21st International Workshop on Advanced Computing and Analysis Techniques in Physics Research

From Galaxy Clusters to Cosmic Evolution with Simulation-Based Inference

Presented by: Moonzarin Reza[1,*]

Collaborators: Y. Zhang[2], B. Nord[3,4], L. Strigari[1], A. Ćiprijanović[3], J. Poh[4], S. Shevchek[3]

1. Department of Physics and Astronomy, Texas A&M University, USA
2. NSF's NOIRLab, Tucson, Arizona
3. Fermi National Accelerator University, USA
4. Department of Astronomy and Astrophysics, University of Chicago, USA

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

Galaxy Cluster SMACS 0723

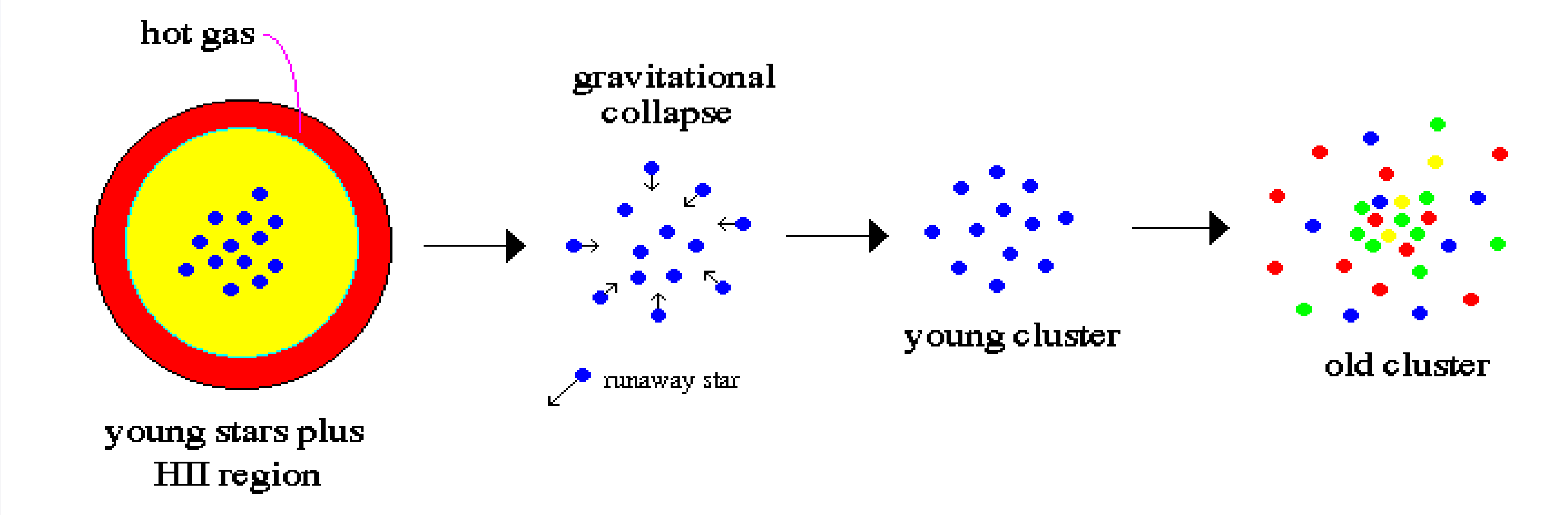
Credit: James Webb Space Telescope

Cosmological Parameters

Parameters: Ω_m & σ_8

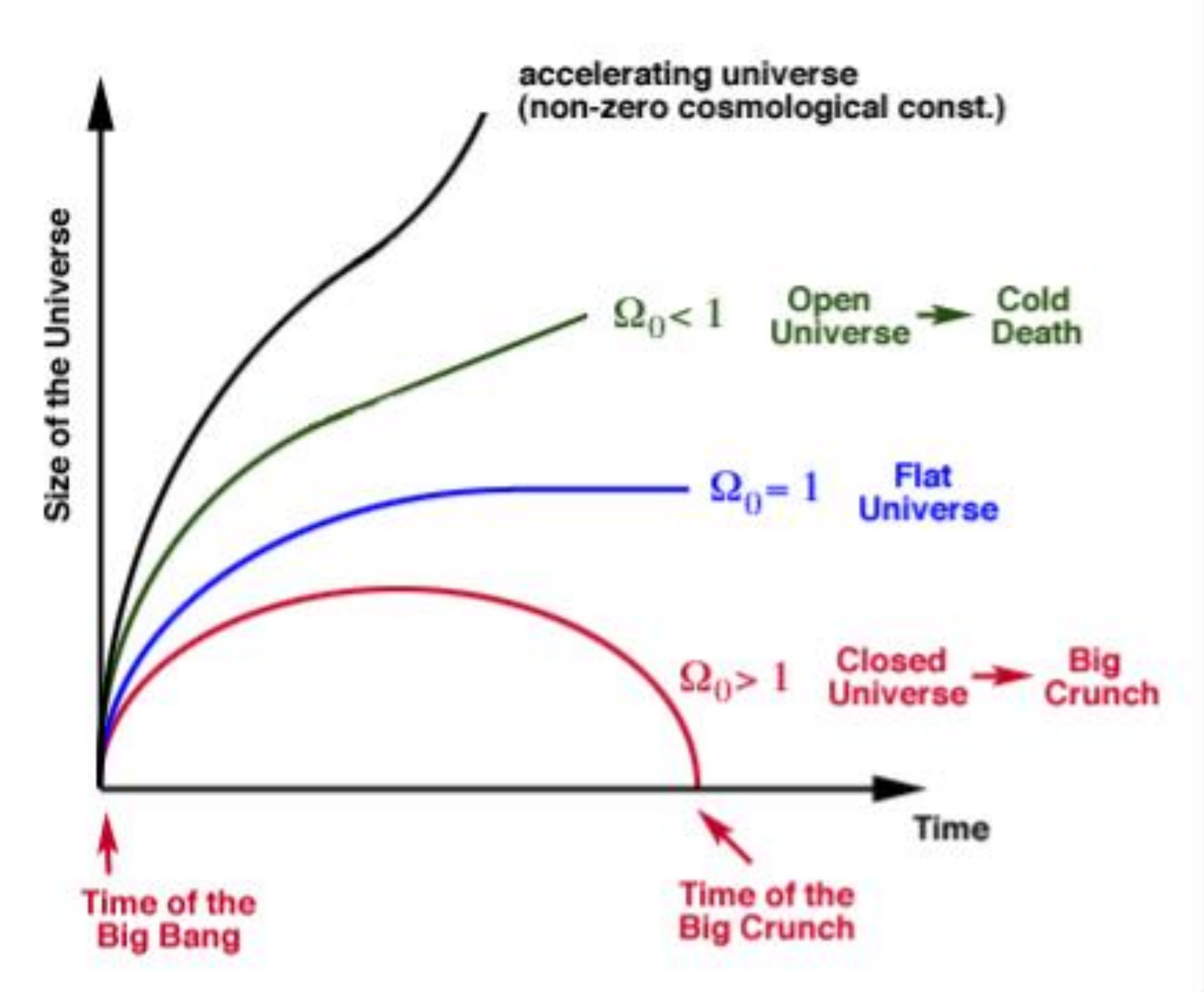
Galaxy Cluster Formation and Evolution

Credit: James Webb Space Telescope



Credit: abyss.uoregon.edu

Cosmological Parameters



Parameter: H_0

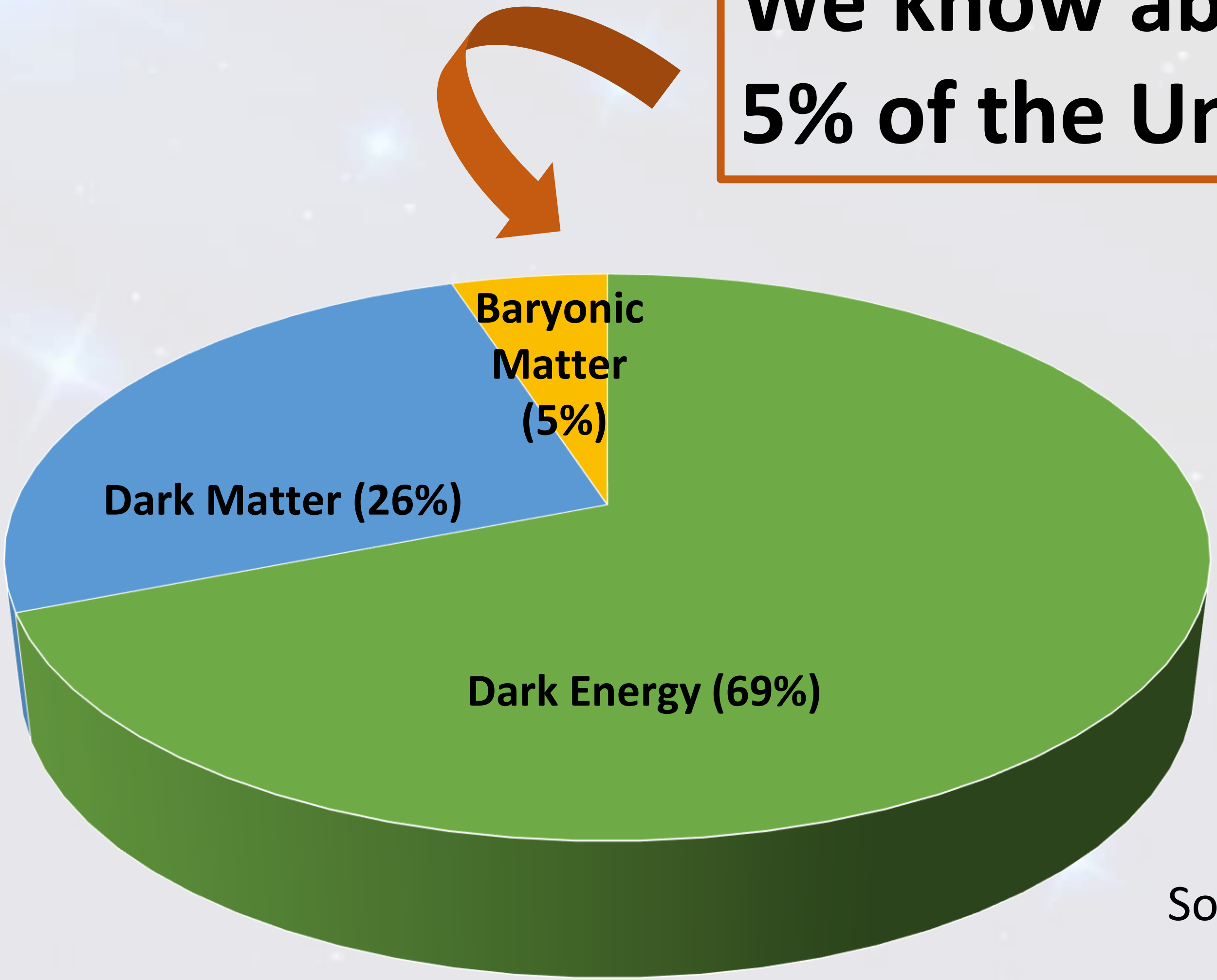
Ultimate Fate of the Universe

$$\Omega_0 = \Omega \,/\, \Omega_c$$

Credit: https://astro.umass.edu

Mass-Energy Distribution of the Universe

We know about only 5% of the Universe!



