Discrete wavelet transform power spectrum estimator

Jesús Pando
Department of Physics, University of Arizona and UMR 7550 CNRS, Observatoire de Strasbourg, France

Li-Zhi Fang
Department of Physics, University of Arizona, Tucson, Arizona 85721

(Received 24 July 1997; revised manuscript received 10 November 1997)

A method for measuring the spectrum of a density field by the discrete wavelet transform (DWT) is studied. We show how the Fourier power spectrum can be detected by using the wavelet function coefficients (WFC) of the DWT. This method can successfully measure the power spectrum in samples for which traditional methods often fail because the samples are finite sized, have a complex geometry, or are varyingly sampled. We demonstrate that the spectrum features, such as the power law index, the magnitude, and the typical scales can be determined by the DWT reconstructed spectrum. We apply this method to analyze the power spectrum of the spatial distribution of the Ly-α clouds. The two popular data sets used for the spectrum detection have quite different geometries and samplings, yet the one-dimensional (1D) power spectra and their 3D reconstruction given by the DWT estimator show the same features. The analysis makes clear that the DWT estimator is a sensitive tool in revealing common and physical properties from diverse data sets. [S1063-651X(98)05603-7]

PACS number(s): 02.70.Hm, 02.70.Rw

I. INTRODUCTION

The power spectrum or its Fourier transform pair, the autocorrelation function, are probably the most commonly used techniques to detect structure in distributions such as galaxies, photons, bacteria, hadrons, etc. Although the power spectrum is only the lowest order statistical measure of the deviations of the random density field from homogeneity, it directly reflects the physical scales of the processes that affect structure formation. Moreover, the positive definiteness of the power spectrum is useful for constraining the parameter space in comparing predictions with data. As a result, the estimation of the power spectrum is often the first statistical description of a distribution attempted.

Even though there are a myriad of ways to estimate the power spectrum, efforts continue on developing new estimators for the spectrum. The main reason for this is that the traditional estimators have several weaknesses, especially when dealing with samples that are finite sized, have a complex geometry, are sampled irregularly, or in which the mean density is uncertain or varying. We discuss these problems in more detail next.

It is well known that a classical spectrum estimator, the Fourier transform of the autocorrelation function, depends essentially on a good measure of the mean density [1]. A two-point correlation analysis cannot detect any correlations with amplitudes comparable to the uncertainty in the mean density. If the spectrum is determined via the two-point correlation function, uncertainties on all the scales on which the correlation amplitude is comparable to the uncertainty in the mean density will occur. This problem is more severe for cases in which the mean density is changing with the size of the sample. It is difficult, or even impossible, to accurately determine the mean density in these distributions because of the finiteness in the sample size. This is sometimes called the infrared (long-wavelength) uncertainty in the Fourier spectrum estimator.

If, on the other hand, the spectrum for a finite sample is estimated directly from the Fourier transform, this generally only gives a convolution of the true power with the window function $W(\mathbf{x})$, which is 1 inside the sample volume and 0 elsewhere. Because the basis functions of the Fourier transform are not orthonormal over a finite nonperiodic volume, the convolved spectrum depends on the shape of the sample volume. In order to subtract the contribution of the window function to each Fourier mode, an estimation of the mean density of the objects is still needed.

Furthermore, the mean density is generally estimated from the sample itself. Therefore, the problem becomes more serious when the sampling rate, or the object density, is different for different samples. Forming an ensemble from such samples in order to detect the spectrum becomes problematic. The effect of the finite size of the distribution cannot be eliminated because the Fourier basis is not localized.

This difficulty can be overcome by using the count in cell (CIC) technique or the Fourier transform on a finite domain (the Gabor transform, for instance) [2]. The CIC detects the variance σ^2 of the density fluctuations in cubic cell windows with side l or of Gaussian spheres with radius R_G. Since the CIC variance is less dependent on the geometry of the sample volume, the behavior of the distribution outside the sample is not needed. This reduces the uncertainties caused by finite sized samples. It is believed that the variance in cell l (or sphere of radius R_G) is mainly given by perturbations on scale $\sim l$ (or R_G). As such the variances are considered to be a measure of the power spectrum on scale l [3].

The problem with the CIC measure is that its basis functions (windows) are not orthogonal with respect to scale. The variances $\sigma^2(l)$ obtained from the cells at scale l contain contributions from scales larger than l. We will show that this scale mixing becomes a serious problem when a power law spectrum has a negative index (see Sec. II). As a consequence, the resulting errors are not easy to interpret because the errors for different l are not independent.

Thus, in order to effectively detect the power spectrum from samples in which the above mentioned problems occur, the basis functions for the spectrum should be orthogonal and localized in both physical and scale spaces. This motivated us to explore the spectrum estimator based on the discrete wavelet transform (DWT) because its basis functions have just these properties [4–8].

The basis of the DWT is constructed from two sets of localized functions: the scaling functions and the wavelet functions [4]. To expand a function in a wavelet basis, only the scaling function coefficients (SFC), measuring the local mean density, and the wavelet function coefficients (WFC), measuring the differences between the *local* mean densities at adjoining scales, are needed [8]. As a consequence, the mean density on length scales larger than the sample size is not needed in calculating the WFC. Thus, aside from an additive constant, the wavelet spectrum (see Sec. II) will not be affected by the uncertainty in the mean density.

The Karhuen-Loéve (KL) transform given by diagonalizing the covariance matrix, then optimizing over the set of orthogonal transformations of the covariance matrix, is another way to detect the spectrum that avoids the pitfalls of finite sized samples [9–12]. However, finding the KL eigenvectors of a matrix of order f has computing complexity $O(f^3)$. In addition, the KL basis is not admissible. Even after finding the eigenvectors of a data set, updating the basis with some extra samples will cost an additional $O(f^3)$ operations. Finally, the KL transform is only available for second order statistics. On the other hand, the DWT can also quasidiagonalize the covariance matrix, and thus the KL transform can be approximately represented by wavelets, which leads to less computing complexity [13,14]. More importantly, the DWT can be generalized to higher order statistics [15].

In Sec. II, we develop all the necessary formalism to determine the power spectrum of a distribution using the DWT. In Sec. III, we demonstrate the capability of the DWT in reconstructing power spectra of various types. In Sec. IV, as an application of the DWT estimator, we detect the power spectrum for real samples of the Lyman-α (Ly-α) absorption clouds. We end with a discussion of our findings.

