

# CORE COSMOLOGY LIBRARY: PRECISION COSMOLOGICAL PREDICTIONS FOR LSST

HUSNI ALMOUBAYYED,<sup>1</sup> DAVID ALONSO,<sup>2,3</sup> JONATHAN BLAZEK,<sup>4,5</sup> PHILIP BULL,<sup>6,7</sup> JEAN-ERIC CAMPAGNE,<sup>8</sup>  
N. ELISA CHISARI,<sup>3</sup> ALEX DRLICA-WAGNER,<sup>9</sup> ZILONG DU,<sup>10</sup> TIM EIFLER,<sup>11,12</sup> JOHN ELLISON,<sup>10</sup> RENÉE HLOZEK,<sup>13</sup>  
MUSTAPHA ISHAK,<sup>14</sup> SHAHAB JOUDAKI,<sup>3</sup> MATTHEW KIRBY,<sup>15</sup> DAVID KIRKBY,<sup>16</sup> ELISABETH KRAUSE,<sup>17</sup>  
C. DANIELLE LEONARD,<sup>1</sup> CHRISTIANE S. LORENZ,<sup>3</sup> PHIL MARSHALL,<sup>18</sup> THOMAS MCCLINTOCK,<sup>15</sup> SEAN MCCLAUGHLIN,<sup>19</sup>  
ALEXANDER MEAD,<sup>20</sup> JÉRÉMY NEVEU,<sup>8</sup> STÉPHANE PLASZCZYNSKI,<sup>8</sup> JAVIER SANCHEZ,<sup>16</sup> SUKHDEEP SINGH,<sup>1,21</sup>  
ANŽE SLOSAR,<sup>22</sup> TILMAN TRÖSTER,<sup>23</sup> ANTONIO VILLARREAL,<sup>24</sup> MICHAL VRASIL,<sup>25</sup> ERIKA L. WAGONER,<sup>15</sup> AND  
JOE ZUNTZ<sup>23</sup>

(LSST DARK ENERGY SCIENCE COLLABORATION)

<sup>1</sup>*McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

<sup>2</sup>*School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, CF24 3AA, United Kingdom*

<sup>3</sup>*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, United Kingdom*

<sup>4</sup>*Center for Cosmology and Astroparticle Physics, Ohio State, Columbus, OH 43210, USA*

<sup>5</sup>*Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland*

<sup>6</sup>*Department of Astronomy, University of California Berkeley, Berkeley, CA 94720, USA*

<sup>7</sup>*Radio Astronomy Laboratory, University of California Berkeley, Berkeley, CA 94720, USA*

<sup>8</sup>*Laboratoire de l'Accélérateur Linéaire, Université Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France*

<sup>9</sup>*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*

<sup>10</sup>*Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA*

<sup>11</sup>*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA*

<sup>12</sup>*Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA*

<sup>13</sup>*Dunlap Institute for Astronomy and Astrophysics & Department for Astronomy and Astrophysics, University of Toronto, ON M5S 3H4*

<sup>14</sup>*Department of Physics, The University of Texas at Dallas, Richardson, TX 75083, USA*

<sup>15</sup>*Department of Physics, University of Arizona, Tucson, AZ 85721, USA*

<sup>16</sup>*Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*

<sup>17</sup>*Kavli Institute for Particle Astrophysics and Cosmology, Stanford, CA 94305-4085, USA*

<sup>18</sup>*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*

<sup>19</sup>*Stanford University, Stanford, CA, 94305, USA*

<sup>20</sup>*Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada*

<sup>21</sup>*Berkeley Center for Cosmological Physics and Department of Physics, University of California, Berkeley, California*

<sup>22</sup>*Brookhaven National Laboratory, Physics Department, Upton, NY 11973, USA*

<sup>23</sup>*Institute for Astronomy, Royal Observatory Edinburgh, Edinburgh EH9 3HJ, UK*

<sup>24</sup>*Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh PA 15260*

<sup>25</sup>*Institute of Physics CAS, Prague, 182 21, CZ*

## ABSTRACT

The Core Cosmology Library (CCL) provides routines to compute basic cosmological observables with validated numerical accuracy. In the current public version, predictions are provided for distances and background quantities, angular auto- and cross-spectra of cosmic shear, lensing of the microwave background and clustering, and the halo mass function through state-of-the-art modeling prescriptions available in the literature. These predictions have been validated to a numerical accuracy level documented herein against the results of independent implementations. Fiducial specifications for the expected galaxy distributions and clustering bias for the Large Synoptic Survey Telescope are also included, together with the capability of computing redshift distributions for a user-defined photometric redshift model. CCL is written in C, with a Python interface. **JAB: General comment: Do we want to make more clear that we are always referring to numerical accuracy rather than accuracy with “reality?” Or is this obvious?****EC: Made some additions in the intro to make it more clear.**

Contents		2.2. Growth of perturbations	5
		2.3. Matter power spectrum	6
1. Introduction	2	2.4. Two-point correlators	7
		2.4.1. Angular power spectra	7
2. Cosmological models and observables	3	2.4.2. Correlation functions	9
2.1. Background cosmology	4		

2.4.3. 3-dimensional spatial correlation function.	10
2.5. Halo mass function	10
2.6. Halo bias	11
2.7. Photometric redshifts	12
3. Implementation of high-accuracy cosmological functions	12
3.1. Background functions & growth of perturbations	12
3.2. Matter power spectrum	12
3.3. Angular power spectra	13
3.3.1. Limber approximation	13
3.3.2. Beyond Limber: <b>Angpow</b>	13
3.4. Correlation functions	14
3.5. Halo mass function	14
3.6. Massive neutrinos	15
4. Validation over the $\Lambda$ CDM parameter space	15
4.1. Background quantities & growth of perturbations	16
4.2. Matter power spectra	16
4.2.1. Analytic expressions	16
4.2.2. Validation of interpolation schemes	16
4.2.3. Generalized validation of the power spectrum over $\Lambda$ CDM parameter space	20
4.2.4. Validation of the Cosmic Emulator implementation	22
4.3. Halo bias and halo mass function	22
4.4. Two-point statistics	23
4.4.1. <b>Angpow</b>	25
5. Usage	25
6. Outlook	25

## 1. INTRODUCTION

Starting in the next decade, large-scale galaxy surveys will drive a new era of high precision cosmology (LSST Dark Energy Science Collaboration 2012; Green et al. 2011; Laureijs et al. 2011). Their main goal is arguably to answer the question of the origin of cosmic acceleration, in other words, to elucidate the nature of “dark energy”, broadly understood as a family of potential models: from a cosmological constant to a dynamical field and modifications of gravity (see for example Carroll (2001); Peebles & Ratra (2003); Padmanabhan (2003); Copeland et al. (2006); Ishak (2007); Weinberg et al. (2013) and references therein). These data will also allow us to shed light on a number of open questions in fundamental physics, such as the nature of dark matter, the mass of neutrinos or the level of non-Gaussianity in the initial conditions **To do: maybe add some references here.**

High precision constraints on dark energy models will be achieved by probing at the same time the expansion and growth history of the Universe over a long redshift baseline. For this purpose, it will be crucial to combine measurements of multiple cosmological probes: weak and strong gravitational lensing, the clustering of galaxies, distances to supernovae, and the abundance, clustering and gravitational lensing of galaxy clusters. Current weak lensing surveys have started to take this approach already (Joudaki et al. 2018; van Uitert et al. 2018; DES Collaboration et al. 2017; Krause et al. 2017). From a theoretical perspective, there are two challenges faced by the next generation of galaxy surveys.

The first one is to ensure that all probes are modeled accurately from a *physical* point of view, including cosmological, astrophysical, and observational effects, to avoid potential biases in the final cosmological results. In the context of weak gravitational lensing, for example, phenomena that can lead to biases include the impact of baryons on the distribution of matter and the intrinsic alignments of galaxies (e.g. van Daalen et al. 2011; Semboloni et al. 2011; Troxel & Ishak 2015; Krause et al. 2016; Blazek et al. 2017). In the context of galaxy clustering, the most relevant astrophysical systematic is the galaxy-matter bias relation on small scales (Zhao et al. 2013; Desjacques et al. 2016), and effects such as magnification of number counts and redshift space distortions need to be included in the models (Alonso & Ferreira 2015; Ghosh et al. 2018).

Second, even standard cosmological quantities in the simplest models, such as distances in a  $\Lambda$ CDM cosmology, have to be predicted to a validated degree of *numerical* accuracy. Achieving this objective is not always trivial, as computing these quantities generally requires, for

example, numerical integration or interpolation, both of which are prone to numerical error. In this work, we establish the numerical accuracy of our predictions by performing comparisons between independent implementations of the same models.

Faced with these challenges, the Dark Energy Science Collaboration (DESC), one of the science collaborations of the Large Synoptic Survey Telescope (LSST), has built a comprehensive software tool that satisfies the needs of the next generation of cosmological analysis: the Core Cosmology Library<sup>1</sup> (CCL). CCL is a software library providing the infrastructure to make theoretical predictions that are validated to a well-documented numerical accuracy for the purpose of constraining cosmology. While CCL was built with LSST in mind, the goal is to produce a publicly available, user-friendly, well-documented, adaptable software that can be used in any theoretical modeling work in cosmology.

In its current form, CCL computes standard cosmological functions including the Hubble parameter, cosmological distances, density parameters, the halo mass function, halo bias and linear growth functions. It calculates the matter power spectrum using various methods including common approximations, by calling external software such as CLASS (Blas et al. 2011), or emulators, such as the “Cosmic Emulator” of Lawrence et al. (2017). It computes 2-point angular power spectra and correlation functions from various probes, going beyond the Limber approximation. While CCL incorporates state-of-the-art models available in the literature, this manuscript is mainly concerned with documenting their implementation and *numerical* accuracy, but does not address the *physical* accuracy of each model, for which we point the reader to the relevant references in the following sections.

CCL’s overall structure is illustrated in Figure 1. Our implementation has support for spatially flat and curved  $\Lambda$ -Cold Dark Matter ( $\Lambda$ CDM) cosmologies, and  $w$ CDM cosmologies with the option of using a time-dependent equation of state. It also allows for cosmologies with multiple massive neutrino species and can be linked to external software for modified gravity predictions (hi\_CLASS, Zumalacárregui et al. 2017).

This paper is organized as follows. Section 2 describes the cosmological models and observables supported by CCL. In Section 3, we describe the details of the implementation of the quantities introduced in Section 2. Section 4 provides details of the validation tests performed and the accuracy achieved. Section 5 gives brief

guidelines for the usage of CCL, although we direct the reader to the software online repository, documentation and user manual for further information. We conclude in Section 6 with an outlook towards the integration of CCL in the LSST DESC pipelines, and we outline future additions to the software.

## 2. COSMOLOGICAL MODELS AND OBSERVABLES

The overarching goal of CCL is to allow seamless integration of different cosmological models of interest to LSST.

The cosmological components include the matter density parameter  $\Omega_m$ , the dark energy density  $\Omega_\Lambda$ , the radiation density  $\Omega_g$ , the curvature density  $\Omega_K$ , the neutrino density of both massless and massive neutrinos, given by  $\Omega_{\nu,\text{rel}}$  and  $\Omega_{\nu,m}$  respectively, and parameters for the equation of state of dark energy defined below. The normalization of the density fluctuations is established either in terms of the amplitude of the primordial power spectrum,  $A_s$ , which is a power-law of index  $n_s$ , or in terms of the RMS variance in spheres of  $8\ h/\text{Mpc}$  today,  $\sigma_8$ .

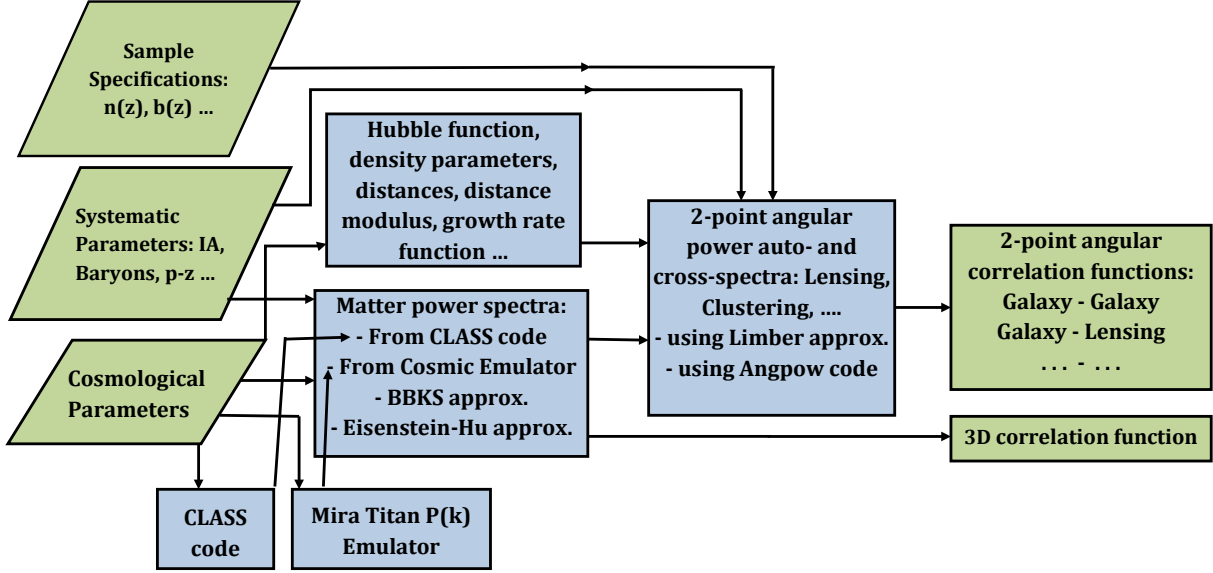
Currently, the following families of models are supported:

- Flat  $\Lambda$ CDM cosmology with parameters  $\Omega_b$ ,  $n_s$ ,  $A_s$  or  $\sigma_8$ ,  $\Omega_m$ , and a cosmological constant dark energy model with equation-of-state  $w = -1$ .
- $w$ CDM and the Chevallier-Polarski-Linder (CPL) model for dark energy and its redshift dependence, which adopts the following parametrization for  $w$  as a function of the scale factor,  $a$  (Chevallier & Polarski 2001 and Linder 2003),

$$w(a) = w_0 + w_a(1 - a). \quad (1)$$

- A universe with non-zero curvature ( $K$ ) so that the curvature energy density parameter is the difference of the sum of the energy densities of the other components with respect to unity (flatness), i.e.  $\Omega_K = 1 - \sum_i \Omega_i$ , including  $\Omega_\Lambda$ , the dark energy density.
- All of the above, plus an arbitrary, user-defined modified growth function (see description in Section 2.2).
- Relativistic neutrinos in combination with any of the above, specified by  $N_{\text{eff}}$  (the number of relativistic neutrino species in the early Universe) and by the number of massive neutrinos today (see next item).

<sup>1</sup> publicly available at <https://github.com/LSSTDESC/CCL>



**Figure 1.** CCL structure flowchart. CCL is written in C with a Python interface. CCL routines calculate basic cosmological functions such as the Hubble function, density parameters, distances and growth function. It uses various methods to compute the matter-power spectrum including CLASS, “Cosmic Emulator” developed by Lawrence et al. (2017), or common approximations. It computes 2-point angular power spectra and correlation functions from various probes. CCL is designed to accommodate multiple methods to calculate cosmological observables.

- Massive neutrino species in combination with any of the above except the user-defined modified growth function, specified by either  $\Sigma m_\nu$  (which maps on to the density  $\Omega_{\nu,m}$  above), or by the individual masses of each of three neutrino species.

In the particular case of cosmologies with massive neutrinos, CCL allows the user to specify either a sum of masses,  $\Sigma m_\nu$ , or the individual mass of each of three neutrino species. In the former case, CCL will by default split  $\Sigma m_\nu$  into three neutrino masses which are consistent with the normal hierarchy (see, e.g. Lattanzi & Gerbino (2017) for a review). However, the user can alternatively ask for  $\Sigma m_\nu$  to be split either into masses consistent with the inverted hierarchy, or into equal masses. Each neutrino species is then checked for whether it is non-relativistic (massive) at  $z = 0$ , and this information is used in combination with the user-provided value of  $N_{\text{eff}}$  (the effective number of relativistic neutrino species in the early universe) to set the number of relativistic neutrino species.

Not all features of CCL are available for all models. For a guide to which predictions are available for each model, see Table 1. Note that if users install their own version of CLASS, CCL can then make predictions for a more extended set of cosmologies. Users should take care to understand the validity of the CCL assumptions for their own models.

