Observing relativistic features in large-scale structure surveys – I: Multipoles of the power spectrum

C. Guandalin1 *, J. Adamek2,3, P. Bull2,4, C. Clarkson2,4, L. R. Abramo1 and L. Coates2

1Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, Rua do Matão 1371, CEP 05314-970, São Paulo, Brazil
2School of Physics & Astronomy, Queen Mary University of London, 327 Mile End Road, London E1 4NS, United Kingdom
3Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
4Department of Physics and Astronomy, University of Western Cape, Cape Town 7535, South Africa

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Planned efforts to probe the largest observable distance scales in future cosmological surveys are motivated by a desire to detect relic correlations left over from inflation, and the possibility of constraining novel gravitational phenomena beyond General Relativity (GR). On such large scales, the usual Newtonian approaches to modelling summary statistics like the power spectrum and bispectrum are insufficient, and we must consider a fully relativistic and gauge-independent treatment of observables such as galaxy number counts in order to avoid subtle biases, e.g. in the determination of the $f_{NL}$ parameter. In this work, we present an initial application of an analysis pipeline capable of accurately modelling and recovering relativistic spectra and correlation functions. As a proof of concept, we focus on the non-zero dipole of the redshift-space power spectrum that arises in the cross-correlation of different mass bins of dark matter halos, using strictly gauge-independent observable quantities evaluated on the past light cone of a fully relativistic $N$-body simulation in a redshift bin $1.7 \leq z \leq 2.9$. We pay particular attention to the correct estimation of power spectrum multipoles, comparing different methods of accounting for complications such as the survey geometry (window function) and evolution/bias effects on the past light cone, and discuss how our results compare with previous attempts at extracting novel GR signatures from relativistic simulations.

Key words: cosmology: relativistic effects, power spectrum, redshift-space distortions; methods: numerical, statistical

1 INTRODUCTION

The next generation of galaxy surveys – such Euclid, VRO/LSST, and SKA – will be both wide and deep, covering a broad range of redshifts as well as large areas of the sky, therefore mapping out an unprecedentedly large volume of space and time. On the one hand, this will significantly increase the amount of information available for existing types of cosmological analyses, reducing the same variance uncertainties on observables such as the BAO scale, redshift-space distortions, and the lensing shear power spectrum. On the other hand, the sheer size of these surveys will also allow qualitatively different cosmological observations to be made. In particular, they will be large enough to access modes on the order of the matter-radiation equality scale $k_{eq}$ (e.g. Philcox et al. 2020), and possibly even up to the comoving horizon scale $k_{h} \sim \langle aH \rangle$. These represent the very largest observable scales in the Universe, where novel observational features of inflationary and gravitational physics arise that cannot be constrained on the smaller scales probed by existing surveys (e.g. Liguori et al. 2010; Camera et al. 2015; Baker & Bull 2015; Alonso et al. 2015; Alonso & Ferreira 2015; Fonseca et al. 2015; Raccanelli et al. 2016; Gomes et al. 2020; Weltman et al. 2020).

On such large scales, corrections to the standard flat-sky/distant-observer approach to modelling effects such redshift space distortions emerge (c.f. Kaiser 1984), leading to so-called relativistic corrections or relativistic effects. They have been shown to be an important source of systematic error on large scales, especially for a potential detection of the scale-dependent bias in the galaxy distribution that would be caused by primordial non-Gaussianity (Camera et al. 2015; Raccanelli et al. 2016; Wang et al. 2020). This manifests as an additional $k^{-2}$ scaling in the bias of dark matter tracers (Dalal et al. 2008), which comes from non-linear corrections to the primordial Bardeen potential due to primordial non-Gaussianities of the local type (Komatsu & Spergel 2001). Relativistic terms with similar $k^{-2}$ scalings also become important on comparable scales (e.g. see Alonso et al. 2015; Abramo & Bertacca 2017), and so an accurate accounting of them is crucial if we are to recover an unbiased estimate of the non-Gaussianity parameter $f_{NL}$, for example.

Relativistic effects are not only a complicating factor, but contain novel information on the nature of gravity in their own right. Within the context of GR, several unique non-Newtonian features emerge due to such effects. For example, McDonald
(2009) has shown that relativistic effects induce odd multipoles to appear in the cross power spectrum of dark matter tracers, a characteristic with no Newtonian counterpart (Bonvin, Hui & Gaztanaga 2016; Gaztanaga, Bonvin & Hui 2017; De Weerd et al. 2020). This is, by itself, a new cosmological observable allowing us to probe the equivalence principle at cosmological scales via the Euler equation (Bonvin & Fleury 2018), as well as the gravitational redshift effect (McDonald 2009; Bonvin, Hui & Gaztanaga 2014). Moreover, by bearing a strong dependence on the Weyl potential, this provides an alternative test for theories of gravity, while the dependence on astrophysical parameters like the magnification and evolution biases opens a new window to a better understanding of the LSS. Other approaches to constraining deviations from GR via the behaviour of the relativistic effects have also been considered, e.g. Lombriser, Yoo & Koyama (2013); Baker & Bull (2015).

In this paper, we develop the basic building blocks of an analysis pipeline that is capable of extracting the relativistic effect signatures from large-scale structure data. As previously mentioned, standard LSS analysis techniques often rely on Newtonian assumptions or the distant-observer approximation, and so it is necessary to adapt them in order to account for the relativistic effects. Relativistic effects also introduce additional dependencies on the astrophysical properties of the source galaxy population(s) that must be accounted for, such as the magnification bias and evolution bias. Using a mock dark matter halo catalogue extracted from the past light cone of a fully-relativistic RayGal simulation1 and has the advantage of having previously been detected in the tipoles. The dipole is the most straightforward to model and detect, and has the advantage of having previously been detected in the

https://cosmo.obspm.fr/raygalgroupsims-relativistic-halo-catalogs/

Fourier Transform (FFT) estimator for the power spectrum multipoles (Bonvin & Durrer 2011; Bonvin 2014). The number overdensity of some tracer $\alpha$ can thus be defined as

$$\delta_\alpha^{(s)}(s) = \frac{N_\alpha(z, \hat{n}) - \bar{N}_\alpha(z)}{\bar{N}_\alpha(z)} + \frac{\delta V(z, \hat{n})}{V(z)},$$

where the equality is obtained by relating the number counts with the number density as $n(z, \hat{n}) = N(z, \hat{n})/V(z)$. In the above equation, $\bar{N}_\alpha(z)$ is the selection function of the tracer $\alpha$, obtained by angular averaging over the tracer number count.

The quantities defined in equation (1) are in redshift space, meaning that they are characterized by the observed (comoving) coordinates $s = (s, \theta, \phi)$, with the radial comoving coordinate $s$ being connected to the observed redshift by some cosmological model2. The standard treatment (Kaiser 1984), relating the number of sources in a perfect Friedmann-Lemaître-Robertson-Walker (FLRW) universe with the truly observed density field via the conservation of number counts, gives rise to the so-called redshift-space distortions. This allows us to relate the theoretical predictions in a homogeneous universe with the observed quantities with the addition of departures from the perfect FLRW metric.

