THE ASTROPHYSICAL JOURNAL, 561:22–34, 2001 November 1

PROBABILITY DISTRIBUTION FUNCTION OF COSMOLOGICAL DENSITY FLUCTUATIONS FROM A GAUSSIAN INITIAL CONDITION: COMPARISON OF ONE-POINT AND TWO-POINT LOGNORMAL MODEL PREDICTIONS WITH *N*-BODY SIMULATIONS

ISSHA KAYO, ATSUSHI TARUYA, AND YASUSHI SUTO
Department of Physics and Research Center for the Early Universe (RESCEU), School of Science, University of Tokyo, Tokyo 113-0033, Japan;
kayo@utap.phys.s.u-tokyo.ac.jp, ataruya@utap.phys.s.u-tokyo.ac.jp, suto@phys.s.u-tokyo.ac.jp
Received 2001 April 2; accepted 2001 July 3

ABSTRACT

We quantitatively study the probability distribution function (PDF) of cosmological nonlinear density fluctuations from *N*-body simulations with a Gaussian initial condition. In particular, we examine the validity and limitations of one-point and two-point lognormal PDF models against those directly estimated from the simulations. We find that the one-point lognormal PDF very accurately describes the cosmological density distribution even in the nonlinear regime (rms variance $\sigma_{\rm nl} \lesssim 4$, overdensity $\delta \lesssim 100$). Furthermore, the two-point lognormal PDFs are also in good agreement with the simulation data from linear to fairly nonlinear regimes, while they deviate slightly from the simulation data for $\delta \lesssim -0.5$. Thus, the lognormal PDF can be used as a useful empirical model for the cosmological density fluctuations. While this conclusion is fairly insensitive to the shape of the underlying power spectrum of density fluctuations $P(k)$, models with substantial power on large scales, i.e., $n \equiv d\ln P(k)/d\ln k \lesssim -1$, are better described by the lognormal PDF. On the other hand, we note that the one-to-one mapping of the initial and evolved density fields, consistent with the lognormal model, does not approximate the broad distribution of their mutual correlation even on average. Thus, the origin of the phenomenological lognormal PDF approximation still remains to be understood.

Subject headings: cosmology: theory — dark matter — galaxies: clusters: general — large-scale structure of universe — methods: numerical

1. INTRODUCTION

The probability distribution function (PDF) of the cosmological density fluctuations is the most fundamental statistic characterizing the large-scale structure of the universe. In the standard picture of gravitational instability, the PDF of the primordial density fluctuations that are responsible for the current structures in the universe is assumed to obey a random Gaussian distribution. Therefore, it is fully specified by the two-point correlation function $\xi(r)$ or, equivalently, the power spectrum $P(k)$. As long as the density fluctuations are in the linear regime, their PDF remains Gaussian. Once they reach the nonlinear stage, however, their PDF significantly deviates from the initial Gaussian shape because of the strong nonlinear mode coupling and the nonlocality of the gravitational dynamics. The functional form for the resulting PDFs in nonlinear regimes are not known exactly, and a variety of phenomenological models have been proposed (Saslaw 1985; Suto, Itoh, & Inagaki 1990; Lahav et al. 1993; Gaztañaga & Yokoyama 1993; Suto 1993; Ueda & Yokoyama 1996). Once such a one-point PDF is specified, one can characterize the clustering of the universe with the higher order statistics such as skewness and kurtosis. Moreover, the two-point PDF is useful in estimating the errors in the one-point statistics due to finite sampling since the measurement at different positions is not independent, and their correlations are supposed to be dominated by the two-point correlation function (Colombi, Bouchet, & Schaeffer 1995; Szapudi & Colombi 1996). Also, the two-point PDF plays an important role in analytical modeling of dark halo biasing on two-point statistics.

From an empirical point of view, Hubble (1934) first noted that the galaxy distribution in angular cells on the celestial sphere may be approximated by a lognormal distribution rather than a Gaussian one. More recent analysis of the three-dimensional distribution of galaxies indeed confirmed this (e.g., Hamilton 1985; Bouchet et al. 1993; Kofman et al. 1994). Interestingly, several *N*-body simulations in cold dark matter (CDM) models also indicated that the PDF of density fluctuations is fairly well approximated by the lognormal (e.g., Coles & Jones 1991; Kofman et al. 1994; Taylor & Watts 2000), at least in a weakly nonlinear regime.

Those observational and numerical indications have not yet been understood theoretically; Bernardeau (1992, 1994) showed that the PDF computed from the perturbation theory in a weakly nonlinear regime approaches the lognormal form only when the primordial power spectrum is proportional to k^n with $n = -1$. On the basis of this result, Bernardeau & Kofman (1995) argued that the successful fit of the lognormal PDF in the CDM models should be interpreted as accidental and simply results from the fact that the power spectra of those models are well approximated by $k^{n_{\rm eff}}$ with $n_{\rm eff} \simeq -1$ on scales of cosmological interest. They claimed that the lognormal PDF may fail either in a highly nonlinear regime or in models with power spectrum with $n_{\rm eff} \neq -1$. In fact, Ueda & Yokoyama (1996) conclude that the lognormal PDF does not fit well the PDF in a highly nonlinear regime, from the analysis of CDM simulations by Suginohara & Suto (1991) employing $N = 64^3$ particles in a 100 Mpc box.

The aim of this paper is to study the extent to which the lognormal model describes the PDF in weakly and highly nonlinear regimes using the high-resolution *N*-body simulations with $N = 256^3$. In particular, we extend our analysis to the two-point PDF in addition to the one-point PDF discussed previously. Bernardeau (1996) analytically computed the two-point PDF using the perturbation technique

and compared it somewhat indirectly with N-body simulations in a weakly nonlinear regime. In contrast, we focus on the highly nonlinear regime and examine the validity of the empirical lognormal model.

This paper is organized as follows: § 2 describes the lognormal PDF derived through one-to-one mapping between the linear and nonlinear density fields. The detailed comparison between the lognormal predictions and N-body results is presented in § 3. Finally, § 4 is devoted to conclusions and discussion.

2. PROBABILITY DISTRIBUTION FUNCTIONS FROM THE LOGNORMAL TRANSFORMATION

In this section we briefly outline the derivation of the lognormal PDF assuming a one-to-one correspondence between the linear and evolved density fluctuations. Throughout the paper, we consider the mass density field $\rho(\boldsymbol{x}; R)$ at the position $\boldsymbol{x}$ smoothed over the scale R. This is related to the *unsmoothed* density field $\rho(\boldsymbol{x})$ as

$$\rho(\boldsymbol{x}; R) = \int d^3\boldsymbol{y} W(|\boldsymbol{x}-\boldsymbol{y}|; R)\rho(\boldsymbol{y}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, \tilde{W}(kR)\tilde{\rho}(\boldsymbol{k})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} . \quad (1)$$

In the above expression, W denotes the window function and $\tilde{W}$ and $\tilde{\rho}$ represent the Fourier transforms of the corresponding quantities. In what follows, we adopt the two conventional windows

$$\tilde{W}(x) = \begin{cases} e^{-x^2/2} & \text{Gaussian} , \\ 3(\sin x - x\cos x)/x^3 & \text{top hat} . \end{cases} \quad (2)$$

Then, the density contrast at the position $\boldsymbol{x}$ is defined as $\delta(\boldsymbol{x}; R) \equiv [\rho(\boldsymbol{x}; R) - \bar{\rho}]/\bar{\rho}$, with $\bar{\rho}$ denoting the spatial average of the smoothed mass density field. For simplicity we use δ to denote $\delta(\boldsymbol{x}; R)$ unless otherwise stated.

2.1. *One-Point Lognormal PDF*

The one-point lognormal PDF of a field δ is defined as

$$P_{\rm LN}^{(1)}(\delta) = \frac{1}{(2\pi\sigma_1^2)^{1/2}} \exp\left\{-\frac{[\ln(1+\delta)+\sigma_1^2/2]^2}{2\sigma_1^2}\right\}\frac{1}{1+\delta} . \quad (3)$$

The above function is characterized by a single parameter σ_1 that is related to the variance of δ. Since we use δ to represent the density fluctuation field smoothed over R, its variance is computed from its power spectrum $P_{\rm nl}$ explicitly as

$$\sigma_{\rm nl}^2(R) \equiv \frac{1}{2\pi^2}\int_0^\infty P_{\rm nl}(k)\tilde{W}^2(kR)k^2\, dk . \quad (4)$$

Here and in what follows we use subscripts "lin" and "nl" to distinguish the variables corresponding to the primordial (linear) and the evolved (nonlinear) density fields, respectively. Then σ_1 depends on the smoothing scale R alone and is given by

$$\sigma_1^2(R) = \ln[1+\sigma_{\rm nl}^2(R)] . \quad (5)$$

Given a set of cosmological parameters, one can compute $\sigma_{\rm nl}(R)$ and thus $\sigma_1(R)$ very accurately using a fitting formula for $P_{\rm nl}(k)$ (e.g., Peacock & Dodds 1996, hereafter PD). In this sense, the above lognormal PDF is completely specified and allows the definite comparison against the numerical simulations (see § 3).

