Assessing theoretical uncertainties for cosmological constraints from weak lensing surveys

Ting Tan, Dominik Zürcher, Janis Fluri, Alexandre Refregier, Federica Tarsitano, and Tomasz Kacprzak

1 Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Énergies, LPNHE, 4 Place Jussieu, F-75252 Paris, France
2 Institute for Particle Physics and Astrophysics, Department of Physics, ETH Zürich, Wolfgang Pauli Strasse 27, 8093 Zürich, Switzerland
3 Data Analytics Lab, Department of Computer Science, ETH Zurich Universitätsstrasse 6, 8006 Zürich, Switzerland

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Weak gravitational lensing is a powerful probe which is used to constrain the standard cosmological model and its extensions. With the enhanced statistical precision of current and upcoming surveys, high accuracy predictions for weak lensing statistics are needed to limit the impact of theoretical uncertainties on cosmological parameter constraints. For this purpose, we present a comparison of the theoretical predictions for the non-linear matter and weak lensing power spectra, based on the widely used fitting functions (mead and rev-halofit), emulators (EuclidEmulator, EuclidEmulator2, BaccoEmulator and CosmicEmulator) and N-body simulations (Pkdgrav3). We consider the forecasted constraints on the $\Lambda$CDM and $w$CDM models from weak lensing for stage III and stage IV surveys. We study the relative bias on the constraints and their dependence on the assumed prescriptions. Assuming a $\Lambda$CDM cosmology, we find that the relative agreement on the $S_8$ parameter is between 0.2 – 0.3σ for a stage III-like survey between the above predictors. For a stage IV-like survey the agreement becomes 1.4 – 3.0σ. In the $w$CDM scenario, we find broader $S_8$ constraints, and agreements of 0.18 – 0.26σ and 0.7 – 1.7σ for stage III and stage IV surveys, respectively. The accuracies of the above predictors therefore appear adequate for stage III surveys, while the fitting functions would need improvements for future stage IV weak lensing surveys. Furthermore, we find that, of the fitting functions, mead provides the best agreement with the emulators. We discuss the implication of these findings for the preparation of the future weak lensing surveys.

Key words: weak gravitational lensing – cosmological parameters from LSS – large-scale structure of Universe

1 INTRODUCTION

The next generation of wide field cosmological surveys, such as LSST, Euclid, and NGRST will map the matter distribution of the local Universe with an unprecedented accuracy. These high precision measurements present a challenge for the theoretical modeling of cosmological observables. Cosmic shear is a cosmological observable that relies on the distortions of galaxy shapes caused by weak gravitational lensing (eg. Bartelmann & Schneider 2001). This effect is due to the gravitational deflection of photons by the matter density field along the line of sight. Cosmic shear measures the inhomogeneities in the cosmic density field with high precision and can be used as an unbiased tracer of the matter distribution.

It is sensitive to both, the matter distribution of the Universe and the growth of cosmic structure, which is important for the understanding of the expansion history of the Universe. A commonly used cosmic shear summary statistic is the cosmic shear angular power spectrum, which can be predicted from the matter power spectrum. The modeling of the matter power spectrum on large scales can be derived using perturbation theory (Blas et al. 2014; Bernard et al. 2002; Crocce & Scoccimarro 2006; Croce et al. 2012; Blas et al. 2016; Nishimichi et al. 2016; Baumann et al. 2012; Cataneo et al. 2019; Foreman & Senatore 2016; Beutler et al. 2017; d'Amico et al. 2020), where the structure formation of the Universe is linear. However, at non-linear, small scales with $k \gtrsim 1$Mpc$^{-1}$, nonlinear processes have a strong impact on the matter power spectrum, and perturbation theory is no longer valid.

In this work, we compare the theoretical predictions of the nonlinear matter power spectrum, and the associated theoretical uncertainties on cosmological parameters from mea-
measurements of the cosmic shear angular power spectrum. The comparison includes some widely used models fitted from N-body simulations using analytical halo models: halofit (Smith et al. 2003) is fitted to low resolution, gravity-only N-body simulations, which is known to exhibit a non-negligible mismatch with current state-of-the-art hydrodynamic N-body simulations: rev-halofit (Takahashi et al. 2012), developed as the revisited version of halofit is used in the analysis of the Dark Energy Survey (DES) (Trroxel et al. 2018a); and mead (Mead et al. 2015), which is used in the analysis of the Kilo-Degree Survey (KiDS) combined with the VISTA Kilo-Degree Infrared Galaxy Survey (VIKING) (Hildebrandt et al. 2020). Apart from the halo model fitting method, emulators are generated from the interpolation of a suite of N-body simulations, e.g. CosmicEmulator (Heitmann et al. 2009; Lawrence et al. 2017; Heitmann et al. 2013), BaccoEmulator (Angulo et al. 2020; Aricò et al. 2021), EuclidEmulator (Knabenhans et al. 2019) and its updated version EuclidEmulator2 (Collaboration et al. 2020), COSMOPower (Mancini et al. 2021) and GP emulator (Giblin et al. 2019). In this study, CosmicEmulator, BaccoEmulator, EuclidEmulator and EuclidEmulator2 are representatively selected in the comparison at the level of the matter power spectrum, and a comparison between rev-halofit, mead and EuclidEmulator is also shown in Knabenhans et al. (2021). In order to estimate the theoretical uncertainties, we look at the weak lensing cosmological parameter constraints, by generating a forecast for a stage III, DES-like survey and a stage IV, Euclid-like survey. We take into account the parameters described by the standard ΛCDM cosmological model and the extended w/CDM model.

This paper is organised as follows. In Section 2 we describe the theoretical framework, including three halo-model based fitting functions, mead, halofit and rev-halofit: four power spectrum emulators extracted from N-body simulations: CosmicEmulator, BaccoEmulator, EuclidEmulator and EuclidEmulator2, and one N-body simulation code Pkdgrav3 (Potter et al. 2017). In Section 3 we present the method and the relevant codes used in this study. We summarize our results in Section 4 and our conclusions in Section 5.

2 THEORY

In this section, we describe the theoretical background of the matter power spectrum, weak lensing and its angular power spectrum, as well as the different predictors of the matter power spectrum that we include in the comparisons.

2.1 Weak Lensing

Considering the cosmic density field ρ(⃗r) at the position ⃗r, the density contrast δ(⃗r) is defined as the relative difference of ρ(⃗r) to the average density $\bar{\rho}$

$$\delta(\vec{r}) = \frac{\rho(\vec{r}) - \bar{\rho}}{\bar{\rho}}.$$  (1)

In Fourier space, the density contrast takes the following form

$$\delta(\vec{k}) = \int \delta(\vec{r}) \exp(i\vec{k} \cdot \vec{r})d^3r.$$  (2)

Furthermore, the matter power spectrum $P(\vec{k})$ is defined as the correlation of the density contrast in Fourier space (Peebles 1980):

$$\langle \delta(\vec{k})\delta(\vec{k}') \rangle = (2\pi)^3 \delta^3_D(k + k') P(k),$$  (3)

where $\delta_D$ is the three dimensional Dirac delta function.

