

# Inferring the timing and morphology of cosmic reionization through forward models

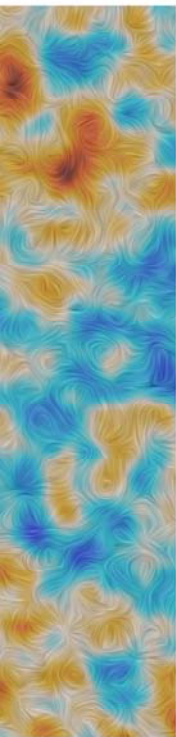
Andrei Mesinger



# First billion years - birth of structure and Cosmic Dawn

Image: ESA

**CMB**



$$z \approx 10^3$$



$$4 \cdot 10^5$$

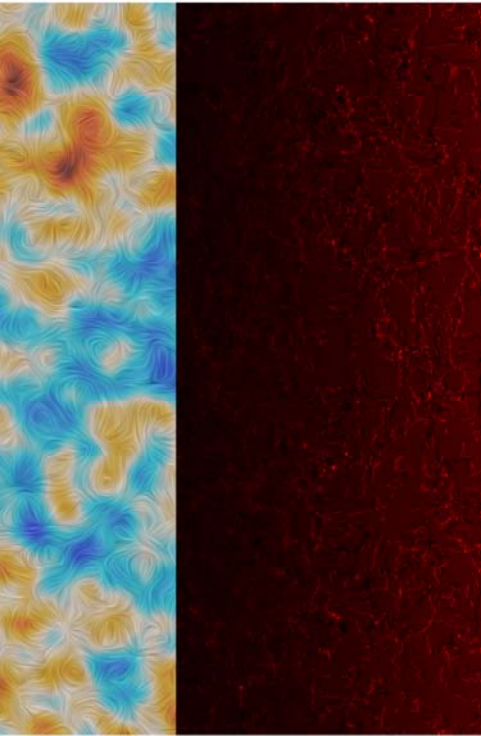
cosmic time [yr]

# First billion years - birth of structure and Cosmic Dawn

Image: ESA

AM+2016

## CMB Dark Ages



$z \approx 10^3$

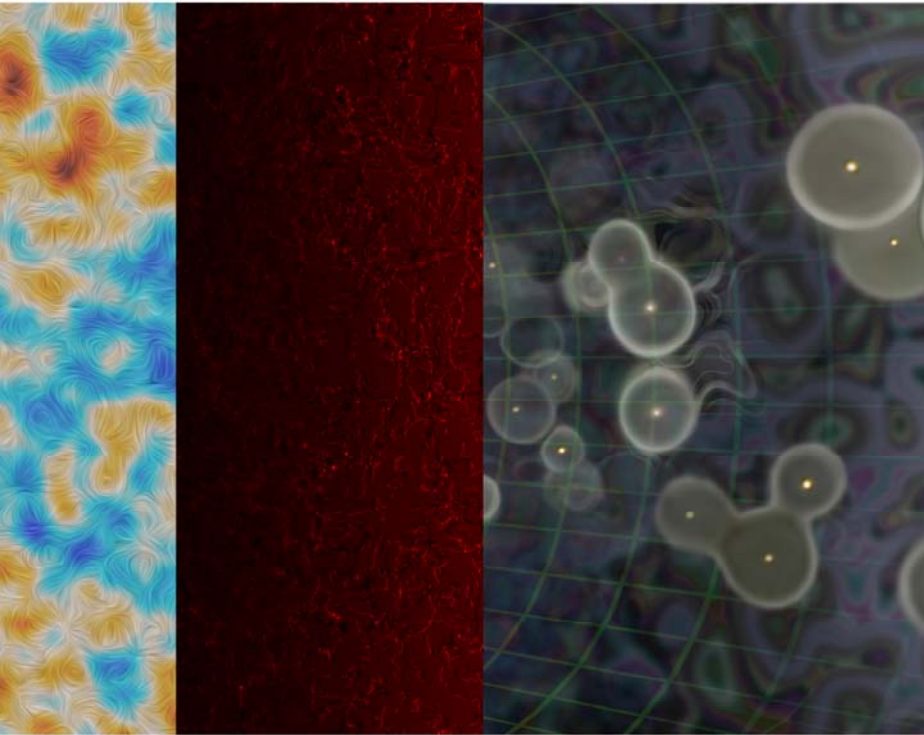
$z \approx 30$

4 · 10<sup>5</sup>                      10<sup>8</sup>                      cosmic time [yr]

# First billion years - birth of structure and Cosmic Dawn

Image: NASA/  
CXC/M.WEISS  
AM+2016; J. Munoz

CMB Dark Ages **Cosmic Dawn**



$z \approx 10^3$

$z \approx 30$

$4 \cdot 10^5$

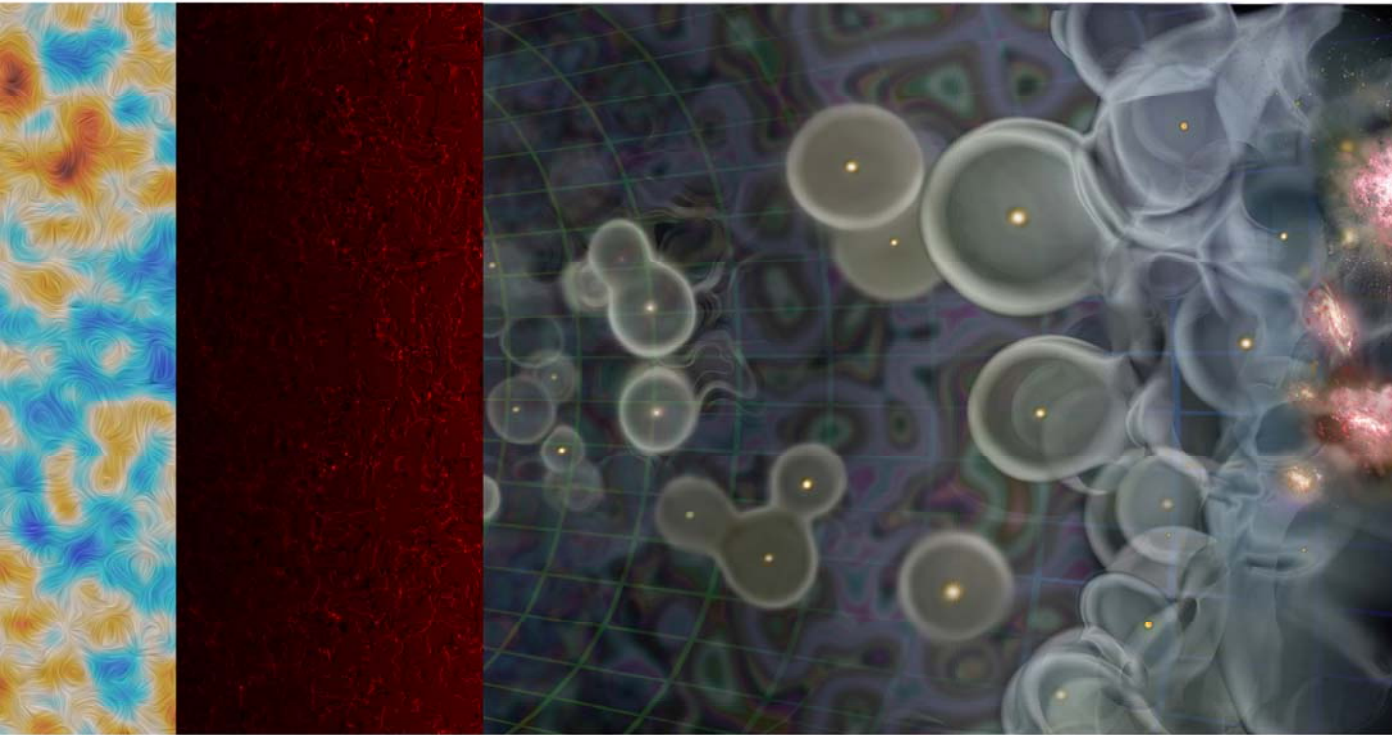
$10^8$

cosmic time [yr]

# First billion years - birth of structure and Cosmic Dawn

Image: NASA/  
CXC/M.WEISS  
AM+2016; J. Munoz

CMB Dark Ages Cosmic Dawn Reionization



$z \approx 10^3$

$z \approx 30$

$z \approx 5$

$4 \cdot 10^5$

$10^8$

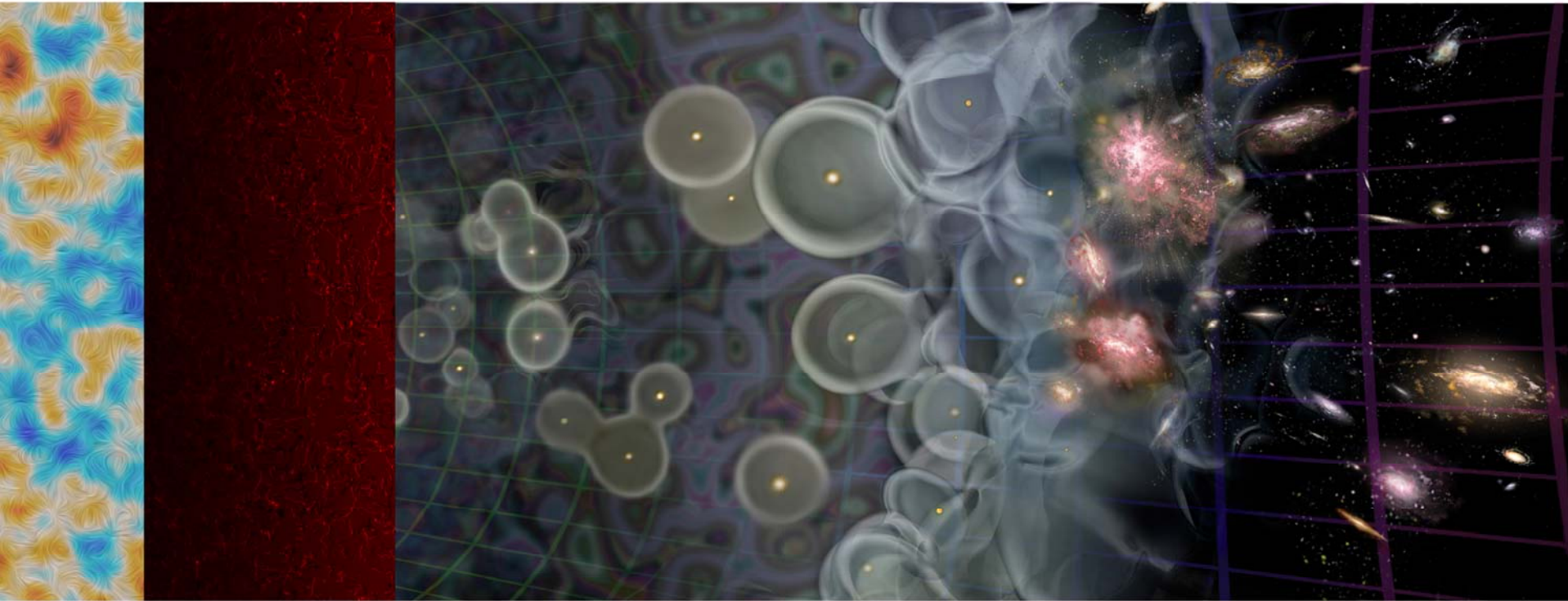
$10^9$

cosmic time [yr]

# First billion years - birth of structure and Cosmic Dawn

Image: NASA/  
CXC/M.WEISS  
AM+2016; J. Munoz

CMB    Dark Ages    Cosmic Dawn    Reionization    Late Universe



$z \approx 10^3$                        $z \approx 30$                        $z \approx 5$                        $z = 0$

cosmic time [yr]

$4 \cdot 10^5$

$10^8$

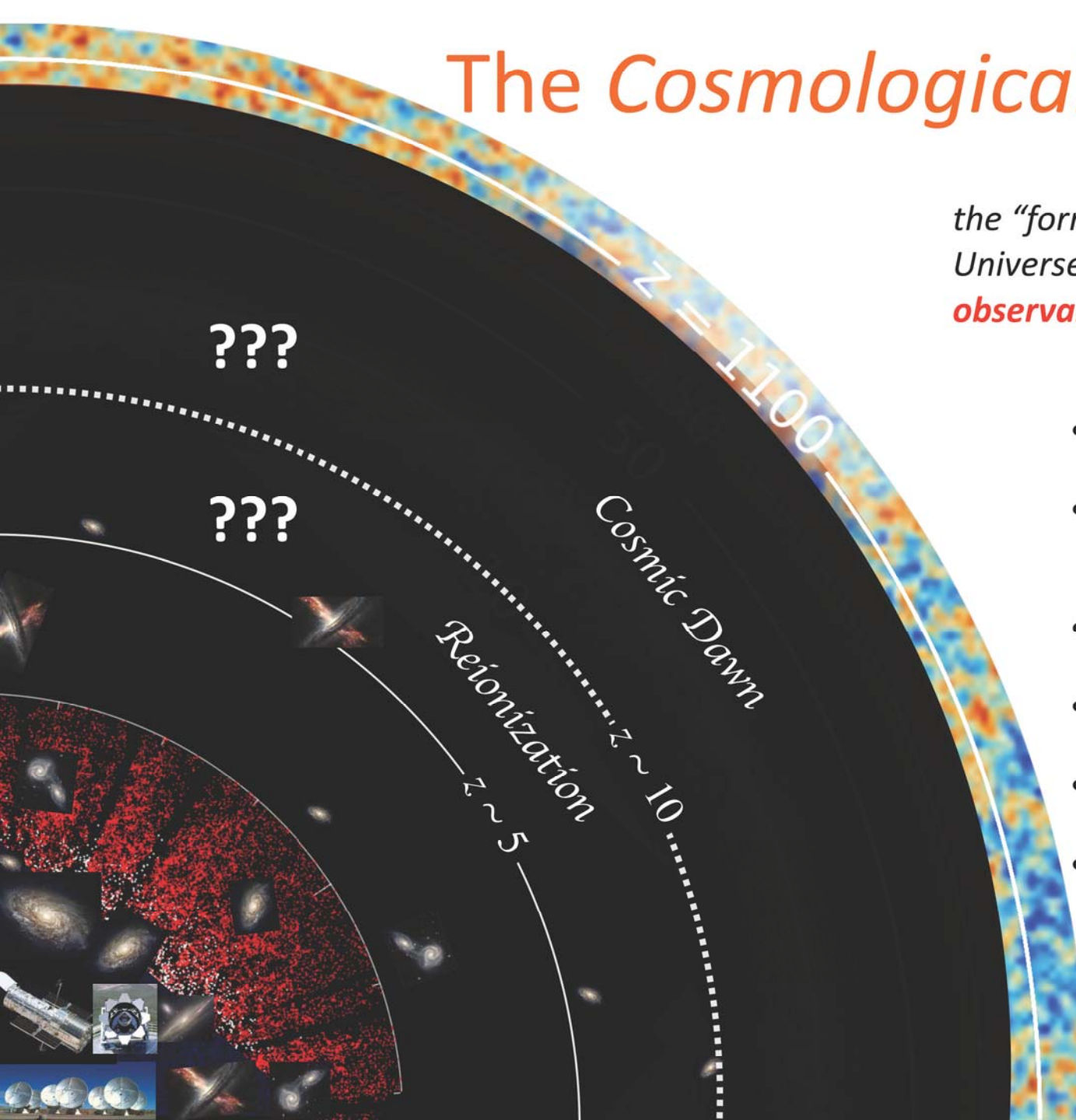
$10^9$

$10^{10}$

***The Epoch of Reionization:  
the final phase change of our Universe***

# The *Cosmological Frontier*...

the “formative childhood” of the Universe, yet the **majority of the observable volume**

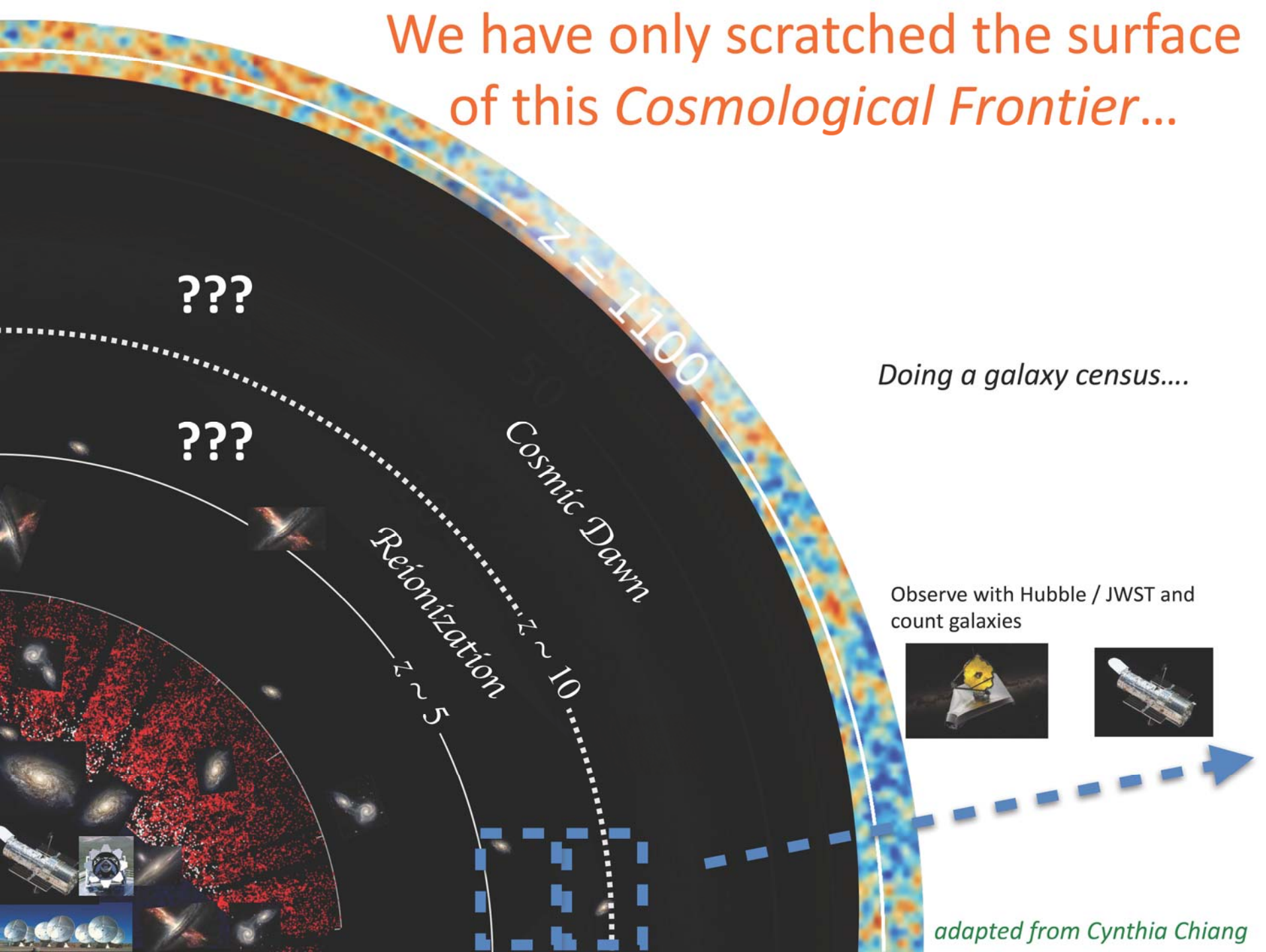


- When and how did the first galaxies form?
- How did they impact each other and their surroundings?
- What are the dominant feedback mechanisms?
- Can we learn about Dark Matter properties?
- How does the Hubble parameter evolve?
- What are the properties of the first stars and black holes?

adapted from Cynthia Chiang

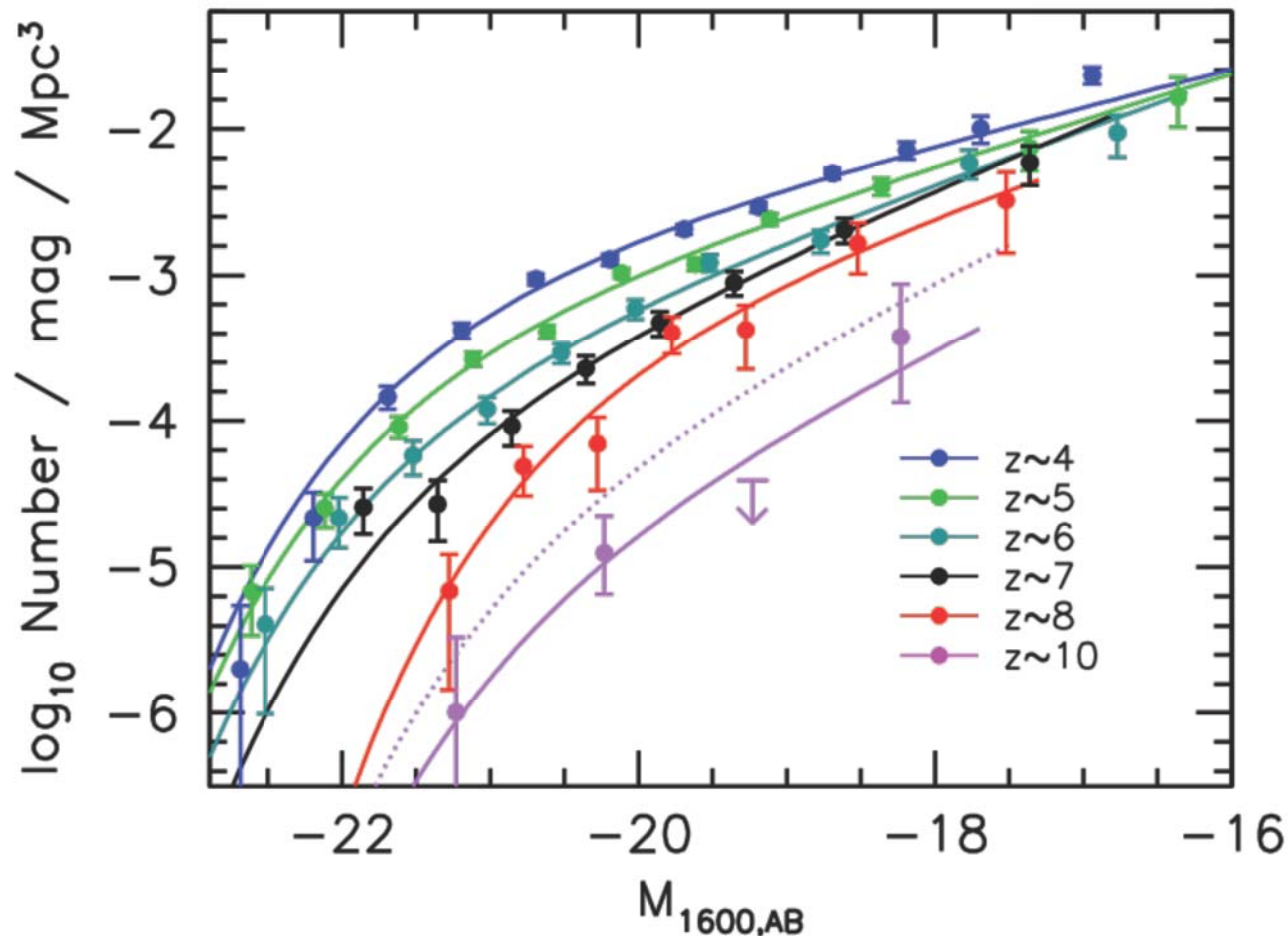


# We have only scratched the surface of this *Cosmological Frontier*...



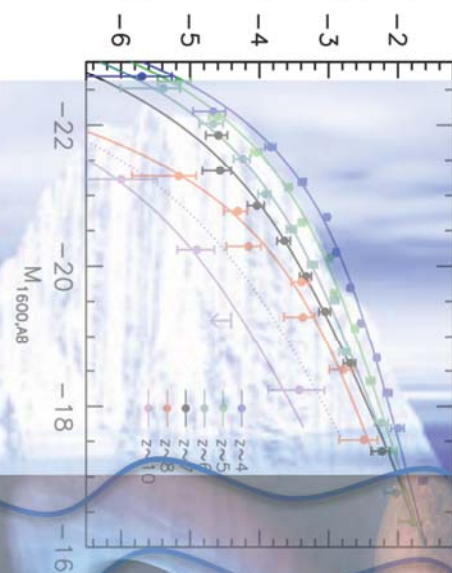
# Galaxies during the first billion years

- Telescopes like Hubble and ALMA have enabled detailed studies of *the brightest galaxies*



Bouwens+ (2015)

$\log_{10}$  Number / mag / Mpc<sup>3</sup>



Hubble limit  
(no lensing)

JWST limit  
(no lensing)

$M_{AB} = -22$

$M_{AB} = -18$

$M_{AB} = -14$

$M_{AB} = -10$

$M_{AB} = -6$

*>99.9% of the first galaxies  
will not be seen even with JWST*

*The first stars and black holes*

hidden population of  
abundant, faint galaxies??

H<sub>2</sub> cooling

H-cooling threshold



$\log_{10}$  Number / mag / Mpc<sup>3</sup>

9 5 4 3 2

$M_{AB} = -22$

$M_{AB} = -18$

$M_{AB} = -14$

$M_{AB} = -10$

$M_{AB} = -6$

Hubble limit  
(no lensing)

JWST limit  
(no lensing)

H-cooling threshold

H<sub>2</sub> cooling





$\log_{10}$  Number / mag / Mpc<sup>3</sup>

9 5 4 3 2

$M_{AB} = -22$

$M_{AB} = -18$

$M_{AB} = -14$

$M_{AB} = -10$

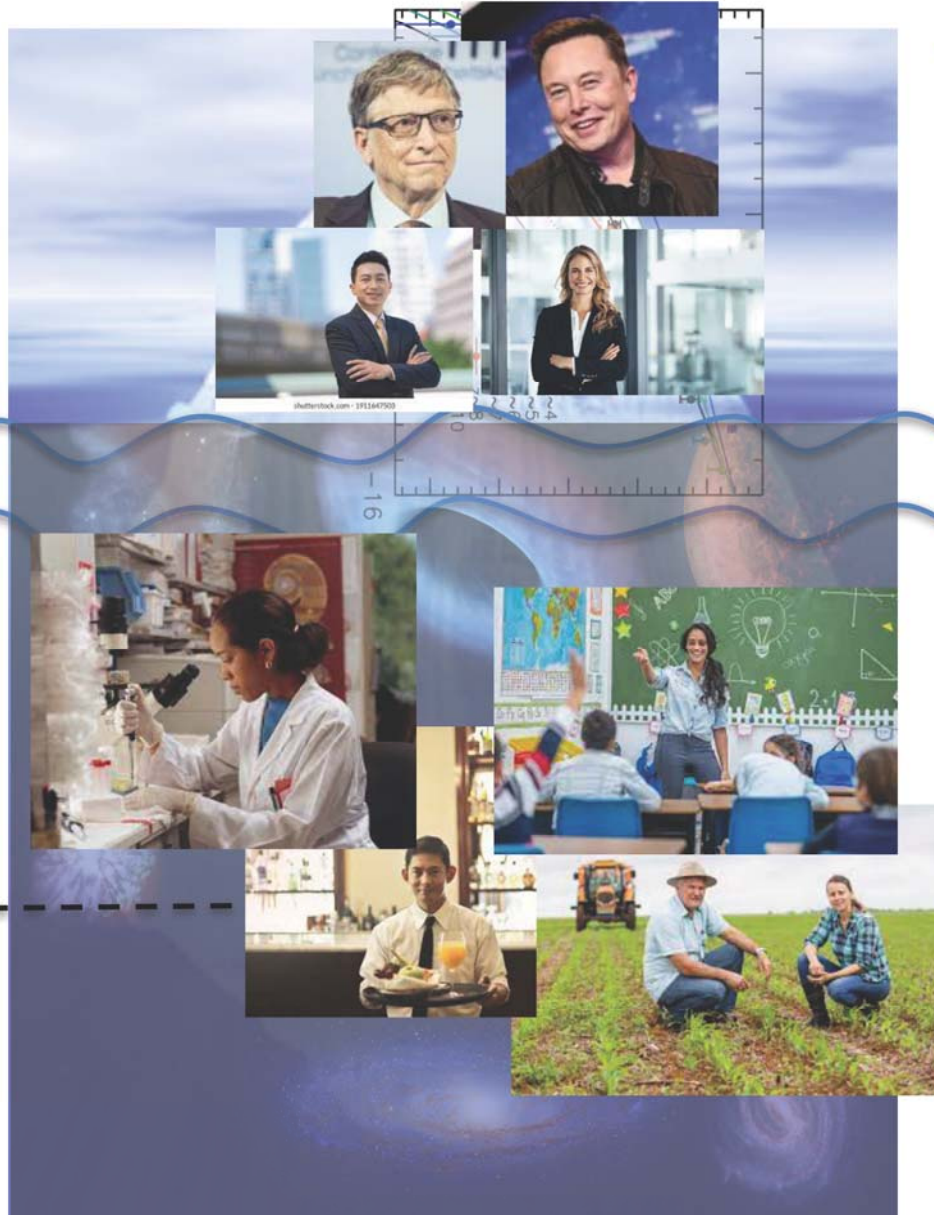
$M_{AB} = -6$

Hubble limit  
(no lensing)

JWST limit  
(no lensing)

