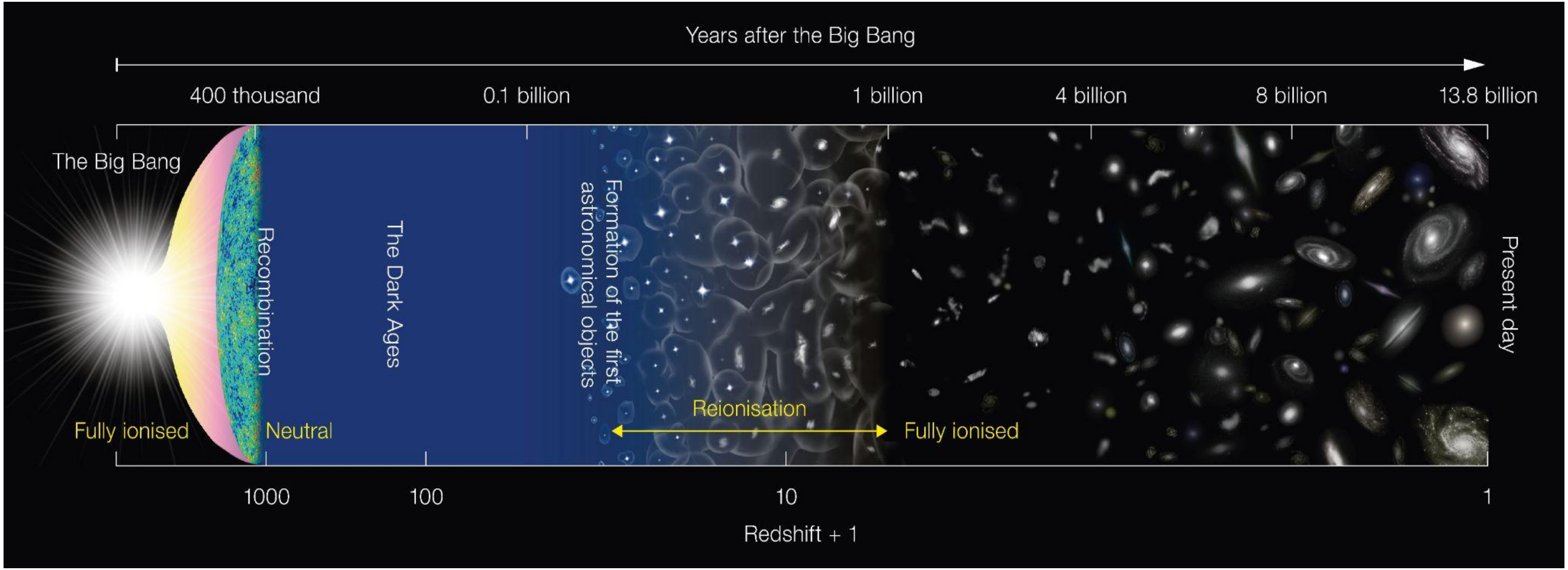
Constraining Reionization using BNN Emulated 21-cm Bispectrum

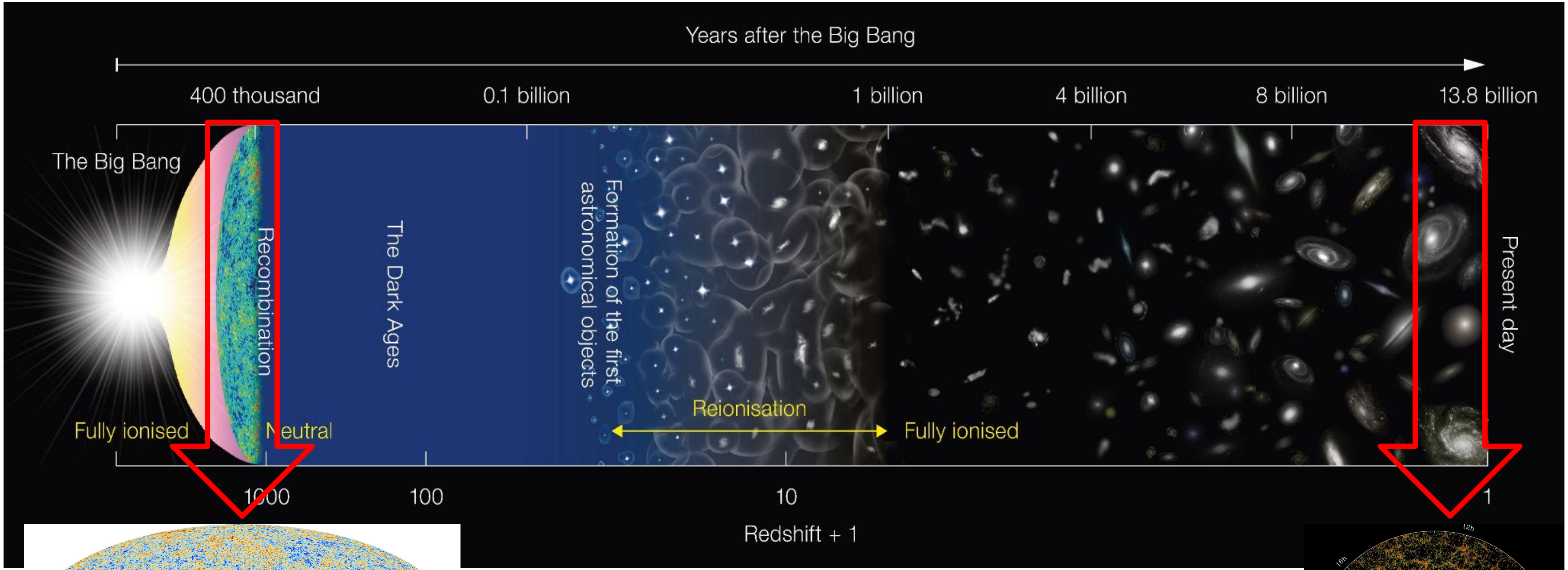
Yashrajsinh Mahida

Department of Astronomy, Astrophysics and Space Engineering
IIT Indore

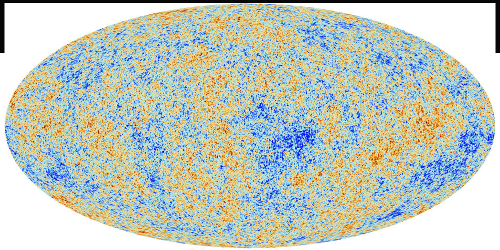
Collaborators: Suman Majumdar, Sanjay Kumar Yadav, Leon Noble


Years after the Big Bang
400 thousand
0.1 billion
1 billion
4 billion
8 billion
13.8 billion
The Big Bang
Recombination
The Dark Ages
Formation of the first astronomical objects
Present day
Fully ionised
Neutral
Reionisation
Fully ionised
1000
100
10
1
Redshift + 1

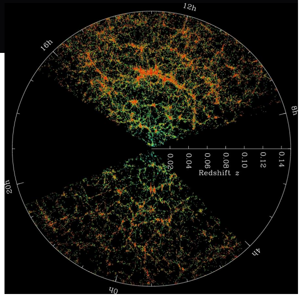
Credit: NAOJ



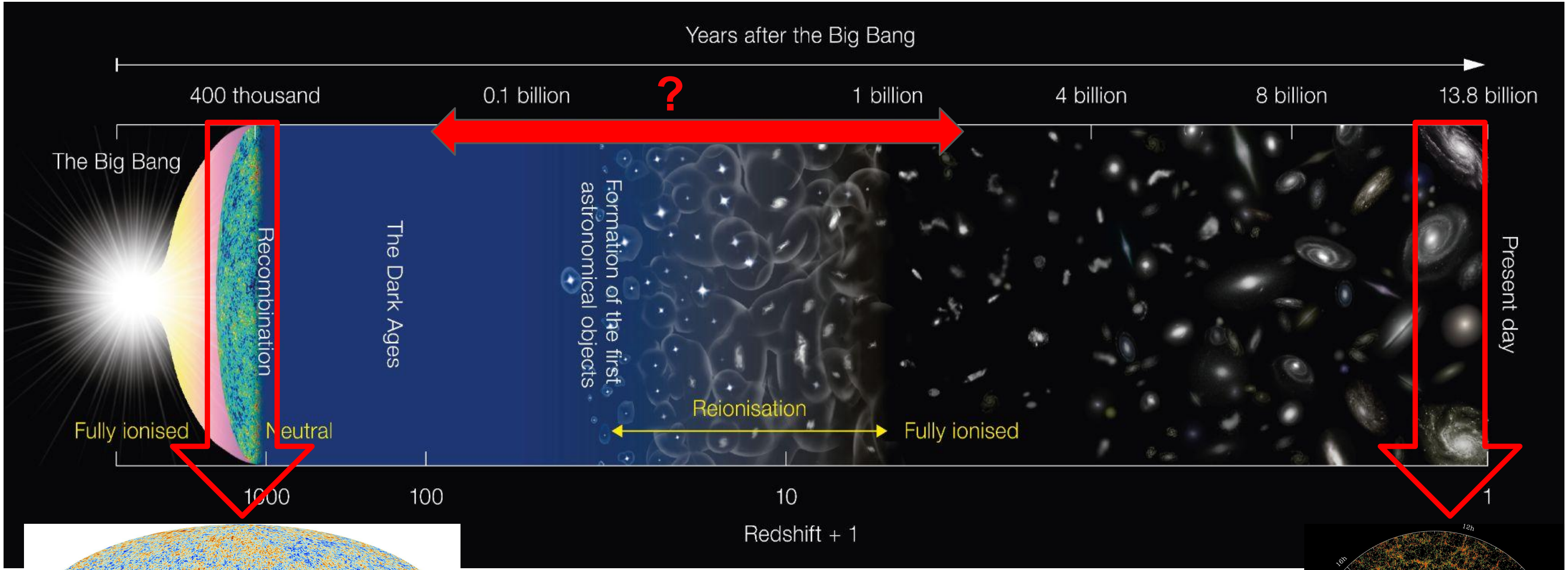
Credit: NAOJ

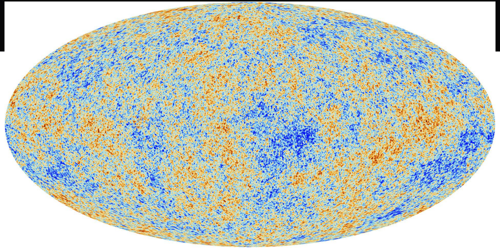


Credit: Planck Collaboration



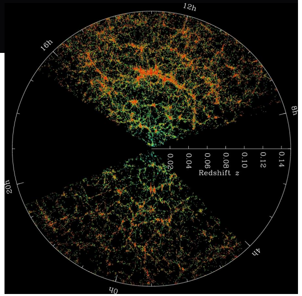
Credit: SDSS





Credit: Planck Collaboration

Credit: NAOJ



Credit: SDSS

Epoch of Reionization (EoR)

- It is the least known period of the early Universe.

