Universiteit
Leiden
The Netherlands

Playing dice with the universe: Bayesian statistical analyses of cosmological models and new observables

Cañas Herrera, G.

Citation

Cañas Herrera, G. (2022, October 19). *Playing dice with the universe: Bayesian statistical analyses of cosmological models and new observables*. *Casimir PhD Series*. Retrieved from https://hdl.handle.net/1887/3483658

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3483658

Note: To cite this publication please use the final published version (if applicable).

Part III

Data Science for the Euclid mission

Chapter 6

Cosmological Likelihood for Observables in *Euclid*

We can use observations of the large scale structure (LSS) of the universe, as galaxy clustering (GC) and weak lensing (WL), to extract cosmological information (see chapter 1). If we want to exploit all the information encoded in LSS, we need large data sets of galaxies that contain not only their positions in the sky and their shapes but also their redshifts. These large data sets form the so-called galaxy redshift surveys. Some examples of well-known galaxy redshifts surveys are the *Center for Astrophysics Redshift Survey* (CfA) (Tonry & Davis, 1979) and the *Sloan Digital Sky Survey* (SDSS) (Lundgren et al., 2015). These surveys have successfully shown that we can use the information contained in the large scale structure of the universe to improve our cosmological knowledge and constrain not only cosmological parameters but also be able to discern among several cosmological models.

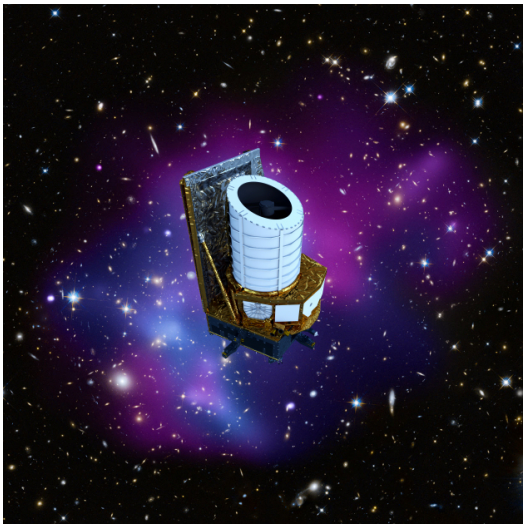

Figure 6.1: Artist's impression of the *Euclid* spacecraft.

In this chapter, we will focus on the future *Euclid* mission, a near-infrared space telescope currently under development by the European Space Agency and the Euclid Consortium, and for which the author of this thesis has dedicated a vast percentage of her doctoral time. In particular, the author works as one of the main *software developers* for the official Bayesian analysis pipeline that will be eventually used to produce the constraints on the cosmological parameters once the data is available: the *Cosmological Likelihood for Observables in Euclid* (`CLOE`).

This chapter is organized as follows. In section 6.1 we present the mission, main goals, structure and technical details of the telescope. In section 6.2 we introduce the features of the future *Euclid* surveys and we show briefly how the photometric and spectroscopic techniques works. In section 6.3 we explain in mathematical detail the theoretical description and modelling of the primary observational probes. In section 6.4 we detail how the fiducial benchmark data and covariance matrices are generated and later used to validate the Bayesian analysis pipeline. In section 6.6 we formally introduce `CLOE`: structure, features and dependencies. In section 6.7 we show the results of the validation of `CLOE` in its earliest version release. Finally, we conclude with the outline of the future plans for `CLOE` in section 6.8.

6.1 The *Euclid* Mission

6.1.1 Telescope, goals and the Euclid Consortium

Euclid[1] is an European Space Agency (ESA) medium-class mission part of the *Cosmic Vision* 2015-2025 program. According to the main scientific objectives listed in the mission's study definition report[2] (Laureijs et al., 2011), the ultimate goals of *Euclid* are to understand the physical origin of the accelerated expansion of the Universe, whose responsible agent is thought to be Dark Energy, and to understand the nature of Dark Matter. Moreover, *Euclid* is meant to study the initial conditions that seeded the early universe and were responsible for the formation of the cosmic structure and ultimately, it will also try to understand the range of validity of General Relativity. To achieve its goals, *Euclid* will map the LSS of the universe by creating one of the largest galaxy catalogues ever, and this is why *Euclid* is considered primarily a cosmological mission.

Euclid satellite's launch is planned for 2024[3] using an Ariane 62 and will travel up to the L2 Sun-Earth Lagrangian point to be on duty for a 6-years mission. It will explore the expansion history of the Universe and the evolution of the LSS by measuring the position, shapes and redshifts of galaxies. For that, it will incorporate a 1.2-meter telescope and two scientific instruments that will observe in the optical and near-infrared bands: a high-quality panoramic optic visible imager (VIS), a near-infrared 3-filter (Y, J and H) photometer (NISP-P) with a slitless spectrograph (NISP-S). The spectrograph consists on one 'blue' grism ($920 - 1250\ nm$) and three 'red' grisms ($1250 - 1850\ nm$ placed on three different orientations). The optical and near-infrared instruments share a common field-of-view of 0.53 deg^2. *Euclid* will cover eventually 15000 deg^2 of the sky up to redshift $z \approx 2$. The primary result product of the mission will be a complete redshift galaxy catalogue up to magnitude 24 in the Y, J, and H bands. It will collect enough information to study the mission's primary observational probes: 30 million spectroscopic redshifts used for spectroscopic galaxy clustering measurements and 2 billion photometric galaxy images, which will be used

[1] Up-to-date information about the mission can be found at `https://www.Euclid-ec.org`

[2] The study definition report receives the popular name of the "Red Book".

[3] The launch of the satellite is yet unknown due to consequences of the war conflict in Ukraine.

for weak lensing and photometric galaxy clustering observations (Amendola, Appleby, et al., 2018; Euclid Collaboration et al., 2020).

Apart from the primary probes, *Euclid* will provide additional cosmological observables such as the cross-correlation with CMB observations, strong lensing, abundance and properties of galaxy clusters and even luminosity distance measurements through supernovae Ia. Moreover, *Euclid* will complement the cosmological observables with astrophysical and astronomical probes of high scientific interest (i.e: cool brown dwarfs, stellar populations in the galaxy and the universe nearby, High-z Lyman Break Galaxies...).

In June 2012, ESA named the *Euclid Consortium* (EC) as the team responsible for the mission, the data production and the scientific instruments (VIS and NISP). The scientific exploitation and interpretation of the massive *Euclid* data collection will be led by the scientists of the EC as well. It comprises approximately 1500 scientists with various backgrounds (i.e: astronomers and astrophysicists, theoretical physicists, engineers, technicians, and managers...). The EC contains researchers from 14 European countries working in their national space agencies or associated research institutes (Austria, Belgium, Denmark, Finland, France, Germany, Italy, the Netherlands, Norway, Portugal, Romania, Spain, Switzerland and the United Kingdom), scientists from Canada and USA (through NASA and other US laboratories) and few scientists from Japan. The EC is led by the Euclid Consortium Lead (ECL) and the Euclid Consortium Board (ECB), which have also the role of acting as contact points between the EC members and ESA. The activity of the EC is organized in several groups (Tutusaus Lleixa, 2018, September):

- the EC Science Ground Segment (SGS), which is responsible for the design, test, integration and operation of the data processing tools and pipelines and whose activities are sub-arranged in Organizational Units (OUs). Each of the OUs is dedicated to one specific task, for instance, providing the measurements of the positions of galaxies.

- the EC Science Working Groups (SWGs), which have the responsibility of the scientific production and delivery of Euclid data releases and their scientific exploitation. There are three different types of SWGs: the cosmology SWGs (weak lensing, galaxy clustering, galaxy cluster, theory...), the legacy SWGs (exoplanets, Milky Way...) and the cosmological simulations SWG. The work within the SWGs is further structured in Working Packages (WPs).

With *Euclid's* launch fast approaching, it is necessary to improve and update the current cosmological forecasts for the scientific performance of the satellite to assess the impact of the design and technical description of the mission. The first *Euclid* forecasts were shown in the definition study report (Laureijs et al., 2011). Years later, the EC multidisciplinary group "Inter-Science Working Group Taskforce Forecast" (also known as "IST:Forecast") (Euclid Collaboration et al., 2020) unified the implemented definitions of the different *Euclid* primary probes: updating the

information presented in the Red Book by including the latest astrophysical knowledge. Besides, they also validated the data analysis methodology and computational tools to produce Fisher forecasts and verify their robustness. The work carried out by "IST:Forecast", which focused their analysis on the ΛCDM model, remains the current state-of-the-art for the scientific performance of the mission.

Still, as mentioned in chapter 1, section 1.7, it is customary to use Bayesian statistics as the analysis framework for testing cosmological models against data to infer the probability distributions of the parameters of a model. Therefore, the construction of a *software* that allows us to perform a full Bayesian statistical analysis not to only go beyond the "Fisher forecast" scenario but to analyse the incoming data is crucial for the *Euclid* mission. Using a full Bayesian statistics approach will open the possibility of forecasting the scientific performance of the satellite not only for the Standard Cosmological Model but also for extensions of ΛCDM, where, for instance, the posterior distributions of the parameters of interest sometimes cannot be easily approximated by a Gaussian function. To be able to move towards a full Bayesian statistical analysis, it is necessary to construct a software that computes the *likelihood* distribution given the model and its parameters, and some (fiducial) data. Furthermore, we need to model the theoretical observables to make predictions: weak lensing (WL), photometric galaxy clustering (GCph), spectroscopic galaxy clustering (GCsp) and their combinations.

Additionally, to unleash the full power of *Euclid* primary observational probes, it is essential to exploit all the information contained in the non-linear regime of matter density perturbations; that is, at late times and small scales ($k > 0.1$ Mpc). At these scales, the matter distribution is affected by the non-linear evolution of density fluctuations, which induces changes in the shape of the matter power spectrum beyond the predictions of linear perturbation theory (see subsection 1.4.1 for more information). Therefore, the EC is giving special attention to the need of modelling with precision the theoretical predictions of *Euclid* primary observables for the non-linear scales to make the best use of the future *Euclid* data.

6.1.2 The role of IST:L and IST:Non-Linear

The work presented in this chapter, which is a summary of the science contained in a series of upcoming *Euclid* key-project papers, represents the results of the intensive activity of code design, development, review and testing carried out by several scientists from different Science Working Groups within the EC. The work has been conducted by a novel and special multidisciplinary group known as the "Inter-Science Taskforce Likelihood" (IST:L). This group is in charge of developing the official likelihood and *Euclid* software that will produce the theoretical predictions for *Euclid* primary probes and the computation of the *Euclid* likelihood, and will also produce the official constraints on the probability distributions of the cosmological parameters. The results of this work activity, which currently has been taking place for approximately three years, is the software *Cosmological Likelihood for Observables in Euclid* (also known as `CLOE`).

`CLOE`[4], which is currently available only for EC members, can compute the theoretical predictions of the primary probes given the implementation recipes provided by the cosmology Science Working Groups. Furthermore, `CLOE` can be attached to the Bayesian analysis framework `Cobaya` as an *external likelihood* to perform full sampling of the posterior probability distributions of the parameters of interest (see section 6.6). In May 2021 `CLOE` v.1.0 was released, including the first implementation for the theoretical observables' recipe only up to the linear regime. A version 1.1 appeared soon after after fixing a series of bugs and implementing functionalities in the user interface. By the end of 2022, `CLOE` v.2.0 will be available within the EC including a revisited recipe of the observables including corrections for the non-linear scales.

The multidisciplinary group in charge of correcting the primary observables for the non-linear scales is the "Inter-Science Working Group Taskforce Non-Linear" (IST:NL). If these corrections are not implemented, the theoretical models will be incomplete and will eventually bias the the statistical analysis. This is why the work of IST:NL is crucial for `CLOE` to succeed and provide correct constraints of the cosmological parameters. IST:NL is responsible for the non-linear modelling as well as the theory covariances, and they will create the "IST:NL model library" that interfaces with `CLOE`. So far, IST:NL has implemented external codes able to provide non-linear corrections: `bacco` (Aricò, Angulo, & Zennaro, 2021; Angulo et al., 2021; Aricò, Angulo, Contreras, et al., 2021), `euclid emulator 2` (Euclid Collaboration et al., 2021) and `fastPT` (McEwen, Fang, Hirata, & Blazek, 2016; Fang, Blazek, McEwen, & Hirata, 2017). Therefore, IST:L is responsible for the computation of the theoretical predictions for the primary observables and the likelihood, and IST:NL will include the non-linear corrections by implementing the non-linear codes within `CLOE`.

The work in IST:L and IST:NL is organized following an *Agile* inspired philosophy where the working tasks are split into smaller sub-tasks that are assigned to different sub-groups, which use a `SCRUM` framework for developing, delivering, and sustaining software (see (Abrahamsson, Salo, Ronkainen, & Warsta, 2017) for more details about *Agile*). Each sub-group is composed by a *Scrum master, code developer, code reviewer* and several theory *experts*. The software `CLOE` contains unit testing and combines the practices of continuous integration (CI) and continuous delivery (CD), enforcing automation in building, testing and deployment of `CLOE` (see (*SCRUM: practical guide*, 2022) and (Sane, 2021) for more information about good practices in coding). The author of this chapter works as one of the main *code developers* for `CLOE` within IST:L and as a consultant for IST:NL.

6.2 Photometric and Spectroscopic surveys

Galaxy redshift surveys can be classified based on the technique used to extract the redshift measurements. In this sense, we can mainly classify redshift surveys as

[4]For EC members `CLOE` can be downloaded from `https://gitlab.euclid-sgs.uk/pf-ist-likelihood/likelihood-implementation`.

photometric and spectroscopic ones. Spectroscopic redshift surveys use spectrographs to obtain redshifts. These devices can obtain the spectrum of a cosmic object by separating the incoming light into several narrow bins using dispersion. As seen during chapter 1, section 1.1, the light coming from distance objects is redshifted. To extract the redshift measurements, we need to compare the obtained spectra to the one known from an object of the same class at rest. This approach requires a large collection time and previous information about the angular position of the galaxies for targeting. Photometric redshift surveys, on the other hand, use images to extract low-resolution spectra, where the redshift is still inferred by comparing the obtained spectra with other spectra from similar objects at rest. This method does not require target galaxies and it needs less exposition time than the spectroscopic technique, but provides worse redshift estimates.

