Ellipsoidal collapse and the redshift-space probability distribution function of dark matter

Tsz Yan Lam⋆ and Ravi K. Sheth⋆

Department of Physics & Astronomy, University of Pennsylvania, 209 S. 33rd Street, Philadelphia, PA 19104, USA

Accepted 2008 June 20. Received 2008 June 12; in original form 2008 May 8

ABSTRACT
We use the physics of ellipsoidal collapse to model the probability distribution function (PDF) of the smoothed dark matter density field in real and redshift space. We provide a simple approximation to the exact collapse model which clearly shows how the evolution can be thought of as a modification of the spherical evolution model as well as of the Zeldovich approximation. In essence, our model specifies how the initial smoothed overdensity and shear fields can be used to determine the shape and the size of the region at later times. We use our parametrization to extend previous work on the real-space PDF so that it predicts the redshift-space PDF as well. Our results are in good agreement with the measurements of the PDF in simulations of clustering from Gaussian initial conditions down to scales on which the rms fluctuation is slightly greater than unity. We also show how the highly non-Gaussian non-linear redshifted density field in a numerical simulation can be transformed so that it provides an estimate of the shape of the initial real-space PDF. When applied to our simulations, our method recovers the initial Gaussian PDF, provided the variance in the non-linear smoothed field is less than 4.

Key words: methods: analytical – dark matter – large scale structure of the Universe.

1 INTRODUCTION
The probability distribution function (PDF; some authors prefer to call this the probability density function) specifies the probability that a randomly placed cell of specified size and shape contains a certain specified density. As gravitational instability alters the spatial distribution of dark matter, the PDF of the dark matter density field evolves. Broadly speaking, there are two approaches to modelling this evolution: one is to study the evolution of the lower order moments of the PDF and the other studies the shape of the PDF in its entirety.

Typically, the first approach attempts to describe the evolution exactly, using methods derived from perturbation theory, but then the PDF is often computed approximately, upon truncating higher order terms in the perturbation theory (Bernardeau 1994; Scoccimarro & Frieman 1999; Bernardeau et al. 2002). The second uses simple approximations to the exact evolution to derive tractable results; in almost all cases, these specify local transformations which describe how quantities in the initial field determine the evolved density. The best studied of these approximations are the spherical evolution model (Bernardeau 1994; Protogeros & Scherrer 1997; Fosalba & Gaztanaga 1998; Gaztañaga & Croft 1999; Betancort-Rijo & López-Corredoira 2002), or the Zeldovich approximation (Betancort-Rijo 1991; Padmanabhan & Subramanian 1993; Kofman et al. 1994) and its extension to the ellipsoidal collapse model (Ohta, Kayo & Taruya 2004; Lam & Sheth 2008). In principle, there is a limit to how well approaches based on deterministic transformations can fare (Sheth 1998), but, provided one restricts attention to large enough scales, this is a small effect (Lam & Sheth 2008).

The associated PDF in redshift, rather than real space, is less well-studied. In this case too, the PDF is described in terms of its moments (Hivon et al. 1995; Bernardeau et al. 2002), using the spherical model (Scherrer & Gaztangaga 2001), or using the Zeldovich approximation (Hui, Kofman & Shandarin 2000). In addition, Watts & Taylor (2001) describe a calculation of the PDF based on the second-order perturbation theory. The main goal of the present work is to present a calculation of the PDF which is based on the ellipsoidal collapse model.

Our work is motivated by the fact that the spherical collapse model is inconsistent with a linear theory analysis of the second moment of the PDF by Kaiser (1987). However, Kaiser showed that the ratio of the redshift-space variance to that in real space should be $(1+2f/3+f^2/5)$ (here $f = d\ln D/d\ln a$ with $D$ the linear theory growth factor and $a$ the expansion factor), a finding which was confirmed by Fisher (1995) using a very different approach; the spherical model predicts that this factor is $(1+f/3)^2$ (Scherrer & Gaztangaga 2001). An analysis based on the Zeldovich approximation does yield Kaiser’s factor...
2 The redshift-space PDF

In what follows, the variance of the initial density fluctuation field when smoothed on scale \( R_M \) plays an important role. It is denoted by

\[
\sigma^2_L(M) \equiv \int \frac{dk}{k} \frac{k^3 P_s(k)}{2\pi^2} |W(k R_M)|^2, \tag{1}
\]

where \( P_s(k) \) is the power spectrum of the initial field, extrapolated using linear theory to the present time, \( W(x) \) is the Fourier transform of the smoothing window function and \( R_M = (3M/4\pi \bar{\rho})^{1/3} \). In addition, we use \( D(t) \) to denote the linear growth factor and \( f \equiv d \ln D/d \ln a \approx \Omega^{5/6} \).

