An accurate linear model for redshift space distortions in the void-galaxy correlation function

Seshadri Nadathur* and Will J. Percival

Institute of Cosmology and Gravitation, University of Portsmouth, Burnaby Road, Portsmouth PO1 3FX, UK

ABSTRACT

Redshift space distortions within voids provide a unique method to test for environmental dependence of the growth rate of structures in low density regions, where effects of modified gravity theories might be important. We derive a linear theory model for the redshift space void-galaxy correlation that is valid at all pair separations, including deep within the void, and use this to obtain expressions for the monopole $\xi_0$ and quadrupole $\xi_2$ contributions. Our derivation highlights terms that have previously been neglected but are important within the void interior. As a result our model differs from previous works and predicts new physical effects, including a change in the sign of the quadrupole term within the void radius. We show how the model can be generalised to include a velocity dispersion. We compare our model predictions to measurements of the correlation function using mock void and galaxy catalogues modelled after the BOSS CMASS galaxy sample using the Big MultiDark $N$-body simulation, and show that the linear model with dispersion provides an excellent fit to the data at all scales, $0 \leq s \leq 120 \, h^{-1}\text{Mpc}$. While the RSD model matches simulations, the linear bias approximation does not hold within voids, and care is needed in fitting for the growth rate. We show that fits to the redshift space correlation recover the growth rate $f(z = 0.52)$ to a precision of 2.7% using the simulation volume of $(2.5 \, h^{-1}\text{Gpc})^3$.

Key words: gravitation – large-scale structure of Universe –cosmology: observations – methods: data analysis – methods: analytical

1 INTRODUCTION

Galaxy redshift surveys provide a map of the large-scale structure of the Universe containing anisotropic distortions of the clustering caused by gravitationally-induced peculiar velocities that contribute to the galaxy redshifts, as first predicted by Kaiser (1987). Measurement of these redshift space distortions (RSD) (e.g. Peacock et al. 2001; Guzzo et al. 2008; Beutler et al. 2012; Samushia et al. 2012; Reid et al. 2012; Howlett et al. 2015; Beutler et al. 2017) can be used to determine the growth rate of structure, which can provide strong tests of General Relativity (GR). The theory of RSD in the galaxy clustering is however complicated by significant non-linear contributions which are important even at quite large pair separation scales, requiring sophisticated modelling (e.g. Scoccimarro 2004; Matsubara 2008; Taruya et al. 2010; Reid & White 2011; Jennings et al. 2011).

The RSD modelling could in principle be simplified in underdense regions where the dynamics is closer to linear, such as voids. Cosmic voids are large underdensities in the galaxy distribution which can be used to trace stationary points of the gravitational potential (Nadathur et al. 2017), where velocities are dominated by coherent bulk flows. Voids have been much studied recently in other contexts, including their action as weak gravitational lenses (Krause et al. 2013; Melchior et al. 2014; Clampitt & Jain 2015; Sánchez et al. 2016), the secondary CMB anisotropies they generate through the integrated Sachs-Wolfe effect (e.g. Granett et al. 2008; Hotchkiss et al. 2015; Nadathur & Crittenden 2016; Cai et al. 2017; Kovács et al. 2017) and the thermal Sunyaev-Zeldovich effect (Alonso et al. 2017), and the Baryon Acoustic Oscillation (BAO) peak in their clustering (Kitaura et al. 2016). Velocity dynamics around voids have been studied using RSD in the void-galaxy cross-correlation (Paz et al. 2013; Hamaus et al. 2015; Cai et al. 2016; Hamaus et al. 2017; Achitouv et al. 2017) and that of the void positions themselves (Chuang et al. 2017).

In addition to possible simplifications in the linear modelling, an important advantage of studying RSD around voids is that it presents an opportunity to study the growth of density perturbations specifically in low-density regions. Several popular alternatives to GR, such as chameleon $f(R)$ gravity (Hu & Sawicki 2007), Dvali-Gabadadze-Porrati (DGP) (Dvali et al. 2000), and Galileon (Nicolis et al. 2008) models, may be important in voids where the dynamics is expected to be closer to linear. A linear theory approximation for the RSD in the void-galaxy cross-correlation is thus a natural way to test for environmental dependence in low-density regions.

In this paper, we derive a linear theory model for the redshift space void-galaxy correlation that is valid at all pair separations, including deep within the void, and use this to obtain expressions for the monopole $\xi_0$ and quadrupole $\xi_2$ contributions. Our derivation highlights terms that have previously been neglected but are important within the void interior. As a result our model differs from previous works and predicts new physical effects, including a change in the sign of the quadrupole term within the void radius. We show how the model can be generalised to include a velocity dispersion. We compare our model predictions to measurements of the correlation function using mock void and galaxy catalogues modelled after the BOSS CMASS galaxy sample using the Big MultiDark $N$-body simulation, and show that the linear model with dispersion provides an excellent fit to the data at all scales, $0 \leq s \leq 120 \, h^{-1}\text{Mpc}$. While the RSD model matches simulations, the linear bias approximation does not hold within voids, and care is needed in fitting for the growth rate. We show that fits to the redshift space correlation recover the growth rate $f(z = 0.52)$ to a precision of 2.7% using the simulation volume of $(2.5 \, h^{-1}\text{Gpc})^3$.

Key words: gravitation – large-scale structure of Universe –cosmology: observations – methods: data analysis – methods: analytical

* seshadri.nadathur@port.ac.uk
2009; Deffayet et al. 2009) gravity models, among others, make use of screening mechanisms in order to suppress fifth force effects in high density regions such as the Solar System; theories with such screening effects therefore predict environment-dependent differences in the growth rate \( f \), which could be probed by the RSD effects within voids.

In this paper we study the void-galaxy cross-correlation function \( \xi_{vg} \) in redshift space. Cai et al. (2016) have previously studied the same problem and provided a linear model for \( \xi_{vg} \), which was subsequently also used by Hamiaux et al. (2017). However, this model correctly described simulation results only for void-galaxy pair separations greater than \( r > R_c \), outside the low-density region of greatest interest. By extending the model using a form of the quasi-linear streaming model, Cai et al. (2016) were able to extend the region of validity slightly to \( r > 0.5R_c \), though it was still not correct in the void centres. In fact in the void centre region the model can lead to unphysical predictions of \( \xi_{vg} < -1 \).

We revisit the derivation of the void-galaxy RSD model from first principles and identify linear-order terms that were neglected in the expression obtained by Cai et al. (2016) but are important in the void centre regions. These terms do not have counterparts in the linear Kaiser model for RSD in the galaxy correlation, and arise because of the restriction to void regions. They have important physical effects, in particular causing a change in sign of the quadrupole term within void regions. These terms become unimportant outside the void radius, where our expression matches that of Cai et al. (2016).

We show how the model can be extended to include a velocity dispersion without changing the linear nature of the theory. This is not the same as the form of the Gaussian streaming model used in several other studies of the void-galaxy correlation (e.g. Paz et al. 2013; Hamiaux et al. 2015; Achitouv et al. 2017; Achitouv 2017). We show why it is not correct to use the standard streaming model result for the galaxy correlation in the void-galaxy case.

We then compare our theoretical results to data from mock void and galaxy catalogues from a large \( N \)-body simulation and show that our linear dispersion model provides an excellent fit to the data at all scales and performs significantly better than alternatives. We highlight the fact that the approximation of a linear galaxy bias does not hold within voids. We discuss the consequences of these results for obtaining an unbiased estimate of the growth rate within void environments.

The layout of the paper is as follows. Section 2 describes the simulation data we use and the creation of galaxy and void mocks. In Section 3 we derive the linear theory model for the void-galaxy correlation and its multipoles, and discuss differences with previous works. In Section 4 we compare theoretical predictions to simulation data and in Section 5 we discuss strategies for measurement of the growth rate based on these results. We sum up and draw conclusions in Section 6.

2 DATA

2.1 Simulation and galaxy mocks

We use data from the \( z = 0.52 \) redshift snapshot from the Big MultiDark (BigMD) \( N \)-body simulation (Klypin et al. 2016) from the MultiDark simulation project (Prada et al. 2012). This simulation follows the evolution of 3840\(^3\) particles in a box of side \( L = 2500 \ h^{-1}\)Mpc using the GADGET-2 (Springel 2005) and Adaptive Refinement Tree (Kravtsov et al. 1997; Gottlober & Klypin 2008) codes, with cosmological parameters \( \Omega_M = 0.307, \Omega_B = 0.048, \Omega_{\Lambda} = 0.693, n_s = 0.95, \sigma_8 = 0.825 \) and \( h = 0.7 \). Initial conditions for the simulation were set using the Zeldovich approximation at starting redshift \( z_i = 100 \).

