Faculty Journal club @ 2024/4/10

Simulation-Based Inference of Reionization Parameters From 3D Tomographic 21 cm Lightcone Images

Xiaosheng Zhao,[1] Yi Mao,[1] Cheng Cheng,[2] and Benjamin D. Wandelt[3,4,5]

[1]Department of Astronomy, Tsinghua University, Beijing 100084, China
[2]School of Chemistry and Physics, University of KwaZulu-Natal, Westville Campus, Durban, 4000, South Africa
[3]Sorbonne Université, CNRS, UMR 7095, Institut d'Astrophysique de Paris (IAP), 98 bis bd Arago, 75014 Paris, France
[4]Sorbonne Université, Institut Lagrange de Paris (ILP), 98 bis bd Arago, 75014 Paris, France
[5]Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA

Astro-ph/ 2105.03344

Hayato Shimabukuro

(Take home messages)

- **Simulation-based inference (SBI)** constructs Likelihood through forward simulation by using **Density estimation likelihood-free inference (DELFI)**

- **SBI** + **machine learning** + **Bayesian inference** → **posterior distribution**

Theory

(e.g) cosmological model

$$\Omega_{\rm tot}, \Omega_b, \Omega_m, \Omega_\Lambda, H_0, n_s \ldots$$

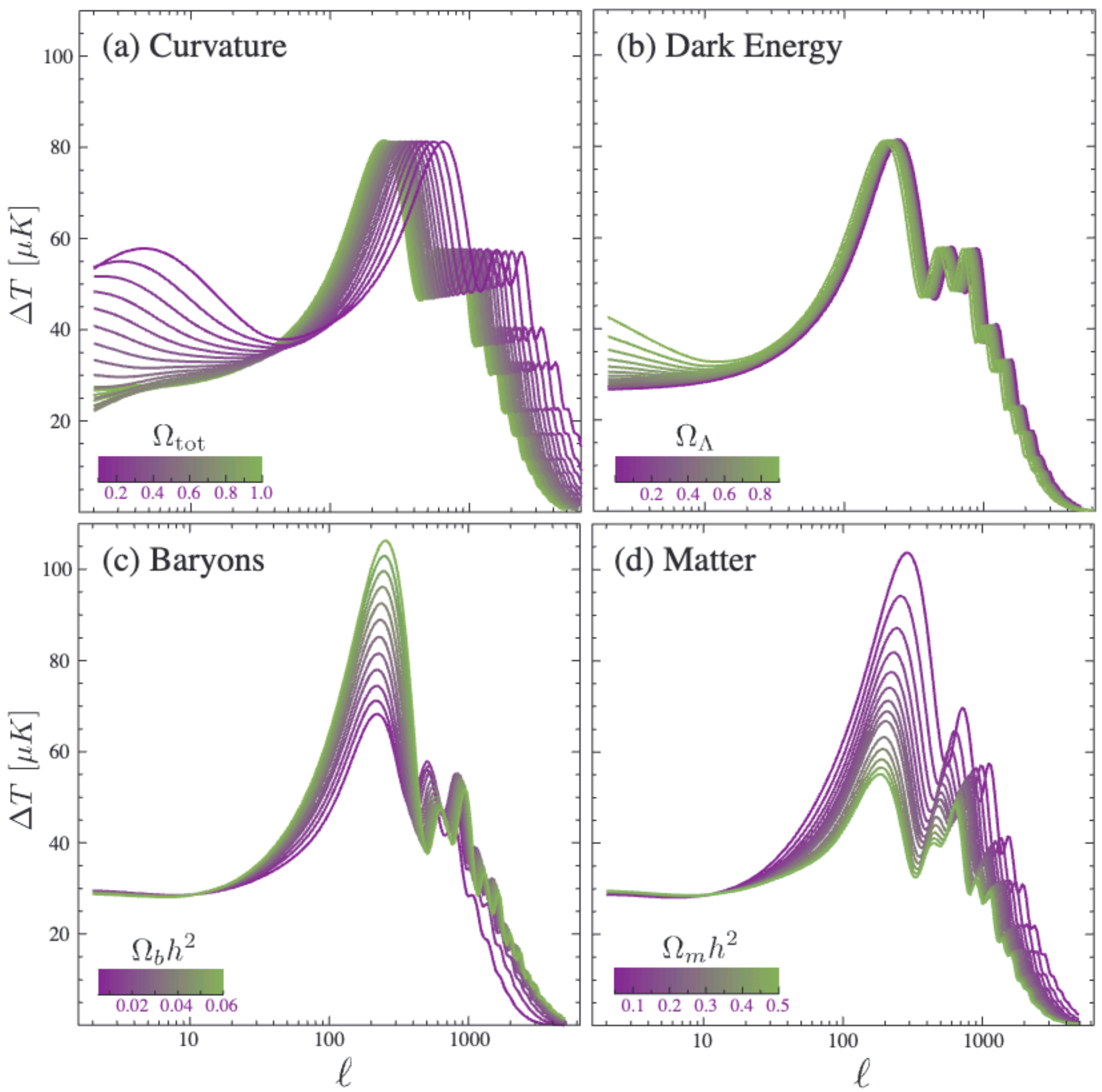
•A model is characterized by parameters.

Theory

(e.g) cosmological model

$$\Omega_{\text{tot}}, \Omega_b, \Omega_m, \Omega_\Lambda, H_0, n_s \ldots$$

•A model is characterized by parameters.

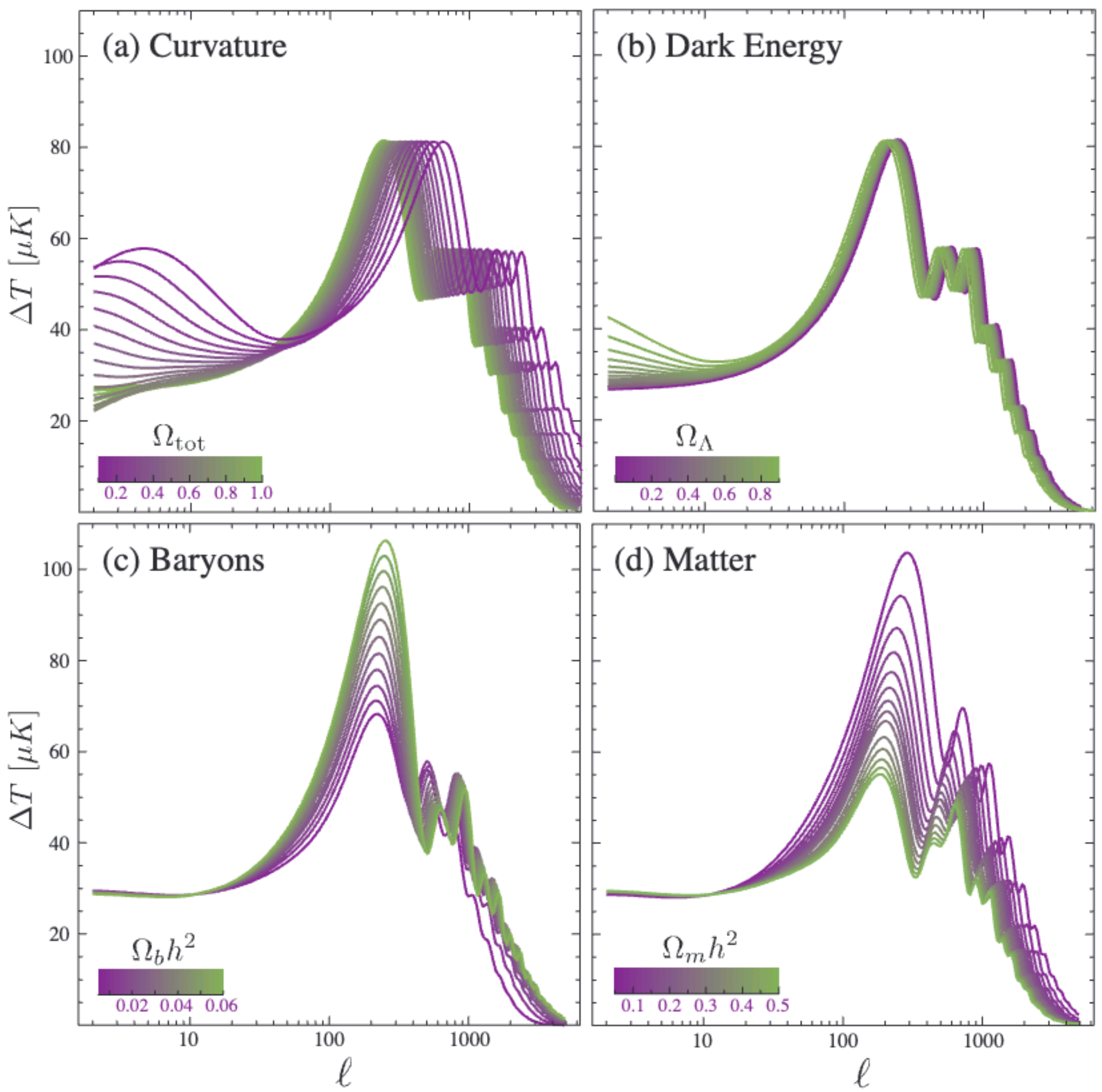


Theory

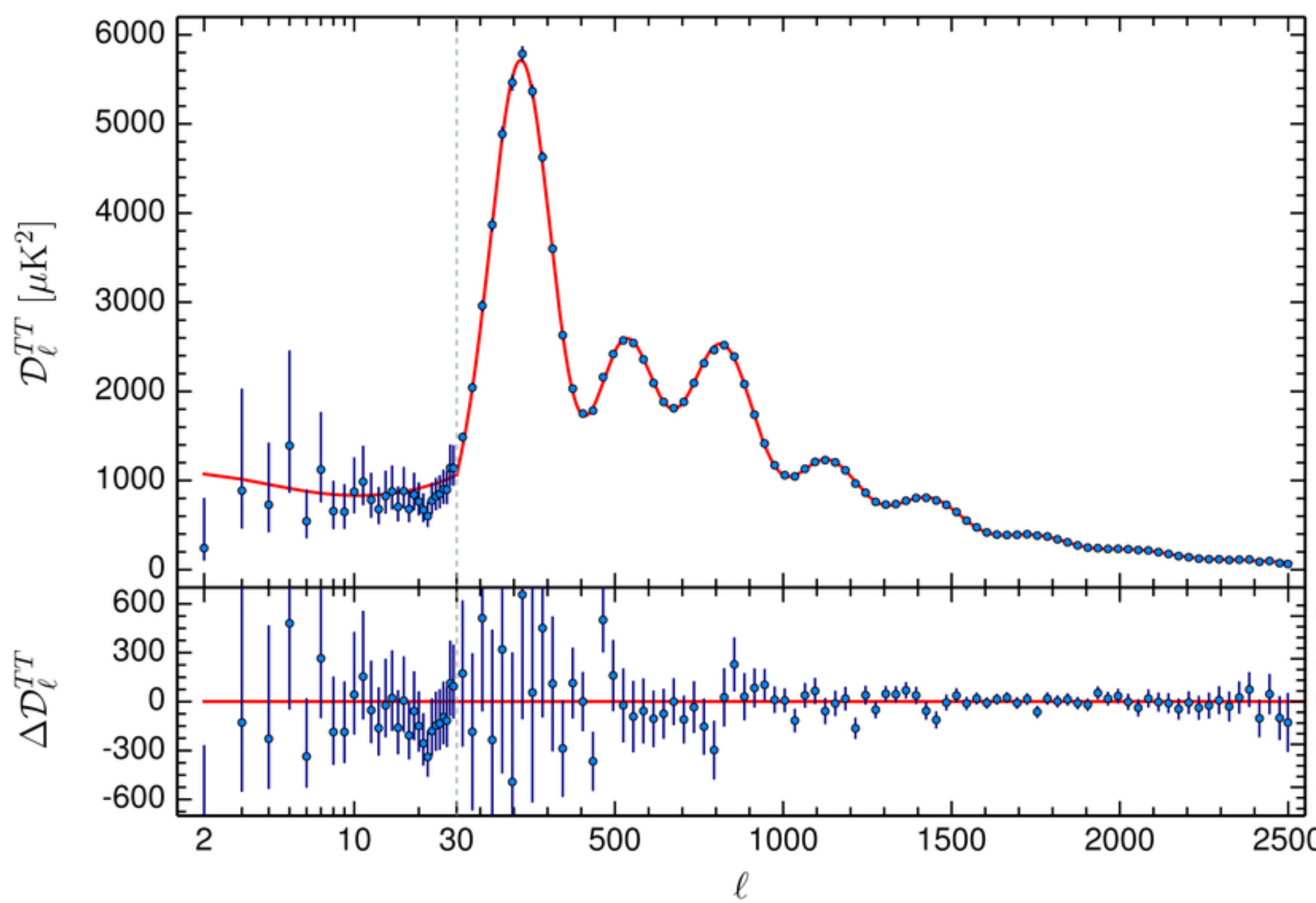
(e.g) cosmological model

$$\Omega_{\rm tot}, \Omega_b, \Omega_m, \Omega_\Lambda, H_0, n_s \ldots$$

•A model is characterized by parameters.



