Mon. Not. R. Astron. Soc. **357,** 1387–1398 (2005) doi:10.1111/j.1365-2966.2005.08776.x

Covariance of the one-dimensional mass power spectrum

Hu Zhan[1★]† and Daniel Eisenstein[2]†
[1]*Department of Physics, University of Arizona, Tucson, AZ 85721, USA*
[2]*Steward Observatory, University of Arizona, Tucson, AZ 85721, USA*

Accepted 2004 December 20. Received 2004 December 20; in original form 2004 July 19

ABSTRACT

We analyse the covariance of the one-dimensional mass power spectrum along lines of sight. The covariance reveals the correlation between different modes of fluctuations in the cosmic density field and gives the sample variance error for measurements of the mass power spectrum. For Gaussian random fields, the covariance matrix is diagonal. As expected, the variance of the measured one-dimensional mass power spectrum is inversely proportional to the number of lines of sight that are sampled from each random field. The correlation between lines of sight in a single field may alter the covariance. However, lines of sight that are sampled far apart are only weakly correlated, so that they can be treated as independent samples. Using N-body simulations, we find that the covariance matrix of the one-dimensional mass power spectrum is not diagonal for the cosmic density field due to the non-Gaussianity and that the variance is much higher than that of Gaussian random fields. From the covariance, one will be able to determine the cosmic variance in the measured one-dimensional mass power spectrum as well as to estimate how many lines of sight are needed to achieve a target precision.

Key words: cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

The one-dimensional mass power spectrum (PS) and its relation to the three-dimensional mass PS has been frequently utilized to recover the linear mass PS from the Lyα forest (Croft et al. 1998, 1999, 2002; Gnedin & Hamilton 2002; Kim et al. 2004). This opens a great window for studying the large-scale structure of the Universe over a wide range of redshift. As one starts to attempt precision cosmology using the Lyα forest (Croft et al. 2002; Mandelbaum et al. 2003; Spergel et al. 2003), it becomes necessary to quantify systematic uncertainties of the PS analysis including the covariance.

For an ensemble of isotropic fields, the one-dimensional mass PS $P_{1\mathrm{D}}(k)$ is a simple integral of the three-dimensional mass PS $P_{3\mathrm{D}}(k)$ (Lumsden, Heavens & Peacock 1989)

$$P_{1\mathrm{D}}(k) = \frac{1}{2\pi}\int_k^\infty P_{3\mathrm{D}}(k')k'\,\mathrm{d}k', \tag{1}$$

where k is the line-of-sight (LOS) wavenumber. However, there is only one observable Universe. One has to replace the ensemble average with a spatial average. For instance, one may sample multiple LOS densities from the three-dimensional cosmic density field and use the average PS of the one-dimensional densities in place of the ensemble-average quantity.

The covariance of the spatial-average PS may differ from that of the ensemble-average PS for at least two reasons. First, LOSs are no longer independent of each other. Correlations between close LOSs could increase the covariance of the one-dimensional mass PS, although they may also be used to extract cosmological parameters from the Lyα forest (Hui, Stebbins & Burles 1999; McDonald & Miralda-Escudé 1999; Viel et al. 2002; McDonald 2003; Rollinde et al. 2003). Secondly, the length of each LOS is always much less than the size of the Universe, so that false correlations between different modes are introduced in the covariance (for the three-dimensional case, see Feldman, Kaiser & Peacock 1994). Because the cosmic density field is highly non-Gaussian on scales of interest, N-body simulations are helpful for quantifying the covariance.

The rest of the paper is organized as follows. In Section 2 we briefly describe the notation and the convention of Fourier transforms, the three-dimensional mass PS, and its covariance. The spatial-average one-dimensional mass PS and its covariance are derived in Sections 3 and 4, respectively. The effects of the LOS length are discussed in Section 5. In Section 6 we present numerical results of the covariance from

★Current address: Department of Physics, University of California, One Shields Ave, Davis, CA 95616, USA.
†E-mail: zhan@physics.ucdavis.edu (HZ); deisenstein@as.arizona.edu (DE)

N-body simulations with different box sizes. The conclusions and discussions are given in Section 7. Unless specified otherwise, the PS refers to the mass PS.

2 THREE-DIMENSIONAL POWER SPECTRUM

We assume the following convention of Fourier transforms for a cubic density field of volume $V = B^3$:

$$\hat{\delta}(\boldsymbol{n}) = \int_V \delta(\boldsymbol{x})\, \mathrm{e}^{-2\pi \mathrm{i}\boldsymbol{n}\cdot\boldsymbol{x}/B}\, \mathrm{d}\boldsymbol{x}, \qquad \delta(\boldsymbol{x}) = \frac{1}{V}\sum_{\boldsymbol{n}=-\infty}^{\infty} \hat{\delta}(\boldsymbol{n})\, \mathrm{e}^{2\pi \mathrm{i}\boldsymbol{n}\cdot\boldsymbol{x}/B}. \tag{2}$$

Here, $\delta(\boldsymbol{x})$ and $\hat{\delta}(\boldsymbol{n})$ are the overdensity and its Fourier counterpart, respectively, the summation $\sum_{\boldsymbol{n}=-\infty}^{\infty}$ is an abbreviation for $\sum_{n_1,n_2,n_3=-\infty}^{\infty}$, and the wave vector $\boldsymbol{k} = 2\pi\boldsymbol{n}/B$. With the understanding that Fourier modes exist only at discrete wavenumbers (i.e. n_1, n_2 and n_3 are integers), one may use $\boldsymbol{k}$ and $\boldsymbol{n}$ interchangeably for convenience. To be complete, the orthonormality relations are

$$\frac{1}{V}\int_V \mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{n}-\boldsymbol{n}')\cdot\boldsymbol{x}/B}\mathrm{d}\boldsymbol{x} = \delta^{\mathrm{K}}_{\boldsymbol{n},\boldsymbol{n}'}, \qquad \frac{1}{V}\sum_{\boldsymbol{n}=-\infty}^{\infty} \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{n}\cdot(\boldsymbol{x}-\boldsymbol{x}')/B} = \delta^{\mathrm{D}}(\boldsymbol{x}-\boldsymbol{x}'), \tag{3}$$

where $\delta^{\mathrm{K}}_{\boldsymbol{n},\boldsymbol{n}'}$ is the three-dimensional Kronecker delta function, and $\delta^{\mathrm{D}}(\boldsymbol{x}-\boldsymbol{x}')$ the Dirac delta function.

The three-dimensional PS of the Universe is defined through

$$\langle \hat{\delta}(\boldsymbol{k})\hat{\delta}^*(\boldsymbol{k}')\rangle = P_{3\mathrm{D}}(\boldsymbol{k}) V \delta^{\mathrm{K}}_{\boldsymbol{n},\boldsymbol{n}'}, \tag{4}$$

where $\langle \ldots \rangle$ denotes an ensemble average. One may define an observed PS, $\mathcal{P}_{3\mathrm{D}}(\boldsymbol{k}) = |\hat{\delta}(\boldsymbol{k})|^2/V$, so that $P_{3\mathrm{D}}(\boldsymbol{k}) = \langle \mathcal{P}_{3\mathrm{D}}(\boldsymbol{k})\rangle$. Note that a shot-noise term should be included if the PS is measured from discrete objects, and it is inversely proportional to the mean number density of the objects. We neglect the shot noise in this paper. For an isotropic universe, $P_{3\mathrm{D}}(\boldsymbol{k})$ is a function of the length of $\boldsymbol{k}$ only, i.e. $P_{3\mathrm{D}}(\boldsymbol{k}) \equiv P_{3\mathrm{D}}(k)$. The four-point function of $\hat{\delta}(\boldsymbol{k})$ is

$$\langle \hat{\delta}(\boldsymbol{k})\hat{\delta}^*(\boldsymbol{k}')\hat{\delta}(\boldsymbol{k}'')\hat{\delta}^*(\boldsymbol{k}''')\rangle = V^2\left[P_{3\mathrm{D}}(k)P_{3\mathrm{D}}(k'')\delta^{\mathrm{K}}_{\boldsymbol{n},\boldsymbol{n}'}\delta^{\mathrm{K}}_{\boldsymbol{n}'',\boldsymbol{n}'''} + P_{3\mathrm{D}}(k)P_{3\mathrm{D}}(k')\delta^{\mathrm{K}}_{\boldsymbol{n},\boldsymbol{n}'''}\delta^{\mathrm{K}}_{\boldsymbol{n}',\boldsymbol{n}''}\right] + VT(\boldsymbol{k},-\boldsymbol{k}',\boldsymbol{k}'',-\boldsymbol{k}'''), \tag{5}$$

where T is the trispectrum, and we have restricted the wave vectors to be in the same hemisphere so that the term $P_{3\mathrm{D}}(k)P_{3\mathrm{D}}(k')\delta^{\mathrm{K}}_{\boldsymbol{n},-\boldsymbol{n}''}\delta^{\mathrm{K}}_{\boldsymbol{n}',-\boldsymbol{n}'''}$ does not appear. There is a redundancy in the variables of the trispectrum, which has only six degrees of freedom arising from relative coordinates of the four points under the constraint of homogeneity (Peebles 1980). It is evident from the four-point function that the covariance of the three-dimensional PS is (see also Meiksin & White 1999; Cooray & Hu 2001)

$$\sigma^2(\boldsymbol{k},\boldsymbol{k}') = \langle[\mathcal{P}_{3\mathrm{D}}(\boldsymbol{k}) - P_{3\mathrm{D}}(\boldsymbol{k})][\mathcal{P}_{3\mathrm{D}}(\boldsymbol{k}') - P_{3\mathrm{D}}(\boldsymbol{k}')]\rangle = P^2_{3\mathrm{D}}(\boldsymbol{k})\delta^{\mathrm{K}}_{\boldsymbol{n},\boldsymbol{n}'} + V^{-1}T(\boldsymbol{k},-\boldsymbol{k},\boldsymbol{k}',-\boldsymbol{k}'). \tag{6}$$

For Gaussian random fields (GRFs), the trispectrum vanishes, and $\sigma^2(\boldsymbol{k},\boldsymbol{k}') = P^2_{3\mathrm{D}}(\boldsymbol{k})\ \delta^{\mathrm{K}}_{\boldsymbol{n},\boldsymbol{n}'}$. In reality, the survey volume is always smaller than the observable Universe, and the survey geometry is more complex than the simple case we have assumed. These lead to a modification of the covariance for GRFs (Feldman et al. 1994)

$$\sigma^2(\boldsymbol{k},\boldsymbol{k}') \simeq P^2_{3\mathrm{D}}(\boldsymbol{k})\left|\frac{\int u^2(\boldsymbol{x})\,\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}}{\int u^2(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}}\right|^2, \tag{7}$$

where $u(\boldsymbol{x})$ is a weight function that depends on the survey volume and geometry, and the shot noise is neglected. For the low-redshift cosmic density field, the non-Gaussianity can dramatically boost the elements of the covariance matrix through the trispectrum.

