The Astrophysical Journal, 620:559–563, 2005 February 20

CORRECTING FOR THE ALIAS EFFECT WHEN MEASURING THE POWER SPECTRUM USING A FAST FOURIER TRANSFORM

Y. P. Jing
Shanghai Astronomical Observatory, The Partner Group of the Max-Planck-Institut für Astrophysik, Nandan Road 80, Shanghai 200030, China
Received 2004 September 10; accepted 2004 October 28

ABSTRACT

Because of mass assignment onto grid points in the measurement of the power spectrum using a fast Fourier transform (FFT), the raw power spectrum $\langle|\delta^f(k)|^2\rangle$ estimated with the FFT is not the same as the true power spectrum $P(k)$. In this paper we derive a formula that relates $\langle|\delta^f(k)|^2\rangle$ to $P(k)$. For a sample of N discrete objects, the formula reads $\langle|\delta^f(k)|^2\rangle = \sum_{\boldsymbol{n}}[|W(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n})|^2 P(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n}) + 1/N|W(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n})|^2]$, where $W(\boldsymbol{k})$ is the Fourier transform of the mass assignment function $W(\boldsymbol{r})$, $k_{\rm N}$ is the Nyquist wavenumber, and $\boldsymbol{n}$ is an integer vector. The formula is different from that in some previous works in which the summation over $\boldsymbol{n}$ is neglected. For the nearest grid point, cloud-in-cell, and triangular-shaped cloud assignment functions, we show that the shot-noise term $\sum_{\boldsymbol{n}}(1/N)|W(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n})|^2$ can be expressed by simple analytical functions. To reconstruct $P(k)$ from the alias sum $\sum_{\boldsymbol{n}}|W(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n})|^2 P(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n})$, we propose an iterative method. We test the method by applying it to an N-body simulation sample and show that the method can successfully recover $P(k)$. The discussion is further generalized to samples with observational selection effects.

Subject headings: galaxies: clusters: general — large-scale structure of universe — methods: data analysis — methods: statistical

1. INTRODUCTION

The power spectrum $P(k)$ of the spatial cosmic density distribution is an important quantity in galaxy formation theories. On large scales, $P(k)$ is a direct measure of the primordial density fluctuation, and $P(k)$ on small scales carries information on later nonlinear evolution; therefore, measuring $P(k)$ can serve to distinguish between different theoretical models.

The power spectrum $P(k)$ as a clustering measure has already been applied by many authors to observational samples of galaxies and clusters of galaxies, including the Center for Astrophysics (CfA) and Perseus-Pisces redshift surveys (Baumgart & Fry 1991), a radio galaxy survey (Peacock & Nicholson 1991), the *IRAS* QDOT survey (Kaiser 1991), the galaxy distribution in nearby superclusters (Gramann & Einasto 1992), the Southern Sky Redshift Survey (SSRS) sample (Park et al. 1992), the CfA and SSRS extensions (Vogeley et al. 1992; da Costa et al. 1994), the 2 Jy *IRAS* survey (Jing & Valdarnini 1993), the 1.2 Jy *IRAS* survey (Fisher et al. 1993), and redshift samples of Abell clusters (Jing & Valdarnini 1993; Peacock & West 1992). The power spectrum is also widely measured for cosmological N-body simulations, since it can easily characterize the linear and nonlinear evolution of the density perturbation (e.g., Davis et al. 1985).

All these workers except Fisher et al. (1993), have used the fast Fourier transformation (FFT) technique to make the Fourier transforms (FTs). In fact, one can use the direct summation (eq. [5] below) to measure the power spectrum for a sample of a few thousand objects (Fisher et al. 1993). However, it appears to be more convenient for most workers to use FFT packages (now available on many computers) to analyze the power spectrum. This is probably the reason why most of the previous statistical studies of the power spectrum have used FFT. For N-body simulations, one has to use FFT to obtain the power spectrum, since normally there are more than a million particles. When using FFT, one needs to collect density values $\rho(\boldsymbol{r}_g)$ on a grid from a density field $\rho(\boldsymbol{r})$ (or a particle distribution), which is usually called "mass assignment." The mass assignment is equivalent to convolving the density field by a given assignment function $W(\boldsymbol{r})$ and sampling the convolved density field on a finite number of grid points. The FT of $\rho(\boldsymbol{r}_g)$ generally is not equal to the FT of $\rho(\boldsymbol{r})$, and the power spectrum estimated directly from the FT of $\rho(\boldsymbol{r}_g)$ is a biased one. The smoothing effect has already been considered in many previous works (e.g., Baumgart & Fry 1991; Jing & Valdarnini 1993; Scoccimarro et al. 1998), but the sampling effect has not.

In this paper we derive the formulae that express these effects of the mass assignment on the estimated power spectrum. A procedure is proposed to correct for these effects, in order to recover the true power spectrum. The procedure is tested and shown to be very successful with a simulation sample.