### 2.1 The non-linear density in real space

The non-linear overdensity of a region of volume \( V \) containing mass \( M \) is

\[
\rho_s = 1 + \delta = \frac{M}{\bar{\rho} V}, \tag{2}
\]

where \( \bar{\rho} \) is the mean density. If the evolved Eulerian region is a sphere, then the evolved overdensity in real space is well-approximated by

\[
\rho_s = \frac{(1 - \delta_s/3)^3}{(1 - \delta_s/\delta)} \left( \frac{3}{1 - \lambda_j} \right)^{-1} \tag{3}
\]

(see Appendix A). Here, \( \delta_s \) is the critical value of spherical collapse model and its exact value depends weakly on the background cosmology. In a \( \Lambda \)CDM cosmology \( \delta_s \approx 1.66 \), while in an Einstein–de Sitter model \( \delta_s \approx 1.686 \). The \( \lambda_j \)'s are the eigenvalues of the deformation tensor, extrapolated using linear theory to the present time, and

\[
\delta_i = \sum_{j=1}^3 \lambda_j \text{ is the linear theory overdensity.}
\]

This approximation has considerable intuitive appeal: the factor in the product sign is the Zeldovich approximation to the evolution, and the factor in front is the ratio of the evolution of a sphere in the Zeldovich approximation to that in the spherical collapse model. The expression actually uses a simple approximation to the evolution of the density in the spherical model; in principle, one could use the exact evolution instead.

### 2.2 The non-linear density in redshift space

The redshift-space overdensity \( \rho_s \) is related to the real-space one by the mapping from real space to redshift space

\[
s = x + v \cdot e, \tag{4}
\]

where \( s \) is the redshift-space coordinate, \( x \) is the real-space coordinate, \( v \) is the peculiar velocity at \( x \), \( H \) is the Hubble parameter and \( e \) is the line-of-sight direction unit vector. Then,

\[
\rho_s = \rho_i \left| 1 + \frac{1}{H} \frac{\partial v_3}{\partial x_3} \right|^{-1}, \tag{5}
\]

where we assume the line-of-sight direction is along the third axis. The velocity of the ellipsoidal region along its principal axis can be computed using equation (A10),

\[
\frac{dR_i}{dt} \equiv 1 - f \left\{ R_i^3 \lambda_k - A_i \delta_l \left[ 1 - (1 - \delta_i/\delta) \delta_l^{(3-1)/3} \right] / 3 \right\}, \tag{6}
\]

where \( f = d \ln D/d \ln a \) and \( D(t) \) is the linear growth factor. \( R_i^3 \) is the initial axis length and it is related to \( \lambda \) by equation (A9). For a randomly oriented ellipsoid, we label the Euler angle components \( (\epsilon_1, \epsilon_2, \epsilon_3) \) = (cos \( \psi \) sin \( \theta \), sin \( \psi \) sin \( \theta \), cos \( \theta \)) for rotation from the line-of-sight coordinate to the coordinate of the principal axes of the ellipsoid. As a result, the redshift-space overdensity is

\[
\rho_s = \frac{(1 - \delta_s/3)^3}{(1 - \delta_s/\delta)} \left( \frac{3}{1 - \lambda_j} \right) \times \left[ 1 - \sum_{k=1}^{\infty} f \left\{ R_k^3 \lambda_k - A_k \delta_l \left[ 1 - (1 - \delta_i/\delta) \delta_l^{(3-1)/3} \right] / 3 \right\} \right]^{-1}. \tag{7}
\]

Note that \( \rho_s \to \rho_i \) when \( f \to 0 \), as it should.

### 2.3 The PDF from ellipsoidal collapse

Our redshift-space perturbation theory calculation makes use of the same assumptions as in real space: there is a local mapping from the eigenvalues \( \lambda_i \) of the deformation tensor to the non-linear overdensity \( \rho_s \), and the smoothing scale associated with \( \rho_i \) in the initial field is the one containing the same mass with the correct shape. The most important difference in redshift-space calculation is that the mapping from \( \lambda_i \) to \( \rho_s \) is no longer deterministic: the orientation of the ellipsoid with respect to the line of sight introduces stochasticity that must be integrated over. As a result, the evolved (Eulerian) PDF at a fixed \( V \) is related to the initial (Lagrangian) PDF at fixed mass-scale \( M \) by

\[
\rho_s p(\rho_s|V) d\rho_s = \int d\mathbf{e} \, d\mathbf{\lambda} \, p(\mathbf{\lambda}|\sigma) \delta_D \left[ \rho_s = \rho_s'(\mathbf{\lambda}, e) \right]. \tag{8}
\]

where \( \rho_s \equiv M/M \) and \( \rho_s' \) is the redshift-space overdensity given in equation (7).

In practice, the joint distribution of \( p(\mathbf{\lambda} | \sigma) \) is a function of \( \lambda_i/\sigma \) (Doroshkevich 1970), where

\[
\sigma = \sigma_L^{\text{up}}(M) \exp \left\{ - B \sum_{k<j} \left[ \ln \left( \frac{R_i^3}{R_k^3} \right) \right]^2 \right\}. \tag{9}
\]

Here, \( R_i^3 \) denotes the initial size of the \( j \)th principal axis, and \( B = 0.0486 \) (Betancort-Rijo & López-Corredoira 2002). As a result, it is useful to rewrite equation (8) as

\[
\rho_s p(\rho_s|V) d\rho_s = \int d\mathbf{e} \, d\mathbf{\lambda} \, p(\mathbf{\lambda}|1) \delta_D \left[ \rho_s = \rho_s'(\mathbf{\lambda}, e) \right]. \tag{10}
\]
As happens for the real-space calculation (Kofman et al. 1994; Scherrer & Gaztañaga 2001; Ohta et al. 2004), this estimate of the redshift-space PDF is not correctly normalized:

$$N \equiv \int_0^\infty d\rho_\nu \rho_\nu p(\rho_\nu) \neq 1.$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} (11)

(On the other hand, the integral of equation 8 over all $\rho_\nu$ does equal unity.) Therefore, we set

$$\rho^*_\nu \equiv N \rho_\nu \quad \mathrm{and} \quad \rho^*_\nu p^*(\rho^*_\nu) = \rho^2 p(\rho_\nu).$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} (12)

This ensures that $\int d\rho_\nu \rho^*_\nu p^*(\rho^*_\nu)$ and $\int d\rho_\nu \rho^2 p(\rho_\nu)$ both equal unity (Lam & Sheth 2008). In practice, for the scales of interest in what follows (cells of radius 4, 8 and 16 h$^{-1}$ Mpc), $N$ is only slightly larger than unity: 1.264, 1.104 and 1.052.

