Beyond the traditional line-of-sight approach of cosmological angular statistics

Journal of Cosmology and Astroparticle Physics
An IOP and SISSA journal

Beyond the traditional line-of-sight approach of cosmological angular statistics

Nils Schöneberg,[a] **Marko Simonović,**[b,c] **Julien Lesgourgues**[a] **and Matias Zaldarriaga**[b]

[a]RWTH Aachen University, Institute for Theoretical Particle Physics and Cosmology (TTK), Sommerfeldstraße 16, Aachen, D-52074 Germany
[b]School of Natural Sciences, Institute for Advanced Study, 1 Einstein Dr, Princeton, NJ, 08540 U.S.A.
[c]Theoretical Physics Department, CERN, Geneva 23, CH-1211 Switzerland

E-mail: schoeneberg@physik.rwth-aachen.de, marko.simonovic@cern.ch, lesgourg@physik.rwth-aachen.de, matiasz@ias.edu

Received August 9, 2018
Accepted October 15, 2018
Published October 25, 2018

Abstract. We present a new efficient method to compute the angular power spectra of large-scale structure observables that circumvents the numerical integration over Bessel functions, expanding on a recently proposed algorithm based on FFTlog. This new approach has better convergence properties. The method is explicitly implemented in the `CLASS` code for the case of number count C_ℓ's (including redshift-space distortions, weak lensing, and all other relativistic corrections) and cosmic shear C_ℓ's. In both cases our approach speeds up the calculation of the exact C_ℓ's (without the Limber approximation) by a factor of order 400 at a fixed precision target of 0.1%.

Keywords: cosmological perturbation theory, power spectrum

ArXiv ePrint: 1807.09540

https://doi.org/10.1088/1475-7516/2018/10/047

JCAP10(2018)047

Contents

JCAP10(2018)047

1 Introduction

One important pathway to constrain cosmological parameters is the measurement of the two-point statistics of large-scale structure. For galaxy redshift surveys, this measurement is often cast in a three-dimensional redshift-dependent Fourier power spectrum $P(k, z)$. However, this quantity is not directly observable, because redshift surveys do not map objects in four-dimensional spacetime, but rather in the three-dimensional past-light-cone, with one redshift and two angular coordinates. Thus $P(k, z)$ can only be reconstructed by assuming a fiducial cosmology and by going through some rather complicated steps. A more natural and straightforward way to describe the two-point statistics of a galaxy catalogue is through the angular number count power spectra C_ℓ's for a list of redshift bins, including auto- and cross-correlations between bins. The bin width usually reflects the measurement error of the (photometric or spectroscopic) survey.

Although the angular spectra are a much more direct representation of galaxy redshift survey data, analysis pipelines based on the matter power spectrum are often preferred for computational reasons. Indeed, an accurate calculation of the angular spectra in many redshift bins (beyond the Limber approximation) is very time-consuming compared to the calculation of the matter power spectrum. This problem is worsened when all the General

Relativity (GR) corrections to the number count C_ℓ's, presented in [1–6] and implemented in the codes `CAMB_sources`[1] [4, 7] and `CLASS`[2] [6, 8], are taken into account. Then, even some idealized Fisher matrix forecasts can become computationally expensive [9, 10], and even more so MCMC parameter extractions. For instance, reference [9] presented a Fisher matrix forecast of the sensitivity of a future hypothetical spectroscopic survey to cosmological parameters, taking all GR corrections into account. This forecast had to be limited to $N_z \leq 32$ redshift bins in order to avoid computing times exceeding the timescale of one week on a small computing cluster, while the figure of merit for such a survey was expected to saturate for $N_z \geq 100$. The solution to this computing issue is not just to switch off all GR corrections: while some of them can always be safely neglected, the Redshift-Space Distortions (RSDs) will have to be taken into account in the analysis of future high-precision galaxy redshift surveys. The weak lensing contributions can also be significant, especially in the case of cross-correlation spectra between redshift bins [9]. In conclusion, in order to efficiently analyze data from future galaxy redshift surveys with the angular spectrum method, we really need a breakthrough on the computational side.

For weak lensing (i.e. cosmic shear) surveys, the data is more often expressed in terms of angular power spectra C_ℓ's in a list of redshift bins, with an important role of the cross-correlation spectra even for non-adjacent bins (since weak lensing effects are correlated for sources laying at different redshifts). The Limber approximation based on the matter power spectrum $P(k, z)$ is often used at high ℓ's in order to speed up the calculation, but the low-ℓ part of the calculation remains slow, and the matching between the exact spectrum and the Limber spectrum at some intermediate ℓ value is a source of inaccuracy.

For both redshift and cosmic shear surveys, there exist some alternative ways to represent the two-point statistics of the data, such as two-point correlation functions in redshift shells (see [11]). An algorithm efficiently computing these functions has recently been described in [12]. The purpose of this paper is, however, to stick to the angular spectra and to re-visit the numerical method used for calculating C_ℓ's, with a new way to separate the cosmological and geometrical information, even more efficient than the traditional line-of-sight approach of [13]. The main ideas of this method were suggested recently in [14, 15], and are based on a power law decomposition of the source functions, which allows to separate the integrals into one part depending on cosmology and one part depending on geometry. The integration of spherical Bessel functions, which are quickly oscillating and only slowly damped, can then be performed analytically. The new method speeds up the calculation of the number count and cosmic shear C_ℓ's by one to two orders of magnitude (when weak lensing corrections are taken into account, and also for number count). Thus our work contains a step towards practical implementations of angular power spectra in the analysis pipeline of future galaxy redshift surveys and cosmic shear surveys.

In section 2, we explain the general approach, including new developments with respect to reference [14]. In section 3 we discuss our practical implementation of this method in `CLASS`, including all GR corrections to the number count spectra C_ℓ's, and introducing the idea of "approximate separability". Finally, in section 4, we give a report on the performance of the new code (for a fixed level of precision) and conclude. The reader is encouraged to look at [14] and [15] for more theoretical and technical details on this method and its application to other types of spectra (such as the CMB power spectrum or general bispectra).

[1] https://camb.info/sources/.
[2] http://class-code.net/.

JCAP10(2018)047

2 Method

2.1 The traditional line-of-sight approach

Any observable $\mathcal{O}_\alpha(\hat{\boldsymbol{n}}, z)$ can be expanded for every redshift z in spherical harmonics $Y_{\ell m}(\hat{\boldsymbol{n}})$ with respect to its angular position on the sky $\hat{\boldsymbol{n}}$,

$$\mathcal{O}_\alpha(\hat{\boldsymbol{n}}, z) = \sum_{\ell,m} a^\alpha_{\ell m}(z) Y_{\ell m}(\hat{\boldsymbol{n}})\,, \qquad \text{where} \qquad a^\alpha_{\ell m}(z) = \int d\Omega_{\hat{\boldsymbol{n}}} Y^*_{\ell m}(\hat{\boldsymbol{n}}) \mathcal{O}_\alpha(\hat{\boldsymbol{n}}, z)\,. \tag{2.1}$$

The angular expansion coefficients $a^\alpha_{\ell m}$ can be used to define the angular power spectrum for the observables $\mathcal{O}_\alpha$, or more generally any cross-spectrum

$$C^{\alpha\beta}_\ell(z_1, z_2) \equiv \left\langle a^\alpha_{\ell m}(z_1)\; a^\beta_{\ell m}(z_2)^* \right\rangle\,. \tag{2.2}$$

It is well-known that the theoretical prediction for the coefficients $a^\alpha_{\ell m}(z)$ is given by a convolution of the Fourier transform of the observable in real space $\mathcal{O}_\alpha(\boldsymbol{k}, z)$ with a spherical Bessel function $j_\ell(k\chi(z))$,

$$a^\alpha_{\ell m}(z) = 4\pi i^\ell \int \frac{d^3k}{(2\pi)^3}\, j_\ell(k\chi(z))\, \mathcal{O}_\alpha(\boldsymbol{k}, z)\, Y^*_{\ell m}(\hat{\boldsymbol{k}})\,. \tag{2.3}$$

Here $\boldsymbol{k}$ is a Fourier wave number, $k = |\boldsymbol{k}|$ its modulus, $\hat{\boldsymbol{k}} = \boldsymbol{k}/k$ the corresponding unit vector, $\chi(z)$ is the comoving distance, given by

$$\chi(z) = \int_0^z \frac{dz'}{a_0 H(z')}\,, \tag{2.4}$$

and $H(z)$ is the Hubble parameter. In the following we will neglect the explicit dependence and write $\chi \equiv \chi(z)$. The quantity $\mathcal{O}_\alpha(\boldsymbol{k}, z)$ is usually computed by Einstein-Boltzmann solvers like `CAMB` or `CLASS`. It can be split into a transfer function $T_\alpha(k, z)$ and a primordial curvature perturbation $\mathcal{R}(\boldsymbol{k})$,[3]

$$\mathcal{O}_\alpha(\boldsymbol{k}, z) = T_\alpha(k, z)\mathcal{R}(\boldsymbol{k})\,. \tag{2.5}$$

The primordial curvature perturbation is randomly distributed according to a (nearly scale-independent) primordial spectrum $\mathcal{P}_\mathcal{R}(k)$, defined in the following way

$$\langle \mathcal{R}(\boldsymbol{k})\mathcal{R}(\boldsymbol{k}')\rangle \equiv \frac{2\pi^2}{k^3}\mathcal{P}_\mathcal{R}(k)\,\delta(\boldsymbol{k} - \boldsymbol{k}')\,. \tag{2.6}$$

Plugging (2.5) into (2.3) and finally into (2.2), we obtain

$$C^{\alpha\beta}_\ell(z_1, z_2) = 4\pi \int_0^\infty \frac{dk}{k}\; \mathcal{P}_\mathcal{R}(k)\; T_\alpha(k, z_1) j_\ell(k\chi_1)\; T_\beta(k, z_2) j_\ell(k\chi_2)\,. \tag{2.7}$$

Furthermore, in concrete observables like galaxy number count or cosmic shear, redshift measurement errors are described by averaging $\mathcal{O}_\alpha(\hat{\boldsymbol{n}}, z)$ over some window functions $\mathcal{W}^i(z)$ normalised to one (i is the index of a given redshift bin). The normalised window function in

[3] Some contributions (like RSD, Doppler effects) cannot *immediately* be written in this form, but we will show how they can be rewritten as such after using various tricks listed in section 3, like equation (3.6).

JCAP10(2018)047

comoving distance space is then $W^i(\chi) = \mathcal{W}^i(z)a_0H(z)$, where z is understood as a function of χ, such that

$$\int_0^\infty W^i(\chi)d\chi = \int_0^\infty \mathcal{W}^i(z)a_0H(z)d\chi = \int_0^\infty \mathcal{W}^i(z)dz = 1\,. \tag{2.8}$$

Averaging equation (2.7) over these windows leads to the expression

$$\begin{aligned} C_\ell^{\alpha\beta,ij} = 4\pi \int_0^\infty d\chi_1\, W^i(\chi_1) \int_0^\infty d\chi_2\, W^j(\chi_2) \\ \times \int_0^\infty \frac{dk}{k}\, \mathcal{P}_\mathcal{R}(k)\; T_\alpha(k,\chi_1) j_\ell(k\chi_1)\; T_\beta(k,\chi_2) j_\ell(k\chi_2)\,. \end{aligned} \tag{2.9}$$

The usual way to solve such an integral is to calculate

$$C_\ell^{\alpha\beta,ij} = 4\pi \int_0^\infty \frac{dk}{k}\, \mathcal{P}_\mathcal{R}(k)\, \Delta_\ell^{\alpha,i}(k)\, \Delta_\ell^{\beta,j}(k)\,, \tag{2.10}$$

where for every wavenumber k and every window function i the integral

$$\Delta_\ell^{\alpha,i}(k) = \int_0^\infty d\chi\, W^i(\chi) T_\alpha(k,\chi) j_\ell(k\chi) \tag{2.11}$$

has to be computed. The Spherical Bessel functions $j_\ell(x)$ introduce fast and slowly damped oscillations, making it hard to compute the integrals efficiently.

2.2 Separating cosmology and geometry

It has been argued in [14] that there is a simple way to circumvent the issue with the integrals over spherical Bessel functions. The main idea is to approximate the k-dependence of the product $\mathcal{P}_\mathcal{R}(k)T_\alpha(k,\chi_1)T_\beta(k,\chi_2)$ in the relevant wavenumber range $[k_{\rm min}, k_{\rm max}]$ using some set of simple basis functions, chosen such that the integral in k can be solved analytically. The proposal of [14] is to use the FFTlog to perform the power-law decomposition of the relevant functions

$$\mathcal{P}_\mathcal{R}(k)T_\alpha(k,\chi_1)T_\beta(k,\chi_2) = \sum_n c_n^{\alpha\beta}(\chi_1,\chi_2)\, k^{\nu_n}\,. \tag{2.12}$$

In this expansion we will refer to $c_n^{\alpha\beta}$ as Fourier coefficients and ν_n as Fourier frequencies for a given Fourier mode n. The details and the practical implementation of this decomposition will be described in the following sections. Let us for the moment assume that such approximation is possible and rewrite the k-integral in the following way

$$\begin{aligned} \int_0^\infty \frac{dk}{k} \mathcal{P}_\mathcal{R}(k)\, T_\alpha(k,\chi_1) j_\ell(k\chi_1)\, T_\beta(k,\chi_2) j_\ell(k\chi_2) = \\ \sum_n c_n^{\alpha\beta}(\chi_1,\chi_2) \int_0^\infty \frac{dk}{k}\, k^{\nu_n}\, j_\ell(k\chi_1) j_\ell(k\chi_2)\,. \end{aligned} \tag{2.13}$$

If we define

$$I_\ell(\nu,t) \equiv 4\pi \int_0^\infty \frac{du}{u}\; u^\nu\, j_\ell(u) j_\ell(ut)\,, \tag{2.14}$$

the angular power spectrum can be rewritten as

$$C_\ell^{\alpha\beta,ij} = \sum_n \int_0^\infty d\chi_1\, W^i(\chi_1) \int_0^\infty d\chi_2 W^j(\chi_2)\, c_n^{\alpha\beta}(\chi_1,\chi_2)\, \chi_1^{-\nu_n}\, I_\ell\left(\nu_n, \frac{\chi_2}{\chi_1}\right)\,. \tag{2.15}$$

JCAP10(2018)047

Finally, changing the integration variables to $\chi = \chi_1$ and $t = \chi_2/\chi_1$ we obtain a simplified expression

$$C_\ell^{\alpha\beta,ij} = \sum_n \int_0^\infty d\chi\, W^i(\chi) \int_0^\infty dt\, W^j(\chi t)\, c_n^{\alpha\beta}(\chi, \chi t)\, \chi^{1-\nu_n}\, I_\ell\left(\nu_n, t\right) . \tag{2.16}$$

Notice that this expression is formally the same as in (2.10). However, the spherical Bessel functions are now analytically integrated. The resulting function $I_\ell\left(\nu, t\right)$ has two important properties: first, it has a simple analytical form in terms of hypergeometric functions; secondly, it is smooth, which allows for straightforward numerical integration in t. These simplifications are crucial for the efficient numerical evaluation of $C_\ell^{\alpha\beta,ij}$.