A halo catalogue was created for the given snapshot using the Bound Density Maximum algorithm (Klypin & Holtzman 1997; Riebe et al. 2013). We populated these halos with mock galaxies using the Halo Occupation Distribution (HOD) model of Zheng et al. (2007), assigning central and satellite galaxies to halos according to a distribution based on the halo mass. Details of the algorithm and HOD model parameters used are described more fully in Nadathur et al. (2017), who used the same mock catalogue: these parameters were taken from Manera et al. (2013) and are designed to approximately reproduce the clustering and mean number density for galaxies in the Baryon Oscillation Spectroscopic Survey (BOSS) CMASS galaxy sample.

To measure dark matter (DM) densities in the simulation, we used a cloud-in-cell interpolation scheme to determine the DM density on a 2500\(^3\) grid using the full particle output of the simulation. We used this grid density field to determine the dark matter power spectrum \( P(k) \). We then measured the galaxy power spectrum \( P_{gg}(k) \) for the mocks; by fitting for the ratio \( P_{gg}(k)/P(k) = b^2 \) at large scales, \( k \lesssim 0.05 \ h\)Mpc\(^{-1} \), we determine the linear bias value for the galaxy mocks, \( b = 1.88 \).

Using the real space galaxy positions \( \mathbf{x} \) and velocities \( \mathbf{v} \), we determine their redshift space positions in the plane-parallel approximation assuming the line of sight direction to be along the \( z \)-axis of the simulation box,

\[
\mathbf{s} = \mathbf{x} + \frac{\mathbf{v} \cdot \hat{z}}{aH} \cdot \hat{z}.
\]  

(1)

2.2 Void finding

We identify voids in the real space galaxy mocks using the ZOBOV watershed void-finding algorithm (Neyrinck 2008). The ZOBOV algorithm uses a Voronoi tessellation field estimator (VTFE) technique to reconstruct the galaxy density field from the discrete distribution, and then identifies local minima in this field and the watershed basins around them, to form a non-overlapping set of voids corresponding to local density depressions. As in previous works (Nadathur et al. 2016; Nadathur et al. 2017), we define each individual density basin as a distinct void, without any additional merging of neighbouring regions. A fuller description of the algorithm and void properties can be found in Nadathur (2016) and Nadathur et al. (2017).

Although voids are of arbitrary shape and are in general far from spherically symmetric, it is convenient to define an effective spherical radius, \( R_v = (3V/4\pi)^{1/3} \), where \( V \) is the
total volume of the void. We determine the centre of each void to be the centre of the largest sphere completely empty of galaxies that can be inscribed within the void (Nadathur & Hotchkiss 2015a; Nadathur et al. 2017). In Appendix A we consider the effect of defining the void centre as the volume-weighted barycentre of void member galaxies, as is also popular in the literature. Such a redefinition does not alter any of the qualitative conclusions in the following sections, but it decreases the available signal-to-noise for the RSD measurement and worsens the agreement with linear dynamics due to bulk velocity flows.

It is important to stress that we apply the void-finding algorithm to the real space galaxy mocks and not to the shifted version in redshift space. As we discuss in the next section, this is crucial because none of the theoretical models for RSD in the void-galaxy cross-correlation discussed in this paper or elsewhere in the literature are applicable unless the real space void positions are known. In a companion publication (Nadathur, Carter & Percival, in prep.), we show how this practical difficulty can be overcome when using survey data where the real space galaxy positions are not known.

Approximately 33000 voids are identified in the simulation box. As the void-finding algorithm is space-filling, these voids cover the entire box volume, and undoubtedly include some spurious identifications that do not correspond to genuine matter underdensities. We also do not expect a linear model of coherent velocity outflow from a void to successfully describe the RSD around all voids. In particular it is not expected to work where the local environment of structures outside the void is important in determining the velocity field. This is expected to be the case for small voids, and indeed Cai et al. (2016) find that RSD models for the void-galaxy correlation do not work for small voids. However, the distinction between ‘large’ and ‘small’ voids is somewhat ambiguous, and the numerical value of the cut on void size depends both on the bias and number density of the galaxies in question (Nadathur & Hotchkiss 2015a,b) as well as on the particular features of the voidfinder.

We therefore restrict our void sample to the half of all voids with effective radius greater than the median radius. This is an easily reproducible criterion. For the mocks and voids used in this work, this means selecting \( R_v \geq 43 \, h^{-1}\text{Mpc} \), which leaves 16 421 voids. Throughout the rest of this paper, we exclusively use this sample of voids. The mean effective void radius for this sample is \( \overline{R}_v = 55.6 \, h^{-1}\text{Mpc} \). In some figures in later sections we show distances from the void centre in units of this mean radius, for context. However in our analysis we do not rescale distances in units of individual void sizes. Such a rescaling requires a strong implicit assumption of self-similarity of void profiles dependent only on void size (Nadathur et al. 2015), which is not justified (Nadathur & Hotchkiss 2015a,b; Nadathur et al. 2017). In addition, such a rescaling would effectively weight galaxy counts in the same volume differently depending on the assigned void effective size, which complicates the error determination.

Note that restricting our sample to the largest 50% of voids need not necessarily be the optimal choice to ensure validity of linear theory. Nadathur et al. (2017) show that environmental effects around voids are more strongly correlated with a combination of void size and density than with void size alone, so this may provide a better selection criterion. However, the median size cut implemented here has the advantage of simplicity, and as we show later, is sufficient that a purely linear RSD model already provides an excellent fit to the data.

3 THEORY

3.1 The void-galaxy cross-correlation in redshift space

Let \( \mathbf{X} \) denote the comoving location of a void centre, and \( \mathbf{x} \) the location of a galaxy in its vicinity. The real-space separation vector for the void-galaxy pair, \( \mathbf{r} = \mathbf{x} - \mathbf{X} \), is transformed to \( \mathbf{s} \) in redshift space. Assuming that the total number of void-galaxy pairs is unchanged by the shift to redshift-space, we require that

\[
(1 + \xi_{sg}(s)) \, d^3 s = (1 + \xi_{sg}(r)) \, d^3 r,
\]

where \( \xi_{sg} \) denotes the void-galaxy cross-correlation (in what follows we will suppress the subscript where there is no risk of confusion), and the superscripts ‘a’ and ‘r’ denote redshift and real space respectively. \( \xi_{sg} \) may also be viewed as the galaxy number density profile around the void, and is sometimes denoted \( \delta_g \) in this context.

If we further assume that the void centre is stationary, so that only the galaxy peculiar velocity \( \mathbf{v} \) is relevant to the RSD pattern, in the distant-observer approximation we may write

\[
\mathbf{s} = \mathbf{r} + \frac{\mathbf{v} \cdot \mathbf{X}}{aH} \mathbf{X},
\]

where \( a \) is the scale factor and \( H \) is the Hubble rate.

Note that the assumptions that void centres remain stationary and that their numbers are conserved between real and redshift space are both known to fail. That is, if voids are identified using the redshift space galaxy field, different numbers and centres of voids are obtained than if they are identified using the same algorithm applied to real space galaxies (see also the discussion in Chuang et al. 2017). In a separate work, we will show that these effects are very important, producing strong contributions to the observed RSD pattern even within the scale of the mean void radius. Given the complex and non-linear operation of void-finding algorithms, accounting for these effects in the modelling is challenging. In the present work we restrict ourselves to developing the theory for the more manageable case where both assumptions are valid. To ensure this, we identify voids using the real space galaxy positions in the simulation, and consider the RSD pattern about the centres thus defined. In a companion paper (Nadathur, Carter & Percival, in prep.) we discuss practical measures for subtracting this additional distortion from data so that a fair comparison with theory is possible.