In Kaiser (1984), corrections to the angular pair of coordinates $(\theta, \phi)$ are not considered, and perturbations to the radial coordinate $s$ come solely from the peculiar velocities of the sources. Even though it describes satisfactorily observations limited to sub-horizon scales, where the Newtonian treatment is well suited, this is not a truly observed quantity, as it is gauge-dependent. Furthermore, future galaxy surveys and cosmological observations that rely on the largest (near-horizon) scales demand a proper treatment of the LSS clustering. At smaller scales, the improved sensitivity will also hold the potential for a detection of subleading corrections (for an example see Saga et al. 2020).

Relativistic corrections that appear by considering the co-variant definition of redshift have been widely developed in the past decade, and became a paradigm to study large cosmological scales. In addition to solving well-known gauge issues manifested at these scales, it accounts for a number of effects with no Newtonian counterpart. For instance, gravitational redshift and lensing effects are concisely included in equation (1), and we refer the reader to equation (3.23) of Yoo (2014) and equation (16) of Bonvin (2014) for its full expression.

By collecting the terms proportional to $v \cdot n$ we end up with (Bonvin 2014; Clarkson et al. 2019):

$$\delta_\alpha^{(s)}(r) = b_\alpha \delta^{(s)}(r) - \frac{1}{k} \hat{H}_e \cdot (v \cdot n) + A_\alpha (v \cdot n),$$

where

$$A_\alpha = \frac{5s_\alpha - 2}{H_l} + b_\alpha + \frac{\hat{H}_e}{H^2} - 5s_\alpha,$$

is called Doppler term, $H^{-1} \hat{H}_e (v \cdot n)$ is the standard Kaiser term,

3 The radial comoving coordinate in redshift space $s$, obtained from the observed redshift, should not be confused with the magnification bias $s_\alpha$ of some tracer $\alpha$, which will carry a Greek index throughout this work. We also draw the reader’s attention to the radial comoving coordinate denoted by $v$ in real space, obtained from the unperturbed (Hubble flow) redshift of a perfect FLRW universe.

1 https://github.com/gevolution-code

2 https://cosmo.obspm.fr/raygalgroupsims-relativistic-halo-catalogs/
\( \mathcal{H} = aH \) is the comoving Hubble factor, \( s_\alpha \propto \partial_\ln r \ln(r^2 \phi_\alpha) \) is called magnification bias, 
\[
\beta_r = -(1 + z) \frac{\partial \ln \bar{n}}{\partial z}
\]  
(4) is the evolution bias and \( \beta_s \) is the linear bias. With the exception of the true density perturbation \( \delta_\alpha \), all other terms appear due to departures from a perfect FLRW universe.

Within the linear theory, we can relate quantities in configuration space with their Fourier counterpart to arrive at the main equation
\[
\delta_a(\mathbf{k}) = \delta^f(\mathbf{k}) \left[ b_\alpha + f \frac{\mu_\mathbf{k}}{k} + i f(\mathcal{H} k^{-1}) A_\alpha \mu_\mathbf{k} \right],
\]  
(5)

with \( \mu_\mathbf{k} \equiv (\mathbf{k} \cdot \hat{r}) \) to keep the explicit dependence with the LOS.

Assuming that all objects in the survey possess the same LOS, i.e. \( \mathbf{k} \cdot \hat{r} = \mu \) is a constant (flat-sky approximation), the cross-spectrum \( P_{\alpha \beta}^{(s)}(k) = \langle \delta_\alpha(k) \delta_\beta^*(k) \rangle \) of two tracers \( \alpha \) and \( \beta \) is given by:
\[
P_{\alpha \beta}^{(s)}(k) = P^{(r)}(k) \left[ (b_\alpha + f \mu^2)(b_\beta + f \mu^2) + A_\alpha A_\beta f_\mu^2 \frac{\mathcal{H}^2}{k^2} ight]
\]
\[+ i f \mu \left[ (b_\beta + f \mu^2) A_\alpha - (b_\alpha + f \mu^2) A_\beta \right] \frac{\mathcal{H} \mathcal{H}_k}{k}.
\]  
(6)

In this equation, \( \alpha \) and \( \beta \) refers to distinct tracers, which could be different types of galaxies or dark matter halos of different masses, \( f \) is the growth rate, parametrised by \( f(z) \sim \Omega_m(z)^{\gamma} \), with \( \gamma \) being the growth index, and \( P^{(r)}(k) \) is the matter power spectrum in real space.

In this case, isotropy is broken by the choice of LOS and we can expand \( P_{\alpha \beta}^{(s)}(k) = P_{\alpha \beta}^{(s)}(k, \mu) \) in a Legendre series:
\[
P^{(s)}(k, \mu) = \sum_{l=0}^{\infty} P^{(s)}_l(k) L_l(\mu),
\]  
(7)

where
\[
P^{(s)}_l(k) = f^{(r)}_l(k) c_f.
\]  
(8)

Neglecting the quadratic terms \( O(\mathcal{H}^2/k^2) \), the coefficients of the expansion are given by
\[c_0(f, b) = b_\alpha b_\beta + \frac{1}{3} f(b_\alpha + b_\beta) + \frac{1}{3} f^2 \]  
(9)
\[c_1(k, f, b, A) = \frac{1}{5} f^2 \mathcal{H}^2 k \left[ A_\alpha(3f + 5b_\alpha) - A_\beta(3f + 5b_\alpha) \right] \]  
(10)
\[c_2(f, b) = \frac{2}{3} (b_\alpha + b_\beta) + \frac{4}{7} f^2 \]  
(11)
\[c_3(k, f, A) = \frac{2}{5} i f^2 \mathcal{H}^2 k (A_\alpha - A_\beta) \]  
(12)
\[c_4(f) = \frac{8}{35} f^2 \]  
(13)

In the absence of these quadratic corrections, the monopole, quadrupole and hexadecapole are the same as in the Newtonian case. Still, the imaginary term appearing from the relativistic corrections in equation (2) gives rise to the dipole term manifested in the cross-spectrum of LSS tracers:
\[
P_{1 \beta}^{(s)}(k) = f^2 \mathcal{H}^2 k \left[ A_\alpha(3f + 5b_\alpha) - A_\beta(3f + 5b_\alpha) \right] P^{(r)}(k).
\]  
(14)

While it scales as \( \mathcal{H}/k \) for the cross-correlation of LSS tracers, a fact that makes this signal a smoking gun for relativistic effects in the galaxy clustering, it is identically zero for the auto-correlation. We also call the reader’s attention to the fact that this dipole term is anti-symmetric, meaning that \( \langle \delta_\alpha(k) \delta_\beta^*(k) \rangle = -\langle \delta_\beta(k) \delta_\alpha^*(k) \rangle \).

