Simulation-based cosmological inference from 3D maps of multiple tracers

Giulio Scelfo

(SISSA, INFN)

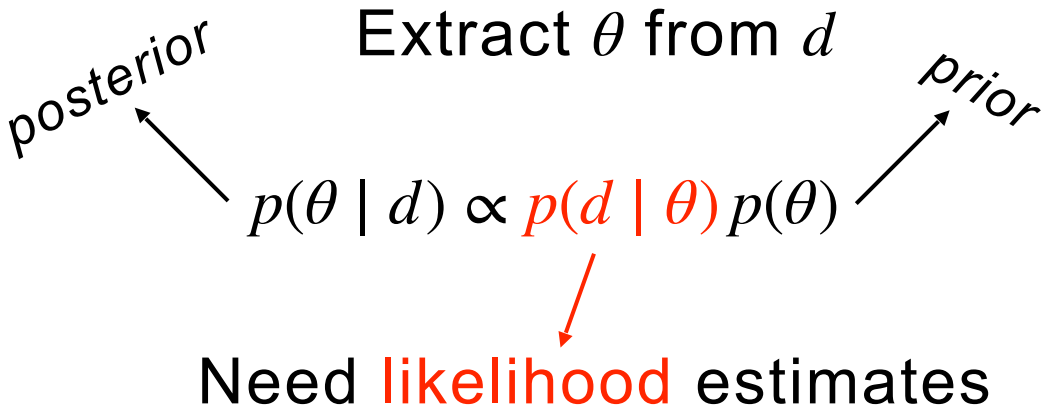
2nd SMASHING workshop, Vipava, October 22nd 2026





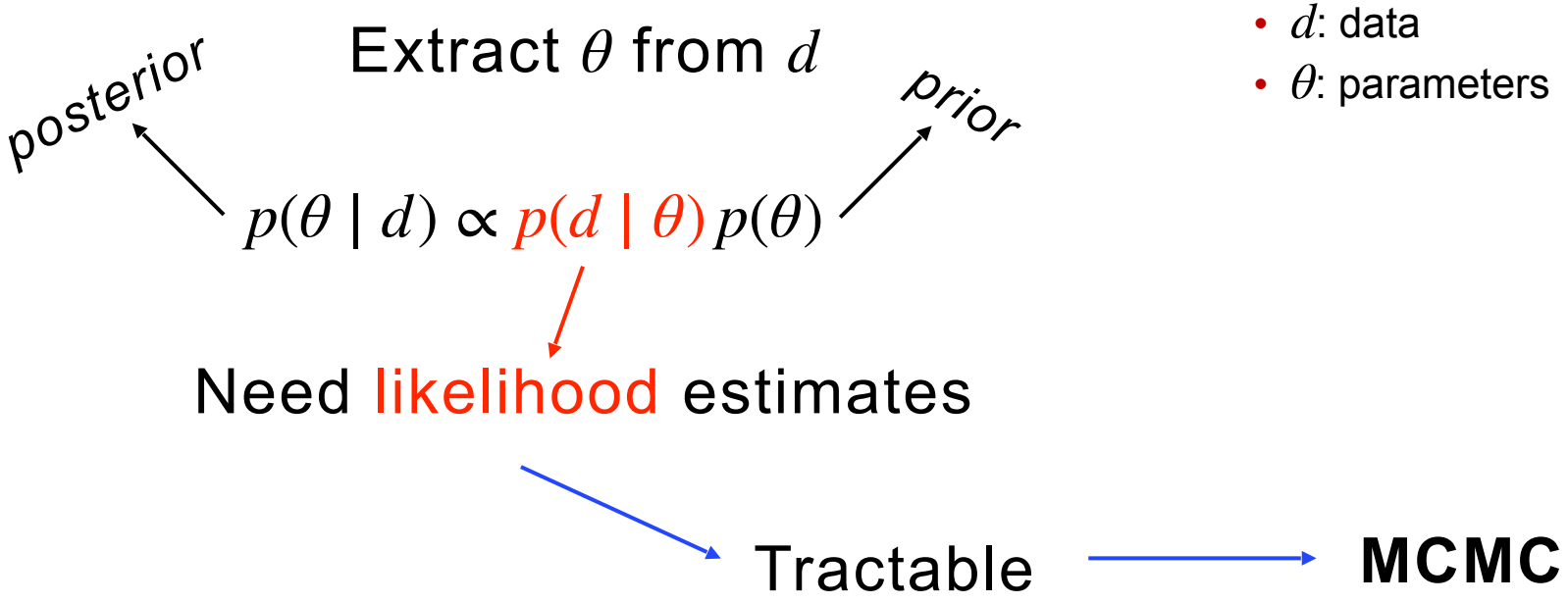

In collaboration with S. Mishra, M. Rigo, R. Trotta, M. Viel [arXiv: 2605.26210 (accepted on JCAP)]