It is known that the above lognormal function may be obtained from the one-to-one mapping between the linear random Gaussian and the nonlinear density fields (e.g., Coles & Jones 1991). We define a linear density field g smoothed over R obeying the Gaussian PDF,

$$P_{\rm G}^{(1)}(g) = \frac{1}{(2\pi\sigma_{\rm lin}^2)^{1/2}}\exp\left(-\frac{g^2}{2\sigma_{\rm lin}^2}\right), \quad (6)$$

where the variance is computed from its linear power spectrum

$$\sigma_{\rm lin}^2(R) \equiv \frac{1}{2\pi^2}\int_0^\infty P_{\rm lin}(k)\tilde{W}^2(kR)k^2\, dk . \quad (7)$$

If one introduces a new field δ from g as

$$1+\delta = \frac{1}{(1+\sigma_{\rm nl}^2)^{1/2}}\exp\left\{\frac{g}{\sigma_{\rm lin}}[\ln(1+\sigma_{\rm nl}^2)]^{1/2}\right\}, \quad (8)$$

the PDF for δ is simply given by $(dg/d\delta)P_{\rm G}^{(1)}(g)$, which reduces to equation (3).

At this point, the transformation equation (8) is nothing but a mathematical procedure to relate the Gaussian and lognormal functions. Thus, there is no physical reason to believe that the new field δ should be regarded as a nonlinear density field evolved from g, even in an approximate sense. In fact, it is physically unacceptable since the relation, if taken at face value, implies that the nonlinear density field is completely determined by its linear counterpart locally. We know, on the other hand, that the nonlinear gravitational evolution of cosmological density fluctuations proceeds in a quite nonlocal manner and is sensitive to the surrounding mass distribution.

Nevertheless, the fact that the lognormal PDF provides a good fit to the simulation data empirically as discussed in § 1 implies that the transformation equation (8) somehow captures an important aspect of the nonlinear evolution in the real universe. In § 3 we present detailed discussion on this problem. Before that, we derive the two-point lognormal PDF by applying this transformation in § 2.2.

2.2. *Two-Point Lognormal PDF*

Consider two density fields, $\delta_1 = \delta(\boldsymbol{x}_1; R)$ and $\delta_2 = \delta(\boldsymbol{x}_2; R)$, located at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, respectively, and smoothed over R. We denote the two-point PDF by $P^{(2)}(\delta_1, \delta_2; r)$ for the two fields with a specified separation r, i.e., satisfying the condition $|\boldsymbol{x}_1 - \boldsymbol{x}_2| = r$.

In the case of the Gaussian fields g_1 and g_2, this two-point PDF is given by the bivariate Gaussian (e.g., Bardeen et al. 1986)

$$P_{\rm G}^{(2)}(g_1, g_2; r) = \frac{1}{2\pi\sqrt{\det M}}\exp\left[-\frac{1}{2}(g_1, g_2)M^{-1}\begin{pmatrix} g_1 \\ g_2\end{pmatrix}\right], \quad (9)$$

where

$$M \equiv \begin{pmatrix} \langle g_1^2\rangle & \langle g_1 g_2\rangle \\ \langle g_1 g_2\rangle & \langle g_2^2\rangle \end{pmatrix} = \begin{bmatrix} \sigma_{\rm lin}^2 & \xi_{\rm lin}(r) \\ \xi_{\rm lin}(r) & \sigma_{\rm lin}^2 \end{bmatrix}, \quad (10)$$

$$\xi_{\rm lin}(r; R) = \frac{1}{2\pi^2}\int_0^\infty P_{\rm lin}(k)\tilde{W}^2(kR)\frac{\sin(kr)}{kr}k^2\, dk . \quad (11)$$

TABLE 1

SIMULATION PARAMETERS FOR THE CDM MODELS

Model	Ω_0	λ_0	Γ^a	σ_8	L_{box} (h^{-1} Mpc)	Realizations
SCDM	1.0	0.0	0.50	0.6	100	3
LCDM.....	0.3	0.7	0.21	1.0	100	3
LCDM300...	0.3	0.7	0.21	1.0	300	3
OCDM	0.3	0.0	0.25	1.0	100	3

[a] Shape parameter of the power spectrum.

From an analogy of equation (8), let us assume that the transformation from (g_1, g_2) to (δ_1, δ_2) is given by the form

$$1 + \delta_i = \alpha e^{\beta g_i} \quad (i = 1, 2) . \tag{12}$$

The coefficients α and β are determined by the following conditions:

$$\langle \delta_1 \rangle = \langle \delta_2 \rangle = 0 , \tag{13}$$

$$\langle \delta_1^2 \rangle = \langle \delta_2^2 \rangle = \sigma_{nl}^2 , \tag{14}$$

$$\langle \delta_1 \delta_2 \rangle = \xi_{nl}(r) . \tag{15}$$

In the above expressions, angular brackets denote the average over the two-point PDF, which in the present model reduces to

$$\begin{aligned} \langle \mathcal{F}(\delta_1, \delta_2) \rangle &\equiv \int_{-1}^{\infty}\int_{-1}^{\infty} \mathcal{F}(\delta_1, \delta_2) P^{(2)}(\delta_1, \delta_2; r) d\delta_1 \, d\delta_2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{F}[\delta_1(g_1), \delta_2(g_2)] \\ &\quad \times P_G(g_1, g_2; r) dg_1 \, dg_2 . \end{aligned} \tag{16}$$

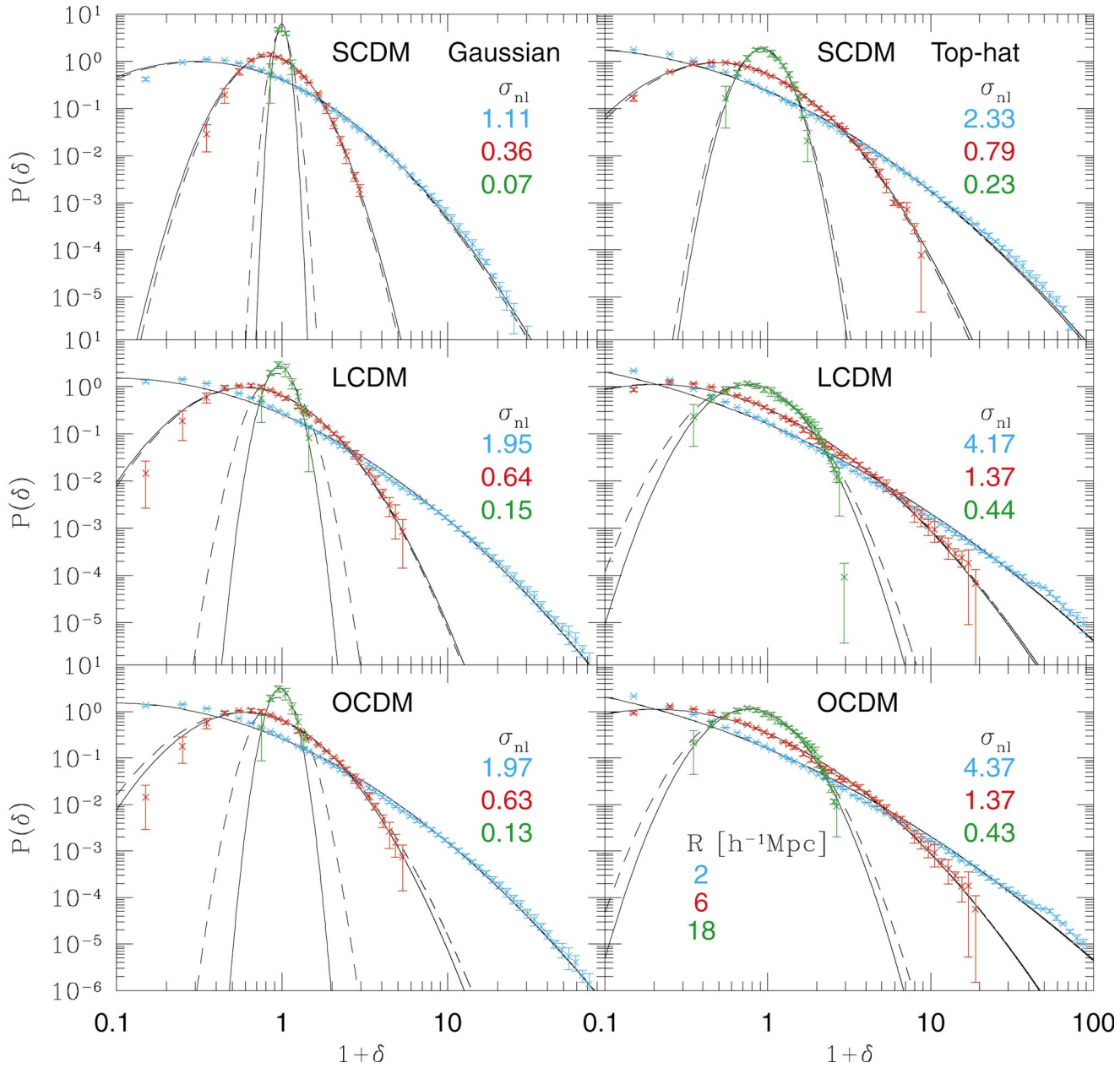

FIG. 1.—One-point PDFs in CDM models with Gaussian (*left panels*) and top-hat (*right panels*) smoothing windows; $R = 2$ (*cyan*), 6 (*red*), and 18 h^{-1} Mpc (*green*). The top, middle, and bottom panels correspond to the PDFs in SCDM, LCDM, and OCDM, respectively. *Solid lines*: the lognormal PDF adopting σ_{nl} calculated directly from the simulations. *Long-dashed lines*: the lognormal PDF estimated from the nonlinear fitting formula of Peacock & Dodds (1996). Values of σ_{nl} in each panel are estimated from the simulations.

