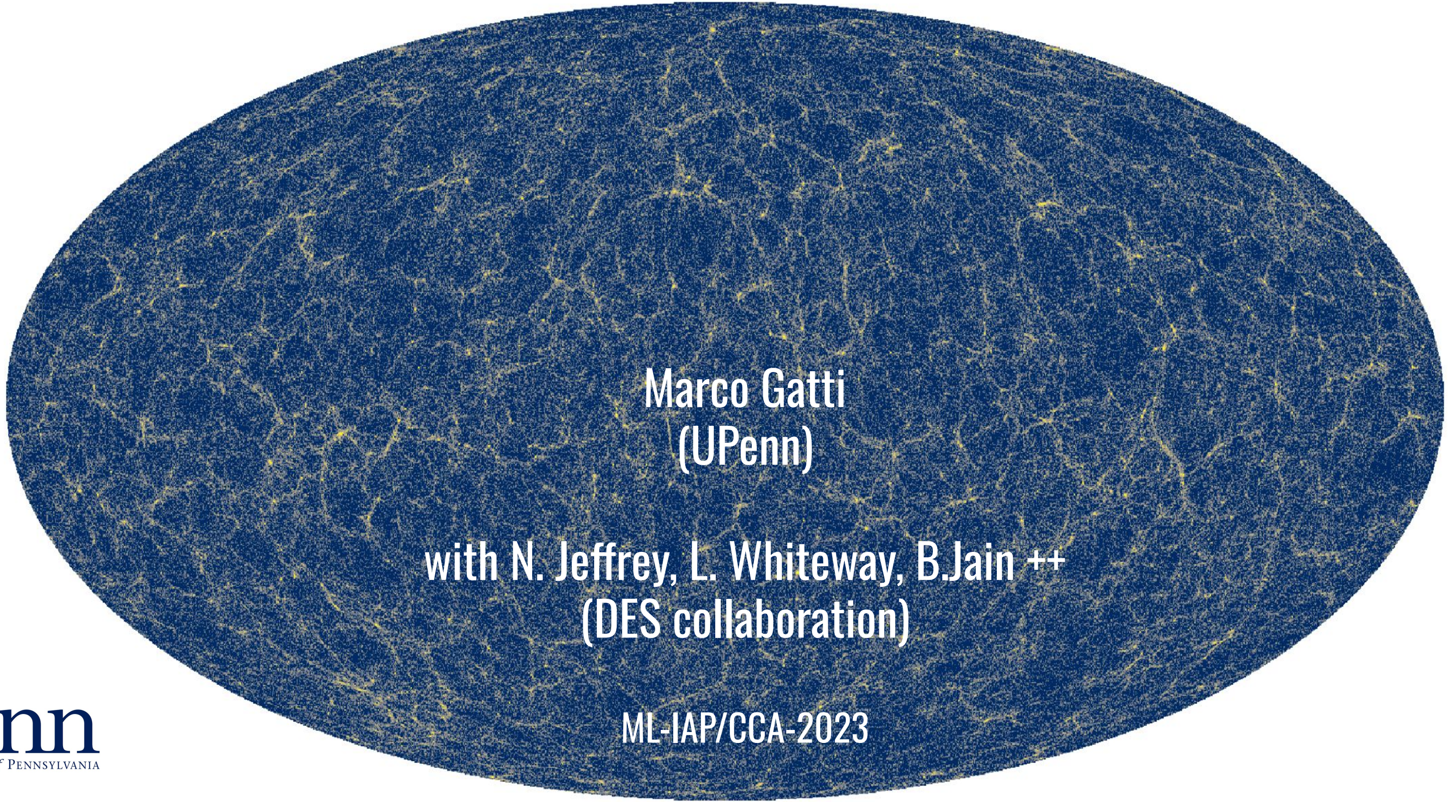
Likelihood-free Inference with Non-Gaussian statistics of weak lensing observables & DES Y3 data



