Field-level inference of weak lensing

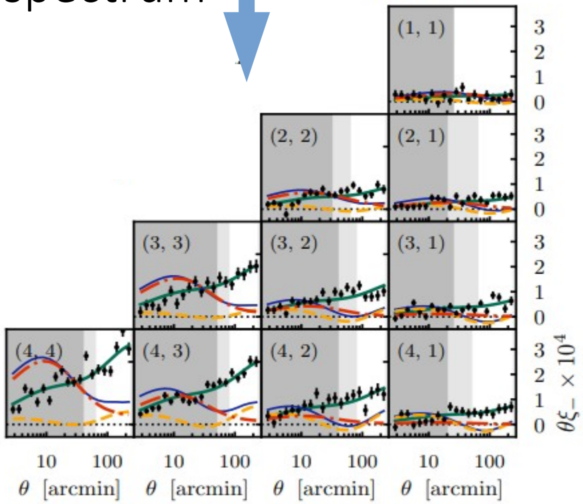
Natalia Porqueres

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

MIAPbP – 14/4/2025

The standard analysis of weak lensing

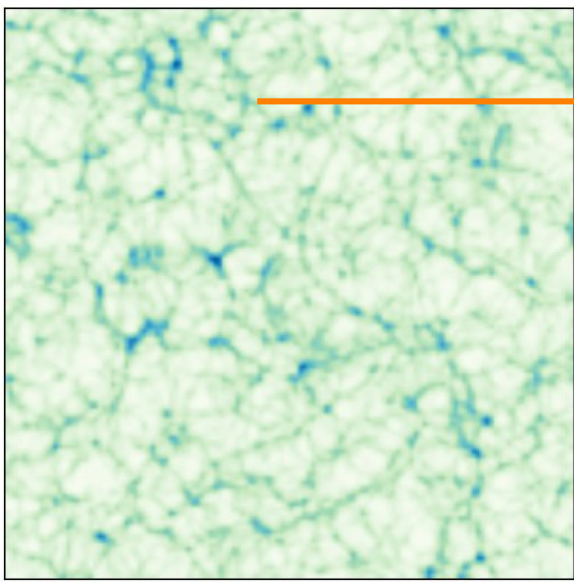
Compression to
power spectrum



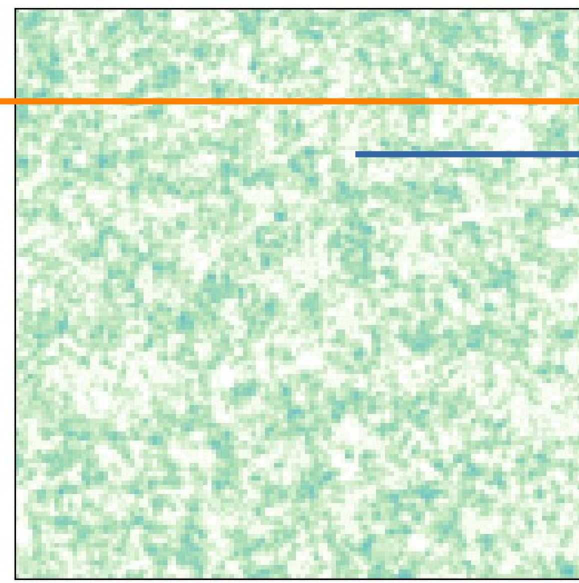
Standard analysis discards information

Is the standard approach enough?

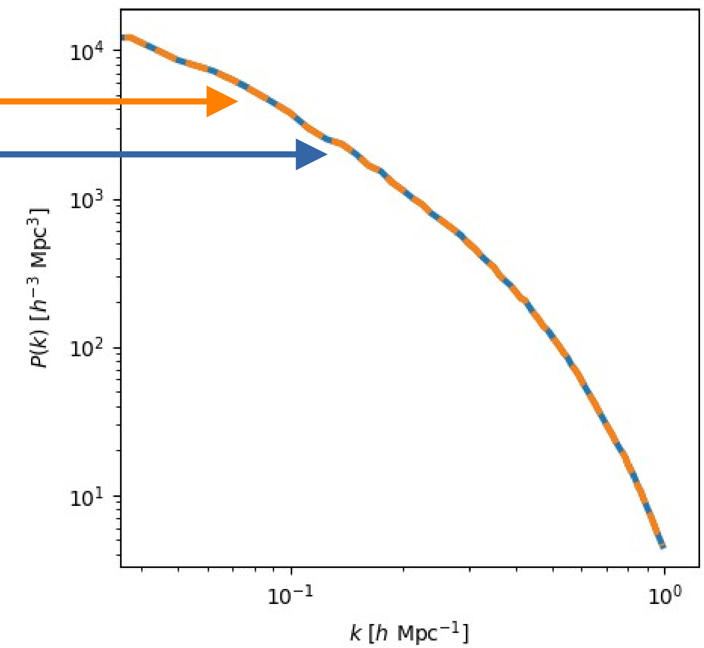
The standard approach misses the non-Gaussian information



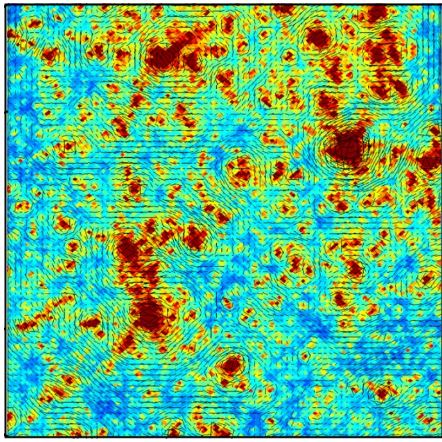
Cosmic web



Gaussian random field



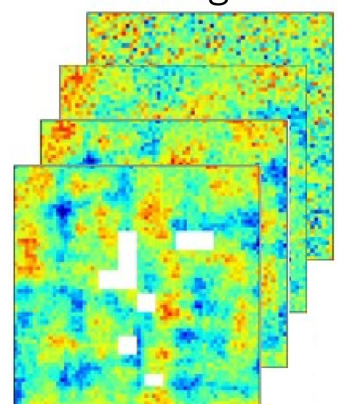
Extracting more information from the data



Data compression → High-order statistics — Arka's talk

Data compression → Simulation-based inference (neural network statistics) — Kai's talk

Cosmological parameters —Forward model (simulations)→ Predicted lensing fields



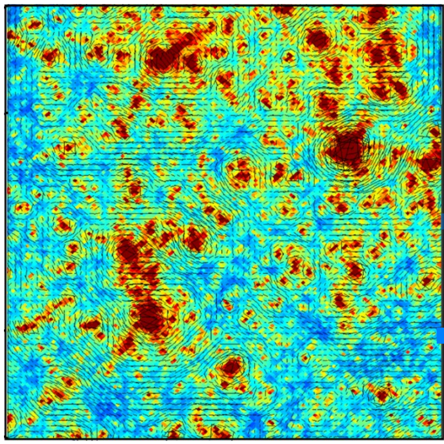
Predicted lensing fields —Data compression→ Summary statistics

data —Data compression→ Summary statistics

Summary statistics ⇄ Summary statistics

Natalia Porqueres

Extracting more information from the data

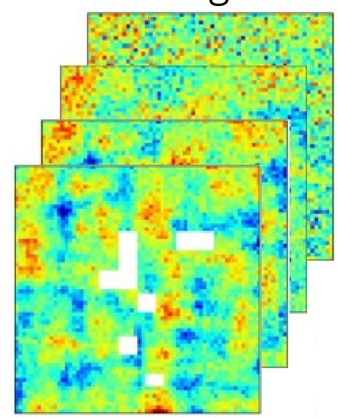


Data compression → High-order statistics — Arka's talk

Data compression → Simulation-based inference (neural network statistics) — Kai's talk

→ **Field-level approach**

Cosmological parameters —Forward model (simulations)→ Predicted lensing fields ⇄ data



