Capitalising on Artificial Intelligence for Large-Scale Structure Cosmology

Tomasz Kacprzak (Swiss Data Science Center and ETH Zurich)

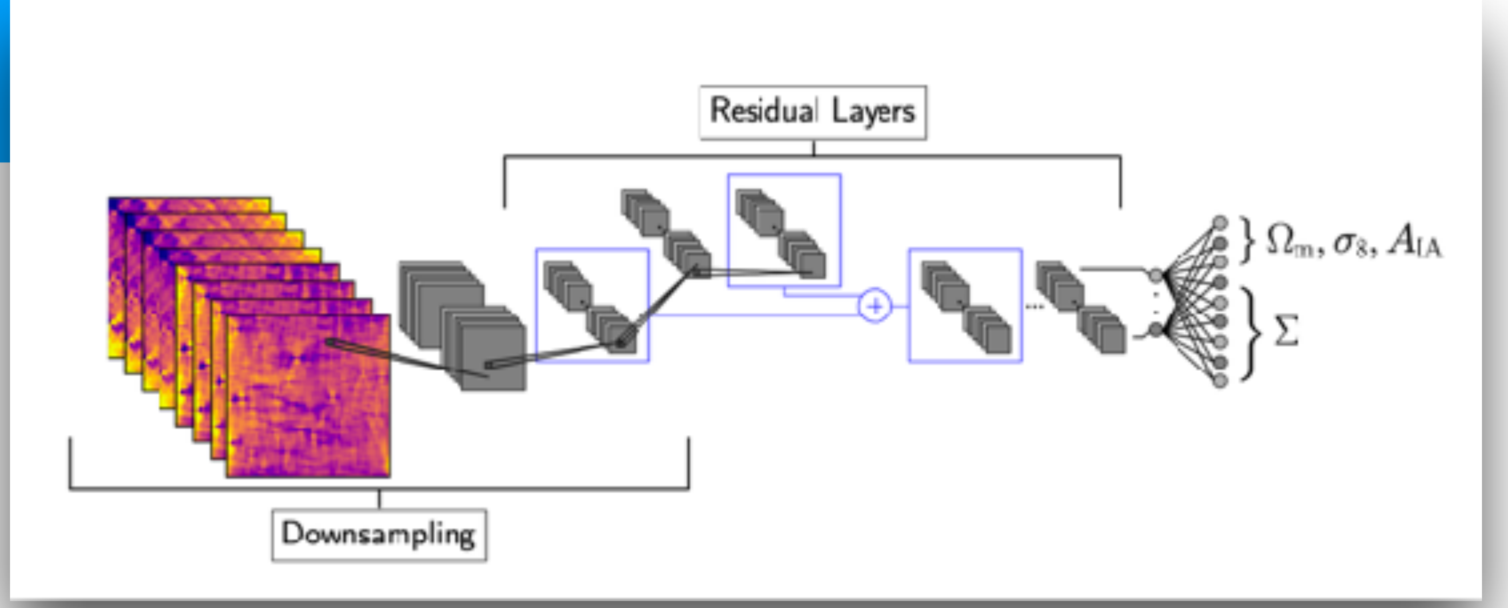
ML-IAP/CCA-2023, Paris, 30 Nov 2023

ETH zürich

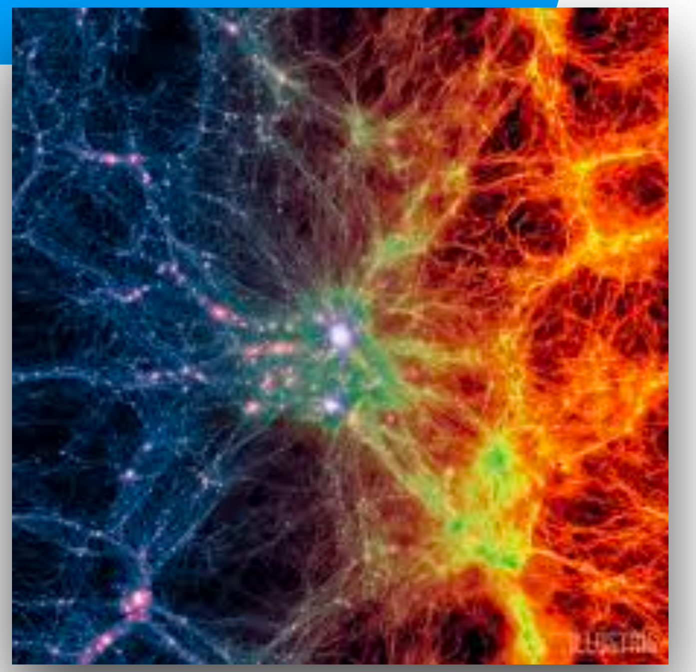
SDSC

How can AI open new possibilities in cosmological analysis of LSS?

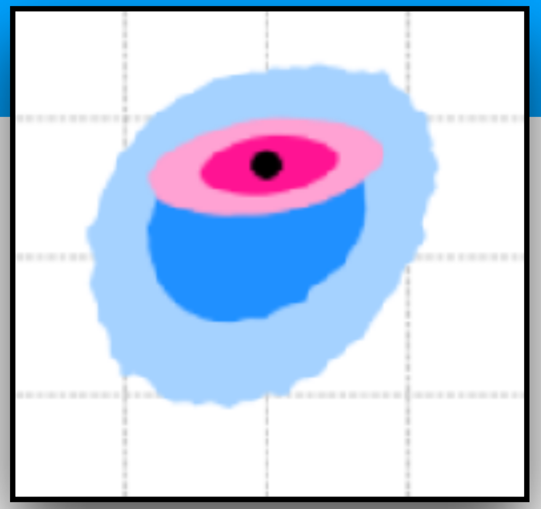
Reaching the information floor of the data



Accelearating simulations

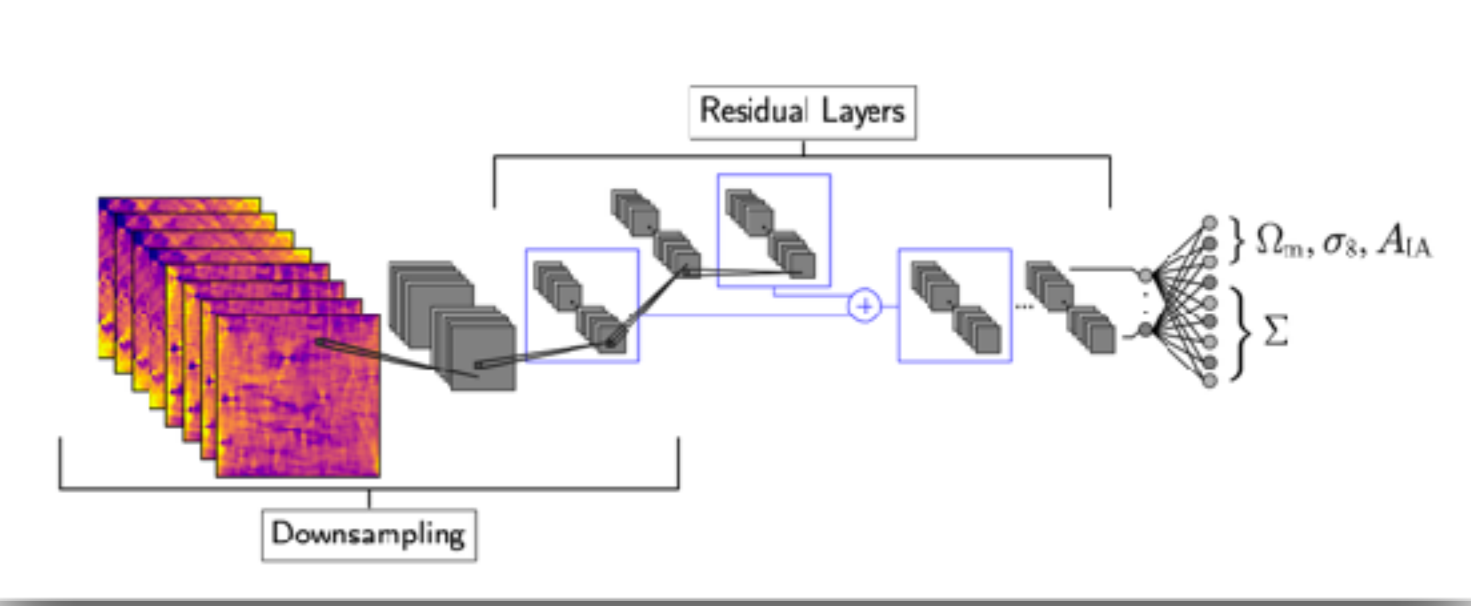


Breaking degeneracies between cosmology and systematics

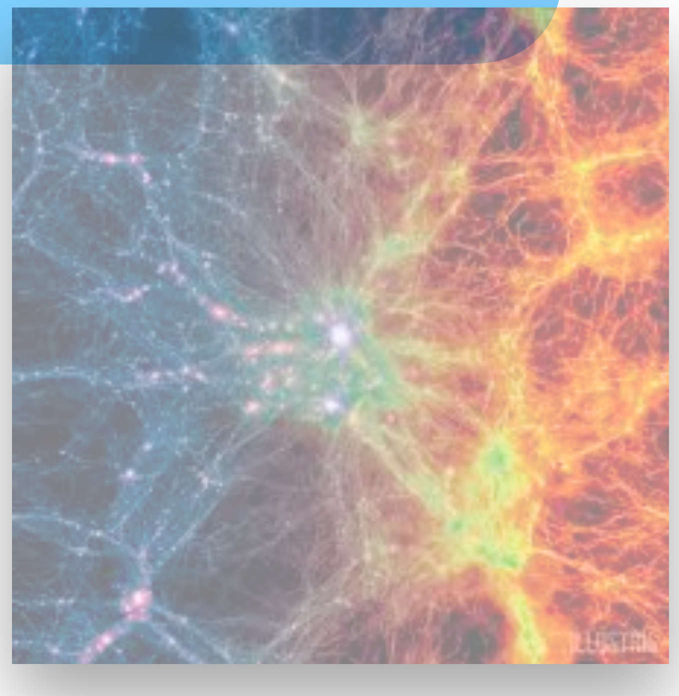


How can AI open new possibilities in cosmological analysis?

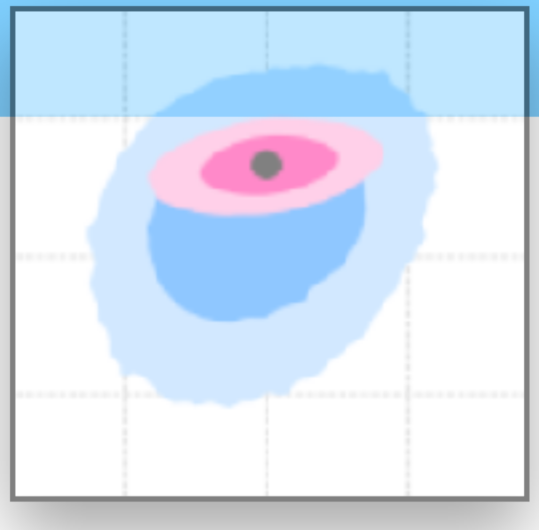
Reaching the information floor of the data



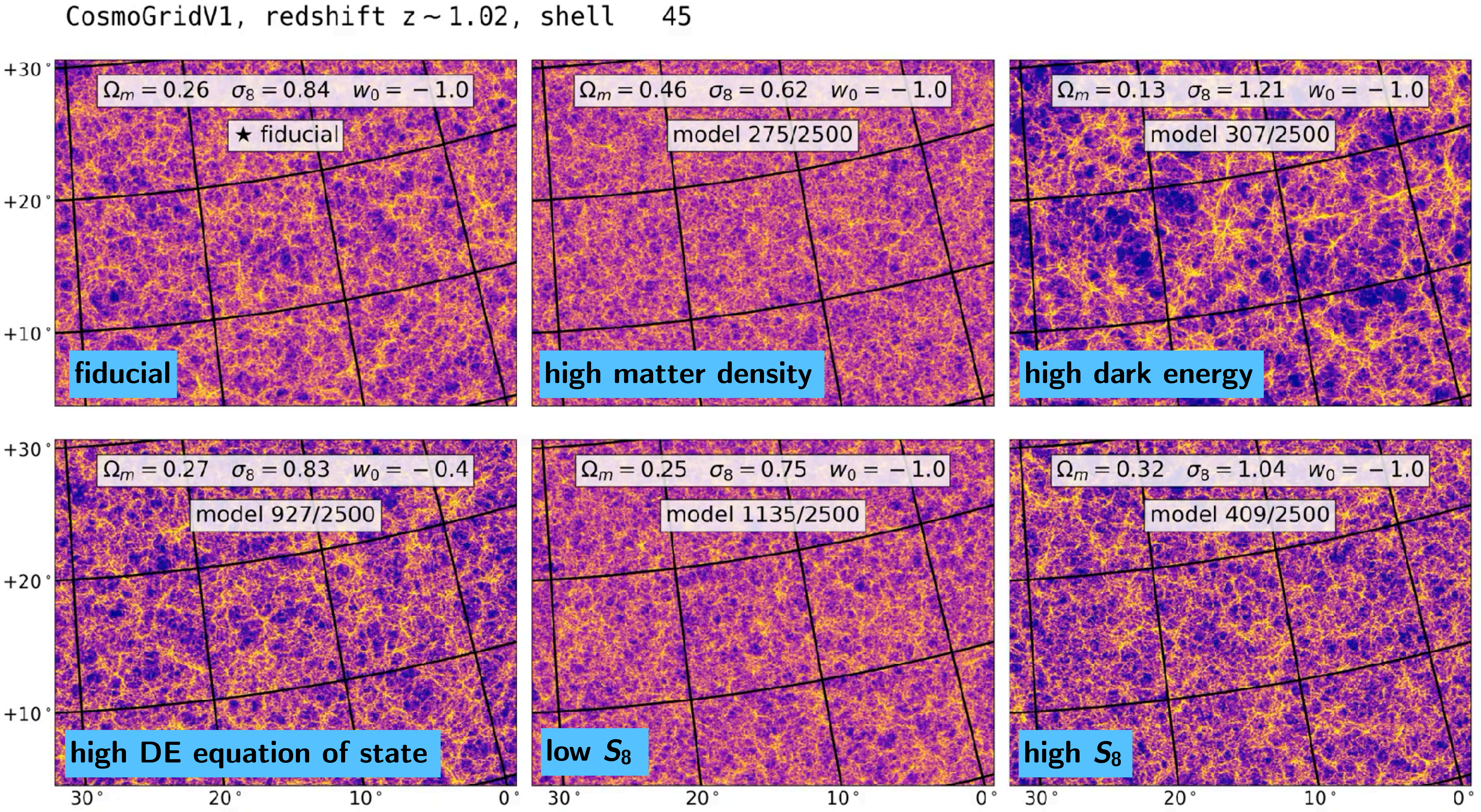
Accelearating simulations



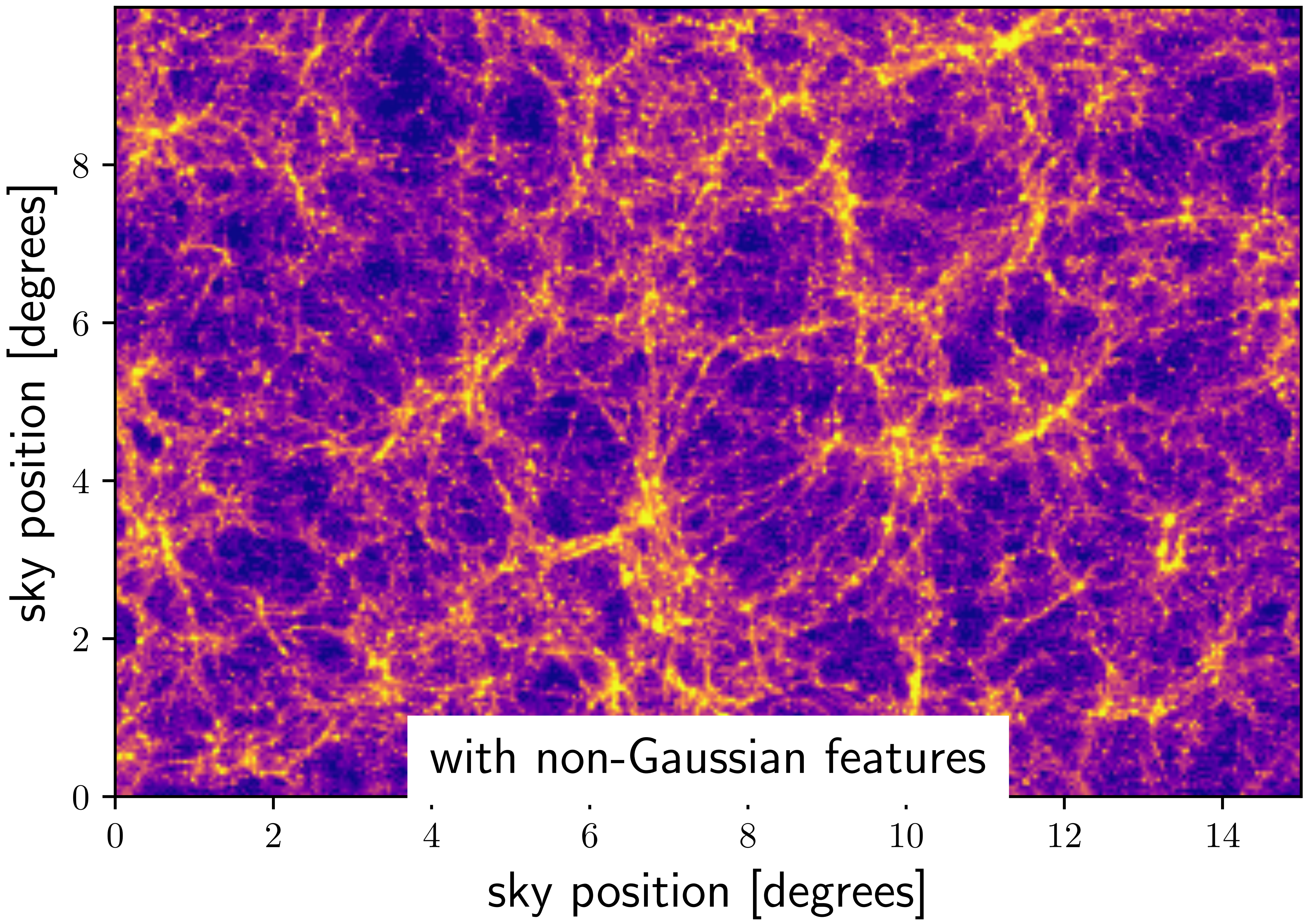
Breaking degeneracies between cosmology and systematics



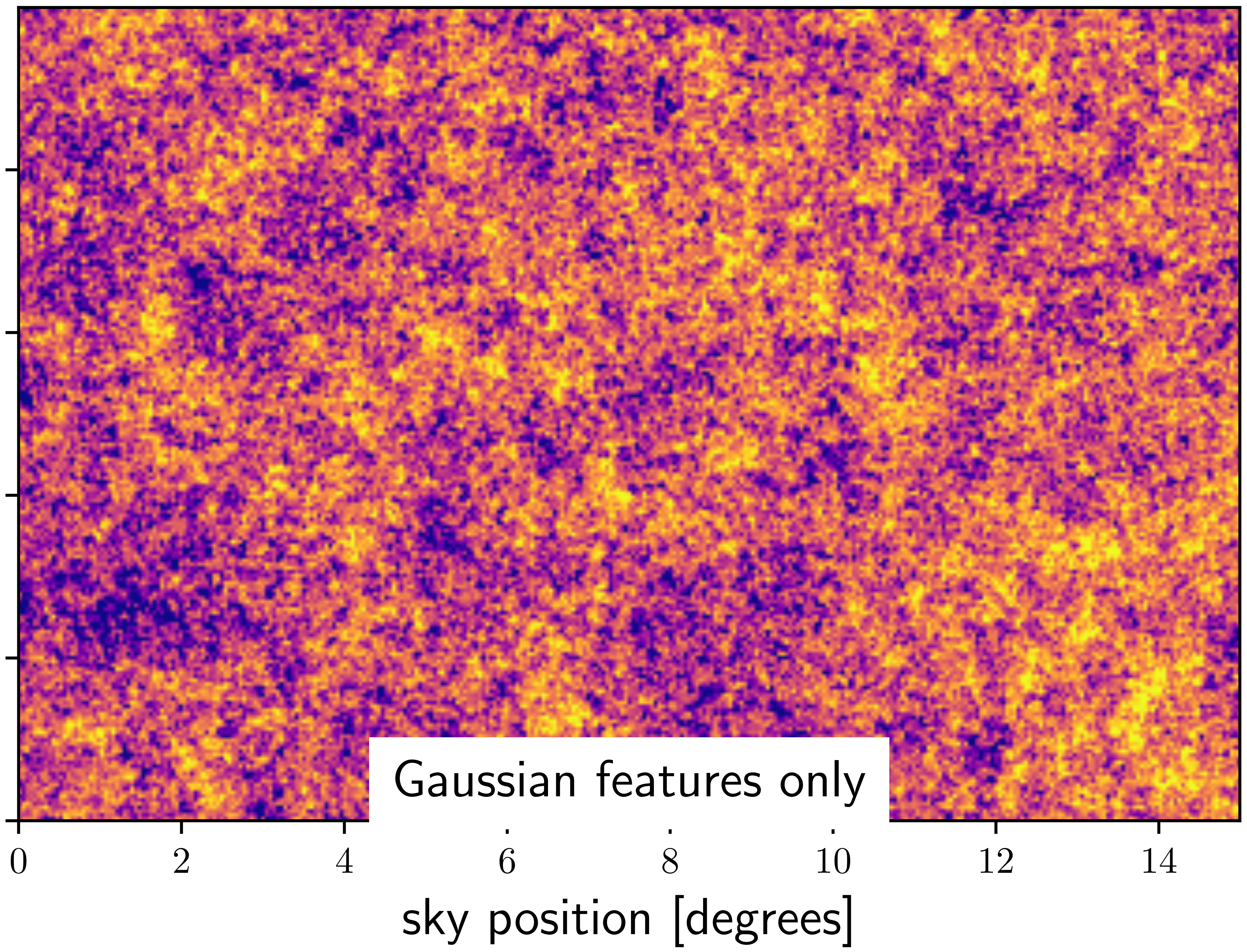
Dark matter distributions carries information about cosmological parameters



Large Scale Structure is highly non-Gaussian



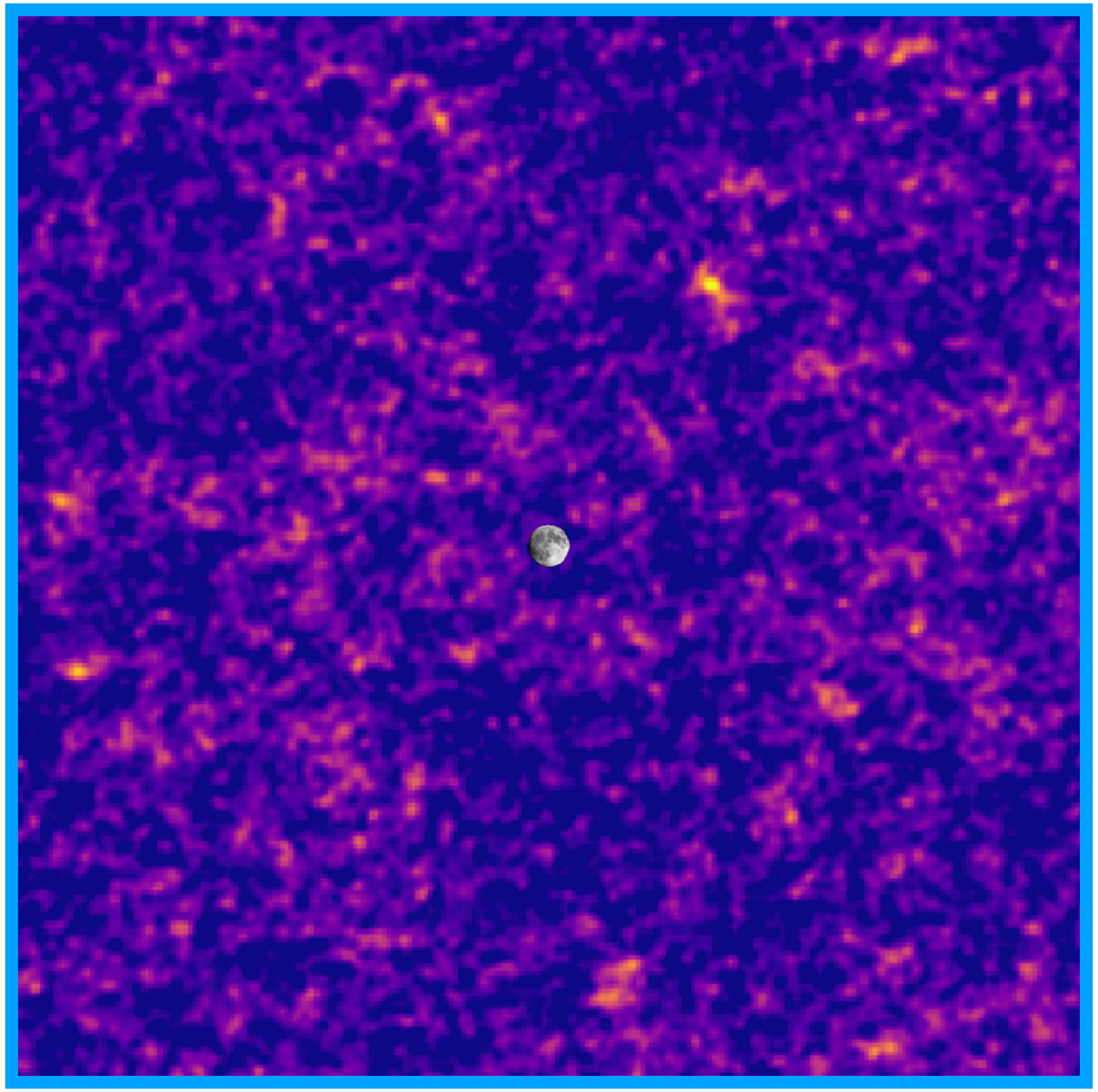
N-body simulation slice



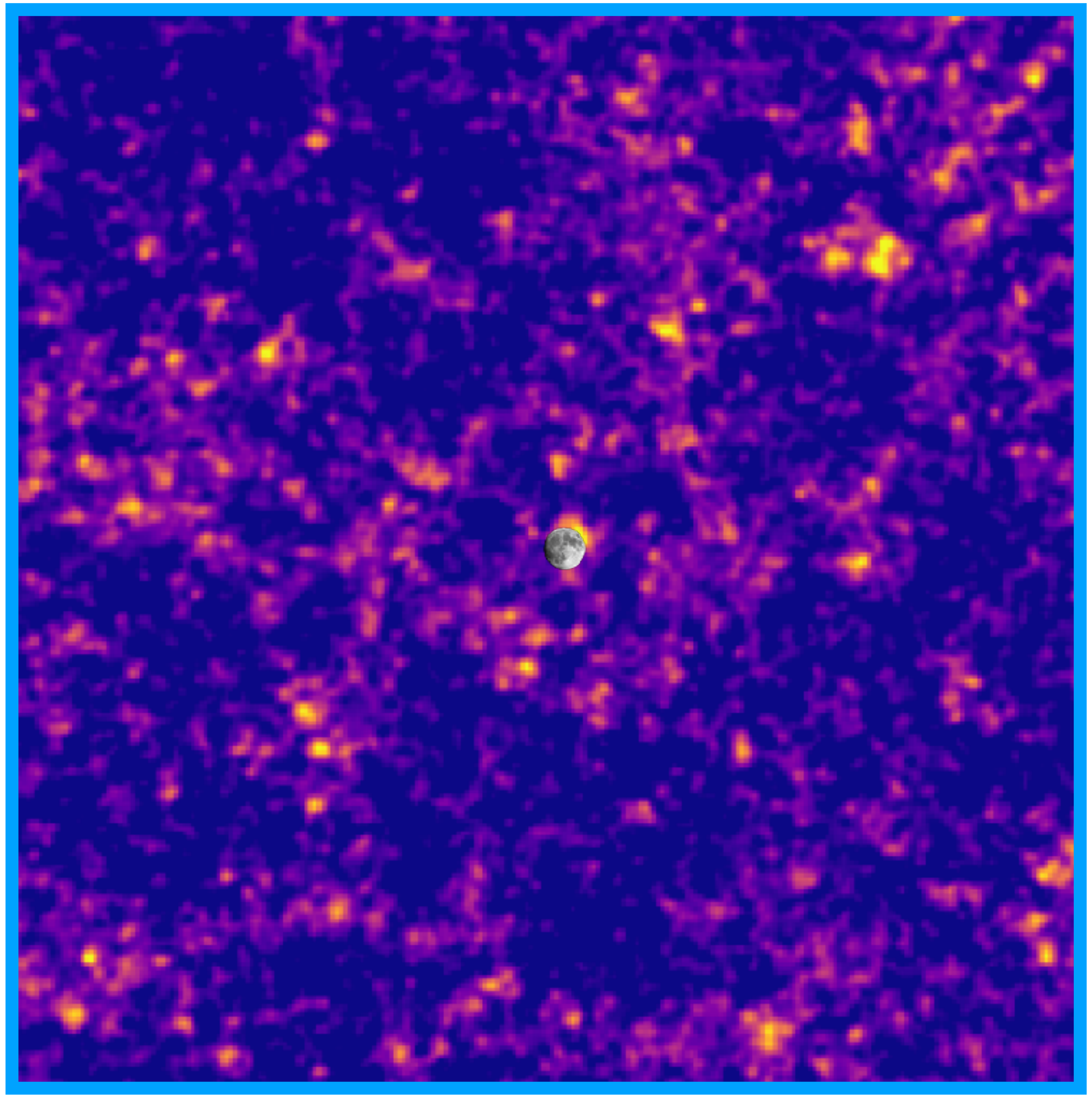
Gaussian Random Field with the same power spectrum as the N-body slice

these maps have the same power spectra

Weak lensing matter mass maps: projected matter distribution

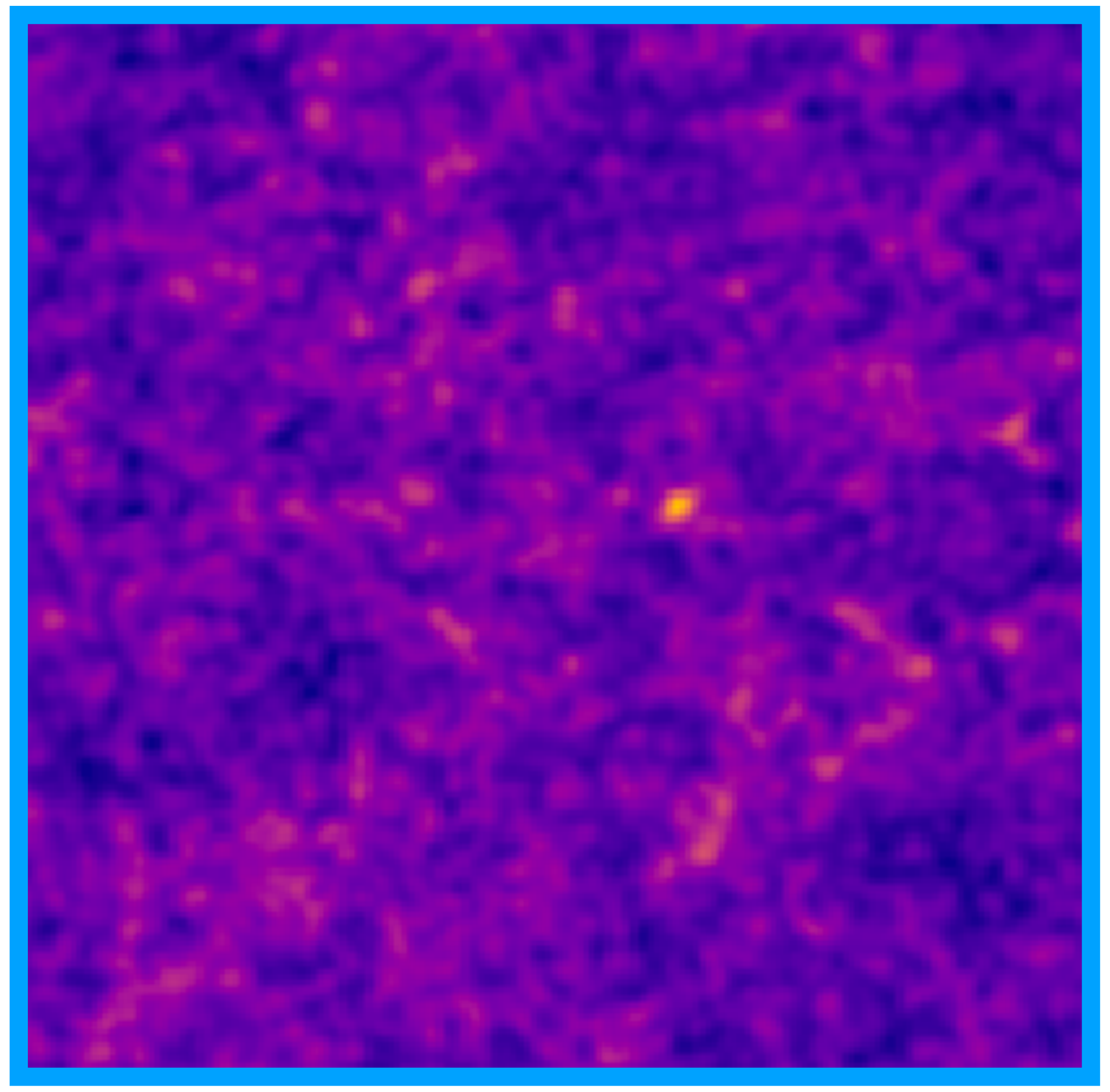


low σ_8 low Ω_m



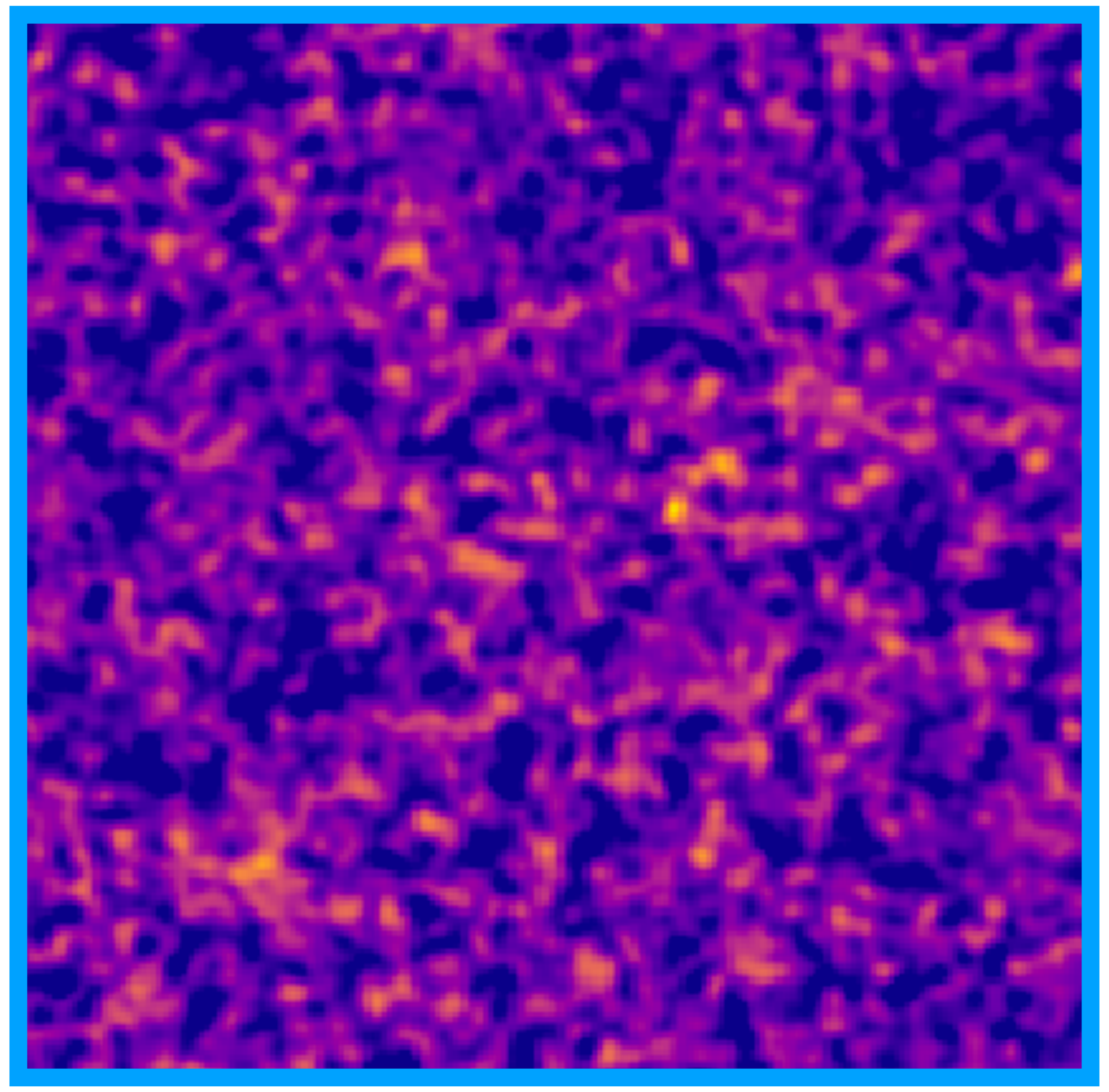
high σ_8 high Ω_m

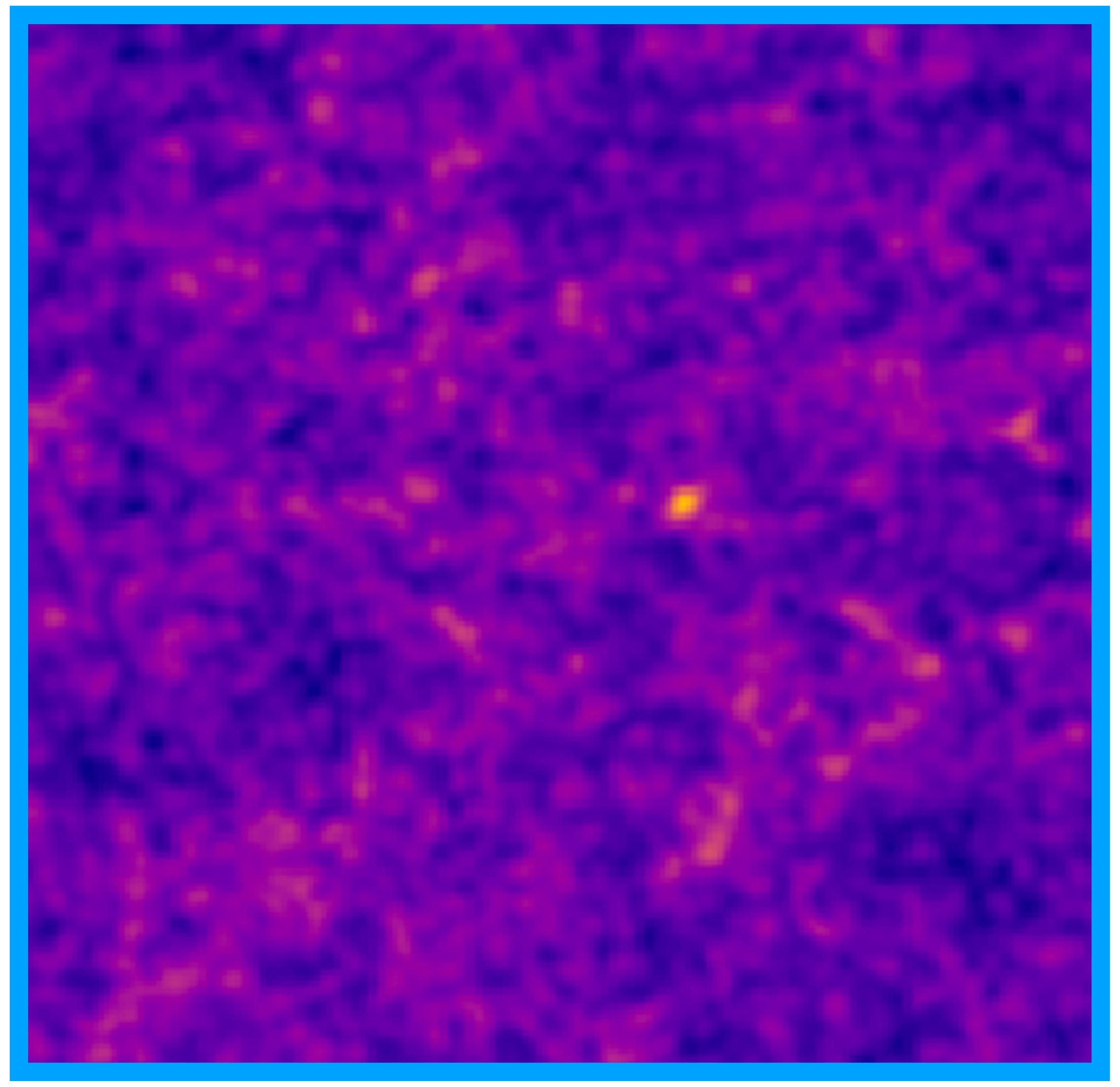
What is the advantage of deep learning for current and upcoming data?



add noise →

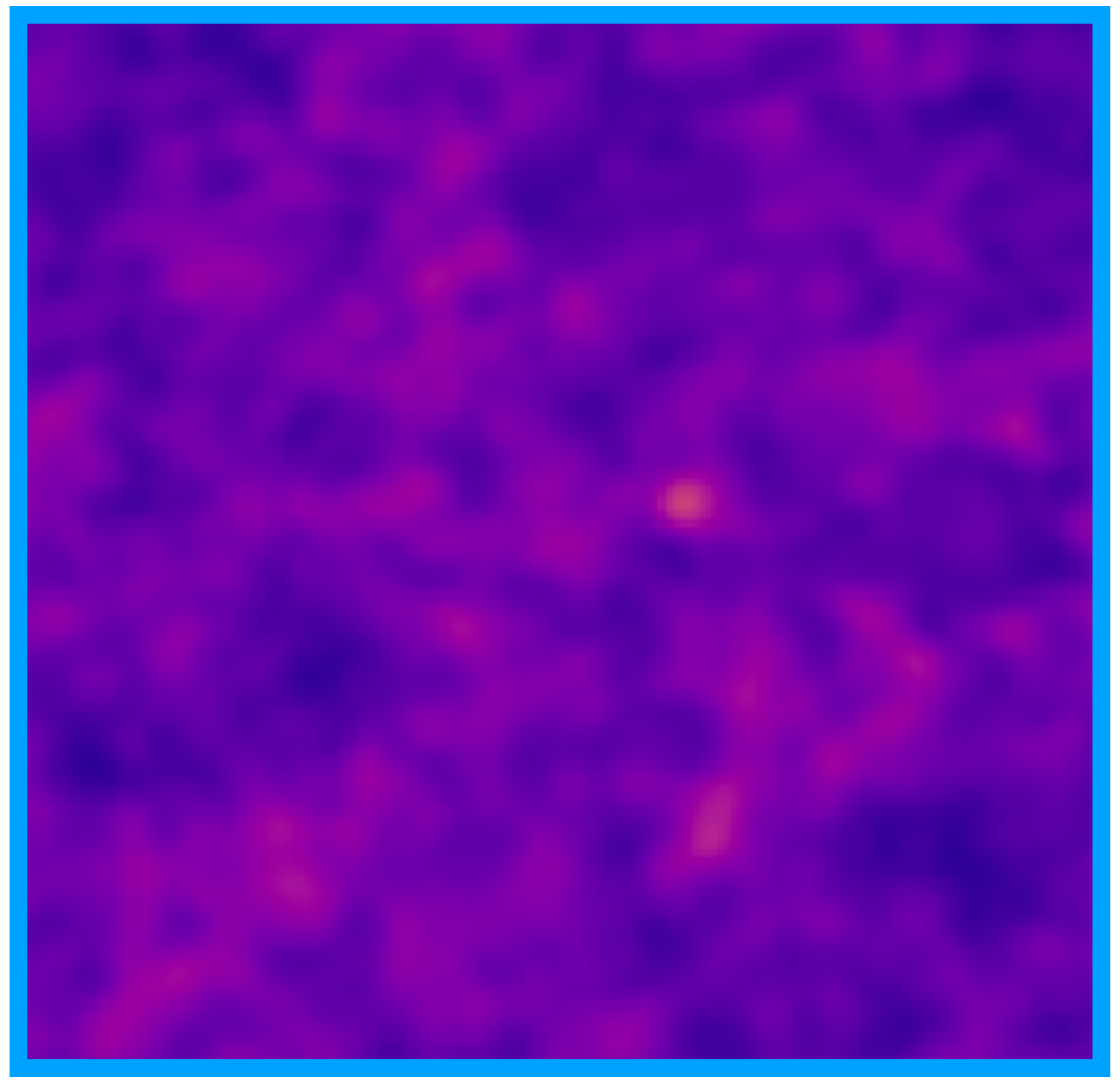
quality of
observations





add smoothing →

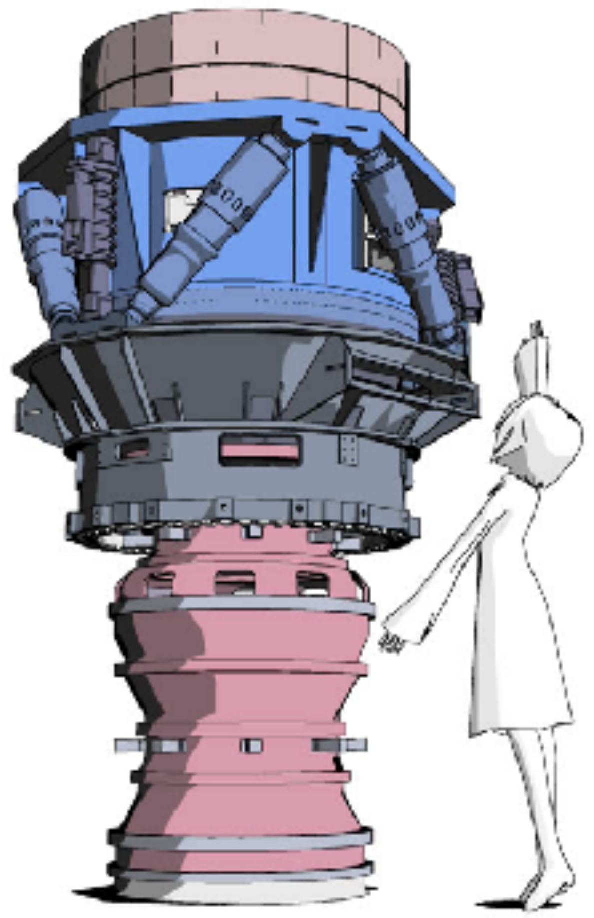
quality of
simulations



How much more information can we gain with deep learning for Stage-III and Stage-IV surveys?

