DETECTING SCALE-DEPENDENCE OF BIAS FROM APM-BGC GALAXIES

LI-ZHI FANG,1 ZU-GAN DENG,2,3 AND XIAO-YANG XIA3,4

Received 1997 September 8; accepted 1998 May 18

ABSTRACT

We present an investigation of the scale-dependence of bias described by the linear model:

\[ \frac{\delta \rho (x)}{\rho} = \frac{b \delta \rho (x)}{\rho}, \]

where \( \delta \rho (x) \) and \( \rho (x) \) are the galaxy number density and mass density, respectively. Using a discrete wavelet decomposition, we show that the behavior of bias scale-dependence cannot be described by one parameter \( b \). In the linear bias model, the scale-dependence should be measured by the \( j \)-spectra of wavelet-coefficient--represented bias parameters \( b^j \), \( n \) being positive integers. Because \( b^j \) with different \( n \) are independent from each other, a systematic analysis of the \( j \)-spectra of \( b^j \) is necessary. We performed a \( j \)-spectrum analysis for samples of elliptical and lenticular (EL) and spiral (SP) galaxies listed in the APM Bright Galaxy Catalogue. We found that, for statistics of two-point correlation functions or discrete wavelet transform power spectrum, the scale-independence holds within 1 \( \sigma \). However, the bias scale-dependence becomes substantial when phase-sensitive statistics (e.g., \( b^j \) with \( n > 2 \) or \( b^j \)) are applied. These results indicate that the bias scale-dependence has the same origin as the non-Gaussianity of galaxy distributions. This is generally consistent with the explanation that the bias scale-dependence originated from nonlinear and nonlocal relationship between galaxy formation and their environment.

Subject headings: cosmology: theory — galaxies: statistics — large-scale structure of universe

1. INTRODUCTION

Bias is introduced to reconcile the amplitude of fluctuations inferred from clustering of galaxies with that derived from mass distributions. Bright galaxies seem to have a stronger clustering than that of the underlying mass. Therefore, it is generally believed that galaxies are biased tracers of the mass density field; i.e., it follows the clustering of (dark) matter, but with an enhanced amplitude.

Generally, bias is phenomenologically modeled by a linear relation as

\[ \delta(x)_g = b \delta(x)_m, \]

where \( \delta(x)_g \) and \( \delta(x)_m \) are the galaxy and mass number density perturbations, respectively, and \( \rho \) and \( \bar{\rho} \) are the average of \( \rho(x) \) and \( \bar{n}(x) \). The phenomenological bias parameter \( b \) is assumed to be constant but may be different for different galaxy types, say, \( b_{\text{early}}, b_{\text{late}} \) for early and late types of galaxies.

Equation (1) and its variants are widely used in the determination of cosmological parameters from samples of redshift surveys of galaxies. However, we have really very little idea about which the bias parameter \( b \) should be. In fact, both bias \( (b > 1) \) and antibias \( (b < 1) \) are employed in current data analysis (e.g., Mo, Jing, & White 1996). This prevents unambiguous measures of cosmological parameters, giving only bias-contaminated results.

Theoretically, the physical mechanism responsible for equation (1) is far from clear. The first analytic model of bias, in which objects are identified with high peaks or collapsed halos of the density field, offers a plausible explanation of the bias of galaxy clusters: the correlation amplitudes of clusters are strong functions of cluster richness (Kaiser 1984). In this case, bias is mainly caused by the mass of collapsed halos. The larger the mass of the halo, the higher the richness of the cluster.

Yet, the formation of galaxies does not depend only on the mass or local mass density of collapsed halos, but is substantially modulated by various environmental effects, such as the suppression of star formation in neighboring protogalaxies (Rees 1985), the stimulation of the formation of nearby galaxies (Dekel & Rees 1987), dynamical friction effects (Couchman & Carleberg 1992), and so on. All these environmental effects are beyond local density and lead to a nonlocal relation between the number density of galaxies and the background mass field (Bower et al. 1993). Moreover, the rate of star formation is most likely nonlinearly dependent on local mass density. A common result of the nonlocal and/or nonlinear relation between \( \delta(n) \) and \( \delta_m \) is the scale-dependence of parameter \( b \). Therefore, to have a deep understanding of the mechanism of galaxy bias, searching for the scale-dependence of parameter \( b \) is necessary (Coles 1993; Catelan et al. 1994).

So far, the results of detecting \( b \) scale-dependence are quite scattered. For instance, the values of the bias-contaminated density parameter \( \beta = \Omega^{0.6}/b \) are found to be in the range of 0.4–1, equal to about 0.5–0.6 at Gaussian-smoothing scales of 3–6 h\(^{-1}\) Mpc, and \( \sim 1 \) on scales of 12 h\(^{-1}\) Mpc (e.g., Dekel, Burstein, & White 1997). That is, \( b \) is probably scale-dependent from 6 to 12 h\(^{-1}\) Mpc. On the other hand, some studies conclude that, for galaxies of all types and luminosities, the scale-dependence of the bias parameter is weak (e.g., Kauffmann, Nusser, & Steinmetz 1997). Why have different detections given different conclusions? This question motivated us to study the physics involved in equation (1).

In the first part of this paper, we show that even in a linear bias model, the behavior of scale-dependence cannot be described by one parameter \( b \) but by a series of \( j \)-spectra of wavelet-coefficient–represented bias parameters \( b^j \) and...
$b_j^{(n)}$, $n$ being positive integers. The $j$-spectra of $b_j^{(n)}$ are all statistically independent. Different detections may actually measure different parameters $b_j^{(n)}$. Therefore, it should not be surprising that some detections are positive and some are negative. The different behavior of bias scale-dependence given by different detection does not cause confusion, but may greatly be helpful to reveal the physics behind the bias model (1). Therefore, to have a complete picture of bias scale-dependence, a systematic analysis of the $j$-spectra of $b_j^{(0)}$ and $b_j^{(1)}$ is necessary.

