The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Evolution of higher-order correlations demonstrated with Minkowski Functionals

James M. Sullivan,[1,2]★ Alexander Wiegand,[1,3] Daniel J. Eisenstein[1]

[1] *Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA, 02138, USA*
[2] *University of Texas at Austin, 110 Inner Campus Dr, Austin, TX 78705, USA*
[3] *Max–Planck–Institut für Astrophysik, Karl–Schwarzschild–Str. 1, D–85741 Garching, Germany*

29 November 2017

ABSTRACT
We probe the higher-order galaxy clustering in the final data release (DR12) of the Sloan Digital Sky Survey Baryon Oscillation Spectroscopic Survey (BOSS) using germ-grain Minkowski Functionals (MFs). Our data selection contains $979,430$ BOSS galaxies from both the northern and southern galactic caps over the redshift range $z = 0.2 - 0.6$. We extract the higher-order part of the MFs, detecting the deviation from the purely Gaussian case with $\chi^2 \sim O(10^3)$ on 24 degrees of freedom across the entire data selection. We measure significant redshift evolution in the higher-order functionals for the first time. We find $15 - 35\%$ growth, depending on functional and scale, between our redshift bins centered at $z = 0.325$ and $z = 0.525$. We show that the structure in higher order correlations grow faster than that in the two-point correlations, especially on small scales where the excess approaches a factor of 2. We demonstrate how this trend is generalizable by finding good agreement of the data with a hierarchical model in which the higher orders grow faster than the lower order correlations. We find that the non-Gaussianity of the underlying dark matter field grows even faster than the one of the galaxies due to decreasing clustering bias. Our method can be adapted to study the redshift evolution of the three-point and higher functions individually.

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

1 INTRODUCTION

Cosmic large-scale structure is largely understood by characterizing and modeling observed galaxy distributions, which strongly influence the accepted cosmological model. Recent galaxy redshift surveys contain spectroscopic redshifts for of order $10^5 - 10^6$ galaxies, and their size makes them ideal testing grounds for exploring non-Gaussian features in this structure. Any quantitative analysis of such features must be firmly rooted in precise statistical measures. The standard analysis employs the two-point correlation function ξ_2, but the condensed information offered by this function fails to fully account for fundamental observed structures such as sheets and filaments. To accurately describe large-scale structure we must then look beyond two-point statistics to higher-order correlation information.

The simplest approach is to calculate the higher-order

correlation functions directly. While this has been widely achieved for the two-point function, and recently quite successfully for the three-point function (e.g. by Slepian & Eisenstein (2015) and Slepian et al. (2017)), fourth- and higher-order functions have yet to be well-determined. As it turns out, computing these functions directly is computationally infeasible. There are various alternatives, and one of the most useful and rigorous is Minkowski Functional (MF) analysis. MFs quantitatively describe the geometry of extended bodies by mapping the shape of a body to the real numbers. These functionals uniquely characterize the geometry and topology of a galaxy distribution, and contain information about all higher-order correlation functions.

MFs were first used to characterize large-scale structure by Mecke et al. (1994) in the form of the germ-grain model. This model pins down the morphology of the galaxy distribution by treating the survey galaxies as points (the germs) and decorating them with balls (the grains) whose scale-probing radius is the only model parameter. The union
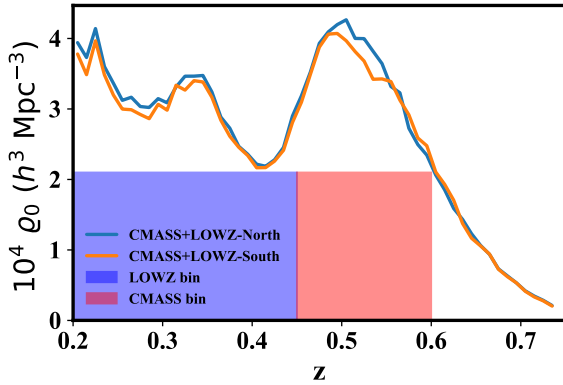
★ E-mail: jsull3@utexas.edu

Figure 1. Minimum number density for the combined full CMASS, LOWZ, and LOWZE2 catalogs. By considering only values of ϱ_0 that are at or below the minimum value over the combined sample, we need not restrict our analysis to differing density ranges at different redshifts. The bins we call CMASS and LOWZ in this work are shown here.

Table 1. Reduced set of basic parameters of the SDSS DR12 CMASS and LOWZ samples. The parent samples from Reid et al. (2016) yield (their Table 2) N_{used} galaxies and the listed effective areas. We then opt to split the combined sample into two redshift bins, $0.45 \leqslant z \leqslant 0.60$ and $0.20 \leqslant z \leqslant 0.45$, which we call CMASS and LOWZ, respectively (see also Figure 1). N_{gal} is the number of galaxies in these two redshift slices.

Sample	CMASS			LOWZ		
Property	NGC	SGC	total	NGC	SGC	total
N_{used}	568,776	208,426	777,202	248,237	113,525	361,762
Effective area (deg^2)	6,851	2,525	9,376	5,836	2,501	8,337
N_{gal}(our bins)	410,617	151,003	561,620	294,091	123,719	417,810

derstanding. Therefore, we start in Section 2 by describing our usage of the BOSS dataset. Section 3 and Section 4 then guide through the germ-grain model, its application to the data and our transformations of the MFs. Section 5 describes higher-order correlations and relates our analysis of redshift evolution. Section 6 presents our conclusions.

2 BOSS DATA

Our sample was collected with the 2.5m Sloan Telescope Gunn et al. (2006). For specifics on photometry and instruments, refer to Gunn et al. (1998), Fukugita et al. (1996), Lupton et al. (2001), Smith et al. (2002), Pier et al. (2003), Padmanabhan et al. (2008), Doi et al. (2010), as well as the eighth data release (Eisenstein et al. 2011). For details on spectroscopic redshift determination see Smee et al. (2013) and Bolton et al. (2012). Our samples are drawn from the third phase of the SDSS (York et al. 2000) Luminous Red Galaxy (LRG) catalog (Eisenstein et al. 2001) of the BOSS (Dawson et al. 2016) in DR12 (Alam et al. 2015). For specifics regarding the data, see Reid et al. (2016).

2.1 Redshift Samples - CMASS and LOWZ

By using the DR12 dataset, we access an unprecedented number of spectroscopic redshifts. We use a combination of the CMASS and LOWZ samples, described in detail in Reid et al. (2016). The CMASS sample was designed to expand upon the color cuts towards the blue in the SDSS-I and SDSS-II LRG samples, and has a redshift range of $0.40 \lesssim z \lesssim 0.70$. The LOWZ sample was designed to decrease the lower redshift bound of SDSS-I and SDSS-II down to $z = 0.2$., increasing the effective galaxy count by roughly a factor of 3. Merging those two selections gives the "complete sample" which has been used in a number of recent analyses (see Alam et al. (2017) and references therein). We consider two redshift bins (Figure 1), which we will use CMASS ($0.45 \leqslant z \leqslant 0.60$) and LOWZ ($0.20 \leqslant z \leqslant 0.45$) to refer to from now on. This naming convention means our bins do not correspond to the separation into LOWZ and CMASS used in earlier analyses. The bins we use give a redshift range near triple that used in the DR12 paper. We use both the northern and the southern galactic caps (N- and SGCs) of CMASS and LOWZ for an effective area of $9,376 \ deg^2$. We also make

of these balls creates a set of extended bodies to which methods from integral geometry can be applied (For a review see Schmalzing (1999) or Schmalzing et al. (1996), or Buchert (1995) for a short review.). We discuss this model further in the context of our analysis.