Intro: Parameter inference



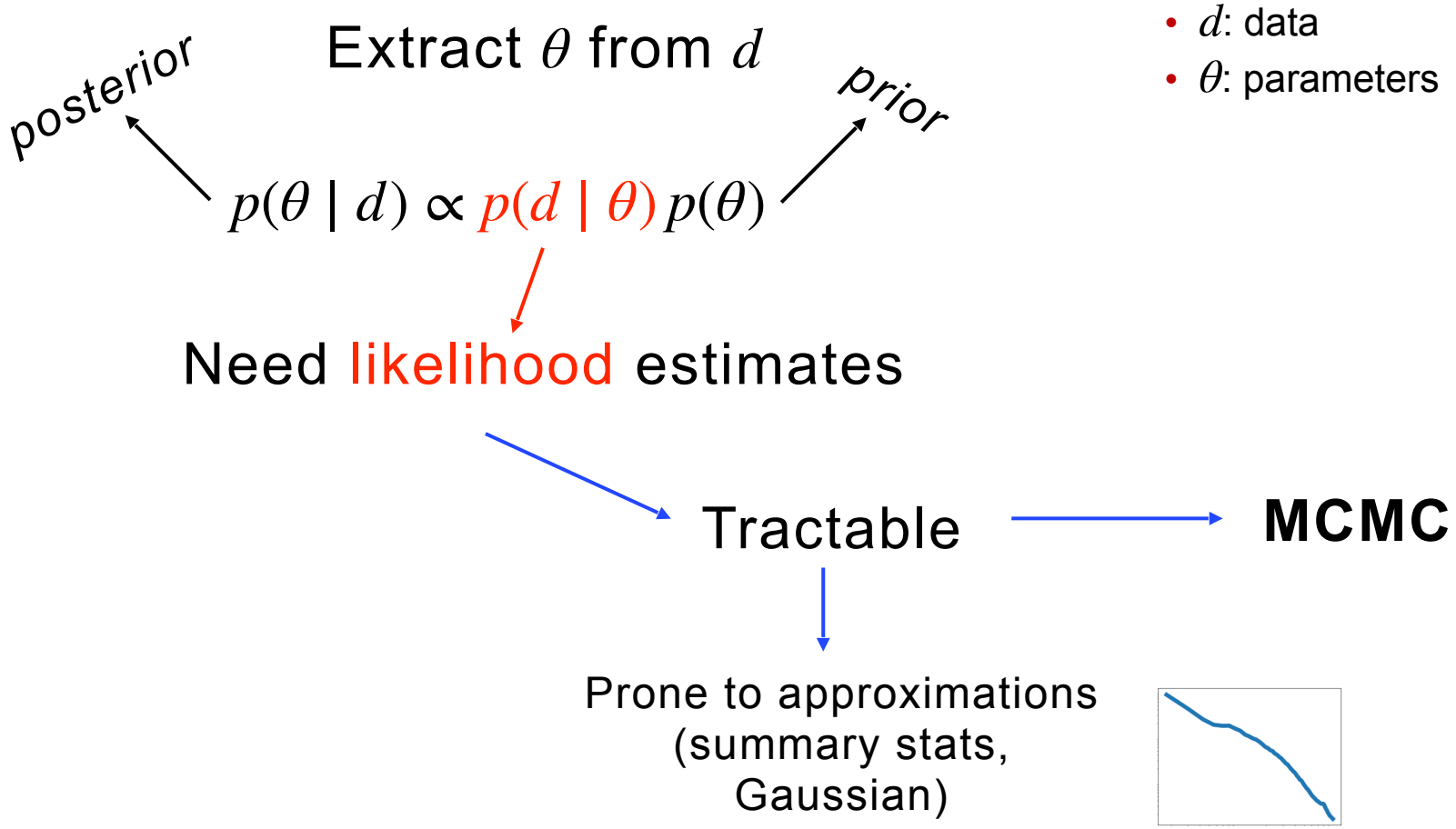
- d: data
- θ: parameters

Intro: Parameter inference



- d: data
- θ: parameters

Intro: Parameter inference



- d: data
- θ: parameters

Intro: Parameter inference

- d: data
- θ: parameters

Extract θ from d

posterior

prior

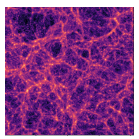
$$p(\theta \mid d) \propto p(d \mid \theta)\, p(\theta)$$

Need likelihood estimates

Intractable

Tractable → **MCMC**

Prone to approximations (summary stats, Gaussian)



Field-level!

Simulation-based inference (SBI)

Aim of the work

Let's say we want to perform **cosmological inference**:

$$\theta_{\mathscr{C}} = \{\Omega_m, \sigma_8\}$$

Present day total matter fraction

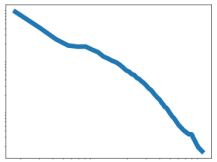
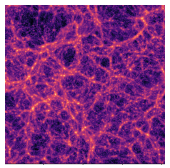
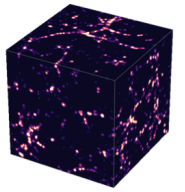
Amplitude of matter density fluctuations

And we'd like to perform **SBI**.

