Forward modeling weak lensing fields with KaRMMa and GANSky

Supranta S. Boruah

In collaboration with Eduardo Rozo, Pier Fiedorowicz, Rafael Garcia, Will Coulton, Gary Bernstein

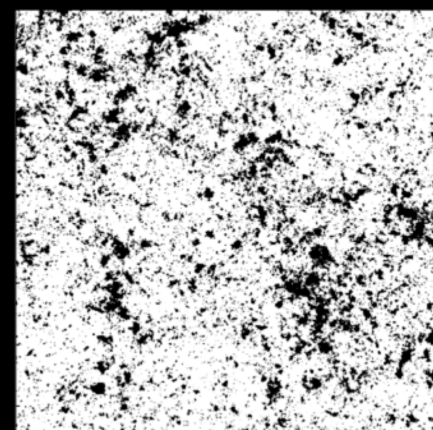
22nd May, 2024

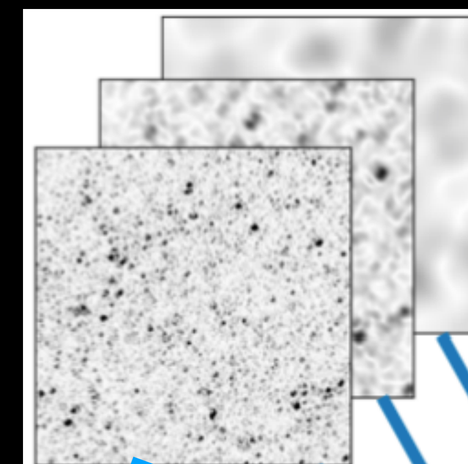
Statistical Challenges in 21st Century Cosmology, Chania

Penn
UNIVERSITY *of* PENNSYLVANIA

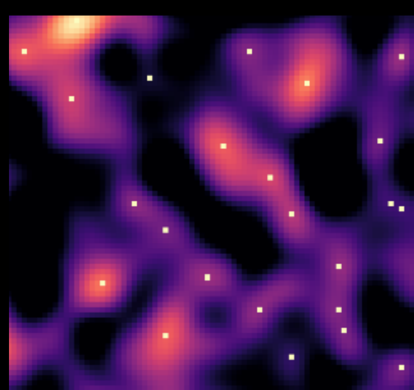
Optimal cosmology inference

Topological estimators

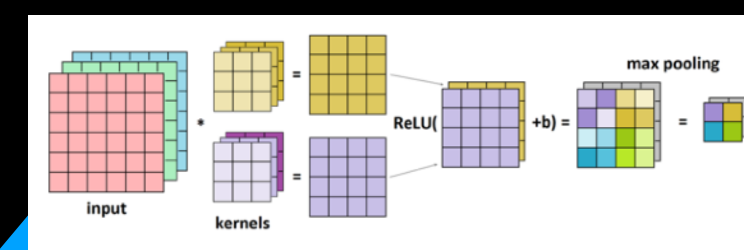




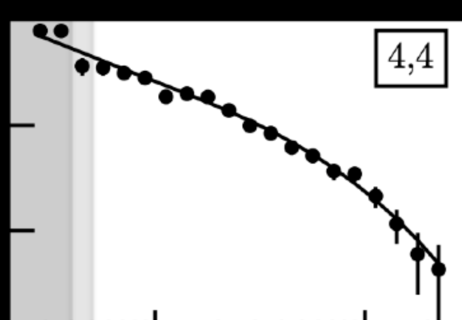
Scattering transform



Peak Counts

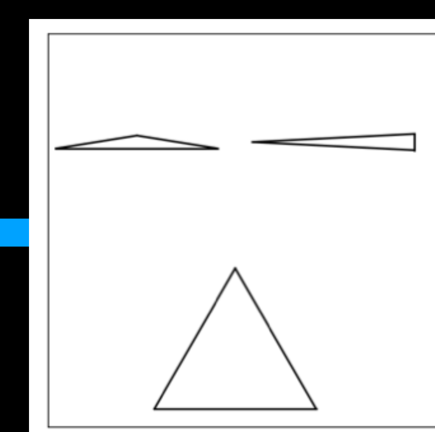


CNN



Power Spectrum

How do we **optimally** extract information from cosmological data?



Bispectrum

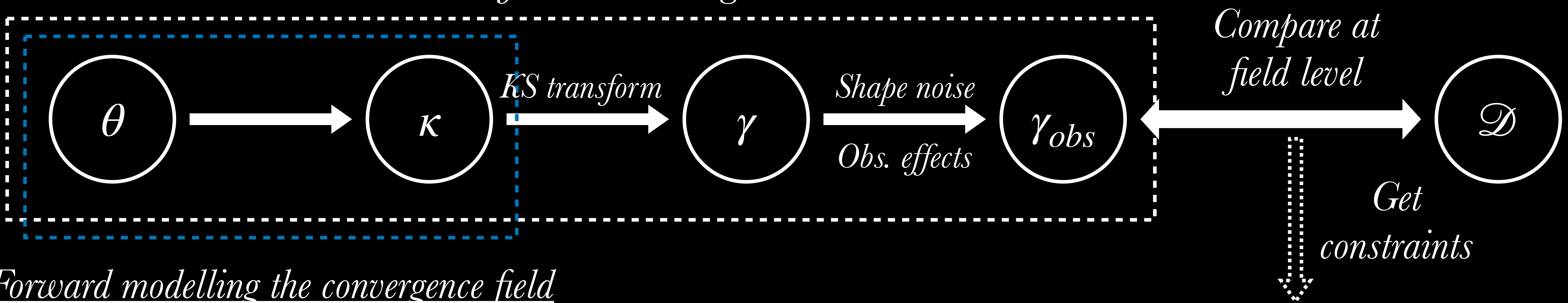
Images: DES, Cheng+ 2022, Li+ 2016, Petri+ 2013,

Solutions for optimal weak lensing inference

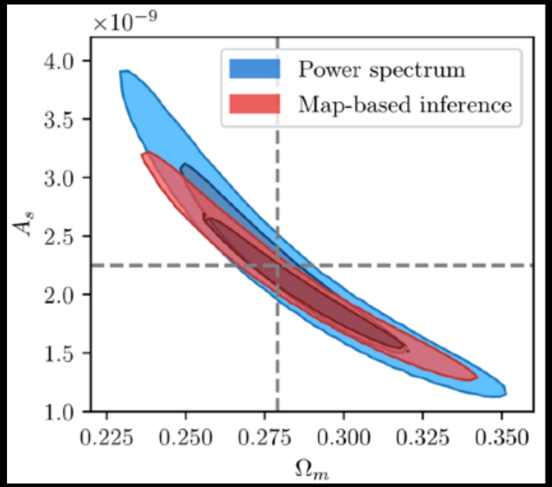
Method	Advantages	Outstanding issues	
Bayesian forward-modelled field-based inference	*This session* - Guaranteed to be optimal provided the correct model - Principled & Interpretable - ….	- Computationally expensive - Sampling over very high-dimensional parameters - Multimodality in the posteriors? - ….	
(inverse) Field-based inference with CNNs	- Given large enough NN, it can extract all information (Universal approximation theorem) - Simpler to implement - ….	- Is the compression optimal? - Interpretability - ….	*Niall, Bhuv, Lucas's talk*
SBI with all possible summary statistics	- Less compute cost - Unbiased results if we have the correct forward model - Relatively easy to implement - ….	- Can we achieve optimal compression in practice - How many statistics to include? - Dependence on mass mapping algorithm / survey mask, etc. - ….	*Judit, Maximilian, Lucas' talk*

Forward modelling weak lensing data

Forward model for weak lensing data



Forward modelling the convergence field
How to model $P(\kappa|\theta)$?



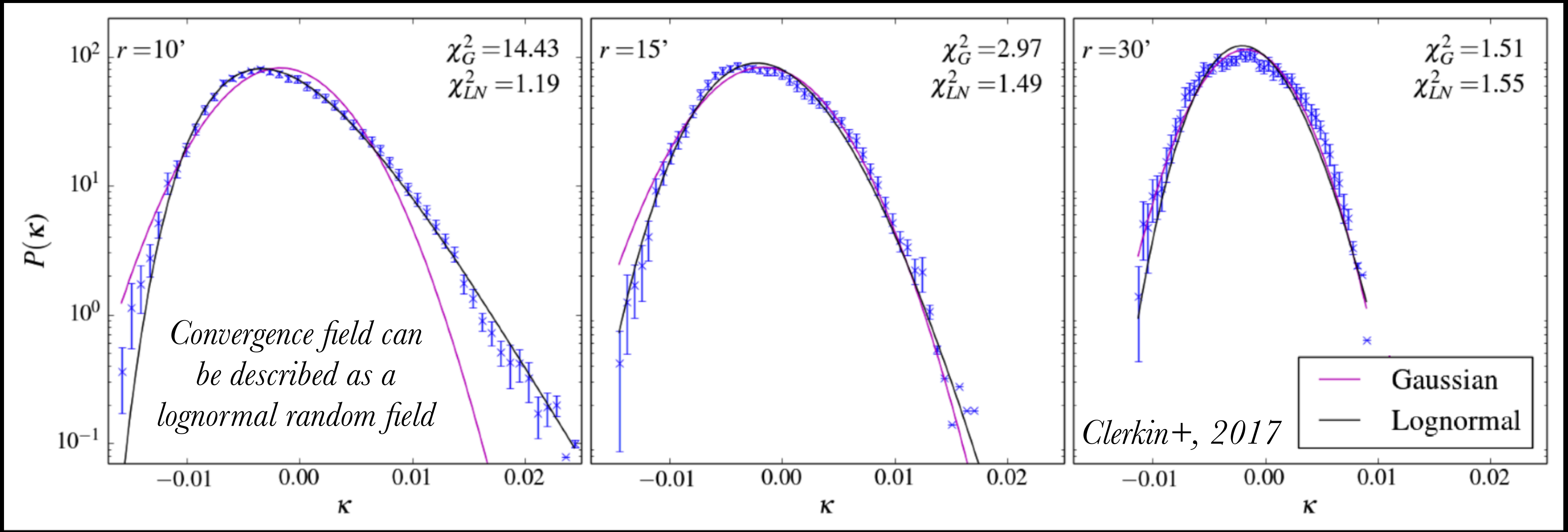
Different forward models of the WL data:

- 3D density + ray tracing *(BORG-WL)*
- 2D convergence field modeling:
  - Gaussian prior *(ALMANAC)*
  - Lognormal prior (*KaRMMa/MIKO*)
  - ML-based prior

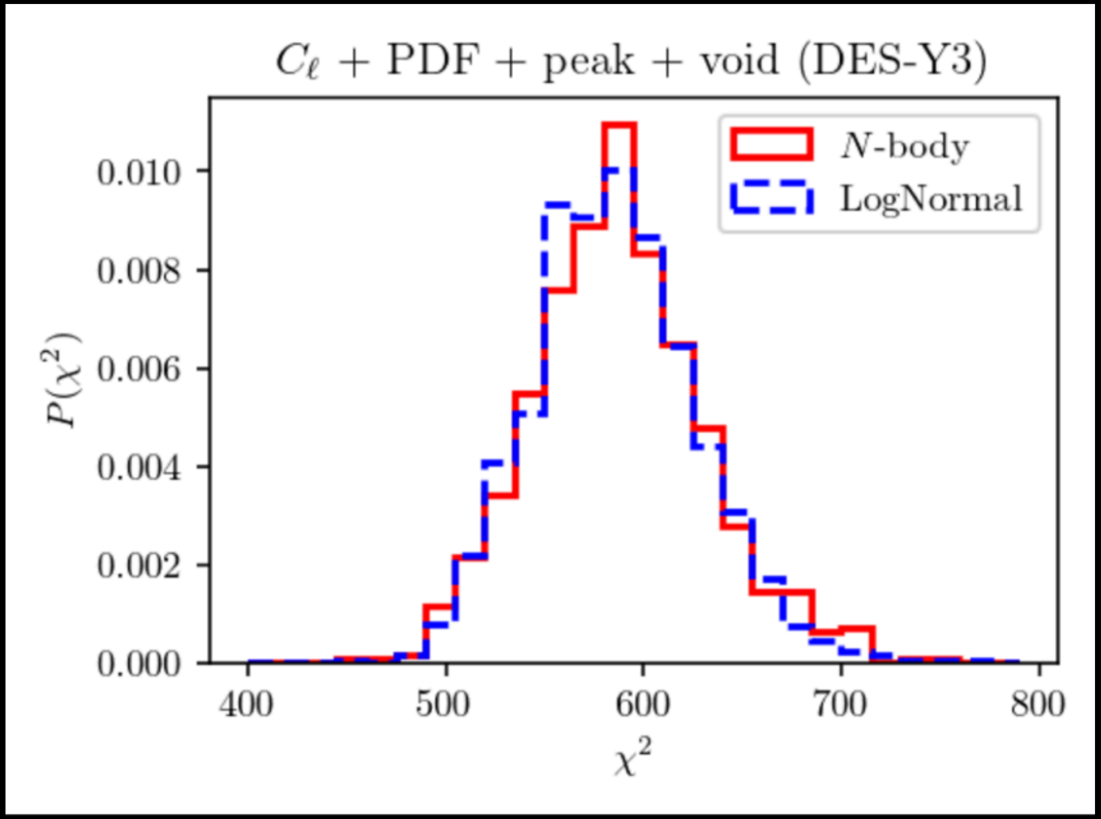
This talk

- ……

Lognormal model for convergence field



- Lognormal prior is analytic, $\kappa = e^y - \Lambda$, where, $y \sim \mathcal{N}(\mu, \sigma^2)$
- Correct 2-pt and 1-pt function
- Adds non-Gaussian information



KaRMMa — map inference with lognormal priors

Sample mass maps from the posterior: $P(\kappa|D) \propto P(D|\kappa)P(\kappa)$

Assume lognormal prior on κ

***Other features*:**

- Full-sky forward model *(w/ HEALPIX maps)*
- Tomographic mass mapping (Include cross-correlations in the prior)

KaRMMa

KaRMMa - Kappa Reconstruction for Mass Mapping

KaRMMa is a library for curved-sky mass map reconstruction using a lognormal prior. For more information, see our paper.