3 ONE-DIMENSIONAL POWER SPECTRUM

For a LOS density that is along the x_3-axis and sampled at (x_1, x_2), the one-dimensional Fourier transform gives

$$\hat{\delta}(\boldsymbol{x}_\perp, n_3) = \int_0^B \delta(\boldsymbol{x})\, \mathrm{e}^{-2\pi\mathrm{i}n_3x_3/B}\,\mathrm{d}x_3 = \frac{1}{B^2}\sum_{\boldsymbol{n}_\perp=-\infty}^{\infty}\hat{\delta}(\boldsymbol{n}_\perp, n_3)\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{n}_\perp\cdot\boldsymbol{x}_\perp/B}, \tag{8}$$

where the subscript $\perp$ signifies the first two components of a vector (i.e. $\boldsymbol{x}_\perp = (x_1, x_2)$). Similar to the three-dimensional PS, the one-dimensional PS is expected to follow

$$\langle \hat{\delta}(\boldsymbol{x}_\perp, n_3)\hat{\delta}^*(\boldsymbol{x}_\perp, n_3')\rangle = P_{1\mathrm{D}}(n_3) B \delta^{\mathrm{K}}_{n_3,n_3'}. \tag{9}$$

Substituting equation (8) in equation (9) and making use of equation (4), one finds the relation between the one-dimensional PS and the three-dimensional PS

$$P_{1\mathrm{D}}(n_3) = \frac{1}{B^2}\sum_{\boldsymbol{n}_\perp=-\infty}^{\infty} P_{3\mathrm{D}}(\boldsymbol{n}_\perp, n_3), \tag{10}$$

which is a discrete analogue of equation (1).

Practically, one measures PSs of LOS densities sampled at some locations, e.g. $\mathcal{P}_{1\mathrm{D}}(\boldsymbol{x}_\perp, k_3) = |\hat{\delta}(\boldsymbol{x}_\perp, k_3)|^2/B$. A simple estimator of the one-dimensional PS may be constructed by a spatial average over many LOSs, i.e. $\mathcal{P}_{1\mathrm{D}}(k_3) = \langle \mathcal{P}_{1\mathrm{D}}(\boldsymbol{x}_\perp, k_3)\rangle_{\boldsymbol{x}_\perp}$, where $\langle \ldots \rangle_{\boldsymbol{x}_\perp}$ runs over all LOSs sampled in a single universe. To assess the performance of the estimator, two questions need to be addressed. (i) How does $\mathcal{P}_{1\mathrm{D}}(k_3)$ relate

itself to $\mathcal{P}_{3D}(\boldsymbol{k})$? (ii) What is the covariance of $\mathcal{P}_{1D}(k_3)$ with respect to $P_{1D}(k_3)$? The rest of this section answers the former, and Section 4 the latter.

For simplicity, we assume that LOSs are sampled regularly in transverse directions at an interval of $b = B/m$, where m is an integer. Each LOS has a length of B, and $\boldsymbol{x}_\perp = \boldsymbol{l}_\perp\, b$ with $l_1, l_2 = 0, \ldots, m-1$. The estimated one-dimensional PS is then

$$\mathcal{P}_{1D}(n_3) = \frac{1}{m^2 B^5} \sum_{\boldsymbol{l}_\perp=0}^{m-1} \left| \sum_{\boldsymbol{n}_\perp=-\infty}^{\infty} \hat{\delta}(\boldsymbol{n}_\perp, n_3)\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{n}_\perp \cdot \boldsymbol{l}_\perp / m} \right|^2 = \frac{1}{B^5} \sum_{\boldsymbol{n}_\perp,\, \boldsymbol{j}_\perp=-\infty}^{\infty} \hat{\delta}(\boldsymbol{n}_\perp, n_3)\, \hat{\delta}^*(\boldsymbol{n}_\perp + m\boldsymbol{j}_\perp, n_3), \tag{11}$$

where the equality

$$\frac{1}{m} \sum_{l=0}^{m-1} \mathrm{e}^{2\pi \mathrm{i}(n-n')l/m} = \delta^{\mathrm{K}}_{n,n'+mj} \quad j = 0, \pm 1, \ldots, \pm\infty \tag{12}$$

has been used to obtain the last line in equation (11). It is easy to show with equation (11) that

$$P_{1D}(k_3) = \langle \mathcal{P}_{1D}(k_3) \rangle = \langle \mathcal{P}_{1D}(\boldsymbol{x}_\perp, k_3) \rangle. \tag{13}$$

Note that there is no spatial average on the far right side of equation (13). It seems that only in the limit $m \to \infty$ do $\mathcal{P}_{1D}(k_3)$ and $\mathcal{P}_{3D}(\boldsymbol{k})$ follow the same relation as equations (1) and (10), i.e. $\mathcal{P}_{1D}(k_3) = \sum_{\boldsymbol{n}_\perp=-\infty}^{\infty} \mathcal{P}_{3D}(\boldsymbol{k}_\perp, k_3)/B^2$, where we have assumed $\hat{\delta}(\infty) = 0$ so that only $\boldsymbol{j}_\perp = (0, 0)$ terms contribute. In other words, to reduce the uncertainties of the recovered $\mathcal{P}_{3D}(\boldsymbol{k})$ and $P_{3D}(\boldsymbol{k})$, one has to increase the sampling rate in the transverse direction.

Without losing generality, one may choose m to be even, so that equation (11) can be rearranged into

$$\mathcal{P}_{1D}(n_3) = \frac{1}{B^2} \sum_{\boldsymbol{n}_\perp=-m/2}^{m/2} \mathcal{P}^{\mathrm{a}}_{3D}(\boldsymbol{n}_\perp, n_3) = \frac{1}{B^2} \sum_{\boldsymbol{n}_\perp=-\infty}^{\infty} \mathcal{P}_{3D}(\boldsymbol{n}_\perp, n_3) + A(n_3), \tag{14}$$

where

$$\mathcal{P}^{\mathrm{a}}_{3D}(\boldsymbol{n}_\perp, n_3) = \frac{1}{V} \left| \sum_{\boldsymbol{j}_\perp=-\infty}^{\infty} \hat{\delta}(\boldsymbol{n}_\perp + m\boldsymbol{j}_\perp, n_3) \right|^2, \tag{15}$$

and

$$A(n_3) = \frac{1}{B^5} \sum_{\substack{\boldsymbol{j}_\perp, \boldsymbol{j}'_\perp=-\infty \\ \boldsymbol{j}'_\perp \neq 0}}^{\infty} \sum_{\boldsymbol{n}_\perp=-m/2}^{m/2} \hat{\delta}(\boldsymbol{n}_\perp + m\boldsymbol{j}_\perp, n_3)\, \hat{\delta}^*[\boldsymbol{n}_\perp + m(\boldsymbol{j}_\perp + \boldsymbol{j}'_\perp), n_3]. \tag{16}$$

Equation (14) provides two equivalent views of the sampling effect. On one hand, the one-dimensional PS is a sum of the aliased three-dimensional PS, $\mathcal{P}^{\mathrm{a}}_{3D}(\boldsymbol{k})$, that is sampled by a grid of $m \times m$ LOSs. On the other hand, it is a complete sum of the underlying three-dimensional PS with an extra term $A(k_3)$ that is determined by the sampling rate and properties of the density field.

The discrete Fourier transform cannot distinguish a principal mode at $|k| \leqslant k_{\mathrm{Nyq}}$ from its aliases at $k \pm 2k_{\mathrm{Nyq}}$, $k \pm 4k_{\mathrm{Nyq}}$, ..., where k_{Nyq} is the sampling Nyquist wavenumber. For example, $k_{\mathrm{Nyq}} = \pi/b$ if one samples the field at an equal spacing of b. The alias modes will be added to the principal mode if they are not properly filtered out before sampling (for details, see Hockney & Eastwood 1981). Because the cosmic density field is not band-limited, aliasing can distort statistics of the field. The distortion on the three-dimensional PS cannot be quantified a priori, because it depends on relative phases between the principal mode and its aliases. The alias effect is less pronounced, if amplitudes of the alias modes are much smaller than that of the principal mode. Because the three-dimensional PS decreases toward small scales, a high sampling rate (or k_{Nyq}) can suppress aliasing for modes with $k \ll k_{\mathrm{Nyq}}$.

If the Fourier modes of the density field are uncorrelated, the term $A(k_3)$ may be neglected even for a finite number of LOSs and therefore validate equation (10). Strictly speaking, $A(k_3)$ vanishes only as an ensemble average over many GRFs, but because there are so many independent modes in a shell of radius around k – especially at large wavenumbers – the summation in $A(k_3)$ will tend to vanish even for a single GRF. The cosmic density field is more Gaussian at higher redshift, so equation (10) may be a good approximation then. At low redshift, however, the non-Gaussianity alters the statistics of one-dimensional fields (see, for example, Amendola 1994), such that one might only be able to recover a heavily aliased three-dimensional PS from a sparse sample of LOSs.

