Euclid Theoretical Predictions Specification Document (TPSD)

Document Num.:	EC-???-???
Issue:	0
Date:	February 25, 2016
Authors:	This document is under the custodianship of the Euclid Inter-Science Working Group Taskforce for Forecasting.
Authorised by:	TWG

Abstract: *This document contains the specification of the remit of the inter-science working group taskforce for forecasting and probe combination. Details of how this group relates to the SWGs is defined in the PPSAID.*

Document Change History

Title: *Euclid Theoretical Predictions Specification Document (TPSD)*			
Issue	**Rev.**	**Date**	**Reason for Change**
Draft 0	0	2015-09	Creation

Approved by:

NAME	**Organization**	**Date**	**Signature**

Contents

1 Aim

This document aims at responding to the general need of a common forecasting pipeline, developed in agreement with all SWGs. It outlines the approaches that Euclid cosmological parameter forecasting will take during the implementation phase (2015–) and up until launch. This will cover standard 'Fisher' forecasting methodology, as well as more complex predictions. Specifications for the primary probe forecasts are defined, as well as how these should be combined together and with external cosmological parameter constraints, in particular CMB data. This document will also define the general code infrastructure that the forecasting pipeline will use, in preparation for the theoretical calculations that will enter into the Euclid likelihood pipeline (that will take output from the SGS LE3, and theoretical computations outlined here).

Therefore this document covers:

1. Definition of the primary probe theoretical computation data flow
2. The specification of code comparisons at each stage of the data flow – in building a single common pipeline
3. Specifications for code comparison for existing non-trunk or external forecast codes
4. How to compare Fisher matrices for a minimal Fisher matrix code
5. Definition of coding practices – including language, licensing and publication policy when using IST products etc.
6. How the code interacts with the SGS
7. How to deal with new input provided by the SWGs (for example if there is a new parameter or a new forecasting method developed by the SWG, how does the IST react).

Applicable Documents

- PPSAID (Post Processing Science Analysis Implementation Document)
- Theory Parameter Defintion Document
- SGS Data Flow Down Description Data
- GDPRD (Ground Data Processing Requirements Document)
- OU RSDs (Requirement Specification Documents)
- MPD (Mission Performance Document)

2 Relation to the Euclid Likelihood Pipeline

This document describes how the *theoretical* calculations for the Euclid primary probes are made. These calculations are initially, before launch, to be used to create Euclid 'forecasts'. A forecast is a performance prediction for Euclid that makes a statement about the accuracy and precision that Euclid will be able to achieve the Level 0 objectives outlines in the Mission Performance Document (MPD). This document also describes how the primary probes will be combined.

The construction of the Euclid likelihood pipeline is described conceptually in the PPSAID. This describes how there are three elements that need to come together in order for Euclid to achieve its science objectives. These are 1) The output from SGS LE3 - i.e. statistics computed on the data, 2) Output from the cosmological simulations Science Working Group - who will create realisations of Universe that can be used to create cosmology-dependent

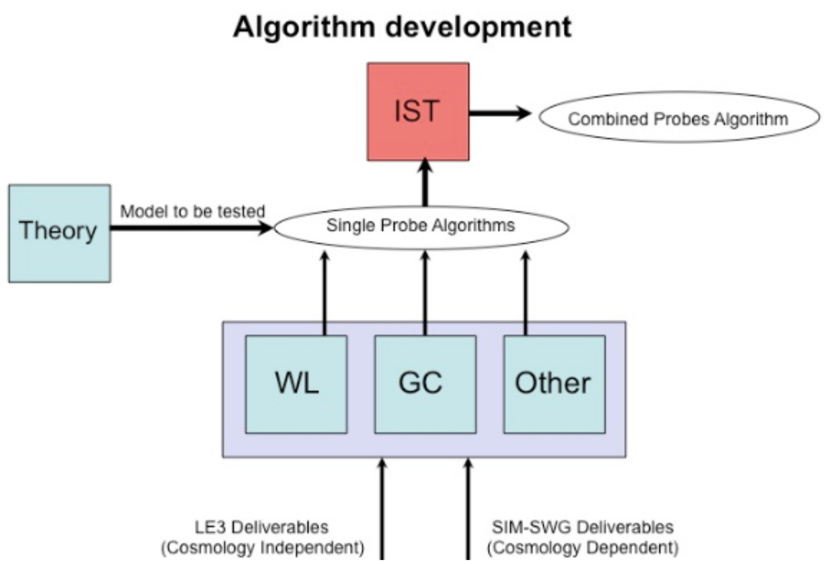


Figure 1: Figure from PPSAID. A schematic of the flow of information to create the algorithms for the individual probes and the combined probes. Note that the development of the algorithms is a collaborative effort between the cosmology SWGs and SGS.

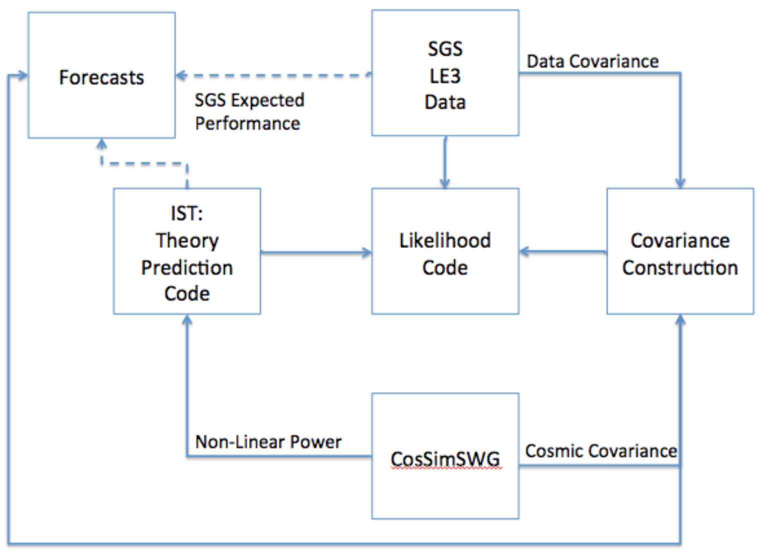


Figure 2: How the theory calculations relate to the SGS, CosmoSIMSWG and Euclid likelihood codes. The IST makes theoretical predictions that are incorporated into both forecasts and ultimately into likelihoods.

covariance estimates, and predictions for matter clustering, 3) Predictions of the measured statistics from theories against which the data can be tested (i.e. "forecasts" before launch).

