Mon. Not. R. Astron. Soc. **290,** 193–202 (1997)

Modelling the evolution of correlation functions in gravitational clustering

Dipak Munshi[1]★ and T. Padmanabhan[2]★

[1] *Queen Mary and Westfield College, Mile End Road, London E1 4NS*
[2] *Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, India*

Accepted 1997 May 13. Received 1997 April 17; in original form 1996 July 1

ABSTRACT

Padmanabhan has suggested a model to relate the non-linear two-point correlation function to the linear two-point correlation function. In this paper, we extend this model in two directions: (1) by averaging over the initial Gaussian distribution of density contrasts, we estimate the spectral dependence of the scaling between non-linear and linear correlation functions; (2) by using a physically motivated Ansatz, we generalize the model to N-point correlation functions and relate the non-linear, volume averaged N-point correlation function $\bar{\xi}_N(x, a)$ with the linearly extrapolated volume averaged two-point correlation function $\bar{\xi}_2(l, a)$ evaluated at a different scale. We compare the point of transition between the different regimes obtained from our model with that from numerical simulations, and show that the spectral dependence of the scaling relations seen in the simulations can be easily understood. A comparison of the calculated form of $\bar{\xi}_N$ with the simulations shows reasonable agreement. We discuss several implications of the results.

Key words: methods: analytical – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

There is growing evidence that the large-scale structure in the Universe is formed through the gravitational amplification of small inhomogeneities. Semi-analytic modelling of the gravitational clustering of collisionless, non-relativistic dark matter particles will be of significant utility in understanding the formation of large-scale structures. Such modelling is straightforward when the density contrasts are small and perturbation theory, based on a suitably chosen small parameter, is applicable (see, e.g., Fry 1984; Moutarde et al. 1991; Buchert 1992; Bernardeau 1992). At the other extreme, highly non-linear regimes can be handled if one is prepared to make some extra assumptions like stable clustering (Peebles 1980) or those which underlie the Press–Schecter formalism, Peaks formalism, etc. (Bardeen et al. 1986). The intermediate regime is considerably more difficult but some progress has been made recently even in this case (Padmanabhan 1996; also see Hamilton et al. 1991 and Nityananda & Padmanahban 1994), using the scale-invariant spherical infall models. These papers give an expression for the non-linear mean correlation function in terms of the linear mean correlation function in both the intermediate and non-linear regimes. To do so, Padmanabhan (1996) concentrated on a typical spherical region and ignored the effects arising out of averaging over peaks of different sizes. Also, no attempt was made to model higher-order correlation functions. In this paper, we generalize these ideas in two directions.

(1) We consider peaks of different heights and average over them, along the lines suggested in Padmanabhan et al. (1996). Such an averaging will introduce spectrum dependence in the relation connecting non-linear and linear correlation functions which have been noted in numerical simulations (Padmanabhan et al. 1996; Jain, Mo & White 1995; Peacock & Dodds 1996). We will show that this dependence can be understood from our model.

(2) We shall postulate an Ansatz for the higher-order correlation functions, generalizing the result for the two-point correlation function. Using this Ansatz, we shall calculate the S_N parameters for both the intermediate and non-linear regimes and compare them with the simulations.

★E-mail: D.Munshi@qmw.ac.uk (DM);
nabhan@iucaa.ernet.in (TP)

We will see that there is reasonable – though not excellent – agreement between the model and simulations, suggesting that the Ansatz we have proposed is along the right lines.

2 THE MODEL AND THE ANSATZ

The basic idea behind the model used in Padmanabhan (1996) can be described as follows. Consider the evolution of density perturbations starting from an initial configuration which we take to be a realization of a Gaussian random field with variance σ. A region with initial density contrast δ_i will expand to a turnaround radius $x_{\rm ta}=x_i/\delta_i$ and will finally collapse to an object of radius $x_{\rm f}$ which will contribute to the two-point correlation function an amount proportional to $(x_i/x_{\rm f})^3$. The initial density contrast within a *randomly* placed sphere of radius x_i will be $\nu\sigma(x_i)$ with a probability to exp $(-\nu^2/2)$. On the other hand, the initial density contrast within a sphere of radius x_i, *centred around a peak in the density field*, will be proportional to the two-point correlation function and will be $\nu^2\bar{\xi}(x_i)$ with a probability proportional to exp $(-\nu^2/2)$. (We have obtained the quadratic scaling in ν based on the assumption that $\bar{\xi}$ scales in the same way as mean square fluctuations in the mass, which – in turn – will scale as the mean square of the Gaussian density field. In general, one expects the scaling to be ν^α with $\alpha\approx 2$–1. The results are easily generalizable to any value of α. We will stick to $\alpha=2$ since it gives reasonable agreement with simulations and is based on simple considerations.) It follows that the contribution from a *typical* region will scale as $\bar{\xi}_{\rm nl}\propto\bar{\xi}_i^{3/2}$ while that from *higher peaks* will scale as $\bar{\xi}_{\rm nl}\propto\bar{\xi}_i^{3}$. In the intermediate phase, most dominant contribution arises from high peaks and we find the scaling to be $\bar{\xi}_{\rm nl}\propto\bar{\xi}_i^{3}$. The non-linear virialized regime is dominated by the contribution from several typical initial regions and has the scaling $\bar{\xi}_{\rm nl}\propto\bar{\xi}_i^{3/2}$. This was essentially the feature pointed out in Padmanabhan (1996), though in that work it was assumed that $\nu=1$. To take into account the statistical fluctuations of the initial Gaussian field we can average over different ν with a Gaussian probability distribution. (Strictly speaking, there will be deviations from pure Gaussian distribution because our averaging requires a mapping from Lagrangian to Eulerian coordinates; we shall ignore this because it is a higher-order effect.) We shall do this calculation in the next section.

To generalize the above ideas for higher-order correlation functions is more non-trivial. In general N-point correlation functions will depend on shapes but volume averaging will remove this shape dependence. The 'S_N parameters' are then defined as dimensionless ratios of $\bar{\xi}_N(x,a)$ and $\bar{\xi}_2(x,a)^{(N-1)}$. Such volume-averaged N-point functions (which can be directly related to counts-in-cells) and the S_N parameters have been studied extensively in the literature (White 1979; Balian & Schaeffer 1984; Bouchet & Hernquist 1992). The S_N parameters show a fairly simple pattern of behaviour both in the perturbative and non-linear regimes. It can be shown that all S_Ns can be evaluated from the spherical collapse model in the limit $\bar{\xi}_2\to 0$. In this limit, they are constant and depend only on the initial spectral index when smoothing is taken into account. They are also expected to be constants in the non-linear regime. These results indicate that the hierarchical pattern, which is generally assumed to describe non-linear $\bar{\xi}_N$ functions, could have a larger range of validity. We shall exploit this possibility to estimate the S_N in the intermediate and non-linear regimes along the following lines.

