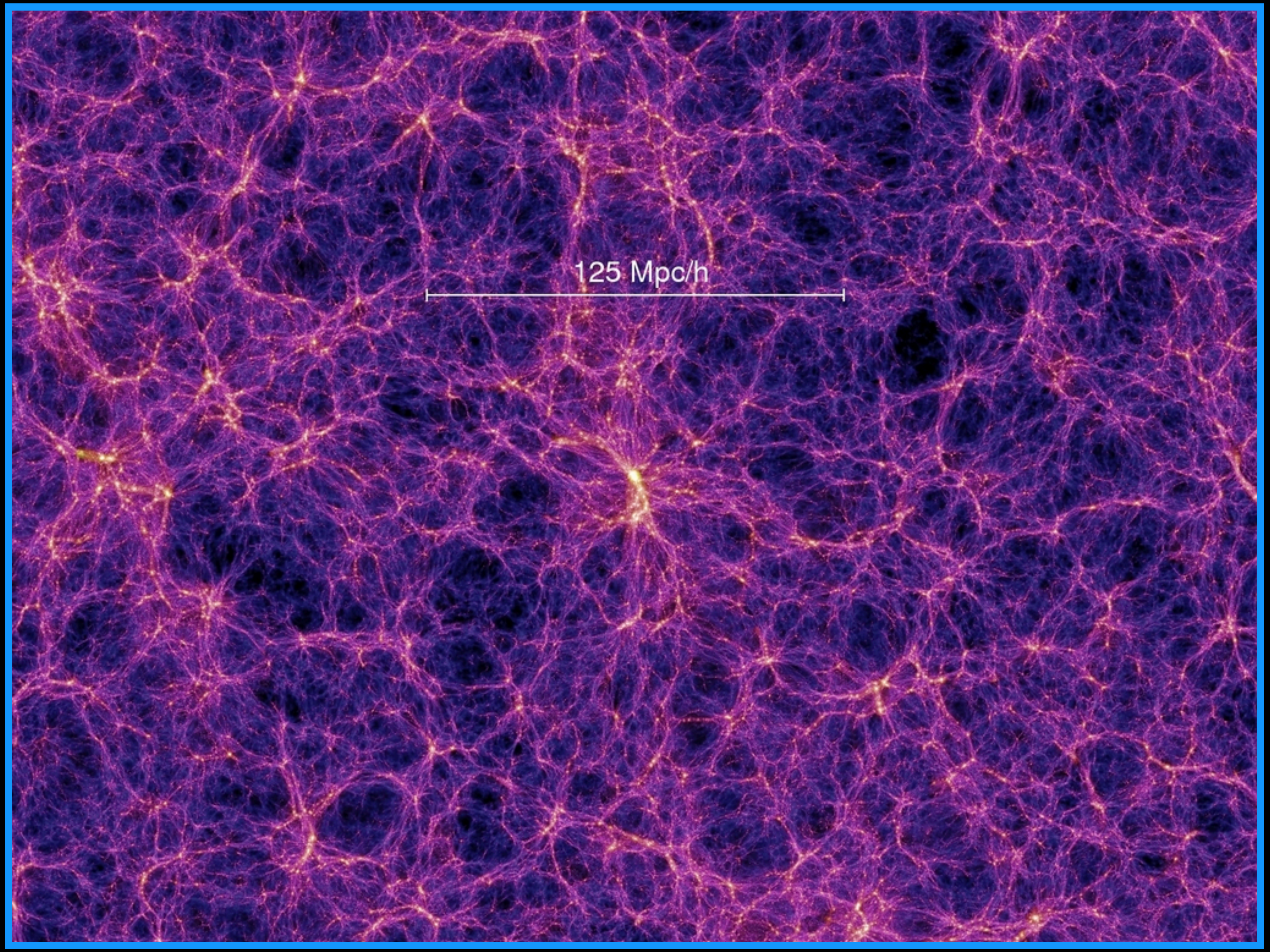
AI analyses dark matter maps:

improving the precision of cosmological measurements with CNNs

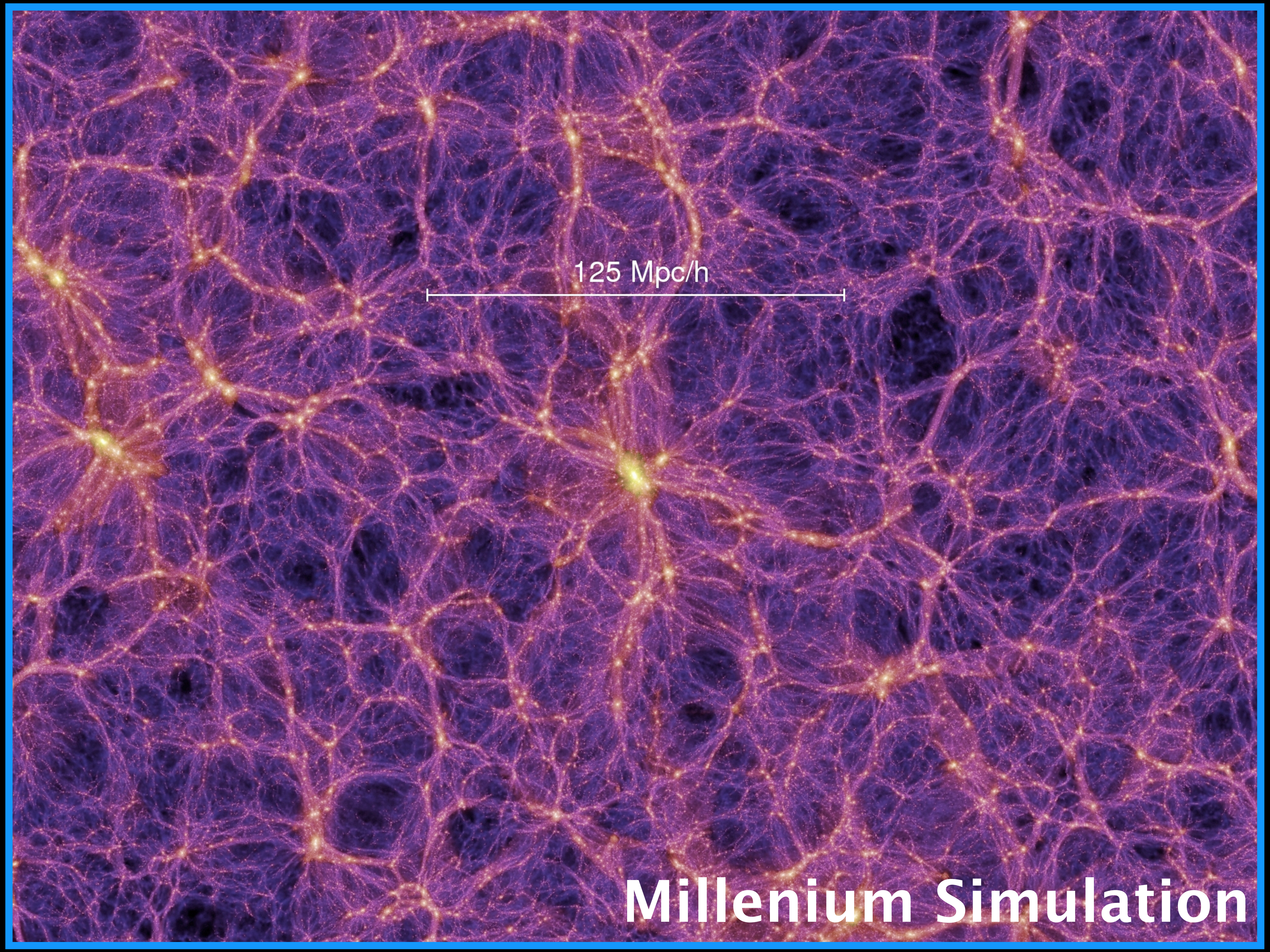
Tomasz Kacprzak (ETH Zurich),
Janis Fluri, Aurelien Lucchi, Nathanaël Perraudin,
Thomas Hoffman, Alexandre Refregier, Adam Amara

AI & Space, AMLD2020, 27 January 2020

Distribution of dark matter in the universe



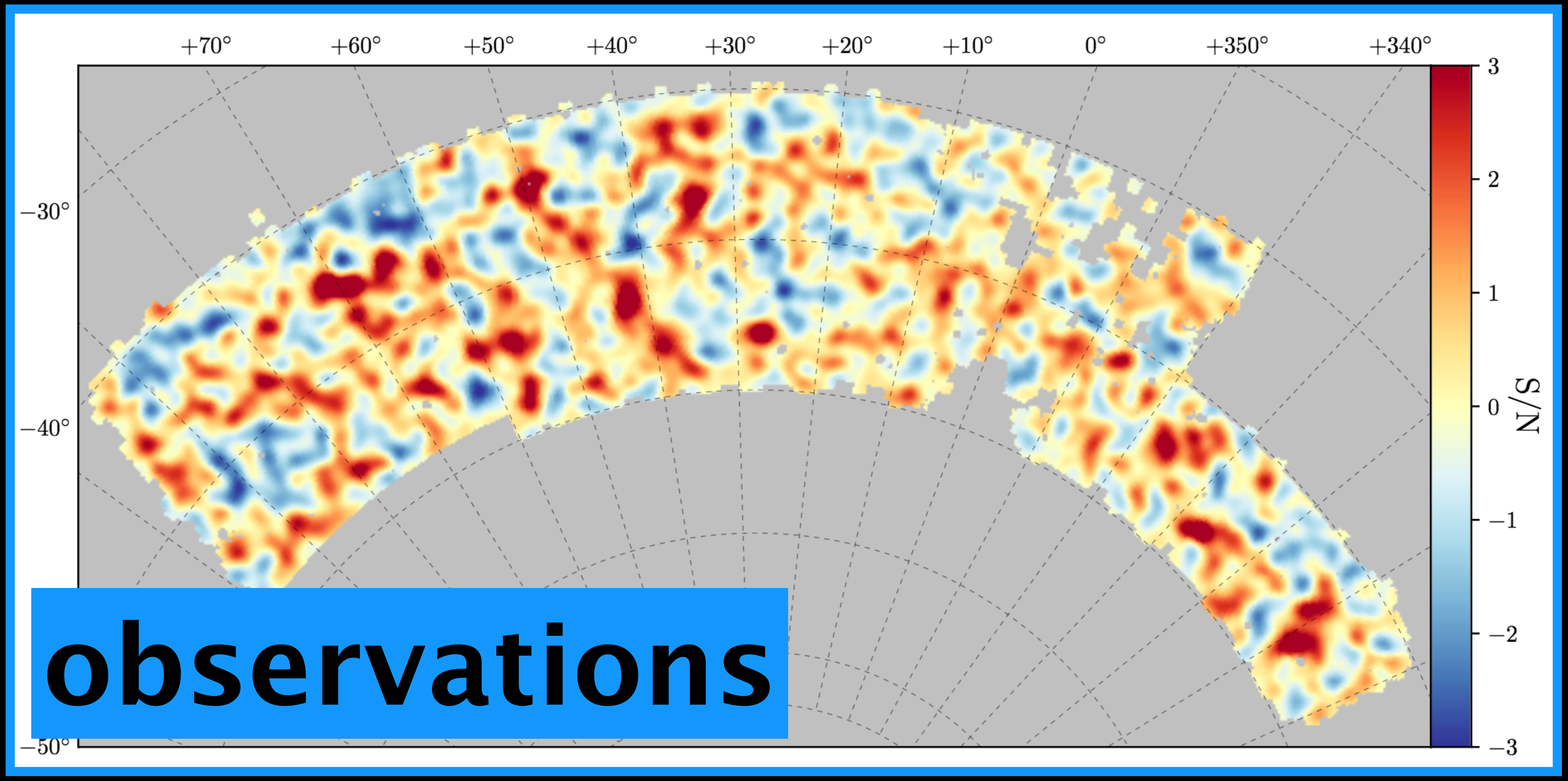
4.7 billion years after the big bang



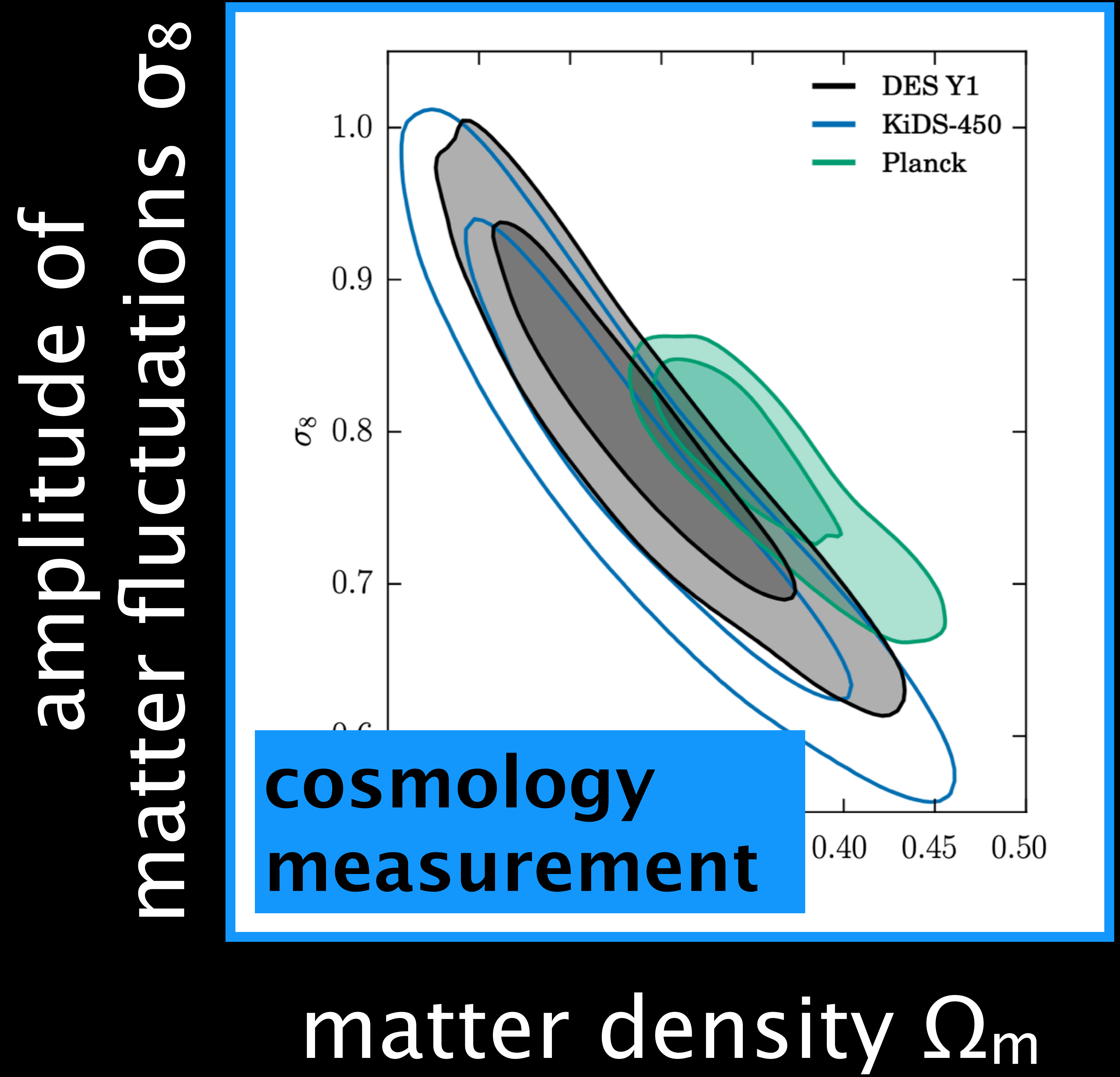
today, 13.6 billion years

Dark matter distribution can teach us about the laws of physics: gravity, dark energy, dark matter