II. THE DISCRETE WAVELET TRANSFORM POWER SPECTRUM

The use of the continuous wavelet transform (CWT) has gained increasing importance in many fields in analyzing the texture of matter distributions. However, the machinery we will develop here to estimate the power spectrum is based on the *discrete* wavelet transform [16–21]. Unlike the case with the Fourier transform, in which the discrete Fourier transform is just the continuous Fourier transform estimated from a discrete grid of points, the difference between the CWT and DWT is more fundamental. The DWT is not constructed by estimating the CWT at discrete points. That is, one does not merely replace the integral with a sum and a continuous variable with a discrete variable. Generally, the basis functions of the CWT are overcomplete and nonorthogonal, whereas the DWT basis functions are complete and orthogonal. Using the CWT can lead to correlations that are not in the sample, but due to the correlations among the wavelet coefficients. The DWT allows for an orthogonal projection on a complete set of modes and thus, the wavelet coefficients are independent.

A. The power spectrum

Let us consider a one-dimensional random density field $\rho(x)$. It is not difficult to extend all results into two and three dimensions. In studying structures in a density field, it is convenient to use the density contrast (or perturbations) defined by

$$\epsilon(x)=\frac{\rho(x)-\overline{\rho}}{\overline{\rho}}, \tag{2.1}$$

where $\overline{\rho}$ is the mean density of the field. The Fourier expansion of ϵ is

$$\epsilon(x)=\sum_{n=-\infty}^{\infty}\epsilon_n e^{i2\pi nx/L} \tag{2.2}$$

with the coefficients given by

$$\epsilon_n=\frac{1}{L}\int_0^L \epsilon(x)e^{-i2\pi nx/L}dx. \tag{2.3}$$

Parseval's theorem relates the power for a distribution to the coefficients of the Fourier expansion. For the density contrast this yields

$$\frac{1}{L}\int_0^L |\epsilon(x)|^2 dx=\sum_{n=-\infty}^{\infty}|\epsilon_n|^2, \tag{2.4}$$

which shows that the perturbations can be decomposed into domains, n, by the orthonormal Fourier basis functions. The power spectrum of perturbations on the scale L/n is then defined as

$$P(n)=|\epsilon_n|^2. \tag{2.5}$$

We proceed similarly using the DWT basis functions. To subject a finite sample of extent $0<x<L$ to a DWT expansion, we define a density distribution $\epsilon(x)$ that is equal to the sample in the region $0<x<L$, and is L periodic on $-\infty<x<\infty$. The wavelet expansion of $\epsilon(x)$ is [8,16]

$$\epsilon(x)=\sum_{j=0}^{\infty}\sum_{l=-\infty}^{\infty}\tilde{\epsilon}_{j,l}\psi_{j,l}(x), \tag{2.6}$$

where $\psi_{j,l}(x)$ is the wavelet function, defined as

$$\psi_{j,l}(x)=\left(\frac{2^j}{L}\right)^{1/2}\psi(2^j x/L-l). \tag{2.7}$$

The real function $\psi(\eta)$ is called the generating wavelet and is localized in the interval $0\leqslant\eta\leqslant 1$ and centered at $\eta=1/2$.

Equation (2.7) shows that the wavelet functions $\psi_{j,l}(x)$ are families of functions constructed by dilating the generating function $\psi(x/L)$ by a factor of 2^j, and by translating $\psi(x/L)$ by l. The wavelet functions $\psi_{j,l}(x)$ are thus on scale $L/2^j$ and are centered at $lL/2^j$. Furthermore the $\psi_{j,l}(x)$ are orthogonal with respect to both indices, that is, the $\psi_{j,l}(x)$

are orthogonal to both translation and dilation. The wavelet function coefficients (WFC), $\tilde{\epsilon}_{j,l}$, in Eq. (2.6) are simply the inner product

$$\tilde{\epsilon}_{j,l}=\int_{-\infty}^{\infty}\epsilon(x)\psi_{j,l}(x)dx. \tag{2.8}$$

The limits of the integral here are formally $-\infty$ and ∞. In practice, $\epsilon(x)$ only in the interval $L/2^j$ centered at $lL/2^j$ is needed to calculate $\tilde{\epsilon}_{j,l}$ because $\psi_{j,l}(x)$ is localized.

Finally, Parseval's theorem for the DWT can be shown to be [8]

$$\frac{1}{L}\int_0^L|\epsilon(x)|^2dx=\sum_{j=0}^{\infty}\frac{1}{L}\sum_{l=0}^{2^j-1}|\tilde{\epsilon}_{j,l}|^2. \tag{2.9}$$

The existence of Parseval's identity for the DWT strongly suggests that the second order statistical behavior of $\epsilon(x)$ can be completely described by the $|\tilde{\epsilon}_{j,l}|^2$, i.e., the DWT power spectrum. Comparing Eqs. (2.4) and (2.9), one can relate $(1/L)\Sigma_{l=0}^{2^j-1}|\tilde{\epsilon}_{j,l}|^2$ to the power of perturbations on length scale $L/2^j$, and $|\tilde{\epsilon}_{j,l}|^2/L$ to the power of the perturbations on scale $L/2^j$ and position $lL/2^j$. Thus, the power spectrum with respect to the wavelet basis should be defined as

$$P_j=\frac{1}{L}\sum_{l=0}^{2^j-1}|\tilde{\epsilon}_{j,l}|^2. \tag{2.10}$$

B. Relationship between ϵ_n and $\tilde{\epsilon}_{j,l}$

Both the DWT and Fourier basis functions are complete so it is possible to express the WFC $\tilde{\epsilon}_{j,l}$ in terms of the Fourier coefficients ϵ_n, and vice versa. Substituting expansion (2.2) into Eq. (2.8) yields

$$\tilde{\epsilon}_{j,l}=\sum_{n=-\infty}^{\infty}\epsilon_n\int_{-\infty}^{\infty}e^{i2\pi nx/L}\psi_{j,l}(x)dx=\sum_{n=-\infty}^{\infty}\epsilon_n\hat{\psi}_{j,l}(-n), \tag{2.11}$$

where $\hat{\psi}_{j,l}(n)$ is the Fourier transform of $\psi_{j,l}(x)$, i.e.,

$$\hat{\psi}_{j,l}(n)=\int_{-\infty}^{\infty}\psi_{j,l}(x)e^{-i2\pi nx/L}dx. \tag{2.12}$$

Using Eq. (2.7), one can rewrite Eq. (2.11) as

$$\tilde{\epsilon}_{j,l}=\sum_{n=-\infty}^{\infty}\left(\frac{2^j}{L}\right)^{1/2}\epsilon_n\int_{-\infty}^{\infty}e^{i2\pi nx/L}\psi(2^jx/L-l)dx. \tag{2.13}$$

Letting $\eta=2^jx/L-l$ gives

$$\tilde{\epsilon}_{j,l}=\sum_{n=-\infty}^{\infty}\left(\frac{2^j}{L}\right)^{-1/2}\epsilon_ne^{i2\pi nl/2^j}\int_{-\infty}^{\infty}e^{i2\pi n\eta/2^j}\psi(\eta)d\eta \tag{2.14}$$

or

$$\tilde{\epsilon}_{j,l}=\sum_{n=-\infty}^{\infty}\left(\frac{2^j}{L}\right)^{-1/2}\epsilon_n\hat{\psi}(-n/2^j)e^{i2\pi nl/2^j}, \tag{2.15}$$

where $\hat{\psi}(n)$ is the Fourier transform of the generating wavelet $\psi(\eta)$

$$\hat{\psi}(n)=\int_{-\infty}^{\infty}\psi(\eta)e^{-i2\pi n\eta}d\eta. \tag{2.16}$$

Equation (2.15) is the expression for the WFC in terms of the Fourier amplitudes.

