New Perspectives: 2022
FERMILAB-SLIDES-22-044-SCD

Estimating the Cosmological Parameters from Galaxy Cluster Abundance: an SBI approach

M. Reza[1,*], Y. Zhang[1] and B. Nord[2,3,4]

1. Department of Physics and Astronomy, Texas A&M University, USA
2. Fermi National Accelerator University, USA
3. Department of Astronomy and Astrophysics, University of Chicago, USA
4. Kavli Institute for Cosmological Physics, University of Chicago, USA

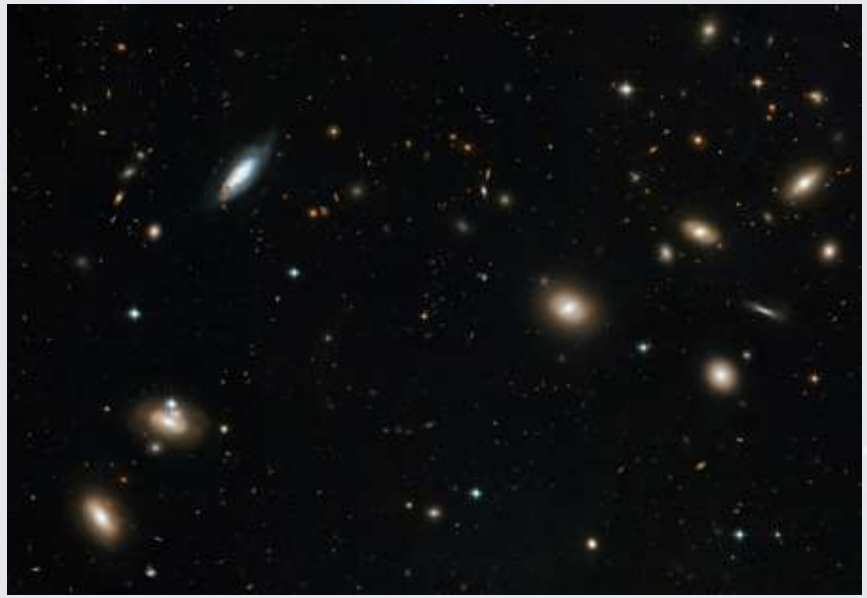
*Research supported by Texas A&M University and Deep Skies Lab
*Funding by Texas A&M University
Contact: moonzarin@tamu.edu

GALAXY CLUSTERS

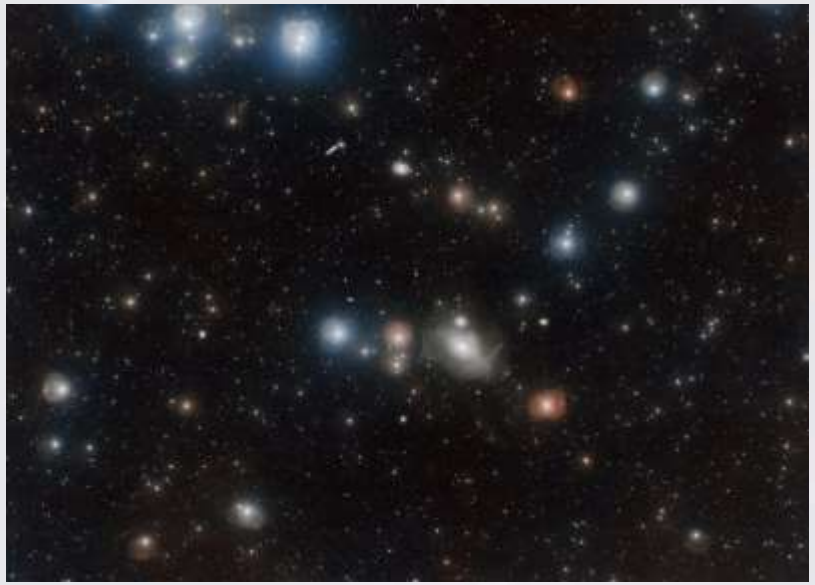
Largest known gravitationally-bound structures in the Universe