Cosmological parameter inference

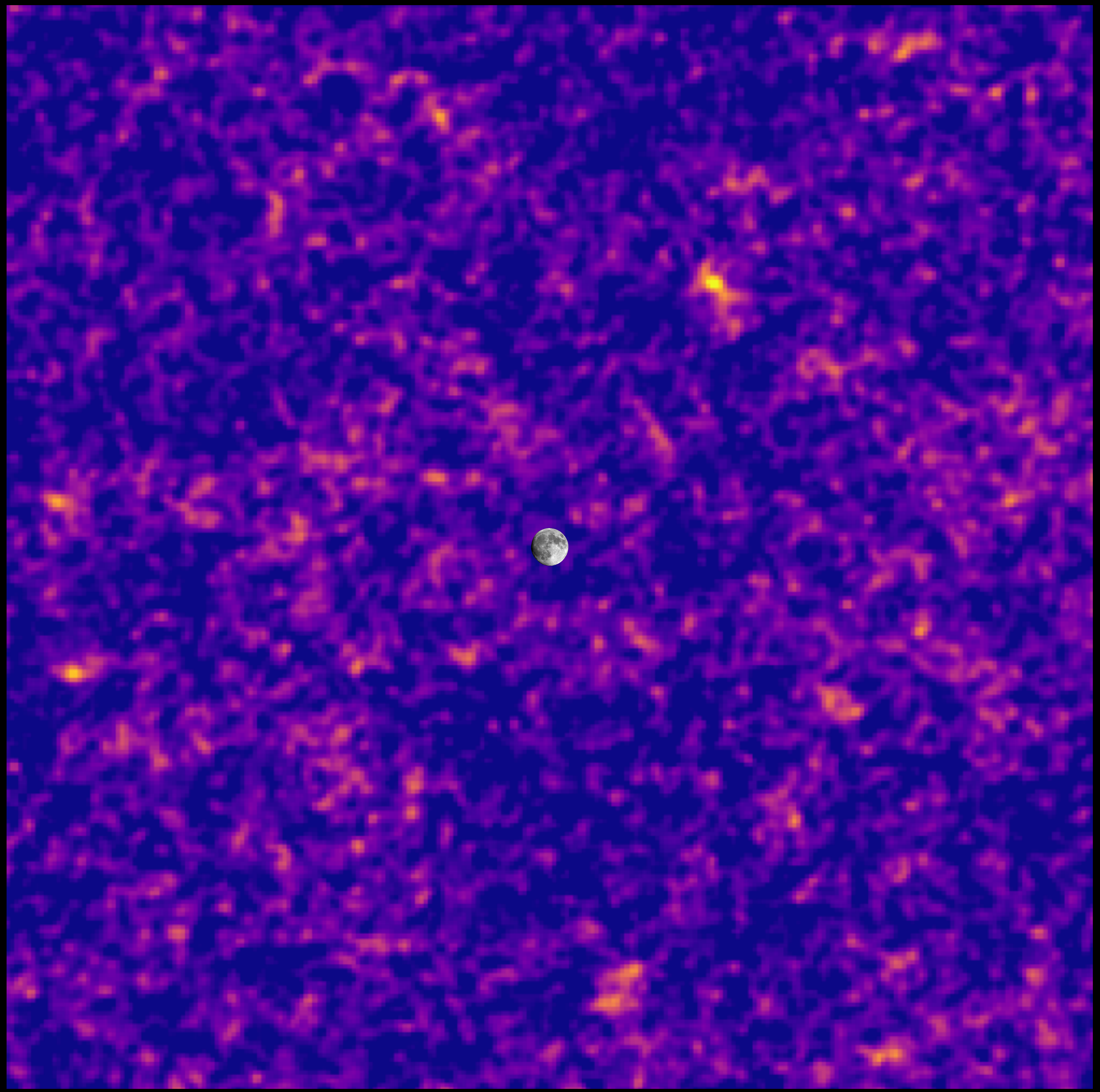


DES Collaboration (+TK) 2017 1708.01535

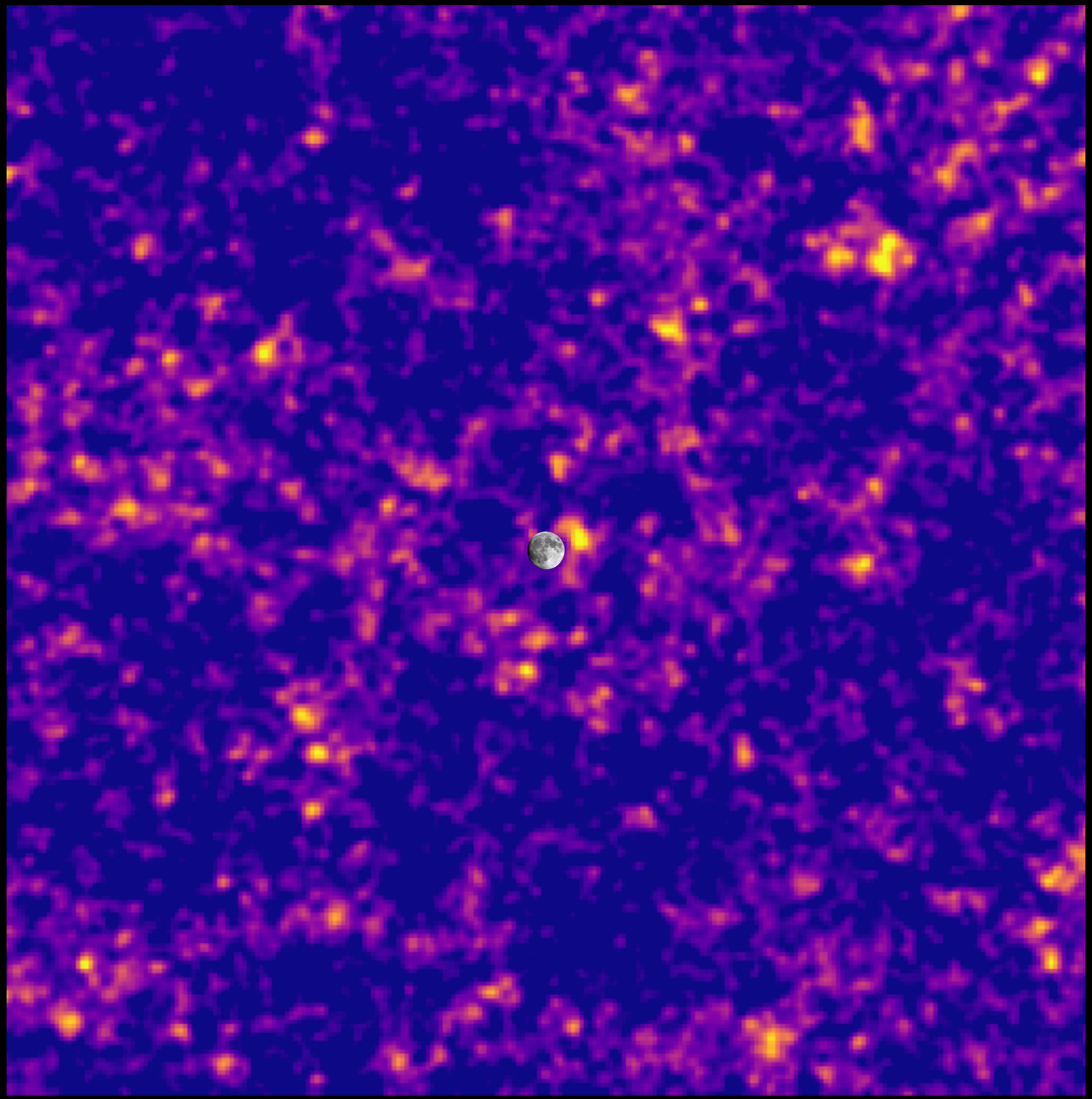


DES Collaboration (+TK) 2017 1708.01530

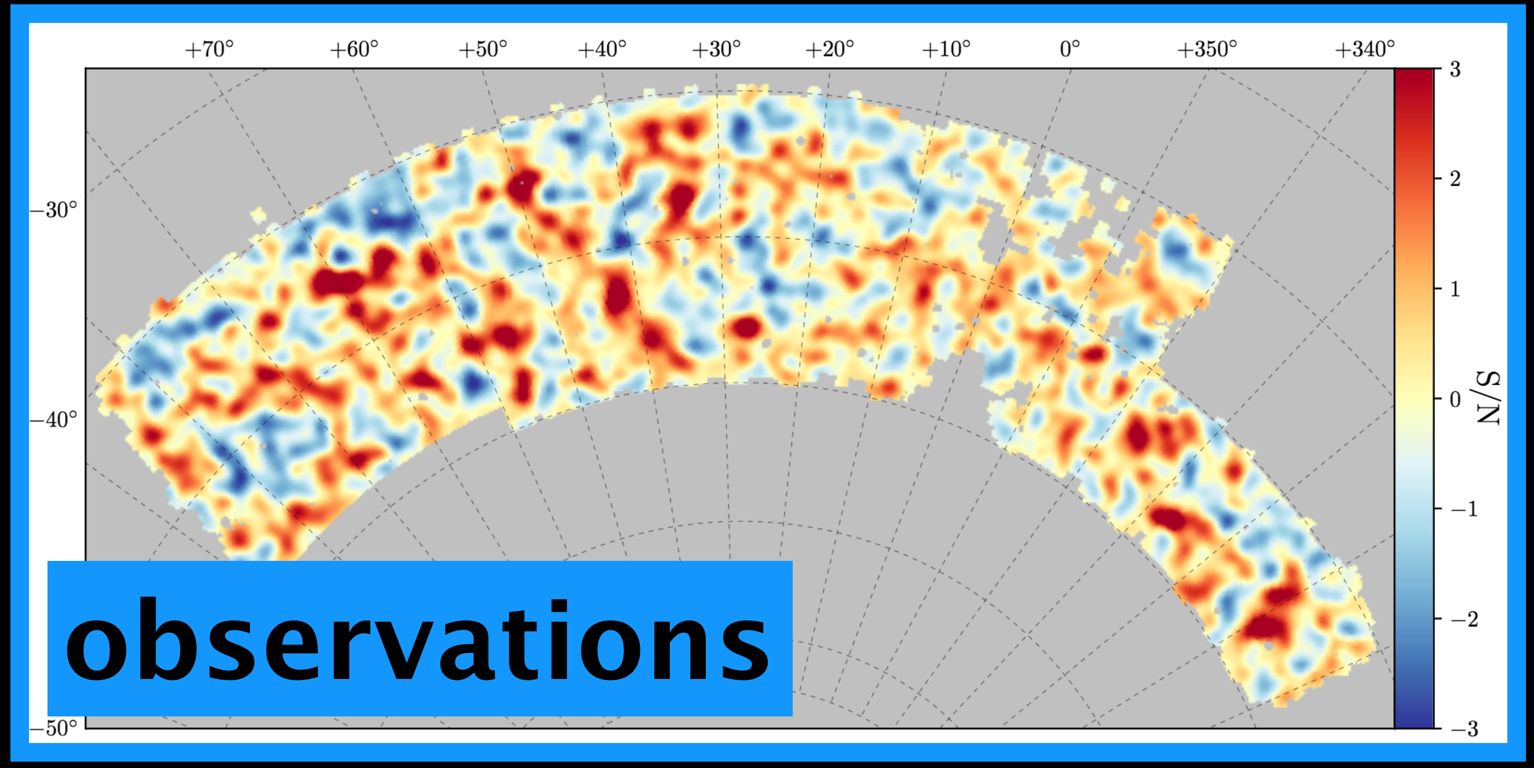
Dark matter distribution carries non-Gaussian information about the cosmological model



low σ_8 low Ω_m



high σ_8 high Ω_m



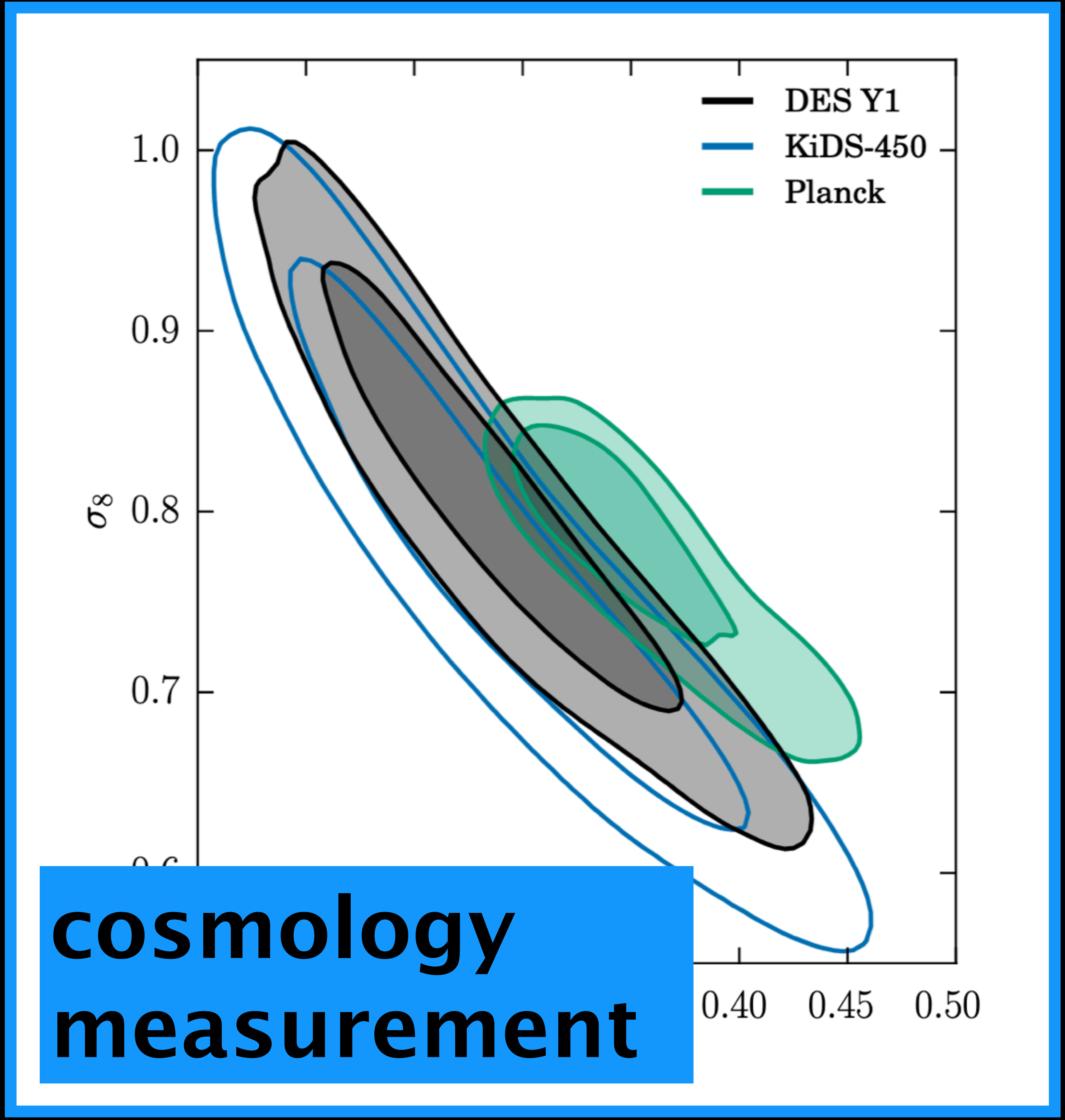
observations

comparison method?

