Beyond Yamamoto: Anisotropic Power Spectra and Correlation Functions with Pairwise Lines-of-Sight

Oliver H.E. Philcox1,2* and Zachary Slepian3,4†

1Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
2School of Natural Sciences, Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA
3Department of Astronomy, University of Florida, 211 Bryant Space Science Center, Gainesville, FL 32611, USA
4Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

ABSTRACT

Conventional estimators of the anisotropic power spectrum and two-point correlation function (2PCF) adopt the ‘Yamamoto approximation’, fixing the line-of-sight of a pair of galaxies to that of just one of its members. Whilst this is accurate only to first-order in the characteristic opening angle \( \theta_{\text{max}} \), it allows for efficient implementation via Fast Fourier Transforms (FFTs). This work presents practical algorithms for computing the power spectrum and 2PCF multipoles using pairwise lines-of-sight, adopting either the galaxy midpoint or angle bisector definitions. Using newly derived infinite series expansions for spherical harmonics and Legendre polynomials, we construct estimators accurate to arbitrary order in \( \theta_{\text{max}} \), though note that the midpoint and bisector formalisms themselves differ at fourth order. Each estimator can be straightforwardly implemented using FFTs, requiring only modest additional computational cost relative to the Yamamoto approximation. We demonstrate the algorithms by applying them to a set of realistic mock galaxy catalogs, and find both procedures produce comparable results for the 2PCF, with a slight preference for the bisector power spectrum algorithm, albeit at the cost of greater memory usage. Such estimators provide a useful method to reduce wide-angle systematics for future surveys.

Key words: cosmology: large-scale structure of Universe, theory – methods: statistical, data analysis

1 INTRODUCTION

We have now entered the epoch of ‘precision cosmology’. In the coming years, the volume of cosmological data available will increase at a prodigious rate, thanks to the advent of large spectroscopic surveys such as DESI (DESI Collaboration et al. 2016), Euclid (Laureijs et al. 2011) and SPHEREx (Doré et al. 2014). As the number of observed galaxies grows, so too does the precision on fundamental parameters such as the growth rate, energy densities and Hubble parameter. Given that we will soon be able to measure summary statistics at the sub-percent level, it is vital to understand also their systematics to this precision, else we risk biasing our inference or losing effective survey volume.

Whilst there is growing interest in more complicated statistics (e.g., Gil-Marín et al. 2017; Slepian et al. 2017; Chudaykin & Ivanov 2019; Philcox et al. 2020; Samushia et al. 2021), the information content of future surveys will be dominated by the two-point correlator, masquerading either as the configuration-space two-point correlation function (2PCF), \( \xi(r) \), or the Fourier-space power spectrum, \( P(k) \) (e.g., Beutler et al. 2017; Alam et al. 2017; eBOSS Collaboration et al. 2020). In a statistically isotropic universe, both of these will depend only the distance between galaxies, be it \( r = |r| \) or the momentum-space equivalent \( k = |k| \). In our Universe this is not the case, since redshift-space distortions (RSD) impart a preferred origin to the observer (Kaiser 1987), sourcing additional cosmological information (e.g., Lesgourgues & Pastor 2006; Weinberg et al. 2013).

To fully encapsulate RSD, two-point statistics should depend on the position vectors to the two galaxies in question, rather than just a single length. Taking into account the various rotational symmetries, such a configuration can be specified by three degrees of freedom; options include the separation \( r \) (or \( k \)) and the angles between the separation vector and the line-of-sight (LoS) to each galaxy (e.g., Pápai &
Szapudi 2008; Yoo & Seljak 2015; Castorina & White 2018) or the separation, the mean distance to the galaxy pair, and a single angle (e.g., Reimberg et al. 2016; Beutler et al. 2019). Unless our interest lies in the largest-possible scales (for instance in $f_{NL}$-analyses), it is usually sufficient to parametrize the two-point correlators by just two variables; the inter-galaxy distance $r$ or $k$, and the angle of the galaxy separation vector to a joint LoS, $\mu$. In this case, the functions can be robustly expanded as a Legendre series in $\mu$, and theory and observations simply compared. Of course, any such approximation necessarily induces wide-angle effects on the largest scales, which are the subject of extensive discussion in the literature (e.g., Hamilton 1992; Hamilton & Culhane 1996; Hamilton 1998; Zaroubi & Hoffman 1999; Szalay et al. 1998; Szapudi 2004; Datta et al. 2007; Pápai & Szapudi 2008; Shaw & Lewis 2008; Bonvín & Durrer 2011; Raccanelli et al. 2014; Yoo & Seljak 2015; Slepian & Eisenstein 2015; Reimberg et al. 2016; Castorina & White 2018; Beutler et al. 2019). In general, the error in these approaches depends on the characteristic opening angle $\theta_{\text{max}}$, defined either as the galaxy pair opening angle in the maximum radial bin (2PCF) or the survey opening angle (power spectrum). The dichotomy arises since the power spectrum depends on an integral over all galaxy pairs, whilst the 2PCF only requires pairs separated by the scale of interest.

If the two-parameter formalism is adopted (as has become commonplace), an important question must be asked: how should one choose the joint LoS to the galaxy pair? The early literature adopted a single LoS for the whole survey (e.g., Kaiser 1987; Hamilton 1992), which, whilst simple to implement, incurs significant errors if the survey is wide. A more accurate prescription is to fix the LoS to the direction vector of a single galaxy, in the ‘Yamamoto approximation’ (Yamamoto et al. 2006). Whilst this gives an error at $O(\theta_{\text{max}}^2)$ for characteristic size $\theta_{\text{max}}$, which becomes important for DESI volumes (Sugiyama et al. 2019), it is straightforward to implement using Fast Fourier Transforms (FFTs), thus is the approach found in most recent analyses (e.g., Scoccimarro 2015; Bianchi et al. 2015; Hand et al. 2017; Beutler et al. 2017; Hand et al. 2018). Beyond this approximation, there are two appealing LoS choices: the angle between the galaxy midpoint and separation vector (cf. Yamamoto et al. 2006; Scoccimarro 2015; Samushia et al. 2015; Bianchi et al. 2015) and the angle bisector (cf. Szalay et al. 1998; Matsubara 2000; Szapudi 2004; Yoo & Seljak 2015). Both are consistent to third order in the opening angle (demonstrated in Slepian & Eisenstein 2015), but incur an error at fourth order, and, for $\theta_{\text{max}} < 10^\circ$, lose little information compared to the double LoS approach (Beutler et al. 2012; Samushia et al. 2012; Yoo & Seljak 2015). However, their naïve implementation scales as $O(N^2)$ for $N$ galaxies, rather than the $O(N_g \log N_g)$ dependence enjoyed by algorithms based on FFTs with $N_g$ grid cells.

In this work, we will demonstrate that the midpoint and bisection power spectrum and 2PCF estimators may be efficiently computed in $O(N_g \log N_g)$ time using FFTs. In both cases, it is necessary to perform a series expansion of the angular dependence in the galaxy pair opening angle $\theta$ (with $\theta < \theta_{\text{max}}$); however, we give the explicit form of these corrections at arbitrary order. The different LoS definitions require different mathematical treatments for greatest efficiency; for the bisection case, we implement the suggestion of Castorina & White (2018) and extend it to arbitrary order, whilst for the midpoint approach, we provide novel formulae based on a newly-derived spherical harmonic shift theorem (similar to that of Garcia & Slepian 2020 for the three-point function). This paper is an extension also of Slepian & Eisenstein (2015), which gave the lowest-order corrections for the 2PCF, but did not consider the Fourier-space counterpart (which carries somewhat more subtleties). In contrast, Samushia et al. (2015) considered the power-spectrum in the midpoint formalism, but only applied to a simplified spherical cap geometry. Our work goes beyond the previous by giving a full catalog of arbitrary-order expressions for the midpoint and bisection formalism in real- and Fourier-space. We further consider their application to realistic data using the MultiDark-\texttt{parsec} mock catalogs (Kitaura et al. 2016), and make the analysis code publicly available.\footnote{github.com/oliverphilcox/BeyondYamamoto}

The remainder of this work is structured as follows. We begin in §2 by recapitulating the basic two-point correlator estimators. §3 & §4 present our implementations of the power spectrum and 2PCF algorithms in the midpoint formalism, before the same is done in the bisection formalism in §§ §5 & §6. §7 considers the application of the algorithms to data, before we conclude in §8. A list of useful mathematical identities is given in Appendix A, with Appendices B-F giving mathematical derivations of results central to this work, in particular, a shift theorem for spherical harmonics and Legendre polynomials. For the reader less interested in mathematical derivations, we recommend skipping §3.3 and the Appendices, and note that the key equations in this work are boxed.

2 ESTIMATORS FOR THE TWO-POINT CORRELATORS

We begin by stating the Fourier conventions used throughout this work. We define the Fourier and inverse Fourier transforms by

$$X(k) \equiv \mathcal{F}[X](k) = \int \! dx \ e^{-ik \cdot x} X(x), \quad X(x) \equiv \mathcal{F}^{-1}[X](x) = \int \frac{dk}{(2\pi)^3} e^{ik \cdot x} X(k), \quad (2.1)$$

leading to the definition of the Dirac delta function, $\delta_D$, as

$$\frac{1}{(2\pi)^3} \delta_D(k_1-k_2) = \int \! dx \ e^{i(k_1-k_2) \cdot x}. \quad (2.2)$$

The correlation function and power spectrum of the density field, $\delta$, are defined as

$$\xi(r) = \langle \delta(x) \delta(x + r) \rangle, \quad (2\pi)^3 \delta_D(k+k') P(k) = \langle \delta(k) \delta(k') \rangle, \quad (2.3)$$

\footnote{github.com/oliverphilcox/BeyondYamamoto}
with the power spectrum as the Fourier transform of the correlation function. Additionally, we use the shorthand

\[ \int \equiv \int \frac{dk}{(2\pi)^3}, \quad \int \Omega_k \equiv \int \frac{d\Omega_k}{4\pi}. \]  

(2.4)

2.1 Power Spectrum

The conventional estimator for the galaxy power spectrum multipoles, \( P_\ell (k) \), is defined as a Fourier transform of two density fields \( \delta (r_1) \) and \( \delta (r_2) \):

\[ P_\ell (k) = \frac{2\ell + 1}{V} \int_{\Omega_k} \int dr_1 dr_2 e^{-ik(r_2 - r_1)} \delta (r_1) \delta (r_2) L_\ell (\hat{k} \cdot \hat{n}) \]  

(2.5)

(e.g., Hand et al. 2017), where \( \hat{n} \) is the joint line-of-sight (LoS) to the pair of galaxies at \( (r_1, r_2) \), \( L_\ell \) is the Legendre polynomial of order \( \ell \), \( V \) is the survey volume and hats denote unit vectors. Whilst we have not included \( k \)-space binning, this is a straightforward addition, requiring an additional integral over \( |k| \). In spectroscopic surveys, we do not have access to \( \delta \) directly, only a set of galaxy and random particle positions with associated density fields \( n_g (r) \) and \( n_r (r) \) respectively. In this context, we replace

\[ \delta (r) \rightarrow \frac{w(r) [n_g (r) - \alpha r n_r (r)]}{l^{1/2}}, \quad I = \int dr w^2 (r) \tilde{n}^2 (r) \]  

(2.6)

(Yamamoto et al. 2006), where \( w(r) \) are some weights (accounting for systematics, optimality and completeness), \( \tilde{n} \) is the mean galaxy density and \( \alpha_r \) is the ratio of randoms to galaxies. (2.5) is strictly an estimator for the window-convolved power spectrum,\(^2\) and includes a shot-noise term. Since the latter does not depend on the LoS, we will ignore it henceforth, and additionally denote the (windowed) density field simply by \( \delta (r) \).

In the simplest (‘plane-parallel’) approximation, we take \( \hat{n} \) to be fixed across the survey, thus the estimator simplifies to the form familiar from \( N \)-body estimators, involving the Fourier-space density field \( \delta (k) \):\(^3\)

\[ \hat{P}_\ell (k) = \frac{2\ell + 1}{V} \int_{\Omega_k} L_\ell (\hat{k} \cdot \hat{n}) |\delta (k)|^2. \]  

(2.7)

(Kaiser 1987; Hamilton 1992; Feldman et al. 1994). This necessarily incurs an error at \( O(\theta_{\max}) \) where \( \theta_{\max} \) is the survey opening angle,\(^4\) and is valid only for the smallest surveys.

At the next order in approximation is the ‘Yamamoto formalism’ used by most current estimators; this approximates the LoS as the

---

\(^2\) One may remove the window function by judicious use of quadratic estimators (e.g., Philcox 2020).

\(^3\) In full, \( \delta (k) \) should be multiplied by a compensation function \( M (k) \) to account for the mass assignment scheme window function.

\(^4\) Note this depends on the survey opening angle not the pair opening angle, since the power spectrum is an integral over all galaxy pairs. Its importance increases as the wavenumber becomes small. We note that this is a rough guide to the size of the error, rather than a formally-defined scaling.
position vector of a single galaxy, i.e. \( \mathbf{n} = \mathbf{r}_1 \) or \( \mathbf{n} = \mathbf{r}_2 \), as shown in Fig. 1. In this case, one may expand \( L^{\ell}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}}) \) via the addition theorem for Legendre polynomials (A5), and arrive at the estimator

\[
\beta^{\text{Yama}}_\ell(k) = \frac{4\pi}{V} \int_{\Delta k} \sum_{m=-\ell}^{\ell} Y^m_\ell(\hat{\mathbf{k}}) \int d\mathbf{r}_1 \ e^{i\mathbf{k} \cdot \mathbf{r}_1} \delta(\mathbf{r}_1) \int d\mathbf{r}_2 \ e^{-i\mathbf{k} \cdot \mathbf{r}_2} Y^m_\ell(\hat{\mathbf{r}}_2) \delta(\mathbf{r}_2) \quad (2.8)
\]

\[
= \frac{4\pi}{V} \int_{\Delta k} \left[ \sum_{m=-\ell}^{\ell} Y^m_\ell(\hat{\mathbf{k}}) Y^m_\ell(\hat{\mathbf{r}}) \right] \delta^*(\mathbf{k})
\]

(\text{Yamamoto et al. 2006; Scoccimarro 2015; Bianchi et al. 2015; Hand et al. 2017, 2018). This incurs an error roughly scaling as \( O(\theta_{\text{max}}^4) \).)

Beyond the Yamamoto approximation, there are multiple options for how to proceed. Clearly, setting the LoS to the direction of just one galaxy not an optimal strategy on large scales, and a full treatment would include the positions of both galaxies, since their pairwise velocity cannot be simply described by a single LoS. This results in a higher-dimensional data-vector however, thus is generally disfavored.

Two primary options exist for defining a single LoS correct to \( O(\theta_{\text{max}}) \), both of which are shown in Fig. 1: (a) using the position vector of the midpoint of the two galaxies, \( \mathbf{n} \propto (\mathbf{r}_1 + \mathbf{r}_2) \) (cf. Scoccimarro 2015; Samushia et al. 2015; Bianchi et al. 2015), or (b) using the bisector of the galaxy-observer-galaxy triangle, \( \mathbf{n} \propto (\mathbf{r}_1 + \mathbf{r}_2) \) (cf. Szalay et al. 1998; Marulli 1999; Szapudi 2004; Yoo & Seljak 2015). These differ only at higher-order,\(^5\) and we will consider both in this work. In full, these are given by

\[
\beta^{\text{midpoint}}_\ell(k) = \frac{2\ell + 1}{V} \int_{\Delta k} \int d\mathbf{r}_1 d\mathbf{r}_2 e^{-i(\mathbf{r}_2 - \mathbf{r}_1) \cdot \mathbf{k}} \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) L^{\ell}_\ell(\hat{\mathbf{k}} \cdot \mathbf{r}_1 + \mathbf{r}_2) \quad (2.9)
\]

\[
\beta^{\text{bisector}}_\ell(k) = \frac{2\ell + 1}{V} \int_{\Delta k} \int d\mathbf{r}_1 d\mathbf{r}_2 e^{-i(\mathbf{r}_2 - \mathbf{r}_1) \cdot \mathbf{k}} \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) L^{\ell}_\ell(\hat{\mathbf{k}} \cdot \mathbf{r}_1 + \mathbf{r}_2).
\]

Notably, neither straightforwardly factorizes into pieces depending only on \( \mathbf{r}_1 \) and \( \mathbf{r}_2 \), making their computation more involved than that of the Yamamoto estimator. A naïve implementation would involve counting all pairs of galaxies individually; this results in an estimator with \( O(N^2) \) complexity for \( N \) galaxies; matching that of the original Yamamoto et al. (2006) estimator before the work of Bianchi et al. (2015).

For upcoming galaxy surveys, such estimators will be prohibitively slow, though shown to be remarkably efficient on small scales in Philcox & Eisenstein 2020 and Philcox 2021. As shown below, we can derive a separable, and hence efficient, estimator using convergent series expansions.

