

# Lawrence Berkeley National Laboratory

## Recent Work

### Title

A practical computational method for the anisotropic redshift-space three-point correlation function

### Permalink

<https://escholarship.org/uc/item/2pj878mw>

### Journal

Monthly Notices of the Royal Astronomical Society, 478(2)

### ISSN

0035-8711

### Authors

Slepian, Z  
Eisenstein, DJ

### Publication Date

2018-08-01

### DOI

10.1093/mnras/sty1063

Peer reviewed

# A practical computational method for the anisotropic redshift-space three-point correlation function

Zachary Slepian<sup>1,2†</sup> and Daniel J. Eisenstein<sup>3</sup>

<sup>1</sup>*Einstein Fellow, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

<sup>2</sup>*Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA*

<sup>3</sup>*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*

Accepted 2018 April 19. Received 2018 April 19; in original form 2017 November 3

## ABSTRACT

We present an algorithm enabling computation of the anisotropic redshift-space galaxy three-point correlation function (3PCF) scaling as  $N^2$ , with  $N$  the number of galaxies. Our previous work showed how to compute the isotropic 3PCF with this scaling by expanding the radially binned density field around each galaxy in the survey into spherical harmonics and combining these coefficients to form multipole moments. The  $N^2$  scaling occurred because this approach never explicitly required the relative angle between a galaxy pair about the primary galaxy. Here, we generalize this work, demonstrating that in the presence of azimuthally symmetric anisotropy produced by redshift-space distortions (RSD), the 3PCF can be described by two triangle side lengths, two independent total angular momenta, and a spin. This basis for the anisotropic 3PCF allows its computation with negligible additional work over the isotropic 3PCF. We also present the covariance matrix of the anisotropic 3PCF measured in this basis. Our algorithm tracks the full 5D redshift-space 3PCF, uses an accurate line of sight to each triplet, is exact in angle, and easily handles edge correction. It will enable use of the anisotropic large-scale 3PCF as a probe of RSD in current and upcoming large-scale redshift surveys.

**Key words:** cosmology: observations – distance scale – large-scale structure of Universe.

## 1 INTRODUCTION

Measuring the clustering of galaxies is a standard cosmological probe, revealing the Universe’s contents and laws while simultaneously illuminating the process of galaxy formation. This clustering is often quantified via correlation functions, which measure the excess above random of e.g. pairs (two-point correlation function, 2PCF) or triplets (three-point correlation function, 3PCF) of galaxies (Peebles 1980). The Universe is homogeneous and isotropic on large scales, so in principle the correlation functions should be direction independent (isotropic). However, in practice spectroscopic surveys differ in their method for measuring objects’ line-of-sight positions as opposed to the two transverse coordinates, so the correlation functions become direction dependent (anisotropic).

In particular, an object’s angular position is directly observable. In contrast, an object’s line-of-sight position is inferred from its redshift by presuming that the redshift is due solely to the object’s recession as it comoves with the background expansion of the Universe. But large-scale structure grows by convergence of

matter on to overdense regions, generating peculiar velocities relative to the background expansion and rendering this assumption inaccurate. Furthermore, on smaller scales, virial motions of galaxies inside clusters generate additional peculiar velocities. The resulting patterns in the observed clustering, which on small scales ( $\lesssim 20$  Mpc) are parallel to the line of sight (‘fingers of God’; Jackson 1972) and on larger scales are transverse (‘Kaiser pancakes’; Kaiser 1987), are termed redshift-space distortions (RSD; Hamilton 1998, for a review).

Despite rendering cosmological parameter inference from galaxy clustering more difficult, RSD contain additional information on the laws of physics and the Universe’s contents. In particular, they scale as  $f = d \ln D / d \ln a \approx \Omega_m^{-1}$ , with  $D$  the linear growth rate,  $a$  the scale factor,  $\Omega_m$  the matter density observed at a particular redshift, and  $\gamma = 0.56$  in General Relativity (GR), but with other values for alternative theories of gravity (Linder 2005). If GR is assumed, then RSD probe the matter density, and if the matter density is assumed, RSD test GR (e.g. Raccanelli et al. 2013).

RSD generate dependence of the observed clustering on angle with respect to the line of sight or lines of sight to a given  $N$ -tuple of galaxies. In the simplest treatment of RSD, which uses linear perturbation theory and assumes a single line of sight to the entire survey (‘flat sky’), RSD produce a quadrupole and hexadecapole ( $\ell = 2$  and 4) in this angle’s cosine in the 2PCF or its Fourier analogue the power spectrum (Kaiser 1987; Hamilton 1993).

\* E-mail: [zslepian@gmail.com](mailto:zslepian@gmail.com), [zslepian@lbl.gov](mailto:zslepian@lbl.gov) (ZS); [deisenstein@cfa.harvard.edu](mailto:deisenstein@cfa.harvard.edu) (DJE)

† Einstein Fellow

The 2PCF and power spectrum multipoles have been used with considerable success to additionally constrain the parameters already probed by the isotropic ( $\ell = 0$ ) 2PCF as well as to measure  $f$  (e.g. Ross et al. 2017; Beutler et al. 2017). Further, the multipoles can be integrated over and summed to produce clustering wedges in angle to the line of sight, offering an alternative route to parameter constraints (Kazin, Sánchez & Blanton 2012; Grieb et al. 2017; Hand et al. 2017). Models for RSD that go beyond linear perturbation theory have also been developed (Taruya, Nishimichi & Saito 2010; Wang, Reid & White 2014; Jeong et al. 2015; Bianchi, Percival & Bel 2016; Vlah, Castorina & White 2016; see White et al. 2015 for comparison of recent models). Considerable work has also been done on wide-angle effects (e.g. Raccanelli et al. 2016; Samushia, Branchini & Percival 2015; Slepian & Eisenstein 2015a; Pápai & Szapudi 2008; Reimberg, Bernardeau & Pitrou 2016).

The situation for the 3PCF, measuring excess triplets of galaxies above random, is more complicated. The RSD now depend on the angles of two triangle sides to the line of sight (the angle of the third will be fixed by these two). Including the three parameters required to describe the triangle itself, there are now five free parameters even in the simplest flat-sky case (Scoccimarro, Couchman & Frieman 1999). In principle, the anisotropic 3PCF contains rich additional information on galaxy biasing and the growth rate, but in practice, it is difficult to measure. Further, it is challenging to report and visualize, as it depends on five parameters.

The isotropic 3PCF already is computationally expensive, scaling as  $N^3$  in the simplest approach, with  $N$  the number of galaxies. To reduce this computational cost, different acceleration schemes have been proposed based on kd-trees and introducing approximations in how the clustering is measured (Zhang & Pen 2005; Gardner et al. 2007; March 2013).

Our own recent work presented a 3PCF algorithm exploiting properties of the multipole basis that we believe is transformatively fast for large-scale 3PCF work, scaling as  $N^2$  rather than  $N^3$  (Slepian & Eisenstein 2015c). The algorithm further accelerates to become order  $N_g \log N_g$ , with  $N_g$  the number of grid points if Fourier transforms (FTs) are used (Slepian & Eisenstein 2016b). We showed that the multipole framework also allows accurate modelling of the covariance on large scales. We applied this framework to the Baryon Oscillation Spectroscopic Survey (BOSS) Data Release (DR) 12 Constant Mass (CMASS) sample of  $\sim 800\,000$  luminous red galaxies to detect the baryon acoustic oscillations (BAO, Slepian et al. 2017; Slepian et al. 2016a; for pedagogical presentation of the BAO physics, see Slepian & Eisenstein 2016c; Eisenstein, Seo & White 2007) as well as to constrain novel forms of biasing (Slepian et al. 2016a,b). Additional work explored the theoretical predictions for the isotropically averaged 3PCF in the multipole basis (Slepian & Eisenstein 2016a) and provided the models fit in these observational works.

In this paper, we extend our 3PCF algorithm to track the line-of-sight dependence of the 3PCF. We show that a simple promotion of the Legendre coefficients relevant for the isotropic 3PCF to mixed spherical harmonic coefficients depending on two total angular momenta  $l$  and  $l'$  and one spin  $m$  fully captures anisotropic clustering. These coefficients can be easily obtained using the same procedure previously developed for the isotropic 3PCF, which centred on obtaining spherical harmonic expansions of the density field on spherical shells around every galaxy in the survey. Further, adding a simple rotation of coordinates so that the position vector of the central galaxy serves as the line of sight to the triplet allows use of a varying line of sight, which tracks the anisotropic clustering more accurately than assuming a single line of sight to the entire

survey. Additionally, as shown in Slepian & Eisenstein (2016b), the spherical harmonic coefficients can be obtained in  $N_g \log N_g$  time using FTs, and so if the density field is gridded, the anisotropic 3PCF algorithm of this paper can be accelerated even further.

In addition to the speed of measurement it enables, the basis advanced in this paper has two other important advantages. First, the parametrization of the 3PCF it involves, as coefficients depending on angular momenta  $l, l'$ , spin  $m$  and triangle side lengths  $r_1, r_2$ , can be easily sliced for analysis and visualization. One might fix the angular momenta and spin and show a colour plot versus  $r_1$  and  $r_2$ , generalizing what was done in Slepian & Eisenstein (2015c), Slepian & Eisenstein (2015b), and Slepian & Eisenstein (2016a) to look at the scale-dependent structure. Alternatively, one could examine the angular structure by fixing  $r_1, r_2$ , and  $l$  and showing the dependence on  $l'$  and  $m$ .

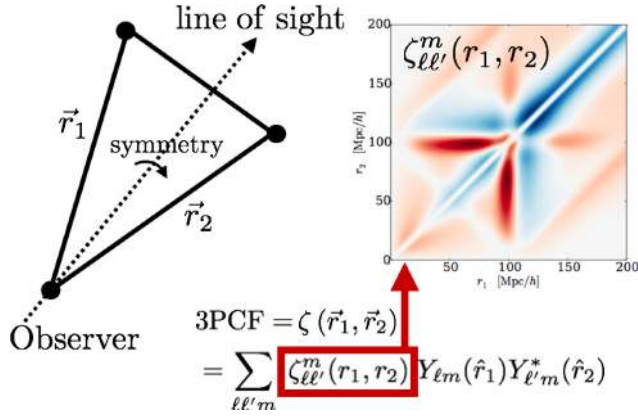
Second, our parametrization permits straightforward handling of the covariance matrix of the anisotropic 3PCF. We extend the work of Slepian & Eisenstein (2015c) to compute the anisotropic covariance matrix assuming a boundary-free survey whose density perturbations follow a Gaussian random field (GRF) with an anisotropic power spectrum given by the Kaiser formula for RSD. We leave the survey volume and the number density as free parameters to be fit from mock catalogues. This calculation supplies an important advance: having a smooth covariance matrix means it can be inverted, a known difficulty for covariance matrices estimated from large numbers of mock catalogues (Percival et al. 2014). In particular, the inverse of the covariance depends on the smallest eigenvalue and to accurately obtain this one requires many mock catalogues per dimension of the covariance matrix. Since the dimension can be large, one often requires the 3PCF for thousands of mocks, adding significant time to any analysis.

In contrast, given our template with only two free parameters, one does not require many mocks to fit for them and obtain a smoothly invertible covariance. The basis we propose permits straightforward computation of this template covariance. In particular, expanding in angular momentum eigenstates (which the spherical harmonics are) means that the angular integrals to bring the covariance from Fourier space (where the GRF calculation is easiest) to configuration space (where the measurement is done) simplify greatly. These integrals, formally 12D, can be reduced to 1D and 2D integrals of the power spectrum, enabling fast evaluation on a grid rather than using more complicated higher dimensional integration techniques.

The paper is laid out as follows. Section 2 presents the basis we will use for our algorithm and shows how it emerges from imposing symmetry about the line of sight (taken to be the  $z$  axis) on the most general representation for the two vectors defining a given triangle. Section 3 outlines the algorithm, showing that only a slight generalization of Slepian & Eisenstein (2015c) is needed to obtain the anisotropic 3PCF and that the speed remains  $\mathcal{O}(N^2)$ . Section 4 describes how a varying line of sight that follows each galaxy triplet can be incorporated. Section 5 discusses edge correction in the algorithm's basis. In Section 6, we compute the covariance of the anisotropic 3PCF in the limited but useful approximation described above. Section 7 concludes and is followed by two brief appendices, one with identities used in the work's main body (Appendix A) and a second showing the impossibility of using a triple Legendre series for the anisotropic 3PCF (Appendix B).

## 2 BASIS

Consider a triplet of galaxies at positions  $\mathbf{x}$ ,  $\mathbf{x} + \mathbf{r}_1$ , and  $\mathbf{x} + \mathbf{r}_2$ . An estimate of the full 3PCF about  $\mathbf{x}$  including any possible dependence



**Figure 1.** Here, we schematically show the approach of this paper. The 3PCF is parametrized by two vectors  $\mathbf{r}_1$  and  $\mathbf{r}_2$ , which encode the triangle's side lengths and its orientation with respect to the line of sight, shown in dashed black. The azimuthal symmetry about the line of sight is the reason only a single spin  $m$  enters the basis. The basis is shown in the red-boxed equation: we expand the anisotropic 3PCF in radial coefficients  $\zeta_{\ell\ell'}^m$  times spherical harmonics in  $\hat{\mathbf{r}}_1$  and  $\hat{\mathbf{r}}_2$ . We seek to measure the radial coefficients and can then plot them at fixed  $l$ ,  $l'$ , and  $m$  versus  $r_1$  and  $r_2$  to reveal the spatial structure. Alternatively, we might also plot the coefficients at fixed  $r_1$ ,  $r_2$ , and  $l$  versus  $l'$  and  $m$  to show the angular structure. The features are BAO, and the plot is adopted from Slepian & Eisenstein (2016a) just to suggest how these plots might look.

on the triangle configuration as well as on its orientation is

$$\hat{\zeta}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{x}) = \frac{4\pi}{\sqrt{(2l+1)(2l'+1)}} \times \sum_{lm} \sum_{l'm'} \hat{\zeta}_{ll'}^{mm'}(r_1, r_2; \mathbf{x}) Y_{lm}(\hat{\mathbf{r}}_1) Y_{l'm'}^*(\hat{\mathbf{r}}_2). \quad (1)$$

The  $Y_{lm}$  are spherical harmonics and the pre-factor is a normalization to recover the 3PCF's expansion into Legendre polynomials in the isotropic limit as in Slepian & Eisenstein (2015c).

Since RSD are due to the difference between how the line of sight and transverse positions are computed from a survey, they must be symmetric under rotations about the line of sight  $\hat{\mathbf{n}}$ . Here, we take  $\hat{\mathbf{n}} = \hat{\mathbf{x}}$ , i.e. that there is a single line of sight to the entire triangle, given by the line of sight to the galaxy at the vertex where the triangle's opening angle is defined. This geometry is shown in Fig. 1, and we further discuss this choice for the line of sight in Section 4. We work in a coordinate system where the  $z$ -axis is along  $\hat{\mathbf{x}}$ . Averaging over rotations around  $\hat{\mathbf{n}} = \hat{\mathbf{x}} = \hat{\mathbf{z}}$ , the azimuthally averaged 3PCF, denoted with subscript 'azi', is

$$\begin{aligned} \hat{\zeta}_{\text{azi}}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{x}) &= \int_0^{2\pi} d\phi \hat{\zeta}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{x}) \\ &= \frac{4\pi}{\sqrt{(2l+1)(2l'+1)}} \sum_{lm} \sum_{l'm'} \hat{\zeta}_{ll'}^{mm'}(r_1, r_2; \mathbf{x}) \\ &\quad \times \int_0^{2\pi} d\phi Y_{lm}(\mathbf{R}_z(\phi)\hat{\mathbf{r}}_1) Y_{l'm'}^*(\mathbf{R}_z(\phi)\hat{\mathbf{r}}_2), \end{aligned} \quad (2)$$

where  $\mathbf{R}_z(\phi)$  represents a rotation by an angle  $\phi$  about  $\hat{\mathbf{z}}$ . The spherical harmonics are defined

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\phi}, \quad (3)$$

so the only  $\phi$  dependence is in the exponential.