For full-sky surveys, the cosmic shear angular power spectrum is approximately identical to the convergence power spectrum (Bartelmann & Maturi 2016), which can be defined as a weighted integration along the line-of-sight over the matter power spectrum (Bartelmann & Schneider 2001), and simplified using the Kaiser-Limber approximation (LoVerde & Afshordi 2008; Limber 1953; Kaiser 1992, 1998). We follow the formalism of (LoVerde & Afshordi 2008; Kitching et al. 2017; Kilbinger et al. 2017; Tarsitano et al. 2020; Giannantonio et al. 2012) to compute the cross-correlated shear power spectrum with tomographic redshift bins i and j:

$$C_{ij}(\ell) = \frac{9}{16} \left(\frac{H_0}{c}\right)^4 \Omega_m^2 \int_0^{\chi_h} d\chi P_{NL}(\ell, \chi) \frac{g_i(\chi)g_j(\chi)}{(ar(\chi))^2}.$$  (4)

Here $P_{NL}$ is the non linear matter power spectrum, $\chi$ is the comoving distance, $\chi_h$ is the comoving horizon distance, $\Omega_m$ is the total matter density, $a = (1 + z)^{-1}$ is the scale factor and $g(\chi)$ is the lensing efficiency function defined as:

$$g(\chi) = 2 \int_0^{\chi_h} d\chi' n_i(\chi') \frac{r(\chi')}{r(\chi)},$$  (5)

with $n_i(\chi)$ being the normalized number density of the observed galaxies at a comoving distance $\chi$.

2.2 Matter Power Spectrum

The matter power spectrum is a fundamental statistics to study the large scale structure of the Universe. As seen above, it is, in particular, useful to predict the cosmic shear angular power spectrum. Therefore, it is necessary to have an accurate theoretical model for the matter power spectrum on all scales. On large scales and mildly non-linear scales, the matter power spectrum can be modeled using perturbation theory and some extended theories (Martinelli et al. 2021). On small scales, which are in the non-linear regime, these approaches are not suited to predict the power spectrum with the necessary precision, while other methods are developed with the use of halo model or simulations.

2.2.1 Analytical Predictions

A common way to model the matter power spectrum on these small scales is to empirical fit physically motivated formulas to measurements from N-body simulations, e.g. as done in Hamilton et al. (1991). Furthermore, modelling the density field as a collection of virialized halos, the matter power spectrum can be approximated analytically using the statistics of halos, and fitted to simulations or emulators (Ma & Fry 2000; Seljak 2000; Cooray & Sheth 2002).

In this study, we compare 3 halo-model based fitting functions: mead, halofit, and rev-halofit. halofit was built using a series of N-body simulations with a total of $N = 256^3$ particles and the box size from 84Mpc/h to 240Mpc/h. Using

$$MNRAS 000, 1–16 (2021)$$
the halo model, the matter power spectrum is constructed with two terms, the one-halo term proposed by Peacock & Smith (2000); Ma & Fry (2000); Seljak (2000); Scoccimarro et al. (2001) and a two-halo term (Ma & Fry 2000; Seljak 2000; Scoccimarro et al. 2001) to describe the exclusion effects between dark matter halos. The one-halo term indicates the correlation of the matter field of one single halo, which dominates on small scales, while the two-halo term describes the cross-correlation between different halos, that has a strong impact on larger scales. Assuming that the halos are distributed according to the halo mass function (Press & Schechter 1974; Sheth & Tormen 1999), the matter power spectrum modelled with this approach can achieve a high precision on large scales. However, due to the lack of baryons and the relatively low resolution of the N-body simulations, giving an accuracy at the level for $k = 1 h Mpc^{-1}$ (Heitmann et al. 2010), and larger differences for $k > 1 h Mpc^{-1}$, which is insufficient for the non-linear regime. *rev-halofit* is a revised prescription of *halofit*, which provides a more accurate prediction of the matter power spectrum for $k < 30 h Mpc^{-1}$ and $z < 10$, with a 5% level accuracy at $k = 1 h Mpc^{-1}$ and 10% level accuracy at $k = 10 h Mpc^{-1}$. *rev-halofit* uses high resolution N-body simulations for 16 cosmological models around the Wilkinson Microwave Anisotropy Probe (WMAP) best-fit cosmological parameters. The N-body simulations were run with the Gadget-2 N-body code (Springel et al. 2001; Springel 2005a), 1024$^3$ particles in total, and the box size from 3200Mpc$/h$ to 20000Mpc$/h$. The power spectrum is fitted using an improved fitting formula with 5 more model parameters as compared to *halofit*. Several extended methods have been proposed to improve the halo model (Bird et al. 2012; Mohammed & Seljak 2014; Seljak & Vlah 2015). Here we only consider *mead* (Mead et al. 2015), which reaches an accuracy at the 5% percent level for $k = 10 h Mpc^{-1}$ and $z < 2$. *mead* introduces more physical parameters in addition to the halo model, and is fitted to the “Coyote Universe” (Heitmann et al. 2013) suite of high resolution simulations, the same simulations used for the generation of CosmicEmulator. It also includes massive neutrinos (Mead et al. 2016) and baryonic effects e.g. active galactic nuclei (AGN) feedback, supernovae explosions, and gas cooling. However, we only consider the dark-matter-only case in this study.

### 2.2.2 Emulators

The fitting functions based on halo models described in Section 2.2.1 can provide accurate non-linear power spectrum predictions for large $k$-modes and a wide redshift range, which can be used to predict cosmological observables. However, they also have limitations as the precision is not uniform for different cosmological parameters, and it is difficult for fitting functions to give a high precision below the 1% level compared to high resolution simulations. Power spectrum emulators are constructed following a different approach in which one interpolates the power spectrum from a set of N-body simulations within a certain range of relevant parameters, using interpolation methods, e.g. Gaussian Processes Regression (Heitmann et al. 2010, 2013; Angulo et al. 2020) or polynomial chaos expansion (Knabenhans et al. 2019; Collaboration et al. 2020). Compared to fitting functions, emulators usually provide consistent precision of the predictions for different $k$-modes. However, emulators also have limitations: Firstly, the covered parameter space is limited, thus making it difficult to perform a likelihood analysis, for which one needs to explore a wide range of parameter values. Secondly, the ranges of $k$ and redshift are also limited, making it difficult to compute the weak lensing cosmic shear observables for high $\ell$s, which requires an integration over a large $k$ range.