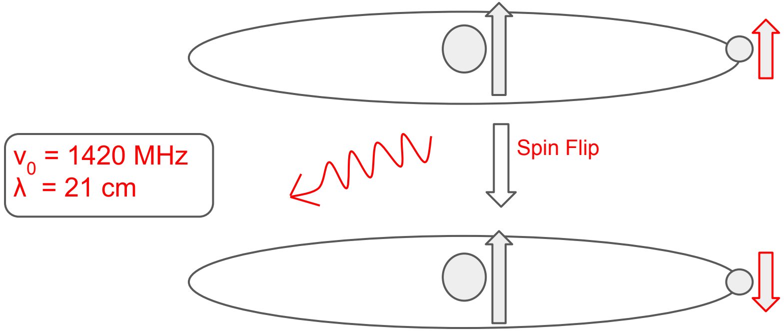
- It is a period when Intergalactic medium (IGM) was being ionized due to the formation of the first luminous sources.

- It will provide the information of the evolution of the Large-Scale Structure of the Universe.

21-cm Line

- The 21 cm line is produced by the hyperfine splitting of the neutral hydrogen caused by the interaction between electron and proton magnetic moments

- Hydrogen was the most abundant element in the IGM during EoR. So, this 21-cm line can work as the perfect tracer for IGM during early Universe

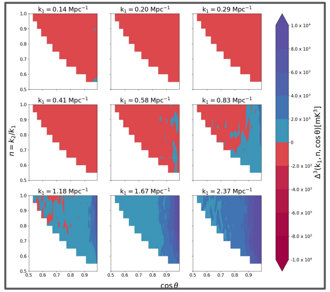


Forward

Forward Modeling of Signal Statistics

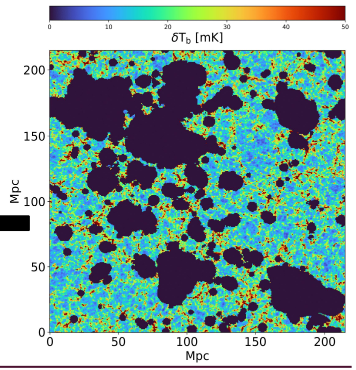
$\theta_m \rightarrow$ **Simulation** $\rightarrow$ **Summary Statistics**

$|||$



Observation

$$D \equiv B_{obs}$$



$$\mathcal{L}(D|M, \theta_m) \longrightarrow p(\theta_m|D)$$

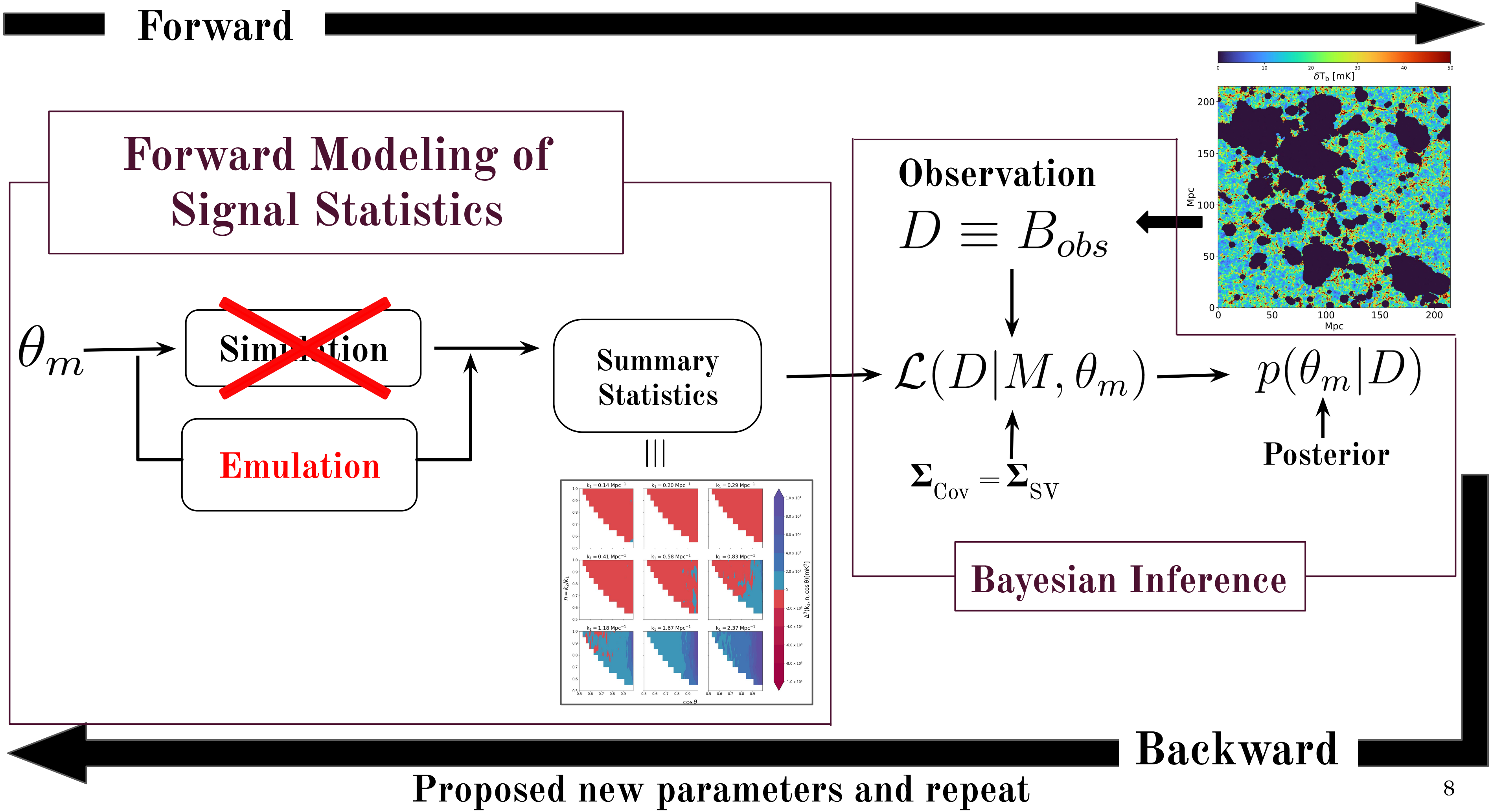
$$\mathbf{\Sigma}_{\rm Cov} = \mathbf{\Sigma}_{\rm SV}$$

