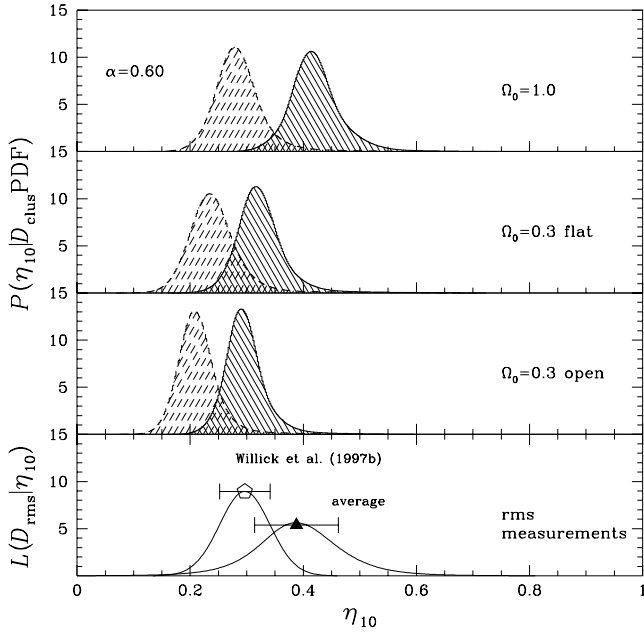


FIG. 6.—Cluster-inferred  $\eta_{10}$  compared to peculiar velocity/redshift distortion-inferred  $\eta_{10}$  at  $10 h^{-1}$  Mpc for  $\alpha = 0.60$ . The solid-filled distribution functions are for a Gaussian PDF, and the dashed-filled distribution functions are for a texture PDF.

derived value of  $y$ , and obtain a value of  $\sigma_{R, \text{clus}} \equiv \delta_c / y$  as a function of the value of  $R$  inferred from each temperature bin. Multiplying by  $\Omega_0^{0.6}$  gives the cluster-inferred  $\eta_{R, \text{clus}}$ , which can be compared directly to the values from the rms measurements.

A clear comparison with rms-inferred  $\eta$  can be made if we extrapolate all values of  $\eta$  to a common scale, using the power-law model we have assumed throughout. We choose  $10 h^{-1}$  Mpc because this scale corresponds to a tem-

perature within the range of the observations for  $\Omega_0 = 0.3$ – $1.0$ : 8.6 keV for  $\Omega_0 = 1$  and 3.4–4.0 keV (depending on  $\phi$ ) for  $\Omega_0 = 0.3$ . For  $\Omega_0 = 0.3$ , the scale  $R = 8 h^{-1}$  Mpc corresponds to a temperature of 2.2–2.6 keV, outside the range of observations we consider.

For each temperature bin, we therefore extrapolate the derived value of  $\eta_R$  to  $10 h^{-1}$  Mpc using  $\eta_R \propto \sigma_R \propto R^{-\alpha}$ , and obtain a measurement of  $\eta_{10}$  for each bin. We wish to find the resulting probability distribution in  $\eta_{10}$ . We find that the statistical errors in  $\eta_{10}$  from equations (32) and (33) are smaller than the scatter of the values of  $\eta_{10}$ , although the discrepancies are not as drastic as in the case of the rms measurements. Certainly, then, there remains some type of systematic error. We sum the distribution functions of each extrapolated  $\eta_{10}$  to obtain the total distribution function, assuming each error is Gaussian:

$$\mathcal{P}(\eta_{10} | D_{\text{clus}} \text{ PDF}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\sqrt{2\pi\sigma_i^2}} \times \exp \left[ -\frac{1}{2} \left( \frac{\eta_{10,i} - \eta_{10}}{\sigma_i} \right)^2 \right], \quad (34)$$

where the sum is over the temperature bins labeled by  $i$ . Note that we are summing over the individual distributions because the individual temperature bins are not independent measures of  $\eta_{10}$ ; i.e., the data set as a whole is taken as a single measurement, with the error given by the total distribution  $\mathcal{P}(\eta_{10} | D_{\text{clus}} \text{ PDF})$ .