Similarly, it is possible to express the Fourier coefficients, ϵ_n, in terms of the WFC as

$$\epsilon_n=\frac{1}{L}\sum_{j=0}^{\infty}\sum_{l=0}^{2^j-1}\tilde{\epsilon}_{j,l}\hat{\psi}_{j,l}(n),\quad n\neq 0 \tag{2.17}$$

or

$$\epsilon_n=\sum_{j=0}^{\infty}\sum_{l=0}^{2^j-1}\left(\frac{1}{2^jL}\right)^{1/2}\tilde{\epsilon}_{j,l}e^{-i2\pi nl/2^j}\hat{\psi}(n/2^j),\quad n\neq 0. \tag{2.18}$$

Equations (2.15) and (2.18) show that, in principle, the information contained in each transform is equivalent.

C. The wavelet spectrum estimator

We now look at the wavelet power spectrum defined by Eq. (2.10) in more detail. The first thing to note is that the wavelet functions, $\psi_{j,l}(x)$, are localized in Fourier space. For instance, the Fourier transform of the Battle-Lemarié wavelet, which is constructed with fourth order spline functions, is nonzero only in two symmetric narrow ranges centered, respectively, at $n=+1$ and -1 with widths $\Delta n\ll 1$. For the Daubechies 4 (D4) wavelet, $\hat{\psi}(n)$ also has two symmetric peaks centered at $n=\pm n_p$ ($n_p>0$) and of width Δn_p. The sum over n in Eq. (2.15) need only be taken on the two intervals $(n_p-0.5\Delta n_p)2^j\leqslant n\leqslant(n_p+0.5\Delta n_p)2^j$ and $-(n_p+0.5\Delta n_p)2^j\leqslant n\leqslant-(n_p-0.5\Delta n_p)2^j$. Equation (2.15) can be approximately rewritten as

$$\begin{aligned}\tilde{\epsilon}_{j,l}&\simeq\left(\frac{L}{2^j}\right)^{1/2}\left[\hat{\psi}(-n_p)\sum_{n=(n_p-0.5\Delta n_p)2^j}^{(n_p+0.5\Delta n_p)2^j}\epsilon_ne^{i2\pi nl/2^j}+\hat{\psi}(n_p)\sum_{n=-(n_p+0.5\Delta n_p)2^j}^{-(n_p-0.5\Delta n_p)2^j}\epsilon_ne^{i2\pi nl/2^j}\right]\\&=\left(\frac{L}{2^j}\right)^{1/2}\sum_{n=(n_p-0.5\Delta n_p)2^j}^{(n_p+0.5\Delta n_p)2^j}[\hat{\psi}(-n_p)\epsilon_ne^{i2\pi nl/2^j}+\hat{\psi}(n_p)\epsilon_{-n}e^{-i2\pi nl/2^j}].\end{aligned} \tag{2.19}$$

Equation (2.19) shows that the WFC on scale j are mainly determined by the Fourier components δ_n with n centered at

$$n = \pm n_p 2^j, \tag{2.20}$$

where $n = \pm n_p$ are the positions of the peaks of $\hat{\psi}(n)$.

Because both $\psi(x)$ and $\epsilon(x)$ are real, we have $\hat{\psi}(-n_p) = \hat{\psi}^*(n_p)$ and $\epsilon_{-n} = \epsilon_n^*$. Equation (19) then becomes

$$|\tilde{\epsilon}_{j,l}|^2 \simeq \frac{L}{2^j}\left| 2 \sum_{n=(n_p-0.5\Delta n_p)2^j}^{(n_p+0.5\Delta n_p)2^j} \mathrm{Re}\{\hat{\psi}(n_p)\epsilon_n e^{i2\pi nl/2^j}\}\right|^2. \tag{2.21}$$

In the case of Gaussian perturbations, the distribution of the phases of ϵ_n is uniformly random, and Eq. (2.21) reduces to

$$|\tilde{\epsilon}_{j,l}|^2 \simeq \frac{L}{2^{j-1}}|\hat{\psi}(n_p)|^2 \sum_{n=(n_p-0.5\Delta n_p)2^j}^{(n_p+0.5\Delta n_p)2^j} |\epsilon_n|^2. \tag{2.22}$$

According to the central limit theorem, the distribution of ϵ_n will be Gaussian if the density field is a superposition of non-Gaussian perturbations [22]. Therefore, Eq. (2.22) also holds for many kinds of non-Gaussian perturbations. For instance, if the density field consists of a large number of randomly distributed non-Gaussian clumps, the Fourier amplitudes, ϵ_n, will be given by the superposition of a large number of non-Gaussian contributions. The central limit theorem then guarantees that the ϵ_n will be Gaussian, and that the phases of ϵ_n will be uniformly randomly distributed. Even when the clumps are correlated, the central limit theorem still holds as long as the two-point correlation function of the clumps approaches zero sufficiently fast [23].

Moreover, many physically interesting density fields are assumed to be ergodic, that is, the average over an ensemble is equal to the spatial average taken over one realization. A homogeneous Gaussian field with a continuous power spectrum is certainly ergodic [22]. In some non-Gaussian cases, such as homogeneous and isotropic turbulence [1], ergodicity also approximately holds. Roughly, the ergodic assumption 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. This property can effectively be used by DWT because the $\psi_{j,l}(x)$ are orthogonal with respect to the position index l. The 2^j wavelet coefficients for a given j are statistically independent and the $\tilde{\epsilon}_{j,l}$ can be treated as independent realizations. That is, for ergodic random fields the WFC form a valid statistical ensemble on scale j. This ensemble is very useful.

In standard Fourier techniques, it is known that if the density fluctuation field is a homogeneous random process, the average of the Fourier amplitudes over an *ensemble* of fluctuation fields with finite extent (zero outside) will be the same as that over a fluctuation field of infinite extent [22]. Unfortunately, no such ensemble is available if there are only a few (or, as in cosmology, only 1) realizations. However, because of the locality of the DWT basis functions, ensembles are generated at each scale with the DWT decomposition.