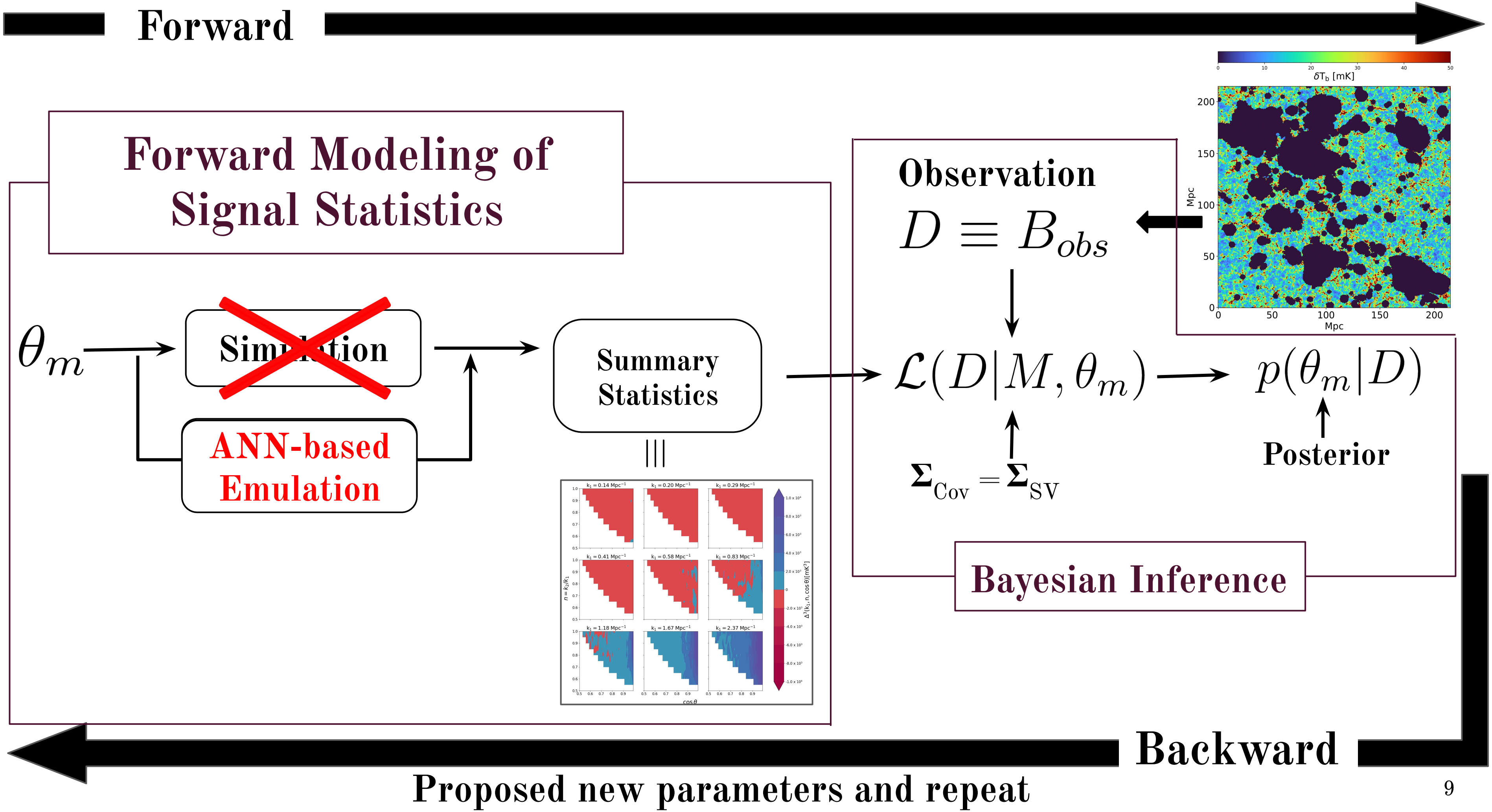
Posterior

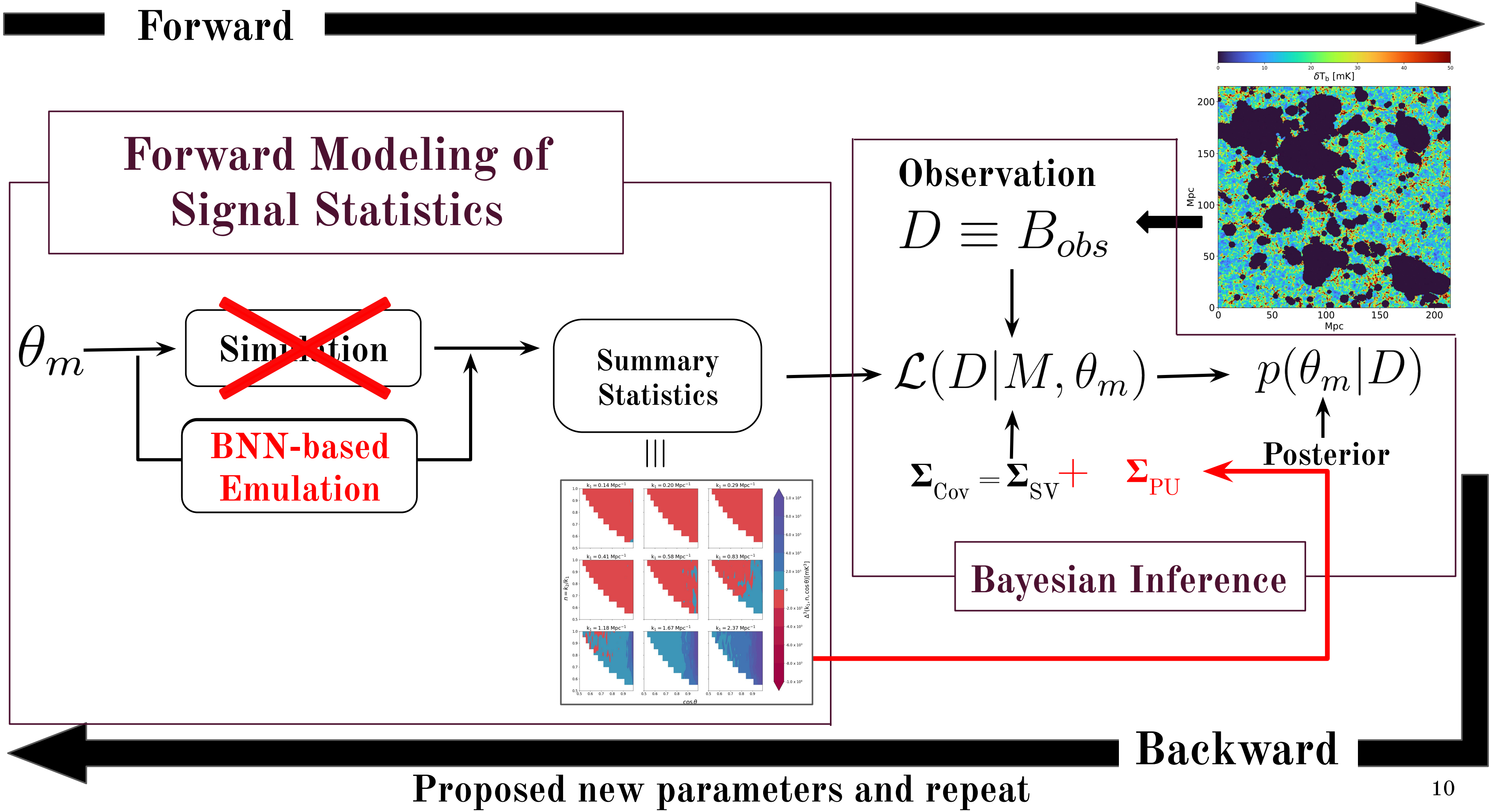
Bayesian Inference

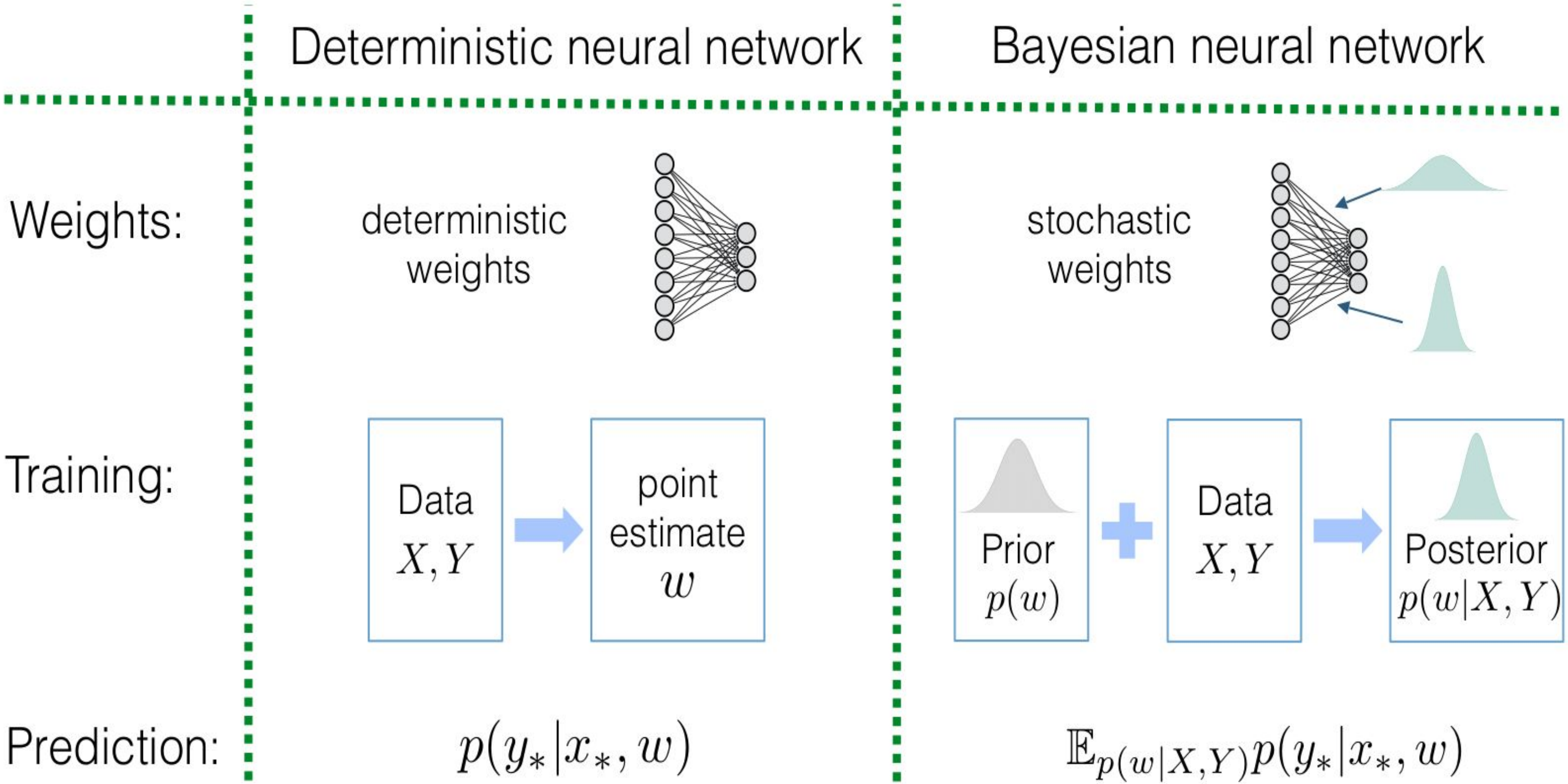
Backward

Proposed new parameters and repeat


Forward
Forward Modeling of Signal Statistics
θ_m
~~Simulation~~
Emulation
Summary Statistics
|||
$\cos\theta$
Observation
$D \equiv B_{obs}$
δT_b [mK]
Mpc
Mpc
$\mathcal{L}(D|M, \theta_m)$
$p(\theta_m|D)$
$\mathbf{\Sigma}_{\mathrm{Cov}} = \mathbf{\Sigma}_{\mathrm{SV}}$
Posterior
Bayesian Inference
Backward
Proposed new parameters and repeat


Forward
Forward Modeling of Signal Statistics
θ_m
~~Simulation~~
ANN-based Emulation
Summary Statistics
|||
$n = k_2/k_1$
$\cos\theta$
Observation
$D \equiv B_{obs}$
δT_b [mK]
Mpc
Mpc
$\mathcal{L}(D|M, \theta_m)$
$p(\theta_m|D)$
Posterior
$\mathbf{\Sigma}_{\mathrm{Cov}} = \mathbf{\Sigma}_{\mathrm{SV}}$
Bayesian Inference
Backward
Proposed new parameters and repeat


Forward
Forward Modeling of Signal Statistics
θ_m
~~Simulation~~
BNN-based Emulation
Summary Statistics
|||
$\cos\theta$
Observation
$D \equiv B_{obs}$
δT_b [mK]
Mpc
Mpc
$\mathcal{L}(D|M, \theta_m)$
$p(\theta_m|D)$
Posterior
$\mathbf{\Sigma}_{\rm Cov} = \mathbf{\Sigma}_{\rm SV} + \mathbf{\Sigma}_{\rm PU}$
Bayesian Inference
Backward
Proposed new parameters and repeat



Credit: Nadezhda Chirkova

Generating The Training Dataset

21-cm Brightness Temperature Map for given set of Parameters

Reionization Parameters:

$M_{(halo,min)}$ = Minimum halo mass that can host the ionizing sources

N_{ion} = Dimensionless parameter which relates the host halos and ionizing photons produced by them.

R_{mfp} = The mean free path of the ionizing photons.

Reionization Code: https://github.com/rajeshmondal18/ReionYuga
Ref: Choudhury et al. 2009 (http://adsabs.harvard.edu/abs/2009MNRAS.394..960C),
Majumdar et al. 2014 (http://adsabs.harvard.edu/abs/2014MNRAS.443.2843M)
Mondal et al. 2017 (http://adsabs.harvard.edu/abs/2017MNRAS.464.2992M),

Power spectrum as a summary statistic

ANN-emulator Architecture

Layers	Neurons	Activation Function
Input Layer	3	-
Hidden Layer 1	1024	ELU
Hidden Layer 2	512	ELU
Hidden Layer 3	256	ELU
Hidden Layer 4	128	ELU
Hidden Layer 5	64	ELU
Hidden Layer 6	32	ELU
Output Layer	7	-