Another popular use of MFs is to apply them to the isodensity contours of density fields, including galaxy and cluster surveys (e.g. Schmalzing et al. (1996), Kerscher et al. (1997), Kerscher et al. (1998), Kerscher et al. (2001a), and Kerscher et al. (2001b)), dark matter overdensity fields (Platzöder & Buchert (1996), Schmalzing & Buchert (1997), Sahni et al. (1998), Sathyaprakash et al. (1998), Schmalzing et al. (1999), Hikage et al. (2003), Nakagami et al. (2004), Choi et al. (2013), and Blake et al. (2014)), and other astrophysical settings (Petri et al. (2013), Gleser et al. (2006), Einasto et al. (2014), and Yoshiura et al. (2017)). Still further work with MFs has recently used the isotemperature contour maps of the CMB to constrain its Gaussianity (Ducout et al. (2013), Planck Collaboration et al. (2014a), and Planck Collaboration et al. (2014b), Planck Collaboration et al. (2016)). Clearly MFs are a widely-used tool in a variety of cosmological subfields.

Here we focus on the germ-grain model, because recent work using the 7th and 12th data releases of SDSS-III (DR7 and DR12) of the Baryon Oscillation Spectroscopic Survey (BOSS - Dawson et al. (2013)) has shown its use in accessing correlation information that cannot be calculated directly (DR7 paper - Wiegand et al. (2014), DR12 paper - Wiegand & Eisenstein (2017)), SDSS-III - Eisenstein et al. (2011)). We use the largest spectroscopic redshift survey to date to access unprecedented accuracy in the higher-order correlation functions. We explore the growth of structure by measuring redshift evolution of the higher-order correlations, which are expected to grow nonlinearly in time with varying behavior for different orders.

This paper is organized in a non-standard way. We try to guide the reader through the uncommon quantities by introducing them at the place where we first use them on the data. In this way, there is a more immediate visualization of the mathematical expressions which hopefully promotes un-
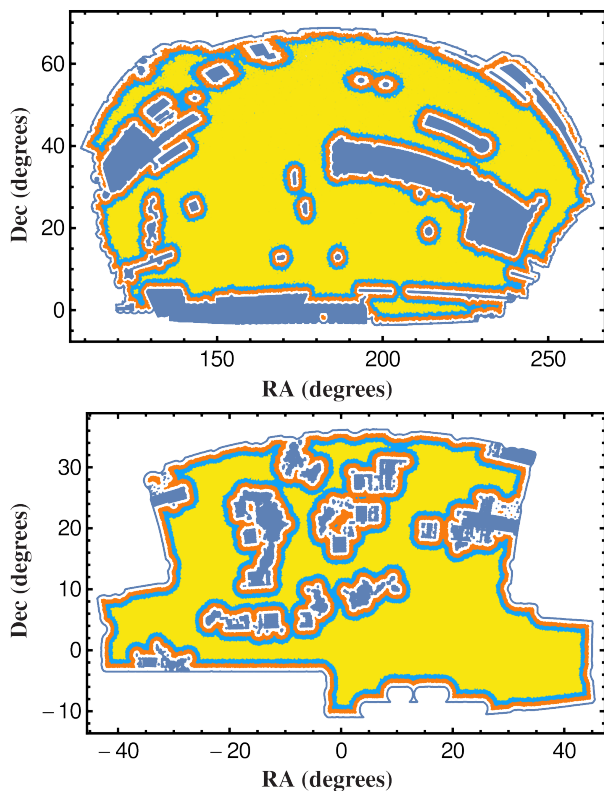
Figure 2. External and internal boundaries for the northern and southern patches analyzed. In blue the regions that we chose to define our boundary. In orange, light blue and yellow a projection is shown of the regions 18, 36 and $54h^{-1}\mathrm{Mpc}$ away from the nearest boundary. The boundary for the northern high-z bin is the same as in Wiegand et al. (2014). The boundary shown here is for LOWZ North excluding the EARLY3 region.

use of the LOWZE2 sample (chunk 2 of the original LOWZ footprint), though it contributes a relatively small amount to the effective area and sample count ($131\ \mathrm{deg}^2$, $2,985$ galaxies). There are a total of $979,430$ galaxies in our two-bin sample, more than double the $410,615$ used in the DR12 paper. The CMASS-South sample is approximately a factor of 2 smaller than its northern counterpart (Table 1). This has some effect on the behavior of the MFs, especially at large scales.

2.2 Survey and Corrections

We used a similar masking process to that of the DR12 paper. In correcting for bad and missing data, we used veto masks to remove galaxies based on effects such as poor spectral seeing, extinction, areas without observations and poor photometric conditions. We applied these veto masks to the mocks in order to make them as similar to the data as possible. Of the remaining galaxies, we considered only those in areas of the survey that were more than a distance of $2R$ from any inner or outer boundary of the data, where R is the scale at which we probe the MFs (see Section 3 for details). This restricted our use of the data, moreso for large values of R, but the alternative method of using boundary corrections and random distributions is not effective for MFs.

The masks for both caps are provided in Figure 2. The

LOWZ sample especially has a comparatively large number of holes and covers a smaller area of the sky. We see the effects of this on our analysis, especially at low densities. The masks presented in Figure 2 are a compromise between the need for a contiguous sample and the wish to exclude regions of bad or missing data. We did not check again the extent to which we succeeded in providing an unbiased estimate of the MFs, but the results of the DR12 paper indicate that the periodic box value should be within our error bars. We do some cross check that this is indeed the case in Section 5.2.

2.3 MD Patchy Mocks

We use mock data realizations to compare the correlations in the BOSS data to those of the cosmological concordance model. The mocks used in our analysis are drawn from the MultiDark(MD)-Patchy mock galaxy catalogs (Kitaura et al. 2016). These mocks were specifically produced to reflect the number density of the data, and were generated by referencing a large N-body simulation (BigMultiDark - Klypin et al. (2016)). We used the same ΛCDM cosmology as in the simulation, which the authors called the Planck1 cosmology, with $\Omega_M = 0.307115$, $\Omega_\Lambda = 0.692885$, $\Omega_b = 0.048$, $\sigma_8 = 0.8288$, and $h = 0.6777$. As indicated in the DR12 paper, these mocks are very well suited for comparison when calculating the higher-order parts of the MFs, although there is light tension between the data and the MD-Patchy power spectrum that may be relieved by a 5% reduction in the mock power spectrum amplitude. We used 997 mock survey realizations for CMASS-North, 399 mocks for CMASS-South and LOWZ-North, and 383 mocks for LOWZ-South in our calculations using the CHIPMINK code (Wiegand et al. 2014). These large numbers of mock realizations allow us to derive correlation-conscious uncertainties when comparing the data to the concordance model.

3 THE GERM-GRAIN MODEL

In this Section, we provide an abbreviated description of Minkowski Functionals. For a more complete explanation of the germ-grain model, see the DR7 paper.

3.1 Minkwoski Functionals

We apply a result from integral geometry (Hadwiger 1957) and define a set of 4 base functionals for the class of functionals over poly-convex bodies in three-dimensional Euclidean space. For these four Minkowski Functionals, we choose a normalization (in which they are usually denoted $V_0 - V_3$) that relates to familiar geometric quantities as follows:

$$V_0 = V \ ; \ V_1 = \frac{S}{6} \ ; \ V_2 = \frac{H}{3\pi} \ ; \ V_3 = \chi \ . \tag{1}$$

Here V is the volume of the extended body, S is its surface area, H is the integral mean curvature of the surface, and χ is the topological Euler characteristic or integral Gaussian curvature. To refresh the reader, H is the average of the curvature of a surface over all angles and χ is the sum of independent components and cavities minus the number of holes.
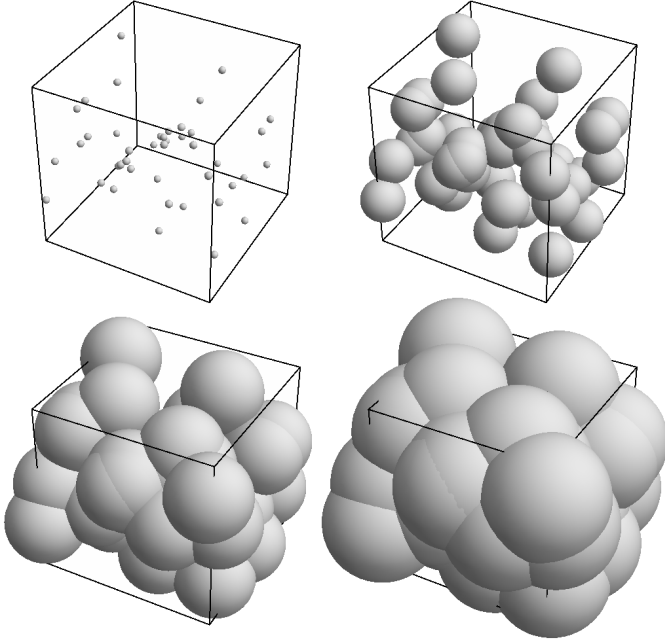
The germ-grain model connects MFs and observed

