Probing EoR with SKA

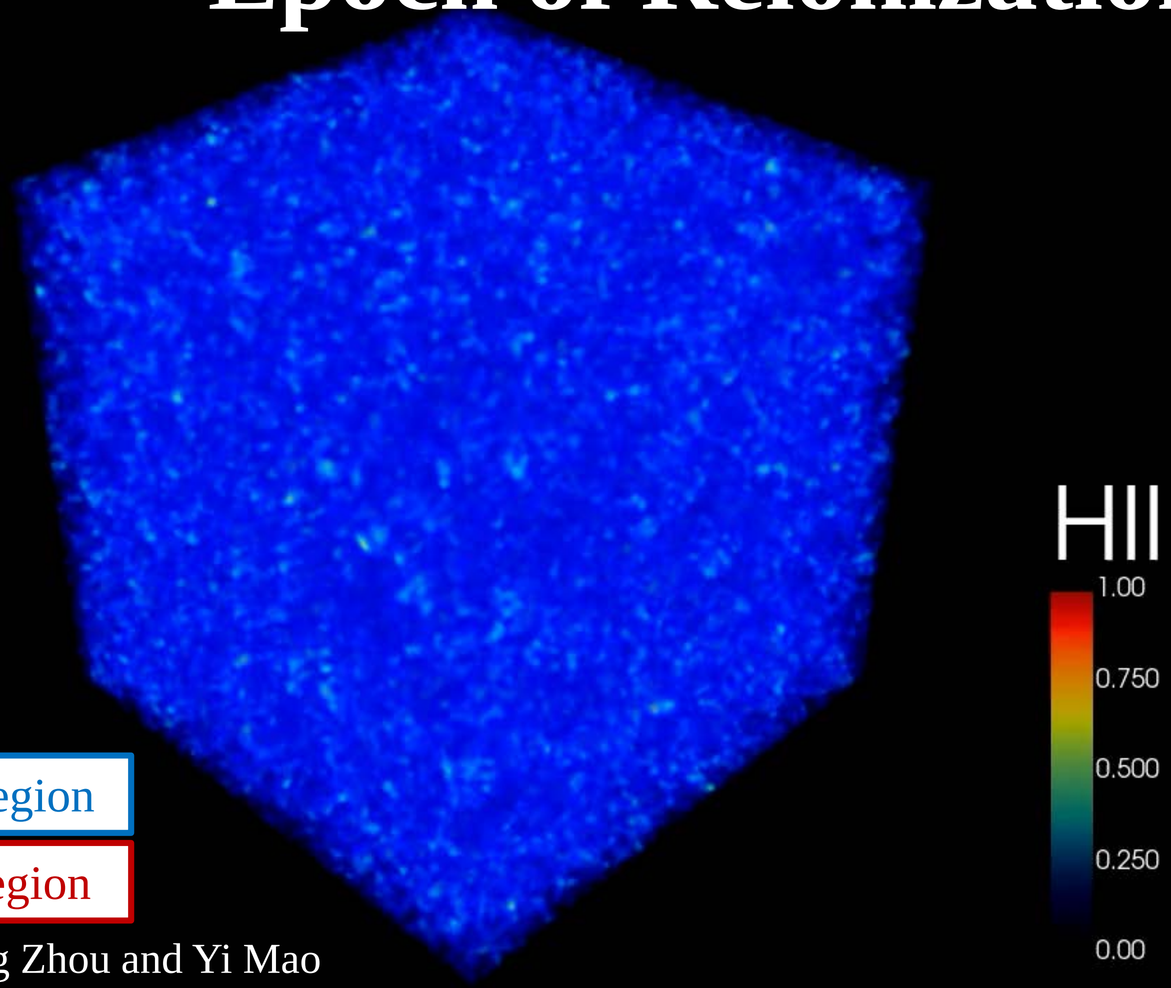
Yi Mao（茅奕）
Tsinghua University

21 cm Cosmology Workshop 2023
Northeastern University & NAOC
July 17, 2023. Shenyang

清華大学天文系
Department of Astronomy, Tsinghua University

z=15.062

Epoch of Reionization

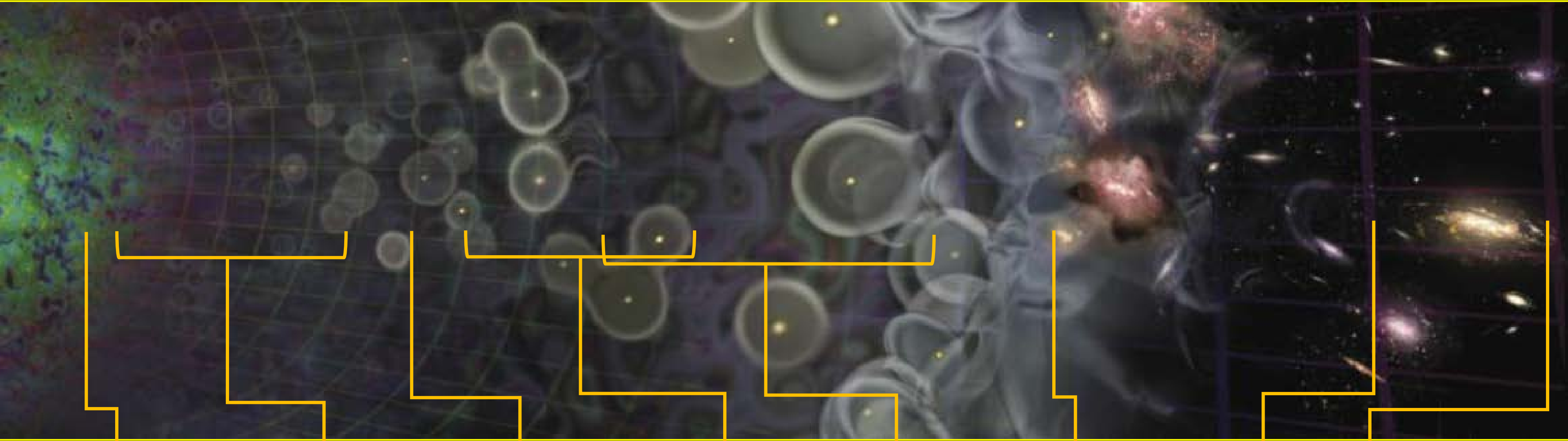


Blue: H I region

Red: H II region

Movie: Meng Zhou and Yi Mao

Epoch of Reionization



宇宙38万年

今天宇宙138亿年

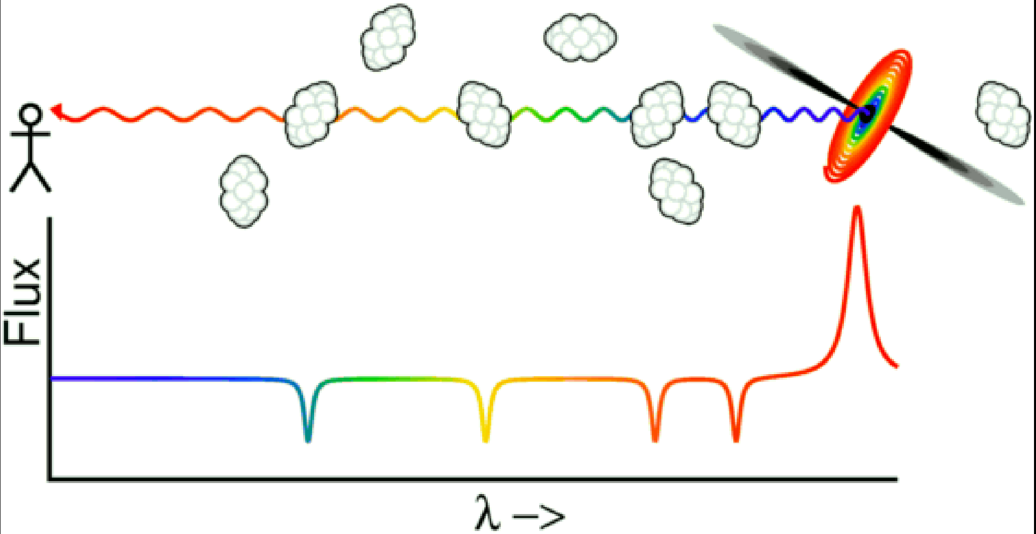
Cosmic Dawn and Epoch of Reionization

（cosmic age of ~ 0.1-1 Billion years）

宇宙黎明和再电离时代的*间接观测*

宇宙再电离时代的探针–1：Lyα 森林

来自遥远类星体的光谱在Lyα**波长的**吸收线强度反映了吸收体的中性氢密度。

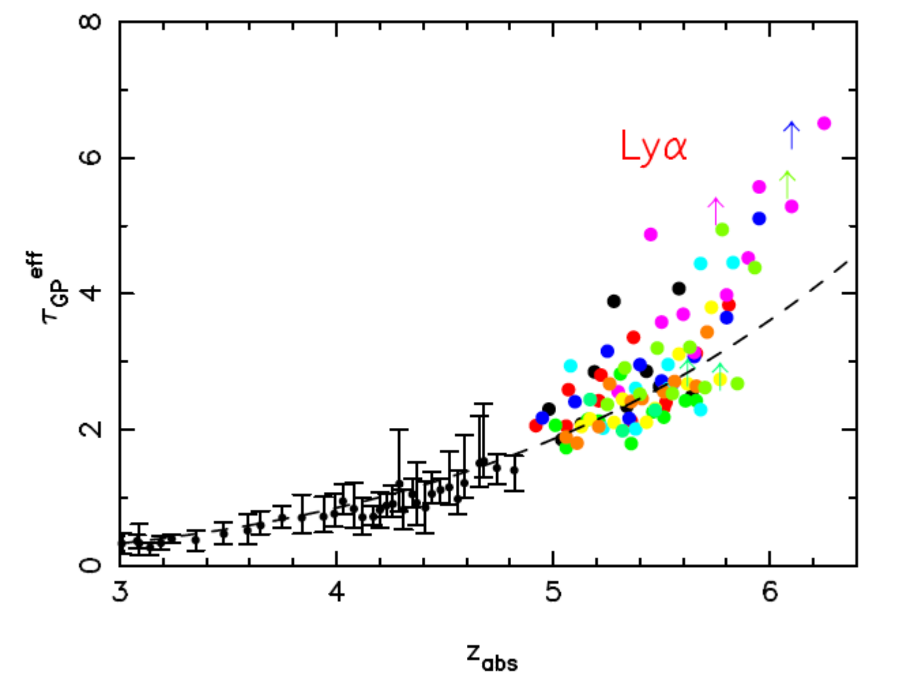


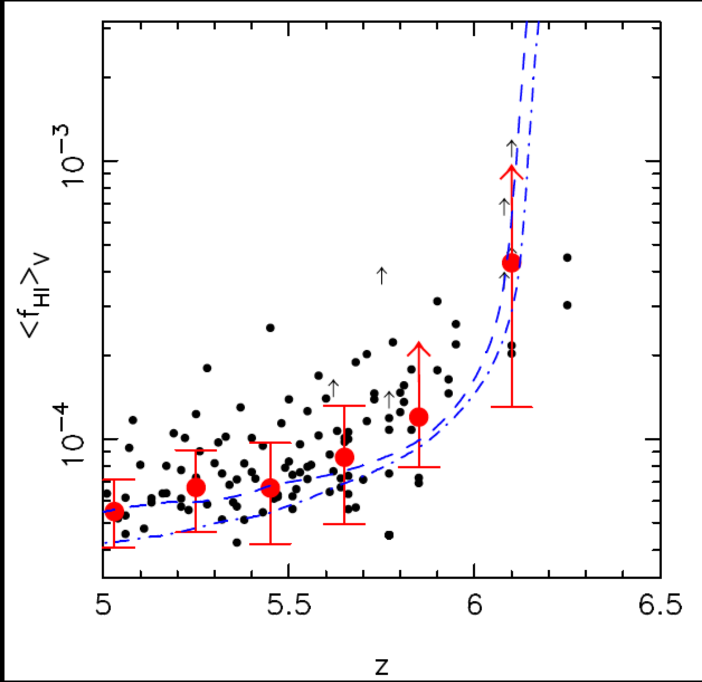
$$\lambda_{\rm obs} = \lambda_0(1+z)$$

$$\tau_{\rm GP}(z) = 1.8\times10^5 h^{-1}\Omega_m^{-1/2}\left(\frac{\Omega_b h^2}{0.02}\right)\left(\frac{1+z}{7}\right)^{3/2}\left(\frac{n_{\rm HI}}{\langle n_{\rm H}\rangle}\right)$$

宇宙再电离时代的探针-1：Lyα 森林

斯隆巡天测量类星体的Lyα吸收线通过星系际介质的光学深度，发现其在红移6附近迅速提高。这对应了中性氢在星系际介质中的比率迅速提高。这表明宇宙再电离时代在红移6时结束。



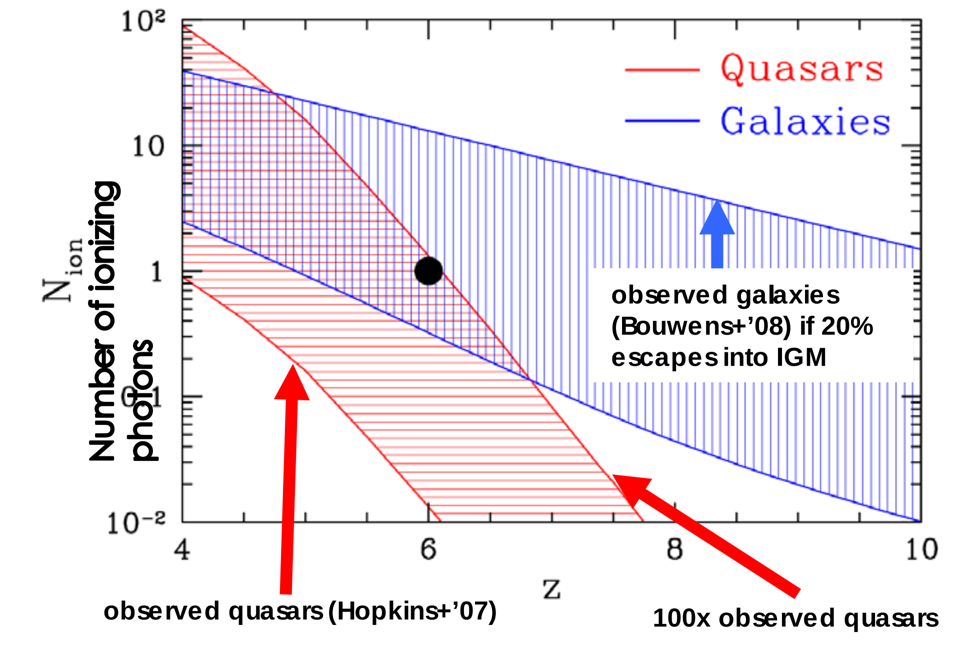


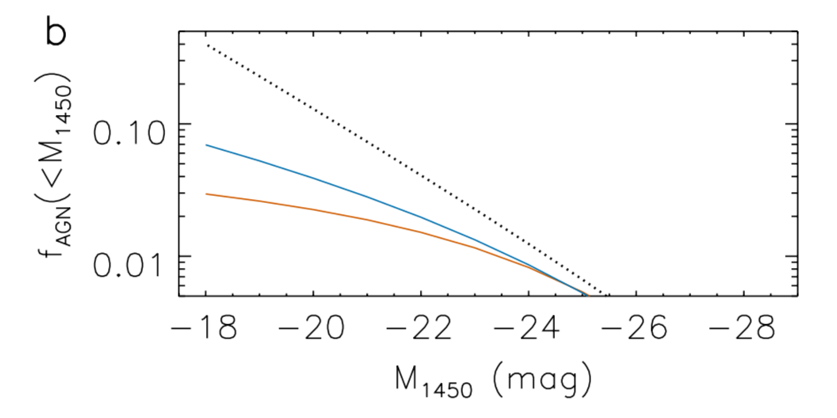
Fan et al (2006)