In this study, we compare 4 emulators: CosmicEmulator (Heitmann et al. 2016), BaccoEmulator (Angulo et al. 2020), EuclidEmulator (Knabenhans et al. 2019), and EuclidEmulator2 (Collaboration et al. 2020). CosmicEmulator is fitted using a set of the “Coyote Universe” simulations and the “Mira-Titan Universe” (Lawrence et al. 2017) simulations. We use the latest version of the emulator (Heitmann et al. 2016), for which the “Mira-Titan Universe” simulations were run with 3200$^3$ particles and a simulation volume of $(2100h^{-1} h Mpc)^3$. The CosmicEmulator successfully achieves high precision predictions of the power spectrum within the 4% level for $k_{max} = 5 h Mpc^{-1}$ and $z < 2$. It allows for the variation of various parameters, including the matter density $\Omega_m$, the amplitude of density fluctuations $\sigma_8$, the baryon density $\Omega_b$, the scalar spectral index $n_s$, the dark energy equation of state parameters $w_0$ and $w_a$, the dimensionless Hubble parameter $h$, the neutrino density $\Omega_\nu$, and the redshift $z$. EuclidEmulator uses a different emulator method using N-body simulations generated with the PkdGrav3 code (Potter et al. 2017). It uses 100 simulations with 2048$^3$ particles in a $(1250 h^{-1} h Mpc)^3$ simulation volume. The non-linear correction is encoded as a boost factor adding up to the input linear power spectrum, achieving a precision at the 1% level for predictions within the ranges $k < 1 h Mpc^{-1}$ and $z < 1$. Knabenhans et al. (2019) demonstrated that EuclidEmulator agrees with *rev-halofit* at the 8% level. As an updated version of EuclidEmulator, EuclidEmulator2 is extended with dynamical dark energy and massive neutrinos, created with a larger parameter space and a modified version of the PkdGrav3 N-body code. EuclidEmulator2 provides a consistent accuracy with simulations at the 2% level up to $k_{max} = 10 h Mpc^{-1}$ for $z < 2$, and slightly lower accuracy for higher redshift $z \sim 3$. However, as EuclidEmulator2 uses the amplitude of the primordial power spectrum $A_s$ instead of $\sigma_8$ as input parameter, we use the following formula to transfer $\sigma_8$ into $A_s$: \[
A_s = \left(\frac{\sigma_8}{\sigma_{8,0}}\right)^2 \times A_{s,0}
\]
in our comparison, where $\sigma_{8,0} = 0.826$ and $A_{s,0} = 2.184 \times 10^{-9}$. BaccoEmulator is another state-of-the-art emulator using an updated version of the L-Gadget3 code (Springel 2005b; Angulo et al. 2012) with 4320$^3$ particles in a $(1440 h^{-1} h Mpc)^3$ simulation volume. It has a 2% level accuracy over the redshift range $0 < z < 1.5$ and $k < 5 h Mpc^{-1}$.

### 2.2.3 N Body Simulations

We also include in this study a comparison with a dark-matter-only N-body simulation run with PkdGrav3, which is based on a binary tree algorithm. This code uses 5th order interpolation.
Table 1. Parameter settings for mock surveys: The Stage IV survey is created using a 4 times larger survey area and galaxy density compared to the Stage III Survey. A deeper Smail redshift distribution is also used in the Stage IV Survey.

| Survey | Stage III | Stage IV |
|--------|-----------|----------|
| Survey Area [deg²] | 5’000 | 20’000 |
| Galaxy Density [arcmin⁻²] | 5 | 20 |
| Redshift Distribution | Smail | Smail |
| Redshift Bins | 4 | 4 |
| Redshift Range⁴ | 0.025 ~ 3.0 | 0.025 ~ 3.0 |

multipole expansions of the gravitational potential between particles and can achieve fast computational speeds with hardware acceleration. A comparison between PkdGrav3 and the N-body codes, Gadget-3, Gadget-4 and Ramses is presented in Schneider et al. (2016) and Springel et al. (2020). The PkdGrav3 simulations are the same as the ones used for EuclidEmulator, with 2048³ particles in total and the box size of \( L = 1250h^{-1}\)Mpc. The details are presented in Knabenhans et al. (2019).

3 METHOD

In this work, we perform a comparison of predictors of the nonlinear matter power spectrum, i.e. halo-model based fitting functions and emulators. We estimate the theoretical uncertainties of these predictors on the parameter constraint level by looking at the weak lensing cosmological parameter constraints from a stage III survey and a stage IV survey. For each survey, we perform a comparison using the standard ΛCDM cosmological model and the extended wCDM model.

3.1 Survey parameters

The estimate of the theoretical uncertainties for cosmological parameters is realised by forecasting the constraints for a stage III survey and a stage IV survey. The covariance matrix is estimated from simulations, using the NGSF code described in Zürcher et al. (2021) and DES Collaboration et al. (2021). We refer the reader to Zürcher et al. (2021) for a detailed description of the method. Table 1 shows the parameter settings used for the generation of the mock galaxy surveys. Martinelli et al. (2021) suggests using \( \ell_{\text{max}} = 5000 \) for stage IV-like surveys to probe deep into non-linear regime. However, in this study we use a more conservative limit of \( \ell_{\text{max}} = 1000 \), and do not take into account baryonic effects.

We use Smail et al. (1995) distributions to model the global redshift distribution of the source galaxies for both the stage

³ The presented redshift range refers to the considered range used in the generation process of the covariance matrix for the mock surveys. The range differs from the redshift range used for the predictors of the weak lensing power spectrum in Section 4.2, where we use [0.08, 2] for the stage III survey and [0.08, 3.0] for the stage IV survey.

Figure 1. The redshift distributions of the source galaxies. One can see the four tomographic distributions for the stage III survey and the stage IV survey. The global distributions, that follows the Smail et al. (1995) model, is shown by the dashed lines.

III survey and the stage IV survey. The corresponding formulas and parameter settings for these two distributions are as follows

\[
n(z)_{\text{stageIII}} = z^\alpha \exp\left[-(\frac{z}{z_0})^\beta\right],
\]

with \( \alpha = 1.5, \beta = 1.1 \) and \( z_0 = 0.31 \) and

\[
n(z)_{\text{stageIV}} = \left(\frac{z}{z_0}\right)^\alpha \exp\left[-(\frac{z}{z_0})^\beta\right],
\]

with \( \alpha = 2.0, \beta = 1.5 \) and \( z_0 = 0.64 \) (Martinelli et al. 2021). In both cases the source galaxies are divided into four tomographic bins with equal number of galaxies in each bin. As a result of the auto- and cross-combinations of these four redshift bins, we have 10 combinations of auto- and cross-correlations for the cosmic shear measurements (4 auto-correlations and 6 cross-correlations). Figure 1 shows the global and tomographic redshift distributions used in this study.