H-cooling threshold

H<sub>2</sub> cooling





$\log_{10}$  Number / mag / Mpc<sup>3</sup>

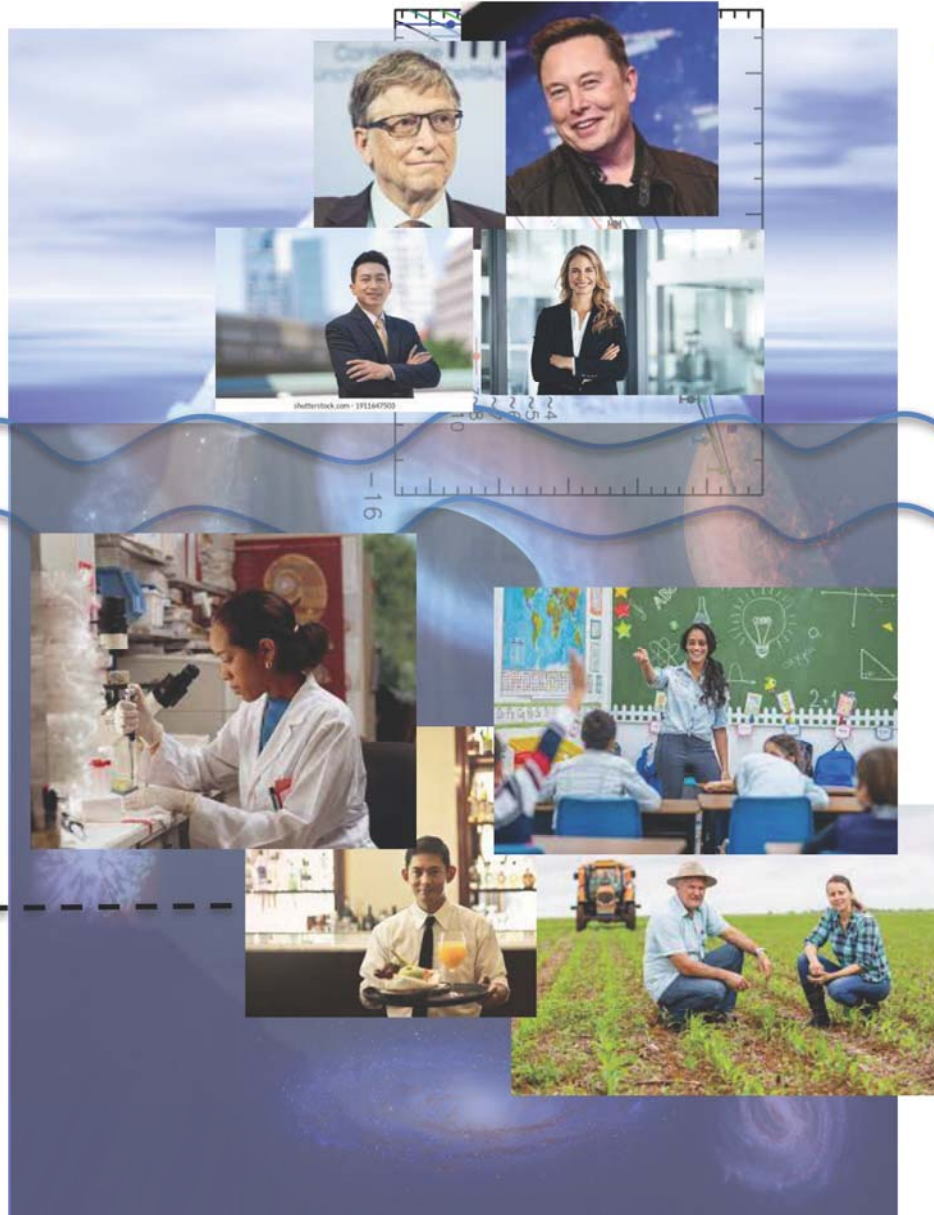
6 5 4 3 2

Hubble limit  
(no lensing)

JWST limit  
(no lensing)

H-cooling threshold

H<sub>2</sub> cooling



$M_{AB} = -22$

$M_{AB} = -18$

$M_{AB} = -14$

$M_{AB} = -10$

$M_{AB} = -6$

Direct studies of **INDIVIDUAL** galaxies

Indirect studies of **POPULATION AVERAGES** through IGM imprints

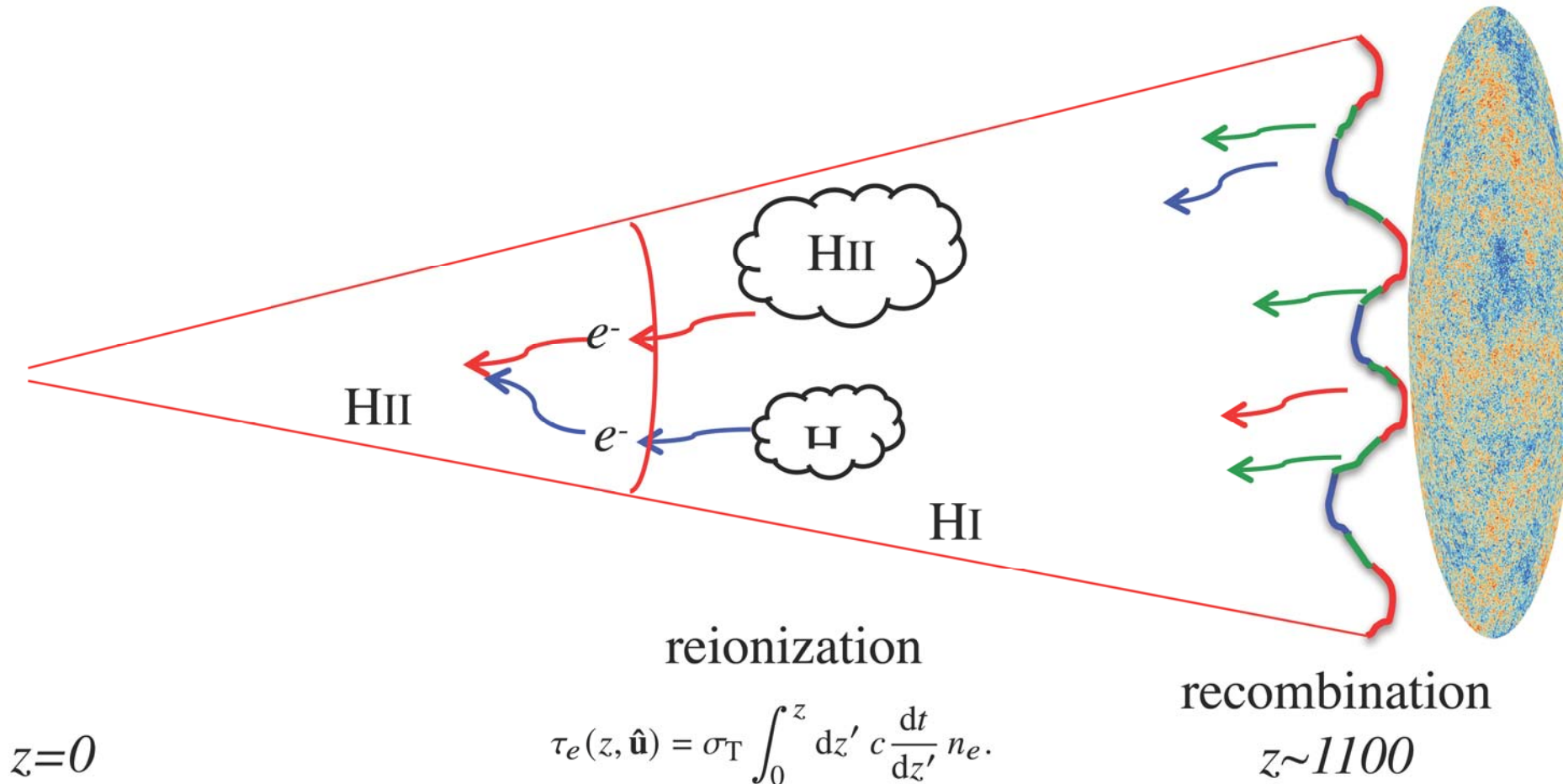
# Why use the IGM to learn about the EoR?

- The IGM temperature and density is shaped by the cumulative radiation from *all* galaxies → a *democratic* probe of all photons
- Lots of statistics
- Not biased by proximity to sources

# Studying the EoR

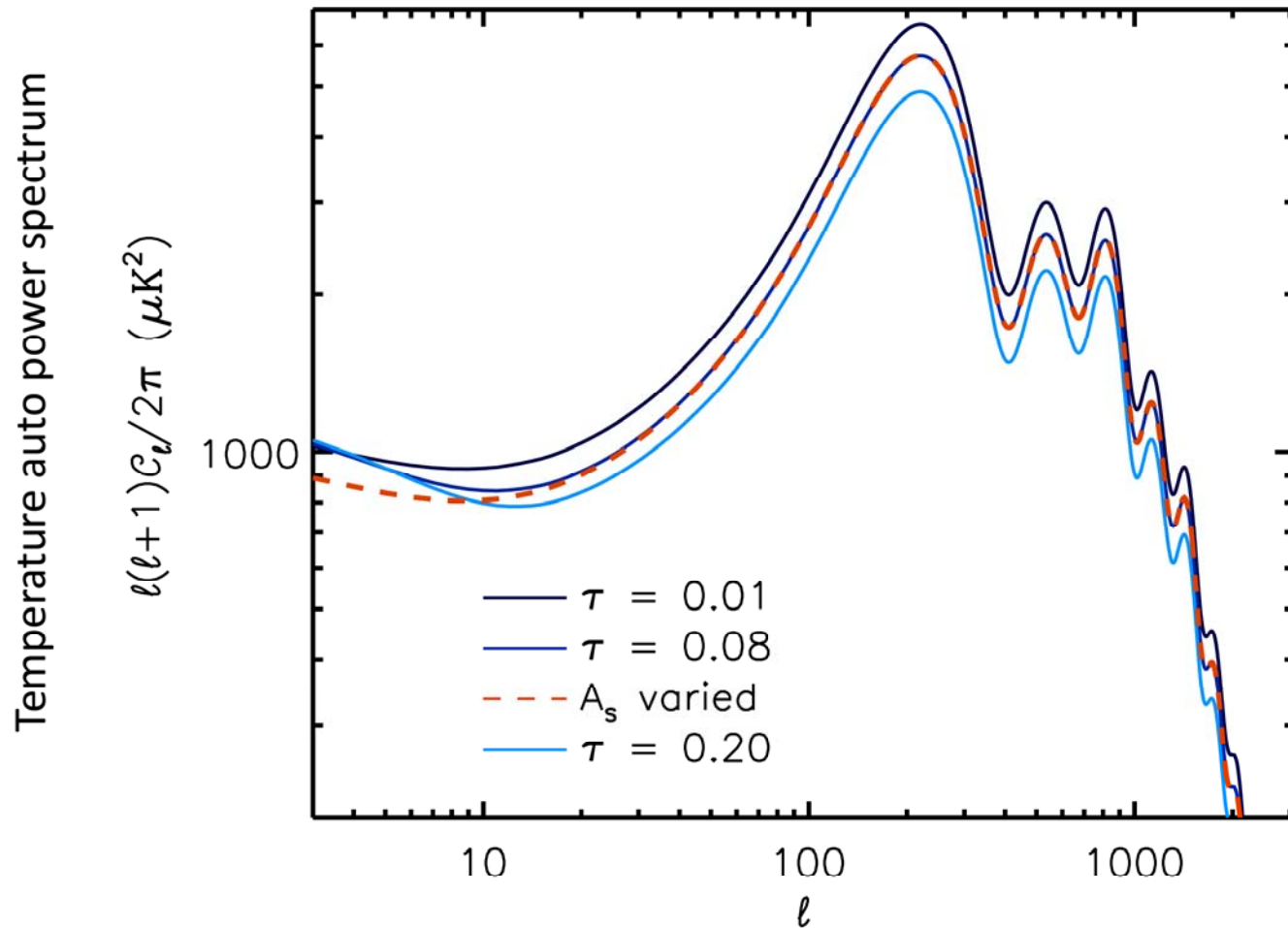
- Two main classes of probes

## 1. Integral CMB constraints (e.g. $\tau_e$ , kinetic SZ)



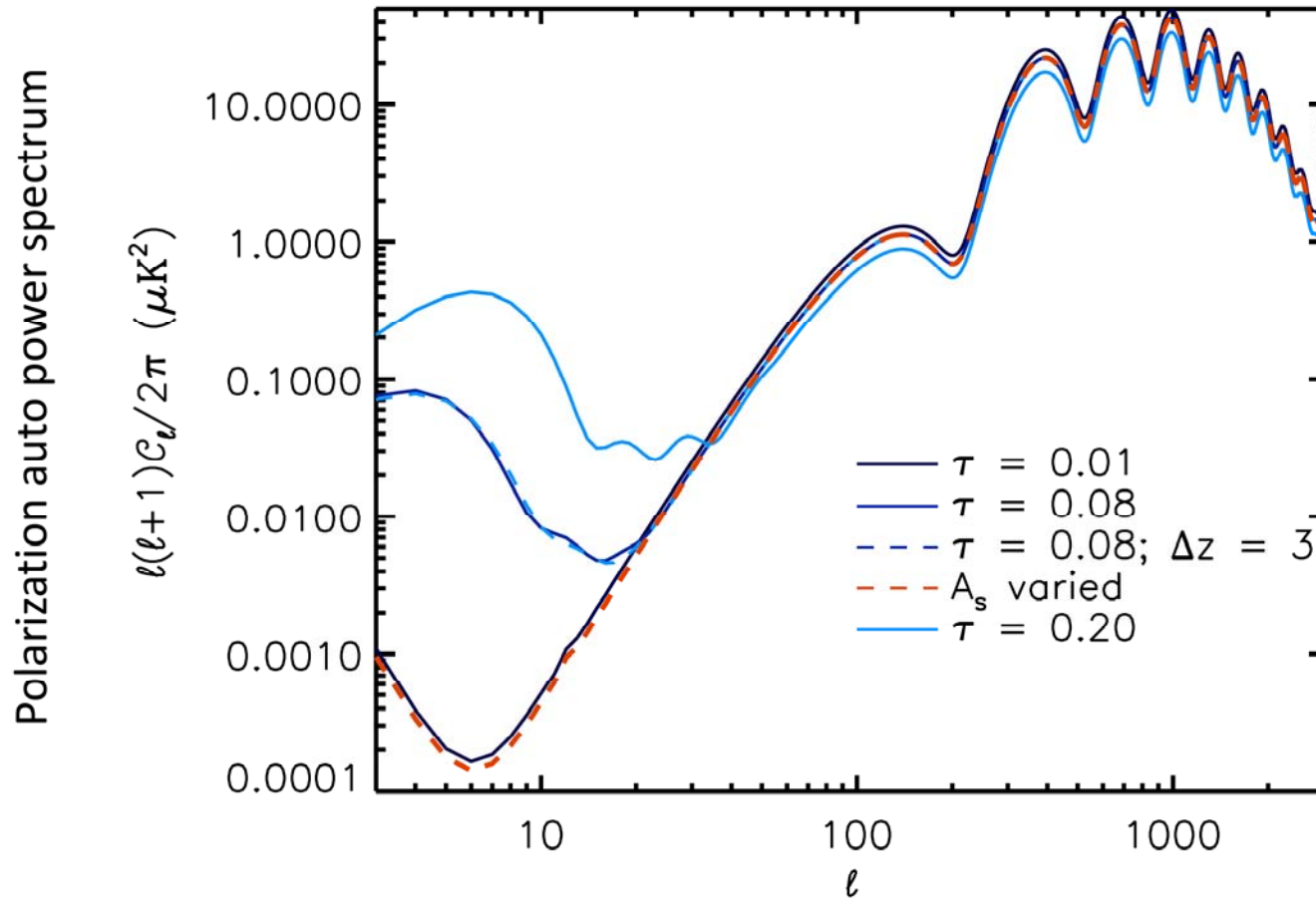


# Optical depth to the CMB



$$\tau_e(z, \hat{\mathbf{u}}) = \sigma_T \int_0^z dz' c \frac{dt}{dz'} n_e.$$

# Optical depth to the CMB

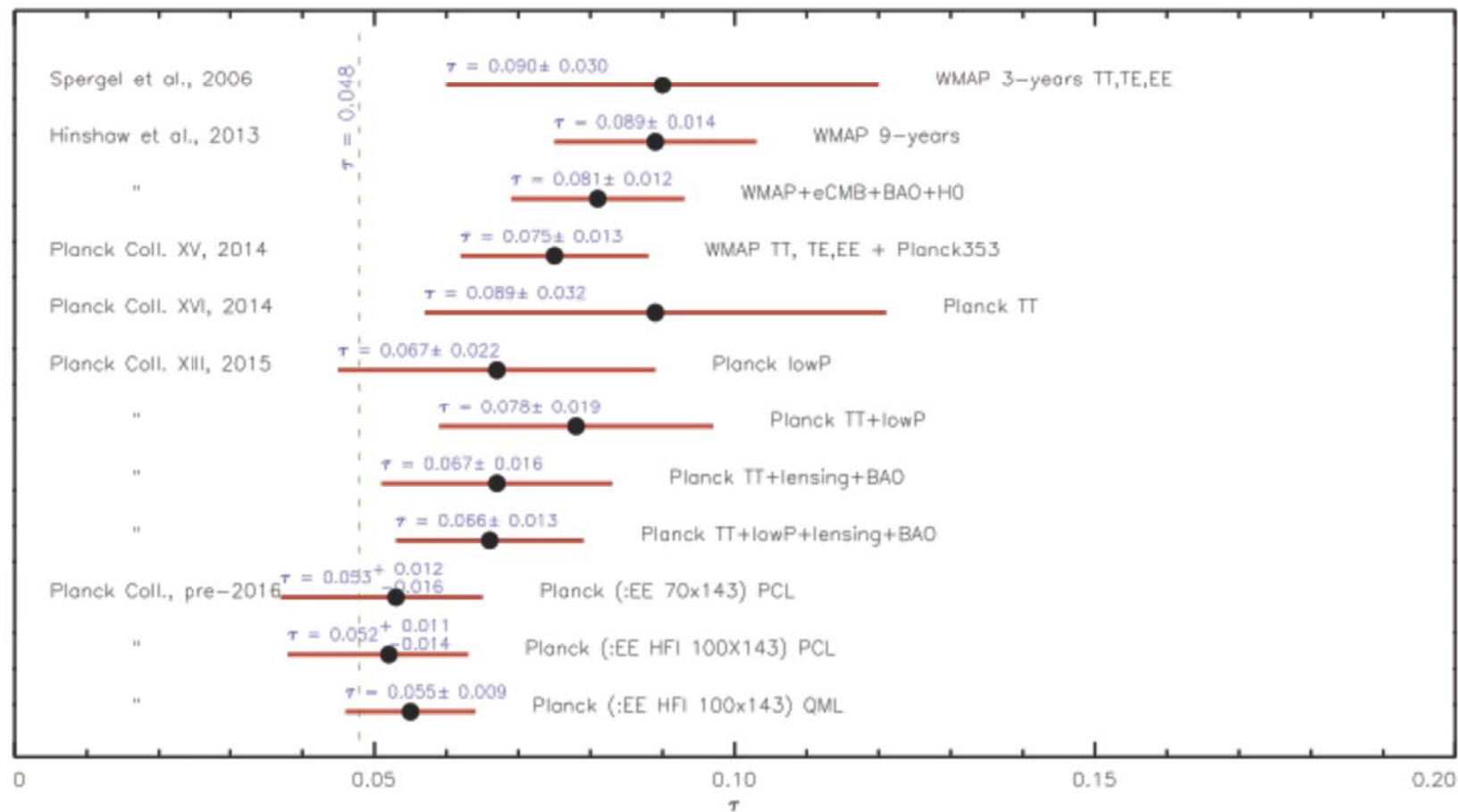


*note the size of the signal*

$$\tau_e(z, \hat{\mathbf{u}}) = \sigma_T \int_0^z dz' c \frac{dt}{dz'} n_e.$$

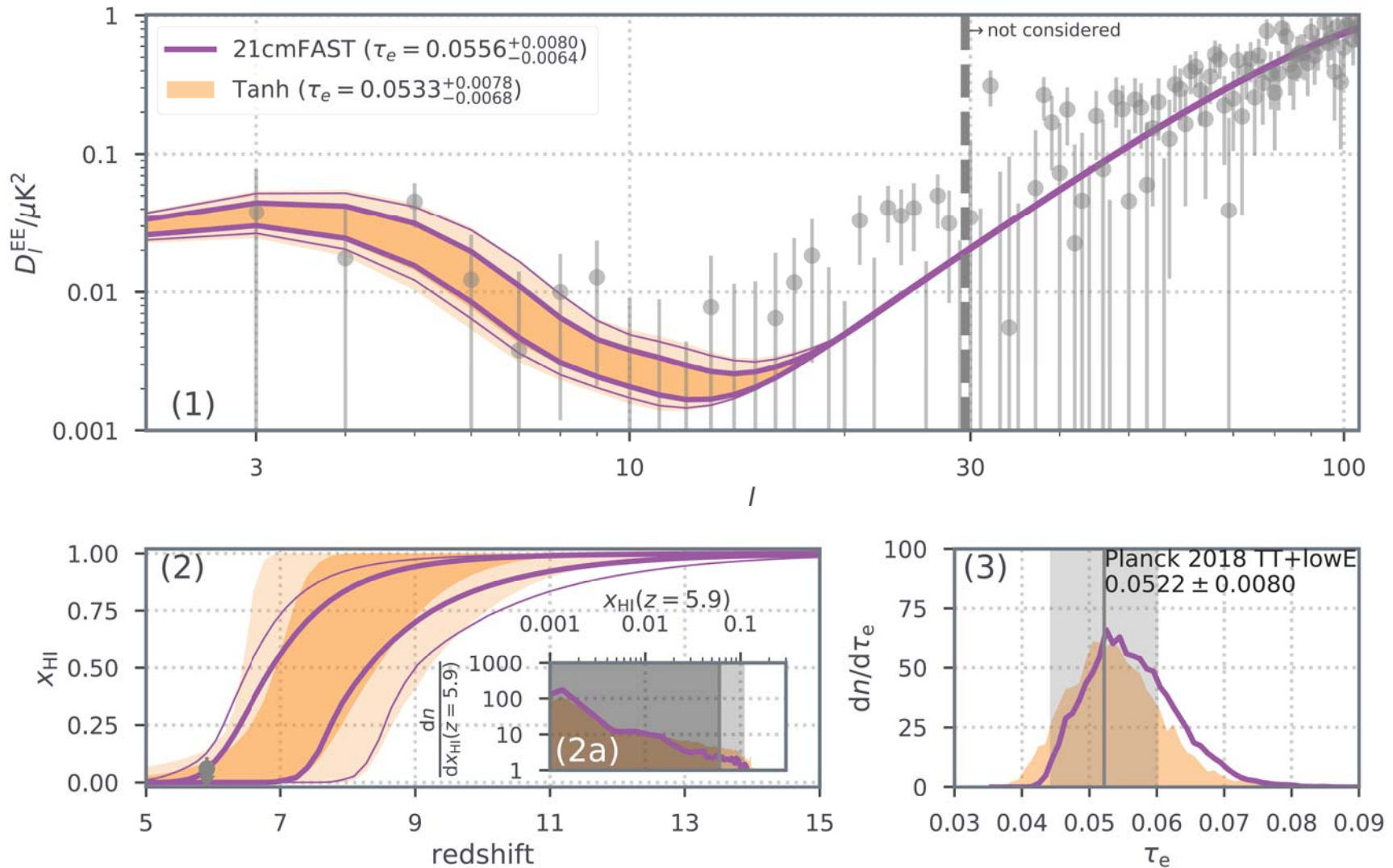
# History of Thompson scattering optical depth measurements

WMAP1 2003



Planck 2016

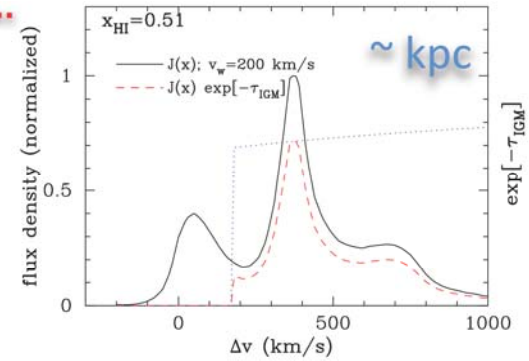
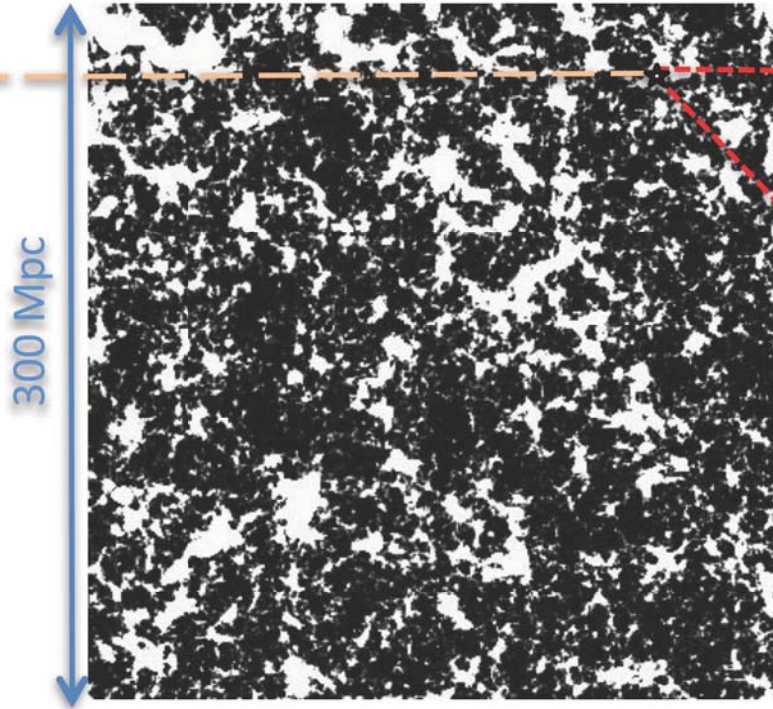
# Current EoR constraints from the CMB



# When?

- Two main classes of probes
  1. Integral CMB constraints (e.g.  $\tau_e$ , kinetic SZ)
  2. Astrophysical 'flashlights' (e.g. high- $z$  galaxies, QSOs)

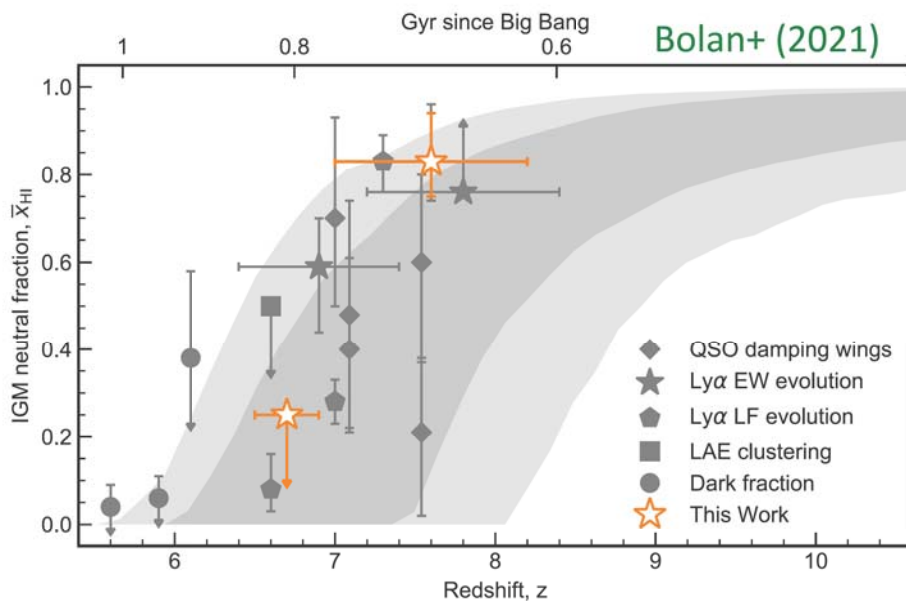
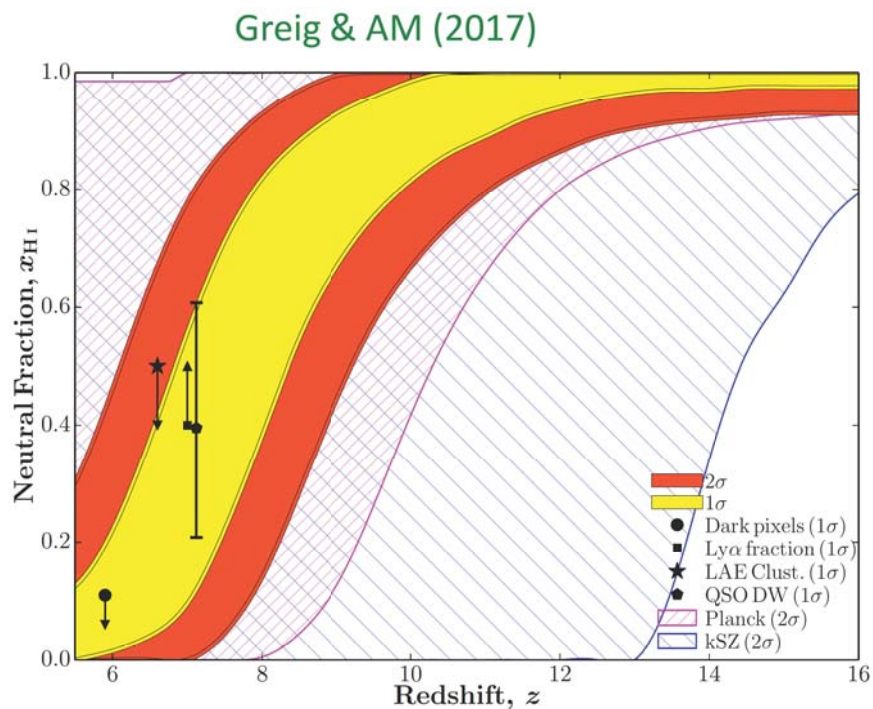
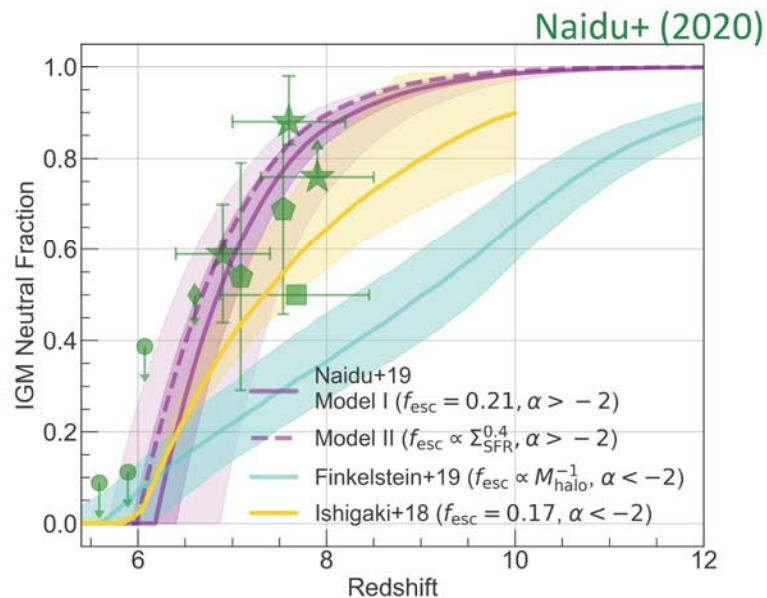
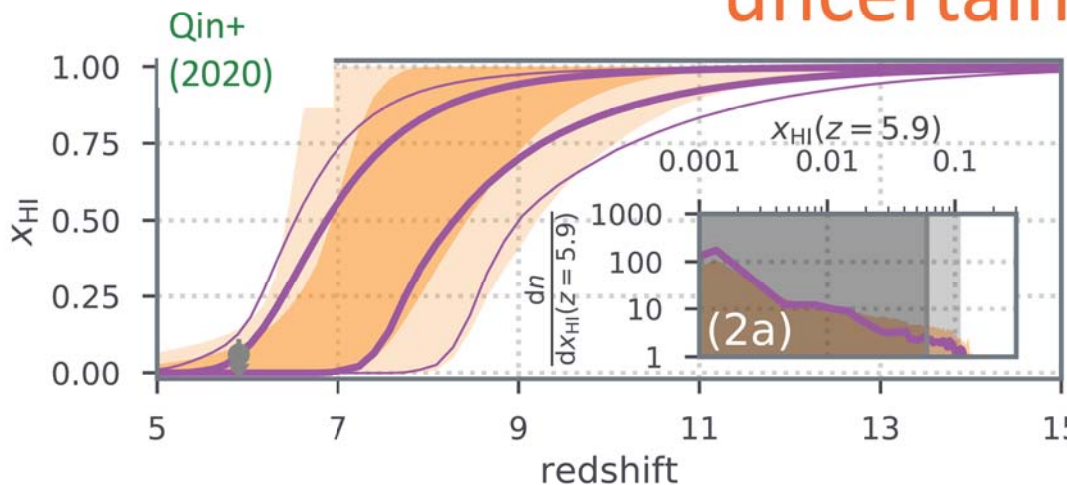
# Ly $\alpha$ damping wing absorption as a probe of the EoR



during reionization, **cosmic HI patches** absorb Ly $\alpha$  photons in the damping wing of the line