When *Euclid* arrives at the L2 Lagrangian point, it will start operating for a 6-years long mission, creating two different surveys. The first survey is the *Euclid* Wide Survey covering 15000deg^2 of the sky up to a magnitude of approximately 24 for both VIS and NISP instruments; thus, it will contain both spectroscopic and photometric redshifts. The contamination coming from the Milky Way as well as the contamination coming from the Solar System will be removed. This survey will be used for the study of the *Euclid* primary probes: weak lensing and galaxy clustering. The other survey will consist of three *Euclid* Deep Fields going to 2 magnitudes deeper than the wide survey but covering only, in total, 40 deg^2. The *Euclid* Deep Fields will be used to calibrate the wide survey, apart from providing precious information to study Active Galactic Nuclei, high redshift galaxies and other objects. It is expected that after the 6 operation years, Euclid will have measured the shape and photometric redshifts of 1.5×10^9 galaxies, creating the photometric *Euclid* survey, and about 5×10^7 galaxy spectroscopic redshifts, used for the spectroscopic *Euclid* survey.

The photometric measurements of the *Euclid* wide survey will be used to study weak lensing (WL) and photometric galaxy clustering (GCph). WL is sensitive to both baryonic and dark matter (BM and DM) and the expansion rate of the universe. The spectroscopic measurements of the *Euclid* wide survey will be employed for the spectroscopic galaxy clustering (GCsp), including baryonic acoustic oscillations (BAO) and redshift space distortions (RSD). GCsp is sensitive to the distribution of matter as well as the expansion of the universe and the growth rate of structures. Combining all the probes, as we will see in section 6.7, will allow us to constrain different cosmological parameters and possibly discern among cosmological models once the real data is available.

Moreover, when the redshift of the galaxies is known, we can extract more cosmological information using *tomography*: we can create slices of two-dimensional projected images to recover the three-dimensional distribution of matter in the universe (Hu, 1999). By dividing the galaxies into redshift tomographic bins, $n_i(z)$, we can obtain the time evolution of LSS. For the *Euclid* mission, the distribution of galaxies of both the photometric and spectroscopic surveys will be divided into redshift bins, which will be used to compute the *Euclid* primary probes. In this first approach,

10 redshift bins for the photometric probes and 4 redshift bins for the spectroscopic probe are used[5].

6.3 Theoretical predictions for *Euclid* primary observables

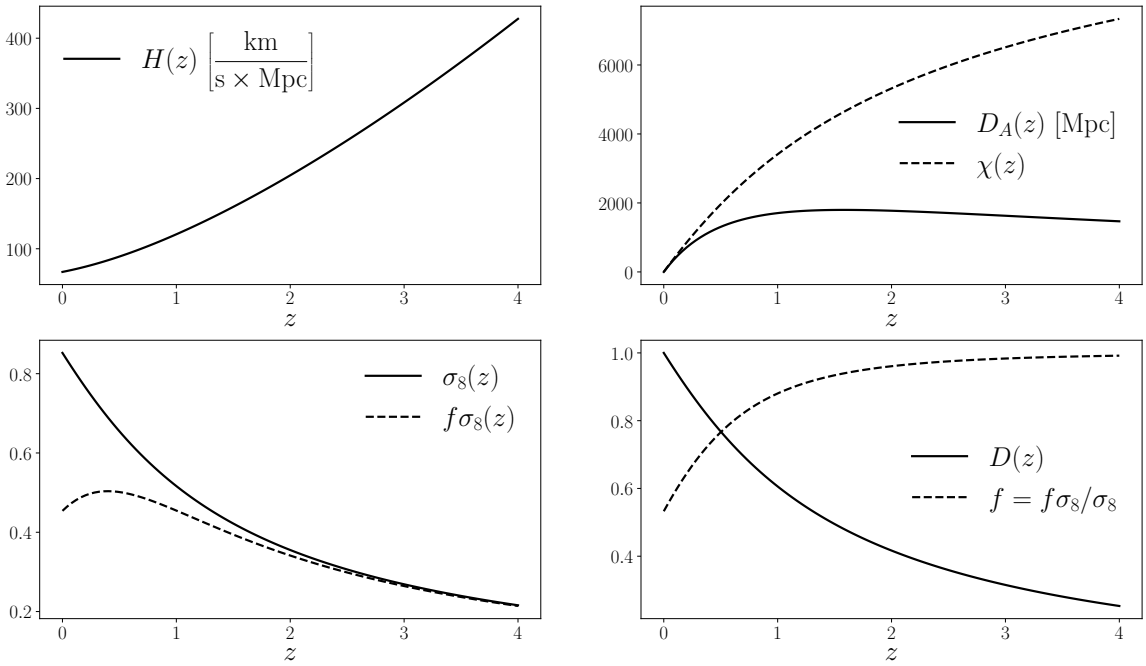

Figure 6.2: Cosmological background quantities requested by `CLOE` to `CAMB` via `Cobaya`, to be used as ingredients of the computation of the *Euclid* primary theoretical observables described in section 6.3. **Top-left** is the Hubble parameter $H(z)$, **top-right** are the comoving $\chi(z)$ and the angular diameter $D_{\mathrm{A}}(z)$ distances, **bottom-left** are the root mean square mass fluctuations amplitude on 8 Mpc/h scales $\sigma_8(z)$ (where h is the reduced Hubble parameter), and the product of σ_8 times the growth rate $f(z)$, and finally, **bottom-right** are the growth factor $D(z)$ and growth rate $f(z)$. All quantities are represented as a function of redshift z. Within `CLOE`, these quantities are obtained in `cobaya_interface.py` and further processed in `cosmo.py`. All plots have been generated taken the fiducial cosmological values presented in section 6.9. All the definitions of these background functions can be found at chapter 1.

In this section, the modelling of the *Euclid* main scientific probes is presented: weak lensing and photometric galaxy clustering for the photometric catalogue with the cross-correlation between the two, and spectroscopic galaxy clustering for the spectroscopic catalogue. For simplicity and completeness, the modelling described

[5]The number of redshift bins is decided by the SWGs and the corresponding OUs, and can be changed at later times

Cosmological probe	Output
Spectroscopic galaxy clustering (GCsp)	
Spectroscopic galaxy bias	$b^{\rm GCsp}(z)$
Redshift-space galaxy-galaxy power spectrum	$P_{\rm gg}^{\rm GCsp}(z,k)$
Perpendicular scaling factor	$q_\perp(z)$
Parallel scaling factor	$q_\parallel(z)$
Cosine of the angle between $\vec{k}$ and the line-of-sight	μ_k
Legendre multipole power spectra ($\ell = m$)	$P_\ell(k,z)$
Weak lensing (WL)	
Galaxy redshift density distribution	$n_i^{\rm WL}$
Intrinsic alignment - intrinsic alignment power spectrum	$P_{\rm IAIA}(z,k)$
Matter density - intrinsic alignment power spectrum	$P_{\delta{\rm IA}}(z,k)$
Intrinsic alignment function	$f_{\rm IA}(z)$
Intrinsic alignment nuisance parameters	$A_{\rm IA}$, $\eta_{\rm IA}$
Shear window function	$W_i^\gamma(z,k)$
Intrinsic alignment window function	$W_i^{\rm IA}(z)$
Intrinsic alignment angular power spectra	$C_{ij}^{\rm IAIA}(\ell)$
Shear angular power spectra	$C_{ij}^{\gamma\gamma}(\ell)$
Weak lensing angular power spectra	$C_{ij}^{\rm WL}(\ell)$
Photometric galaxy clustering (GCph)	
Galaxy redshift density distribution	$n_i^{\rm GCph}$
Galaxy-galaxy power spectrum	$P_{\rm gg}^{\rm GCph}(z,k)$
Photometric galaxy window function	$W_i^{\rm GCph}(z)$
Photometric galaxy bias	$b^{\rm GCph}(z)$
Photometric galaxy clustering angular power spectra	$C_{ij}^{\rm GCph}(\ell)$
Weak Lensing - Photometric Galaxy Clustering Cross-Correlation (XC)	
Galaxy-matter power spectrum	$P_{\delta{\rm g}}^{\rm GCph}(z,k)$
Galaxy-intrinsic alignment power spectrum	$P_{\rm gAI}^{\rm GCph}(z,k)$
Photometric galaxy-matter angular power spectrum	$C_{ij}^{\delta{\rm g}}(\ell)$
Photometric galaxy-intrinsic alignment angular power spectrum	$C_{ij}^{\rm gIA}(\ell)$
Photometric cross-correlation angular power spectrum	$C_{ij}^{\rm XC}(\ell)$

Table 6.1: Summary of the different quantities and symbols used to write the recipe for the *Euclid* primary probes: weak lensing (WL), photometric galaxy clustering (GCph), spectroscopic galaxy clustering (GCsp) and the weak lensing - photometric galaxy clustering cross-correlation (XC). The recipe is written in terms of several different power spectra $P(z,k)$ to simplify the interface with IST:NL, which will provide IST:L with the non-linear corrections for the observables. Note that the notation for the angular power spectra $C_{ij}(\ell)$ is written differently with respect to the notation used in the previous chapters, where here we have decided to write explicitly the dependence on the tomographic redshift bins i,j as subscripts and the symbol of the probe as superscript. The colours used to highlight the different cosmological probes are in concordance with the colours used in the plots of the observables or cosmological constraints.

below corresponds to the recipe v.1.0 implemented in `CLOE` version 1.0 and version 1.1. The primary *Euclid* observables are written in terms of basic cosmological functions described in chapter 1, such as the *comoving distance* $\chi(z)$, the *Hubble parameter* $H(z)$, the *matter power spectrum* $P_{\delta\delta}$, the *growth factor* $D(z)$ and growth rate $f(z)$, the *angular diameter distance* $D_{\rm A}$, and the *Modified Gravity phenomenological function* $\Sigma_{\rm MG}$[6] (see Figure 6.2). In this section, the *speed of light* c factors in the mathematical expressions are kept.

6.3.1 Galaxy Clustering: Spectroscopic (GCsp)

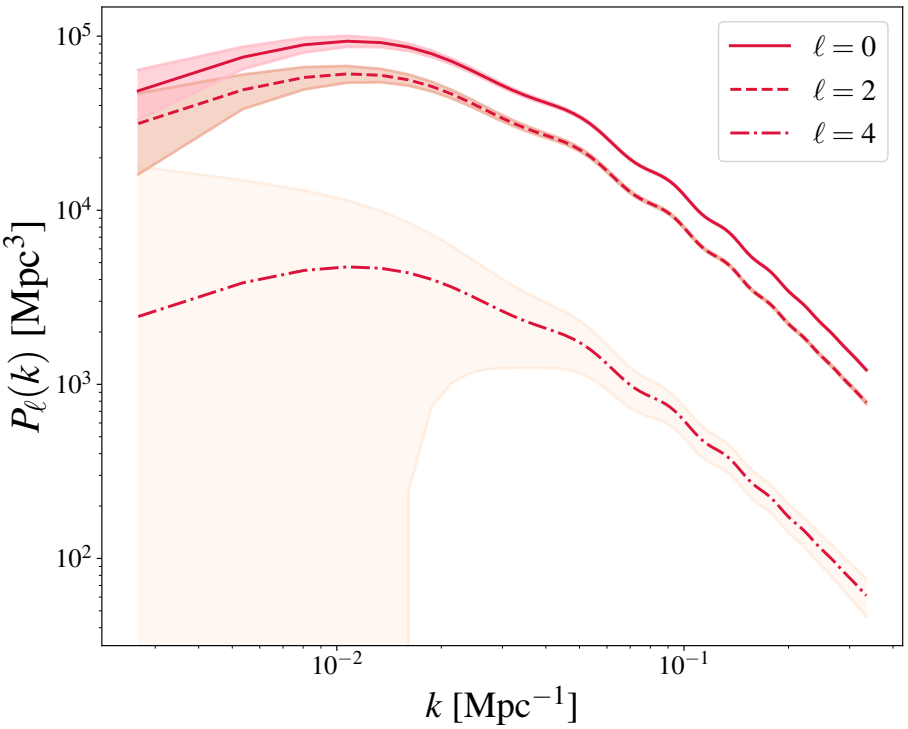

Figure 6.3: Legendre multipoles $P_\ell(k)$ for the spectroscopic galaxy clustering probe (GCps) as a function of scale k for multipoles $\ell = 0$ (solid line), $\ell = 2$ (dashed line) and $\ell = 4$ (dashed-dotted line), corresponding to the second tomographic redshift bin centred at $z = 1.2$. The shades regions correspond to the data uncertainties given the fiducial covariance matrix as explained in subsection 6.4.4. Within `CLOE`, the spectroscopic legendre multipoles are obtained in `spectro.py`. The plot has been generated taking the fiducial cosmological values presented in section 6.9.