DARK ENERGY SURVEY

Stage III

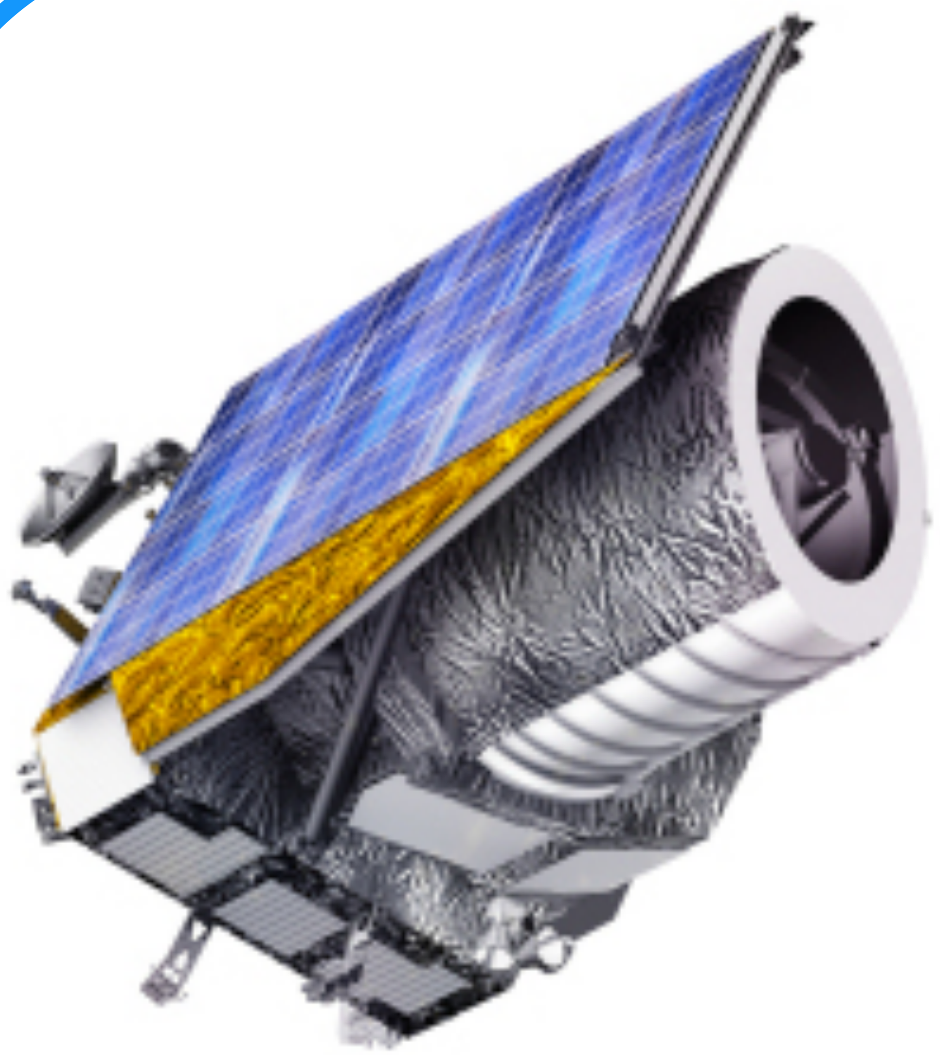


Hyper Suprime Cam

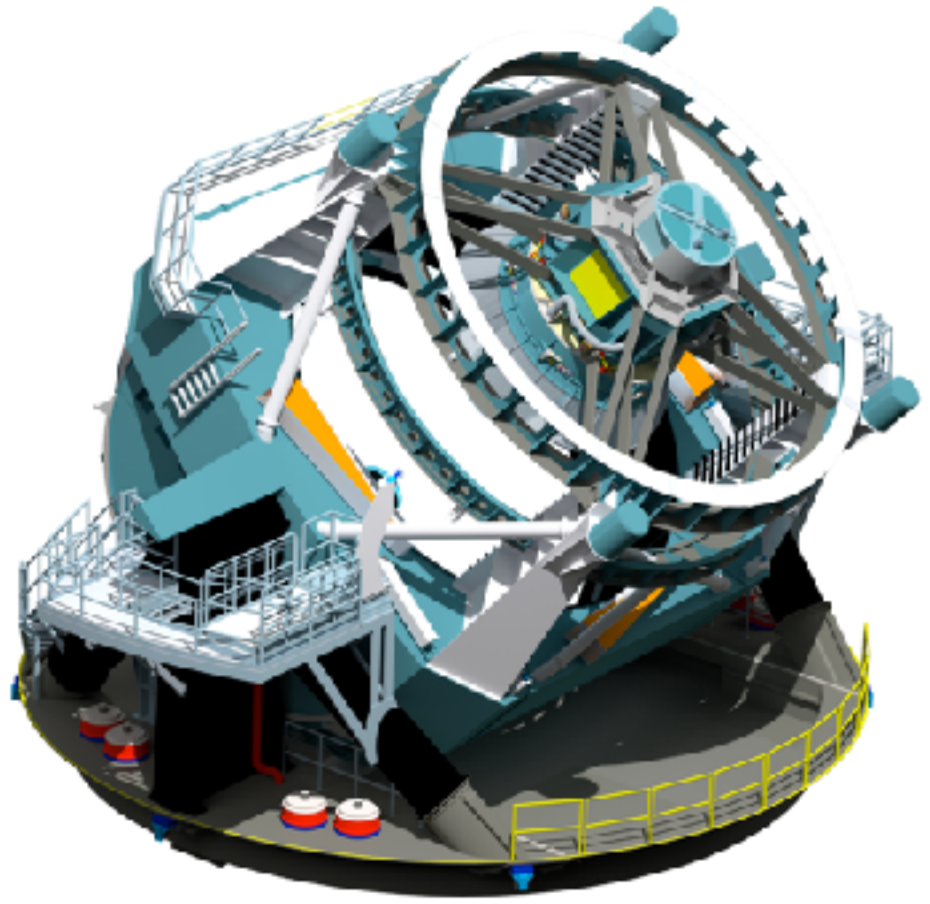
KiDS

Kilo Degree Survey

Stage IV



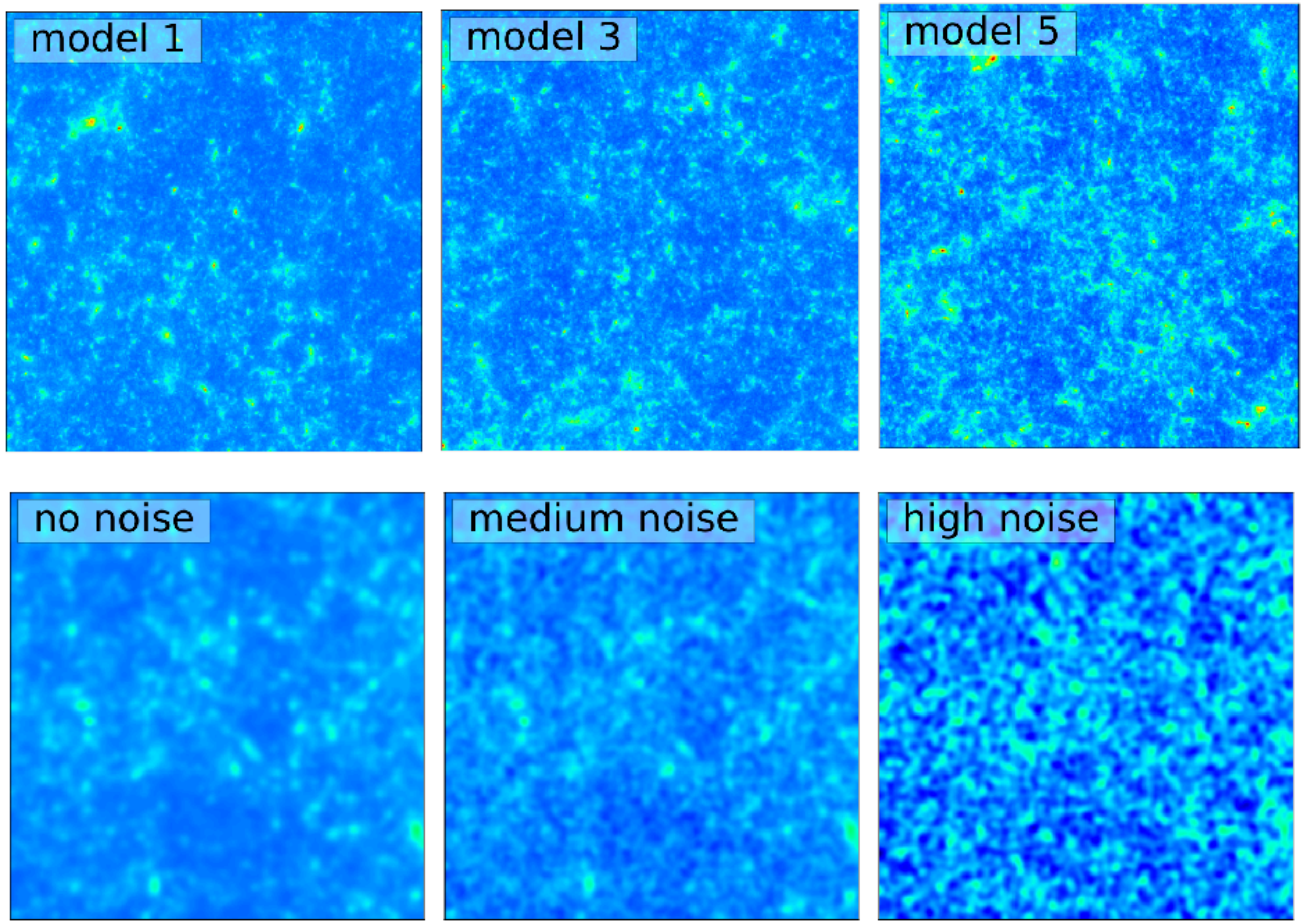
Euclid

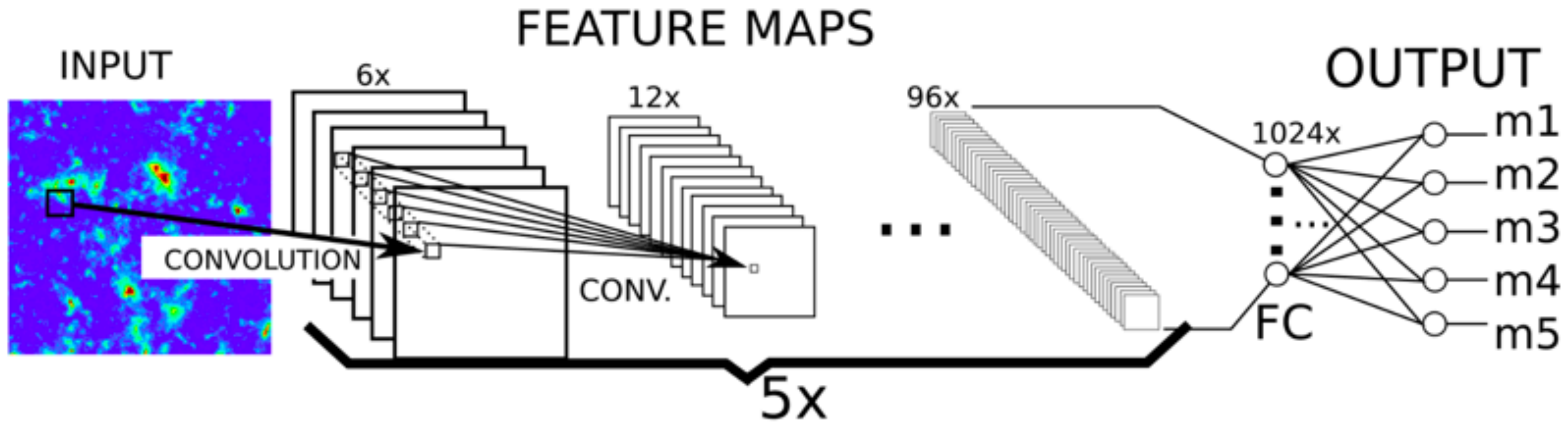


Rubin/LSST

First results for CNN vs 2-pt on lensing convergence: classification

- First application of CNNs to weak lensing maps for a classification problem
- Discriminating between five cosmologies in Ω_m, σ_8 that have the same power spectra
- Smoothing scale $\sigma = 0.9$ arcmin
- Realistic noise level for Stage-3 surveys
- Biggest challenge: make the network handle very noisy data
- Classification accuracy CNN 90%, compared to skewness and kurtosis (70%)

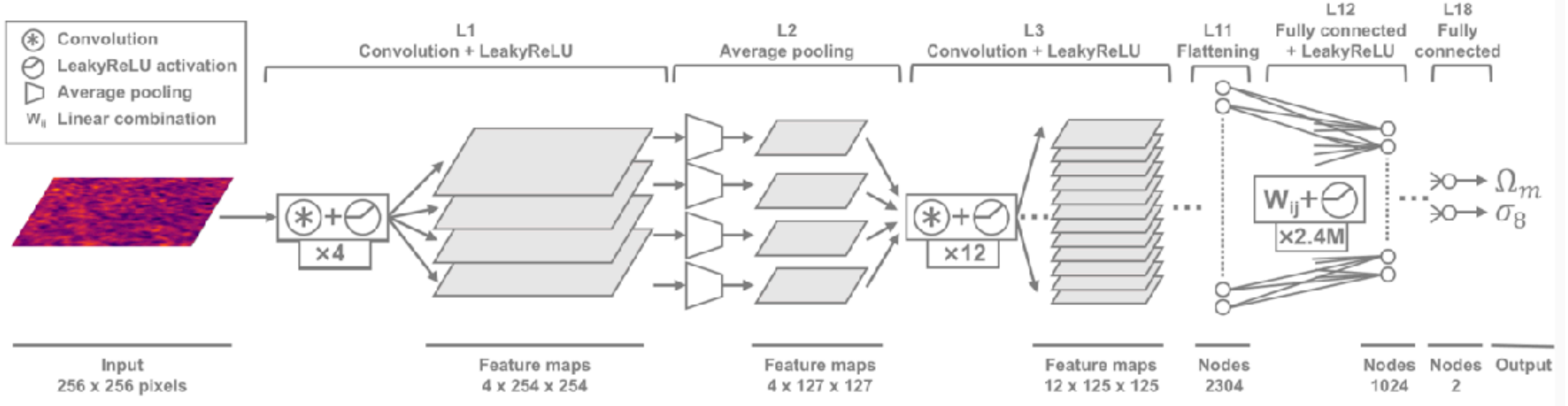


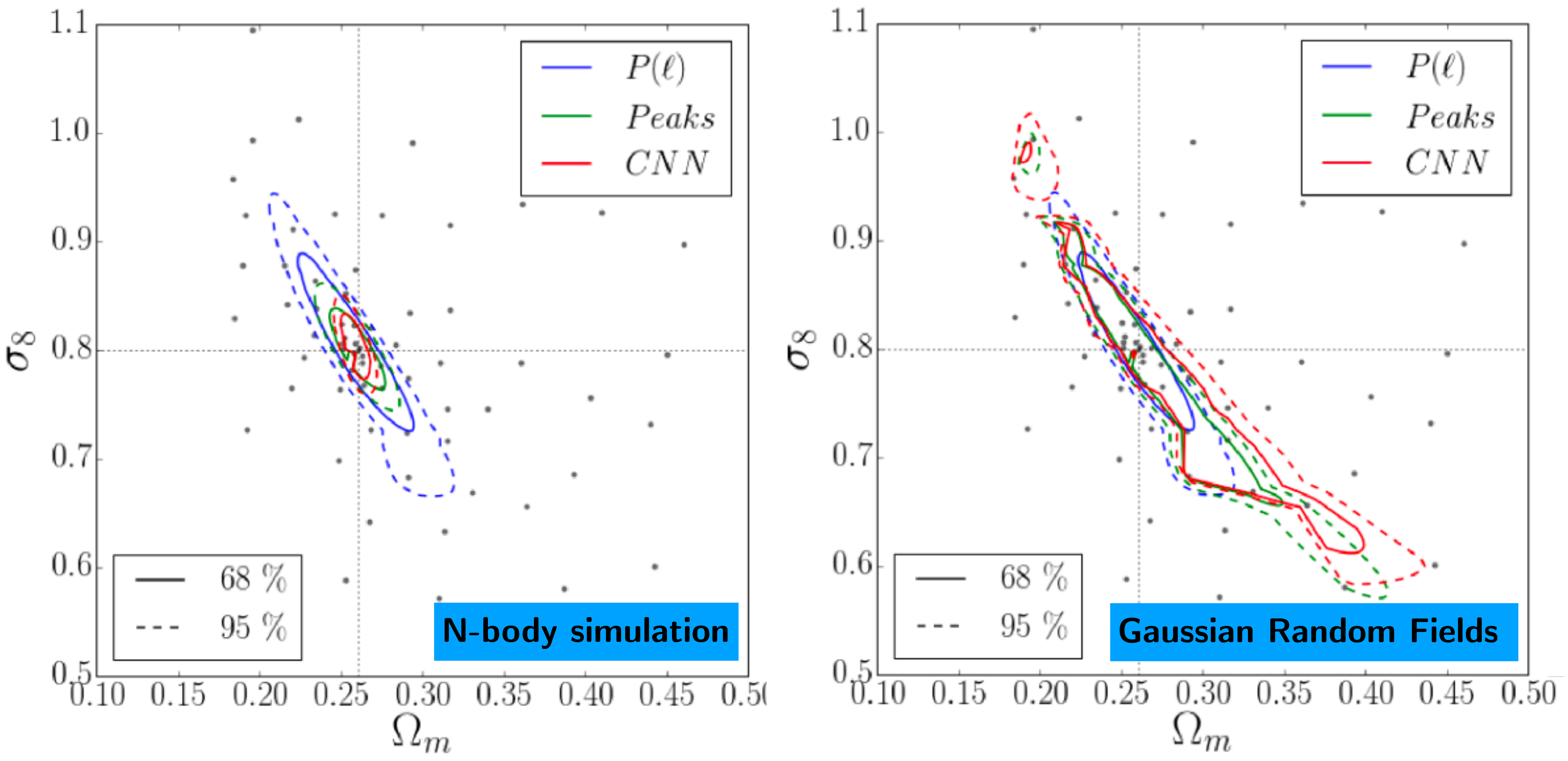


Schmelzle, +TK, et al. 1707.05167

First results for CNN vs 2-pt

- First comparison between CNN and 2-pt on the constraints level noise- free N-body sims
- Greatly improved precision by CNN vs 2-pt, also beating peak counts
- Same results for CNN as for 2-pt for Gaussian Random Fields → reassuring!

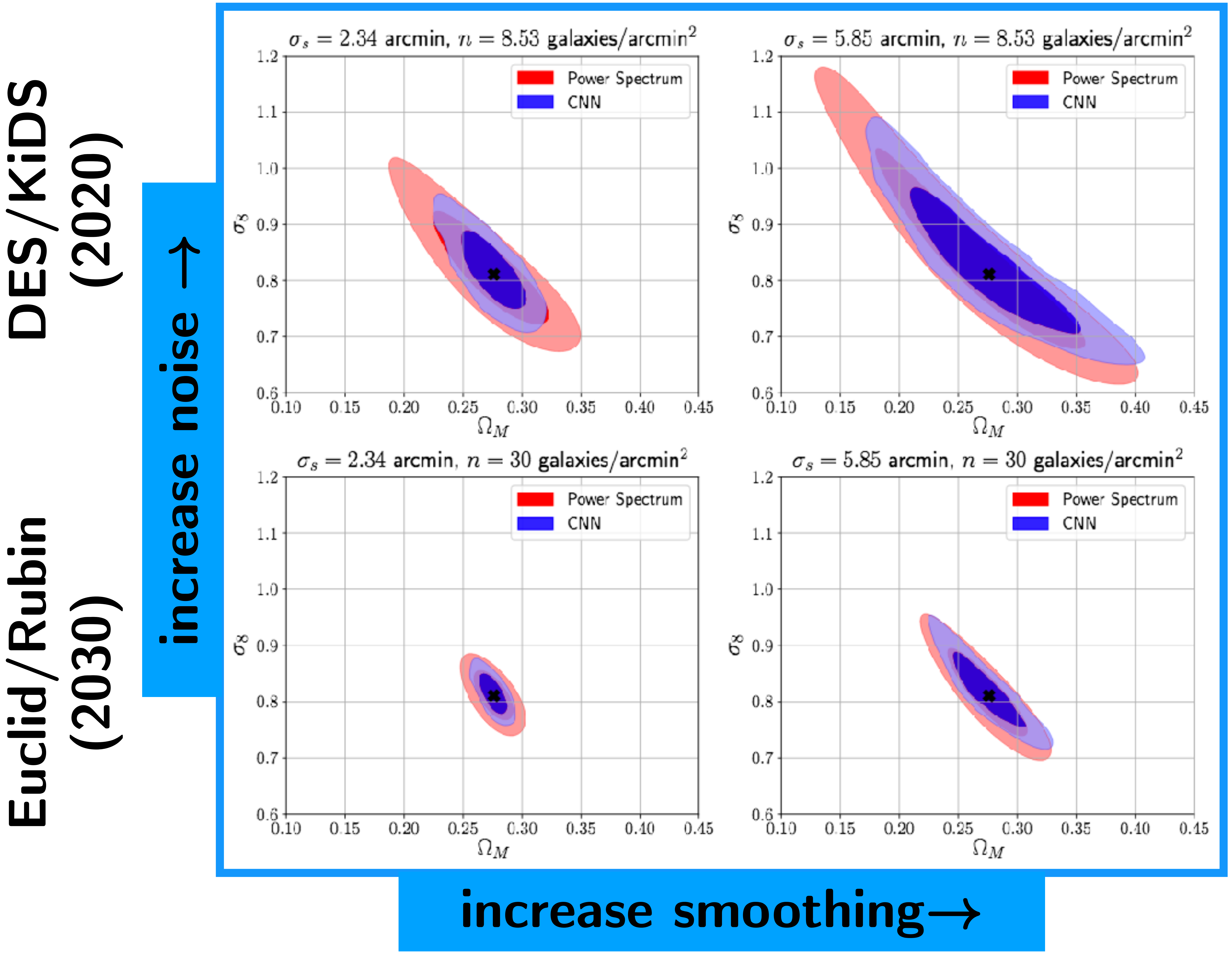




Gupta et al. 2018 1902.03663

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

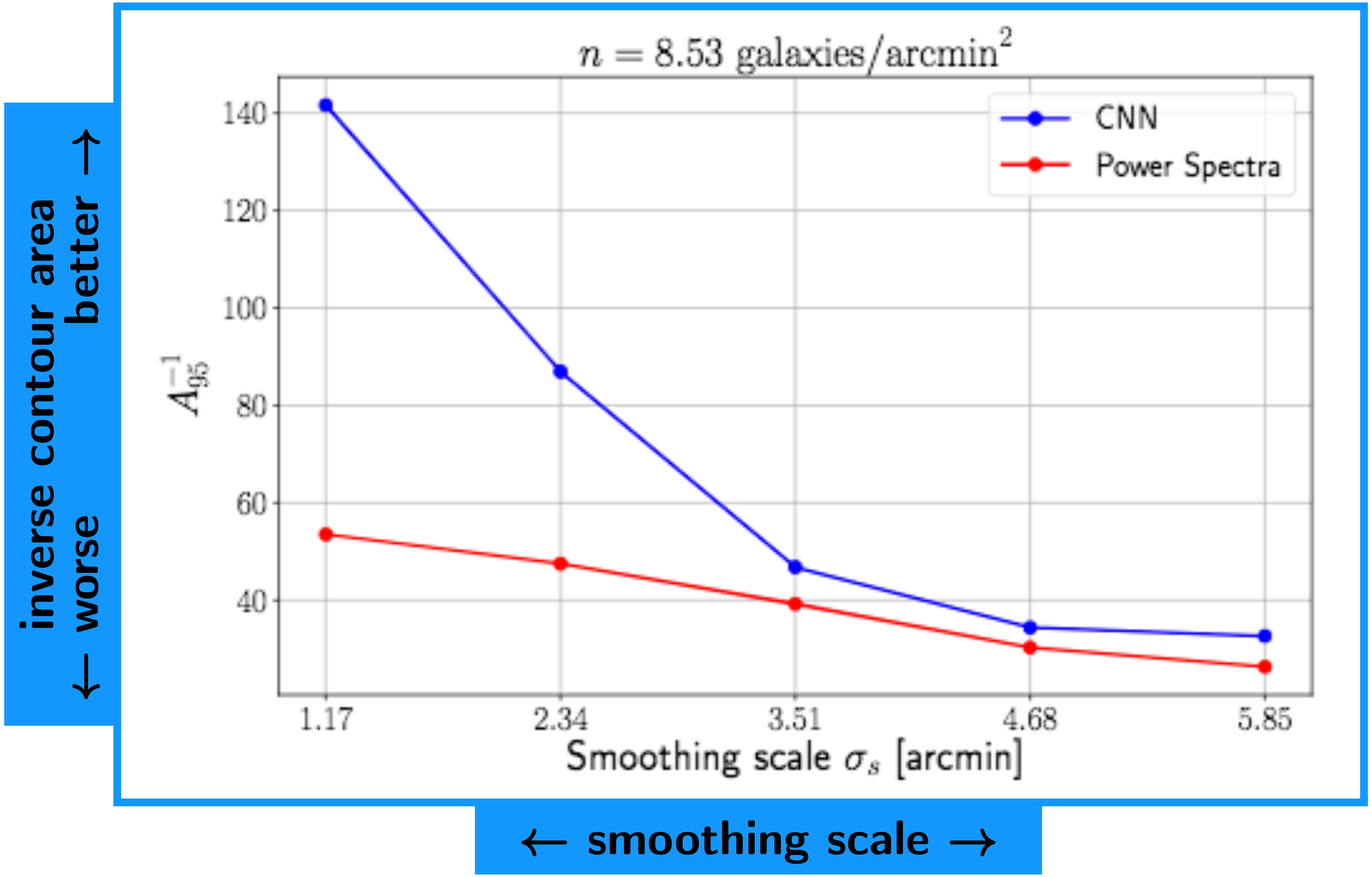
- Work led by Janis Fluri, interdisciplinary PhD (2022) with the Cosmology Group, the ETHZ Data Analytics Lab and Swiss Data Science Center
- The advantage of deep learning is preserved for high noise levels
- Advantage of deep learning starts at intermediate scales, around $\ell < 1000$
- This is the regime already affected by baryonic feedback
- The advantage increased greatly if small scales included



Fluri, TK, et al. 1807.08732

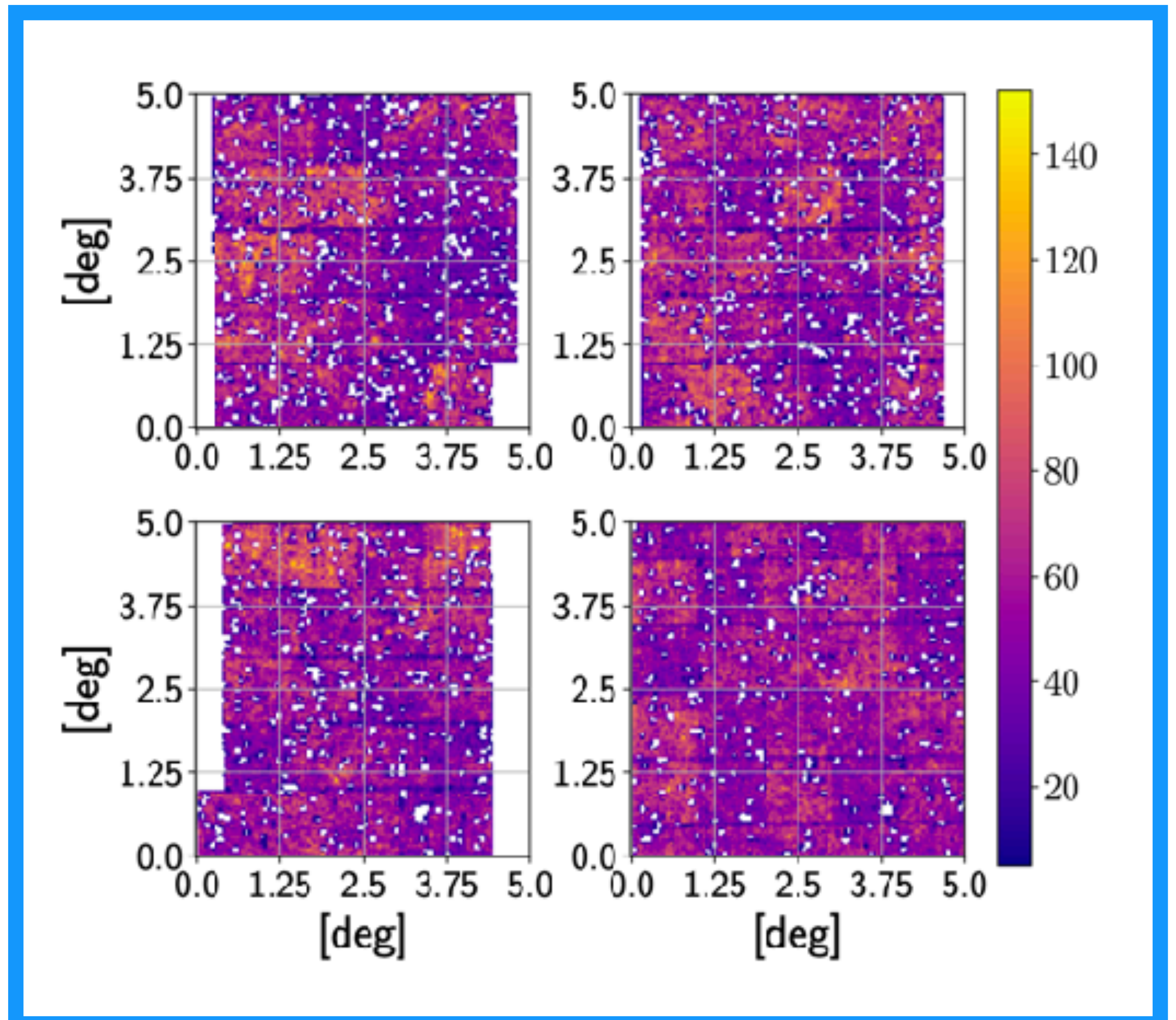
What is the advantage of deep learning for current and upcoming data?

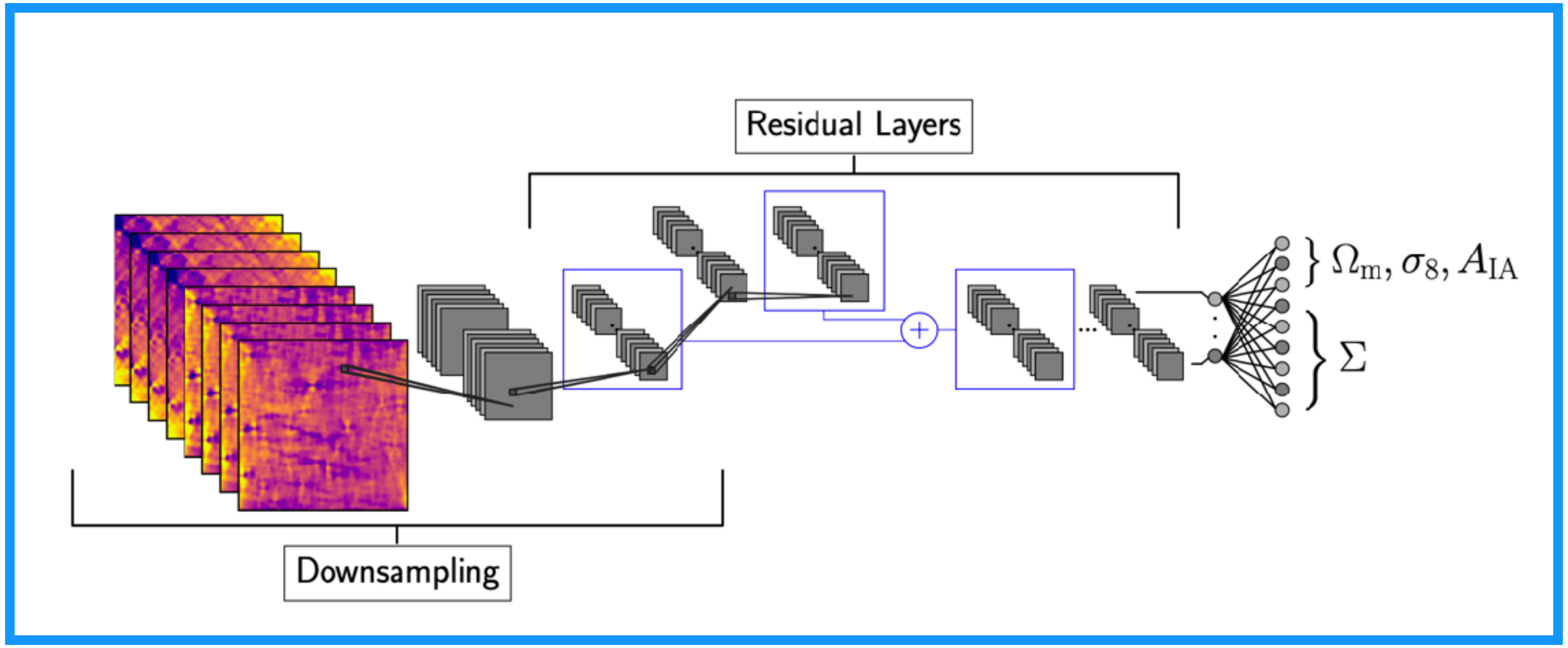
- The advantage of deep learning is preserved for high noise levels
- Advantage of deep learning starts at intermediate scales, around $\ell = 1000$
- This is the regime already affected by baryonic feedback
- The advantage increased greatly if small scales included



Fluri, TK, et al. 1807.08732

First CNN measurement on data: analysis of KiDS-450 with deep learning





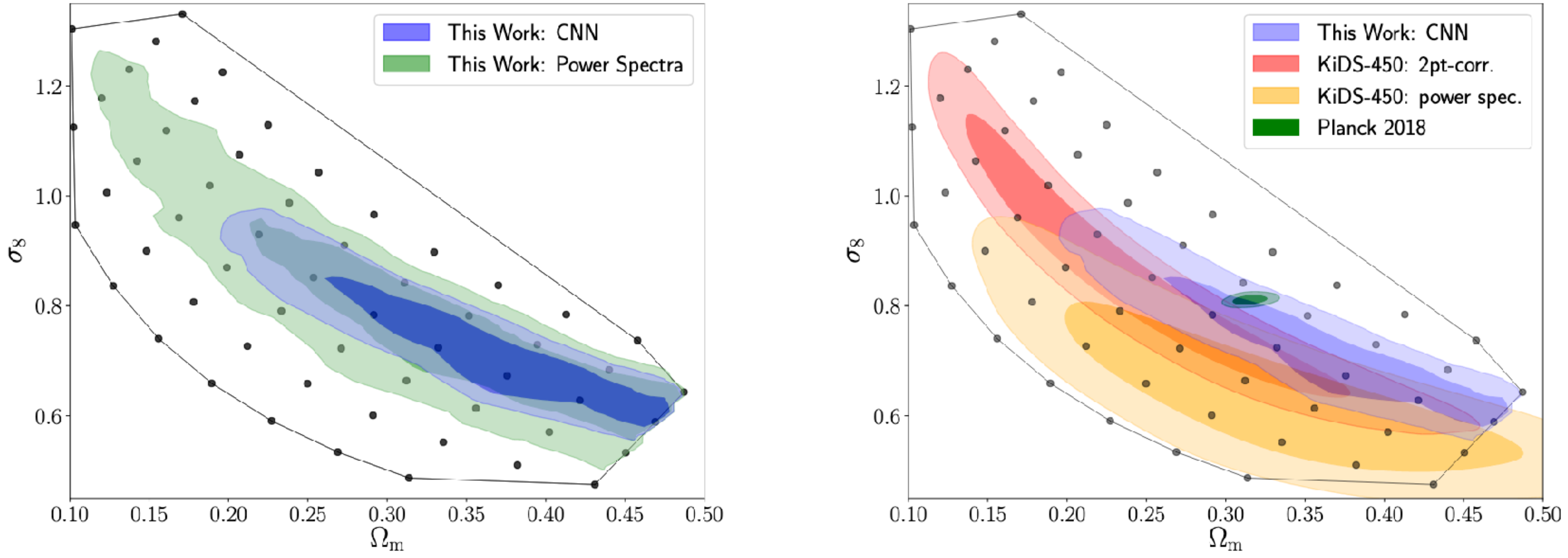
data:
20 x 4 tomographic shear maps

network: 3 parameter output

likelihood analysis

Fluri, TK, et al. 1906.03156

Analysis of KiDS-450 with deep learning



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

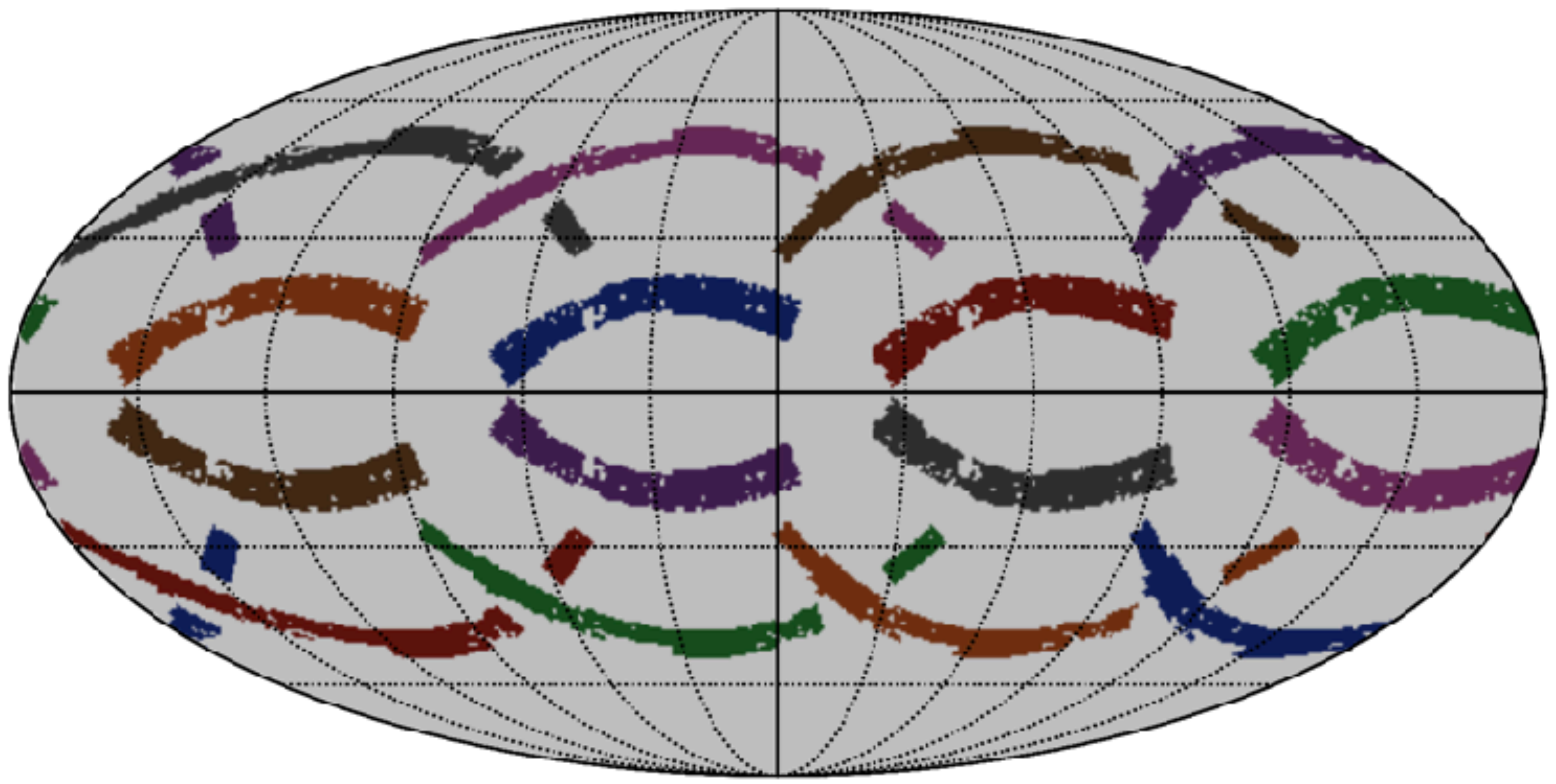
First results using machine learning inference in LSS cosmology

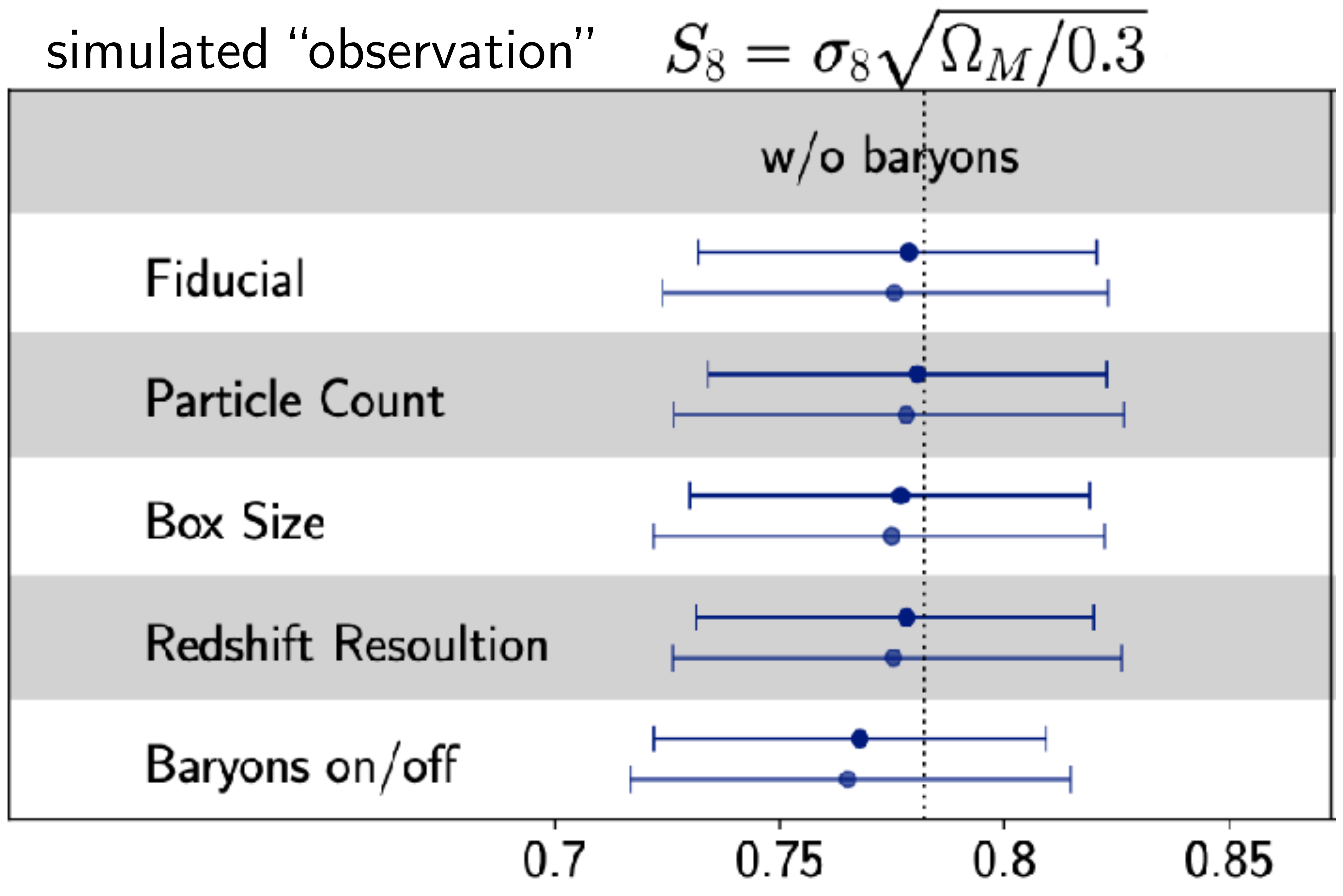
Blinded analysis

Fluri, TK, et al. 1906.03156

KiDS-1000 constraints with deep learning

- Demonstration of the scalability of the deep learning approach
- Full KiDS-1000 survey analysis of the 1000 deg^2
- Low-res analysis at nside=512 due to processing power limitations, pixel size 7 arcmin
- Using full CosmoGridV1 simulation volume
- Constrained: Ω_m, σ_8, w_0, A_{IA}
- Marginalized: H_0, Ω_b, n_s, +baryons M_c, ν, +sys
- Improved results compared to power spectra $\sim$25%
- Blinded analysis with results consistent with main KiDS results

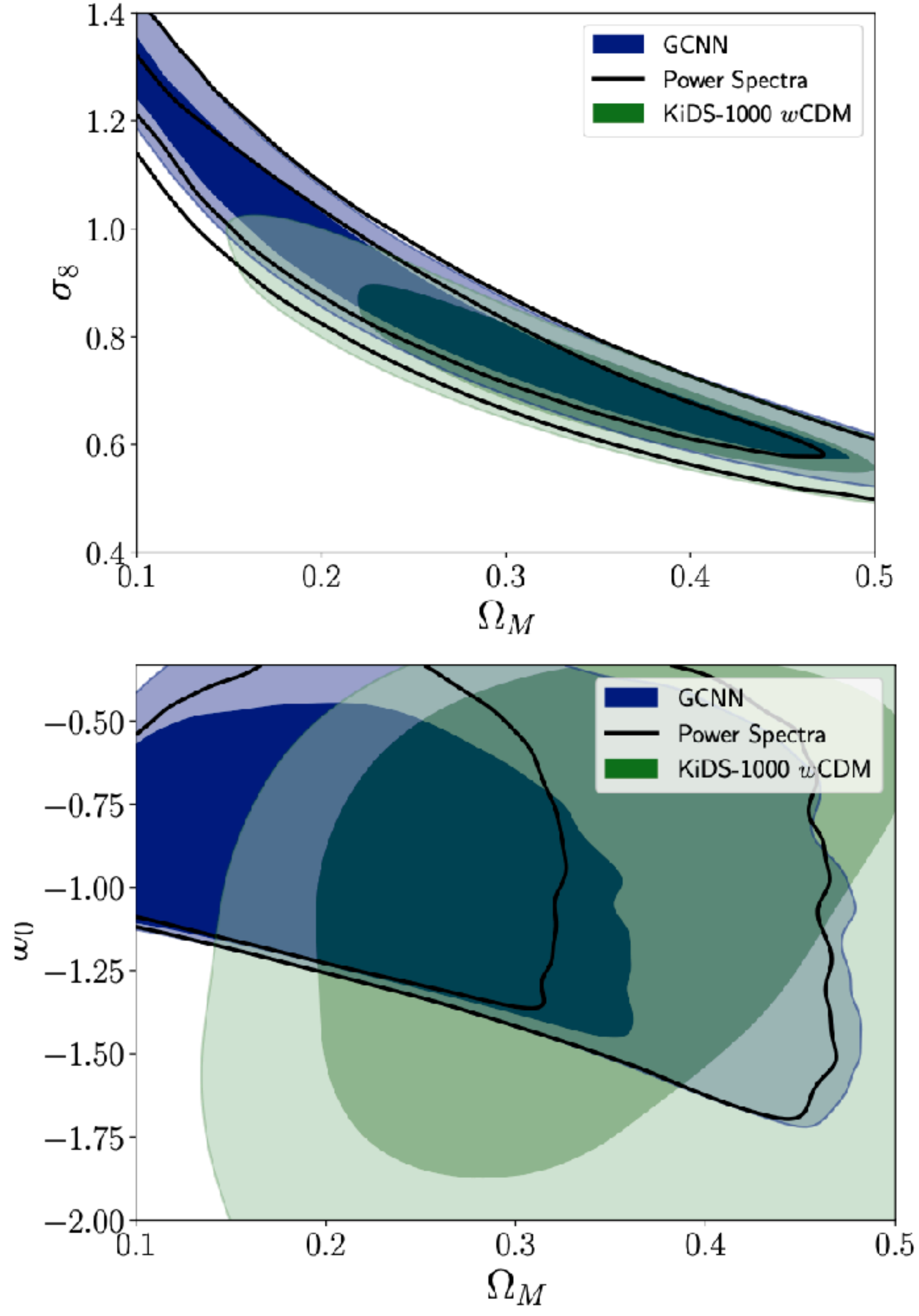




Fluri, TK, et al. 2022, 2201.07771

KiDS-1000 constraints with deep learning

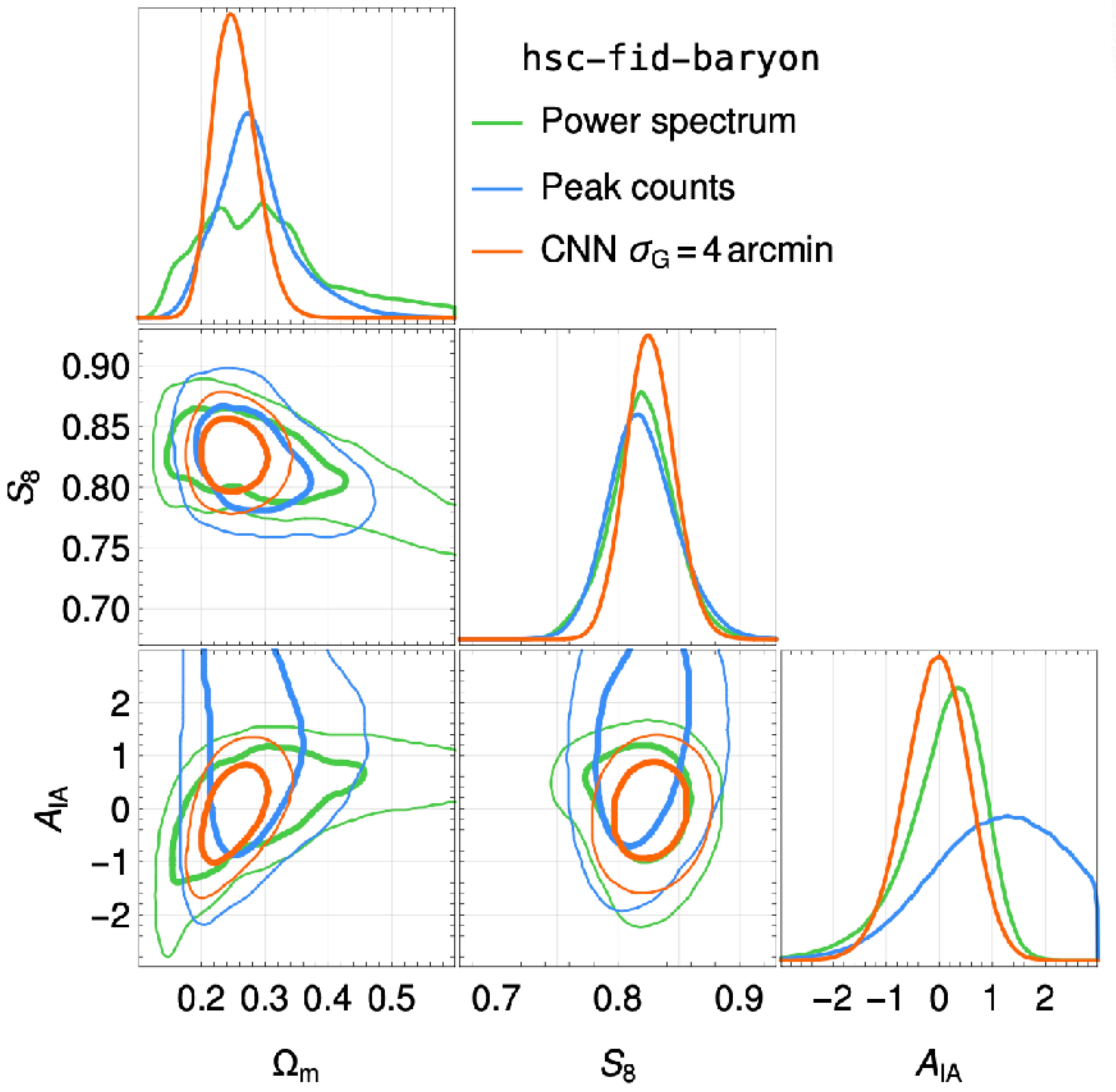
- Demonstration of the scalability of the deep learning approach
- Full KiDS-1000 survey analysis of the 1000 deg^2
- Low-res analysis at nside=512 due to processing power limitations, pixel size 7 arcmin
- Using full CosmoGridV1 simulation volume
- Constrained: Ω_m, σ_8, w_0, A_{IA}
- Marginalized: H_0, Ω_b, n_s, +baryons M_c, ν, +sys
- Improved results compared to power spectra $\sim$25%
- Blinded analysis with results consistent with main KiDS results



Fluri, TK, et al. 2022, 2201.07771

Cosmology from HSC Y1 with deep learning

- 171 deg[2] lensing data from HSC
- Ω_m, σ_8, A_{IA}, +baryons: M_c, $M_{1,0}$, η, β
- Marginalized systematics
- No marginalization of other cosmological parameters in wCDM
- for the first time: forward-modelling of PSF leakage residuals on the map level
- CNNs deliver improved constraints: 5–24% for S_8 and a factor of 2.5–3.0 smaller for Ω_m
- Greatly improved constraint on A_{IA}, although IA fixed with redshift
- No blinding