2. FORMULAE

Let us first recall the definition of the power spectrum $P(k)$. Let $\rho(\boldsymbol{r})$ be the cosmic density field and $\bar{\rho}$ the mean density. The density field can be expressed with a dimensionless field $\delta(\boldsymbol{r})$ (which is usually called the density contrast):

$$\delta(\boldsymbol{r}) = \frac{\rho(\boldsymbol{r}) - \bar{\rho}}{\bar{\rho}}. \tag{1}$$

Based on the cosmological principle, one can imagine that $\rho(\boldsymbol{r})$ is periodic in some large rectangular volume V_μ. The FT of $\delta(\boldsymbol{r})$ is then defined as

$$\delta(\boldsymbol{k}) = \frac{1}{V_\mu}\int_{V_\mu} \delta(\boldsymbol{r}) e^{i\boldsymbol{r}\cdot\boldsymbol{k}}\, d\boldsymbol{r}, \tag{2}$$

and the power spectrum $P(k)$ is simply related to $\delta(\boldsymbol{k})$ by

$$P(k) \equiv \left\langle |\delta(\boldsymbol{k})|^2 \right\rangle, \tag{3}$$

where $\langle \ldots \rangle$ means the ensemble average.

In a practical measurement of $P(k)$ either for an extragalactic catalog or for simulation data, the continuous density field $\rho(\boldsymbol{r})$ is sampled by only a finite number N of objects. In these cases one needs to deal with the "discreteness" effect arising from the Poisson shot noise. To show this, let us consider an ideal case in which no selection effect has been introduced into the sample. (In fact, some kind of selection effects must exist in extragalactic catalogs. We discuss this in § 3.) The number density distribution of objects can be expressed as $n(\boldsymbol{r}) = \sum_j \delta^{\rm D}(\boldsymbol{r} - \boldsymbol{r}_j)$, where $\boldsymbol{r}_j$ is the coordinate of object j and $\delta^{\rm D}(\boldsymbol{r})$ is the Dirac δ-function. In analogy with the continuous case, the FT of $n(\boldsymbol{r})$ is defined as

$$\delta^d(\boldsymbol{k}) = \frac{1}{V_\mu \bar{n}} \int_{V_\mu} n(\boldsymbol{r}) e^{i\boldsymbol{r}\cdot\boldsymbol{k}}\, d\boldsymbol{r} - \delta^{\rm K}_{\boldsymbol{k},0}, \tag{4}$$

where $\bar{n}$ is the global mean number density, the superscript d represents the discrete case of $\rho(\boldsymbol{r})$, and $\delta^{\rm K}$ is the Kronecker delta. Following Peebles (1980, §§ 36–41), we divide the volume V_μ into infinitesimal elements $\{dV_i\}$ with n_i objects inside dV_i. Then the above equation can be written as

$$\delta^d(\boldsymbol{k}) = \frac{1}{N} \sum_i n_i e^{i\boldsymbol{r}_i\cdot\boldsymbol{k}} - \delta^{\rm K}_{\boldsymbol{k},0}, \tag{5}$$

where N is $\bar{n} V_\mu$, the number of objects in V_μ. Since dV_i is taken so small that n_i is either 0 or 1, we have $n_i = n_i^2 = n_i^3 = \ldots$, $\langle n_i \rangle = \bar{n}\, dV_i$, and $\langle n_i n_j \rangle_{i\neq j} = \bar{n}^2\, dV_i\, dV_j\, [1 + \langle \delta(\boldsymbol{r}_i)\delta(\boldsymbol{r}_j)\rangle]$. We can find the ensemble average of $\delta^d(\boldsymbol{k}_1)\delta^{d*}(\boldsymbol{k}_2)$:

$$\begin{aligned}
\left\langle \delta^d(\boldsymbol{k}_1)\delta^{d*}(\boldsymbol{k}_2) \right\rangle &= \frac{1}{N^2} \sum_{i,j} \left\langle n_i n_j \right\rangle e^{i\boldsymbol{r}_i\cdot\boldsymbol{k}_1 - i\boldsymbol{r}_j\cdot\boldsymbol{k}_2} - \delta^{\rm K}_{\boldsymbol{k}_1,0}\delta^{\rm K}_{\boldsymbol{k}_2,0} \\
&= \frac{1}{N^2} \sum_{i\neq j} \left\langle n_i n_j \right\rangle e^{i\boldsymbol{r}_i\cdot\boldsymbol{k}_1 - i\boldsymbol{r}_j\cdot\boldsymbol{k}_2} \\
&\quad + \frac{1}{N^2} \sum_i \langle n_i \rangle e^{i\boldsymbol{r}_i\cdot(\boldsymbol{k}_1 - \boldsymbol{k}_2)} - \delta^{\rm K}_{\boldsymbol{k}_1,0}\delta^{\rm K}_{\boldsymbol{k}_2,0} \\
&= \langle \delta(\boldsymbol{k}_1)\delta^*(\boldsymbol{k}_2)\rangle + \frac{1}{N}\delta^{\rm K}_{\boldsymbol{k}_1,\boldsymbol{k}_2}.
\end{aligned} \tag{6}$$

The last equation assumes $\boldsymbol{k}_1 \neq 0$ or $\boldsymbol{k}_2 \neq 0$. The true power spectrum is then

$$P(k) \equiv \left\langle |\delta(\boldsymbol{k})|^2 \right\rangle = \left\langle \left|\delta^d(\boldsymbol{k})\right|^2 \right\rangle - \frac{1}{N}. \tag{7}$$

Therefore, the discreteness (or shot noise) effect is to introduce an additional $1/N$ term to the power spectrum $\langle|\delta^d(\boldsymbol{k})|^2\rangle$. This fact is already well known to cosmologists. We present the above derivation because this method is useful in the following derivations.