Source: National Radio Astronomy Observatory

Inputs and Outputs

Inputs (Observables)

- Average mass of galaxy clusters
- Total number of galaxy clusters

Outputs (Parameters)

- Cosmic (Simplified Halo)
- Cosmic and Astrophysical (Cluster)

Cosmological Parameters

- ❑ Baryonic density (Ω_b)
- ❑ Matter density (Ω_m)
- ❑ Hubble's constant (H_0)
- ❑ Power law index (n_s)
- ❑ Amplitude fluctuation (σ_8)

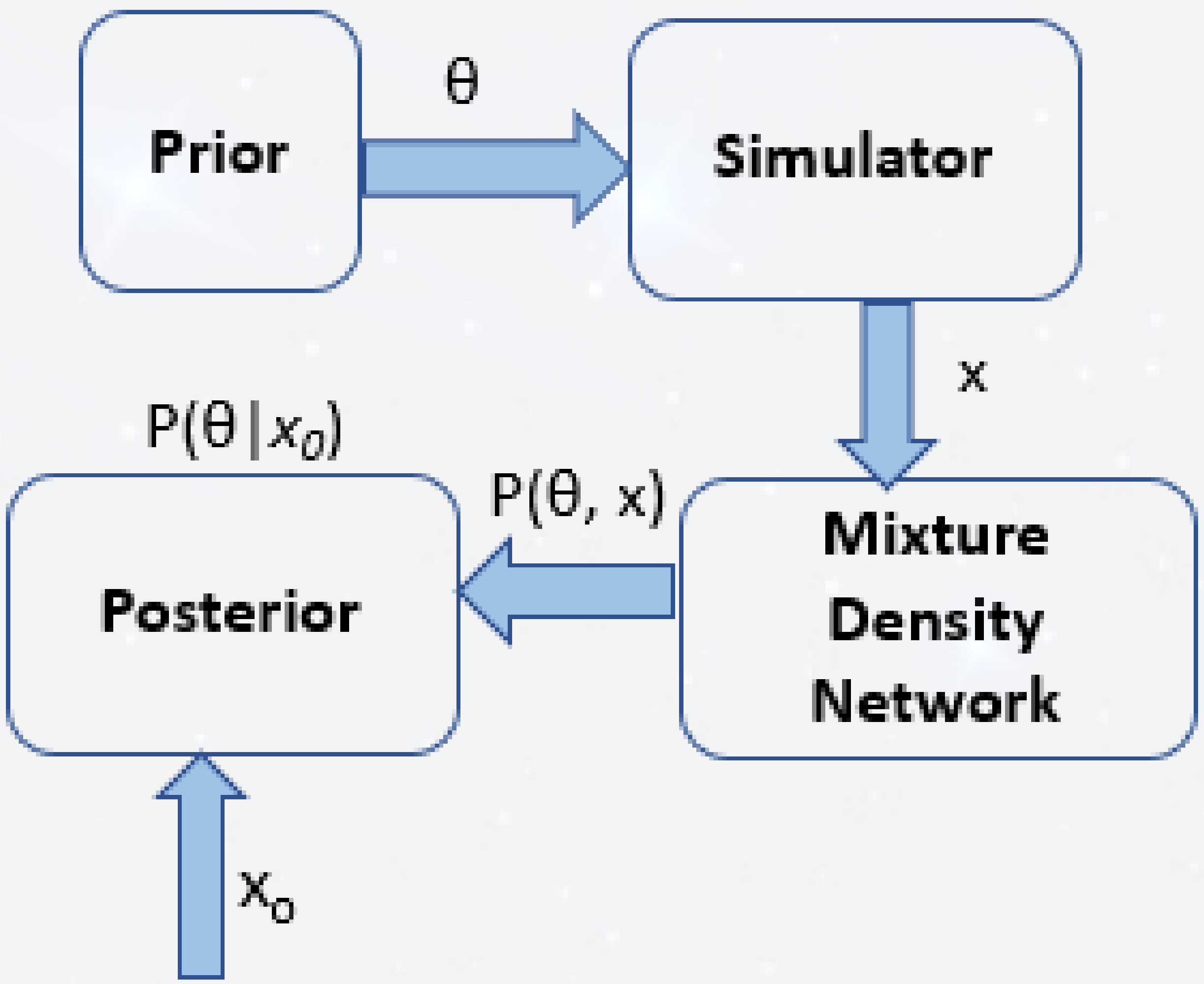
Astrophysical Parameters

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

- M_A
- M_B
- σ

(Murata et al. 2017)