### 2.1. Background cosmology

The models that are specified above map directly onto cosmological observables such as the expansion rate of the Universe, which is parameterized through the Hubble parameter as

$$\frac{H(a)}{H_0} = a^{-3/2} \left( \Omega_m + \Omega_\Lambda a^{-3(w_0+w_a)} \exp[3w_a(a-1)] + \Omega_K a + (\Omega_g + \Omega_{\nu,\text{rel}}) a^{-1} + \tilde{\Omega}_{\nu,m}(a) a^3 \right)^{\frac{1}{2}}, \quad (2)$$

which is a function of the energy density in the different components today and the scale factor. In this expression, we have assumed the CPL parameterization described above for the dark energy equation of state, and  $\tilde{\Omega}_{\nu,m}(a) = \rho_{\text{crit},0}^{-1} \rho_{\nu,m}(a)$  is the fractional energy density of massive neutrinos, calculated via

$$\tilde{\Omega}_{\nu,m}(a) = \frac{7}{8} \sum_{i=1}^{N_\nu} \frac{4\sigma_B}{c\rho_{\text{crit},0}} \left( \frac{T_\nu^{\text{eff}}}{a} \right)^4 \int_0^\infty dx x^2 \frac{\sqrt{x^2 + (\mu^i)^2}}{\exp(x) + 1}. \quad (3)$$

Here,  $\sigma_B$  is the Stefan-Boltzmann constant,  $c$  is the speed of massless particles,  $\rho_{\text{crit},0}$  is the critical density at  $z = 0$ , and  $T_\nu^{\text{eff}}$  is the present effective temperature of the massive neutrinos.  $T_\nu^{\text{eff}}$  is related to the temperature of the CMB via  $T_\nu^{\text{eff}} = T_{\text{CMB}} T_{\text{NCDM}}$ , where  $T_{\text{NCDM}}$  is a dimensionless factor ( $\simeq 1$ ) used by e.g. CLASS to set the ratio  $\Sigma m_\nu / \Omega_{\nu,m}$  to its experimentally measured value. Note that  $T_{\text{NCDM}}$  is used to modulate the effective temperature of massive neutrinos only; the temperature of

**Table 1.** Cosmologies implemented in CCL and observables supported in each of them. Note that the only reason why angular power spectra appear not to be supported in non-flat cosmologies is that the hyperspherical Bessel functions are currently not implemented, although their impact is fairly limited. Likewise, number counts power spectra are strictly not supported in the presence of massive neutrino cosmologies due to the scale-dependent growth rate that affects the redshift-space distortions term, even though the impact of this is also small for wide tomographic bins. Finally, notice that in addition CCL can make predictions for the growth of perturbations for some modified gravity models through a user defined  $\Delta f(a)$  as detailed in Section 2.2, and that other extensions are supported via integration of external modified gravity codes.

Observable/Model	flat $\Lambda$ CDM	$\Lambda$ CDM+ $K$	$\Lambda$ CDM + $m_\nu$	$w$ CDM	$w_0 + w_a$
Distances	✓	✓	✓	✓	✓
Growth	✓	✓	$X$	✓	✓
$P_m(k, z)$	✓	✓	✓	✓	✓
Halo Mass Function	✓	✓	$X$	✓	✓
$C_l$ , number counts	✓	$X$	$X$	✓	✓
$C_l$ , weak/CMB lensing	✓	$X$	✓	✓	✓
Correlation function	✓	$X$	✓	✓	✓

relativistic neutrinos follows the usual relation. Finally,  $\mu^i$  is a per-species mass-dependent dimensionless constant, given by  $\mu^i = m_\nu^i c^2 a / (k_B T_\nu^{\text{eff}})$  where  $k_B$  is the Boltzmann constant.

The density parameters  $\Omega_X(a)$  of a given species  $X$  at a given time are defined in terms of the physical background densities  $\bar{\rho}_X(a)$  via  $\Omega_X(a) \equiv \rho_{\text{crit}}^{-1}(a) \bar{\rho}_X(a)$ , where the critical density  $\rho_{\text{crit}}(a) = (8\pi G)^{-1} 3c^2 H^2(a) = \rho_{\text{crit},0} H_0^{-2} H^2(a)$ . As an example, the physical density of matter is given by  $\bar{\rho}_m(a) = \bar{\rho}_{m,0} a^{-3} = \rho_{\text{crit},0} \Omega_{m,0} a^{-3}$ , and its density parameter is  $\Omega_m(a) = \Omega_{m,0} H_0^2 a^{-3} H^{-2}(a)$ . CCL moreover allows for comoving physical densities  $\bar{\rho}_{X,\text{com}}(a) = \bar{\rho}_X(a) a^3$  to be extracted, which in the case of matter reduces to  $\bar{\rho}_{m,\text{com}} = \rho_{\text{crit},0} \Omega_{m,0}$ . We include bars for  $\rho_X$  to distinguish from spatially-varying densities in later sections.

Fitting models to cosmological observables requires predicting cosmological distances for a given model. We consider the comoving radial distance, which is calculated via a numerical integral as

$$\chi(a) = c \int_a^1 \frac{da'}{a'^2 H(a')}. \quad (4)$$

The comoving angular diameter distance is then computed in terms of the comoving radial distance,

$$r(\chi) = \begin{cases} k^{-1/2} \sin(k^{1/2} \chi) & k > 0 \\ \chi & k = 0 \\ |k|^{-1/2} \sinh(|k|^{1/2} \chi) & k < 0 \end{cases} \quad (5)$$

where  $k = (\sum_X \Omega_X - 1) c^{-2} H_0^2$  is the curvature. This function can be written directly in terms of the present

curvature density parameter,  $\Omega_{K,0}$ , as

$$r(\chi) = \sqrt{\Omega_{K,0}} \sinh(\sqrt{\Omega_{K,0}} \chi), \quad (6)$$

which is valid for all forms of the curvature density. The angular diameter distance is given by  $d_A = a r(a)$ , and the luminosity distance is  $d_L = r(a)/a$ , leading to the familiar relation  $d_A = a^2 d_L$  which is valid in general for all metric theories of gravity. The CCL suite also has the functionality to compute the distance modulus, defined as

$$\mu = 5 \log_{10}(d_L/\text{pc}) - 5, \quad (7)$$

along with  $a(\chi)$ , the inverse function of  $\chi(a)$ .

## 2.2. Growth of perturbations

**Once the 2 questions below are addressed this Sub-section is ready for review.**

Probing the growth history of the Universe can allow us to distinguish between cosmological models. To compute the linear growth factor of matter perturbations,  $D(a)$ , CCL solves the following differential equation:

$$\frac{d}{da} \left( a^3 H(a) \frac{dD}{da} \right) = \frac{3}{2} \Omega_M(a) a H(a) D, \quad (8)$$

using a Runge-Kutta Cash-Karp algorithm.

In doing this, CCL simultaneously computes the logarithmic growth rate  $f(a)$ , defined as:

$$f(a) \equiv \frac{d \ln D}{d \ln a}. \quad (9)$$

CCL provides functions that return the growth normalized to  $D(a=1) = 1$  and to  $D(a \ll 1) \rightarrow a$ . It employs an accelerated spline that is linearly spaced in the scale factor to interpolate the growth functions.



The growth calculations cover flat and curved  $\Lambda$ CDM,  $w$ CDM, and  $(w_0 + w_a)$ CDM cosmologies. However, it should be noted that the above treatment is ill-defined in the presence of massive neutrinos, and attempts to compute the growth rate in cosmologies with massive neutrinos will get an inconsistent result.

Finally, CCL allows for an alternative ‘modified gravity’ cosmological model defined by a regular curved  $(w_0 + w_a)$ CDM background as well as a user-defined  $\Delta f(a)$ , such that the true growth rate in this model is given by  $f(a) = f_0(a) + \Delta f(a)$ , where  $f_0(a)$  is the growth rate in the background model. Note that this model is only consistently implemented with regards to the computation of the linear growth factor and growth rates. All other CCL functions (including the non-linear power spectrum) will ignore these modifications. This model, and the interpretation of the predictions given by CCL, should therefore be used with care. A version of the growth code and transfer functions depending on modified gravity parameters is under development.

### 2.3. Matter power spectrum

Theoretical predictions for cosmological observables such as galaxy clustering, gravitational lensing and cluster mass functions rely on knowledge of the distribution of matter from small to large scales in the Universe. The quantity most frequently used to describe the distribution of matter at a given wavenumber and redshift is the matter power spectrum,  $P(k, z)$ , defined as

$$\langle \tilde{\delta}(\mathbf{k}, z) \tilde{\delta}(\mathbf{k}', z) \rangle = (2\pi)^3 P(k, z) \delta_D^3(\mathbf{k} - \mathbf{k}') \quad (10)$$

where  $\tilde{\delta}(\mathbf{k})$  is the Fourier component of the overdensity field at a given wavenumber and  $\delta_D^3$  is the Dirac delta function.  $P(k, z)$  has units of volume and a dimensionless analogue is often defined as

$$\Delta^2(k, z) \equiv \frac{k^3}{2\pi^2} P(k, z). \quad (11)$$

CCL implements several different methods for making predictions for the matter power spectrum. Two of those methods, the BBKS (Bardeen et al. 1986) and Eisenstein & Hu (1998) approximations, are only accurate to within a few percent and are implemented for validation purposes only. These approximations provide analytical expressions for the transfer function,  $T(k)$ , which is related to the matter power spectrum by  $\Delta(k)^2 \propto T^2(k) k^{3+n_s}$ , where  $n_s$  is the power-law index of the primordial power spectrum. The normalization of the power spectrum is either defined at  $z = 0$  by setting  $\sigma_8$  to its value today, or by setting the amplitude of primordial fluctuations,  $A_s$ .

The default CCL implementation uses the CLASS algorithm Blas et al. (2011) to obtain predictions for  $P(k, z)$ .

In addition, CCL can also generate  $P(k, z)$  predictions by emulation of cosmological numerical simulations, in particular, using the ‘Cosmic Emulator’ developed by Lawrence et al. (2017).

None of the above methods account for the impact of baryonic physics on the distribution of matter, which is known to exceed the percent level at scales  $k \gtrsim 1/\text{Mpc}$  (van Daalen et al. 2011; Vogelsberger et al. 2014; Hellwing et al. 2016; Springel et al. 2017) and can affect the extraction of cosmological parameters (Semboloni et al. 2011, 2013; Mohammed & Seljak 2014; Eifler et al. 2015; Mohammed & Gnedin 2017). To account for this effect, we incorporate in CCL an effective parametrization (Schneider & Teyssier 2015) of the redistribution of matter as a consequence of feedback from Active Galactic Nuclei and adiabatic cooling. We give an overview of each method to predict the matter power spectrum in what follows.

**BBKS approximation.** CCL implements the analytical BBKS approximation to the transfer function (Bardeen et al. 1986), given by

$$T(q \equiv k/\Gamma h \text{Mpc}^{-1}) = \frac{\ln[1 + 2.34q]}{2.34q} \times [1 + 3.89q + (16.2q)^2 + (5.47q)^3 + (6.71q)^4]^{-0.25} \quad (12)$$

where  $\Gamma = \Omega_m h$ . The BBKS power spectrum option is primarily used as a precisely-defined input for testing the numerical accuracy of CCL routines (as described in Sec. 3), and it is not recommended for other uses.

**Eisenstein & Hu approximation.** CCL also provides an approximation to the matter power spectrum as implemented by Eisenstein & Hu (1998) (we refer the reader to this paper for a detailed discussion of the fitting formulae).<sup>2</sup>

**CLASS.** The default configuration of CCL adopts predictions for the linear and nonlinear matter power spectrum from the publicly available software (Blas et al. 2011). CLASS currently computes the non-linear power spectrum using the HaloFit prescription of Takahashi et al. (2012).

**Cosmic emulator.** To do: In the section on the HMF and halo bias below, there is a lot of emphasis on the fact that these quantities are not themselves without significant uncertainty. Should we include a sentence or two to that effect here as well? - DL JAB: my general comment in the abstract is related to this. TM: I think we should avoid all discussions of physical accu-

<sup>2</sup> Note that the implementation in CCL modifies Eq. 5 of Eisenstein & Hu (1998) using  $a^{-1} = 1 + z$  instead of the approximation  $a^{-1} \sim z$ . The difference in the resulting power spectra is negligible, but larger than 1 part in  $10^4$  for  $k < 10 h \text{Mpc}^{-1}$ .

racy and near the top say “all models presented in the paper (P, HMF, bias, etc.) are themselves physically uncertain. The reader should consult those papers to learn about the physical accuracy.” The emulator (Lawrence et al. 2017) provides accurate predictions for the non-linear matter power spectrum, at the 1% level for  $z \leq 2$  and in the wavenumber range  $k = [10^{-3}, 5] \text{ Mpc}^{-1}$ . The allowed range of cosmological parameters that can be passed to the emulator is as follows:

$$\begin{aligned} 0.12 &\leq \Omega_m h^2 \leq 0.155, \\ 0.0215 &\leq \Omega_b h^2 \leq 0.0235, \\ 0.7 &\leq \sigma_8 \leq 0.9, \\ 0.55 &\leq h \leq 0.85, \\ 0.85 &\leq n_s \leq 1.05, \\ -1.3 &\leq w_0 \leq -0.7, \\ -1.73 &\leq w_a \leq -0.7,^3 \\ 0.0 &\leq \Omega_{\nu, \text{m}} h^2 \leq 0.01. \end{aligned} \quad (13)$$

In the case of the emulator, the effective number of relativistic neutrino species is set to  $N_{\text{eff}} = 3.04$ . In Lawrence et al. (2017), the neutrino component of the power spectrum is not simulated, but either linearly evolved and added to the simulated power spectra at low redshift, or accounted for by a scale-dependent correction to the growth function. More details on this method and its accuracy can be found in (Upadhye et al. 2014; Castorina et al. 2015; Heitmann et al. 2016).

**Baryonic correction model (BCM).** CCL incorporates the impact of baryons on the total matter power spectrum via the BCM of Schneider & Teyssier (2015). The main consequences of baryonic processes are: to suppress the power spectrum at intermediate scales ( $k \sim$  a few  $h/\text{Mpc}$ ) due to the ejection of gas by Active Galactic Nuclei feedback, and to enhance it at smaller scales due to adiabatic cooling. To account for these effects, BCM uses an effective decomposition for the impact of gas ejection ( $G$ ) and the enhancement of the small scale profile due to star formation ( $S$ ) to estimate the fractional effect of baryonic processes on the dark matter-only power spectrum ( $P_{\text{DMO}}$ ):

$$P_{\text{BCM}}(k, z) = P_{\text{DMO}}(k, z) G(k|M_c, \eta_b, z) S(k|k_s) \quad (14)$$

Three effective parameters govern the contribution of baryonic processes to modifying the total matter power spectrum:

- $\log_{10}[M_c/(M_\odot/h)]$ : the mass of the clusters responsible for feedback, which regulates the amount of suppression of the matter power spectrum at intermediate scales;

- $\eta_b$ : a dimensionless parameter which determines the scale at which suppression peaks;
- and  $k_s [h/\text{Mpc}]$ : the wavenumber that determines the scale of the stellar distribution of matter in the center of halos.

If these parameters are not specified by the user, CCL assumes the default parameters of Schneider & Teyssier (2015), calibrated through different comparisons with observations and simulations in that work.

#### 2.4. Two-point correlators

This section describes the theoretical expectations for the two-point correlators of pairs of quantities (fields) defined on the sky. These fields can be classified in terms of their spin  $s$  under rotations on the plane tangent to the sphere. In general a spin- $s$  field is defined by two real-valued functions of the spherical coordinates  $a_1(\hat{\mathbf{n}})$  and  $a_2(\hat{\mathbf{n}})$  (e.g.  $\gamma_1$  and  $\gamma_2$  for weak lensing or the Stokes parameters  $Q$  and  $U$  in the case of polarized intensity), from which one can form the complex field  $a = a_1 + ia_2$ .

Spin- $s$  quantities can be decomposed into their harmonic coefficients  ${}_s a_{\ell m}$  through a spherical harmonic transform:

$${}_s a_{\ell m} = \int d\hat{\mathbf{n}} a(\hat{\mathbf{n}}) {}_s Y_{\ell m}^*(\hat{\mathbf{n}}), \quad a(\hat{\mathbf{n}}) = \sum_{\ell m} {}_s a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}})$$

where  ${}_s Y_{\ell m}$  are the spin-weighted spherical harmonics. The harmonic coefficients can then be associated with parity-even and parity-odd components ( $E$ -modes and  $B$ -modes respectively) as<sup>4</sup>

$$\begin{aligned} E_{\ell m} &= -\frac{1}{2} [{}_s a_{\ell m} + (-1)^s {}_{-s} a_{\ell m}] \\ iB_{\ell m} &= -\frac{1}{2} [{}_s a_{\ell m} - (-1)^s {}_{-s} a_{\ell m}] \end{aligned}$$

In what follows we will focus on scalar ( $s = 0$ ) quantities such as the overdensity of source number counts or the lensing convergence, and on spin-2 fields such as the lensing shear. We will also distinguish between *tracers* (fields observed on the sky, such as number counts in a redshift bin, shear, or CMB temperature fluctuations) and *contributions* to the total observed fluctuations of these tracers (such as the biased matter density term in number counts, redshift-space distortions, magnification, etc.).

##### 2.4.1. Angular power spectra

<sup>4</sup> We note that for spin-0 quantities the minus sign preceding these equations is usually omitted.

The angular power spectrum  $C_\ell^{ab}$  between two tracers  $a$  and  $b$  is defined as

$$\langle a_{\ell m} b_{\ell m}^* \rangle \equiv C_\ell^{ab} \delta_{\ell\ell'} \delta_{mm'}, \quad (15)$$

where  $a_{\ell m}$  and  $b_{\ell m}$  can be either the  $E$ -mode or  $B$ -mode component of the corresponding field. In what follows we will only deal with fields for which the  $B$ -modes are identically 0, and therefore all equations refer to the  $E$ - $E$  power spectrum. In general, this power spectrum can be written as:

$$C_\ell^{ab} = 4\pi \int_0^\infty \frac{dk}{k} \mathcal{P}_\Phi(k) \Delta_\ell^a(k) \Delta_\ell^b(k), \quad (16)$$

where  $\mathcal{P}_\Phi(k)$  is the dimensionless power spectrum of the primordial curvature perturbations, and  $\Delta^a$  and  $\Delta^b$  are the transfer functions corresponding to these tracers. Each transfer function will receive contributions from different terms. Currently CCL supports three types of tracers: number counts, galaxy shape distortions and lensing convergence, with the following contributions:

**Number counts.**—The transfer function for number counts can be decomposed into three contributions:  $\Delta^{\text{NC}} = \Delta^{\text{D}} + \Delta^{\text{RSD}} + \Delta^{\text{M}}$ , where

- $\Delta^{\text{D}}$  is the standard density term proportional to the matter density:

$$\Delta_\ell^{\text{D}}(k) = \int dz p_z(z) b(z) T_\delta(k, z) j_\ell(k\chi(z)), \quad (17)$$

where  $j_\ell(x)$  is  $\ell$ -th order spherical Bessel function,  $T_\delta$  is the matter transfer function,  $b(z)$  is the linear clustering bias for this tracer and  $p_z(z)$  is the normalized distribution of sources in redshift (i.e. the selection function). The fluctuations in the number density of sources in different redshift bins are therefore treated by CCL as different tracers. Note that CCL currently does not support non-linear or scale-dependent bias, but future releases will do so under a number of schemes, including perturbative approaches as implemented in e.g. [McEwen et al. \(2016\)](#).