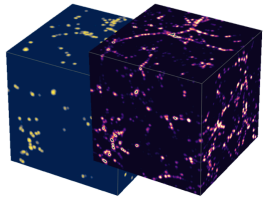
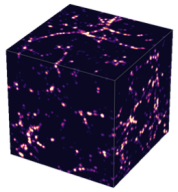
Aim of the work

Systematic exploration of the **information gain in cosmological inference when**:

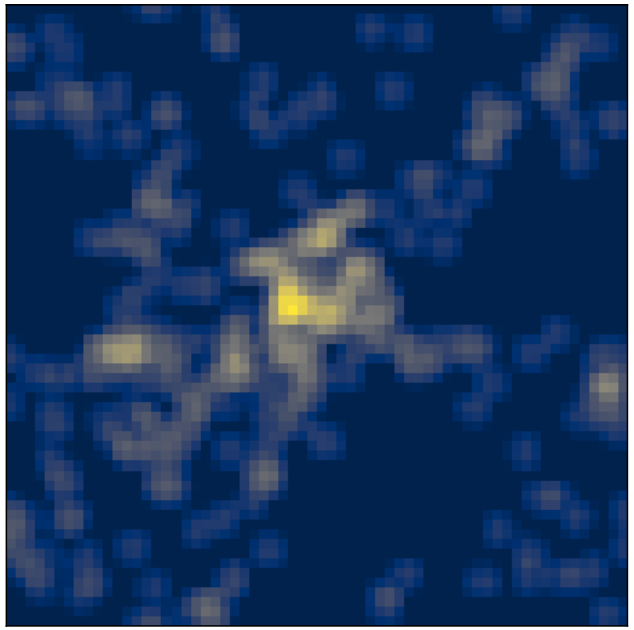
- switching **from summary stats to** (2D or 3D) **full field** SBI?

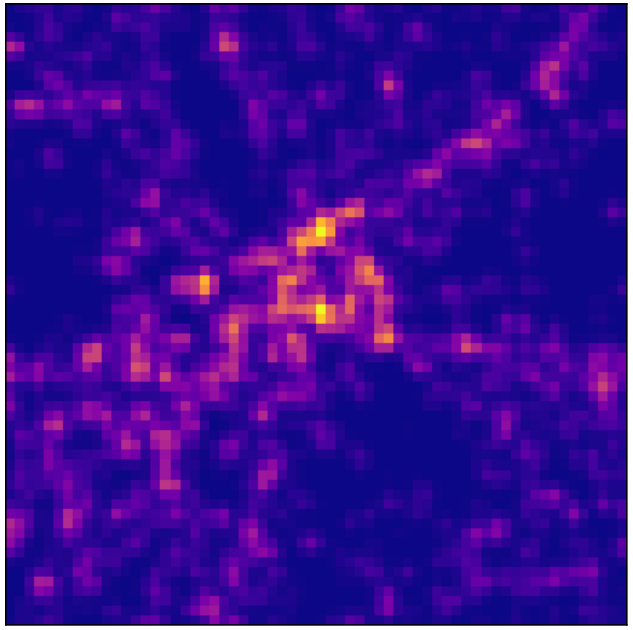
- adopting a **multi-tracers** setup instead of single probes?

Tracers



Galaxies number counts density



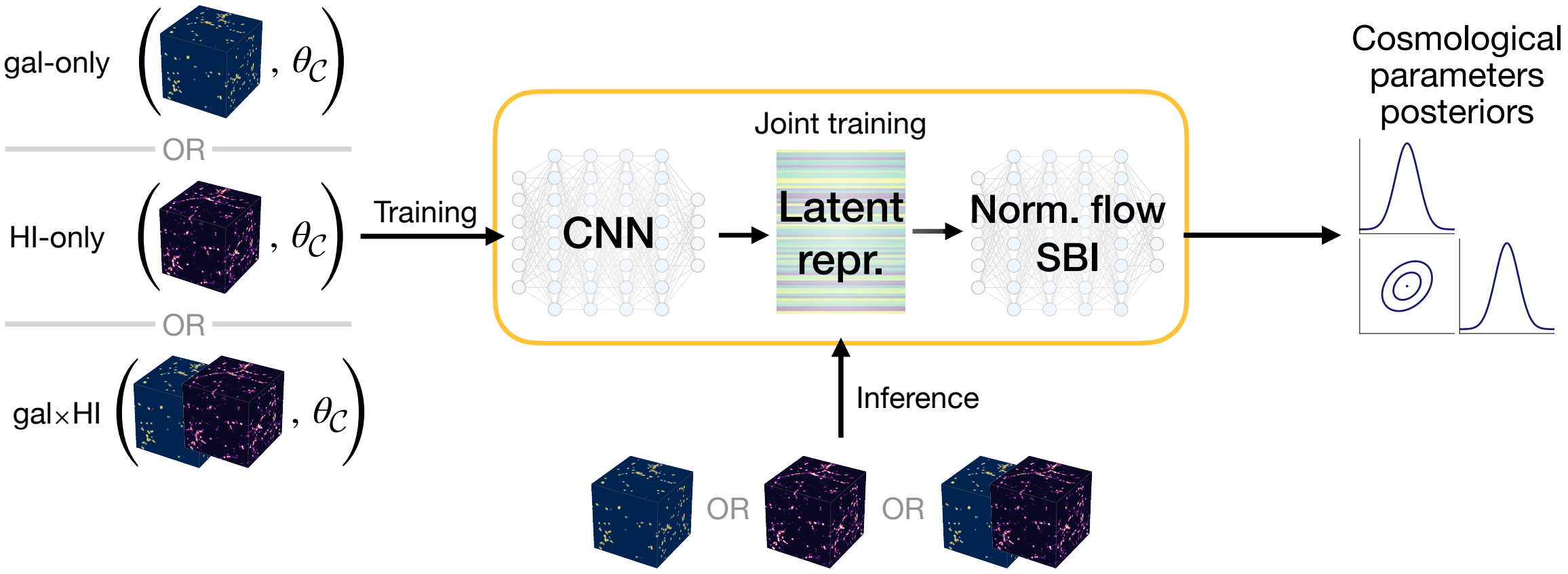
HI (neutral hydrogen)