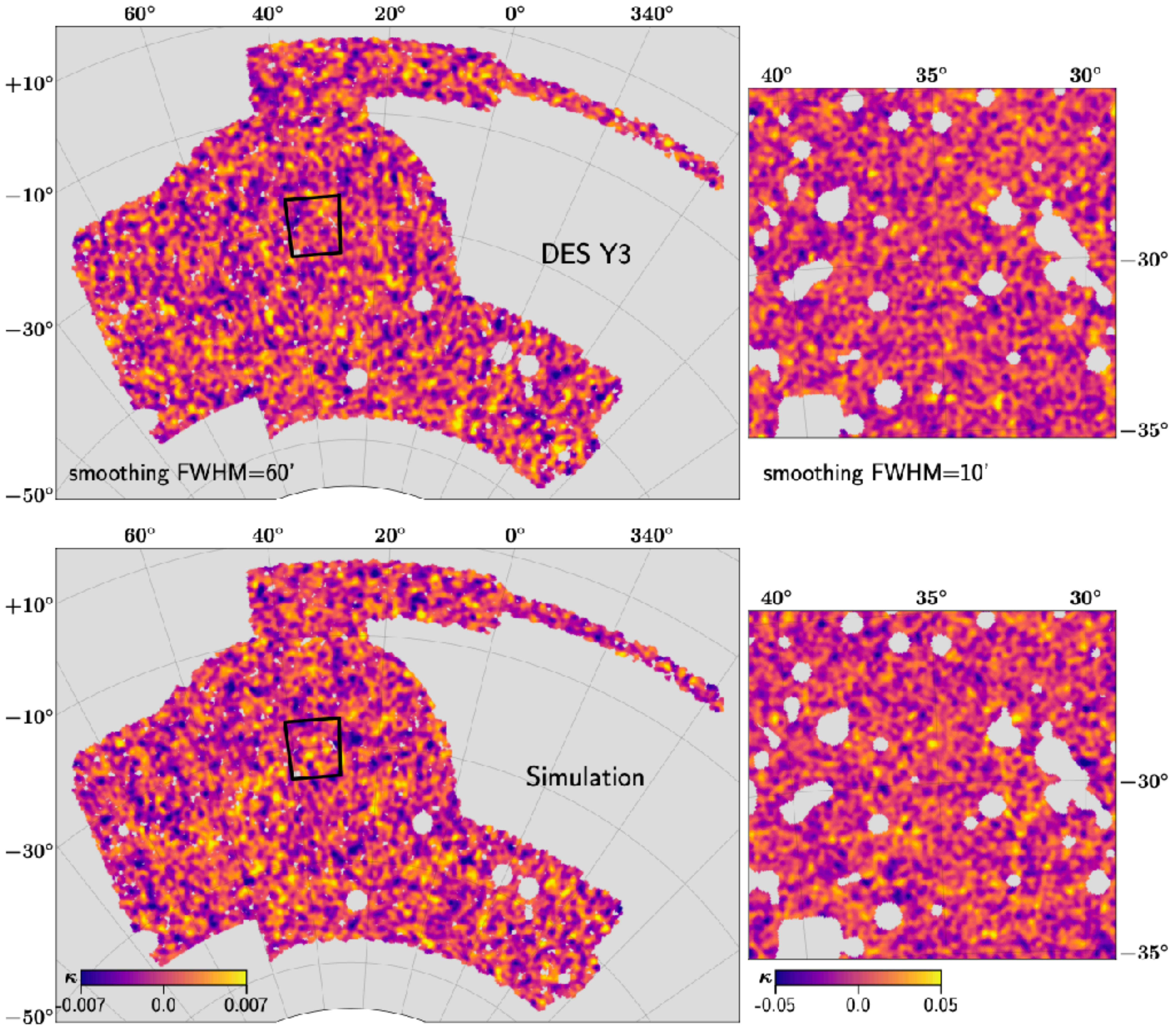


Lu, Haiman, Li 2023 2301.01354

Human vs machine: peaks statistics for DES Y3

Image credit: Samantha Bond (SKIM Group)

Human intuition statistics: peaks for DES-Y3



DES Y3 data

- 5000 deg^2
- Up to redshift z=1.5
- ≈6 galaxies/arcmin^2

Simulations

Constrain:

- σ_8 : clustering strength
- Ω_m : matter density
- A_{IA} : intrinsic alignment

Marginalize:

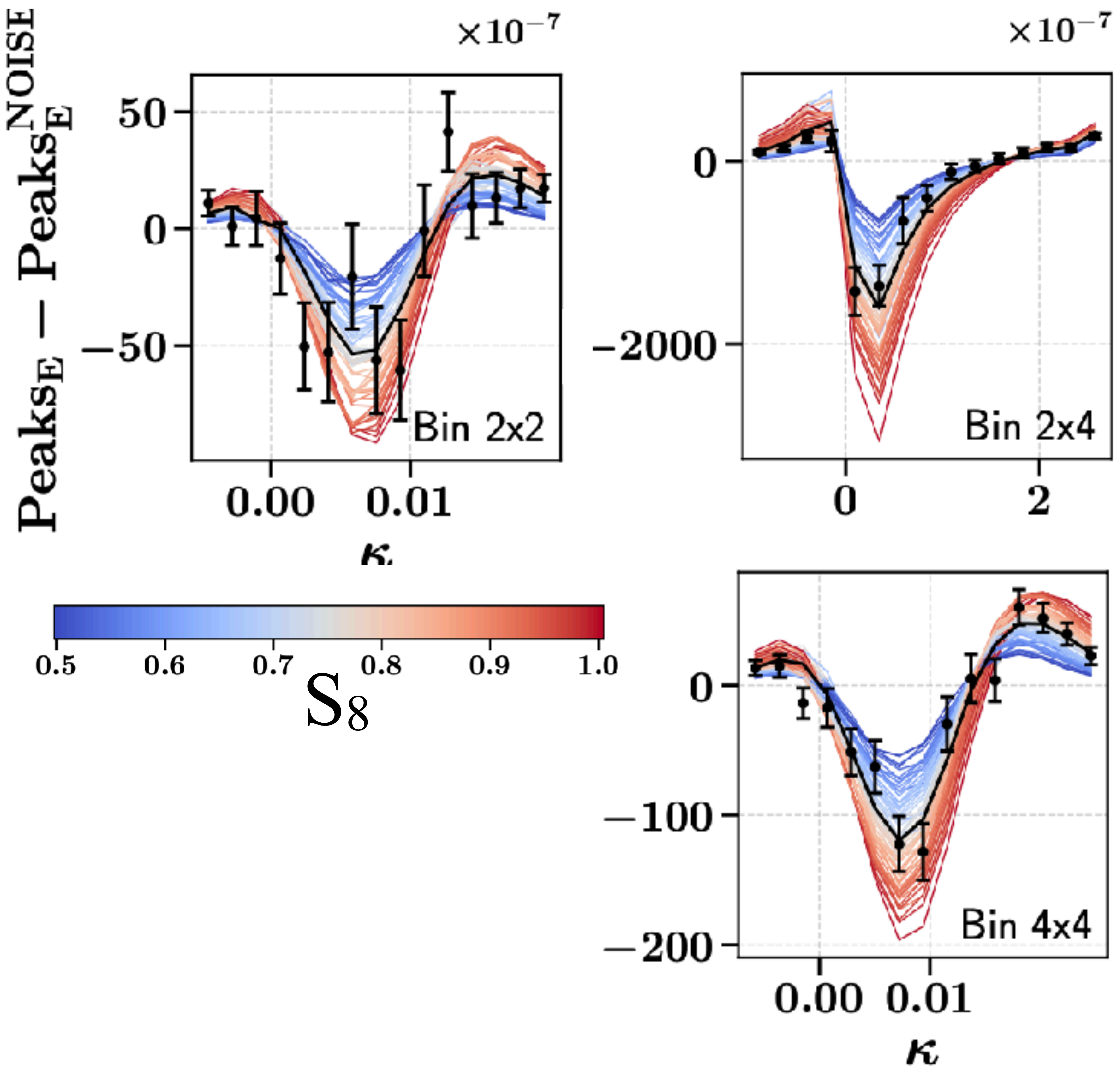
➡ n_s, Ω_b, h

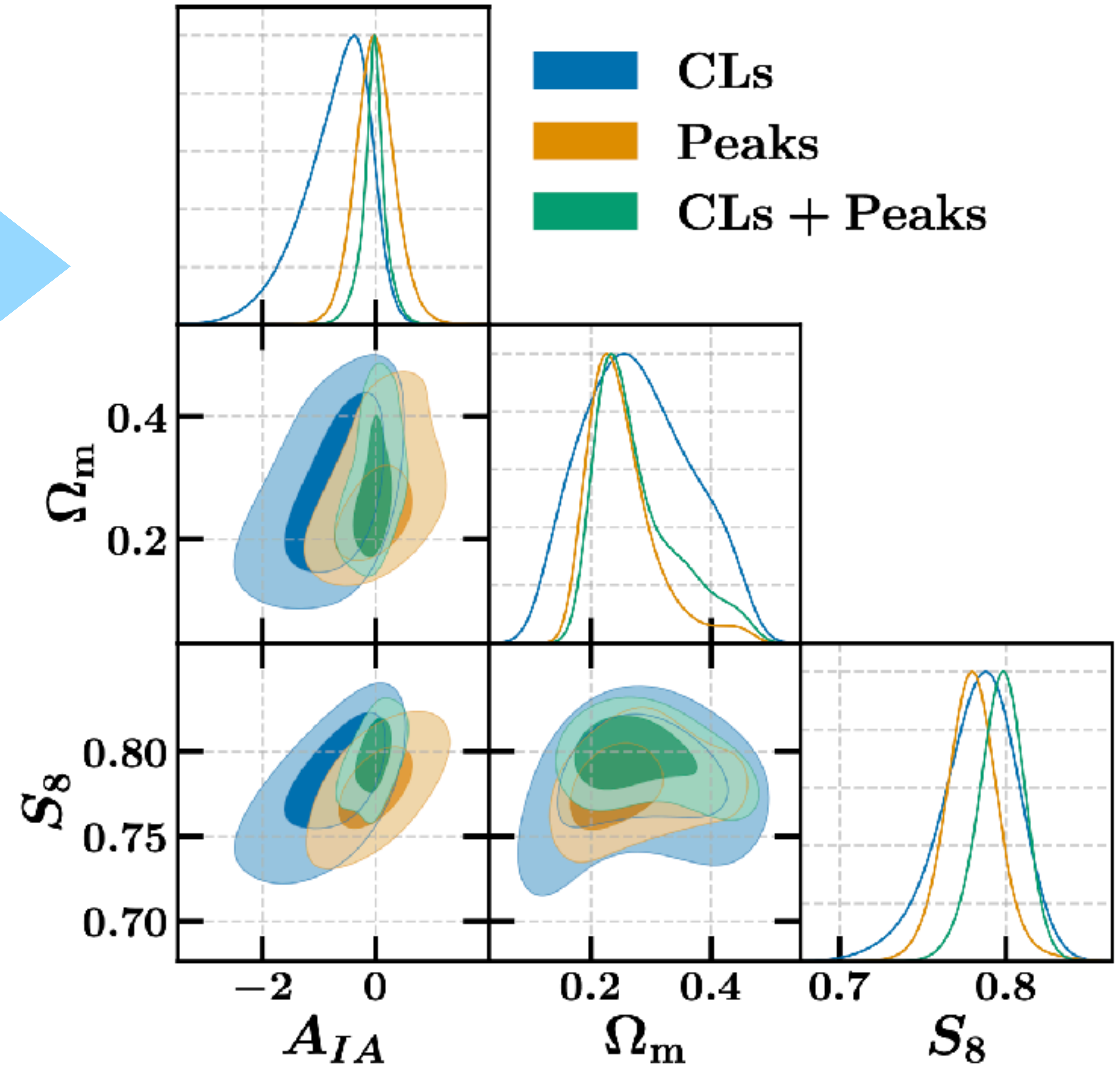
➡ n(z) error

➡ Shear calibration error

Zuercher, +TK, +DES Collaboration 2110.10135

Human intuition statistics: peaks for DES-Y3





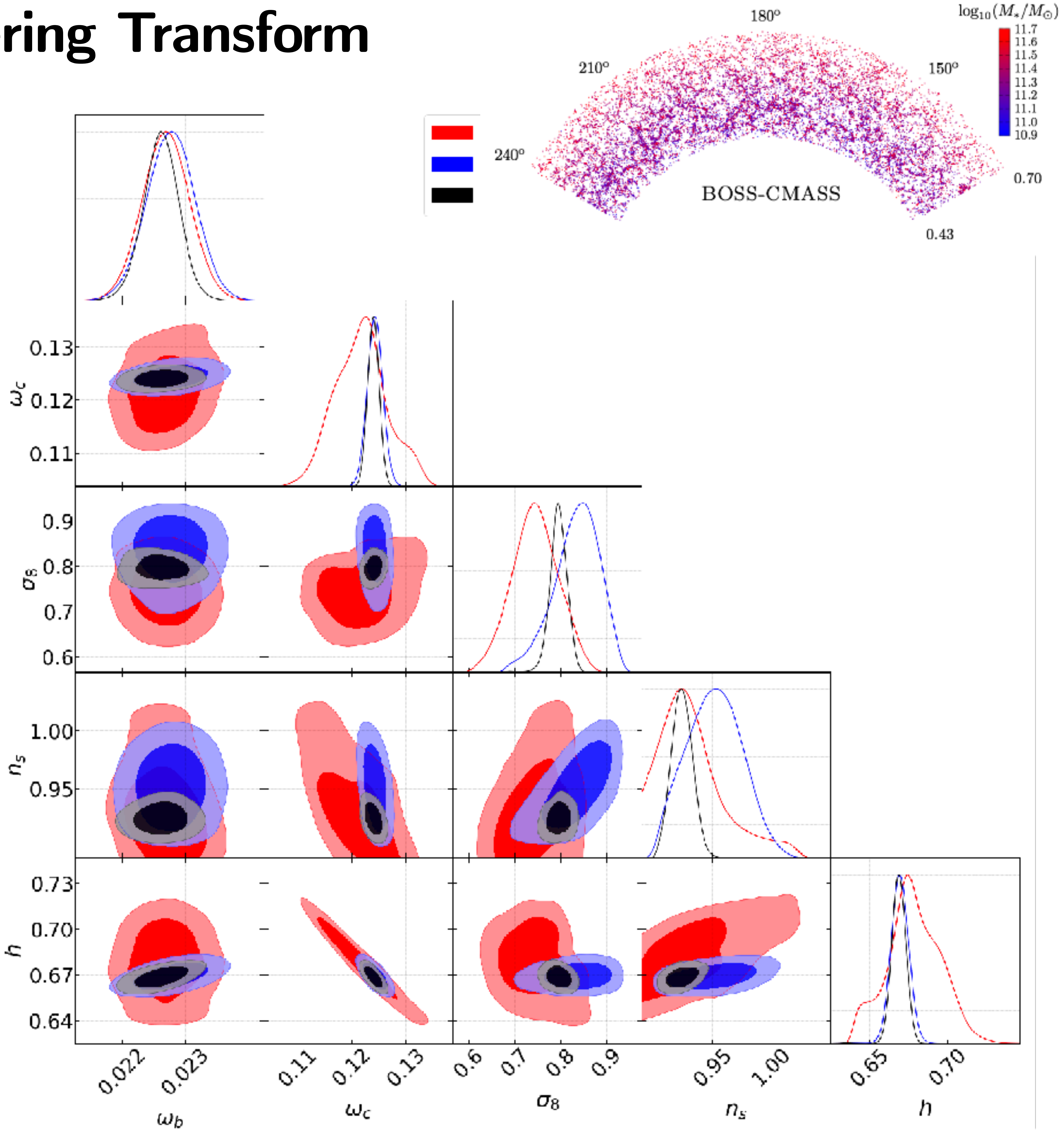
- Tomographic peaks measurement
- Conservative scales, smoothing $\sigma > 8$ arcmin
- Using a Peaks + Cl emulator
- 40% improvement for combined analysis
- Blinded analysis

$$S_8 = 0.797^{+0.015}_{-0.013}$$
$$A_{IA} = -0.03 \pm 0.23$$

Zuercher, +TK, +DES Collaboration 2110.10135

BOSS galaxy clustering with Wavelet Scattering Transform (24 Oct 2023)

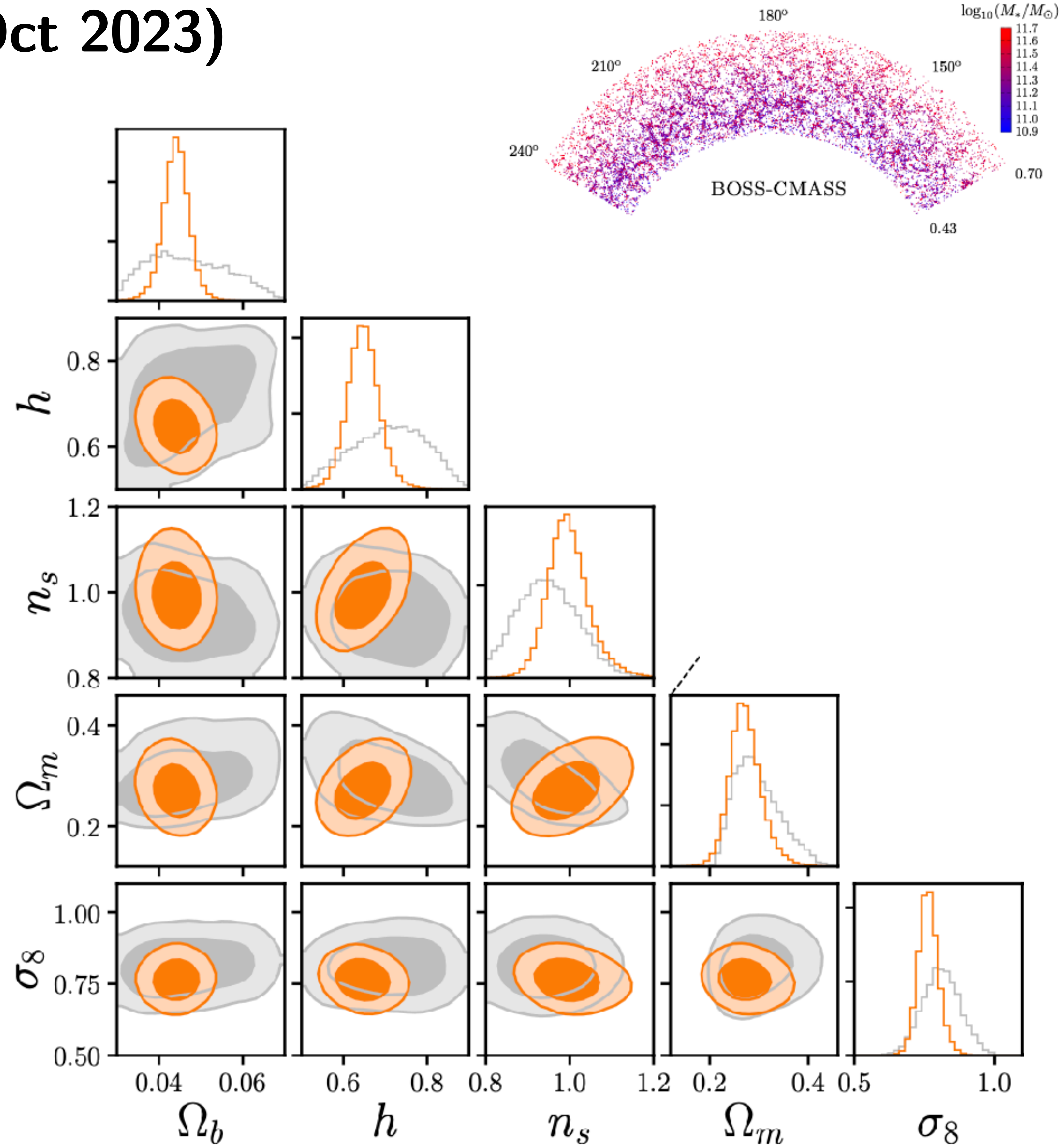
- Emulator created using ABACUS-SUMMIT simulations
- Using BOSS spectroscopic LRG galaxy clustering $z \in [0.46, 0.57]$ (CMAS sample)
- 4 cosmology parameters: ω_c, σ_8, n_s, h
- Smallest smoothing scale: 8 Mpc/h
- 6 halo occupation distribution parameters:
  - M_{cut} minimum halo mass to host a central galaxy
  - M_1 typical halo mass that hosts one satellite galaxy
  - σ steepness of the error function upturn in the
  - α is the power-law index on the number of satellite galaxies
  - κM_{cut} minimum mass of a halo that can host a satellite
  - $\bar{n}_{cent}^{LRG}$ modulation to satellite occupation function to disfavor satellites from halos without centrals
- Wavelet Scattering Transform + correlation gives improvement $2.5 - 6\times$ compared to correlations only
- No blinding / under review



Valogiannis, Yuan, Dvorkin 2023 2310.1611

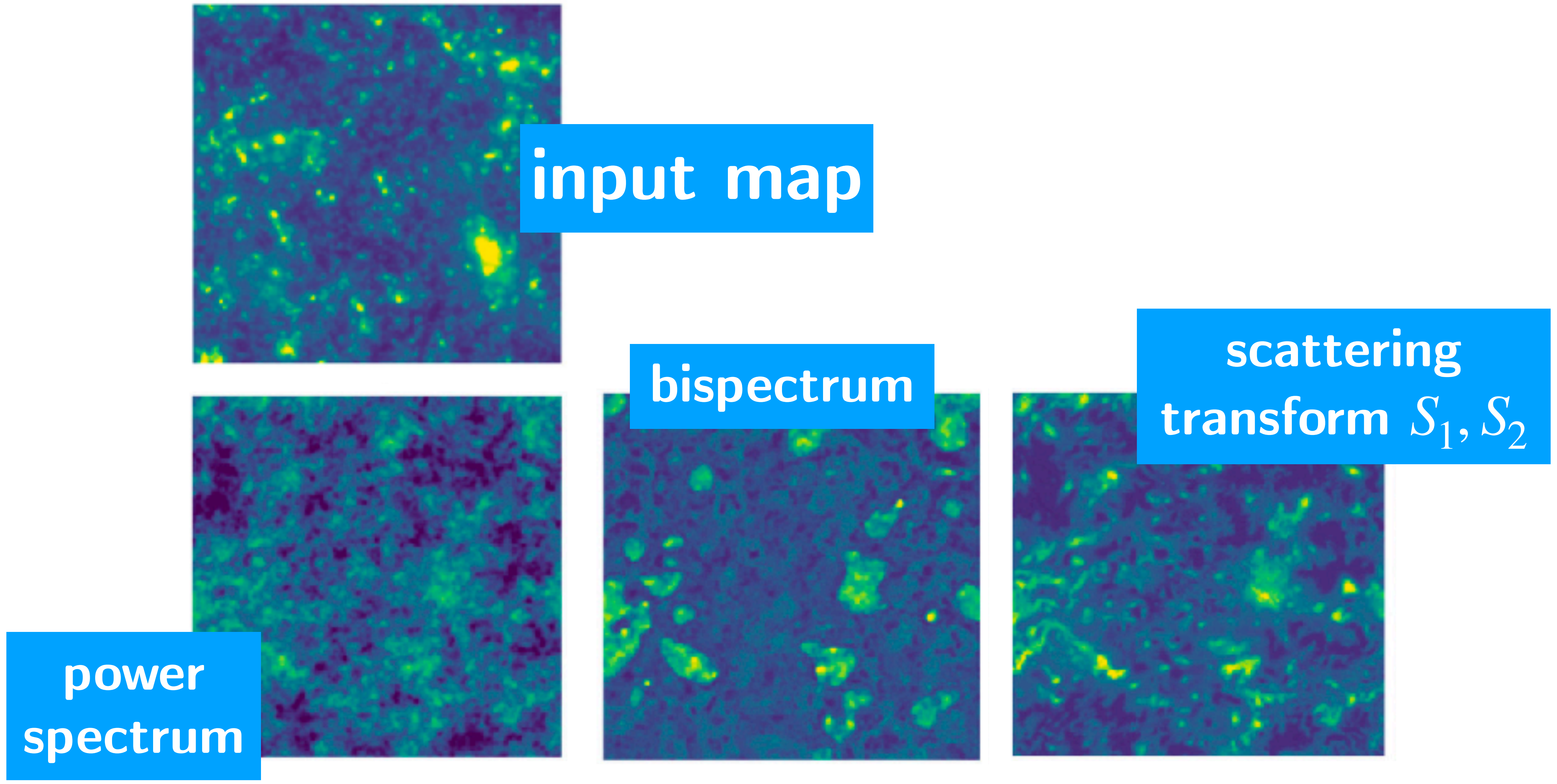
BOSS galaxy clustering with CNNs (23 Oct 2023)

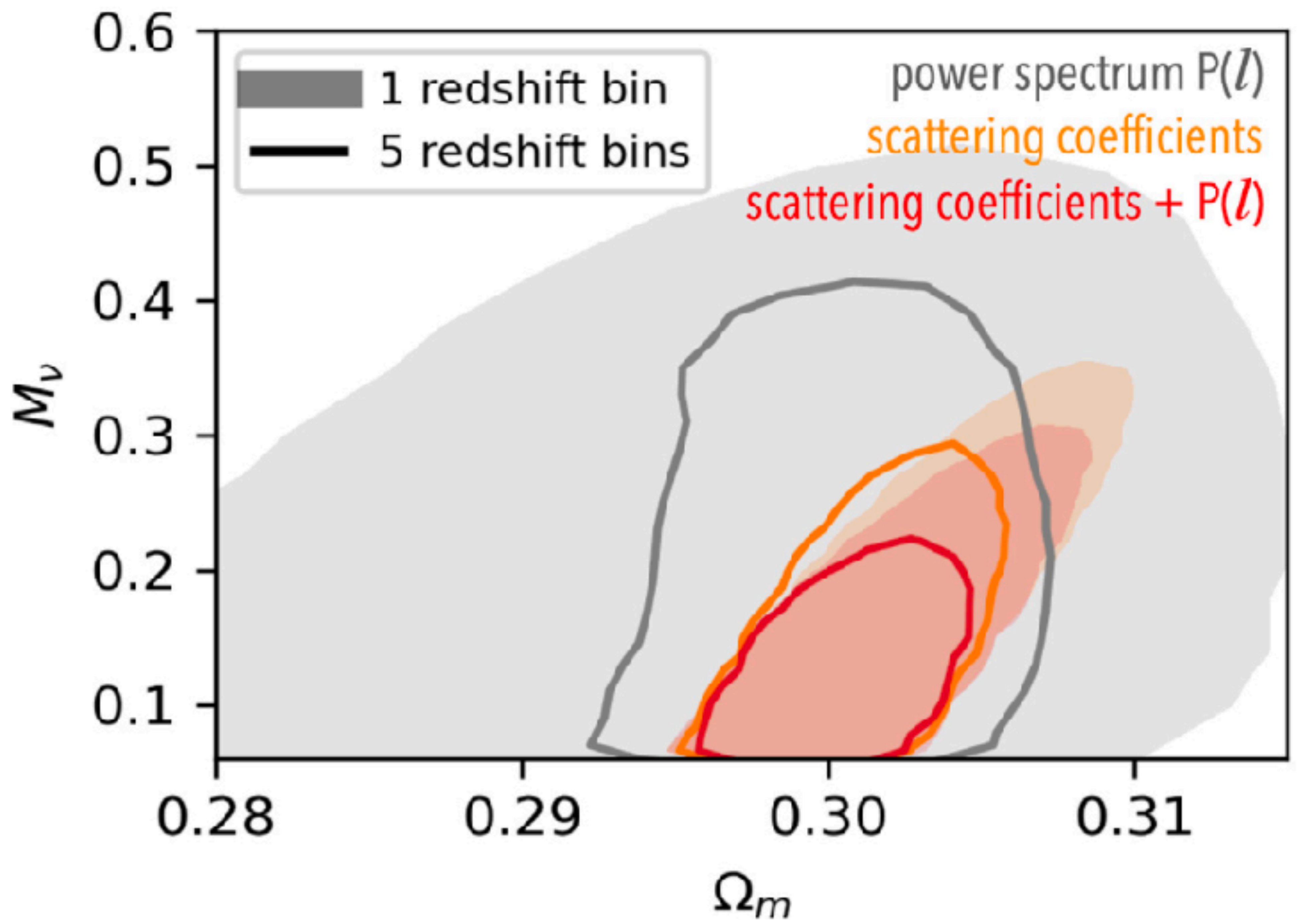
- Using the Qiuchote Simulations suite
- Using BOSS spectroscopic LRG galaxy clustering $z \in [0.46, 0.57]$ (CMAS sample)
- 5 cosmology parameters: Ω_m, σ_8, n_s, h, Ω_b
- Marginalizing over 9 HOD parameters
- Using the SimBIG forward modelling framework
- Pixelizing the clustering data into voxel box with $64 \times 128 \times 128$ voxels
- Smoothing scale: voxel size of $11\ Mpc/h$
- Training CNNs with $3 \times 3 \times 3$ learnable kernels
- CNNs give improvement $2.65 \times$ compared to correlations only
- Gives low $H_0 = 64.5 \pm 3.8\ km/s/Mpc$ values → interesting for H_0 tension?
- No blinding / under review



SBI for neutrino constraints

- Forecast for Rubin/LSST full data
- High resolution: pixel size 0.4 arcmin
- Improvement in M_ν constraint by $\times 2$, but depend on the noise level
- Using the MassiveNus simulation suite
- Potential to detect neutrino mass from Stage-3 surveys?

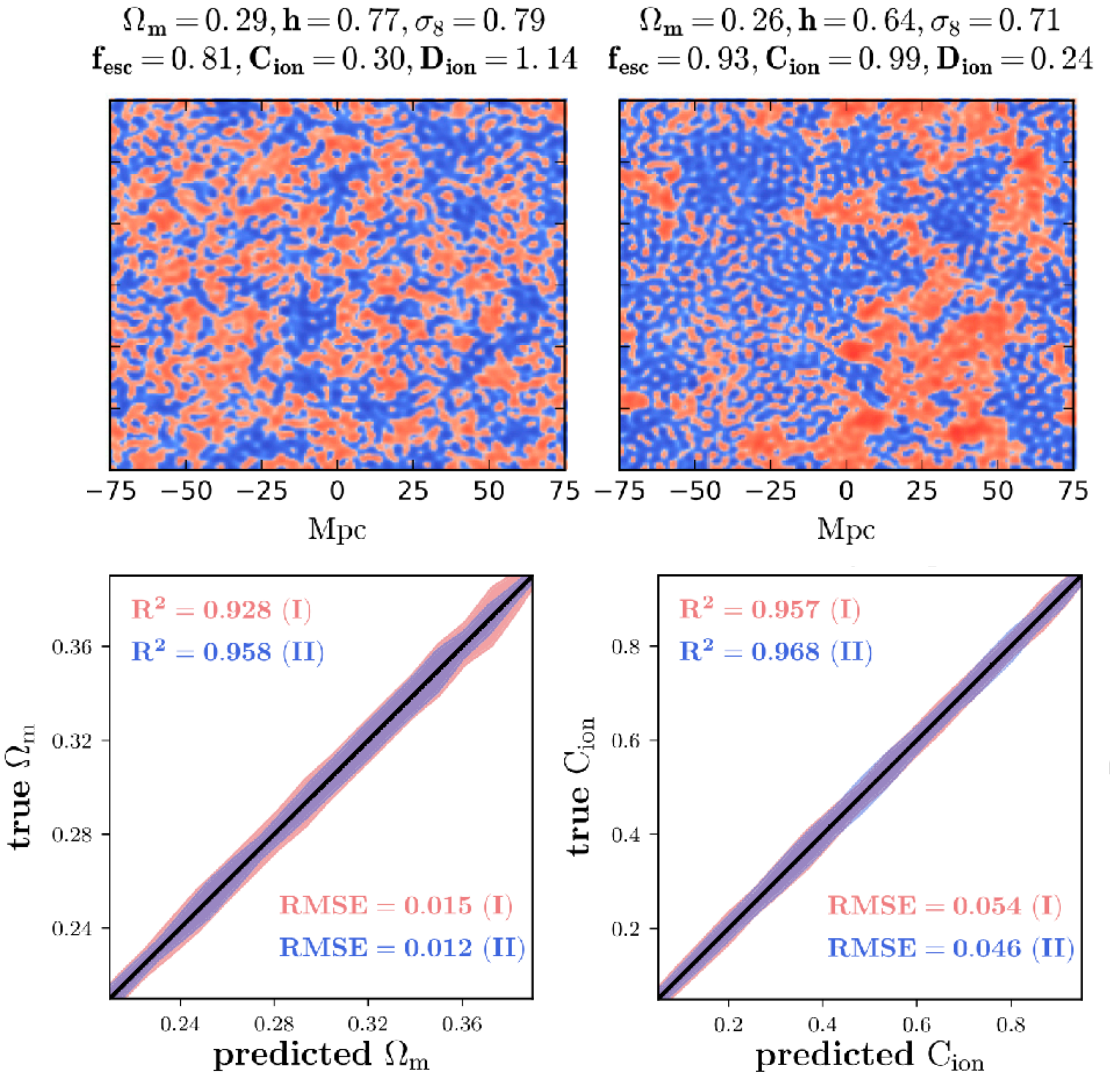




Cheng, Ménard 2021 2103.09247

AI cosmology with 21cm maps from SKA

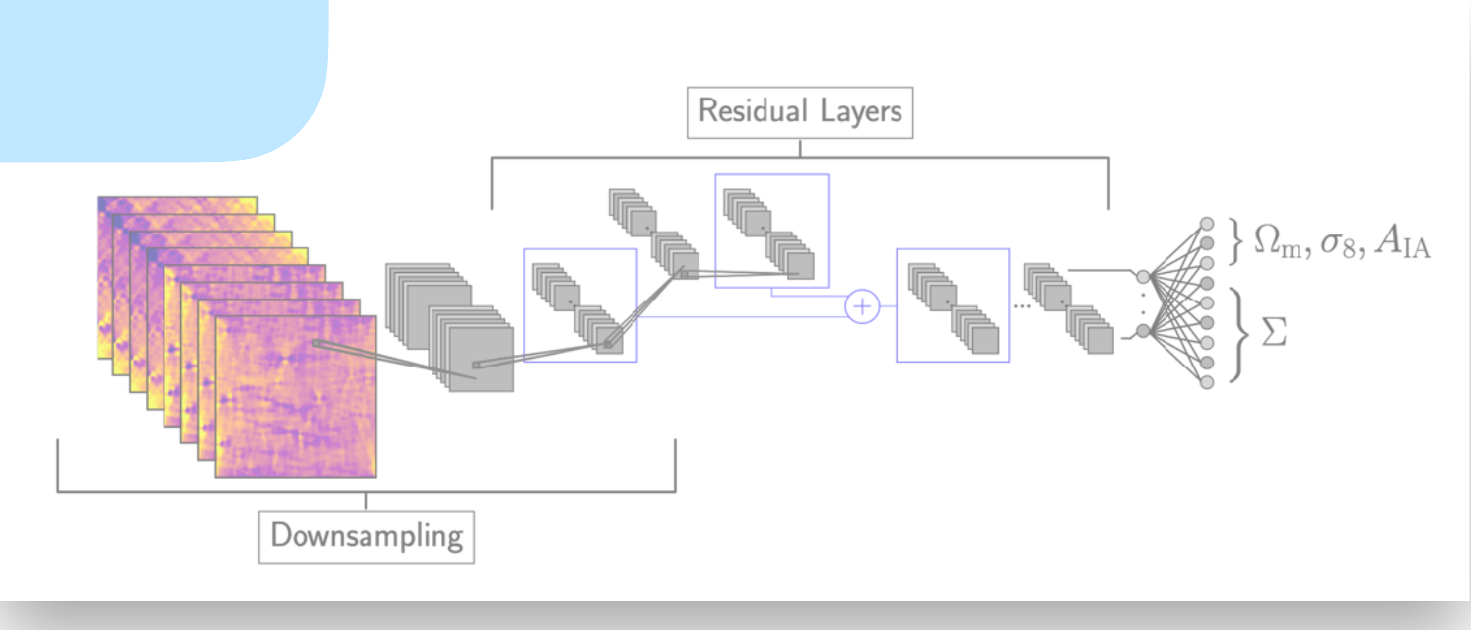
- Use the SKA 21-cm instrument model, including noise, angular resolution, foreground cleaning
- Using the SIMFAST21 simulation code
- Using CNN architectures: VGGNet, ResNet
- Simultaneously Ω_m, σ_8, h, and astrophysics:
  - Photon escape fraction f_{esc}
  - Ionizing emissivity power dependence on halo mass C_{ion}
  - Ionizing emissivity redshift evolution index D_{ion}
- Very good accuracy!



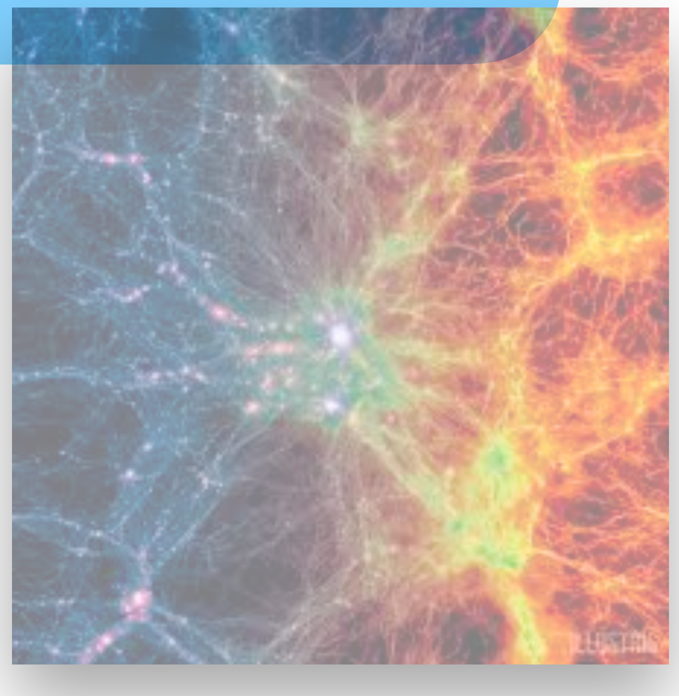
Hassan Andrianomena Doughty 2020 1907.07787

How can AI open new possibilities in cosmological analysis?

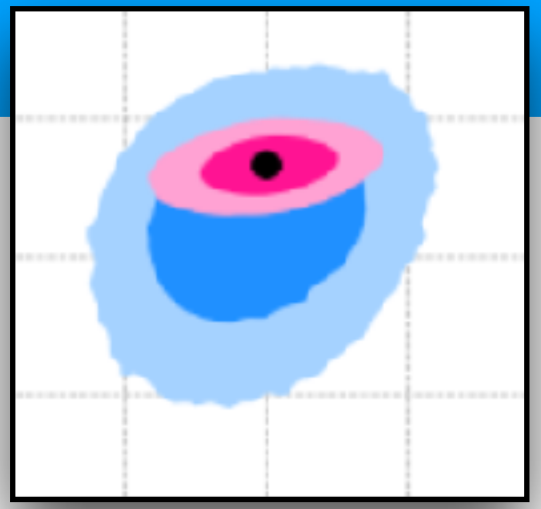
Reaching the information floor of the data



Accelearating simulations

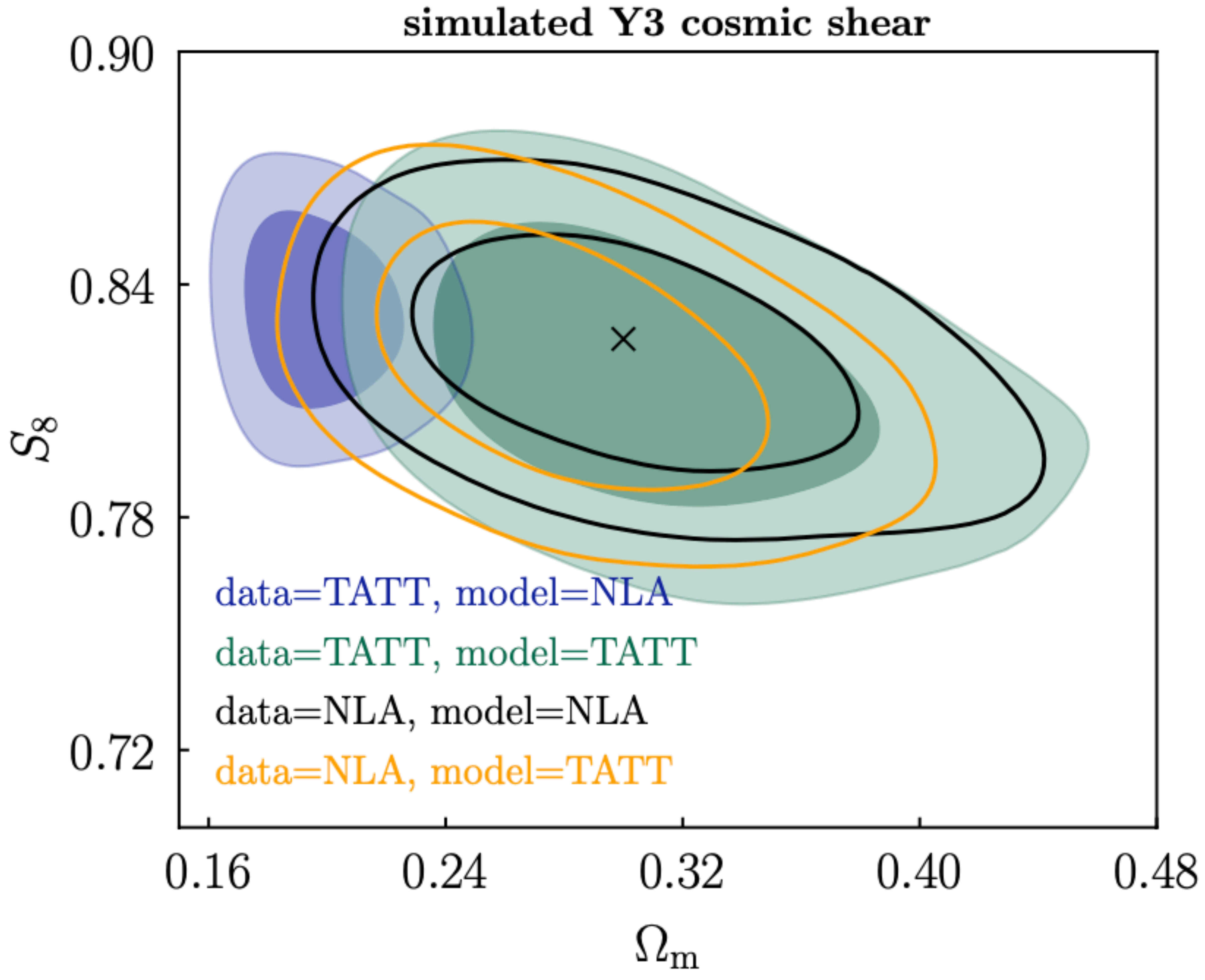


Breaking degeneracies between cosmology and systematics



Robust cosmology ↔ well constrained astrophysics

- Cosmology and astrophysics can be degenerate (IA, baryon, galaxy evolution)
- Measurements have a limited ability to constrain the models
- Too simple astrophysical models can lead to underfitting bias in results
- Marginalizing over complex astrophysical models can lead to loss of precision and to prior volume effects, which bias results
- Simulation-Based Analysis can extract more information, which enables constraining more complex astrophysical models without loss of precision in cosmology parameters
- For this, especially powerful is map-level probe combination

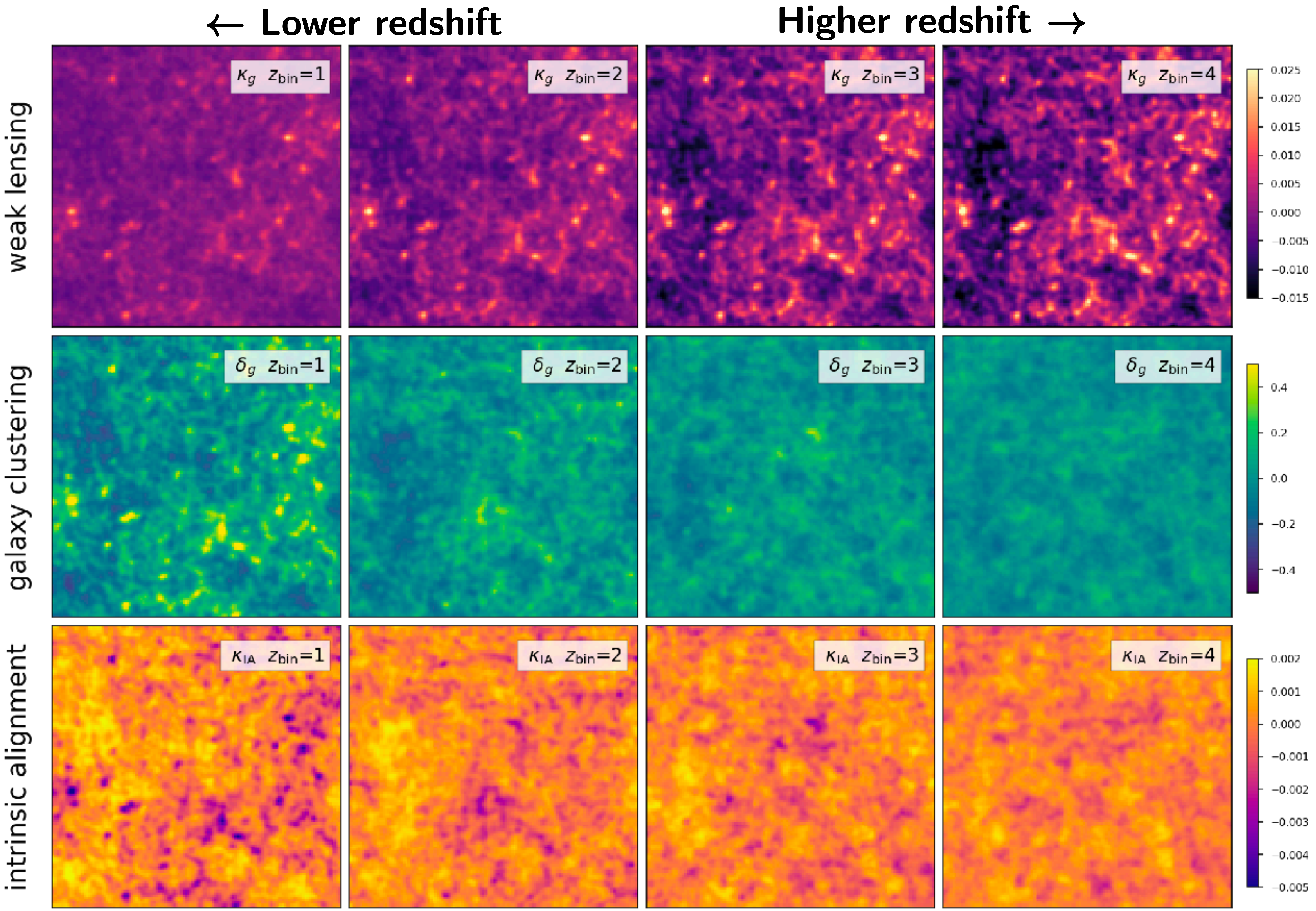


Secco+TK, DES Collab [2105.13544]

DeepLSS: combined probes with deep learning

Kacprzak and Fluri 2022, arXiv:2203.09616, Phys. Rev. X 12, 031029

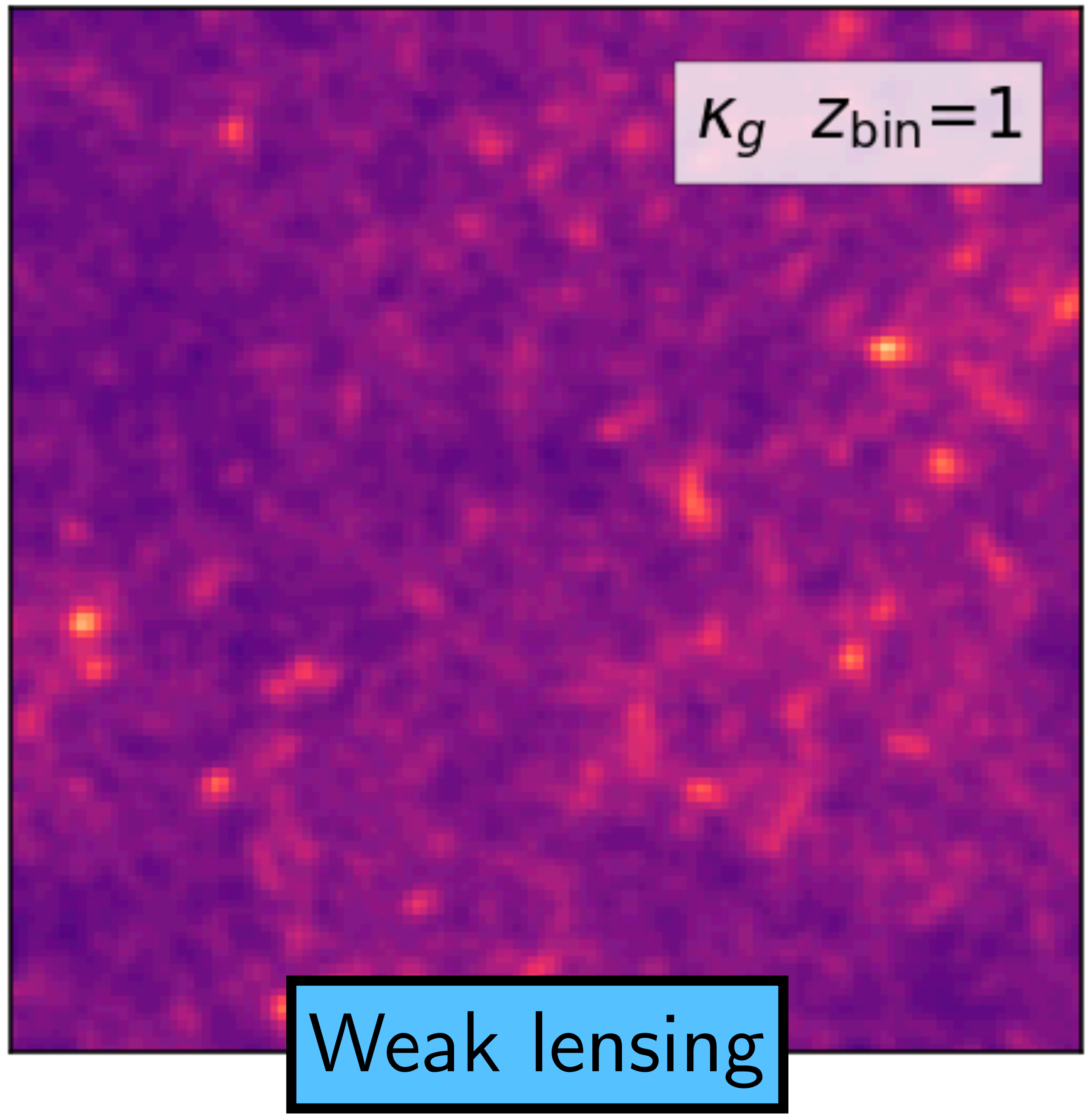
- Consistent simulations of galaxy shapes and positions
- Using Stage-III configuration with 4 tomographic bins
- Avoiding non-linear bias and baryons by smoothing the maps
- Code public: github.com/tomaszkacprzak/DeepLSS