宇宙再电离时代的探针-1：Lyα 森林

宇宙再电离时代在红移6时结束意味着完成再电离的主要电离光源不可能是类星体，而只能是星系。这一理论上的推论最近也被高红移类星体观测所证实（Jiang et al. 2022）

Figure from Trac and Gnedin (2009)

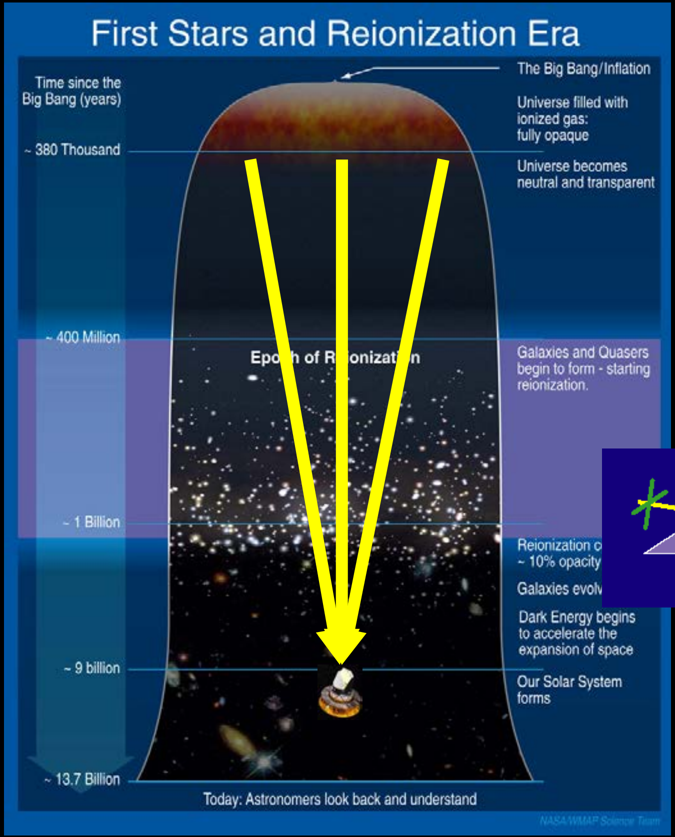




(b) Quasar contribution to reionization. The three curves represent the fractions (f_{AGN}) of the cumulative quasar emissivity to the total photon emissivity required to ionize the universe. They indicate a negligible quasar contribution. Figure from Jiang et al. (2022)

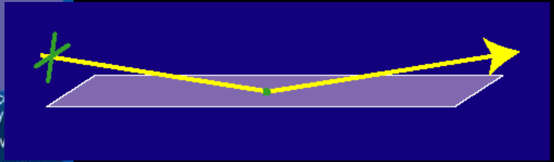
宇宙再电离时代的探针-2：宇宙微波背景辐射

宇宙微波背景辐射光子被处于再电离时代及之后的自由光子所散射(汤普逊散射)



Credit: NASA/WMAP Science Team

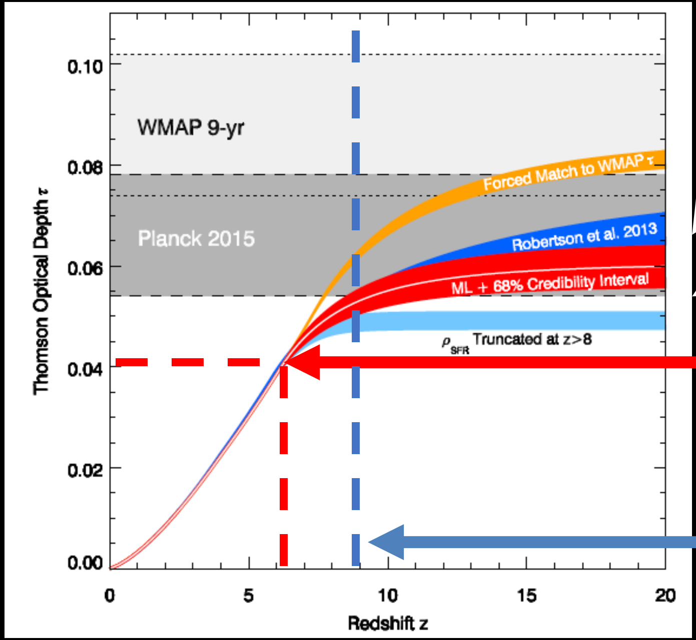
$$\tau_{\rm es} = \int \sigma_T \, n_e \, d\ell$$



宇宙再电离时代的探针-2：宇宙微波背景辐射

大约6%的宇宙微波背景辐射光子会被自由光子所散射，表明宇宙再电离是一个渐变的过程

Robertson et al. (2015)



$\tau = 0.066 \pm 0.013$

Planck 2015 (TT + lowP/LFI + lensing + BAO)

$\tau = 0.058 \pm 0.012$

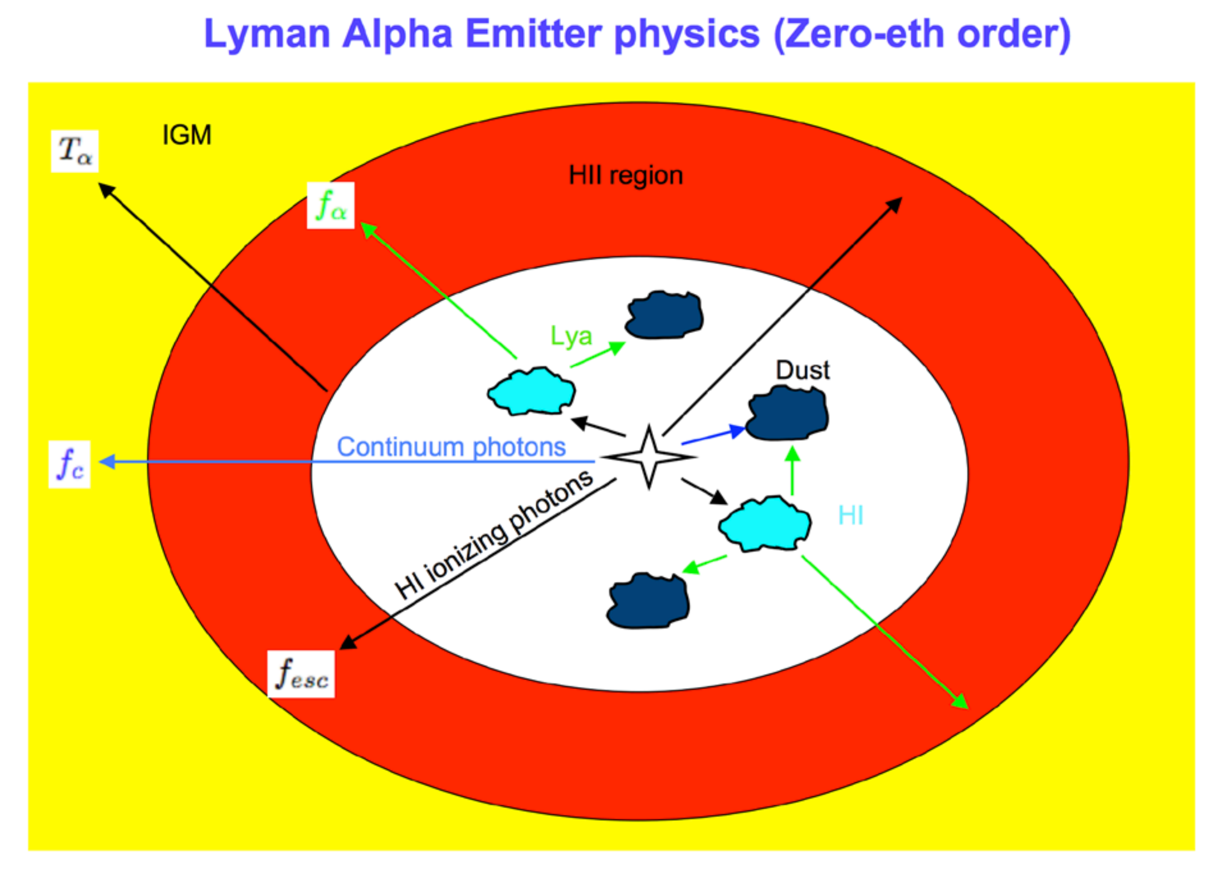
Planck 2016 (TT + lowP/HFI)

4% CMB photons could have been scattered by the fully ionized IGM at z<6

If reionization would be instantaneous, then $z_{\rm re} = 8.8^{+1.3}_{-1.2}$ (Planck 2015), or $z_{\rm re} = 8.5^{+1.0}_{-1.1}$ (Planck 2016).

宇宙再电离时代的探针-3：Lyα 发射体

第一代星系附近发出Lyα波长的光子，会被周围的中性氢气体散射



宇宙再电离时代的探针-3：Lyα 发射体

第一代星系附近发出Lyα波长的光子，会被周围的中性氢气体散射。 Lyα光度函数在高光度端的突起表明，**宇宙再电离是成泡状结构、非均匀的过程**。

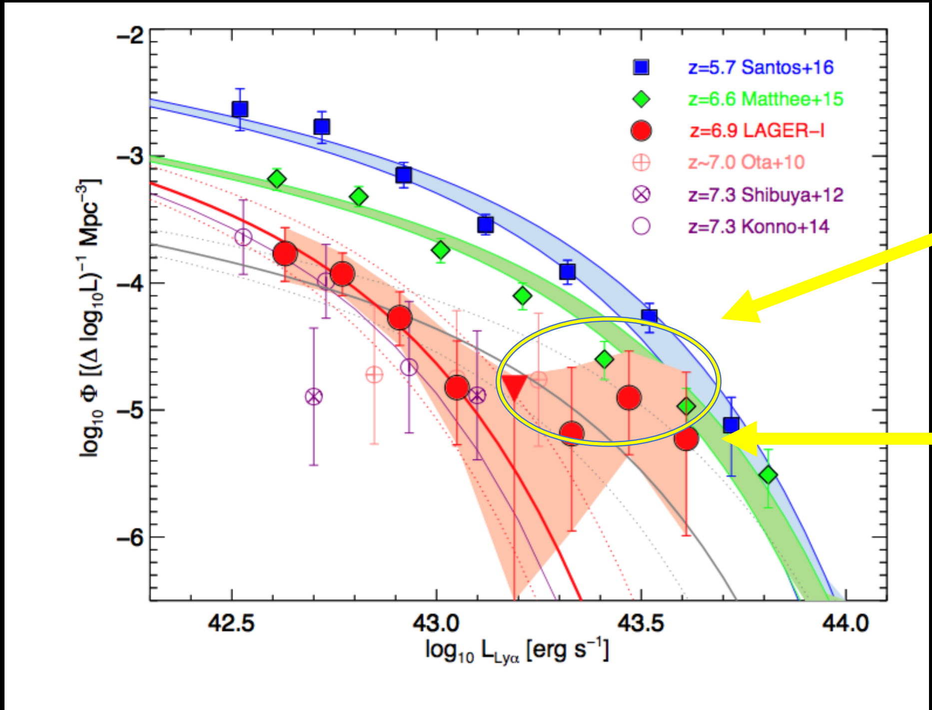


Figure 2. The Lyα LFs of LAE surveys at $z \sim [5.7, 6.6, 7, 7.3]$.

LAGER collaboration (Zheng et al. 2017)

Evolution of LF indicates $x_{HI} \sim 0.4 - 0.6$ at $z = 6.9$ (however, model dependent).

Bump at bright end indicates large HII bubbles, where Hubble flow, galactic inflow/outflow can bring Lyα photons out of resonance.

