Bayesian hierarchical models

Advanced sampling

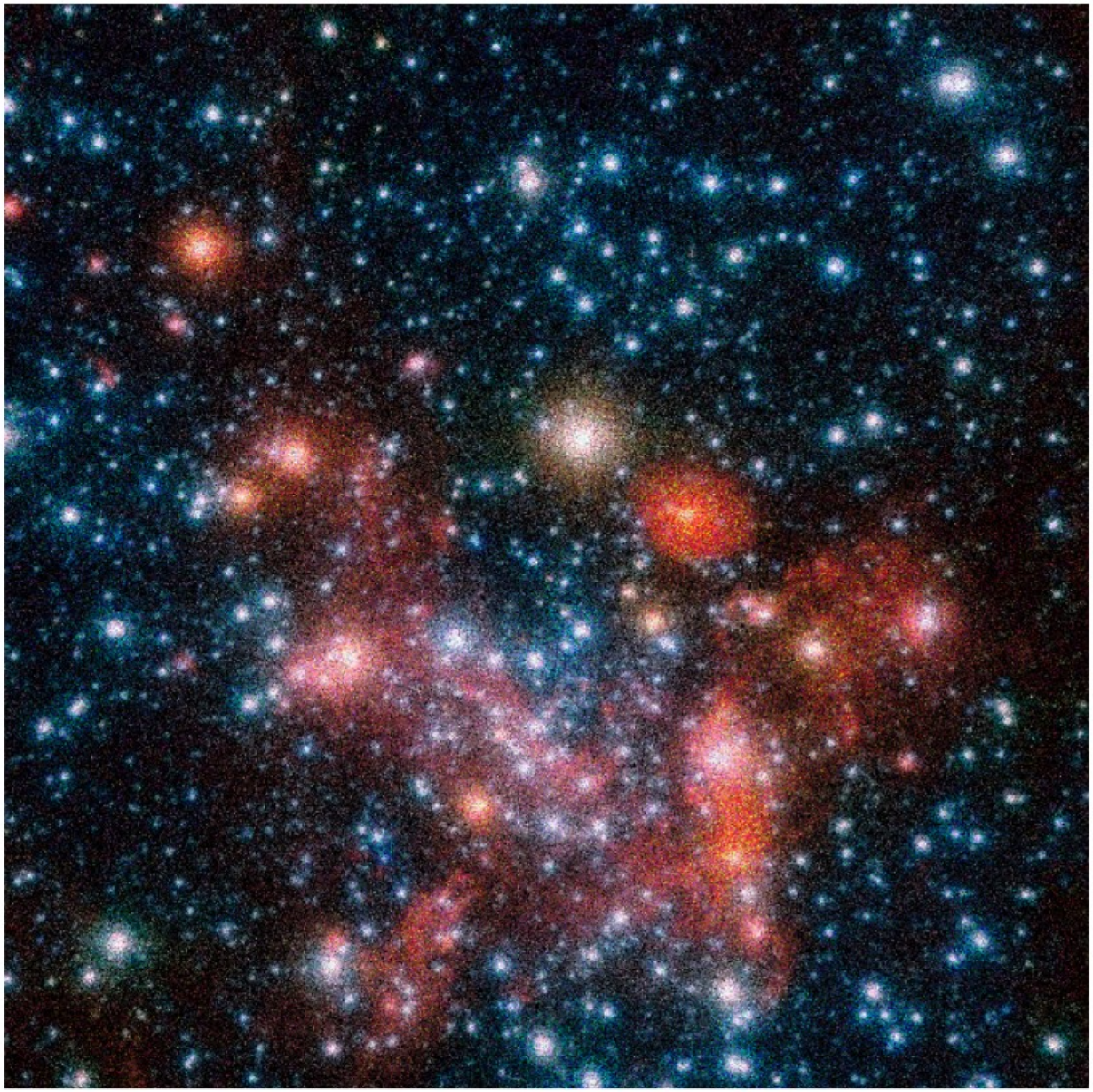
Natalia Porqueres

ICIC school – Sep 2024

True stellar field

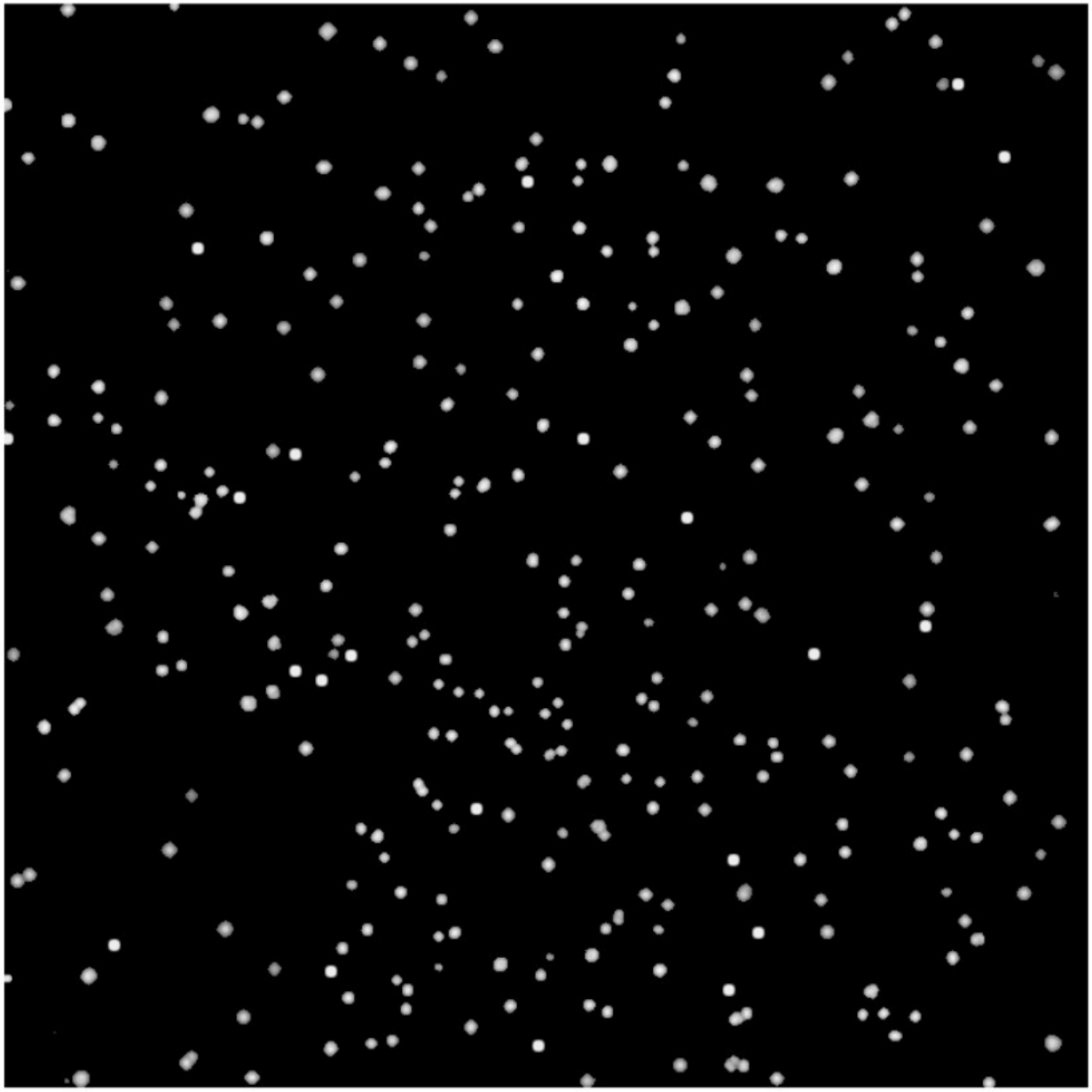
ESO

Stellar field with noise detector



ESO + ES

Assuming chi-square

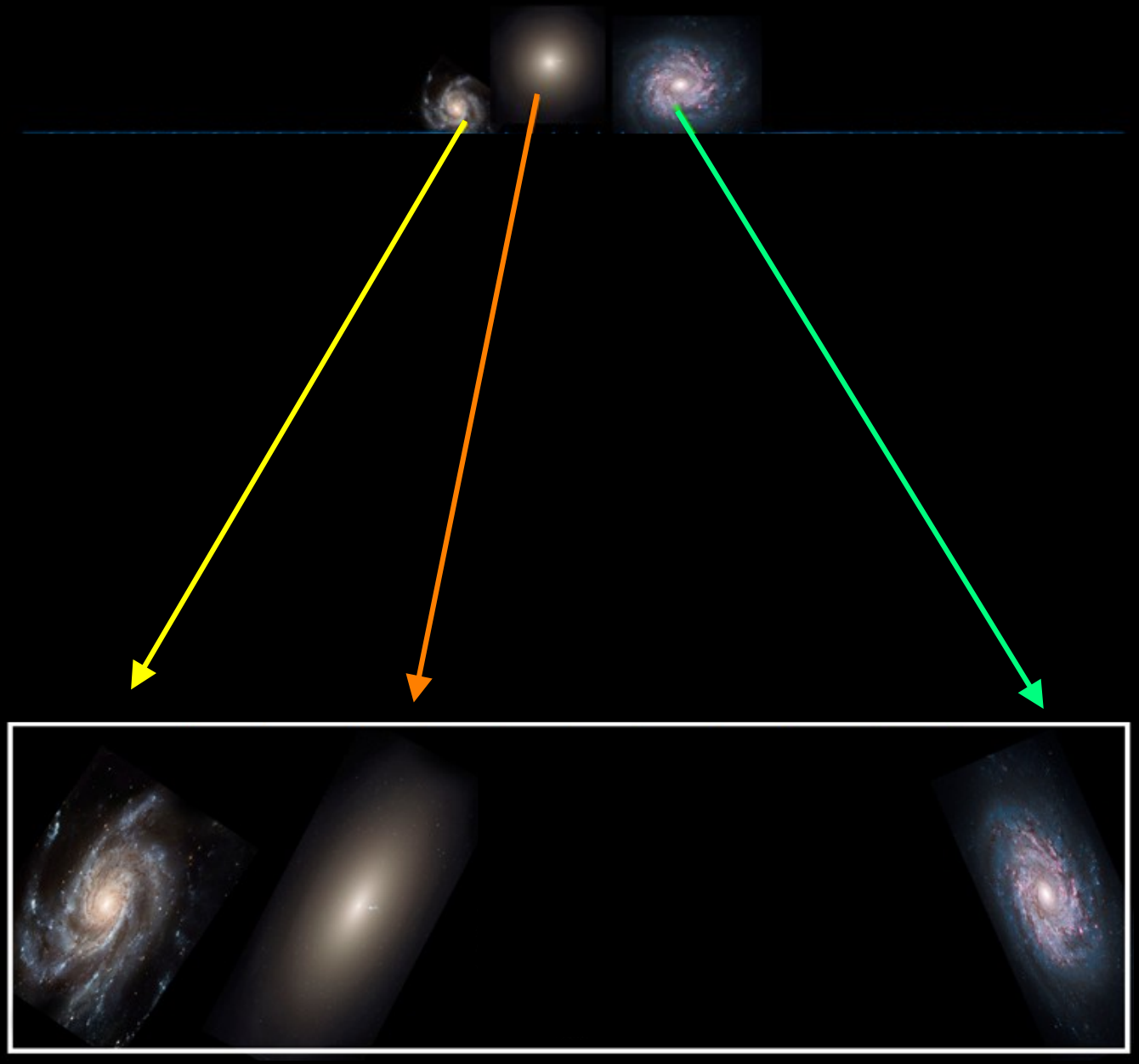


ES

Weak gravitational lensing

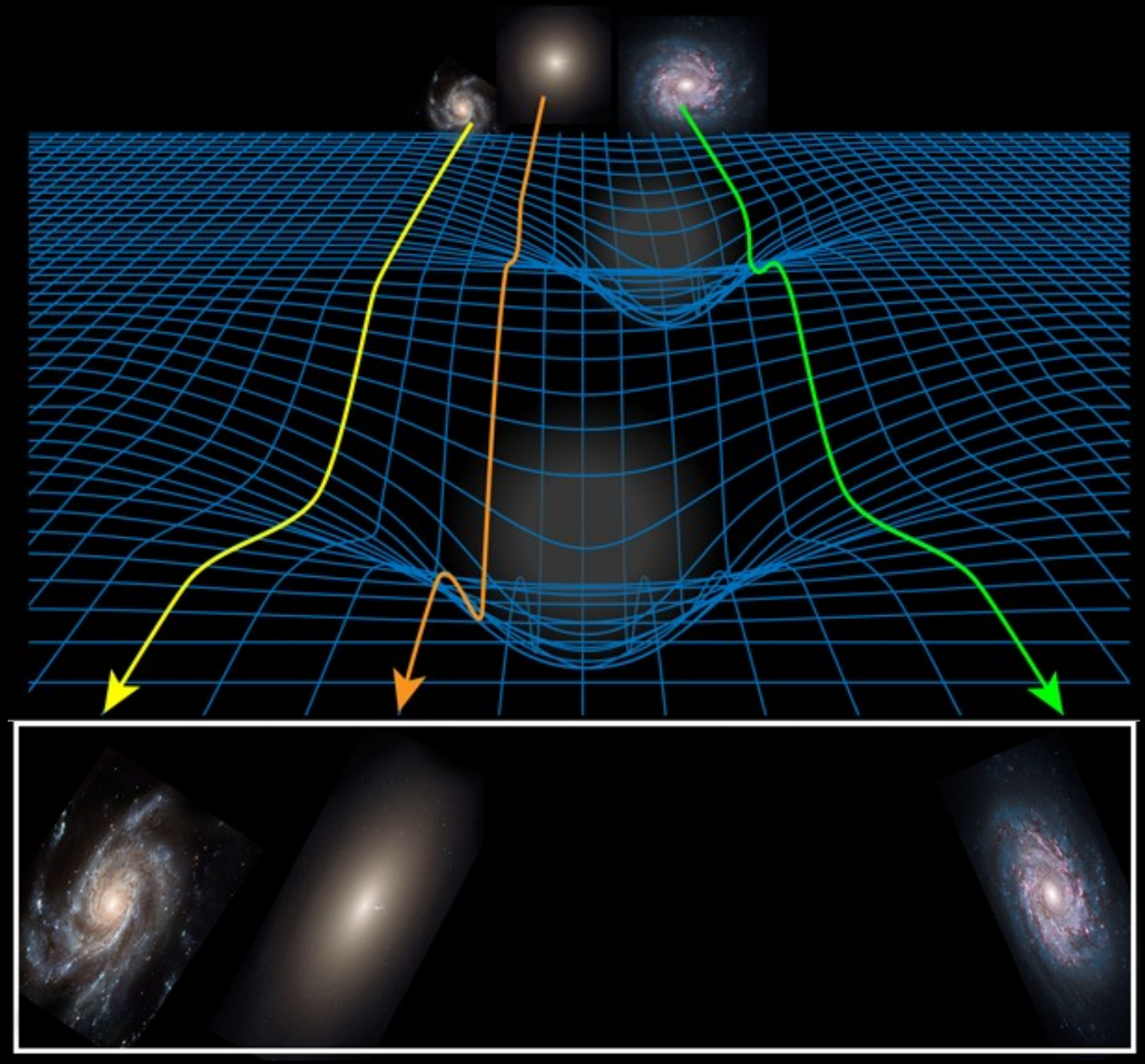
Natalia Porqueres

Weak gravitational lensing



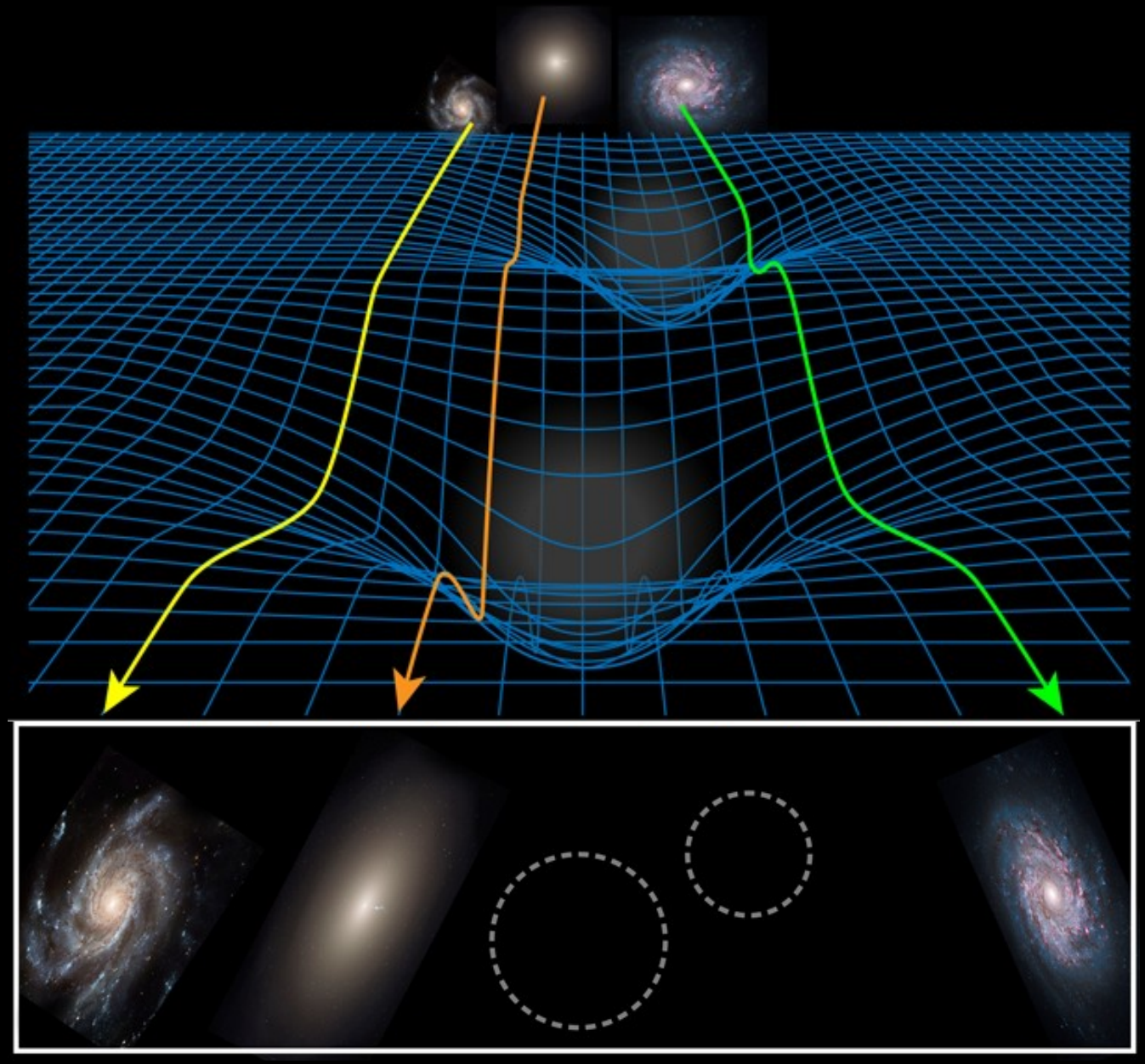
Natalia Porqueres

Weak gravitational lensing