Likelihood at field level

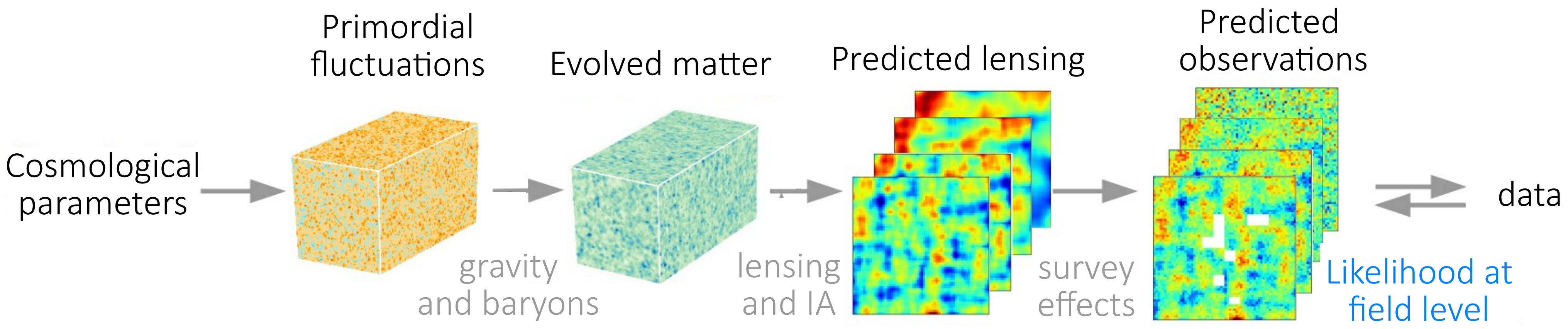
Captures all the information

Easier to separate cosmology and systematics

Provides a map of the matter distribution

Natalia Porqueres

BORG-WL: field-level inference of weak lensing



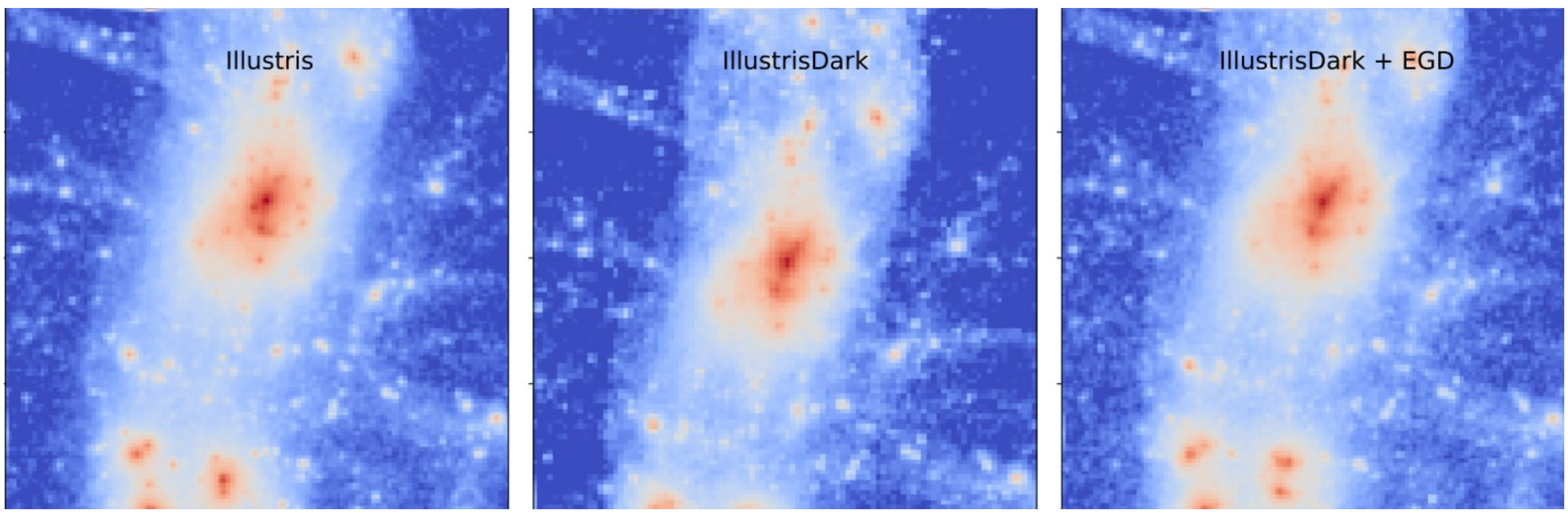
Natalia Porqueres

Baryon feedback: Enthalpy gradient descent model

Displaces dark matter particles with

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

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



Dai & Seljiak, 2018

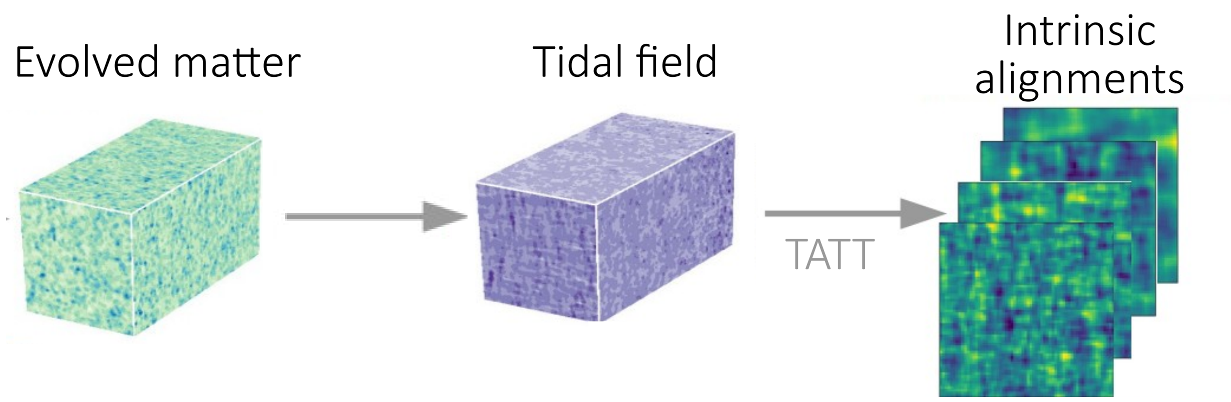
Self calibration of model parameters.

Displaces particles independently of whether they are in a halo.

Fully differentiable.

Intrinsic alignments: Tidal alignment and tidal torquing

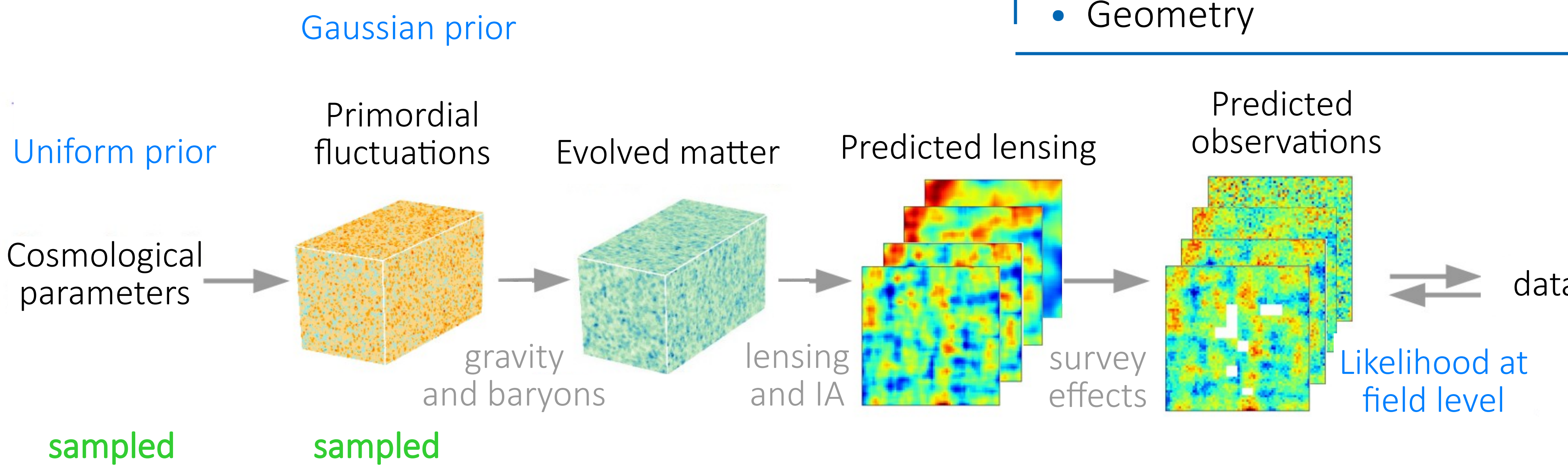
TATT model: tidal alignment and torquing from the gravitational potential



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

BORG-WL: field-level inference of weak lensing

Sensitivity to cosmology:
- Initial matter power spectrum
- Growth of structures
- Geometry



Likelihood: Assuming Gaussian noise in pixelised shear

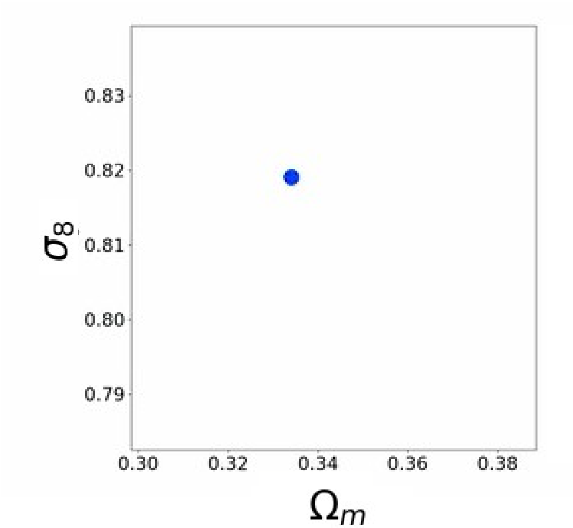
$$\log \mathcal{L} = -\frac{1}{2} \sum_{b} \sum_{mn} \frac{(\epsilon^b_{1,mn} - \hat{\epsilon}^b_{1,mn})^2 + (\epsilon^b_{2,mn} - \hat{\epsilon}^b_{2,mn})^2}{\sigma_b^2}$$

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

Natalia Porqueres

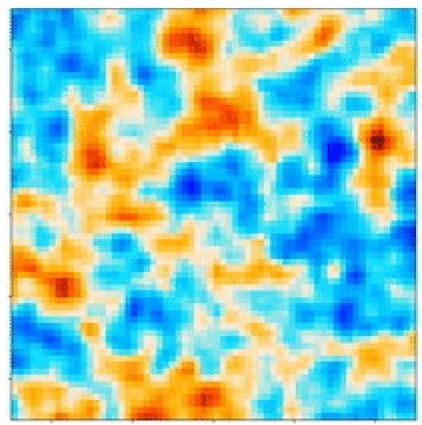
The samples

Cosmological parameters



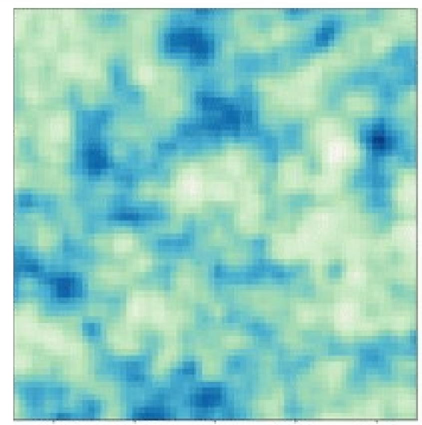
sampled

Primordial fluctuations

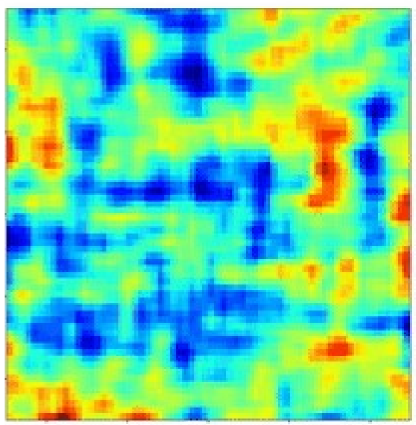


sampled
10^7 parameters!