宇宙再电离时代的观测证据

- **已有探针：**

☑ **Lyα 森林光深测量：**再电离结束于红移6

☑ **宇宙微波背景辐射：**再电离渐变

☑ **Lyα 发射体：**再电离非均匀

- **未来探针：**

☐ **中性氢21厘米谱线强度映射**

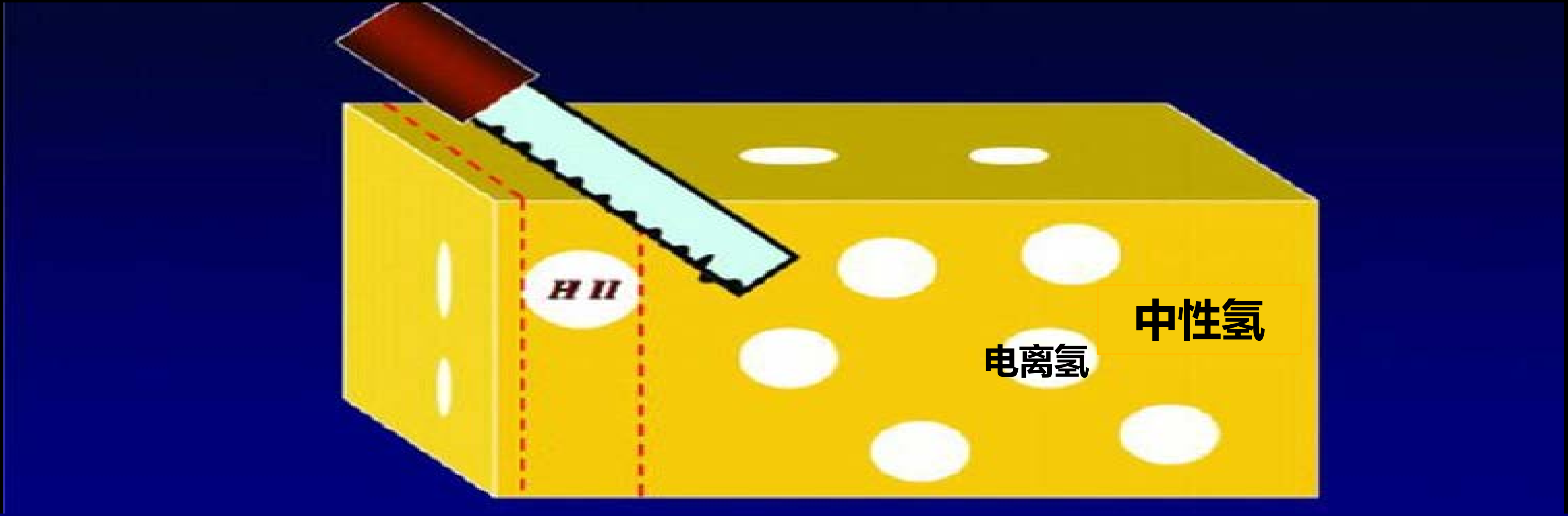
☐ **分子谱线强度映射，及其与21厘米交叉关联**

☐ **中性氢21厘米森林**

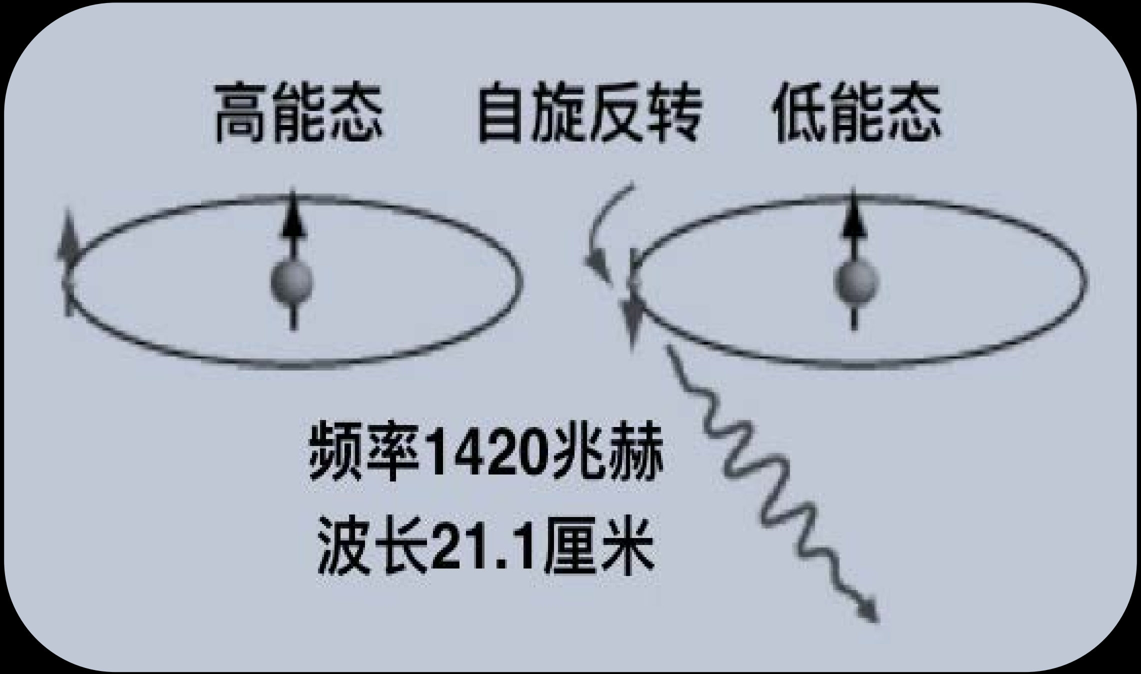
☐ **低红移Lyα 森林功率谱**

宇宙黎明和再电离时代的*直接*观测

中性氢



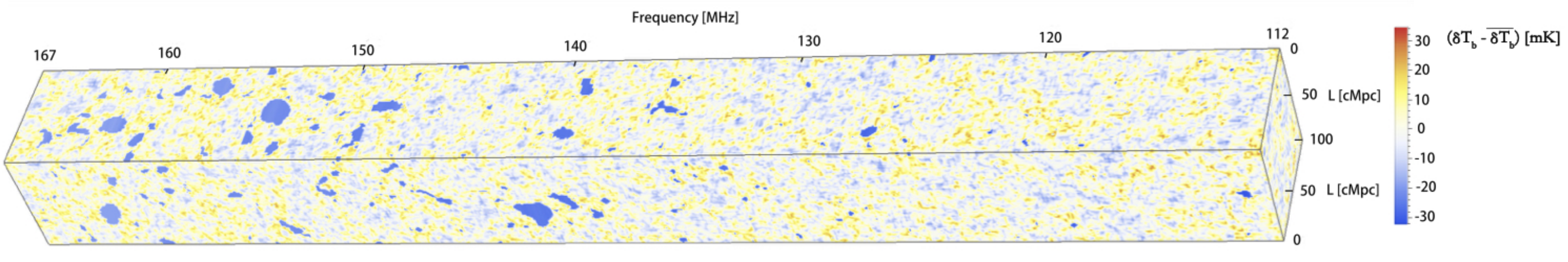
中性氢21厘米谱线

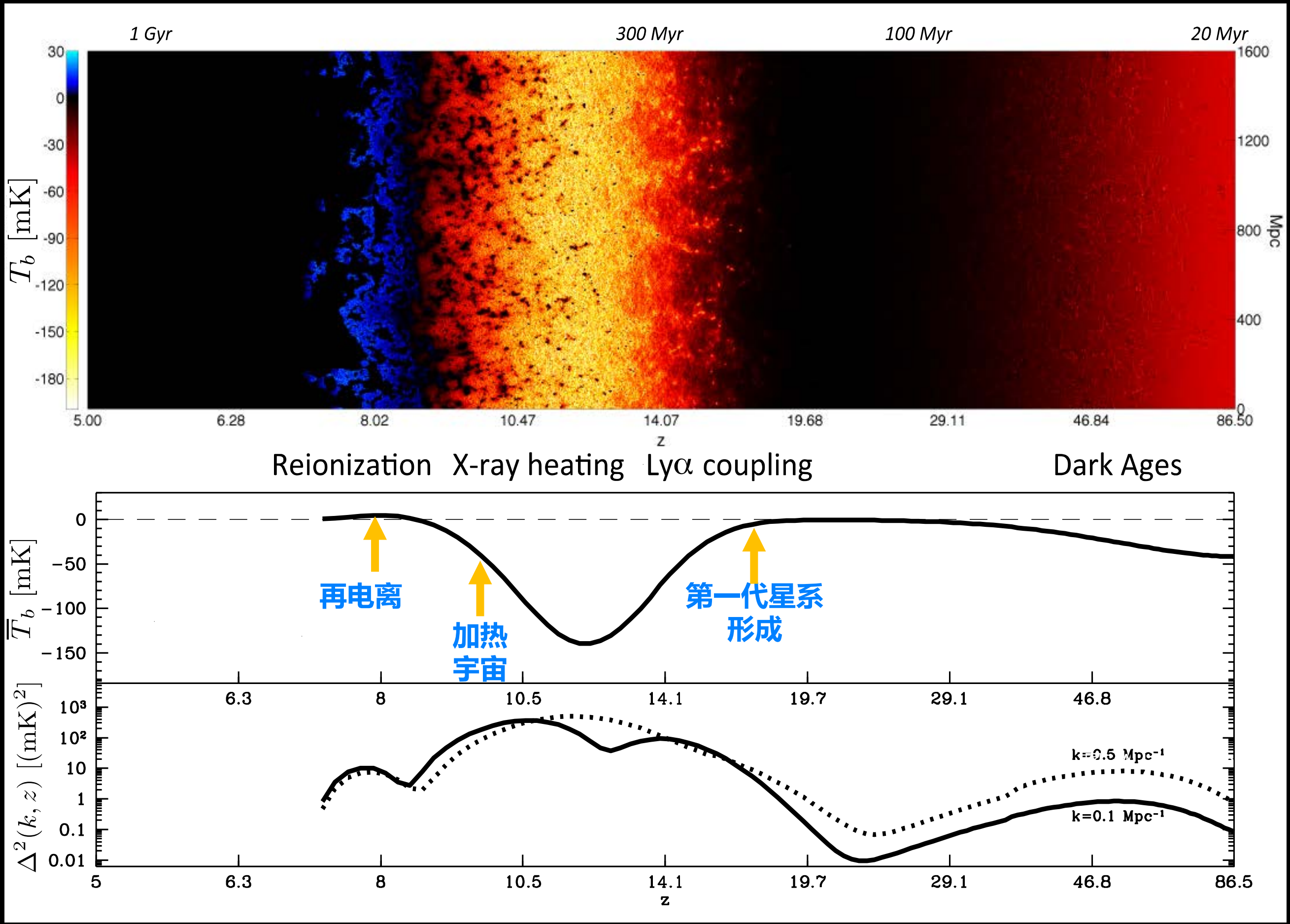


爱因斯坦受激辐射系数和自发跃迁系数 $\left(A_{10} = 2.85 \times 10^{-15}\ \mathrm{s}^{-1}\right)$

(对应于32Myrs才会自发跃迁一次！) So 21 cm line is optically thin!

通过21厘米谱线看到的宇宙（模拟图）





图像信号

平均信号

统计起伏信号

单个天线测量平均信号

EDGES
高频天线
美国/澳洲

EDGES
低频天线
美国/澳洲

PRIZM
低频天线
南非

PRIZM
高频天线
南非

LEDA
美国

SARAS-3
印度


21CMA 中国·新疆

LOFAR 荷兰/欧洲


MWA 澳大利亚

PAPER 美国/南非


GMRT 印度

第一代射电干涉阵列（试图）测量信号的统计起伏

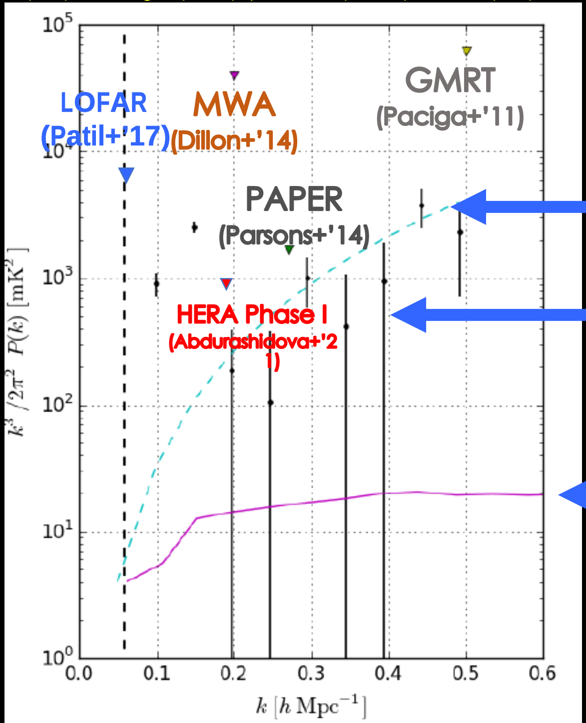
Direct Constraints on Reionization: 21-cm

PAPER, MWA, LOFAR and GMRT measure the **power spectrum** of 21-cm brightness temperature fluctuations.

PAPER reported a 2σ upper limit on 21-cm power spectrum of $(22.4\ \mathrm{mK})^2$ at $k=0.15 - 0.5\ h\ \mathrm{Mpc}^{-1}$ at $z=8.4$

$$\langle \delta T_b(\mathbf{k})\, \delta T_b^{\star}(\mathbf{k}') \rangle = (2\pi)^3 P(\mathbf{k}) \delta^{(3)}(\mathbf{k} - \mathbf{k}')$$

Ali et al. (2015)



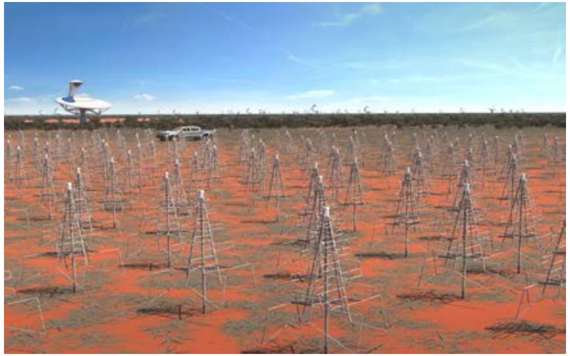
Predicted noise 2σ level

PAPER 2015 observation (black dots with 2σ error bars)

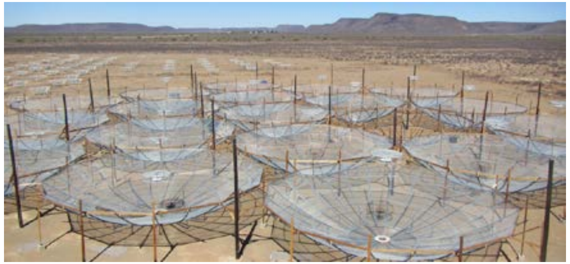
Predicted signal from a theoretical model at 50% ionization

Also see Yinzhe Ma's talk

第二代射电干涉阵列将测量信号的统计起伏和图像



SKA（平方公里阵列射电望远镜）
一期工程～157亿欧元
14国国际合作
（澳/加/中/法/德/印/意/荷/新/南非/西/瑞典/瑞士/英）



HERA
美国
（+加/英/南非/意）

世界巨眼：平方公里阵列射电望远镜（SKA）

特点

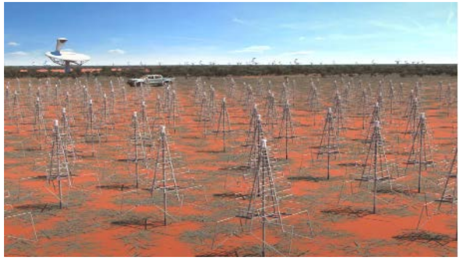
- SKA将是世界最大综合孔径射电望远镜。
- 革命性特点：高分辨率、高灵敏度、大视场、宽频段、大数据