Figures 5, 6, and 7 show the cluster-inferred  $\eta_{10}$  distributions for power laws  $\alpha = 0.83, 0.60$ , and  $1.10$ , and for different values of  $\Omega_0$  and the PDF. Also shown are the likelihoods for mean peculiar velocity/redshift distortion-derived rms values discussed in § 2, and the Willick et al. (1997b) result alone. The means and standard deviations of the resulting distribution of  $\eta_{10}$  measurements are given in Table 2, for various values of  $\Omega_0$  and  $\alpha$ . Note that lower values of  $\alpha$  seem to give a “better fit,” with  $\alpha = 0.60$  or  $0.83$  having the smallest fractional errors in  $\eta$  in every case.

The most prominent feature of the inferred rms from clusters is its strong dependence on the PDF: Gaussian and texture models give predictions of  $\eta$  that differ at the  $\sim 3\sigma$  level. The degree to which the cluster-inferred value of  $\eta_{10}$  differs between Gaussian and texture PDFs is directly related to the degree to which the tail of the texture PDF is distinguishable from a Gaussian.

#### 4. DISCUSSION

##### 4.1. Which PDF is More Likely?

We quantify the comparison between the rms and cluster-inferred determinations of  $\eta_{10}$  with a Bayesian model comparison analysis (Loredo 1990). For now, we fix the background cosmology and the power-law index  $\alpha$ , and suppress the dependence on these assumptions. We ultimately seek the probability that a certain PDF is correct, given the cluster data  $D_{\text{clus}}$  and the rms data  $D_{\text{rms}}$ :

$$\mathcal{P}(\text{PDF} | D_{\text{clus}} D_{\text{rms}}) = \frac{\mathcal{P}(\text{PDF})}{\mathcal{P}(D_{\text{clus}} D_{\text{rms}})} \mathcal{P}(D_{\text{clus}} D_{\text{rms}} | \text{PDF}), \quad (35)$$

where we have used Bayes Theorem. The fraction in front consists of a prior (which we take as uniform) and a normalization factor. Using the product rule, we separate the last

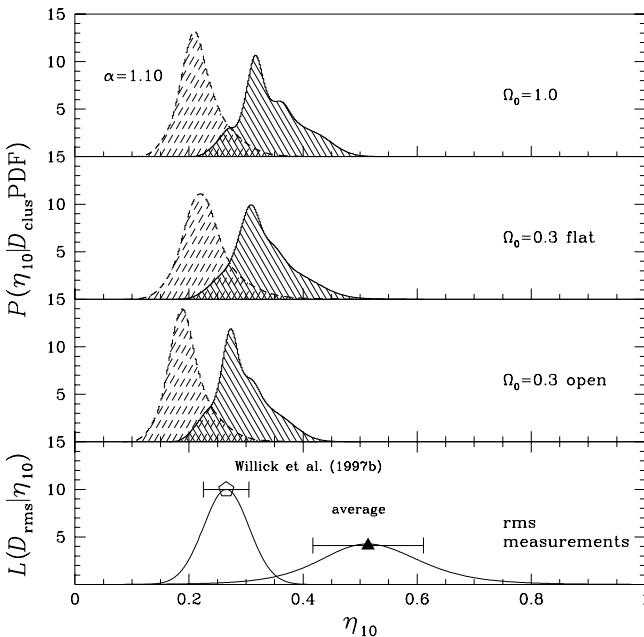


FIG. 7.—Cluster-inferred  $\eta_{10}$  compared to peculiar velocity/redshift distortion-inferred  $\eta_{10}$  at  $10 h^{-1}$  Mpc for  $\alpha = 1.10$ . The solid-filled distribution functions are for a Gaussian PDF, and the dashed-filled distribution functions are for a texture PDF.



probability

$$\mathcal{P}(D_{\text{clus}} D_{\text{rms}} | \text{PDF}) = \mathcal{P}(D_{\text{clus}} | \text{PDF}) \mathcal{P}(D_{\text{rms}} | D_{\text{clus}} \text{PDF}) . \quad (36)$$

Since we are primarily interested in the rms-versus-tail comparison, we take the first probability in equation (36) to be independent of the PDF. A more sophisticated analysis would evaluate the “goodness of fit” of the cluster data for each PDF, and thereby estimate the absolute probability  $\mathcal{P}(D_{\text{clus}} | \text{PDF})$  (e.g., using a Kolmogorov-Smirnov-type test). However, as mentioned in § 3, the scatter in the inferred values of  $\eta_{10, \text{clus}}$  is greater than the calculated Poisson errors, so any such analysis will have to take into account unknown systematic errors (e.g., using the model described in § 2.2).