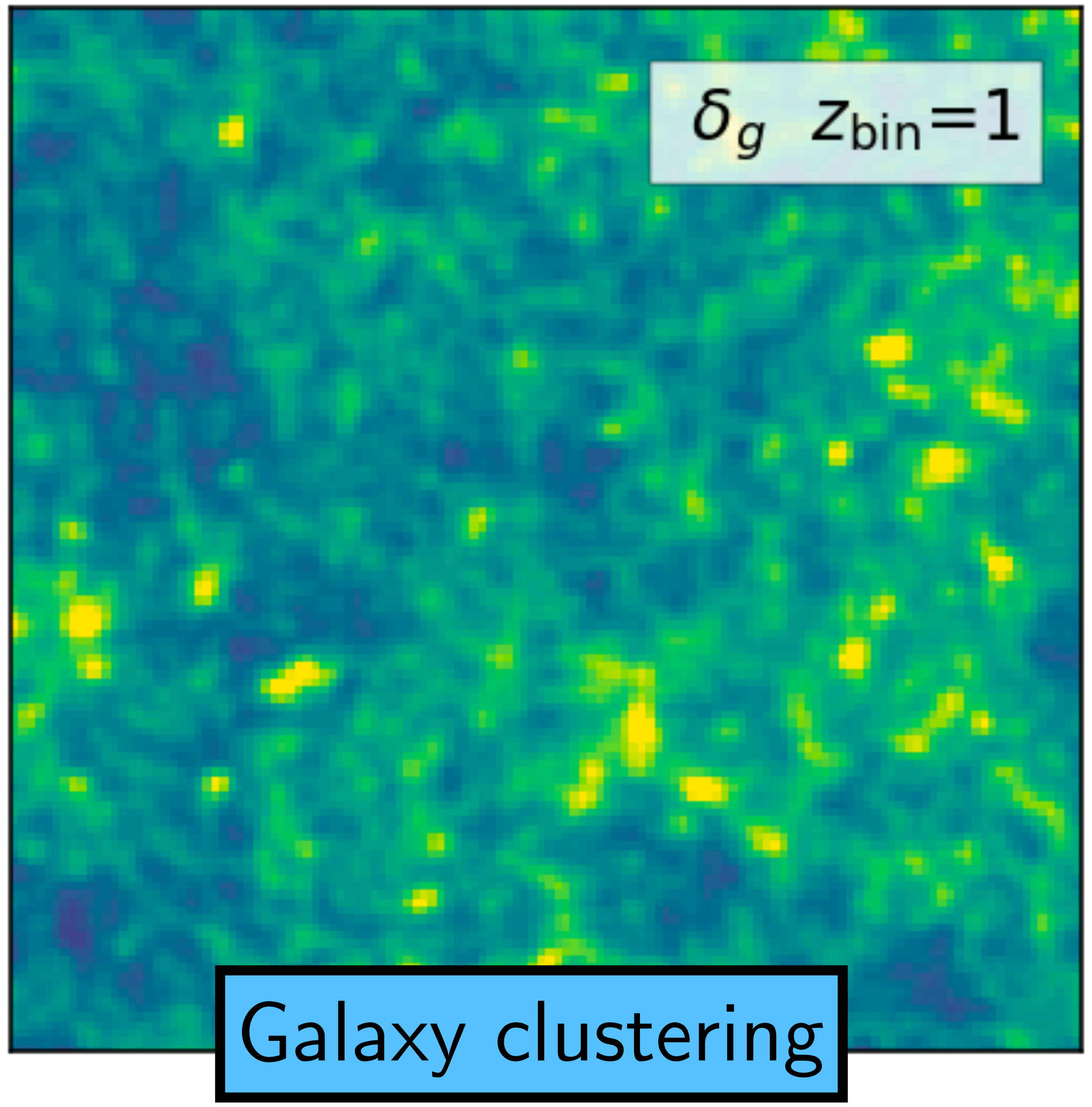


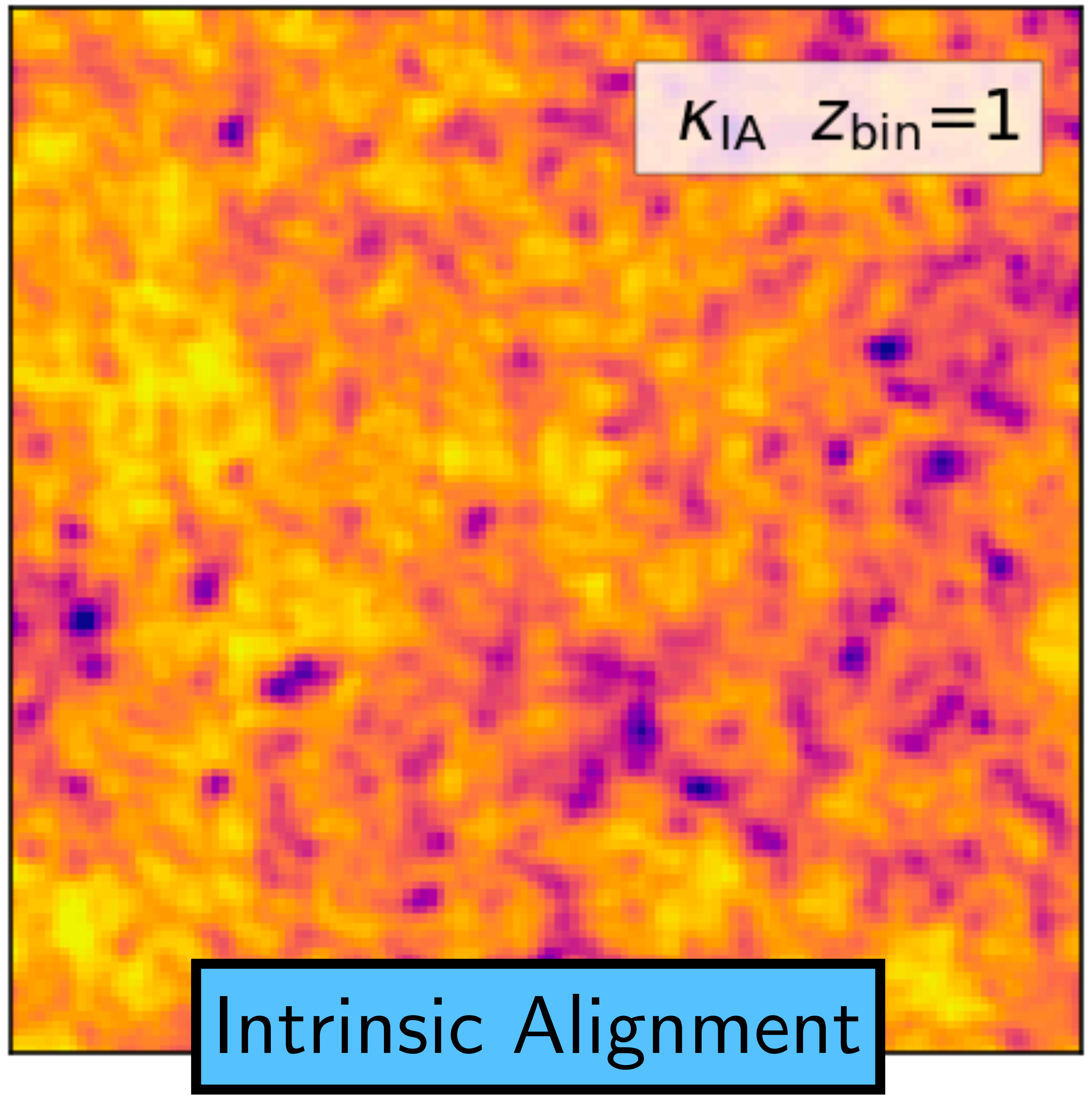
DeepLSS: combined probes with deep learning

Kacprzak and Fluri 2022, 2203.09616, Phys. Rev. X 12, 031029

Physical fields:



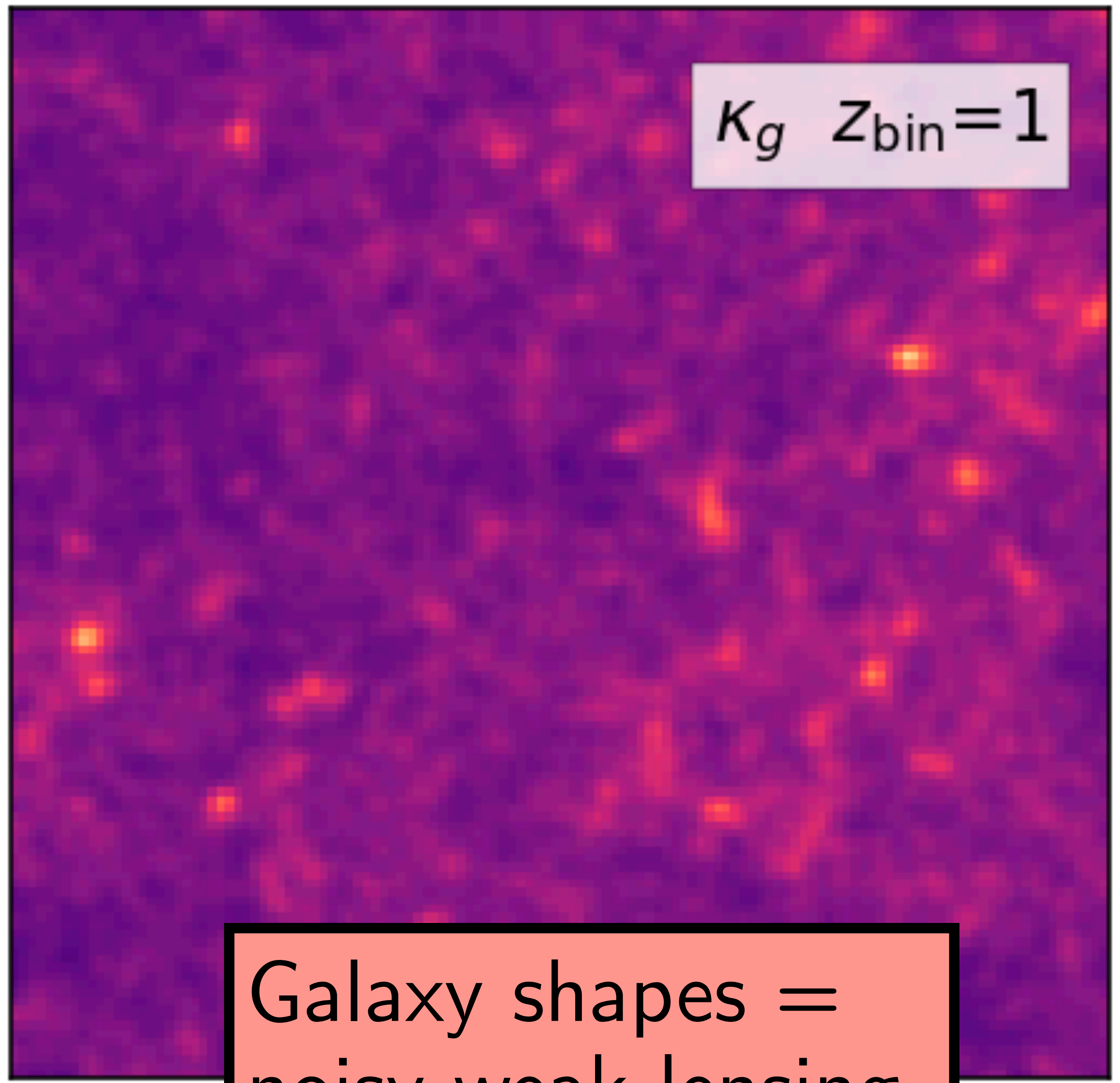




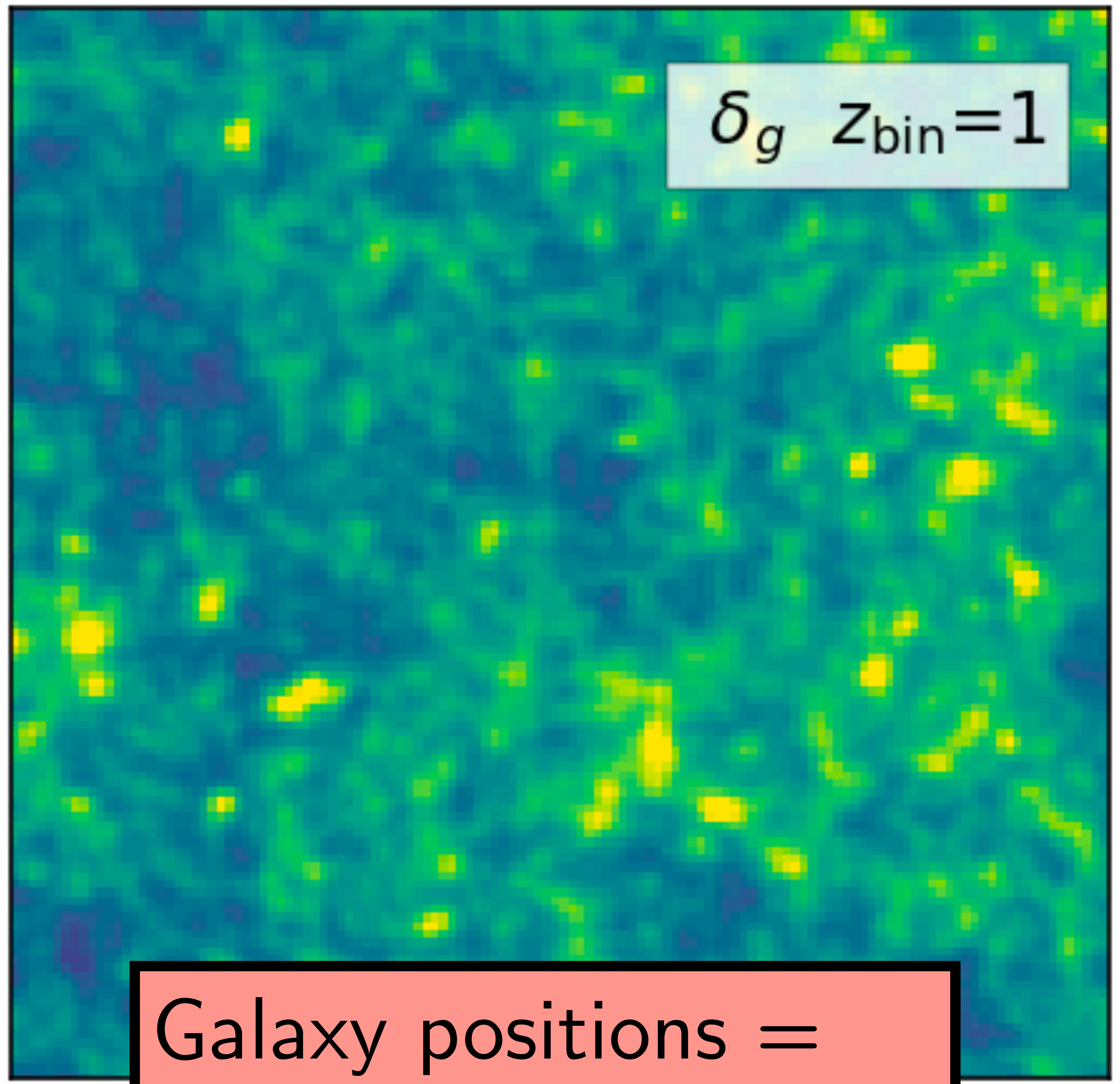
DeepLSS: combined probes with deep learning

Kacprzak and Fluri 2022, 2203.09616, Phys. Rev. X 12, 031029

Observables:



Galaxy shapes = noisy weak lensing **and** intrinsic alignment



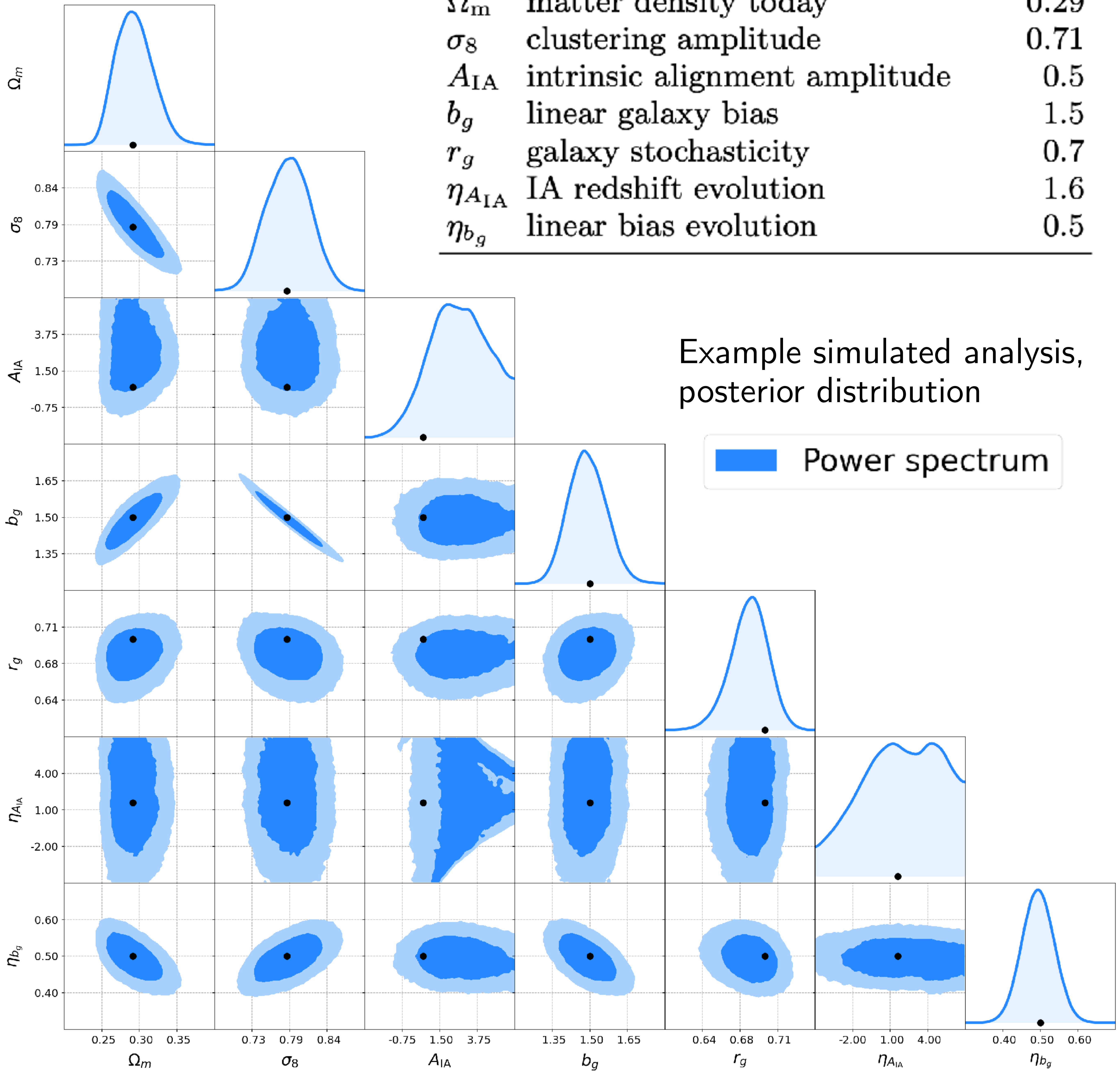
Galaxy positions = **biased** galaxy clustering

DeepLSS: combined probes with deep learning

- Apples-to-apples comparison between power spectra and deep learning
- Biasing sector constraints also improved by 30-40%
- Deep learning analysis breaks several key degeneracies
- Intrinsic alignment measurement is greatly de-correlated from cosmology
- Galaxy biasing evolution is also de-correlated from cosmology
- Cosmology constraints improved due to breaking degeneracy with IA

	description	fiducial
Ω_m	matter density today	0.29
σ_8	clustering amplitude	0.71
A_IA	intrinsic alignment amplitude	0.5
b_g	linear galaxy bias	1.5
r_g	galaxy stochasticity	0.7
η_{A_IA}	IA redshift evolution	1.6
η_{b_g}	linear bias evolution	0.5

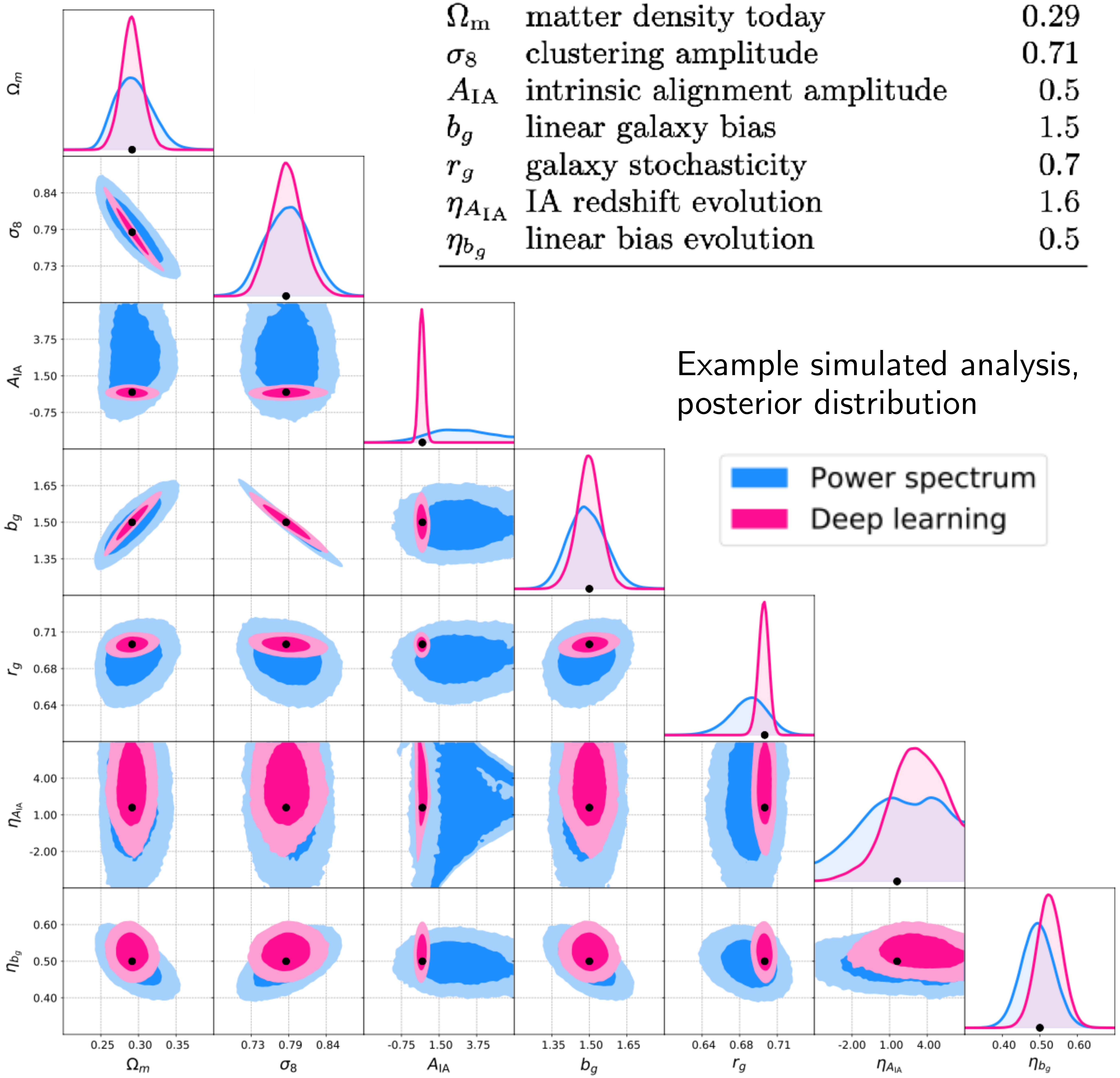
Example simulated analysis, posterior distribution



DeepLSS: combined probes with deep learning

- Apples-to-apples comparison between power spectra and deep learning
- Biasing sector constraints also improved by 30-40%
- Deep learning analysis breaks several key degeneracies
- Intrinsic alignment measurement is greatly de-correlated from cosmology
- Galaxy biasing evolution is also de-correlated from cosmology
- Cosmology constraints improved due to breaking degeneracy with IA

	description	fiducial
Ω_m	matter density today	0.29
σ_8	clustering amplitude	0.71
A_IA	intrinsic alignment amplitude	0.5
b_g	linear galaxy bias	1.5
r_g	galaxy stochasticity	0.7
η_{A_IA}	IA redshift evolution	1.6
η_{b_g}	linear bias evolution	0.5

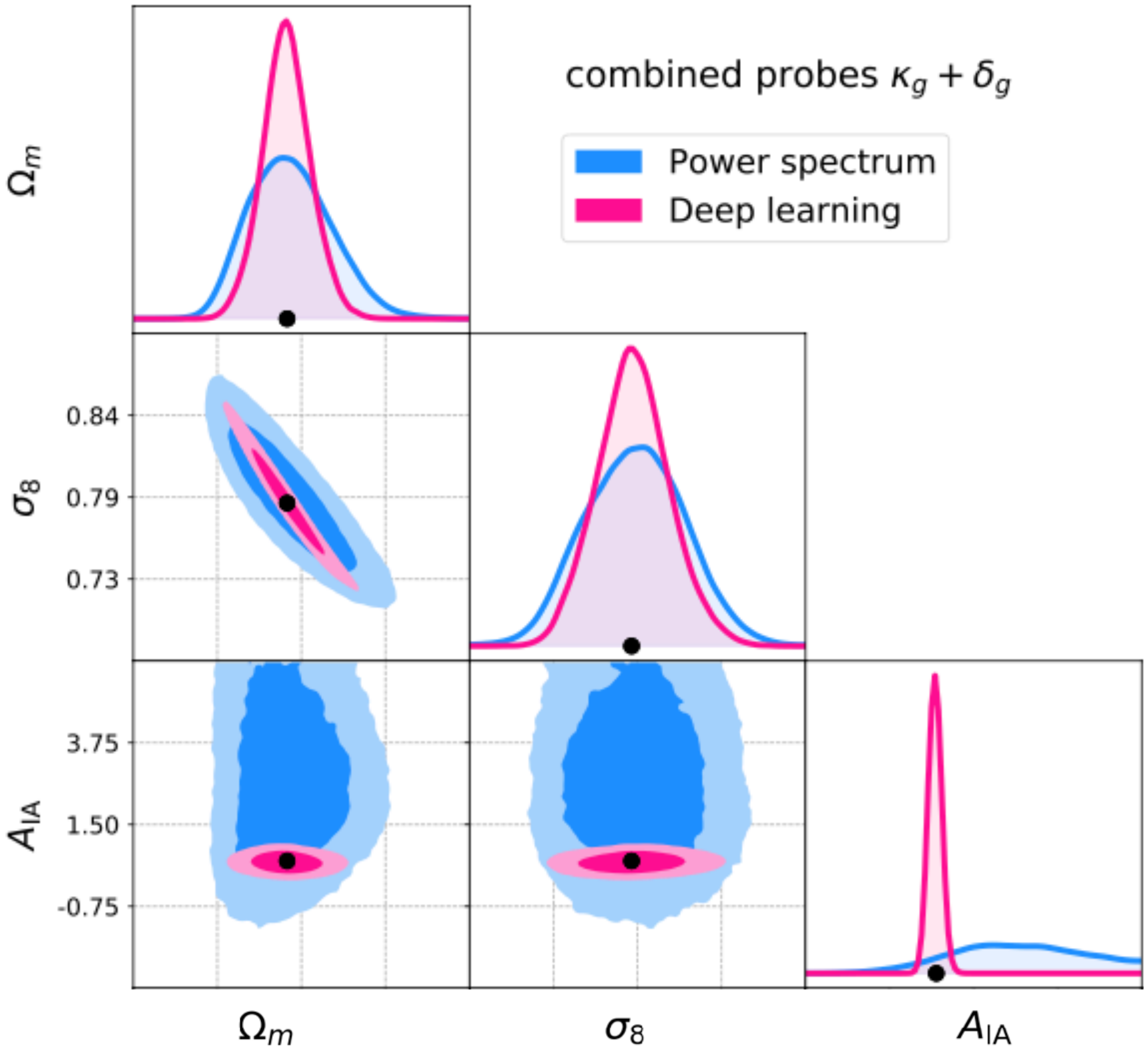


Example simulated analysis, posterior distribution

Kacprzak and Fluri 2022, arXiv:2203.09616

DeepLSS: combined probes with deep learning

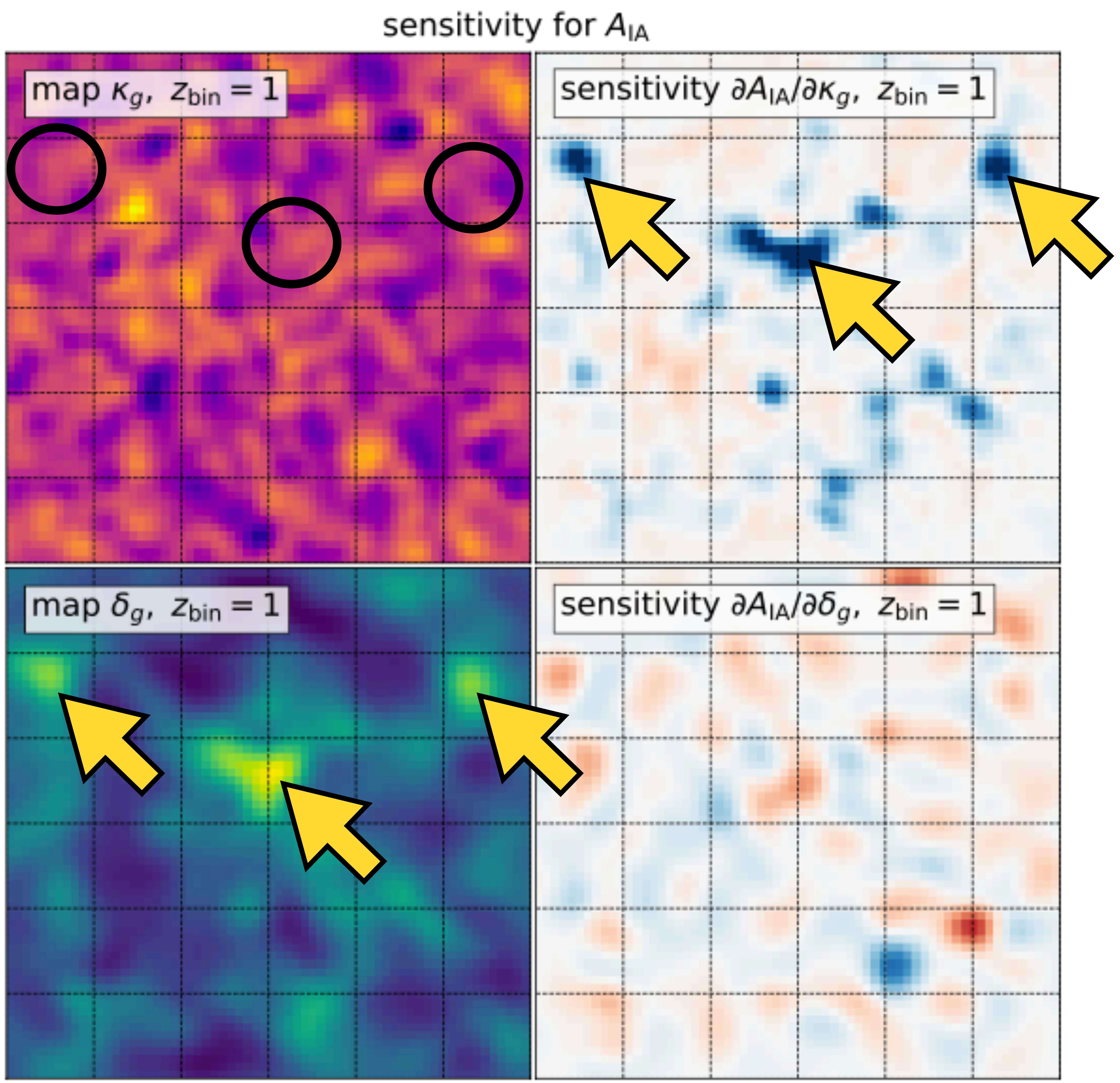
- Apples-to-apples comparison between power spectra and deep learning
- Deep learning analysis breaks several key degeneracies
- Intrinsic alignment measurement is greatly de-correlated from cosmology
- Cosmology constraints improved due to breaking degeneracy with IA



Kacprzak and Fluri 2022, Phys. Rev. X 12, 031029

Where is the additional information coming from?

- Sensitivity maps show which pixels have the most impact on the final prediction of the network
- The networks focuses on very specific regions in the galaxy positions and lensing maps
- Deep learning weights the data in a way that maximises information gain



Kacprzak and Fluri 2022, Phys. Rev. X 12, 031029

What is the network really learning?

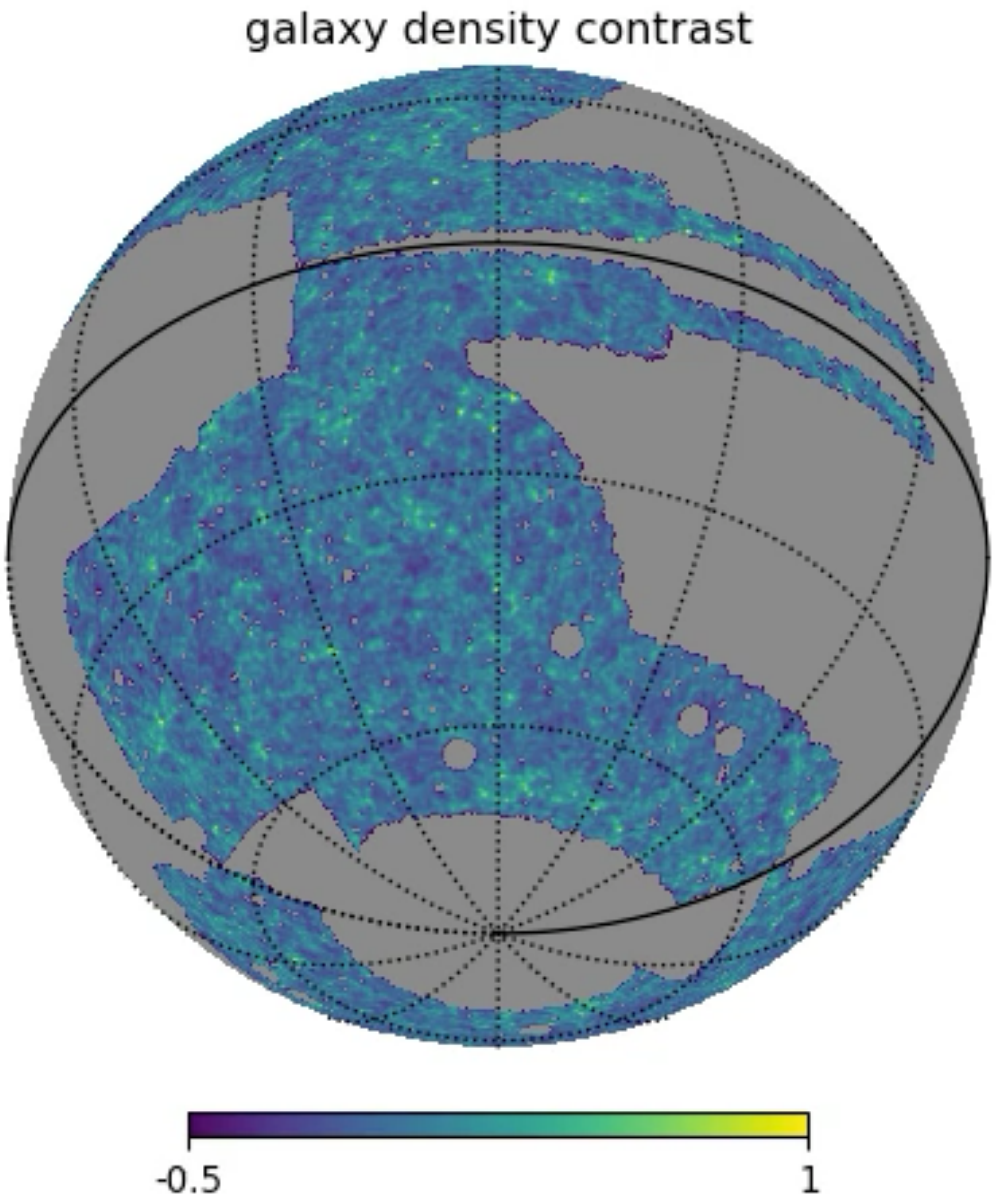
- It is important to understand where the information is coming from
- One way to understand it: which areas of the map are being used to make a decision about the output cosmology?
- A number of interpretability measures are present in literature for computer vision
- First paper on their application to lensing: Zorrilla Matilla et al. 2020
- Different interpretability measures give different insights, maps for Ω_m output neuron

input lensing map **gradient** **guided backprop** **Taylor decomposition** **LRP-$\alpha\beta$**

Zorrilla Matilla et al. 2020 2007.06529

Simulations-Based Inference in Dark Energy Survey

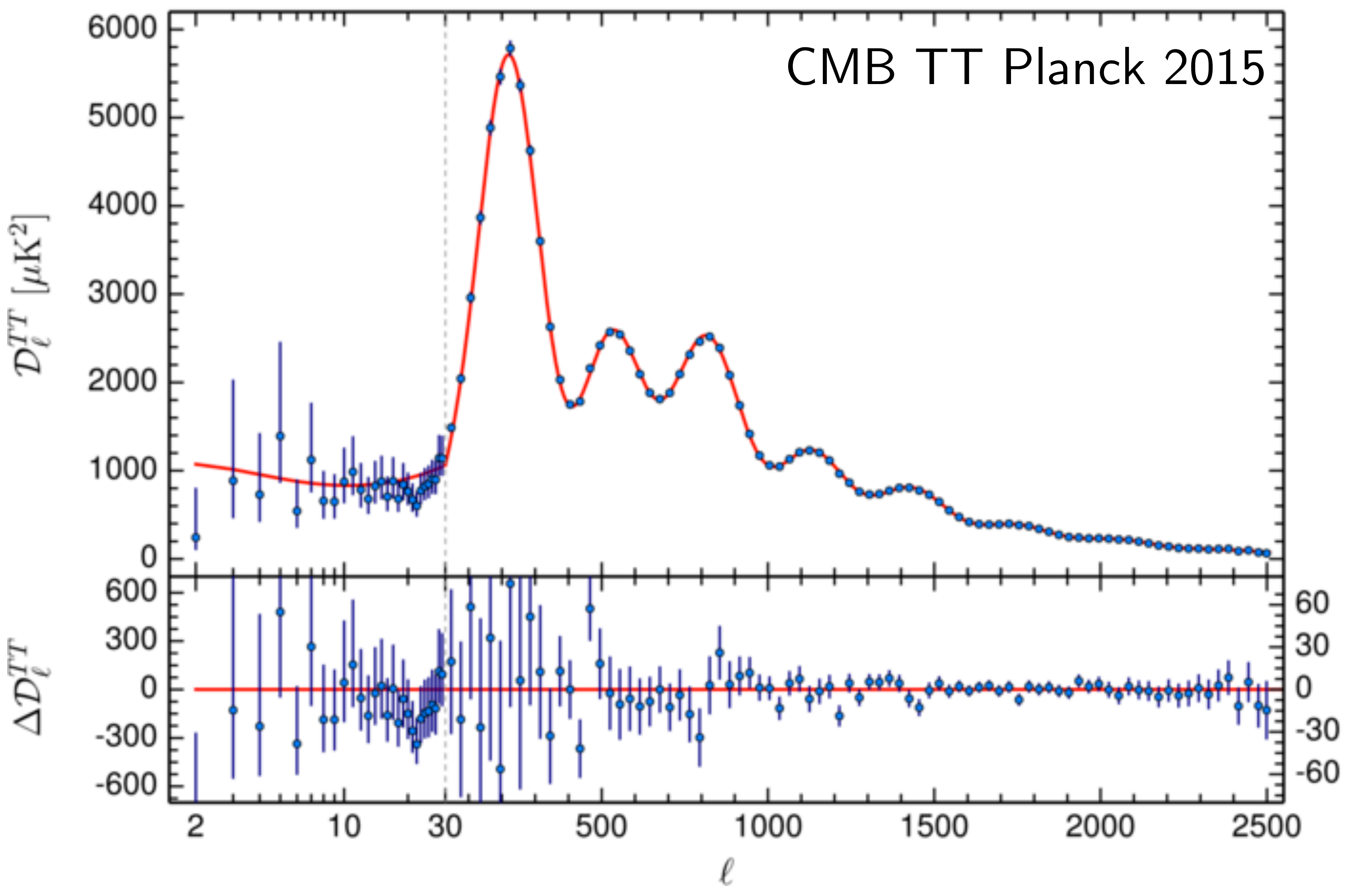
- First multiprobe map-level SBI analysis with deep learning and peak counts
- Traninig set size ∼50 TB + on-the-fly augmentations: noise addition, systematic effects
- Currently training low-resoultion version (batch size 100m pixels)
- Running on 2 NERSC Perlmutter nodes (8 GPUs)
- NERSC Science Acceleration Program (NESAP) is helping us with scaling to high-res, 4x more pixels
- See poster by Arne Thomsen



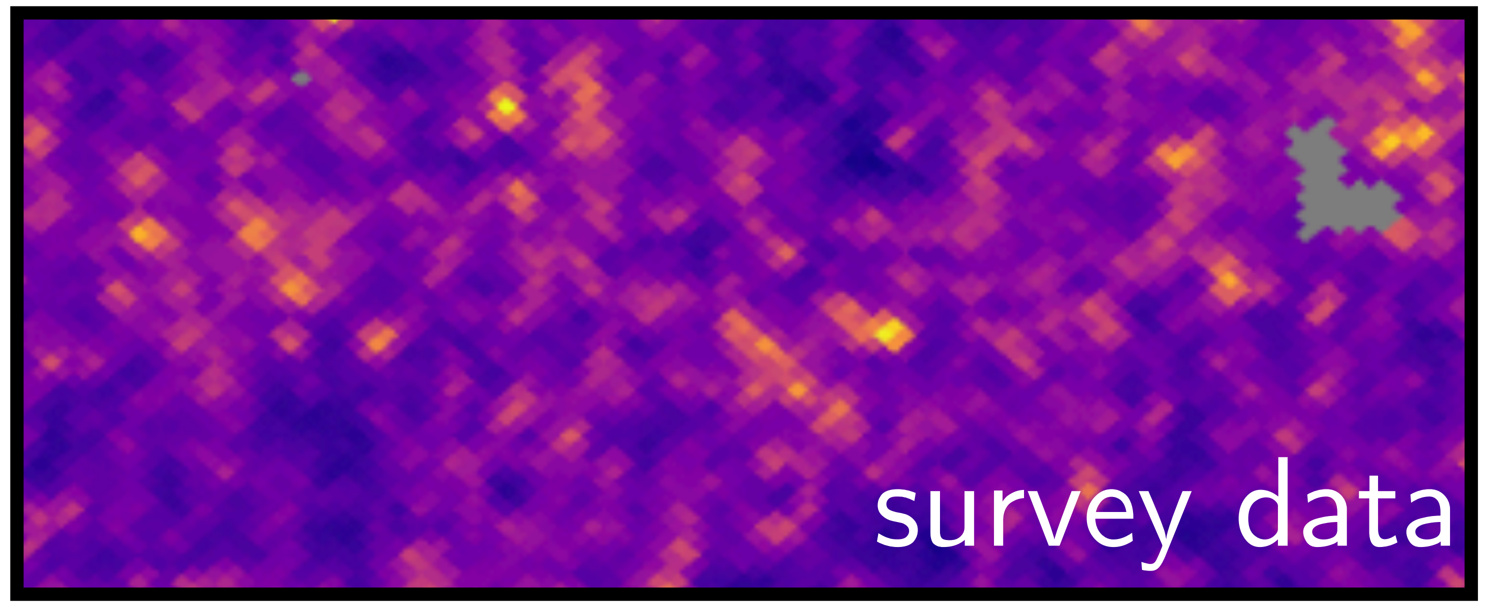
Data Science Challenge: map-level goodness of fit

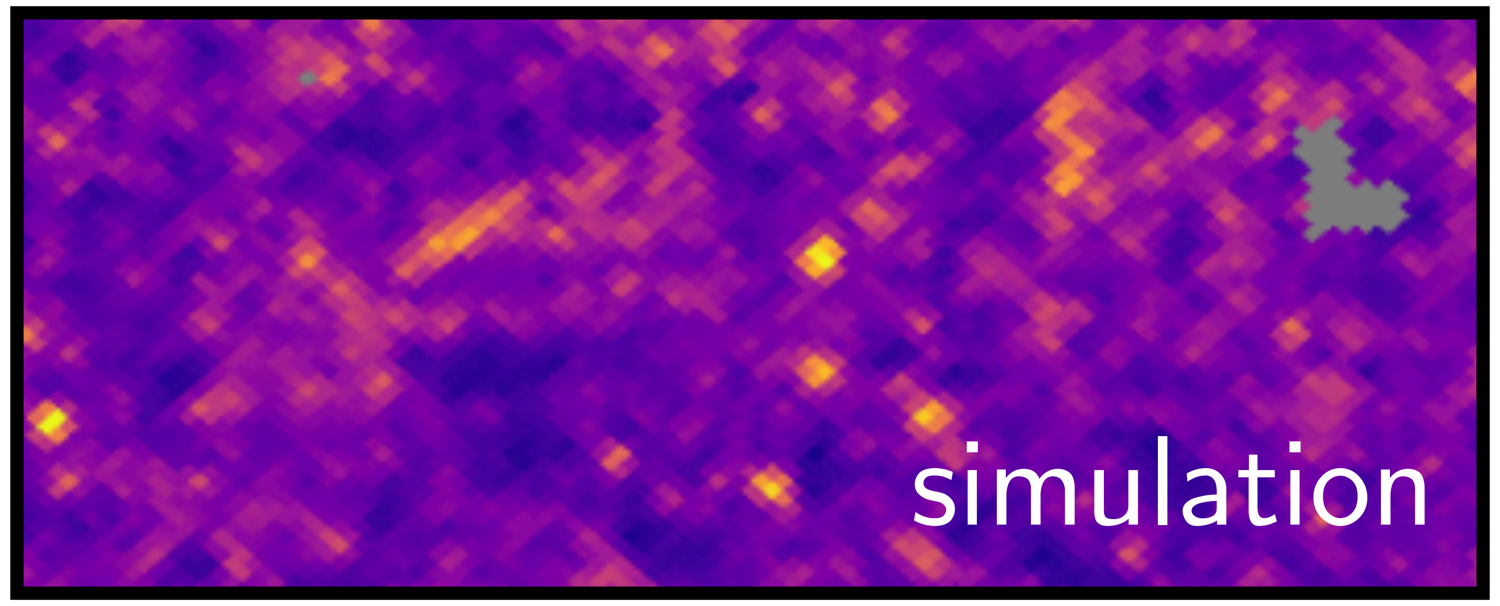
- Goal: simulations-based inference without bias stemming from a wrong simulation model
- In traditional inference with summaries, simple metrics exist to quantify this, simplest being the reduced χ^2
- Problem: how to evaluate goodness of fit for a 1m-dimensional map vector
- No established methods existing yet
- Possible approach: decompose data and check stability of results:
  - different areas of the survey
  - scale decomposition: naturally included in normalizing flows (Dai Seljak 2023 2306.04689)
- Another proposed method: foundation models

Distribution of residuals tells us about goodness of fit, for example χ^2_{red}



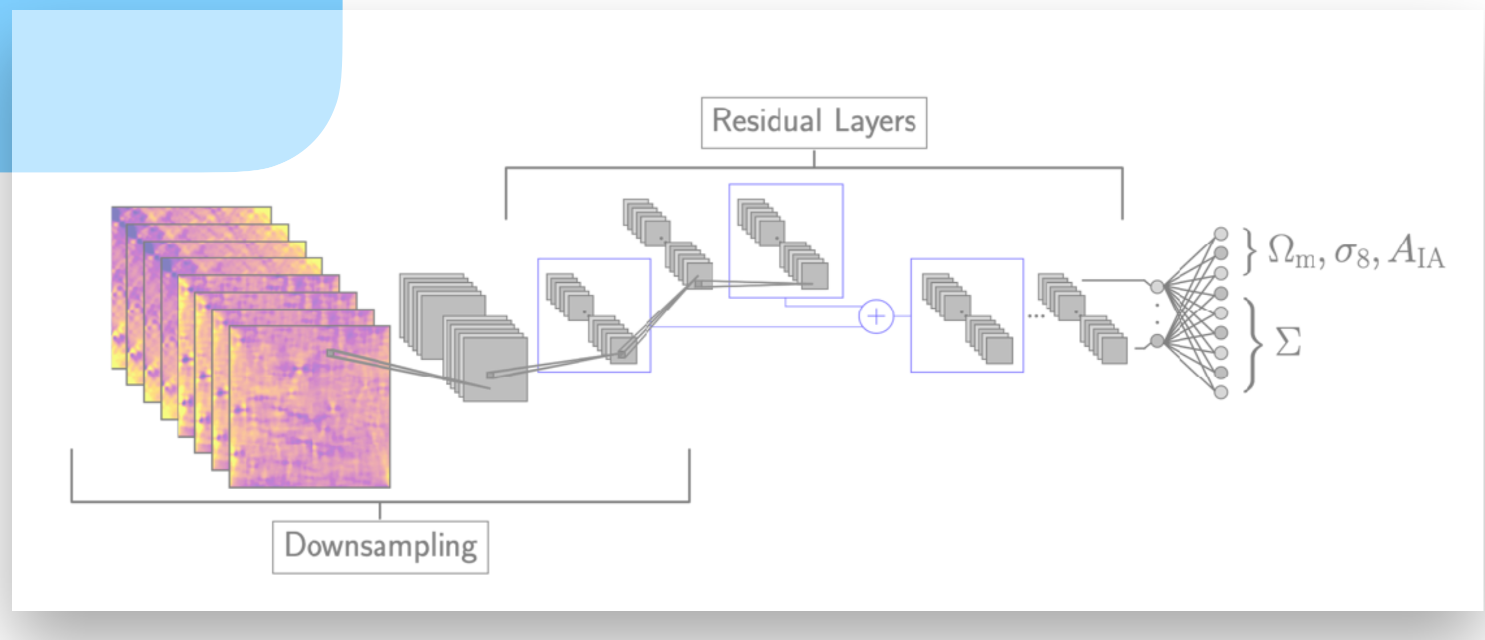
How to quantify goodness of fit for stochastic maps?