### 2.2 Two-Point Correlation Function

Similar estimators may be derived for the multipoles of the two-point correlation function, \( \xi^{\ell}(r) \). Analogous to (2.5), the general form is given by

\[
\hat{\xi}^{\ell}(r) = \frac{2\ell + 1}{V} \int d\mathbf{r}_1 d\mathbf{r}_2 \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) L^{\ell}_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}}) \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right] , \quad (2.10)
\]

where \( \hat{\mathbf{n}} \) is again the LoS and \( \Delta = \mathbf{r}_2 - \mathbf{r}_1 \) is the separation vector. Here, the square brackets pick out a particular value of the galaxy separation \( \Delta = r \), and, if we substitute

\[
\left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right] \rightarrow \frac{1}{v^a} \Theta^a(r), \quad (2.11)
\]

where \( \Theta^a(r) \) is some binning function with volume \( v^a \), (2.10) becomes the estimator for the 2PCF in a finite bin \( a \). Just as for the power spectrum, we cannot access the overdensity field \( \delta \) directly, and must instead work with galaxies and random particle catalogs. Conventionally, this leads to the 2PCF being computed via the Landy-Szalay estimator (Landy & Szalay 1993), using \( DD, DR \) and \( RR \) counts, each of which is estimated via (2.10). Since such complexities do not depend on our choice of \( \hat{\mathbf{n}} \), we ignore them here, alongside the intricacies of edge-correction.

Two types of correlation function algorithms abound in the literature. Firstly, they are often computed by exhaustive pair counting, scaling as \( N^2 \) (e.g., Sinha & Garrison 2020). Since we explicitly consider each pair of galaxies, any pairwise LoS can be simply included, rendering moot the analysis of this work. However, direct pair-counting can be exceedingly slow for large-datasets, thus it is commonplace to

\(^5\) Specifically, due to the symmetry under \( \mathbf{r}_1 \leftrightarrow \mathbf{r}_2 \) permutations for even \( \ell \), any odd \( \theta \) contribution must vanish, thus the difference between midpoint and bisector formalisms starts at \( O(\theta_{\text{max}}^4) \).
compute the 2PCF via FFT-based approaches. In the plane-parallel approximation of fixed \( \hat{n} \), the estimator may be written

\[
\hat{\xi}^\text{obs.}_\ell(r) = \frac{2\ell+1}{V} \int d\Delta \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right] L_\ell(\hat{\Delta} \cdot \hat{n}) \int dr \delta(r) \delta(r + \Delta)
\]

\[
= \frac{2\ell+1}{V} \int d\Delta \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right] \frac{1}{\sqrt{\Delta}} \left[ \delta(k) \right] (\Delta), \tag{2.12}
\]

where we switch variables to \( r = r_1, \Delta = r_2 - r_1 \) in the first line, and write the \( r \) integral as a convolution in the second. Following the inverse Fourier transform, (2.12) is a weighted real-space summation, which is easy to compute. This again incurs an error at \( O(\theta_{\text{max}}^2) \), where \( \theta_{\text{max}} \) is now the characteristic opening of galaxy pairs separated by the maximum value of \( \tau \) considered.

Similarly to (2.8), the 2PCF may be estimated in the Yamamoto formalism by expanding the Legendre polynomial in spherical harmonics and writing the result as a convolution integral:

\[
\hat{\xi}^\text{Yam.}_\ell(r) = \frac{4\pi}{V} \int d\Delta \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right] \sum_{m=-\ell}^{\ell} Y^m_\ell(\hat{\Delta}) \int dr \delta(r) \delta(r + \Delta) Y^m_\ell(r + \Delta)
\]

\[
= \frac{4\pi}{V} \int d\Delta \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right] \sum_{m=-\ell}^{\ell} Y^m_\ell(\hat{\Delta}) F^{-1} \left[ \delta^* (k) F \left[ Y^m_\ell \right] (k) \right] (\Delta), \tag{2.13}
\]

(e.g., Slepian & Eisenstein 2016), which is computable in a similar manner to the plane-parallel estimator, now requiring \( 2(2\ell+1) \) FFTs for a given \( \ell \). It again suffers an \( O(\theta_{\text{max}}^2) \) error.

The \( O(\theta_{\text{max}}^2) \) pairwise 2PCF estimates can be written in an analogous form to (2.10):

\[
\hat{\xi}^\text{midpoint}_\ell(r) = \frac{2\ell+1}{V} \int d\Delta \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right] \delta_D(r_1 + \Delta) \int dr_2 \delta(r_2) L_\ell(\hat{\Delta} \cdot \hat{r}_1 + \hat{r}_2),
\]

\[
\hat{\xi}^\text{bisector}_\ell(r) = \frac{2\ell+1}{V} \int d\Delta \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right] \delta_D(\hat{r}_1 + \hat{r}_2) \int dr_2 \delta(r_2) \int\frac{d\Delta}{V} \delta_D(r_1 + \Delta) \int dr_2 \delta_D(r_2) L_\ell(\hat{\Delta} \cdot \hat{r}_1 + \hat{r}_2),
\]

Unlike the Yamamoto estimator, neither of these can be straightforwardly recast as a convolution, and thus evaluated via FFTs. It is however possible via series expansions, as will be discussed below.

3 MIDPOINT FORMALISM: POWER SPECTRUM

3.1 Series Expansion

To obtain an efficient power spectrum algorithm within the midpoint formalism, it is necessary to perform a series expansion on the angular dependence, \( Y^m_\ell(r_1 + r_2) \), such that the estimator can be recast in a form conducive to FFT application. To motivate this, we start with the definition (2.9) in terms of the new variables \( \Delta = r_2 - r_1 \) and \( R = r_1 + r_2 \), using the addition theorem (A5) to write the Legendre polynomial in terms of spherical harmonics:\footnote{We note that it is formally possible to first perform the \( \Omega_k \) integral analytically, which leads to the replacement \( \int\frac{d\Omega_k}{4\pi} e^{-iK \cdot \hat{k}} L_\ell(\hat{\Delta} \cdot \hat{k}) \rightarrow (-i)^\ell \delta_{2\ell} (k\Delta) L_\ell(\hat{\Delta} \cdot \hat{k}) \) using (A11) \& (A12). However, this is not exact when the finite nature of the \( k \)-space grid is considered, thus will not be adopted here.}

\[
\tilde{p}^\text{pair}(\ell) = \frac{4\pi}{V} \sum_{m=-\ell}^{\ell} \int\frac{d\Omega_k}{4\pi} Y^m_\ell(\hat{k}) \int dr_1 \int dr_2 \delta(r_1) \delta(r_2) \delta_D(r_1 + \Delta) Y^m_\ell(\hat{R}). \tag{3.1}
\]

If the spherical harmonic factor \( Y^m_\ell(\hat{R}) \) can be written in a form separable in \( \Delta \) and \( r_1 \), the above expression can be evaluated as a convolution, allowing for acceleration by way of Fourier transforms (analogous to the 2PCF manipulations in (2.12) \& (2.13)). Such an expansion is indeed possible, via the spherical harmonic shift theorem, which states

\[
Y^m_\ell(a + A) = \sum_{\alpha=0}^{\ell} \sum_{J_1=0}^{\ell_1} \sum_{J_2=0}^{\ell_2} \sum_{M=-J_1}^{J_1} \sum_{M=-J_2}^{J_2} \varphi_{J_1J_2M}^{a,\ell_1\ell_2m} \left( \frac{a}{A} \right) \alpha J_1^M J_2^M (\hat{a}) Y^m_\ell(\hat{A}). \tag{3.2}
\]

(B7), for arbitrary vectors \( a, A \) with \( a \ll A \). This is proved in Appendix B and is a major new result of this work. Essentially, (3.2) is an infinite expansion in terms of two spherical harmonics and the (small) ratio of \( |a| \) and \( |A| \), giving an arbitrarily accurate approximation of \( Y^m_\ell(a + A) \) if truncated at sufficiently large \( \alpha \). This uses the numerical coefficients \( \varphi_{J_1J_2M}^{a,\ell_1\ell_2m} \) given in (B8), which may be pre-computed and obey the relations

\[
\begin{align*}
\varphi_{J_1J_2M}^{a,\ell_1\ell_2m} &= 0 \quad \text{if } J_1 + \alpha + \text{ either } J_2 + \alpha = \ell + \ell_2 + \ell_1 = \ell_2 = 0 \quad \text{or } J_1 + \alpha = 0 \\
\varphi_{J_1J_2M}^{0,\ell_1\ell_2m} &= \sqrt{4\pi} \delta^K_{0,0} \delta^K_{0,j} \delta^K_{0,M_0} \\
\varphi_{J_1J_2M}^{\alpha,0,0} &= \sqrt{4\pi} \delta^K_{0,0} \delta^K_{0,j} \delta^K_{0,M_0} \\
\end{align*}
\]

(where \( \delta^K_{ij} \) is the Kronecker delta, equal to unity if \( i = j \) and zero else), as proved in Appendix B2.
In our context, we may use (3.2) to expand $Y^m_\ell(\hat{\mathbf{R}})$ using $\textbf{R} = 2\textbf{r}_1 + \Delta \text{ or } \textbf{R} = 2\textbf{r}_2 - \Delta$, recalling $\Delta \ll \textbf{R}$, i.e. that $r_1 \approx r_2$. Suppressing summation limits for clarity, let us

$$Y^m_\ell(\hat{\mathbf{R}}) = \sum_{\alpha J_1 J_2 M} \varphi_{J_1 J_2 M} \left( \frac{\Delta}{2 \textbf{r}_1} \right)^\alpha Y^M_{J_1} (\hat{\lambda}) Y^{m-M}_{J_2} (\hat{\textbf{r}}_1) = \sum_{\alpha J_1 J_2 M} \varphi_{J_1 J_2 M} (-1)^J \left( \frac{\Delta}{2 \textbf{r}_1} \right)^\alpha Y^M_{J_1} (\hat{\lambda}) Y^{m-M}_{J_2} (\hat{\textbf{r}}_2),$$

(3.4)

using $Y^M_{J_1} (\hat{\lambda}) = (-1)^J Y^M_{J_1} (\hat{\lambda})$. The two may be combined to give the symmetrized form:

$$Y^m_\ell(\hat{\mathbf{R}}) = \frac{1}{2} \sum_{\alpha J_1 J_2 M} \varphi_{J_1 J_2 M} (\hat{\lambda}) \left[ \varepsilon_1 Y^{m-M}_{J_1} (\hat{\textbf{r}}_1) + (-1)^J \varepsilon_2 Y^{m-M}_{J_2} (\hat{\textbf{r}}_2) \right],$$

(3.5)

where $\varepsilon_i = \Delta/(2r_i)$. Each term is fully separable in $\hat{\lambda}$ and $\hat{\textbf{r}}$, and the expansion is simply a power series in $\varepsilon_i$, which is closely related to the pair opening angle $\theta = \Delta/R$. Indeed, including all terms up to $\alpha = K$ gives an approximation incurring an error only at $O(\varepsilon^{K+1})$. At lowest order ($\alpha = 0$), we have $\varphi_{J_1 J_2 M} = \delta_{J_1 0} \delta_{J_2,s} \delta_{M,0} \times \sqrt{4\pi} (3.3)$, thus

$$Y^m_\ell(\hat{\mathbf{R}}) = \frac{1}{2} \left[ Y^m_\ell(\hat{\textbf{r}}_1) + (-1)^J Y^m_\ell(\hat{\textbf{r}}_2) \right];$$

(3.6)

this simply yields the Yamamoto estimator in symmetrized form.

As an example, we consider the expansion of $Y^2_2(\hat{\mathbf{R}})$, including all terms up to $\alpha = 2$. From the above expressions, we obtain

$$2 Y^2_2(\hat{\mathbf{R}}) = \left\{ Y^2_2(\hat{\textbf{r}}) \right\}$$

(3.7)

with zeroth-, first- and second-order pieces shown in red, orange and green. Whilst this is lengthy, it is nonetheless computationally tractable.

3.2 Implementation

The above series expansion may be used to write the pairwise power spectrum estimator as a series of convolutions. Inserting (3.5) into (3.1) gives

$$\tilde{p}_\ell^{\text{midpoint}}(k) = \frac{4\pi}{V} \sum_{m=-\ell}^{\ell} \int_{\mathbf{k}_1} Y^m_\ell(\hat{\mathbf{k}}) \sum_{\alpha J_1 J_2 M} \varphi_{J_1 J_2 M} \int d\Delta e^{-i\mathbf{k} \cdot \Delta} \frac{Y^{m-M}_{J_1}(\hat{\Delta})}{\Delta^\alpha} \right] + (-1)^J \int d\textbf{r}_2 \frac{Y^{m-M}_{J_2}(\hat{\textbf{r}}_2)}{(2\textbf{r}_2)^\alpha} \right].$$

Relabelling $\Delta \rightarrow -\Delta$, $\textbf{r}_2 \rightarrow \textbf{r}_1$ and $\mathbf{k} \rightarrow -\mathbf{k}$ in the second term shows that the two are equivalent up to a factor $(-1)^J$. Here and henceforth we will assume even $\ell$, allowing us to absorb the symmetrization:

$$p_\ell^{\text{midpoint}}(k) = \frac{4\pi}{V} \sum_{m=-\ell}^{\ell} \int_{\mathbf{k}_1} Y^m_\ell(\hat{\mathbf{k}}) \sum_{\alpha J_1 J_2 M} \varphi_{J_1 J_2 M} \left[ (-1)^J \int d\Delta e^{-i\mathbf{k} \cdot \Delta} \frac{Y^{m-M}_{J_1}(\hat{\Delta})}{\Delta^\alpha} \right] + \int d\textbf{r}_2 \frac{Y^{m-M}_{J_2}(\hat{\textbf{r}}_2)}{(2\textbf{r}_2)^\alpha} \right].$$

(3.8)

Next, we note that the $\textbf{r}_2$ integral can be written as a convolution:

$$\delta_{J_2(m-M)}(\mathbf{k}) \equiv (-1)^J \int d\textbf{r}_2 \frac{\delta(\textbf{r}_2)(2\textbf{r}_2)^{-\alpha} Y^{m-M}_{J_2}(\hat{\textbf{r}}_2)}{\delta(\textbf{r}_2) - \Delta} = (-1)^{J+1} \int d\textbf{r}_2 \frac{1}{(2\textbf{r}_2)^{-\alpha} Y^{m-M}_{J_2}(\hat{\textbf{r}}_2)}.$$

(3.9)

(3.10)

An alternative approach would be to find an expansion of $Y^m_\ell(\hat{\mathbf{R}})$ that is separable in $r_1$ and $r_2$ rather than $r_1$ and $\Delta$. In this case, the power spectrum could be computed in the same manner as in the Yamamoto approximation, however, this is more difficult to obtain since $r_1/r_2$ is order unity, so one cannot simply apply (3.2). An approach similar to this will prove useful for the bisector formalism however ($\S5$).

Note the distinction between $\theta$, the opening angle for a particular pair, and $\theta_{\text{max}}$, the characteristic opening angle, which is fixed for a particular analysis.

We replace $(-1)^J$ with $(-1)^J$ for later convenience; for even $\ell$, these must have the same sign, due to the parity-rules on $\varphi_{J_1 J_2 M}$.  

---

7 An alternative approach would be to find an expansion of $Y^m_\ell(\hat{\mathbf{R}})$ that is separable in $r_1$ and $r_2$ rather than $r_1$ and $\Delta$. In this case, the power spectrum could be computed in the same manner as in the Yamamoto approximation, however, this is more difficult to obtain since $r_1/r_2$ is order unity, so one cannot simply apply (3.2). An approach similar to this will prove useful for the bisector formalism however ($\S5$).

8 Note the distinction between $\theta$, the opening angle for a particular pair, and $\theta_{\text{max}}$, the characteristic opening angle, which is fixed for a particular analysis.

9 We replace $(-1)^J$ with $(-1)^J$ for later convenience; for even $\ell$, these must have the same sign, due to the parity-rules on $\varphi_{J_1 J_2 M}$.
and the $\Lambda$ integral is then just a Fourier transform:

$$
\hat{p}_{\text{midpoint}}(k) = \frac{4\pi}{V} \sum_{m=-\ell}^{\ell} \int_{\Omega_k} y_{m,\ell}^{\text{mid}}(k) \sum_{\alpha \in J_1, J_2, M} \psi_{\alpha, \ell, m}^J(\Lambda) \Lambda^\alpha s_{J_1, J_2, (m-M)}(\Lambda) (k).
$$

(3.11)

Calculation of the spectra is thus reduced to computing a convolution for each $\{\alpha, J_1, J_2\}$ triplet, then performing a Fourier transform for each $(\ell m)$ pair. We note that the summation over $\{\alpha, J_1, J_2\}$ can be moved inside the $k$-space Fourier transform; we separate it here for clarity. For $\alpha = 0$, we require $J_1 = 0$, $J_2 = \ell$, $M = 0$ as before; this implies

$$
g_{\ell m}^0(\Lambda) = (-1)^{\ell - f - 1} \left[ \delta^o(\Lambda) F \left[ Y_{\ell m}^o \right](k) \right](\Lambda),
$$

(3.12)

and, since $Y_{\ell j}^M(\Lambda) \to (4\pi)^{-1/2}$, the estimator is equal to that of Hand et al. (2017), as expected.