Figure 3. Illustration of the germ-grain procedure to transform a set of galaxies into an extended body. The galaxy positions are surrounded by balls of a common radius R. Then the MFs of the body formed by the union of all the balls are studied as a function of R. R effectively sets the scale at which correlations are probed. From Wiegand et al. (2014).

galaxy distributions. The germ-grain procedure is to take a point distribution of "germs" and surround each point with a convex body, or "grain." By taking a galaxy distribution with positions $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ as the point distribution, and using balls (filled spheres) of radius R as the convex bodies, we create an extended body from the union of the balls. The MFs can now be used to describe the relatively simple extended body that is transformed from the observed galaxy distribution (Figure 3).

The MFs clearly depend on the common ball scaling radius R, but they also depend on the sample density ϱ_0 of the observed distribution due to the germ-grain construction. The radius R is a tunable parameter that we alter in our analysis to probe different scales of structure. When exploring the density parameter space, we randomly sample from either the mock or data distribution at different values of ϱ_0. These values are determined by downsampling at fixed percentages from the height of our bins. At low densities, this tends to erase structure and the MF's approach Poissonity.

4   MEASURING MFS

In this Section, we describe the quantities we measure for each sample.

4.1   Functional Densities

To measure the MFs for our samples, we must calculate the MFs of the individual balls. These *partial MFs* are then taken as an ensemble of statistical quantities. Averaging the partial functionals leads to the *global* functionals used in our analysis. Rather than to work with the extrinsic quantities V_μ, we use the survey volume V_{Survey} to define the intrinsic functional densities v_μ:

$$v_\mu \equiv V_\mu / V_{\text{Survey}} \; ; \; \mu \in \{0, \ldots, 3\} \; . \tag{2}$$

The analytic expressions for these v_μ are given by:

$$
\begin{aligned}
\langle v_0 \rangle &= 1 - e^{-\varrho_0 \overline{V}_0} \, , \\
\langle v_1 \rangle &= \varrho_0 \overline{V}_1 e^{-\varrho_0 \overline{V}_0} \, , \\
\langle v_2 \rangle &= \left(\varrho_0 \overline{V}_2 - \frac{3\pi}{8} \varrho_0^2 \overline{V}_1^2 \right) e^{-\varrho_0 \overline{V}_0} \, , \\
\langle v_3 \rangle &= \left(\varrho_0 \overline{V}_3 - \frac{9}{2} \varrho_0^2 \overline{V}_1 \overline{V}_2 + \frac{9\pi}{16} \varrho_0^3 \overline{V}_1^3 \right) e^{-\varrho_0 \overline{V}_0} \, ,
\end{aligned}
\tag{3}
$$

where $\langle v_\mu \rangle$ refer to the average functional densities over the distribution, and $\overline{V}_\mu$ refer to the unnormalized modified MFs (Wiegand et al. 2014). The v_μ as measured for the mocks are shown in Figure 4. Their behavior as a function of R (or diameter D in this case) and ϱ_0 is reasonably intuitive. For v_0, as R increases the volume of the survey box occupied by the balls is saturated and v_0 approaches unity, and a decrease in density delays this saturation. The other functionals are slightly less transparent, but show that there is some maximum in the uncovered surface area (v_1) when R is increased, and a similar maximum and minimum in curvature (v_2) and Euler characteristic (v_3). For the surface area, a decrease in density will delay the maximum as a function of radius, and for the remaining two MF densities, a decrease in density decreases the amplitude of the functional. These v_μ plainly exhibit non-Poissonity at higher densities, but are relatively opaque to non-Gaussianity, motivating a transformation of these quantities.

4.2   Transforming the Densities

Observed galaxy distributions have significant structure, and thus must employ the full form of $\overline{V}_\mu$ in Section 4.1. But, for the case of a Poisson distribution, the unnormalized modified MFs $\overline{V}_\mu$ are given by the more familiar MFs of a ball B, i.e. $\overline{V}_\mu = V_\mu(B)$ with

$$
V_0(B) = \frac{4\pi}{3} R^3 \; ; \; V_1(B) = \frac{2}{3} \pi R^2 \; ;
$$
$$
V_2(B) = \frac{4}{3} R \; ; \; V_3(B) = 1 \; . \tag{4}
$$

These are natural and convenient to work with, and we use the Poisson case as a reference in analyzing distributions with more structure. Following from this choice, we adopt the dimensionless transformed MFs η_μ:

$$\eta_\mu \equiv \overline{V}_\mu / V_\mu(B) \; . \tag{5}$$

The mapping from v_μ to η_μ retains the information about the MFs as a function of R and ϱ_0, but the latter provides a better perspective in terms of the higher-order correlations. To see this, we can rewrite the dimensionless MFs as the following power series:

$$\eta_\mu = \sum_{n=0}^{\infty} \frac{c_{\mu, n+1}}{(n+1)!} \left(-\varrho_0 V_0(B) \right)^n \tag{6}$$
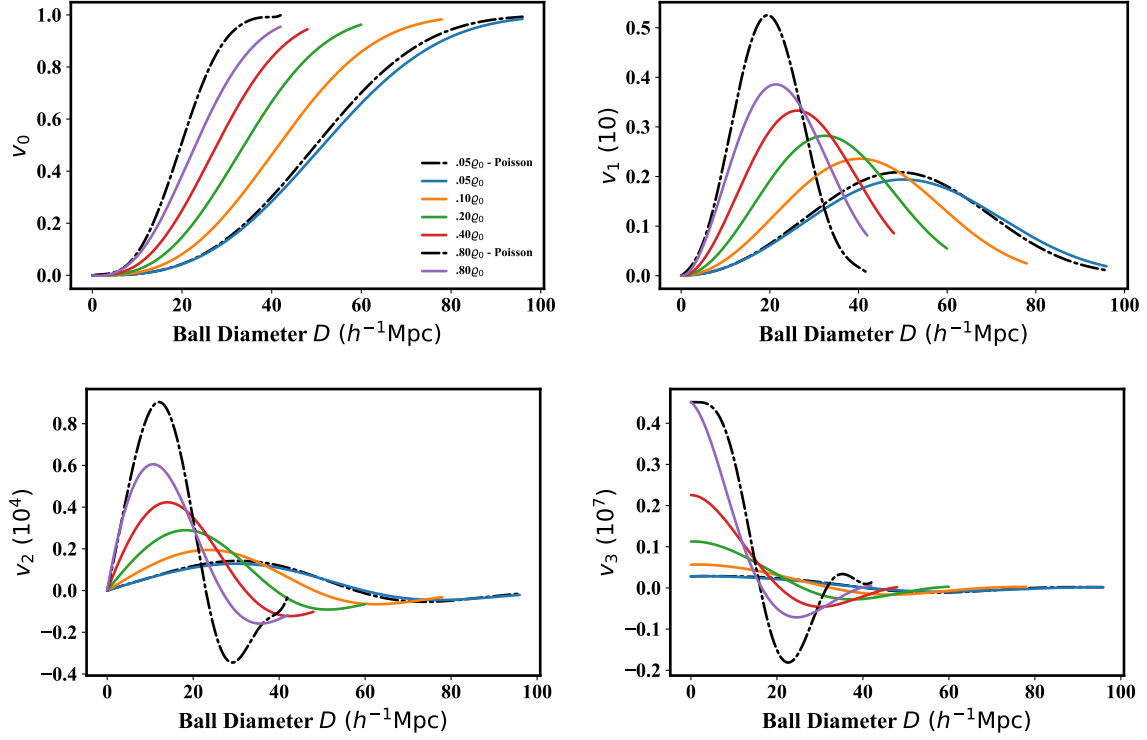
Figure 4. MF densities ν_μ for the CMASS-North MD-Patchy mock realizations within a sample density range of 5% to 80% (in factor of 2 steps) of the reference density $\varrho_0 = 2.08 \times 10^{-4} h^3 \text{Mpc}^{-3}$. The most extreme densities are plotted along with their theoretical Poisson point distributions (black dash-dotted lines). At the lowest density the MFs approach the Poisson curves, and at the highest density show the greatest difference from the Poisson curves.