In principle one can use the direct summation of equation (5) to measure the power spectrum for a sample of discrete objects. However, as described in § 1, most of the previous statistical studies of the power spectrum have used an FFT. Moreover, it would be impossible to use direct summation to measure the power spectrum for an N-body simulation. The quantity computed by the FFT is

$$\delta^f(\boldsymbol{k}) = \frac{1}{N} \sum_{\boldsymbol{g}} n^f(\boldsymbol{r}_g) e^{i\boldsymbol{r}_g\cdot\boldsymbol{k}} - \delta^{\rm K}_{\boldsymbol{k},0}, \tag{8}$$

where the superscript f denotes quantities in the FFT, $n^f(\boldsymbol{r}_g)$ is the convolved density value on the $\boldsymbol{g}$th grid point $\boldsymbol{r}_g = \boldsymbol{g}H$ ($\boldsymbol{g}$ is an integer vector; H is the grid spacing),

$$n^f(\boldsymbol{r}_g) = \int n(\boldsymbol{r}) W(\boldsymbol{r} - \boldsymbol{r}_g)\, d\boldsymbol{r}, \tag{9}$$

and $W(\boldsymbol{r})$ is the mass assignment function.

Following Hockney & Eastwood (1981), equation (8) can be expressed in a more compact way by using the so-called sampling function. The sampling function $\Pi(\boldsymbol{r})$ is defined as the sum of the Dirac δ-functions spaced at unit length in all three spatial directions, i.e., $\Pi(\boldsymbol{r}) = \sum_{\boldsymbol{n}} \delta^{\rm D}(\boldsymbol{r} - \boldsymbol{n})$, where $\boldsymbol{n}$ is an integer vector. Defining

$$n'^f(\boldsymbol{r}) \equiv \Pi\left(\frac{\boldsymbol{r}}{H}\right) \int n(\boldsymbol{r}_1) W(\boldsymbol{r}_1 - \boldsymbol{r})\, d\boldsymbol{r}_1 \tag{10}$$

and constructing

$$\delta'^f(\boldsymbol{k}) = \frac{1}{N} \int n'^f(\boldsymbol{r}) e^{i\boldsymbol{r}\cdot\boldsymbol{k}}\, d\boldsymbol{r} - \delta^{\rm K}_{\boldsymbol{k},0}, \tag{11}$$

one can easily prove

$$\delta'^f(\boldsymbol{k}) = \delta^f(\boldsymbol{k}). \tag{12}$$

Therefore, one can express $\delta^f(\boldsymbol{k})$ as (see eqs. [10]–[12])

$$\delta^f(\boldsymbol{k}) = \frac{1}{N} \int_{V_\mu} \Pi\left(\frac{\boldsymbol{r}}{H}\right) \sum_i n_i W(\boldsymbol{r}_i - \boldsymbol{r}) e^{i\boldsymbol{r}\cdot\boldsymbol{k}}\, d\boldsymbol{r} - \delta^{\rm K}_{\boldsymbol{k},0}. \tag{13}$$

The ensemble average of $\delta^f(\boldsymbol{k}_1)\delta^{f*}(\boldsymbol{k}_2)$ then reads

$$\begin{aligned}
\left\langle \delta^f(\boldsymbol{k}_1)\delta^{f*}(\boldsymbol{k}_2) \right\rangle &= \frac{1}{N^2} \int_{V_\mu} \Pi\left(\frac{\boldsymbol{r}_1}{H}\right) \Pi\left(\frac{\boldsymbol{r}_2}{H}\right) \\
&\quad \times \left[\sum_{i\neq j} \langle n_i n_j \rangle W(\boldsymbol{r}_{i1}) W(\boldsymbol{r}_{j2}) \right. \\
&\quad \left. + \sum_i \langle n_i \rangle W(\boldsymbol{r}_{i1}) W(\boldsymbol{r}_{i2}) \right] \\
&\quad \times e^{i\boldsymbol{r}_1\cdot\boldsymbol{k}_1 - i\boldsymbol{r}_2\cdot\boldsymbol{k}_2}\, d\boldsymbol{r}_1\, d\boldsymbol{r}_2 \\
&\quad - \frac{1}{N} \int_{V_\mu} \Pi\left(\frac{\boldsymbol{r}_1}{H}\right) \sum_i \langle n_i \rangle W(\boldsymbol{r}_{i1}) e^{i\boldsymbol{r}_1\cdot\boldsymbol{k}_1} \delta^{\rm K}_{\boldsymbol{k}_2,0}\, d\boldsymbol{r}_1 \\
&\quad - \frac{1}{N} \int_{V_\mu} \Pi\left(\frac{\boldsymbol{r}_2}{H}\right) \sum_i \langle n_i \rangle W(\boldsymbol{r}_{i2}) e^{-i\boldsymbol{r}_2\cdot\boldsymbol{k}_2} \delta^{\rm K}_{\boldsymbol{k}_1,0}\, d\boldsymbol{r}_2 \\
&\quad + \delta^{\rm K}_{\boldsymbol{k}_1,0}\delta^{\rm K}_{\boldsymbol{k}_2,0},
\end{aligned} \tag{14}$$