3.2 Covariance Matrix

An accurate estimate of the survey covariance matrix is crucial for the correct calculation of the likelihood function. We estimate the covariance matrices for the stage III and stage IV survey setups described in Table 1 from numerical simulations. We generate a large number (\( N = 2000 \)) of realisations of the angular power spectra for each survey setup following the methodology outlined in Zürcher et al. (2021). In the following, we introduce the used N-Body simulations, briefly summarize the forward modelling procedure used to generate the angular power spectra and describe the estimation of the covariance matrix. We refer the reader to Zürcher et al. (2021) for a more detailed description of the methodology.

We utilise the 50 independent PkdGrav3 (Potter et al. 2017) N-Body simulations at the fiducial cosmology that were previously used in Zürcher et al. (2021); DES Collaboration et al. (2021) and generated using the state-of-the-art dark-matter-only N-body code PkdGrav3. The cosmological parameters in the used simulations are fixed to the (ΛCDM, TT, TE, EE+lowE+lensing) results of Planck 2018.
(Aghanim et al. 2020), except for $\Omega_m$ and $\sigma_8$ which are set to the values found in Troxel et al. (2018b). This setup results in $\Omega_m = 0.26$, $\sigma_8 = 0.84$, $\Omega_b = 0.0493$, $n_s = 0.9649$, $w = -1$ and $h = 0.6736$. We include three massive neutrino species in all simulations. The neutrinos are modelled as a relativistic fluid (Tram et al. 2019) and a degenerate mass hierarchy with a minimal neutrino mass of $m_{\nu} = 0.02$ eV per species was chosen. The dark energy density $\Omega_\Lambda$ is adapted for each cosmology to achieve a flat geometry.

Each simulation was run using a unit box with a side-length of 900 Mpc$/h$ and $768^3$ simulated particles. In order to achieve a simulation volume large enough to cover the redshift range up to $z = 3.0$ the unit box was replicated up to 14 times per dimension depending on the cosmology. While such a replication scheme is known to underpredict the variance of very large, super-box modes (Fluri et al. 2019), it has been demonstrated by DES Collaboration et al. (2021) that the simulations accurately recover the angular power spectra predicted by the theory code CLASS (Lesgourgues 2011) for $\ell \in [30, 2048]$.

The particle shells from each PKDGRAV3 simulation are combined into tomographic full-sky mass maps using the U\textsc{Falcon} software (Sgier et al. 2019). The particle shells are weighted according to the tomographic redshift distributions shown in Figure 1. The U\textsc{Falcon} software uses the HEALPix (Gorski et al. 2005) pixelization scheme to pixelize the sphere. A resolution of NSIDE = 1024 was chosen. U\textsc{Falcon} also makes use of the Born approximation, which is known to deteriorate the accuracy of the produced mass maps. However, Petri et al. (2017) have demonstrated that the introduced bias is negligible for stage III-like and stage IV-like surveys.

The spherical Kaiser-Squires mass mapping technique (Kaiser & Squires 1993; Wallis et al. 2017) is used to obtain the cosmic shear signal from the simulated mass maps. To forward-model a realistic weak lensing survey a shape noise signal must then be added to the cosmic shear signal and an appropriate survey mask must be applied. The survey masks are chosen such that we obtain eight stage III surveys and two stage IV surveys from each full-sky map.

The shape noise signal is obtained in the same way as described in Zürcher et al. (2021). We randomly sample galaxy positions within the survey region until the target source density is reached. The intrinsic ellipticities of the galaxies are then drawn from a probability distribution that was fit to the observed galaxy ellipticities in Troxel et al. (2018b) (see Zürcher et al. (2021)). The ellipticity of each individual galaxy is rotated by a random phase. Using five and twenty shape noise realisations per survey patch, we achieve the desired number of $N = 2000$ survey realisations for the stage III and stage IV survey setup, respectively.

The tomographic angular power spectra realisations $C_{\ell,i}$ are then measured from the forward-modelled surveys using the anafast routine of the healpy software (Zonca et al. 2019) using 20 bins from $\ell_{\text{min}} = 100$ to $\ell_{\text{max}} = 1000$, the same as Sgier et al. (2019), where the index $i$ runs over the number of survey realisations $N$. The covariance matrix $\Sigma$ is estimated according to
\begin{equation}
\hat{\Sigma} = \frac{1}{N-1} \sum_{i=1}^{N} (C_{\ell,i} - \bar{C}_{\ell})(C_{\ell,i} - \bar{C}_{\ell})^T, \tag{9}
\end{equation}
where $\bar{C}_{\ell}$ indicates the mean of the angular power spectra realisations $C_{\ell,i}$. The estimated correlation matrices $C_{n,m} \equiv \Sigma_{n,m}/\sqrt{\Sigma_{n,n}\Sigma_{m,m}}$ are presented in Figure 2.

### 3.3 Likelihood Analysis

We use a Bayesian likelihood approach to evaluate the cosmological parameter constraints of different predictors. We assume a Gaussian error model and the likelihood is realized by:
\begin{equation}
\log \mathcal{L} = -\frac{1}{2} \sum_{ij} (C_{\ell,\text{truth}} - C_{\ell,\text{compare}})^T \hat{\Sigma}^{-1} (C_{\ell,\text{truth}} - C_{\ell,\text{compare}}) \tag{10}
\end{equation}

Here $C_{\ell,\text{truth}}$ stands for the value of the observable, computed using PyCosmo (Refregier et al. 2017; Tarsitano et al. 2020; Moser et al. 2022) with a chosen predictor and the fiducial cosmological parameters, measured by the Wilkinson Microwave Anisotropy Probe satellite (WMAP) 9 (Hinshaw et al. 2013), presented in Table 2. $C_{\ell,\text{compare}}$ is predicted using another predictor for comparison. The cosmology for the observable is different from what is used for the covariance matrix. However, this effect is neglected assuming the covariance matrix parameter independent (Kodwani et al. 2018). $\Sigma^{-1}$ is the unbiased estimate of the inverse covariance matrix (Hartlap, J. et al. 2007; Percival et al. 2014) represented as:
\begin{equation}
\hat{\Sigma}^{-1} = \frac{N - N'}{N - 1} \Sigma^{-1}, \tag{11}
\end{equation}

$N$ is the number of realisations generated from the simulations and $N'$ is the total number of data bins, which is given by
\begin{equation}
N' = N_{\text{redshift}} \times N_{\ell}. \tag{12}
\end{equation}

Here we have $N = 2000$, $N_{\ell} = 20$ and $N_{\text{redshift}} = 10$. 