mean brightness temperature

- Start from **CAMELS** simulations (IllustrisTNG) of $(50\ \mathrm{Mpc/h})^3$
- 512^3 elements
- z=0, smoothing 0.2 Mpc/h
- astrophysics described by 4 parameters ($A_{\mathrm{SN1}}, A_{\mathrm{SN2}}, A_{\mathrm{AGN1}}, A_{\mathrm{AGN2}}$)
- We need to **emulate** maps (Mishra+25)

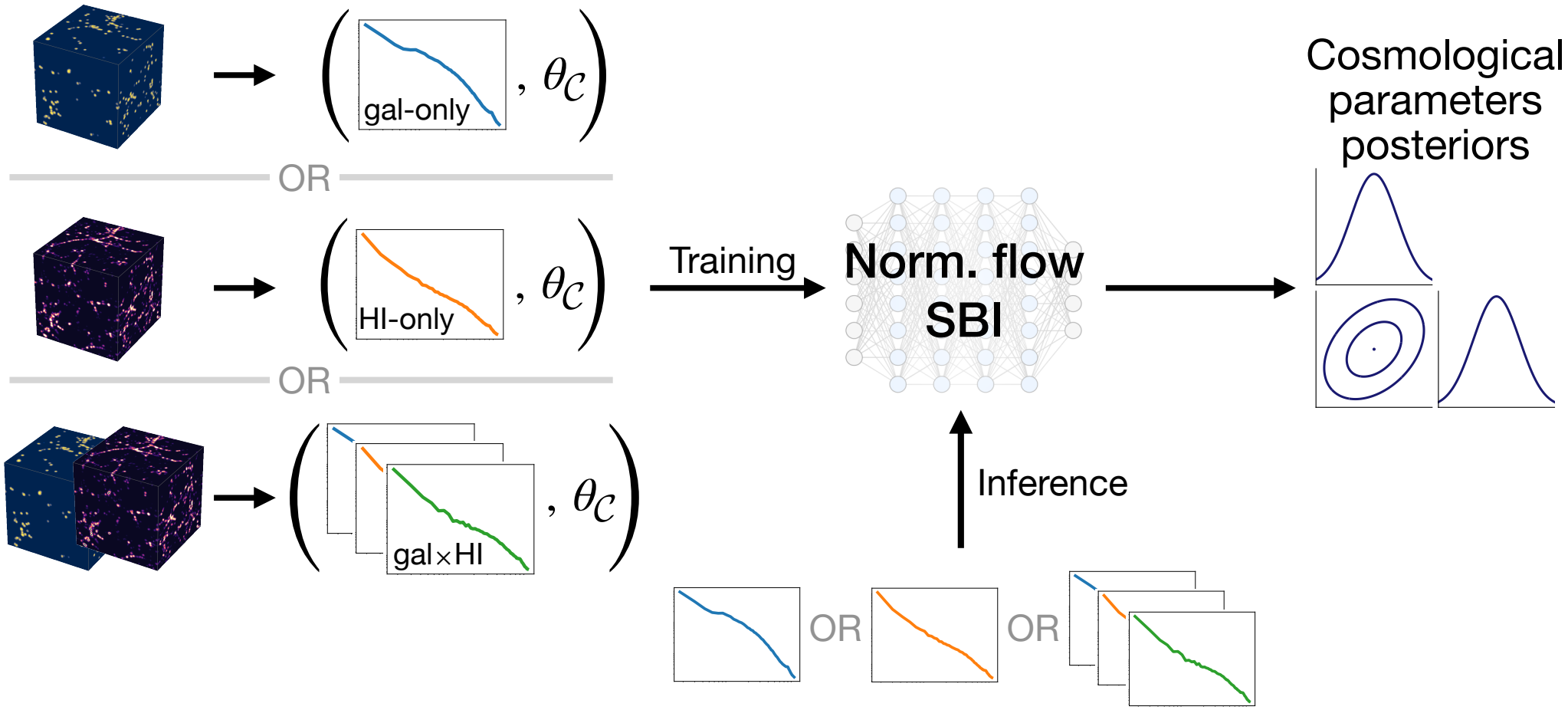
SBI on maps

Neural Posterior Estimation (NPE) with Masked Autoregressive Flows (MAFs)