When implementing (3.11), we must be aware of a certain subtlety. Our formalism requires the Fourier transform of a function of $\Lambda$ weighted by $\Delta^\alpha$ where $\alpha \geq 0$. If $s_{J_1, J_2, (m-M)}(\Lambda)$ decays slower than $\Delta^{-\alpha}$, we will obtain an integrand that, in the limit of infinite survey volume, is not square integrable, i.e. it diverges at large $\Lambda$. This is particularly clear when the mean survey distance, $\bar{r}$ is large; in this case $\Delta/(2r) \approx \Delta/(2\bar{r})$, which is a monotonically increasing function of $\Delta$. Whilst the infinite volume limit is somewhat academic (since our expansion parameter $\theta_{\text{max}}$ cannot be assumed small), for finite volumes, we note that the magnitude of $\Delta^\alpha s_{J_1, J_2, (m-M)}$ towards the survey edges for $\alpha > 0$. As such, it is important that we consider the full extent of the $\Lambda$-space function. For a survey of characteristic width $L$, the convolved width is $2L$, thus this requires us to use a grid at least twice the survey width when painting particles. This restriction reduces the efficiency of the algorithm, since we must double the number of grid cells per dimension to obtain the same Nyquist frequency.

The selection rules on $(J_1, J_2, M)$ allow us to quantify the method’s complexity. In general, if one wishes to compute all even power spectrum multipoles up to $\ell_{\text{max}}$ using even $(\text{odd})$ $\alpha = 0, \alpha$ diagonal $g_{\ell m}^\alpha$, functions with even (odd) $J'$ up to $J' = \ell_{\text{max}} + \alpha$; a total of $(1 + (\ell_{\text{max}} + \alpha)2^2$, each of which requires two Fourier transforms. The Fourier transform over $\Lambda$ can then be performed just once per $(\ell m)$ pair $(1 + \ell_{\text{max}}/2)^2$ times. As a concrete example, computing the spectra up to $\ell_{\text{max}} = 4$ requires $21 g_{\ell m}^{\alpha, J}$ coefficients for $\alpha = 1$, and $28$ for $\alpha = 2$.

### 3.3 Parity-Even Form

Whilst the above estimator is mathematically valid, closer inspection reveals a curious property; it contains terms both odd and even in the small angles $\epsilon_i$. Since (for even $\ell$) the power-spectrum definition is symmetric under permutation of $\mathbf{r}_1$ and $\mathbf{r}_2$, we would expect any contributions of $O(\theta^\ell)$ to vanish for odd $K$ (recollecting $\theta = \Delta/R$). In fact, this is the case, and the above expansion can be recast in a manner to make this manifest. Below, we consider a straightforward way to achieve this, based on iterated infinite sequences. An alternative method, which is less obvious $a$ priori, but simpler to implement, is described in Appendix D.

In order to demonstrate that the odd terms in $\theta$ vanish, we first consider the term

$$
e_1^{\alpha y_{\ell m}^m}(\mathbf{r}_1) = (-1)^{J}\epsilon_1^{\alpha y_{\ell m}^m}(\mathbf{r}_2) \right) + (1\to J^{y_{\ell m}^m}(\mathbf{r}_1) - (1\to J^{y_{\ell m}^m}(\mathbf{r}_2)
$$

appearing in (3.5). At lowest order in $\theta$, $e_1 \approx e_2 \approx \theta$ and $r_1 \approx r_2 \approx r$, giving

$$
e_1^{\alpha y_{\ell m}^m}(\mathbf{r}_1) = (-1)^{J}e_1^{\alpha y_{\ell m}^m}(\mathbf{r}_2) \right) + (1\to J^{y_{\ell m}^m}(\mathbf{r}_1) - (1\to J^{y_{\ell m}^m}(\mathbf{r}_2)
$$

(3.13)

For even $J_1$, the contribution is $O(\theta^o)$, yet for odd $J_1$, the term starts only at $O(\theta^{o+1})$. Since $J_1 + \alpha$ must be even for $\psi_{\alpha, J_1, \ell m}^J$ to be non-zero, we find that even $\alpha$ terms in (3.5) (i.e. those with even powers of $\epsilon_1$) begin to contribute at $O(\theta^o)$, whilst the leading-order piece of odd $\alpha$ terms vanishes, and their contribution starts at $O(\theta^{o+1})$. This is not sufficient to demonstrate that there are no terms odd in $\theta$ however, since, the term containing $e_1^{\alpha}$ could include a non-cancelling $e_2^{\alpha}$ contribution for example.

To obtain a manifestly parity-even expansion, we first split the $\alpha$ summation of (3.5) into even and odd pieces, noting that the latter contributions must start at $O(\theta^{o+1})$, viz. the above discussion:

$$
Y_{\ell m}^o(\mathbf{R}) = \sum_{\text{even } \alpha \in J_1, J_2, M} \psi_{\alpha, \ell, m}^J(\Lambda) \left[ e_1^{\alpha y_{\ell m}^m}(\mathbf{r}_1) + e_2^{\alpha y_{\ell m}^m}(\mathbf{r}_2) \right] + \sum_{\text{odd } \alpha \in J_1, J_2, M} \psi_{\alpha, \ell, m}^J(\Lambda) \left[ e_1^{\alpha y_{\ell m}^m}(\mathbf{r}_1) - e_2^{\alpha y_{\ell m}^m}(\mathbf{r}_2) \right] \right).
$$

(3.14)

Next, we rewrite the odd-parity piece by expanding the $\mathbf{r}_2$ term around $\Lambda = 0$, which allows us to explicitly cancel the lowest-order piece. To do so, we require a generalized version of the spherical harmonic shift theorem, proved in appendix C:

$$
\left( \frac{1}{|A + a|} \right)^{\alpha} Y_{J_1}^l(\hat{\mathbf{a}}) Y_{J_2}^{-l}(\mathbf{a} + \hat{\mathbf{A}}) = \sum_{\beta \in S_{12} \in T} \Omega_{\beta}^{J_1 J_2 m M} \left( \frac{\alpha}{\beta} \right) Y_{J_1}^l S_1 S_2 T^T(\hat{\mathbf{a}}),
$$

(3.15)

where the $\Omega$ coefficients are given in (C3), and the $S_i$ summation is limited to the range $[\max(0, J_i - \beta), J_i + \beta]$. Notably, $\Omega_{\beta}^{J_1 J_2 m M} = \delta^K_{J_1 J_2} \delta^K_{S_1 S_2 T} \delta^K_{S_1 S_2 T}$ from (C4). Applying this to $e_2^{\alpha y_{\ell m}^m}(\mathbf{r}_1) Y_{\ell m}^m(\mathbf{r}_2)$ with $\mathbf{a} = \Lambda$, $\mathbf{A} = \mathbf{r}_1$ yields

$$
e_2^{\alpha y_{\ell m}^m}(\Lambda) Y_{\ell m}^m(\mathbf{r}_2) = e_1^{\alpha y_{\ell m}^m}(\Lambda) Y_{\ell m}^m(\mathbf{r}_1) + \sum_{\beta \in S_{12} \in T} \Omega_{\beta}^{J_1 J_2 m M} \epsilon^{\alpha + \beta} \epsilon_2^{\alpha y_{\ell m}^m}(\Lambda) Y_{\ell m}^m(\mathbf{r}_1).
$$

MNRAS 000, 1–27 (2021)
As before, we may expand the relevant, i.e. of convolutions that need to be performed, since there is no longer a requirement to compute removing the asymmetric piece at lowest order. In practice, this means that we have removed all terms proportional to \(e_i\), reducing the total of convolutions that need to be performed, so there is no longer a requirement to compute \(a = 1\).

The pudding is not yet proved, however, since we have not removed terms with, for example, \(e_i\). This is possible via a similar prescription, first noting that the lowest-order piece of the summand in (3.18) (i.e. that with \(r_1 = r_2\)) is non-zero only for odd \(\beta\); a consequence of the parity rules given in Appendix C, restricting \(\beta + J + S_1\) to be even. For even \(\beta\), we again have a cancellation between terms involving \(r_1\) and \(r_2\).

As before, we may expand the relevant \(r_2\) piece in terms of \(r_1\), and cancel the lowest-order contribution. This procedure of “split according to parity, expand, symmetrize” may be iterated, and leads to the following form:

\[
y_{\ell}^{m}(\hat{R}) = \frac{1}{2} \sum_{\text{even } a} \sum_{J,J',M} \Phi_{a,J,-\ell}^{m} + \frac{1}{4} \sum_{a} \sum_{J,J',M} \frac{a \ell}{M} y_{\ell}^{m}(\hat{R}) \left[ e_{1}^{\alpha \beta m}(\hat{R}) + e_{2}^{\alpha \beta m}(\hat{R}) \right]
\]

(3.19)

which contains only even powers of \(e_i\). Whilst each successive line requires more work to evaluate, we note that the only terms up to the second (third) line are required for an expansion correct to third (fifth) order in \(\beta\). Furthermore, we may simplify the above by introducing redefined coefficients \(\Phi_{a,J,M}^{\alpha,\ell}\), such that

\[
y_{\ell}^{m}(\hat{R}) = \frac{1}{2} \sum_{\text{even } a} \sum_{J,J',M} \Phi_{a,J,-\ell}^{m} y_{\ell}^{m}(\hat{R}) \left[ e_{1}^{\alpha \beta m}(\hat{R}) + e_{2}^{\alpha \beta m}(\hat{R}) \right].
\]

(3.20)

for even \(J_1, J_2, \alpha\), using

\[
\Phi_{a,J,M}^{\alpha,\ell} = \omega_{a,J,M}^{\alpha,\ell}
\]

(3.21)

\[
\frac{1}{2} \sum_{a < \alpha} \sum_{J,J',M} \frac{a \ell}{M} y_{\ell}^{m}(\hat{R}) \left[ e_{1}^{\alpha \beta m}(\hat{R}) + e_{2}^{\alpha \beta m}(\hat{R}) \right]
\]

(3.22)

As an example, for \(\ell = 2, m = 1\), the parity-even expansion truncating at \(a = 2\) gives

\[
2y_{2}^{1}(\hat{R}) = \{y_{2}^{1}(\hat{R})
\]

marking zeroth- (second-)order terms in red (green). At fixed maximum order in \(\beta\), this contains significantly fewer terms than (3.7), since all odd powers vanish. We note that \(\Phi_{2,J,M}^{\alpha,\ell} = -\omega_{2,J,M}^{\alpha,\ell}\) here.

Adopting this notation, the full parity-even power spectrum estimator is given by

\[
P^{\text{midpoint}}_{\ell}(k) = \frac{4\pi}{V} \sum_{m=0}^{\ell} \int_{\Omega_{k}} dY_{\ell}^{m}(\hat{k}) \sum_{\alpha} \sum_{J,J',M} \Phi_{a,J,M}^{\alpha,\ell} \int d\Delta e^{-\bar{k} \cdot \bar{\Delta}} Y_{\ell}^{m}(\hat{\Delta}) \int d\bar{r}_{1} \delta(\bar{r}_{1}) \delta(\bar{r}_{1} + \bar{\Delta}) \frac{Y_{\ell}^{m}(\hat{r}_{1})}{(2\bar{r}_{1})^{\alpha}}
\]

(3.23)

This arises naturally in the alternate derivation given in Appendix D.
\[ p_{\text{midpoint}}^\ell (k) = \frac{4\pi}{V} \sum_{m=-\ell}^{\ell} \int_{\Omega_k} Y_m^m (\hat{k}) \sum_{\text{even } \alpha} \sum_{J_1 J_2 M} \Phi_{J_1 J_2 M}^{\alpha, \ell} \left[ Y^M_{J_1} (\Delta) \Delta^\alpha \Phi_{J_2 (m-M)} (\Delta) \right] (k), \] (3.23)

analogous to (3.9), but now requiring only even \( \alpha \) (and thus even \( J_1, J_2 \)), significantly reducing the necessary number of \( \Phi^\alpha_{J_1 J_2 M} \) fields (3.10).

For \( \ell_{\text{max}} = 2 \), we require 28 (45) functions for \( \alpha = 2 \) (\( \alpha = 4 \)). It is important to note that the parity-even estimator will exhibit somewhat slower convergence than the all-parity form, since we capture only a subset of the pieces containing odd \( \alpha \) (i.e. those that contribute to even \( \alpha < \alpha_{\text{max}} \)), and the expansion formally requires \( 2\epsilon_{\text{max}} \approx 2\delta_{\text{max}} \approx \eta_{\text{max}} \approx 1 \). If the latter conditions are met, both forms are fully convergent.

### 4 MIDPOINT FORMALISM: 2PCF

#### 4.1 Series Expansion

Just as for the power spectrum, we may construct an efficient 2PCF estimator via series expansions coupled with FFTs. In this instance, the angular dependence appears through \( L_{\ell}(\hat{\alpha} \cdot \hat{R}) \); our goal therefore is to expand this in a form separable in \( \Delta \) and \( r_1 \). To this end, we use the Legendre polynomial shift theorem, which states

\[ L_{\ell}(\hat{a} + \hat{\Lambda}) = \sum_{\alpha=0}^{\ell} \sum_{J=\text{max}(\ell-\alpha,0)}^{\ell+\alpha} \frac{2J+1}{2\ell+1} f^\alpha_J \left( \frac{a}{A} \right) \alpha L_J(\hat{a} \cdot \hat{\Lambda}), \] (4.1)

(cf. E4), for \( a \ll A \). This is proved in Appendix E1, and uses the coefficients \( f^\alpha_J \) presented in (E11). In brief, the derivation proceeds by noting that \( L_{\ell}(\hat{a} + \hat{\Lambda}) \) can be written as a sum of spherical harmonics in \( \hat{a} \) and \( \hat{\Lambda} \) using the addition theorem (A5), then expanding \( Y^\alpha_J (\hat{a} + \hat{\Lambda}) \) using the spherical harmonic shift theorem of Appendix B, and simplifying the resulting coefficients. An alternative approach would be to write the Legendre polynomial as a power series then perform a Taylor expansion; this gives the same results.

For our context, we set \( \hat{a} = \hat{\Delta}, \hat{A} = 2\hat{r}_1 \) or \( \hat{a} = -\hat{\Delta}, \hat{A} = 2\hat{r}_2 \) in (4.1) to yield the symmetrized form

\[ L_{\ell}(\hat{\Delta} \cdot \hat{R}) = \frac{1}{2} \sum_{\alpha=0}^{\ell} \sum_{J=\text{max}(\ell-\alpha,0)}^{\ell+\alpha} \frac{2J+1}{2\ell+1} f^\alpha_J \left( \frac{\epsilon_1}{\epsilon_2} \right) \alpha L_J(\mu_1) + (-1)^{\ell+J} \epsilon_2^\alpha L_J(\mu_2), \] (4.2)

where \( \mu_1 = \hat{\Delta} \cdot \hat{r}_1 \) and \( \epsilon_i = \Delta/(2r_i) \). This is markedly simpler than the result for the spherical harmonic shift theorem, and can be equivalently derived by an expansion of \( L_{\ell}(\hat{\Delta} \cdot \hat{R}) \) in powers of \( \hat{\Delta} \cdot \hat{R} \), which is then expanded via the binomial theorem. Key properties of this expansion are discussed in Appendix E1; here we note that the coefficients are non-zero only for even \( \alpha + \ell + J \) and \( f^\alpha_J = \delta^\alpha_J \), such that \( L_{\ell}(\hat{\Delta} \cdot \hat{R}) = L_{\ell}(\mu_1) = L_{\ell}(\mu_2) \) at lowest order, i.e. the Yamamoto approximation.

As an example, we consider the expansion of the \( \ell = 2 \) moment. At fourth order in \( \epsilon_i \) (and hence the opening angle \( \theta \))

\[ 2 L_2(\hat{\Delta} \cdot \hat{R}) = L_2(\mu_1) + \frac{6}{5} \epsilon_1 \left[ L_1(\mu_1) - L_3(\mu_1) \right] + \frac{1}{385} \epsilon_1^2 \left[ 7L_0(\mu_1) - 55L_2(\mu_1) + 48L_4(\mu_1) \right] \]
\[ - \frac{4}{105} \epsilon_1^3 \left[ 9L_1(\mu_1) - 49L_3(\mu_1) + 40L_5(\mu_1) \right] + \frac{1}{385} \epsilon_1^4 \left[ 11L_0(\mu_1) + 165L_2(\mu_1) - 816L_4(\mu_1) + 640L_6(\mu_1) \right] \]
\[ + O(\epsilon_1^5) + (\epsilon_1 \leftrightarrow \epsilon_2), \]

where colors separate the different orders, as before. For even (odd) \( \alpha \), the expansion simply consists of Legendre polynomials of even (odd) order up to \( \alpha + \ell \), weighted by powers of \( \epsilon_i \).