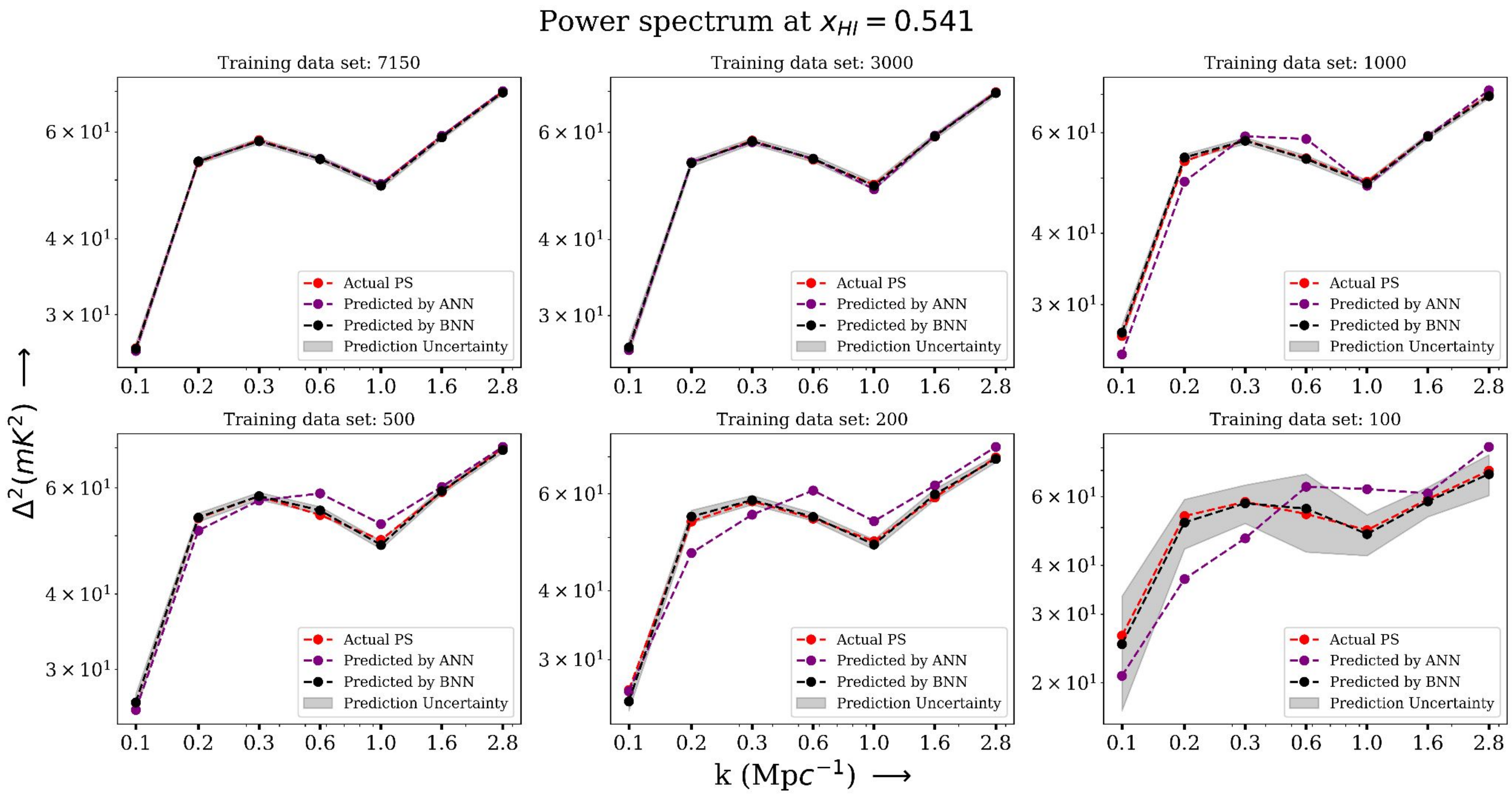
ANN-emulator Architecture

Layers	Neurons	Activation Function
Input Layer	3	-
Hidden Layer 1	1024	ELU
Hidden Layer 2	512	ELU
Hidden Layer 3	256	ELU
Hidden Layer 4	128	ELU
Hidden Layer 5	64	ELU
Hidden Layer 6	32	ELU
Output Layer	7	-

BNN-emulator Architecture

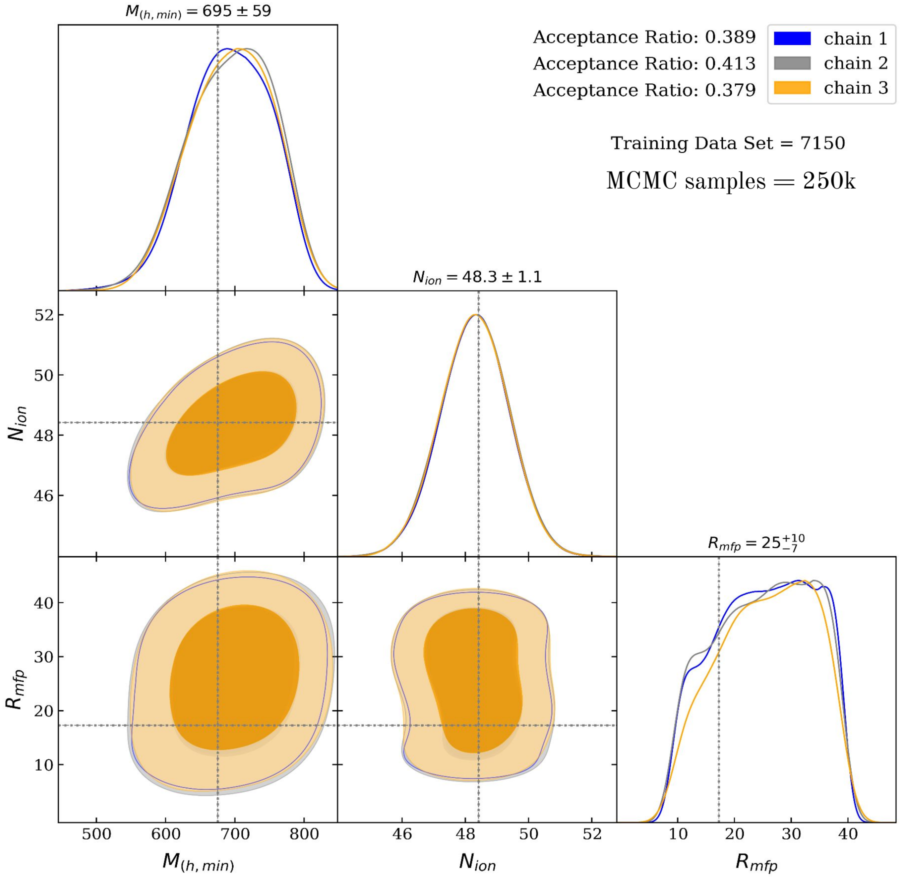
Layers	Neurons	Activation Function
Input Layer	3	-
Hidden Layer 1	50	ELU
Hidden Layer 2	100	ELU
Hidden Layer 3	50	ELU
Output Layer	7	-

Comparison of ANN & BNN-emulator prediction with different training data sets sizes

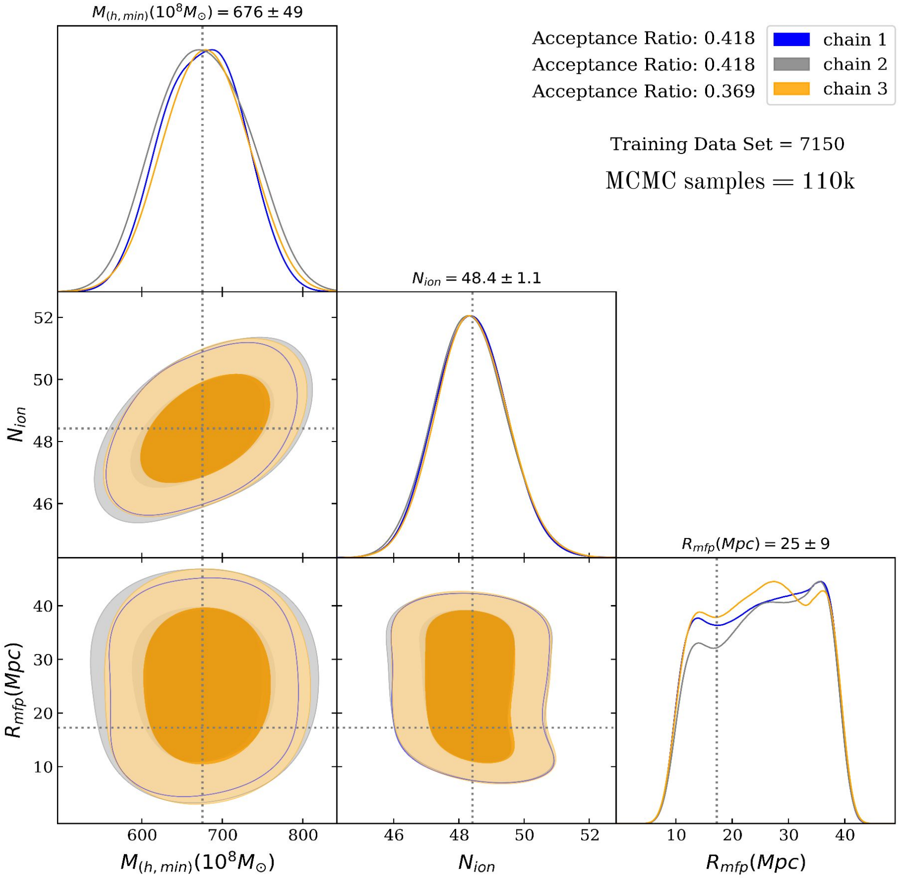


Constraints on Parameters: Test set-1

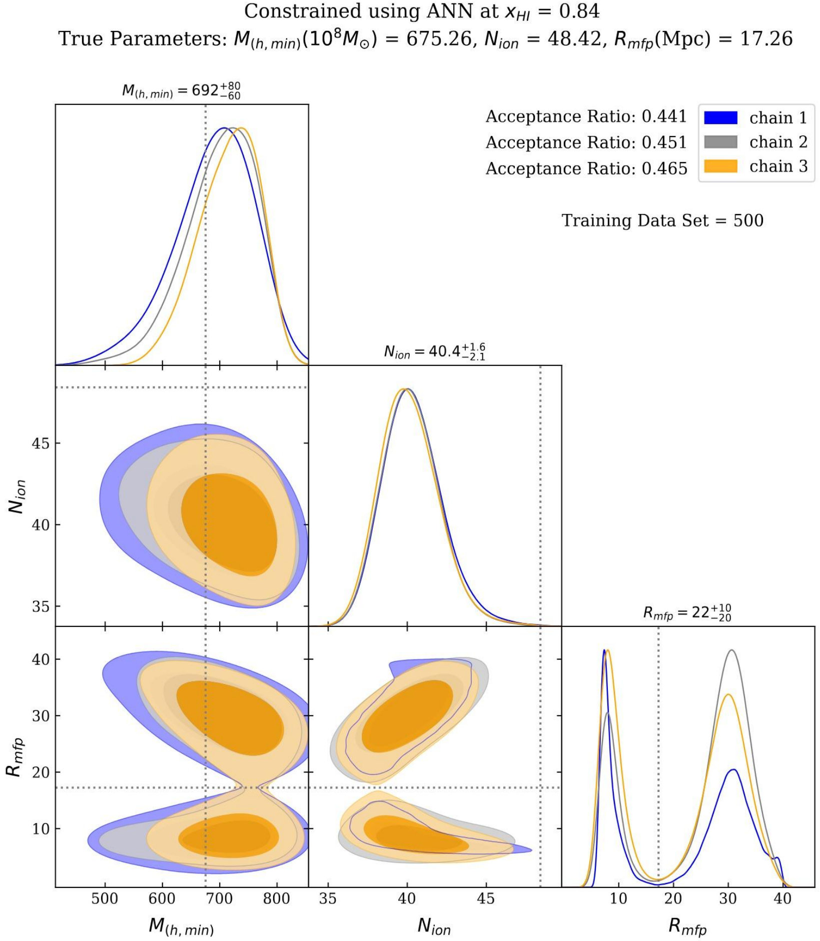
Constrained using ANN at X_{HI} = 0.84
True Parameters: $M_{(h,min)}(10^8 M_\odot)$ = 675.26, N_{ion} = 48.42, R_{mfp}(Mpc) = 17.26