2.4 Other approximations

We will compare our model for the PDF with two other useful approximations which are less physically motivated. One is the lognormal,

$$\rho_\nu p(\rho_\nu) = \exp \left[-(\ln \rho_\nu + \mu_\nu)^2/2\sigma^2_\nu\right]/\sqrt{2\pi}\sigma_\nu,$$

where $\mu_\nu = \sigma^2_\nu/2$ and $\sigma^2_\nu = \ln(\rho^2_\nu)$ and $\sigma^2_\nu$ is given by multiplying the Kaiser factor by the Smith et al. (2003) fitting formula for the real-space variance $\sigma^2_\nu$.

The second is based on the finding that in simulations the distribution of $(\delta_1/\sigma_\nu)$ is almost the same as that of $(\delta_1/\sigma_\nu)$ (Scherrer & Gaztañaga 2001). Since the spherical model provides a good description of the $(\delta_1/\sigma_\nu)$ in real space (Lam & Sheth 2008), we use it together with the Kaiser factor to set

$$p(\rho_\nu) = \frac{\sigma_\nu}{\sigma_\nu} p_\mathrm{sw}(\rho_\nu),$$

where $p_\mathrm{sw}(\rho)$ is the real-space PDF with the spherical collapse model. In essence, this can be thought of as a model in which the redshift-space overdensity is a linearly biased version of the real-space overdensity, the constant of proportionality (the bias factor) being $(\sigma_\nu/\sigma_r)$. Kaiser (1987) suggests that this should be an excellent approximation on large scales; although Scoccimarro (2004) suggests that it is not a well-motivated approximation on intermediate or small scales, measurements of the redshift-space PDF by Scherrer & Gaztañaga (2001) suggest that it is nevertheless reasonably accurate.

3 COMPARISON WITH SIMULATION

We compare our models with the measurements in a simulation which was run and analysed by Smith, Sheth & Scoccimarro (2007). The simulation box is a periodic cube 512 $h^{-1}$ Mpc on a side in a cosmological model where $(\Omega_m, \Omega_\Lambda, h, \sigma_8) = (0.27, 0.73, 0.72, 0.9)$. The simulation followed the evolution of 400$^3$ particles, each of mass $1.57 \times 10^{11} h^{-1} M_\odot$. Haloes were identified using the FOF algorithm with link length 0.2 times the mean interparticle separation.

To measure the Eulerian real-space PDFs, we place spheres of volume $V$ randomly in the simulation box, and record the number of particles in each sphere, taking care to account for periodic boundary conditions. We measure the redshift-space PDFs similarly, after transforming the particle positions as follows. We assume that one of the sides of the box lies parallel to the line of sight. The redshift-space coordinate is then given by adding the peculiar velocity (divided by the Hubble constant) in this direction to the real-space coordinate, taking care to account for the periodicity of the box.

Technically, our model of the PDF is for a continuous density field rather than for discrete particle counts. In practice, we are interested in large enough volumes that discreteness effects are negligible, so we simply set $\rho \equiv 1 + \delta = N/N$. However, there is another effect which is more pernicious. The ‘virial’ motions of particles within haloes give rise to long fingers-of-god – this makes haloes stretch about seven times longer along the line of sight than across – an effect first noted by Jackson (1972), which is entirely absent in the perturbation theory calculation of the redshift-space PDF. To illustrate their effect on the PDF, we have also computed redshift-space positions after first setting all virial motions to zero, so that dark matter particles have the same speed as their associated haloes. (That is, we assign all particles in a halo the same velocity as that of the halo centre of mass.) In the figures that follow, filled black squares show the true redshift-space measurements, and open orange squares show results when fingers-of-god are absent.

Fig. 1 shows the redshift-space PDF of the dark matter in spheres of radius 8 $h^{-1}$ Mpc. The main panel shows this using a log scale to emphasize the behaviour of the low-probability tails, and the inset shows the same data using linear axes to highlight the behaviour near the peak of the distribution. Filled and open squares show the measurements in the simulation (i.e. before and after compressing the fingers-of-god). These indicate that fingers-of-god are most important in the denser cells; this is not unexpected, since dense cells host the most massive haloes which have the largest virial motions. Our perturbation theory based prediction is represented by the dotted (blue) curve – note that it provides a significantly better fit to the simulations when the fingers-of-god have been compressed.
The solid (magenta) curve is the perturbation theory based prediction with $B = 0$ in equation (9), which corresponds to using our ellipsoidal collapse model but ignoring the effect on the variance associated with differences from a spherical shape. The dotted curve ($B = 0.0486$) is always closer to the measurements than is the solid one ($B = 0$), consistent with the real-space results presented in Lam & Sheth (2008), although the difference between the two curves is not large. In what follows, we only show results for the case when $B = 0.0486$. The dot–dashed (green) curve shows the lognormal (equation 13) and the long-dashed (red) curve shows equation (14).

