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

EuMocks: Mocking the Universe for Euclid++

Pierluigi Monaco, UniTS (+INAF, INFN, IFPU, ICSC),
M. Lepinzan, L. Tornatore, G. Taffoni, C. Carbone

Spoke 3 Technical Workshop, Trieste October 9 / 11, 2023

Scientific Rationale

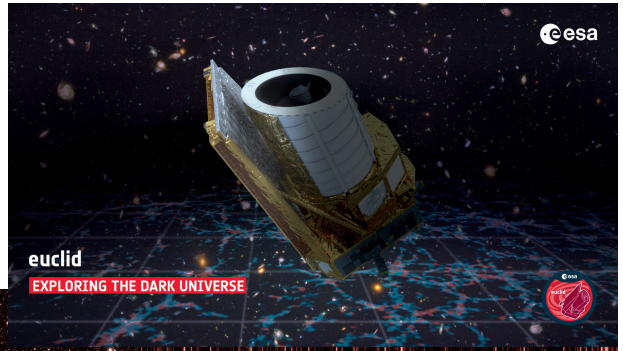
Euclid will survey the universe down to redshift z~2, **mapping the large-scale structure** to measure its geometry and growth rate to shed light on dark sector

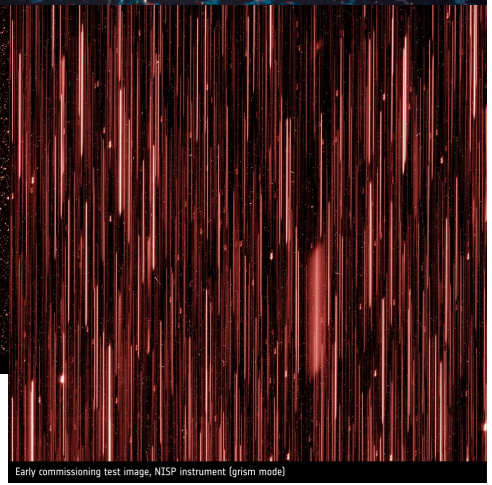
A spectroscopic sample will be based on **slitless spectroscopy** of ~15,000 sq deg of the sky, detecting the Halpha line at 0.9<z<1.8

Control of **systematic errors** will be the issue to tackle to provide convincing and potentially groundbreaking results.

To this aim we need **thousands of simulations** of the Universe observable by Euclid...

and by **SKAO**, DESI, LSST, Roman...



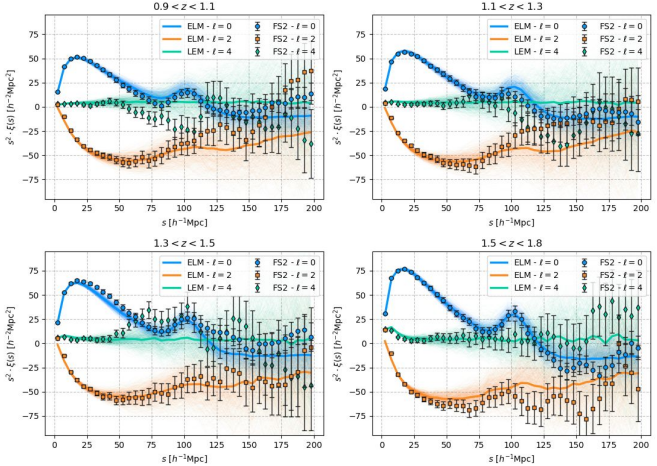


Euclid's first light images

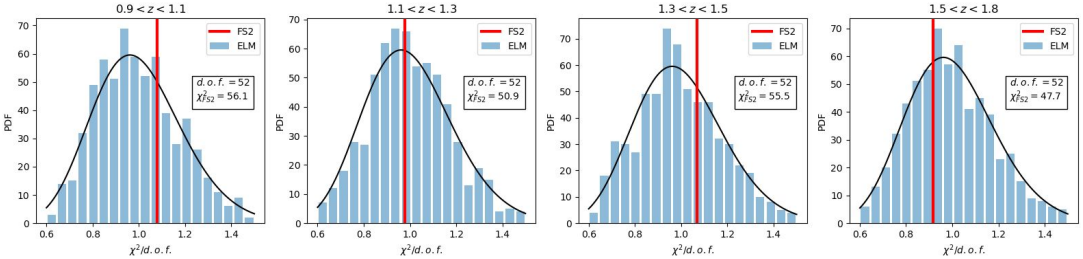
Technical Objectives, Methodologies and Solutions

We aim at producing **3500 simulations of a volume of ~4Gpc with ~10^{12} particles, resolving halos of ~10^{11} M_{sun}/h**, with output on a past-light-cone covering **half of the sky** and starting at **z=3**.