Observations/experiments



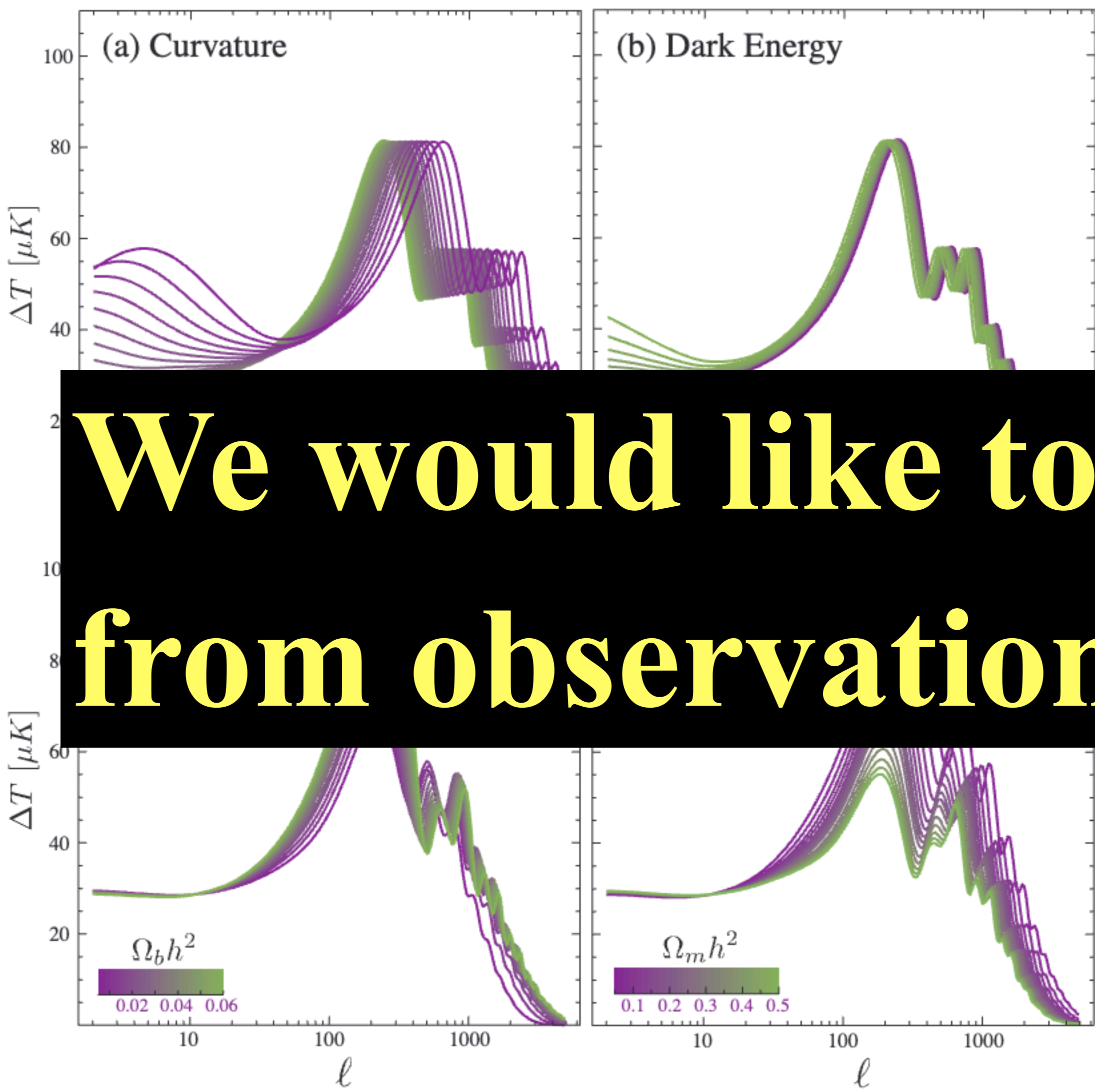
Theory

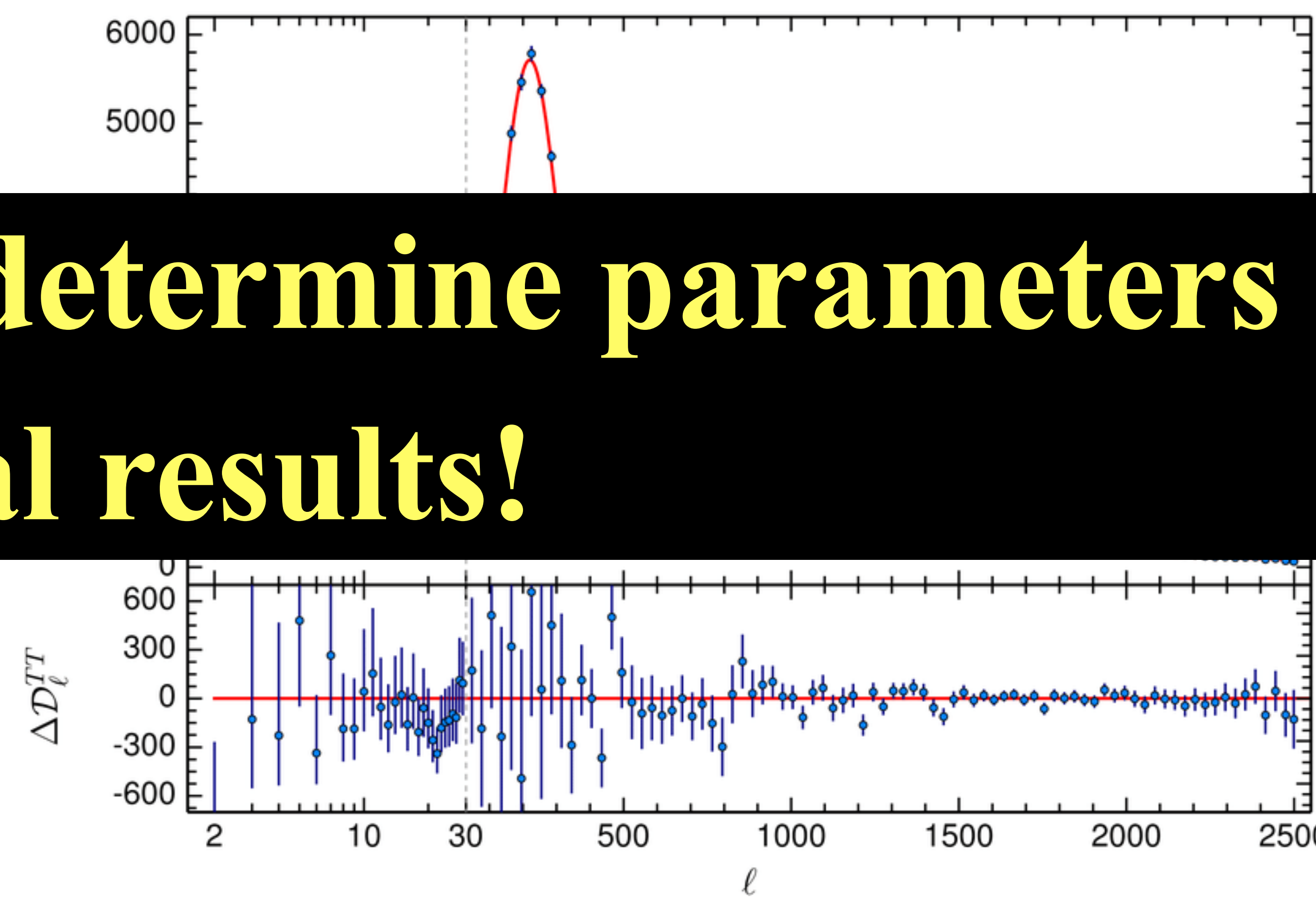
(e.g) cosmological model

$$\Omega_{\rm tot}, \Omega_b, \Omega_m, \Omega_\Lambda, H_0, n_s \ldots$$

•A model is characterized by parameters.

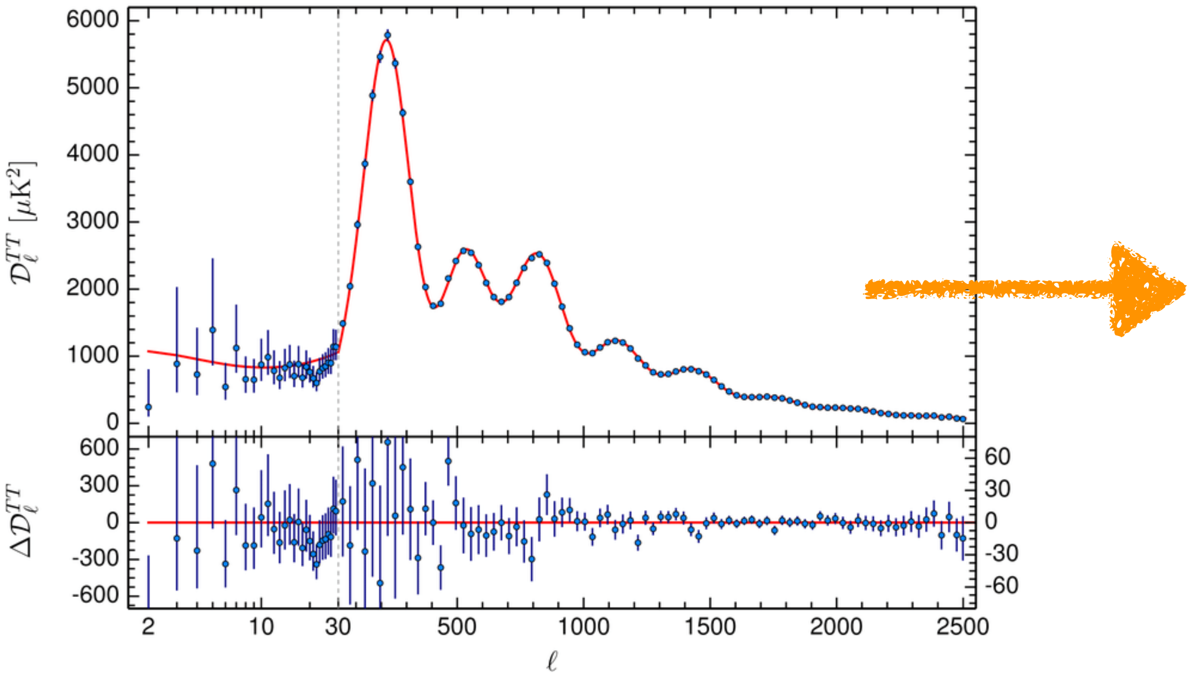
Observations/experiments





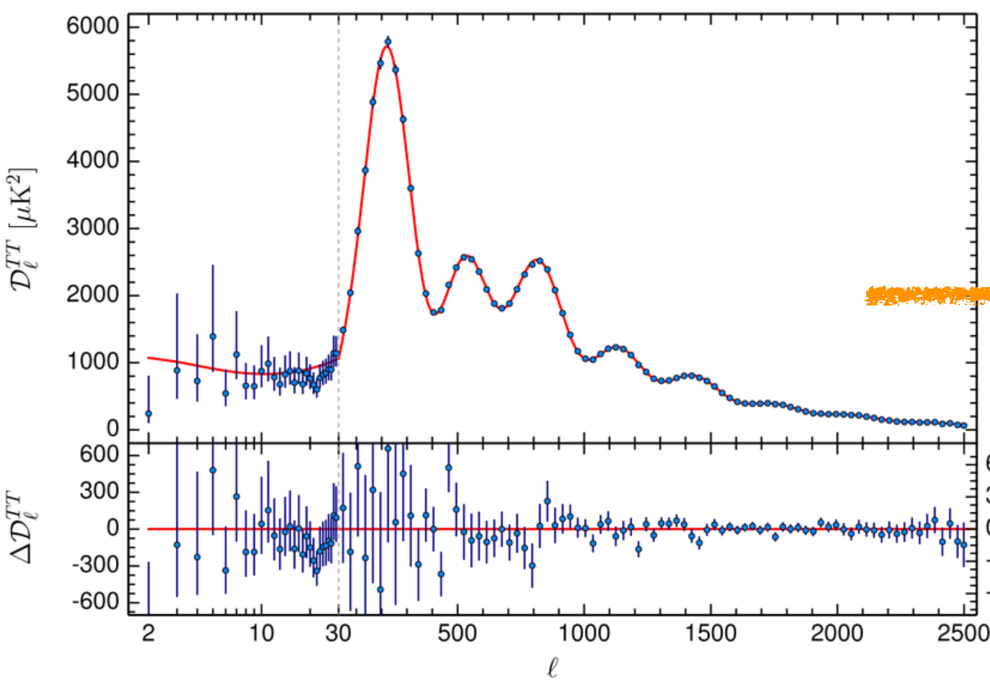
We would like to determine parameters from observational results!

•**For given observational results, what parameter sets are preferred??**



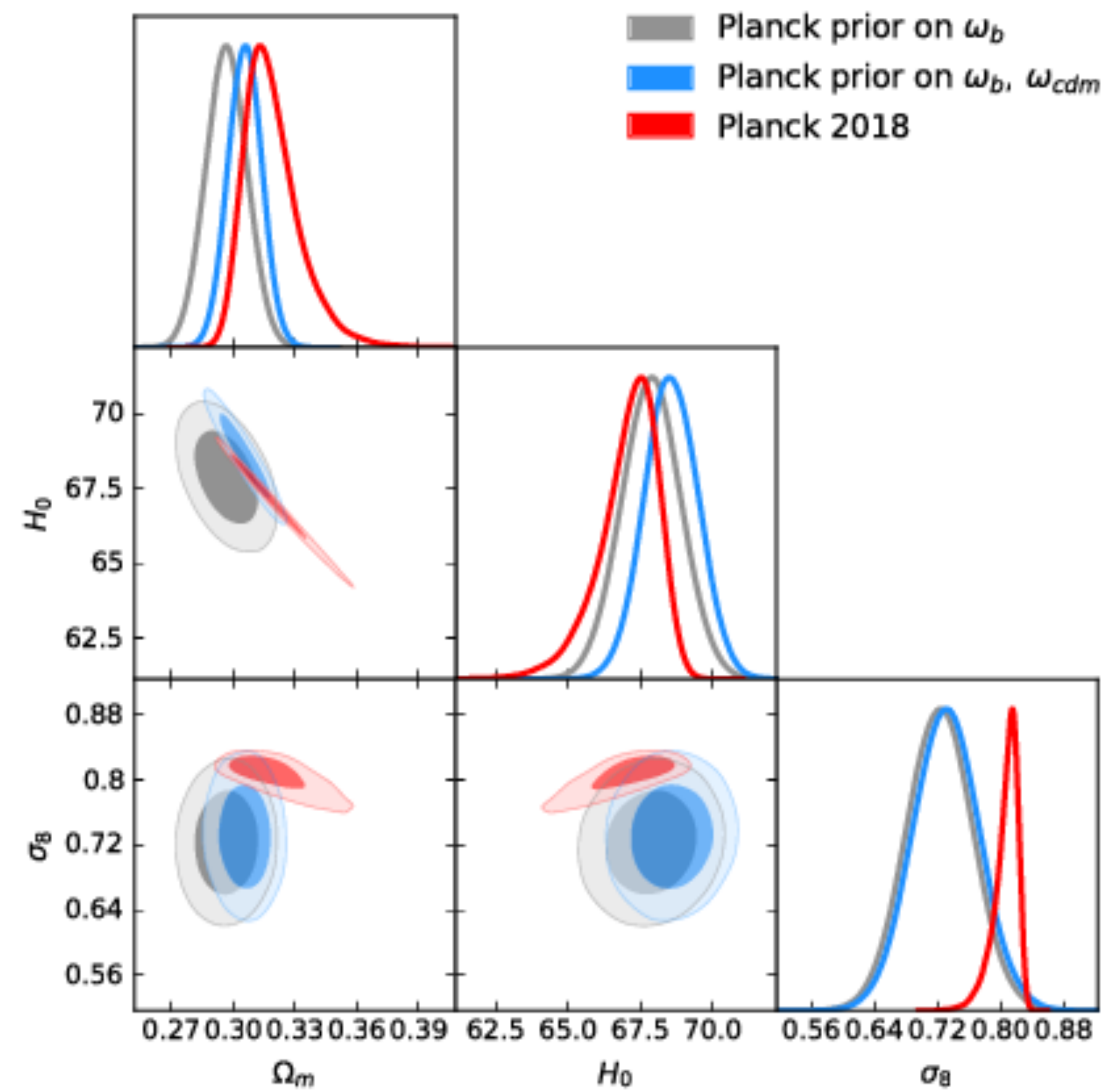
$$\Omega_{\rm tot}, \Omega_b, \Omega_m, \Omega_\Lambda, H_0, n_s = ??$$

•**For given observational results, what parameter sets are preferred??**

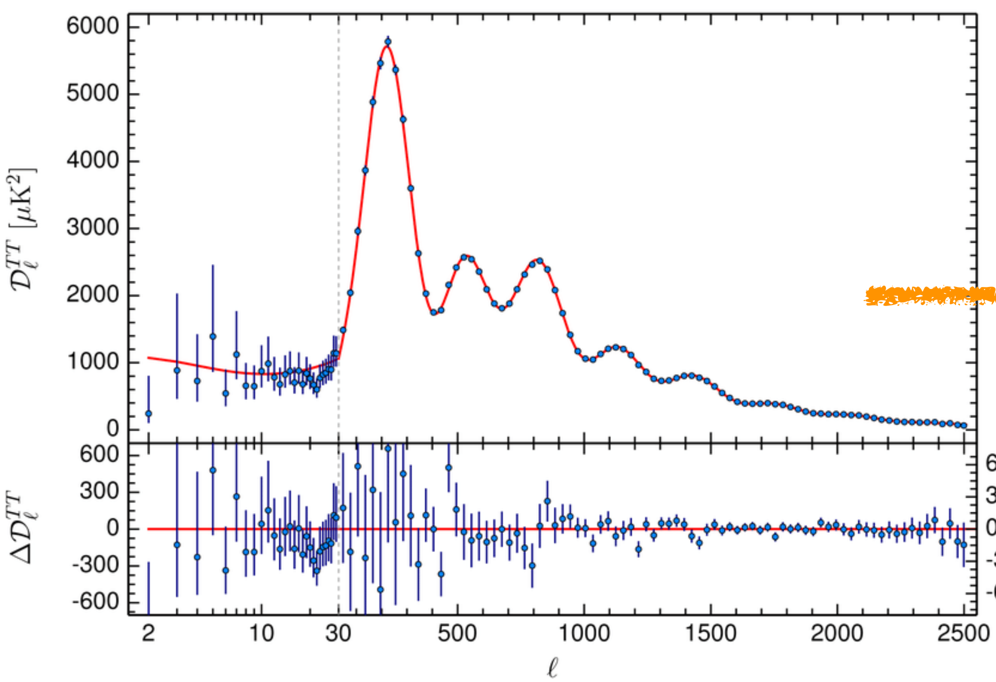


$$\Omega_{\rm tot}, \Omega_b, \Omega_m, \Omega_\Lambda, H_0, n_s = ??$$

•In **Bayesian inference**, we evaluate the **posterior distribution** of parameters with MCMC.



