DYNAMICAL FIELDS AND Ω SIMULTANEOUSLY FROM REDSHIFT AND PECULIAR VELOCITY SURVEYS

A. Nusser [1] and A. Dekel [2]
[1] *Institute of Astronomy, Madingley Rd, Cambridge, CB3 0HA, England*
[2] *Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel*

Abstract

We present a method (termed SIMPOT) for a simultaneous reconstruction of the large-scale density, velocity and potential fields and the parameter $\beta \equiv \Omega^{0.6}/b$ (where Ω is the cosmological density parameter and b is the linear biasing parameter) from the observed radial peculiar velocities and the distribution of galaxies in redshift space. SIMPOT models the velocity potential field by means of an expansion in spherical harmonics and radial Bessel functions, and, assuming a value for β, relates it to the distribution of galaxies in redshift space using linear theory of gravitational instability. The method is tested using realistic mock catalogs of Mark III peculiar velocities and IRAS 1.2Jy redshifts. The mock data are extracted from an $\Omega = 1$ N-body simulation which mimics the nearby universe. If linear biasing prevails in the universe, SIMPOT provides an unbiased estimate for β with an error of ± 0.06. SIMPOT can be applied with different weights to the velocity and redshift data. With high weight assigned to velocities, the recovered fields are mainly determined by the velocities, but properly regulated by the distribution of galaxies in regions of poor velocity data. SIMPOT can be viewed as a general method for recovering large scale dynamical fields from a multitude of heterogenous data.

1 Introduction

Three related objectives of current cosmological research are (a) determination of the universal density parameter Ω, (b) recovery the underlying large scale fields of peculiar velocity and mass density in the local universe, and (c) establishing the relationship between the distribution of galaxies and mass.

Two complementary kinds of data are available: (i) complete redshift surveys of galaxies, which serve as (perhaps biased) tracers of the mass distribution, and (ii) samples of galaxies with redshifts and inferred distances, which serve as tracers of the peculiar velocity field. These data are commonly analyzed separately and compared in retrospect [4]. On one hand, redshift surveys yield the mass distribution and the corresponding velocity field under a given biasing relation and Ω. The apparent anisotropy of clustering in redshift surveys can be used to estimate Ω assuming a biasing relation [10] [8] [2]. Peculiar velocity samples, on the other hand, yield the velocity field based on the potential nature of gravitating flows, and then the corresponding density field using gravitational instability and an assumed Ω. The recovered velocity field is used to estimate Ω independently of the distribution of galaxies and the biasing relation [17] [7] [1]. At a later stage, the recovered fields from the two inputs are tested for mutual consistency according to gravitational instability and linear biasing, and then compared to yield a value of β [6] [11] [15] [3] [19].

Here we make a first attempt at combining the data from redshift surveys and peculiar velocity samples in recovering the dynamical fields and simultaneously determining β. The simultaneity is advantageous because the data are complementary: the gravitational velocities from the redshift surveys help filling in regions which are poorly sampled by peculiar-velocity data, while the observed velocities help constraining the biasing for the redshift-data analysis, and the bulk velocity which arises from outside the volume covered by the survey. A simultaneous fit is more likely to converge to the true solution than a successive fit. The fact that a simultaneous fit allows applying exactly the same variable effective smoothing to both data sets is important especially because the biasing could be scale dependent.

In §2 we outline the method. In §3 we test the method using mock data. In §4 we discuss the method and its application.

2 The method

The data consist of two independent samples of galaxies: a whole-sky collection of radial peculiar velocities inferred by Tully-Fisher-like distance indicators, and a complete redshift survey with a known selection function. Our goal is to find the velocity potential field $\Phi(\boldsymbol{r})$, and the parameter β, which simultaneously fit the two data sets. We formulate a parametric model $\tilde{\Phi}$ for the velocity potential field by means of a finite expansion in a set of spherical harmonic and radial Bessel functions [18] [9],

$$\tilde{\Phi}(\boldsymbol{r}) = -\frac{1}{2}\mathcal{H}r^2 - \boldsymbol{B}\cdot\boldsymbol{r} + \sum_{l=1}^{l_{max}}\sum_{m=-l}^{l}\sum_{n=1}^{n_{max}} \Phi_{lmn}\, j_l(k_n r)\, Y_{lm}(\hat{r}). \tag{1}$$

