The future of Cosmology with Euclid:
hopes & challenges

Guadalupe Cañas-Herrera, Matteo Martinelli, Stéphane Ilic

Simulation-based inference with Swyft

Thursday 26th January, 2023

Disclaimer:

Amazing work of Stefano Camera, Matteo Martinelli, Santiago Casas, Sergio Di Domizio, Andrea Pezzotta, Stefano Davini, Amandine Le Brun, Linda Blot, Davide Sciotti,

Konstantinos Tanidis, Isaac Tutusaus, Marco Bonici, Ziad Sakr, Stephane Ilic, Virginia Ajani, Domenico Sapone,

Sam Farrens,

Vincenzo Cardone, Shahab Joudaki, Valeria Pettorino

Overview of data analysis for testing cosmological models

DATA

Model
Define the cosmological model we aim to test against data (i.e: ΛCDM, extensions...)

Cosmological observables
Compute theoretical predictions for the observational probes, to be compared to data

Likelihood
Construct a likelihood to compare theory with data

Bayesian Inference
Obtain the posterior distributions of the parameters of interest

Euclid:

A mission to map the Dark Universe

Euclid in a nutshell

Euclid Spacecraft in TAS-F
(Nov 2022)

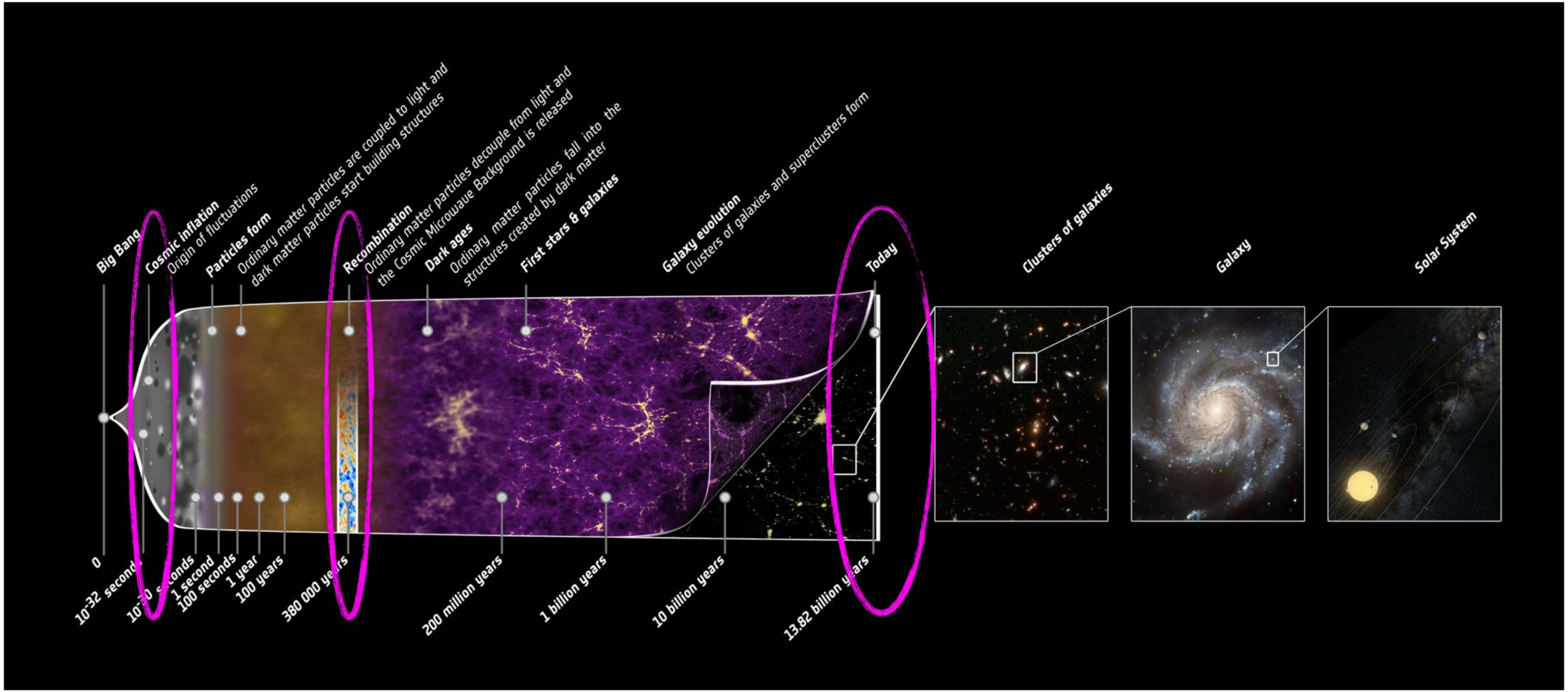
- Launch in summer 2023!
- Using a (~~Soyuz)~~ (~~Ariane 6~~) Falcon 9
- operations from L2
- M-class mission of the ESA “Cosmic Vision” Science programme
- Manufacturer: Thales Alenia Space & Airbus
- Joint of two proposed missions: DUNE and SPACE

Euclid in a nutshell

Euclid Spacecraft in TAS-F
(Nov 2022)

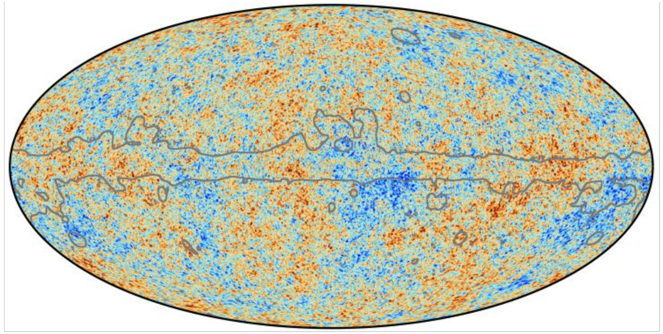
- A mirror telescope of 1.2m diameter
- Two instruments: VIS (visible) and NISP (near infrared)
- Area of 1500 deg^2
- Photometric catalogue ($0 < z < 2$) using $\sim 10^9$ galaxies
- Spectroscopic catalogue ($0.7 < z < 1.8$) using $\sim 10^6$ galaxies

Cosmology 101



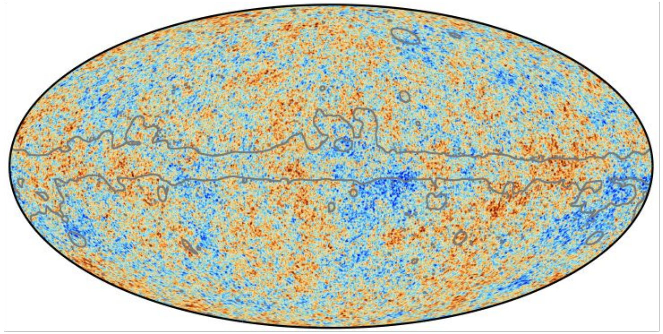
[NASA - Planck 14]

Euclid Primary Observables



CMB

Euclid Primary Observables

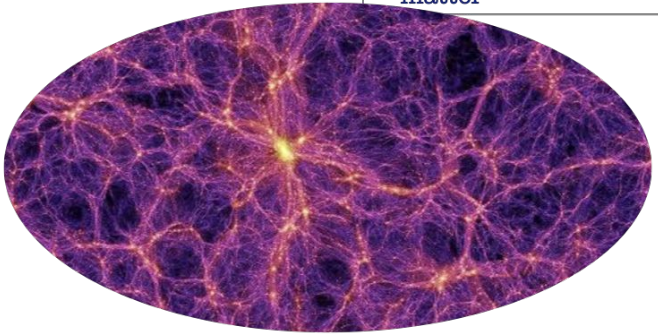


CMB

➡ Apparent change in the shape of galaxies due to the deflection of light as it travels from a source to an observer

Weak Lensing

➡ Distribution of positions of galaxies in the sky to study the relation between baryonic and dark matter



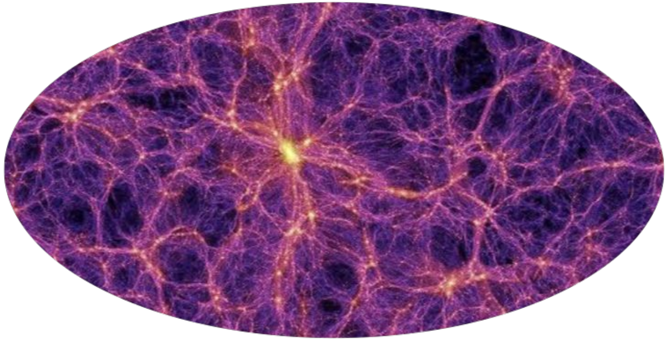
Galaxy Clustering

Euclid Primary Observables

Towards the Era of LSS Cosmology

The Large Scale Structure of our Universe contains precious information about the **main components** of the underlying **Cosmological model** and the **initial conditions** of the universe

Weak Lensing

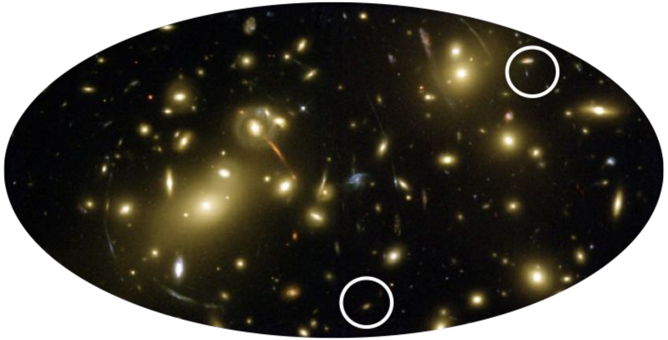


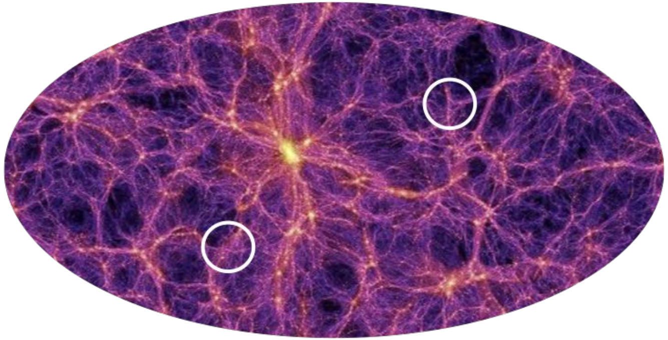
Galaxy Clustering

Euclid Primary Observables

Towards the Era of LSS Cosmology

The Large Scale Structure of our Universe contains precious information about the **main components** of the underlying **Cosmological model** and the **initial conditions** of the universe

