Lifting weak lensing degeneracies with a field-based likelihood

Natalia Porqueres

with Alan Heavens, Daniel Mortlock, Guilhem Lavaux, Lucas Makinen

Bucaramanga – 27/10/2023

Outline

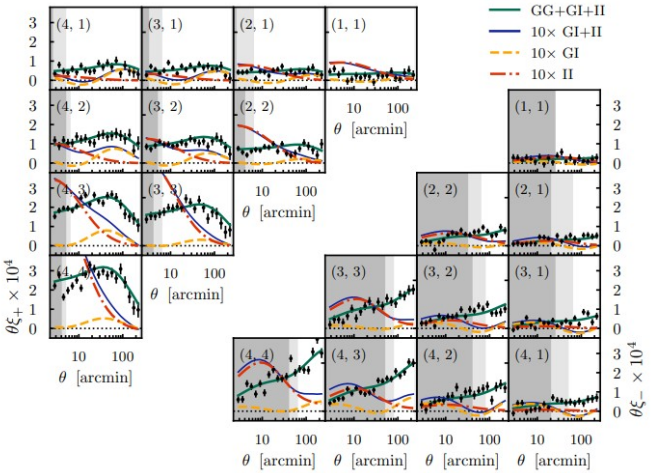
The method A field-based approach of weak lensing

The proof of concept Results with synthetic data

The challenge Getting ready for real data

The standard approach to shear analysis

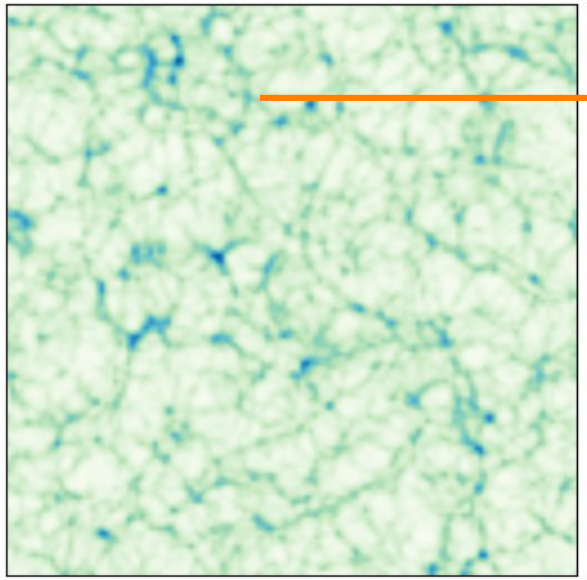
Summarise



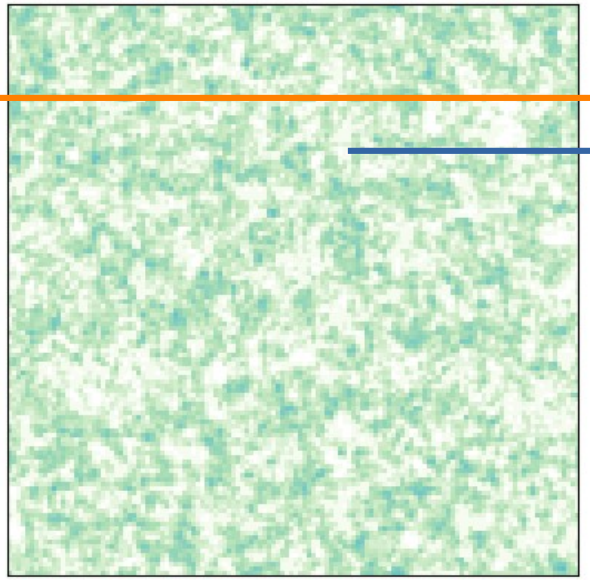
Does this capture all the information?

Is the standard approach enough?

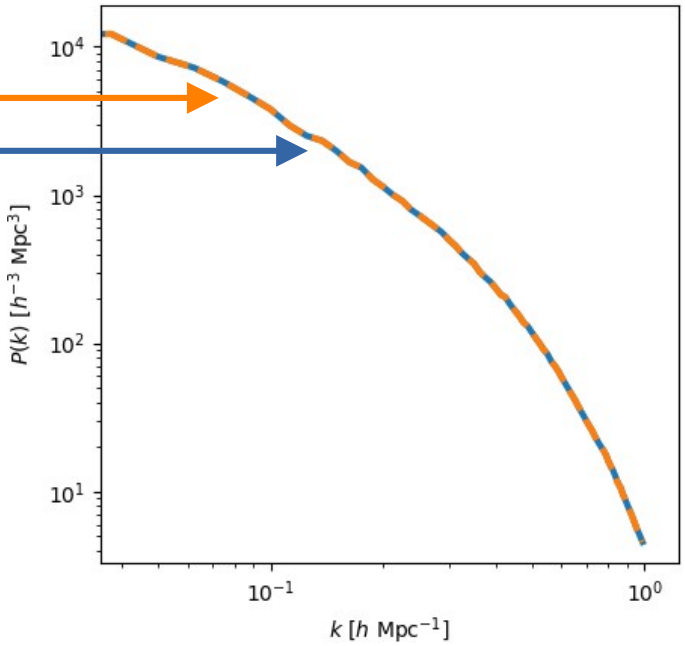
The standard approach misses information!



Cosmic web



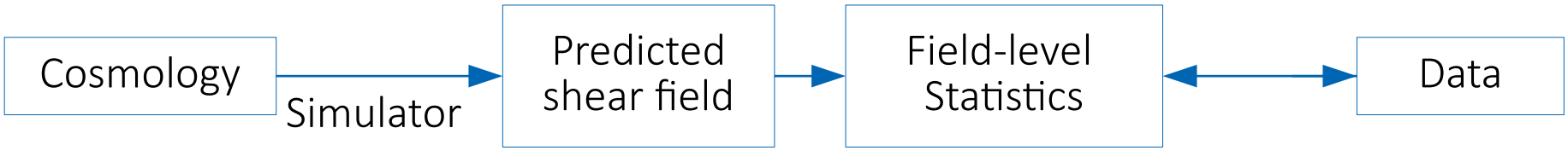
Gaussian random field



- Higher-order statistics: what are the sampling distributions? And covariance matrix?

- Field-based approach without data compression of the pixelised shear.

The field-level approach



Statistics pixel by pixel can **access all the information**

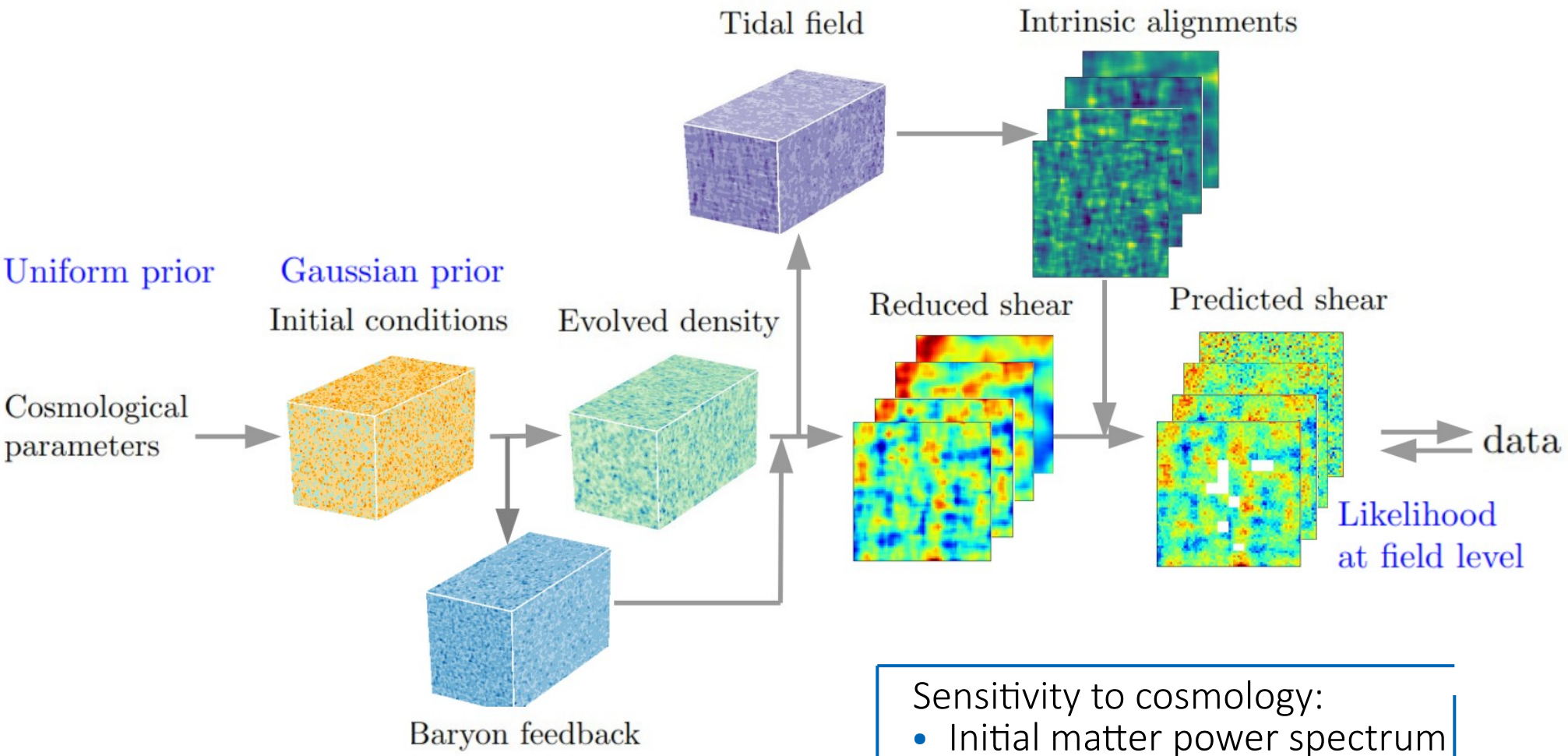
Two approaches:

- **explicit inference** (aka Bayesian hierarchical model)
  (e.g. Jasche+2013, Huiyuan+2014, Seljak+2017, Modi+2018, Schmidt+2019, Porqueres+2022, Boruah+2022)
  Treat the simulator as a probabilitic model to sample full posterior distribution
  Simulations are constrained by the data.