After a straightforward calculation, one obtains

$$\alpha = \frac{1}{(1+\sigma_{\rm nl}^2)^{1/2}}\,, \quad \beta = \left[\frac{\ln(1+\xi_{\rm nl})}{\xi_{\rm lin}}\right]^{1/2}. \tag{17}$$

Then this procedure yields the two-point lognormal PDF,

$$P_{\rm LN}^{(2)}(\delta_1, \delta_2; r) = \frac{1}{2\pi(S^2 - X^2)^{1/2}} \times \exp\left[-\frac{S(L_1^2 + L_2^2) - 2XL_1L_2}{2(S^2 - X^2)}\right] \times \frac{1}{(1+\delta_1)(1+\delta_2)}\,, \tag{18}$$

where

$$X \equiv \ln(1+\xi_{\rm nl})\,, \tag{19}$$

$$S \equiv \ln(1+\sigma_{\rm nl}^2)\,, \tag{20}$$

$$L_i \equiv \ln[(1+\delta_i)\sqrt{1+\sigma_{\rm nl}^2}]\,, \quad (i = 1, 2)\,. \tag{21}$$

Again the nonlinear two-point correlation function $\xi_{\rm nl}(r; R)$ can be computed as

$$\xi_{\rm nl}(r; R) = \frac{1}{2\pi^2}\int_0^\infty P_{\rm nl}(k)\tilde{W}^2(kR)\,\frac{\sin(kr)}{kr}\,k^2\,dk\,, \tag{22}$$

and thus equation (18) can be fully specified using the PD nonlinear power spectrum.

3. THE LOGNORMAL PDFs AGAINST N-BODY SIMULATIONS

The previous section discussed a prescription to derive one-point and two-point lognormal PDFs assuming one-to-one mapping between the linear and nonlinear density fields. As remarked, however, the assumption does not seem to be justified in reality. So in this section we compare the lognormal PDFs extensively with the results of cosmological N-body simulations and discuss their validity and limitations. The analysis for the one-point PDF below significantly increases the range of δ compared with several previous works. As far as we know, the direct estimation of the two-point PDF in the nonlinear regime from simulations has not been performed before, and this is the first attempt.

3.1. *N-Body Simulations*

For the present analysis, we use a series of cosmological N-body simulations in three CDM models (SCDM, LCDM, and OCDM for standard, lambda, and open CDM models, respectively; Jing & Suto 1998) and four scale-free models with the initial power spectrum $P(k) \propto k^n$ ($n = 1$, 0, -1, and -2; Jing 1998). All the models employ $N = 256^3$ dark matter particles in a periodic comoving cube, $L_{\rm box}^3$, and are evolved using the P^{3}M code. The gravitational softening length is $L_{\rm box}/2560$ ($3L_{\rm box}/5120$) for the CDM (scale free) models and is kept fixed in the comoving length. The amplitude of the fluctuations in CDM models, σ_8, is normalized according to the cluster abundance (e.g., Kitayama & Suto 1997). The scale-free models assume an Einstein–de Sitter universe (density parameter $\Omega_0 = 1$, cosmological constant $\lambda_0 = 0$). The other parameters of the CDM models are summarized in Table 1.

The mass density fields are computed on 512^3 grids in the simulation box. First, we assign particles to each grid point using the cloud-in-cell interpolation. Then, we apply the smoothing kernel in the Fourier space and obtain the *smoothed* density fields after the inverse Fourier transform. Note that the density fields on those grids are not completely independent for $R > L_{\rm box}/512$, and we do heavy oversampling in this sense. Nevertheless, the error bars quoted in our results below are estimated from the variance among the three different realizations for each model (except the $n = -1$ model, which has only two realizations) and thus are free from the oversampling.

3.2. *The One-Point PDF*

Consider first the one-point PDFs in CDM models (Fig. 1). The PDFs are constructed by binning the data with $\Delta\delta = 0.1$, but we do not plot all the data points just for an illustrative purpose. We compute the density fields smoothed over Gaussian (*left panels*) and top-hat (*right panels*) windows with different smoothing lengths; $R = 2$, 6, and 18 h^{-1} Mpc are plotted in cyan, red, and green symbols, respectively, with error bars. The corresponding values of $\sigma_{\rm nl}$ are summarized in Table 2 and are also shown on each panel. Solid lines show the lognormal PDFs adopting $\sigma_{\rm nl}$ directly evaluated from simulations. The agreement between the lognormal model and the simulation results is quite impressive. A small deviation is noticeable only for $\delta \lesssim -0.5$.

We also show the lognormal PDFs in dashed lines, adopting $\sigma_{\rm nl}$ calculated from the nonlinear fitting formula of PD (values in parentheses in Table 2). Therefore, the predictions do not use the specific information of the current simulations and are completely independent in this sense. While these predictions are in good agreement with simulation data for $R \lesssim 6\ h^{-1}$ Mpc, the results for $R = 18\ h^{-1}$ Mpc are rather different. Actually, this discrepancy should be ascribed to the simulations themselves, not to the model predictions; Table 2 indicates that the $\sigma_{\rm nl}$ in the current simulations become systematically smaller than the PD predictions for larger R. This is because the simulations assume (incorrectly) no fluctuations beyond the scale of the simulation box size, $L_{\rm box}$. This constraint systematically reduce the fluctuations as the smoothing scale R approaches $L_{\rm box}$. We made sure that this is indeed the case by repeating the same analysis using the CDM simulations evolved in $L_{\rm box} = 300\ h^{-1}$ Mpc (Jing & Suto 1998); the variance of fluctuations at $R \le 18\ h^{-1}$ Mpc from the simulations agrees with the PD prediction to within 2% accuracy. Thus, we conclude that the lognormal PDF with the PD formula accurately reproduces the simulation results in the CDM models.

Next, turn to the scale-free models. Figure 2 shows plots corresponding to those in Figure 1, but for $n = 1$ to -2 models. In this figure, we compare the simulation data (*symbols with error bars*) with the lognormal PDF predictions (*solid lines*) adopting $\sigma_{\rm nl}$ from simulations (Table 3). Generally, their agreement is good also in these models. A closer look at Figure 2, however, reveals that the simulation results start to deviate from the lognormal predictions at both high- and low-density regions and that the deviation seems to systematically increase as n becomes larger. While this tendency is qualitatively consistent with the earlier claim by Bernardeau (1994) and Bernardeau & Kofman (1995) on the basis of the perturbation theory, our fully

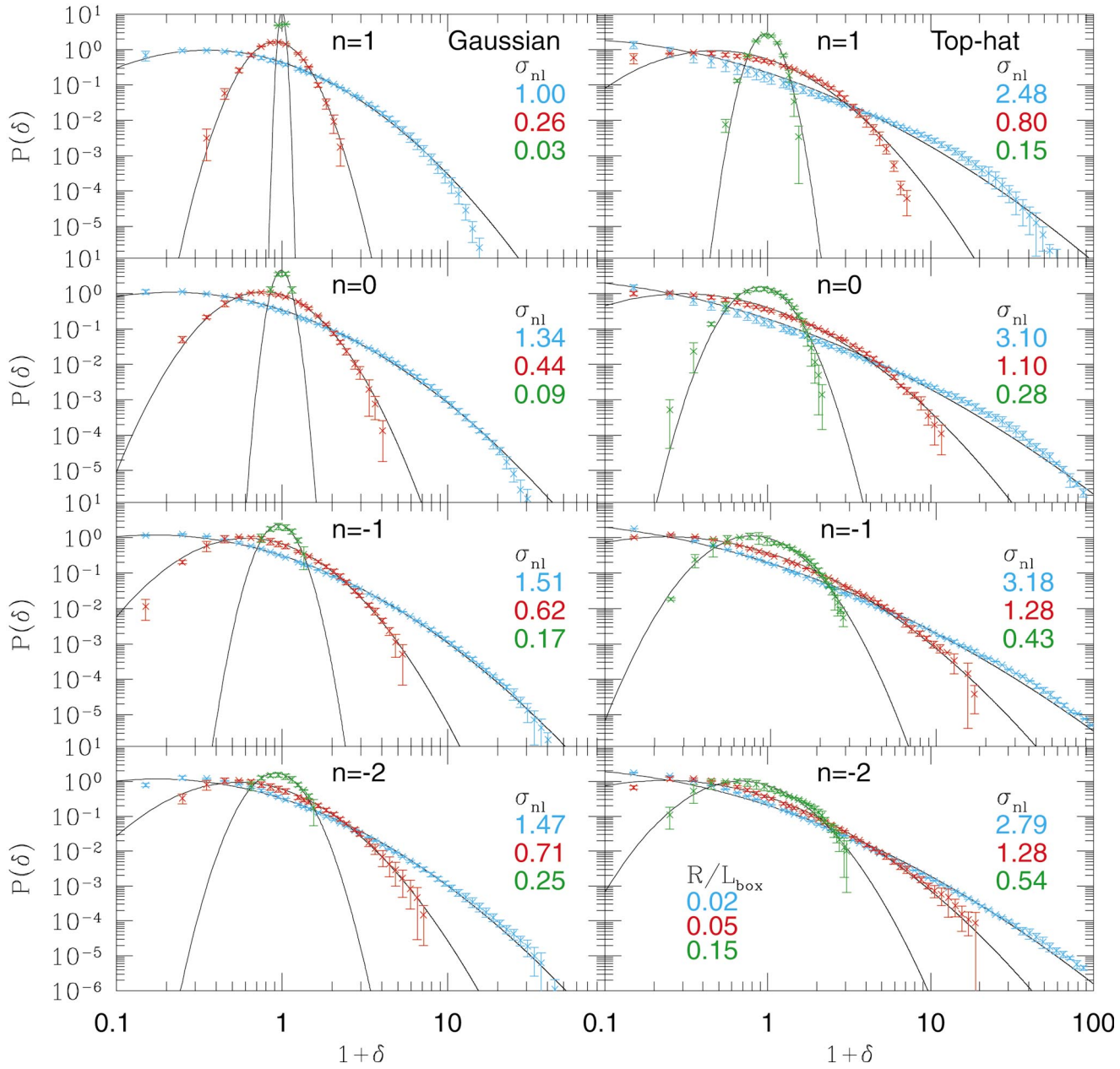

FIG. 2.—Same as Fig 1, but in the scale-free models ($n = 1$ to -2, *top to bottom*); $R = 0.02L_{\rm box}$ (*cyan*), $0.05L_{\rm box}$ (*red*), and $0.15L_{\rm box}$ (*green*)

nonlinear simulations show that the deviation from the lognormal PDF is not so large even in these scale-free models.