We can set the initial azimuthal angle  $\phi_1$  of  $\hat{\mathbf{r}}_1$  to zero, and the initial angle of  $\hat{\mathbf{r}}_2$  we denote  $\phi_2$ . Applying the rotation  $\mathbf{R}_z(\phi)$  then

simply adds an angle  $\phi$  to each azimuthal angle, so the azimuthal average of the spherical harmonics in equation (2) scales as

$$\begin{aligned} &\int_0^{2\pi} d\phi Y_{lm}(\mathbf{R}_z(\phi)\hat{\mathbf{r}}_1) Y_{l'm'}^*(\mathbf{R}_z(\phi)\hat{\mathbf{r}}_2) \\ &\propto \int_0^{2\pi} d\phi e^{im\phi} e^{-im'(\phi_2+\phi)} \\ &\propto \int_0^{2\pi} d\phi e^{i(m-m')\phi} = \delta_{mm'}^K, \end{aligned} \quad (4)$$

where in the second line, we took the exponential in  $\phi_2$  outside the integral and then dropped it.  $\delta_{mm'}^K$  is a Kronecker delta, unity if the subscripted arguments are equal and zero otherwise.

Inserting this result in equation (1), we see that only spherical harmonic combinations where  $m = m'$  can enter the azimuthally averaged 3PCF. We note that  $l$  need not equal  $l'$ , in contrast to the isotropic case where averaging the spherical harmonics over full 3Drotations forces  $l = l'$  as well (Slepian & Eisenstein 2015c). Thus, the azimuthally averaged 3PCF estimate is fully described as

$$\begin{aligned} \hat{\zeta}_{\text{azi}}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{x}) &= \frac{4\pi}{\sqrt{(2l+1)(2l'+1)}} \\ &\quad \times \sum_{l''} \sum_m \hat{\zeta}_{ll''}^m(r_1, r_2; \mathbf{x}) Y_{lm}(\hat{\mathbf{r}}_1) Y_{l''m}^*(\hat{\mathbf{r}}_2); \end{aligned} \quad (5)$$

moving forward, we discuss only the azimuthally averaged 3PCF and so we suppress the subscript 'azi.' Integrating over  $d^3\mathbf{x}$  yields the full 3PCF  $\zeta$  as the translation average of the estimate about a particular galaxy at  $\mathbf{x}$  given by equation (5).

We now discuss an additional symmetry relevant for this basis. We need only compute the coefficients  $\zeta_{ll'}^m$  for  $m \geq 0$  as those for  $m < 0$  are the complex conjugate of the  $m > 0$  coefficients. There are several ways to understand this point.

First, mathematically, we have

$$\zeta_{ll'}^m(r_1, r_2; \mathbf{x}) = \int d\Omega_1 d\Omega_2 \hat{\zeta}(r_1, r_2) Y_{lm}^*(\hat{\mathbf{r}}_1) Y_{l'm}(\hat{\mathbf{r}}_2) \quad (6)$$

and

$$\zeta_{ll'}^{-m}(r_1, r_2; \mathbf{x}) = \int d\Omega_1 d\Omega_2 \hat{\zeta}(r_1, r_2) Y_{l-m}^*(\hat{\mathbf{r}}_1) Y_{l'-m}(\hat{\mathbf{r}}_2) \quad (7)$$

Using the identity that  $Y_{lm}^*(\hat{\mathbf{r}}) = (-1)^m Y_{l-m}(\hat{\mathbf{r}})$ , equation (7) becomes

$$\begin{aligned} \zeta_{ll'}^{-m}(r_1, r_2; \mathbf{x}) &= (-1)^{2m} \int d\Omega_1 d\Omega_2 \hat{\zeta}(r_1, r_2) Y_{lm}(\hat{\mathbf{r}}_1) Y_{l'm}^*(\hat{\mathbf{r}}_2) \\ &= (\zeta_{ll'}^m(r_1, r_2; \mathbf{x}))^* \end{aligned} \quad (8)$$

where the second equality follows by noting that  $(-1)^{2m} = 1$ , conjugating equation (6), and recalling that  $\zeta(\mathbf{r}_1, \mathbf{r}_2; \mathbf{x})$  is real.

Physically, the redundancy of the  $-m$  coefficients occurs because flipping  $m \rightarrow -m$  is equivalent to flipping the azimuthal angle  $\phi \rightarrow -\phi$ , modulo a factor of  $(-1)^m$ .<sup>1</sup> Considering  $\phi$  to be defined as the angle swept out as one moves from the  $x$ -axis towards the  $y$ -axis in the  $xy$ -plane, this transformation is equivalent to flipping  $y \rightarrow -y$ , while keeping the  $x$ - and  $z$ -axes fixed.<sup>2</sup> This flip corresponds to

<sup>1</sup> Flipping  $m$  also affects the associated Legendre polynomial in equation (3) because  $P_l^{-m}(\cos\theta) = (-1)^m [(l-m)!/(l+m)!] P_l^m(\cos\theta)$ , but the factorial piece here cancels the factorials in the spherical harmonic's definition after  $m \rightarrow -m$  is taken there. Consequently, flipping  $m$  in the associated Legendre polynomial only contributes an overall factor of  $(-1)^m$  to the flipped- $m$  spherical harmonic.

<sup>2</sup> This interpretation is easily manifested by writing the spherical harmonics in the Cartesian basis, where they are proportional to powers of  $(x + iy)/r$

reversing the handedness of the coordinate system, as it now satisfies a left-hand rule ( $\hat{x} \times \hat{y} = -\hat{z}$ ) rather than a right-hand rule. We see then that the redundancy of the  $-m$  coefficients reflects the physical symmetry that the galaxy distribution about any point  $\mathbf{x}$  is insensitive to the handedness of the coordinate system one chooses around that point.

As a result of this symmetry, in the sum over  $m$  of equation (5), the negative and positive  $m$  can always be paired to give a real result that is  $2 \operatorname{Re} \zeta_{ll'}^m$ . Consequently, in what follows, we will have this in mind that the final results reported from our approach to the anisotropic 3PCF will be real and symmetrized (denoted by a bar) over positive and negative spins as

$$\bar{\zeta}_{ll'}^m \equiv \zeta_{ll'}^m + \zeta_{ll'}^{-m} (1 - \delta_{m0}^K) = (2 - \delta_{m0}^K) \operatorname{Re} \zeta_{ll'}^m \quad (9)$$

where we take  $m \geq 0$  above. However, throughout the paper, we will often find it convenient to perform analytic calculations in terms of the unsymmetrized  $\zeta_{ll'}^m$  and symmetrize at the final step.

We now discuss the behaviour of  $\zeta_{ll'}^m$  under parity.  $\zeta_{ll'}^m$  behaves as  $(-1)^{l+l'}$  under this transformation. For indistinguishable points, such as galaxies from a single population or survey, the anisotropic 3PCF must be symmetric under parity, meaning  $l + l'$  must be even.  $l + l'$  need not be even for e.g. a 3PCF formed from taking two points from one galaxy population and a third from a different population at higher redshift, as this choice would introduce a preferred orientation.

Finally, in the isotropic limit,  $l' = l$  and the coefficient  $\zeta_{ll'}^m$  becomes  $m$ -independent, i.e.  $\zeta_{ll'}^m \rightarrow \zeta_l$ . We can then sum the spherical harmonics in equation (5) over spins using the spherical harmonic addition theorem (Arfken, Weber & Harris 2013, equation 16.57) to recover that  $\zeta(r_1, r_2; \hat{r}_1 \cdot \hat{r}_2; \mathbf{x}) = \sum_l \zeta_l(r_1, r_2; \mathbf{x}) \mathcal{L}_l(\hat{r}_1 \cdot \hat{r}_2)$  as in Slepian & Eisenstein (2015c), with  $\mathcal{L}_l$  a Legendre polynomial of order  $l$ .

### 3 ALGORITHM

We now show how the anisotropic 3PCF coefficients in the spherical harmonic basis may be obtained in  $\mathcal{O}(N^2)$  time with  $N$  the number of galaxies. Here, we focus on the 3PCF of an arbitrary density field given by  $\delta$ ; in Section 5, we will discuss applying this approach to the anisotropic analogue of the Szapudi & Szalay (1998) estimator for the isotropic 3PCF.

We again begin with a ‘primary’ galaxy at  $\mathbf{x}$  and denote the positions of two ‘secondary’ galaxies as  $\mathbf{x} + \mathbf{r}_1$  and  $\mathbf{x} + \mathbf{r}_2$ . This configuration forms a triangle; we bin its side lengths into radial bins denoted  $r_1$  and  $r_2$ . We denote the radially binned density field

$$\bar{\delta}(r_i; \hat{r}_i; \mathbf{x}) = \int r^2 dr \Phi(r; r_i) \delta(\mathbf{x} + \mathbf{r}), \quad (10)$$

where  $\Phi$  is a binning function that ensures  $r$  is in the bin denoted by  $r_i$ .

We now desire an estimate of the anisotropic binned 3PCF coefficients about  $\mathbf{x}$ . We have

$$\hat{\zeta}_{ll'}^m(r_1, r_2; \mathbf{x}) = \delta(\mathbf{x}) \int d\Omega_1 d\Omega_2 \bar{\delta}(r_1; \hat{r}_1; \mathbf{x}) \times \bar{\delta}(r_2; \hat{r}_2; \mathbf{x}) Y_{lm}(\hat{r}_1) Y_{l'm}^*(\hat{r}_2) \quad (11)$$

for  $m > 0$  and  $(x - iy)/r$  for  $m < 0$ ; this of course just comes from applying Euler’s formula to  $\exp[i m \phi]$  and identifying  $\cos \phi = x/r$  and  $\sin \phi = y/r$ .

It is immediate that this double integral factorizes; defining spherical harmonic coefficients of the binned density as

$$a_{lm}(r_i; \mathbf{x}) = \int d\Omega Y_{lm}^*(\hat{r}) \bar{\delta}(r_i; \hat{r}_i; \mathbf{x}) \quad (12)$$

we may write

$$\hat{\zeta}_{ll'}^m(r_1, r_2; \mathbf{x}) = \delta(\mathbf{x}) a_{lm}(r_1; \mathbf{x}) a_{l'm}^*(r_2; \mathbf{x}). \quad (13)$$

Symmetrizing and using that  $\hat{\zeta}_{ll'}^{-m} = (\hat{\zeta}_{ll'}^m)^*$ , we find that

$$\hat{\zeta}_{ll'}^m(r_1, r_2; \mathbf{x}) = \delta(\mathbf{x}) [a_{lm}(r_1; \mathbf{x}) a_{l'm}^*(r_2; \mathbf{x}) + a_{l'm}^*(r_1; \mathbf{x}) a_{lm}(r_2; \mathbf{x})], \quad (14)$$

which is manifestly symmetric under flipping the conjugate signs, and so our choice to place the conjugate on  $a_{lm}(r_1)$  in equation (13) for the  $\hat{\zeta}_{ll'}^m$  used to construct  $\hat{\zeta}_{ll'}^m$  did not matter.

Importantly, computing the  $a_{lm}$  about a given primary galaxy at  $\mathbf{x}$  on a given radial bin scales as the number of galaxies in that bin. In total if we compute correlations out to a radius  $R_{\max}$ , obtaining the  $a_{lm}$  on all bins scales as  $nV_{\max}$  with  $n$  the survey number density and  $V_{\max}$  the volume of a sphere with radius  $R_{\max}$ . The  $a_{lm}$  must be obtained around every galaxy, so the total work scales as  $N(nV_{\max})$ .

In detail, around each primary, at each  $l$  we have  $l + 1$  distinct  $a_{lm}$ , as we require only the  $m \geq 0$  spherical harmonic coefficients due to the symmetry that  $\hat{\zeta}_{ll'}^{-m} = (\hat{\zeta}_{ll'}^m)^*$  as discussed in Section 2. The total number of coefficients up to  $l_{\max}$  is then  $(l_{\max} + 2)(l_{\max} + 1)/2$ . Each of these must be computed on each bin, so the total number of coefficients to store is  $N_{\text{bins}}(l_{\max} + 2)(l_{\max} + 1)/2$ . We note that the total work of obtaining these scales as  $(nV_{\max})(l_{\max} + 2)(l_{\max} + 1)/2$ , as the first factor accounts for all of the bins out to  $R_{\max}$ .

We now briefly compute the number of combinations of these coefficients that must be formed around each primary. While forming these combinations is a negligible fraction of the total work (we find  $\sim 2$  per cent for the implementation discussed in Section 7), for completeness we discuss the combinatoric computation as it involves several steps. We consider  $l \geq l'$  without loss of generality and focus on  $l > l'$  first. At each  $l$ , we have  $l$  allowed values of  $l'$ , and at each  $l'$ , we have  $l' + 1$  values of  $m$ . We must now sum this over  $l'$  up to  $l$  and then  $l$  from 0 to  $l_{\max}$ . We find

$$\begin{aligned} N_{\text{combs}, l > l'} &= N_{\text{bins}}^2 \sum_{l=0}^{l_{\max}} \sum_{l'=0}^{l-1} (l' + 1) \\ &= \frac{N_{\text{bins}}^2}{2} l_{\max} (l_{\max} + 1)^2, \end{aligned} \quad (15)$$

noting that we include  $N_{\text{bins}}^2$  rather than  $(N_{\text{bins}} + 1)N_{\text{bins}}/2$  because there is no switch symmetry of the bins as  $l \neq l'$ .

For the  $l = l'$  piece, we have just one allowed  $l'$  at each  $l$ , but still  $l + 1$  allowed spins  $m$ , but there is now a switch symmetry between the bins, so we compute

$$\begin{aligned} N_{\text{combs}, l = l'} &= \frac{(N_{\text{bins}} + 1)N_{\text{bins}}}{2} \sum_{l=0}^{l_{\max}} (l + 1) \\ &= \frac{(N_{\text{bins}} + 1)N_{\text{bins}}}{4} (l_{\max} + 1)^2. \end{aligned} \quad (16)$$

The full number of harmonic coefficient combinations required is then the sum of equations (15) and (16).

## 4 DETERMINATION OF THE LINE OF SIGHT

### 4.1 A generalized Yamamoto estimator

In this work, we take the line of sight to a given triangle of galaxies to be the position vector  $\mathbf{x}$  of the galaxy at which the triangle's opening angle is defined. This galaxy is the 'primary,' and serves both to define the origin of coordinates and the  $z$ -axis for computing the spherical harmonic expansion of the radially binned 'secondary' galaxies around it.

Defining the line of sight as the position vector of one triplet member is the natural generalization of the Yamamoto et al. (2006) estimator for the anisotropic 2PCF or power spectrum (further discussed in this latter context in Bianchi et al. 2015; Scoccimarro 2015; Slepian & Eisenstein 2016b). As shown in Slepian & Eisenstein (2015a) for the anisotropic 2PCF, wide-angle effects enter this estimator only at  $\mathcal{O}(\theta^2)$ , where  $\theta$  is the opening angle of the triangle formed by the observer and the galaxy pair.