Our model provides a better fit than these other more empirical curves, except in the highest density tail.

Fig. 2 shows similar results for spheres of radius 4 and 16 $h^{-1}$ Mpc. In the larger cells, fingers-of-god are irrelevant, and the lognormal provides a slightly better description than does our model, which again overshoots at large densities. The approach which assumes a simple mapping from real- to redshift-space overdensity fairs reasonably well at large densities, but underestimates the height of the peak of the distribution, and overestimates the PDF at lower densities. Our model is substantially better than the other two in the smaller 4 $h^{-1}$ Mpc cells.

4 RECONSTRUCTION OF THE INITIAL DISTRIBUTIONS

In Lam & Sheth (2008), we described how the spherical collapse model could be used as a basis for a method of reconstructing the shape of the initial PDF from the Eulerian one. In principle, the method can be used as a consistency check for Gaussian initial conditions. We showed that, in real space, it can reconstruct the Gaussian form of the initial distribution from the highly skewed non-linear PDF on scales where the rms fluctuation is about 2 or less.

Extending this reconstruction method to redshift space is complicated by the fact that we know the spherical model is inadequate (it does not reproduce the Kaiser factor). However, the ellipsoidal collapse model, which we have shown works quite well, has many

more parameters – the lengths and orientations of the three principal axes – so that the final density is determined by a stochastic rather than deterministic mapping.

So, we have tried a simpler approach which is based on the fact that equation (14) provides a reasonable description of the redshift-space PDF. Namely, we assume that

$$\delta_s = \delta_r (\sigma_s/\sigma_r),$$  

and use the (square root of the) Kaiser factor in place of $\sigma_s/\sigma_r$; we have found that using our ellipsoidal collapse calculation of this quantity instead produces slightly worse results. Then, we set

$$\nu \equiv \frac{1 - [(1 + \delta_s \sigma_s/\sigma_r)/N_c]^{-1/k}}{\sigma_r (M/N_c)/\delta_c},$$  

where $\delta_c$ is the critical linear density associated with collapse in the spherical model and $N_c$ is the normalization factor (the analogue of equation 11) in the spherical model calculation of the real-space PDF. Finally, we make a histogram of $\nu$ by weighting each cell which contributes to $p(\delta_s)$ by its value of $(1 + \delta_s \sigma_s/\sigma_r)/N_c$.

Before showing how well this works, Fig. 3 shows a test of our assumption that equation (15) is accurate, meaning that $\delta_s$ is linearly related to $\delta_r$, with little scatter around the mean relation. Evidently, these assumptions are reasonable on large scales (bottom and middle panels), but become increasingly worse on smaller scales (top).

Fig. 4 shows the reconstruction results for 4, 8 and 16 $h^{-1}$ Mpc spheres. The left- and right-hand panels show the result of transforming the PDFs shown by filled and open squares in the previous figures (i.e. full and compressed fingers-of-god). Smooth solid curves show a Gaussian distribution for comparison. The reconstruction method works well for the 8 and 16 $h^{-1}$ Mpc scales. This method even works for the smaller 4 $h^{-1}$ Mpc scales if the fingers-of-god have been removed. The dotted curves show the corresponding non-linear distributions (equation 8) to demonstrate that our reconstruction method transforms significantly skewed distributions back to the original symmetric Gaussian shape.
Comparison of real- and redshift-space overdensities in cells of \( \delta \) Mpc (top), 8 h^{-1} Mpc (middle) and 16 h^{-1} Mpc (bottom). Filled and open symbols are before and after fingers-of-god have been removed. They show the mean, and error bars show the rms, for narrow bins in \( \delta/\sigma_r \). If equation (15) is accurate, then \( \langle \delta/\sigma_r \rangle = 0 \) with little scatter.

Figure 3. Comparison of real- and redshift-space overdensities in cells of radius 4 h^{-1} Mpc (top), 8 h^{-1} Mpc (middle) and 16 h^{-1} Mpc (bottom). Filled and open symbols are before and after fingers-of-god have been removed. They show the mean, and error bars show the rms, for narrow bins in \( \delta/\sigma_r \). If equation (15) is accurate, then \( \langle \delta/\sigma_r \rangle = 0 \) with little scatter.

5 DISCUSSION

We have extended our previous work on the real-space dark matter PDF to redshift space. To do so, we provided a simple parametrization of the evolution of the non-linear density in the ellipsoidal collapse model. This explicitly shows how the model extends both the spherical collapse model and the Zeldovich approximation (equations (A10) and 7). This parametrization, with our prescription for relating spatial statistics in the initial and evolved fields (equation 8), results in an excellent agreement with measurements from a numerical simulation, in both real (Fig. A1) and redshift space (Figs 1 and 2). Our parametrization fares significantly better than analyses which are based on exact second- and third-order perturbative expansions of the ellipsoidal collapse evolution – these produce PDFs with a power-law tail at high densities which is a signature that the analysis is breaking down – a similar tail was seen for the Zeldovich approximation (Hui et al. 2000). Our approach also fares better than the lognormal model, and one which assumes that redshift-space overdensities are linearly proportional to those in real-space: \( \delta = (\sigma/\sigma_i) \delta_r \) where the constant of proportionality is given by the (square root of the) ratio of the variances. The latter is known to be a good approximation on large scales – we find that it is accurate on small scales, except in underdense regions.