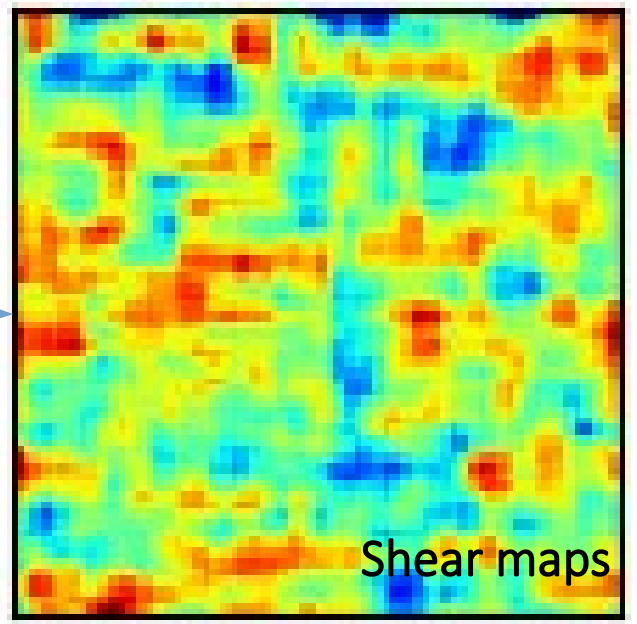
Natalia Porqueres

Weak gravitational lensing



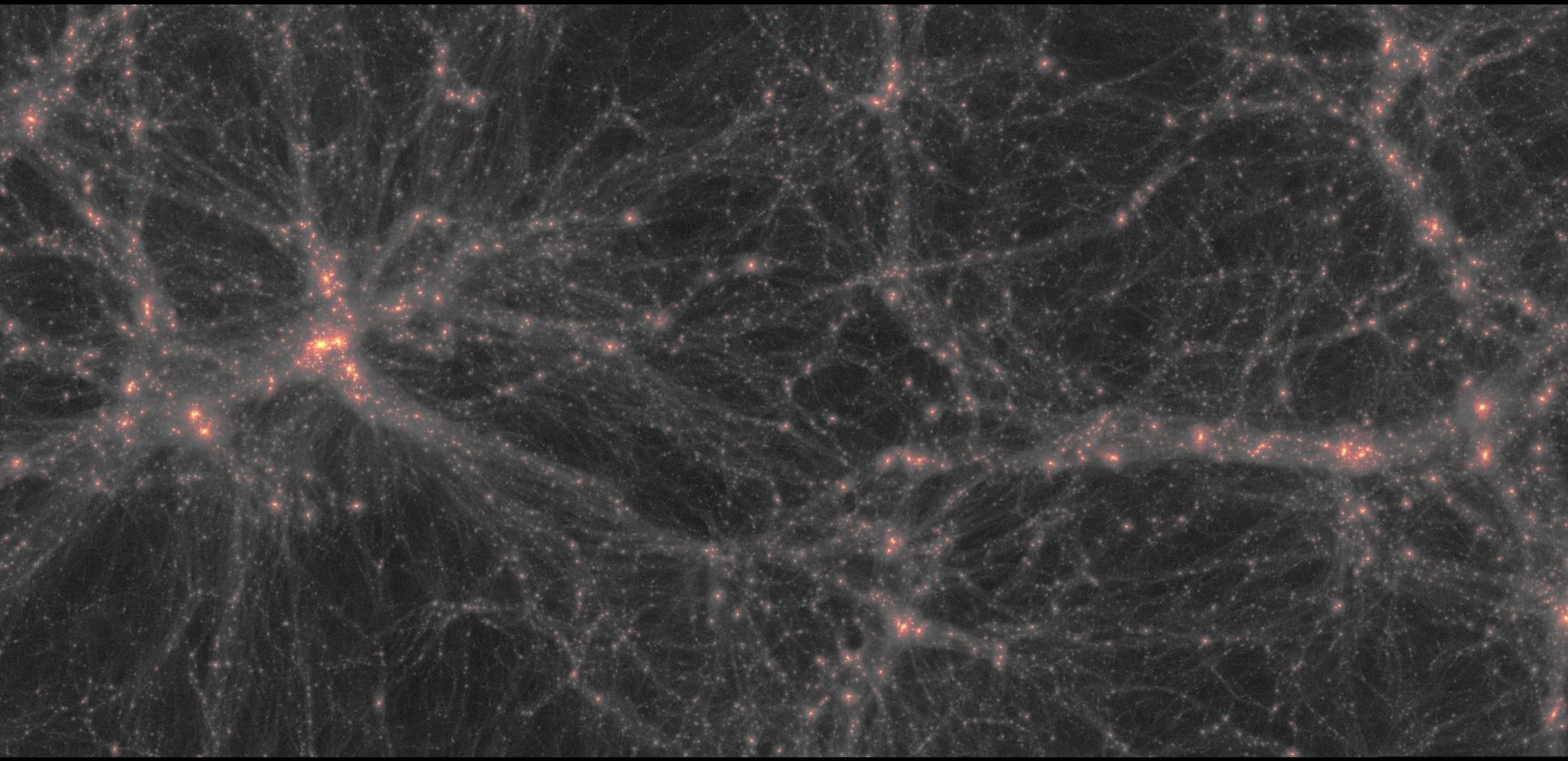
Natalia Porqueres

Building shear maps from the clean data



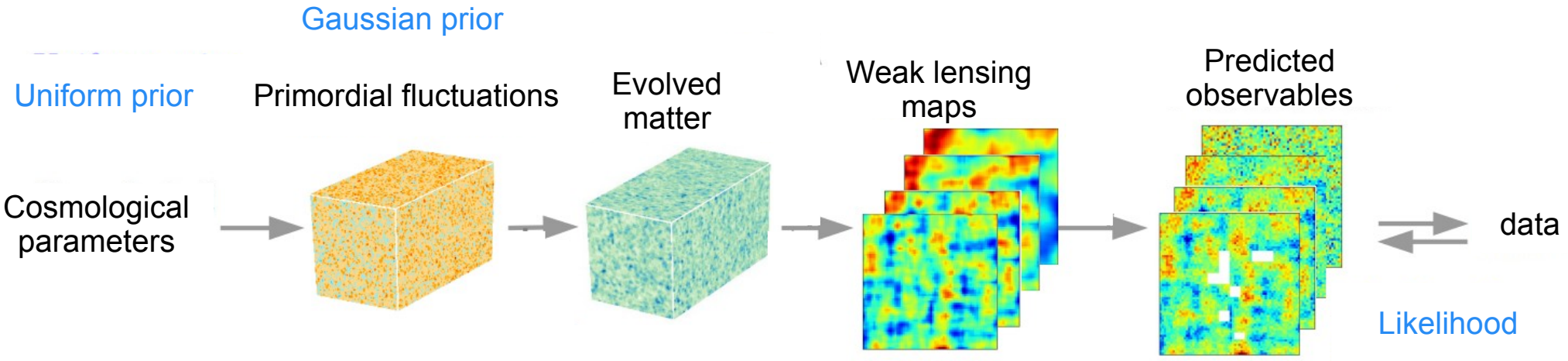
Average galaxies shape in pixels

The cosmic structures

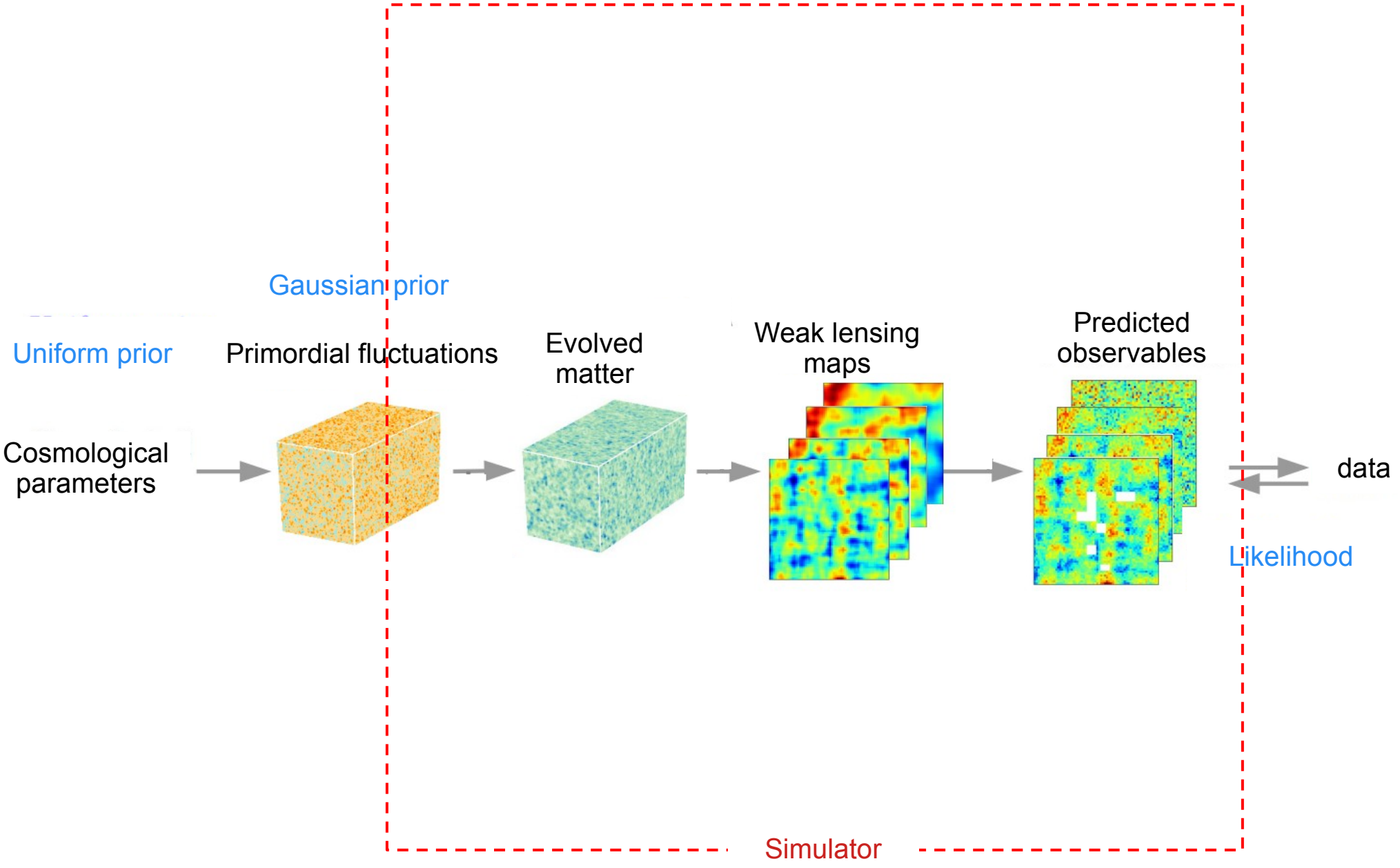


Natalia Porqueres

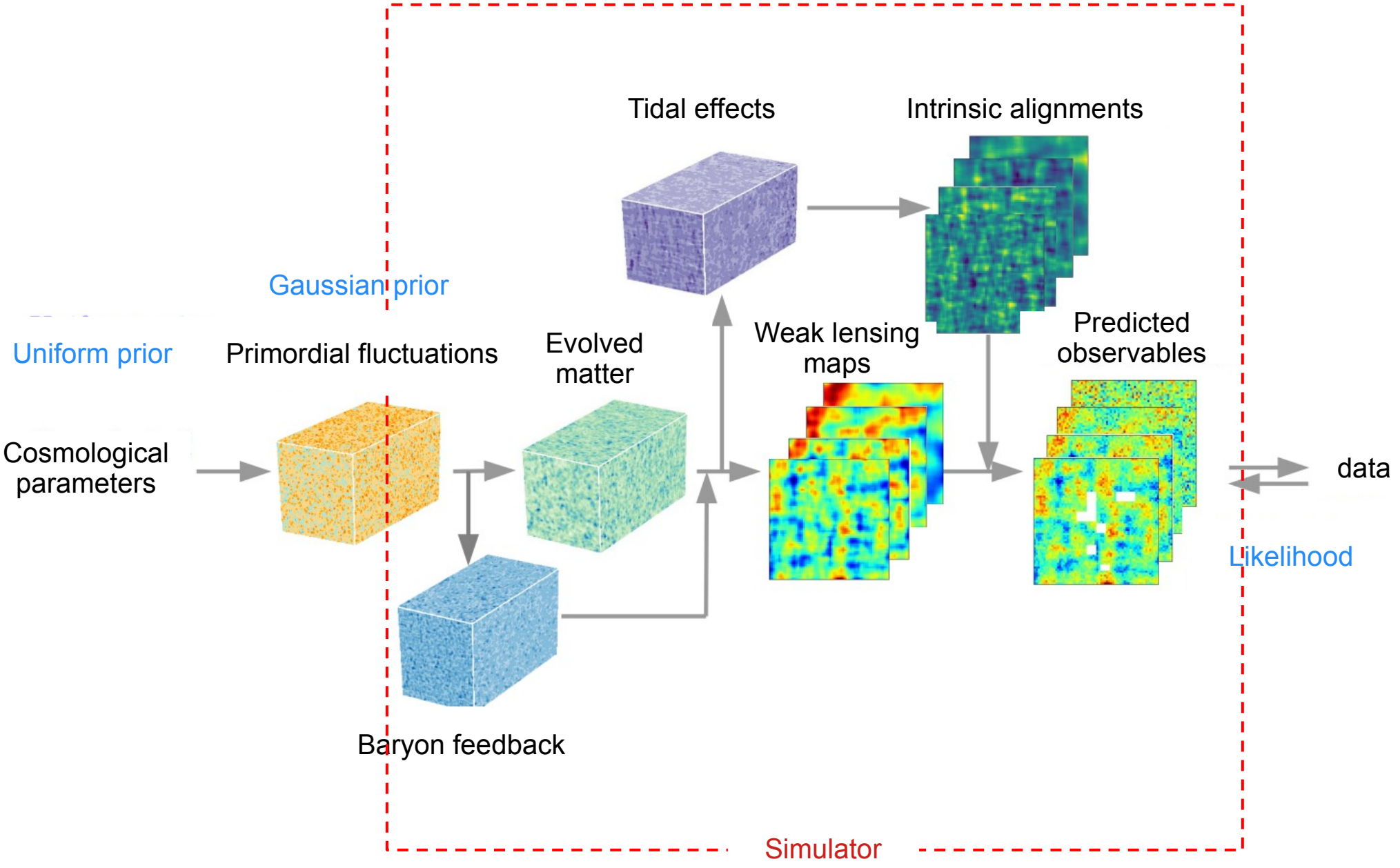
A BHM with simulations



A BHM with simulations

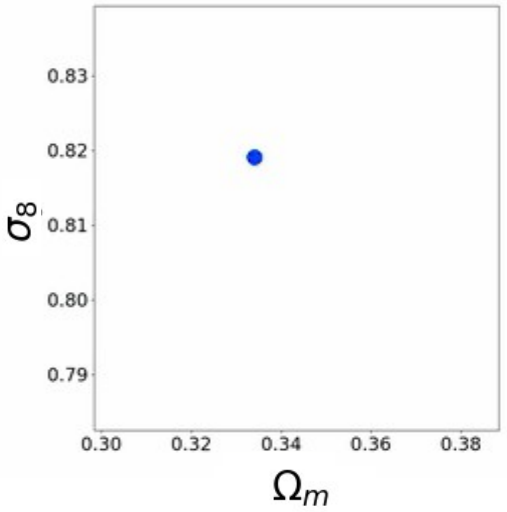


A BHM with simulations



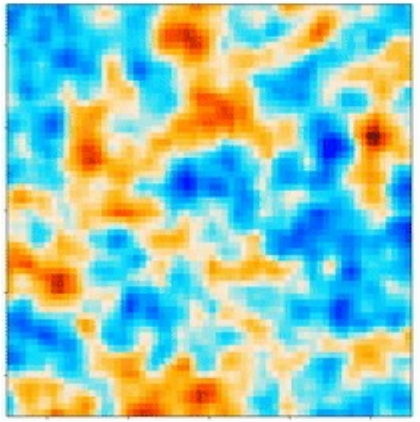
