NON-GAUSSIANITY AND THE RECOVERY OF THE MASS POWER SPECTRUM FROM THE Lyα FOREST

LONG-LONG FENG
Center for Astrophysics, University of Science and Technology of China, Hefei, Anhui 230026, P.R. China; fengll@physics.arizona.edu

AND

LI-ZHI FANG
Department of Physics, University of Arizona, Tucson, AZ 85721; fanglz@physics.arizona.edu

Received 1999 September 9; accepted 2000 January 21

ABSTRACT

We investigate the effect of non-Gaussianity on the reconstruction of the initial mass field from the Lyα forest. We show that the transmitted flux of QSO absorption spectra are highly non-Gaussian in terms of the statistics, the kurtosis spectrum, and the scale-scale correlation. These non-Gaussianities cannot be completely removed by the conventional algorithm of Gaussianization, and the scale-scale correlations are largely retained in the mass field recovered by the Gaussian mapping. Therefore, the mass power spectrum recovered by the conventional algorithm is systematically lower than the initial mass spectrum on scales at which the local scale-scale correlation is substantial. To reduce the non-Gaussian contamination, we present two methods. The first is to perform the Gaussianization scale-by-scale using the discrete wavelet transform (DWT) decomposition. We show that the non-Gaussian features of the Lyα forest basically will no longer exist in the scale-by-scale Gaussianized mass field. The second method is to choose a proper orthonormal basis (representation) to suppress the effect of the non-Gaussian correlations. In the quasi-linear regime of cosmic structure formation, the DWT power spectrum is efficient for suppressing the non-Gaussian contamination. These two methods significantly improve the recovery of the mass power spectrum from the Lyα forest.

Subject headings: cosmology: theory — dark matter — large-scale structure of universe — quasars: absorption lines

1. INTRODUCTION

Since high-resolution QSO spectra became available, the transmitted flux in QSO spectra, or the Lyα forest, offers an unprecedented opportunity to study the large-scale structure of the universe and its evolution at redshifts beyond the galaxy redshift catalog (e.g., Bi & Davidsen 1997, and references therein). A basic goal of this study is to reconstruct the initial mass field. Assuming that these objects trace the underlying matter field in some way, it should be possible to trace the evolution of the mass field back in time. Because the initial mass field is expected to be Gaussian in many models of the origin of fluctuations, reconstructing the initial mass fluctuations is synonymous with recovering the mass power spectrum (e.g., Croft et al. 1999.)

The recovery of power spectra from the Lyα forests relies on two theoretical conjectures. The first is to assume that the transmitted flux of a QSO Lyα absorption spectrum is a point-to-point tracer of the underlying dark matter distribution. The Lyα forest has been successfully modeled by absorption of the ionized intergalactic gas, the distribution of which is continuous, and locally determined by the underlying dark matter distribution (Bi 1993; Fang et al. 1993; Bi, Ge, & Fang 1995; Hernquist et al. 1996; Bi & Davidsen 1997; Hui, Gnedin, & Zhang 1997). Thus, the transmitted flux of a QSO absorption spectrum at a given redshift depends only on the mass density of dark matter at the position corresponding to the redshift.

The second assumption is that the initial mass field can be recovered from the flux of the QSO spectrum by the Gaussianization algorithm (Weinberg 1992; Croft et al. 1998). With this method, the shape of the one-dimensional initial Gaussian density field with an arbitrary normal-

ization can be recovered approximately from the observed flux by a point-to-point Gaussian mapping, if the relation between flux and mass density is monotonic, i.e., the higher the underlying mass density, the stronger the Lyα absorption. The monotonicity would be a good approximation in the weak nonlinear or quasilinear evolutionary regime of the cosmic clustering.

This paper studies the influence of the non-Gaussianity of the Lyα forest on the recovery of the initial power spectrum. It is motivated by the recent systematic detection of non-Gaussianity in the Lyα forests. Despite the well-known fact that the two-point correlation function of the Lyα absorption lines is quite weak, the distribution of these lines does show non-Gaussian behavior. For instance, it was pointed out about 10 yr ago that the distribution of the nearest neighbor Lyα line intervals is different from a Poisson process (Duncan, Ostriker, & Bajtlik 1989; Liu & Jones 1990; Fang 1991). Recently, the detection of the spectrum of higher order cumulants (Pando & Fang 1998a) and the scale-scale correlations (Pando et al. 1998a) of the Lyα forests implies systematic non-Gaussianity on scales as large as about 10 h⁻¹ Mpc. The abundance of the Lyα line “clusters” identified with respect to the richness is also found to be significantly different from a Gaussian process (Pando & Fang 1996).

According to the philosophy of the Gaussianization reconstruction, all the non-Gaussian features of the Lyα forests are not initial. They should be removed by the Gaussianization of the flux of QSO spectra. The recovered mass field should be Gaussian. The algorithm of Gaussian mapping is designed to remove the non-Gaussianities of the flux, and recover a Gaussian mass field.
The idea of Gaussianization is exquisite. However, we will show that even though the current algorithm of Gaussianization maps does map the distribution of the flux value onto a Gaussian probability distribution function (PDF), the above-mentioned non-Gaussianities of the Lyα forests still remain largely in the Gaussianized flux. In other words, the recovered field is not linear and Gaussian, but contaminated by the non-Gaussian behavior of the Lyα forest.

It has been recognized that the estimation of the power spectrum is significantly affected by non-Gaussian behavior, such as the correlation between band-averaged power spectra, which is essentially the scale-scale correlation (e.g., Meiksin & White 1999). Therefore, with the current algorithm of Gaussianization, the recovered power spectrum is distorted by the non-Gaussianity of the Lyα forests. The question is then raised of how to improve the algorithm of Gaussianization in order to recover a mass field exempt from the non-Gaussianity of the Lyα forest, or how to suppress the effect of the non-Gaussian contamination in the estimation of the power spectrum. We investigate these problems in this paper.