Another important advantage of decomposition (2.12) is that all cosmology dependence is in the coefficients $c_n^{\alpha\beta}$. This means that the function $I_\ell(\nu, t)$ is cosmology-independent. This provides a motivation to rewrite the expression for $C_\ell^{\alpha\beta,ij}$ such that universal geometrical factors (like the function $I_\ell(\nu, t)$) and cosmology-dependent factors (like the coefficients $c_n(\chi, \chi t)$ and the window functions) are completely separated. To achieve this we can change the order of integrations

$$C_\ell^{\alpha\beta,ij} = \sum_n \int_0^\infty dt\, I_\ell\left(\nu_n, t\right) \int_0^\infty d\chi\, W^i(\chi) W^j(\chi t) c_n^{\alpha\beta}(\chi, \chi t) \chi^{1-\nu_n} , \tag{2.17}$$

and define a cosmology-dependent function $f_n^{\alpha\beta,ij}(t)$

$$f_n^{\alpha\beta,ij}(t) \equiv \int_0^\infty d\chi\, W^i(\chi) W^j(\chi t) c_n^{\alpha\beta}(\chi, \chi t) \chi^{1-\nu_n} . \tag{2.18}$$

Using this definition, we can write the angular power spectrum as the following integral

$$C_\ell^{\alpha\beta,ij} = \sum_n \int_0^\infty dt\, I_\ell\left(\nu_n, t\right) f_n^{\alpha\beta,ij}(t) . \tag{2.19}$$

The main virtue of this expression is separation of geometry and cosmology. The cosmology-dependent function $f_n^{\alpha\beta,ij}(t)$ depends on the observables $\mathcal{O}_\alpha$ and window functions W_i but it is *independent* of the multipole ℓ. On the other hand, the geometrical function $I_\ell(\nu, t)$ is cosmology-independent and it has to be calculated only once for any cosmology. (In practice it can be stored in a file and loaded when desired.) The function $f_n^{\alpha\beta,ij}(t)$ represent the correlation between the two window functions W^i and W^j weighed by a Fourier mode n represented by $c_n^{\alpha\beta}(\chi, \chi t)\chi^{1-\nu_n}$. The parameter t corresponds to the relative correlation distance in χ. These cosmological Fourier-weighed correlators $f_n^{\alpha\beta,ij}(t)$ are finally convolved with the geometrical Fourier-weighed correlators $I_\ell(\nu_n, t)$ and summed over all Fourier modes n used in the weighing.

If we try to compare this more directly to the traditional line-of-sight approach, the integral over t in (2.18) has similarities with the integral over χ in (2.11), in which one also convolves a cosmological function $W^i(\chi)T(k, \chi)$ with a geometrical function $j_\ell(k\chi)$. The final sum over n in (2.18) plays a role similar to the integral over k in (2.10). However, we will see that the new method offers the advantage of better behaved integrals, which results in increased precision and speed.

Before we discuss details and practical implementation of the power-law decomposition and evaluation of $I_\ell(\nu, t)$, let us finish this section by mentioning some symmetry relations

JCAP10(2018)047

which further simplify the expressions above. There are two such relations, which follow simply from a substitution of the integration variables. The geometrical function $I_\ell(\nu, t)$ satisfies

$$I_\ell\left(\nu, 1/t\right) = t^\nu I_\ell\left(\nu, t\right) , \tag{2.20}$$

and the cosmological function $f_n^{\alpha\beta,ij}(t)$ obeys

$$f_n^{\alpha\beta,ij}(1/t) = t^{2-\nu_n} f_n^{\beta\alpha,ji}(t) . \tag{2.21}$$

This allows us to write

$$\begin{aligned} C_\ell^{\alpha\beta,ij} &= \sum_n \int_0^\infty dt\, I_\ell\left(\nu_n, t\right) f_n^{\alpha\beta,ij}(t) \\ &= \sum_n \left(\int_0^1 dt\, I_\ell\left(\nu_n, t\right) f_n^{\alpha\beta,ij}(t) + \int_1^\infty dt\, I_\ell\left(\nu_n, t\right) f_n^{\alpha\beta,ij}(t) \right) \\ &= \sum_n \left(\int_0^1 dt\, I_\ell\left(\nu_n, t\right) f_n^{\alpha\beta,ij}(t) + \int_0^1 du/u^2\, I_\ell\left(\nu_n, 1/u\right) f_n^{\alpha\beta,ij}(1/u) \right) \\ &= \sum_n \left(\int_0^1 dt\, I_\ell\left(\nu_n, t\right) f_n^{\alpha\beta,ij}(t) + \int_0^1 dt\, I_\ell\left(\nu_n, t\right) f_n^{\beta\alpha,ji}(t) \right) , \end{aligned} \tag{2.22}$$

leading to the following result

$$C_\ell^{\alpha\beta,ij} = \sum_n \int_0^1 dt\, I_\ell\left(\nu_n, t\right) \left(f_n^{\alpha\beta,ij}(t) + f_n^{\beta\alpha,ji}(t) \right) . \tag{2.23}$$

Alternatively one may use

$$C_\ell^{\alpha\beta,ij} = \sum_n \int_0^1 dt\, I_\ell(\nu_n, t) \left(f_n^{\alpha\beta,ij}(t) + t^{\nu_n - 2} f_n^{\alpha\beta,ij}(1/t) \right) . \tag{2.24}$$

The difference between the two is how the functions $f_n^{\alpha\beta,ij}(t)$ have to be calculated. Since $C_\ell^{\alpha\beta,ij} = C_\ell^{\beta\alpha,ji}$, their calculation is required only for $i \geq j$. Then in equation (2.23) one has to calculate the $f_n^{\alpha\beta,ij}$ for all (i, j) combinations, while in equation (2.24) only $i \geq j$ are required.

In practice one has to be careful about evaluating $f_n^{\alpha\beta,ij}(1/t)$ with enough precision. In the integral over χ in (2.18), any grid in χ corresponds to a grid in χ/t, which appears as an argument of the functions W^j and $c_n^{\alpha\beta}$. When t becomes very small, one must check that the grid in χ/t does not become too coarse. Furthermore, we see that $t^{\nu_n - 2}$ is divergent for $t \to 0$ when $\Re[\nu_n] < 2$ (which is usually the case, as we shall see in section 2.3). For the case of $f_n^{\alpha\beta,ji}(t)$ these two problems do not arise, and the grid in χt is sampled even finer with smaller t.

Both problems can be handled by noting that in many cases $f_n^{\alpha\beta,ij}(1/t)$ only needs to be evaluated for t values far from 0, such that $1/t$ is never too large. Indeed, for most contributions to the number count spectra like e.g. density fluctuations or Redshift-Space Distortions, we can assume that the window functions $W^i(\chi)$ have a finite support $[{\chi_{\rm min}}^i, {\chi_{\rm max}}^i]$. Then, (2.18) shows that $f_n^{\alpha\beta,ij}(t)$ is non-zero only in the range $[t_{\rm min}^{ij}, t_{\rm max}^{ij}]$ with

$$t_{\rm min}^{ij} = {\chi_{\rm min}}^j / {\chi_{\rm max}}^i , \qquad t_{\rm max}^{ij} = {\chi_{\rm max}}^j / {\chi_{\rm min}}^i . \tag{2.25}$$

JCAP10(2018)047

Thus the function can even vanish entirely in the range $t \in [0,1]$, but only when ${\chi_{\rm min}}^j > {\chi_{\rm max}}^i$. If the redshift bins are arranged in growing redshift and distance order, this usually does not happen for $i \geq j$. However, $f_n^{\alpha\beta,ij}(1/t)$ and $f_n^{\beta\alpha,ji}(t)$ have a support given by

$$t^{ji}_{\rm min} = {\chi_{\rm min}}^i / {\chi_{\rm max}}^j \,, \qquad t^{ji}_{\rm max} = {\chi_{\rm max}}^i / {\chi_{\rm min}}^j \,, \tag{2.26}$$

and thus vanish entirely for any $t \in [0,1]$ if ${\chi_{\rm min}}^i > {\chi_{\rm max}}^j$, which may happen for some index pairs with $i \geq j$. Even if neither function vanishes entirely, the minimum of $t^{ij}_{\rm min}$ and $t^{ji}_{\rm min}$ provides an overall lower limit on t. We will also see in section 2.4 that the functions $I_\ell(\nu, t)$ will provide additional limits on $t_{\rm min}$. The existence of such a minimum value of t, and thus of a maximum value of $1/t$, tells us that we can perform integrals in the form of (2.24) without worrying too much about sampling or divergence issues in the limit $t \to 0$.

For other types of contributions to the spectra like e.g. weak lensing, we will see that the window functions are defined on a support starting from $\chi = 0$. In that case the previous discussion does not apply and we will go back to the integral in the form of (2.23).

In the practical implementation of the new method, we actually calculate the function $f_n^{\alpha\beta,ij}(t)$ using a rather coarse grid in t since it is very smooth. Only when convolving with the more oscillating function $I_\ell(\nu, t)$ will a more fine-grained grid be required. We will then interpolate within the coarse grid of $f_n^{\alpha\beta,ij}(t)$ values using Cubic Hermite Spline Interpolation.

2.3 The power-law (FFTlog) decomposition

In this section we will give details of the power-law decomposition in (2.12) and the FFTlog algorithm used to achieve it. The power-law expansion of a function $f(\boldsymbol{x}, k)$ is simply a Fourier decomposition in $\log(k)$. To see this let us consider a generic function $f(\boldsymbol{x}; k)$, periodic in the interval $[k_{\rm min}, k_{\rm max}]$ with the period $T \equiv \log(k_{\rm max}) - \log(k_{\rm min}) = \log(k_{\rm max}/k_{\rm min})$. The vector $\boldsymbol{x}$ describes an arbitrary set of extra variables. The coefficients of the logarithmic Fourier expansion of the function $f(\boldsymbol{x}, k)$ are given by

$$c_n(\boldsymbol{x}) = \frac{1}{T} \int_{\log(k_{\rm min})}^{\log(k_{\rm max})} f(\boldsymbol{x}, k) \exp\left(-\frac{2\pi i\, n}{T} \log(k)\right) d\log(k) \,. \tag{2.27}$$

Using a finite set of these Fourier coefficients we can approximate the original function as

$$\bar{f}(\boldsymbol{x}; k) = \sum_n c_n(\boldsymbol{x}) \exp\left[\frac{2\pi i\, n}{\log(k_{\rm max}/k_{\rm min})} \log(k)\right] = \sum_n c_n(\boldsymbol{x})\, k^{\nu_n} \,, \tag{2.28}$$

which can be easily rewritten in the desired form

$$\bar{f}(\boldsymbol{x}; k) = \sum_n c_n(\boldsymbol{x})\, k^{\nu_n} \,, \qquad \text{where} \qquad \nu_n = \frac{2\pi\, i\, n}{\log(k_{\rm max}/k_{\rm min})} \,. \tag{2.29}$$

The more terms are kept in this sum, the closer $\bar{f}(\boldsymbol{x}, k)$ is to the function $f(\boldsymbol{x}, k)$. In practice, the Fourier coefficients $c_n(\boldsymbol{x})$ can be computed binning the function in $\log(k)$ and using the Fast Fourier Transform (FFT) algorithm — hence the name "FFTlog" commonly used in cosmology literature [14–18] (see appendix B for further details on the relation to the FFT).

One immediate question is why this method is applicable to the case we are interested in, where the transfer functions or the primordial power spectrum are not periodic functions

JCAP10(2018)047

in k. The reason is that these smooth functions in k are always multiplied by the spherical Bessel functions which peak for $k \sim \ell/\chi$ and decay both in $k \to 0$ and $k \to \infty$ limit.[4] Therefore, the integral picks up most of its contribution from a finite range of scales and replacing the true function with its approximation (2.28) leads to negligible error even when the limits of integration are taken to be 0 and ∞. However, the asymptotic behavior of the spherical Bessel functions in two different limits is such that they do not approach zero equally fast

$$j_\ell(k\chi) \to k^\ell\,, \quad k \to 0\,, \qquad \text{and} \qquad j_\ell(k\chi) \to k^{-1}\,, \quad k \to \infty\,. \tag{2.30}$$

Depending on the observable and the form of the transfer functions, the result may be much more sensitive to high or low k. Therefore, to ensure better convergence properties, it is convenient in practice to use a slightly more general form of the Fourier transform which allows additional freedom in regulating the asymptotic behavior. This can be simply achieved by writing $k^b \cdot k^{-b} f(\boldsymbol{x}, k)$ and applying the logarithmic Fourier expansion to $k^{-b} f(\boldsymbol{x}, k)$ instead of $f(\boldsymbol{x}, k)$. In this way, one can always choose the "tilt" b to ensure more symmetric behavior of the integrand in the two limits. Effectively, this procedure shifts the Fourier frequencies ν_n in the following way

$$\nu_n = \frac{2\pi\, i\, n}{\log(k_{\rm max}/k_{\rm min})} + b\,. \tag{2.31}$$

We will give more details on the choice of b for different observables below.

The best convergence in (2.29) is reached when the function that we want to Fourier transform (i.e. $k^{-b} f(\boldsymbol{x}, k)$) in the range $[k_{\rm min}, k_{\rm max}]$ has equal values at the boundaries,

$$k_{\rm min}^{-b} f(\boldsymbol{x}, k_{\rm min}) = k_{\rm max}^{-b} f(\boldsymbol{x}, k_{\rm max})\,, \tag{2.32}$$

because it can then be considered as a sample of a periodic and continuous function in the range $[k_{\rm min}, k_{\rm max}]$. Our strategy is choosing b is such that $k^{-b} f(k) \longrightarrow 0$ on both ends of the interval. For the values of $k_{\rm min}$ and $k_{\rm max}$, we should in principle consider the finite range of scales to which the C_ℓ's that we want to compute are actually sensitive through their convolution kernels. Then $k_{\rm min}$ should be of the order of the scale crossing the Hubble radius today, $k_0 \sim a_0 H_0$, and $k_{\rm max}$ should depend on the maximum multipole and minimum redshift considered. Typically, one is interested in multipoles such that the non-linear corrections are either negligible, or small enough to be well under control: this typically limits the sensitivity of the C_ℓ's to maximum wavenumbers of the order of 1 to 10 inverse Megaparsecs. We will come back to the details of choice of $k_{\rm min}$ and $k_{\rm max}$ in later sections.

Given the previous discussion, we must ensure that $k^{-b} f(\boldsymbol{x}, k)$ has vanishing values at the boundaries of the interval $[k_{\rm min}, k_{\rm max}]$, where $f(\boldsymbol{x}, k)$ is the left-hand side of equation (2.12): it is simply the product of the dimensionless primordial spectrum $\mathcal{P}_{\mathcal{R}}(k)$ and two transfer functions $T_\alpha(k, \chi)$. The behaviour of this function in the small-scale and large-scale limits is known at least approximately. Observations confirm that the dimensionless primordial spectrum is equal or close to a power-law of the form $k^{n_s - 1}$. The transfer functions depend on the gauge and on the type of observable being computed. Here we discuss two different relevant cases.

[4] Indeed, this is the reason why in the commonly used Limber approximation [19–23] one can replace the Bessel function by a properly normalized Dirac function picking up only the contribution from the point $k = (\ell + 1/2)/\chi$. In other regions the integrand is either vanishingly small and/or highly oscillatory, thus not contributing significantly to the integral. However the Limber approximation is usually not accurate enough for fitting data from future Large Scale Structure surveys (especially when considering narrow window functions or cross-correlation between redshift bins).