The samples

Cosmological parameters



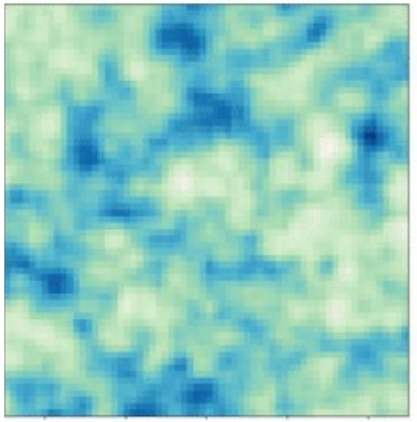
sampled

Primordial fluctuations

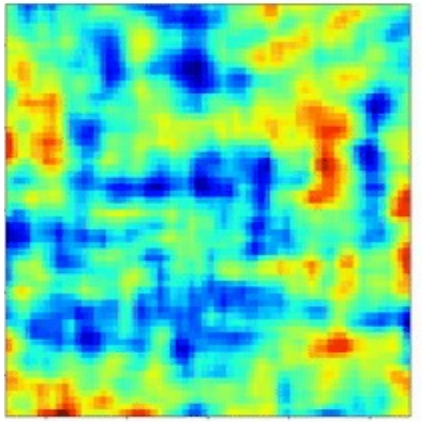


sampled

Evolved matter

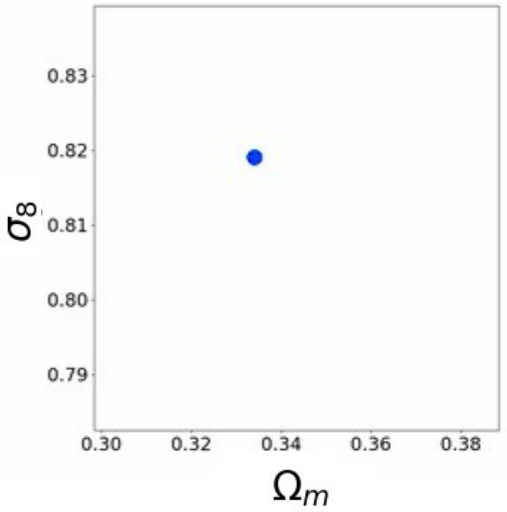


Predicted lensing



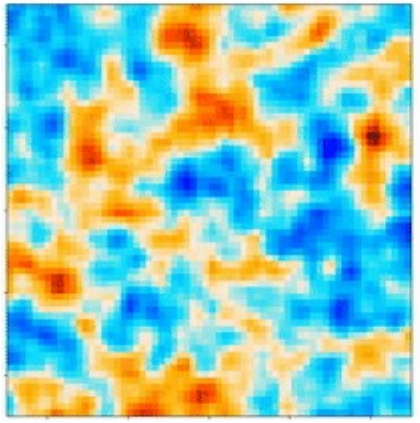
The samples

Cosmological parameters



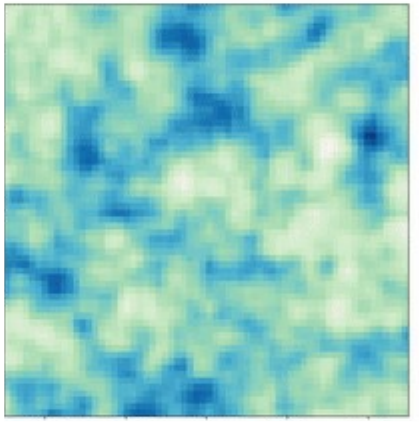
sampled

Primordial fluctuations

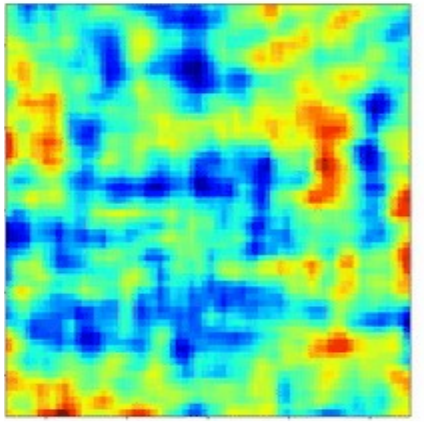


sampled

Evolved matter



Predicted lensing