Producing Bayesian mass maps with DES-Y3 weak lensing data

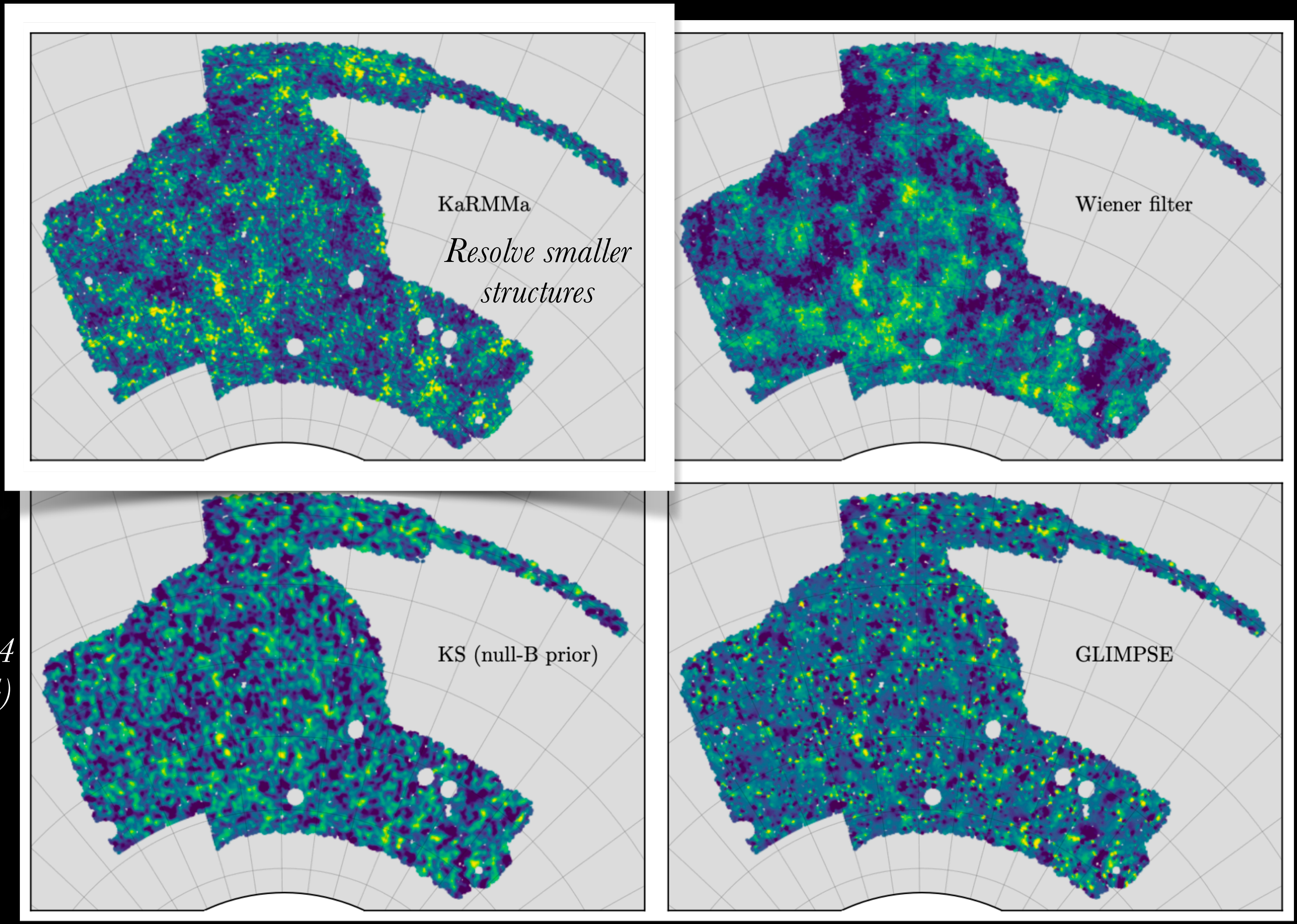
You can use this repository to run KaRMMa on DES-Y3 weak lensing data. The DES-Y3 data used to create KaRMMa mass maps are included in this repository here.

Publicly available on Github. Reproduce KaRMMa DES-Y3 mass-maps yourself!

https://github.com/Supranta/KaRMMa

DES-Y3 KaRMMa mass maps

Used KaRMMa to produce better mass maps with DES-Y3 data



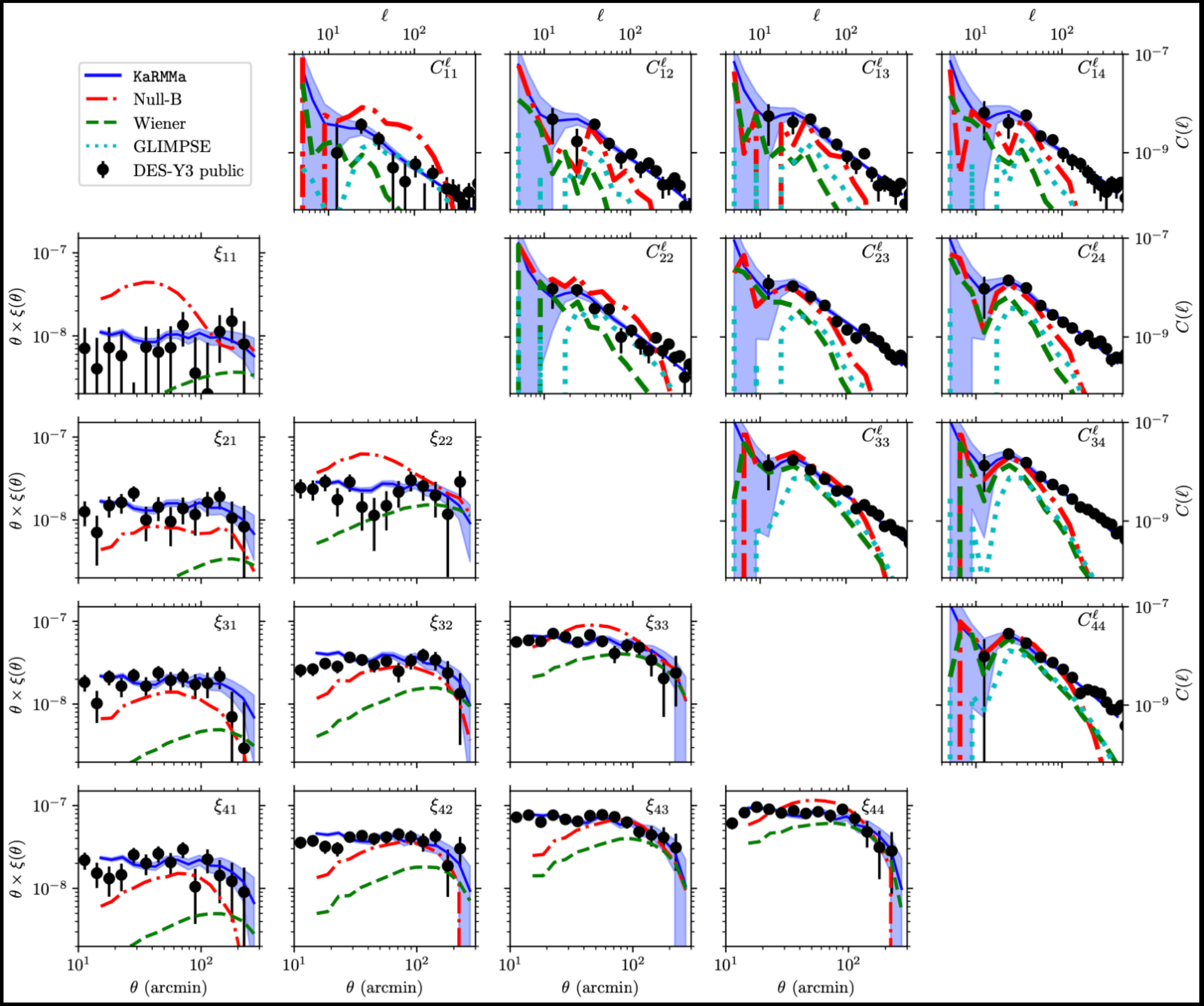
Boruah+ 2024 (2403.05484)

DES-Y3 KaRMMa mass maps

Existing mass mapping methods do not give correct statistics

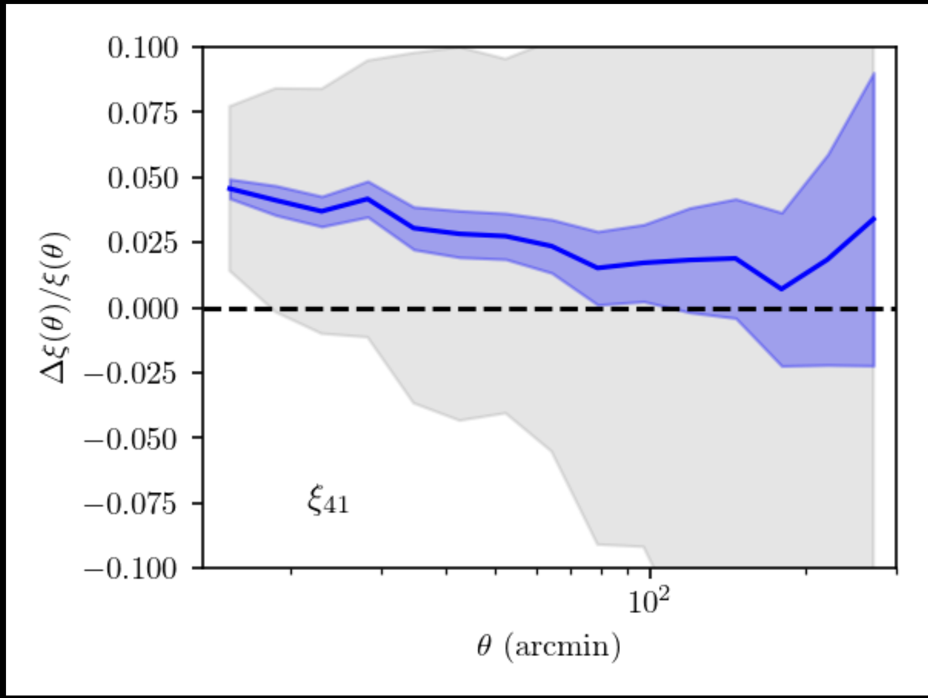
KaRMMa mass maps have the expected two-point functions

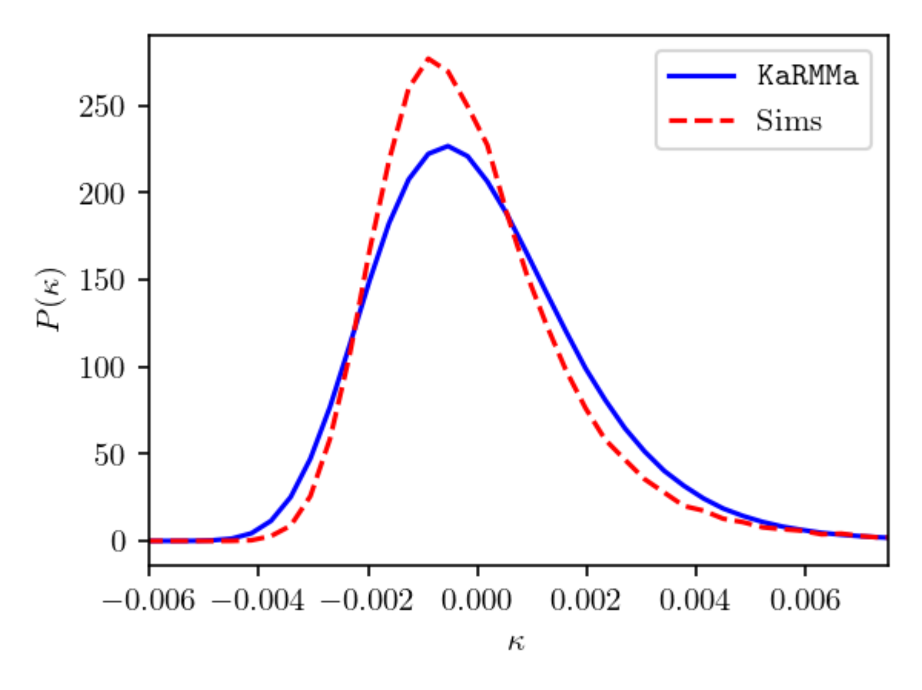
Boruah+ 2024 (2403.05484)