In the second part, we performed a systematic detection of the bias scale-dependence with the samples of galaxies listed in the APM Bright Galaxies Catalogue (APM-BGC). This analysis shows that for the APM-BGC sample, the scale-independence approximately holds if only the two-point correlation function and power spectrum are involved, while the scale-dependence becomes substantial when higher order or phase-sensitive statistics involved. This result is valuable to constrain models of hierarchical clustering of galaxies.

The paper is organized as follows: In § 2 we briefly introduce a method of the space-scale decomposition based on discrete wavelet transform (DWT) analysis. With this method, various statistics of measuring the bias scale-dependence are developed and presented; § 3 describes the sample to be analyzed. In § 4 we apply the DWT method to analyze the galaxy sample of the APM bright galaxies. The implications of these results are discussed in § 5.

2. Linear Bias Model in DWT Representation

2.1. The Discrete Wavelet Transform

Let us briefly introduce the DWT analysis of large-scale structures for the details referred to (Fang & Pando 1997; Pando & Fang 1996, 1998). We consider here a one-dimensional mass density distribution $\rho(x)$ or contrast $\delta(x) = [\rho(x) - \rho]/\rho$, which are mathematically random fields over a spatial range $0 \leq x \leq L$. It is not difficult to extend all results developed in this section into two-dimensions and three-dimensions, because the DWT bases for higher dimensions can be constructed by a direct product of one-dimensional bases.

Like the Fourier expansion of the field $\delta(x)$, the DWT expansion of the field $\delta(x)$ is given by

$$\delta(x) = \sum_{j_0}^{\infty} \sum_{l=0}^{2^j-1} \hat{\delta}_{j,l} \psi_{j,l}(x),$$

where $\psi_{j,l}(x)$, $j = 0, 1, \ldots$, and $l = 0 \ldots 2^j - 1$ are the bases of the DWT. Because these bases are orthogonal and complete, the wavelet function coefficient (WFC), $\hat{\delta}_{j,l}$, is computed by

$$\hat{\delta}_{j,l} = \int \delta(x) \psi_{j,l}(x) dx.$$  \hfill (3)

The wavelet transform bases $\psi_{j,l}(x)$ are generated from the basic wavelet $\psi(x/L)$ by a dilation $2^j$ and a translation $l$; i.e.,

$$\psi_{j,l}(x) = \left(\frac{2^j}{L}\right)^{1/2} \psi\left(\frac{2^j x}{L - l}\right).$$  \hfill (4)

The basic wavelet $\psi$ is designed to be continuous, admissible, and localized. Unlike the Fourier bases $\exp(2\pi i nx/L)$, which are nonlocal in physical space, the wavelet bases $\psi_{j,l}(x)$ are localized in both physical space and Fourier (scale) space. In physical space, $\psi_{j,l}(x)$ is centered at position $L/2^j$, and in Fourier space, it is centered at wavenumber $2\pi \times 2^j/L$. Therefore, the DWT decomposes the density fluctuating field $\delta(x)$ into domains $j$, $l$ in phase space, and for each basis the corresponding area in the phase space is as small as that allowed by the uncertainty principle. WFC $\hat{\delta}_{j,l}$ and its intensity $|\hat{\delta}_{j,l}|^2$ describe, respectively, the fluctuation of density and its power on scale $L/2^j$ at position $L/2^j$.

In order to have reasonable statistics, the cosmic density field is usually assumed to be ergodic: the average over an ensemble is equal to the spatial average taken over one realization. This is the so-called “fair sample hypothesis” (Peebles 1980). A homogeneous Gaussian field with continuous spectrum is certainly ergodic (Adler 1981). In some non-Gaussian cases, such as homogeneous and isotropic turbulence (Vanmarcke 1983), ergodicity also approximately holds. Roughly, the ergodic hypothesis is reasonable if spatial correlations are decreasing sufficiently rapidly with increasing separation. The volumes separated with distances larger than the correlation length are approximately statistically independent. In this case, WFCs, $\hat{\delta}_{j,l}$, at different $l$ from one realization of $\delta(x)$ can be employed as statistically independent measurement. Thus, the WFCs of $\delta(x)$ at different $l$ on a given $j$, $\hat{\delta}_{j,l}$, form an ensemble of the WFCs on the scale $j$. In other words, when the “fair sample hypothesis” holds, an average over ensemble can be fairly estimated by the average over $l$; i.e., $\langle \hat{\delta}_{j,l}\rangle \simeq (1/2^j) \sum_{l=0}^{2^j-1} \hat{\delta}_{j,l}$, where $\langle \ldots \rangle$ denotes ensemble average.

2.2. Reconstruction of Density Field

Using the completeness of the DWT basis, one can reconstruct the original density field from the coefficient $\hat{\delta}_{j,l}$. To achieve this, DWT analysis employs another set of functions consisting of the so-called scaling functions, $\phi_{j,l}$, which are generated from the basic scaling $\phi(x/L)$ by a dilation $2^j$ and a translation $l$; i.e.,

$$\phi_{j,l}(x) = \left(\frac{2^j}{L}\right)^{1/2} \phi\left(\frac{2^j x}{L - l}\right).$$  \hfill (5)

The basic scaling $\phi$ is essentially a window function with width $x/L = 1$. Thus, the scaling functions $\phi_{j,l}(x)$ are also windows, but with width $(1/2^j)L$ and centered at $L/2^j$. The scaling functions $\phi_{j,l}(x)$ are orthogonal with respect to the index $l$, but not for $j$. This is a common property of window functions, which can be orthogonal in physical space but not in Fourier space.