To examine the validity of the lognormal PDF more quantitatively, we compare the normalized skewness $S \equiv \langle\delta^3\rangle/\langle\delta^2\rangle^2$ and the normalized kurtosis $K \equiv (\langle\delta^4\rangle - 3\langle\delta^2\rangle^2)/\langle\delta^2\rangle^3$. The lognormal PDF predicts that

$$S(R) = 3 + \sigma_{\rm nl}^2(R)\,, \tag{23}$$

$$K(R) = 16 + 15\sigma_{\rm nl}^2(R) + 6\sigma_{\rm nl}^4(R) + \sigma_{\rm nl}^6(R)\,, \tag{24}$$

which are plotted in dotted lines in Figure 3 for six models: two LCDM models with $L_{\rm box} = 100$ and $300\ h^{-1}$ Mpc and four scale-free models with $n = 1, 0, -1$, and -2.

In practice, however, the density field δ in numerical simulations does not extend the entire range between -1 and ∞ but rather is limited as $\delta_{\rm min} < \delta < \delta_{\rm max}$ owing to the finite size of the simulation box. Thus, the nth-order moments of δ in simulations may be better related to

$$\langle\delta^n\rangle' = \int_{\delta_{\rm min}}^{\delta_{\rm max}} \delta^n P_{\rm LN}^{(1)}(\delta)d\delta\,. \tag{25}$$

TABLE 2

AMPLITUDE OF $\sigma_{\rm nl}(R)$ EVALUATED FROM THE CDM SIMULATIONS

Smoothing	R (h^{-1} Mpc)	SCDM	LCDM	OCDM
Top hat	2	2.33 (2.24)	4.17 (4.08)	4.37 (4.23)
	6	0.79 (0.77)	1.37 (1.40)	1.37 (1.38)
	18	0.23 (0.24)	0.44 (0.50)	0.43 (0.47)
Gaussian...	2	1.11 (1.08)	1.95 (1.96)	1.97 (1.96)
	6	0.36 (0.35)	0.64 (0.69)	0.63 (0.67)
	18	0.065 (0.090)	0.15 (0.22)	0.13 (0.21)

NOTE.—The values in parentheses are estimated from the nonlinear fitting formula of Peacock & Dodds 1996.

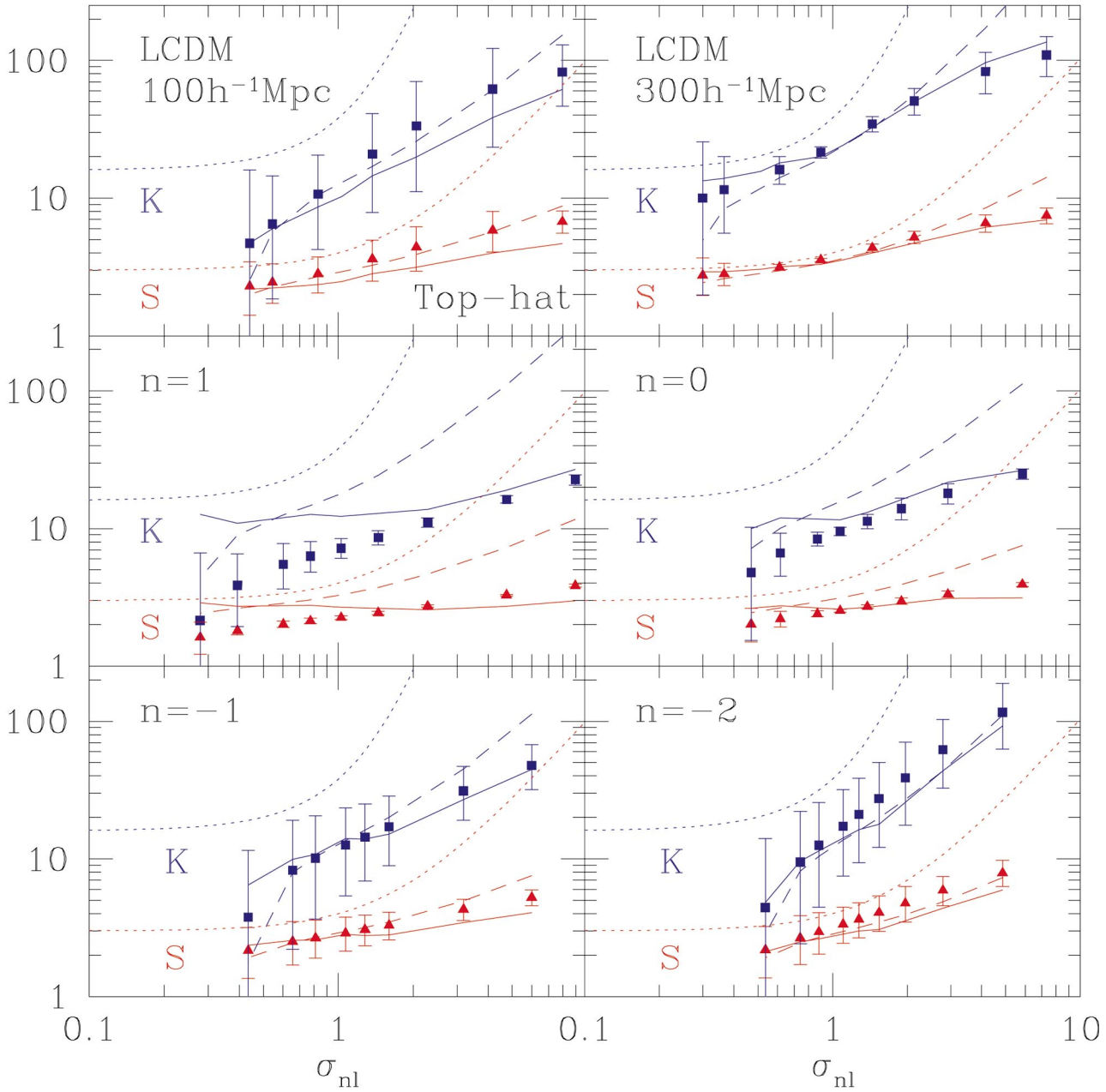

FIG. 3.—Normalized skewness S and normalized kurtosis K against σ_{nl} from simulations and the lognormal PDF predictions. Symbols represent the values estimated from simulations (quoted 1 σ error bars represent the scatter in the realizations). The meaning of predictions plotted in different lines is explained in the text. Top-hat smoothing is assumed.

The specific values for δ_{min} and δ_{max} may be roughly estimated from the condition that the expectation number of independent sampling spheres in the simulation box for $\delta < \delta_{min}$ or $\delta > \delta_{max}$ becomes unity:

$$\frac{L_{box}^3}{4\pi R^3/3} \int_{\delta_{max}}^{\infty} P_{LN}^{(1)}(\delta)d\delta = 1 , \quad (26)$$

$$\frac{L_{box}^3}{4\pi R^3/3} \int_{-1}^{\delta_{min}} P_{LN}^{(1)}(\delta)d\delta = 1 . \quad (27)$$

TABLE 3

THE RMS $\sigma_{nl}(R)$ IN THE SCALE-FREE SIMULATIONS

Smoothing	R (L_{box})	$n = 1$	$n = 0$	$n = -1$	$n = -2$
Top hat	0.02	2.48	3.10	3.18	2.79
	0.05	0.80	1.10	1.28	1.28
	0.15	0.15	0.28	0.43	0.54
Gaussian	0.02	1.00	1.34	1.51	1.47
	0.05	0.26	0.44	0.62	0.71
	0.15	0.03	0.09	0.17	0.25

Dashed lines in Figure 3 show the lognormal PDF predictions based on equations (25)–(27). The filled triangles and squares represent the measurement of S and K from the simulations, and the solid lines indicate the lognormal PDF predictions using equation (25) with the actual values for δ_{min} and δ_{max} in the simulations. Except for the $n = 1$ scale-free model, the predictions in solid lines reproduce the simulation data very well, which indicates that the lognormal PDF is in fact a good approximation. The relatively large