Evolved matter



Predicted lensing



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

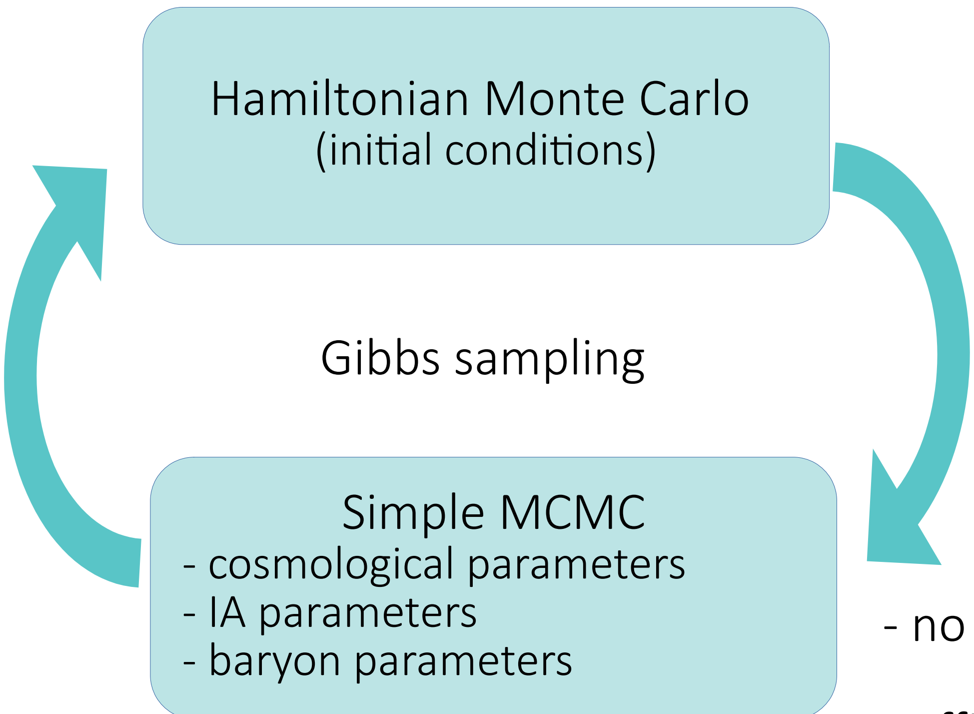
posterior likelihood prior

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

forward model

Natalia Porqueres

Sampling in a high-dimensional space



Addresses curse of dimensionality by:

- exploiting information in gradients

- using conserved quantities

- no analytical derivatives

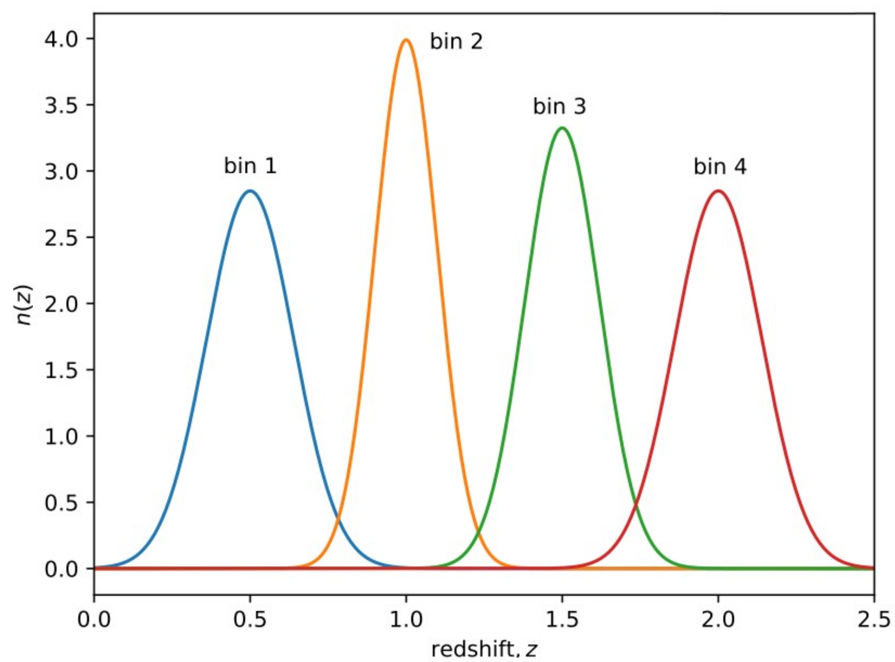
- efficient mass matrix?

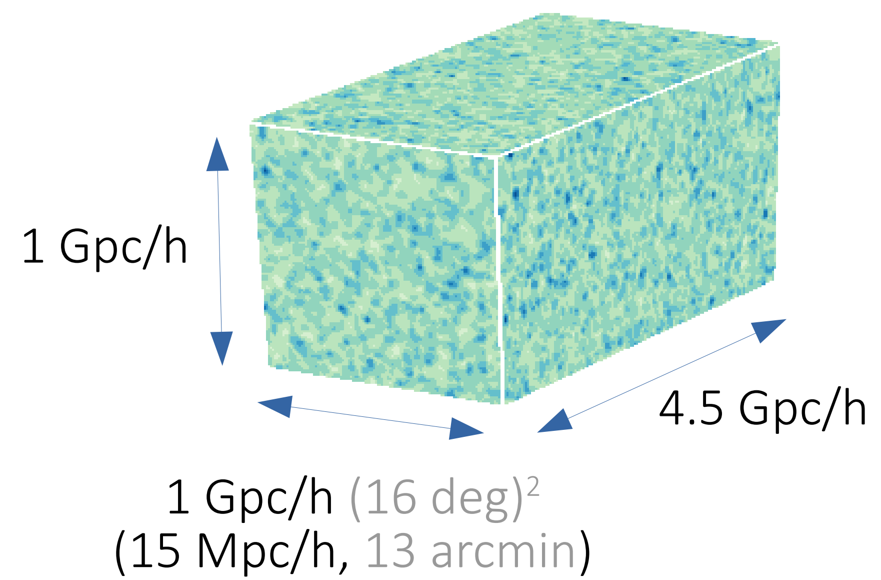
The proof of concept Results with synthetic data

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

Simulated data

Tomographic bins





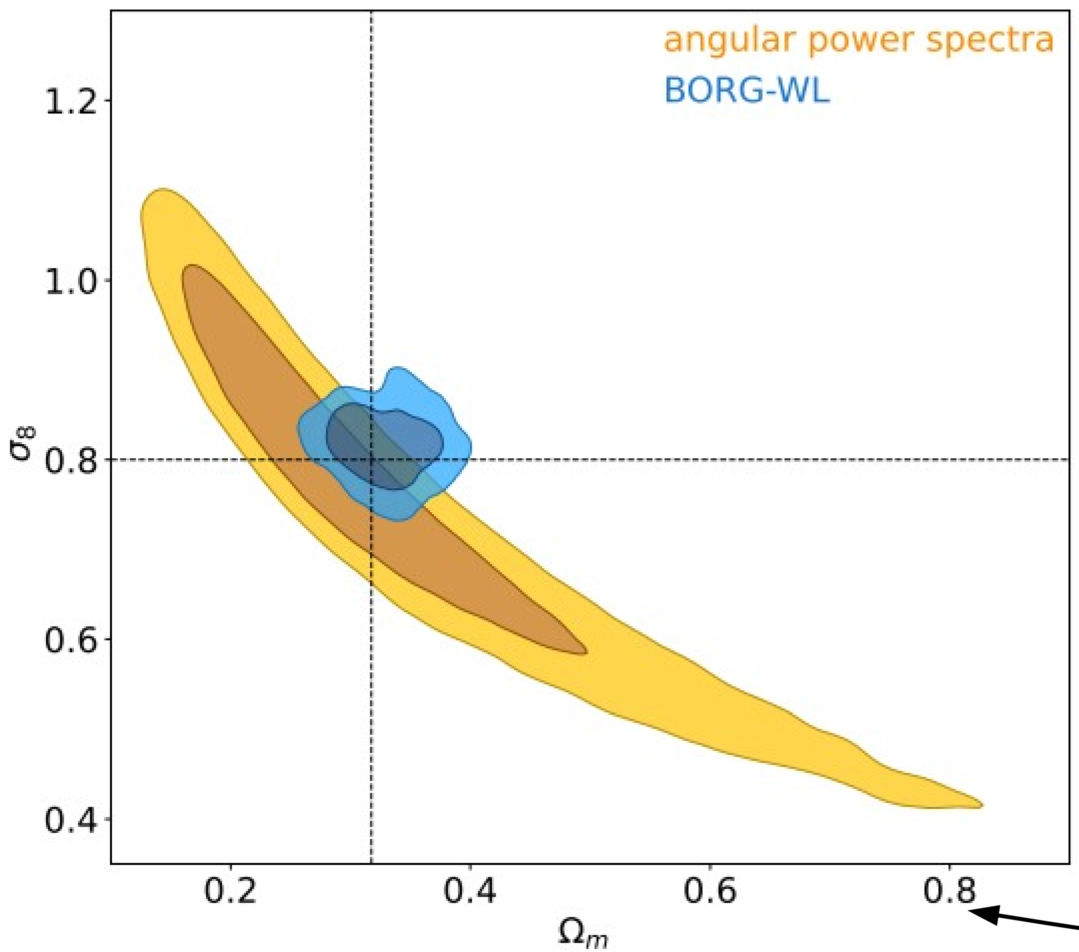
Gravity model: LPT.