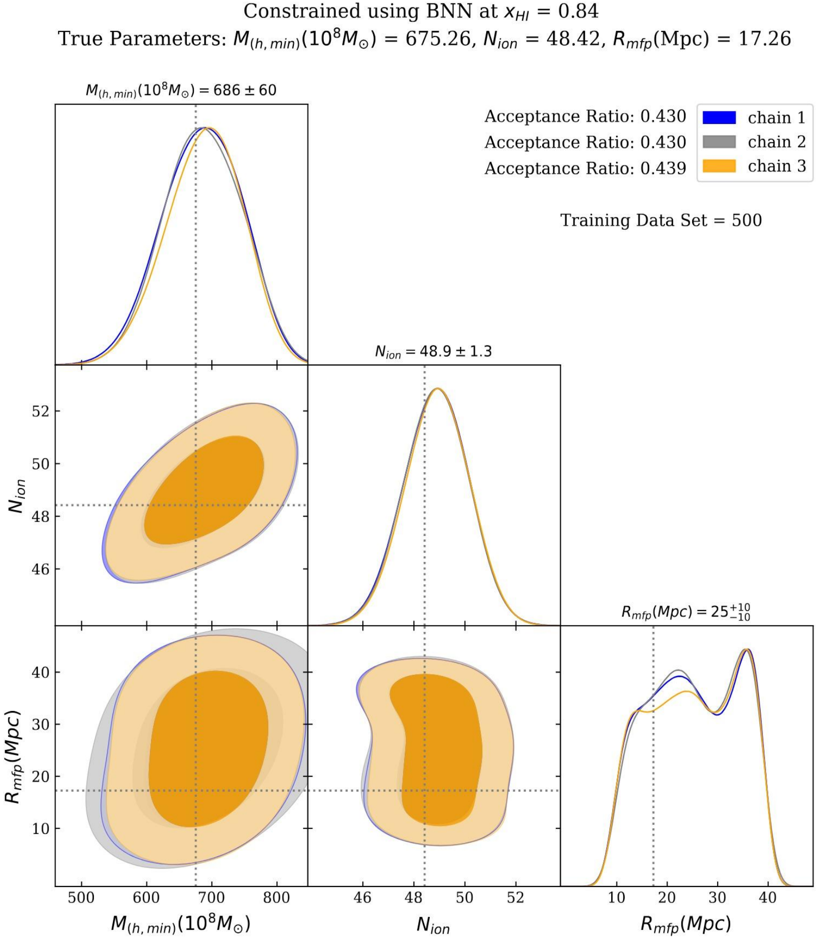


Constrained using BNN at X_{HI} = 0.84
True Parameters: $M_{(h,min)}(10^8 M_\odot)$ = 675.26, N_{ion} = 48.42, R_{mfp}(Mpc) = 17.26



With dataset size of 500: Test set-1

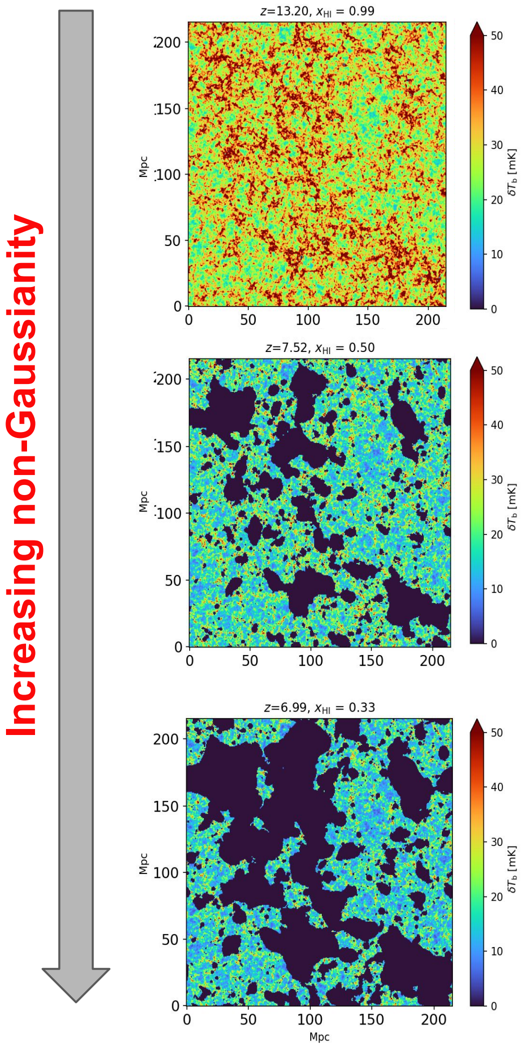




BUT

Power spectrum is NOT an optimal statistic

Probing the Non-Gaussianity

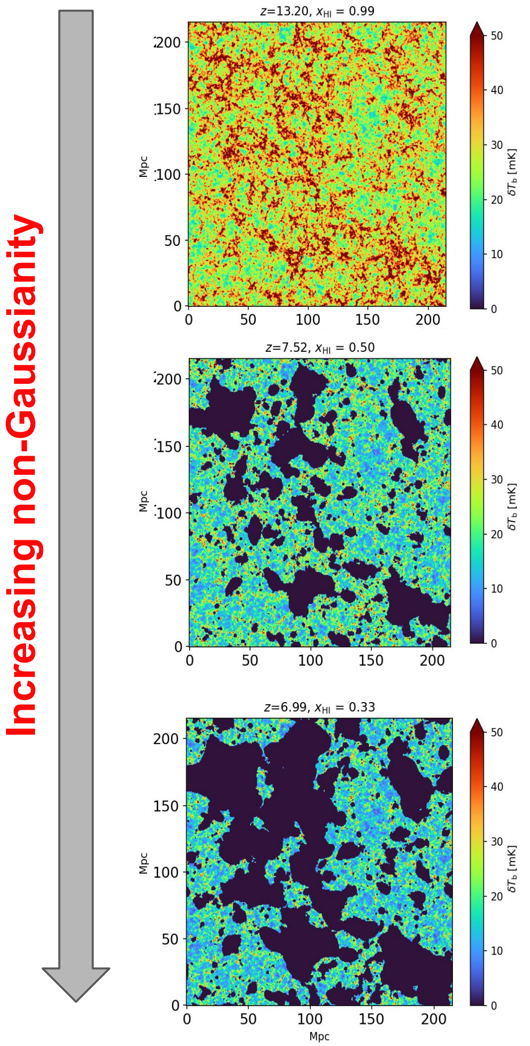


Probing the Non-Gaussianity

- 21-cm signal from EoR is highly non-Gaussian.

> ➢ **Intrinsic**: Underlying matter density distribution
>
> ➢ **Evolving**: Distribution of ionizing regions

- We need higher order signal statistic like Bispectrum to quantify this non-Gaussianity



Bispectrum as a summary statistic

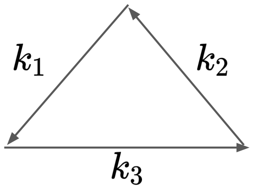
Bispectrum

- Fourier equivalent to three point correlation function.
- It quantifies the correlation between different Fourier modes.

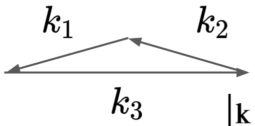
$$\langle \Delta_b(k_1)\Delta_b(k_2)\Delta_b(k_3)\rangle = V\delta^K_{(k_1+k_2+k_3,0)} B_b(k_1,k_2,k_3)$$

- Normalised Bispectrum

$$\Delta^3(k_1,k_2,k_3) = \frac{k_1^3 k_2^3}{(2\pi^2)^2} B(k_1,k_2,k_3)$$

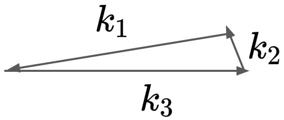


Equilateral $|\mathbf{k}_1| = |\mathbf{k}_2| = |\mathbf{k}_3|$



$|\mathbf{k}_1| = |\mathbf{k}_2| = |\mathbf{k}_3|/2$

Stretched



Squeezed $|\mathbf{k}_1| \approx |\mathbf{k}_3|, |\mathbf{k}_2| \to 0$

Ref: Majumdar et al. 2018, MNRAS, 476, 4007

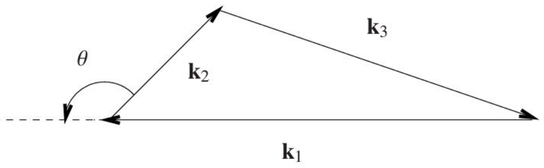
Generating 21-cm Bispectrum Training Dataset

Bispectrum Estimation

- We estimate the bispectrum using the binned bispectrum estimator which is defined as the average B inside the semi-spherical bin corresponding to the mth triangle configuration

$$\boxed{\widehat{B}_m(k_1, k_2, k_3) = \frac{1}{N_{tri} V} \sum_{[k_1 + k_2 + k_3 = 0] \in m} \Delta(k_1)\,\Delta(k_2)\,\Delta(k_3)}$$

- Parametrization for triangle configuration



$$\boxed{\frac{k_2}{k_1} = n} \qquad \boxed{\frac{k_1 \cdot k_2}{k_1 k_2} = -\cos\theta}$$

Ref: Majumdar et al. 2018, MNRAS, 476, 4007
Bharadwaj et al. 2020, MNRAS, 493, 594
Majumdar et al. 2020, MNRAS, 499, 5090

Generating 21-cm Bispectrum Training Dataset

Bispectrum Estimation

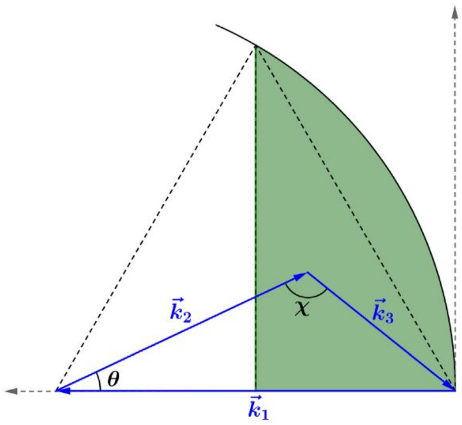
Unique triangle conditions:

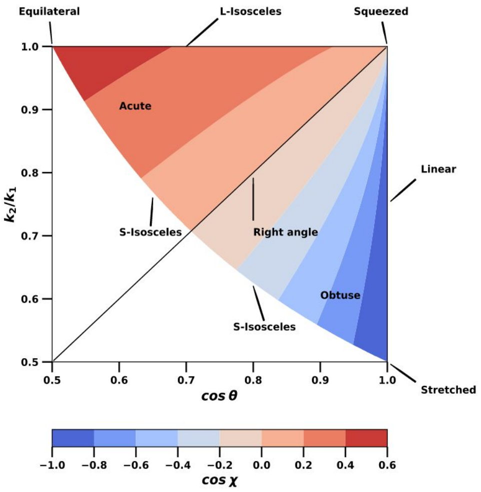
$k_1 \geq k_2 \geq k_3$

$0.5 \leq n \leq 1.0$

$0.5 \leq \cos\theta \leq 1.0$

$n \cos\theta \geq 0.5$





Ref: [Majumdar et al. 2018, MNRAS, 476, 4007](#)
[Bharadwaj et al. 2020, MNRAS, 493, 594](#)
[Majumdar et al. 2020, MNRAS, 499, 5090](#)