The theory contruction formally does not involve the SGS algorithm development, but will be constructed such that the outputs from the SGS LE3 will be inputs into a likelihood pipeline that will also take inputs from the cosmological simulations and input from the theoretical algorithms. This relationship is show in Figure 2.

3 Definition of Cosmological Statistic Theory-Flows

The first step in creating the Euclid forecasting pipeline is to define the data-flow from cosmological parameters, and input survey characteristics, to theoretical predictions of relevant statistics. This is a similar procedure to that used in the Euclid ground segment to define the data-flow from pixels and spectro-photometry to Level-3 products. We adopt a very similar procedure here except that we refer to this as a theory-flow to demark it from SGS activities. The statistics, i.e. 'observables', that we will use as the primary probe are

- The weak lensing tomographic power spectrum
- Galaxy clustering: BAO, RSD, P(k)
- Secondary cosmological probes: Supernovae, clusters, ...

- **discussion point: do we need more e.g. 2nd-priority tiers for LE3, i.e. spherical-Bessel stats from the beginning**
- **discussion point: observables may change with the model: ex. isotropy tests may want different ℓ requirements.**

In the following we will define a theory-flow down for each of these primary probe statistics, and also for the combination of the primary probes, and the combination with CMB data. In doing this each element in the theory-flow will be referred to as a "**Processing Function (PF)**" using exactly the same naming convention as in the SGS for algorithm entities.

The broad grouping of the processing functions are into

- Inputs: Model and parameter definitions, including cosmological mode and parameters, and nuisance parameters.
- Inputs: Survey characteristics, including (expected) distributions of observed quantities, and survey geometry.
- Internal: Functions that act internally to the flow, for example matter power spectrum estimatation.
- Output: These are the expected theoretical statistics, and manipulations of those statistics e.g. derivatives, Fisher matrices etc.

Each processing function is given a unique identifier that refers to the pipeline(s) in which it is embedded, the nature of the processing function, and a short descriptor of its function.

Finally each theory-flow is required to have a matching data-flow in the SGS documentation i.e. that it is expected that the statistics that the theory-flow will compute from a theoretical side will also be computed by the SGS from the data for eventual inclusion in a Euclid likelihood pipeline. In the event that this document identifies missing data-flows then this should be discussed with the SWG leads, and SGS PO.

3.1 Weak Lensing: Tomographic Weak Lensing Theory-Flow

Here we specify the theory-flow for the tomographic weak lensing. In Figure 3 we show the flowdown from cosmological parameters, nuisance and survey parameters to output statistics and Fisher matrices.