Using Eqs. (2.5) and (2.10) in Eq. (2.22) gives

$$P(n)_j \simeq \frac{1}{2^{j+1}\Delta n_p}|\hat{\psi}(n_p)|^{-2} P_j. \tag{2.23}$$

$P(n)_j$ is the average of the Fourier spectrum on scale j given by

$$P(n)_j = \frac{1}{2^j \Delta n_p} \sum_{n=(n_p-0.5\Delta n_p)2^j}^{(n_p+0.5\Delta n_p)2^j} P(n). \tag{2.24}$$

Equations (2.23) and (2.24) are the basic formulas relating the Fourier power spectrum to the DWT power spectrum.

One can also measure the power spectrum of the distribution by using the variance of the $\tilde{\epsilon}_{j,l}$ as

$$P(n)_j \simeq \frac{1}{2^{j+1}\epsilon n_p}|\hat{\psi}(n_p)|^{-2} P_j^{\mathrm{var}} \tag{2.25}$$

and

$$P_j^{\mathrm{var}} = \frac{1}{L}\sum_{l=0}^{2^j-1} (\bar{\tilde{\epsilon}}_{j,l} - \tilde{\epsilon}_{j,l})^2, \tag{2.26}$$

where $\bar{\tilde{\epsilon}}_{j,l}$ is the average of $\tilde{\epsilon}_{j,l}$ over l. Because the mean of WFC $\tilde{\epsilon}_{j,l}$ is zero [Eq. (2.15)], P_j is statistically equal to P_j^{var}.

D. Comparison of DWT spectrum estimator to the CIC

To demonstrate the advantage of a basis function that is orthogonal to scale we compare the DWT spectrum estimator to the CIC technique. The most direct way of describing the differences between the two is by noting that the cubic cell of the CIC corresponds to a window of the form

$$\phi(\eta) = \begin{cases} 1 & \text{if} \quad 0 \leqslant \eta < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.27}$$

The corresponding basis function is

$$\psi(\eta) = \begin{cases} 1 & \text{if} \quad 0 \leqslant \eta < 1/2 \\ -1 & \text{if} \quad 1/2 \leqslant \eta < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.28}$$

This is the lowest order Daubechies wavelet, known as the Haar wavelet whose Fourier transform is

$$\hat{\psi}(n) = \frac{2}{\pi n}[\sin(\pi n) - i\cos(\pi n)]\sin^2(\pi n/2). \tag{2.29}$$

When $n \ll 1$, $\hat{\psi}(n) \sim -i(\pi/2)n$. Therefore, Eq. (2.29) is not localized in Fourier space. In the case of a power law spectrum, $\epsilon_n = Kn^\gamma$, Eq. (2.15) gives

$$\tilde{\epsilon}_{j,l}=\sum_{n<2^j}\left(\frac{2^j}{L}\right)^{-1/2}\epsilon_n\hat{\psi}(-n/2^j)e^{i2\pi nl/2^j}+\text{terms}\quad n\geqslant 2^j$$

$$=\sum_{n<2^j}\left(\frac{2^j}{L}\right)^{-1/2}\frac{i\pi A}{2^{j+1}}n^{1+\gamma}e^{i2\pi nl/2^j}+\text{terms}\quad n\geqslant 2^j. \tag{2.30}$$

The square average $\tilde{\epsilon}_{j,l}$, or similarly the CIC on scale j will be contaminated by perturbations on scales larger than j, and the Haar wavelet WFC (or CIC) on scale j will no longer be a good measure of spectrum on scale l.

III. SPECTRA RECONSTRUCTION

A. A simple example: Power law spectra

We now demonstrate the ability of the DWT spectrum estimator, i.e., Eqs. (2.23) or (2.25), to reconstruct power spectra. We first consider a power law spectrum given by $P(k)\sim k^{\gamma}$. From Eqs. (2.23) and (2.24), we have

$$P(n)_{j+1}=2^{\gamma}P(n)_j \tag{3.1}$$

and

$$P_j\propto 2^{j(\gamma+1)} \tag{3.2}$$

or

$$\log_2 P_j=(\gamma+1)j+\text{const.} \tag{3.3}$$

The slope of $\log_2 P_j$, when plotted against j, is $\gamma+1$. The index of a power law can be directly detected by

$$\gamma=\frac{d\log_2 P_j}{dj}-1. \tag{3.4}$$

Figure 1 shows a DWT reconstructed power law spectrum. The random density field is generated from perturbations with power $P(k)=k^{-2}$, where $k\equiv 2\pi n/L$ is the wave number in length units. The sample is distributed over $2^9=512$ bins. Figure 1(b) shows the spectra $\log_2 P_j$ and $\log_2 P_j^{\rm var}$ plotted against j. The points of $\log_2 P_j^{\rm var}$ in Fig. 1(b) have been shifted down to $\log_2 P_j^{\rm var}-1$ for presentation purposes only. As expected, Fig. 1(b) shows (1) P_j is equal to $P_j^{\rm var}$ and (2) the slopes of the lines $\log_2 P_j$ vs j or $\log_2 P_j^{\rm var}$ vs j are equal to -1, giving a spectral index $\gamma=-2$. The 1σ error bars are for 100 realizations.

B. Normalization factors

In the previous section the spectrum index of a power law was measured. It is equally important to estimate the amplitude of the power spectrum, $\log P(k)_j$, from $\log P_j$. This can be done from Eq. (2.24), which gives