#### 4.2 Implementation

Inserting (4.2) into the 2PCF estimator (2.14) yields

\[ \hat{\xi}_{\ell}^{\text{midpoint}} (r) = \frac{1}{V} \sum_{\alpha} \sum_{J=J_{\text{min}}}^{J_{\text{max}}} f^\alpha_J \int d\hat{r}_1 d\hat{r}_2 \delta(\hat{r}_1) \delta(\hat{r}_2) \left( \frac{\Delta}{2r_2} \right)^{\alpha} \frac{\Delta_{\ell}}{2r_2}, \]

suppressing summation indices for clarity and absorbing the \( \hat{r}_1 \leftrightarrow -\hat{r}_2 \) symmetrization (for even \( \ell \)). Expanding the Legendre polynomial via the addition theorem (A5), we may express the integral in separable form:

\[ \hat{\xi}_{\ell}^{\text{midpoint}} (r) = \frac{4\pi}{V} \sum_{\alpha} \sum_{J=J_{\text{min}}}^{J_{\text{max}}} f^\alpha_J \int d\Delta Y^M_M (\hat{\Delta}) \Delta^{\alpha} \frac{\delta_{\ell}(r-\Delta)}{4\pi r^2} \int d\hat{r}_2 \delta(\hat{r}_2) \delta(r_2 - \Delta) Y^M_M (\hat{r}_2) \frac{Y^M_M (\hat{r}_2)}{2r_2} \] (4.4)

As for the power spectrum, the \( r_2 \) integral is simply a convolution:

\[ \xi_J^\alpha_M (\Delta) = (-1)^J \int d\hat{r}_2 \left[ \delta(\hat{r}_2)(2r_2)^{-\alpha} Y^M_M (\hat{r}_2) \delta(r_2 - \Delta) \right] \]
\[ = (-1)^J \hat{f}^{-1} \left[ \delta^\alpha (k) \hat{f} \right] \left[ (2r)^{-\alpha} Y^M_M (\hat{k}) \delta(k) \right] (\Delta), \] (4.6)
where the $g_{JM}^\alpha(\Delta)$ functions are the same as those in (3.10). In this notation, the 2PCF estimator is given by

$$\hat{g}_{JM}^\alpha(\Delta) = \frac{\Delta_{\text{midpoint}}}{V} \sum_{\alpha J M} f_\alpha f_J \int d\Delta \frac{Y_M(\hat{\Delta}) \Delta^\alpha g_{JM}^\alpha(\Delta)}{4\pi r^2} \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right]. \quad (4.7)$$

which may be computed via a simple summation over real-space pixels, given some $r = |\Delta|$ binning scheme. This is analogous to the estimator obtained within the Yamamoto scheme:

$$\hat{\xi}_\ell^{\text{Yama}}(\Delta) = \frac{4\pi}{V} \sum_{M=-\ell}^{\ell} \int d\Delta \frac{Y_M(\Delta) \Delta^\ell g_{JM}^0(\Delta)}{4\pi r^2} \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right]. \quad (4.8)$$

and simply requires additional 2PCF multipoles to be computed, weighted by powers of $\Delta/(2r)$. Since algorithm this does not require a Fourier transform in $\Delta$ space, we do not need to use an increased boxsize, contrary to the power spectrum case (cf. §3.2). Instead, we require only that the width of the box is at least the maximum galaxy dimension plus the maximum separation of interest, to avoid particle overlap on periodic wrapping. To obtain a 2PCF estimate in multipoles up to $\ell_{\text{max}}$, we need to compute the $g_{JM}^\alpha$ functions with $\alpha \leq K, J \leq \ell_{\text{max}} + \alpha, |M| \leq J$, each of which requires a forward and inverse Fourier transform. For $\ell_{\text{max}} = 4$ case requires 21 (28) $g_{JM}^\alpha$ coefficients for $\alpha = 1$ ($\alpha = 2$) just as for $\hat{P}_\ell(k)$, each of which is computed in $O(N_g \log N_g)$ time for $N_g$ grid points.

### 4.3 Parity-Even Form

We may further simplify the estimator by writing it as a sum over only even $\alpha$ (and thus even multipoles $J$). This is possible via recasting (4.2) in a manifestly parity-even form:

$$L_\ell(\hat{\Delta} \cdot \hat{R}) = \frac{1}{2} \sum_{\alpha = 0}^{\infty} \sum_{J = \text{max}(\ell - \alpha, 0)}^{\ell + \alpha} F_\ell^{\alpha} \left[ e_\ell^{\alpha} L_J(\mu_1) + (-1)^{\ell + J} e_\ell^{\alpha} L_J(\mu_2) \right]. \quad (4.9)$$

(E4), where the corresponding coefficients, $F_\ell^{\alpha}$, are given in (E15). The corresponding estimator is thus

$$\hat{g}_{JM}^{\alpha}(\hat{\Delta}) = \frac{4\pi}{V} \sum_{\alpha J M} F_\ell^{\alpha} \int d\Delta \Delta^\ell Y_M(\Delta) \Delta^\alpha \frac{\delta_D(r - \Delta)}{4\pi r^2} \left[ \frac{\delta_D(r - \Delta)}{4\pi r^2} \right]. \quad (4.10)$$

The derivation of (4.9) is analogous to that of §3.3 and sketched in Appendix E2. As an example, the even-parity expansions of $\ell = 2$ and $\ell = 4$ become

$$2 L_2(\hat{\Delta} \cdot \hat{R}) = L_2(\mu_1) + \frac{1}{35} e_2^4 \left[ -7 L_2(\mu_1) + 55 L_2(\mu_1) - 48 L_4(\mu_1) \right] \quad (4.11)$$

$$2 L_4(\hat{\Delta} \cdot \hat{R}) = L_4(\mu_1) + \frac{10}{77} e_4^4 \left[ 11 L_4(\mu_1) + 165 L_2(\mu_1) - 816 L_4(\mu_1) + 640 L_6(\mu_1) \right] + O \left( e_4^6 \right) + (r_1 \leftrightarrow -r_2)$$

$$- \frac{5}{9009} e_4^4 \left[ 143 L_0(\mu_1) + 4420 L_2(\mu_1) - 52083 L_4(\mu_1) + 110240 L_6(\mu_1) - 62720 L_8(\mu_1) \right] + O \left( e_4^6 \right) + (r_1 \leftrightarrow -r_2),$$

correct to fifth-order in $\theta$. These results are consistent with the second-order calculation of Slepian & Eisenstein (2015, Eqs. 19 & 20), and may be extended to arbitrarily high orders in $\ell$ and $\alpha$, with each order just requiring computation of more $g_{JM}^\alpha(\Delta)$ functions.

### 5 BISECTOR FORMALISM: POWER SPECTRUM

We now consider the estimators for the power spectrum in the bisector formalism. The discussion in this section closely follows that of Castorina & White (2018, Appendix E), but is extended to give closed-form estimators at arbitrary order in $\theta_{\text{max}}$.

#### 5.1 Series Expansion

To derive the midpoint power spectrum estimator (§3), we began by expanding the spherical harmonic $V_\ell^m(\hat{n})$ as a separable function in $\hat{r}_1$ and $\hat{\Delta}$. This was both tractable, since one can write $\hat{n} = (2\hat{r}_1 + \hat{\Delta})$, and favorable, since it led to an explicit expansion in powers of $\varepsilon = \Delta/(2r) \equiv \theta$. When adopting the bisector LoS, we have $\hat{n} \propto \hat{r}_2 + \hat{r}_1$, which does not facilitate as straightforward separation into $\hat{\Delta}$ and $\hat{r}_1$ pieces due to the additional normalization by $|r_1|$ and $|r_2|$. For this reason, we take a different approach, proposed by Castorina & White (2018), whereupon we first expand instead in $r_1$ and $r_2$, then in the small angle $\theta$.

Whilst the explicit details of this can be found in Appendix F, for now we note that the derivation proceeds by writing the bisector vector
\( \mathbf{d} = r_1 + (1 - t)r_2 \), where \( t = r_2/(r_1 + r_2) \). Expanding the spherical harmonic \( Y^m_\ell(\mathbf{d}) \) in terms of \( r_1 \) and \( (1 - t)r_2 \) via the solid harmonic addition theorem (A10) gives the finite series

\[
Y^m_\ell(\mathbf{d}) = \sum_{\ell=0}^{\infty} \sum_{\mu=-\ell}^{\ell} A^\mu_{\ell m} \left( \frac{r_1}{d} \right)^{\ell-\ell'} Y^{m-\mu}_{\ell'}(\mathbf{r}_2),
\]

(5.1)

with \( d = |\mathbf{d}| \) and coupling coefficient \( A^\mu_{\ell m} \) defined in (B1). Due to the \((r_1/d)^{-\ell'} \) factor, this is not directly separable in \( r_1, \mathbf{r}_2 \); however, we can write

\[
\frac{r_1}{d} = \left[ 2(1 + \cos \phi) \right]^{-1/2},
\]

(5.2)

and expand in the small parameter \((1 - \cos \phi)\), where \( \cos \phi = r_1 \cdot \mathbf{r}_2 \). From the geometry of Fig. 1, it is clear that \( \phi = 2\Delta \) for \( \Delta \ll r_1, r_2 \), thus this expansion is effectively one in \( \theta^2 \) (since \( 1 - \cos \phi \approx 2\theta^2 \)). Omitting lengthy algebra, this leads to the series expansion for \( Y^m_\ell(\mathbf{d}) \):

\[
Y^m_\ell(\mathbf{d}) = \sum_{\beta=0}^{\infty} \sum_{J_1=0}^{\infty} \sum_{J_2=0}^{\infty} \sum_{M=-J_1}^{J_1} B^{J_1J_2M}_{\ell J_1M}(\mathbf{r}_1) Y^{m-M}_{J_2}(\mathbf{r}_2)
\]

(cf. F7), where the coupling coefficients \( B^{J_1J_2M}_{\ell J_1M} \) are defined in (F8). These obey the properties

\[
B^{J_1J_2M}_{\ell J_1M} = 0 \quad \text{if} \quad \ell + J_1 + J_2 \text{ is odd}
\]

(5.4)

\[
B^{J_1J_2M}_{\ell J_1J_2M} = 2^{-\ell} \frac{\Delta^{J_1}}{\ell!} \frac{\Delta^{J_2}}{J_2!} \frac{\Delta^{M}}{M!} \cdot B^{000}_{\ell J_1J_2M} = \sqrt{2\pi} \Delta^{\ell} \delta_{J_10} \delta_{J_20} \delta_{M0}
\]

(cf. Appendix F), with the latter indicating that \( Y^0_0(\mathbf{d}) = 1/\sqrt{4\pi} \), as expected.

Notably, (5.3) is an infinite expansion involving only separable functions of \( \mathbf{r}_1 \) and \( \mathbf{r}_2 \) (and no powers of \(|\mathbf{r}_1|\)), which will allow for an efficient power spectrum estimator to be wrought. The approximation order is controlled by the value of \( \beta \); fixing to \( \beta \leq K \) expands the internal angle up to \((1 - \cos \phi) \Delta K \), which scales as \((\sqrt{2} \theta)^{2K}\). Unlike in the midpoint formalism, the expansion is automatically symmetric in \( \phi \) and hence \( \theta \), thus there is no need for the parity-even manipulations of §3.3. In this case, the symmetries of \( B^{J_1J_2M}_{\ell J_1J_2M} \) require both odd and even \( J_1, J_2 \) (albeit with the restriction \( J_1 + J_2 = \text{even for even} \ell \)), unlike for the midpoint estimator.

As an example, we consider the expansion of \( Y^1_2(\mathbf{d}) \) at \( \beta = 0 \) and \( \beta = 1 \), i.e. the \( O(\theta^0) \) and \( O(\theta^2) \) contributions. This yields

\[
y^1_2(\mathbf{d})_{\beta=0} = \sqrt{\frac{3}{2}} \left[ 3 Y^1_1(\mathbf{r}_1) Y^0_0(\mathbf{r}_2) + \sqrt{15} Y^1_0(\mathbf{r}_1) Y^0_1(\mathbf{r}_2) \right] + (\mathbf{r}_1 \leftrightarrow \mathbf{r}_2)
\]

(5.5)

\[
y^1_2(\mathbf{d})_{\beta=1} = \sqrt{\frac{15}{2}} \left[ 3 Y^1_1(\mathbf{r}_1) Y^0_0(\mathbf{r}_2) + 5 \sqrt{2} Y^1_0(\mathbf{r}_1) Y^1_0(\mathbf{r}_2) + 21 \sqrt{15} Y^1_0(\mathbf{r}_1) Y^0_1(\mathbf{r}_2) - 2 \sqrt{10} Y^2_1(\mathbf{r}_1) Y^0_0(\mathbf{r}_2) \right]
\]

(5.5)

\[
+ 3 \sqrt{5} Y^2_0(\mathbf{r}_1) Y^0_1(\mathbf{r}_2) + 7 \sqrt{30} Y^2_0(\mathbf{r}_1) Y^1_0(\mathbf{r}_2) - 7 \sqrt{2} Y^2_1(\mathbf{r}_1) Y^1_1(\mathbf{r}_2) \right] + (\mathbf{r}_1 \leftrightarrow \mathbf{r}_2).
\]

Several points are of note. Firstly, in this formalism we have the same terms appearing at \( \beta = 0 \) and \( \beta = 1 \); for instance there is a \( Y^1_2(\mathbf{r}_1) Y^0_0(\mathbf{r}_2) \) term sourced both at \( \beta = 0 \) and all higher orders. This differs from the midpoint expansion of §3 in which the \( O(\theta^K) \) part contains a factor \( \epsilon_\Delta^K \). Secondly, setting \( \beta = 0 \) does not recover the Yamamoto formalism (which would have \( Y^1_2(\mathbf{d}) = Y^1_2(\mathbf{r}_1) + Y^1_2(\mathbf{r}_2) \)). This is clear from the non-trivial \( \beta = 0 \) coefficients in (5.4), and indicates that the scheme includes a resummation of higher-order terms, due to the non-perturbative solid harmonic expansion carried out in (5.1). Importantly, this implies that the midpoint formalism with \( \alpha = 2 \) will not equal that for the bisector method with \( \beta = 1 \), even though both are \( O(\theta^{2_{\text{max}}}) \) (the order to which the two LoS definitions agree). The difference is due to different higher-order terms (and exists also between the midpoint expansion and its even-parity equivalent).

Given the simple nature of (5.3) (containing only spherical harmonics in \( \mathbf{r}_1 \) and \( \mathbf{r}_2 \)), it is interesting to consider whether the above approach may be applied also to the midpoint estimator. In this case, we use \( \mathbf{R} = r_1 + r_2 \), with (5.1) becoming

\[
Y^m_\ell(\mathbf{R}) = \sum_{\ell=0}^{\infty} \sum_{\mu=-\ell}^{\ell} A^\mu_{\ell m} \left( \frac{r_1}{R} \right)^{\ell-\ell'} Y^{m-\mu}_{\ell'}(\mathbf{r}_2),
\]

(5.6)

where \( s = r_1/r_2 \), using the definition \( R^2 = r_1^2 + r_2^2 + 2r_1r_2 \cos \phi \). As before, this can be expanded around the point \( \cos \phi = 1 \), via

\[
\left[ 1 + s^2 + 2s \cos \phi \right]^{-\ell/2} \sum_{\ell=0}^{\infty} \sum_{\mu=-\ell}^{\ell} A^\mu_{\ell m} \left( \frac{r_1}{R} \right)^{\ell-\ell'} Y^{m-\mu}_{\ell'}(\mathbf{r}_2),
\]

(5.7)

provided that \((1 + s^2)^2 > 4s, i.e. s > 0\). The difficulty arises from the additional factor of \( s \); we require a separable perturbative expansion of \((1 + s)\ell^{\ell-2\mu} \) yet \( s \) is order unity. It is this complexity (not found in the bisector approach) that yields the \((\Delta, r_1)\) separation of §3.3 more useful for the midpoint LoS definition.

Beyond Yamamoto Correlators

11

MNRAS 000, 1–27 (2021)
5.2 Implementation

By inserting the expansion of (5.3) into (3.1), we obtain an estimator for the power spectrum in the bisector formalism:

\[
\hat{p}_{\text{bisector}}(k) = \frac{4\pi}{V} \sum_{m=-\ell}^{\ell} \int_{\Omega_k} Y^m_\ell(\hat{r}_1) \int dr_1 dr_2 e^{-ik(r_2-r_1)} \delta(r_1) \delta(r_2) \sum_{J,J_1,J_2,M} B_{J,J_1,J_2,M}^{\beta,\ell m} Y^M_{J_1}(\hat{r}_1) Y^{-M}_{J_2}(\hat{r}_2) \tag{5.8}
\]

\[
\Rightarrow \hat{p}_{\text{bisector}}(k) = \frac{4\pi}{V} \sum_{m=-\ell}^{\ell} \int_{\Omega_k} Y^m_\ell(\hat{r}) \sum_{\beta,J_1,J_2,M} B_{J,J_1,J_2,M}^{\beta,\ell m} F_{J_1} Y^M_{J_1}(\hat{r}) \delta(-k) F_{J_2} Y^{-M}_{J_2}(\hat{r}).
\]

In the last line, we have written the estimator in terms of the Fourier transforms of \(Y^M_{J_1}\), resulting in a simple-to-implement estimator. This bears strong similarities to the Yamamoto approximation (but now involves two density-weighted spherical harmonics), and does not include any factors of \(\epsilon_i = \Delta/(2r_i)\). In comparison to the midpoint estimator, the lack of \(\epsilon_i\) terms has the advantage that we do not need to use a wider box-size relative to the standard estimators (cf. §3.2).

Practically, one estimates the power spectra by first computing \(\mathcal{F} \{ Y^M_{J_1}\}(\hat{r})\) for all (odd and even) \(J\) up to \(\ell_{\text{max}} + \beta_{\text{max}}\), giving a total of \((1 + \ell_{\text{max}} + \beta_{\text{max}})(2 + \ell_{\text{max}} + \beta_{\text{max}})/2\) terms, using the symmetry \(\mathcal{F} \{ Y^M_{J_1}\}(-\hat{r}) = (-1)^J \mathcal{F} \{ Y^M_{J_1}\}(\hat{r})\). For \(\ell_{\text{max}} = 4\), this requires 15 (21) terms for \(\beta_{\text{max}} = 0\) (\(\beta_{\text{max}} = 1\)). For modest computational resources, it may be impractical to store all possible \(Y^M_{J_1}\) grids, since each requires substantial memory. If these are computed separately, we note that the total number of relevant \(Y^M_{J_1} Y^{-M}_{J_2}\) combinations is significant; 20 for \(\beta = 0\) and 68 for \(\beta = 1\) at \(\ell = 2\). We caution that this approach may thus be quite computationally intensive.