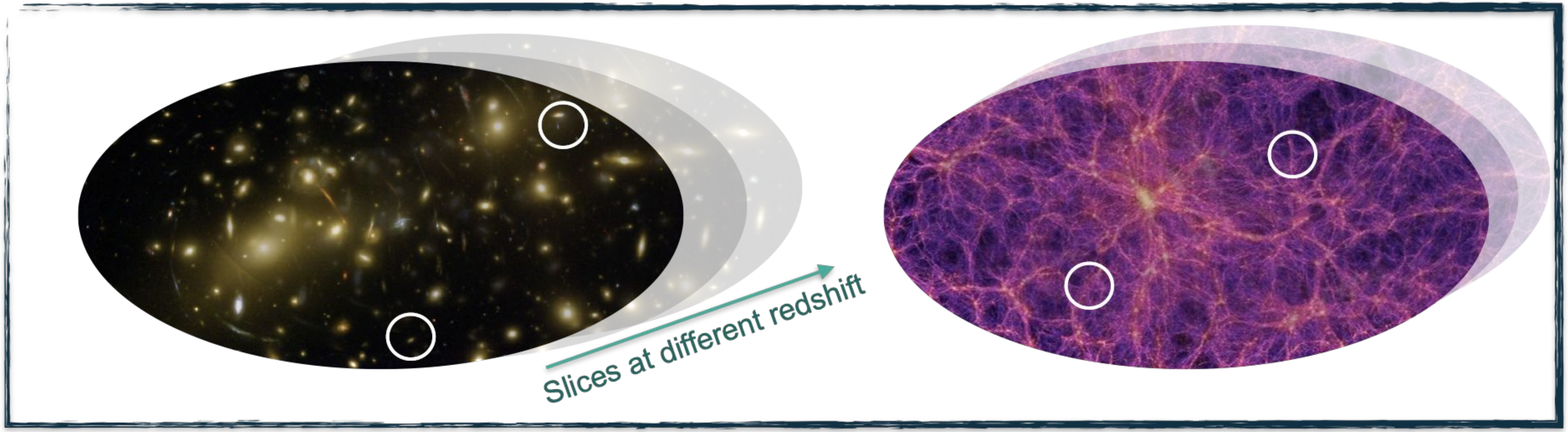




Euclid Primary Observables

Towards the Era of LSS Cosmology

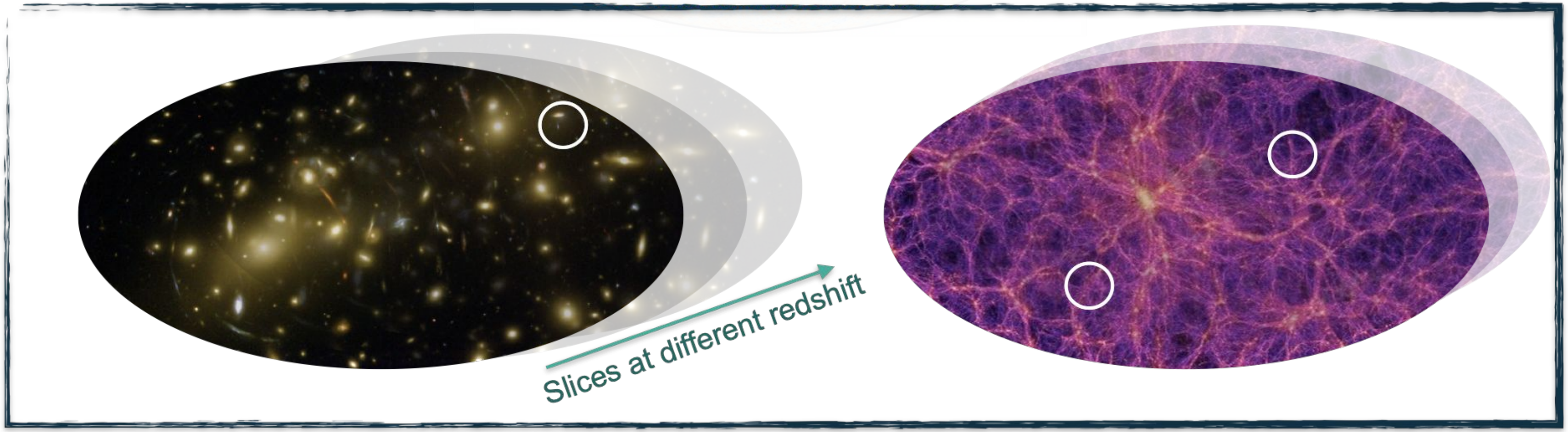
We are insterested in the 3D distribution: the goal is to create **3D maps** using **tomography**



Euclid Primary Observables

Euclid primary cosmological probes

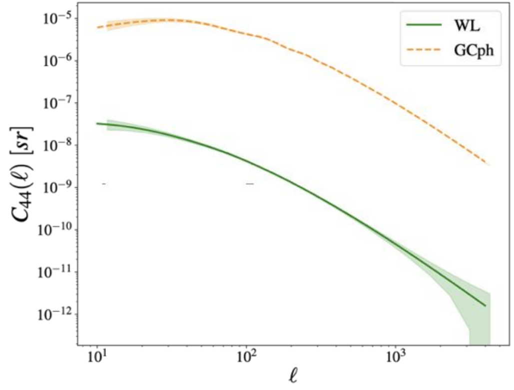
[ESA- Euclid operations 3D model, 19]



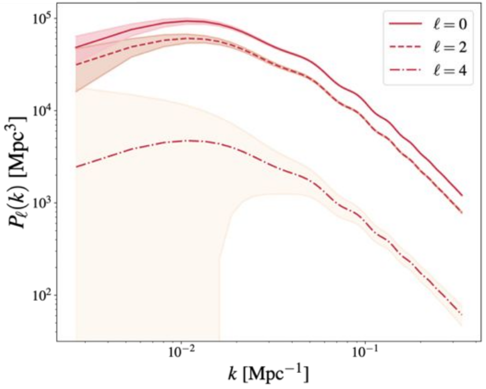
Euclid Primary Observables

Euclid primary cosmological probes

[ESA- Euclid operations 3D model, 19]



The Cosmological Likelihood for Observables in Euclid (**CLOE**) will predict theoretically Euclid observables and will compare them to the data

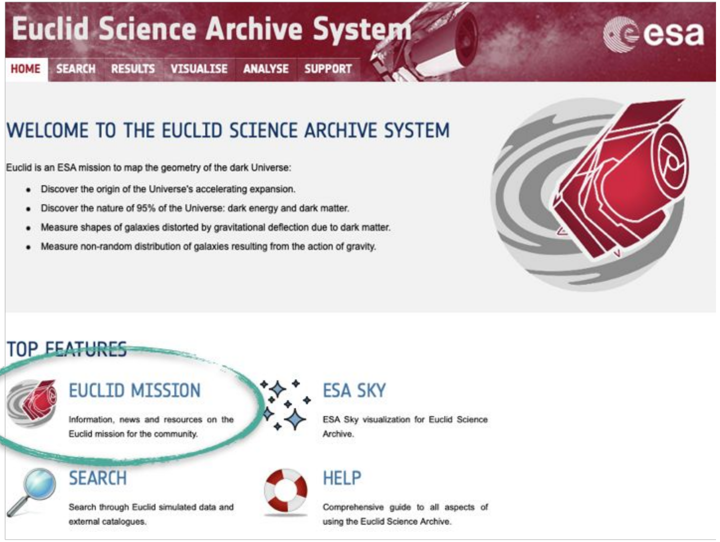


The Euclid Consortium & ESA

“Single official scientific consortium having the responsibility of the scientific instruments, the production of the data and of leading the scientific exploitation of the mission until completion”

CLOE is developed and tested by the Euclid Consortium

CLOE is an official product of the Euclid mission



The goal: analysis software pipeline

Scope: develop the Likelihood and Euclid software that allows to compute theoretical predictions for the primary probes and sample of cosmological parameters

Cosmological **L**ikelihood and **O**bservables in **E**uclid (**CLOE**)

IST:Non-Linear also develops CLOE by integrating the non-linear corrections of the observables

InterScience Taskforces: Likelihood & Non-Linear

A rare picture of (some) members of the two ISTs in the same room (Apr 2023)

InterScience Taskforce: Likelihood

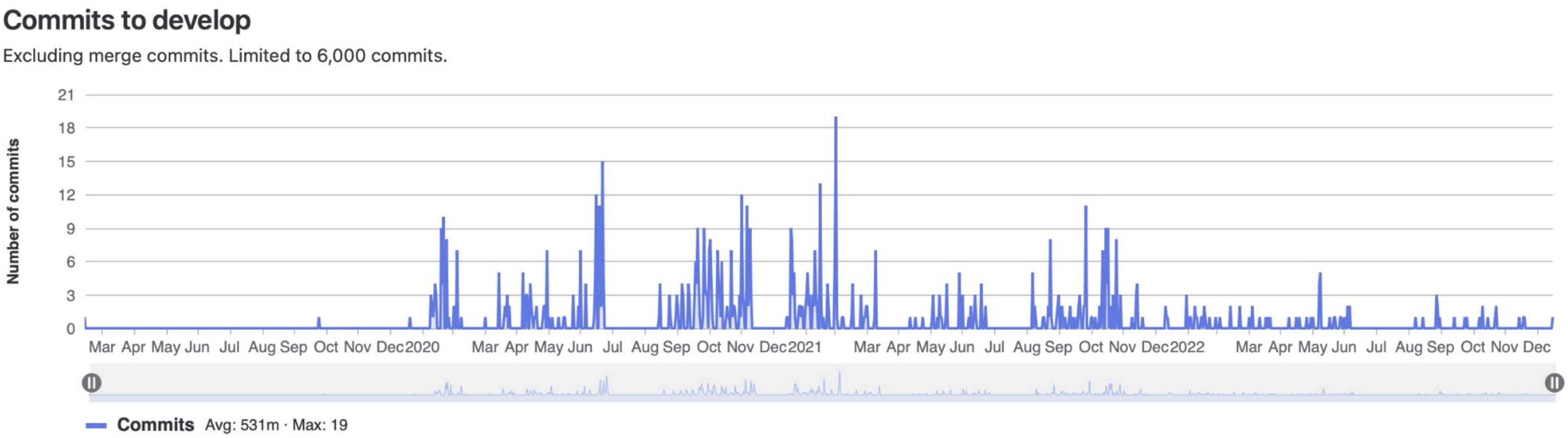
IST:L CLOE meeting in Leiden (Sep 2022)

- Software development following Agile practices
- Training for code development in python
- Implementation of unit tests, CI/CD, git version control system
- Not professional computer scientist… but lots of experience in cosmological data science
- We are aiming for Open Science and new data tools!

CLOE: features in a nutshell

- Fully written in python™ with internal calls to fortran/C
- Easy installation in conda environments ANACONDA
- Code development in GitLab
- Modular, easily extensible, prepared to adapt to future code

CLOE: writing code collaboratively



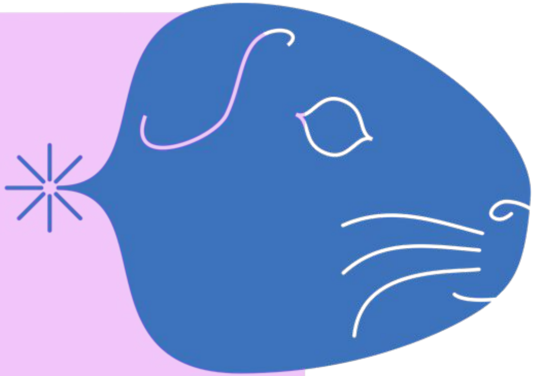
CLOE: structure

Retrieves **background quantities** from a Boltzmann Solver (i.e: CAMB/CLASS)