$$\log P(k)_j=\log P_j-(\log 2)j+A, \tag{3.5}$$

where

$$A=-\log[2\Delta n_p|\hat{\psi}(n_p)|^2]. \tag{3.6}$$

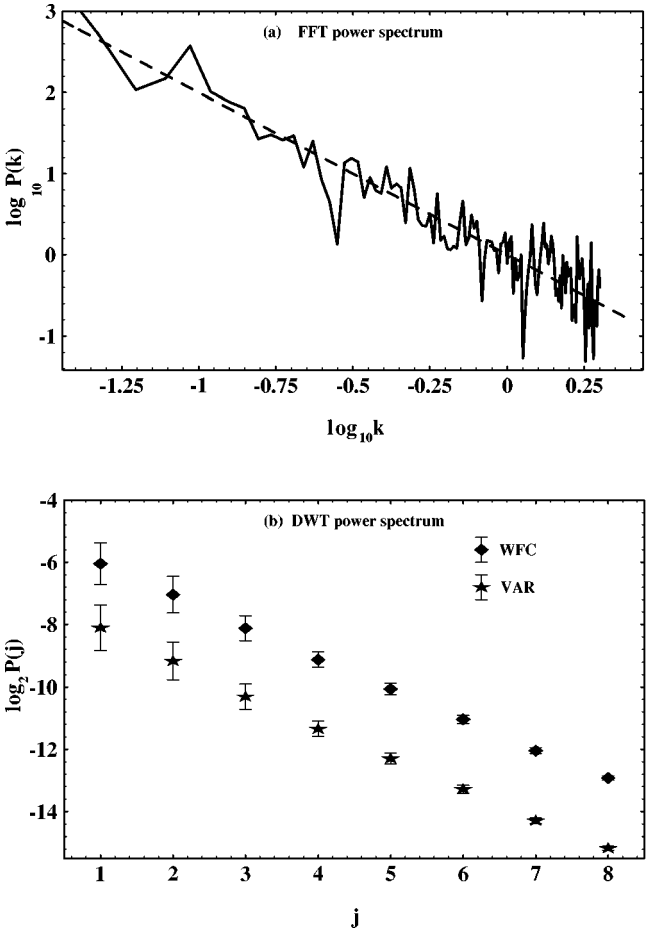

FIG. 1. Reconstructed power spectrum for a distribution generated from a power law $P(k)=k^{-2}$. (a) Reconstructed spectrum from FFT detection, and (b) reconstructed spectrum from the DWT estimator, where $k=2\pi n/L$ is the wave number in a length unit. The points of $\log_2 P^{\rm var}$ (star) in (b) have been shifted down to $\log_2 P_j^{\rm var}-1$ for clarity of presentation only. The slopes of the lines $\log_2 P_j-j$ (diamond) or $\log_2 P_j^{\rm var}-j$ are -1, indicating a spectral index of -2.

If A is known, $\log P(k)_j$ can be found from $\log P_j$. A normalizes the wavelet amplitude to the Fourier amplitude and is dependent on the choice of wavelet. In the case of the D4 wavelet, $A=0.602$.

From Eq. (2.20), we have

$$\log k=(\log 2)j-\log L/2\pi+B, \tag{3.7}$$

where

$$B=\log n_p. \tag{3.8}$$

B normalizes scale j to $\log k$. For the D4 wavelet basis, $B=0.270$. Equations (3.5) and (3.7) transfer P_j or $P_j^{\rm var}$ into the mean Fourier spectrum $P(k)_j$, and *vice versa*. There is no particular reason to normalize the DWT estimator to the Fourier power spectrum. However, we do so here to confirm the results obtained using the DWT estimator with the well established Fourier methods and to relate the wavelet scale parameter j to the Fourier wave number k, which has a better understood physical interpretation.

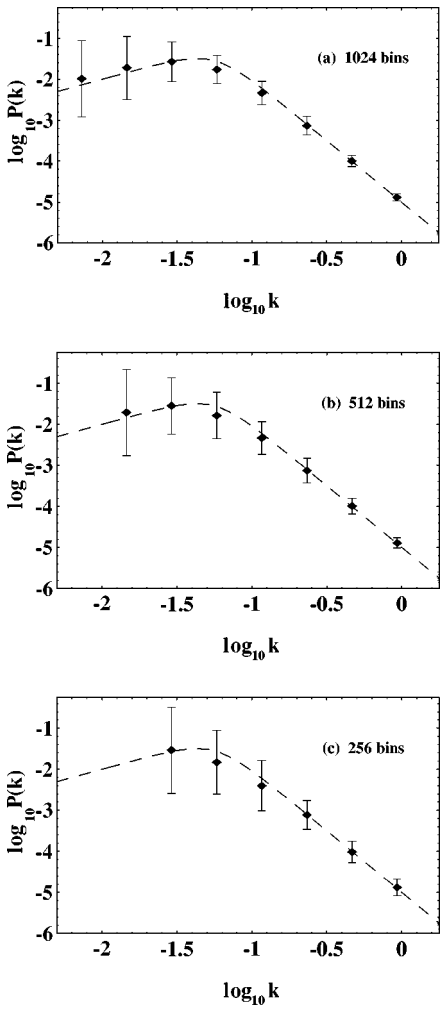

FIG. 2. Reconstructed spectrum for a distribution generated from Eq. (3.9). $k=2\pi n/L$ is the wave number in length units. The dotted line is the theoretical curve and the diamond points are the DWT determined spectrum. The samples in (a), (b), and (c) are generated in (a) 1024, (b) 512, and (c) 256 bins, and the bin size is 2π length units. The peak of the spectrum is at $\log k=-1.37$.

To test these normalizations the following spectrum was used:

$$P(k)=\frac{k}{1+10^5k^4}. \tag{3.9}$$

This has a peak at $\log k\sim-1.37$, or a typical scale $1/k=23.4$ (length) units. Using Eq. (3.9), samples of length L were generated into 256, 512, and 1024 bins. The reconstruction of the spectrum is shown in Fig. 2. The peak and the amplitude of the power spectrum are perfectly detected by the DWT. Either P_j or $P_j^{\rm var}$ can effectively provide information on the shape of the spectrum as well as its amplitude.

C. Boundary conditions

The WFC measure fluctuations from the *local* mean density. This locality property allows the $\tilde{\epsilon}_{j,l}$ to be independent of data outside an ''influence'' cone. That is, if ψ is well localized in the interval Δx, then the wavelet coefficients

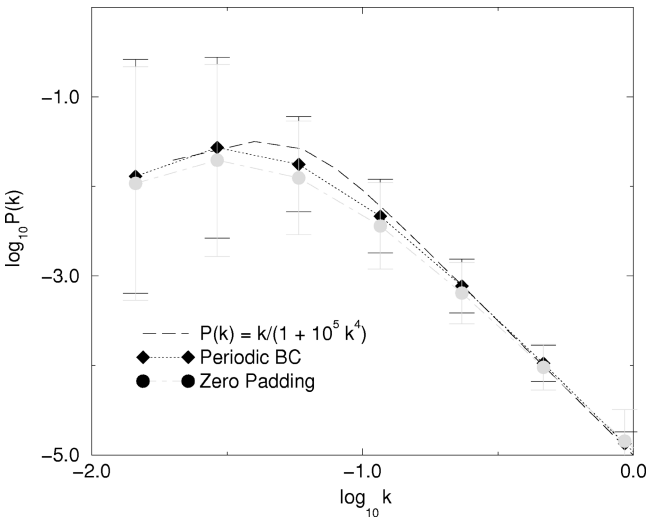

FIG. 3. Reconstruction of a typical scale spectrum [Eq. (3.9) and Fig. 2]. $k=2\pi n/L$ is the wave number in length units. Shown in the figure is the original distribution [Eq. (3.9)] without any boundary conditions (dotted line) and the power spectrum determined using periodic boundary conditions and zero padding.