6 BISECTOR FORMALISM: 2PCF

We finally turn to the two-point correlation function in the bisector formalism. In this case, the angular dependence of the estimator (2.14) is given by \(L_\ell(\hat{A} \cdot \hat{d})\) where \(\hat{d}\) is the bisector vector as before. For this, we do not require a specialized perturbative expansion, instead expanding the Legendre polynomial via the addition theorem (A5) and utilizing the expansion of the previous section (5.3):

\[
L_\ell(\hat{A} \cdot \hat{d}) = \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} Y^m_\ell(\hat{A}) \sum_{\beta=0}^{\ell} \sum_{J_1=0}^{\ell+\beta} \sum_{M=-J_1}^{J_1} B_{J_1,J_2,M}^{\beta,\ell m} Y^M_{J_1}(\hat{r}_1) Y^{-M}_{J_2}(\hat{r}_2), \tag{6.1}
\]

Since we expand \(Y^m_\ell(\hat{d})\) in terms of \(\hat{r}_1\) and \(\hat{r}_2\), contracting with \(Y^m_\ell(\hat{A})\) does not simplify the coefficients (unlike that seen for the midpoint case in §4.1). Whilst this form is appealing since it involves no new derivations or powers of \(\Delta/(2r_i)\), we note that, at least at leading order, it is possible to obtain a series expansion of \(L_\ell(\hat{A} \cdot \hat{d})\) in powers of \(\epsilon_i\) and \(L_\ell(\hat{A} \cdot \hat{r}_i)\) as in the midpoint formalism. This can be done by writing the Legendre polynomial as a power series then performing a Taylor expansion (as in Sleipan & Eisenstein 2015).

To obtain an efficient 2PCF estimator, we insert (6.1) into (2.14), giving

\[
\xi_\ell^{\text{bisector}}(r) = \frac{4\pi}{V} \sum_{m=-\ell}^{\ell} \int dA Y^m_\ell(\hat{A}) \left[ \frac{\delta(p - \Delta)}{4\pi r^2} \right] \sum_{\beta,J_1,J_2,M} B_{J_1,J_2,M}^{\beta,\ell m} \int dr_1 \left[ Y^M_{J_1}(\hat{r}_1) \delta(r_1) \right] \left[ Y^{-M}_{J_2}(\hat{r}_1 + \Delta) \delta(r_1 + \Delta) \right], \tag{6.2}
\]

writing \(r_2 = r_1 + \Delta\). As in §2, the \(r_1\) integral is a convolution and can thus be computed via Fourier methods; specifically

\[
h_{J_1,J_2,M}^{\text{bisector}}(\Delta) = \int dr_1 \left[ Y^M_{J_1}(\hat{r}_1) \delta(r_1) \right] \left[ Y^{-M}_{J_2}(\hat{r}_1 + \Delta) \delta(r_1 + \Delta) \right] = \mathcal{F}^{-1} \{ \mathcal{F} \{ Y^M_{J_1}\}(-\hat{k}) \mathcal{F} \{ Y^{-M}_{J_2}\}(\hat{k}) \} (\Delta). \tag{6.3}
\]

This gives the full form:

\[
\xi_\ell^{\text{bisector}}(r) = \frac{4\pi}{V} \sum_{m=-\ell}^{\ell} \int dA Y^m_\ell(\hat{A}) \left[ \frac{\delta(p - \Delta)}{4\pi r^2} \right] \sum_{\beta,J_1,J_2,M} B_{J_1,J_2,M}^{\beta,\ell m} h_{J_1,J_2,M}^{\text{bisector}}(\Delta). \tag{6.4}
\]

Whilst this does not require powers of \(\epsilon_i\) (as in §4), it is quite computationally expensive to implement since we require separate convolutions for each \((J_1, J_2, m, M)\) quartet; a total of 20 (68) for \(\beta = 0\) (\(\beta = 1\)) at \(\ell = 2\), as before. This may be contrasted from the midpoint case (4.7), which requires estimation of the \(g_1^\beta M\) functions, depending only on a single total angular momentum. However, note that the summation over \(\beta, J_1, J_2, M\) can be performed before the inverse transform, leading to significant expedition.
7 RESULTS

Having established the existence of efficient pairwise power spectrum and 2PCF estimators, we are now ready to implement them. Below, we will consider their application on realistic survey, including a comparison between the two LoS schemes, before first discussing the convergence of the aforementioned series expansions.

7.1 Convergence Tests

Three series expansions have been introduced in this work; one for each of $Y_{\ell}^m(\hat{R})$ and $L_{\ell}^m(\hat{\Delta} \cdot \hat{R})$ (§3.1 & §4.1), relevant to the midpoint estimators, and one for $Y_{\ell}^m(\hat{d})$ (§5.1) for the bisector estimators. To test these, we consider a simple scenario where we fix $\hat{R}$ and vary $\Delta$ such that we scan over the convergence parameter $\theta = \Delta/R$. We compute the three angular statistics using the associated $\mathbf{r}_1, \mathbf{r}_2$ vectors as a function of $\theta$, storing both the true value and the approximation at a given value of $\alpha$ or $\beta$. For the midpoint estimators, we consider both the standard expansions and those after the even-parity transformations of §3.3 & §4.3.

The results are shown in Fig. 2 for $\ell = 2$ and 4. Notably, the fractional error in the spherical harmonic or Legendre polynomial generally decreases as we include more terms in the infinite series, indicating convergence for both the midpoint and bisector formalisms. As expected, the fractional error becomes larger as $\theta$ increases (and thus the survey becomes wider), and, for $\theta = 0.2$ (around twice that of the BOSS BAO scale), the even-parity midpoint expansion is non-convergent. The bisector expansions are found to be highly convergent here, as a consequence of their inherent partial resummations of higher-order terms (cf.§4.1). For the midpoint series, we find that the even parity expansions (involving only spherical harmonics of even order) give somewhat larger fractional errors than their all parity equivalents; this is due to the even parity scheme having more than twice as many even-order terms compared to the odd-parity scheme. The fractional error increases as we include more terms in the infinite series, indicating convergence for both the midpoint and bisector formalisms. As expected, the fractional error becomes larger as $\theta$ decreases when we include more terms in the infinite series, indicating convergence for both the midpoint and bisector formalisms. As expected, the fractional error becomes larger as $\theta$ decreases.

7.2 Application to Mock Galaxy Surveys

To provide a practical demonstration of above algorithms, we apply them to a set of realistic mock galaxy catalogs. For this purpose, we use 24 MultiDark-PATCHY (hereafter ‘PATCHY’) simulations[1] (Rodríguez-Torres et al. 2016; Kitaura et al. 2016), created for the analysis of the twelfth data-release (DR12) (Alam et al. 2017) of the Baryon Oscillation Spectroscopic Survey (BOSS), part of SDSS-III (Eisenstein et al. 2011; DESI Collaboration et al. 2016). The data is split according to the criteria discussed in (Beutler et al. 2017); here we specialize to the patch with the largest number density (and volume); the north Galactic cap in the redshift range $0.2 < z < 0.5$. This has total volume $1.46h^{-3}$Gpc$^3$ and mean redshift $z = 0.38$. The simulations are generated with the cosmology $\{\Omega_m = 0.307115, \Omega_b = 0.048, \sigma_8 = 0.8288, h = 0.6777\}$, and each contains $\approx 4.8 \times 10^9$ simulated galaxies, alongside a random catalog 50× larger.

For each simulation, the mock galaxy positions are painted to a cuboidal grid using random walk (Hand et al. 2018), using FKP weights (Feldman et al. 1994) and triangle-shaped-cell interpolation, with a fiducial value $\Omega_m = 0.31$ used to convert redshifts and angles into comoving coordinates. We use the same grid-generating parameters as in the final BOSS data release but double the box-size (as discussed in §3.2), giving a Nyquist frequency $k_{Nyq} = 0.3h$ Mpc$^{-1}$. All further computations are carried out in Python, making use of the PyFFTW library to implement the algorithms of §3.2, §4.2, §5.2 & §6. When binning spectra, we adopt the parameters $\{k_{\min} = 0.01 h$ Mpc$^{-1}, k_{\max} = 0.25 h$ Mpc$^{-1}, \Delta k = 0.01 h$ Mpc$^{-1}\}, \{r_{\min} = 10 h^{-1}$ Mpc, $r_{\max} = 190 h^{-1}$ Mpc, $\Delta r = 10 h^{-1}$ Mpc$\}$ in Fourier- and configuration-space respectively. In both cases, we use a maximum Legendre multipole of $\ell_{\max} = 4$.[12]

7.2.1 Power Spectrum

Contributions to the power spectrum multipole in the midpoint and bisector formalisms are shown in Fig. 3. The figure displays the contributions as a function of their order in the (assumed small) parameter $\theta_{\max}$, with, for example, the $\beta = 1$ bisector piece giving a contribution which starts at $O(\theta_{\max}^2)$. For the midpoint, the same order is obtained by summing the $\alpha = 1$ and $\alpha = 2$ pieces, or just from the $\alpha = 2$ piece in the even parity formalism of §3.3. As previously mentioned, the $O(\theta_{\max}^2)$ piece is equal to the Yamamoto spectrum in the midpoint formalism, but includes higher-order corrections for the bisector LoS definition. Our first note is that, in all cases, the leading order contribution is subdominant, indicating that the Yamamoto approximation is, in general, a fair approximation for BOSS. Relative to

[11] Publicly available at data.sdss.org/sas/dr12/boss/lss.
[12] Jupyter notebooks containing our analysis pipeline are publicly available at github.com/oliverphilcox/BeyondYamamoto. Our implementation takes ~ 3 hours to analyze the 2PCF and $P_r(k)$ of a single BOSS-like simulation on 4 Intel Skylake processors using $\alpha_{\max} = 4, \beta_{\max} = 2, \text{and } \ell_{\max} = 4$. The analysis requires ~ 150 GB of memory, though this can be significantly reduced at the expense of longer computation time.

MNRAS 000, 1–27 (2021)
the errorbars, any wide-angle corrections represent a few-percent correction at best for this survey.\textsuperscript{13} Whilst they seem somewhat smaller for the bisector LoS, this is primarily due to the additional $O(\theta_{\text{max}}^2)$ terms absorbed into the $\beta = 0$ estimator. For a larger volume survey, the statistical errors shrink, thus these wide-angle effects will grow in importance. Secondly, it is clear that, in the bisector formalism, the quartic terms are of significantly reduced magnitude compared to those at quadratic order; this indicates that the underlying perturbative expansion is well convergent, and matches that found in Fig. 2. For the all parity midpoint case, the convergence is more tenuous; though the results appear convergent for the quadrupole, the case is less clear for the hexadecapole, and the even parity estimators show non-convergence in both cases, with the $\alpha = 4$ term being larger than that with $\alpha = 2$. This may be rationalized by noting that convergence condition in §3.3 is stricter than that of §4.1 ($2(\theta_{\text{max}}^2 \ll 1$ compared with $\theta_{\text{max}} \ll 1$), leading to the poorer behavior. Overall, the plot indicates that our series are only marginally useful for a survey of BOSS width.

To better understand this, we consider a simple test; shifting the \textsc{pathy} mock data radially outwards by $2000h^{-1}\text{Mpc}$, from its initial position, centered at $\sim 1000h^{-1}\text{Mpc}$. This increases the mean galaxy distance by a factor \~3, thus reducing $\theta_{\text{max}}$ by the same factor. This test is of relevance for future surveys such as DESI and Euclid which generally focus on higher redshifts than BOSS. Resulting power spectrum contributions are shown in Fig. 4, and match our expectations; we see a greater difference between the $O(\theta_{\text{max}}^2)$ and $O(\theta_{\text{max}}^4)$ contributions than before, and significantly improved convergence for all estimators. Clearly then, our estimators will perform better in regimes where the characteristic angle $\theta_{\text{max}}$ is smaller. Even in this example, the systematic errors are at the percent level (and again artificially smaller for the

\textsuperscript{13} The fractional error is roughly scale-invariant; the reason for this is not obvious, due to the inherent scale mixing for the power spectrum compared to the 2PCF. This observation is consistent with the conclusions of Castorina & White (2018) however.
Figure 3. Contributions to the pairwise power spectra from the midpoint and bisector formalisms for a BOSS-like sample. The left (middle) panel shows the all-parity (even-parity) midpoint estimator of §3.2 (§3.3), whilst the right gives the bisector approach of §5.2, based on Castorina & White (2018). We show the absolute magnitudes of the relevant terms in the top panel, whilst the bottom gives the ratio to the standard deviation of the Yamamoto spectra. In all cases, we plot the mean of 24 mocks and display the quadrupole (full lines) and hexadecapole (dashed lines), noting that the monopole receives no wide-angle corrections. The red, green and blue lines correspond to corrections of order $\theta_0^{\text{max}}$, $\theta_2^{\text{max}}$ and $\theta_4^{\text{max}}$ in the characteristic angle $\theta^{\text{max}}$ (i.e. $\theta_2^{\text{max}}$ corresponds to $\alpha = 1$ plus $\alpha = 2$ in the all-parity midpoint formalism, $\alpha = 2$ in the even-parity midpoint approach, or $\beta = 1$ for the bisector algorithm). For the midpoint spectra, the $\theta_0^{\text{max}}$ term is the Yamamoto approximation, whilst for the bisector it is the sum of this and higher-order corrections, as discussed in §5.1. In general the corrections are a small fraction of the errorbars, and we find some convergence issues for the midpoint approximations, indicating that the convergence criterion $\theta^{\text{max}} \ll 1$ is not well-satisfied.

The marked difference between the 2PCF and power spectrum begs the question: why does our power spectrum algorithm fare so much worse? Given that we observe similar behavior in the midpoint and bisector formalisms (with a very different expansion scheme), it is unlikely to stem from some algebraic fault. Our justification for this relies on the following observations: (1) the fractional contribution of the higher-order terms to the 2PCF increases as a function of $r$, becoming unity on scales comparable with the survey width, and (2) the power spectrum is an (Bessel function weighted) integral over the 2PCF. The scales on which the small-$\theta$ approximation is most tenuous thus have

14 The method of Slepian & Eisenstein (2016) gives a straightforward approach by which to include these.

7.2.2 2PCF

Fig. 5 displays the analogous results for the 2PCF multipoles of the (unshifted) PATCHY mocks. We reiterate that these do not include window-function corrections for simplicity. The 2PCF presents a very different story to the power spectrum. We see a clear demarcation of the expansion orders, with a high degree of convergence seen for the quadrupole and hexadecapole on all scales tested for both midpoint and bisector estimates. At larger radii the fractional contribution of the higher-order terms increases, and, extrapolating by eye, one would expect the expansions to break down around $300 \sim 400h^{-1}\text{Mpc}$. These results clearly demonstrate that our expansions are performing as expected. As for the power spectrum, the post-Yamamoto corrections are small; $\sim 1\%$ of the error bars on all scales, with a slight reduction at high radius (whereupon cosmic variance dominates).

The method of Slepian & Eisenstein (2016) gives a straightforward approach by which to include these.
an impact on all the $k$-modes of the power spectrum. This is clearly seen from the (more convergent) bisector formalism; whilst the ratio of quartic to quadratic contribution strongly increases as a function of scale for the 2PCF, it is roughly constant for the convergence is significantly more convergent (in the sense that the $O(\theta_{\text{max}}^4)$ term is subdominant to the $O(\theta_{\text{max}}^6)$ piece).

### 7.3 Comparison of Methods

Given that the bisector and midpoint are both valid LoS definitions at $O(\theta_{\text{max}}^2)$ it is important to compare their utility in the context of the algorithms presented in this work. Below we list several differences, both in terms of efficiency and accuracy of their associated power spectrum and 2PCFs.

- **Memory Requirements:** A naïve implementation of the bisector formalism requires holding in memory all possible $F [Y^M \delta](\mathbf{k})$ fields up to a maximum multipole $\ell_{\text{max}} + \beta_{\text{max}}$ (such that we can take their outer product); a total of 21 for the $O(\theta_{\text{max}}^4)$ correction with $\ell_{\text{max}} = 4$. For a modest FFT grid of size 512$^3$ in double precision, this requires 60 GB of memory, which is infeasible on many computer architectures. Whilst one could compute each function ‘on-the-fly’, this will be significantly slower. For the same order of approximation, we require 28 $s_{\text{TM}}^\gamma$ functions (using even multipoles up to $\ell_{\text{max}} + \alpha$), but, since no outer product is required, only one needs to be held in memory at any time. The most efficient CPU-wise implementation of the midpoint algorithm thus requires significantly less memory.

- **Computation Time:** An efficient metric with which to judge the runtime of the two approaches is the number of FFTs required at a given order. Assuming $\ell_{\text{max}} = 4$ and quadratic corrections, the midpoint $P_\ell(k)$ algorithm requires 70 FFTs (two for each $s_{\text{TM}}^\gamma$ function, one for $\delta^\gamma(\mathbf{k})$ and an addition 13 for transforming into $\mathbf{k}$-space), whilst the bisector approach needs just 28 (one per $F [Y^M \delta](\mathbf{k})$ evaluation). The bisector method is thus somewhat faster, though this holds only if high-memory computational resources are available. For the 2PCF, the conclusion is similar, with the midpoint approach requiring 56 FFTs, and the bisector 29 (if memory is no concern).