- We have constraints from:
  1. Ly $\alpha$  forest opacity
  2. QSO damping wings
  3. QSO near zones
  4. Redshift evolution of Ly $\alpha$  EW
  5. Redshift evolution of LAE
  6. Clustering of LAEs

# Until recently, the timing of the EoR was fairly uncertain...





# *Improved inference from the Ly $\alpha$ forest*

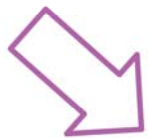
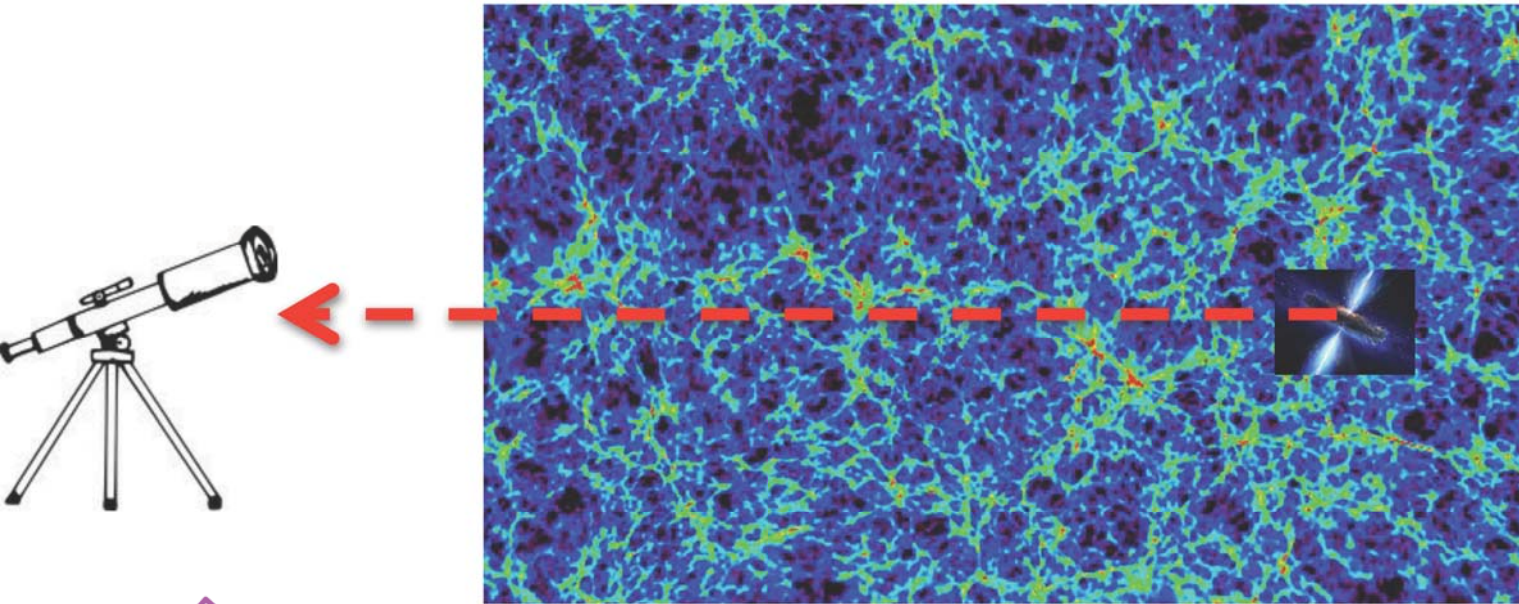
- We have constraints from:
  1. Ly $\alpha$  forest opacity
    - Sensitive to the end stages of the EoR
    - Probes large volumes of the Universe: (i) amazing statistics; (ii) not biased by local environment of sources
    - Unabsorbed continuum can be predicted to  $\sim 1-10\%$  level precision (e.g. Bosman+2021)
    - Beautiful recent dataset from XQR30



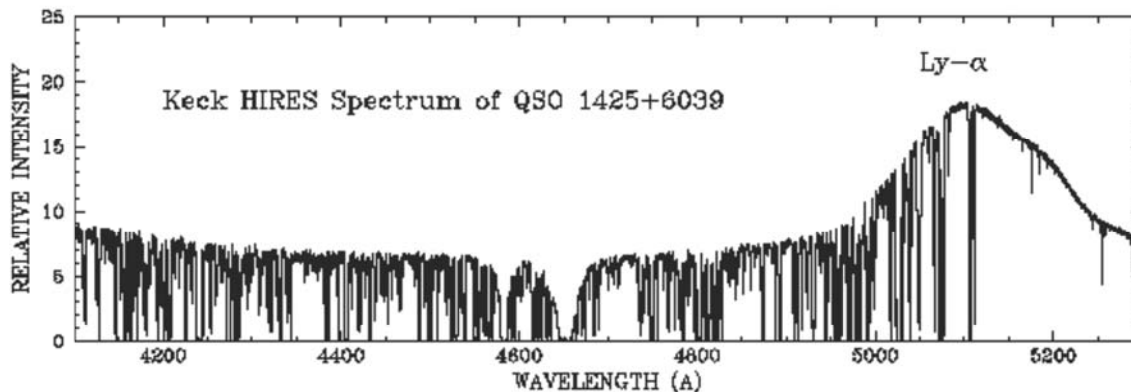
Qin, AM+ in prep

*A bit of history...*

# The Ly $\alpha$ forest is the oldest, most robust probe of the IGM



$$\tau_{\text{GP}} \approx 35.6\kappa \left( \frac{T_0}{7500 \text{ K}} \right)^{-0.724} \left( \frac{\Gamma_{\text{HI}}}{3 \times 10^{-13} \text{ s}^{-1}} \right)^{-1} \Delta^{2-0.724(\gamma-1)} \left( \frac{1+z}{6.6} \right)^{4.5},$$



# The presence of flux in the Ly $\alpha$ forest provided the first solid estimate of the ionization state of the IGM

1965ApJ...142.1633G

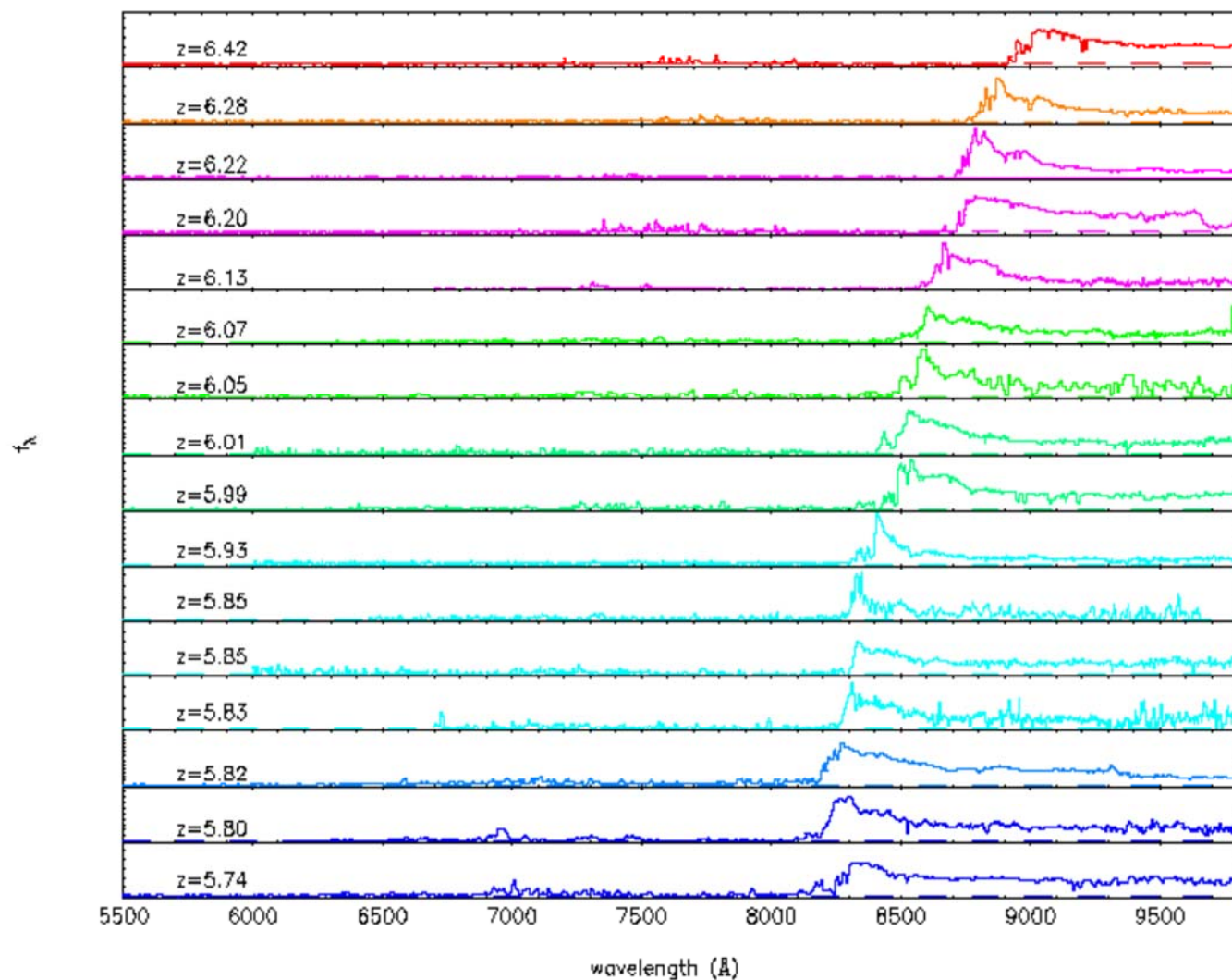
## NOTES

### ON THE DENSITY OF NEUTRAL HYDROGEN IN INTERGALACTIC SPACE

Recent spectroscopic observations by Schmidt (1965) of the quasi-stellar source 3C 9, which is reported by him to have a redshift of 2.01, and for which Lyman- $\alpha$  is in the visible spectrum, make possible the determination of a new **very low value for the density of neutral hydrogen in intergalactic space**. It is observed that the continuum of the source continues (though perhaps somewhat weakened) to the blue of Ly- $\alpha$ ; the line as seen on the plates has some structure but no obvious asymmetry. Consider, however, the fate of photons emitted to the blue of Ly- $\alpha$ . As we move away from the source along the line of sight, the source becomes redshifted to observers locally at rest in the expansion, and for one such observer, the frequency of any such photon coincides with the rest frequency of Ly- $\alpha$  in his frame and can be scattered by neutral hydrogen in his vicinity. The calculation of the size of the effect is very easily performed as follows:

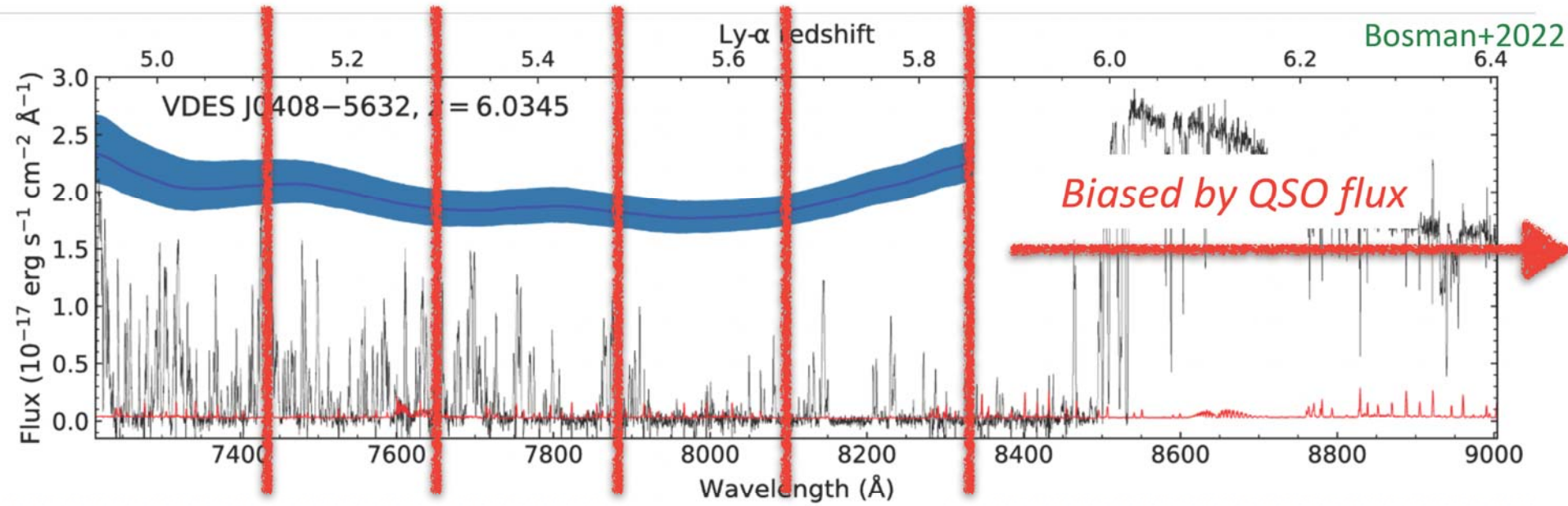
Gunn & Peterson (1965)

But we see less and less flux at higher  $z$



Fan+ (2006)

Usually quantified in terms of an “effective optical depth”

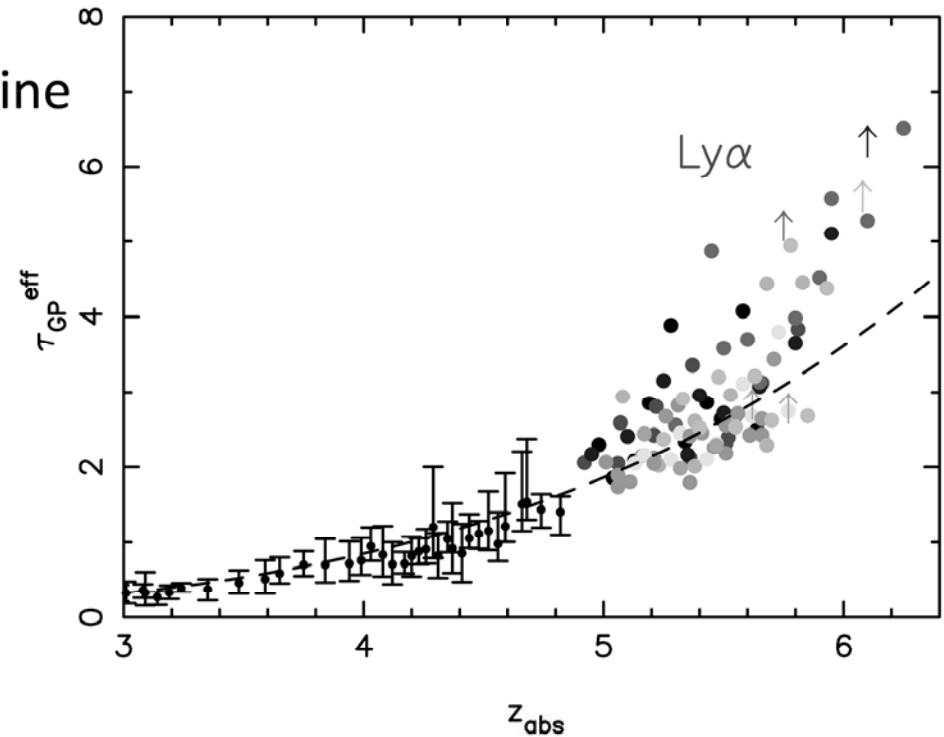


Define “effective optical depth” over each segment:

$$\tau_{\text{eff}} \equiv -\ln\langle f \rangle_{\Delta z=0.1}$$

# Ly $\alpha$ effective optical depth, circa 2006

Accelerated **rise** and sightline to sightline **scatter** of effective optical depth was suggestive of reionization at  $z \sim 6$

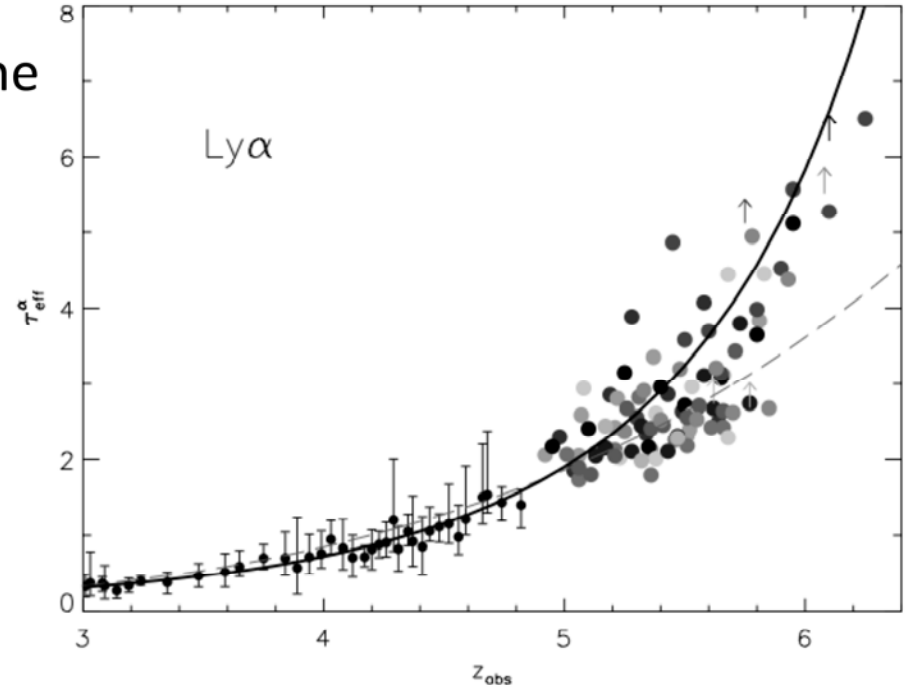


Fan+ 2006

# Ly $\alpha$ effective optical depth, circa 2006

Accelerated **rise** and sightline to sightline **scatter** of effective optical depth was suggestive of reionization at  $z \sim 6$

**\*BUT\***



Distributions of  $\tau_{\text{eff}}$  consistent with an ionized Universe at  $z < 6$   
(e.g. [Becker+2007](#); [Lidz+2007](#); [Gallerani+2008](#))

[Becker+ 2007](#)

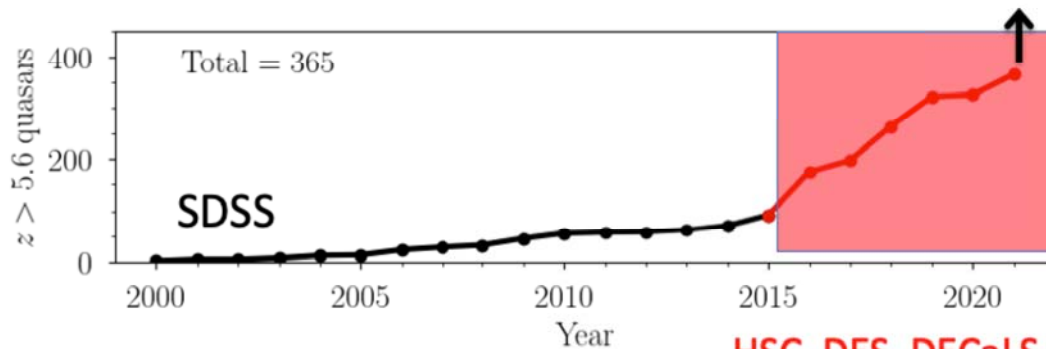
***Only fluctuations in the density field are sufficient to explain the data***

So we can't really tell from the forest data...



# Until we started getting lots more, high quality QSO data

~5x more quasars  
in the last 5 years



HSC, DES, DECaLS  
Pan-STARRS1, VHS, UHS

Andika+2020  
Bañados+2015-2021  
Belladitta+2020  
Carnall+2015  
Fan+2019  
Jiang+2015-2016  
Matsuoka+2016-2019  
Mazzucchelli+2017

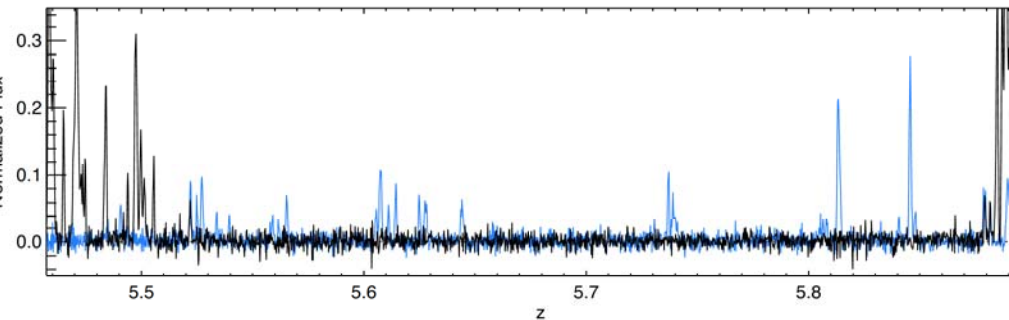
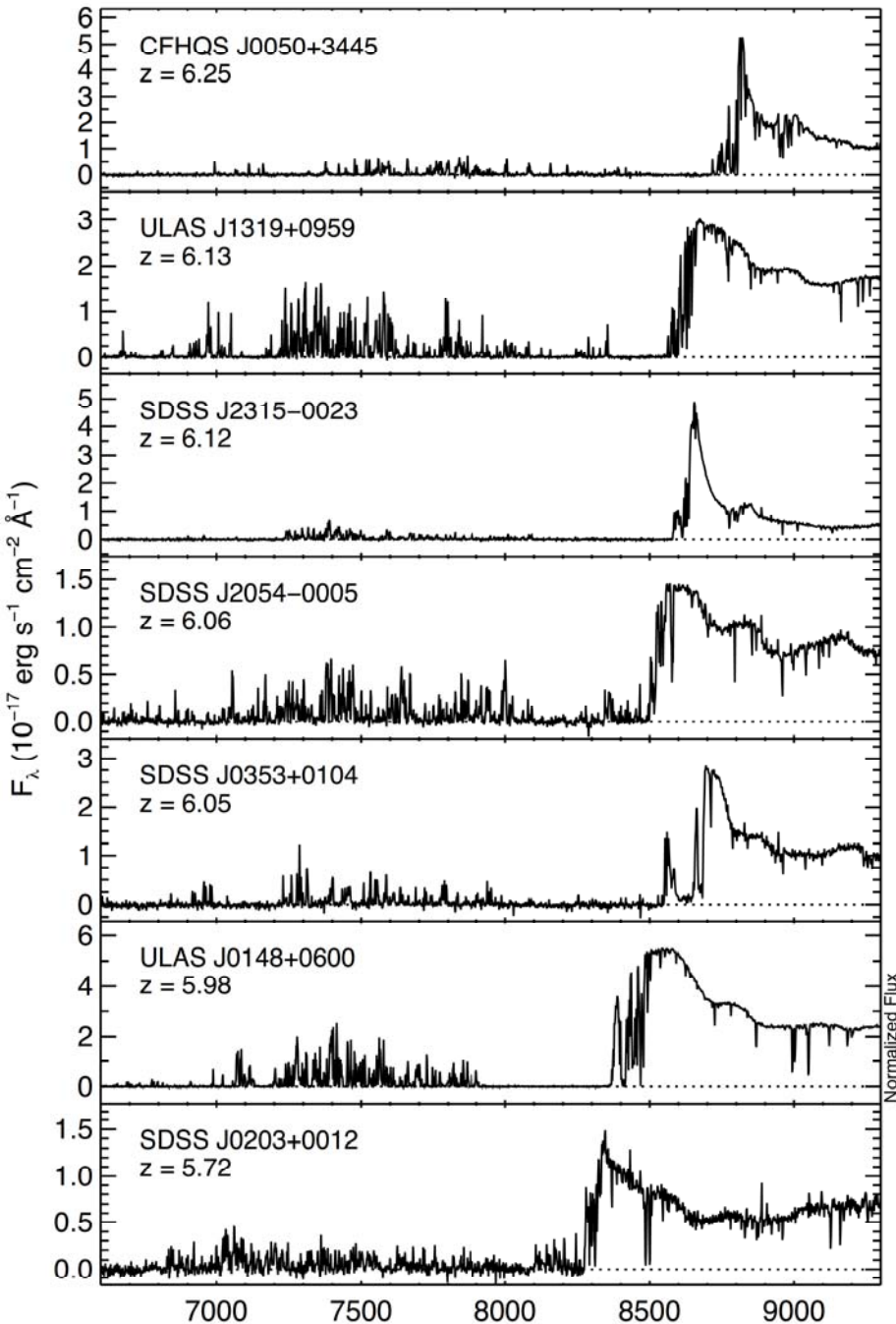
Pons+2019  
Reed+2015-2019  
Tang+2019  
Venemans+ 2015ab  
Wang+2016-2021  
Yang+2017-2020  
Wu+2015  
...

slide courtesy of E. Banados

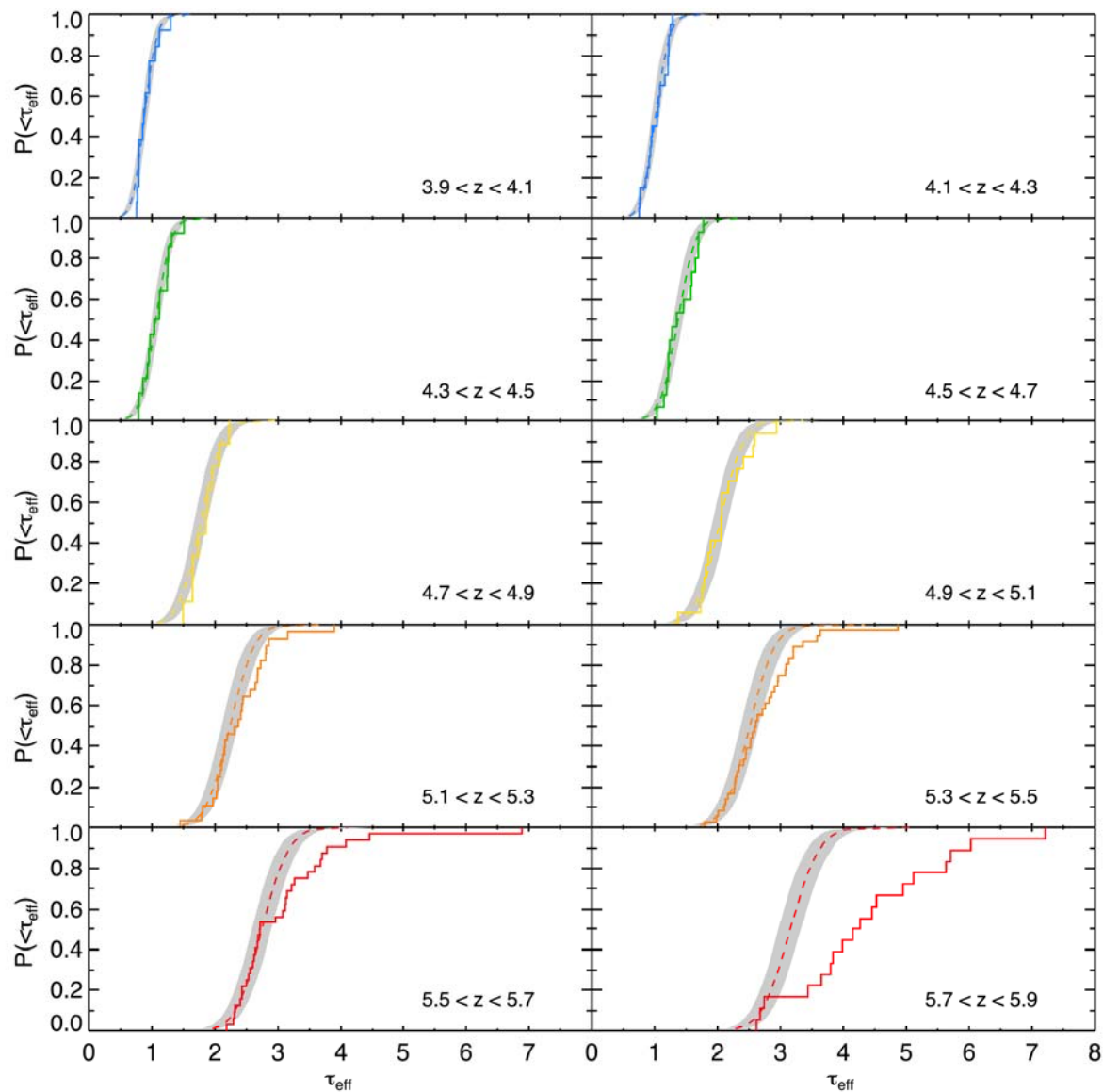
With even larger  
scatter and longer  
dark troughs

Becker+ 2015

GP through that is 160Mpc long!!!