Computes **Theory predictions** for Euclid primary probes

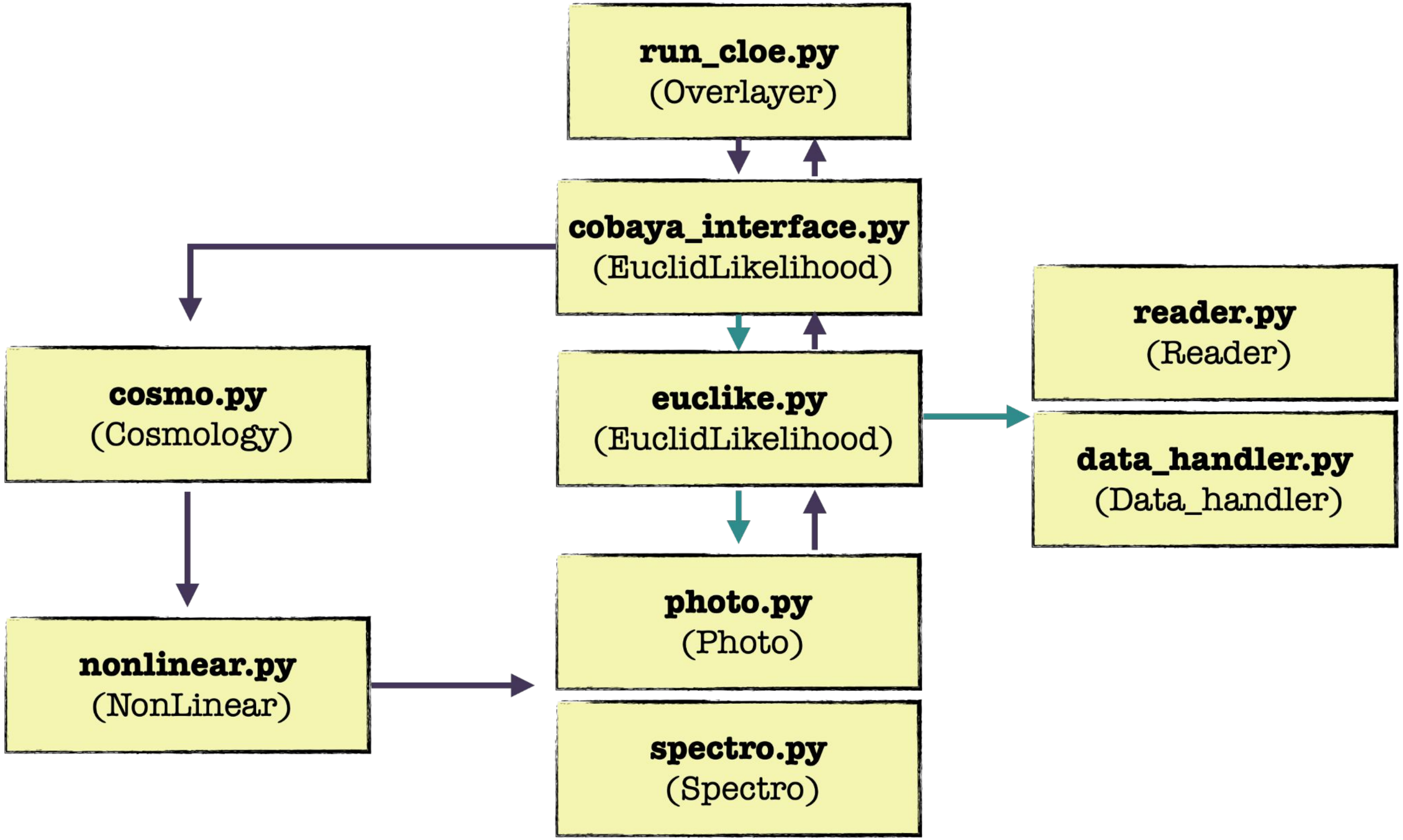
Reads **data** and other input information coming from the Science Ground Segment