In Figures 1 and 2 we illustrate the dipole term in both the Fourier and configuration spaces, respectively, for three linear and evolution bias differences (different colours) at a fixed redshift \( z = 2 \).

In what follows we explore the detection of (14) in a relativistic simulation of a light cone, described in Section 3. Since we will be dealing with dark matter halos, the magnification bias \( s_\alpha \) in the Doppler term vanishes. Therefore, in addition to the linear bias of the halos, the remaining parameter entering the theoretical predictions is the evolution bias (4). The procedure for fitting \( b_\mu \) from the halo samples is described in Appendix A3, with the results discussed in Section 4.

3 SIMULATION

In this work we make use of a large \( N \)-body simulation performed with the relativistic code gevolution (Adamek et al. 2016a,b). The simulation has a comoving volume of \((2.4 \text{ Gpc}/h)^3\) with dark matter particles of mass \(2.64 \times 10^9 \text{(M}_\odot/h)\), and represents a typical \( \Lambda \text{CDM} \) cosmology: \( h = 0.67556, \omega_\text{b} = 0.022032, \omega_\text{cdm} = 0.12038, T_{\text{CMB}} = 2.7255 \text{ K}, A_s = 2.215 \times 10^{-9}, n_s = 0.9619, N_{\text{eff}} = 3.046, \) and \( N_{\text{cdm}} = 0 \). In order to avoid replications in the light cone, the pencil beam was carefully oriented in the periodic domain. The initial conditions for the simulation were set at a redshift of \( z = 127 \).

Unlike the standard approach to building light cones (Merson et al. 2012; Smith et al. 2017; Breton et al. 2019), which consists
of generating many simulation snapshots with a sufficient small redshift step between them to avoid time discretization effects in the final light cone, the light cone output from gevolution records particle positions and velocities on the fly. During the simulation, particles are identified that are within a proper comoving distance interval from a predefined observer that would cause them to be placed in the final catalogue. These particles are then shifted by a fractional time step and recorded on the null FLRW hypersurface given by the past light cone of the observer. Hence, there are no time discretization artefacts and no need to generate an enormous amount of snapshots to build the light cone. In our case, no replications whatsoever were performed in order to cover the whole light cone volume, which has the advantage of removing any concerns about spurious correlations on large scales due to periodicity for example.

The gevolution code does not employ the adaptive mesh refinement (AMR) method and thus has a low accuracy at small scales. However, while AMR can improve the 1-halo term by better resolving halo substructures, it does not significantly impact the large scales dominated by the 2-halo term, which is the focus of this work. As will be pointed out in Section 3.2, all sub-halos are discarded in our analysis in any case.

3.1 Ray tracing

We apply a ray tracing algorithm to our simulation as a post-processing tool. The algorithm was previously described in Lepori et al. (2020), but we give a brief review of it here.

The purpose of the ray tracer is to add extra information on source objects within the simulation to the catalogue, such as

Figure 2. Same as Figure 1, but for the cross-correlation function dipole of different tracers at redshift $z = 1.9$. Differences in the linear and evolution bias are shown in the legend. Solid lines represent the case where there is no magnification bias $\alpha_s = 0$, whereas shaded regions represent the effect of different magnification biases among the tracers. Dotted lines show the limiting case where $\alpha_s$ is smaller than $\alpha_s$ by 40%, whilst the dashed ones shows the opposite case, with $\alpha_s$ larger than $\alpha_s$ by a factor of 40%.

3.2 Halo catalogue

From the real space particles, the halo catalogue was created with the Rockstar halo finder (Behroozi, Wechsler & Wu 2012), using a friends-of-friends (FOF) algorithm with linking length $b = 0.28$ in order to detect over $10^7$ halos in the light cone.

After going through the ray-tracer algorithm, which is crucial to connect the halos and the observer, the perturbed three-dimensional positions of halos were obtained and the resulting file consists of three mock surveys contained within the range $0.0 \lesssim z \lesssim 7.1$, with different survey areas. The survey that will be used in this work spans the range of comoving look-back distance from 275 up to 4560 Mpc/$h$.

We limit ourselves to the high-redshift region between $z_{\text{min}} = 1.7$ and $z_{\text{max}} = 2.9$, with redshift bins of size $\Delta z = 0.4$ which kept their angular diameter distance ($DA$) relative to a specific observer, the respective observed redshift ($z$), or the ellipticity ($\epsilon$) which is closely related to the weak-lensing shear ($\gamma$). In contrast to the more common case where ray tracing is applied to Newtonian N-body simulations, in gevolution the metric perturbations and the source positions are both provided in Poisson gauge, which makes the treatment of gauge issues transparent. Our algorithm also does not rely on the Born approximation to model the light path. Importantly, incorrectly modelling the lensing probability distribution function can lead to errors in estimating cosmological parameters, as shown in (e.g.) Adamek et al. (2019).

The algorithm is similar to the one presented in Breton et al. (2019), and works by integrating the geodesic equations backwards in time from the observer to the source of interest on the observer’s past light cone. A physical definition of source, such as a halo or a dark matter particle, is required, as a 4-velocity vector is needed to define the source’s rest frame. This allows us to get the observed redshift of the source in a gauge-independent way. For each of these sources, we use the background FLRW model to give us the initial direction vector ($n$) for each light ray towards a source. We then integrate backwards in time with the fully perturbed metric until the light ray reaches its closest approach to the event on the light cone. At this point, we can now calculate a “deflection angle” by which the initial $n$ must be corrected to achieve a closer approach to the source. We repeat this process several times until suitable convergence is achieved.

This process works well in the weak-lensing regime, as only a single null ray exists between the observer and each source. In the strong lensing regime, multiple images can be formed, which complicates matters. The number of sources where this phenomenon is observed is negligible however, and so we concentrate only on weak lensing. Since strong lensing will only affect our results on very small scales where an image could be duplicated, this choice has a negligible impact on our analysis.

Ray tracing is the key step in properly incorporating relativistic corrections in our analysis. For example, instead of using the redshift output directly from the halo finder which would only include the background expansion and the Doppler correction, we are able to use the ‘observed’ redshift, which includes all relativistic effects. We can also calculate the perturbed position of sources on the sky, which is important for any n-point correlation calculations done using the catalogue. The algorithm also output $DA$ and both the real and the imaginary parts of the shear component separately ($\gamma_1 + i\gamma_2 \approx -\frac{\alpha}{s}$), although these are not needed in the current analysis.

$\gamma_1 + i\gamma_2 \approx -\frac{\alpha}{s}$. Although these are not needed in the current analysis.
redshift selection, we were left with the population incorporating all halos is referred to as the comparison with the Tinker et al. (2010) fitting function. The halo and dark matter (see Appendix A for a throughout discussion and the respective central redshift. The biases were computed by fitting the monopole of the correlation function and the values shown in the Table as being sensitive, for example, to the different intrinsic brightness of galaxies.