Mass range 10^{14} - $10^{15}\,M_{\odot}$

Contains hundreds to thousands of galaxies



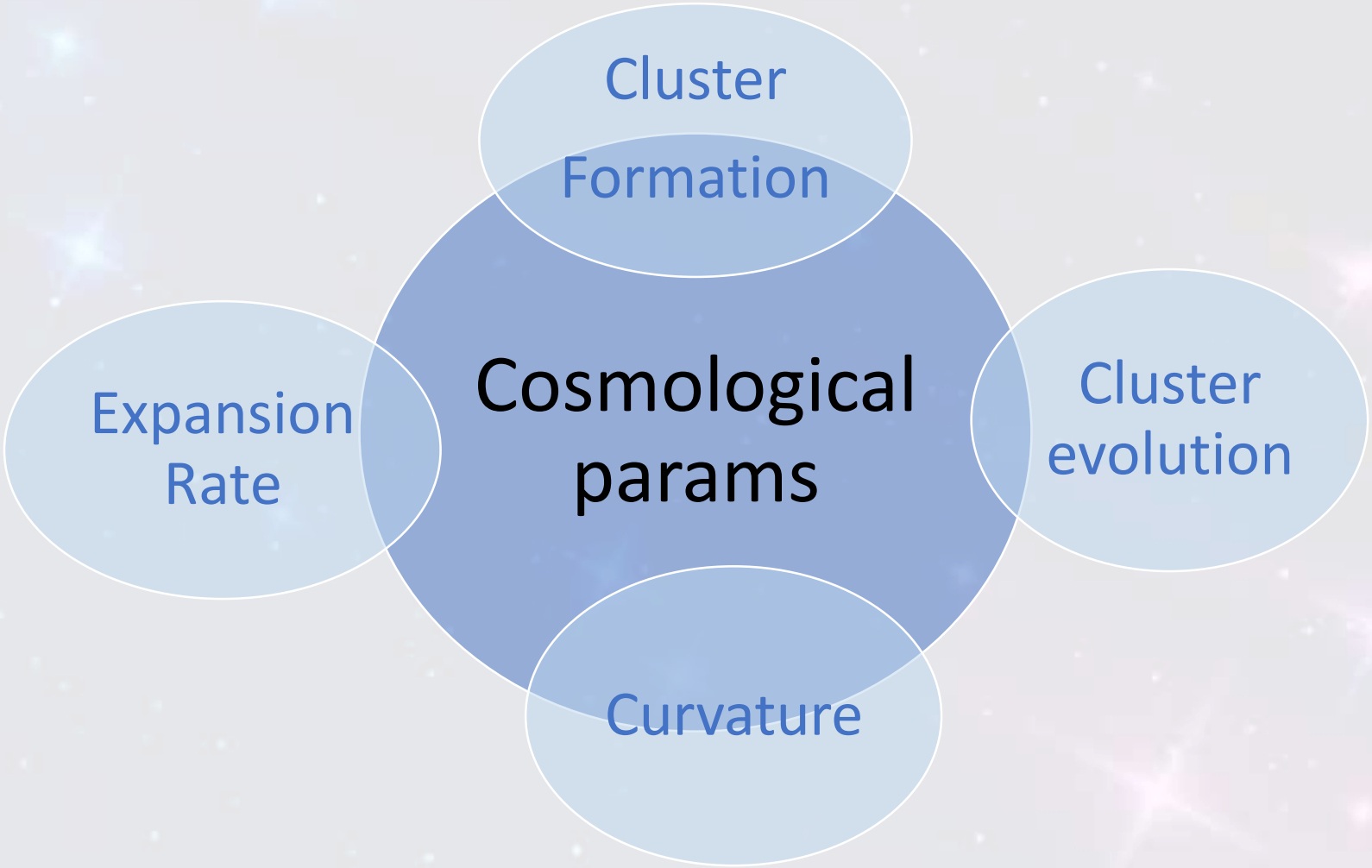
Coma cluster (NASA, ESA and HST)



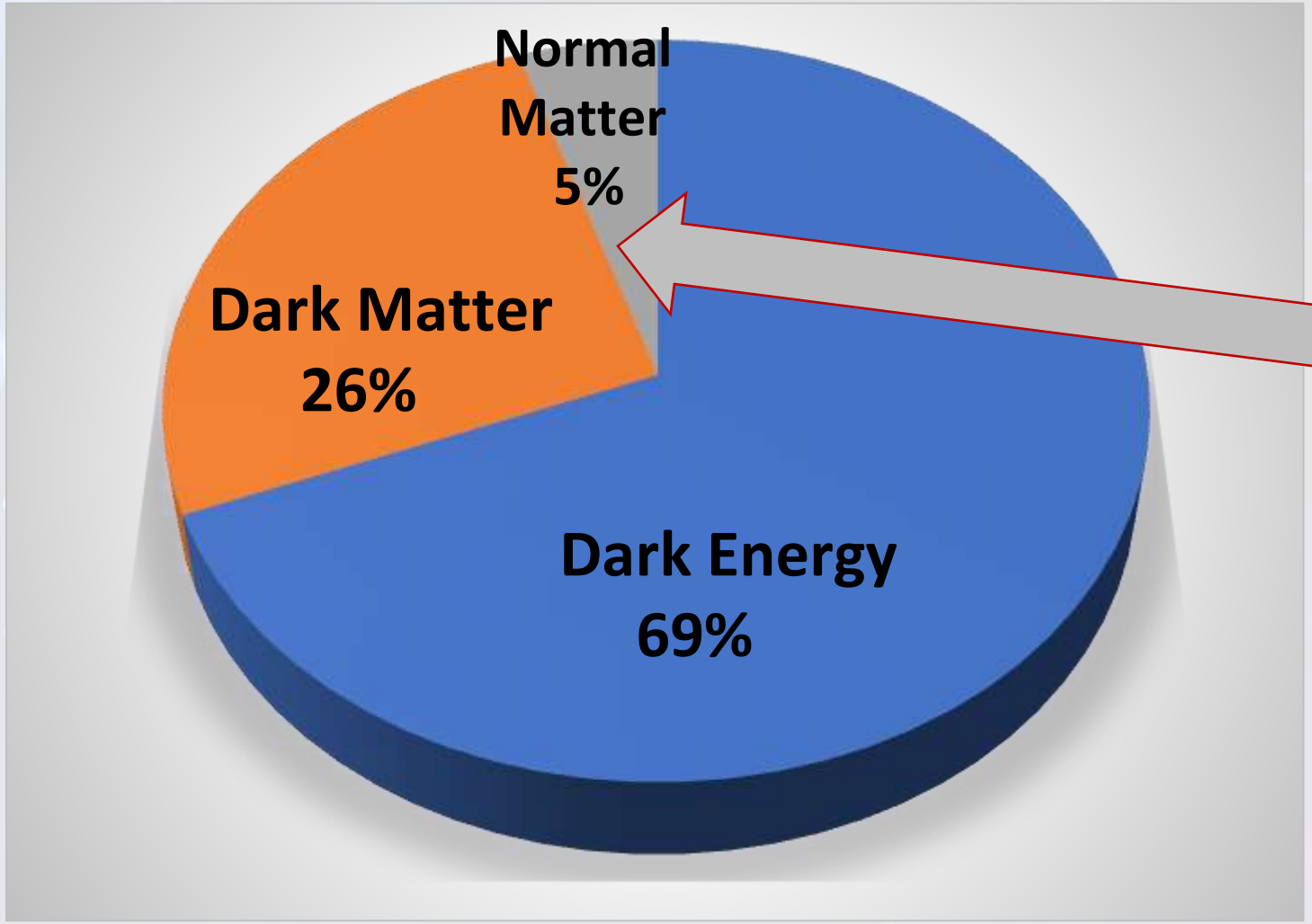
Fornax cluster (VLT survey telescope and ESO)