This paper is organized as follows. In § 2, using popular cold dark matter (CDM) models, we present the non-Gaussian features in the transmitted flux of the Lyα forests. In § 3, we demonstrate that the non-Gaussianity of the mass field recovered by the conventional Gaussianization algorithm is about the same order as the original non-Gaussianity. Two alternatives that yield better Gaussianization are then proposed. In § 4, the distortion of a power spectrum by the non-Gaussianity is shown, and a possible way of suppressing the non-Gaussian effect on the power spectrum detection, i.e., by properly choosing the representation of the power spectrum, is suggested. We conclude the paper with a discussion of our findings in § 5.

2. THE NON-GAUSSIAN FEATURES OF THE Lyα FORESTS

2.1. Samples of the Lyα Forests

To investigate the recovery method for the mass power spectrum, we generate the simulation samples of the Lyα forest in the semianalytic model of the intergalactic medium (IGM) developed by Bi and coworkers (Bi 1993; Fang et al. 1993; Bi et al. 1995; Bi & Davidsen 1997). This model can approximately fit most observed features of the Lyα forest, including the column density distribution and the number density of the Lyα forest lines, the distribution of the equivalent widths and their redshift dependence, the clustering, and the Gunn-Peterson effect. Moreover, in this model, the relations among the dark matter field, the flux of the Lyα absorption, and the power spectrum of reconstructed initial mass fields are under control. It should be very useful in revealing the problems of the reconstruction.

The model is described in detail in the references listed above. We now give a brief account of particularly the fundamental physics underlying in this model. The basic assumption of the model is that the density distribution of the baryonic diffuse matter in the universe, \( n(x) \), is determined by the underlying dark matter density distribution via a lognormal relation as

\[
  n(x) = n_0 \exp \left[ \delta_0(x) - \frac{\langle \delta^2 \rangle}{2} \right],
\]

where \( n_0 \) is the mean number density, and \( \delta_0(x) \) is a Gaussian random field derived from the density contrast, \( \delta_{DM} \), of the dark matter by

\[
  \delta_0(x) = \frac{1}{4\pi x_0^2} \int \frac{\delta_{DM}(x_1)}{|x - x_1|} e^{-|x - x_1|/x_0} dx_1
\]

in the comoving space, or

\[
  \delta_0(k) = \frac{\delta_{DM}(k)}{1 + x_0^2 k^2}
\]

in the Fourier space. To take into account the effect of redshift distortion, the peculiar-velocity field along the line of sight is also calculated by the simulation model (Bi 1993; Fang et al. 1993; Bi & Davidsen 1997).

The Gaussian field, \( \delta_{DM} \), is produced in a CDM model. To account for the baryonic effect on the transfer function, we adopt the fitting formula for power spectra presented by Eisenstein & Hu (1999). Because the goal of this paper is mainly to examine the recovery method of the power spectrum, we will not take into account the variants of the CDM family, but use only the standard one, i.e., the flat model (\( \Omega_0 = 1.0 \)) normalized by the 4 yr COBE data, and \( \Gamma = \Omega_0 \) is taken to be 0.3, where \( h \) denotes the normalized Hubble parameter, and \( \Omega_0 \) is the cosmological density parameter of total mass. This model is compatible with the galaxy correlation observed on scales of \( \sim 10 h^{-1} \text{ Mpc} \) (Efstathiou, Bond, & White 1992). The baryonic fraction in the total mass was fixed by the constraint from the primordial nucleosynthesis of \( \Omega_0 = 0.0125 h^{-2} \) (Walker et al. 1991).

The factor \( x_b \) in equation (2) is the Jeans length of the IGM, given by

\[
  x_b \equiv \left[ \frac{2\pi k T_m}{3\mu m_p \Omega(1 + z)} \right]^{1/2},
\]

where \( T_m \) and \( \mu \) are the density-average temperature and molecular weight of the IGM, respectively, and \( \gamma \) is the ratio of specific heats. The thermal equation of state of the IGM is assumed to be polytropic, \( T \propto n^{\gamma-1} \), with \( \gamma = 4/3 \).

The lognormal relation, equation (1), has the following properties: (1) when fluctuations are small, i.e., \( (n/n_0 - 1) \approx \delta_0 \), equation (1) is just the expected linear evolution of the IGM; and (2) on small scales as \( |x - x_0| \ll x_0 \), equation (1) becomes the well-known isothermal hydrostatic solution, which describes highly clumped structures such as intracluster gas, \( n \propto \exp \left( -\frac{\mu m_p}{\psi_{DM} \gamma kT} \right) \), where \( \psi_{DM} \) is the dark matter potential (Sarazin & Bahcall 1977).

The absorption optical depth at the observed wavelength \( \lambda \) is

\[
  \tau(\lambda) = \int_{t_0}^{t_{QSO}} \sigma \left( \frac{c}{\lambda_z} \frac{1 + z}{1 + z_0} \right) n_{HI}(t) dt,
\]

where \( z_0 = (\lambda/\lambda_z) - 1 \), \( t_0 \) denotes the present time, \( t_{QSO} \) is the time corresponding to the redshift of the QSO, \( z_{QSO} \), and so stands for the relation between \( t \) and \( z \); \( \sigma \) is the absorption cross section at the Lyα transition, and \( \lambda_z = 1216 \text{ Å} \) represents the Lyα wavelength. The density of the neutral hydrogen atoms, \( n_{HI} \), can be found from \( n \) by the cosmic abundance of hydrogen, and photoionization equilibrium (Bi et al. 1995).

