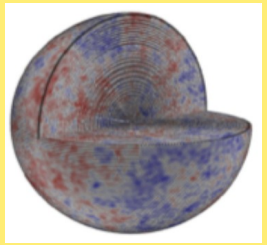
Cosmic scales: integrating Bayesian inference and deep learning to reveal new physics

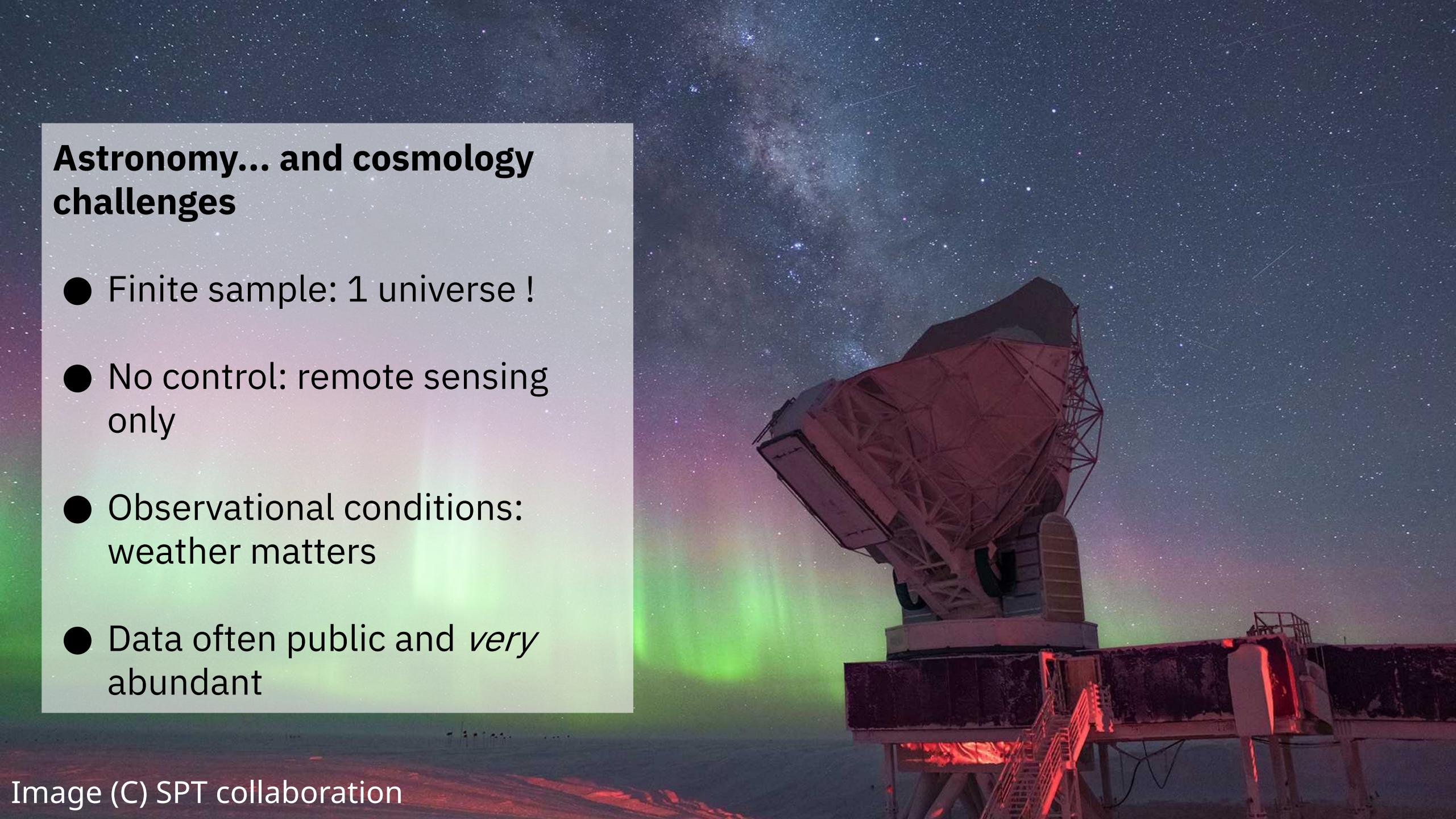
SCAI/AI4Science meeting



Guilhem Lavaux

Institut d'Astrophysique de Paris
CNRS & Sorbonne Université

9 March 2026



Astronomy... and cosmology challenges

- Finite sample: 1 universe !
- No control: remote sensing only
- Observational conditions: weather matters
- Data often public and *very* abundant

Image (C) SPT collaboration

Astronomy... and cosmology challenges

- Finite sample: 1 universe !
- No control: remote sensing only
- Observational conditions: weather matters
- Data often public and *very* abundant

Depending on problems we go for old or new techniques

- Self organizing maps
- Random forests
- Deep Neural network
- Explicit, physics constrained, inferences

Image (C) SPT collaboration

Astronomy... and cosmology challenges

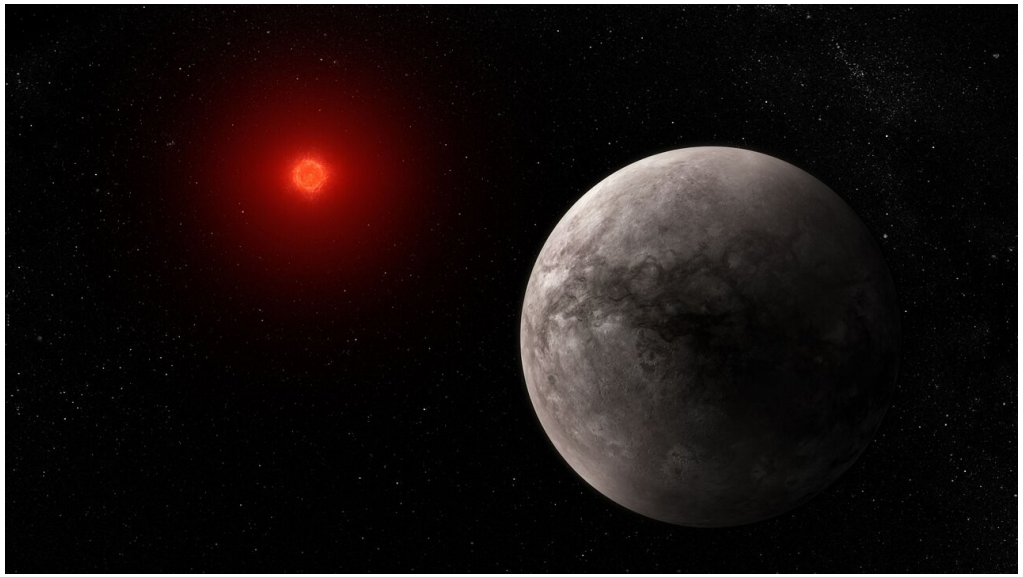
- Finite sample: 1 universe !
- No control: remote sensing only
- Observational conditions: weather matters
- Data often public and *very* abundant

Depending on problems we go for old or new techniques

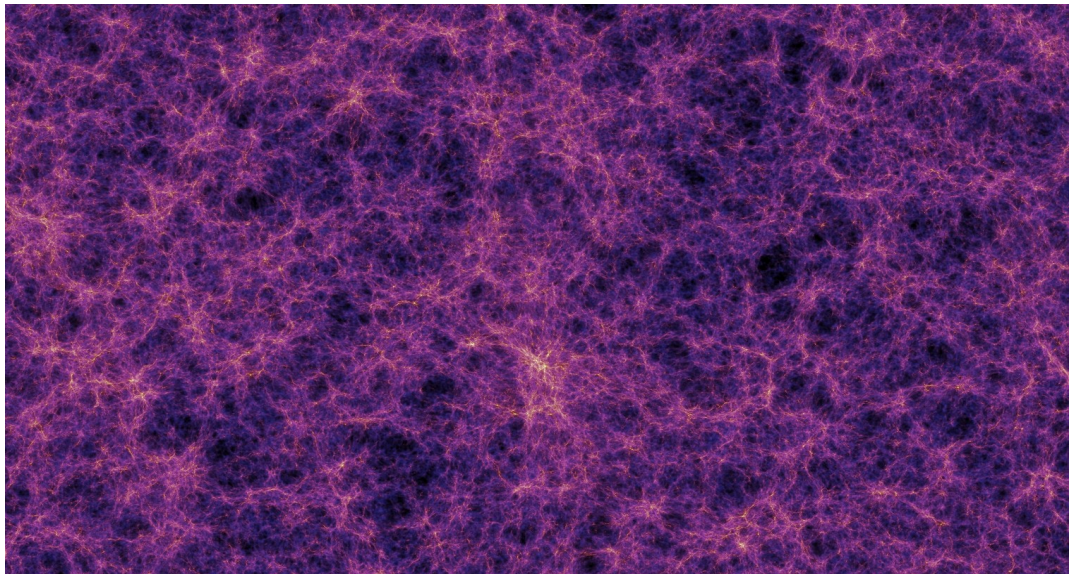
- Self organizing maps
  objects classification
  data selection characterization
- Random forests
  Object detection
- Deep Neural network
  Emulation, Bayesian inference
- Explicit, physics constrained, inferences
  10 to >10^9 parameters

Image (C) SPT collaboration

A biased selection of astronomy challenges



Exoplanets



Cosmology

© NASA, ESA, CSA, J. Olmsted (STScI), T. P. Greene (NASA Ames), T. Bell (BAERI), E. Ducrot (CEA), P. Lagage (CEA)