Calculates the **Likelihood**

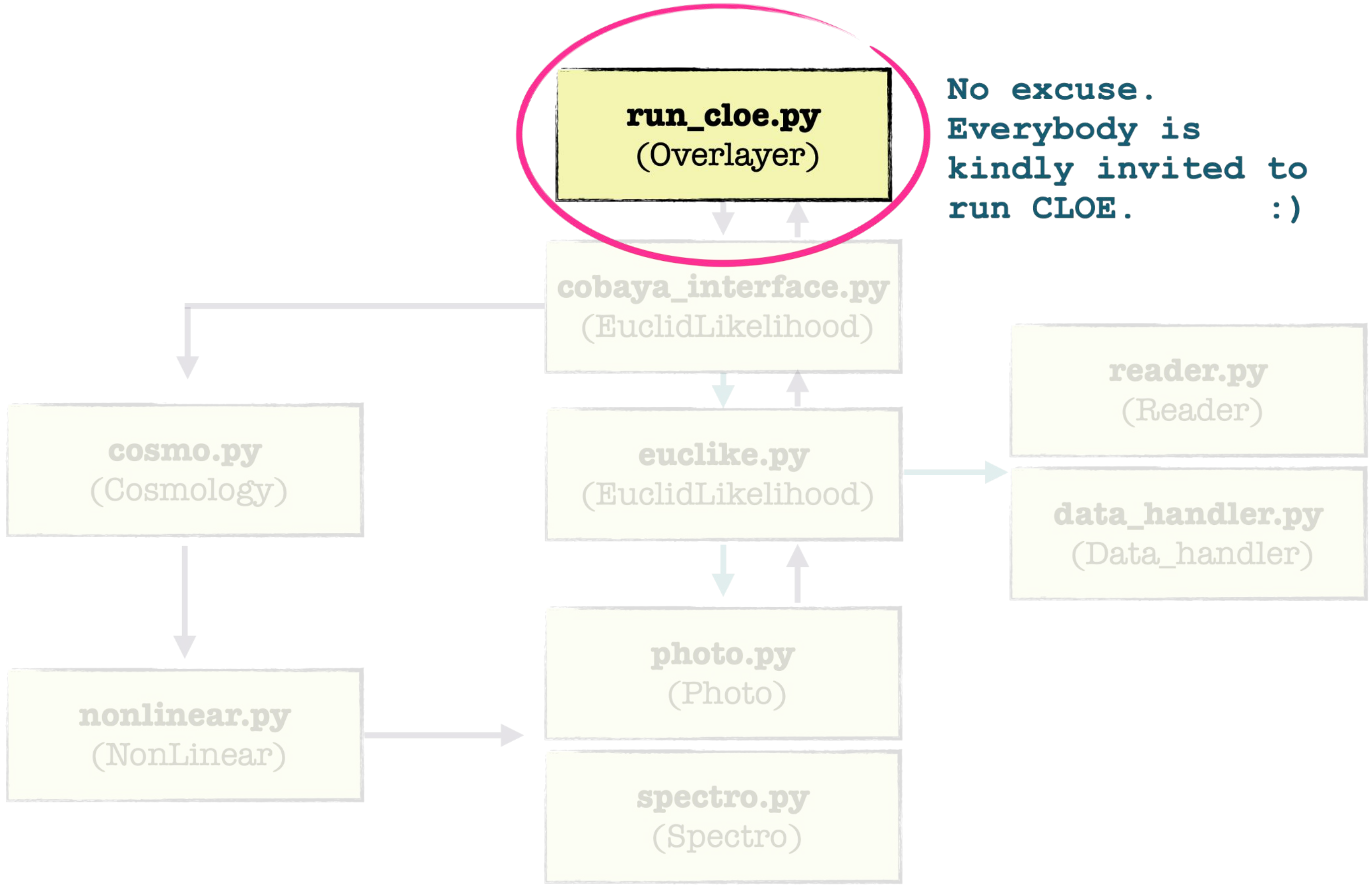


Sampling and processing

CLOE: structure

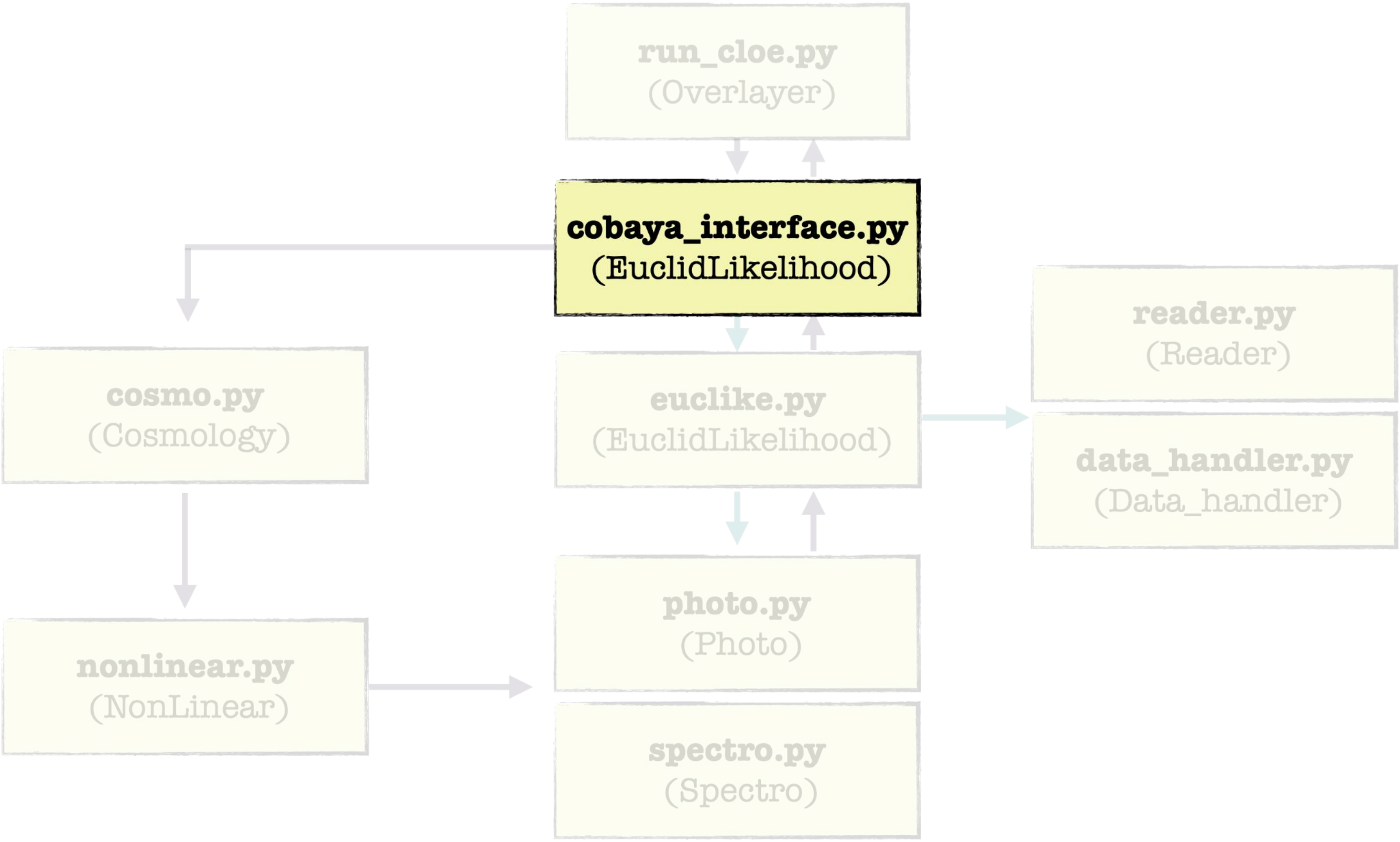


CLOE: structure

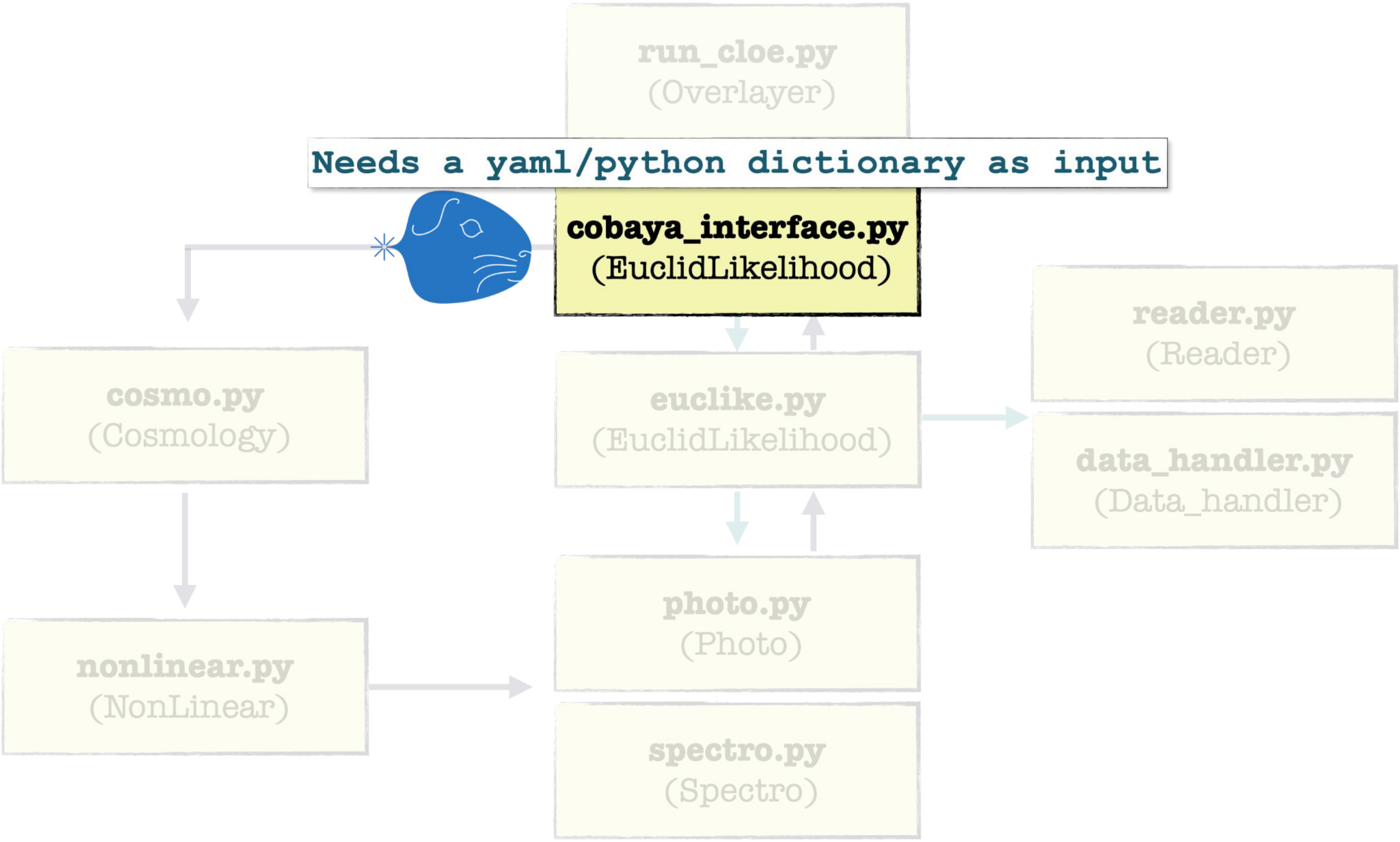


No excuse.
Everybody is
kindly invited to
run CLOE. :)