The evolution of N-point correlation functions is described by momentum moments of BBGKY hierarchical equations, which can be expressed in the form

$$\frac{\partial Q_N}{\partial t}-\frac{1}{\alpha}\sum\frac{\partial}{\partial x_i^\alpha}(Q_N v_i^\alpha)=0. \tag{1}$$

Here α varies from 1 to N, i varies over the Cartesian components and Q_N is the full N-point correlation function given by

$$Q_N(1,2,\ldots,N)=1+\xi_2(1,2)+\ldots+\xi_3(1,2,3)+\ldots+\xi_N(1,2,3,\ldots,N), \tag{2}$$

and ξ_N denotes the reduced part of N-point correlation function. For the two-point correlation function, the resulting equation can be simplified to (Peebles 1980)

$$\frac{\partial\bar{\xi}_2}{\partial t}+\frac{1}{ax^2}\frac{\partial}{\partial x}[x^2(1+\bar{\xi}_2)v]=0, \tag{3}$$

which describes the conservation of pairs. In the integral form, the same result can be expressed as

$$x^3[1+\bar{\xi}_2(x)]=l^3, \tag{4}$$

where $l=\langle x_i^3\rangle^{1/3}$ is the average initial scale from which collapsed structures of size x have formed.

Our aim is to generalize the above result for higher-order correlation functions, but it is obvious that one cannot get such a simple relation for higher-order correlation functions in which N–1 different length-scales are present. To make progress, one needs to assume that, although there are different length-scales present in the reduced N-point correlation function, all of them have to be roughly of the same order to give a significant contribution. This is supported by the fact that, by its very construction, the reduced N-point correlation function vanishes when a single point or group of points from this set of N points is moved to large separation (in that limit, the correlation function is just the product of lower-order reduced correlation functions). For a geometrical picture, one can think of a polyhedron, inscribed in a sphere, with particles at each vertex having their velocities directed towards the centre of the sphere. (This configuration has the relative velocity of particles directed along their relative separation and hence can satisfy the stable clustering Ansatz.) For such a configuration, the scale in the correlation function will be the radius of the sphere circumscribing the polyhedron. If the correlation functions are described by a single scale, then a natural generalization of equation (4) will be

$$\bar{\xi}_N\approx\langle x_i^{3(N-1)}\rangle/x^{3(N-1)}. \tag{5}$$

The validity of such an Ansatz is open to question and we do hope to check it directly in numerical simulations in a future work. In this paper we accept the above Ansatz as a working hypothesis and use it to calculate the S_N parameters in different regimes. Since these parameters have been studied numerically we can directly test predictions of this

Ansatz against existing results to obtain a feel for the validity of the Ansatz. It may be noted that, even though several models have been proposed to predict the values of S_N parameters (Hamilton 1988; Fry 1984; Schaeffer 1984; Balian & Schaeffer 1989; Bernardeau & Schaeffer 1992) they fail to reproduce correct values for lower-order S_N parameters. We shall show that the non-linear values of lower-order S_N parameters are predicted with a fair degree of accuracy in our model.

3 THE CORRELATION FUNCTIONS IN DIFFERENT REGIMES

We shall now consider the implementation of the above ideas in three different regimes of gravitational clustering. We shall call the first one the 'perturbative regime', in which we expect perturbation theory is valid. The second regime (which we call the 'intermediate' regime) is dominated by the scale-invariant radial infall of high peaks. The third regime ('non-linear') is dominated by virialized blobs of matter. While we are mainly interested in the latter two regimes, we shall begin with some important observations regarding the perturbative regime.

Perturbative regime

We divide the density field into two parts at each point, with one part coming from the spherical collapse (which we call the 'monopole' part) and the rest of the contribution coming from higher-order spherical harmonics, characterizing shear, tide and non-linear coupling between them:

$$\delta(x, a) = \delta_{\rm sph}(x, a) + \varepsilon(x, a). \tag{6}$$

It should be noted that one *cannot* assume $\delta(x) = \delta_{\rm sph}(x)$ for each point. While this may be obvious from symmetry considerations, a more formal argument can be given along the following lines.

Let us assume for a moment that we can set $\varepsilon = 0$. Since the growth of density contrast in the spherical collapse model is well known (Peebles 1980), we can expand $\delta_{\rm sph}(x, a)$ in a Taylor series to get

$$\delta(x, a) = \delta_{\rm sph}(x, a) = \sum_{N=1}^{\infty} \delta_{\rm sph}^{(N)}(x, a) = \sum_{N=1}^{\infty} \frac{\mu_N}{N!}\, \delta^{(1)N}, \tag{7}$$

where $\mu_2 = 34/21$, $\mu_3 = 682/189$, $\mu_4 = 446, 440/43, 659\ldots$. It is clear that, in spherical collapse, $\delta^{(N)} \propto \delta^N$. On the other hand, since $\langle\delta\rangle = 0$ we demand $\langle\delta^{(N)}\rangle = 0$ at every order of perturbation. This implies, in the case of the spherical collapse model, that $\langle\delta^N\rangle = 0$ for all N, i.e. moments of δ^N vanish at all orders; hence we must conclude that δ vanishes at each point identically. Clearly, we cannot set $\xi = 0$ in equation (6).

For a generic Gaussian field, we have to work with a δ which has two parts, one coming from pure spherical collapse and the other part, ε, related to deviation from spherical collapse dynamics. The Taylor series will then be

$$\delta(x, a) = \sum_{N=1}^{\infty} \delta^{(N)}(x, a) = \sum_{N=1}^{\infty} \frac{\mu_N}{N!}\, \delta^{(1)N} + \sum_{N=2}^{\infty} \varepsilon_N a^N, \tag{8}$$

where we have expanded ε in a perturbative series with ε_N being of order δ^N. (Note that there is no contribution to ε in the linear order.) Although both the terms will be important for a generic random field, it has already been shown (Bernardeau 1994a) that ε becomes less dominating for rare events, i.e. for large values of $\nu = (\delta/\sigma)$. In the perturbative regime, where σ is small, any deviation from homogeneity is a rare event and hence one can assume that, statistically, ε will be close to zero at most of the points. One can explicitly demonstrate this claim by calculating the parameters $S_N = \langle\delta^N\rangle_{\rm c}/\langle\delta^2\rangle_{\rm c}^{(N-1)}$ in the limit $\varepsilon_N \to 0$ (where the subscript c denotes connected parts) and showing that it will reproduce the well known results of S_N derived earlier by summing up all tree level diagrams in the limit $\sigma \to 0$ (Bernardeau 1992). Consider, for example, the case of S_3. We have

$$S_3 = \langle\delta^3\rangle/\langle\delta^2\rangle^2 = 3\langle\delta^{(1)^2}\delta^{(2)}\rangle/\langle\delta^{(1)^2}\rangle^2 \tag{9}$$

$$= 3(\langle\delta^{(1)^2}\rangle\langle 1/2\mu_2\delta^{(1)^2} + \varepsilon_2\rangle + \langle\varepsilon_2\delta^{(1)^2}\rangle + \mu_2\langle\delta^{(1)^2}\rangle^2)/\langle\delta^{(1)^2}\rangle^2. \tag{10}$$