- **implicit inference** (aka simulation-based inference, likelihood-free inference)
  (e.g. Fluri+2018, Charnock+2018, Makinen+2021, Gupta+2019, Jeffrey+2020, Leclercq2022, Dai+2022)
  Treat the simulator as a black box to draw random samples
  Simulations are random.

BORG framework: **explicit inference** with a **gravity model**

The forward model of BORG-WL

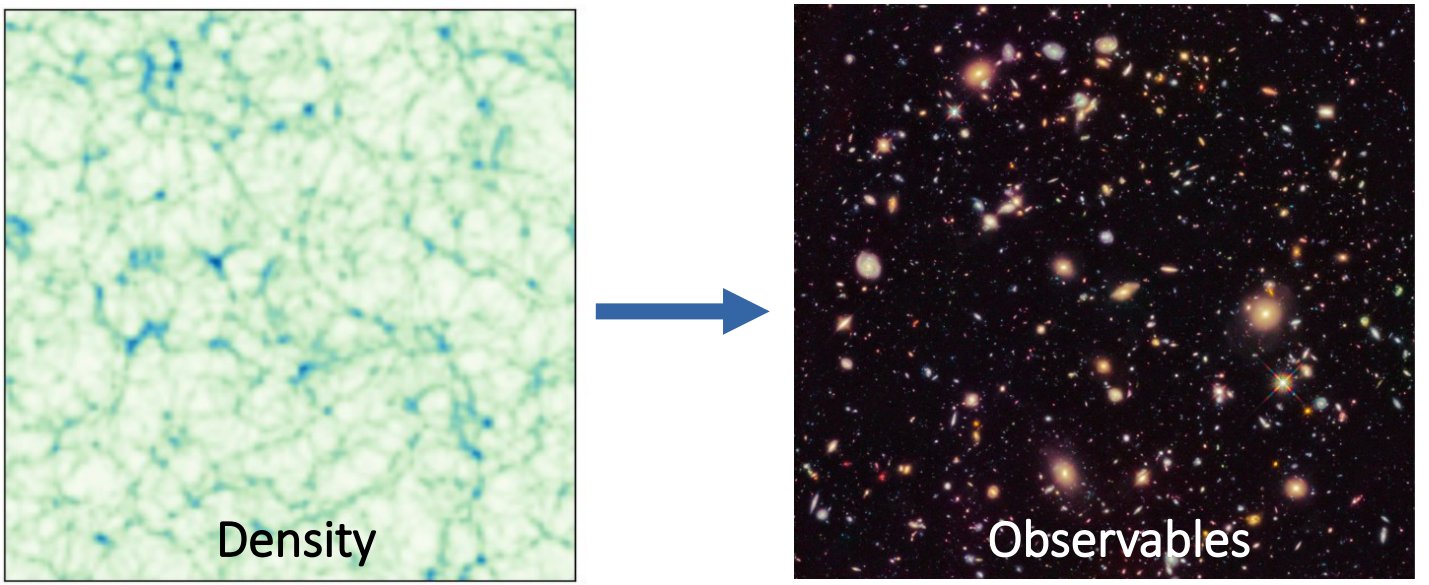


Sensitivity to cosmology:

- Initial matter power spectrum
- Growth of structures
- Geometry

Intrinsic alignments: Tidal alignment and tidal torquing

Blazek+ 2017



TATT model includes linear (tidal alignment) and quadratic (tidal torque) terms

$$\gamma_1^{\rm IA}(\boldsymbol{r},\boldsymbol{\vartheta}) = (C_1 + C_{1\delta}\delta)(s_{xx} - syy) + C_2(s_{xk}s_{xk} - s_{yk}s_{yk})$$
$$\gamma_2^{\rm IA}(\boldsymbol{r},\boldsymbol{\vartheta}) = 2(C_1 + C_{1\delta}\delta)s_{xy} + 2C_2 s_{xk}s_{yk},$$

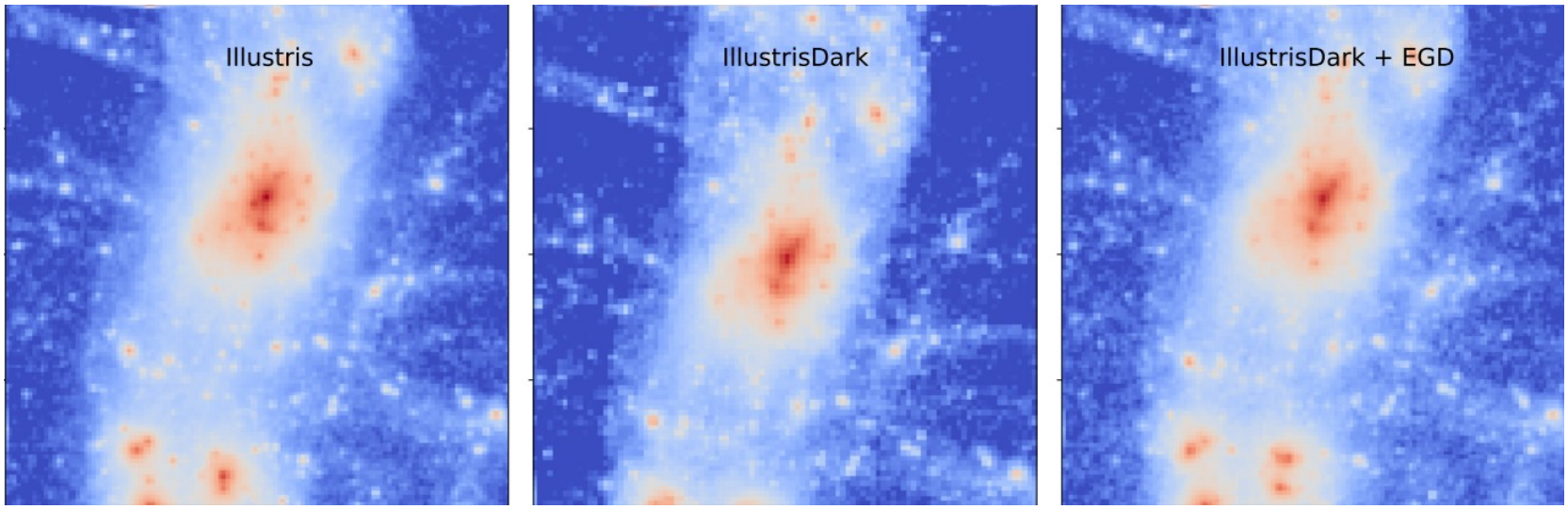
Baryon feedback: Enthalpy gradient descent model

Dai & Seljiak, 2018

Displaces dark matter particles with

$$\mathbf{S}_{\text{baryons}} = -\frac{\beta}{H_0^2}\frac{k_B T_0}{\mu}\frac{\gamma}{\gamma-1}\nabla\left[\mathbf{O}_{\text{J}}(1+\delta)\right]^{\gamma-1}$$

Assumes power law equation of state of baryons $T(\delta) = T_0(1+\delta_b)^{\gamma-1}$