3.2 Galaxy Clustering:

TBW

3.3 Weak Lensing+Galaxy Clustering:

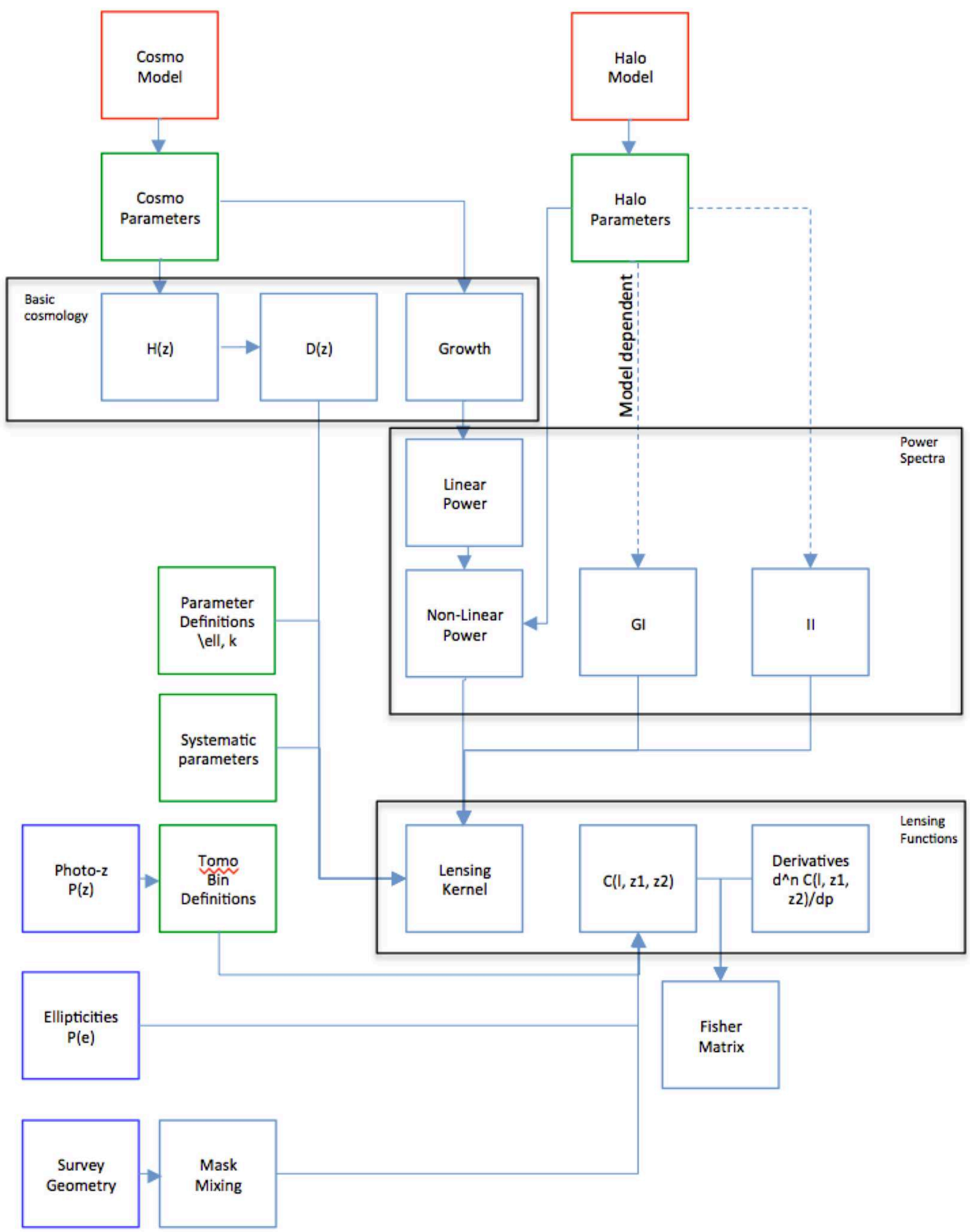
TBW

4 Coding, Requirements & Verification

Here we discuss some aspects of the coding, requirements and verification of the PFs.

4.1 Requirements

Each processing function must have an associated requirement. A requirement must be necessary, unambiguous and tracable. These requirements will be of severable different types. For example


Cosmo Model
Halo Model
Cosmo Parameters
Halo Parameters
Basic cosmology
H(z)
D(z)
Growth
Model dependent
Power Spectra
Linear Power
Non-Linear Power
GI
II
Parameter Definitions \ell, k
Systematic parameters
Lensing Functions
Photo-z P(z)
Tomo Bin Definitions
Lensing Kernel
C(l, z1, z2)
Derivatives d^n C(l, z1, z2)/dp
Ellipticities P(e)
Fisher Matrix
Survey Geometry
Mask Mixing

- Some processing functions will have trivial 'existence' requirements flow that are simply to ensure that parameters exist in the specified header file and that they are read in correctly. These parameters will be expected to be traceable either to the Parameter definition document, GDPRD, OU RSD, or SciRD requirement in which the parameter is defined in exact detail; either a parameter of interest (i.e. cosmological) or a nuisance parameter (i.e. a systematic).
- Some processing functions will have numerical 'budget allocations' that have flow-down from higher level breakdowns with regard to the top-level science objectives. For example an error on the Figure of Merit caused by binning in some quantity may have flow-down from a requirement that has shared a larger error-on-FoM budget with other PFs in the SGS or cosmological simulations PFs.

4.2 Verification

As well as requirements each PF will be expected to have a verification test associated with it. This means a specified set of tests that with known input produce known outputs such that the code can be verified to be computing what it is intended. In the case of theory-flows these verfication test should ideally be of a theoretical (analytic) nature in particular regimes of applicability, and numerical in others. Cross-validation verification tests are also encouraged for each PF, where several (> 2) independently written codes can be compared and the output compared for each on.

In the SGS as well as verification tests there are also 'validation' tests that compare the PF outputs to calibration data. There is no such data in the theory-flow case, however there are retrodictive tests on current data that can be performed. These tests will require, where possible, the output of the functions and statistics to be tested against current data, which should be consistent with the outputs (albiet at much large tolerance than will be required for Euclid). Some PF may be more amenable for this, for example certain parts of PFs that involve CMB cross-correlations can be compared against Planck data.

4.3 Coding Standards

We will adopt the same general coding practices as the SGS, in terms of language, library use and tracability. These can be found in the SGS SIRD document. However, as it is expected to be scientists and not software engineers that create these codes, we do not adopt the next level of engineering i.e. the use of LODEEN and CODEEN environments.

In order to facilitate code production we will use the following tools: **to be dicussed, TBW**

There is also the possibility of using external codes, outside Euclid for forecasting purposes. If this is the case a special committee will be formed and asked to assess the pros and cons of the use of such code, and the amount of time estimated to code a similar Euclid-specific version estimated. In the event that it is voted by the IST to be the best course to adopt such a code then the IST coordinators will form an MOU between the code authors and the Euclid IST should be sought, that should be done independently, and in addition to, the license agreements of the code. The MOU will cover in particular the ownership and license of any Euclid-specific modules that may contain prioritary information, authorship rights incurred, and so forth.

5 Forecasting Procedures

In this section we outline the two mechanisms through which robust forecasting for Euclid will be achieved.

5.1 Holistic Approach – Comparing Currently Existing Pipelines

The first approach will be to use *existing* forecasting codes that are already at a mature stage. In this case the full end-to-end theory-flow is being tested in a single step by comparing, for a given set of well defined inputs, the final Fisher matrix output or similar. In this it is required that

1. The inputs are well defined and common to all forecasting pipelines (e.g. parameter definitions)
2. That the outputs are well defined (e.g two-derivative Fisher matrix elements)
3. The metrics that will be used for the comparison (e.g. the DETF Figure of Merit), and a measure of agreement in these metrics that must be achieved in order to confirm that the codes are consistent.

In this case it is not possible to objective verify that any given code is returning the true performance, but the inter-subjectivity of more than one code can be verified by having many code do the comparison. In the event that there is inconsistency in the code then in the first instance simple debugging is encouraged and in the second instance an atomistic approach should be adopted.

The holistic approach will result in inter-subjectively verified, consistent, forecasts for Euclid, but will not constitute a production of a verified Euclid theory code. For this we will pursue the atomistic approach below.

The IST will provide for each theory-flow a set of templates that can be used to test forecasting pipeline in this holistic manner. These are provide in Appendices **to be defined**.

5.2 Atomistic Approach – Building a Euclid Cosmology Pipeline

Although the holistic approach to forecasting can result in forecasts that are external self-consistent, we also wish to build a theory code that verified to be both internal and externally self consistent. To this end we will develop the atomistic approach to forecast building where each PF in the theory-flow is compared on a case-by-case basis, such that the result is a likelihood code that – by construction – has some objectively verified forecasting capability built into it. This approach is significantly slower than a holistic approach, which we will pursue for more short term forecasting objectives.