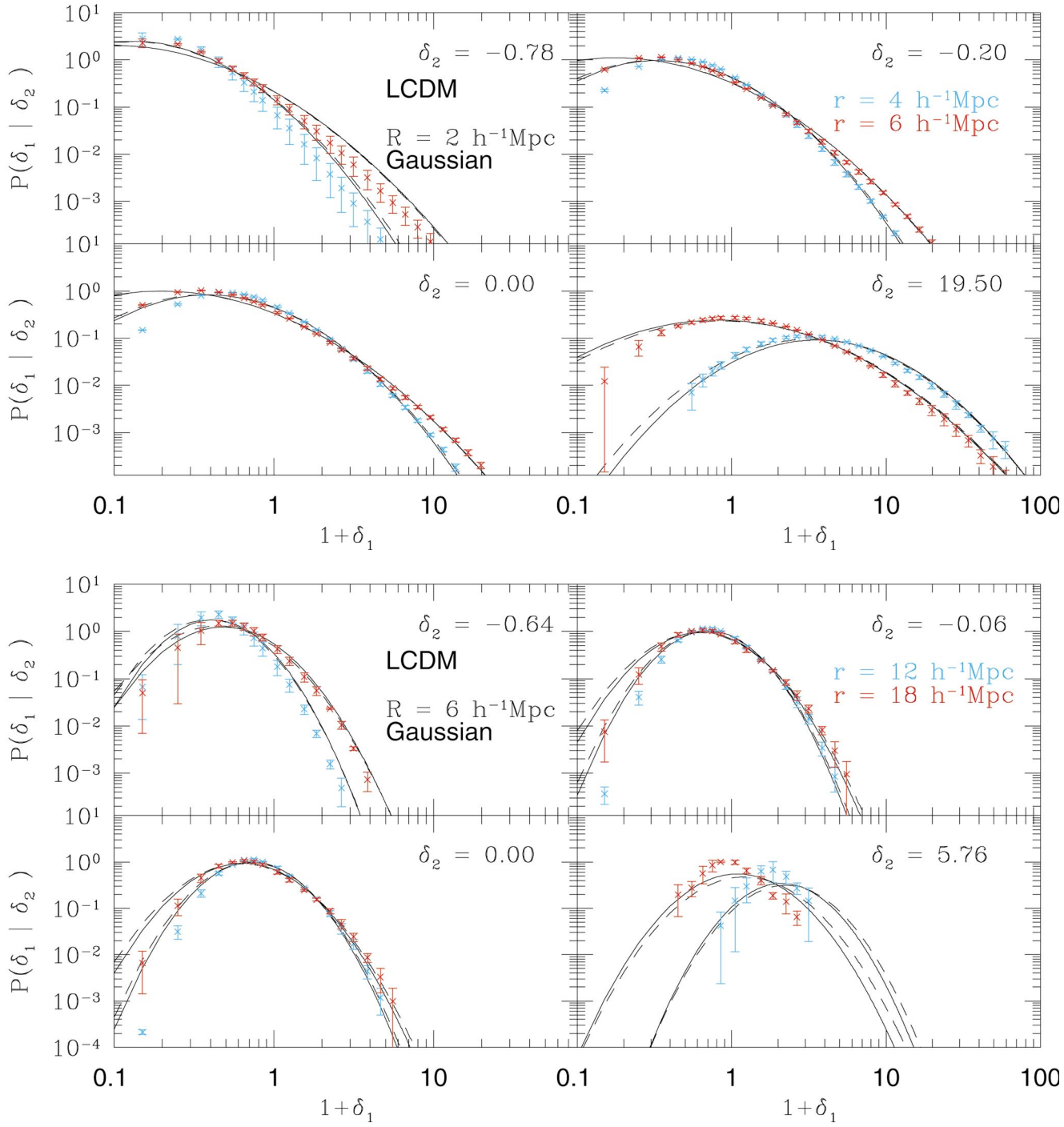

FIG. 4.—Two-point PDFs in the LCDM model with Gaussian smoothing over $R = 2\ h^{-1}$ Mpc (*upper panels*) and $R = 6\ h^{-1}$ Mpc (*lower panels*). The results at separation $r = 2R$ and $3R$ are plotted. *Solid lines*: The lognormal PDF adopting σ_{nl} and ξ_{nl} calculated directly from the simulations. *Long-dashed lines*: The lognormal PDF estimated from the nonlinear fitting formula of Peacock & Dodds (1996).

discrepancy between the lognormal prediction and the simulation in the $n = 1$ model is real since one can clearly recognize the systematic tendency with respect to n; models with smaller n, i.e., with substantial power on large scales, are better described by the lognormal PDF. This is consistent with the discussion by Bernardeau (1994).

Incidentally, both the current simulations and the lognormal PDF approximation confirmed the relatively strong scale dependence of S and K for $\sigma_{nl} > 1$, as pointed out earlier by Lahav et al. (1993) and Suto (1993). In fact, the degree of their scale dependence is also sensitive to the underlying power spectrum. Thus, the hierarchical ansatz for the higher order clustering is not valid in general.

In summary, we find that the one-point lognormal PDF remains a fairly accurate model for the cosmological density distribution even up to $\sigma_{nl} \sim 4$ and $\delta \sim 100$, fairly independently of the shape of the underlying power spectrum of density fluctuations. The range of validity turns out to be significantly broader than ranges from previous studies based on lower resolution simulations, $0.1 \lesssim \sigma \lesssim 0.6$ and $\delta \lesssim 4$ (Kofman et al. 1994) and $0.3 \lesssim \sigma \lesssim 1.5$ and $\delta \lesssim 9$ (Bernardeau & Kofman 1995), for instance.

3.3. *The Two-Point PDF*

While we would like to perform a similar comparison for

the two-point PDFs, it is a function of four variables—δ_1, δ_2, R, and r—and thus the comparison becomes rather complicated. Therefore, we decided to use the conditional two-point PDF for the purpose:

$$P^{(2)}(\delta_1 \mid \delta_2; r) \equiv \frac{P^{(2)}(\delta_1, \delta_2; r)}{P^{(1)}(\delta_2)} . \quad (28)$$

Since we already made sure that the one-point PDF is very accurately approximated by the lognormal, our task is to see if simulation results fit the conditional two-point lognormal PDF:

$$P^{(2)}_{\rm LN}(\delta_1 \mid \delta_2; r) \equiv \left[\frac{S}{2\pi(S^2 - X^2)}\right]^{1/2} \times \exp\left[-\frac{(SL_1 - XL_2)^2}{2S(S^2 - X^2)}\right]\frac{1}{(1+\delta_1)} . \quad (29)$$

The evaluation of the conditional two-point PDFs from simulations is carried out as follows. From the smoothed density fields computed on the 512^3 grid points, we first select those grid points with $\delta_2 - \Delta\delta_2/2 < \delta < \delta_2 + \Delta\delta_2/2$. The bin size $\Delta\delta_2$ is adjusted for each value of δ_2 so that approximately 10^5 grid points satisfy the condition. Next we pick up all grid points separated at $r - \Delta r/2 \sim r + \Delta r/2$ from the above grids. The separation interval Δr is chosen so that $4\pi r^2 \Delta r (512/L_{\rm box})^3 \sim 10^3$. Finally, we compute the conditional two-point PDF with a constant bin size of 0.1 in δ_1.

Figure 4 plots the resulting PDFs for the LCDM model with the Gaussian smoothing window; the four upper panels show the PDFs for the separation $r = 4$ and 6 h^{-1} Mpc and the smoothing length $R = 2$ h^{-1} Mpc, while the four lower panels for $r = 12$ and 18 h^{-1} Mpc and $R = 6$ h^{-1} Mpc. Solid lines indicate the conditional lognormal PDFs adopting the values of $\sigma_{\rm nl}$ and $\xi_{\rm nl}$ from simulations, while dashed lines show those using the PD predictions (Tables 2, 3, 4, and 5). Clearly, the lognormal PDF is a reasonably good approximation. The deviation at $\delta_2 \lesssim -0.5$, on the other hand, seems real and may be an enhanced feature that we noted in the one-point PDF (Fig. 1).

TABLE 4

AMPLITUDE OF $\xi_{\rm nl}(r; R)$ EVALUATED FROM THE CDM SIMULATIONS WITH GAUSSIAN SMOOTHING

R (h^{-1} Mpc)	r (h^{-1} Mpc)	SCDM	LCDM	OCDM
2	4	0.68 (0.64)	2.15 (2.10)	2.05 (2.06)
2	6	0.36 (0.35)	1.10 (1.15)	1.06 (1.10)
6	12	0.058 (0.063)	0.21 (0.26)	0.20 (0.24)
6	18	0.021 (0.028)	0.10 (0.144)	0.087 (0.127)

NOTE.—The values in parentheses are estimated from the nonlinear fitting formula of Peacock & Dodds 1996.

TABLE 5

TWO-POINT CORRELATION $\xi_{\rm nl}(r; R)$ IN THE SCALE-FREE SIMULATIONS WITH GAUSSIAN SMOOTHING

R ($L_{\rm box}$)	r ($L_{\rm box}$)	$n = 1$	$n = 0$	$n = -1$	$n = -2$
0.02	0.04	0.40	0.82	1.29	1.35
0.02	0.06	0.13	0.36	0.67	0.83

Figure 5 indicates that the good agreement is achieved not only in the Gaussian smoothed LCDM model but can also be found in the other models and/or the top-hat smoothing. Figure 6 plots the conditional two-point PDFs at $\delta_2 \lesssim -0.7$ and $\delta_2 \gtrsim 10$ in the scale-free models, where the deviation from the lognormal becomes manifest. Considering the error bars estimated from the different realizations for each model, the deviation seems statistically real.