ANN Emulator Architecture

First Four k1 modes

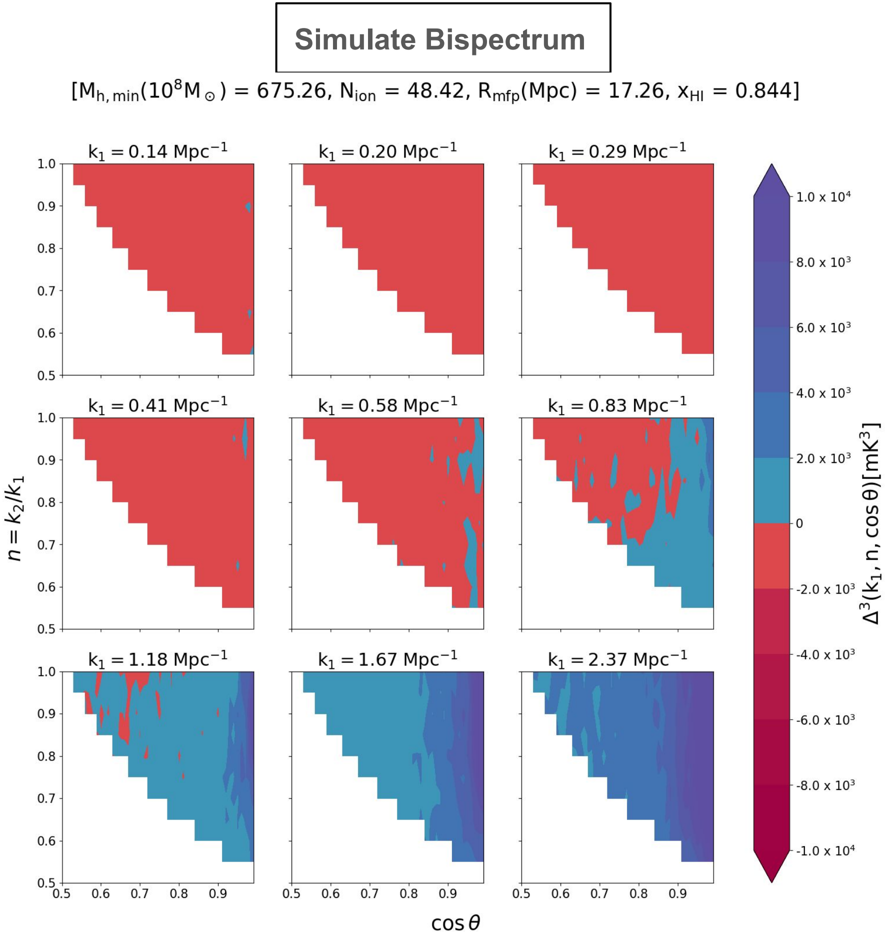
Layers	Nodes	Activation Function
Input Layer	3	
Hidden Layer 1	2255	ReLU
Hidden Layer 2	1524	ReLU
Hidden Layer 3	1008	ReLU
Output Layer 1 [For Magnitude]	328	
Output Layer 2 [For Sign]	328	

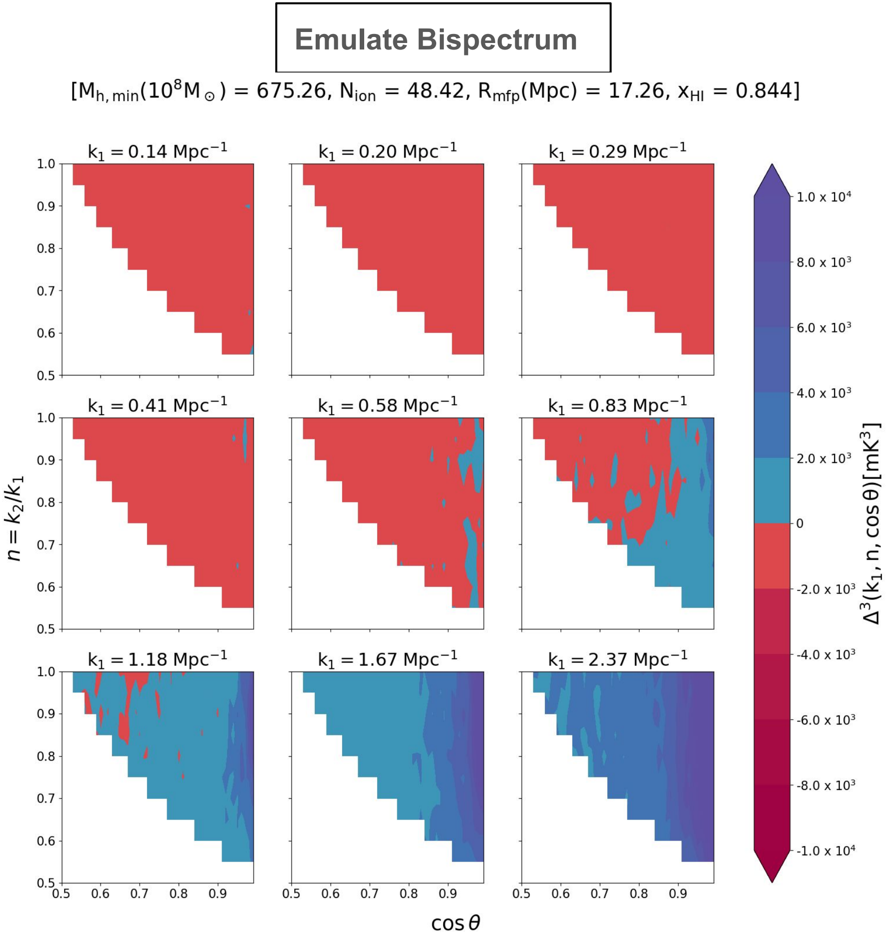
Remaining Five k1 modes

Layers	Nodes	Activation Function
Input Layer	3	
Hidden Layer 1	2255	ReLU
Hidden Layer 2	1524	ReLU
Hidden Layer 3	1008	ReLU
Hidden Layer 4	956	
Output Layer 1 [For Magnitude]	328	
Output Layer 2 [For Sign]	328	

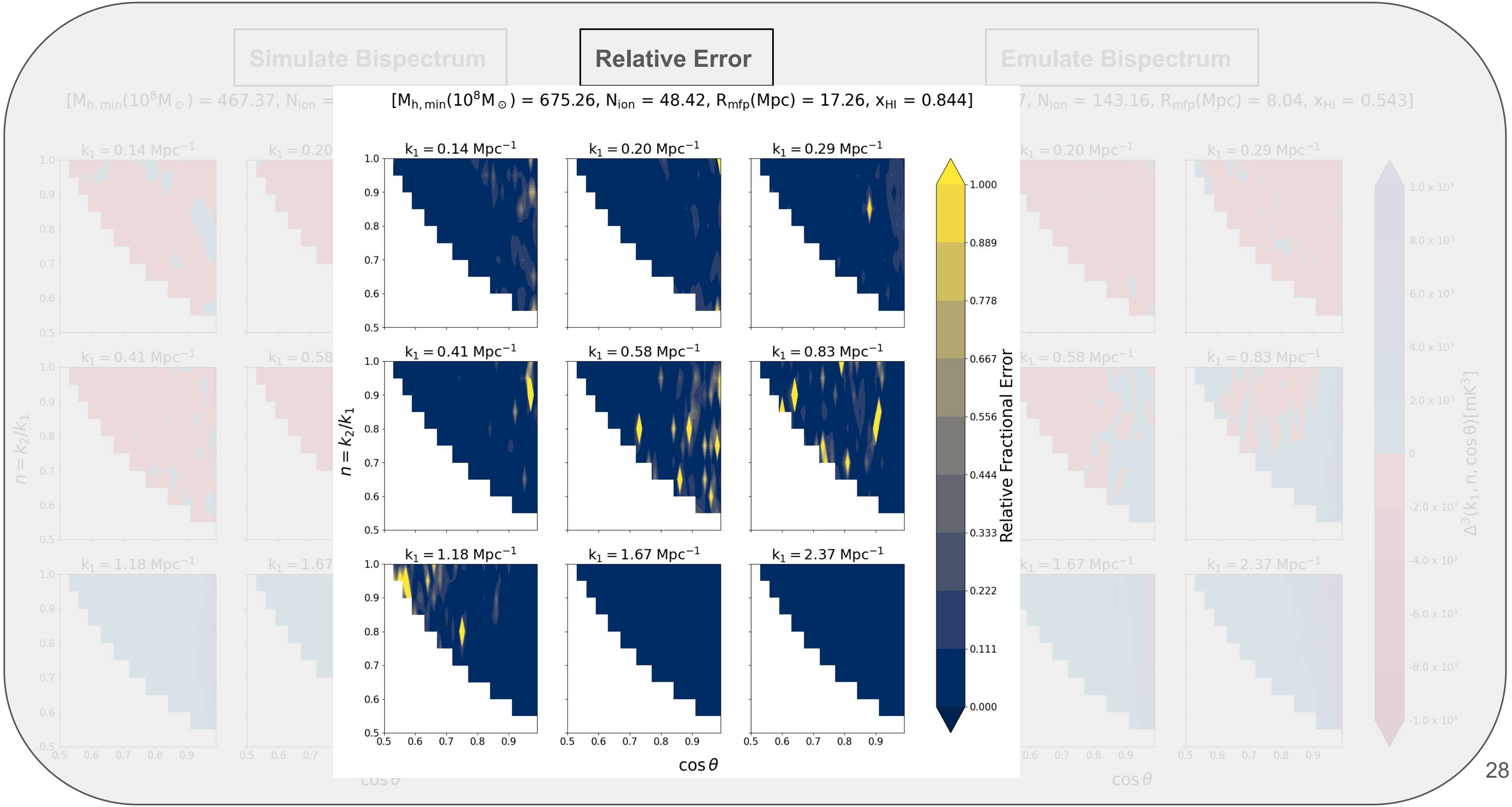
Loss Function = MSE (Mean Squared Error)