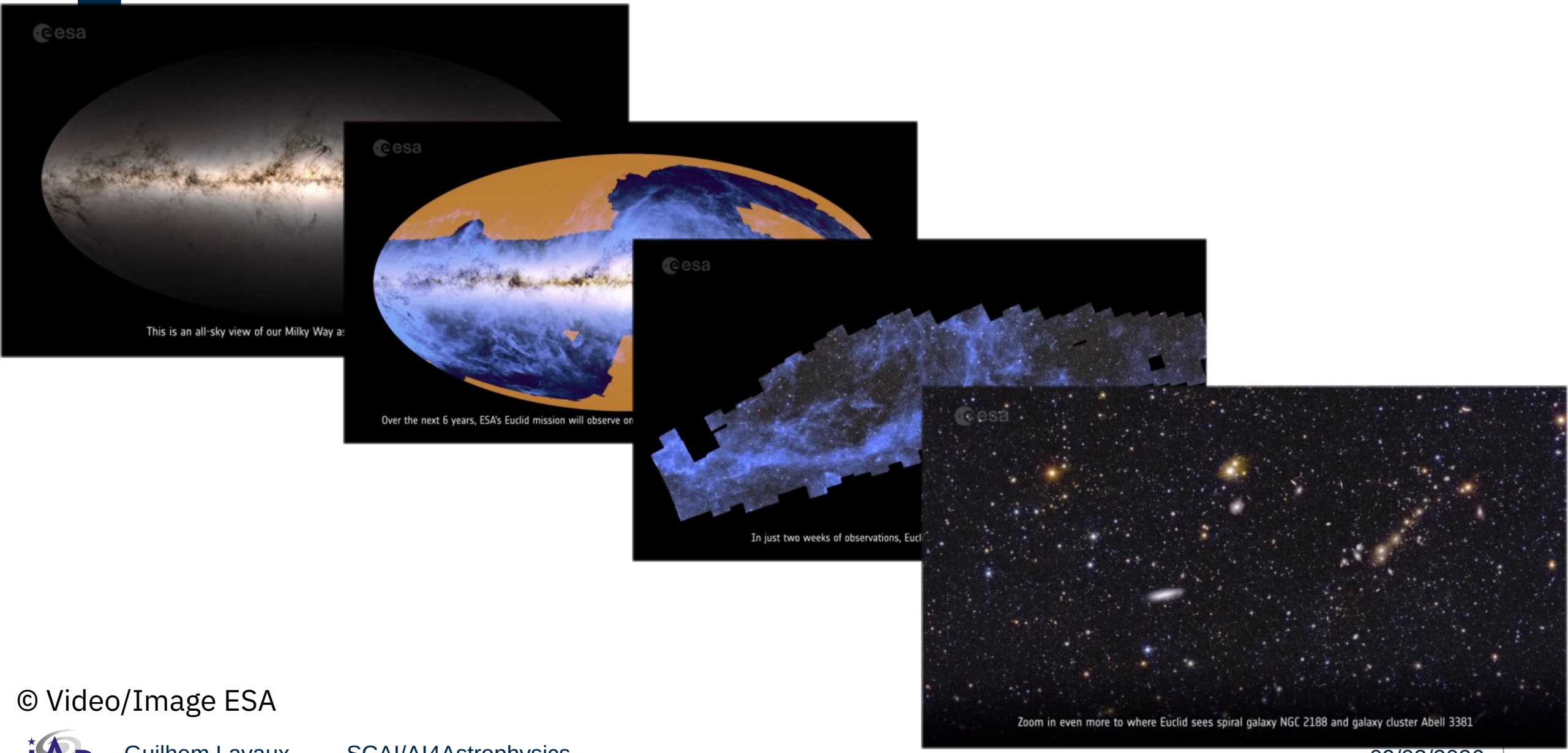
Outline

01 Exoplanets

02 Cosmology: Large scale Bayesian inference with BORG

03 Hybrid inferences

From astronomy surveys to physics

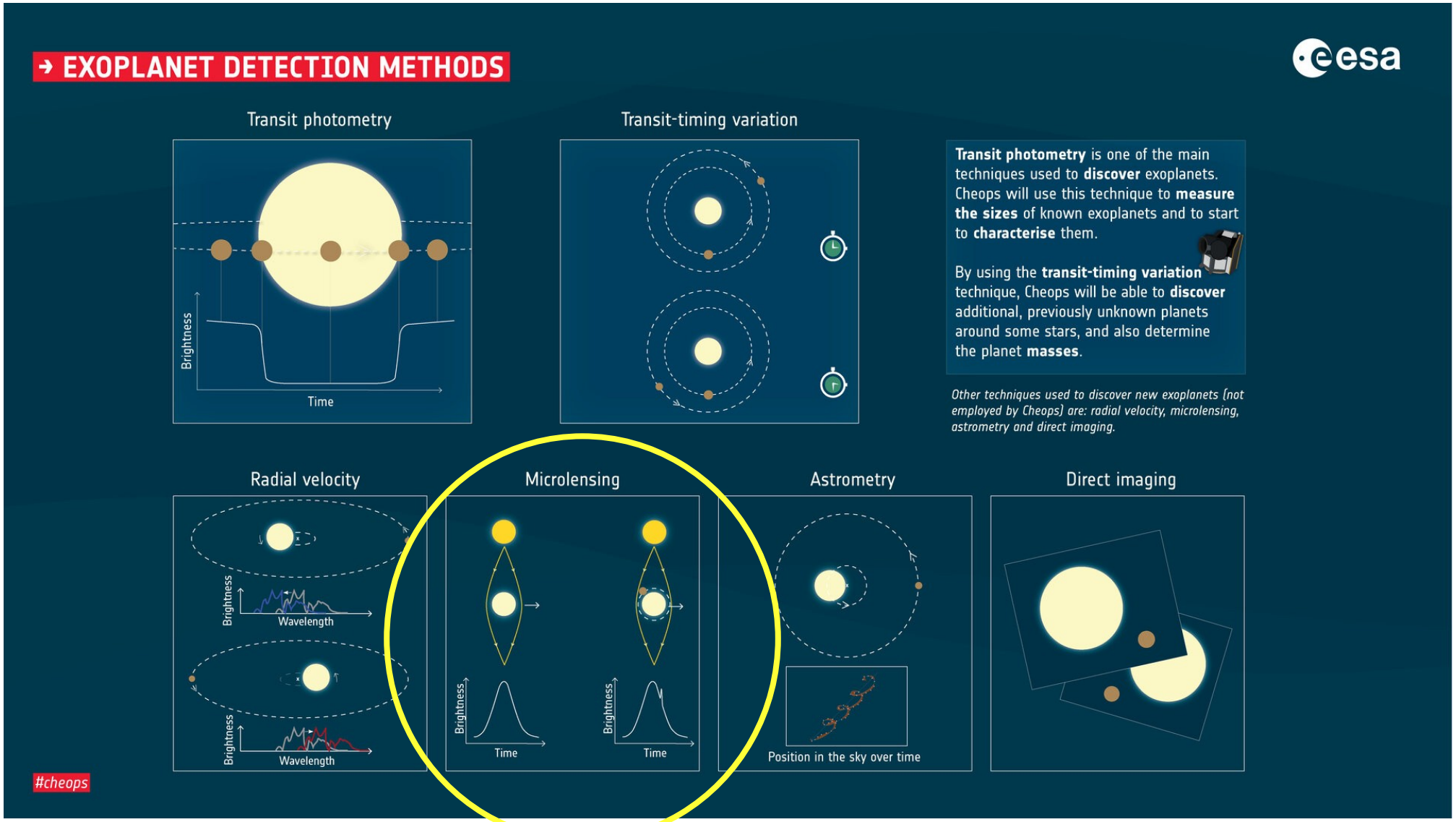


© Video/Image ESA

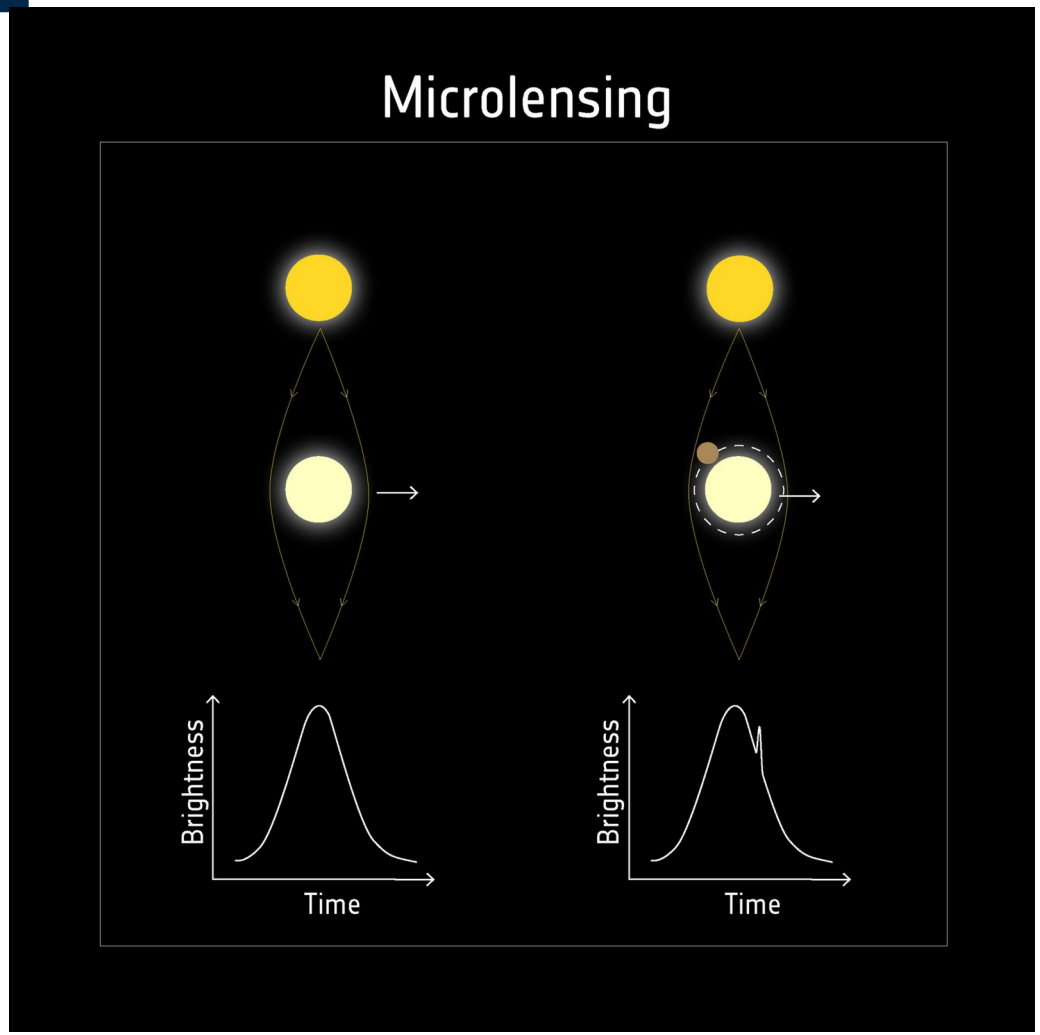
01

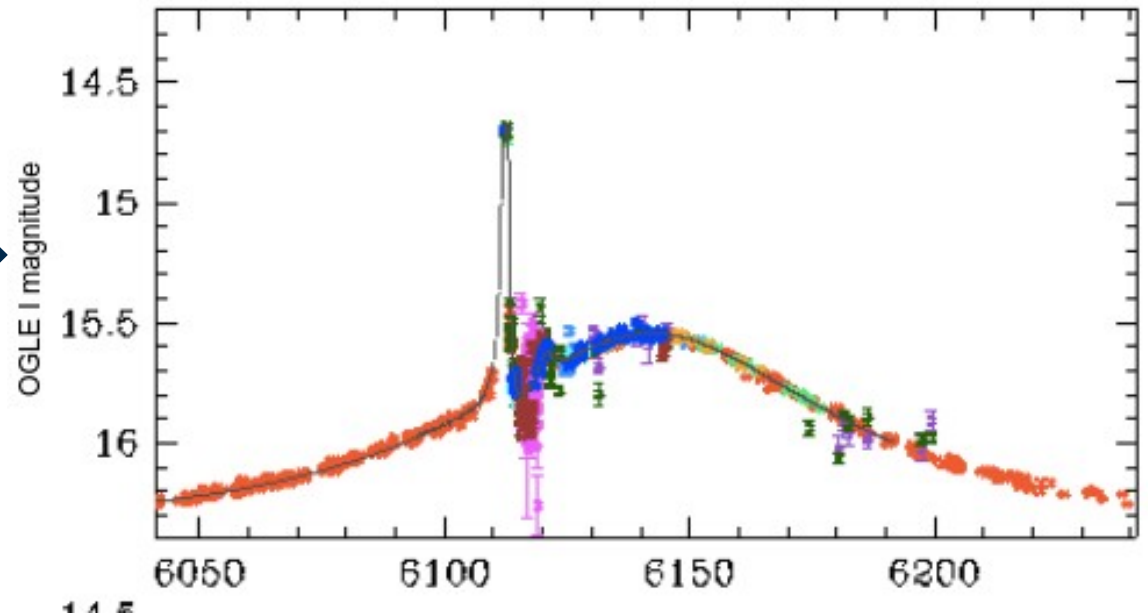
Exoplanets detection

Exoplanets detection challenges



Microlensing effect for planet detection





Microlensing event: OGLE-2012-BLG-0406

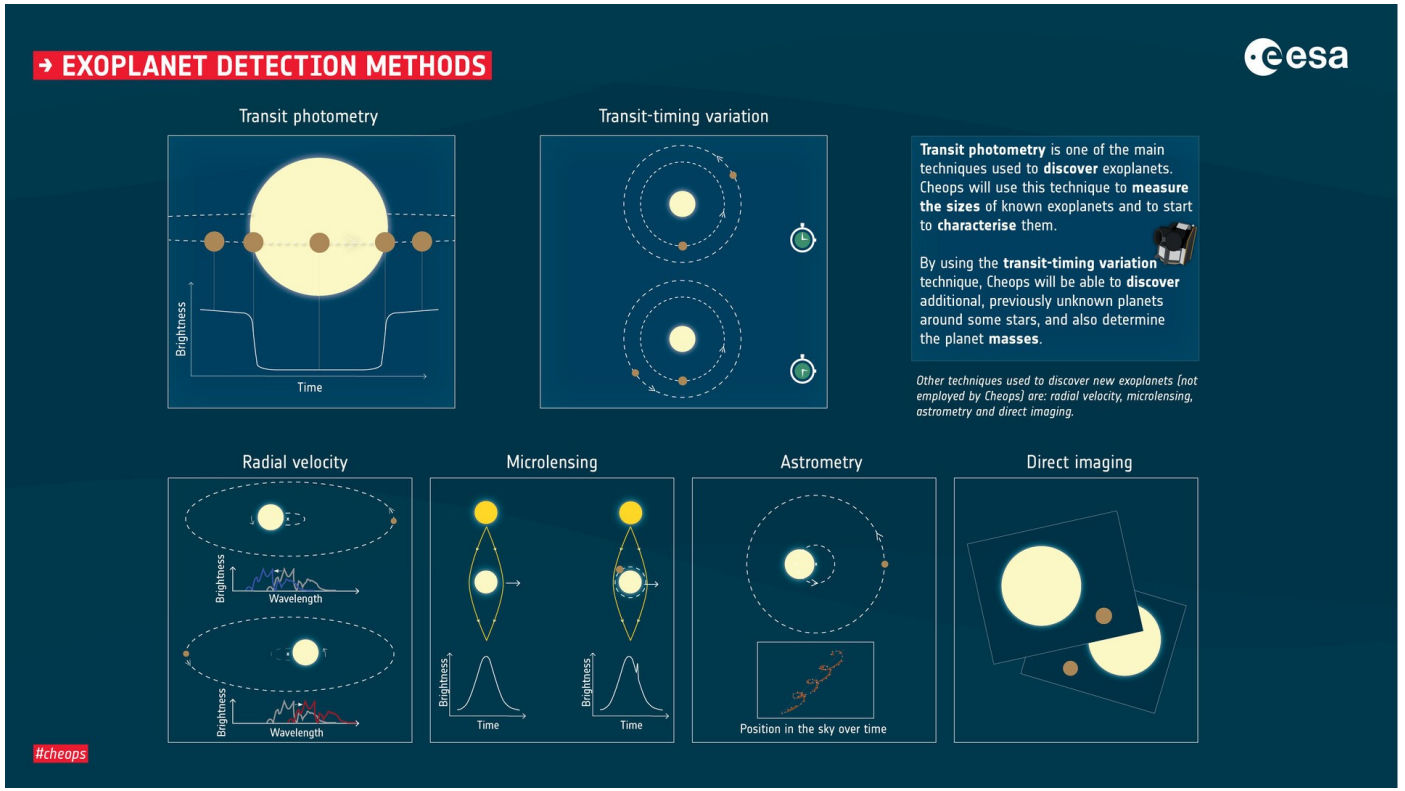
Important degeneracy between large and small sky separation for Star-Planet lensing