The first term vanishes because $\langle\delta^{(2)}\rangle = 0$ and the second term gives us vanishing contribution in the limit $\varepsilon_2 \to 0$ and we get the well known result $S_3 = 3$. A similar calculation for higher-order moments reproduces tree level results of perturbation theory, $S_4 = 4\mu_3 + 12\mu_2^2$ etc. These are the exact values of S_N parameters in the limit $\sigma^2 \to 0$ (i.e. at the tree level of perturbative calculation neglecting all loop corrections) obtained previously by rigorous analytical calculations (Bernardeau 1992, 1994a, b, c, 1995). Our analysis reconfirms that any deviation from spherical dynamics does not alter the values of S_N parameters at tree level in which only the monopole part of the dynamics is relevant. The higher-order harmonics (shear, stress and their couplings) start contributing only from loop level. (This is also true for approximation schemes like the Zeldovich approximation, etc; see Munshi, Sahni & Starobinsky 1994.)

Intermediate regime

In the intermediate regime, we concentrate on the collapse of regions around peaks in which the density contrast scales as the correlation function. [We shall work with a $\Omega = 1$ universe.] Consider a spherical region of initial radius x_i and overdensity $\delta = \nu^2\bar{\xi}_L(x_i) = \sigma_0^2\nu^2 x_i^{-(n_{\rm p}+3)}$ where $n_{\rm p}$ is the index of the initial power spectrum and σ_0 is a constant. This region will expand to a maximum radius $x_{\rm ta} = (x_i/\delta) \propto \nu^{-2}x_i^{(n_{\rm p}+4)}$ and then collapse back to a final radius $x_{\rm f} \propto x_{\rm ta}$. In the scale-invariant radial collapse, the resulting profile will scale with $x_{\rm ta}$ (Fillmore & Goldreich 1984; Bertschinger 1985; Hoffman & Shaham 1985). Taking $x = \lambda x_{\rm ta}$ and using equation (4), it is easy to see that

$$\bar{\xi}_2(x) \simeq \left(\frac{\sigma_0}{\lambda^{1/2}}\right)^z \langle\nu^z\rangle x^{-3(n_{\rm p}+3)/(n_{\rm p}+4)} \tag{11}$$

where we have introduced the notation $z = 6/(n_{\rm p}+4)$.

Evaluating the average $\langle\ldots\rangle$ using the Gaussian distribution we find that the final result can be written as $\bar{\xi}_2(x) = A\bar{\xi}_{2,\rm lin}^3(l)$ where $l \approx x\bar{\xi}_2^{1/3}$ and

$$A = \frac{\langle\nu^z\rangle^{6/z}}{\lambda^3} = \frac{1}{\lambda^3}\left\{\frac{1}{\sqrt{2\pi}}\, 2^{(z-1)/2}\Gamma[(z+1)/2]\right\}^{6/z}. \tag{12}$$

The above result is the generalization (in the intermediate regime) of the analysis presented in Padmanabhan (1996), taking into account the averaging over different $\nu\sigma$ peaks. It shows that the averaging introduces a spectrum-dependent scaling.

Let us now consider the higher moments. Using our Ansatz for higher-order moments (equation 5) we can now compute the result for $\bar{\xi}_N$ to be

$$\bar{\xi}_N(x)=\left(\frac{\sigma_0}{\lambda^{1/2}}\right)^{z(N-1)}\langle \nu^{z(N-1)}\rangle x^{-3(N-1)(n_p+3)/(n_p+4)}. \tag{13}$$

The scaling we get for higher-order moments is clearly hierarchical in nature. Using the definition of S_N parameters we find that, in this intermediate regime,

$$S_N^{\rm int}=\bar{\xi}_N/\bar{\xi}_2^{(N-1)}=\langle \nu^{(N-1)z}\rangle/\langle \nu^z\rangle^{(N-1)} \tag{14}$$

or, equivalently,

$$S_N^{\rm int}=(4\pi)^{(N-2)/2}\frac{\Gamma\{[z(N-1)+1]/2\}}{\Gamma[(z+1)/2]^{(N-1)}}. \tag{15}$$

Using the above results, we can also directly relate the $\bar{\xi}_N(x)$ with $\bar{\xi}_2(l)$ and obtain

$$\bar{\xi}_N(x)=S_N\bar{\xi}_2^{(N-1)}(x)=S_N A^{N-1}\bar{\xi}_{2,\rm lin}^{3(N-1)}(l). \tag{16}$$

Non-linear regime

In this case, we take the initial density contrast to scale as the variance of the Gaussian random field, so that $\delta\propto x_i^{-(n_p+3)/2}$. We assume, as before, that a patch with initial radius x_i will attain a maximum radius $x_{\rm ta}=x/\delta$ which will collapse to form a structure of size $x=\lambda x_{\rm ta}$. Then, a corresponding calculation gives

$$\bar{\xi}_2(x)=\left(\frac{\sigma_0}{\lambda}\right)^{y}\langle \nu^{y}\rangle x^{-3(n_p+3)/(n_p+3)/(n_p+5)}, \tag{17}$$

where we have introduced the notation $y=6/(n_p+5)$. After averaging over the initial Gaussian distribution, this result becomes $\bar{\xi}_2(x)=B\bar{\xi}_{2,\rm lin}(l)^{3/2}$ where

$$B=\frac{\langle \nu^y\rangle^{3/y}}{\lambda^3}=\frac{1}{\lambda^3}\left\{\frac{1}{\sqrt{2\pi}}2^{(y-1)/2}\Gamma[(y+1)/2]\right\}^{6/y}. \tag{18}$$

This generalizes the corresponding result of Padmanabhan (1996) to the non-linear regime by taking into account the initial Gaussian fluctuations. The averaging introduces a spectrum-dependent pre-factor.