From Eq. 3, the line-of-sight component of the separation vector is

\[
\mathbf{s}_l = \mathbf{r}_l + \frac{v_l}{aH}.
\]

Symmetry arguments show that the average coherent velocity flow must be directed along the radial direction, \( \mathbf{v} = v_c \hat{r} \). The determinant of the Jacobian for the coordinate transformation is therefore
\[ J \left( \frac{s}{r} \right) = 1 + \frac{v_r}{r a H} + \frac{(v'_r - v_r/r)}{a H} \mu^2, \]

where \( \mu \) denotes the derivative with respect to the radial distance \( r \), and \( \mu \) is the cosine of the angle between the line-of-sight direction and the separation vector,

\[ \mu = \frac{\mathbf{X} \cdot \mathbf{r}}{X|\mathbf{r}|} = \cos \theta. \]

Eq. 2 can therefore be rewritten as

\[ 1 + \xi^s(s) = (1 + \xi^c(r)) \left[ 1 + \frac{v_r}{r a H} + \frac{(v'_r - v_r/r)}{a H} \mu^2 \right]^{-1}. \]

We will assume that the peculiar velocity field is coupled to the density as in linear dynamics, and is determined by the void alone, so that

\[ \mathbf{v}(r) = -\frac{1}{3} f a H \Delta(r) \mathbf{r}, \]

where \( \Delta(r) \) is the average mass density contrast within radius \( r \) of the void centre,

\[ \Delta(r) \equiv \frac{3}{r^3} \int_0^r \delta(y) y^2 dy, \]

where \( \delta(y) \) is the mass density profile of the void, and \( f = \frac{d \ln D}{d \ln a} \), with \( D \) the growth factor and \( a \) the scale factor, is the linear growth rate of density perturbations.

Combining Eqs. 7 and 8, and using the fact that \( \Delta'(r) = 3/r (\delta(r) - \Delta(r)) \), we obtain

\[ 1 + \xi^s(s) = (1 + \xi^c(r)) \left[ 1 - \frac{f}{3} \Delta(r) - f \mu^2 (\delta(r) - \Delta(r)) \right]^{-1}. \]

The expansion of the second term on the RHS must be performed with care. For the voids in highly biased galaxy tracers that we consider in this paper, the mass density contrast is small enough that we can safely drop terms of second order or higher in \( \delta \) and \( \Delta \). However, note that this is separate from the assumption of the validity of linear dynamics, Eq. 8, and may not hold for other void populations, even if the dynamics is close to linear.

Crucially also, independent of assumptions about the size of \( \delta \) and \( \Delta \), the real space galaxy correlation \( \xi^c(r) \) cannot be assumed to be small. The very fact that voids are identified by selecting regions with very few or no galaxies ensures that in the void interior \( \xi \sim -1 \), and thus that terms proportional to \( \delta \) and \( \Delta \) must be retained to linear order. The left panel of Figure 1 demonstrates this explicitly for the voids in our simulation.

Retaining these terms in the expansion, Eq. 10 reduces to

\[ \xi^s(s, \mu) = \xi^c(r) + \frac{f}{3} \Delta(r) (1 + \xi^c(r)) + f \mu^2 [\delta(r) - \Delta(r)] (1 + \xi^c(r)), \]

where the difference between real space and redshift space coordinates \( r \) and \( s \) is

\[ r = s \left( 1 + \frac{f}{3} \Delta(s) \mu^2 \right), \]

to the same linear order in \( \Delta \). The effect of this coordinate shift is neglected in the linear Kaiser theory for RSD in the galaxy correlation (Kaiser 1987) because it only appears at second order in that derivation (see, e.g., Section 4.2 of Hamilton 1998). However, in our case, the coordinate shift appears in \( \xi \) at linear order, via the Taylor expansion

\[ \xi(r) = \xi(s) + \xi'(s) \frac{r}{3} s \Delta(s) \mu^2 + \ldots. \]

The effect of the terms in the model above can be understood as follows. The velocity outflow from voids introduces a mapping between the real space correlation \( \xi^c \) at separation \( r \) and the redshift space correlation \( \xi^c \) at separation \( s \), where \( r < s \) along the line of sight, for \( \Delta < 0 \). Since \( \xi \) decreases towards the centre, this results in a ‘stretching’ of the void along the line of sight direction, in accordance with naive intuition for the RSD effect around underdensities. This effect is most important where the gradient of \( \xi^c \) is steepest.

On the other hand, as noted and explained by Cai et al. (2016), the terms \( f \Delta/3 + f \mu^2 (\delta - \Delta) \), arising from the Jacobian of the transformation, give rise instead to a ‘squashing’ along the line of sight.1 This effect is suppressed by a factor of \( (1 + \xi^c) \) and is therefore most important around the void edges and outside the void. The competing effects of the stretching and squashing terms can be seen in Figure 4. The combination results in a change in the sign of the quadrupole term within the void, going from negative to positive with increasing \( s \), as will be shown in Section 4.

3.2 Linear bias approximation

Eqs. 11 and 12 together provide the correct expression for the void-galaxy cross-correlation in redshift space assuming linear dynamics and at linear order in \( \delta, \Delta \). They form the base model that we will use to explore the simulation data. However, their use requires knowledge of the true mass density contrast \( \delta(r) \) and its integral version \( \Delta(r) \), which cannot be directly determined for voids in survey data. An approximation that is often made (Cai et al. 2016; Hamaus et al. 2017) is to assume that a simple linear bias relationship holds within the voids, so that \( \xi^c(r) = b \delta(r) \) and \( \bar{\xi}^c(r) = 3/r^3 \int_0^r \xi^c(y) y^2 dy = b \Delta(r) \), where \( b \) is the large-scale linear bias factor determined from the galaxy clustering.

Under this approximation, the full model of Eqs. 11 and 12 can be rewritten as

\[ \xi^s(s, \mu) = \xi^c(r) + \frac{\beta}{3} \bar{\xi}^c(r) (1 + \xi^c(r)) + \beta \mu^2 [\xi^c(r) - \bar{\xi}^c(r)] (1 + \xi^c(r)), \]

where

\[ r = s \left( 1 + \frac{\beta}{3} \bar{\xi}^c(s) \mu^2 \right), \]

and \( \beta \equiv f/b \).

To explore the validity of this linear bias approximation within voids, we determine the linear bias factor for our mock galaxies by computing the (real space) galaxy power spectrum \( P_{gg}(k) \) and the full matter power spectrum \( P(k) \)

1 In principle it is possible for these terms to also lead to a stretching for pathological choices of \( \delta(r) \), but in such a case it is unlikely that a linear model based on Eq. 8 would be viable anyway. In practice, we do not observe such behaviour for any subset of the voids in our simulation.
Importantly, the deviations are (i) scale dependent, and (ii) bias approximation are shown in the right panel of Figure 1. A similar though less pronounced kink is present in Figure 2. A similar though less pronounced kink is present in matter overdensity $\delta$ caused by the void-finding procedure, which we do not consider here. Where a linear bias $b$ characterises the galaxy clustering as above, the relationship between the local galaxy overdensity $\delta_g$ and the matter overdensity $\delta$ is still $\delta_g = b\delta + \epsilon$, where $\epsilon$ represents a stochastic term. If the stochastic term is unbiased, the conditional expectation value over the entire universe or simulation box is indeed
\begin{equation}
\langle \delta_g \rangle = b\delta.
\end{equation}
However, voids are by construction selected as relatively rare regions with a small galaxy density. This necessarily introduces a selection effect in $\langle \delta_g \rangle$: regions with negative fluctuations in $\epsilon$ are more likely to be selected as voids than those with positive $\epsilon$, thus biasing the mean relationship such that $\xi^e/r = \delta$ systematically overestimates the depth of the void where $\xi^e < 0$. Conversely, $\xi^e/b$ overestimates the height of the wall around the void edge where $\xi^e > 0$. This will be true even when $b = 1$, i.e. when the tracers used to identify the voids and determine $\xi$ are a random subset of the DM particles in the simulation.\footnote{Under certain simplifying assumptions about the nature of the stochastic term, this selection effect can be modelled analytically; see, e.g., Gruen et al. (2016), Sec. 3.1. In general, the void selection condition will introduce correlations between the stochastic terms at different radial distances, complicating the modelling. We will not pursue this further in the current work.}

ever be much reduced or even absent if the tracer profile $\xi^e$ is measured using a different set of tracers to those used to select void locations, such as an overlapping galaxy sample with higher mean number density, or the set of all halos in the simulation from which the mock galaxy hosts are drawn.

For our purposes, the consequences of this failure of the linear bias relationship within voids are two-fold:

(i) if the same galaxy tracers are used to identify the voids and to measure the RSD pattern, Eqs. 14 and 15 are necessarily always inaccurate with respect to the true model in Eqs. 11 and 12 in the void interiors, but

(ii) the extent of the discrepancy is somewhat mitigated by the fact that the discrepancy from the linear bias value is largest in the regions where the $(1 + \xi^e(r))$ term in Eq. 14 is small.

For the mocks we have both the matter and galaxy density fields, so we can contrast results for the real to redshift space mapping using the directly measured density field within voids as well as that inferred from the linear bias assumption, as done in Section 4. This allows us to isolate the effect of this assumption on the RSD model.