The selection function gives an estimate of the probability that a galaxy brighter than a certain threshold, at a distance $s$, is included in the sample. Hence, it is intrinsically related to the notion of luminosity function $Φ(L)$ (Martínez & Saar 2002). In Wang et al. (2020) a clear example of such fact is given, with the luminosity function of eBOSS quasars (QSO) being used to fit the QSO number density and derive the evolution and magnification biases.

To resume the construction of the estimator, $N_X(x_{ijk})$ denotes either the count-in-cells of the the random catalogue, $X = r$, or the data (halo) catalogue, $X = h$, where $x_{ijk}$ is the position of each cell in a three-dimensional grid obtained by a mass assignment scheme, e.g. Nearest Grid Point (NGP), Cloud In Cell (CIC), or Triangular Shaped Cloud (TSC). In this analysis we consider the simplest NGP assignment.

We begin by defining the weighted galaxy fluctuation, or the overdensity field, as

$$F(x) = \frac{w(x) [n_h(x) - α n_i(x)]}{N},$$

where $n_h(x)$ is the density number which will be written as a grid, after mass assignment scheme is chosen. Therefore, in practice $n_h(x)$ is $N_h(x_{ijk})$ is the count-in-cells grid and $n_i(x)$ is the corresponding quantity for the random catalogue, which is obtained by randomly sampling $α^{-1}$ times more objects within the survey volume, with the same selection function as the real data.

The results presented here do not employ a weighting scheme, i.e. $w(x) = 1$, and we follow Jeong (2010) for the implementation of the quadratic estimator. The normalization factor will be given by

$$N \approx \alpha^2 \ell_x \ell_y \ell_z \sum_{x_{ijk}} N_i^2(x_{ijk}),$$

and the shot noise, only relevant for the monopole term, will be

$$P_{shot} \approx \ell_x \ell_y \ell_z \left( \frac{1 + α}{α} \right) \sum_{x_{ijk}} N_i^2(x_{ijk}).$$

Table 1. Specifications of the halo samples, selected to match the number density for each population, yielding $n_h = n_i - n_r$. The mean redshift $\bar{z}$ is obtained from all the halos within each redshift bin, as the effective redshift of each halo sample differs from $\bar{z}$ by a sub-percent level ($\lesssim 0.5\%$). The biases have been computed from the monopole of the correlation function by fitting a constant to the ratio between the halo auto-correlation and the matter monopole in real space (see Appendix A2), differing from the Tinker bias by $\sim 5\%$. For the three redshift bins, the volumes are such that the fundamental mode of observation is $k_p = 2\pi/V^{1/3} = 7 \times 10^{-3}$ (h/Mpc).

The high-redshift binning was chosen to deliver a reasonable volume necessary for the observation of the relativistic features at large scales, giving an effective fundamental mode of $k_p = 2\pi/V^{1/3} = 7 \times 10^{-3}$ (h/Mpc). In a future work we will present the results of the same analysis, but in the full-sky case. The current survey area of $\sim 400$ deg$^2$ is compatible with the current survey areas available for a cross-correlation analysis (Zhao et al. 2020).

The high redshift binning was chosen to deliver a reasonable volume necessary for the observation of the relativistic features at large scales, giving an effective fundamental mode of $k_p = 2\pi/V^{1/3} = 7 \times 10^{-3}$ (h/Mpc). In a future work we will present the results of the same analysis, but in the full-sky case. The current survey area of $\sim 400$ deg$^2$ is compatible with the current survey areas available for a cross-correlation analysis (Zhao et al. 2020).

The final halo catalogue was then separated into three halo samples per redshift bin, of different masses such that, at each redshift bin, the number of halos was the same in each sample. The main properties of these samples are detailed in Table 1. Because more massive halos are expected at lower redshifts, the effective redshift $\bar{z}$ of each halo sample varies slightly, but only by less than $0.5\%$; therefore, we considered the values shown in the Table as the respective central redshift. The biases were computed by fitting the ratio between the real space power spectrum of the halos and dark matter (see Appendix A for a thorough discussion and comparison with the Tinker et al. (2010) fitting function). The halo population incorporating all halos is referred to as $H_{hal}$ in what follows.

4 POWER SPECTRUM MULTIPLE ESTIMATOR

To compute the power spectrum multipoles we make use of the standard approach proposed by Yamamoto et al. (2006); Bianchi et al. (2015) and Scoccimarro (2015) (for pioneering work see also Yamamoto, Nishioka & Taruya 2000), which we dub YBS estimator. It is built upon the practical algorithm developed by Feldman, Kaiser & Peacock (1994) to optimally estimate the power spectrum of galaxy surveys with a varying selection function. As mentioned in section 2, the selection function $N$ encodes the spatial modulations of the mean number density of objects. For both spectroscopic and photometric surveys, the selection function accounts for all non-cosmological effects, being sensitive, for example, to the different intrinsic brightness of galaxies.

The selection function gives an estimate of the probability that a galaxy brighter than a certain threshold, at a distance $s$, is included in the sample. Hence, it is intrinsically related to the notion of luminosity function $Φ(L)$ (Martínez & Saar 2002). In Wang et al. (2020) a clear example of such fact is given, with the luminosity function of eBOSS quasars (QSO) being used to fit the QSO number density and derive the evolution and magnification biases.

To resume the construction of the estimator, $N_X(x_{ijk})$ denotes either the count-in-cells of the random catalogue, $X = r$, or the data (halo) catalogue, $X = h$, where $x_{ijk}$ is the position of each cell in a three-dimensional grid obtained by a mass assignment scheme, e.g. Nearest Grid Point (NGP), Cloud In Cell (CIC), or Triangular Shaped Cloud (TSC). In this analysis we consider the simplest NGP assignment.

We begin by defining the weighted galaxy fluctuation, or the overdensity field, as

$$F(x) = \frac{w(x) [n_h(x) - α n_i(x)]}{N},$$

where $n_h(x)$ is the density number which will be written as a grid, after mass assignment scheme is chosen. Therefore, in practice $n_h(x)$ is $N_h(x_{ijk})$ is the count-in-cells grid and $n_i(x)$ is the corresponding quantity for the random catalogue, which is obtained by randomly sampling $α^{-1}$ times more objects within the survey volume, with the same selection function as the real data.

The results presented here do not employ a weighting scheme, i.e. $w(x) = 1$, and we follow Jeong (2010) for the implementation of the quadratic estimator. The normalization factor will be given by

$$N \approx \alpha^2 \ell_x \ell_y \ell_z \sum_{x_{ijk}} N_i^2(x_{ijk}),$$

and the shot noise, only relevant for the monopole term, will be

$$P_{shot} \approx \ell_x \ell_y \ell_z \left( \frac{1 + α}{α} \right) \sum_{x_{ijk}} N_i^2(x_{ijk}).$$

where $\ell_i \equiv L_i/n_i$ is the size of each cell dimension in units of Mpc/h. In this work we choose $\ell_i = 10$ Mpc/h.