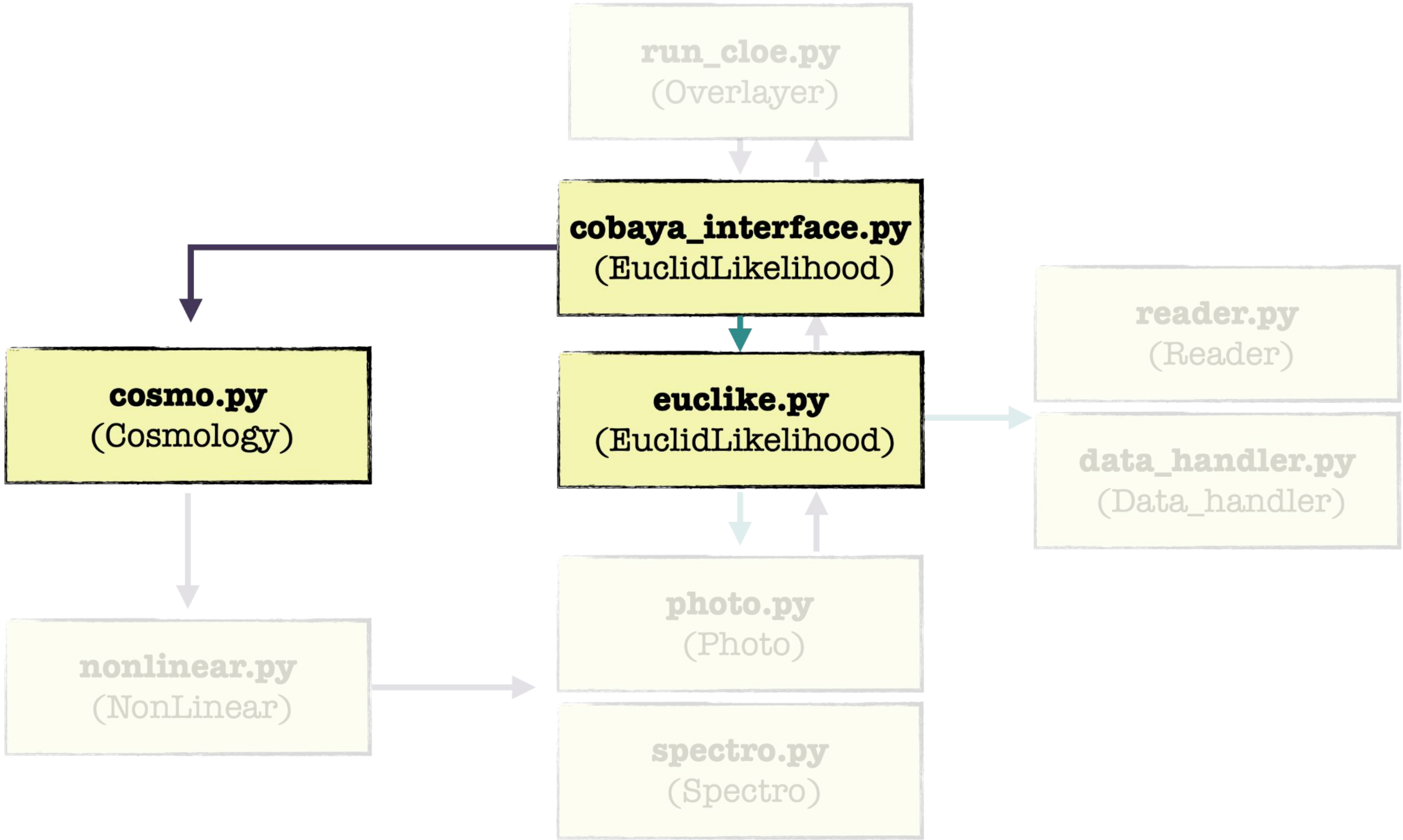
CLOE: structure



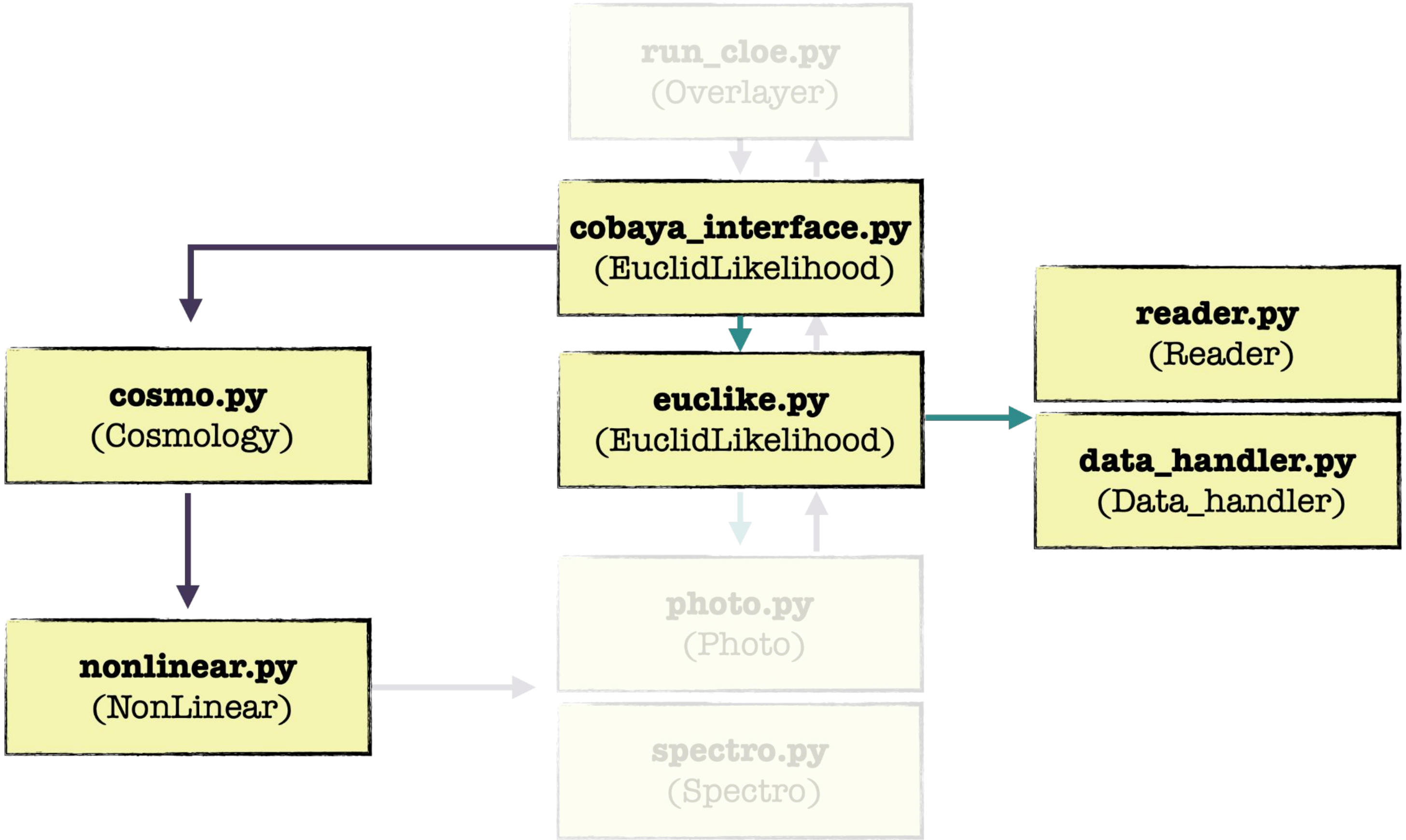
CLOE: structure



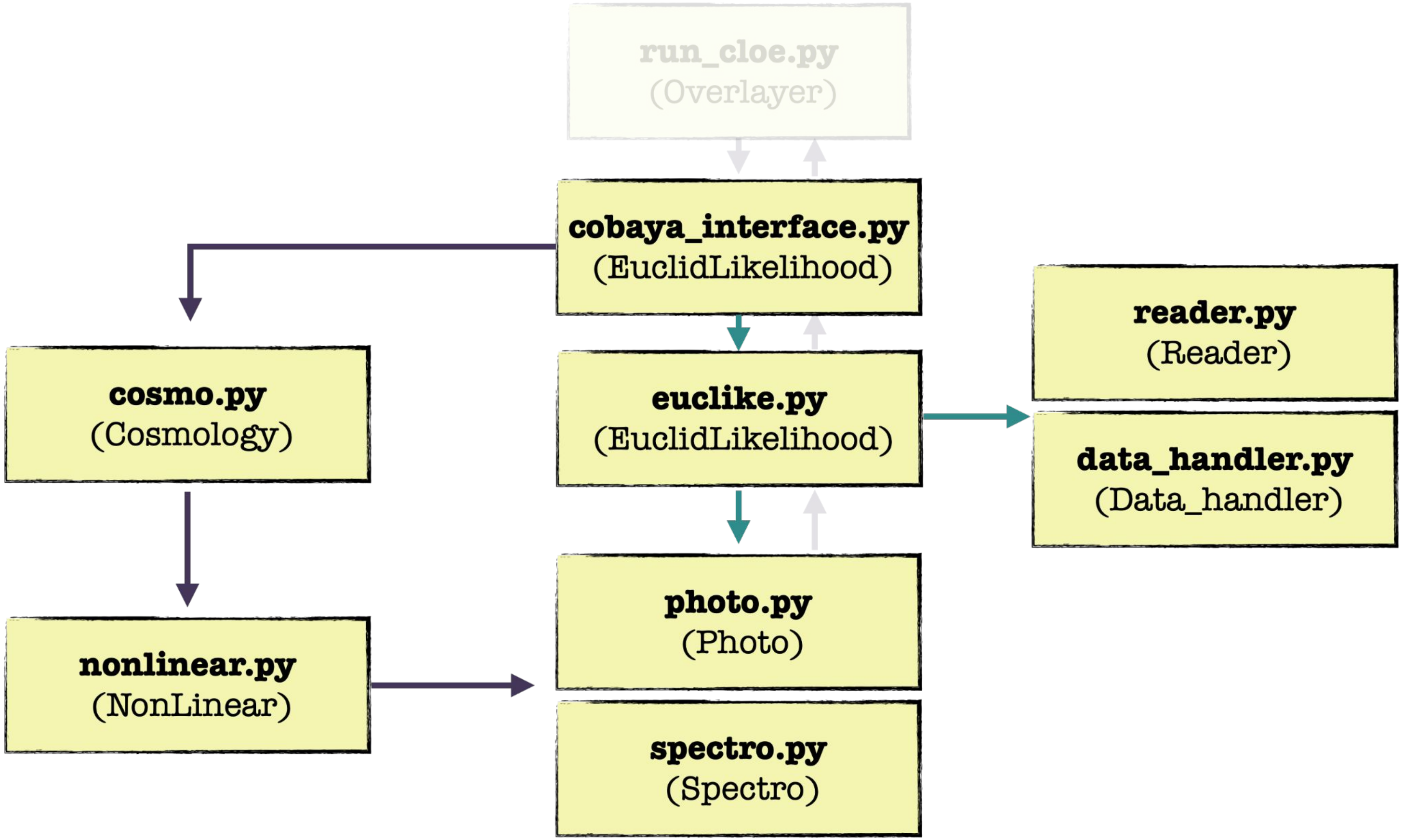
CLOE: structure



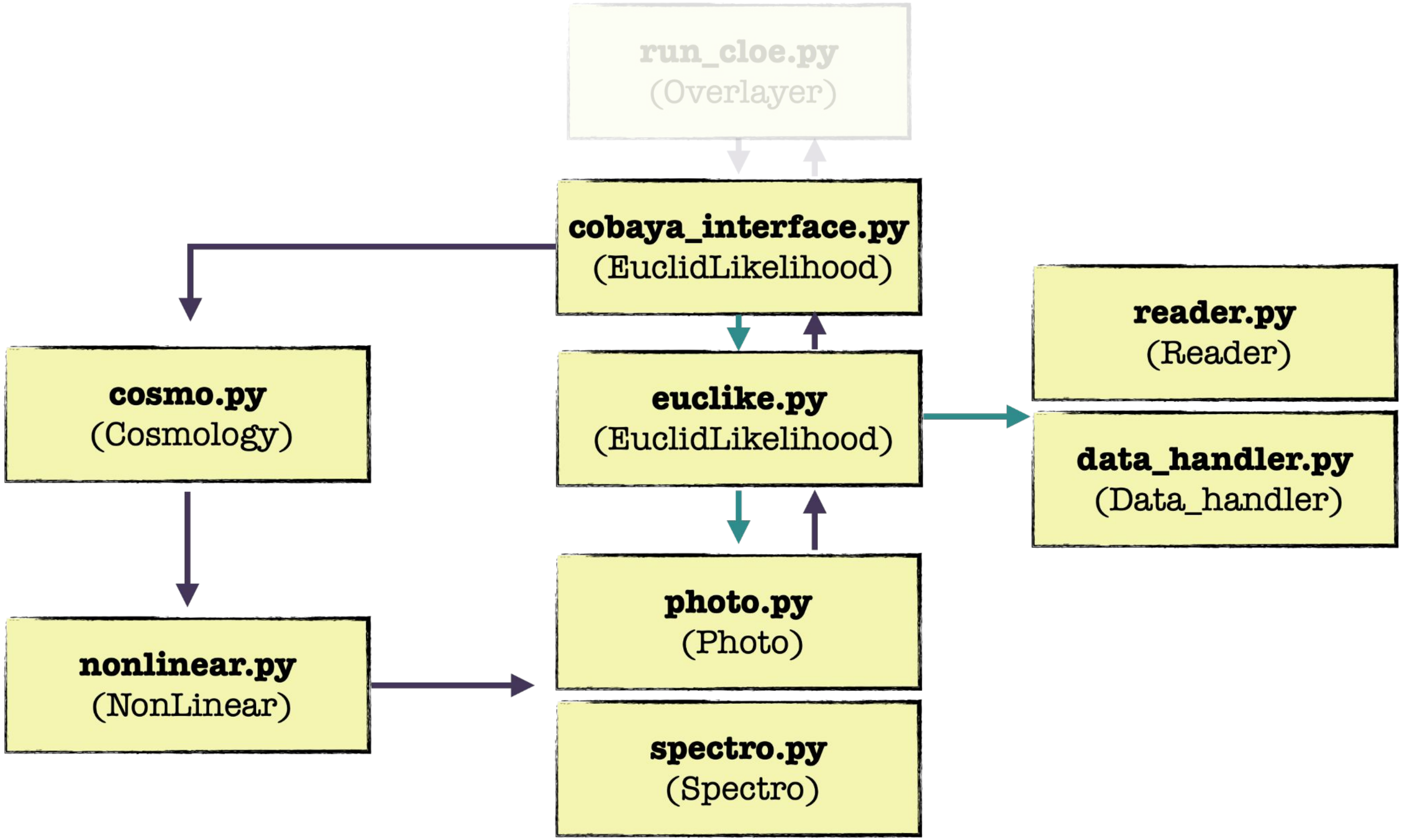
CLOE: structure



CLOE: structure



CLOE: structure



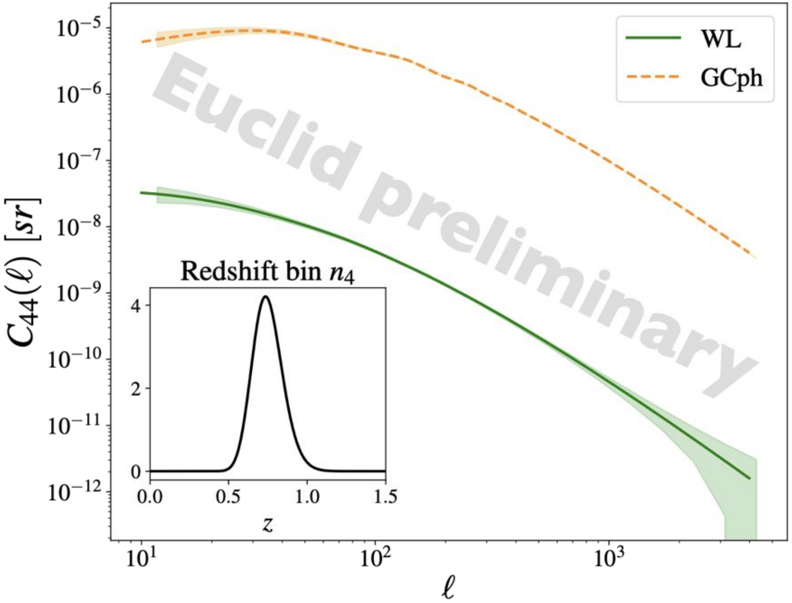
Euclid is designed to test beyond the Standard Cosmological Model

- **How do we model non-linear scales?**
- Systematics of the observations (shear, photometric redshifts purity, completeness…)
- **Computational expensive calculations**
- How do we model the observables when we take non-standard cosmologies into account?

Challenge 1: theoretical predictions

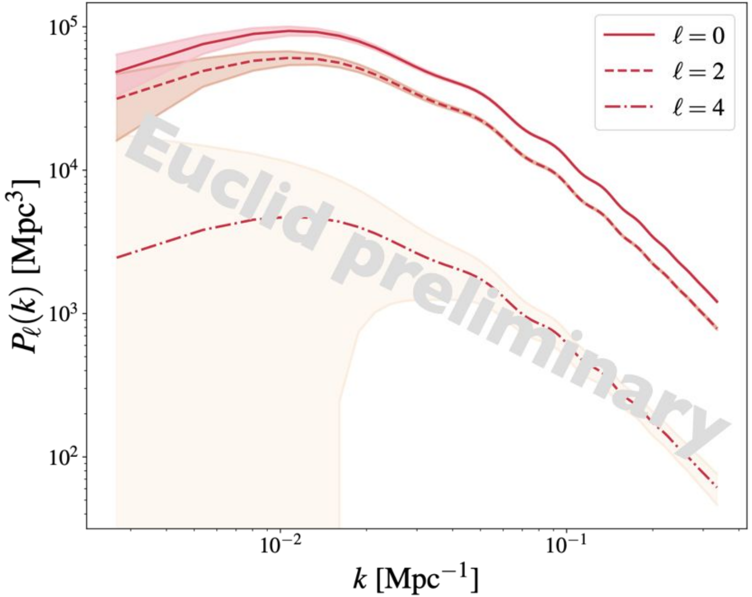
Photometric catalogue

$$C_{ij}^{AB}(\ell) \propto \int W_i(z)^A W_j(z)^B P_{\delta\delta} dz$$



Spectroscopic catalogue

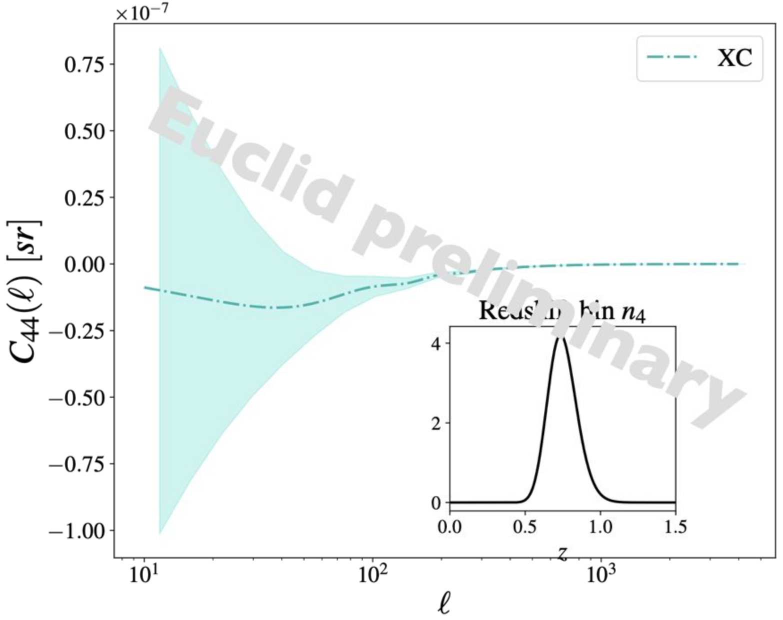
$$P_\ell(k,z) \propto \int_{-1}^{1} d\mu_k\, L_\ell(\mu_k) P_{\delta\delta}$$