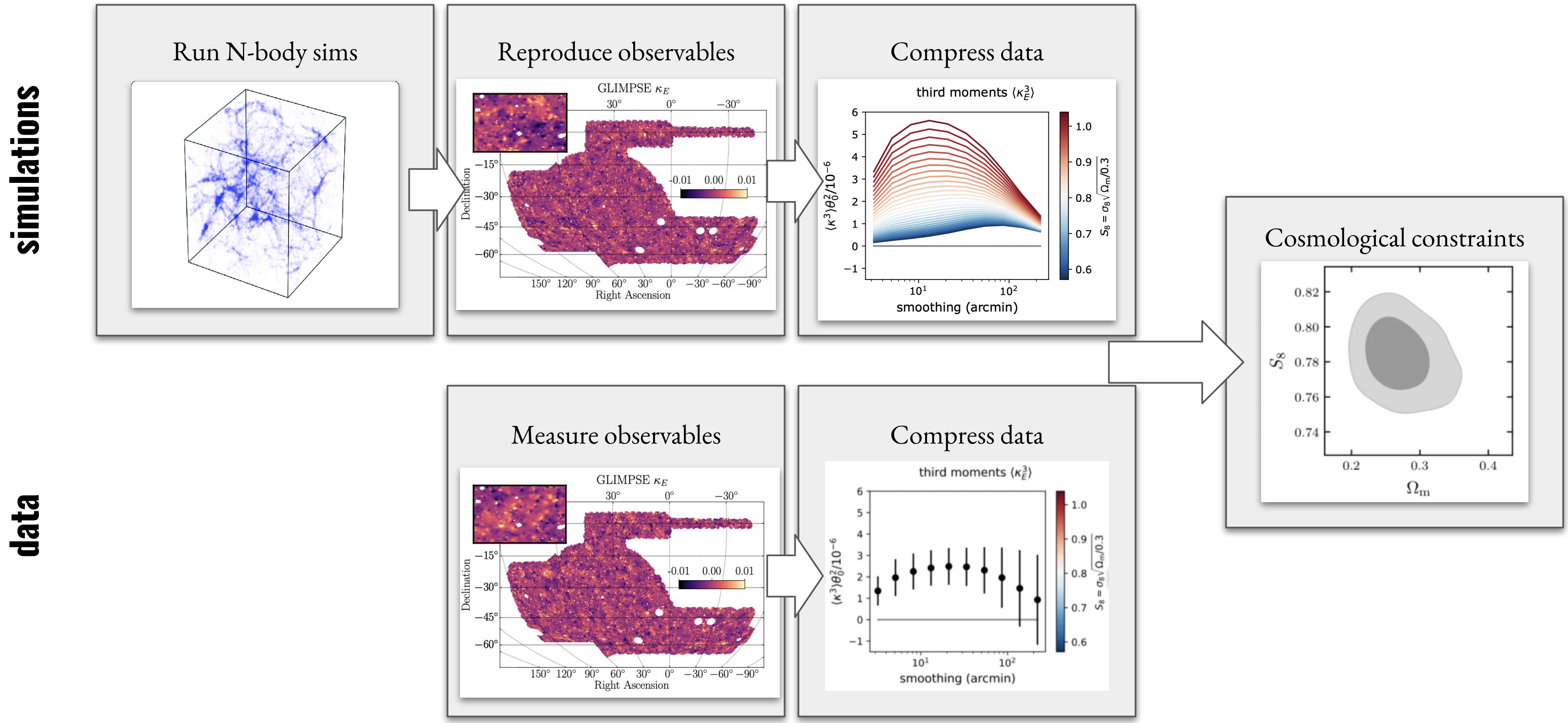
Marco Gatti
(UPenn)

with N. Jeffrey, L. Whiteway, B.Jain ++
(DES collaboration)