After the construction of these quantities, the power spectrum multipoles can be obtained by the YBS estimator, which we now briefly discuss. The whole idea of this method relies in generalizing the power spectrum to local regions in space, where statistical homogeneity may be assumed. These regions are defined by a single middle line-of-sight $d = (s_1 + s_2)/2$, as shown on the left of the volume.

$$F(k) = (2\pi)^{-4} \int \delta^3(q)W(|k - q|) \Phi(q),$$

and one can show that $F(k)F(-k) = (2\pi)^{-4} \int \delta^3(q)W(|k - q|)^2 \Phi(q) + \int \delta^3(q)W(|k - q|) \Phi(q)$. Therefore, in Feldman, Kaiser & Peacock (1994) it is considered the overdensity field divided by the magnitude of the window function, $W = \int \delta^3 W(x) d^3x$, which we called $N$ (Jeong 2010).

This criterion was chosen to keep halos of different evolutionary stages somewhat separated.
Figure 3. Then, the corresponding power spectrum at this region is

\[ P(k_1, k_2) = \int d^3s_1 \int d^3s_2 \xi(s_1, s_2) e^{i\mathbf{k}_1 \cdot \mathbf{s}_1} e^{i\mathbf{k}_2 \cdot \mathbf{s}_2}. \]  

(18)

Notice that

\[ e^{i\mathbf{k}_1 \cdot \mathbf{s}_1} e^{i\mathbf{k}_2 \cdot \mathbf{s}_2} = e^{i\mathbf{d} \cdot \mathbf{q}} e^{-i\mathbf{k} \cdot \mathbf{s}}. \]

(19)

where the Fourier transform of the local configuration is shown on the right of Figure 3. With this change of coordinates, one can see that the local power spectrum can be obtained by taking the Fourier transform of the first component of the local correlation function \( \xi(s_1, s_2) = \xi(s, d) \). Finally, the multipoles of the local power spectrum can be obtained from the Legendre expansion

\[ P(k, d) = \sum_{\ell} P_\ell(k, d) \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{d}}). \]

(20)

Lastly, the inversion of this relation yields the power spectrum multipoles:

\[ \hat{P}_\ell(k) = \left( \frac{2\ell + 1}{2\pi} \right) \int d^3s_1 \int d^3s_2 F(s_1) F(s_2) \times e^{-i(k \cdot s_1 - k \cdot s_2)} S_\ell' - S_\ell \right), \]

(21)

where the brackets correspond to an average over k-shells and \( S_\ell \) is the shot-noise term, only relevant for the monopole.

In order to speed up the computation of the multipoles by means of FFTs, the YBS estimator takes the end-point LOS \( s_1 \). It is worth mentioning that this LOS intrinsically generates odd multipoles which may impact the signal we are trying to measure. Hence, this must be accounted for in the window function, as discussed in Appendix B.

With the adoption of this LOS, the monopole can estimated as

\[ \hat{P}_0(k) = \frac{1}{N} \left[ F_0(k) F_0'(k) - S_0 \right] \]

(22)

where

\[ F_0(k) = \int d^3x F(x) e^{i\mathbf{k} \cdot \mathbf{x}} \]

(23)

is the Fourier transform of the overdensity field with no weight,

\[ F(x) = n_0(x) - a n_r(x), \]

(24)

and the dipole is obtained by

\[ \hat{P}_1(k) = \frac{3}{N} \left( F_0(k) F'_0(k) \right), \]

(25)

where

\[ F_1(k) = \sum_{i=x,y,z} \hat{k}_i f_{i,1}(k), \]

(26)

and

\[ f_{i,1}(k) = \int d^3r \hat{r}_i F(r) e^{i\mathbf{k} \cdot \mathbf{r}}. \]

(27)

For higher order multipoles, we refer the interested reader to Bianchi et al. (2015) and Beutler, Castorina & Zhang (2019) for the even and odd ones, respectively.

Finally, to compute the cross-dipole, \( F_0(k) \) and \( F_1(k) \) are built from the first and second tracers, respectively, and the normalization becomes (Beutler & Di Dio 2020):

\[ N \approx \frac{\alpha_{\text{B}}(1)\beta_{\text{B}}(2)}{\ell_{x}^y \ell_{z}^x} \sum_{ij,k} N_{\ell}^{(1)}(x_{ijk}) N_{\ell}^{(2)}(x_{ijk}). \]

(28)

\[ P(k, d) = \sum_{\ell} P_\ell(k, d) \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{d}}). \]

(20)

5 RESULTS

For the objects under analysis (dark matter halos), the concept of luminosity function can replaced by the halo mass function, \( d\bar{n}/d\ln M \), which gives the probability of having a mean number of halos, within some comoving volume, with mass in the range \([\ln M_i, \ln (M_i + dM)]\). Thereby, the selection function coincides with the comoving mean number density of halos within a certain mass bin \([M_i, M_{i+1}]\), in complete analogy to the definition of selection function from \( \Phi(L) \):

\[ \bar{n}(\Delta M) = \int_{M_i}^{M_{i+1}} \frac{d\bar{n}}{d\ln M} d\ln M. \]

(29)

Still for the specific case of dark matter halos, the magnification bias is identically zero, \( s = 0 \), and the only term accounting for the mass function variations is the evolution bias, which explores its dependency with time: this is related to the fact that halos can merge to form more massive structures and, thereby, their number counts are not conserved. The evolution biases \( b_{\ell}^{(s)}(z) \) of each halo sample \( i = \{H_0, H_1, H_2, H_{\text{all}}\} \) considered in this work, within each redshift bin, is shown in Figure 4. The procedure to compute \( b_{\ell}^{(s)}(z) \) is described in Appendix A and is based on the work of Beutler & Di Dio (2020).

We work with 29 bandpowers (Fourier bins) linearly spaced between \( k_{\text{min}} \approx 0.006 \) and \( k_{\text{max}} \approx 0.157 \) with \( \Delta k \approx 0.005 \). As already mentioned in Section 4, we work with three-dimensional grids containing \( n_x = n_y = n_z = n_x \) cells of side \( \ell = 10 \text{ Mpc}/h \), with the number of cells \( n_i \) varying between the different redshift bins. This binning was chosen to deliver a less noisy measurement at large scales.

Figure 5 shows the monopole estimated from the four halo samples in the first redshift bin (\( \bar{z} \approx 1.89 \)). Apart from the amplitude of the monopole, not much change occurs between different redshifts, hence we only show the first \( \bar{z} \)-bin here. Our error

\[ \text{See Appendix A for a detailed discussion on the mass function of our samples.} \]
bars are estimated from the standard deviation of 100 log-normal mocks generated with the same characteristics of the original data: box dimensions, evolution and linear biases, selection function and survey mask. For this last step, we first generated the log-normal mocks for the whole box encompassing the different redshift bins of the light cone, and then applied the proper mask to select the specific angular region. However, because these error bars only quantify the variance of the estimator, the particular survey features do not matter much, and thus it should be possible to compute the variance of mocks inside the box, with the only caveat of following the mean number density of tracers to properly incorporate the shot noise. The impact of the window function in the standard deviation of the log-normal samples generated minor changes at very large scales, and we do not explore this further.