where $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$. Using

$$\Pi(\boldsymbol{k}) = \frac{1}{V_\mu} \int_{V_\mu} \Pi\left(\frac{\boldsymbol{r}}{H}\right) e^{i\boldsymbol{r}\cdot\boldsymbol{k}}\, d\boldsymbol{r} = \sum_{\boldsymbol{n}} \delta_{\boldsymbol{k},2k_{\rm N}\boldsymbol{n}}, \tag{15}$$

where $k_{\rm N} = \pi/H$ is the Nyquist wavenumber, one can find

$$\begin{aligned}
\left\langle \delta^f(\boldsymbol{k}_1)\delta^{f*}(\boldsymbol{k}_2) \right\rangle = \sum_{\boldsymbol{n}_1,\boldsymbol{n}_2} \Big[\left|W(\boldsymbol{k}_1')\right|^2 P(\boldsymbol{k}_1') \delta^{\rm K}_{\boldsymbol{k}_1',\boldsymbol{k}_2'} \\
+ \frac{1}{N} \left|W(\boldsymbol{k}_1')\right|^2 \delta^{\rm K}_{\boldsymbol{k}_1',\boldsymbol{k}_2'} \Big],
\end{aligned} \tag{16}$$

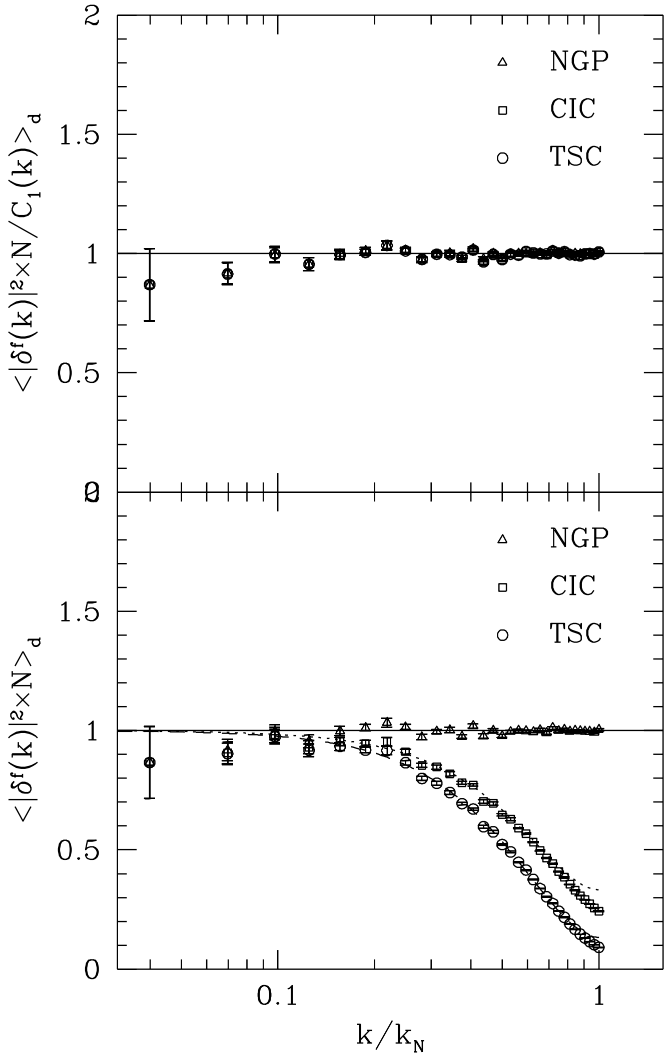

FIG. 1.—Shot noise $D^2(\boldsymbol{k})$ estimated from 10 samples of Poisson-distributed random points. Each symbol represents the result for each mass assignment, as indicated in the figure. *Top*: Result with the function $\langle D^2(\boldsymbol{k})N/C_3(\boldsymbol{k})\rangle_d$, i.e., the $D^2(\boldsymbol{k})$ scaled to $1/NC_1(\boldsymbol{k})$. The estimated result agrees quite well with the analytical prediction $\langle D^2(\boldsymbol{k})N/C_3(\boldsymbol{k})\rangle_d = 1$. *Bottom*: Comparison of the estimated $\langle D^2(\boldsymbol{k})N\rangle_d$ with our analytical predictions for the NGP (*solid line*), CIC (*dotted line*), and TSC (*dashed line*) assignment functions. For CIC and TSC, we have used the approximate formulae of eq. (21).