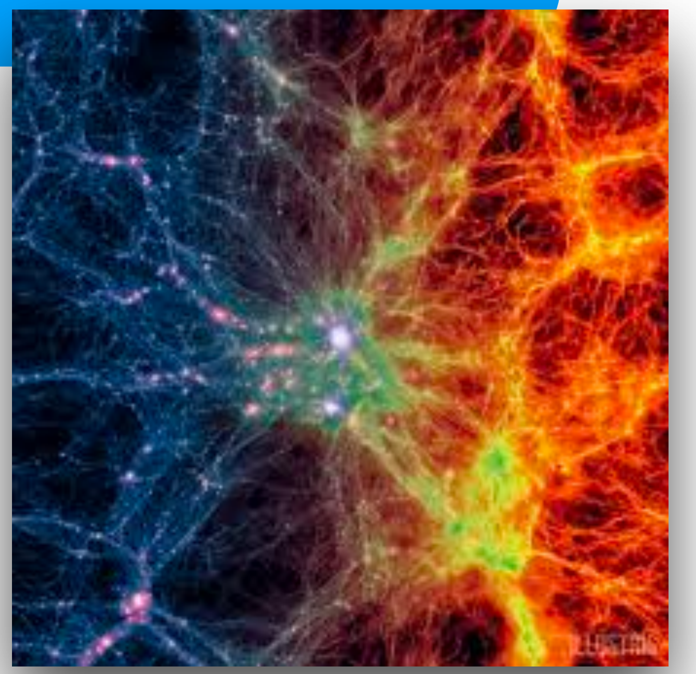


How can AI open new possibilities in cosmological analysis?

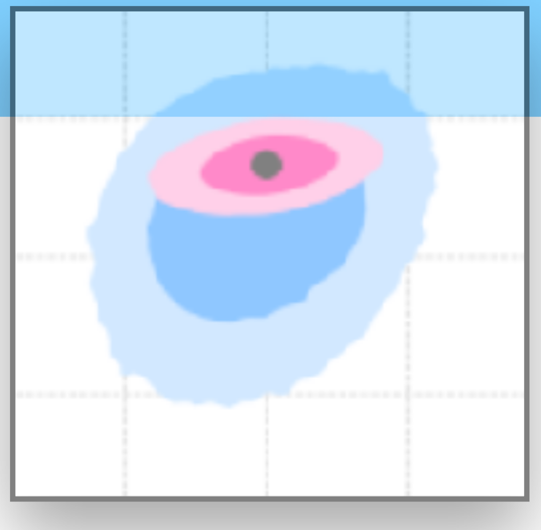
Reaching the information floor of the data



Accelearating simulations

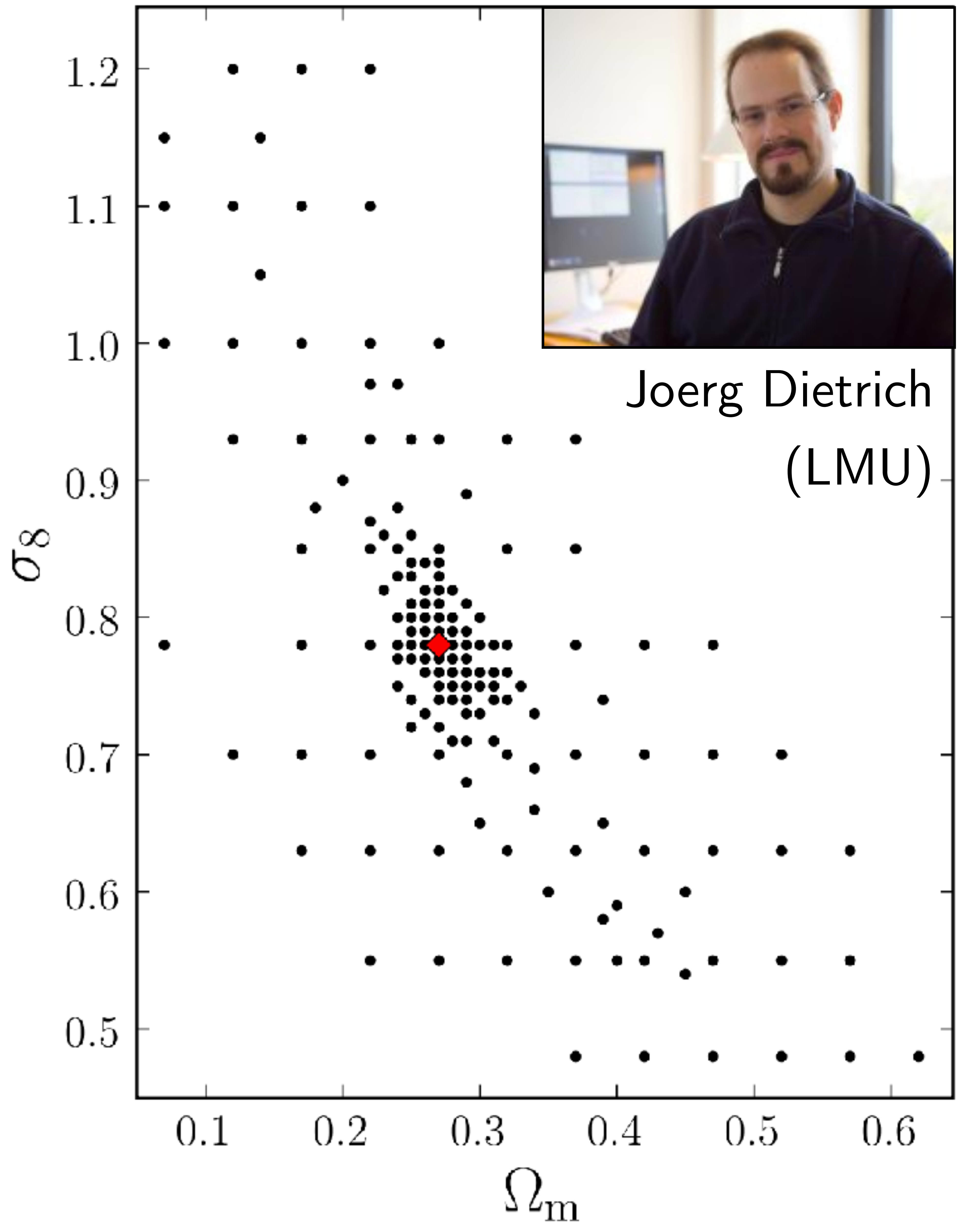


Breaking degeneracies between cosmology and systematics



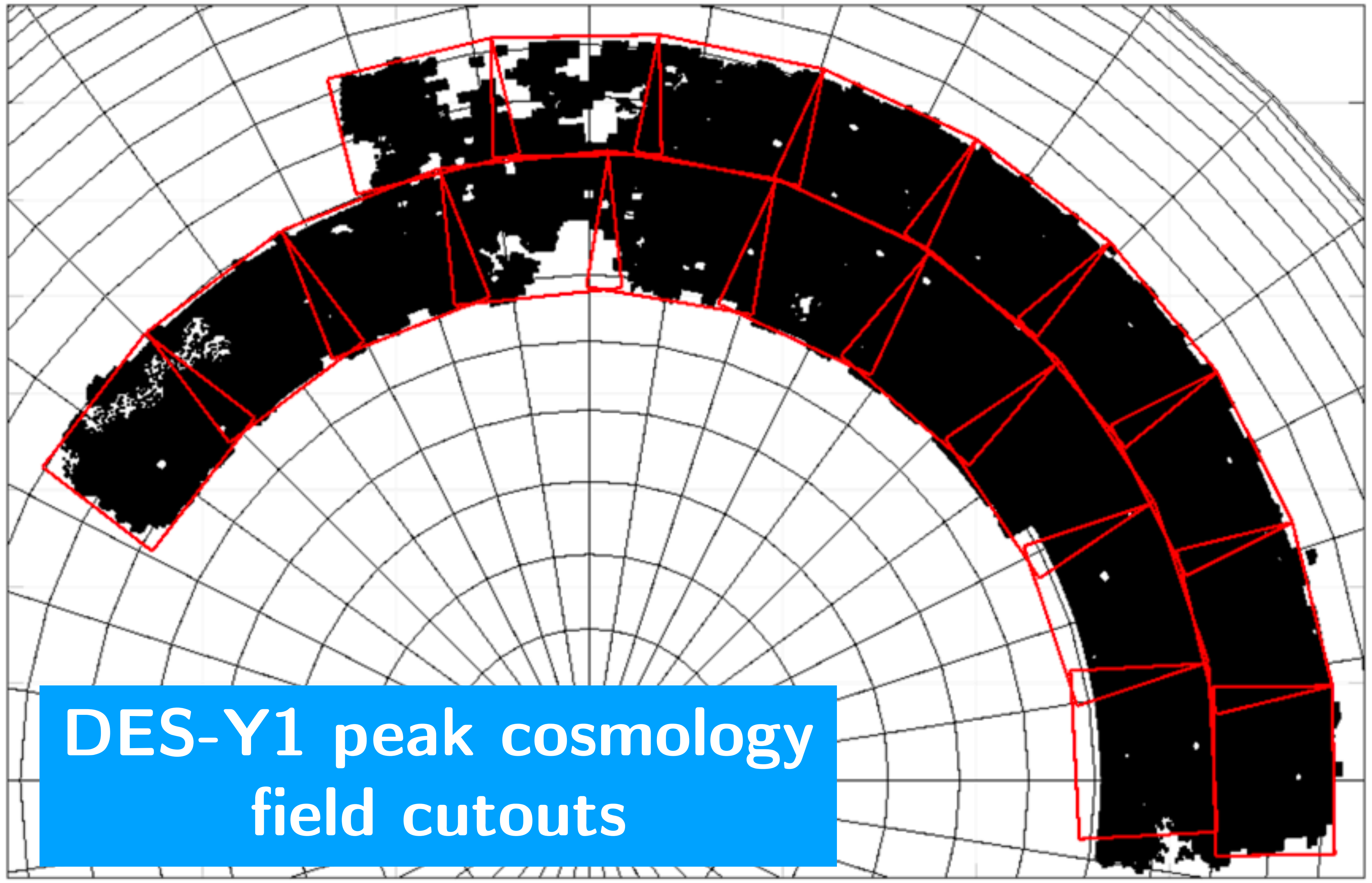
The simulation set that started it all: Dietrich & Hartlap 2009

- First simulations-based inference for cosmology, with two types of peak counts functions
- 158 simulations in $\Omega_m - \sigma_8$ space
- Pencil-beam lightcones from 256^3 particle simulations in $200\ h^{-1}\ Mpc$ boxes
- 6×6 deg^2 convergence maps
- Uniformly sampled background galaxies with shears from fixed n(z)
- Used in first shear peak statistics papers:
  - CFHTLenS: Liu et al. 1412.0757
  - DES-SV: Kacprzak et al. 1603.05040
  - KiDS-450: Martinet et al. 1709.07678



Cosmo-Slics

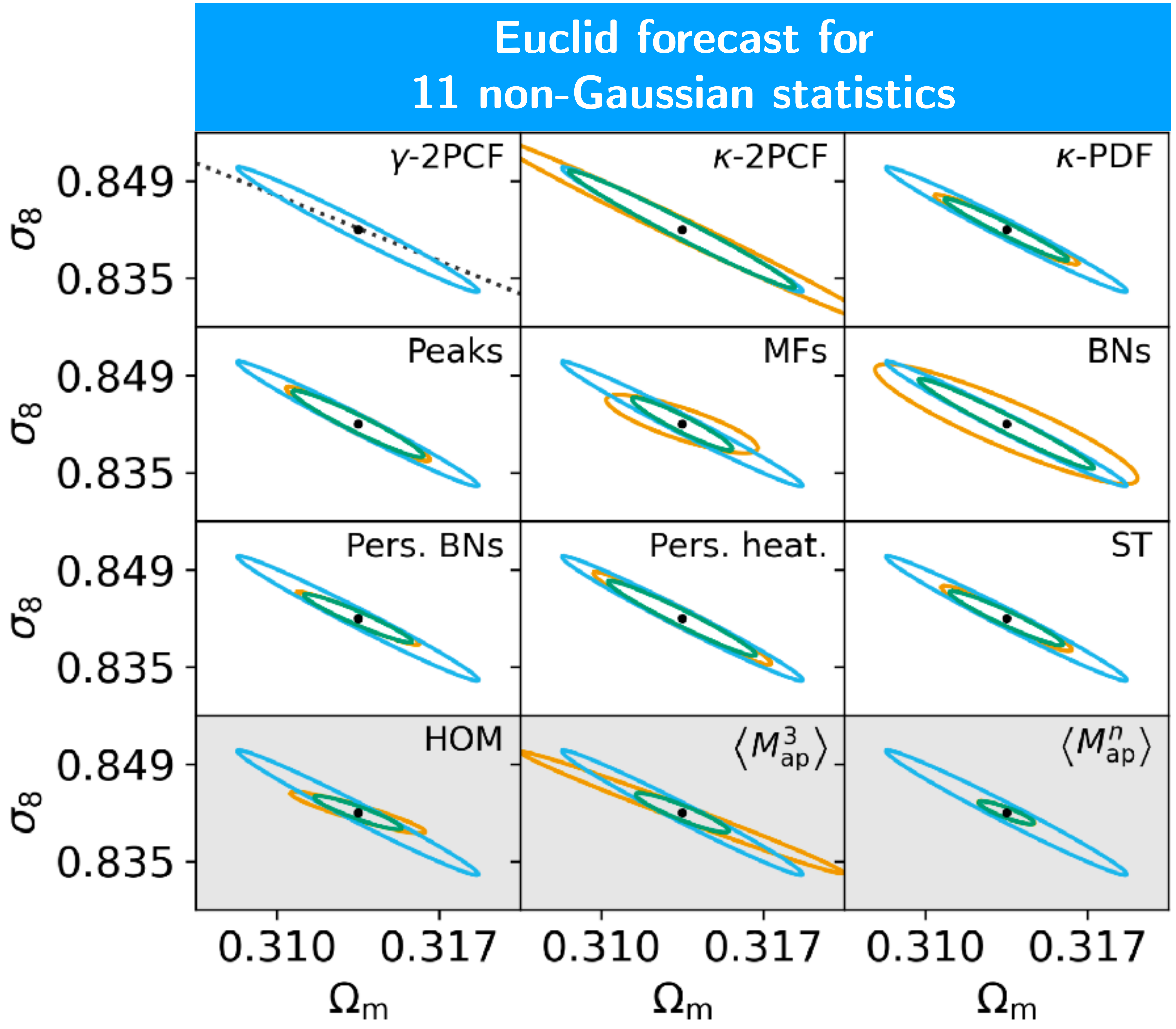
- 26 cosmology parameters spanning Ω_m, $\sigma_8\, h$, w
- 100s of realizations at fiducial cosmology, good for covariance validation
- 10×10 deg² pencil-beam lensing maps
- High resolution convergence: 15362 particles inside $505\ h^{-1}\ Mpc$ boxes
- Great for small scales, but at the moment lacking baryon feedback models
- Used in DES-Y1 peak counts cosmology paper: Harnois-Déraps et al. 2021 2012.02777
- Used in many forecasts for non-Gaussian statistics, latest for Euclid Preparation Key Project
- Available at: https://slics.roe.ac.uk



Harnois-Déraps et al. 2021 2012.02777

Cosmo-Slics

- 26 cosmology parameters spanning Ω_m, σ_8 h, w
- 100s of realizations at fiducial cosmology, good for covariance validation
- 10×10 deg² pencil-beam lensing maps
- High resolution convergence: 15362 particles inside $505\ h^{-1}\ Mpc$ boxes
- Great for small scales, but at the moment lacking baryon feedback models
- Used in DES-Y1 peak counts cosmology paper: Harnois-Déraps et al. 2021 2012.02777
- Used in many forecasts for non-Gaussian statistics, latest for Euclid Preparation Key Project
- Available at: https://slics.roe.ac.uk

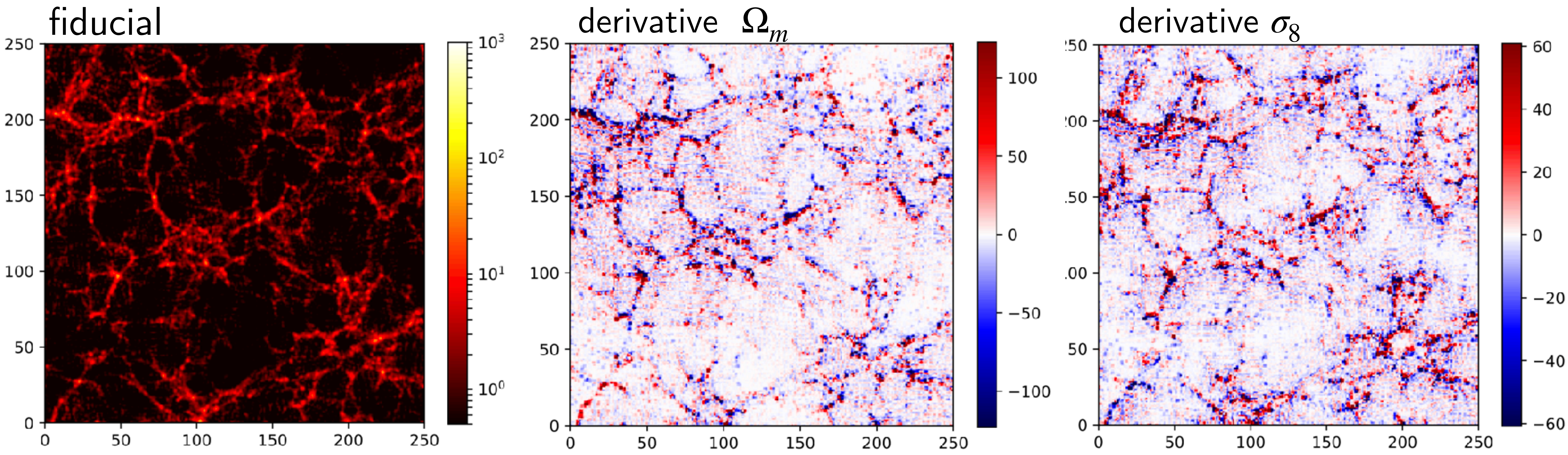


Euclid Collab. 2023 2301.12890

Quichote simulations

- 11000 simulations distributed in a latin hypercube
- Varying parameters Ω_m, $\sigma_8\, \Omega_b\, n_s$, σ_8, M_ν, w
- 7000 cosmological models
- Central simulation and derivatives from the same initial conditions

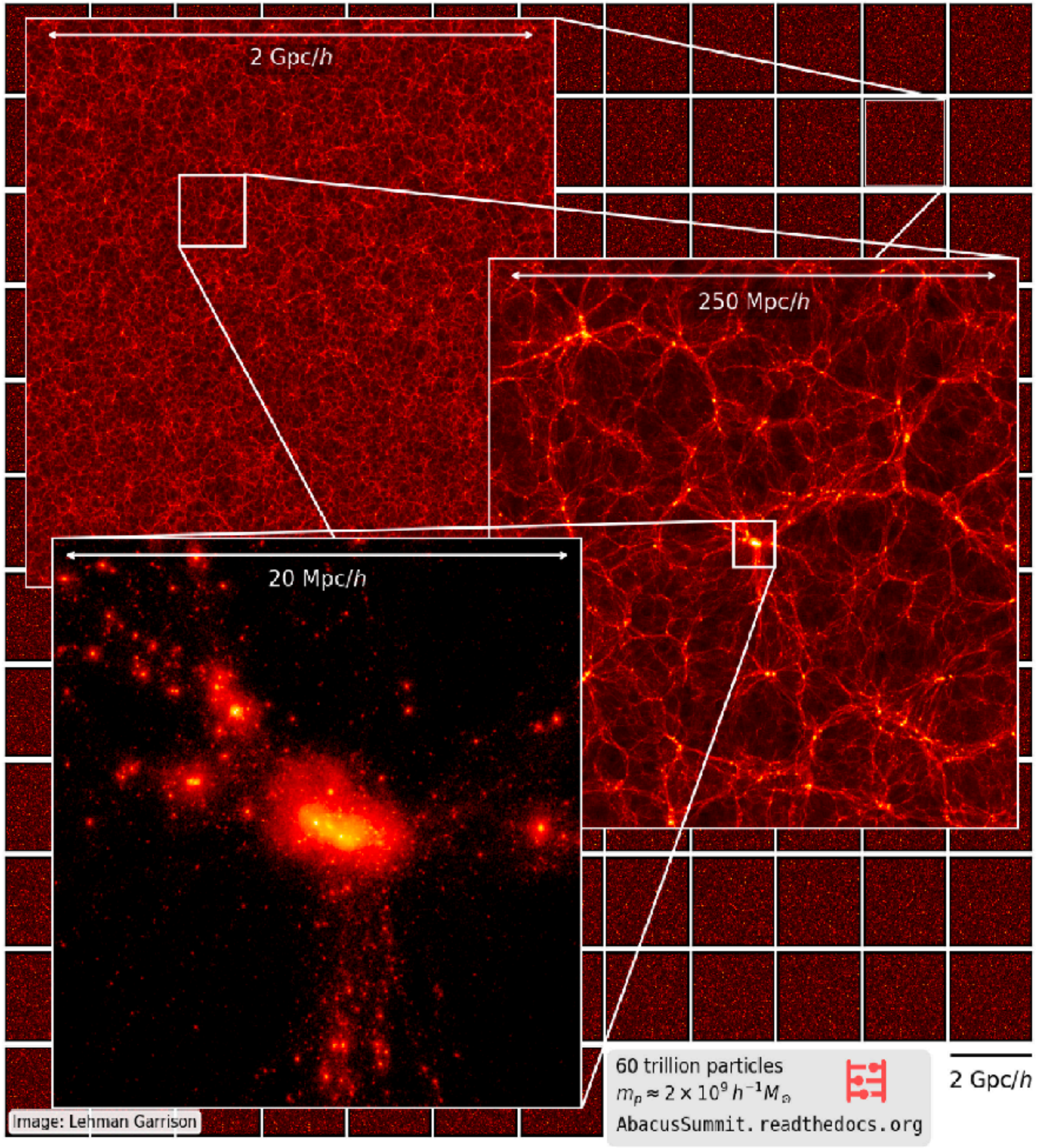
- 1 Gpc/h boxes with 512^3 particles
- 45500 simulations total
- High-res benchmarks
- Snapshots at z=0, 0.5, 1, 2, 3
- Data available at https://quijote-simulations.readthedocs.io/



Villaescusa-Navarro et al. 2019, 1909.05273

The ABACUS-SUMMIT simulations

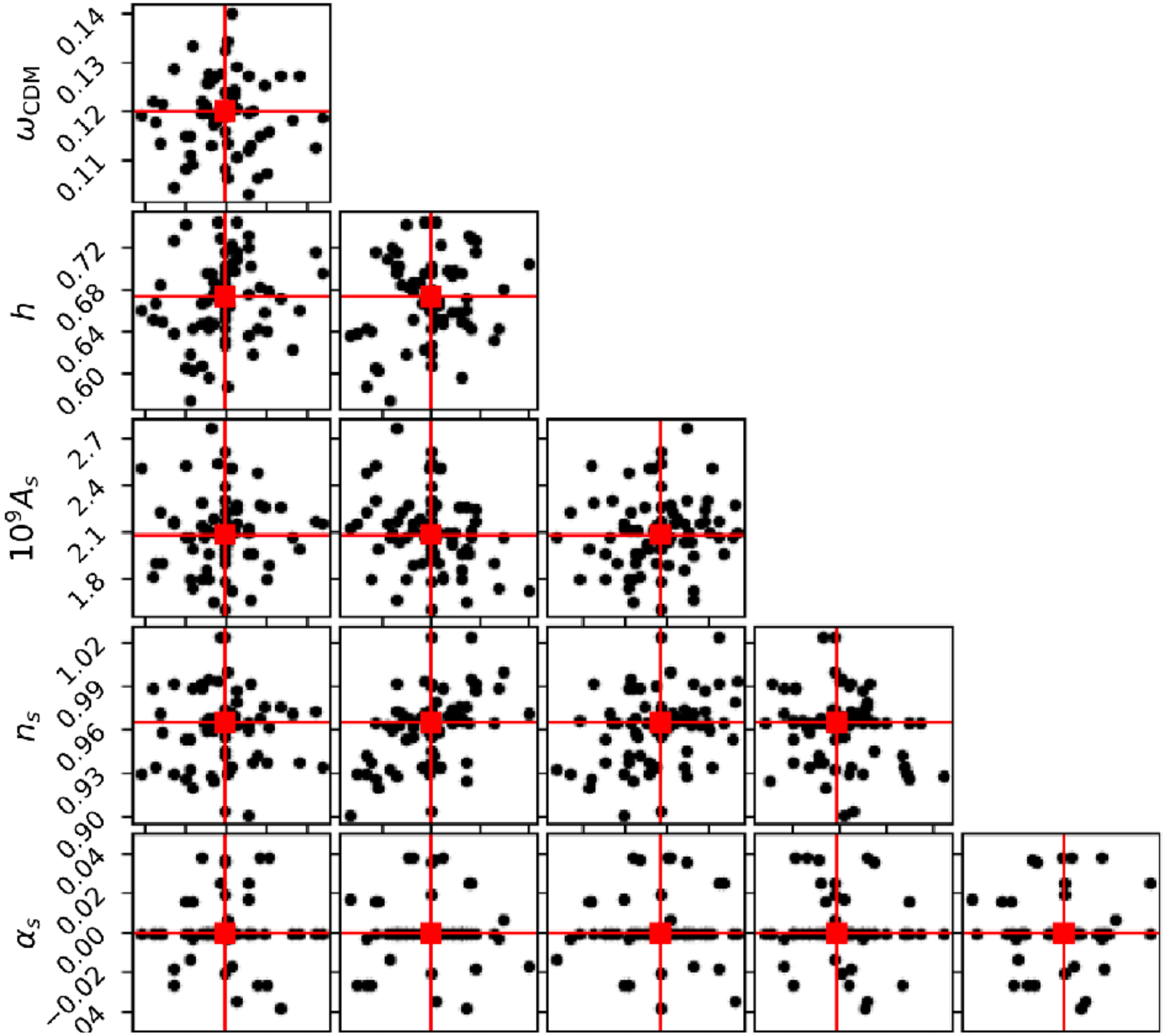
- 139 simulations with 6912^3 particles in $2\ h^{-1}\ Gpc$ boxes
- 97 cosmological models
- Spanning Ω_m, A_s, h, n_s core parameters, +additional "derivatives" for extended parameters
- Stored 33 timesteps in for $z \in [0.1,8]$
- full snapshots, lightcones, halo catalogs, particle subsets, merger trees
- 2PB of data products
- data products support using HODs to create galaxy catalogs
- Ran on Summit, one of the largest supercomputers in the world



Maksimova et al. 2021 2110.11398

The ABACUS-SUMMIT simulations

- 139 simulations with 6912^3 particles in $2\ h^{-1}\ Gpc$ boxes
- 97 cosmological models
- Spanning Ω_m, A_s, h, n_s core parameters, +additional "derivatives" for extended parameters
- Stored 33 timesteps in for $z \in [0.1, 8]$
- full snapshots, lightcones, halo catalogs, particle subsets, merger trees
- 2PB of data products
- data products support using HODs to create galaxy catalogs
- Ran on Summit, one of the largest supercomputers in the world

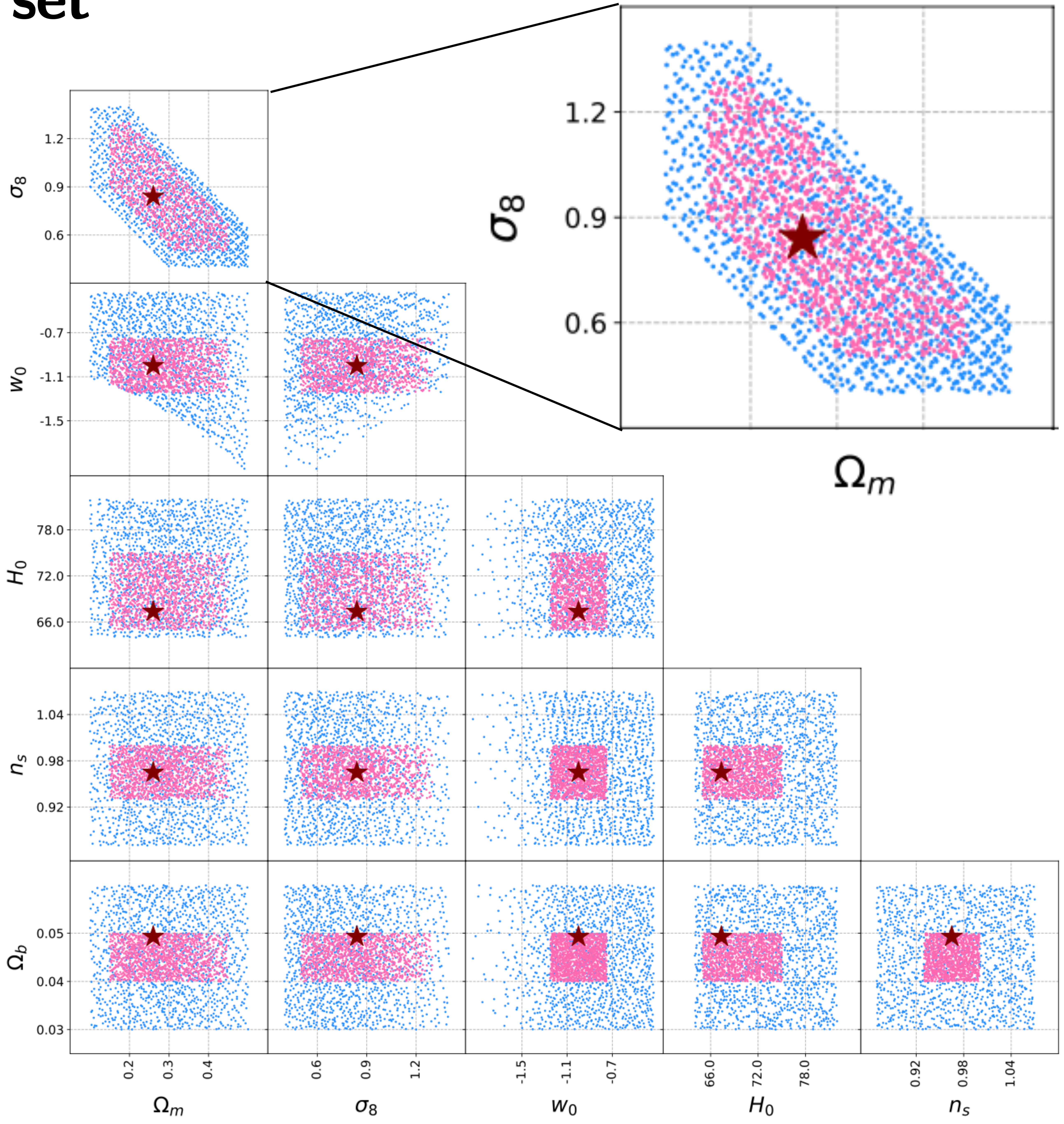


Maksimova et al. 2021 2110.11398

The public CosmoGridV1 simulation set

TK, Fluri, Schneider, Refregier, Stadel 2209.04662

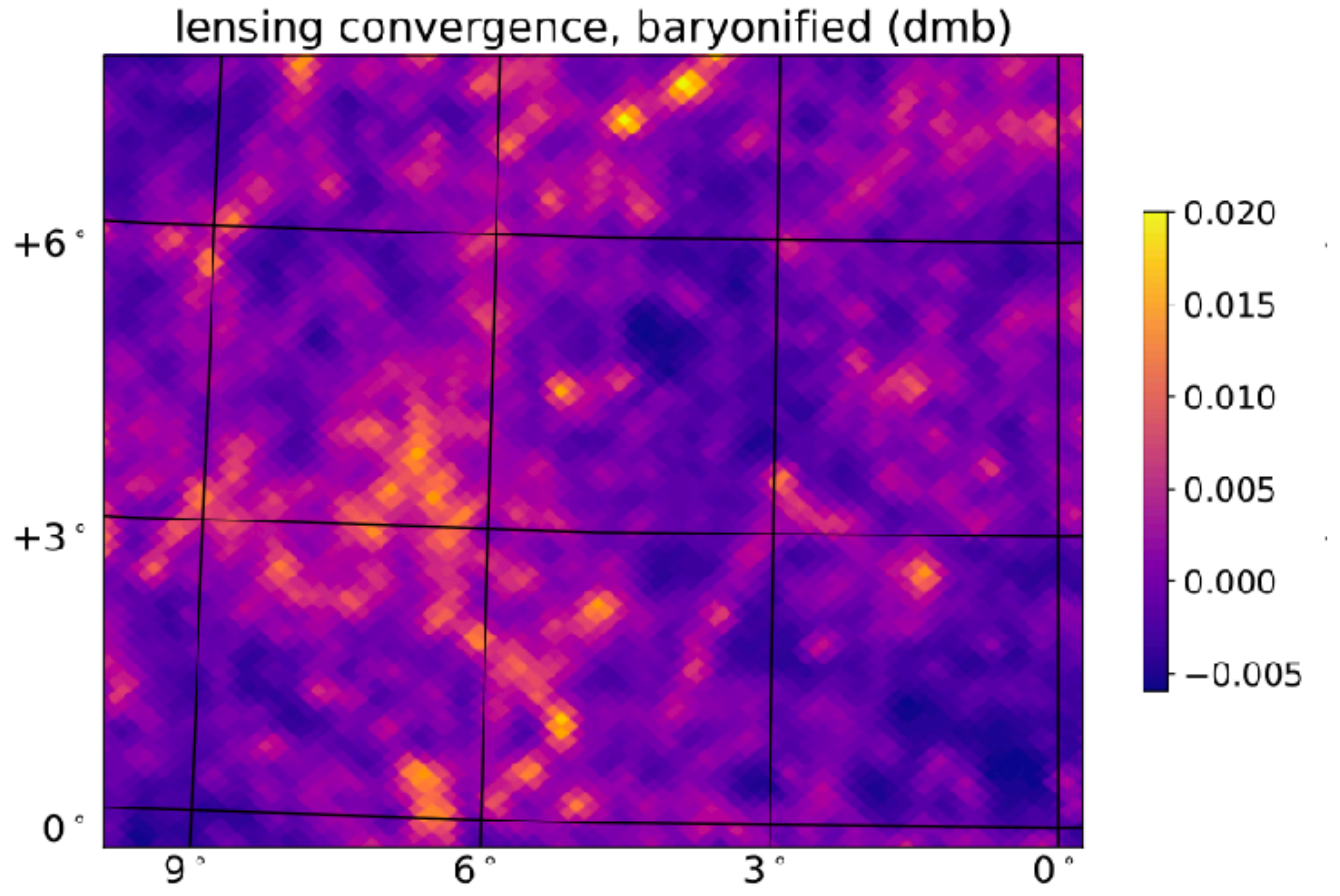
- 2500 unique cosmologies in wCDM $\Omega_m, \sigma_8, w, n_s, H_0, \Omega_b$
- 200 simulations at the fiducial+deltas
- 7 unique initial conditions simulations per cosmology
- Total ~20000 independent N-body simulations
- 900 Mpc/h boxes, 832^3 particles, box replication up to $z = 3.5$
- 70 shells per lightcone, typical shell thickness 60 Mpc/h
- Benchmark simulations available with bigger boxes, more particles and more shells
- Baryonification and NLA intrinsic alignments available
- CSCS Large Production Project, 2020-21
- 300TB of raw simulation data, CC-BY licence

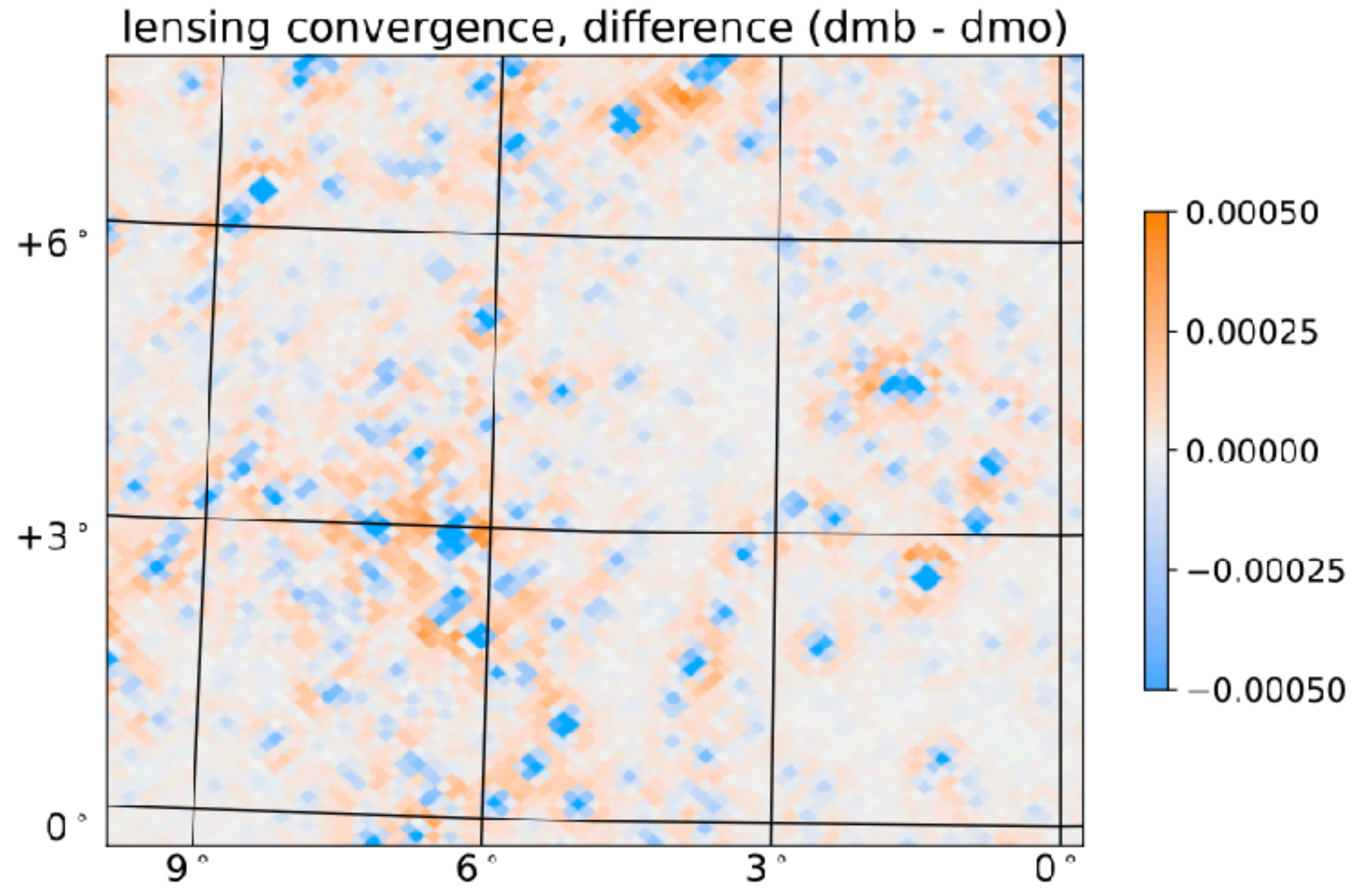


The public CosmoGridV1 simulation set

TK, Fluri, Schneider, Refregier, Stadel 2209.04662

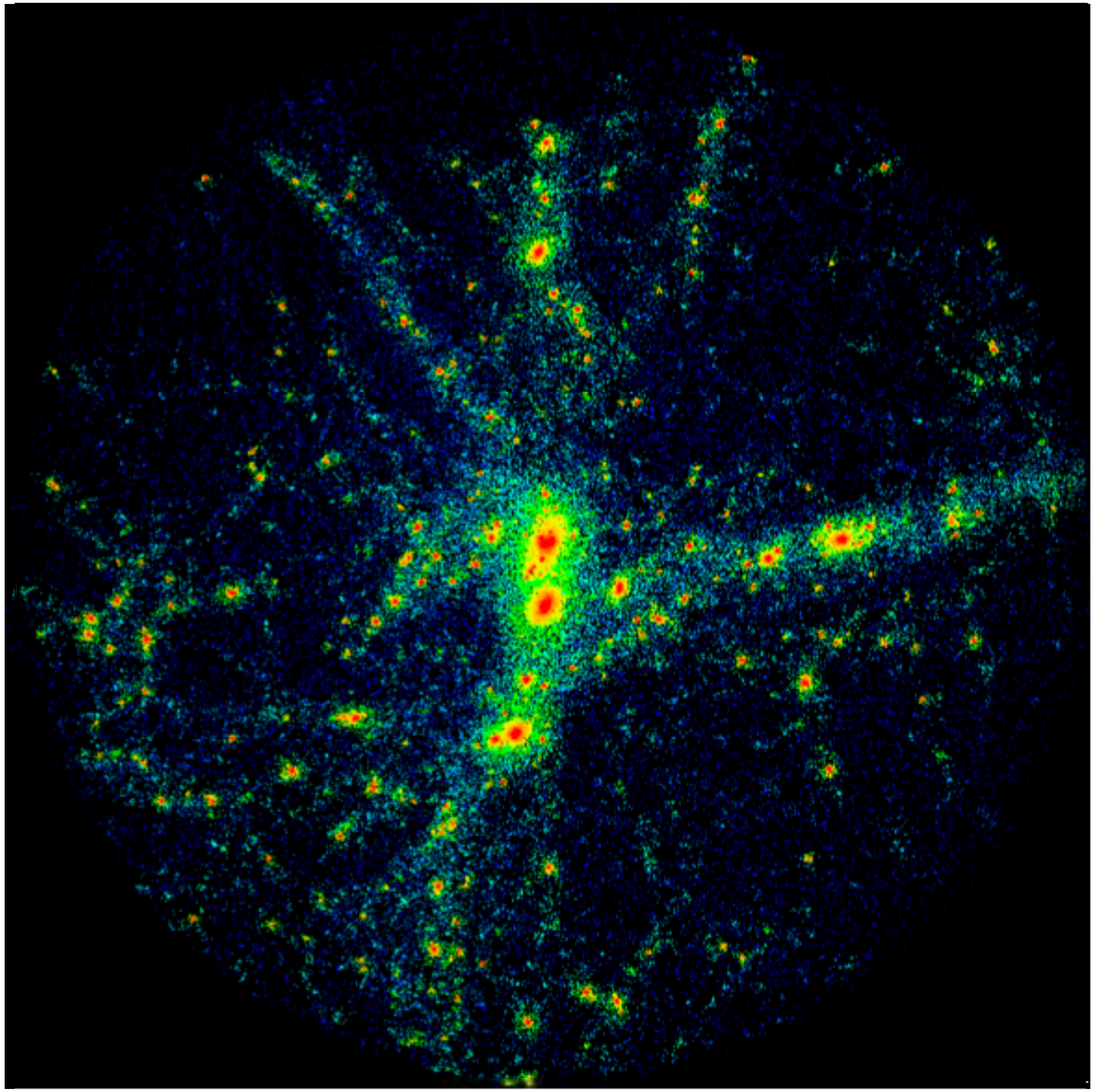
- So far used at Nside=512 for the 1000 dataset (Fluri et al. 2022, 2201.07771)
- Smaller subset used for Nside=1024 for the Stage-4 peaks+non-Gaussian forecast (Zuercher et al. 2022, 2206.01450)
- Baryonification done so far for two parameters controlling the mass dependence of the gas profile: scale and redshift dependence
- Baryonification can be re-done with more parameters
- Possibility of doing a 2048 analysis with more baryon parameters





Baryon Correction Model (BCM), or *baryonification*

- Idea: modify dark matter only simulations so that they have the same distribution as hydrosims
- Use Halo Model framework
- Write a halo profile function including baryon correction terms
- Procedure for each halo:
  - Measure NWF halo parameters and make a parametric profile
  - Modify halo profile according to BCM parameters
  - Find displacement vectors for each particle in a halo
- Advantage: parametric model that can reproduce multiple hydro-sims on the power spectrum level
- Question: does it reproduce hydro-sims on a map level?

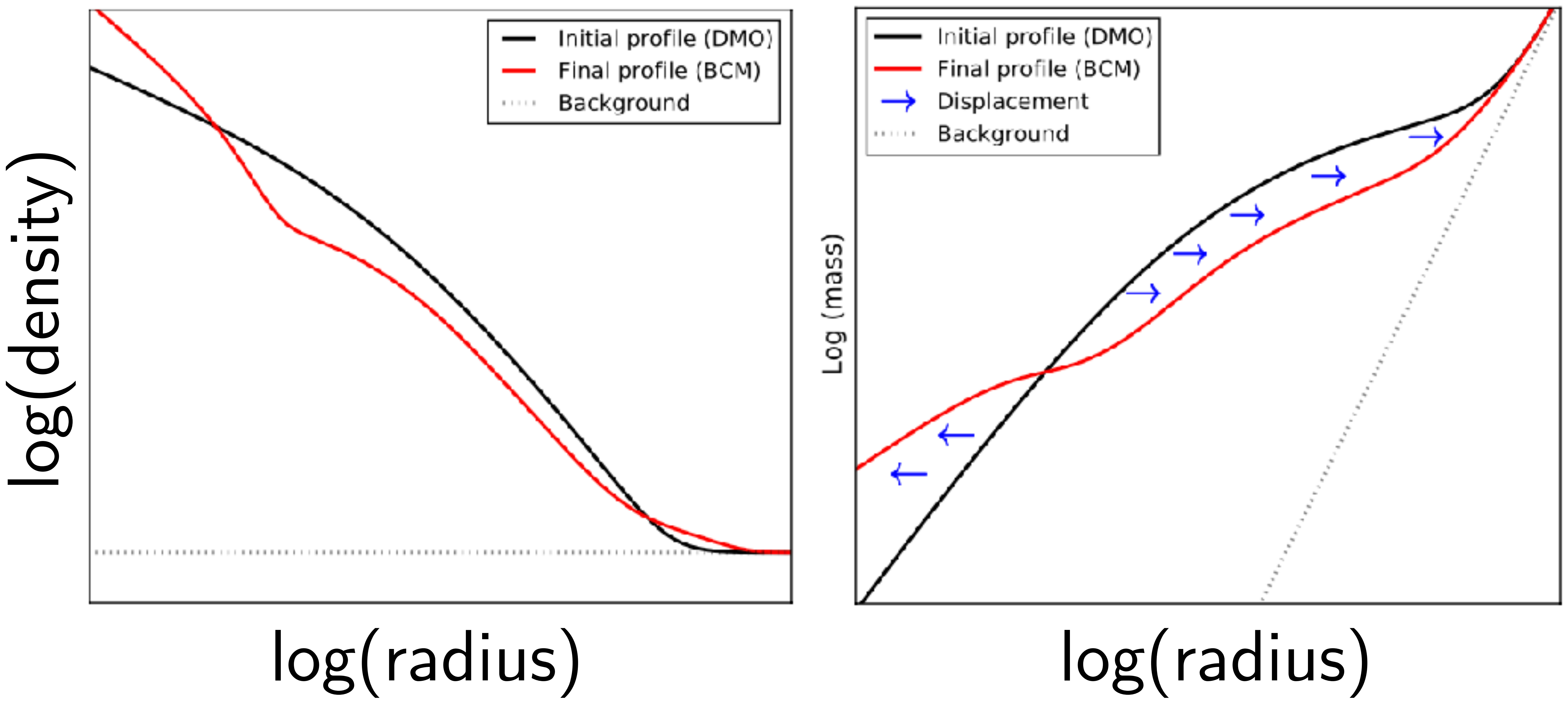


Schneider Teyssier 2015 1510.06034

Baryon Correction Model (BCM), or *baryonification*

- Idea: modify dark matter only simulations so that they have the same distribution as hydrosims
- Use Halo Model framework
- Write a halo profile function including baryon correction terms
- Procedure for each halo:
  - Measure NWF halo parameters and make a parametric profile
  - Modify halo profile according to BCM parameters
  - Find displacement vectors for each particle in a halo
- Advantage: parametric model that can reproduce multiple hydro-sims on the power spectrum level
- Question: does it reproduce hydro-sims on a map level?