JCAP10(2018)047

Density transfer functions. In the important case of density fluctuations, which usually dominate the number count spectra, $T_\alpha(k,\chi)$ has a well-known behaviour as a function of k. For any time after radiation-matter equality, it grows like k^2 between k_0 and the scale that crossed the Hubble radius at Matter-Radiation equality, $k_{\rm eq} \sim 10^{-2} h/{\rm Mpc}$. At the level of linear theory, it then keeps growing logarithmically on scales $k > k_{\rm eq}$. Non-linear corrections slightly increase the slope beyond the scale of non-linearity. Strong non-linear effects like baryonic feedback occur on even larger wavenumbers and are irrelevant for the range of multipoles usually used for cosmological parameter inference. The behavior that we just described is valid in any gauge, since gauge differences only show up on super-Hubble scales which are outside of our interval. Note that this precise behavior of the transfer function in the range $k_{\rm eq} < k < k_{\rm max}$ is not essential in our discussion, since our purpose is just to choose the tilt b such that $k^{-b} f(\boldsymbol{x}, k)$ is suppressed in $k_{\rm max}$ compared to its peak value.

In summary, the asymptotic behavior of the function $k^{-b} f(\boldsymbol{x}, k)$ is typically given by

$$\begin{aligned} k^{-b} f(\boldsymbol{x}, k) &\sim k^{n_s+3-b} && (k \ll k_{\rm eq})\,, \\ k^{-b} f(\boldsymbol{x}, k) &\sim k^{n_s-1-b} \ln(k)^2 && (k \gg k_{\rm eq})\,. \end{aligned} \tag{2.33}$$

Thus any tilt factor in the range $n_s - 1 < b < n_s + 3$ will ensure that (2.32) is fullfiled, with vanishing values at the boundaries compared to the peak value.

The choice of tilt is also important for the calculation of the geometrical function $I_\ell(\nu_n, t)$ given by (2.14). Indeed, the tilt b enters the definition of the coefficients ν_n in equation (2.31). Thus it is important to choose b such that within this range, the u integral of equation (2.14) is well behaved, such that $I_\ell(\nu_n, t)$ can be calculated. This condition was already investigated in [14] and imposes a range $-2\ell < \Re[\nu_n] < 2$. For $\Re[\nu_n] = b$, this overlaps with the previous range $n_s - 1 < b < n_s + 3$, for whatever realistic value of n_s. The optimal choice is found to lay somewhere between 1.5 and 1.9.

However, with $n_s \simeq 1$, $k_{\rm max} \simeq 1\ h/{\rm Mpc}$ and $b \simeq 1.9$, the value of $k^{-b} f(\boldsymbol{x}, k)$ at $k_{\rm max}$ is not very small: it is around 10% of the peak value. Stopping the integral around this scale would imply that we FFT-transform a periodic function with a discontinuity in $k_{\rm max}$. The FFT method is very sensitive to such discontinuities. If nothing is done to further reduce the difference between the values at the edge, this small discontinuity generates a small unphysical contribution to the FFT coefficient c_n and leads to a percent-level error in the final result. To solve this problem, one can introduce additional regulatory mechanisms. Knowing that for the relevant C_ℓ's, the precise form of the transfer function is not important for wavenumbers bigger than about one inverse Megaparsec, one has a lot of freedom for regularizing the function $k^{-b} f(\boldsymbol{x}, k)$ in the large k limit. The choice of reference [14] was to numerically evolve wavenumbers up to a $k_{\rm max}$ of $\sim 52 h/{\rm Mpc}$, and to multiply the transfer functions by an exponential cut-off around $k_{\rm cut} \sim 10 h/{\rm Mpc}$. In the `CLASS` implementation, we choose a computationally faster method: we extend the upper boundary in the integrals to about $k_{\rm max} \simeq 10^3 h/{\rm Mpc}$, but with an analytic extrapolation of the transfer function of the form $T(k) \sim \log(ak)$ above the maximum wavenumber for the numerical evaluation, set by default by `CLASS` for a redhsift bin of mean conformal distance χ_i to $k_{\rm num} = 2.4\,\ell_{\rm max}/\chi_i$ (typically of the order of $1h/{\rm Mpc}$). We stress that this extrapolation is not meant to be physical and is just a regulatory artifact. We tested different values of $k_{\rm max}$ in the range from $10h/{\rm Mpc}$ to $10^3 h/{\rm Mpc}$ and found that the result was very stable, varying by less than one permille.

Other transfer functions. We will present in section 3.3 a list of transfer functions which are useful for the calculation of number count spectra (with all GR corrections) and weak

JCAP10(2018)047

lensing spectra. We will see that the other transfer functions scale with k roughly in the same way as the density transfer function, but with an extra gobal factor k^{-2}. In that case, our strategy consists in doing an FFTlog transformation of $\bar{T}_\alpha(k,\chi) = k^2 T_\alpha(k,\chi)$, which then has the same behavior in k as the density contribution. For this, we write $T_\alpha(k,\chi) = k^{-2}\bar{T}_\alpha(k,\chi)$, and denote the FFT transform of the product $\mathcal{P}_\mathcal{R}(k)\bar{T}_\alpha(k,\chi_1)\bar{T}_\beta(k,\chi_2)$ as $\bar{c}_n^{\alpha\beta}(\chi_1,\chi_2)$. This changes equation (2.13) to

$$\int_0^\infty \frac{dk}{k}\, k^{-4}\, \mathcal{P}_\mathcal{R}(k)\, \bar{T}_\alpha(k,\chi_1) j_\ell(k\chi_1)\, \bar{T}_\beta(k,\chi_2) j_\ell(k\chi_2) = \sum_n \bar{c}_n^{\alpha_\beta}(\chi_1,\chi_2) \int_0^\infty \frac{dk}{k}\, k^{\nu_n - 4}\, j_\ell(k\chi_1) j_\ell(k\chi_2)\,. \tag{2.34}$$

Thus, this redefinition effectively replaces ν_n with $\nu_n - 4$ or $\nu_n - 2$, both in the definition of $f^{ij}(t)$ in equation (2.18) and in the argument of $I_\ell(\nu,t)$ in equations (2.23), (2.24). This implies that the values of the tilt such that the integrals $I_\ell(\nu,t)$ are well behaved now shift to either $-2\ell+2 < b < 4$ or $-2\ell+4 < b < 6$. In all these cases and for $\ell \geq 2$, the tilt value chosen in the density case, $b \simeq 1.9$, is still valid, and the Bessel integrals are well behaved.

Finally, let us note that the reference [14] proposes an additional trick to shift the tilt b by using an integrations by part. In the present implementation in `CLASS` we did not experience the need to use this technique. It is further discussed in section 3.3.

2.4 The analytical Bessel integrals

We conclude the discussion on the method by making a few comments on evaluation of Bessel integrals $I_\ell(\nu,t)$. While [14] and [15] provide explanations of how to obtain $I_\ell(\nu,t)$ in a fast and accurate manner, we focus here on several aspects raised when working with finite precision arithmetics.

The $I_\ell(\nu,t)$'s can be expressed in terms of Hypergeometric functions for $t < 1$,

$$I_\ell(\nu,t) = \frac{2^{\nu-1}\pi^2 \Gamma\left(\ell+\frac{\nu}{2}\right)}{\Gamma\left(\frac{3-\nu}{2}\right)\Gamma\left(\ell+\frac{3}{2}\right)}\; t^\ell\; {}_2F_1\left(\frac{\nu-1}{2}, \ell+\frac{\nu}{2}; \ell+\frac{3}{2}\,\middle|\, t^2\right)\,. \tag{2.35}$$

For $t > 1$ they can be calculated using equation (2.20). From this equation it follows that $I_\ell(\nu,t)$ vanishes for $t \to 0$. For $t = 1$ the hypergeometric function can be expressed in terms of gamma functions. The typical behaviour of $I_\ell(\nu,t)$ is such that its absolute value peaks at $t = 1$. Furthermore, given that $I_\ell(\nu,t) \sim t^\ell$, for high values of ℓ the Bessel integrals have support only in a narrow interval close to $t = 1$. This means that in practice, in order to reach a given precision, we can find some $t_{\rm min}$ such that $|I_\ell(\nu,t_{\rm min})| < \epsilon |I_\ell(\nu,1)|$, where $\epsilon \ll 1$, and approximate the integrals (2.23), (2.24) as ranging from $t_{\rm min}$ to 1. In this way $I_\ell(\nu,t)$ does not need to be evaluated for $t < t_{\rm min}$, and the integrals over t only need to be carried over a small range, which gets narrower with larger ℓ's.

For very large ℓ the result obtained using the Bessel integrals must agree with the Limber approximation [19–21],

$$j_\ell(u) j_\ell(ut) \approx \frac{2\pi^2}{\ell_0}\delta(\ell_0 - u)\delta(\ell_0 - tu)\,, \qquad \text{with} \qquad \ell_0 = \ell + 1/2\,. \tag{2.36}$$

This would imply

$$I_\ell(\nu,t) \approx 2\pi^2\, (\ell_0)^{\nu-3}\, \delta(1-t)\,, \tag{2.37}$$

JCAP10(2018)047

which is indeed consistent with the result of the explicit computation of $I_\ell(\nu,t)$ in this limit, using the fact that $(1-t_{\rm min})$ is proportional to ℓ^{-1} and $I_\ell(\nu,1)$ proportional to $\ell^{\nu-2}$ (see appendix C). However, our method is such that the Limber approximation never has to be used. Once $f_n^{ij}(t)$ has been evaluated in the range $0<t<1$ for the purpose of computing the low multipoles with the full integral of (2.23), (2.24), the computational cost of the integrals for higher multipoles is minor, because the support of integration $[t_{\rm min},1]$ gets smaller. Thus there would be not much gain in using the Limber approximation and in our `CLASS` implementation we currently choose not to use it for any ℓ.

The direct evaluation of the Bessel integrals on the whole interval $0<t<1$ using equation (2.35) or similar other representations (see appendix C) can be relatively slow and inaccurate. The reason is that the power series of the Gauss hypergeometric function ${}_2F_1$ does not converge well in finite precision arithmetics due to large cancellations and/or many terms in the series have to be kept. This problem is particularly relevant for large imaginary parameters. Therefore, it is important to avoid direct evaluation whenever possible. Luckily, there are several alternative methods to efficiently calculate the Bessel integrals which we list below. On the basis of several speed and accuracy tests (performed by comparing with reference results provided by the python library `mpmath`), we have established and implemented in our code a list of threshold values t_i in the range $[0,1]$ at which it is advantageous to switch from one method to another. These thresholds depend both on ν and $\ell_{\rm max}$.

Taylor method. For $t\to 0$, we can use an analytic Taylor expansion of the hypergeometric function in the vicinity of zero. A similar approach can be taken for $t\to 1$, because the hypergeometric function can first be transformed with a change of variable $t^2\to\left((1-t^2)/(1+t^2)\right)^2$, and then expanded in small $\epsilon=1-t$. This transformation only works for tiny ϵ values, otherwise the result depends on precise cancellations between two hypergeometric functions. The different transformations for the $I_\ell(\nu,t)$'s are listed in appendix C. The Taylor method is not as fast as the recursion methods discussed next, but it is more robust in the close vicinity of the edges. We use it in two small ranges $0<t\le t_1$ and $t_2<t<1$.

Recursion methods. The Bessel integral $I_\ell(\nu,t)$ obeys the recursion relation [14]

$$\left(3+\ell-\frac{\nu}{2}\right)I_{\ell+2}(\nu,t)=\frac{1+t^2}{t}\left(\ell+\frac{3}{2}\right)I_{\ell+1}(\nu,t)-\left(\ell+\frac{\nu}{2}\right)I_\ell(\nu,t)\,, \tag{2.38}$$

where the $I_0(\nu,t)$ and $I_1(\nu,t)$ are known analytically:

$$I_0(\nu,t)=2\pi\cos\left(\frac{\pi\nu}{2}\right)\Gamma(\nu-2)t^{-1}\left[(1+t)^{2-\nu}-(1-t)^{2-\nu}\right]\,, \tag{2.39}$$

$$I_1(\nu,t)=2\pi\frac{\cos\left(\frac{\pi\nu}{2}\right)\Gamma(\nu-2)}{(4-\nu)t^2}\left[(1+t)^{2-\nu}((1+t)^2+\nu t)-(1-t)^{2-\nu}((1+t)^2-\nu t)\right]\,.$$

We can follow the recursion relation in forward direction $\ell\to\ell+1$, starting from $\ell=0$ and $\ell=1$. Alternatively the recursion relation can be used in the backward direction $\ell\to\ell-1$, starting from some $\ell_{\rm seed}$ and $\ell_{\rm seed}-1$. The forward and backward recursions are always the fastest methods but they are not always stable.

The method is only stable until a maximum value of ℓ which increases with t. We employ it only above a threshold value t_3 such that stability extends up to at least to the highest multipole value $\ell_{\rm max}$ needed for the angular spectrum computation.

JCAP10(2018)047

For $t < t_3$, we try to use the backward recursion whenever possible, starting from some $I_{\ell,\mathrm{seed}}(\nu,t)$ and $I_{\ell,\mathrm{seed}-1}(\nu,t)$ given by the direct calculation (C.6). We mentioned that this calculation can be slow and inaccurate, but high accuracy is not needed in this context. Indeed, the error made on the initial terms is reduced at each iteration. Therefore, instead of starting from the highest multipole value ℓ_{max} needed for the angular spectrum computation, we start from a value offset by some amount $\delta\ell$; $\ell_{\mathrm{seed}} = \ell_{\mathrm{max}} + \Delta\ell$, based on Miller's Recurrence Algorithm. The error on $I_{\ell,\mathrm{max}}(\nu,t)$ is then reduced by a factor $t^{\Delta\ell}$. The optimal amount of offset $\Delta\ell$ depends strongly on the precision with which the initial values have been computed, but also on t and ν (for instance, we found that it should scale as $(1-t)^{-1}$). We implemented in the code an ansatz for $\Delta\ell(\nu,t)$, keeping in mind that it is usually much faster to take a larger offset than to calculate the initial seeds with higher precision. Furthermore, we can use a trick similar to the one presented in [15]: once we reach the analytically known $\ell = 0$, we can compare the result obtained by backward recursion with the analytic result, calculate the complex ratio

$$\lambda(\nu,t) = \frac{I_0^{(\mathrm{analytical})}(\nu,t)}{I_0^{(\mathrm{recursion})}(\nu,t)}\,, \tag{2.40}$$

and multiply a posteriori all $I_\ell(\nu,t)$ with this ratio.

The backward recursion is more stable for smaller t values (but not very close to zero, which is why we use the Taylor method below t_1). We find that for some values of ν, the backward recursion gives good results in the whole range from Taylor to Forward recursion, $t_1 < t < t_3$, allowing us to switch directly from the backward to the forward method at t_3. For some other values of ν, the backward recursion is only accurate up to another threshold value t_2, and the last remaining range from Backward to Forward recursion, $t_2 < t < t_3$, needs to be covered by a last method.

Helper function method. Finally, in regions $t_2 < t < t_3$ where neither the recursion relations nor the Taylor expansions give a satisfyingly small error, we switch to the method introduced in appendix E of [15], which introduces what we call a "helper function" for the purpose of computing the hypergeometric function. We have adapted their approach for our purposes. For instance, we do not consider different $\ell \neq \ell'$ in the Bessel functions, we use a different method for setting the seeds, and we use a different forward-backward recursion switching criteria. This method is not as fast as the recursion methods, but it appears to be almost universally convergent.

For the Taylor and "helper function" methods, it is important to exploit the recursion relations for the Gamma functions $\Gamma(x+1) = x\Gamma(x)$ as much as possible. Otherwise, the evaluation of a few $\Gamma(x)$ from scratch in each recursion step would be prohibitively slow.

3 Different source terms

Reference [14] already addressed the fact that the transfer function of the number count spectrum involves multiple terms, and focused in particular on Redshift-Space Distortions. Here we want to go into the full details of all the terms and problems involved. The derivation of all the necessary source terms for the galaxy number count is given in [3] and their concrete expressions in `CLASS` are summarized in [6] (on page 19). We also give a summary of these terms (with additional information relevant to the present work) in appendix A, tables 6 and 7.