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

marginal error
5 times smaller

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

NP+2022, 2023

Where is the information coming from?

Comparison with high-order stats

Is there a summary statistic that captures as much information?

Extreme data compression for SBI that preserves information

MOPED for Gaussian data, can we do it for non-Gaussian data?

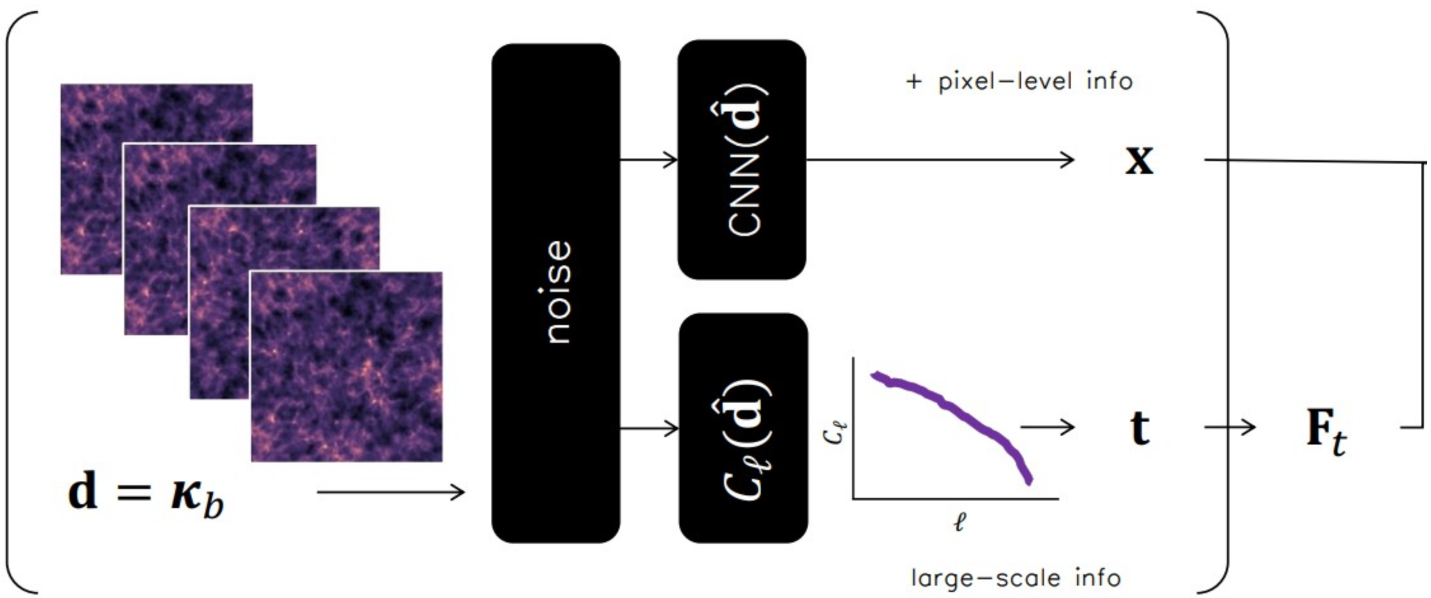
Hybrid summary statistics

Physics-based statistics (e.g. power spectrum) are informative. Use them

More information can be captured with CNN

Ask neural network to capture ***additional*** information

Lucas Makinen
Imperial College



$$\mathbf{F} = \mathbf{F}_t + \frac{1}{s}\mathbf{v}\mathbf{v}^T$$

$$[\mathbf{v}]_i = \langle x \rangle_{,\theta_i} - \boldsymbol{\mu}^T_{,\theta_i}\mathbf{C}_t^{-1}\mathbf{u}$$

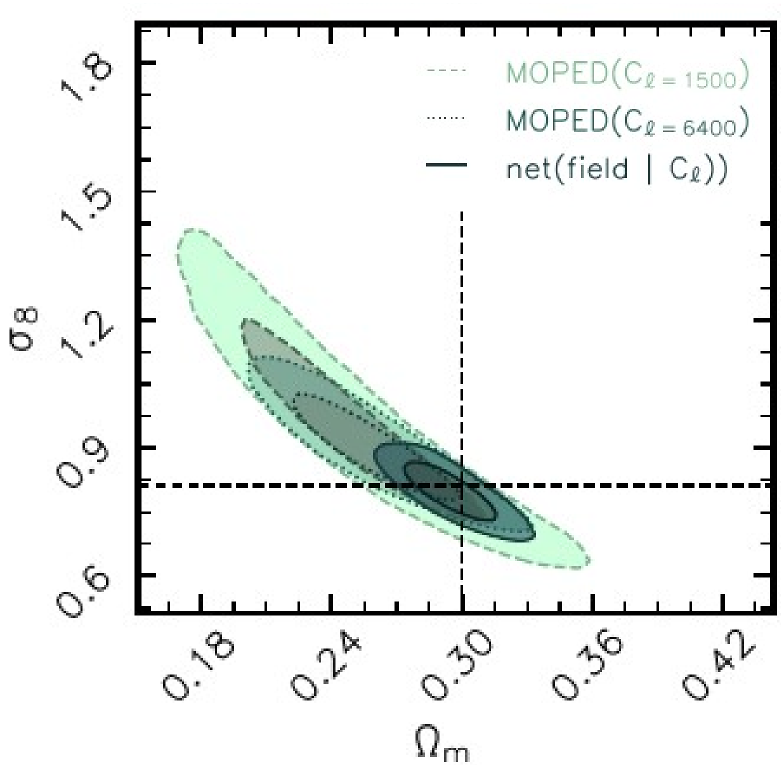
Hybrid summary statistics

Physics-based statistics (e.g. power spectrum) are informative. Use them

More information can be captured with CNN

Ask neural network to capture ***additional*** information

Lucas Makinen
Imperial College



highly complementary summary to 2-point statistics.

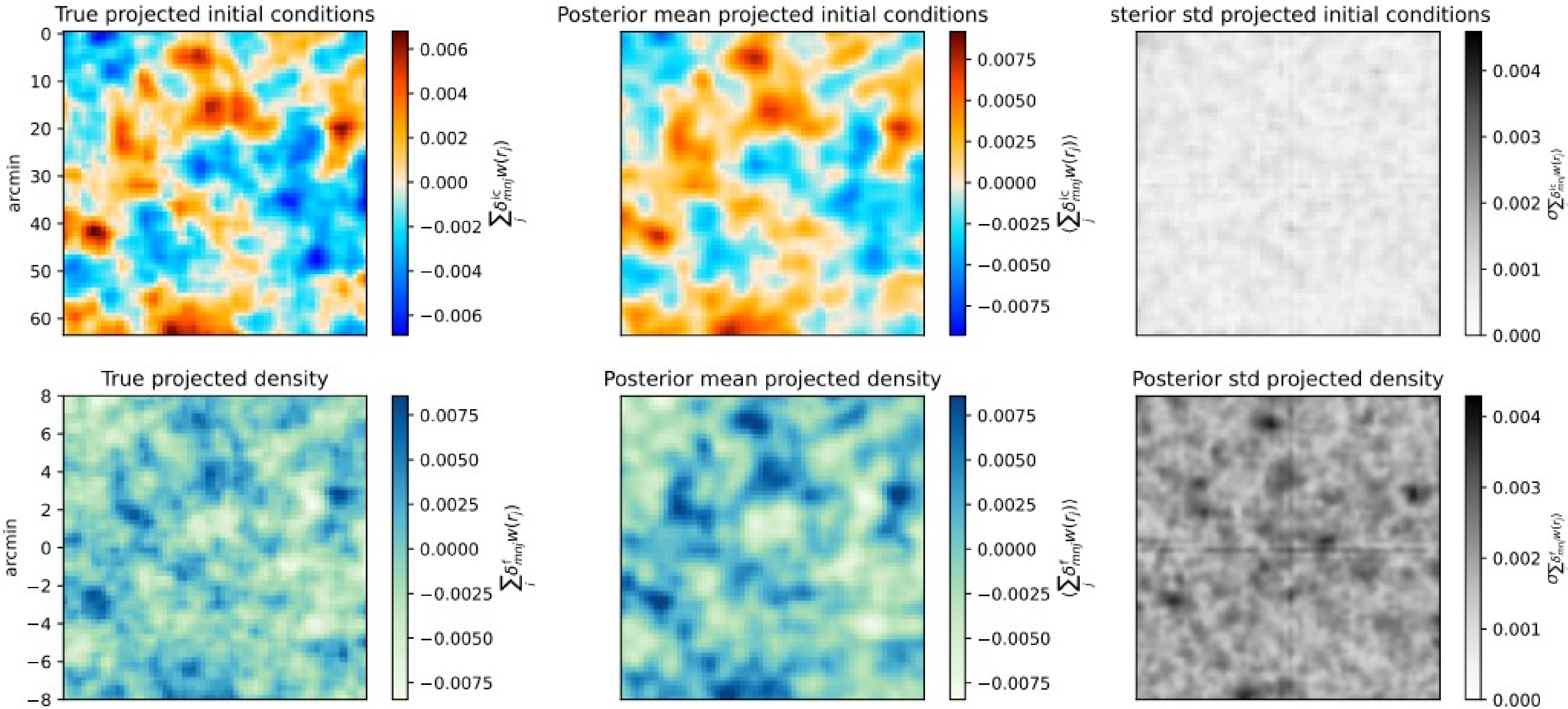
tighter constraints than combining independently-defined statistics.

lightweight NN to match large regression networks with fewer simulations.

