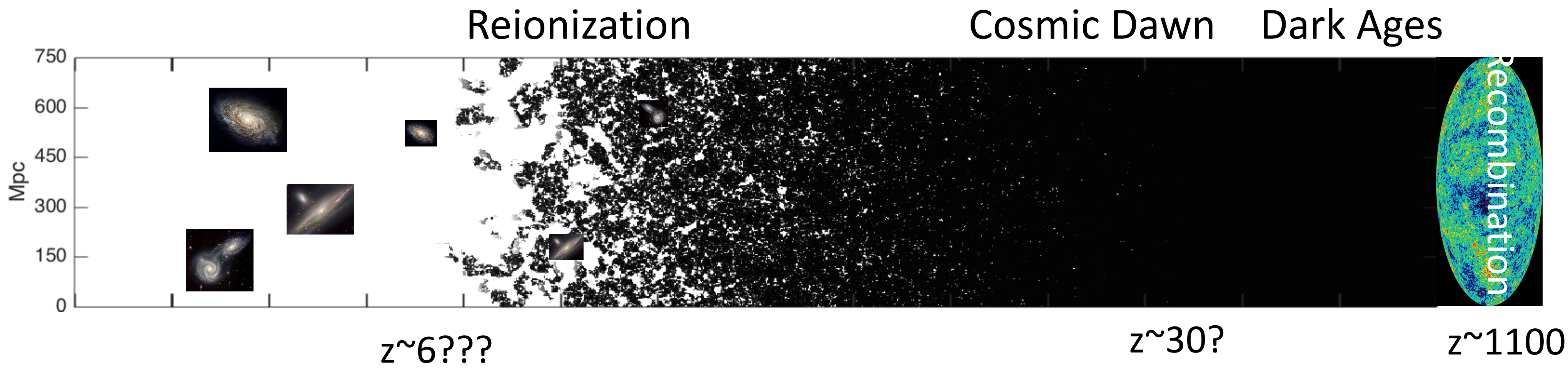
Cosmic Dawn and Reionization:
ongoing activities towards the SKA