For Daubechies wavelets, the basic wavelet and the basic scaling are related by recursive equations as

$$\phi(x/L) = \sum_{l} a_l \phi(2x/L - l), \quad \psi(x/L) = \sum_{l} b_l \phi(2x/L + l).$$  \hfill (6)

(Daubechies 1992), where coefficients $a_l$ and $b_l$ are different for different wavelets. In this paper, we use the Daubechies 4 wavelet (D4), for which $a_0 = [1 + (3)^{1/2}]/4$, $a_1 = [3 + (3)^{1/2}]/4$, $a_2 = [3 + (3)^{1/2}]/4$, $a_3 = [3 - (3)^{1/2}]/4$, $a_4 = [1 - (3)^{1/2}]/4$.

From equation (6), one can show that the scaling functions $\phi_{j,l}(x)$ are always orthogonal to the wavelet bases $\psi_{j',l}(x)$ if $j \leq j'$; i.e.,

$$\int \phi_{j,l}(x) \psi_{j',l}(x) dx = 0, \quad \text{for } j \leq j'.$$  \hfill (7)
Therefore, \( \phi_{j,i}(x) \) can be expressed by \( \psi_{j,l}(x) \) as

\[
\phi_{j,i}(x) = \sum_{j=0}^{\infty} \sum_{l=0}^{2^j-1} c_{j,l} \psi_{j,l}(x) = \sum_{j=0}^{j-1} \sum_{l=0}^{2^j-1} c_{j,l} \psi_{j,l}(x) .
\]

(8)

The coefficients \( c_{j,l} = \int \phi_{j,i}(x) \psi_{j,l}(x) \, dx \) can be determined from \( a_i \) and \( b_i \).

Using \( \phi_{j,i}(x) \), we construct a density field on scale \( j \) as

\[
\rho^j(x) = \sum_{j=0}^{2^j-1} \sum_{l=0}^{2^j-1} w_{j,l} \phi_{j,i}(x) ,
\]

where \( w_{j,l} \) is called the scaling function coefficient (SFC), given by

\[
w_{j,l} = \int_{0}^{L} \rho(x) \phi_{j,i}(x) dx .
\]

(10)

Since the scaling function \( \phi_{j,i}(x) \) is window-like, the coefficient \( w_{j,l} \) is actually a “count-in-cell” in a window on scale \( j \) at position \( l \).

Using equations (2), (8), (9), and (10), one can find

\[
\rho^j(x) = \tilde{\rho} \sum_{j=0}^{j-1} \sum_{l=0}^{2^j-1} \tilde{w}_{j,l} \psi_{j,l}(x) + \tilde{\rho} .
\]

(11)

Namely, \( \rho^j(x) \) contains all terms of density fluctuations \( \tilde{w}_{j,l} \psi_{j,l}(x) \) of \( j' < j \), but not terms of \( j' \geq j \). From equations (2) and (11), we have

\[
\rho(x) = \rho^j(x) + \tilde{\rho} \sum_{j=0}^{\infty} \sum_{l=0}^{2^j-1} \tilde{w}_{j,l} \psi_{j,l}(x) .
\]

(12)

One can also define the smoothed density contrasts on scale \( j \) to be

\[
\delta^j(x) = \frac{\rho^j(x) - \tilde{\rho}}{\tilde{\rho}} = \sum_{j=0}^{j-1} \sum_{l=0}^{2^j-1} \tilde{w}_{j,l} \psi_{j,l}(x) .
\]

(13)

Equations (12) and (13) show that \( \rho^j(x) \) and \( \delta^j(x) \) are the density field smoothed on scale \( j \). Using equations (12) and (13), one can construct the density field \( \rho^j(x) \) or \( \delta^j(x) \) on finer and finer scales by WFCs \( \tilde{w}_{j,l} \) until the precision of the original field is reached. Since the sets of bases \( \psi_{j,l} \) and \( \phi_{j,l} \) are complete, the original field can be reconstructed without lost information.

2.3. Statistical Description of Scale-Dependence of Bias

We subject the DWT decomposition to equation (1), which gives

\[
(\tilde{w}_{j,l})_m = b(\tilde{w}_{j,l})_m .
\]

(14)

Statistically, equation (14) means that the one-point distributions of the WFCs for galaxies should be proportional to the matter, and the proportional coefficient \( b \) is \( j \)-independent. Thus, linear bias models (eq. [1]) require that various statistics based on the WFC one-point distributions for galaxies and matter should satisfy the linear relations given by equation (14).

If we define the cumulant moment of the one-point distribution of \( \tilde{w}_{j,l} \) as

\[
C^n_{j,l} = \langle \tilde{w}_{j,l}^n \rangle = \frac{1}{2^j} \sum_{i=0}^{2^j-1} \tilde{w}_{j,l}^n ,
\]

then model (1) is actually to require that the ratios \( (C^n_{j,l})/(C^n_{j,m}) = \langle (\tilde{w}_{j,l})^n \rangle /\langle (\tilde{w}_{j,m})^n \rangle \) should be \( j \)-scale-independent for all order \( n \). Namely,

\[
\tilde{b}^j = (C^n_{j,l})/(C^n_{j,m})
\]

should be \( j \)-independent. In other words, the \( j \)-spectra of \( \tilde{b}^j \) are flat. For non-Gaussian fields, the parameters \( \tilde{b}^j \) for different \( n \) generally are statistically independent. Therefore, the behavior of bias scale-dependence cannot be described by one parameter \( b \), but by the \( j \)-spectra of \( \tilde{b}^j \) for all \( n \).

From equations (13) and (14), one has

\[
\delta^j(x) = b \delta^j(x)_m .
\]

(17)

Equation (17) can be rewritten as

\[
\rho^j(x) - \tilde{\rho} = b^j (\rho^j(x)_m - \tilde{\rho}_m) ,
\]

(18)

where constant \( b^j = b m / \tilde{\rho}_m \). Considering that \( \int \phi_{j,i}(x) dx \) is independent of index \( l \), the DWT decomposition of equation (18) gives

\[
(\tilde{w}_{j,l} - \tilde{w}_j)_m = b^j (w_{j,l} - \tilde{w}_j)_m ,
\]

(19)

where \( \tilde{w}_j = (1/2^j) \sum_{l=0}^{2^j-1} w_{j,l} \) is the mean of the SFCs on scale \( j \). Similar to equation (16), one can design statistics \( b^j \) as follows:

\[
b^j = \sum_{l=0}^{2^j-1} (\tilde{w}_{j,l} - \tilde{w}_j)_m^2 .
\]

(20)

If bias is scale-independent, the \( j \)-spectra of \( b^j \) should also be flat.