Makinen+ 2024

Matter reconstruction from field-level

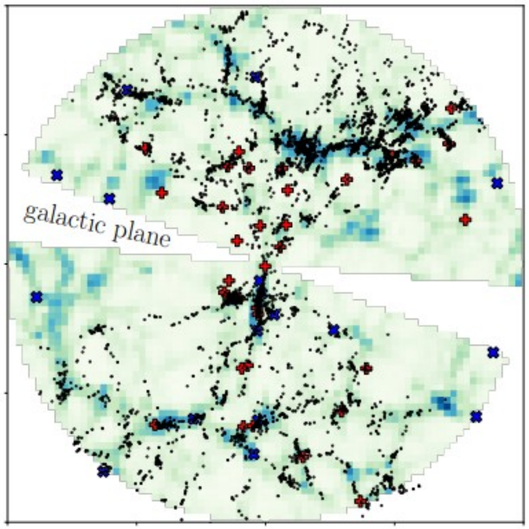
Inferred density fields



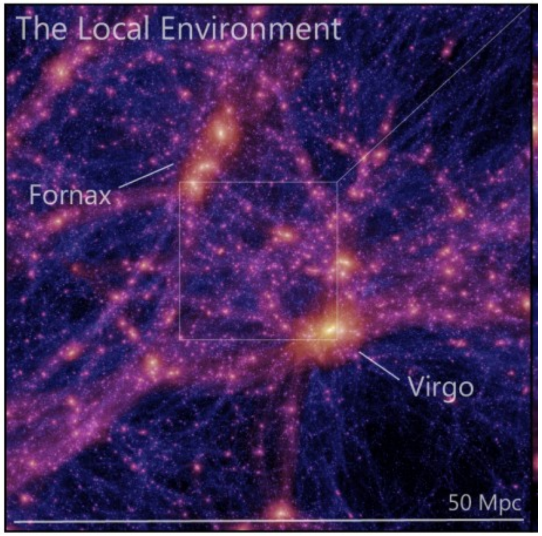
NP+2022, 2023

What physics can we do with the density fields?

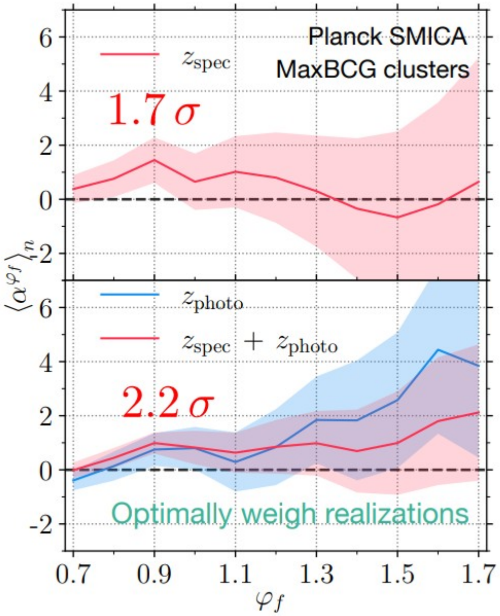
Galaxy environment
(**NP**+ 2017)



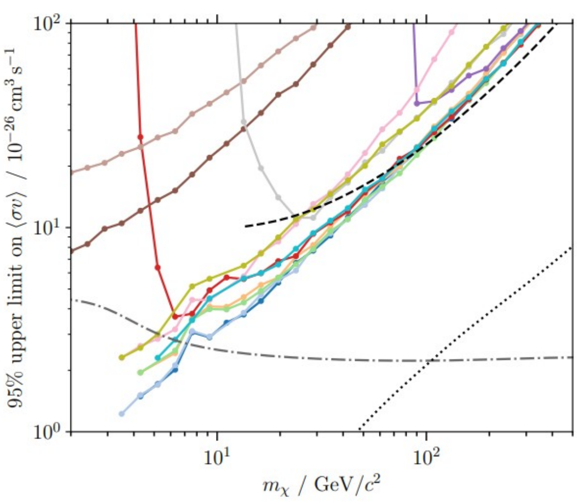
Constrained simulations
(McAlpine+ 2022)



kSZ effect
(Nguyen+ 2020)



Constrain DM annihilation
(Bartlett+ 2022)



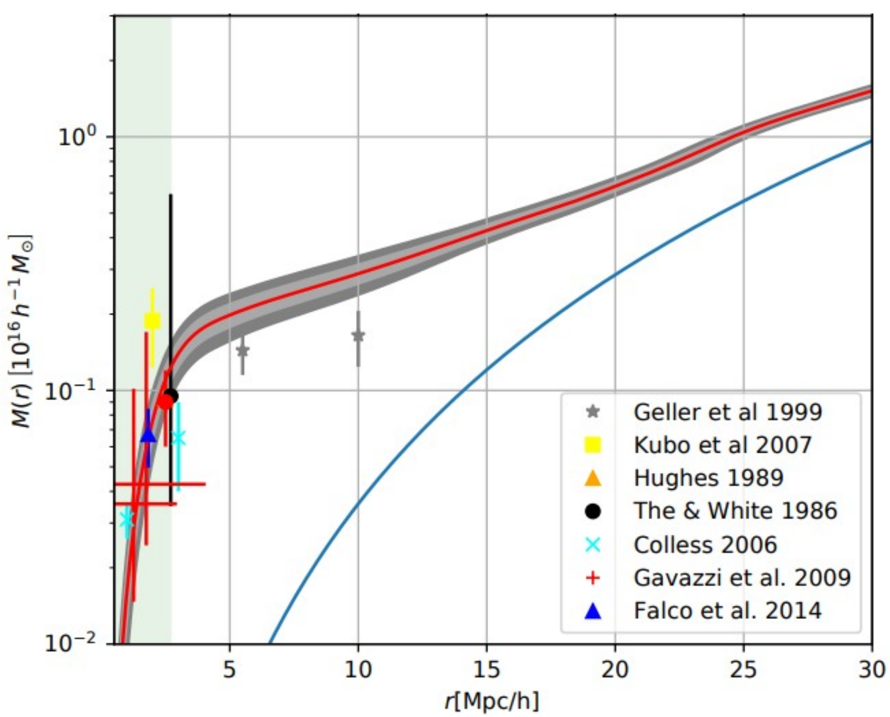
Natalia Porqueres

Density fields as a validation tool

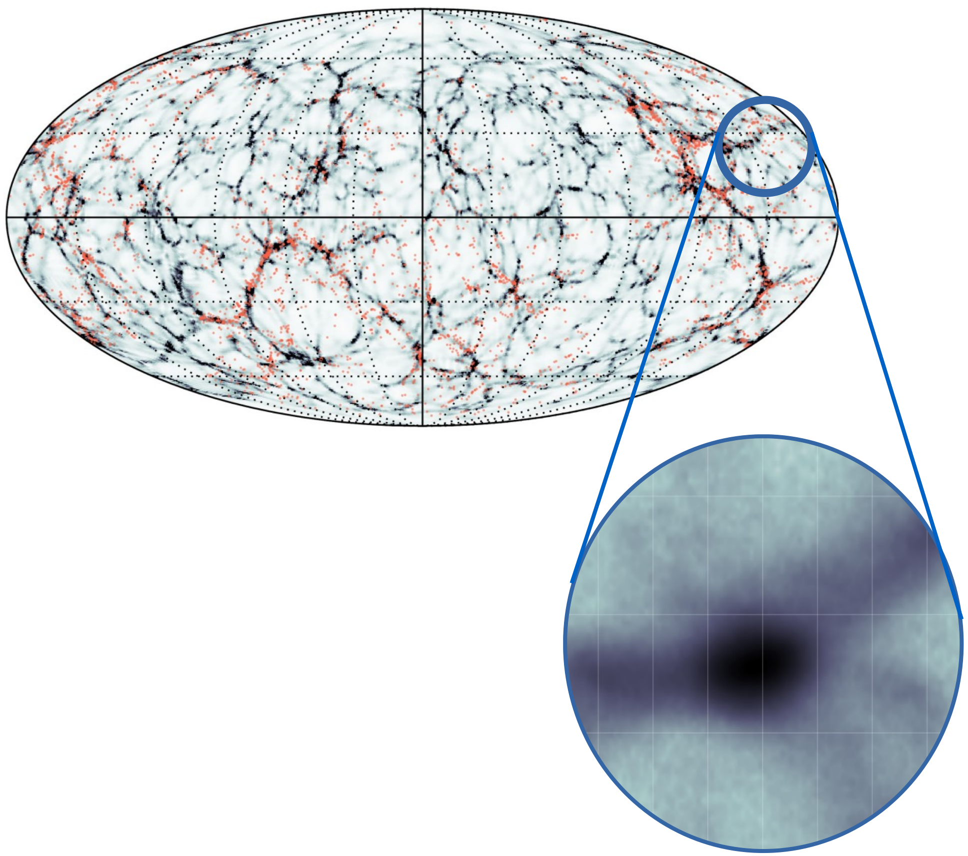
Posterior predictive tests and cross-validation with independent measurements

Inferred mass profile agree with external measurements?