For each PF the IST will create a series of PF requirements that must be met in order for the PF to be agreed to be usable. Each PF can be developed in isolation, although this may cause mis-matches between lower-level PF and upper-level ones when internal PF are written. Therefore we will pursue a 'top-down' construction i.e. from the inputs subsequently down through the theory-flow to the output level. For each PF the requirement and verification will define

1. Analytic examples where the output can be verified against a known truth i.e. systems tests of the PF. That must be reproduced at machine-precision.
2. A series of numerical convergence tests that the PF must be shown to reach stability in. In particular for derivative and integral approximations.
3. A series of metrics against which more than one independently code can be compared. These metrics should cover the whole range of functional output expected to be applicable for Euclid e.g. comoving distance at all redshifts, not only comoving distance 'figure of merit', and such metrics should agree at machine-precision.

We will pursue the construction of several theory pipelines written in several different languages such that language-specific libraries and assumptions are more ready identified as development moves down through the pipeline. At each stage a record must be kept of the requirements and evidence of the PF being passed.

5.3 Definition of Cosmological & Nuisance Parameters

Here we define the common set of cosmological and nuisance parameters that will be used as a starting point for both approaches to forecasting.

need to refer to the theory parameter definition document here.

References for GC: [1]. References for WL: [2] General references on Fisher and Euclid: [3, 4]

In the event that a SWG generates a new set of parameters, e.g. in the case that a new theory or new observations lead to the requirements that additional parameters are included in a forecast, this will be proposed by the SWG in question in a technical note written to the IST. The IST managers will then enter discussion with the SWG in question on the implementation. Any additional parameter will be recorded in the mission data base, and parameter definition documents, and this document will be updated accordingly in the even that a parameter is decided to be included.

6 Management of the IST

The IST will be subdivided into several workpackages to reflect the tasks at hand. These workpackages will be decided between the IST managers and the IST team, and recorded in the SWG Work Package Description document. Workpackages will be reviewed in an organic process, i.e. created and reviewed on a case by case basis as the need arises, on approximately a 6-month to 1-year rotating timetable.

7 End Comments

This document does the following:

1. Defines the scope of the IST with regard to the SGS, and SWGs.
2. Contains and defines the theory-flow for the primary cosmological probes.
3. Defines the procedure through which the primary cosmological probes forecasts will be constructed both through a holistic (end-to-end) approach and an atomistic (pipeline-building approach).
4. In the following appendices describes the defined holistic procedure of the GC and WL primary probes.

The following will be captured in specific separate documents that will refer to this one.

1. Requirements and verification specifications for each of the primary cosmological probe theory-flows (one for each flow).
2. Specifications of the atomistic specifications for each theory flow.
3. Results of the holistic and atomistic forecasts, these will appear documented as a Euclid technical notes, and may also lead to publications.

The above documents will be recorded on the Euclid redmine and uploaded to the Euclid ESA RSSD under the SWG section when ratified.

Appendix: Galaxy Clustering Holistic Forecast Definitions

[1] We recall the main equations to be used. Observed linear spectrum at all redshifts:

$$P_{r,obs}(k_{ref},\mu_{ref};z) = P_s(z) + \frac{D_A(z)^2_{ref}H(z)}{D_A(z)^2H_{ref}(z)}\frac{P(k,z)}{P(k,z=0)}b^2(z)(1+\beta(z)\mu^2)^2P(k,z=0)\,, \tag{1}$$

We use:

- For the linear matter power spectrum $P(k,z) \to$ input files provided by Enea Di Dio and Bin Hu.
- For the growth index $f(z,k)$ (needed for $\beta_d(z) = f(z)/b(z)$), $\to$ input files provided by Enea Di Dio and Bin Hu (k dependence not necessary. k fixed to k = 0.1 h/Mpc). Input files need to be done differently for future tests).
- For the $H(z)$ and $D_A(z)$ we use analytical expressions (see below, sec. 7). SC checked that the analytical derivatives wrt ω_b and ω_c (calculated considering h as a parameter) now match also the numerical ones.

Here $b(z)$ is the bias (provided by column 5 of Tab.(1)), $\beta(z)$ is the redshift-distortion factor, $P(k)$ is the undistorted linear matter spectrum (provided in the input with the correct normalization), μ is the direction cosine, D_A is the angular diameter distance (provided in the input) and H is the Hubble rate at the shell redshift z (provided in the input). $P_s(z)$ is the z- dependent shot-noise correction due to discreteness in the survey. The subscript ref (for 'reference') indicates quantities calculated in the fiducial model. In linear theory we have $\beta_d(z) = f(z)/b(z)$ where $f(z) \equiv d\log G/d\log a$ is the growth rate (f is provided directly in the input for every cosmology, G is the growth factor but you should not need it, the input provides directly the growth rate f).

The Fisher matrix for every redshift bin shell is integrated in k (Right now doing the μ integral from 0 to 1 and multiplying the integrand by 2)):

$$F_{ij} = \frac{1}{8\pi^2}\int_{-1}^{1}\mathrm{d}\mu\int_{k_{\min}}^{k_{\max}}k^2\cdot\mathrm{d}k\frac{\partial\ln P_{obs}(z;k,\mu)}{\partial\theta_i}\frac{\partial\ln P_{obs}(z;k,\mu)}{\partial\theta_j}\cdot\left[\frac{n\,P_{obs}(z;k,\mu)}{n\,P_{obs}(z;k,\mu)+1}\right]^2 V_{\mathrm{s}} \tag{2}$$

where the θ's are the parameters: $\theta = \{\omega_c \equiv \Omega_{c_0}h^2, \omega_b \equiv \Omega_{b_0}h^2, h, n_s, \ln(10^{10}A_s), b_i, P_s\}$ (τ has been fixed to fiducial because it is degenerate when using only WL and GC).