For higher-order moments, analysis can be done in an equivalent way, and the result is

$$\bar{\xi}_N(x)=\left(\frac{\sigma_0}{\lambda}\right)^{y(N-1)}\langle \nu^{y(N-1)}\rangle x^{-3(N-1)(n_p+4)/(n_p+5)}, \tag{19}$$

where $y=6/(n+5)$. So the scaling we get for higher-order moments is again hierarchical in nature and the S_N parameters can be evaluated in exactly the same manner as before, giving

$$S_N^{\rm non}=\bar{\xi}_N/\bar{\xi}_2^{(N-1)}=\langle \nu^{(N-1)y}\rangle/\langle \nu^y\rangle^{(N-1)} \tag{20}$$

or, equivalently

$$S_N^{\rm int}=(4\pi)^{(N-2)/2}\frac{\Gamma\{[z(N-1)+1]/2\}}{\Gamma[(z+1)/2]^{(N-1)}}. \tag{15}$$

This result can also be expressed as

$$\bar{\xi}_N(x)=S_N\bar{\xi}_2^{(N-1)}(x)=S_N B^{N-1}\bar{\xi}_{2,\rm lin}^{3(N-1)/2}(l). \tag{22}$$

The averaging process $\langle\ldots\rangle$ in both quasi-linear and non-linear regimes can be made more sophisticated by introducing an additional weight factor which is proportional to some power of Lagrangian volume of the patch from which the object is collapsing, i.e. x_i^m (see Padmanabhan et al. 1996). In that case the results generalize to

$$\bar{\xi}_N(x)=\langle x_i^{3(N-1)+m}\rangle/2\langle x_i^m\rangle x^{3(N-1)}, \tag{23}$$

which can be simplified to

$$S_N=(2\langle \nu^{m\beta}\rangle)^{N-2}\langle \nu^{3(N-1)\beta+m\beta}\rangle/\langle \nu^{3\beta+m\beta}\rangle^{(N-1)}, \tag{24}$$

where $\beta=x/3$ in the quasi-linear regime and $y/3$ in the non-linear regime. Finally, the S_N parameters are recovered after doing the averaging as before,

$$S_N=\left[2\Gamma\left(\frac{m\beta+1}{2}\right)\right]^{N-2}\frac{\Gamma\{[3\beta(N-1)+m\beta+1]/2\}}{\Gamma[(3\beta+m\beta+1)/2]^{(N-1)}}. \tag{25}$$

The simplest choice is $m=0$, which we shall use in this paper. It may be noted that the model used by Jain et al. (1995) corresponds to $m=3$; the expressions given above can be used to read off the S_N parameters in any other scheme. (As we shall see in the next section, $m=0$ seems to give a fairly good fit to the numerical simulation results.) We may also note the following.

(i) The expressions derived for S_N parameters are valid for $N\geq 2(S_1=1$ by definition).

(ii) Direct comparison with the results of the intermediate and non-linear regimes show $S_N^{\rm int}(n_p)=S_N^{\rm non}(n_p+1)$. Also note that the value of S_N is independent of λ.

(iii) Temporal dependence of ξ_N in both the quasi-linear and non-linear regimes can be derived from the fact that any statistic of the scale-invariant system can be expressed as a function of $x/x_{\rm nl}$, where $x_{\rm nl}$ is the scale defined through the relation $\sigma(x_{\rm nl})=1$. Since $x_{\rm nl}\propto a^{2/(n+3)}$, all correlations will be a function of the single variable $q=xa^{-(n+3)/2}$.

(iv) It is clear that, except for calculating the averages of powers of ν (which is assumed to be distributed normally), nowhere have we actually used the fact that the initial density distribution was Gaussian, which clearly shows that our method of analysis can be generalized in a straightforward manner to calculate S_N parameters for initially non-Gaussian distributions.

(v) For studying gravitational clustering in dimensions other than three, the same method of analysis can be used with the scaling $\bar{\xi}_N(x,a)\propto\bar{\xi}_{2,\rm lin}^{D(N-1)}(l,a)$ in the intermediate regime and $\bar{\xi}_N(x,a)\propto\bar{\xi}_{2,\rm lin}^{D(N-1)/2}(l,a)$ in a highly non-linear regime.

(vi) Given the S_N parameters, one can compute the void probability distribution function and related quantities. This calculation is indicated in the Appendix.

Transition between the regimes

Having determined the behaviour of correlation functions in the three different regimes, one can enquire where the transition between the regimes occurs. Since there exist three distinct phases in gravitational clustering, we have two transition points: (1) the transition from the perturbative to the intermediate regime and (2) the transition from the intermediate to the non-linear regime. Let the first transition occur when $\bar{\xi}_{2,\rm lin}(l)=T^{(2)}_{c_1}$ and the second when $\bar{\xi}_{2,\rm lin}(l)=T^{(2)}_{c_2}$. Finding the value of $\bar{\xi}_{2,\rm lin}(l)$ for which quasi-linear and intermediate $\xi_2(x)$ matches we get $T^{(2)}_{c_1}=1/A^{1/2}$. Similarly, equating the expressions for the intermediate and non-linear regimes gives $T^{(2)}_{c_2}=(b/A)^{2/3}$.

We have used the two-point correlation function to define the transitions since they are most directly related to the density inhomogeneity. It is, of course, possible to repeat the same exercise using higher-order correlation functions. Our results for the higher-order correlation functions can be summarized as

$$\begin{aligned}\bar{\xi}^{\rm pert}_N(x)&=S^{\rm tree}_N\bar{\xi}^{N-1}_{2,\rm lin}(l)\\ \bar{\xi}^{\rm int}_N(x)&=A^{N-1}S^{\rm int}_N\bar{\xi}^{3(N-1)}_{2,\rm lin}(l)\\ \bar{\xi}^{\rm non}_N(x)&=B^{N-1}S^{\rm non}_N\bar{\xi}^{3(N-1)/2}_{2,\rm lin}(l).\end{aligned}\tag{26}$$

Using these, it is easy to see that the transition points defined through the N-point correlation function will give us

$$T^{(N)}_{c_1}=(S^{\rm tree}_N/S^{\rm int}_N)^{1/2(N-1)}T^{(2)}_{c_1}\tag{27}$$

and

$$T^{(N)}_{c_2}=(S^{\rm non}_N/S^{\rm int}_N)^{2/3(N-1)}T^{(2)}_{c_2}.\tag{28}$$

Since $S^{\rm tree}_N<S^{\rm int}_N$ and $S^{\rm int}_N>S^{\rm non}_N$, it is clear that the transition for higher-order moments will occur for smaller and smaller values of $\xi_2(l)$.

It should be noted that, although S_N parameters are insensitive to the modelling parameters like λ, the transition points are sensitive to the choice of these variables. We have taken $\lambda=1/2$ which is close to the value taken by Jain et al. (1995) for their fitting function.

4 UNIFIED ANALYSIS FOR DIFFERENT REGIMES

In the last section, we discussed the intermediate and non-linear regimes separately. It is, however, possible to discuss the two regimes together by using a simple approximation. We shall discuss this approach in this section. The results of last section can be obtained as a special case of this approach. To do this we begin with equation (3) written as

$$\frac{\partial D}{\partial A}-h\frac{\partial D}{\partial X}=3h,\tag{29}$$

where we have introduced the following new variables

$$D=\ln\,[1+\bar{\xi}_2(x,a)],\qquad A=\ln a,\qquad X=\ln x,\tag{30}$$

and written the pair velocity as $v=-h\dot{a}x$. Simulations indicate that $h\approx 2$ in the intermediate regime and $h\approx 1$ in the non-linear regime, which we have used for results obtained in earlier sections. Here we will try to get a unified picture covering both regimes. We shall now assume that we can treat h as approximately constant while integrating this equation. In that case, the general solution is

$$1+\bar{\xi}_2=a^{3h}F(a^hx)\tag{31}$$

where F is an arbitrary function to be determined by the initial condition. If the linear $\bar{\xi}_L$ is a power law, we know that the true $\bar{\xi}_2(a,x)$ can only depend on the variable $q\equiv xa^{-2/(n+3)}$, which is possible only if F is a power law. So we *must* have

$$\xi_2(x,a)\propto a^{3h}(ax^h)^{-\gamma}.\tag{32}$$

The index γ can be determined by matching the above expression with the linear two-point correlation function at a scale $x_c=a^{2/(n+3)}$, which is going non-linear. We then get

$$\gamma=\frac{3h(n+3)}{h(n+3)+2}.\tag{33}$$

Now it is possible to write the two-point correlation function as

$$\bar{\xi}_2(x,a)\propto a^{6h/[1+h(n+3)]}x^{-3h(n+3)/[2+h(n+3)]}.\tag{34}$$

Earlier results of the intermediate and the highly non-linear regimes can be recovered by taking $h=2$ and $h=1$, respectively.