Challenge 1: theoretical predictions

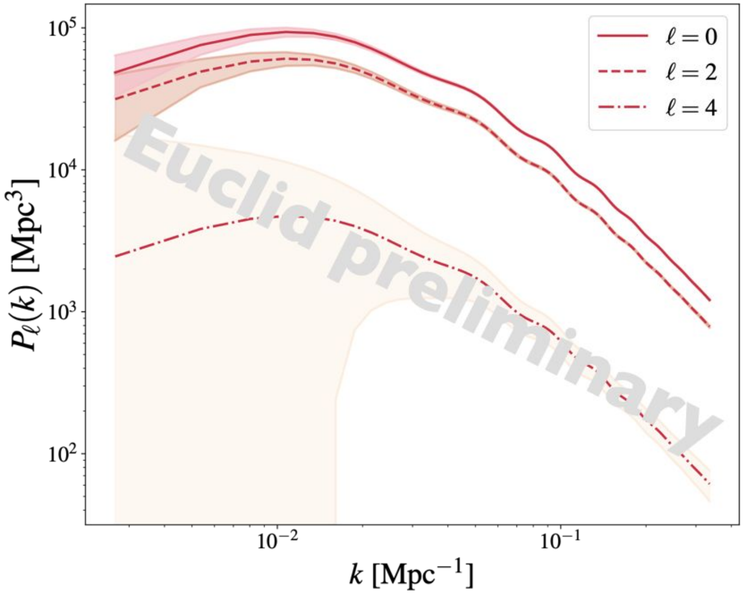
Photometric catalogue

$$C_{ij}^{AB}(\ell) \propto \int W_i(z)^A W_j(z)^B P_{\delta\delta} dz$$



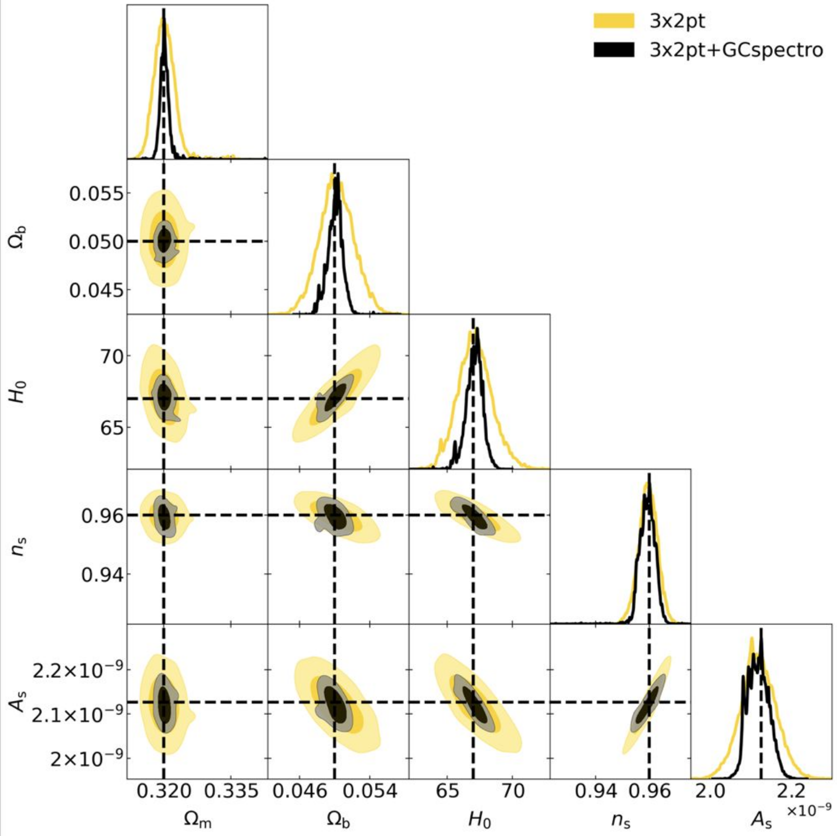
Spectroscopic catalogue

$$P_\ell(k,z) \propto \int_{-1}^{1} d\mu_k \, L_\ell(\mu_k) P_{\delta\delta}$$



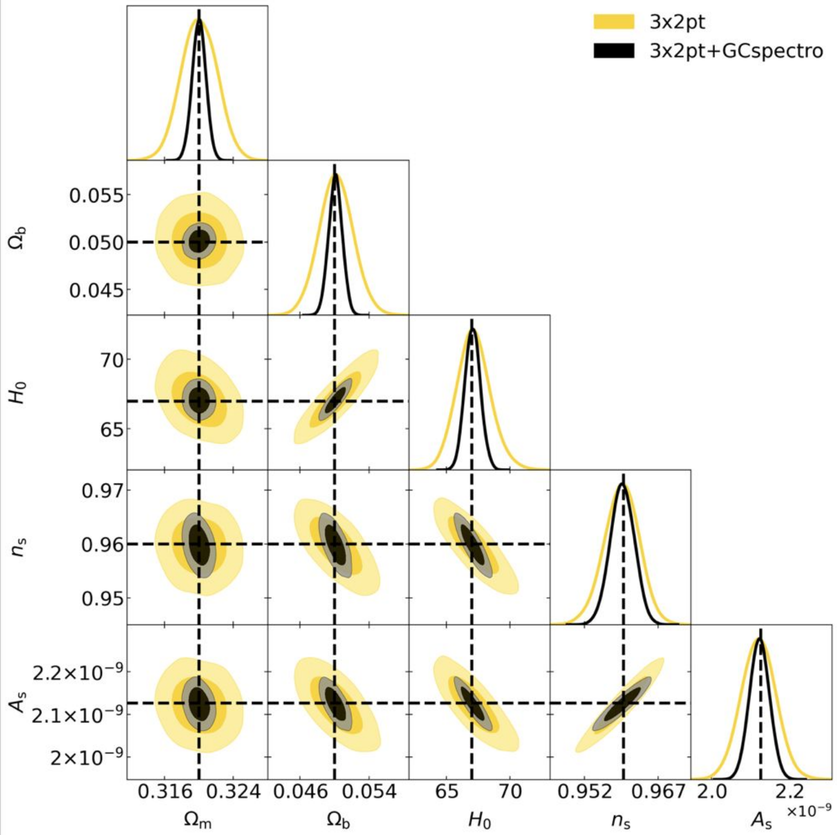
Challenge 2: sampling the posterior

- MCMC approach to sample the posterior distribution
- Convergence is time-consuming, sometimes impossible
- Scientific question: reinventing the wheel?



Challenge 2: sampling the posterior

- About one week to converge (even if it is a very optimistic scenario)
- Forecasting is essential to study how well we can obtain cosmological constraints in the future



Some (computing) details

- **Which parameters?**
  - cosmological + nuisance parameters
- **And… how many?**
  - +30 parameters (and increasing…)
- **And… how are their posteriors?**
  - Gaussian (because we are only playing with ΛCDM so far)
- **Model and data “interaction”**
  - angular power spectra for photo (l) + Legendre multipoles (k)
- **Time**
  - Each call to compute the predictions is around O(1) seconds
- **Bottleneck**
  - Computation of non-linearities/computation of the observables

Sampling issues

Is everything always smooth and Gaussian?

Extending the standard cosmological model

When testing the standard ΛCDM cosmological model one parameterize deviations from it

$$\begin{aligned} w_{\mathrm{DE}} &= -1 \\ k^2\Psi &= -4\pi G a^2 \rho\Delta \\ k^2\left[\Phi+\Psi\right] &= -8\pi G a^2 \rho\Delta \\ \frac{\Phi}{\Psi} &= 1 \end{aligned}$$

$$\begin{aligned} w_{\mathrm{DE}} &= w(a) \\ k^2\Psi &= -4\pi G\mu(a,k) a^2 \rho\Delta \\ k^2\left[\Phi+\Psi\right] &= -8\pi G\Sigma(a,k) a^2 \rho\Delta \\ \frac{\Phi}{\Psi} &= \eta(a,k) \end{aligned}$$

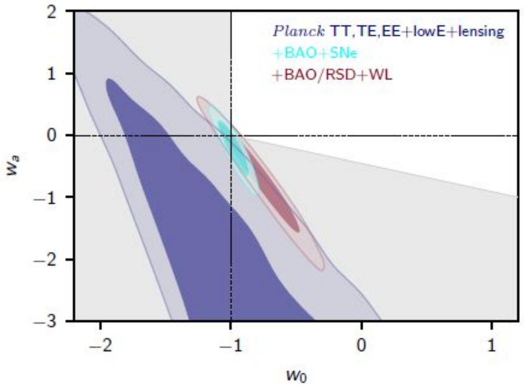
Can be non-parametric
(e.g. binned) functions
=> many additional d.o.f!

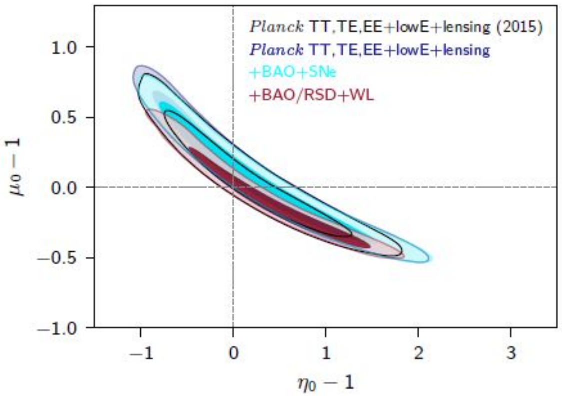
This is a common phenomenological approach, used for many extensions of the standard model, such as Dark Energy and Modified Gravity

$$w_{\mathrm{DE}}(a) = w_0 + w_a(1-a)$$

$$\begin{aligned} \mu(a,k) &= 1 + E_{11}\Omega_{DE}(a) \\ \eta(a,k) &= 1 + E_{22}\Omega_{DE}(a) \end{aligned}$$

Constraints on non-standard models





Using existing data, it was possible to constrain deviations from the standard model.

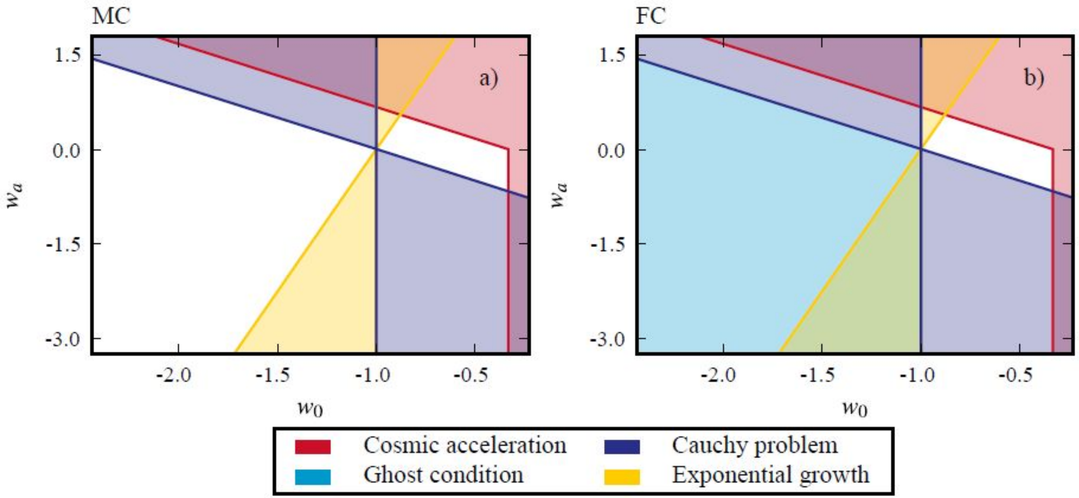
Planck satellite made the first comprehensive constraints on this, but many more followed from LSS surveys (KiDS, DES)

Planck 2018 results VI, cosmological parameters arXiv: 1807.06209

Where is the problem?

This phenomenological approach neglects the physical viability of such deviations.