We incorporate in the model a monopole term corresponding to a Hubble-like peculiar flow $\mathcal{H}r$ (or a constant offset in density), and a bulk flow term corresponding to a constant velocity $\boldsymbol{B}$. The model free parameters are Φ_{lmn}, B_i, and $\mathcal{H}$. The corresponding model for the radial peculiar velocity field is

$$\tilde{u}(\boldsymbol{r}) = -\frac{\partial\Phi}{\partial r} = \mathcal{H}r + \boldsymbol{B}\cdot\hat{r} + \sum \Phi_{lmn}\, k_n \left(j_{l+1}(k_n r) - l\frac{j_l(k_n r)}{k_n r}\right) Y_{lm}(\hat{r}). \tag{2}$$

Away of triple-value zones in redshift space, the density fluctuation field is related to the velocity potential by linear gravitational instability theory [15],

$$\delta_s(\boldsymbol{s}) = \frac{1}{f(\Omega)}\nabla^2\Phi + \frac{1}{s^2}\frac{\partial}{\partial s}\left(s^2\frac{\partial\Phi}{\partial s}\right), \tag{3}$$

where $f(\Omega) \approx \Omega^{0.6}$. The second term on the r.h.s represents the apparent density enhancement in redshift space as a result of the mapping from distance to redshift. Note that this second term breaks the linear scaling of the potential with f.

The "observed" mass density $\delta_s(\boldsymbol{s})$ has to be inferred from the galaxy redshift survey. The relation between galaxies and mass is modeled by "linear biasing", where the local density fluctuations galaxies and mass are related via a constant biasing factor b. In linear theory, the parameters b and Ω in all the dynamical relations combine into $\beta \equiv f/b$. Since the redshift survey is flux limited, the true number density of galaxies is estimated by weighting the observed number density, $n_0(\boldsymbol{s})$, by the inverse of the selection function of the survey, $\phi(\boldsymbol{r})$. The selection function ϕ must be evaluated at the real space position corresponding to $\boldsymbol{s}$, which is unknown a priori. Approximating $\phi(\boldsymbol{r})$ by the measurable $\phi(\boldsymbol{s})$ leads to systematic biases [12]. Instead, we approximate: $\phi(\boldsymbol{r}) \approx \phi(\boldsymbol{s}) - (\partial\phi/\partial s)u(\boldsymbol{s})$. The field we choose to express in terms of the model parameters is the fluctuation in $n_0(\mathbf{s})/\phi(\mathbf{s})$, which can be determined from the observations. We obtain using (1), (3), and (4)

$$\tilde{\delta}_G(\mathbf{s}) = -2\frac{\mathbf{B}\cdot\hat{s}}{s} + \sum_{lmn}\Phi_{lmn}\left[-\beta^{-1}k_n^2 j_l(k_n s) + b_{ln}(s)\right] Y_{lm}(\hat{s}), \tag{5}$$

where,

$$b_{ln} = k_n^2 l\left(l+1+\frac{\mathrm{d}\ln\phi}{\mathrm{d}\ln s}\right)\frac{j_l(k_n s)}{(k_n s)^2} - k_n^2\frac{\mathrm{d}\ln\phi}{\mathrm{d}\ln s}\frac{j_{l+1}(k_n s)}{k_n s} - k_n^2 j_l(k_n s), \tag{6}$$

and we have omitted the term involving $\mathcal{H}$ so that any uncertainty in the form of a constant offset in the density is absorbed into the Hubble-like flow in the velocity model.

The expansion of an arbitrary radial function into a set of Bessel functions requires a choice of boundary conditions appropriate for the problem in question. These boundary conditions determine the wavenumbers k_n in the expansion. In any realistic application, it is necessary to set a maximum radius, R_{max}, within which the data are sufficiently reliable to constrain the potential field. It is at R_{max} where the outer boundary conditions should be specified. We are guided by the asymptotic behavior of a peculiar velocity field generated by fluctuations in the mass-density distribution. If we assume that the density of matter beyond R_{max} is constant, then potential theory dictates that the radial dependence of the radial velocity u [eq. (2)] at $r > R_{max}$ is $u \propto r^{-(l+2)}$ at $r > R_{max}$ [9]. This leads to the condition $j_{l-1}(k_n R_{max}) = 0$.

2.1 Simultaneous Estimation of β and Φ_{lmn}

In the usual sense of the theory of data modeling, the parameters of the model are found by achieving a maximum in the probability that the model produces the observations within their measurement errors.