ML-IAP/CCA-2023

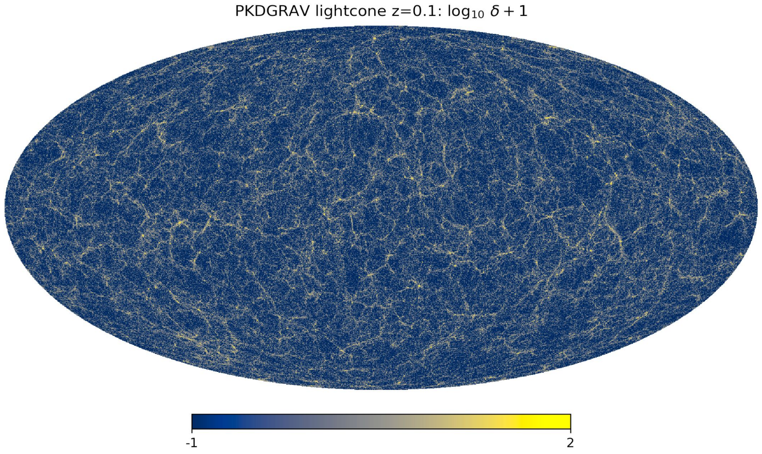
UNIVERSITY of PENNSYLVANIA

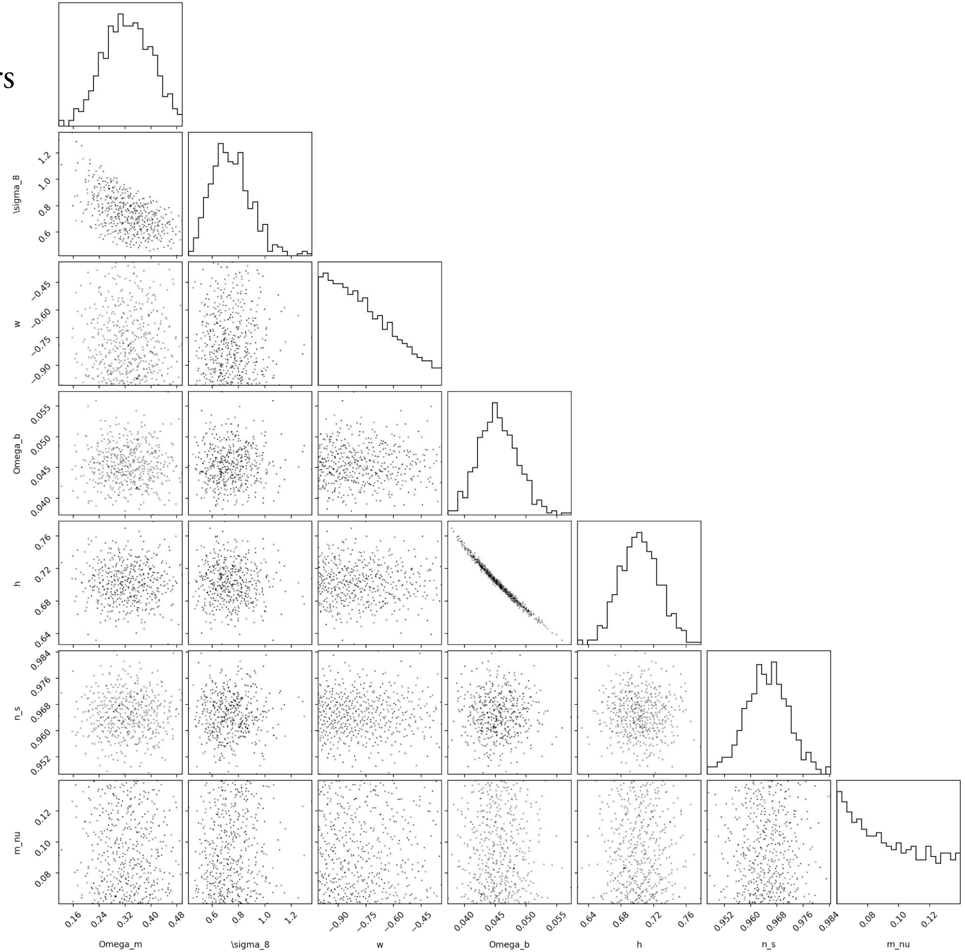
Simulation-based Inference with Non-Gaussian statistics



Dirac-simulations

-750 full-sky N-body simulations, 7 cosmological parameters (wCDM), run with PKDRAV3
- 50k GPU hours to run sims, ~ few TB of storage for the mocks. ~500kCPU hours to make mocks and measure statistics
- density field, convergence & shear maps in ~70 redshifts between z=0 and z =49
- Resolution: ~1.6 arcminutes (ell~4000)





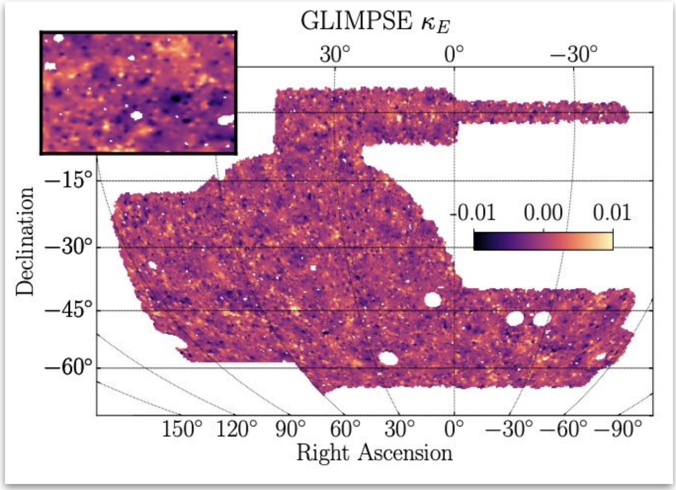
Reproduce observables

Our weak lensing mocks include:

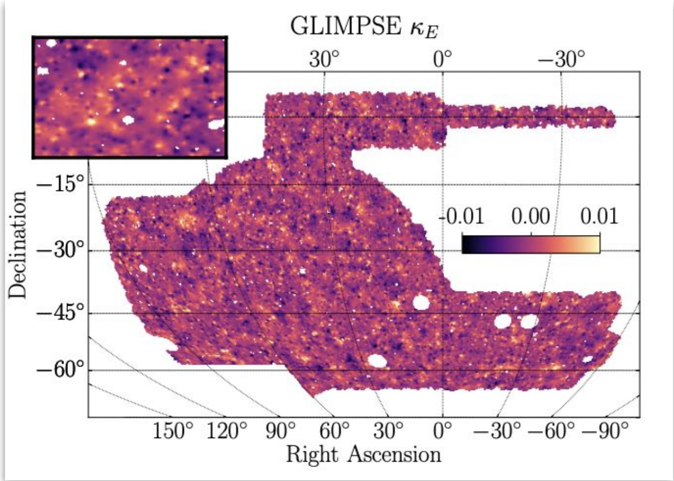
- Tomography (4 bins)
- Realistic shape noise
- Source clustering
- Photo-z uncertainties
- Shear bias
- Intrinsic alignment (NLA+bias)

We marginalise over all the relevant nuisance parameters describing observational & astrophysical systematics during the inference

simulations



data



Data compression & choice of summary statistics

Probes considered (and possibly combined):

Gatti et al, 2310.17557

- 2nd & 3rd moments
- Wavelet-based: Wavelet Phase Harmonics (WPH)
- Scattering Transform (ST)

Jeffrey et al, in prep.

- Power spectra
- Peaks
- CNN

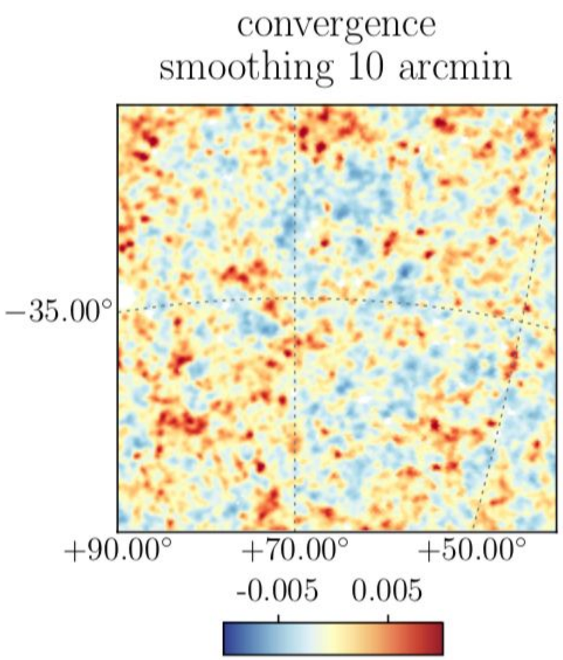
Prat et al, in prep.

- Persistent homology
- Betti numbers

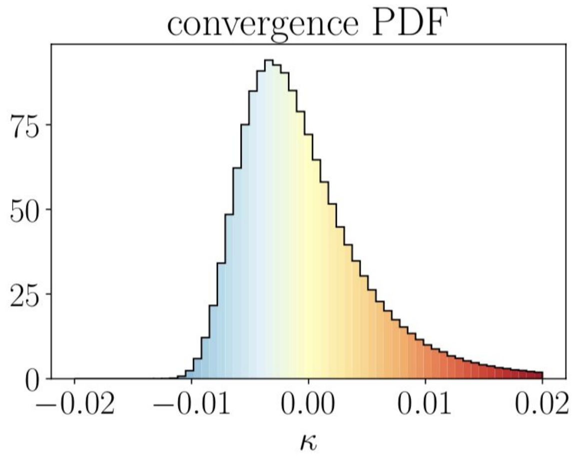
Data compression & choice of summary statistics

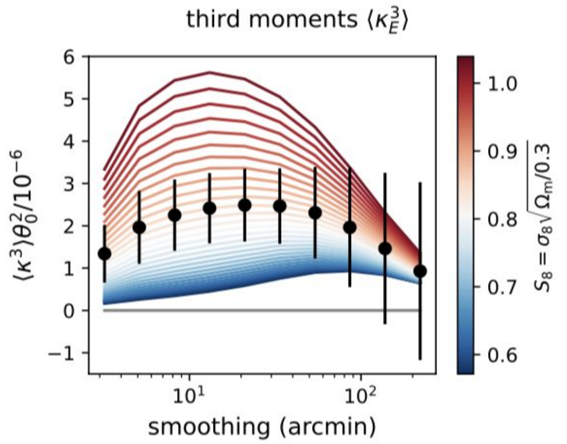
Probes considered (and possibly combined):