Further, the Yamamoto estimator differs from using the angle bisector of this triangle or the mid-point of the pair separation only at  $\mathcal{O}(\theta^2)$ . The cancellation of the  $\mathcal{O}(\theta)$  effect occurs because it has odd parity. Thus, when the effect is summed over the two options for the line of the sight to a galaxy pair (the position vectors of the first and the second pair members), it cancels (Slepian & Eisenstein 2015a).

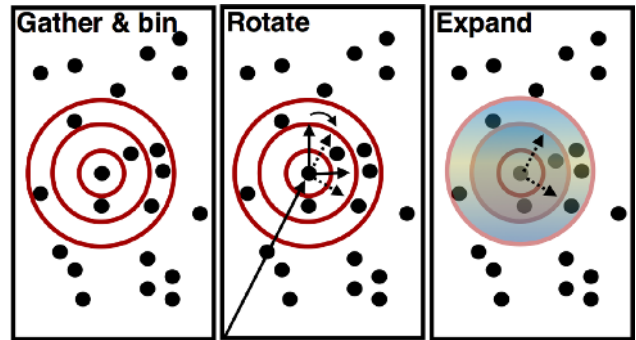
These arguments rely on expanding the anisotropic 2PCF in Legendre polynomials tracking the cosine of the angle between the pair separation and the line of sight. We suspect that in the current work cancellation of the  $\mathcal{O}(\theta)$  error will also occur. In any case, using a rotating line of sight for the anisotropic 3PCF is undoubtedly more accurate than the flat-sky approximation. Finally, for completeness, we note that there has been one other work using a rotating line of sight for three-point clustering: Scoccimarro (2015) presents a rotating line-of-sight estimator for multipole moments of the bispectrum averaged over rotations about one of the wave vectors.

### 4.2 Computing using the rotating line of sight

There are two options for computing the 3PCF using the moving line of sight outlined above. Which option is preferable is likely use-case dependent, so we present both.

The simpler approach is to rotate the coordinates of all the secondaries about a given primary into the frame where the primary lies on the  $z$ -axis. We term this 'pre-rotation.' The computational cost here about a given primary is the number of secondaries,  $nV_{\max}$ , and the rotation must be performed around each of the  $N$  primaries, so the total cost is  $N(nV_{\max}) = \mathcal{O}(N^2)$ , the same scaling for this approach as for the overall 3PCF algorithm. The algorithm flow here is shown in Fig. 2. We note that while the scaling is formally  $\mathcal{O}(N^2)$ , the rotations are still a small amount of work relative to the multipole computation, as we further detail below.

We now outline how the pre-rotation computation proceeds. We rotate the primary and all secondaries to a system where the primary is along the  $z$ -axis, which is the line of sight. We define a triad of unit vectors, with one being the line of sight and the other two orthogonal to it, and then dot them into the separation vectors between the secondaries and the primary. This yields the secondaries' new coordinates in our desired system. This procedure requires just one matrix by vector multiply (nine multiply adds), as compared to the 286 multiply adds required to then obtain the multipole contributions for a given secondary (with  $\ell_{\max} = 10$ ). Thus, pre-rotation



**Figure 2.** Here, we show the algorithm flow around a given primary at a point  $\mathbf{x}$  for the pre-rotation approach. First, we gather all galaxies within  $R_{\max}$  and bin them into spherical shells (also called radial bins). Then we rotate so that the primary lies along the  $z$ -axis. Finally, on each spherical shell we expand the angular dependence of the density into spherical harmonics with coefficients  $a_{lm}(r_i; \mathbf{x})$ , with  $r_i$  designating the radial bin.

only adds 3 per cent more work for a given secondary. Further, it is independent of the number of multipoles and number of radial bins.

We now discuss a second approach to rotation. We observe that rotations will not mix radial bins; a secondary in a given radial bin will be in the same bin post-rotation. Further,  $l_{\max}$  sets the fineness with which the angular structure of the secondaries on a given spherical shell is probed; much as in the cosmic microwave background an angular momentum  $l$  probes angular separations as  $\Delta\theta \sim 180^\circ/l$ . Rotations do not change the underlying angular structure of the secondaries on a shell, meaning that they should not mix information from different  $l$ , and in particular that measuring a given  $l_{\max}$  in any basis offers the same  $l_{\max}$  after rotation. However, the information in the spins  $m$  will mix. Rotating about the  $z$ -axis means a given frequency  $m$  (recall  $Y_{lm} \propto \exp[i m \phi]$ ) in the original system will map to a new frequency  $m'$  in the new, primed system. This mapping occurs because the new plane in which  $\phi'$  is measured will be at some angle to the original plane in which  $\phi$  was measured, and therefore a sinusoid with frequency  $m$  in the original plane will have a new frequency when it is measured in projection. Indeed, in the limiting case of a rotation by  $90^\circ$ , a displacement by  $\Delta\phi$  in the old system would map to a displacement by  $\Delta\phi' = 0$  in the new system, entirely changing the spin structure of the expansion.

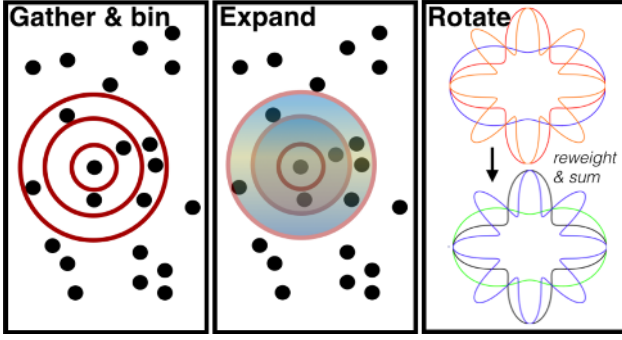
These points suggest that we can reconstruct the spherical harmonic coefficients in our rotated basis from a sum over spins of those measured in any other basis, in particular whatever 'global' basis in which all  $x$ ,  $y$ , and  $z$  galaxy positions in a survey might be specified. In this section only, we adopt the notation superscript 'G' to denote the spherical harmonic coefficients measured in this 'global' basis, and superscript 'L' to denote those in the local basis rotated about a primary at  $\mathbf{x}$  so that  $\mathbf{x}$  is along the  $z$ -axis.

To derive the relation between the 'global' and 'local' coefficients, consider the density field  $\delta$  on a shell denoted by  $r$  about a primary at  $\mathbf{x}$ :

$$\delta(r; \hat{\mathbf{r}}; \mathbf{x}) = \sum_{LM} a_{LM}^G(r; \mathbf{x}) Y_{LM}(\hat{\mathbf{r}}). \quad (17)$$

Under the desired rotation  $\mathbf{R}$ , the density becomes

$$\begin{aligned} \delta(r; \mathbf{R}\hat{\mathbf{r}}; \mathbf{x}) &= \sum_{LM} a_{LM}^G(r; \mathbf{x}) Y_{LM}(\mathbf{R}\hat{\mathbf{r}}) \\ &= \sum_{LM} a_{LM}^G(r; \mathbf{x}) \sum_{M'} D_{MM'}^L(\mathbf{x}) Y_{LM'}(\hat{\mathbf{r}}) \end{aligned} \quad (18)$$



**Figure 3.** Here, we show the algorithm flow when rotation about a given primary is done after the spherical harmonic expansion rather than prior to it, and hence in the space of spherical harmonic coefficients  $a_{lm}$  rather than spatial coordinates. We term this ‘post-rotation.’ The ‘Gather and bin’ step is unchanged from Fig. 2, but we now expand the density’s angular dependence into spherical harmonics on each bin in the second step rather than the third. We then take the  $a_{lm}$  coefficients on a given bin (three are schematically indicated here in the different colours, with colour denoting the coefficient’s amplitude), reweight them as in equation (20, new colours mean new amplitudes), and sum to obtain the new, rotated  $a_{lm}$ .

where in the second line, we expanded the rotated spherical harmonic as a sum over unrotated spherical harmonics;  $D_{MM'}^L$  is a Wigner D-matrix [e.g. Arfken et al. 2013, equation (16.52) or Varshalov, Moskalev & Kersonskii 2013, chapter 4].

Noting that the expansion of the rotated density field in the ‘local’ basis where we desire the coefficients is

$$\delta(r; \mathbf{R}\hat{r}; \mathbf{x}) = \sum_{lm} a_{lm}^L(r; \mathbf{x}) Y_{lm}(\hat{r}), \quad (19)$$

setting the expansions (18) and (19) equal, and invoking orthogonality, we find

$$a_{lm}^L(r; \mathbf{x}) = \sum_M a_{lM}^G(r; \mathbf{x}) D_{Mm}^L(\mathbf{x}). \quad (20)$$

With the relation (20), we can measure the spherical harmonic coefficients in any desired global basis and after finding them around all primaries, recombine locally on each bin about each primary weighted by the Wigner D-matrices. Since the rotations occur after the spherical harmonic coefficients are computed, we term this approach ‘post-rotation.’

The scaling of this post-rotation approach differs from that of the pre-rotation approach outlined earlier, as we now show. We need only obtain  $m \geq 0$  spherical harmonic coefficients as noted in Section 3, and for the same reason on the right-hand side we need only  $a_{lM}$  with  $M \geq 0$ . Further, the D-matrices for  $M < 0$  can be related to those for  $M > 0$ , and using symmetry properties of the D-matrix, we can also ensure that  $m \leq M$ . Thus the entire right-hand side of equation (20) can be cast in terms of  $M \geq 0$  spherical harmonic coefficients and D-matrices and using only D-matrices with  $m \leq M$ . At each  $l$ , we then have  $(l+1)(l+2)/2$  D-matrix elements to compute, leading to 286 matrix elements for  $l_{\max} = 10$ . The algorithm flow for this ‘post-rotation’ approach is shown in Fig. 3.

A cross-check on this result is that all of the spherical harmonics up to  $l_{\max} = 10$  can be computed using 286 combinations of powers of  $x/r$ ,  $y/r$ , and  $z/r$  where  $x$ ,  $y$ , and  $z$  are the relative coordinates of a secondary galaxy in the frame where the primary is at the origin (see Slepian & Eisenstein 2015c). The behaviour of these 286 fundamental power combinations under rotation must completely

determine the rotated spherical harmonic expansion, confirming that we should need 286 matrix elements to perform the rotation.

The computational cost of evaluating the Wigner D-matrices is small for the modest  $l$  we require. Direct evaluation is one option; for instance Varshalov et al. (2013) give easily implemented expressions in terms of Gauss’s hypergeometric function  ${}_2F_1$ . There also exist more efficient methods for their computation, such as the use of recursion relations or pseudo-spectral projection (primarily important going to high  $l$ , Varshalov et al. 2013; Gimbutas & Greengard 2009; Gumerov & Duraiswami 2014; Feng et al. 2015).

We have written  $D_{Mm}^L(\mathbf{x})$  as a function of the primary location  $\mathbf{x}$ . In detail, following the conventions of Varshalov et al. (2013), the D-matrix as a function of the Euler angles  $\alpha = 0$ ,  $\beta = -\theta(\mathbf{x})$ , and  $\gamma = -\phi(\mathbf{x})$  is

$$D_{MM'}^J(\alpha, \beta, \gamma) = e^{-iM\alpha} d_{MM'}^J(\beta) e^{-iM'\gamma}, \quad (21)$$

where  $d_{MM'}^J$  is a little-d matrix (see Varshalov et al. 2013).

In contrast to pre-rotation, for post-rotation, the total computational cost is independent of the number of secondaries about a given primary, but depends on the number of multipoles and number of radial bins, scaling as  $l_{\max}^3 N_{\text{bins}}$ .

### 4.3 Using Fourier transforms for the harmonic coefficients

As shown in Slepian & Eisenstein (2016b), the spherical harmonic coefficients about a given primary galaxy are simply a convolution integral, and so fast Fourier transforms (FFTs) can be used to obtain the coefficients about all primaries in the survey at once scaling as  $N_g \log N_g$ , with  $N_g$  the number of grid points, if the galaxy density field is gridded. While the fineness of grid required will be somewhat application-dependent, there are likely contexts (such as analysis of a large number of mocks for verifying the pipeline or covariance matrix) where a possible loss in precision will be offset by the increase in speed FFTs afford.

Importantly, the post-rotation approach outlined in Section 4.2 is essential for enabling use of FTs to obtain the spherical harmonics. The pre-rotation approach relies on sitting on a given primary galaxy and rotating using its coordinates, which is not possible in an FFT-based approach. However, once the spherical harmonic coefficients are known about each primary galaxy for any given choice of coordinates, as shown in Section 4.2, they can be rotated on a primary-by-primary basis with little additional computation. Thus, the coefficients can be obtained in some arbitrary global frame by FTs and then post-processed to give the desired identification of the local  $z$ -axis with the line of sight to each primary galaxy.

## 5 EDGE CORRECTION

We now discuss edge correction in the spherical harmonic basis. The estimator for the full 3PCF is (Szapudi & Szalay 1998)

$$\hat{\xi} = \frac{NNN}{RRR}, \quad (22)$$

where  $N \equiv D - R$  is the data minus random count with  $D$  indicating data and  $R$  indicating random. The interpretation of this estimator as an optimal weighting in the shot-noise limit and how it should be used to obtain the 3PCF is further discussed in Slepian & Eisenstein (2015c).

### 5.1 Edge correction in the spherical harmonic basis

Here, we show how to implement edge correction using our algorithm. We will begin with all spins, including negative ones, and then show how to cast the edge correction in terms of solely positive semi-definite spins so that we may work with the symmetrized 3PCF coefficients  $\hat{\zeta}_{ll'}^m$ . We multiply each side of the estimator (22) by  $RRR$  and then expand the result into spherical harmonics, obtaining

$$\sum_{jj's} \mathcal{N}_{jj'}^s Y_{js}(\hat{r}_1) Y_{j's}^*(\hat{r}_2) = \sum_{ll'm} \sum_{kk'p} \hat{\zeta}_{ll'}^m \mathcal{R}_{kk'}^p Y_{lm}(\hat{r}_1) \times Y_{kp}(\hat{r}_1) Y_{l'm}^*(\hat{r}_2) Y_{k'p}^*(\hat{r}_2), \quad (23)$$

where the  $\mathcal{N}_{jj'}^s$  are the double spherical harmonic coefficients of  $N$ ,  $\mathcal{R}_{kk'}^p$  those for  $R$ , and  $\hat{\zeta}_{ll'}^m$  those for  $\hat{\zeta}$ . Integrating both sides against  $Y_{js}^*(\hat{r}_1) Y_{j's}(\hat{r}_2)$  and invoking orthogonality, we find

$$\mathcal{N}_{jj'}^s = \sum_{ll'm} \sum_{kk'p} \hat{\zeta}_{ll'}^m \mathcal{R}_{kk'}^p \mathcal{G}_{lkj}^{mp-s} \mathcal{G}_{l'k'j'}^{-m-ps} \quad (24)$$

where  $\mathcal{G}$  is the Gaunt integral, i.e. the integral of three spherical harmonics with the indicated total angular momenta (subscripts) and spins (superscripts); we write it in terms of 3j-symbols in Appendix A. We note that if  $k = 0 = k'$ , then  $l = j$  and  $l' = j'$ : if the randoms have no angular structure, then the coefficients  $\mathcal{N}_{jj'}^s$  are the desired coefficients  $\hat{\zeta}$  of the 3PCF up to normalization. In this case, the edge correction was simply a division by the average background count to convert the galaxy density field into an overdensity field. We further observe that  $k + k'$  and  $j + j'$  must have the same parity, because the Gaunt integrals enforce that  $l + k + j$  and  $l' + k' + j'$  are even, and  $l + l'$  is even, so  $k + k' + j + j'$  is even as well. This rule somewhat reduces the number of random coefficients  $\mathcal{R}$  required to obtain a given 3PCF coefficient from the measured difference field  $\mathcal{N} = N - R$ . Unlike the 3PCF coefficients, the random coefficients need not be symmetric under parity because the survey geometry may not be. Hence, the difference field coefficients need not be either.