Simulation-based Inference (SBI)

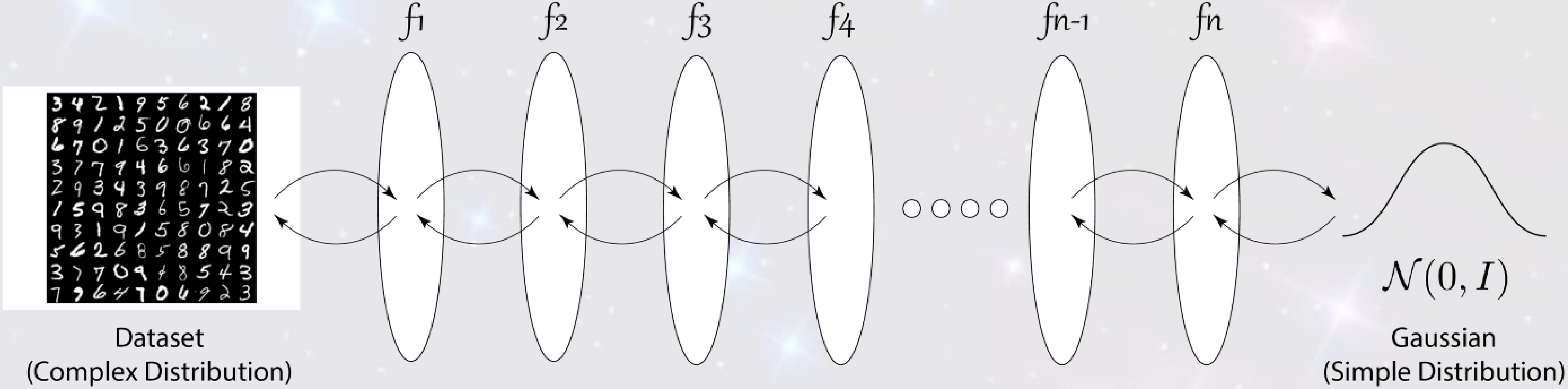


Working principle of SBI method

Normalizing Flow-based SBI

✓Consists of a series of simple invertible functions

✓Used for complex data representation



Markov Chain Monte Carlo (MCMC)

Used to:

draw samples from high-dimensional pdf

Disadvantages:

explicit likelihood dependence

dependent chains

Sampling Strategies:

Metropolis-Hastings

Gibbs Sampling

SBI vs MCMC

SBI

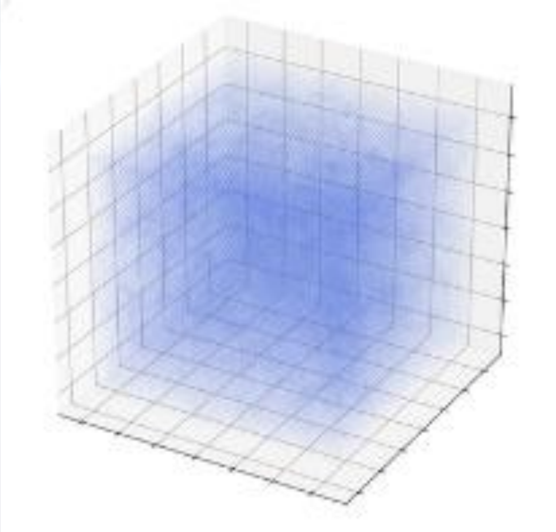
- Easy to incorporate complex physical models
- Precomputes mock observables for training
- Convenient speed

MCMC

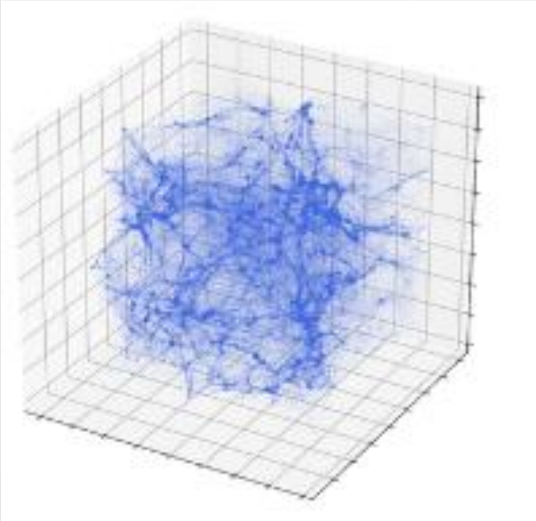
- Difficult to include complex phenomena
- Calculates likelihood during runtime
- Much slower

Dataset 1: Quijote Simulations

N-body simulations: Large Scale Structure Formation



Early Universe



Late Universe

Credit: TensorFlow Blog

Villaescusa-Navarro, F., et al. "The Quijote Simulations, " *arXiv:1909.05273* (2019).