where $c_{\mu,1} = 1$ and

$$c_{\mu,n+1}(R) = V_0^{-n}(B) \int \xi_{n+1}(0, \mathbf{x}_1, \dots \mathbf{x}_n) \times \qquad (7)$$

$$\frac{V_\mu (B \cap B_{\mathbf{x}_1} \cap \cdots \cap B_{\mathbf{x}_n})}{V_\mu (B)} d^3 x_1 \dots d^3 x_n$$

where the weights $V_\mu (B \cap B_{\mathbf{x}_1} \cap \cdots \cap B_{\mathbf{x}_n}) / V_\mu (B)$ in the integrand may be derived by calculating the MFs of the intersection of n balls.

To emphasize the structure: Each term of the series Section 4.2 is the integral of the product of the n-point correlation function ξ_{n+1} and a geometric weighting function that only depends on the balls (and therefore R). The factor that includes the sample density ϱ_0 dependence is independent of the correlation-dependent integral. Now we have a quantity conveniently defined on the unit interval $[0, 1]$ (for standard cosmological structure) that can be more easily compared in terms of our parameters.

Figure 5 shows representative density samples (5% - 80% of $\varrho_0 = 2.08 \times 10^{-4} h^3 \text{ Mpc}^{-3}$) of η_μ as a function of D and gives a clearer measure of the non-Poissonian behavior of the MFs. The MFs deviate further from Poisson behavior as D and ϱ_0 increase, approaching the structure of the full galaxy distribution, and illustrating the increased non-Poissonity at large scales. The MFs are on the order of 1% different between the redshift bins, so already the redshift evolution in the MFs is evident. However, we have yet to isolate higher-order contributions of the MFs, which we discuss in the next section.

5 HIGHER-ORDER CONTRIBUTIONS TO THE MFS

This Section describes our analysis of the isolated higher-order contributions to the MFs. We start with the definition of the higher-order quantity we analyze and then turn to the main results of the paper, i.e. the growth of these higher-order contributions.

5.1 The Two-point Contribution

To measure higher-order correlations, we subtract the two-point correlation function contribution to the MFs. To do this, we must first calculate the two-point term. Ignoring the higher-order terms ($n \geqslant 3$) of equation (6) we are left with:

$$\eta_\mu = 1 - \frac{\varrho_0}{2} \int d^3 x_1 \xi_2 (0, \mathbf{x}_1) \frac{V_\mu (B \cap B_{\mathbf{x}_1})}{V_\mu (B)} . \qquad (8)$$

We assume isotropic correlations, and upon transforming to spherical coordinates equation (8) becomes:

$$\eta_\mu = 1 - 2\pi \varrho_0 \int_0^{2R} \frac{V_\mu (R, r)}{V_\mu (B(R))} \xi_2 (r) r^2 dr , \qquad (9)$$

where the integral is zero if the centers of the balls are more than $2R$ apart. This is the same as equation (6) for $n = 1$ barring the factor of $V_0(B)$.
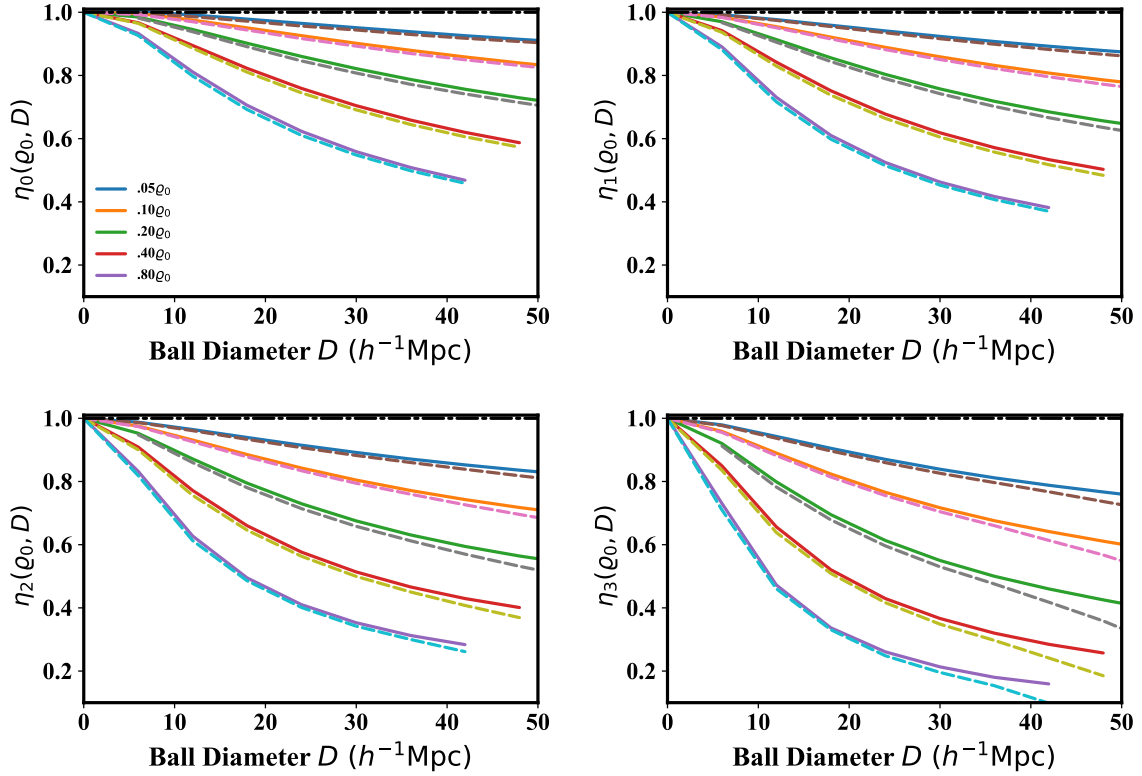
Figure 5. Dimensionless transformed MFs in the Northern Galactic Cap evaluated for 997 (CMASS) and 399 (LOWZ) SDSS DR12 MD-Patchy mocks plotted as a function of ball diameter D for densities 5% to 80% of the reference density $\varrho_0 = 2.08 \times 10^{-4} h^3 \text{Mpc}^{-3}$. CMASS-North values are solid, LOWZ-North values are dashed. These quantities can be more directly related to the correlation functions than the ν_μ can. Significant difference from the Poisson case is shown by deviation in the MFs from the black dash-dotted line at 1. The difference in the functionals between the redshift bins is visible to a slight degree across all the MFs.

We can write the weight functions $V_\mu \left(B \cap B_{\mathbf{x}_1} \right)$ as:

$$V_0 \left(R, r \right) = \frac{1}{12} \pi \left(2R - r \right)^2 \left(r + 4R \right) \ ; \tag{10}$$

$$V_1 \left(R, r \right) = \frac{1}{3} \pi R \left(2R - r \right) \ ; \tag{11}$$

$$V_2 \left(R, r \right) = \frac{2}{3} \left(2R - r \right) + \frac{2}{3} R \sqrt{1 - \left(\frac{r}{2R} \right)^2} \arcsin \left(\frac{r}{2R} \right) \tag{12}$$

$$V_3 \left(R, r \right) = 1 \ , \tag{13}$$

where R is again the radius of the balls and r is the integration variable. The weights then conveniently define integration windows for ξ_2 for the different functionals, and μ increases, so does the effective scale at which the structure is probed (see the DR12 paper for a more detailed discussion).