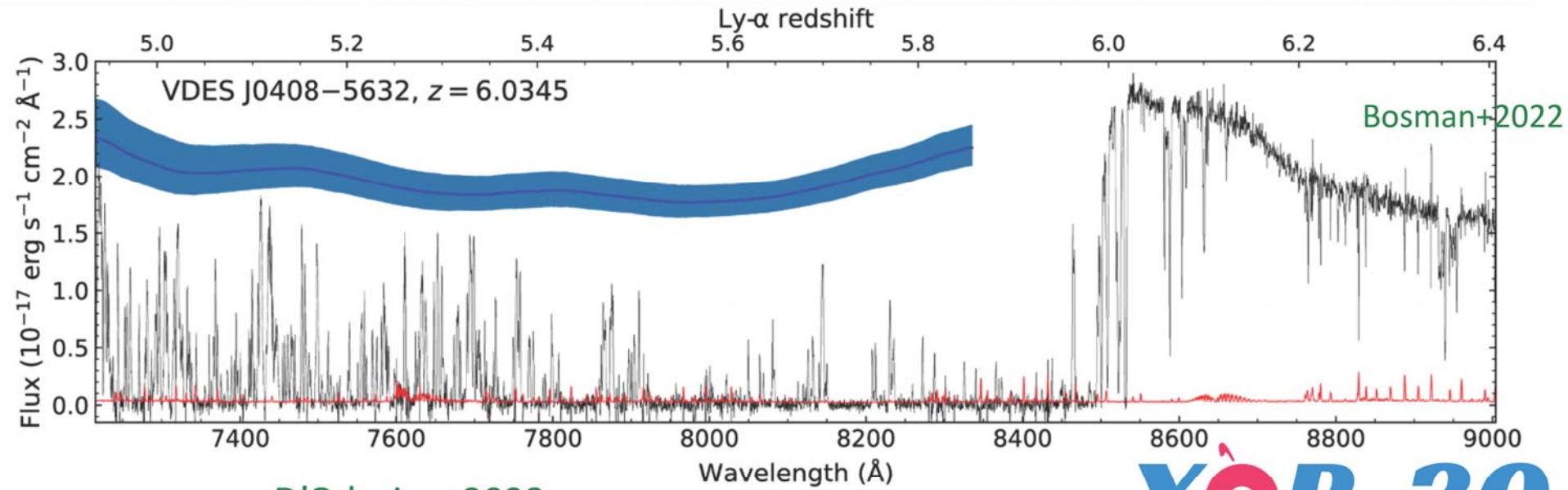


# Large scatter

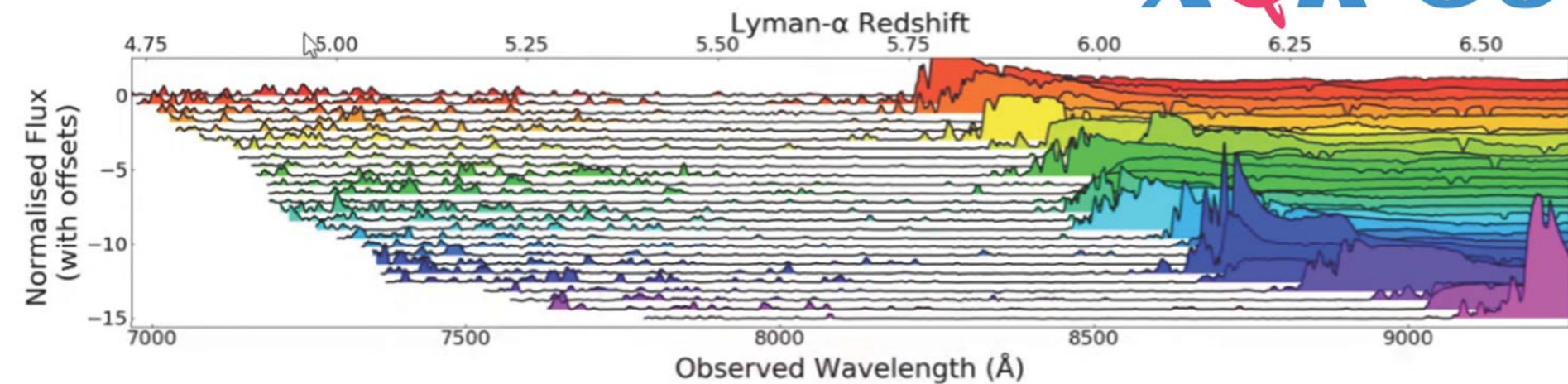


Distributions of  $\tau_{\text{eff}}$  are **TOO BROAD** to be consistent with a uniform, pervasive UVB at  $z > 5.2$

# XQR-30: super high-quality QSO spectra at $z=5.8-6.6$ using X-Shooter on VLT (PI: V. D'Odorico)

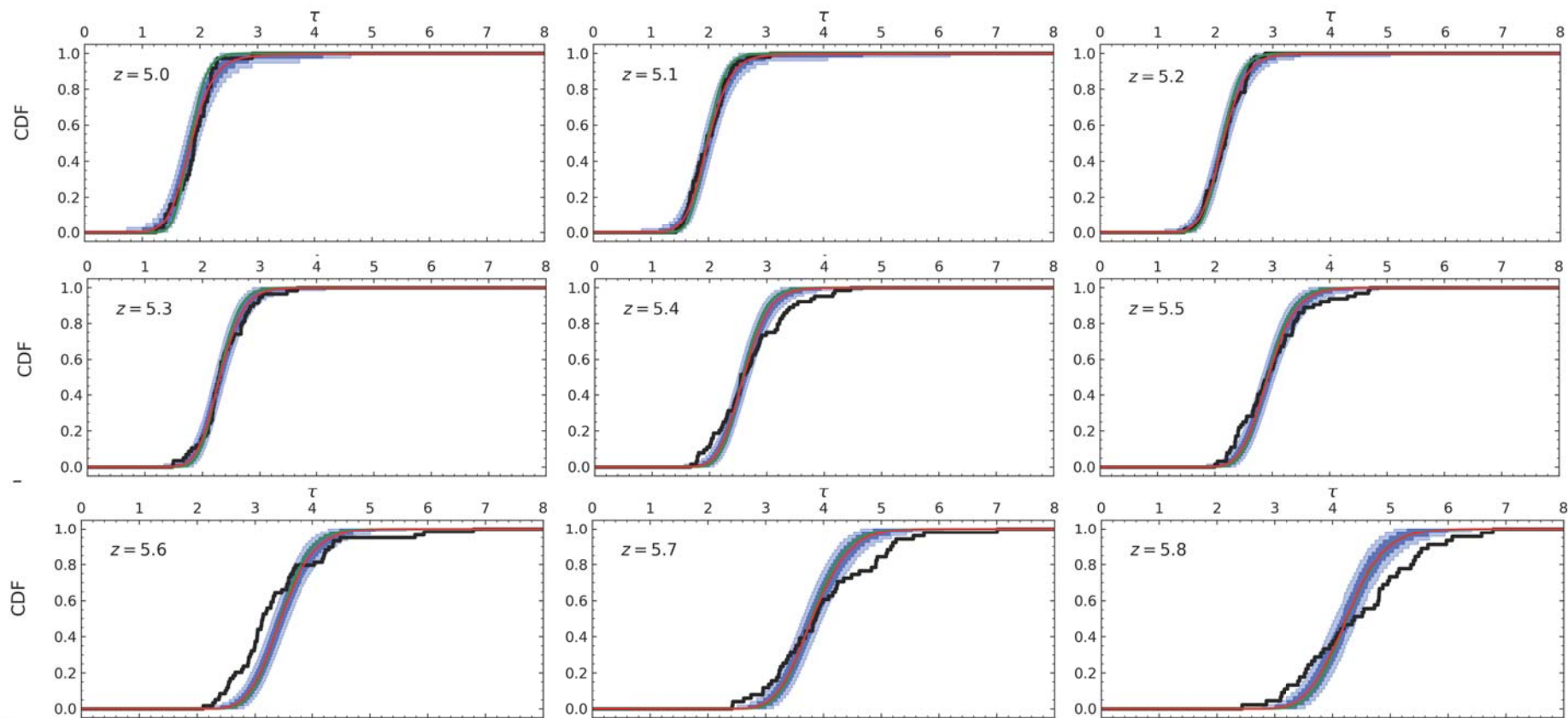


**XQR-30**



total sample >40 spectra probing forest at  $5 < z < 6.4$

# We need an additional source of scatter



Distributions of  $\tau_{\text{eff}}$  are **TOO BROAD** to be consistent with a uniform, pervasive UVB at  $z > 5.2$

Bosman+2021

**XQR-30**

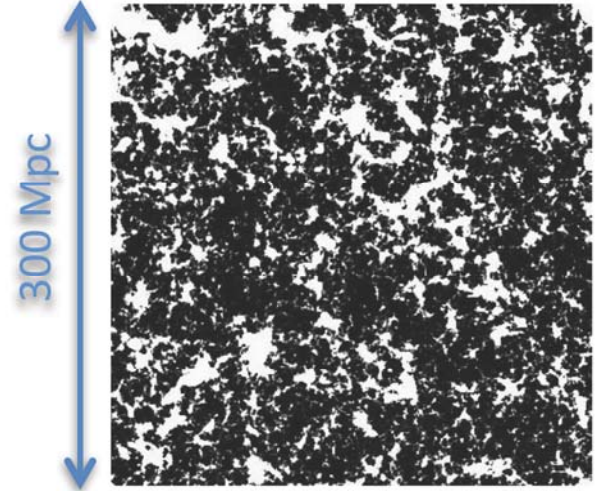
# What can increase fluctuations in $\tau_{\text{eff}}$ ?

Ionized component

$$\tau_{\text{HII}} \propto \frac{\Delta^2 \alpha_B(T)}{\Gamma_{\text{ion}}} \propto \frac{\Delta^2 \alpha_B(T)}{\epsilon_{\text{ion}} \lambda_{\text{mfp}}}$$

Neutral component

$$\tau_{\text{HI}} \sim 10^5 - 10^6$$



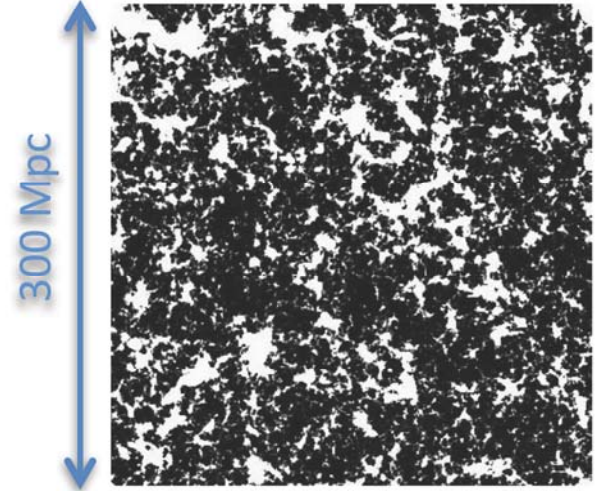
# What can increase fluctuations in $\tau_{\text{eff}}$ ?

Ionized component

$$\tau_{\text{HII}} \propto \frac{\Delta^2 \alpha_B(T)}{\Gamma_{\text{ion}}} \propto \frac{\Delta^2 \alpha_B(T)}{\epsilon_{\text{ion}} \lambda_{\text{mfp}}}$$

Neutral component

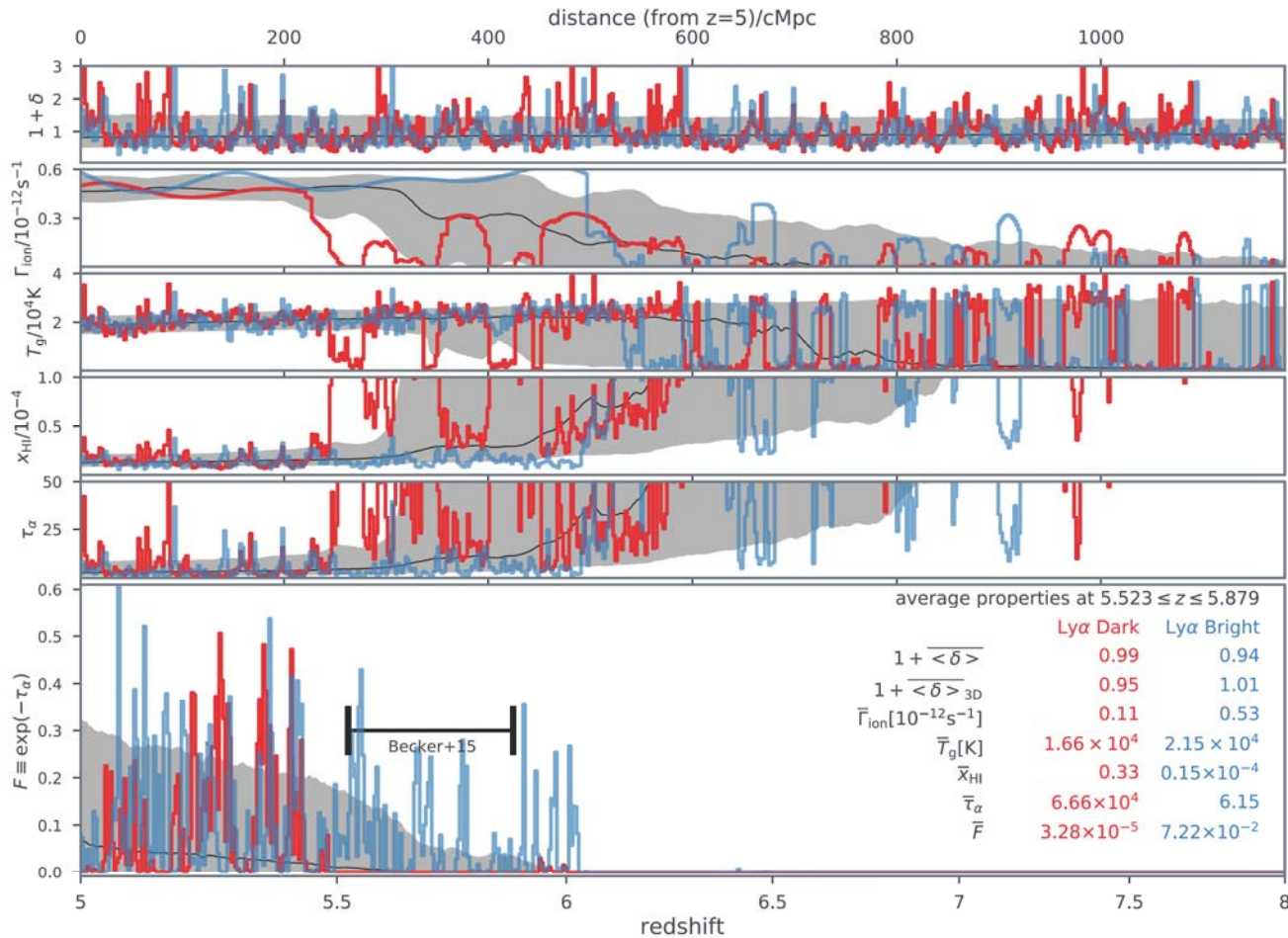
$$\tau_{\text{HI}} \sim 10^5 - 10^6$$



Large-scale fluctuations could be explained by (*tuning*):

- Temperature (e.g. D'Aloisio+2015)
- Rare sources (Chardin+2015,2017; Meiksin+2020)
- Mean free path (Davies+2016; Davies+2023)
- patchy EoR (Kulkarni+2019; Keating+2020; Choudhury+2021; Qin+2021)

# Or a combination of effects — large scale fluctuations from a combination of EoR and UVB

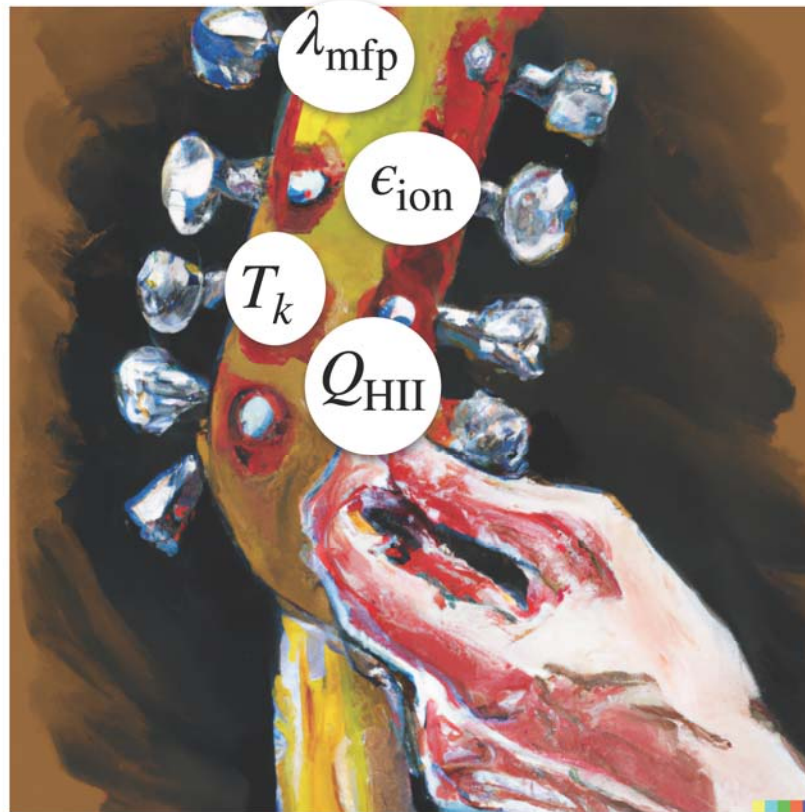


from MAP model

Qin, AM+ 2021  
see also Keating+2020



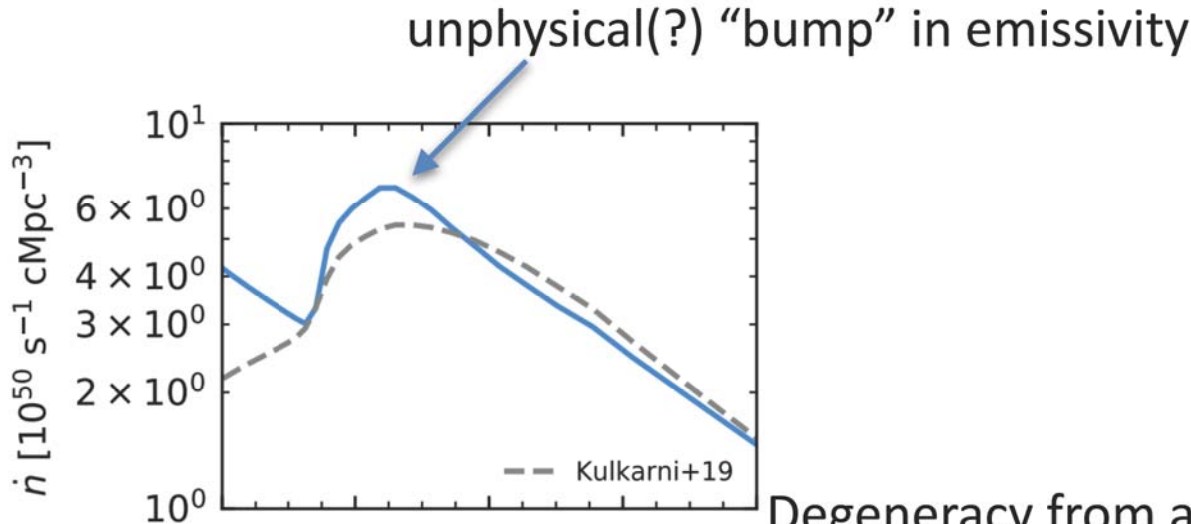
# The problem with *tuning*...



DALL-E 2 OpenAI

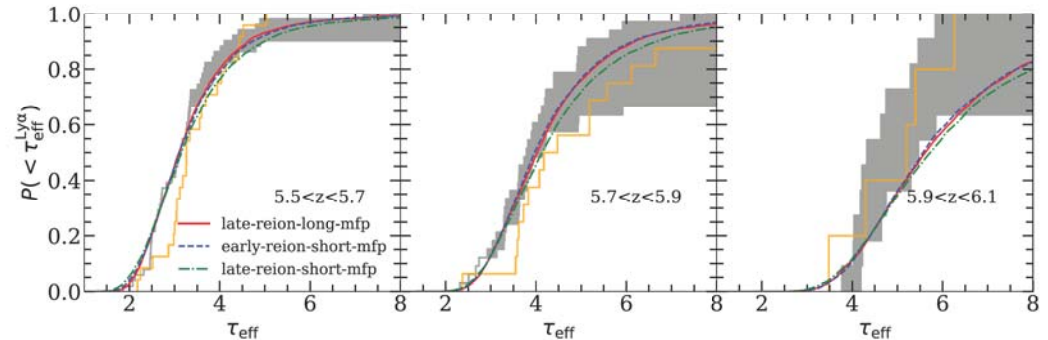
What do we learn about galaxies, if we introduce “effective” parameters and “tune”?

# The problem with *tuning*...



Kulkarni+19; Keating+20

Degeneracy from ad-hoc mpf and EoR tuning



Nasir&D’Aloisio (2019)

What do we learn about galaxies, if we introduce “effective” parameters and “tune”?

The only *correct* way to learn from data is through **Bayesian inference** from **physical forward models**

- We need efficient **simulations** with a parametric **model** for cosmology + astrophysics + systematics
- We need **observations**!
- We need a quantitative way of **comparing** the model(s) to the observations

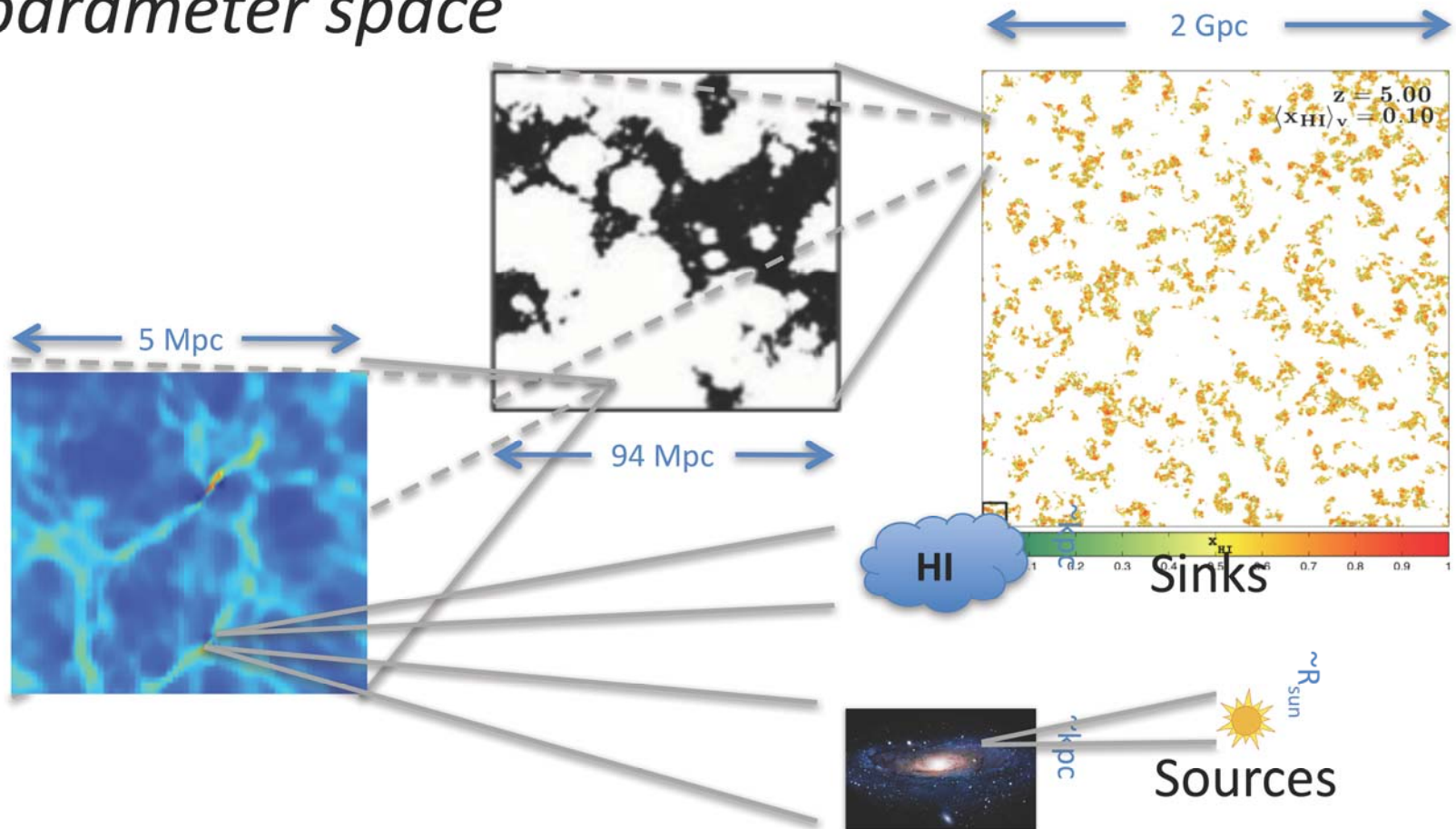
$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

Lots of quality data can really  
stress-test our models...

*Why is it so hard to model?*

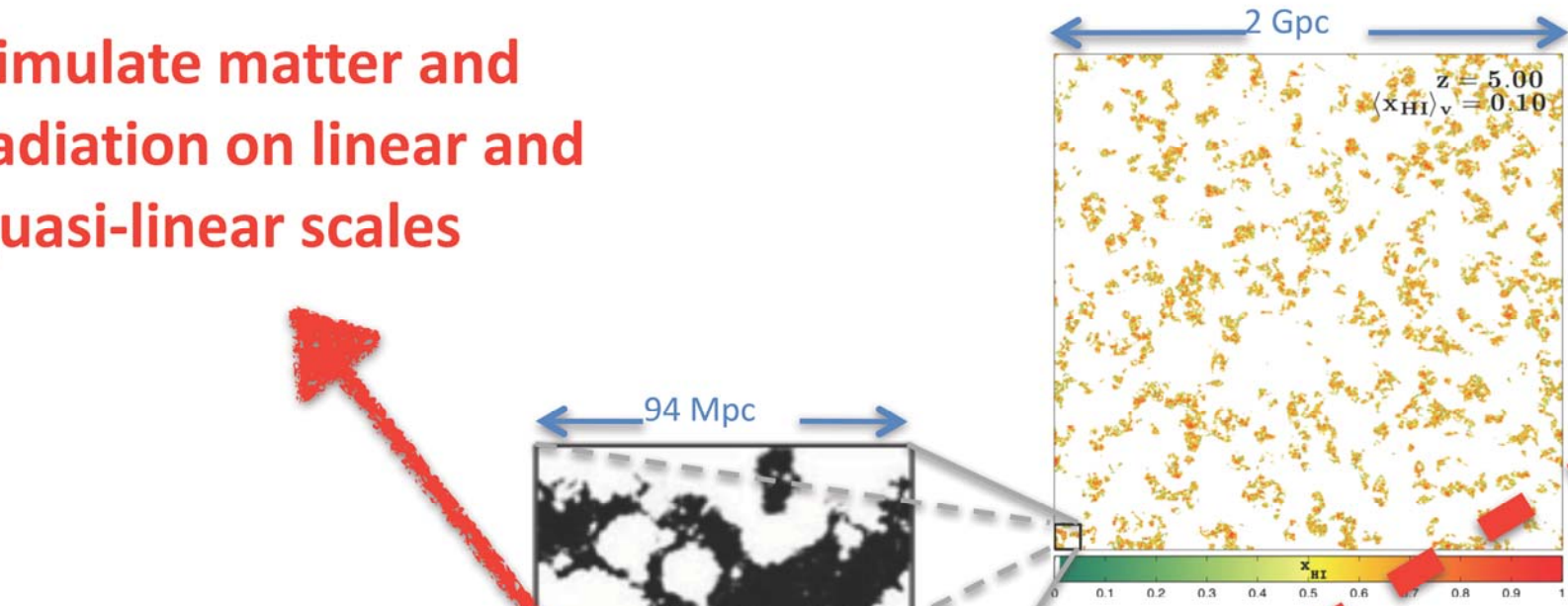
# Challenges in IGM modeling

- (i) Range of scales - *need multi-scale modeling*
- (ii) Complexity - *need efficient sampling of large parameter space*

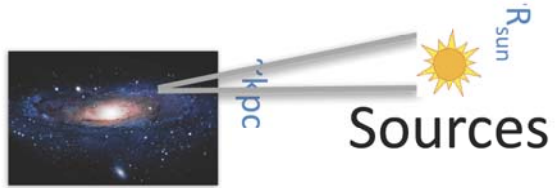


# Forward modeling the Ly $\alpha$ forest and other EoR observables

Simulate matter and radiation on linear and quasi-linear scales



characterize galaxies and sinks with sub-grid models



# The method

- i) For a given sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST

$$\frac{M_*}{M_h} = \mathbf{f}_{*,10} \left( \frac{M_h}{10^{10} M_\odot} \right)^{\alpha_*} \frac{\Omega_b}{\Omega_m}$$
$$L_{1500} \propto \frac{M_*}{\mathbf{t}_{*}^{\text{dyn}}}$$
$$L_{\text{ion}} \propto \mathbf{f}_{\text{esc},10} \left( \frac{M_h}{10^{10} M_\odot} \right)^{\alpha_{\text{esc}}} \left( \frac{1+z}{7} \right)^{\beta_{\text{esc}}} L_{1500}$$
$$f_{\text{duty}} = \exp[-\mathbf{M}_{\text{turn}}/M_h]$$