Standard N-body codes are too slow to produce such a massive set of simulations
Approximate methods: **PINOCCHIO**



Chi^2, Monopole + Quadrupole, 20 < r [Mpc/h] < 150



A. Veropalumbo, E. Sefusatti et al.

Technical Objectives, Methodologies and Solutions

We aim at producing **3500 simulations of a volume of ~4Gpc with ~10^{12} particles, resolving halos of ~10^{11} M_{sun}/h**, with output on a past-light-cone covering **half of the sky** and starting at **z=3**.

Standard N-body codes are too slow to produce such a massive set of simulations
Approximate methods: **PINOCCHIO**

Development is needed (Spoke3 + Spoke1, PRIN PNRR 2022):
SCIENCE:
- improve and optimize the **reconstruction of halos**, especially in filaments
- implement a **Particle-Mesh** code to move halos (instead of 3LPT)
- add **lensing** and relativistic effect

CODING:
- port on **GPUs**
- improve the **fragmentation (deblending, clustering)** algorithm

Technical Objectives, Methodologies and Solutions

We aim at producing **3500 simulations of a volume of ~4Gpc with ~10^{12} particles, resolving halos of ~10^{11} M_{sun}/h**, with output on a past-light-cone covering **half of the sky** and starting at **z=3**.

Standard N-body codes are too slow to produce such a massive set of simulations
Approximate methods: **PINOCCHIO**

Development is needed (Spoke3 + Spoke1, PRIN PNRR 2022):
SCIENCE:
- improve and optimize the **reconstruction of halos**, especially in filaments
- implement a **Particle-Mesh** code to move halos (instead of 3LPT)
- add **lensing** and relativistic effect

CODING:
- port on **GPUs**
- improve the **fragmentation (deblending, clustering)** algorithm

See talk and PhD project by Marius Lepinzan

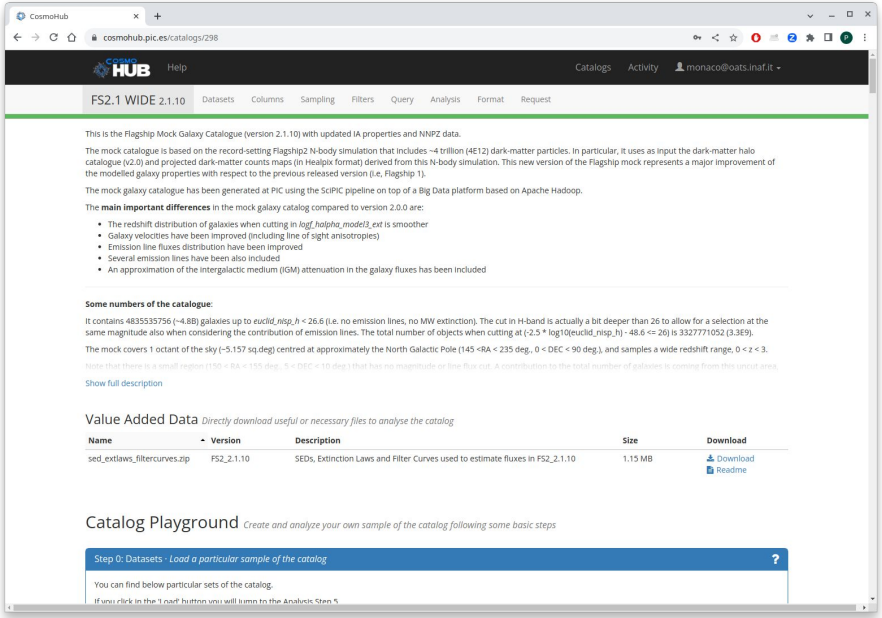
Challenges

Massively parallel code, every step must be optimized before burning so much computing time

Computing time: ~30,000,000 core hours
Memory: ~128 TB
Storage: >~1PB

Petabyte-scale output to be offered to the community
-> National / Interoperable Data Lake

example: Cosmohub.pic.es



Timescale, Milestones... and KPIs?

2023 | **2024** | **2025** | **2026**

launch July 2023

DR1 data Sept 2024

DR1 public Sept 2025

DR2 data Sept 2026

Milestone 6 | Milestone 7 | Milestone 8 | Milestone 9 | Milestone 10

ICSC year 1

ICSC year 2

ICSC **END**

ICSC resources available

code development.....

available simulations.....

Sept 2024: **available simulations** in the data lake?

March 2025 **code ready**

March 2025 data lake ready

August 2025 **first 100 sims**

August 2026 3500 sims

KPI **KPI** **KPI**

Next Steps and Expected Results (by next checkpoint: April 2024)

...well, shouldn't we first approve the project?