Characteristics:

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

Purpose:

- Constrain cosmological models
- Provide statistics to train ML algorithms

Dataset 1: 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

Dataset 2: Analytical Simulations

Analytical simulations are generated by **CosmoSIS:**

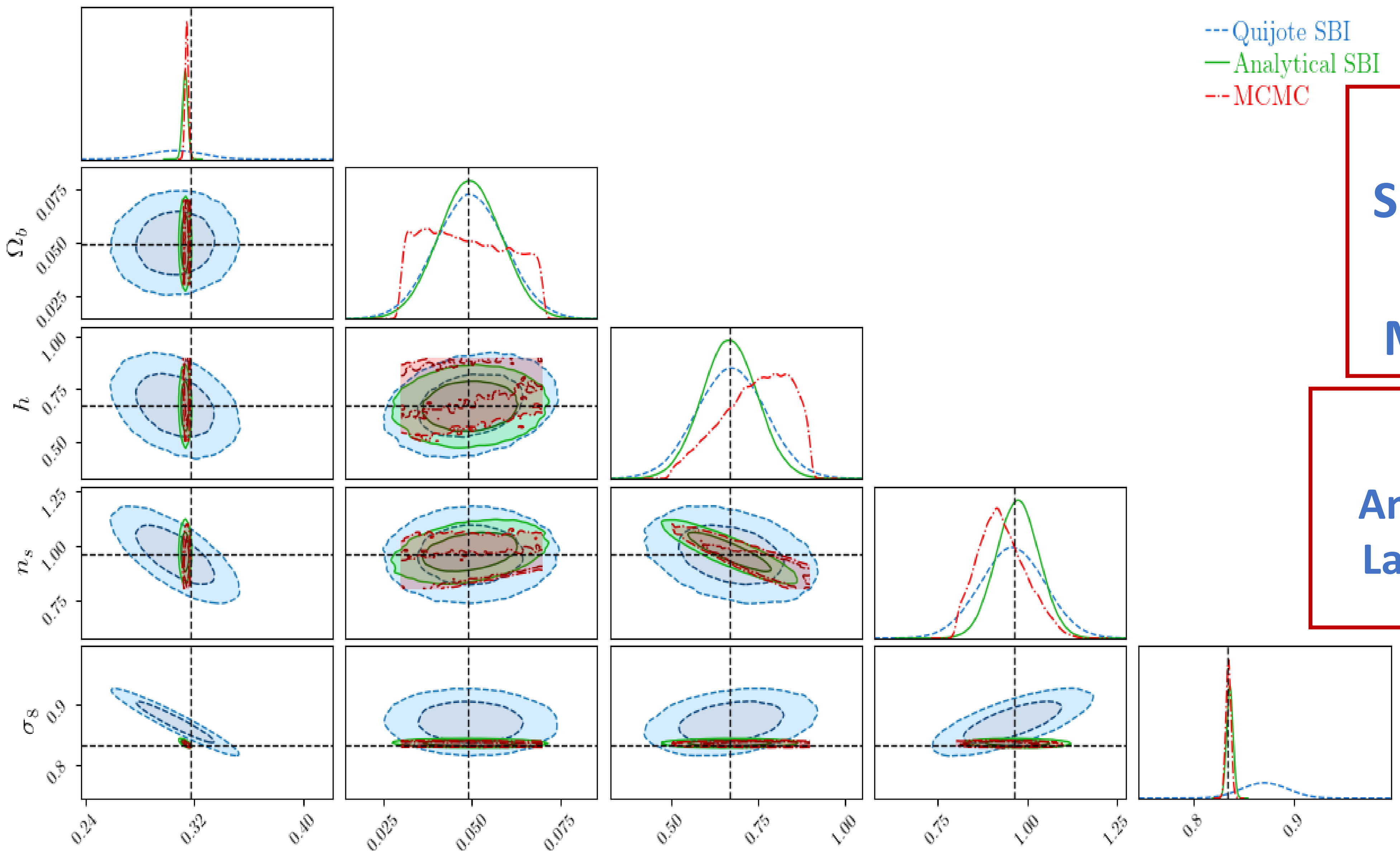
$$\mathbf{N} = \mathrm{v} \int n\,(m, z = 0) |\ \Omega_m\,, \sigma_8)$$

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

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

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

Results [Posterior Distribution]