5.2 Higher-order Functionals and Growth

We isolate the higher-order contributions to the MFs to study the higher order correlations of the observed galaxy distribution. Subtracting the two-point contribution equation (9) from the power series expression for η_μ, we take only the terms with $n \geqslant 2$ in equation (6) as the higher-order part of the functionals, which we define as η_μ^h. The two-point contribution is obtained from a direct measurement of the two-point correlation function monopole by performing the integral in equation (9).

Table 2. Significance of the detection of higher-order clustering in the MFs evaluated at $D = 36$ h^{-1} Mpc. The χ^2 values quantify the significance of the deviation of the higher-order functionals η_μ^h of the data from the zero line in Figure 6. N refers to North, S to South, and CMASS and LOWZ to our bins of Figure 1. There are 24 degrees of freedom, and hence we find a highly significant contribution of the higher-order part of the MFs for all analyzed samples. The decrease with μ is due to the increasing error bars induced by the inversion of Section 4.1. The decrease from left to right reflects the relative sizes of the samples. Having this level of significance is important for our study of the changes of theses quantities with redshift.

μ	CMASS-N	LOWZ-N	CMASS-S	LOWZ-S
0	5385	3431	1796	1267
1	5067	3098	1666	1096
2	3993	2463	1230	851
3	2664	1693	794	601

The higher-order functionals η_μ^h are a significant portion of the calculated functionals η_μ. Figure 6 shows $\eta_\mu^h(D = 36 \ h^{-1}$ Mpc) for both redshift samples and both hemispheres. We measure the significance of the detection of higher-order clustering by computing χ^2 values with 24 degrees of freedom for the mocks and the zero line in Table 2 and find values of order 10^3, indicating a highly significant contri-
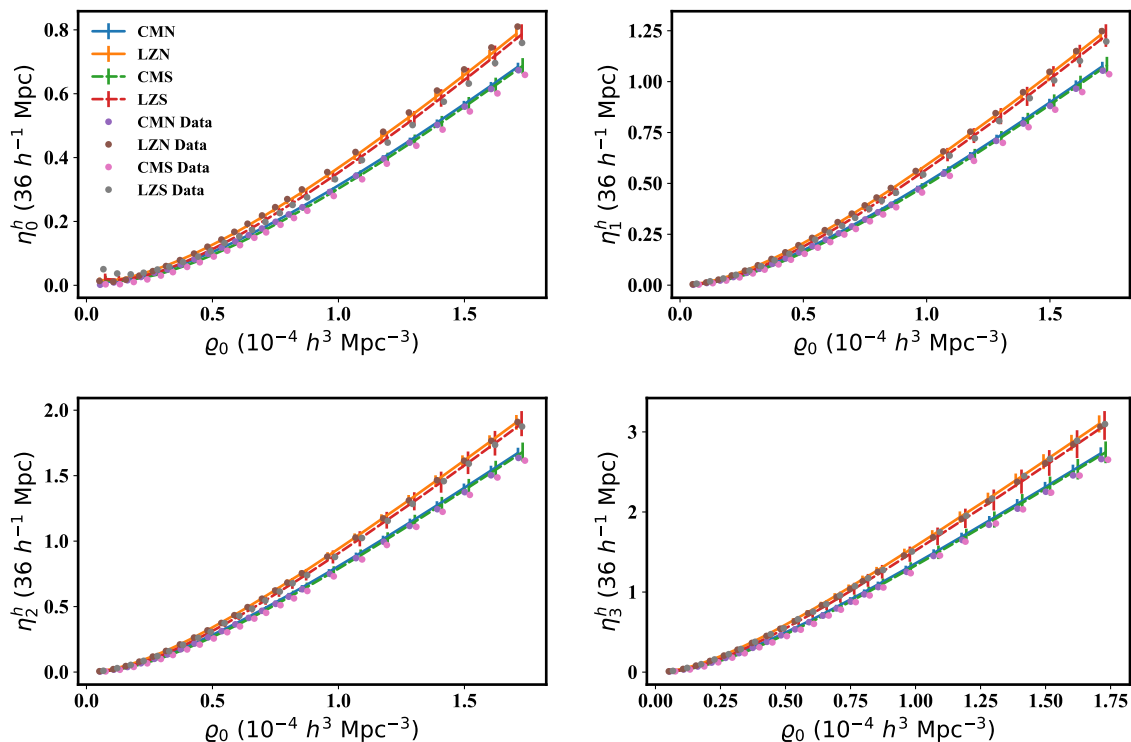
Figure 6. The higher-order contribution to the MFs η_μ^h as a function of density ϱ_0. The contribution of the known two point correlations has been removed by subtracting equation (8) from equation (6). The lines are the MD-Patchy mock values with $1 - \sigma$ error bars on the mean. The dots are the values derived from the data. Already on this figure we can see a clear split between the two redshift bins (see also Table 4). Moreover, both the different regions on the sky and the data and the mocks agree reasonably well (see also Table 3).

bution of higher-order correlations to the structure of the distribution. All χ^2 values are computed using the correction to the estimated inverse covariance matrix of Percival et al. (2014). The clear detection of the higher-order contributions to the MFs is important for this paper's goal of studying the redshift evolution of higher-order clustering. Table 2 also illustrates the enormous increase in our ability to probe higher-order correlations via η_μ^h due to the large increase in sample size and volume relative to DR7.

In the limit of low sampling density in Figure 6, the higher-order part seems negligible, reflecting the near two-point structure characteristic of a completely Gaussian point distribution. Though it seems intuitive to expect this in the low-density limit, there is no theorem that would require a structured point process to pass through a stage of Gaussianity when the density is reduced. Regardless, the two-point behavior increasingly fails to describe the higher-order MFs as the sample density increases, meaning the integrated three and higher point functions must be non-zero, quantitatively capturing the non-Gaussianity of the density field.

The data from the northern and southern sky in Figure 6 shows consistent higher-order clustering. Table 3 quantifies this by computing the difference between the η_μ^h of the data in the south and the north for each of our two redshift bins. Because the analysis of the mocks reveal a small shift between the hemispheres due to the different boundary corrections, we apply the difference between the mean of the mock results as a correction to the hemispheric residual of

Table 3. Comparison of the NGC and SGC in each redshift bin. The χ^2 values of significance are for the deviation in the MFs of the data in the NGC and SGC. We attempt to correct for potential systematics by subtracting the separation of the mean of 399 mocks from the difference in the data. χ^2 values are for η_μ^h in each redshift bin at $D = 36$ h^{-1} Mpc for 24 degrees of freedom.

μ	CMASS		LOWZ	
	χ^2	p-value	χ^2	p-value
0	14.1	0.055	33.3	0.098
1	18.3	0.211	40.2	0.020
2	19.2	0.257	39.4	0.025
3	21.3	0.378	32.2	0.122

the data before computing χ^2. As we are testing a residual difference, we use the sum of the south and north covariance matrices to estimate the significance of this residual. We use the 24 density points in Figure 6 and we evaluate both, mocks and data, at common values for the density. We find that the residuals have χ^2 values ranging from 14.1 to 40.2. Given that we have 24 degrees of freedom, these are unsurprising values, with no case below 2% probability. Therefore, we conclude that the results in the SGC and NGC are statistically consistent with each other.