We now show how equation (24) can be recast in terms of the symmetrized quantities our algorithm tracks. We first rewrite each coefficient separated into its real and imaginary parts, giving

$$\mathcal{N}_{jj'}^{s,R} + i \mathcal{N}_{jj'}^{s,I} = \sum_{ll'm} \sum_{kk'p} \left[ \hat{\zeta}_{ll'}^{m,R} + i \hat{\zeta}_{ll'}^{m,I} \right] \left[ \mathcal{R}_{kk'}^{p,R} + i \mathcal{R}_{kk'}^{p,I} \right] \times \mathcal{G}_{lkj}^{mp-s} \mathcal{G}_{l'k'j'}^{-m-ps} \quad (25)$$

where superscript R denotes the real part and superscript I the imaginary part. Inverting equation (9), we may replace the real part of each coefficient in equation (25) in terms of the symmetrized coefficient, yielding

$$\frac{1}{2 - \delta_{s0}^K} \bar{\mathcal{N}}_{jj'}^{|s|} + i \mathcal{N}_{jj'}^{s,I} = \sum_{ll'm} \sum_{kk'p} \left[ \frac{1}{2 - \delta_{m0}^K} \hat{\zeta}_{ll'}^{|m|} + i \hat{\zeta}_{ll'}^{m,I} \right] \times \left[ \frac{1}{2 - \delta_{p0}^K} \bar{\mathcal{R}}_{kk'}^{|p|} + i \mathcal{R}_{kk'}^{p,I} \right] \mathcal{G}_{lkj}^{mp-s} \mathcal{G}_{l'k'j'}^{-m-ps}; \quad (26)$$

the absolute value signs on the spins are necessary because their range includes negative values here but the symmetrized coefficients are defined only for positive or zero spins.

We seek  $\hat{\zeta}_{ll'}^{|m|}$  and would like to estimate it solely in terms of  $\bar{\mathcal{N}}_{jj'}^{|s|}$  and  $\bar{\mathcal{R}}_{kk'}^{|p|}$ . We see that if we set the imaginary part of the randoms' unsymmetrized coefficient to zero (indeed, we do not wish to track it in our algorithm), this ensures the real part of the right-hand

side of equation (26) involves only our desired  $\hat{\zeta}_{ll'}^{|m|}$ . In contrast, the imaginary part remaining on the right-hand side will involve  $\hat{\zeta}_{ll'}^{m,I} \bar{\mathcal{R}}_{kk'}^{|p|}$  and thus has no information about  $\hat{\zeta}_{ll'}^{|m|}$ . We may now take the real part of the whole equation, yielding

$$\frac{1}{2 - \delta_{s0}^K} \bar{\mathcal{N}}_{jj'}^{|s|} = \sum_{ll'm} \sum_{kk'p} \frac{1}{2 - \delta_{m0}^K} \hat{\zeta}_{ll'}^{|m|} \times \frac{1}{2 - \delta_{p0}^K} \bar{\mathcal{R}}_{kk'}^{|p|} \mathcal{G}_{lkj}^{mp-s} \mathcal{G}_{l'k'j'}^{-m-ps}; \quad (27)$$

Following Slepian & Eisenstein (2015c), we now divide both sides of equation (27) through by  $\bar{\mathcal{R}}_{00}^0$  and separate off the  $kk'p = 000$  term on the right-hand side, finding

$$\frac{\bar{\mathcal{N}}_{jj'}^{|s|}}{(2 - \delta_{s0}^K) \bar{\mathcal{R}}_{00}^0} = \frac{1}{2 - \delta_{s0}^K} \hat{\zeta}_{jj'}^{|s|} + \sum_{ll'm} \frac{1}{2 - \delta_{m0}^K} \hat{\zeta}_{ll'}^{|m|} \times \sum_{kk'p \neq 000} \frac{1}{2 - \delta_{p0}^K} \bar{f}_{kk'}^{|p|} \mathcal{G}_{lkj}^{mp-s} \mathcal{G}_{l'k'j'}^{-m-ps} \quad (28)$$

where  $\bar{f}_{kk'}^{|p|} \equiv \bar{\mathcal{R}}_{kk'}^{|p|} / \bar{\mathcal{R}}_{00}^0$ .

In the isotropic case,  $l = l'$  and  $\bar{\zeta}_{jj'}^{|s|}$  and  $\bar{\zeta}_{ll'}^{|m|}$  become spin-independent. Thus, in this limit, we can sum over all spins to estimate the isotropic 3PCF  $\zeta_l$  of Slepian & Eisenstein (2015c). The only factors inside the sums over  $m$  and  $s$  on the right-hand side then are 3j-symbols, and invoking the orthogonality identity (Olver et al. 2010, 34.3.16), we find that  $k = k'$ . We then sum the resultant  $f_{kk}^p$  over  $p$  weighted by the Gaunt integrals. This sum can be identified with the edge-correction factors  $f_l'$  of Slepian & Eisenstein (2015c), showing that only isotropic edge-correction factors enter an estimate of the isotropic 3PCF.

We now define an edge correction matrix  $\mathbf{M}$  with elements

$$M_{ll'm}^{jj's} = \sum_{kk' \neq 00} \bar{f}_{kk'}^{|p|} \mathcal{G}_{lkj}^{mp-s} \mathcal{G}_{l'k'j'}^{-m-ps}, \quad p = s - m, \quad (29)$$

using that the Gaunt symbol forces  $m + p - s = 0$  to set  $p$ . In terms of this matrix, the edge-correction equation (28) becomes

$$\frac{\bar{\mathcal{N}}}{\bar{\mathcal{R}}_{00}^0} = (\mathbf{I} + \mathbf{M}) \hat{\zeta} \quad (30)$$

where  $\bar{\mathcal{N}}$  is a vector of  $\bar{\mathcal{N}}_{jj'}^{|s|}$ , the double spherical harmonic moments of the counts,  $\hat{\zeta}$  is a vector of the double spherical harmonic moments  $\hat{\zeta}_{ll'}^{|m|}$ , and  $\mathbf{I}$  is the identity matrix. Given the coefficients  $\bar{f}_{kk'}^{|p|}$ , which encode the impact of the survey geometry on the randoms, and  $\bar{\mathcal{R}}_{00}^0$ , the overall normalization of the random triple count, this equation can be solved by matrix inversion for the desired  $\hat{\zeta}$ .

We note that there are two approximations involved in solving the edge correction equation here. First, to obtain a given matrix element, one formally requires all values of the random coefficients  $\mathcal{R}$  to perform the sum in equation (29). However, in practice, as the next subsection will illustrate, these values fall so quickly with rising  $k$  and  $k'$  that we expect the full matrix element is very well approximated by truncating at some  $k$  and  $k'$  of order the maximal  $l$  and  $l'$  to which an anisotropic 3PCF measurement is desired.

Second, the edge correction matrix is formally infinite in dimension, as at fixed  $l$  and  $l'$ , all  $j$  and  $j'$  can enter the correction. Thus, one should also measure an infinite set of coefficients  $\bar{\mathcal{N}}_{jj'}^{|s|}$  of the  $D - R$  field. However, again, in practice, truncating this set at roughly the maximal  $l$  or  $l'$  desired for the 3PCF measurement should be sufficiently accurate to correct for the survey geometry. This point is



detailed further in Slepian & Eisenstein (2015c) for the isotropic 3PCF, and the same outcome is expected for the anisotropic 3PCF.

Summarizing, the rapid fall-off of the edge correction factors means that truncating the sums required for each matrix element is a good approximation, and the expected near-diagonality of the matrix means that the truncation of the matrix itself will not greatly affect the inverse.

## 5.2 A toy model for the edge correction factors

Here, we investigate a toy model of the survey geometry to gain intuition for the values of the edge correction factors. We consider a planar survey boundary perpendicular to the line of sight (i.e. the  $z$ -axis). For a given primary galaxy, the sphere around the primary over which we compute the 3PCF will impinge upon the survey boundary if the primary is less than  $R_{\max}$  away from it. These galaxies will be only a small fraction of the total for realistic survey geometries, so the edge-correction factors we estimate from this model will be diluted by the bulk of the survey where a primary's sphere is fully contained, and we quantify this dilution at the end of the section.

For now we focus on primaries whose surrounding sphere does impinge on the survey boundary. We take it that a given primary is a distance  $z$  away from the survey boundary. This model was explored in Slepian & Eisenstein (2015c), but only the isotropic edge correction factors  $f_l = 4\pi/(2l+1) \sum_m f_{ll}^m$  were computed there, as they were the only factors relevant for correcting the isotropic 3PCF. Here, we seek the more general

$$f_{kk'}^0 \equiv \frac{\mathcal{R}_{kk'}^0}{\mathcal{R}_{00}^0} = \frac{\langle a_{k0} a_{k'0}^* \rangle}{a_{00}^2} \quad (31)$$

and angle brackets represent azimuthal averaging. We note that the azimuthal symmetry of our toy model means the only non-zero edge correction factors are those with zero spin.

For a secondary galaxy  $R_{\max}$  distant from the primary, there will be a critical angle  $\mu_c = z/R_{\max}$  for which it is outside the survey for larger  $\mu$ . As shown in Slepian & Eisenstein (2015c, equation 35), the spherical harmonic coefficients are then

$$a_{l0} = \sqrt{\frac{\pi}{2l+1}} [\mathcal{L}_{l+1}(\mu_c) - \mathcal{L}_{l-1}(\mu_c)] \quad (32)$$

for  $l \geq 1$ . We note that there is no dependence on the triangle side length because these coefficients represent a random density field. We also note that since the only relevant coefficients for this toy model are at spin zero, the symmetrization discussed in Section 2 does not affect our results here.

To obtain  $f_{kk'}^0$ , we now form the product of the spherical harmonic coefficients, equation (31) demands and average over  $\mu_c$ , as

$$\begin{aligned} \langle a_{k0} a_{k'0}^* \rangle &= \frac{\pi}{\sqrt{(2k+1)(2k'+1)}} \\ &\times \int_0^1 d\mu_c [\mathcal{L}_{k+1}(\mu_c) \mathcal{L}_{k'+1}(\mu_c) - \mathcal{L}_{k-1}(\mu_c) \mathcal{L}_{k'+1}(\mu_c) \\ &- \mathcal{L}_{k+1}(\mu_c) \mathcal{L}_{k'-1}(\mu_c) + \mathcal{L}_{k-1}(\mu_c) \mathcal{L}_{k'-1}(\mu_c)]. \quad (33) \end{aligned}$$

We may perform this integral by using the linearization formula for Legendre polynomials given in Appendix A to convert the products of Legendres into sums over a single Legendre and then integrating

using the recursion relation also given in Appendix A. We find

$$\begin{aligned} \langle a_{k0} a_{k'0}^* \rangle &= \frac{\pi}{\sqrt{(2k+1)(2k'+1)}} \sum_J \left[ \begin{pmatrix} k+1 & k'+1 & J \\ 0 & 0 & 0 \end{pmatrix}^2 \right. \\ &- \begin{pmatrix} k-1 & k'+1 & J \\ 0 & 0 & 0 \end{pmatrix}^2 - \begin{pmatrix} k+1 & k'-1 & J \\ 0 & 0 & 0 \end{pmatrix}^2 \\ &\left. + \begin{pmatrix} k-1 & k'-1 & J \\ 0 & 0 & 0 \end{pmatrix}^2 \right] [\mathcal{L}_{J-1}(0) - \mathcal{L}_{J+1}(0)]. \quad (34) \end{aligned}$$

To simplify, we observed that  $\mathcal{L}_J(1) = 1$  for all  $J$  and so the terms from the upper bound of the integral (33) vanish.<sup>3</sup> We note that  $J$  must be odd or the difference of Legendres in the last line of equation (34) vanishes, as  $\mathcal{L}_n(0) = 0$  for  $n$  odd by parity.

We pause to observe that in the limit where  $k' \simeq k$  and  $k \rightarrow \infty$ ,  $J$  is constrained to be small ( $J \leq |k - k'|$ ). We can then use the square of the 3j-symbol's asymptotic, which is  $\simeq d^2/(2k+1)$ , where  $d$  is a Wigner matrix and  $|d^2| \leq 1$ . Combining this with the pre-factor  $\propto 1/(2k+1)$ , we see that the limit of the isotropic edge correction factors scales as  $1/(2k+1)^2$ , going to zero as  $k$  grows. This roughly recovers the limiting behaviour of equation (38) for the isotropic edge correction factor in Slepian & Eisenstein (2015c).

From explicit computation in Slepian & Eisenstein (2015c), we have that  $\mathcal{R}_{00}^0 = (a_{00})^2 = (7\pi)/3$ , so we may form the ratio  $f_{kk'}^0$  of equation (31).<sup>4</sup> We compute a number of values to show that the edge correction factors are indeed small:  $f_{10}^0 = -9.28$  per cent,  $f_{11}^0 = 17.14$  per cent,  $f_{20}^0 = -6.39$  per cent,  $f_{12}^0 = 6.92$  per cent,  $f_{30}^0 = -2.36$  per cent,  $f_{31}^0 = -1.87$  per cent,  $f_{22}^0 = 4.08$  per cent,  $f_{41}^0 = -2.90$  per cent,  $f_{32}^0 = 0.66$  per cent,  $f_{50}^0 = 0.37$  per cent,  $f_{42}^0 = -0.91$  per cent, and  $f_{33}^0 = 1.90$  per cent. The factors where  $k = k'$  correspond to the  $f_k$  of Slepian & Eisenstein (2015c), and we recover the same values here as in that work. We point out that the edge correction factors in this toy model do not satisfy the constraint that  $k + k'$  is even; our survey boundary is a plane perpendicular to the  $z$ -axis and so the random clustering is not invariant under parity.

In a Sloan Digital Sky Survey (SDSS) Baryon Oscillation Spectroscopic Survey (BOSS)-like geometry, only 20 per cent of galaxies lie within  $R_{\max} = 200$  Mpc of a survey boundary, so these factors should be further scaled down by roughly a factor of 5. On the other hand, the true survey geometry is more complicated than our simple planar toy model, so the edge correction factors estimated above should be taken only as a rough guide.

Finally, we note that the calculation of this section could also be applied to estimate the edge-correction factors' values for an angular survey boundary. For this situation, the boundary has rotated  $90^\circ$  from the planar redshift boundary explored above, or equivalently, we can consider that the  $z$ -axis of coordinates has rotated by this amount. Thus, we can compute the harmonic coefficients of this angular boundary by rotating all of the redshift-boundary coefficients prior to squaring and taking their expectation value over  $\mu_c$ . This rotation can be accomplished using Wigner D-matrices. Since the D-matrices are rotations, and thus unitary, we expect that an angular boundary would not lead to significantly different amplitudes of the

<sup>3</sup>This sum of four 3j-symbols can be further simplified using recursions 8.6.4.21 and 8.6.4.23 in Varshalov et al. (2013) to a pre-factor times a single 3j-symbol if desired.