### 3.3 Multipole expansion

It is convenient to expand $\xi^e$ in terms of its multipoles,
\begin{equation}
\xi^e(s) = \int_0^1 \xi^e(s, \mu) (1 + 2\ell) P_\ell(\mu) d\mu
\end{equation}
where $P_\ell(\mu)$ are the Legendre polynomials of order $\ell$. At linear order in $\delta$ and $\Delta$, only the monopole and quadrupole terms are non-zero. Using $P_0(\mu) = 1$ and $P_2(\mu) = (3\mu^2 - 1)/2$, these can be calculated by direct integration for the model using either Eqs. 11 and 12 or Eqs. 14 and 15. For all numerical calculation of model multipoles presented in this paper, we use this approach.

However, approximate analytical forms can also be obtained by using the first two terms of the expansion in...
Eq. 13, to rewrite Eq. 11 as
\[
\xi'(s, \mu) \simeq \xi(s) + \frac{f}{3} \delta(s) (1 + \xi(s)) + \frac{f \mu^2}{3} \left[ \delta(\tau) - \Delta(\tau) \right] (1 + \xi(\tau)),
\]
and the quadrupole
\[
\xi''(s) = \frac{2 f}{9} s \xi'(s) \Delta(s) + \frac{2 f}{3} \left[ \Delta(\tau) - \delta(\tau) \right] (1 + \xi(\tau)).
\]
Corresponding versions of these equations can be obtained when also including the linear bias approximation.

3.4 Velocity dispersion and the streaming model

So far we have used a pure Kaiser model to describe \(\xi'\); that is, we assumed that velocities around void centres exactly follow the coherent outflow described by the linear relationship in Eq. 8. In Section 4 we will show that this is a surprisingly good approximation for the mean outflow velocity on all scales, so the model of Eq. 11 provides a good qualitative description of the void-galaxy correlation seen in simulation.

However, to enable a quantitative fit to the data, a more realistic model must account for the dispersion of galaxy velocities around this mean. To do this, we introduce a dispersion in the line-of-sight galaxy velocities, such that
\[
v = v_i + \hat{r} v_i, \tag{21}
\]
where \(v_i\) is the coherent radial component given by Eq. 8, and \(v_i\) is a zero-mean random variable with probability distribution function \(P(v_i)\). This results in an integral for \(\xi'\):
\[
1 + \xi'(\sigma, \pi) = \int d\sigma |v_i| P(v_i) \left( 1 + \xi'(\sigma, \pi - v_i/aH) \right), \tag{22}
\]
where \(\sigma\) and \(\pi\) are respectively distances transverse to and along the line of sight, \(r = \sqrt{\sigma^2 + \pi^2}\), with \(r_\sigma = \sigma\) and \(r_\pi = \pi - (v_i + v_i/\mu)/aH\), and \(J(\xi)\) is as in Eq. 5. The effect of the dispersion in \(v_i\) is primarily through shuffling the radial and transverse distances contributing to \(\xi'\) in the integral. The dispersion has a negligible effect on the Jacobian, as the contribution to the radial outflow averages to zero.

We will take the probability distribution function \(P(v)\) to have a Gaussian form,
\[
P(v) = \frac{1}{\sqrt{2\pi} \sigma_v} \exp \left( -\frac{v^2}{2\sigma_v^2} \right). \tag{23}
\]
The dispersion is most generally a function of the radial separation scale, \(\sigma_v = \sigma_v(r)\). The validity of the Gaussian assumption for \(P(v)\) may also vary with \(r\). We consider these questions in more detail in Section 4.

Note that Eq. 22 is not quite the same as an adaptation of the streaming model (e.g., Fisher 1995; Scoccimarro 2004; Reid & White 2011) to the void-galaxy case. In particular, as the coherence of density and velocity fields is accounted for by the inverse Jacobian, this is still fundamentally a linear model in all senses. We still assume the coherent outflow \(v_0\) is determined by linear dynamics as before; in evaluating the Jacobian we still drop terms higher than linear order in \(\delta\) and \(\Delta\); and the convolution with \(P(v)\) merely broadens the coordinate shift effect already present at linear order in Eqs. 11 and 12. In the limit of zero dispersion, where \(P(v_i) \rightarrow \delta_D(v_i)\), this model reduces to Eq. 11. Eq. 22 is thus similar to the dispersion model used for the galaxy correlation (Hamilton 1998) before the development of the full streaming model, except that the width of the velocity distribution is allowed to depend on scale.

Unlike in the case for the galaxy correlation, this linear dispersion model is already sufficient to provide an excellent fit to the data at all scales, as we show in Section 4.3. We may therefore safely leave the development of a full streaming model for the void-galaxy case to future work.

3.5 Comparison with previous results in the literature

The basic linear model we have presented above differs in several key respects to other models for the void-galaxy correlation in the literature (e.g., Paz et al. 2013; Cai et al. 2016; Hamaus et al. 2017; Achitouv et al. 2017; Achitouv 2017). In Section 4 we will compare these models to \(\xi'\) measured in the simulation data and show that our model provides a significantly better description of the true void-galaxy correlation.

Cai et al. (2016) (and subsequently Hamaus et al. 2017) follow a derivation similar to that described in Section 3.1, with three important differences: (1) they keep terms of order \(\xi\) and \(\xi\Delta\) in the expansion of Eq. 10, (2) they approximate \(s = r\) contrary to Eq. 12, and (3) they assume the linear bias approximation (Section 3.2) holds within voids. They therefore obtain the substantially simpler result
\[
\xi'(s, \mu) = \xi'(s) + \frac{\beta}{3} \xi'(s) + \beta \mu^2 \left[ \xi'(s) - \xi'(s) \right], \tag{24}
\]
which can be compared to Eqs. 11 and 14. This model gives expressions for the monopole
\[
\xi_0(s) = \left( 1 + \frac{\beta}{3} \right) \xi'(s), \tag{25}
\]
and the quadrupole
\[
\xi_2(s) = \frac{2 \beta}{3} \left[ \xi'(s) - \xi'(s) \right]. \tag{26}
\]

A simple way to see that these cannot be correct within the void interior is to consider the limiting case where \(\xi' \sim -1\) close to the void centre, at \(s \sim 0\). Here Eq. 25 predicts a redshift space monopole \(\xi_0 < -1\), which is unphysical. More generally, Eq. 24 is a poor description of

\[3\] Cai et al. (2016) actually provide expressions assuming \(\xi' = \delta\) and \(\xi' = \Delta\), which is equivalent to using the linear bias assumption with \(b = 1\).
\(\xi'(s)\) everywhere that \((1 + \xi')\) is small, which in practice for the voids considered here means at least for all radial separations within the mean void radius.

In addition, Eq. 24 entirely misses the important stretching effect along the line of sight due to the mapping of void-galaxy separations from real to redshift space. As we show in the next section, this causes a change in the sign of the quadrupole term within the void radius which is not captured by Eq. 26. In fact this change of sign was already seen in simulation data by Cai et al. (2016) (see Figures 1-3 in their paper), although not satisfactorily explained. This failure of the model is the primary reason why the growth rate estimator proposed in that paper fails for \(s < \Pi_v\), leading to negative reconstructed values of the growth rate within voids (see the discussion in Cai et al. 2016).

On the other hand, in contrast to both our results and those of Cai et al. (2016), Hamaus et al. (2017) do not observe any change of sign in the quadrupole \(\xi'\) measured in BOSS galaxy data and associated galaxy mocks. This is because they use voids identified using the galaxy positions in redshift space—when this is done, neither of the fundamental assumptions used to obtain Eqs. 2 and 3 are valid, and so none of the theoretical models for \(\xi'\) discussed in this paper or elsewhere in the literature are applicable. We discuss this problem and its resolution in a separate work (Nadathur, Carter & Percival, in prep.).

Finally, a number of different authors (e.g. Paz et al. 2013; Hamaus et al. 2015, 2016; Cai et al. 2016; Achitouv et al. 2017; Achitouv 2017) have used an analogy with the Gaussian streaming model for the galaxy autocorrelation to model \(\xi'\):  
\[
1 + \xi'(\sigma, \pi) = \int \frac{(1 + \xi'')}{\sqrt{2\pi}\sigma_v} \exp\left(-\frac{(v_{||} - v_r(r)\mu)^2}{2\sigma_v^2}\right) dv_{||}, \tag{27}
\]

with \(v = \sqrt{\sigma^2 + (\pi - v_l/aH)^2}\). Comparison with Eq. 22 shows that this expression differs from our dispersion model. Unlike our dispersion model, it does not reduce to the linear expression, Eq. 11, in the limit of zero dispersion when \(v = v_r\). Also unlike our dispersion model, it provides a poor fit to the data, as we will show in Section 4.