$\tilde{\epsilon}_{j,l}$, corresponding to position x_0, will only measure fluctuations in the interval $[(x_0-\Delta x)/2^{j+1},(x_0+\Delta x)/2^{j+1}]$ [16]. As a result the DWT spectrum estimator is less sensitive to boundary conditions.

To test this point, spectrum (3.9) was used to generate samples over a finite length L in 512 bins. Two different boundary conditions were studied: (A) periodic boundary conditions; (B) zero padding. Figure 3 shows the results of reconstructing the power spectrum from distributions generated using these boundary conditions. The results show that the spectrum can be correctly reconstructed by the DWT regardless of the boundary conditions.

A similar test is shown in Fig. 4, in which the density distribution over 512 bins is generated from a power law spectrum $P(k)=k^{-2}$, and then the last quarter bins set to zero. When the spectrum is determined using the fast Fourier transform (FFT), the usual procedure is to convolve the Fourier coefficients with a window function. The spectrum shown in Fig. 4 was computed using a trianglelike window function. The spectra are reconstructed with window sizes, 4, 8, and 16 bins, and the power law index is found to be -1.73, -1.93, and -2.14, respectively. As shown, the power law index from the FFT reconstruction sensitively depends on the window size even within a specific choice of a window function. The ''arbitrary'' selection of the window size leads to an uncertainty in the index. On the other hand, for a given wavelet, there is no arbitrariness in the selection of the window size. The wavelet basis functions on various scales are constructed by the requirements of orthonormality and completeness. The DWT reconstructed power spectrum gives the correct values for the spectrum index, -1.86.

D. Sampling

That the DWT estimator is less sensitive to sampling can already be seen in Fig. 2, in which the spectrum is correctly reconstructed regardless the number of bins. Binning can be thought of as a kind of sampling.

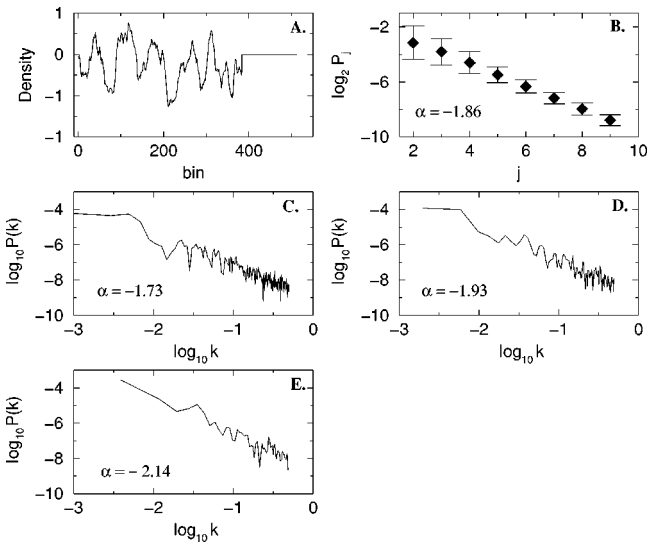

FIG. 4. Reconstruction of power law spectrum $P(k)=k^{\alpha}$ with $\alpha=-2$. Shown in the figure are (a) the original distribution with last quarter bins set to zero, (b) the DWT power spectra from $\log_2 P_j$, (c), (d), and (e) the spectrum reconstructed by the window Fourier transform with window size 4, 8, and 16 bins, respectively. Once again, $k=2\pi n/L$ is the wave number in length units.

A more realistic example of sampling is the identification of objects from a density field. For instance, one can identify overdense regions in the density field as objects. This new distribution is a sampling of the original density field. Obviously, the new distribution will not be distributed in the same way as the original density field. However, it is known that distributions constructed in this way are not different from the original field on scales larger than the characteristic scale of the sampling [1]. One can expect that a proper spectrum estimator will be able to reconstruct the original spectrum from sampled data on scales larger than the characteristic scale. Yet, different criteria for identifying objects, i.e., different samplings, will give a different mean density. This inevitably leads to the sampling problem for spectrum estimators depending on the mean density.

Moreover, for real data, it is often necessary to sample data in 1D and then try to reconstruct the 3D power spectrum. If the density perturbations in 3D are statistically isotropic with a spectrum $P_3(k)$, we have

$$P(k)=2\pi\int_k^{\infty}P_3(q)q\,dq. \tag{3.10}$$

This shows that the contributions from long wavelength perturbations, or wave number less than k, are not important. If the 1D spectrum can be approximated as a power law,

$$P(k)\propto k^{-\alpha}, \tag{3.11}$$

with $\alpha>0$, the 3D spectrum can be reconstructed from Eq. (3.10) as

$$\log P_3(k)=\log P(k)-2\log k+\log(\alpha/2\pi). \tag{3.12}$$

Using the DWT spectrum P_j to replace $P(k)$, we have

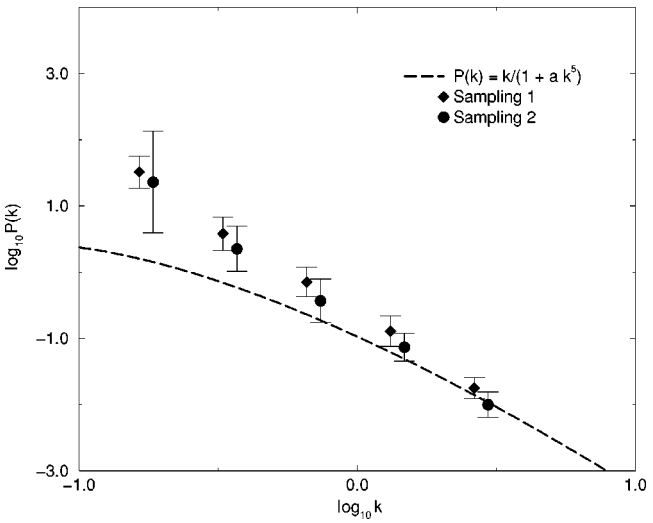

FIG. 5. Reconstruction of the 3D power spectrum for samples with samplings (1) and (2). k is now in inverse length units.

$$\log P_3(k)=\log P_j-2\log k-(\log 2)j+A+\log(\alpha/2\pi). \tag{3.13}$$

Figure 5 shows an example, in which the 3D spectrum is given by Eq. (3.9). 20 1D distributions are generated from the 3D perturbations. The density fields are sampled by identifying the peaks in the density field as objects. Two different criteria are used. In sampling (1) the number density is greater than that in sample (2) by a factor of between 2 and 3.

Figure 5 shows that the 3D spectra reconstructed from samplings 1 and 2 are the same. Figure 5 also shows that the reconstructed spectra generally agree with the original spectrum, but have a faster increase than the theoretical spectrum as the scale increases, especially on scales close to $\log k=-0.75$. This is because we have used a simple power law to describe the entire 1D spectrum [Eq. (3.11)]. The long wavelength perturbations are overestimated by a power law approximation of the 1D spectrum.