Drawbacks of KaRMMa

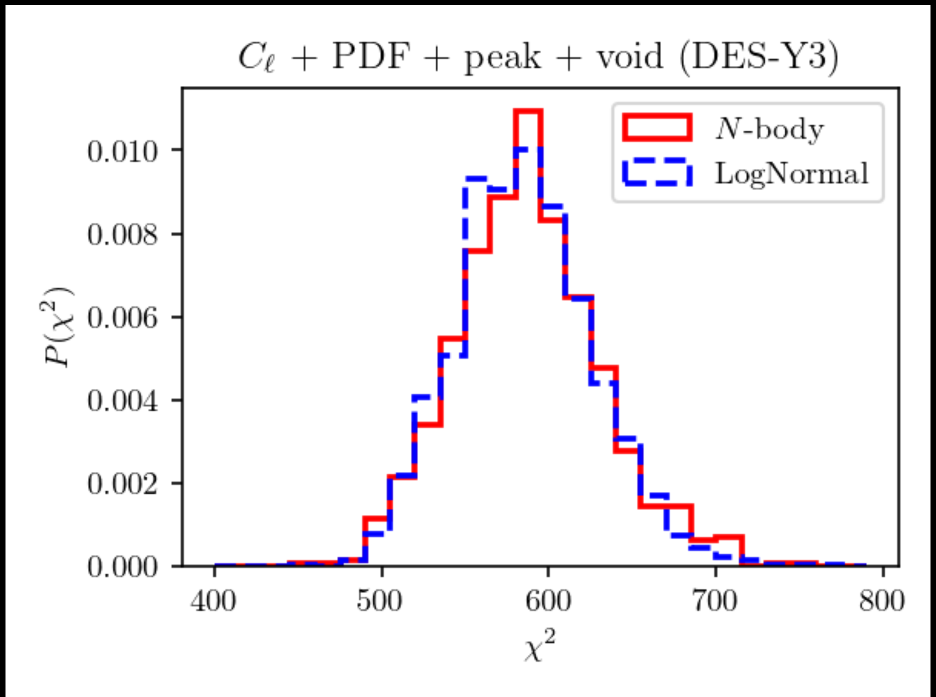
$\mathcal{O}(5\%)$ *bias in the recovered power spectrum*





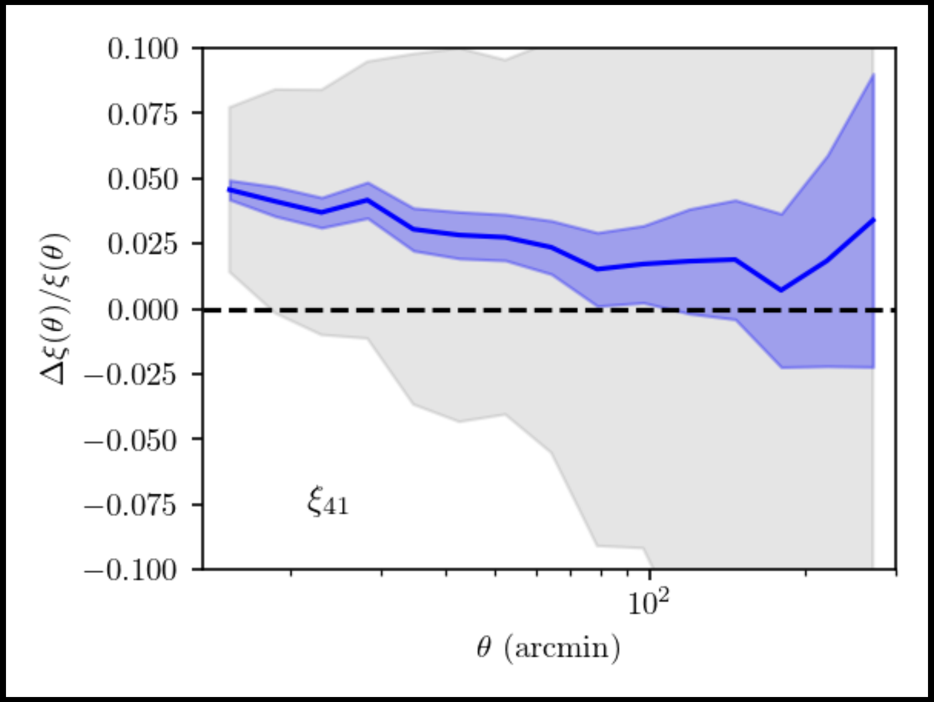
Bias in the recovered 1-pt function

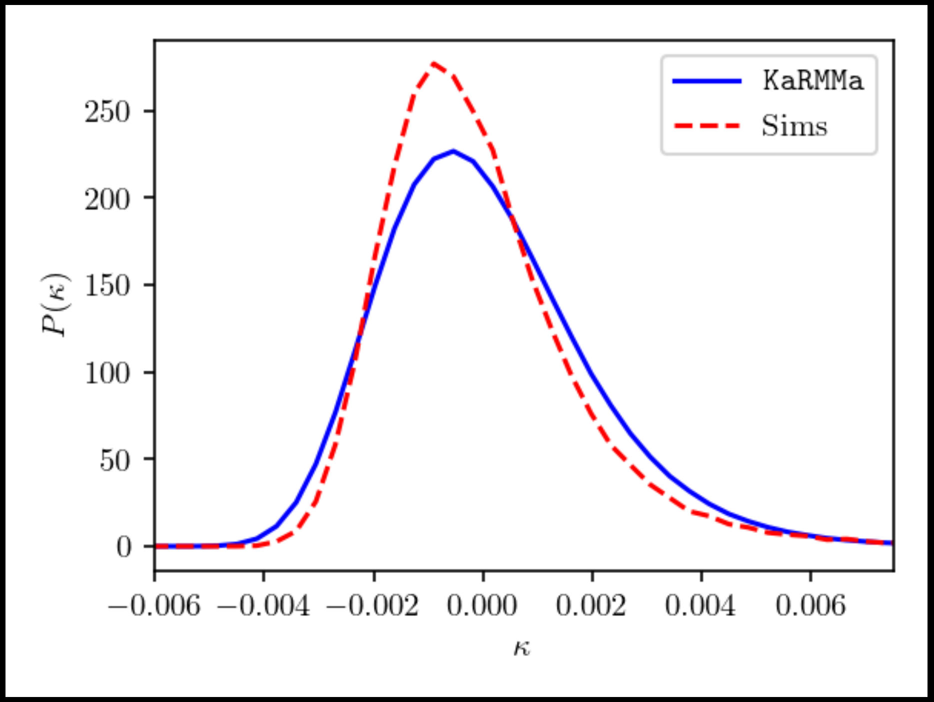
Biases not significant for DES-Y3



Drawbacks of KaRMMa

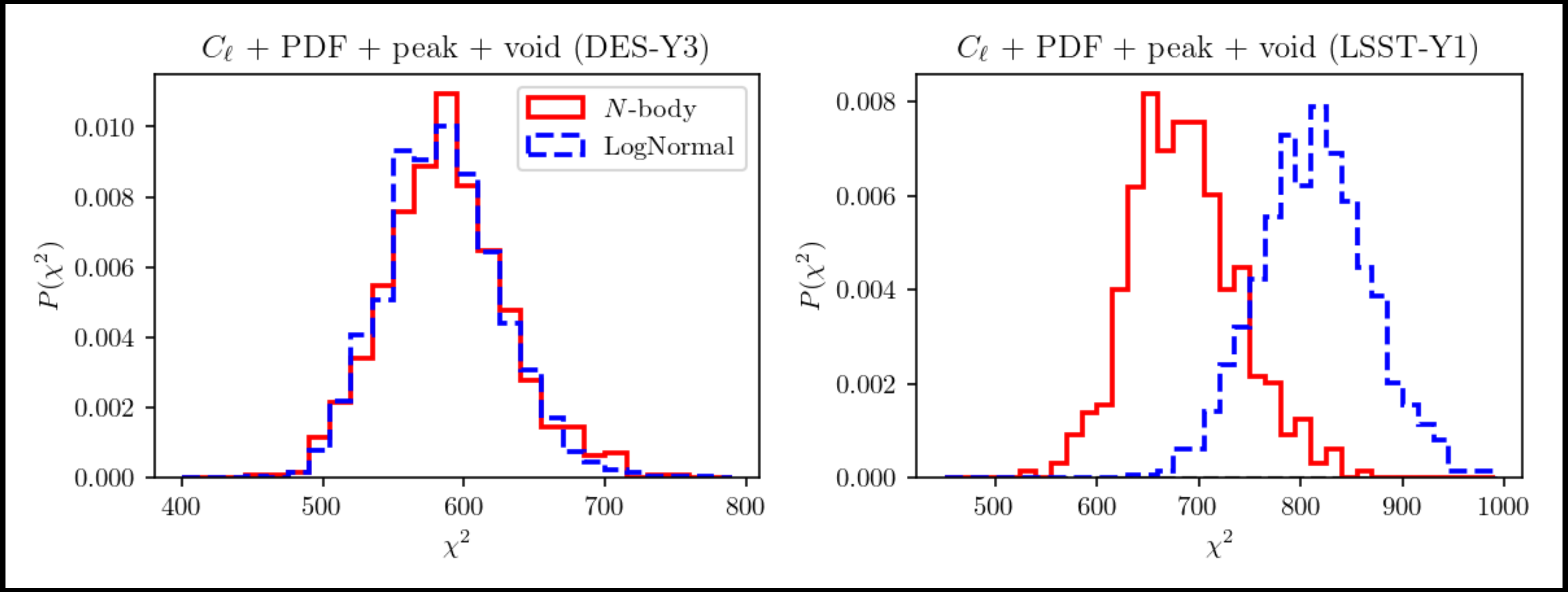
$\mathcal{O}(5\%)$ bias in the recovered power spectrum