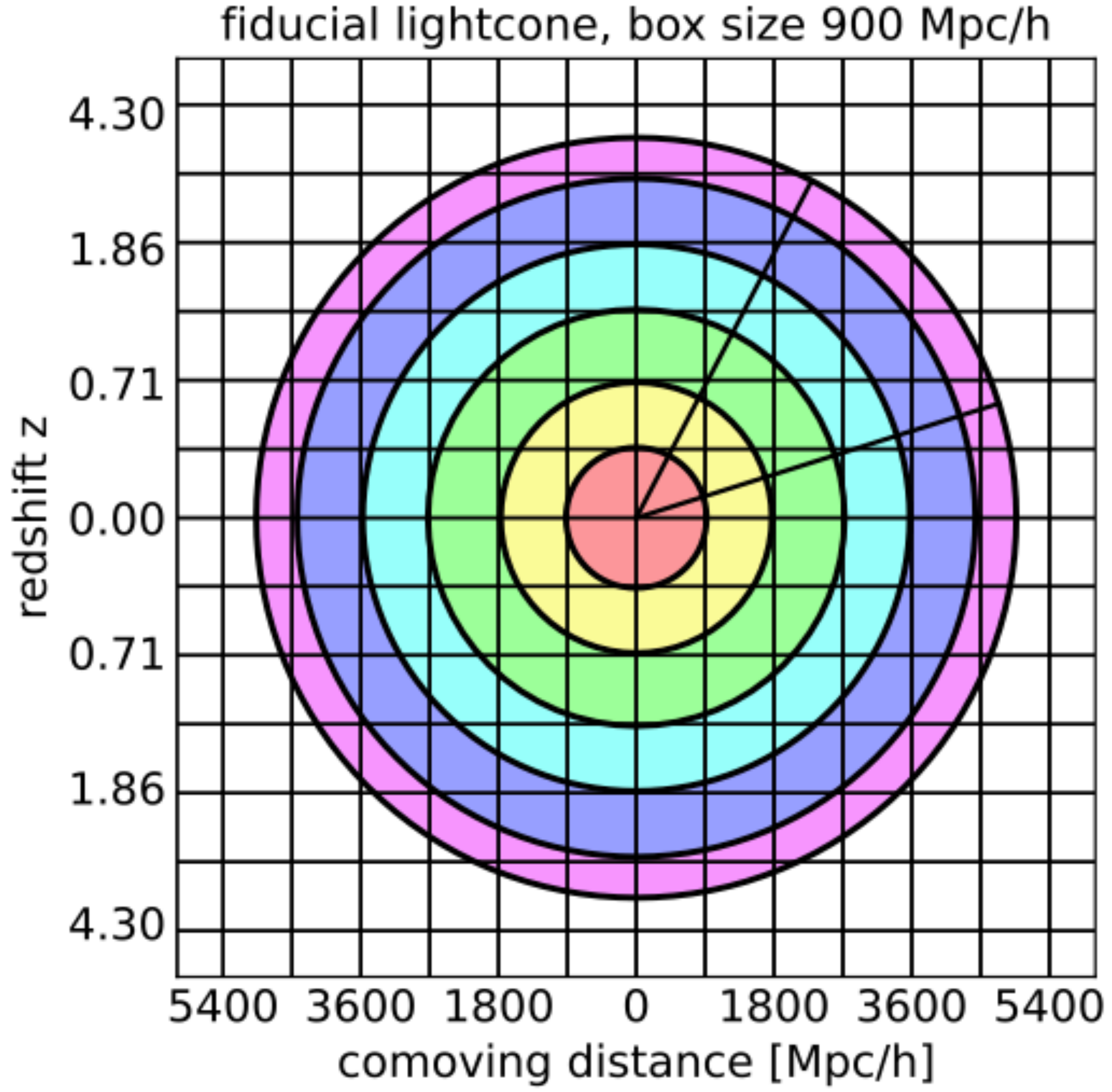
$$M_i = M_{\rm nfw}(r_i),$$
$$M_f = f_{\rm cdm} M_{\rm nfw}(r_i) + f_{\rm cgal} Y_{\rm cgal}(r_f) + f_{\rm bgas} Y_{\rm bgas}(r_f) + f_{\rm egas} Y_{\rm egas}(r_f),$$

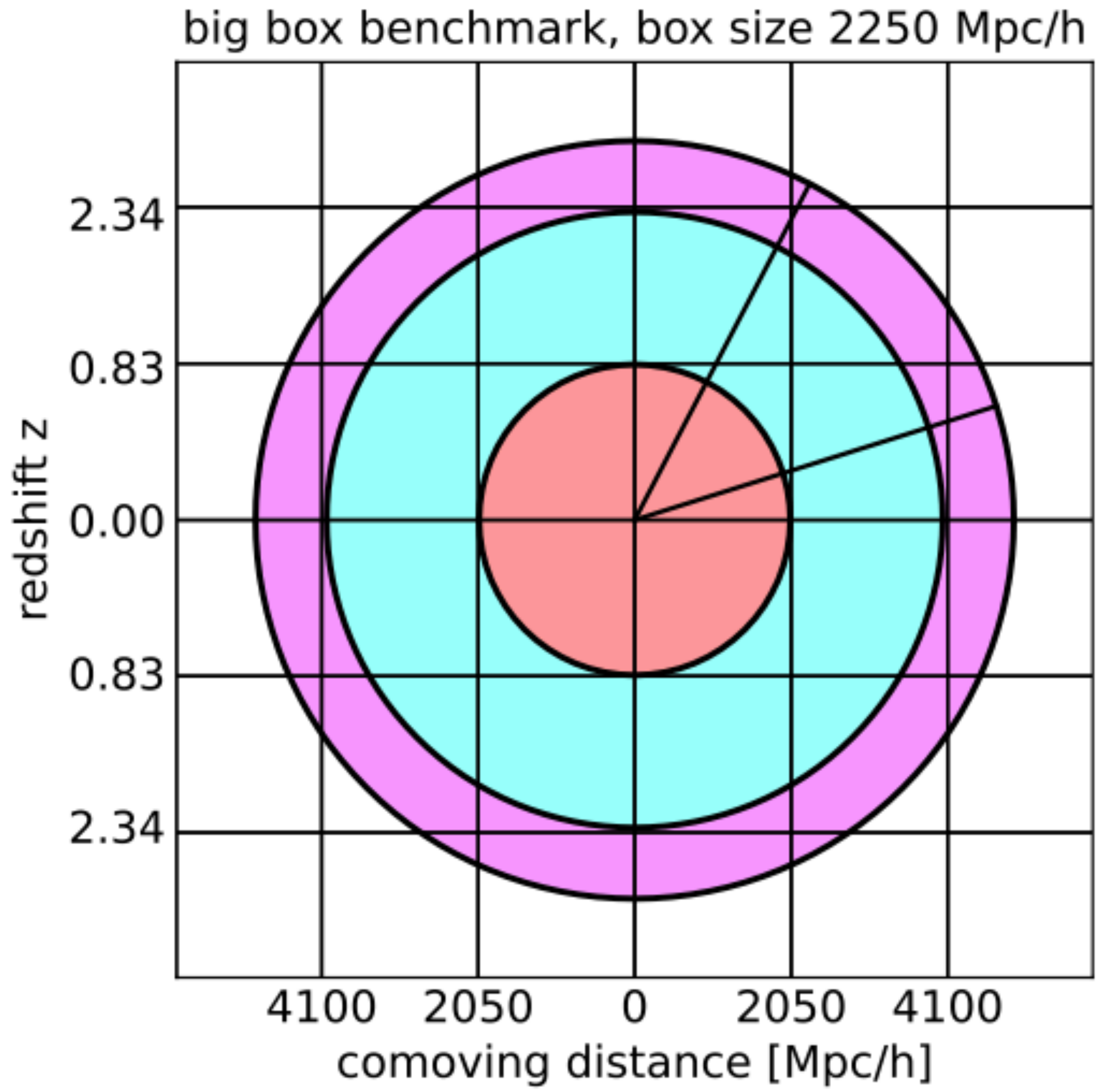


Schneider Teyssier 2015 1510.06034

CosmoGridV1 shell permutations

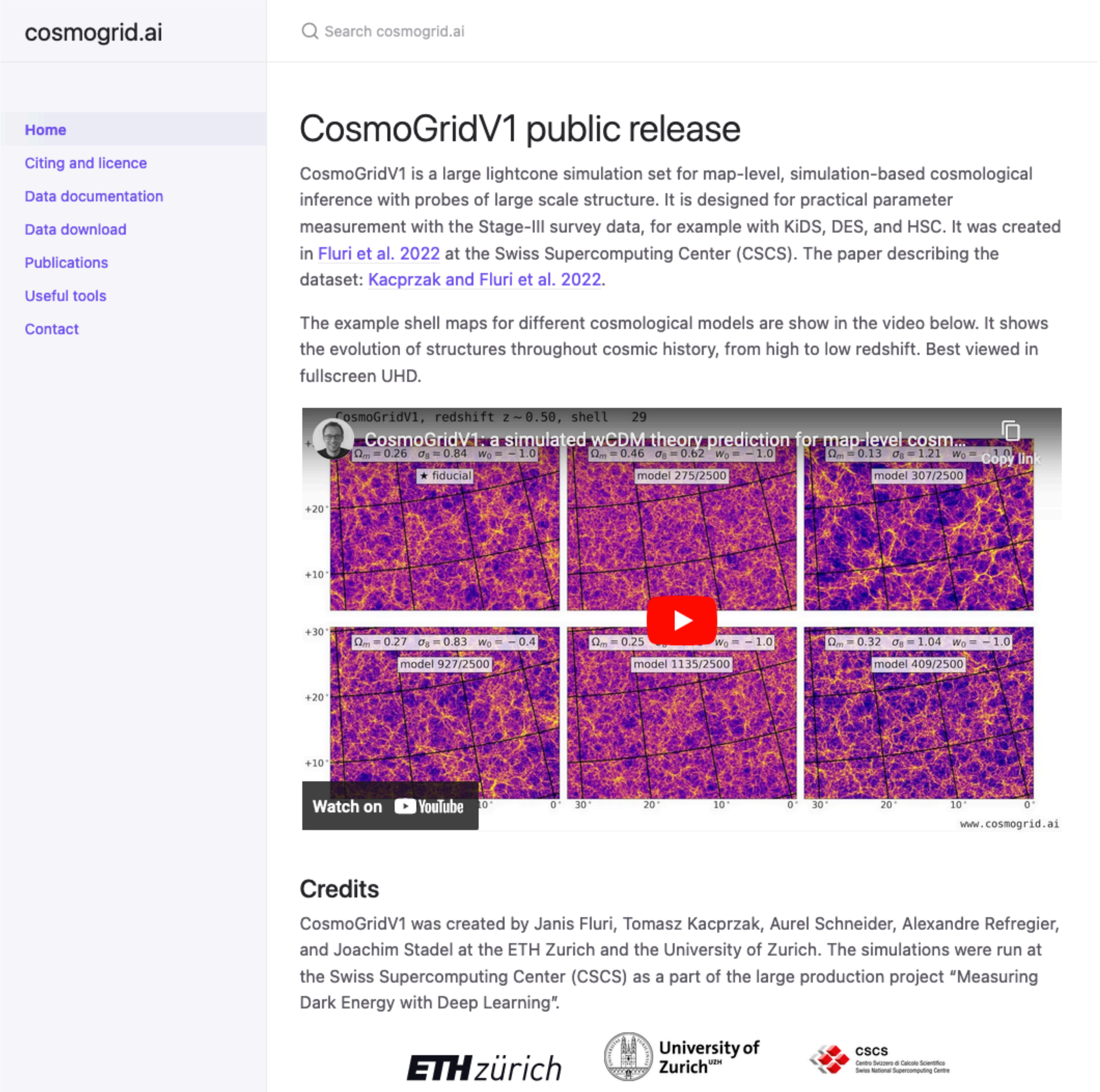
- Box replication schemes do not introduce discontinuities in the lightcone, but can lead to repeated structures along the line of sight
- The line of sight repetition will happen for every 45 deg x N, (N=integer) angle
- To avoid this, we introduced a "shell permutation scheme", where each shell group (of thickness corresponding to the simulation box) comes from independent N-body simulations
- Up to redshift of z=2 we have 5 shells
- This introduces slight discontinuities in the lightcone, but generally <1% of the volume (in 4 Mpc voxels)
- Big box benchmark is composed of 2 shell groups





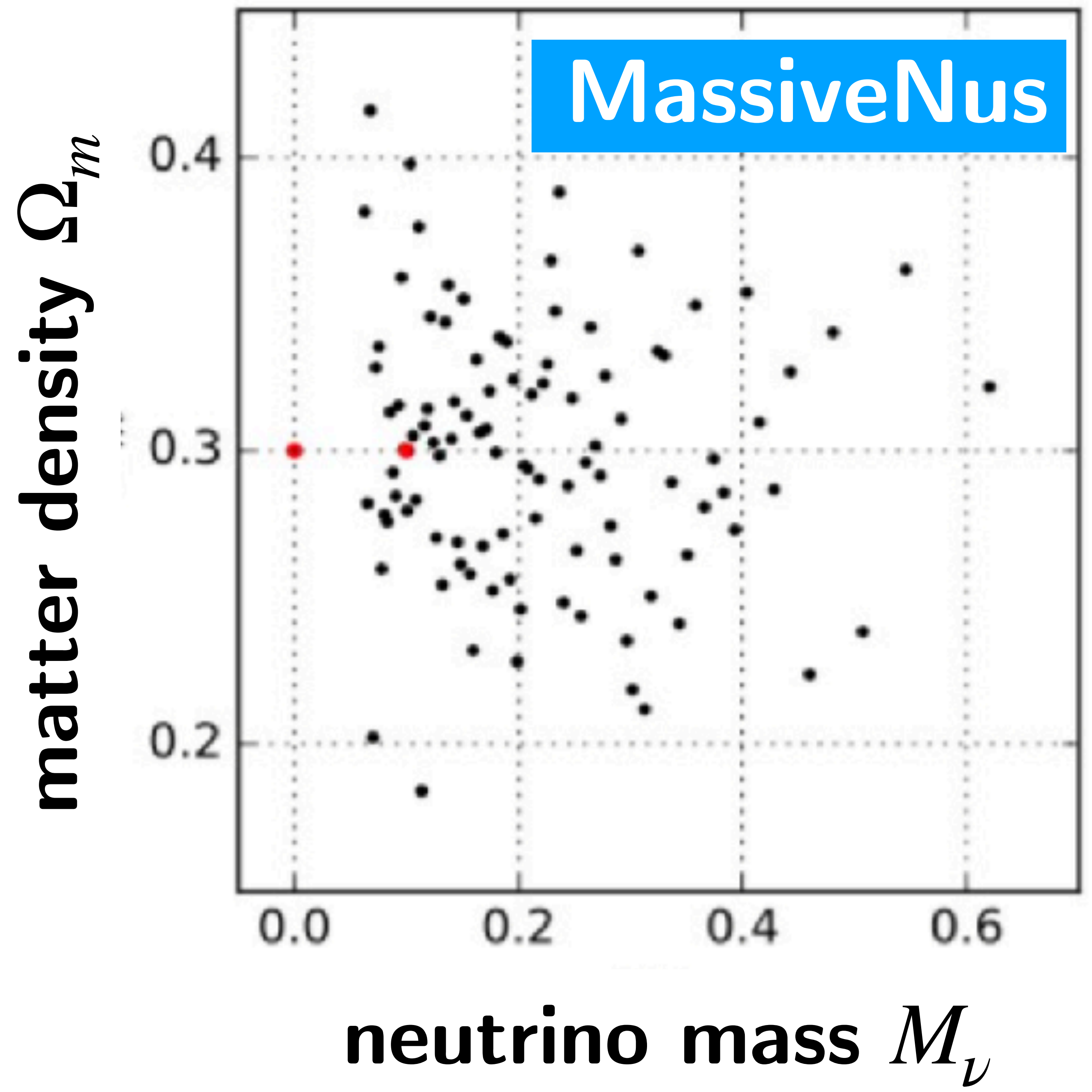
The data

- → Data available at **www.cosmogrid.ai**
- → Full data documentation available
- → Fast transfer with Globus
- → Creative Commons BY licence
- → CosmoGridV1.1 in preparation to include ISW and CMB lensing maps

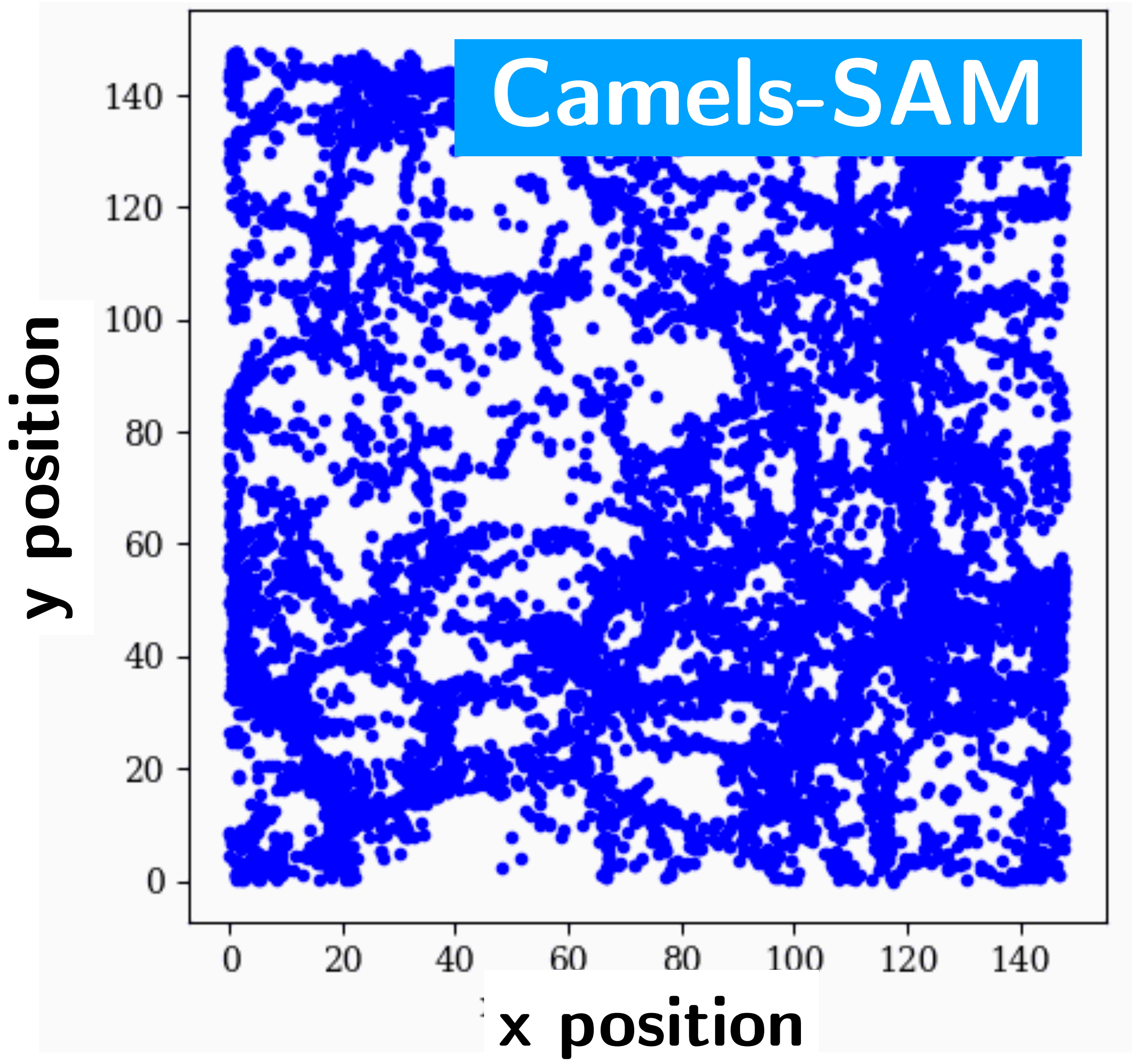


Other amazing publicly available simulation suites

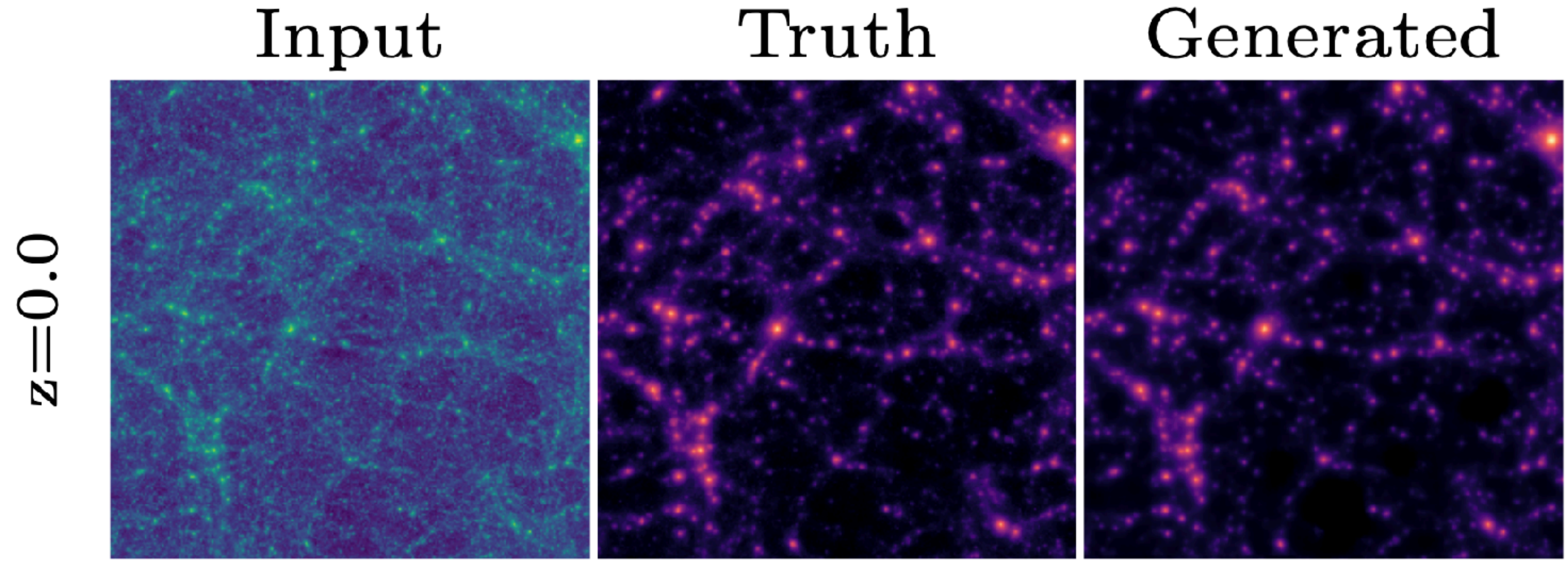
- MassiveNus (Liu et al. 2018 1711.10524)
  - Designed to explore simulations-based inference for constraining cosmological neutrino masses, 101 simulations in Ω_m, A_s, M_ν, high-resolution pencil beams
  - Data available at http://columbialensing.org



- The CAMELS-SAM Suite (Perez et al. 2022 2204.02408)
  - 1000 dark matter simulations at $100\ h^{-1}\ Mpc$, cosmology parameters: Ω_m, σ_8, three semi-analytical baryon feedback parameters: A_{SN1}, A_{SN2}, A_{AGN}
  - Data available at https://camels-sam.readthedocs.io/



“Painting with baryons: augmenting N-body simulations with gas using deep generative models”



Dark matter map

Gas pressure map

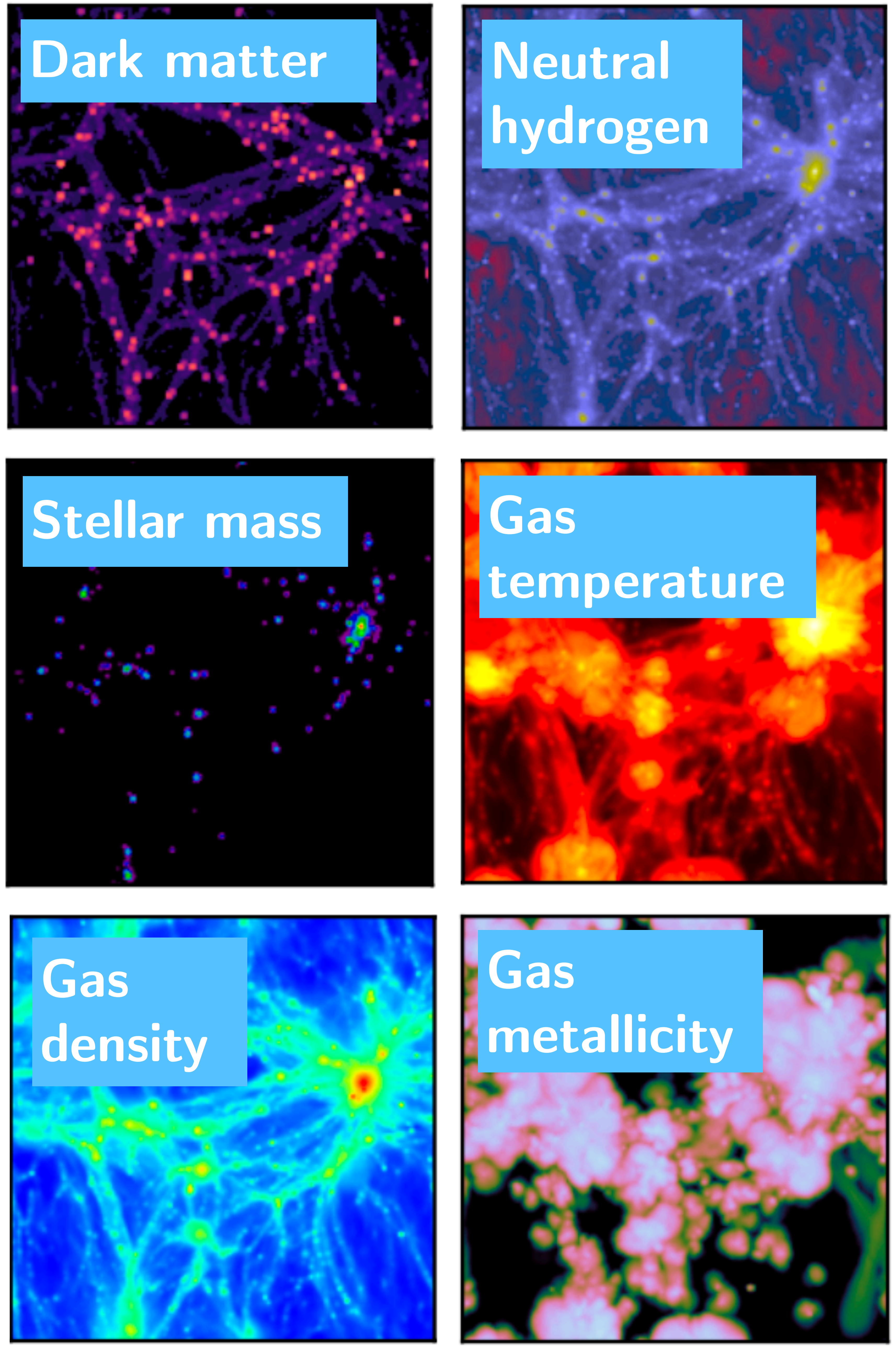
- Using BAHAMAS simulations to create gas pressure maps for the corresponding dark matter maps
- Using Generative Adversarial Nets and Variational Autoencoders to create the gas pressure maps based on the dark matter map only

Tröster et al. 2019, 1903.12173

CAMELS: Cosmology and Astrophysics with MachinE Learning Simulations

General, precise simulations including all of the important effects

- Magneto-hydrodynamic simulations using AREPO and GIZMO, employing baryonic subgrid physics as IllustrisTNG and SIMBA
- Dataset used to demonstrate the possibilities of machine learning to understand astrophysics and cosmology jointly
- 4233 small boxes $(25\ h^{-1}\ Mpc)^3$ spanning the wCDM cosmological model and different AGN feedback models
- 20+ methods papers for various problems in the last 2 years
- Data publicly available at https://camels.readthedocs.io

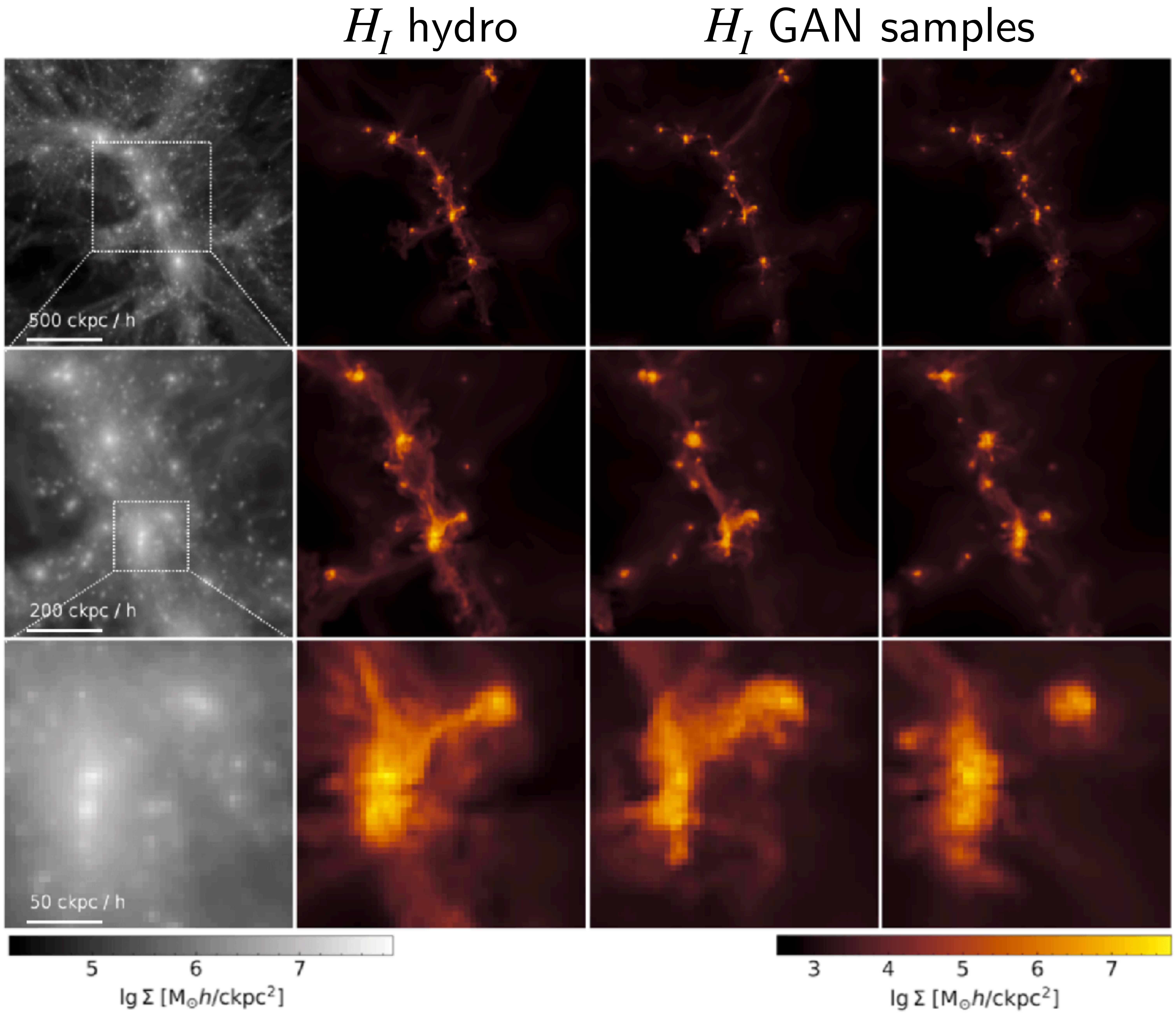


Villaescusa-Navarro et al. 2022 2201.01300

From EMBER to FIRE: predicting high-res baryon fields from matter-only

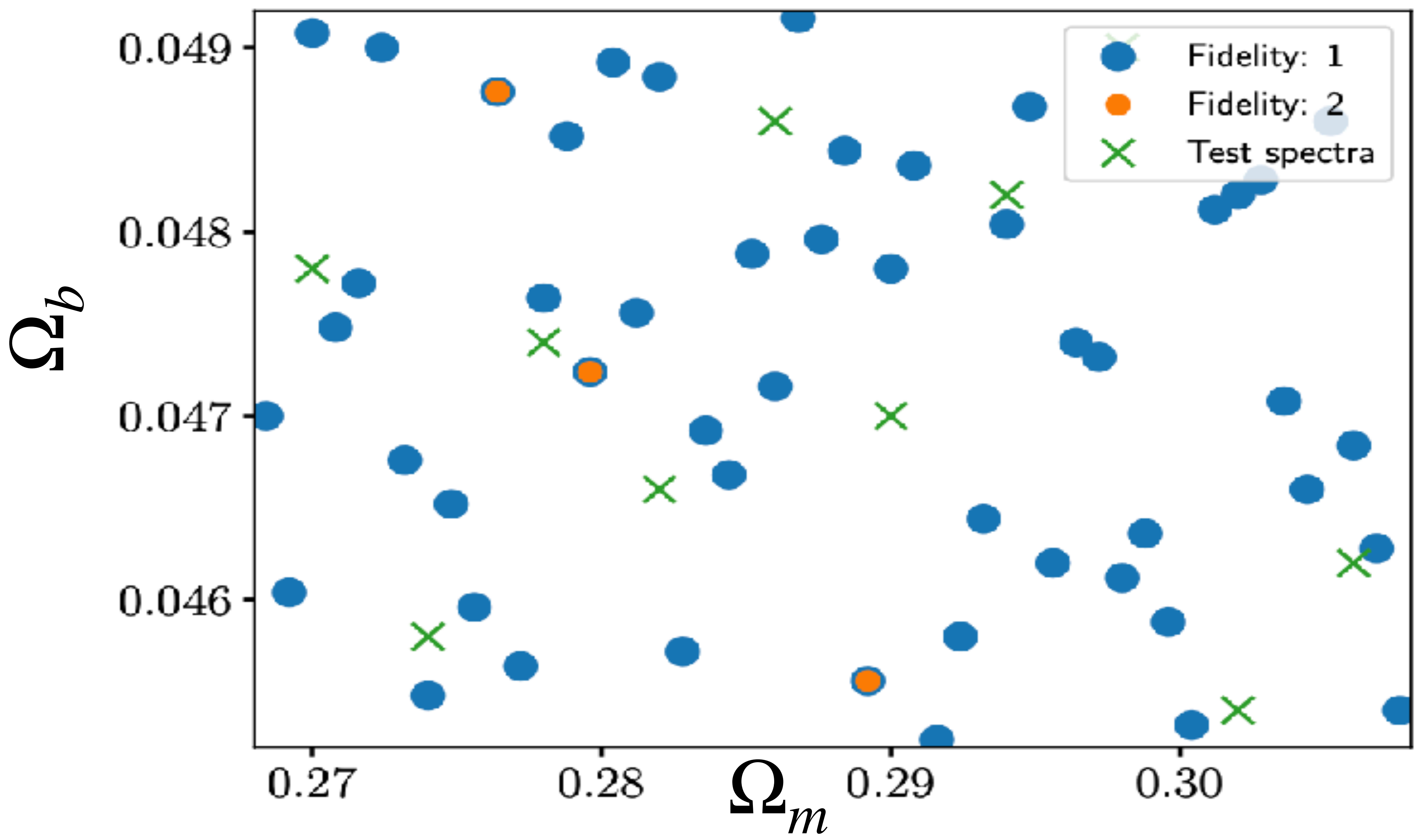
Bernardini et al. 2022 2110.11970

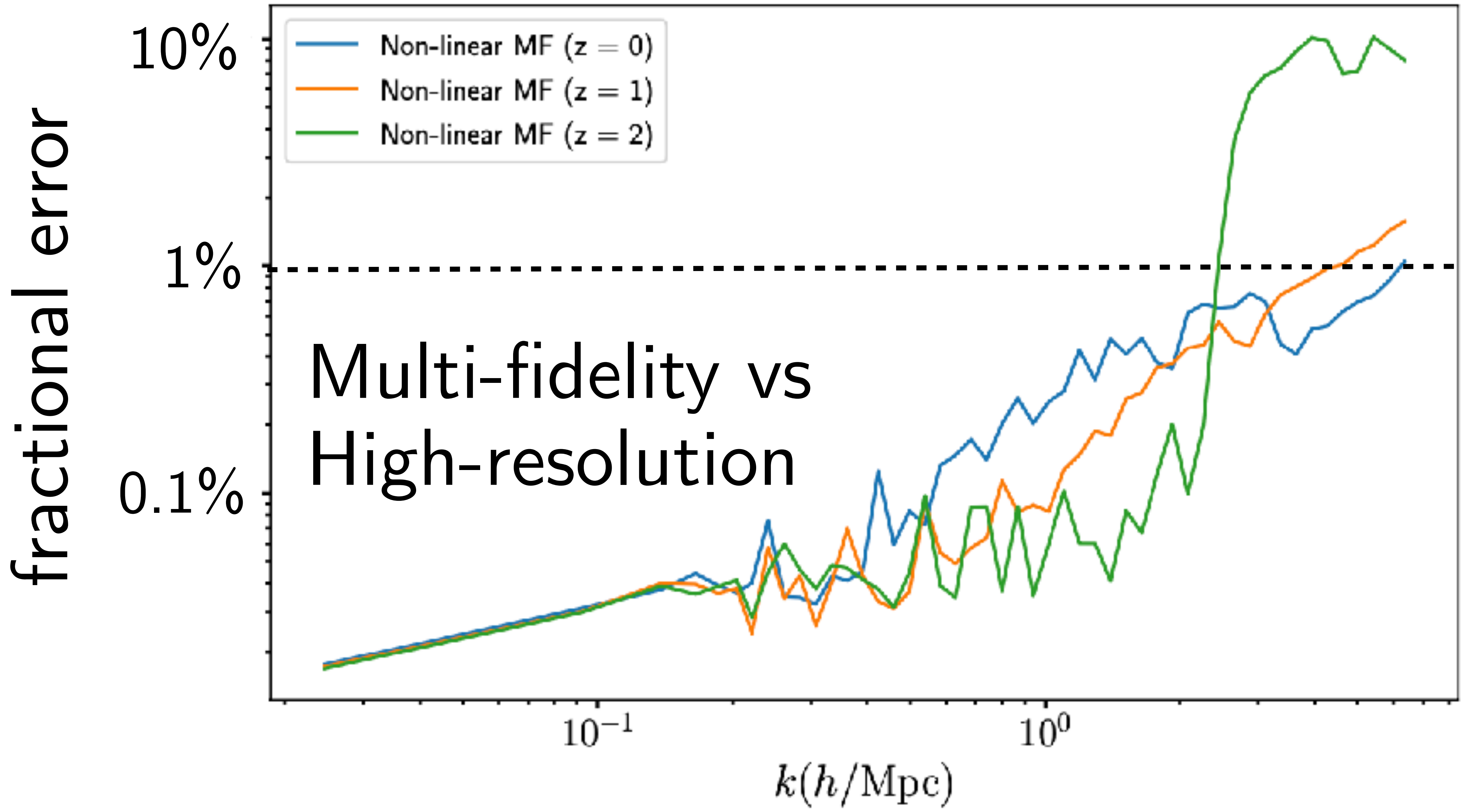
- Improved accuracy of baryon painting, using WGANs
- Zoom-in hydro-simulations FIRE representing large range of scales
- Learning the H_I power specra with 10% accuracy at $\sim kpc$ scales
- Multi-scale application, hydro at 15 h^{-1} Mpc, dark matter -only 100 $h^{-1}Mpc$



Data Science Challenge: simulations-based inference with multi-fidelity simulations

- Goal: build simulations for AI-based map-level inference that include realistic small scale effects
- Problem: small hydrosimulations are expensive, cosmological-scale hydro-sims are out of our current reach
- Idea: combine expensive small scale simulations with cheap large scale simulation
- This technique is called **multi-fidelity** and has recently been demonstrated on the level of summary statistics (power spectra)
- **Challenge:** build AI systems that can learn from small-scale simulations and correctly "augment" large scale simulations and produce consistent probe maps for SBI
- Possible answers: foundation models





MF Ho, Bird, Shelton 2021 2105.01081

Where are we with SBI/AI?

→ **Present gains of SBI/AI analysis:** increasing the precision of results, typically 30-50% for lensing for current scale limit, recently gains of $2-6\times$ for clustering

→ **Upcoming gains of SBI/AI analysis:** constraining more complex astrophysics and make our cosmology measurements more robust

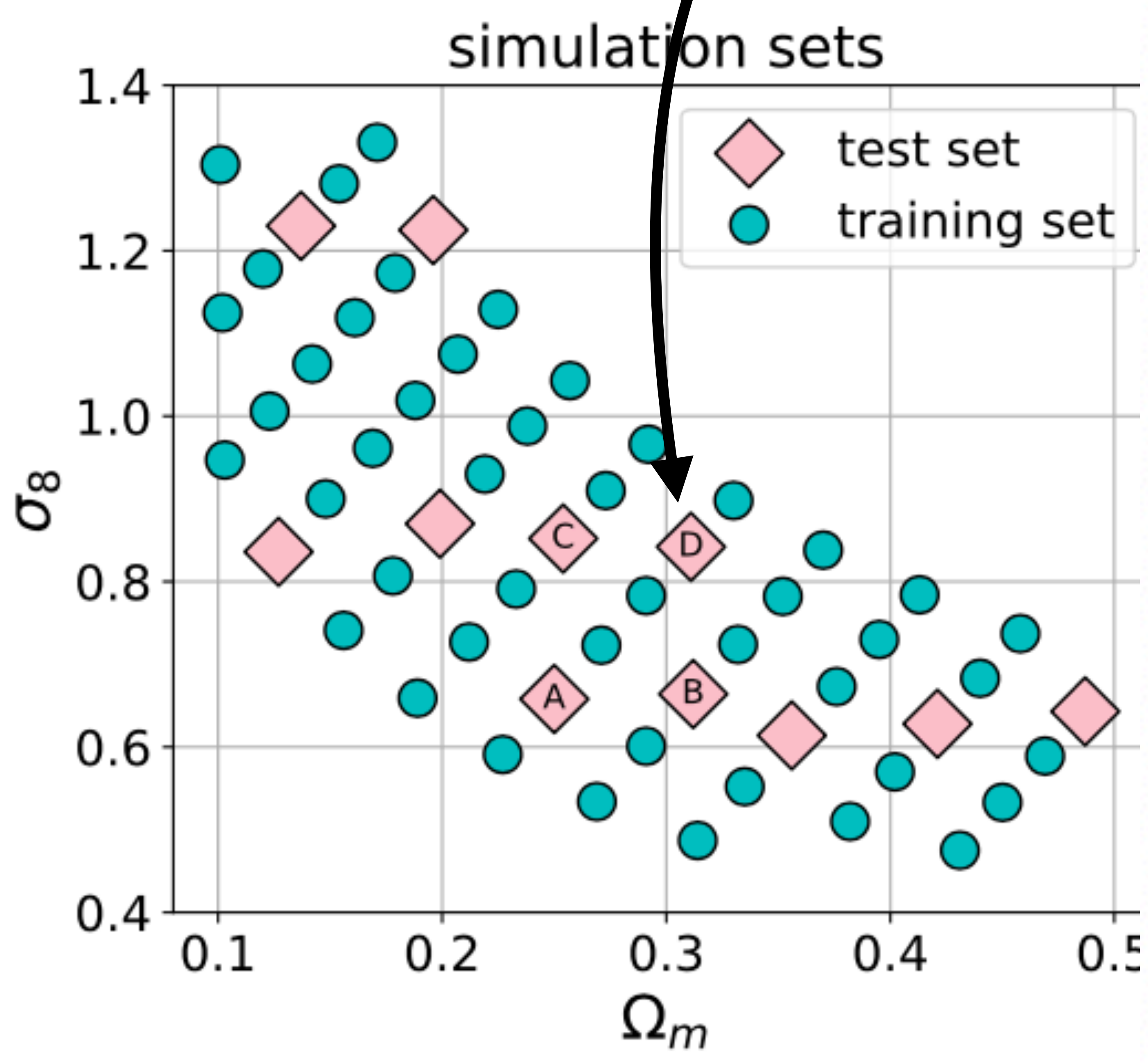
→ **Future gains of SBI/AI analysis:** increased information sensitivity enables constraining models beyond wCDM, testing modified DE/DM/Gravity theories

Backup slides

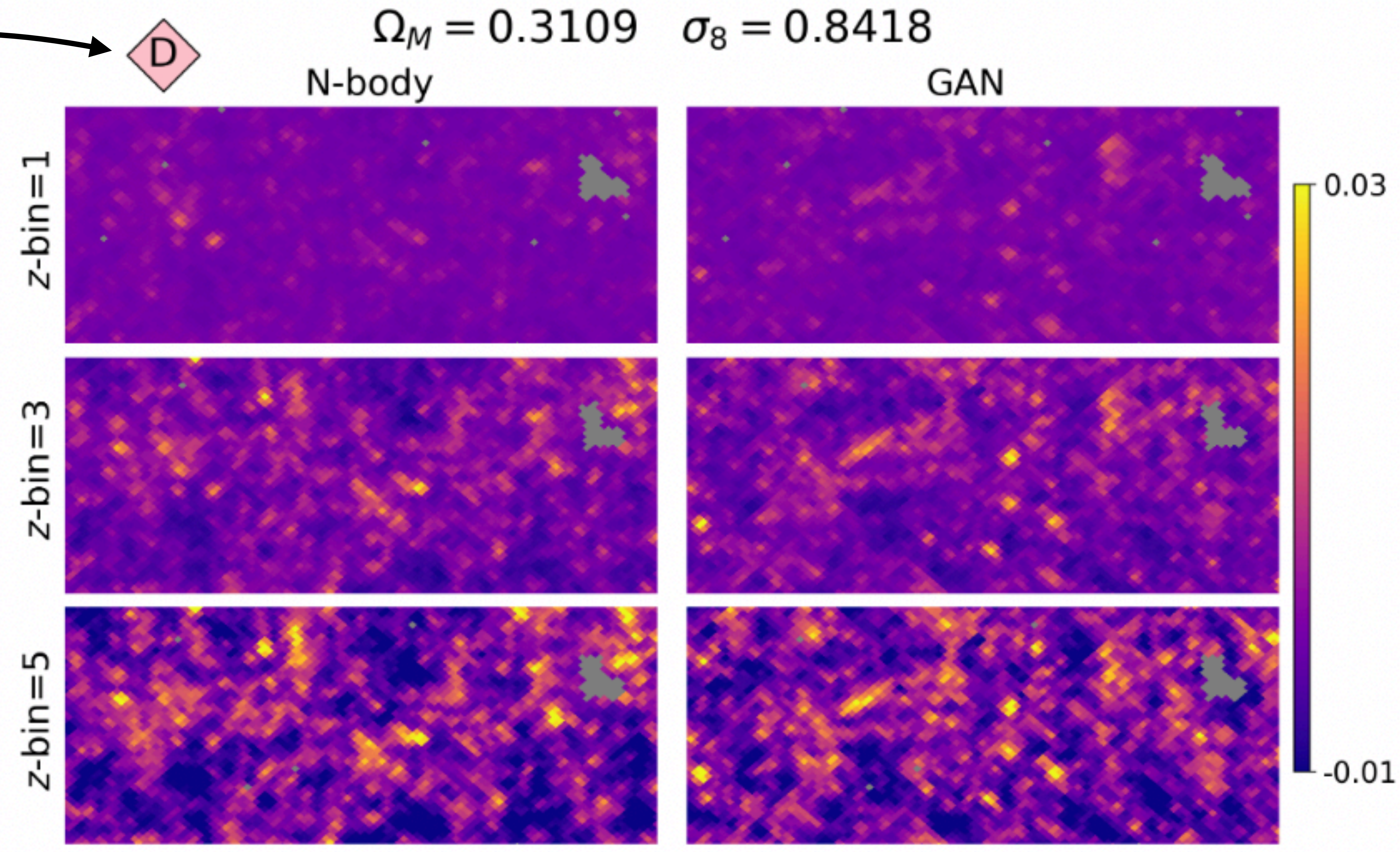
List of our papers on AI in Cosmology

- **Cosmology from Galaxy Redshift Surveys with PointNet**, S. Anagnostidis, A. Thomsen, T. Kacprzak, T. Tröster, L. Biggio, A. Refregier, T. Hofmann, NeurIPS 2022, 2211.12346
- **CosmoGridV1: a simulated wCDM theory prediction for map-level cosmological inference,** T. Kacprzak, J. Fluri, A. Schneider, A Refregier, J Stadel, JCAP 2023, 2, 50, 2209.04662
- **DeepLSS: breaking parameter degeneracies in large scale structure with deep learning of combined probes,** T. Kacprzak, J. Fluri, PhysRevX, 2022, 2, 031029, 2203.09616,
- **A Full wCDM Analysis of KiDS-1000 Weak Lensing Maps using Deep Learning,** J. Fluri, T. Kacprzak, A. Lucchi, A. Schneider, A. Refregier, T. Hofmann PhysRevD, 2022, 105, 8, 083518, 2201.07771,
- **Cosmological constraints with deep learning from KiDS-450 weak lensing maps**, J. Fluri, T. Kacprzak, A. Lucchi, A. Refregier, A. Amara, T. Hofmann, A. Schneider PhysRevD, 2019, 100, 6, 1906.03156
- **DeepSphere: Efficient spherical convolutional neural network with HEALPix sampling for cosmology**, N. Perraudin, M. Defferrard, T. Kacprzak, R. Sgier, Astronomy and Computing, 2019, 27, 130, 1810.12186,
- **Cosmological constraints from noisy convergence maps through deep learning**, J. Fluri, T. Kacprzak, A. Lucchi, A. Refregier, A. Amara, T. Hofmann PhysRevD, 2018, Vol. 98, 12, 1807.08732
- **Cosmological model discrimination with deep learning**, J. Schmelzle, A. Lucchi, T. Kacprzak, A. Amara, R. Sgier, A. Refregier, T. Hofman, 1707.05167
- **Fast Cosmic Web Simulations with Generative Adversarial Networks**, A. C. Rodriguez, T. Kacprzak, A. Lucchi, A. Amara, R. Sgier, J.Fluri, T. Hofmann, A. Refregier, CompAst, 2018, 5, 1, 4, 11, 1801.09070,
- **Cosmological N-body simulations: a challenge for scalable generative models**, N. Perraudin, A. Srivastava, Ankit, A. Lucchi, T. Kacprzak, T. Hofmann, A. Réfrégier, CompAst 2019, 6, 1, 1908.05519
- **A tomographic spherical mass map emulator of the KiDS-1000 survey using conditional GANs**, T. W. H. Yiu, J. Fluri, T. Kacprzak JCAP, 2022, 12, 013, 2112.12741
- **Fast Point Spread Function Modeling with Deep Learning**, J. Herbel, T. Kacprzak, A. Amara, A. Refregier, A. Lucchi, JCAP, 2018, 07, 54, 1801.07615