$\Omega_m \quad \sigma_8 \quad H_0 \quad \ldots$

theory prediction

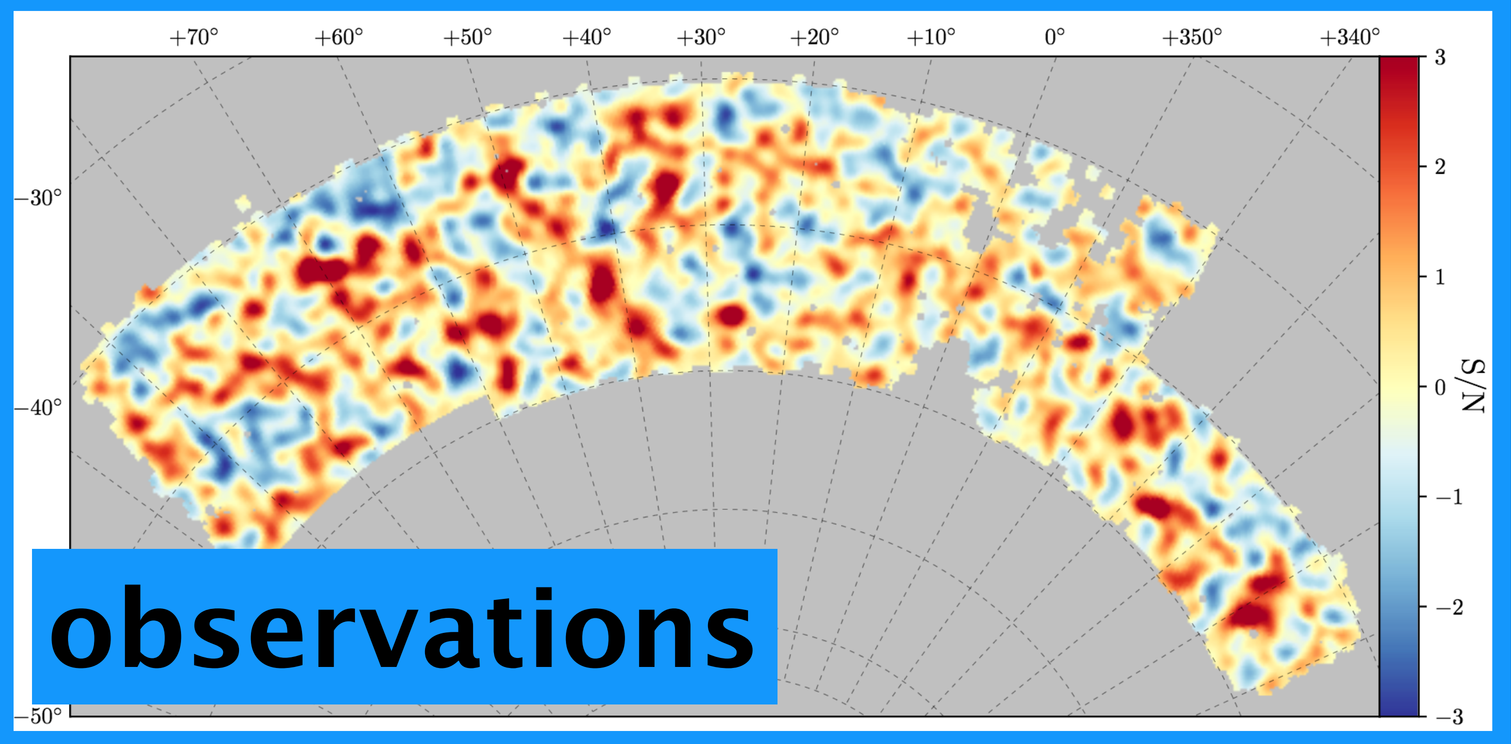
σ8

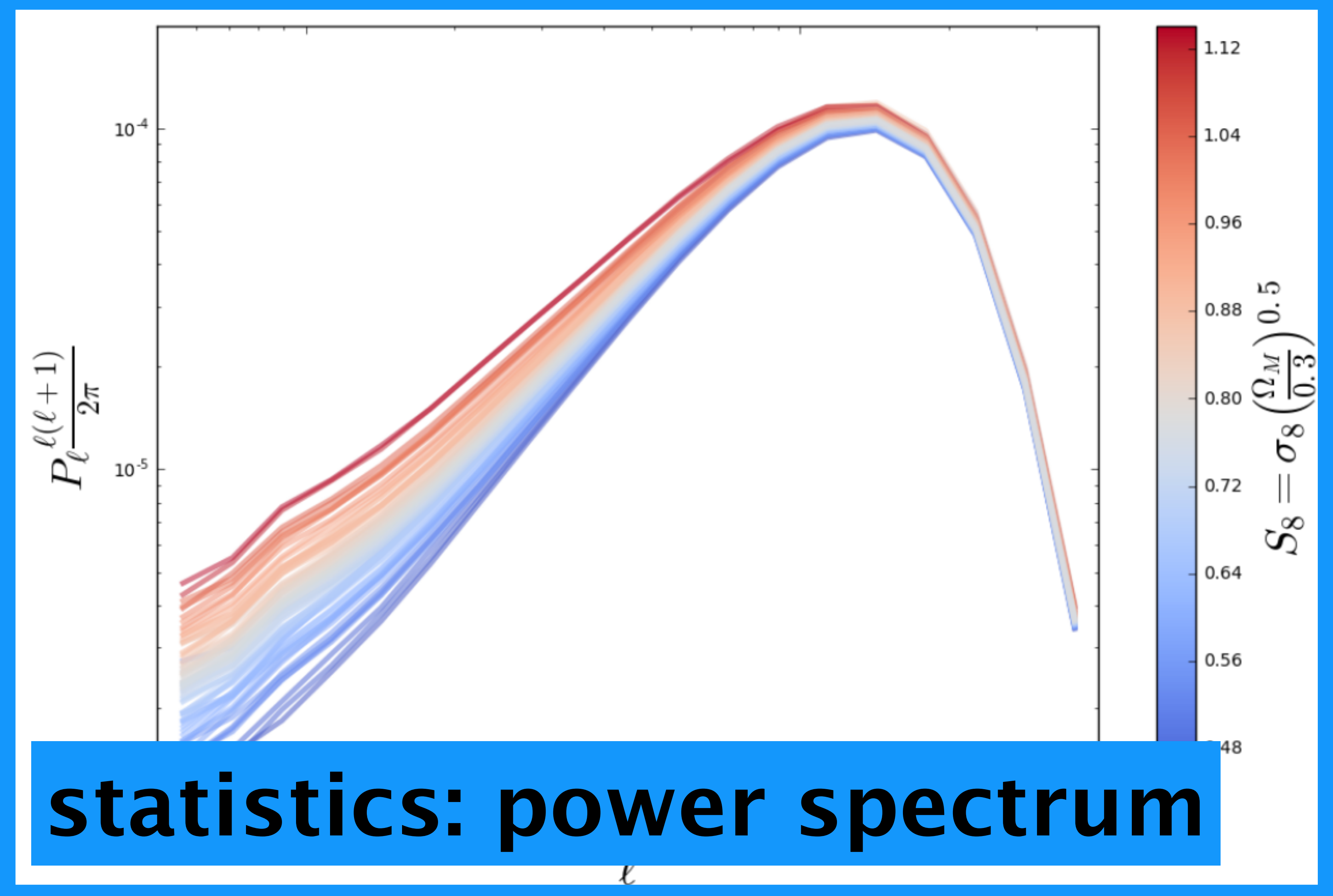


cosmology measurement

Ωm

Traditional inference

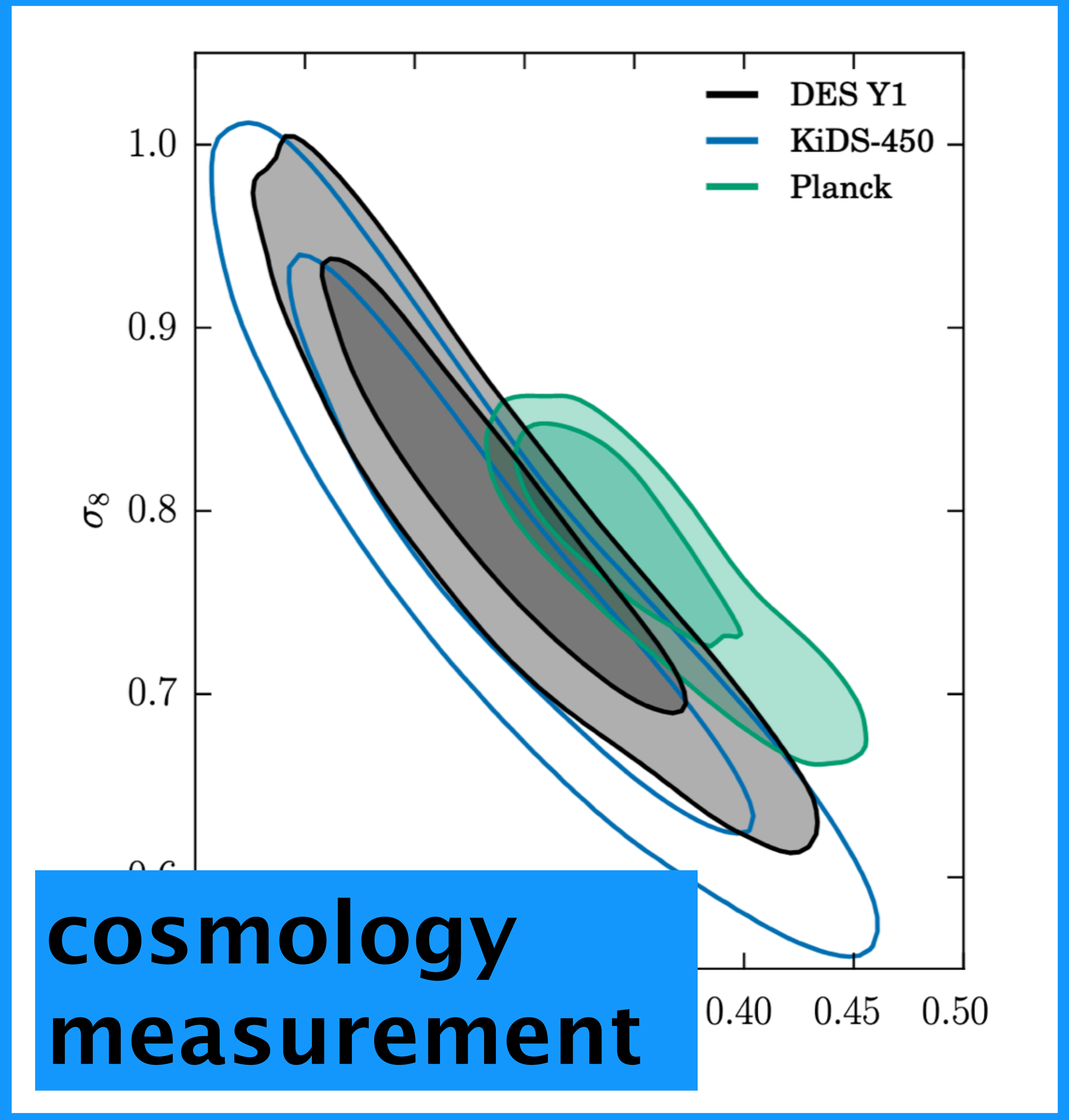




$$C_l = \frac{9}{16}\left(\frac{H_0}{c}\right)^4 \Omega_m^2 \int_0^{\chi_h} d\chi \left[\frac{g(\chi)}{ar(\chi)}\right]^2 P\left(\frac{l}{r},\chi\right),$$