As we have seen in subsection 1.5.3, galaxies are biased tracers of the total matter in the Universe. In the linear regime, we can assume a linear deterministic bias

[6]The Modify Gravity function $\Sigma_{\rm MG}$ is implemented in the code, but in this version, it is set equal to 1.

model and write the relation between the galaxy power spectrum $P_{\rm gg}^{\rm GCsp}$ and the matter power spectrum $P_{\delta\delta}(k,z)$ as

$$P_{\rm gg}^{\rm GCsp}(k,\mu_k,z) = \left[b^{\rm GCsp}(z) + f(z)\mu_k^2\right]^2 P_{\delta\delta}(k,z)\,, \tag{6.1}$$

where we know that our observations are affected by redshift space distortions (RSD), whose effect in the power spectrum can be described at the linear level by $f(z)\mu_k$. Also $k \equiv |\vec{k}|$ is the module of the wavevector $\vec{k}$, μ_k is the cosine of the angle between $\vec{k}$ and the line-of-sight direction, and $f(z)$ is the growth rate. The $b^{\rm GCsp}(z)$ is the redshift-dependent linear bias parameter and will take different values for each tomographic redshift bin. The analysis of tomographic galaxy clustering probe using spectroscopic redshifts is expressed in Legendre multipoles. In Fourier space, the Legendre multipoles $P_\ell(k)$ are given by

$$P_\ell(k,z) \equiv \frac{2\ell+1}{2}\int_{-1}^{1} {\rm d}\mu_k\, L_\ell(\mu_k) P_{\rm gg}^{\rm GCsp}(k,\mu_k,z)\,, \tag{6.2}$$

where L_ℓ is the Legendre polynomial. As we have explained already subsection 1.5.3, the true underlying power spectrum is not direct observable. All estimates of these functions based on galaxy surveys require the assumption of a fiducial cosmology to transform the observed redshifts into physical separations. A difference between the fiducial and true cosmologies leads to a re-scaling of the components parallel and perpendicular to the line-of-sight direction, $s_\parallel$ and $s_\perp$, of the total separation vector $\mathbf{s}$ between two galaxies in the survey, defined as

$$s_\perp = q_\perp s_\perp^{\rm fid}, \tag{6.3}$$

$$s_\parallel = q_\parallel s_\parallel^{\rm fid}, \tag{6.4}$$

where 'fid' denotes the quantities in the fiducial cosmology and the scaling factors q are given by the ratios of the comoving angular diameter distance, $D_{\rm A}(z)$, and the Hubble parameter, $H(z)$, respectively, in the true and fiducial cosmologies at the mean redshift of the sample:

$$q_\perp = \frac{D_{\rm A}(z)}{D_{\rm A}^{\rm fid}(z)}, \tag{6.5}$$

$$q_\parallel = \frac{H^{\rm fid}(z)}{H(z)}. \tag{6.6}$$

Equation (6.3) and equation (6.4) can be written in terms of k and μ_k as (Ballinger et al., 1996)

$$k(k^{\rm fid},\mu_k^{\rm fid},z) = k^{\rm fid}\left[q_\parallel^{-2}(z)\,(\mu_k^{\rm fid})^2 + q_\perp^{-2}(z)\left(1-(\mu_k^{\rm fid})^2\right)\right]^{1/2}, \tag{6.7}$$

$$\mu_k(\mu_k^{\rm fid},z) = \mu_k^{\rm fid}\, q_\parallel^{-1}(z)\left[q_\parallel^{-2}(z)\,(\mu_k^{\rm fid})^2 + q_\perp^{-2}(z)\left(1-(\mu_k^{\rm fid})^2\right)\right]^{-1/2}. \tag{6.8}$$

In Fourier space, the fiducial cosmology also leads to a rescaling of the power spectrum amplitude by a factor $\left(q_\perp^2 q_\parallel\right)^{-1}$. These geometric distortions must be applied to

all model predictions before they are compared with real galaxy clustering measurements. Using equations (6.7, 6.8), the prediction for the observed Legendre multipoles $P_{\mathrm{obs},\ell}(k^{\mathrm{fid}})$ including geometric Alcock-Pacyznski distortions can be obtained as

$$P_\ell(k^{\mathrm{fid}}, z) = \frac{1}{q_\perp^2(z) q_\parallel(z)} \frac{2\ell+1}{2} \int_{-1}^{1} L_\ell(\mu_k^{\mathrm{fid}}) P_{\mathrm{gg}}^{\mathrm{GCsp}} \left(k(k^{\mathrm{fid}}, \mu_k^{\mathrm{fid}}), \mu_k(\mu_k^{\mathrm{fid}}); z \right) \mathrm{d}\mu_k^{\mathrm{fid}} \,. \tag{6.9}$$

The plotted expression for equation (6.9) can be found in Figure 6.3.

6.3.2 Weak Lensing (WL)

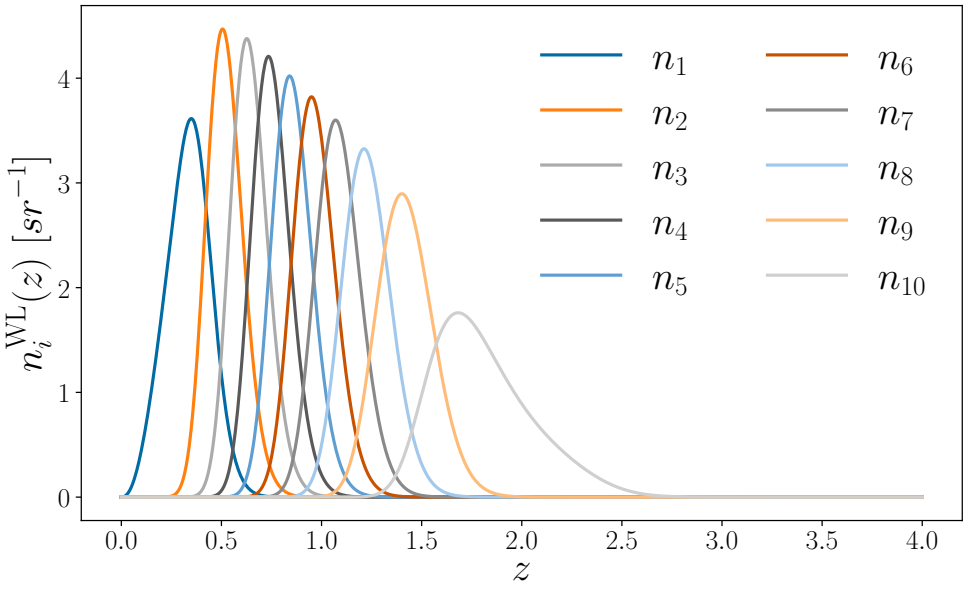

Figure 6.4: Galaxy redshift density distribution $n_i^{\mathrm{WL}}(z)$ as a function of redshift for the weak lensing (WL) probe in the 10 different tomographic redshift bins i. These distributions are read by `CLOE` as external files in `reader.py` and further processed in `redshift_distribution.py`. The benchmark distributions have been generated as explained in subsection 6.4.1. In `CLOE` v.1.1, the density distributions for both weak lensing and photometric galaxy clustering are considered equal: $n_i^{\mathrm{WL}}(z) = n_i^{\mathrm{GCph}}(z)$.

As it was already introduced in subsection 1.5.4, the light coming from distant galaxies bends towards us due to the massive objects that it can find along its trajectory, inducing distortions in their observed shape, allowing us to map the distribution of matter in space. This effect is called gravitational lensing, and in their weakest limit, we can extract cosmological information by studying the tiny distortions in galaxy shapes statistically. The weak lensing probe is written in terms of the tomographic angular power spectra in harmonic space, $C_{ij}^{\mathrm{WL}}(\ell)$, with i, j labelling the corresponding redshift bins, and WL standing for Weak Lensing. The angular power spectrum C_{ij}^{WL} is the Fourier transform of the two-point correlation function, which

can be written as a line-of-sight integral of the matter power spectrum $P_{\delta\delta}$ and a window function $W_i(z)$, which contains information about the probability distribution of observing a source galaxy at a given redshift z and about the physics of gravitational lensing. The weak lensing effect is modelled by taking into account not only the shear power spectrum (γ) but also the intrinsic alignment (IA) contribution, which is an astrophysical systematic error on the shapes of galaxies due to tidal forces applied by the surrounding large scale structure. In total, the weak lensing angular power spectrum $C_{ij}^{\rm WL}(\ell)$ is:

$$C_{ij}^{\rm WL}(\ell) = C_{ij}^{\gamma\gamma}(\ell) + C_{ij}^{\rm IA\gamma}(\ell) + C_{ij}^{\rm IAIA}(\ell). \tag{6.10}$$

We start by introducing the WL galaxy source redshift density distribution convolved with the photo-z uncertainties of the i-th redshift bin $n_i^{\rm WL}(z)$ (see Figure 6.4). The total galaxy density in the i-th redshift bin is given by the equation

$$\bar{n}_i^{\rm WL} = \int_{z_{\rm min}}^{z_{\rm max}} {\rm d}z\, n_i^{\rm WL}(z), \tag{6.11}$$

where $z_{\rm min}$ and $z_{\rm max}$ are the minimum and the maximum redshift value of the redshift bin distribution, respectively. In this recipe, we assumed that the analysis will be done with 10 different tomographic redshift bins. Note that we have explicitly added a superscript 'WL' to remind us that, in general, WL and GCph samples can be different, i.e. $n_i^{\rm WL}(z) \neq n_i^{\rm GCph}(z)$, although, in this recipe so far, both are considered equal. We can now define the window functions $W_i(z)$ in terms of the WL galaxy source redshift distribution $n_i^{\rm WL}(z)$ for the two main contributions included in the modelling of the weak lensing effect (see Figure 6.5): the shear window function W_i^{γ} and the intrinsic alignment $W_i^{\rm IA}$ one, defined as:

$$W_i^{\gamma}(k,z) = \frac{3}{2}\left(\frac{H_0}{c}\right)^2 \Omega_{\rm m}(1+z)\Sigma_{\rm MG}(k,z)\chi(z)\int_z^{z_{\rm max}} {\rm d}z' n_i^{\rm WL}(z')\left[\frac{\chi(z')-\chi(z)}{\chi(z')}\right], \tag{6.12}$$

$$W_i^{\rm IA}(z) = \frac{n_i^{\rm WL}(z)}{\bar{n_i}^{\rm WL}}\frac{H(z)}{c}, \tag{6.13}$$

where $\Sigma_{\rm MG}(k,z)$ a modified gravity function. Moreover, we can relate the window functions W_i with the matter density distribution in the universe encoded in the matter power spectrum $P_{\delta\delta}$. For the case of IA, we need to introduce an intrinsic alignment model. In version 1.0, we model the intrinsic alignment using the redshift-dependent non-linear alignment (zNLA) model explained in (Euclid Collaboration et al., 2020), which is written in terms of the intrinsic alignment function $f_{\rm IA}(z)$:

$$f_{\rm IA}(z) = -A_{\rm IA}\mathcal{C}_{\rm IA}\frac{\Omega_m}{D(z)}(1+z)^{\eta^{\rm IA}}[\langle L\rangle(z)/L_\star(z)]^{\beta^{\rm IA}}, \tag{6.14}$$

where $\langle L\rangle(z)$ is the luminosity redshift-dependent mean, $L_\star(z)$ the characteristic luminosity of source galaxies as computed from the luminosity function, the parameter

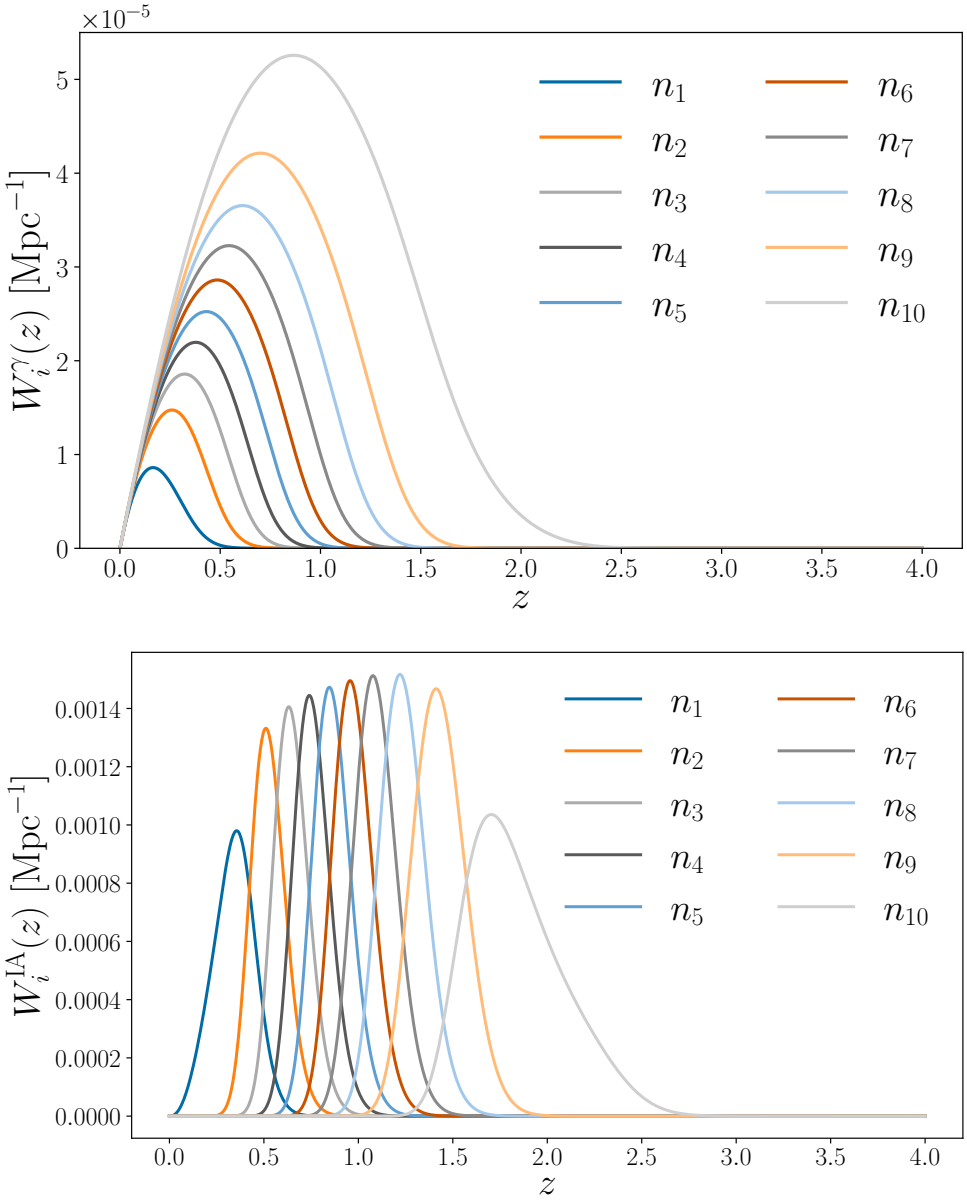

Figure 6.5: Weak lensing window functions: (**top**) the shear window function $W_i^\gamma(z)$ and (**bottom**) the intrinsic alignment $W_i^{\rm IA}(z)$ as a function of redshift z for the 10 photometric tomographic redshift bins i. The plots have been generated taking the fiducial cosmological values presented in section 6.9. Within `CLOE`, these window functions are calculated in `photo.py`.