COSMOLOGICAL PARAMETERS

Galaxy Clusters can constrain **cosmology**



MASS-ENERGY DISTRIBUTION OF THE UNIVERSE



We see only 5% of the Universe!

INPUTS AND OUTPUTS

Halo (simplified galaxy cluster problem)

- ❑ Baryonic density (Ω_b)
- ❑ Matter density (Ω_m)
- ❑ Hubble's constant (h)
- ❑ Power law index (n_s)
- ❑ Amplitude fluctuation (σ_8)

Cluster (observational applications, includes mass-richness relation)

$$\ln \boldsymbol{\lambda} = M_A + M_B \ln(M), \text{ with scatter } \sigma$$

- M_A
- M_B
- lnsigma (in addition to the 5 params)

INPUTS (Observables)

- Mass of galaxy clusters
- Number of galaxy clusters

OUTPUTS

Cosmological parameters

MARKOV CHAIN MONTE CARLO (MCMC)

Used to:

draw samples from high-dimensional pdf

Disadvantages:

Dependent chains

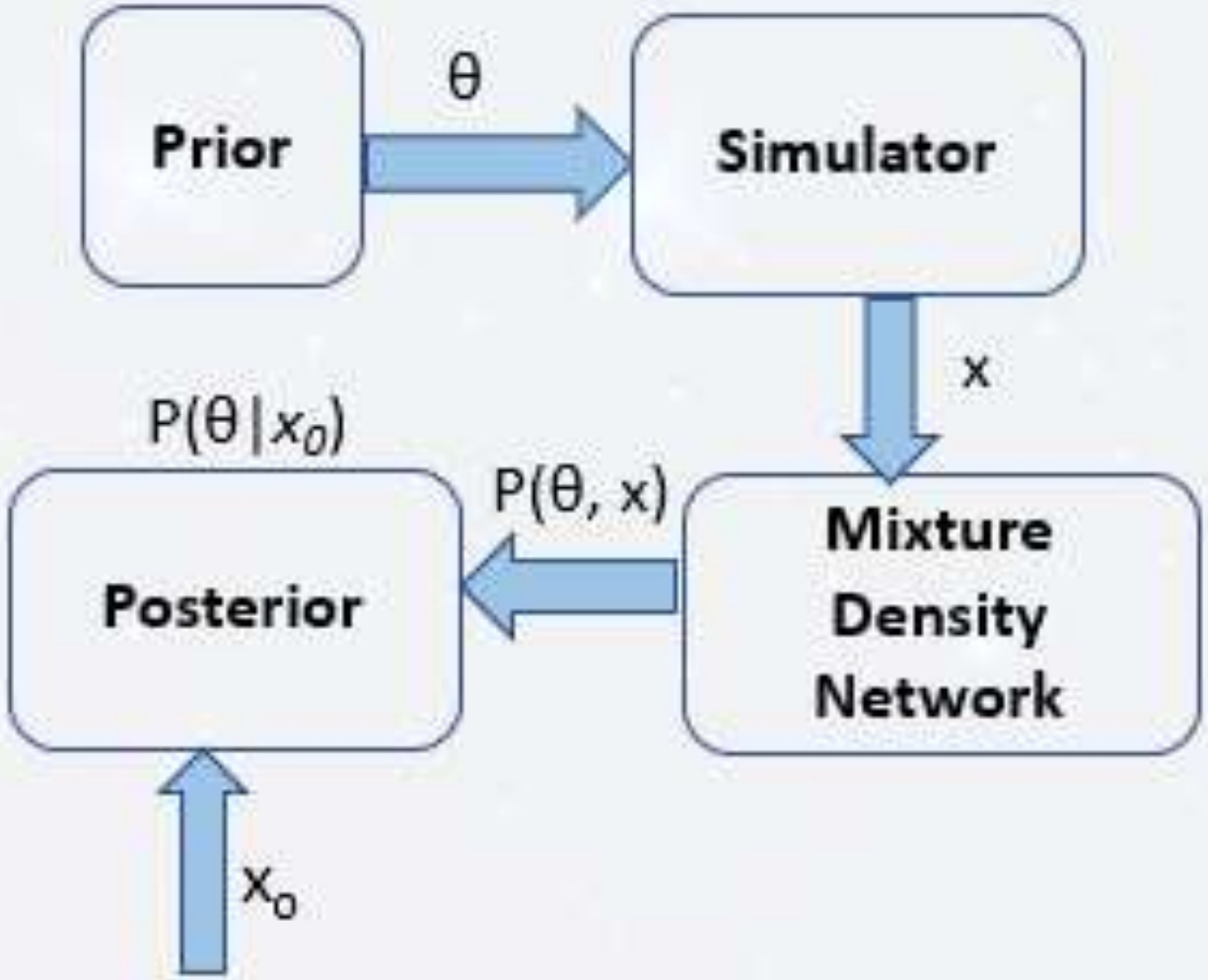
Intractable

Sampling Strategies:

Metropolis-Hastings

Gibbs Sampling

SIMULATION-BASED INFERENCE