KiDS-1000 mass map emulator



Grid of simulations as train/test set

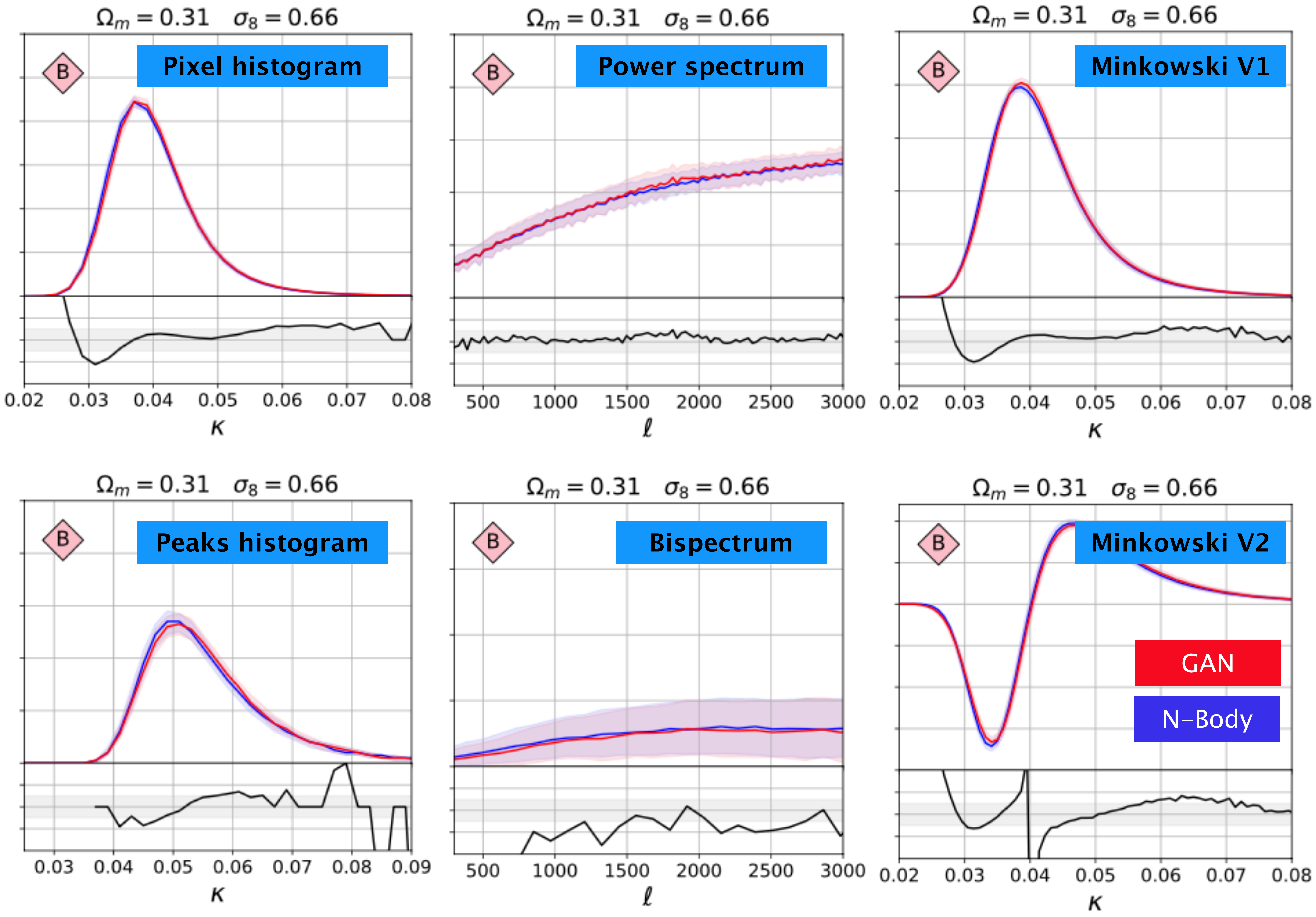


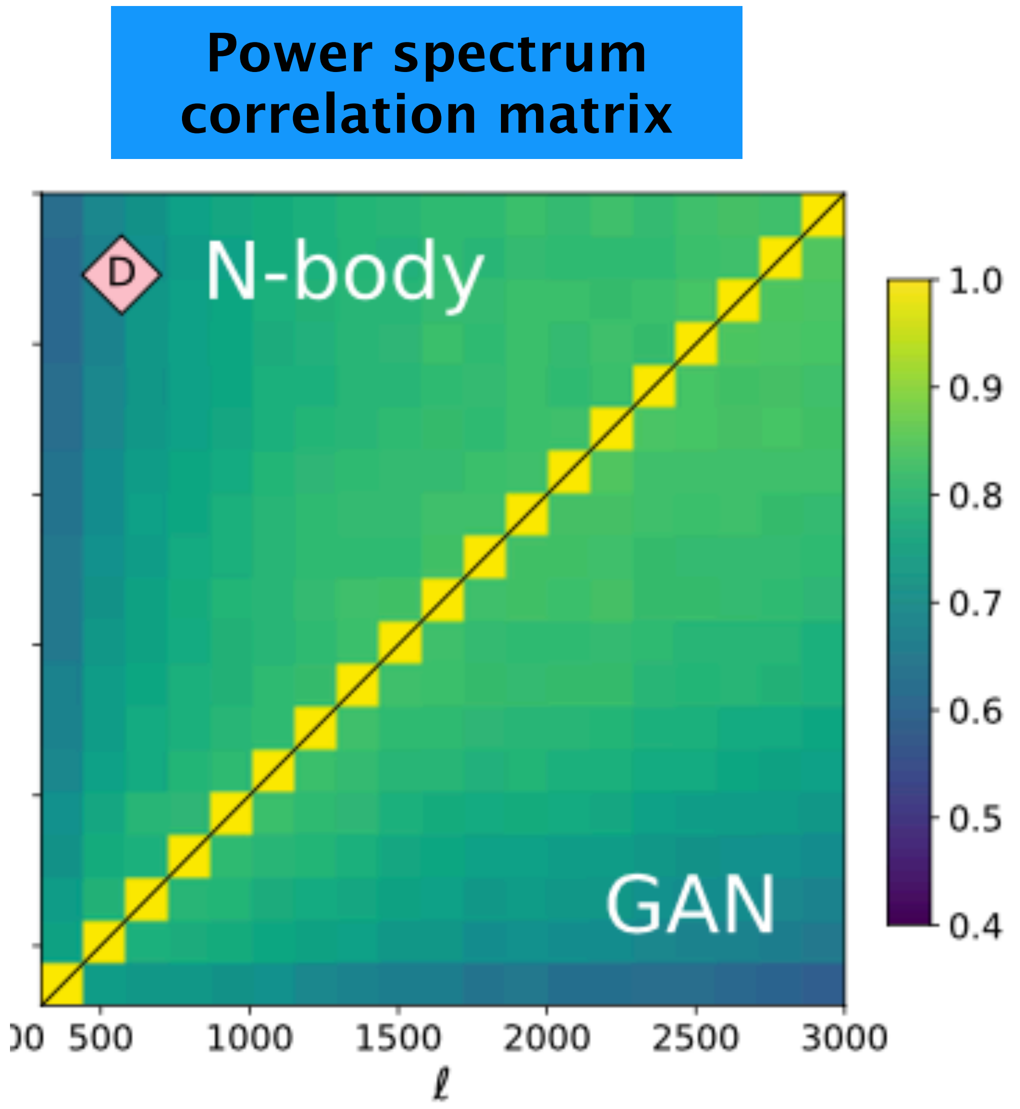
Visual comparison between original N-body and GAN maps

Very fast generator publicly available: https://tfhub.dev/cosmo-group-ethz/models/kids-cgan/1

Emulation of cosmological mass maps with conditional GANs

Perraudin, TK, et al. 2020, 2004.08139

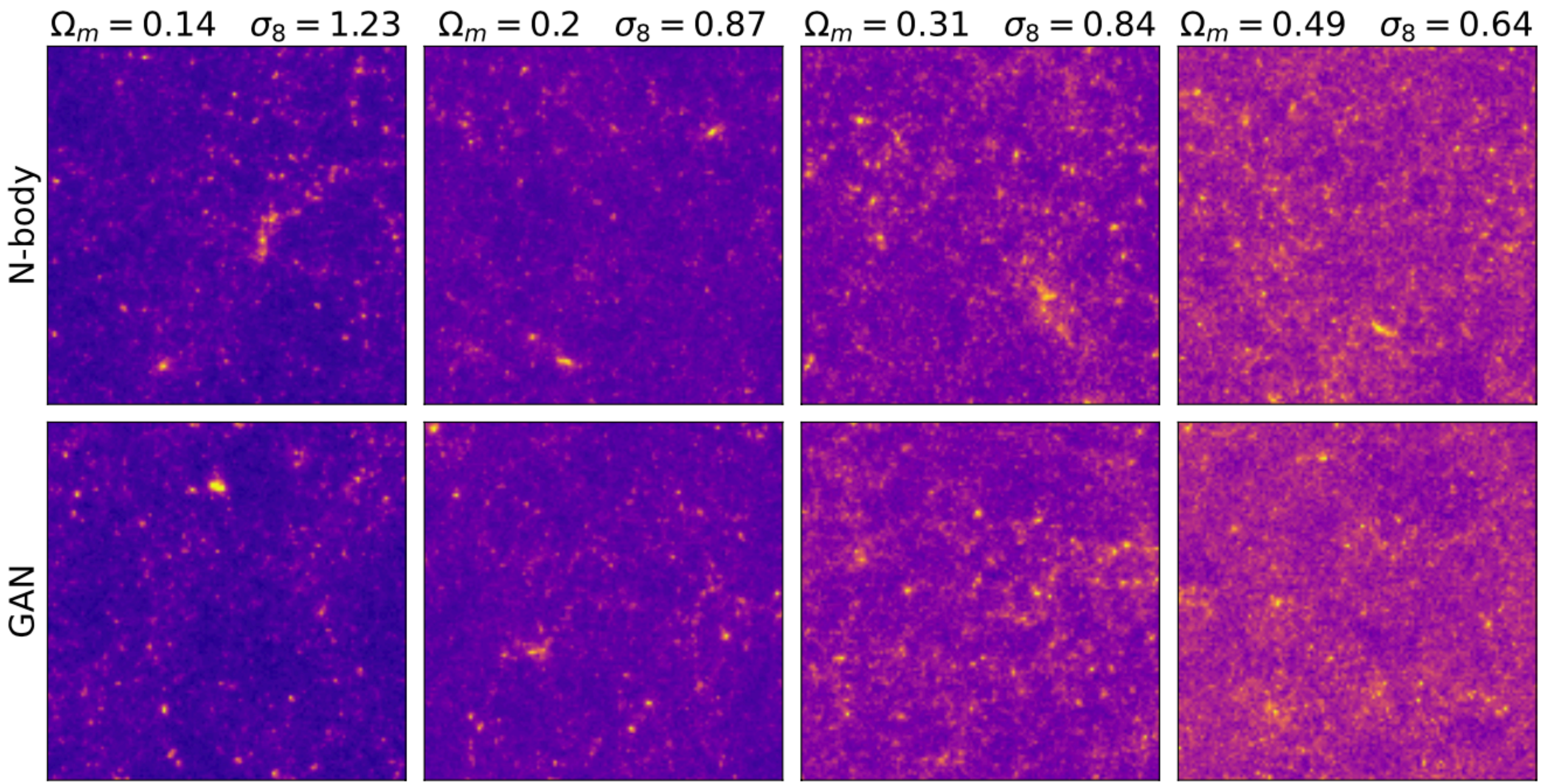




Quantitative comparison: a very good match of summary statistics

Emulation of cosmological mass maps with conditional GANs

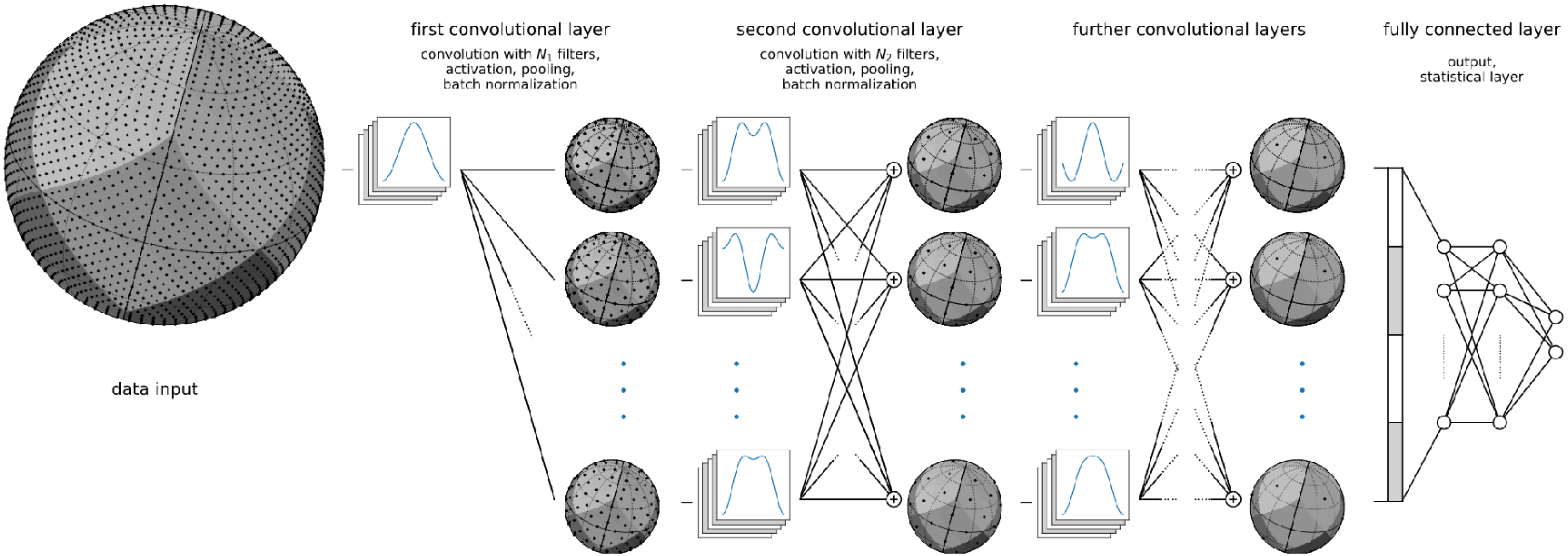
Perraudin, TK, et al. 2020, 2004.08139



Comparison between the N-body and GAN-generated mass maps for varying cosmological parameters

Deep learning on the sphere: a tool for large area sky maps

Perraudin, TK, et al. 1810.12186



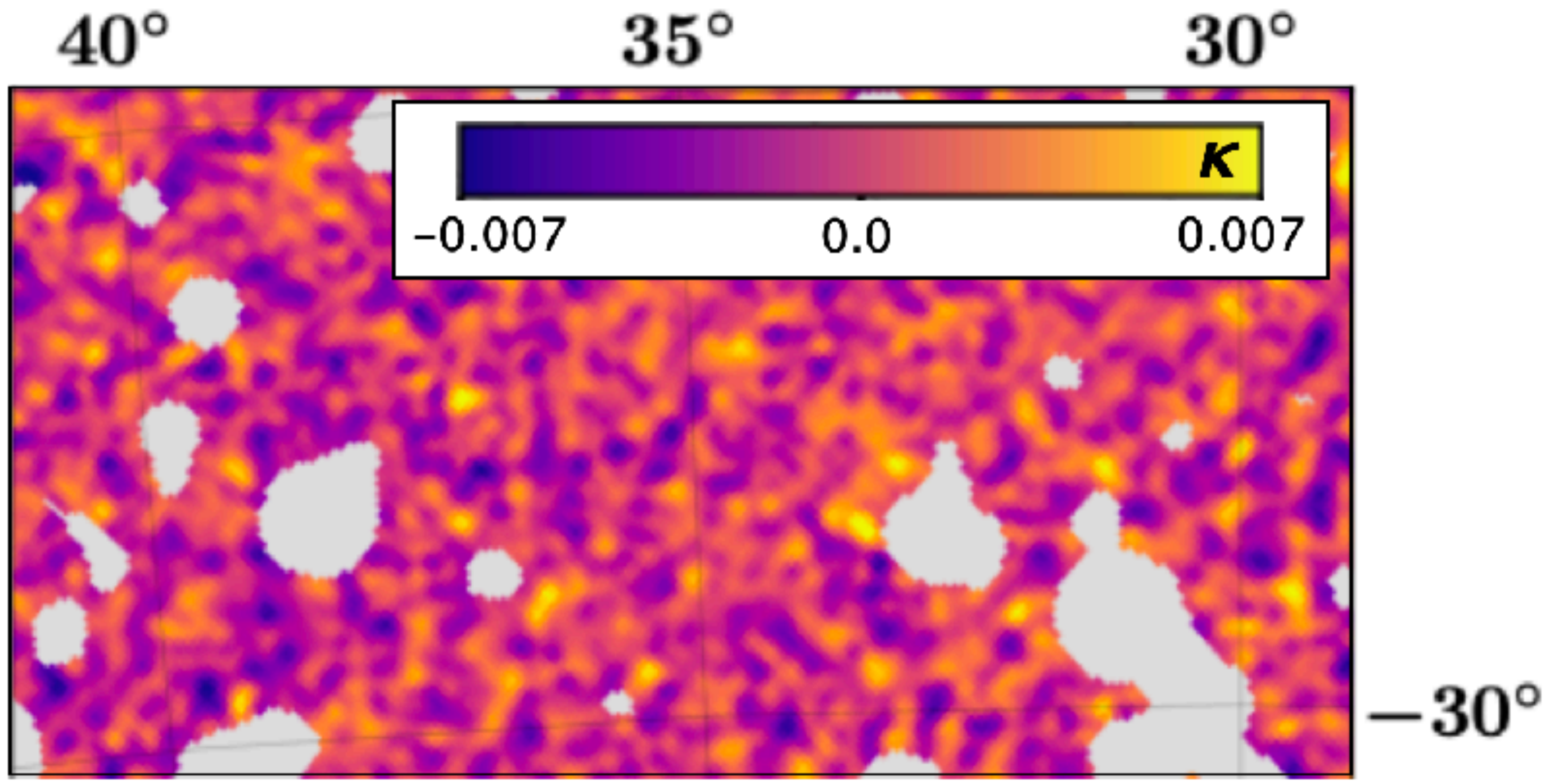
- Various CNN/Transformer architectures on the sphere with Healpix sampling
- Using graph representation, useful for analysis of data on part of the sphere
- One of the fastest sphere convolutions available (but slightly approximate)
- Used by other domains: weather, geo-sciences
- Tensorflow and PyTorch interfaces

github.com/deepsphere

ETH zürich

SDSC

LSS observations



Zuercher, +TK, +DES, 2110.10135

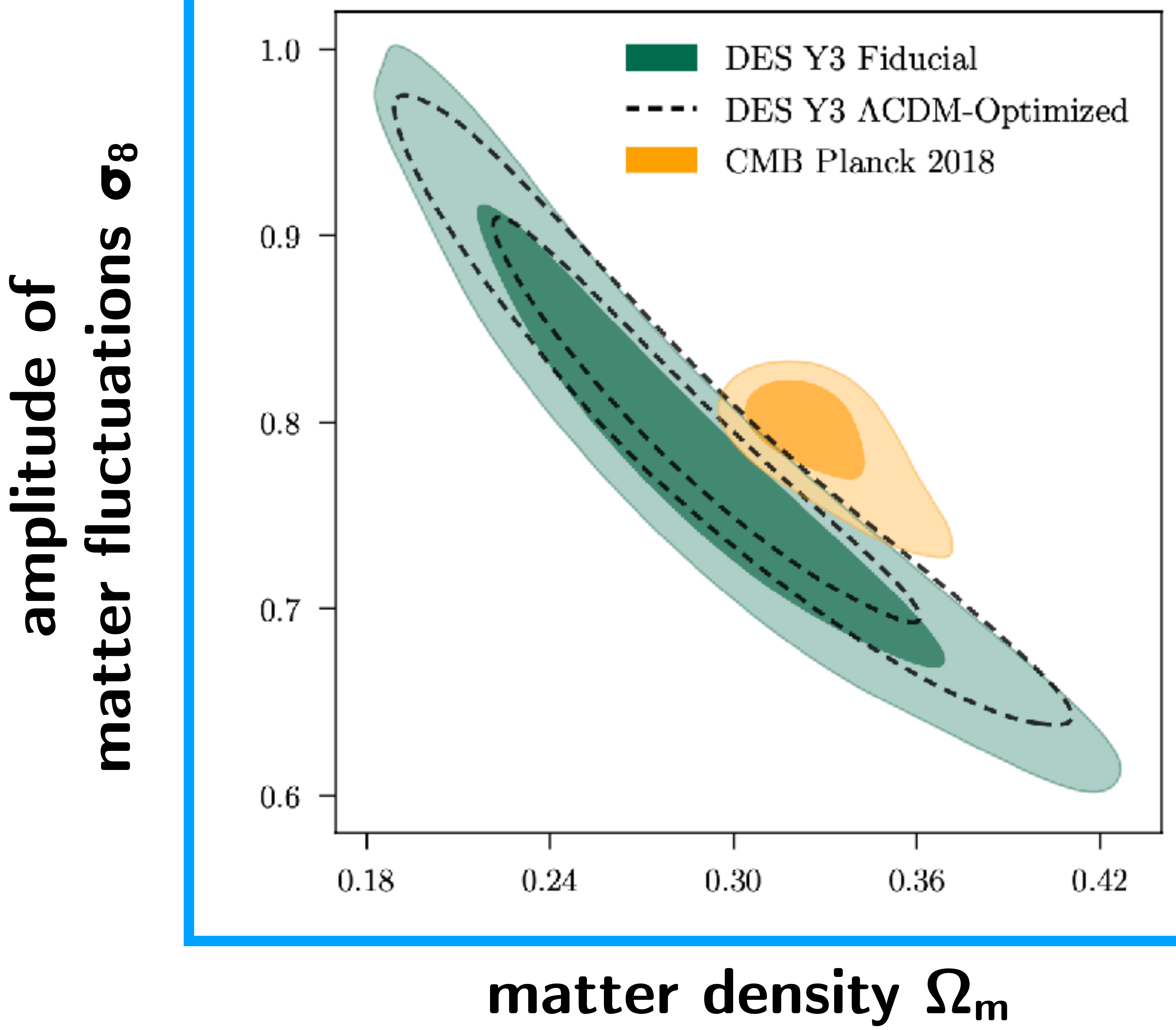
Cosmological parameter inference

Assume a model with parameters

Assume priors on parameters

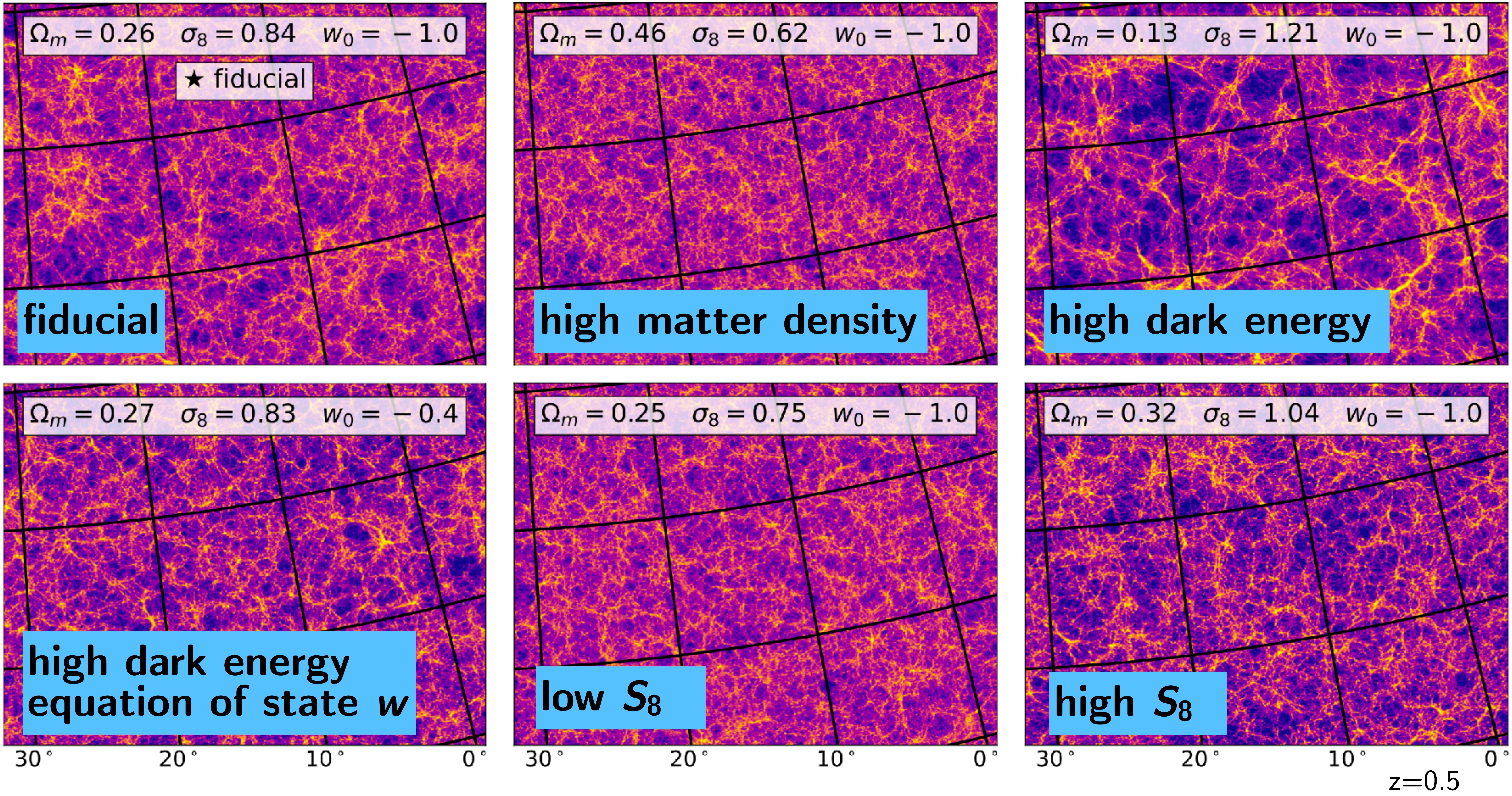
Compare with observations

parameter measurement



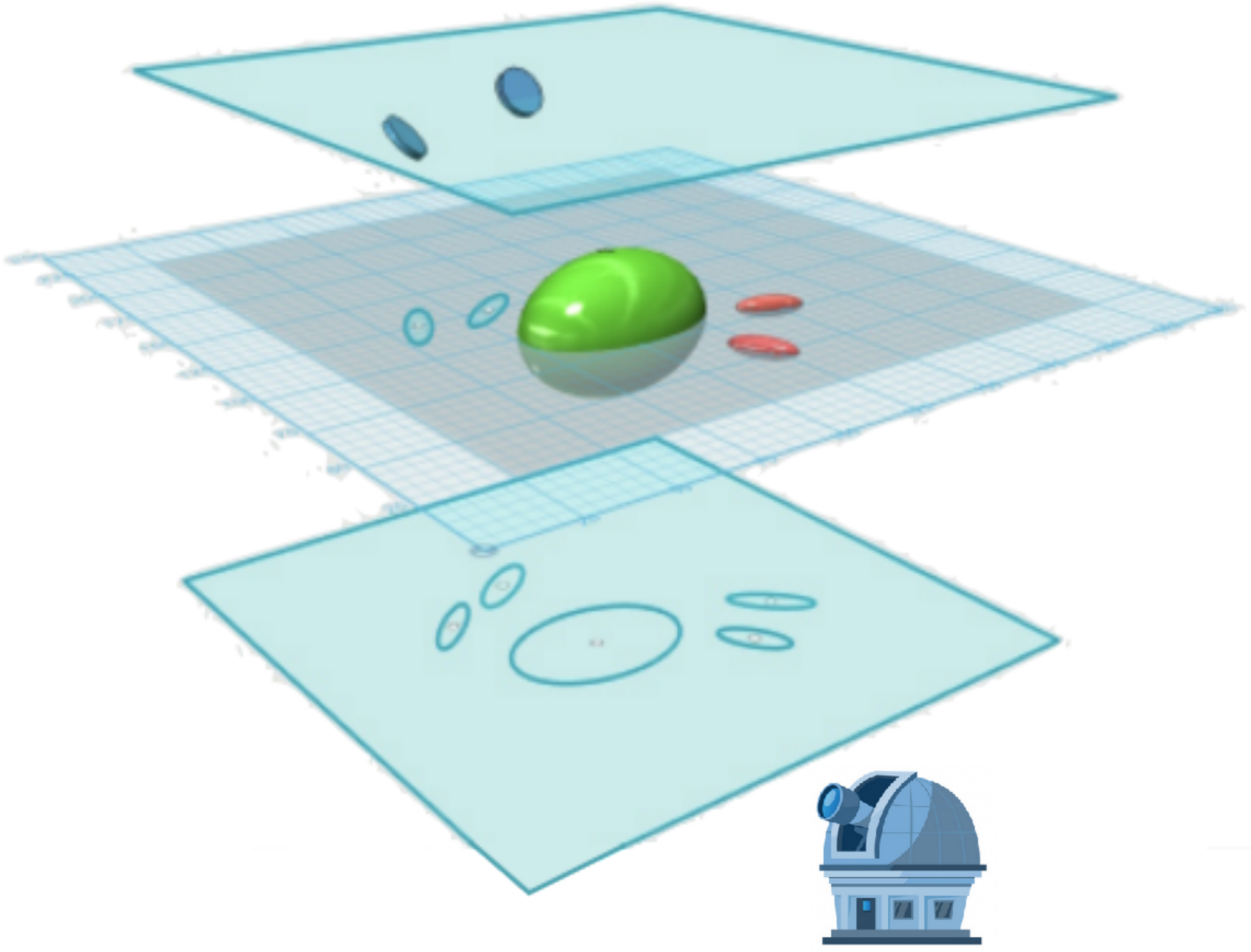
Secco, +DES. +TK, 2105.13544

Dark matter distributions carries information about cosmological parameters



Tomographic combined probes of LSS

- Photometric surveys take images of the sky for a few filters
- Photo-z is used to make maps of galaxy positions and shapes



background galaxies

foreground galaxies

telescope image

Image: Joachimi et al. 2015

Tomographic combined probes of LSS

- **Weak lensing: galaxy shapes are unbiased tracer of dark matter between the galaxy and the observer**
- Galaxy clustering: galaxy positions are a biased tracer of underlying dark matter
- Intrinsic alignmnets: galaxy shapes are also aligned with their local density environment

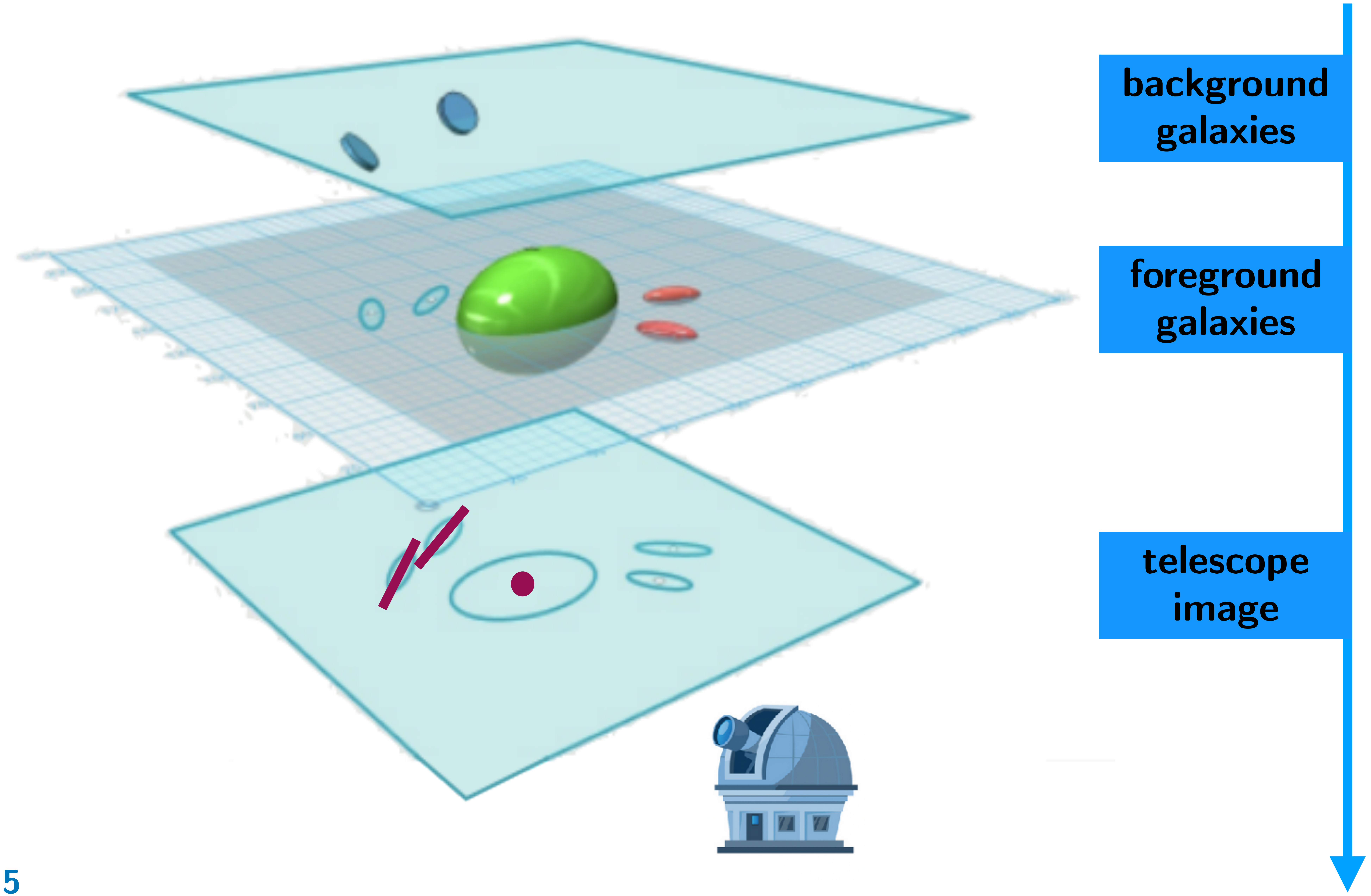


Image: Joachimi et al. 2015

Tomographic combined probes of LSS

- Weak lensing:
  galaxy shapes are unbiased tracer of dark matter between the galaxy and the observer
- **Galaxy clustering: galaxy positions are a biased tracer of underlying dark matter**
- Intrinsic alignmnets:
  galaxy shapes are also aligned with their local density environment

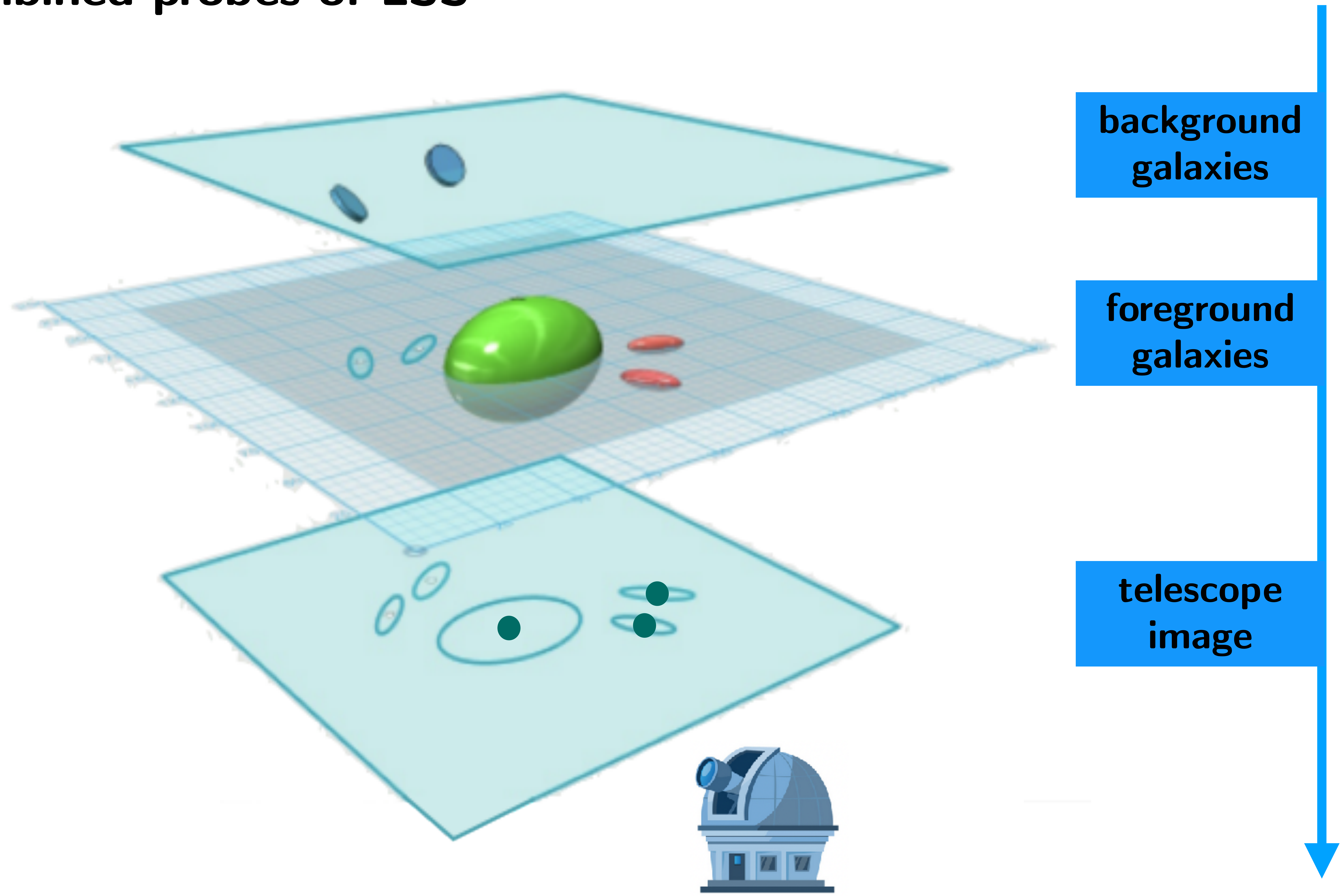


Image: Joachimi et al. 2015

Tomographic combined probes of LSS

- Weak lensing:
  galaxy shapes are unbiased tracer of dark matter between the galaxy and the observer
- Galaxy clustering:
  galaxy positions are a biased tracer of underlying dark matter
- **Intrinsic alignmnets:
  galaxy shapes are also aligned with their local density environment**

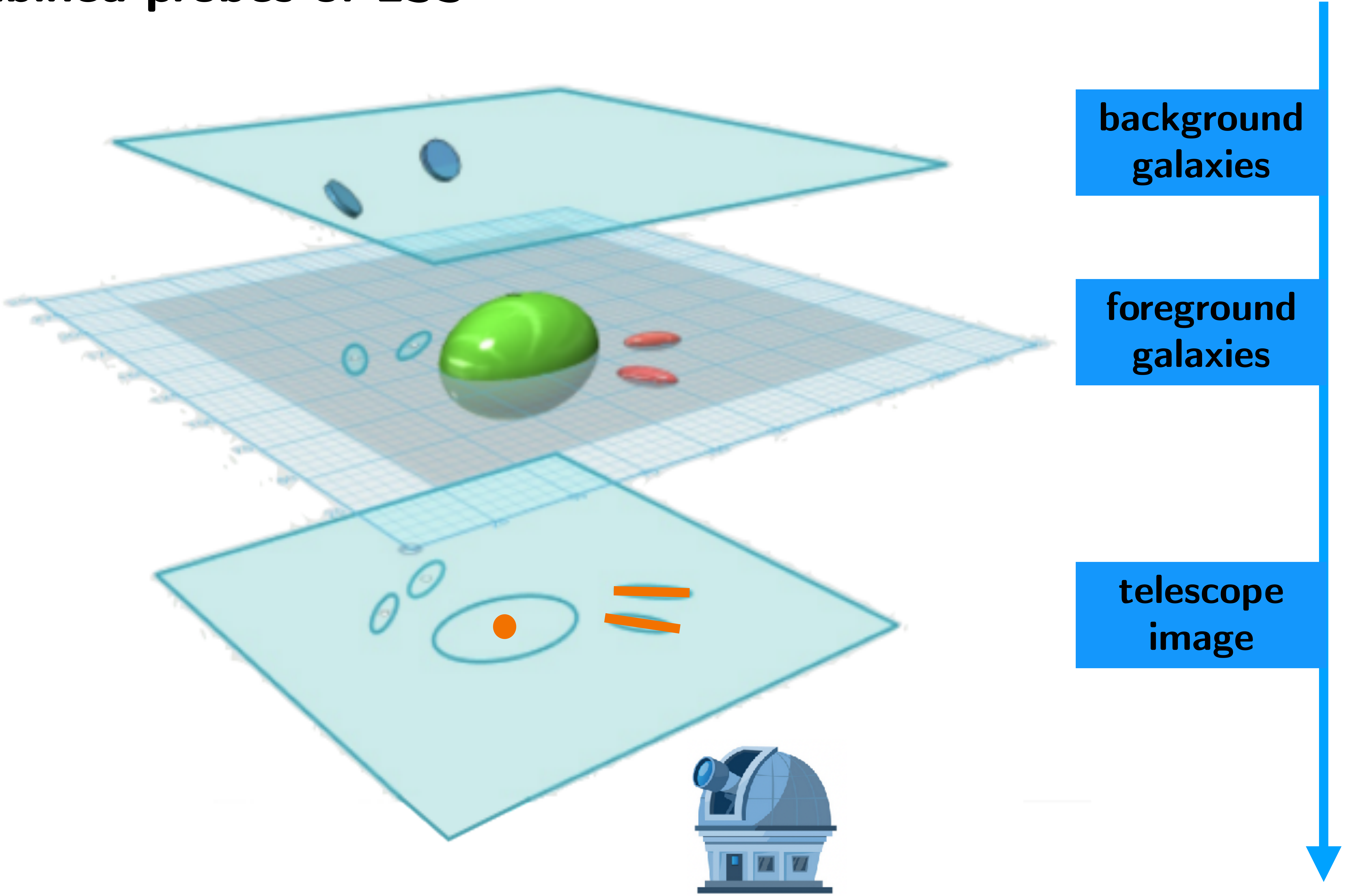
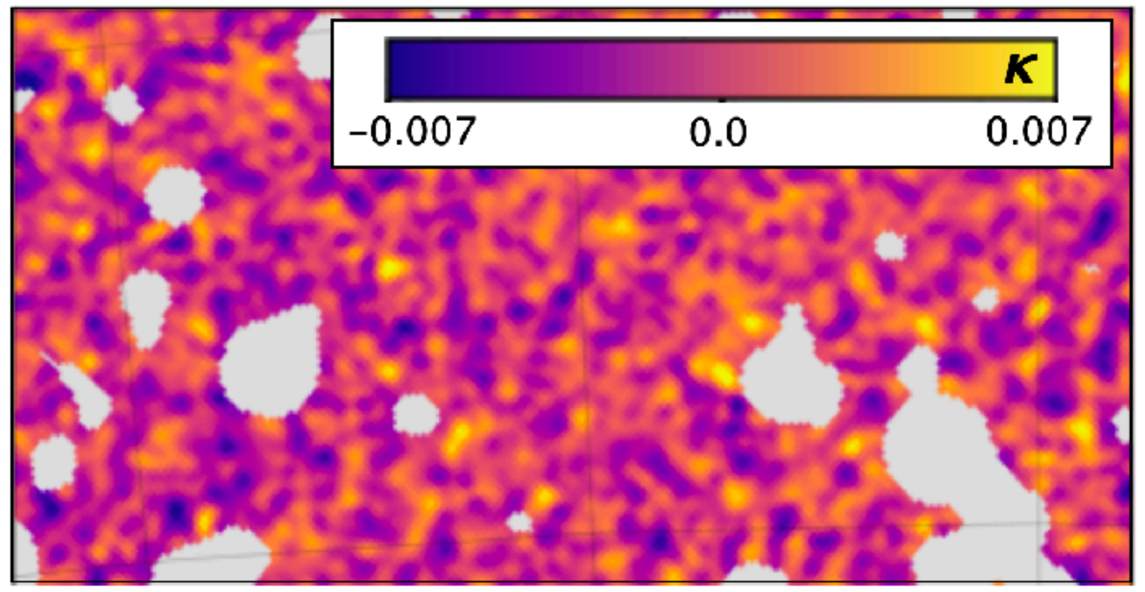


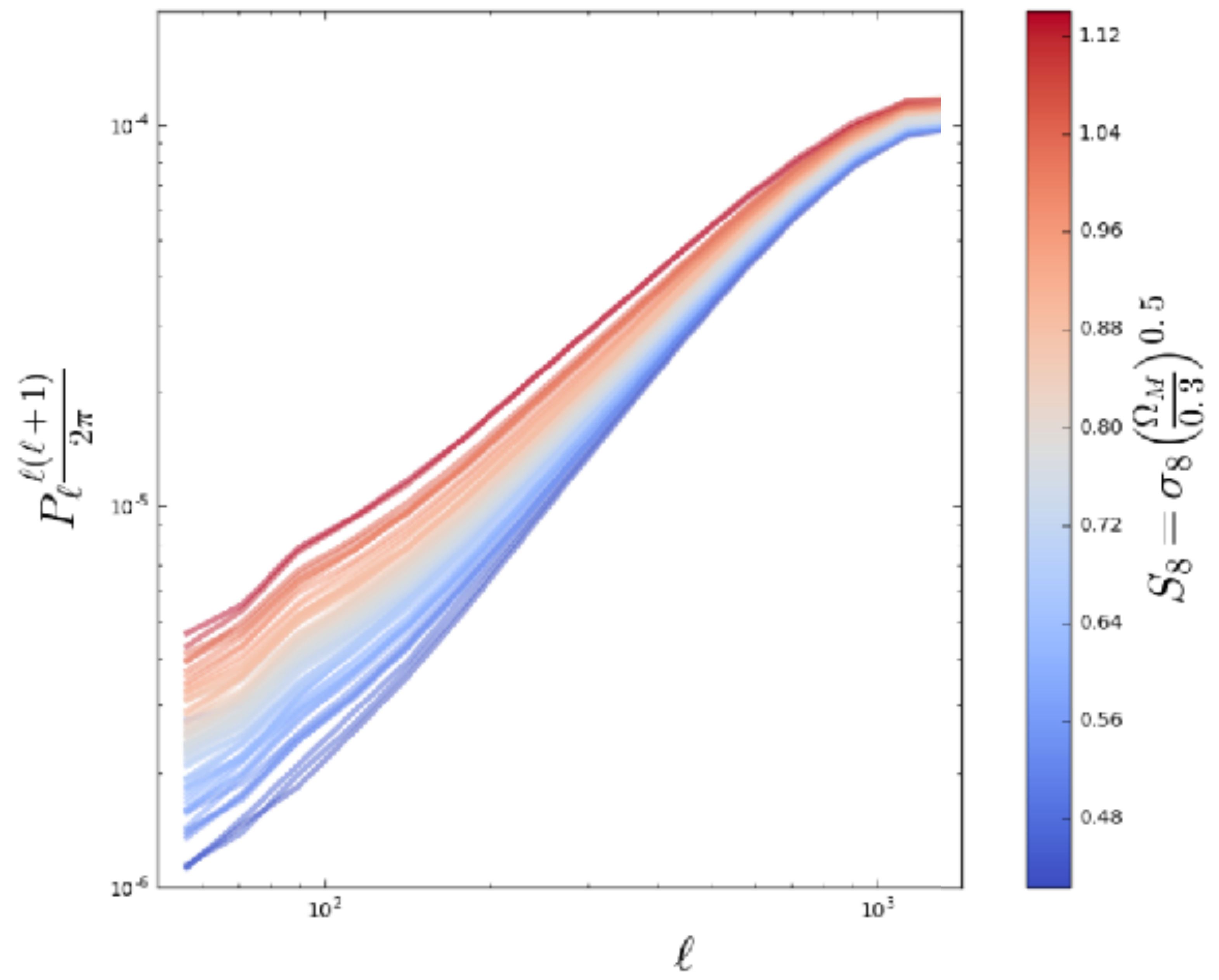
Image: Joachimi et al. 2015

LSS observations

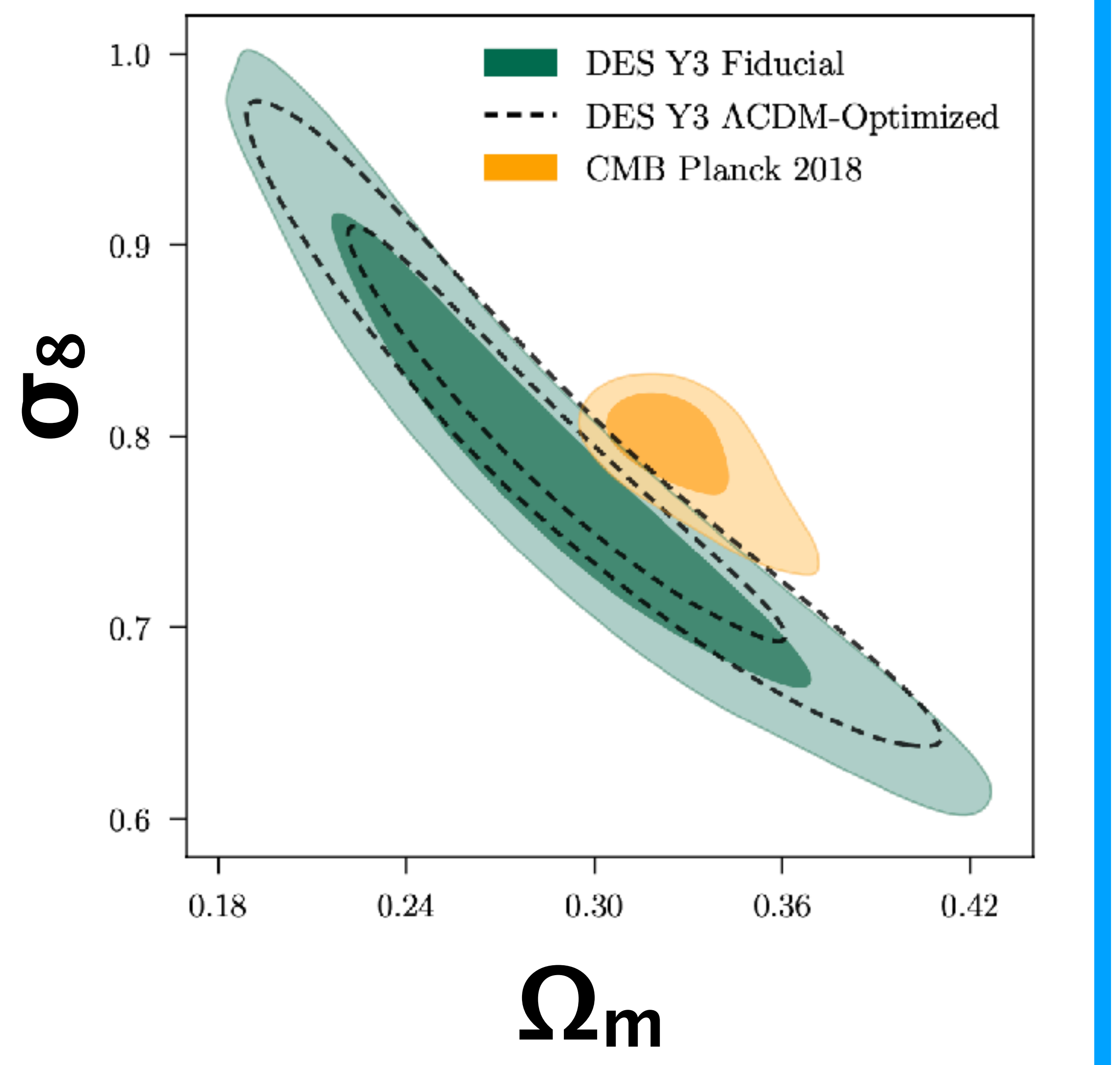


Traditional inference

statistics: 2pt functions



parameter measurement

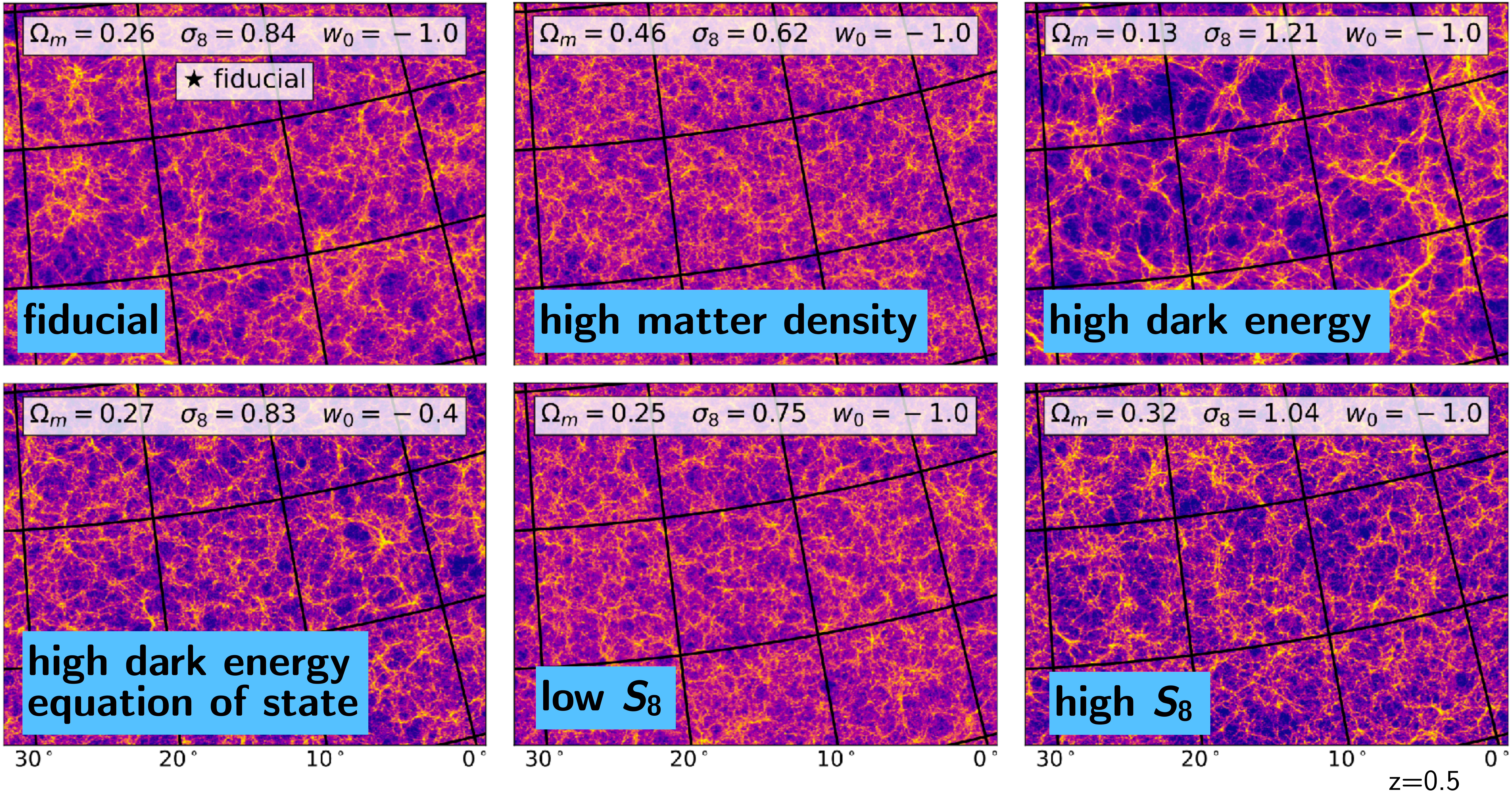


$$C_l = \frac{9}{16}\left(\frac{H_0}{c}\right)^4 \Omega_m^2 \int_0^{\chi_h} d\chi \left[\frac{g(\chi)}{ar(\chi)}\right]^2 P\left(\frac{l}{r}, \chi\right)$$