$\beta^{\rm IA} = 0$, and $\mathcal{C}_{\rm IA} = 0.0134$ is kept fixed as it is degenerated with the rest of the parameters. The parameters $\eta^{\rm IA}$ and $A^{\rm IA}$ are free parameters of the model and should be determined by fitting the data and/or via carefully designed simulations; this is why they are defined as free nuisance parameters in this recipe so that we can marginalize over them. Given the intrinsic alignment model we can define the corresponding power spectra that relates the matter power spectrum $P_{\delta\delta}$ with the intrinsic alignment systematic distortions:

$$P_{\delta{\rm IA}}(k,z) = f_{\rm IA}(z) P_{\delta\delta}(k,z), \tag{6.15}$$

$$P_{\rm IAIA}(k,z) = [f_{\rm IA}(z)]^2 P_{\delta\delta}(k,z). \tag{6.16}$$

Finally, we can write the full expressions for each of the WL angular power spectra components present in equation (6.10) using the Limber approximation:

$$C_{ij}^{\gamma\gamma}(\ell) = c \int {\rm d}z \frac{W_i^{\gamma}\left[k_\ell(z), z\right] W_j^{\gamma}\left[k_\ell(z), z\right]}{H(z)\chi^2(z)} P_{\delta\delta}\left[k_\ell(z), z\right], \tag{6.17}$$

$$C_{ij}^{\gamma{\rm IA}}(\ell) = c \int dz \frac{W_i^{\gamma}\left[k_\ell(z), z\right] W_j^{\rm IA}(z) + W_i^{\rm IA}(z) W_j^{\gamma}\left[k_\ell(z), z\right]}{H(z)\chi^2(z)} P_{\delta{\rm IA}}\left[k_\ell(z), z\right], \tag{6.18}$$

$$C_{ij}^{\rm IAIA}(\ell) = c \int dz \frac{W_i^{\rm IA}(z) W_j^{\rm IA}(z)}{H(z)\chi^2(z)} P_{\rm IAIA}\left[k_\ell(z), z\right]. \tag{6.19}$$

The different components of $C_{ij}^{\rm WL}(\ell)$ are evaluated at $k = k_\ell(z) = \frac{\ell+1/2}{\chi(z)}$. An example plot for the total weak lensing angular power spectrum $C_{ij}^{\rm WL}(\ell)$ can be found in Figure 6.6.

6.3.3 Photometric Galaxy Clustering (GCph)

Similarly to the weak lensing case (section 6.3.2), to analyze the photometric galaxy clustering probe we still focus on the tomographic galaxy clustering angular power spectra $C_{ij}^{\rm GCph}(\ell)$ in the Fourier space using the Limber approximation and remaining up to linear scales. We start by defining the GCph galaxy source redshift density distribution $n_i^{\rm GCph}(z)$ in each tomographic bin i of the photometric survey by $n_i^{\rm CGph}(z)$, being the total angular galaxy density $\bar{n}_i^{\rm GCph}$ in such redshift bin given by

$$\bar{n}_i^{\rm GCph} = \int_{z_{\rm min}}^{z_{\rm max}} {\rm d}z\, n_i^{\rm GCph}(z). \tag{6.20}$$

We can define the window function $W_i^{\rm GCph}$ for a given tomographic bin i as

$$W_i^{\rm GCph}(z) = \frac{n_i^{\rm GCph}(z)}{\bar{n}_i^{\rm GCph}} \frac{H(z)}{c}, \tag{6.21}$$

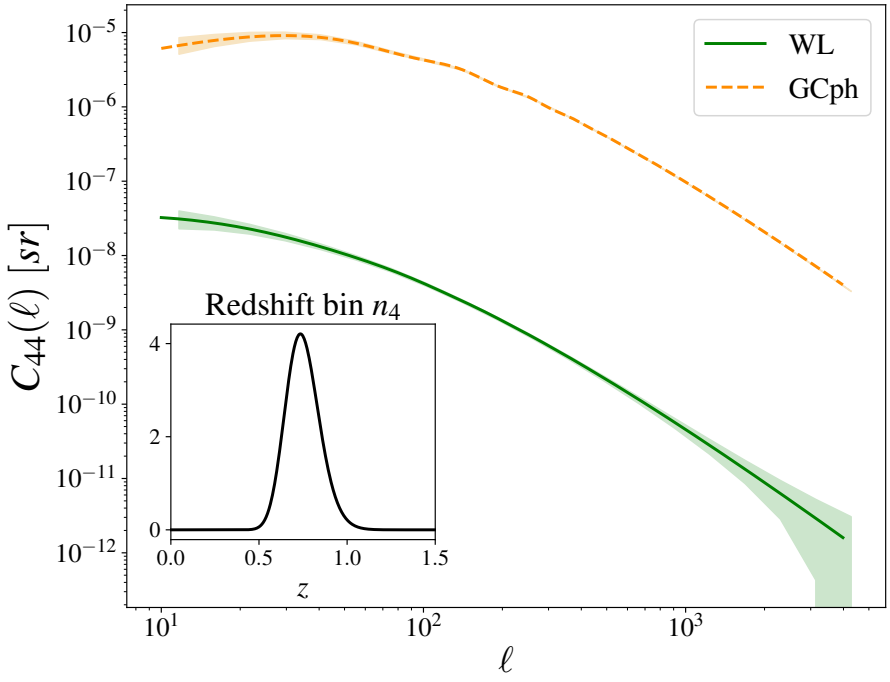

Figure 6.6: Photometric angular power spectra $C(\ell)$ for both weak lensing (WL, in green solid line) and photometric galaxy clustering (GCph, in yellow dashed lines) probes and for only one tomographic redshift bin $n_4(z)$ (that can be seen in the subplot in solid black line). The WL angular power spectrum includes both the contributions from shear and intrinsic alignment. The shades regions (WL, green and GCph, yellow) correspond to the data uncertainties given the fiducial covariance matrix as explained in subsection 6.4.1 and subsection 6.4.2. The plots have been generated taking the fiducial cosmological values presented in section 6.9. Within `CLOE`, these angular power spectra are calculated in `photo.py`.

which contains information about the probability density distribution of source galaxies at a given redshift. Similarly to the weak lensing case, we can relate the matter density power spectrum $P_{\delta\delta}$ with the galaxy-galaxy photometric power spectrum $P_{\rm gg}^{\rm GCph}(k,z)$, defined in this recipe as

$$P_{\rm gg}^{\rm GCph}(k,z) = [b^{\rm GCph}(z)]^2 P_{\delta\delta}(k,z), \tag{6.22}$$

where $b^{\rm GCph}$ is the linear photometric galaxy clustering bias and we expect to have different values for each redshift bin, so that we treat them as nuisance parameters to be sampled and marginalized over. Joining all the information together, the GCph angular power spectrum reads

$$C_{ij}^{\rm GCph}(\ell) = \int {\rm d}z \frac{W_i^{\rm GCph}(z) W_j^{\rm GCph}(z)}{H(z)\chi^2(z)} P_{\rm gg}^{\rm GCph}\left[k_\ell(z), z\right], \tag{6.23}$$

and it is evaluated at $k = k_\ell(z) = \frac{\ell+1/2}{\chi(z)}$. An example plot of equation (6.3.3) can be found at Figure 6.6.

6.3.4 Weak Lensing - Photometric Galaxy Clustering Cross-Correlation (XC)

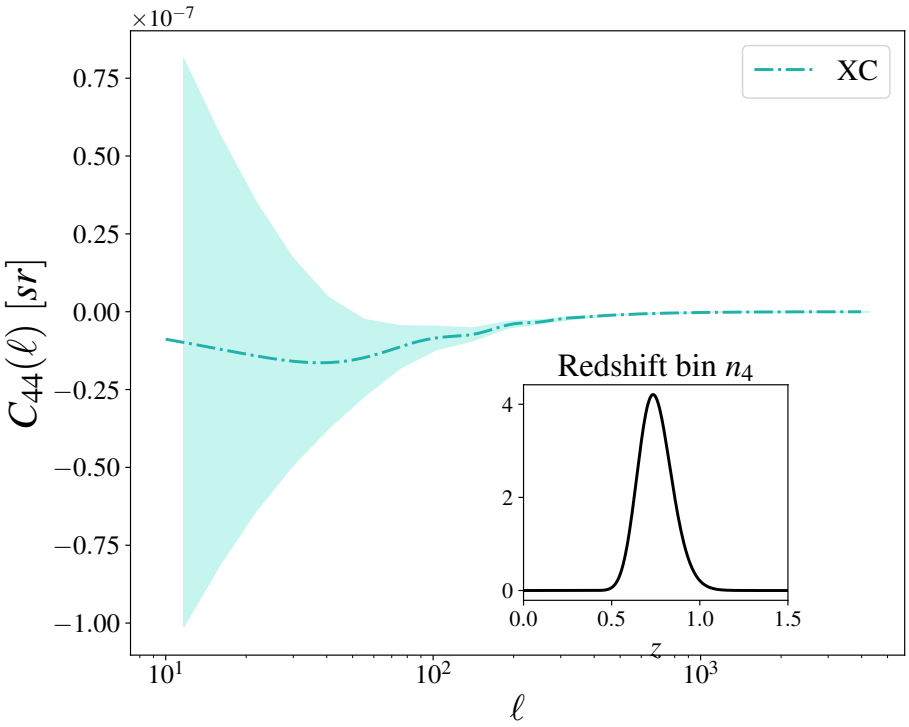

Figure 6.7: Angular power spectrum $C(\ell)$ for the photometric weak lensing - photometric Galaxy Clustering cross-correlation (XC, dashed-dotted blue line) and for only one tomographic redshift bin $n_4(z)$ (that can be seen in the subplot). The shade region (XC, blue) corresponds to the data uncertainties given the fiducial covariance matrix as explained in subsection 6.4.3. The plot has been generated taking the fiducial cosmological values presented in section 6.9. Within `CLOE`, the angular power spectrum is calculated in `photo.py`.

As for the other probes of the photometric catalogue, we focus on the angular cross-correlation between WL and GCph in Fourier space, $C_{ij}^{\rm XC}(\ell)$, with i, j labelling redshift bins, and WL and GCph respectively denoting lensing and (photometric) galaxy clustering. For each single probe that we cross-correlate, we refer to their dedicated sections (see subsection 6.3.2 and subsection 6.3.3). To calculate the cross-correlation, we need to define all the cross power spectra needed to build the corresponding angular power spectra. To take into account the shear-position correlation

we employ the following definition for the galaxy-matter power spectrum:

$$P_{\delta \mathrm{g}}^{\mathrm{GCph}}(k,z) = b^{\mathrm{GCph}}(z) P_{\delta\delta}(k,z), \tag{6.24}$$

where the photometric galaxy bias b^{photo} are those described in subsection 6.3.3. Similarly, to take into account the galaxy-IA correlation we employ the following definition of the density-intrinsic cross-spectrum:

$$P_{\mathrm{gIA}}^{\mathrm{GCph}}(k,z) = [f_{\mathrm{IA}}(z) b^{\mathrm{GCph}}(z)] P_{\delta\delta}(k,z), \tag{6.25}$$

where $f_{\mathrm{IA}}(z)$ is the intrinsic alignment function defined in equation (6.14). The weak lensing - position angular cross-spectrum $C_{\mathrm{ij}}^{\mathrm{XC}}(\ell)$ will be then given by the sum of the galaxy-shear correlation and the galaxy-IA correlation:

$$C_{ij}^{\mathrm{XC}}(\ell) = C_{ij}^{\delta \mathrm{g}}(\ell) + C_{ij}^{\mathrm{gIA}}(\ell), \tag{6.26}$$

which in practice is implemented as

$$\begin{aligned} &C_{ij}^{\mathrm{XC}}(\ell) \\ &\quad = c \int \frac{dz}{H(z) r(z)} \Big[W_i^{\gamma}\left[k_\ell(z), z\right] W_j^{\mathrm{GCph}}(z) P_{\delta \mathrm{g}}^{\mathrm{GCph}}(k_\ell, z) \\ &\qquad + W_i^{\mathrm{IA}}(z) W_j^{\mathrm{GCph}}(z) P_{\mathrm{gIA}}^{\mathrm{GCph}}(k_\ell, z) \Big], \end{aligned} \tag{6.27}$$

resulting in a single integration of the two different integrands, where $W_i^{\mathrm{GCph}}(z)$, $W_i^{\gamma}\left[k_\ell(z), z\right]$ and $W_i^{\mathrm{IA}}(z)$ are respectively defined in equations (6.21), (6.12) and (6.14). An example plot of the full observable in equation (6.27) can be found at Figure 6.7.

6.4 Benchmark data and covariance matrices

After introducing the modelling of the theoretical predictions for the *Euclid* primary observables, in this section, we introduce how the fiducial benchmark data, as well as the covariance matrices, have been generated to validate `CLOE`.