In a word, the scale-dependence of bias is described by the \( j \) spectra of parameters \( \tilde{b}^j \) and \( b^j \). It is possible that some \( j \)-spectra are flat, but others are nonflat. Therefore, to study bias scale-dependence, it is necessary to take a systematic detection of the \( j \)-spectra of \( \tilde{b}^j \) and \( b^j \).

2.4. Non-Gaussianity and Spectra of \( \tilde{b}^j \) and \( b^j \)

A physical reason for choosing \( \tilde{b}^j \) and \( b^j \) to detect scale-dependence of bias stems from the non-Gaussianity of mass density fields.

The structure formation due to gravitational clustering can roughly be divided into three stages: linear, quasi-linear, and fully developed nonlinear. It is obvious from equation (1) that the linear evolution of \( \delta(x)_j \) and \( \delta(x)_m \) does not cause bias scale-dependence. It is generally believed that the evolution of cosmic clustering on very large scales probably still remains in linear stage at present. Therefore, the bias parameter should be scale-independent on very large scales if the initial \( b \) is scale-independent.

On scales from about 5~10 \( h^{-1} \) Mpc, the cosmic gravitational clustering may also be not fully developed yet, but partially developed in the quasi-linear regime. In this case, the power spectrum of density perturbations is not significantly different from the linear power spectrum, but it is characterized by power transfer via mode coupling. The power transfer among perturbations on different scales will give rise to a scale-dependence of bias. On the other hand, power transfer also leads to non-Gaussianity of the density fields. Therefore, the bias scale-dependence in the range of 5~10 \( h^{-1} \) Mpc may not easily be detected by methods insensitive to non-Gaussianity, such as power spectrum or two-point correlation function, but by statistics sensitive to...
non-Gaussianity. The statistics \( \tilde{b}^{(n)} \) and \( b^{(n)} \) are non-
Gaussian-sensitive and therefore suitable to detect the bias
scale-dependence caused by non-Gaussian process.

It has been shown that the \( n = 2 \) statistics of \( (\tilde{w}_{j,i}) \), i.e.,
\( \sigma_j^2 = C_j^2 \), essentially is the Fourier power spectrum. There-
fore, a Gaussian random field can be completely described
by its DWT power spectrum \( \sigma_j^2 = C_j^2 \) (Pando & Fang 1998).
Higher order \( (n > 2) \) statistics of \( (\tilde{w}_{j,i})^n \) are sensitive to non-
Gaussianity of the density fields. For instance, the \( n = 3, 4 \)
statistics of \( (\tilde{w}_{j,i})^n \) can be rewritten as

\[
S_j = \frac{1}{2\sigma^3} C_j \quad K_j = \frac{1}{2\sigma^2} C_j^2 - 3 .
\]

(21)

Therefore, \( S_j \) and \( K_j \) are the \( j \)-spectra of \( S \) and \( K \).

For Gaussian fields \( \delta(x) \), all the cumulant moments \( C^n \) or
\( C_j^n \) are equal to zero, only except second-order \( C^2 \) or \( C_j^2 \)
(Fang & Pando 1997). Thus, all statistics \( b^{(n)} \) of \( n > 2 \) are good for
detecting the scale-dependence of bias related to non-Gaussian
processes.

The statistics of the SFCs \( w_{j,i} \) are different from the
WFCS \( \tilde{w}_{j,i} \). As mentioned in § 2.3, the scaling functions
\( \phi_{j}(x) \) are window-like functions. SFCs are similar to the
statistics of counts in cell (CIC). The DWT scaling functions
\( \phi_{j}(x) \) correspond to the CIC window functions, and the
SFCs of the DWT correspond to the counts of CIC.
Because \( \phi_{j}(x) \) and the window functions of CIC are not
orthogonal with respect to \( j \) (or scale), the SFCs, \( w_{j,i} \), are
given by a superposition of WFCS, \( \tilde{w}_{j,i} \) and \( j \leq j \) (see eqs.
[8]–[10]). Namely, for any order of \( n \), statistics based on
SFCs are always sensitive to the location or the phases of
perturbations on different scales; i.e., they are sensitive to non-
Gaussianity. Thus, regardless whether \( n \) is larger than
2, \( b^{(n)} \) are always useful to detect the scale-dependence of
bias related to non-Gaussian processes.

3. SAMPLES OF GALAXIES

Neither \( \tilde{b}^{(n)} \) nor \( b^{(n)} \) are directly testable, because we do
not know the distribution of dark matter \( \delta(x) \). But bias
scale-dependence can be revealed from the difference between
the scale behaviors of the galaxies with different
internal properties. It has been very well known that the
relative abundances of galaxies with different morphology
are environment-dependent (Dressler 1980), the clustering
strength of galaxies is luminosity dependent (Xia, Deng, &
Zhou 1987), and the correlation amplitudes of optically
selected galaxies are different from that of infrared-selected
(Saunders, Rowan-Robinson, & Lawrence 1992; Strauss et
al. 1992). Therefore, it is most likely that galaxies with differ-
ent morphology possess different scale-dependence of bias.
Thus, the bias scale-dependence might be detected by com-
paring the \( j \)-spectra of \( \tilde{b}^j \) and \( b^j \) for galaxies with different
morphology.