Working principle of SBI method

SBI VS MCMC

SBI

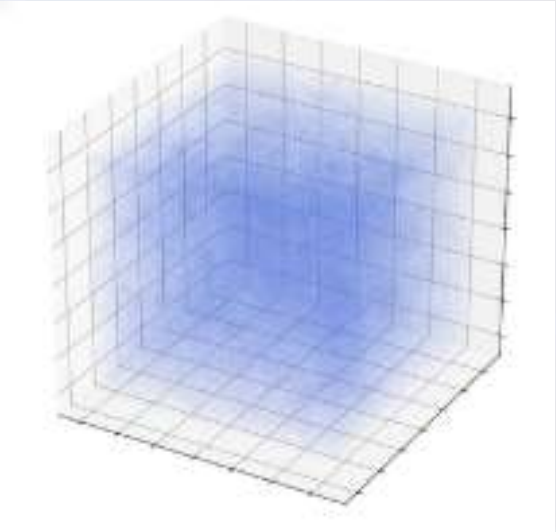
- Easy to incorporate complex physical models
- Precomputes mock observables for training
- Convenient speed

MCMC

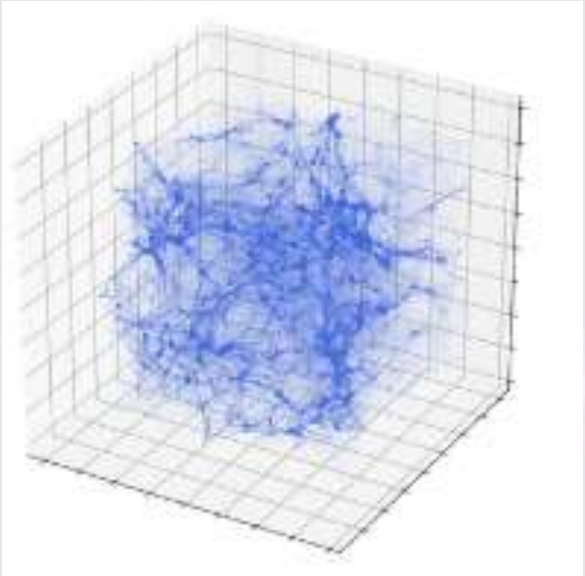
- Difficult to include complex phenomena
- Calculates likelihood during runtime
- Much slower

QUIJOTE SIMULATIONS

N-body simulations: large scale structure formation



Initial conditions of the Universe



Large scale structures

Credit: TensorFlow Blog

Characteristics:

- 44,100 full N-body simulations
- Snapshots at multiple redshifts

Purpose:

- Constrain cosmological models
- Provide statistics to train ML algorithms

QUIJOTE SIMULATIONS

Training Set:

Latin Hypercube Simulations

(Variable Cosmology)

Ω_m : [0.1 − 0.5]
Ω_b : [0.03 − 0.07]
h : [0.5 − 0.9]
n_s : [0.8 − 1.2]
σ_8 : [0.6 − 1.0]

Test Set:

Fiducial Planck Cosmology

(Fixed Cosmology)