Figure 6 clearly suggests that the MF values are larger at lower redshift. Figure 7 shows the percentage growth from the higher redshift CMASS bin to the lower redshift LOWZ
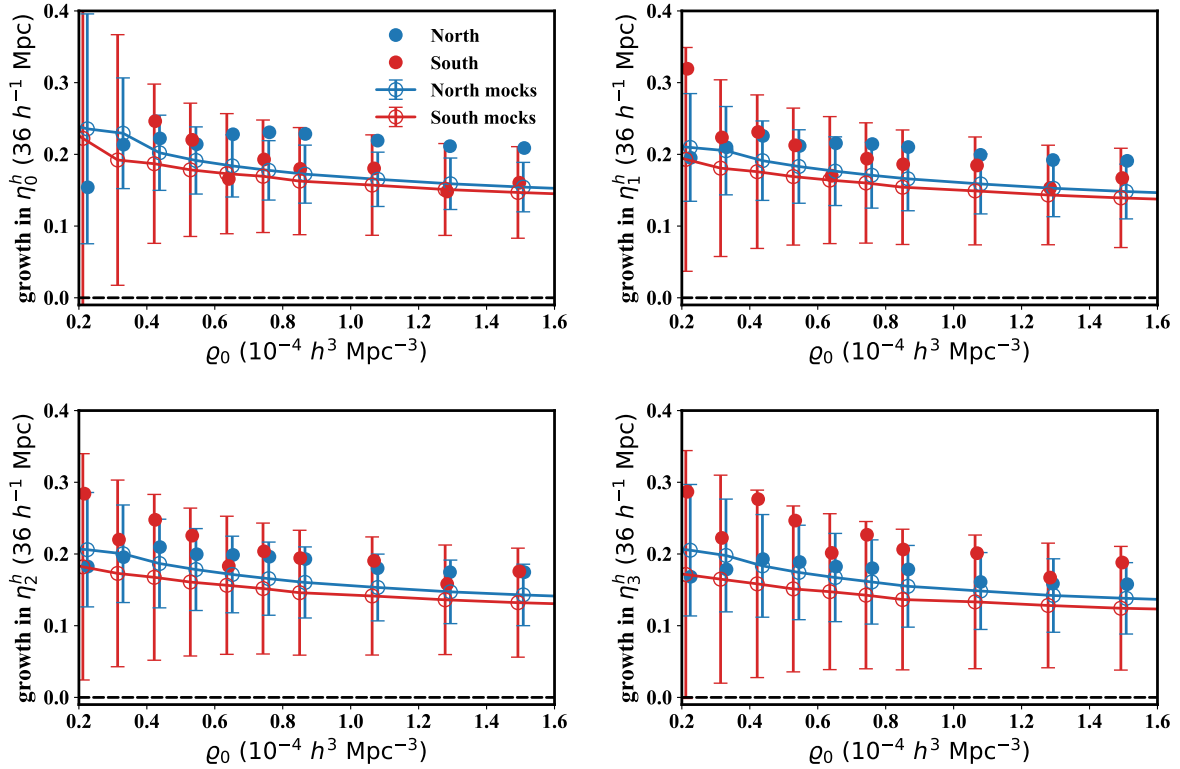
Figure 7. The percentage growth of the higher-order term of the functionals η_μ^h between redshift bins. We find approximately $15\% - 25\%$ growth from the higher redshift bin to the lower redshift bin. The MD-Patchy mock values are shown by the solid lines with hollow points and $1 - \sigma$ error bars for one mock realization each, and the data by the solid points. The zero line is shown in black for comparison to the no-growth case.

for a selection of 10 of our 24 density values. This shows that the growth is fairly constant as a function of density, with a mild decrease as the density increases. Importantly, the percentage growth of 20–25% in the MF is about a factor of two larger than the $\approx 10\%$ growth in the two-point correlation function between these two samples.

The percentage growth values for the data in the NGC and SGC seem to agree with the mocks relatively well. Barring the limit of low density, and the case of η_0^h, the values are close to, if not within, the $1 - \sigma$ error bars. However, it is clear that the growth for the data is systematically higher than the mean growth of the mocks.

5.3 Model comparison

Having detected significant growth in the higher-order contributions to the MFs, we trace its origin in this section. To this end we analyzed a set of 16 N-body simulations of the same Planck cosmology, but with varying phases (Garrison et al. 2016). For these simulations we determined the two-point correlation functions and MFs for the underlying dark matter field and for the galaxies derived from a self-adjusted HOD model at two different redshifts. The redshifts that we use correspond to a snapshot of the simulation at $z = 0.3$ for our LOWZ bin and $z = 0.5$ for our CMASS bin.

The HOD model we fit is of the form used in Zheng et al. (2009) and Kwan et al. (2015). We derive its five parameters using code developed by Yuan et al. (2017), performing

a fit to our measured redshift space two-point correlation function monopole of the data. More specifically, we use the FoF halos provided by the simulation, select them with the HOD prescription, and redshift the final selection in the z-direction. We used this approach instead of the standard one, as parameters derived from a fit to the projected two-point function of the data as found in the literature (although based on a different reference cosmological model) did not give a good match for the amplitude of the redshift space two-point monopole. For LOWZ for example, we needed a slightly higher M_{cut} and lower σ than the best fit of Parejko et al. (2013). However, the final parameters of our model lie well within their "Mean Full" error bars.

The philosophy of using the redshift space two-point correlation function of the data that we also used in the rest of our analysis was to generate a galaxy distribution from the HOD model that best reflects the two point properties we measure. Every deviation in the higher-order only MFs of the resulting distribution from those of the data is then maximally informative of the properties of the higher orders.

Figure 8 shows the growth of the two-point correlation function between our CMASS and our LOWZ bin. The growth is largest for the underlying dark matter field in yellow which agrees well with the linear growth of the amplitude of the dark matter density field given by the ratio of their linear growth factors $D(z = 0.3)/D(z = 0.5) \approx 1.228$.

From the comparison of the final HOD galaxy two-point correlations to those of their dark matter field, we
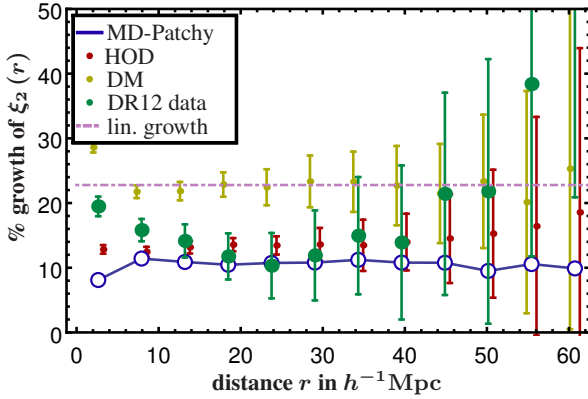
Figure 8. Comparison of the growth of two-point correlations between our two redshift bins for different cases. We use only data from the NGC here. Blue hollow points joined by the line show the mean of the MD-Patchy mocks. The red points are for the galaxy samples derived from our own HOD model, and the yellow points show the growth of the underlying dark matter field. The green solid points are the values for the galaxies from the final DR 12 sample. Their $1 - \sigma$ error bars are derived from the fluctuations between the MD-Patchy mocks. The dashed-dotted line indicates the ratio of the linear growth factors of our ΛCDM model.

Table 4. Significance of growth with redshift in each galactic cap. The χ^2 values of significance are for the growth in the MFs for the data between the CMASS and LOWZ redshift bins for η_μ^h in each galactic cap at $\varrho_0 = 0.5\varrho_{max}$ for 7 degrees of freedom. These values quantify the growth shown in Figure 9.

μ	NGC		SGC	
	χ^2	p-value	χ^2	p-value
0	96.4	10^{-18}	13.4	0.06
1	111.0	10^{-20}	16.1	0.02
2	108.4	10^{-20}	15.2	0.03
3	106.0	10^{-19}	14.0	0.05

find that the approximate linear bias drops from $b_1^2 = 4.93$ to $b_1^2 = 4.60$. This 7% decrease in the correlation amplitude counteracts the increase of the clustering from the non-linear evolution of the dark matter density field and lets the red points in Figure 8 lie below those for the underlying dark matter field.

The growth derived from the data shown as the green filled circles lies in the expected range for both HOD galaxies (in red) and mock galaxies from the MD-Patchy mocks (in blue), except at very small scales. For MD-Patchy this may be related to the fundamental resolution of the underlying density field from which the mocks were created. For our HOD model it probably is related to the use of FoF halos and the resolution of light dark matter halos.

Figure 9 shows the growth of the higher-order contribution to the MFs. It follows the general behavior of the two point case in Figure 8, with the strongest growth being realized in the dark matter field. Also here the data lies within the range of the simulated galaxies, except at the smallest scales. Again, the scale dependence of the growth is fairly weak, but more important than in the two-point case.