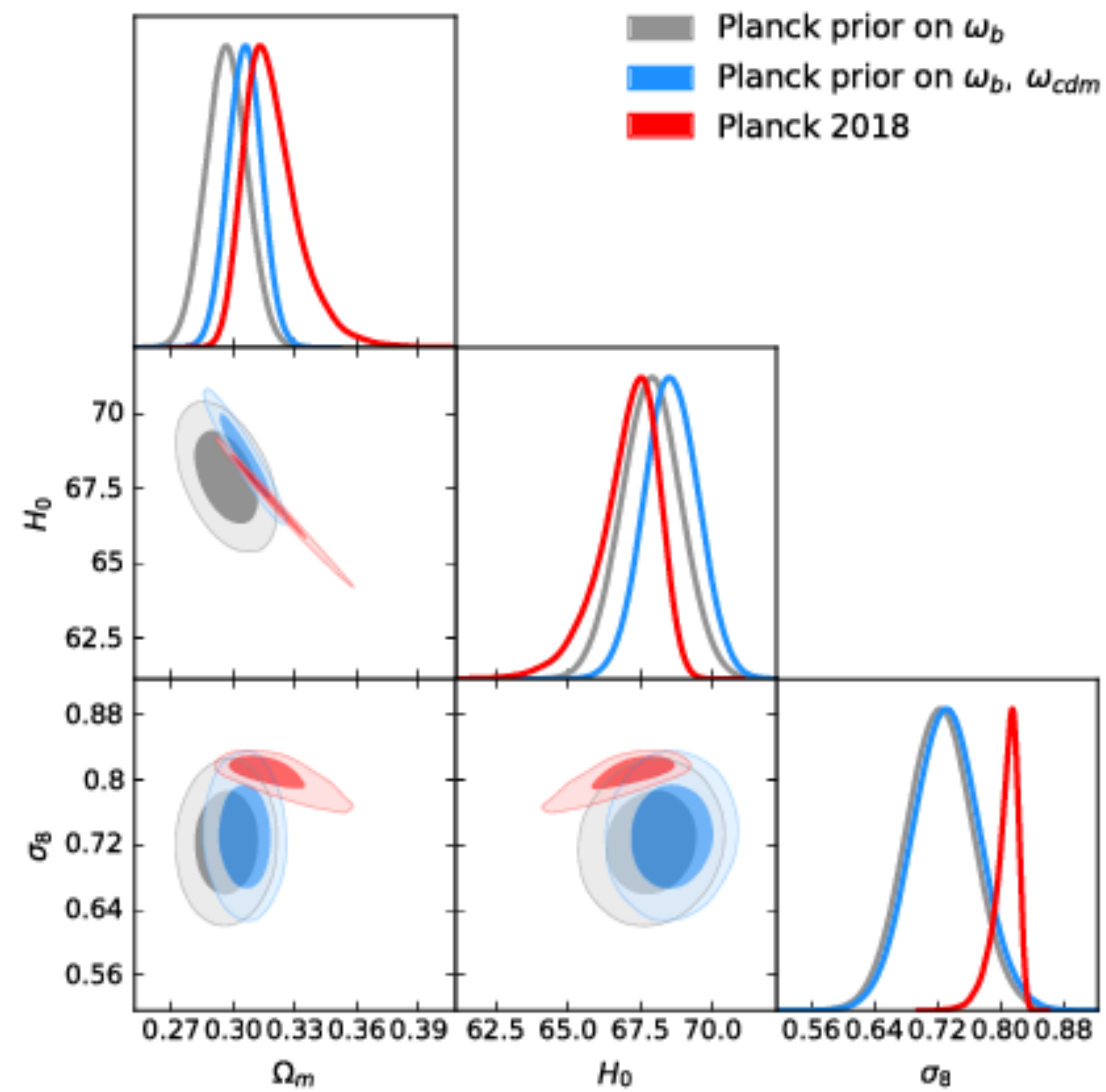
•For given observational results, what parameter sets are preferred??

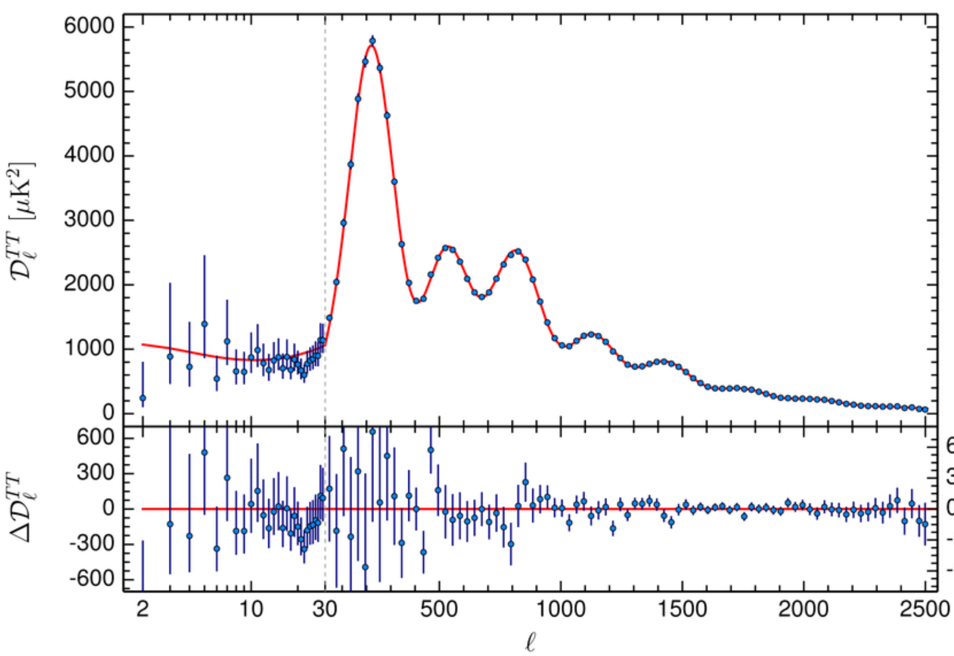


$$\Omega_{\rm tot}, \Omega_b, \Omega_m, \Omega_\Lambda, H_0, n_s = ??$$

•In **Bayesian inference**, we evaluate the **posterior distribution** of parameters with MCMC.

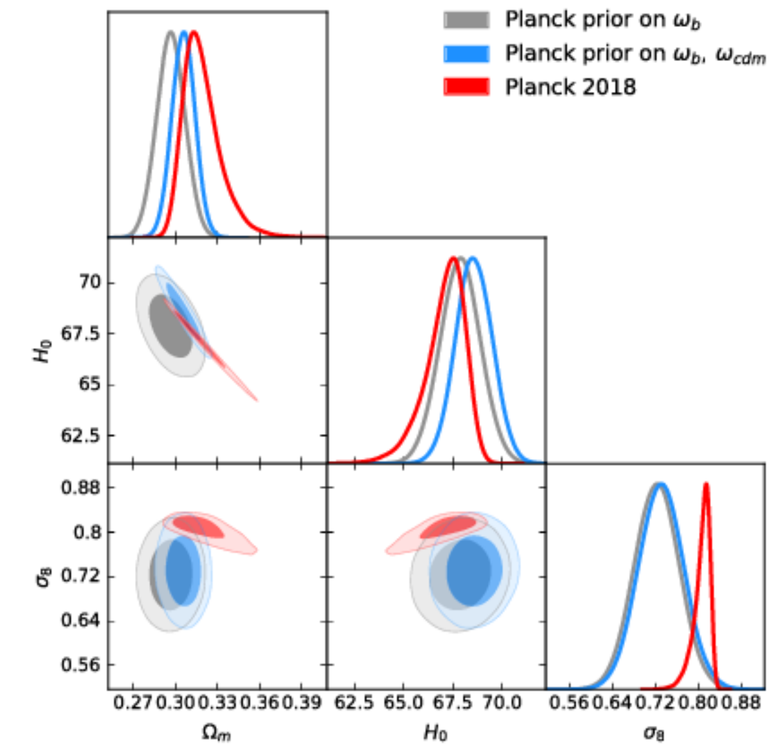
How do you obtain posterior distributions?





posterior

$$p(\text{model} \,|\, \text{data})$$



For given data, we evaluate which theoretical parameter values explain the data well.

Posterior **Likelihood** **Prior**

$$p(\theta \,|\, \mathbf{t_0}) \propto \mathscr{L}(\mathbf{t_0} \,|\, \theta) p(\theta)$$

(Key question 1)

We usually assume Gaussian as likelihood. Can we use a more flexible likelihood?

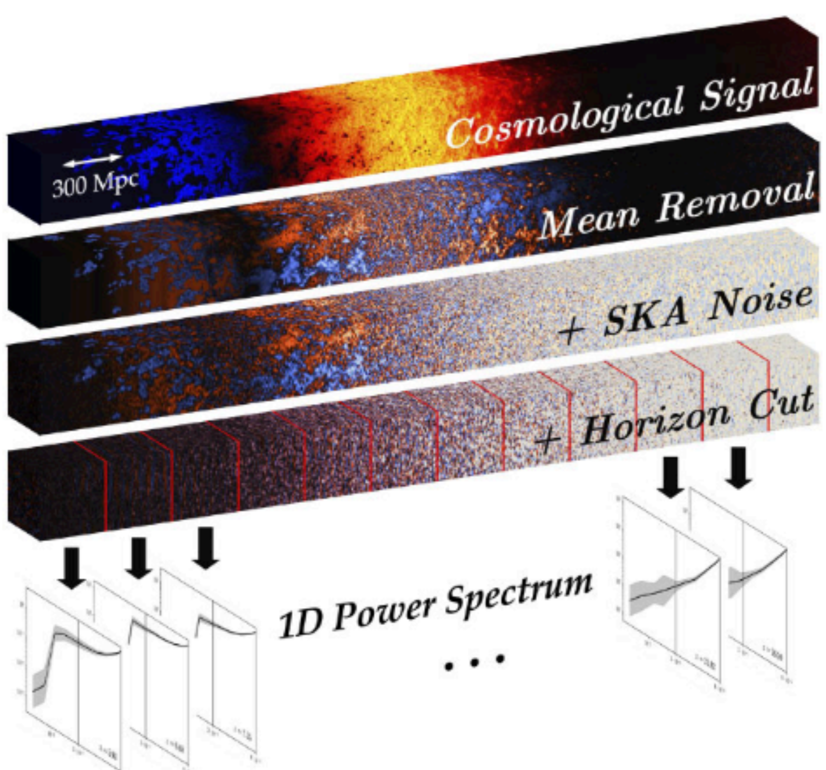
$$\mathscr{L}(t_0 \mid \theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Another parameter estimation approach

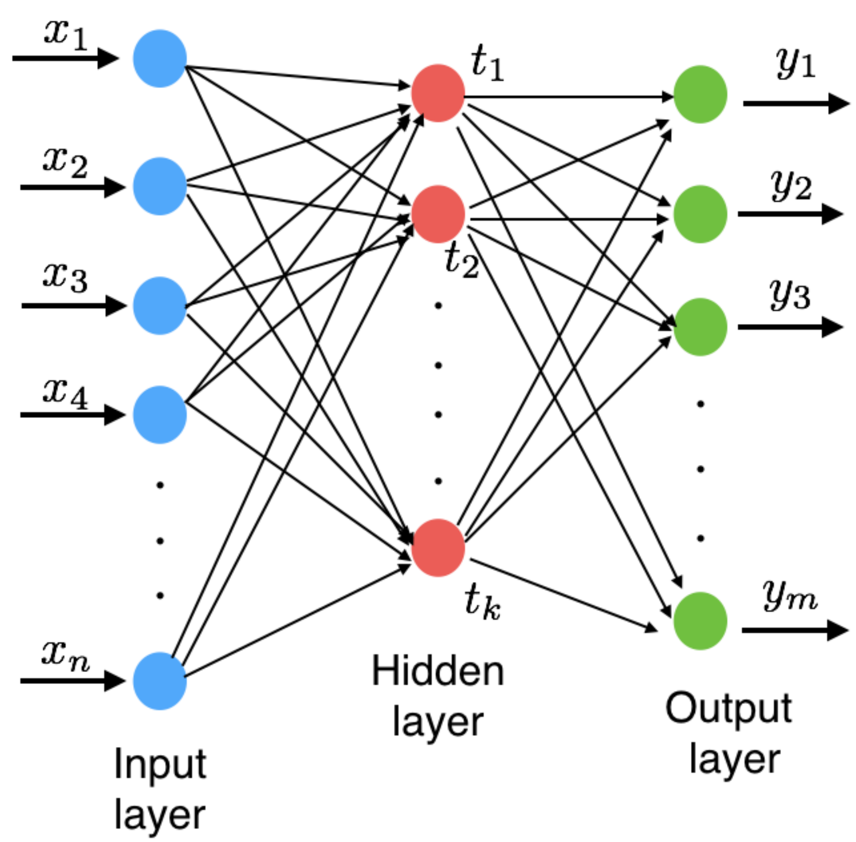
Artificial Neural Network(ANN)

•The ANN is one of the machine learning techniques. Once we train the architecture of ANN by training the dataset, we can apply the trained network to unknown data sets.