JCAP10(2018)047

We write the total transfer function of number count in the form

$$T(k,\chi) = \sum_x T_x(k,\chi)\,, \tag{3.1}$$

where x labels the different contributions: Doppler terms and Redshift-Space Distortions, lensing terms, and other gravitational terms (accounting for small GR corrections).[5] Consequently we have to introduce a double sum in the definition of the function $f_n^{ij}(t)$,

$$f_n^{ij}(t) = \int_0^\infty d\chi\, W^i(\chi) W^j(\chi t) \sum_{x,y} c_n^{xy}(\chi,\chi t)\,, \tag{3.2}$$

with $c_n^{xy}(\chi,\chi t)$ defined analogously to equation (2.12), using

$$\mathcal{P}_\mathcal{R}(k) T_x(k,\chi_1) T_y(k,\chi_2) = \sum_n c_n^{xy}(\chi_1,\chi_2)\, k^{\nu_n}\,. \tag{3.3}$$

Since the C_ℓ's depend linearly on the Fourier coefficients c_n^{xy}, they can also be decomposed as $C_\ell^{ij} = \sum_{x,y} C_\ell^{ij,xy}$.

While most terms can be cast in the general form of equations (3.2) and (2.23), (2.24) easily, there are two kinds of source terms for which this is possible, but not straightforward. These are terms for which the transfer function $\Delta_\ell^i(k)$ would involve not a spherical Bessel function, but its first or second derivative (Doppler and Redshift-Space Distortion terms); and those for which it contains one additional integral over time or comoving radius (weak lensing and some other GR corrections). In what follows we give details of how to deal with each of these complications separately, since they involve different challenges and solutions.

Any factors of ℓ appearing in the source terms can immediately and trivially be factorized out from each $C_\ell^{ij,xy}$, in such a way that the cosmological functions $f_n^{ij,xy}(t)$ remain independent of the multipole index ℓ. These factors only need to be taken into account later, when doing the final summation over the (x,y) indices.

3.1 Derivatives of Bessel functions

The Doppler terms and Redshift-Space Distortions include derivatives of the Bessel functions. They give rise to contributions of the form

$$C_\ell^{ij,xy} = 4\pi \sum_n \int_0^\infty d\chi_1 d\chi_2\, W^i(\chi_1) W^j(\chi_2)\, c_n^{xy}(\chi_1,\chi_2) \int_0^\infty \frac{dk}{k}\, k^{\nu_n}\, j_\ell^{(n_1)}(k\chi_1) j_\ell^{(n_2)}(k\chi_2)\,, \tag{3.4}$$

where $n_1, n_2 = 0, 1, 2$ and $j_\ell^{(0)}$, $j_\ell^{(1)}$, $j_\ell^{(2)}$ denote respectively the zero-, first- and second-order derivatives of the spherical Bessel function. We will make use of integrations by parts to obtain an expression without derivatives of the Bessel functions.[6] The boundary terms in

[5] Note that we have used labels α and β on the transfer function in the previous sections to indicate different kind of observables in the general cross-spectra. The latin indices x and y here indicate different contributions to a single observable. For clarity we will suppress Greek indices in this section, but it is always assumed that our formulas apply to a generic cross spectrum.

[6] Note that, in principle, we could avoid any integration by parts and instead use the relation between derivatives of Bessel functions and linear combinations of $j_\ell(x)$ and $j_{\ell+1}(x)$. This would require the introduction of additional analytical integrals, which do not obey the same transformation properties that make $I_\ell(\nu,t)$ relatively easy to calculate (see appendix C for a more thorough discussion). Thus, the integration by parts method turns out to be a simpler approach.

JCAP10(2018)047

the integration by parts vanish since

$$\lim_{x\to 0} j_\ell(x) = \lim_{x\to\infty} j_\ell(x) = 0 \qquad \text{for} \quad l \geq 1\,. \tag{3.5}$$

In the traditional line-of-sight approach, one could write that

$$\Delta_\ell^{i,x}(k) = \int_0^\infty d\chi W^i(\chi) T_x(k,\chi) \frac{\partial j_\ell(k\chi)}{\partial k\chi} = -\int_0^\infty d\chi \frac{\partial \left(W^i(\chi) T_x(k,\chi)\right)}{k\,\partial\chi} j_\ell(k\chi)\,. \tag{3.6}$$

We see that each derivative of a Bessel function both gives us a power of k and a time derivative of the product of the window and transfer function. We can use exactly the same approach in the FFT formalism and, for instance, in the case $(n_1, n_2) = (1, 0)$ write

$$C_\ell^{ij,xy} = 4\pi \sum_n \int_0^\infty d\chi_1 d\chi_2 W^j(\chi_2) \frac{\partial \left(W^i(\chi_1) c_n^{xy}(\chi_1,\chi_2)\right)}{\partial\chi_1} \int_0^\infty \frac{dk}{k}\, k^{\nu_n - 1}\, j_\ell(k\chi_1) j_\ell(k\chi_2)\,. \tag{3.7}$$

While taking derivatives of the Fourier coefficients $c_n^{xy}(\chi_1,\chi_2)$ is possible, it would be very slow in practice, because these derivatives should be evaluated in each point of the discrete (χ_1,χ_2) grid. This problem is amplified by the fact that we also need all other derivatives of the form $(\partial_{\chi_1})^{n_1}(\partial_{\chi_2})^{n_2} c_n^{xy}(\chi_1,\chi_2)$. These issues could be avoided if the coefficients c_n^{xy} were separable in χ_1 and χ_2, and of the form

$$c_n^{xy}(\chi_1,\chi_2) = D^x(\chi_1) D^y(\chi_2) c_n\,, \tag{3.8}$$

because then the derivatives could be calculated independently. This "separability" ansatz was implicitly suggested in [14, 15]. It relies on the separability of the transfer function into time and wavenumber dependence, $T_x(k,\chi) = D_x(\chi) T_x(k,0)$. This assumption is only fulfilled in a ΛCDM universe without massive neutrinos, in the sub-Hubble limit and within linear theory. However, the deviation from the "separable" limit remains small even in the presence of massive neutrinos, non-linear corrections or other physical ingredients leading to a scale-dependent growth factor. We can take advantage of this and define rescaled Fourier coefficients

$$\tilde{c}_n^{xy}(\chi_1,\chi_2) = \frac{c_n^{xy}(\chi_1,\chi_2)}{D_x(\chi_1) D_y(\chi_2)}\,, \tag{3.9}$$

which are only weakly dependent on conformal time. We obtain the effective growth factor $D_x(\chi)$ by extracting it from the exact transfer function $T_x(k,\chi)$. To focus on the scales of interest, we weigh the contributions from different positions in the k-grid differently using weights w_m,

$$D_x(\chi) = \frac{\sum_m w_m T_x(k_m,\chi)/T_x(k_m,0)}{\sum_m w_m}\,. \tag{3.10}$$

For the weights w_m we choose a Gaussian in $\log(k)$, centered around some scale of interest k_0 ($= 1h/\text{Mpc}$ in our implementation), with a standard deviation of half-a-decade. The precise form of the weights and the scale k_0 do not influence our results beyond a relative deviation of 10^{-6}.

The rescaled Fourier coefficients of equation (3.9) allow us to work in a "semi-separable" approximation in which we assume that the dependence of the factors $\tilde{c}_n^{xy}(\chi_1,\chi_2)$ on conformal time is much weaker than that of the growth factors $D^x(\chi)$. In this limit the derivative

JCAP10(2018)047

is now *approximated* by

$$\begin{aligned}\frac{\partial}{\partial\chi_1}c_n^{xy}(\chi_1,\chi_2) &= \frac{\partial}{\partial\chi_1}\Big(D_x(\chi_1)D_y(\chi_2)\tilde{c}_n^{xy}(\chi_1,\chi_2)\Big)\\ &\approx D_y(\chi_2)\frac{\partial D_x(\chi_1)}{d\chi_1}\tilde{c}_n^{xy}(\chi_1,\chi_2)\,,\end{aligned} \tag{3.11}$$

effectively giving us the same factorization as when assuming full separability. This allows us to write the terms with derivatives of the spherical Bessel functions in a simple form. For instance, in the example in equation (3.7), the expression for the angular power spectrum can be written in the form

$$\begin{aligned}C_\ell^{ij,xy} = \sum_n \int_0^\infty d\chi_1 d\chi_2 \frac{\partial\left(W^i(\chi_1)\,D_x(\chi_1)\right)}{\partial\chi_1} W^j(\chi_2)D_y(\chi_2)\\ \times\ \tilde{c}_n^{xy}(\chi_1,\chi_2)\,\chi_1^{-(\nu_n-1)} I_\ell\left(\nu_n-1,\frac{\chi_2}{\chi_1}\right)\,.\end{aligned} \tag{3.12}$$

This can be evaluated very efficiently, since the derivatives of the products $W^i D_x$ can be pre-calculated and later interpolated at some χ and χt (and additionally χ/t for equation (2.24)). This approach applies to any combination of derivatives. The precision tests presented in section 4 prove that the semi-separable approximation of equation (3.11) does not introduce any sizable inaccuracy, even when non-linear corrections are implemented (with fitting formulas like Halofit [24, 25]), or when massive neutrinos are introduced (see section 4.2).

3.2 Integrals over Bessel functions

Some source terms accounting for weak lensing and other GR corrections involve an additional integral over conformal time, such that in the traditional line-of-sight approach their harmonic transfer function reads

$$\Delta_\ell^{i,x}(k) = \int_0^{\tau_0} d\chi'\, W^i(\chi') \int_0^{\chi'} d\chi\, f(\chi,\chi')\, T_x(k,\chi)\, j_\ell(k\chi)\,, \tag{3.13}$$

where τ_0 is the conformal age of the universe, and $f(\chi,\chi')$ is a convolution kernel. In the case of lensing, χ' would be the distance to the source and χ the distance to the lens. To obtain the same functional form as other harmonic transfer functions we can swap the order of integration [6],

$$\Delta_\ell^{i,x}(k) = \int_0^{\tau_0} d\chi \int_\chi^{\tau_0} d\chi'\, W^i(\chi')\, f(\chi,\chi') T_x(k,\chi)\, j_\ell(k\chi)\,, \tag{3.14}$$

leading to the definition of an new window function

$$\tilde{W}(\chi) = \int_\chi^{\tau_0} d\chi'\, W^i(\chi')\, f(\chi,\chi')\,. \tag{3.15}$$

Once again, the same approach can be readily transposed to the FFT method. For instance, the cross-correlation between one integrated and one non-integrated term reads

$$C_\ell^{ij,xy} = \sum_n \int_0^{\tau_0} d\chi_1\, \tilde{W}^i(\chi_1) \int_0^{\tau_0} d\chi_2\, W^j(\chi_2)\, c_n^{xy}(\chi_1,\chi_2)\, \chi_1^{-\nu_n}\, I_\ell\left(\nu_n,\frac{\chi_2}{\chi_1}\right)\,, \tag{3.16}$$

with $\tilde{W}^i(\chi)$ defined like in equation (3.15).

JCAP10(2018)047

The important difference compared to non-integrated terms comes from the broad support of the integrated window function $\tilde{W}^i(\chi)$. If we assume $W^i(\chi)$ is non-zero in the range $[\chi^i_{\rm min}, \chi^i_{\rm max}]$, then $\tilde{W}^i$ has support in the whole $[0, \chi^i_{\rm max}]$ range. Thus $\tilde{W}^i(\chi)$ has to be sampled with more points in χ, an effect that is worsened by the oscillatory nature of the $\chi^{1-\nu_n}$ factor appearing in the definition of $f_n^{ij}(t)$ in equation (2.18).

For contributions involving one integrated and one non-integrated term, the definition of χ becomes important. It is better to define χ as the conformal distance of the non-integrated term. In that case χ still has a restricted support, allowing for evaluation with a sparser grid. The range of t also has to be adjusted; the new range is $[0, 1]$, as can be seen from equations (2.25) and (2.26).

For contributions involving two integrated terms, χ has a broad support in any case. Not only are more sampling points required, but since small χ values are allowed, the oscillatory nature of the $\chi^{1-\nu_n}$ factor in the definition of f_n^{ij} in equation (2.18) could become problematic. We decide to integrate instead over $\log(\chi)$, since this is the characteristic oscillation length of the factor $\chi^{1-\nu_n} = \cos(\log(\chi)(1-\nu_n)) + i\sin(\log(\chi)(1-\nu_n))$.

We also implement an algorithm estimating a value $\chi_{\rm cut}$ below which $\tilde{W}^i(\chi)$ is sufficiently small and contributions to the integral over χ are negligible. Even if $\chi_{\rm cut}$ is small (e.g. $\chi_{\rm cut} \sim 1{\rm Mpc}/h \;\Rightarrow\; z_{\rm cut} \sim 2\cdot 10^{-4}$), the lower integration bound in $\log(\chi)$ space is still drastically cut.

3.3 Behaviour in k

The transfer functions $T_x(k,\chi)$ of the various contributions to number count and lensing power spectra involve the gauge-invariant perturbations D, Θ, Φ and Ψ defined in [6] and summarized in appendix A, table 7. These account respectively for density, velocity divergence and metric fluctuations. They appear in the transfer functions in combination with various powers of k. However, after performing the transformation of section 3.1 in order to eliminate derivatives of Bessel functions, one is left only with contributions from D, Θ/k^2, Φ, Ψ, Φ' and Ψ'.

The gauge-invariant variables D, Θ, Φ and Ψ behave roughly like the perturbations $\delta_{\rm m}$, $\theta_{\rm m}$, ϕ and ψ of the Newtonian gauge (where the index m stands for non-relativistic matter). It is easy to show with analytical arguments or by looking at the output of Boltzmann codes that the following quantities share roughly the same k-dependence, at least at low redshift: $D/k^2 \propto \Theta/k^2 \propto \phi \propto \psi \propto \Phi' \propto \Psi'$. This implies that all the transfer functions involved in the problem scale in roughly the same way except for the density one: $T_D(k,\chi)/k^2 \propto T_{\Theta/k^2}(k,\chi) \propto T_\Phi(k,\chi) \propto \cdots$.

In section 2.3 we have explained how we deal with this situation for non-density terms: instead of performing the FFTlog transformation of $T_x(k,\chi)$, we apply it to $\bar{T}_x(k,\chi) = k^2 T_x(k,\chi)$. This is done at the expense of changing the real part of the complex frequencies ν for which the integrals $I_\ell(\nu,t)$ must be evaluated. While we need to compute $I_\ell(\nu_n,t)$ always for all n's, for non-density types we need additional computations with $\Re[\nu_n]$ equal to b, $b-2$ or $b-4$, with our chosen tilt $b \approx 1.9$.

Another approach would be to use a trick from [14], and to transform the density source term $T_D(k,\chi)$ using

$$k^2 j_\ell(k\chi) = -\left[\frac{\partial^2}{\partial\chi^2} + \frac{2}{\chi}\frac{\partial}{\partial\chi} - \frac{\ell(\ell+1)}{\chi^2}\right] j_\ell(k\chi)\,. \tag{3.17}$$

JCAP10(2018)047

In that way, one can bring all the terms to the same behaviour in k, and compute all $I_\ell(\nu_n, t)$ with a unique value of $\Re[\nu_n]$ equal to $b - 4$. Since the calculation of the $I_\ell(\nu, t)$ can be done only once and stored for later use, we did not adopt this trick here. In the last two lines of table 7, we also included the new source terms that need to be computed if one uses equation (3.17) for the density. We implemented this method in `CLASS` for the sake of comparison, and the user can decide to switch to it if desired.