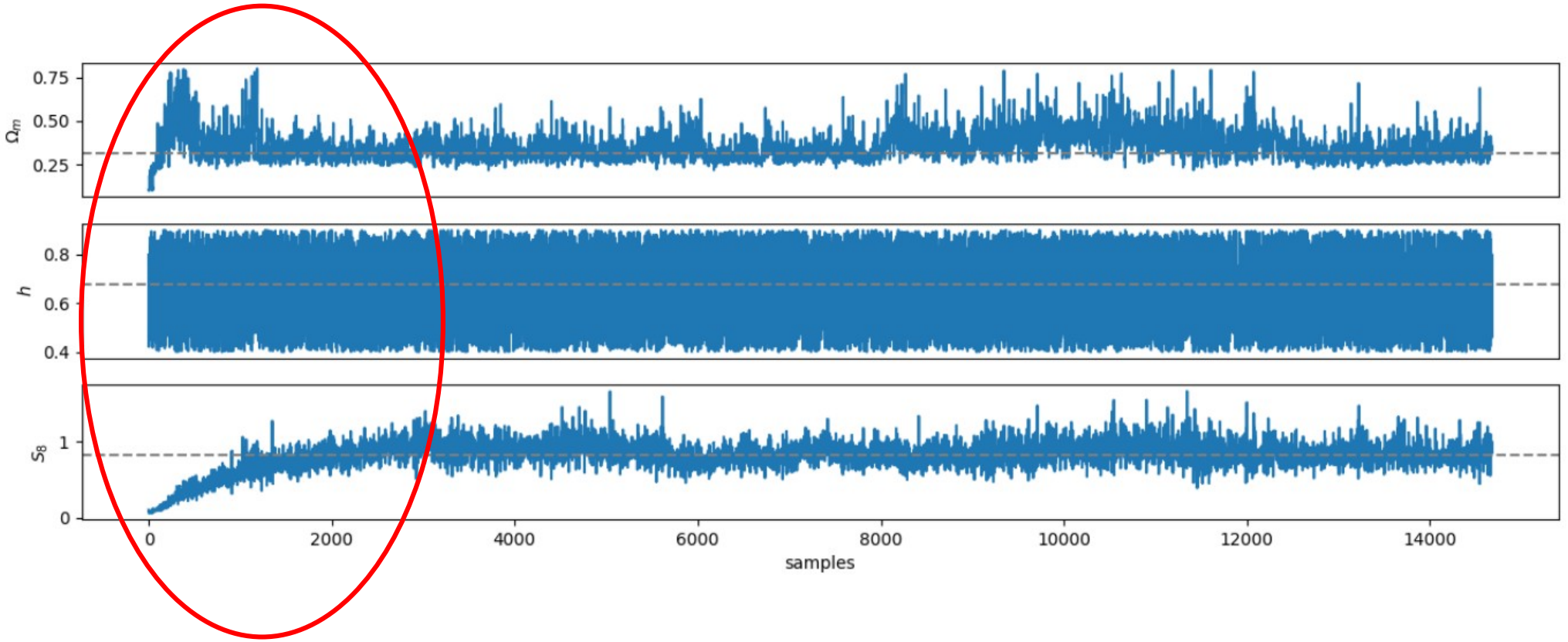


10^6 parameters!

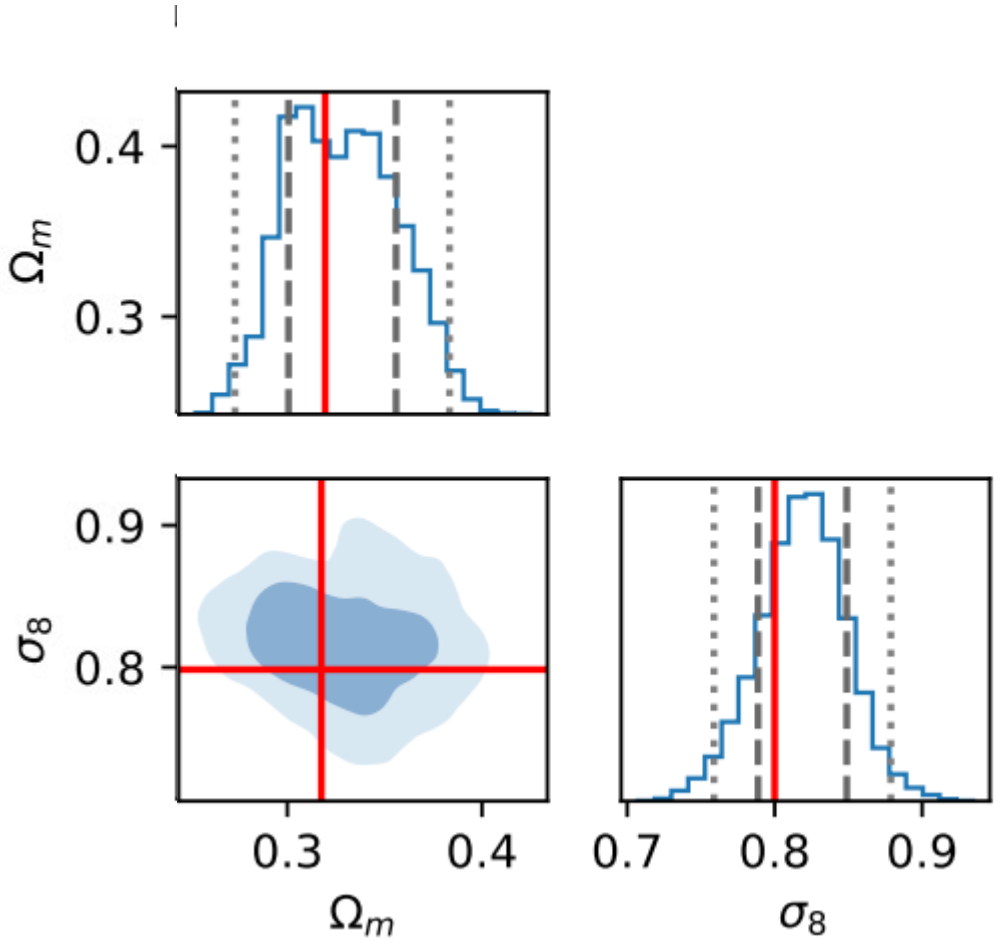
Bayesian Hierarchical Models

- Split the problem into steps, where the full data description is made of sub-models
- BHM links the sub-models and propagates the uncertainty from one step to the next
- We know the conditional probability of each step
- **BHM builds a complete model of the data and propagate the uncertainties in a principled way**
- In practice: sub-models introduce “latent” variables
- BHM can have many parameters.