To what concerns the variance of our measurements, the window function had a negligible impact. Nonetheless, to avoid any sort of complications, we explore the asymmetry of the relativistic signal as suggested in Beutler & Di Dio (2020): by computing $\Delta P_1 = P^{\beta\alpha}_1 - P^{\alpha\beta}_1$ it is possible to isolate the relativistic contribution and get rid of the impact of the window function, which is symmetric. As pointed in Section 2, the Doppler term is anti-symmetric, $\langle \delta_\alpha(k) \delta_\beta^* (k) \rangle = -\langle \delta_\beta(k) \delta_\alpha^* (k) \rangle$, and thus $\Delta P_1 \sim 2 \langle \delta_\alpha(k) \delta_\beta^* (k) \rangle$.

Consistently, the theoretical prediction for the last cross-correlation $H_2 \times H_{\text{all}}$ is positive, as $\Delta b = b_2 - b_{\text{all}} > 0$. However, as shown in Figure 6, we conclude that no detection can be claimed with this pencil-beam light cone. Very similar results were obtained for the other two redshift samples $z_1 \approx 2.29$ and $z_2 \approx 2.69$.

Still, we point out the possibility of exploring optimal weighting schemes to enhance the signal in the light of the work carried out by Castorina et al. (2019). Lastly, the advantages of employing different tracers, coupled with the low densities of the catalogues, as well as the need for robust statistics, suggests the use of optimal weights for a more efficient combination of tracers (Abramo, Secco & Loureiro 2016; Montero-Dorta et al. 2020).

6 CONCLUSIONS

We have performed a power spectrum multipole analysis on data from a light cone generated from a fully-relativistic N-body simulation. We focused on the dipole signal in the cross-correlation between different dark matter halo sub-populations, which is a purely relativistic (non-Newtonian) effect. The simulation was generated by the gevolution code, which employs a novel ray-tracing method to connect the halos with the observer, and which is capable of incorporating all relevant general relativistic effects on cosmologically-relevant distance scales. We showed in detail how the survey window function and quantities such as the evolution bias can be estimated on the past light cone, allowing a rigorous comparison with gauge-invariant theoretical calculations at linear order.

Similar studies of relativistic observables in simulations have
been made in the past. For example, Breton et al. (2019) and Beutler & Di Dio (2020) used the full-sky RayGal simulation, which is limited to the redshift range of $0.05 < z < 0.465$, with an effective $\bar{z} \sim 0.341$. While the simulation that we based our study on in principle covers the redshift range $0 \leq z \leq 7.1$, our analysis focused on a particular high redshift bin in the range $1.7 \leq z \leq 2.9$, covering a sky fraction of only $f_{\text{sky}} = 0.01$. This is a similar sky area to the overlap region between different LSS tracers (luminous red galaxies and emission line galaxies) in the multi-tracer analysis of the final eBOSS data (see Table 2 of Zhao et al. 2020), although these data are from lower redshift, $z \sim 1$.

While we were able to robustly test our analysis methods using these simulated data, no conclusive detection of the dipole signature was possible due to the limited volume of the redshift bin, a challenge that is of paramount importance for current surveys too. Beutler & Di Dio (2020) studied the possibility of subtracting various contributions to the total signal in order to isolate the Doppler contribution and remove sample variance. Since the Doppler term is expected to increase in amplitude with redshift, one could also consider developing an optimal weighting scheme to enhance the signal and improve the prospects of detection (Castorina et al. 2019). We leave this, and other schemes (Abramo, Secco & Loureiro 2016; Abramo & Bertacca 2017; Montero-Dorta et al. 2020) to enhance detectability of the signal, to be explored in future work however.

We did not incorporate wide-angle effects in our modelling, as they are not relevant for the solid angle and redshift range of our analysis. A careful account of these effects should also be explored in the context of wider survey areas however, particularly in the case of future surveys such as Euclid, LSST, and SKA, which are expected to cover an appreciable fraction of the sky.

Similarly, integrated effects (e.g. lensing), while fully included in our mock data, were neglected in our analytical model, but are known to impact large angular scales. Despite the Doppler term being the largest contribution to the relativistic effects for our particular setup, non-local terms should also be modelled and properly included for analyses that go to larger scales.

In this paper, we have limited our analysis to a single high-redshift bin with a relatively narrow survey area, and have pursued only a limited set of observables, i.e. the multipoles of the relativistic power spectrum. In future work, we will relax these limitations by moving to larger survey volumes more representative of the next generation of large-scale structure surveys, while also including wide-angle and integrated effects, and extending our analysis to two-point correlation functions and multipoles of the relativistic bispectrum.

ACKNOWLEDGEMENTS

We thank David Alonso for making CUTE public. CG thanks Rodrigo Voivodic, Florian Beutler, and Michel-Andrés Breton for comments and fruitful discussions. In particular, she owes special thanks to Marcos Lima for all the support during the first stages of this work. She would also like to thank Queen Mary University of London for hospitality. CG was supported by QMUL (PHY2420B) and FAPESP (2018/10396-2) grants. JA acknowledges funding by STFC Consolidated Grant ST/P000592/1. CC is supported by STFC Consolidated Grant ST/P000592/1. LRA acknowledges financial support from CNPq (306696/2018-5) and FAPESP (2015/17199-0 and 2018/04683-9). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s710.

Carbon footprint: In this work we re-used existing data from a simulation that consumed about 8000 kWh of electrical energy. This has an estimated impact of 1600 kg CO$_2$ when we use the conversion factor of 0.2 kg CO$_2$ kWh$^{-1}$ suggested by Vuarnoz & Jusselme (2018) (see Table 2 therein, assuming Swiss mix). The additional energy used during the numerical analysis of the data is insignificant in comparison. This work also included a round trip São Paulo ↔ London economy flight, emitting approximately 900 kg CO$_2$.

REFERENCES

Abramo, L. R., Secco, L. F. & Loureiro, A., 2016, MNRAS, 455(4), 3871
Abramo, L. R. & Bertacca, D., 2017, Phys. Rev. D, 96(12), 123535

---

8 ICAO Carbon Emissions Calculator, 25 August (2020).
Appendix A: Halo Properties

A1 Halo mass function

For completeness we computed the mass function of our full halo sample between $z = 0.05$ and $z = 0.465$ (effective redshift $\bar{z} = 0.34$) for comparison with the full-sky RayGal simulation$^9$ employed in the analysis of Breton et al. (2019). The halo mass function describes the probability of having a comoving number density of halos at redshift $z$ in the range $[\ln M, \ln M + d\ln M]$:

$$d\bar{n}(M, z) = \frac{\bar{\rho}_m(z)}{M} f(\sigma) d\ln \sigma, \quad \frac{d\ln M}{d\ln M}.$$  (A1)

with $\bar{\rho}_m(z)$ the comoving background matter density today, $f(\sigma)$ the multiplicity function, and $\sigma$ the overdensity variance smoothed in a sphere of radius $R$. The multiplicity function can be computed analytically from the spherical collapse model (Press & Schechter 1974) or the ellipsoidal collapse (Sheth & Tormen 1999), or from numerical fits (Tinker et al. 2008).