<sup>4</sup>We use the first line of equation (35) of Slepian & Eisenstein (2015c) with  $l = 0$  and then average over  $\mu_c$ .

edge correction factors from those in the redshift-boundary case; thus our overall claim that these factors should be small still holds.

## 6 COVARIANCE OF THE ANISOTROPIC 3PCF

### 6.1 Adapting the real-space calculation to redshift space

Here, we adapt the isotropic 3PCF covariance calculation of Slepian & Eisenstein (2015c) for the anisotropic case. We use a tilde to denote a Fourier-space quantity, use a negative exponential for inverse FTs, and always use  $d^3k/(2\pi)^3$  when going from Fourier space to configuration space. Adapting Slepian & Eisenstein (2015c, equation 45) by replacing the real-space density perturbation  $\delta$  by its redshift-space analogue  $\delta_s$ , we see that in equation (47) of that work, the Fourier-space density perturbations  $\tilde{\delta}$  can be replaced by their redshift-space analogues  $\tilde{\delta}_s(\mathbf{k})$ .

Following Slepian & Eisenstein (2015c), we use Wick's Theorem to contract the six Fourier-space density fields implied by the 3PCF covariance. We then adapt equation (49) of that work by replacing the isotropic power spectra with their multipole analogues,

$$P(k; \mu) = \sum_k P_{l_k}(k) \mathcal{L}_{l_k}(\mu) \quad (35)$$

We emphasize that this expansion for the anisotropic power spectrum is fully general, though in the flat-sky ('Kaiser') approximation and under linear theory, the series reduces further to have terms only at  $l = 0, 2$ , and  $4$ , as we will discuss further in Section 6.5. Following through to the analogues of Slepian & Eisenstein (2015c, equations 50 and 51), we find

$$\begin{aligned} \text{Cov} \equiv \langle \hat{\xi}(\mathbf{r}_1, \mathbf{r}_2) \hat{\xi}(\mathbf{r}'_1, \mathbf{r}'_2) \rangle &= \frac{1}{V} \int \frac{d^3q d^3p d^3k}{(2\pi)^9} \\ &\times \sum_{l_q l_p l_k} P_{l_q}(q) P_{l_p}(p) P_{l_k}(k) \mathcal{L}_{l_q}(\mu_q) \mathcal{L}_{l_p}(\mu_p) \mathcal{L}_{l_k}(\mu_k) \\ &\times (2\pi)^3 \delta_D^{[3]}(\mathbf{q} + \mathbf{p} + \mathbf{k}) e^{-i[\mathbf{q} \cdot \mathbf{r}_1 + \mathbf{p} \cdot \mathbf{r}_2]} \\ &\times \left\{ e^{-i[\mathbf{q} \cdot \mathbf{r}'_1 + \mathbf{p} \cdot \mathbf{r}'_2]} + e^{-i[\mathbf{p} \cdot \mathbf{r}'_1 + \mathbf{q} \cdot \mathbf{r}'_2]} + e^{-i[\mathbf{k} \cdot \mathbf{r}'_1 + \mathbf{p} \cdot \mathbf{r}'_2]} \right. \\ &\left. + e^{-i[\mathbf{p} \cdot \mathbf{r}'_1 + \mathbf{k} \cdot \mathbf{r}'_2]} + e^{-i[\mathbf{k} \cdot \mathbf{r}'_1 + \mathbf{q} \cdot \mathbf{r}'_2]} + e^{-i[\mathbf{q} \cdot \mathbf{r}'_1 + \mathbf{k} \cdot \mathbf{r}'_2]} \right\}. \quad (36) \end{aligned}$$

### 6.2 Projection on to our basis

We now show how to obtain the covariance of our symmetrized harmonic coefficients for the 3PCF from the full covariance. We will begin by projecting the covariance on to the full spherical harmonic basis including negative spins and at the end of the calculation show how to obtain the covariance of the symmetrized coefficients from these results.

The covariance projected on to our full spherical harmonic basis is

$$\begin{aligned} \text{Cov}_{l_1 l_2 m, l'_1 l'_2 m'}(r_1, r_2; r'_1, r'_2) &= \frac{\sqrt{(2l_1 + 1)(2l_2 + 1)(2l'_1 + 1)(2l'_2 + 1)}}{(4\pi)^2} \\ &\times \int d\Omega_{r_1} d\Omega_{r_2} d\Omega_{r'_1} d\Omega_{r'_2} Y_{l_1 m}^*(\hat{r}_1) Y_{l_2 m}(\hat{r}_2) \\ &\times Y_{l'_1 m'}^*(\hat{r}'_1) Y_{l'_2 m'}(\hat{r}'_2) \text{Cov}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2). \quad (37) \end{aligned}$$

The pre-factor comes from two copies of the inverse of the pre-factor in equation (1), as we wish to extract the covariance of the coefficients  $\zeta_{ll'}^m$ . Noticing that the only dependence on real-space

variables in the full covariance is in the exponentials, we require the projection integrals

$$\int d\Omega_r e^{-i\mathbf{k} \cdot \mathbf{r}} Y_{lm}^*(\hat{r}) = (4\pi)(-i)^l j_l(kr) Y_{lm}^*(\hat{k}) \quad (38)$$

and

$$\int d\Omega_r e^{-i\mathbf{k} \cdot \mathbf{r}} Y_{lm}(\hat{r}) = (4\pi)(-i)^l j_l(kr) Y_{lm}(\hat{k}) \quad (39)$$

where we used the plane wave expansion (Arfken et al. 2013 equation 16.61) to expand the exponential into spherical Bessel functions and spherical harmonics and then invoked orthogonality.

The projection of all of the exponentials in equation (36) is then

$$\begin{aligned} \mathcal{E}_{l_1 l_2 m, l'_1 l'_2 m'} &\equiv \sqrt{(2l_1 + 1)(2l_2 + 1)(2l'_1 + 1)(2l'_2 + 1)} \\ &\times (-i)^{l_1 + l_2 + l'_1 + l'_2} \mathcal{J}_{l_1 l_2}(qr_1; pr_2) Y_{l_1 m}^*(\hat{q}) Y_{l_2 m}(\hat{p}) \\ &\times \{ \mathcal{J}_{l'_1 l'_2}(qr'_1; pr'_2) Y_{l'_1 m'}^*(\hat{q}) Y_{l'_2 m'}(\hat{p}) \\ &+ \mathcal{J}_{l'_1 l'_2}(pr'_1; qr'_2) Y_{l'_1 m'}^*(\hat{p}) Y_{l'_2 m'}(\hat{q}) \\ &+ \mathcal{J}_{l'_1 l'_2}(kr'_1; pr'_2) Y_{l'_1 m'}^*(\hat{k}) Y_{l'_2 m'}(\hat{p}) \\ &+ \mathcal{J}_{l'_1 l'_2}(pr'_1; kr'_2) Y_{l'_1 m'}^*(\hat{p}) Y_{l'_2 m'}(\hat{k}) \\ &+ \mathcal{J}_{l'_1 l'_2}(kr'_1; qr'_2) Y_{l'_1 m'}^*(\hat{k}) Y_{l'_2 m'}(\hat{q}) \\ &+ \mathcal{J}_{l'_1 l'_2}(qr'_1; kr'_2) Y_{l'_1 m'}^*(\hat{q}) Y_{l'_2 m'}(\hat{k}) \}, \quad (40) \end{aligned}$$

where we have used equation (38) and simplified. We have defined

$$\mathcal{J}_{l_1 l_2}(qr_1; pr_2) \equiv j_{l_1}(qr_1) j_{l_2}(pr_2). \quad (41)$$

The symmetry structure of these terms can be easily checked: within the curly brackets, the first and second terms are equal under  $q \leftrightarrow p$ , the third and fourth under  $k \leftrightarrow p$ , and the fifth and sixth under  $k \leftrightarrow q$ .

Returning to equation (36) and using Slepian & Eisenstein (2015c, equation 58) to expand the Dirac delta function into an integral over plane waves and thence into spherical harmonics and spherical Bessel functions, and the spherical harmonic addition theorem to expand the Legendre polynomials (Arfken et al. 2013 equation 16.57), the projected covariance is

$$\begin{aligned} \text{Cov}_{l_1 l_2 m, l'_1 l'_2 m'}(r_1, r_2; r'_1, r'_2) &= \frac{1}{V} \int \frac{d^3q d^3p d^3k}{(2\pi)^9} \sum_{l_q l_p l_k} P_{l_q}(q) P_{l_p}(p) P_{l_k}(k) \\ &\times \frac{(4\pi)^3}{(2l_q + 1)(2l_p + 1)(2l_k + 1)} \\ &\times \sqrt{\frac{(2l_q + 1)(2l_p + 1)(2l_k + 1)}{(4\pi)^3}} \\ &\times Y_{l_q 0}(\hat{q}) Y_{l_p 0}(\hat{p}) Y_{l_k 0}(\hat{k}) \\ &\times (4\pi)^3 \sum_{J_1 J_2 J_3} \sum_{S_1 S_2 S_3} \mathcal{D}_{J_1 J_2 J_3} \mathcal{C}_{J_1 J_2 J_3} \mathcal{R}_{J_1 J_2 J_3}(q, p, k) \\ &\times \begin{pmatrix} J_1 & J_2 & J_3 \\ S_1 & S_2 & S_3 \end{pmatrix} Y_{J_1 S_1}^*(\hat{q}) Y_{J_2 S_2}^*(\hat{p}) Y_{J_3 S_3}^*(\hat{k}) \\ &\times \mathcal{E}_{l_1 l_2 m, l'_1 l'_2 m'}. \quad (42) \end{aligned}$$

The first factor in the third line stems from the spherical harmonic addition theorem applied to the three Legendre polynomials entering the power spectrum multipoles. The second factor comes from the spin-zero values of the spherical harmonics of the line of sight (taken to be the  $z$ -axis); these are the only spherical harmonics entering the power spectrum multipole decomposition. The fourth line comes from the spherical harmonic expansion of the power spectrum's multipole moments, and the fifth and sixth lines come from the Dirac delta function's expansion into spherical harmonics, with

$$\begin{aligned} \mathcal{D}_{J_1 J_2 J_3} &\equiv i^{J_1+J_2+J_3} \\ \mathcal{C}_{J_1 J_2 J_3} &\equiv \sqrt{\frac{(2J_1+1)(2J_2+1)(2J_3+1)}{4\pi}} \\ \mathcal{R}_{J_1 J_2 J_3}(q, p, k) &\equiv \int r^2 dr j_{J_1}(qr) j_{J_2}(pr) j_{J_3}(kr). \end{aligned} \quad (43)$$

For each Fourier-space unit vector in equation (42), there is one spherical harmonic contributed by the expansion of the Legendre polynomial from the power spectrum multipoles, another from the Dirac delta function, and either zero, one, or two from  $\mathcal{E}_{\text{proj}, l_1 l_2 m, l'_1 l'_2 m'}$ . Thus, when we perform the integrations over angles  $d\Omega_q d\Omega_p d\Omega_k$ , we will have integrals of two, three, or four spherical harmonics.

In particular, the first two terms in the curly brackets in equation (40) will lead to four spherical harmonics in  $\hat{p}$  and  $\hat{q}$  and two in  $\hat{k}$ ; the final four terms in the curly brackets will give four in  $\hat{p}$  and three in  $\hat{q}$  or vice versa and always three in  $\hat{k}$ . We thus work in terms of the Gaunt integral  $\mathcal{G}$  of three spherical harmonics (all unconjugated) and a generalized Gaunt integral  $\mathcal{H}$  of four spherical harmonics (again all unconjugated). Explicit expressions for  $\mathcal{G}$  and  $\mathcal{H}$  are given in Appendix A. Performing the angular integrals, the covariance becomes

$$\begin{aligned} \text{Cov}_{l_1 l_2 m, l'_1 l'_2 m'}(r_1, r_2; r'_1, r'_2) &= (-i)^{l_1+l_2+l'_1+l'_2} \\ &\times \frac{1}{V} \int \frac{k^2 dk p^2 dp q^2 dq}{(2\pi^2)^3} \sum_{l_q l_p l_k} P_{l_q}(q) P_{l_p}(p) P_{l_k}(k) \\ &\times (4\pi)^{3/2} \sqrt{\frac{(2l_1+1)(2l_2+1)(2l'_1+1)(2l'_2+1)}{(2l_q+1)(2l_p+1)(2l_k+1)}} \\ &\times \sum_{J_1 J_2 J_3} \sum_{S_1 S_2 S_3} \mathcal{D}_{J_1 J_2 J_3} \mathcal{C}_{J_1 J_2 J_3} \mathcal{R}_{J_1 J_2 J_3}(q, p, k) \\ &\times \begin{pmatrix} J_1 & J_2 & J_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} J_1 & J_2 & J_3 \\ S_1 & S_2 & S_3 \end{pmatrix} \mathcal{J}_{l_1 l_2}(qr_1, pr_2) \\ &\times (-1)^{m+m'} \left\{ \delta_{J_3 l_k}^K \delta_{S_3 0}^K \left[ \mathcal{J}_{l'_1 l'_2}(q, p) \mathcal{H}_{l_q J_1 l_1 l'_1}^{0-S_1-m-m'} \mathcal{H}_{l_p J_2 l_2 l'_2}^{0S_1 m m'} \right. \right. \\ &+ \mathcal{J}_{l'_1 l'_2}(p, q) \mathcal{H}_{l_q J_1 l_1 l'_2}^{0-S_1-m-m'} \mathcal{H}_{l_p J_2 l_2 l'_1}^{0S_1 m-m'} \left. \right] \\ &+ \mathcal{J}_{l'_1 l'_2}(k, p) \mathcal{G}_{l_q J_1 l_1}^{0-S_1-m} \mathcal{H}_{l_p J_2 l_2 l'_1}^{0-S_2 m m'} \mathcal{G}_{l_k J_3 l'_1}^{0-S_3-m'} \\ &+ \mathcal{J}_{l'_1 l'_2}(p, k) \mathcal{G}_{l_q J_1 l_1}^{0-S_1-m} \mathcal{H}_{l_p J_2 l_2 l'_1}^{0-S_2 m-m'} \mathcal{G}_{l_k J_3 l'_2}^{0-S_3 m'} \\ &+ \mathcal{J}_{l'_1 l'_2}(k, q) \mathcal{H}_{l_q J_1 l_1 l'_2}^{0-S_1-m-m'} \mathcal{G}_{l_p J_2 l_2}^{0-S_2 m} \mathcal{G}_{l_k J_3 l'_1}^{0-S_3-m'} \\ &+ \left. \mathcal{J}_{l'_1 l'_2}(q, k) \mathcal{H}_{l_q J_1 l_1 l'_1}^{0-S_1-m-m'} \mathcal{G}_{l_p J_2 l_2}^{0-S_2 m} \mathcal{G}_{l_k J_3 l'_2}^{0-S_3 m'} \right\} \end{aligned} \quad (44)$$

We emphasize that we never do arithmetic in the indices of  $\mathcal{G}$  or  $\mathcal{H}$ , so a minus sign on an index should always be interpreted to mean the negative of the indicated variable. Relative to equations (40) and (42), we used the identity that  $Y_{lm}^* = (-1)^m Y_{l-m}$  to ensure all spherical harmonics were unconjugated prior to integration, and we then used that  $S_1 + S_2 + S_3 = 0$  to simplify (required by the 3j-symbol). In the first and second terms, this condition means that  $S_1 + S_2 = 0$  because the Kronecker delta sets  $S_3 = 0$ . For Gaunt integrals, the spins must sum to zero, and for  $\mathcal{H}$ , the first two spins must have sum equal and opposite to the sum of the last two spins.<sup>5</sup> These rules set  $S_1, S_2$ , and  $S_3$  in equation (44). In fact, they overconstrain it, and so ensuring both rules can be satisfied for each term acts as a consistency check.