The reason for this is simple: the streaming model in Eq. 27 has been derived for the case of linear, Gaussian fluctuations in the galaxy or matter autocorrelation (Fisher 1995; Scoccimarro 2004), and cannot be extended to the void-galaxy cross-correlation. Not only is the moment generating function \(Z(\lambda, r)\) (see Scoccimarro 2004) different for the void-galaxy case, the derivation of the Gaussian streaming model explicitly assumes that the only non-zero cumulant of the density field is the second. However, the void-galaxy case involves constrained averages of the fields at specially selected locations rather than over all space, so that by definition the expectation value \(\langle \delta \rangle \neq 0\) within the void. As a result the analogy is invalid and Eq. 27 does not describe the correct streaming model for the void-galaxy cross-correlation.

4 This model does approximately match Eqs. 11 and 22 when the conditions \(|\xi''|, |\delta|, |\Delta| \ll 1\) are satisfied and the dispersion is small, \(\sigma_v \ll raH\) (Cai et al. 2016). In practice this holds only outside the void boundary, \(s \gtrsim 1.5\Pi_v\).

\section*{4 SIMULATION RESULTS}

4.1 Void density profiles

We measure the density profiles \(\delta_i(r)\) for each void \(i\) using the full resolution DM particle output of the simulation at the \(z = 0.52\) snapshot in 50 equally-spaced radial bins out to separations of 120 \(h^{-1}\)Mpc. We use the same binning for each void and do not rescale separations by the void radius; therefore, these individual profiles can be simply averaged over all voids to give the stack average profile \(\delta(r)\). We measure \(\xi'(r)\), which is equivalent to the average real space galaxy density profile around voids, in the same way after substituting galaxy tracers for the DM particles. To determine \(\Delta(r)\) and \(\xi'(r)\) we interpolate \(\delta(r)\) and \(\xi'(r)\) respectively and numerically evaluate the corresponding integrals.

The profiles obtained are shown in Figure 1, where for context we have shown radial separations in units of the mean radius of all voids in the stack, \(\Pi_v = 55.6\ h^{-1}\)Mpc. The importance of the relative sizes of the different terms and the failure of the linear bias approximation have already been discussed in Section 3.

4.2 Velocity profiles and dispersion

We measure the average (stacked) galaxy velocity profile around the void centres in the simulation as
where $r_{ij}$ is the real space separation vector between the $i$th void and $j$th galaxy and $v_j$ the velocity of that galaxy, and the sum runs over all $N_{vg}$ void-galaxy pairs with $|r_{ij}|$ in the range $(r - dr, r + dr)$. As before, we use 50 equally-spaced bins out to a maximum separation of $120$ $h^{-1}$Mpc. We also measure the dispersion in the line-of-sight velocity component, $\sigma_{v_{ij}}^{gal}$, defined by

$$
\sigma_{v_{ij}}^{gal}(r) = \left[ \frac{1}{N_{vg}} \sum_{i,j} \left( v_j \cdot \hat{r}_{ij} - \bar{v}_{ij}(r) \right)^2 \right]^{1/2},
$$

where $\hat{r}_{ij}$ is the line of sight direction, which is the same for all galaxies in the plane-parallel approximation and taken to be along the $z$-axis of the simulation box, $\bar{v}_{ij}(r)$ is the mean line of sight velocity component at $r$, and the sum is over all void-galaxy pairs in the given separation bin as before.

Note that Eqs. 28 and 29 weight each void-galaxy pair equally. An alternative way to define these quantities, which has sometimes been used in the literature, is to calculate the average and dispersion separately for each void, and then average the results over all voids. This is equivalent to weighting each void equally: this procedure leads to smaller measured values of $\sigma_{v_{ij}}^{gal}$ and $\sigma_{v_{ij}}^{gal}$ at small separations, since most voids do not have any galaxies in interior bins and thus contribute zero to the average. At large $r$ both methods will largely agree. However, for comparison with the void-galaxy cross-correlation, which also weights each void-galaxy pair equally, our method is the more appropriate.

Finally, to control for possible statistical effects due to the void selection criterion as seen in the relationship between $\xi(r)$ and $\delta(r)$, we also measure the average halo velocity profile around void centres, using the equivalent of Eq. 28 but for the full halo catalogue. This has many more objects than the mock galaxy catalogue and therefore avoids the problem of low tracer statistics close to void centres. Under the reasonable assumption that the halo velocity field is unbiased with respect to the dark matter velocity, we refer to this quantity as $v_{rDM}(r)$.

Figure 2 plots the results as a function of distance from the void centre, together with the linear velocity model for $v_{r}(r)$ obtained from Eq. 8 (solid black line). The linear model is calculated using the fiducial value of the growth rate for the given simulation and redshift, $f = 0.761$. The first conclusion that can be drawn from this is that linear dynamics gives an extremely good description for the mean $v_{rDM}(r)$, with deviations at $\lesssim 5\%$ over the entire range of scales tested (see also Hamaus et al. 2014). It is also a good model for the mean galaxy radial velocity profile $v_{r}^{gal}(r)$ at distances $r \gtrsim 30$ $h^{-1}$Mpc, where $v_{r}^{DM}(r)$ and $v_{r}^{gal}(r)$ coincide.

Closer to the void centres, $r < T_{min}$, $v_{r}^{gal}(r)$ starts to deviate from both the linear model and $v_{r}^{DM}(r)$. While some scatter is expected where the errors in $v_{r}^{gal}(r)$ become large due to the small number of void-galaxy pairs (note that the two interior-most bins have no galaxies at all), in the range $10 \lesssim r \lesssim 30$ $h^{-1}$Mpc the difference is statistically significant. The fact that it coincides with the ‘kink’ visible in the $\xi(r)$ profile in Figure 1 and a similar, though less pronounced, kink that is also present at the same distance in $\delta(r)$, suggests that this is a physical consequence of some aspect of the void selection algorithm. We leave a fuller explanation of this effect for future work.

Recently, Achitouv (2017) proposed a semi-empirical non-linear correction term to modify the linear velocity re-
the measured dispersions derived in Section 3.1 above. Close to the void centre, Hamaus et al. (2017), Eq. 24. The qualitative features visua-

2 The measured values of the monopole $\xi^0(s)$ and quadrupole $\xi^2(s)$ from the simulation data, com-

pared to theory predictions derived in the previous section evaluated using the fiducial growth rate $f = 0.761$. For visual clarity, the monopole values are shown as the difference with respect to the real space version, $\xi^0 - \xi^2$. Two important features are immediately apparent: the quadrupole changes sign within the void interior, with a negative dip at intermediate values of $s$ changing to a broad $\xi^2(s) > 0$ feature at $s \approx \mathcal{R}_v = 55\ h^{-1}\text{Mpc}$; and the redshift space monopole correlation function for line of sight velocities in four different bins of $\sigma_v\|\xi^0\|$, compared to the Gaussian distribution assumed in Eq. 23

4.3 Redshift space correlation function

We measure the void-galaxy cross-correlation function in the simulation data using a modified version of the CUTE correlation function code (Alonso 2012). Note that, as discussed in Section 3, in order to ensure compatibility between data and the theoretical modelling, void positions are identified using the real space galaxy positions in the simulation.

Figure 4 shows the measured redshift space cross-correlation $\xi^r(\sigma, \pi)$ as a function of the transverse and line-of-sight distances, compared to the prediction from our model, Eq. 11, and that of the model of Cai et al. (2016) and Hamaus et al. (2017), Eq. 24. The qualitative features visible in the data can be understood on the basis of the equations derived in Section 3.1 above. Close to the void centre, the measured $\xi^r$ shows approximate spherical symmetry. At intermediate distances, the effect of the coordinate shift discussed in Section 3.1 causes an elongation effect, stretching the contours of equal $\xi^r$ along the line of sight direction. At large distances, $s \approx \mathcal{R}_v = 55\ h^{-1}\text{Mpc}$, this effect is reversed, with a relative squashing of the contours of $\xi^r(\sigma, \pi)$ along the line of sight direction.

To better understand the relative merits of the two theoretical models, Figure 5 shows the residual difference between the measured $\xi^r(\sigma, \pi)$ and the predicted correlation functions in the two cases. It is clear that the model of Cai et al. (2016); Hamaus et al. (2017) is a very poor fit to the data at all radial separations within the mean void radius $\mathcal{R}_v$, and particularly so very close to the centre, where Eq. 24 predicts unphysical values of $\xi^r < -1$. By contrast, the inclusion of the $(1 + \xi(r))$ factors in Eq. 11 correctly ensures that the theory prediction matches the measured value at the void centre—where in fact RSD effects are absent, $\xi^r(s) \approx \xi^r(s)$—and also produces a much better fit to the data at all distances within the void interior.