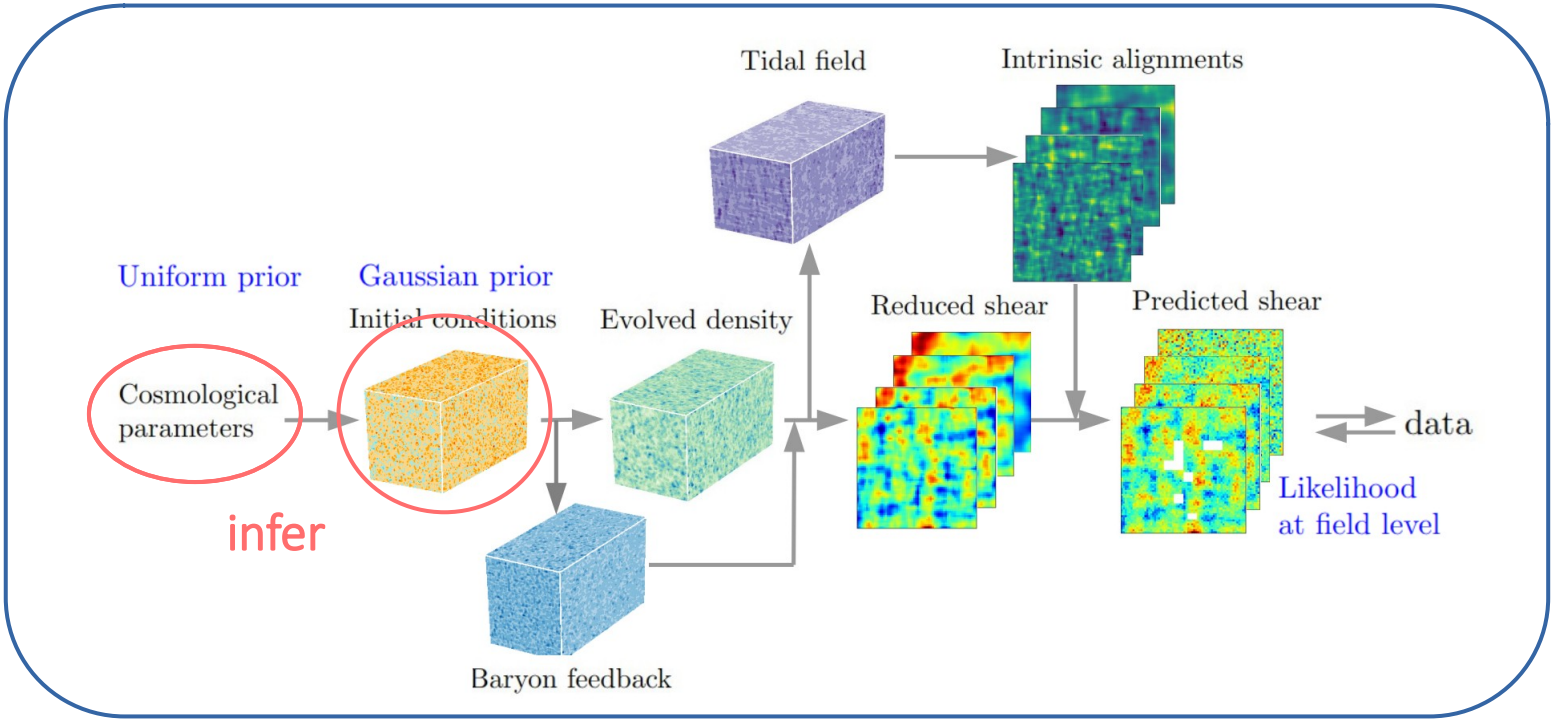
Dai & Seljiak, 2018

- Self calibration of model parameters.

- Displaces particles independently of whether they are in a halo.

- Fully differentiable.

The likelihood



Likelihood: Assuming Gaussian noise in pixelised shear

$$\log \mathcal{L} = -\frac{1}{2} \sum_b \sum_{mn} \frac{(\epsilon^b_{1,mn} - \hat{\epsilon}^b_{1,mn})^2 + (\epsilon^b_{2,mn} - \hat{\epsilon}^b_{2,mn})^2}{\sigma_b^2}$$

Doesn't need covariance matrix! $\sigma_b = \sigma_\epsilon / \sqrt{N_b}$

Getting the full posterior distribution

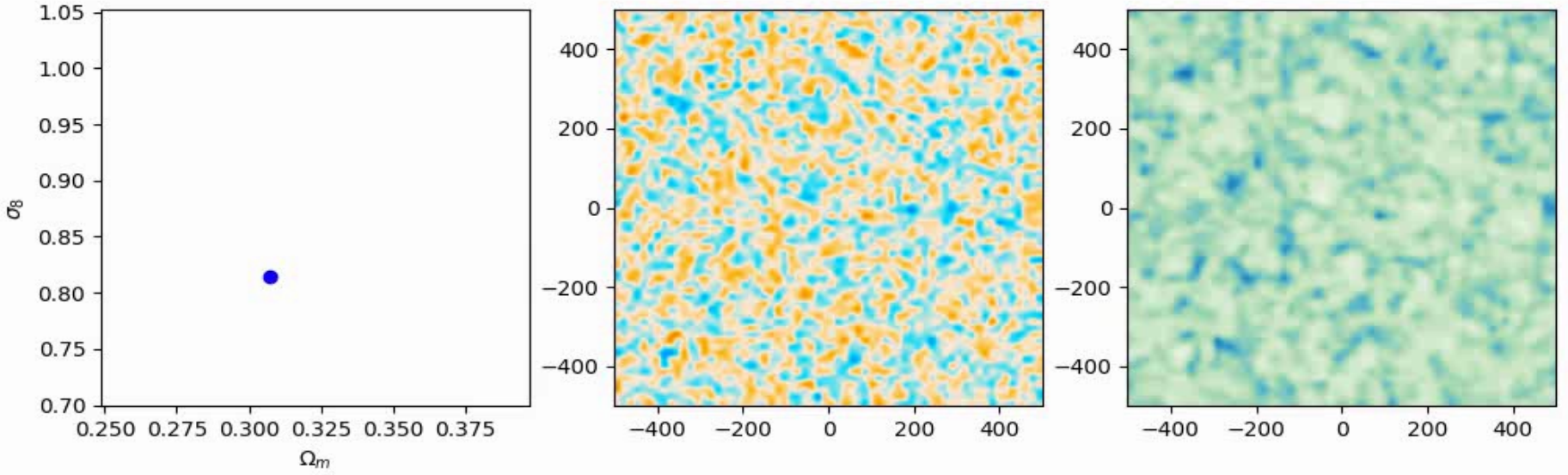
$$P(\theta, \delta^{\rm ic}|\hat{\epsilon}) \propto P(\hat{\epsilon}|\epsilon) P(\epsilon|\delta^{\rm ic}) P(\delta^{\rm ic}|\theta) P(\theta)$$

posterior — likelihood — prior

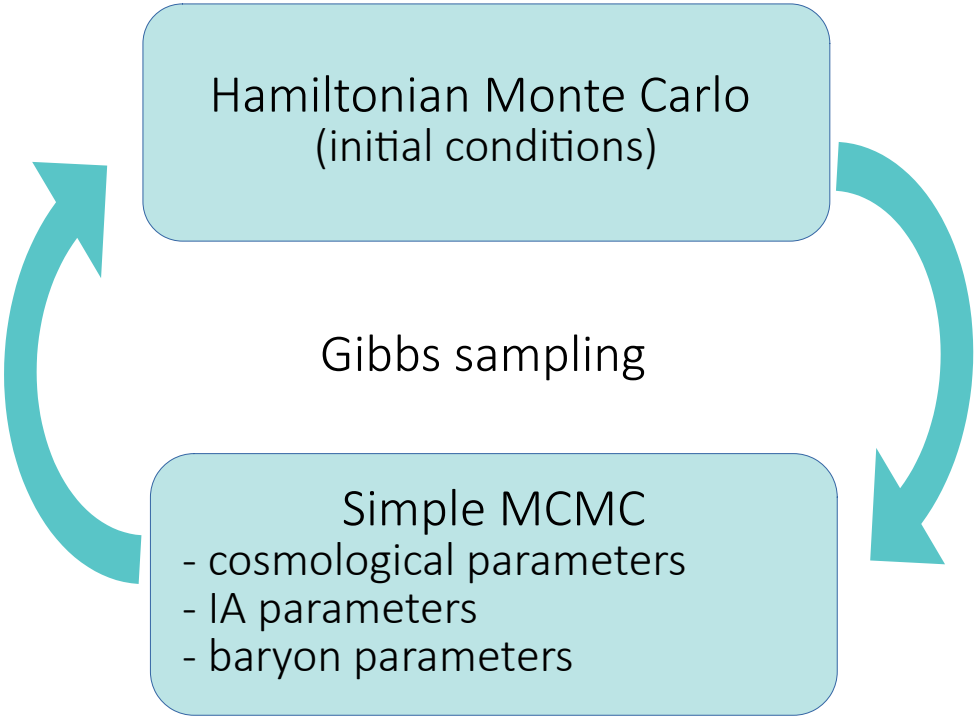
$$\delta^D \left[\epsilon - \mathcal{F}(\delta^{\rm ic})\right] \; G(0, S(\theta))$$

forward model — prior

10^7 parameters!



Sampling efficiently in a high-dimensional space



Addresses curse of dimensionality by:

- exploiting information in gradients

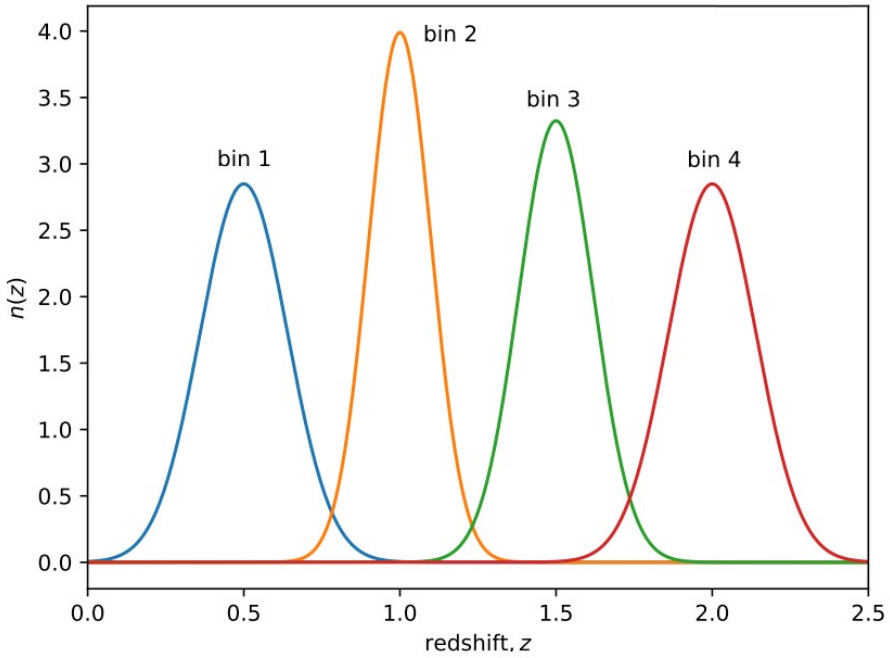
- using conserved quantities

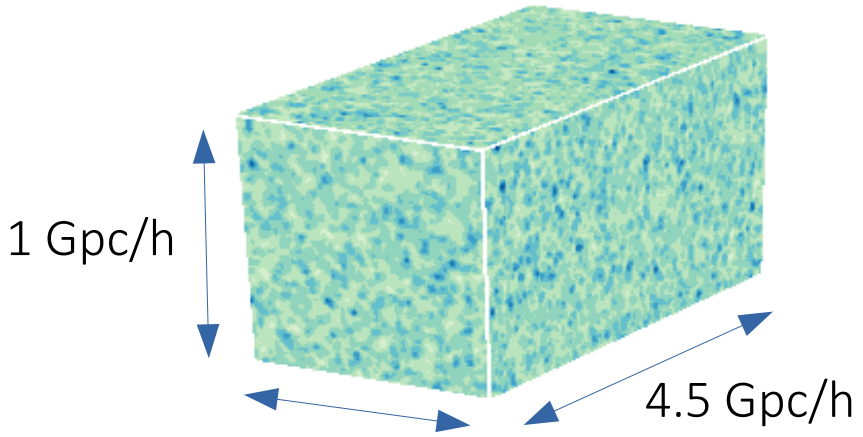
The proof of concept Results with synthetic data

How much do we get from going to the field level?