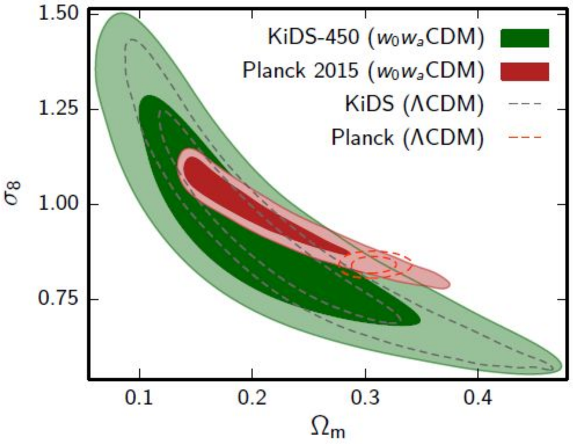
Imposing theoretical conditions removes areas of parameter space

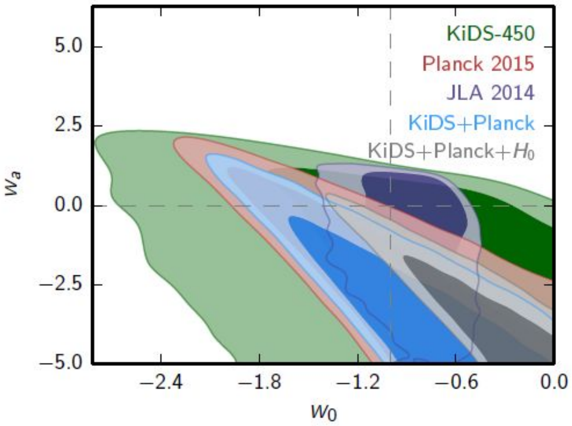


Peirone, MM, Raveri, Silvestri (2017) arXiv: 1702.06526

Usual methods can have issues

Such conditions can have significant impact on the results, but standard methods such as Metropolis-Hastings can suffer.

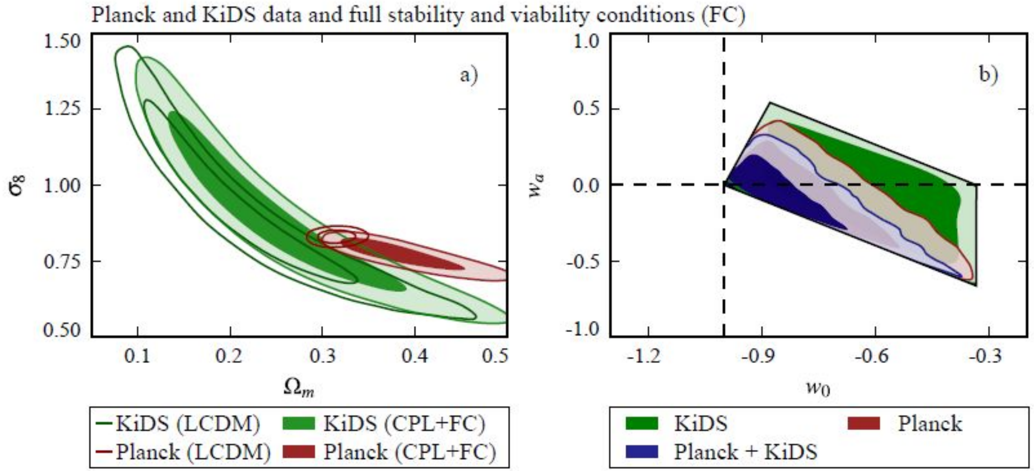




Joudaki et al. (2016) arXiv: 1610.04606

Usual methods can have issues

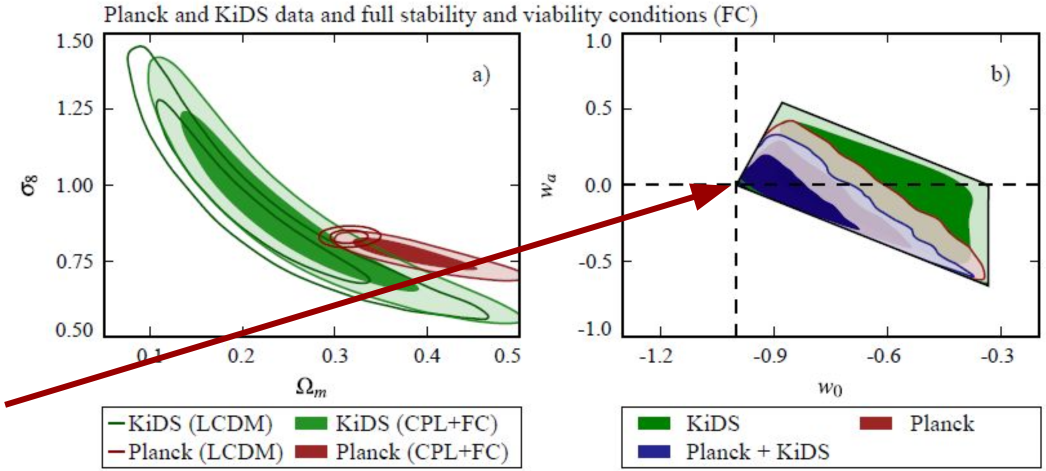
Such conditions can have significant impact on the results, but standard methods such as Metropolis-Hastings can suffer.



[Peirone, MM, Raveri, Silvestri (2017) arXiv: 1702.06526](#)

Usual methods can have issues

Such conditions can have significant impact on the results, but standard methods such as Metropolis-Hastings can suffer.



Standard model is at the edge of the allowed space. Difficult to sample!

Peirone, MM, Raveri, Silvestri (2017) arXiv: 1702.06526

Investigating extended models we face non-trivial volumes of the parameter space.

The standard cosmological model often lies at the edge of such volumes.

- Metropolis-Hastings struggles to efficiently sample the posteriors;
- Nested sampling can help, but we have a very high number of parameters
- Other methods might be unfeasible, e.g. HMC

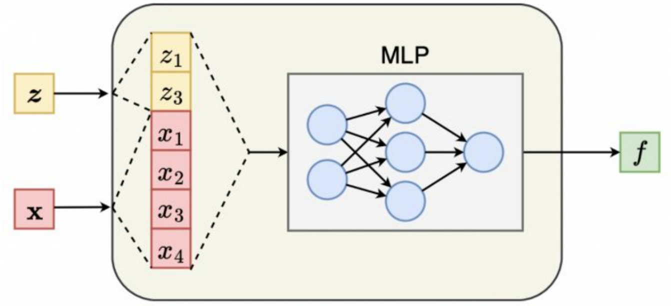
SWYFT can help!

Eventually include SWYFT within the Euclid pipeline (aka: CLOE)?

How and why Swyft can help:

- Already proved its potential in a (fairly) similar context: CMB analysis
- Next-gen surveys -> (many) more parameters (exotic models and/or nuisance)
- We are (almost always) only interested in 1D & 2D posteriors
- At first glance, a simulator should be straightforward to devise
- SBI could allow us to go beyond our standard statistics

However:



- We expect N_data / N_parameters >> 1

  -> Data summaries may be unavoidable

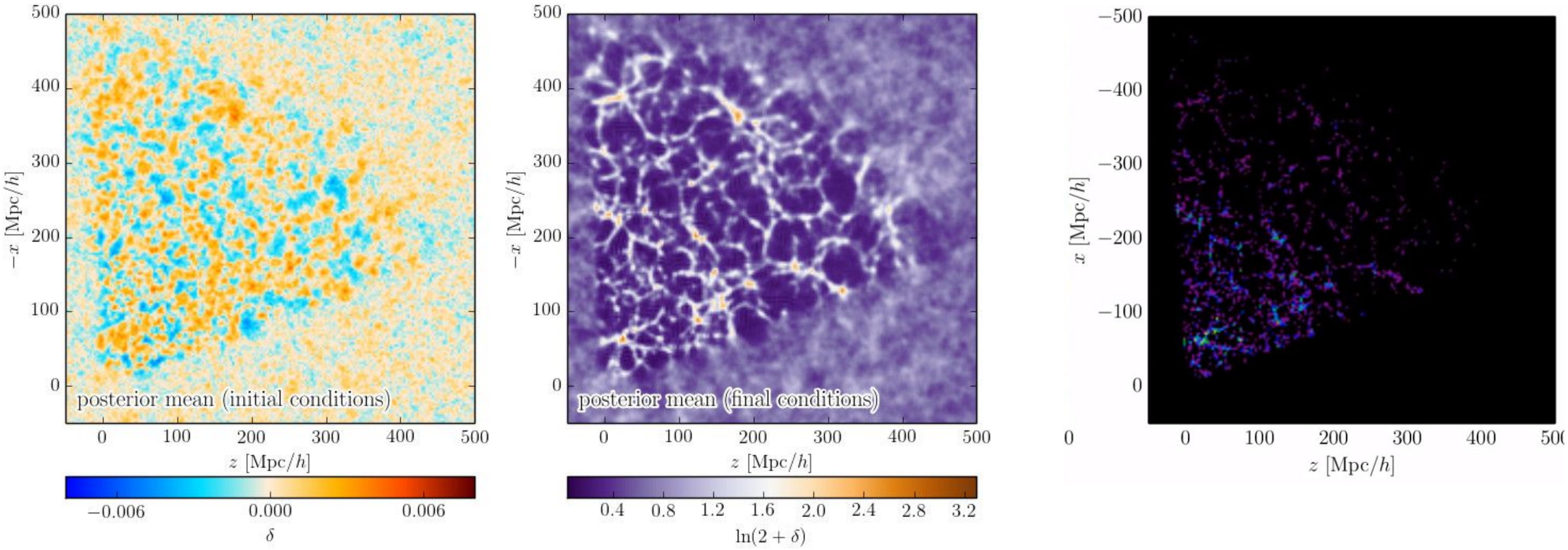
- Our observables (C_ells) are already summary statistics

  -> any addition compression is necessary lossy (how much?)

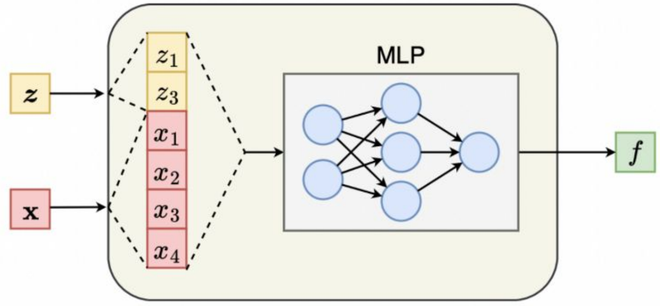
- Ability to deal with “big” and “twisted” parameter spaces? “Corners”?
- Rigorous simulator may not be that easy (non-linearities, higher-order stats)

BORG (Bayesian Origin Reconstruction from Galaxies)

(Jasche & Wandelt 2013, Jasche, Leclercq & Wandelt 2015, Lavaux & Jasche 2016)



However:



- We expect N_data / N_parameters >> 1

  -> Data summaries may be unavoidable

- Our observables (C_ells) are already summary statistics

  -> any addition compression is necessary lossy (how much?)

- Ability to deal with “big” and “twisted” parameter spaces? Corners?

- Rigorous simulator may not be that easy (non-linearities, higher-order stats)

- Expected reluctance from the community regarding such new tool

We need to test SWYFT in LSS forecasts

OTHER slides

An update on CLOE status

20/12/2022

Cosmological Likelihood for Observables in Euclid

What does CLOE do?

- CLOE gets cosmological quantities (H(z), P(k,z)...) from Boltzmann solvers based on CAMB or CLASS;
- It computes the theoretical predictions for Euclid observables using these quantities (and some more computed internally);
- It compares these with data and computes the likelihood;
- It produces MCMC thanks to its interface with samplers.

Who is working on CLOE?

The construction of the code architecture and the implementation of the observables recipes are responsibility of IST:L

IST:L works with the ESA science goals as the main objective.
Further extensions of the recipe are not within the IST:L mandate.

IST:L leads:
Vincenzo Cardone, Shahab Joudaki, Valeria Pettorino

IST:L core team:
Virginia Ajani, Linda Blot, Marco Bonici, Stefano Camera, Guadalupe Cañas-Herrera, Santiago Casas, Stefano Davini, Sergio Di Domizio, Sam Farrens, Stephane Ilic, Amandine Le Brun, Matteo Martinelli, Andrea Pezzotta, Ziad Sakr, Domenico Sapone, Davide Sciotti, Konstantinos Tanidis, Isaac Tutusaus

IST:NL and CLOE

In addition to what is provided by IST:L, CLOE relies also on the contribution of IST:NL.