The dependence of  $D_{\text{rms}}$  on  $D_{\text{clus}}$  in the second probability in equation (36) is through the value of  $\eta_{10}$ ; using the summation rule gives

$$\mathcal{P}(D_{\text{rms}} | D_{\text{clus}} \text{PDF}) = \int d\eta_{10} \mathcal{P}(D_{\text{rms}} | \eta_{10} \text{PDF}) \times \mathcal{P}(\eta_{10} | D_{\text{clus}} \text{PDF}) . \quad (37)$$

The resulting expression for equation (35), suppressing all factors assumed to be independent of the PDF, is

$$\mathcal{P}(\text{PDF} | D_{\text{clus}} D_{\text{rms}}) \propto \int d\eta_{10} \mathcal{P}(D_{\text{rms}} | \eta_{10} \text{PDF}) \times \mathcal{P}(\eta_{10} | D_{\text{clus}} \text{PDF}) . \quad (38)$$

The cluster distribution function in  $\eta_{10}$ ,  $\mathcal{P}(\eta_{10} | D_{\text{clus}} \text{PDF})$ , is given by equation (34). The rms probability depends only on the power-law slope  $\alpha$  and the weighting of the different analyses, and is independent of the PDF. If we use the error model in § 2.2, the probability density  $\mathcal{P}(D_{\text{rms}} | \eta_{10} \text{PDF})$  is given by the likelihood function from equation (11) and Figure 2,

$$\mathcal{P}(D_{\text{rms}} | \eta_{10} \text{PDF}) = \mathcal{L}_{\text{rms}}(\eta_{10}) . \quad (39)$$

If we are considering only the Willick et al. (1997b) data,  $D_{\text{w}}$ , and we assume normal errors, we use

$$\mathcal{P}(D_{\text{w}} | \eta_{10} \text{PDF}) = \frac{1}{\sqrt{2\pi\epsilon_{\text{w}}^2}} \exp \left[ -\frac{1}{2} \left( \frac{\eta_{10} - \eta_{\text{w}}}{\epsilon_{\text{w}}} \right)^2 \right] , \quad (40)$$

where  $\eta_{\text{w}}$  and  $\epsilon_{\text{w}}$  are given by the values inferred from their analysis, also listed in Table 1.

Since we do not know the full range of possible PDFs, we take the Bayesian approach of considering the likelihood ratio (the “odds,” or “Bayes factor” for uniform priors,

Loredo 1990) of a Gaussian versus a texture PDF. The values of the likelihood ratio  $\mathcal{P}(\text{Gaussian})/\mathcal{P}(\text{texture})$  are listed in Table 3 for various background cosmologies and values of  $\alpha$ . If we take the average of the rms measurements, a Gaussian PDF is favored for any background cosmology. The probability of a Gaussian PDF is always more than 2 times that of a texture PDF. The probability that the Gaussian PDF is correct is 70%–88%, roughly 1–1.5  $\sigma$ . If we only use the Willick et al. (1997b) measurement, however, then the PDF that is favored depends strongly on the background cosmology, and somewhat on the power-law index  $\alpha$ . In the case of  $\alpha = 0.83$ , if we live in an open  $\Omega_0 = 0.3$  universe, then the Gaussian PDF is 3 times more likely than the texture PDF; if we have a flat  $\Omega_0 = 0.3$  universe ( $\Lambda = 0.7$ ), then a Gaussian and texture PDF are roughly equally likely; and if  $\Omega_0 = 1$ , a texture PDF is 3 times more likely than a Gaussian.