IV. AN APPLICATION: THE Ly-α ABSORPTION CLOUDS

A. Problem in Ly-α absorption cloud detection

Ly-α absorption line systems in QSO spectra come from intervening neutral hydrogen absorbers, or clouds. The distribution of the Ly-α absorption lines in redshift space is a direct measure of the spatial distribution of neutral hydrogen clouds in the Universe [24].

Our goal here is not to study the physics of the formation of the clouds, but to show that the power spectrum of the mass distribution traced by the hydrogen clouds can be detected by the DWT estimator.

The Ly-α lines are much more numerous than other high redshift objects, tend to be relatively less affected by selection effects, and cover a large range in redshift or spatial space. These objects should serve as good candidates for studying the cosmic mass distribution on large scales. Yet, so far no power spectrum has been detected, especially on scales larger than $5\,h^{-1}$ Mpc, where h is the Hubble constant in units of 100 km s^{-1} Mpc^{-1}. This particular data set pre-

sents several problems in detecting the spectrum using traditional methods.

The first problem is that the mean number density of the Ly-α clouds significantly varies with redshift. This variation can be approximately described as

$$\frac{dN}{dz}=\left(\frac{dN}{dz}\right)_0(1+z)^{\gamma}, \tag{4.1}$$

where $(dN/dz)_0$ is the number density extrapolated to zero redshift, and $\gamma\sim 2$ is the evolution index. In other words, no mean density is available for calculating the two-point correlation function. This makes detection of Ly-α absorption clouds by the two point correlation function extremely difficult [24,25].

The second problem is the complex geometry of the Ly-α absorption clouds. Ly-α absorption line samples from different QSO spectra distribute over different ranges in redshift space, or different ranges in physical space. Moreover, the total number of the Ly-α lines are different for different forest samples. How do we obtain estimates of the spectrum from these very irregular samples?

As discussed above, both problems can be solved by the DWT spectrum estimator. One can regularize all samples into the same range, say $(D_{\min},D_{\max})$, in physical space as follows. For a forest sample in a spatial range (D_1,D_2), one can extend it to a larger range $(D_{\min},D_{\max})$ by adding zero to the data in ranges $(D_{\min},D_1)$ and $(D_2,D_{\max})$. Thus, all samples distribute on $(D_{\min},D_{\max})$. Since wavelets are localized in both configuration and Fourier spaces, the wavelet transform in an interval (D_1,D_2) will not be affected by the addition of data in the regions $(D_{\min},D_1)$ and $(D_2,D_{\max})$.

B. Results

In order to check whether the above-mentioned problems are really surmounted by the DWT estimator, two popular and independent data sets of the Ly-α absorption clouds were studied. The first was compiled by Lu, Wolfe, and Turnshek ([26] hereafter LWT). It contains $\sim$950 lines from the spectra of 38 QSOs that exhibit neither broad absorption line nor metal line systems. The second is from Bechtold ([27], hereafter JB), which contains a total $\sim$2800 lines from 78 QSOs spectra, in which 34 high redshift QSOs were observed at moderate resolution. These two data sets are sampled differently. The LWT samples contain only lines whose equivalent width W is $\geqslant$0.36 Å. JB contains samples with linewidths $W>0.16$ Å and $W>0.32$ Å. Their redshift dependences are also different: LWT showed that $(dN/dz)_0\simeq 3$ and $\gamma=2.75\pm 0.29$ for lines with $W>0.36$ Å while JB found $\gamma=1.89\pm 0.28$ for $W>0.32$ Å and $\gamma=1.32\pm 0.24$ for $W_{\rm th}>0.16$ Å.

Because the linewidth W depends only on the local environment, statistical features, such as the power spectrum, on large scales should be the same for the distributions sampled by linewidths $W>0.36$, 0.32, 0.16 Å.

All samples were regularized as described above into range of $D_{\min}=2{,}300\ h^{-1}$ Mpc to $D_{\max}=3{,}300\ h^{-1}$ Mpc, which corresponds to a redshift range of 1.7–4.1 in a flat universe. The 1D spectra for LWT ($W>0.36$ Å), JB ($W>0.32$ Å), and JB ($W>0.16$ Å) are plotted in Fig. 6. The error bars are calculated from the ensemble of the QSOs absorption spectra.

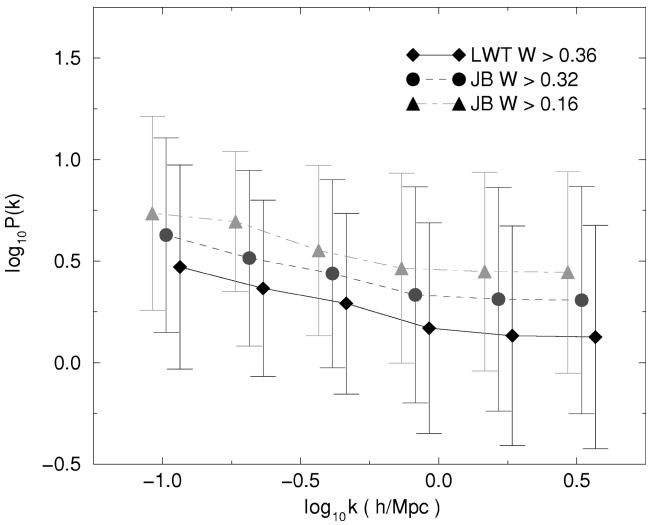

FIG. 6. The 1D power spectrum for samples of LWT $W>0.36$ Å, JB $W>0.32$ Å, and JB $W>0.16$ Å. For clarity, the spectra for JB $W>0.32$ Å and JB $W>0.16$ Å have been shifted in the negative k direction.

The 3D spectra can be constructed from the 1D spectra of the LWT and JB samples as done in Eq. (3.13). The results are shown in Fig. 7. As in Fig. 5, because the 1D spectra are approximated as a power law, the reconstructed 3D spectra should be a good estimate on small scales, but will have large deviations on large scales.