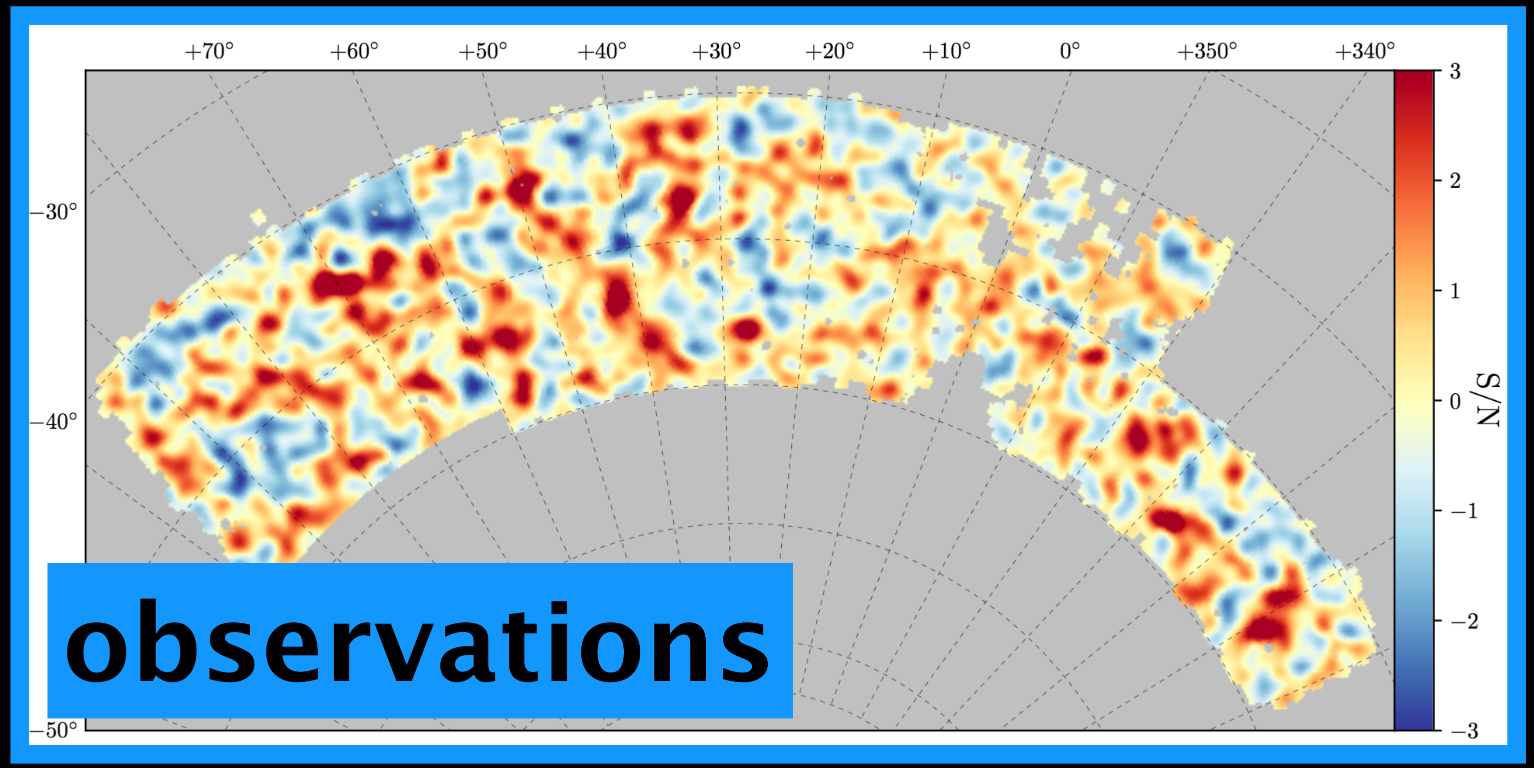
theory prediction: analytical

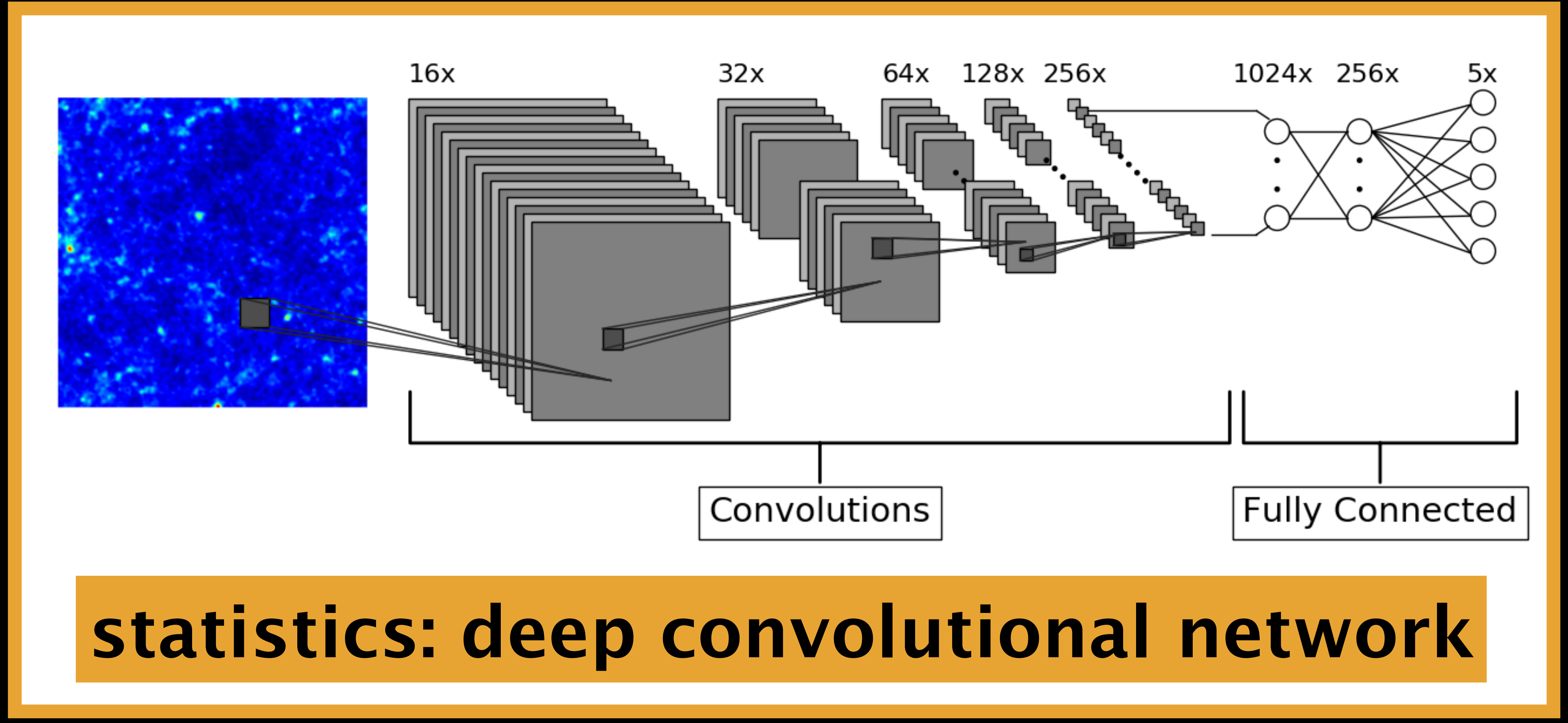
σ_8

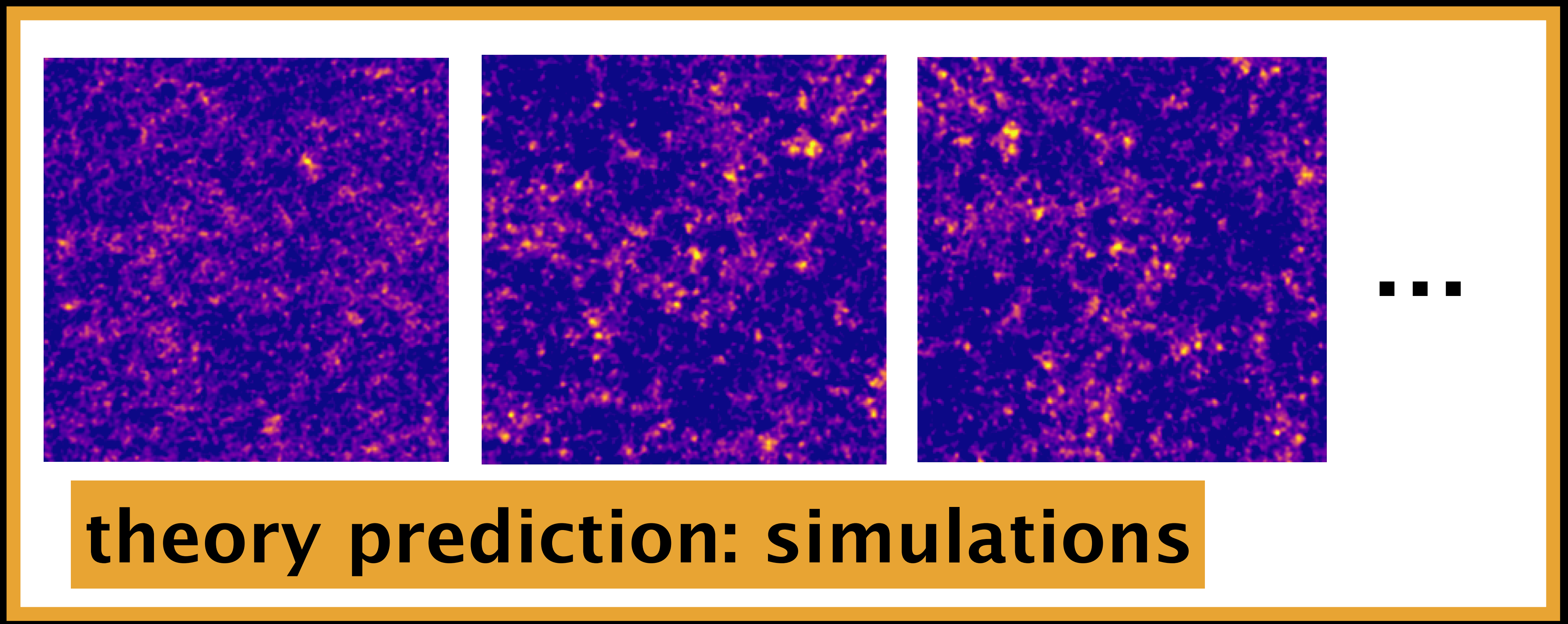


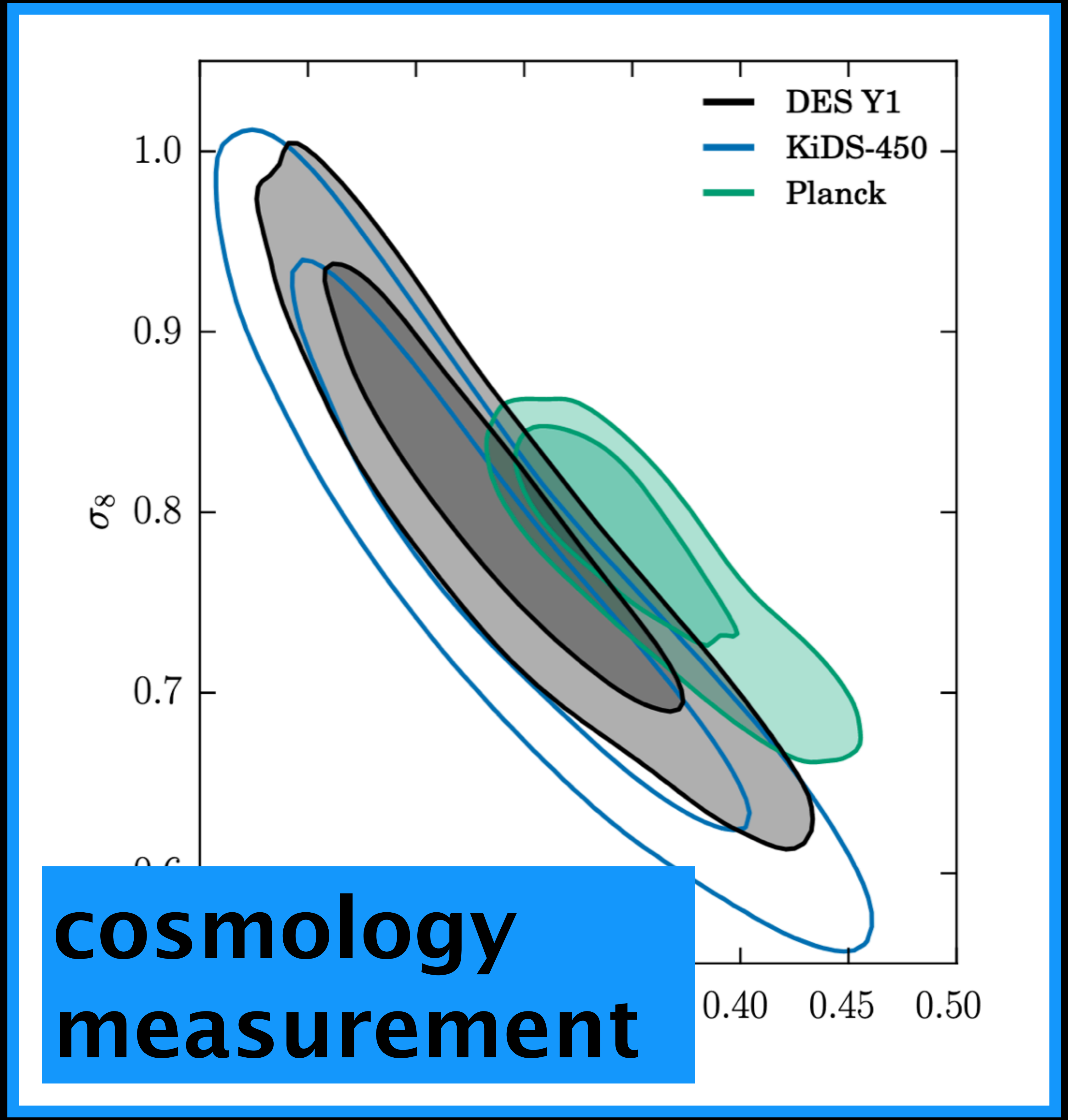
Ω_m

Inference with Deep Learning





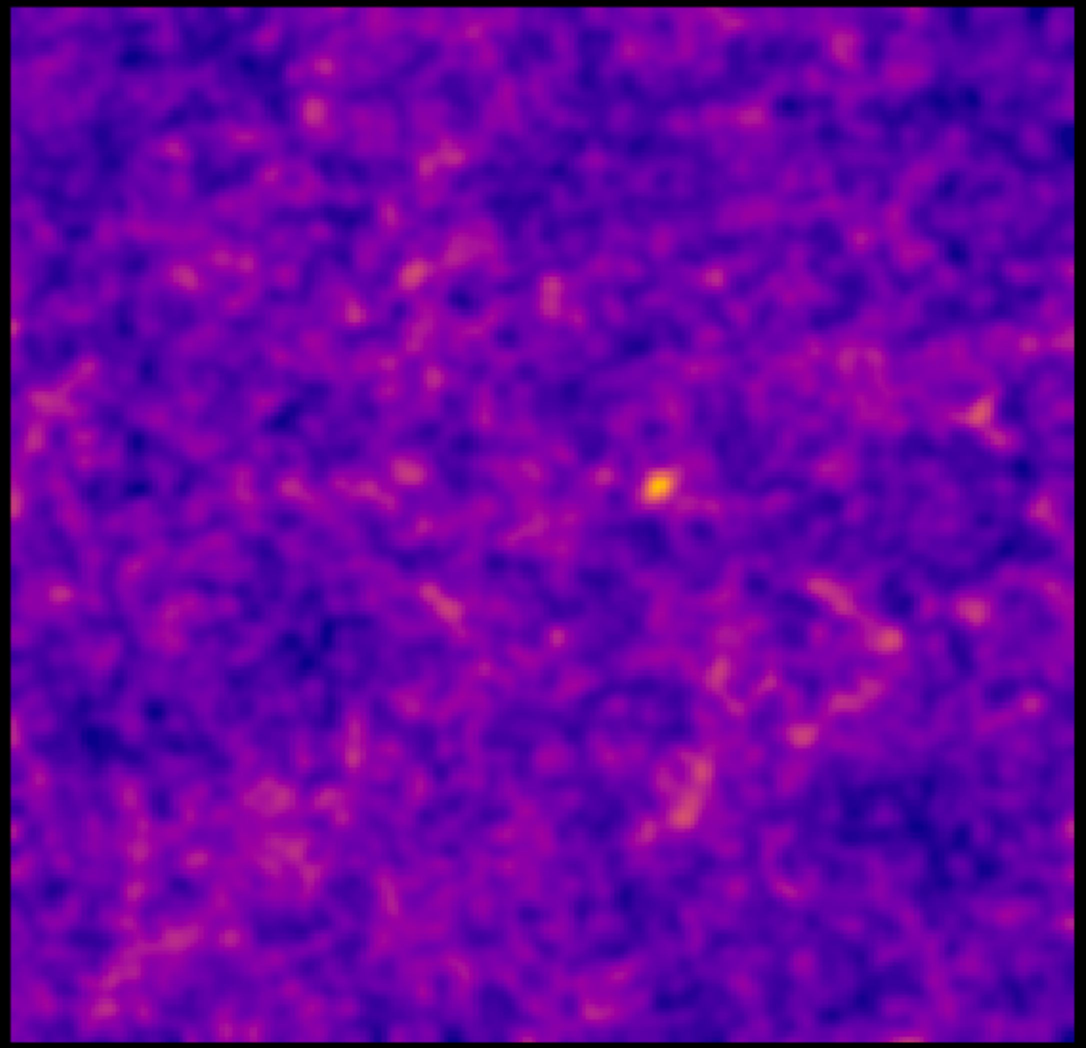




σ_8

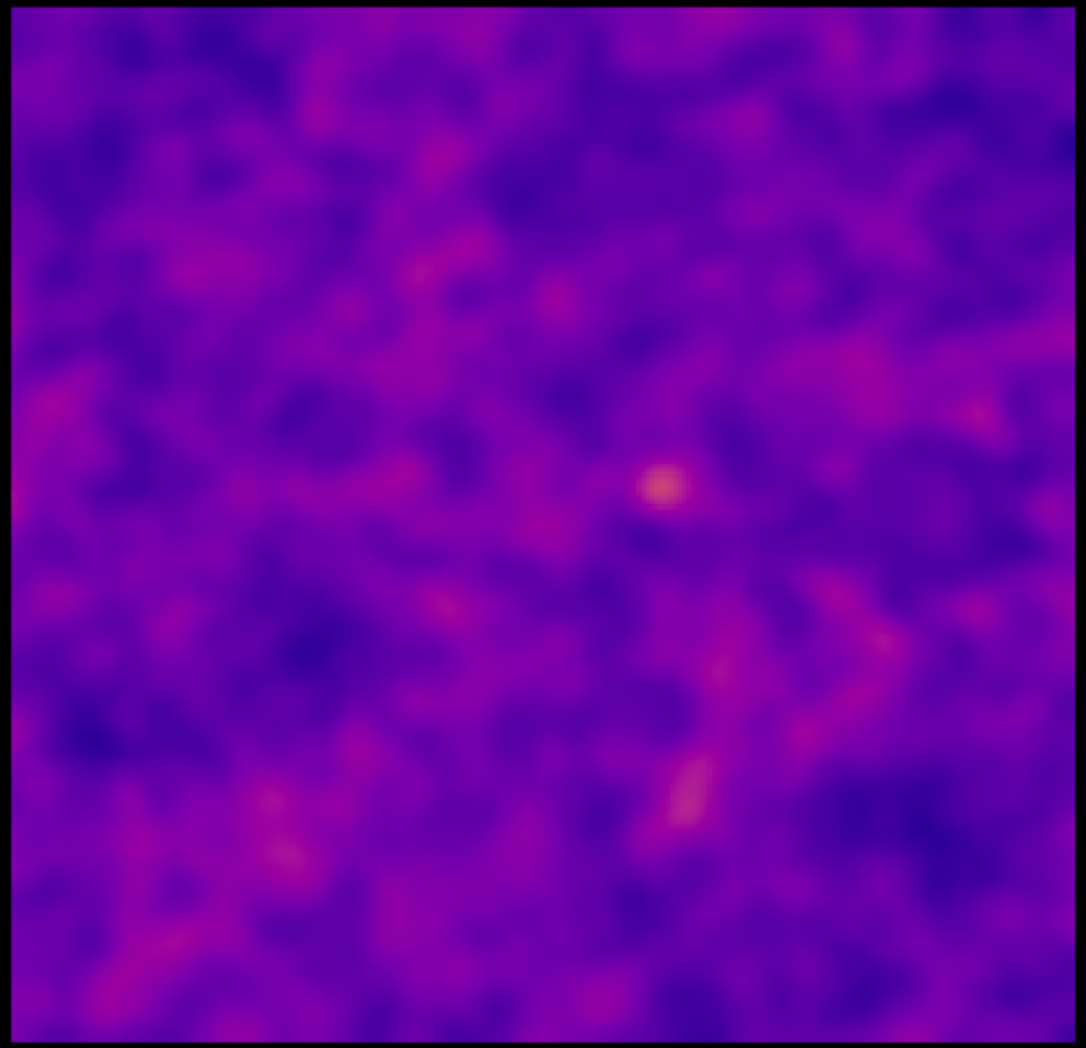
Ω_m

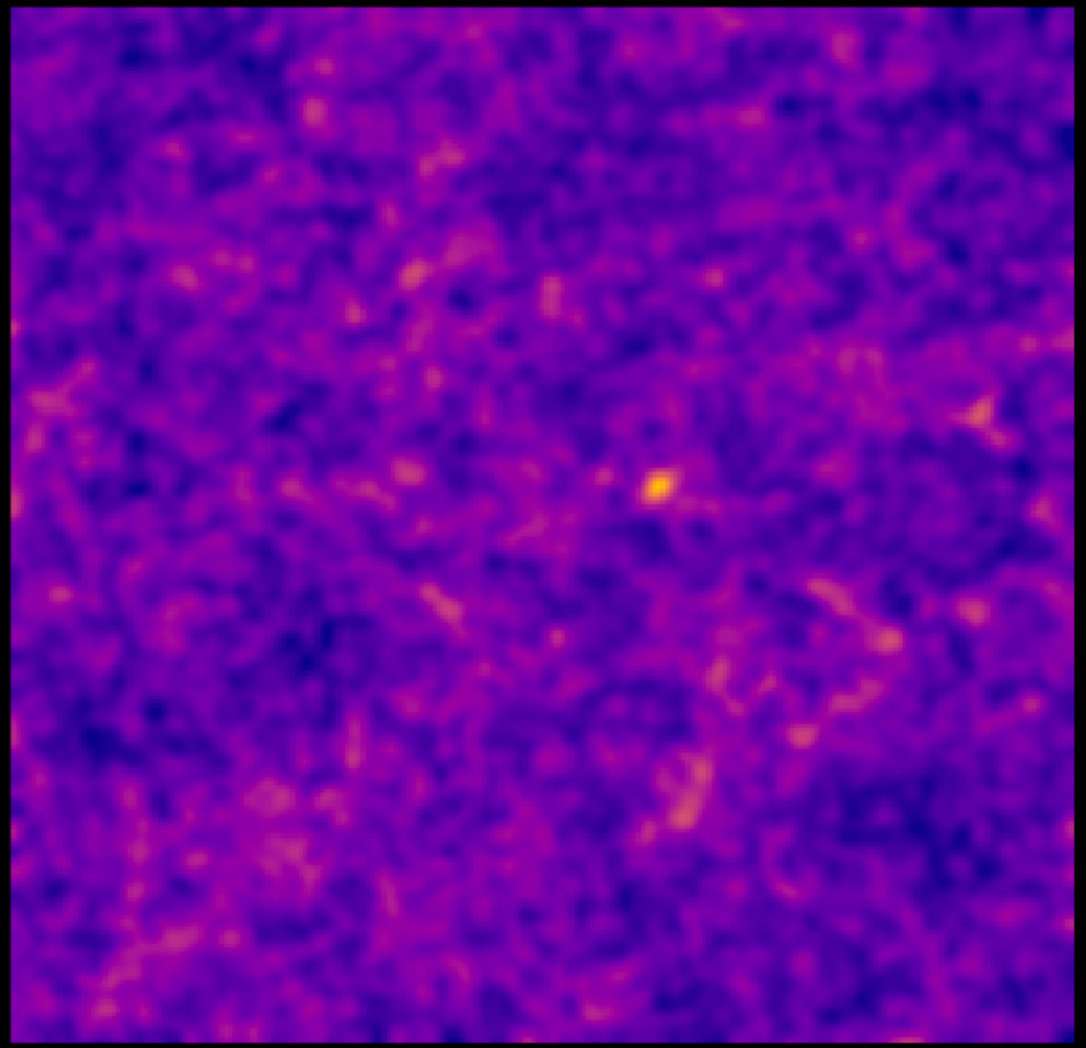