- Power spectra
- **2nd & 3rd moments**
- Peaks
- Wavelet-based: Wavelet Phase Harmonics (WPH) and Scattering Transform (ST)
- CNN



2nd moments probes the **variance** of the field,
3rd moments the **skewness**

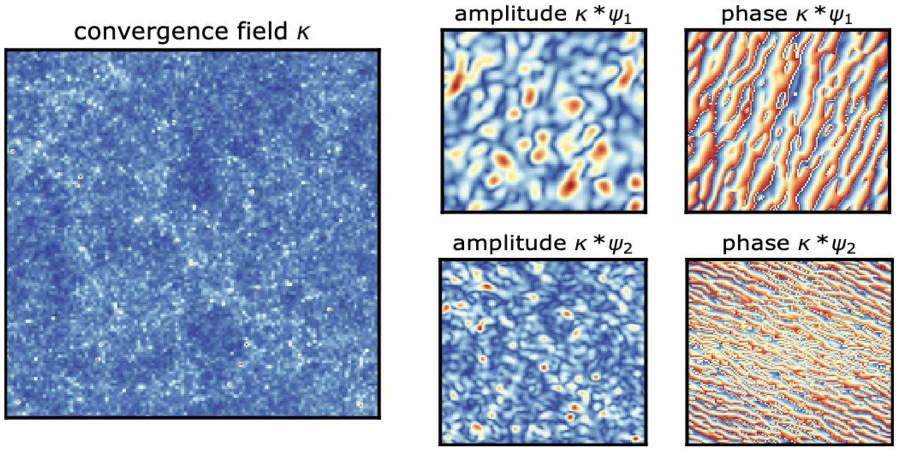




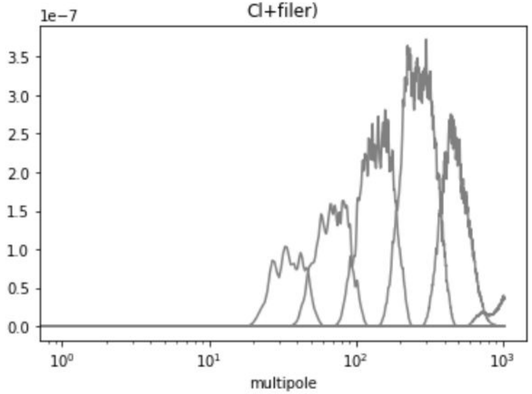
Data compression & choice of summary statistics

Probes considered (and possibly combined):

- Power spectra
- 2nd & 3rd moments
- Peaks
- **Wavelet-based: Wavelet Phase Harmonics (WPH) and Scattering Transform (ST)**
- CNN



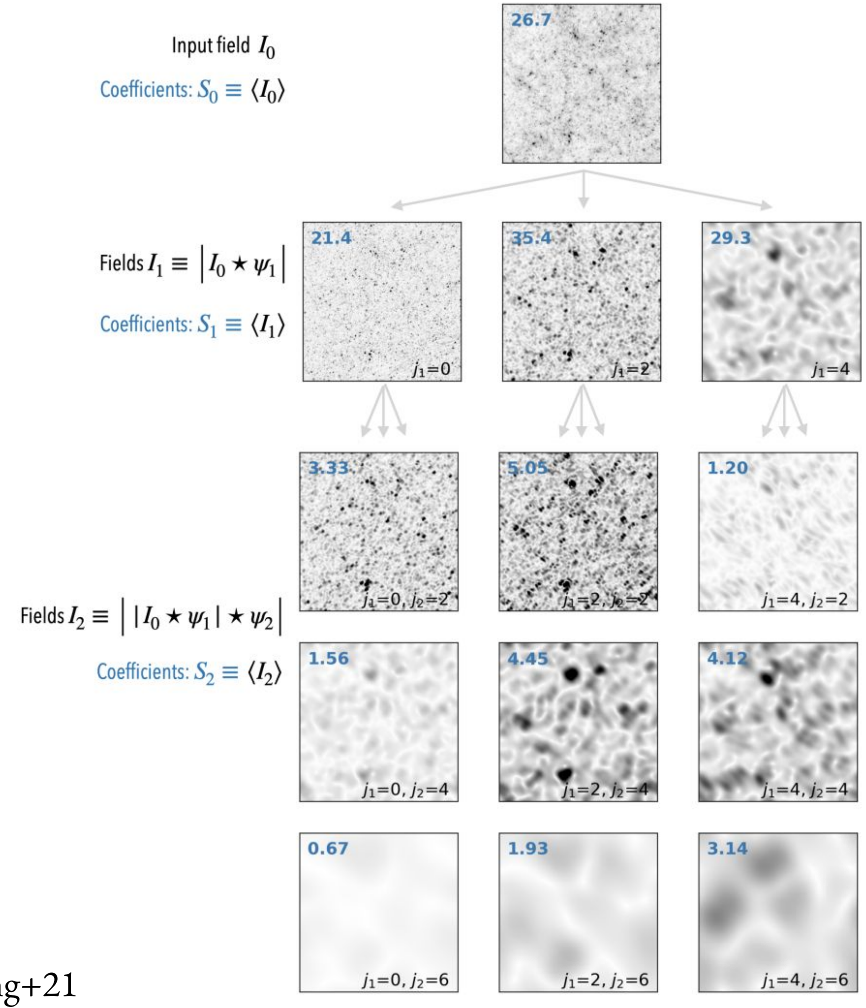
Smoothed maps well localised in real space and Fourier space