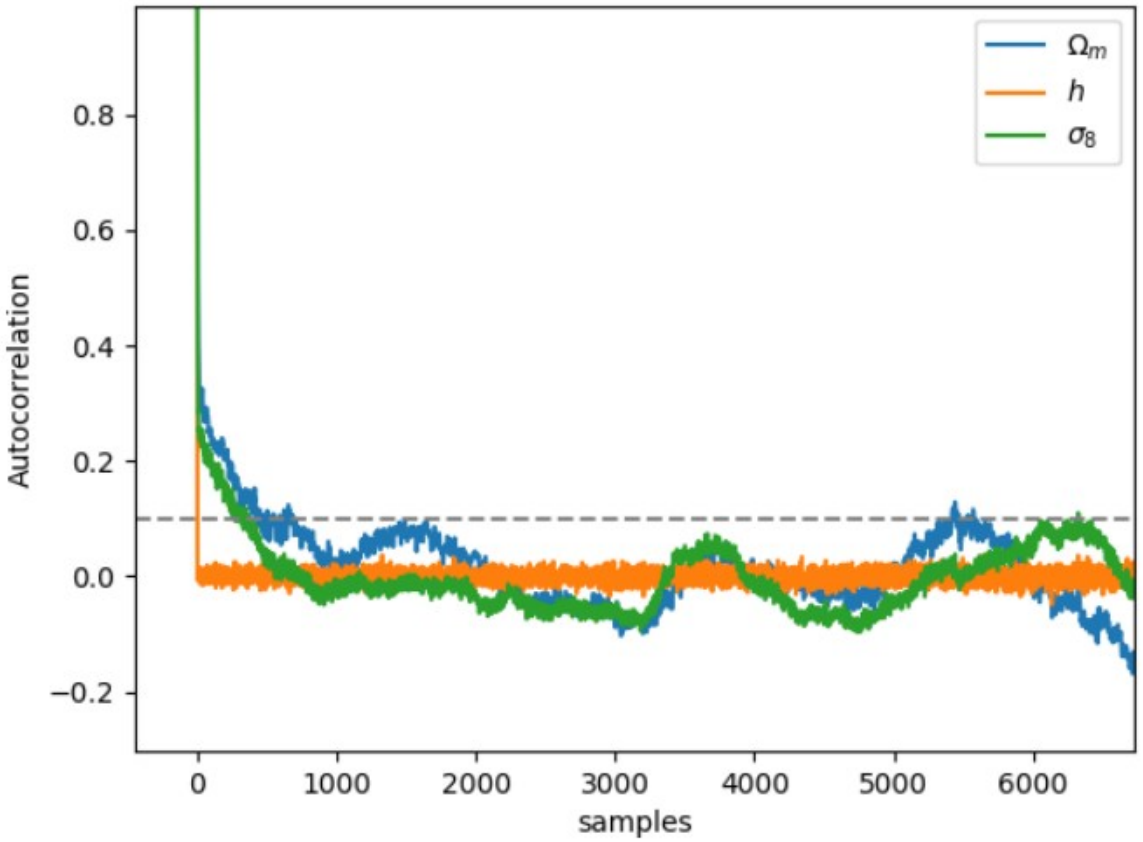
Testing results: trace plots



Testing results: Corner plots



Testing results: Correlation length



$$C_n(\theta) = \frac{1}{N-n}\sum_{i=0}^{N-n}\frac{(\theta^i - \langle\theta\rangle)(\theta^{i+n} - \langle\theta\rangle)}{\mathrm{Var}(\theta)}$$

Effective sample size:

$$N_{\mathrm{eff}} = \frac{N}{\sum_{t=-\infty}^{\infty}\rho_t} = \frac{N}{1+2\sum_{t=1}^{\infty}\rho_t}.$$

Testing results: Convergence tests

Gelman-Rubin test:

Run M chains with different starting point and N samples.