Nevertheless, Figure 5 shows that even for the improved model, when the velocity dispersion term is not included the theory residuals are still potentially significant. To perform a quantitative analysis, we instead perform a multipole expansion of the angular correlation function $\xi^r(\sigma, \mu)$ from the simulation data to extract the monopole and quadrupole terms using Eq. 17 for comparison with theory. The multipoles are measured in 50 radial bins in the range $0 \leq s \leq 120\ h^{-1}\text{Mpc}$, and we use 100 bins in $\mu$. Note that we do not rescale radial distances by the void radius, thus using the same bin sizes for each void. Such a rescaling would constitute a strong assumption of self-similarity in void profiles, which is not justified (Nadathur & Hotchkiss 2015a,b; Nadathur et al. 2017).
Figure 4. Left: The 2D void-galaxy cross-correlation function $\xi^s(\sigma, \pi)$ in redshift space measured for voids in the simulation. Dashed curves indicate contour lines for $\xi^s = -0.9$, $-0.6$ and $-0.3$, the solid curve is the contour $\xi^s = 0$. Centre: The corresponding theoretical prediction for our linear model. Eq. 11. Right: The model prediction for the model of Cai et al. (2016) and Hamaus et al. (2017), Eq. 24. Note that in this case the colour scale is saturated at the centre, as the model predicts $\xi^s < -1$.

Figure 5. Residual differences $\xi^s(\sigma, \pi) - \xi^s_{th}(\sigma, \pi)$ between the measured void-galaxy cross-correlation and theoretical predictions. Left: For the model in Eq. 11. Right: For the model in Eq. 24. The thick dashed circle in both plots indicates the mean void radius $R_v$.

diffrs from the real space version only at intermediate $s$, with $\xi^s_0 - \xi^r \simeq 0$ both at the very centre (where $\xi^r \simeq -1$) and around the mean void radius.

The dashed curves in each panel show the predictions of the basic linear model derived from Eq. 11 without accounting for velocity dispersion: these capture the main qualitative features above but do not provide a good fit to the data, as expected from visual inspection of the residuals in Figure 5. The solid curves show the prediction for the linear dispersion model of Eqs. 22 and 23, where we have used the radial dependence of the dispersion $\sigma_v(r)$ measured in the data, shown in Figure 2. This model provides an excellent fit to the data at all scales: the $\chi^2$/d.o.f. values for this model are 41.1/48 for the monopole alone, 54.3/48 for the quadrupole alone, and 121.3/98 for both combined.

Finally, the dotted curves in Figure 7 show the predictions for $\xi_0^s$ and $\xi_2^s$ according to the model presented by Cai et al. (2016); Hamaus et al. (2017), Eqs. 25 and 26. Unsurprisingly, this model fails qualitatively and quantitatively to describe the data at any point within the void interior. Outside the mean void radius $R_v = 55$ $h^{-1}$Mpc, the multipoles from this model approach the correct expression: this explains the observation of Cai et al. (2016) that the linear growth rate estimator proposed in that paper works only in the region $s > R_v$.

Table 1 compares the goodness of fit for the various models discussed: here ‘lin.+dispersion’ refers to the full linear dispersion model above, ‘lin.+dispersion+bias’ to this model with the additional simplifying linear bias assumption $\xi^r(r)/b = \delta(r)$, ‘lin. only’ to the model without dispersion, Eq. 11, and ‘lin.+bias’ to this model with the addition of the bias assumption, Eq. 14. The fit for the streaming model of Eq. 27 uses $\delta(r)$ (without the linear bias assumption) to obtain $v_s(r)$, and $\sigma_v(r)$ measured from simulation, so can be directly compared to the dispersion model without bias. The fit for the Cai et al. (2016) model includes the bias assumption, as in Eq. 24.
been described in the literature, e.g. Hamaus et al. (2014); Ricciardelli et al. (2014); Nadathur et al. (2015); Nadathur & Hotchkiss (2015a,b); Barreira et al. (2015); Cautun et al. (2016). However, there is no consensus on any particular fitting form, and in any case the fit will strongly depend on the void-finding and stacking algorithms as well as on the properties of the galaxy sample in question (Nadathur & Hotchkiss 2015b), so the calibration needs to be performed on a case-by-case basis. Alternative methods to obtain $\xi'$ are by deprojecting the redshift space monopole $\xi_0$ (Pisani et al. 2014) or by using reconstruction techniques (Nadathur, Carter & Percival, in prep.).

In principle the matter overdensity profile $\delta'(r)$ can be determined independently of $\xi'(r)$ through void lensing measurements (Krause et al. 2013; Melchior et al. 2014; Clampitt & Jain 2015; Sánchez et al. 2016). Alternatively, it could also be calibrated from simulations as for $\xi'$. Use of the linear bias assumption $\delta'(r) = \xi'(r)/b$ is however not a good approximation if the void-galaxy correlation is to be determined for the same galaxy tracers used to identify the voids. This is because of the selection effect described in Section 3.2, that introduces a systematic scale-dependent shift to the inferred relationship between $\xi'(r)$ and $\delta'(r)$. For instance, using the linear bias approximation for our simulation data leads to errors in the predicted quadrupole $\xi_2$ of up to 20%.

The other piece of information required to perform a fit for $f$ is the dispersion relation $\sigma_{\delta^2}(r)$. It is possible that future work on a complete streaming model for the void-galaxy case may allow this function to be determined theoretically from $\xi'(r)$ and $\delta(r)$, or it could be calibrated from simulation in the same way as $\delta(r)$. At present we take its value directly from the simulation measurements. This leaves only a single free parameter in our model: the growth rate $f$.

Allowing $f$ to vary and fitting to the measured multipole data using Eq. 22, we obtain the posterior likelihoods shown in Figure 8. This gives a recovered value of $f = 0.78 \pm 0.02$ (68% c.l.), consistent with fiducial value $f = 0.761$ for the simulation, and corresponding to a 2.7% precision for our simulation volume of $(2\, h^{-1}\text{Gpc})^3$. The constraints obtained from the monopole and quadrupole taken individually are also shown, and are consistent with each other and with the combined result.

Table 2 shows how the constraints on $f$ change when only data within different separation ranges are used for the fits. In all cases the constraints obtained are consistent with each other and the final value. It is also clear that almost all of the constraining power of the data comes from the contribution to $\xi'$ from galaxies within the void interiors, $s < R_v$, which is the region where our model performs significantly better than the alternatives.

Finally, we tested the effect on the measurement of the growth rate of using the linear bias assumption $\delta(r) = \xi'(r)/b$. Due to the failures described above, this leads to a strongly biased estimator of $\beta$, with the reconstructed value being more than 3\sigma smaller than the fiducial for fits to both the quadrupole and monopole.

A popular alternative methodology (e.g. Hamaus et al. 2015; Cai et al. 2016; Achitouv 2017) is not to fix $\sigma_{\delta^2}(r)$ from simulation but to parametrise it using some functional form with one or more free parameters, which are to be...
Figure 7. Measured multipoles of the redshift space void-galaxy cross-correlation $\xi^s$ as a function of the radial void-galaxy separation $s$ (data points). Error bars are derived from diagonal elements of the estimated covariance matrix. Left: The monopole $\xi^s_0(s)$. For visual clarity, this is plotted as the difference $\xi^s_0 - \xi^r$ compared to the real space monopole, shown in Figure 1. Right: The quadrupole $\xi^s_2(s)$.

In both panels, the solid curve shows the theoretical prediction for our full linear model including velocity dispersion, using Eq. 22; the dashed curve is the prediction for the linear model without the velocity dispersion term, using Eq. 11; and the dotted curve is the prediction for the incorrect model of Eq. 24.

Figure 8. Likelihood for the growth rate $f$, obtained from fits of Eq. 22 to the measured multipoles of the void-galaxy cross-correlation, corresponding to a recovered value $f = 0.78 \pm 0.02$, a 2.7% constraint. The dashed and dotted curves show the likelihoods for fits to the monopole and quadrupole separately. The functions $\delta(r)$ and $\sigma_{v||}(r)$ are calibrated from the simulation. The vertical line indicates the fiducial value of the growth rate.

Figure 9. 68% and 95% confidence limit contours on the two free parameters $(f, \sigma_0)$, where $\sigma_0$ is a constant velocity dispersion, from fits to the quadrupole (blue contours) and monopole (red) from simulation data. The vertical line shows the fiducial value of $f$ for the simulation. Introducing $\sigma_0$ as a free parameter weakens the constraints on $f$. The quadrupole still provides an unbiased estimate of $f$, with 5.3% precision, but the monopole estimator is biased.