The RayGal simulation consists of a set of high-resolution Newtonian N-body simulations, whose halos have been ray-traced to the redshift space position, rendering them with almost all properties of our halos. The RayGal light cone was built from 300 snapshots to avoid time discretization effects.

Our analysis was based on halo masses $M_{200b} \equiv M$ defined within the density thresholds of $\Delta = 200$, whose correspondence with the parameters fit of the Tinker mass function (Tinker et al. 2008) (solid lines in Figure A1) is straightforward. In Corasaniti et al. (2018), RayGal halo mass functions were computed from the snapshots and were based on the Sheth-Tormen (Sheth & Tormen 1999) fit, with the halo identified with the spherical overdensity (SO) method, and thus the halos are more closely connected to the ellipsoidal collapse employed in the Sheth-Tormen fit (Desjacques, Jeong & Schmidt 2018).

However, for the RayGal light cone, halos were identified via a friends-of-friends (FOF) algorithm, just like in our catalogue. In Smith et al. (2017), differences with numerical fits seem at the low-mass end are also present, and they conclude that such discrepancies are associated with the comparison between different halo finder methods (SO and FOF). They computed the mass function using the SO correspondent, and just as in our case, found the same behaviour at low masses. In the catalogue employed in our analysis, $\sim 1.1 \times 10^6$ halos with $M_{200b} \in [0.518, 4.862] \times 10^{12} (M_\odot/h)$ were discarded for having their respective $M_{200b}$ null. As pointed out in Smith et al. (2017), small overdensities in large FOF groups might be identified as part of the larger group, leading to a lack of such structures.

Such discrepancies are important if one wishes to paint galaxies to the halos via, e.g., a halo occupation distribution. For our current purposes, the lack of a proper function to describe the light cone halo mass function impacted only our ability to predict the linear halo bias (see next section), and thus did not pose an issue for the analysis.

A2 Halo bias

The halo mass function describes the fraction of matter inside dark matter halos. So in order to obtain the correct halo statistics, we must account for their position in space. The halo bias, which is defined by the ratio of the halo power spectrum, $P_{hh}(k)$, to the linear dark matter power spectrum, $P_{ll}(k)$ (Tinker et al. 2010),

$$b^2(k) = \frac{P_{hh}(k)}{P_{ll}(k)},$$  (A2)

is best understood within the context of the Peak-Background Split (PBS), where the long-wavelength modes enhances the probability

$^9$ The RayGal simulation is contained within $0.05 < z < 0.465$, their effective redshift is the same as the low redshift considered in this section for comparative purposes.
of forming halos by decreasing the threshold \(\delta_c(\bar{z} = 0) = 1.686\) for overdensities which are located at the peak of large-scale (background) fluctuations. It can be either derived from analytical mass functions, giving the Press-Schechter and Sheth-Tormen halo biases, or from equation A2 via numerical simulations. From Tinker et al. (2010), the bias is given by

\[
b(v) = 1 - A \frac{\nu^B}{\nu^C + \nu^D} + Bv^B + Cv^C, \tag{A3}
\]

where \(\nu = \delta_c/\sigma\) and \(A, B, C, a, b,\) and \(c\) are parameters fitted from simulations, depending on the matter perturbations at virialization which is chosen to be \(\Delta = 200\). This phenomenological fit proved to be unsatisfactory for our halo samples, for the reasons described in the previous section.

We proceeded then to the definition of equation (A2), with the polynomial fit

\[
b^*(k) = b_1^* + b_2^* k, \tag{A4}
\]

and considered the linear term as the fit for the linear halo biases, neglecting the scale dependence emerging from non-linear effects in the power spectrum. Notice that the estimated spectra \(P_{bh}\) employed in this fit are for the real space halos. The results are shown in Table A1 for each halo sample.

This approach worked reasonably well for the first redshift slice, but proved to be inaccurate for the higher redshift samples, where the impact of the window function was larger. The most likely reason is the mode coupling with the window function, present in our real space light cone.

Hence, instead of convolving the theory with the window function, as done in Breton et al. (2019), we adopted a similar approach to Breton et al. (2019): the linear bias was computed by fitting a constant function to the ratio

\[
\frac{d\bar{n}}{d\ln M} = \frac{\bar{n}}{M^{1/3} \bar{H}} \left( \frac{d\bar{n}}{d\ln M} \right)_{\text{RayGal}} \tag{A5}
\]

where \(\bar{\nu}_0\) is the monopole of the matter auto-correlation function, computed from

\[
\bar{\nu}_0(x) = \int \frac{dk}{2\pi^2} j_0(k_{\parallel})P^r(k_{\parallel}), \tag{A6}
\]

and \(\bar{\nu}_{hh}^0\) is monopole of the halo-halo auto-correlation, computed from the real space catalogues with CUTE\(^\text{10}\) (Alonso 2012). The real space linear matter power spectrum\(^\text{11}\) was obtained from the CLASS Boltzmann solver (Blas, Lesgourgues & Tram 2011) with the input parameters of the simulation.

This method is not perfect though: we observed a shift in the BAO peak scale for the halo samples if compared to the matter power spectrum. Also, we limited ourselves to the range \(28 < r_{\text{fit}} < 68\) in units of Mpc/h. Despite being time consuming, it is safe from the mode coupling induced by the window function. The results from this fit is shown in Table A1, last column, and proved to be a better fit for all redshift slices.

### A3 Evolution bias

The evolution bias of LSS tracers quantifies the intrinsic variation in the number of sources in the Universe, and thus gives information about the time evolution of tracers. It is defined in equation (4), and depends on the comoving number density of sources in real space \(\bar{n}\). In the case of dark matter halos, this parameter

---

\(^\text{10}\) https://github.com/damonge/CUTE/

\(^\text{11}\) Differences of using the non-linear matter power spectrum were below the percent level.
describes the mean number density of halos. The evolution bias account for the true tracer: in the case of a halo catalogue, coming from a simulation, the underlying dark matter distribution is known and the comoving number density of halos is complete (i.e. all halos that were supposed to be found are included in the catalogue), and the evolution bias is completely faithful to the intrinsic cosmological variations of the number density.

In Figure A2 we show the comoving number density of real space halos, inside each redshift bin considered in the main analysis, normalized by the total number of halos for each mass bin defined for \( \tilde{H}_0 \approx \tilde{H}_1 \approx \tilde{H}_2 \). As one can see, there is a large variation explained by the fact that more massive halos are more common at lower redshifts, which can be explicitly seen by the slope of the curves. This intrinsic variation is captured by the evolution bias parameter \( b_\epsilon \), and is a major parameter entering the relativistic corrections.