4 Accuracy and performance

4.1 Accuracy

We first check whether the default `CLASS` implementation and the new method converge to the same results in the limit of high accuracy settings. We increased the most relevant precision parameters up to the point of saturating the memory limits of a 16-dual-core workstation with 32GB of RAM. In figure 1, we can see that the agreement on the C_ℓ's is in the 0.1% range for the number count spectra (when all terms and GR corrections are taken into account). The same applies to the cosmic shear C_ℓ's, shown in figure 2. When only the density source terms are involved, the level of agreement further improves to about 0.01% (figure 3). See table 9 in appendix D for the cosmological parameters used for the comparisons. These results validate the accuracy of the new method, since an error level of 0.1% on the full number count or cosmic shear spectra is sufficient for fitting the experimental results of future surveys. In order to test the "semi-separable" approximation implemented in our code, we must repeat this exercise in presence of nonlinear corrections and/or massive neutrinos, which both introduce a scale dependence in the density fluctuation growth factor. To this end, we switch on Halofit corrections [24–26] and/or degenerate massive neutrinos with a total mass $M_\nu = 1\,\mathrm{eV}$. This mass choice is rather extreme given that current cosmological upper bounds are in the ballpark of $M_\nu \sim 0.1\,\mathrm{eV}$ to $0.3\,\mathrm{eV}$, and as such, any lower mass should be captured even better. The effect of these ingredients at redshift $\bar{z} = 1$ is visible in the upper panels of figure 4.

We first check that the "full separability" approximation would introduce large errors. The "full separability" approximation is implemented in the code as follows: instead of performing the FFTlog transformation at each time step, we only do it once at redshift zero to derive the c_n's of equation (3.8). We then rescale them at each time using the effective growth factor defined in section 3.1. Unsurprisingly, we obtain in this case a large error in the number count spectrum, of the order of 50% (bottom left panel in figure 4).

However, our "semi-separability" method brings the result back in very good agreement with the traditional line-of-sight result. The residual difference between the old and new method is better seen in the bottom right panel of figure 4, in the case of number count with only density terms. Adding non-linear corrections leaves the residual at a level below 0.01%, while switching on a total neutrino mass $M_\nu = 1\,\mathrm{eV}$ raises it to around 0.04%. When considering number count spectra with all GR terms, we find that adding non-linear corrections and/or massive neutrinos keeps the difference at the level of 0.1%, which is sufficient for future surveys.

Thus the new method — including all the choices and approximations presented in the previous sections — is sufficiently accurate. Still, it would be very interesting to investigate what dominates the residual $\sim 0.1\%$ differences found for all the spectra including a lensing effect. If there is a clear reason to believe that most of it comes from errors in the traditional line-of-sight approach (as implemented in the default `CLASS` code), then the new method

JCAP10(2018)047

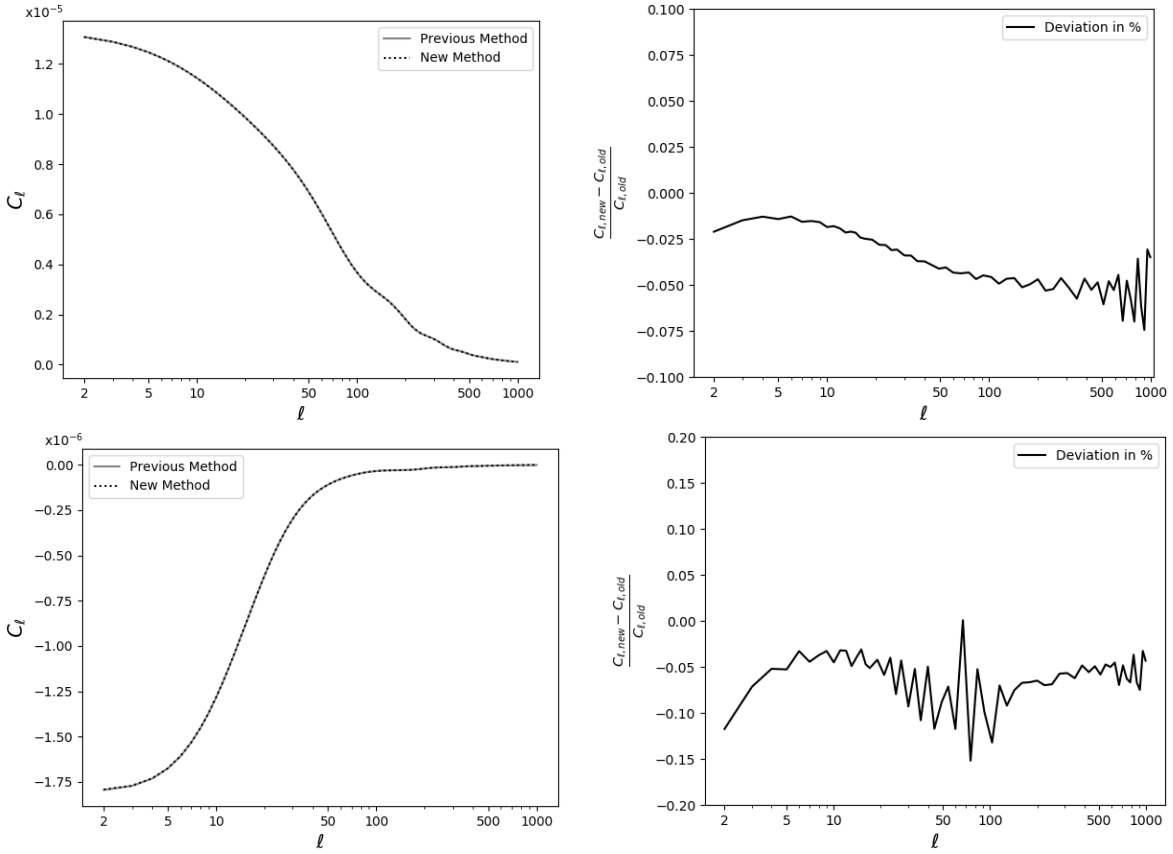

Figure 1. *(Top)* Auto-correlation spectrum of number count (involving all source contributions) in one redshift bin defined by a Gaussian window function with mean redshift $\bar{z} = 1.0$ and width $\delta z = 0.05$. *(Bottom)* Cross-correlation between two redshift bins defined by two Gaussian windows with $(\bar{z}_1, \delta z_1) = (1.0, 0.05)$ and $(\bar{z}_2, \delta z_2) = (1.25, 0.05)$.

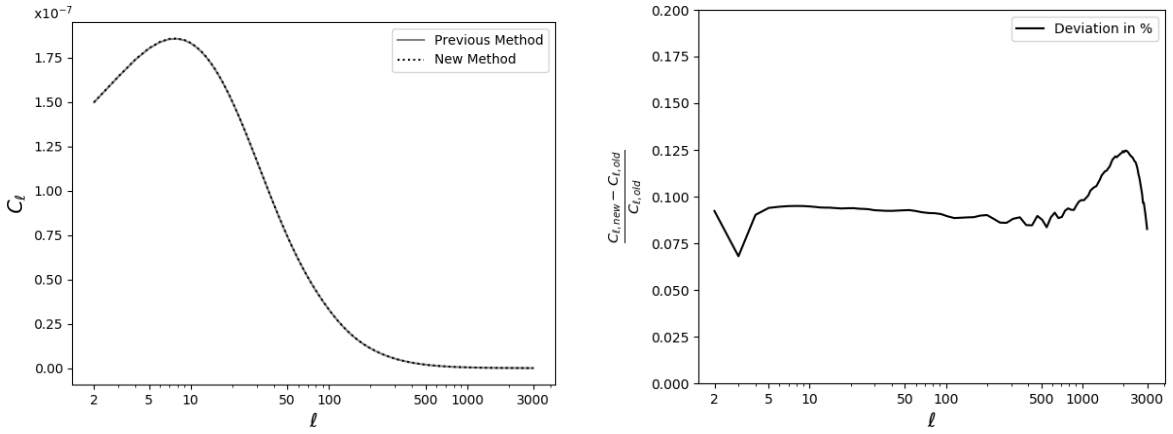

Figure 2. Auto-correlation spectrum of cosmic shear (or more precisely of the lensing potential $C_\ell^{\phi\phi}$) in one redshift bin defined by a Gaussian window function with mean redshift $\bar{z} = 1.0$ and width $\delta z = 0.05$.

could actually be significantly more precise than 0.1%. Indeed, we believe that this the case, because when we increase the precision settings, the lensing spectrum obtained with the old

JCAP10(2018)047

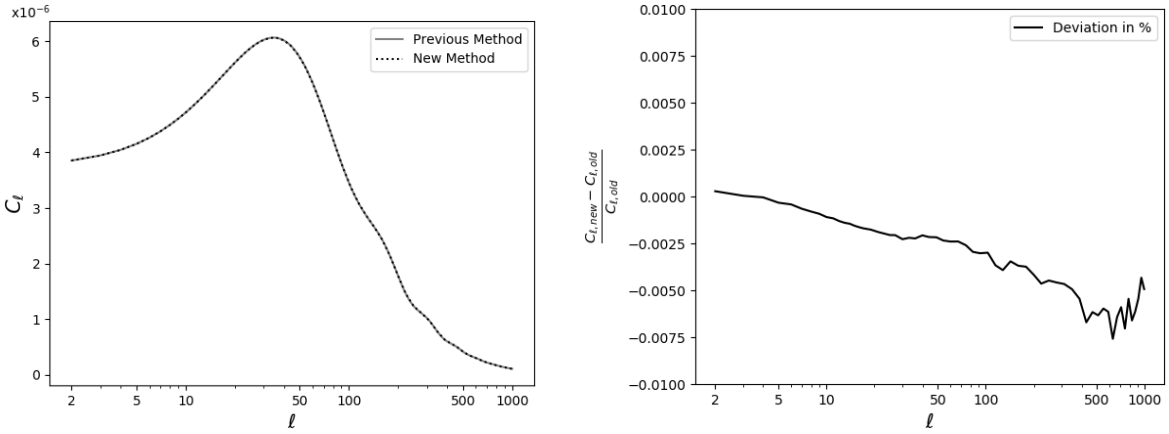

Figure 3. Auto-correlation spectrum of number count (involving only the density source term) in one redshift bin defined by a Gaussian window function with mean redshift $\bar{z} = 1.0$ and width $\delta z = 0.05$.

CLASS code converges very slowly towards a stable result. It is not even fully converged at the sub-percent level when the memory limit is reached. This is not the case with the new method, which already gives stable results for nearly the same precision settings in the case of density and lensing contributions, while using a small amount of memory. The calculation of the density contribution with the old code is also much more stable than that of lensing. Thus it seems that the true level of general agreement between the old and new method might be 0.01% rather than 0.1% due to the fact that the old method has a specific problem to compute lensing spectra with the very high level accuracy.

This problematic feature of the default CLASS implementation can be easily explained. It is related to the fact that in the lensing case, the support of the window function $\tilde{W}^i(\chi)$ defined in equation (3.15) reaches $\chi_{\rm min} = 0$. Indeed, in principle, the lensing of a source standing at some finite redshift has contributions from modes located arbitrarily close to us, for which a given angle corresponds to infinitely small wavelengths and thus infinitely large wavenumbers. This means that in the traditional line-of-sight approach, the harmonic transfer functions $\Delta_\ell^{\alpha,i}(k)$ of lensing converge very slowly as a function of k. Mathematically this is seen by noticing that these transfer functions are obtained by integrating the product $\tilde{W}^i(\chi)T_\alpha(k,\chi)j_\ell(k\chi)$ over χ. For arbitrarily large k, there is always a non-vanishing contribution to the integral from a distance of $\chi \sim \ell/k$. The convergence of $\Delta_\ell^{\alpha,i}(k)$ is only ensured by the fact that the lensing transfer function $T^{\phi+\psi}(k,\chi)$ scales like k^{-2} in the large k limit. This is very different from non-integrated cases, for which convergence arises much earlier simply because the lower edge $\chi^i_{\rm min}$ of the support of $W^i(\chi)$ forces $\Delta_\ell^{\alpha,i}(k)$ to vanish above $k \sim \ell/\chi^i_{\rm min}$. This slow convergence of $\Delta_\ell^{\alpha,i}(k)$ for lensing term is indeed problematic because the period of oscillation of $j_\ell(k\chi)$ as a function of χ is given by $2\pi/k$. Thus, in order to get precise results, one needs to increase a lot the number of sampled values in the integral over the line-of-sight. This leads to a saturation of the memory before obtaining a sampling that would ensure the convergence of the C_ℓ's at the 0.1% level.

The new method avoids such problems because the integral over slowly-damped oscillating functions at large k is done analytically. The integral over large k is performed within the FFTlog transformation, which does not involve any Bessel function and converges without problems. There are no issues of divergences for small χ either, since the product $\tilde{W}(z)\chi^{1-\nu_n}$ always approaches zero in the limit $\chi \to 0$ for valid choices of $\Re[\nu_n]$ for which the Bessel

JCAP10(2018)047

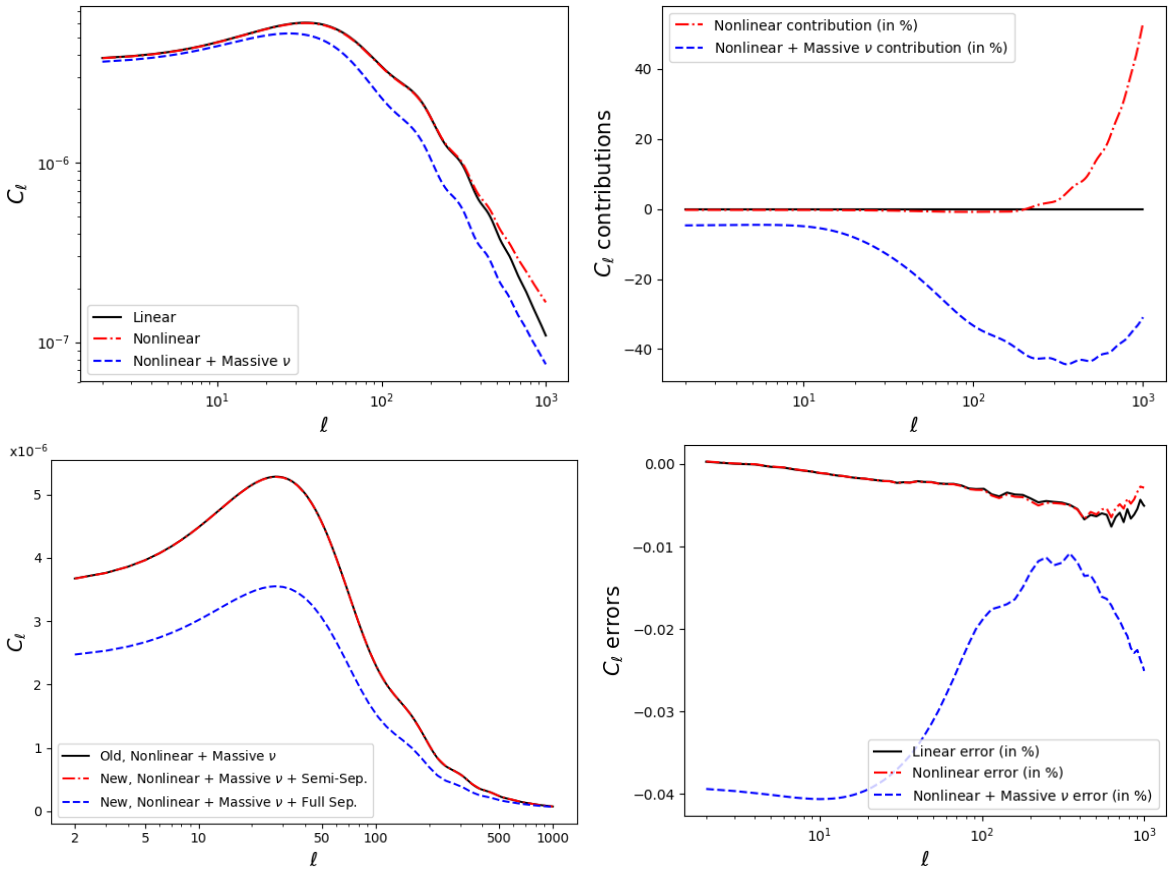

Figure 4. Number count spectra involving only density terms for a redshift bin centered at $\bar{z} = 1.0$ with width $\delta z = 0.05$. *(Top Left)* Total spectra w/o nonlinear corrections from Halofit and massive neutrinos with $M_\nu = 1\,\mathrm{eV}$. *(Top Right)* Impact of these two corrections on the power spectrum, computed as a relative difference (in %) with respect to the linear spectrum of the massless neutrino model. *(Bottom left)* Result of the new method with either the "full separability" or "semi-separability" approximations compared to the traditional line-of-sight approach. *(Bottom right)* Relative difference (in %) between the new and old methods. One can immediately see that the additional effects are well captured and the error remains at the sub-permille level.

integral $I_\ell(\nu_n, t)$ is well behaved (see section 2.3). Thus, the lensing spectra can be computed as accurately as the density spectra without requiring significantly more memory. Besides any considerations based on performance, this is a true advantage of the FFTlog method.