4 COVARIANCE OF THE ONE-DIMENSIONAL POWER SPECTRUM

The covariance of the PS is of interest because it tells us how likely an estimated PS is to represent the true PS and how much the modes on different scales are correlated. For the one-dimensional PS, the covariance is defined as

$$\sigma^2_{1D}(k_3, k'_3) = \langle [\mathcal{P}_{1D}(k_3) - P_{1D}(k_3)][\mathcal{P}_{1D}(k'_3) - P_{1D}(k'_3)] \rangle. \tag{17}$$

The covariance of the mean PS of N LOSs that are sampled in a single field can be expanded into a sum of pairwise covariances between two LOSs separated by $\boldsymbol{s}_\perp^{jl} = \boldsymbol{x}_\perp^j - \boldsymbol{x}_\perp^l$, e.g.

$$\sigma_{\rm 1D}^2(k_3, k_3') = \frac{1}{N^2}\sum_{j,l=1}^{N} \sigma_{\rm 1D}^2\left(k_3, k_3'; \boldsymbol{s}_\perp^{jl}\right), \tag{18}$$

where

$$\sigma_{\rm 1D}^2\left(k_3, k_3'; \boldsymbol{s}_\perp^{jl}\right) = \left\langle \left[\mathcal{P}_{\rm 1D}\left(\boldsymbol{x}_\perp^j, k_3\right) - P_{\rm 1D}(k_3)\right]\left[\mathcal{P}_{\rm 1D}\left(\boldsymbol{x}_\perp^l, k_3'\right) - P_{\rm 1D}(k_3')\right]\right\rangle,$$

and $\boldsymbol{x}_\perp^j$ is the location of the jth LOS in the x_1–x_2 plane.

For GRFs, the four-point function (equation 5) helps reduce the pairwise covariance to

$$\begin{aligned}\sigma_{\rm 1D}^2(n_3, n_3'; \boldsymbol{s}) &= \frac{1}{B^{10}}\sum_{\boldsymbol{n}_\perp, \boldsymbol{n}_\perp', \boldsymbol{n}_\perp'', \boldsymbol{n}_\perp'''=-\infty}^{\infty} \left\langle \hat{\delta}(\boldsymbol{n}_\perp, n_3)\hat{\delta}^*(\boldsymbol{n}_\perp', n_3)\hat{\delta}(\boldsymbol{n}_\perp'', n_3')\hat{\delta}^*(\boldsymbol{n}_\perp''', n_3')\right\rangle \\ &\quad \times {\rm e}^{2\pi{\rm i}[(\boldsymbol{n}_\perp - \boldsymbol{n}_\perp')\cdot\boldsymbol{x}_\perp^j + (\boldsymbol{n}_\perp'' - \boldsymbol{n}_\perp''')\cdot\boldsymbol{x}_\perp^l]/B} - P_{\rm 1D}(n_3)P_{\rm 1D}(n_3') \\ &= \left|\frac{1}{B^2}\sum_{\boldsymbol{n}_\perp=-\infty}^{\infty} P_{\rm 3D}(\boldsymbol{n}_\perp, n_3)\,{\rm e}^{2\pi{\rm i}\boldsymbol{n}_\perp\cdot\boldsymbol{s}/B}\right|^2 \delta^{\rm K}_{n_3,n_3'} \equiv |\xi(\boldsymbol{s}, n_3)|^2\delta^{\rm K}_{n_3,n_3'},\end{aligned} \tag{19}$$

where the subscript and superscript are dropped for $\boldsymbol{s}$. Fourier transforms of $\xi(\boldsymbol{s}, k_3)$ will give the three-dimensional PS and correlation function of the density field, and $P_{\rm 1D}(k_3) = \xi(0, k_3)$ (see also Viel et al. 2002). Because of isotropy, the pairwise covariance $\sigma_{\rm 1D}^2(k_3, k_3'; \boldsymbol{s})$ and $\xi(\boldsymbol{s}, k_3)$ depend only on the magnitude of the separation, i.e. $\sigma_{\rm 1D}^2(k_3, k_3'; \boldsymbol{s}) \equiv \sigma_{\rm 1D}^2(k_3, k_3'; s)$ and $\xi(\boldsymbol{s}, k_3) \equiv \xi(s, k_3)$.

If one LOS is sampled from each GRF, the variance of the measured one-dimensional PS $\mathcal{P}_{\rm 1D}(k_3)$ is $\sigma_{\rm 1D}^2(k_3, k_3) = \sigma_{\rm 1D}^2(k_3, k_3; 0) = P_{\rm 1D(\it k3)}^2$, analogous to the three-dimensional case. If $N = m^2$ LOSs are sampled in each GRF on a regular grid as in Section 3, e.g. $\boldsymbol{s} = (\boldsymbol{l}_\perp - \boldsymbol{l}_\perp')\,b$, the covariance becomes

$$\begin{aligned}\sigma_{\rm 1D}^2(n_3, n_3') &= \frac{1}{N^2}\sum_{\boldsymbol{l}_\perp, \boldsymbol{l}_\perp'=0}^{m-1} \sigma_{\rm 1D}^2[n_3, n_3'; (\boldsymbol{l}_\perp - \boldsymbol{l}_\perp')b] = \frac{1}{N^2B^4}\sum_{\boldsymbol{n}_\perp, \boldsymbol{n}_\perp'=-\infty}^{\infty} P_{\rm 3D}(\boldsymbol{n}_\perp, n_3)P_{\rm 3D}(\boldsymbol{n}_\perp', n_3)\left|\sum_{\boldsymbol{l}_\perp=0}^{m-1}{\rm e}^{2\pi{\rm i}(\boldsymbol{n}_\perp - \boldsymbol{n}_\perp')\cdot\boldsymbol{l}_\perp/m}\right|^2 \delta^{\rm K}_{n_3,n_3'} \\ &= \frac{1}{B^4}\sum_{\boldsymbol{n}_\perp, \boldsymbol{j}_\perp=-\infty}^{\infty} P_{\rm 3D}(\boldsymbol{n}_\perp, n_3)P_{\rm 3D}(\boldsymbol{n}_\perp + m\boldsymbol{j}_\perp, n_3)\,\delta^{\rm K}_{n_3,n_3'},\end{aligned} \tag{20}$$

where we have used equation (12) to reach the last line. This result can be easily obtained using equations (5) and (11) as well. It is seen that the summation in the last line of equation (20) runs over only one mode out of every N modes in the Fourier space. If the three-dimensional PS were constant, the variance of the mean PS of the N LOSs would be N times smaller than the variance of the PS of a single LOS. This is coincident with the theory of the variance of the mean. Because the cosmic density field is not a GRF in general, equation (20) is not expected to give an accurate estimate.

If the N LOSs are sampled randomly in each GRF, the summation over LOSs in the second line of equation (20) should be recast to read

$$\sigma_{\rm 1D}^2(n_3, n_3') = \frac{1}{N^2B^4}\sum_{\boldsymbol{n}_\perp, \boldsymbol{n}_\perp'=-\infty}^{\infty} P_{\rm 3D}(\boldsymbol{n}_\perp, n_3)P_{\rm 3D}(\boldsymbol{n}_\perp', n_3)\sum_{j,l=1}^{N}{\rm e}^{2\pi{\rm i}(\boldsymbol{n}_\perp - \boldsymbol{n}_\perp')\cdot(\boldsymbol{x}_\perp^j - \boldsymbol{x}_\perp^l)/B}\,\delta^{\rm K}_{n_3,n_3'}. \tag{21}$$

Because $\boldsymbol{x}_\perp^j$ is randomly distributed, the second sum in equation (21) tends to vanish for a large number of LOSs except when $j = l$. Thus, one obtains

$$\sigma_{\rm 1D}^2(n_3, n_3') \simeq \frac{1}{NB^4}\sum_{\boldsymbol{n}_\perp, \boldsymbol{n}_\perp'=-\infty}^{\infty} P_{\rm 3D}(\boldsymbol{n}_\perp, n_3)P_{\rm 3D}(\boldsymbol{n}_\perp', n_3)\,\delta^{\rm K}_{n_3,n_3'} = \frac{1}{N}P_{\rm 1D}^2(n_3)\,\delta^{\rm K}_{n_3,n_3'}. \tag{22}$$

Again, it shows that the variance of $\mathcal{P}_{\rm 1D}(k_3)$ is inversely proportional to the number of LOSs, but this is valid only for a large number of LOSs randomly sampled in a GRF.

Through the Gaussian case one can find the lowest bound of the uncertainty for a measured one-dimensional PS and the minimum number of LOSs needed for a target precision. For example, to measure the one-dimensional PS of GRFs accurate to 5 per cent on every mode, one needs at least 400 LOSs. However, the cosmic density field has a far more complex covariance due to its non-vanishing trispectrum, and it can have a much larger variance of the measured one-dimensional PS.

The trispectrum introduces an extra term to the covariance of the one-dimensional PS in addition to the Gaussian piece (equation 20), i.e.

$$\sigma_{\rm 1D}^2(n_3, n_3') = \frac{1}{B^4}\sum_{\boldsymbol{n}_\perp, \boldsymbol{j}_\perp=-\infty}^{\infty} P_{\rm 3D}(\boldsymbol{n}_\perp, n_3)P_{\rm 3D}(\boldsymbol{n}_\perp + m\boldsymbol{j}_\perp, n_3)\,\delta^{\rm K}_{n_3,n_3'} + \frac{1}{B}T_{\rm 1D}(n_3, n_3'), \tag{23}$$

where

$$T_{\rm 1D}(n_3, n_3') = \frac{1}{B^6}\sum_{\boldsymbol{n}_\perp, \boldsymbol{j}_\perp, \boldsymbol{n}_\perp', \boldsymbol{j}_\perp'=-\infty}^{\infty} T(\boldsymbol{n}, -\boldsymbol{n} - m\boldsymbol{j}, \boldsymbol{n}', -\boldsymbol{n}' - m\boldsymbol{j}'), \tag{24}$$

$\boldsymbol{n}_\perp$, $\boldsymbol{j}_\perp$, $\boldsymbol{n}'_\perp$, and $\boldsymbol{j}'_\perp$ are the transverse components of $\boldsymbol{n}$, $\boldsymbol{j}$, $\boldsymbol{n}'$, and $\boldsymbol{j}'$, respectively, and $j_3 = j'_3 = 0$. In deriving equation (23), we have made use of equations (5) and (11).