Ω_m : 0.318
Ω_b : 0.049
h : 0.671
n_s : 0.962
σ_8 : 0.834

ANALYTICAL SIMULATIONS

Analytical simulations are generated by **Cosmosis:**

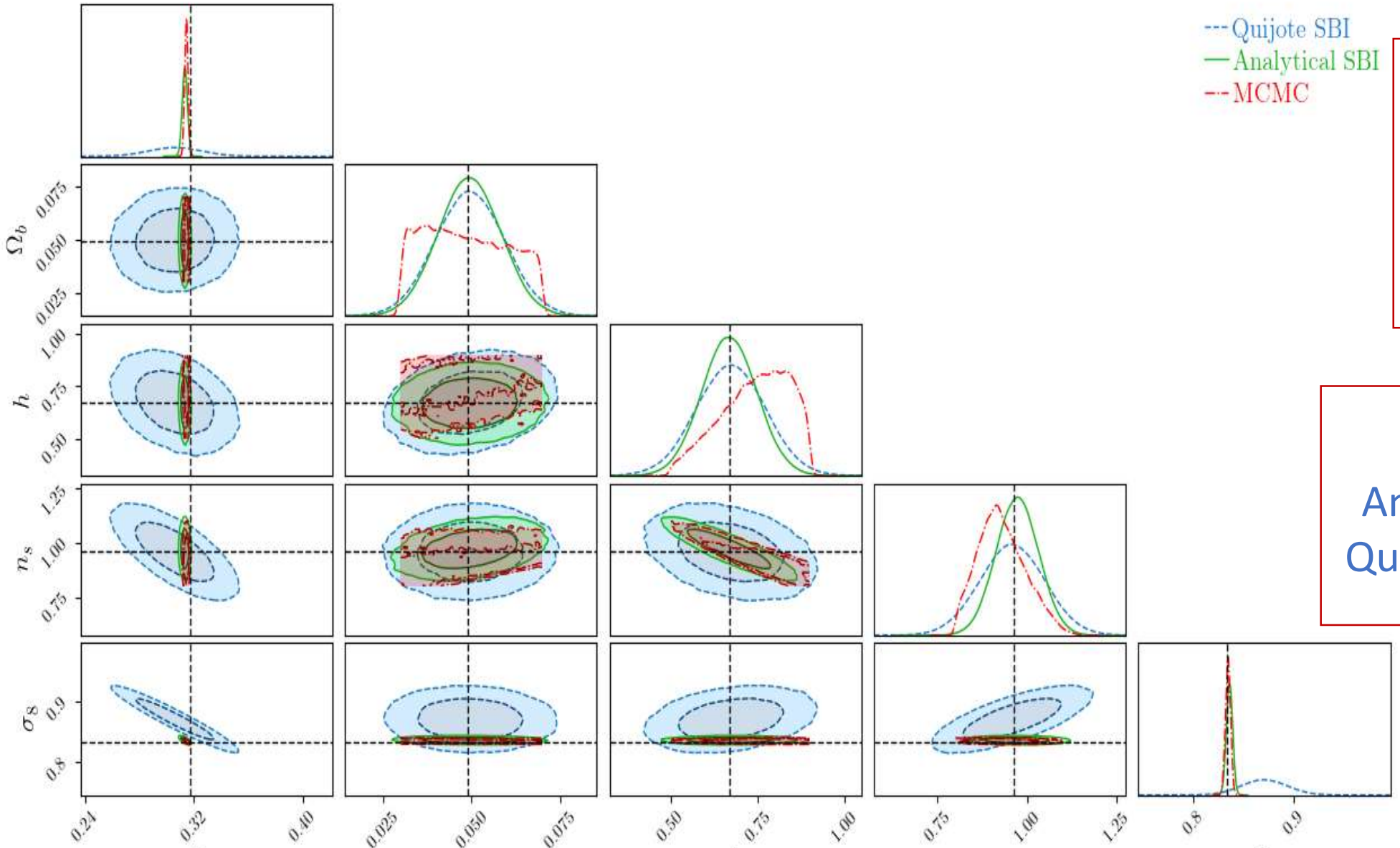
$$\mathbf{N} = \mathrm{v} \int n\,(m, z = 0) |\ \Omega_m, \sigma_8)$$

$$\mathbf{M} = \mathrm{v} \int n\,(m, z = 0) |\ \Omega_m, \sigma_8) * m$$

--- N and M are the mock observables used for training

--- n is the cluster number density function (depends on the model)

RESULTS [Posterior Distribution]