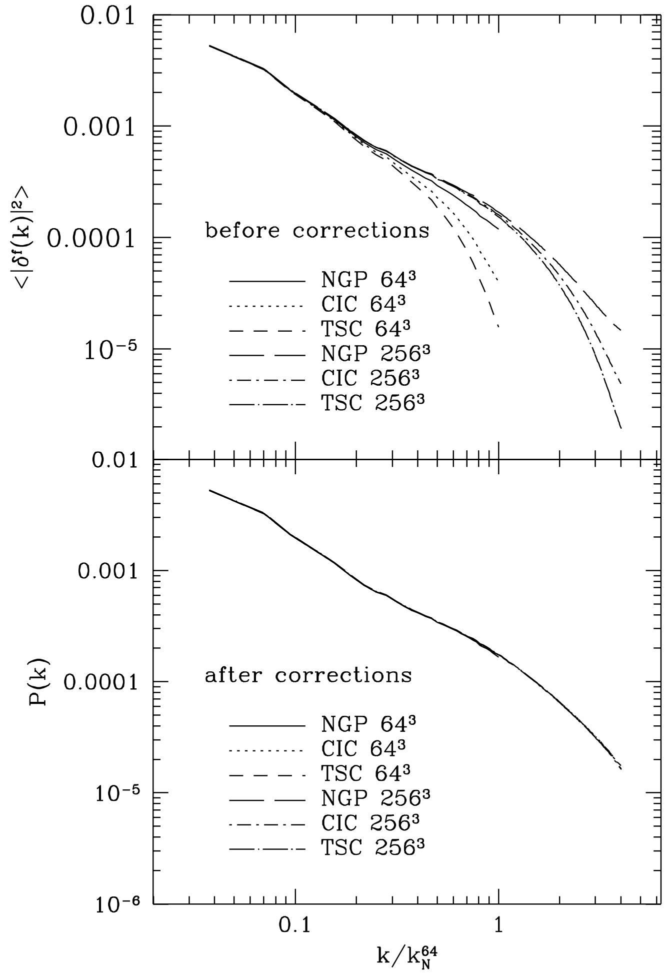

FIG. 2.—Six power spectra that we measured for an N-body simulation sample using three mass assignments (NGP, CIC, and TSC) and two grids (64^3 and 256^3 grid points), where k_N^{64} is the Nyquist wavenumber for 64^3 grid points. *Top*: Raw power spectra $\langle|\delta^f(k)|^2\rangle$ estimated directly from the FFT (eq. [8]). *Bottom*: True power spectra $P(k)$ reconstructed from the $\langle|\delta^f(k)|^2\rangle$ following the procedure described in the text. The six reconstructed $P(k)$ agree so well that their curves overlay each other.

where $\boldsymbol{k}_i' = \boldsymbol{k}_i + 2k_N\boldsymbol{n}_i$ and $W(\boldsymbol{k})$ is the FT of $W(\boldsymbol{r})$. For $\boldsymbol{k}_1 = \boldsymbol{k}_2 = \boldsymbol{k}$ we obtain our desired result:

$$\left\langle|\delta^f(\boldsymbol{k})|^2\right\rangle = \sum_{\boldsymbol{n}} |W(\boldsymbol{k}+2k_N\boldsymbol{n})|^2 P(\boldsymbol{k}+2k_N\boldsymbol{n}) + \frac{1}{N}\sum_{\boldsymbol{n}} |W(\boldsymbol{k}+2k_N\boldsymbol{n})|^2, \tag{17}$$

where the summation is over all three-dimensional integer vectors $\boldsymbol{n}$. The meaning of the above equation is very clear. The density convolution introduces the factor $W^2(\boldsymbol{k})$ both to the power spectrum and to the shot noise ($1/N$). The finite sampling of the convolved density field results in the "alias" sums (i.e., the sums over $\boldsymbol{n}$). The alias effect is well known in Fourier theory but has not been taken seriously in the power spectrum analysis of large-scale clustering in observational cosmology. The effects of both the convolution and the alias are significant near the Nyquist wavenumber k_N (see Figs. 1 and 2 in § 3).

3. A PROCEDURE TO RECOVER $P(k)$

In the practical measurement of $P(k)$ using an FFT, one should first choose a mass assignment function. The nearest grid point (NGP; $p = 1$), clouds-in-cell (CIC; $p = 2$), and triangular-shaped cloud (TSC; $p = 3$) assignment functions are the most popular functions for this purpose. For these schemes, we have (Hockney & Eastwood 1981)

$$W(\boldsymbol{k}) = \left[\frac{\sin(\pi k_1/2k_N)\sin(\pi k_2/2k_N)\sin(\pi k_3/2k_N)}{(\pi k_1/2k_N)(\pi k_2/2k_N)(\pi k_3/2k_N)}\right]^p, \tag{18}$$

where k_i is the ith component of $\boldsymbol{k}$.