Calculate the mean and std in each chain: $\hat{\theta}_m = \frac{1}{N_m} \sum_i^{N_m} \theta_i^m$

Calculate the meand and std of all chains: $\hat{\theta} = \frac{1}{M} \sum_m^M \hat{\theta}_m$

Compute how the means vary around the joint

$$B = \frac{N}{M-1} \sum_{m=1}^{M} (\hat{\theta}_m - \hat{\theta})^2$$

Compute the average variance of the chains

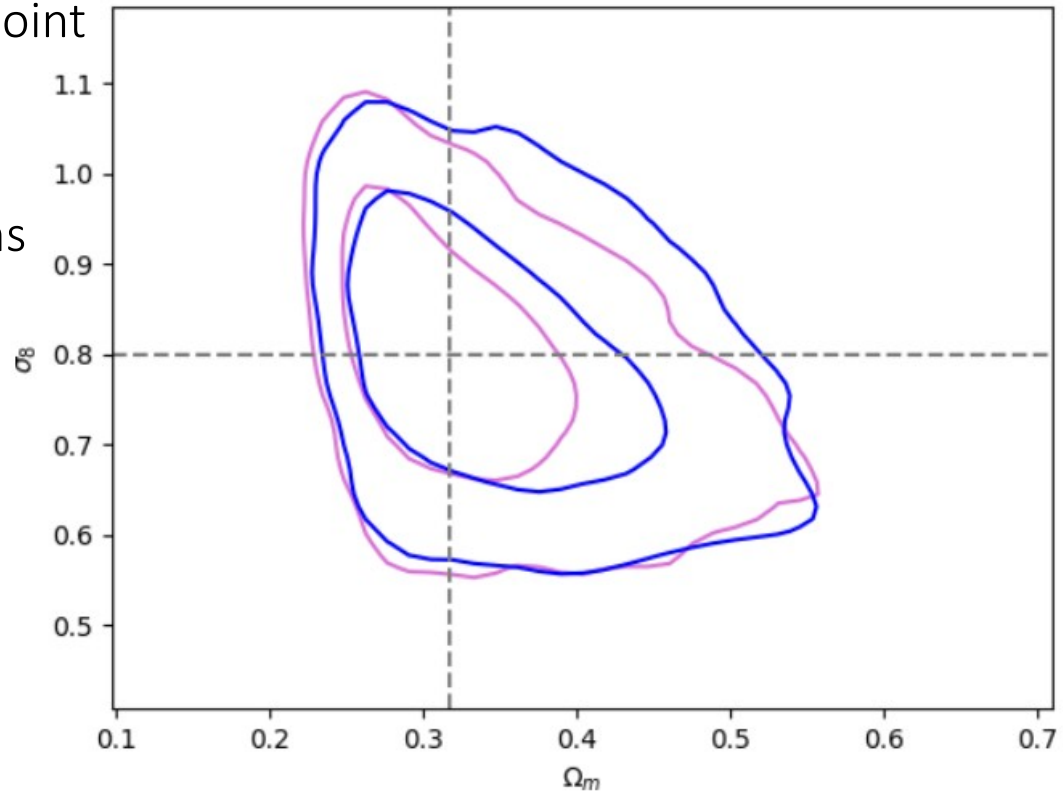
$$W = \frac{1}{M} \sum_{m=1}^{M} \sigma_m^2$$

Compute unbiased estimator of the true variance

$$\hat{V} = \frac{N-1}{N} W + \frac{M+1}{MN} B$$

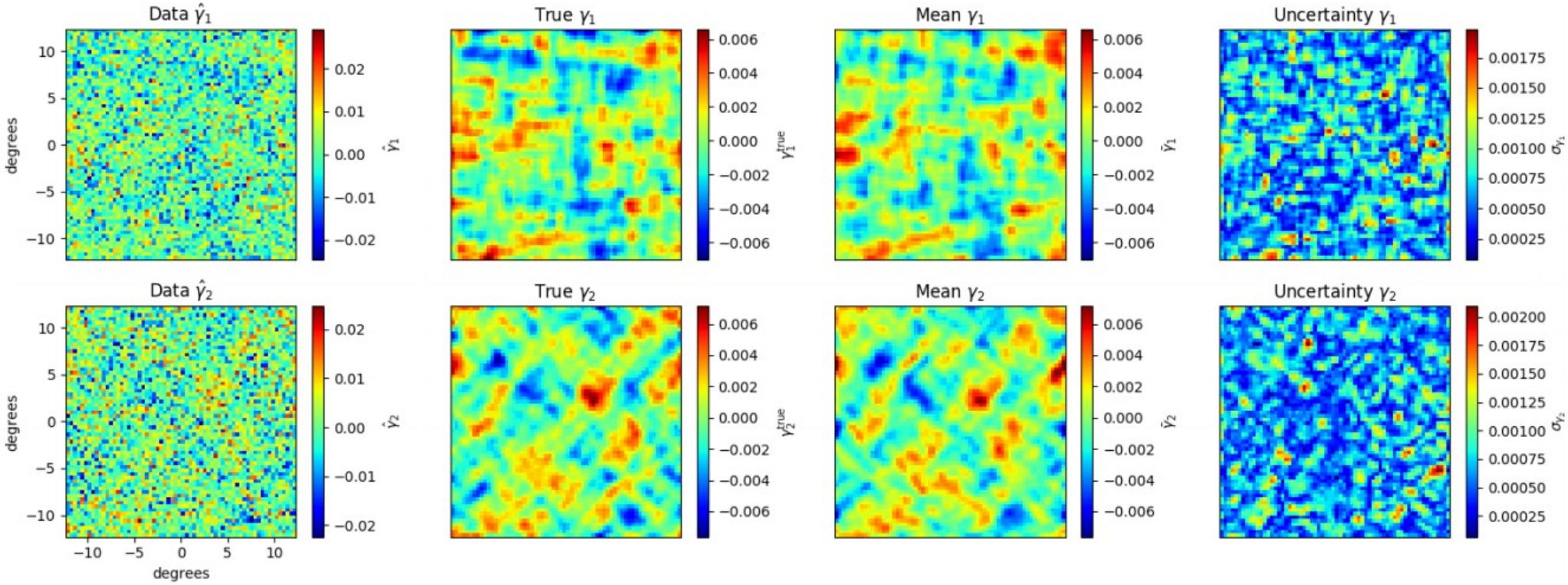
The chain has converged if

$$R = \sqrt{\hat{V}/W} \approx 1$$



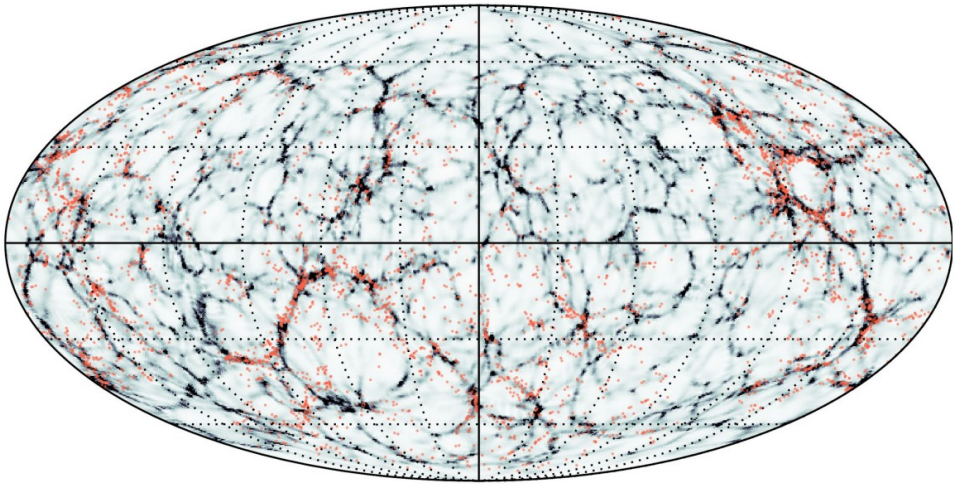
Posterior predictive tests

Posterior predictive tests: Can the inferred quantities explain the data?