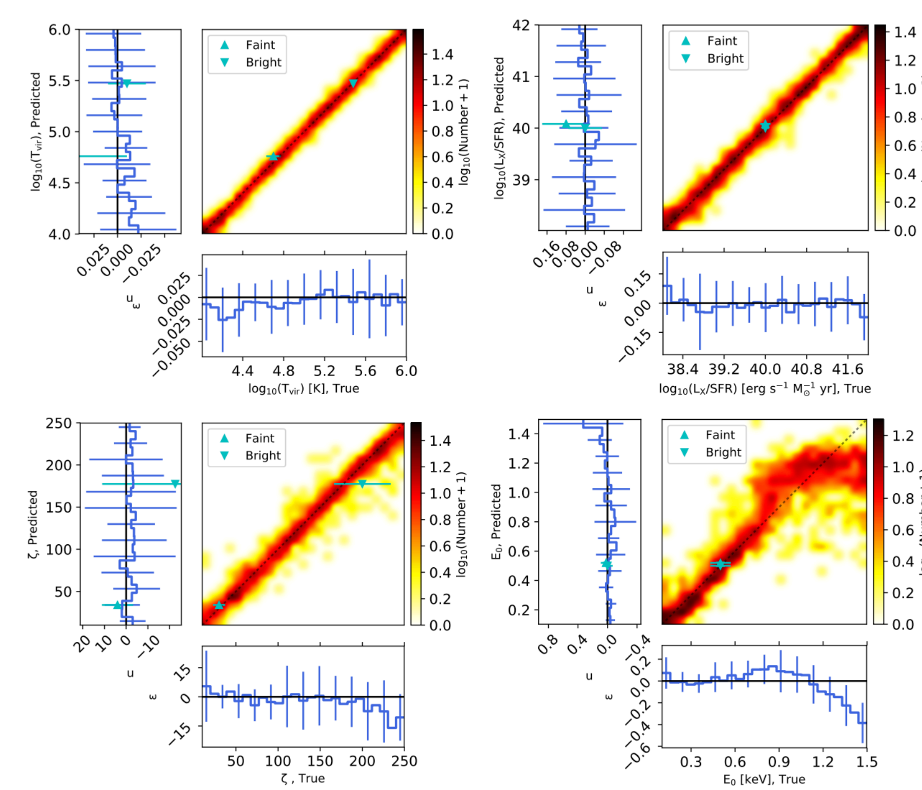
Input



ANN

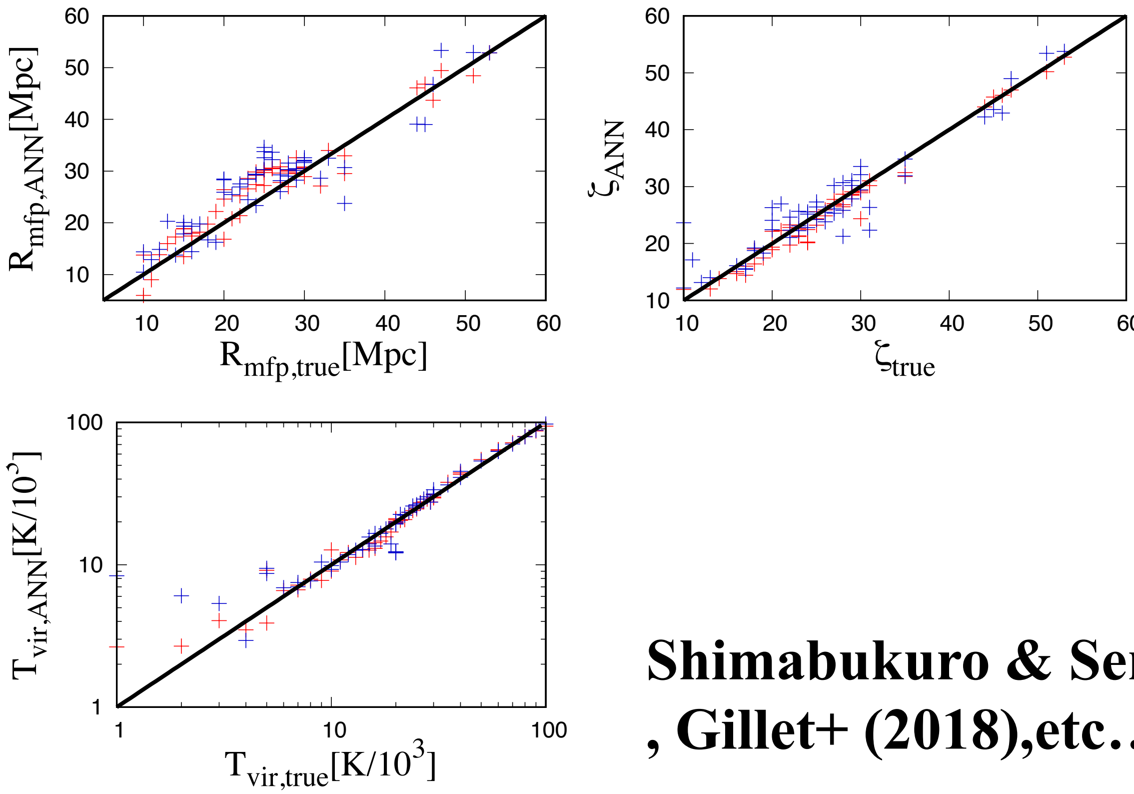


Output

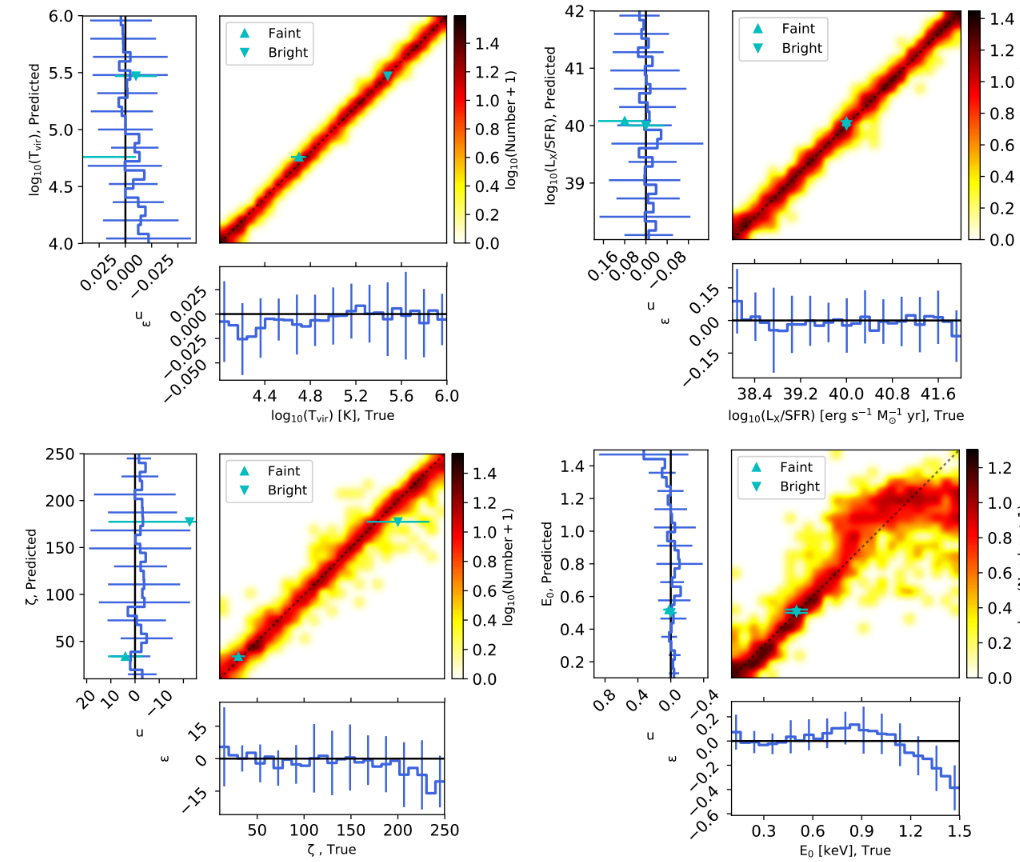


Power spectrum, images

Parameters

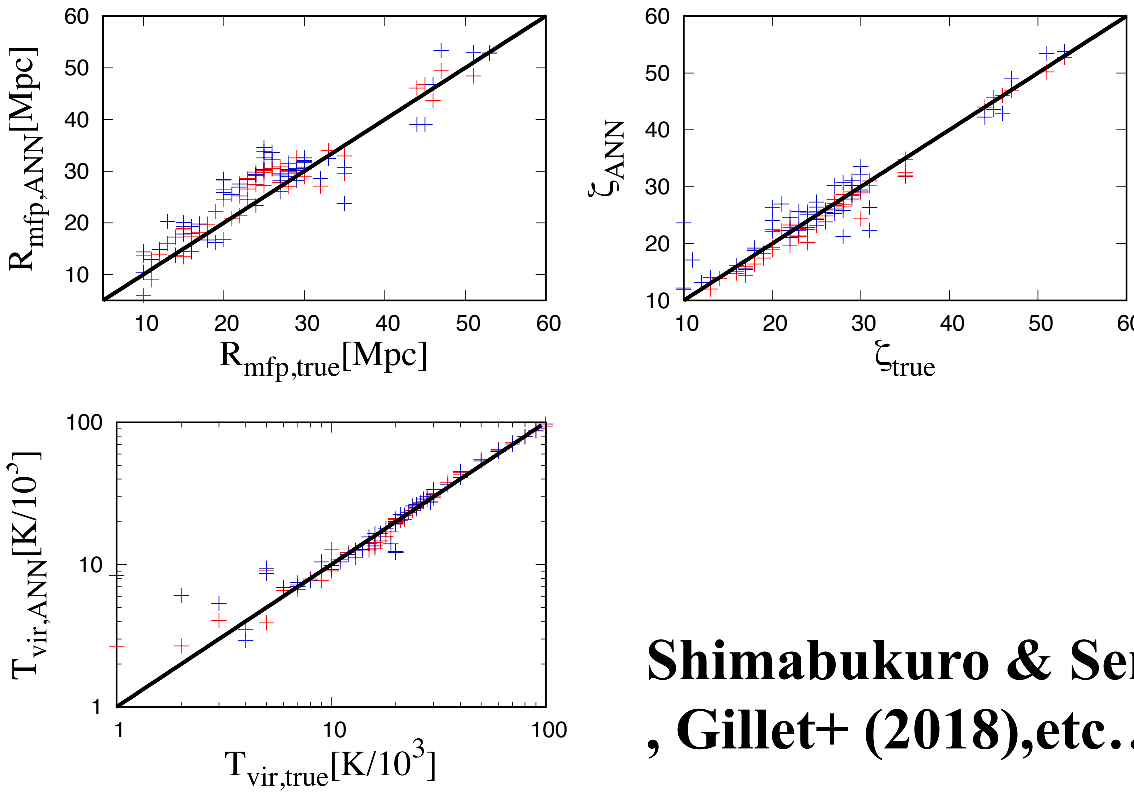


Shimabukuro & Semelin(2017), Gillet+ (2018),etc…

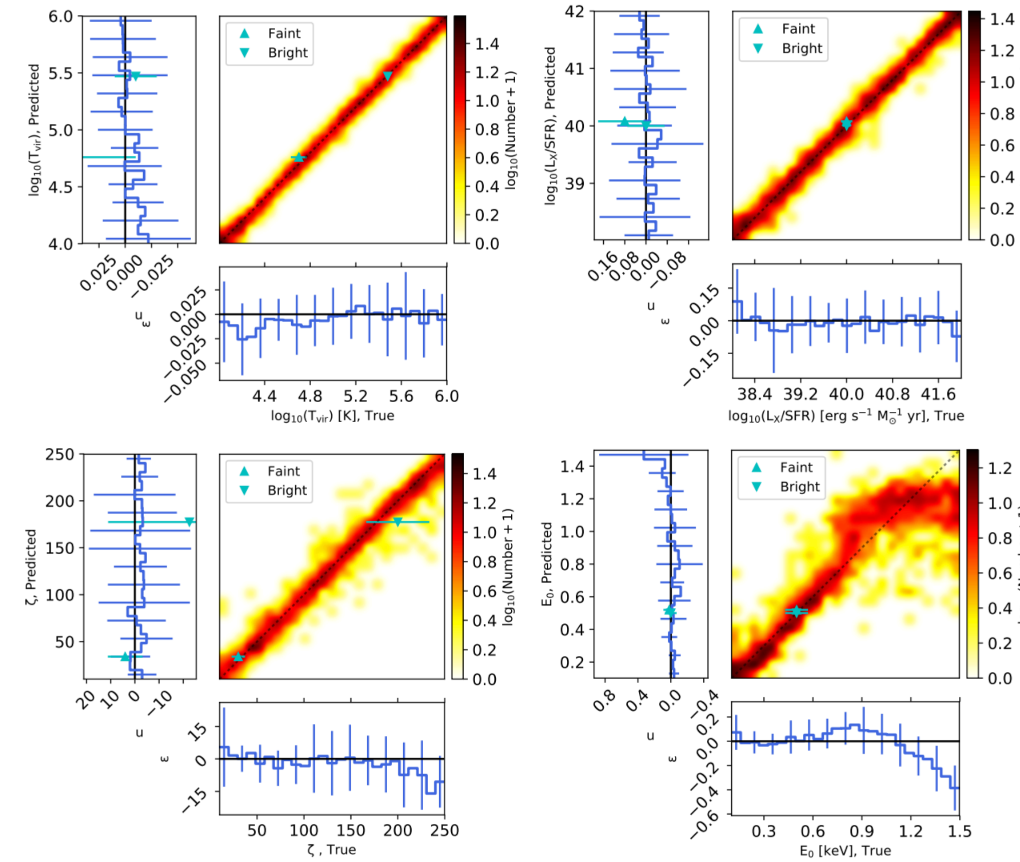


•We usually **DO NOT** evaluate the **uncertainty** of machine learning itself. **ANN just returns "points"**.

•In Bayesian inference with MCMC, we need to calculate likelihood (and prior) to obtain the posterior of parameters.



Shimabukuro & Semelin(2017), Gillet+ (2018),etc…



•We usually **DO NOT** evaluate the **uncertainty** of machine learning itself. **ANN just returns "points"**.

•In Bayesian inference with MCMC, we need to calculate likelihood (and prior) to obtain the posterior of parameters.

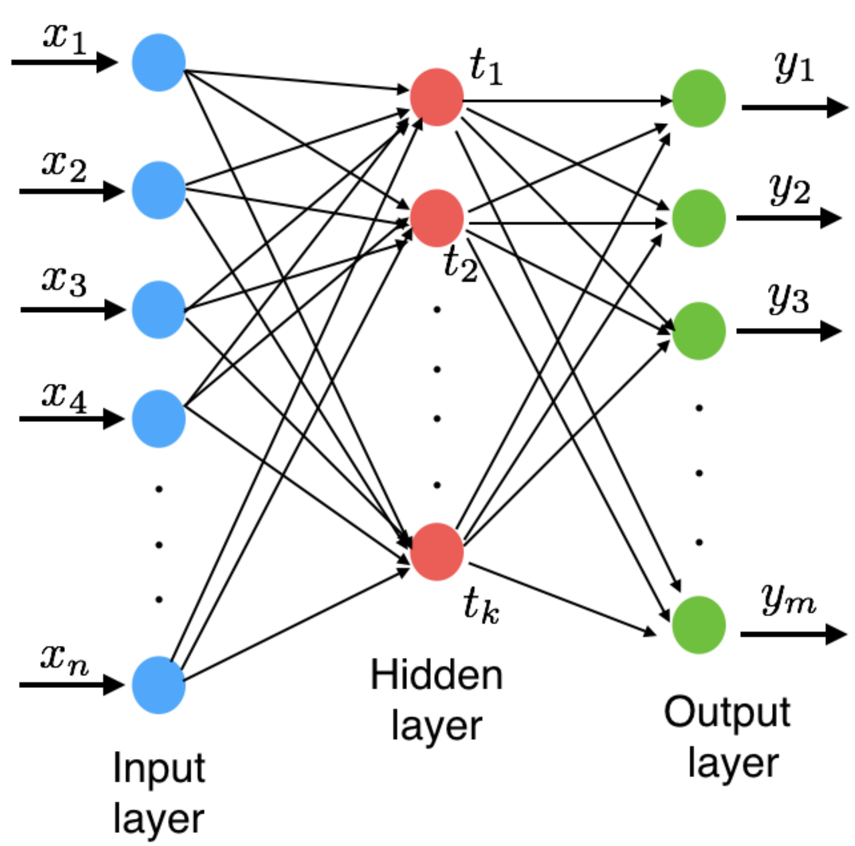
Key question 2.
Can we obtain posterior with ANN direct parameter estimate?