Andrei Mesinger

SCUOLA NORMALE SUPERIORE

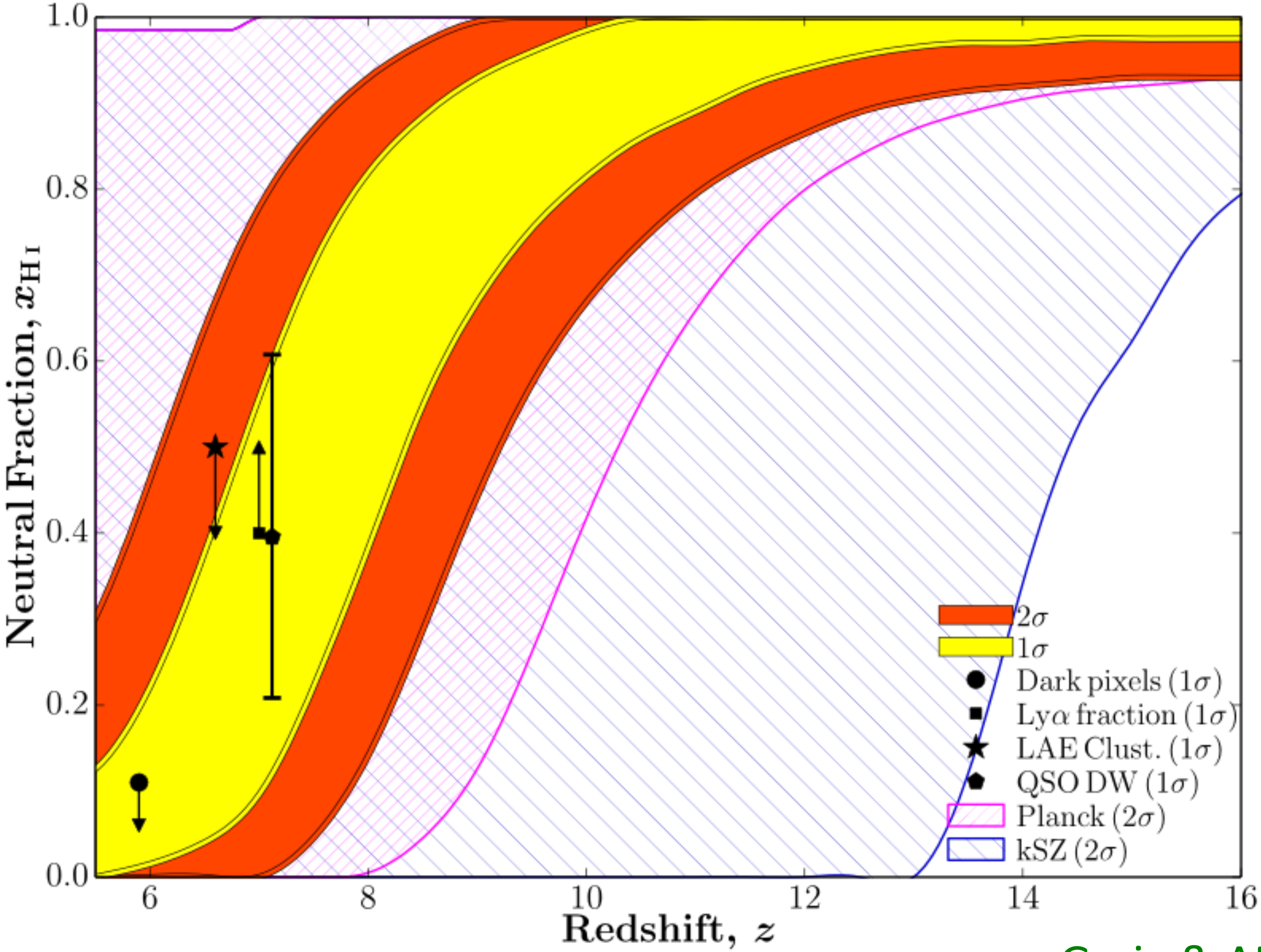
European Research Council

Why Cosmic Dawn?



Potentially some fundamental questions: **When** did the first generations of galaxies form? **What** were their properties? **How** did they interact with each other and the intergalactic medium? What is the structure of the intergalactic medium? What is the thermal and ionization history of the baryons?

current state of knowledge:
When did the Universe reionize?



Greig & AM (2017)
see also Planck 2016;
Price+2016; Mitra+2016

We now have a reasonable handle on when…

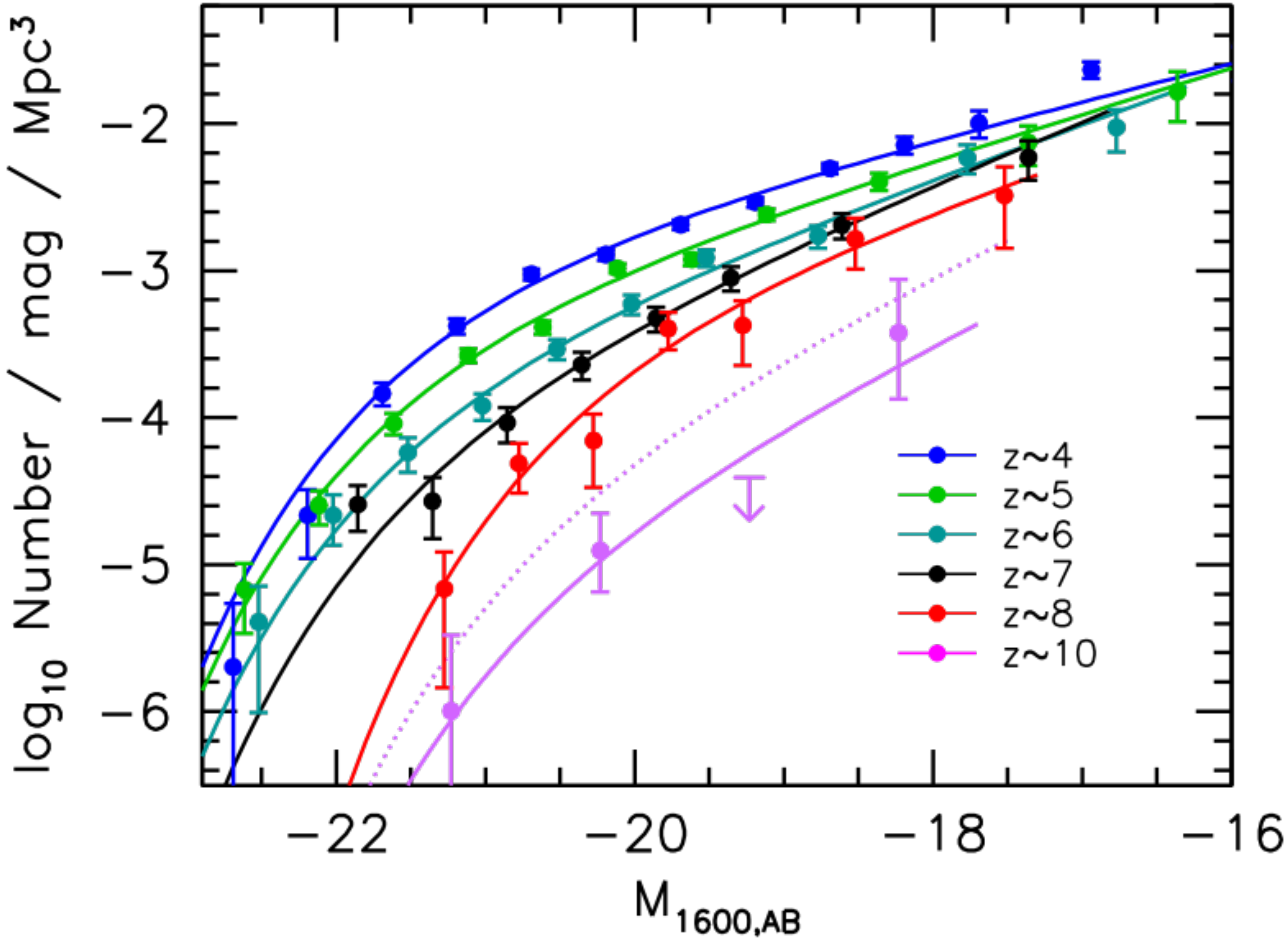
What and how??

stellar populations vs AGN, IMF in first galaxies, role of SNe and radiative feedback, metal pollution, efficiency of star formation, IGM structures, UVB evolution etc..