Validating the results of a real data analysis

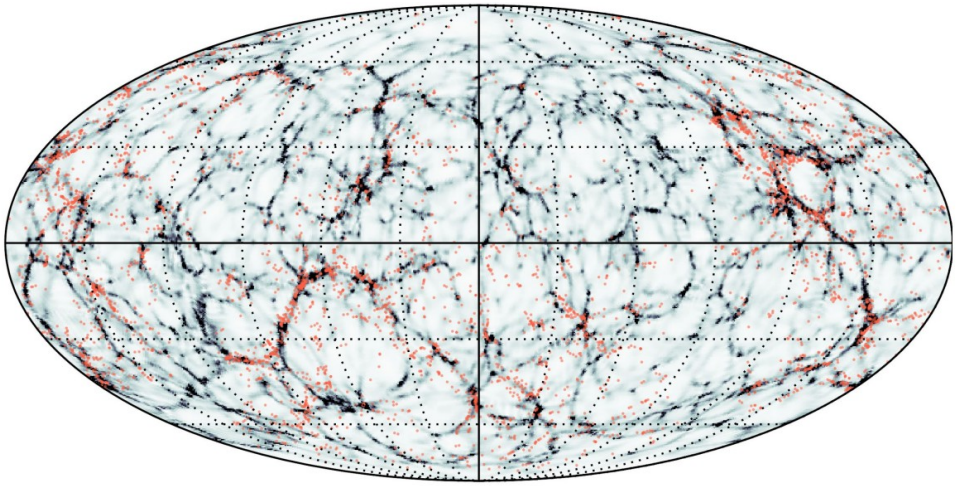
Posterior predictive tests and cross-validation with independent measurements



Validating the results of a real data analysis

Posterior predictive tests and cross-validation with independent measurements

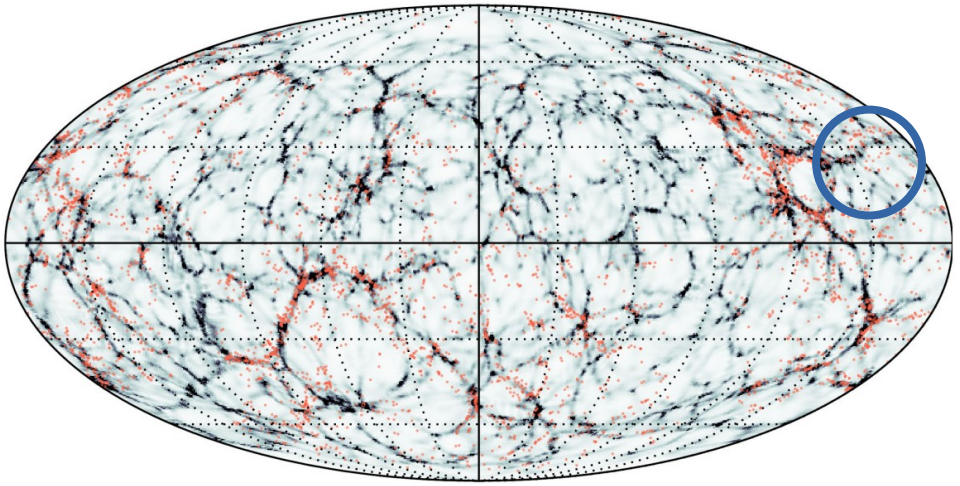
- do we get the clusters we know are there?



Validating the results of a real data analysis

Posterior predictive tests and cross-validation with independent measurements

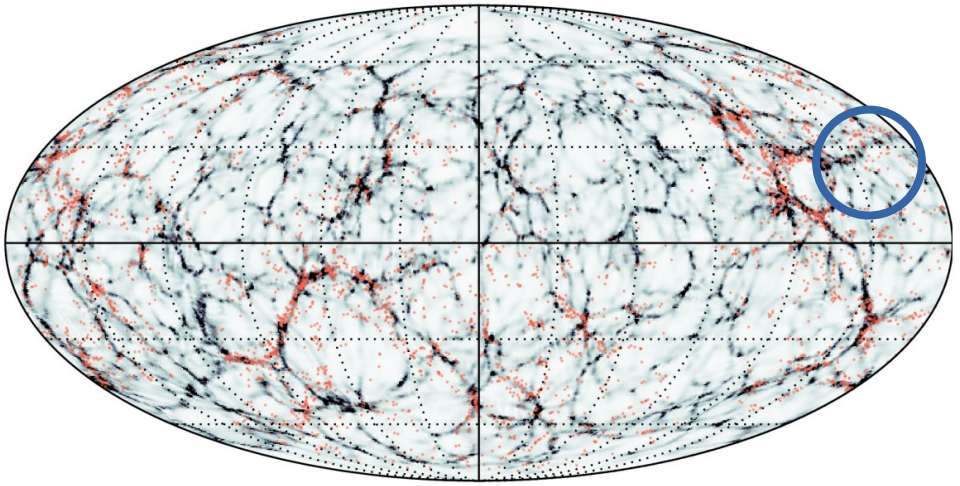
- do we get the clusters we know are there?

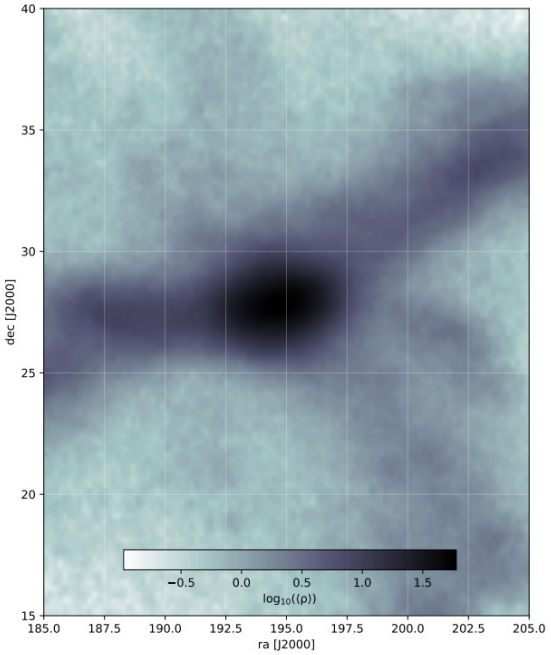


Validating the results of a real data analysis

Posterior predictive tests and cross-validation with independent measurements

- do we get the clusters we know are there?





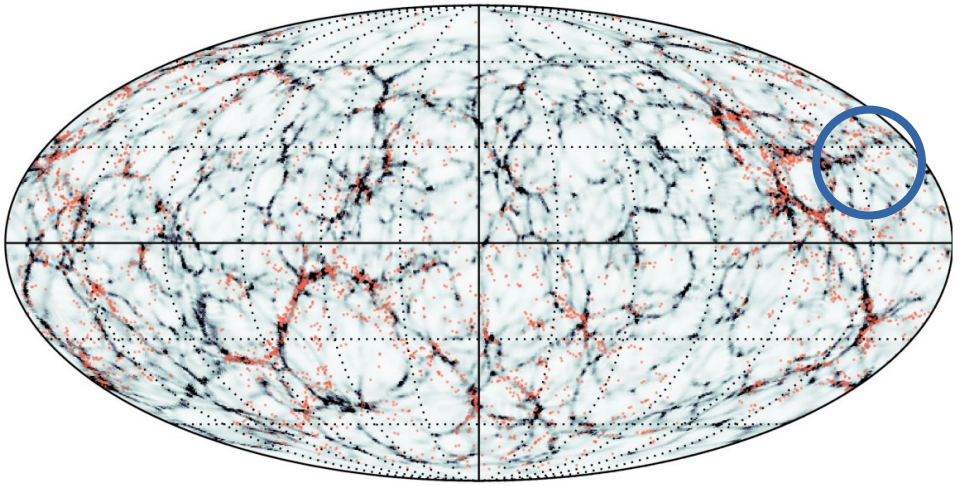
Jasche & Lavaux 2019

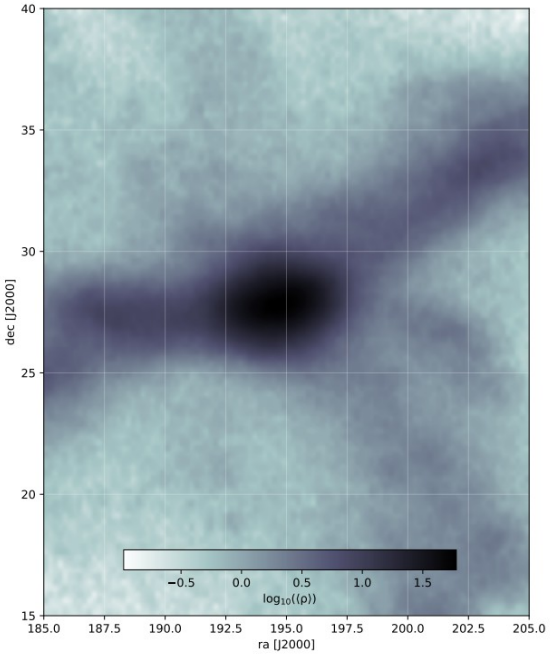
Validating the results of a real data analysis

Posterior predictive tests and cross-validation with independent measurements

- do we get the clusters we know are there?

- do we get the expected mass profiles?





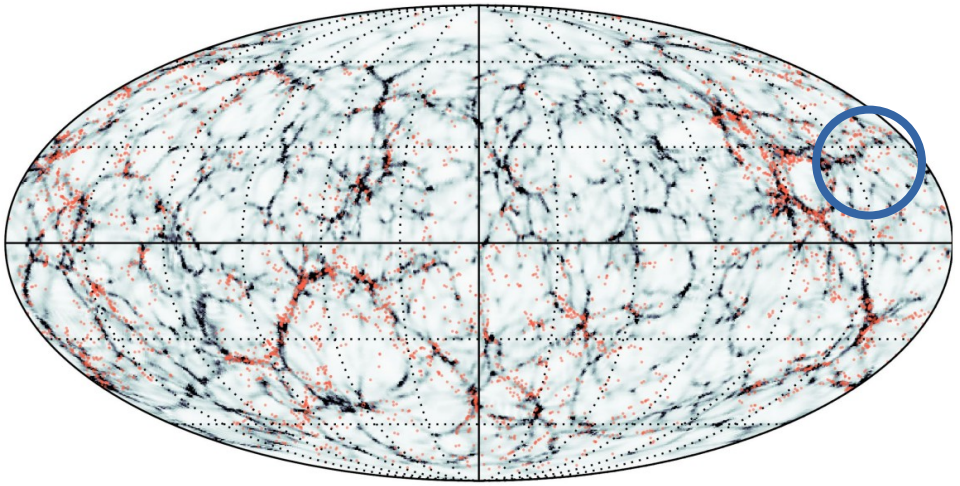
Jasche & Lavaux 2019

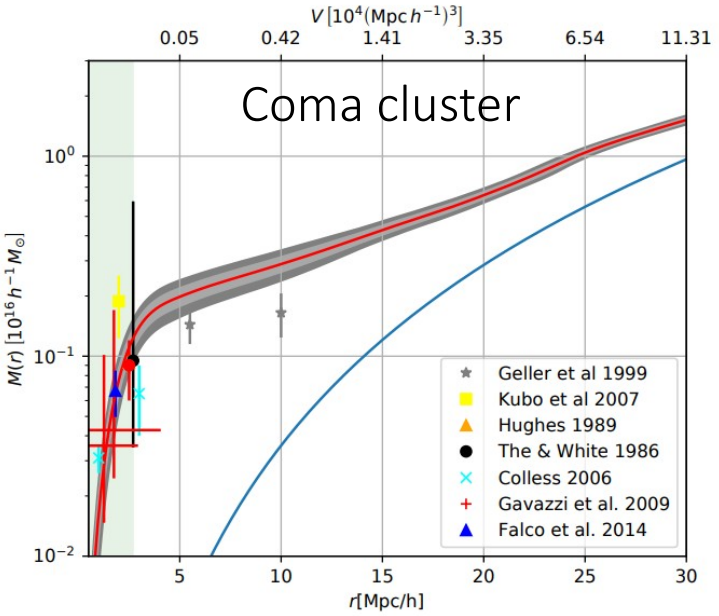
Validating the results of a real data analysis