5 LINE-OF-SIGHT LENGTH

Observationally, the length of a LOS is always much less than the size of the observable Universe. For example, in the case of a pencil-beam survey, the length is limited by the depth of the survey. For the Lyα forest, it is often determined by the distance between the starting redshift of Lyα and Lyβ lines. Damped Lyα systems and other astrophysical or instrumental factors can break the spectrum into yet shorter chunks. Even if a very long LOS were obtained, one might still wish to measure the PS on smaller segments to avoid evolutionary effects. To account for the length of LOSs, the one-dimensional PS and its covariance must be reformulated.

For a LOS from $(\boldsymbol{x}_\perp, 0)$ to $(\boldsymbol{x}_\perp, L)$, its Fourier transform is

$$\tilde{\delta}(\boldsymbol{x}_\perp, \tilde{n}_3) = \int_0^L \delta(\boldsymbol{x}_\perp, x_3) \mathrm{e}^{-2\pi \mathrm{i} \tilde{n}_3 x_3/L} \mathrm{d}x_3 = \frac{L}{B^3} \sum_{\boldsymbol{n}=-\infty}^{\infty} \hat{\delta}(\boldsymbol{n}) \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{n}_\perp \cdot \boldsymbol{x}_\perp / B} \frac{\mathrm{e}^{2\pi \mathrm{i}(n_3 L/B - \tilde{n}_3)} - 1}{2\pi \mathrm{i}(n_3 L/B - \tilde{n}_3)}, \tag{25}$$

so that

$$\left\langle \tilde{\delta}(\boldsymbol{x}_\perp, \tilde{n}_3) \tilde{\delta}^*(\boldsymbol{x}_\perp, \tilde{n}'_3) \right\rangle = \frac{L^2}{B^3} \sum_{\boldsymbol{n}=-\infty}^{\infty} P_{3\mathrm{D}}(\boldsymbol{n}) w(n_3, \tilde{n}_3) w(n_3, \tilde{n}'_3), \tag{26}$$

where

$$w(n_3, \tilde{n}_3) = (-1)^{\tilde{n}_3} \frac{\sin[(n_3 L/B - \tilde{n}_3)\pi]}{(n_3 L/B - \tilde{n}_3)\pi}. \tag{27}$$

The function $w(n_3, \tilde{n}_3)$ is a window function that mixes Fourier modes of the density field along the n_3 direction into the one-dimensional Fourier mode at $\tilde{n}_3$. Because $w(n_3, \tilde{n}_3)$ and $w(n_3, \tilde{n}'_3)$ are not orthogonal to each other, the term $\langle \tilde{\delta}(\boldsymbol{x}_\perp, \tilde{n}_3) \tilde{\delta}^*(\boldsymbol{x}_\perp, \tilde{n}'_3) \rangle$ is no longer diagonal with respect to $\tilde{n}_3$ and $\tilde{n}'_3$. However, it remains diagonal-dominant.

One may define the observed one-dimensional PS at $\boldsymbol{x}_\perp$ as

$$\widetilde{\mathcal{P}}_{1\mathrm{D}}(\boldsymbol{x}_\perp, \tilde{n}_3) = \left\langle |\tilde{\delta}(\boldsymbol{x}_\perp, \tilde{n}_3)|^2 \right\rangle \Big/ L. \tag{28}$$

The ensemble-averaged one-dimensional PS is

$$\widetilde{P}_{1\mathrm{D}}(\tilde{n}_3) = \langle \widetilde{\mathcal{P}}_{1\mathrm{D}}(\boldsymbol{x}_\perp, \tilde{n}_3) \rangle = \frac{L}{B^3} \sum_{\boldsymbol{n}=-\infty}^{\infty} P_{3\mathrm{D}}(\boldsymbol{n}) w^2(n_3, \tilde{n}_3). \tag{29}$$

For GRFs, the covariance of $\widetilde{\mathcal{P}}_{1\mathrm{D}}(\tilde{n}_3)$ is

$$\begin{aligned} \tilde{\sigma}^2_{1\mathrm{D}}(\tilde{n}_3, \tilde{n}'_3) &= \langle [\widetilde{\mathcal{P}}_{1\mathrm{D}}(\boldsymbol{x}_\perp, \tilde{n}_3) - \widetilde{P}_{1\mathrm{D}}(\tilde{n}_3)][\widetilde{\mathcal{P}}_{1\mathrm{D}}(\boldsymbol{x}_\perp, \tilde{n}'_3) - \widetilde{P}_{1\mathrm{D}}(\tilde{n}'_3)] \rangle \\ &= \frac{L^2}{B^6} \sum_{\boldsymbol{n}, \boldsymbol{n}'=-\infty}^{\infty} P_{3\mathrm{D}}(\boldsymbol{n}) P_{3\mathrm{D}}(\boldsymbol{n}') w(n_3, \tilde{n}_3) w(n_3, \tilde{n}'_3) w(n'_3, \tilde{n}_3) w(n'_3, \tilde{n}'_3). \end{aligned} \tag{30}$$

As expected, the covariance is not diagonal because of the window function $w(n_3, \tilde{n}_3)$.

6 *N*-BODY TESTS

Even at moderately high redshift, the cosmic density field is already quite non-Gaussian on scales below $10\,h^{-1}$ Mpc. When projected in one dimension, the small-scale non-Gaussianity will obviously affect the measured one-dimensional PS on much larger scales. The non-vanishing trispectrum introduces an extra term, $T_{1\mathrm{D}}(k_3, k'_3)$, to the covariance of the one-dimensional PS (see equation 23). Although it is possible to derive an approximate trispectrum based on the halo model (e.g. Cooray & Hu 2001), the one-dimensional projection, unfortunately, obscures the contribution of the trispectrum to the one-dimensional PS. Therefore, numerical simulations are necessary for the study.

Three N-body simulations of 256^3 cold dark matter (CDM) particles are used to quantify the covariance of the one-dimensional PS. The model parameters are largely consistent with *Wilkinson Microwave Anisotropy Probe* (*WMAP*) results (Spergel et al. 2003), e.g. $(\Omega, \Omega_\mathrm{b}, \Omega_\Lambda, h, \sigma_8, n) = (0.27, 0.04, 0.73, 0.71, 0.85, 1)$, where Ω is the cosmic matter density parameter, Ω_b is the baryon density parameter, Ω_Λ is the energy density parameter associated with the cosmological constant, σ_8 is the rms density fluctuation within a radius of $8\,h^{-1}$ Mpc, and n is the power spectral index. The box sizes of the simulations are $128\,h^{-1}$ Mpc (labelled as B128), $256\,h^{-1}$ Mpc (B256) and $512\,h^{-1}$ Mpc (B512). The baryon density parameter Ω_b is used only for the purpose of calculating the transfer function using LINGER (Ma & Bertschinger 1995), which is then read by GRAFIC2 (Bertschinger 2001) to generate the initial condition. The CDM particles are evolved from $z = 44.5$ to the present using GADGET (Springel, Yoshida & White 2001). Additionally, GRFs are generated with the same box sizes as those of N-body simulations for comparison and are labelled as R128, R256 and R512, respectively.

The simulations produce snapshots at $z = 3$ and 0. For each snapshot, the particles are assigned to a density grid of 512^3 nodes using the triangular-shaped cloud scheme (Hockney & Eastwood 1981). LOSs are then sampled on the x_1–x_2 plane, and each LOS is extracted with full resolution along the x_3-axis. Because the density grid has a finite resolution, both the coordinates and the wavenumbers are discrete

and finite. The equations in Sections 2–5 need to be modified accordingly, so that they do not sum over non-existing modes. For example, equation (10) becomes

$$P_{\rm 1D}(n_3) = \frac{1}{B^2} \sum_{\boldsymbol{n}_\perp=-M/2}^{M/2} P_{\rm 3D}(\boldsymbol{n}_\perp, n_3), \tag{31}$$

where $M = 512$ is the number of nodes of the density grid in transverse directions.

For brevity and the purpose of comparing the covariance, we introduce the following three covariance matrices.

(i) The normalized pairwise covariance between two LOSs

$$C(k_3, k_3'; s) = \sigma^2_{\rm 1D}(k_3, k_3'; s)[P_{\rm 1D}(k_3)P_{\rm 1D}(k_3')]^{-1}, \tag{32}$$

where s is the transverse separation between the two LOSs.

(ii) The normalized covariance of the estimated one-dimensional PS

$$C(k_3, k_3') = \sigma^2_{\rm 1D}(k_3, k_3')[P_{\rm 1D}(k_3)P_{\rm 1D}(k_3')]^{-1}, \tag{33}$$

where $\sigma^2_{\rm 1D}(k_3, k_3')$ is the covariance of the mean PS of N LOSs (e.g. equation 11), and we have omitted its dependence on the number of LOSs.

(iii) The reduced covariance

$$\hat{C}(k_3, k_3') = C(k_3, k_3')[C(k_3, k_3)C(k_3', k_3')]^{-1/2}. \tag{34}$$

For GRFs, all these covariances are diagonal in matrix representation, if the length of the LOS is the same as the size of the simulation box. In addition, we have $C(k_3, k_3; 0) = 1$ and $C(k_3, k_3) = N^{-1}$, where N is the number of LOSs that are sampled for estimating the one-dimensional PS. The advantage of $\hat{C}(k_3, k_3')$ is that $\hat{C}(k_3, k_3) = 1$ for all fields, so that they can be compared with each other in a single (grey) scale.