Let us consider two types of galaxies, \( I \) and \( II \), both of
which obey linear model (1) with different parameter \( b \). All

statistics of equations (15) and (20) are available by repla-
ging \( (g, m) \) by \( (I, II) \). Namely, the scale-dependence of bias
can be revealed by \( j \)-spectrum of \( \tilde{b}^{(n)} \) and \( b^{(n)} \) defined as

\[
\tilde{b}^{(n)} = \frac{(C^n)_I}{(C^n)_II},
\]

(23)

\[
b^{(n)} = \frac{\sum_{n=0}^{2n} (w_{j,i} - \tilde{w}_{j,i})^n}{\sum_{n=0}^{2n} (w_{j,i} - \tilde{w}_{j,i})^n}.
\]

(24)

Any nonflatness of these \( j \)-spectra indicates that the bias is scale-
dependent for at least one type of the considered gal-
axies.

We analyzed the samples of galaxies listed in the
APM-BGC (Loveday 1996), which gives positions, magni-
tudes, and morphological types of 14,681 galaxies brighter
than \( 16^{m}44 \) over a 4180 deg\(^2 \) area in 180 Schmidt survey
fields of southern sky. The completeness is about 96.3\%,
with a standard deviation of 1.9% inferred from 12 carefully
checked fields. Therefore, it is large and uniform enough for
a two-dimensional DWT analysis.

We choose the early type or elliptical and lenticular (EL)
galaxies as \( I \) and late type or spiral (SP) galaxies as \( II \). All
the elliptical and lenticular galaxies are compiled into one sample containing 4439 ELs. The SP sample contains 8217 SP
galaxies.

In order to do a two-dimensional DWT analysis, we
chose three fields S1, S2, and S3 from the entire survey area.
The three fields are selected to be square on an equal area
projection of the sky. The equal area projection keeps the
surface number density of the galaxies in the plane the same
as on the sky. The sizes of S1, S2, and S3 are taken to be as
large as possible, covering the whole area of the survey.
The whole survey and the three fields, S1, S2, and S3, are plotted
in Figure 1. Each field has angular size of about 37 deg\(^2 \),
S1, S2, and S3 contain 1095, 1039, and 1055 ELs and 2186,
2188, and 2092 SPs, respectively. The galaxies in the regions
S1 and S3 are completely independent. S2 has some over-
laps with S1 and S3.

We divide each square region into \( 2^{10} \times 2^{10} (1024^2) \) cells
labeled by \( l = (l_1, l_2) \), \( l_1, l_2 \) being an integer from 0 to 1023.
Using the estimation of the mean depth of the sample given
by the luminosity function from the Stromlo-APM redshift
survey (Loveday et al. 1992), one finds that the cell size is
about 65 h\(^{-1} \) Kpc, which is fine enough to detect statistical
features on scales larger than 1 h\(^{-1} \) Mpc.

The distribution of the galaxies in the sample can be
be described by a two-dimensional density field \( \rho(x) \) or con-
trast \( \delta(x) \), where \( x = (x_1, x_2) \). In doing two-dimensional
DWT analysis of \( \rho(x) \) or \( \delta(x) \), the wavelet functions and the
scaling functions are constructed from the direct product of the
one-dimensional bases; i.e., \( \psi_{j,k}(x) \equiv \psi_{j_1,k_1}(x_1)\psi_{j_2,k_2}(x_2) \),
and \( \phi_{j,k}(x) \equiv \phi_{j_1,k_1}(x_1)\phi_{j_2,k_2}(x_2) \), where \( j = (j_1, j_2) \) and
\( l = (l_1, l_2) \). For scale \( j \), the angular scale is 2\(^{2-\ell} \) in units of a

cell. In this paper, we are interested in scales of \( j = 8, 7, 6, 5,
4, and 3 \), corresponding to spatial scales of about 0.26, 0.52,
1.04, 2.08, 4.16, and 8.32 h\(^{-1} \) Mpc, respectively.

In doing statistical analyses of the APM-BGC, the effect of
the 1456 holes drilled around big bright objects should be
taken into account. For instance, random samples are gen-

erated in the area with the same drilled holes as that in the
original survey. We also removed a few galaxies placed in
the holes drilled.

A common problem of statistics for a discrete sample is
sampling error, which is, in particular, serious for statistics
of non-Gaussianity. Even if the original matter field is Gaussian, the sampled data must be non-Gaussian on scales for which the mean number in one cell is small. This is the non-Gaussianity of shot noise. Any non-Gaussian behavior of the density fields will be contaminated by the shot noise. Because bias might be related to non-Gaussian process, the detected result is inevitably contaminated by sampling. Fortunately, the DWT spectrum method is found to be effective for suppressing the contamination of shot noise. It has been shown that the non-Gaussianity of shot noise is significant only on scales comparable with the mean distance of the nearest neighbors of the objects. Because the wavelet bases are localized, the central limit theorem guarantees the non-Gaussianity of shot noise rapidly and monotonously approaching zero on larger scales (Greiner, Lipa, & Carruthers & Pando Therefore,1995; Fang 1997). besides small scales, the behavior of the DWT \( j \)-spectrum will not be affected by the shot noise of sampling.

4. DETECTION OF \( j \)-SPECTRUM

4.1. Two-Point Correlation Function

Before calculating \( b_I^0 \), we consider the two-point angular correlation functions of the ELs and SPs of the APM-BGC, because the two-point angular correlation function analysis of Stromlo-APM redshift survey has been done by Loveday et al. (1995). One can show the reliability of the DWT method by comparing our result with the earlier one.

Figure 2 shows the angular correlation functions, \( w_{\text{EL}}(\theta) \) and \( w_{\text{SP}}(\theta) \), which are obtained from whole sample of the ELs and SPs. The dotted lines give the 1 \( \sigma \) error calculated from 20 random samples, which are produced by randomizing the positions of the galaxies in the same field with the same drilled holes as the samples.