Figure: left © ESA, right: C. Ranc (2015)

Exoplanets detection challenges

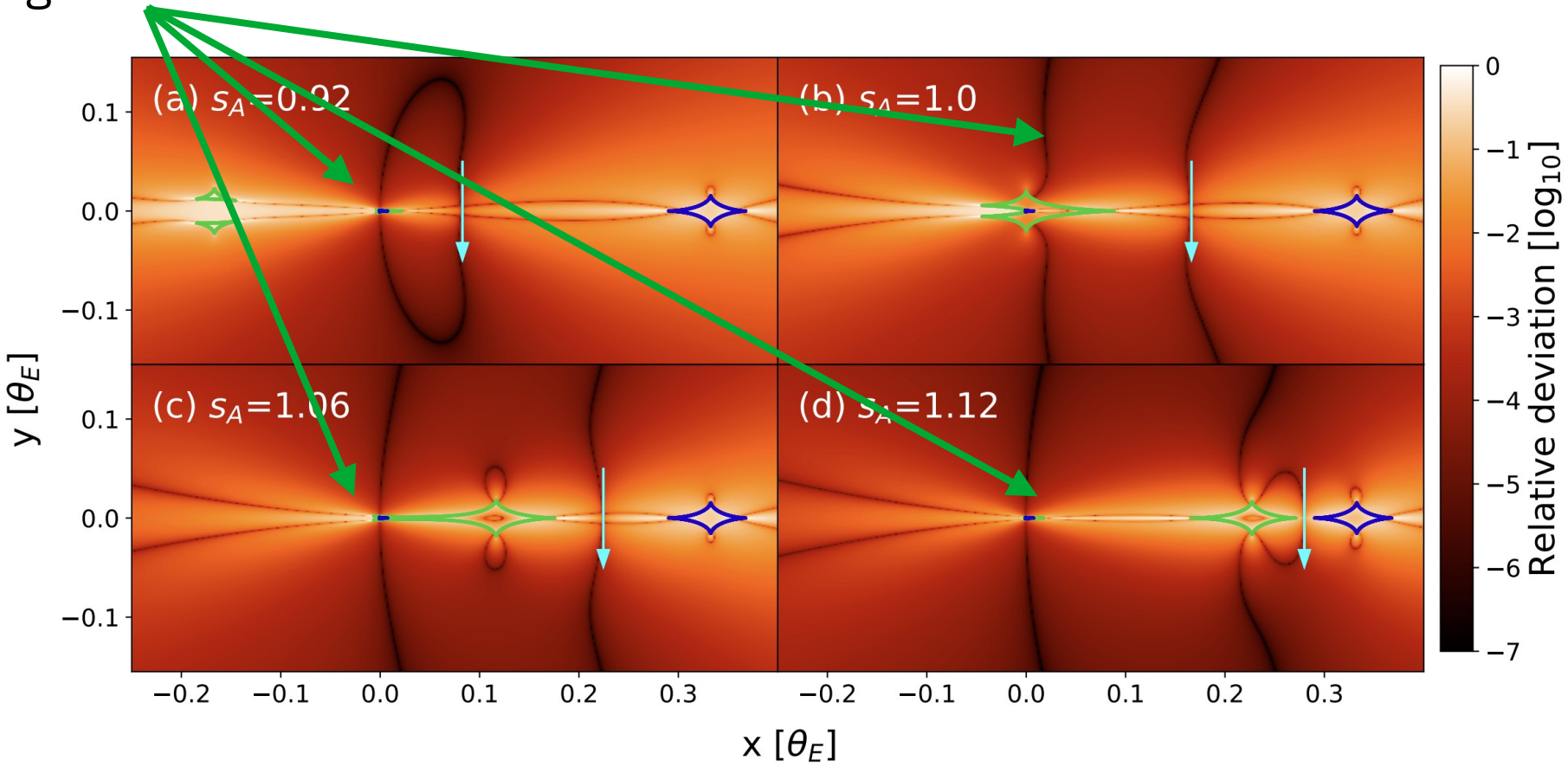
Convolutional Neural Network → at work for detection/characterization

Neural Posterior Estimator → inferences of microlensing orbital parameter



Exoplanets theoretical degeneracies discovered through AI

- Microlensing detection of planets relying on amplitude magnification
- New degeneracies found with Neural Posterior Estimation

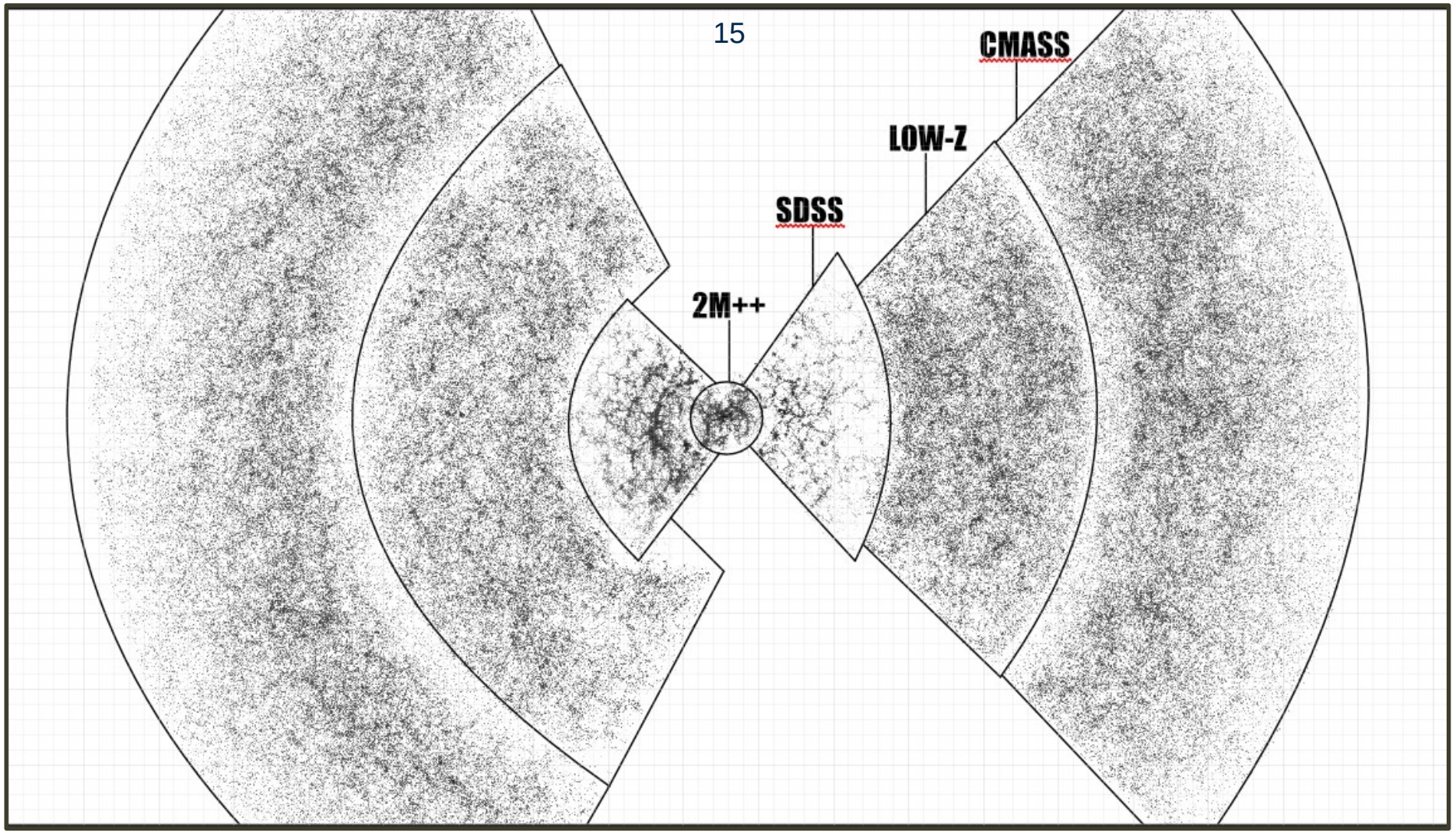


02

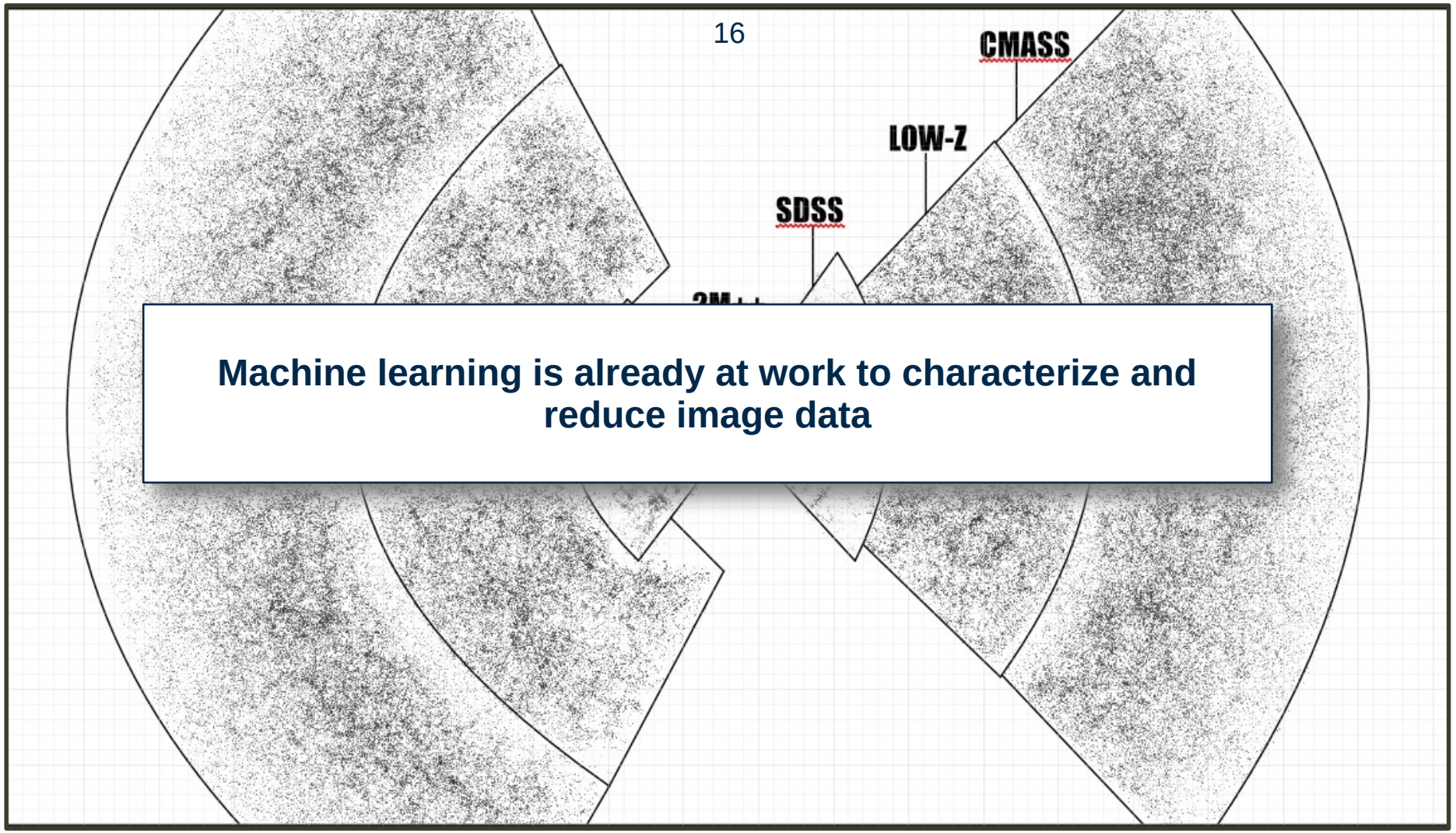
Cosmology: Large scale Bayesian inference with BORG