4.2 Performance

Precision measure. We will now estimate the CPU time needed to reach a precision of $Q = 0.1\%$, where we define the precision measure as

$$Q = \sqrt{\frac{1}{N_\ell} \sum_\ell \left(\frac{C_\ell - C_\ell^{\mathrm{ref}}}{C_\ell^{\mathrm{ref}}} \right)^2}, \tag{4.1}$$

with N_ℓ discrete samples in ℓ.

To give a more fair comparison, the reference spectra are calculated for each method separately. This means that we disregard in these tests the 0.1% level residuals between the

JCAP10(2018)047

reference spectra of the two methods. Indeed, as explained in the previous section, we have strong hints that these residuals come from issues of insufficient sampling in the traditional method, but in absence of a definite proof, we do not want to bias our conclusions by this assumption. Thus our analysis quantifies the numerical error coming from the degradation of the precision settings in each method, relative to the reference spectra of the same method.

ℓ sampling. The default `CLASS` implementation does not compute every value of ℓ separately, but interpolates with a cubic spline interpolation algorithm between a reasonably chosen set of discrete ℓ values at which the C_ℓ's are actually calculated. We adopt the same strategy in the new algorithm. The list of ℓ values at which the C_ℓ's are calculated by default (and in our tests) can be found in appendix D, tables 10 and 11. Of course, the precision measure Q takes only these ℓ values into account.

Approximations. All our performance calculations use the "full separability" approximation, but we avoid using the Limber approximation, since we are interested also in effects for small ℓ for which the Limber approximation deviates most from the correct result. We looked at three important test cases: the calculation of cosmic shear spectra, of number count spectra involving only density contributions, and of number count spectra involving all contributions, except for the small gravitational corrections labeled as G1–G5 here and in [6]. These are known to be time-consuming but mostly negligible in the final results.

Setup. The computation time for the geometrical integrals $I_\ell(\nu, t)$ is explicitly taken out from the direct comparison, since these could be stored as small binary files on the computer, and thus would not need to be calculated every time. The execution times correspond to a single core and thread on an Intel i5-6200U Quadcore (2.3 GHz).

Precision parameters. In each test, all relevant precision parameters in the previous `CLASS` method and the new method are tuned to achieve maximum speed while remaining at the desired accuracy of $Q \simeq 0.1$ The tuned parameters of the old method include `hyper_sampling_flat`, guiding the sampling of the spherical Bessel functions, `selection_sampling_bessel` giving the χ sampling, and `q_linstep` and `q_logstep_spline`, both guiding the k sampling. As described in section 2.2, the cosmological function $f_n^{ij}(t)$ is calculated in a rough grid and interpolated for the final integration. Thus, the tuned parameters of the new method include the number of FFT coefficients N_c, the coarse number of t values $N_{t,\mathrm{spline}}$, and the number of t values for the final integration N_t. In the new method, it is possible to use different χ samplings for integrated or normal effects, leading to the two additional parameters $N_{\chi,\mathrm{integrated}}$ and $N_{\chi,\mathrm{normal}}$. The values of the parameters used for the tests can be found in table 1.

Nomenclature. In the two methods, we call N_z the number of redshift bins of the survey, and N_{tot} the total number of redshift bin combinations. When considering contributions from all cross correlations between redshift bins, this number is $N_{\mathrm{tot}} = N_z(N_z+1)/2$ and grows quadratically. If we include only the first M closest neighbours, the number is instead given by $N_{\mathrm{tot}} = (M+1)N_z - M(M+1)/2$ and grows linearly. In our tests, for the cosmic shear spectra, we calculated all possible cross correlations between different redshift bins. For the number count spectra, we limit ourselves to the first two neighboring bins ($M = 2$), because other correlations are expected to be strongly noise-dominated in future galaxy redshift surveys.

JCAP10(2018)047

	Cosmic Shear	NC, density only	NC, all (no gr)
Ranges			
$\ell_{\max}$	3000	1000	1000
Number of ℓ	112	62	62
Non-Diagonals	All	2	2
δz	0.05	0.01	0.01
Old parameters			
`hyper_sampling_flat`	7.5	7	8
`selection_sampling_bessel`	1.5	0.9	1.7
`q_logstep_spline`	120	21	21
`q_linstep` (ignored)	100000	100000	100000
New parameters			
$N_{\chi,\text{integrated}}$	75	—	50
$N_{\chi,\text{normal}}$	—	15	25
$N_{t,\text{spline}}$	35	20	70
N_t	130	50	100
N_c	100	95	95

Table 1. Parameter values for the different timing tests. NC = Number count. The parameter `q_linstep` is relevant for small wavenumbers and thus mainly for CMB observables: in our case it generally did not affect the results.

		Old		New		
N_{tot}	N_z	T [s]	T/N_z [s]	T [s]	T/N_{tot} [ms]	$I_\ell(\nu,t)$ [s]
5050	100	850	8.5	20	4.0	2.2
2556	71	620	8.7	11	4.4	2.6
1326	51	450	8.8	6.1	4.6	2.5
351	26	240	9.2	1.6	4.6	2.6
105	14	130	9.4	0.50	4.8	2.7
10	4	42	10.5	0.12	12	2.1

Table 2. Cosmic Shear. A comparison of the computation times T for each method, revealing that the old method scales as N_z, while the new method scales as N_{tot}. The new method is up to 350 times faster. The last column gives the time needed to calculate the $I_\ell(\nu,t)$ integrals, which is of minor importance, given that this calculation can be done once and for all.

Speedup. The performance comparison between the two methods for different numbers of observable redshift bins is given in tables 2, 3, and 4. The new method is always faster in our tests. The speedup of the new method ranges in our tests from factors of around 2 (density only) to 400 (cosmic shear or lensed density). As mentioned in section 3.3, if multiple sources are calculated, multiple $\Re[\nu]$ are required for the $I_\ell(\nu,t)$, explaining the slightly longer times in the last column of table 4 compared those in tables 2 and 3 (8 s instead of 2.5 s). This would not be the case if we used the "tilt reduction" method of section 2.3, but this issue is

JCAP10(2018)047

		Old		New		
$N_{\rm tot}$	N_z	T [s]	T/N_z [s]	T [s]	$T/N_{\rm tot}$ [ms]	$I_\ell(\nu,t)$ [s]
297	100	13	130	5.5	19	2.4
210	71	9.3	130	3.8	18	2.5
150	51	6.7	130	2.8	19	2.5
75	26	3.4	130	1.2	16	2.5
39	14	2.0	140	0.60	15	2.5
9	4	0.60	150	0.24	27	2.1

Table 3. Number Count, density contributions only. A comparison of the computation times T for each method, revealing that the old method scales as N_z, while the new method scales as $N_{\rm tot}$. The new method is up to 2.5 times faster.

		Old		New		
$N_{\rm tot}$	N_z	T [s]	T/N_z [s]	T [s]	$T/N_{\rm tot}$ [ms]	$I_\ell(\nu,t)$ [s]
297	100	7500	75	27	90	7.8
210	71	5400	76	19	90	8.0
150	51	3800	74	14	93	7.8
75	26	2000	77	7.2	96	7.9
39	14	1100	79	3.2	82	7.9
9	4	300	75	0.80	89	7.9

Table 4. Number Count, all contributions except for gravitational terms. A comparison of the computation times T for each method, revealing that the old method scales as N_z, while the new method scales as $N_{\rm tot}$. The new method is up to 380 times faster.

anyway unimportant, give that the $I_\ell(\nu,t)$'s calculation can be done once and for all. The results can be stored in a binary file and quickly retrieved in later executions.

Scaling with N_z and $N_{\rm tot}$. The new method scales with $N_{\rm tot}$ instead of N_z (like the old one) due to the different arrangement of integrations. Thus, when we consider cosmic shear with all cross-correlation spectra, the new method scales quadratically with N_z, while the old one scales linearly. This could be a problem for the new method when many bins are considered, but fortunately the computation time per spectrum is much smaller in this method, of the order of $T/N_{\rm tot} \simeq 5\,$ms, instead of $T/N_z \simeq 9\,$s in the old method for the same accuracy. Thus the new method is about 350 faster for 4 bins, 42 times faster for one hundred bins, and the old method would only overtake for ~ 4000 redshift bins (8 million total combinations). One should keep in mind that cosmic shear analyses will always be optimal when analysed with a rather small number of redshift bins, in order to get good statistics in each bin: thus the most interesting cases are the ones with a dozen of bins at most, for which we find the best gains with the new method.

When considering instead cluster counts with only two non-diagonal terms ($M = 2$), both method scale linearly with N_z, and the gain is roughly independent of the number of bins. This gain is still very large when including integrated terms like lensing, since the new method gives accurate spectra without requiring integrals over slowly-converging oscillatory

JCAP10(2018)047

functions in k space. We find a speed up by a factor of the order of 380 in that case. When including only the leading density term, the new method is still faster, by about 2.5. This comparison is valid for $M = 2$: the new method performs even better when considering $M = 1$ or when calculating only auto-correlation spectra, while the old method would overtake for $M \geq 8$ when computing density terms only.

Scaling with precision parameters. We have also qualitatively evaluated how the new method scales with the different precision parameters. Most scaling relations are approximately linear, although with some minimal constant offset. This is likely due to other parts of the code becoming the dominant time consuming contributions in those cases. The parameters N_c, N_χ, N_t, $N_{t,\mathrm{spline}}$, and N_ℓ show this scaling. The calculation of the geometrical integrals $I_\ell(\nu, t)$ on the other hand is linearly proportional to N_c, ℓ_{max}, and the number of t values for which they are calculated. It also depends non-trivially on $Re[\nu]$.

Scaling with bin width. In the old method, the χ integration over a very narrow redshift bin around some χ^* results in a particular $j_\ell(k\chi^*)$ effectively being selected, thus the k sampling has to be rather precise to effectively capture the oscillatory nature of the spherical Bessel function. Schematically, we can describe this by approximating the narrow Gaussian window function as a Dirac-delta distribution,

$$\begin{aligned}\Delta^{\alpha,i}(k) &= \int_0^\infty d\chi\; T_\alpha(k,\chi)\; j_\ell(k\chi)\; W^i(\chi) \\ &\approx \int_0^\infty d\chi\; T_\alpha(k,\chi)\; j_\ell(k\chi)\; \delta(\chi - \chi^*) \\ &= T_\alpha(k,\chi^*)\; j_\ell(k\chi^*)\,.\end{aligned} \tag{4.2}$$

Thus the final integral in k oscillates like the Spherical Bessel function, meaning it has to be evaluated with a rather fine grid. A very broad redshift bin instead effectively results in averaging the $j_\ell(k\chi)$ over several oscillations, canceling out some of the oscillatory nature of the Bessel functions. Of course a broader redshift window requires more sampling points in χ instead. Thus there is a trade-off, and for number count the old method is quicker for intermediate bin widths ($\delta z \sim 0.01$) than for very wide or thin bins. For the new method, the Bessel functions are already integrated to obtain $I_\ell(\nu, t)$, always canceling out most of the oscillations. However, as the window width is shrunk, a smaller range of t values is allowed, leading to a speedup of the new method. Table 5 shows that the new method is always faster, but the gain is minimal for $\delta z \sim 0.01$, and grows on both sides of this value. Note that the previous performance tests of this section assumed with $\delta z \sim 0.01$, which is the most unfavorable case for the new method, with a factor two improvement for number count with density terms only. With $\delta z \sim 0.001$ the speed up factor reaches 40. Thus the new method will perform very well for spectroscopic surveys with typical bin widths $\delta z \sim 0.001(1 + z)$.

5 Conclusions

In this paper we have further developed and implemented in the `CLASS` code a method proposed in [14] (see also [15]) to efficiently evaluate the angular power spectra for large-scale structure observables. We particularly focus on the angular power spectrum of the number counts including all relevant terms such as redshift-space distortions, weak lensing, relativistic corrections etc., and the angular power spectrum of cosmic shear. The method is

JCAP10(2018)047

	Old					New (without $I_\ell(\nu,t)$)					
N_z	100	71	51	26	14	100	71	51	26	14	
Width											Speedup
0.5	62	48	33	17	9.0	7.8	5.5	4.0	2.0	1.0	8
0.1	42	33	22	11	6	8.4	6.0	4.4	2.2	1.1	5
0.05	37	31	20	10	5.6	7.4	5.2	3.8	1.86	0.96	5
0.01	13	9.3	6.7	3.4	2.0	5.5	3.8	2.8	1.2	0.6	2
0.005	60	48	31	16	8	3.6	2.5	1.8	0.86	0.41	20
0.001	80	45	33	17	9	1.9	1.3	0.93	0.45	0.22	40

Table 5. The scaling of the different methods with the redshift bin width for number count spectra. While the old method takes longer for either very broad or very thin redshift bins, the new method always increases in speed for thinner bins. The precision parameters for these runs can be found in the appendix, table 8. Our previous tests were conservatively performed in the case $\delta z = 0.01$ leading to minimal speedup.

based on the power-law decomposition of the k-dependent function integrated over spherical Bessel functions which allows for the analytical solutions of the momentum integrals. This power-law decomposition is achieved using the FFTlog algorithm. Compared to earlier studies we make one step further in separating the functions that depend on cosmology form cosmology-independent geometrical factors. We also discuss and test many important aspects of practical implementation of the algorithm. All these improvements lead to significant gain in speed and accuracy.

In order to test our code we evaluate the angular power spectra for number counts and cosmic shear for many different configurations of window functions and redshift bins and compare it to the default `CLASS` implementation. We find that the new method is always more efficient, with the speedup up to a factor of $\mathcal{O}(400)$. At the same time the accuracy of the results remains very good, with the relative error below 0.1%. There are reasons to believe that this error is due to numerical integration in the standard approach and that the accuracy of the new method can be easily an order of magnitude better. Given that the relative error of 0.1% is sufficient for all practical purposes for current and planned large-scale structure surveys, we leave the further investigation of this issue for future work. We plan to publicly release the CLASS modifications discussed in this work within a few months, after further polishing and minor improvements in the new parts of the code.