As usual, these correlation functions show a power law, and the amplitude of the correlation for the ELs is higher than that of the SPs. This is the well-known segregation of morphology: the clustering of EL galaxies is stronger than that of SP (Davis & Geller 1976). Actually, this is an evidence of bias. However, Figure 2 shows that the ratio \( w_{\text{EL}}(\theta)/w_{\text{SP}}(\theta) \) is approximately constant in the angular range of \(-0.4 < \log \theta < 0.6\). Namely, the segregation can be explained by linear bias if \( b_I \) and \( b_{II} \) are taken to be different constants. In other words, in terms of two-point correlation functions, no scale-dependent bias is needed, or at most a very weak scale-dependence. With the DWT analysis, we are able to measure this scale-dependence more quantitatively.

This result of Figure 2 will not be changed if the error is estimated by bootstrap samples. This is expected, because the number of galaxy pairs is large; the bootstrap-resampling error is only larger than Poissonian error by a factor of 1.7 (Mo, Jing, & Börner 1992).

4.2. Statistics \( B_I^{(2)} \) and \( B_I^{(3)} \)

Power spectrum is the Fourier counterpart of two-point correlation function. As mentioned in § 2.4, the second-order statistic \( \langle \tilde{w}_I^2 \rangle \) is equivalent to the measure of Fourier power spectrum (Pando & Fang 1998). Therefore, one can expected that the second-order statistics \( \langle B_I^{(2)} \rangle \) should show the same results as that given by two-point correlation function.

![Fig. 1.—Equal project of the APM bright galaxy distribution and three regions S1, S2, and S3, in which the wavelet analysis has been done. The top and bottom panels are for EL and SP galaxies, respectively.](image1)

![Fig. 2.—Two-point angular correlation functions of EL (solid lines) and SP (dashed lines) galaxies in APM-BGC.](image2)
The two-dimensional extension of the $n = 2$ statistic $\langle \hat{B}^{(2)} \rangle$ of equation (23) is

$$\hat{B}^{(2)}_{j_1,j_2} = \frac{\sum_{l_1=0}^{2^{2j_1}-1} \sum_{l_2=0}^{2^{2j_2}-1} w_{j_1,l_1}^2}{\sum_{l_1=0}^{2^{2j_1}-1} \sum_{l_2=0}^{2^{2j_2}-1} w_{j_1,l_1}^2}. \quad (25)$$

In order to estimate the error of $\hat{B}^{(2)}_{j_1,j_2}$ and to avoid the effect of the drilled holes and boundary, we generated 500 randomized samples of the EL and SP and calculated the average $\hat{B}^{(2)}_{j_1,j_2}$ and variance from these random samples. Figure 3 plots $\hat{B}^{(2)}_{j_1,j_2}$, defined as

$$\hat{B}^{(2)}_{j_1,j_2} = \frac{\hat{B}^{(3)}_{j_1,j_2}}{\hat{B}^{(3)}_{j_1,j_2}}. \quad (26)$$

The variance from the random samples is also plotted in Figure 3. The $j$-spectrum of $\hat{B}^{(2)}_{j_1,j_2}$, indeed, shows flatness within 1 $\sigma$; only field S2 has a little higher $\hat{B}^{(2)}_{j_1,j_2}$ on large-scale $j = 3$. This result is in good agreement with the detection of two-point correlation functions (Loveday et al. 1995). This shows that the DWT power spectrum estimator is reliable (Pando & Fang 1998).

It has been emphasized in § 2.4 that the flatness of the $j$-spectrum of the second-order statistics $\hat{B}^{(2)}_{j_1,j_2}$ does not imply that other $j$-spectra will also be flat, especially in the case that the bias scale-dependence is caused by non-Gaussian process. Therefore, non-Gaussian sensitive statistics are necessary. Figure 4 shows the result of statistics $\hat{B}^{(3)}_{j_1,j_2}$, which is defined by

$$\hat{B}^{(3)}_{j_1,j_2} = \frac{\hat{B}^{(3)}_{j_1,j_2}}{\hat{B}^{(3)}_{j_1,j_2}}, \quad (27)$$

where

$$\hat{B}^{(3)}_{j_1,j_2} = \frac{\sum_{l_1=0}^{2^{2j_1}-1} \sum_{l_2=0}^{2^{2j_2}-1} w_{j_1,l_1}^3}{\sum_{l_1=0}^{2^{2j_1}-1} \sum_{l_2=0}^{2^{2j_2}-1} w_{j_1,l_1}^2}. \quad (28)$$

Comparing Figure 4 with Figure 3, it is clear that the $j$-spectra of $\hat{B}^{(3)}_{j_1,j_2}$ show more deviation from a flat spectrum. Therefore, the bias scale-dependence might really originate from non-Gaussian processes.

4.3. Statistics $A_j$

A simplest phase-sensitive statistic is the reconstructed distribution $\rho(x)$ (eq. [9]), which is given by SFCs. The two-dimensional extension of equation (9) is

$$\rho(x) = \sum_{l_1=0}^{2^{2j_1}-1} \sum_{l_2=0}^{2^{2j_2}-1} w_{j_1,l_1} \phi_{j_1,l_1}(x). \quad (29)$$

The expression $j = (j_1, j_2)$ means a reconstruction of the density field on scale $j$ in the dimension $x_1$ and $j_2$ in $x_2$. Therefore, when $j = (j_1, 0)$, $\rho(x)$ is a projection on axis $x_1$, and when $j = (0, j_2)$, it is a projection on axis $x_2$. In the case that $j_1 = j_2 = j$, the field is smoothed on the scale $j$ in both directions, $x_1$ and $x_2$.