Mass profile



Jasche & Lavaux 2019

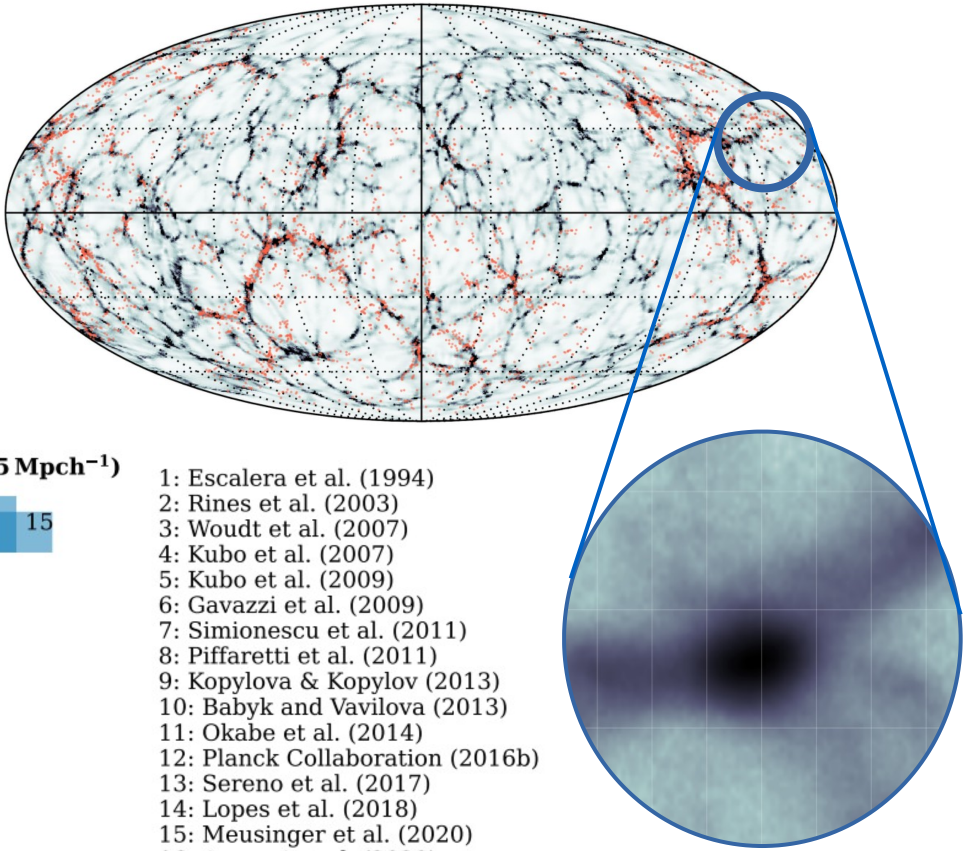


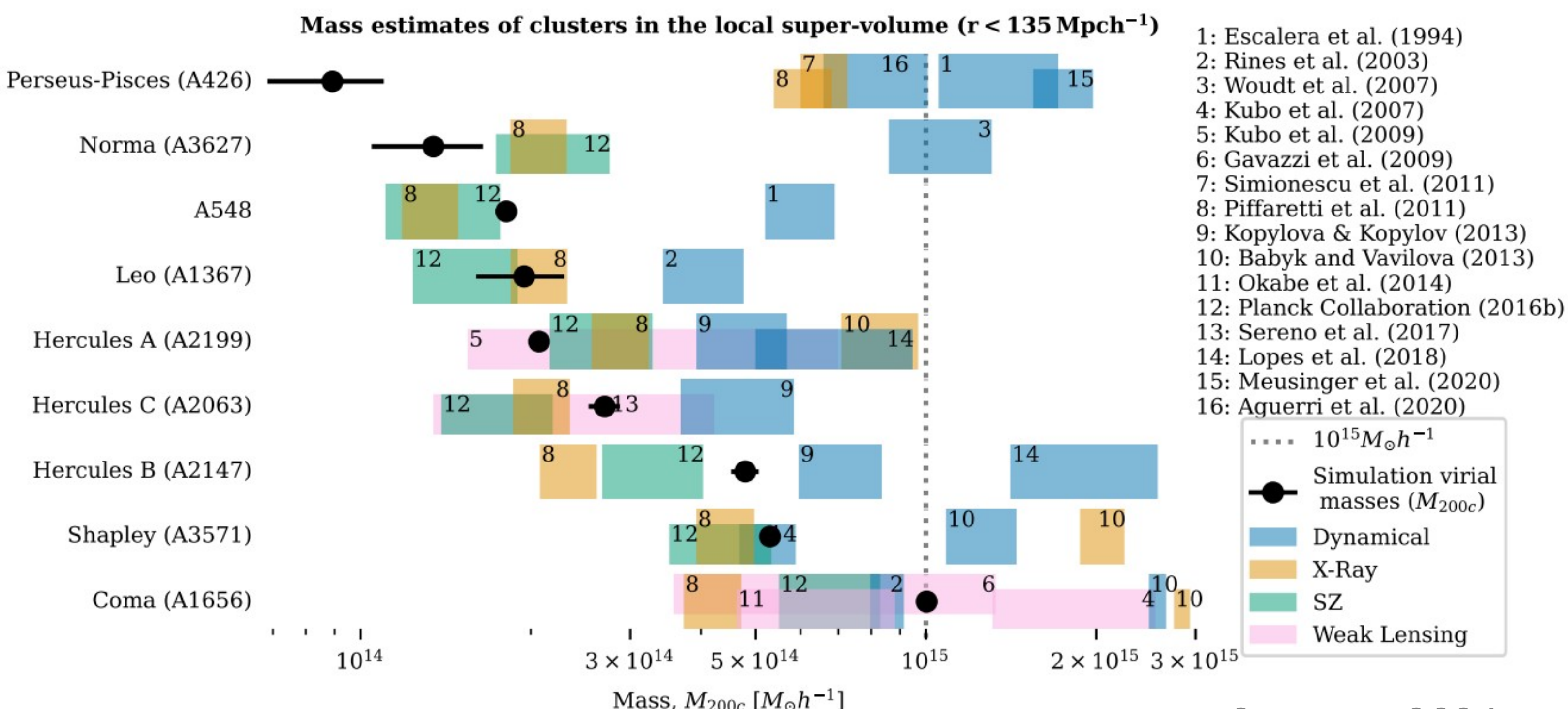
Natalia Porqueres

Density fields as a validation tool

Posterior predictive tests and cross-validation with independent measurements

Inferred mass profile agree with external measurements?