With such a speed up, it should become possible to perform Fisher matrix or MCMC forecasts of the sensitivity of future Large Scale Structure surveys to cosmological parameters, even with a very high number N_z of redshifts bins in the case of spectroscopic redshift surveys. In the future, we plan to perform such forecasts, and to study the feasibility of concretely using angular power spectra in the analysis of future galaxy redshift surveys.

Another clever way of speeding up Boltzmann codes has been proposed by the authors of [27]. This approach, called Angpow, sticks to the same sequence of operations as the traditional line-of-sight approach until the calculation of the transfer functions $\Delta_\ell^{\alpha,i}(k)$. However the latter are sampled in particular k values, allowing to compute the $C_\ell^{\alpha\beta,ij}$ efficiently with a Clenshaw-Curtis-Chebyshev algorithm. This allows to reduce the number of sampled k values by a significant amount. We note that this method is interesting per se, but cannot

JCAP10(2018)047

be combined with the current approach since we infer the $C_\ell^{\alpha\beta,ij}$ through a very different integral.

Throughout this paper we have focused on the power spectrum only and we have discussed only the large-scale structure observables. However, the method we presented is equally applicable to other cases such as for example the CMB primary anisotropies. It would be very interesting to see whether similar improvements are possible for the CMB observables. Another direction for future work is the application of our method to higher point correlation functions. In particular, as shown in [14], the case of the CMB or large-scale structure bispectra seems promising. Yet another possible application of our method is efficient evaluation of the covariance matrix for angular power spectra. All these examples are quite relevant and it would be worth exploring them in more details in the future.

Acknowledgments

M.S. gratefully acknowledges support from the Institute for Advanced Study and the Raymond and Beverly Sackler Foundation. M.Z. is supported by NSF grants AST-1409709 and PHY-1521097, the Canadian Institute for Advanced Research (CIFAR) program on Gravity and the Extreme Universe and the Simons Foundation Modern Inflationary Cosmology initiative.

A Types of source terms

We summarize in table 6 the different source functions contributing to the number count spectra.

The column "Window function" corresponds to the combination $D(\chi)W^i(\chi)$ or $D(\chi) \cdot \tilde{W}^i(\chi)$ of the main text (e.g. equation (3.12)), and the column "Source" to what is called $T_\alpha(k,\chi)$. The "Prefactor" is an overall factor for the window function which comes from the different operations in section 3, like the integration by parts leading to factors of $\partial/\partial\chi$ in section 3.1, possibly some extra ℓ factors, and the k behavior intrinsic to the window function as in section 3.3. For convenience, we also list the resulting overall behavior of the product of source and window functions in the column "k scaling". The source function contributing to the cosmic shear spectra is identical to the "lensing" source function (second entry of the table).

The transfer functions refer to gauge-independent variables whose expression in the newtonian and synchronous gauge is explicitely given in table 7 below. The last two entries in table 6 are not used in our default implementation. We switch from the "Density" to "Density Split" terms when testing the method that uses equation (3.17) to reduce the tilt. These sources use the operator $\mathcal{D}$ given by

$$\mathcal{D} = \left[-\frac{\partial^2}{\partial\chi^2} + \frac{2}{\chi}\frac{\partial}{\partial\chi} - \frac{2}{\chi^2} \right] . \tag{A.1}$$

In table 7 we summarize the different gauge invariant source functions $T_\alpha(k,\chi)$ used throughout this paper.

JCAP10(2018)047

Name	Prefactor	Window function	Source	k scaling
Density	1	bW	T^D	1
Lensing	$\ell(\ell+1)$	$\int_\chi^\infty d\chi' \left(\frac{2-5s}{2}\right) \frac{\chi'-\chi}{\chi\chi'} W$	$T^{\phi+\psi}$	$1/k^2$
Doppler 1	$-\frac{1}{k^2}\frac{\partial}{\partial\chi}$	$W\left(1+\frac{H'}{aH}+\frac{2-5s}{\chi aH}+5s-f_{\rm evo}\right)$	T^θ	$1/k^2$
Doppler 2	$\frac{1}{k^2}$	$W(f_{\rm evo}-3)aH$	T^θ	$1/k^2$
RSD	$\frac{1}{k^2}\frac{\partial^2}{\partial\chi^2}$	$\frac{1}{aH}W$	T^θ	$1/k^2$
GR 1	1	$W\left(2+\frac{H'}{aH^2}+\frac{2-5s}{\chi aH}+5s-f_{\rm evo}\right)$	T^ψ	$1/k^2$
GR 2	1	$-W(2-5s)$	T^ψ	$1/k^2$
GR 3	1	$W/(aH)$	$T^{\psi'}$	$1/k^2$
GR 4	1	$\int_\chi^\infty d\chi' \left(\frac{2-5s}{\chi'}\right) W$	$T^{\phi+\psi}$	$1/k^2$
GR 5	1	$\int_\chi^\infty d\chi' \left(1+\frac{H'}{aH}+\frac{2-5s}{\chi' aH}+5s-f_{\rm evo}\right) W$	$T^{(\phi+\psi)'}$	$1/k^2$
Density Split 1	$\mathcal{D}$	bW	T^D	$1/k^2$
Density Split 2	$\frac{\ell(\ell+1)}{\chi^2}$	bW	T^D	$1/k^2$

Table 6. Different terms appearing in the total source functions with their respective prefactors. b is the (linear) bias, s is the magnification bias, $f_{\rm evo}$ is a possible evolution of dN/dz, all in accordance with [6], where they are defined more precisely. The layout of the table is discussed further in the text below.

Source Name & Symbol	Newtonian Gauge	Synchronous Gauge
Density T^D	$\delta_m + 3\frac{aH}{k^2}\theta_m$	$\delta_m + 3\frac{aH}{k^2}\theta_m$
Velocity T^Θ	θ_m	$\theta_m + k^2\alpha$
Bardeen Potential T^Φ	ϕ	$\eta - \mathcal{H}\alpha$
Bardeen Potential T^Ψ	ψ	$\mathcal{H}\alpha + \alpha'$

Table 7. Gauge invariant quantities in Newtonian and Synchronous Gauge. These correspond to the source functions in `CLASS`. For the notation of Synchronous Gauge see [28]. See also [6].

JCAP10(2018)047

B Details on the power law decomposition

In this section we want to discuss how the power law decomposition in equation (2.29) is done in practice on a finite-precision computer. We want to perform an expansion of the type

$$f(\boldsymbol{x};k) = \sum_n c_n(\boldsymbol{x})\; k^{\nu_n}\,, \quad \text{where } \nu_n = \frac{2\pi i\, n}{\log(k_{\rm max}/k_{\rm min})} + b\,. \tag{2.29}$$

First, we should note that the precise form of ν_n depends on our choice of sampling in $\log(k)$. In our code, we use a slightly modified definition,

$$\tilde{\nu}_n = \frac{2\pi i\, n}{T}\frac{N-1}{N} + b\,, \tag{B.1}$$

where N is the number of samples in $\log(k)$ and $T = \log(k_{\rm max}/k_{\rm min})$.

This choice is based on the following argument: for N evenly sampled values of $\log(k_m)$ with $m \in \{0, \ldots, N-1\}$ we obtain

$$\log(k_m) = \log(k_{\rm min}) + \frac{m}{N-1}\left(\log(k_{\rm max}) - \log(k_{\rm min})\right)\,, \tag{B.2}$$

such that the sampling is linear, starts at $k_{\rm min}$ and ends at $k_{\rm max}$.

If we first set the tilt b to zero, the FFT coefficients calculated by the direct FFT algorithm are simply given as

$$\tilde{c}_n = \sum_m f(k_m)\exp(-2\pi i\, n\, m/N)\,, \tag{B.3}$$

and transforming them according to

$$c_n = \frac{1}{N}(k_{\rm min})^{-\tilde{\nu}_n}\tilde{c}_n = \frac{1}{N}\exp\left(-\log(k_{\rm min})2\pi i\, n \Big/ T\frac{N-1}{N}\right)\tilde{c}_n\,, \tag{B.4}$$

gives us the coefficients with our desired back-transformation properties

$$\begin{aligned}
\sum_n c_n k_a^{\tilde{\nu}_n} &= \sum_m\sum_n f(k_m)\exp(-2\pi i n m/N)\exp\left(-\log(k_{\rm min})2\pi i\, n \Big/ T\frac{N-1}{N}\right)k_a^{\tilde{\nu}_n}\Big/ N \\
&= \sum_m\sum_n f(k_m)\exp(-2\pi i\, n\, m/N)\exp\left((\log(k_a)-\log(k_{\rm min}))2\pi i\, n \Big/ T\frac{N-1}{N}\right)\Big/ N \\
&= \sum_m\sum_n f(k_m)\exp(-2\pi i\, n\, m/N)\exp\left(\frac{a}{N-1}2\pi i\, n\frac{N-1}{N}\right)\Big/ N \\
&= \sum_m\sum_n f(k_m)\exp(2\pi\, i\, n(a-m)n/N)/N \\
&= \sum_m f(k_m)\delta_{am} = f(k_a)\,.
\end{aligned} \tag{B.5}$$

The factor $(N-1)/N$ in the definition of the $\tilde{\nu}_n$ can thus be explicitly seen as a "correction" for our choice of sampling in $\log(k)$, which involved a scaling $m/(N-1)$ instead of m/N for

JCAP10(2018)047

the usual FFT. This sampling was chosen in such a way as to have $k_{\rm max}$ be the maximum sampled value.

Transforming instead $f(k)k^{-b}$ introduces the tilt b in our definition of the Fourier mode ν_n. If we choose to transform precisely $f(k)(k/k_{\rm min})^{-b}$, the relation $c_n = \frac{1}{N}(k_{\rm min})^{-\tilde{\nu}_n}\tilde{c}_n$ remains valid even for general $b \neq 0$. Overall, the coefficients c_n can be obtained by a fast implementation using the FFT algorithm and a simple multiplication by $(k_{\rm min})^{-\tilde{\nu}_n}/N$.

C Transformations of the Bessel integrals

Analytic limit. Many properties used in the following section can be found on http://functions.wolfram.com/HypergeometricFunctions/Hypergeometric2F1. Further discussion on hypergeometric functions can also be found in [14].

First we note that

$$ {}_2F_1\left(a,b;c \mid 1\right) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}\,, \qquad \text{where} \quad \mathrm{Re}[c-a-b] > 0\,. \tag{C.1}$$

In our case $\mathrm{Re}[c-a-b] = 2-\mathrm{Re}[\nu_n]$, requiring $\mathrm{Re}[\nu_n] < 2$. This corresponds to the analytical result

$$ I_\ell(\nu,1) = \frac{\pi^{\frac{3}{2}}\Gamma\left(\ell+\frac{\nu}{2}\right)\Gamma\left(1-\frac{\nu}{2}\right)}{\Gamma\left(\frac{3-\nu}{2}\right)\Gamma\left(\ell+2-\frac{\nu}{2}\right)}\,. \tag{C.2}$$

This is the analytic limit for $t \to 1$.

Limber limit. We now want to discuss the Limber limit, not used in our implementation but mentioned in equation (2.37) for self-consistency checks. First, we want to note that one has to be careful not to double-count the point $t = 1$ when using equation (2.37). Thus, the replacement for the $I_\ell(\nu,t)$ for the integration from t in $[0,1]$ is actually

$$ I_\ell(\nu,t) \xrightarrow[\ell\to\infty]{} \pi^2\,(\ell_0\,)^{\nu-3}\,\delta(1-t)\,, \tag{C.3}$$

which explicitly counts the point $t = 1$ only once per integration. Using the Sterling approximation,

$$ \lim_{z\to\infty}\frac{\sqrt{2\pi z}\,(z/e)^z}{\Gamma(z+1)} = 1\,, \tag{C.4}$$

we find that for large ℓ this result converges towards

$$ I_\ell(\nu,1) \xrightarrow[\ell\to\infty]{} \pi^{\frac{3}{2}}\frac{\Gamma\left(1-\frac{\nu}{2}\right)}{\Gamma\left(\frac{3-\nu}{2}\right)}\frac{\ell^{\ell+\nu/2}}{\ell^{\ell+2-\nu/2}}\,, \tag{C.5}$$

which supports the statement that for large ℓ we have $I_\ell(\nu,1) \sim \ell^{\nu-2}$ in the Limber approximation. If we assume that the dominant t dependence stems from the t^ℓ factor for large ℓ (this is a good approximation), we can immediately see that the condition $t_{\rm min}^\ell = \epsilon$ implies $t_{\rm min} = \epsilon^{1/\ell} \approx 1 + log(\epsilon)/\ell + \mathcal{O}(1/\ell^2)$, and thus $1 - t_{\rm min} \approx log(1/\epsilon)/\ell + \mathcal{O}(1/\ell^2)$, which is proportional to $1/\ell$ for $\epsilon \ll 1$. Since the support of the function shrinks like $1/\ell$ while its amplitude grows as $\ell^{\nu-2}$, we recover the argument that the Limber limit reads $\ell^{\nu-3}\delta(1-t)$ for large ℓ. This is further supported by figures 5 and 6. We have additionally confirmed using `Mathematica` and `mpmath` in Python, that the integrated area behaves like $\ell^{\nu-3}$ for large ℓ as expected.

JCAP10(2018)047

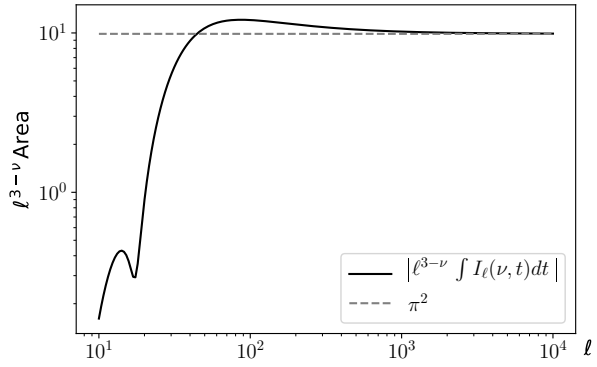

Figure 5. An illustration of the Limber limit: for large ℓ the area under the curve $I_\ell(\nu, t)$ approaches $\pi^2 \ell^{\nu-3}$ when integrated from 0 to 1. We see that the factor $\ell^{3-\nu}\, I_\ell(\nu, t)$ approaches the constant π^2, which is an equivalent statement.

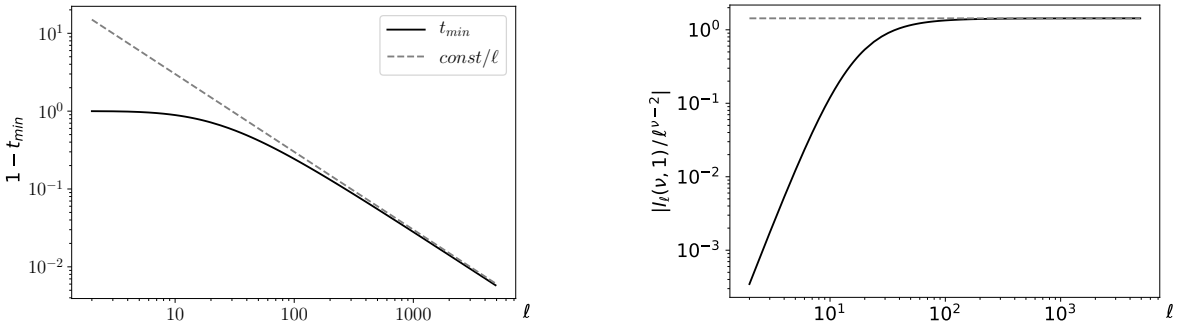

Figure 6. Another consequence of the Limber limit: for large ℓ the $t_{\min}$ parameter behaves as $1/\ell$ (left), and the $|I_\ell(\nu, 1)/\ell^{\nu-2}|$ is constant as in equation (C.5) (right). Note that the oscillations due to imaginary ν are correctly captured and the relative size approaches the correct constant. The black lines indicate the behavior for $\nu = -2.1 + 30i$ and $\epsilon = 10^{-4}$, while the grey lines specify asymptotes. On the left, the grey line is ℓ^{-1} times an arbitrary constant (here $30/\ell$), while on the right side the constant is fixed by C.5. The constant for $t_{\min}$ is not exactly $\log(1/\epsilon)$ because of the influence of the hypergeometric function.