As in the case of the one-point PDF, we illustrate the validity of the two-point lognormal PDF using the moments. Specifically we evaluate $\langle(\delta_1\,\delta_2)^2\rangle$ and $\langle(\delta_1\,\delta_2)^3\rangle$ according to

$$\langle(\delta_1\,\delta_2)^n\rangle(r) \equiv \int\int_{\mathscr{C}(\delta_1,\,\delta_2)} (\delta_1\,\delta_2)^n P^{(2)}_{\rm LN}(\delta_1, \delta_2; r) d\delta_1\, d\delta_2 , \quad (30)$$

where we select the range of the integration as

$$\mathscr{C}(\delta_1, \delta_2) = [(\delta_1, \delta_2) \mid \delta_{\rm min} \le \delta_1 \le \delta_{\rm max},\ \delta_{\rm min} \le \delta_2 \le \delta_{\rm max}] , \quad (31)$$

from the values of $\delta_{\rm min}$ and $\delta_{\rm max}$ directly measured from each simulation model. The results are summarized in Table 6, which indicates again that the two-point lognormal PDF predictions reproduce the simulation data except for $n \ge 0$. Thus, we also conclude that the two-point lognormal model describes fairly well the PDF of cosmological fluctuations for most δ_1 and δ_2 regions of of interest; the small but finite deviations exist only in $\delta_1 \lesssim -0.5$ and/or $\delta_2 \lesssim -0.5$, and also in $\delta_1 \gtrsim 10$ and $\delta_2 \gtrsim 10$.

3.4. *Does the Lognormal Transformation Approximate the Gravitational Evolution of the Density Fluctuations?*

The agreement between the lognormal predictions and the simulation results might be interpreted as indirect evidence that the lognormal transformation equation (8) is a good approximation for the nonlinear gravitational growth of the cosmological density fluctuations, at least on average.

In order to see if this is really the case, we consider the relation of the smoothed density fields at the same comoving position but at different redshifts, z_1 and z_2. For this purpose, we use one realization from the LCDM model evolved in $L_{\rm box} = 300\ h^{-1}$ Mpc. If the lognormal transform-

TABLE 6

TWO-POINT MOMENTS: $\langle(\delta_1\,\delta_2)^2\rangle$ AND $\langle(\delta_1\,\delta_2)^3\rangle$

Model	$\langle(\delta_1\delta_2)^2\rangle$	$\langle(\delta_1\delta_2)^3\rangle$	$\delta_{\rm min}$	$\delta_{\rm max}$
LCDM	$(0.76 \pm 0.4) \times 10^3$	$(1.5 \pm 1.4) \times 10^6$	...	...
Lognormal...	0.64×10^3	1.1×10^6	-0.96	90
LCDM300 ...	$(1.4 \pm 0.2) \times 10^3$	$(4.9 \pm 1.5) \times 10^6$	...	...
Lognormal...	1.3×10^3	6.2×10^6	-0.98	165
$n = 1$	3.1 ± 0.1	73 ± 11	...	...
Lognormal...	4.0	140	-0.98	14
$n = 0$	24 ± 3	$(2.6 \pm 0.6) \times 10^3$	...	...
Lognormal...	30	5.6×10^3	-0.99	30
$n = -1$.....	88 ± 18	$(2.2 \pm 1.0) \times 10^4$	...	...
Lognormal...	102	3.9×10^4	-0.98	42
$n = -2$.....	120 ± 40	$(4.1 \pm 2.1) \times 10^4$	...	...
Lognormal...	120	5.3×10^4	-0.96	46

NOTE.—From simulations and the lognormal PDF predictions. The LCDM models adopt the Gaussian smoothing with $R = 2\ h^{-1}$ Mpc and the moments are evaluated at the pair separation of $r = 4\ h^{-1}$ Mpc. For the scale-free models, $R = 0.02 L_{\rm box}$ and $r = 0.04 L_{\rm box}$.

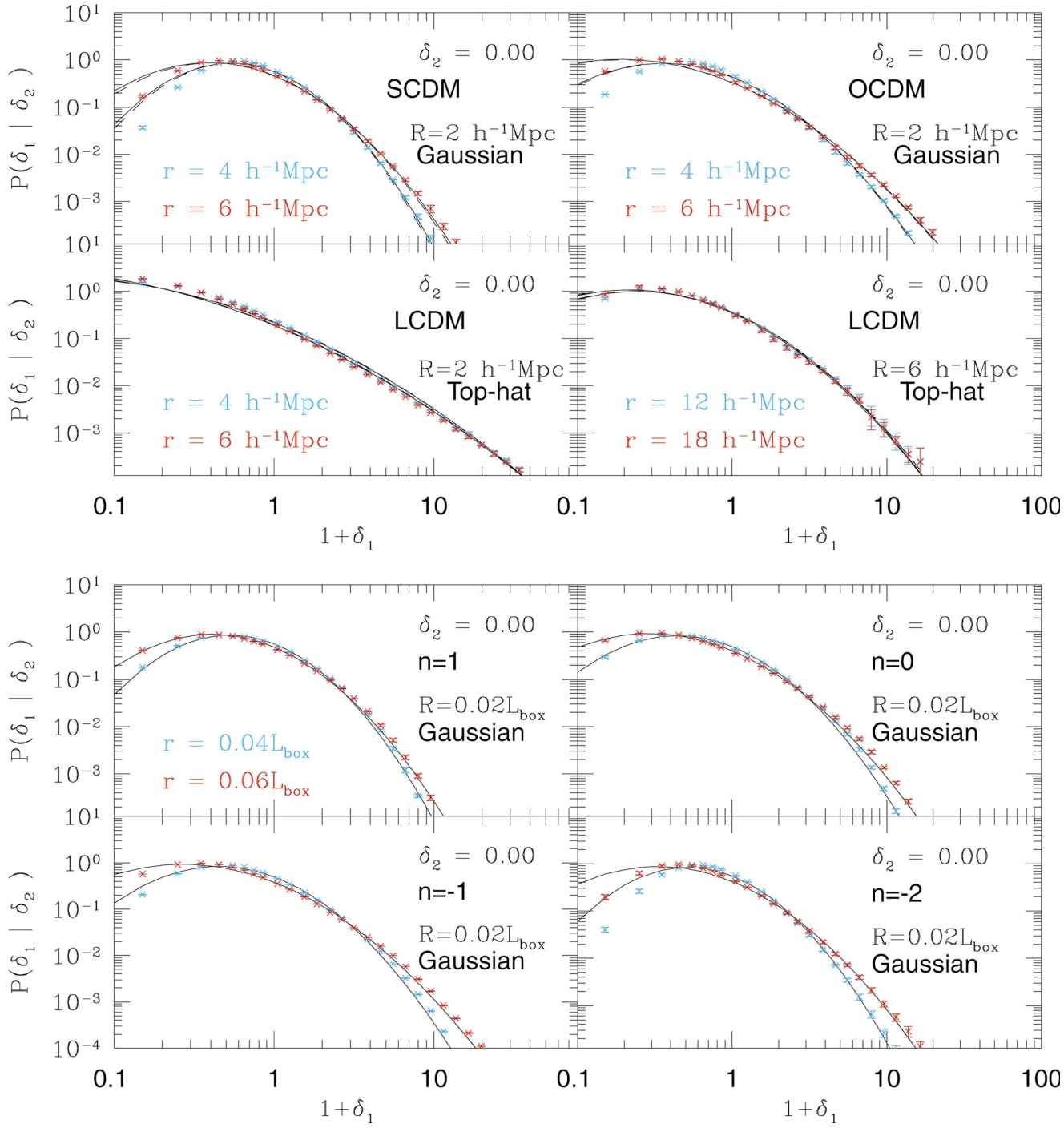

FIG. 5.—Two-point PDFs for different models and smoothing window functions. The upper panels plot the results in CDM models, the lower panels in scale-free models.

ation equation (8) is exact, the density fluctuations, $\delta(z_1)$ and $\delta(z_2)$, should satisfy

$$1+\delta(z_2)=\frac{1}{[1+\sigma_{\rm nl}^2(z_2)]^{1/2}}\times\exp\left(\frac{\{\ln[1+\sigma_{\rm nl}^2(z_2)]\}^{1/2}}{\{\ln[1+\sigma_{\rm nl}^2(z_1)]\}^{1/2}}\times\ln\{[1+\sigma_{\rm nl}^2(z_1)]^{1/2}[1+\delta(z_1)]\}\right). \quad (32)$$

Figure 7 plots the color contour of the joint probability $P[\delta(z), \delta(z=9)]$ of densities at $z=2.2$, 1.0, and 0 against that at $z=9$ on the same grid points in the LCDM model. We adopt Gaussian smoothing with $R=6\ h^{-1}$ Mpc (*left panels*) and $2\ h^{-1}$ Mpc (*right panels*). The solid lines in white and magenta represent the lognormal transformation equation (32) and the conditional mean from simulations for a fixed $\delta(z=9)$. The lognormal transformation traces the mean relation of simulations to some extent only when the nonlinearity is weak (see higher z and larger R cases). On the other hand, in the nonlinear region, the transformation equation (32) starts to deviate from the mean relation of the simulations significantly, and the distribution around the mean relation becomes broad. A similar tendency was