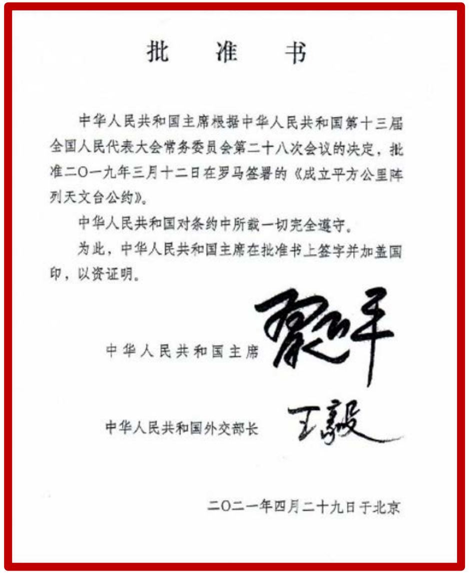
重要性

- 宇宙再电离时代的21厘米观测是SKA最重要的科学目标之一。
- 这个领域将迎来宇宙学的重大突破！

国际合作

- 2019年，中、英、澳、南非、意、荷、葡签署SKA天文台公约。
- 2021年，全国人大批准SKA天文台公约。
- 2028年，SKA预计完成建设。





批　准　书

中华人民共和国主席根据中华人民共和国第十三届全国人民代表大会常务委员会第二十八次会议的决定，批准二〇一九年三月十二日在罗马签署的《成立平方公里阵列天文台公约》。

中华人民共和国对条约中所载一切完全遵守。

为此，中华人民共和国主席在批准书上签字并加盖国印，以资证明。

中华人民共和国主席

中华人民共和国外交部长

二〇二一年四月二十九日于北京

习近平主席批准《成立平方公里阵列天文台公约》

Data Analysis in 21cm observations

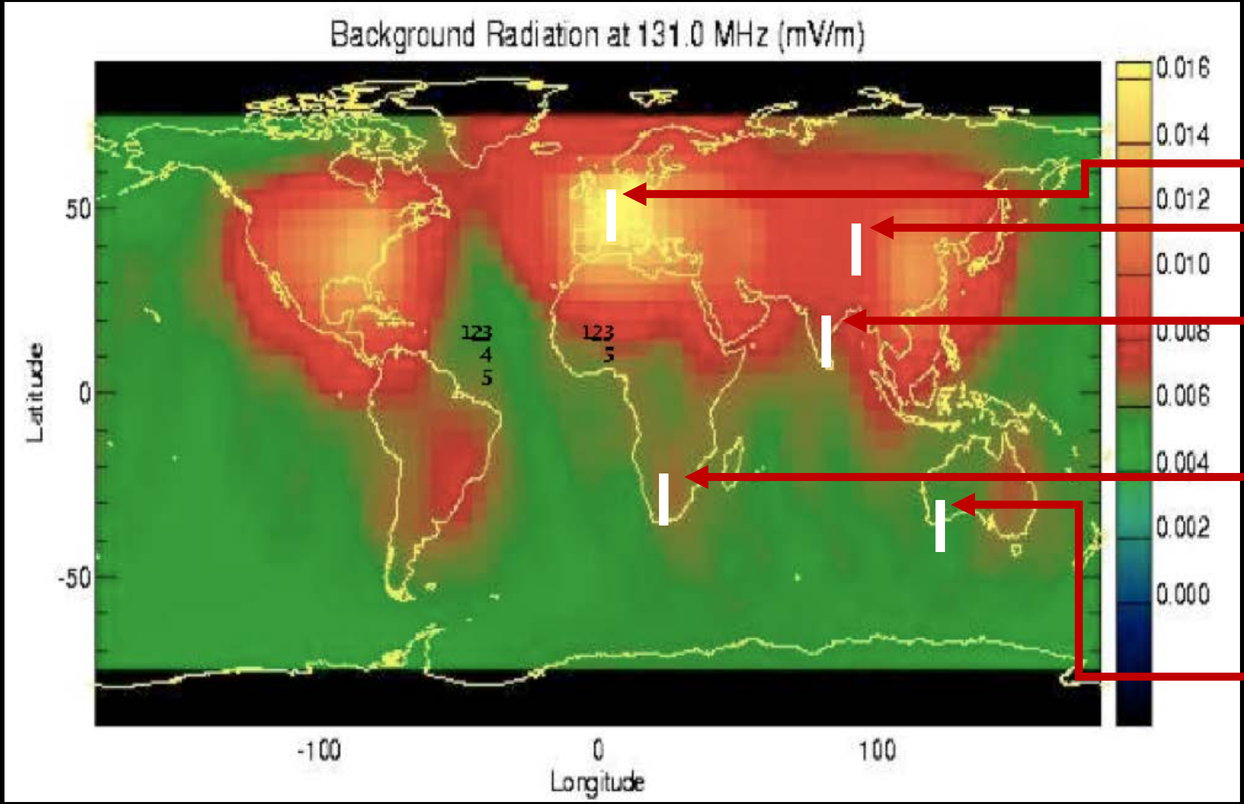
- Calibration (see Xin Wang's talk)
- RFI flagging (in visibility measurement)
- Image making (see Le Zhang's talk)
- Foreground subtraction

(also see the talks of John Podczerwinksi, Cunnington Steve, Shulei Ni, Feng Shi)

- Scientific interpretation (parameter inference)

(also see Hayato Shimabukuro's talk)

Radio Frequency Inference (RFI)



LOFAR

21CMA

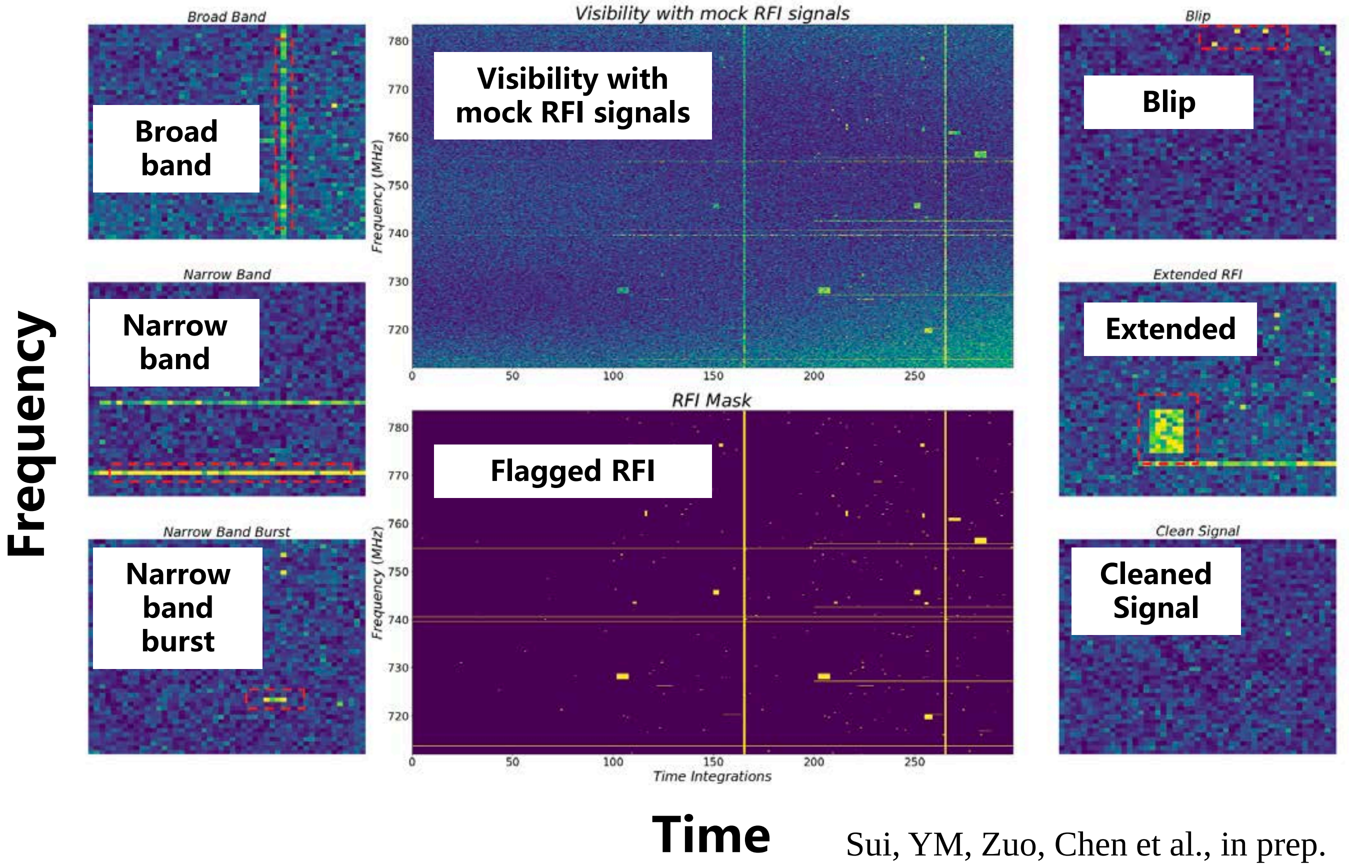
GMRT
SARAS-3

PRIZM
PAPER
HERA
SKA-mid

EDGES
MWA
SKA-low

RFI Flagging with U-Net

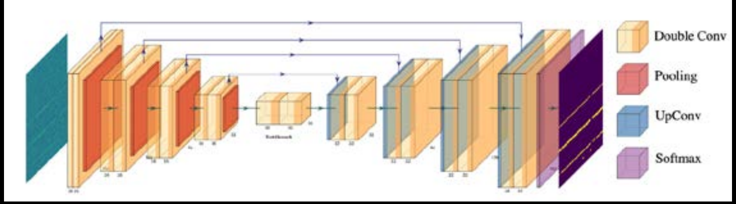
Characteristic RFI Classification

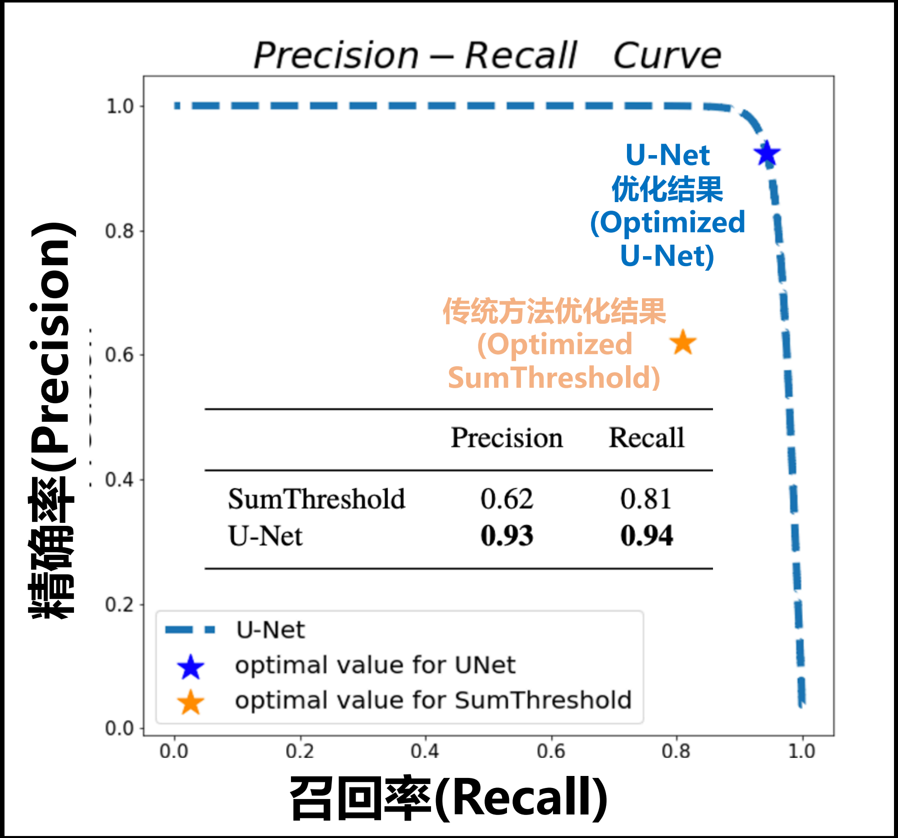


Sui, YM, Zuo, Chen et al., in prep.

Ce Sui

RFI Flagging with U-Net





	Precision	Recall
SumThreshold	0.62	0.81
U-Net	**0.93**	**0.94**

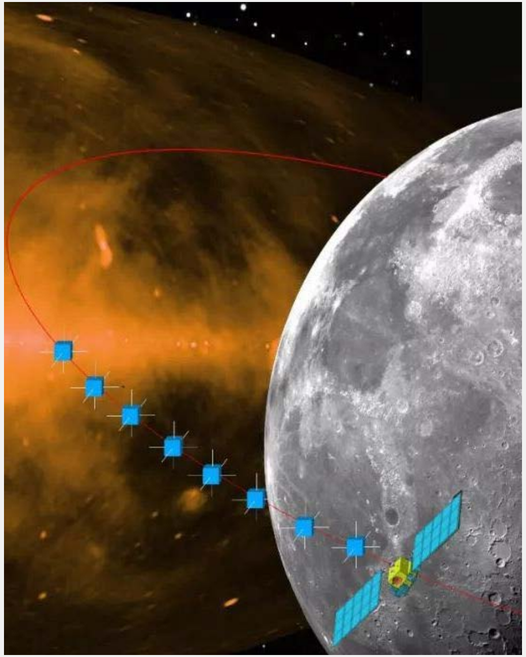
Ce Sui

See also Sun, Deng, Wang, et al., MNRAS, 2022

Sui, YM, Zuo, Chen et al., in prep.

远离地面无线电干扰——到月球背面去

DAPPAR

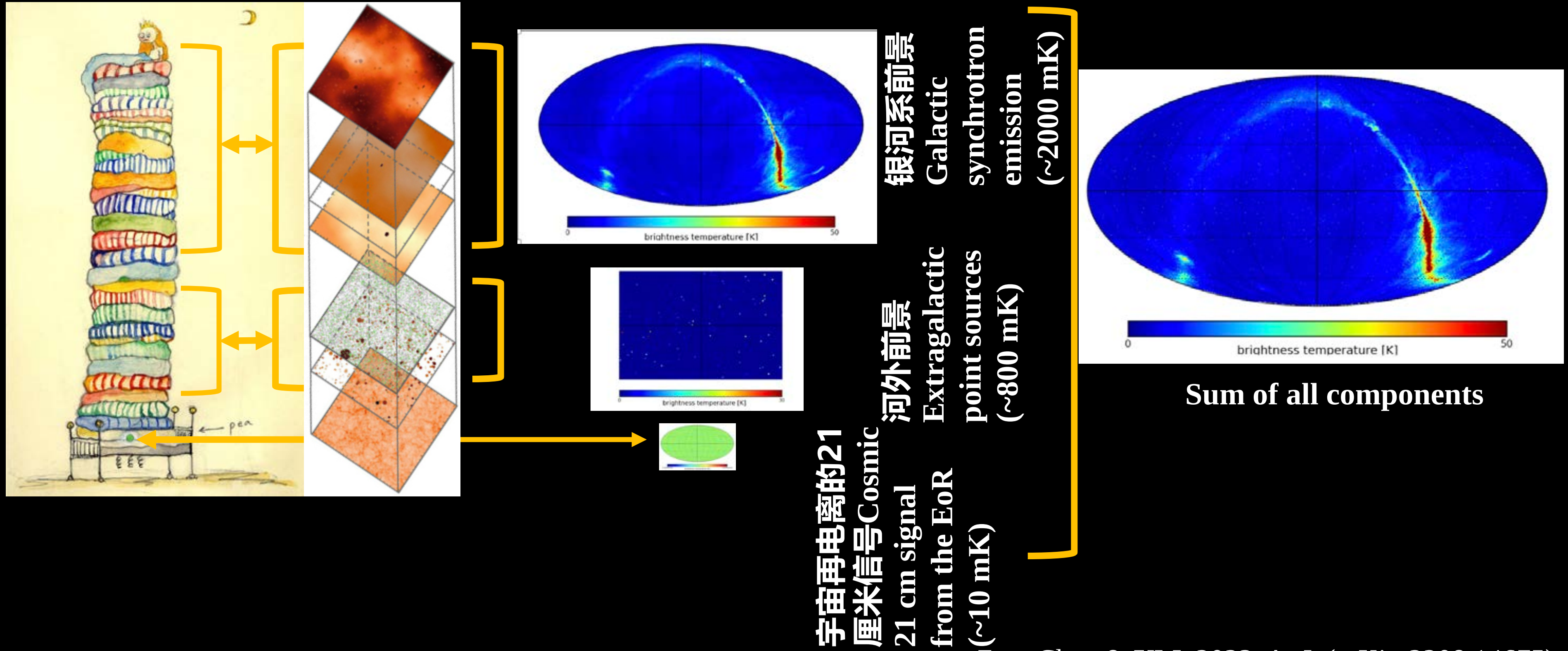


鸿蒙计划

(超长波天文观测阵列，DSL)