Data compression & choice of summary statistics

Probes considered (and possibly combined):

- Power spectra
- 2nd & 3rd moments
- Peaks
- **Wavelet-based: Wavelet Phase Harmonics (WPH) and Scattering Transform (ST)**
- CNN

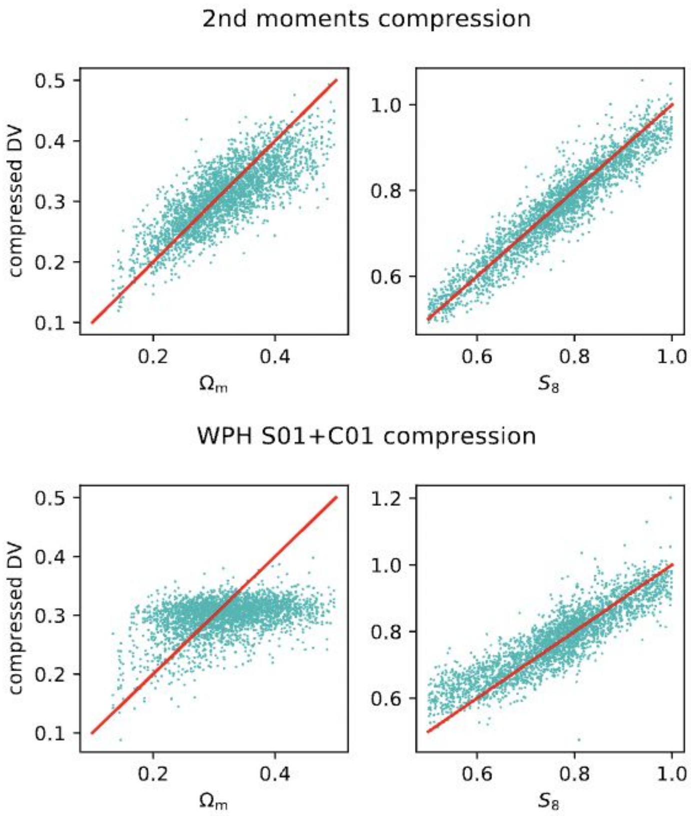


Cheng+21

Data compression : neural compression

Table 3. Neural Network Layers and number of parameters

Layer (type)	output shape	number of parameters
Dense	900	900*(length DV+1)
LeakyReLU	900	0
Dense	800	720800
LeakyReLU	800	0
Dense	100	80100
ReLU	100	0
Dense	100	10100
ReLU	100	0
Dense	1	101



Likelihood inference with neural density estimators (NDE)