#### 4.2. Constraints from Cluster Evolution?

We have only considered the present-day abundances of clusters, and we have a degeneracy between the assumed PDF and the background cosmology. The evolution of cluster abundances, however, offers an opportunity to break this degeneracy. In particular, high- $\Omega_0$  universes evolve more rapidly than those with low  $\Omega_0$ . Gaussian models with  $\Omega_0 = 1$  are already heavily disfavored because they predict too rapid an evolution in cluster abundances out to  $z \geq 0.3$  (Henry 1997; Donahue et al. 1997). The fact that the Willick et al. (1997b) measurement leads to the  $\Omega_0 = 1$  Gaussian scenario being disfavored relative to a texture scenario simply reinforces the case against a flat matter-dominated universe for Gaussian fluctuations.

The formalism of § 3 relating cluster abundances to the PDF and the value of  $\eta$  is fully generalizable to  $z > 0$ . As an example, we show in Figure 8 the cumulative number density of clusters  $\geq 10$  keV as a function of redshift for various scenarios. These curves are normalized to a present-day number density of  $1 \times 10^{-8} h^3 \text{ Mpc}^{-3}$  (see Fig. 4). For fixed background cosmology, the texture scenarios evolve more slowly than the Gaussian ones because of the shallower texture PDF. For fixed PDF, the evolution of the number density evolves most quickly with  $\Omega_0 = 1$  and most slowly with open  $\Omega_0 = 0.3$ . A Gaussian PDF with a flat  $\Omega_0 = 0.3$  background cosmology has roughly the same evolutionary history as a texture PDF with  $\Omega_0 = 1$ . If evolution is observed to be significantly slower or more rapid than this, then an  $\Omega_0 = 1$  texture scenario can be ruled out.

Henry (1997) reports that the number density of  $z = 0.3$  clusters at the hot end of the temperature function ( $\gtrsim 6.5$  keV) is a factor of  $\sim 2.5$  lower than today (his Fig. 2), although this is still within the calculated scatter due to

TABLE 3  
PROBABILITY OF GAUSSIAN (G) RELATIVE TO TEXTURE (T) PDF

ROOT MEAN SQUARE AVERAGING	$\Omega_0$ AND GEOMETRY	$\mathcal{P}(\text{G})/\mathcal{P}(\text{T})$ FOR $\alpha =$		
		0.60	0.83	1.10
Maximum likelihood .....	1.0	2.28	<b>6.52</b>	7.00
	0.3 Flat	3.28	<b>4.18</b>	4.41
	0.3 Open	3.98	<b>4.71</b>	4.62
Willick et al. 1997b only .....	1.0	0.124	<b>0.318</b>	0.704
	0.3 Flat	1.53	<b>1.25</b>	0.789
	0.3 Open	3.28	<b>3.01</b>	2.13

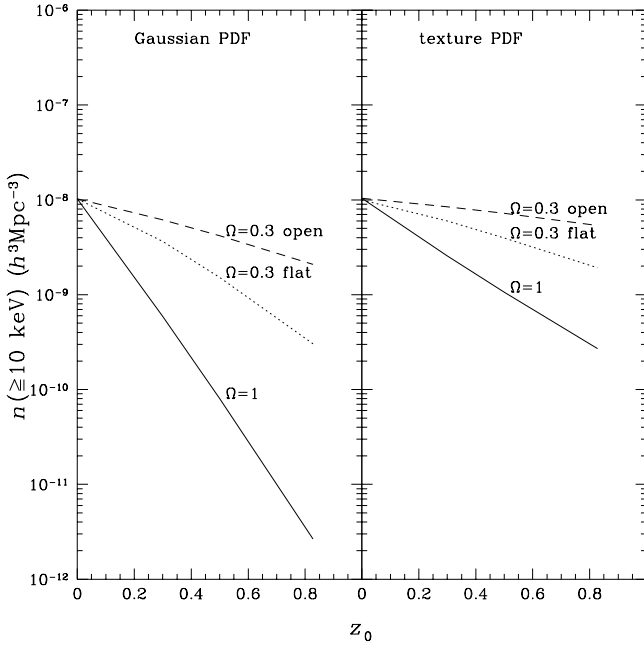


FIG. 8.—Evolution of the number density of clusters with temperature  $\geq 10$  keV, normalized to the present-day abundance, for Gaussian and texture PDFs, and various background cosmologies.