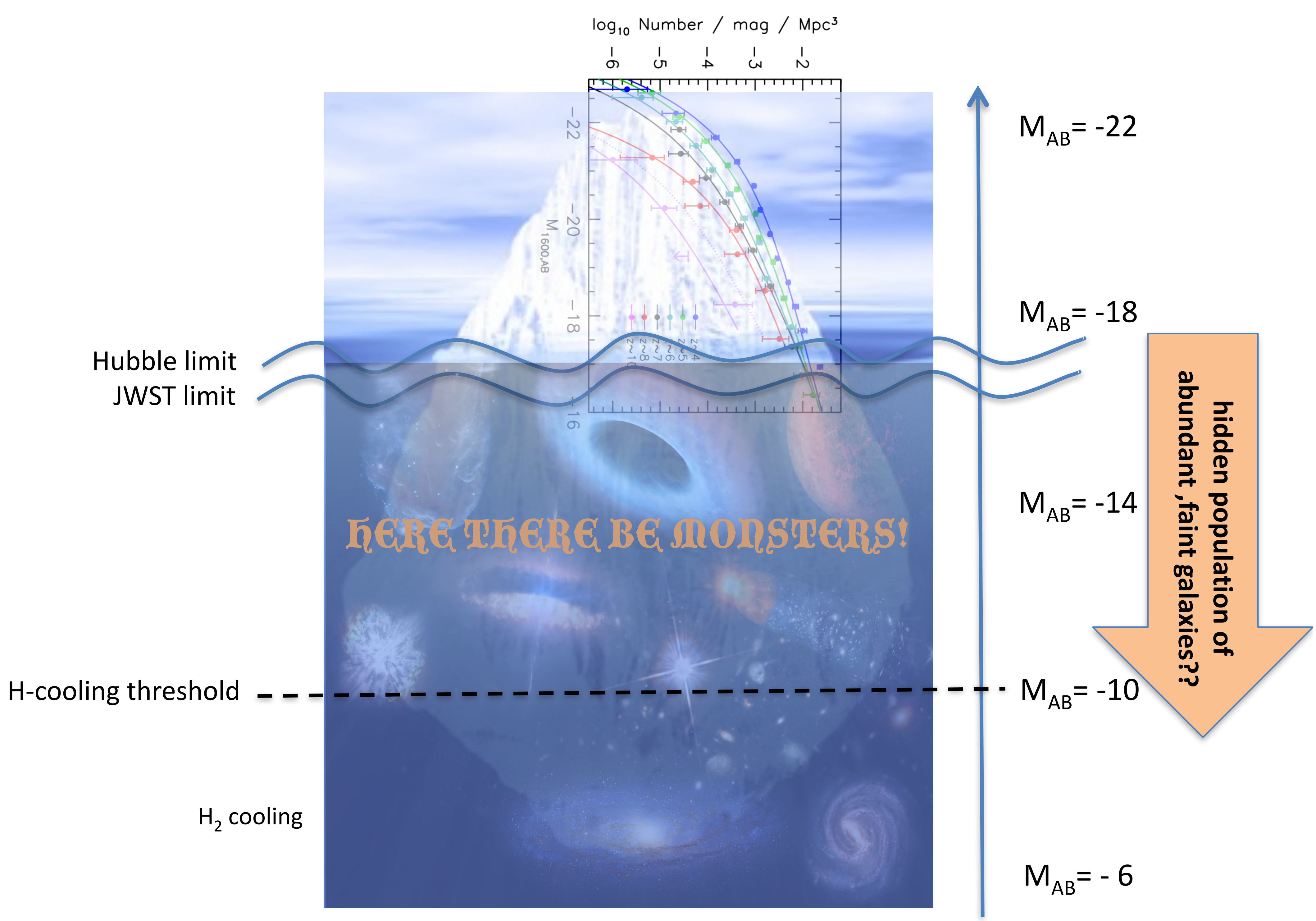
we don't really know…

What and how?