We also used our results to motivate a method for reconstructing the shape of the initial PDF from the non-linear evolved redshift-space PDF. This method works well on scales where the rms fluctuation is less than 2 when fingers-of-god have been removed (Fig. 4). We are in the process of applying this method to reconstruct the baryonic acoustic oscillation (BAO) peak in the correlation function (Eisenstein et al. 2007). Instead of using equation (16), it is also possible to formulate the reconstruction using equations (A10) or (7); this is also the subject of the work in progress.

While this provides a nice graphic reconstruction of the Gaussianity of the initial conditions, more work is necessary before it can be used to distinguish between a truly Gaussian model and the very mildly non-Gaussian models which are not yet excluded by the CMB (Komatsu et al. 2008). In this respect, it may be interesting to extend our method for estimating the non-linear PDF to the family of non-Gaussian models which are currently studied in the context of using cluster abundances and clustering to constrain primordial non-Gaussianity (Dalal et al. 2008; Lo Verde et al. 2008; Slosar et al. 2008).

ACKNOWLEDGMENTS

We thank R. Scherrer for pointing out that there was little understanding of why, when scaled by their variances, the real and redshift-space PDFs were so similar, L. Hui for urging us to go beyond an order-by-order analysis, R. E. Smith for the simulation data and the referee, V. Desjacques, for a thoughtful report. This work was supported in part by NSF Grant AST-0507501.

REFERENCES

Barrow J. D., Silk J., 1981, ApJ, 250, 432
Bernardeau F., 1994, A&A, 291, 697
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Betancort-Rijo J., 1991, MNRAS, 251, 399
Betancort-Rijo J., Lópe-Corredor M., 2002, ApJ, 566, 623
Bond J. R., Myers S. T., 1996, ApJS, 103, 1
Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, A&A, 296, 575
Dalal N., White M., Bond J. R., Shirokov A., 2008, preprint (arXiv:0803.3453)
Desjacques V., 2008, MNRAS, 388, 638
Desjacques V., Smith R. E., 2008, preprint (arXiv:0805.2145)
Doroshkevich A. G., 1970, Astrofizika, 6, 581
APPENDIX A: RELATION BETWEEN INITIAL AND FINAL DENSITY IN THE ELLIPSOIDAL COLLAPSE MODEL

In this Appendix, we derive relatively simple analytic approximations to the evolution of the density within a triaxial perturbation. The exact evolution of the perturbation is assumed to be described by the ellipsoidal collapse model of Bond & Myers (1996). Shen et al. (2006) provide a simple but useful approximation to this evolution which is based on a beautiful analysis of the collapsing ellipsoid problem by White & Silk (1979). It is this approximate solution that we exploit.

A1 The collapse model

Gravity evolves the initial axes \( (R_1, R_2, R_3) \) of a triaxial object to \( (R_1, R_2, R_3) \) because

\[
\frac{d^2 R_i}{dt^2} = H_o^2 \Omega_i R_i - 4\pi G \bar{\rho} R_i \left( \frac{1 + \delta \lambda_i}{3} + \frac{b'_i}{2} \delta + \lambda_{\text{ext}, i} \right),
\]

where the interior and exterior tidal forces are, respectively,

\[
b'_i(t) = \left( \prod_{j=1}^{3} R_j(t) \right) \int_0^\infty \frac{dr}{R_i^2 + \tau} \prod_{j=1}^{3} \left[ R_j^2 + \tau \right]^{1/2},
\]

\[
\lambda_{\text{ext}, i}(t) = \frac{D(t)}{D(t)} \left[ \lambda_i(t) - \frac{\delta_i}{3} \right],
\]

(Bond & Myers 1996). Here, \( \lambda_i \) are the initial eigenvalues of the deformation tensor and the initial conditions of equation (A1) are set by the Zeldovich approximation:

\[
R_i(t) = a_i[1 - \lambda_i(t)],
\]

where \( f_i = d \ln D / d \ln a \) and \( D(t) \) is the linear growth factor.

Although equation (A1) in general must be solved numerically, Shen et al. (2006) showed that the following approximation was rather accurate:

\[
R_i(t) = \frac{a(t)}{a_i} R_i(t_i) \left[ 1 - D(t) \lambda_i(t) \right] - \frac{a(t)}{a_i} R_i(t_i) \left[ 1 - \frac{D(t)}{3} \delta_i - \frac{a(t)}{a(t_i)} \right],
\]

where \( R_i(t_i) = 3/\sum_j R_j(t_i)^{-1} \) and \( a(t_i) \) is the expansion factor of a spherical universe with initial overdensity \( \delta_i = \sum_j \lambda_j(t_i) \). This expression is inspired by an analysis of collapsing ellipsoids by White & Silk (1979). The first term in the expression above is the Zeldovich approximation to the evolution and the second term is a correction to this approximation. When \( \lambda_2 = \delta_i/3 \), then the evolution of the second axis is almost exactly like that of a sphere with initial overdensity \( \delta_i \).
A2 Evolution to a spherical volume