Simulated data

Tomographic bins





1 Gpc/h (16 deg)^2
(15 Mpc/h, 13 arcmin)

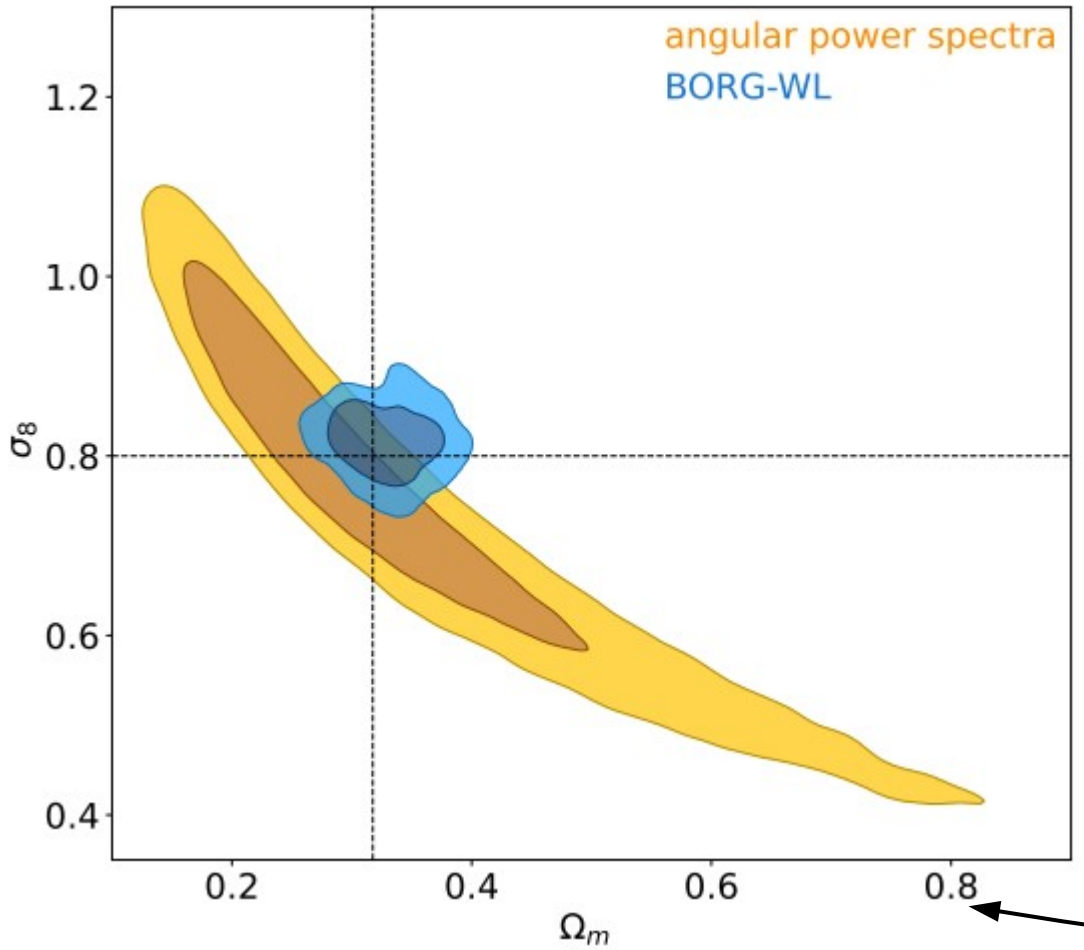
Gravity model: LPT.

Gaussian pixel-noise: 30 sources per square arcmin.

Intrinsic alignment amplitudes: DES values (0.18, 0.8, 0.1)

Comparison of cosmology constraints

marginal error
3 times smaller



$(1\ \mathrm{Gpc/h})^2$, 4.5 Gpc/h
$(16\ \mathrm{deg})^2$, 13 arcmin
LPT

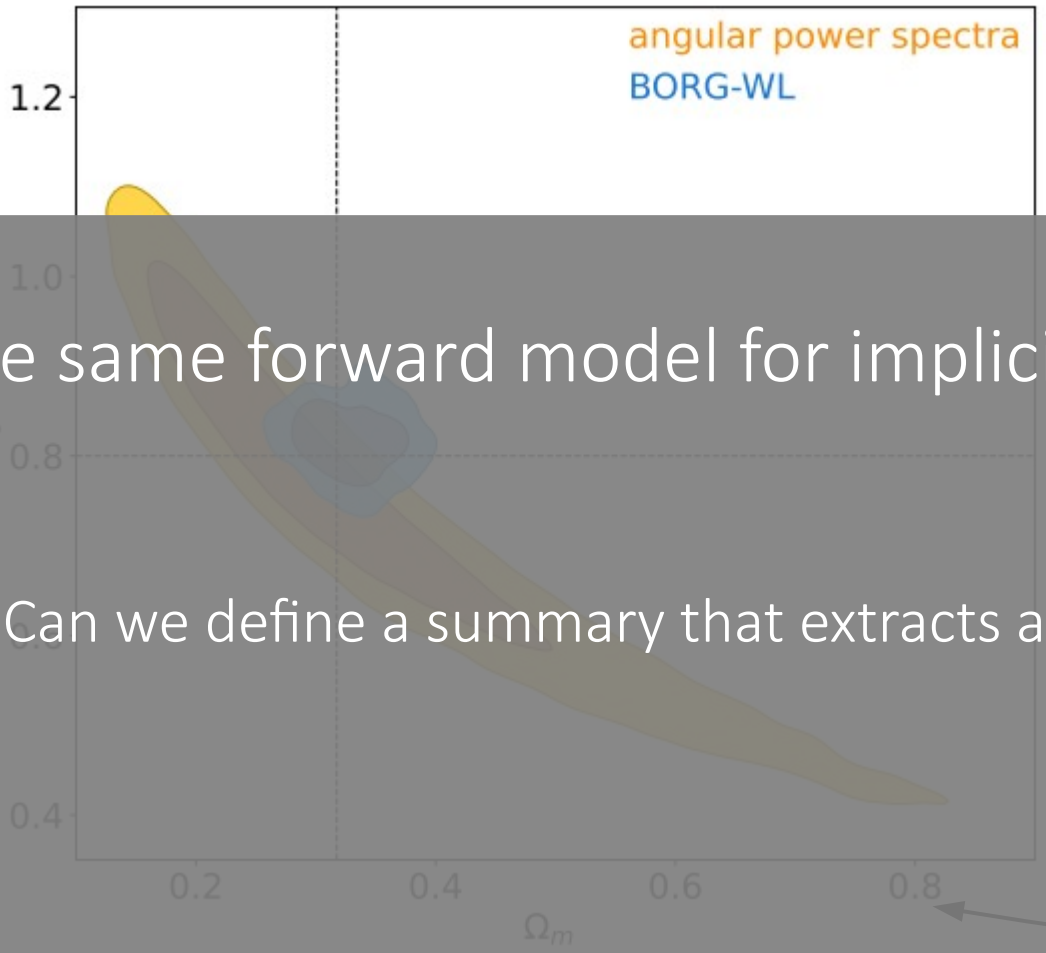
marginal error
5 times smaller

Field-level approach **lifts degeneracy** by extracting more information from the data

Porqueres+2023

Comparison of cosmology constraints

marginal error
3 times smaller



(1 Gpc/h)^2, 4.5 Gpc/h
(16 deg)^2, 13 arcmin
LPT

Using the same forward model for implicit inference

Lucas Makinen
(Imperial College)

Can we define a summary that extracts as much information?