Truth

SBI within 2σ

Bias

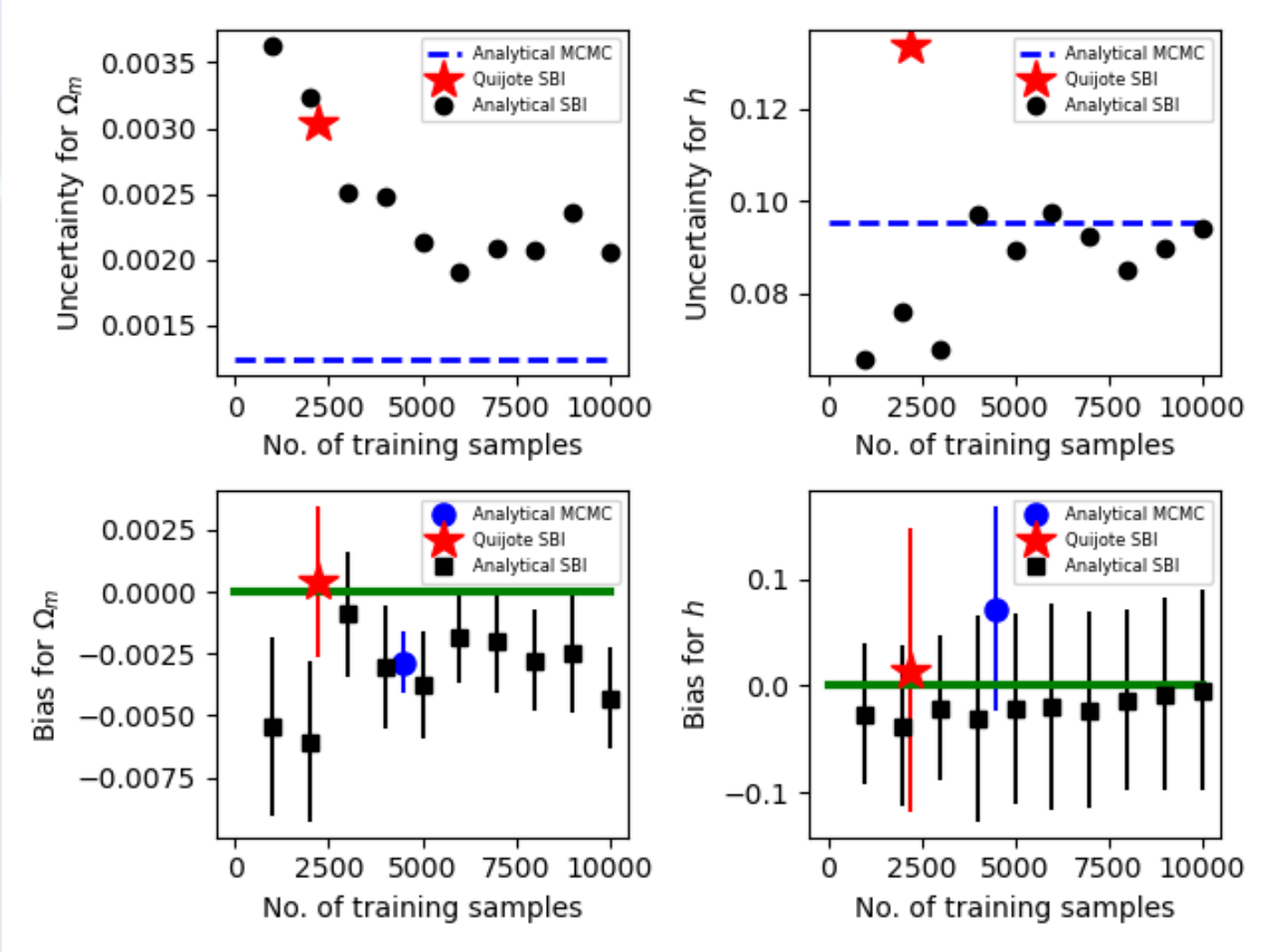
MCMC > SBI

Uncertainty

Analytical ≈ MCMC

Quijote is the largest

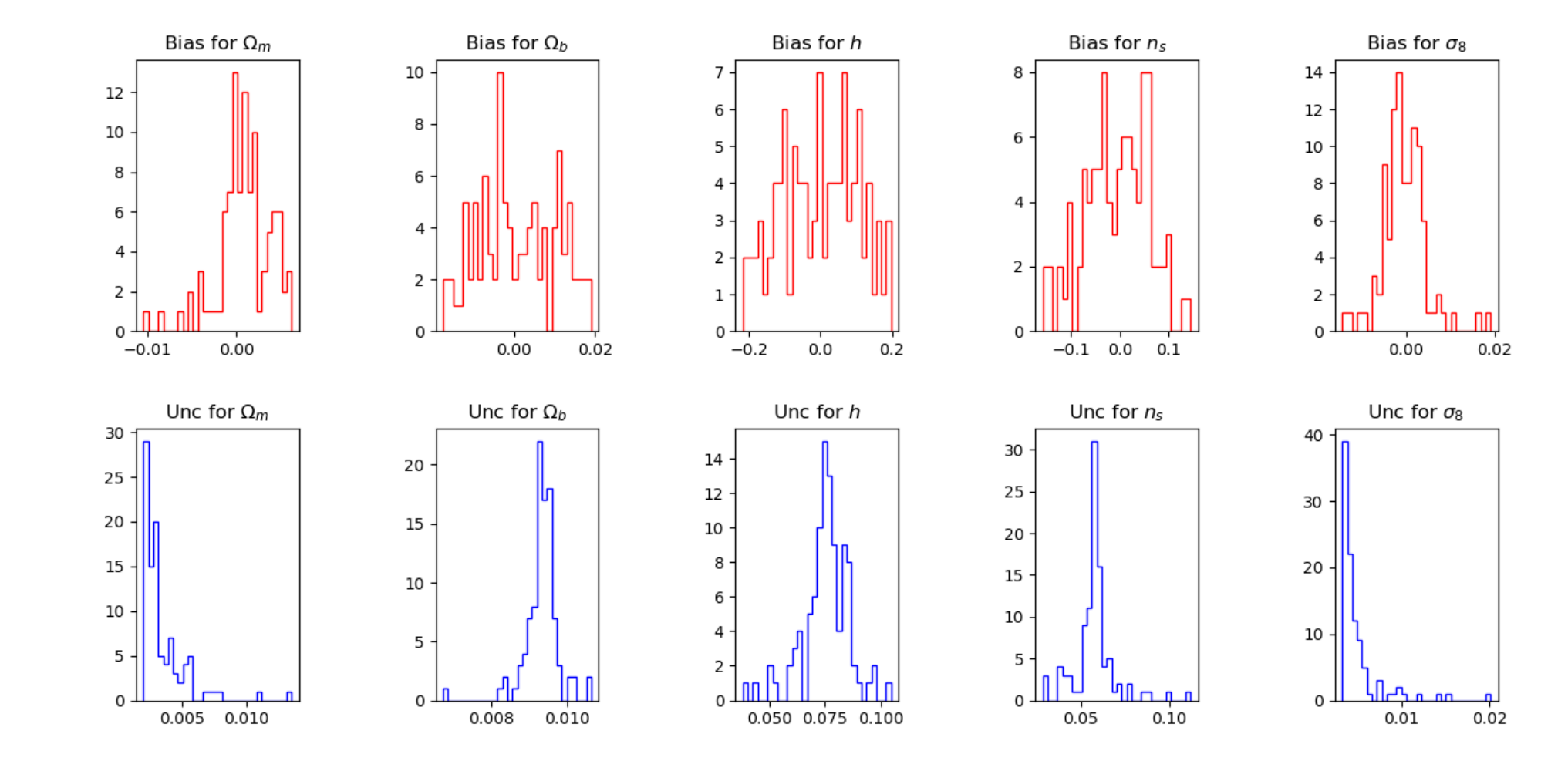
RESULTS [Training Sample Size]