The normalized pairwise covariance $C(k_3, k_3'; s)$ is shown in Fig. 1 for GRFs R512 and the simulation B512 at $z = 3$. The covariances for GRFs are averaged over an ensemble of 2000 random realizations, while those for the simulation are averaged over 2000 pairs of LOSs from a single field. The behaviour of the covariances is consistent with the expectation. Namely, $C(k_3, k_3'; 0) \simeq \delta^{\rm K}_{n_3,n_3'}$ and $C(k_3, k_3; s)$ decreases as the separation s increases. The simulation does deviate from GRFs because of the non-Gaussianity, which increases the variance $C(k_3, k_3; s)$. Fig. 2 compares the diagonal elements $C(k_3, k_3; s)$ with those calculated using equation (19). We note in passing that the expected values of $C(k_3, k_3; 8\,h^{-1}\,{\rm Mpc})$ are so close to 0 that they are even below the rms value of the off-diagonal elements of $C(k_3, k_3'; 8\,h^{-1}\,{\rm Mpc})$ for 2000 GRFs. Hence, it is practically difficult to recover three-dimensional statistics from $\sigma_{\rm 1D}(k_3, k_3; s)$ or $\xi(s, k_3)$ if the LOSs are too far apart. In theory, the modes of two LOSs are always correlated as $k_3 \to 0$, regardless of their separation. However, large-scale LOS modes are greatly affected by small-scale three-dimensional perturbations, which are nearly uncorrelated for large separations. As such, the correlation for $s \gtrsim 8\ h^{-1}$ Mpc is so weak that it will not be easily detected against statistical uncertainties even for a sizable sample of 2000 pairs of LOSs.

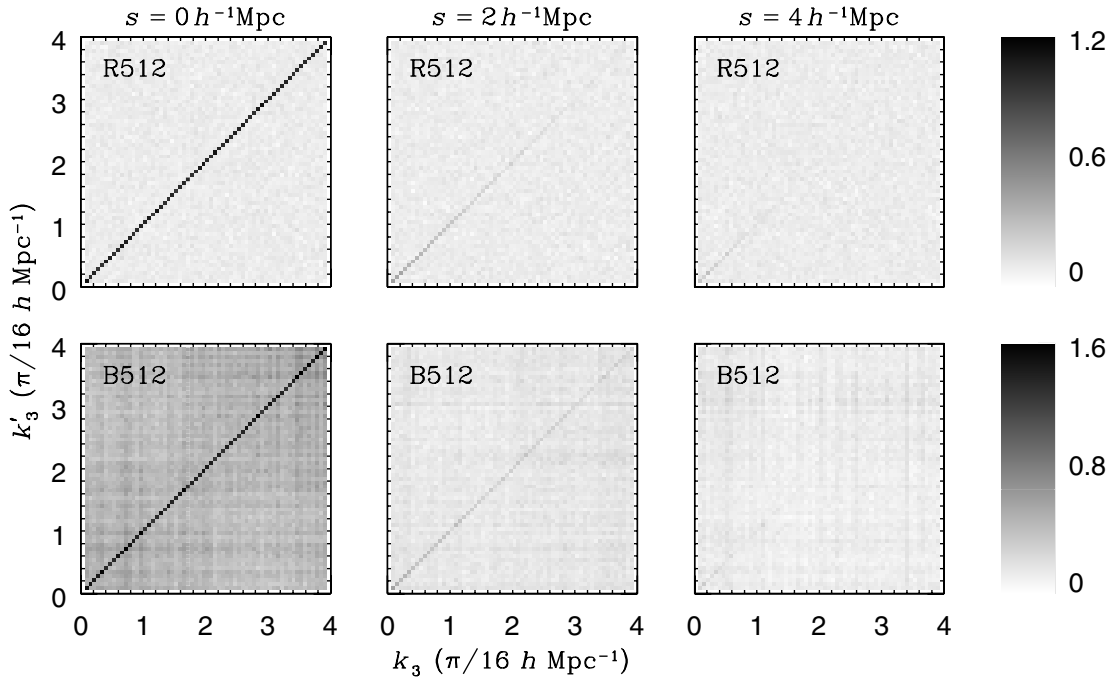

Figure 1. Normalized pairwise covariance matrices $C(k_3, k_3'; s)$ (see equation 32) in grey-scale. The covariances are calculated over 2000 pairs of LOSs sampled at a fixed separation s from the simulation B512 at $z = 3$ (lower panels) and from 2000 GRFs (R512) that have the same box size and three-dimensional mass PS as the simulation (upper panels). The grey-scale extends to -0.1, which corresponds to white.

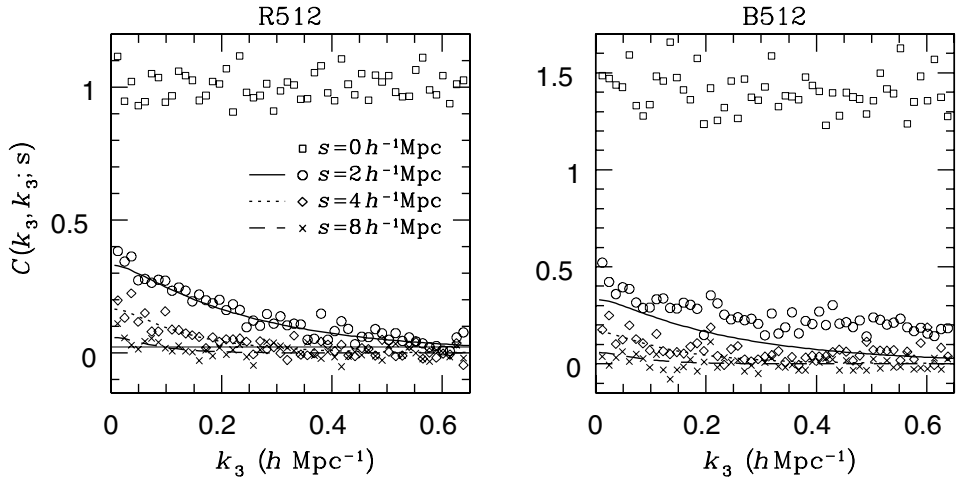

Figure 2. Diagonal elements of the covariance matrices $C(k_3, k_3; s)$. The horizontal solid line in the left panel marks the rms value of the off-diagonal elements in the $s = 8\ h^{-1}$ Mpc case. Other lines are from direct summations of the three-dimensional mass PS using equation (19). The symbols are measured from GRFs R512 (left) and the simulation B512 at $z = 3$ (right). For GRFs, $C(k_3, k_3; 0)$ is expected to be unity.

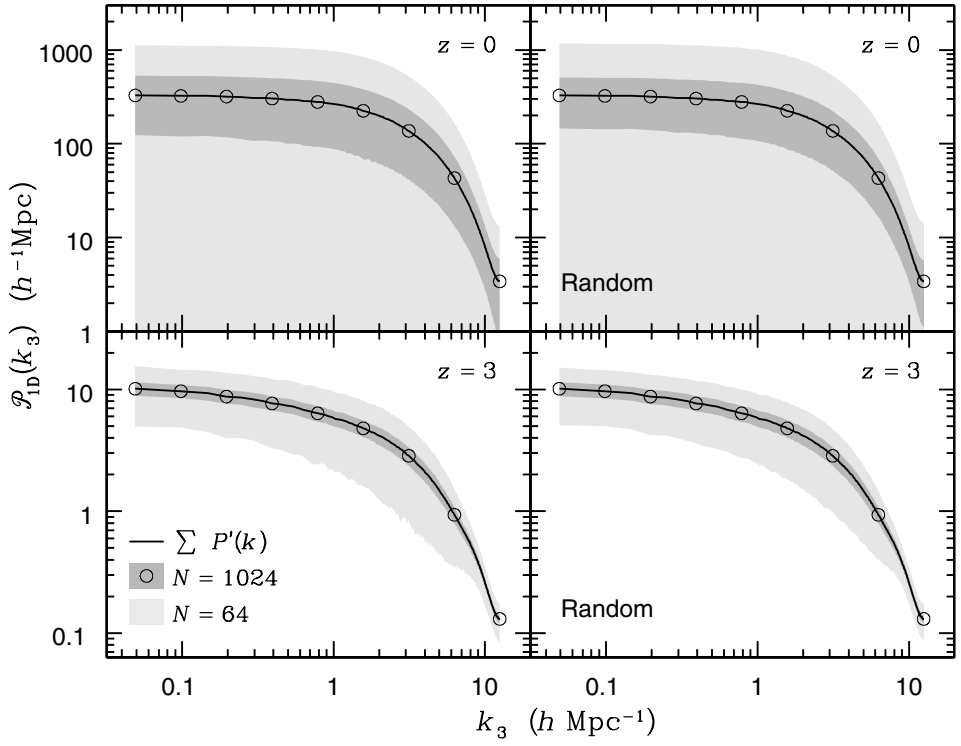

Figure 3. Estimated one-dimensional mass PSs $\mathcal{P}_{1D}(k_3)$. The PSs are measured by averaging over 64 (light grey) and 1024 (dark grey) LOSs that are drawn from the B128 simulation. Shaded areas mark 1 σ dispersions of the PSs among 2000 (light grey) and 256 (dark grey, as 256×1024 exhausts all the 512^2 LOSs) distinct drawings. Circles are the mean PSs for dark grey areas, i.e. the mean of 512^2 LOSs. Solid lines are results of a direct summation of the three-dimensional mass PS using equation (31). The LOSs are sampled on a grid with fixed spacing in the left panels but drawn randomly in the right panels.

Hence, Figs 1 and 2 suggest that LOSs sampled in a single cosmic density field are practically independent of each other as long as they are well separated. This condition is easily met. For example, McDonald et al. (2004) use 3035 $z > 2.3$ quasar spectra from the Sloan Digital Sky Survey (SDSS; over 2627 deg^2) to determine the Lyα flux PS. Assuming *WMAP* parameters, one will find that the mean separation between LOSs is $\sim$60 comoving h^{-1} Mpc such that the covariance arisen from the correlation between LOSs can be safely neglected.

