Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

PBJ: preparing for the analysis of galaxy clustering data from Stage-IV surveys

Chiara Moretti, R. Trotta, M. Viel, E. Sefusatti, E. Sarpa, +

Spoke 3 General Meeting, Elba 5-9 / 05, 2024

SISSA DATASCIENCE Machine Learning for the Natural Sciences

SISSA

Scientific Rationale

Stage-IV spectroscopic galaxy surveys → precision cosmology by mapping Universe with billions objects

Dark energy? Dark matter? Neutrino mass?

Use **summary statistics** (two-point function, higher order) to extract information from galaxy distribution

→ need **fast** and **accurate** tools (theoretical modelling and Bayesian analysis)

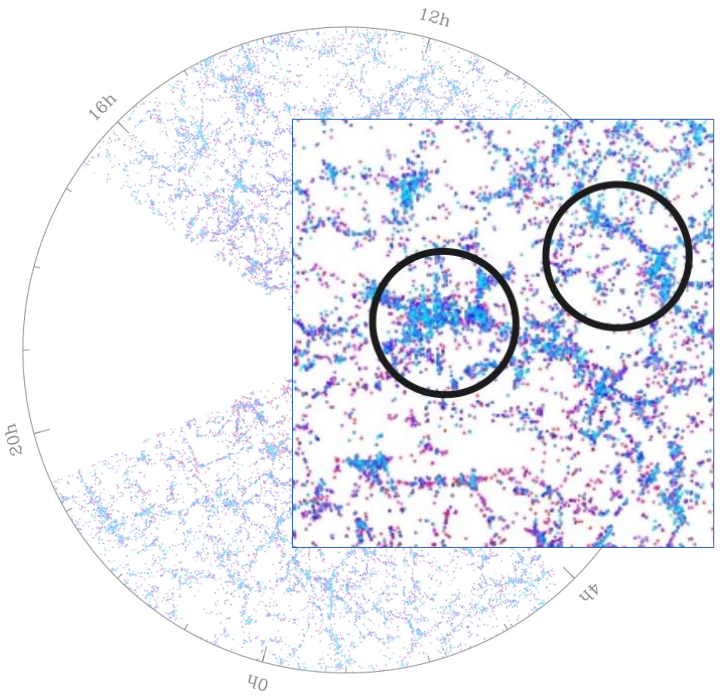
→ robust **validation** of the pipeline

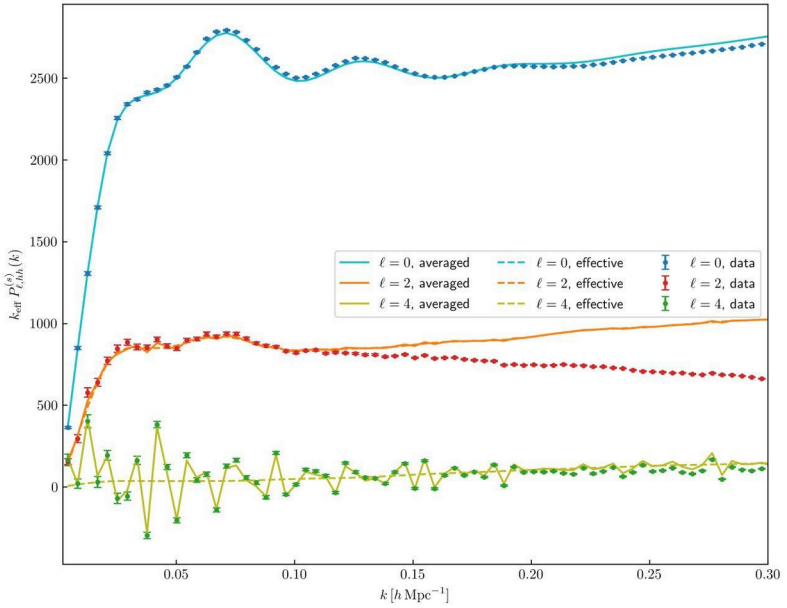
→ Modelling **systematics**

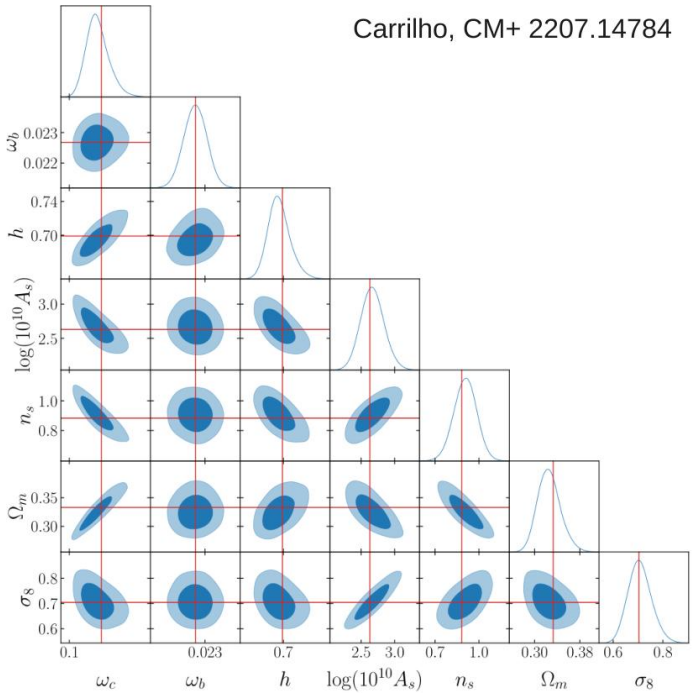
Technical Objectives, Methodologies and Solutions

Goal: produce a Bayesian analysis pipeline to extract cosmology from galaxy distribution

PBJ: power spectrum and bispectrum joint analysis







PBJ: the code

- Fully in python – still extremely fast! Pgg in ~0.04 s, Bggg in ~0.1 s
  - Convergence on Euclid-like datavector in O(10) cpu hours
- State of the art **nonlinear model** for power spectrum (EFT of Large Scale Structure) + nonlinear bias treatment → **ported to official Euclid likelihood**
- Loop corrections computed with **FastPT**
- Tree-level **bispectrum**
- Emulators for linear power spectrum (100x faster)
- Analytic marginalisation (10x faster)
- Several **samplers** (Metropolis-Hastings, affine invariant, nested sampling + machine learning powered)

PBJ: the code

- **Modular structure**
  - Theory module
  - Likelihood module
  - Binning
  - Main module/class

- **Code restructuring→ user-friendly**
  - Only needs two parameter files + minimal python script to run

```
import PBJ
from tools.param_handler import read_file

init_dict = read_file("paramfile.yaml")

pbj = PBJ.pbj(init_dict)
pbj.initialise_full()

pbj.run_sampler(NmaxP=33.5, nsteps=20000, nwalker=200)
```

Timescale, Milestones and KPIs

MILESTONE 6

Target: re-analysis of BOSS data for beyond-LCDM (γ + massive neutrinos), forecasts for Stage-IV surveys

KPI: paper submission (arxiv:2306.09275)

MILESTONE 7

Target: : analysis of Flagship simulation in real space

KPI: Euclid key paper submitted (arXiv:2312.00697)