For the data shown in Figure 9 we quantify the significance with which we measure the growth in the higher orders in Table 4. We determine the χ^2 values by computing the difference of the values of η_μ^h in our CMASS and LOWZ bins for the seven radii plotted in Figure 9. As we determine the significance for the difference of two independent values, we use the inverse of the sum of the covariance matrices of the CMASS and LOWZ data for obtaining the χ^2. The data used for η_3^h can be found in Table 5 at the end of the text. We do not use the fraction shown in Figure 9 directly, as for a ratio distribution the calculation of the significance is less straightforward than for the difference. We find a very significant detection of growth. Even if we remove the first two radial points for which it is not entirely clear how well the mocks and their fluctuations describe the data, we are still left with a 5.5σ detection. Due to the strong correlation in ϱ_0, adding in more points in ϱ_0 does not change this picture. We investigated a compression of the ϱ_0 direction into a single value by a transformation into the eigenbasis of the covariance matrix. The leading two eigenvectors give a good description of the data as a function of density and using them instead of only the value at $\varrho_0 = 0.5\varrho_{max}$ as in Table 4 does not increase the significance by much.

In order to test if we understand the growth of the higher orders, we added a line to Figure 9 that shows the predicted growth for the mean of the mocks derived from the growth of their two-point function. This has been calculated following the observation of the DR12 paper that in a perturbative approach to the higher-order connected correlation functions the linear power spectrum enters the series expansion in equation (6) with the same power as the density. So the leading effect of increasing the linear power spectrum amplitude can be obtained by a simple rescaling of the density. We derive this rescaling factor at each radius from the measured integrated two-point correlation function. Applying this correction to the higher-order terms as well, by evaluating the mean MD-Patchy CMASS η_μ^h at the shifted ϱ_0 position and taking the ratio, gives the dashed-dotted line in Figure 9. The good match with the measured growth in the mocks indicates that it is indeed the leading terms at each correlation order that grow proportionally to their respective amplitude factor. This is true within the errors also for the data (their predicted growth is not shown on the plot for clarity) and illustrates how structure grows hierarchically.

Finally, we note that the reasonable agreement in the shape of the growth for our HOD galaxies and the MD-Patchy mocks provides a basic test of the influence of the boundaries. As the HOD galaxies are derived from a boundaryless periodic box and the MD-Patchy mocks include all effects of the our boundary treatment, we can be confident that the main conclusions of this paper regarding the detection of growth in the higher orders are not affected by possibly remaining issues with the mask.

6 SUMMARY & CONCLUSIONS

In this paper, we perform an unprecedented analysis of higher-order correlation information on the largest spectroscopic redshift survey to date. We use Minkowski Functional analysis along with independently calculated two-point cor-
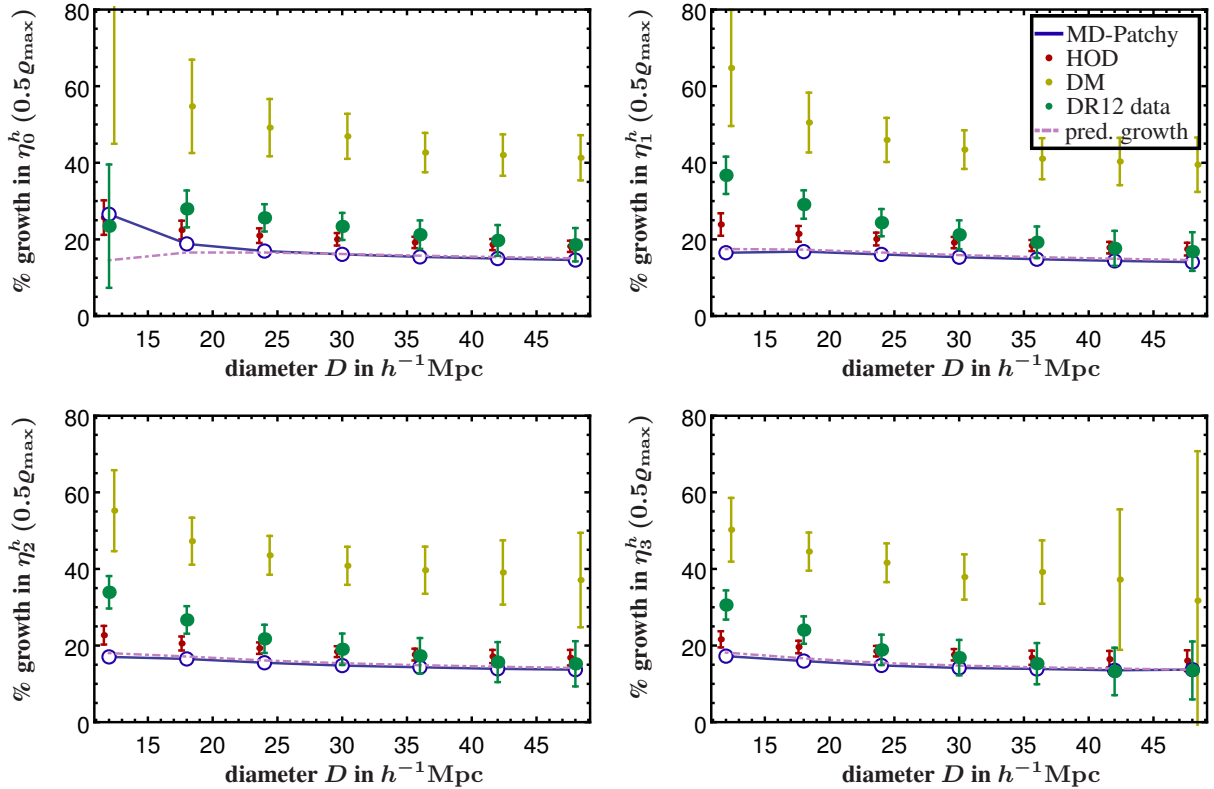
Figure 9. Comparison of the growth in the higher-order correlations, measured through η_μ^h, between our two redshift bins for different cases. We use only data from the NGC here. Blue hollow points joined by the line show the mean of the MD-Patchy mocks. The red points are derived from the galaxy samples of our own HOD model, and the yellow points show the growth of the underlying dark matter field. In those two cases the $1-\sigma$ error bars are derived from realizations with different phases of the initial modes. The green solid points are the values measured for the galaxies from the final DR 12 sample. Their $1-\sigma$ error bars are derived from the fluctuations between the MD-Patchy mocks. The dashed-dotted line shows the expected growth when scaling all higher-order functions of the MD-Patchy mocks by their corresponding powers of the linear amplitude growth (see text). η_μ^h is measured at 50% of the maximal density, i.e. at $1.07 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$, but density dependence is relatively modest for large scales, cf. Figure 7.

relation functions to measure non-Gaussianity in large-scale structure. Our detection of non-Gaussianity is highly significant, with χ^2 values of $O(10^3)$ for 24 degrees of freedom across the NGC and SGC of the SDSS-III DR12.

We measure the percentage growth in redshift of the higher-order part of the MFs for the first time. This growth is measured to be $15\% - 35\%$, depending on functional, density and scale. We show the redshift evolution of non-Gaussianity is greater in the higher-order contribution to the MFs than in the two-point term. When comparing the integrated two-point correlations to the higher-order part of the MFs, the excess of the higher orders approaches a factor of 2 on small scales.

Using simulations we show that the growth of higher-order structure of the underlying dark matter field is even larger, exceeding 60% on small scales for some of the functionals, but that it is counterbalanced in the galaxy field by a decrease in the bias which reduces the growth of the two-point amplitude by 7%.

Finally we show how measuring the higher-order MFs as a function of density allows us to perform a simple test of hierarchical clustering. Rescaling the density corresponds to increasing the correlation strength in the higher-orders

proportional to their order and reproduces the results for the growth very well.

Our results motivate multiple avenues for future analysis of higher-order correlations, including measurement of the growth of the integrated three-point and four-point functions. The growth in the higher-order correlations may also be applied in relation to the linear growth rate.