For an Eulerian sphere of $R_E$, equation (A6) implies that

$$R_i^k[1 - D(t)\lambda_1] = R_i^k[1 - D(t)\lambda_2] = R_i^k[1 - D(t)\lambda_3].$$  \hspace{1cm} (A7)

Substitute into the definition of $R_k(t)$,

$$R_k(t) = \frac{3}{\sum_j R_j(t)^{-1}} = \frac{R_k(t)[1 - D(t)\lambda_k]}{1 - D(t)\delta_i/3},$$  \hspace{1cm} (A8)

where $k = 1, 2, 3$. Therefore, we can write the evolved $R_k(t)$ as

$$\frac{R_E}{R_k} = \frac{a_i(t)/a(t)}{1 - D(t)\delta_i/3}[1 - D(t)\lambda_k],$$  \hspace{1cm} (A9)

where $R_k = R_k(t_i)$. This writes the evolution as that in the Zeldovich approximation, times a correction factor which is simply the ratio of the evolved and initial radii of a Zeldovich sphere $(1 - D\delta_i/3)$ divided by the same ratio in the exact spherical evolution model. Thus, the evolved overdensity is

$$\rho = 1 + \delta = \prod_{j=1}^{3} \frac{R_j}{R_E} \approx \frac{(1 - D(t)\delta_i/3)^3}{(a_i(t)/a(t))^3} \prod_{j=1}^{3} \frac{1}{1 - D(t)\lambda_j}$$

$$\approx \frac{(1 - D(t)\delta_i/3)^3}{(1 - D(t)\delta_j/\delta_c)^3} \prod_{j=1}^{3} \frac{1}{1 - D(t)\lambda_j}.$$  \hspace{1cm} (A10)

The final expression uses a simple approximation for the evolution in the spherical collapse model:

$$\frac{a(t)}{a_i(t)} \approx \left(1 - D(t)\frac{\delta_i}{\delta_c}\right)^{-\delta_i/3}$$  \hspace{1cm} (A11)

(Bernardeau 1994; Sheth 1998), where $\delta_c \approx 5/3$. Note that when all the eigenvalues are identical, then equation (A10) reduces to the spherical collapse model.

Insertion of equation (A11) or (A10) into equation (8) of the main text yields models for the redshift-space PDF associated with spherical and ellipsoidal collapse, respectively. The corresponding models for the real-space PDF are given by setting $f = 0$ in that expression. These real-space models are shown by the dashed (red) or solid (blue) curves; both are in excellent agreement with the measured real-space PDF (symbols).
A3 Evolution to an arbitrary shape

When the Eulerian volume in consideration is triaxial, the computation of the overdensity is less trivial. In general, one must solve a system of non-linear equations for \( x_k \equiv a(t)R_j^3/a_jR_k(t) \) in

\[
x_k[1 - D(t)\lambda_k] = 1 + 3 \left( 1 - \frac{D(t)\delta_i}{3} - \frac{a_i(t)}{a(t)} \right) \left( \sum_{j=1}^{3} \frac{R_k}{R_j} x_j \right)^{-1}.
\]

The overdensity is \( 1 + \delta = \prod_{j=1}^{3} x_j \). The above equation can be solved numerically or iteratively with

\[
x_k^{(n+1)} = \left[ 1 + 3 \left( 1 - \frac{D(t)\delta_i}{3} - \frac{a_i(t)}{a(t)} \right) \left( \sum_{j=1}^{3} \frac{R_k}{R_j} x_j^{(n)} \right)^{-1} \right]^{-1} \prod_{j=1}^{3} x_j^{(n)}.
\]

An analytical expression for the overdensity of the ellipsoid can be obtained if we make the approximation that the collapse of the second axis is very close to that in the spherical model. Shen et al. (2006) show why this is almost always an excellent approximation. One can then approximate the collapse of the second axis using either of the two expressions below:

\[
\frac{R^2_2}{R^2(t)} \approx \left\{ \begin{array}{l} \frac{1 - D(t)\delta_i/3}{a_i(t)/a(t)} \frac{1}{1 - D(t)\lambda_2}, \\ 1 - D(t)\lambda_2. \end{array} \right. \]

The first expression is our approximate description of evolution to an Eulerian sphere and the second is simply the spherical collapse model. Substituting these approximations into equation (A6) yields

\[
\frac{R^2_2}{R^2(t)}[1 - D(t)\lambda_k] = \prod_{j=1}^{3} \frac{R_j}{R_j(t)} \approx \left\{ \begin{array}{l} \frac{a_i(t)/a(t)}{1 - D(t)\lambda_2}, \\ \frac{R_j}{R_j(t)} - 1 + \frac{1 - D(t)\delta_i/3}{a_i(t)/a(t)} \frac{1}{1 - D(t)\lambda_2}. \end{array} \right.
\]

for which the associated overdensities are

\[
\rho = \prod_{j=1}^{3} \frac{R_j}{R_j(t)} \approx \left\{ \begin{array}{l} \prod_{j=1}^{3} \frac{a_i(t)}{a(t)} \frac{1 - D(t)\delta_i/3}{1 - D(t)\lambda_2}, \\ \prod_{j=1}^{3} \frac{a_i(t)}{a(t)} \frac{1 - D(t)\delta_i/3}{1 - D(t)\lambda_2} \left( \frac{R_j}{R_j(t)} + \frac{a_i(t)}{a(t)} \frac{1 - D(t)\delta_i/3}{1 - D(t)\lambda_2} \right). \end{array} \right.
\]

where \( a_i(t) \) is approximated by equation (A11). Compared to equation (A10), the last term in the first approximation is the extra factor accounting for the non-spherical shape in the final volume. For \( R_1 = R_2 = R_3 \), this extra factor is unity and equation (A17) has the same expression as in the spherical case.