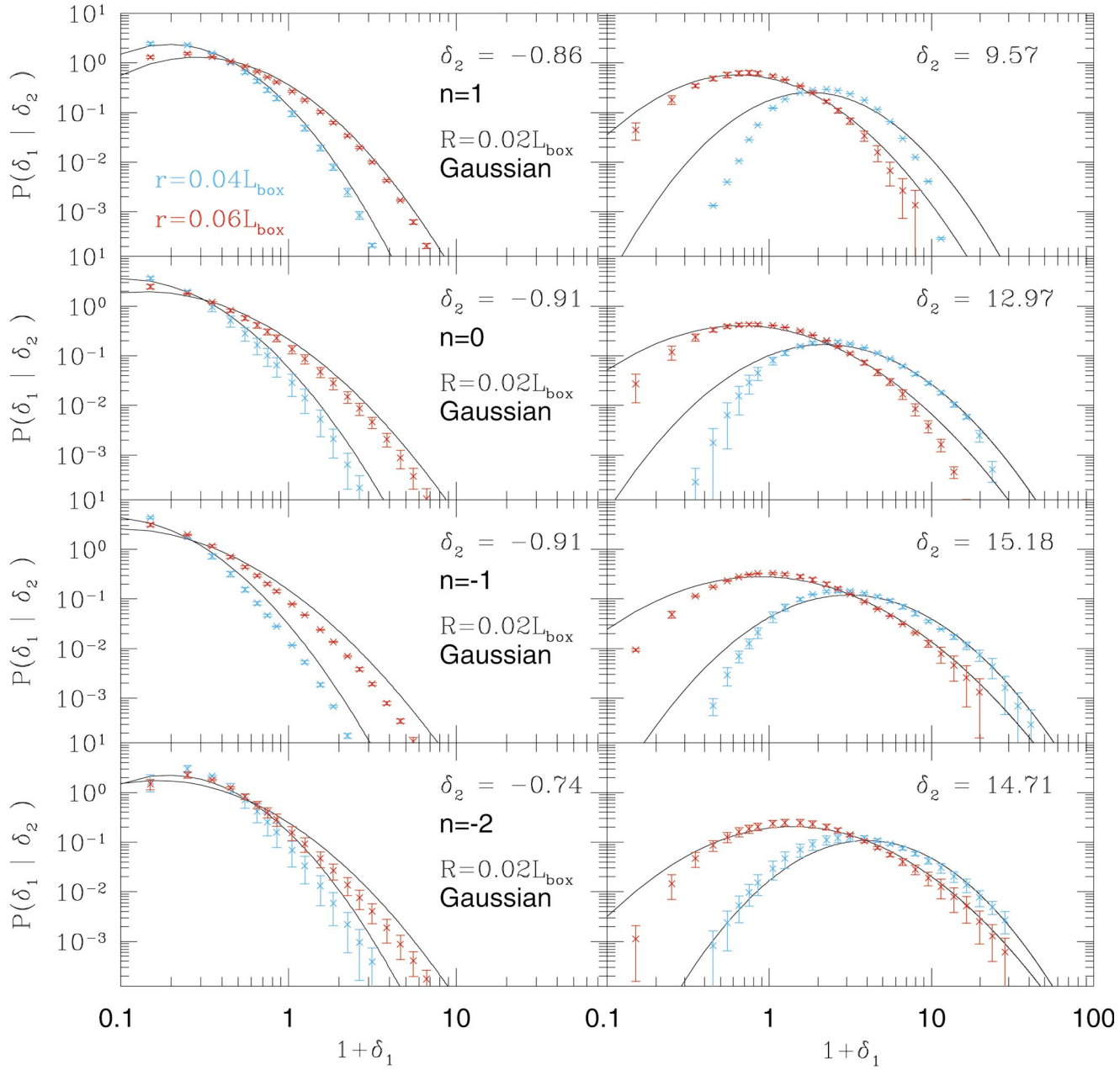

FIG. 6.—Two-point PDFs for the scale-free models at negative (*left panels*) and positive (*right panels*) tails of the distribution of δ_2

found in a somewhat different analysis by Coles, Melott, & Shandarin (1993). In a sense, this is a physically natural and expected result, but then it makes it even more difficult to account for the good agreement between the lognormal and simulation PDFs in those scales.

To better understand the distribution of the linear and its evolved density fields, we compute the conditional probability $P[\delta(z) \mid \delta(z = 9)]$, i.e., the slice of Figure 7 at a given $\delta(z = 9)$. The results are plotted in Figure 8 and exhibit some regularity in the distribution. The peak positions seem to show some scaling with respect to the value of $\delta(z = 9)$, and also the tail of the distribution asymptotically approaches a single power law. While we do not yet fully understand the behavior, this regularity in the distribution function may be useful in explaining our findings that the one-point and two-point lognormal PDFs work well empirically.

4. CONCLUSIONS AND DISCUSSION

In the present paper we have estimated the probability distribution functions of cosmological density fluctuations from the high-resolution N-body simulations with a Gaussian initial condition. In particular, we have critically examined the validity of the lognormal models for the one- and two-point PDFs in both weakly and strongly nonlinear regimes.

We have shown that the one-point lognormal PDF is a fairly accurate model not only in weakly nonlinear regimes as claimed previously but also in more strongly nonlinear regimes, even up to $\sigma_{nl} \sim 4$ and $\delta \sim 100$. Furthermore, we extended the analysis to the two-point PDF and found that the lognormal PDF serves also as an empirically accurate model for the range of densities of interest. This is the case fairly independently of the shape of the underlying power

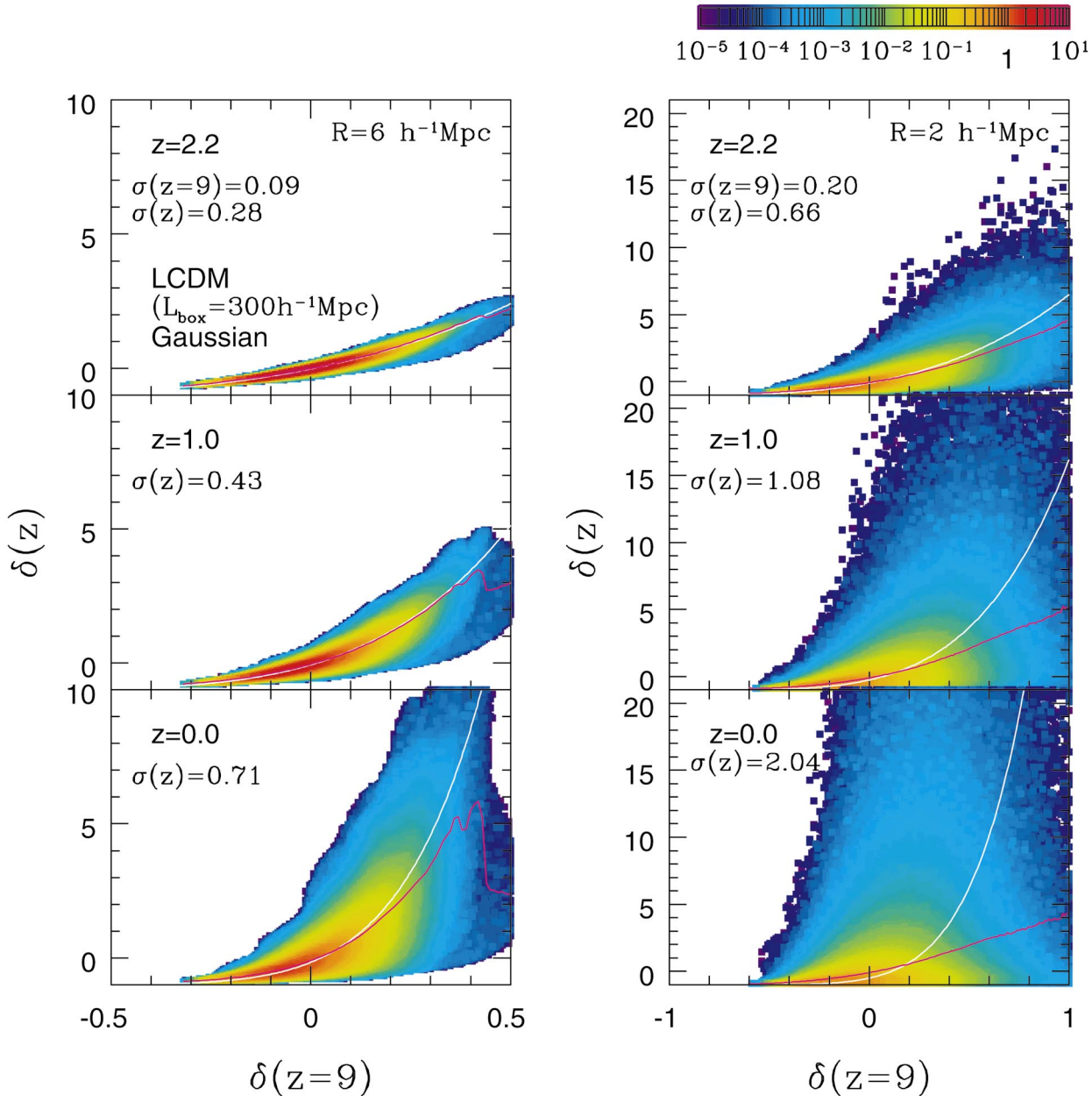

FIG. 7.—Contour plots of the joint probability $P[\delta(z)\,|\,\delta(z=9)]$ in the LCDM model ($L_{\rm box} = 300\ h^{-1}$ Mpc) with Gaussian smoothing window; $R = 6\ h^{-1}$ Mpc (*left panels*) and $R = 2\ h^{-1}$ Mpc (*right panels*). Top, middle, and bottom panels correspond to correlations of $\delta(z = 2.2)$, $\delta(z = 1)$, and $\delta(z = 0)$ against $\delta(z = 9)$, respectively. White lines represent the lognormal transformation equation (32); magenta lines are the conditional mean at a fixed $\delta(z = 9)$.

spectrum of density fluctuations, although models with large power on small scales (e.g., $n \geq 0$ scale-free models in our examples) seem to show a small deviation from the lognormal prediction at the tails of the distribution, especially for $\delta \lesssim -0.5$. In particular, the lognormal PDF reproduces very well the skewness and kurtosis measured from the simulation data when the finite size of the simulation volume is properly taken into account.