Obviously, in this model, the relation between the transmitted flux, \( F(\lambda) = e^{-\tau} \), and \( n(x) \) or \( \delta_0(x) \) is basically local. The nonlocality is only caused by the width of the absorption cross section \( \sigma \) and the peculiar velocity of the neutral hydrogen. Therefore, \( F \) is approximately a point-to-point
from one independent measure of the flux field. The ensemble. In other words, when the fair-sample hypothesis
at a given index, information on the flux.
Since the DWT bases are complete, the WFCs contain all
basis is used.
particular choice as long as a compactly supported wavelet
functions are localized in both physical space and the Fourier (scale)
localized in both physical space and the Fourier (scale)
and complete set of the DWT basis (for details of the DWT,
where
point distribution of the WFCs at a given scale
The distribution of represents approximately the one-
tracer of the mass fluctuation, Moreover, the
of the Ly4 forest lines (Pando & Fang 1998a). The skewness and kurtosis spectra of the transmit-
ted flux in 100 simulated samples are shown in Figures 1 and 2, respectively. To assess the statistical significance, the
95% confidence range from 100 realizations of Gaussian
noise are shown by the bars and gray band, respectively. The skewness
Gaussian distribution. For this purpose, we calculate the
cumulant moments defined by
\begin{align}
I_j^3 &= M_j^3, \\
I_j^4 &= M_j^4 - 3M_j^2M_j^2, \\
I_j^5 &= M_j^5 - 10M_j^3M_j^2.
\end{align}
where
\begin{equation}
M_n^j \equiv \frac{1}{2^n} \sum_{l=0}^{2^n-1} (\bar{F}_{l,j} - \bar{F}_{l,j})^n.
\end{equation}
The second-order cumulant moment gives the DWT
power spectrum (§ 4; see also Pando & Fang 1998b). For Gaussian fields, all the cumulant moments higher than
order 2 are zero. Thus, one can measure the non-
Gaussianity by \(I_4^2\) with \(n > 2\). We call \(I_4^2\) the DWT spectrum of nth cumulant. The cumulant measures \(I_4^2\) and \(I_4^4\) are related to the well-known skewness and kurtosis, respectively, defined by
\begin{align}
S_j &\equiv \frac{1}{(I_j^2)^{3/2}} I_j^3, \\
K_j &\equiv \frac{1}{(I_j^2)^2} I_j^4.
\end{align}
Using the skewness and kurtosis spectra as statistical
indicators, a significant non-Gaussian behavior has been
found in the distribution of Ly4 forest lines (Pando & Fang 1998a). The skewness and kurtosis spectra of the transmitted
flux in 100 simulated samples are shown in Figures 1 and 2, respectively. To assess the statistical significance, the
95% confidence range from 100 realizations of Gaussian
noise are also displayed in these figures. Clearly, the kurtosis spectrum of simulated \(F\) shows differences from the
Gaussian noise spectra on scales \(j \geq 8\) (or \(\leq 1.5 h^{-1} \text{Mpc}\))

![Fig. 1.—Skewness spectrum of the Ly4 forests. The 95% confidence ranges of the skewness spectrum of the simulated samples and Gaussian noise are shown by the bars and gray band, respectively. The skewness spectrum of the Keck data of HS 1700 + 64 is shown by squares connected by the solid line. The physical scale related to \(j\) is 189.8 × 2ependent measures of the \(F\) field. The corresponding simulation size in the CDM model is 189.84 \(h^{-1}\) Mpc in comoving space, which is long enough to incorporate most of the fluctuation power. The selection of this redshift range is in order to compare the simulation with the Keck spectrum of HS 1700 + 64. The spectrum of HS 1700 + 64 ranges from 3723.012 to 5523.554 Å, with a resolution of 3 km s\(^{-1}\), or a total of 55,882 pixels, of which the first 2\(^{14}\) pixels are chosen here. These data have been

2.2. The Skewness and Kurtosis Spectra of the Transmitted Flux

If appropriate parameters of the intergalactic UV background are adopted, the lognormal IGM model described
in § 2.1 could successfully explain many observed properties of the Ly4 forest and their evolution from redshift 2 to 4.
Now we show that it also works against tests of the non-
Gaussian features.

We use the wavelet transform to analyze the non-
Gaussian behavior of the transmitted flux \(F\). As a one-
dimensional field, the flux \(F(\lambda)\) in the wavelength range \(L = \lambda_{\text{max}} - \lambda_{\text{min}}\) is subject to a discrete wavelet transform
(DWT) as
\begin{equation}
F = \hat{F} + \sum_{j=0}^{\infty} \sum_{l=0}^{2^j-1} \hat{e}_{j,l} \psi_{j,l}(\lambda),
\end{equation}
where \(\psi_{j,l}(x), j = 0, 1, \ldots, l = 0 \ldots 2^j - 1\), is an orthogonal and complete set of the DWT basis (for details of the DWT,
see, e.g., Fang & Thews 1998). The wavelet basis \(\psi_{j,l}(x)\) is localized in both physical space and the Fourier (scale)
space. The function \(\psi_{j,l}(x)\) is centered at the position \(IL/2^j\) of the physical space, and at wavenumber \(2\pi \times 2^j/L\) of the
Fourier space. Therefore, the wavelet function coefficients
(WFCs), \(\hat{e}_{j,l}\), have two subscripts, \(j\) and \(l\). These describe the fluctuation of the flux on the scale \(L/2^j\) at position \(IL/2^j\). To be more specific, we use the Daubechies 4-wavelet in this paper, although all conclusions are not affected by this particular choice as long as a compactly supported wavelet basis is used.

The WFC, \(\hat{e}_{j,l}\), is computed by the inner product of
\begin{equation}
\hat{e}_{j,l} = \langle F \ast \psi_{j,l} \rangle.
\end{equation}
Since the DWT bases are complete, the WFCs contain all
information on the flux.

Note that \(\psi_{j,l}(x)\) are orthogonal with respect to the position
index \(l\), and therefore, for an ergodic field, the 2\(^j\) WFCs
at a given \(j\), i.e., \(\hat{e}_{j,l}, l = 0, 1, \ldots, 2^j - 1\), can be treated as independent measures of the flux field. The 2\(^j\) WFCs, \(\hat{e}_{j,l}\), from one realization of \(F(\lambda)\) can be employed as a statistical
ensemble. In other words, when the fair-sample hypothesis
holds (Peebles 1980), an ensemble average can be estimated equivalently by averaging over \(l\), i.e., \(\langle \hat{e}_{j,l} \rangle \simeq (1/2^j) \sum_{l=0}^{2^j-1} \hat{e}_{j,l}\), where \(\langle \cdots \rangle\) denotes the ensemble average. The distribution of \(\hat{e}_{j,l}\) represents approximately the one-
point distribution of the WFCs at a given scale \(j\). The non-Gaussianity of the flux \(F(\lambda)\) can be directly measured by the deviation of the one-point distribution from a Gaussian distribution. For this purpose, we calculate the cumulant moments defined by
\begin{align}
I_j^3 &= M_j^3, \\
I_j^4 &= M_j^4 - 3M_j^2M_j^2, \\
I_j^5 &= M_j^5 - 10M_j^3M_j^2.
\end{align}
where
\begin{equation}
M_n^j \equiv \frac{1}{2^n} \sum_{l=0}^{2^n-1} (\bar{F}_{l,j} - \bar{F}_{l,j})^n.
\end{equation}