| Parameters | Fiducial values | Priors (stage-III survey) | Priors (stage-IV survey) |
|------------|----------------|---------------------------|--------------------------|
| $\Omega_m$ | 0.291          | [0.0, 0.6]                | [0.2, 0.4]               |
| $n_s$      | 0.969          | [0.3, 2.0]                | [0.7, 1.2]               |
| $h$        | 0.69           | [0.1, 2.5]                | [0.4, 0.9]               |
| $\sigma_8$ | 0.826          | [0.3, 1.4]                | [0.7, 0.95]              |
| $w_0$      | -1.0           | [-3.5, 0.5]               | [-2.5, 0.5]              |
| $\Omega_b$ | 0.0473         |                           |                         |
3.4 Parameter Inference

The posterior is sampled efficiently using the Markov Chain Monte Carlo (MCMC) ensemble sampler, emcee (Foreman-Mackey et al. 2013). We vary 4 cosmological parameters ($\Omega_m, \sigma_8, n_s, h$) for the $\Lambda$CDM cosmological model and an additional parameter $w_0$ for the extended $\omega$CDM model, where we fix $w_a \equiv 0$. Table 2 shows the priors used for these parameters. We run the MCMC chains with 100 walkers per parameter and cut the burn in phase for each run as one third of the chain length. Each individual chain has more than 100,000 samples. For the visualisation of the marginalised posteriors, we use the public Getdist (Lewis 2019).

4 RESULTS

We present the results of our comparison of different predictors in this section, including the analysis of the matter power spectrum, the weak lensing power spectrum, and the cosmological parameter constraints based on the stage III and stage IV weak lensing surveys.

4.1 Power Spectrum

We use the linear power spectrum predicted by PyCosmo and generated following Eisenstein & Hu (1999) as the input for all predictors. Figure 3 shows the comparison of dark-matter-only non-linear $P(k)$ predictions from different predictors at redshift $z = 0$, and the comparison for different redshifts ranging from $z = 0$ to $z = 5$ in Appendix A. The results are shown for $k$ ranging from $k = 0.01 h$Mpc$^{-1}$ to $9h$Mpc$^{-1}$ using 10000 bins. BaccoEmulator and CosmicEmulator are not valid for $z > 3$, so we do not present their comparison for the higher redshift at $z = 5$. Figure 3 and Figure A1 indicate that:

- All the predictors except for halofit are within the 5% level of accuracy compared to rev-halofit for $z < 2$ and $k < 7h$Mpc$^{-1}$ (BaccoEmulator is valid for $z < 1.5$ and $k < 5h$Mpc$^{-1}$, see the details in Figure A1). Note that this is consistent with the comparison of mead, rev-halofit and halofit in Mead et al. (2015).
- halofit shows stronger discrepancies compared with the other predictors at small scales for $k > 0.1h$Mpc$^{-1}$ and this discrepancy can reach 20% for $k \sim 10h$Mpc$^{-1}$.
- mead and rev-halofit show close agreement with the emulators at the 5% level for $k < 9h$Mpc$^{-1}$ and $z < 0.5$. However, at higher redshifts $1 < z < 5$, the discrepancies between mead and the emulators can reach 10% for $k > 3h$Mpc$^{-1}$, while rev-halofit provides a more consistent precision within 5%.
- All the emulators yield an agreement within the 2 – 3% level compared with the Pkdgrav3 simulation for $k < 9h$Mpc$^{-1}$ and $z < 1.5$. However, this is not valid at higher redshifts.
- For large scales with $k < 0.5h$Mpc$^{-1}$, the different predictors show a better agreement at higher redshifts.

4.2 Weak Lensing Power Spectrum

We compute the weak lensing shear power spectrum $C_\ell$ for the Stage III survey and the Stage IV survey with different predictors. Limited by the range of $k_{max}$ of the emulators, the $C_\ell$s are computed using 20 $\ell$-bins spaced linearly between $\ell_{\min} = 100$ and $\ell_{\max} = 1000$. The integrated redshift range is [0.08, 2.0] for the stage III survey and [0.08, 3.0] for the stage IV survey. This setting was chosen in order to avoid the instability of emulators for low redshifts, where we found that EuclidEmulator and EuclidEmulator2 predict the $C_\ell$s with a discrepancy larger than 10% at $z < 0.08$. This choice differs from the setting used for the generation of the covariance matrix. However, we find that this only changes the discrepancies between different predictors for $C_\ell$s by 0.1%, since only 1% of the low-redshift galaxies are missed for the stage III sur-
In the individual panels, we present the left-hand panels showing the results for the Stage III survey and the right-hand side showing the Stage IV survey results. For each survey, we perform a comparison for the stage IV survey as it allows only up to this redshift range, we have to exclude rev-halofit as reference.

Figure 3. Comparison of dark-matter-only, non-linear $P(k)$ predictions for different predictors at redshift $z = 0$, subtracted and divided by rev-halofit as reference.

We present the two-dimensional 68% and 95% confidence level contours of the posterior distributions for the ΛCDM model in Figure 5 and Figure 6 for the Stage 3 and Stage 4 survey setup, respectively. The parameters $\{\Omega_m, \sigma_8, n_s, h\}$ are varied in the MCMC analysis. We additionally compute the constraints on $S_8$ and summarise the shifts in $S_8$ in Figure 9, presenting the median values of the posteriors and the error bars indicating the 68% confidence limits of the constraints. One can infer from the posterior distributions in Figure 9 and Table B1 that the agreement on $S_8$ between different predictors is less than 0.6σ for the stage III survey (0.2 – 0.3σ if halofit excluded), while being much larger for the stage IV survey. This is caused by the higher constraining power of the IV survey. More specifically, the agreements are generally on the 1.4 – 6.1σ level (1.4 – 3.0σ if halofit excluded). mead shows good agreement with CosmicEmulator, EuclidEmulator2 and EuclidEmulator, while rev-halofit exhibits a larger discrepancy.

The comparison of $C_\ell$ for different predictors does not show a significant difference between the stage III survey and the stage IV survey.