It is actually possible to relax the power-law requirement and still obtain a general result. From the characteristics of equation (29) we can show that $\bar{\xi}_2(x,a)$ can be expressed as a function of $\bar{\xi}_{\rm L}(l,a)$ where $l^3\equiv x^3[1+\bar{\xi}_2(x,a)]$. That is,

$$\bar{\xi}_2(x,a)=U[\bar{\xi}_{\rm L}(l,a)]\tag{35}$$

where U is some function. Combining with (31) we have $a^{3h}F(a^hx)=U[\bar{\xi}_{\rm L}(l,a)]$ or, equivalently, $a^{3h}F(r)=U[a^2Q(r)]$ where we have used the fact that we can write $l^3\cong x^3\bar{\xi}_2=r^3F(r)$ with $r=a^hx$. However, the above expression can be valid for arbitrary a at fixed r only if

$$U(z)\propto z^{3h/2}.\tag{36}$$

So in terms of correlations functions we must have

$$\bar{\xi}_2(x,a)\propto[\bar{\xi}_{\rm L}(l,a)]^{3h/2}.\tag{37}$$

This generalizes the relations $\bar{\xi}_2\propto\bar{\xi}^3_{\rm L}$, $\bar{\xi}_2\propto\bar{\xi}^{3/2}_{\rm L}$ we used for the quasi-linear and non-linear regimes in the last section.

To perform the averages over regions with different peak heights, we only have to do the rescaling $\bar{\xi}_{\rm L}\to\nu^2\bar{\xi}_{\rm L}$ and note that $\bar{\xi}_2\propto(x_i/x)^3\propto\bar{\xi}^{3h/2}_{\rm L}(x_i)$ implies $x\propto x_i\bar{\xi}^{h/2}_{\rm L}(x_i)$. Simple algebra then gives

$$\xi_2(x,a)\propto\frac{\langle x_i^3\rangle}{x^3}\propto\langle\nu^{6h/[2+h(n+3)]}\rangle x^{-3h(n+3)/[1+h(n+3)]}.\tag{38}$$

This generalizes relations (11) and (17) of the last section. Assuming that ν is a Gaussian variable and performing the average, we get up to a normalization

$$\xi_2(x,a)\propto 2^{\alpha/2}\Gamma\left(\frac{\alpha+1}{2}\right)x^{-3h(n+3)/[2+h(n+3)]}\tag{39}$$

where $\alpha=6h/[2+h(n+3)]$. Normalizing the expression properly, we can write the final result as

$$\xi_2(a, x) = \mathscr{A}(h, n)\,[\xi_2(l, a)]^{3h/2}, \tag{40}$$

where

$$\mathscr{A}(h, n) = \left(\frac{2}{\lambda}\right)^{3h/2} \left[\frac{\Gamma[(\alpha+1)/2]}{2\sqrt{\pi}}\right]^{3h/\alpha}. \tag{41}$$

To obtain the earlier results for the intermediate regime we have to set $h=2$ which gives $\alpha=6/(n+4)$ and

$$A = \left(\frac{2}{\lambda}\right)^{3} \left[\frac{\Gamma[(\alpha+1)/2]}{2\sqrt{\pi}}\right]^{6/\alpha}. \tag{42}$$

similarly, for the non-linear regime we have $h=1$, $\alpha=6/(n+5)$ and

$$\mathrm{B} = \left(\frac{2}{\lambda}\right)^{3/2} \left\{\frac{\Gamma[(\alpha+1)/2]}{2\sqrt{\pi}}\right\}^{3/\alpha}. \tag{43}$$

These results match with earlier expressions.

Using our Ansatz in (5) it is possible to generalize the result for higher-order correlation functions. We find

$$S_N(h, n) = \langle v^\alpha (N-1)\rangle / \langle v^\alpha \rangle^{N-1}, \tag{44}$$

which can be explicitly written as

$$S_N(h, n) = (4\pi)^{(N-2)/2}\,\Gamma\left(\frac{\alpha(N-1)+1}{2}\right) \Big/ \Gamma\left(\frac{\alpha+1}{2}\right). \tag{45}$$

This allows the calculation of $S_N[\xi_2(x, a)]$, given $h[\xi_2(x, a)]$.

5 COMPARISON WITH SIMULATIONS

The results obtained in the previous section can be compared with the simulations as regards two essential aspects.

First, the results show that the spectral dependence of the scaling relations between the non-linear and linear correlation functions can arise due to averaging peaks of different heights. This, in turn, implies that the scales at which the transition from the perturbative to the intermediate regime, or from the intermediate to the non-linear regime, takes place, depend on the power spectrum index. By comparing the predicted values for these transitions with the results of simulations, we can test the validity of our averaging scheme and the basic model for the two-point correlation function. Secondly, we can compare the values of $\bar{\xi}_N$ obtained from the model with those of simulation in both intermediate and non-linear regimes. Since $\bar{\xi}_2$ is related to $\bar{\xi}_{2L}$ in a non-local manner, and our Ansatz relates $\bar{\xi}_N$ to $\bar{\xi}_2$ (in an indirect manner), we would expect some non-local relationship between $\bar{\xi}_N$ and $\bar{\xi}_2$. This will test the validity of our Ansatz regarding higher-order correlation functions. Note that these two comparisons test the two distinct generalizations of the work in Padmanabhan (1996), introduced in this paper.

The comparison of predicted values for the transition is shown in Fig 1 along with the results of the numerical simulation given in Jain et al. (1995). We see that there is good agreement between theory and simulations, suggesting (i) that the basic picture for the evolution of gravitational clustering, developed in Padmanabhan (1996), is correct, and (ii) that the spectrum dependence of the scaling relation can be understood by averaging over the initial fluctuations. The analysis also shows that, as n_p changes from -2 to 0, the lower transition point varies between 0.25 and 1.0 while the upper one varies between 2.0 and 7.0.