Sample size ↑,
Unc ↓ for Ωm
No trend for h

Sample size ↑,
Bias ↓ for Ωm
Constant bias for h

RESULTS [Analytical Simulations]



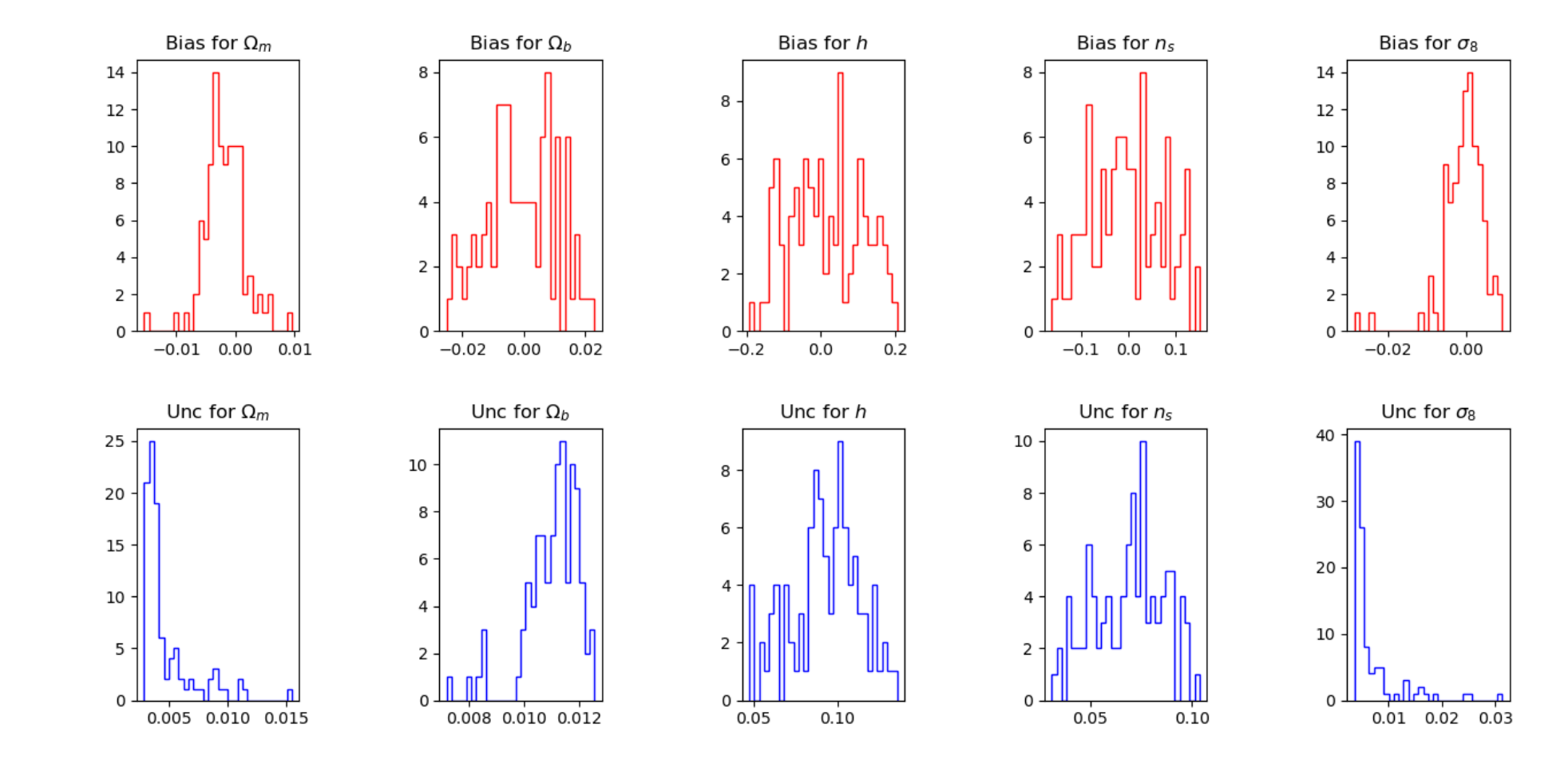
Bias

Single peak around zero for Ω_b and σ_8

Uncertainty

Lowest uncertainty for Ω_m and σ_8

RESULTS [Quijote Simulations]