Key Questions

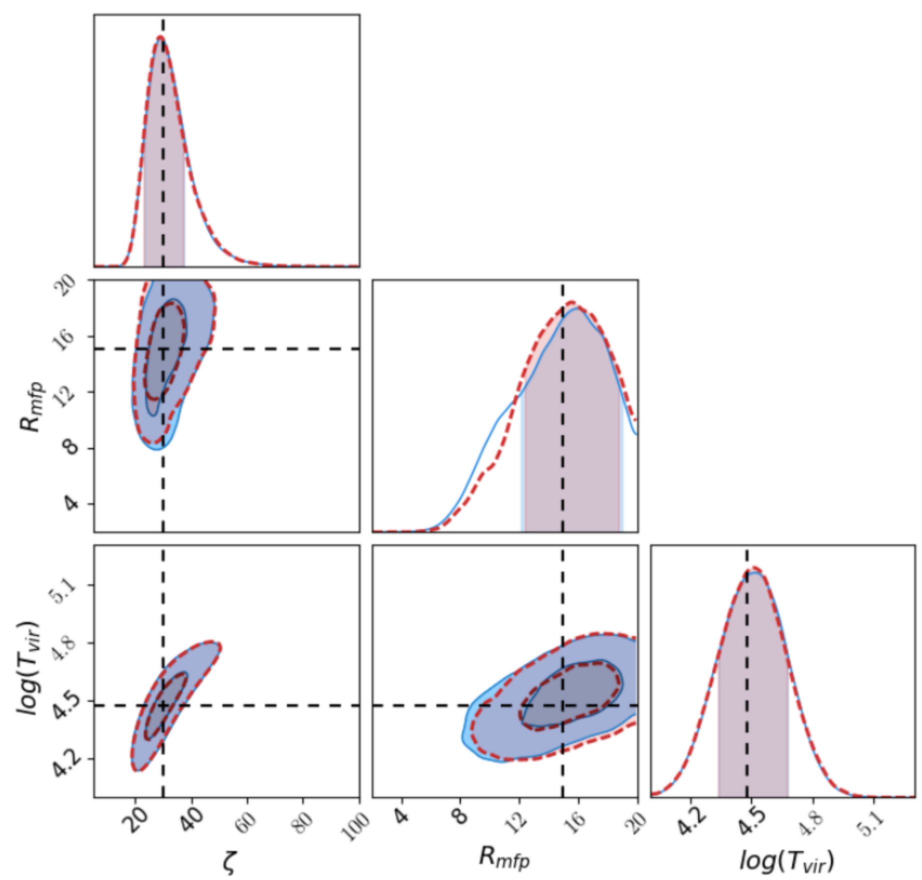
Q1.We usually assume Gaussian as likelihood. Can we use a more flexible likelihood?

$$p(\theta\,|\,\mathbf{t_0}) \propto \mathscr{L}(\mathbf{t_0}\,|\,\theta)p(\theta) \qquad \mathscr{L}(t_0 \mid \theta) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Q2. Can we obtain the posterior with an ANN direct parameter estimate?



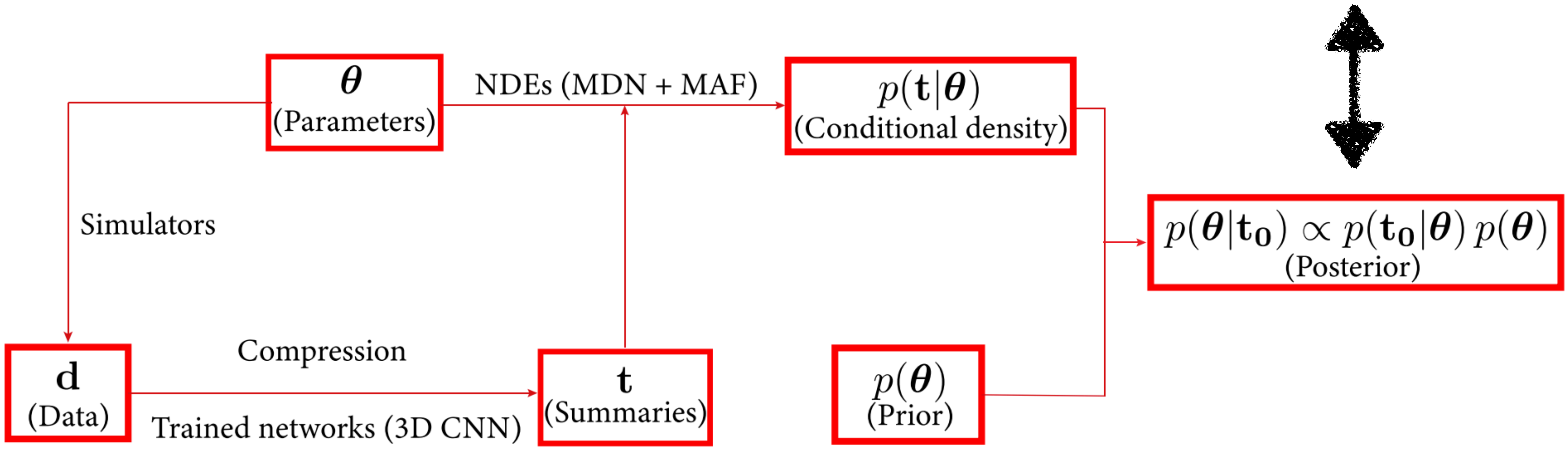
?



Posterior inference with machine learning

- This paper suggested a "**likelihood-free**" approach (**DELFI**, *Density estimation likelihood-free inference*) or **simulation-based inference** in 21cm study. They consider **conditional density distribution** instead of likelihood.

$$p(\theta \,|\, \mathbf{t_0}) \propto \mathscr{L}(\mathbf{t_0} \,|\, \theta) p(\theta)$$

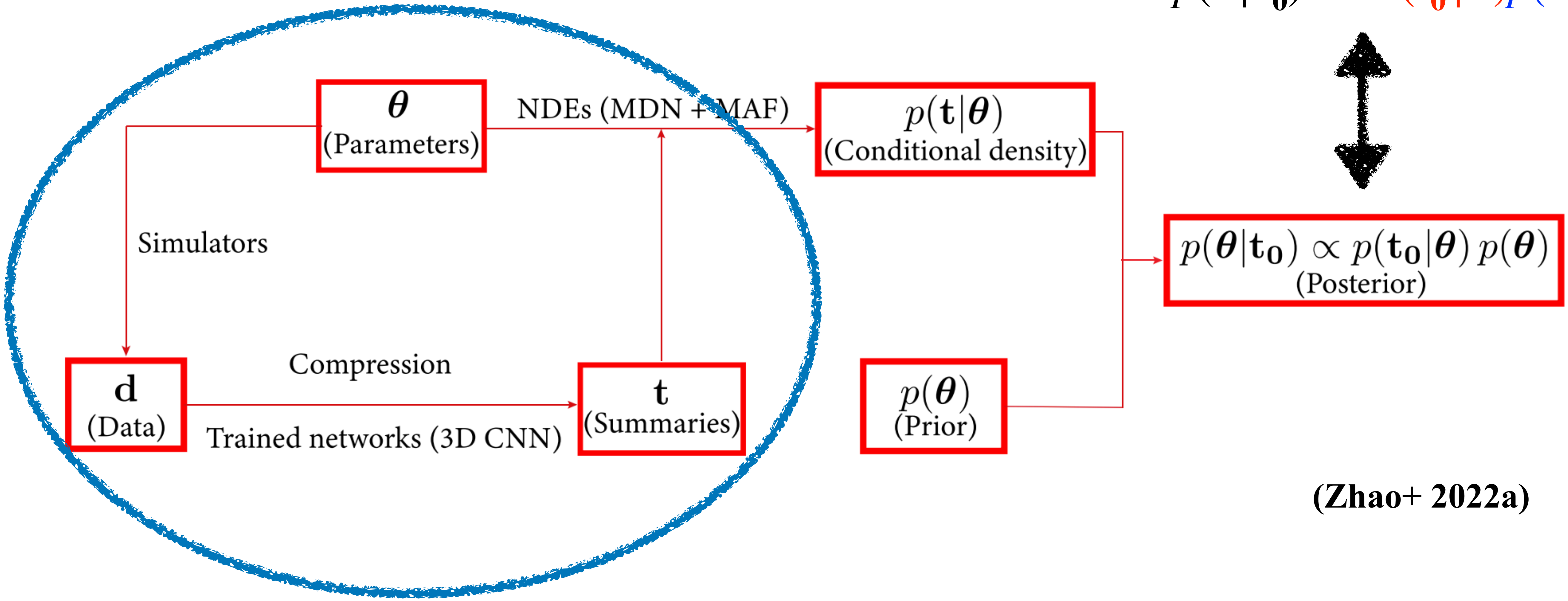


(Zhao+ 2022a)

Posterior inference with machine learning

•This paper suggested a "**likelihood-free**" approach (**DELFI**, *Density estimation likelihood-free inference*) or **simulation-based inference** in 21cm study. They consider **conditional density distribution** instead of likelihood.

$p(\theta\,|\,\mathbf{t_0}) \propto \mathscr{L}(\mathbf{t_0}\,|\,\theta)p(\theta)$

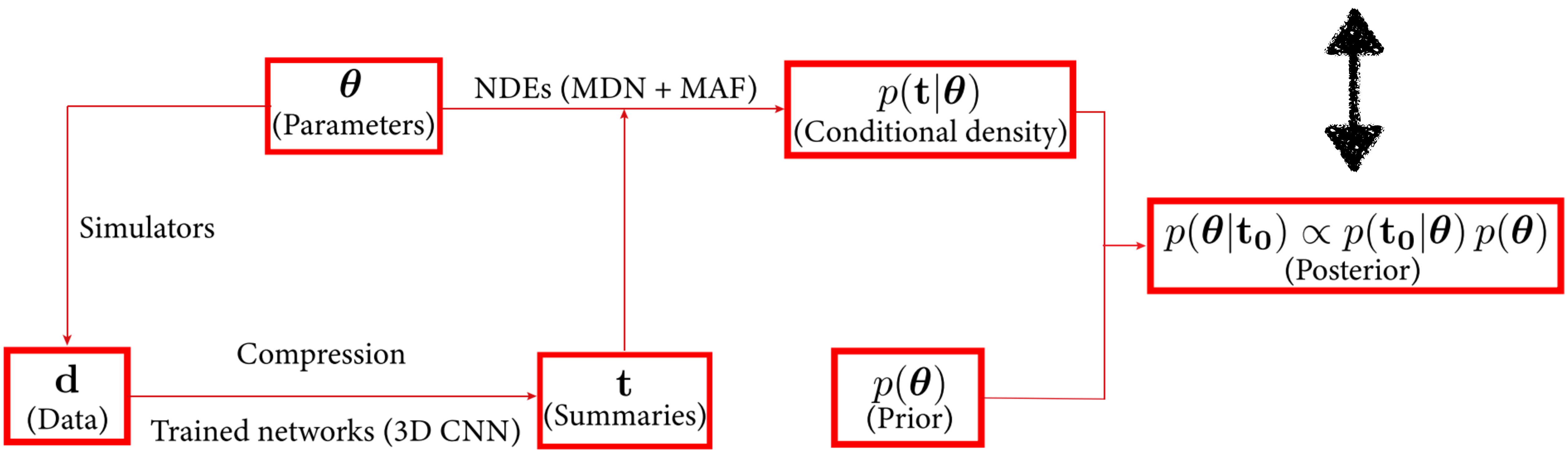


(Zhao+ 2022a)

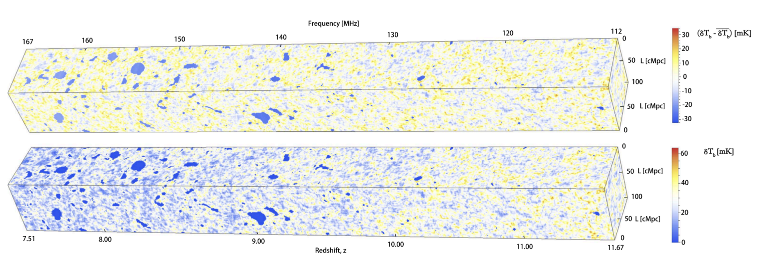
Posterior inference with machine learning

•This paper suggested a "**likelihood-free**" approach (**DELFI**, *Density estimation likelihood-free inference*) or **simulation-based inference** in 21cm study. They consider **conditional density distribution** instead of likelihood.

$$p(\theta \,|\, \mathbf{t_0}) \propto \mathscr{L}(\mathbf{t_0} \,|\, \theta) p(\theta)$$



(Zhao+ 2022a)

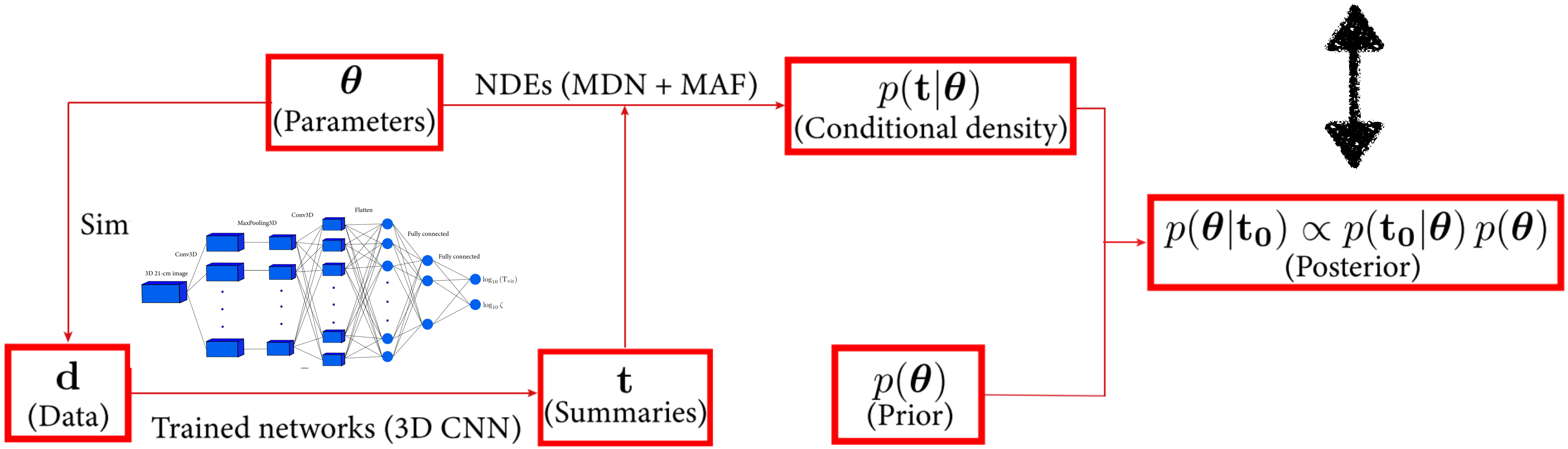


Data: 21cm image

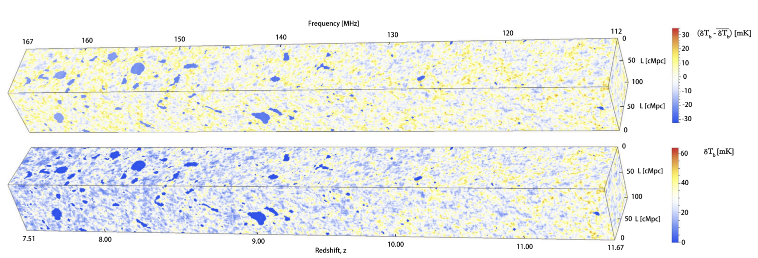
Posterior inference with machine learning