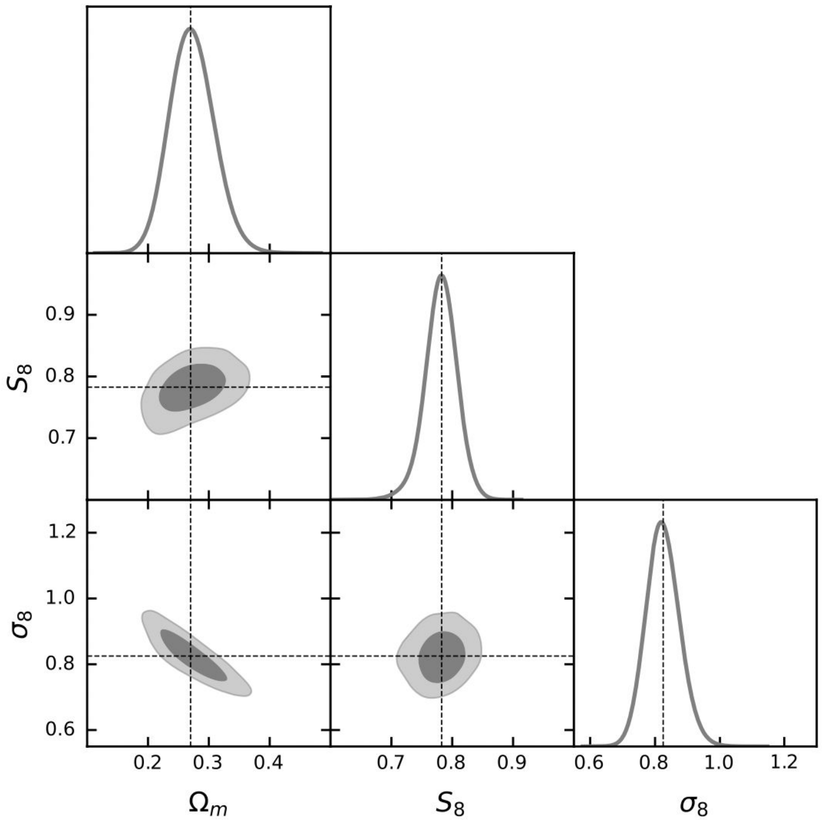
Simulation-based inference allows us to infer unknown cosmological parameters by directly comparing observed data with forward-simulated mock data.

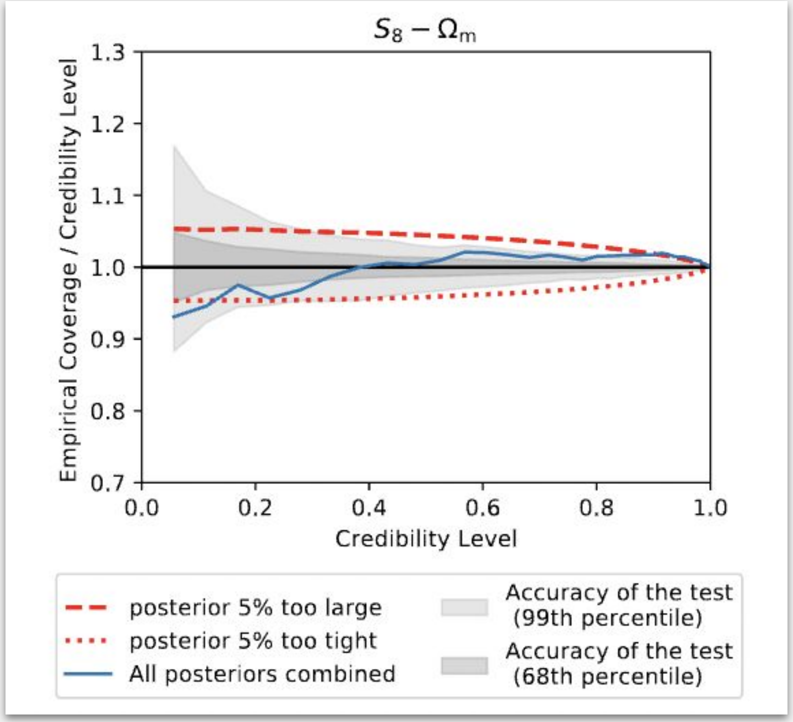
1 - draw simulated data vectors with noise {d_i,theta_i}

2 - use neural density estimators to estimate p(d|theta)

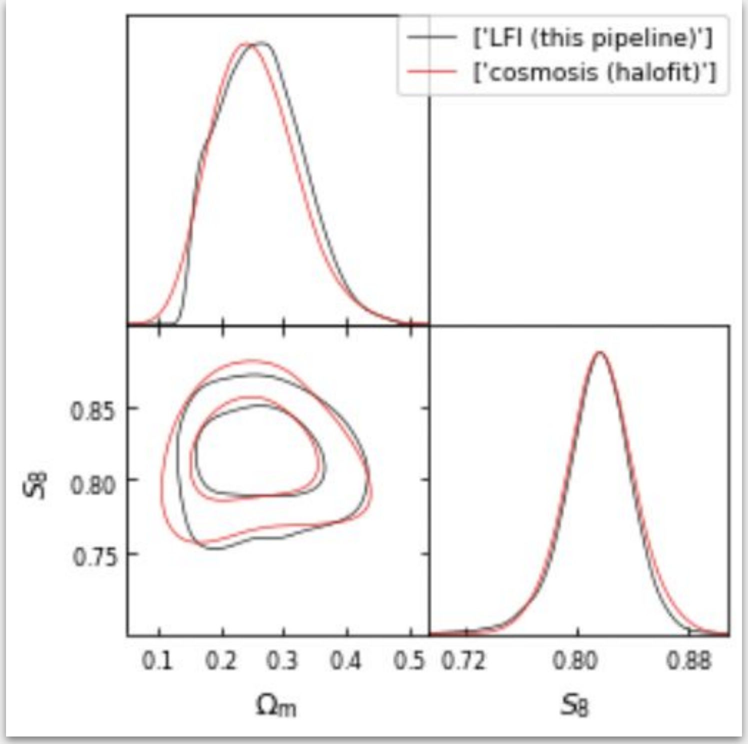
3 - use your observed data d_o to get the likelihood L = p(d_o|theta)