A DWT reconstruction on the three fields S1, S2 and S3 was performed. Since the SFCs, $w_{j_1,l_1}$ are proportional to the
Fig. 5.—Reconstructed distributions of S1 on scales $j_i = j_j = 6, 5, 4, 3$ from top to bottom. Results on the left are for EL galaxies, and results on the right are for SPs. The bold contours are for $w_{S1}^{j_i,j_j} - w_{S1}^{j_j} > 2, 3, 4 ...$, from outside to inside, successively, and the light contours denote the area with $w_{S1}^{j_i,j_j} / w_{S1}^{j_j} \leq 1/2, 1/8, 1/32, 1/128$ from outside to inside, successively.
Fig. 6.—Simplified representation of Fig. 5. Dark area denote the overdense regions, $w_{j_1,j_2,j_3} > \overline{w}_j$, and the blank area the underdense regions, $w_{j_1,j_2,j_3} < \overline{w}_j$ where $j_1 = j_2 = j$. 
density at position \( l \), the mean density of the reconstructed density field is proportional to

\[
\overline{w}_j = \frac{1}{2^i_2} \sum_{l_2=0}^{2^{2j}-1} w_{j,l_2}.
\]  

The cells with \( w_{j,l} > \overline{w}_j \) are dense regions (clumps), and cells with \( w_{j,l} < \overline{w}_j \) are underdense regions (voids). Actually, for all \( j \), \((\overline{w}_j)_I \) and \((\overline{w}_j)_II \) are always proportional to the mean number density of galaxies EL and SP, respectively, and therefore, the ratio \((\overline{w}_j)_I/((\overline{w}_j)_II \) does not depend on \( j \).

Figure 5 shows reconstructed distributions of S1 on scales \( j_1 = j_2 = j = 6, 5, 4, 3 \), in which the bold contours denote the area with \( w_{j_1,j_2,l_1,l_2} - \overline{w}_j > 2, 3, 4 \ldots \sigma_j \) from outside to inside successively, and the 1 \( \sigma \) line is calculated from 500 random samples for each field. Figure 5 also shows the underdense regions by light contours corresponding to \( w_{j_1,j_2,l_1,l_2}/\overline{w}_j \leq 1/2, 1/8, 1/32, 1/128 \) successively from outside to inside.

A simplified representation of these fields is shown in Figure 6, in which the dark area represents the overdense regions, \( w_{j_1,j_2,j} > \overline{w}_j \), and the blank area represents the underdense regions, \( w_{j_1,j_2,j} < \overline{w}_j \). The ratio between the clustering strengths of the EL and SP on scale \( j \) can be measured from Figure 6 by

\[
a_j = \frac{\text{(dark area)}_\text{EL}}{\text{(dark area)}_\text{SP}}.
\]  

Similar to the statistic \( B_j^{(2)} \), to estimate the error of \( a_j \), we generated 500 randomized samples of the ELs and SPs and calculated the average \( \overline{a}_j \) of these random samples. Figure 7 plots the result of statistics \( A_j \) defined as

\[
A_j = \frac{a_j}{\overline{a}_j}.
\]

The 1 \( \sigma \) lines of both randomized and bootstrap resampling are also shown in Figure 7. The two error estimates gave about the same confidence level. The \( j \)-spectrum of \( A_j \) is found to be significantly nonflat.

4.4. Statistics \( B_j \)

Since \( A_j \) is based on statistic \( w_{j,l} \), one can expected that the statistic of \( b_j^{(1)} \) of equation (21) will give the same conclu-
where we use absolute value because for odd \(m\) limit magnitude \(m\). To study the influence of sample depth on our statistics, we (redshift) projection of their three-dimensional distribution.

The two-dimensional extension of equation (21) is

\[
b_{j_1,j_2}^{(1)} = \frac{\left( \sum_{l_1=0}^{2^{j_1}-1} \sum_{l_2=0}^{2^{j_2}-1} |w_{j_1,l_1} - \bar{w}_{j_1}| \right)}{\left( \sum_{l_1=0}^{2^{j_1}-1} \sum_{l_2=0}^{2^{j_2}-1} |w_{j_1,l_1} - \bar{w}_{j_1}| \right)_{I}},
\]

where we use absolute value \(|w_{j_1,l_1} - \bar{w}_{j_1}|\), because for odd \(n\), \(\langle (w_{j_1,l_1} - \bar{w}_{j_1})^n \rangle\) has larger relative errors. We calculated \(B_j^{(1)}\) defined as

\[
B_j^{(1)} = \frac{b_{j_1,j_2}^{(1)}}{\sigma_j},
\]

where \(\bar{b}_{j_1,j_2}^{(1)}\) is the average of \(b_{j_1,j_2}^{(1)}\) over 500 randomized samples of the ELs and SPs. The results of \(B_j^{(1)}\) for the three fields are given in Figure 8. The three fields S1, S2, and S3 show the same feature of the \(j\)-dependence. The mean SFCs and their variance from the randomized samples are also plotted in Figure 8. Like the statistics \(A_j\), the spectrum of \(B_j^{(1)}\) is nonflat with significance larger than \(2\sigma\).

Since the two-dimensional sample of galaxies is a \(z\) (redshift) projection of their three-dimensional distribution. To study the influence of sample depth on our statistics, we calculated \(B_j^{(1)}\) for samples of ELs and SPs with different limit magnitude \(m\). Figure 9 plots the result of \(B_j^{(1)}\), which is the same as \(B_j^{(1)}\) but for sample with limit magnitude \(m = 16.20\). In this case, the numbers of EL and SP galaxies are about 20%-25% less than that of \(m = 16.44\). The \(j\)-spectra in Figure 9 have completely the same features as Figure 8. Therefore, conclusions drawn from Figure 8 are insensitive to the depth of the samples.

In Figures 8 and 9, the \(j\)-dependence of bias is prominent on scales larger than \(1 \h^{-1}\) Mpc, i.e., larger than the mean distance of nearest neighbors of galaxies; therefore, the effect of shot noise is negligible.