Truth

SBI within 2σ

Bias

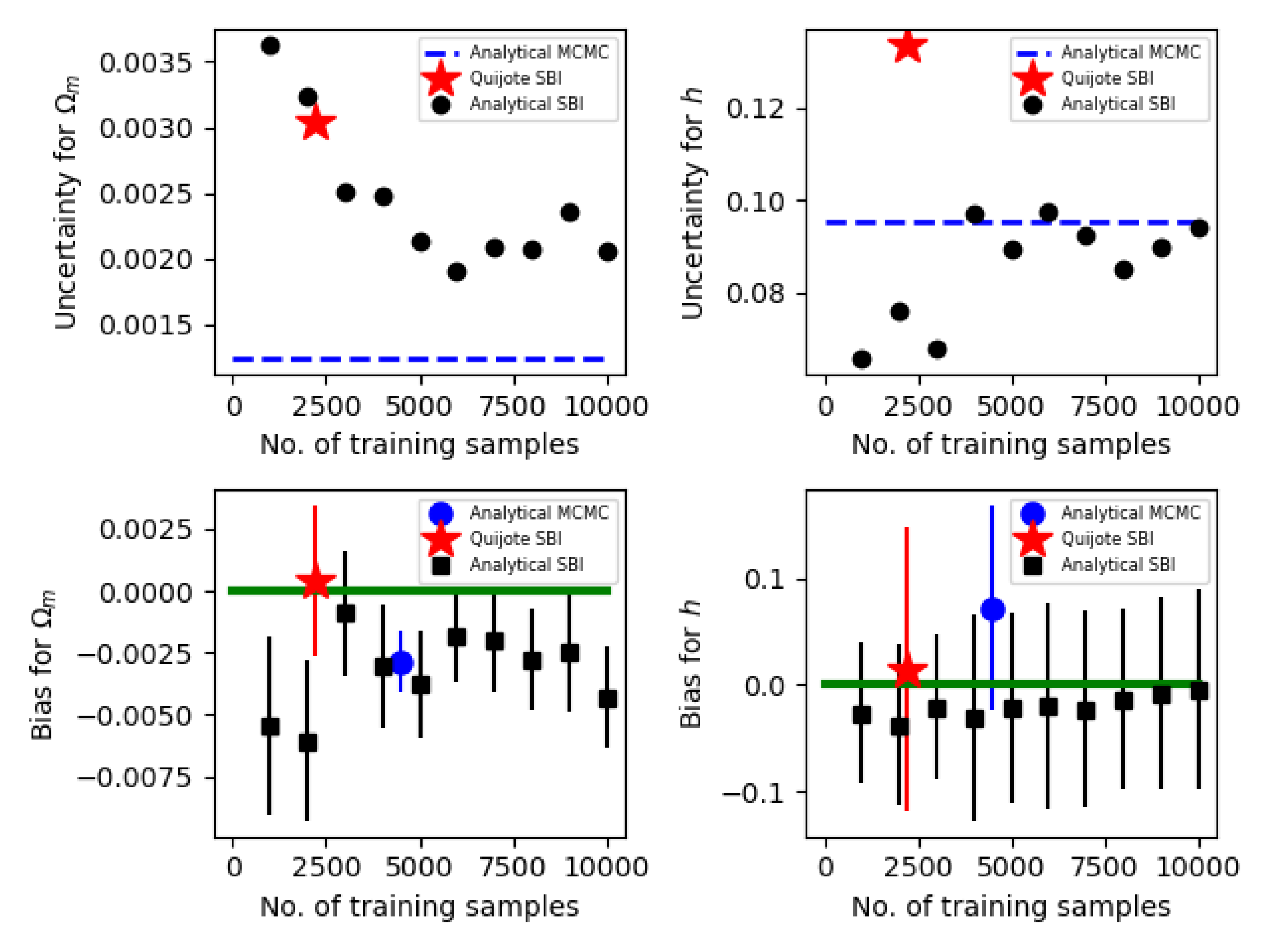
MCMC > SBI

Uncertainty

Analytical ≈ MCMC

Largest for Quijote

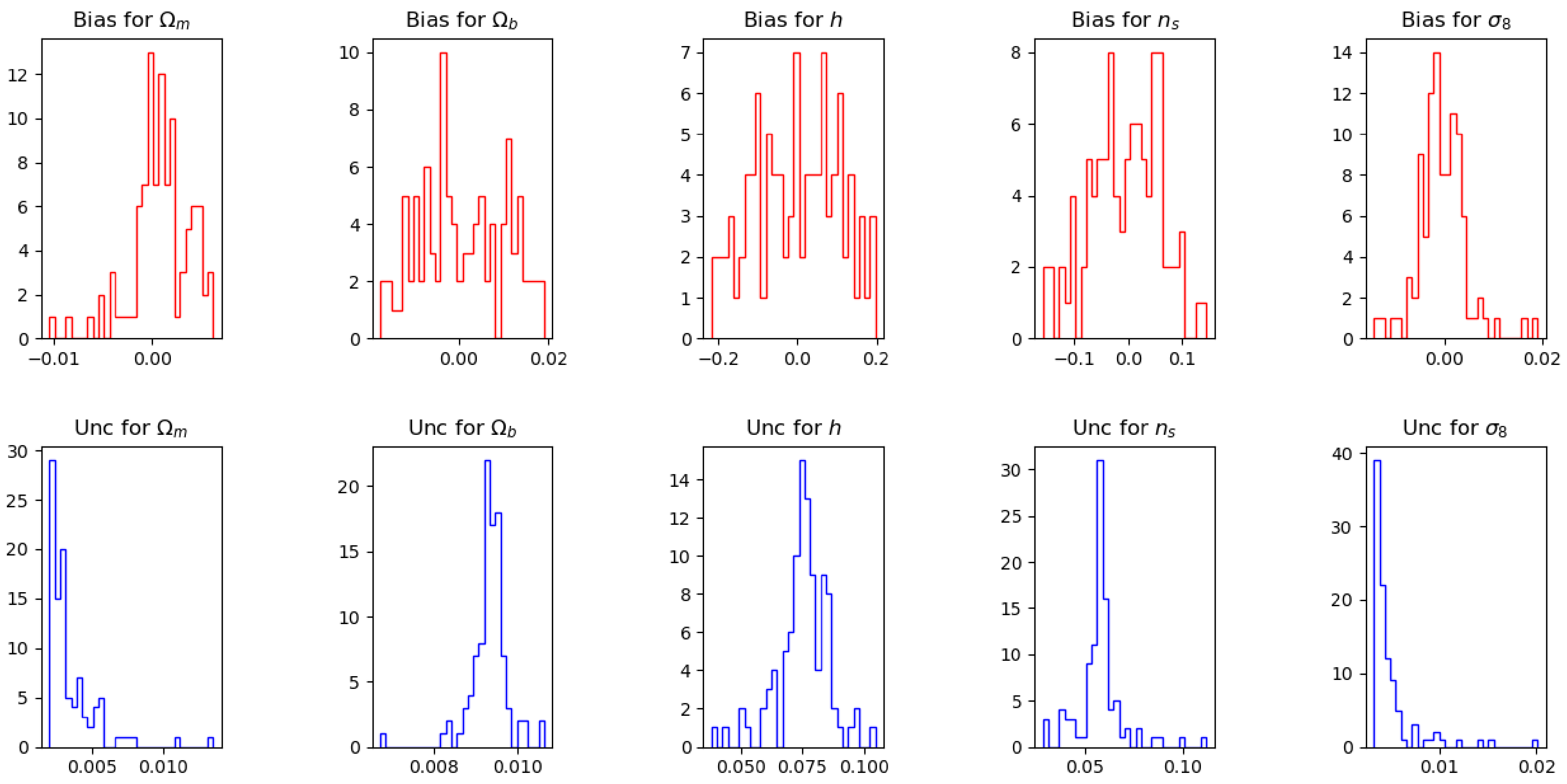
Results [Training Sample Size]