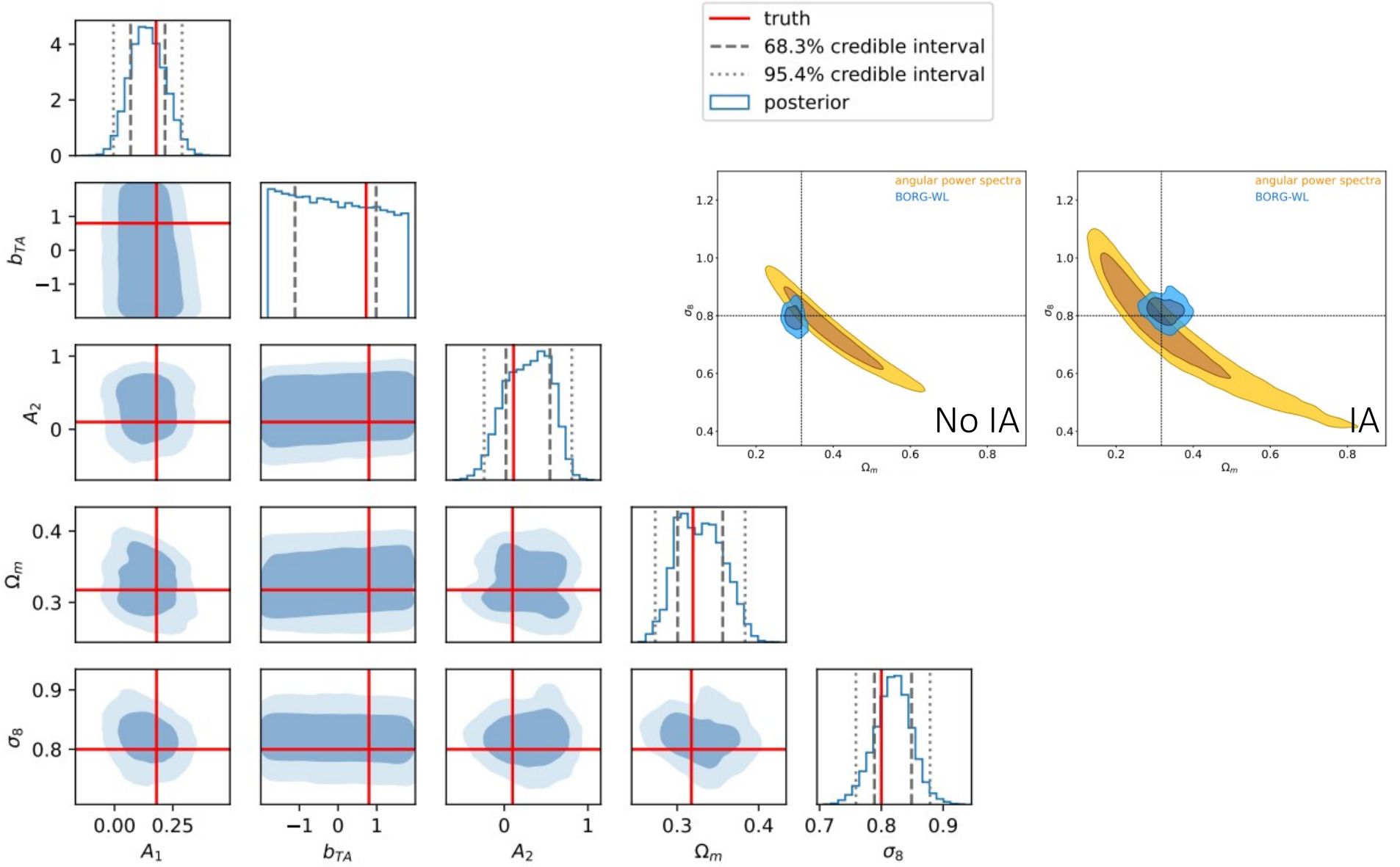
marginal error
5 times smaller

Field-level approach **lifts degeneracy** by extracting more information from the data