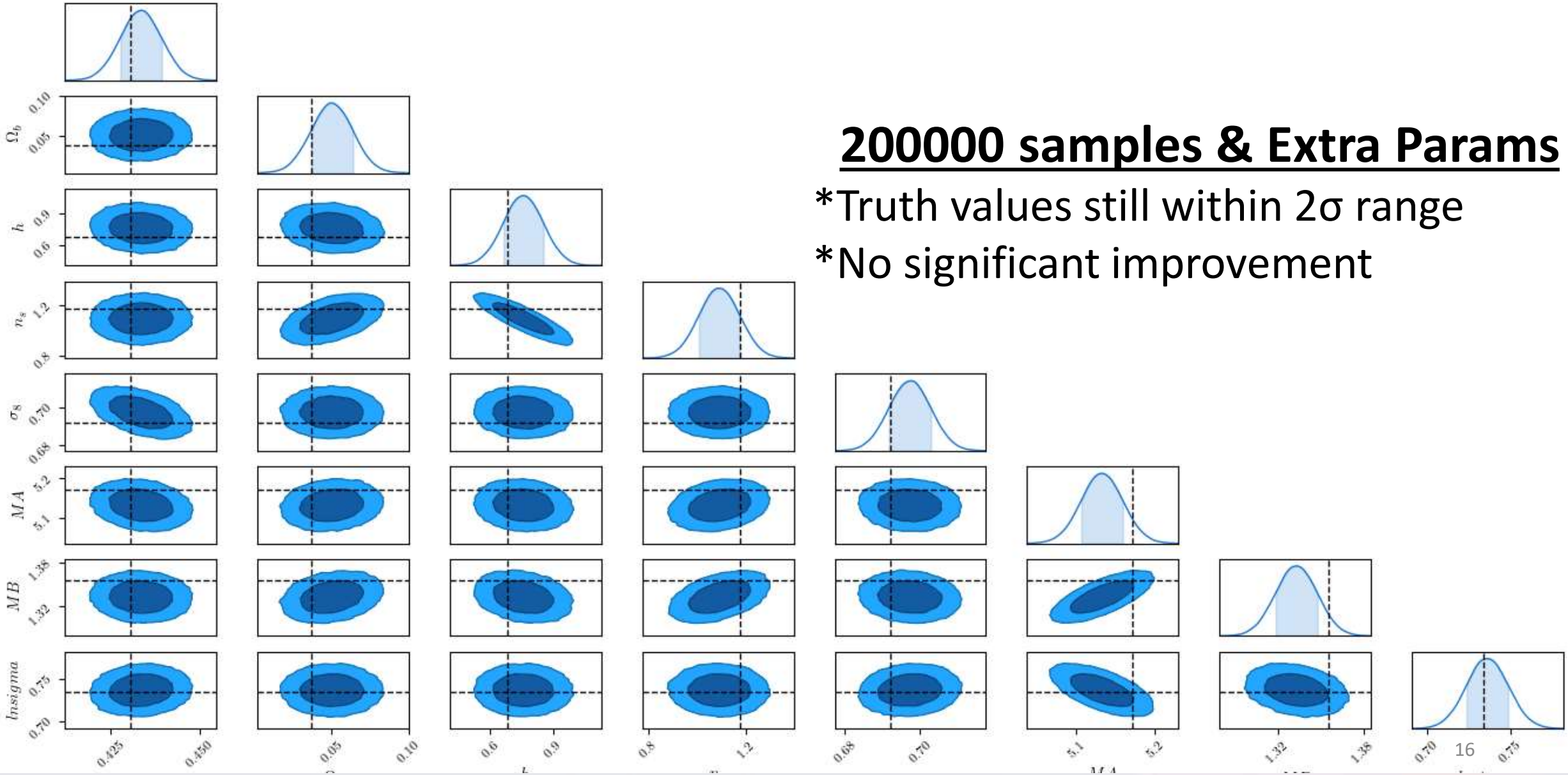
Bias

Dominant peak around zero for Ω_m, h & σ_8

Uncertainty

Lowest uncertainty for Ω_m & σ_8

RESULTS [Cluster]



200000 samples & Extra Params

*Truth values still within 2σ range

*No significant improvement

Conclusions

Simulation-based Inference (SBI)

- Recovers cosmological params within 95% CI
- Smaller bias than MCMC
- Faster than MCMC

Thank You

Any Questions??

Contact: moonzarin@tamu.edu