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

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

Results [Analytical Simulations]



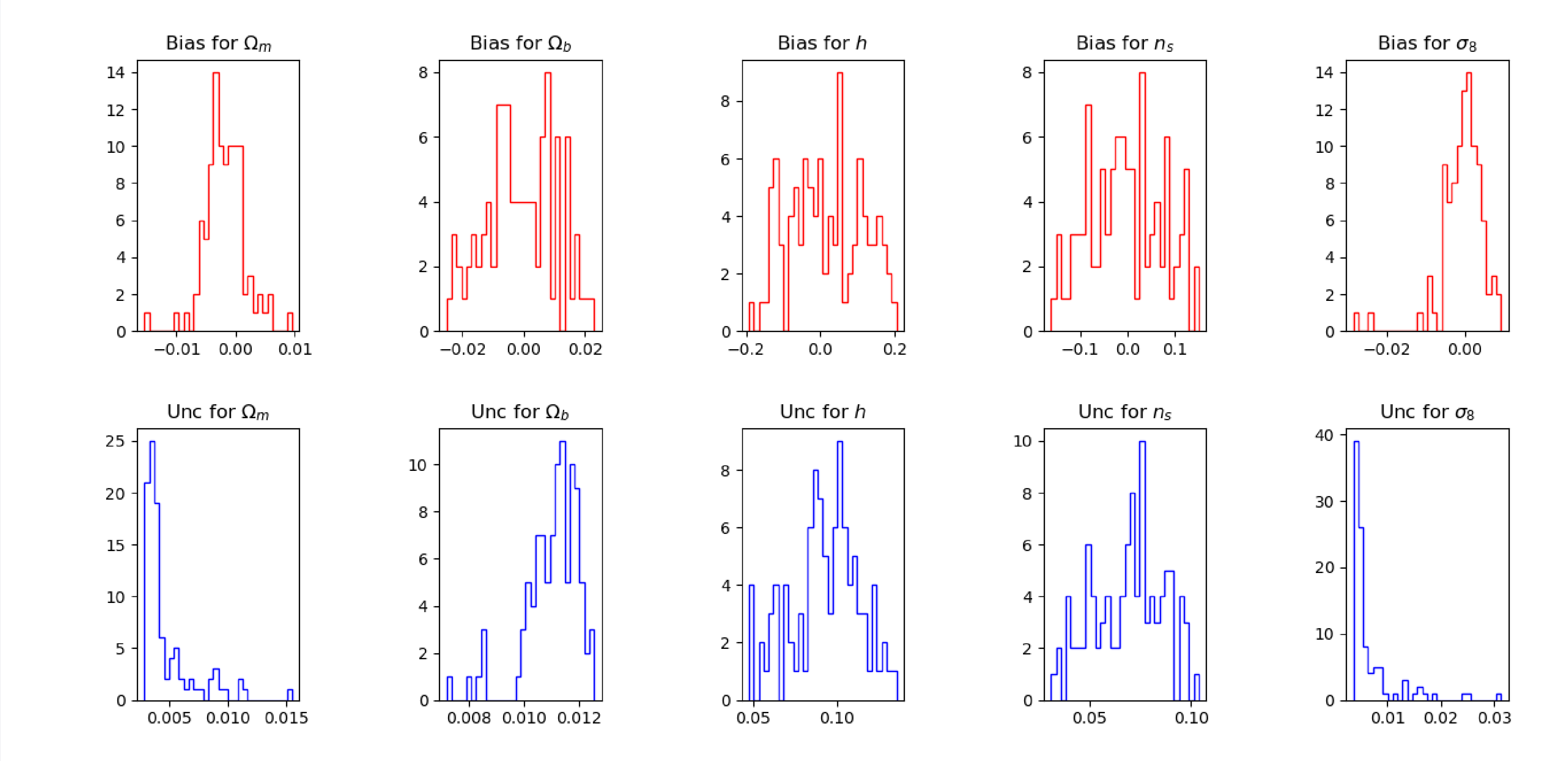
Bias

Single peak around zero for Ω_b and σ_8

Uncertainty

Smallest for Ω_m and σ_8

Results [Quijote Simulations]