Bispectrum Prediction using ANN





Bispectrum Prediction using ANN



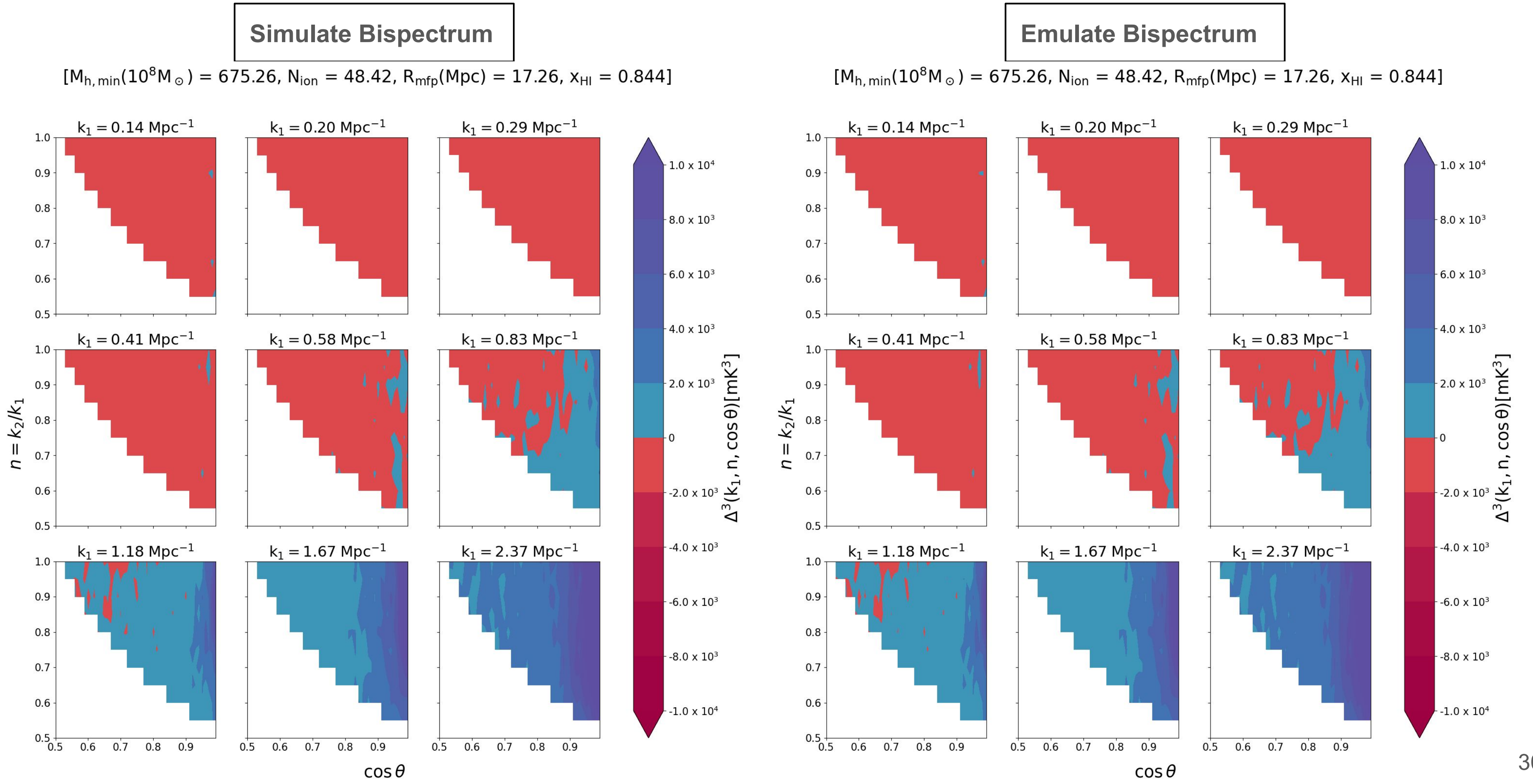
BNN Emulator Architecture

<u>For All k1 modes</u>

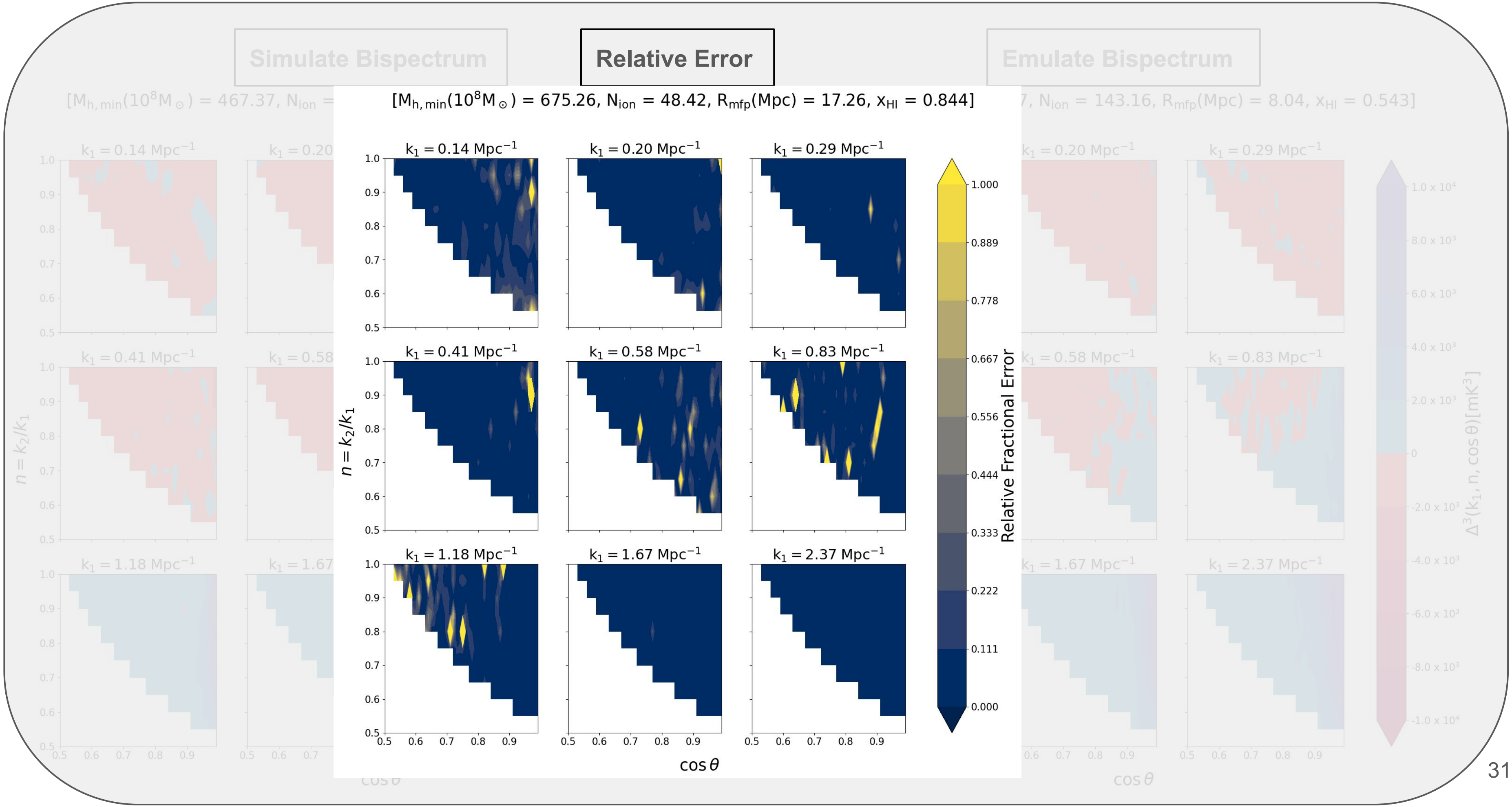
Layers	Nodes	Activation Function
Input Layer	3	
Hidden Layer 1	128	ELU
Hidden Layer 2	256	ELU
Hidden Layer 3	512	ELU
Output Layer	328	

- Prior distribution of weights & biases $\Rightarrow$ Gaussian
- Likelihood function $\Rightarrow$ Multivariate Gaussian