- $\Delta^{\text{RSD}}$  is the linear contribution from redshift-space distortions (RSDs):

$$\Delta_\ell^{\text{RSD}}(k) = \int dz \frac{(1+z)p_z(z)}{H(z)} T_\theta(k, z) j_\ell''(k\chi(z)), \quad (18)$$

where  $T_\theta(k, z)$  is the transfer function of  $\theta$ , the divergence of the comoving velocity field, and  $j_\ell''$  is the second order derivative of the spherical Bessel function,  $j_\ell$ . Note that the RSD contribution to

number counts is currently computed by CCL assuming a linear-theory relation between the matter overdensity and peculiar velocity fields, mediated by the scale-independent growth rate  $f$ . While this should not be problematic for wide photometric redshift bins and standard cosmological models, users should exercise care when interpreting results for narrow window functions or exotic cosmologies. Additionally, number count tracers with RSD in cosmologies with massive neutrinos are not currently supported.

- $\Delta^{\text{M}}$  is the contribution from lensing magnification:

$$\Delta_\ell^{\text{M}}(k) = -\ell(\ell+1) \int \frac{dz}{H(z)} W^{\text{M}}(z) T_{\phi+\psi}(k, z) j_\ell(k\chi(z)), \quad (19)$$

where  $T_{\phi+\psi}$  is the transfer function for the Newtonian-gauge scalar metric perturbations, and  $W^{\text{M}}$  is the magnification window function:

$$W^{\text{M}}(z) \equiv \int_z^\infty dz' p_z(z') \frac{2-5s(z')}{2} \frac{r(\chi' - \chi)}{r(\chi')r(\chi)}. \quad (20)$$

Here,  $s(z)$  is the magnification bias, given as the logarithmic derivative of the number of sources with magnitude limit, and  $r(\chi)$  is the angular comoving distance (see Eq. 5).

Note that CCL currently does not compute relativistic corrections to number counts ([Challinor & Lewis 2011](#); [Bonvin & Durrer 2011](#)). Although these will be included in the future, their contribution to the total fluctuation is largely subdominant (see [Alonso et al. \(2015\)](#) and the two references above), and therefore it is safe to ignore them for our purposes.

**Correlated galaxy shapes.**—The transfer function for correlated galaxy shapes (intrinsic and lensed) is currently decomposed into two terms:  $\Delta^{\text{SH}} = \Delta^{\text{WL}} + \Delta^{\text{IA}}$ , where

- $\Delta^{\text{L}}$  is the standard lensing (“cosmic shear”) contribution:

$$\Delta_\ell^{\text{L}}(k) = -\frac{1}{2} \sqrt{\frac{(\ell+2)!}{(\ell-2)!}} \int \frac{dz}{H(z)} W^{\text{L}}(z) T_{\phi+\psi}(k, z) j_\ell(k\chi(z)), \quad (21)$$

where  $W^{\text{L}}$  is the lensing kernel, given by

$$W^{\text{L}}(z) \equiv \int_z^\infty dz' p_z(z') \frac{r(\chi' - \chi)}{r(\chi')r(\chi)}. \quad (22)$$



- $\Delta^{\text{IA}}$  is the transfer function for intrinsic galaxy alignments. CCL currently supports the so-called “non-linear alignment model”, in which the intrinsic galaxy inertia tensor is proportional the local tidal tensor [Hirata & Seljak \(2004\)](#); [Hirata et al. \(2007\)](#):

$$\Delta_\ell^{\text{IA}}(k) = \sqrt{\frac{(\ell+2)!}{(\ell-2)!}} \int dz p_z(z) b_{\text{IA}}(z) f_{\text{IA}}(z) T_\delta(k, z) \frac{j_\ell(k\chi(z))}{[k\chi(z)]^2}. \quad (23)$$

Here,  $b_{\text{IA}}$  is the so-called alignment bias, and  $f_{\text{IA}}$  is the fraction of aligned galaxies in the sample. Notice that  $b_{\text{IA}}(z)$  absorbs the typical normalization factors used in the literature for intrinsic alignment amplitude and redshift evolution. It is thus not to be confused with  $C_1$  or  $A_{\text{IA}}$ , typical parameters for the linear alignment model, as adopted in works such as [Blazek et al. \(2017\)](#); [van Uitert et al. \(2018\)](#); [Joudaki et al. \(2018\)](#); [Hildebrandt et al. \(2017\)](#).

**Lensing convergence.**—The transfer function for the lensing convergence of a given source plane at redshift  $z_*$  receives only one contribution, given by

$$\Delta_\ell^\kappa(k) = -\frac{\ell(\ell+1)}{2} \int_0^{\chi_*} \frac{dz}{H(z)} \frac{r(\chi_* - \chi)}{r(\chi)r(\chi_*)} T_{\phi+\psi}(k, z), \quad (24)$$

where  $\chi_* \equiv \chi(z_*)$ .

It is worth noting that the equations above should be modified for non-flat cosmologies by replacing the spherical Bessel functions  $j_\ell$  with their hyperspherical counterparts ([Kamionkowski & Spergel 1994](#)). These are currently not supported by CCL, and their impact is mostly relevant on low multipoles. The library also assumes a factorizable matter power spectrum at unequal times  $P_\delta(k, z_1, z_2) = T_\delta(k, z_1)T_\delta(k, z_2) 2\pi^2 \mathcal{P}_\Phi(k)$  (see [Kitching & Heavens 2017](#)). Furthermore, CCL currently assumes a relation between transfer functions  $T_\delta$ ,  $T_\theta$  and  $T_{\phi+\psi}$  that is strictly only valid in vanilla  $\Lambda\text{CDM}$ <sup>5</sup>:

$$T_\delta = -\frac{1+z}{H(z)f(z)} T_\theta = -\frac{k^2}{3H_0^2\Omega_M} \frac{T_{\phi+\psi}}{1+z}. \quad (25)$$

These approximations will be revisited in future versions of the library.

#### 2.4.2. Correlation functions

<sup>5</sup> Note that the transfer functions are defined here for the full non-linear density field, as opposed to the more common linear transfer functions.

Fields can also be correlated in configuration space, and the corresponding correlators are called correlation functions. Let  $a$  and  $b$  be two fields with spins  $s_a$  and  $s_b$ . We start by defining  $\tilde{a}(\hat{\mathbf{n}}_1)$  and  $\tilde{b}(\hat{\mathbf{n}}_2)$  as the fields  $a$  and  $b$  rotated such that the  $x$ -axis of the tangential coordinate systems at  $\hat{\mathbf{n}}_1$  and  $\hat{\mathbf{n}}_2$  become aligned with the vector connecting both points. We can then define two correlation functions:

$$\xi_+^{ab}(\theta) \equiv \langle \tilde{a}(\hat{\mathbf{n}}_1) \tilde{b}^*(\hat{\mathbf{n}}_2) \rangle, \quad \xi_-^{ab}(\theta) \equiv \langle \tilde{a}(\hat{\mathbf{n}}_1) \tilde{b}(\hat{\mathbf{n}}_2) \rangle,$$

where  $\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 \equiv \cos \theta$ .

$\xi_\pm$  can be related to the power spectra as

$$\xi_\pm^{ab} = \sum_\ell \frac{2\ell+1}{4\pi} (\pm 1)^{s_b} C_\ell^{ab\pm} d_{s_a, \pm s_b}^\ell(\theta), \quad (26)$$

where  $d_{mm'}^\ell$  are the Wigner- $d$  matrices ([Ng & Liu 1999](#); [Chon et al. 2004](#)) and we have defined the power spectra

$$C_\ell^{ab\pm} \equiv \left( C_\ell^{aEbE} \pm C_\ell^{aBbB} \right) + i \left( C_\ell^{aBbE} \pm C_\ell^{aEbB} \right), \quad (27)$$

which reduces to the  $EE$  power spectrum when all  $B$ -modes are 0.

Note that, as scalar quantities are real, any correlation involving at least one spin-0 field only has one non-unique correlation function. In these cases, the Wigner- $d$  matrices can also be expressed in terms of associated Legendre polynomials  $P_\ell^m$ , and therefore Eq. 26 becomes

$$\xi^{ab}(\theta) = \sum_\ell \frac{2\ell+1}{4\pi} C_\ell^{ab} \sqrt{\frac{(\ell-s_a)!}{(\ell+s_a)!}} P_\ell^{s_a}(\cos \theta) \quad (28)$$

$$= \sum_\ell \frac{2\ell+1}{4\pi} C_\ell^{ab} P_\ell(\cos \theta), \quad (29)$$

where the last equality holds for  $s_a = 0$ , and both expressions assume  $s_b = 0$ .

In the flat-sky approximation we can take the small-scale limit  $\ell \gg s_1, s_2$  and approximate

$$d_{s_1 s_2}^\ell(\theta) \longrightarrow J_{s_1-s_2}(\ell\theta), \quad (30)$$

where  $J_\alpha(x)$  is the Bessel function of order  $\alpha$ . Eq. 26 then becomes<sup>6</sup>

$$\xi_\pm^{ab}(\theta) = (\pm 1)^{s_b} \int \frac{d\ell \ell}{2\pi} C_\ell^{ab\pm} J_{s_a \mp s_b}(\ell\theta). \quad (31)$$

In summary, for spins 0 and 2, the three relevant cases for the cosmological observables supported by CCL are:

<sup>6</sup> See the weak lensing review by [Bartelmann & Schneider \(2001\)](#), page 44 and [Joachimi & Bridle \(2010\)](#).

- $s_a = s_b = 0$  (e.g. galaxy-galaxy, galaxy- $\kappa$  and  $\kappa$ - $\kappa$ ):

$$\xi^{ab}(\theta) = \sum_{\ell} \frac{2\ell+1}{4\pi} C_{\ell}^{ab} P_{\ell}(\cos \theta) \quad (\text{full-sky}) \quad (32)$$

$$= \int_0^{\infty} \frac{d\ell}{2\pi} C_{\ell}^{ab} J_0(\ell\theta) \quad (\text{flat-sky}) \quad (33)$$

- $s_a = 2, s_b = 0$  (e.g. galaxy-shear,  $\kappa$ -shear):

$$\xi^{ab}(\theta) = \sum_{\ell} \frac{2\ell+1}{4\pi} C_{\ell}^{ab} d_{2,0}^{\ell}(\theta) \quad (\text{full-sky}) \quad (34)$$

$$= \int_0^{\infty} \frac{d\ell}{2\pi} C_{\ell}^{ab} J_2(\ell\theta) \quad (\text{flat-sky}) \quad (35)$$

- $s_a = s_b = 2$  (e.g. shear-shear):

$$\xi_{\pm}^{ab}(\theta) = \sum_{\ell} \frac{2\ell+1}{4\pi} C_{\ell}^{ab} d_{2,\pm 2}^{\ell}(\theta) \quad (\text{full-sky}) \quad (36)$$

$$= \int_0^{\infty} \frac{d\ell}{2\pi} C_{\ell}^{ab} J_{2\mp 2}(\ell\theta) \quad (\text{flat-sky}) \quad (37)$$

#### 2.4.3. 3-dimensional spatial correlation function.

In addition to the angular correlation functions, **CCL** can also be used to compute the three-dimensional spatial correlation function,  $\xi(r)$ , from the transform of the matter power spectrum:

$$\xi(r) = \frac{1}{2\pi^2} \int dk k^2 P(k) \frac{\sin(kr)}{kr} \quad (38)$$

In the future **CCL** will be expanded to incorporate the calculation of the higher-order multipoles needed to characterize the redshift-space three-dimensional correlation function in the presence of RSDs.

#### 2.5. Halo mass function

Being able to calculate the halo abundance as a function of mass is a necessary step to being able to constrain cosmology with probes such as galaxy clusters. While analytic functions are the traditional means of predicting evolution of halo abundances, calibration is frequently required against cosmological simulations. In order to reach the high precision required for cosmological constraints in a self-consistent fashion, it is ultimately necessary to use cosmological simulations; we implement this with halo mass functions with parameters fit to these simulations. The calculation of the halo mass function focuses around the spherical overdensity method of halo finding, in which the size of a halo can be defined with:

$$\bar{\rho}(r_{\Delta}) = \Delta \times \bar{\rho}_m, \quad (39)$$

where a spherical halo with radius  $r_{\Delta}$  has an average density  $\bar{\rho}$  equal to the overdensity parameter  $\Delta$  times the mean background density of the universe at a given redshift,  $\bar{\rho}_m(z)$ . Within the literature, the choice of  $\Delta$  can vary considerably, as observations focusing on the compact cores of halos often take much larger values of  $\Delta$  than the fiducial definition in most halo clustering studies,  $\Delta = 200$ . We note that another common definition exists which utilizes the critical density of the universe,  $\rho_{\text{crit}}$ ; this introduces a simple conversion factor between the two definitions that must be accounted for. Currently, **CCL** only accepts overdensity parameters with respect to the mean matter density, but we plan to allow for self-consistent handling of critical density based definitions in the future.

The halo mass function is defined as

$$\frac{dn}{dM} = f(\sigma) \frac{\bar{\rho}_m}{M} \frac{d \ln \sigma^{-1}}{dM}, \quad (40)$$

where  $n$  is the number density of halos of a given mass  $M$  associated with the RMS variance of the matter density field  $\sigma^2$  at a given redshift and  $f$  is the derived fitting function<sup>7</sup>. In **CCL** calling this function returns the mass function in logarithmic mass bins,  $dn/d \log_{10} M$ , where the input is the halo mass  $M$  and scale factor  $a$ .

The halo mass  $M$  is related to  $\sigma$  by first computing the radius  $R$  that would enclose a mass  $M$  in a homogeneous Universe at  $z = 0$ :

$$M = \frac{H_0^2}{2G} R^3 \rightarrow \frac{M}{M_{\odot}} = 1.162 \times 10^{12} \Omega_M h^2 \left( \frac{R}{1 \text{ Mpc}} \right)^3. \quad (41)$$

The RMS density contrast in spheres of radius  $R$  can then be computed as

$$\sigma_R^2 = \frac{1}{2\pi^2} \int dk k^2 P(k) \tilde{W}_R^2(k) \quad (42)$$

where  $P(k)$  is the linear matter power spectrum and  $\tilde{W}(kR)$  is the Fourier transform of a spherical top hat window function,

$$\tilde{W}_R(k) = \frac{3}{(kR)^3} [\sin(kR) - kR \cos(kR)]. \quad (43)$$

This is commonly related in terms of the mass inside of the Lagrangian scale of the halo, using the following transformation:

$$R = (3M/4\pi\bar{\rho}_m)^{1/3}. \quad (44)$$

<sup>7</sup> Not to be confused with the linear growth rate of structure defined in Eq. 9

As a consequence, one can also define  $\sigma_M$  as the RMS variance of the density field smoothed on some scale  $M$ , analogously to Eq.(42).

One commonly used halo mass function definition within the literature is the Tinker et al. (2010) fitting function. This fitting function has been developed using collisionless  $N$ -body simulation data, using halos identified by spherical overdensities. This is an extension of the Tinker et al. (2008) halo mass function, which is also included within CCL as a comparative option. This fitting function assumes little change with respect to cosmological parameters. Further, it includes a redshift scaling which is assumed to sharply end at a redshift of  $z = 3$ . This halo mass function is calibrated within the range of  $10^{10.5}hM_\odot \leq M \leq 10^{15.5}hM_\odot$  at a redshift of  $z = 0$ .

For comparison purposes, we also have included the results of Angulo et al. (2012), which uses the Millenium XXL simulation in order to study galaxy cluster scaling relations. As part of this study, they have calculated their own fit to the Tinker et al. (2010) fitting function. While this additional halo mass function is available, it has not been extended to a broad range of overdensity parameter  $\Delta$ , nor has it been extended beyond a redshift of  $z = 0$ .

The Tinker et al. (2008) fitting function uses the following parameterisation:

$$f(\sigma) = A \left[ \left( \frac{\sigma}{b} \right)^{-a} + 1 \right] e^{-c/\sigma^2}, \quad (45)$$

where  $A$ ,  $a$ ,  $b$ , and  $c$  are fitting parameters that have additional redshift scaling. This basic form is modified for the Angulo et al. (2012) formulation. The resulting form is

$$f(\sigma) = A \left[ \left( \frac{b}{\sigma} + 1 \right)^{-a} \right] e^{-c/\sigma^2}, \quad (46)$$

where the only change is in the formulation of the second term. Note that the fitting parameters in the Angulo et al. (2012) formulation do not contain any redshift dependence and the use of it is primarily for testing and benchmark purposes.

The Tinker et al. (2010) model parameterizes the halo mass function in terms of the peak height,  $\nu = \delta_c/\sigma_M$ , where  $\delta_c = 1.686$  is the critical density for collapse. The function is then re-expressed as

$$f(\nu) = \alpha [1 + (\beta\nu)^{-2\phi}] \nu^{2\eta} e^{(-\gamma\nu^2/2)}. \quad (47)$$

We note that these halo mass functions, while implemented to high *numerical* accuracy through comparison against a flat  $\Lambda$ CDM cosmology (see ‘CCL1’ in Table

3), carry their own uncertainties. It has not been significantly studied whether the halo mass function is universal with respect to changes in dark energy parameterisation. Tinker et al. (2008, 2010) quote 5% accuracy of their mass functions. This result is consistent with the work of Watson et al. (2013), which also finds a 5% level difference in comparison to the Tinker fitting function. Further study will be required in the future in order to gain percent level accuracy in determining the halo mass function. **To do: For what cosmologies was the Tinker mass function tested? I.e., what should go in Table 1?**

## 2.6. Halo bias

JAB: I’m a bit confused by this section. This is a quantity that CCL can return. But currently the clustering observables can only handle linear (scale-independent) bias? And should we say that the user can specify a linear bias that does not come from this fitting formula? ASV: Including the Tinker fitting function for the halo bias is a fairly natural step considering we are also including the Tinker mass function (on which this is heavily related to). I do not think the fact that the clustering section of the code is currently unready to handle more than linear bias impacts the usefulness of having this built in. That being said, I will specify that currently we do not have a linear bias model built in as an explicit framework, though it can certainly be passed in to other sections of the code.