We now show that this expression requires only finite sums over angular momenta. The Gaunt integrals  $\mathcal{G}$  require that the total momenta form a closed triangle ('closure condition'), and the spins in turn are bounded by the total momenta. Similarly, the generalized Gaunt integrals  $\mathcal{H}$  require that the total momenta form a closed quadrilateral, and the spins are again bounded by the total momenta. Of the momenta, only the  $J_i$  are free, and these appear at most once in each  $\mathcal{G}$  or  $\mathcal{H}$ ; thus the other, constrained sides along with the closure condition will bound the  $J_i$ , so that all the sums over momenta are finite. In particular, in the Kaiser limit for the anisotropic power spectrum,  $l_q, l_p$ , and  $l_k$  take on values 0, 2, and 4. The other momenta,  $l_1, l_2, l'_1$ , and  $l'_2$ , are fixed by the covariance matrix element desired.

### 6.3 Reduction to integrals of the power spectrum's multipole moments

While equation (44) appears involved, it is actually a considerable simplification of the covariance as regards calculation. Recalling that the  $\mathcal{J}$  are simply products of spherical Bessel functions with arguments given by the Fourier-space magnitude noted ( $p, q$ , or  $k$ ) times a configuration space side length, we will show that equation (44) enables computation of the covariance with 1D and 2D integral transforms. In particular, all of the wave-vector magnitude dependence in the projected covariance can be written in terms of  $f$ -tensors

$$\begin{aligned} f_{nm}^l(r; r_i) &= \int \frac{k^2 dk}{2\pi^2} P_l(k) j_n(kr) j_m(kr_i) \\ f_{nmj}^l(r; r_i, r'_j) &= \int \frac{k^2 dk}{2\pi^2} P_l(k) j_n(kr) j_m(kr_i) j_j(kr'_j). \end{aligned} \quad (45)$$

Given that the subscripted variables will be binned into of order  $N_{\text{bins}} = 10$  bins and the spherical Bessel functions of these arguments replaced with their bin-averaged analogues, these tensors need be computed only on a fine grid in  $r$ , not in the  $r_i$  or the  $r'_j$ . Recall that  $r$  is a dummy variable to be integrated over to enforce the Dirac delta function constraint.

We now show how these  $f$ -tensors emerge. Focusing solely on the dependence on the Fourier-space wave-vectors' magnitudes,

<sup>5</sup>This will always be true for two pairs of spins but the fact that here the pairing is first two and last two is specific to how we chose to couple spherical harmonics when evaluating  $\mathcal{H}$ .

the first term in equation (44) scales as

$$\int \frac{k^2 dk}{2\pi^2} P_{l_k}(k) j_{l_k}(kr) \int \frac{q^2 dq}{2\pi^2} P_{l_q}(q) j_{l_q}(qr) j_{l_1}(qr_1) j_{l'_1}(qr'_1) \times \int \frac{p^2 dp}{2\pi^2} P_{l_p}(p) j_{l_p}(pr) j_{l_2}(pr_2) j_{l'_2}(pr'_2), \quad (46)$$

and the first factor further simplifies to

$$\int \frac{k^2 dk}{2\pi^2} P_{l_k}(k) j_{l_k}(kr) = \xi_{l_k}(r), \quad (47)$$

where  $\xi_{l_k}(r)$  is the  $l_k^{\text{th}}$  multipole moment of the anisotropic 2PCF. The second term in equation (44) has the same form as the first term but with  $r'_1 \leftrightarrow r'_2$ .

Again focusing solely on the dependence on wave-vector magnitudes, the third term in equation (44) scales as

$$\int \frac{q^2 dq}{2\pi^2} P_{l_q}(q) j_{l_q}(qr) j_{l_1}(qr_1) \times \int \frac{k^2 dk}{2\pi^2} P_{l_k}(k) j_{l_3}(kr) j_{l'_1}(kr'_1) \times \int \frac{p^2 dp}{2\pi^2} P_{l_p}(p) j_{l_2}(pr) j_{l_2}(pr_1) j_{l'_2}(pr'_2). \quad (48)$$

The fourth term is given by equation (48) with  $r'_1 \leftrightarrow r'_2$ .

Finally, the wave-vector magnitude dependence of the fifth term in equation (44) scales as

$$\int \frac{p^2 dp}{2\pi^2} P_{l_p}(p) j_{l_2}(pr) j_{l_2}(pr_2) \int \frac{k^2 dk}{2\pi^2} P_{l_k}(k) j_{l_3}(kr) j_{l'_1}(kr'_1) \times \int \frac{q^2 dq}{2\pi^2} P_{l_q}(q) j_{l_1}(qr) j_{l_1}(qr_1) j_{l'_2}(qr'_2), \quad (49)$$

with the sixth term given by switching  $r'_1 \leftrightarrow r'_2$  above.

#### 6.4 Elimination of the delta function's spins

We observe that the  $f$ -tensors depend only on the total angular momenta, not the spins. Moreover, as noted earlier, the integrals  $\mathcal{G}$  and  $\mathcal{H}$  determine  $S_1$ ,  $S_2$ , and  $S_3$  in equation (44), so we may eliminate the sum over  $S_1$ ,  $S_2$ , and  $S_3$ . We would expect one can resum without these spins because they enter only the Dirac delta function and it is in total spin-independent. Thus, new weights  $w_i$ ,  $i = 1-6$  for each term can be defined that depend only on the total angular momenta and the free spins  $m$  and  $m'$ . All weights have the same argument:  $w_i = w_i(l_q l_p l_k; l_1 l_2; l'_1 l'_2; mm')$ ; below we suppress it for brevity. The weights are

$$\begin{aligned} w_1 &= \mathcal{H}_{l_q J_1 l_1 l'_1}^{0-S_1-m-m'} \mathcal{H}_{l_p J_2 l_2 l'_2}^{0S_1 mm'}, & S_1 &= -m - m' \\ w_2 &= \mathcal{H}_{l_q J_1 l_1 l'_2}^{0-S_1-mm'} \mathcal{H}_{l_p J_2 l_2 l'_1}^{0S_1 m-m'}, & S_1 &= m' - m \\ w_3 &= \mathcal{G}_{l_q J_1 l_1}^{0m-m} \mathcal{H}_{l_p J_2 l_2 l'_1}^{0-S_2 mm'} \mathcal{G}_{l_k J_3 l'_1}^{0m'-m'}, & S_2 &= m + m' \\ w_4 &= \mathcal{G}_{l_q J_1 l_1}^{0m-m} \mathcal{H}_{l_p J_2 l_2 l'_1}^{0-S_2 m-m'} \mathcal{G}_{l_k J_3 l'_2}^{0-m'm'}, & S_2 &= m - m' \\ w_5 &= \mathcal{H}_{l_q J_1 l_1 l'_2}^{0-S_1-mm'} \mathcal{G}_{l_p J_2 l_2}^{0-mm} \mathcal{G}_{l_k J_3 l'_1}^{0m'-m'}, & S_1 &= m' - m \\ w_6 &= \mathcal{H}_{l_q J_1 l_1 l'_1}^{0-S_1-m-m'} \mathcal{G}_{l_p J_2 l_2}^{0-mm} \mathcal{G}_{l_k J_3 l'_2}^{0-m'm'}, & S_1 &= -m - m'. \end{aligned} \quad (50)$$

We have explicitly replaced the spins  $S_i$  that can be written simply as  $\pm m$  or  $\pm m'$ , but indicated above at right, the values of spins written as sums or differences of  $m$  and  $m'$  to avoid ambiguous superscripts.

Written using these weights (with arguments suppressed), the covariance becomes

$$\begin{aligned} \text{Cov}_{l_1 l_2 m, l'_1 l'_2 m'}(r_1, r_2; r'_1, r'_2) &= \frac{(4\pi)^{3/2}}{V} (-1)^{m+m'} (-i)^{l_1+l_2+l'_1+l'_2} \\ &\times \int r^2 dr \sum_{l_q l_p l_k} \frac{1}{\sqrt{(2l_q+1)(2l_p+1)(2l_k+1)}} \\ &\times \sum_{J_1 J_2 J_3} \mathcal{D}_{J_1 J_2 J_3} \mathcal{C}_{J_1 J_2 J_3} \begin{pmatrix} J_1 & J_2 & J_3 \\ 0 & 0 & 0 \end{pmatrix} \\ &\times \left\{ \xi_{l_k}(r) \left[ w_1 f_{J_1 l_1 l'_1}^{l_q}(r; r_1, r'_1) f_{J_2 l_2 l'_2}^{l_p}(r; r_2, r'_2) \right. \right. \\ &+ w_2 f_{J_1 l_1 l'_2}^{l_q}(r; r_1, r'_2) f_{J_2 l_2 l'_1}^{l_p}(r; r_2, r'_1) \left. \right] + \begin{pmatrix} J_1 & J_2 & J_3 \\ S_1 & S_2 & S_3 \end{pmatrix} \\ &\times \left\{ f_{J_1 l_1}^{l_q}(r; r_1) \left[ w_3 f_{J_2 l_2 l'_2}^{l_p}(r; r_2, r'_2) f_{J_3 l'_1}^{l_k}(r; r'_1) \delta_{S_1-m, S_3-m'}^K \right. \right. \\ &+ w_4 f_{J_2 l_2 l'_1}^{l_p}(r; r_2, r'_1) f_{J_3 l'_2}^{l_k}(r; r'_2) \delta_{S_1-m, S_3 m'}^K \left. \right] \\ &+ f_{J_2 l_2}^{l_p}(r; r_2) \left[ w_5 f_{J_1 l_1 l'_2}^{l_q}(r; r_1, r'_2) f_{J_3 l'_1}^{l_k}(r; r'_1) \delta_{S_2 m, S_3-m'}^K \right. \\ &+ w_6 f_{J_1 l_1 l'_1}^{l_q}(r; r_1, r'_1) f_{J_3 l'_2}^{l_k}(r; r'_2) \delta_{S_2 m, S_3 m'}^K \left. \right] \left. \right\}. \end{aligned} \quad (51)$$

In the terms proportional to  $w_3$  through  $w_6$ , we fix two of the spins using the Kronecker deltas, and the third is given in equation (50) to avoid showing arithmetic in the index of a Kronecker delta. We emphasize that equation (51) is fully general because the power spectrum can always be expanded in a Legendre series.

We close by emphasizing that equation (51), though it appears complicated, substantially reduces the computational burden of obtaining the covariance matrix. The weights are simply an enumerable set of constant coefficients, and the main work is computing the  $f$ -tensors. However, these are simply 2D and 3D integral transforms of the power spectrum, and once a set of them is computed, it can be combined to form the full covariance as above. Furthermore, since  $r_1$ ,  $r_2$ ,  $r'_1$ , and  $r'_2$  are binned, one can replace the spherical Bessel functions by their bin-averaged values, and the  $f$ -tensors need then only be obtained at the  $\sim N_{\text{bins}}^2$  combinations of bin centres rather than on a full grid in any of the  $r_i$ . A fine grid in  $r$  is required, since  $r$  is subsequently integrated over. Thus, in essence the  $f$ -tensors are not even truly 2D or 3D transforms, but rather 1D cross a small set of bin centre combinations. Equation (51) therefore represents the reduction of the 12D integral naively required for the covariance to a sum over a number of roughly 1D integrals.

#### 6.5 Covariance with the Kaiser formula power spectrum

We now observe that in the Kaiser approximation, where there is a single line of sight to the entire survey (Kaiser 1987), and using linear perturbation theory, the redshift-space density rescales the real-space density as

$$\tilde{\delta}_s(\mathbf{k}) = (1 + \beta \mu^2)^2 \tilde{\delta}(\mathbf{k}), \quad (52)$$

where  $\beta = f/b_1$ , with  $f = d \ln D / d \ln a$ , the logarithmic derivative of the linear growth rate  $D$  with respect to scale factor  $a$  and  $b_1$  the linear bias.  $\mu = \hat{k} \cdot \hat{n}$ , where  $\hat{n}$  is the line of sight to the survey. In this approximation, the  $k$ -dependence of all three moments of the anisotropic 2PCF is the same: it is just the power spectrum. The

multipole moments simply have different constant pre-factors, i.e.

$$\begin{aligned} c_0 &= 1 + \frac{2\beta}{3} + \frac{\beta^2}{5} \\ c_2 &= \frac{4\beta}{3} + \frac{4\beta^2}{7} \\ c_4 &= \frac{8\beta^2}{35}, \end{aligned} \quad (53)$$

where  $c_0$  is for the monopole,  $c_2$  for the quadrupole, and  $c_4$  for the hexadecapole.

This approximation permits considerable simplification of the covariance equation (51). The upper index may be eliminated from all  $f$ -tensors,  $\xi_{l_k}$  becomes a transform of the power spectrum against  $j_{l_k}(kr)$  with a pre-factor of  $c_{l_k}$ , and so a factor of  $c_{l_q} c_{l_p} c_{l_k}$  is inserted within the sum over  $l_q, l_p$ , and  $l_k$ . This step substantially reduces the number of  $f$ -tensors necessary to evaluate.

### 6.6 Incorporating shot noise

For any discrete sampling of the underlying continuous density field, a shot-noise term inversely proportional to the survey number density  $n$  will enter the covariance. As discussed in more detail in Slepian & Eisenstein (2015c), this can be incorporated in the  $f$ -tensor framework simply by mapping  $P(k) \rightarrow P(k) + 1/n$ , where in this case, the shot-noise term enters only the monopole moment of the power spectrum with respect to the line of sight. The shot noise can only enter the monopole because a discrete sampling effect should have no direction dependence. When the combinations  $P + 1/n$  are multiplied out, some of the terms will involve only  $1/n$ . As also detailed in Slepian & Eisenstein (2015c), the  $f$ -tensors proportional to these terms can then be evaluated in closed form; this might be used to accelerate covariance matrix evaluations if desired.

### 6.7 Reduction to the isotropic covariance

As a check, we reduce the result equation (51) to the isotropic 3PCF covariance by setting  $l_1 = l_2 = l, l'_1 = l'_2 = l'$ , and  $l_q = l_p = l_k = 0$ , as well as summing over  $m$  and  $m'$ . In this limit, we find

$$\begin{aligned} \mathcal{H}_{0J_1 l l'}^{0S_1 m m'} &= \frac{1}{4\pi} \sqrt{(2J_1 + 1)(2l + 1)(2l' + 1)} \\ &\times \begin{pmatrix} J_1 & l & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} J_1 & l & l' \\ -S_1 & -m & -m' \end{pmatrix} \end{aligned} \quad (54)$$

where we used Olver et al. (2010, 34.3.1) to evaluate two 3j-symbols that had zeros in one column (the full form of  $\mathcal{H}$  is in Appendix A). Inserting this result (and its analogue of the same form for different spins) into  $w_1$ , employing 3j-symbol identities Olver et al. (2010, 34.3.10 and 34.3.18), and summing over  $m$  and  $m'$ , we find

$$w_1 = \begin{pmatrix} J_1 & l & l' \\ 0 & 0 & 0 \end{pmatrix}^2. \quad (55)$$

Applying the same manipulations, we find  $w_2 = w_1$ .