What is the advantage of deep learning for current and upcoming data?



add smoothing→

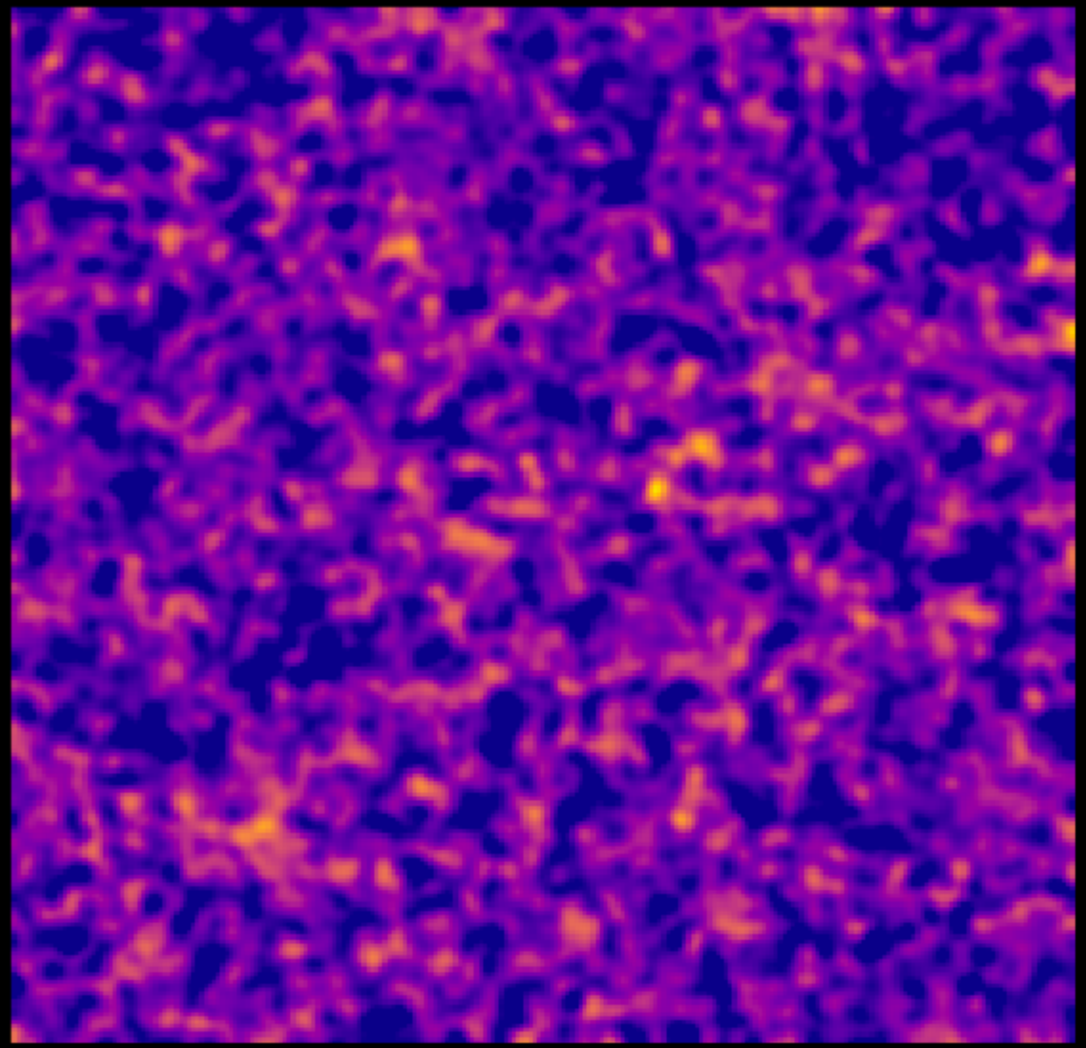
quality of simulations



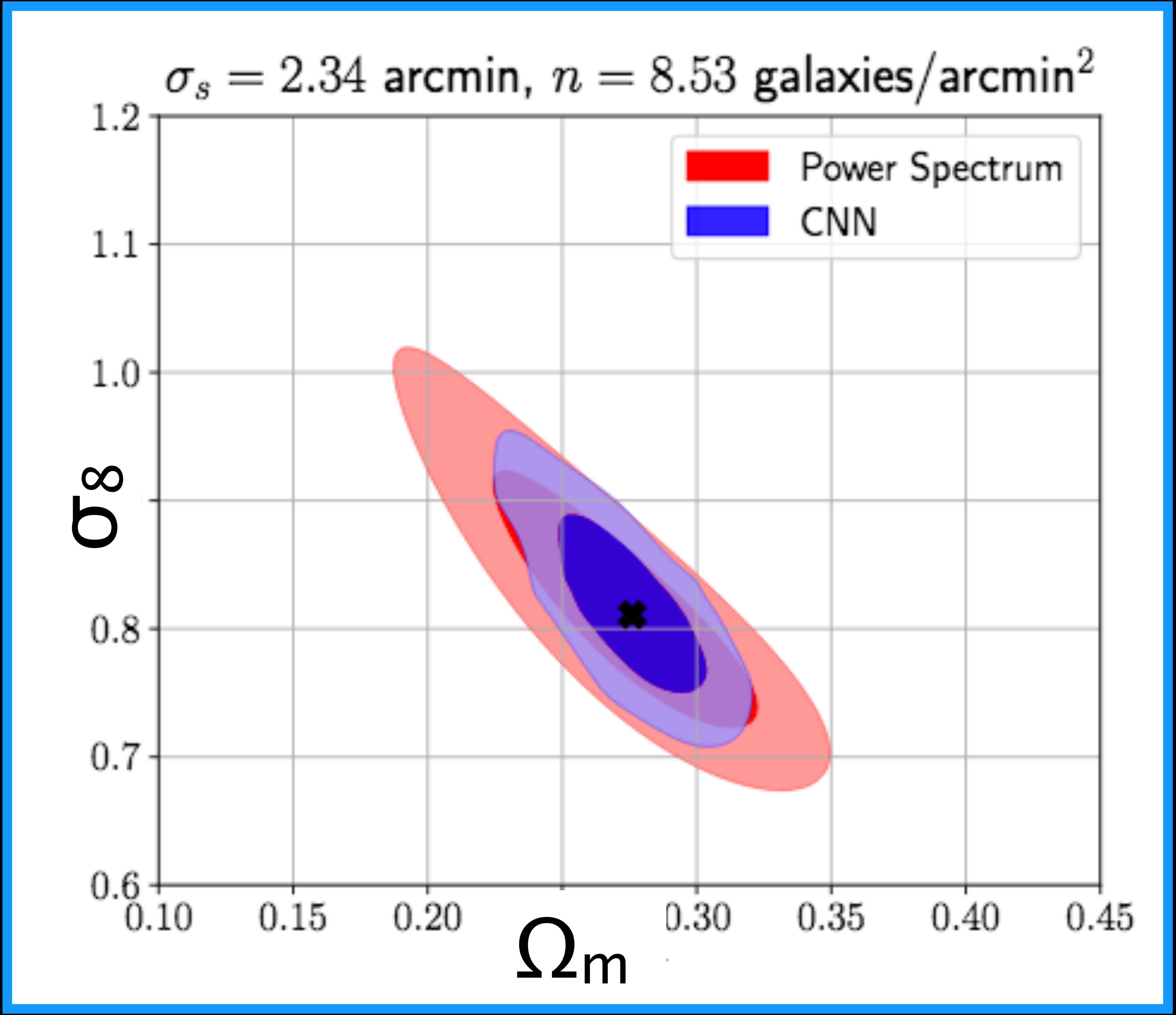


add noise →

quality of observations



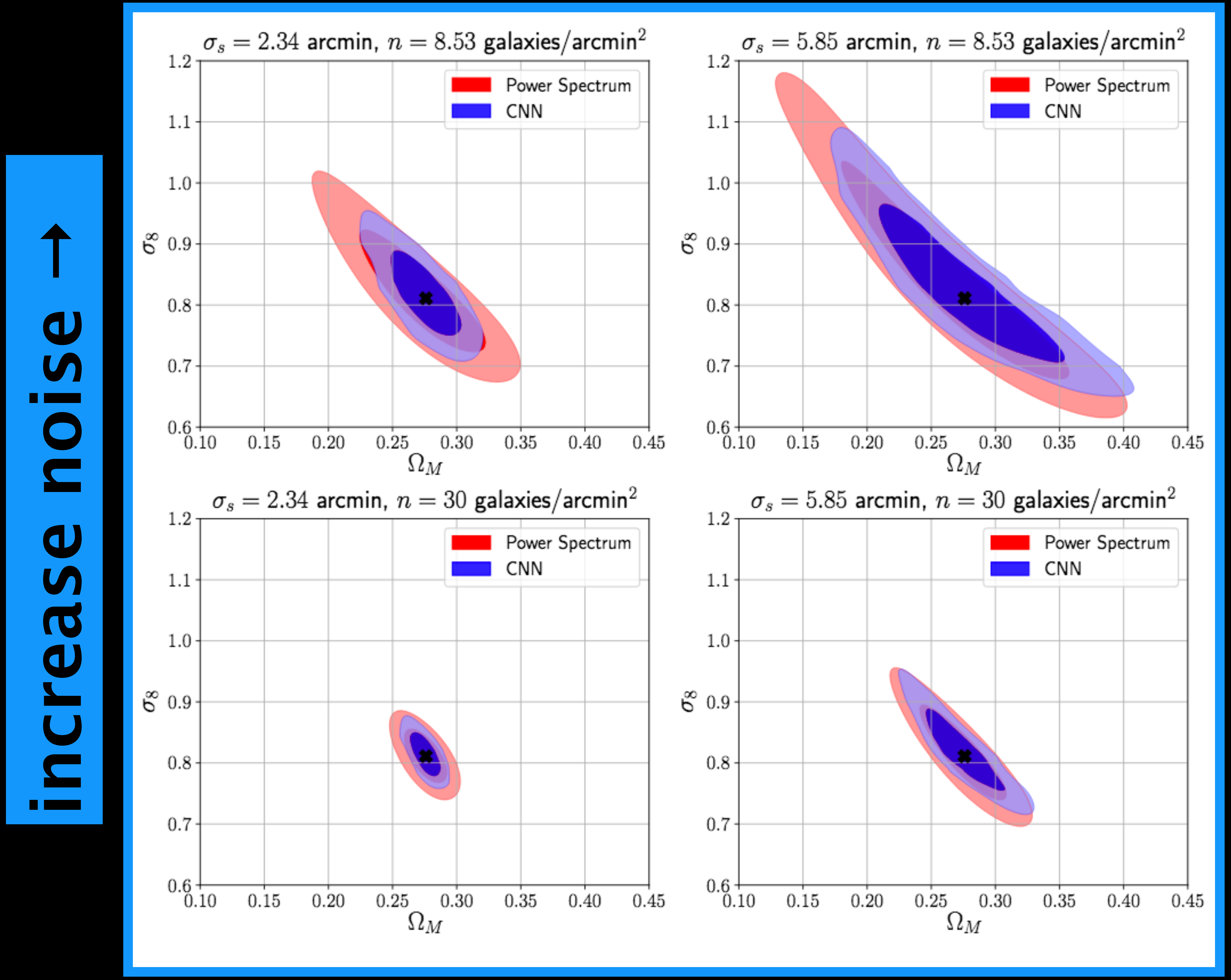
Deep learning captures more information



40% increase in constraining power!