- Galaxy candidates have been found out to z~10. Are these the stellar populations responsible for the Cosmic Dawn and reionization? Estimates suggest they are too few, with too few ionizing photos escaping



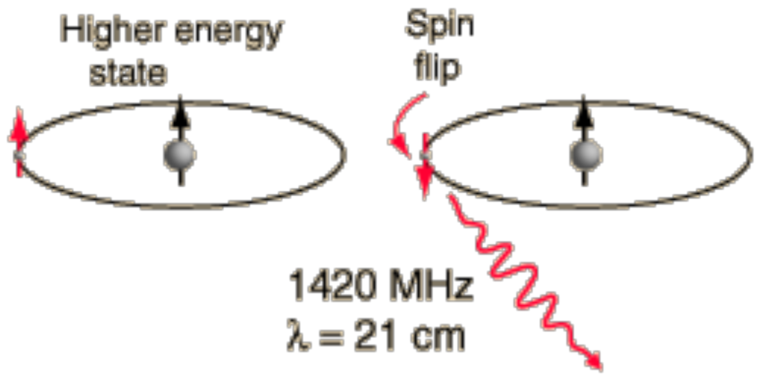
Bouwens+ (2015)



log10 Number / mag / Mpc³
M1600,AB
z~4
z~5
z~6
z~7
z~8
z~10
HERE THERE BE MONSTERS!
Hubble limit
JWST limit
H-cooling threshold
H2 cooling
MAB= -22
MAB= -18
MAB= -14
MAB= -10
MAB= - 6
hidden population of
abundant ,faint galaxies??

Get ready for the revolution: the cosmic 21 cm signal

21 cm line from neutral hydrogen

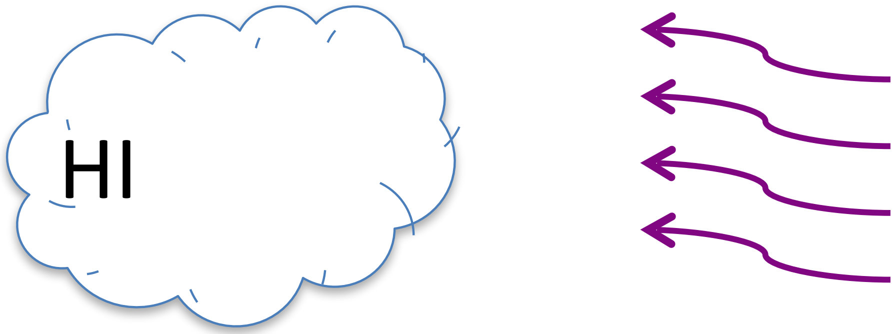


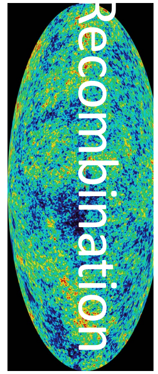
Hyperfine transition in the ground state of neutral hydrogen produces the 21cm line.

Cosmic 21-cm signal

SKA (202x —

z = 0





z~1100

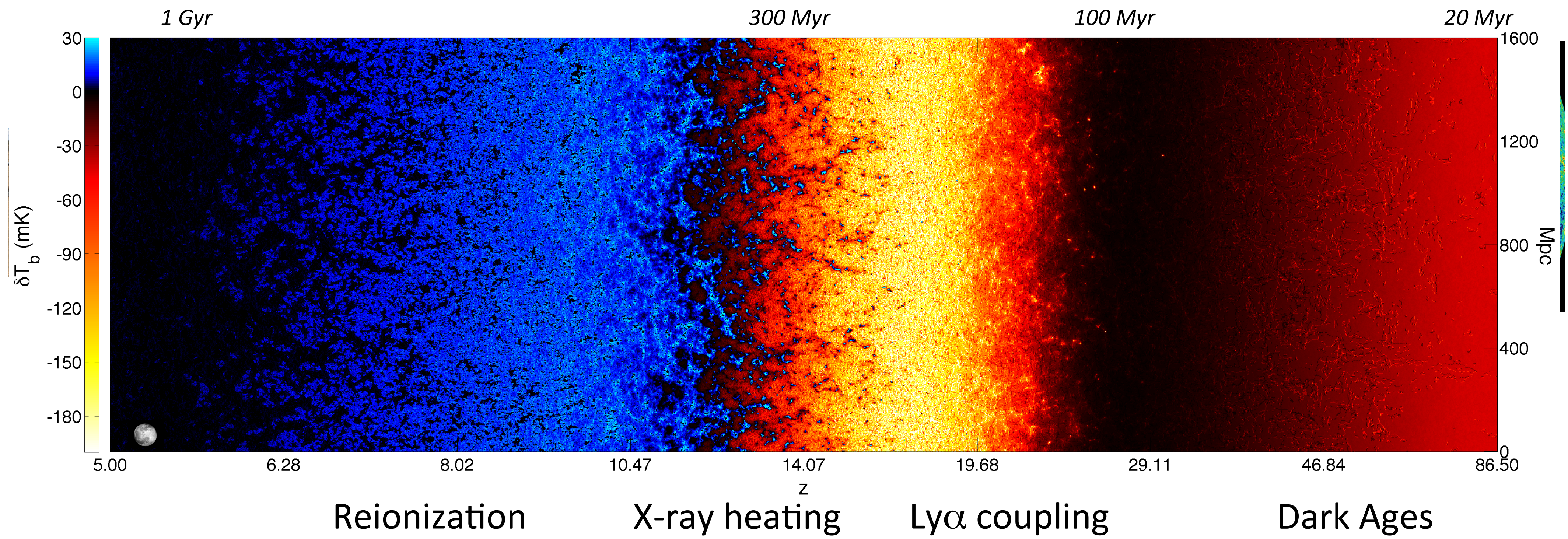
use the CMB as a background. measure the difference in intensities of the CMB and the cosmic HI, the so-called brightness temperature offset from the CMB:

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

Signal contains both **ASTROPHYSICAL** and **COSMOLOGICAL** terms

Cosmic 21-cm signal

AM+ 2016

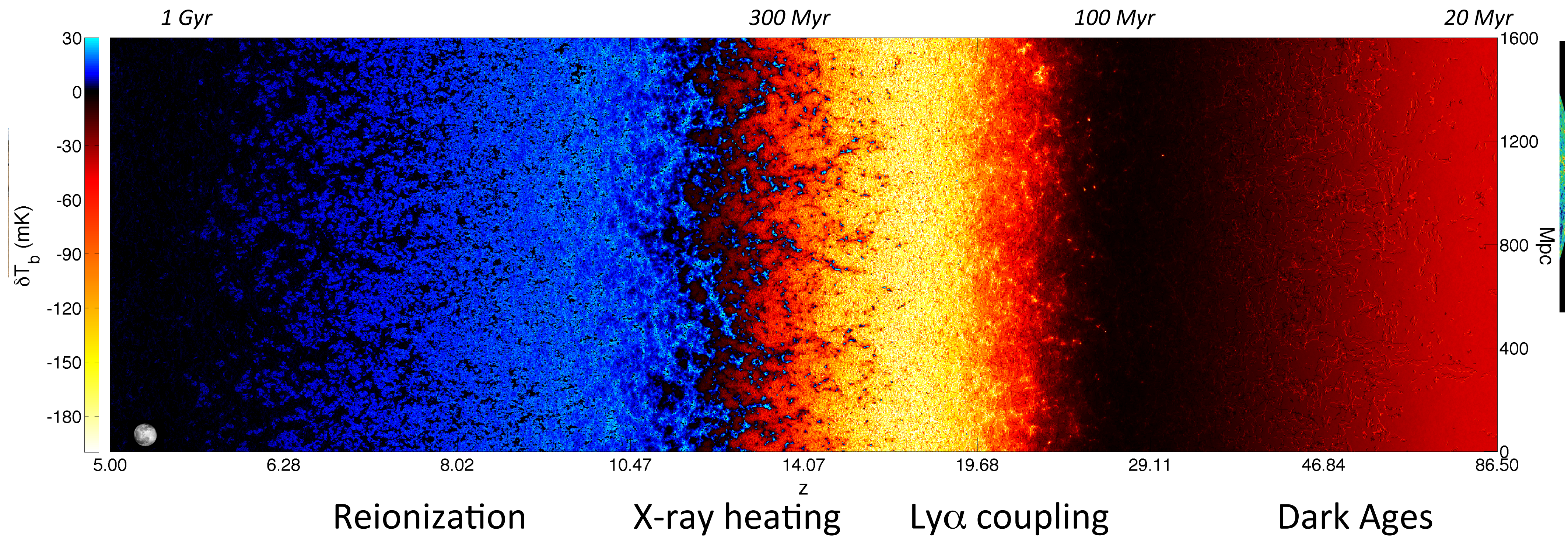


$$\delta T_b(\nu) \approx 27 x_{HI}(1+\delta_{nl})\left(\frac{H}{dv_r/dr + H}\right)\left(1-\frac{T_\gamma}{T_S}\right)\left(\frac{1+z}{10}\frac{0.15}{\Omega_M h^2}\right)^{1/2}\left(\frac{\Omega_b h^2}{0.023}\right) \mathrm{mK}$$