FIG. 1.—Skewness spectrum of the Ly4 forests. The 95% confidence ranges of the skewness spectrum of the simulated samples and Gaussian noise are shown by the bars and gray band, respectively. The skewness spectrum of the Keck data of HS 1700 + 64 is shown by squares connected by the solid line. The physical scale related to \(j\) is 189.8 × 2\(^{-1}\) \(h^{-1}\) Mpc in the CDM model.
with 95% confidence. The skewness spectrum does not show significant difference from the Gaussian noise until \( j = 11 \) (\( \sim 100 \, h^{-1} \, \text{kpc} \)). These results are qualitatively consistent with that for the observed forest line distributions. The skewness and kurtosis spectra of the flux of HS 1700 + 64 are also presented in Figures 1 and 2. It is apparent that the CDM model is in excellent agreement with the observation.

2.3. The Scale-Scale Correlations of the Transmitted Flux

The scale-scale correlations measure the correlations between the fluctuations on different scales (Pando et al. 1998a, 1998b; Feng, Deng, & Fang 2000). This non-Gaussianity is independent of the higher order cumulants (§ 2.2), which are only \( j \)-dependent. A simplest measure of the scale-scale correlation is given by

\[
C_{j,p}^\text{pp} = \frac{2 \sum_{l=0}^{j+p-1} \sum_{l'=0}^{j+p-1} \bar{F}_l^p \bar{F}_{l'}^{p+1}}{\sum \bar{F}_l^p \sum \bar{F}_{l'}^{p+1}},
\]

(15)

where \( p \) is an even integer, and square brackets denote the integer part of the quantity. Because \( L/2^j = L2^{j+1} \), the position \( l \) at scale \( j \) is the same as the positions \( 2l \) and \( 2l + 1 \) at scale \( j + 1 \). Therefore, \( C_{j,p}^\text{pp} \) measures the correlation between fluctuations on scales \( j \) and \( j + 1 \) at the same physical point. For Gaussian fields, \( C_{j,p}^\text{pp} = 1 \). Values of \( C_{j,p}^\text{pp} > 1 \) correspond to the positive scale-scale correlation, and \( C_{j,p}^\text{pp} < 1 \) to the negative case. One variant of the above definition is

\[
C_{j,p,\Delta}^\text{pp} = \frac{2 \sum_{l=0}^{j+p-1} \sum_{l'=0}^{j+p-1} \bar{F}_l^p \bar{F}_{l'}^{p+1}}{\sum \bar{F}_l^p \sum \bar{F}_{l'}^{p+1}}.
\]

(16)

This statistics is for measuring the correlations between fluctuations on scales \( j \) and \( j + 1 \), but at different positions, i.e., the fluctuation at scale \( j \) is displaced from the \( j + 1 \) fluctuation by a distance \( \Delta L/2^j \).

The scale-scale correlation, \( C_{j,2,2} \), calculated from the simulated transmitted flux and HS 1700 + 64 are shown in Figure 3. Clearly, the values of \( C_{j,2,2} \) are significantly larger than unity and well above the Gaussian noise spectra on all scales \( j \geq 7 \). This result is also qualitatively in agreement with the scale-scale correlation of the Ly\( \alpha \) forests (Pando et al. 1998a). Figure 3 also indicates that the model of § 2.1 is still in good shape for fitting the observed non-Gaussian correlation.

Similar to equation (15), for the correlation between scales \( |j - j'| = 1 \), one can define, in principle, the correlation between two arbitrary scales with \( |j - j'| > 1 \). However, for hierarchical clustering, the scale-scale correlation is quantified mainly by \( |j - j'| = 1 \). Therefore, we will not calculate the scale-scale correlations for \( |j - j'| > 1 \).

3. THE NON-GAUSSIAN FEATURES OF THE GAUSSIAN-RECOVERED MASS FIELDS

3.1. Non-Gaussianity after Gaussianization

The purpose of the cosmological reconstruction is to extract the power spectrum of the initial linear mass fluctuations from the observed distribution of various tracers of the evolved density field. The algorithm of Gaussianization was designed to recover the primordial density fluctuations from an observed galaxy distribution (Weinberg 1992). This method has recently been applied to recovering the linear density field and its power spectrum from the observed transmitted flux \( F \) of QSO absorption spectra (Croft et al. 1998, 1999).

The key step of the Gaussianization algorithm is a pixel-to-pixel mapping from an observed flux \( F \) onto the density contrast \( \delta \). The PDF of the observed transmitted flux \( F \) is generally non-Gaussian, while the PDF of the initial density contrasts, \( \delta \equiv (n/n_0) - 1 \), is assumed to be Gaussian in a large variety of galaxy formation models. The relation between \( F = \exp (-\tau) \) and \( \delta \) is monotonic, i.e., high initial density \( \delta \) pixels evolved into high-\( \tau \) pixels, and low initial density pixels into low-\( \tau \) pixels. Thus, using the observed \( F \),
one can sort out the total number $N$ of pixels by the amount of $F$ in ascending order: the pixel with the lowest $F$ is labeled 1, the next higher $F$ pixel is labeled 2, and so on. For the $n$th pixel, we then assign the density contrast $\delta$, which is given by the solution of the equation $(2\pi)^{-1/2} \int_{-\infty}^{\infty} \exp\left(-x^2/2\right) dx = n/N$. Thus, the Gaussian mapping produces a mass field with the same rank order as the flux, but with a Gaussian PDF of $\delta(x)$. The overall amplitude of the recovered power spectrum should be determined by a separate procedure. For instance, we can set up the initial condition by using the recovered spectrum, evolve the simulation to the observed redshift, and then normalize the spectrum by requiring that the simulation reproduce the observed power spectrum of the transmitted QSO flux. This amplitude normalization is model-dependent.