Bispectrum Prediction using BNN

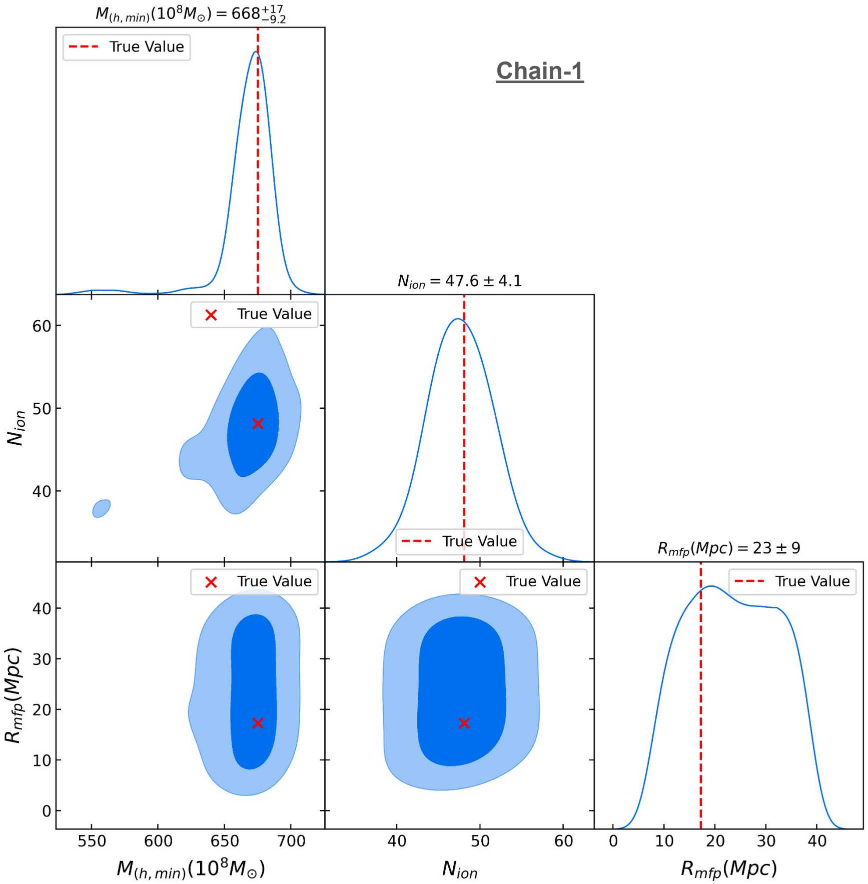


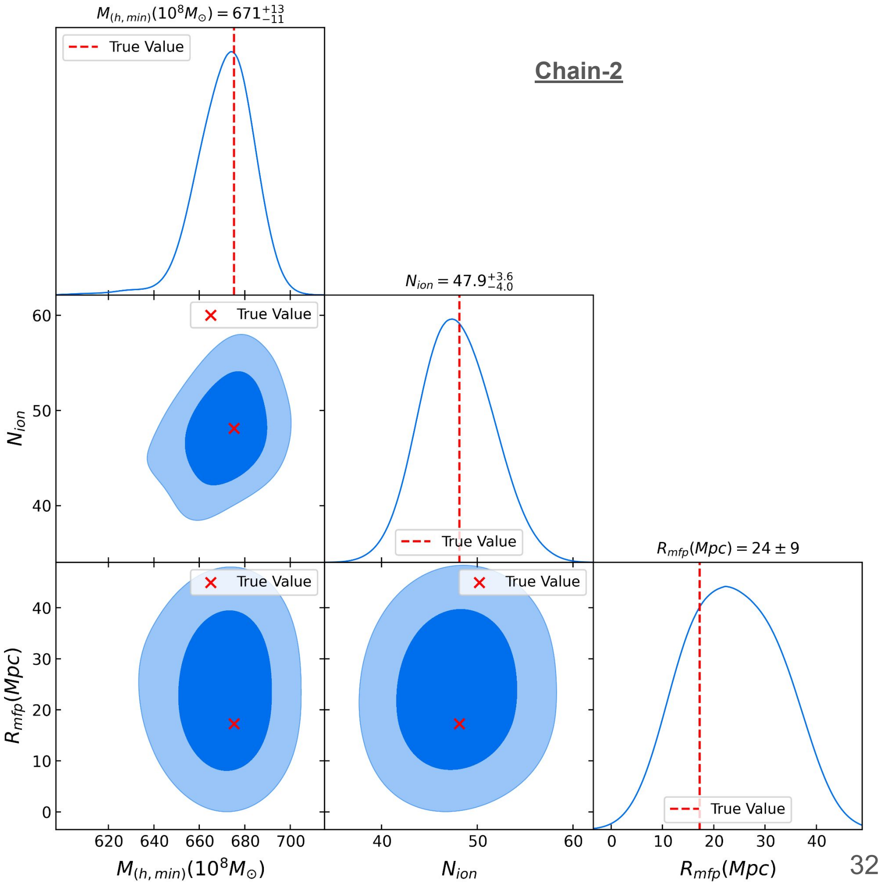
Bispectrum Prediction using BNN



Parameters Constraint using ANN

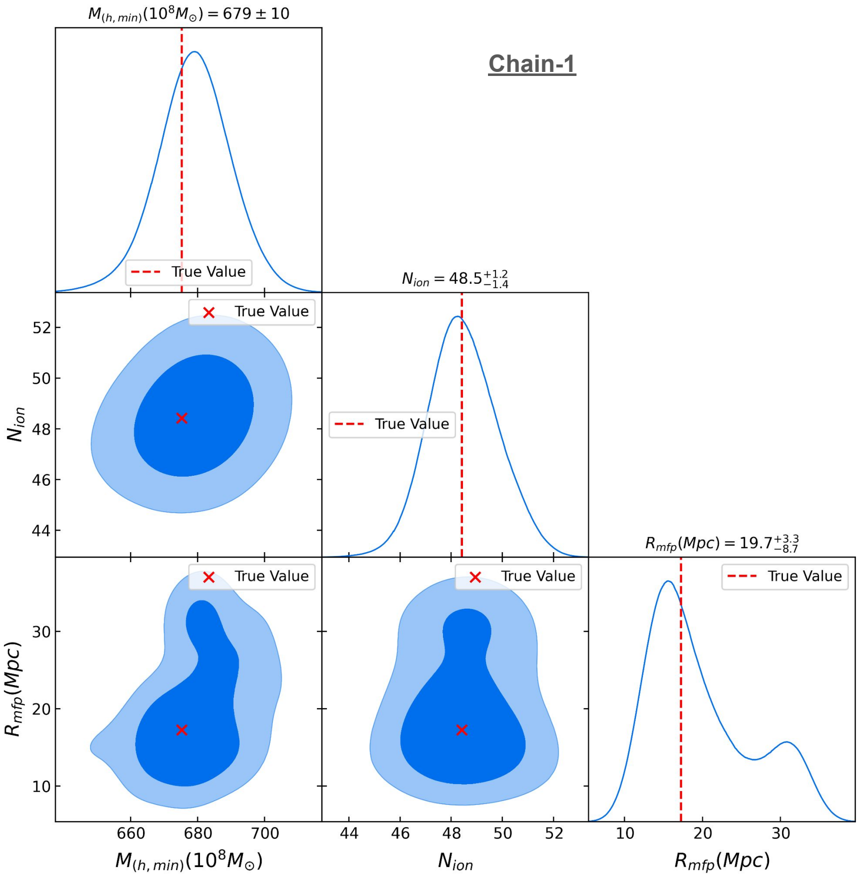
$[M_{h,min}(10^8 M_\odot) = 675.26,\ N_{ion} = 48.42,\ R_{mfp}(Mpc) = 17.26,\ x_{HI} = 0.844]$

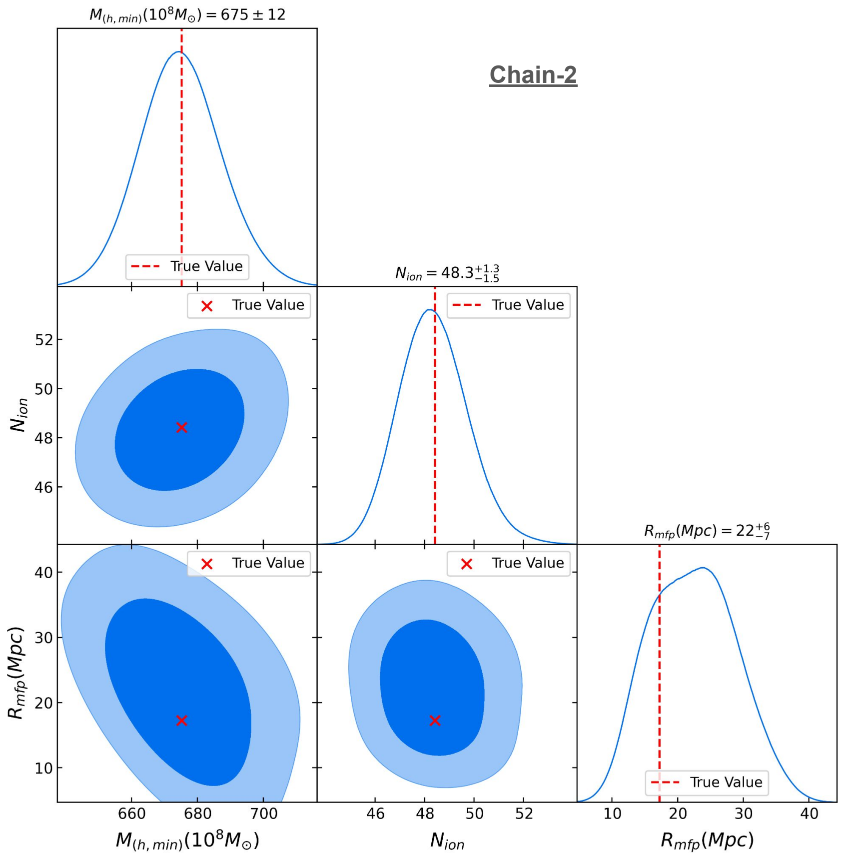




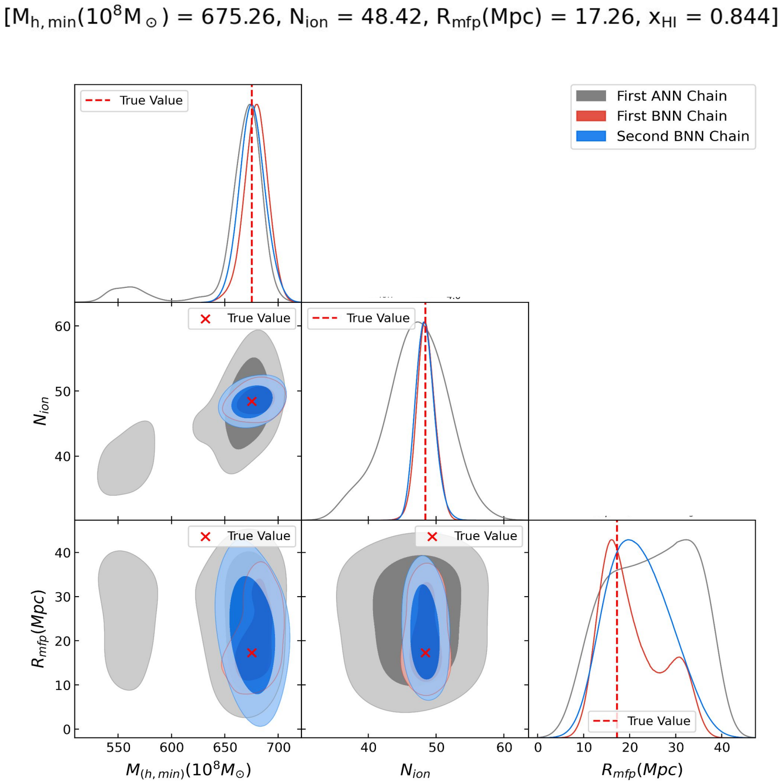
Parameters Constraint using BNN

$$[M_{h,min}(10^8 M_\odot) = 675.26,\ N_{ion} = 48.42,\ R_{mfp}(Mpc) = 17.26,\ x_{HI} = 0.844]$$

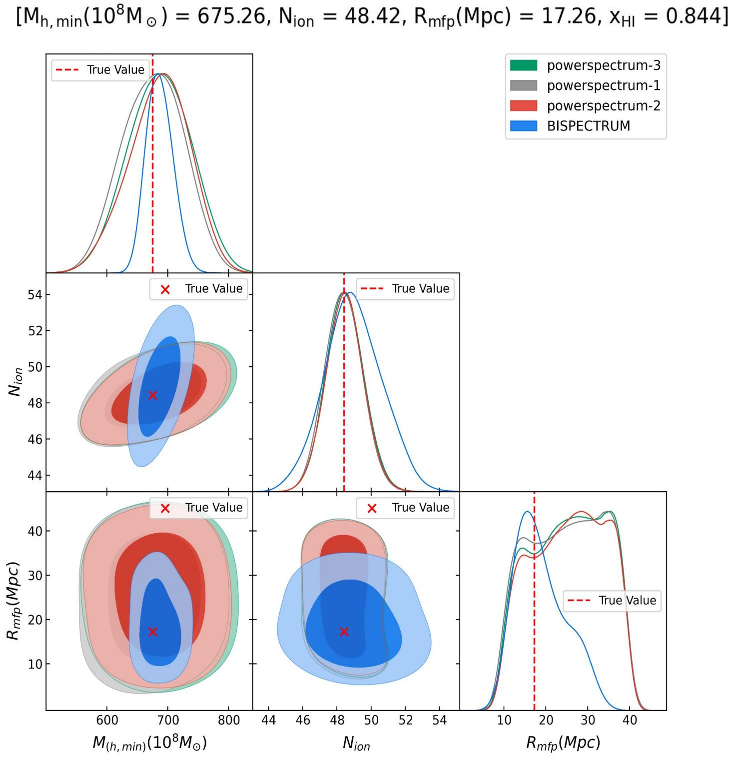




Comparing Parameter Constraints of BNN with ANN



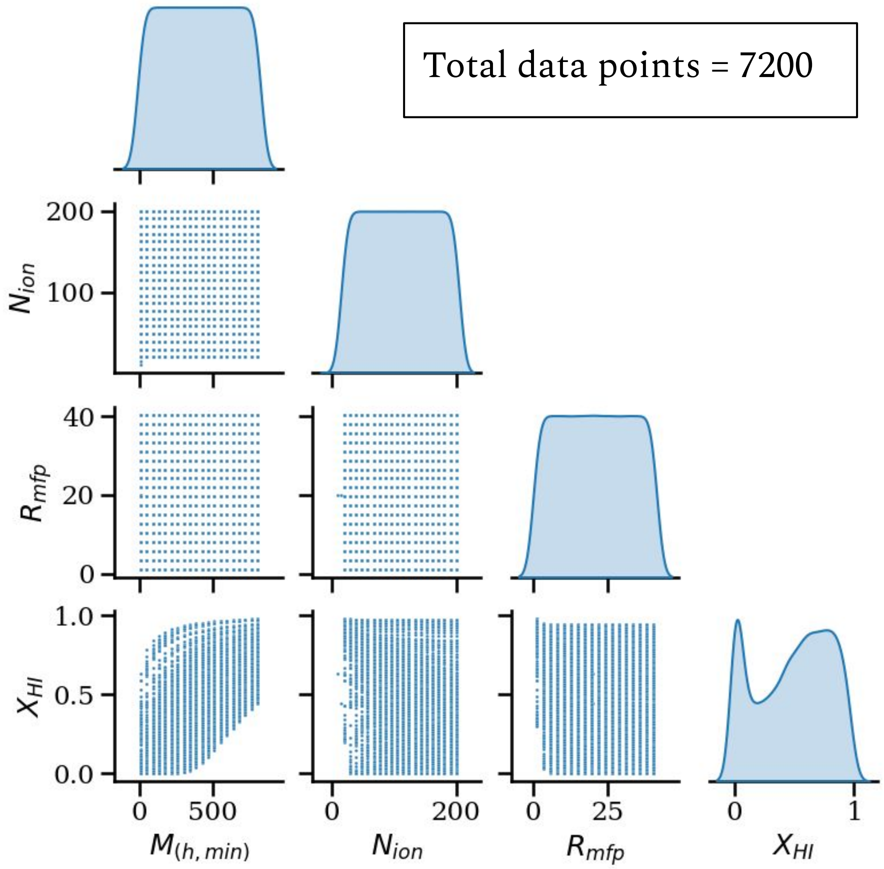
Comparing Parameter Constraints of Bispectrum with Power Spectrum



Summary

- 21-cm signal is highly non-Gaussian. So, we use bispectrum for the parameter estimation
- ANNs provide only point value prediction. So, we cannot quantify the uncertainty in their prediction.
- We use BNNs, which provide the posterior distribution of predicted signal. So, they are better at quantifying the model uncertainty.
- We show that BNNs give better constraints of the EoR parameters than ANNs.
- Bispectrum provides tighter constraints than power spectrum.

- New data set with large number of samples.
- In this set samples are uniformly distributed.



Total data points = 7200