Signal contains both **ASTROPHYSICAL** and **COSMOLOGICAL** terms

Cosmic 21-cm signal

AM+ 2016



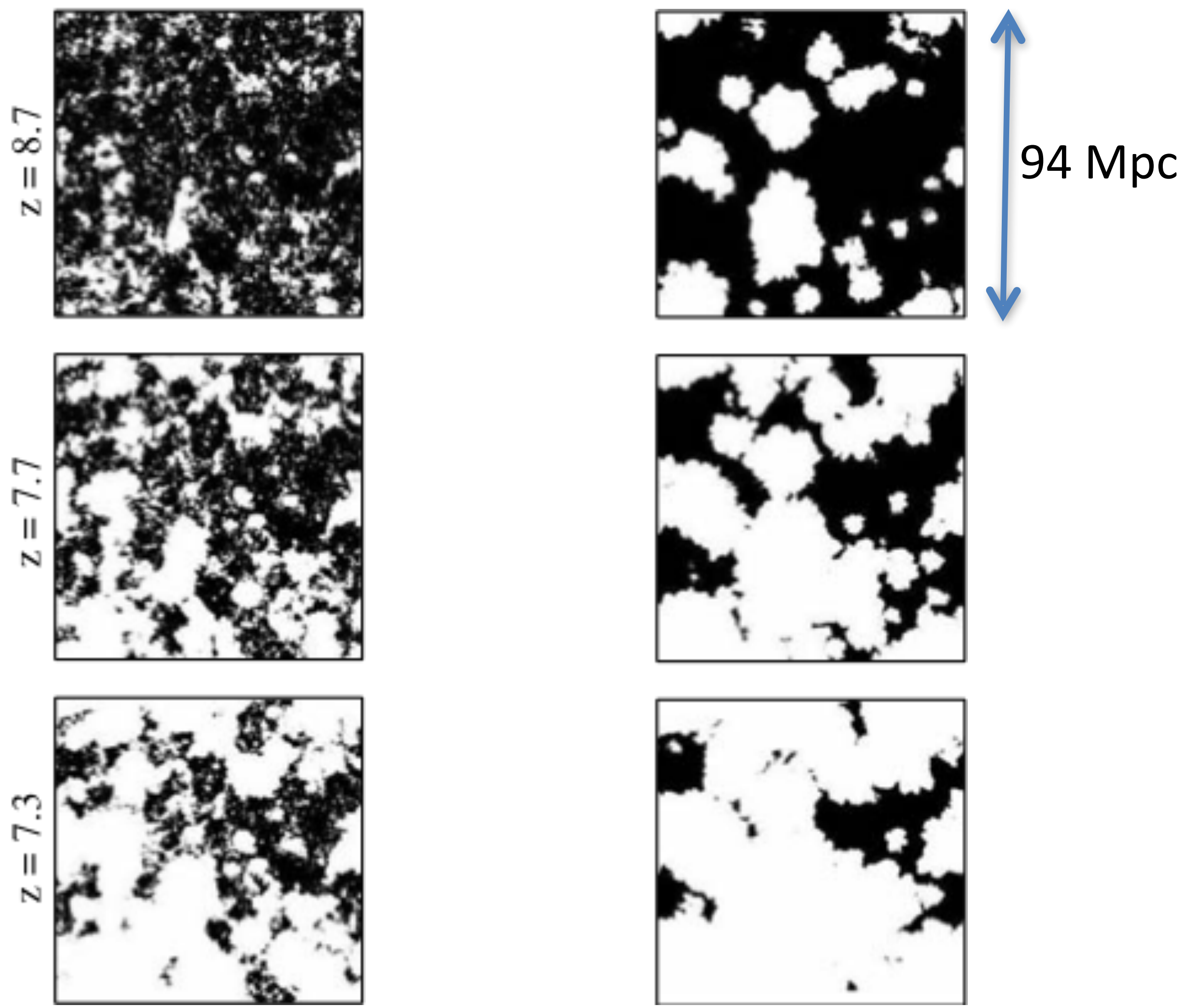
- **3D** signal with **> 10 orders of magnitude** more independent modes than in the CMB!
- data collection with upcoming Square Kilometre Array (SKA) will surpass **10x current global internet traffic**!
- even the narrowest fields will contain >billion of unseen galaxies
- **BIG DATA REVOLUTION**!

Signal contains both **ASTROPHYSICAL** and **COSMOLOGICAL** terms

So how do we learn about the unseen first galaxies?

Its all in the patterns!

- Galaxy clustering + stellar properties → *evolution of large-scale EoR/CD structures*