Fig. 3 shows the one-dimensional PS measured by averaging over 64 and 1024 LOSs from the B128 simulation. Although the mean PS of all the 512^2 LOSs agrees with the result of a direct summation of the three-dimensional PS using equation (31), the deviation of the PS for any particular group of 64 or 1024 LOSs is substantial, especially at $z = 0$. The variance is smaller at $z = 3$ than at $z = 0$ because the cosmic density field is more Gaussian earlier on, and it is roughly inversely proportional to the number of LOSs. This can be seen better by comparing

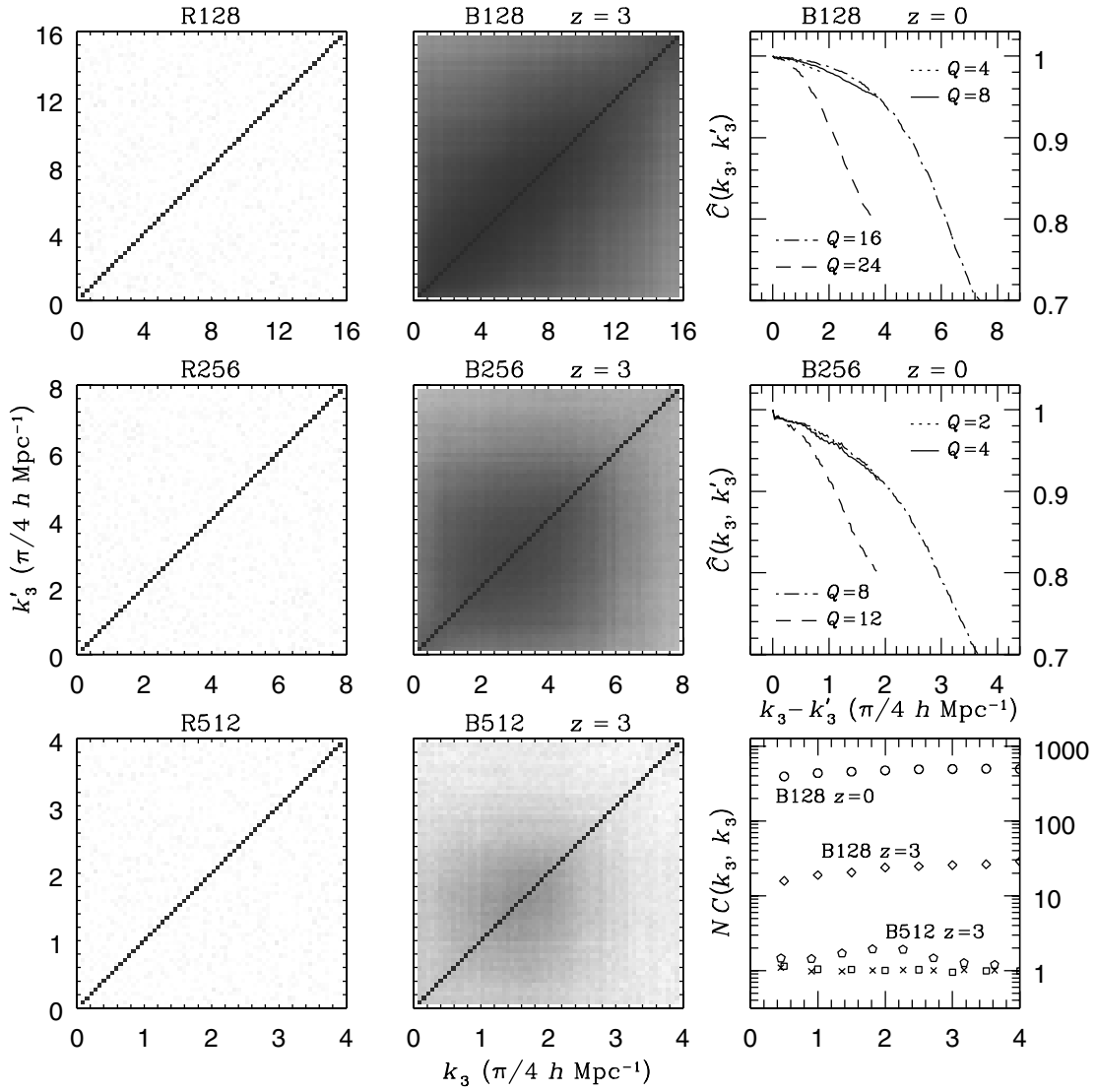

Figure 4. Reduced covariances $\hat{C}(k_3, k_3')$ (see equation 34) averaged over 2000 groups, each of which consists of 64 LOSs ($N = 64$). For each panel in the left, the LOSs are randomly drawn from a single GRF that has a box size of $128\,h^{-1}$ Mpc (R128), or $256\,h^{-1}$ Mpc (R256) or $512\,h^{-1}$ Mpc (R512). The GRFs have the same three-dimensional mass PS as their corresponding simulations at $z = 0$, but note that $\hat{C}(k_3, k_3')$ is independent of redshift for GRFs. Similarly, the middle column is for simulations at $z = 3$. The covariances are shown in a linear grey-scale with black being 1.2 and white less than or equal to 0. At $z = 0$, the normalized covariances become much less diagonally dominant than the B128 $z = 3$ panel in the same grey-scale, so only four cross-sections along $Q = (k_3 + k_3')/(\pi/4\,h\ \mathrm{Mpc}^{-1})$ are plotted for B128 and B256. Diagonal elements $C(k_3, k_3)$ (see equation 33) of R128 (squares), B128 at $z = 0$ (circles), B128 at $z = 3$ (diamonds), R512 (crosses) and B512 at $z = 3$ (pentagons) are shown with a multiplication factor N in the lower-right panel. For GRFs, $NC(k_3, k_3) = 1$.

the lower-right panel of Fig. 4 with that of Fig. 5. However, even at $z = 3$ the variance of the one-dimensional PS is still much higher than the variance for GRFs, $N^{-1}P_{\mathrm{1D}}^2(k_3)$, which indicates a heavy contribution from the trispectrum. The formulae in Sections 3–5 often assume that LOSs are sampled on a grid with fixed spacing, which may not be applicable to realistic data such as inverted densities from the Lyα forest (Nusser & Haehnelt 1999; Zhan 2003). Therefore, we sample the LOSs in two ways in Fig. 3: grid sampling and random sampling. Because no significant difference is observed, random sampling can be safely applied in the rest of this paper.

The normalized covariances $\hat{C}(k_3, k_3')$ and $C(k_3, k_3')$ are quantified in Figs 4 ($N = 64$) and 5 ($N = 1024$). The left column in each figure shows the covariances of the spatially averaged one-dimensional PS from a single GRF that has the same box size and three-dimensional PS as its corresponding simulation. Clearly, the covariances based on spatial average are nearly diagonal with unity diagonal elements. This is in agreement with the expectations based on ensemble average for GRFs and is consistent with the ergodicity argument. The middle column is similar to the left column except that the density fields are from simulations at $z = 3$. The modes in the simulated density fields are strongly correlated, so that the covariances are no longer diagonal. In other words, the trispectrum is non-vanishing for the cosmic density field, as it is the only term that contributes to off-diagonal elements in the covariance. The cosmic density field becomes so non-Gaussian at $z = 0$ that grey-scale figures of the covariances will not be readable. Hence, we only plot four cross-sections perpendicular the diagonal with $Q =$