Bias in the recovered 1-pt function

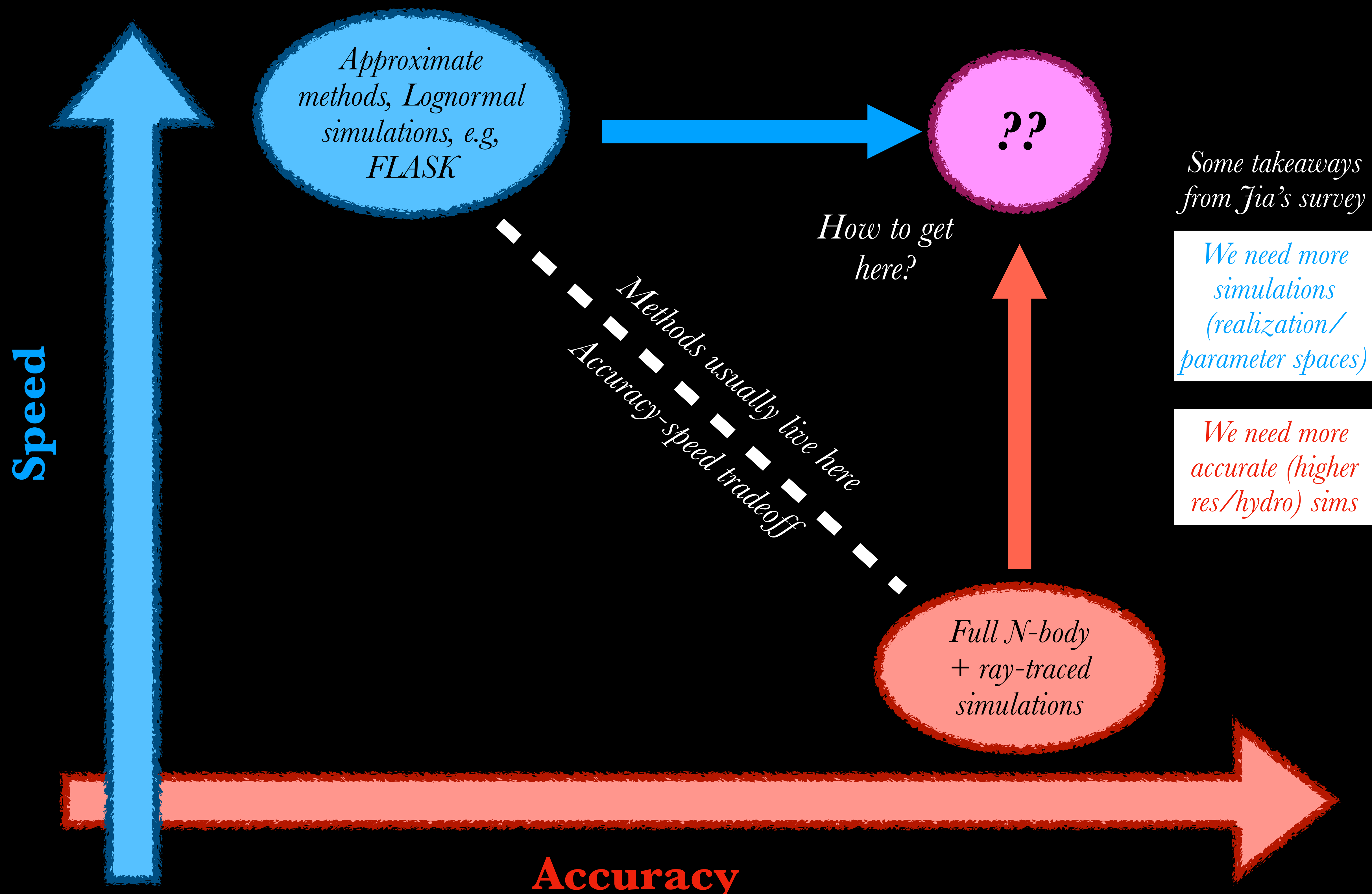
Biases not significant for DES-Y3



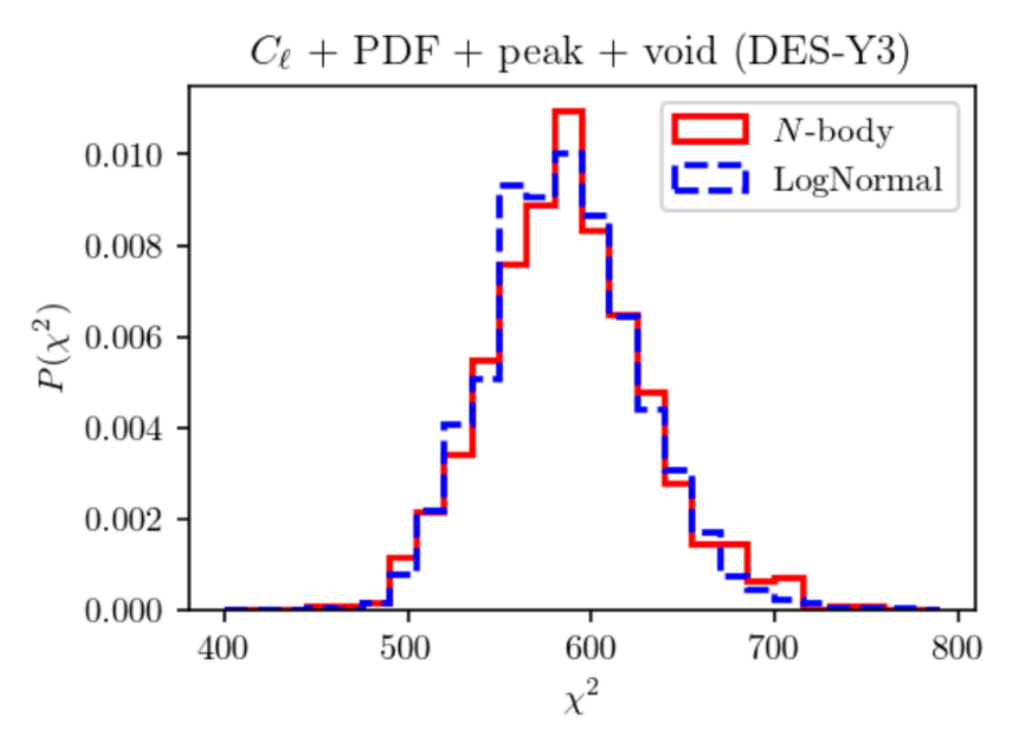
But lognormal model not sufficient for Stage-IV surveys

We need a better forward model for Stage-IV surveys

WL simulations for Stage-IV surveys

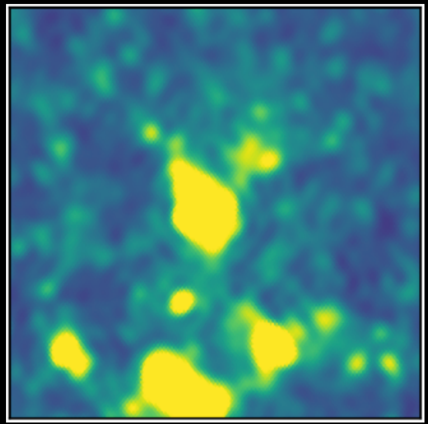


Improving the lognormal prior using ML

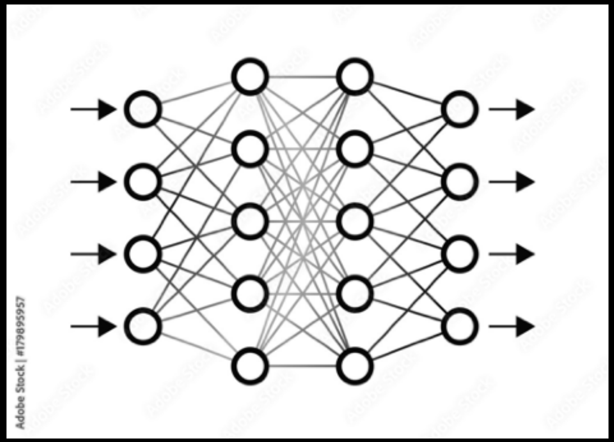


Lognormal model already described the convergence field very well.

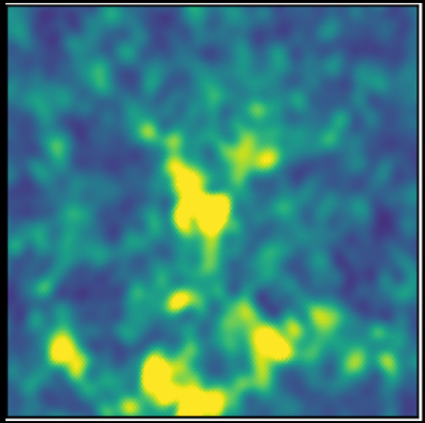
Can we make small changes to the lognormal maps to emulate simulation quality maps?



Lognormal



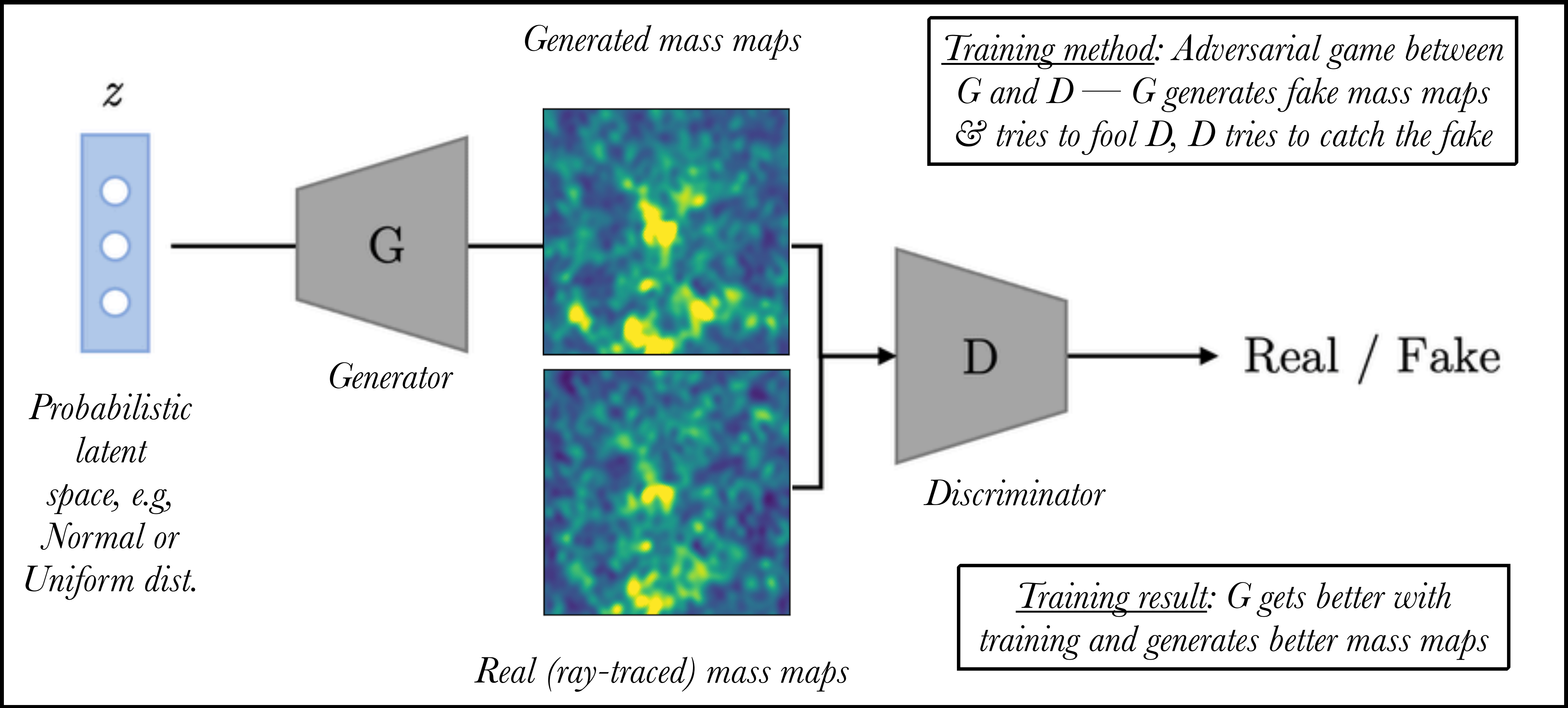
ML blackbox



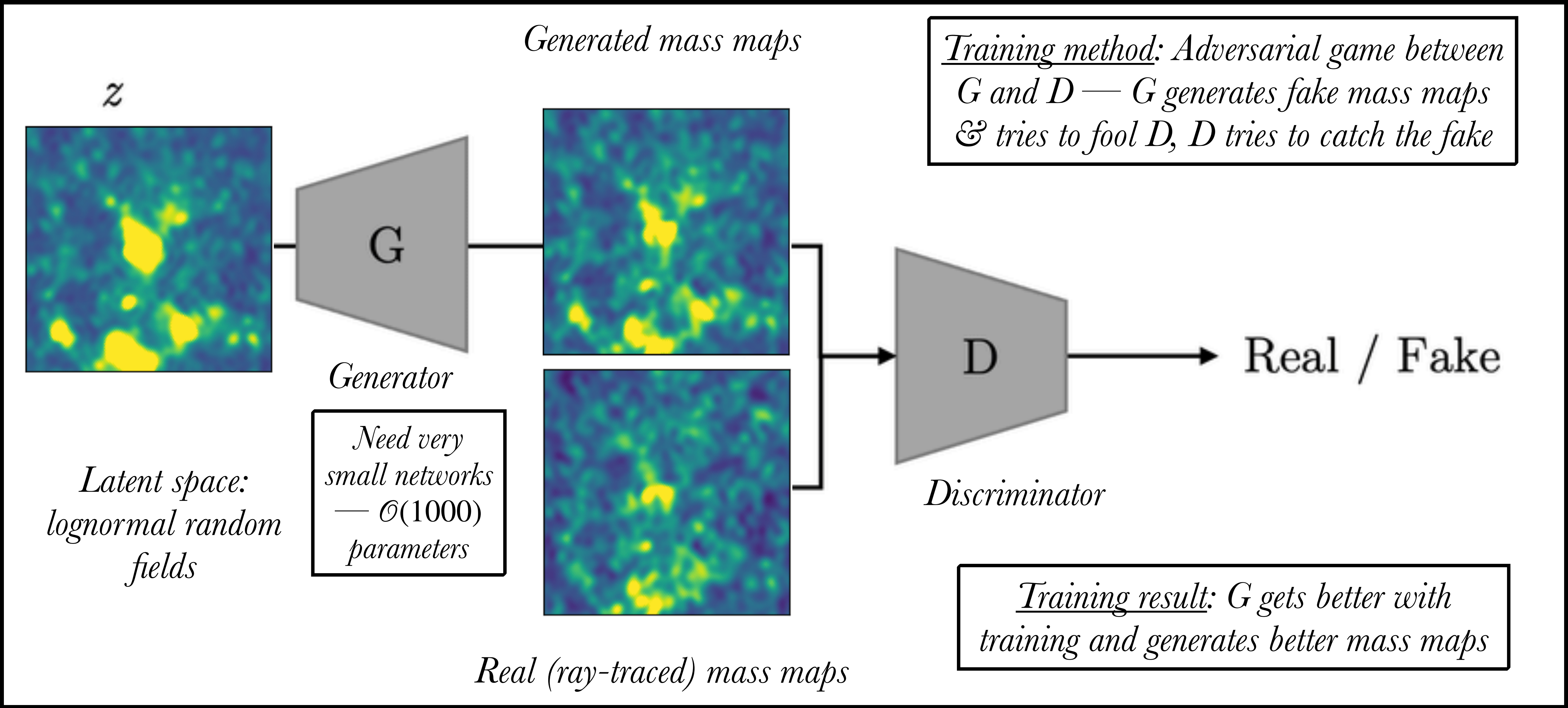
Full simulation

<u>Main idea</u>: Use ML to learn the mapping from the lognormal simulations to more accurate simulations!