The input provides the spectra $P(z,k)$ (already corresponding to the A_s normalization) and all other functions $f(z)$, $D_A(z)$, $H(z)$ for each model.

$n(z)$ is the galaxy number density at redshift z and V_s is the survey volume of the entire (so all 4π) redshift shell dz. The power spectra are all obtained numerically by solving the background and perturbation equations of the system for each value of the parameters and they are provided as input.

The derivative of the galaxy power spectrum, at each redshift (so z is fixed in the middle of the bin), with respect to the parameters (in this case ω_c) is:

$$\begin{aligned}\left.\frac{\mathrm{d}\ln P_{\mathrm{obs}}(\bar z,k,\mu;\theta_i)}{\mathrm{d}\theta_i}\right|_{fid} &= \frac{\partial\ln P_{obs}(\bar z,k,\mu;\theta_i)}{\partial\ln P_s(\bar z)}\frac{\partial\ln P_s(\bar z)}{\partial\theta_i} + \frac{\partial\ln P_{obs}(\bar z,k,\mu;\theta_i)}{\partial\ln f(\bar z)}\frac{\partial\ln f(\bar z)}{\partial\theta_i}\\ &+ \frac{\partial\ln P_{obs}(\bar z,k,\mu;\theta_i)}{\partial\ln H(\bar z)}\frac{\partial\ln H(\bar z)}{\partial\theta_i} + \frac{\partial\ln P_{obs}(\bar z,k,\mu;\theta_i)}{\partial\ln D_A(\bar z)}\frac{\partial\ln D_A(\bar z)}{\partial\theta_i}\\ &+ \frac{\partial\ln P_{obs}(\bar z,k,\mu;\theta_i)}{\partial\ln P(k,\bar z)}\frac{\partial\ln P(k,\bar z)}{\partial\theta_i} + \frac{\partial\ln P_{obs}(\bar z,k,\mu;\theta_i)}{\partial\ln b(\bar z)}\frac{\partial\ln b(\bar z)}{\partial\theta_i}\\ &+ \frac{\partial\ln P_{obs}(\bar z,k,\mu;\theta_i)}{\partial k}\frac{\partial k}{\partial\theta_i} + \frac{\partial\ln P_{obs}(\bar z,k,\mu;\theta_i)}{\partial\mu}\frac{\partial\mu}{\partial\theta_i}\end{aligned} \tag{3}$$

[1] For the moment this part has been extracted from the TWG documentation of the code comparison, whose contributors were: L.Amendola, S. Camera, Santiago Casas, Enea Di Dio, Bin Hu, Elisabetta Majerotto, Valeria Pettorino, Domenico Sapone.

In general, one needs to include a redshift error σ_z and further multiply: $P_{obs} = P_{obs} e^{-k^2\mu^2\sigma_z^2/H(z)^2}$. The last two terms are non-vanishing if one takes the Alcock-Paczynski effect into account for k and μ, since they are also affected by geometrical terms.

The derivatives are:

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln P_s(\bar{z})} = \frac{1}{P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)} \tag{4a}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln f(\bar{z})} = \frac{2\beta(\bar{z})\mu^2}{1+\beta(\bar{z})\mu^2} \tag{4b}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln b(\bar{z})} = \frac{2}{1+\beta(\bar{z})\mu^2} \tag{4c}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln P(k, \bar{z})} = 1 \tag{4d}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln H(\bar{z})} = 1 + \frac{4\beta(z)\mu}{1+\beta(z)\mu^2}(-\mu(\mu^2-1)) + \frac{\partial \ln P\left(\bar{z}, k; \theta_i\right)}{\partial k} k\mu^2 \tag{4e}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln D_A(\bar{z})} = -2 + \frac{4\beta(z)\mu}{1+\beta(z)\mu^2}(-\mu(\mu^2-1)) + \frac{\partial \ln P\left(\bar{z}, k; \theta_i\right)}{\partial k} k(\mu^2-1) \tag{4f}$$

All these terms are set as follows:

- $\frac{\mathrm{d}\ln P(\bar{z},k)}{\mathrm{d}\theta_i}$ are numerical derivatives, coming from the input files (of course $P(\bar{z}, k)$ does not depend on bias or P_s). The input files are however provided at different k's; therefore, when calculating the derivative you need to interpolate in k first and then calculate the difference at fiducial $(1 \pm \epsilon)$.
- $f(\bar{z}, k)$ is given by the input file: while in the input f is a function of k, for the comparison, we fix k = 0.1 h/Mpc.
- $\frac{\mathrm{d}\ln f(\bar{z})}{\mathrm{d}\theta_i}$ are numerical derivatives, coming from the input files (k fixed as above).
- $\frac{\mathrm{d}\ln H(\bar{z})}{\mathrm{d}\theta_i}$ and $\frac{\mathrm{d}\ln D_A(\bar{z})}{\mathrm{d}\theta_i}$ are given by the analytical expressions described below (sec. 7)

We tested three different methods:

- Direct derivative: you integrate directly all terms of equation 3 into eq.2;
- Numerical derivative: evaluate the numerical derivative of eq.2 just calculating $P_{obs}(\bar{z}, k)$ at $\theta_{i,fid}(1 \pm \epsilon)$ and then divide by $2\epsilon\theta_{i,fid}$;
- Jacobian method (or extended BAO Seo-Eisenstein method, see [1])

All three methods match when there is no AP effect. With AP effect, the Jacobian method differs from the other two. For the following comparison, fisher matrices are provided using the direct derivative method.