Fluri, TK, et al. 1807.08732

Deep learning captures more information



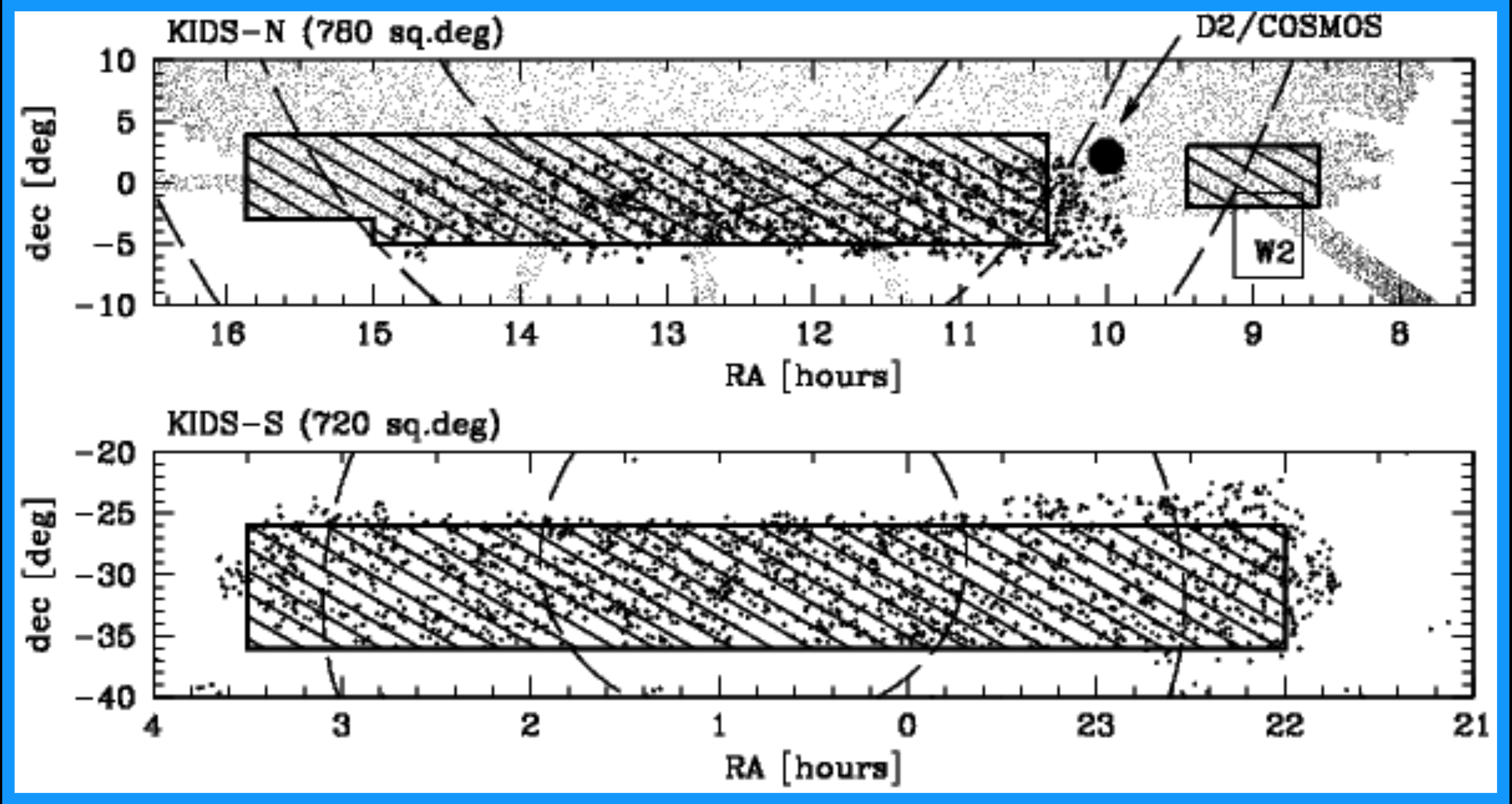
increase smoothing→

Gain in constraining power depends on noise level and smoothing scale

Fluri, TK, et al. 1807.08732

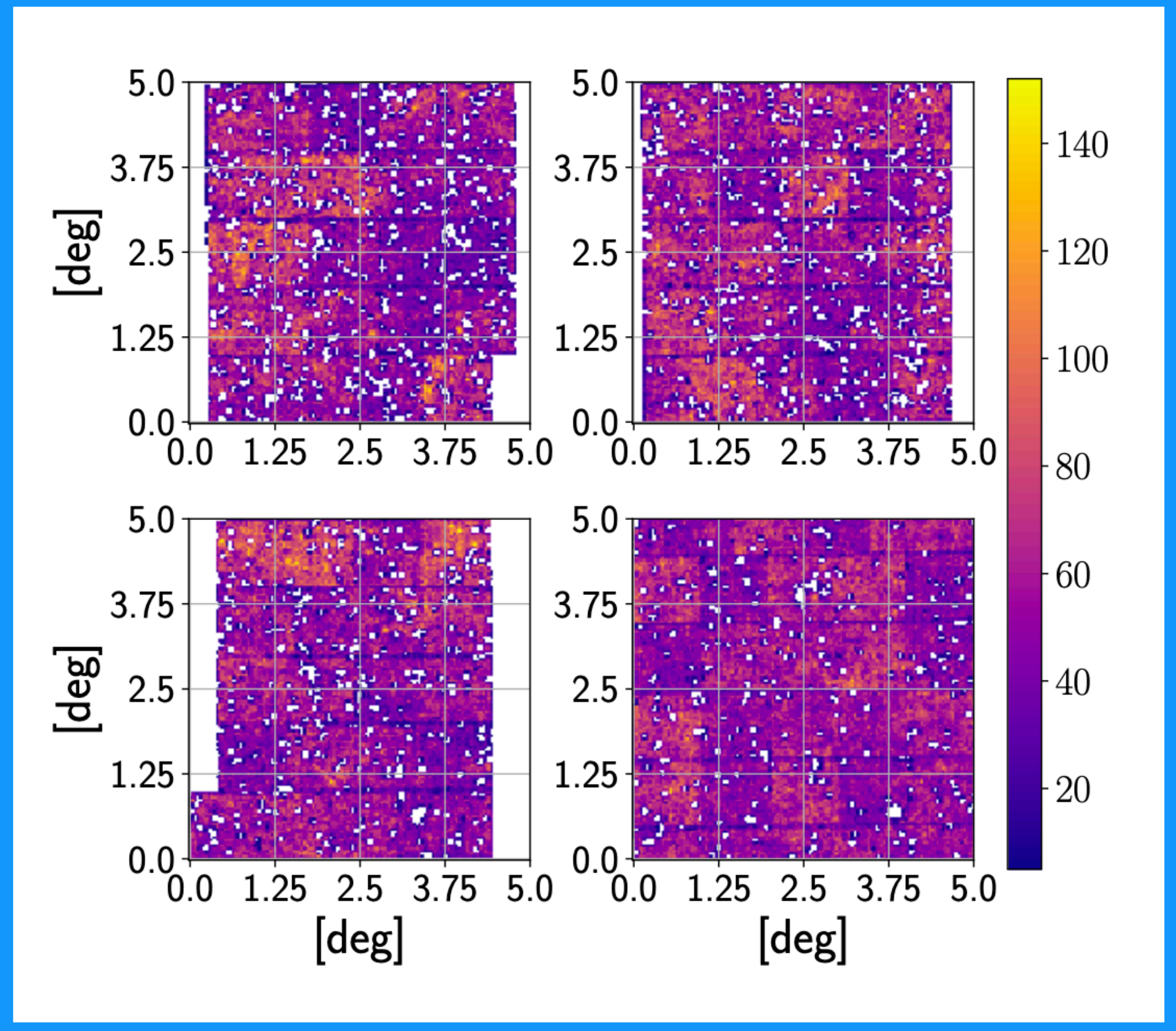
Analysis of KiDS-450 with deep learning

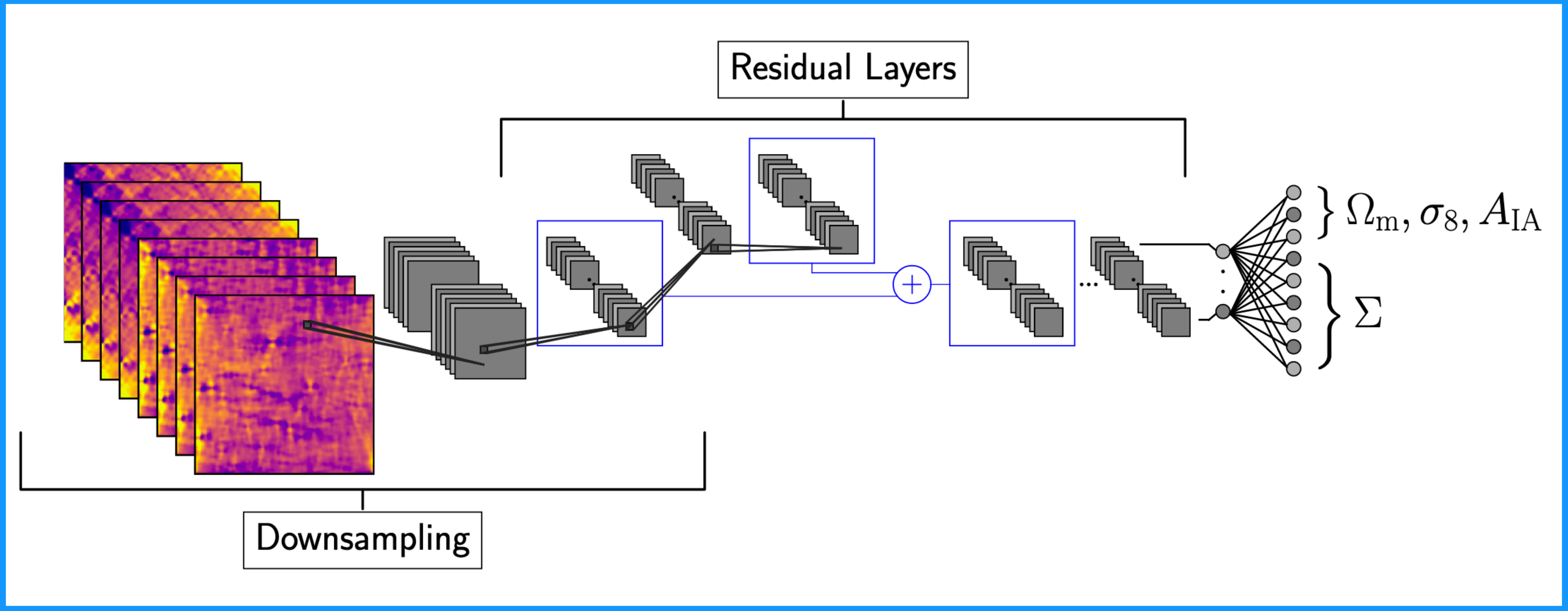
- Kilo-Degree Survey intermediate data set
- Using weak gravitational lensing to create the dark matter maps
- Low-noise dataset
- Taken by VLT Survey Telescope using OmegaCAM