Porqueres+2023

Natalia Porqueres

Intrinsic alignment parameters

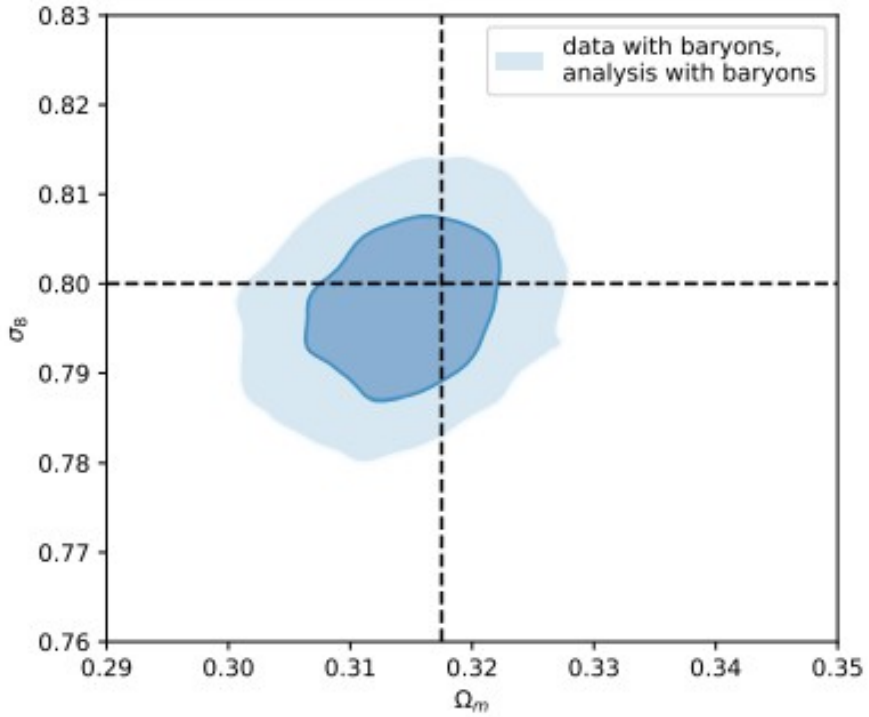


Porqueres+2023

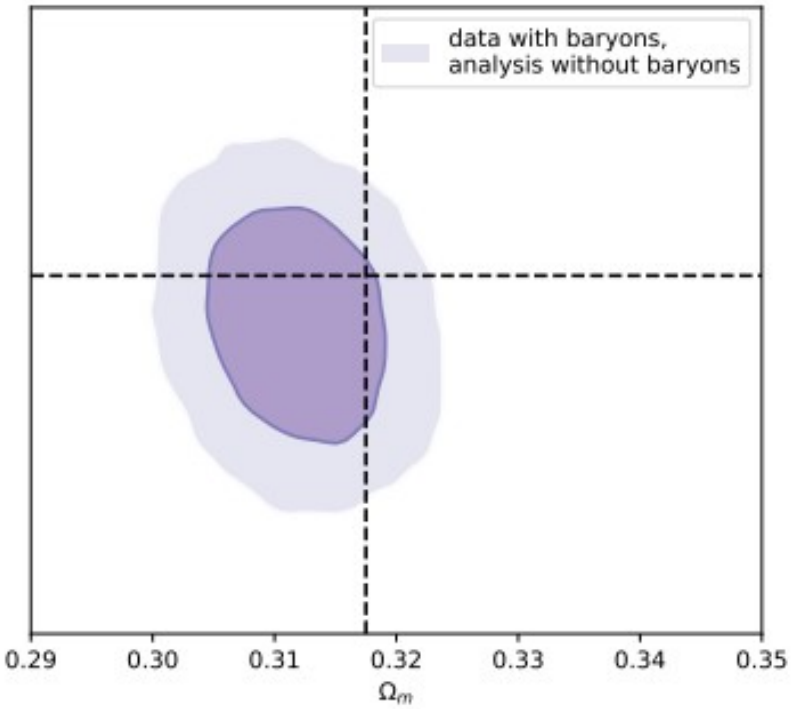
Natalia Porqueres

Baryon feedback parameters

$(0.1\ \mathrm{Gpc}/h)^2$,2.5 Gpc/h
$(3\ \mathrm{deg})^2$, 5 arcmin
LPT



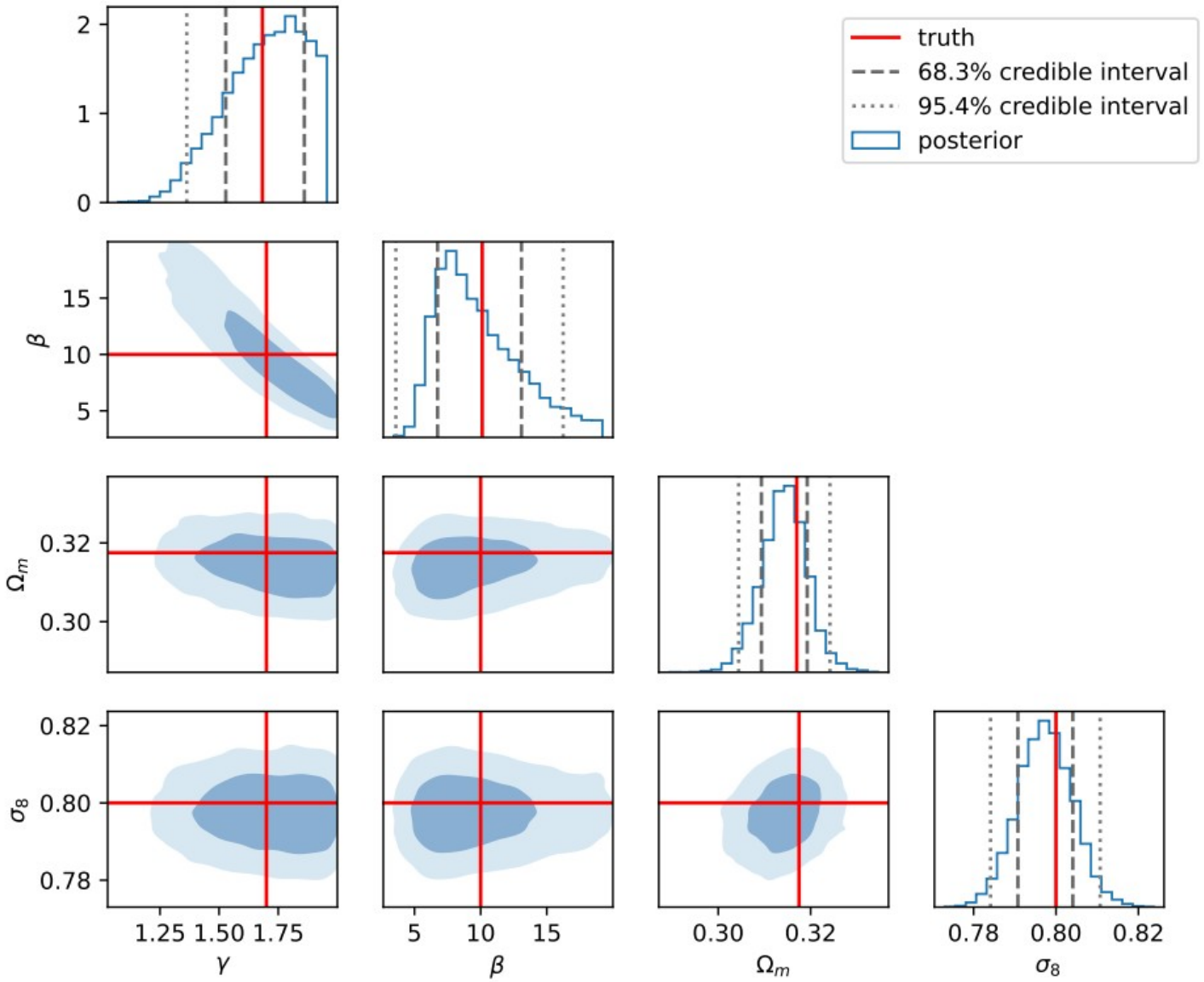
self-consistent



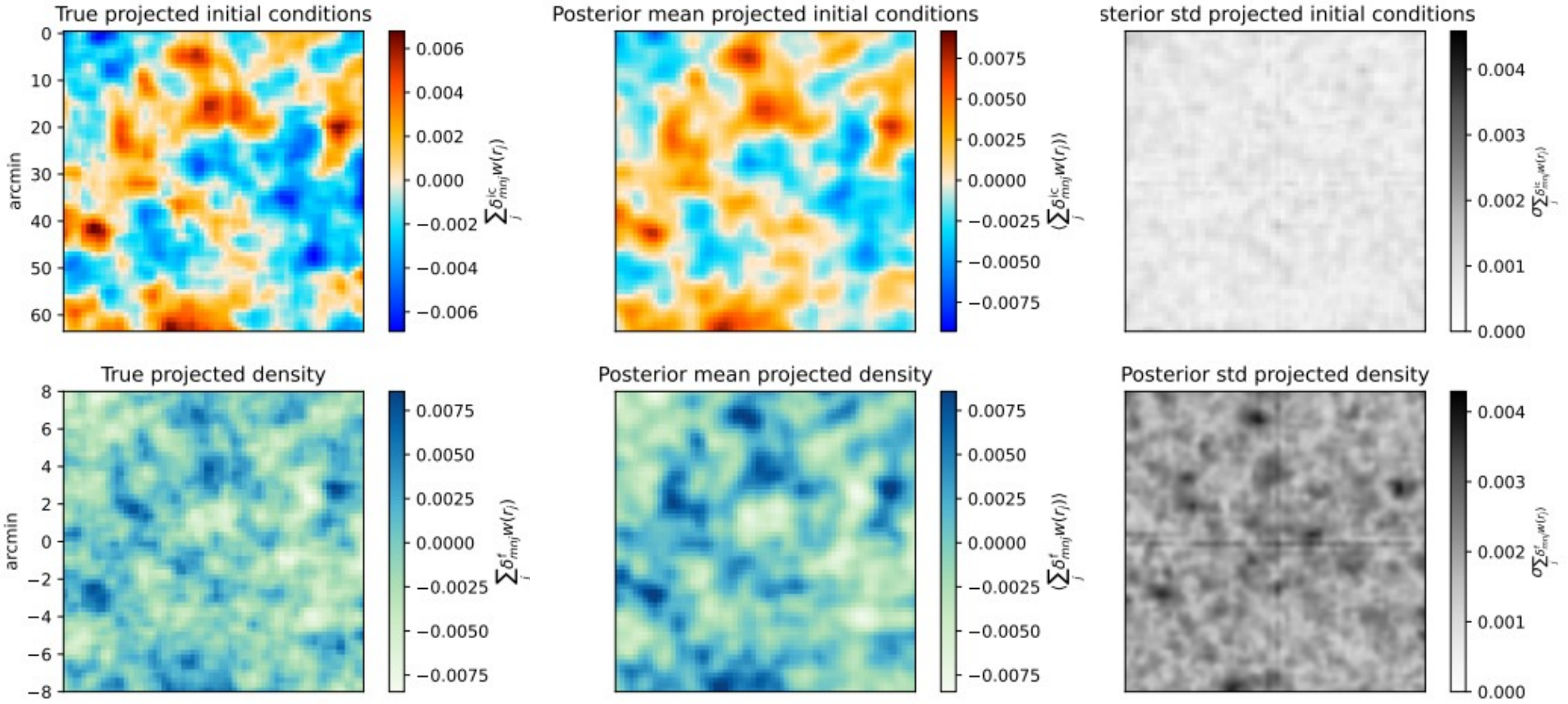
model misspecification

Porqueres+2023

Posterior distribution baryon parameters

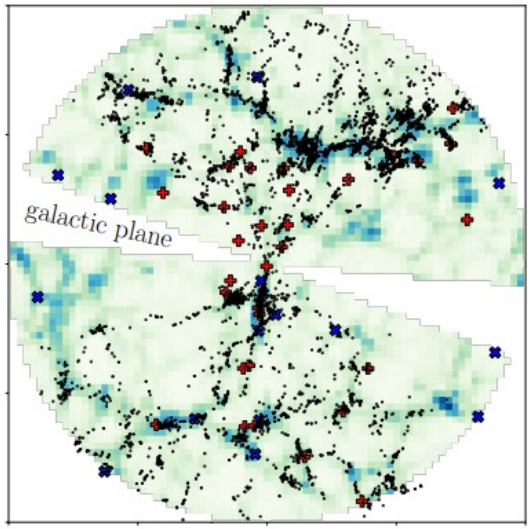


Projection of inferred density fields

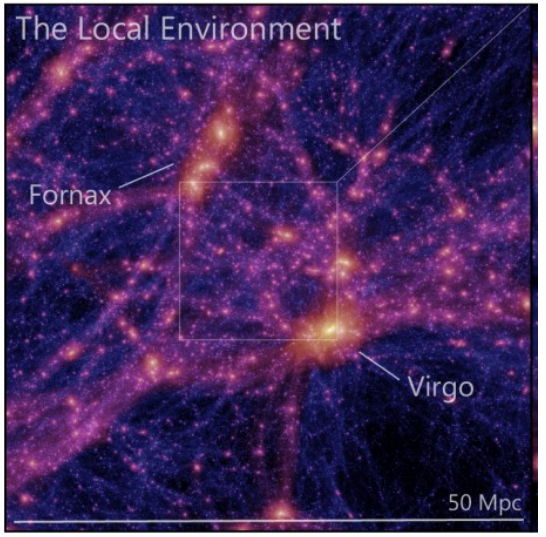


What physics can we do with the density fields?

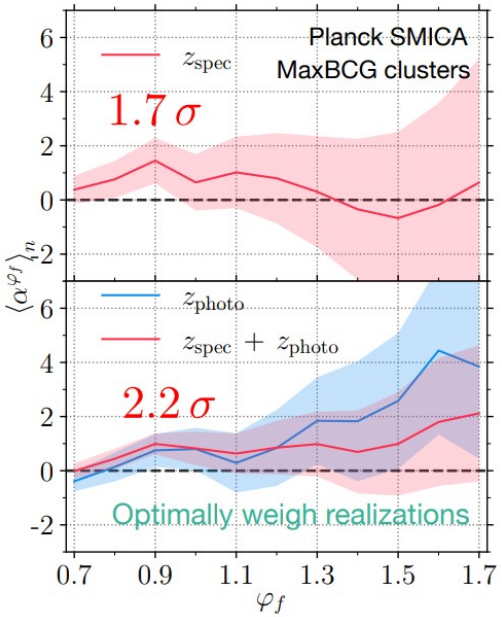
Galaxy or SN environment
(Porqueres et al. 2017)
(Tsaprazi et al. 2021)



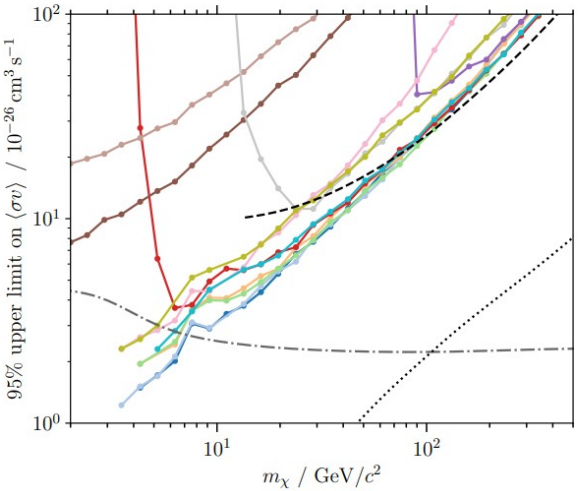
Constrained simulations
(McAlpine et al 2022)



kSZ effect
(Nguyen et al 2020)