6.4.1 Weak Lensing (WL)

As a test case, we consider 10 equally populated redshift bins over the range $0.001 \leq z \leq 2.5$. We model the photometric redshift uncertainty as the sum of two Gaussian distributions: one for the well determined photometric redshifts and another for the outliers (see (Euclid Collaboration et al., 2020)), parameterized by $1 - f_{out}$ with σ_b uncertainty, and another f_{out} with σ_o uncertainty, respectively:

$$\begin{aligned} p_{\mathrm{ph}}(z_{\mathrm{p}}|z) = &\frac{1 - f_{\mathrm{out}}}{\sqrt{2\pi}\sigma_{\mathrm{b}}(1+z)} \exp\left\{ -\frac{1}{2} \left[\frac{z - c_{\mathrm{b}} z_{\mathrm{p}} - z_{\mathrm{b}}}{\sigma_{\mathrm{b}}(1+z)} \right]^2 \right\} \\ &+ \frac{f_{\mathrm{out}}}{\sqrt{2\pi}\sigma_{\mathrm{o}}(1+z)} \exp\left\{ -\frac{1}{2} \left[\frac{z - c_{\mathrm{o}} z_{\mathrm{p}} - z_{\mathrm{o}}}{\sigma_{\mathrm{o}}(1+z)} \right]^2 \right\}. \end{aligned} \tag{6.28}$$

The number density distribution $n_i(z)$ of the observed galaxies in the ith bin is given by:

$$n_i(z) = \frac{\int_{z_i^-}^{z_i^+} \mathrm{d}z_\mathrm{p}\, n(z) p_\mathrm{ph}(z_\mathrm{p}|z)}{\int_{z_\mathrm{min}}^{z_\mathrm{max}} \mathrm{d}z \int_{z_i^-}^{z_i^+} \mathrm{d}z_\mathrm{p}\, n(z) p_\mathrm{ph}(z_\mathrm{p}|z)}, \tag{6.29}$$

where (z_i^-, z_i^+) are the edges of the ith redshift bin and

$$n(z) \propto \left(\frac{z}{z_0}\right)^2 \exp\left[-\left(\frac{z}{z_0}\right)^{3/2}\right] \tag{6.30}$$

is the galaxy density distribution as defined in the "Red Book" (Laureijs et al., 2011), where $z_0 = z_\mathrm{m}/\sqrt{2}$ with z_m being the median redshift. The values for the different parameters are provided in (Euclid Collaboration et al., 2020). As mentioned in the description of the WL angular power spectra, concerning the Intrinsic Alignment effect, for this recipe we have used a zNLA model (or eNLA with $\beta_\mathrm{IA} = 0$) for the benchmark data array as well as for the covariance matrices. The following fiducial values are used $\{\mathcal{A}_\mathrm{IA}, \eta_\mathrm{IA}, \beta_\mathrm{IA}\} = \{1.72, -0.41, 0.0\}$, while $\mathcal{C}_\mathrm{IA}$ =0.0134 remains fixed in the analysis, as it is degenerated with $\mathcal{A}_\mathrm{IA}$.

Concerning the covariance matrix, we will consider only Gaussian terms and it is given by

$$\begin{aligned}
&\mathrm{Cov}[C_{ij}^\mathrm{WL}(\ell), C_{mn}^\mathrm{WL}(\ell')] \\
&\qquad = \frac{\delta^\mathrm{K}_{\ell\ell'}}{(2\ell+1) f_\mathrm{sky} \Delta\ell} \Big\{ \Big[C_{im}^{WL}(\ell) + N_{im}^\mathrm{WL}(\ell) \Big] \\
&\qquad \Big[C_{jn}^\mathrm{WL}(\ell') + N_{jn}^\mathrm{WL}(\ell') \Big] \\
&\qquad + \Big[C_{in}^\mathrm{WL}(\ell) + N_{im}^\mathrm{WL}(\ell) \Big] \\
&\qquad \Big[C_{jm}^\mathrm{WL}(\ell') + N_{jm}^\mathrm{WL}(\ell') \Big] \Big\} ,
\end{aligned} \tag{6.31}$$

where δ^K is the Kronecker delta symbol, the indexes i, j, m, n run over all tomographic bins, $\Delta\ell$ is the width of the multipoles bins and f_sky is the fraction of the sky covered by the survey (see (Euclid Collaboration et al., 2020) for details). The shot noise term due to the uncorrelated part of the intrinsic alignment ellipticity field is defined as

$$N_{ij}^\mathrm{WL}(\ell) = \frac{\sigma_\epsilon^2}{\bar{n}_i^\mathrm{WL}} \delta_{ij}^\mathrm{K} \tag{6.32}$$

with $\sigma_\epsilon = 0.3$ being the variance of the intrinsic ellipticity. It is worth mentioning that the code implementation reads the observed galaxy distribution (already convolved with the photometric redshift PDF) from an external file since it will be provided by the corresponding SGS OU as well as the covariance matrix as an external file as it will be computed (analytically or from simulations) by an external IST:NL/SWG/OU team.

Any of the validated codes used in the IST:Forecast work can give as outputs the $C_{ij}^\mathrm{WL}(\ell)$ using the galaxy density distribution in equation (6.29), as well as the

Gaussian covariance matrix. We have used the IST:Forecast codes to produce the benchmark data, which is read by `CLOE` as external data. The external benchmark data $C_{ij}^{\rm WL}(\ell)$ is computed by considering 20 logarithmically equispaced values in ℓ over the range $\ell = [10, 4000]$.

6.4.2 Photometric Galaxy Clustering (GCph)

Like in the case for external benchmark for $C_{ij}^{\rm WL}(\ell)$, we consider 10 equally populated redshift bins from $z = 0.001$ up to $z = 2.5$ with 20 multipole bins ℓ logarithmically equally spaced over the range $\ell = [10, 4000]$. The photometric bias $b^{\rm GCph}$ for the benchmark is obtained by linearly interpolating the bias values for the redshift bins $0.001 \leq z \leq 2.5$, and for redshifts above the final bin ($z > 2.5$), we use the bias obtained for the final redshift bin value. Similarly, for redshifts below the first bin ($z < 0.001$), we use the bias obtained for the first redshift bin value. In practice, this is computed as

$$b^{\rm GCph}(z) = \begin{cases} \mathcal{L}(z = 0.001) & \text{for} \quad z < 0.001 \\ \mathcal{L}(z) & \text{for} \quad 0.001 \leq z \leq 2.5 \\ \mathcal{L}(z = 2.5) & \text{for} \quad z > 2.5 \end{cases} \tag{6.33}$$

where $\mathcal{L}$(z) is obtained by linearly interpolating the points

$$[z_1, (1+z_1)^{1/2}, z_2, (1+z_2)^{1/2}, z_3, (1+z_3)^{1/2}, ..., z_{10}, (1+z_{10})^{1/2}], \tag{6.34}$$

with z_i the centre of the i-th redshift bin.

Concerning the data covariance matrix, we consider a theoretical Gaussian covariance matrix as in the Weak Lensing case in equation (6.31), but with a different shot-noise term given by

$$N_{ij}^{\rm GCph}(\ell) = \frac{\delta_{ij}^{\rm K}}{\bar{n}_i^{\rm GCph}}, \tag{6.35}$$

which is scale-independent. Similarly to the weak lensing case, `CLOE` reads both the observed galaxy distribution (already convolved with the photometric redshift PDF, which in this case is assumed to be equal to the one used for the weak lensing observable) and the covariance matrix as external files. The $C_{ij}^{\rm GCph}(\ell)$ as well as the covariance matrix are obtained by using IST:Forecast validated codes.

6.4.3 Weak Lensing - Photometric Galaxy Clustering Cross-Correlation (XC)

The benchmark data is computed as given in sections 6.4.2 and 6.4.1. Concerning the covariance matrix of the XC observable, we consider again a theoretical Gaussian

covariance matrix, which in this case reads

$$\begin{aligned}\mathrm{Cov}[C_{ij}^{\mathrm{XC}}(\ell), C_{kl}^{\mathrm{XC}}(\ell')] \\ &= \frac{\delta^{\mathrm{K}}_{\ell\ell'}}{(2\ell+1) f_{\mathrm{sky}} \Delta\ell} \Big\{ C_{il}^{\mathrm{XC}}(\ell)\, C_{jk}^{\mathrm{XC}}(\ell') + \\ &\Big[C_{ik}^{\mathrm{WL}}(\ell) + N_{ik}^{\mathrm{WL}}(\ell)\Big] \\ &\Big[C_{jl}^{\mathrm{GCph}}(\ell') + N_{jl}^{\mathrm{GCph}}(\ell')\Big]\Big\}\,,\end{aligned} \tag{6.36}$$

where the noise terms are defined in equations (6.32) and (6.35). The covariance matrix as well as the fiducial benchmark data can be read as external files and provided as an input to `CLOE`, and the $C_{ij}^{\mathrm{GCph}}(\ell)$ as well as the covariance matrix is obtained by using IST:Forecast validated codes.

6.4.4 Spectroscopic Galaxy Clustering (GCsp)

As an external benchmark, we consider predictions of the Legendre multipoles in Fourier space computed under the same recipe described in subsection 6.3.1, with covariance matrices computed within the Gaussian approximation as described in (Grieb, Sánchez, Salazar-Albornoz, & Vecchia, 2016). These predictions assume the same underlying cosmology[7] and survey specifications in terms of volume, galaxy number density and bias as of the IST:Forecast paper (Euclid Collaboration et al., 2020) (see tables 1, 2, and 3 of this paper). The analysis of the spectroscopic sample assumes 4 redshift bins from $z = 0.9$ up to $z = 1.8$, whose redshift range is summarised in Table 6.2. All theory predictions are computed at the mean redshift $z = (z_{max} - z_{min})/2$ of each bin. The input linear power spectra used to compute these external benchmarks were obtained with `CAMB`, using the parameter file stored in the `gitlab` repository of the IST:Forecast (which corresponds to the cosmology specified in Table 1 of (Euclid Collaboration et al., 2020), except for the value of σ_8 that here is 0.8156, corresponding to $A_s = 2.12605 \cdot 10^{-9}$).

The Gaussian prediction for the statistical uncertainty on the power spectrum multipoles $P_\ell(k)$ depends on the choice of the k-bin width, Δk. As we are implicitly assuming that measurements of $P_\ell(k)$ are uncorrelated under the Gaussian assumption, the minimal value for Δk should be the effective fundamental frequency defined as $k_{\mathrm{f}}^{\mathrm{eff}} \equiv 2\pi/V^{1/3}$, V being the volume sample (we assume an ideal cubic or otherwise compact volume). A value of Δk smaller than $k_{\mathrm{f}}^{\mathrm{eff}}$ would result in correlated bins even under the Gaussian assumption since $k_{\mathrm{f}}^{\mathrm{eff}}$ is the smallest difference in k we can resolve. The values of $k_{\mathrm{f}}^{\mathrm{eff}}$ for the four redshift bins are reported in Table 6.2. We assume a bin width of $\Delta k = 0.004$, a value larger but close to the largest value for $k_{\mathrm{f}}^{\mathrm{eff}}$ across the four redshift bins. The power spectra are computed in the range $0.002 < k < 0.502$ and are evaluated at the centre of the wavenumber bin. To compute the predictions for the gaussian variance we average over the bins.

[7]Except for the value of σ_8 that is 0.8156, corresponding to $A_s = 2.12605 \cdot 10^{-9}$.

redshifts	$k_{\rm f}^{\rm eff}$
$0.9 < z < 1.1$	0.0031
$1.1 < z < 1.3$	0.0030
$1.3 < z < 1.5$	0.0029
$1.5 < z < 1.8$	0.0025

Table 6.2: Values of $k_{\rm f}^{\rm eff}$ for the four spectroscopic redshift bins.

With respect to the covariance matrix, if we define the multipole expansion of the per-mode covariance in a volume $V_{\rm sur}(z)$ as (Grieb et al., 2016)

$$\sigma^2_{\ell_1\ell_2}(k^{\rm fid};\,z) \equiv \frac{(2\ell_1+1)\,(2\ell_2+1)}{V_{\rm sur}(z)}\times \int_{-1}^{1}\left[\frac{1}{q_\perp^2 q_\parallel}P_{\rm gg}\left(k(k^{\rm fid},\mu_k^{\rm fid}),\mu_k(\mu_k^{\rm fid})\right)+\frac{1}{\bar{n}}\right]^2_{\ell_1}(\mu_k^{\rm fid})\,_{\ell_2}(\mu_k^{\rm fid})\mathrm{d}\mu_k^{\rm fid}\ , \quad (6.37)$$

the Gaussian covariance of the power spectrum multipoles $P_{{\rm obs},\ell}(k^{\rm fid};\,z)$ is given by

$$\mathrm{Cov}\left[P_{{\rm obs},\ell_1}(k_i^{\rm fid};\;z),P_{{\rm obs},\ell_2}\left(k_j^{\rm fid};\,z\right)\right]=\frac{2\,(2\pi)^4}{V^2_{k_i^{\rm fid}}}\delta_{ij}\int_{k_i^{\rm fid}-\Delta k^{\rm fid}/2}^{k_i^{\rm fid}+\Delta k^{\rm fid}/2}\sigma^2_{\ell_1\ell_2}(k^{\rm fid})\,(k^{\rm fid})^2\mathrm{d}k^{\rm fid}\ , \quad (6.38)$$

where the volume of the bin in k-space is $V_{k_i^{\rm fid}} = 4\pi[(k_i^{\rm fid}+\Delta k^{\rm fid}/2)^3-(k_i^{\rm fid}-\Delta k^{\rm fid}/2)^3]/3$. We will neglect any cosmology dependence of the covariance matrix C, which is kept fixed during the analysis. For our linear theory tests, we will compute input covariance matrices using the Gaussian approximation of (Grieb et al., 2016) specifically, the covariance matrices of the Legendre multipoles are given by equations (16) and (18) of (Grieb et al., 2016). All the spectroscopic benchmark data as well as the covariance matrices are pre-computed and given as external input to `CLOE`.