### 4.3 Cosmological Parameters Constraints

#### 4.3.1 ΛCDM cosmology constraints

We consider the constraining power of weak lensing surveys on dark energy parameters by adopting a time dependent dynamical dark energy equation of state, the CPT-parameterisation (Chevallier & Polarski 2001; Linder 2003),

#### 4.3.2 $w$CDM cosmology constraints

The comparison of the weak lensing cosmological parameter constraints for different predictors is present in this section. As indicated in Section 3, we consider a stage III survey and a stage IV survey. For each survey, we perform a comparison using the standard ΛCDM cosmological model and the extended $w$CDM model. A summary of the constraints on $\{S_8, \Omega_m, \sigma_8\}$ is presented in Table 3, and the constraints on $\{S_8, \Omega_m, n_s, h, w_0\}$ in Table B1.
as an extension to the $\Lambda$CDM model. The equation of state parameter is given by
\begin{equation}
    w(a) = w_0 + w_a(1-a).
\end{equation}

where we use a fixed $w_a = 0$ and a free $w_0$. We present the two-dimensional marginal posterior distributions for the $w$CDM cosmology parameters in Figure 7 and Figure 8, for the stage III survey and the stage IV survey, respectively. Taking into account the dark energy model changes the shape and the contour size of the posterior distributions, decreasing the constraining power on the cosmological parameters. The discrepancies in $S_8$ between predictors are generally smaller compared with the $\Lambda$CDM model due to the decrease in constraining power: $0.18 - 0.34\sigma$ for the stage III survey and $0.7 - 2.4\sigma$ for the stage III survey ($0.18 - 0.26\sigma$ and $0.7 - 1.7\sigma$ if halofit is excluded, respectively). mead still shows good agreement with EuclidEmulator and EuclidEmulator2 for both the stage III survey and the stage IV survey. rev-halofit agrees with all the predictors within $0.3\sigma$ for the stage III survey, and shows discrepancies at the $0.7 - 2.4\sigma$ level for the stage IV survey.

5 CONCLUSIONS

The different halo-model based fitting functions and emulators have been widely used for the prediction of non-linear power spectrum to study the large scale structure of the Universe. It is essential to understand their advantages, limitations, and theoretical uncertainties for different surveys and cosmologies. From our results, we conclude that:

- Compared with Pkdgrav3 simulations, the halo-model based fitting functions, except halofit, yield a $5 - 10\%$ level accuracy for the matter power spectrum $P(k)$ for $k < 9h$Mpc$^{-1}$ and $z < 2$, while emulators show better precision at the $2\%$ level. For the weak lensing shear power spectrum $C_{\ell}$, all the predictors, except for halofit, show a $5\%$ level mutual agreement.

- For the stage III survey with a $\Lambda$CDM cosmology, the agreement on $S_8$ between different predictors are within $0.6\sigma$, and within $0.2\sigma$ for other cosmological parameters ($0.3\sigma$ and $0.2\sigma$ if we exclude halofit, respectively). This indicates the applicability of the studied predictors for the stage III surveys.

- For the stage IV survey using a $\Lambda$CDM cosmology, the disagreements on $S_8$ are increased to several $\sigma$s, with the largest
Figure 5. Cosmological parameter constraints for the stage III survey in the $\Lambda$CDM model. For each constraint, $C_{\ell,\text{truth}}$ is predicted using the first predictor shown in the legend, and $C_{\ell,\text{compare}}$ computed using the second predictor, as indicated in Section 3.3. For the stage III survey, we set $C_{\ell,\text{truth}}$ with the halo-model based fitting functions (rev-halofit, mead and halofit) and 3 emulators (EuclidEmulator, EuclidEmulator2 and CosmicEmulator), and compare with predictions from only the fitting functions (in this figure only rev-halofit).

- The accuracy of the current fitting function models and emulators therefore appear sufficient for stage III surveys. However, for the future IV surveys, our results suggest that the fitting function models are currently not sufficiently accurate, and would need further improvements in the future. For emulators, it is required to explore wider ranges of cosmological parameters, $k$-modes, and redshifts, while pursuing the discrepancy of 6.1σ between rev-halofit and halofit, and the best agreement between mead and EuclidEmulator2.

- If $w_0$ is taken into account for the $w$CDM cosmology, we get weaker constraints on $S_8$, and the discrepancies between different predictors are reduced to $0.2 - 0.3\sigma$ and $0.7 - 2.4\sigma$ for the stage III survey and the stage IV survey respectively ($0.18 - 0.26\sigma$ and $0.7 - 1.7\sigma$ if we exclude halofit, respectively).
consistent precision with reliable hydrodynamic N-body simulations.

Note that, in this study, we include dark-matter-only predictions, without any consideration of baryonic effects, which can have a strong impact on small scales (Jing et al. 2006; Rudd et al. 2008). Current studies of halo-model based fitting functions already include other systematics, i.e. massive neutrino and baryonic effects like AGN feedback and gas cooling. The inclusion of these systematics will significantly reduce the constraining power, and might alleviate the discrepancies between the predictors. There are also other sources of uncertainties in weak lensing experiments that we did not include in this work and that could affect the our results, e.g. photometric redshift uncertainty (Hildebrandt et al. 2020; Huterer et al. 2006; Choi et al. 2016), shear bias (Bernstein & Jarvis 2002; Hirata et al. 2004; Bernstein 2010; Refregier et al. 2012; Melchior & Viola 2012) and galaxy intrinsic alignment (Heavens et al. 2000; Fluri et al. 2019; Hirata & Seljak 2004; Bridle & King 2007; Joachimi et al. 2011).
Theoretical uncertainties for weak lensing

Figure 7. Cosmological parameter constraints of the stage III survey in the \( \Lambda \)CDM cosmological model. Including \( w_0 \) reduces significantly the constraining power, yielding much broader contours than the \( \Lambda \)CDM model.

ACKNOWLEDGMENTS

This work was supported in part by grant 200021_192243 from the Swiss National Science Foundation. We thank Mischa Knabenhans from University of Zürich for the distribution of Pkdgrav-3. We further thank Aurel Schneider from the University of Zürich for the useful discussions regarding this project and the covariance matrix for a stage IV survey. We would also like to thank Uwe Schmitt from ETH Zürich for his support with the GitLab server and development of PyCosmo. The Collaborating Institutions are the Eidgenössische Technische Hochschule (ETH) Zürich, Ecole Polytechnique, the Laboratoire de Physique Nucléaire et des Hautes Energies of Sorbonne University.
Figure 8. Cosmological parameter constraints of the stage IV survey in the $w$CDM cosmological model. The discrepancies between the predictors are alleviated, taking into account a simple $w$CDM cosmological model with a varying $w_0$.

DATA AVAILABILITY

Most of the analysis in this work is done on the Euler cluster\(^5\) operated by ETH Zurich. Here follows the computational codes used in this study: PyCosmo (Refregier et al. 2017; Tarsitano et al. 2020; Moser et al. 2022) is used as the main tool where all the non linear codes are implemented for the computation of auto (cross) power spectra, galaxy reshift distribution counts, and observable of cosmic shear. It is also extended to include interfaces with the emulators. Anafast is used for computation of power spectra from simulations, and all the the maps (masks, weight, shear, mass) in pipeline are in HealPix format. We use Emcee-3.0.2 (Foreman-Mackey et al. 2013) for the sampling of parameter space and Getdist (Lewis 2019) for the plotting of likelihood contours and Uhammer for the simplification of Emcee running. Some of the results in this paper have been derived using the healpy and HEALPix packages (Gorski et al. 1999). In this

\(^5\) https://scicomp.ethz.ch/wiki/Euler
Figure 9. Deviations of the parameter constraints on $S_8$. The upper plot shows the result for the stage III survey, for the $\Lambda$CDM model (black) and the $\omega$CDM model (green), respectively. The lower plot shows the stage IV survey, for the $\Lambda$CDM (red) and $\omega$CDM (blue), respectively.
study, we made use of the functionalities provided by numpy \citep{vanderWalt2011}, scipy \citep{Virtanen2020} and matplotlib \citep{Hunter2007}.