LFI requires significantly less simulations than the emulator approach! (~20 times less)





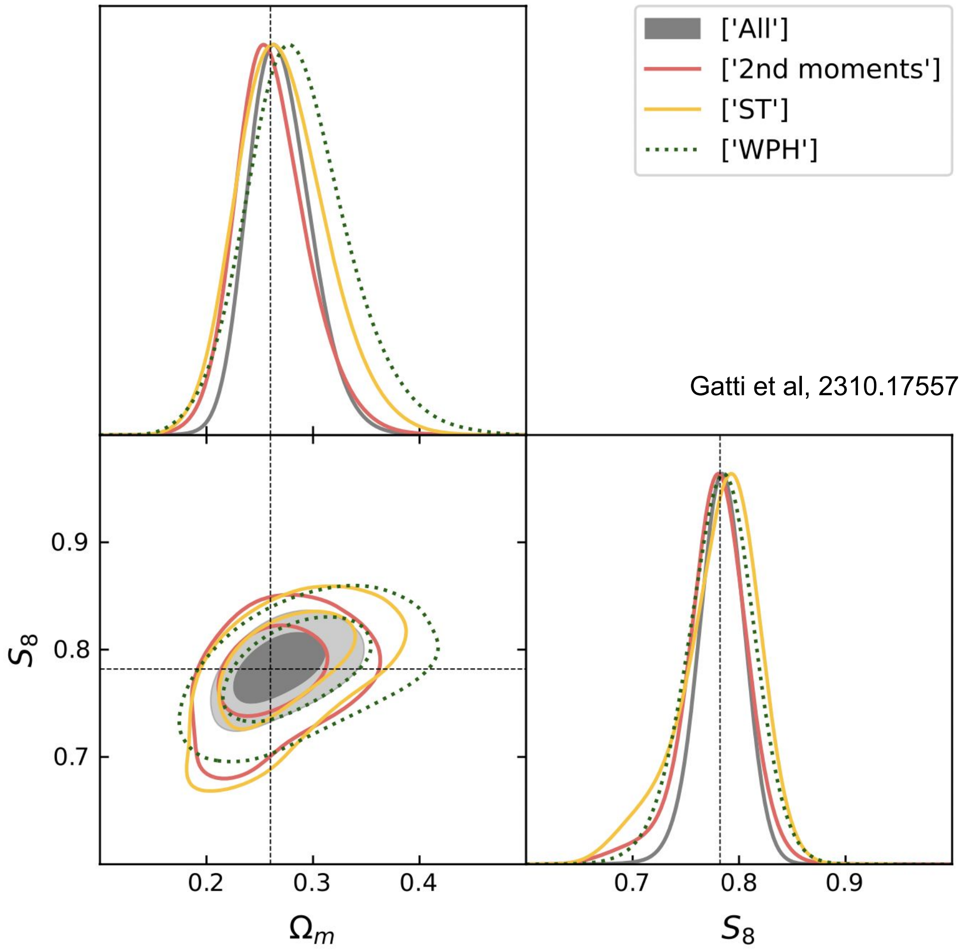
Check 1: **is the posterior size correct?**



Check 2: **simulation based analysis of power spectra (LFI)** vs **standard Gaussian Likelihood + theory modeling**

Summary Statistic(s)	$\sigma(S_8)$ [x100]	$\sigma(\Omega_\mathrm{m})$ [x100]	$\mathrm{FoM}(S_8, \Omega_\mathrm{m})$ -
2nd moments	2.7	3.4	904
2nd+3rd moments	2.6(+ 3%)	3.4(0%)	1035(+15%)
2nd moments + ST	2.7(+ 2%)	3.0(+12%)	1245(+38%)
2nd moments + WPH	2.4(+11%)	2.9(+15%)	1385(+53%)
2nd+3rd moments+ST+WPH	2.1(+21%)	2.9(+14%)	1733(+92%)

Validation with independent sims



Gatti et al, 2310.17557