•This paper suggested a “**likelihood-free**” approach (**DELFI**, *Density estimation likelihood-free inference*) or **simulation-based inference** in 21cm study. They consider **conditional density distribution** instead of likelihood.

$$p(\theta\,|\,\mathbf{t}_0) \propto \mathscr{L}(\mathbf{t}_0\,|\,\theta)p(\theta)$$



(Zhao+ 2022a)

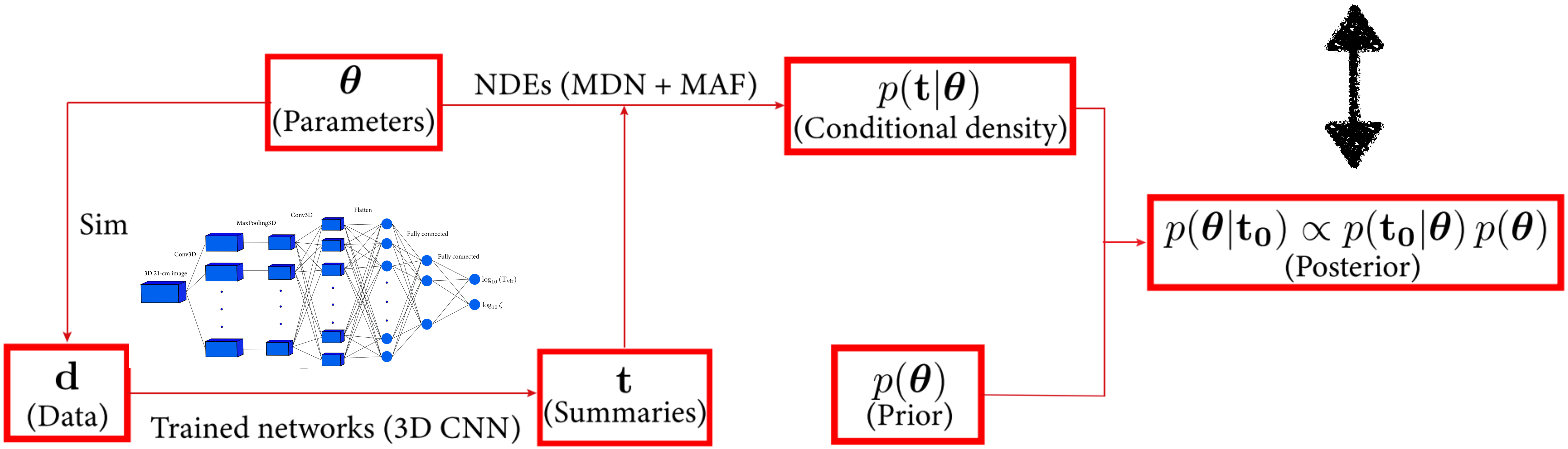


Data: 21cm image

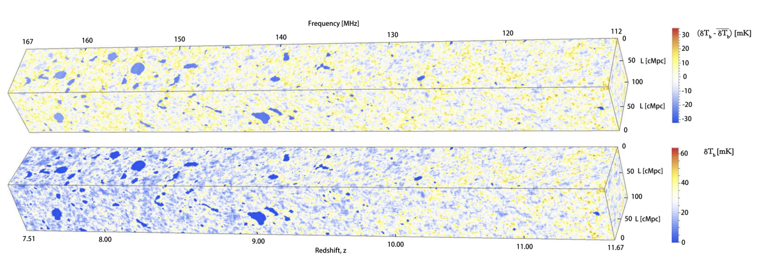
Posterior inference with machine learning

• This paper suggested a "**likelihood-free**" approach (**DELFI**, *Density estimation likelihood-free inference*) or **simulation-based inference** in 21cm study. They consider **conditional density distribution** instead of likelihood.

$$p(\theta\,|\,\mathbf{t_0}) \propto \mathscr{L}(\mathbf{t_0}\,|\,\theta)p(\theta)$$



(Zhao+ 2022a)



Summaries:

EoR parameters(T_{vir}, ζ)

Data: 21cm image

Posterior inference with machine learning

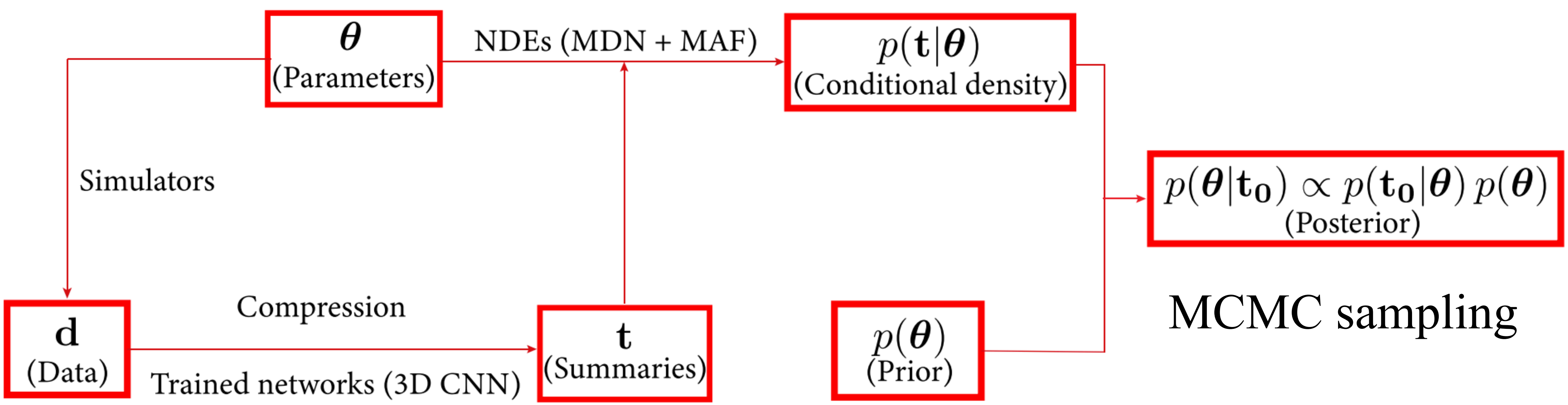
We train neural networks with $\{\theta, t\}$ and obtain conditional density $p(\mathbf{t} \mid \boldsymbol{\theta})$ based on simulations.

•Mixture density networks(**MDN**) (**Bishop 1994**)

•Masked Autoencoder for Density Estimation (**MADE**) (**Papamakarios+ 2017**)

(See also **Alsing+2019**)

Training dataset $\{\theta, \mathbf{t}\}$



MCMC sampling

(**Zhao+ 2022a**)

Posterior inference with machine learning

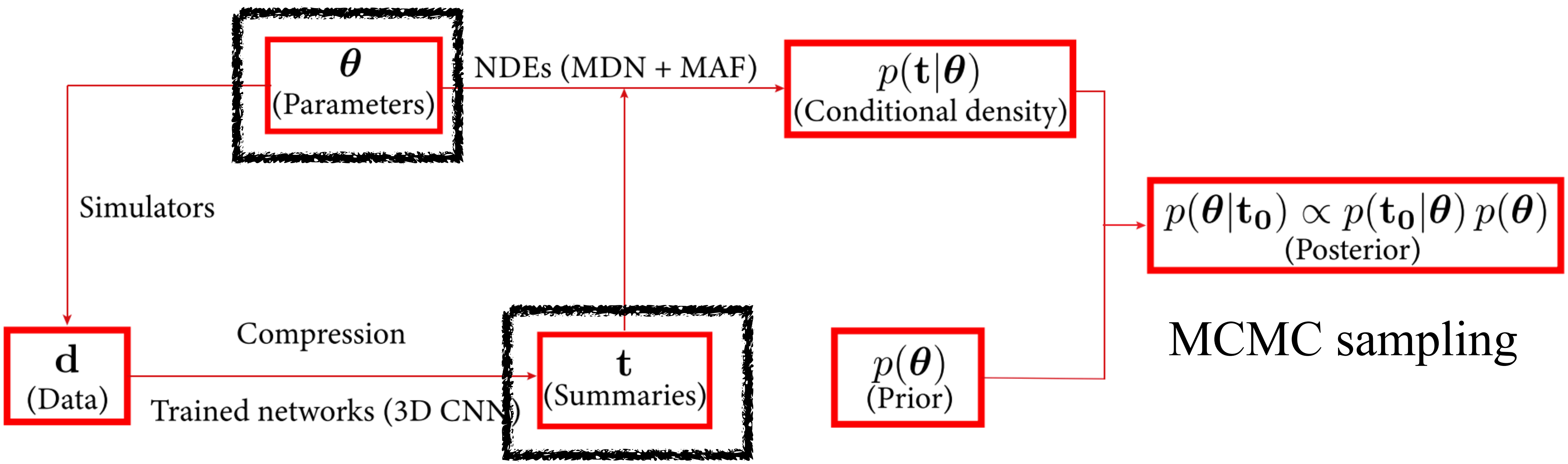
We train neural networks with $\{\theta, t\}$ and obtain conditional density $p(\mathbf{t} \mid \boldsymbol{\theta})$ based on simulations.

•Mixture density networks(**MDN**) (**Bishop 1994**)

•Masked Autoencoder for Density Estimation (**MADE**) (**Papamakarios+ 2017**)

(See also **Alsing+2019**)

Training dataset $\{\theta, \mathbf{t}\}$



MCMC sampling

(Zhao+ 2022a)

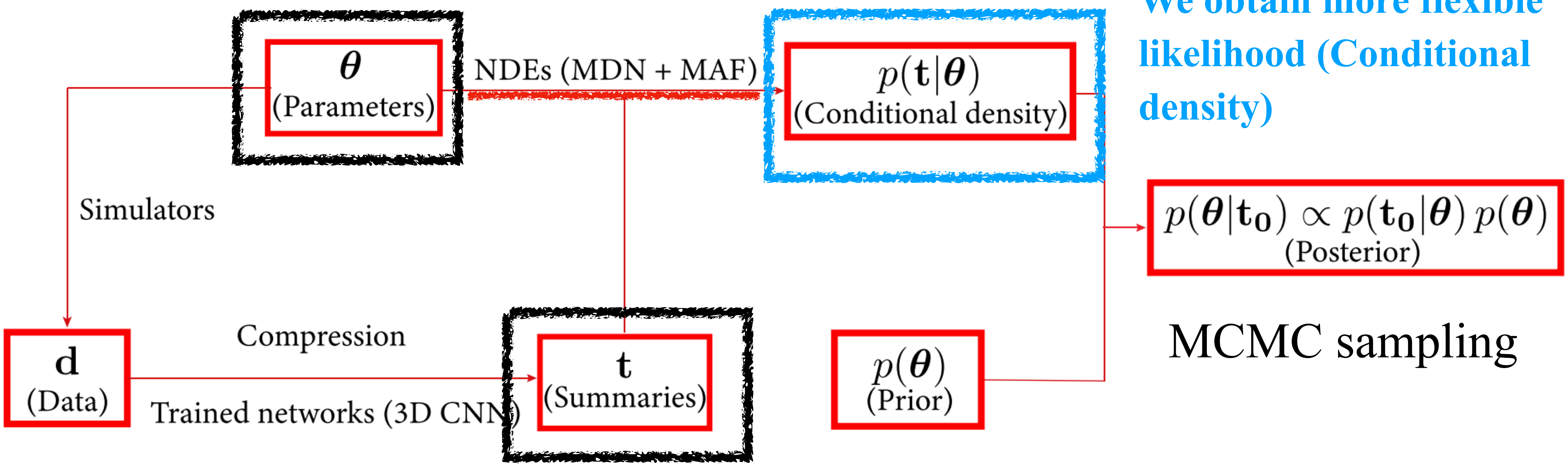
Posterior inference with machine learning

We train neural networks with $\{\theta, t\}$ and obtain conditional density $p(\mathbf{t} \mid \boldsymbol{\theta})$ based on simulations.