There is no ambiguity in formulating the probability for observing the peculiar velocities given the model. If the errors, σ_{ui}, in the measurements of the peculiar velocities, u_{0i}, are uncorrelated and Gaussian, then the probability of the occurrence of the observations given the model $\tilde{\Phi}$ is,

$$\Pr[u_0|\tilde{\Phi}] \propto \exp\left[-\frac{\chi_u^2}{2}\right] \equiv \exp\left\{-\frac{1}{2}\sum_{i=1}^{N_g'}\sigma_{ui}^{-2}\left[u_{0i} - \tilde{u}(\mathbf{r}_i)\right]^2\right\}, \tag{8}$$

where the sum is over all galaxies in the sample of peculiar velocities. The errors, σ_{ui}, are mainly a result of the intrinsic scatter of the distance indicator used to derive the peculiar velocities and partly of observational uncertainties, e.g., inclination effects in measuring rotational velocities in the case of the Tully-Fisher distance indicator.

Less obvious is how an "observed" density field. Here we define a density-contrast field in redshift space by [9],

$$\delta_{0s}(\mathbf{s}) = \frac{1}{\bar{N}}\sum_{i=1}^{N_g}\frac{1}{\phi(\mathbf{s}_i)}\delta_D(\mathbf{s}-\mathbf{s}_i) - 1, \tag{10}$$

where $\bar{N}$ is the true mean number density of galaxies and the sum is over all galaxies in the survey. Since the contribution of each galaxy is weighted by the selection function at its redshift rather than its distance, the definition (10) is a biased estimate of the true density. However, this bias is properly mitigated by the model (5). In order to estimate the shot-noise errors in (10), we divide space into cubic cells of volume ΔV centered at positions $\mathbf{s}_\alpha$ of a uniform grid. We approximate the Dirac delta function by $\delta_D(\mathbf{s}-\mathbf{s}_i) \approx 1/\Delta V$ if $\mathbf{s}$ and $\mathbf{s}_i$ fall in the same cell and zero otherwise. Then (10) implies that the number density excess, $\delta(\mathbf{s}_\alpha)_{0\Delta V}$, inside a cell α is constant. In this scheme, the variance in the determination of (10) due to shot-noise is the inverse of the number density expected in each cell, i.e., $\sigma_\delta^2(\mathbf{s}_\alpha) = \frac{1}{\Delta V \bar{N}\phi(\mathbf{s}_\alpha)}$. Clearly the shot-noise errors in different cells do not correlate. Therefore, if we limit the analysis to modeling the the large scale density field, the probability that $\delta_{0\Delta V}$ is realized in a discrete sampling of $\tilde{\delta}_G$ is Gaussian which, in the limit $\delta V \to 0$, becomes,

$$\Pr[\delta_{\bullet\Delta V}|\tilde{\Phi},\beta] \propto \exp\left[-\frac{\chi_\delta^2}{2}\right] \equiv \exp\left\{-\frac{1}{2}\bar{N}\int d^3s\phi(\mathbf{s})\left[\delta_{0s}(\mathbf{s})-\tilde{\delta}_G(\mathbf{s})\right]^2\right\}. \quad (11)$$

Now we can write the probability for reproducing the observed peculiar velocities and density field as,

$$\Pr[u_0,\delta_0|\tilde{\Phi},\beta] = \Pr[u_0|\tilde{\Phi}]\Pr[\delta_0|\tilde{\Phi},\beta] = A\exp\left[-\frac{1}{2}\left(\chi_u^2+\chi_\delta^2\right)\right], \quad (14)$$

where A is a normalization factor which is independent of β and the model parameters.

The problem of rendering the probability (14) stationary reduces to minimizing the function $\chi^2 = \chi_\delta^2 + \chi_u^2$ subject to variations in β and Φ_{lmn}. At the minimum the derivatives $\partial\chi^2/\partial\beta$ and $\partial\chi^2/\partial\Phi_{lmn}$ vanish, yielding a set of equations involving nonlinear coupling between β and Φ_{lmn}. However, for a given β, the equations $\partial\chi^2/\partial\Phi_{lmn} = 0$ are linear in Φ_{lmn}. Therefore we choose to determine $\Phi_{lmn}(\beta)$ using standard techniques for solving linear equations. Then we adjust the best estimate for β simply by examining the curve of $\chi^2\left[\Phi_{lmn}(\beta),\beta\right]$.