An important step in many interpretations of the halo model is to have a measure of the bias of dark matter halos, defined as the ratio of the halo power spectrum to the linear dark matter power spectrum,

$$b^2(k) = \frac{P_h(k)}{P_{lin}(k)}. \quad (48)$$

As with measures of the halo mass function, high accuracy cosmological constraints requires the use of numerical simulations to develop fitting functions and emulators. We note that we will define haloes as in the above subsection focusing on the halo mass function. CCL currently implements the halo bias fitting function results in Tinker et al. (2010), though future improvements will likely require the use of emulator methods.

The Tinker et al. (2010) model parameterizes the halo mass function and the halo bias in terms of the peak height and the critical density for collapse as

$$b(\nu) = 1 - A \frac{\nu^a}{\nu^a + \delta_c^a} + B\nu^b + C\nu^c, \quad (49)$$

$$f(\nu) = \alpha [1 + (\beta\nu)^{-2\phi}] \nu^{2\eta} e^{(-\gamma\nu^2/2)}. \quad (50)$$

Again, while high *numerical* accuracy has been verified, there is a remaining uncertainty. Tinker et al.

(2010) found a  $\sim 6\%$  scatter when determining the halo bias due to differences in simulations alone. In addition, this parameterization does not include a careful exploration of any impact due to changes in the dark energy equation of state. As with the halo mass function, studies will be required to reach accuracy at the percent level for any cosmological predictions.

We note that many observables are currently tuned to utilizing a linear halo bias (e.g. scale independent). While the current functionality of CCL is to return the Tinker et al. (2010) model for halo bias, it is entirely possible to generate simpler models of halo bias and pass this into further functions.

### 2.7. Photometric redshifts

Redshifts of LSST galaxies will be obtained via photometry. Therefore, performing any cosmological analysis which incorporates redshift information requires a model for the probability of measuring a photometric redshift  $z_{\text{ph}}$  for an object with hypothetical spectroscopic redshift  $z_s$ . In order to maintain agnosticism towards the optimal model, and hence to allow for the future inclusion of advancements from ongoing research, CCL allows the user to flexibly input a photometric redshift model. In addition, for ease of use, CCL provides the option of using a built-in function for a simple Gaussian photometric redshift probability distribution. The photometric redshift model can then be used, for example, when computing  $\frac{dN}{dz}^i$  in photometric redshift bin  $i$ , as given by:

$$\frac{dN^i}{dz} = \frac{\frac{dN}{dz} \int_{z_i}^{z_{i+1}} dz' p(z, z')}{\int_{z_{\min}}^{z_{\max}} dz \frac{dN}{dz} \int_{z_i}^{z_{i+1}} dz' p(z, z')} \quad (51)$$

where  $p(z, z')$  is the photometric redshift probability distribution, and  $z_i$  and  $z_{i+1}$  are the photo- $z$  edges of the bin in question. In the case of the simple Gaussian photometric redshift model for which support is included in CCL out of the box,  $p(z, z')$  is given as

$$p(z, z') = \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{(z - z')^2}{2\sigma_z^2}\right). \quad (52)$$

## 3. IMPLEMENTATION OF HIGH-ACCURACY COSMOLOGICAL FUNCTIONS

In this section, we note some of the assumptions and implementation details that are relevant when making accurate cosmological predictions. The validation tests performed for CCL are described in detail in Section 4.

### 3.1. Background functions & growth of perturbations

**Once we agree on the last paragraph, this Subsection is ready for review.**

Cosmological predictions require making assumptions on the values of several physical constants, as defined in the previous sections. We have performed a comparison of the physical constants used in CCL to those used in GSL and CLASS as well as published sources such as the NIST Handbook and PDG Review of Particle Physics (Beringer et al. 2012). Where possible, we have set constants to the values that are used internally in CLASS. This includes the value of the gravitational constant, the Boltzmann constant, the Planck constant, the speed of light, and the electron charge. CLASS does not include a definition of the solar mass or the Stefan-Boltzmann constant so we use the values used by GSL. After comparison between the physical constants used in CCL and those of the sources mentioned above, we have found better than  $10^{-4}$  agreement with everything except the gravitational constant.

The value of the gravitational constant,  $G$ , enters into the critical density. We found that failure to define  $G$  with sufficient precision initially resulted in lack of convergence at the  $10^{-4}$  level between the CCL outputs and benchmarks (and among benchmarks). Importantly, note that CAMB barely has  $10^{-4}$  precision in  $G$  (and similarly, there might be other constants within CAMB/CLASS for which one should check the precision level). For CCL predictions and benchmarks, we are using the value from CLASS. **To do: Was this the  $G$  we used for benchmarks?**

### 3.2. Matter power spectrum

For speed, the initialization of a cosmological model within CCL performs initial computations of the linear and nonlinear matter power spectra, which are then interpolated whenever required. The spline is performed in two variables: the logarithmically-spaced wavenumber and the logarithmically and linearly-spaced scale factor. Splining the power spectra output leads to some precision loss (compared to, for example, direct outputs from CLASS or the cosmic emulator) which is quantified in Section 4.

We introduce a maximum value  $k$  (in units of  $1/\text{Mpc}$ ) up to which we evaluate the power spectra for interpolation; we call this parameter `K_MAX_SPLINE`. A separate `K_MAX` parameter sets the limit of evaluation of the matter power spectrum. The range between `K_MAX_SPLINE`  $< k < \text{K\_MAX}$  is evaluated by performing a second order Taylor expansion in  $\ln k$ .

The Taylor expansion is implemented as follows: first, we compute the first and second derivative of  $\ln P(k, z)$  at  $k_0 = \text{K\_MAX} - 2\Delta \ln k$  via finite difference derivatives using GSL. The fiducial choice for  $\Delta \ln k$  is  $10^{-2}$ . We then apply a second order Taylor expansion to extrapolate

the matter power spectrum to  $k > \text{K\_MAX\_SPLINE}$ . The Taylor expansion gives

$$\ln P(k, z) \simeq \ln P(k_0, z) + \frac{d \ln P}{d \ln k} (\ln k_0, z) (\ln k - \ln k_0) + \frac{1}{2} \frac{d^2 \ln P}{d \ln k^2} (\ln k_0, z) (\ln k - \ln k_0)^2. \quad (53)$$

We also extrapolate the power spectrum at small wavenumbers. In this case, we introduce the parameter  $\text{K\_MIN}$ , the wavenumber below which the power spectra are obtained by a power-law extrapolation with index  $n_s$ :

$$\log P(k < \text{K\_MIN}, z) = \log P(\text{K\_MIN}, z) + n_s (\log k - \log \text{K\_MIN}) \quad (54)$$

The value adopted for  $\text{K\_MIN}$  depends on the choice of power spectrum method. Note that an additional parameter,  $\text{K\_MIN\_DEFAULT}$ , sets the minimum  $k$  for integrations. This is set to  $\text{K\_MIN\_DEFAULT} = 5 \times 10^{-5} / \text{Mpc}$ . For **CLASS** and the nonlinear power spectrum, we adopt  $\text{K\_MIN}$  that coincides with the smallest wavenumber output by **CLASS**,  $\text{K\_MIN} = 7 \times 10^{-6} / \text{Mpc}$ .<sup>8</sup> As a consequence, when using **CCL** with **CLASS**, no extrapolation will occur at low wavenumbers unless the user modified the default values of  $\text{K\_MIN\_DEFAULT}$  and/or  $\text{K\_MIN}$ .

### 3.3. Angular power spectra

Different numerical approaches have been implemented in the library in order to expedite the computation of angular power spectra. We describe these here.

#### 3.3.1. Limber approximation

As shown in Section 2.4.1, computing each transfer function contributing to a given power spectrum involves a radial projection (i.e. an integral over redshift or  $z$  or  $\chi$ ), and thus computing full power spectra consists of a triple integral for each  $\ell$ . This can be computationally intensive, but can be significantly simplified in certain regimes by using the Limber approximation, given by:

$$j_\ell(x) \simeq \sqrt{\frac{\pi}{2\ell+1}} \delta\left(\ell + \frac{1}{2} - x\right). \quad (55)$$

This eliminates the integrals associated with each of the two transfer functions, massively accelerating the calculation.

<sup>8</sup> For BBKS, the power spectrum is computed analytically at all  $k$ , i.e., there is no extrapolation. For the Eisenstein & Hu implementation, the splines of the power spectrum span  $\text{K\_MIN\_DEFAULT} < k < \text{K\_MAX\_SPLINE}$ , so there is only extrapolation at high  $k$ .

Thus for each  $k$  and  $\ell$  we can define a radial distance  $\chi_\ell \equiv (\ell+1/2)/k$ , with corresponding redshift  $z_\ell$ . Substituting this in the expressions presented in Section 2.4.1, the power spectrum can be computed as a single integral:

$$C_\ell^{ab} = \frac{2}{2\ell+1} \int_0^\infty dk P_\delta(k, z_\ell) \tilde{\Delta}_\ell^a(k) \tilde{\Delta}_\ell^b(k). \quad (56)$$

where

$$\tilde{\Delta}_\ell^D(k) = p_z(z_\ell) b(z_\ell) H(z_\ell) \quad (57)$$

$$\tilde{\Delta}_\ell^{\text{RSD}}(k) = \frac{1+8\ell}{(2\ell+1)^2} p_z(z_\ell) f(z_\ell) H(z_\ell) - \quad (58)$$

$$\frac{4}{2\ell+3} \sqrt{\frac{2\ell+1}{2\ell+3}} p_z(z_{\ell+1}) f(z_{\ell+1}) H(z_{\ell+1})$$

$$\tilde{\Delta}_\ell^M(k) = 3\Omega_{M,0} H_0^2 \frac{\ell(\ell+1)}{k^2} \frac{(1+z_\ell)}{\chi_\ell} W^M(z_\ell) \quad (59)$$

$$\tilde{\Delta}_\ell^L(k) = \frac{3}{2} \Omega_{M,0} H_0^2 \sqrt{\frac{(\ell+2)!}{(\ell-2)!}} \frac{1}{k^2} \frac{1+z_\ell}{\chi_\ell} W^L(z_\ell) \quad (60)$$

$$\tilde{\Delta}_\ell^{\text{IA}}(k) = \sqrt{\frac{(\ell+2)!}{(\ell-2)!}} \frac{p_z(z_\ell) b_{\text{IA}}(z_\ell) f_{\text{red}}(z_\ell) H(z_\ell)}{(\ell+1/2)^2}. \quad (61)$$

The Limber approximation works best for wide radial kernels and high  $\ell$ .

#### 3.3.2. Beyond Limber: *Angpow*

**CCL** incorporates natively routines to compute the  $C_\ell^{ab}$  angular power spectra as described above without the Limber approximation. The algorithm performs first the integrals over  $z$  for both tracers, and ends with the  $k$  integral. This computation is much slower than using the Limber approximation, but results in precise angular power spectra at low  $\ell$ , and correct cross-correlations between tracers. The integration of these routines has been tested against the **CLASS** code and recovers the same angular power spectra when precision parameters are set to high values.

However, the computation of the  $C_\ell^{ab}$  without the Limber approximation is too costly in terms of computing time using this method, if one wants to explore extensively a full cosmological parameter space. The aim of the **Angpow** software (Campagne et al. 2017a) is to compute the angular power spectra  $C_\ell^{ab}$  without any Limber numerical approximation in a faster way but still accurately. **CCL** has been linked to the **Angpow** code, which is briefly described here.

The angular power spectrum for two tracers  $C_\ell^{ab}$  is computed in **Angpow** according to the following expression

$$C_\ell^{ab} = \int_0^\infty \int_0^\infty dz dz' p_{z_1}(z_1) p_{z_2}(z') \times \int_0^\infty dk f_\ell(z, k) f_\ell(z', k). \quad (62)$$



The auxiliary function  $f_\ell(z, k)$  can be defined without loss of generality as

$$f_\ell(z, k) \equiv \sqrt{\frac{2}{\pi}} k \sqrt{P_\delta(k, z)} \tilde{\Delta}_\ell(z, k) \quad (63)$$

with  $\tilde{\Delta}_\ell(z, k)$  the function describing the physical processes such as matter density fluctuations, redshift-space distortions as described for instance in references [Durrer \(2008\)](#); [Yoo et al. \(2009\)](#); [Yoo \(2010\)](#); [Challinor & Lewis \(2011\)](#); [Bonvin & Durrer \(2011\)](#). Currently, the **Angpow** version delivered with CCL can only deal with galaxy clustering tracers (no lensing), and this without the magnification lensing term (equation 19). The incorporation of those transfer functions is left for future work, but in principle **Angpow** has the capability to treat them. For now, for galaxy clustering tracers we define  $\tilde{\Delta}_\ell(z, k)$  as

$$\tilde{\Delta}_\ell(z, k) \approx b(z)j_\ell(k\chi(z)) - f(z)j_\ell''(k\chi(z)) \quad (64)$$

with  $j_\ell(x)$  and  $j_\ell''(x)$  the spherical Bessel function of order  $\ell$  and its second derivative, and  $f(z)$  the growth rate of structure.

In **Angpow**, the inner integral in  $k$  is computed first. To conduct the computation of such an integral of highly oscillating functions, the 3C-algorithm described in details in ([Campagne et al. 2017a](#)) is used. In brief, it relies on the projection of the oscillating  $f_\ell(z, k)$  onto a Chebyshev series of order  $2^N$ , the product of the two Chebyshev series is performed with a  $2^{2N}$  Chebyshev series; then the integral is computed using Clenshaw-Curtis quadrature. Finally, the integrals over  $z$  are performed via again an optimised Clenshaw-Curtis quadrature. All the Chebyshev expansions and the Clenshaw-Curtis quadrature are performed via the *Discrete Cosine Transform* of type I from the DCT-I fast transform of the FFTW library ([Frigo & Johnson 2012](#)).

As in the general case the Limber approximation is valid at high  $\ell$  values, the CCL user can define an  $\ell$  threshold to switch from the non-Limber slow computation to the faster Limber approximation.

### 3.4. Correlation functions

The exact Equations (32, 34, 36) relating the angular correlation functions and power spectra involve carrying out  $N_\theta \times \ell_{\max}$  operations, where  $\ell_{\max} \sim O(10^{4-5})$  is the maximum multipole needed to achieve convergence and  $N_\theta$  is the number of angular scales  $\theta$  at which the angular correlation function needs to be computed. Thus, evaluating these expressions directly becomes prohibitively slow and should be avoided except in regimes where other approximations are not valid. In particular CCL

only supports the brute-force evaluation of these equations for correlations involving at least one spin-0 field. The currently preferred method in CCL is to use the flat-sky approximation and evaluate the Hankel transforms (Eqs. 33, 35, 37).

CCL currently provides two methods to compute Hankel transforms:

- **Brute-force integration.** CCL allows users to compute Hankel transforms by brute-force integration over the Bessel functions. The oscillating nature of these functions makes this method slow and not appropriate for likelihood-sampling. The preferred method to compute correlation functions is through the use of **FFTlog** (see below), and we support the brute-force method only for testing and validation.
- **FFTlog.** The public code **FFTlog**<sup>9</sup> is able to compute fast Hankel transforms through the assumption that the kernels of these transforms are periodic functions in logarithmic space. The Hankel transform can then be solved using Fast Fourier Transforms at a much lower computational expense than brute-force integration ([Hamilton 2000](#); [Talman 2009](#)). CCL incorporates a version of the **FFTlog** method with only minor modifications from the original. The only potential drawback of this methods is the need to sample the kernels (i.e. the  $C_\ell$ ) on very small scales to ensure the convergence of the method. To do this, CCL extrapolates the power spectrum as a power law ( $C_\ell \propto \ell^\beta$ ) **To do: How is  $\beta$  estimated?** We have verified that this method agrees with the brute-force integration well within cosmic-variance uncertainties.

We should also note that other approaches relating the correlation functions directly with the 3D matter power spectrum (e.g. [Campagne et al. 2017b](#)) could be useful in accelerating this computation, and we will explore these avenues in the future.

### 3.5. Halo mass function

**To do: Some of this seems would be better in the validation section. Can we restrict to implementation here?** To achieve  $10^{-4}$  precision in  $\sigma(M)$  and the normalisation of the power spectrum, one should check that the integral of  $\sigma_8$  and  $\sigma(M)$  has converged for the chosen values of  $\{k_{\min}, k_{\max}\}$ . After checking convergence, we achieved the desired precision.

<sup>9</sup> <http://casa.colorado.edu/~ajsh/FFTLog/>

Also note that for  $\sigma(M)$ , it is important to set the desired precision level correctly for the numerical integrator. The integral usually yields  $\sigma^2(M)$ , and not  $\sigma(M)$ . Hence, one has to set the desired precision taking the exponent into account.

Derivatives are calculated utilizing a spline that is built off of the previously described  $\sigma(M)$  spline. As such, these splines cover the range from  $10^6$  to  $10^{17}M_\odot$ . For each value of  $\log(M)$  in our spline evaluation, we calculate the value of  $\sigma(M)$  half a step in either direction. We use the difference compared to the mass spacing to calculate an approximate derivative, which is then used in the spline interpolation. This has been tested to meet our necessary precision for the halo mass function within the mass range explored by Tinker et al. (2010). We note that there the accuracy is reduced at the edges of these splines and exploring extreme mass ranges may require changes in the parameters to initialize these splines.