The degree of agreement of the lognormal models that we have shown is amazing considering the fact that the underlying mapping between the initial and the evolved density fields differs significantly from the simulation results even in an averaged sense. We have explicitly shown the probability distribution of the initial and evolved density fields from simulations, although we were not able to provide a physical explanation for the origin of the lognormal PDF. This should be left as our future work, and we would like to come back to it in a future paper. For this purpose, other theoretical approaches based on perturbation theory (Bernardeau 1992, 1994) and the spherical collapse model (Fosalba & Gaztañaga 1998) may be helpful.

Nevertheless, our present work provides an empirical justification for the use of the lognormal PDF in a variety of theoretical model predictions. For instance, Matsubara & Yokoyama (1996) proposed to evaluate the effect of the nonlinear gravitational evolution on the genus statistics using the lognormal mapping. Taruya & Suto (2000) constructed an analytical model for halo biasing on the basis of the one-point lognormal PDF of underlying mass density field. Hikage, Taruya, & Suto (2001) applied this biasing model in their predictions of the genus for clusters of galaxies.

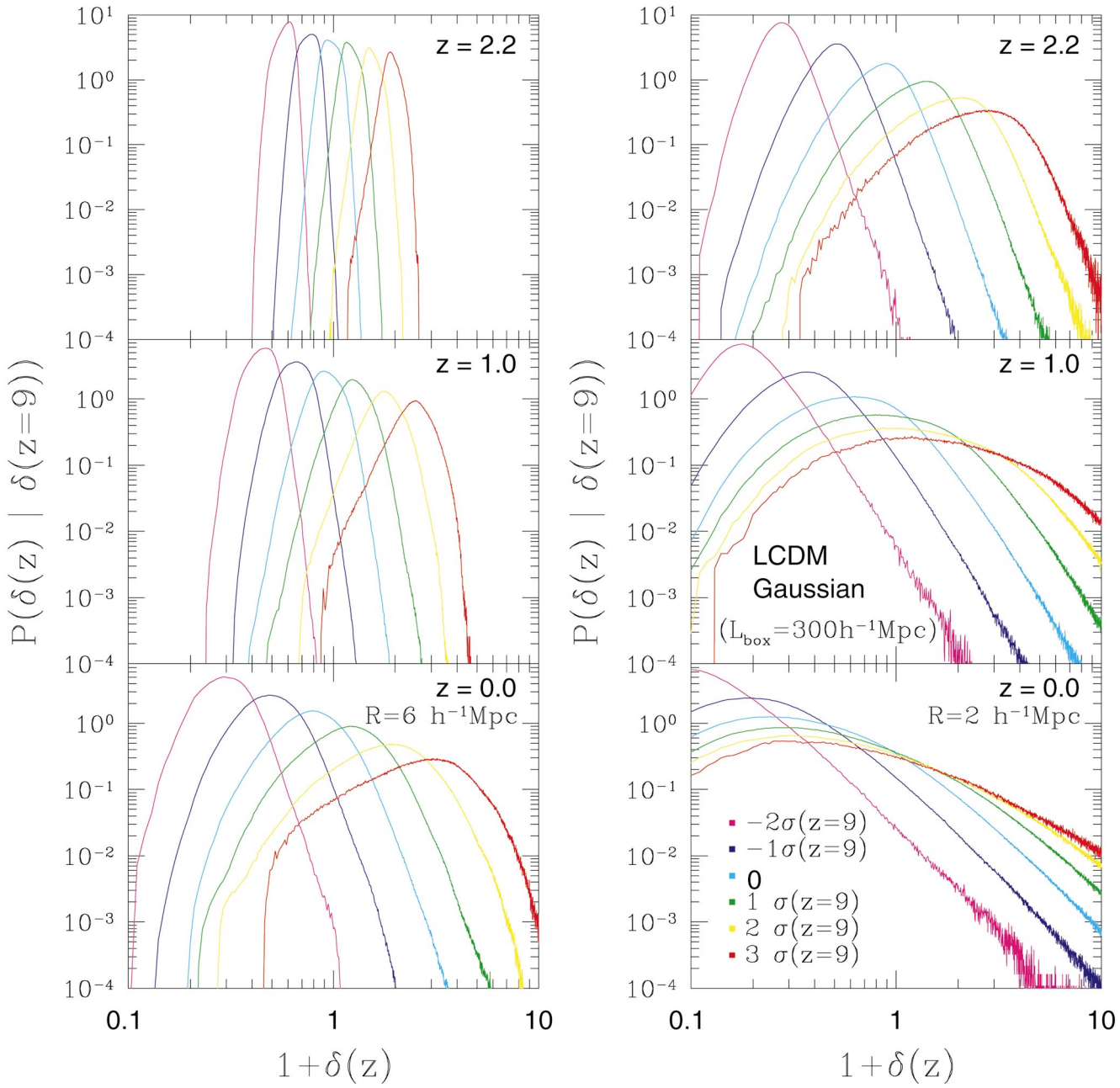

FIG. 8.—Conditional probability $P[\delta(z)|\delta(z=9)]$ for a fixed $\delta(z=9)$ corresponding to each panel of Fig. 7. In each panel, results for $\delta(z=9) = -2, -1, 0, 1, 2$, and 3 times the $\sigma(z=9)$, the rms of $\delta(z=9)$, are plotted.

Finally, the present results might be useful in considering the prediction of weak lensing statistics (Valageas 2000; Munshi & Jain 2000). To construct a model for PDF in redshift space is another important topic (e.g., Watts & Taylor 2001; Hui, Kofman, & Shandarin 2000), which is relevant in discussing Lyα forests (Gaztañaga & Croft 1999).

We thank Y. P. Jing for kindly providing us his N-body data and an anonymous referee for constructive comments. I. K. and A. T. gratefully acknowledge support from the Takenaka-Ikueikai fellowship and a Japan Society for the Promotion of Science fellowship, respectively. Numerical computations were carried out at RESCEU (Research Center for the Early Universe, University of Tokyo), ADAC (the Astronomical Data Analysis Center) of the National Astronomical Observatory, Japan, and at KEK (High Energy Accelerator Research Organization, Japan). This research was supported in part by the grant-in-aid from the Ministry of Education, Science, Sports, and Culture of Japan (07CE2002, 12640231) and by the Supercomputer project 00-63 of KEK.

REFERENCES

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Bernardeau, F. 1992, ApJ, 392, 1
———. 1994, A&A, 291, 697
———. 1996, A&A, 312, 11
Bernardeau, F. & Kofman, L. 1995, ApJ, 443, 479
Bouchet, F., Strauss, M. A., Davis, M., Fisher, K. B., Yahil, A., & Huchra, J. P. 1993, ApJ, 417, 36
Coles, P., & Jones, B. 1991, MNRAS, 248, 1
Coles, P., Melott, A. L., & Shandarin, S. F. 1993, MNRAS, 260, 765
Colombi, S., Bouchet, F. R., & Schaeffer, R. 1995, ApJS, 96, 401
Fosalba, P., & Gaztañaga, E. 1998, MNRAS, 301, 503
Gaztañaga, E., & Croft, R. A. C. 1999, MNRAS, 309, 885
Gaztañaga, E., & Yokoyama, J. 1993, ApJ, 403, 450
Hamilton, A. J. S. 1985, ApJ, 292, L35
Hikage, C., Taruya, A., & Suto, Y. 2001, ApJ, 556, 641
Hubble, E. P. 1934, ApJ, 79, 8
Hui, L., Kofman, L., & Shandarin, S. 2000, ApJ, 537, 12
Jing, Y. P. 1998, ApJ, 503, L9
Jing, Y. P., & Suto, Y. 1998, ApJ, 494, L5
Kitayama, T., & Suto, Y. 1997, ApJ, 490, 557
Kofman, L., Bertschinger, E., Gelb, J. M., Nusser, A., & Dekel, A. 1994, ApJ, 420, 44
Lahav, O., Itoh, M., Inagaki, S., & Suto, Y. 1993, ApJ, 402, 387
Matsubara, T., & Yokoyama, J. 1996, ApJ, 463, 409
Munshi, D., & Jain, B. 2000, MNRAS, 318, 109
Peacock, J. A., & Dodds, S. J. 1996, MNRAS, 280, L19 (PD)
Saslaw, W. C. 1985, Gravitational Physics of Stellar and Galactic Systems (Cambridge: Cambridge Univ. Press)
Suginohara, T., & Suto, Y. 1991, PASJ, 43, L17
Suto, Y. 1993, Prog. Theor. Phys., 90, 1173
Suto, Y., Itoh, M., & Inagaki, S. 1990, ApJ, 350, 492
Szapudi, I., & Colombi, S. 1996, ApJ, 470, 131
Taruya, A., & Suto, Y. 2000, ApJ, 542, 559
Taylor, A. N., & Watts, P. I. R. 2000, MNRAS, 314, 92
Ueda, H., & Yokoyama, J. 1996, MNRAS, 280, 754
Valageas, P. 2000, A&A, 356, 771
Watts, P. I. R., & Taylor, A. N. 2001, MNRAS, 320, 139