6.5 Standard likelihood analysis

The estimation of constraints on a given set of cosmological parameters, θ, based on the *Euclid* data, $\vec{d}$, will follow a Bayesian framework. According to Bayes' theorem (see chapter 1, section 1.7.1), the key ingredient in the estimation of the posterior distribution of the parameters, $P(\theta|\vec{d},M)$, is the likelihood function $\mathcal{L}(\vec{d}|\theta,M)$, which describes the plausibility of a certain parameter value θ after observing a particular outcome. For all *Euclid* primary probes, in this version of the recipe, we assume that the likelihood function $\mathcal{L}(\vec{d}|\theta,M)$ of these measurements is Gaussian with a cosmology independent covariance matrix C:

$$-2\log\mathcal{L}(\vec{d}|\theta,M)\propto\left(\vec{d}-\vec{T}(\theta)\right)^t C^{-1}\left(\vec{d}-\vec{T}(\theta)\right), \quad (6.39)$$

where C^{-1} is the inverse of the covariance matrix, also known as the precision matrix, $\vec{d}$ is the data vector constructed with the fiducial benchmark data explained in section 6.4 and $\vec{T}(\theta)$ is the theory vector constructed with the predictions for the *Euclid* observables whose recipes are depicted in section 6.3.

6.6 CLOE: specifications and structure

The code *Cosmological Likelihood for Observables in Euclid* (CLOE) is designed to produce theoretical predictions of the primary *Euclid* observables (see section 6.3) as well as the computation of the likelihood (see section 6.5) given some fiducial benchmark data (detailed in section 6.4). The specifications and structure presented in this section are referred to CLOE v.1.1, which contains the theoretical recipe version 1.0. explained in section 6.3. The code CLOE is fully written in python (Van Rossum & Drake Jr, 1995) using a set of common packages described in Table 6.3. It is designed to work as an external likelihood class for the Bayesian Analysis Framework code Cobaya[8]. To build and run CLOE, IST:L provides a dedicated conda environment (*Anaconda Software Distribution*, 2020) with development tools.

Package	Version
astropy	5.0.1
matplotlib	3.5.1
scipy	1.8.0
numpy	1.22.2
pandas	1.4.1
PyYAML	6.0
Cobaya	3.1.1
getdist	1.3.3

Table 6.3: Summary of the python packages that CLOE uses in all its different modules. The versions correspond to those requested by the conda environment provided by IST:L to run CLOE v.1.1.

The structure of CLOE is very modular, as it is designed keeping in mind the flexibility and future expansion of the code. The philosophy behind CLOE's design is to make it as user-friendly as possible so that many scientists within the EC can use it when the first release of data arrives[9]. As all *Euclid* data will be publicly released after a relatively short proprietary period and will constitute for many years the ultimate survey for Cosmology, we also aim CLOE to be very easy to learn and use so that everybody within the Cosmology community can exploit its capacity and features soon.

[8]See the documentation of how to create a customized external likelihood for Cobaya at (*Creating your own cosmological likelihood class*, 2022).

[9]Note that CLOE will eventually be used within different Working Packages of the SWGs. However, IST:L will be the ultimately responsible team for producing the official constraints on the cosmological parameters.

CLOE is the result of a study and analysis of the current state-of-the-art open-science Cosmology codes available in the literature. The first goal of IST:L was to investigate the features, portability, pros and cons of the most widely used cosmological software for Bayesian Statistical analyses: CosmoMC (Lewis & Bridle, 2002; Lewis, 2013), MontePython (Audren, Lesgourgues, Benabed, & Prunet, 2013), CosmoSIS (Zuntz et al., 2015) and Cobaya (Torrado & Lewis, 2021, 2019). In the end, given the collection of requirements requested by different Science Working Groups (SWGs) as well as the goals of the *Euclid* mission, IST:L decided to interface the likelihood code with Cobaya (and in future versions, with CosmoSIS), because:

- both codes allow us to obtain computational theoretical background functions from the most widely used Boltzmann Solvers CAMB (Lewis et al., 2000; Howlett et al., 2012) and CLASS (Essinger-Hileman et al., 2014). Moreover, Cobaya is prepared to interact with modified versions of CAMB and CLASS without further hacking Cobaya source code.

- both codes contain a large number of different samplers (the most important of them explained in chapter 1: nested sampler Polychord (W. J. Handley et al., 2015a, 2015b), Metropolis Hastings, evaluate...)

- both codes accept external self-defined likelihoods with minimal modification of their source codes.

The structure of CLOE v.1.1 can be seen in figure 6.8. The software CLOE consists of a main python class inherited from Cobaya that works as an external likelihood, called cobaya_interface.py (the class being called EuclidLikelihood). This class initializes the class Euclike and collects the main cosmological background functions requested to the Boltzmann Solver (i.e: CAMB or CLASS) through Cobaya. These ingredients are saved into an instance of the class Cosmology (within the cosmo module), where they are subsequently passed on to the nonlinear module. Given the requests selected by the user, the class Euclike invokes the initialization and reading routines of the corresponding selected data and covariance matrices to create the **data vector** (using the classes Reader and Masking in figure 6.8). Given the read data, CLOE computes the theoretical predictions of the observables and produces the **theory vector** calculating the predictions for the *Euclid* primary observables coded in the modules photometric_survey and spectroscopic_survey. Finally, Euclike computes the "log-like" value (see equation 6.39), which is returned to the Cobaya class in cobaya_interface.py. The selection of the observables and further features of the theoretical predictions can be done by the user either following the standard Cobaya input files (i.e: yaml files, python scripts or even using a python interpreter as jupyter) or using the so-called CLOE *overlayer* (a series of scripts provided in CLOE to camouflage the use of Cobaya so that the user is agnostic to the whole CLOE structure). CLOE understands "on-the-fly" the nuisance parameters that are introduced in the theoretical predictions, so that cobaya_interface.py source code does not need to be further modified. The advanced structure of CLOE, seen in detail with

its several python classes, inheritance and dependencies, can be found at the IST:L gitlab repository[10].

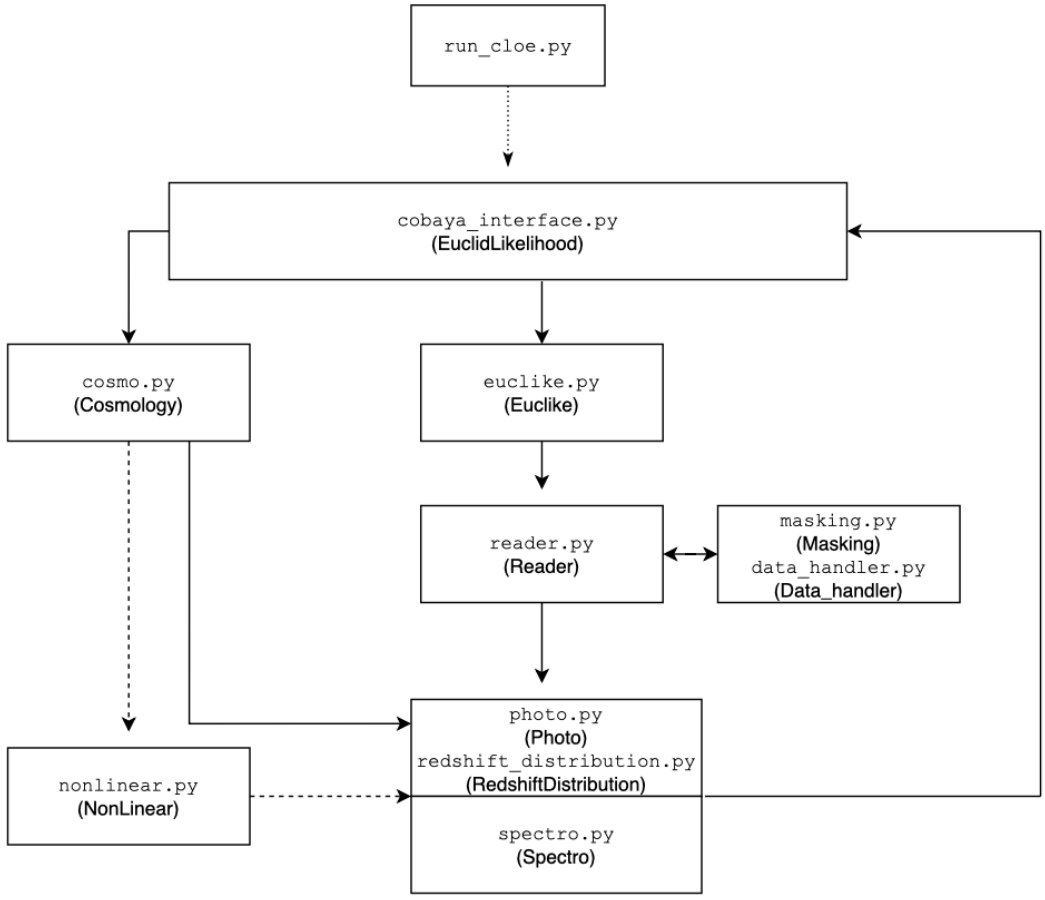

Figure 6.8: Simple diagram of the structure of CLOE in version 1.1, generated using draw.io. In the diagram it can be seen how the different modules within CLOE interact. The module run_cloe.py is informally known as the *overlayer* and the module nonlinear.py is currently implemented within the structure of CLOE but not yet used in this version (IST:NL is the responsible of the structure of this module and in this particular version, the theoretical observables are computed up to linear regime).

The software CLOE contains a README, an automatic generated documentation of the python software[11] as well as a series of scripts and notebooks. In particular, CLOE offers a DEMO jupyter notebook that shows how to compute the *Euclid* primary observables according to the corresponding recipe, internal background functions retrieved by Cobaya, the values of the likelihood, posterior and priors, the redshift

[10]Visit https://gitlab.euclid-sgs.uk/pf-ist-likelihood/likelihood-implementation/-/wikis/uploads/083b51a7591215dffecdbe511446d8b5/classes.png

[11]The automated generated documentation is created using "Read the docs" https://docs.readthedocs.io/en/stable/index.html.

distributions of the galaxies for both the photometric and spectroscopic catalogues and even the photometric window functions and various matter power spectra explained in subsection 6.3.1, subsection 6.3.2, subsection 6.3.3, subsection 6.3.4 and in table 6.3. All the figures of section 6.3 have been produced using this DEMO jupyter notebook.

6.7 Validation of the benchmark data using CLOE

To validate the theoretical recipe implemented in CLOE of the *Euclid* primary probes, we carry out a Monte Carlo Markov Chain (MCMC) sampling technique to retrieve the posterior distributions of the cosmological and nuisance parameters given the fiducial benchmark data explained in section 6.4. We take a flat ΛCDM model as the baseline cosmological model and we use the theoretical prediction of the *Euclid* primary probes up to linear order (non-linear corrections are not included). The sampler used is the Metropolis-Hastings mcmc included in Cobaya. The Boltzmann Solver is CAMB. In all cases, we use the default Cobaya convergence criterion (i.e: R-1 (means) < 0.01 and R-1 (standard deviations) < 0.2, see section 1.7.2 for more details about the definition of the convergence criterion R). We show in Table 6.4 the time required for each run to reach convergence according to these conditions while letting the nuisance parameters free. These runs were done in the xMaris cluster at the Lorentz Institute. xMaris runs on CentOS v7.6 and has as architecture x86_64. To run CLOE in this cluster, the conda environment it is not used and python v.3.7.4, gcc v.8.30, OpenMPI v.3.1.4 and mpi4py v.3.1.1 are used instead. It has been verified that very similar running times were obtained when a different cluster was used (Alice cluster at Leiden University).

The list of free sampled parameters included 5 cosmological parameters ($\Omega_{\rm b}$, $\Omega_{\rm m}$, H_0, n_s, A_s, where σ_8 and S_8 are obtained as derived parameters) and up to 16 nuisance parameters (the 10 different photometric bias parameters $b^{\rm GCph}$ for the 10 photometric redshift bins, the 4 different spectroscopic bias parameters $b^{\rm GCsp}$ for the 4 spectroscopic redshift bins, and the 2 intrinsic alignment nuisance parameters $A^{\rm IA}$ and $\eta^{\rm IA}$). The running scripts can be found in the *Euclid* gitlab repository of CLOE.

All the runs consisted of 8 chains, each of them crossing into 4 cores. The chains achieved a convergence of $R - 1 < 0.1$ in approximately 3 days, where some of them were visually converged, in particular, those runs that included the combinations of different observational probes. However, to reach the standard convergence criterion, the runs needed several days to end. The average computation time for CAMB was approximately 4.2 seconds per evaluation and that of CLOE around 7.1 seconds per evaluation. Note that this average time for CLOE is the sum of both the computation time of the theoretical observables as well as the computation time of the likelihood, and therefore, the average time for the theoretical predictions computed by CLOE is around 3 seconds per evaluation. Each chain contains around $\mathcal{O}(5)$ accepted steps.

After reaching convergence, the analysis and post-processing of the chains is performed with GetDist (Lewis, 2019). The list of best fit values, together with their 68% confidence intervals, are obtained after marginalizing the posterior distributions.