We apply the Gaussianization to 100 simulation samples of the QSO transmitted flux, and measure the skewness and kurtosis spectrum as well as the scale-scale correlation. The results are displayed in Figures 4–6. For comparison, the non-Gaussian spectra of the flux in Figures 1–3 are also plotted correspondingly. Figures 4–6 show that the Gaussianized flux still largely exhibits non-Gaussian features. In particular, the scale-scale correlation of the Gaussianized field is as strong as the pre-Gaussianized flux on scales $j \geq 10$. That is, the recovered density field is seriously contaminated by the non-Gaussianities in the original flux.

### 3.2. The Efficiency of the Conventional Gaussianization

The reason for the lower efficiency of the conventional Gaussian mapping (§ 3.1) is simple. The initial Gaussian random mass field is assumed to be a superposition of independent modes, of which the PDFs are Gaussian. For instance, in the Fourier representation, all Fourier modes of a Gaussian mass field are Gaussian, i.e., they have a Gaussian PDF of the amplitudes and randomized phases. The conventional algorithm considers only the Gaussianization of one variable, $\delta$. It does not guarantee the Gaussianization of the amplitudes and phases of all relevant modes. In other words, the Gaussian mapping algorithm will work perfectly for a system with one stochastic variable, but not for a field.

Alternatively, this problem can also be seen via the DWT representation. Using equation (1), any one-dimensional mass field given by density contrast $\delta(x) (\delta = 0)$ can be decomposed with respect to a DWT basis as

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

(17)

where the superscript $M$ means mass. Equation (17) represents a linear superposition of modes $\psi_{j,l}$. As has been pointed out in § 2.2, for a given $j$, the $2^j$ WFCs $\hat{e}_M^{j,l}$ form a statistical ensemble. The distribution of the $2^j$ WFCs gives

---

**Fig. 4.** Skewness spectrum of the mass field, recovered by the three Gaussianization methods. *Top*: Transmitted flux samples free from the effect of peculiar velocity; *bottom*: samples affected by the peculiar velocity. The 95% confidence ranges of the skewness spectrum are shown by the symbols as indicated in figure. The gray band shows the 95% confidence ranges for the Gaussian noise samples.

**Fig. 5.** Same as Fig. 4, but for the kurtosis spectrum.
between the distributions of WFCs with different index and therefore the local relations equation (18) remains. The Gaussian mapping changes all the WFCs at a given position. Moreover, this correlation cannot be eliminated by the Gaussianization of $\tilde{e}^M_{j,1}$. The Gaussian mapping changes all the WFCs at a given position (pixel) by the same amplifying or reducing factor, and therefore the local relations equation (18) remains.

The scale-scale correlations only depend on the statistical behavior of the fluctuation distribution with respect to the index $j$. Therefore, a Gaussian field requires a uncorrelation between the distributions of WFCs with different $j$. This uncorrelation corresponds to decorrelating the band-average Fourier modes, which will be discussed in detail in § 4.

3.3. Algorithms of Scale-by-Scale Gaussianization

Based on the considerations in the last section, we can design an algorithm that is capable of reducing the contamination of the non-Gaussianity, and produce fields with less non-Gaussianity. The new method is based on the scale-by-scale decomposition of flux and mass fields. From equation (6), we have

$$ F = F^j + \sum_{f=0}^{\infty} \sum_{l=0}^{2^j-1} \tilde{e}_{f,l}^M \psi_{f,l} $$

(19)

and

$$ F^j = \tilde{F} + \sum_{f=0}^{j-1} \sum_{l=0}^{2^j-1} \tilde{e}_{f,l}^M \psi_{f,l}, $$

(20)

where $F^j$ is $F$ smoothed by a filter on the scale of $j$. There is a recursion relation in $F^j$, given by

$$ F^{j+1} = F^j + \sum_{l=0}^{2^j-1} \tilde{e}_{j,l}^M \psi_{j,l}, $$

(21)

namely, the flux $F^{j+1}$ can be reconstructed from the flux $F^j$ and $2^j$ WFCs, $\tilde{e}_{j,l}$, at the scale $j$. Similarly, for a mass distribution, we have

$$ \delta = \delta^j + \sum_{f=0}^{\infty} \sum_{l=0}^{2^j-1} \tilde{e}_{f,l}^M \psi_{f,l}, $$

(22)

and

$$ \delta^{j+1} = \delta^j + \sum_{l=0}^{2^j-1} \tilde{e}_{j,l}^M \psi_{j,l}. $$

(23)

(24)

Since the relations between $F$ and $\delta$ are local and monotonic, the smoothed flux $F^{j+1}$ depends only on the smoothed mass field $\delta^{j+1}$, and one can perform a local and monotonic mapping between $F^{j+1}$ and $\delta^{j+1}$. Thus, we can implement the reconstruction of the mass field $\delta^{j+1}$ from $F^{j+1}$ by a scale-by-scale Gaussianization algorithm (hereafter referred to as algorithm I):

1. Suppose that the reconstruction down to the scale $j$ has been done, i.e., the $\delta^j$ is already known.
2. Calculate the WFCs of the flux $F$ on the scale $j$.
3. Perform the Gaussian mapping of the $2^j$ WFCs $\tilde{e}_{j,1}$, and assign the Gaussianized result, $\hat{e}_{j,1}$, to the $2^j$ pixels according to the rank order. The distribution of $\hat{e}_{j,1}$ is Gaussian with zero mean and variance 1.
4. Find the $2^j$ WFCs of the mass field by

$$ \tilde{e}_{j,l}^M = \nu \hat{e}_{j,l}^M, $$

(25)

where the parameter $\nu$ is a normalization factor to be determined. The one-point distribution of the WFCs of mass field at the scale $j$, $\tilde{e}_{j,l}^M$, is then Gaussianized.
5. Reconstruct the mass field $\delta^{j+1}$ on the scale $j+1$ by the recursion relation, equation (24).
6. To determine the parameter $\nu$, we require that the DWT power spectrum of the flux $F^{j+1}$ simulated from $\delta^{j+1}$...
reproduces the observed flux $F^{j+1}$. We then have $\delta^{j+1}$. The reconstruction of the mass field on the scale $j + 1$ is done.