Table 3. Numerical constraints on the cosmological parameters corresponding to the contours in Figure 5, 6, 7, and 8. For each predictor, the $\sigma$s show the theoretical discrepancies for each parameter, compared to the reference one.

| Survey Cosmology | Predictor | $S_N$ ($\sigma$) | $\Omega_m$ ($\sigma$) | $w_0$ ($\sigma$) |
|------------------|-----------|------------------|----------------------|-----------------|
| stage-III $\Lambda$CDM | rev-halofit | 0.841$^{0.0241}_{-0.0203}$ | 0.298$^{+0.0617}_{-0.0648}$ | 0.11 |
| | mead | 0.803$^{0.0260}_{-0.0202}$ | 0.2996$^{+0.0731}_{-0.0698}$ | 0.025 |
| | halofit | 0.7946$^{0.0292}_{-0.0201}$ | 0.2884$^{+0.074}_{-0.074}$ | 0.125 |
| | euclid | 0.8083$^{0.0256}_{-0.0201}$ | 0.2987$^{+0.0759}_{-0.0759}$ | 0.1 |
| | cosmicemu | 0.8042$^{0.0285}_{-0.018}$ | 0.2916$^{+0.0789}_{-0.0789}$ | 0.03 |
| | euclid2 | 0.8031$^{0.0269}_{-0.0177}$ | 0.2887$^{+0.0835}_{-0.0679}$ | 0.01 |
| stage-III $w$CDM | rev-halofit | 0.8165$^{0.0433}_{-0.0061}$ | 0.2846$^{+0.092}_{-0.09}$ | $-0.9242^{+0.4704}_{-0.224}$ |
| | mead | 0.7947$^{0.0495}_{-0.0058}$ | 0.315$^{+0.075}_{-0.102}$ | $-1.139^{+0.6581}_{-0.224}$ |
| | halofit | 0.7879$^{0.0517}_{-0.0062}$ | 0.2968$^{+0.0787}_{-0.101}$ | $-1.1333^{+0.6581}_{-0.224}$ |
| | euclid | 0.7977$^{0.0544}_{-0.0052}$ | 0.3049$^{+0.083}_{-0.101}$ | $-1.1886^{+0.7187}_{-0.224}$ |
| | cosmicemu | 0.7982$^{0.0564}_{-0.0062}$ | 0.2931$^{+0.0921}_{-0.102}$ | $-1.1408^{+0.6926}_{-0.224}$ |
| | euclid2 | 0.8018$^{0.0641}_{-0.00627}$ | 0.2896$^{+0.0692}_{-0.0677}$ | $-1.0254^{+0.5498}_{-0.2745}$ |
| stage-IV $\Lambda$CDM | rev-halofit | 0.8135$^{0.0023}_{-0.0024}$ | 0.2915$^{+0.0077}_{-0.0084}$ | $-0.1171^{+0.1171}_{-0.1171}$ |
| | mead | 0.8024$^{0.0027}_{-0.0026}$ | 0.3008$^{+0.0094}_{-0.0074}$ | $-0.1046^{+0.1171}_{-0.1171}$ |
| | halofit | 0.7941$^{0.0029}_{-0.0018}$ | 0.2856$^{+0.0064}_{-0.0079}$ | $-0.1046^{+0.1171}_{-0.1171}$ |
| | euclid | 0.8094$^{0.0033}_{-0.0032}$ | 0.2917$^{+0.0084}_{-0.0079}$ | $-0.1046^{+0.1171}_{-0.1171}$ |
| | euclid2 | 0.8058$^{0.0032}_{-0.0032}$ | 0.2926$^{+0.0084}_{-0.0084}$ | $-0.1046^{+0.1171}_{-0.1171}$ |
| stage-IV $w$CDM | rev-halofit | 0.8127$^{0.0079}_{-0.0063}$ | 0.2909$^{+0.0099}_{-0.0099}$ | $-1.0217^{+0.1171}_{-0.1171}$ |
| | mead | 0.7960$^{0.0067}_{-0.0059}$ | 0.2979$^{+0.0106}_{-0.0099}$ | $-1.106^{+0.1171}_{-0.1171}$ |
| | halofit | 0.7900$^{0.0077}_{-0.0072}$ | 0.2856$^{+0.0094}_{-0.0094}$ | $-1.0646^{+0.1171}_{-0.1171}$ |
| | euclid | 0.8060$^{0.0073}_{-0.0072}$ | 0.2908$^{+0.0088}_{-0.0088}$ | $-1.040^{+0.1171}_{-0.1171}$ |
| | euclid2 | 0.7990$^{0.0078}_{-0.0069}$ | 0.2901$^{+0.0099}_{-0.0094}$ | $-1.0965^{+0.1288}_{-0.1288}$ |

REFERENCES

Aghanim N., et al., 2020, Astronomy & Astrophysics, 641, A6
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, Monthly Notices of the Royal Astronomical Society, 426, 2046
Angulo R. E., Zennaro M., Contreras S., AricÃ­ G., Peellejo-IbaÃ­ez M., StÃ¡jek J., 2020, The BACCO Simulation Project: Exploiting the full power of large-scale structure for cosmology (arXiv:2004.06245)
Aricó G., Angulo R. E., Contreras S., Ondaro-Mallea L., Pellejero-Ibanae M., Zennaro M., 2021, Monthly Notices of the Royal Astronomical Society, 506, 4070
Bartelmann M., Maturi M., 2016, Weak gravitational lensing (arXiv:1612.06535)
Bartelmann M., Schneider P., 2001, Physics Reports, 340, 291
Baumann D., Nicolis A., Senatore L., Zaldarriaga M., 2012, Journal of Cosmology and Astroparticle Physics, 2012, 051
Bernardeau F., Colombi S., GaztaÃ±aga E., Scoccimarro R., 2002, Physics Reports, 367, 1
Bernstein G. M., 2010, Monthly Notices of the Royal Astronomical Society, 406, 2793
Bernstein G., Jarvis M., 2002, The Astronomical Journal, 123, 583
Beutler F., et al., 2017, Monthly Notices of the Royal Astronomical Society, 464, 3409
Bird S., Viel M., Haehnelt M. G., 2012, Monthly Notices of the Royal Astronomical Society, 420, 2551
Blas D., Garny M., Konstandin T., 2014, Journal of Cosmology and Astroparticle Physics, 2014, 010
Blas D., Garny M., Ivanov M. M., Sibiryakov S., 2016, Journal of Cosmology and Astroparticle Physics, 2016, 052
Brdile S., King L., 2007, New Journal of Physics, 9, 444
Cataneo M., Lombriser L., Heymans C., Mead A., Barreira A.,
APPENDIX A: POWER SPECTRUM COMPARISON

In this section we present the comparison of the non-linear power spectrum for all redshifts, as shown in Figure A1.