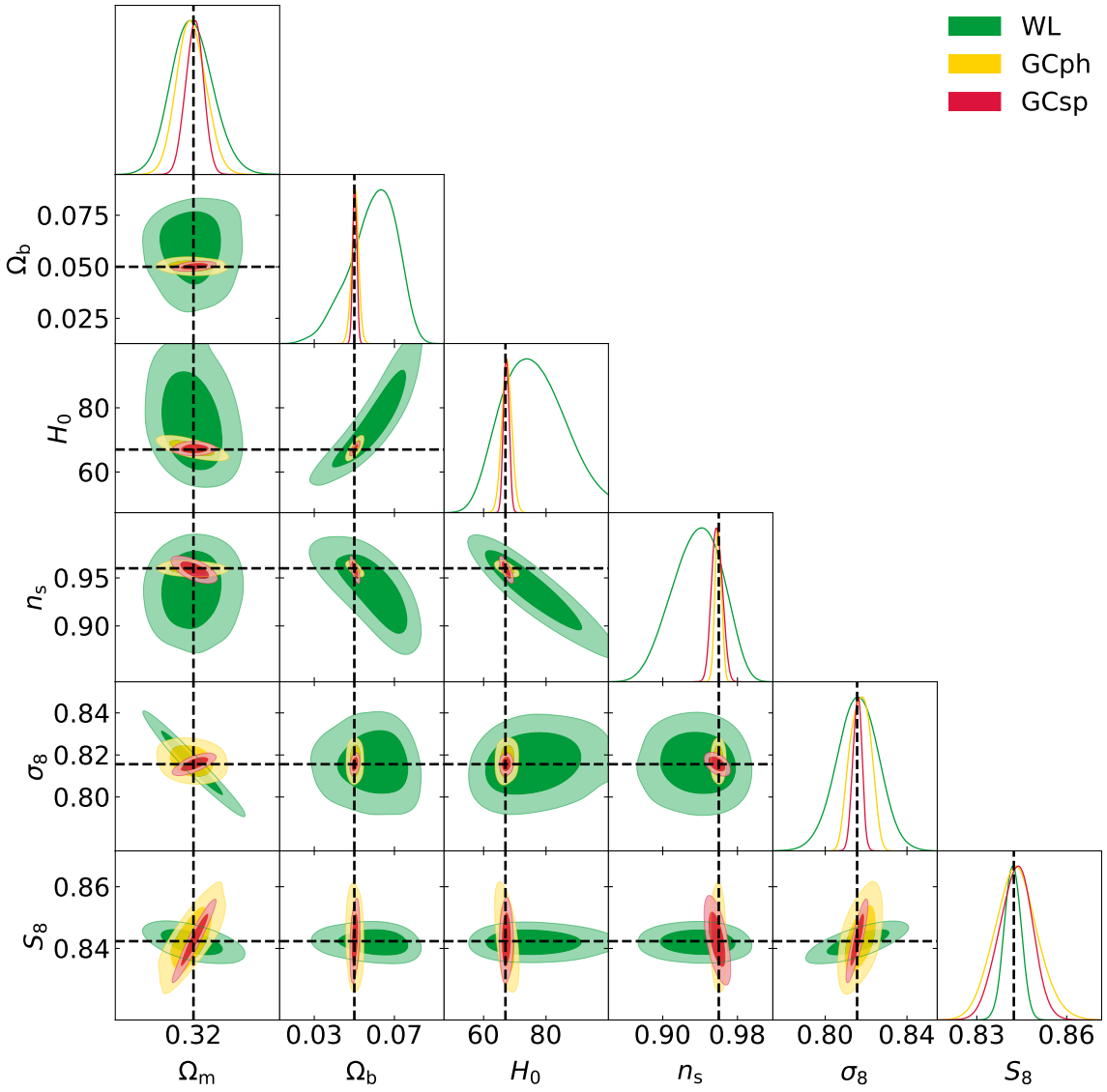

Figure 6.9: Constraints on the cosmological parameters using *Euclid* primary observational probes. This plot shows the different observational primary probes independently, i.e. weak lensing (WL) in green, photometric galaxy clustering (GCph) in yellow and spectroscopic galaxy clustering (GCsp) in red. The red dashed lines show the fiducial values of the parameters.

These best fits and uncertainties are gathered in the appendix section 6.9, in Table 6.5, Table 6.6, Table 6.7, Table 6.8 and Table 6.9. The benchmark data have been generated using a list of fiducial values for the cosmological and nuisance parameters that are also present in those tables. We also use the fiducial values for the cosmological parameters in plots of the posterior distributions, Figure 6.9 and Figure 6.10 for visualisation purposes.

As expected, the individual *Euclid* primary probes alone (WL, GCph and GCsp) do not show a big constraining power on the cosmological parameters. In fact, photo-

Data combination	Node	Elapsed time
GCsp	maris069	14 days, 18 hours, 52 minutes
GCph	maris070	42 days, 5 hours, 32 minutes
WL	maris072	16 days, 3 hours, 24 minutes
XC	maris073	17 days, 19 hours, 34 minutes
GCph+WL	maris070	14 days, 23 hours, 1 minutes
3x2pt	maris073	10 days, 4 hours, 3 minutes
GCph+WL+GCsp	maris073	15 days, 17 hours, 43 minutes
3x2pt+GCsp	maris072	7 days, 20 hours, 14 minutes

Table 6.4: Computation time to reach convergence for each of the considered *Euclid* probe combinations in the case where the nuisance parameters are free to vary. All nodes have as hardware an opteron6376 with 1Gb,R815,highmem. The combination 3x2pt refers to WL+GCph+XC together.

metric galaxy clustering (GCph) shows some degeneracies between the nuisance bias parameters b^{GCph} and the amplitude of the power spectrum A_s, and therefore, also between σ_8 (see Figure 6.9). This is the main reason why the computation time of this probe to reach convergence is the largest with respect to the other observational probes. To achieve convergence, the proposal step of the photometric nuisance bias parameters' priors have been adjusted to the optimal value to avoid the chains to get stuck. Weak lensing (WL) also shows degeneracies in the Hubble parameter H_0 and in the density of baryons Ω_{b}. Moreover, its constraining power on the spectral index n_s is more limited than that of GCph. Contrarily, WL constrains σ_8, due to the fact that WL is very sensitive to the matter density Ω_{m}. The spectroscopic galaxy clustering is overall the primary observational probe that has a good constraining power on all the cosmological parameters.

When combined, we can unleash the full power of *Euclid* primary probes (see Figure 6.10). The combination of both photometric primary probes (WL and GCph) breaks the degeneracies previously shown by the individual photometric probes (in particular, in σ_8, H_0 and Ω_{b}). Including the cross-correlation between weak lensing and photometric galaxy clustering (XC) in the analysis; in order words, the full 3x2pt photometric combination, increases the constraining power in σ_8 and Ω_{m}, although it does not improve the constraining power on the rest of parameters. However, when GCsp is included in the analysis, the confidence intervals of all cosmological parameters reduce significantly, showing how the combination of all *Euclid* primary probes will be the future state-of-the-art of new LSS cosmological analysis, being at the level of constraints obtained by Planck (if not successfully beating it).

For the case of the 3x2pt with GCsp, the analysis included 16 free nuisance parameters. The 12 photometric nuisance parameters, bias and and intrinsic alignment ones, are shown in Figure 6.11. It can be observed how adding the cross-correlation XC between weak lensing and photometric galaxy clustering improves the constraining power on the nuisance parameters, in particular, the intrinsic alignment parameters. In the case of the analysis of the spectroscopic bias parameters (see Figure 6.12),

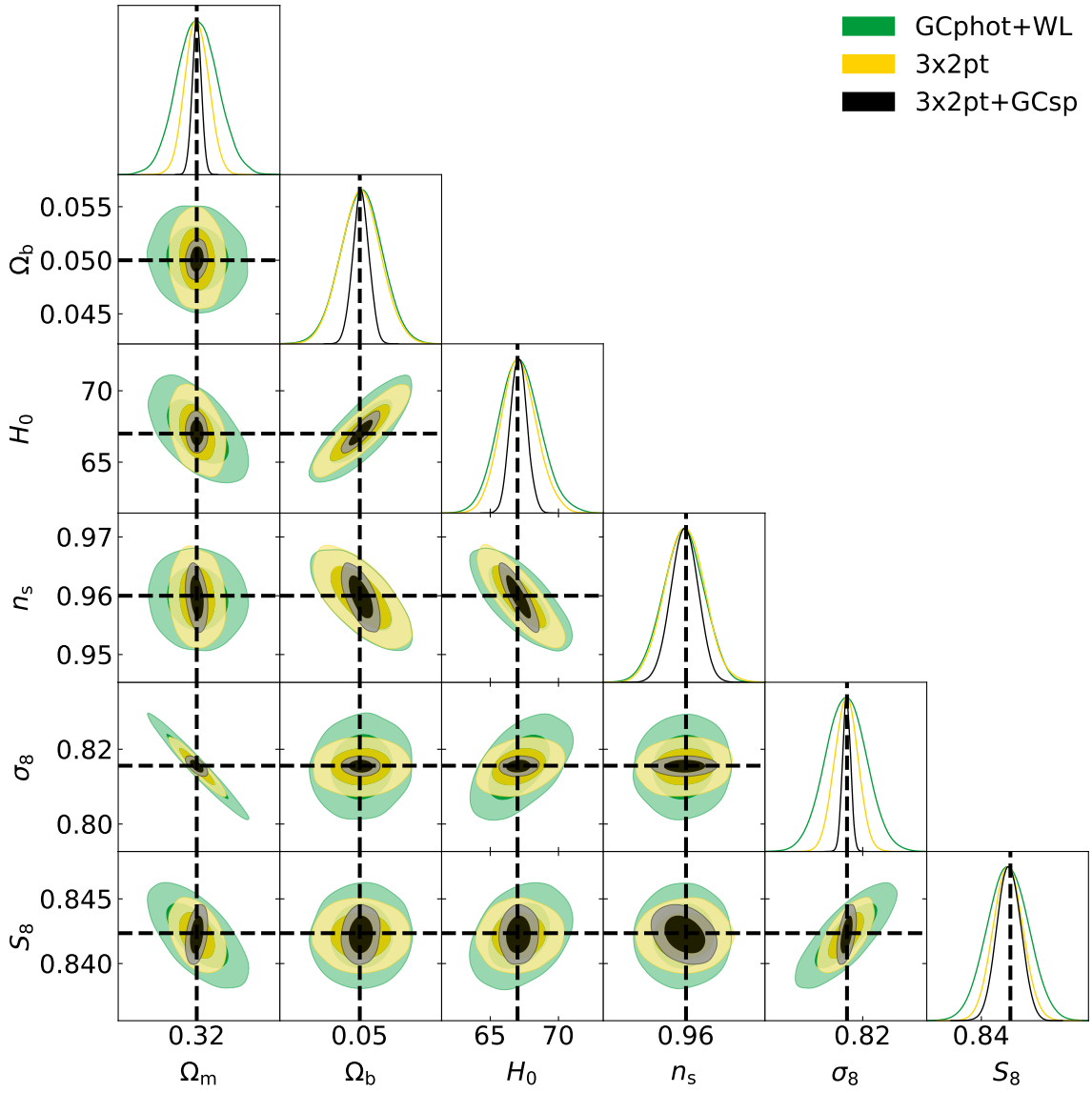

Figure 6.10: Constraints on the cosmological parameters using all possible (available) combinations of *Euclid* primary probes. The green contours show the combination of the two photometric probes, GCph and WL, yellow contours add to this combination the cross-correlation of these two probes (XC), producing the standard 3x2pt combination, and in the black contours, GCsp is added to this last combination as an independent probe (3x2pt+GCsp). The black dashed lines show the fiducial values of the parameters.

the spectroscopic bias parameters get further constrained when the 3x2pt probe is included in the analysis. We conclude that the improvement on all the nuisance parameters is a natural consequence of the improvement in constraining power of the cosmological parameters as more *Euclid* probes are combined.

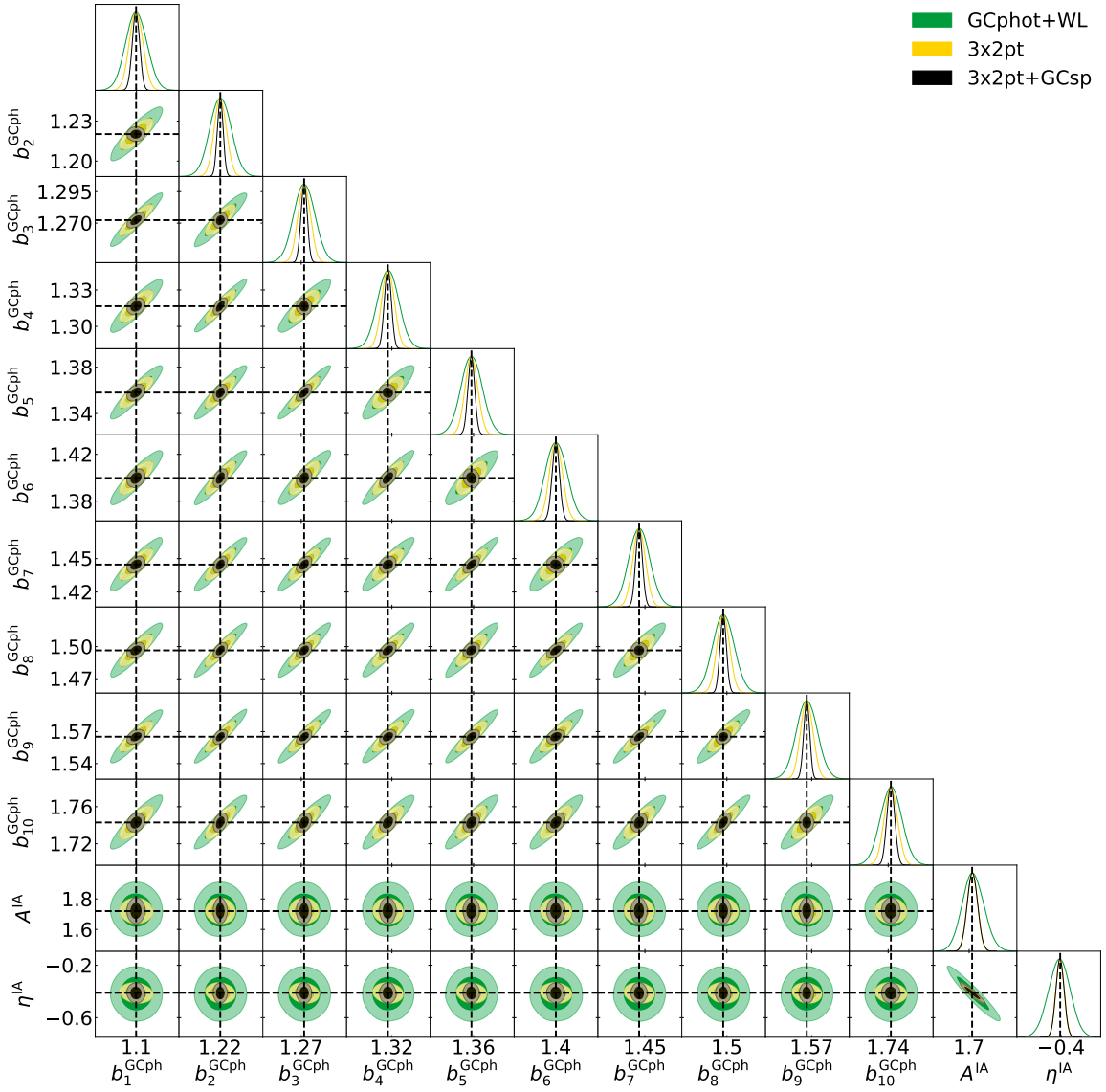

Figure 6.11: Constraints on the nuisance parameters of the photometric observational probes of *Euclid*: WL, GCph and XC. All photometric nuisance parameters of *Euclid* are shown, including the 10 different photometric galaxy clustering bias parameters b^{GCph} and the two weak lensing intrinsic alignment parameters A^{IA} and η^{IA}. The green contours show the combination of the two photometric probes, GCph and WL, yellow contours add to this combination the cross-correlation of these two probes, producing the standard 3x2pt combination, and in the black contours, GCsp is added to this last combination as an independent probe (3x2pt+GCsp). The black dashed lines show the fiducial values of the nuisance parameters.