Repeating steps 1–6, one can reconstruct the mass field on scales from large to small until the scale of the resolution of the flux $F$, or the scale on which the relation between $F$ and $\delta$ is no longer local.

Figure 7 illustrates the transmitted flux, the initial density field, and the density field recovered by algorithm I. The recovered one-dimensional density field is in excellent agreement with the original density field scale-by-scale. The non-Gaussianities of the recovered fields by algorithm I are shown in Figures 4–6. The skewness and kurtosis spectra exhibit almost nothing but Gaussianity. The scale-scale correlation is also significantly reduced.

The Gaussianization algorithm I is conceptually clear. However, it needs to determine the normalization factor $v$ at each scale. Therefore, it is rather cumbersome to do the numerical calculation. Moreover, there is still some residual scale-scale correlation in the recovered mass field. In fact, algorithm I does ensure the Gaussian PDF of $v_{i,j}$, but it is unable to remove all the correlation between different modes, just as the simple example (eq. [18]) demonstrated in §3.2.

To avoid the multiple normalizations and preserve the virtues of scale-by-scale Gaussianization, we design an alternative algorithm as follows (hereafter referred to as algorithm II):

1. Use the conventional Gaussianization ($§$3.1) to reconstruct the mass field, i.e., to perform Gaussian mapping of the density contrast $\delta$ and normalize the mass field by requiring that the evolved simulations reproduce the power spectrum of the observed flux.
2. Calculate the WFCs, $\delta_{j,l}$, of the recovered mass field, $\delta^M$, on each scale $j$.
3. Similar to step 3 of algorithm I, perform the Gaussianization of $\delta_{j,l}$ for each scale $j$ to produce unnormalized WFCs, $\delta_{j,l}$.
4. Normalize the WFCs, $\delta_{j,l}$, on scale $j$ by requiring that the variance of $\delta_{j,l}$, i.e., the second cumulant moment $I_j^2$ (eq. [7]), is the same as those for the WFCs $\delta_{j,l}$.
5. For each scale $j$, randomize the spatial sequence of the Gaussianized WFCs $\delta_{j,l}$, i.e., make a random permutation among the index $l$.
6. Using these WFCs $\delta_{j,l}$, one can reconstruct the mass density field by equation (24) up to the scale given by the resolution of the flux.

Algorithm II is still scale-by-scale in nature. However, the normalization is done only once for the recovered $\delta$. Step 4 ensures that the normalization is unchanged after step 3, which eliminates the non-Gaussianities of the skewness and kurtosis spectra. Step 5 eliminates the residual scale-scale correlations by a randomization of the spatial index $l$ of $\delta_{j,l}$. Namely, it changes only the position of $\delta_{j,l}$, but not the values. Therefore, it is similar to a randomization of phases of the Fourier modes, and will not change the normalization of the amplitude and power spectrum of the fields.

Figures 4–6 show that the Gaussianized field found by algorithm II contains almost none of the non-Gaussian features considered. However, it should be pointed out that the field given by algorithm II is no longer a point-to-point reconstruction, due to the randomization of $l$. Namely, the recovered field will not be point-to-point the same as the field shown in Figure 7. Nonetheless, since the purpose of the Gaussianization is to recover the power spectrum of the primordial density fluctuations, algorithm II is a valuable approach. As will be shown in the next section, algorithm II gives a more unbiased estimation of the power spectrum by the standard FFT technique.

In order to illustrate the effect of peculiar velocities on the Gaussianization, each of Figures 4–6 contains two panels: one employed the simulation samples including the effects of peculiar velocities, and the other did not. All the figures show that for algorithm I, the effect of peculiar velocities is significant only on small scales, $j > 9$ or $k >
10 h Mpc⁻¹; while for algorithm II, the effect of peculiar velocities appears on smaller scales. Therefore, our proposed scale-by-scale Gaussianization methods would not be affected by the peculiar velocities as least up to the scale \( j = 9 \).

4. RECOVERY OF MASS POWER SPECTRUM FROM THE TRANSMITTED QSO FLUX

4.1. The Power Spectrum in Different Representations

As a preparation for measuring the non-Gaussian effects on power spectrum recovery, we first discuss the representation of the power spectrum. Principally, a random field can be described by any complete orthonormal basis (representation). Although the default representation of the power spectrum is defined on a Fourier basis, one can define the power spectrum with respect to different representations. This is due to the fact that Parseval’s theorem holds for any complete and orthonormal-basis decomposition.

In the Fourier representation, the power spectrum of a one-dimensional density field, \( \delta(x) \), is given by

\[
P(n) = |\hat{\delta}_n|^2,
\]

where \( \hat{\delta}_n \) is the Fourier transform of \( \delta(x) \). The value of \( |\hat{\delta}_n|^2 \) measures the power of mode \( n \) because of Parseval’s theorem,

\[
\frac{1}{L} \int_0^L \delta^2(x) \, dx = \sum_{n = -\infty}^{\infty} |\hat{\delta}_n|^2.
\]

Similarly, we have Parseval’s theorem for the DWT transform, given by (Fang & Thews 1998; Pando & Fang 1998b)

\[
\frac{1}{L} \int_0^L \delta^2(x) \, dx = \sum_{j=0}^{2^{j-1}-1} \sum_{l=0}^{2^j} |\hat{\epsilon}_{j,l}|^2.
\]

(For simplicity, we ignore the superscript \( M \) on \( \epsilon_{j,l} \)). Therefore, the term \( |\hat{\epsilon}_{j,l}|^2 \) describes the power of mode \((j, l)\), and the total power on the scale \( j \) is

\[
P_j = \frac{1}{L} \sum_{l=0}^{2^{j-1}-1} |\hat{\epsilon}_{j,l}|^2,
\]

which defines the DWT power spectra \( P_j \).