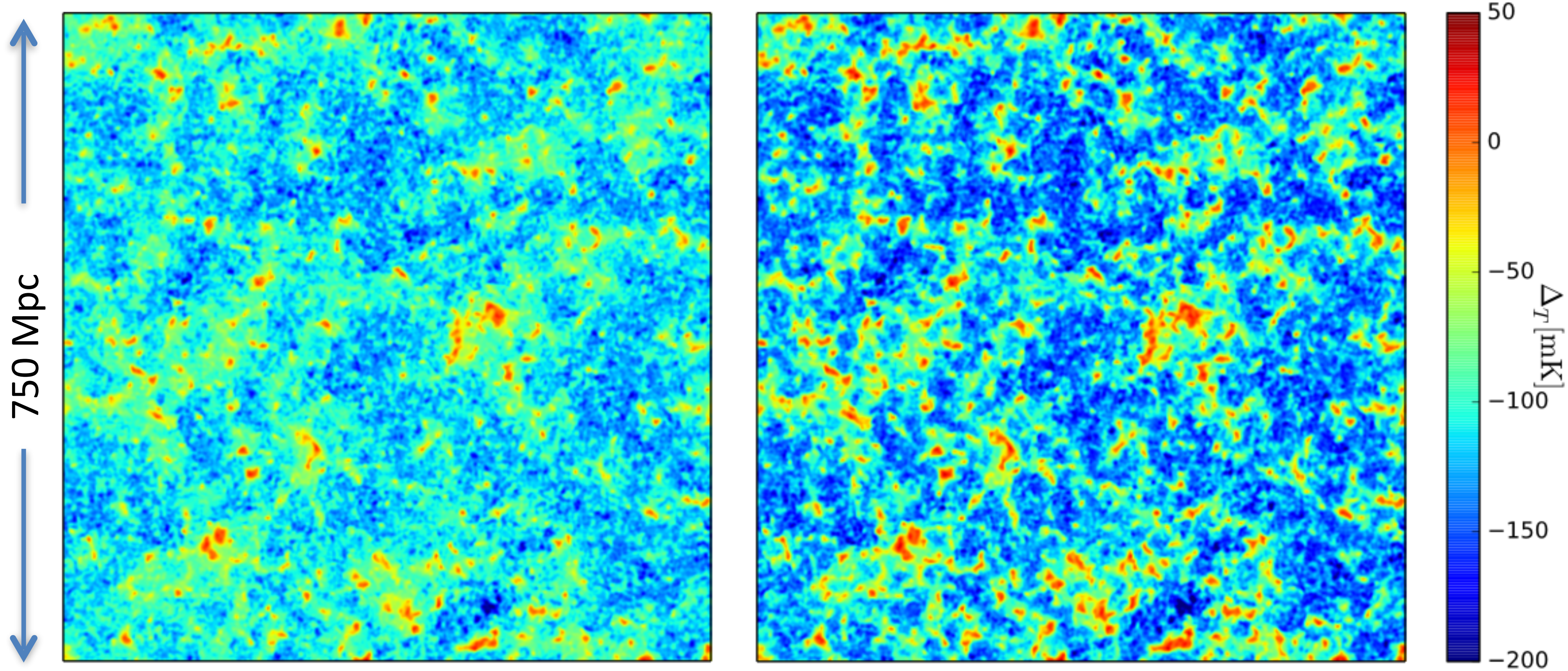
Abundant, faint galaxies vs **Rare, bright galaxies**

McQuinn+ 2007

Patterns in the Epoch of Heating

High-energy processes in the first galaxies are also encoded in the cosmic 21-cm signal