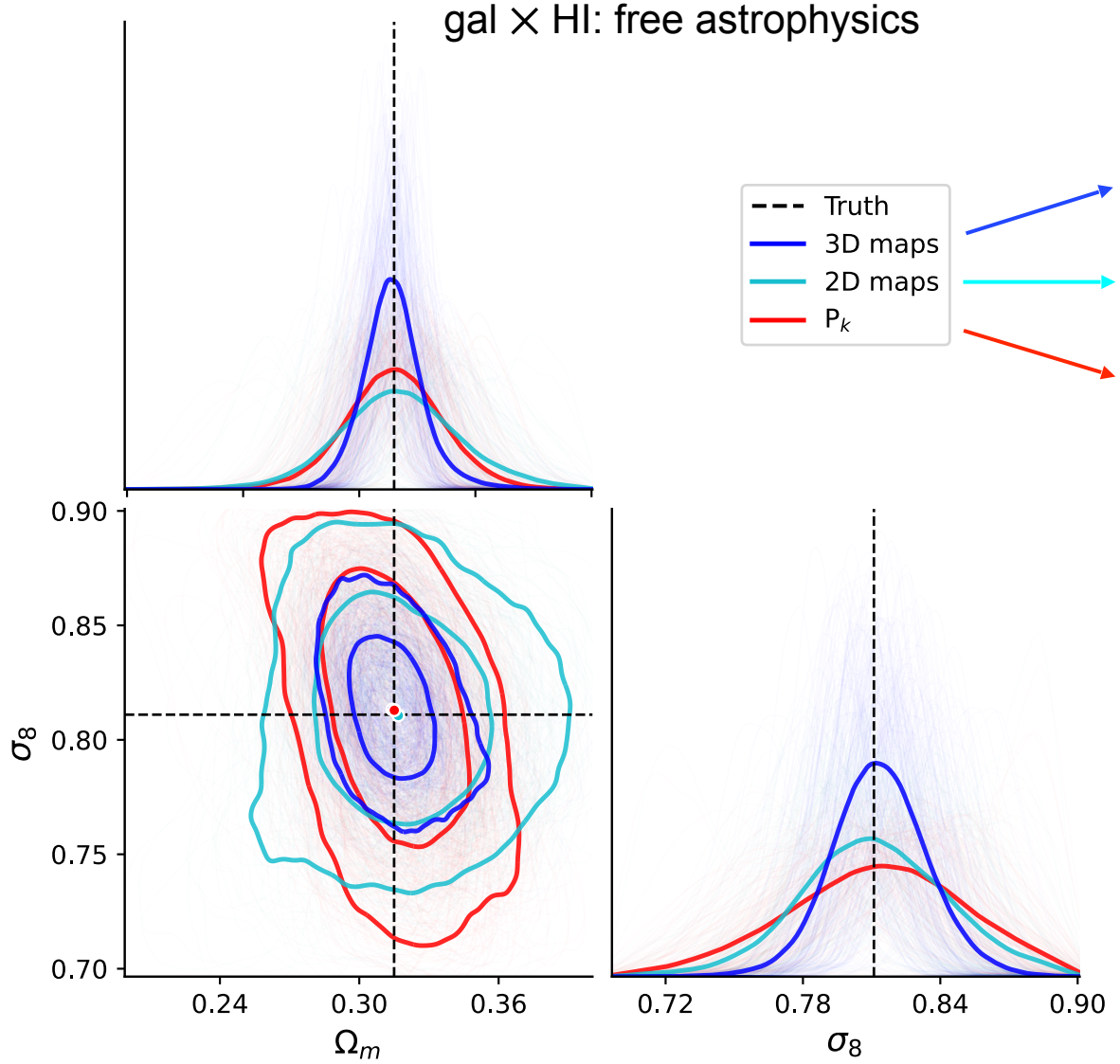
SBI on power spectra

Neural Posterior Estimation (NPE) with Masked Autoregressive Flows (MAFs)