Once one has selected the assignment function, the shot-noise effect (the second term on the right-hand side of eq. [17]) is easy to correct. For the NGP, CIC, and TSC assignments, the shot-noise term can be expressed by

$$D^2(\boldsymbol{k}) \equiv \frac{1}{N}\sum_{\boldsymbol{n}} W^2(\boldsymbol{k}+2k_N\boldsymbol{n}) = \frac{1}{N}C_1(\boldsymbol{k}), \tag{19}$$

where $C_1(\boldsymbol{k})$ are simple analytical functions:

$$C_1(\boldsymbol{k}) = \begin{cases} 1, & \text{NGP}, \\ \Pi_i \left[1 - \frac{2}{3}\sin^2\left(\frac{\pi \boldsymbol{k}_i}{2k_{\rm N}}\right)\right], & \text{CIC}, \\ \Pi_i \left[1 - \sin^2\left(\frac{\pi \boldsymbol{k}_i}{2k_{\rm N}}\right) + \frac{2}{15}\sin^4\left(\frac{\pi \boldsymbol{k}_i}{2k_{\rm N}}\right)\right], & \text{TSC}. \end{cases} \quad (20)$$

Furthermore, one can easily show that the $C_1(\boldsymbol{k})$ of the CIC and TSC schemes are approximately isotropic for $k \le k_{\rm N}$, i.e.,

$$C_1(\boldsymbol{k}) \approx \begin{cases} 1 - \frac{2}{3}\sin^2\left(\frac{\pi k}{2k_{\rm N}}\right), & \text{CIC}, \\ 1 - \sin^2\left(\frac{\pi k}{2k_{\rm N}}\right) + \frac{2}{15}\sin^4\left(\frac{\pi k}{2k_{\rm N}}\right), & \text{TSC}. \end{cases} \quad (21)$$

We have tested equations (19)–(21) by calculating $\langle|\delta^f(\boldsymbol{k})|^2\rangle$ for 10 random simulation samples, each of which consists of $N = 10^5$ points randomly distributed in a unit cube. In this case, $D^2(\boldsymbol{k}) = \langle|\,\delta^f(\boldsymbol{k})\,|^2\rangle$. In Figure 1 (*top*) we show the average values and 1 σ errors of $\langle D^2(\boldsymbol{k})N/C_1(\boldsymbol{k})\rangle_d$ estimated from the 10 samples, where (there and below) $\langle\ldots\rangle_d$ means an average over all directions of $\boldsymbol{k}$. In this calculation we have used the NGP, CIC, and TSC assignment functions and have used equation (20) for $C_1(\boldsymbol{k})$. The result for each assignment is shown by one symbol in the figure. From equation (19), one expects $\langle D^2(\boldsymbol{k})N/C_1(\boldsymbol{k})\rangle_d = 1$ for all three assignment functions. Clearly, the simulation results agree very well with equation (20). In Figure 1 (*bottom*), we show the results in a slightly different way, i.e., the averages and 1 σ errors of $\langle D^2(\boldsymbol{k})N\rangle_d$. The solid line shows $\langle D^2(\boldsymbol{k})N\rangle_d = 1$ for the NGP assignment. The dotted and dashed lines are the approximate expressions of equation (21) for the CIC and TSC assignments. Again we find a very good agreement between equation (21) and the results from the random sample. This means that in most applications, one can use equation (21) to correct the shot noise in the FFT measurement of the power spectrum.

After correcting the shot noise, our central problem becomes how to extract $P(k)$ from the first term of equation (17). Let us consider the correction factor $C_2(k)$, which is defined as

$$C_2(k) = \frac{\left\langle \sum_{\boldsymbol{n}} W^2(\boldsymbol{k} + 2k_{\rm N}\boldsymbol{n})P(\boldsymbol{k} + 2k_{\rm N}\boldsymbol{n})\right\rangle_d}{P(k)}. \quad (22)$$

Since $W(\boldsymbol{k})$ is a decreasing function and $P(k)$ is also a decreasing function on the scales $k \gtrsim k_{\rm N}$, we expect the alias contribution of large $|\boldsymbol{n}|$ to $C_2(k)$ to be small. In particular, for $k \ll k_{\rm N}$, any alias contribution is small, and we have $C_2(k) \approx W^2(k) \approx 1$ independent of $P(k)$. For $k \sim k_{\rm N}$, the alias contribution to $C_2(k)$ becomes important, most of which is owing to the alias of $|\boldsymbol{k} + 2k_{\rm N}\boldsymbol{n}| \sim \boldsymbol{k}$. Therefore, the dependence of $C_2(k)$ on $P(k)$ is only on the shape of $P(k)$ at $k \sim k_{\rm N}$, i.e., the local slope $\alpha_{\rm N}$ of $P(k)$ at $k \sim k_{\rm N}$, $\alpha_{\rm N} = [\ln P(k)/\ln k]_{k \sim k_{\rm N}}$.