Generative adversarial networks (GAN)

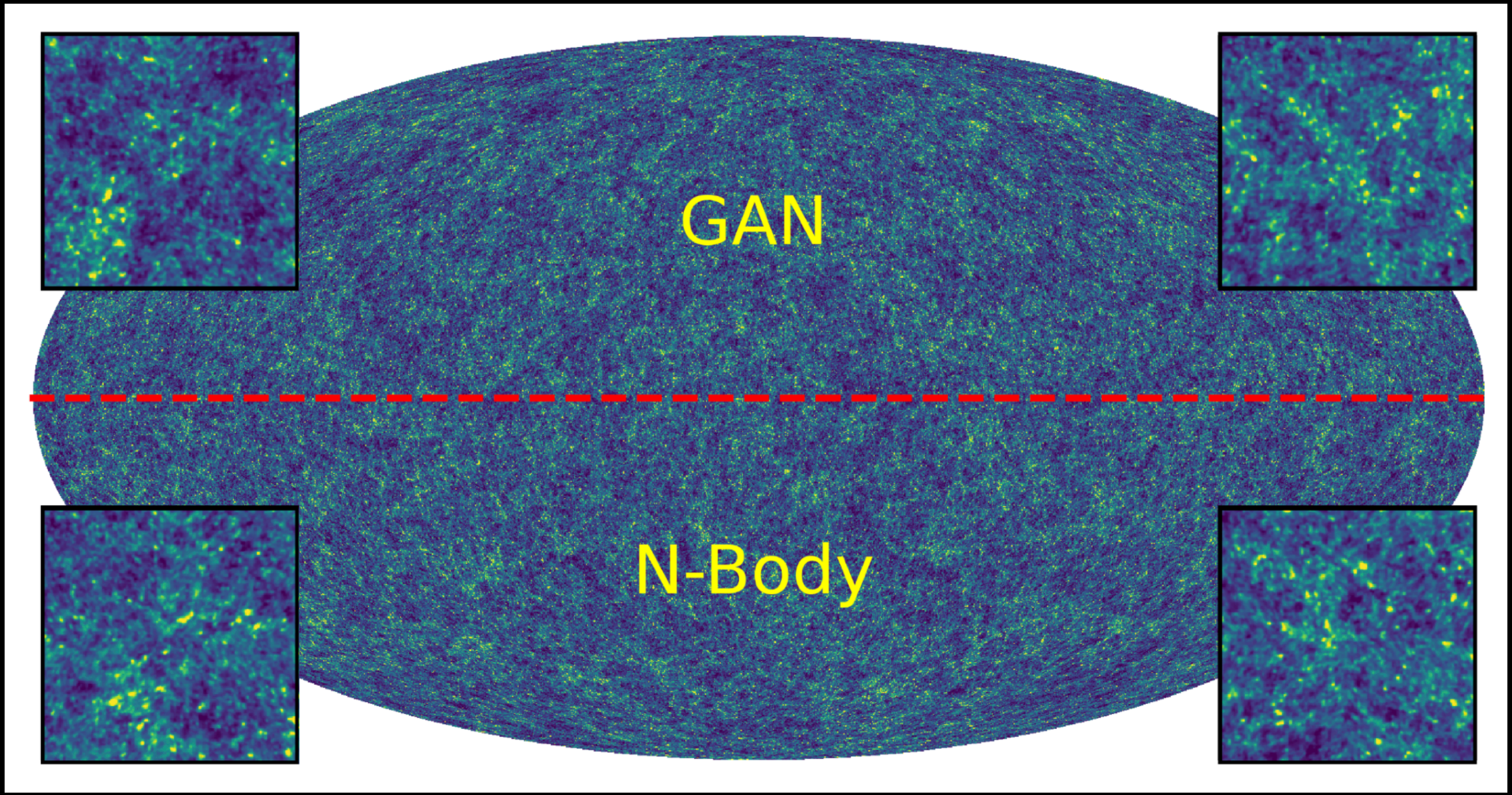


Generative adversarial networks (GAN)



Realistic latent space — lognormal simulations — ML needs to learn the small scale redistribution

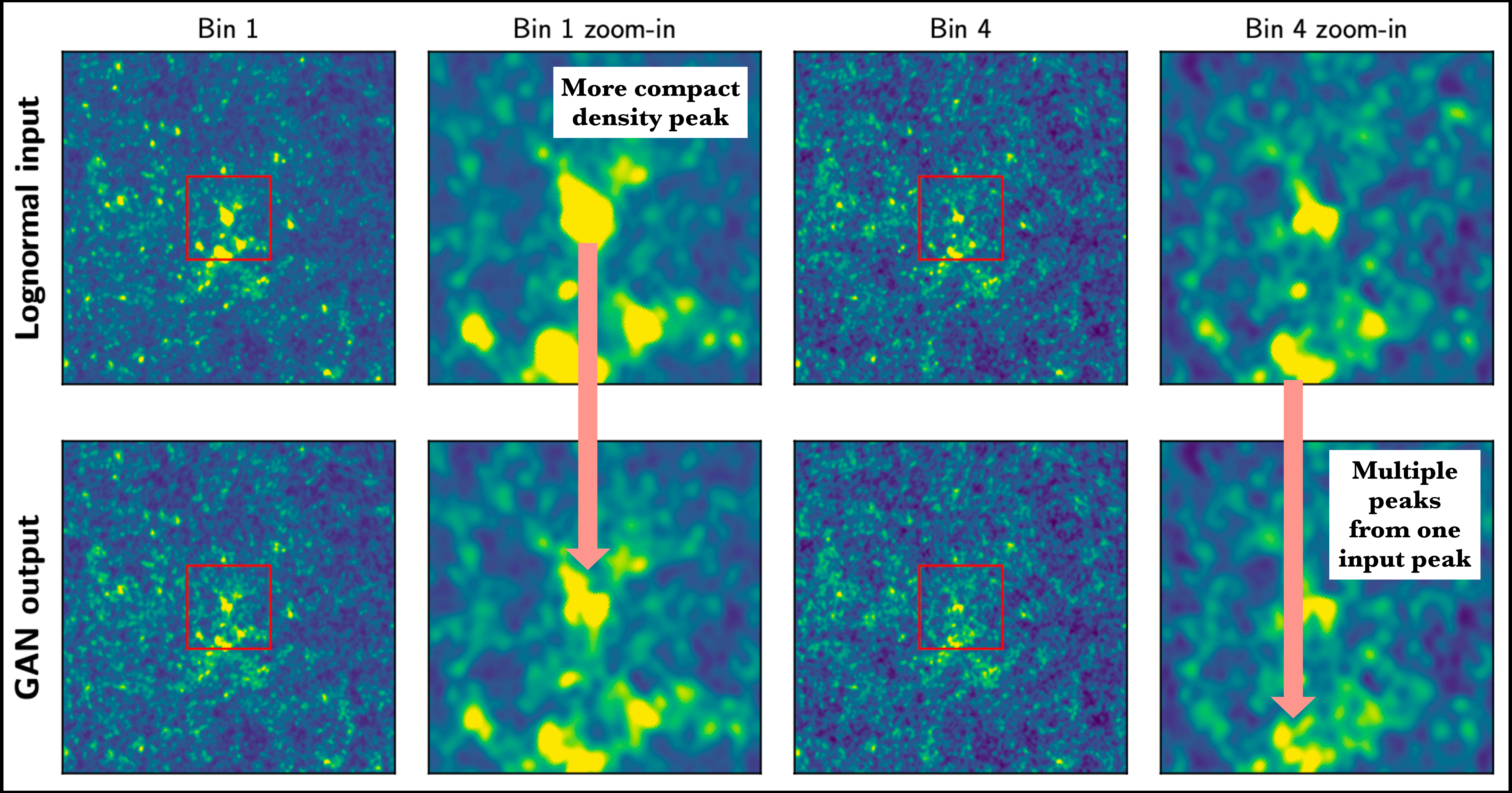
Curved sky WL maps w/ GANSky



We produce HEALPIX maps with GANSky

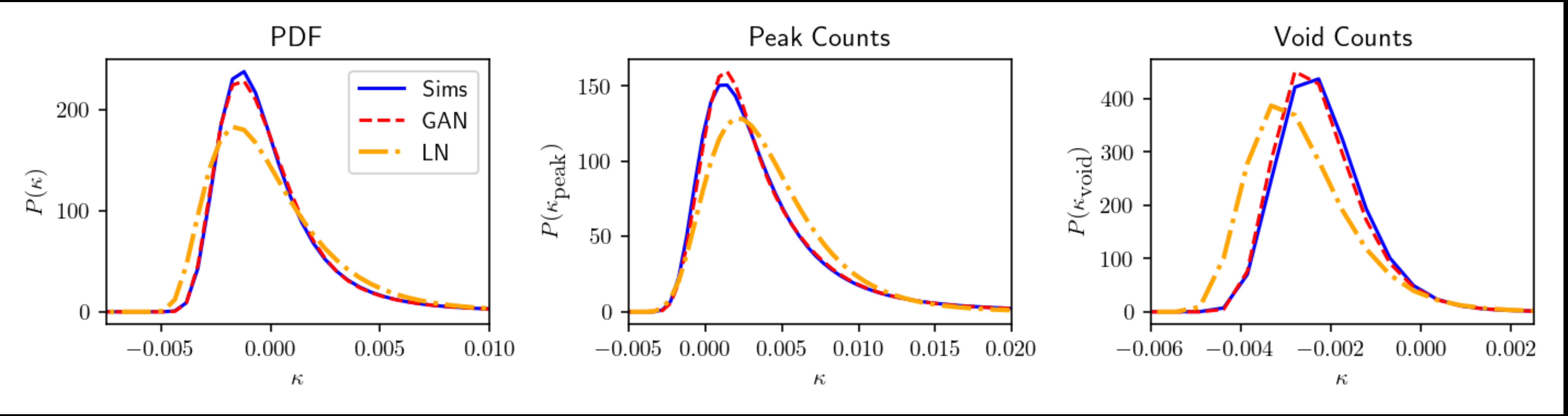
Explainable & physics-informed machine learning

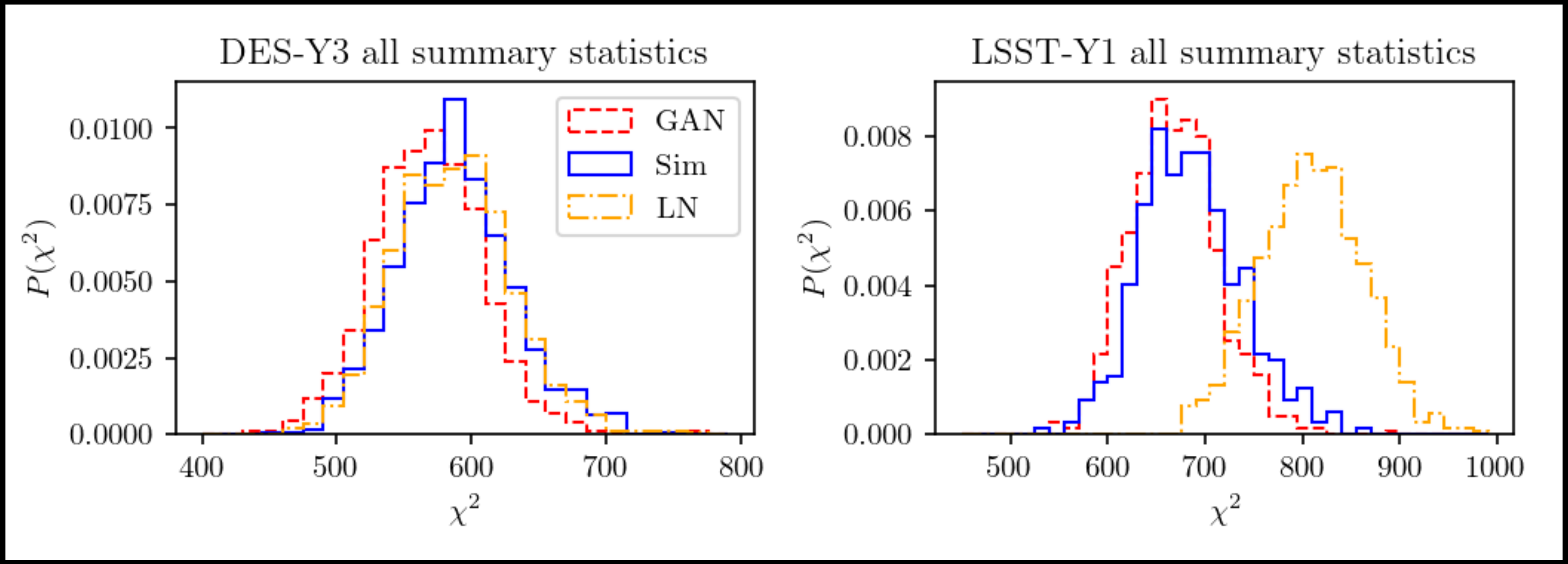
Rotational equivariance built into the neural network



The GAN only makes small changes to the map,
i.e, output maps are almost similar to input maps