In order to accomodate a wide range of values of the overdensity parameter  $\Delta$ , we have generated a spline interpolation between best fit values as defined by Tinker et al. (2008) and Tinker et al. (2010). This covers a dynamic range from  $\Delta = 200$  to  $3200$ . Within this range, we interpolate in the space of the fit parameter and  $\log \Delta$  using the Akima interpolation built from piecewise third order polynomials. We have chosen this rather than the fitting formulas utilized in Tinker et al. (2010) in order to assure high precision match to the Tinker halo mass function when choosing a value of  $\Delta$  directly from the paper.

### 3.6. Massive neutrinos

When initializing a cosmology with massive neutrinos within CCL, the user can provide either a single value for  $m_\nu$ , corresponding to a sum of the masses of three neutrinos, or a set of three values, corresponding directly to the three masses. In the former case, one can also specify how the sum of masses should be split for calculations. The default behavior of CCL is to split the sum into three masses which are consistent with the normal neutrino mass hierarchy, but an inverted hierarchy or equal splitting can be requested. (For a review of the neutrino mass hierarchies and relevant particle physics results, see for example Lattanzi & Gerbino (2017); Lesgourgues & Pastor (2012).)

For equal splitting, it is clearly trivial to compute the three neutrino masses. If splitting with respect to the normal or inverted hierarchy is desired, it is slightly more complicated. The relevant known quantity which has been determined via particle physics experiments is the square of the difference of neutrino masses (up to a sign for one of the differences, hence the two pos-

sible hierarchies). Because we know the square of the differences rather than the differences themselves, we must solve a set of quadratic equations for the neutrino masses. In CCL this is accomplished by making use the knowledge that the mass difference between two of the neutrino species is small. This allows us to Taylor expand about this difference, and greatly simplifies solving for the three neutrino masses. This methodology represents an approximation (related to discarding higher-order terms in this Taylor expansion), however it produces discrepancies from the exact solution which result in cosmological predictions which differ at levels far below our required validation levels (detailed below in Section 4).

Having then a set of three neutrino masses, we check which of the corresponding neutrino species is non-relativistic today ( $m_\nu > 0.00017$ , Lesgourgues & Pastor 2012), and obtain the number of massive neutrinos in the cosmology. We use this, along with  $N_{\text{eff}}$ , to set the number of relativistic neutrinos species, which is required in computing  $\Omega_g$  and  $\Omega_{\nu,\text{rel}}$ . We must be slightly careful in doing so, as for massive neutrinos only we modify the relationship between the temperature of the CMB and the neutrino temperature as described below equation 3 above. In order to compute  $N_{\nu,\text{rel}}$  consistent with the  $N_{\text{eff}}$  passed by the user, we do:

$$N_{\nu,\text{rel}} = N_{\text{eff}} - (T_{\text{NCDM}})^4 \left( \frac{4}{11} \right)^{-\frac{4}{3}} N_{\nu,\text{m}}. \quad (65)$$

It may sometimes be preferable or necessarily to specify a cosmology in terms of  $\Omega_{\nu,\text{m}}$  instead of  $m_\nu$ . To facilitate this, CCL includes a convenience function which returns the latter given the former.

In equation 3 above, we specify how  $\Omega_{\nu,\text{m}}$  is computed for a given cosmology with massive neutrinos. Within this expression is a phase-space integral:

$$\int_0^\infty dx x^2 \frac{\sqrt{x^2 + (\mu)^2}}{\exp(x) + 1}. \quad (66)$$

At high and low  $\mu$ , corresponding to high and low mass neutrinos, this integral need not be evaluated numerically. At high  $\mu$ , we set the integral equal to  $\frac{5\zeta(3)}{18\pi^4}\mu$ , while at low  $\mu$  it goes to  $\frac{7}{8}$ . The  $\mu$  values at which these approximations are taken can be set by the user. Outside of the regime in which these approximations are valid, the integral is computed numerically using `gsl`, splined, and stored such that for a single cosmology it must only be computed once.

## 4. VALIDATION OVER THE $\Lambda$ CDM PARAMETER SPACE

Our goal in building CCL was to ensure that all outputs are validated to a well established accuracy level. This was achieved by comparing these outputs against one or multiple independent benchmarks obtained in the same configuration (i.e., for the same cosmology). In this section we document the accuracy achieved for each observable. In some cases, the achieved level of accuracy is higher than the impact of other systematics which have not yet been considered in this version of CCL. **To do: Not sure what this previous sentence refers to.** In the cases where this applies, we make it clear below. Table 2 summarises all the CCL validation tests discussed in this section.

CCL benchmarks are available and documented through the code repository. All plots presented in this section can be reproduced by means of a `python` notebook available with the code. Accuracy checks can also be run automatically upon installation of the software.

#### 4.1. Background quantities & growth of perturbations

Comoving radial distances, the growth factor and distance moduli were compared against independently produced benchmarks at a set of low,  $z = 0, 1, 2, 3, 4, 5$ , and high redshifts,  $10 \leq z \leq 10^3$ . These comparisons were performed for the cosmologies listed in Table 3. The accuracy metric was the defined as the fractional difference between the prediction made by CCL and by an independent implementation (labeled  $i$ ), i.e., for the comoving radial distance,

$$\mathcal{A} \equiv \frac{|D_{\text{CCL}}(z) - D_i(z)|}{D_{\text{CCL}}(z)} \quad (67)$$

**To do: Check for all  $\mathcal{A}$  definitions whether CCL is in the denominator in the code.** and analogously defined for the growth factor and distance moduli. The target level of accuracy was  $10^{-4}$  and this was achieved or surpassed for all the background quantities. Figure 2 summarizes our results for cosmologies without massive neutrinos. The left panels show the distance accuracy achieved for different cosmological models (curves of different thickness) as a function of redshift, which is always better than  $10^{-6}$ . Similarly, the right panels show the growth factor accuracy, which is  $< 10^{-5}$ .

Analogous comparisons of the comoving radial distances were made to independently produced benchmarks at low and high redshifts for cosmologies containing massive neutrinos. (Notice that the growth function with massive neutrinos is not supported by CCL because it is scale-dependent and therefore ill defined in our current framework.) These comparisons are displayed in Figure 3. The cosmological models used in tests of back-

ground quantities for cosmologies with massive neutrinos are given in Table 4. In this case, the benchmarks were produced by the `astropy` library (Astropy Collaboration et al. 2013). We increase the maximum tolerated fractional difference between quantities produced by CCL and benchmarks in this case to  $10^{-3}$ , due to the fact that while CCL uses the full phase-space integral in computing the massive neutrino density as defined above, `astropy` uses a fitting function which is itself not accurate at a  $10^{-3}$  level. As can be seen in Table 4, we perform comparisons in cosmologies with  $N_{\text{eff}} = 3$  rather than the more standard  $N_{\text{eff}} = 3.046$  in order to ensure the same cosmologies can easily be defined in CCL and `astropy`. **To do: Tilman, add comments on CLASS-based neutrino tests when merged.**

### 4.2. Matter power spectra

#### 4.2.1. Analytic expressions

As discussed in Section 2.3, several power spectrum methods are implemented in CCL. Two of them, the BBKS (Bardeen et al. 1986) and the Eisenstein & Hu methods are implemented for validation purposes only and feed into the tests for observables such as angular power spectra and correlation functions, as we will see in subsequent sections. These two implementations have been validated against independent implementations. The accuracy in this case was defined as the absolute fractional difference between the CCL and the independent predictions:

$$\mathcal{A} \equiv \frac{|P_{\text{CCL}}(k, z) - P_i(k, z)|}{P_{\text{CCL}}(k, z)} \quad (68)$$

For BBKS, this test was performed at  $0 \leq z \leq 5$  in the wavenumber range  $10^{-3} \leq k \leq 10h/\text{Mpc}$  with 10 bins per decade, and yielded an accuracy level of  $10^{-4}$ .

<sup>10</sup> For the Eisenstein & Hu matter power spectrum, we obtained similar accuracy at  $z = 0$  for the same wavenumbers. The cosmologies for which the tests were implemented are specified in Table 2.

The BCM implementation for the impact of baryons on the matter power spectrum, described in Section 2.3 is also analytical. Following Equation (68), we found it to be accurate to  $10^{-12}$ .

#### 4.2.2. Validation of interpolation schemes

<sup>10</sup> We noticed that there are 2 typographical errors for the BBKS transfer function in “Modern Cosmology” (Dodelson 2004) compared to the original BBKS paper. The quadratic term should be  $(16.1q)^2$  and the cubic term should be  $(5.46q)^3$ . The BBKS equation is correct in Peacock (1999). Using the wrong equation can give differences in the results above the  $10^{-4}$  level.

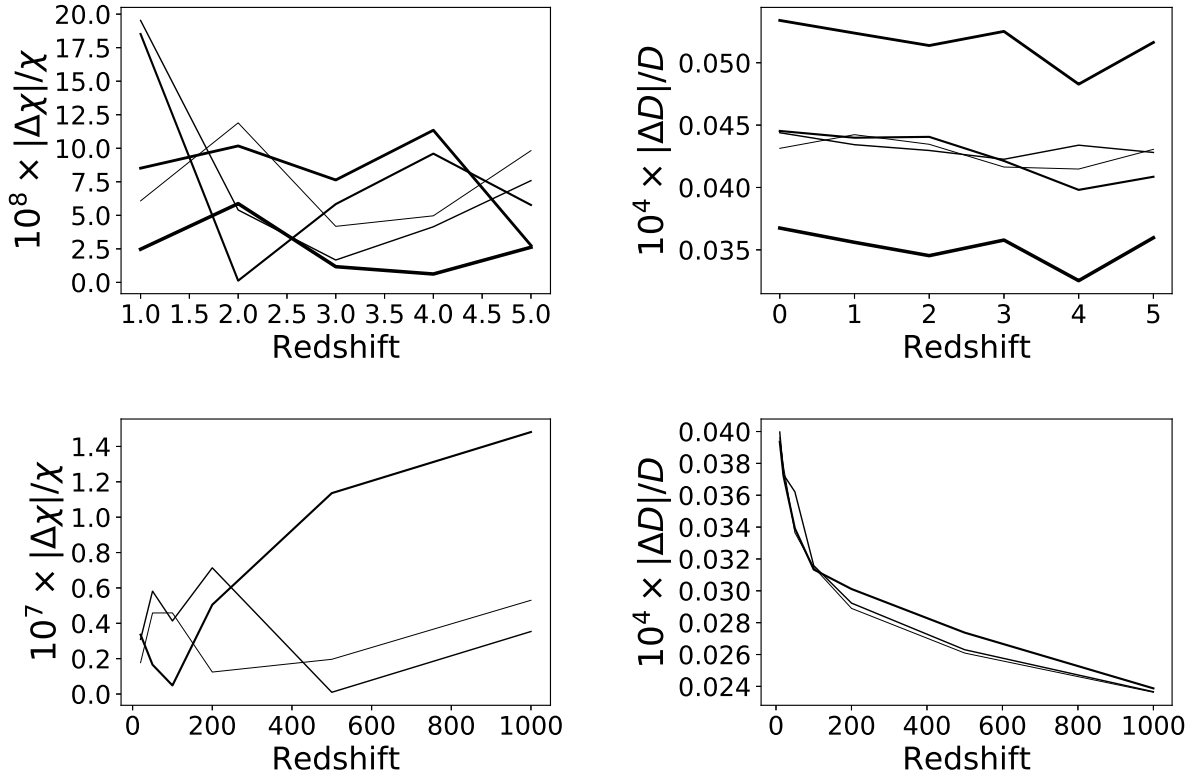
Quantity	Equation/Reference	Cosmologies	Range	Accuracy, $\mathcal{A}$
Comoving radial distance (low $z$ ), $\chi$	(4)	CCL1-5	$0 \leq z \leq 5$	$10^{-4}$
Comoving radial distance (high $z$ ), $\chi$	(4)	CCL1-3	$10 \leq z \leq 10^3$	$10^{-4}$
Growth (low $z$ ), $D$	(8)	CCL1-5	$0 \leq z \leq 5$	$10^{-4}$
Growth (high $z$ ), $D$	(8)	CCL1-3	$10 \leq z \leq 10^3$	$10^{-4}$
$\sigma(M)$ (BBKS)	(42)	CCL1-3	$10^6 \leq M/(M_\odot/h) \leq 10^{16}$	$10^{-4}$
$dn/dM$	(40)	CCL1	$10^{10} \leq M/(M_\odot) \leq 10^{16}$ & $z = 0$	$5 \times 10^{-3}$
$P(k)$ (BBKS)	(12)	CCL1-3	$10^{-3} \leq k/(h/\text{Mpc}) \leq 10$ & $0 \leq z \leq 5$	$10^{-4}$
$P(k)$ (Eisenstein & Hu)	Eisenstein & Hu (1998)	CCL1	$10^{-3} \leq k/(h/\text{Mpc}) \leq 10$ & $z = 0$	$10^{-4}$
$P(k)$ (CosmicEmu $w$ CDM)	Lawrence et al. (2017)	M1,M3,M5,M6,M8,M10	$10^{-3} \leq k/\text{Mpc}^{-1} \leq 5$ & $z = 0$	$3 \times 10^{-2}$
$P(k)$ (CosmicEmu $\nu$ CDM)	Lawrence et al. (2017)	M38,M39,M40,M42	$10^{-3} \leq k/\text{Mpc}^{-1} \leq 5$ & $z = 0$	$3 \times 10^{-2}$
$C_\ell$ clustering	(16),(17)	CCL6	$1 \leq \ell \leq 30000$	$10^{-3}$
$C_\ell$ weak lensing	(16),(21)	CCL6	$1 \leq \ell \leq 30000$	$10^{-3}$
$C_\ell$ CMB lensing	(16),(24)	CCL6	$1 \leq \ell \leq 3000$	$10^{-3}$
$\xi_+, \xi_-, \xi$	(37),(33)	CCL6	$0.01 < \theta/\text{deg} < 5$	$0.5\sigma$
3D correlation	(38)	CCL1-5	$1 < r/\text{Mpc} < 200$ & $0 \leq z \leq 5$	$4 \times 10^{-3}$
$C_\ell$ clustering non-Limber	(16),(17),(18)	CCL1	$500 \leq \ell < 1000$	$2 \times 10^{-2}$
$C_\ell$ clustering Angpow	(16),(17),(18)	CCL1	$1 \leq \ell < 1000$	$1 \times 10^{-2}$

To do: **CLASS?**

**Table 2.** Summary of CCL validation tests and accuracy achieved. These tests can be reproduced by the user and are integrated into the CCL repository via TravisCI. To do: Some questions about the HMF test. We are checking the HMF but also  $\sigma$  and  $\sigma^{-1}$  with different tolerances. In what way are those checks different from the  $\sigma(M)$  check? Also, the units of mass in the  $\sigma(M)$  check and HMF check are inconsistent - do we want to change that? For non-Limber tests, the quoted precision stands for the ratio between the relative errors between the computed  $C_\ell$ s and the cosmic variance. We should also clarify here what the difference is between the two  $C_\ell$  non-Limber tests.

Cosmological models for code comparison project									
Acronym	Model	$\Omega_m$	$\Omega_b$	$\Omega_\Lambda$	$h_0$	$\sigma_8$	$n_s$	$w_0$	$w_a$
CCL1	flat $\Lambda$ CDM	0.3	0.05	0.7	0.7	0.8	0.96	-1	0
CCL2	$w_0$ $\Lambda$ CDM	0.3	0.05	0.7	0.7	0.8	0.96	-0.9	0
CCL3	$w_a$ $\Lambda$ CDM	0.3	0.05	0.7	0.7	0.8	0.96	-0.9	0.1
CCL4	open $w_a$ $\Lambda$ CDM	0.3	0.05	0.65	0.7	0.8	0.96	-0.9	0.1
CCL5	closed $w_a$ $\Lambda$ CDM	0.3	0.05	0.75	0.7	0.8	0.96	-0.9	0.1
CCL6	flat $\Lambda$ CDM	0.3	0	0.7	0.7	0.8	0.96	-1	0

**Table 3.** Cosmological models for code comparison project. To do: Model 6 was introduced because we use it in some of the checks, it needs more explanation in the main text

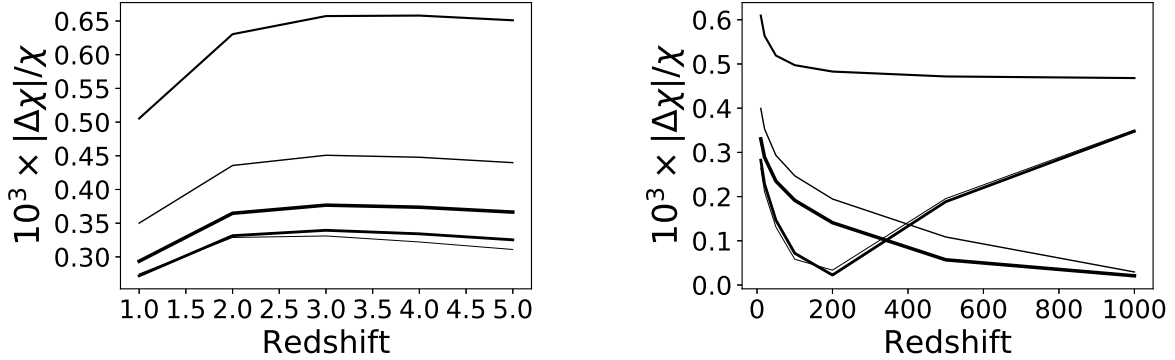


**Figure 2.** Accuracy achieved by CCL in the prediction of distances (left panels) and growth factor (right panels) for models CCL1-5 documented in Table 3. Different models correspond to different thickness curves. The top panels represent the accuracy checks at low redshift, while the bottom panels demonstrate accuracy up to high redshift.