Transformations. For high t we can transform our original $I_\ell(\nu, t)$ using the properties of hypergeometric functions to speed up convergence, by bringing the argument closer to 0, where the hypergeometric function's Taylor expansion converges faster. We find for high t the transformation

$$I_\ell(\nu,t) = \frac{\pi^{\frac{3}{2}} t^\ell \left(\frac{2}{1+z}\right)^{\ell+\nu/2}}{\Gamma\left(\frac{3-\nu}{2}\right)} \left[\frac{\Gamma\left(\frac{2\ell+\nu}{2}\right)\Gamma\left(\frac{2-\nu}{2}\right)}{\Gamma\left(\frac{4+2\ell-\nu}{2}\right)}\, {}_2F_1\left(\frac{2\ell+\nu}{4}, \frac{2\ell+\nu+2}{4}, \frac{\nu}{2}, \left(\frac{1-z}{1+z}\right)^2\right)\right.$$
$$\left. + \Gamma\left(\frac{\nu-2}{2}\right)\left(2\left(\frac{1+z}{1-z}\right)\right)^{\nu-2} {}_2F_1\left(\frac{2\ell+6-\nu}{4}, \frac{2\ell+4-\nu}{4}, 2-\frac{\nu}{2}, \left(\frac{1-z}{1+z}\right)^2\right)\right], \tag{C.6}$$

where $z = t^2$. For small t, this formula depends on a precise cancellation between the two hypergeometric functions, disfavoring its use in this regime. Instead, for small t we find

JCAP10(2018)047

another transformation that allows for the efficient calculation of $I_\ell(\nu,t)$

$$I_\ell(\nu,t) = \frac{2^{\nu-1}\pi^2 t^\ell\,(1+z)^{-\left(\ell+\frac{\nu}{2}\right)}}{\Gamma\left(\frac{3-\nu}{2}\right)}\frac{\Gamma\left(\frac{2\ell+\nu}{2}\right)}{\Gamma\left(\ell+3/2\right)}\,{}_2F_1\left(\frac{2\ell+\nu}{4},\frac{2\ell+\nu+2}{4},\ell+\frac{3}{2},\frac{4z}{(1+z)^2}\right). \tag{C.7}$$

Notice that on the right hand side of this formula the imaginary frequencies ν in the arguments of the hypergeometric function are divided by a factor of 4. This is very helpful for faster convergence of the numerical evaluation. Similarly to [14] we prescribe to switch from version (C.7) to (C.6) above some t^*. One can easily perform numerical test to find the most optimal choice of t^* is roughly $t^* \approx 0.9$.

Derivatives of the Bessel functions. The integration by parts method in section 3.1 works well most of the time, but there are some situations in which it can fail:

1. When the window function is a Dirac distribution, which makes the operation of taking derivatives very imprecise.
2. When the window function is a function similar to a top-hat, for which the second derivatives are very localized around the edges of the function. This results in narrow peaks in the $f_n^{ij}(t)$ function at values corresponding to the edge ratios.

In our current implementation, we just discard these two cases. We can give however some hints of possible solutions. First of all, it is possible to use the following equation to translate a derivative into a change of ℓ,

$$\frac{\partial}{\partial x}j_\ell(x) = \frac{\ell}{x}j_\ell(x) - j_{\ell+1}(x)\,. \tag{C.8}$$

Thus, for each derivative of the Bessel function we obtain the two integrals, $\ell I_\ell(\nu-1,t)$ and

$$4\pi\int_0^\infty du u^{\nu-1} j_{\ell+1}(u) j_\ell(ut) \equiv J_\ell^{(1)}(\nu,t)\,. \tag{C.9}$$

Thus the integrals $I_\ell(\nu,t)$ and $J_\ell^{(1)}(\nu,t)$ provide a basis for all derivatives of Bessel functions. Sadly, the $J_\ell^{(1)}(\nu,t)$ does not have nice symmetry properties with respect to t as $I_\ell(\nu,t)$ does. In order to deal only with $t<1$, we could decide to define

$$J_\ell^{(2)}(\nu,t) \equiv 4\pi\int_0^\infty du u^{\nu-1} j_\ell(u) j_{\ell+1}(ut)\,, \tag{C.10}$$

and relate $J^{(1)}(\nu,t)$ for $t>1$ to $J^{(2)}(\nu,1/t)$ and vice versa.

Secondly, one can relate all derivatives of the Bessel functions to the derivatives of $I_\ell(\nu,t)$, e.g.

$$\begin{aligned} 4\pi\int_0^\infty du u^{\nu-1} j_\ell(u)\left.\frac{\partial j_\ell(x)}{\partial x}\right|_{x=ut} &= \partial_t\, I_\ell(\nu-1,t)\,, \\ 4\pi\int_0^\infty du u^{\nu-1}\left.\frac{\partial j_\ell(x)}{\partial x}\right|_{x=u} j_\ell(ut) &= -t\,\partial_t\, I_\ell(\nu-1,t) + (1-\nu) I_\ell(\nu-1,t)\,. \end{aligned} \tag{C.11}$$

JCAP10(2018)047

δz	0.5	0.1	0.05	0.01	0.005	0.01
q logstep	80	50	33	21	5	0.05
selection sampling bessel	1.0	1.9	2.5	1.0	1.0	9.0
hyper sampling flat	7	7.5	7	7	7	7

Table 8. Precision parameters used for the tests presented in table 5.

	Name	Value	Name	Value
Common parameters	h	0.6711	T_{cmb}	2.726K
	ω_b	0.02207	Y_{He}	0.25
	N_{eff}	3.04	A_s	$2.22 \cdot 10^{-9}$
	z_{reio}	10.	n_s	0.97
Massless neutrinos	ω_{cdm}	0.12029		
Massive neutrinos	ω_{cdm}	0.10965534872	N_{ncdm}	1
	m_{ncdm}	$3 \cdot 0.33$eV		

Table 9. The cosmological parameters used in the comparison study of section 4.

Of course, since the derivatives of the hypergeometric functions are known analytically, these expressions are also known from analytic formulas. In either case, there is a strong computational effort involved compared to the simple integration by parts method that we use for smooth window functions.

If one wants to use the relations from appendix G of [15], one would also have to expand the Doppler terms involving only first derivatives using the following equation

$$k j_\ell(k\chi) = \frac{\ell - 1}{\chi} j_{\ell-1}(k\chi) - \frac{\partial}{\partial \chi} j_{\ell-1}(k\chi)\,, \tag{C.12}$$

to keep the separation of appearing Bessel functions to $\Delta\ell = 2$, which allows the given recursion relations in $\Delta\ell$ to be used. The Redshift-Space Distortion term results only in $\delta\ell = -2, 0, 2$.

D Further notes

Width measurement parameters. We give below the precision parameters used in our tests involving different bin widths, presented in table 5.

Cosmological parameters. In all our tests, we used a fixed set of ΛCDM model parameters summarized in table 9. The value of Ω_Λ is adjusted to cancel the spatial curvature. Y_{He} is the Helium fraction, z_{reio} is the redshift of reionization, N_{eff} is the effective neutrino number. The index ncdm refers to non-cold Dark Matter, in our case massive neutrinos. We assumed 3 neutrino species with degenerate mass of 0.33 eV , for which we held constant N_{eff} , H_0 , Ω_Λ , and ω_b , implying fixed ω_m.

Discrete ℓ values. The list of ℓ values used in section 4 for testing purposes is given below.

JCAP10(2018)047

2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	19	21	23	25	27	30	33	36
40	44	49	54	60	67	75	83	92	103	115	128
143	160	179	200	223	249	278	311	348	388	428	468
508	548	588	628	668	708	748	788	828	868	908	948
988	1000										

Table 10. The ℓ values for the calculation of the number count spectra ($\ell_{\rm max} = 1000$, $N_\ell = 62$).

2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	19	21	23	25	27	30	33	36
40	44	49	54	60	67	75	83	92	103	115	128
143	160	179	200	223	249	278	311	348	388	428	468
508	548	588	628	668	708	748	788	828	868	908	948
988	1028	1068	1108	1148	1188	1228	1268	1308	1348	1388	1428
1468	1508	1548	1588	1628	1668	1708	1748	1788	1828	1868	1908
1948	1988	2028	2068	2108	2148	2188	2228	2268	2308	2348	2388
2428	2468	2508	2548	2588	2628	2668	2708	2748	2788	2828	2868
2908	2948	2988	3000								

Table 11. The ℓ values for the calculation of the Cosmic shear spectra ($\ell_{\rm max} = 3000$, $N_\ell = 112$).

References

[1] J. Yoo, A.L. Fitzpatrick and M. Zaldarriaga, *A New Perspective on Galaxy Clustering as a Cosmological Probe: General Relativistic Effects*, *Phys. Rev.* **D 80** (2009) 083514 [arXiv:0907.0707] [INSPIRE].

[2] J. Yoo, *General Relativistic Description of the Observed Galaxy Power Spectrum: Do We Understand What We Measure?*, *Phys. Rev.* **D 82** (2010) 083508 [arXiv:1009.3021] [INSPIRE].

[3] C. Bonvin and R. Durrer, *What galaxy surveys really measure*, *Phys. Rev.* **D 84** (2011) 063505 [arXiv:1105.5280] [INSPIRE].

[4] A. Challinor and A. Lewis, *The linear power spectrum of observed source number counts*, *Phys. Rev.* **D 84** (2011) 043516 [arXiv:1105.5292] [INSPIRE].

[5] D. Bertacca, R. Maartens, A. Raccanelli and C. Clarkson, *Beyond the plane-parallel and Newtonian approach: Wide-angle redshift distortions and convergence in general relativity*, *JCAP* **10** (2012) 025 [arXiv:1205.5221] [INSPIRE].

[6] E. Di Dio, F. Montanari, J. Lesgourgues and R. Durrer, *The CLASSgal code for Relativistic Cosmological Large Scale Structure*, *JCAP* **11** (2013) 044 [arXiv:1307.1459] [INSPIRE].

[7] A. Lewis, A. Challinor and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473 [astro-ph/9911177] [INSPIRE].

[8] D. Blas, J. Lesgourgues and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes*, *JCAP* **07** (2011) 034 [arXiv:1104.2933] [INSPIRE].

[9] E. Di Dio, F. Montanari, R. Durrer and J. Lesgourgues, *Cosmological Parameter Estimation with Large Scale Structure Observations*, *JCAP* **01** (2014) 042 [arXiv:1308.6186] [INSPIRE].

[10] E. Di Dio, F. Montanari, A. Raccanelli, R. Durrer, M. Kamionkowski and J. Lesgourgues, *Curvature constraints from Large Scale Structure*, *JCAP* **06** (2016) 013 [arXiv:1603.09073] [INSPIRE].

[11] V. Tansella, C. Bonvin, R. Durrer, B. Ghosh and E. Sellentin, *The full-sky relativistic correlation function and power spectrum of galaxy number counts. Part I: theoretical aspects*, *JCAP* **03** (2018) 019 [arXiv:1708.00492] [INSPIRE].

[12] V. Tansella, G. Jelic-Cizmek, C. Bonvin and R. Durrer, *COFFE: a code for the full-sky relativistic galaxy correlation function*, arXiv:1806.11090 [INSPIRE].

[13] U. Seljak and M. Zaldarriaga, *A Line of sight integration approach to cosmic microwave background anisotropies*, *Astrophys. J.* **469** (1996) 437 [astro-ph/9603033] [INSPIRE].

[14] V. Assassi, M. Simonović and M. Zaldarriaga, *Efficient evaluation of angular power spectra and bispectra*, *JCAP* **11** (2017) 054 [arXiv:1705.05022] [INSPIRE].

[15] H.S. Grasshorn Gebhardt and D. Jeong, *Fast and accurate computation of projected two-point functions*, *Phys. Rev.* **D 97** (2018) 023504 [arXiv:1709.02401] [INSPIRE].

[16] A.J.S. Hamilton, *Uncorrelated modes of the nonlinear power spectrum*, *Mon. Not. Roy. Astron. Soc.* **312** (2000) 257 [astro-ph/9905191] [INSPIRE].

[17] M. Schmittfull, Z. Vlah and P. McDonald, *Fast large scale structure perturbation theory using one-dimensional fast Fourier transforms*, *Phys. Rev.* **D 93** (2016) 103528 [arXiv:1603.04405] [INSPIRE].

[18] J.E. McEwen, X. Fang, C.M. Hirata and J.A. Blazek, *FAST-PT: a novel algorithm to calculate convolution integrals in cosmological perturbation theory*, *JCAP* **09** (2016) 015 [arXiv:1603.0482].

[19] D.N. Limber, *The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field.*, *Astrophys. J.* **117** (1953) 134.

[20] N. Kaiser, *Weak lensing and cosmology*, *Astrophys. J.* **498** (1998) 26 [astro-ph/9610120] [INSPIRE].

[21] P. Lemos, A. Challinor and G. Efstathiou, *The effect of Limber and flat-sky approximations on galaxy weak lensing*, *JCAP* **05** (2017) 014 [arXiv:1704.01054] [INSPIRE].

[22] J. Lesgourgues and T. Tram, *Fast and accurate CMB computations in non-flat FLRW universes*, *JCAP* **09** (2014) 032 [arXiv:1312.2697] [INSPIRE].

[23] M. LoVerde and N. Afshordi, *Extended Limber Approximation*, *Phys. Rev.* **D 78** (2008) 123506 [arXiv:0809.5112] [INSPIRE].

[24] VIRGO CONSORTIUM collaboration, R.E. Smith et al., *Stable clustering, the halo model and nonlinear cosmological power spectra*, *Mon. Not. Roy. Astron. Soc.* **341** (2003) 1311 [astro-ph/0207664] [INSPIRE].

[25] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya and M. Oguri, *Revising the Halofit Model for the Nonlinear Matter Power Spectrum*, *Astrophys. J.* **761** (2012) 152 [arXiv:1208.2701] [INSPIRE].

[26] S. Bird, M. Viel and M.G. Haehnelt, *Massive Neutrinos and the Non-linear Matter Power Spectrum*, *Mon. Not. Roy. Astron. Soc.* **420** (2012) 2551 [arXiv:1109.4416] [INSPIRE].

[27] J.E. Campagne, J. Neveu and S. Plaszczynski, *Angpow: a software for the fast computation of accurate tomographic power spectra*, *Astron. Astrophys.* **602** (2017) A72 [arXiv:1701.03592] [INSPIRE].

[28] C.-P. Ma and E. Bertschinger, *Cosmological perturbation theory in the synchronous versus conformal Newtonian gauge*, *Astrophys. J.* **455** (1995) 7 [astro-ph/9401007] [INSPIRE].

JCAP10(2018)047