# The method

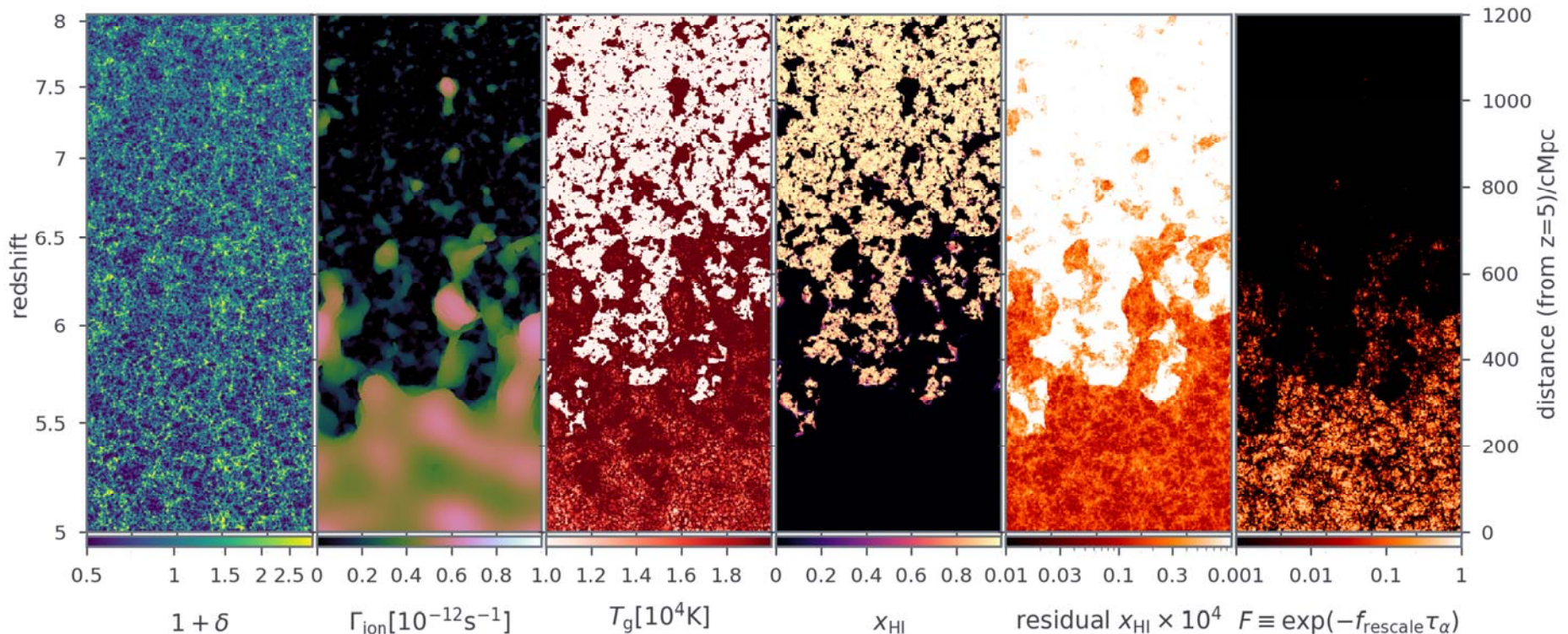
- i) For a given sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST

$$\frac{M_*}{M_h} = f_{*,10} \left( \frac{M_h}{10^{10} M_\odot} \right)^{\alpha_*} \frac{\Omega_b}{\Omega_m}$$

$$L_{1500} \propto \frac{M_*}{t_* t_{\text{dyn}}}$$

$$L_{\text{ion}} \propto f_{\text{esc},10} \left( \frac{M_h}{10^{10} M_\odot} \right)^{\alpha_{\text{esc}}} \left( \frac{1+z}{7} \right)^{\beta_{\text{esc}}} L_{1500}$$

$$f_{\text{duty}} = \exp[-M_{\text{turn}}/M_h]$$



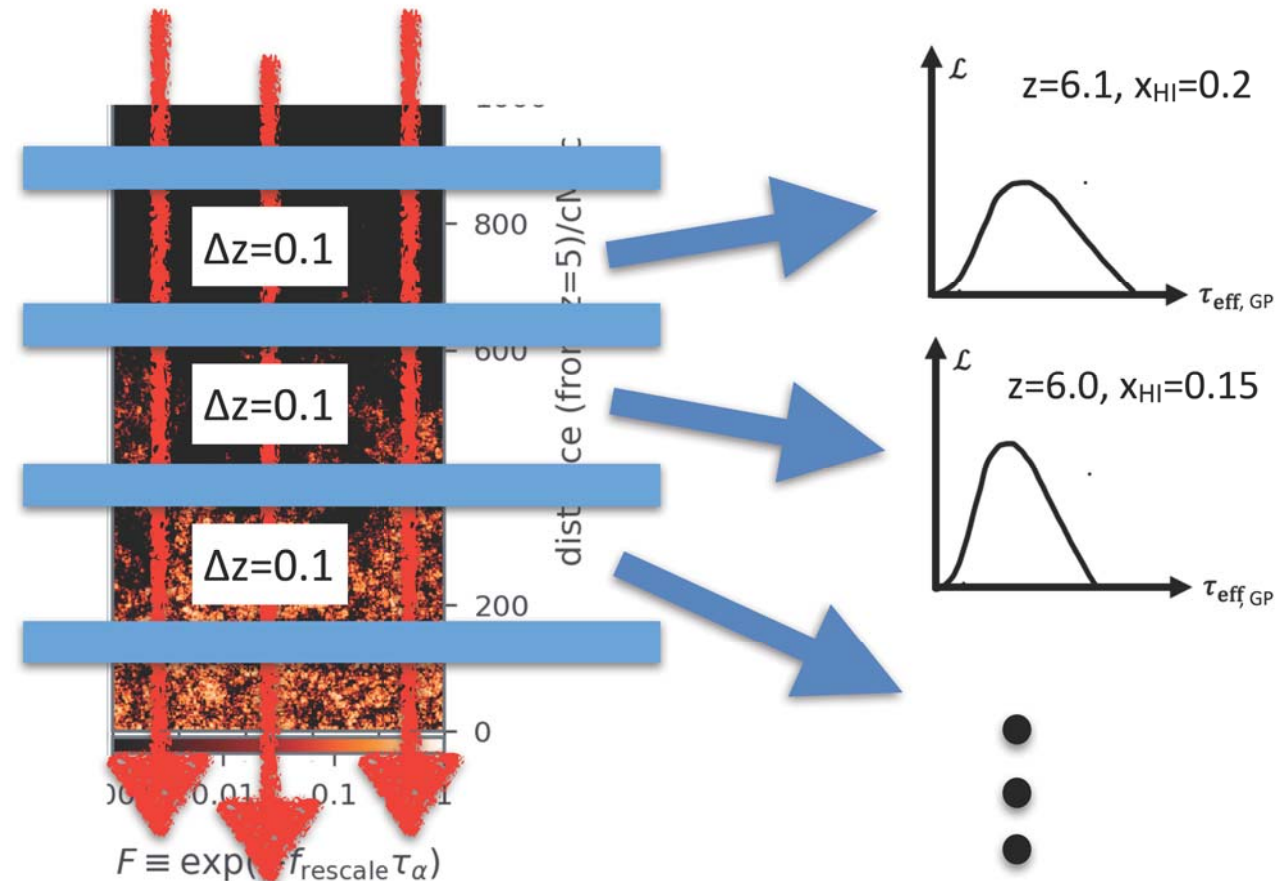


# The method

i) For a given sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST

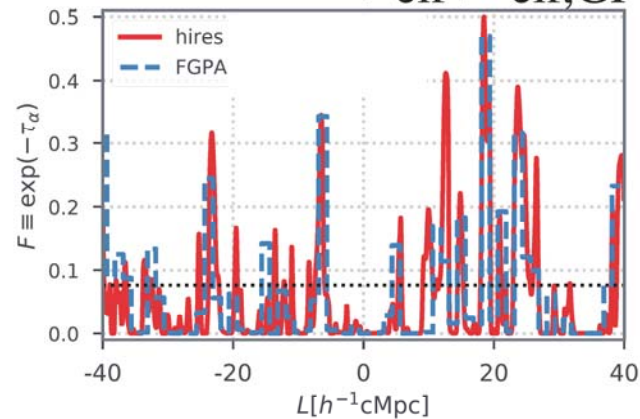
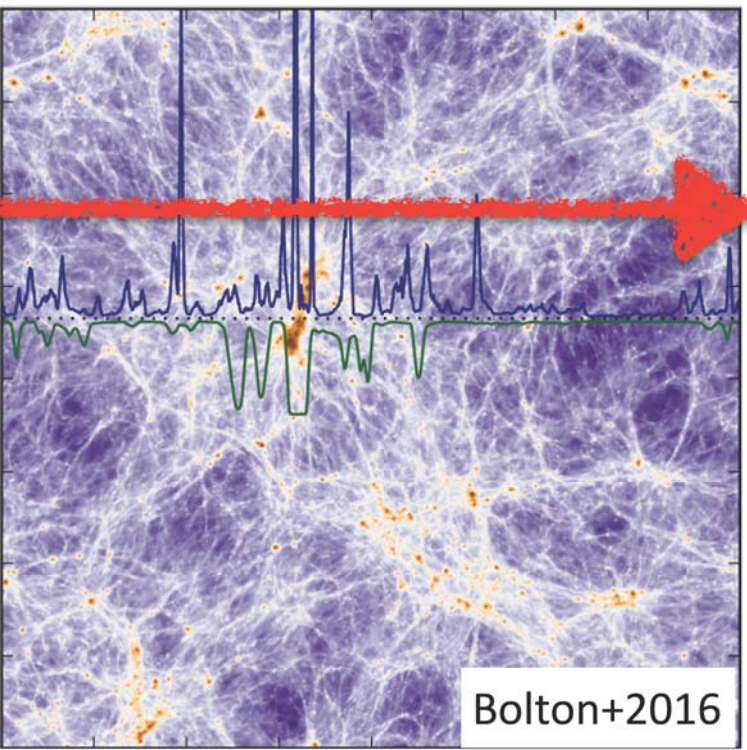
ii) Draw lines of sight and bin to obtain a distribution of

$$\tau_{\text{eff,GP}} \equiv -\ln\langle f \rangle_{\Delta z=0.1} \text{ for each redshift bin}$$



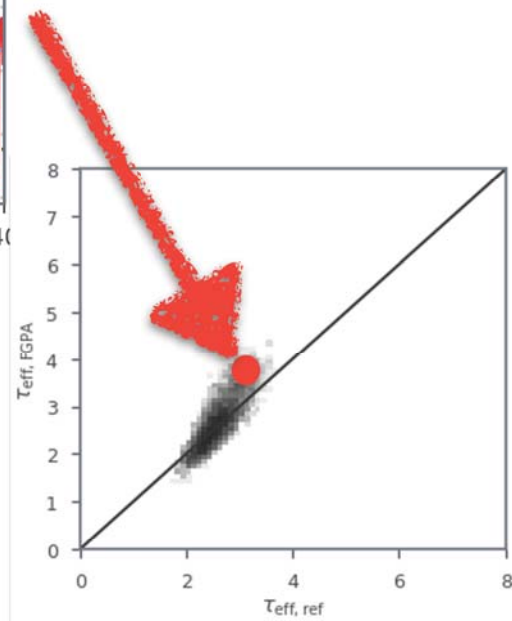
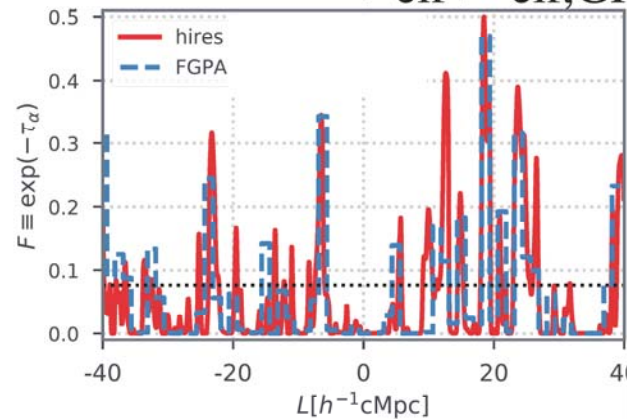
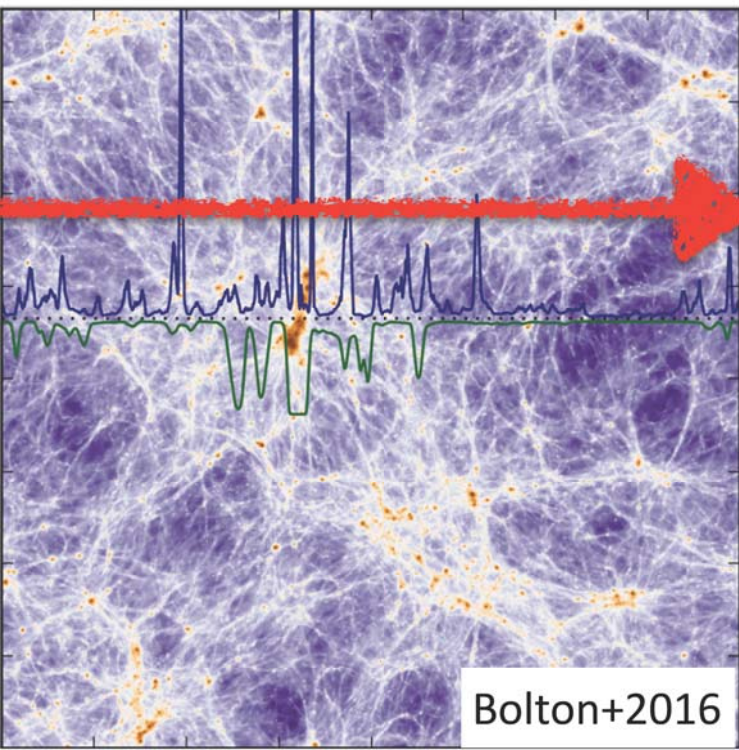
# The method

- i) For a given sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST
- ii) Draw lines of sight and bin to obtain a distribution of  $\tau_{\text{eff,GP}} \equiv -\ln\langle f \rangle_{\Delta z=0.1}$  for each redshift bin
- iii) **Account for *bias* from missing small-scales by calibrating to high-res Sherwood hydrodynamic simulations,  $P(\tau_{\text{eff}} | \tau_{\text{eff,GP}}, z, \bar{x}_{\text{HI}})$**



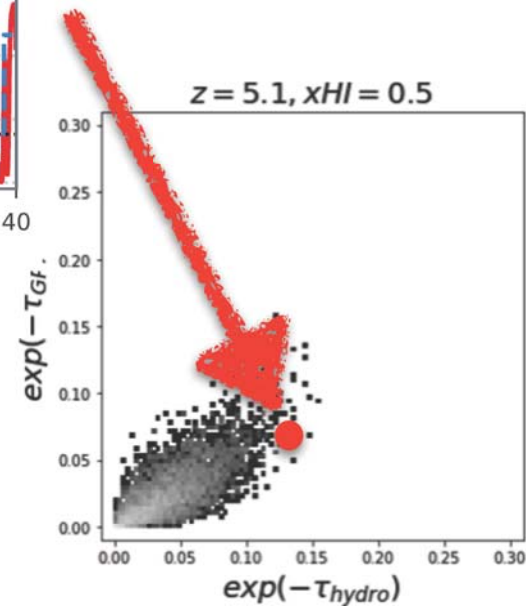
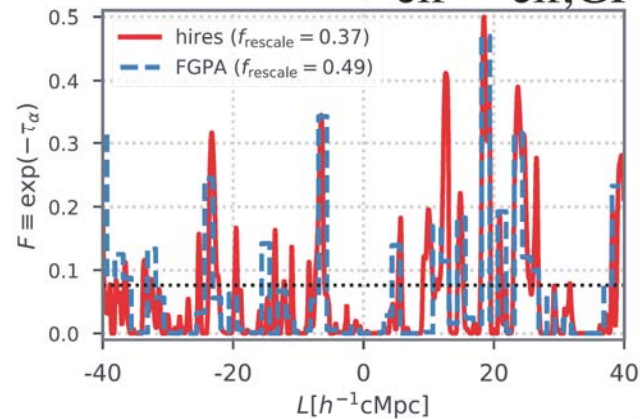
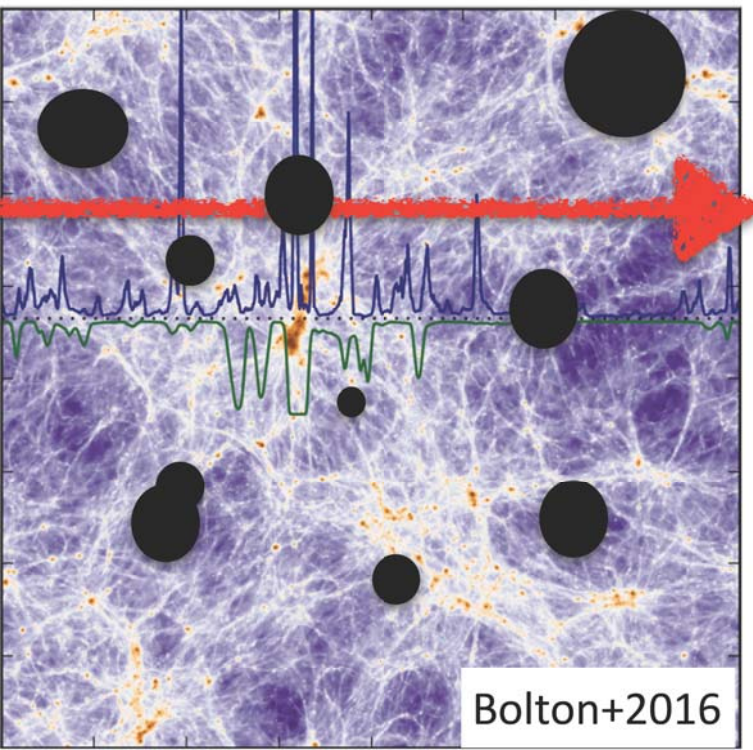
# The method

- i) For a given sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST
- ii) Draw lines of sight and bin to obtain a distribution of  $\tau_{\text{eff,GP}} \equiv -\ln\langle f \rangle_{\Delta z=0.1}$  for each redshift bin
- iii) Account for *bias* from missing small-scales by calibrating to high-res Sherwood hydrodynamic simulations,  $P(\tau_{\text{eff}} | \tau_{\text{eff,GP}}, z, \bar{x}_{\text{HI}})$



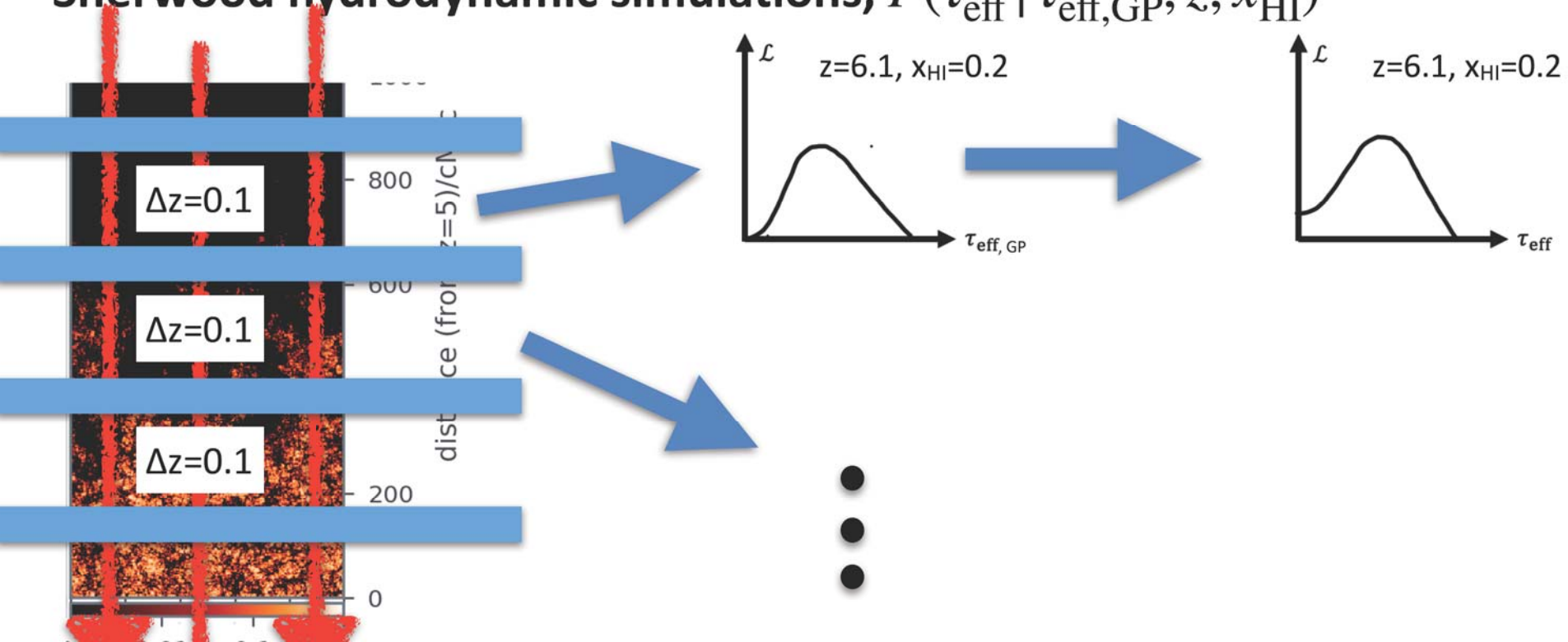
# The method

- i) For a given sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST
- ii) Draw lines of sight and bin to obtain a distribution of  $\tau_{\text{eff,GP}} \equiv -\ln\langle f \rangle_{\Delta z=0.1}$  for each redshift bin
- iii) Account for *bias* from missing small-scales by calibrating to high-res Sherwood hydrodynamic simulations,  $P(\tau_{\text{eff}} | \tau_{\text{eff,GP}}, z, \bar{x}_{\text{HI}})$



# The method

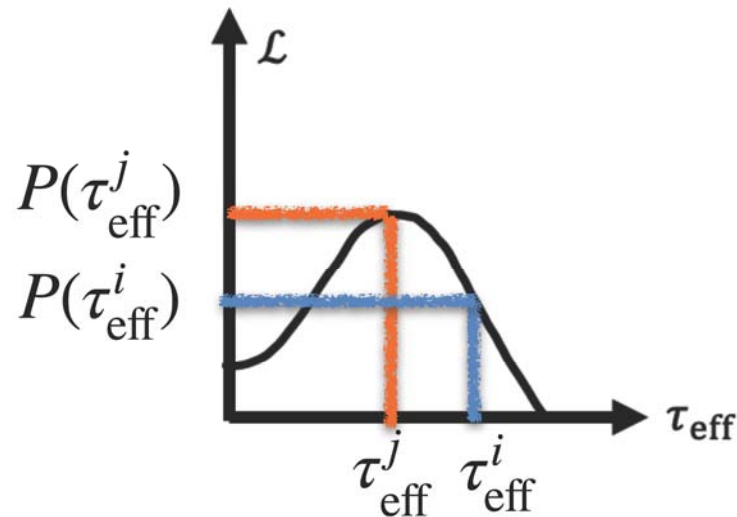
- i) For a given sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST
- ii) Draw lines of sight and bin to obtain a distribution of  $\tau_{\text{eff,GP}} \equiv -\ln\langle f \rangle_{\Delta z=0.1}$  for each redshift bin
- iii) Account for *bias* from missing small-scales by calibrating to high-res Sherwood hydrodynamic simulations,  $P(\tau_{\text{eff}} | \tau_{\text{eff,GP}}, z, \bar{x}_{\text{HI}})$



# The method

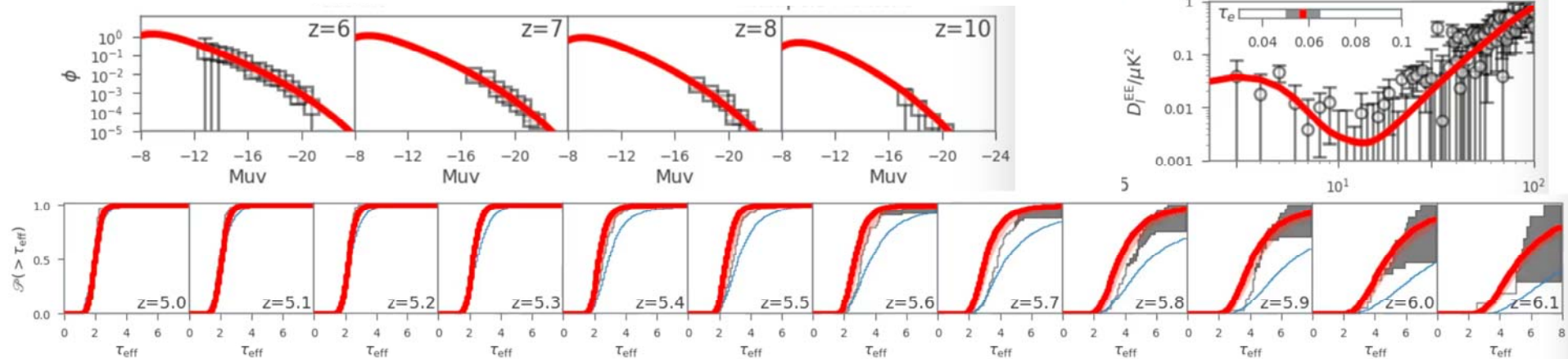
- i) For a given sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST
- ii) Draw lines of sight and bin to obtain a distribution of  $\tau_{\text{eff,GP}} \equiv -\ln\langle f \rangle_{\Delta z=0.1}$  for each redshift bin
- iii) Account for bias from missing small-scales by calibrating to high-res Sherwood hydrodynamic simulations,  $P(\tau_{\text{eff}} | \tau_{\text{eff,GP}}, z, \bar{x}_{\text{HI}})$
- iv) **Evaluate the forest likelihood. For each observed QSO,  $i$ , the likelihood is just  $P(\tau_{\text{eff}} = \tau_{\text{eff}}^i, z)$ . The total forest likelihood is then**

$$\prod_{z=5.3}^{z=6.2} \prod_{i=1}^{N_{\text{los}}(z)} \mathcal{L}_z(\tau_{\text{eff},i})$$



# The method

- i) For a given sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST
- ii) Draw lines of sight and bin to obtain a distribution of  $\tau_{\text{eff,GP}} \equiv -\ln\langle f \rangle_{\Delta z=0.1}$  for each redshift bin
- iii) Account for bias from missing small-scales by calibrating to high-res Sherwood hydrodynamic simulations,  $P(\tau_{\text{eff}} | \tau_{\text{eff,GP}}, z, \bar{x}_{\text{HI}})$
- iv) Evaluate the forest likelihood.
- v) **Compute the total likelihood, including also: (i) HST UV LFs at  $z=6,7,8,10$ ; (ii) Planck  $\tau_e$**