The data sets contain large errors, and therefore, the DWT estimated spectra contain uncertainties. Nevertheless, some conclusions can already be drawn. Figures 6 and 7 clearly show that the amplitude and shape of the DWT estimated 1D and 3D spectra for all data sets are very similar. This strongly implies that these features are common properties of the spectra of the Ly-α absorption clouds. That is, these features should be given by the physics of the Ly-α cloud

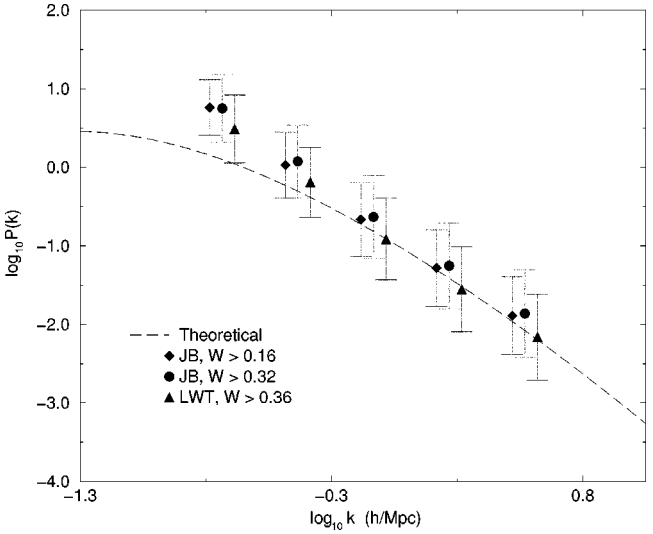

FIG. 7. The 3D power spectrum reconstructed from data LWT $W>0.36$ Å, JB $W>0.32$ Å, and JB $W>0.16$ Å. The theoretical curve is the standard cold-dark matter model spectrum.

formation, and not by the geometry or by the sampling of the data.

The errors at large scale are about the same as that on small scales. This means that the DWT estimator can uniformly detect the spectrum on almost all scales on which data are available. This cannot be achieved by the two-point correlation function, which generally overlooks structures on large scales because autocorrelation amplitudes on large scales are small, and structure cannot be detected when its amplitude is comparable with the uncertainty of the mean density.

V. CONCLUSION

Throughout most of this work the conventional definition of the spectrum was used, i.e., the spectrum with respect to the Fourier basis. There is no particular reason to do this and one can define the spectrum with respect to any complete and orthonormal basis.

Theoretically, the spectrum based on any complete orthonormal basis decomposition is equivalent as long as there is an ensemble of realizations of the random density field. Practically, they may not be completely equivalent because real data are often incomplete as it is constrained by its finite size and, additionally, an ensemble of realizations may not be available. Thus, different bases might measure different aspects of the density fields.

A good way to see this is to look at how different bases *chop* up phase space. In phase space, all orthonormal bases decompose the space into elements with product $\Delta x\ \Delta k \geqslant 2\pi$. However, different bases distribute the product differently. The ordinary Fourier transform has $\Delta k \to 0, \Delta x \to \infty$, windowed Fourier transforms have $\Delta x = \Delta k = \text{const}$, while for the wavelet basis, the elements $\Delta x, \Delta k$ adapt to the scale being analyzed [16]. The wavelet basis is the best choice if the localization of both k and x is crucial.

Indeed, we have shown that the DWT spectrum estimator is an efficient and reliable tool for detecting the spectrum of density perturbations from samples with a complex geometry. We have also shown that the DWT can estimate the power spectrum regardless of how the underlying density field is sampled. We applied the DWT estimator to real samples of the Ly-α absorption clouds, for which the spectrum has not been previously detected. The results show that the power spectrum of Ly-α absorption clouds can stably be detected on scales 2–100 h^{-1} Mpc by the DWT estimator. Therefore, the discrete wavelet transform power spectrum estimator is an important and necessary supplement to currently existing spectrum estimators.

[1] E. H. Vanmarke, *Random Fields, Analysis and Synthesis* (MIT Press, Cambridge, MA, 1983).

[2] F. Takagi, Phys. Rev. Lett. **53**, 427 (1984); Phys. Rev. C **32**, 1799 (1985).

[3] W. Saunders, C. Frenk, M. Rowen-Robinson, G. Efstathiou, A. Lawrence, N. Kaiser, R. Ellis, J. Crawford, X. Xia, and L. Parry, Nature (London) **349** , 32 (1991).

[4] I. Daubechies, *Ten Lectures on Wavelets* (SIAM, Philadelphia, 1992).

[5] C. K. Chui, *Wavelets: A Tutorial in Theory and Applications* (Academic Press, Boston, 1992).

[6] Y. Meyer, *Wavelets: Algorithms and Applications* (SIAM, Philadelphia, 1993).

[7] G. Kaiser, *A Friendly Guide to Wavelets* (Birkhäuser, Boston, 1994).

[8] L. Z. Fang and Jesús Pando, *The 5th Current Topics of Astrofundamental Physics*, edited by N. Sanchez and A. Zichichi (World Scientific, Singapore, 1997).

[9] C. W. Thertien, *Discrete Random Signals and Statistical Signal Processing* (Prentice-Hall, Englewood Cliffs, 1992).

[10] H. V. Poor, *An Introduction to Signal Detection and Estimation* (Springer, New York, 1994).

[11] J. L. Lumley, *Transition and Turbulence*, edited by R. E. Meyer (Academic, New York, 1981).

[12] N. Aubry, P. Holmes, J. L. Lumley, and E. Stone, J. Fluid Mech. **192**, 115 (1988).

[13] M. V. Wickerhauser, *Adapted Wavelet Analysis From Theory to Software* (A.K. Peters, Wellesley, MA, 1994).

[14] P. Carruthers, in *Proceedings of the XXIII International Symposium on Multiparticle Dynamics, Aspen, Colorado* (World Scientific, Singapore, 1995).

[15] J. Pando, M. Greiner, P. Lipa, and L. Z. Fang, e-print astro-ph/9710204.

[16] M. Farge, Annu. Rev. Fluid Mech. **24**, 395 (1992).

[17] M. Yamada and K. Ohkitani, Prog. Theor. Phys. **86**, 799 (1991).

[18] M. Greiner, P. Lipa, and P. Carruthers, Phys. Rev. E **51**, 1948 (1995).

[19] J. Pando and L. Z. Fang, Astrophys. J. **459**, 1 (1996).

[20] Z. Huang, I. Sarcevic, R. Thews, and X. N. Wang, Phys. Rev. D **54**, 750 (1996).

[21] M. Greiner, J. Giesemann, P. Lipa, and P. Carruthers, Z. Phys. C **69**, 305 (1996).

[22] R. J. Adler, *The Geometry of Random Fields* (Wiley, New York, 1981).

[23] Z. H. Fan and J. M. Bardeen, Phys. Rev. D **51**, 6714 (1995).

[24] R. J. Weyman, *The Environment and Evolution of Galaxies* (Kluwer Academic, Boston, 1993).

[25] J. K.Webb, *Observational Cosmology*, edited by A. Hewitt, G. Burbidge, and L. Z. Fang (Dordrecht, Boston, 1987).

[26] L. Lu, A. M. Wolfe, and D. A. Turnshek, Astrophys. J. **367**, 19 (1991).

[27] J. Bechtold, Astrophys. J., Suppl. **91**, 1 (1994).