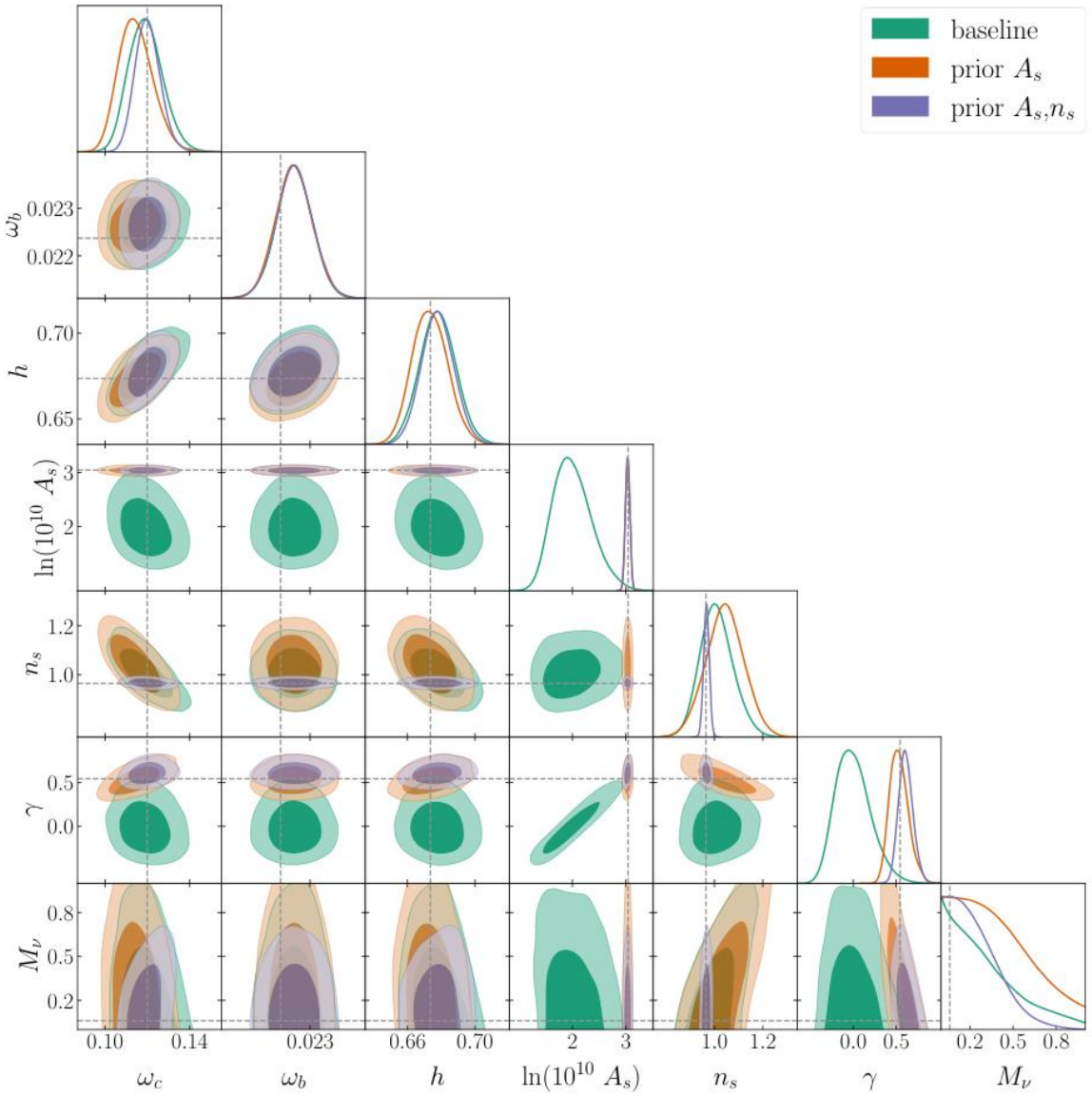
Target: modelling beyond-LCDM models

KPI: Euclid key paper submitted (arXiv:2311.13529)