6.8 The future of CLOE

In the previous sections, we have seen how we can use the *Cosmological Likelihood for Observables in Euclid* (CLOE) to plot *Euclid* primary observables and other internal

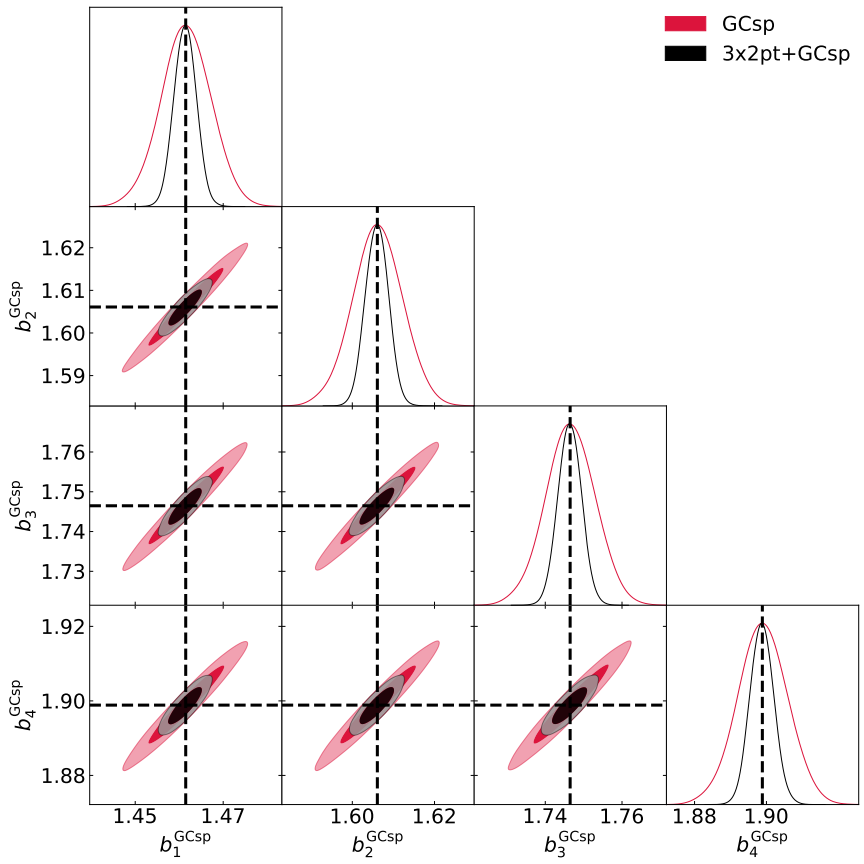

Figure 6.12: Nuisance parameters of the spectroscopic observational probes of *Euclid*. The 4 bias nuisance parameters of the spectroscopic *Euclid* probe are shown. The red contours show the bias parameters when only the Spectroscopic Galaxy Clustering is used, whereas, in the black contours, the photometric 3x2pt probe is added to the analysis. The black dashed lines show the fiducial values of the nuisance parameters.

ingredients needed for the computation of the theoretical recipe. Moreover, we have demonstrated that we can use `CLOE` as an external likelihood for `Cobaya` to perform parameter inference. Nevertheless, to do meaningful science, `CLOE` needs to evolve. In fact, versions 1.0 and 1.1 of `CLOE` can be considered a proof-of-concept where we set up the general structure of code, the interaction among the modules and the interface with `Cobaya` as well as with IST:NL. However, for real scientific performance, several changes are needed.

First, we need to update the recipe for the *Euclid* primary probes to include a more realistic modelling of the observables. For the spectroscopic survey, in addition to the observed Legendre multipoles $P_\ell(k)$, we aim to include the observed two-point correlation $\xi(s)$ function multipoles, where s is the separation between two galaxies in the survey. Furthermore, for the photometric survey, in particular photometric galaxy clustering, we will incorporate new angular power spectra terms in $C_{ij}^{\rm GCph}$ to account for contributions coming from redshift space distortions (RSD) and magnification. For the weak lensing observable, the WL galaxy redshift distribution $n_i^{\rm WL}(z)$ will be

replaced by an effective one to take into account the shear multiplicative bias and the blending terms. Furthermore, due to the sensitivity of the weak lensing probe on the non-linear model used to correct the matter power spectrum, the Bernaredeau - Nishimichi - Taruya (BNT) transform (see (Bernardeau, Nishimichi, & Taruya, 2014)) will be also included in the lensing window functions to minimise the overlap between them so that each tomographic bin only receives contribution from a well defined range in scale. For that, we will follow the recipes from (P. L. Taylor, Bernardeau, & Kitching, 2018; P. L. Taylor et al., 2021). For both cases, WL and GCph, we will allow testing non-flat cosmological models by including the corresponding changes into the equations and requesting new background functions to the Boltzmann solvers through `Cobaya`. Besides, the equations for the photometric observables will be updated to include the Modify Gravity function $\mu_{\rm MG}$ too.

Second, we will need to work towards a complete `CLOE` version that contains the necessary implementations to correct *Euclid* primary probes at non-linear scales. In this sense, a merge of IST:NL software development into `CLOE` is essential for the release of v.2.0 as well as for the validation of the current `nonlinear` module interface. Finally, we need to validate the new theoretical recipe against fiducial benchmark data and we need to check the overall performance of `CLOE`, giving particular attention to the speed of the code, and introduce any software optimisation if required.

`CLOE` is planned to be used for the third *Euclid* Science Performance Verification (SPV3) exercise. This activity aims to assess whether the expected performances of the *Euclid* project are in line with the core science objectives of the nominal mission, and furthermore, it will ensure gaining insights into the performances of the project and identify any critical aspect for the realization of the core goals in terms of systematic errors (either of instrumental, astrophysical of theoretical origin). For this exercise, IST:L will use `CLOE` to produce the official constraints on the cosmological parameters given the w_0w_aCDM model. For this reason, the release of `CLOE` v.2.0, including the above-mentioned modifications on the theoretical recipe of the observables, non-linear corrections and required speed optimisations, is planned before the end of 2022.

6.9 Appendix: fiducial and best fit values for the cosmological and nuisance parameters

	$\Omega_{\rm b}$	$\Omega_{\rm m}$	H_0	$n_{\rm s}$	σ_8
Fiducial	0.05	0.32	67	0.96	0.8156
GCsp	0.0502 ± 0.0020	0.3197 ± 0.0048	$67.3^{+1.5}_{-1.8}$	0.9594 ± 0.0036	$0.700^{+0.031}_{-0.057}$
GCph	0.0504 ± 0.0018	0.3194 ± 0.0047	67.4 ± 1.6	0.9595 ± 0.0035	0.8172 ± 0.0047
WL	$0.0593^{+0.014}_{-0.0089}$	$0.3196^{+0.0062}_{-0.0070}$	76^{+9}_{-10}	$0.937^{+0.029}_{-0.024}$	0.816 ± 0.010
XC	0.0513 ± 0.0053	0.3180 ± 0.0095	$68.7^{+3.9}_{-5.3}$	0.9572 ± 0.0098	0.8172 ± 0.0056
GCph+WL	0.0502 ± 0.0020	0.3201 ± 0.0036	67.2 ± 1.5	0.9595 ± 0.0036	0.8153 ± 0.0058
3x2pt	0.0502 ± 0.0020	0.3200 ± 0.0021	$67.2^{+1.2}_{-1.4}$	0.9596 ± 0.0036	0.8154 ± 0.0033
GCph+WL+GCsp	0.05007 ± 0.00079	0.32000 ± 0.00086	67.06 ± 0.62	0.9598 ± 0.0025	0.8154 ± 0.0012
3x2pt+GCsp	0.05009 ± 0.00079	0.32001 ± 0.00080	67.07 ± 0.60	0.9598 ± 0.0024	0.8155 ± 0.0011

Table 6.5: Recovered best fit values and 68% errors for the cosmological parameters for the considered probe (and probe combinations) when the nuisance parameters were let free.

	$b_1^{\rm GCph}$	$b_2^{\rm GCph}$	$b_3^{\rm GCph}$	$b4^{\rm GCph}$	$b5^{\rm GCph}$
Fiducial	1.0997727037892875	1.220245876862528	1.2723993083933989	1.316624471897739	1.35812370570578
GCph	1.0976 ± 0.0067	1.2177 ± 0.0071	1.2697 ± 0.0076	1.3139 ± 0.0072	1.3549 ± 0.0076
XC	1.101 ± 0.017	1.220 ± 0.011	1.271 ± 0.012	1.3152 ± 0.0083	1.353 ± 0.010
GCph+WL	1.1000 ± 0.0077	1.2205 ± 0.0082	1.2726 ± 0.0087	1.3169 ± 0.0089	1.3583 ± 0.0094
3x2pt	1.0998 ± 0.0050	1.2203 ± 0.0048	1.2724 ± 0.0051	1.3167 ± 0.0054	1.3581 ± 0.0056
GCph+WL+GCsp	1.0998 ± 0.0028	1.2203 ± 0.0024	1.2724 ± 0.0028	1.3167 ± 0.0030	1.3581 ± 0.0032
3x2pt+GCsp	1.0997 ± 0.0027	1.2203 ± 0.0022	1.2724 ± 0.0026	1.3166 ± 0.0029	1.3581 ± 0.0030

Table 6.6: Recovered best fit values and 68% errors for the first five photometric bias nuisance parameters for the considered probe (and probe combinations) when the cosmological parameters were let free.

	$b_6^{\rm GCph}$	$b_7^{\rm GCph}$	$b_8^{\rm GCph}$	$b9^{\rm GCph}$	$b10^{\rm GCph}$
Fiducial	1.3998214171814918	1.4446452851824907	1.4964959071110084	1.5652475842498528	1.7429859437184225
GCph	$1.3971^{+0.0083}_{-0.0095}$	1.4416 ± 0.0085	1.4933 ± 0.0089	1.5621 ± 0.0091	1.739 ± 0.011
XC	1.397 ± 0.013	1.442 ± 0.013	1.496 ± 0.016	1.560 ± 0.016	1.750 ± 0.029
GCph+WL	1.4000 ± 0.0095	1.4449 ± 0.0098	1.497 ± 0.010	1.565 ± 0.010	1.743 ± 0.012
3x2pt	1.3999 ± 0.0058	1.4446 ± 0.0059	1.4965 ± 0.0061	1.5653 ± 0.0060	1.7429 ± 0.0074
GCph+WL+GCsp	1.3999 ± 0.0034	1.4446 ± 0.0033	1.4965 ± 0.0031	1.5653 ± 0.0030	1.7430 ± 0.0044
3x2pt+GCsp	1.3998 ± 0.0033	1.4446 ± 0.0031	1.4965 ± 0.0030	1.5652 ± 0.0029	1.7429 ± 0.0043

Table 6.7: Recovered best fit values and 68% errors for the first five photometric bias nuisance parameters for the considered probe (and probe combinations) when the cosmological parameters were let free.

	$b_1^{\rm GCsp}$	$b_2^{\rm GCsp}$	$b_3^{\rm GCsp}$	$b4^{\rm GCsp}$
Fiducial	1.4614804	1.6060949	1.7464790	1.8988660
GCsp	1.4615 ± 0.0057	1.6061 ± 0.0061	1.7465 ± 0.0065	1.8989 ± 0.0070
3x2pt+GCsp	1.4614 ± 0.0025	1.6060 ± 0.0028	1.7464 ± 0.0031	1.8988 ± 0.0033

Table 6.8: Recovered best fit values and 68% errors for the spectroscopic bias nuisance parameters for the considered probe (and probe combinations) when the cosmological parameters were let free.

	$A^{\rm IA}$	$\eta^{\rm IA}$
Fiducial	1.72	-0.41
WL	1.724 ± 0.076	-0.412 ± 0.094
XC	1.722 ± 0.038	-0.404 ± 0.046
GCph+WL	$1.748^{+0.079}_{-0.064}$	$-0.441^{+0.069}_{-0.093}$
3x2pt	1.722 ± 0.035	-0.411 ± 0.032
GCph+WL+GCsp	1.725 ± 0.073	-0.415 ± 0.086
3x2pt+GCsp	1.722 ± 0.035	-0.411 ± 0.032

Table 6.9: Recovered best fit values and 68% errors for the intrinsic alignment nuisance parameters for the considered probe (and probe combinations) when the cosmological parameters were let free.