Since the local slope $\alpha_{\rm N}$ is unknown a priori in practical measurement of $P(k)$, we propose an iterative method to get the correction factor $C_2(k)$. Suppose that we have measured the power spectrum $P_r(k)$:

$$\begin{aligned} P_r(k) &= \left\langle \left\langle |\delta^f(\boldsymbol{k})|^2 \right\rangle - D^2(\boldsymbol{k}) \right\rangle_d \\ &= \left\langle \sum_{\boldsymbol{n}} W^2(\boldsymbol{k} + 2k_{\rm N}\boldsymbol{n})P(\boldsymbol{k} + 2k_{\rm N}\boldsymbol{n}) \right\rangle_d \end{aligned} \quad (23)$$

for $k \le k_{\rm N}$. The local slope α_0 of $P_r(k)$ at $k \sim k_{\rm N}$ is calculated by a power-law fitting to $P_r(k)$ at $0.5k_{\rm N} \le k \le k_{\rm N}$. Assuming the power-law form k^{α_0} for $P(k)$ of equation (22), we calculate $C_2(k,\ \alpha_0)$ and get $P_0(k) = P_r(k)/C_2(k,\ \alpha_0)$. Using the local slope α_0 at $0.5k_{\rm N} \le k \le k_{\rm N}$ of the $P_0(k)$ just obtained, we calculate $C_2(k,\ \alpha_0)$ and $P_0(k) = P_r(k)/C_2(k,\ \alpha_0)$ again. This calculation is repeated until $P_0(k)$ (or α_0) converges to some defined accuracy. The converged $P_0(k)$ is our desired $P(k)$.

Success of the above iteration procedure is shown by Figure 2, in which we present our measurement of $P(k)$ for a simulation sample of Jing et al. (1995). The sample consists of five realizations of a particle-particle-particle-mesh simulation of a low-density flat universe with $\Omega_0 = 0.2$, $\Lambda_0 = 0.8$, and $h = 1$. The simulation box size is 128 h^{-1} Mpc, and the number of simulation particles is $N = 64^3$. Since the details of the simulation are unimportant for our result here, we do not discuss them anymore. Figure 2 (*top*) shows the raw power spectra $\langle|\delta^f(\boldsymbol{k})|^2\rangle$, which are calculated with the NGP, CIC, or TSC mass assignment and with 256^3 or 64^3 grid points. As expected, the raw power spectrum depends on the scheme of the mass assignment: a higher order mass assignment gives a smaller $\langle|\delta^f(\boldsymbol{k})|^2\rangle$ near the Nyquist frequency. At $k \sim k_{\rm N}^{64}$, the Nyquist wavenumber of 64^3 grid points, the raw power spectra calculated with 256^3 grid points are expected be very close to the true power spectrum $P(k)$ because the effect of neither the shot noise $[D^2(k) \ll P(k)]$, the convolution $[W(k) \approx 1]$, nor the alias $[C_2(k) \approx 1]$ is important. The differences at $k \sim k_{\rm N}^{64}$ between the power spectra calculated with 64^3 grid points and those with 256^3 grid points show the importance of the effects discussed in this paper. Figure 2 (*bottom*) plots the power spectra $P(k)$, which are corrected following the procedure prescribed previously in this section. In this example we use equation (21) to correct for the shot noise, and we require that the slope α_0 converge to the accuracy of $|\Delta\alpha_0| \le 0.02$ in the iterative method of deconvolving the alias summation. For this accuracy, only fewer than five iterations are needed in each calculation. The six power spectra $P(k)$, measured with different mass assignments and with different numbers of grid points, agree so well that their curves overlay each other in the figure. The biggest difference between the 64^3 and 256^3 $P(k)$ is at $k = k_{\rm N}^{64}$, which is less than 4%. The result is very encouraging, and it tells us that using the correction procedure prescribed in this paper, one can obtain the true $P(k)$ for $k \le k_{\rm N}$ in the FFT measurement, independent of which mass assignment is used.

In the above derivations and discussions, we have assumed for simplicity that the sample is uniformly (within Poisson fluctuation) constructed in a cubic volume. However, the above method is also valid for a sample with selection effects (e.g., an extragalactic catalog). To show this, let us introduce a selection function $S(\boldsymbol{r})$ defined as the observable rate of the sample at position $\boldsymbol{r}$. If the underlying density distribution is $n(\boldsymbol{r})$, the density distribution $n_s(\boldsymbol{r})$ of the sample is then

$$n_s(\boldsymbol{r}) = S(\boldsymbol{r})n(\boldsymbol{r}). \quad (24)$$

In this case, we define the following transformation for the FFT:

$$\delta_s^f(\boldsymbol{k}) = \frac{1}{N}\sum_g \left[n_s^f(\boldsymbol{r}_g) - \bar{n}S(\boldsymbol{r}_g)\right]e^{i\boldsymbol{r}_g \cdot \boldsymbol{k}}, \quad (25)$$

where $n_s^f(\boldsymbol{r}_g)$ is the convolved density of $n_s(\boldsymbol{r})$ at $\boldsymbol{r}_g$ (see eq. [9]) and $\bar{n}$ is the mean underlying number density. Following the derivation of equation (16), we can easily find

$$\left\langle \delta_s^f(\boldsymbol{k}_1)\delta_s^{f*}(\boldsymbol{k}_2)\right\rangle = \sum_{\boldsymbol{n}_1,\boldsymbol{n}_2}\Big[W(\boldsymbol{k}_1')W(-\boldsymbol{k}_2') \times \sum_{\boldsymbol{k}_3'} S(\boldsymbol{k}_1'+\boldsymbol{k}_3')P(\boldsymbol{k}_3')S(-\boldsymbol{k}_2'-\boldsymbol{k}_3') + \frac{1}{N}S(\boldsymbol{k}_1'-\boldsymbol{k}_2')W(\boldsymbol{k}_1')W(-\boldsymbol{k}_2')\Big], \quad (26)$$

where $S(\boldsymbol{k})$ is defined as

$$S(\boldsymbol{k}) = \frac{1}{\int S(\boldsymbol{r})\,d\boldsymbol{r}}\int S(\boldsymbol{r})e^{i\boldsymbol{r}\cdot\boldsymbol{k}}\,d\boldsymbol{r}. \quad (27)$$