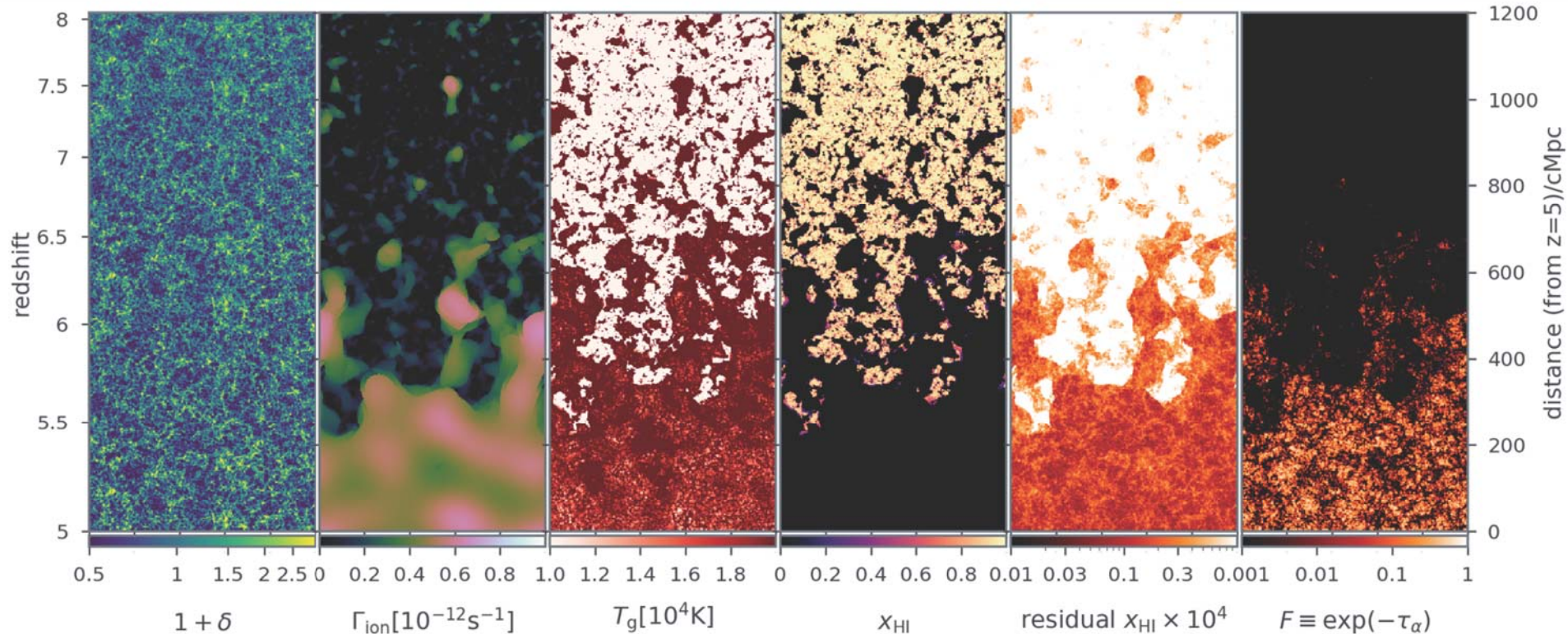


# The method

- i) For a **NEW** sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST

*Repeat sampling  $\sim 100k$  times with 21cmMC to build a posterior*

*Lightcone image from Qin, AM+ (2021)*



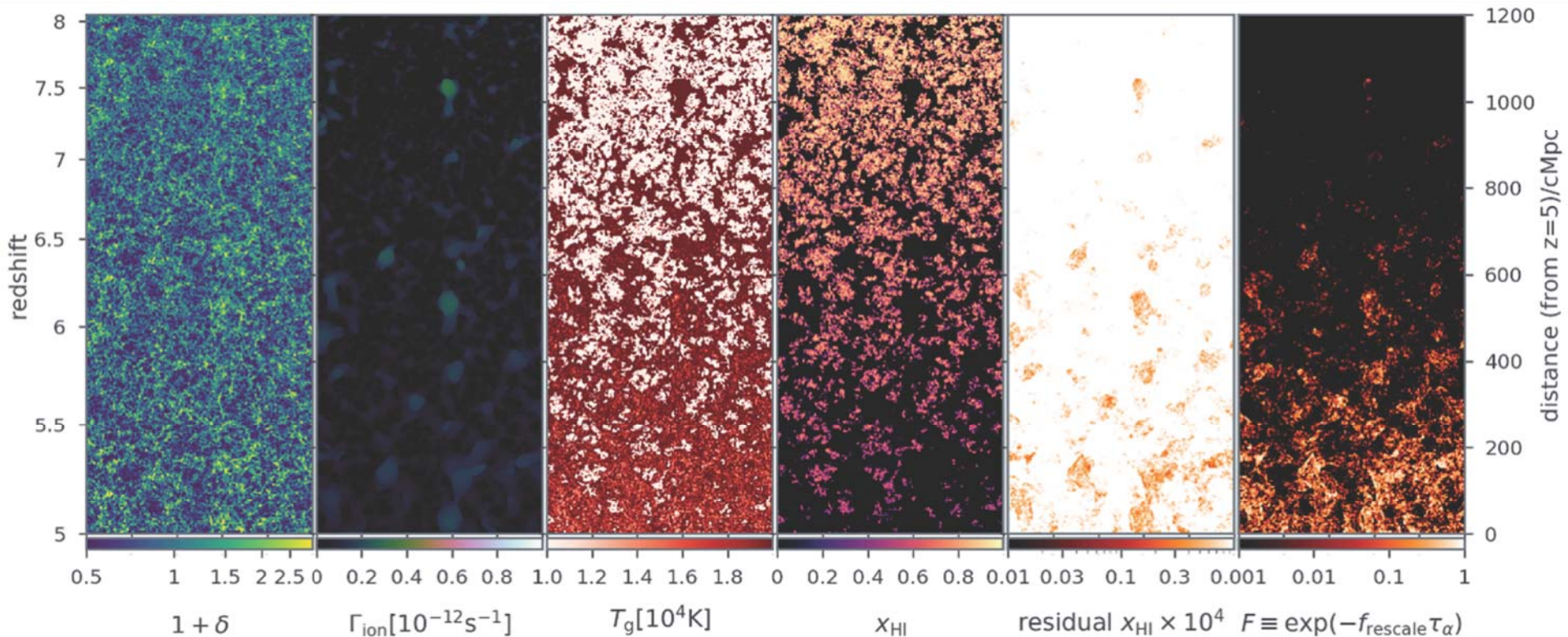


# The method

- i) For a **NEW** sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST

*Repeat sampling  $\sim 100k$  times with 21cmMC to build a posterior*

*Lightcone image from Qin, AM+ (2021)*

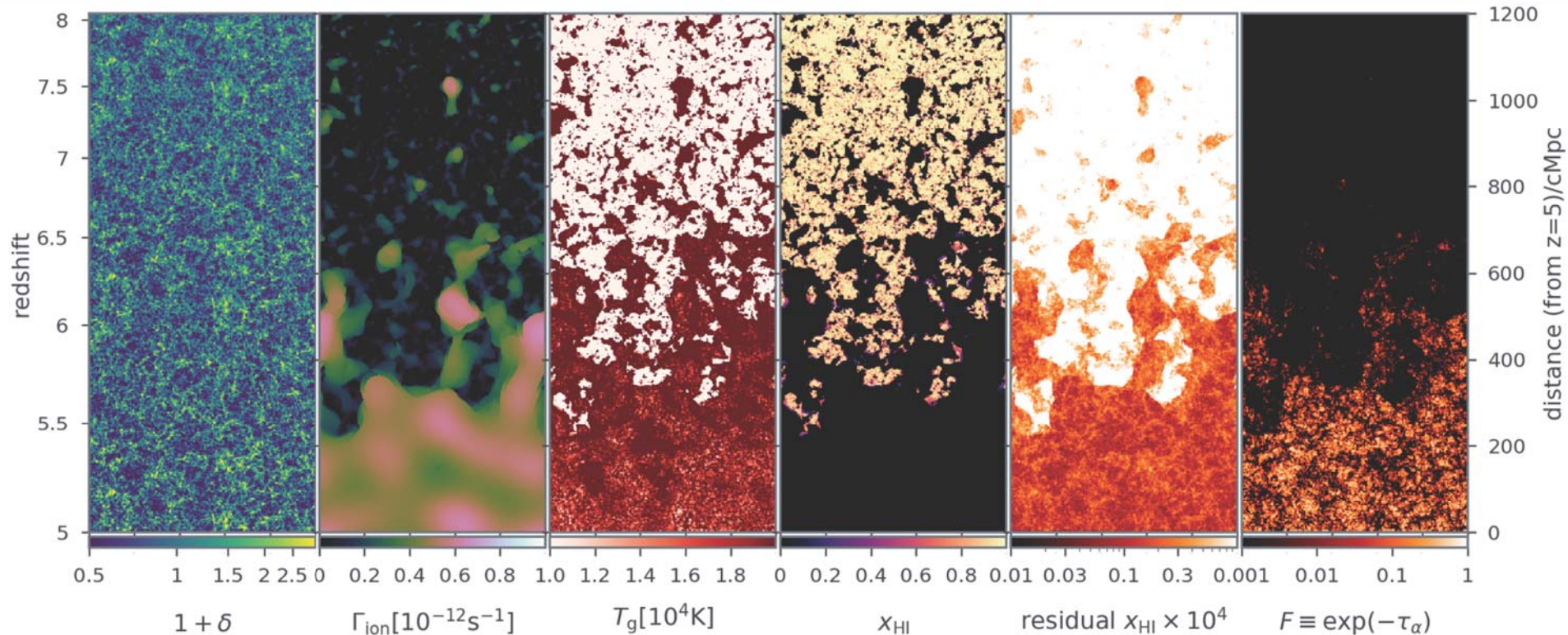


# The method

- i) For a **NEW** sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST

*Repeat sampling  $\sim 100k$  times with 21cmMC to build a posterior*

*Lightcone image from Qin, AM+ (2021)*

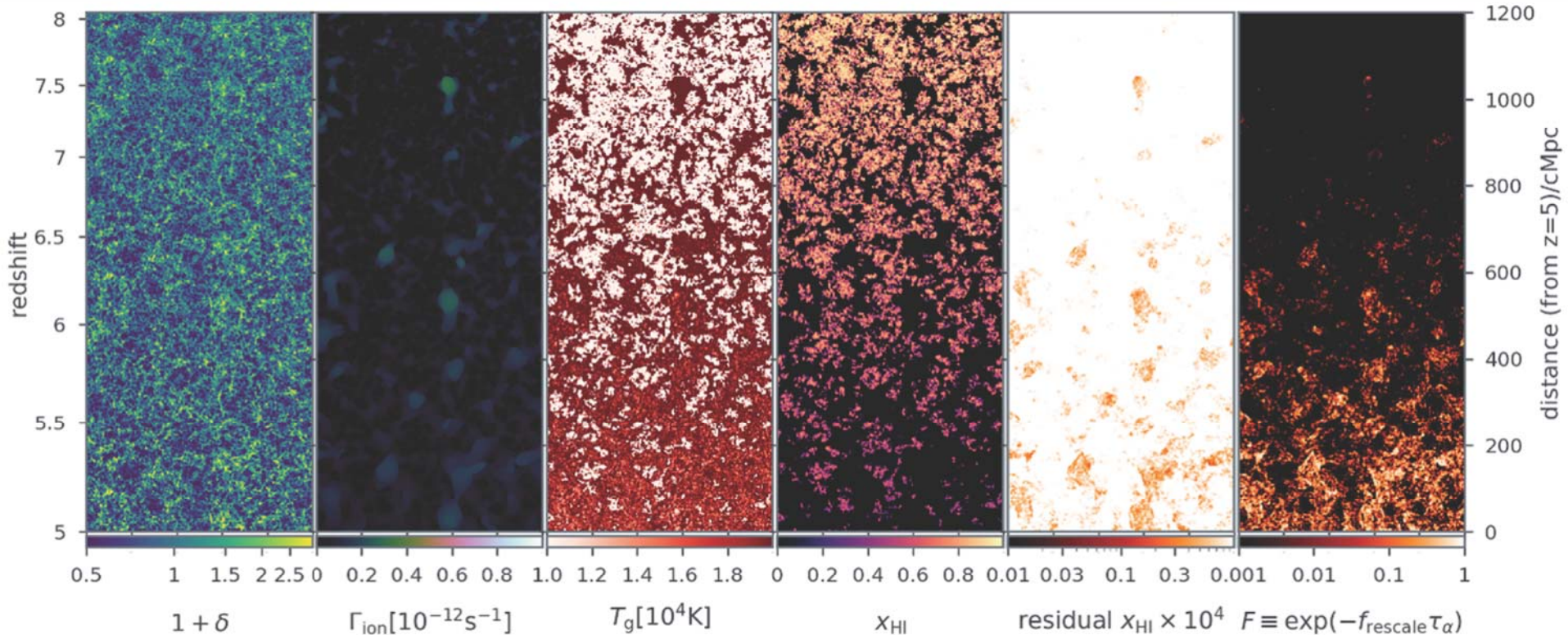


# The method

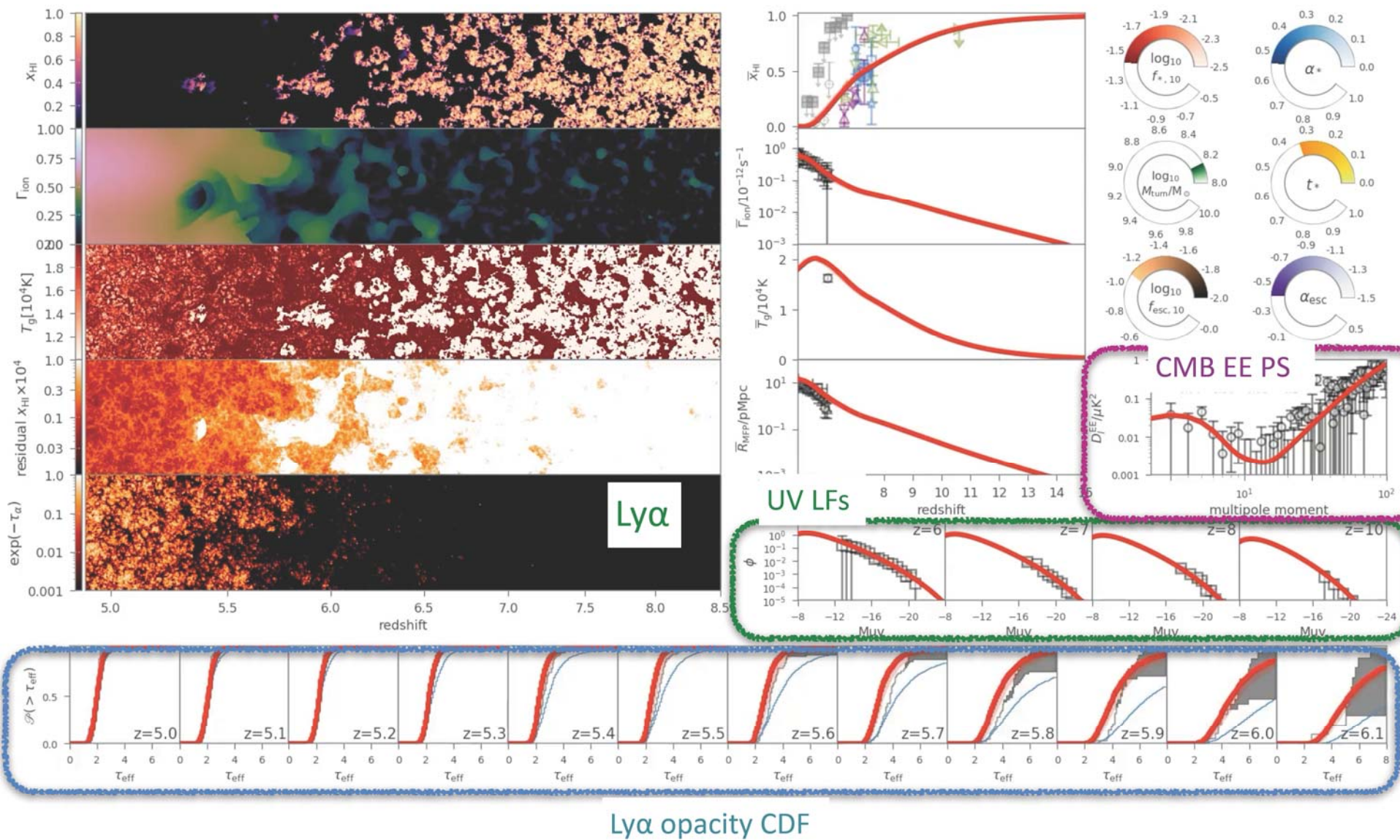
- i) For a **NEW** sample of galaxy parameters, simulate light-cones of Ly $\alpha$  optical depth with 21cmFAST

*Repeat sampling  $\sim 100k$  times with 21cmMC to build a posterior*

*Lightcone image from Qin, AM+ (2021)*

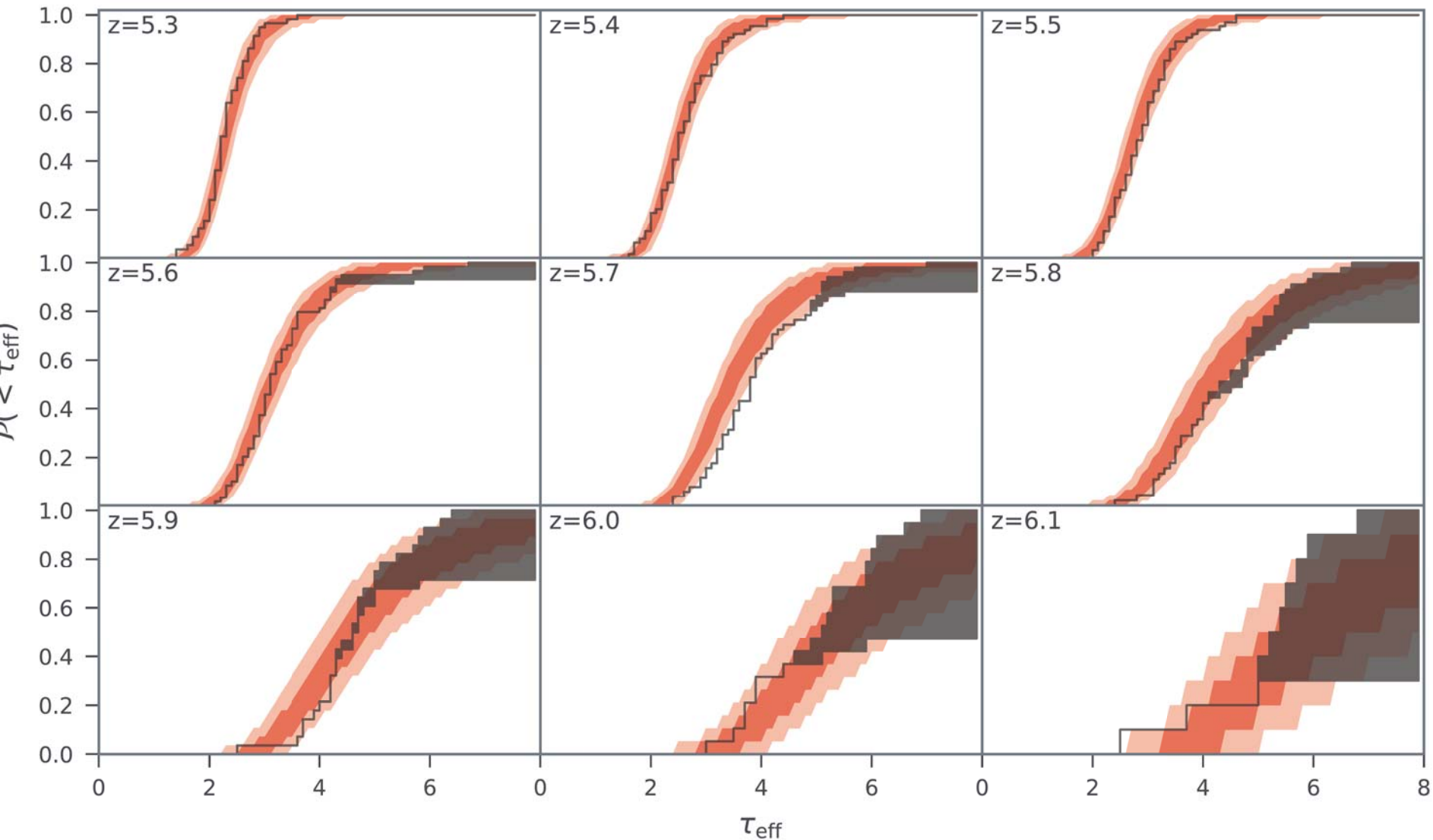


# Forward modeling: Ly $\alpha$ forest + UV LFs + CMB $\tau$



# Results (*preliminary!*)

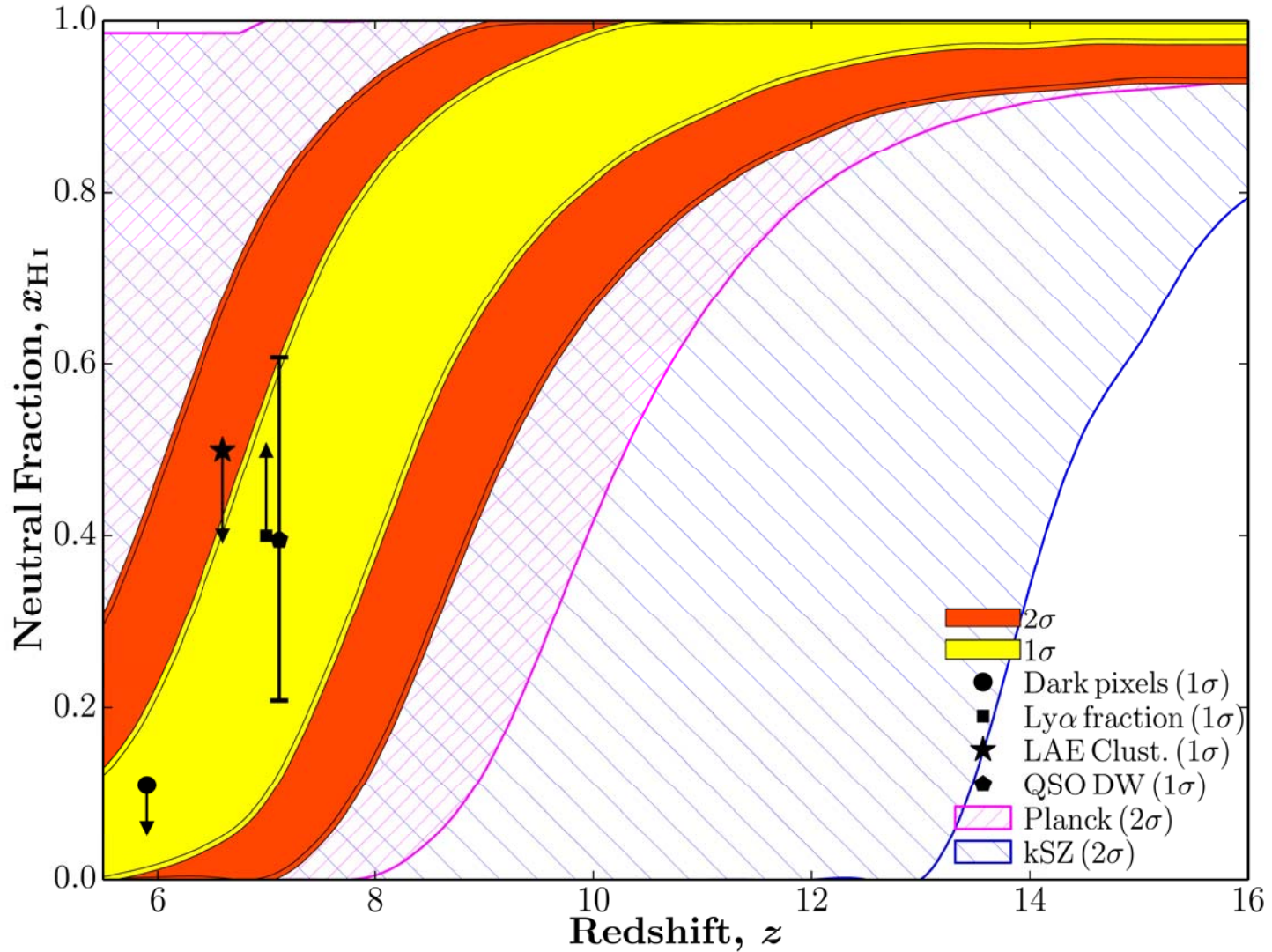
# Forest $\tau_{\text{eff}}$ distributions



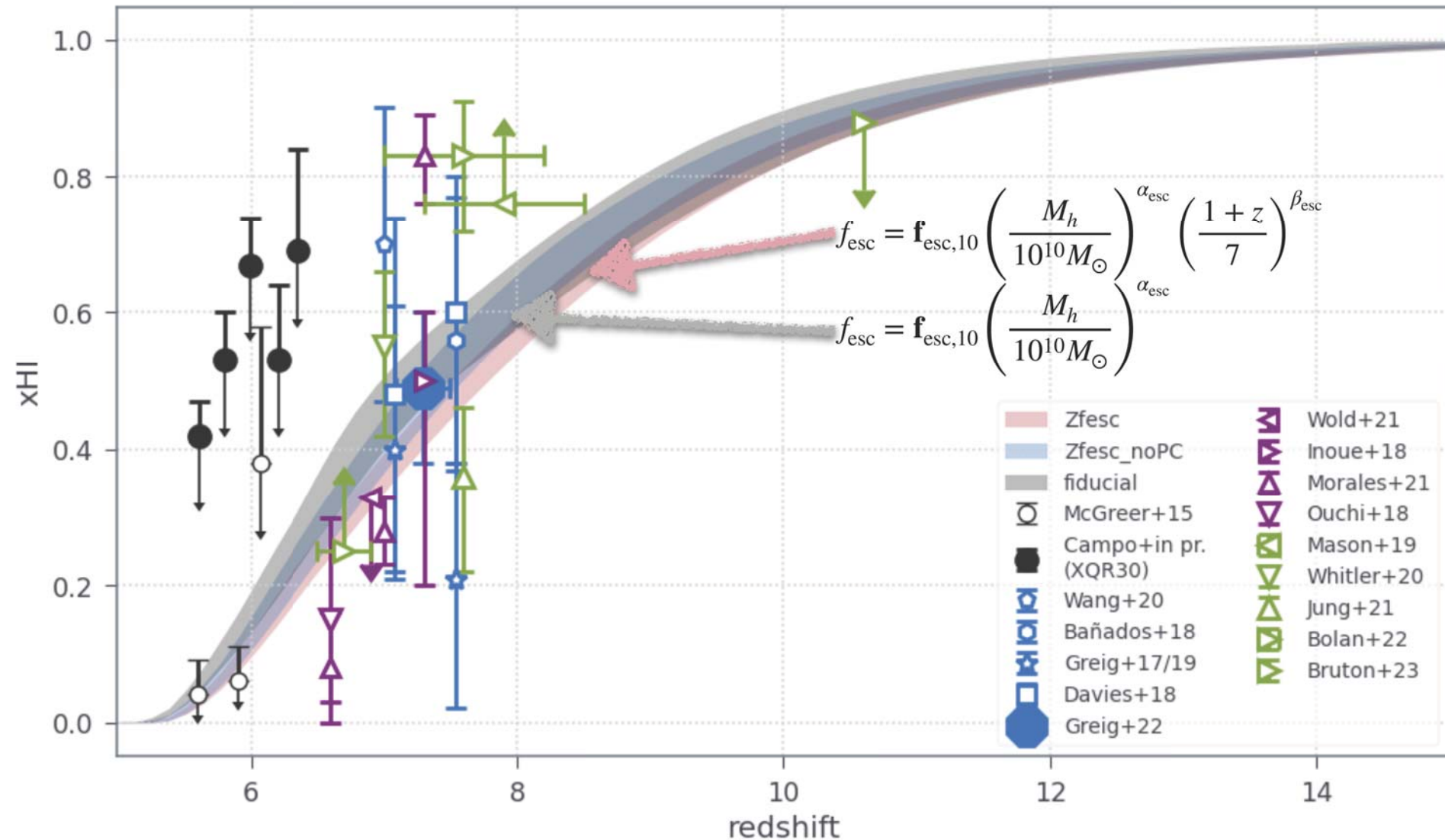
**Joint fit over all redshift bins!** no recalibration at each redshift, hyperparameters, removing mean flux, ad-hoc tuning / effective parameters, artificially scaling mfp or emissivity vs  $z$ , etc.

# Until recently, the timing of the EoR was fairly uncertain...

Greig & AM (2017)



# Inferred EoR history

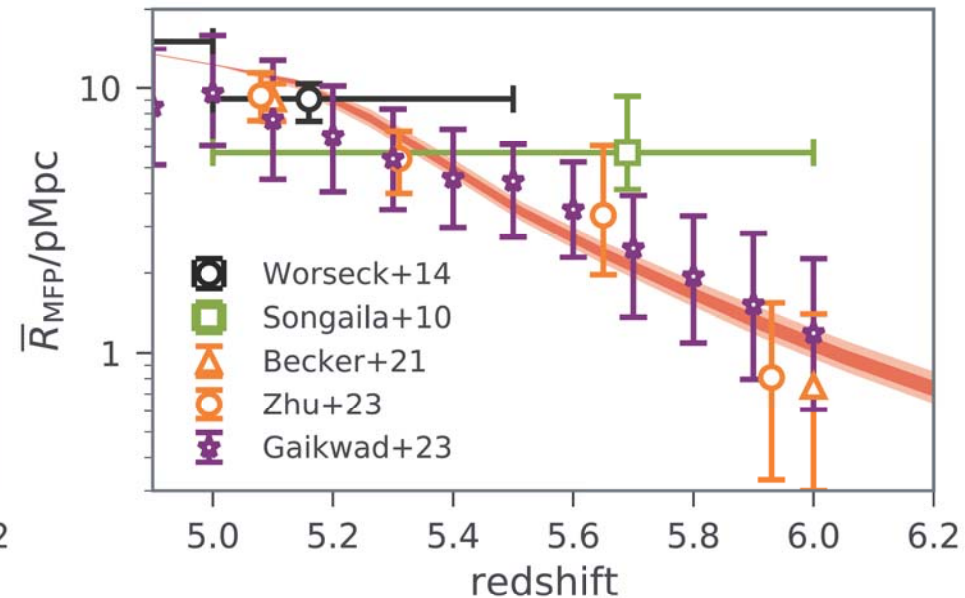
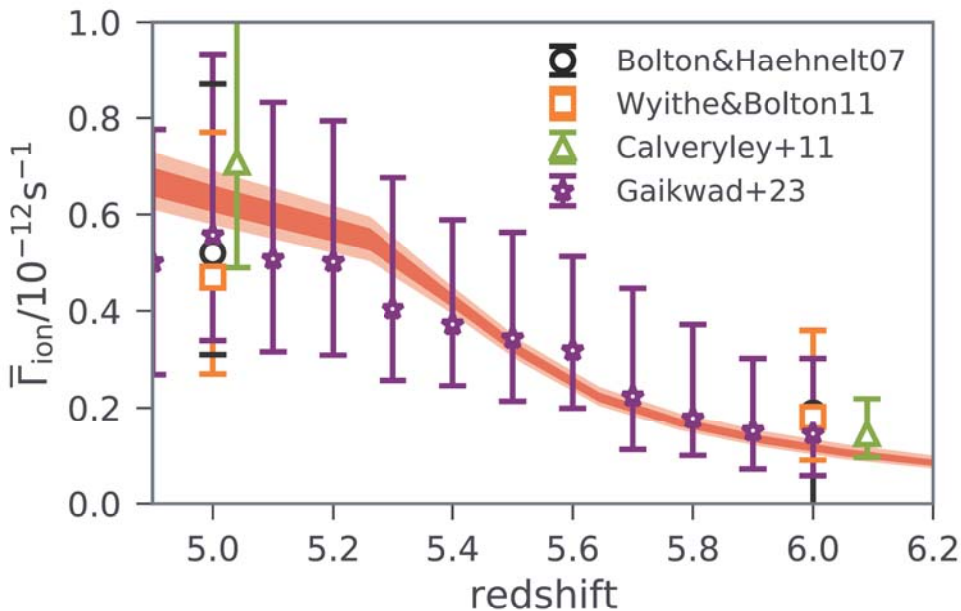


*None of these data points are used in the likelihood!*

Qin, AM+ in prep



# Average IGM properties

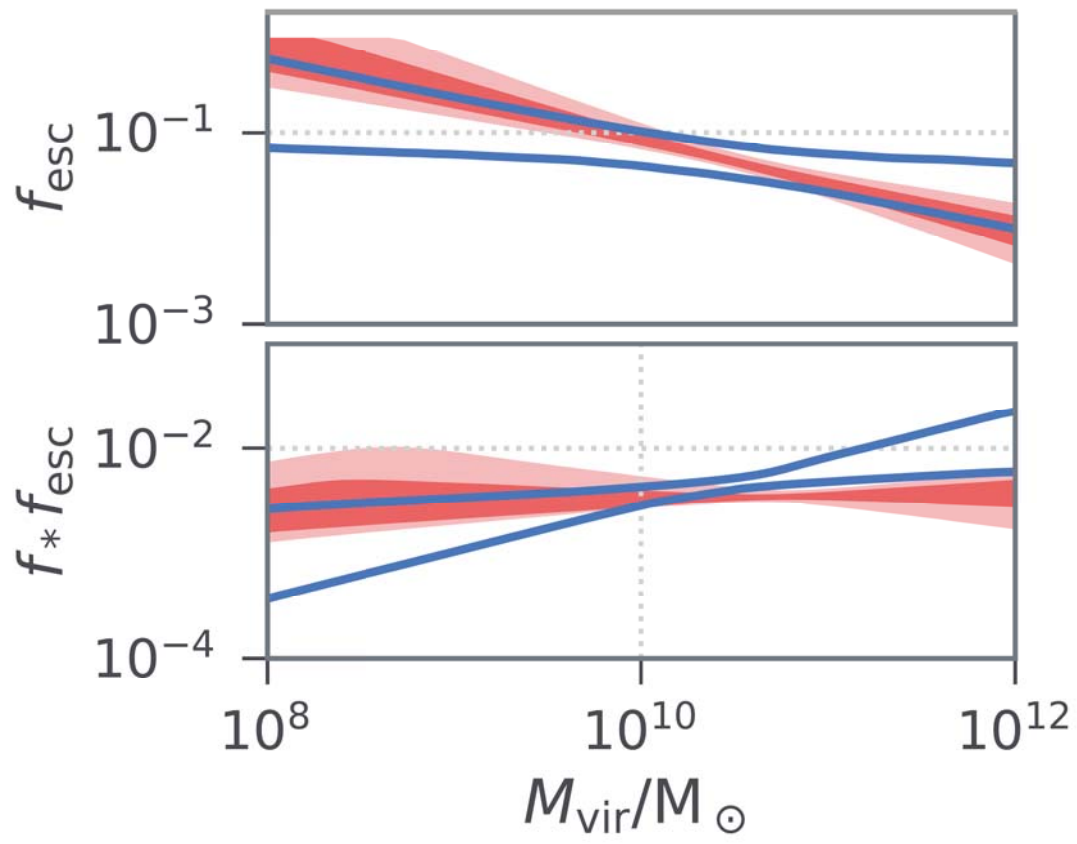


Qin, AM+ in prep

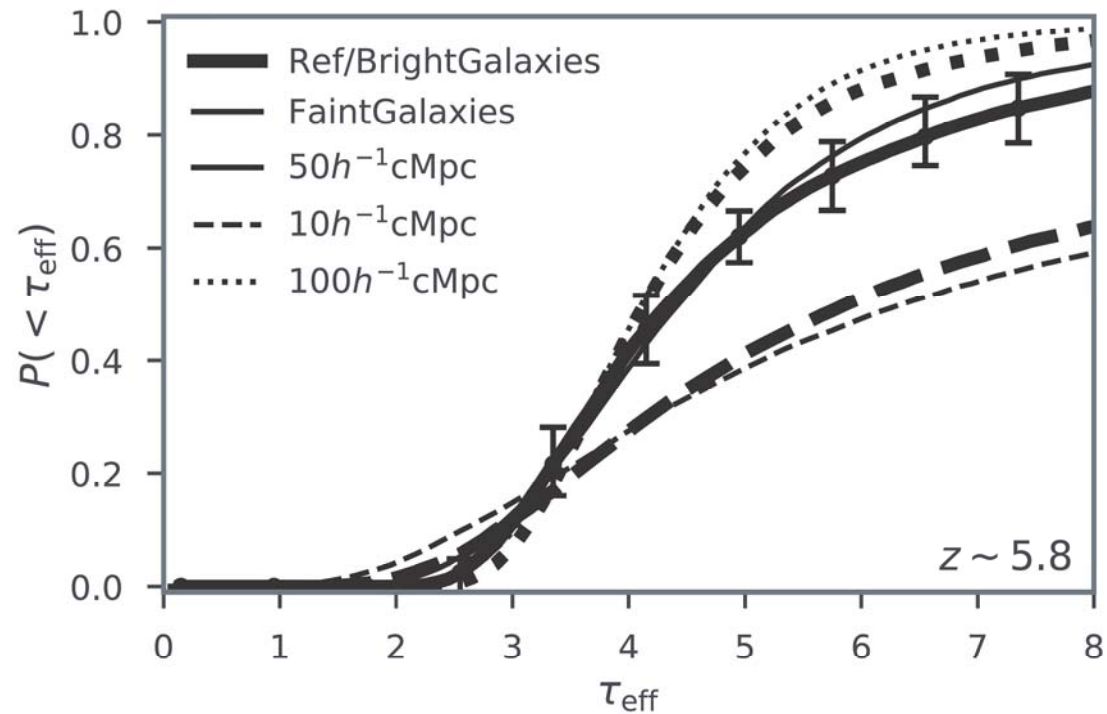
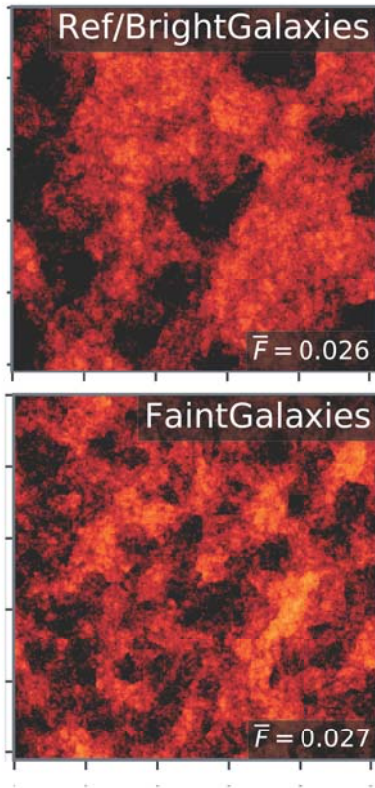
*None of these data points are used in the likelihood!*