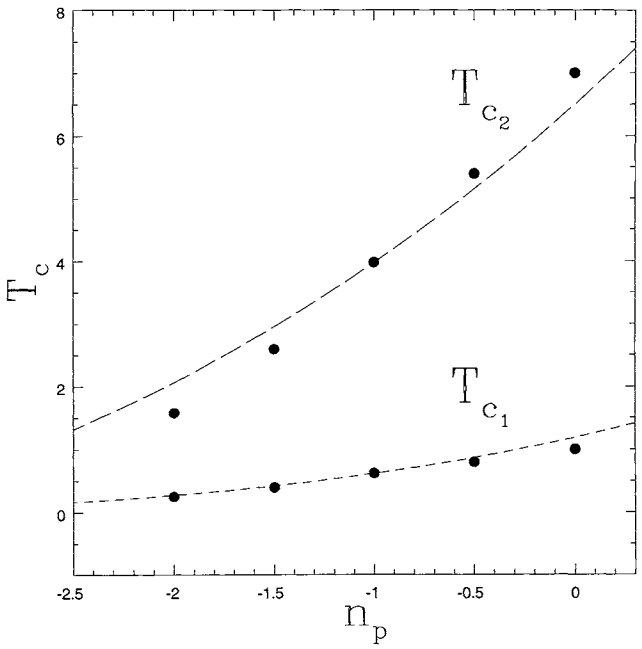

Figure 1. Transition points T_{c_1} and T_{c_2} are plotted as a function of spectral index n_p. The short-dashed curve corresponds to transition points from the perturbative regime to the intermediate regime, while the long-dashed curve corresponds to the transition from the intermediate to the non-linear regime, predicted from our model. The curves are normalized to match Jain et al.'s simulation result for $n=-1$. Circles correspond to values obtained in the simulation by Jain et al. (1995). Transition points were obtained by finding the intersection of straight lines which we fit to represent various regimes.

We shall next turn to the comparison of $\bar{\xi}_N$ predicted by our model with the results of numerical simulations, in order to test the validity of our Ansatz. In doing so, one should be aware of several effects which could 'corrupt' the values of the S_N parameters in the simulations, and take adequate precautions to correct for them. At small σ, i.e. in the perturbative regime, the main contribution to error comes from cosmic variance, i.e. as a result of the presence of a small number of large cells containing completely independent samples. This error can be reduced only by increasing the size of the N-body computation box. On the other hand, in the highly non-linear regime one is restricted by the resolution of the N-body simulation for probing a very small scale. Poisson shot noise also starts playing an increasingly dominant role as soon as the average occupancy of cells becomes comparable to unity. The usual procedure used for computing S_N parameters is by taking moments of cell counts $P(n)$ for different cell sizes. In general, cell counts show a power-law behaviour in the highly non-linear regime up to $n=n_c(=\bar{n}\bar{\xi}_2)$ followed by an exponential tail at $n>n_c$ (Balian & Schaeffer 1989). In an ideal (infinite) catalogue this exponential tail would be extended to very small values

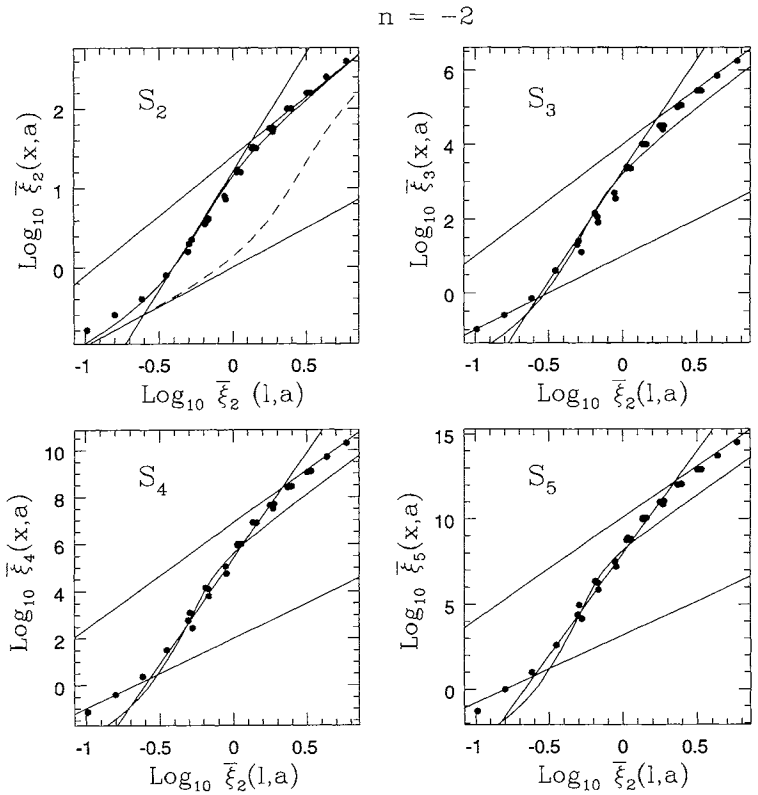

Figure 2. $\bar{\xi}_N(x, a)$ has been plotted against $\bar{\xi}_2(l, a)$ for $n = -2$ spectra. The solid curve in panel S_2 corresponds to the fit by Jain et al. (1995) and the dashed curve for Hamilton et al. (1991); in other panels solid curves correspond to prediction from our model. Straight lines in different panels correspond to slopes $(N-1)$, $3(N-1)$ and $3(N-1)/2$ for the perturbative, intermediate and highly non-linear regimes respectively. Dots represent N-body simulation data from Colombi, Bouchet & Hernquist (1996).

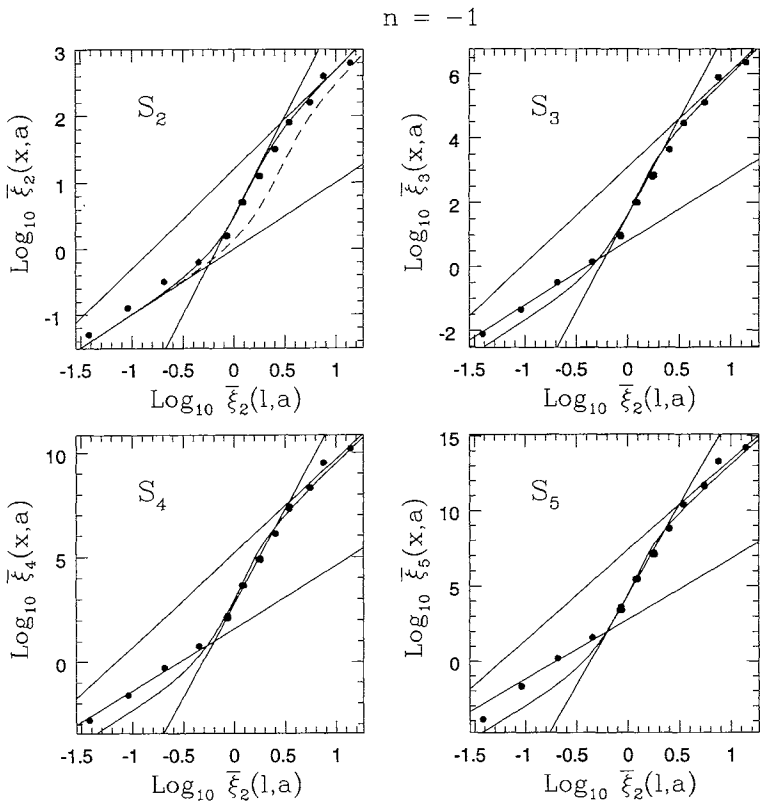

Figure 3. As Fig. 2 for $n = -1$ spectrum.