Statistical precision

gal × HI: free astrophysics



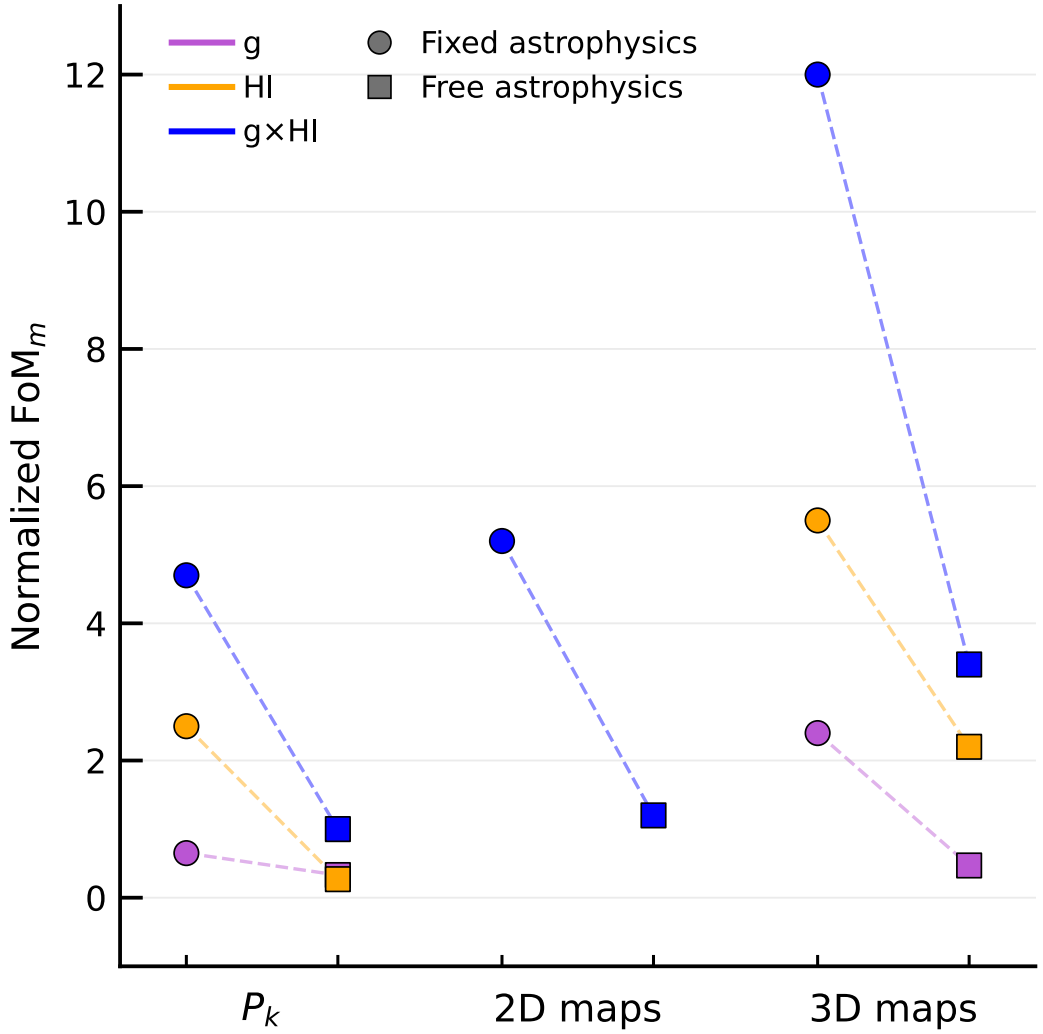
SBI training time & train set size:

- days (GPU), ∼ 500 GB
- 40 mins (GPU), ∼ 5 GB
- < 5 minutes, ∼ 1 MB

Test set:
200 seeds
LCDM-Planck
emulated maps

Statistical precision

$$\mathrm{FoM} = 1/A_{68\%}$$



- Multi-tracing improves by a factor $\sim$ 2 - 6
- 3D field-level improves by $\sim$ 3
- 2D does not add significant gain wrt Pk

(Accuracy: bias always < 1.5%)

Conclusions

- 3D field-level SBI outperforms the $P(k)$ (FoM factor $\sim$ 2 - 6),
  but at a much higher computational cost
- 2D field-level SBI does not provide significant advantages wrt $P(k)$
- Combining different tracers outperforms single-tracers analyses (FoM factor $\sim$ 3)
- Next steps: move towards real surveys specifics (model noises, volumes, etc.)!

THANK YOU

FOR THE ATTENTION!

BACKUP SLIDES

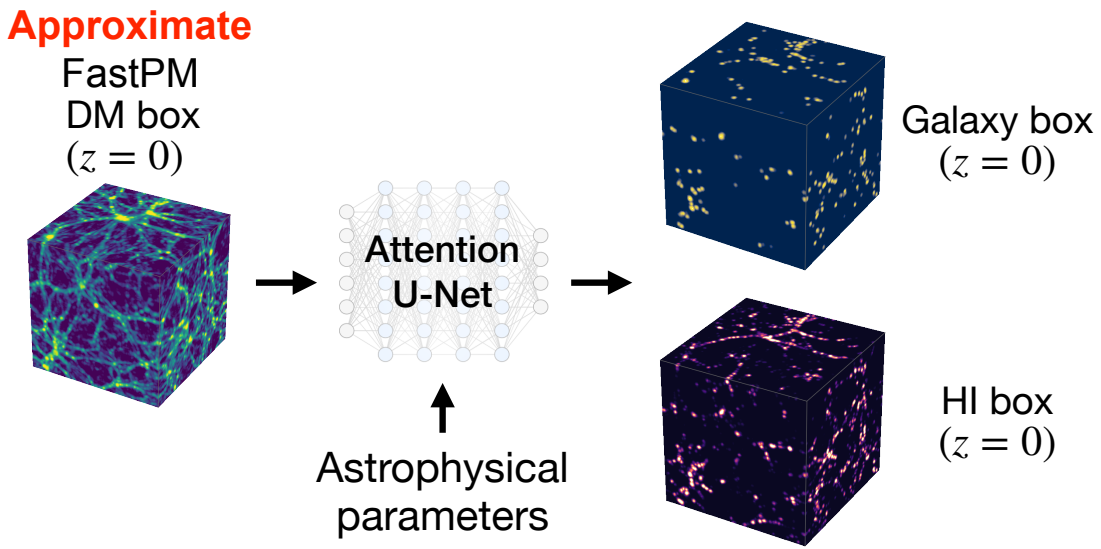
Emulating our sets

- How many simulations do we need for SBI in cosmology? [Bairagi+25]