# Conclusions about galaxy properties are model dependent however...

PRELIMINARY



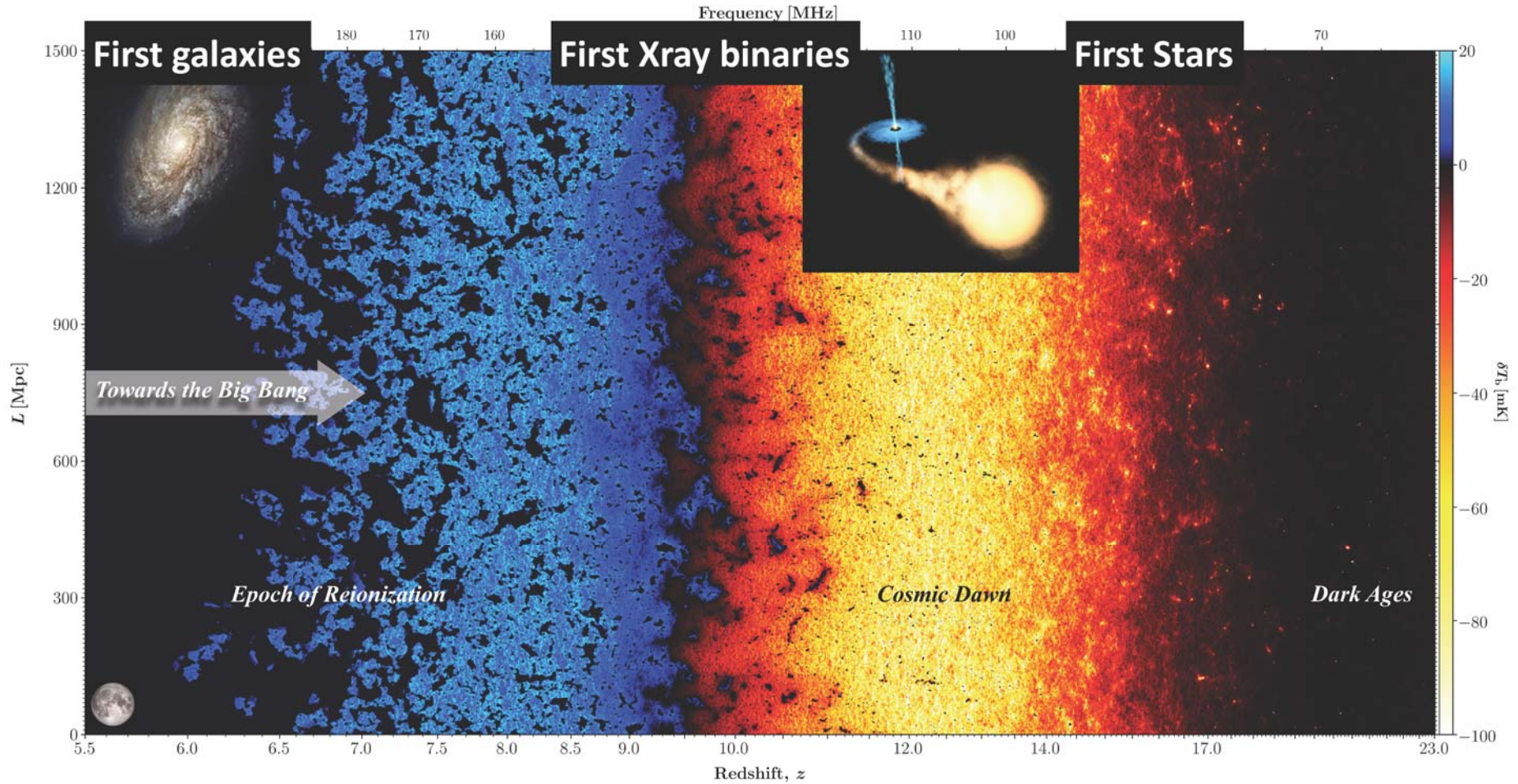
# The Ly $\alpha$ forest does not have the dynamic range to probe EoR morphology



Qin+2021

**We need other probes to understand morphology!**

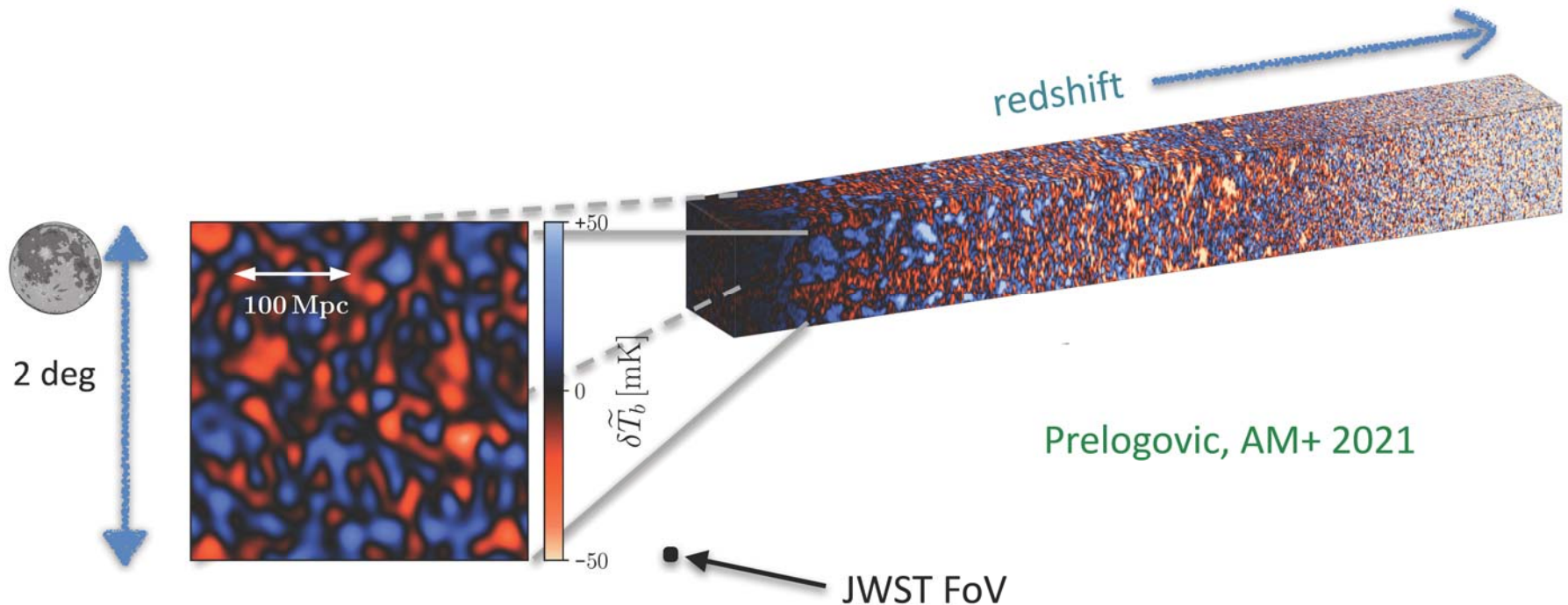
# We will have reionization topology with the 21cm signal with SKA



$$\delta T_b(\nu) \approx 27 \chi_{\text{HI}} (1 + \delta_{\text{nl}}) \left( \frac{H}{dv_r/dr + H} \right) \left( 1 - \frac{T_\gamma}{T_S} \right) \left( \frac{1+z}{10} \frac{0.15}{\Omega_M h^2} \right)^{1/2} \left( \frac{\Omega_b h^2}{0.023} \right) \text{mK}$$

# However, this is on large scales...

And is unlikely to come for another ~decade at least...



Prelogovic, AM+ 2021

single frequency slice with  
noise and foreground excision

# It would be great to study topology sooner and also on smaller scales...

zoom-in comparable to JWST FoV

Ionization morphology allows us to:

- determine **which galaxies drive reionization**
- connect the **growth of individual cosmic HII regions to the properties of galaxies** inside them
- connect the JWST-detectable galaxies to the **thousands of surrounding galaxies too faint to detect**

## Cosmic Dawn III

P. OCVIK, Observatoire astronomique de Strasbourg  
Cosmic Dawn & CLUES collaborations  
Summit / Oak Ridge Leadership Computing Facility

$z=25.0$

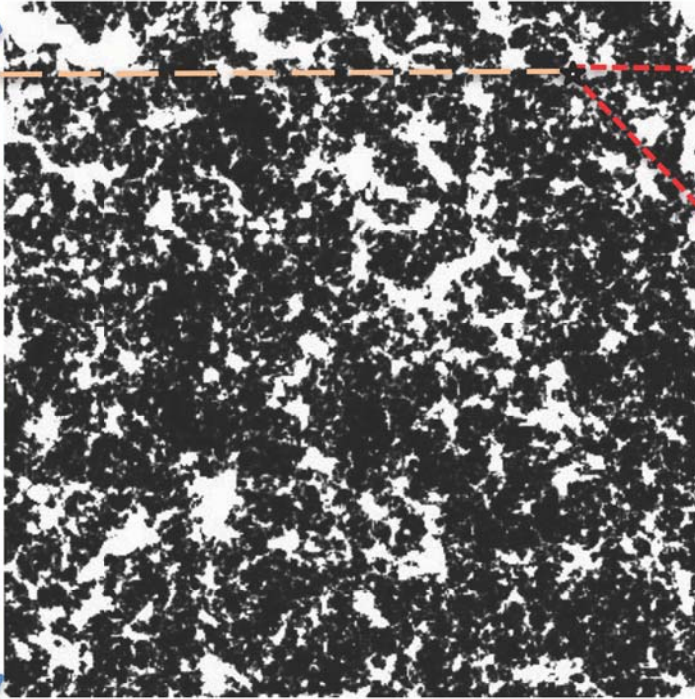
8 cMpc/h sub-region

CODAIII, credit: P. Ocvirk

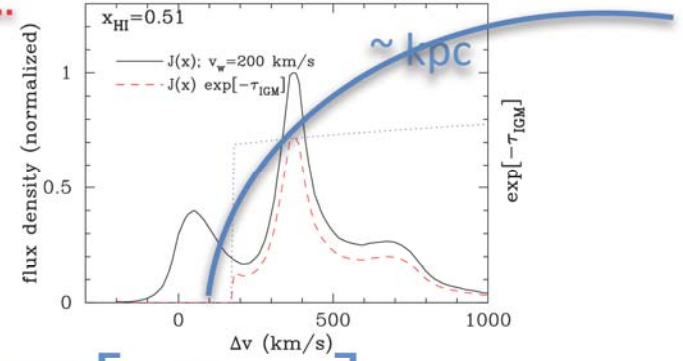
# Lyman alpha from galaxies is a great tool to study EoR topology NOW



300 Mpc



AM+2015



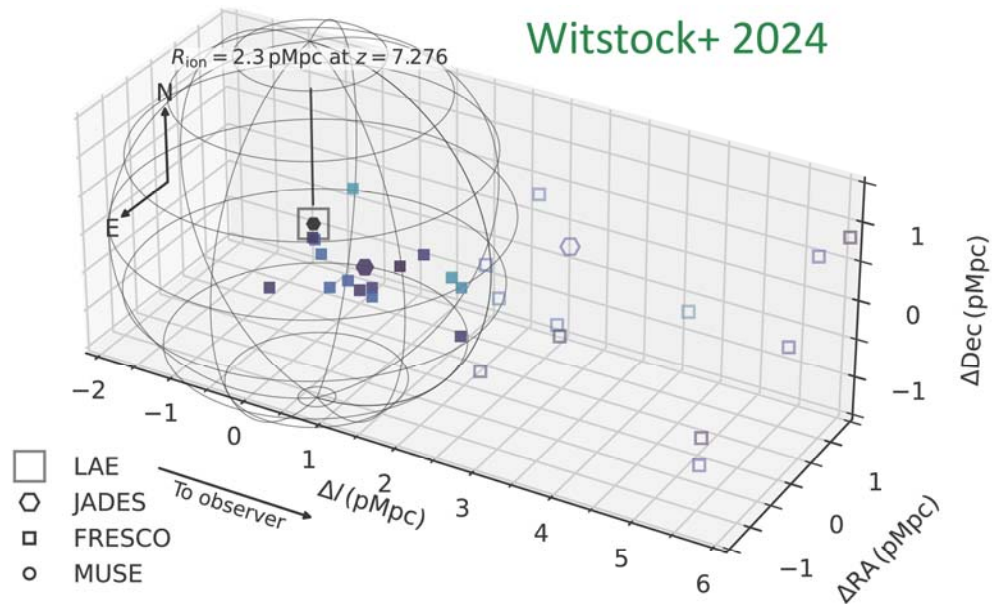
$\exp[-\tau_{\text{reion}}]$

during reionization, **cosmic HI patches** absorb Ly $\alpha$  photons in the damping wing of the line

Ly $\alpha$  damping wing absorption is an especially powerful probe of the **EARLY STAGES OF THE EOR**

# Exciting recent observations of galaxy groups

e.g. Tilvi+20; Endsley & Stark 22; Jung+22; Saxena+23; Whitler+23; Hayes & Scarlatta 23; Umeda+23; Witstock+24





# Exciting recent observations of galaxy groups

e.g. Tilvi+20; Endsley & Stark 22; Jung+22; Saxena+23; Whitley+23; Hayes & Scarlatta 23; Umeda+23; Witstock+24

but...

## Analysis of surrounding HII topology is very approximate / qualitative...

- typically observed galaxies are treated **individually?!?** (*HII regions come from the cumulative radiation of **thousands of fainter galaxies***)
- assume **uniform reionization?!?** (*reionization is patchy -> **scatter and bias** in estimates Mesinger & Furlanetto 08*)
- ignores or simplifies many sources of **stochasticity** when assuming **intrinsic emission?!?**

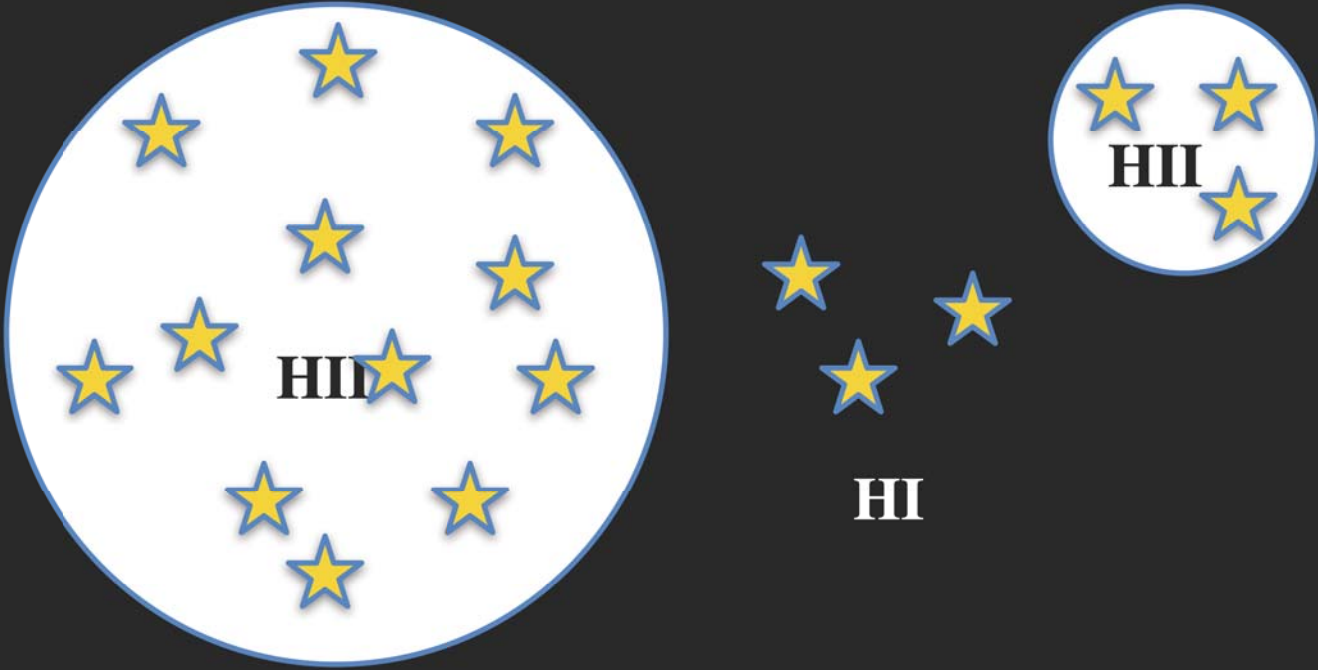
# New frameworks for studying EoR topology using galaxy “groups”



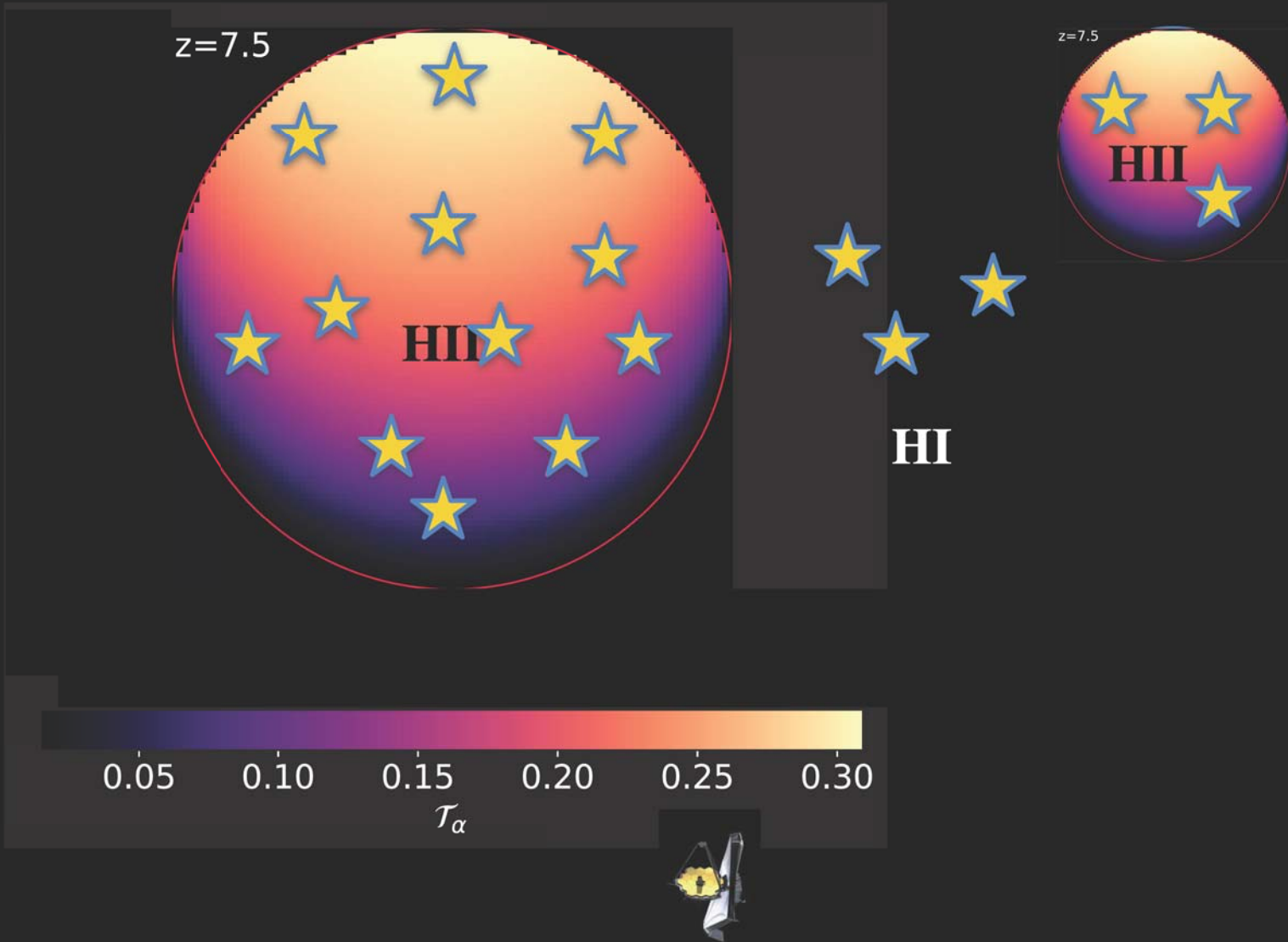
in collaboration with **Ivan Nikolić** (SNS), **Ting-Yi Lu** (DAWN), Charlotte Mason (DAWN)



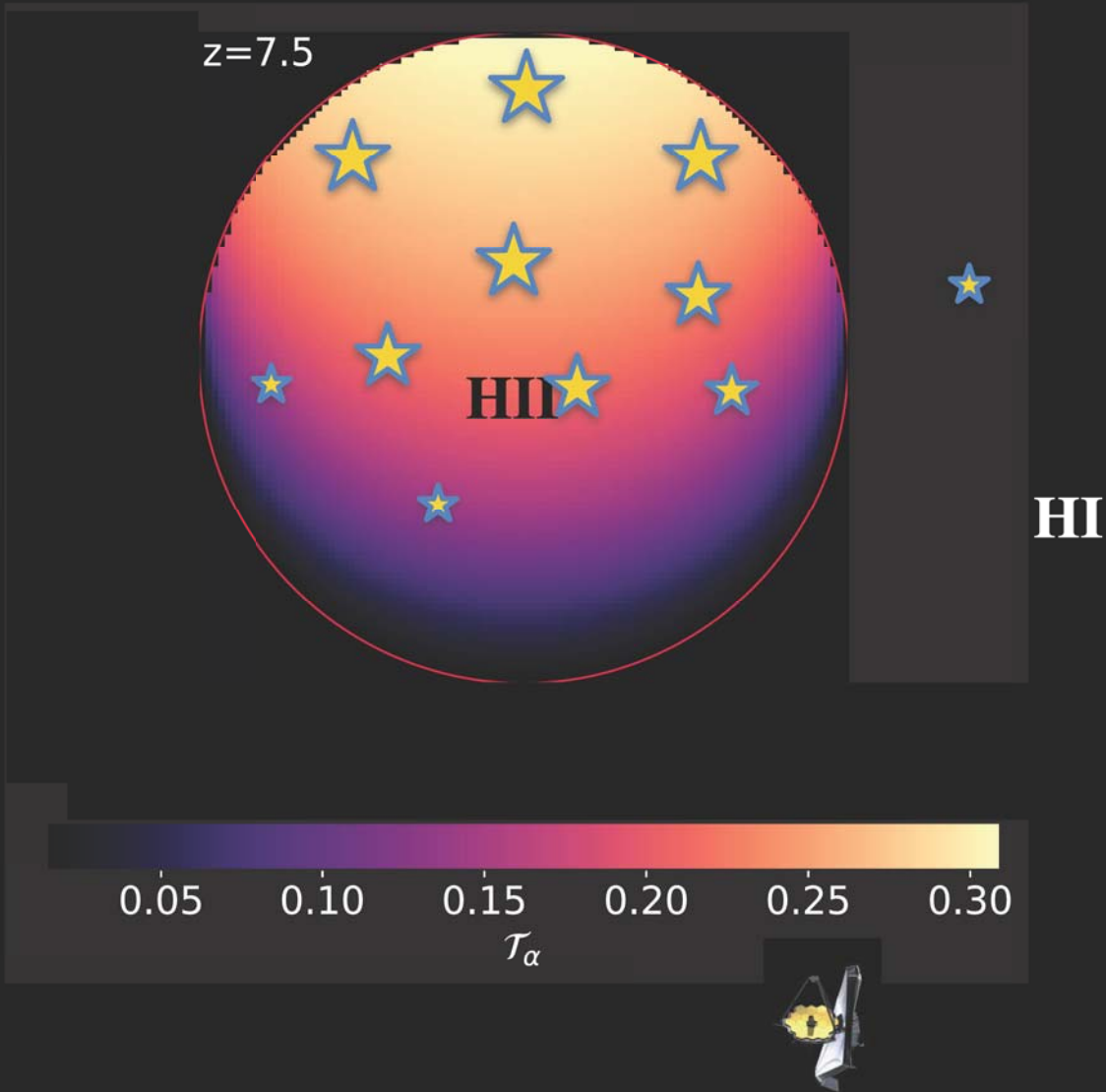
# Combining information from individual galaxies



# Combining information from individual galaxies



# Combining information from individual galaxies



# Combining information from individual galaxies



# Combining information from individual galaxies



# Inferring bubbles around galaxies

**GOAL:** Infer the position and size of an HII region, given galaxy observations

$$P(\mathcal{O}, R_b | \mathbf{x}^i, f_\alpha^i(\lambda), M_{\text{uv}}^i, z)$$

## Forward model sources of stochasticity:

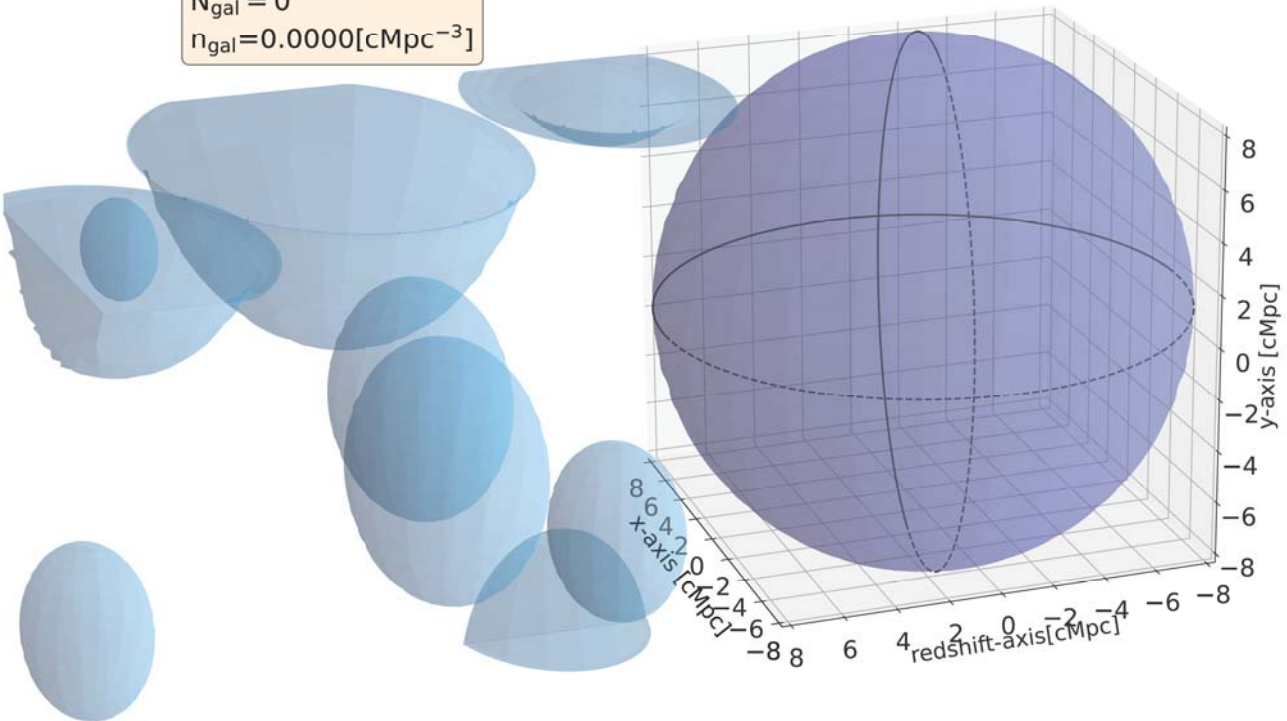
- global neutral fraction
- surrounding patchy EoR topology
- galaxy location
- Muv
- Ly $\alpha$  intrinsic flux ( $\Delta v$ , EW)
- NIRSPEC noise