- construction of the module computing nonlinear corrections necessary for theoretical predictions of Euclid's observables
- produce covariance matrices for 3x2pt observables

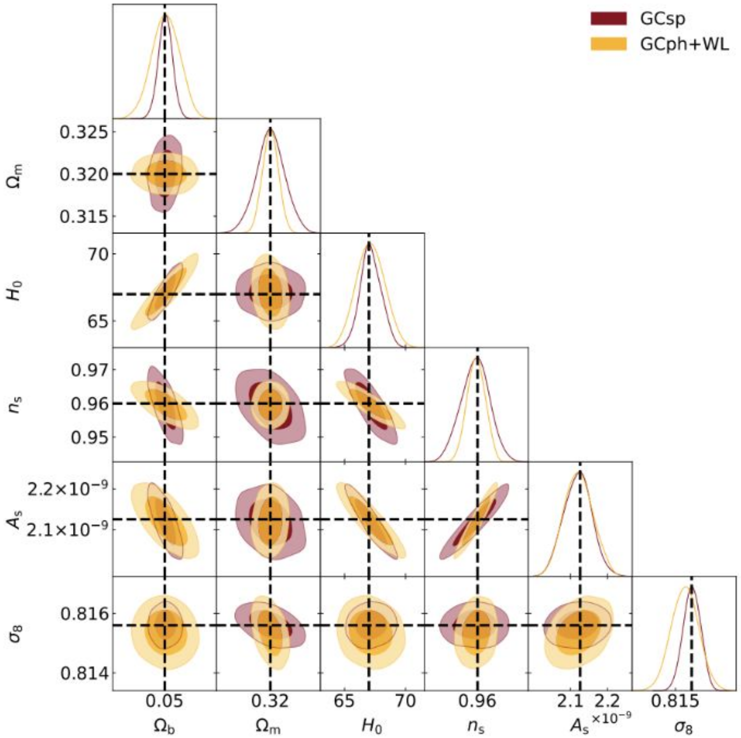
A rare picture of (some) members of the two ISTs in the same room

CLOE v1

Released to the SWGs in March 2021

- Only flat models
- Nuisance effects included:
    - intrinsic alignment
    - linear galaxy bias
- Linear power spectra
- Angular power spectra for photometric
- Legendre multipoles for spectroscopic

CLOE v1 works and it’s validated, but it could not be used for meaningful results due to its limitations.



What will be new in v2?

IST:L meeting in Leiden (09/22)

CLOE v2 is now close to completion and will be released “soon”

- Improved recipes for theoretical predictions
  - nonlinear corrections
  - non-flat models
  - redshift space distortions in GCph
- Several systematic effects are now included
  - magnification bias
  - sample purity
  - shear calibration
  - shift in mean bin redshift
- New observables
  - Pseudo-Cl
  - BNT transformation
  - 2pt correlations in configuration space
- Non-Gaussian likelihood

Models	GR	• flat and non flat: implemented at background level; non flat projection to be added) • massive neutrinos: scale independent growth factor
	MG	• γ - Linder parameterisation: modified growth factor and growth rate • (μ, Σ) - default to 1 but can be read as external file or imported from Boltzmann solvers
3x2pt	probes	• angular power spectra: tomography with different samples for lenses and sources • 2pt correlation functions: same as above • Pseudo Cl: mask mixing matrix imported as external input
	options	• scale cuts: selection of the multipole range and redshift bins • BNT: data vector and covariance matrix transformed using internally computed BNT matrix
GC spectro	probes	• Legendre multipoles • 2pt correlation function
	options	• convolution with window function: window function provided as external input
Likelihood	options	• Gaussian: covariance matrix including SSC provided as external input • non Gaussian: implementation in accordance with Percival et al. 2022
User choices	Boltzmann solver	• CAMB • CLASS • any MG variant with python wrapper (e.g., MGCLASS, HiCLASS, EFTCAMB)
User choices	miscellaneous	• selecting individual or combined probes • running from command line • configuration files for popular models • GUI to set configuration files

Can we do cosmology with v2?

With the implemented modifications we can start obtaining meaningful results

- **flexibility with probes**: 3x2pt and GCspectro (in combination or standalone)
- **extended models**: ΛCDM, CPL dark energy, MG (only γ parameterization), curvature
- **systematic effects**: modelling of several possible effects for marginalization
- **nonlinear corrections**: emulators and PT

CLOE supports (through Cobaya) also the use of modified Boltzmann solvers. If one has a CAMB/CLASS version for any model this can be interfaced easily.

Warning:

- Range of validity of nonlinear corrections (IST:NL)
- Recipes provided by SWG contain ΛCDM assumptions (IST:L recipe Key Project)
- Extensions of ΛCDM are the target of TWG work

What is left to do for v2?

v2

PF-IST-Likelihood / Likelihood-implementation

233 Issues · 51 Merge requests   78% complete

IST:L still has some open tasks, but we are wrapping up.

Most of the task remaining aim at optimizing the code and improving the documentation.

Some parts of the recipe are still being finalized:

- BNT and pseudo-Cls;
- non-Gaussian likelihood;
- easy switch between CAMB and CLASS;
- Cls calculation without Limber approximation (might move to v2.1);

Originally requested recipe is already implemented. IST:L working on recently requested improvements.

Further requirements for v2 completion

IST:L is currently defining the procedure for validation and for the work on KPs.
However, there is no defined timeline yet.

Other than the leftover implementation from IST:L, the products of IST:NL still need to be integrated in the main CLOE code.

After successfully merging the new nonlinear module in CLOE, the code will support emulators to be used for 3x2pt and 1-loop PT for GCspectro, with new parameters included

This step is necessary to start the validation of v2, which will lead to the release.

There is also the possibility of further requests from the SWGs, which could cause additional delay.

A look forward to v3

v3

PF-IST-Likelihood / Likelihood-implementation

17 Issues · 0 Merge requests 0% complete

Our work is not done! Significant work is expected for v3

- **Interface with CosmoSIS**
  Possibility to sample parameters from both Cobaya and CosmoSIS
- **Merging of SWGs contributions**
  new contributions to the likelihood (XCMB, Clusters, TWG, ...)
- **Further optimization**
  might include some restructuring of the code. Automatic number of parameters.
- **Changes in recipe**
  Switch off Limber (if not in v2), further changes from SWGs
- **Extend systematic effects**
  beyond linear bias and eNLA intrinsic alignment.
- **Prepare for data**
  improve interface with OU-LE3 products

CLOE nonlinear – LCDM

WL + photo probes:

- implemented halofit, HMcode, euclid emu2, bacco for nonlinear matter

Spectro probe:

- EFTofLSS model ported from PBJ (EdS approximation, WnW split for IR resummation, Eulerian bias expansion)

→ Emulators already merged to IST:L, EFT model under review

→ Running chains on synthetic data for validation and testing slides by Lucia

Beyond LCDM: nDGP in CLOE

- Overleaf Euclid KP-JC-6: Paper 3
- Linear growth function

$$\tilde{D}'' + \left(\frac{2}{a} + \frac{\mathcal{H}'}{\mathcal{H}}\right)\tilde{D}' - \mu_\Phi(a)\frac{3\Omega_m(a)}{2a^2}\tilde{D} = 0$$

- Modified Newton Constant

$$\mu_\phi(a) = 1 + \frac{1}{3\beta}$$

$$\beta = 1 + \frac{H(a)}{\sqrt{\Omega_{\rm rc}}}\left(1 + \frac{\dot{H}}{3}\right)$$

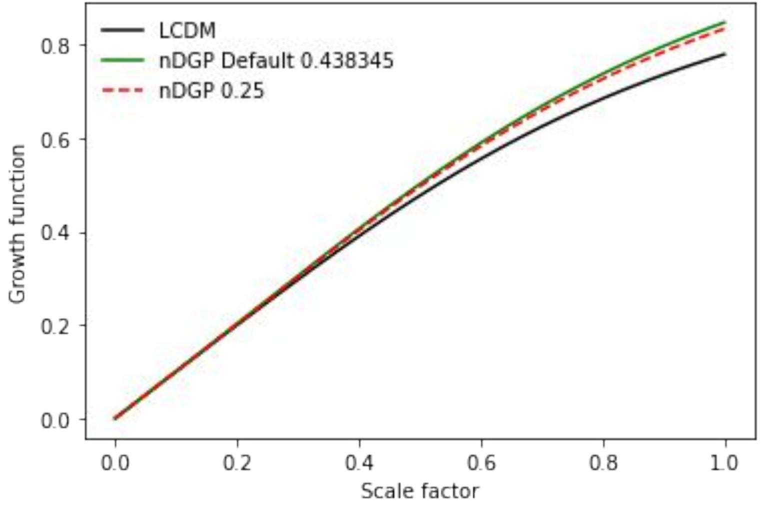
where $\cdot \equiv d/dt$, $\Omega_{\rm rc} \equiv 1/4H_0^2 r_c^2$ and r_c is the cross-over scale.

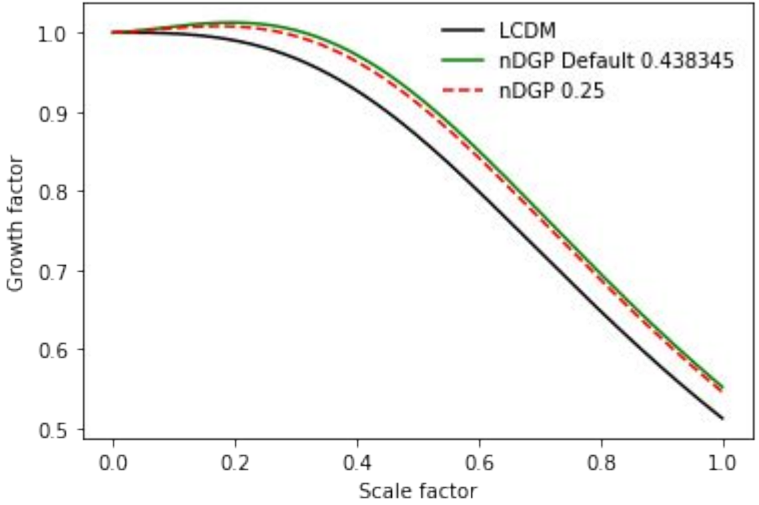
slides by Lucia

Code development

- Code + unit-tests + example GitLab IST-NL-implementation/develop_NL_eft

```
from modified_newton_constant_generator import *
from modified_growth_solver import ModifiedGrowth as MG
from modified_growth_equation import modified_linear_ode as MODE
```





- NOTE: growth should be normalised to LCDM at z=0
- Accepts other EFToDE models!

CLOE implementation

- P(k, z=0) from Boltzmann code
- Rescaling it to z_obs by D(z)/D(0)
- RSD requires f(z_obs)

EFTofLSS model implemented accepts f,D as input parameters for the rescaling → easily extendable

f(R) emulators implemented for WL/photo (used in SP-photo)

Working on gamma

Discussion:

- Which other models?
- Do we want to add a module that uses PyBird? (if EdS approx + rescaling is not good enough to recover parameters – output of KP-JC-6 p3)

Proyecto de muestra

Escribe aquí tu texto Escribe aquí tu texto.

Escribe aquí tu texto Escribe aquí tu texto Escribe aquí tu texto Escribe aquí tu texto.

Formación

Nombre del centro

Ciudad (Provincia)

MBA: mayo del 20XX

Nombre del centro

Ciudad (Provincia)

Grado en Informática: mayo del 20XX

Nombre del centro

Ciudad (Provincia)

Laboratorio AdTech: enero del 20XX

Contacto

Tu nombre

no_reply@example.com

www.example.com