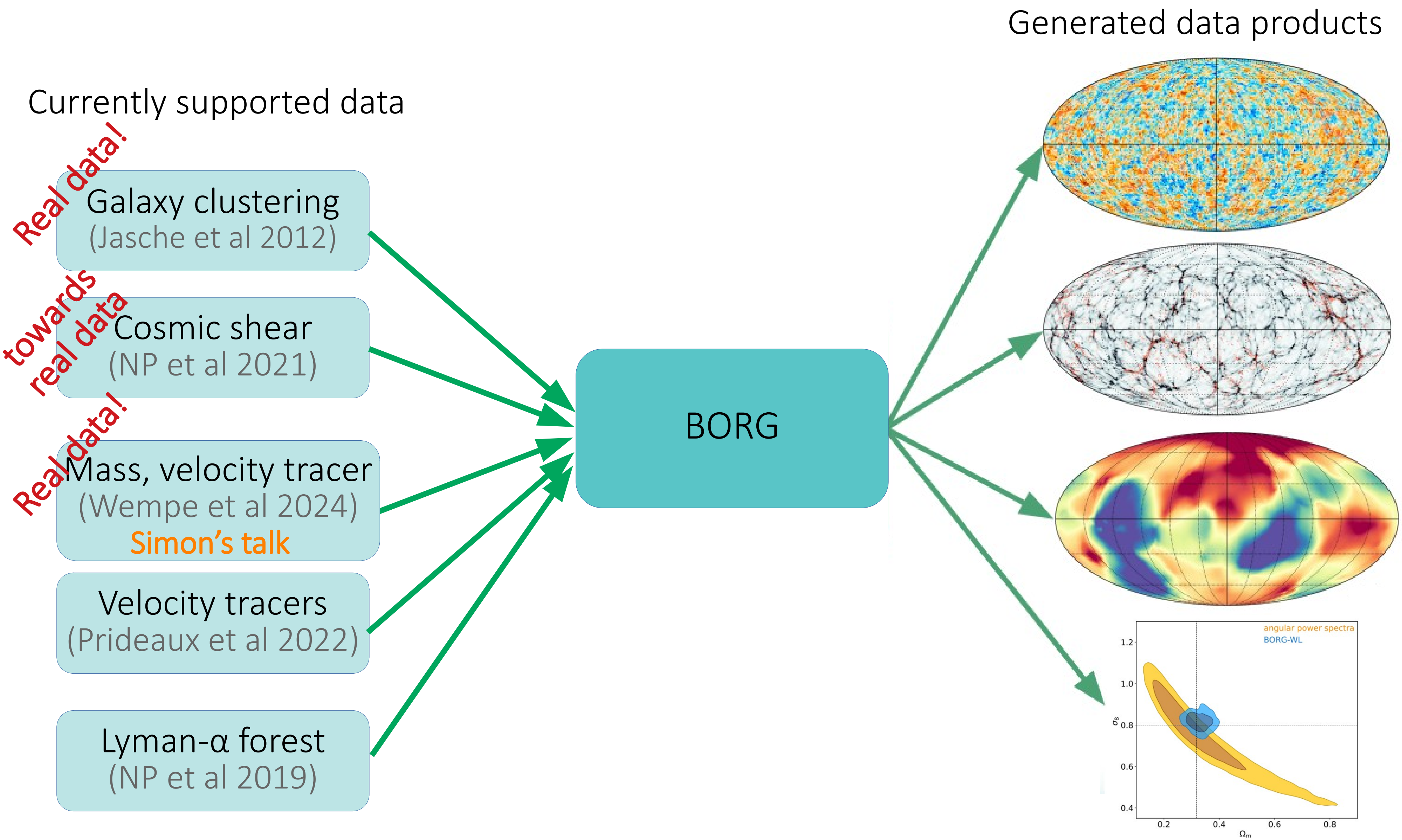




Stopyra+ 2024

Natalia Porqueres

Toward multiprobe analysis



Multiprobe analysis at the field-level to capture physical correlations

The challenge Getting ready for Euclid data

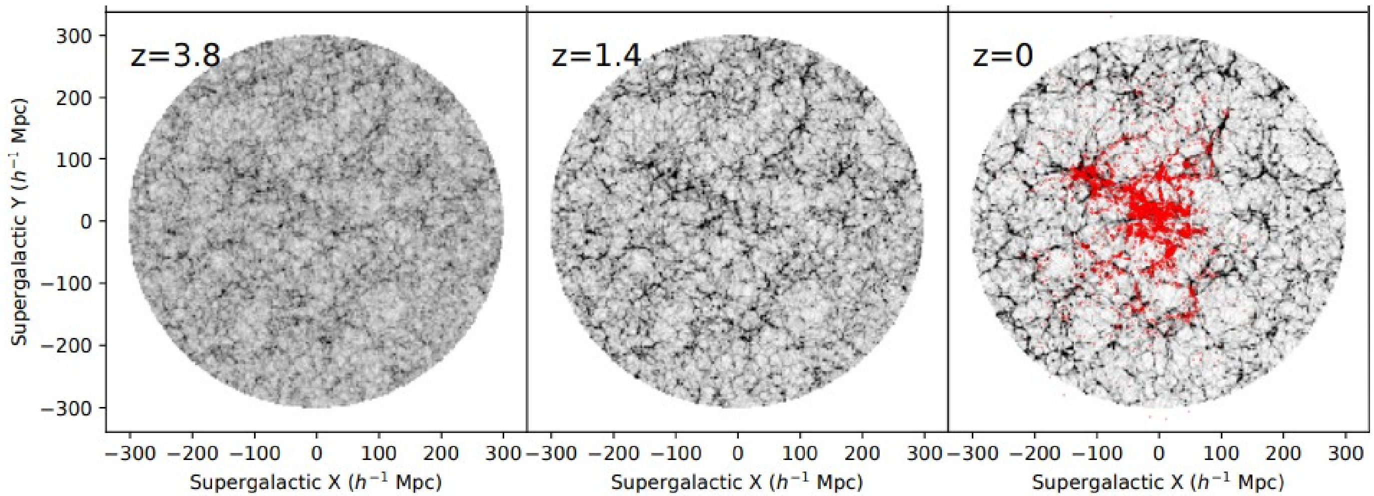
More realistic gravity model

Systematics (redshifts, model accuracy)

Scalability to Euclid data

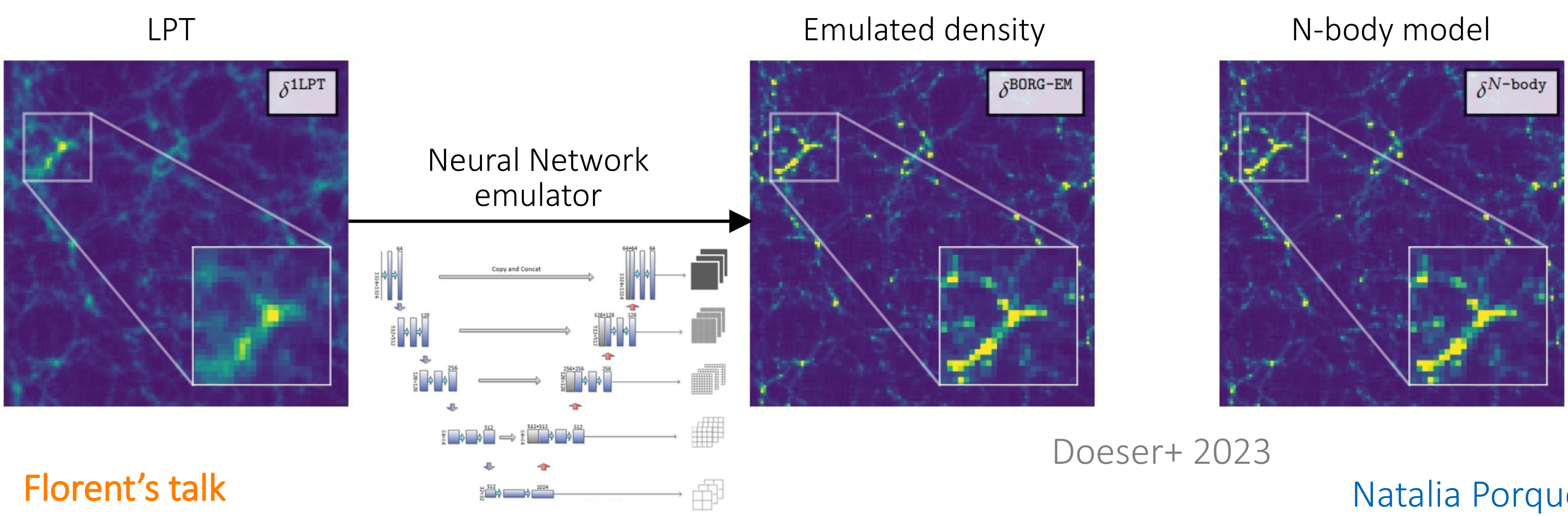
More realistic gravity model in BORG

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



Jasche & Lavaux 2019

N-body emulator: percentage-level agreement (4x faster than PM, 1.6s/evaluation)

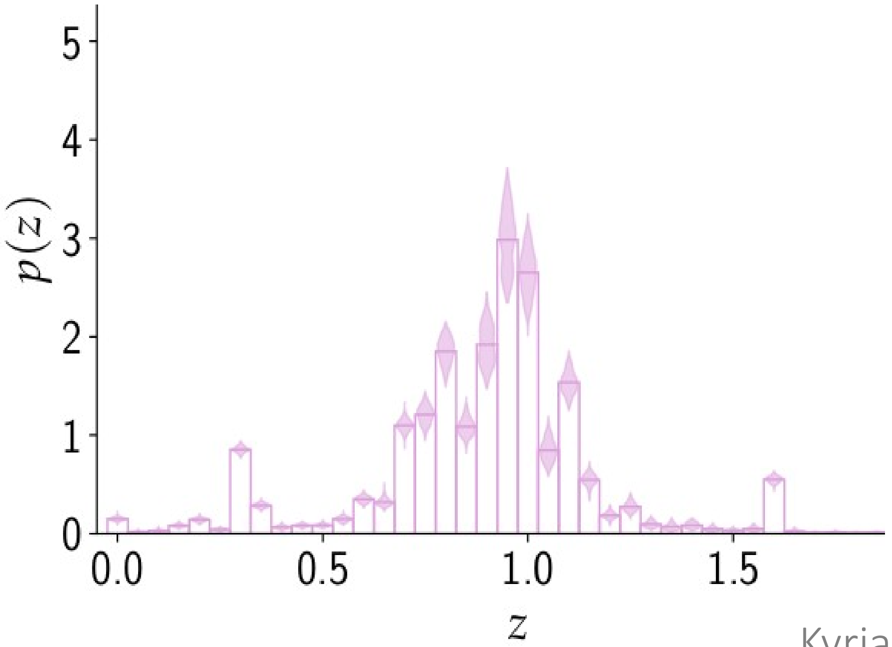


Doeser+ 2023

Florent's talk

Natalia Porqueres

Systematics: Redshift distribution



Leistedt+ 2016;
Kyriacou, Heavens+ in prep

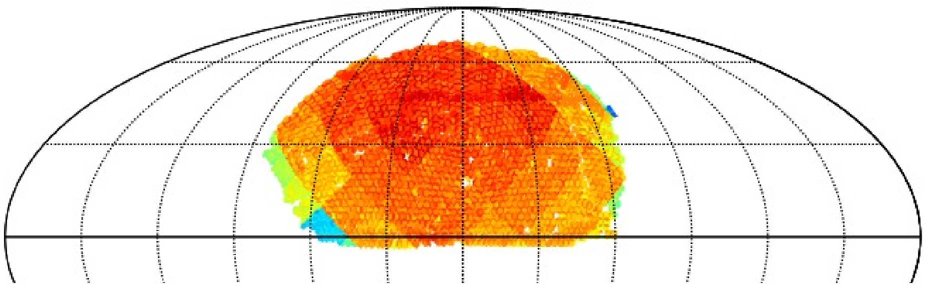
Hybrid physics-informed and machine-learning sampler

Sample n(z) as parametrised histograms

Marginalise over n(z) parameters

Systematics: Unknown contaminations

Unknown contaminations → Physical models do not exist



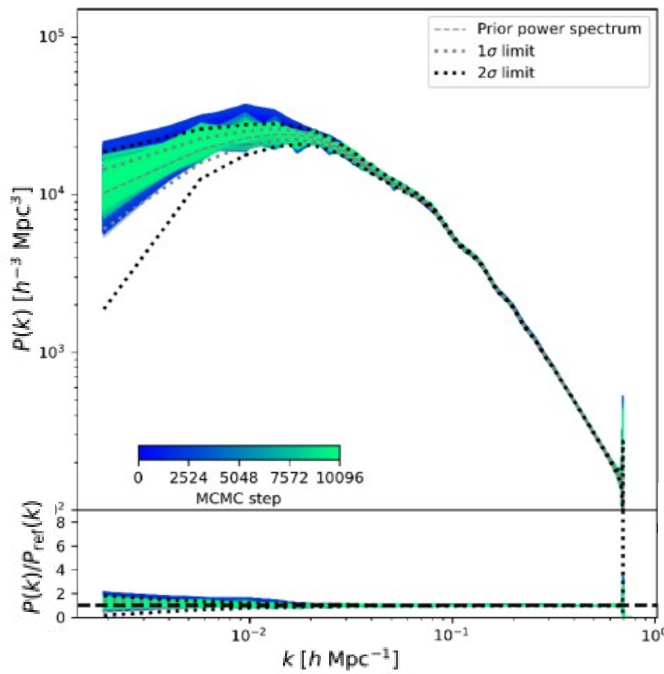
$$\lambda_c = A_c \bar{\lambda}_c$$

$$\downarrow$$

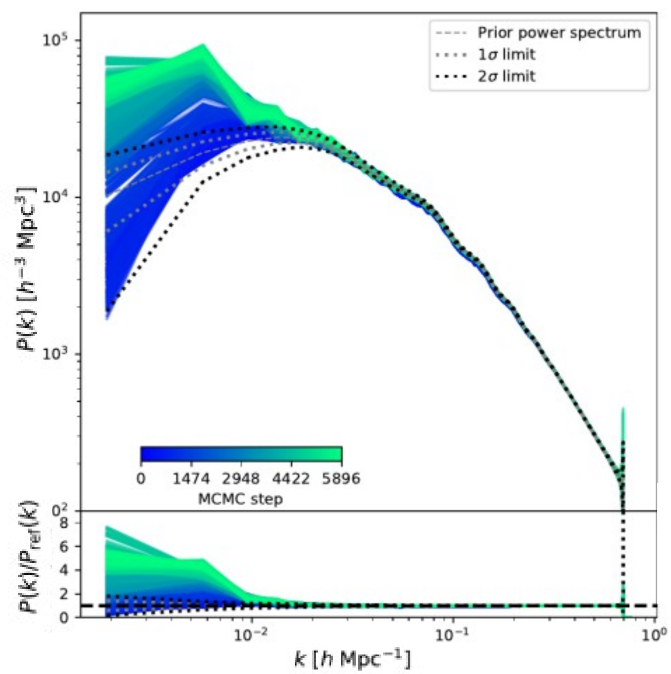
$$P(N|\lambda) = \prod_i \frac{e^{-\lambda_i}(\lambda_i)^{N_i}}{N_i!}$$