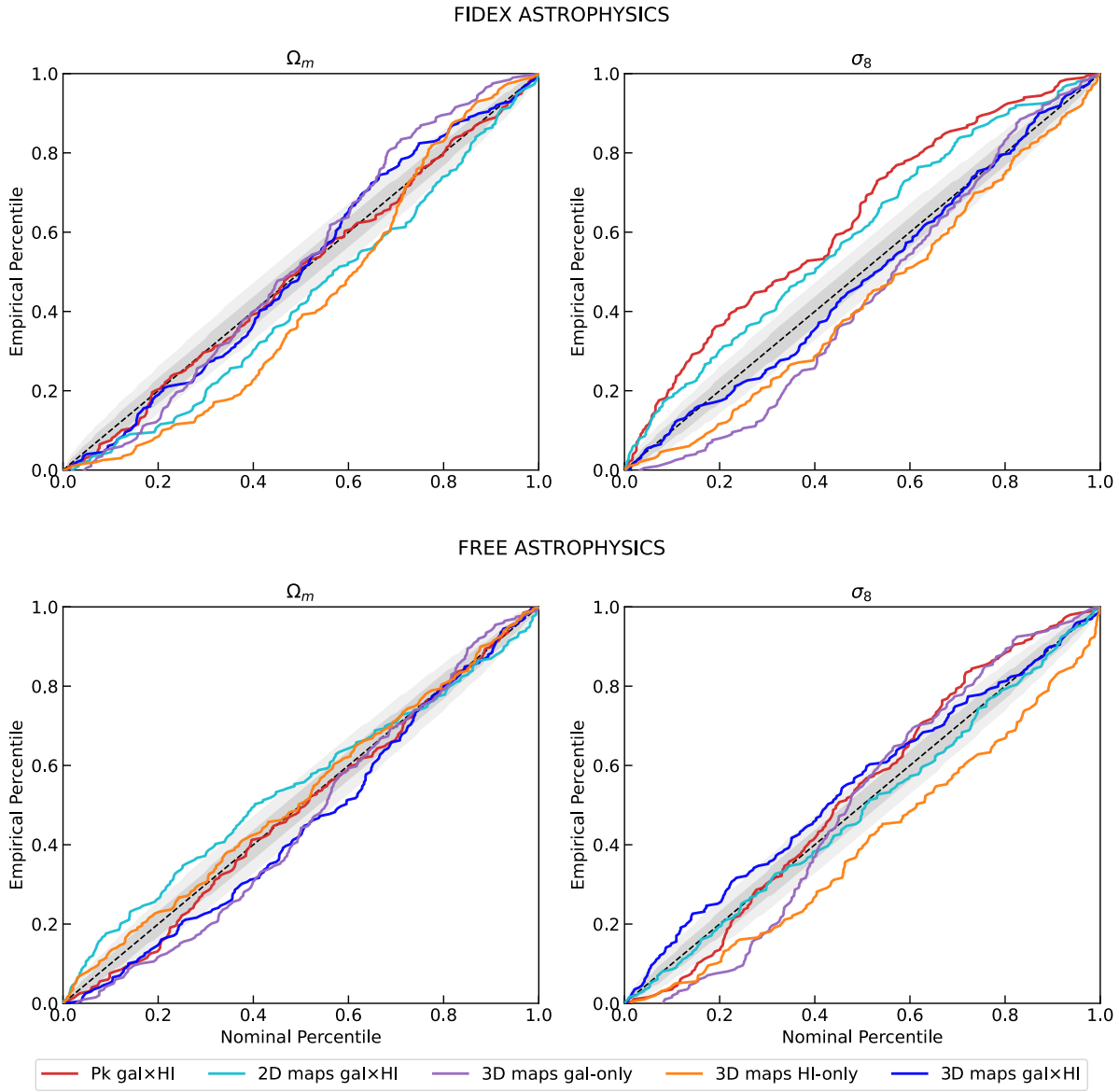
- Currently available simulation suites do not provide sufficient training data for a generic NN to reach optimal regime!

- **Emulators** can come into play!