(See Xuelei Chen's talk)

Foreground Contamination in 21 cm Observations



Zuo, Chen & **YM**, 2023, ApJ (arXiv:2208.14675)

银河系和河外射电源前景污染——去除前景办法研究

Shifan Zuo

- 传统使用经典的主成分分析方法（PCA），效果不佳。
- 在PCA基础上，我们提出半盲的”奇异矢量投影” （SVP）方法，改进效果明显。

$$D\,(\text{image}) = F\,(\text{foreground}) + N\,(\text{signal} + \text{noise})$$

PCA/SVD

$$\boldsymbol{D} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}$$

Singular Vector Projection (SVP)

$$\boldsymbol{F} = \boldsymbol{U}_f\boldsymbol{S}_f\boldsymbol{V}_f^{\mathrm{T}} \qquad \text{(foreground modelling)}$$

$$\boldsymbol{N}_{\mathrm{L}} = \boldsymbol{D} - \boldsymbol{U}_f\boldsymbol{U}_f^{\mathrm{T}}\boldsymbol{D}, \qquad \text{(left)}$$

$$\boldsymbol{N}_{\mathrm{R}} = \boldsymbol{D} - \boldsymbol{D}\boldsymbol{V}_f\boldsymbol{V}_f^{\mathrm{T}}, \qquad \text{(right)}$$

$$\boldsymbol{N}_{\mathrm{B}} = \boldsymbol{D} - \boldsymbol{U}_f\boldsymbol{U}_f^{\mathrm{T}}\boldsymbol{D}\boldsymbol{V}_f\boldsymbol{V}_f^{\mathrm{T}}, \qquad \text{(both)}$$

$$\boldsymbol{N}_{\mathrm{D}} = \boldsymbol{D} - \boldsymbol{U}_f\left(\boldsymbol{U}_f^{\mathrm{T}}\boldsymbol{D}\boldsymbol{V}_f\right)_{\mathrm{diag}}\boldsymbol{V}_f^{\mathrm{T}}. \qquad \text{(diagonal)}$$

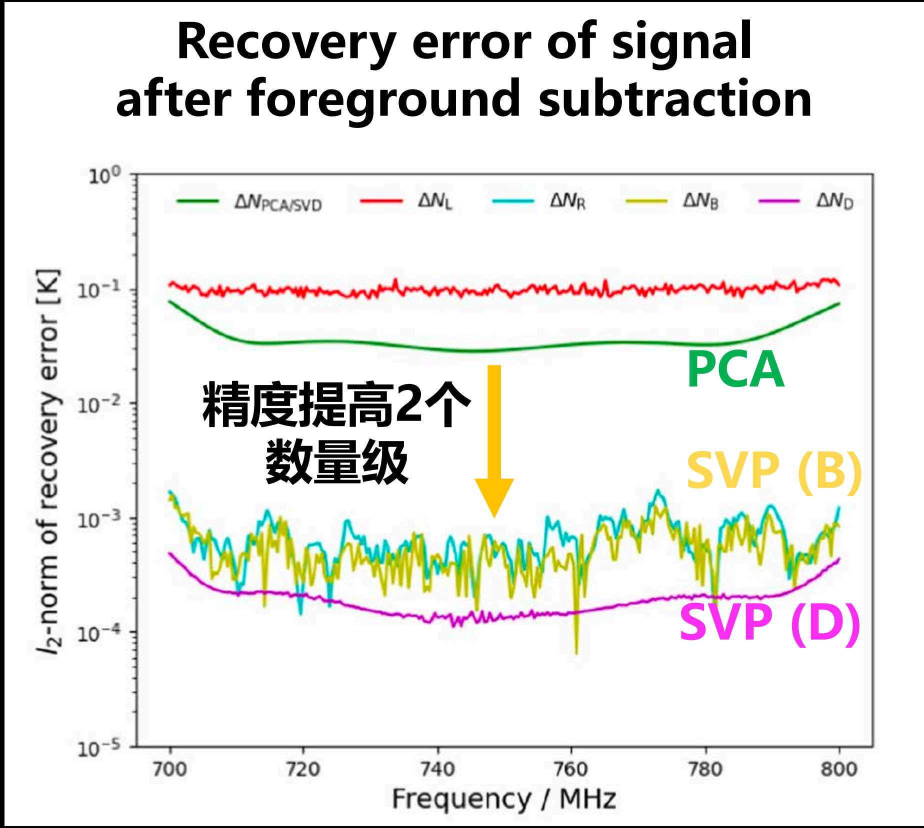
SVP estimators are semiblind, in the sense that **they are independent of absolute strength of foreground**, but depend on amplitude change in frequency direction (left vectors) and/or sky plane (right vectors).

Zuo, Chen & **YM**, 2023, ApJ (arXiv:2208.14675)

Semiblind Foreground Subtraction: SVP

Left singular vectors contain the frequency info.

Right singular vectors contain the pixel-wise info.



Shifan Zuo

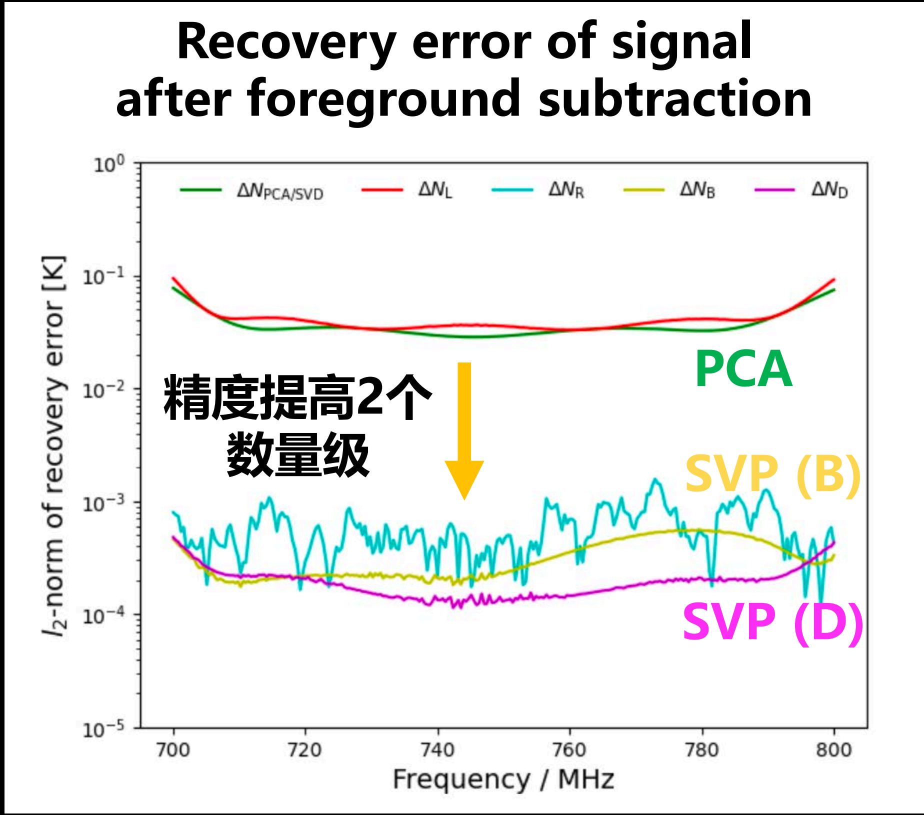
Zuo, Chen & **YM**, , ApJ (arXiv:2208.14675)

SVP with *Incomplete* Singular Vectors

Only the largest five left and/or right singular vectors of the foregrounds are exploited here.

Left singular vectors contain the frequency info.

Right singular vectors contain the pixel-wise info.

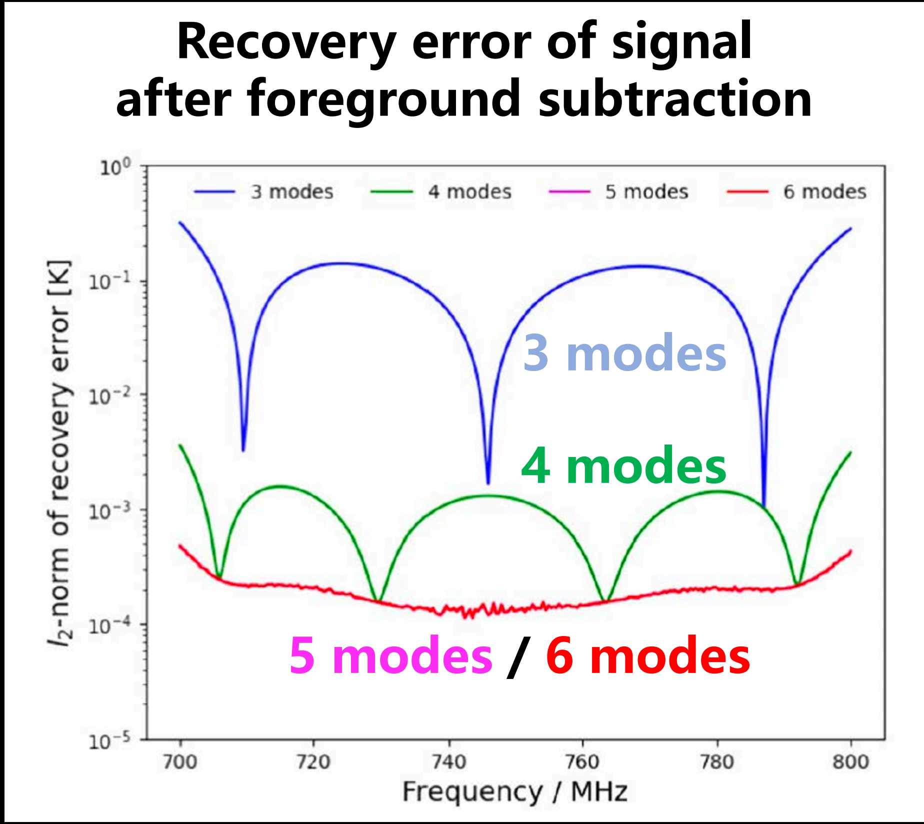


Shifan Zuo

Zuo, Chen & **YM**, 2023, ApJ (arXiv:2208.14675)

SVP with *Incomplete* Singular Vectors

Convergence test:
Only the largest 3, 4, 5, or 6 left and right singular vectors of the foregrounds are exploited here (with SVP-D estimator).

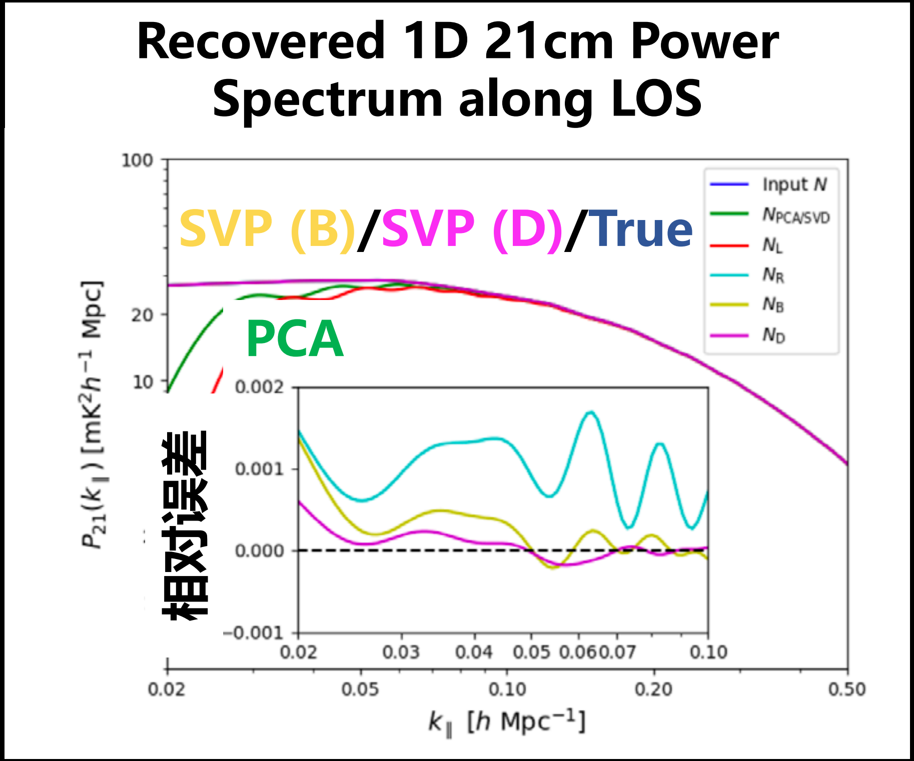


Shifan Zuo

Zuo, Chen & **YM**, 2023, ApJ (arXiv:2208.14675)

Semi-blind Foreground Subtraction: SVP

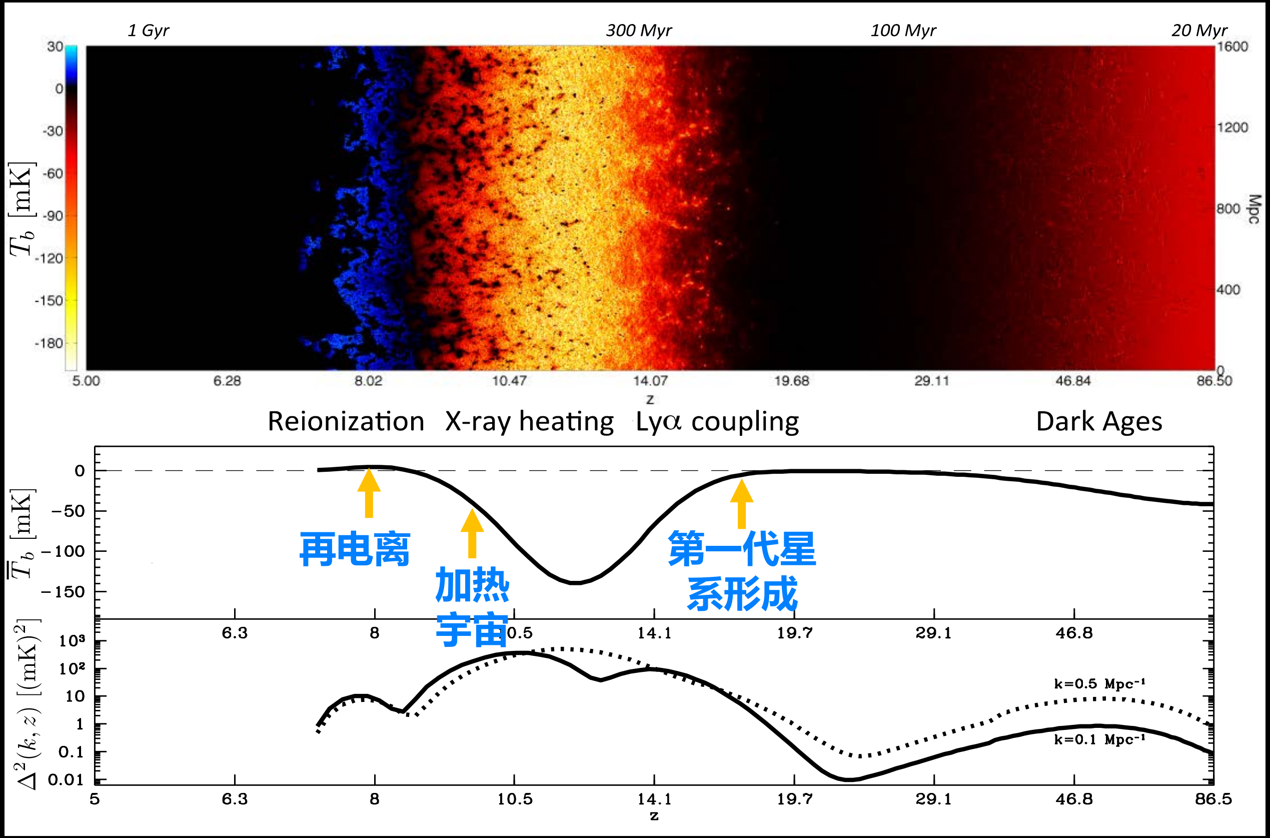
Only the largest five left and/or right singular vectors of the foregrounds are exploited here.