The Hubble parameter and the angular diameter distance

We here assume that we consider h as a parameter, rather than Ω_Λ. The Hubble Parameter, for the cosmology we are using, is

$$H(z) = 100h\sqrt{\frac{\Omega_c h^2 + \Omega_b h^2 + \Omega_\nu h^2}{h^2}(1+z)^3 + 1 - \frac{\Omega_c h^2 + \Omega_b h^2 + \Omega_\nu h^2}{h^2}} \equiv 100E(z) \;\; [\text{h km/s/Mpc}] \tag{5}$$

We then take into account the variation of $\omega_c = \Omega_{c,0}h^2$ and $\omega_b = \Omega_{b,0}h^2$. The derivatives of the logarithm of Hubble parameter with respect to ω_c, ω_b and h are:

$$\frac{\partial \ln H}{\partial \omega_c}|_{\omega_b,h} = \frac{(1+z)^3-1}{2h^2E^2(z)} \tag{6}$$

$$\frac{\partial \ln H}{\partial \omega_b}|_{\omega_c,h} = \frac{(1+z)^3-1}{2h^2E^2(z)} \tag{7}$$

$$\frac{\partial \ln H}{\partial h}|_{\omega_b,\omega_c} = \frac{1}{hE^2(z)} - \frac{1}{h} \tag{8}$$

and the other derivatives of the Hubble parameters with respect to n_s, A_s and τ are clearly zero. Note that had we chosen Ω_i as variables instead of ω_i, the derivative of the Hubble function with respect to h in units of [h km/s/Mpc] would have been zero. The angular diameter distance is

$$D_A(z) = \frac{1}{H_0}\frac{c}{1+z}\int_0^z \frac{dx}{E(x)} \tag{9}$$

where c is the speed of light and it is important to have distances in Mpc/h and

$$H_0 = \frac{100\,h\,\mathrm{Km/s/Mpc}}{c} = \frac{1}{2997.92}\frac{h}{\mathrm{Mpc}}\,. \tag{10}$$

The derivatives of the angular diameter distance are:

$$\frac{\partial \ln D_A}{\partial \omega_c} = \frac{1}{D_A(z)}\frac{c}{1+z}\int_0^z -\frac{\partial \ln H(x)}{\partial \omega_c}\frac{dx}{H(x)} \tag{11}$$

$$\frac{\partial \ln D_A}{\partial \omega_b} = \frac{1}{D_A(z)}\frac{c}{1+z}\int_0^z -\frac{\partial \ln H(x)}{\partial \omega_b}\frac{dx}{H(x)} \tag{12}$$

$$\frac{\partial \ln D_A}{\partial h} = \frac{1}{D_A(z)}\frac{c}{1+z}\int_0^z -\frac{\partial \ln H(x)}{\partial h}\frac{dx}{H(x)} \tag{13}$$

All these derivatives will be used in the Fisher matrix. Galaxy density: take column 4 of Table (1), (dn3) as dn/dz/$d\Omega$ (with Ω in square degrees). There is no need to add in an extra redshift success rate (e.g. epsilon=0.35) as before - this is already included in the numbers, which now account for redshift- and flux-dependent success rate, in addition to using an improved mock catalogues and redshift-measurement technique. There is also no need to include any scaling. These numbers are for the standard integration time (540 sec per dither, standard dither pattern). As an example, to get the total number of galaxies in a bin of width $0.95 < z < 1.05$, one would calculate:

(Number of galaxies $0.95 < z < 1.05$) = $4825.80 * 0.1 * 15000 = 7,239,000$

So in a Euclid survey over $15,000 deg^2$ we would get good redshifts for $7,239,000$ galaxies in the redshift range $0.95 < z < 1.05$. Note the 0.1 factor required to turn dn/dz from a derivative to a number in a bin of width $Dz = 0.1$.

This gives you the total number of galaxies. Then the number density that you need in the fisher for GC, n in eq.2 in the formula for the fisher matrix, is the density per volume in each bin, therefore equal to:

$$n = dn3/dV_{dz} * dz \tag{14}$$

where the dz is the width of the redshift shell (i.e 0.1 but see the file) and

$$dV_{dz} = \frac{\frac{4\pi}{3}\left[(1+z_2)^3 D_A[z_2;\theta]^3 - (1+z_1)^3 D_A[z_1;\theta]^3\right]}{4\pi\left(\frac{180}{\pi}\right)^2} \tag{15}$$

is the volume in the redshift shell per degree squared. The survey Volume is

$$V_s = \text{Area(i.e. } 15,000 \text{ sq deg)} * dV_{dz} \tag{16}$$

The third column (dVol) should not be used anywhere.

Table 1: Specifications of the redbook, (derived from Orsi et al. 2009).

zmin	zmax	dVol	dn3	b3
0.35	0.45	3.60	0.000	0.992
0.45	0.55	5.10	0.000	1.024
0.55	0.65	6.40	0.000	1.053
0.65	0.75	7.90	2434.280	1.083
0.75	0.85	9.20	4364.812	1.125
0.85	0.95	10.30	4728.559	1.104
0.95	1.05	11.70	4825.798	1.126
1.05	1.15	12.30	4728.797	1.208
1.15	1.25	13.30	4507.625	1.243
1.25	1.35	14.00	4269.851	1.282
1.35	1.45	14.60	3720.657	1.292
1.45	1.55	15.10	3104.309	1.363
1.55	1.65	15.60	2308.975	1.497
1.65	1.75	16.00	1541.831	1.486
1.75	1.85	16.20	1474.707	1.491
1.85	1.95	16.50	893.716	1.573
1.95	2.05	16.70	497.613	1.568

GC comparison: where to start

In Case 1,2,3 the following common options are fixed:

1. k_{min} : 0.001 h/Mpc
2. $k_{max} = 0.2$ h/Mpc
3. k-binning: use directly the one of the spectra provided in the input
4. No relativistic corrections.
5. Area : 15,000 sq deg
6. Redshift range : $0.7 < z < 2.0$
7. Binning: use all bins as in the table, starting from 0.65 (the first non-zero in dn3).

GC case 1: minimal setting.

In addition to the common settings above, use these settings for your first comparison.