The other weights are slightly more complicated. Thus, we first present them prior to summing over  $m$  and  $m'$ . For  $w_3$ , we find

$$\begin{aligned} w_3 &= \frac{(-1)^{-m'-m}}{(4\pi)^2} \sqrt{(2J_2 + 1)(2l + 1)(2l' + 1)} \\ &\times \begin{pmatrix} J_2 & l & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} J_2 & l & l' \\ -S_2 & m & m' \end{pmatrix}; \end{aligned} \quad (56)$$

$w_4$  is the same save for switching  $m' \rightarrow -m'$  in the above. For  $w_5$ , we find

$$\begin{aligned} w_5 &= \frac{(-1)^{m'+m}}{(4\pi)^2} \sqrt{(2J_1 + 1)(2l + 1)(2l' + 1)} \\ &\times \begin{pmatrix} J_1 & l & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} J_1 & l & l' \\ -S_1 & -m & m' \end{pmatrix}; \end{aligned} \quad (57)$$

$w_6$  is the same save for switching  $m' \rightarrow -m'$  in the above.

Inserting  $w_3$  through  $w_6$  into equation (51), we note that the pre-factors of  $(-1)^{\pm m \pm m'}$  in the weights cancel with the overall pre-factor  $(-1)^{m+m'}$ . We now sum over  $m$  and  $m'$  with the appropriate replacements for  $S_1, S_2$ , and  $S_3$  in the 3j-symbol that is a pre-factor of these terms. We then invoke orthogonality of the 3j-symbols summed over spins (Olver et al. 2010, 34.3.18) to find that the sums of products of spin-dependent symbols yield unity. We identify  $J_1$  or  $J_2$  as appropriate with  $l_2$  of Slepian & Eisenstein (2015c, equation 65) for the isotropic covariance (they are simply dummy momenta coupling  $l$  and  $l'$ ). Finally, the pre-factor  $(-i)^{l_1+l_2+l'_1+l'_2}$  in equation (51) becomes  $(-1)^{l+l'}$ , and we multiply our covariance by  $(4\pi)^2/[(2l + 1)(2l' + 1)]$ , incorporating the pre-factor in equation (1) for each of the two 3PCFs forming the covariance. Simplifying what results reduces our anisotropic covariance to the isotropic covariance of Slepian & Eisenstein (2015c).

### 6.8 Symmetrization

We now discuss how to cast the covariance matrix of the 3PCF's symmetrized harmonic coefficients in terms of the covariance of the unsymmetrized coefficients.

We desire  $\langle \bar{\zeta}_{l_1 l_2}^m \bar{\zeta}_{l'_1 l'_2}^{m'} \rangle$  where here we take it that  $m \geq 0$  and  $m' \geq 0$ . Rewriting the symmetrized coefficients in terms of the unsymmetrized ones using equation (9) and multiplying out, we obtain

$$\begin{aligned} \langle \bar{\zeta}_{l_1 l_2}^m \bar{\zeta}_{l'_1 l'_2}^{m'} \rangle &= \langle \zeta_{l_1 l_2}^m \zeta_{l'_1 l'_2}^{m'} \rangle + (1 - \delta_{m'0}^K) \langle \zeta_{l_1 l_2}^m \zeta_{l'_1 l'_2}^{-m'} \rangle \\ &+ (1 - \delta_{m0}^K) \langle \zeta_{l_1 l_2}^{-m} \zeta_{l'_1 l'_2}^{m'} \rangle \\ &+ (1 - \delta_{m0}^K) (1 - \delta_{m'0}^K) \langle \zeta_{l_1 l_2}^{-m} \zeta_{l'_1 l'_2}^{-m'} \rangle. \end{aligned} \quad (58)$$

We notice that the last term is the complex conjugate of the first term for  $m$  and  $m' > 0$ , and that for  $m$  and  $m' = 0$ , it drops out (in this case the first term is its own complex conjugate). Similarly, the third term is the complex conjugate of the second term for non-zero spins. This confirms our expectation that the covariance be real given that the symmetrized 3PCF coefficients are real. These observations mean that the complex conjugate pairs can be added to reduce the symmetrized coefficients' covariance to

$$\begin{aligned} \langle \bar{\zeta}_{l_1 l_2}^m \bar{\zeta}_{l'_1 l'_2}^{m'} \rangle &= [2 - \delta_{m'0}^K - \delta_{m0}^K + \delta_{m0}^K \delta_{m'0}^K] \text{Re} \langle \zeta_{l_1 l_2}^m \zeta_{l'_1 l'_2}^{m'} \rangle \\ &+ [2 - \delta_{m'0}^K - \delta_{m0}^K] \text{Re} \langle \zeta_{l_1 l_2}^m \zeta_{l'_1 l'_2}^{-m'} \rangle, \end{aligned} \quad (59)$$

where to add the second and third terms of equation (58), we used that

$$\text{Re} \langle \zeta_{l_1 l_2}^m \zeta_{l'_1 l'_2}^{-m'} \rangle = \text{Re} \langle \zeta_{l_1 l_2}^{-m} \zeta_{l'_1 l'_2}^{m'} \rangle \quad (60)$$

because the product within the expectation value on the left-hand side above is the complex conjugate of that on the right-hand side above.

## 7 CONCLUSIONS

We have presented an algorithm for tracking the full 5D anisotropic 3PCF. Assuming the RSD have azimuthal symmetry about the line of sight, the 3PCF depends on the three ‘internal’ triangle parameters  $r_1$ ,  $r_2$ , and  $\hat{r}_1 \cdot \hat{r}_2$ , and the angles of two of the triangle sides to the line of sight. Here, we have traded these parameters for an equivalent 5D representation – two side lengths  $r_1$  and  $r_2$ , two total angular momenta  $l$  and  $l'$ , and a spin  $m$  – and constructed the mixed spherical harmonic coefficients of the anisotropic 3PCF to capture its angle and orientation dependences. Our reformulation of the problem renders the density field integrals for these coefficients factorizable, fundamentally reducing the scaling of the problem from an  $\mathcal{O}(N^3)$  triplet count to an  $\mathcal{O}(N^2)$  pair count, with  $N$  the number of galaxies in the survey.

In addition to its speed, the algorithm presented here has three other significant advantages. First, it allows use of a rotating line of sight for the 3PCF, more accurate than assuming a single line of sight to the entire survey. The line of sight to the galaxy triplet is taken to be the vector to one of the three triplet members; this is the analogue of the Yamamoto estimator for the anisotropic 2PCF or power spectrum. Furthermore, we have shown how these rotations can be done after computation of the spherical harmonic moments, enabling use of FFTs to evaluate the anisotropic 3PCF. Second, the spherical harmonic basis permits straightforward edge correction, an essential step to remove spurious signal generated by the survey geometry rather than the underlying galaxy clustering. Third, the basis enables computation of the covariance matrix under the assumption of a GRF density described by a power spectrum with multipole moments with respect to the line of sight. In the Kaiser approximation of a flat sky, single line of sight and linear perturbation theory, the power spectrum’s multipole moments become particularly simple, further accelerating evaluation of the covariance.

We note that the basis of spherical harmonic moments advocated here is a compression of the full redshift-space anisotropic 3PCF, as formally an infinite number of  $\ell$  and  $\ell'$  are required to model an arbitrary function of two directions. However, as shown in Slepian & Eisenstein (2015b, 2016a), in practice for the isotropic 3PCF, a finite, small number of multipoles contains the bulk of the information, at least on scales sufficiently large to be well modelled by perturbation theory. Given the structure of the anisotropic bispectrum in perturbation theory (e.g. Rampf & Wong 2012), we expect this conclusion will hold for the anisotropic 3PCF as well. Thus, we believe that the spherical harmonic basis is a parsimonious yet effective compression of the full anisotropic 3PCF.

There has been some work on modelling the anisotropic component of the bispectrum: Scoccimarro et al. (1999) compute the tree-level redshift-space bispectrum via Eulerian standard perturbation theory. Notably, the tree-level Eulerian and Lagrangian predictions for the redshift-space bispectrum (and hence 3PCF, since it is just the inverse FT) agree (Rampf & Wong 2012). This latter work presents the full, unaveraged tree-level bispectrum prediction. In contrast, Scoccimarro et al. (1999) modelled an average of the 5D redshift-space bispectrum over rotations of one wave vector about the other, reducing it to a 4D function. That work began with the basis of spherical harmonics for the bispectrum’s orientation dependence times coefficients depending on the three wave vector magnitudes. The averaging then reduced this to a basis of Legendre polynomials in the angle between the unaveraged side and the line of sight times the same wavevector-magnitude-dependent coefficients.

More recently, Gagrani & Samushia (2017) compared the information content of the full redshift-space bispectrum to that of its

4D reduction computed in this way, showing that at the level of the Fisher matrix most of the information is retained after this averaging. The basis we propose in this work could easily be used for the bispectrum as well, but it is sufficiently different from the basis used in Scoccimarro et al. (1999) and Gagrani & Samushia (2017) that it is not clear how one would average over rotations about one triangle side if so desired.

We note that two important papers in the development of spherical harmonics and Legendre polynomials for the bispectrum and 3PCF are Verde, Heavens & Matarrese (2000) and Szapudi (2004). The former proposed expansion of the projected galaxy density field in spherical harmonics, showed that the isotropic projected bispectrum can be written as a Legendre series, and noted that this formalism also covers the full 3D case. However, it did not discuss the anisotropic 3PCF or bispectrum. The latter proposed expanding the full 3D 3PCF or bispectrum in Legendre polynomials, but again did not discuss anisotropy.

Here, we have presented the mathematical formalism for our anisotropic 3PCF algorithm; in two companion papers, we discuss in detail an implementation suitable for massive-scale high-performance computing (Friesen et al. 2017). In particular, we modified a codebase originally developed in Slepian & Eisenstein (2015c) to track the anisotropic clustering as discussed here. We optimized this code to run on the Cray XC40 system Cori at Lawrence Berkeley National Laboratory’s National Energy Research Supercomputing Center, which comprises roughly 10k nodes, each with 68 compute cores. Running it on the largest available galaxy simulation, Outer Rim (Habib et al. 2016), with 2 billion haloes, we computed the anisotropic 3PCF out to  $200 \text{ Mpc} h^{-1}$  in 1070 s on 9636 nodes, achieving 5.06 PFLOPS sustained. At peak, the code achieved 9.8 PFLOPS, roughly 39 per cent of peak performance, but 80 per cent given the instruction mix the algorithm requires.<sup>6</sup> This speed means that any anisotropic 3PCF computation for galaxy surveys of sizes available in the next decade is practical even on current computing resources. Indeed, the anisotropic 3PCF out to  $R_{\text{max}} = 200 \text{ Mpc}$  for all galaxies in the observable Universe ( $\sim 100$  billion) is computable in a few days with the algorithm on Cori. The speed is further important because a full 3PCF analysis requires computing many random catalogs’ 3PCF for edge correction, and many mocks’ 3PCF for pipeline testing, model and covariance matrix verification, and fitting the free parameters (volume and shot noise) of the covariance matrix.

Regarding speed, we note that several works have presented acceleration schemes for the neighbour-finding piece of the 3PCF computation. Bernardeau, van Waerbeke & Mellier (2003) use Delaunay triangulation to accelerate neighbour finding in the context of computing the weak shear 3PCF. March (2013) uses kd-trees for the neighbour finding, but does not include any further acceleration in computing the actual 3PCF, leading to an algorithm that still scales as the number of secondaries squared around a given primary, for an overall  $N^3$  scaling (see his fig. 21). Indeed, our high-performance computing implementation of the algorithm, presented in the companion paper Friesen et al. (2017), uses kd-trees for the neighbour finding as well. However, as fig. 4 of Friesen et al. (2017) shows,

<sup>6</sup>For this implementation, we pre-rotated the galaxies, which while computationally less efficient is simpler to code than the post-rotation approach also presented here. We expect incorporating this latter optimization will only improve runtimes, though we do not expect significant gains because the pre-rotation is a small part of the total work in the current code, as discussed in Section 4.2.

the neighbour finding is a subdominant part of the total time cost as compared to the spherical harmonic computation (22 per cent versus 55 per cent). Consequently, the neighbour finding is not the primary driver of the algorithm's fundamental scaling. If further acceleration of the algorithm were the paramount consideration, the FT-based approach that computes the harmonic coefficients scaling as  $N_g \log N_g$ , discussed in Section 4.3 would likely be the appropriate choice.

Regarding the algorithm's scaling, as  $(nV_{\max})N$ , with  $V_{\max}$  the spherical volume within a radius of  $R_{\max}$ ,  $n$  the survey number density, and  $N$  the number of galaxies, we note that this applies in the limit that one works on large enough scales that the density averaged within  $V_{\max}$  has returned to nearly its overall average value. This occurs because on sufficiently large-scale galaxies are essentially unclustered. In particular, the number of neighbours within a given sphere about a primary galaxy scales as the volume of the sphere plus the integral of the 2PCF over the sphere, whereas the number of neighbours about a randomly chosen point in the survey scales as the volume of the sphere. The 2PCF integrates to zero over sufficiently large scales ( $\gtrsim 150 \text{ Mpc}^{-1}$ ). Thus, the number of neighbours about a primary converges to the average number of galaxies within a randomly placed sphere in the survey, rendering valid our estimate of the number of neighbours as  $nV_{\max}$ .

Future work will be translating the predictions of Rampf & Wong (2012) into the spherical harmonic basis for direct comparison with the output of the algorithm. Indeed, for this purpose, one could also use the predictions of Hashimoto, Rasera & Taruya (2017), which include additional terms representing further perturbative corrections to the redshift-space bispectrum. What is clear by inspection is that the isotropic part of the 3PCF can only generate  $l = l'$  couplings, so any 'off-diagonal' couplings  $l \neq l'$  isolate  $\mathcal{O}(f)$  contributions. Thus, in principle, these couplings provide a robust window on the growth rate. However, in practice, anisotropies in the survey mask can couple isotropic coefficients of the measured 3PCF to anisotropic coefficients of the edge-corrected 3PCF, and vice versa. More detailed analysis with mock catalogues and a realistic survey geometry will thus need to be conducted to fully quantify this coupling.

None the less, we believe the algorithm presented here will enable precise, robust measurement of the growth rate of structure with the anisotropic 3PCF much as is already done with the anisotropic 2PCF and power spectrum. Thus far, the anisotropic 3PCF has not been measured. Therefore, the next step for future work is applying this algorithm to data. Numerous suitable samples exist already, such as the SDSS DR12 BOSS CMASS sample used in Slepian et al. (2016a) and Gil-Marín et al. (2017) or, if smaller volume but higher number density were desired, the VIMOS Public Extragalactic Redshift Survey (VIPERS) (Scodreggio et al. 2016). Ongoing and future surveys such as eBOSS (Dawson et al. 2016) and Dark Energy Spectroscopic Instrument (DESI) (Levi et al. 2013) will provide even larger, richer data sets to which to apply this algorithm.

Measuring the growth rate of structure via RSD is both an important lever on the cosmological parameters and a key test of our theory of gravity. While the anisotropic 2PCF and power spectrum already will probe it to extremely high precision with next-generation surveys, any additional sources of information can only strengthen our understanding of these two fundamental areas.

Further, breaking the degeneracy between  $\sigma_8$ ,  $f$ , and galaxy biasing is a challenging problem that requires several different observables to fully address. In addition to its importance for cosmology,

extracting precise bias measurements will shed new light on galaxy formation.