Shifan Zuo

Zuo, Chen & **YM**, 2023, ApJ (arXiv:2208.14675)

Extract astrophysical information from cleaned data



图像信号

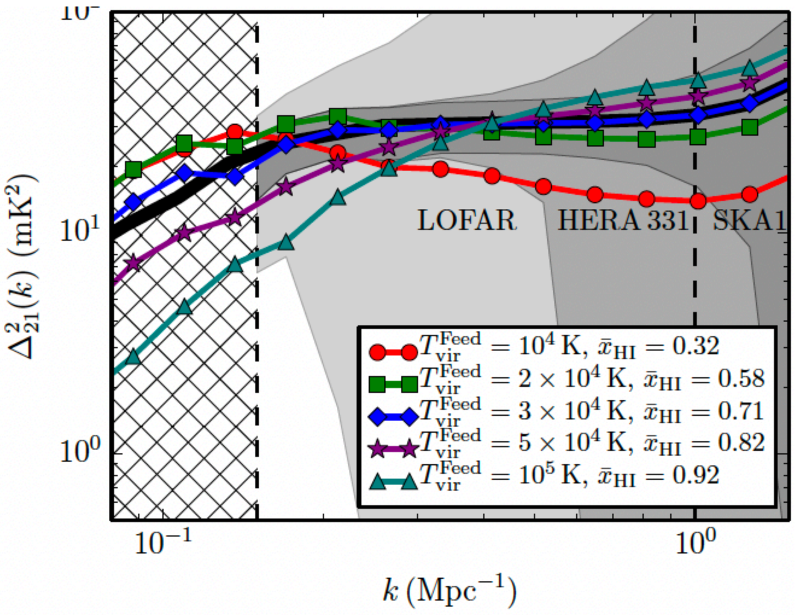
平均信号

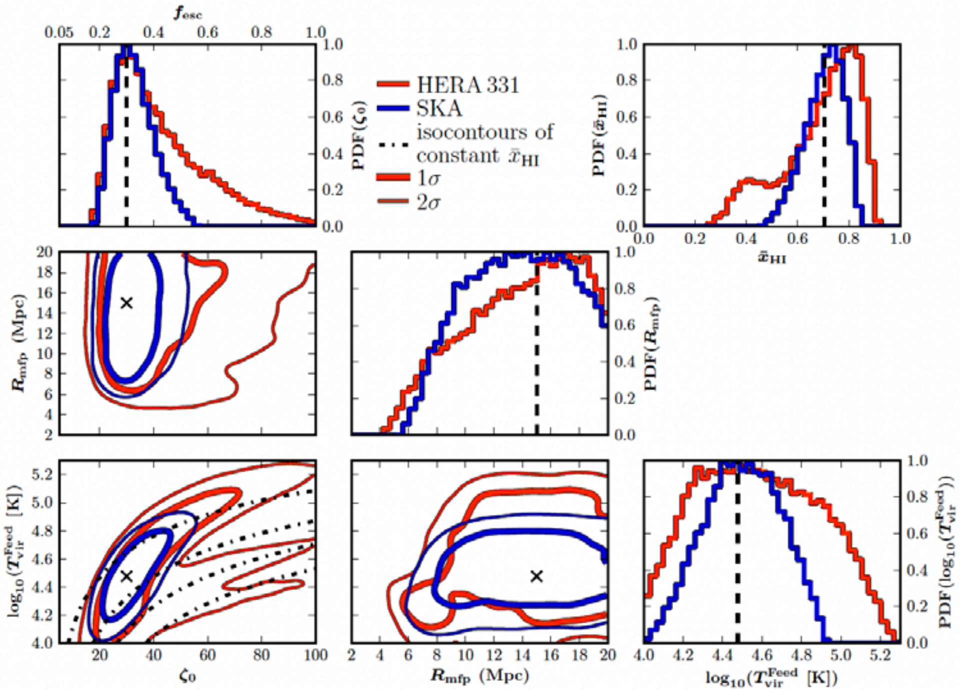
统计起伏信号（功率谱）

图：A. Liu & R. Shaw

Parameter Estimation using 21 cm Power Spectrum

21cm Power Spectrum





Reionization Parameters

ζ the ionizing efficiency

T_{vir} the minimum virial temperature of halos that host ionizing sources

Bayesian inference of reionization model parameters with conventional MCMC method (`21CMMC` code)

Greig & Mesinger, 2015, MNRAS

Parameter Estimation using 21 cm Power Spectrum

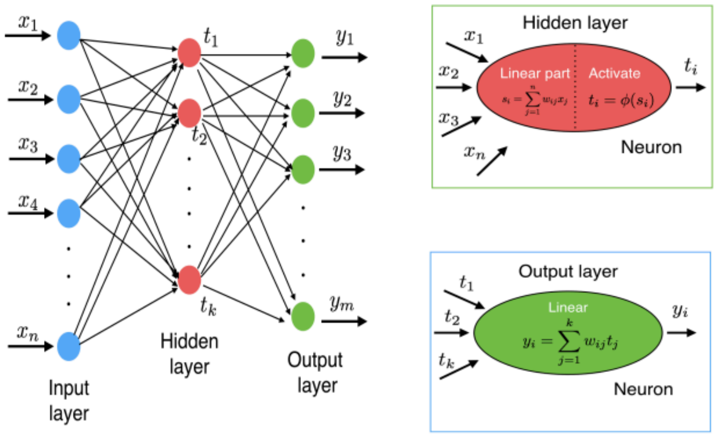
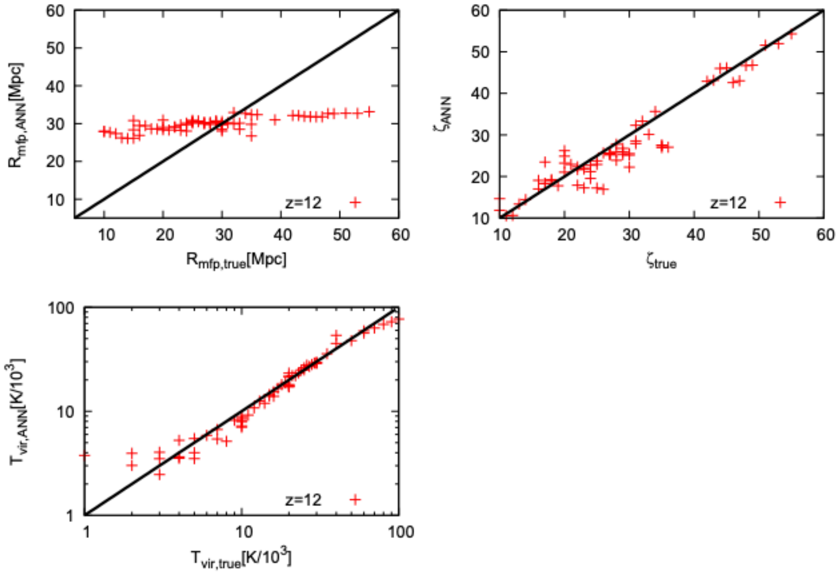


Figure 1. Typical architecture of an artificial neural network. The architecture of the ANN consists of an input layer, a hidden and an output layer of neurons. Each neuron connects the neurons in the next layer.



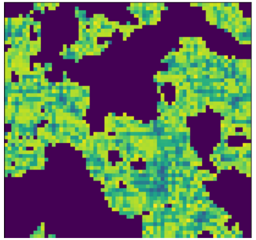
Estimation of reionization model parameters with artificial neural networks
<u>(note: point estimate, not posterior inference)</u>

Figure 4. The EoR model parameter values computed by the ANN from the PS against the values used in the simulation at z=12. Note that the result for the Virial temperature is plotted

Shimabukuro & Semelin, 2017, MNRAS

Statistical Inference in Cosmology

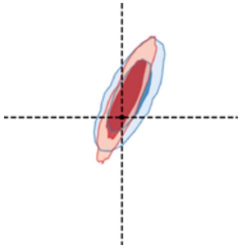
Observation



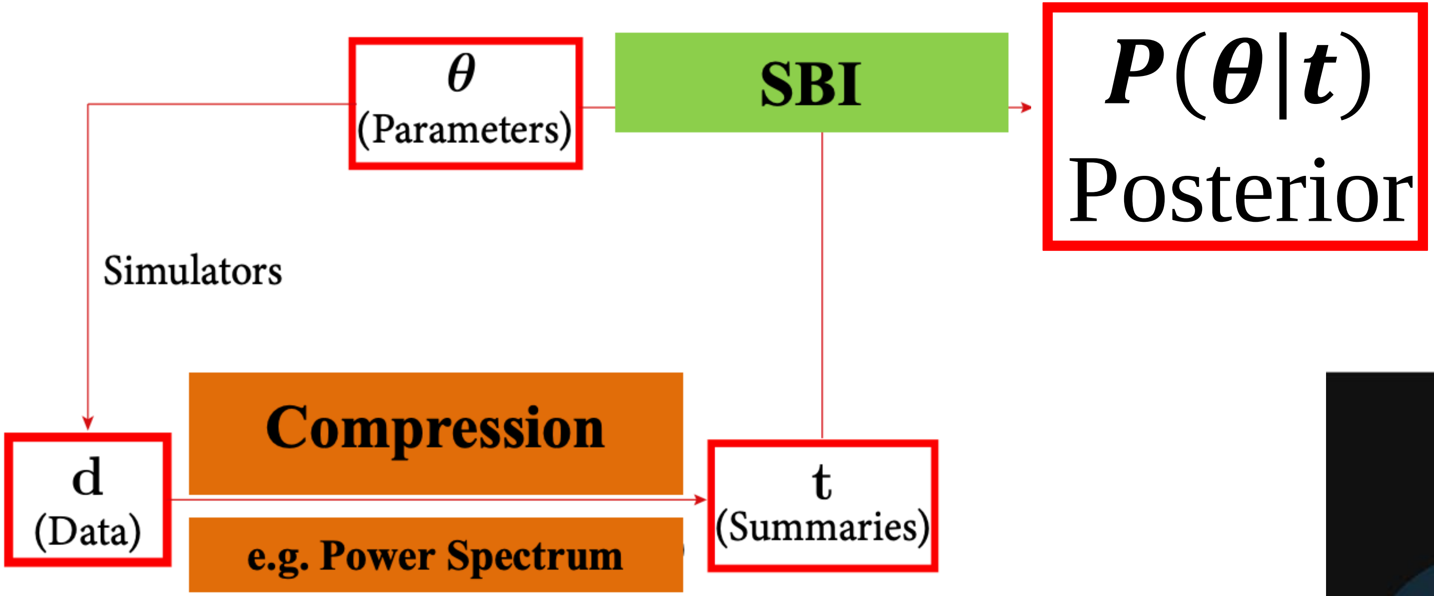
Summary Statistics ⇒ Likelihood ⇒ Sampling

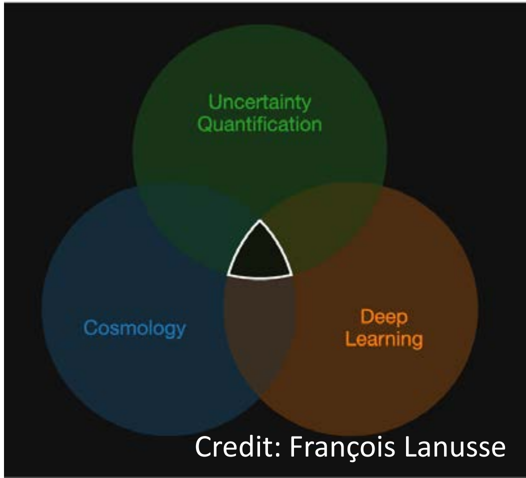
Statistical Inference

Parameter Posterior

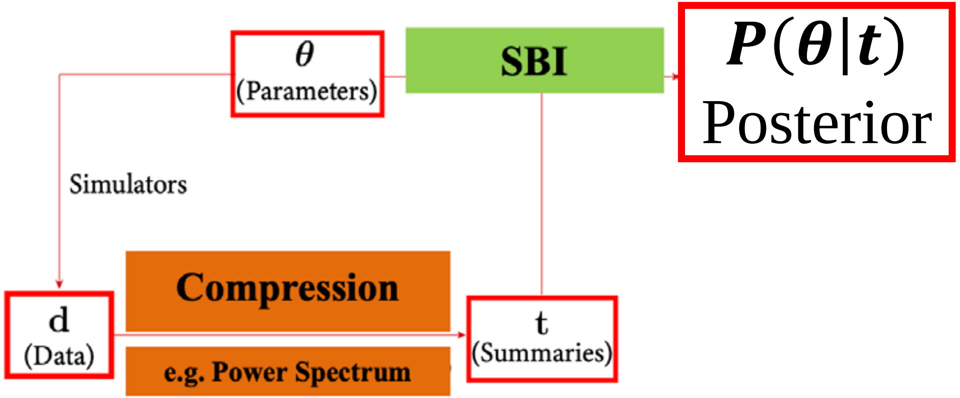


Simulation-Based Inference (SBI)