Poisson statistics. For  $\Omega_0 = 1$ , the Gaussian model predicts a factor  $\sim 8$  decrease in the density of such clusters at  $z = 0.3$  while the texture model predicts a factor  $\sim 3$  decrease. For the  $\Omega_0 = 0.3$  scenarios, the reduction factor is 1.1–2. Henry (1997) reports that the cluster evolutionary data rule out the Gaussian  $\Omega_0 = 1$  at 99% confidence.

Donahue et al. (1997) report an extremely hot X-ray temperature of  $14.7 \pm 4$  keV for cluster MS 1054–03, at redshift 0.828. They indicate that this result, along with data from other high-redshift clusters they have studied, implies that, within errors of  $\sim 50\%$ , the number density of clusters with  $T > 10$  keV is *essentially unchanged* from  $z = 0$  to  $z = 0.8$ . If this result is confirmed, even a texture PDF would be ruled out for  $\Omega_0 = 1$ , since it predicts a decrease in the number density by a factor of  $\sim 30$ . Further observations at redshift  $z = 0.5$ – $1.0$  will almost certainly resolve the ambiguity between models with different  $\Omega_0$ .

## 5. SUMMARY AND CONCLUSIONS

We show that comparing the amplitude of tail fluctuations, indicated by the abundances of clusters of galaxies, to the amplitude of the rms mass fluctuation, indicated by galaxy peculiar velocities and redshift-space distortions, can be a strong test of the Gaussianity of the PDF of the initial

density field with a weak dependence on the slope of the density fluctuation spectrum.

Observations of peculiar velocities and redshift distortions in the linear regime can measure the rms level of gravitational potential fluctuations present in the universe in the combination  $\eta_R \equiv \sigma_R \Omega_0^{0.6}$ , independent of galaxy bias. Various existing measurements of this quantity are not consistent with one another within their stated errors. We assume the existence of a systematic error in each measurement and use a maximum likelihood technique to combine these recent measurements of  $\eta_R$ . The systematic error leads to a  $\sim 20\%$  uncertainty in the value of  $\eta_R$ . This procedure gives  $\eta_{10, \text{rms}} = 0.39 \pm 0.07$  to  $0.51 \pm 0.10$ , depending on the assumed slope of the mass fluctuation spectrum. We use this value and one recent measurement by Willick et al. (1997b), which indicates  $\eta_{10} = 0.27$  to  $0.30 \pm 0.04$ , to compare to the cluster-inferred measures of  $\eta_{10}$ . We chose the Willick et al. (1997b) measurement because it is the most recent analysis that explicitly filters close to the  $10 h^{-1}$  Mpc scale on which clusters form.

The value of  $\eta_{10}$  inferred from cluster abundances is calculated using the Press-Schechter approach for both a Gaussian and a texture PDF. We make the standard assumptions: that there is a power-law dependence of  $\eta_R \propto R^{-\alpha}$ , and that clusters form by undergoing spherical collapse to a singular isothermal sphere. We find a simple relation for the typical redshift of formation of clusters in both models. This is particularly important in the texture scenario, where the shallower tail of the PDF leads to significant early formation of clusters. This typical redshift is important in the relation between the virial temperature of the clusters and the comoving radius of the initial perturbation. For a texture PDF, the observed abundances of X-ray clusters implies  $\eta_{10} = 0.20 \pm 0.03$  to  $0.25 \pm 0.04$  for  $\Omega_0 = 0.3$ , and  $\eta_{10} = 0.22$  to  $0.29 \pm 0.03$  for  $\Omega_0 = 1.0$ . Given a Gaussian PDF, on the other hand, the cluster abundances indicate  $\eta_{10} = 0.29 \pm 0.04$  to  $0.33 \pm 0.05$  for  $\Omega_0 = 0.3$  and  $\eta_{10} = 0.34 \pm 0.05$  to  $0.42 \pm 0.04$  for  $\Omega_0 = 1$ .