Table 2. Constraints on $f$ obtained from fits to data in different separation ranges relative to the mean void radius $R_v = 55.6 \, h^{-1}\text{Mpc}$.

| Data range      | $f$       |
|-----------------|-----------|
| $s < 0.5R_v$    | $0.79 \pm 0.03$ |
| $s < R_v$       | $0.77 \pm 0.02$ |
| all $s < 120 \, h^{-1}\text{Mpc}$ | $0.78 \pm 0.02$ |

marginalised over. The simplest possible such parametrisation is to take it to be a constant,

$$\sigma_{v||}(r) = \sigma_0. \tag{32}$$

Figure 2 suggests that this might be a reasonable approximation separately in the void interior and exterior regions. As sensitivity to the dispersion also decreases in the void exterior, where $\sigma_v \ll rh$, it might be hoped that this approximation is sufficient to reconstruct $f$.

To test this, we fit the model to the measured multipoles from the simulation data with $f$ and $\sigma_0$ as two free parameters. $\xi^s(r)$ and $\delta(r)$ are taken from the fits to the simulation data as before. Figure 9 shows the resulting 1 and 2$\sigma$ confidence level contours determined from fitting the quadrupole (blue) and monopole (red) separately. It is clear that the addition of the additional parameter loosens the constraints obtained considerably. The fit to the quadrupole provides an unbiased reconstruction of the fiducial growth rate, with the reconstructed value $f = 0.72 \pm 0.04$ at 68% c.l. after
marginalizing over $\sigma_0$, corresponding to a 5.3% constraint lower than but consistent with the fiducial value $f = 0.761$. However, fitting to the monopole $\xi_m^0$ gives a biased reconstruction of the growth rate which is more than 3$sigma$ from the fiducial value. This indicates a failure of the constant dispersion model and limits the amount of information that can be extracted from measurement of $\xi^s$ if a constant $\sigma_0$ is assumed.

6 CONCLUSIONS
Measurement of redshift space distortions in the void-galaxy correlation function $\xi^s(s)$ is an important tool that can be used to test for possible environmental dependence of the growth rate of structures. In particular, the growth rate in the lowest density regions close to void centres might contain information about possible non-standard theories of gravity. However, reconstructing the growth rate in voids requires a model for $\xi^s$ which can be trusted in these regions.

We have derived a configuration-space model for the void-galaxy correlation in redshift space using linear theory, which we characterise in terms of the multipole moments of the correlation. Our model accounts for several terms that are important within voids but have previously been neglected. As a result we are able to account for important physical effects that had not been appreciated, including the change in sign of the quadrupole term within the void, indicating a turnover point between stretching and squashing of the contours of the correlation function along the line of sight direction. The model can be broadened to include a dispersion in galaxy velocities along the line of sight; the dispersion model thus obtained differs from the streaming model used in previous studies, which was based on an inappropriate application of the formula for the Gaussian streaming model derived for the galaxy autocorrelation.

Comparing our model predictions to measurements of $\xi^s$ using purpose-built void and galaxy catalogues at redshift $z = 0.52$ in the Big MultiDark simulation shows that the linear dispersion model provides an excellent fit to the data for all values of the void-galaxy separation down to the minimum bin width used, $2.4\,h^{-1}$Mpc. This is an important contrast to the modelling of RSD in the galaxy correlation, where non-linear effects are very important at small pair separations and complicate the use of small-scale data. Comparison with the data shows that our linear model for the void-galaxy correlation performs significantly better than others in the literature, especially within the low density region of interest.

A consequence of our results is that the ratio of the redshift space monopole and quadrupole cannot be used as an estimator for the growth rate within voids, contrary to (Cai et al. 2016; Hamaus et al. 2017). This is because the linear model proposed in these papers fails at $s < R_v$. Determination of the growth rate using the correct expression we provide requires knowledge of the real space correlation function $\xi^s(r)$, which can be taken from fits to simulations or possibly reconstructed from redshift space data. Using functional forms for $\delta(r)$ and $\sigma_{\|}(r)$ determined from the simulation data, we show that fitting for the growth rate yields $f = 0.78 \pm 0.02$, a 2.7% constraint in good agreement with the fiducial value $f = 0.761$. Assuming a constant dispersion with amplitude $\sigma_{\|}(r) = \sigma_0$ taken as a free parameter, fits to the quadrupole alone still provide an unbiased estimate of the true growth rate $f$ after marginalising over $\sigma_0$, though with a reduced precision of 5.3%. However, the assumption of a constant linear bias leads to a biased reconstruction of the true $f$ from the monopole data.

Our results also highlight two very important points which have relevance to practical applications of this method to measure RSD effects within voids. Firstly, if the galaxy tracers used to measure the RSD effects are the same as those used to identify the voids, an assumption of linear galaxy bias within the void interior is not appropriate due to statistical selection effects on the void size scale. Assuming a constant linear bias leads to incorrect predictions from the model, which possibly require the addition of empirical non-linear correction terms (Achitouv 2017), even though in reality linear theory continues to provide a good description of the dynamics. In order to overcome this problem, in practical terms it would be better to use two differently biased galaxy tracers covering the same volume to identify void regions and to measure the RSD effects.

Secondly, we highlighted how the assumptions inherent in the derivation of the model require knowledge of the real space positions of voids, such that all RSD effects are due to motions of the galaxies only. In general if the voids are identified using galaxies in redshift space, this will not be the case (Chuang et al. 2017). Practical methods to reconstruct the real space void positions in data will be discussed in a companion paper.

Finally, a comment on the relative merits of using the void-galaxy correlation and the galaxy autocorrelation to measure RSD is in order. It is true that unlike for the galaxy RSD case, the model presented in this work is based only on linear theory, and fits the simulation data extremely well at all pair separations. It is particularly noteworthy that dispersion effects can be easily accommodated within a linear model. These appear to be significant advantages over the galaxy RSD case, for which quasi-linear modelling is required and small-scale data is excluded when performing fits. However, since most of the information in the void-galaxy correlation comes from those galaxies in the void interiors, this method also discards some of the available data. One therefore should not necessarily expect it to outperform the standard analysis, as has sometimes been suggested (Hamaus et al. 2017). Instead an advantage of the void-galaxy RSD measurement is specifically to constrain possible environmental dependence of the growth rate in low-density regions. It also provides a complementary measurement technique that is sensitive to different systematics. In particular, if the real space $\xi^s$ can be reconstructed directly from data rather than fit from simulations, our model depends directly on $f$ alone rather than the usual combination $f\sigma_8$, although further investigation is required to determine the dependence on $\sigma_8$ through the functions $\delta(r)$ and $\sigma_{\|}(r)$.

7 ACKNOWLEDGEMENTS
We thank Yanchuan Cai for helpful correspondence, and Davide Bianchi, Florian Beutler, Hans Winther and Rossana Ruggeri for stimulating discussions. SN acknowledges fund-
REFERENCES