Generally, the second-order correlation functions of \( \hat{\delta}_n \) or \( \hat{\epsilon}_{j,l} \) can be converted from each other by

\[
\langle \epsilon_{j',l'} \epsilon_{j,l} \rangle = \sum_{n',n} \langle \delta_{n'} \delta_n \rangle \psi_{j',l'}(n') \psi_{j,l}(n),
\]

\[
\langle \delta_{n'} \delta_n \rangle = \sum_{j',l'=0}^{2^{j-1}-1} \sum_{j,l=0}^{2^j} \langle \epsilon_{j',l'} \epsilon_{j,l} \rangle \psi_{j',l'}(n') \psi_{j,l}(n),
\]

where \( \psi_{j,l}(n) \) is the Fourier transform of \( \delta_{j,l}(x) \). For an homogeneous random field, \( \langle \delta_{n'} \delta_n \rangle = \langle |\hat{\delta}_n|^2 \rangle \delta_{n,n'} \), we then have

\[
\langle \epsilon_{j,l}^* \rangle = \sum_{n = -\infty}^{+\infty} |\hat{\delta}_n|^2 \langle \psi_{j,l}(n) \rangle^2
\]

or

\[
P_j = \sum_{n = -\infty}^{+\infty} P(n) |\psi(n/2)|^2,
\]

where \( \psi(n/2) \) is the Fourier transform of the generating wavelet \( \psi(x) \) (Pando & Fang 1998b). In equation (33), the function \( |\psi(n/2)|^2 \) plays the role of window function in the wavenumber \( n \) space. The function \( \psi(n) \) is localized in \( n \)-space. For the Daubechies 4-wavelet, \( |\psi(n)| \) is peaked at \( n = \pm n_0 \), with the width of \( \Delta n \). Therefore, the DWT spectrum \( P_j \) gives an estimator of the “band-averaged” Fourier power spectrum within the band centered at

\[
\log n = (\log 2)j + \log n_p,
\]

with the bandwidth

\[
\Delta \log n = \Delta n_p/n_p.
\]

Since the mean of the WFCs \( \tilde{\epsilon}_{j,l} \) over \( j \) is zero, the second cumulant moment \( I_j^2 \) is related to the DWT spectrum \( P_j \) by \( I_j^2 = (L/2^j)P_j \), and we use the variance \( I_j^2 \) as the estimator of the DWT power spectrum instead of \( P_j \).

For a Gaussian field, the statistical behavior is completely determined by the second-order statistics of the Gaussian variables \( \delta_n \) or \( \tilde{\epsilon}_{j,l} \). Theoretically, the power-spectrum estimators \( P(n) \) and \( P_j \) present an equivalent description. However, as will be shown below, once non-Gaussianity appears, these estimators will no longer be equivalent.

4.2. Effect of Non-Gaussianity on the Recovery of the Mass Power Spectrum

Using the 100 realizations of the mass density fields recovered by the conventional algorithm, and algorithms I and II of the Gaussianization, we calculated the power spectra by the standard FFT technique. To reveal the effect of non-Gaussianity on the power spectrum estimation, we do not include the effects of instrumental noise and continuum fitting in the synthetic spectra. The dominant sources of error in estimating the power spectrum would be the cosmic variance and the non-Gaussian effects.

Figure 8 compares the power spectra obtained by different Gaussianization methods. The one-dimensional linear power spectrum of equation (3) is also shown by a solid line. These power spectra are normalized to the present. In general, the recovered power spectrum can match the shape of the linear theory over a wide range of wavelengths, particularly on larger wavelengths. Yet the recovered spectra show a somewhat systematic departure from the initial mass power spectrum with the increase of wavenumbers. For the conventional Gaussianization, the recovered power spectrum falls below the initial power spectrum on scales of \( j \geq 8 \) or \( k \sim 1.5 h^{-1} \) Mpc. The power spectrum recovered by algorithm I is better than the conventional Gaussianization, and the recovery by algorithm II gives the best one, which is almost the same as the initial power spectrum on all scales.

Comparing Figure 8 with Figures 4–6, we can see that the scales on which the depression of the recovered power spectrum appears is always the same as the scale on which the scale-scale correlations become significant. Moreover, the less the scale-scale correlation (Fig. 6), the less the depression. This indicates that the recovered spectrum is substantially affected by the non-Gaussianities, especially the scale-scale correlations. Actually, this effect has already been recognized by Meiksin & White (1999) in analyzing \( N \)-body simulation samples. Namely, the goodness of a
power spectrum estimation is significantly dependent on the correlation between the Fourier power spectra averaged at different scale bands.

Recalling from the definition of the scale-scale correlation equation (15) that the average over an ensemble is equivalent to the spatial average taken over one realization, equation (15) can be rewritten as

$$C_{j_2,j_2} = \frac{\langle P_j P_{j+1} \rangle}{\langle P_j \rangle \langle P_{j+1} \rangle}. \quad (36)$$

Hence, the scale-scale correlation is actually a measure of the correlation between the Fourier power spectra averaged at different scale bands. It can also be seen from equation (31) that the Fourier power spectrum around $n$ depends on the fluctuations on different $j$, and therefore on their non-Gaussian correlations.

Because algorithm II is most effective for eliminating the scale-scale correlations, the resulting power spectrum shows the best recovery of the linear model.

### 4.3. Suppression of Non-Gaussian Correlations by Representation

In the DWT representation, the power spectrum given by equation (29) does not depend on modes at the scales different from $j$, and therefore the scale-scale correlation will not affect the estimation of $P_j$. One can expect that the DWT spectrum estimator, $P_j$, will give a better recovery of the initial power spectrum.

Figure 9 displays the DWT power spectrum $P_j$ for mass fields given by the different Gaussianization methods. The DWT power spectrum in the linear CDM model is also shown by a solid line, which is calculated from the Fourier linear power spectrum by equation (33) in the continuous limit of $n$. This figure indicates that even for the mass field recovered from conventional Gaussianization, the DWT power spectrum is in good agreement with the initial DWT mass power spectrum up to the scale $j = 9$. This is already much better than its counterpart in the Fourier representation, for which the power spectrum shows significant difference from the linear spectrum on scale $j \approx 8$. For algorithms I and II, the DWT power spectrum also gives the good results. In addition, the errors due to the cosmic variance and normalization in the DWT spectrum are manifestly smaller than for the Fourier spectrum.