kids.strw.leidenuniv.nl

Analysis of KiDS-450 with deep learning





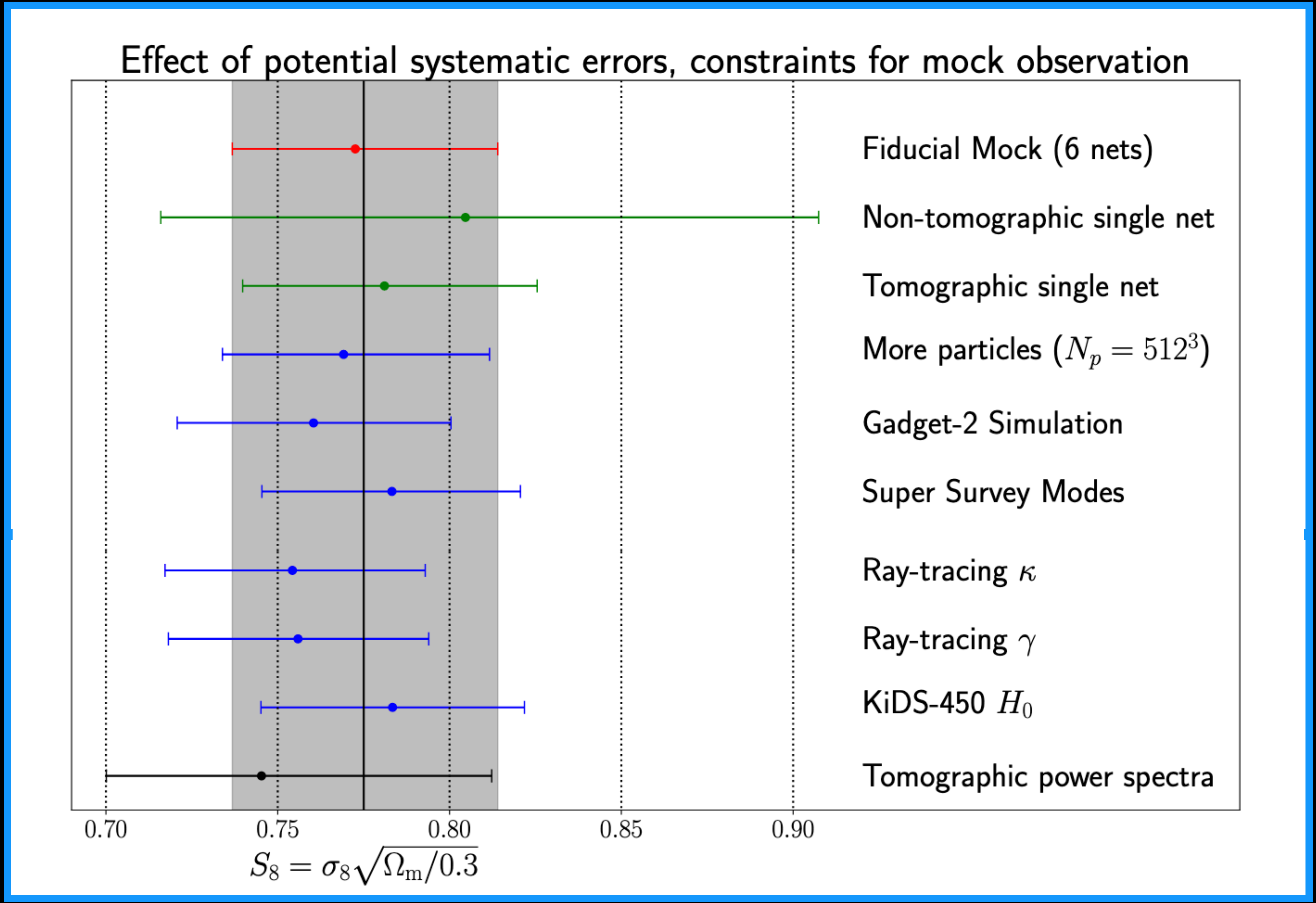
**data:
20 x 4
tomographic mass
maps**

network: 3 parameter output

likelihood analysis

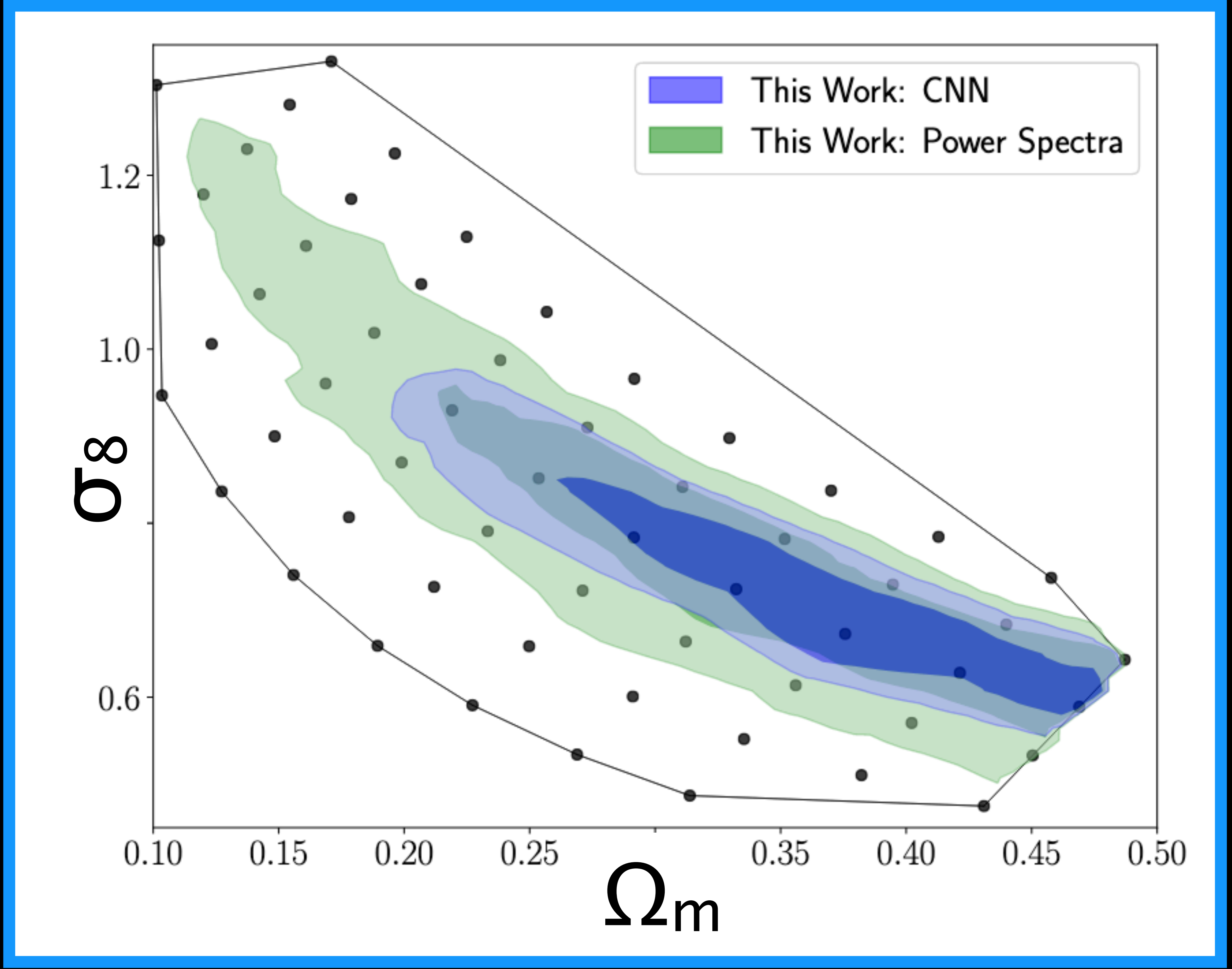
Fluri, TK, et al. 1906.03156

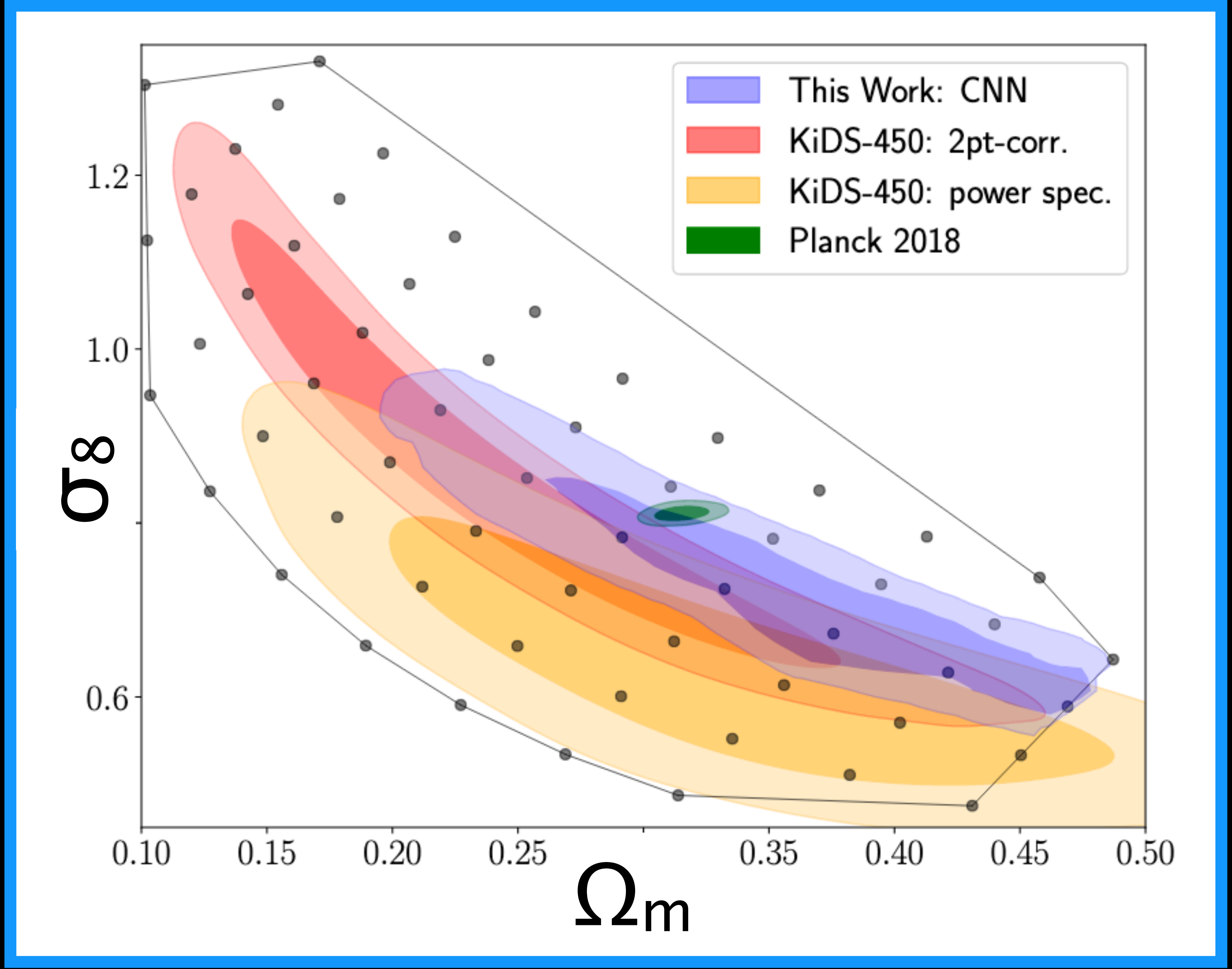
KiDS-450: robustness to simulation details



Fluri, TK, et al. 1906.03156

Analysis of KiDS-450 with deep learning





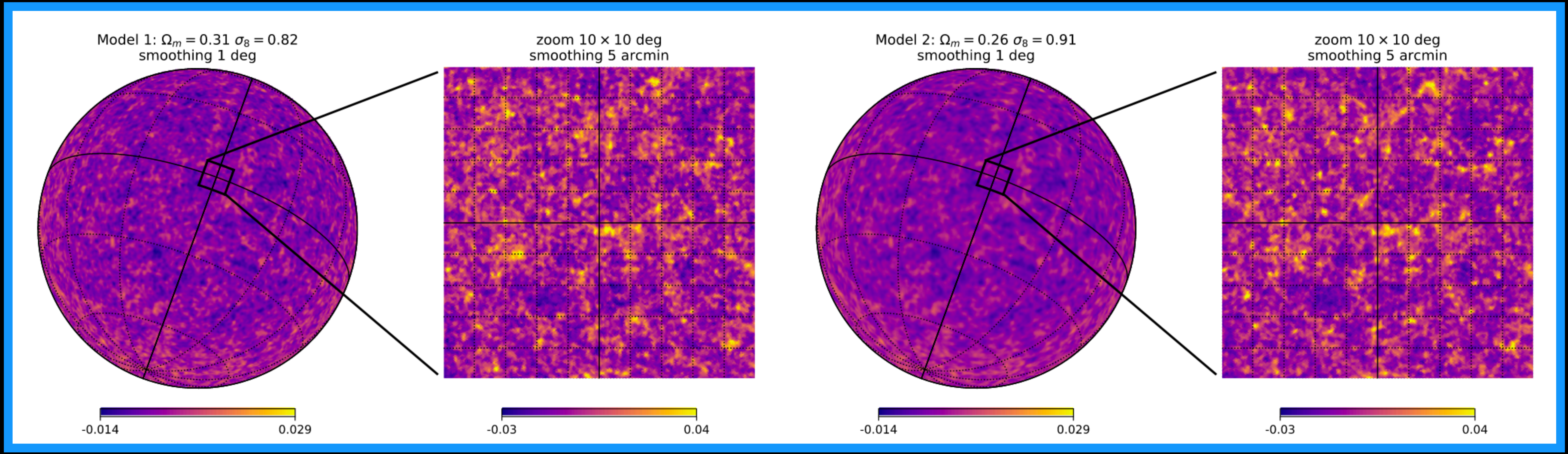
$S_8 = \sigma_8(\Omega_m/0.3)^{0.5} = 0.777 +/- 0.037$

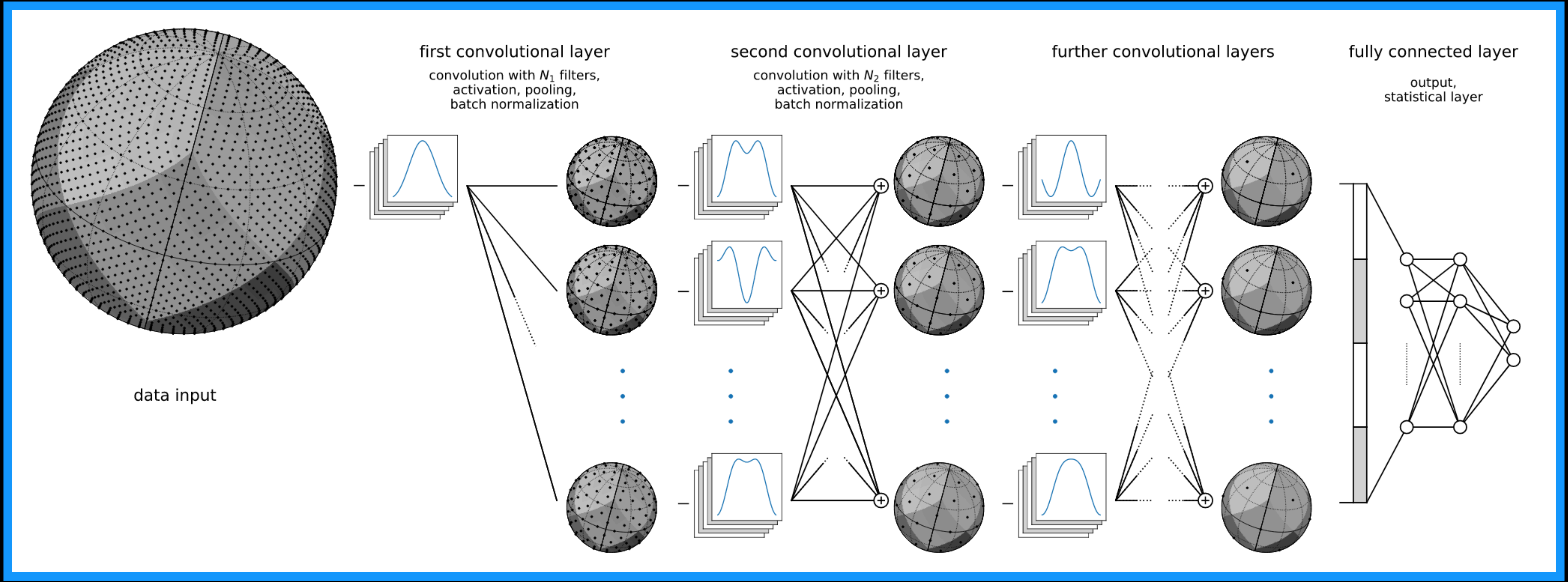
30% increase of constraining power compared to equivalent analysis with the power spectra

blinding strategy used in the analysis

Fluri, TK, et al. 1906.03156

Deep learning on the sphere: a tool for wide area surveys

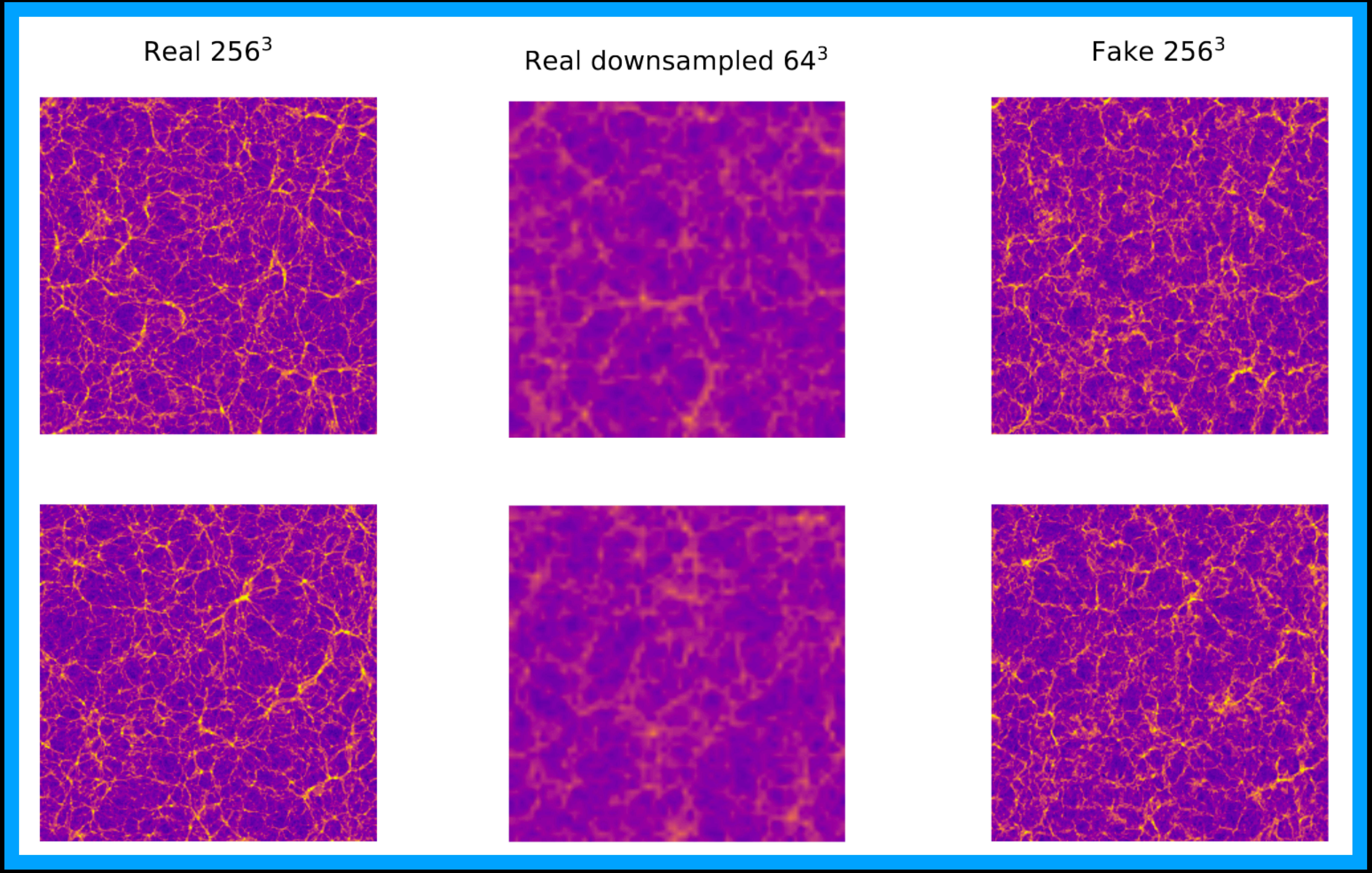




github.com/SwissDataScienceCenter/DeepSphere

Perraudin, TK, et al. 1810.12186

Experimenting with GANs for N-body simulations



256^3 voxel cubes
Sequential Patch-GAN approach for upsampling
Challenge: data and codes available → beat our score!

github.com/nperraud/3DcosmoGAN

Perraudin, TK, et al. 1908.05519

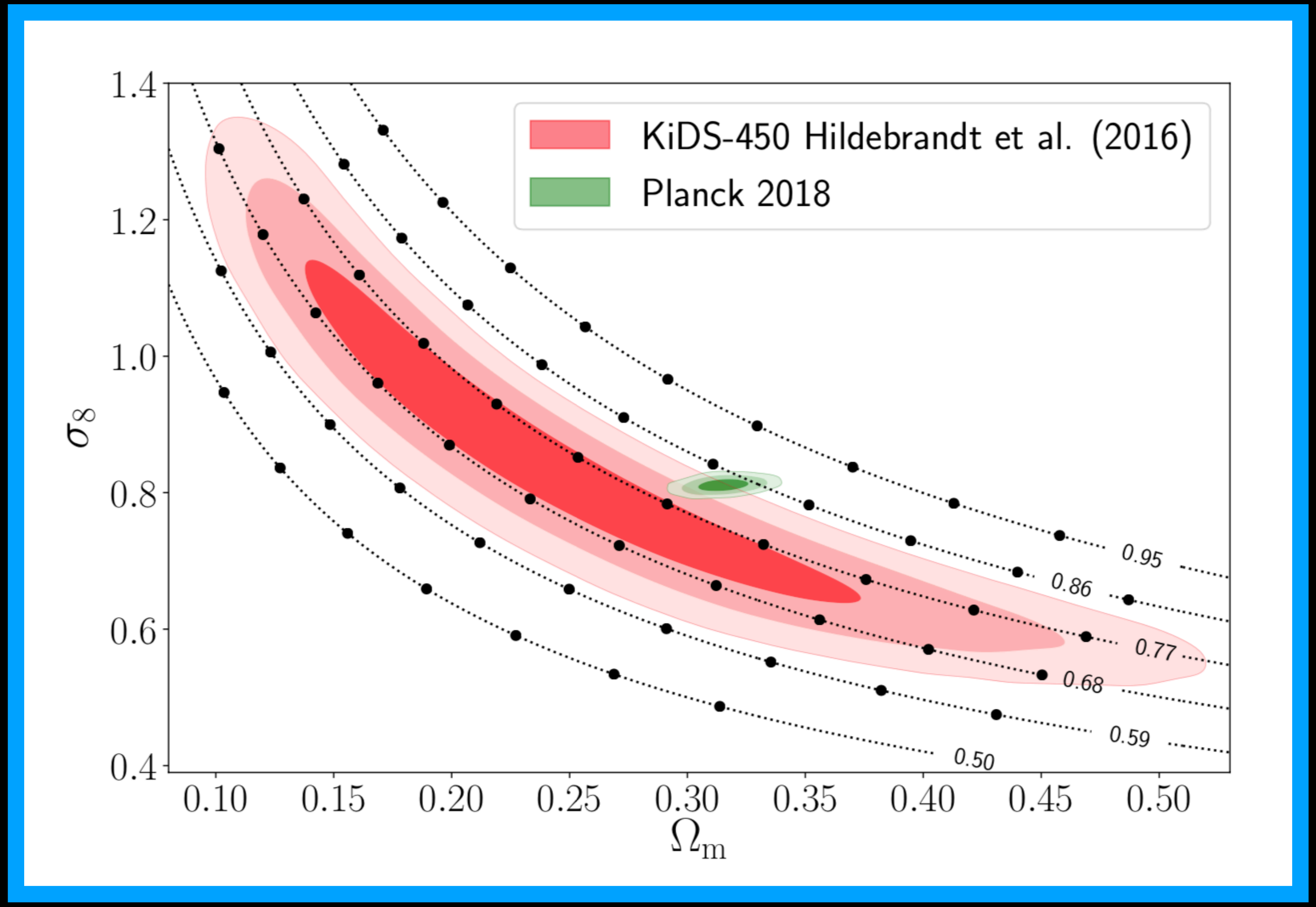
Summary

- Deep learning can improve the information gain from existing, non-linear cosmological data
- Moving towards simulation-based inference
- Currently limited by theory modelling and simulations