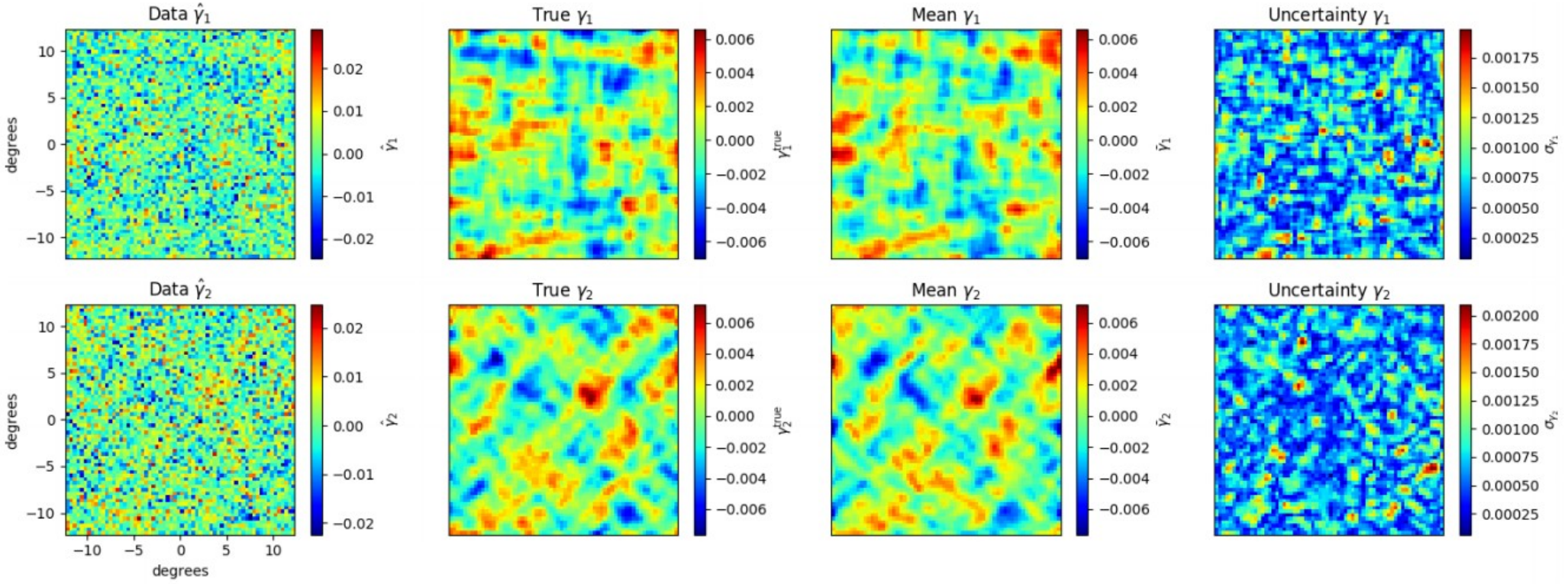
Constrain DM anihilation
(Bartlett et al 2022)



Natalia Porqueres

How can we test the results with such high dimensionality?

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



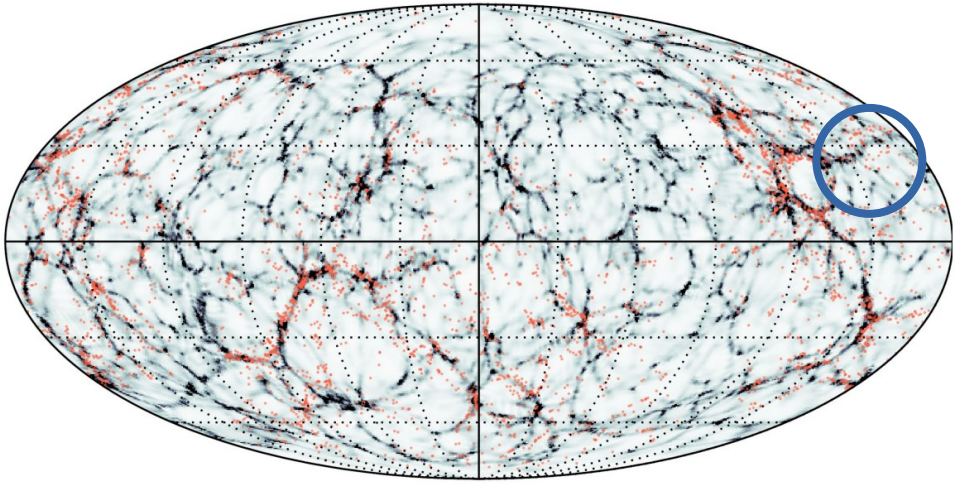
Validating the results of a real data analysis

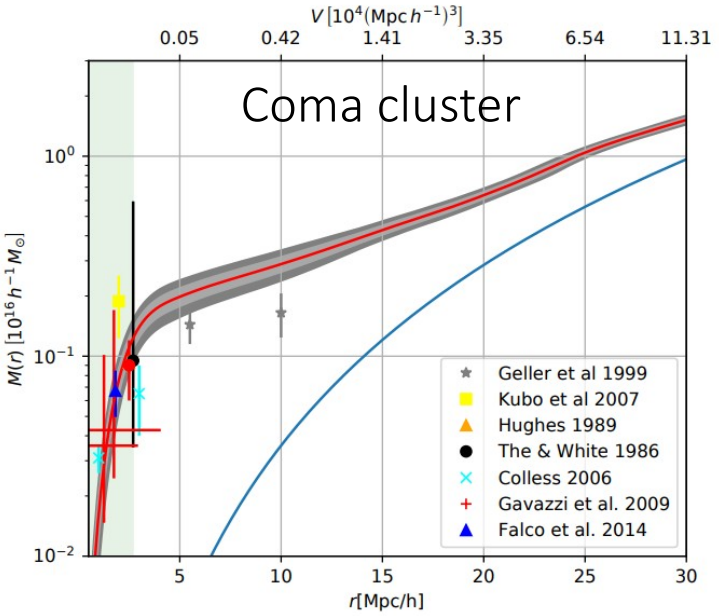
Posterior predictive tests and cross-validation with independent measurements

- do we get the clusters we know are there?

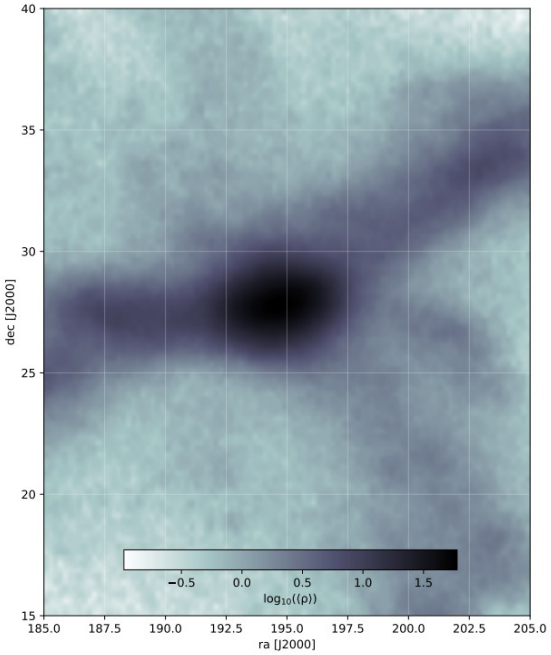
- do we get the expected mass profiles?

- are the inferred IC compatible with CMB?





Jasche & Lavaux 2019



Jasche & Lavaux 2019

Natalia Porqueres

Toward multiprobe analysis

Generated data products

Currently supported data

Real data!

towards real data

Proof of concept

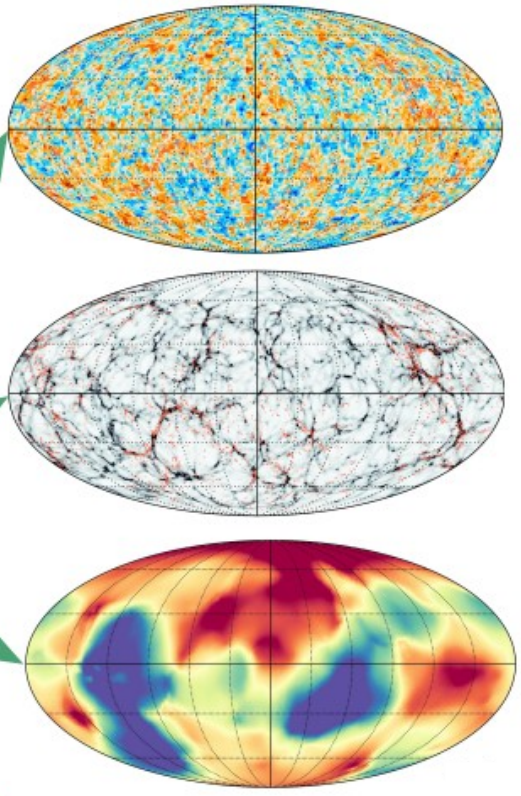
Galaxy clustering
(Jasche et al 2012)

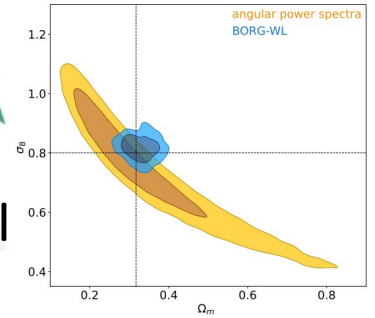
Cosmic shear
(NP et al 2021)

Velocity tracers
(Prideaux et al 2022)

Lyman-α forest
(NP et al 2019)