From image surveys to galaxy surveys

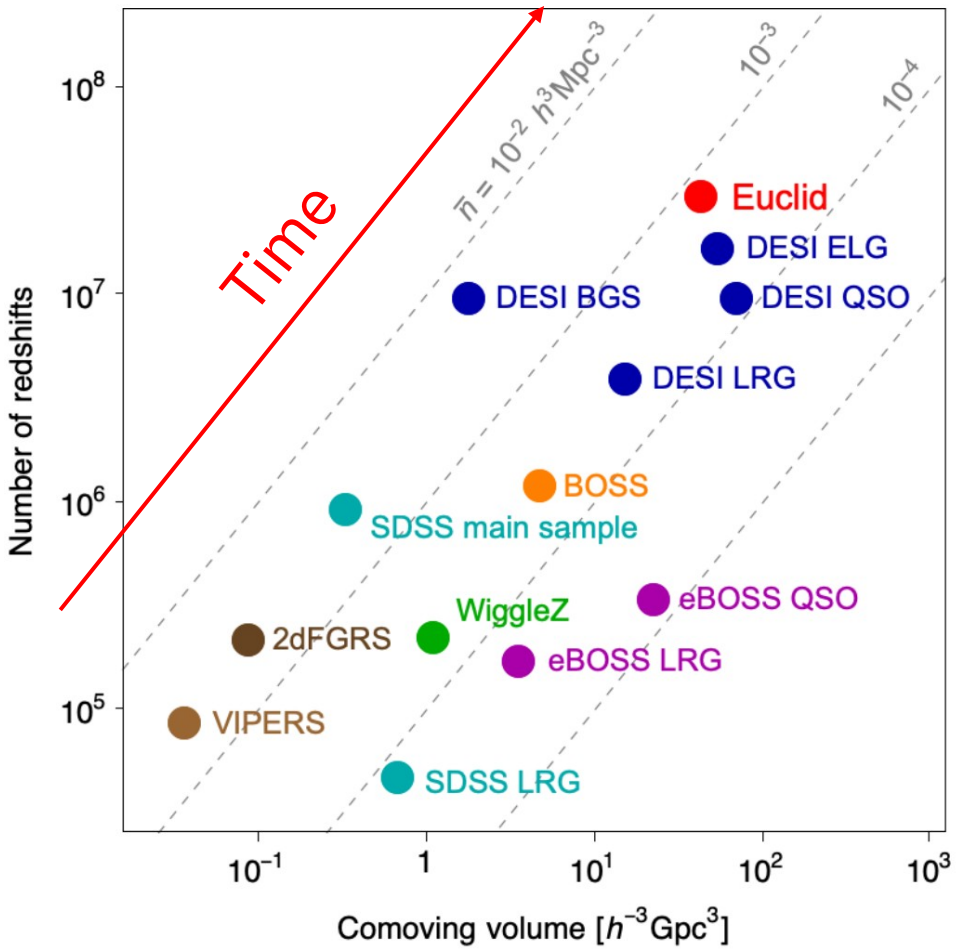
Reduced datasets: how astronomical surveys looked in 2010



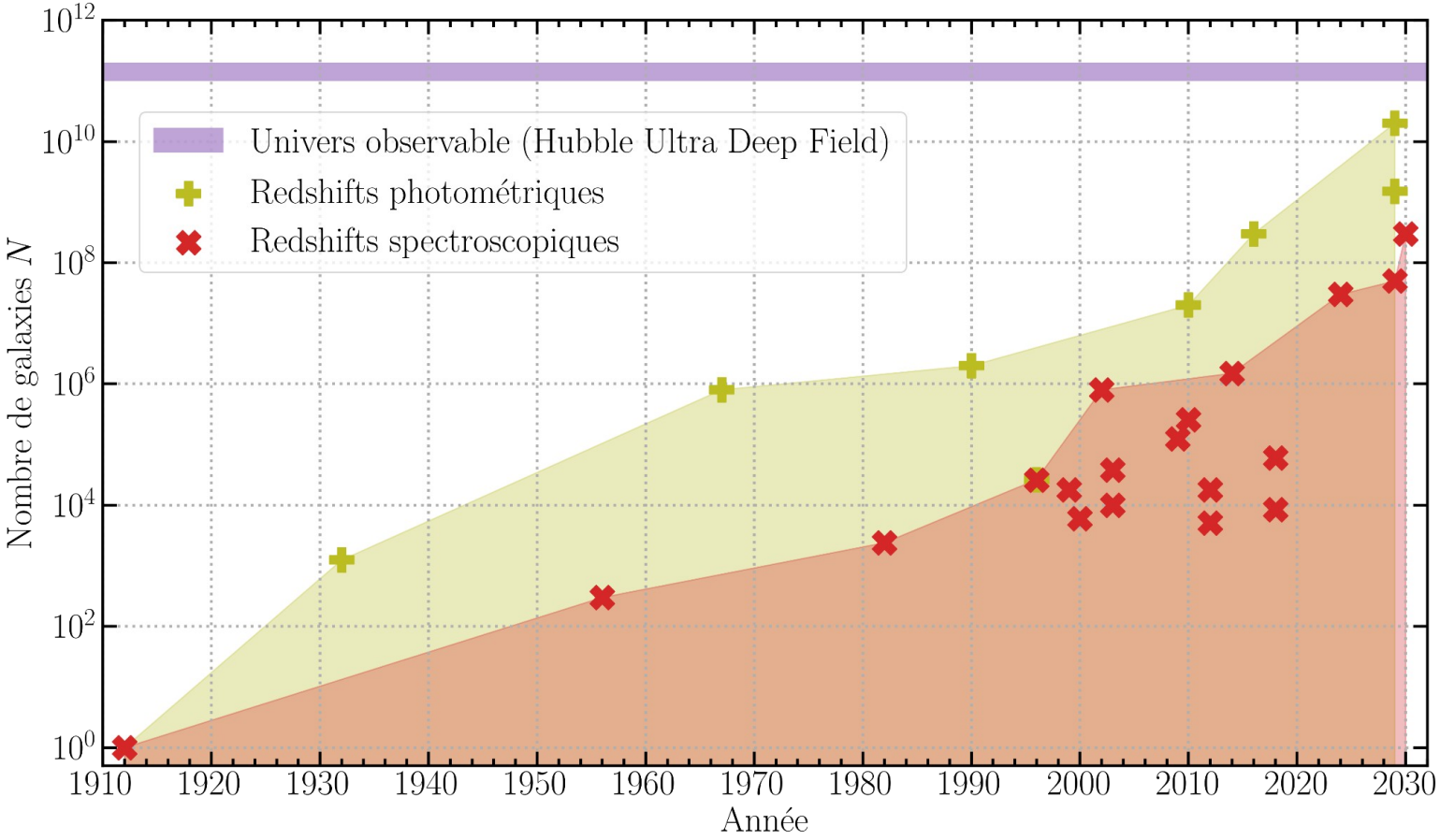
Reduced datasets: how astronomical surveys looked in 2010