APPENDIX B: PERTURBATIVE TREATMENT OF REDSHIFT-SPACE DISTORTIONS

Kaiser (1987) derives a simple relation between the real- and redshift-space power spectra. Fisher (1995) provides a very different derivation of Kaiser’s expression. The Zeldovich approximation provides yet another derivation of this relation (Ohta et al. 2004). We show that our approximation for ellipsoidal collapse also reproduces this expression, and then use it to derive the next-order corrections. This analysis is useful because Scoccimarro (2004) has shown that Kaiser’s approximation is accurate only on rather large scales.

Before we present the algebra, it is worth noting that the approximation used in the main text for the real-space evolution of a collapsing ellipsoid follows from (i) writing the full ellipsoidal collapse as the Zeldovich approximation times a correction factor which is the ratio of the evolution of a sphere in the Zeldovich approximation to that in the spherical collapse model and (ii) using our convenient approximation to the exact evolution in the spherical collapse model. In the perturbative analysis that follows, we have chosen notation which illustrates what happens if we keep approximation (i), but treat the spherical collapse model exactly rather than approximately.

We begin by writing the relation between the linear and non-linear overdensities in the spherical evolution model as

\[
1 + \delta = \left( \frac{a(t)}{a_i(t)} \right)^3 = 1 + \sum_{n=1}^{\infty} \frac{v_n}{n!} \delta^n.
\]

Then, the exact result has

\[
\begin{align*}
v_2 &= \frac{2}{3} \left( 2 - \frac{g_2}{g_1^2} \right), \\
v_3 &= \frac{2}{9} \left( 10 - 12 \frac{g_2}{g_1^2} + \frac{g_{10}}{g_1} + \frac{g_{12}}{g_1^2} \right).
\end{align*}
\]
where $g_i$ is the $i$th-order growth factor (e.g. Fosalba & Gaztañaga 1998). For $\Lambda \neq 0 \ (g_2/g_1^2) \approx -(3/7) \Omega^{-1/143}, (g_3/g_1^3) \approx -(1/3) \Omega^{-4/727}$ and $(g_4/g_1^4) \approx (10/21) \Omega^{-269/17875}$ (Bouchet et al. 1995). The dependence on $\Omega$ is so weak that we can use the $\Omega = 1$ values to find $v_2 \approx 1.62$ and $v_3 \approx 3.926$. In contrast, our previous approximation has

\[
v_\nu \approx \prod_{i=0}^{n-1} \frac{1}{\delta_i} + i
\]

so $v_2 \approx 1.6$ and $v_3 \approx 3.52$.

With this notation, our approximation for the evolution of an ellipsoid is

\[
1 + \delta_i = \left(1 - \frac{\delta_i}{3}\right)^3 \left(1 + \sum_{n=1}^{\infty} \frac{v_\nu}{n!} \delta_i^n \right) \prod_{k=1}^{3} \left(1 - \lambda_k\right)^{-1}
= 1 + \delta_i^{(1)} + \delta_i^{(2)} + \delta_i^{(3)} + \ldots,
\]

(B4)

where

\[
\delta_i^{(1)} = \sum_{j=1}^{3} \lambda_j
\]

\[
\delta_i^{(2)} = \frac{v_2}{2} \delta_i^2 + \frac{\delta_i^3}{3} - \sum_{j \neq k} \lambda_j \lambda_k
\]

\[
\delta_i^{(3)} = \frac{v_3}{6} \delta_i^3 + \frac{17}{27} \delta_i^3 - 2\delta_i \sum_{j \neq k} \lambda_j \lambda_k + \lambda_1 \lambda_2 \lambda_3.
\]

(B5)

In each expression above, the first terms are those associated with the spherical model, and the others are the corrections which come from the ellipsoidal collapse. The terms associated with a perturbation theory analysis of the exact ellipsoidal collapse model are given by Ohta et al. (2004). The above expressions show the terms that are associated with our simple approximation to the exact evolution. They differ slightly from those associated with the exact analysis; e.g. the second-order expression associated with the exact analysis is $(4/17)(\delta_i^2 - 3f_i)$, but they otherwise have the same form, suggesting that our simple expression captures the essence of the ellipsoidal collapse evolution.