BORG





Multiple probe analysis will become possible **at field level**

The challenge Getting ready for real data

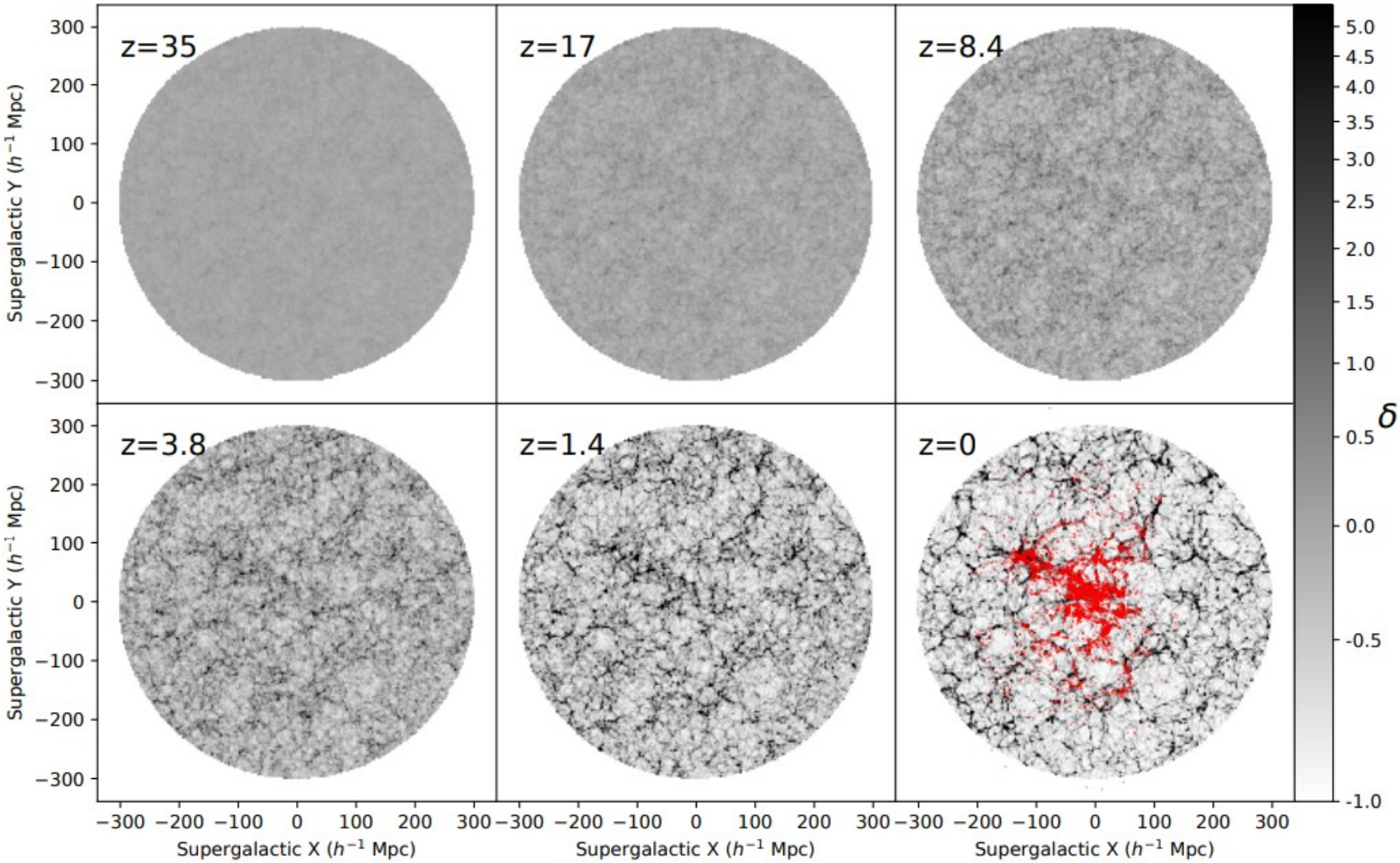
- More realistic gravity model

- Robust n(z)

- What if the models of systematics are not good enough?

More realistic gravity model: Particle-mesh in BORG

Particle mesh is **ready** and has been **used in real data** BORG analysis (2M++)



Jasche & Lavaux 2019

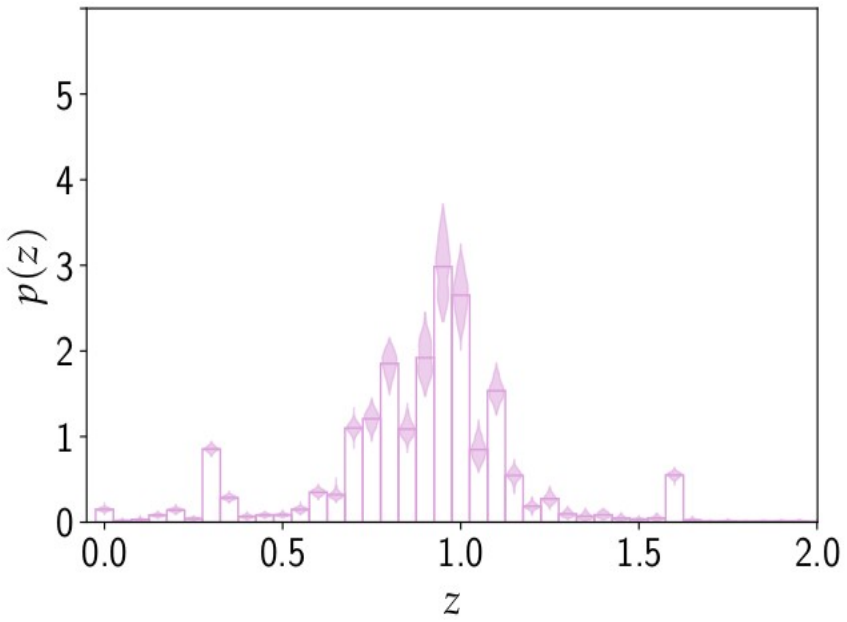
Natalia Porqueres

Calibration systematics: robust n(z)

We need to **propagate the redshift uncertainty** to the cosmology results

Bayesian hierarchical model to sample n(z) and marginalise over redshift uncertainties

Using forward model of the photometric fluxes (Leistedt+ 2016)



Kyriacou, Heavens+ in preparation

What if the models of systematics are not good enough?

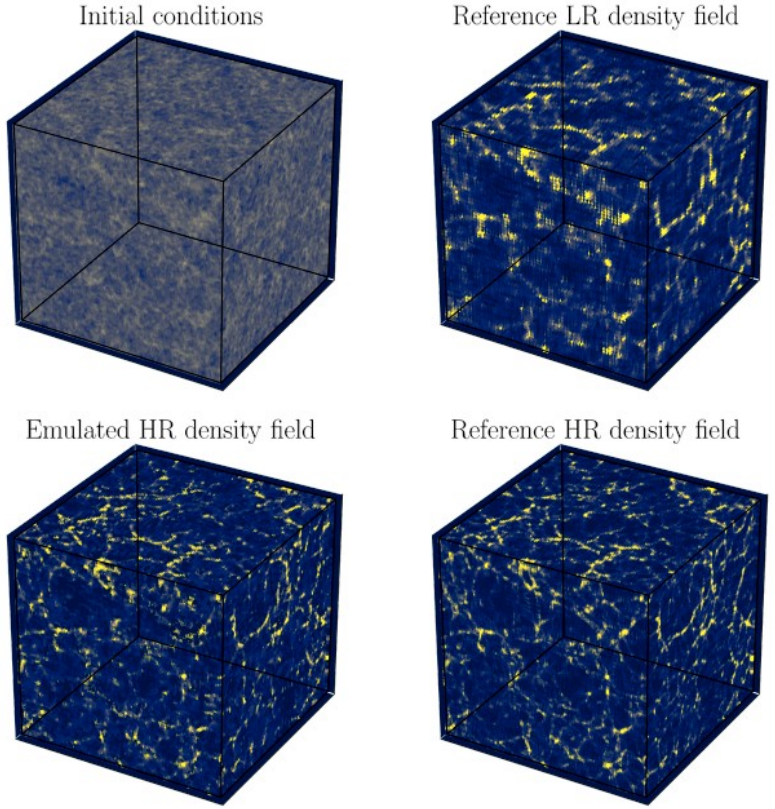
Machine learning can help connecting physical models to data

Neural physical engines (Charnock+ 2019): fully Bayesian way of using NN in MCMC inference

- encoding symmetries to reduce dimensionality

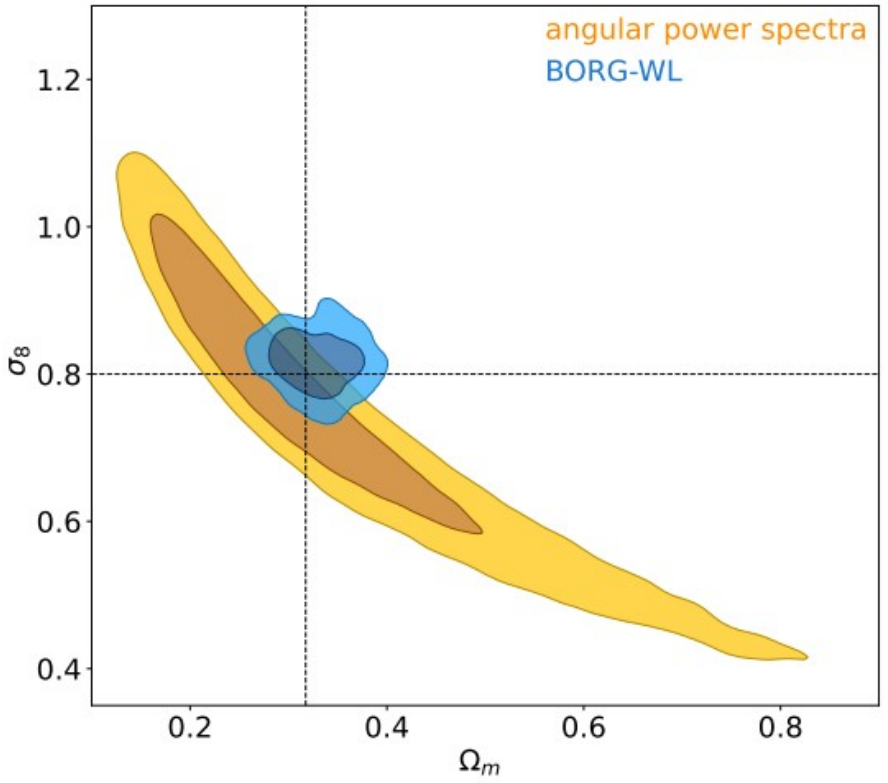
- inferring NN parameters from data. No training data!

An example in BORG: Emulating high resolution N-body



Charnock+ 2019

Summary

- There is more information in the data that the 2-point summary statistics do not capture.

- Field-based approach lifts degeneracy and reduces marginal error up to a factor of 5.