Importantly, measuring the anisotropic 3PCF requires no additional data over what is already used for the anisotropic 2PCF; the challenge is purely computational, not observational. Thus, the algorithm presented here should enhance the scientific value of redshift survey data per dollar spent on telescope time.

## ACKNOWLEDGEMENTS

ZS especially thanks Shirley Ho for encouragement on the computing side of this study and Anya Nugent and Martin White for careful reads of the manuscript. We also thank Douglas Finkbeiner, Stephen Portillo, Natalie Roe, Roman Scoccimarro, Uroš Seljak, and David Weinberg for useful discussions. Support for this work was provided by the National Aeronautics and Space Administration through Einstein Postdoctoral Fellowship Award Number PF7-180167 issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of the National Aeronautics Space Administration under contract NAS8-03060. ZS also acknowledges support from a Chamberlain Fellowship at Lawrence Berkeley National Laboratory (held previously to the Einstein) and from the Berkeley Center for Cosmological Physics. DJE acknowledges support as a Simons Foundation Investigator and from U.S. Department of Energy grant DE-SC0013718.

## REFERENCES

- Adams J., 1878, *Proc. R. Soc.*, **455**, 27, 63  
 Arfken G. B., Weber H. J., Harris F. E., 2013, *Mathematical Methods for Physicists*, seventh edn. Academic Press, Boston  
 Bernardeau F., van Waerbeke L., Mellier Y., 2003, *A&A*, **397**, 405  
 Beutler F. et al., 2017, *MNRAS*, **466**, 2242  
 Bianchi D., Gil-Marín H., Ruggeri R., Percival W. J., 2015, *MNRAS*, **453**, L11  
 Bianchi D., Percival W. J., Bel J., 2016, *MNRAS*, **463**, 3783  
 Dawson K. S. et al., 2016, *AJ*, **151**, 44  
 Eisenstein D. J., Seo H.-J., White M., 2007, *ApJ*, **664**, 660  
 Feng X. M., Wang P., Yang W., Jin G. R., 2015, *Phys. Rev. E*, **92**, 043307  
 Friesen B. et al., 2017, preprint (arXiv:1709.00086)  
 Gagrani P., Samushia L., 2017, *MNRAS*, **467**, 928  
 Gardner J. P., Connolly A., McBride C., 2007, in Shaw R. A., Hill F., Bell D. J., eds, *ASP Conf. Ser. Vol. 376, Astronomical Data Analysis Software and Systems XVI*, Astron. Soc. Pac., San Francisco. p. 69  
 Gil-Marín H., Percival W. J., Verde L., Brownstein J. R., Chuang C.-H., Kitaura F.-S., Rodríguez-Torres S. A., Olmstead M. D., 2017, *MNRAS*, **465**, 1757  
 Gimbutas Z., Greengard L., 2009, *J. Comput. Phys.*, **228**, 5621  
 Grieb J. N. et al., 2017, *MNRAS*, **467**, 2085  
 Gumerov N. A., Duraiswami R., 2014, preprint (arXiv:1403.7698)  
 Habib S. et al., 2016, *New Astron.*, **42**, 49  
 Hamilton A. J. S., 1993, *ApJ*, **406**, L47  
 Hamilton A. J. S., 1998, in Hamilton D., ed., *Astrophysics and Space Science Library*, Vol. 231, *The Evolving Universe*, Kluwer Academic Publishers, Dordrecht. p. 185  
 Hand N., Li Y., Slepian Z., Seljak U., 2017, *J. Cosmol. Astropart. Phys.*, **07**, 002  
 Hashimoto I., Rasera Y., Taruya A., 2017, *Phys. Rev. D*, **96**, 043526  
 Jackson J. C., 1972, *MNRAS*, **156**, 1P  
 Jeong D., Dai L., Kamionkowski M., Szalay A. S., 2015, *MNRAS*, **449**, 3312  
 Kaiser N., 1987, *MNRAS*, **227**, 1  
 Kazin E. A., Sánchez A. G., Blanton M. R., 2012, *MNRAS*, **419**, 3223  
 Levi M. et al., 2013, preprint (arXiv:1308.0847)

- Linder E. V., 2005, *Phys. Rev. D*, 72, 043529  
 March B., 2013, PhD thesis, unpublished  
 Olver F. W. J., Lozier D. W., Boisvert R. F., Clark C. W., 2010, NIST Handbook of Mathematical Functions. Cambridge Univ. Press, New York  
 Pápai P., Szapudi I., 2008, *MNRAS*, 389, 292  
 Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ  
 Percival W. J. et al., 2014, *MNRAS*, 439, 2531  
 Raccanelli A. et al., 2013, *MNRAS*, 436, 89  
 Raccanelli A., Bertacca D., Jeong D., Neyrinck M. C., Szalay A. S., 2016, *Phys. Dark Univ.*, 19, 109,  
 Rampf C., Wong Y. Y. Y., 2012, *J. Cosmol. Astropart. Phys.*, 6, 018  
 Reimberg P., Bernardeau F., Pitrou C., 2016, *J. Cosmol. Astropart. Phys.*, 1, 048  
 Ross A. J. et al., 2017, *MNRAS*, 464, 1168  
 Samushia L., Branchini E., Percival W. J., 2015, *MNRAS*, 452, 3704  
 Scoccimarro R., 2015, *Phys. Rev. D*, 92, 083532  
 Scoccimarro R., Couchman H. M. P., Frieman J. A., 1999, *ApJ*, 517, 531  
 Scodreggio M. et al., 2016, *A&A*, 609, A84  
 Slepian Z., Eisenstein D. J., 2015a, preprint (arXiv:1510.04809)  
 Slepian Z., Eisenstein D. J., 2015b, *MNRAS*, 448, 9  
 Slepian Z., Eisenstein D. J., 2015c, *MNRAS*, 454, 4142  
 Slepian Z., Eisenstein D. J., 2016a, *MNRAS*, 469, 2059,  
 Slepian Z., Eisenstein D. J., 2016b, *MNRAS*, 455, L31  
 Slepian Z., Eisenstein D. J., 2016c, *MNRAS*, 457, 24  
 Slepian Z. et al., 2016a, *MNRAS*, 469, 1738  
 Slepian Z. et al., 2017, *MNRAS*, 474, 2109,  
 Slepian Z. et al., 2017, *MNRAS*, 468, 1070  
 Szapudi I., 2004, *ApJ*, 605, L89  
 Szapudi I., Szalay A. S., 1998, *ApJ*, 494, L41  
 Taruya A., Nishimichi T., Saito S., 2010, *Phys. Rev. D*, 82, 063522  
 Varshalov Moskaliev Kersonskii 2013, Quantum Theory of Angular Momentum, seventh edn. Academic Press, Boston  
 Verde L., Heavens A. F., Matarrese S., 2000, *MNRAS*, 318, 584  
 Vlah Z., Castorina E., White M., 2016, *J. Cosmol. Astropart. Phys.*, 12, 007  
 Wang L., Reid B., White M., 2014, *MNRAS*, 437, 588  
 White M., Reid B., Chuang C.-H., Tinker J. L., McBride C. K., Prada F., Samushia L., 2015, *MNRAS*, 447, 234  
 Yamamoto K., Nakamichi M., Kamino A., Bassett B. A., Nishioka H., 2006, *PASJ*, 58, 93  
 Zhang L. L., Pen U.-L., 2005, *New Astron.*, 10, 569

## APPENDIX A

In this appendix, we collect some important definitions and properties of Legendre polynomials and spherical harmonics. A product of two Legendre polynomials can be linearized into a sum as (Adams 1878)

$$\mathcal{L}_k(\mu)\mathcal{L}_{k'}(\mu) = \sum_J (2J+1) \begin{pmatrix} k & k' & J \\ 0 & 0 & 0 \end{pmatrix}^2 \mathcal{L}_J(\mu), \quad (\text{A1})$$

where the sum's range is set by  $|k-k'| \leq J \leq k+k'$  (the triangularity condition on the 3j-symbol).

The integral of a Legendre polynomial can be obtained using the recursion

$$\mathcal{L}_n(\mu) = \frac{1}{2n+1} \frac{d}{d\mu} [\mathcal{L}_{n+1}(\mu) - \mathcal{L}_{n-1}(\mu)] \quad (\text{A2})$$

to rewrite  $\mathcal{L}_n(\mu)$  as an exact differential.

The Gaunt integral is defined

$$\begin{aligned} \mathcal{G}_{l_1 l_2 l_3}^{m_1 m_2 m_3} &\equiv \int d\Omega Y_{l_1 m_1}(\hat{r}) Y_{l_2 m_2}(\hat{r}) Y_{l_3 m_3}(\hat{r}) \\ &= C_{l_1 l_2 l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \end{aligned} \quad (\text{A3})$$

where  $C_{l_1 l_2 l_3}$  is defined in equation (43).

The integral of four spherical harmonics can be obtained by first linearizing two spherical harmonics into a sum over single spherical harmonics using the Gaunt integral. We begin with

$$Y_{l_1 m_1}(\hat{r}) Y_{l_2 m_2}(\hat{r}) = \sum_{LM} c_{LM}(l_1, l_2; m_1, m_2) Y_{LM}(\hat{r}), \quad (\text{A4})$$

where the coefficients  $c_{LM}(l_1, l_2; m_1, m_2)$  are given by integrating both sides against  $Y_{LM}^*(\hat{r})$  and invoking orthogonality, so that

$$c_{LM}(l_1, l_2; m_1, m_2) = (-1)^M \mathcal{G}_{l_1 l_2 L}^{m_1 m_2 -M}. \quad (\text{A5})$$

We then have

$$\begin{aligned} \mathcal{H}_{l_1 l_2 l_3 l_4}^{m_1 m_2 m_3 m_4} &\equiv \int d\Omega Y_{l_1 m_1}(\hat{r}) Y_{l_2 m_2}(\hat{r}) Y_{l_3 m_3}(\hat{r}) Y_{l_4 m_4}(\hat{r}) \\ &= \sum_L (-1)^M \mathcal{G}_{l_1 l_2 L}^{m_1 m_2 -M} \mathcal{G}_{L l_3 l_4}^{M m_3 m_4} \end{aligned} \quad (\text{A6})$$

by inserting equations (A4) and (A5) into the first line above and then integrating. We note that there is no sum over  $M$  because it is set by the zero-sum rule on the spins enforced by the Gaunt integrals:  $m_1 + m_2 = M$ . We further note that the sum over  $L$  has compact support because of the triangle rules on total angular momenta:  $|l_1 - l_2| \leq L \leq l_1 + l_2$  and the same constraint holds replacing  $l_1 \rightarrow l_3$  and  $l_2 \rightarrow l_4$ .

## APPENDIX B

Having focused this paper on the  $\zeta_{ll'}$  parametrization, we here explain why an alternative, seemingly attractive parametrization of the anisotropic 3PCF does not work. Specifically, by analogy with the anisotropic 2PCF, one might expect that a triple Legendre series in the angle of each triangle side to the line of sight and the internal angle enclosed by the triangle would be the correct basis for the anisotropic 3PCF. However, this approach is not workable, as we show below.

First, we write out the basis about a particular galaxy at  $\mathbf{x}$ ; the full anisotropic 3PCF would then be the average over  $\mathbf{x}$  of these coefficients. We have

$$\begin{aligned} \hat{\zeta}(r_1, r_2; \hat{r}_1 \cdot \hat{r}_2; \hat{r}_1 \cdot \hat{n}, \hat{r}_2 \cdot \hat{n}; \mathbf{x}) \\ = \sum_{l_1 l_2} \zeta_{l_1 l_2}(r_1, r_2; \mathbf{x}) \mathcal{L}_l(\hat{r}_1 \cdot \hat{r}_2) \mathcal{L}_{l_1}(\hat{r}_1 \cdot \hat{n}) \mathcal{L}_{l_2}(\hat{r}_2 \cdot \hat{n}), \end{aligned} \quad (\text{B1})$$

with  $\hat{n}$  the line of sight.

There are two ways to see the flaw in this basis. First, consider a triangle with zero opening angle, so that  $\hat{r}_1 \cdot \hat{r}_2 = 1$ . Then, specifying the orientation of  $\hat{r}_1$  with respect to the line of sight fully specifies that of  $\hat{r}_2$ : the Legendre polynomials in  $l_1$  and  $l_2$  are no longer independent. This lack of independence means that they do not form a basis.

A more formal way to see this issue is by considering how we would obtain the coefficients  $\zeta_{l_1 l_2}(r_1, r_2; \mathbf{x})$ . Placing a particular galaxy at  $\mathbf{x}$  and measuring the coefficients around it, we would attempt to invoke orthogonality by integrating over  $d\Omega_{r_1} d\Omega_{r_2}$ . To do so, we would need to separate the Legendre polynomials on the right-hand side of equation (B1) as

$$\begin{aligned} \mathcal{L}_l(\hat{r}_1 \cdot \hat{r}_2) \mathcal{L}_{l_1}(\hat{r}_1 \cdot \hat{n}) \mathcal{L}_{l_2}(\hat{r}_2 \cdot \hat{n}) \\ = \frac{(4\pi)^2}{(2l+1)\sqrt{(2l+1)(2l_2+1)}} \\ \times \sum_{L_1 L_2} \mathcal{G}_{l l_1 L_1}^{000} \mathcal{G}_{l l_2 L_2}^{000} Y_{L_1 0}(\hat{r}_1) Y_{L_2 0}^*(\hat{r}_2). \end{aligned} \quad (\text{B2})$$



We used the spherical harmonic addition theorem to expand each Legendre polynomial, and then that  $\hat{n} = \hat{z}$  to eliminate the sums over  $m_1$  and  $m_2$ , since only the  $m_1 = m_2 = 0$  modes contribute for a spherical harmonic evaluated along the  $z$ -axis. We then linearized the resulting products of two spherical harmonics in  $\hat{r}_1$  and  $\hat{r}_2$  into a sum over one spherical harmonic in each using equation (A4) from Appendix A, and substituted the explicit form for the linearization coefficients as it is particularly simple, since all spins are zero.

Inserting equation (B2) into equation (B1) and integrating both sides against spherical harmonics over  $d\Omega_{r_1} d\Omega_{r_2}$ , we see that

$$\int d\Omega_{r_1} d\Omega_{r_2} \hat{\zeta}(r_1, r_2; \mathbf{x}) Y_{L_1 M_1}^*(\hat{r}_1) Y_{L_2 M_2}(\hat{r}_2) \propto \sum_{l_1 l_2 m} \mathcal{G}_{l_1 L_1}^{000} \mathcal{G}_{l_2 L_2}^{000} \zeta_{l_1 l_2}(r_1, r_2; \mathbf{x}). \quad (\text{B3})$$

We see that we have failed to extract the desired coefficients  $\zeta_{l_1 l_2}$ , but have only succeeded in measuring some weighted sum of them.

If the  $l$ ,  $l_1$ , and  $l_2$  support of the anisotropic 3PCF were finite, we could measure a large number of integrals as on the left-hand side above and solve for each term in the sum. However, it is known that the multipole expansion of the isotropic 3PCF is formally infinite (Slepian & Eisenstein 2015b, 2016a), so  $l$  ranges from zero to infinity and this approach is not possible. We note that this point is not at odds with our claim in Section 7 that the multipoles are a parsimonious basis for the 3PCF. While the multipole expansion is formally infinite, as shown in Slepian & Eisenstein (2016a, fig. 8), the higher multipoles do not have much new information over the lower (their side-length dependence largely converges to look similar for  $l \gtrsim 5$ ), so the information content remains compact.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.