GANSky produces correct non-Gaussian summary statistics

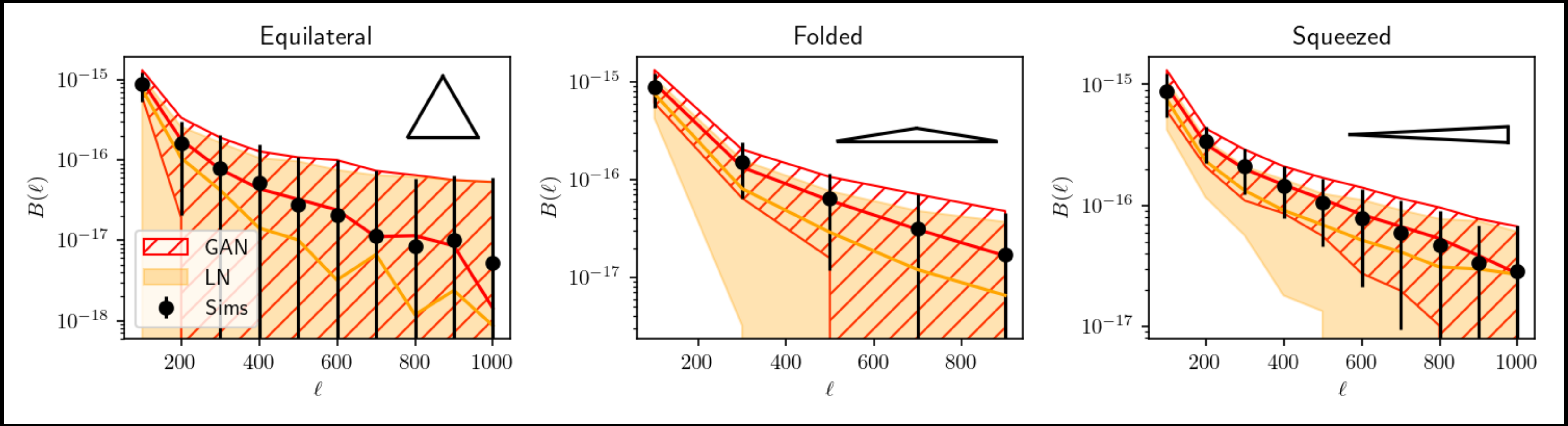




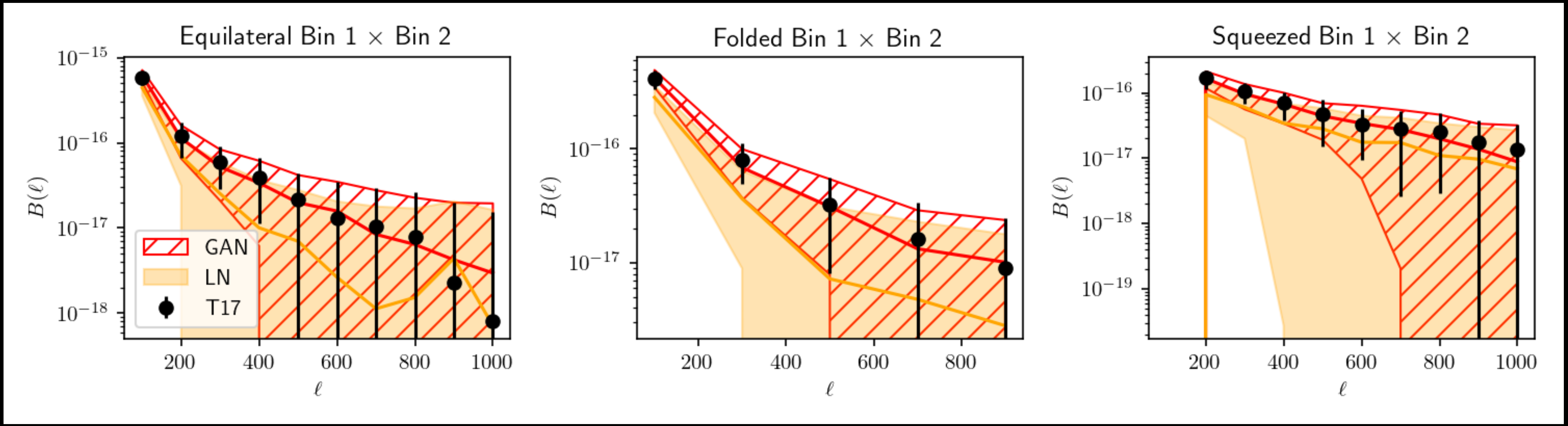
We can reproduce the non-Gaussian summary statistics of full ray-traced simulations with GANs — at a fraction of the cost

GANSky maps have the second moment of the non-Gaussian stats, even at LSST-Y1 level noise

GANSky produces correct non-Gaussian summary statistics



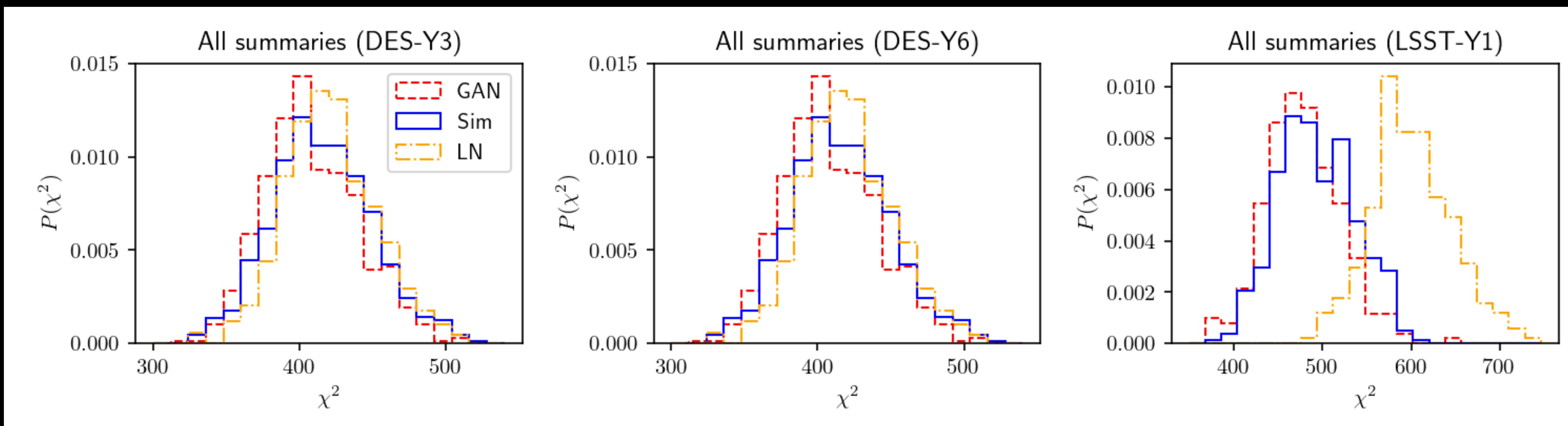
GANSky reproduce the bispectra of these maps correctly



Including correct cross-bin bispectra

Training with only 1 simulation!

One full sky simulation have many many patches!



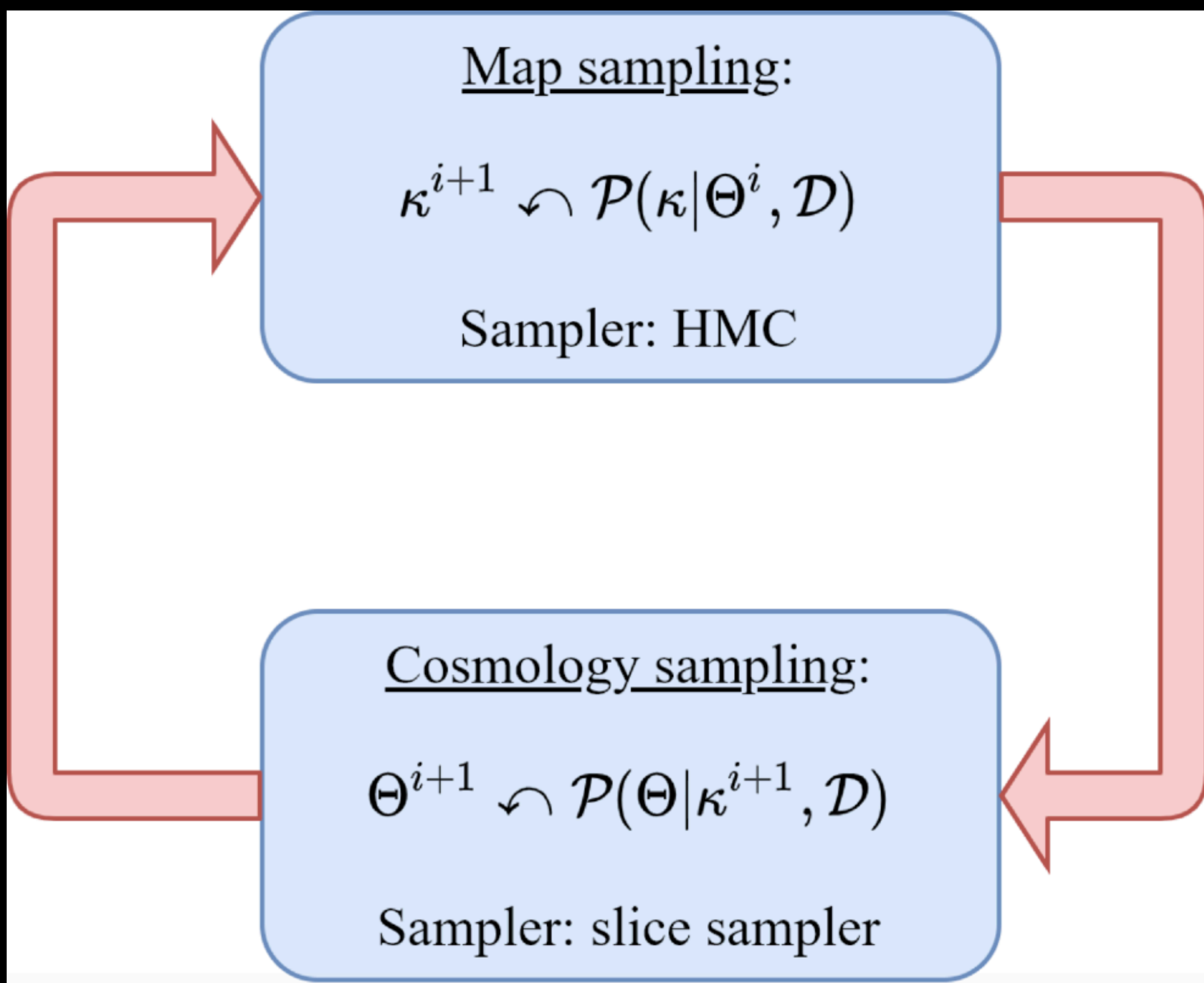
Training w/ only 1 full-sky simulation, achieves accuracy of ray-tracing sims

Our method is not data hungry — interpretable latent space does the trick!

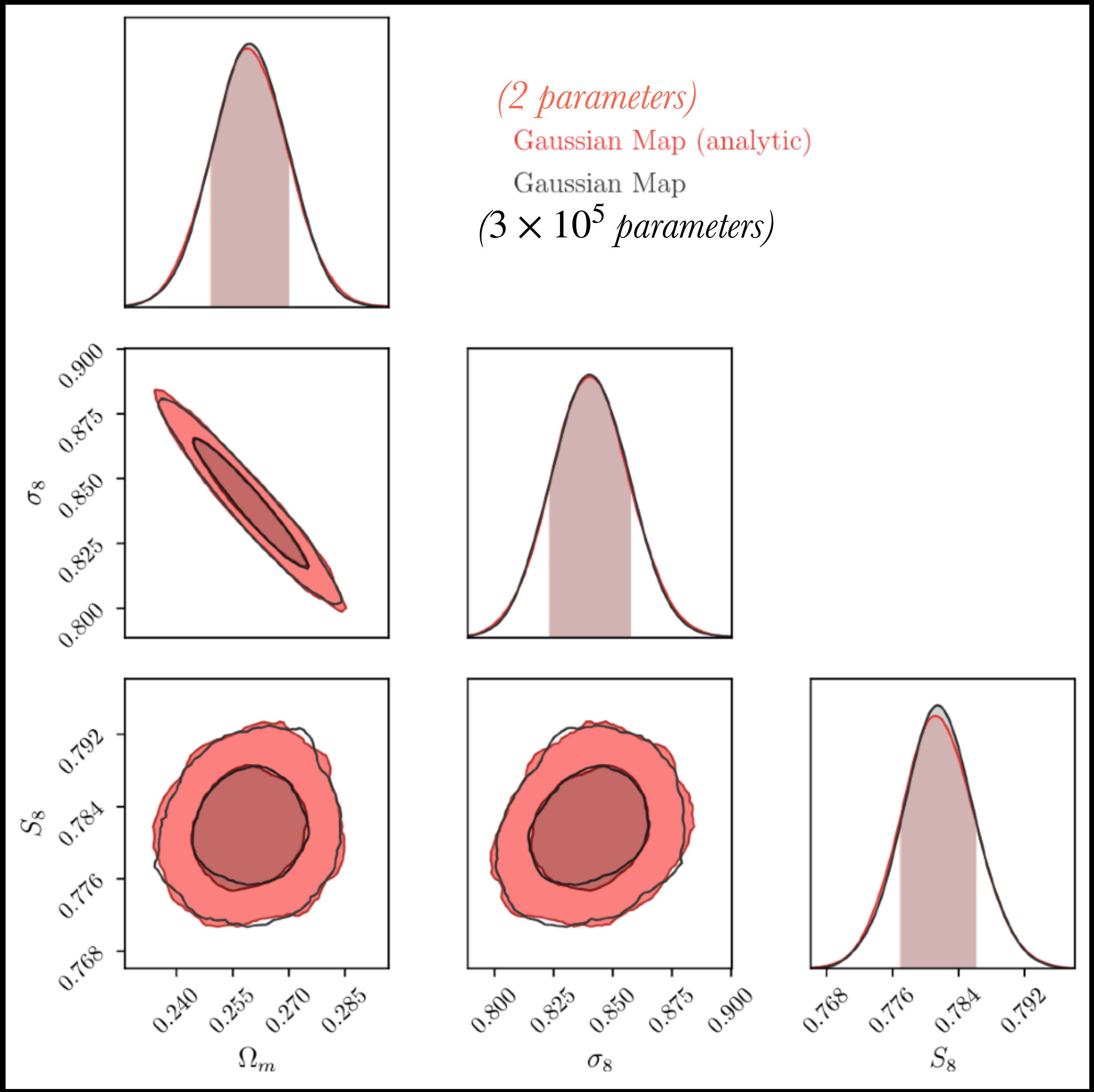
Will be useful for computing covariances of NG stats for Stage-IV surveys — hugely reduce the computation cost