The DWT power spectrum $P_j$ (eq. [29]) is given by the summation of $|\epsilon_{j,l}|^2$ over $j$ at a given scale. Therefore, the non-Gaussian effect on the estimation of $P_j$ mainly arises from the correlation between the WFCs, $\epsilon_{j,l}$, at different $l$, which can be measured by

$$Q_{j,j+1}^{2,2} = \frac{2l \sum_{j=0}^{j+l-1} \epsilon_{j,l}^2 \epsilon_{j,l+1}^2}{\sum_{l} \epsilon_{j,l}^2 \sum_{l} \epsilon_{j,l+1}^2}, \quad (37)$$

where $Q_{j,j+1}^{2,2}$ gives the correlation between the density fluctuations on the same scale $j$ at different places $l$ and $l + \Delta l$.

Figure 10 displays the correlations $Q_{j,j+1}^{2,2}$ with $\Delta l = 1$. It shows that this non-Gaussianity can be ignored until $j = 10$. On the other hand, the scale-scale correlation $C_{j_2,j_2}$ had been significant at $j = 8$ (Fig. 3). As a result, the Fourier power spectrum is contaminated by the non-Gaussianity on scales $j \geq 8$, while the DWT power spectrum is less biased until $j = 9$.

In a word, the non-Gaussian correlations are effectively suppressed in the DWT representation. The DWT spec-
Figure 9.—As in Fig. 8, but for the DWT power spectra.

CONCLUSIONS

In the cosmological reconstruction of the initial Gaussian mass power spectrum, a serious obstacle is the non-Gaussianity of the evolved field. The quality of the recovery of the power spectrum is affected by the non-Gaussian correlations. The precision to which the mass power spectrum can be measured relies on how we treat the non-Gaussianity of the evolved mass field.

In the quasi-nonlinear regime of cosmic gravitational clustering (like that traced by the Lyα forests), the dynamical evolution is characterized by the power transfer from large-scale perturbations to small ones (Suto & Sasaki 1991). This is the mode-mode coupling that produces the scale-scale correlations. Using perturbation theory in the DWT representation, one can further show that the mode-mode coupling at the same position (local coupling) is much stronger than coupling between modes at different positions (nonlocal coupling; Pando, Feng, & Fang 2000). On the other hand, the power spectrum in the quasi-nonlinear regime does not significantly differ from that of the linear regime. Therefore, the algorithm for recovering the initial mass power spectrum from the Lyα forests should be designed to eliminate the local scale-scale correlations of the evolved mass field.

Using simulations in semianalytical models of the Lyα forests, we show that the conventional algorithm of the Gaussianization is not sufficient to recover a Gaussian field. The local scale-scale correlations of the Lyα forests are still retained in the Gaussianized mass field. Based on the DWT scale-space decomposition, we proposed two algorithms of the Gaussianization, which are effective in eliminating the non-Gaussian features.

We showed that representation selection is important for the recovery of the power spectrum. A representation that can effectively suppress the contamination of local scale-scale correlations would be good for extracting the initial linear spectrum. We compared the Fourier and DWT representations for estimating the power spectrum. We demonstrated that, at least in the quasi-nonlinear regime, the DWT power spectrum estimator is better, because it can avoid the major contamination, the local scale-scale correlations. We also showed that the peculiar velocities of gas will not affect the DWT power spectrum recover up to at least the scale $j = 9$. 

Figure 10.—Nonlocal correlations $Q^2_j r$ with $\Delta l = 1$ of the recovered density field by the three Gaussianization methods. The symbols displayed in the figure have the same meanings as in Fig. 6.
We thank D. Tytler for kindly providing the data of the Keck spectrum HS 1700 + 64. We also thank Wei Zheng, Hongguang Bi, and Wolung Lee for useful discussion. L. L. F. acknowledges support from the National Science Foundation of China (NSFC) and a World Laboratory scholarship. This project was done during L. L. F.’s visit at the Department of Physics, University of Arizona. This work was supported in part by the LWL foundation.

REFERENCES

Bi, H. G. 1993, ApJ, 405, 479
Bi, H. G., & Davidsen, A. F. 1997, ApJ, 479, 523
Bi, H. G., Ge, J., & Fang, L. Z. 1995, ApJ, 452, 90
Croft, R. A. C., Weinberg, D. H., Katz, N., & Hernquist, L. 1998, ApJ, 495, 44
Croft, R. A. C., Weinberg, D. H., Pettini, M., Hernquist, L., & Katz, N. 1999, ApJ, 520, 1
Duncan, R. C., Ostriker, J. P., & Bajtlik, S. 1989, ApJ, 345, 39
Efstathiou, G., Bond, J. R., & White, S. D. M. 1992, MNRAS, 258, 1P
Eisenstein, D. J., & Hu, W. 1999, ApJ, 511, 5
Fang, L. Z. 1991, A&A, 244, 1
Fang, L. Z., Bi, H. G., Xiang, S. P., & Börner, G. 1993, ApJ, 413, 477
Fang, L. Z., & Thews, R. 1998, Wavelets in Physics (Singapore: World Scientific)
Feng, L. L., Deng, Z. G., & Fang, L. Z. 2000, ApJ, 530, 53
Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, ApJ, 457, L57

Hui, L., Gnedin, N., & Zhang, Y. 1997, ApJ, 486, 599
Liu, X. D., & Jones, B. J. T. 1990, MNRAS, 242, 678
Meiksin, A., & White, M. 1999, MNRAS, 308, 1179
Pando, J., & Fang, L. Z. 1996, ApJ, 459, 1
———. 1998a, A&A, 340, 335
———. 1998b, Phys. Rev. E, 57, 3593
Pando, J., Feng, L. L., & Fang, L. Z. 2000, ApJ, submitted
Pando, J., Greiner, M., Lipa, P., & Fang, L. Z. 1998a, ApJ, 496, 9
Pando, J., Valls-Gabaud, D., & Fang, L. Z. 1998b, Phys. Rev. Lett., 81, 4568
Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Sarazin, C. L., & Bahcall, J. N. 1977, ApJS, 34, 451
Suto, Y., & Sasaki, M. 1991, Phys. Rev. Lett., 66, 264
Walker, T. P., Steigman, G., Schramm, D. N., Olive, K. A., & Kang, H. S. 1991, ApJ, 376, 51
Weinberg, D. H. 1992, MNRAS, 254, 315