SIMPOT circumvents a major drawback of direct comparisons, say, between the 1.2JyIRAS density and the velocity divergence as derived from POTENT. Were β to be estimated in such comparisons from the regression of the velocity divergence on the density or vice versa, it would be biased due to the presence of errors in the density or the velocity divergence. The reason for the bias in regression lines, is that the dynamical relation between velocity and density is forced to hold between the measured fields which are contaminated by errors. Direct methods usually allude to extensive Monte Carlo simulations in order to take account of the biases in the estimates of β [6]. SIMPOT, however, imposes the dynamical relations in the model. It then seeks best estimates for β and the model parameters by maximizing the probability (14). The dynamical fields which are contrived by SIMPOT are modulated by the observed peculiar velocities and the density field of galaxies. In regions where the velocity data are poor the fields are mainly constructed out of the density field. In regions where the errors in the observed density field are large due to shot-noise while the velocity data are reliable, the fields are affected more by the observed velocities than by the density field.

The constraint on β in SIMPOT is two-faced. On the one hand, the modulation of the potential model by the velocities given β affects the value of χ_δ^2. To see that, consider an unrealistic case of accurate measurements of peculiar velocities and very dilute redshift survey so that the model is solely determined by the velocities independently of β. But then the value of χ_δ^2 strongly depends on β. For β lower the actual value, the density fluctuations as given in terms of potential model by (5) are unrealistically large, causing a large χ_δ^2. For a higher β, larger the actual value, the density fluctuations are underestimated leading to a large χ_δ^2. Note however that according to (5), the density fluctuations do not vanish in the limit $\beta \to \infty$ because the redshift distortion correction terms are independent of β. On the other hand, the value of χ_v^2 is affected by the observed density field in redshift space which introduces an implicit dependence of χ_v^2 on β.

An important question which arises when analyzing two different kinds of data is how well are they consistent given the hypothesis that they describe the same underlying fields? In the rigorous statistical sense, the value of χ^2 per d.o.f is a good indication of a goodness of fit and hence of the agreement between the two kinds of data. However, in the methodology presented here the actual value of the χ^2 is not a good discriminator for the following reasons. First, the χ_δ^2 contains a constant term which is an integration of squares of Dirac delta functions arising from our description of the observed density. This term is infinite but it does not affect the solution for the fit parameters. Clearly, the χ_δ^2 contains a divergent term because the number of d.o.f in the description of the observed density field is infinite. Second, our model does not incorporate small scale velocity scatter present in the peculiar velocities. Third, SIMPOT neglects nonlinear effects in relating the density to the potential model. To some extent these effects might be important in the real universe in some regions such as the Great Attractor.

SIMPOT provides an answer to this question by reconstructing the potential field with different weights assigned to the different data. By inspection of the differences in the reconstructions with different weights one may assess the consistency between the data sets. Thus let us define a function Ξ^2 by,

$$\Xi^2 = \lambda\chi_u^2 + (1-\lambda)\chi_\delta^2, \tag{15}$$

where λ is a free parameter between 0 and 1. A reconstruction of the potential field with $\lambda \to 0$ heavily relies on the observed density field. If $\lambda \to 1$ the reconstruction makes little use of the density field and the recovered potential is dictated by the peculiar velocities. In the reconstruction with $\lambda \neq 0.5$ the value of β derived from minimizing χ^2, i.e. Ξ^2 with $\lambda = 0.5$, must be used. The reason is simply that an estimate β obtained by, $\partial\Xi^2/\partial\beta = 0$ with $\lambda \neq 0.5$, is biased. For instance if $\lambda = \to 1$ then the model is determined by the peculiar velocities and χ_u^2 depends very weakly on *beta*. Therefore, the dependence of Ξ^2 on β is due to χ_δ^2 which is mearly a measure of error weighted scatter of the difference between the density as inferred by the velocities given β and the observed density. Therefore estimating β by $\partial\Xi^2/\partial\beta = 0$ is similar, for $\lambda \to 1$, to computing the regression line of the inferred on observed densities. This estimate is clearly biased. Note however that once β is fixed by minimizing χ^2 the potential model is linear in the remaining parameters which can be determined without any bias by minimizing Ξ^2 with $\lambda \neq 0.5$.