Map-based cosmology inference

Extensive code testing

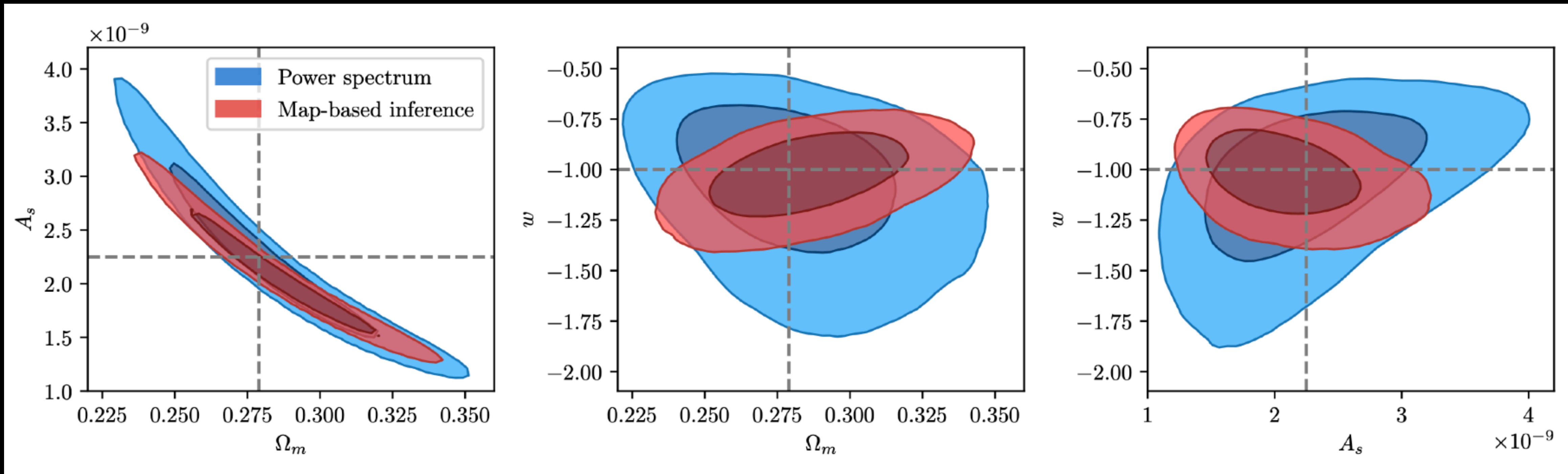


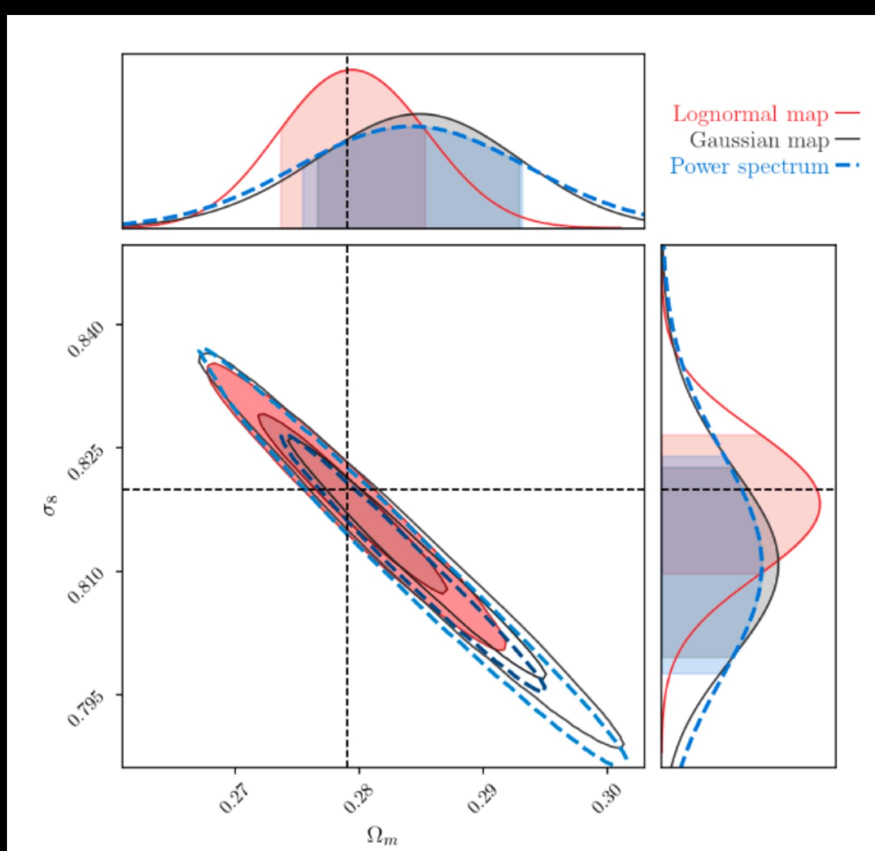
Two-step sampling for cosmology posteriors



Tested the correctness of cosmology posteriors for simple analytic cases

Map-based cosmology inference





<u>Forecast for LSST-Y1</u>*: Upto 2.5x improvement in the constraints on the dark energy equation using map-based inference (equivalent to statistical power of 6 LSST-like surveys)*

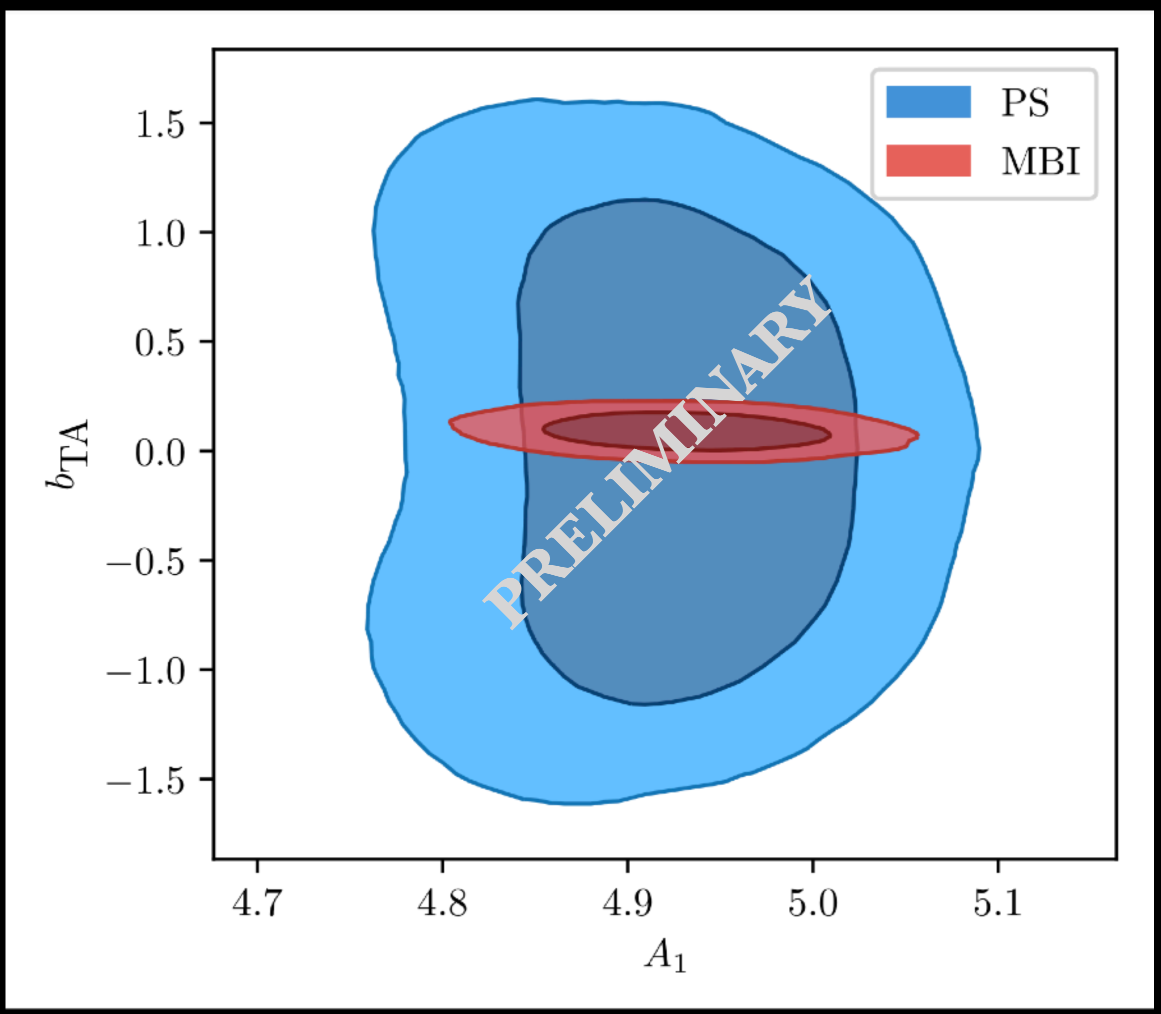
However, no significant gain for ΛCDM analysis! Some degeneracy breaking!

Field-level intrinsic alignment constraints

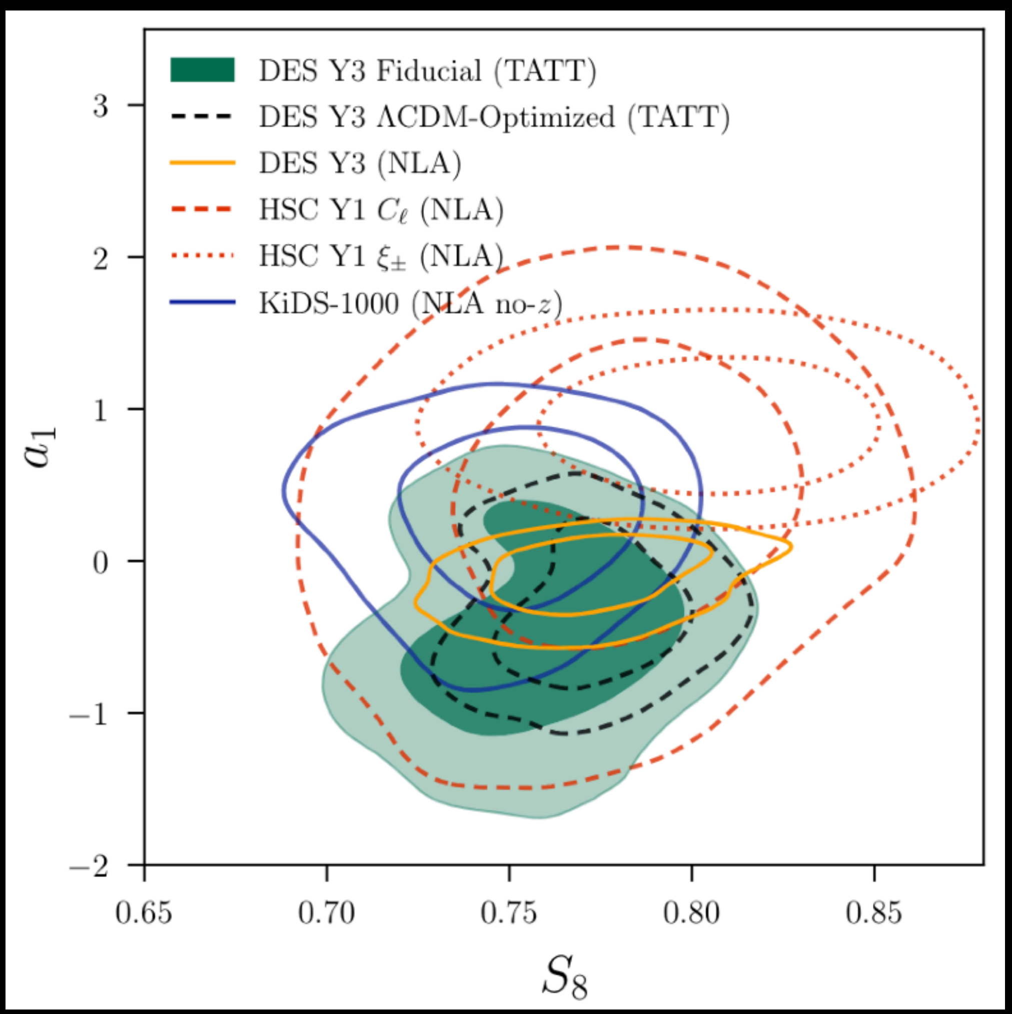
$$\bar{\gamma}^{\mathrm{IA}}_{ij} = A_1 s_{ij} + A_{1\delta}\delta s_{ij} + A_2 \sum_k s_{ik}s_{kj} + \cdots,$$

These terms are non-Gaussian even if δ is Gaussian

We expect to get stronger IA constraints using field-level inference



5x stronger constraints on higher order IA parameters using MBI



TATT parameters dominate the error budget for DES-Y3.