Achitouv I., 2017, Phys.Rev.D, 96, 083506
Achitouv I., Blake C., Carter P., Koda J., Heitler F., 2017, Phys.Rev.D, 95, 083502
Alonso D., 2012, ArXiv e-prints, arXiv:1210.1833
Alonso D., Hill J. C., Hložek R., Spergel D. N., 2017, ArXiv e-prints
Barreira A., Cautun M., Li B., Baugh C. M., Pascoli S., 2015, J. Cosmol. Astropart. Phys., 8, 028
Beutler F. et al., 2012, MNRAS, 423, 3430
Beutler F. et al., 2017, MNRAS, 466, 2242
Cai Y.-C., Neyrinck M., Mao Q., Peacock J. A., Szapudi I., Berlind A. L., 2017, MNRAS, 466, 3364
Cai Y.-C., Taylor A., Peacock J. A., Padilla N., 2016, MNRAS, 462, 2465
Cautun M., Cai Y.-C., Frenk C. S., 2016, MNRAS, 457, 2540
Chuang C.-H., Kitaura F.-S., Liang Y., Font-Ribera A., Zhao C., McDonald P., Tao C., 2017, Phys.Rev.D, 95, 063528
Clampitt J., Jain B., 2015, MNRAS, 454, 3357
Deffayet C., Deser S., Esposito-Farèse G., 2009, Phys.Rev.D, 80, 064015
Dvali G., Gabadadze G., Porrati M., 2000, Physics Letters B, 485, 208
Fisher K. B., 1995, ApJ, 448, 494
Gottloeber S., Klypin A., 2008, ArXiv e-prints
Granett B. R., Neyrinck M. C., Szapudi I., 2008, ApJ, 683, L99
Gruen D. et al., 2016, MNRAS, 455, 3367
Guzzo L. et al., 2008, Nature, 451, 541
Hamaus N., Cousins M.-C., Pisani A., Aubert M., Escoffier S., Weller J., 2017, J. Cosmol. Astropart. Phys., 7, 014
Hamaus N., Pisani A., Sutter P. M., Lavaux G., Escoffier S., Wandelt B. D., Weller J., 2016, Physical Review Letters, 117, 091302
Hamaus N., Sutter P. M., Lavaux G., Wandelt B. D., 2015, J. Cosmol. Astropart. Phys., 11, 036
Hamaus N., Sutter P. M., Wandelt B. D., 2014, Phys. Rev. Lett., 112, 251302
Hamilton A. J. S., 1998, in Hamilton D., ed., Astrophysics and Space Science Library Vol. 231, The Evolving Universe. p. 185
Hotchkiss S., Nadathur S., Gottlöber S., Iliev I. T., Knebe A., Watson W. A., Yepes G., 2015, MNRAS, 446, 1321
Howlett C., Ross A. J., Samushia L, Percival W. J., Manera M., 2015, MNRAS, 449, 848
Hu W., Sawicki I., 2007, Phys.Rev.D, 76, 064004
Jennings E., Baugh C. M., Pascoli S., 2011, MNRAS, 410, 2081
Kaiser N., 1987, MNRAS, 227, 1
Kitaura F.-S. et al., 2016, Physical Review Letters, 116, 171301
Klypin A., Holtzman J., 1997, ArXiv e-prints
Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340
Kovács A. et al., 2017, MNRAS, 465, 4166
Krause E., Chang T.-C., Doré O., Umetsu K., 2013, ApJ, 762, L20
Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73
Manera M. et al., 2013, MNRAS, 428, 1036
Matsubara T., 2008, Phys.Rev.D, 78, 083519
Melchior P., Sutter P. M., Sheldon E. S., Krause E., Wandel B. D., 2014, MNRAS, 440, 2922
Nadathur S., 2016, MNRAS, 461, 358
Nadathur S., Crittenden R., 2016, ApJ, 830, L19
Nadathur S., Hotchkiss S., 2015a, MNRAS, 454, 2228
Nadathur S., Hotchkiss S., 2015b, MNRAS, 454, 889
Nadathur S., Hotchkiss S., Crittenden R., 2017, MNRAS, 467, 4067
Nadathur S., Hotchkiss S., Diego J. M., Iliev I. T., Gottlöber S., Watson W. A., Yepes G., 2015, MNRAS, 449, 3997
Neyrinck M. C., 2008, MNRAS, 386, 2101
Nicolis A., Rattazzi R., Trinercherini E., 2009, Phys.Rev.D, 79, 064036
Paz D., Lares M., Ceccarelli L., Padilla N., Lambas D. G., 2013, MNRAS, 436, 3480
Peacock J. A. et al., 2001, Nature, 410, 169
Pisani A., Lavaux G., Sutter P. M., Wandel B. D., 2014, MNRAS, 443, 3238
Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
Reid B. A. et al., 2012, MNRAS, 426, 2719
Reid B. A., White M., 2011, MNRAS, 417, 1913
Ricciardelli E., Quilis V., Varela J., 2014, MNRAS, 440, 1171
Samushia L., Percival W. J., Raccanelli A., 2012, MNRAS, 423, 2801
Samushia L., Percival W. J., Raccanelli A., 2012, MNRAS, 423, 2801
Spergel D., 2003, Phys.Rev.D, 67, 043505
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2005, MNRAS, 364, 1105
Taruya A., Nishimichi T., Saito S., 2010, Phys.Rev.D, 82, 063522
Taruya A., Nishimichi T., Saito S., 2010, Phys.Rev.D, 82, 063528
Thaler J., Primack J., Oguri M., Finoguenov A., 2011, MNRAS, 417, 2081
Tinker J., Coil A. L., Zehavi I., 2007, ApJ, 667, 760
Zheng Z., Cooray A. L., Zehavi I., 2007, ApJ, 667, 760

APPENDIX A: EFFECTS OF VOID CENTRE DEFINITION

Throughout this work, we have used the void centre definition introduced by Nadathur & Hotchkiss (2015a), which places the void centre at the point furthest removed from the galaxy positions, i.e. the point of minimum galaxy density.
within the void. The choice of void centre is however not unambiguous, and alternative definitions have previously been used in the literature. In particular, Hamaus et al. (2017) define the void centre to be the volume-weighted barycentre of the positions of galaxies within the void. The choice of centre for stacking can affect details of the measured density within a void (Nadathur & Hotchkiss 2015a,b; Nadathur et al. 2017). In particular, the choice of centre.

(i) the terms proportional to $\xi \delta$ and $\xi \Delta$ are of linear order and must be retained in the expansion of Eq. 10; (ii) the linear bias relationship $\xi'(r) = b \delta(r)$ does not apply within voids; (iii) the model of Eq. 24 fails to describe the measured $\xi^s$ within void interiors; but (iv) the model proposed in this paper, Eq. 22, provides a better description of the simulation data, are independent of the choice of void centre.

The first two of these statements are illustrated by Figure A1, which shows the same profiles and residuals as in Figure 1 except with distances measured relative to the void barycentres. The left panel shows that the measured $\xi'(r)$ and $\delta(r)$ profiles differ somewhat from those in Figure 1, in particular showing lower density contrasts at the void centres. This is expected, because the barycentre position is known to be a worse tracer of the region of underdensity within a void (Nadathur & Hotchkiss 2015a,b; Nadathur et al. 2017). In particular, $\xi'(r)$ shows a characteristic increase at small $r$ as noted in several previous works. However, the terms $\delta(r)\xi'(r)$ and $\Delta(r)\xi'(r)$ are still clearly of the same order as $\delta(r)$ within the void interior. The right panel shows that the linear bias approximation also does not apply for barycentre stacks, with a similar pattern of large and scale-dependent residuals from the bias relationship being observed. Neglecting both these effects, as done by Cai et al. (2016) and Hamaus et al. (2017) in deriving Eq. 24, therefore still results in an incomplete model independent of the choice of centre.

To demonstrate this explicitly, Figure A2 shows the multipoles $\xi^s_0(s)$ and $\xi^s_2(s)$ measured from the simulation for the same voids as in Figure 7, but for stacks based around the void barycentres. The curves shown on the plot correspond to the same theoretical models as before, i.e. Eq. 22 (solid line), Eq. 11 (dashed), and Eq. 24 (dotted). The theoretical curves differ from those shown in Figure 7 because the input $\xi'(r)$, $\delta(r)$ etc. differ for barycentre stacks.

Comparison of theory and data shows the same qualitative conclusions hold as before: i.e., Eq. 24 completely fails to describe the measured multipoles, whereas Eq. 22 provides a much better description. However, Figures 7 and A2 also reveal import differences. Firstly, for both the monopole and
the quadrupole, the amplitude of the RSD effect is smaller for the barycentre stacks, and the data is significantly noisier, with much larger error bars, as we have also checked by comparison of the full covariance matrices. This means that stacking around the void barycentre is sub-optimal for void RSD measurements. Secondly, the linear dispersion model of Eq. 22, while qualitatively still correctly capturing the physics, now provides a significantly worse quantitative fit to the data: the reduced $\chi^2$ values are now 1.63, 1.37 and 1.76 for fits to monopole, quadrupole, and both combined. These should be compared to the first line of Table 1.

The reason for this deterioration of the model fit can be understood by reference to Figure A3. This shows that for barycentre stacks, the measured velocity profiles $v_{\text{gal}}(r)$ (data points) and $v_{\text{DM}}(r)$ (thin solid curve) both differ much more strongly from the linear theory prediction based on Eq. 8 (thick solid curve) at small $r$ than for the stacks around the minimum density centre (Figure 2). This means that the assumption of the validity of the linear dynamics governed by the void density profile $\Delta(r)$ alone is less valid for stacks around the barycentre. A likely reason for this is that since void barycentre positions do not trace the position of the true minimum density within the void, they also are worse tracers of the stationary points of the velocity field; that is, bulk velocities are more significant at the void barycentre positions, making Eq. 8 a worse approximation. This explanation is consistent with the results of Nadathur et al. (2017), who show that void barycentre positions are worse tracers of maxima of the gravitational potential, where $\nabla \Phi = 0$ and thus the linear velocity field is zero.