Cosmological models with massive neutrinos											
Acronym	Model	$\Omega_m$	$\Omega_b$	$\Omega_\Lambda$	$h_0$	$\sigma_8$	$n_s$	$w_0$	$w_a$	$N_{eff}$	$m_\nu$ (eV)
CCL7	flat $\Lambda$ CDM, $m_\nu$	0.3	0.05	0.7	0.7	0.8	0.96	-1	0	3	{0.04, 0, 0}
CCL8	$w_0$ $\Lambda$ CDM, $m_\nu$	0.3	0.05	0.7	0.7	0.8	0.96	-0.9	0	3	{0.05, 0.01, }
CCL9	$w_a$ $\Lambda$ CDM, $m_\nu$	0.3	0.05	0.7	0.7	0.8	0.96	-0.9	0.1	3	{0.03, 0.02, 0.04}
CCL10	open $w_a$ $\Lambda$ CDM, $m_\nu$	0.3	0.05	0.65	0.7	0.8	0.96	-0.9	0.1	3	{0.05, 0, 0}
CCL11	closed $w_a$ $\Lambda$ CDM, $m_\nu$	0.3	0.05	0.75	0.7	0.8	0.96	-0.9	0.1	3	{0.03, 0.02, 0}

**Table 4.** Cosmological models with massive neutrinos used in testing CCL against independently produced benchmarks



**Figure 3.** Accuracy achieved by CCL in the prediction of distances for models CCL7-11 with massive neutrinos, documented in Table 4. Different models correspond to different thickness curves. The left panel represent the accuracy checks at low redshift, while the right panel demonstrates accuracy up to high redshift. Note that we allow up to  $10^{-3}$  fractional difference between benchmarks and CCL in this case, because benchmarks utilize a fitting function for the neutrino phase space integral which is itself not accurate to  $10^{-4}$ .

In its default configuration, CCL adopts the HaloFit (Takahashi et al. 2012) implementation by interpolating CLASS power spectra outputs to model the matter power spectrum. The computation of the power spectrum from CLASS can be significantly sped up by interpolating the matter power spectra in the range  $K\_MIN < k < K\_MAX\_SPLINE$  and extrapolating beyond it, as described in Section 3. In this section, we describe the loss of accuracy due to this method. The tests presented are performed in a flat  $\Lambda$ CDM cosmology with  $\Omega_c = 0.25$ ,  $\Omega_b = 0.05$ ,  $A_s = 2.1 \times 10^{-9}$ ,  $h = 0.7$  and  $n_s = 0.96$ .

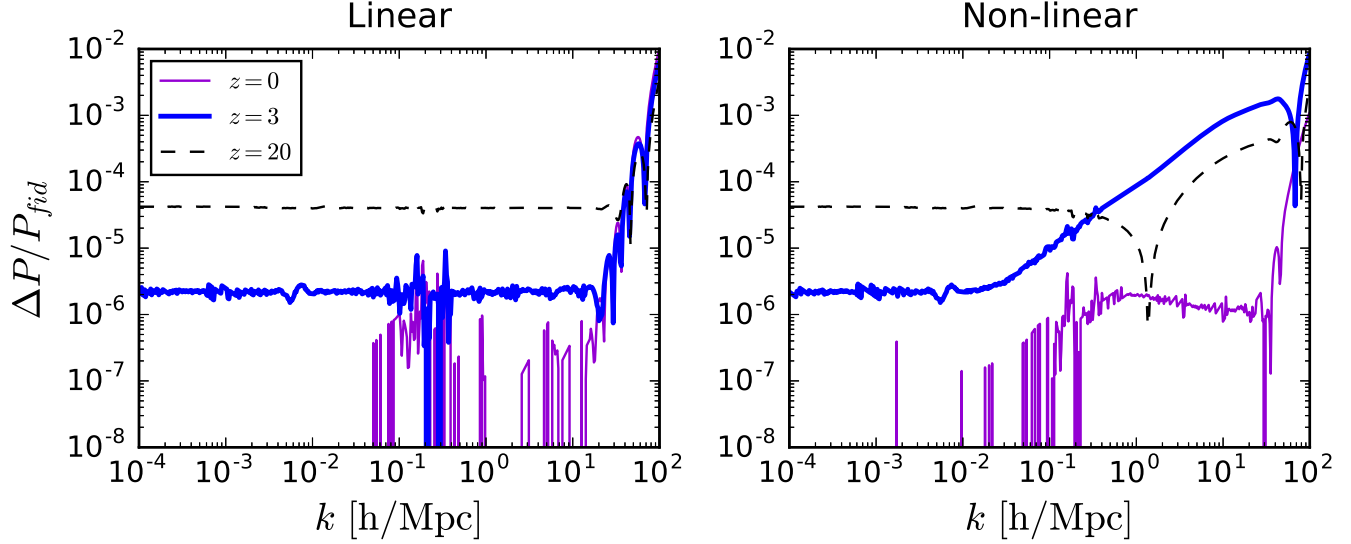
The accuracy of this approximation is shown in Figure 4 for redshifts  $z = 0$ ,  $z = 3$  and  $z = 20$ . We compare the nonlinear matter power spectrum at these redshifts, computed with the previously described approximation, to the matter power spectrum obtained by setting the power spectrum splines to high-accuracy values. We find that for typical values of  $\Delta \ln k = 10^{-2}$  and  $K\_MAX\_SPLINE = 50/\text{Mpc}$ ,  $\ln P$  has converged to an accuracy that surpasses the expected impact of baryonic effects on the matter power spectrum at  $k > 10/\text{Mpc}$ .

To do: Christiane, what was exactly the reference fidu-

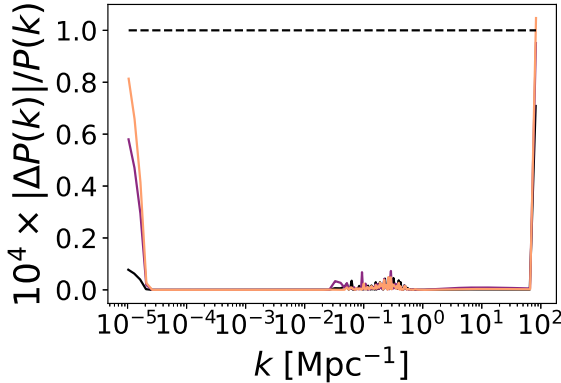
cial power? (For an estimate of the impact of baryons on the total matter power spectrum, see Schneider & Teyssier 2015.)

With the implementation described above, the power spectrum splines are initialized up to  $K\_MAX\_SPLINE$ . This is also true for the linear matter power spectrum, which is used within CCL in particular to obtain  $\sigma_8$  (see Eq. 42). We have tested how this procedure affects the convergence of the linear matter power spectrum. We compare the fiducial CCL output to the case where we set  $K\_MAX\_SPLINE = 5 \times 10^3/\text{Mpc}$ . The result is shown in Figure 4. For some applications that use the linear power spectrum, the user might need to increase the value of  $K\_MAX\_SPLINE$ . While CCL adopts certain fiducial values of the number of scale factor and wavenumber values to use in the interpolation, we have tested that increasing the sampling does not change the results presented in Figures 4.

In addition to the above tests in  $\Lambda$ CDM cosmologies without massive neutrinos, we have checked the impact of using splines at low and high  $k$  in cosmologies CCL7, CCL8, and CCL9 with massive neutrinos, defined in Table 4. We compare the linear and nonlinear mat-



**Figure 4.** The relative error compared to power spectra produced with high values of the power spectrum splines,  $P_{fid}$ , produced by splining the matter power spectrum up to  $K_{MAX\_SPLINE}$  and extrapolating beyond this value with a second order Taylor expansion the natural logarithm of the matter power spectrum. The left panel shows the relative errors for the linear matter power spectrum at  $z = 0$ ,  $z = 3$  and  $z = 20$ . The right panel shows the results for the non-linear matter power spectrum at the same redshifts. The standard CCL parameters adopted are those corresponding to the black dashed curve. For comparison, the impact of baryonic physics on the matter power spectrum is  $\sim 10\%$  at  $k = 1/\text{Mpc}$  (Schneider & Teyssier 2015).



**Figure 5.** Fractional difference between the nonlinear matter power spectrum as computed directly via **CLASS** with that computed using **CLASS** via **CCL** in cosmologies CCL7, CCL8, and CCL9 with massive neutrinos.

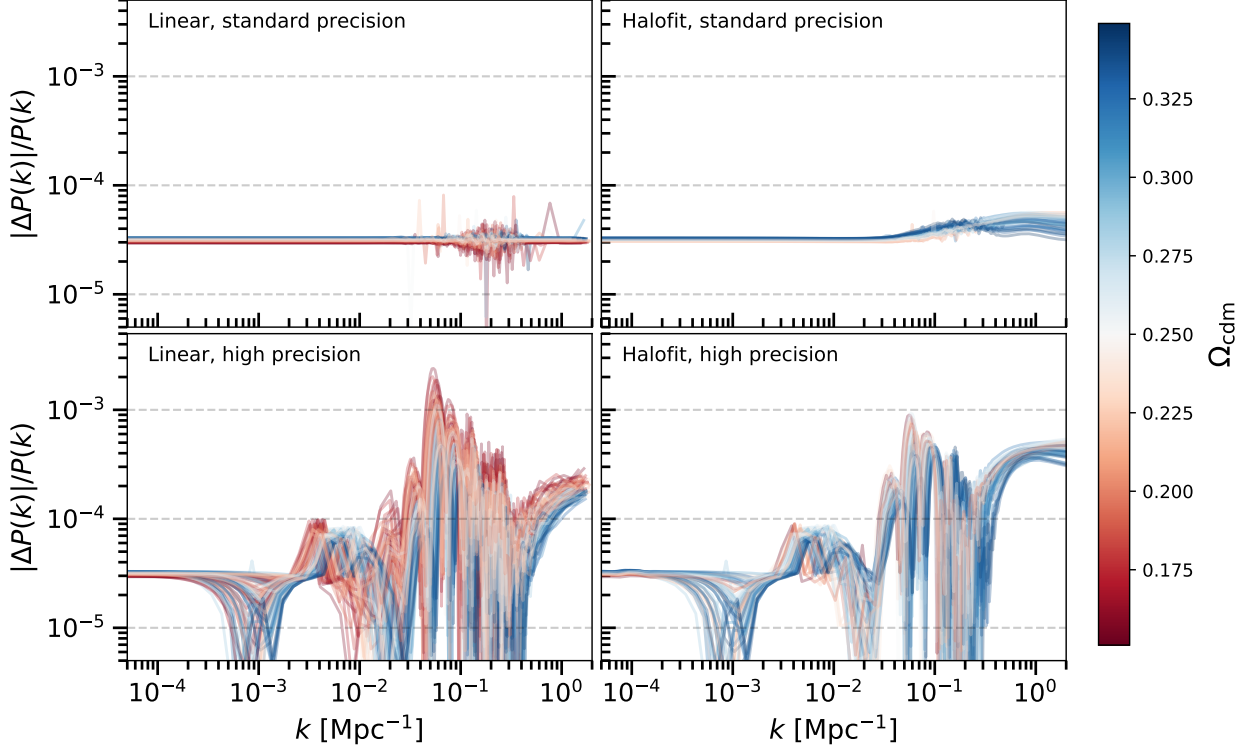
ter power spectrum as computed directly via **CLASS** to that computed using **CLASS** via **CCL** (i.e. using splines at high and low  $k$ ). We find that for  $k$  between  $K_{MIN}$  and  $K_{MIN\_SPLINE}$ , the two power spectra agree to better than  $10^{-4}$  in all models. For  $k$  between  $K_{MAX\_SPLINE}$  and  $K_{MAX}$ , agreement is better than  $10^{-3}$ , which we deem sufficient given the significant uncertainties introduced at these small scales, e.g. baryonic effects. The fractional difference between the two nonlinear power spectra is plotted in Figure 5.

#### 4.2.3. Generalized validation of the power spectrum over $\Lambda$ CDM parameter space

**Volunteer: Phil Bull**

While concentrating on individual points in cosmological parameter space allows us to perform detailed validation tests, as above, it is important for CCL to also be validated across a wide range of cosmological parameter values, e.g. to ensure validity for MCMC analyses. In this section, we present a set of validation tests for the CCL linear and non-linear matter power spectrum functions that spans a broad range of  $\Lambda$ CDM parameters.

Covering a full range of all 5  $\Lambda$ CDM parameters on a regular grid would be prohibitively expensive, so an alternative method for fairly (but more sparsely) sampling the parameter space is needed. We use Latin Hypercube Sampling to determine a tractably-sized set of sample points. This splits the parameter space into a regular grid with a given number of cells per dimension. The sample points are then chosen by going through each dimension in turn and choosing a cell at random without replacement, so that a given cell in each dimension is only ever chosen once. This is repeated until all cells in all dimensions contain a single sample (or until a maximum number of sample points has been reached). This has the effect of covering the space uniformly but sparsely. The exact location of the sample within each cell can be chosen from a uniform distribution within that cell, but for simplicity we put each sample at the



**Figure 6.** Absolute fractional difference between the matter power spectra at  $z = 0$  calculated using CLASS via CCL, and CLASS directly, for a range of cosmological parameter values, and different CLASS precision settings (standard vs. high-precision) and power spectrum types (linear vs. Halofit). The lines are colored according to the value of  $\Omega_c$  for each set of cosmological parameters (see Table 5 for the ranges of other parameters).

cell centre. We use 100 sample points in total, with the ranges for each parameter given in Table 5. For the purposes of this exercise, we allow only massless neutrinos ( $N_{\text{eff}} = 3.046$ ), and set  $T_{\text{CMB}}$  to the same value in CCL and CLASS.

Parameter	Range
$h$	[0.55, 0.8]
$\Omega_c$	[0.15, 0.35]
$\Omega_b$	[0.018, 0.052]
$A_s$	$[1.5, 2.5] \times 10^{-9}$
$n_s$	[0.94, 0.98]

**Table 5.** Ranges of  $\Lambda$ CDM parameters used for the generalised CCL validation tests of the matter power spectrum.

For each set of parameters, we then calculate the linear and non-linear (Halofit) power spectra using CCL for a range of redshifts. A corresponding set of reference power spectra is then produced using CLASS directly, i.e. using a regular installation of CLASS (v2.6.3) that is completely independent of the CCL code. We run this with either default precision settings (‘stan-

dard precision’), or settings intended to produce high-precision CMB results (‘high precision’), taken from the `pk_ref.pre` precision file that is bundled with CLASS.

Fig. 6 shows the fractional difference between the CCL and CLASS matter power spectra at  $z = 0$  for 100 sample points over the parameter space, with each line colored according to the value of  $\Omega_c$  for that sample. Results for different power spectrum types (linear vs. Halofit) and CLASS precision settings are shown for comparison.

As shown in the top two panels of Fig. 6, CCL reproduces the standard CLASS results well across a broad range of parameter values, always remaining well within the target fractional precision of  $10^{-4}$ . This demonstrates the robustness of our choice of spline parameters to different cosmological parameter values.

The lower panels in Fig. 6 show the fractional deviation between CCL (which always uses the ‘standard’ CLASS precision) and CLASS with high precision settings. These deviations are more significant, especially around the wavenumbers where the BAO feature is most prominent. The precision is still generally better than  $10^{-3}$  however, and is only worse than that for the very highest values of  $\Omega_c$ .



**Figure 7.** Absolute fractional difference between CCL and CLASS matter power spectra, as plotted in Fig. 6, but now at  $z = 2$ .

Fig. 7 shows the same comparison, but now for  $z = 2$ . The precision on the linear matter power spectrum is almost an order of magnitude better than at  $z = 0$  for the standard precision settings, and about the same as at  $z = 0$  for the high precision settings. The picture is quite different for the Halofit power spectrum however, where large deviations are seen regardless of the precision settings for models with small values of  $\Omega_c$ . Larger values of  $\Omega_c$  produce only slightly worse precision than at  $z = 0$  however.

These results show that the CLASS-based CCL power spectrum calculations are robust across a broad range of cosmological parameters, especially for the linear power spectrum, but that some caution must be taken when using the Halofit power spectrum in MCMC studies for example. Since other, more realistic, non-linear power spectrum methods are available (e.g. see Sect. 4.2.4), and tend to have larger theoretical errors than the deviations shown in Fig. 7, this is unlikely to be a serious concern for most users however.

#### 4.2.4. Validation of the Cosmic Emulator implementation

To do: I think this section should go before the one on generalized validation of the power spectrum over LCDM space, because that one refers to 'the sections above' as being at specific cosmologies, which is the case

here. - DLEC: The only caveat is that the generalized validation scheme doesnt account for the emulator.

Lawrence et al. (2017) quantified the accuracy of the matter power spectrum from their emulation scheme by comparing their predictions to the resulting power spectra from the numerical simulations used for the calibration of their scheme. We repeated this procedure making the emulator predictions via CCL for a subset of emulator cosmologies. In this section, we quantify the accuracy of the CCL predictions by estimating

$$\mathcal{A} \equiv \frac{|P_{\text{CCL}}(k, z) - P_{\text{L17}}(k, z)|}{P_{\text{CCL}}(k, z)} \quad (69)$$

where the label L17 refers to the smoothed simulated power spectra from Lawrence et al. (2017).

Our results are shown in Figure 8. For cosmologies without neutrinos, we required the matter power spectrum at  $z = 0$  to be within 1% of the smoothed simulated power spectrum from Lawrence et al. (2017) (see their Figure 6). Similarly, we required 3% accuracy for cosmologies with neutrinos (their Figure 5). The cosmologies that were tested are the ones listed in Table 3, which are described in detail in Lawrence et al. (2017). EC: Need to explain 1 – 3% better.