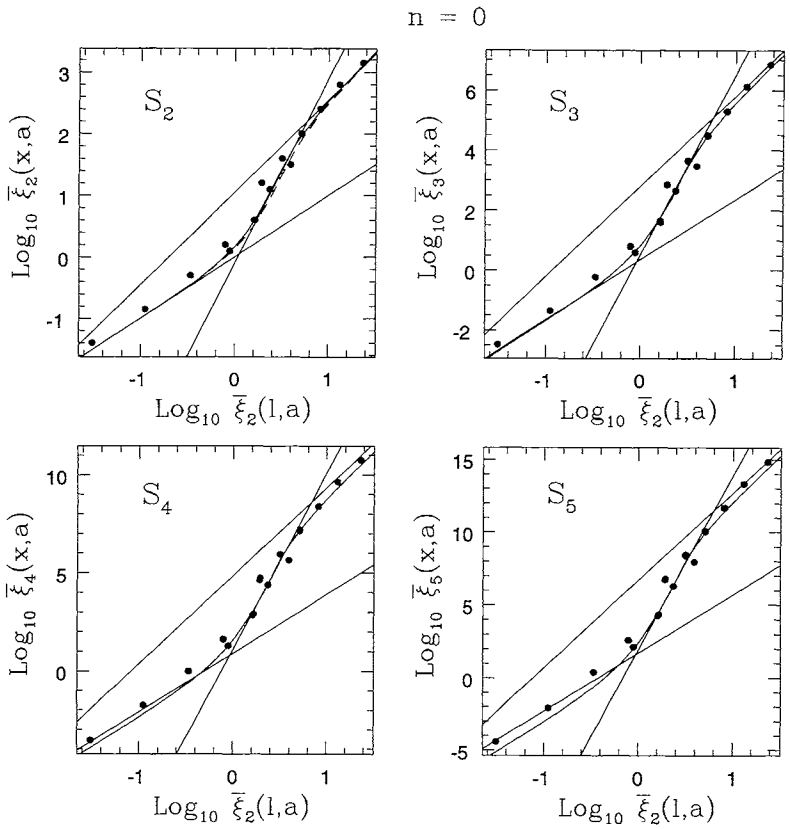

Figure 4. As Fig. 2 for $n=0$ spectrum.

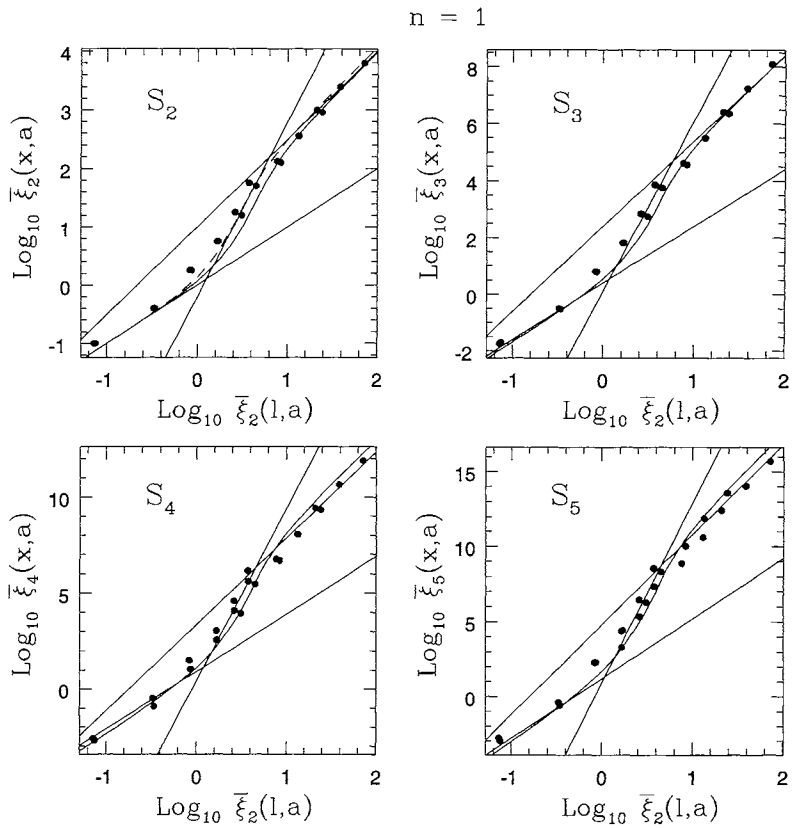

Figure 5. As Fig. 2 for $n=1$ spectrum.

of $P(n)$; but in practice there is a sharp cut-off around $n=n_{\max}$ which is the most dense cell present in the N-body catalogue. Higher moments of $P(n)$ – and hence higher S_N parameters – are more sensitive to the large n tail of $P(n)$. So it is extremely important to extend the exponential tail to infinity and then normalize the corrected $P(n)$ again before calculating the S_N parameters. This way of correcting for measured S_N has been extensively studied and is expected to give a correct – or at least more reliable – result (Colombi, Bouchet & Schaeffer 1992, 1994, 1995; Lucchin et al. 1994). Recently Colombi et al. (1995) have done a careful analysis of high-resolution N-body data with a large dynamic range, correcting for all the errors mentioned above. Their study covers power-law models $n_p=1, 0, -1, -2$, and they study first three non-trivial S_N parameters, i.e. S_3, S_4, and S_5.

We have computed $\bar{\xi}_N$ for $N=2, 3, 4, 5$ for the power laws $n=-2, -1, 0, 1$, using the published data in Colombi et al. (1995) and our Ansatz. The results are shown in Figs 2–5. The theoretical prediction based on our Ansatz is shown by the solid, S-shaped line, and the simulation data is shown by points. The straight lines show the asymptotic theoretical slopes.

It is clear from the graphs that our basic Ansatz is an attempt in the right direction. The overall agreement between the theory and simulations is good, especially when we consider the simplicity of our model. We will now comment on several details related to the comparison between theory and simulations.