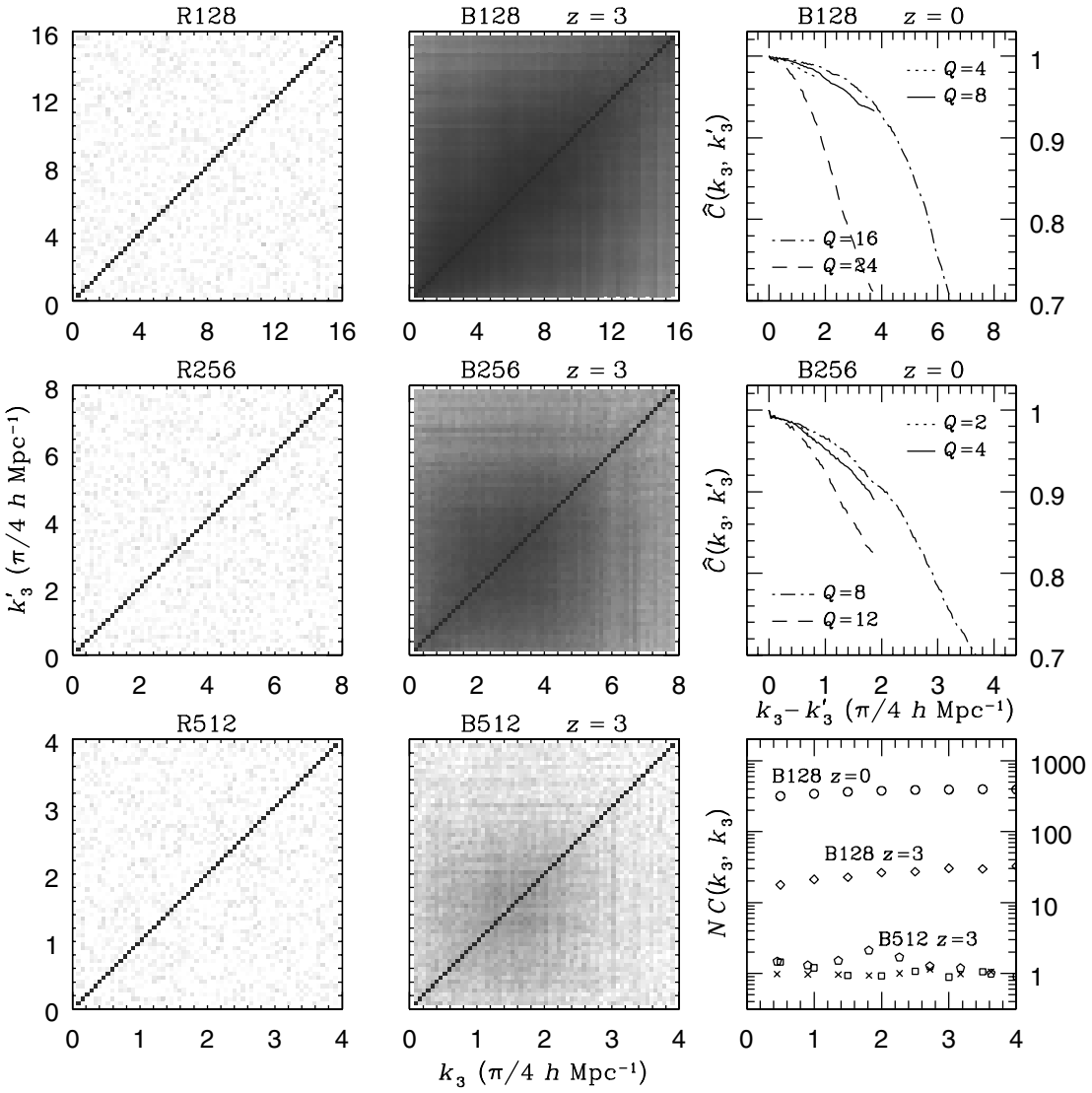

Figure 5. The same as Fig. 4, except that the covariances $\hat{C}(k_3, k_3')$ and $C(k_3, k_3)$ are averaged over 256 groups, each of which consists of 1024 LOSs ($N = 1024$).

$(k_3 + k_3')/(\pi/4\,h\ \mathrm{Mpc}^{-1})$ for B128 and B256 in the right column. The dominance of the diagonals suggested by these cross-sections is actually weaker than that in the middle column, which can be seen by contrasting the cross-sections for B128 at $z = 0$ in Fig. 4 with that for B128 at $z = 3$ in Fig. 6.

The non-Gaussianity is reflected not only in the correlations between different modes but also in the variance of the one-dimensional PS, i.e. the diagonal elements of the normalized covariance $C(k_3, k_3')$. The lower-right panels of Figs 4 and 5 compare $C(k_3, k_3)$ for five different density fields. The variance from the simulation B128 is orders of magnitude higher than the Gaussian value of $N^{-1}P_{\mathrm{1D}}^2(k_3)$, and it grows as the non-Gaussianity becomes stronger toward $z = 0$. As a result, the sample variance error estimated for GRFs is much lower than what can be actually measured from the cosmic density field. According to equation (23), both aliasing and the trispectrum contribute to the variance. Because the GRFs have the same three-dimensional PS and are sampled in the same way as the simulations, their near-Gaussian variances of spatially averaged one-dimensional PSs suggest that the contribution of the aliasing effect is negligible. Comparisons of $C(k_3, k_3)$ for the same density field but with different sizes of sample ($N = 64$ and 1024) confirm the observation in Fig. 3 that the variance of the one-dimensional PS scales roughly as N^{-1}. This means that even though the non-Gaussianity drives up the sample variance error of the measured one-dimensional PS, one can still reduce the error by sampling a large number of LOSs.

The statistics of a large sample may sometimes be dominated by a small fraction of extreme cases. It is inevitable that some of the LOSs in our sample go through high-density halo structures, which may have significant contributions to the covariance matrices. To determine if a LOS falls in this category, one has to assess the density profile intercepted by the LOS. We follow a much simpler approach by imposing a threshold to the mass column densities ρ_{col} of the LOSs. The one-point distribution function of column density contrast $\rho_{\mathrm{col}}/\bar{\rho}_{\mathrm{col}}$ from the simulation B128 (upper-left panel of Fig. 6) peaks at $\rho_{\mathrm{col}}/\bar{\rho}_{\mathrm{col}} = 1$ as expected, and there clearly is a long-tail distribution (as compared to

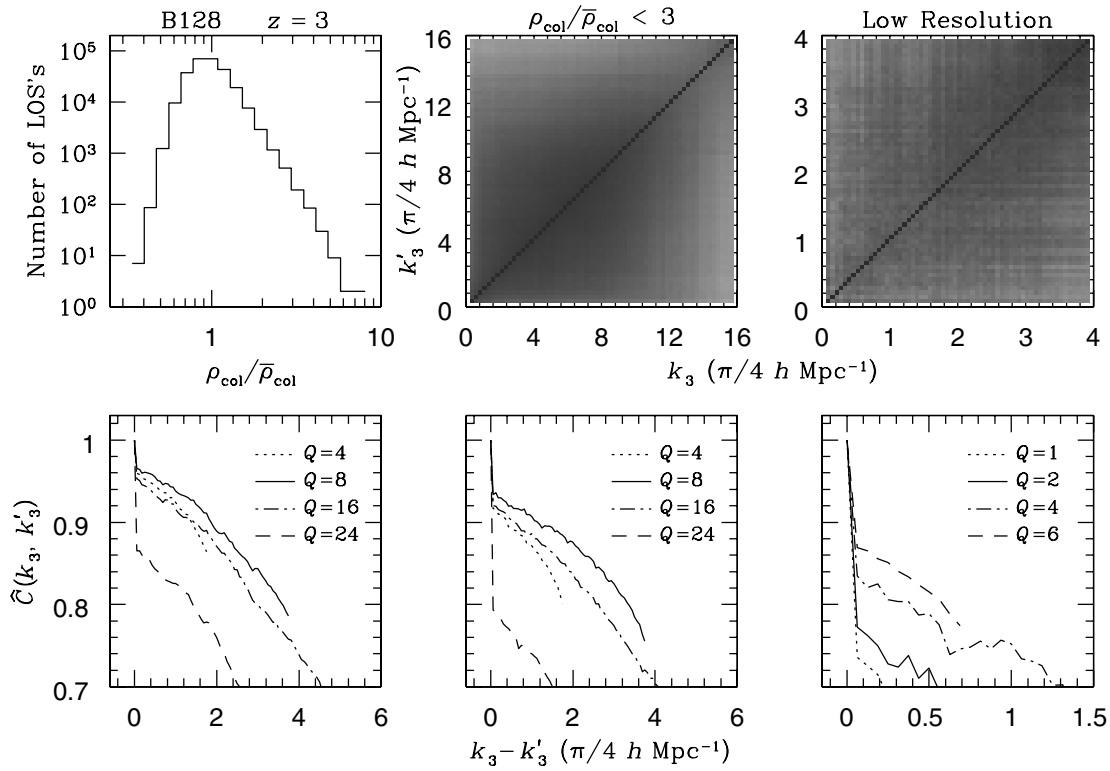

Figure 6. The effect of the non-Gaussianity and resolution on the covariance. The left column shows the distribution function of column density $\rho_{\rm col}/\bar{\rho}_{\rm col}$ (upper panel) and cross-sections of the reduced covariance $\hat{C}(k_3, k_3')$ along $Q = (k_3 + k_3')/(\pi/4\,h\ {\rm Mpc}^{-1})$ (lower panel) for B128 at $z = 3$. The middle column shows $\hat{C}(k_3, k_3')$ (upper panel) and its cross-sections (lower panel) for the same simulation output but with an exclusion of 318 LOSs that have $\rho_{\rm col}/\bar{\rho}_{\rm col} \geqslant 3$. The reduced covariances in the first two columns are calculated in the same way as in Fig. 4, i.e. with 2000 groups and $N = 64$. Thus, about half of the 318 LOSs would be selected without the criterion $\rho_{\rm col}/\bar{\rho}_{\rm col} < 3$. The right column is similar to the middle column, but the density is assigned on a grid of 128^3 nodes. All the 128^2 LOSs are selected and divided into 256 groups with $N = 64$. The grey-scale is the same as Fig. 4.

Gaussian or lognormal distributions) of $\rho_{\rm col}/\bar{\rho}_{\rm col} \gtrsim$ a few. We recalculate the covariance $\hat{C}(k_3, k_3')$ with a selection criterion that the column density of each LOS $\rho_{\rm col}/\bar{\rho}_{\rm col} < 3$. This criterion excludes 318 LOSs from the selection of LOSs for the statistics. As a total of 2000×64 LOSs are randomly selected from 512^2 candidates, only half of the 318 LOSs (i.e. 0.1 per cent of the sample) would be selected without the criterion. The result is shown in the upper-middle panel of Fig. 6, which is visually similar to the B128 $z = 3$ panel in Fig. 4. However, the cross-sections of $\hat{C}(k_3, k_3')$ with the selection criterion (lower-middle panel of Fig. 6) do demonstrate a several per cent reduction of the off-diagonal elements in contrast to those without the selection criterion (lower-left panel). In other words, the 0.1 per cent $\rho_{\rm col}/\bar{\rho}_{\rm col} \geqslant 3$ LOSs contribute at least 10 times as much to the covariance per LOS as $\rho_{\rm col}/\bar{\rho}_{\rm col} < 3$ LOSs do. It is also useful to have a comparison of the normalized covariance $C(k_3, k_3')$ because, unlike the reduced covariance $\hat{C}(k_3, k_3')$ that only tells the relative strength between the diagonal and off-diagonal elements, $C(k_3, k_3')$ reflects absolute values of the covariance up to a factor of the product of PSs at k_3 and k_3'. Fig. 7 shows the diagonal elements of the normalized covariance for B128. It is seen that the variance $C(k_3, k_3)$ of the one-dimensional PS is indeed reduced by roughly a factor of 2 on all scales when the 0.1 per cent $\rho_{\rm col}/\bar{\rho}_{\rm col} \geqslant 3$ LOSs are excluded.