#### 4.3. Halo bias and halo mass function

Independent codes were utilized to test the accuracy of halo mass function predictions. These predictions are



**Figure 8.** Absolute fractional accuracy in the matter power spectra, Eq. (69), obtained by calling the cosmic emulator from CCL and the smoothed simulated spectra from Lawrence et al. (2017). The left panel shows the results for cosmologies without neutrinos; the right panel, results for cosmologies with neutrinos.

computed from power spectra obtained using the BBKS approximation. For the halo mass function, we compare the value of  $\sigma$ ,  $\tilde{\sigma} \equiv \log(\sigma^{-1})$ , and the value of the halo mass function in the form used in Tinker et al. (2008),

$$\mathcal{H} \equiv \log[(M^2/\bar{\rho}_m)dn/dM]. \quad (70)$$

These defines three new accuracy metrics:

$$\mathcal{A}_{hmf1} \equiv \frac{|\sigma_{\text{CCL}} - \sigma_i|}{\sigma_{\text{CCL}}}, \quad (71)$$

$$\mathcal{A}_{hmf2} \equiv \frac{|\tilde{\sigma}_{\text{CCL}} - \tilde{\sigma}_i|}{\tilde{\sigma}_{\text{CCL}}}, \quad (72)$$

$$\mathcal{A}_{hmf3} \equiv \frac{|\mathcal{H}_{\text{CCL}} - \mathcal{H}_i|}{\mathcal{H}_{\text{CCL}}}. \quad (73)$$

We note that while we maintain the  $10^{-4}$  for our evaluations of  $\sigma$ , the accuracy degrades to a value of  $5 \times 10^{-3}$  for the halo mass function evaluation, primarily at the high halo mass and high redshift domains. We find that this increased error is acceptable, as the level of precision is significantly better than the accuracy of current halo mass function models. This is demonstrated in Figure 9, where this calculation has been run for a single cosmology using the Tinker 2010 halo mass function. While there is a degradation in accuracy due to our current spline treatment of the log inverse of  $\sigma(M)$ , we note that it does not significantly degrade our halo mass function determination. While improvement on this remains a task for the future, the halo mass function varies between fitting functions significantly more than this remaining error. As of this time, we currently do not have independent benchmarks for the halo bias function, though it should be noted that this calculation does not involve any additional functions beyond  $\sigma(M)$  and should hold to the  $10^{-4}$  tolerance level.

#### 4.4. Two-point statistics

We used the BBKS linear matter power spectrum to compare two-point statistics for two redshift bins, resulting in three tomography combinations,  $(1-1), (1-2), (2-2)$ . We computed the following quantities:

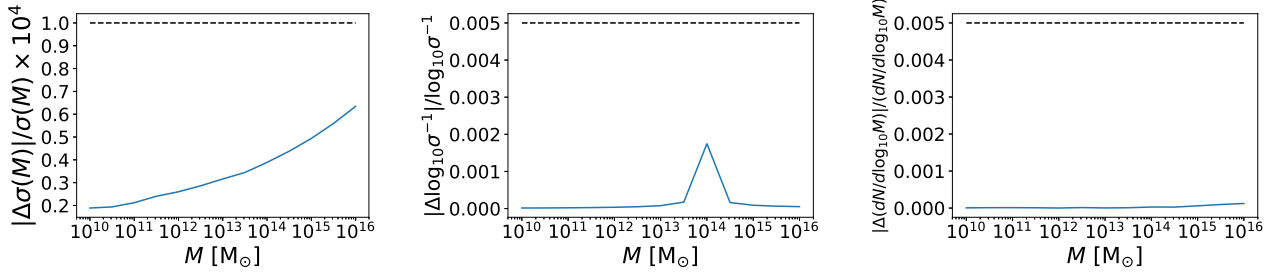
- Number counts angular power spectra: density term only (no magnification, RSD, etc.) with non-evolving linear bias  $b(z) = 1$ , in the range  $10 < \ell < 10000$ , using 5 bins per decade,
- Lensing convergence angular power spectra: leading order convergence term only (no magnification or intrinsic alignments), in the same range and with the same resolution as the case above,
- Number counts angular correlation functions in the range  $0.01 \text{ deg} < \theta < 5 \text{ deg}$ , using 5 bins per decade, and
- Lensing shear angular correlation functions  $(\xi_+, \xi_-)$ , similarly to above.

We adopted the following analytic redshift distributions: a Gaussian with  $\sigma = 0.15$ , centered at  $z_1 = 1$ ; and another Gaussian with the same dispersion but centered at  $z_2 = 1.5$ . We repeated the exercise for two redshift distribution histograms shown in Figure 10. Besides these two-point functions, we also validated the computation of the CMB lensing power spectrum, using the same  $\ell$  binning described above.

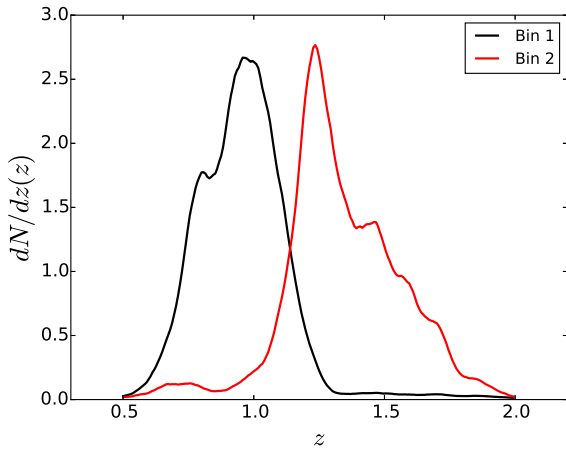
For  $C_\ell$  computations, we achieved a relative difference between CCL and the benchmarks of  $< 10^{-3}$ . This applied both to analytic redshift distributions and histograms, for auto- and cross-correlations between bins for both galaxy clustering and weak lensing, as shown in Figure 11. The CMB lensing result is shown in Figure 12.

To obtain realistic targets for the convergence of projected correlation function computations for LSST anal-





**Figure 9.** Three different numerical tests of the halo mass function calculation. In each line, the blue line is the fractional error in the function, while the black dashed line represents our error tolerance. The first panel demonstrates the robust calculation of  $\sigma(M)$ . The second panel demonstrates a numerical quirk in our spline treatment that is currently not addressed, but does reduce the numerical accuracy in returning the log inverse of  $\sigma(M)$ . We note that this does not significantly impact the error in the halo mass function in the final panel.



**Figure 10.** Binned redshift distributions used for code comparison project.

yses, we calculated the expected statistical uncertainty of the clustering and lensing correlation functions of the LSST gold sample (LSST Science Collaboration et al. 2009) assuming an effective source galaxy density of  $n_{\text{eff}} = 26$  gal/sq arcmin for galaxy shape distortions (Chang et al. 2013), and galaxy density of  $n_{\text{gold}} = 45$  gal/sq arcmin for number counts. Specifically, we calculated the Gaussian covariance of angular correlation functions following the formalism of Joachimi et al. (2008), and note that leaving out the non-Gaussian covariance terms makes our accuracy criterion more conservative. We split the galaxy samples into 10 tomography bins, defined to contain equal numbers of galaxies. The accuracy test then proceeded as follows. We compared the difference between CCL calculated lensing and clustering correlations and the benchmarks, for the analytic redshift distributions and for auto-correlations of redshift bins only. To pass the benchmark test, we required that this difference be smaller than half of the value of the error bar derived from the covariance for

each correlation function computed. Specifically, we took the value of the covariance in the bins centered at  $z = 1$  and  $z = 1.5$  to compare to the benchmarks.

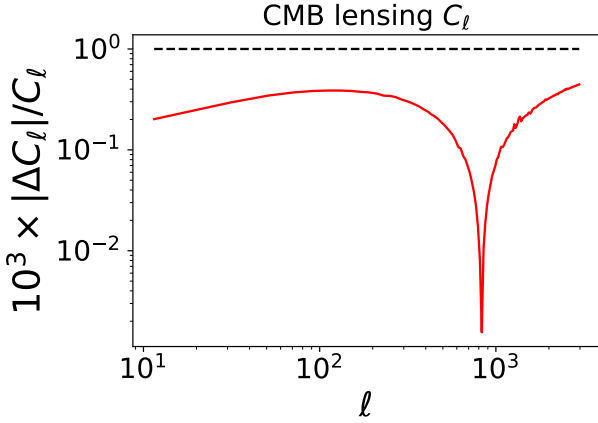
The results of this validation procedure for the projected correlation function are shown in Figure 13. Our results suggest that the convergence between CCL predictions and benchmarks is below the expected statistical uncertainty.

The 3-dimensional spatial correlation function  $\xi(r)$  was validated by comparing with an independent, precise numerical transform. We calculated  $\xi(r)$  by transforming the CCL non-linear HaloFit power spectrum using this independent code for the five cosmologies listed in Table 3 at redshifts  $z = 0, 1, 2, 3, 4, 5$ . We then compared with the  $\xi(r)$  from CCL with a sampling of  $P(k)$  equal to N\_K\_3DCOR bins per decade. The default value of N\_K\_3DCOR = 100,000 results in a relative agreement at the level of  $\Delta\xi(r)/\xi(r) < 2.5 \times 10^{-3}$  for  $0.1 < r < 250$  Mpc and  $z = 0$ . The agreement was better for higher redshifts. We also compared the absolute value of  $r^2\xi(r)$  and find a maximum difference of  $\Delta(r^2\xi(r)) < 3.0 \times 10^{-2}$  for the range  $r = 0.1 - 250$  Mpc. This corresponds to approximately 0.08% of the Baryon Acoustic Oscillation peak value of  $r^2\xi(r)$ . At the peak, the difference is only  $9.0 \times 10^{-3}$ , or 0.024% of the peak height. The results are shown in Fig. 14

To further validate the  $P(k) \rightarrow \xi(r)$  transform we performed a test using an analytical function  $\xi(r) = (r/r_0)^a$ , whose inverse transform  $P(k)$  has a known analytic form. We used  $r_0 = 5h^{-1}$  Mpc $^{-1}$  and  $a = -1.67$ , which approximates the actual three-dimensional correlation function. We then compared the CCL calculation of  $\xi(r)$  to the known analytic result. The relative difference  $\Delta\xi(r)/\xi(r)$  was found to be less than 0.4% in the range  $1 < r < 200$  Mpc rising to about 5% at  $r = 1000$  Mpc (see Fig. 15). For  $r = 0.1 - 0.8$  Mpc the relative difference is  $\approx 8\%$ . The accuracy at low and



**Figure 11.** Tests of the angular power spectrum accuracy. Benchmark comparisons for number counts (solid) and weak lensing (dashed) power spectra are shown for analytic and histogram-based redshift distributions in the left and right panels respectively.



**Figure 12.** Same as Fig. 11 for the CMB lensing power spectrum.

high distances can be improved by increasing the range over which the power spectrum splines are evaluated.

#### 4.4.1. Angpow

CCL performs non-Limber computations of angular power spectra through the **Angpow** library as detailed in Section 3.3.2. The **Angpow** software was tested against **CLASS** and the native CCL computation and can perform the same computations approximately an order of magnitude faster ( $\mathcal{O}(1s)$ ). Its precision and speed parameters were optimised so that the relative numerical error compared with a high precision computation is two orders of magnitude below the relative cosmic variance  $\sqrt{2/(2\ell+1)}$ , from  $\ell = 2$  to  $\ell = 1000$ . We demonstrate this in Figure 16, where we plot the angular clustering power spectrum for a sample of galaxies with  $\langle z \rangle = 1$  and a Gaussian redshift distribution that extends between  $|z - \langle z \rangle| < 5\sigma_z$ , where  $\sigma_z = 0.02$ , for a CCL6 cosmology. The non-Limber prediction deviates from the Limber case at low  $\ell$  as expected. The right panel

shows the fractional difference between the non-Limber curves and the Limber case, demonstrating the accuracy of the **Angpow** prediction for our choice of precision and speed parameters. Also the native non-Limber computation and **Angpow** were tested to recover the Limber approximated curve at high  $\ell$  for a wide Gaussian window ( $\sigma_z = 0.1$ ). The relative errors with respect to the Limber result at high  $\ell$  are below two orders of magnitudes with respect to cosmic variance.

## 5. USAGE

CCL is a public tool developed by the members of the Dark Energy Science Collaboration of LSST and can be downloaded from the collaboration’s github repository at the URL provided in the abstract. Installation instructions are provided in a README file available in that same repository. Instructions on how to generate a Docker<sup>11</sup> image are provided for portability to different architectures. CCL dependencies include the GNU Scientific Library<sup>12</sup> and FFTW3<sup>13</sup>. A suite of tests can be run to ensure installation was successful and all features perform nomrmally. Accuracy checks are performed in C, while a suite of unit tests is available in python.

CCL is documented using Doxygen<sup>14</sup>. For convenience, the repository includes multiple example files in C and several Jupyter notebooks.

CCL is released under terms consistent with BSD 3-Clause licensing<sup>15</sup>.

## 6. OUTLOOK

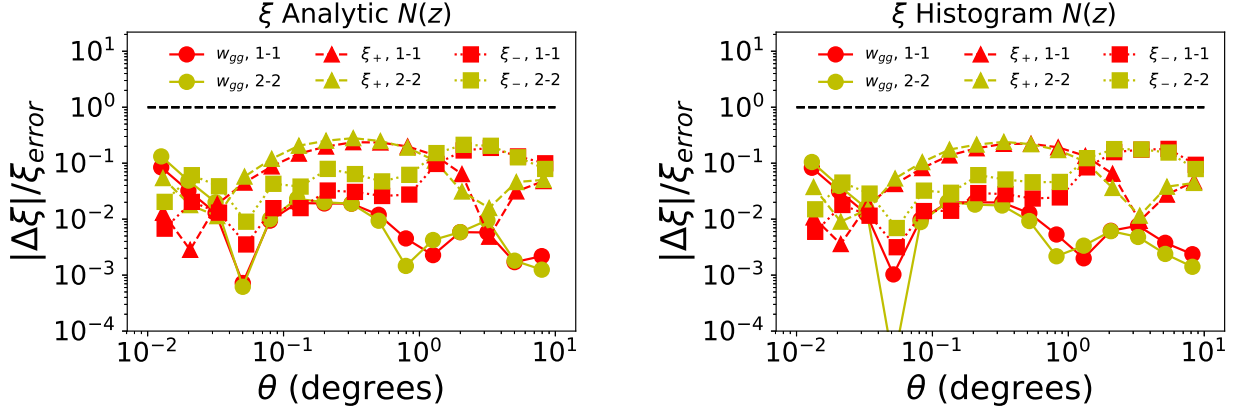
<sup>11</sup> <https://www.docker.com/>

<sup>12</sup> <https://www.gnu.org/software/gsl/>

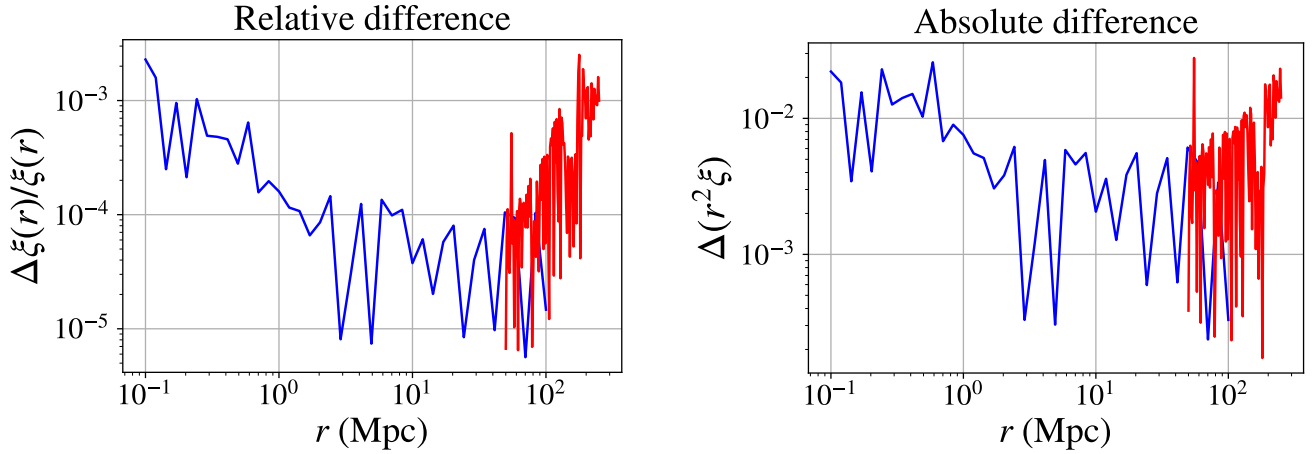
<sup>13</sup> <http://www.fftw.org/>

<sup>14</sup> [www.doxygen.org/](http://www.doxygen.org/)

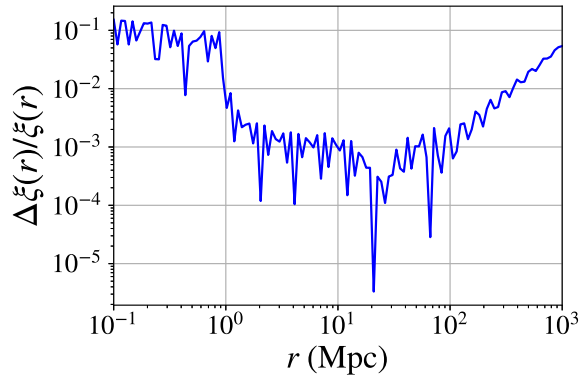
<sup>15</sup> <https://github.com/LSSTDESC/CCL/blob/master/LICENSE>



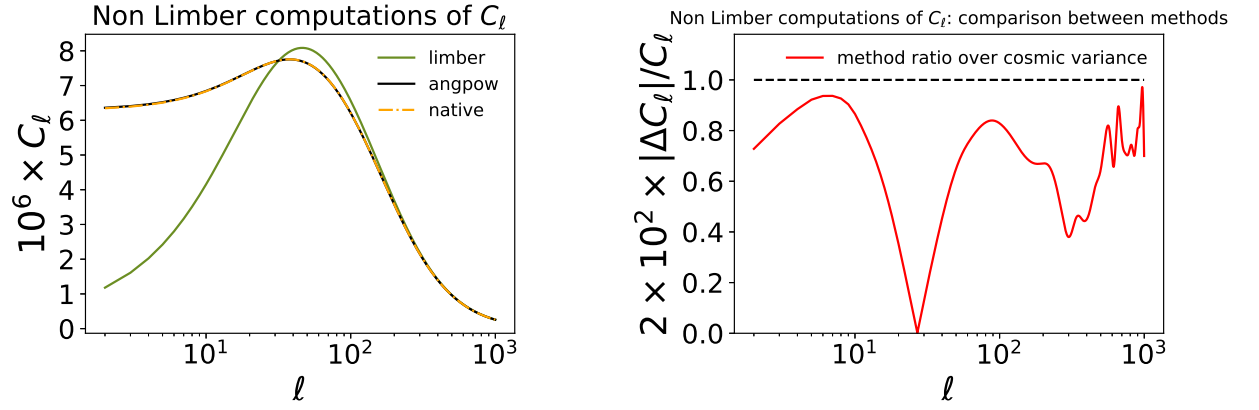
**Figure 13.** Comparison between the predicted projected correlation functions and the expected uncertainties for LSST. The left panel shows predictions for the analytic redshift distributions, while the right panel shows the case of the redshift histograms. The different markers and colors indicate clustering ( $w_{gg}$ , filled circles) or lensing ( $\xi_{\pm}$ , filled triangles and squares) auto-correlations of the 1-1 or 2-2 redshift bin combinations.