Plots of $\bar{\xi}_N(x, a)$ versus $\bar{\xi}_2(x, a)$ show that our basic claim regarding three phases in gravitational clustering seems to be correct, and one does get a non-local scaling relation for the N-point correlation function similar to scaling for the two-point correlation function as suggested from the characteristics. To some extent, the scaling in higher-order correlation functions reflects underlying scaling in the two-point correlation function; but it can also be argued that if $\bar{\xi}_N(x, a)$ can be expressed as a smoothly varying function of $\bar{\xi}_2(x, a)$, $\bar{\xi}_N(x, a)=T_N[\bar{\xi}_2(x, a)]$ then we can write $\bar{\xi}_N(x, a)=T_N\{F_n[\bar{\xi}_2(l, a)]\}=G_{n,N}[\bar{\xi}_2(l, a)]$ where $G_{n,N}=T_N\otimes F_n$, where $\otimes$ denotes convolution, and thus define a scaling relation between $\bar{\xi}_N(x, a)$ and $\bar{\xi}_2(l, a)$.

The relation between $\bar{\xi}_2(a, x)$ and $\bar{\xi}_2(l, a)$ shows good agreement with the analytical fit suggested by Jain et al. (1995). The scatter in the data increases with n, which can be understood in following manner. We used x and $\bar{\xi}_2(x, a)$ to recover the Lagrangian radius $l=x[1+\bar{\xi}_2(x, a)]^{1/3}$, which was then used to get the linearly extrapolated $\bar{\xi}_2(l, a)=\sigma_0^2 a^2 l^{-(n+3)}$. The error in estimation of x or $\bar{\xi}_2(x, a)$ from the published data of Colombi et al. (1995) gets reflected finally in the error of $\bar{\xi}_2(l, a)$. We can relate fractional error $\Delta_{\bar{\xi}_2(l,a)}$ with fractional error Δ_l by $|\Delta_{\bar{\xi}_2(l,a)}|=(n+3)|\Delta_l|$, which shows that for the same Δ_l, $\Delta_{\bar{\xi}_2(l,a)}$ increases with n. This explains (partly) why we get more scatter for $n=+1$ spectra than for $n=-2$ spectra.

ACKNOWLEDGMENTS

It is a pleasure for DM to acknowledge his thesis supervisor, Varun Sahni, for constant encouragement and active support during the course of work. DM thanks Francis Bernardeau and Richard Schaeffer for many useful discussions and warm hospitality during his stay at CEA (Saclay). DM was financially supported by the Council of Scientific and Industrial Research, India, under its SRF scheme.

REFERENCES

Balian R., Schaeffer R., 1989, A&A, 220, 1
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Bernardeau F., 1992, ApJ, 392, 1
Bernardeau F., 1994a, ApJ, 427, 51
Bernardeau F., 1994b, ApJ, 433, 1
Bernardeau F., 1994c, A&A, 291, 697
Bernardeau F., 1995, A&A, 301, 309
Bernardeau F., Schaeffer R., 1992, A&A, 255, 1
Bertschinger E., 1985, ApJS, 58, 39
Bouchet F. R., Hernquist L., 1992, ApJ, 400, 25
Buchert T., 1992, MNRAS, 254, 729
Colombi S., Bouchet F. R., Schaeffer R., 1992, A&A, 263, 1
Colombi S., Bouchet F. R., Schaeffer R., 1994, A&A, 281, 301
Colombi S., Bouchet F. R., Schaeffer R., 1995, ApJS, 96, 401
Colombi S., Bouchet F. R., Hernquist L., 1996, ApJ, 465, 14
Fillmore J., Goldreich P., 1984, ApJ, 281, 1
Fry J., 1984, ApJ, 279, 499
Hamilton A. J. S., 1988, ApJ, 332, 67
Hamilton A. J. S., Kumar P., Lu E., Matthews A., 1991, ApJ, 374, L1
Hoffman Y., Shaham J., 1985, ApJ, 297, 16
Jain B., Mo H., White S. D. M., 1995, MNRAS, 276, L25
Lucchin F., Matarrese S., Melott A. L., Moscardini L., 1994, ApJ, 422, 430
Munshi D., Sahni V., Starobinsky A. A., 1994, ApJ, 436, 517
Moutarde F., Alimi J. M., Bouchet F. R., Pellat R., Ramani A., 1991, ApJ, 382, 377
Nityananda R., Padmanabhan T., 1994, MNRAS, 271, 976
Peebles P. J. E., 1980, The large scale structure of the universe. Princeton Univ. Press, Princeton, NJ
Peacock J. A., Dodds S. J., 1996, MNRAS, 280, L19
Padmanabhan T., 1996, MNRAS, 278, L29
Padmanabhan T., Cen R., Ostriker J. P., Summers F. J., 1996, ApJ, 466, 604
Schaeffer R., 1984, A&A, 134, L15
White S. D. M., 1979, MNRAS, 186, 145

APPENDIX A

Given the S_N parameters, one can compute the void probability function which is of some theoretical and practical importance. In scaling models (which assume S_N parameters are constant over some length-scale), the void probability function can be written as (White 1979)

$$P_0=\exp[-\phi(n_c)/\bar{\xi}_2] \tag{46}$$

where ϕ is a generating function for S_N parameters and is defined as

$$\phi(n_c)=-\sum_{p=1}^{\infty}\frac{S_p}{p!}(-n_c)^p \tag{47}$$

where $n_c=\bar{n}\bar{\xi}_2$ is a scaling variable. Substituting our expression for S_N in the above equation we get

$$\phi(n_c)=n_c-\sum_{p=2}^{\infty}(4\pi)^{(p-2)/2}\frac{\Gamma\{[(p-1)s+1]/2\}}{\Gamma[(s+1)/2]^{p-1}p!}(-n_c)^p, \tag{48}$$

where $s=z$ for the intermediate regime and $s=y$ for the

non-linear regime. Using the definition of the gamma function it is easy to write down the sum as

$$\phi(n_c)=n_c-\frac{\Gamma[(s+1)/2]}{4\pi}\sum_{p=2}^{\infty}\int_0^{\infty}\frac{\mathrm{d}t t^{(p-1)s/2}\exp(-t)}{\sqrt{t}p!}\times\left\{-\frac{\sqrt{4\pi}n_c}{\Gamma[(s+1)/2]}\right\}^p. \quad (49)$$

Interchanging the sum and the integral we get

$$\phi(n_c)=n_c-\frac{\Gamma[(s+1)/2]}{4\pi}\int_0^{\infty}\mathrm{d}t t^{-(s+1)/2}e^{-t}\times[\exp(-ft^{s/2})+ft^{s/2}-1] \quad (50)$$

where

$$f=\left\{\frac{\sqrt{4\pi}n_c}{\Gamma[(s+1)/2]}\right\}, \quad (51)$$

and taking the limit $n_c\to\infty$ shows that $\phi(n_c)\approx n_c/2$ for large n_c. In the scaling model proposed by Balian & Schaeffer, ϕ is expected to scale as $n_c^{1-\omega}$ asymptotically; hence our Ansatz leads to the $\omega=0$ model. (Note that other extreme case is $\omega=1$ which is the negative binomial model proposed earlier.) The counts-in-cell $P(n)$ is proportional to $n^{\omega-2}$ showing that, in our model, $P(n)\propto n^{-2}$.