- Mixture density networks(**MDN**) (**Bishop 1994**)
- Masked Autoencoder for Density Estimation (**MADE**) (**Papamakarios+ 2017**)

(See also **Alsing+2019**)

Training dataset $\{\theta, \mathbf{t}\}$



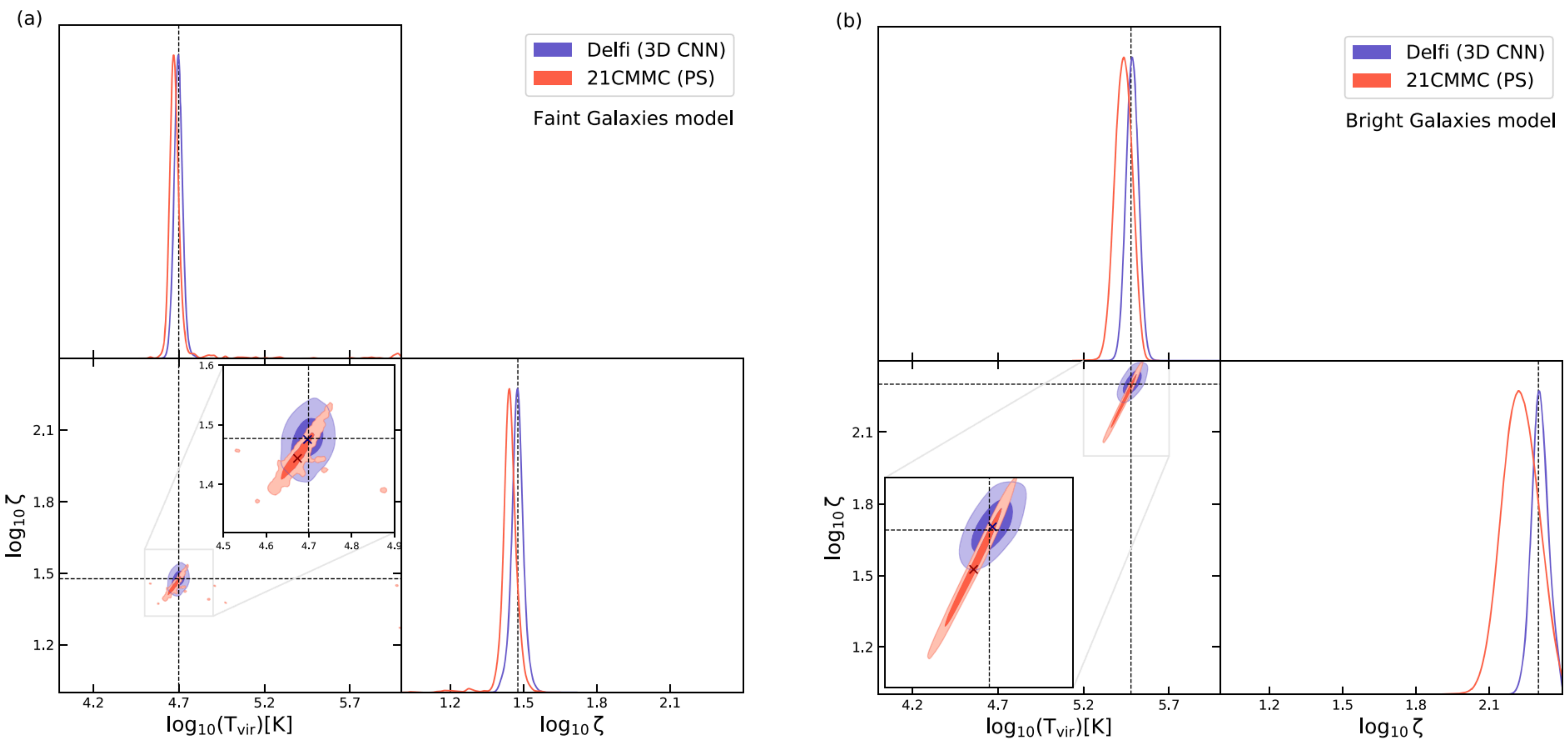
We obtain more flexible likelihood (Conditional density)

MCMC sampling

(Zhao+ 2022a)

Posterior from 21cm image and PS

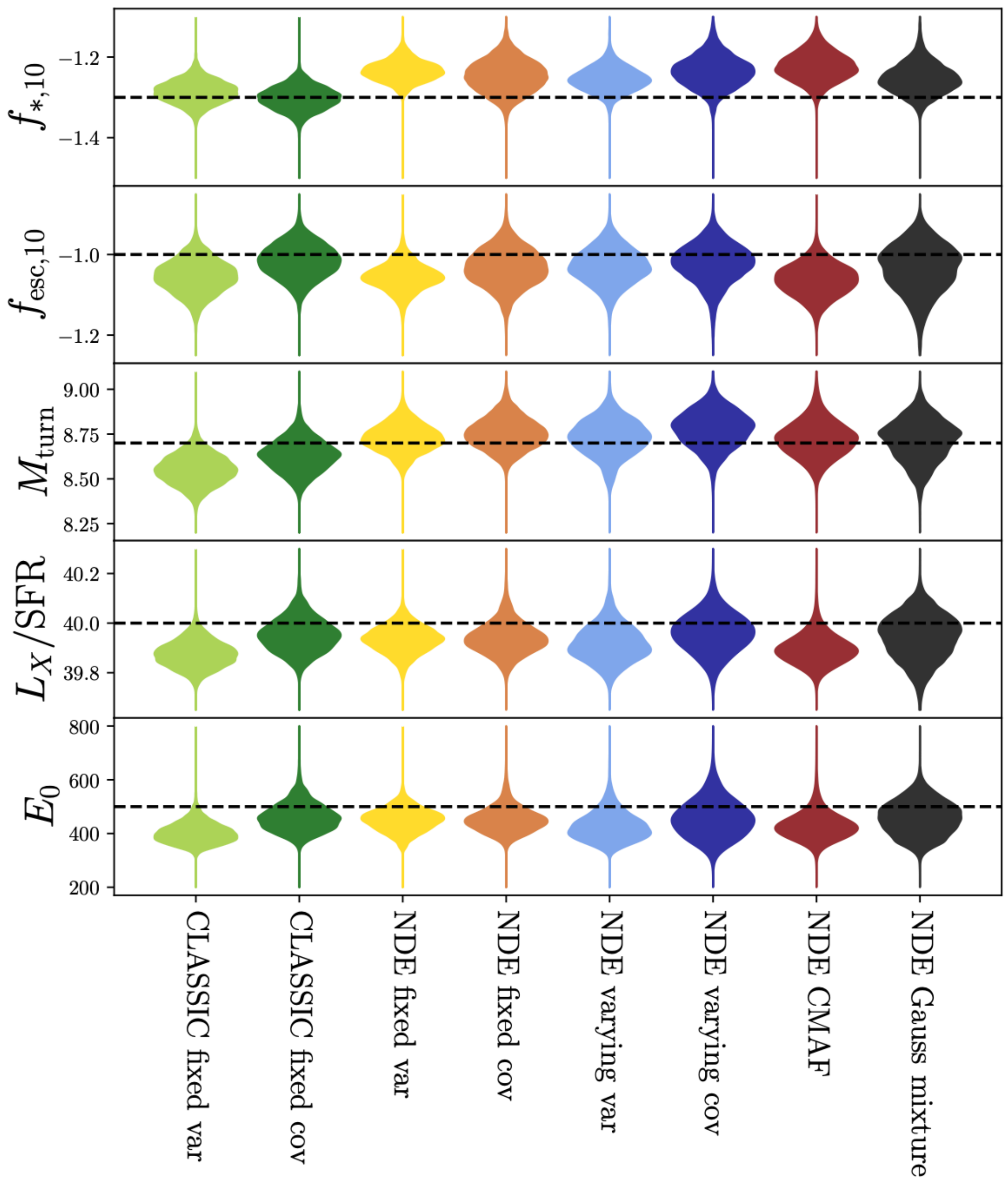
•We can directly compare the posterior obtained from 21cm image map with the posterior obtained from 21cm PS with MCMC.



(Zhao+ 2022a)

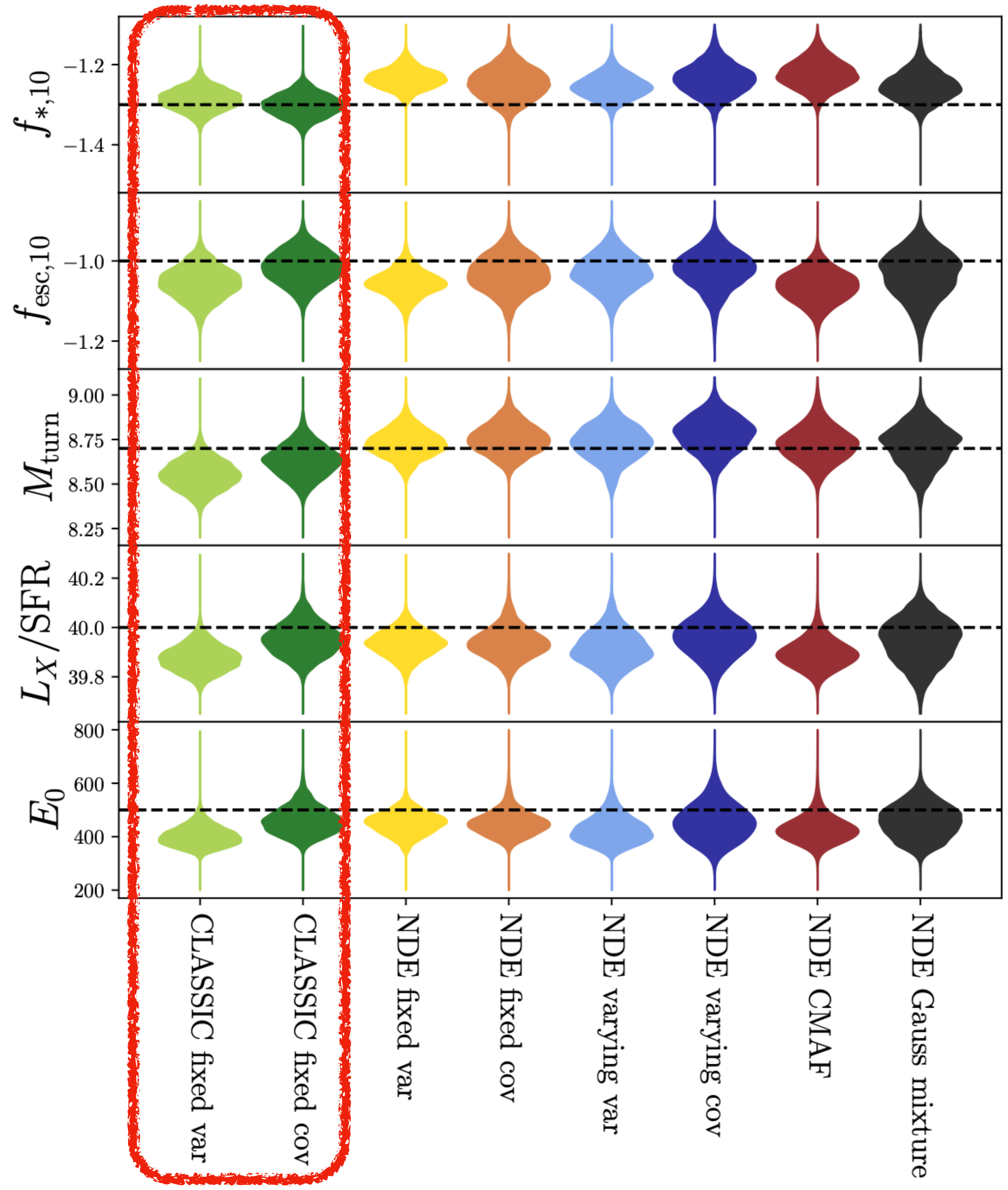
21cm image map can provide tighter constraints on EoR parameters than 21cm PS

Exploring the likelihood



Prelogovic & Mesinger(2023)

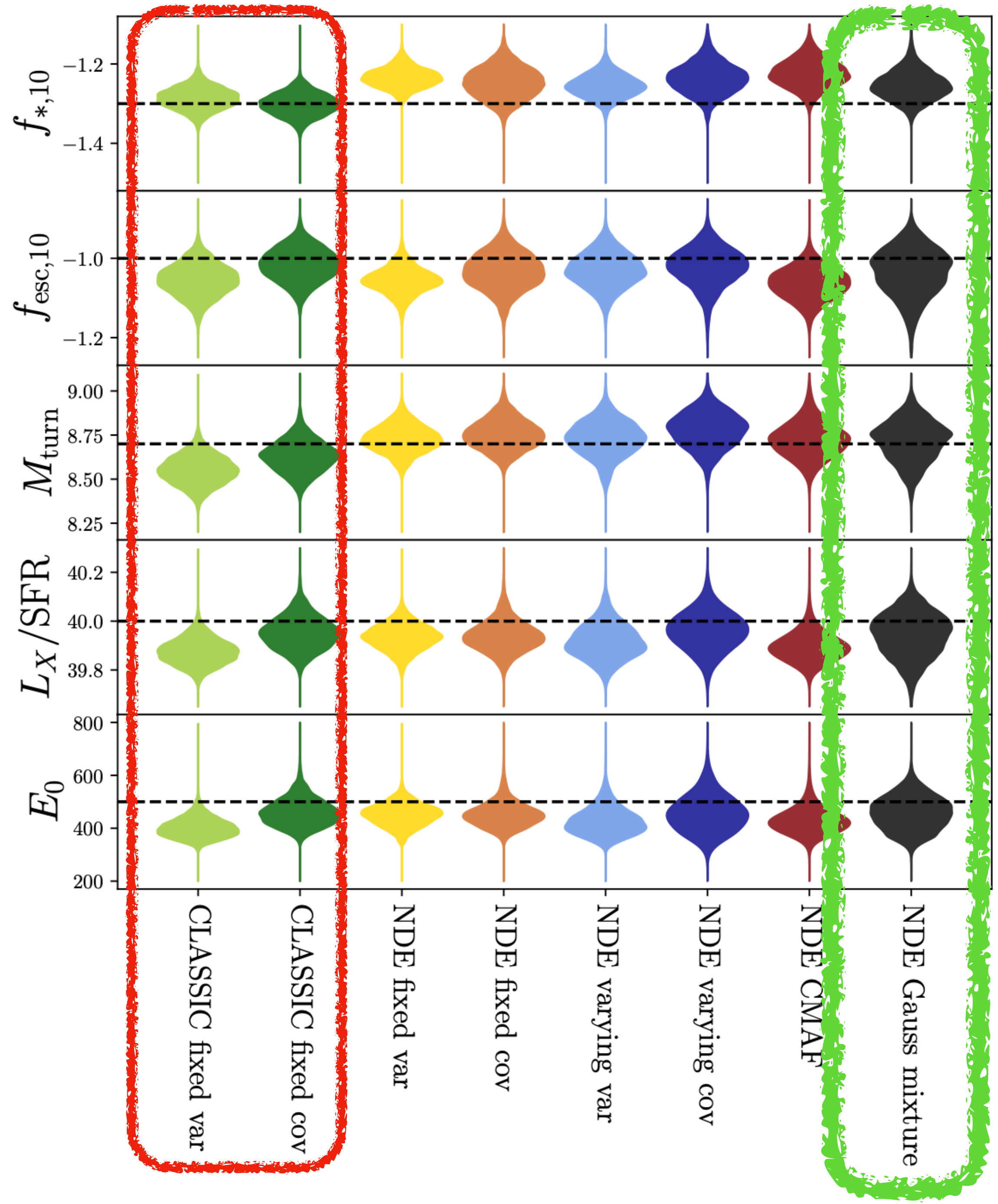
Exploring the likelihood



•If **we include a covariance matrix** in Gaussian-likelihood, parameter inference is improved.