For LSST, IA expected to contribute $\gtrsim 40\,\%$ of the error budget

Summary

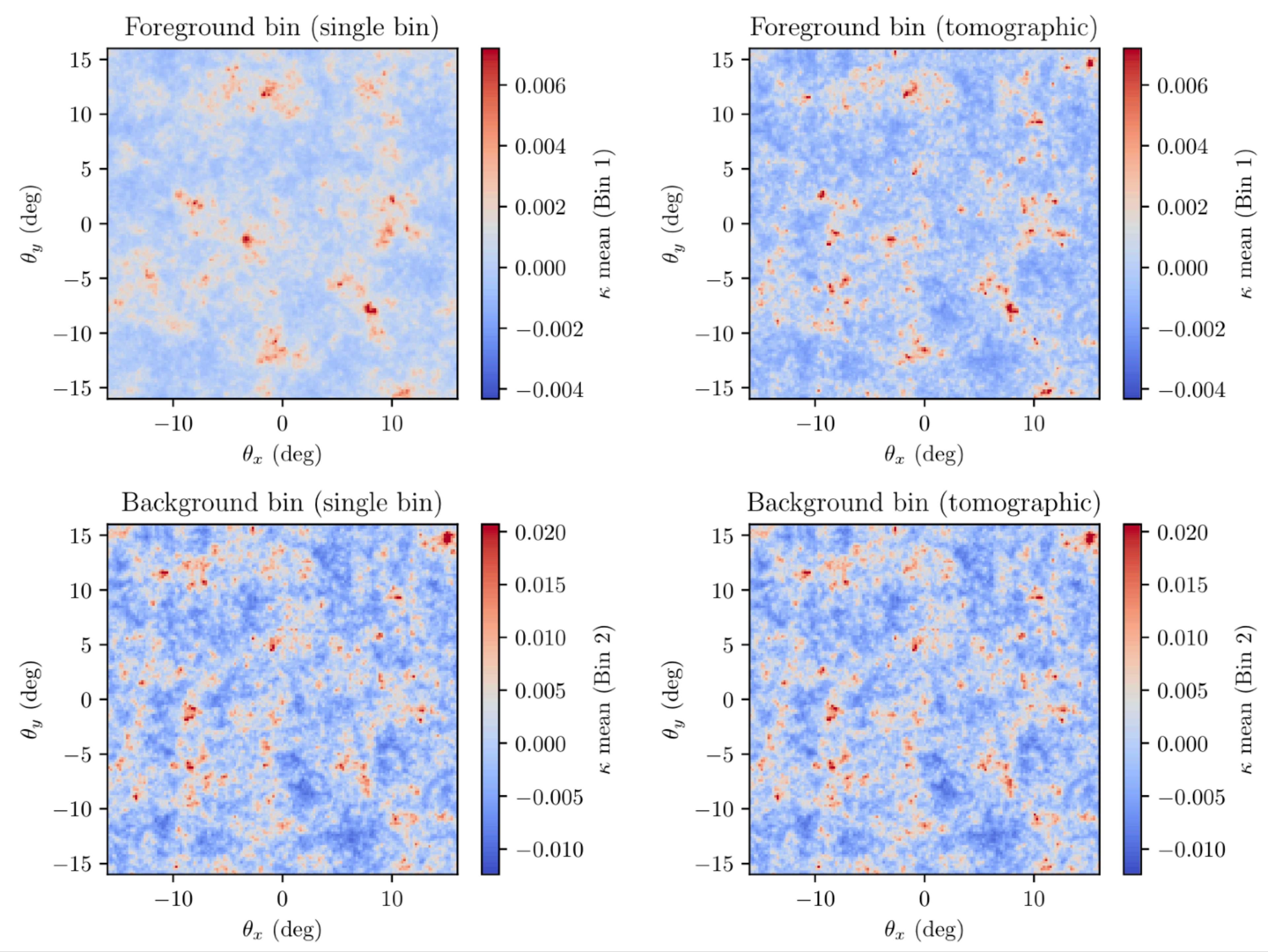
- Forward modelled field-level inference is the optimal way to extract cosmological information from a data set
- We have made multiple breakthroughs to enable field-level inference with weak lensing data:
  - *Improved mass maps from DES-Y3 weak lensing data using KaRMMa*
  - *Developed a GAN-based simulator to produce accurate weak lensing maps*
  - *2.5x gain in cosmological constraints with WL using MBI*
  - *MBI can put much stronger constraints on IA parameters*

Extra Slides

Tomographic mass-mapping

Mass-mapping without tomography

Mass-mapping with tomography

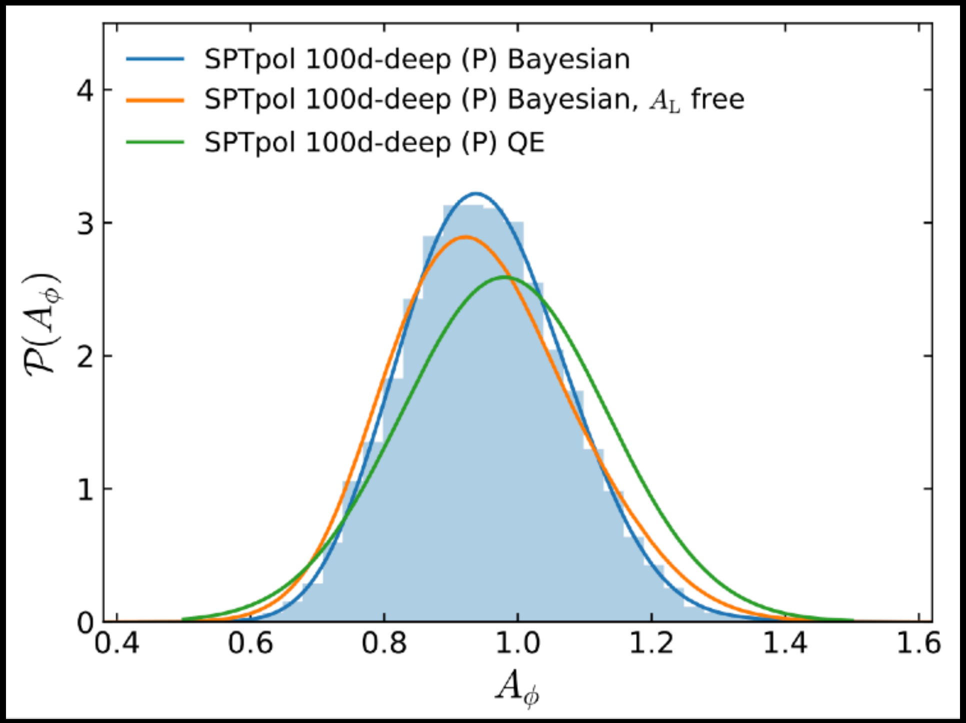


Resolve smaller structures in low-redshift bins — improvement most drastic in the low SNR maps

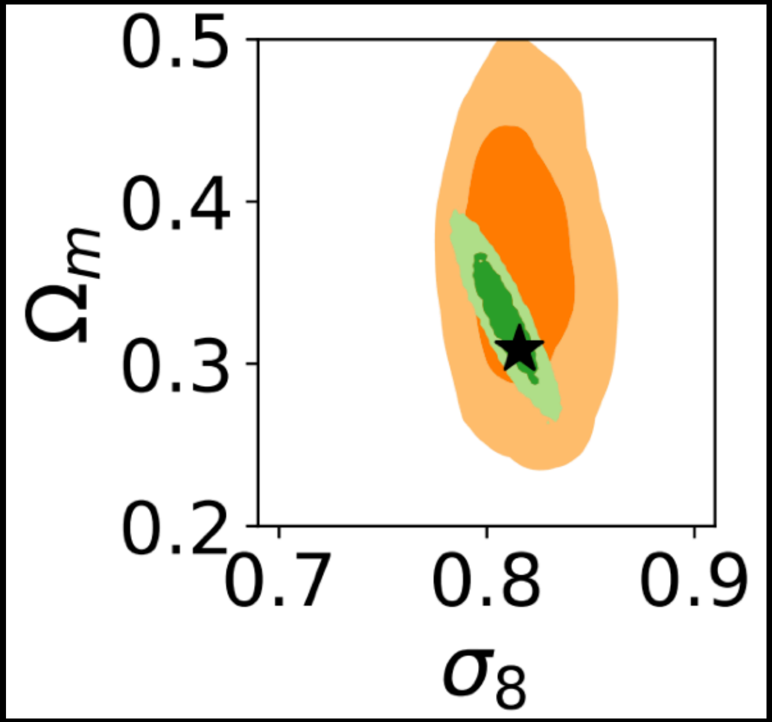
Weak-lensing data in different bins trace the same large-scale structures

High-redshift bins inform the structures in the lower redshift bins

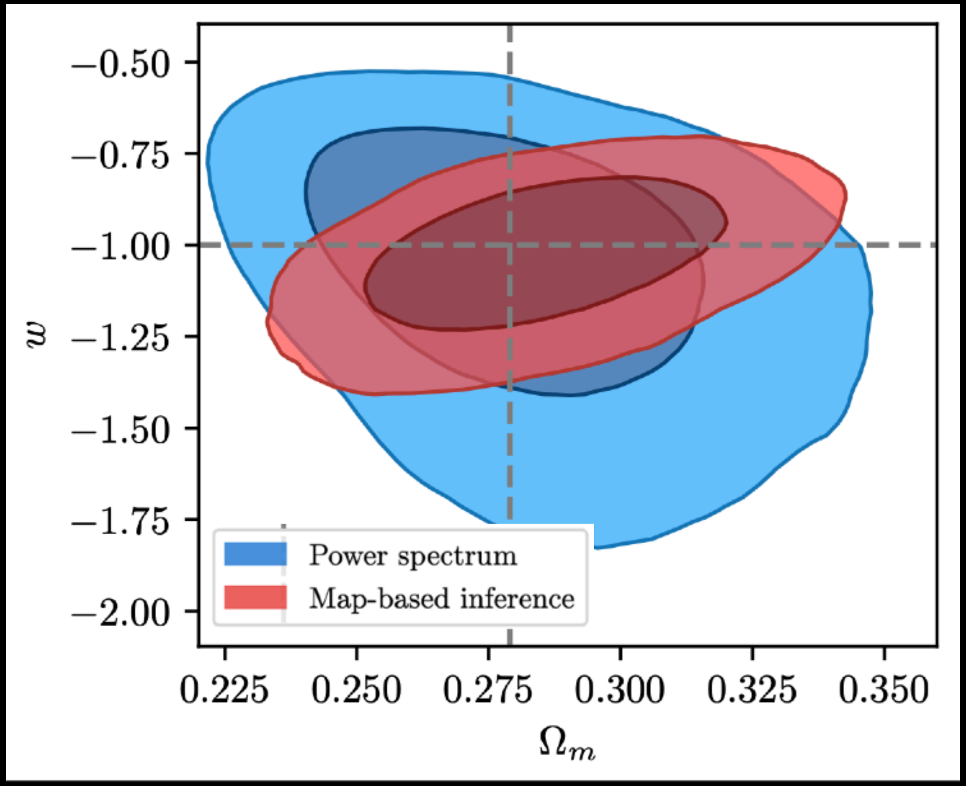
The promise of field-level cosmology inference



CMB lensing: Upto 50% stronger constraints on A_ϕ for CMB-S4 (Millea+, 20)



Projected galaxy clustering: Upto 5x stronger constraints (Dai+, 22)



Weak lensing: Upto 2.5 times stronger constraints w/ LSST-Y1 (Boruah+ 23)

Map-based cosmology inference shows the promise to improve cosmological information across different probes