Extra slides

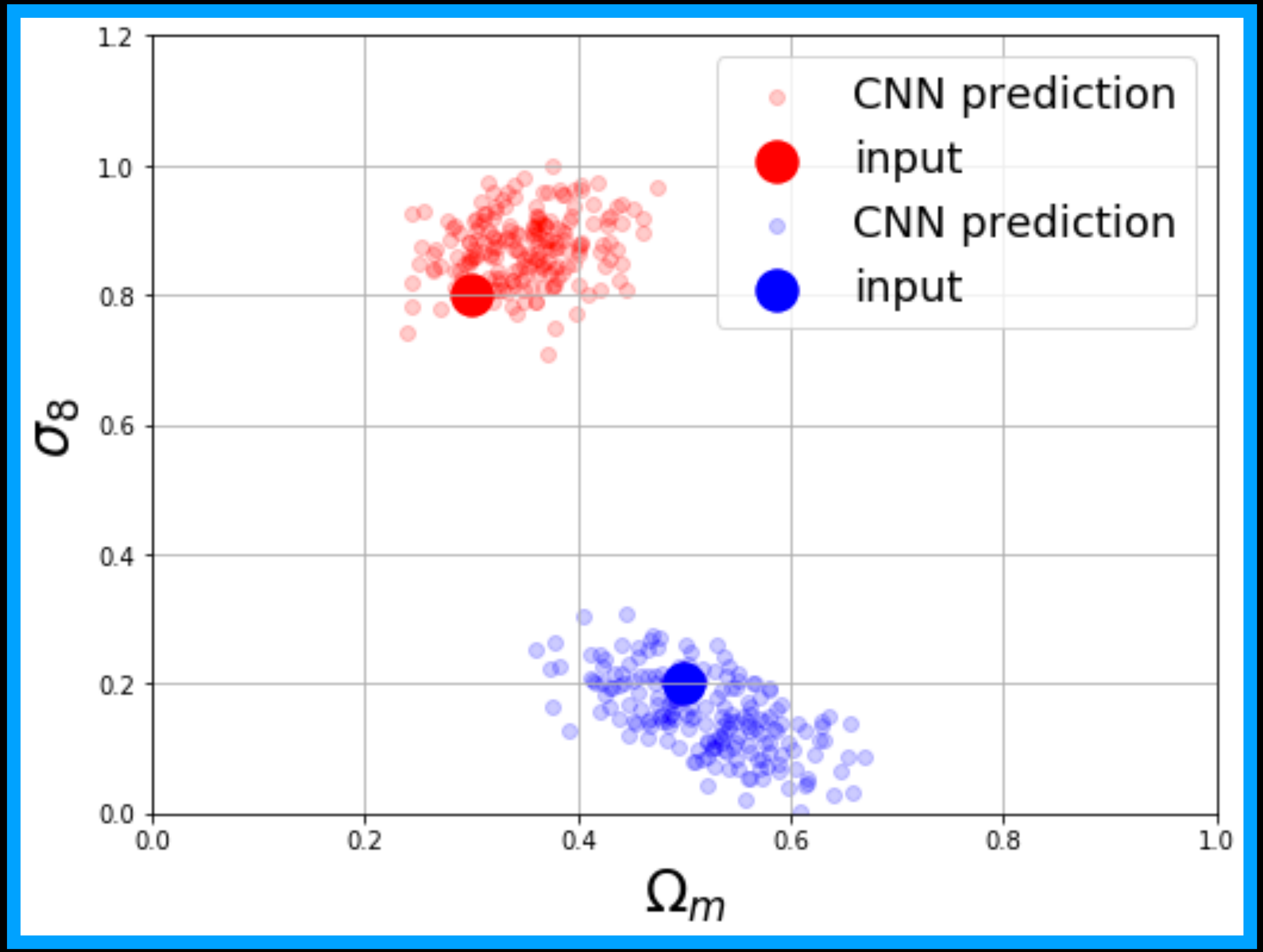
Simulations

- Very heavy simulation load
- PkdGrav3 on GPUs
- 90% of time for simulations, 10% for CNN training
- 57 cosmologies x 12 N-body simulations, convergence and IA map creation
- Inclusion of effects: baryons, advanced intrinsic alignment models



Likelihood modelling

- CNN is a feature extractor
- we create likelihood of features given true value
- we model the likelihood as a Gaussian with mean and covariance varying as a function of cosmological parameters
- we calculate the posterior with a prior across the spaces spanned by the training set



Training details

- ResNets with 10, 15 and 25 layers
- On the fly noise addition
- Likelihood loss used in training

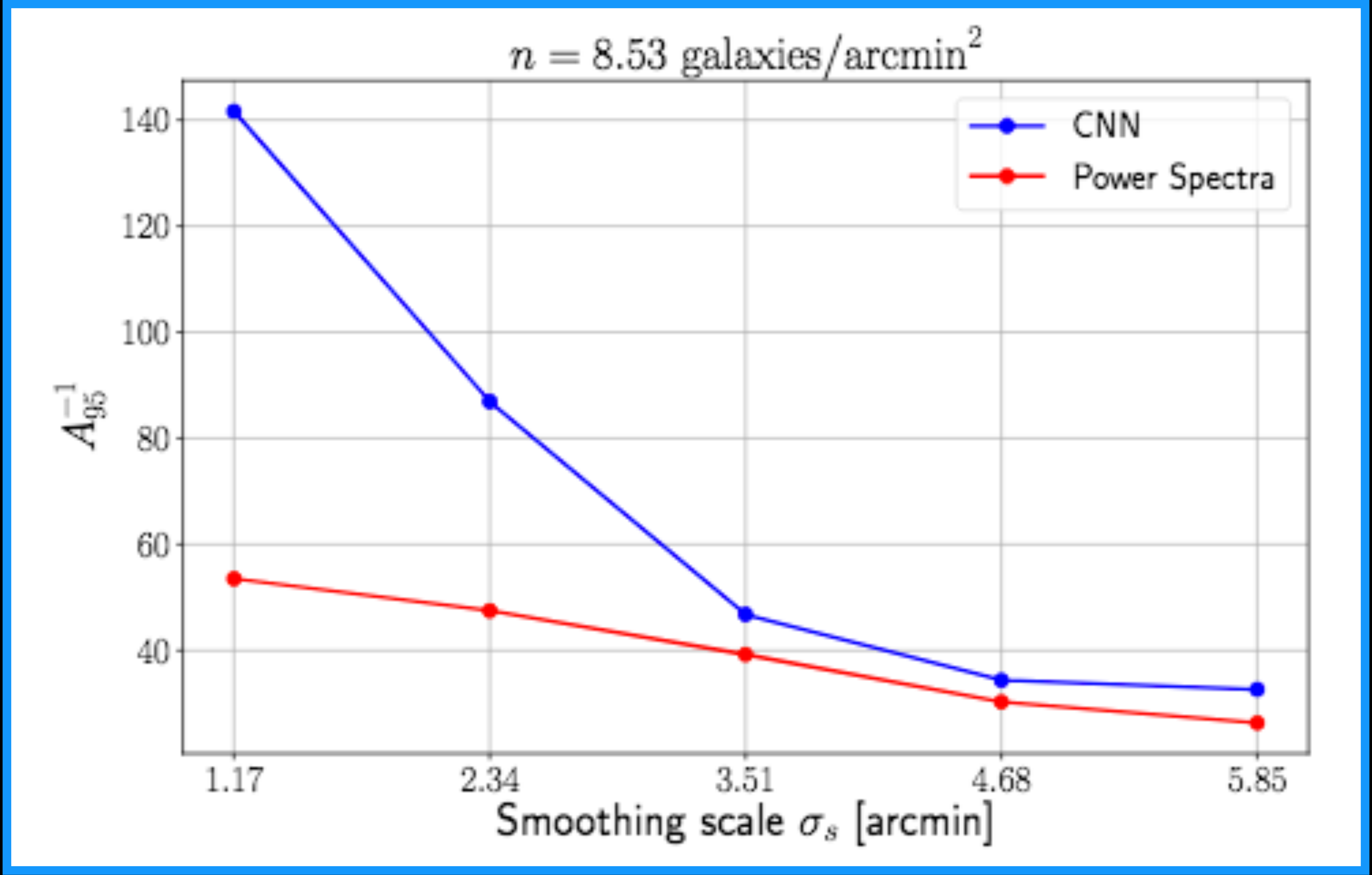
$$L = \frac{1}{2}\left(\ln\left(|\Sigma|\right) + \left(\theta_p - \theta_t\right)^T \Sigma^{-1} \left(\theta_p - \theta_t\right)\right),$$

Σ **prediction of covariance matrix learned by the network (6 parameters)**

θ_p **predicted parameters (3)**

θ_t **target parameter (3)**

Deep learning captures more information



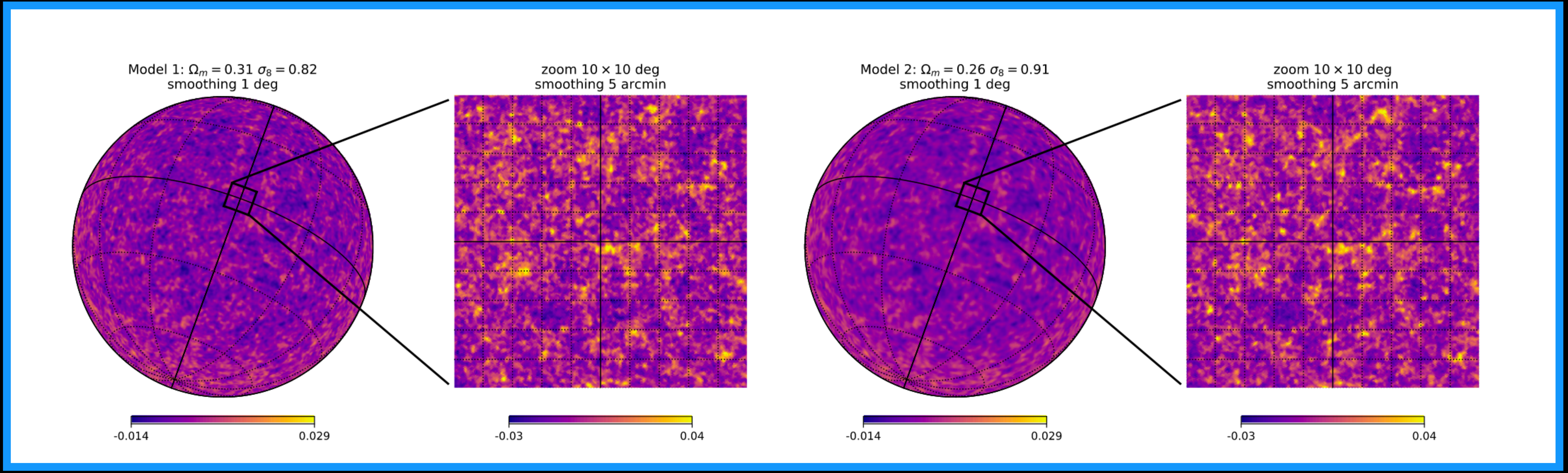
← worse inverse contour area better →

← smoothing scale →

Fluri, TK, et al. 1807.08732

Deep learning on the sphere

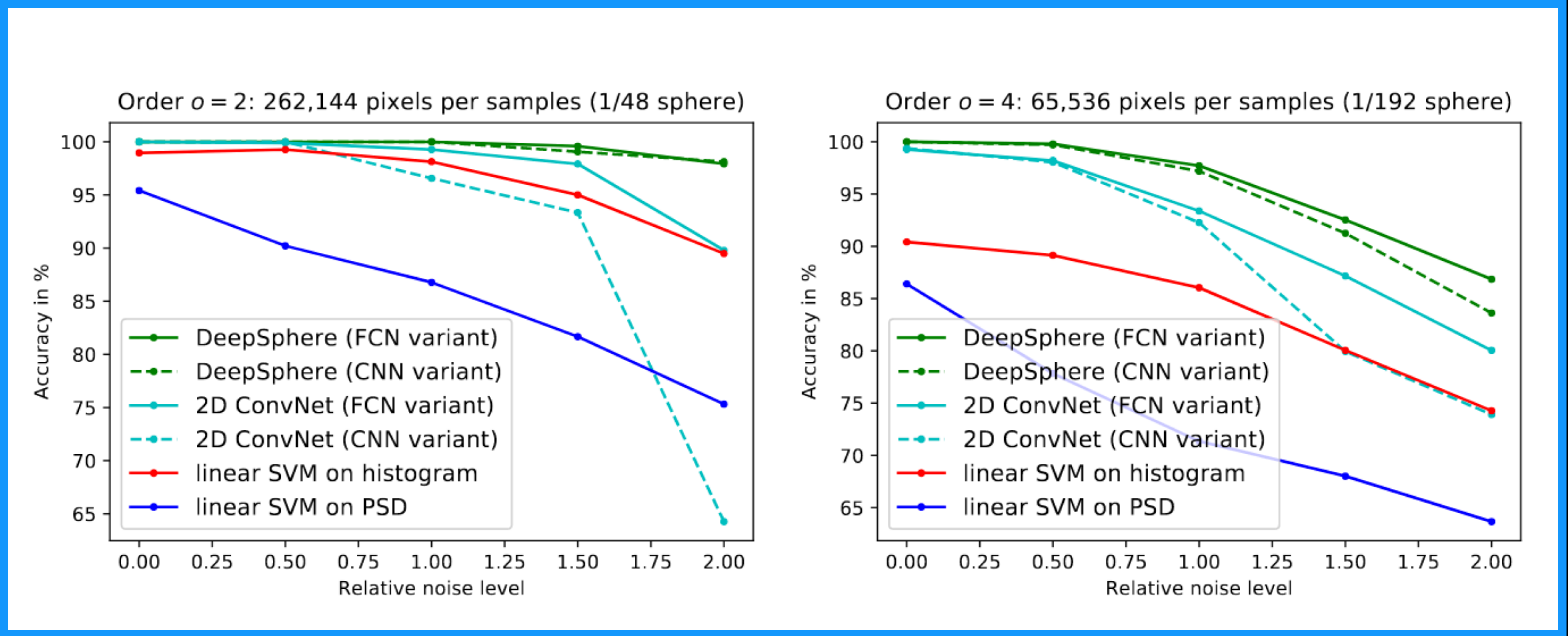
σ_8=0.82, Ω_m=0.31 | **σ_8=0.91, Ω_m=0.26**



two models with the same power spectrum

Perraudin, TK, et al. 1810.12186

Deep learning on the sphere



DeepSphere maintains advantage over baselines
for higher noise levels

github.com/SwissDataScienceCenter/DeepSphere

Perraudin, TK, et al. 1810.12186