Simulation-Based Inference



1. Generative Way (density estimation likelihood-free inference, DELFI):

Zhao, **YM**, et al 2022a,2022b

$$\{\theta, t\} \xrightarrow{\text{Density Estimator}} P(t|\theta) \xrightarrow{\text{Prior}} \boldsymbol{P(\theta|t) \propto P(t|\theta)P(\theta)}$$

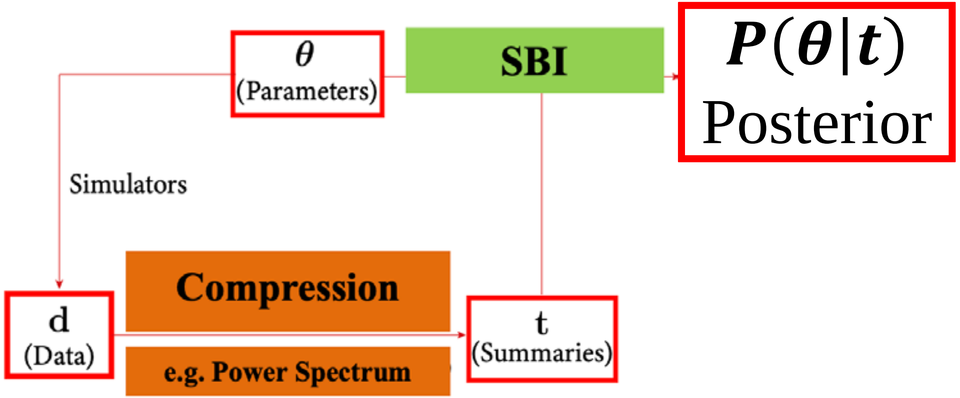
Bayes' Theorem

2. Discriminative Way (Neural Ration Estimation, NRE):

Ce, **YM**, et al
In prep

$$\{\theta, t\} \xrightarrow{\text{Ratio estimation}} r(\theta, t) = \frac{P(t|\theta)}{P(t)} \xrightarrow{\text{Prior}} \boldsymbol{P(\theta|t) = r(\theta, t)P(\theta)}$$

Simulation-Based Inference



1. Generative Way (density estimation likelihood-free inference, DELFI):

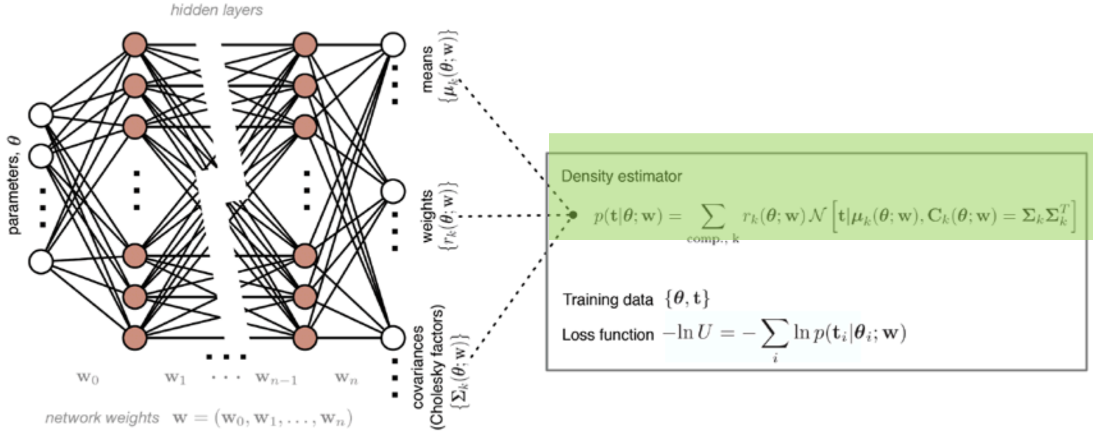
Zhao, **YM**, et al
2022a,2022b

An example of Neural Density Estimators (NDEs)

$$\{\theta, t\} \xrightarrow{\text{Density Estimator}} P(t|\theta) \xrightarrow{\text{Prior}} \boldsymbol{P(\theta|t) \propto P(t|\theta)P(\theta)}$$

Mixture Density Network (MDN)

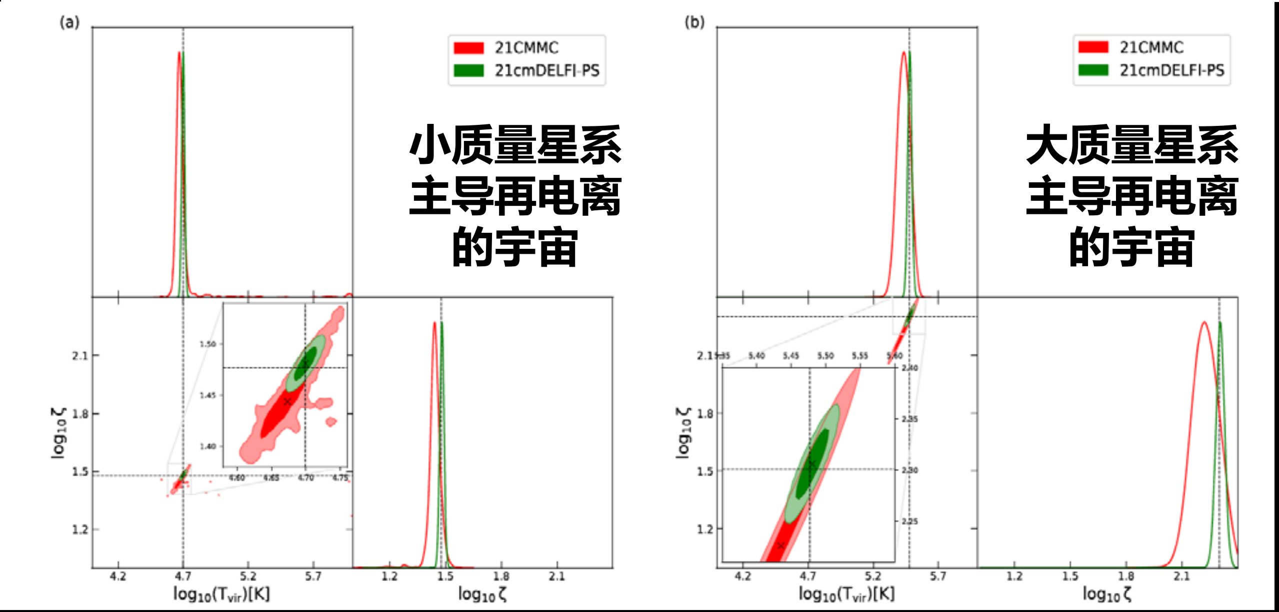
Bayes' Theorem



从21厘米功率谱测量出发限制宇宙再电离理论模型

- 传统使用经典的马尔可夫链蒙特卡罗算法（MCMC），需要做特定假设。
- 发展了基于深度学习的贝叶斯统计推断的新方法，开发了新软件21cmDELFI-PS

对宇宙再电离模型参数进行贝叶斯统计推断的示例

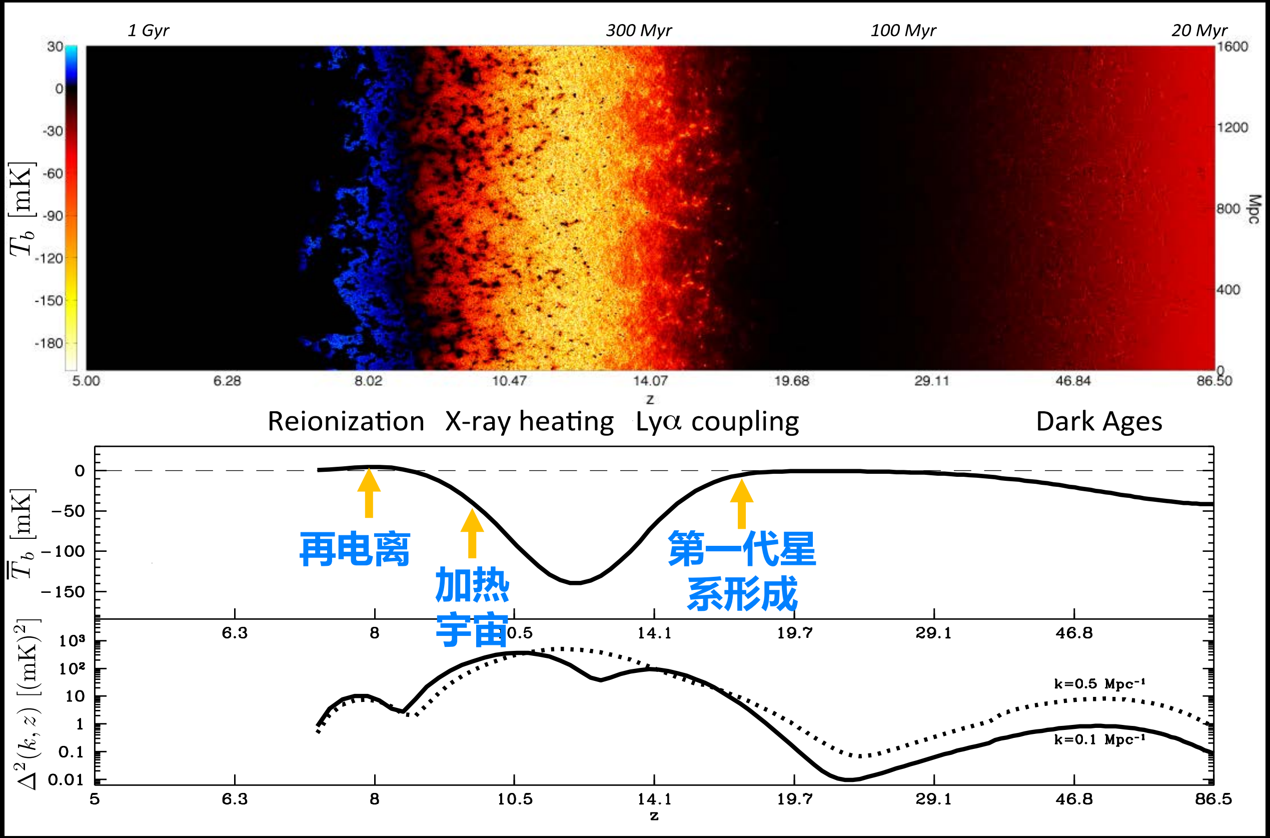


MCMC得到的置信区间

21cmDELFI-PS得到的置信区间

Zhao, **YM** & Wandelt, 2022, ApJ (arXiv:2203.15734)

Extract astrophysical information from cleaned data



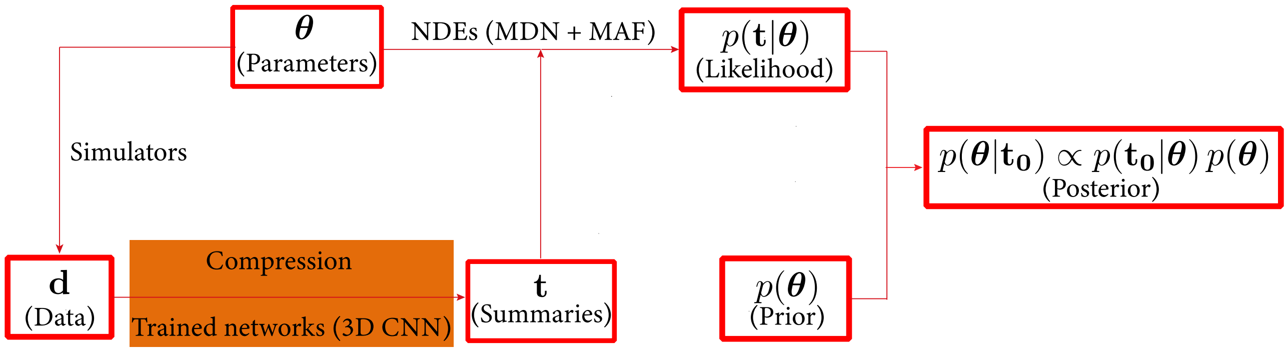
图像信号

平均信号

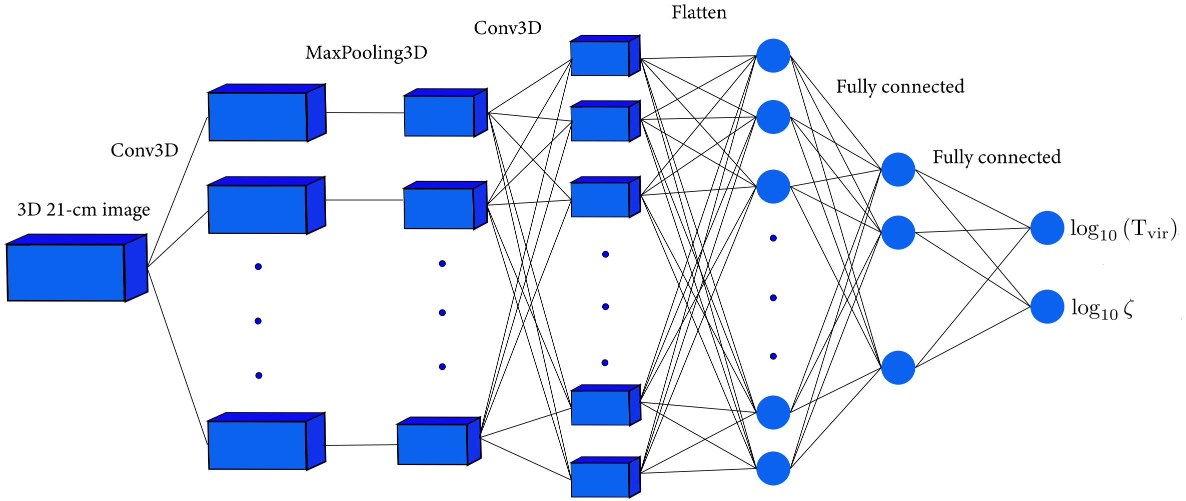
统计起伏信号（功率谱）

图：A. Liu & R. Shaw

Likelihood-free Bayesian inference



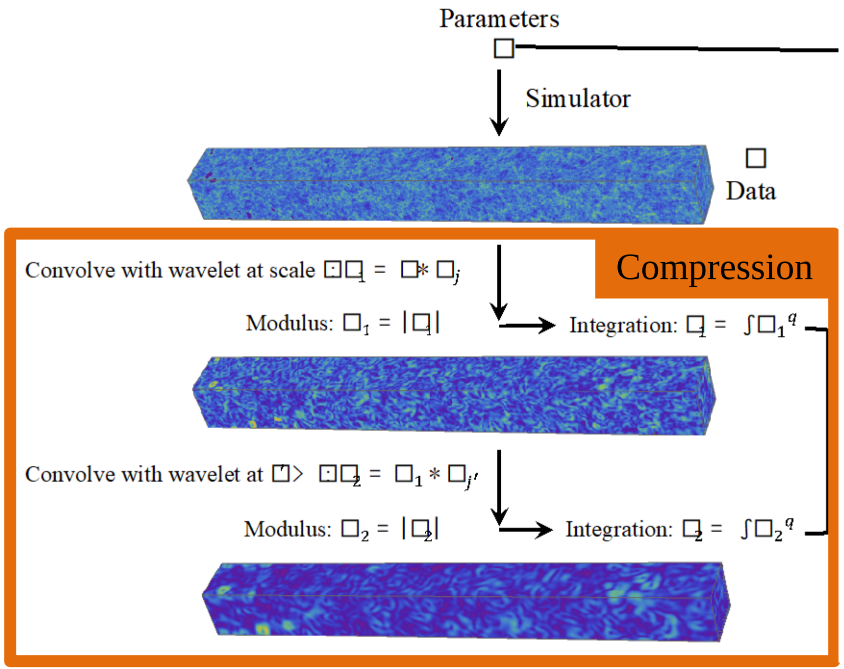
Structure of Convolutional Neural Networks (CNN)