For a fixed number of particles, the simulation box sets a cut-off scale, below which fluctuations cannot be properly represented in low-density regions such as the Lyα forest. In other words, the number of particles and the size of the simulation determine the highest-wavenumber modes that are included in calculations of the one-dimensional PS and its covariance. Because the non-Gaussianity is stronger at smaller scales, a larger simulation box cuts off more small-scale fluctuations and may cause the correlation to appear weaker in Figs 4 and 5. To test this, we assign the density field of the simulation B128 at $z = 3$ on a grid of 128^3 nodes. The spatial resolution is the same as the simulation B512 on a grid of 512^3 nodes. The covariance $\hat{C}(k_3, k_3')$ is calculated in the same way as those in Fig. 4 but with fewer groups of LOSs. Each group still has 64 LOSs. The results are shown in the right column of Fig. 6 and in Fig. 7. Evidently, there is a significant reduction of the correlations between different modes as well as the variance of the one-dimensional PS. Indeed, the variance from the low-resolution calculation is close to that from the B512 simulation. Thus, the apparent resemblance between B512 and the GRF R512 in Figs 4 and 5 is mostly due to the low resolution of the large-box simulation.

It is expected from equation (30) that the covariance matrix will not be diagonal if the length of the LOS is less than the size of the simulation box. Fig. 8 shows the covariances that are calculated in the same way as those in Fig. 4 for the GRF R512 and the simulation B512, except that each LOS is only $128\,h^{-1}$ Mpc long. The effect of the length is not visible for the GRF, but it does increase the correlation between different modes and doubles the variance of the one-dimensional PS for the simulation B512. For real observations, the LOS length

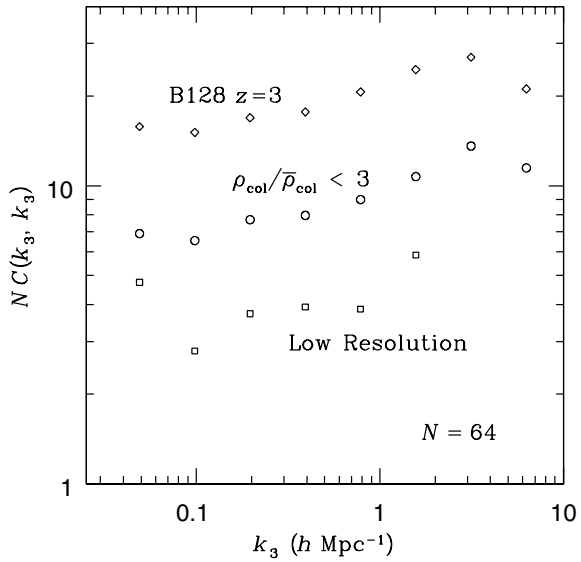

Figure 7. The effect of the non-Gaussianity and resolution on the variance, i.e. diagonal elements of the covariance $\sigma^2_{\rm 1D}(k_3, k'_3)$. Diamonds are the normalized variance $C(k_3, k_3)$ (multiplied by N) for B128 at $z = 3$. Circles correspond to the middle column of Fig. 6, which imposes the selection criterion $\rho_{\rm col}/\bar{\rho}_{\rm col} < 3$. Squares are from the low-resolution calculation in the right column of Fig. 6.

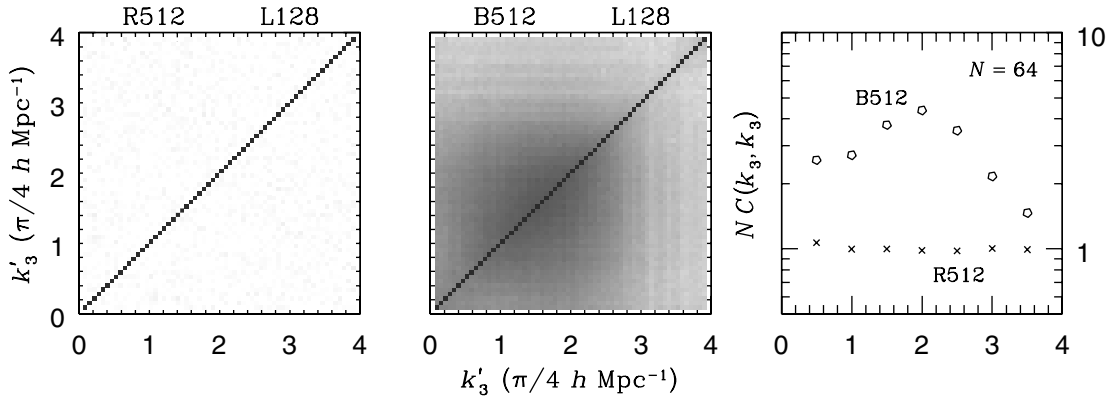

Figure 8. The same as the bottom row of Fig. 4, but with length of LOSs $L = 128\,h^{-1}$ Mpc.

is always much less than the size of the observable Universe, so that the window function $w(k_3, \tilde{k}_3)$ in the LOS direction will cause stronger mixing of modes and more pronounced increase of correlation and variance.

7 CONCLUSIONS

We have investigated the covariance of the one-dimensional mass PS for both GRFs and simulated density fields. The non-linear evolution and the non-Gaussianity of the cosmic density field on small scales (see also Zhan, Jamkhedkar & Fang 2001; Smith et al. 2003; Zaldarriaga, Scoccimarro & Hui 2003; Zhan 2003) have caused the correlation between the fluctuations on different scales and increased the cosmic variance of the one-dimensional PS. Because of this, a large number of LOSs are needed to accurately measure the one-dimensional PS and recover the three-dimensional PS.

The length of LOSs introduces a window function in the LOS direction, which mixes neighboring Fourier modes in the cosmic density field. Fig. 8 has demonstrated for simulations that the length of the LOSs (a quarter of the simulation box size) does affect the covariance of the one-dimensional PS. The covariance of the observed one-dimensional PS will receive more contributions from this effect, because in practice the length of LOSs suitable for study is always much less than the size of the observable Universe.

One may reduce the cosmic variance by binning independent Fourier modes. However, because the modes of the cosmic density field are strongly correlated, binning will be less effective in reducing the sample variance error. On the other hand, the non-Gaussian behaviour of the covariance provides important information of the field such as the trispectrum.

Our analysis shows that even with 64 LOSs at $z = 0$, the cosmic variance on the PS can be several times of the PS itself (see Fig. 4). Thus, the extra power on the scale of 128 h^{-1} Mpc ($\Omega = 1$, $\Omega_\Lambda = 0$) measured by Broadhurst et al. (1990) from four samples of pencil-beam galaxy

surveys can still be consistent with standard cosmogonies (with considerations of effects such as redshift distortion and non-Gaussianity; see, for example, Kaiser & Peacock 1991; Park & Gott 1991; van de Weygaert 1991; Amendola 1994). One may also infer the one-dimensional density field and measure the PS from the Lyα forest. Because the Lyα flux is the direct observable, many works have been based on flux statistics, which are then used along with simulations to constrain cosmology. Using N-body simulations and the pseudo-hydro technique, Zhan (2003) find that the mapping between the nearly-Gaussian Lyα flux PS (Zhan 2004) and the underlying one-dimensional mass PS (in redshift space) has a much larger scatter than that between the flux PS and the one-dimensional linear mass PS. This is expected because the one-dimensional PS of the cosmic density field has a variance much higher than the Gaussian variance. Therefore, one can measure the Lyα flux PS precisely with a small sample of LOSs, but to recover the mass PS to the same precision more LOSs are needed.

ACKNOWLEDGMENTS

We thank R. Davé for useful discussions and the referee for helpful comments. DJE was supported by National Science Foundation AST-0098577 and an Alfred P. Sloan Research Fellowship.

REFERENCES

Amendola L., 1994, ApJ, 430, L9
Bertschinger E., 2001, ApJS, 137, 1
Broadhurst T. J., Ellis R. S., Koo D. C., Szalay A. S., 1990, Nat, 343, 726
Cooray A., Hu W., 2001, ApJ, 554, 56
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
Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, ApJ, 581, 20
Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23
Gnedin N. Y., Hamilton A. J. S., 2002, MNRAS, 334, 107
Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles. McGraw-Hill, New York, p. 152
Hui L., Stebbins A., Burles S., 1999, ApJ, 511, L5
Kasier N., Peacock J. A., 1991, ApJ, 379, 482
Kim T.-S., Viel M., Haehnelt M. G., Carswell R. F., Cristiani S., 2004, MNRAS, 347, 355
Lumsden S. L., Heavens A. F., Peacock J. A., 1989, MNRAS, 238, 293
Ma C.-P., Bertschinger E., 1995, ApJ, 455, 7
McDonald P., 2003, ApJ, 585, 34
McDonald P., Miralda-Escudé J., 1999, ApJ, 518, 24
McDonald P. et al., 2004, ApJ, submitted (astro-ph/0405013)
Mandelbaum R., McDonald P., Seljak U., Cen R., 2003, MNRAS, 344, 776
Meiksin A., White M., 1999, MNRAS, 308, 1179
Nusser A., Haehnelt M., 1999, MNRAS, 303, 179
Park C., Gott J. R. III, 1991, MNRAS, 249, 288
Peebles P. J. E., 1980, The Large Scale Structure of the Universe. Princeton Univ. Press Princeton, p. 150
Rollinde E., Petitjean P., Pichon C., Colombi S., Aracil B., D'Odorico V., Haehnelt M. G., 2003, MNRAS, 341, 1279
Smith R. E. et al., 2003, MNRAS, 341, 1311
Spergel D. N. et al., 2003, ApJS, 148, 175
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
van de Weygaert R., 1991, MNRAS, 249, 159
Viel M., Matarrese S., Mo H. J., Haehnelt M. G., Theuns T., 2002, MNRAS, 329, 848
Zaldarriaga M., Scoccimarro R., Hui L., 2003, ApJ, 590, 1
Zhan H., 2003, MNRAS, 344, 935
Zhan H., 2004, PhD Thesis, Univ. Arizona (astro-ph/0408379)
Zhan H., Jamkhedkar P., Fang L.-Z., 2001, ApJ, 555, 58

This paper has been typeset from a TEX/LATEX file prepared by the author.