We calculate the relative likelihoods of each PDF for a range of cosmological parameters. Using the maximum likelihood average of the rms measurements with the cluster abundances implies the Gaussian model is favored for any  $\Omega_0$ . However, using the Willick et al. (1997b) rms measurement with the clusters implies that a Gaussian is favored for open  $\Omega_0 = 0.3$ , a Gaussian and a texture PDF of roughly equal likelihood for flat  $\Omega_0 = 0.3$ , and a texture PDF is favored for  $\Omega_0 = 1$ . In each case, the accepted hypothesis is accepted at a  $\sim 1.5 \sigma$  level, with likelihood ratio  $\gtrsim 3$ . Preliminary constraints from cluster evolution indicate that low- $\Omega_0$  models are probably more viable than  $\Omega_0 = 1$  models. Our conclusions are summarized in Table 4, where

TABLE 4  
SUMMARY OF TESTS OF GAUSSIAN (G) VERSUS TEXTURES (T)

TESTS	$\Omega_0 = 1$		$\Omega_0 = 0.3$ FLAT		$\Omega_0 = 0.3$ OPEN	
	G	T	G	T	G	T
Root mean square average and $z = 0$ clusters .....	F	D	F	D	F	D
Willick et al. 1997b and $z = 0$ clusters .....	D	F	? =	? =	F	D
Cluster evolution .....	D	?D	?F	F	F	F

NOTE.—An F (D) indicates the PDF is favored (disfavored) over the alternative by  $\gtrsim 3$  times or  $\gtrsim 1.5 \sigma$ ; a question mark and equal sign indicate roughly equal likelihood; a question mark next to the F or D indicates a  $\lesssim 3$  times or  $\lesssim 1.5 \sigma$  result.

an F (D) indicates the PDF is favored (disfavored) over the alternative by  $\gtrsim 3$  times. A question mark next to the F or D indicates the factor is  $\lesssim 3$ .

Our results imply that the textures + CDM scenario faces some challenges from observations on  $5\text{--}25 h^{-1}$  Mpc scales, *completely independent* of CMB measurements. Although the current state of the measurements of the rms prevents us from firmly distinguishing the Gaussian and texture models, the principle behind this method is clear. The observed cluster abundance at  $z = 0$  predicts significantly different values for  $\eta_{10}$  for the two models. A direct determination of the measured rms at the  $\lesssim 10\%$  level, either from future redshift surveys and peculiar velocity measurements, or from an improved understanding of the systematic effects

that are present in the current analyses, will surely enable us to definitively test any distribution function of initial density perturbations. Furthermore, upcoming high-redshift X-ray cluster data will likely break any remaining degeneracy due to the assumed background cosmology. Soon, perhaps, we will finally be able to distinguish different models for the initial PDF of the universe at  $10 h^{-1}$  Mpc.

We thank Avishai Dekel and Ed Turner for useful conversations, Patrick Henry for an advanced preprint of his paper, and Tsafir Kolatt and referee Adrian Melott for useful comments. M. A. S. acknowledges the support of the Alfred P. Sloan Foundation, Research Corporation, and NSF grant AST 96-16901.

## APPENDIX

### WHAT IS THE CORRECT CORRESPONDENCE BETWEEN TOP-HAT AND GAUSSIAN FILTERING?

One of the traditional ways of relating different filtering schemes is by matching the volume over which they filter. For top hat and Gaussian, then, this would correspond to  $4\pi R_T^3/3 = (2\pi R_G^2)^{3/2}$ , or  $R_T/R_G \approx 1.6$ . While this might be appropriate when relating the *mass* of objects picked out by these filter scales, it is not the best way of relating the wavelengths of fluctuations measured by each filter. Here we derive a method of relating top-hat and Gaussian filtered fields by their maximum correlation.