Bias

Dominant peak around zero for Ω_m, h & σ_8

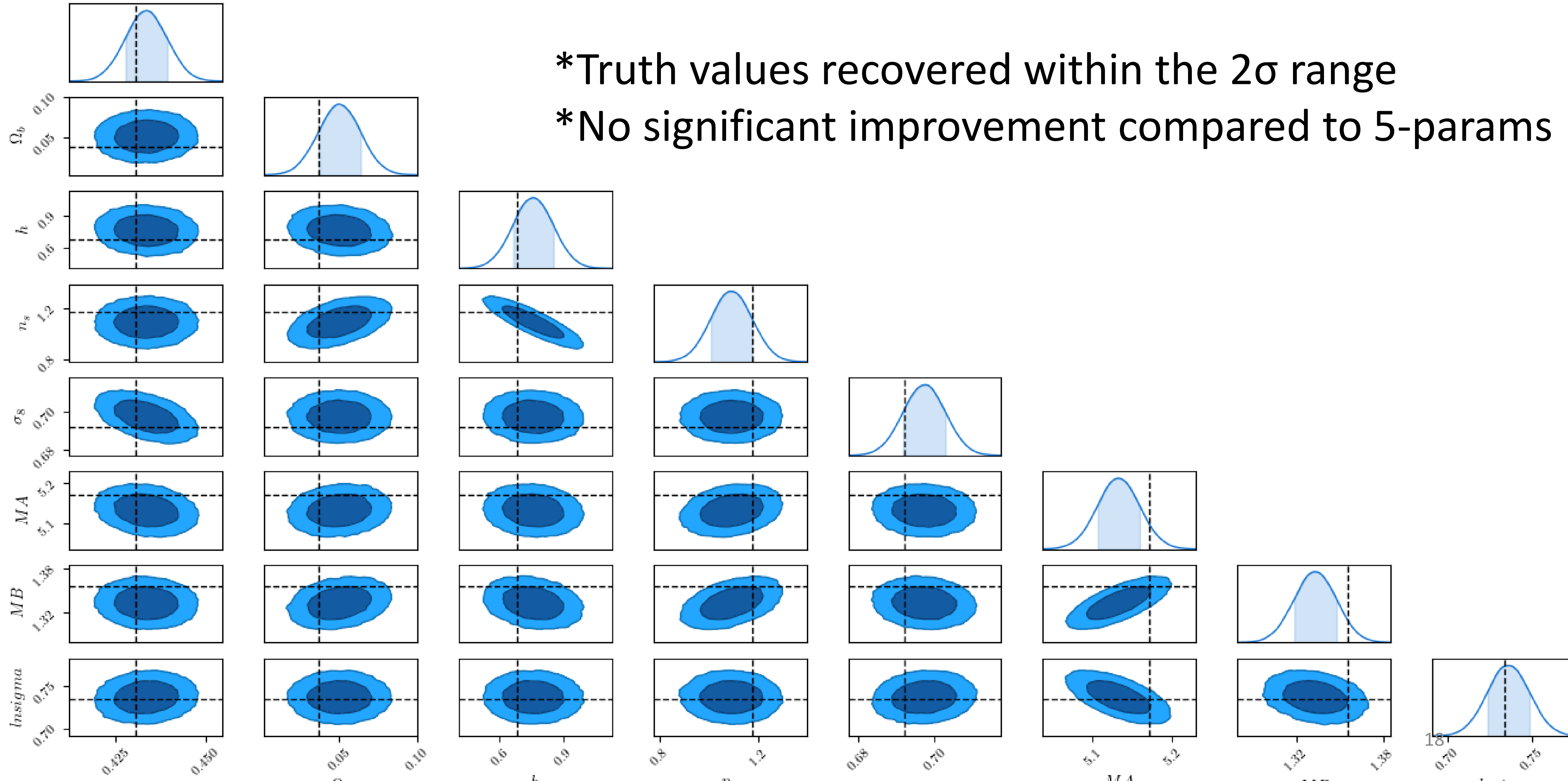
Uncertainty

Smallest for Ω_m & σ_8

Results [Cosmological & Astrophysical]

*Truth values recovered within the 2σ range

*No significant improvement compared to 5-params



Conclusions

Simulation-based Inference (SBI)

- Comparable performance to state-of-the-art-method
- Satisfying results for the simulated Universe
- Promising tool for large-scale (real) cosmic surveys
- Answer fundamental cosmological questions

Contact: moonzarin@tamu.edu