D	12	18	24	30	36	42	48
$\eta_{3,\mathrm{d}}^{h}$	0.262	0.602	0.967	1.338	1.689	2.005	2.287
$\eta_{3,\mathrm{m}}^{h}$	0.236	0.570	0.954	1.343	1.710	2.045	2.351
σ	0.0055	0.013	0.026	0.043	0.065	0.091	0.124
12	1	0.9	0.797	0.711	0.653	0.599	0.545
18	...	1	0.953	0.88	0.815	0.742	0.663
24	...	...	1	0.967	0.91	0.834	0.74
30	...	...	...	1	0.972	0.909	0.811
36	...	...	...	...	1	0.969	0.884
42	...	...	...	...	...	1	0.95
48	...	...	...	...	...	...	1

D	12	18	24	30	36	42	48
$\eta_{3,\mathrm{d}}^{h}$	0.199	0.482	0.806	1.135	1.452	1.755	1.997
$\eta_{3,\mathrm{m}}^{h}$	0.200	0.487	0.823	1.165	1.488	1.785	2.051
σ	0.0038	0.0093	0.018	0.030	0.045	0.065	0.09
12	1	0.89	0.743	0.636	0.556	0.489	0.41
18	...	1	0.937	0.851	0.769	0.696	0.598
24	...	...	1	0.965	0.902	0.828	0.721
30	...	...	...	1	0.973	0.914	0.813
36	...	...	...	...	1	0.969	0.882
42	...	...	...	...	...	1	0.948
48	...	...	...	...	...	...	1

Table 5. Extracted higher-order part of the MF for the data $\eta_{3,\mathrm{d}}^{h}$ and the MD-Patchy mocks $\eta_{3,\mathrm{m}}^{h}$ using NGC only as a function of diameter D at $\varrho_0 = 0.5\varrho_{\mathrm{max}}$, i.e. $1.07 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$. Left table for our LOWZ bin, right table for our CMASS bin. The large tables show the symmetric correlation matrices, with the values on the left indicating the respective diameter of the row. σ^2 gives the diagonal elements of the covariance matrix and $\sigma \otimes \sigma$ multiplies the correlation matrix to give the covariance matrix. The covariances have been derived from 399 MD-Patchy mocks for each of the tables. Dividing the $\eta_{3,\mathrm{d}}^{h}$ values on the left by those on the right gives the green data line in Figure 8 and the covariances can be used to derive those of the MD-Patchy mocks shown in the same figure. As the different density values are strongly correlated and the same is true for the different functionals, the one set of tables presented here already contains a large fraction of the information contained in our measurements.

REFERENCES

Alam S., Ho S., Vargas-Magaña M., Schneider D. P., 2015, MNRAS, 453, 1754
Alam S., et al., 2017, MNRAS, 470, 2617
Blake C., James J. B., Poole G. B., 2014, MNRAS, 437, 2488
Bolton A. S., et al., 2012, AJ, 144, 144
Buchert T., 1995, in Mücket J. P., Gottloeber S., Müller V., eds, Large Scale Structure in the Universe. p. 156 (arXiv:astro-ph/9412061)
Choi Y.-Y., Kim J., Rossi G., Kim S. S., Lee J.-E., 2013, ApJS, 209, 19
Dawson K. S., et al., 2013, AJ, 145, 10
Dawson K. S., et al., 2016, AJ, 151, 44
Doi M., et al., 2010, AJ, 139, 1628
Ducout A., Bouchet F. R., Colombi S., Pogosyan D., Prunet S., 2013, MNRAS, 429, 2104
Einasto M., Lietzen H., Tempel E., Gramann M., Liivamägi L. J., Einasto J., 2014, A&A, 562, A87
Eisenstein D. J., et al., 2001, AJ, 122, 2267
Eisenstein D. J., et al., 2011, AJ, 142, 72
Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748
Garrison L. H., Eisenstein D. J., Ferrer D., Metchnik M. V., Pinto P. A., 2016, MNRAS, 461, 4125
Gleser L., Nusser A., Ciardi B., Desjacques V., 2006, MNRAS, 370, 1329
Gunn J. E., et al., 1998, AJ, 116, 3040
Gunn J. E., et al., 2006, AJ, 131, 2332
Hadwiger H., 1957, Vorlesungen über Inhalt, Oberfläche und Isoperimetrie. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete ; 93, Springer
Hikage C., et al., 2003, PASJ, 55, 911
Kerscher M., et al., 1997, MNRAS, 284, 73
Kerscher M., Schmalzing J., Buchert T., Wagner H., 1998, A&A, 333, 1
Kerscher M., Mecke K., Schmalzing J., Beisbart C., Buchert T., Wagner H., 2001a, A&A, 373, 1
Kerscher M., et al., 2001b, A&A, 377, 1
Kitaura F.-S., et al., 2016, MNRAS, 456, 4156
Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340
Kwan J., Heitmann K., Habib S., Padmanabhan N., Lawrence E., Finkel H., Frontiere N., Pope A., 2015, ApJ, 810, 35
Lupton R., Gunn J. E., Ivezić Z., Knapp G. R., Kent S., 2001, in Harnden Jr. F. R., Primini F. A., Payne H. E., eds, Astronomical Society of the Pacific Conference Series Vol. 238, Astronomical Data Analysis Software and Systems X. p. 269 (arXiv:astro-ph/0101420)
Mecke K. R., Buchert T., Wagner H., 1994, A&A, 288, 697
Nakagami T., Matsubara T., Schmalzing J., Jing Y., 2004, ArXiv Astrophysics e-prints,
Padmanabhan N., et al., 2008, ApJ, 674, 1217
Parejko J. K., et al., 2013, MNRAS, 429, 98
Percival W. J., et al., 2014, MNRAS, 439, 2531
Petri A., Haiman Z., Hui L., May M., Kratochvil J. M., 2013, Phys. Rev. D, 88, 123002

Pier J. R., Munn J. A., Hindsley R. B., Hennessy G. S., Kent S. M., Lupton R. H., Ivezić Ž., 2003, AJ, 125, 1559

Planck Collaboration et al., 2014a, A&A, 571, A23

Planck Collaboration et al., 2014b, A&A, 571, A25

Planck Collaboration et al., 2016, A&A, 594, A16

Platzöder M., Buchert T., 1996, in Weiss A., Raffelt G., Hillebrandt W., von Feilitzsch F., Buchert T., eds, Astro-Particle Physics. p. 251 (`arXiv:astro-ph/9509014`)

Reid B., et al., 2016, MNRAS, 455, 1553

Sahni V., Sathyaprakash B. S., Shandarin S. F., 1998, ApJ, 495, L5

Sathyaprakash B. S., Sahni V., Shandarin S., 1998, ApJ, 508, 551

Schmalzing J., 1999, in Banday A. J., Sheth R. K., da Costa L. N., eds, Evolution of Large Scale Structure : From Recombination to Garching. p. 127

Schmalzing J., Buchert T., 1997, ApJ, 482, L1

Schmalzing J., Kerscher M., Buchert T., 1996, in Bonometto S., Primack J. R., Provenzale A., eds, Dark Matter in the Universe. p. 281 (`arXiv:astro-ph/9508154`)

Schmalzing J., Buchert T., Melott A. L., Sahni V., Sathyaprakash B. S., Shandarin S. F., 1999, ApJ, 526, 568

Slepian Z., Eisenstein D. J., 2015, MNRAS, 454, 4142

Slepian Z., et al., 2017, MNRAS, 469, 1738

Smee S. A., et al., 2013, AJ, 146, 32

Smith J. A., et al., 2002, AJ, 123, 2121

Wiegand A., Eisenstein D. J., 2017, MNRAS, 467, 3361

Wiegand A., Buchert T., Ostermann M., 2014, MNRAS, 443, 241

York D. G., et al., 2000, AJ, 120, 1579

Yoshiura S., Shimabukuro H., Takahashi K., Matsubara T., 2017, MNRAS, 465, 394

Yuan S., Eisenstein D. J., Garrison L. H., 2017, MNRAS, 472, 577

Zheng Z., Zehavi I., Eisenstein D. J., Weinberg D. H., Jing Y. P., 2009, ApJ, 707, 554

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