APPENDIX B: COSMOLOGICAL PARAMETER CONSTRAINTS

The summary of constraints on \( \{S_8, \Omega_m, n_s, h, w_0\} \) is concluded in this section, shown in Table B1.

This paper has been typeset from a \TeX/L\LaTeX file prepared by the author.
Figure A1. The comparison of the dark-matter-only non linear $P(k)$ of different predictors at different redshifts ($z = 0, 0.5, 1, 1.5, 2$ and $5$), subtracted and divided by rev-halofit as reference. BaccoEmulator and CosmicEmulator are not valid for $z > 3$, so we do not take them into comparison for $z = 5$. 
| Survey Cosmology | Predictor | $S_8$ ($\sigma$) | $\Omega_m$ ($\sigma$) | $n_s$ ($\sigma$) | $h$ ($\sigma$) | $w_0$ ($\sigma$) |
|-----------------|-----------|-----------------|----------------------|-----------------|----------------|----------------|
| **stage-III**<br>ACDM | rev-halo fit | 0.8147 $\pm$ 0.0241 | 0.286 $\pm$ 0.0178 | 0.9741 $\pm$ 0.2489 | 0.6736 $\pm$ 0.5642 | 0.4772 |
| | mead | 0.8053 $\pm$ 0.0269 | 0.33 $\pm$ 0.0084 | 0.11 $\pm$ 0.2425 | 0.6855 $\pm$ 0.5779 | 0.02 |
| | halofit | 0.7946 $\pm$ 0.0201 | 0.57 $\pm$ 0.0743 | 0.00 $\pm$ 0.2566 | 0.6777 $\pm$ 0.6698 | 0.01 |
| | euclid | 0.8083 $\pm$ 0.0256 | 0.2 $\pm$ 0.0709 | 0.1 $\pm$ 0.2428 | 0.6644 $\pm$ 0.5612 | 0.01 |
| | cosmicemu | 0.8047 $\pm$ 0.0285 | 0.29 $\pm$ 0.0741 | 0.03 $\pm$ 0.2613 | 0.7452 $\pm$ 0.5542 | 0.1 |
| | euclid2 | 0.8031 $\pm$ 0.0269 | 0.34 $\pm$ 0.0835 | 0.01 $\pm$ 0.2496 | 0.7467 $\pm$ 0.5503 | 0.1 |
| **stage-III**<br>uCDM | rev-halo fit | 0.8165 $\pm$ 0.0433 | 0.2868 $\pm$ 0.0026 | 0.9164 $\pm$ 0.5799 | 0.7868 $\pm$ 0.9823 | 0.9242 $\pm$ 0.4704 |
| | mead | 0.7947 $\pm$ 0.0578 | 0.26 $\pm$ 0.0082 | 0.18 $\pm$ 0.5012 | 0.6527 $\pm$ 0.4233 | 0.11 |
| | halofit | 0.7879 $\pm$ 0.0517 | 0.34 $\pm$ 0.1022 | 0.09 $\pm$ 0.4514 | 0.6199 $\pm$ 0.3779 | 0.13 |
| | euclid | 0.7977 $\pm$ 0.0545 | 0.22 $\pm$ 0.0985 | 0.15 $\pm$ 0.4723 | 0.6209 $\pm$ 0.3873 | 0.13 |
| | cosmicemu | 0.7982 $\pm$ 0.0572 | 0.22 $\pm$ 0.1017 | 0.06 $\pm$ 0.5093 | 0.6688 $\pm$ 0.4301 | 0.08 |
| | euclid2 | 0.8018 $\pm$ 0.0461 | 0.18 $\pm$ 0.0978 | 0.04 $\pm$ 0.5729 | 0.7461 $\pm$ 0.5134 | 0.04 |
| **stage-IV**<br>ACDM | rev-halo fit | 0.8135 $\pm$ 0.0023 | 0.2915 $\pm$ 0.0077 | 0.9696 $\pm$ 0.0178 | 0.6894 $\pm$ 0.0481 | 0.9453 |
| | mead | 0.8028 $\pm$ 0.0026 | 2.96 $\pm$ 0.0093 | 0.87 $\pm$ 0.203 | 0.7181 $\pm$ 0.0495 | 0.5 |
| | halofit | 0.7944 $\pm$ 0.0029 | 6.11 $\pm$ 0.0072 | 0.46 $\pm$ 0.203 | 0.7134 $\pm$ 0.0094 | 0.05 |
| | euclid | 0.8094 $\pm$ 0.0018 | 1.37 $\pm$ 0.0084 | 0.02 $\pm$ 0.198 | 0.7058 $\pm$ 0.0050 | 0.00 |
| | cosmicemu | 0.8058 $\pm$ 0.0017 | 2.62 $\pm$ 0.0084 | 0.1 $\pm$ 0.195 | 0.6958 $\pm$ 0.0019 | 0.00 |
| | euclid2 | 0.8127 $\pm$ 0.0039 | 2.9097 $\pm$ 0.0095 | 0.045 $\pm$ 0.437 | 0.6848 $\pm$ 0.0509 | 0.01 |
| | mead | 0.7968 $\pm$ 0.0061 | 1.73 $\pm$ 0.0100 | 0.53 $\pm$ 0.0441 | 0.6905 $\pm$ 0.0494 | 0.13 |
| | halofit | 0.7902 $\pm$ 0.0073 | 2.39 $\pm$ 0.0096 | 0.42 $\pm$ 0.0466 | 0.6958 $\pm$ 0.0494 | 0.13 |
| | euclid | 0.8061 $\pm$ 0.0072 | 0.68 $\pm$ 0.0049 | 0.01 $\pm$ 0.0553 | 0.6977 $\pm$ 0.0094 | 0.00 |
| | cosmicemu | 0.7996 $\pm$ 0.0059 | 1.31 $\pm$ 0.0094 | 0.06 $\pm$ 0.0515 | 0.6771 $\pm$ 0.0548 | 0.21 |

Table B1. Complete numerical constraints on the cosmological parameters corresponding to the contours in Figure 5, 6, 7, and 8. For each predictor, the $\sigma$s show the theoretical discrepancies for each parameter, compared to the reference one.