The overdensity in redshift space is

\[
1 + \delta_i = \frac{1 + \delta_i}{1 - \sum_{k=1}^{3} \tilde{\delta}_k e_k^2}
= 1 + \tilde{\delta}_i^{(1)} + \tilde{\delta}_i^{(2)} + \tilde{\delta}_i^{(3)} + \ldots,
\]

(B6)

where

\[
\tilde{\delta}_i = \left\{ R_i \lambda_i - A_i \left[ f_i \delta_i - (1 + \sum_{n=1}^{\infty} \frac{1}{n!} \delta_i^n) \right] \right\}^{-1/3} / \delta_i^{1/3}
= f_i \lambda_i + \frac{1}{3} \left( \frac{v_2}{2} f_2 - \frac{4}{3} f_1 \right) \delta_i^2 + f_i \lambda_i^2
\]

\[
+ \frac{\delta_i^2}{3} \left[ \left( \frac{v_2}{2} f_2 - \frac{4}{3} f_1 \right) \lambda_i + \frac{v_3}{6} f_3 - \frac{2v_2}{3} f_1 - \frac{5v_3}{6} f_2 + 2f_1 \right] \delta_i
+ \frac{1}{3} \left( \frac{v_2}{2} f_2 - \frac{4}{3} f_1 \right) \lambda_i \delta_i^2 + f_i \lambda_i \left[ \frac{\lambda_i^2 + \delta_i^2}{2} \left( \frac{v_2}{6} - \frac{2}{9} \right) \right] + \ldots,
\]

(B7)

and $f_i = d \ln g_i / da$. Thus,

\[
\delta_i^{(1)} = \delta_i^{(1)} + \Delta_i^{(1)}
\]

\[
\delta_i^{(2)} = \delta_i^{(2)} + \Delta_i^{(2)} + \delta_i^{(1)} \Delta_i^{(1)}
\]

\[
\delta_i^{(3)} = \delta_i^{(3)} + \Delta_i^{(3)} + \delta_i^{(2)} \Delta_i^{(2)} + \delta_i^{(1)} \Delta_i^{(1)}
\]

(B8)
Figure B1. Four approximations to the redshift-space PDF. Solid line shows the prediction associated with the Zeldovich approximation, but assuming that the associated $\sigma$ is the same for all cells of volume $V$ (Hui et al. 2000). Dashed and dot–dashed curves show the first- and third-order exact ellipsoidal collapse based analyses; these do account for the fact that $\sigma$ depends on the Lagrangian size and shape (e.g. Ohta et al. 2004). Dotted curve shows the PDF associated with our approximation to the exact evolution.

and

\[ \Delta^{(1)}_i = f_1 \sum_{k=1}^{3} \lambda_k \, \epsilon_k^2 \]

\[ \Delta^{(2)}_i = f_1 \sum_{k=1}^{3} \left[ \frac{v_2}{f_1} \cdot \frac{f_2}{f_1} - \frac{4}{3} \right] \delta_k^2 \, \epsilon_k^2 \]

\[ + f_1^2 \sum_{k=1}^{3} \lambda_k^2 \, \epsilon_k^2 + f_1^2 \sum_{j,k=1}^{3} \lambda_j \lambda_k \, \epsilon_j^2 \epsilon_k^2 \]

\[ \Delta^{(3)}_i = f_1 \sum_{k=1}^{3} \left[ \left( \frac{v_2}{f_1} \cdot \frac{f_2}{f_1} - \frac{4}{3} \right) \lambda_k + \left( \frac{v_3}{f_1} \cdot \frac{f_3}{f_1} - \frac{2v_2}{f_1} - \frac{5v_2}{6} \right) \lambda_k \right] \delta_i^2 \, \epsilon_k^2 \]

\[ + 2f_1^2 \sum_{j,k=1}^{3} \lambda_j \left( \left( \frac{v_2}{f_1} \cdot \frac{f_2}{f_1} - \frac{4}{3} \right) \delta_j^2 \right) \epsilon_j^2 \]

\[ + f_1^3 \sum_{i,j,k=1}^{3} \lambda_i \lambda_j \lambda_k \, \epsilon_i^2 \epsilon_j^2 \epsilon_k^2. \]  

(B9)

Our leading-order term, $\delta^{(1)}$, is the same as that of Ohta et al. (2004), who showed that $\sigma_s^2 \equiv \langle (\delta^{(1)}_s)^2 \rangle$, where the angle brackets denote the result of averaging over the Euler angles as well as the distribution of the $\lambda_i$, leads to Kaiser’s formula. The logic that leads to our second-order term is consistent with that of Watts & Taylor (2001), except that because our collapse model considers motions relative to the centre of the object, it has no term which accounts for the motion of the centre of mass. There are additional small differences which arise from the fact that our description of ellipsoidal collapse is based on our approximate model – had we computed our perturbation series expansion using the exact collapse model, we would have reproduced their expressions.

Fig. B1 illustrates one reason why we have used our approximation for the non-linear dynamics, rather than worked with exact perturbative expansions to higher and higher orders. The dotted line shows the redshift-space PDF associated with our approximation; the main text shows that it provides a good description of the simulations. The solid line shows the result of using the Zeldovich approximation for the evolution, and setting $\sigma = \sigma(\bar{\rho} V)$ whatever the shape and overdensity of the initial object (Hui et al. 2000). This results in a PDF with a tail that scales as $\rho^{-3}$ at high densities – a signature that the approach has broken down. The dashed curve shows the result of using equation (9) for $\sigma$ instead. This helps somewhat – although the onset of the $\rho^{-3}$ tail is shifted to larger $\rho_s$, the shape of the PDF in underdense regions is adversely affected. The dot–dashed curve shows the result of going to third-order in the exact ellipsoidal collapse model (recall that, in this context, the Zeldovich approximation is like the first-order term). This modifies the shape of the PDF slightly, but the problem at small $\rho_s$ remains.