The astronomy data revolution for cosmological analysis

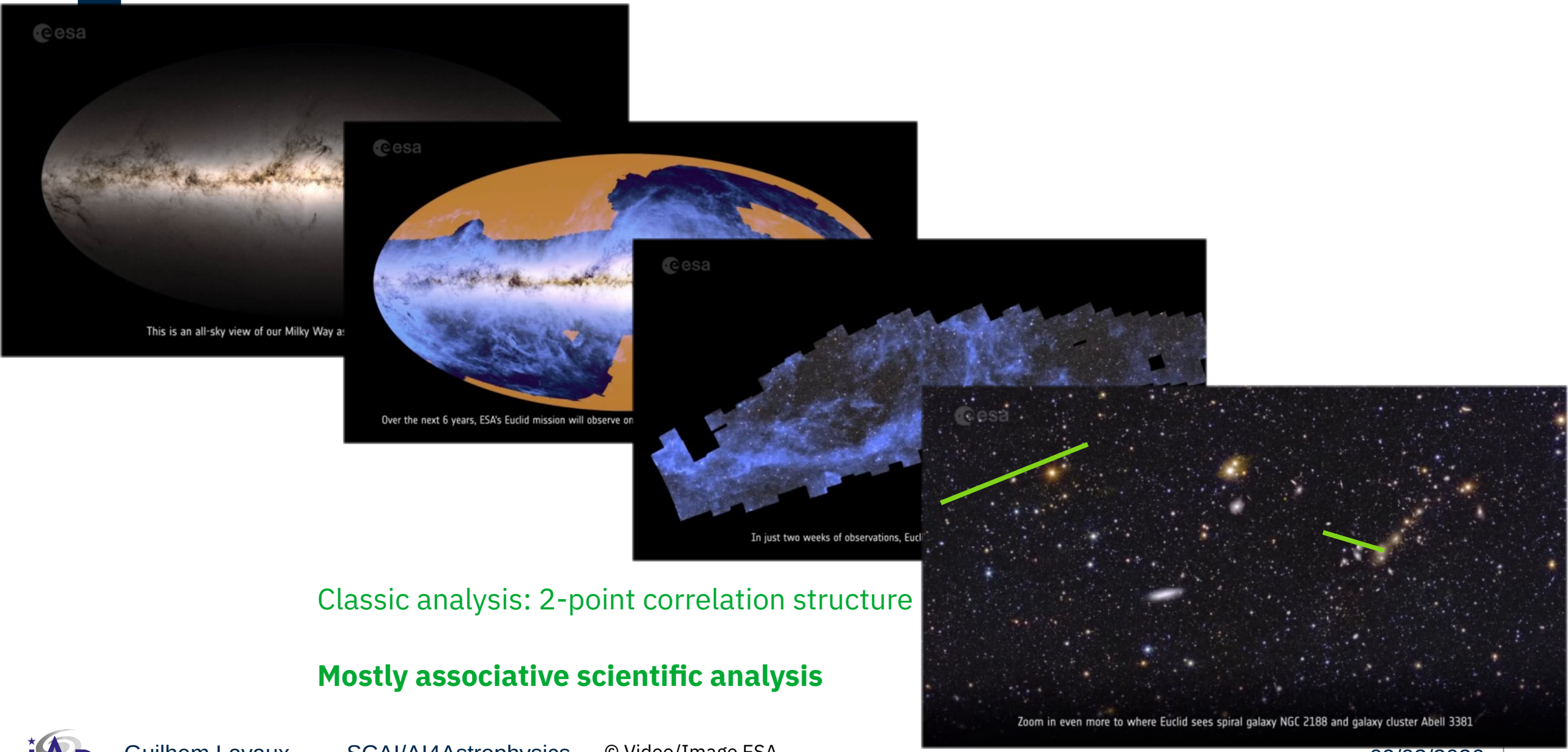


Euclid Overview, Euclid Collaboration, A&A 2025



Credit: F. Leclercq, J. Jasche (data), inspired from J. Peacock plot

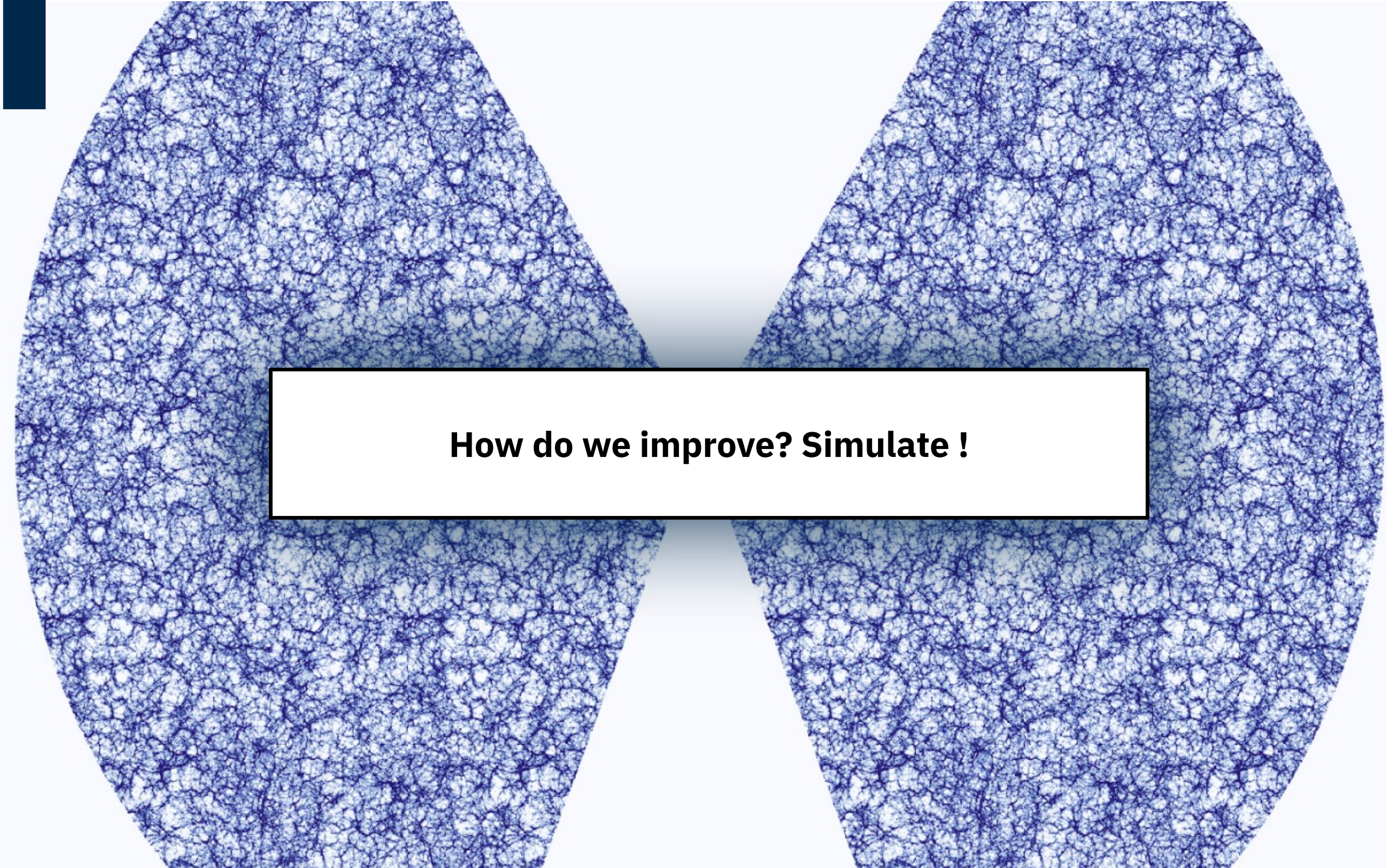
From astronomy surveys to physics



Classic analysis: 2-point correlation structure

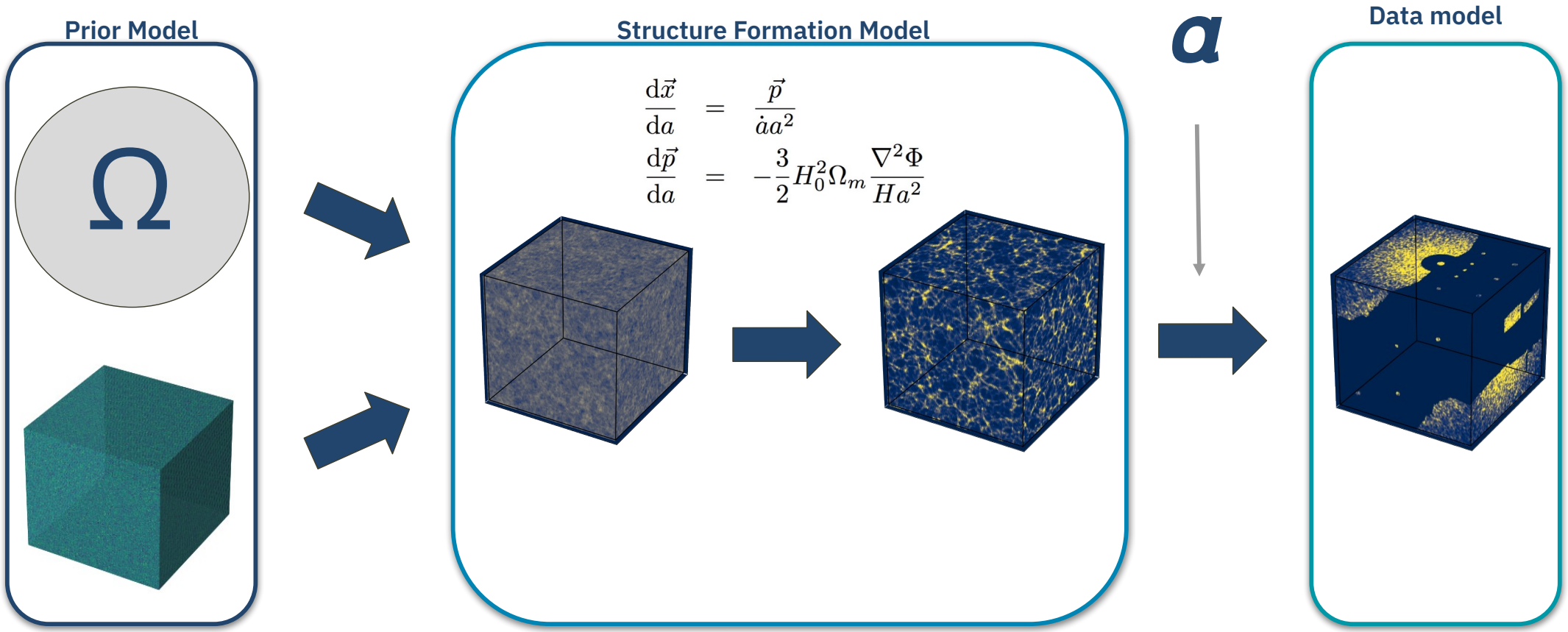
Mostly associative scientific analysis

© Video/Image ESA



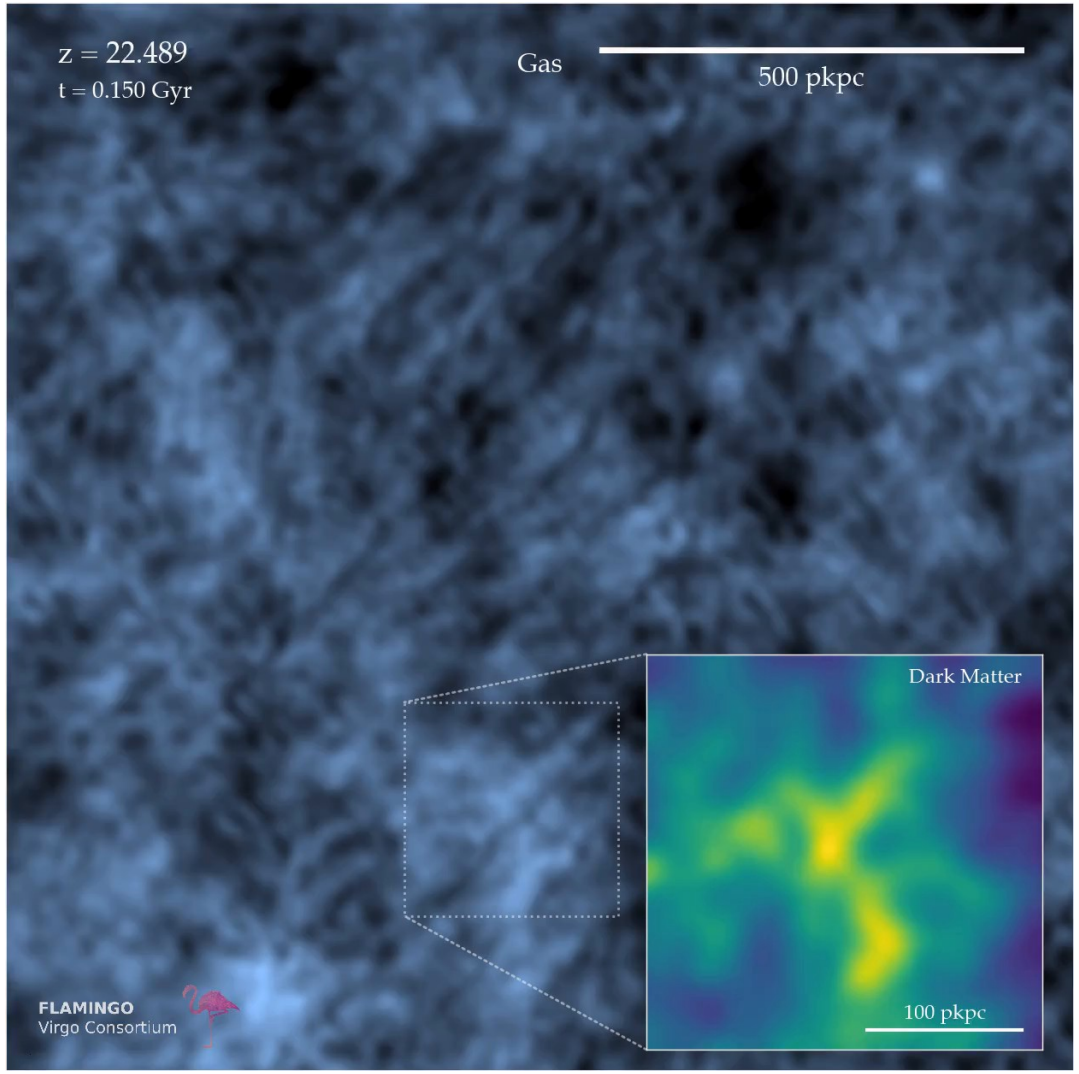
How do we improve? Simulate !

A typical simulation pipeline of the universe in a nutshell



$$\pi\left(\mathbf{x}, \mathbf{\Omega}\right) \qquad \pi\left(\rho_\mathbf{m} | \mathbf{x}, \mathbf{\Omega}\right) \qquad \pi\left(\mathbf{N_g} | \rho_\mathbf{m}, \alpha, \mathbf{\Omega}\right)$$

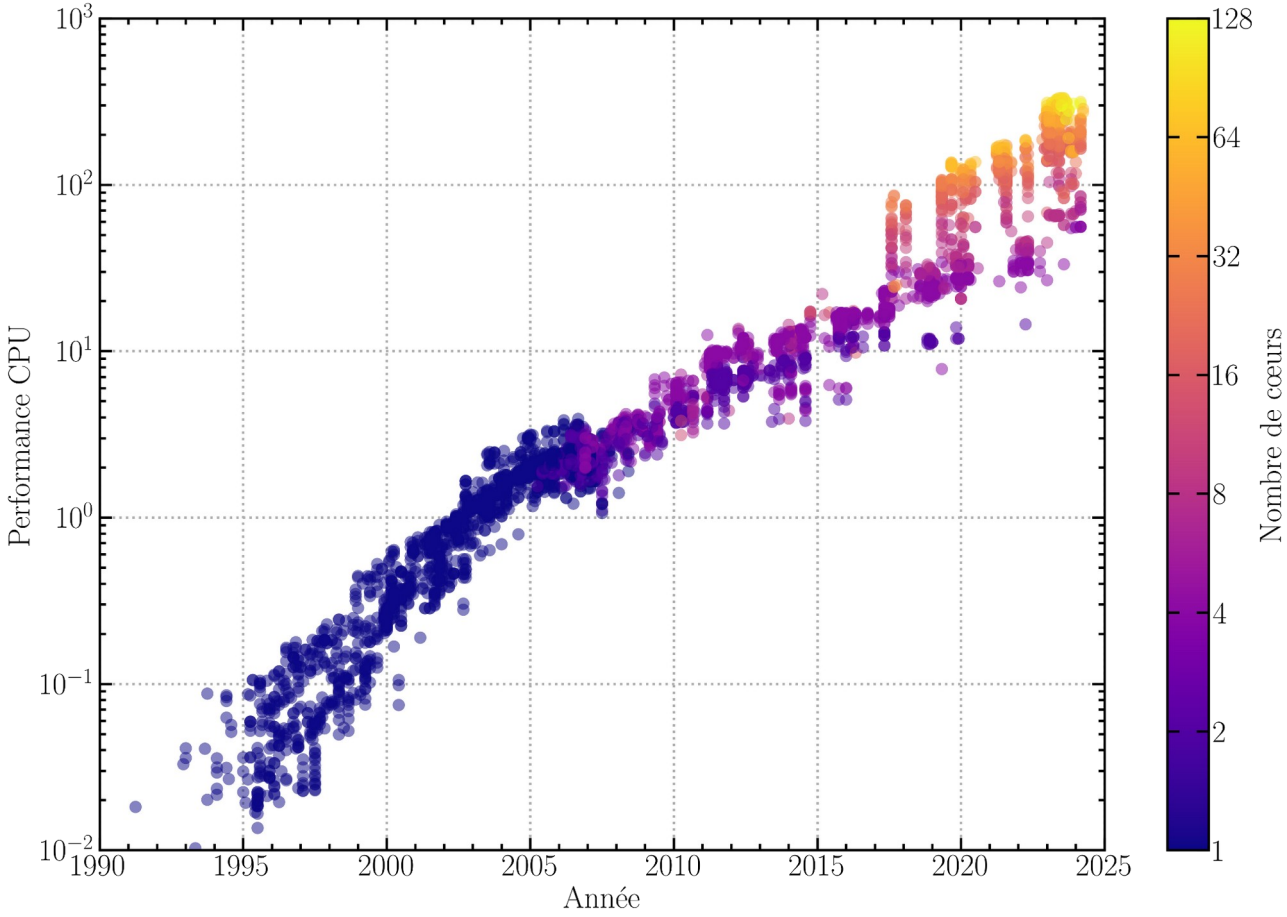
Simulations become increasingly complex with detailed physics