- **Convergence:** As demonstrated in previous sections, the bisector $P_\ell(k)$ algorithms exhibit stronger convergence than those using the

---

**Figure 4.** As Fig. 3 but for a BOSS-like galaxy sample shifted radially outwards by an additional $2000 h^{-1}\text{Mpc}$ to reduce the $\theta_{\text{max}}$ parameter and enforce greater convergence. In this instance, we use the mean of 5 $\text{SDSS}$ catalogs. As expected, the higher-order contributions are suppressed relative to Fig. 3, and the expansion is significantly more convergent (in the sense that the $O(\theta_{\text{max}}^4)$ term is subdominant to the $O(\theta_{\text{max}}^6)$ piece).
midpoint LoS. This is particularly true for the even-parity approach, which breaks down on BOSS survey scales (also indicating that the Yamamoto approximation works poorly there). For the 2PCF, all approaches converge well on scales of interest.

- **Grid Size**: As discussed in §3.2, the midpoint power spectrum algorithm requires the particles to be placed on an FFT grid at least twice as wide as the survey itself to avoid errors. This requires a finer cell-size for the same Nyquist frequency, and thus slower (and more expensive) computation. This is not a concern for the bisector formalism or the 2PCF algorithms.

- **Zeroth-Order Contribution**: The midpoint formalisms have the useful property that the zeroth-order ($\alpha = 0$) contribution is equal to the well-known Yamamoto approximation. In contrast, the $\beta = 0$ bisector algorithms discussed herein have a more complicated form. This is additionally slower to compute, involving four pairs of spherical harmonics for each $Y_\ell^m(\hat{k})$ function rather than one.

- **Legendre Expansion**: The midpoint 2PCF algorithm can be simply expressed in terms of Legendre multipoles (4.4) unlike the bisector approach, which must use spherical harmonics. This is far simpler to interpret, and requires significantly fewer coupling coefficients.

Overall, it is clear that both approaches can be implemented in an efficient manner, and give sensible results if the characteristic size is not too large. The optimal choice of LoS is left to the user, following the above considerations. In general, both approaches agree at $O(\theta^4_{\text{max}})$ thus this choice is not of particular importance, particularly since any single LoS definition necessarily incurs an $O(\theta^4_{\text{max}})$ error.

8 SUMMARY AND OUTLOOK

This work has presented efficient and practical algorithms for computing the two-point correlators using pairwise lines-of-sight; an extension to the usual single-particle Yamamoto approximation. Considering both the bisector and midpoint angle definitions, we have shown how convergent series expansions may be used to write the 2PCF and power spectrum estimators in a form allowing for implementation via FFTs. Any such correction is a function of the characteristic size, $\theta_{\text{max}}$, which we have treated as a perturbation variable. Utilizing newly derived

---

15 An analogous expression for the bisector is possible (and discussed in Slepian & Eisenstein 2015), but requires a different series expansion.
spherical harmonic and Legendre polynomial shift theorems, we have computed the midpoint corrections at arbitrary order in $\theta_{\text{max}}$, paying close attention to the existence (and excision) of odd-parity terms. For the bisector LoS definition, the work of Castorina & White (2018) has been extended, including higher-order corrections and an efficient 2PCF algorithm. To demonstrate our approach, the methodology has been applied to a set of realistic galaxy catalogs, and shown to give good results for the 2PCF. For the power spectrum, the bisector approach is similarly effective, though the midpoint algorithms suffer somewhat with convergence issues as the expansion variable becomes large. Such series expansions perform best when the angular survey size is relatively small. In the opposing limit, convergence is difficult to achieve, yet also the Yamamoto approximation itself is poor.

For BOSS, the size of the post-Yamamoto effects is small; ~1% of the statistical error. For future surveys such as DESI, the errorbars shrink, though the mean redshift also increases. Following Castorina & White (2018, Fig. 6), we forecast that the estimator-induced wide angle effects discussed herein can be marginally important, especially on larger scales than considered herein. Since much information regarding primordial non-Gaussianity is found at low-$\ell$, implementing estimators beyond the Yamamoto approximation should be seriously considered. As shown above, there is freedom in how best to do this, since both the midpoint and bisector approaches fix the primordial non-Gaussianity is found at low-$\ell$.

Pairwise lines-of-sight are not the perfect solution however. To fully extract information from the large-scale modes, we ought to parametrize the 2PCF by two lines-of-sight (e.g., Pápai & Szapudi 2008; Yoo & Seljak 2015), though this presents difficulties since the power spectrum becomes ill-defined and the dimensionality increases. For surveys below $\sim 10^7$, the pairwise approximations are valid (Samushia et al. 2015), yet we should bear in mind their sub-optimality for future surveys such as SPHEREx (Doré et al. 2014). Even with improved estimators, it is necessary to model the wide-angle effects arising from other sources such as the galaxy selection function, and much work has been done to achieve this goal. Indeed, it may prove more straightforward to forward-model also the post-Yamamoto effects, and use a simpler estimator; given the results of this work, there is little justification for this on efficiency grounds. On a more philosophical note, it is interesting to see how even on the largest scales, where the Universe is most linear, the two-point function is still highly non-trivial. Despite being the most basic cosmological observable, the two-point function still has secrets up its sleeve.

ACKNOWLEDGEMENTS

We thank Emanuele Castorina, Karolina Garcia, David Spergel, and Martin White for insightful discussions. We are additionally grateful to the anonymous referee for an insightful report. OHEP acknowledges funding from the WFIRST program through NNG26P030C and NNN12AA01C. The authors are pleased to acknowledge that the work reported on in this paper was substantially performed using the Princeton Research Computing resources at Princeton University which is consortium of groups led by the Princeton Institute for Computational Science and Engineering (PICSciE) and Office of Information Technology’s Research Computing.

APPENDIX A: USEFUL MATHEMATICAL RELATIONS

We present a number of mathematical results used in this work. Firstly, we give the explicit form for $L_\ell(\mu)$ as a finite polynomial in $\mu$ (with $|\mu| \leq 1$):

$$L_\ell(\mu) = \sum_{n=0}^{[\ell/2]} c_n^\ell \times \mu^{\ell-2n}, \quad c_n^\ell = \frac{(-1)^n (2\ell - 2n)!}{2^n n! (\ell - n)! (\ell - 2n)!}$$

(A1)

(Abramowitz & Stegun 1964, Eq. 22.3.8), where the $2 \times 1$ vectors are binomial coefficients and $[\ell/2]$ indicates the largest integer $\leq \ell/2$. The inverse relation also proves useful:

$$\mu^n = \sum_{\ell=n/2}^{[\ell/2]} c_n^\ell \times L_\ell(\mu), \quad c_n^\ell = \frac{(2\ell + 1)}{2^{n+1}} \frac{\sqrt{n} \Gamma(1+n)}{\Gamma((n-\ell)/2+1) \Gamma((n+\ell+3)/2)} = \frac{(2\ell + 1)n!(\ell + n + 2)!!}{(\ell + n + 2)!!(n-\ell)!!}$$

(A2)

(derived from Gradsteyn & Ryzhik 1994, Eq. 7.126.1 invoking Legendre polynomial orthogonality), where the summation is over all $\ell$ downwards from $n$ in steps of two, and we use the double factorial $n!! = n(n-2)(n-4)\ldots$ for positive integer $n$. Note that Legendre polynomials of even (odd) order depend only on even (odd) powers of $\mu$ (and vice versa). To derive the simplified coefficients on the right-hand-side we have assumed $n$ to be integral, and noted that the summation rules imply that $\ell, n$ have the same sign. Inserting (A2) into (A1) and using orthogonality gives the relation

$$L_\ell(\mu) = \sum_{n=0}^{[\ell/2]} \sum_{L=(\ell-2n)}^\ell c_n^\ell \tilde{c}_{\ell-2n}^L L_L(\mu) \Rightarrow \sum_{n=0}^{[\ell/2]} c_n^\ell \tilde{c}_{\ell-2n}^L = \delta^K_{\ell L},$$

(A3)

where $\delta^K$ is the Kronecker delta.

Legendre polynomials have the generating function

$$\left[1 - 2xt + t^2\right]^{-1/2} = \sum_{N=0}^{\infty} L_L(s)t^N$$

(A4)
Beyond Yamamoto Correlators

(Gradshteyn & Ryzhik 1994, Eq. 8.921), and may related to spherical harmonics via the addition theorem:

\[ L_\ell(\hat{x} \cdot \hat{y}) = \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} Y_\ell^m(\hat{x})Y_\ell^m(\hat{y}) \]  

(A5)

(NIST DLMF, Eq. 14.30.9), where * indicates a complex conjugate and \( Y_\ell^m \) is the spherical harmonic of order \((\ell, m)\), which obeys the symmetries \( Y_\ell^m(-\hat{x}) = (-1)^\ell Y_\ell^m(\hat{x}) \) and \( Y_\ell^m(\hat{x}) = (-1)^m Y_\ell^{-m}(\hat{x}) \) and \( Y_\ell^0 = (4\pi)^{-1/2} \). Furthermore, the spherical harmonics are orthonormal:

\[ \int d\Omega_s Y_\ell^m(\hat{s})Y_\ell'^m(\hat{s}) = \delta_{\ell\ell'} \delta_{mm'} \]  

(A6)

(NIST DLMF, Eq. 14.30.8). An important relation is the product-sum rule for spherical harmonics:

\[ Y_{L_1}^{M_1}(\hat{x})Y_{L_2}^{M_2}(\hat{y}) = \sum_{L=|L_1-L_2|}^{L_1+L_2} \sum_{M=-L}^{L} G_{LM_1L_2LM_2} (-1)^M Y_{L}^{M}(\hat{x})Y_{L}^{-M}(\hat{y}) \]  

(A7)

where \( G \) is the Gaunt integral obeying several selection rules including \( M_1 + M_2 + M_3 = 0 \). This can be derived from spherical harmonic orthogonality (A6) and the definition of \( G \) as the integral over three spherical harmonics (NIST DLMF, Eq. 34.3.22). The Gaunt factor may be written explicitly in terms of Wigner 3-j symbols as

\[ G_{LM_1L_2LM_2} = \frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{4\pi} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \]  

(A8)

(NIST DLMF, Eq. 34.3.22).

Closely related to spherical harmonics are the regular solid harmonics, defined by

\[ R_\ell^m(\mathbf{r}) = \sqrt{\frac{4\pi}{2\ell + 1}} r_\ell^m (\hat{r}) \]  

(A9)

These obey an addition theorem:

\[ R_\ell^m(\mathbf{r} + \mathbf{a}) = \sum_{\lambda=0}^{\ell} \sum_{\mu=-\lambda}^{\lambda} \begin{pmatrix} \ell + m & \ell - m \\ \lambda & \lambda \end{pmatrix}^{1/2} \begin{pmatrix} \ell - m & \ell + m \\ \lambda & \lambda \end{pmatrix}^{1/2} R_\lambda^\mu(\mathbf{r})R_{\lambda \ell - m - \mu}(\mathbf{a}) \]  

(A10)

(Tough & Stone 1977), which is a finite series.

The Rayleigh plane wave expansion gives

\[ e^{i \mathbf{x} \cdot \mathbf{y}} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} i^{\ell} (2\ell + 1) j_\ell(xy) L_\ell(\hat{x} \cdot \hat{y}) = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} i^{\ell} j_\ell(xy) Y_\ell^m(\hat{x}) Y_\ell^{-m}(\hat{y}) \]  

(A11)

(Arlik et al. 2013, Eq. 16.63), where we have used (A5) to arrive at the second equality. Using this, we can prove the following identity:

\[ \int d\Omega_s e^{i \mathbf{x} \cdot \mathbf{y}} L_\ell(\mathbf{x} \cdot \mathbf{z}) = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} i^{\ell} j_\ell(xy) \int d\Omega_s Y_\ell^m(\hat{x}) Y_\ell^{-m}(\hat{y}) Y_\ell^m(\hat{z}) Y_\ell^{-m}(\hat{x}) \]  

(A12)

\[ = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{i^{2\ell} j_\ell(xy)}{2\ell + 1} \sum_{m=-\ell}^{\ell} Y_\ell^m(\hat{x}) Y_\ell^{-m}(\hat{z}) = 4\pi i^{2\ell} j_\ell(xy) L_\ell(\hat{x} \cdot \hat{z}) \]

using (A5) & (A11) in the first line, and (A6) to obtain the second.

An additional class of orthogonal polynomials are Gegenbauer (or ultra-spherical) polynomials, with the generating function

\[ \left[ 1 - 2x + t^2 \right]^{-\alpha} = \sum_{N=0}^{\infty} C_N^{(\alpha)}(x) t^N \]  

(A13)

(Gradshteyn & Ryzhik 1994, Eq. 8.930), where \( C_N^{(\alpha)} \) is the Gegenbauer polynomial (for integer order \( N \)), and \( \alpha = 1/2 \) recovers the Legendre polynomials. These have the explicit form

\[ C_N^{(\alpha)}(x) = \sum_{k=0}^{[N/2]} \binom{N}{k} x^{N-2k} \]  

(A14)

(Abramowitz & Stegun 1964, Eq. 22.3.4), with the special case of \( C_N^{(0)} \) as \( \alpha = 0 \). Note that this is a sum of even (odd) powers of \( x \) for even (odd) \( N \), as for the Legendre polynomials.
APPENDIX B: SPHERICAL HARMONIC SHIFT THEOREM

B1 Derivation

We present the derivation of a useful series expansion for spherical harmonics. First, we consider the function $Y^m_\ell(a + \hat{A})$ for $\epsilon \equiv a/A \ll 1$. Converting this into the solid harmonic $R^m_\ell(a + A)$ then using the addition theorem (A10), we obtain

$$Y^m_\ell(a + \hat{A}) = \frac{\ell}{\ell!} \sum_{m=0}^{\ell} \frac{\epsilon^m}{m!} \frac{\ell}{\ell + m} \left( \begin{array}{c} \ell + m \\ \ell \end{array} \right) \left( \begin{array}{c} \ell + m \\ m \end{array} \right) Y^m_\ell(\hat{a}) Y^{m-\mu}_\ell(\hat{A}),$$

(B1)

defining the coefficients $A^\mu_{\ell m}$ in the second line. This is a finite expansion, allowing the composite spherical harmonic to be expressed in terms of the harmonics of $a$ and $\hat{A}$. However, due to the angular dependence of $|\hat{A} + e\hat{a}|^2$, we require an infinite expansion to separate $a$ and $A$ fully. To proceed, we recognize that the denominator is the generating function for a Gegenbauer series (A13) with $\alpha = \ell/2, \tau = e, s = -\hat{a} \cdot \hat{A}$:

$$\frac{1}{|\hat{A} + e\hat{a}|^2} = \frac{1}{|1 + 2e\hat{a} \cdot \hat{A} + e^2|^2} = \sum_{N=0}^{\infty} \sum_{k=0}^{[N/2]} C_N^{(\ell/2)}(-2\hat{a} \cdot \hat{A}) e^N = \sum_{N=0}^{\infty} \sum_{k=0}^{[N/2]} \frac{C^{(\ell/2)}_{Nk}}{N!} \times (-2\hat{a} \cdot \hat{A}) N^{-2k} e^N,$$

(B2)

where the coefficients $C^{(\ell/2)}_{Nk}$ are defined in (A14). Inserting into (B1) gives:

$$Y^m_\ell(a + \hat{A}) = \sum_{\ell=0}^{\infty} \sum_{m=0}^{\ell} A^\mu_{\ell m} \sum_{N=0}^{[N/2]} \frac{c_{Nk}^{(\ell/2)}}{N!} (-2N^{-2k}) \times e^{2N} Y^m_\ell(\hat{a}) Y^{m-\mu}_\ell(\hat{A}) (\hat{a} \cdot \hat{A}) N^{-2k},$$

(B3)

To simplify this further, we express $(\hat{a} \cdot \hat{A}) N^{-2k}$ in spherical harmonics, using (A2):

$$(\hat{a} \cdot \hat{A}) N^{-2k} = \sum_{L'=(N-2k)/2} \sum_{\ell'=(N-2k)/2}^{\ell} \sum_{M'=-L'}^{L'} \frac{2\ell' + 1}{L'=L} c_{Nk}^{(\ell/2)} Y^M_\ell(\hat{a}) Y^{M'}_\ell(\hat{A}),$$

(B4)

expanding the Legendre polynomial into spherical harmonics via the addition theorem (A5). Finally, we can combine the two spherical harmonics in $\hat{a}$ and $\hat{A}$ using the product-to-sum relation (A7):

$$Y^m_\ell(\hat{a}) Y^{M'}_\ell(\hat{a}) = \sum_{J=|L-L'|}^{\ell+L'} \sum_{M=|M-M'|}^{M+M'} Y^r_\ell J J_1 Y^M_\ell J_1$$

(B5)

$$Y^{m-\mu}_\ell(\hat{A}) Y^{M*}_\ell(\hat{A}) = \sum_{J=|L-L'|}^{\ell-|L-L'|} \sum_{M=|M-M'|}^{M+|M-M'|} Y^r_\ell J J_1 Y^M_\ell J_1,$$

Combining results, we obtain the shift theorem for spherical harmonics:

$$Y^m_\ell(a + \hat{A}) = \sum_{\ell=0}^{\infty} \sum_{m=0}^{\ell} A^\mu_{\ell m} \sum_{N=0}^{[N/2]} \frac{c_{Nk}^{(\ell/2)}}{N!} (-2N^{-2k}) \sum_{L'=N-2k}/2 \frac{2\ell' + 1}{L'=L'} \frac{4\ell' + 1}{2\ell' + 1} c_{Nk}^{(\ell/2)}$$

(B6)

$$\times g^{(m-\mu)}(\ell-\mu)(\ell'-\mu) g^{(m-\mu)}(\ell\mu)(\ell'-\mu) Y^M_\ell J J_1 Y^{M'}_\ell J_1 Y^{m-M}_\ell(\hat{A}),$$

where we have noted that the Gaunt integrals imply $M_1 = \mu + M'$, $M_2 = \mu - M' = m - M_1$ and relabelled $M' \rightarrow M$. Defining the coefficients $\tilde{g}_{J_1 J_2 M}$ where $\alpha = \ell + \lambda$, this can be written more succintly as

$$Y^m_\ell(a + \hat{A}) = \sum_{\alpha=0}^{[N/2]} \sum_{\ell'=\alpha}^{\ell+\alpha} \sum_{M'=\alpha}^{M+\alpha} \sum_{J=|\ell-\alpha|}^{\ell+\alpha} \sum_{J_1=|M-M'|}^{M+M'} g^{(m-\mu)}(\ell\mu)(\ell'-\mu) Y^M_\ell J J_1 Y^{m-M}_\ell(\hat{A}),$$

(B7)

where the summation limits will be elaborated upon in the next section. Truncating (B7) at $\alpha = K$ incurs an error of $O(\epsilon^{K+1})$, thus for $\epsilon < 1$ this is a valid perturbative expansion.