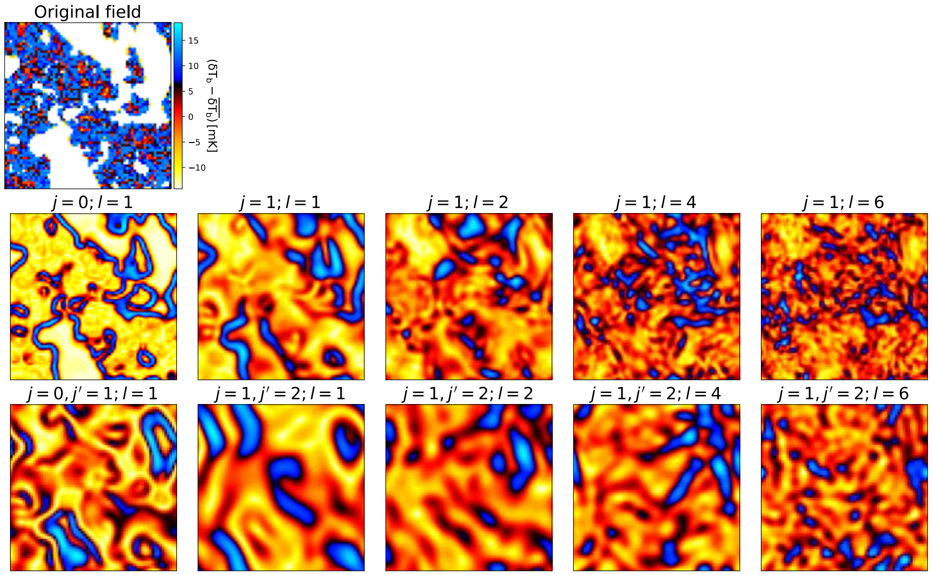


$\mathrm{T_{vir}}$ Minimum virial temperature of haloes that host star-forming galaxies

ζ UV Ionizing efficiency of galaxies.

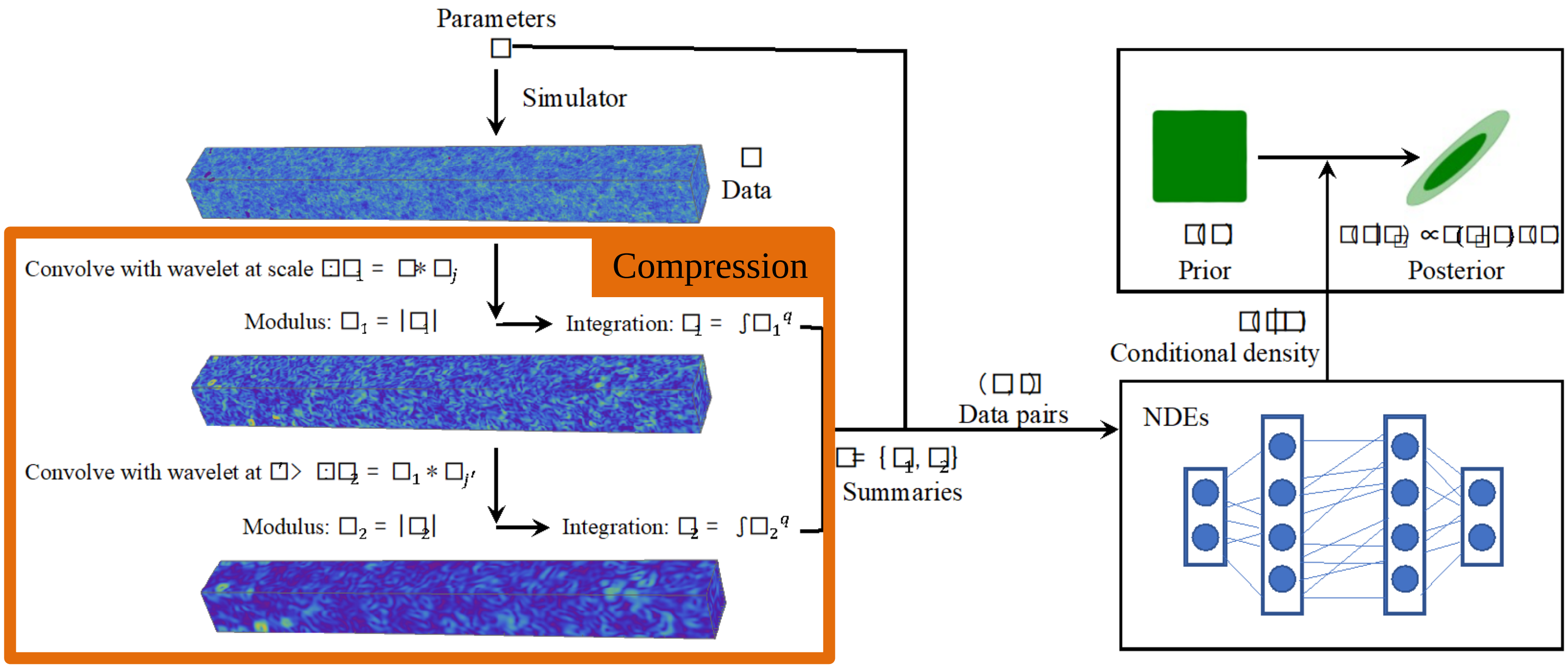
Likelihood-free Bayesian inference





Solid harmonic wavelet scattering transform (**SHWST**)

Likelihood-free Bayesian inference

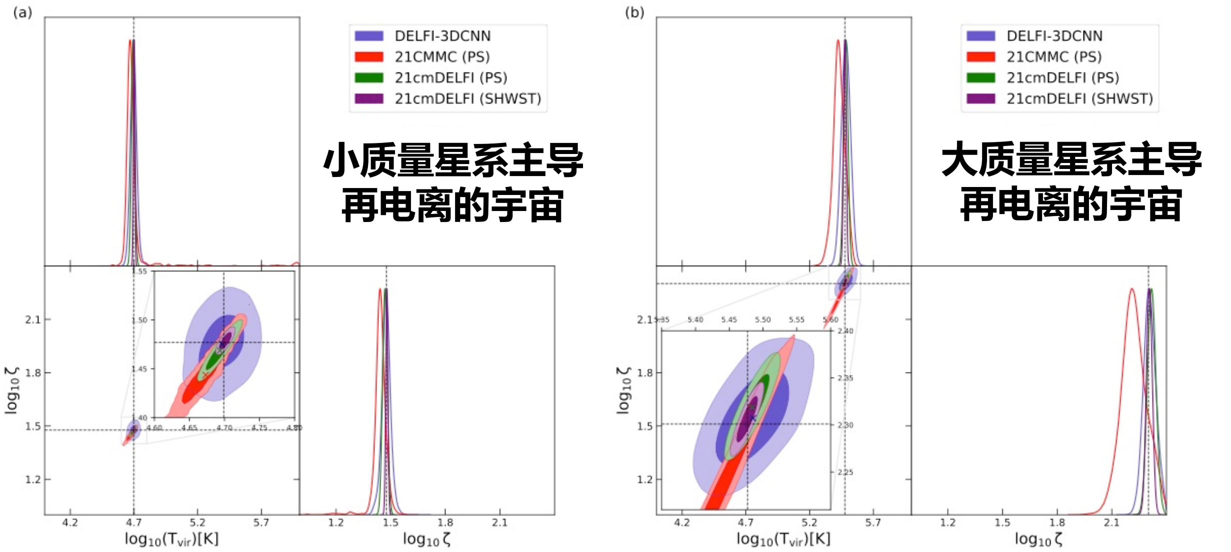


Solid harmonic wavelet scattering transform (**SHWST**)

从21厘米图像测量出发限制宇宙再电离理论模型

- 对21厘米图像信号进行降维（球谐小波散射变换），发展了基于深度学习进行贝叶斯统计推断的新方法，开发了新软件21cmDELFI-ST

对宇宙再电离模型参数进行贝叶斯统计推断的示例



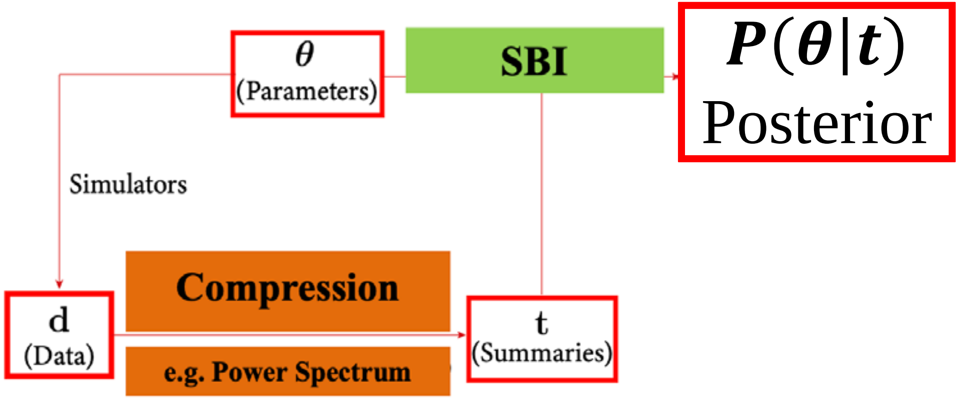
Zhao, **YM**, Zuo & Wandelt, in prep.

MCMC得到的置信区间

21cmDELFI-PS得到的置信区间

21cmDELFI-ST得到的置信区间

Simulation-Based Inference



2. Discriminative Way (Neural Ration Estimation, NRE):

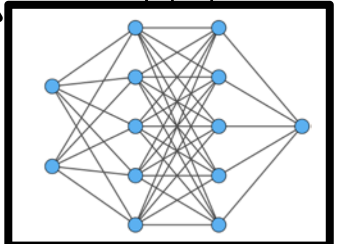
Ce, **YM**, et al
In prep

$$\{\theta, t\} \xrightarrow{\text{Ratio estimation}} r(\theta,t) = \frac{P(t|\theta)}{P(t)} \xrightarrow{\text{Prior}} \boldsymbol{P(\theta|t) = r(\theta,t)P(\theta)}$$

$$\boxed{P(\theta,t) = P(t|\theta)P(\theta)}$$

$$\boxed{P(\theta)P(t)}$$

$$\{\theta, t\} \xrightarrow[D(t,\theta)]{\text{Classifier}} \max_{D} E_{p(\theta,t)}[\log(D(t,\theta)] + E_{p(\theta)p(t)}[\log(1 - D(t,\theta))]$$



$$r(\theta,t) = \frac{P(x|t)}{P(t)}$$

$$\boxed{D^*(t,\theta) = \frac{r(\theta,t)}{r(\theta,t)+1}}$$

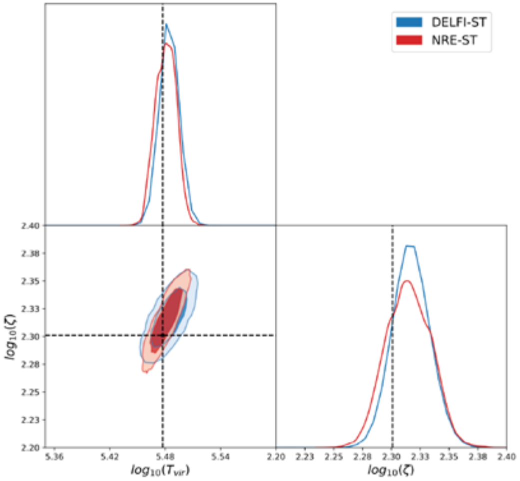
Comparison between two SBI methods

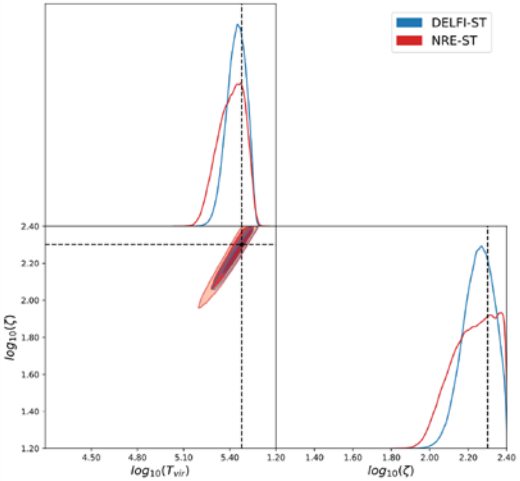
Pure Signal

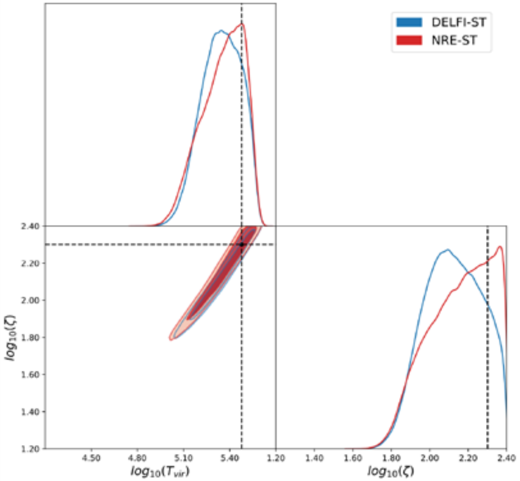
+ Thermal Noise

+Residual Foreground

Bright Galaxy Model







Ce, **YM**, et al. in prep

宇宙再电离时代的观测证据

- **已有探针：**

☑ Lyα **森林光深测量：** 再电离结束于红移6
☑ **宇宙微波背景辐射：** 再电离渐变
☑ Lyα **发射体：** 再电离非均匀 (also see Fengshan Liu’s talk)

- **未来探针：**

☐ **中性氢21厘米谱线强度映射 (e.g. bispectrum - see Siyi Zhao’s talk; polarization)**
☐ **分子谱线强度映射，及其与21厘米交叉关联 (Also see Olivier Perdereau’s talk)**
☐ **中性氢21厘米森林（See the talks of Yue Shao, Wenkai Hu）**
☐ **低红移Lyα 森林功率谱 (Montero-Camacho & Mao, MNRAS, 2020, 2021, 2023)**