3 Testing with with a mock-universe simulation

The systematic biases in the data depend on the morphology of the universe. Therefore it is essential to test the method using a numerical simulation which mimics the large scale structure of the local universe [13]. Such a simulation is generated by integrating forward in time the initial density fluctuations obtained from the 1.2Jy IRAS survey using the Zeldovich-Bernoulli time machine [16]. The simulation was halted at $\sigma_8 = 0.7$. Mock 1.2Jy IRAS and Mark III catalogs were then generated from the simulation. To further improve the realism of the application of SIMPOT and discern the effects of various sources of systematic biases on the estimation of β, 1.2Jy IRAS catalogs with and without attenuation of galaxies in clusters have been generated; the Malmquist bias in the mock Mark III catalog has been corrected for according to the density field from the mock 1.2Jy IRAS using the recipe of Dekel *et al.* (1993). For comparison a mock Mark III free of Malmquist bias has been used as well.

Figure 1 shows maps of the density fluctuations in real space from the mock unbiased 1.2Jy IRAS and Malmquist bias corrected Mark III. The maps are smoothed with a Gaussian window of width 1000km/s. Linear theory was used to derive the density from the potential field. The lower right-hand panel shows a reconstruction which is mainly based on the velocity data. The velocity data, being very noisy and sparse, are not sufficient to constrain the density field in space. Hence SIMPOT was applied with $\lambda = 0.99$, giving some weight to mock IRAS data in order to regularize the minimization procedure. Here the reconstruction overestimates the density fluctuations as expected in linear theory [14]. The lower left panel is a reconstruction from the mock IRAS data with $\lambda = 0$. This is basically a test of the recovery of the density field in real space from the redshift distribution of galaxies. The recovered density field greatly resembles the true field. The reconstruction by SIMPOT with $\lambda = 0.5$ is shown in the upper left panel. The density field in this case correlates very well with the density recovered with $\lambda = 0$ and is consistent, within the errors, with the reconstruction with $\lambda = 0.99$. Given the sparse sampling and the large errors in peculiar velocity surveys the recovery with $\lambda = 0.5$ is in general more sensitive to the density data.

Figure 2 shows curves of $\chi^2(\beta)$ from SIMPOT with $\lambda = 0.5$ for the various variants of the mock Mark III and 1.2Jy IRAS as explained in the figure caption. We see that SIMPOT is successful at recovering the true value of β with random errors of ± 0.06. A key point here is that the systematic uncertainties in the determination of β due to Malmquist bias corrections and the bias of IRAS galaxies around clusters do not amount to a serious discrepancy.

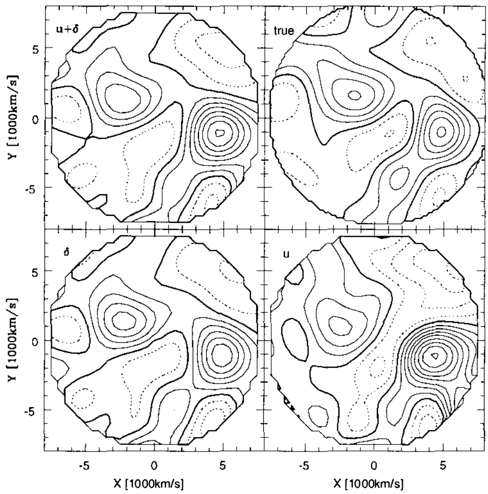

Figure 1: Maps of $-\nabla \cdot \boldsymbol{v}$ in the "supergalactic" plane in the universe simulation. Upper left panel is the true $-\nabla \cdot \boldsymbol{v}$. The three remaining panels are the reconstruction by SIMPOT with $\lambda = 0.5$ (upper right), $\lambda = 0$ (lower left) and $\lambda = 0.99$ (lower right). Contour spacing is 0.2. Heavy, solid and dashed contours denote average, positive and negative fluctuations.