Prelogovic & Mesinger(2023)

Exploring the likelihood

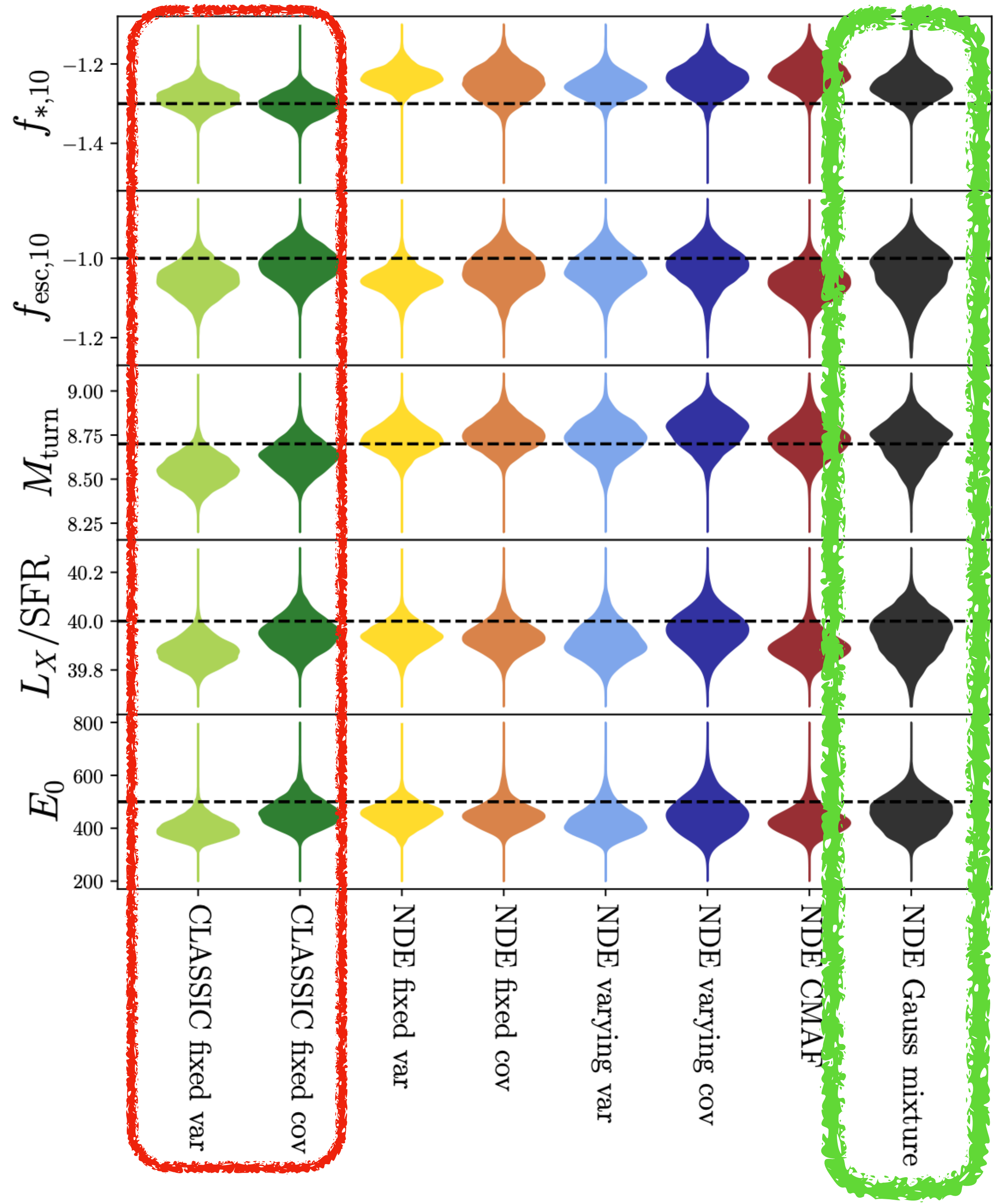


•If **we include a covariance matrix** in Gaussian-likelihood, parameter inference is improved.

•**NDE(neural density estimator) with Gaussian-mixture** provides best constraints on parameters.

Prelogovic & Mesinger(2023)

Exploring the likelihood



•If **we include a covariance matrix** in Gaussian-likelihood, parameter inference is improved.

•**NDE(neural density estimator) with Gaussian-mixture** provides best constraints on parameters.

The **non-Gaussian likelihood including covariance matrix** is better than the Gaussian likelihood for parameter inference from 21cm power spectrum.

Prelogovic & Mesinger(2023)

Faculty Journal club @ 2024/4/10

Simulation-Based Inference of Reionization Parameters From 3D Tomographic 21 cm Lightcone Images

Xiaosheng Zhao,[1] Yi Mao,[1] Cheng Cheng,[2] and Benjamin D. Wandelt[3,4,5]

[1] *Department of Astronomy, Tsinghua University, Beijing 100084, China*
[2] *School of Chemistry and Physics, University of KwaZulu-Natal, Westville Campus, Durban, 4000, South Africa*
[3] *Sorbonne Université, CNRS, UMR 7095, Institut d'Astrophysique de Paris (IAP), 98 bis bd Arago, 75014 Paris, France*
[4] *Sorbonne Université, Institut Lagrange de Paris (ILP), 98 bis bd Arago, 75014 Paris, France*
[5] *Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*

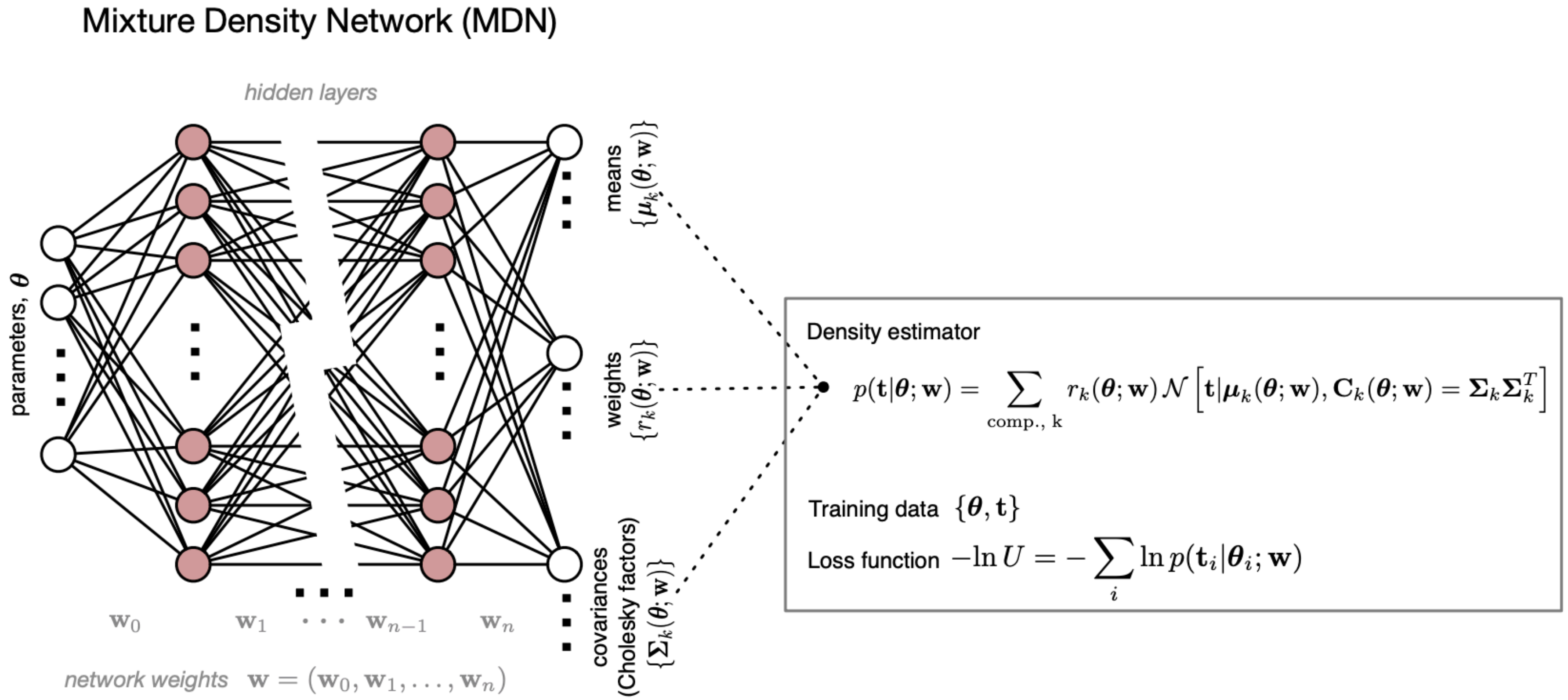
Astro-ph/ 2105.03344

(Take home messages)

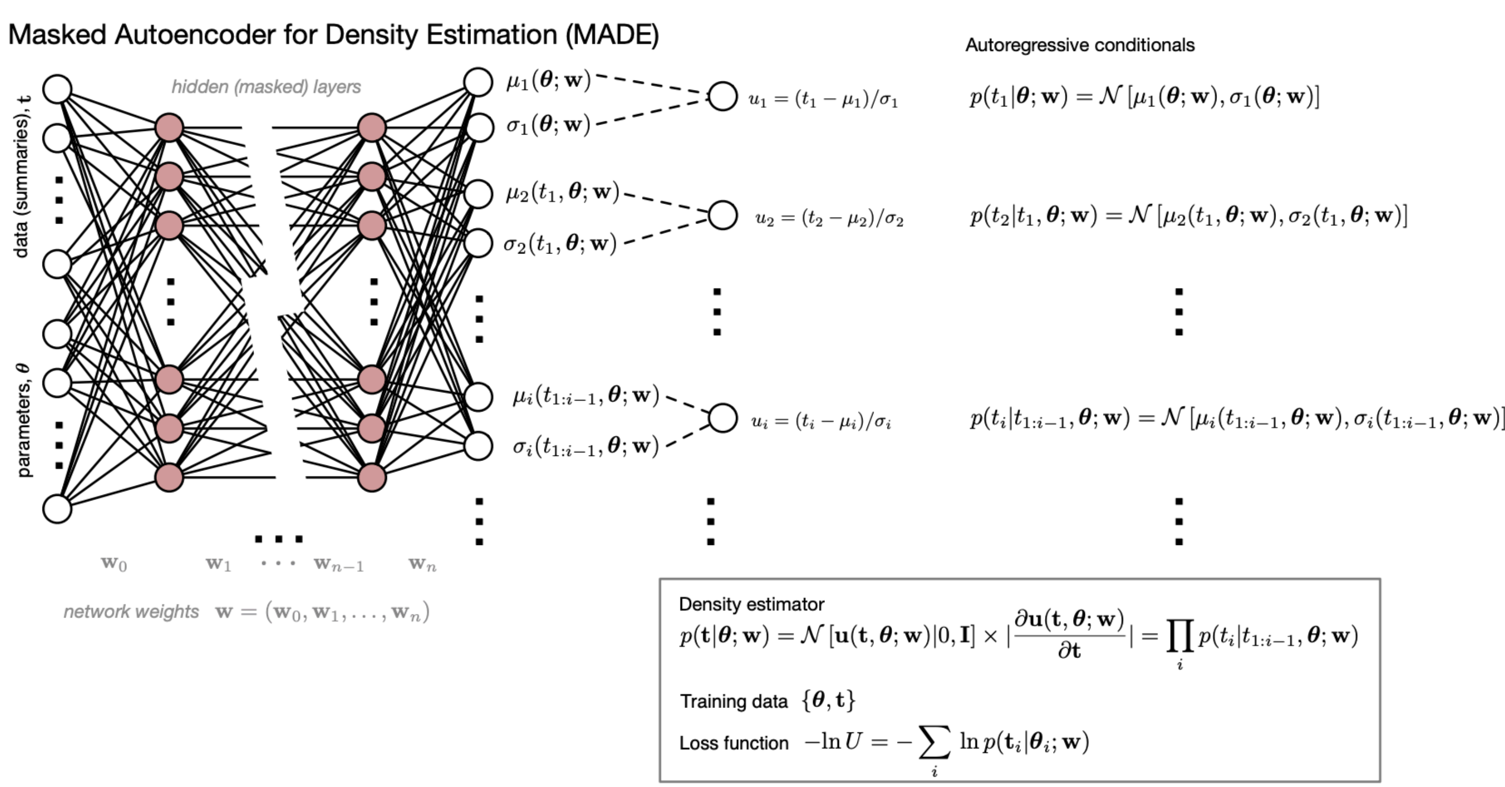
- **Simulation-based inference (SBI)** constructs Likelihood through forward simulation by using **Density estimation likelihood-free inference (DELFI)**

- **SBI** + **machine learning** + **Bayesian inference** → **posterior distribution**

Mixture density networks(MDN) (Bishop 1994)



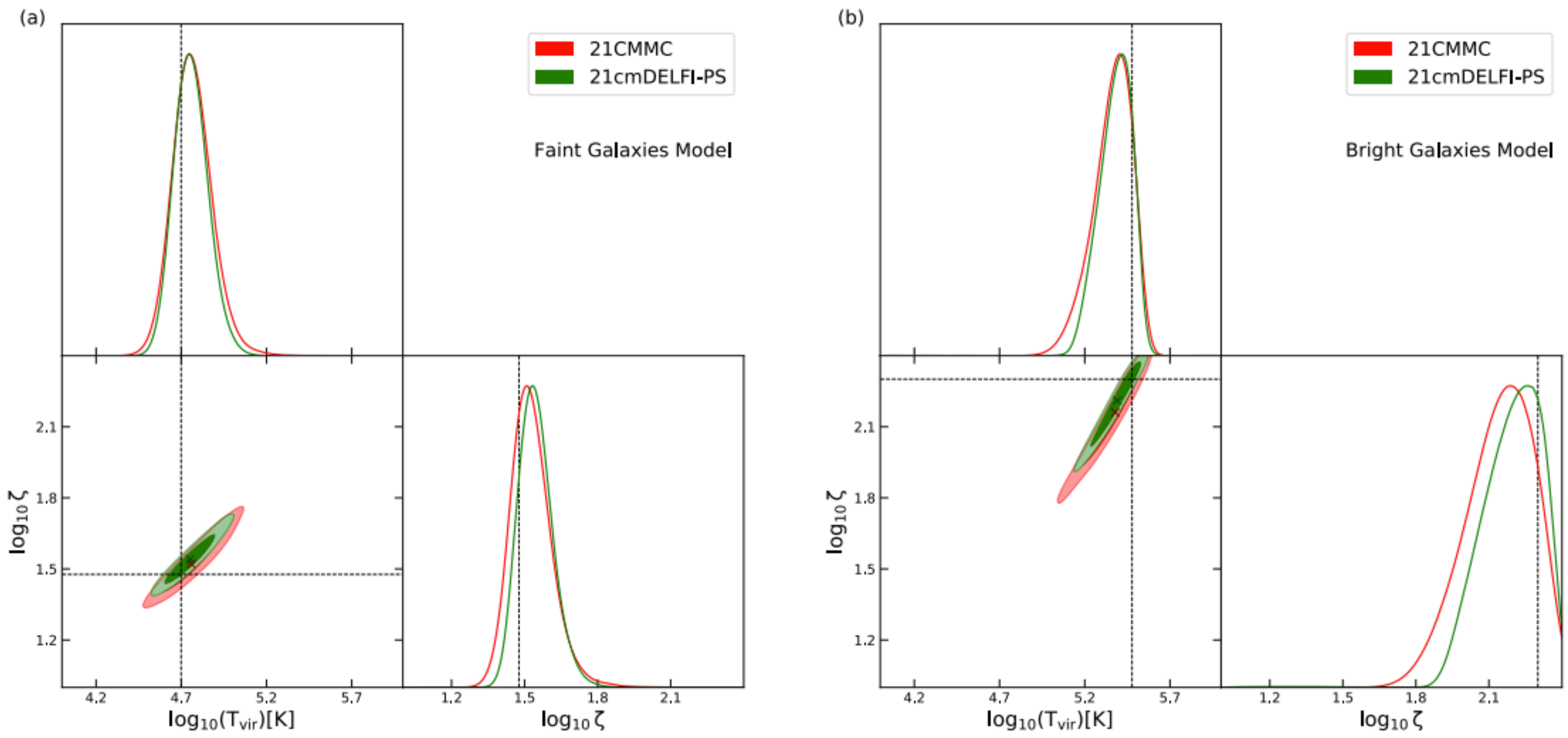
Masked Autoencoder for Density Estimation (MADE) (Papamakarios et al 2017)



(See also Alsing+2019,Wang+2020)

Comparing posteriors

•**We can also compare the posterior obtained from 21cm PS by MCMC with posterior obtained by machine learning based approach (DELFI).**



(Zhao+ 2022b)

•The posterior probability distribution can be obtained with the same accuracy when MCMC is performed and when DELFI is applied.