Credit: Flamingo collaboration / Virgo consortium

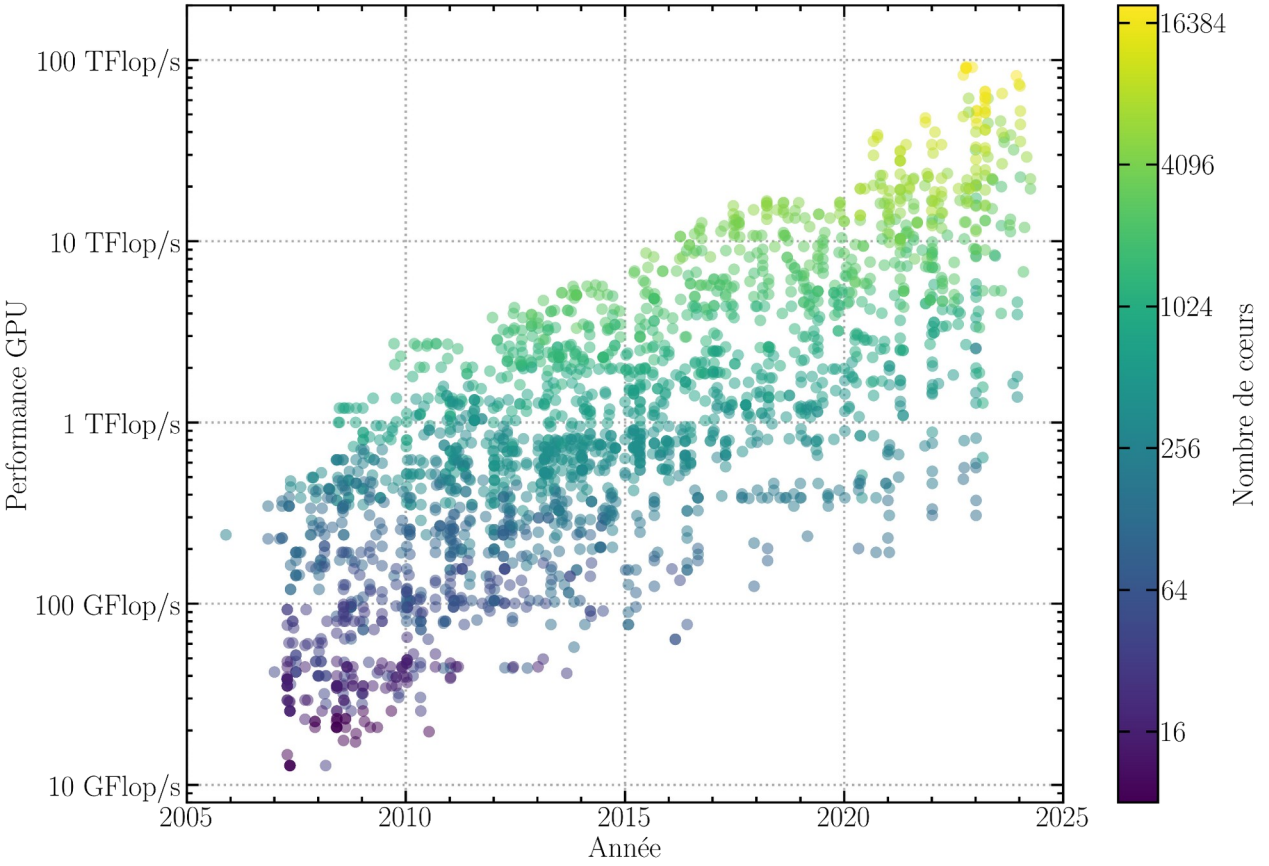
High-performance computing has exploded... but now plateauing

CPU



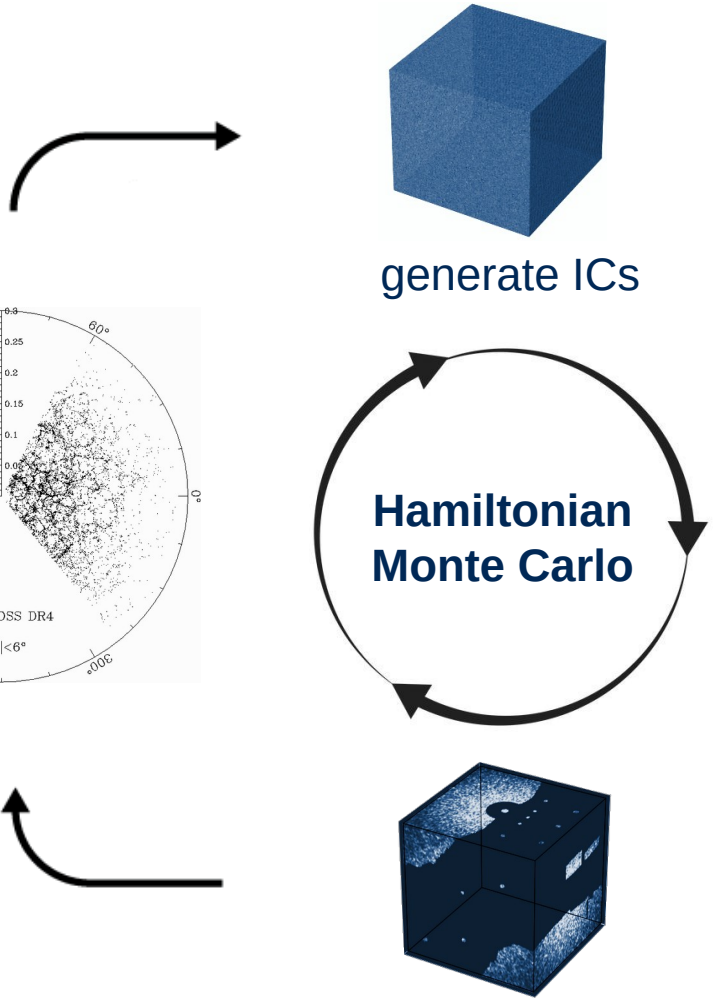
Credit: F. Leclercq, data from spec.org

GPU

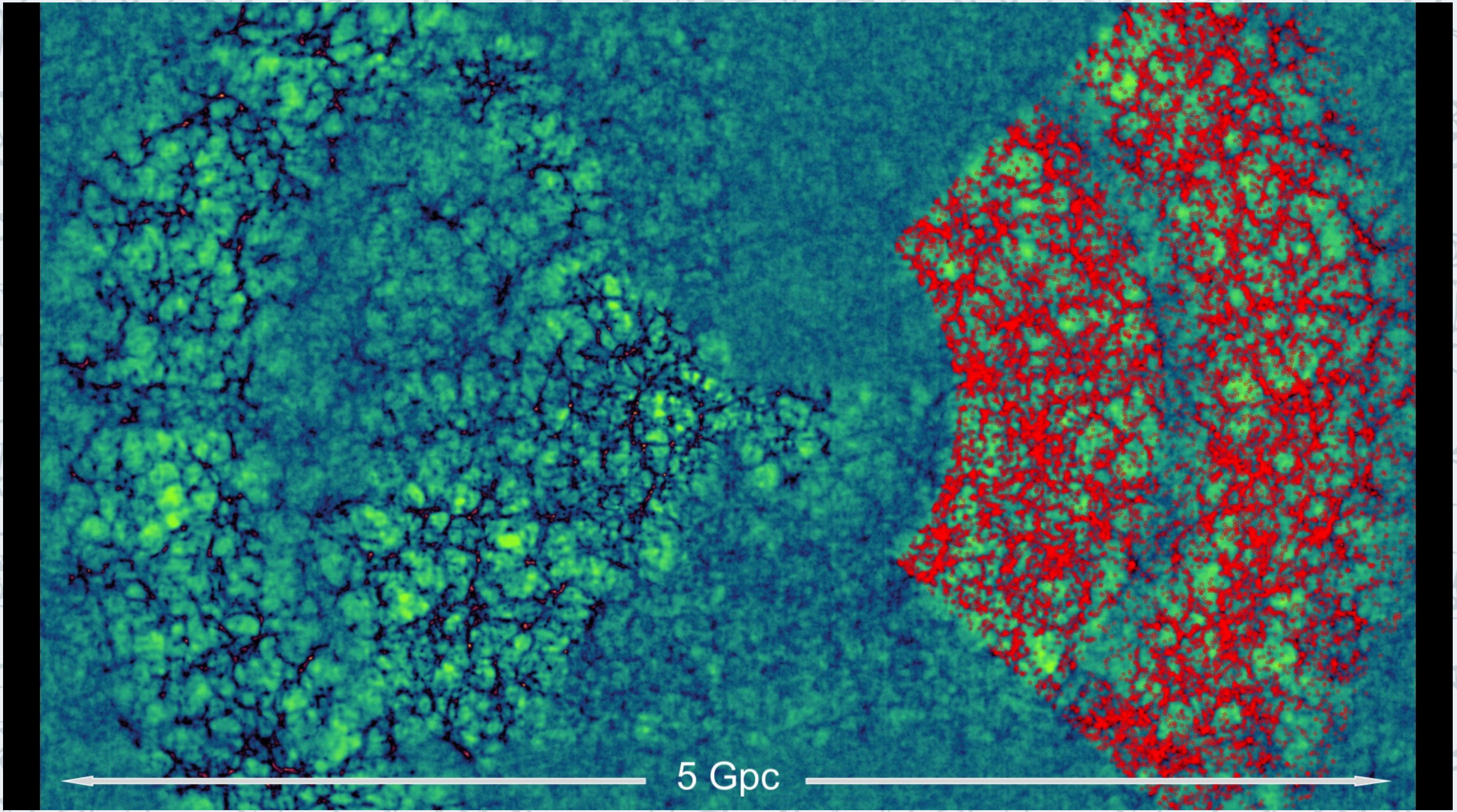


Credit: F. Leclercq, data from techpowerup.org

BORG: a fully Bayesian density inference loop



J. Jasche, F. Leclercq, B. D. Wandelt 2015
G. Lavaux, J. Jasche, F. Leclercq 2019
G. Lavaux, J. Jasche 2016
Jasche & Wandelt 2014
Jasche & Lavaux 2019