1. Bias: fixed to the fiducial value (column 5, b3 in the table 1). All derivatives wrt the bias are zero.
2. The AP effect is zero: the four derivatives of k and μ with respect to H and D_A are set to zero.
3. $P_{shot} = 0$ (fixed)
4. $\sigma_z = 0$

The Fisher matrix to compare with for this case, is case1.dat. If it matches, proceed to case2.

GC case 2: including the bias

In addition to the common settings above, case2 fixes:

1. Bias: marginalization over bias, derivatives as in 4c.
2. The AP effect is zero: the four derivatives of k and μ with respect to H and D_A are set to zero.
3. $P_{shot} = 0$ (fixed)
4. $\sigma_z = 0$

The Fisher matrix to compare with for this case, is case2.dat.
If it matches, proceed to case3.

GC case 3: including the AP effect

In addition to the common settings above, case3 fixes:

1. Bias: marginalization over bias, derivatives as in 4c.
2. The AP effect is included.
3. $P_{shot} = 0$ (fixed)
4. $\sigma_z = 0$

The Fisher matrix to compare with for this case, is case3.dat.
If it matches, proceed to case4.

GC case 4: including the shot noise

In addition to the common settings above, case4 fixes:

1. Bias: marginalization over bias, derivatives as in 4c.
2. The AP effect is included.
3. P_{shot}: marginalisation over P_{shot}; the fiducial is still zero; derivatives are calculated according to 4a
4. $\sigma_z = 0$

GC case 5: Constraining $\ln D_A(z)$, $\ln H(z)$, $f\sigma_8(z)$, $b\sigma_8(z)$, $P_s(z)$ at each redshift bin

In this case, we want to constrain some cosmological functions that depend on redshift, especially $f\sigma_8(z)$, $b\sigma_8(z)$.
For doing this, we rewrite eqn. 1 as:

$$P_s(z) + \frac{D_A(z)^2_{\text{ref}} H(z)}{D_A(z)^2 H_{\text{ref}}(z)} \left(b\sigma_8(z) + f\sigma_8(z)\mu^2\right)^2 \frac{P(k,z)}{\sigma_8^2(z)} \tag{17}$$

The full list of parameters is the following:

- n cosmological parameters, in our case $n = 5$: $\theta_i = \{\Omega_c h^2,\, \Omega_b h^2,\, h,\, n_s,\, \ln 10^{10} A_s\}$
- $5 \times n_{bin}$ redshift-dependent parameters: $\{\ln D_A(z),\, \ln H(z),\, f\sigma_8(z),\, b\sigma_8(z),\, P_s(z)\}$

Now we will have small modifications on the derivatives:

$$\frac{\partial \ln P_{obs}\left(\bar{z},k,\mu;\theta_i\right)}{\partial \ln P_s(\bar{z})} = \frac{1}{P_{obs}\left(\bar{z},k,\mu;\theta_i\right)} \tag{18a}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z},k,\mu;\theta_i\right)}{\partial f\sigma_8(\bar{z},k)} = \frac{2\mu^2}{b\sigma_8(\bar{z})+\mu^2 f\sigma_8(\bar{z},k)} \tag{18b}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z},k,\mu;\theta_i\right)}{\partial b\sigma_8(\bar{z})} = \frac{2}{b\sigma_8(\bar{z})+\mu^2 f\sigma_8(\bar{z},k)} \tag{18c}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z},k,\mu;\theta_i\right)}{\partial \ln H(\bar{z})} = 1+\mu^2\left(1-\mu^2\right)\frac{f\sigma_8(\bar{z},k)}{b\sigma_8(\bar{z})+\mu^2 f\sigma_8(\bar{z},k)} + k\mu^2\frac{\partial \ln P(k,z)}{\partial k} \tag{18d}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z},k,\mu;\theta_i\right)}{\partial \ln D_A(\bar{z})} = -2+4\mu^2\left(1-\mu^2\right)\frac{f\sigma_8(\bar{z},k)}{b\sigma_8(\bar{z})+\mu^2 f\sigma_8(\bar{z},k)} - k(1-\mu^2)\frac{\partial \ln P(k,z)}{\partial k} \tag{18e}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z},k,\mu;\theta_i\right)}{\partial \theta_i} = \frac{\partial \ln P(k,z)}{\partial \theta_i} - \frac{2}{\sigma_8(\bar{z})}\frac{\partial \sigma_8(\bar{z})}{\partial \theta_i}. \tag{18f}$$

The right hand side of eqn.18f is zero, when $\theta_i = \ln 10^{10} A_s$, so we should not add this parameter when constraining at the same time $\sigma_8(z)$. A Fisher matrix for comparison, will be added once this detail is solved.

Appendix: Weak Lensing Holistic Forecast Definitions

7.1 Lensing projections

[Martin K: Generalized equations for non-flat Universe.]

The simplest form of the convergence power spectrum with (correlated) sources in redshift bins centered around z_i and z_j is

$$P_{ij}(\ell) = \frac{9}{4}\Omega_{\rm m}^2\left(\frac{H_0}{c}\right)^2\int_0^{z_{\rm lim}} {\rm d}z\,\frac{W_i(z)W_j(z)}{(1+z)^4}P_{\rm m}\left(\frac{\ell}{f_K[r(z)],z}\right), \tag{19}$$

where $W_i(z)$ is the window function and r(z) is the comoving distance, which in flat space, is equal to the comoving angular diameter distance: This approximate expression uses the Limber approximation (collecting only modes that lie in the plane of the sky, thereby neglecting correlations along the line of sight, [5, 6]), the small-angle approximation (expanding to first order trigonometric functions of the angle) and the flat-sky limit (replacing spherical harmonics by Fourier transforms)

$$r(z) = D_{\rm A}(1+z) \tag{20}$$

This is provided in the input, column 3 of table VII. Please, for this comparison, do not use other methods to calculate it.