NP+ 2018

With unknown contaminations model



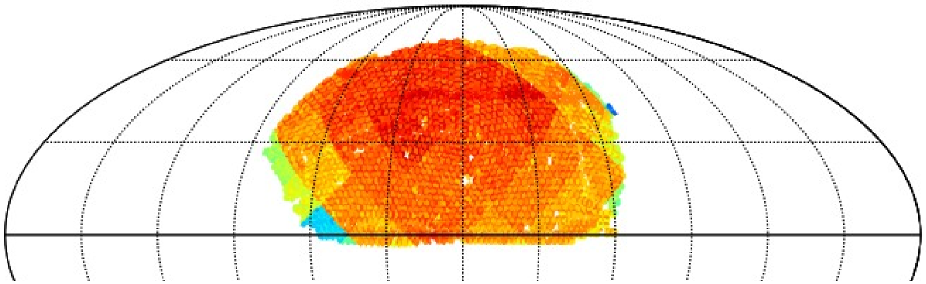
Without the model



NP+ 2018

Systematics: Unknown contaminations

Unknown contaminations → Physical models do not exist



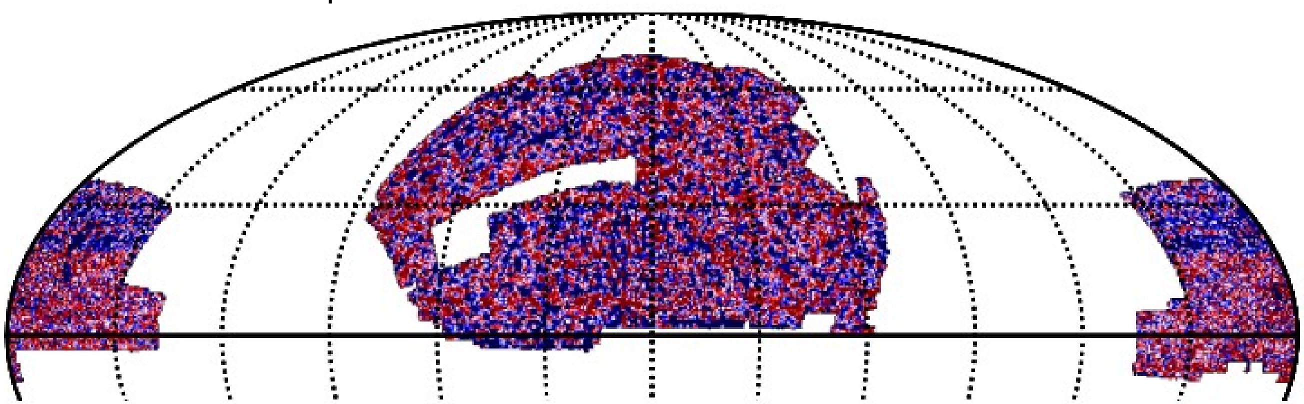
$$\lambda_c = A_c \bar{\lambda}_c$$

$$\downarrow$$

$$P(N|\lambda) = \prod_i \frac{e^{-\lambda_i}(\lambda_i)^{N_i}}{N_i!}$$

NP+ 2018

Maps of the unknown contamination



Lavaux+ 2019

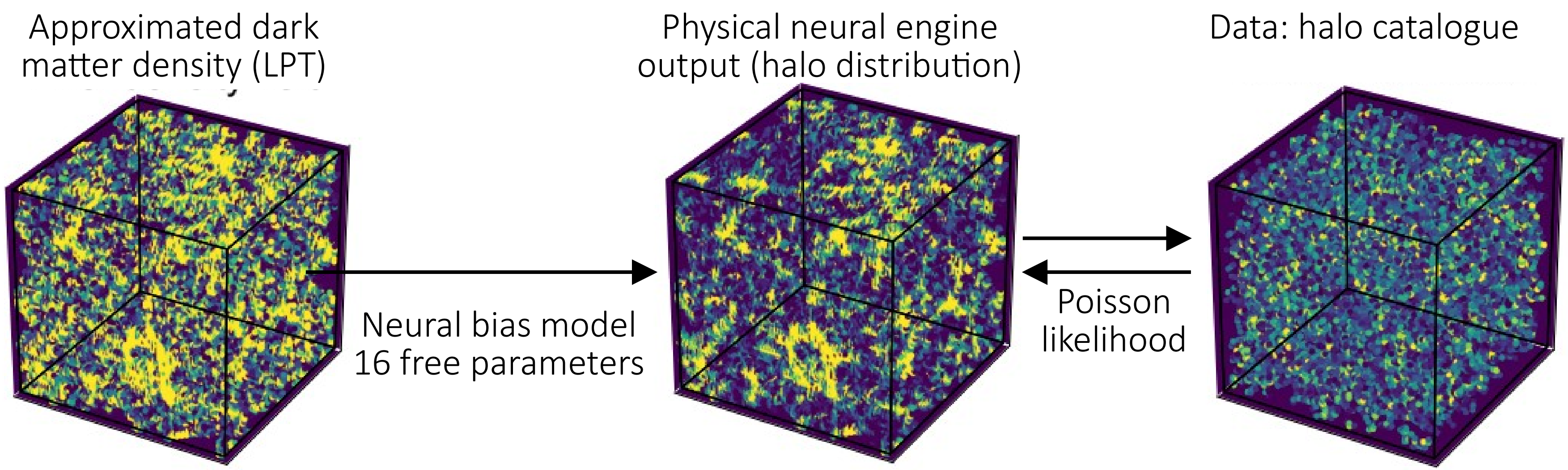
Field-level inference to detect unknown contaminations

Systematics: What if systematics models are not accurate?

Neural networks as bias models to connect data and theory

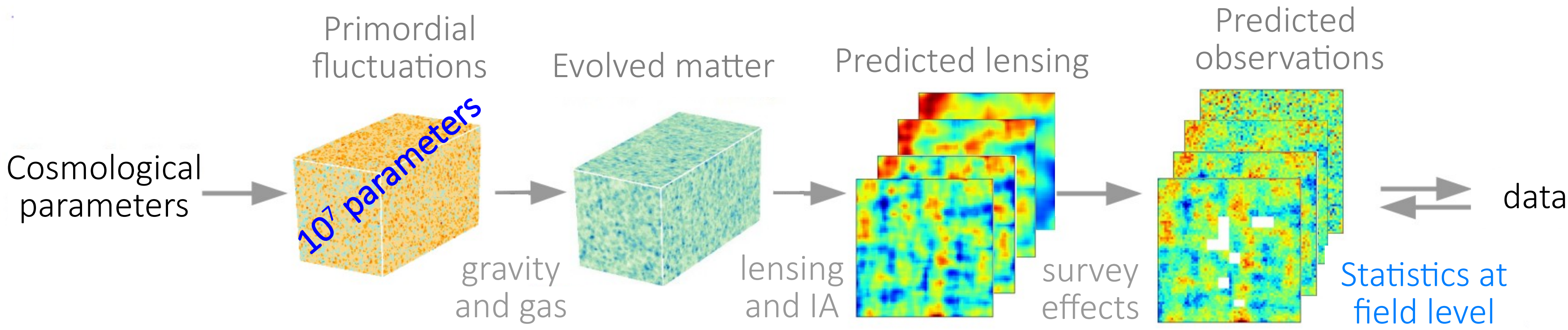
Neural physical engines: Physics-motivated neural networks

- Encode physical symmetries intro NN → less parameters
- Infer NN parameters directly from the data → **no training dataset required!**



(Charnock+ 2019)

Scaling to the Euclid volume

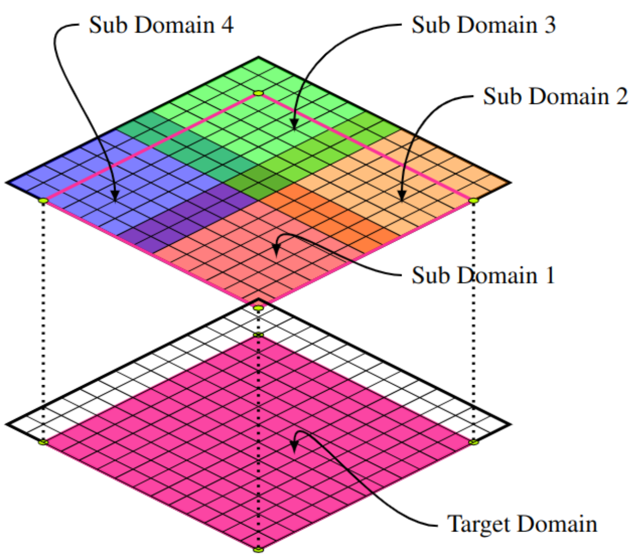


1. Gravity emulator

2. Efficient sampler

Bierkens+2023

3. Sub-volume analysis



Natalia Porqueres

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.
- What’s next?

  Testing with Euclid’s Flagship simulations

  Get ready for first real data application