---

16 This is equivalent to that introduced in Garcia & Slepian (2020) for the three-point correlation function except with one position vector set to zero, affording significant simplification.
B2 Coefficient Properties

From (B6), the shift coefficients $\varphi$ are given by

$$
\varphi_{J_1J_2M}^{\alpha,\ell m} = \sum_{\lambda=0}^{2\ell+1} \sum_{\ell+\mu = \lambda} A_{\ell m}^{J_1J_2} \sum_{N=0}^{\infty} \delta^K_{Nk} \sum_{k=0}^{[N/2]} C_{Nk}^{\ell/2} (-2)^{N-2k} \sum_{L=(N-2k)/2} c^L N^{-2k} \frac{4\pi}{2L'+1} \times G_{L'(J_1)}^{M(-\mu)} G_{L'(-J_2)}^{(m-\mu)\mu(M)(M-m)} (-1)^{m+\mu} \mu (\alpha+\mu) (\alpha-\mu-\ell-2k)$$

(B8)

where the Kronecker delta in the first line enforces that $\lambda + N = \alpha$. Including the various coefficient definitions gives the explicit form

$$
\varphi_{J_1J_2M}^{\alpha,\ell m} = \sqrt{4\pi(2\ell+1)(2J_1+1)(2J_2+1)} \frac{\min(\alpha,\ell)}{\max(\ell+\mu,\lambda-\mu)} \frac{1}{(2\ell'+1)\sqrt{\Gamma(\lambda + \alpha - \lambda - 2k)}} \times (-1)^{k+\alpha+\lambda+3m+\mu} 2^{\ell/2} \Gamma(\lambda + \alpha - 2k - L')/2 + 1 \Gamma((\alpha - \lambda - 2k + L')/2 + 1) \Gamma((\alpha - \lambda - 2k + L' + 3)/2) k! \times \frac{\lambda \mu}{L' - 1 \lambda \mu}$

(B9)

\begin{align*}
\text{We consider two special cases. For } \alpha = 0, \text{ we require } \lambda = N = 0, \text{ and thus, the summation limits, } \mu = k = L' = 0. \text{ This gives}
\varphi_{J_1J_2M}^{0,0} = (4\pi)^{3/2} A_{00}^{00} G_{00}^{00} G_{00}^{00} (-1)^{m} M, \quad (B10)
\end{align*}

using $A_{00}^{00} = \sqrt{4\pi}, \, c_{00}^{00} = 1$ and $C_{00}^{00} = 0$. The Gaunt integrals obey 3-j symmetries, such that $G_{L_1L_2L_3}^{M_1M_2M_3}$ is zero unless $|L_1 - L_2| < L_3 < L_1 + L_2$, implying $J_1 = 0$ (and thus $M = 0$), $J_2 = \ell$. In full, we obtain

$$
\varphi_{J_1J_2M}^{0,\ell m} = c^K_{M0} c^K_{M0} \sqrt{4\pi}, \quad (B11)
$$

This gives $Y_0^\ell (a + \hat{A}) \big|_{\alpha = 0} = Y_0^\ell (\hat{A})$, as expected.

Secondly, we consider $\ell = m = 0$. The summation limits enforce $\lambda = \mu = 0$, and, since $C_{Nk}^{00} = \delta^K_{N0} (A14), N = k = L' = 0$ (and thus $\alpha = \lambda + N = 0$), giving

$$
\varphi_{J_1J_2M}^{\alpha,00} = (4\pi)^{3/2} A_{00}^{00} G_{00}^{00} G_{00}^{00} (-1)^{m} M, \quad (B12)
$$

In this case, the triangle conditions imply $J_1 = J_2 = 0$, thus $M = 0$, and

$$
\varphi_{J_1J_2M}^{0,0} = c^K_{00} c^K_{00} \sqrt{4\pi}, \quad (B13)
$$

giving $Y^\ell_0 (a + \hat{A}) = (4\pi)^{-1/2}$, as expected.

It is useful to consider properties relating to parity. Firstly, consider the parity transformation $a \rightarrow -a, \hat{A} \rightarrow -\hat{A}$:

$$
Y_0^\ell (a + \hat{A}) \rightarrow (-1)^{\ell} Y_0^\ell (a + \hat{A}), \quad Y_0^\ell (a + \hat{A}) \rightarrow (-1)^{J_1+J_2} Y_0^\ell (\bar{a} + \bar{\hat{A}}), \quad (B14)
$$

APPENDIX C: GENERALIZED SPHERICAL HARMONIC SHIFT THEOREM

Below, we prove a useful generalization of the spherical harmonic shift theorem discussed in Appendix B. To begin, consider the expression

$$
\left( \frac{|a|}{|A + \hat{a}|} \right)^{\alpha} Y_0^M (\bar{a}) \bar{Y}_0^M (a + \hat{A}) = \frac{1}{|1 + 2\alpha a \cdot A + \epsilon^2|^{\ell/2}} Y_0^M (\bar{a}) \bar{Y}_0^M (a + \hat{A}), \quad (C1)
$$

for some $\bar{a}, \bar{\hat{A}}$ with $\epsilon = a/A \ll 1$. The spherical harmonic in $a + \hat{A}$ can be expanded identically to Appendix B1; with the prefactor (which is again a generator for Gegenbauer polynomials), we obtain the same expression except that $C_{Nk}^{\ell/2}$ is replaced with $C_{Nk}^{\ell/2}$ in (B2) and the
succeeding. In full, we obtain
\[ \left( \frac{|a|}{|A+a|} \right)^{\ell \ell'} J_{\ell 1} (\hat{a}) Y_\ell^m (\hat{a} + \hat{A}) = \sum_{\beta J_1 J_1'} \omega_{\beta J_1 J_1'}^{\ell \ell'} e^{i \alpha m \beta J_1 J_1'} Y_{\ell 1} (\hat{a}) Y_{\ell' 1} (\hat{a}) Y_{\ell 1}^m (\hat{A}), \] (C2)
where \( \omega_{\beta J_1 J_1'} \) is equal to \( \phi_{J_1 J_1'} \) (B8), except that \( \ell \ell' \rightarrow (\ell + \ell') / 2 \), implying that \( \omega_{\beta J_1 J_1'} = \phi_{J_1 J_1'} \). Next, we can combine the two spherical harmonics in \( a \) using the product-to-sum relation (A7) to yield
\[ \left( \frac{|a|}{|A+a|} \right)^{\ell \ell'} J_{\ell 1} (\hat{a}) Y_\ell^m (\hat{a} + \hat{A}) = \sum_{\beta S_1 S_1'} \omega_{\beta S_1 S_1'}^{\ell \ell'} e^{i \alpha m \beta S_1 S_1'} Y_{\ell 1} (\hat{a}) Y_{\ell 1}^m (\hat{A}), \] (C3)
liberally relabeling variables and defining the new coefficient set \( \Omega \). The expression is now in the form of a power series combined with two spherical harmonics, just as in (B7).

By analogous arguments to before, the \( \omega_{\beta_{J_1 J_1}^{0 J_1 J_1'}} \) coefficient requires \( 0 \leq \beta \leq \beta' \) and \( \max(0, J_2 - \beta) \leq J_2 \leq J_2 + \beta \). Coupled with the triangle conditions on the additional Gaunt integral, we obtain \( 0, J_1 \leq J_1 \leq J_1 + \max(0, J_2 - \beta) \leq J_2 \leq J_2 + \beta \), and \( T < S_1 \). Furthermore, we require \( \beta + J_2 \) and \( J_2 + \beta \) to be even, and, from the Gaunt sum, the same \( \beta \) will also be of use. As in (B11), \( \omega_{\beta_{J_1 J_1}^{0 J_1 J_1'}} = \sqrt{\pi} \delta_{J_1 J_1} K(J_1 J_1')T M \), thus \( \Omega_{\beta_{J_1 J_1}^{0 J_1 J_1'}} = \sqrt{\pi} \delta_{J_1 J_1} K(J_1 J_1')T M \). The Gaunt integral requires \( S_1 = J_1 \), thus we obtain
\[ \Omega_{\beta_{J_1 J_1}^{0 J_1 J_1'}} = \delta_{J_1 J_1'} K(J_1 J_1')T M. \] (C4)

APPENDIX D: ALTERNATIVE DERIVATION FOR THE PARITY EVEN EXPANSION

We briefly present an alternative derivation of the parity-even expansion for spherical harmonics, which, whilst less conceptual, results in an easier-to-implement formalism. To begin, consider the function
\[ \Phi = \sum_{\beta \ell m} \beta \ell m a \beta \ell m \theta \ell m (\hat{a}) \theta \ell m (\hat{A}) \theta \ell m (\hat{A}), \]
where \( a \beta \ell m \theta \ell m \) is non-zero unless \( \beta = 0 \) and \( \ell = \ell' \) and \( m = m' \). The case \( \beta = 0 \) \( \ell = \ell' \) \( m = m' \) is non-zero unless \( J_1 = J_1' \) \( \ell = \ell' \) \( m = m' \). The expression is now in the form of a power series combined with two spherical harmonics, just as in (B7).

By analogous arguments to before, the \( a \beta \ell m \theta \ell m \) coefficient requires \( 0 \leq \beta \leq \beta' \) and \( \max(0, J_2 - \beta) \leq J_2 \leq J_2 + \beta \). Coupled with the triangle conditions on the additional Gaunt integral, we obtain \( 0, J_1 \leq J_1 \leq J_1 + \max(0, J_2 - \beta) \leq J_2 \leq J_2 + \beta \), and \( T < S_1 \). Furthermore, we require \( \beta + J_2 \) and \( J_2 + \beta \) to be even, and, from the Gaunt sum, the same \( \beta \) will also be of use. As in (B11), \( a \beta \ell m \theta \ell m = \sqrt{\pi} \delta_{\beta J_1 J_1'} K(J_1 J_1')T M \), thus \( \Omega_{a \beta \ell m \theta \ell m} = \sqrt{\pi} \delta_{\beta J_1 J_1'} K(J_1 J_1')T M \). The Gaunt integral requires \( S_1 = J_1 \), thus we obtain
\[ \Omega_{a \beta \ell m \theta \ell m} = \delta_{J_1 J_1'} K(J_1 J_1')T M \]. (C4)
assuming even $\ell$, recalling that $\omega_{\ell,M} = \varphi_{\ell,M}$ and contracting the spherical harmonics via (A7). We have additionally inserted the definition of $\Omega$ from (C3). As before, this can be written as a simple summation over even $\alpha$:

$$2 \gamma^m_\ell(\mathbf{R}) = \sum_{\text{even } \alpha} \sum_{J_1, J_2, M} \phi_{\alpha, \ell, m}^{J_1, J_2, M} \gamma^M_{J_1} (\hat{\alpha}) \left[ c_1^{\alpha} \gamma^{M-M'}_{J_2} (\hat{\alpha}) + c_2^{\alpha} \gamma^{M-M'}_{J_2} (\hat{\alpha}) \right].$$

with

$$\phi_{\alpha, \ell, m}^{J_1, J_2, M} = \sqrt{4\pi \delta^K_{\ell,0} \delta^K_{M0}}$$

These are equivalent to the $\Phi$ coefficients of §3.3, through derived in a different manner. This is straightforward to extend to higher order, and somewhat faster to compute than the former procedure, since summations are only over even $\alpha$.

**APPENDIX E: LEGENDRE POLYNOMIAL SHIFT THEOREM**

**E1 Derivation**

A similar relation to that of Appendix B may be derived for the Legendre polynomial $L_\ell(\hat{\alpha} \cdot \mathbf{a} + \mathbf{A})$. This can be derived in one of two ways; either by expanding in powers of $\hat{\alpha} \cdot \mathbf{a} + \mathbf{A}$ directly, or via the previously proved spherical harmonic shift theorem. For consistency with the above, we adopt the latter approach, though we note that the two yield consistent results.

To perform such an expansion, we first rewrite the Legendre polynomial in terms of spherical harmonics using the addition theorem (A5) and apply (B7), yielding

$$L_\ell(\hat{\alpha} \cdot \mathbf{a} + \mathbf{A}) = \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} (-1)^m Y^m_\ell(\hat{\alpha}) Y^m_\ell(\mathbf{a} + \mathbf{A})$$

$$= \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} \sum_{J_1, J_2, M} (-1)^m \phi_{\alpha, \ell, m}^{J_1, J_2, M} \left( \frac{a}{\Lambda} \right)^\alpha Y^M_{J_1} (\hat{\alpha}) Y^{M-M'}_{J_2} (\hat{\alpha}),$$

additionally using $Y^m_\ell = (-1)^m Y^{-m}_\ell$. The two spherical harmonics in $\hat{\alpha}$ may be contracted using the product-to-sum relation (A7), giving

$$L_\ell(\hat{\alpha} \cdot \mathbf{a} + \mathbf{A}) = \frac{4\pi}{2\ell + 1} \sum_{\alpha J J'M'} \sum_{m J} (-1)^m \phi_{\alpha, \ell, m}^{J_1, J_2, (m-M')} \gamma^{M-M'}_{J_1, J_2, \mathbf{a}} \left( \frac{a}{\Lambda} \right)^\alpha Y^M_{J_1} (\hat{\alpha}) Y^M_{J_2} (\hat{\alpha}),$$

where we have relabelled $M' = m - M$, and defined the new coefficient $f_{\alpha, \ell, J J'M'}$. Whilst this may seem complicated, one can in fact show that $f_{\alpha, \ell, J}$ is independent of $M'$ and non-zero only for $J = J_2$. Defining reduced coefficients $f_{\alpha, \ell, J}$ via

$$f_{\alpha, \ell, J} = \delta^K_{J_1} \times f_{\alpha, \ell, J},$$

and using (A5), this leads to

$$L_\ell(\hat{\alpha} \cdot \mathbf{a} + \mathbf{A}) = \sum_{\alpha J} \sum_{J = \text{max}(\ell-\alpha, 0)}^{\ell+\alpha} \frac{2J+1}{2\ell+1} f_{\alpha, \ell, J} \left( \frac{a}{\Lambda} \right)^\alpha L_J (\hat{\alpha} \cdot \hat{\Lambda});$$

a far more tractable expansion. Whilst the $(2J+1)/(2\ell+1)$ factor could be absorbed into $f_{\alpha, \ell, J}$, we retain it here for later convenience.
To see this, we first write out $f_{J,J;M'}$, explicitly:

$$f_{J,J;M'} = \sum_{m,J,J',M} A'^{\mu}_{\ell m} \alpha^{\kappa}_{(\ell+\mu)} C_{N,k}^{2} (-2)^{N-2k} e_{N-2k}^{2} \frac{4\pi}{2L'+1}$$

$$\times G_{J, J', M}^{\mu} G_{\ell}^{(m-M')\mu} (\mu_{m-M'}) G_{\ell}^{(n-M')M'} (-1)^{m-M'-\mu}$$

$$= \sum_{\alpha K \ell} \frac{(2\ell+1)^{3/2} \sqrt{(2\ell'+1)(2\ell+1)}}{2L'} \left\{ \sum_{J,J',M} (-1)^{\ell_{J} + \ell_{J'} + \ell_{M} - \mu_{M}} \right\}$$

$$\times \sum_{J} \frac{2\ell_{J} + 1}{2L' + 1} \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ 0 & 0 & 0 \end{array} \right\}$$

$$\times \sum_{J'} \frac{2\ell_{J'} + 1}{2L' + 1} \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ 0 & 0 & 0 \end{array} \right\}$$

$$\times \left\{ \sum_{J''} (-1)^{M'-\mu} \left( \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ \mu & m - \mu & m - \mu \end{array} \right) \sum_{J''} \left( \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ m - m & m - m & m - m \end{array} \right) \right\}$$

where, in the second line, we have written the Gaunt integrals in terms of Wigner 3-j symbols (A8), and noted that

$$A'^{\mu}_{\ell m} = (-1)^{\ell - m} \sqrt{\frac{4\pi (2\ell + 1)^2}{(2\ell + 1)(2\ell + 2)(2\ell + 3)}} \frac{2L'}{2L'} \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ \mu & m - \mu & m - \mu \end{array} \right\}$$