Consider a homogeneous, isotropic scalar field  $F(x)$  and its Fourier transform in the following convention

$$F(x) = \sum_{\mathbf{k}} \tilde{F}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} = \frac{V}{(2\pi)^3} \int d^3k \tilde{F}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}}, \quad (\text{A1})$$

$$\tilde{F}(\mathbf{k}) = \frac{1}{V} \int_V d^3x F(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}}. \quad (\text{A2})$$

Let the power spectrum  $P(k) \equiv \langle |\tilde{F}(\mathbf{k})|^2 \rangle$ . The (dimensionless) top-hat and Gaussian filters and their Fourier transforms are then

$$W_{R_T}(x) = V \frac{3}{4\pi R_T^3} \quad \text{if } |x| < R_T, \quad 0 \text{ otherwise}, \quad (\text{A3})$$

$$\tilde{W}_{R_T}(\mathbf{k}) = \frac{3j_1(kR_T)}{kR_T} \equiv \tilde{W}_T(kR_T), \quad (\text{A4})$$

$$W_{R_G}(x) = V \frac{1}{(2\pi R_G^2)^{3/2}} \exp\left[-\frac{x^2}{(2R_G^2)}\right], \quad (\text{A5})$$

$$\tilde{W}_{R_G}(\mathbf{k}) = \exp(-k^2 R_G^2/2) \equiv \tilde{W}_G(kR_G), \quad (\text{A6})$$

where  $j_1(x)$  is the spherical Bessel function of order 1. The filtered fields are  $F_{R_T,G}(x) = V^{-1} \int d^3y F(x-y) W_{R_T,G}(y)$ . Using the convolution theorem, the variance for top-hat and Gaussian filters is

$$\sigma_{T,G}^2(R_{T,G}) = \frac{V}{(2\pi)^3} \int dk 4\pi k^2 P(k) \tilde{W}_{T,G}^2(kR_{T,G}). \quad (\text{A7})$$

How do we relate results with a top-hat filter and those with a Gaussian filter? We maximize the correlation between the two filtered fields. The correlation  $\rho$  is defined as the covariance divided by the product of the standard deviations (Stuart & Ord 1994)

$$\rho(R_T, R_G) \equiv \frac{\langle F_{R_T}(x) F_{R_G}(x) \rangle}{\sigma_T(R_T) \sigma_G(R_G)} = \frac{1}{\sigma_T(R_T) \sigma_G(R_G)} V \int \frac{dk 4\pi k^2}{(2\pi)^3} P(k) \tilde{W}_T(kR_T) \tilde{W}_G(kR_G). \quad (\text{A8})$$

Note that this prescription does *not* assume that the field  $F$  is a Gaussian random field, only that it is homogeneous and isotropic. In Table 5, we list the peak correlations for power-law power spectra  $P(k) \propto k^n$  for  $n = -2, -1, 0$ , and  $1$ . For decreasing power spectra ( $n < 0$ ), the correlation between top-hat and Gaussian filtering is very good, greater than 87%, with the ratio  $R_T/R_G = 1.75\text{--}2$ .

TABLE 5  
CORRESPONDENCE BETWEEN TOP-HAT AND GAUSSIAN  
FILTERING

Index $n$ [ $P(k) \propto k^n$ ]	$R_T/R_G$ (Maximum Correlation)	$\rho(R_T, R_G)$
-2 .....	1.94	0.99
-1 .....	1.85	0.97
0 .....	1.75	0.87
1 .....	1.66	0.57

It should be noted that we are not comparing the *value* of  $\sigma$  between the two filtering schemes; we are seeking the scales at which the two filters are most correlated. For example, for a CDM,  $\Omega_0 = 1$ ,  $n = 1$ , power spectrum, when the maximum correlation is used,  $12 h^{-1}$  Mpc Gaussian smoothing corresponds to  $22.7 h^{-1}$  Mpc top-hat smoothing; but  $\sigma_G(12)/\sigma_T(22.7) = 0.86$ , while  $\sigma_G(12)/\sigma_T(25.4) = 1.0$ .