Following Beutler & Di Dio (2020), we fit a linear function to the (unnormalized) comoving number density, \( \tilde{n}(z) = a + b_\epsilon \tilde{z} \), which leads to the analytical expression for \( b_\epsilon \),

\[
\tilde{b}_\epsilon(z) \equiv \frac{c + 1}{e - \tilde{z}} - 1.
\]

where \( c \equiv a/(b) \).

Notice that, even though the comoving mean number density is the same for all the three samples \( \tilde{H}_0, \tilde{H}_1 \) and \( \tilde{H}_2 \), the different evolution with redshift between the different halo populations, defined by different halo masses, leads to distinct evolution biases, as can be seen in Figure 4.

**APPENDIX B: WINDOW FUNCTION**

In this appendix we describe how the window function is obtained. It is employed to compute the observed power spectra multipoles from the theoretical predictions for comparative purposes and imparts substantial effects on large scales and on the odd multipole moments. Therefore, its inclusion is mandatory.

We begin by recalling that the observed density field \( \hat{\delta}(x) \) is given by

\[
\hat{\delta}(x) = W(x)\delta(x),
\]

where \( \delta(x) \) is the true underlying density field and \( W(x) = w(x)\tilde{b}(x) \) accounts for the survey geometry and local weighting \( w \) scheme (Feldman, Kaiser & Peacock 1994). Therefore, the observed correlation function is given by

\[
\hat{\xi}(s_1, s_2) = W(s_1)W(s_2)\xi(s_1, s_2).
\]

Notice that we can write \( s_2 = s_1 + s \), where \( s \) is the pair separation, so that the correlation function may also be written as \( \hat{\xi}(s, x) \).

In Fourier space we obtain the well-known convolution result for the overdensity field,

\[
\hat{\delta}(k) = \int \frac{d^3k'}{(2\pi)^3} W(k-k')\hat{\delta}(k').
\]

yielding the three-dimensional observed power spectrum:

\[
\hat{P}(k) = \int \frac{d^3q}{(2\pi)^3} |W(k-q)|^2 P(q),
\]

where we made use of the fact that \( \delta \) and \( W \) are real quantities.

One possible way to compare theory and estimates is to deconvolve the survey window from \( \hat{P} \), however, since convolution in Fourier space destroys information, the deconvolution of the window is an attempt to recover this intrinsic information loss in the signal analysis. The standard procedure (Wilson et al. 2017; Beutler et al. 2017) consists, instead, in computing the multipoles of \( |W(q)|^2 \) to convolve the theoretical power spectrum to obtain \( \hat{P} \), where

\[
|W(q)|^2 = \int d^3s e^{-i q \cdot s} W^2(s),
\]

and

\[
W^2(s) = \int d^3s_1 W(s_1)W(s_1 + s).
\]

We shall write \( W^2(s) = Q(s) \) and \( |W(q)|^2 = Q(k) \). Notice that this depends on the local LOS, which is taken to be \( s_1 \) (end-point LOS) in the case of the YBS multipole estimator. Hence, the multipoles of the “window function”, with respect to a LOS \( \hat{d} = \hat{s}_1 \), are given by

\[
Q_{\ell}(s) = \frac{2\ell + 1}{4\pi} \int d^3d \int d\Omega_\xi(s, d)P_\ell(\hat{s} \cdot \hat{d})
\]

in configuration space and

\[
Q_{\ell}(k) = \frac{2\ell + 1}{4\pi} \int d^3d \int d\Omega_k Q(s, d)P_\ell(\hat{k} \cdot \hat{d})
\]

in Fourier space. The integrals on \( d\Omega_\xi \) and \( d\Omega_k \) run over the angles between, respectively, \( \hat{s} \) and \( \hat{k} \) with the LOS \( \hat{d} \): \( \int d\Omega = \int_{\Omega_0} d\theta \sin \theta \int_{\Omega_0} d\phi \). After the integration over all angles, \( Q_{\ell}(s) \) can be obtained by the final integration over all possible LOS.

With the theoretical window function, obtained from its definition in equation (B7), differences between halo populations as a result of different selection functions were below the 2% level for the auto-correlation case, with the same behaviour being observed at the estimated windows from the random catalogues. In contrast, differences between redshift bins are more relevant, as can be seen in Figure B1, and are fully included in our analysis.

If we plug in the expressions for \( Q(k, d) \) and \( Q(s, d) \), just as in the local power spectrum case, we see that to estimate the multipoles \( Q_{\ell}(k) \) and \( Q_{\ell}(s) \) we just apply the usual power spectrum and correlation function estimators. Since the explicit convolution of equation (B4) is computationally expensive, and so is the computation of \( Q_{\ell}(s) \) directly from the random pair correlation, as the
Figure A2. The comoving number density of halos, normalized by the total amount $N_0$ of objects inside each mass bin. Solid lines represent the linear fit of equation (A7), whereas dashed lines represent the true $\bar{n}(z)$ of the simulation, computed by dividing each redshift bin into 50 points. The smooth curves are the result of a cubic spline interpolation, for visual reasons.

Figure B1. Multipoles of the window function computed from the inverse Hankel transform of YBS estimator for the random catalogues of the halo population $H_0$. Different line styles represent different multipoles, dots depict the even and vertical bars the odd ones. Different colours show the window function at different redshift slices. Differences between the tracers as a result of different selection functions are negligible, and thus we only show the auto-correlation of the $H_0$; however, the intrinsic fluctuations of the FFT estimator makes the window function very noisy at large scales, causing a variation between redshifts (and also tracers) larger than what is expected from equation (B7).

survey window function (random catalogue) contains $10^8$ particles to completely fill the survey region, one possibility is to compute the power spectrum multipoles of the random catalogues. We adopt this approach, obtaining $Q_\ell(k)$ by means of FFTs, which are then taken to configuration space.

From the straightforward product of equation (B2), the Legendre expansion of (B2) results in (Wilson et al. 2017; Beutler et al. 2017; Beutler, Castorina & Zhang 2019; Beutler & Di Dio 2020):

$$\hat{\xi}_0(s) = \xi_0(s)Q_0(s) + \frac{1}{5}\xi_2(s)Q_2(s) + \frac{1}{5}\xi_4(s)Q_4(s) + \ldots$$

(B9)

and

$$\hat{\xi}_1(s) = \xi_0(s)Q_1(s) + \xi_2(s)\left[\frac{2}{5}Q_1(s) + \frac{9}{35}Q_3(s)\right] + \frac{4}{2T} \xi_4(s)Q_3(s) + \ldots$$

(B10)

From an inverse Hankel (1D Fourier) transform of equations (B9) and (B10) we finally obtain the convolved power spectrum multipoles of equation (B4) which we use for comparison with the estimated quantities.

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