Posterior predictive tests and cross-validation with independent measurements

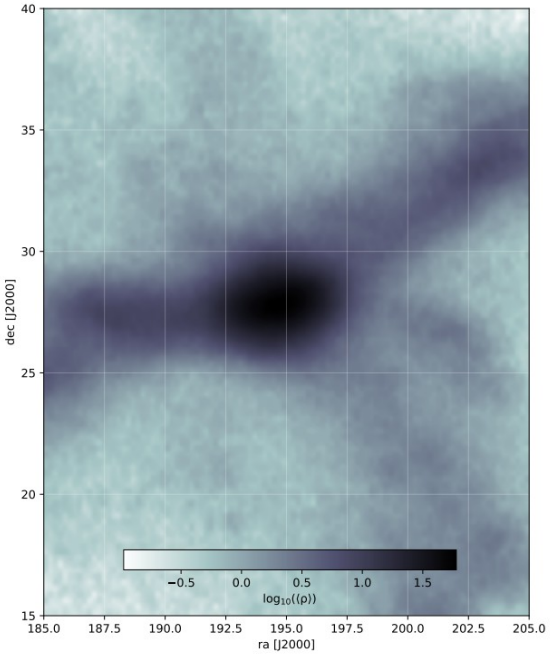
- do we get the clusters we know are there?

- do we get the expected mass profiles?





Jasche & Lavaux 2019



Jasche & Lavaux 2019

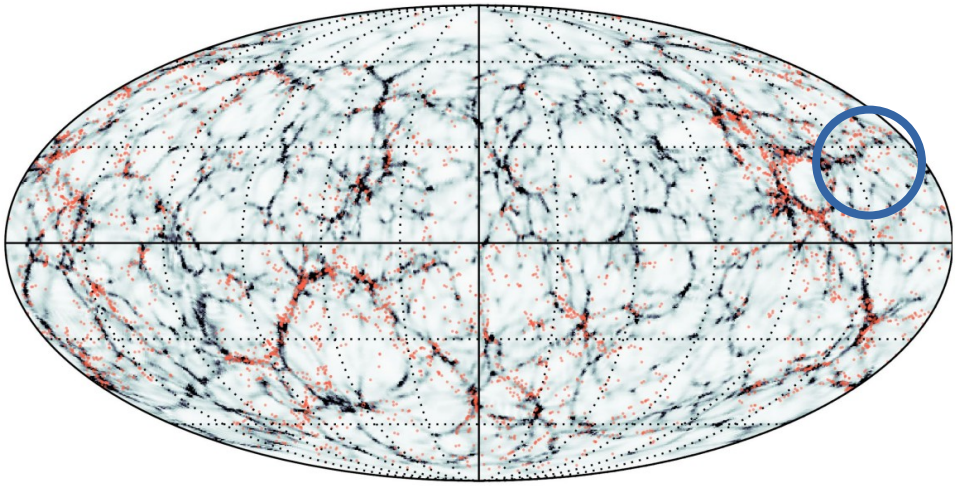
Validating the results of a real data analysis

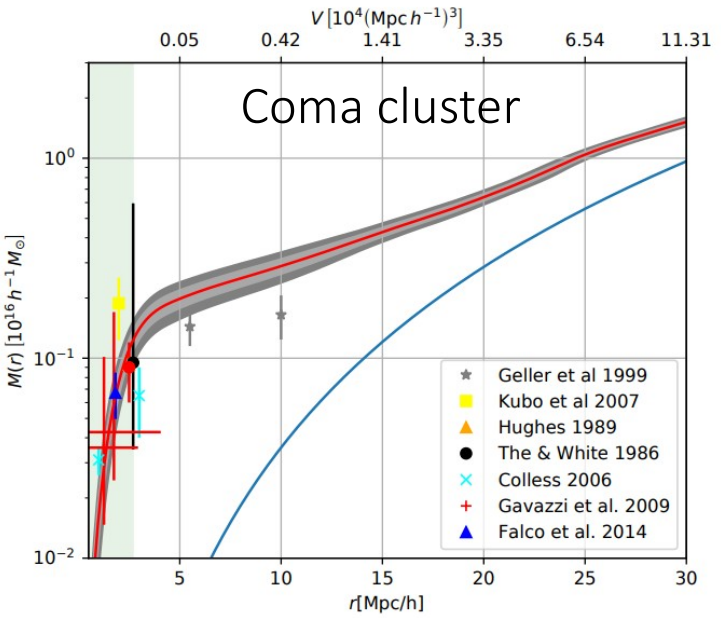
Posterior predictive tests and cross-validation with independent measurements

- do we get the clusters we know are there?

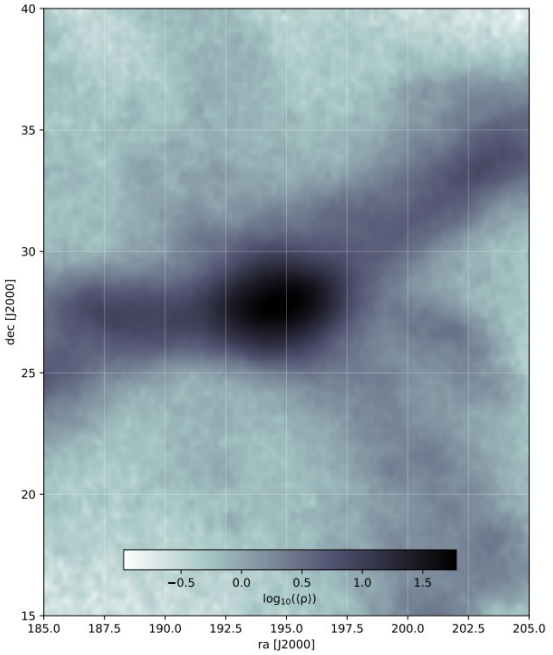
- do we get the expected mass profiles?

- are the inferred IC compatible with CMB?





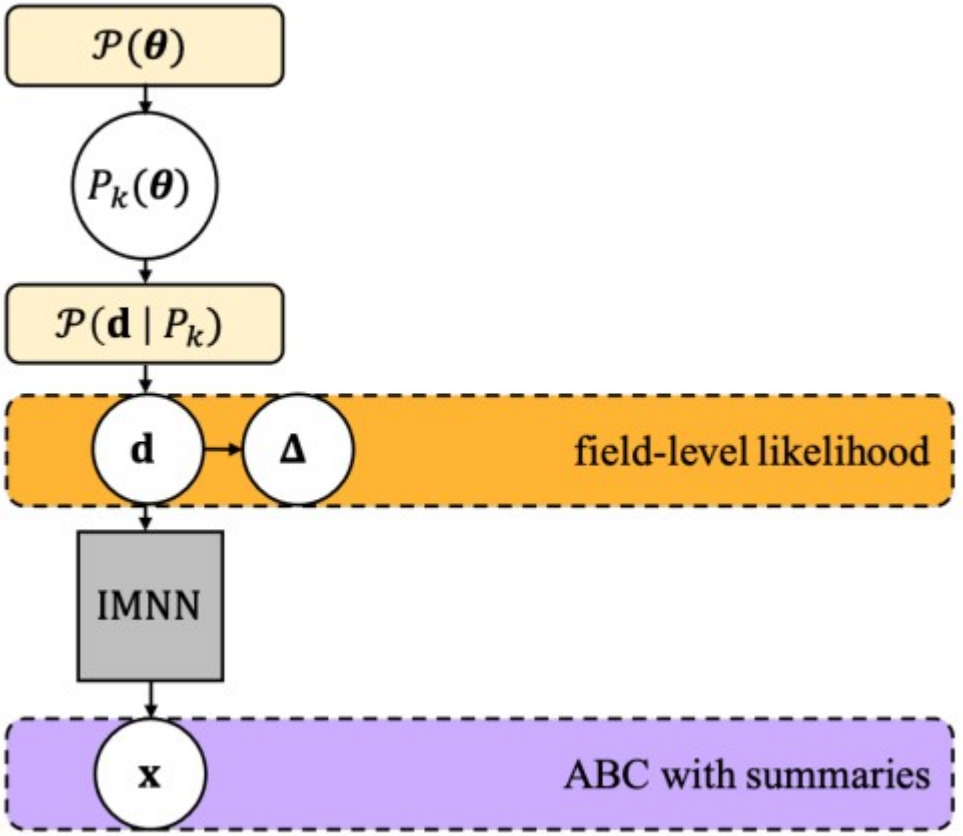
Jasche & Lavaux 2019



Jasche & Lavaux 2019

Natalia Porqueres

Can we find an alternative with the same info?



Makinen+ 2021

Two approaches to inference with simulators

Bayesian Hierarchical Model (aka explicit inference)

Treat the simulator as a probabilitic model to sample full posterior distribution

(e.g. Jasche+2013, Huiyuan+2014, Seljak+2017, Modi+2018, Schmidt+2019, Porqueres+2022, Boruah+2022)

Simulation-based inference (aka implicit inference, likelihood-free inference)

Treat the simulator as a black box to draw random samples

(e.g. Fluri+2018, Charnock+2018, Makinen+2021, Gupta+2019, Jeffrey+2020, Leclercq2022, Dai+2022)