To simplify, we first employ a summation identity for four 3-j symbols, given in Lohmann (2008, Eq. A.24):

$$\sum_{m_{J} m_{J'} m_{M}} (-1)^{j_{J} + j_{J'} + j_{M} + m_{J} + m_{J'} + m_{M}} \left( \begin{array}{ccc} j_{J} & j_{J'} & j_{M} \\ m_{J} & m_{J'} & m_{M} \end{array} \right)$$

$$= \delta_{j_{J} j_{J'}} \delta_{j_{M} j_{m}} \frac{1}{2L' + 1} \left\{ \begin{array}{ccc} j_{J} & j_{J'} & J_{2} \\ m_{J} & j_{m} & J_{2} \end{array} \right\}$$

where the quantity in curly parentheses is a Wigner 6-j symbol (see NIST DLMF, §34.4). Notably, the RHS of the above equation is independent of $m_{J}$ and enforces $J_{2} = J$. Applying this to the final line of (E5), denoted by [...], with $\{ j_{J}, j_{J'}, J_{1}, J_{2}, L', L, \ell - \lambda, \lambda \}$ with some massaging via the 3-j symmetries gives

$$[...] = (-1)^{J_{1}+1+\ell_{J}} \frac{1}{2L' + 1} \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ L' & L' & L' \end{array} \right\}$$

enforcing $J_{2} = J$ and with no dependence on $M'$, as previously mentioned. We may further perform the $J_{1}$ summation analytically, using the 6-j summation relation:

$$\sum_{J_{1}} (-1)^{j_{J} + j_{J'} - j_{J} - j_{J'} + j_{M} - m_{J} - m_{J'}} (2J_{1} + 1) \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ J_{1} & J_{1} & J_{1} \end{array} \right\}$$

$$\left( \begin{array}{ccc} j_{J} & j_{J'} & J_{2} \\ m_{J} & j_{m} & J_{2} \end{array} \right)$$

$$= (-1)^{J'_{2} + J} \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ J'_{2} & \ell & \ell - \lambda \end{array} \right\}$$

(E8)

(Lohmann 2008, Eq. A.26b), which, following application of 6-j symmetries, can be used to show

$$\sum_{J_{1}} (2J_{1} + 1) (-1)^{J_{1}} \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ J_{1} & J_{1} & J_{1} \end{array} \right\}$$

$$\left( \begin{array}{ccc} j_{J} & j_{J'} & J_{2} \\ m_{J} & j_{m} & J_{2} \end{array} \right)$$

$$= (-1)^{J_{2} + J} \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ J_{2} & \ell & \ell - \lambda \end{array} \right\}$$

(E10)

Inserting into (E5) gives the final form for the reduced coefficients $f_{J}^{a,\ell}$:

$$f_{J}^{a,\ell} = (-1)^{\ell} (2\ell + 1)^{3/2} \left( \frac{1}{2L' + 1} \right) \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ J & J & J \end{array} \right\}$$

$$\times \sum_{L_{u}=0}^{L_{u}} \frac{e_{L_{u} L_{u}}}{2L' + 1} \left\{ \begin{array}{ccc} \ell & \ell & \ell - \lambda \\ 0 & 0 & 0 \end{array} \right\}$$

$$\left( \frac{4\pi}{2L' + 1} \right)$$

As before, this satisfies $f_{J}^{a,0} = \delta_{J}^{K} f_{J}^{a,0}$, and $f_{J}^{a,0} = \delta_{J}^{K} \delta_{J}^{a,0}$, which, when inserted into (E4) gives $L_{\ell}(\hat{a} \cdot \mathbf{a} + \hat{A}) = L_{\ell}(\hat{A} \cdot \hat{A})$ at leading order and $L_{0}(\hat{a} \cdot \mathbf{a} + \hat{A}) = 1$. Furthermore, we have the parity-rule $a + \ell + J = \text{even}$ for non-zero $f_{J}^{a,\ell}$ coefficients. We also note a curious property; the sum of $f_{J}^{a,\ell}$ over all $J$ is equal to zero unless $a = 0$, as demonstrated in (4.3) (4.11). To prove this, we set $\lambda = a \mathbf{a}$ in (E4) and recall $L_{\ell}(1) = 1$, giving

$$1 = \sum_{a} \sum_{J} f_{J}^{a,\ell} \left( \frac{1}{k} \right)^{a} = 1 + \sum_{a \geq 0} k^{-a} \sum_{J} f_{J}^{a,\ell}$$

(E12)

For this to be true for arbitrary $k > 1$, we require $\sum_{J} f_{J}^{a,\ell} = 0$ for $a > 0$. 

MNRAS 000, 1–27 (2021)
E2 Parity-Even Form

For the 2PCF estimators considered in §3.3, we use the Legendre polynomial shift theorem (E4) in the following symmetrized form:

\[
L_{\ell}(\hat{A} \cdot \mathbf{r}_1 + \mathbf{r}_2) = \sum_{\alpha=0}^{\ell} \sum_{J=\max(\ell-\alpha,0)}^{\ell+\alpha} \frac{2J+1}{2\ell+1} f_{J,\alpha}^{\ell}(\hat{a} / A) L_{J}(\hat{A} \cdot \hat{A}),
\]

(E13)

for \( \epsilon_i \equiv \Delta/(2r_i) \). Just as for the spherical harmonic expansion (§3.3), these may be recast in a manifestly parity-even form, i.e. one only involving even \( J \) and \( \alpha \) (assuming even \( \ell \)). The derivation of this is analogous to the above, and leads to

\[
L_{\ell}(\hat{A} \cdot \mathbf{r}_1 + \mathbf{r}_2) = \sum_{\alpha=0}^{\ell} \sum_{J=\max(\ell-\alpha,0)}^{\ell+\alpha} \frac{2J+1}{2\ell+1} f_{J,\alpha}^{\ell}(\hat{a} / A) L_{J}(\hat{A} \cdot \hat{A}),
\]

(E14)

for \( 2\epsilon_i \ll 1 \), where the parity-even coefficients are defined recursively via

\[
f_{J,\alpha}^{\ell} = f_{J,\alpha}^{\ell+2} - \frac{1}{2} \sum_{\beta,\gamma} 2^{\alpha-\beta-\gamma} f_{J,\gamma}^{\beta,\gamma} h_{J,\gamma}^{\beta,\gamma} h_{J,\gamma}^{\beta,\gamma} + \frac{1}{4} \sum_{\beta,\gamma} 2^{\alpha-\beta-\gamma} f_{J,\gamma}^{\beta,\gamma} \sum_{\beta,\gamma} \sum_{\gamma=0}^{\alpha-\beta-\gamma} h_{J,\gamma}^{\beta,\gamma} (\beta+\gamma)^T + \ldots\]

(E15)

using the method of §3.3, or equivalently

\[
F_{J,\alpha}^{\ell} = \delta_{J,\alpha}^{\ell},\]

\[
F_{J,\alpha}^{\ell+2} = -f_{J,\alpha}^{\ell+2},\]

\[
F_{J,\alpha}^{\ell+4} = -f_{J,\alpha}^{\ell+4} + \sum_{S} f_{J,\alpha}^{\ell+2} h_{J,\alpha}^{\ell+2S},
\]

et cetera, using that of Appendix D. Here, the \( h_{J,\alpha}^{\beta,\gamma} \) functions are identical to \( f_{J,\alpha}^{\beta,\gamma} \) (E11), but with \( \ell/2 \to (\ell + \alpha)/2 \) in the \( C_{J,\alpha}^{\ell,\ell} \) coefficient.

APPENDIX F: BISECTOR SERIES EXPANSION

To derive the series expansion of §5.1, we start from the bisector vector definition \( \mathbf{d} = t \mathbf{r}_1 + (1-t) \mathbf{r}_2 \) where \( t = r_2/(r_1 + r_2) \), and perform the solid-harmonic expansion (A10):

\[
\psi_{m}(\hat{d}) = \sum_{\mu=0}^{A} \sum_{\ell=0}^{A} A_{\mu}^{\ell}(\hat{r}_1) \hat{r}_1 \psi_{m-\ell}^{\mu}(\hat{r}_2),
\]

(F1)

using \( tr_1 = (1-t) r_2 \), and defining \( A_{\mu}^{\ell} \) as in (B1). Whilst this is not immediately separable in \( r_1, r_2 \) due to the \( d = |\mathbf{d}| \) denominator, we follow Casterina & White (2018) and note that

\[
d^2 = \frac{r_1^2 + r_2^2}{(r_1 + r_2)^2} |\hat{r}_1 + \hat{r}_2|^2 \approx 2r_1^2 (1 + \cos \phi),
\]

(F2)

where \( \cos \phi = \hat{r}_1 \cdot \hat{r}_2 \). Combining with the above:

\[
\psi_{m}(\hat{d}) = [2(1 + \cos \phi)]^{-\ell/2} \sum_{\mu=0}^{A} A_{\mu}^{\ell}(\hat{r}_1) Y_{m-\ell}^{\mu}(\hat{r}_2).
\]

(F3)

The former work proceeded to expand the prefactor to second order in \( \cos \phi \) around \( \cos \phi = 1 \); here, we note that

\[
[2(1 + \cos \phi)]^{-\ell/2} = 2^{-\ell} \left[ 1 + \frac{\cos \phi - 1}{2} \right]^{-\ell/2} = 2^{-\ell} \sum_{\beta=0}^{\ell} \left( -\ell/2 \right) \binom{\ell}{\beta} 2^\beta (\cos \phi - 1)^\beta,
\]

(F4)

via Newton’s generalized binomial theorem around \( \cos \phi = 1 \) (where the \( 2 \times 1 \) vector is a generalized binomial coefficient with \( \binom{n}{a} = a! / (a-a)! \) for falling-factorial Pochhammer symbol \( (a)_n \), defined in NIST DLMF, §5.2). Using the standard binomial theorem, this gives

\[
[2(1 + \cos \phi)]^{-\ell/2} = 2^{-\ell} \sum_{\beta=0}^{\ell} \sum_{n=0}^{\ell} \binom{\ell/2}{\beta} \binom{n}{\beta} (\cos \phi - 1)^\beta\cos^n \phi.
\]

Finally, we may convert \( \cos^n \phi \) into a sum over Legendre polynomials via (A2) and thus spherical harmonics using (A5). This yields

\[
[2(1 + \cos \phi)]^{-\ell/2} = 2^{-\ell} \sum_{\beta=0}^{\ell} \sum_{n=0}^{\ell} \binom{\ell/2}{\beta} \binom{n}{\beta} \cos^n \phi = 2^{-\ell} \sum_{n=0}^{\ell} \sum_{L=\ell/2}^{\ell} \sum_{M=L}^{\ell} \sum_{n=0}^{\ell} \sum_{L=\ell/2}^{\ell} \sum_{M=L}^{\ell} Y_{L}^{M}(\hat{r}_1) Y_{M}^{L}(\hat{r}_2).
\]

(F5)
Combining the spherical harmonics with those in (F1) via (A7) gives the final form for the expansion:

\[
Y^m_\ell (\mathbf{d}) = 2^{-\ell} \sum_{\lambda=0}^{\ell} \sum_{\mu=-\lambda}^{\lambda} A^\lambda_\ell m \left( -\ell/2 \right) \beta J_\lambda \beta \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \frac{2L+1}{2\ell+1} c_n 2 L + 1 \sum_{M=-L}^{L} (-1)^{M-m} M \gamma_{\ell M L} J_{\ell M L} \mathbf{d} (\ell \lambda M) (M-m) M \gamma_{\ell M L} (\mathbf{d}) Y^m_\ell (\mathbf{d}),
\]

defining the coefficients

\[
B^\beta_\ell J_1 J_2 M = 2^{-\ell} \sum_{\lambda=0}^{\ell} \sum_{\mu=-\lambda}^{\lambda} A^\lambda_\ell m \left( -\ell/2 \right) \beta J_\lambda \beta \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \frac{2L+1}{2\ell+1} c_n 2 L + 1 \sum_{M=-L}^{L} (-1)^{M-m} M \gamma_{\ell M L} J_{\ell M L} \mathbf{d} (\ell \lambda M) (M-m) M \gamma_{\ell M L} (\mathbf{d}) Y^m_\ell (\mathbf{d}).
\]

The summation limits on \( J_1, J_2 \) come from the consideration of the Gaunt integral triangle conditions, and the constraints \( L \leq n \leq \beta \), as in Appendix B2. Furthermore, the Gaunt integrals require even \( \lambda + L + J_1 \) and \( \ell - \lambda + L + J_2 \), thus \( \ell + J_1 + J_2 \) is even, but we do not require \( J_1 \) and \( J_2 \) themselves to be even.

For the special case of \( \beta = 0 \), we obtain

\[
B^\beta_\ell J_1 J_2 M = 2^{-\ell} \sum_{\lambda=0}^{\ell} \sum_{\mu=-\lambda}^{\lambda} A^\lambda_\ell m \left( -\ell/2 \right) \beta J_\lambda \beta \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \frac{2L+1}{2\ell+1} c_n 2 L + 1 \sum_{M=-L}^{L} (-1)^{M-m} M \gamma_{\ell M L} J_{\ell M L} \mathbf{d} (\ell \lambda M) (M-m) M \gamma_{\ell M L} (\mathbf{d}).
\]

using \( \left( -\ell/2 \right) = 1, \beta^0_0 = 1 \) and the explicit forms of the Gaunt integral, assuming \( J_1, J_2 \leq \ell \) and \( |M| \leq J_1 \). Notably, \( Y^m_\ell (\mathbf{d}) \) still involves a summation over \( J_1, J_2, M \) at zeroth order in \( \beta \). Considering \( \ell = m = 0 \), the coefficient becomes

\[
B^{\beta=0}_\ell J_1 J_2 M = A^{00}_\ell \beta \sum_{n=0}^{\infty} \frac{2L+1}{2\ell+1} c_n 2 L + 1 \sum_{M=-L}^{L} (-1)^{M-m} M \gamma_{00 L M} J_{00 L M} \mathbf{d} (00) M (M-m) M \gamma_{00 L M} (\mathbf{d}).
\]

using \( A^{00}_\ell = \sqrt{4\pi} \) and \( \beta^0_0 = \delta^\ell_0 \). As expected, this recovers \( Y^0_\ell (\mathbf{d}) = 1/\sqrt{4\pi} \).

REFERENCES

Abramowitz M., Stegun I. A., 1964, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover, New York City

Alam S., et al., 2017, MNRAS, 470, 2617

Arlen G., Weber H., Harris F., 2013, Mathematical Methods for Physicists: A Comprehensive Guide. Elsevier Science

Beutler F., et al., 2012, MNRAS, 423, 3430

Beutler F., et al., 2017, MNRAS, 466, 2242

Beutler F., Castorina E., Zhang P., 2019, J. Cosmology Astropart. Phys., 2019, 040

Bianchi D., Gil-Marín H., Ruggeri R., Percival W. J., 2015, MNRAS, 453, L11

Bonvin C., Durrer R., 2011, Phys. Rev. D, 84, 063505

Castorina E., White M., 2018, MNRAS, 476, 4403

Chudaykin A., Ivanov M. M., 2019, J. Cosmology Astropart. Phys., 2019, 034

DESI Collaboration et al., 2016, arXiv e-prints, p. arXiv:1611.00036

Datta K. K., Choudhury T. R., Bharadwaj S., 2007, MNRAS, 378, 119

Doré O., et al., 2014, arXiv e-prints, p. arXiv:1412.4872

Eisenstein D. J., et al., 2011, AJ, 142, 72

Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23

Garcia K., Slepian Z., 2020, arXiv e-prints, p. arXiv:2011.03503

Gil-Marín H., Percival W. J., Verde L., Brownstein J. R., Chiang C.-H., Kitaura F.-S., Rodríguez-Torres S. A., Omstead M. D., 2017, MNRAS, 465, 1757

Gradshteyn I. S., Ryzhik I. M., 1994, Table of integrals, series and products, Elsevier Science

Hamilton A. J. S., 1992, ApJ, 385, 15

Hamilton A. J. S., 1998, Linear Redshift Distortions: a Review. Kluwer Academic, p. 185, doi:10.1007/978-94-011-4960-0_17

Hamlet A. J. S., Collom C., Munn J., 1996, MNRAS, 278, 73

Hand N., Li Y., Slepian Z., Seljak U., 2017, J. Cosmology Astropart. Phys., 2017, 002

Hand N., Feng Y., Beutler F., Li Y., Modi C., Seljak U., Slepian Z., 2018, AJ, 156, 160

Kaiser N., 1987, MNRAS, 227, 1

Kitaura F.-S., et al., 2016, MNRAS, 456, 4156
This paper has been typeset from a \LaTeX file prepared by the author.