# Inferring individual bubbles around galaxy groups

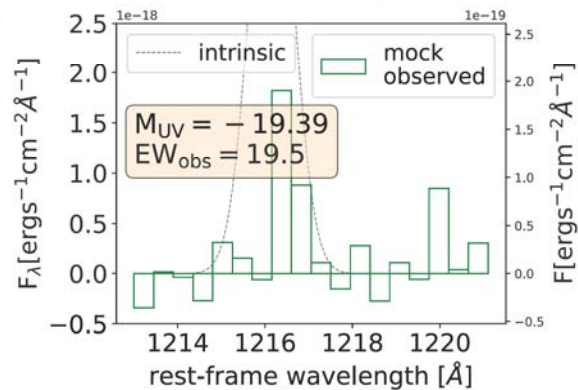
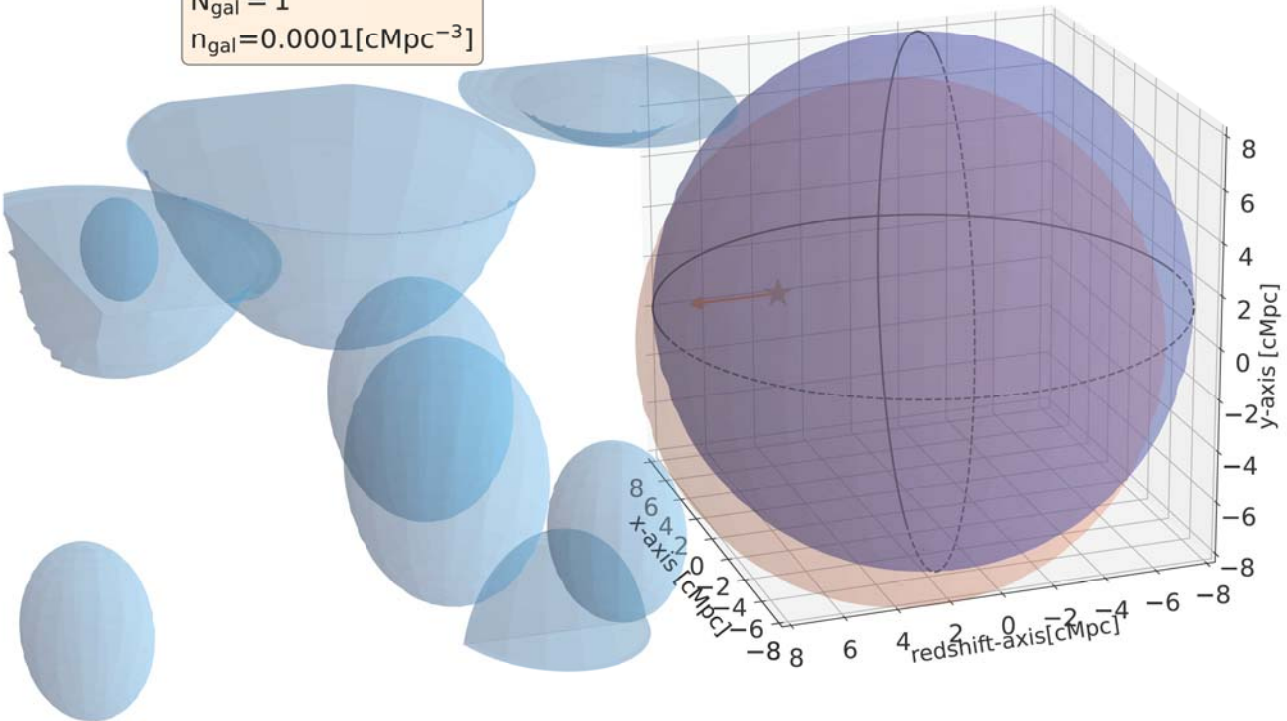
PRELIMINARY

$N_{\text{gal}} = 0$   
 $n_{\text{gal}} = 0.0000 [\text{cMpc}^{-3}]$



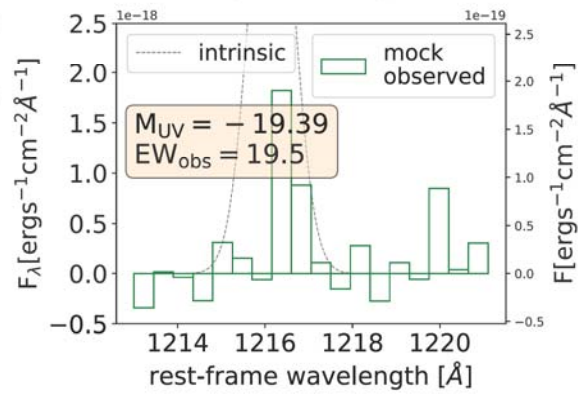
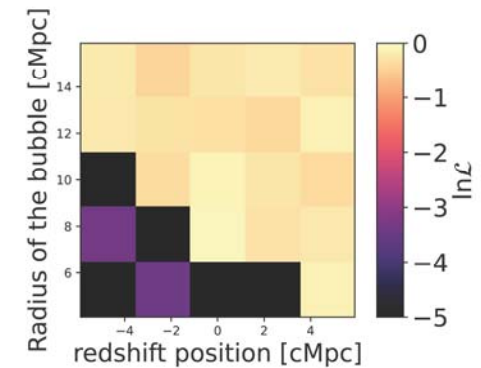
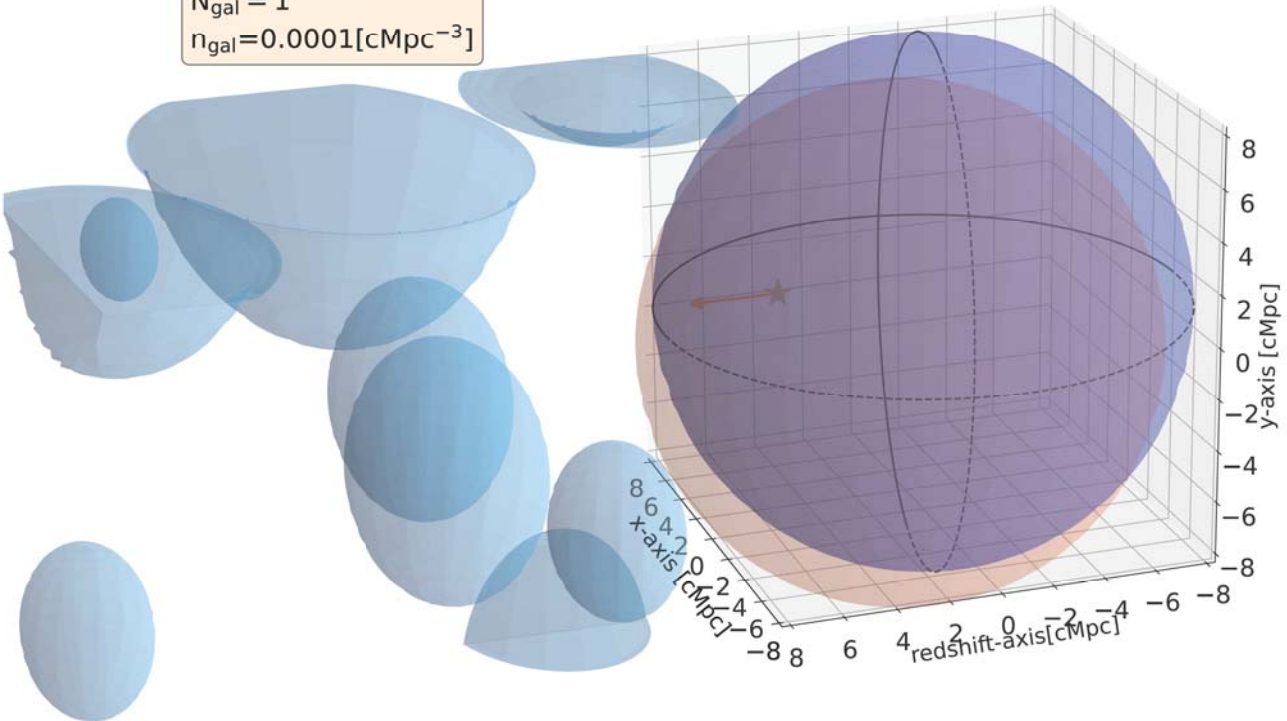
# Inferring individual bubbles around galaxy groups

$N_{\text{gal}} = 1$   
 $n_{\text{gal}} = 0.0001 [\text{cMpc}^{-3}]$



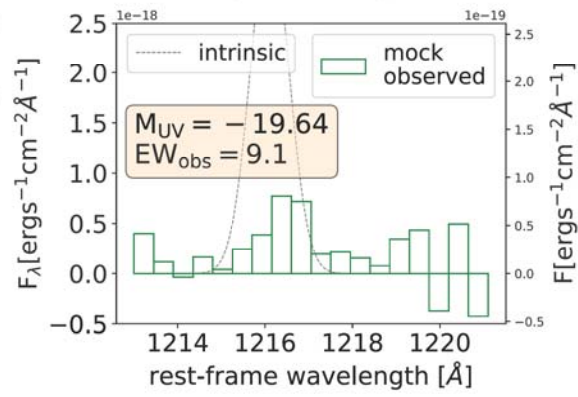
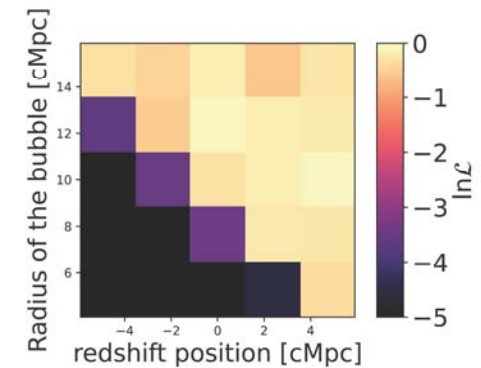
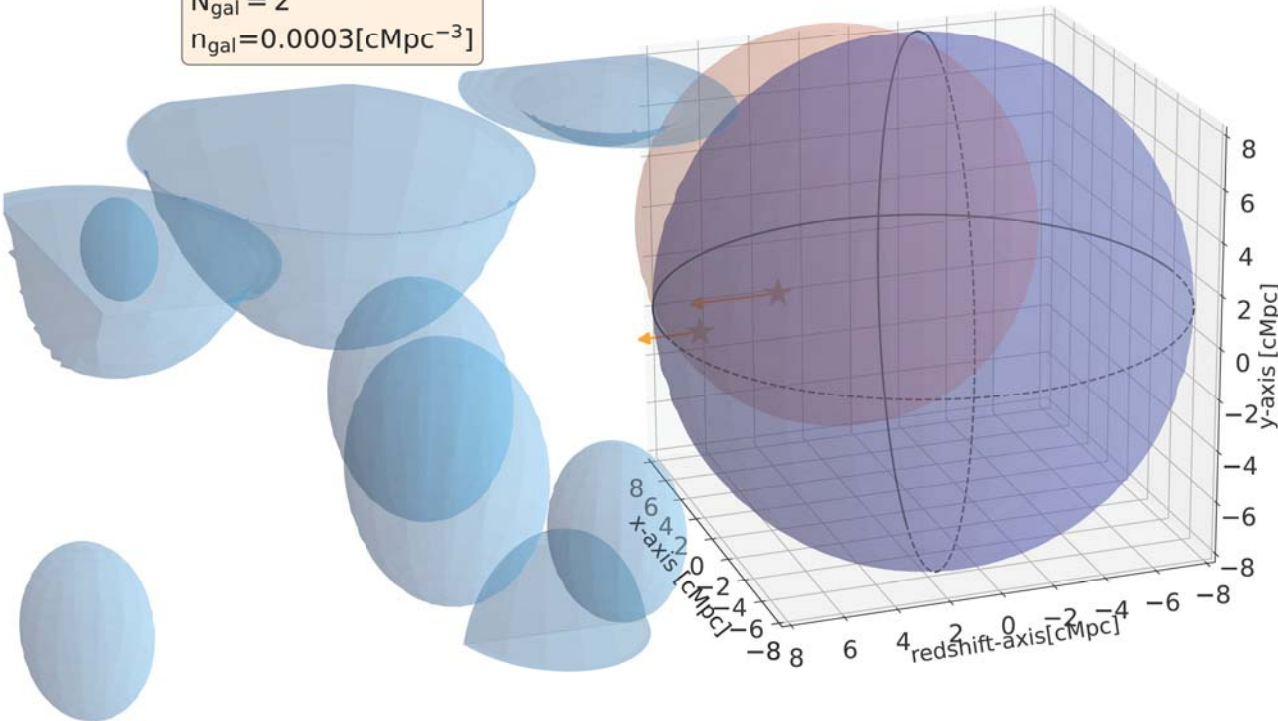
# Inferring individual bubbles around galaxy groups

$N_{\text{gal}} = 1$   
 $n_{\text{gal}} = 0.0001 [\text{cMpc}^{-3}]$



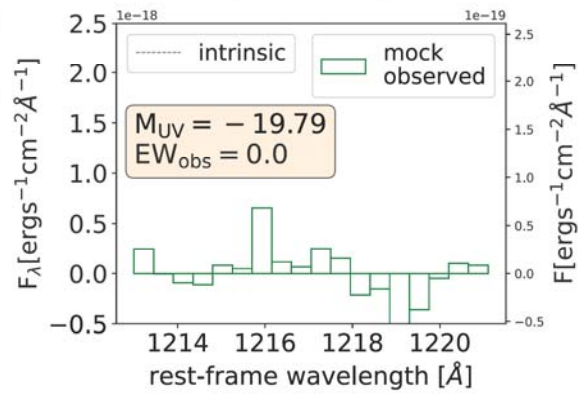
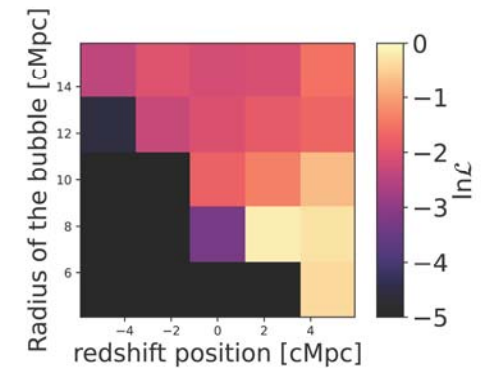
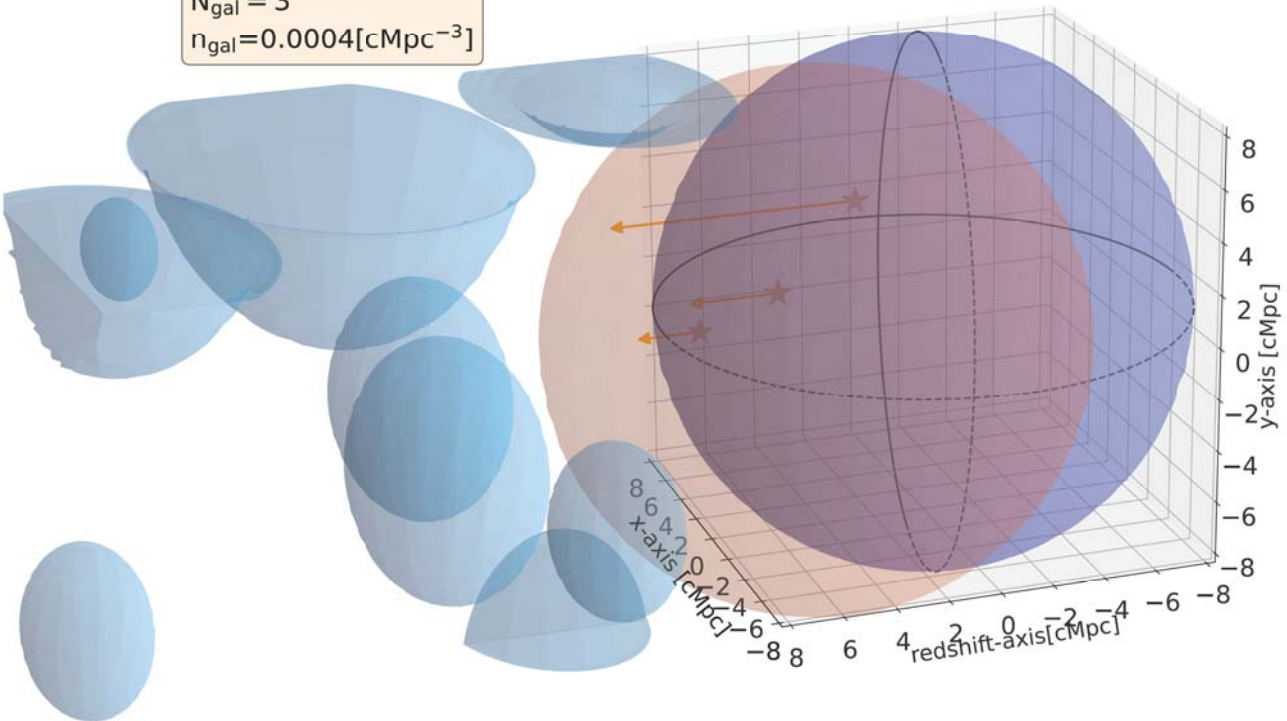
# Inferring individual bubbles around galaxy groups

$N_{\text{gal}} = 2$   
 $n_{\text{gal}} = 0.0003 [\text{cMpc}^{-3}]$



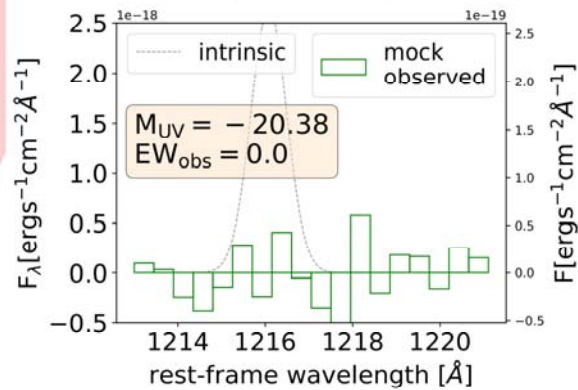
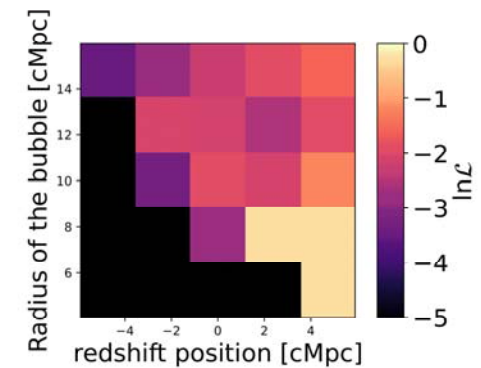
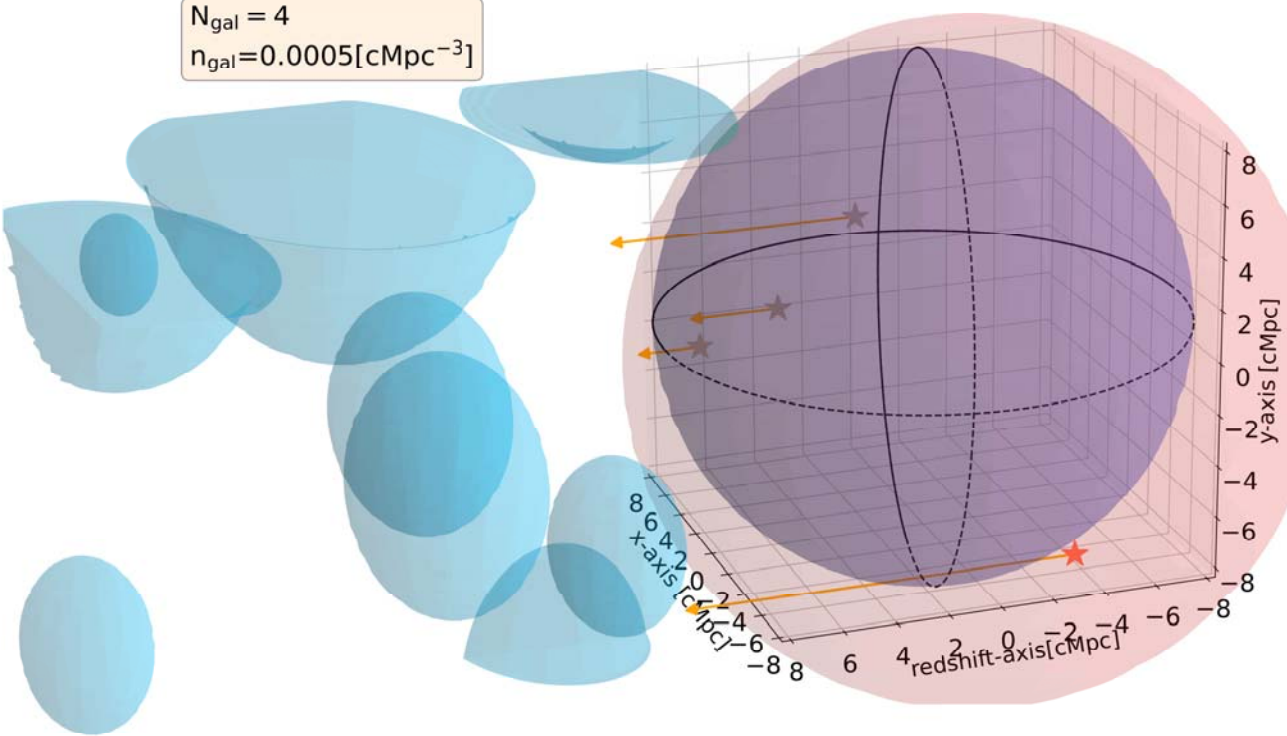
# Inferring individual bubbles around galaxy groups

$N_{\text{gal}} = 3$   
 $n_{\text{gal}} = 0.0004 [\text{cMpc}^{-3}]$

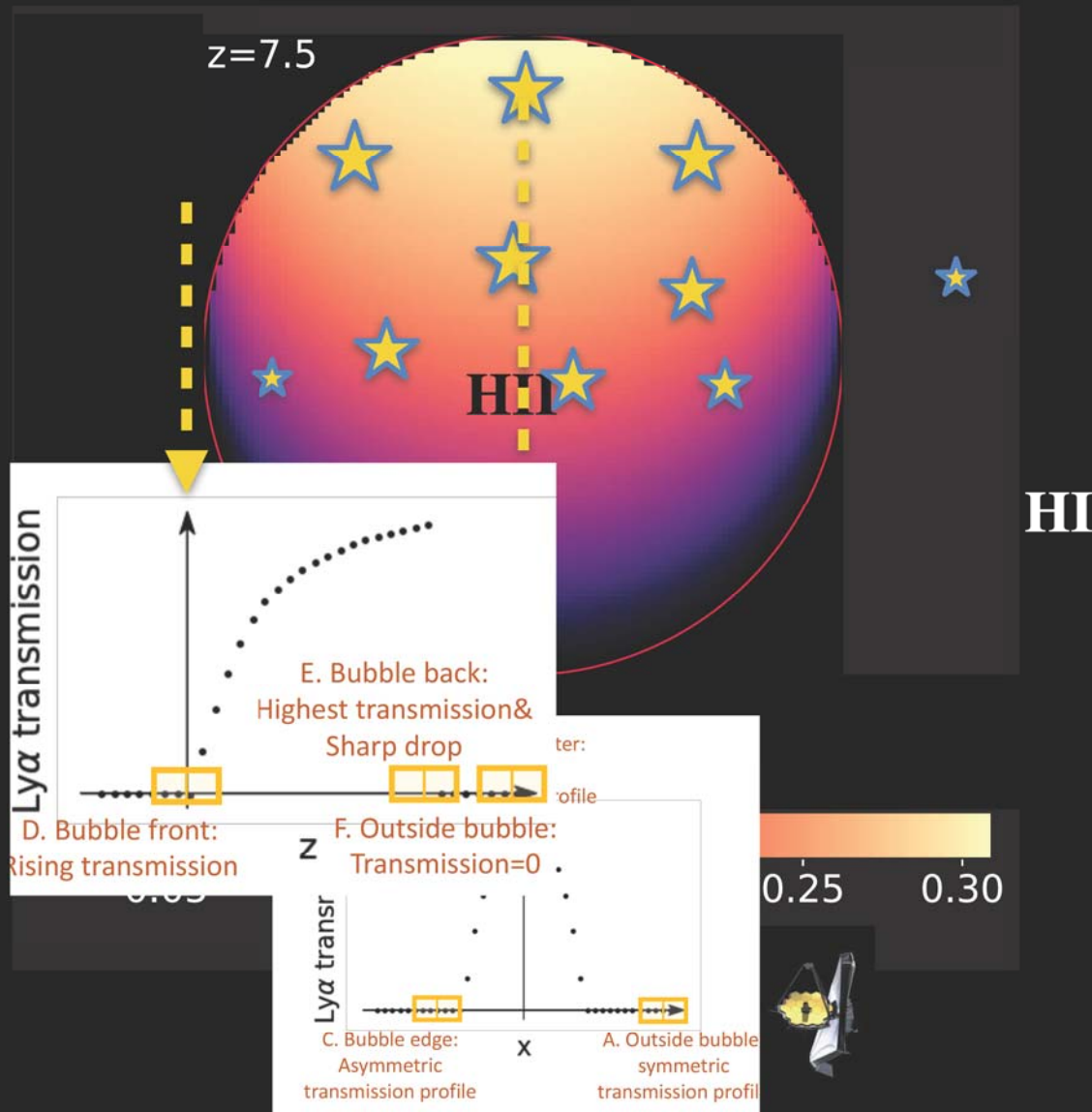


# Inferring individual bubbles around galaxy groups

$N_{\text{gal}} = 4$   
 $n_{\text{gal}} = 0.0005 [\text{cMpc}^{-3}]$



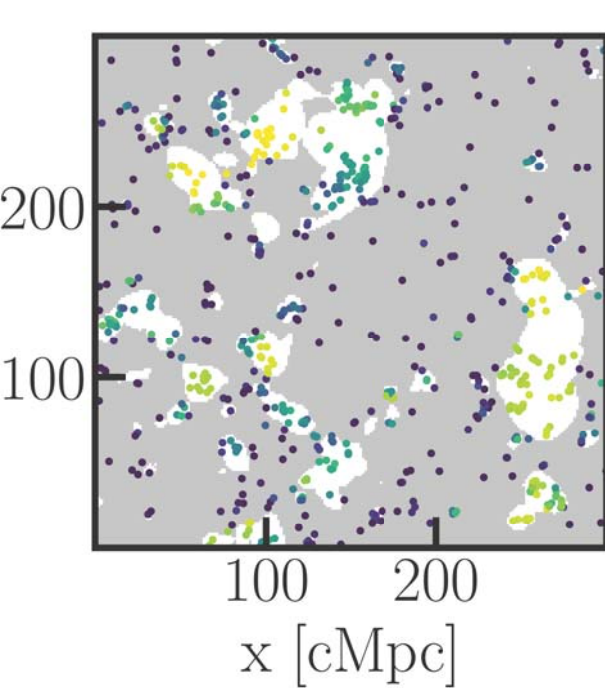
# Large-scale topology: *edge detection through asymmetry*



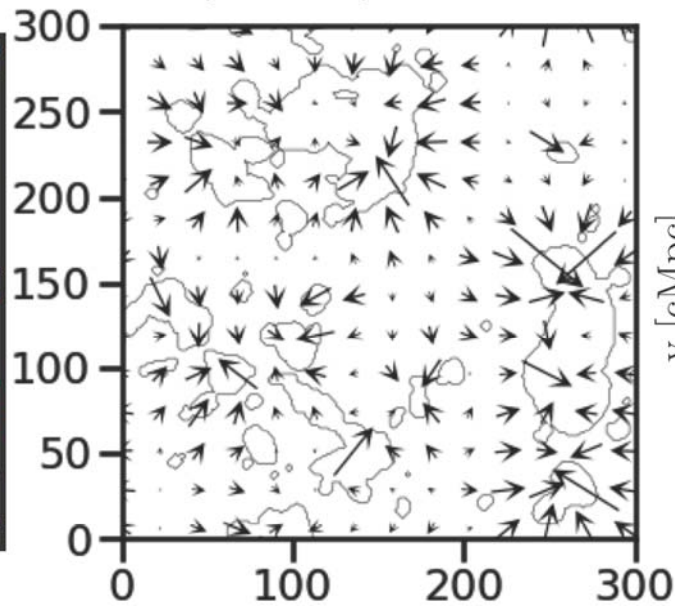
# Large-scale topology: edge detection through asymmetry

PRELIMINARY

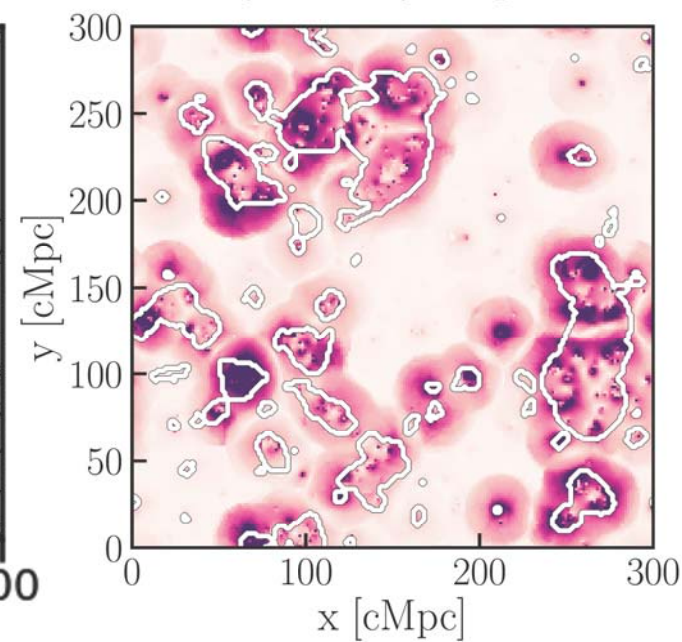
21cmFast simulation



Asymmetry vector field



Asymmetry magnitude





# Conclusions

- Our new **multi-scale forward models** recover Ly $\alpha$  forest data, in combination with UV LFs and CMB  $\tau_{\text{eff}}$ , without *ad-hoc tuning* or introducing *effective* parameters
- **Reionization is late: ends at  $z \sim 5.3 - 5.4$  with midpoint at  $z \sim 7 - 7.5$**
- Estimates of IGM mean free path and ionizing background are *converging between different analysis approaches* -> data is very constraining
- We do not know anything about the topology of this inhomogeneous process
  - (i) tells us **which galaxies drove the EoR**
  - (ii) allows us to connect **galaxy properties to local HII** environment; contribution of **faint galaxies**
- Ly $\alpha$  from galaxies can map local reionization topology during early EoR
  - (i) **Bayesian inference of HII bubbles is possible with  $\sim 3 \times 10^{-3}$  galaxies / cMpc $^3$**
  - (ii) **Asymmetry** of large-scale Ly $\alpha$  EW maps can find **edges of bubbles**

# Apply for a PhD in Physics and/or Computational Astrophysics at Scuola Normale Superiore in Pisa, Italy

- Scuola Normale Superiore (SNS) is Italy's premier university, consistently ranked top in the country and among the top in the world, when normalized by size
- The **Cosmology Group** at SNS has won numerous awards, including **3 ERC grants** and **3 Marie Curie fellowships**