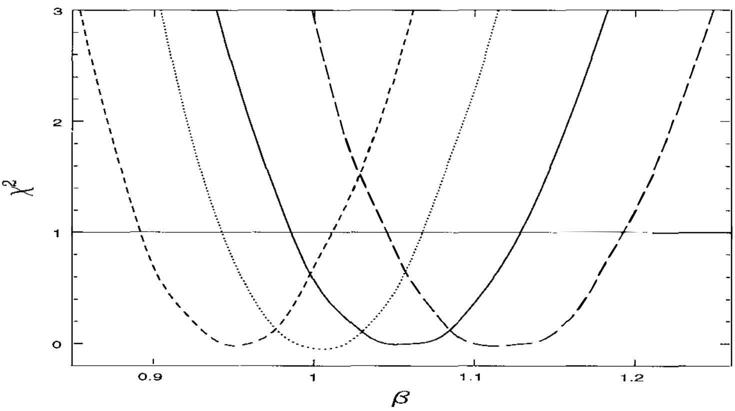

Figure 2: Curves of $\chi^2(\beta) - \chi^2_{min}$ for $\lambda = 0.5$. Short-dashed and dotted lines respectively correspond to unbiased and biased 1.2Jy IRAS mock catalogs with only random errors in the peculiar velocities of mock Mark III . Solid and dotted lines correspond to unbiased and biased mock 1.2Jy IRAS but with random errors and Malmquist bias in Mark III. The Malmquist bias has been corrected for by the density of mock 1.2Jy IRAS.

4 Discussion

In general SIMPOT can be viewed as a method for recovering the underlying large scale fields in the universe given an ensemble of heterogeneous data samples. We have demonstrated that the method, when applied to mock Mark III and 1.2Jy IRAS catalogs, successfully recovers the underlying fields and provides an unbiased estimate for the parameter β. The recovered fields are smoothed independently of the details of individual data sets. This is greatly desirable when assessing the general consistency of different kinds of data on different scales. SIMPOT recovers fields which are locally regulated in space according to the quality of the various data. For example, in the regions where the distribution of galaxies is very dilute but measurements of peculiar velocities are ample, the recovered fields are mainly determined by the velocities. For the purpose of recovering the underlying fields, SIMPOT can in practice be applied with varying weights to the different data. When high weight is given to the velocity data, SIMPOT is essentially a method for recovering the large scale peculiar velocity fields. The interpolation from regions where the velocity data are of good quality to empty regions is modulated self-consistently by the peculiar velocity field as predicted from the distribution of galaxies in redshift space.

Acknowledgements. AN thanks Y.Hoffman for stimulating discussions. AN acknowledges the support of a PPARC postdoctoral fellowship.

References

[1] Bernardeau, F., Juszkiewicz, Dekel, A. and Bouchet, F. 1995, *M.N.R.A.S.* **274** 20

[2] Cole, S., Fisher, K.B., and Weinberg, D. 1994, *M.N.R.A.S.* **267** 785

[3] Davis, M., and Nusser, A. 1995 in *The Maryland Meeting on Dark Matter*

[4] Dekel, A. 1994, Ann. Rev. of A & A

[5] Dekel, A., Bertschinger, E., & Faber, S. M. 1990, *Astrophys. J.* **364** 349

[6] Dekel, A., Bertschinger, E., Yahil, A., Strauss, M., Davis, M., & Huchra, J. 1993, *Astrophys. J.* **412** 1

[7] Dekel, A. and Rees, M. 1993, *Astrophys. J. (Lett.)* **422** L1

[8] Fisher, K.B., Scharf, C.A., and Lahav, O. 1994, *M.N.R.A.S.* **266** 219

[9] Fisher, K.B., Lahav, O., Hoffman, Y., Lynden-Bell, D., and Zaroubi, S. 1995, *M.N.R.A.S.* **272** 885.

[10] Hamilton, A.J.S. 1993, *Astrophys. J. (Lett.)* **406** L47

[11] Hudson, M.J., Dekel, A., Courteau, S., Faber, S.M., and Willick, J. 1995, *M.N.R.A.S.* **274** 305

[12] Kaiser, N. 1987, *M.N.R.A.S.* **227** 1

[13] Kolatt, T., Dekel, A. 1995, to be submitted.

[14] Nusser, A., Dekel, A., Bertschinger, E., & Blumenthal, G. R. 1991,

[15] Nusser, A, & Davis, M. 1994, *Astrophys. J. (Lett.)* **421** L1

[16] Nusser, A, & Dekel, A. 1992, *Astrophys. J.* **391** 443

[17] Nusser, A, & Dekel, A. 1993, *Astrophys. J.* **405** 437

[18] Regös, E., & Szalay, A. 1989, *Astrophys. J.* **345**, 627

[19] Willick, J. *et al.* 1994, in preparation.