## REFERENCES

- Bahcall, N. 1996, 13th Jerusalem Winter School in Theoretical Physics: Formation of Structure in the Universe, in press  
Bahcall, N., & Cen, R. 1992, ApJ, 398, 81  
Bartlett, J. G. 1997, First Moroccan International Astronomy School: From Quantum Fluctuations to Cosmological Structures, in press  
Bartlett, J. G., Gooding, A. K., & Spergel, D. N. 1993, ApJ, 403, 1  
Bernardeau, F. 1992, ApJ, 392, 1  
———. 1994, A&A, 291, 697  
Bertschinger, E., Dekel, A., Faber, S. M., Dressler, A., & Burstein, D. 1990, ApJ, 364, 370  
Cen, R., et al. 1998, in preparation  
Cole, S., Fisher, K. B., & Weinberg, D. H. 1994, MNRAS, 267, 785  
———. 1995, MNRAS, 275, 515  
Davis, M., & Peebles, P. J. E. 1983, ApJ, 372, 394  
Dekel, A., Bertschinger, E., & Faber, S. M. 1990, ApJ, 364, 349  
Donahue, M., Gioia, I., Luppino, G., Hughes, J. P., & Stocke, J. T. 1997, preprint (astro-ph/9707010)  
Ebeling, H., Voges, W., Bohringer, H., Edge, A. C., Huchra, J. P., & Briel, U. G. 1996, MNRAS, 281, 799  
Eke, V. R., Cole, S., & Frenk, C. S. 1996, MNRAS, 282, 263  
Evrard, A. E., Metzler, C. A., & Navarro, J. F. 1996, ApJ, 469, 494  
Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. P. 1994a, MNRAS, 266, 50  
———. 1994b, MNRAS, 267, 927  
Fisher, K. B., Scharf, C. A., & Lahav, O. 1994, MNRAS, 266, 219  
Fry, J. N. 1984a, ApJ, 277, L5  
———. 1984b, ApJ, 279, 499  
Gooding, A. K., Park, C., Spergel, D. N., Turok, N., & Gott, J. R., III. 1992, ApJ, 393, 42  
Gooding, A. K., Spergel, D. N., & Turok, N. 1991, ApJ, 372, L5  
Heavens, A. F., & Taylor, A. N. 1995, MNRAS, 275, 483  
Henry, J. P. 1997, ApJ, 489, L1  
Henry, J. P., & Arnaud, K. A. 1991, ApJ, 372, 637  
Kaiser, N. 1987, MNRAS, 227, 1  
Kim, R., & Strauss, M. A. 1998, ApJ, 493, 39  
Kolatt, T., & Dekel, A. 1997, ApJ, 479, 592  
Lacey, C. G., & Cole, S. 1993, MNRAS, 262, 627  
Loredo, T. J. 1990, in Maximum Entropy and Bayesian Methods, ed. P. F. Fougère (Dordrecht: Kluwer), 81  
Park, C., Spergel, D. N., & Turok, N. 1991, ApJ, 372, L53  
Peebles, P. J. E. 1993, Principles of Physical Cosmology (Princeton: Princeton Univ. Press)  
Pen, U. 1997, ApJ, in press  
Pen, U., Seljak, U., & Turok, N. 1997, Phys. Rev. Lett., 79, 1611  
Pen, U., & Spergel, D. N. 1995, Phys. Rev. D, 51, 4099  
Pen, U., Spergel, D. N., & Turok, N. 1994, Phys. Rev. D, 49, 692  
Press, W. H. 1996, in Unsolved Problems in Astrophysics, ed. J. P. Ostriker (Princeton: Princeton Univ. Press), 49  
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425 (PS)  
Seljak, U., & Bertschinger, E. 1994, ApJ, 427, 523  
Sigad, Y., Eldar, A., Dekel, A., Strauss, M. A., & Yahil, A. 1997, ApJ, in press  
Strauss, M. A., & Willick, J. A. 1995, Phys. Rep., 261, 271  
Stuart, A., & Ord, J. K. 1994, Kendall's Advanced Theory of Statistics, vol. 1 (6th ed.; New York: Wiley)  
Walker, T. P., Steigman, G., Schramm, D. N., Olive, K. A., & Kang, H.-S. 1991, ApJ, 376, 51  
White, S. D. M., Efstathiou, G., & Frenk, C. S. 1993, MNRAS, 262, 1023  
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  
Zaroubi, S., Zehavi, I., Dekel, A., Hoffman, Y., & Kolatt, T. 1997, ApJ, 486, 21