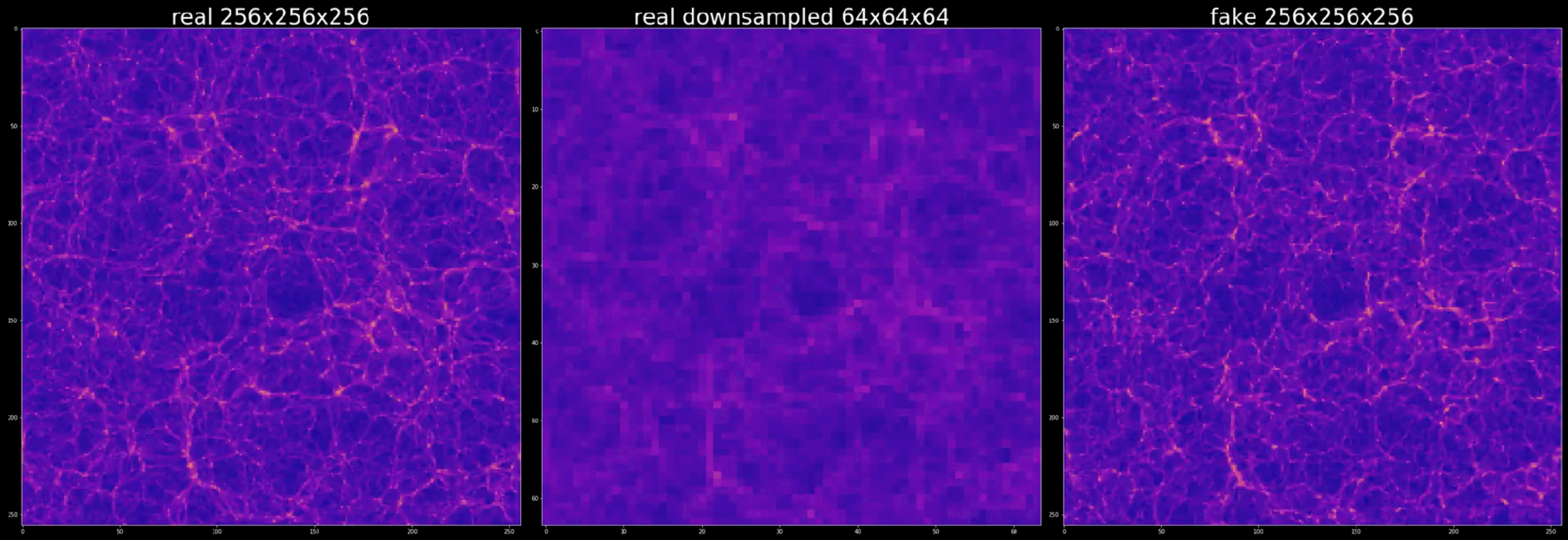
theory prediction: analytical

Inference with Deep Learning
LSS observations
κ
−0.007
0.0
0.007
statistics: deep convolutional network
16x
32x
64x
128x
256x
1024x
256x
5x
Convolutions
Fully Connected
theory prediction: simulations
...
parameter measurement
DES Y3 Fiducial
DES Y3 ΛCDM-Optimized
CMB Planck 2018
σ8
1.0
0.9
0.8
0.7
0.6
0.18
0.24
0.30
0.36
0.42
Ωm

Dark matter distributions carries information about cosmological parameters



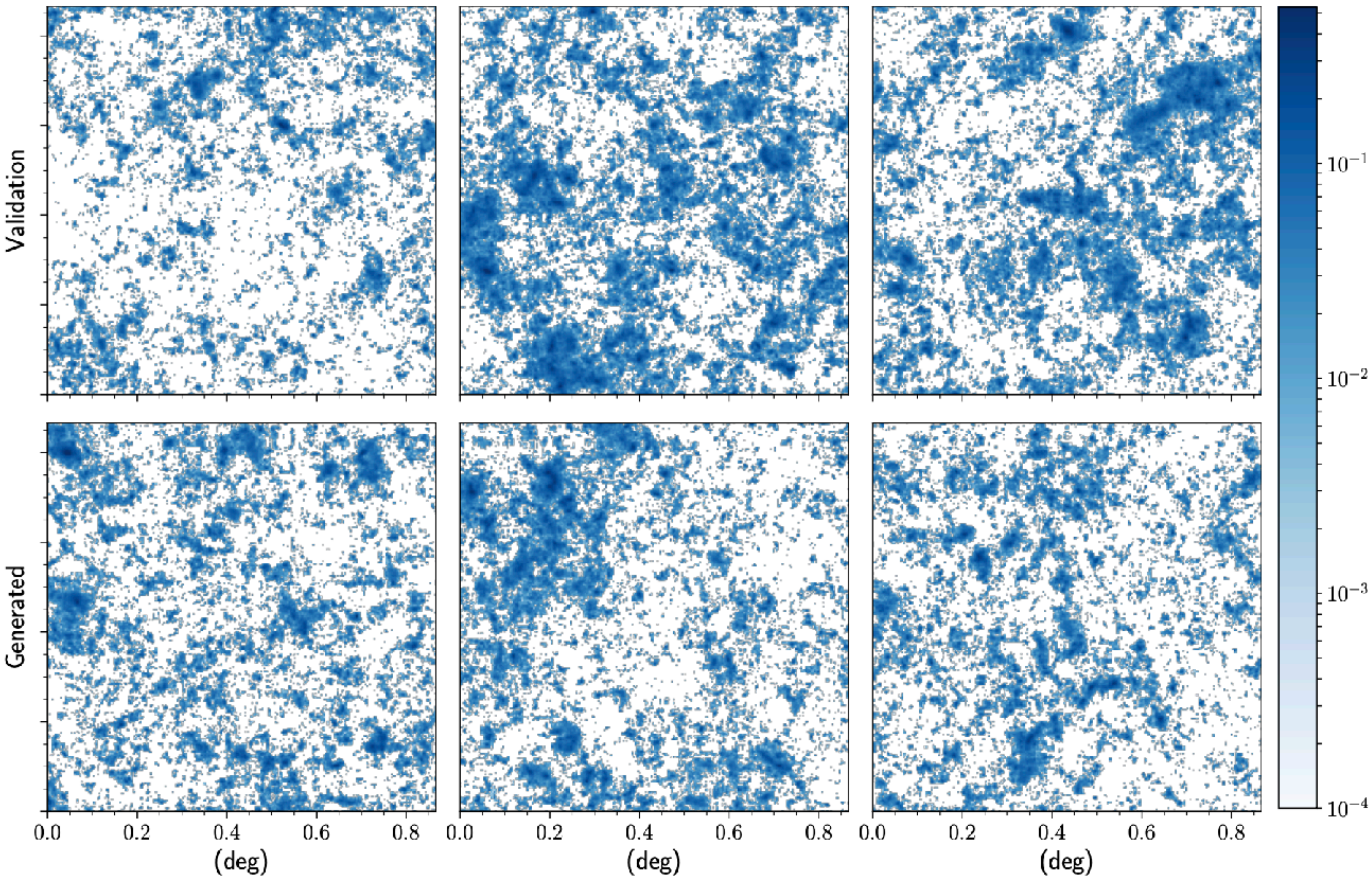
N-body cube generation in 3D



Srivastava, TK, et al., submitted

First Generative Model for cosmological mass maps

- First generative model trained on simulations applied to cosmological fields
- N-body vs GAN visually indistinguishable
- Excellent agreement on (non-Gaussian) summary statistics
- Very simple networks, worked out-of-the-box

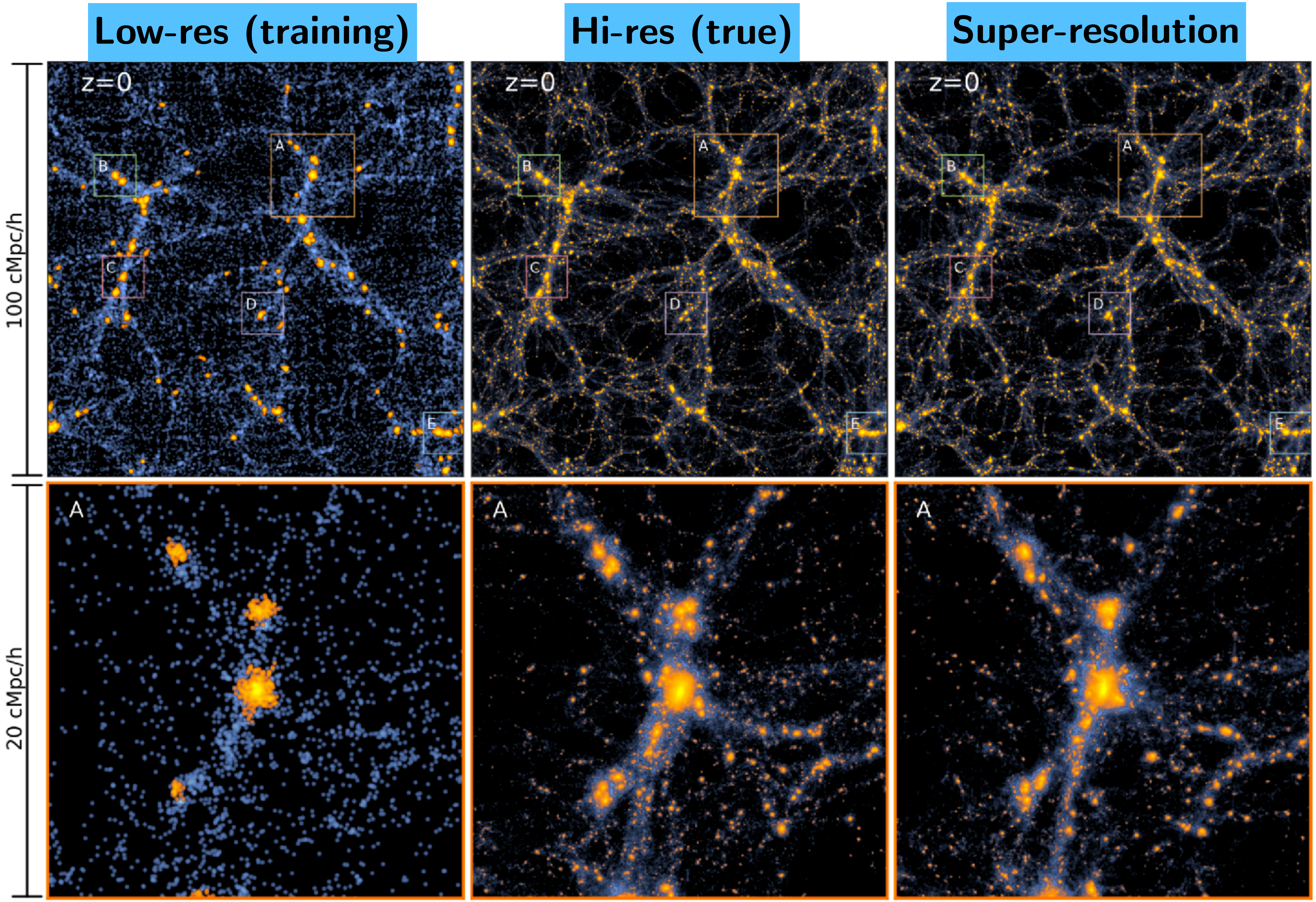


Mustafa et al. 2017 1706.02390

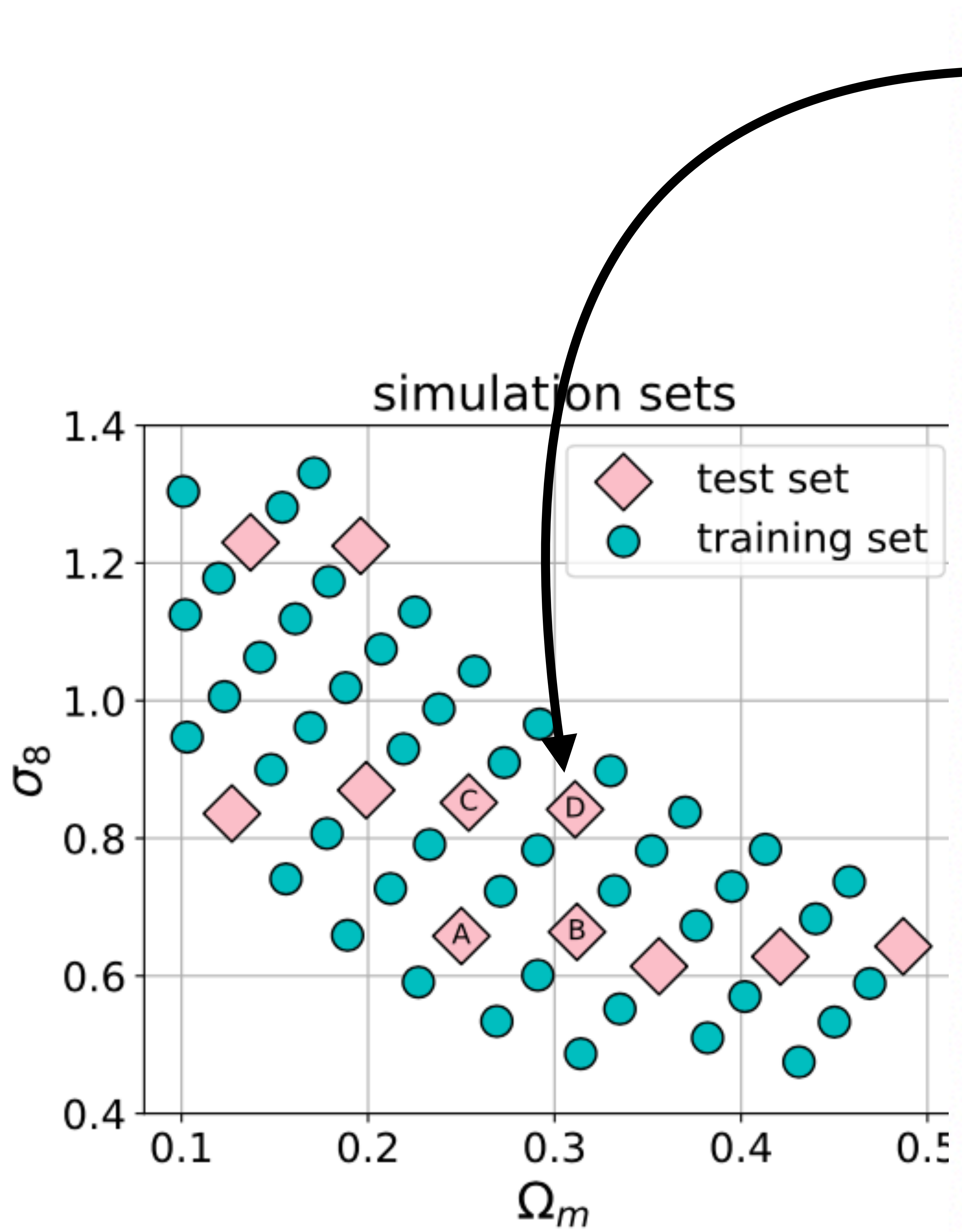
AI super-resolution of N-body simulations

Li et al. 2021 2010.06608

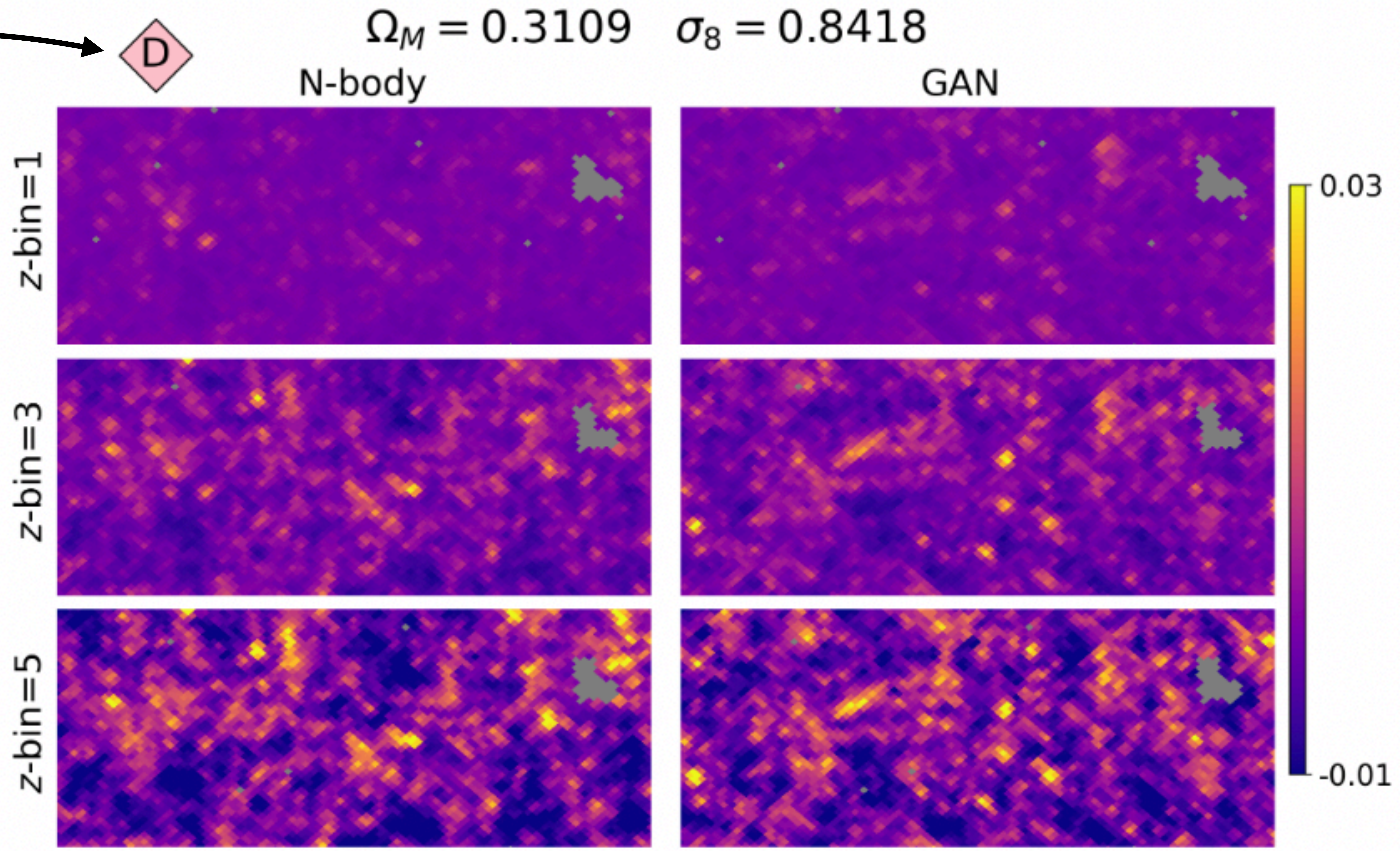
- Learn the mapping from the low to high resolution simulations
- Works on 3D volumes!
- Using Wasserstein GANs with gradient penalty on 3D volumes
- Increase of resolution by a factor of 8
- Super-resolution is extremely fast
- Reproduces well the halo mass function (10^{11}-10^{14} $M_{\odot}$) and power spectra (k between 0.1 - 10)
- Works for a single cosmology, separate GAN for each redshift



KiDS-1000 mass map emulator



Grid of simulations as train/test set



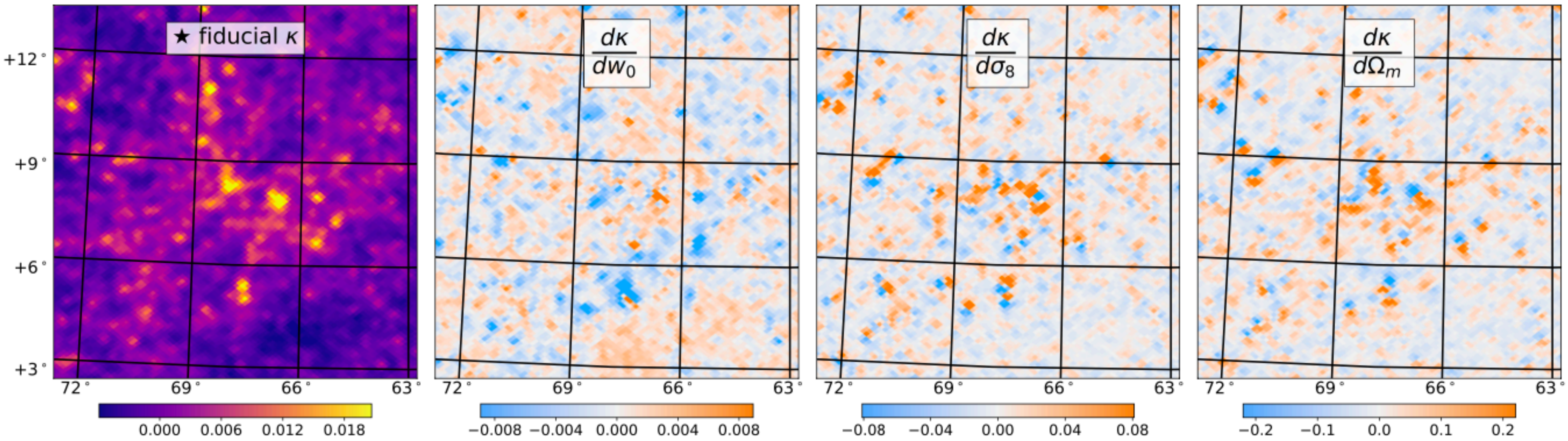
Visual comparison between original N-body and GAN maps

Very fast generator publicly available: https://tfhub.dev/cosmo-group-ethz/models/kids-cgan/1

Fiducial cosmology and derivatives

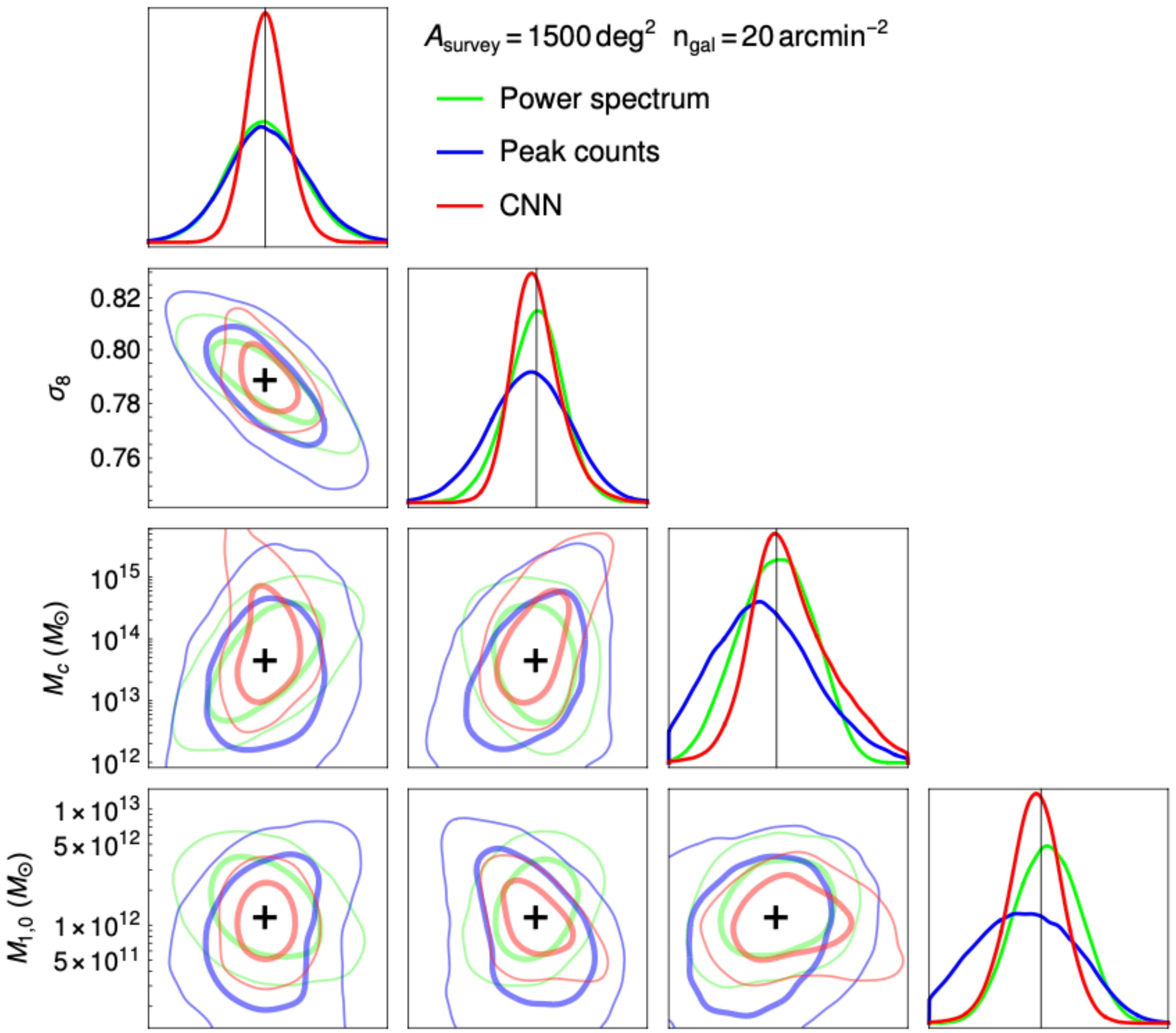
TK, Fluri, Schneider, Refregier, Stadel 2209.04662

- There are 200 independent simulations at the fiducial cosmology
- Each has a +/- Δ simulation with the same initial conditions
- Can be used to create map derivatives with respect to the cosmological parameters
- Useful for Information Maximising Neural Networks



Deep learning helps with constraining baryons

- Forecast for Stage-3 survey with 20 galaxies/arcmin²
- High-resolution maps, pixel size 0.4 arcmin
- Using baryonic correction model (BCM) with parameters: M_c, $M_{1,0}$, η, β
- Improvement over power spectrum for $\Omega_m - \sigma_8$ figure of merit: 1.66× with baryons marginalized
- CNN improves constraints on M_c, $M_{1,0}$, but does not constrain η, β

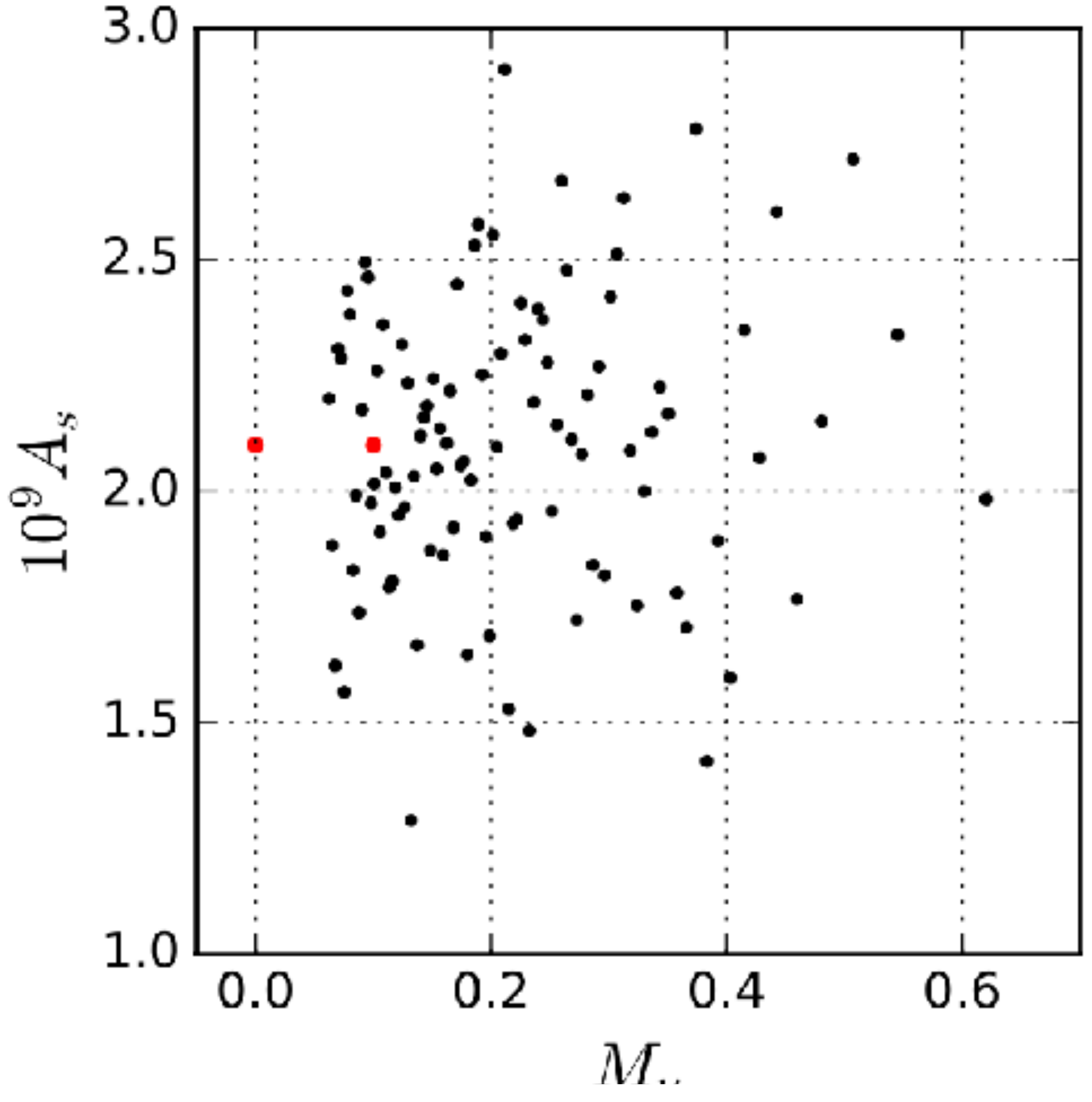


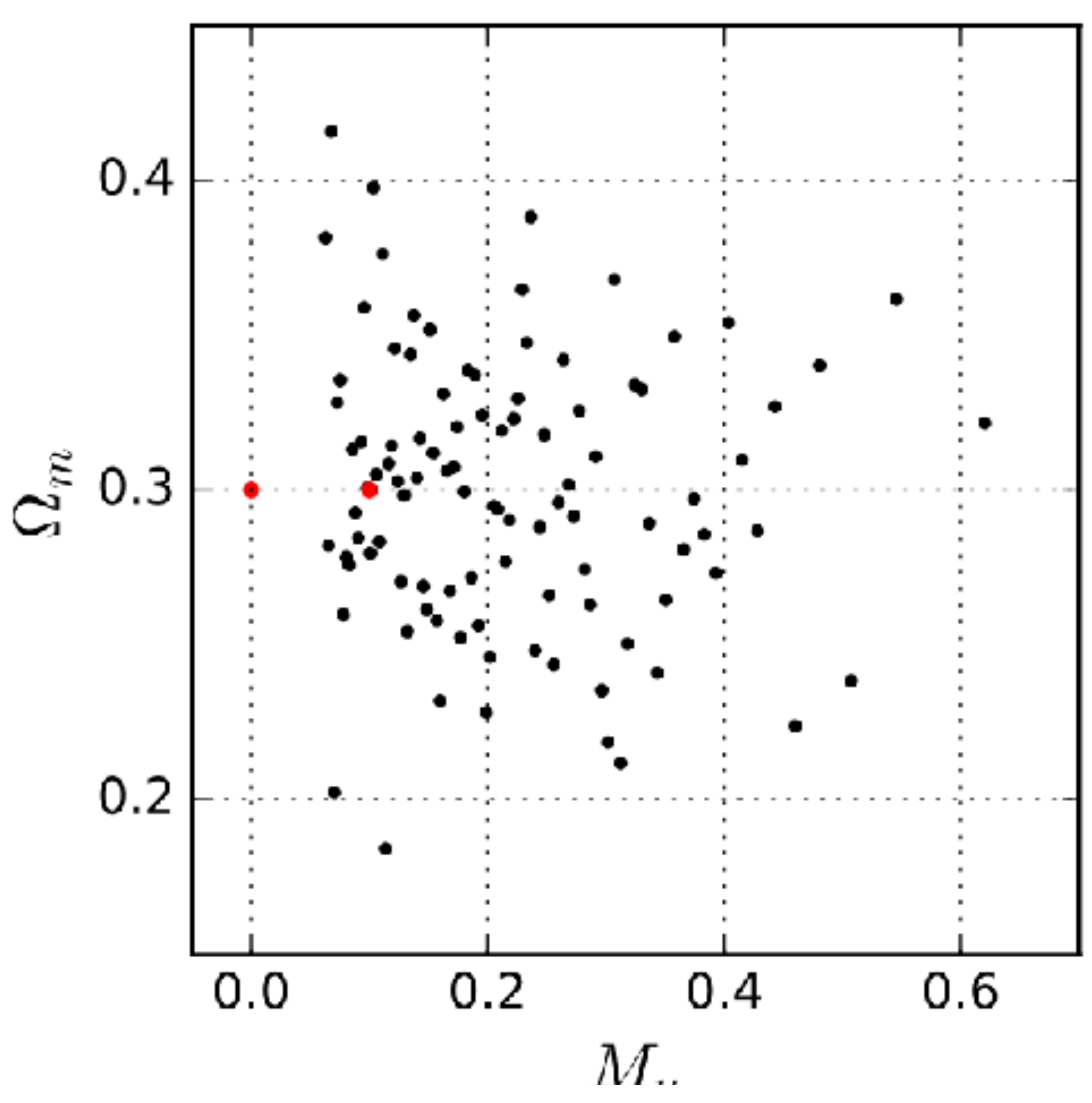
+2 more baryon parameters: ν, β

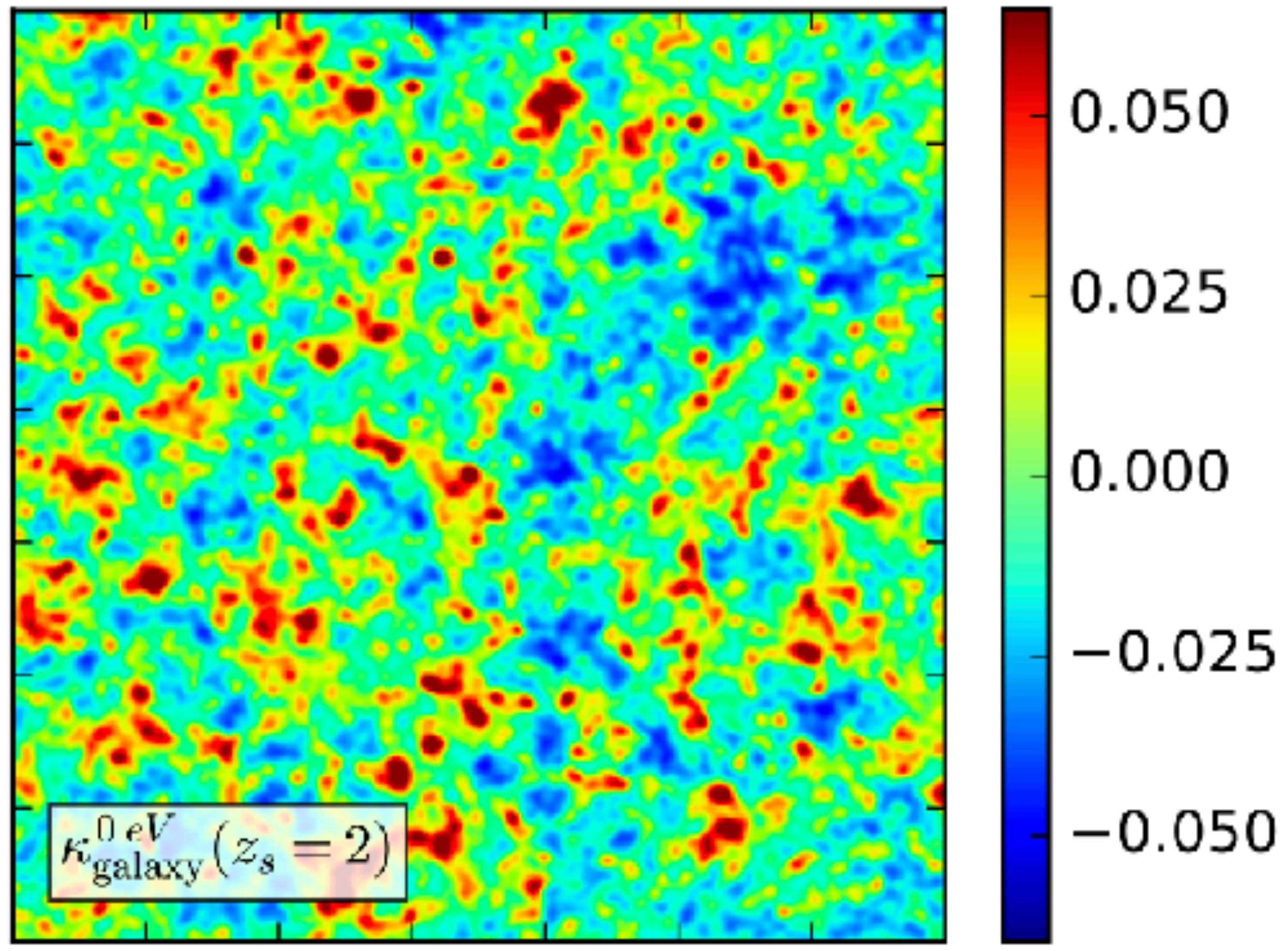
Lu, Haiman, Zorilla-Matilla 2022 2109.11060

Massive-Nus

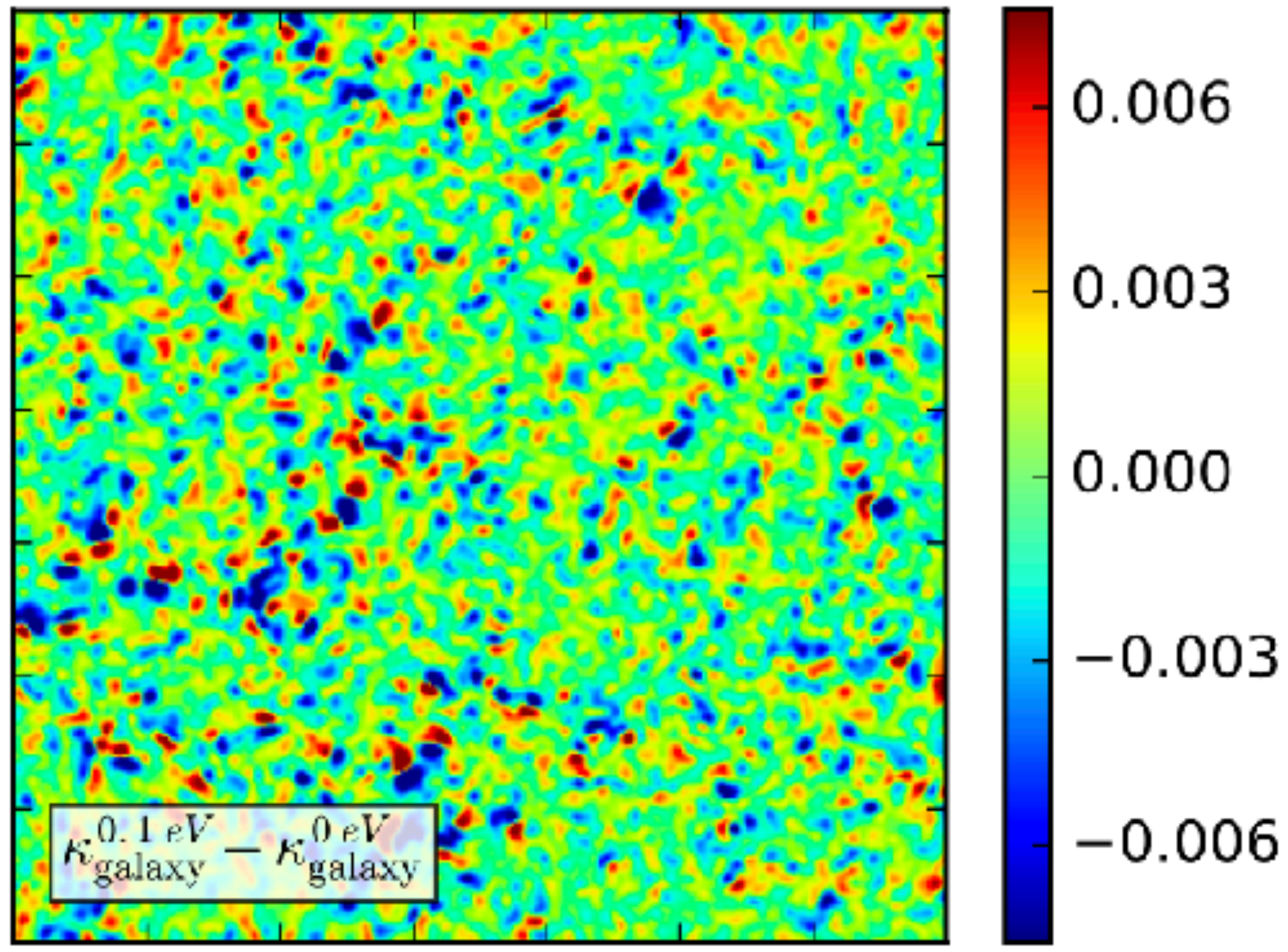
- Designed to explore simulations-based inference for constraining cosmological neutrino masses
- 101 simulations in Ω_m, A_s, M_ν
- Halo catalogs and merger trees for all simulations stored
- Lensing convergence maps 12.25^2 deg^2
- Relatively high resolution: $512\ h^{-1}\ Mpc$ boxes with 1024^3 particles
- Large neutrino mass range: $M_\nu \in [0,\ 0.6]\ eV$
- Data available at: http://columbialensing.org







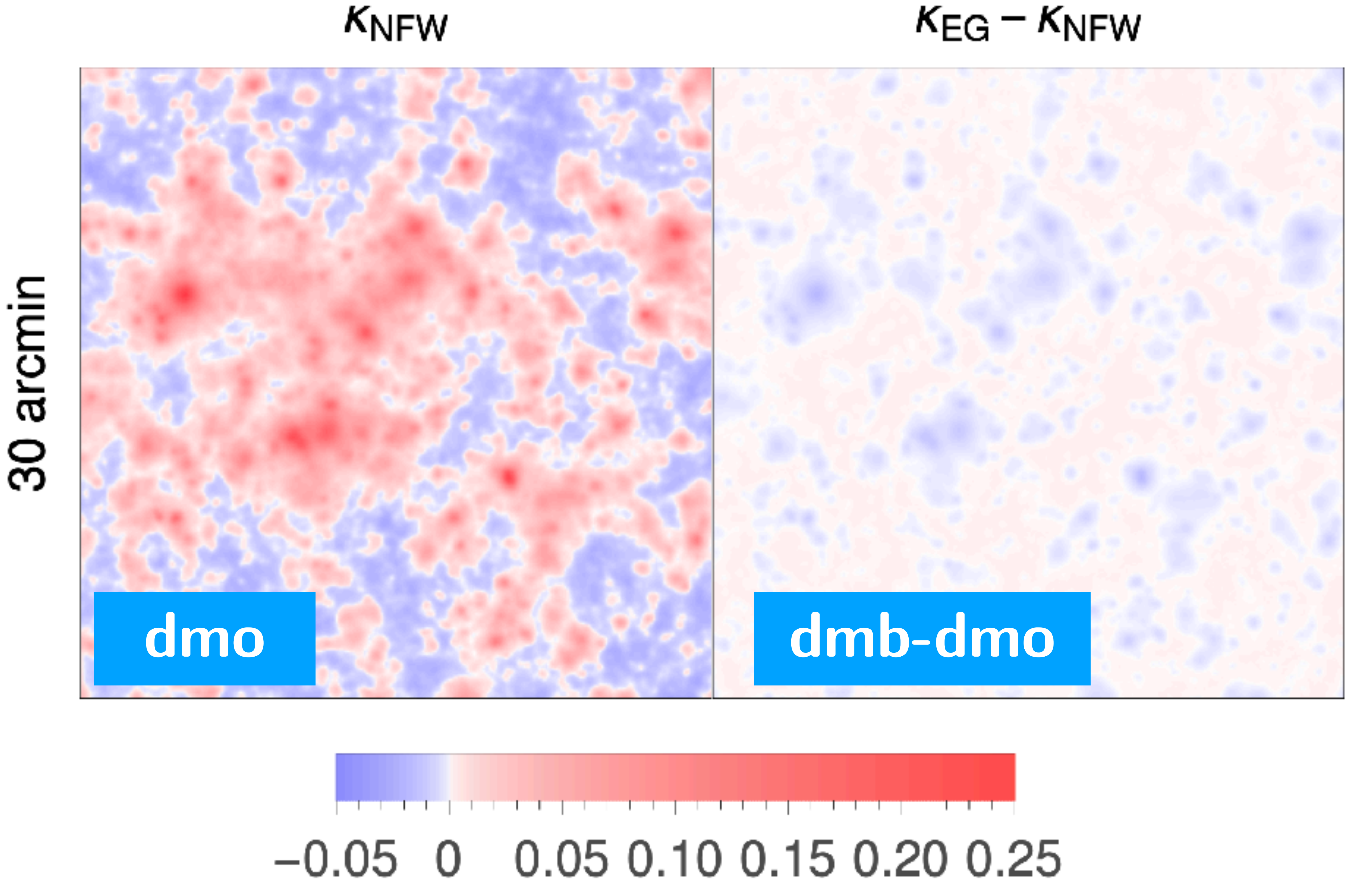
no neutrinos



difference with neutrinos at $M_\nu = 0.1\ eV$

Deep learning helps with constraining baryons

- Forecast for Stage-3 survey with 20 galaxies/arcmin^2
- High-resolution maps, pixel size 0.4 arcmin
- Using baryonic correction model (BCM) with parameters: M_c, $M_{1,0}$, η, β
- Improvement over power spectrum for $\Omega_m - \sigma_8$ figure of merit: 1.66× with baryons marginalized
- CNN improves constraints on M_c, $M_{1,0}$, but does not constrain η, β



Parameter	Meaning
M_c $[h^{-1}\mathrm{M}_\odot]$	characteristic halo mass for retaining half of the total gas
$M_{1,0}$ $[h^{-1}\mathrm{M}_\odot]$	characteristic halo mass for a galaxy mass fraction of 0.023
η	maximum distance of the ejected gas from the parent halo
β	logarithmic slope of the gas fraction *vs.* the halo mass

Lu, Haiman, Zorilla-Matilla 2022 2109.11060