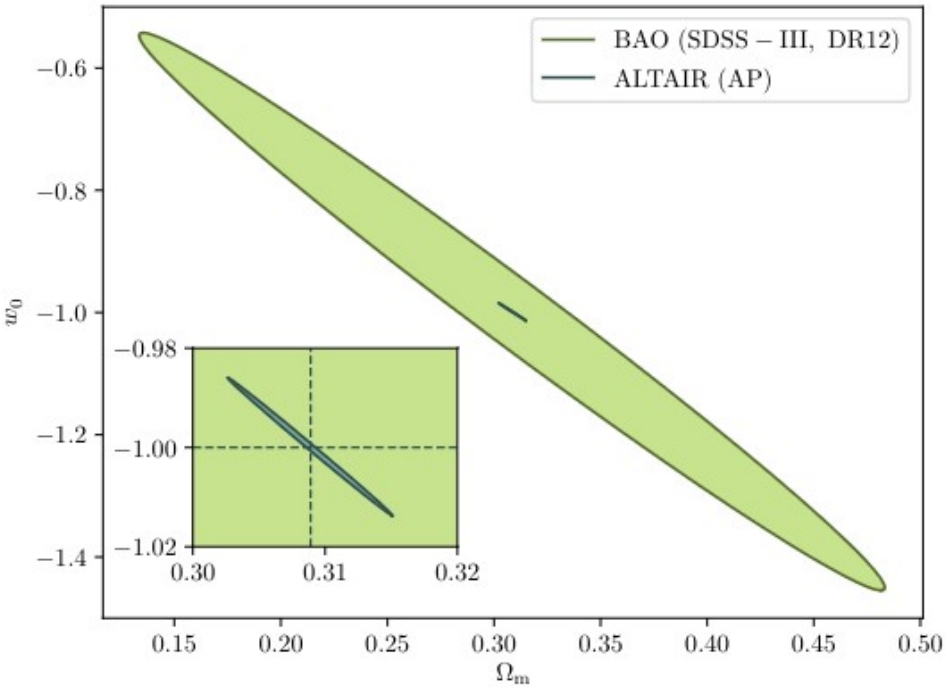


Credit: S. McAlpine (Stockholm University), Manticore project to build a Digital Twin

Testing joint inferences with physical parameters

Dark Energy / Matter content

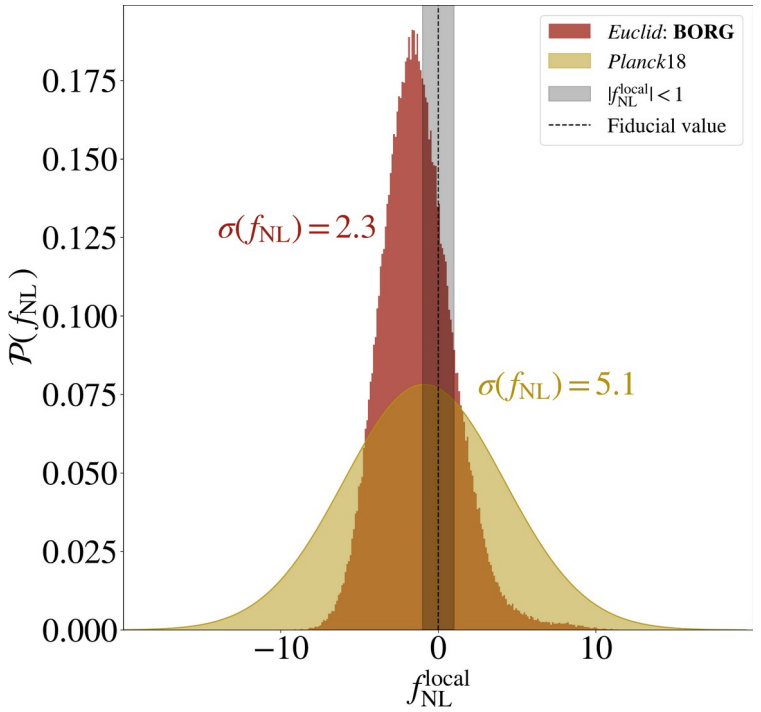
(w, Ω_m)



Kodi Ramanah et al. 2019

Non-gaussianity of initial conditions

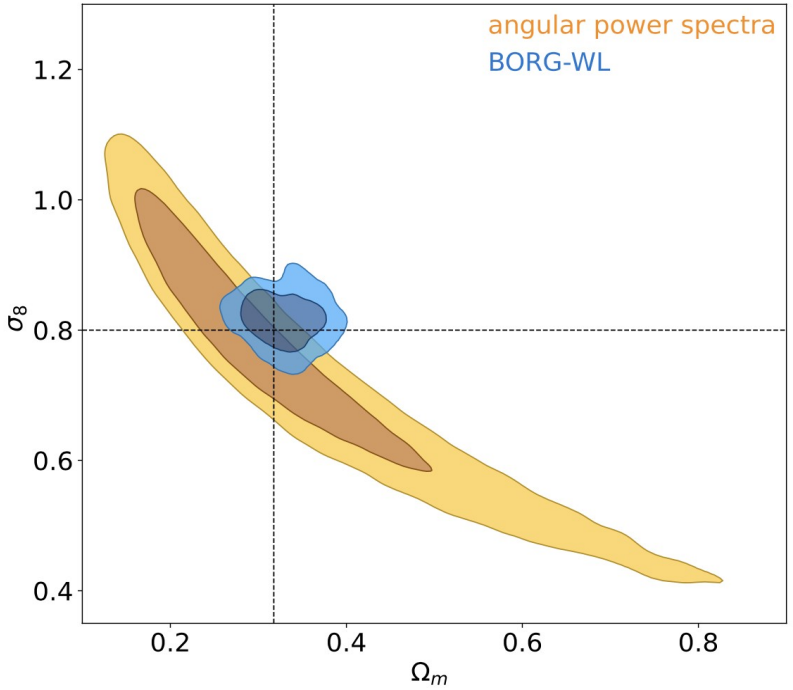
f_{NL}



Andrews et al. 2022, 2024

Amplitude of initial conditions

(σ_8, Ω_m)

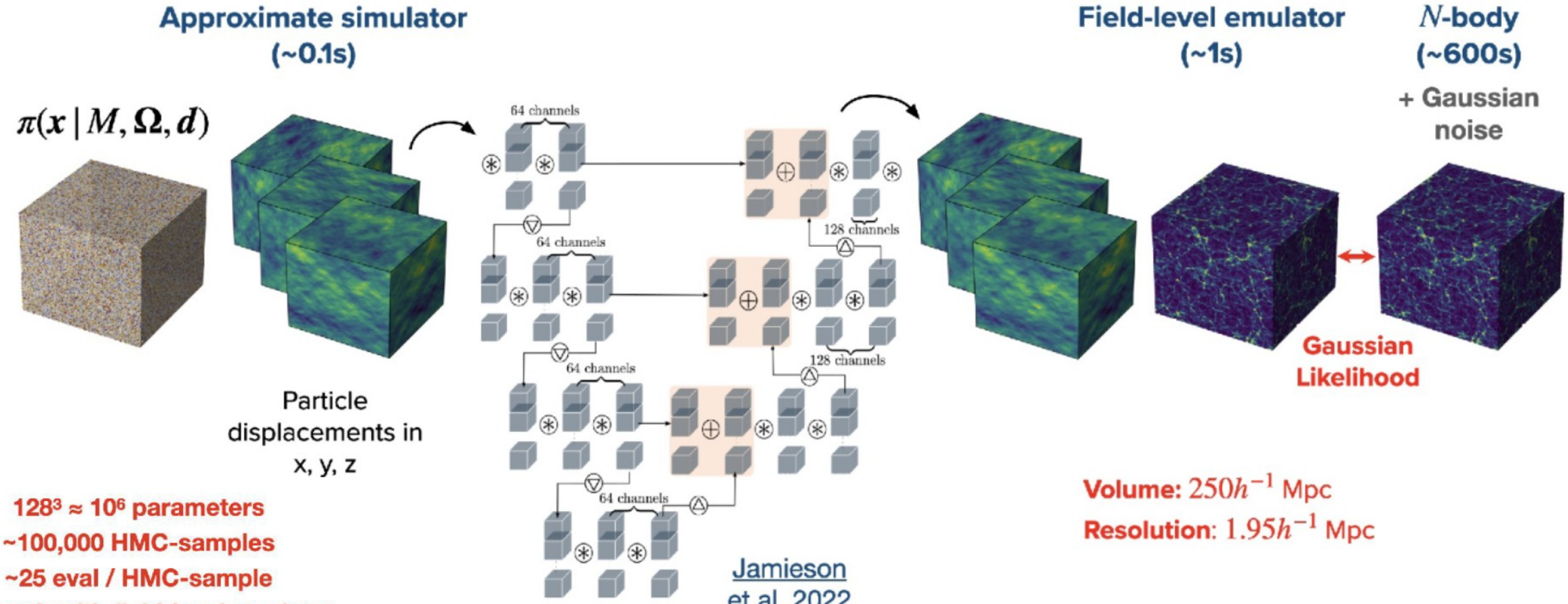


Porqueres et al. 2021, 2022, 2023

03

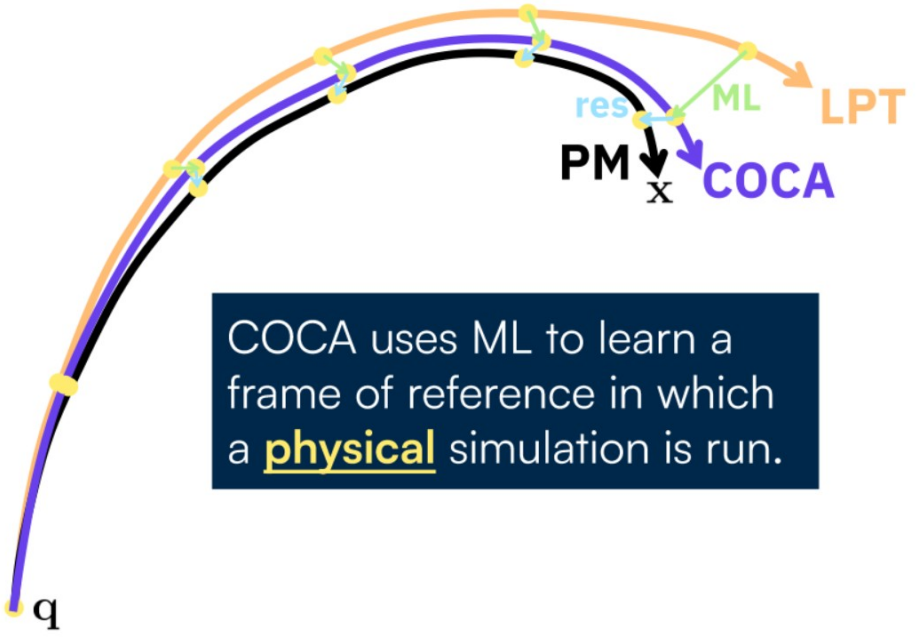
Hybrid inferences: accelerating simulations