Because $S(\boldsymbol{k})$ peaks at $\boldsymbol{k} \approx 0$, the coupling of the selection and the power spectrum (the first term of eq. [26]) is important only at small k. How to treat this coupling is a nontrivial task and beyond the scope of this paper (but see, e.g., Peacock & Nicholson 1991; Jing & Valdarnini 1993). However, for larger k for which the effects discussed here become significant, we have

$$\left\langle |\delta^f(\boldsymbol{k})|^2\right\rangle \approx \sum_{\boldsymbol{k}_3} S^2(\boldsymbol{k}_3)\sum_{\boldsymbol{n}} W^2(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n})P(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n}) + \frac{1}{N}\sum_{\boldsymbol{n}} W^2(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n}). \quad (28)$$

The difference between equations (28) and (17) is only the factor $\sum_{\boldsymbol{k}_3} S^2(\boldsymbol{k}_3)$. Therefore, our procedure for correcting the effects of the mass assignment is still valid for observational samples with selection effects. Equation (28) is derived on the assumption that $S(\boldsymbol{k})$ is compact in $\boldsymbol{k}$-space. This assumption is valid for many redshift surveys, e.g., the CfA and *IRAS* redshift surveys, but it is not valid for surveys of irregular boundaries, e.g., pencil-beam surveys. In addition, current redshift surveys of galaxies already cover a volume of $(1000\ h^{-1}\ {\rm Mpc})^3$. Because one is also interested in clustering information down to scales of 0.1 h^{-1} Mpc, a $(10{,}000)^3$ grid point FFT is required to explore the clustering on all scales, which is still difficult to realize on modern supercomputers. Other estimators, e.g., the FT of the two-point correlation function (Jing & Börner 2004), should be used to determine $P(k)$ on small scales.

4. SUMMARY

In this paper we have derived for the first time the formula (eq. [17]) that relates the raw power spectrum $\langle|\delta^f(k)|^2\rangle$ estimated with an FFT to the true power spectrum $P(k)$. The formula shows clearly how the mass assignment modifies the power spectrum. The convolution of the density field with an assignment function $W(\boldsymbol{r})$ reshapes the power spectrum (including the shot-noise spectrum) by multiplying by the factor $|W(\boldsymbol{k})|^2$; the finite sampling of the convolved density field leads to the alias sum. We have described how to reconstruct $P(k)$ from $\langle|\delta^f(k)|^2\rangle$. For the NGP, CIC, and TSC assignment functions, the shot noise $D^2(\boldsymbol{k})$ can be expressed by simple analytical functions; therefore, the shot noise can be easily corrected. To extract $P(k)$ from the alias sum $\sum_{\boldsymbol{n}} W^2(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n})P(\boldsymbol{k}+2k_{\rm N}\boldsymbol{n})$, we propose an iterative method. The method has been tested by applying it to an N-body simulation sample. Using different numbers of grid points, we have shown that the method can very successfully recover $P(k)$ for all the NGP, CIC, and TSC assignment functions.

The author thanks Bhuvnesh Jain for stimulating discussion on an earlier version of the draft. This work is supported in part by the NKBRSF (G19990754) and by the Natural Science Foundation of China (10125314 and 10373012).

REFERENCES

Baumgart, D. J., & Fry, J. N. 1991, ApJ, 375, 25
da Costa, L. N., Vogeley, M. S., Geller, M. J., Huchra, J. P., & Park, C. 1994, ApJ, 437, L1
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
Fisher, K., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. P. 1993, ApJ, 402, 42
Gramann, M., & Einasto, J. 1992, MNRAS, 254, 453
Hockney, R. W., & Eastwood, J. W. 1981, Computer Simulations Using Particles (New York: McGraw-Hill)
Jing, Y. P., & Börner, G. 2004, ApJ, 617, 782
Jing, Y. P., Mo, H. J., Börner, G., & Fang, L. Z. 1995, MNRAS, 276, 417
Jing, Y. P., & Valdarnini, R. 1993, ApJ, 406, 6
Kaiser, N. 1991, in Texas/ESO-CERN Symp. Relativistic Astrophysics, ed. J. Barrow et al. (New York: New York Acad. Sci.), 295
Park, C., Gott, J. R., III, & da Costa, L. N. 1992, ApJ, 392, L51
Peacock, J. A., & Nicholson, D. 1991, MNRAS, 253, 307
Peacock, J. A., & West, M. 1992, MNRAS, 259, 494
Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Scoccimarro, R., Colombi, S., Fry, J. N., Frieman, J. A., Hivon, E., & Melott, A. 1998, ApJ, 496, 586
Vogeley, M. S., Park, C., Geller, M. J., & Huchra, J. P. 1992, ApJ, 391, L5