$$W_i(z) = \int_z^{z_{\rm lim}} {\rm d}\tilde{z}\,\frac{f_K[r(z)-r(\tilde{z})]}{f_K[r(\tilde{z})]}n_i(\tilde{z}) \tag{21}$$

where $r(z)$ is the coming distance. The redshift distribution for bin i is denoted with n_i, and it is normalized to the number of galaxies belonging to that bin,

$$\int_0^{z_{\rm lim}} {\rm d}z\, n_i(z) = 1. \tag{22}$$

[Martin K: Modified, n should depend on $\tilde{z}$, not on r]

7.2 Redshift distribution

The window function depends on $n_i(z)$, the galaxy distribution in the i-th redshift bin: this is convolved with a Gaussian to account for photometric redshift errors σ_z (value specified below), i.e.

$$n_i(z) = A \int_{i\text{-th bin}} n(\tilde{z}) \exp\left(\frac{-(\tilde{z}-z)^2}{2\sigma_z^2}\right) \mathrm{d}\tilde{z} \tag{23}$$

where the integral is done over $\tilde{z}$ for the single i-th bin. $A = \frac{1}{\sqrt{2\pi}\sigma_z}$ a normalization factor. In eq. (23), the density is given by:

$$n(z) = z^2 exp(-(z/z_0)^{3/2}) \tag{24}$$

where $z_0 = z_{mean}/1.412$ is the peak of $n(z)$ and z_{mean} is the median redshift (value specified below).

7.3 Likelihood function

7.4 Covariance matrix

7.5 Real-space forecasts

7.6 Systematics

7.6.1 Instrumental and observational systematics

7.6.2 Astrophysical systematics

7.7 Intrinsic alignment

7.8 Non-linear power spectrum

7.8.1 Baryonic corrections

7.9 Fisher matrix

Then the Fisher matrix is summed over all multiples:

$$F_{\alpha\beta} = f_{\text{sky}} \sum_{\ell,i,j,k,m} \frac{(2\ell+1)\Delta\ell}{2} \frac{\partial P_{ij}(\ell)}{\partial\theta_\alpha} C^{-1}_{jk} \frac{\partial P_{km}(\ell)}{\partial\theta_\beta} C^{-1}_{mi} \tag{25}$$

with the covariance matrix

$$C_{jk} = P_{jk} + \delta_{jk}\gamma^2_{\text{int}} {q_j}^{-1}, \tag{26}$$

where γ_{int} is the intrinsic galaxy shear and

$$q_j = 3600 \left(\frac{180}{\pi}\right)^2 n_\theta \int_0^\infty n_j(z) dz\,. \tag{27}$$

Here n_θ is the galaxy density per arcmin^2 and $n_j(z)$ is the galaxy density for the j-th redshift bin, as defined in eq.(23).

Since we consider multipoles up to $\ell_{\text{max}} = 5000$, we need to apply non-linear corrections to the matter power spectrum, for which we use the halofit (provided in the input spectra).

Specifications for the weak lensing:

1. Area: 15,000 sq deg
2. $f_{sky} = Area/(4\pi(180/\pi)^2)$
3. Median Redshift: $z_{mean} = 0.9$
4. $\gamma_{\rm int} = 0.22$
5. $n_\theta = 30$ galaxies per $\rm arcmin^2$
6. Error on photometric redshift: σ_z = 0.05 (1+z)
7. Linear Power Spectrum: CAMB (provided in the input, you shouldn't use it)
8. Method to correct for non-linearities: Halo model (however the input files are now linear; it does not matter for now as we are not looking for the correct Fisher, just comparing. For this comparison we are using linear spectra).
9. Binning: 10 bins from z = 0 - 2.5. We assume equipopulated bins: the redshift bins chosen such that each contain the same amount of galaxies. need to add formula
10. k-range: $k_{min} = 0.001$ h/Mpc, $k_{max} = 0.5$ h/Mpc
11. k-binning: use directly the one provided in the input
12. ℓ-range: $\ell_{min} = 5$, ℓ_{max} fixed as a function of z, using the table provided by Enea (uploaded in the wiki). No cut in z.
13. ℓ-binning: 100 bins (with log spaced binning)

References

[1] H.-J. Seo and D. J. Eisenstein. Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys. *ApJ*, 598:720–740, December 2003.

[2] Z. Ma, W. Hu, and D. Huterer. Effects of Photometric Redshift Uncertainties on Weak-Lensing Tomography. *ApJ*, 636:21–29, January 2006.

[3] Luca Amendola and Shinji Tsujikawa. *Dark Energy: Theory and Observations*. Cambridge University Press, 2010.

[4] L. Amendola, S. Appleby, D. Bacon, T. Baker, M. Baldi, N. Bartolo, A. Blanchard, C. Bonvin, S. Borgani, E. Branchini, C. Burrage, S. Camera, C. Carbone, L. Casarini, M. Cropper, C. de Rham, C. Di Porto, A. Ealet, P. G. Ferreira, F. Finelli, J. García-Bellido, T. Giannantonio, L. Guzzo, A. Heavens, L. Heisenberg, C. Heymans, H. Hoekstra, L. Hollenstein, R. Holmes, O. Horst, K. Jahnke, T. D. Kitching, T. Koivisto, M. Kunz, G. La Vacca, M. March, E. Majerotto, K. Markovic, D. Marsh, F. Marulli, R. Massey, Y. Mellier, D. F. Mota, N. Nunes, W. Percival, V. Pettorino, C. Porciani, C. Quercellini, J. Read, M. Rinaldi, D. Sapone, R. Scaramella, C. Skordis, F. Simpson, A. Taylor, S. Thomas, R. Trotta, L. Verde, F. Vernizzi, A. Vollmer, Y. Wang, J. Weller, and T. Zlosnik. Cosmology and Fundamental Physics with the Euclid Satellite. *Living Reviews in Relativity*, 16:6, September 2013.

[5] D. N. Limber. The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. *ApJ*, 117:134–+, January 1953.

[6] N. Kaiser. Weak gravitational lensing of distant galaxies. *ApJ*, 388:272–286, April 1992.