Inference acceleration through density emulators: BORG-EMU

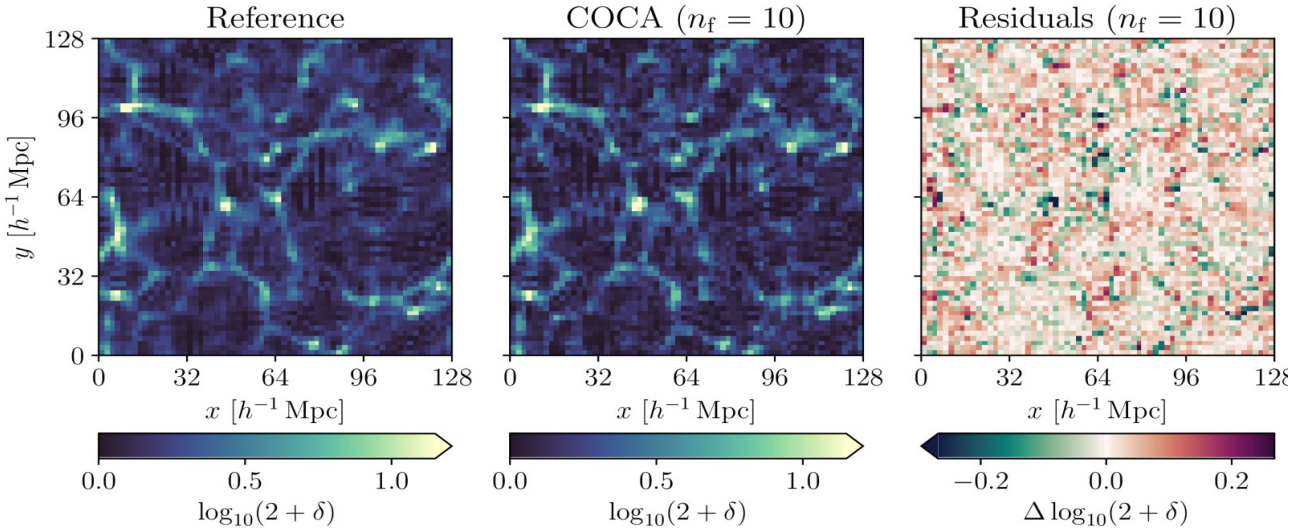


Acceleration with ML-safety: COCA

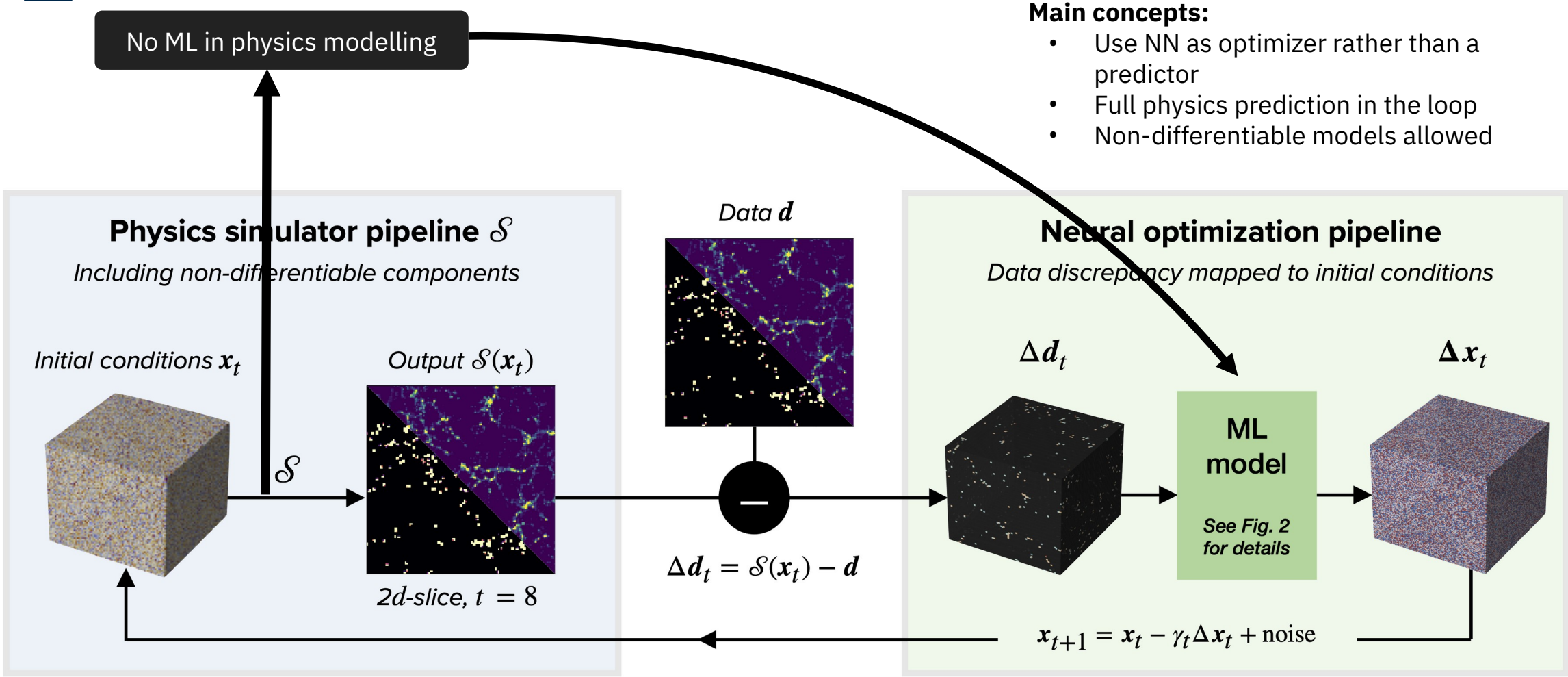
- Objective: design **simulation accelerator** with **guaranteed** physical accuracy asymptotically
- Frame of reference expansion for N-body integration
  - COLA: analytic perturbation theory
  - COCA: Deep learning emulator
- Design:
  - Skip force evaluation / Keep timestep integration
- Performances:
  - 1% accurate for 10x less force evaluation
  - Resilient to out-of-parameter training data



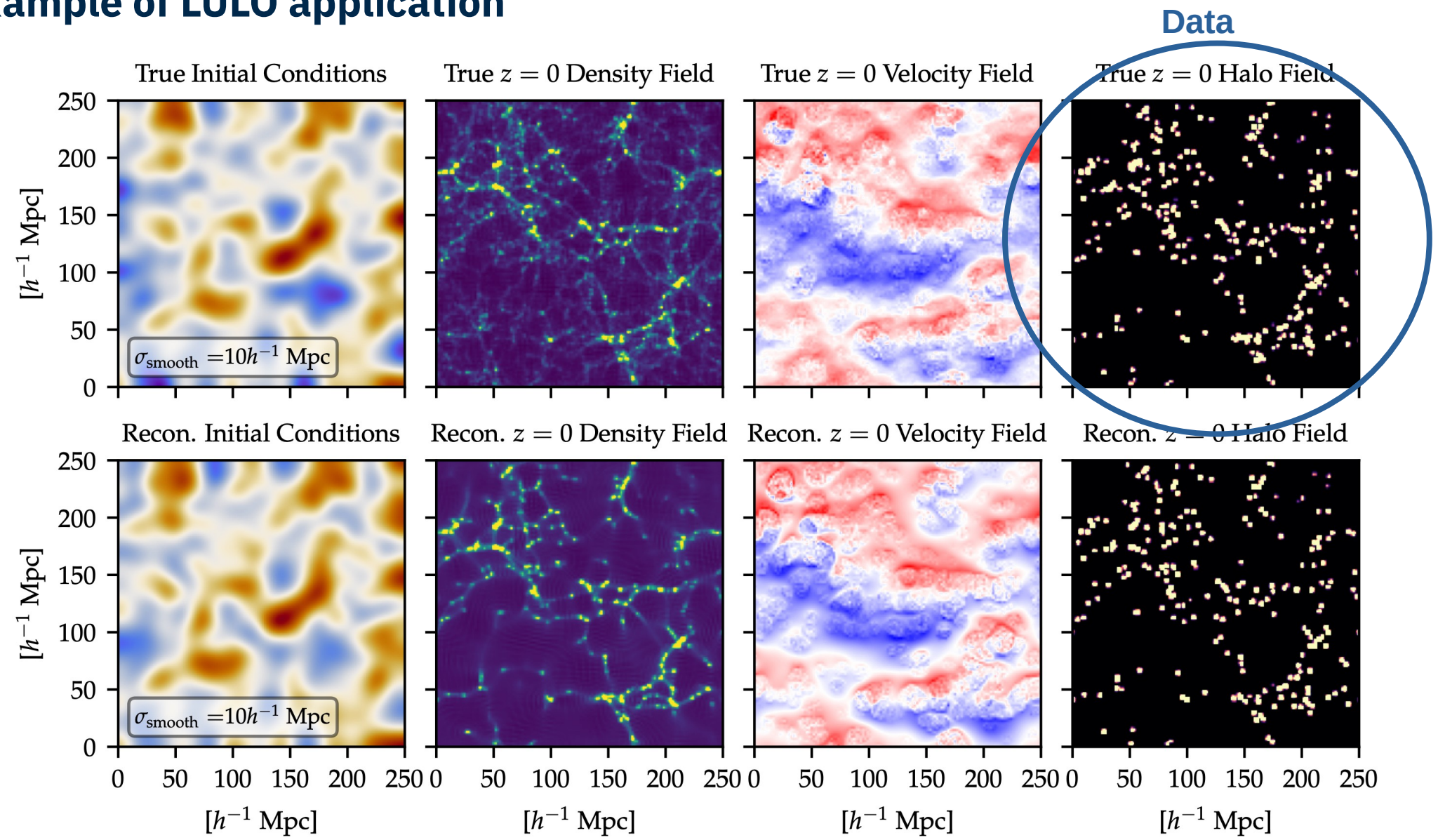
COCA uses ML to learn a frame of reference in which a **physical** simulation is run.



ML safety in inference: Learning the Universe by Learning to Optimize (LULO)

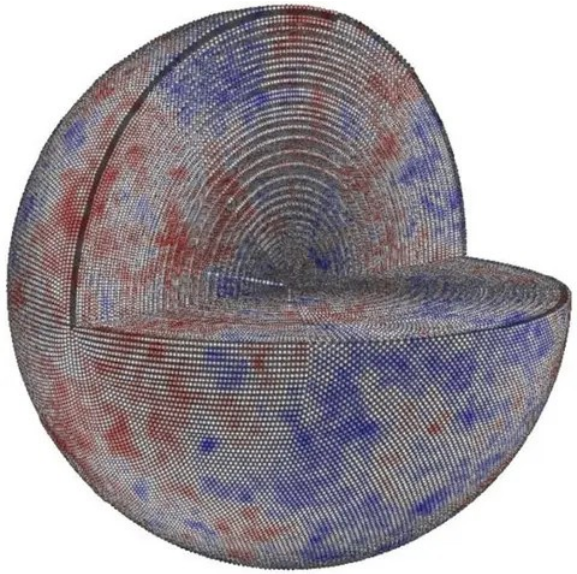


Example of LULO application



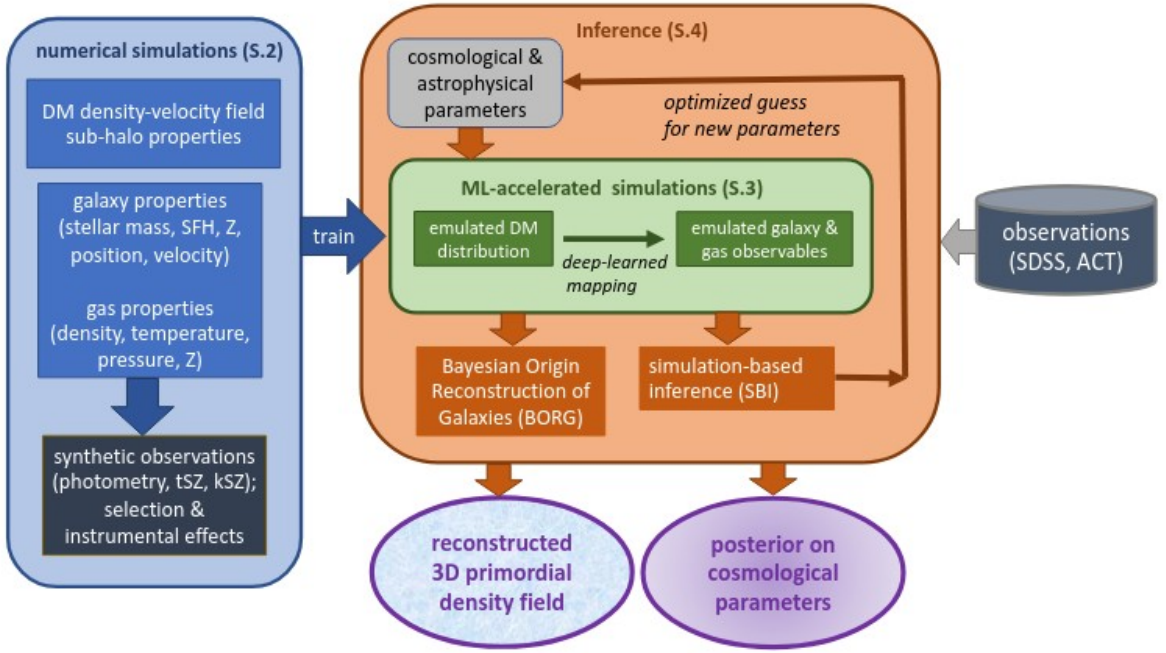
04 Conclusion

Simons Foundation: the "Learning the Universe" collaboration



Formed & funded (2021) → large scale cosmological inference with AI techniques

- USA:
  - Flatiron/CCA
  - Columbia University
  - Johns Hopkins University
  - Pennstate
  - University of California at Berkeley
  - Harvard University
- Europe:
  - CNRS/IAP
  - MPG/MPA Garching
  - Stockholm University



Astronomy requires AI/ML/DeepLearning in all its aspects

- Data:
  - Signal detection (Random Forests, ...)
  - Data filtering (e.g. SOMs)
- Simulation:
  - Simulation acceleration / emulation (BORG-EMU, COCA, ...)
  - Deep Learning boosted optimization (LULO)
- Automatic summary statistics:
  - Information Max. Neural Networks
- Inference:
  - Kernel density estimate (Implicit Likelihood Estimation)
  - Evidence computation (Harmonic Normalizing Flow estimator)
- Software development & Foundation models in development