5. DISCUSSIONS AND CONCLUSIONS

We showed that, instead of one parameter \(b\), equation (1) introduces a series of parameters \(b_j^{(n)}\) and \(b_j^{(n)}\) to describe the bias scale-dependence. The statistics of \(b_j^{(n)}\) with different \(n\) are independent from each others. The DWT analysis provides a simple and effective tool of systematically detecting the scale-dependence of bias parameters on various orders. This method is effectively employed for analyzing galaxy samples, which show morphology segregation and/or luminosity segregation.

With this method we detected the \(j\)-spectra of \(b_j^{(n)}\) and \(b_j^{(n)}\) for the distributions of EL and SP galaxies of the APM-BGC samples. The general results indicate that for second-order statistics, i.e., two-point correlation function and power spectrum, the bias is approximately scale-independent, but not so for higher order or phase-sensitive (non-Gaussian) statistics. The result is consistent with the following fact: most evidence for weak scale-dependence of bias are from statistics of the two-point correlation functions and power spectrum (Kaufmann, Nusser, & Steinmetz 1997), while the evidence for scale-dependence is from phase-sensitive statistics (Sigad et al. 1998). Therefore, the scale-dependence of galaxy bias may have the same origin as the non-Gaussianity of galaxy distribution. Linear evolution cannot cause non-Gaussianity of mass distributions if the initial perturbations are Gaussian. The scale-dependence of bias of galaxy distribution is most unlikely because of the nonlinear evolution of gravitational clustering and the nonlocal relationship between galaxy formation and their environment.

With this in mind, the information of bias scale-dependence is valuable for developing models of galaxy formation. Indeed, despite the fact that the current detection of bias scale-dependence is still very preliminary, the result is already able to set useful constraint on models of hierarchical clustering. In these models, galaxy correlations are generally assumed to be described by the hierarchical relation \(\xi_n = Q_n \xi_{n-2}\), where \(\xi_n\) is the \(n\)-th order correlation function, and \(Q_n\) are constants (White 1979). If these hierarchical relations hold exactly, the second-order (two-point) correlation function plus all constants \(Q_n\) (which may be different for different types of galaxies) completely characterize the clustering of galaxies, including their higher order correlations.

This is, all \(b_j^{(n)}\) and \(b_j^{(n)}\) can be represented by second-order correlation function plus all (scale-independent) constants \(Q_n\). Hence, if bias is scale-independent on second-order, it will be scale-independent on all orders. Therefore, the nonflatness of \(j\)-spectra of \(b_j^{(n)}\) and \(b_j^{(n)}\) for sample APM-BGC implies that the hierarchical relations may not hold exactly, or the coefficients, \(Q_n\), are scale-dependent. Similar conclusions have also been drawn from the detection of the scale-scale correlations of the Ly\(\alpha\) forests of QSO absorption spectrum (Pando et al. 1998). Thus, the detection of \(b_j^{(n)}\) and \(b_j^{(n)}\), joining with other higher order statistics, is effective in revealing the details of the hierarchical clustering scenario.
We thank J. Pando and Y. P. Jing for many helpful comments. Z.-G. D. and X. X. Y. were supported by the National Science Foundation of China.

REFERENCES

Adler, R. J. 1981, The Geometry of Random Fields (New York: Wiley)
Bower, R. G., Coles, P., Frenk, C. S., & White, S. D. M. 1993, ApJ, 405, 403
Catelan, P., Coles, P., Matarrese, S., & Moscardini, L. 1994, MNRAS, 268, 966
Coles, P. 1993, MNRAS, 262, 1065
Couchman, H. M. P., & Carlberg, R. G. 1992, ApJ, 389, 453
Daubechies, I. 1992, Ten Lectures on Wavelets (Philadelphia: SIAM)
Davis, M., & Geller, M. J. 1976, ApJ, 208, 13
Dekel, A., Burstein, D., & White, S. D. M. 1997, in Critical Dialogues in Cosmology, ed. N. Turok (Singapore: World Scientific), 67
Dekel, A., & Rees, M. 1987, Nature, 326, 455
Dressler, A. 1980, ApJ, 236, 351
Fang, L. Z., & Pando, J. 1997, in 5th Current Topics of Astrofundamental Phys., ed. N. Sanchez & A. Zichchi (Singapore: World Scientific), 66
Greiner, M., Lipa, P., & Carruthers, P. 1995, Phys. Rev. E, 51, 1948
Kaiser, N. 1984, ApJ, 284, L9
Kauffmann, G., Nusser, A., & Steinmetz, M. 1997, MNRAS, 286, 795
Loveday, L. 1996, MNRAS, 278, 1025
Loveday, L., Maddox, S. J., Efstathiou, G., & Peterson, B. A. 1995, ApJ, 442, 457
Loveday, L., Peterson, B. A., Efstathiou, G., & Maddox S. J. 1992, ApJ, 390, 338
Mo, H. J., Jing, Y. P., & Börner, G. 1992, ApJ, 392, 452
Mo, H. J., Jing, Y. P., & White, S. D. M. 1996, MNRAS, 284, 189
Pando, J., & Fang, L. Z. 1996, ApJ, 459, 1
———. 1998, Phys. Rev. E, 57, 3593
Pando, J., Lipa, P., Greiner, M., & Fang, L. Z. 1997, ApJ, 469, 9
Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Rees, M. 1985, MNRAS, 213, 75
Saunders, W., Rowan-Robinson, M., & Lawrence, A. 1992, MNRAS, 258, 134
Sigad, Y., Elder, A., Dekel, A., Strauss, M. A., & Yahil, A. 1998, ApJ, 495, 516
Strauss, M. A., Davis, M., Yahil, A., & Huchra, J. P. 1992, ApJ, 385, 421
Vanmarcke, E. 1983, Random Fields, Analysis, and Synthesis (Cambridge: MIT Press)
White, S. D. M. 1979, MNRAS, 186, 145
Xia, X. Y., Deng, Z. G., & Zhou, Y. Y. 1987, in Observational Cosmology, ed. A. Hewitt, G. Burbidge & L. Z. Fang (Dordrecht: Reidel), 363