**Figure 14.** Comparison of the CCL calculation of the three-dimensional spatial correlation function  $\xi(r)$  with a precise, numerical transform of the CCL non-linear Halofit power spectrum. The left panel shows the relative error  $\Delta\xi(r)/\xi(r)$ . The right panel shows the absolute error in  $r^2\xi(r)$ . Both panels are for the CCL1 model of Table 3 at redshift zero.



**Figure 15.** The relative error in the three-dimensional spatial correlation function computed using the CCL algorithm compared to an analytic function  $\xi(r) = (r/r_0)^a$  whose inverse transform  $P(k)$  is known analytically. In this validation test, the known  $P(k)$  was transformed with the CCL algorithm and compared to the known analytic result for  $\xi(r)$ .



**Figure 16.** Different predictions for the clustering angular power spectrum of a sample of galaxies with a Gaussian redshift distribution centered on  $\langle z \rangle = 1$ . The left panel shows the  $C_\ell$  predictions from the Limber case (green), the non-Limber case native to CCL (yellow dot-dashed) and **Angpow** (black). The right panel shows the fractional difference in the predicted clustering angular power spectrum between **Angpow** and the native CCL non-Limber computation. The relative numerical error compared with a high precision computation is two orders of magnitude below the relative cosmic variance:  $\sqrt{2/(2\ell + 1)}$ .

Science software development to facilitate the cosmological inference from LSST data is one of the critical tasks of the Dark Energy Science Collaboration (DESC). Recent cosmological analyses of the Dark Energy Survey (DES) relied on CosmoSIS (Zuntz et al. 2014) and CosmoLike (Krause et al. 2017), the Kilo-Degree Survey (KiDS) uses CosmoLSS Joudaki et al. (2018) and CosmoMC (Lewis & Bridle 2002). All of these frameworks employ CLASS, CAMB (Challinor & Lewis 2005), or the cosmic emulator to compute the density power spectra. Compared to the analyses of DES, KiDS and the Hyper-Suprime Cam Survey (HSC), future data sets (e.g. LSST, *Euclid* and *WFIRST*) have substantially higher demands on analysis frameworks. Analyses are becoming more complex in terms of cosmological physics that is included in the analyses (neutrinos, modified gravity, and dark matter models) and in terms of modeling astrophysical and observational systematics at the required precision.

It is the primary goal of the CCL to become the backbone of all cosmological analyses carried out by the Dark Energy Science Collaboration of the LSST. This unified approach of a validated CCL will ensure that DESC results that are both consistent (in that they will all be based on the same theory framework) and accurate (in that this framework has undergone a rigorous numerical validation).

The implementation of CCL in realistic analysis pipelines has already begun: all likelihood module prototypes under development use it as its backbone, and the first of these, cosmological analysis of angular galaxy clustering cross-correlations, will serve as a model for the design of the joint-probes likelihood of the DESC. This work has allowed us to validate the performance of CCL in a realistic analysis scenario, verifying its accuracy and efficiency in the context of computationally demanding MCMC runs.

Beyond its usefulness in the DESC, the flexible design of CCL makes it a useful tool for the analysis of other cosmological datasets, as well as for the cross-correlation of different experiments. To this end, and to allow a generic and flexible analysis of the LSST data, further functionality will be added to CCL. Plans are in place to extend the range of standard and non-standard cosmological models covered by the code. Work is already underway to add a comprehensive implementation of the halo-model calculation of two-point functions (Peacock & Smith 2000) with sufficient flexibility to include generic parametrizations of observable halo profiles and mass-observable relations. The current simplified treatment of the galaxy-matter connection for galaxy clustering and intrinsic alignments will be improved by im-

plementing generic perturbation-theory approaches to structure formation (McDonald & Roy 2009; McEwen et al. 2016). A more complete implementation of all relevant cross-correlations between large-scale structure observables and other cosmological probes (e.g. CMB integrated Sachs-Wolfe effect -Sachs & Wolfe 1967-, and other secondary anisotropies) will also soon be included.

In general, although this document presents the functionality and performance of CCL shortly after its release, we expect the library to be a continuously evolving piece of software. This will allow CCL to satisfy the analysis needs of future large datasets, as well as more accurate and sophisticated models for a broad range of cosmological and astrophysical observables.

## ACKNOWLEDGMENTS

This paper has undergone internal review in the LSST Dark Energy Science Collaboration. **Thank the reviewers.**

The DESC acknowledges ongoing support from the Institut National de Physique Nucléaire et de Physique des Particules in France; the Science & Technology Facilities Council in the United Kingdom; and the Department of Energy, the National Science Foundation, and the LSST Corporation in the United States. DESC uses resources of the IN2P3 Computing Center (CC-IN2P3–Lyon/Villeurbanne - France) funded by the Centre National de la Recherche Scientifique; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; STFC DiRAC HPC Facilities, funded by UK BIS National E-infrastructure capital grants; and the UK particle physics grid, supported by the GridPP Collaboration. This work was performed in part under DOE Contract DE-AC02-76SF00515. MI acknowledges support by NSF under grant AST-1517768.

We would like to thank the organisers of the the DESC collaboration meetings at: Oxford (July 2016), SLAC (February 2018, March 2016), and ANL (2015), and the LSST-DESC Hack Week organisers (CMU, November 2016), where this work was partly developed. We would also like to acknowledge the contribution of the participants of the TJP Code Comparison Project, some of whom are among the CCL contributors, for providing the benchmarks for testing CCL. We also acknowledge Louis Penafiel and Elizabeth Kimura, who developed the VARRIC code to compare and visualize power spectra calculated by CCL and CLASS. Finally, we are grateful for the feedback received from other working groups of DESC, including Strong Lensing, Supernovae and Photometric Redshifts.



**And any acknowledgements you would like to add.** D. Alonso is supported by the Science and Technology Facilities Council (STFC) through an Ernest Rutherford Fellowship, grant reference ST/P004474/1. E. Chisari acknowledges support from a Beecroft fellowship and a Royal Astronomical Society Research Fellowship.

Author contributions are listed below.

Husni Almoubayyed: Reviewed code/contributed to issues.

David Alonso: Co-led project; developed structure for angular power spectra; implemented autotools; integrated into LSS pipeline; contributed to: background, power spectrum, mass function, documentation and benchmarks; reviewed code

Jonathan Blazek: Planning capabilities and structure; documentation and testing.

Philip Bull: Implemented and documented the Python wrapper; extensive infrastructure work, including maintenance of the build system and unit tests, bug fixes, and code review; enhanced the error handling system; performed power spectrum comparison across parameter space; various architecture/design contributions.

Jean-Eric Campagne: Angpow main builder.

N. Elisa Chisari: Co-led project, coordinated hack projects & communication, contributed to: correlation function & power spectrum implementation, documentation, and comparisons with benchmarks.

Alex Drlica-Wagner: Helped with document preparation.

Zilong Du: Implemented the 3d correlation function.

Tim Eifler: Reviewed/tested code.

John Ellison: Implemented the 3d correlation function; documentation of 3d correlation function.

Renée Hlozek: Contributed initial code for error handling structures, reviewed other code edits.

Mustapha Ishak: Contributed to planning of code capabilities and structure; reviewed code; identified and fixed bugs; wrote text for this note.

Shahab Joudaki: Contributed to implementation of background functions and documentation.

Matthew Kirby: Performed comparison of physical constants.

David Kirkby: Writing, testing and reviewing code. Asking questions.

Elisabeth Krause: Initiated and co-led project; developed CLASS interface and error handling; contributed to other code; reviewed pull requests.

C. Danielle Leonard: Wrote and tested code for LSST specifications, user-defined photo-z interface, and support of neutrinos; reviewed other code; wrote text for this note.

Christiane S. Lorenz: Contributed to accurate high-redshift cosmological background quantities and benchmarked background splines.

Phil Marshall: Helped with document preparation.

Thomas McClintock: Wrote Python and doxygen documentation.

Sean McLaughlin: Wrote doxygen documentation and fixed bugs/added functionality to distances.

Alexander Mead: Wrote halo-model code and documentation

Jérémy Neveu: Contributed to Angpow, built the interface with CCL and the comparisons with benchmarks.

Stéphane Plaszczynski: Contributed to Angpow.

Javier Sanchez: Modified setup.py to allow pip installation and uninstall.

Sukhdeep Singh: Contributed to the correlation functions code.

Anže Slosar: Wrote and reviewed code.

Tilman Tröster: Wrote code for user-changable precision parameters and distance tests, found and fixed bugs.

Antonio Villarreal: Contributed to initial benchmarking, halo mass function code, and general code and issues review.

Michal Vrástl: Wrote documentation and example code, reviewed code.

Erika L. Wagoner: Set up logarithmic binning in wave number.

Joe Zuntz: Wrote initial infrastructure, C testing setup, and reviewed code.

## REFERENCES

- Alonso, D., Bull, P., Ferreira, P. G., Maartens, R., & Santos, M. G. 2015, *ApJ*, 814, 145
- Alonso, D., & Ferreira, P. G. 2015, *PhRvD*, 92, 063525
- Angulo, R. E., Springel, V., White, S. D. M., et al. 2012, *MNRAS*, 426, 2046
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, *ApJ*, 304, 15
- Bartelmann, M., & Schneider, P. 2001, *PhR*, 340, 291
- Beringer, J., et al. 2012, *Phys. Rev.*, D86, 010001
- Blas, D., Lesgourgues, J., & Tram, T. 2011, CLASS: Cosmic Linear Anisotropy Solving System, Astrophysics Source Code Library, ascl:1106.020

- Blazek, J., MacCrann, N., Troxel, M. A., & Fang, X. 2017, ArXiv e-prints, arXiv:1708.09247
- Bonvin, C., & Durrer, R. 2011, PhRvD, 84, 063505
- Campagne, J.-E., Neveu, J., & Plaszczynski, S. 2017a, A&A, 602, A72
- Campagne, J.-E., Plaszczynski, S., & Neveu, J. 2017b, ApJ, 845, 28
- Carroll, S. M. 2001, Living Reviews in Relativity, 4, 1
- Castorina, E., Carbone, C., Bel, J., Sefusatti, E., & Dolag, K. 2015, JCAP, 7, 043
- Challinor, A., & Lewis, A. 2005, PhRvD, 71, 103010
- . 2011, PhRvD, 84, 043516
- Chang, C., Jarvis, M., Jain, B., et al. 2013, MNRAS, 434, 2121
- Chevallier, M., & Polarski, D. 2001, International Journal of Modern Physics D, 10, 213
- Chon, G., Challinor, A., Prunet, S., Hivon, E., & Szapudi, I. 2004, MNRAS, 350, 914
- Copeland, E. J., Sami, M., & Tsujikawa, S. 2006, International Journal of Modern Physics D, 15, 1753
- DES Collaboration, Abbott, T. M. C., Abdalla, F. B., et al. 2017, ArXiv e-prints, arXiv:1708.01530
- Desjacques, V., Jeong, D., & Schmidt, F. 2016, ArXiv e-prints, arXiv:1611.09787
- Dodelson, S. 2004, Modern Cosmology, Vol. 57, 60–61, doi:10.1063/1.1784308
- Durrer, R. 2008, The Cosmic Microwave Background (Cambridge University Press)
- Eifler, T., Krause, E., Dodelson, S., et al. 2015, MNRAS, 454, 2451
- Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
- Frigo, M., & Johnson, S. G. 2012, FFTW: Fastest Fourier Transform in the West, Astrophysics Source Code Library, ascl:1201.015
- Ghosh, B., Durrer, R., & Sellentin, E. 2018, ArXiv e-prints, arXiv:1801.02518
- Green, J., Schechter, P., Baltay, C., et al. 2011, ArXiv e-prints, arXiv:1108.1374
- Hamilton, A. J. S. 2000, MNRAS, 312, 257
- Heitmann, K., Bingham, D., Lawrence, E., et al. 2016, ApJ, 820, 108
- Hellwing, W. A., Schaller, M., Frenk, C. S., et al. 2016, MNRAS, 461, L11
- Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454
- Hirata, C. M., Mandelbaum, R., Ishak, M., et al. 2007, MNRAS, 381, 1197
- Hirata, C. M., & Seljak, U. 2004, PhRvD, 70, 063526
- Ishak, M. 2007, Foundations of Physics, 37, 1470
- Joachimi, B., & Bridle, S. L. 2010, A&A, 523, A1
- Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43
- Joudaki, S., Blake, C., Johnson, A., et al. 2018, MNRAS, 474, 4894
- Kamionkowski, M., & Spergel, D. N. 1994, ApJ, 432, 7
- Kitching, T. D., & Heavens, A. F. 2017, PhRvD, 95, 063522
- Krause, E., Eifler, T., & Blazek, J. 2016, MNRAS, 456, 207
- Krause, E., Eifler, T. F., Zuntz, J., et al. 2017, ArXiv e-prints, arXiv:1706.09359
- Lattanzi, M., & Gerbino, M. 2017, ArXiv e-prints, arXiv:1712.07109
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints, arXiv:1110.3193
- Lawrence, E., Heitmann, K., Kwan, J., et al. 2017, ApJ, 847, 50
- Lesgourgues, J., & Pastor, S. 2012, ArXiv e-prints, arXiv:1212.6154
- Lewis, A., & Bridle, S. 2002, PhRvD, 66, 103511
- Linder, E. V. 2003, Physical Review Letters, 90, 091301
- LSST Dark Energy Science Collaboration. 2012, ArXiv e-prints, arXiv:1211.0310
- LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, ArXiv e-prints, arXiv:0912.0201
- McDonald, P., & Roy, A. 2009, JCAP, 8, 020
- McEwen, J. E., Fang, X., Hirata, C. M., & Blazek, J. A. 2016, JCAP, 9, 015
- Mohammed, I., & Gnedin, N. Y. 2017, ArXiv e-prints, arXiv:1707.02332
- Mohammed, I., & Seljak, U. 2014, MNRAS, 445, 3382
- Ng, K.-W., & Liu, G.-C. 1999, International Journal of Modern Physics D, 8, 61
- Padmanabhan, T. 2003, PhR, 380, 235
- Peacock, J. A. 1999, Cosmological Physics, 704
- Peacock, J. A., & Smith, R. E. 2000, MNRAS, 318, 1144
- Peebles, P. J., & Ratra, B. 2003, Reviews of Modern Physics, 75, 559
- Sachs, R. K., & Wolfe, A. M. 1967, ApJ, 147, 73
- Schneider, A., & Teyssier, R. 2015, JCAP, 12, 049
- Semboloni, E., Hoekstra, H., & Schaye, J. 2013, MNRAS, 434, 148
- Semboloni, E., Hoekstra, H., Schaye, J., van Daalen, M. P., & McCarthy, I. G. 2011, MNRAS, 417, 2020
- Springel, V., Pakmor, R., Pillepich, A., et al. 2017, ArXiv e-prints, arXiv:1707.03397
- Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, ApJ, 761, 152
- Talman, J. 2009, Computer Physics Communications, 180, 332
- Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709

- Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, *ApJ*, 724, 878
- Troxel, M. A., & Ishak, M. 2015, *PhR*, 558, 1
- Upadhye, A., Biswas, R., Pope, A., et al. 2014, *PhRvD*, 89, 103515
- van Daalen, M. P., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011, *MNRAS*, 415, 3649
- van Uitert, E., Joachimi, B., Joudaki, S., et al. 2018, *MNRAS*, arXiv:1706.05004
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *Nature*, 509, 177
- Watson, W. A., Iliev, I. T., D’Aloisio, A., et al. 2013, *MNRAS*, 433, 1230
- Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, *PhR*, 530, 87
- Yoo, J. 2010, *PhRvD*, 82, 083508
- Yoo, J., Fitzpatrick, A. L., & Zaldarriaga, M. 2009, *PhRvD*, 80, 083514
- Zhao, G.-B., Saito, S., Percival, W. J., et al. 2013, *MNRAS*, 436, 2038
- Zumalacárregui, M., Bellini, E., Sawicki, I., Lesgourgues, J., & Ferreira, P. G. 2017, *JCAP*, 8, 019
- Zuntz, J., Paterno, M., Jennings, E., et al. 2014, *CosmoSIS: Cosmological parameter estimation, Astrophysics Source Code Library*, ascl:1409.012