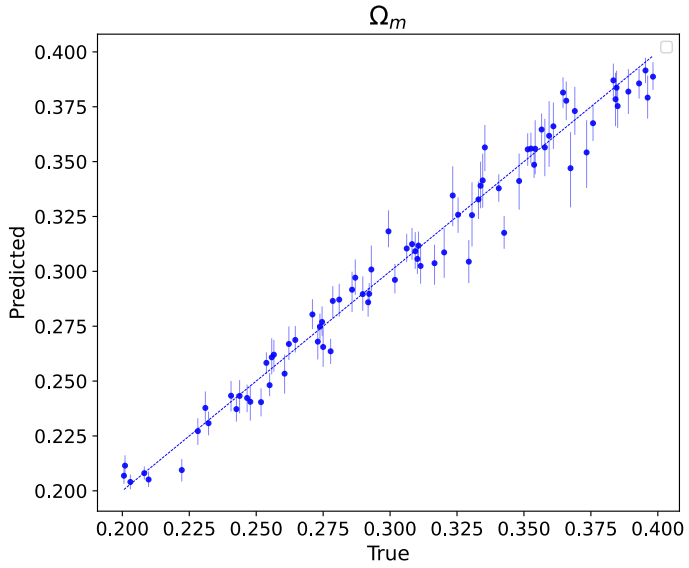
- We use: emulator adapted from Mishra+25, **from approximate DM maps to galaxy/HI fields**

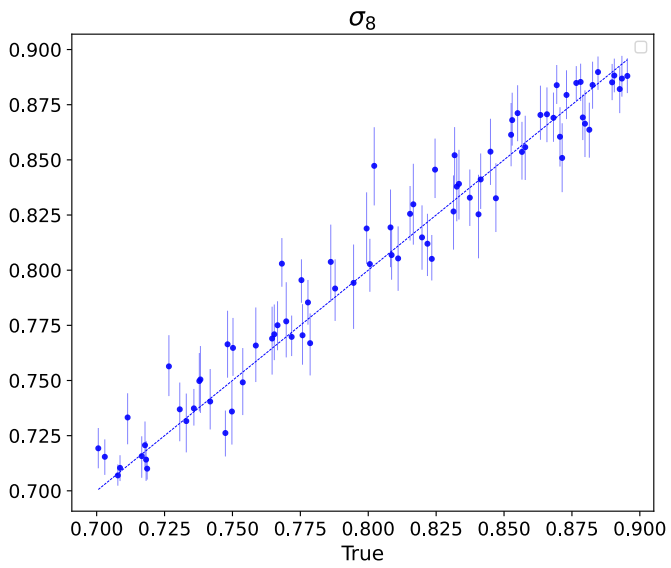


Calibration

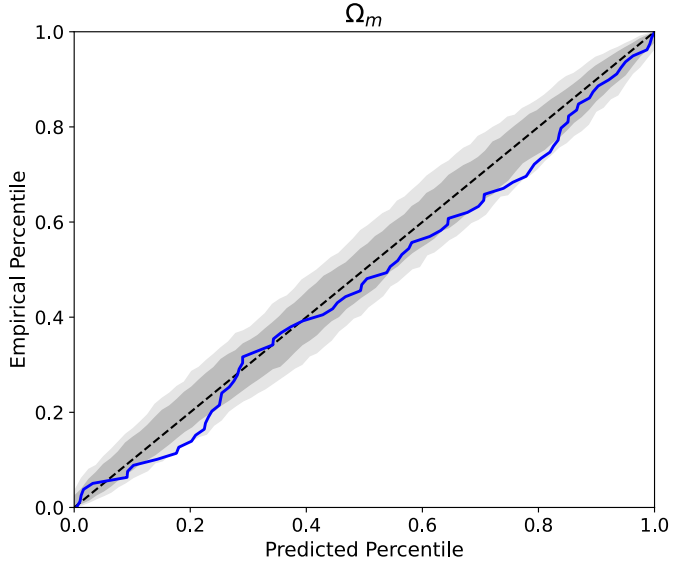


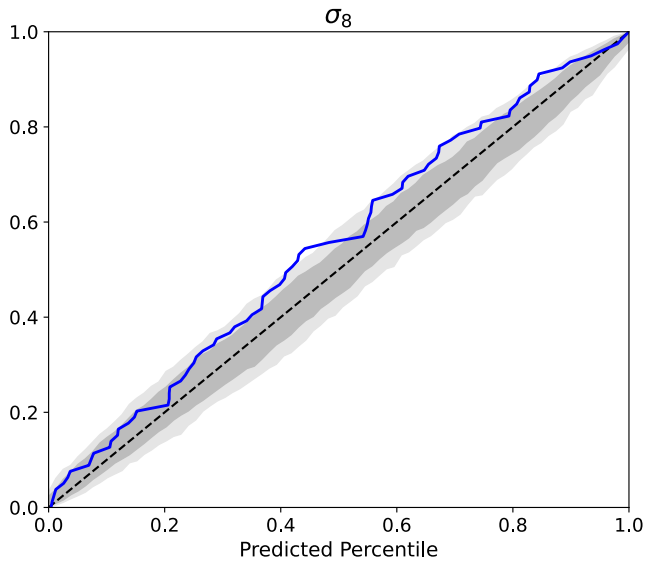
Different ground truths





gal × HI
3D maps
Free Astrophysics





We expect previous results to hold also for non-Planck ground truths