Accomplished Work, Results

Re-analysis of BOSS data + Stage-IV forecasts

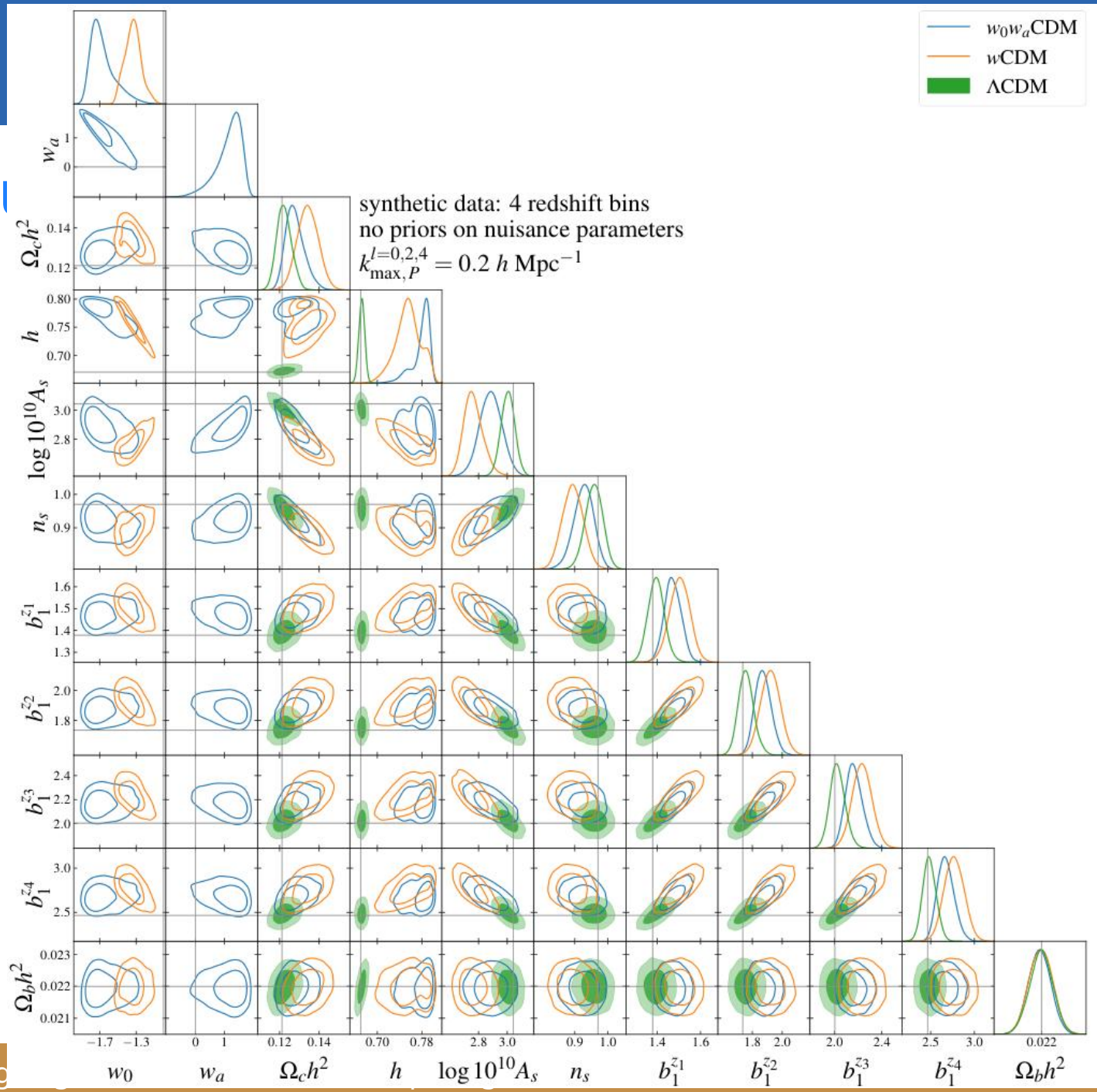
- Constraints on γ + neutrino mass from full-shape power spectrum analysis
- BOSS DR12 data
- Projection effects! Huge problem for beyond-LCDM models → profile likelihood to mitigate
- Forecasts for DESI-like data



Accomplished Work, Resu

Euclid preparation

- Analysis of Flagship simulation: power spectrum in real space (Pezzotta+23)
- Beyond-LCDM models: χ^2 analysis (Bose+23)
- **Both featured PBJ as one of the main codes used for the analyses!**
- Forecasts within Science Performance Verification 3 of Euclid



Next Steps and Expected Results

Work in progress:

- **Public code release**
- Inclusion of post-reconstruction **BAO** (see E. Sarpa's talk)
- **Window** convolution (with J. Salvalaggio)
- Treatment of **interlopers** (with M. Barberi Squarotti)
- **Euclid key papers:**

  → IST:Nonlinear, forecast for spectroscopic probe

  → Galaxy clustering SWG, P analysis in redshift space, P+B analysis in redshift space
- Analysis of simulations with **massive neutrinos** (with E. Bellini)