‘hard’ SED ~ HMXBs　　‘soft’ SED ~ hot ISM

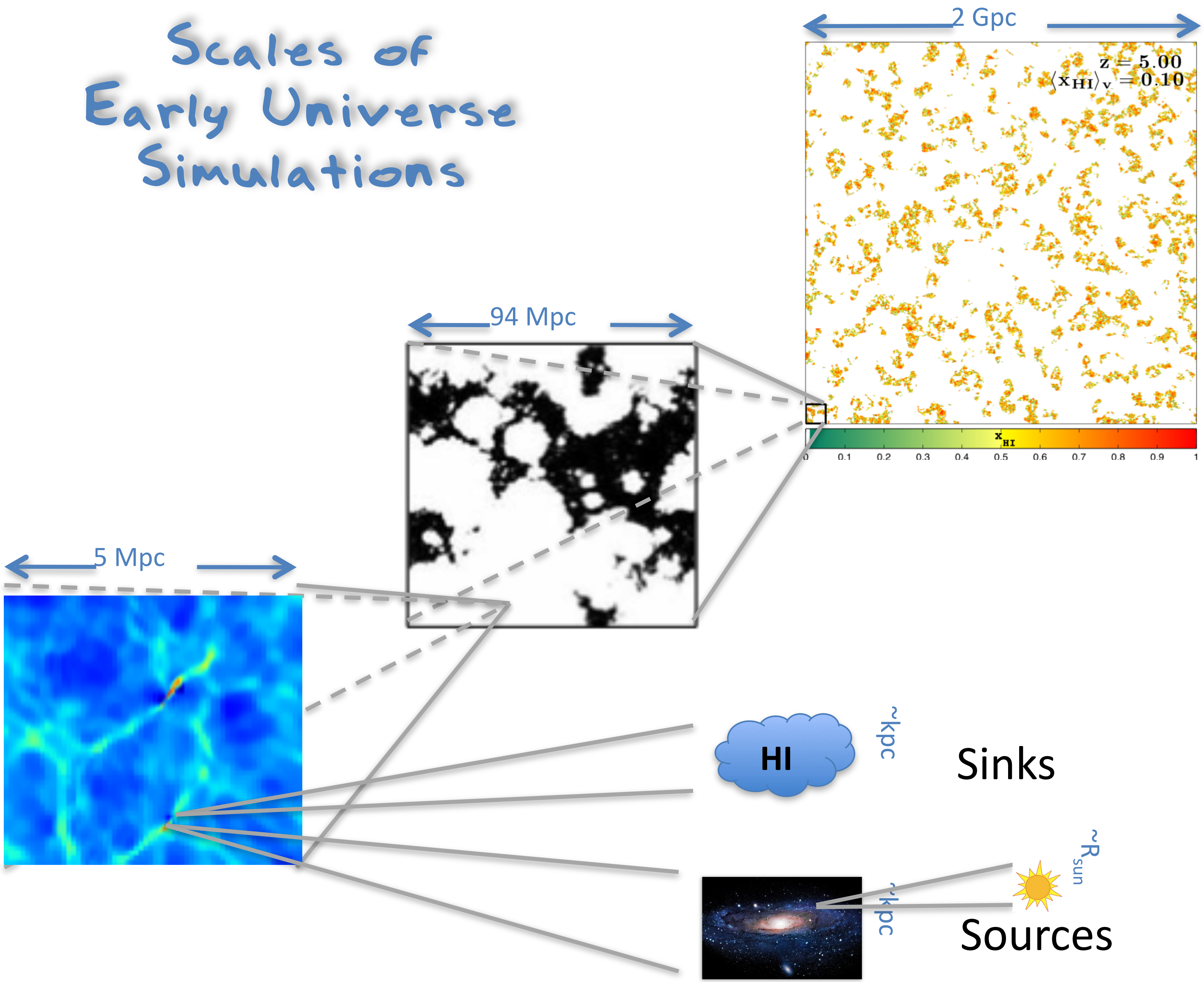


differences are easily detectable with HERA and the SKA

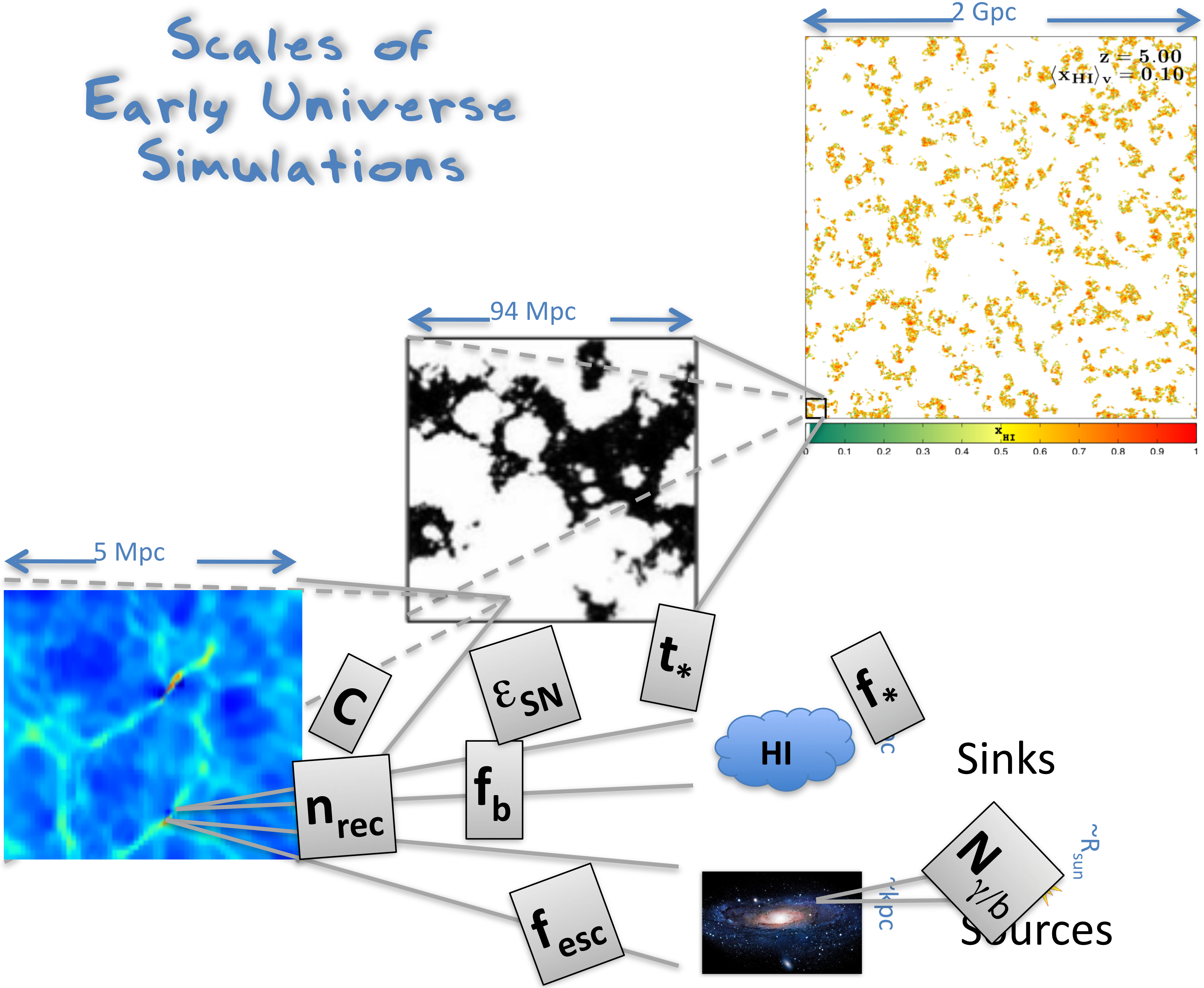
Pacucci+ 2014

How to quantify what we will learn??

Scales of Early Universe Simulations

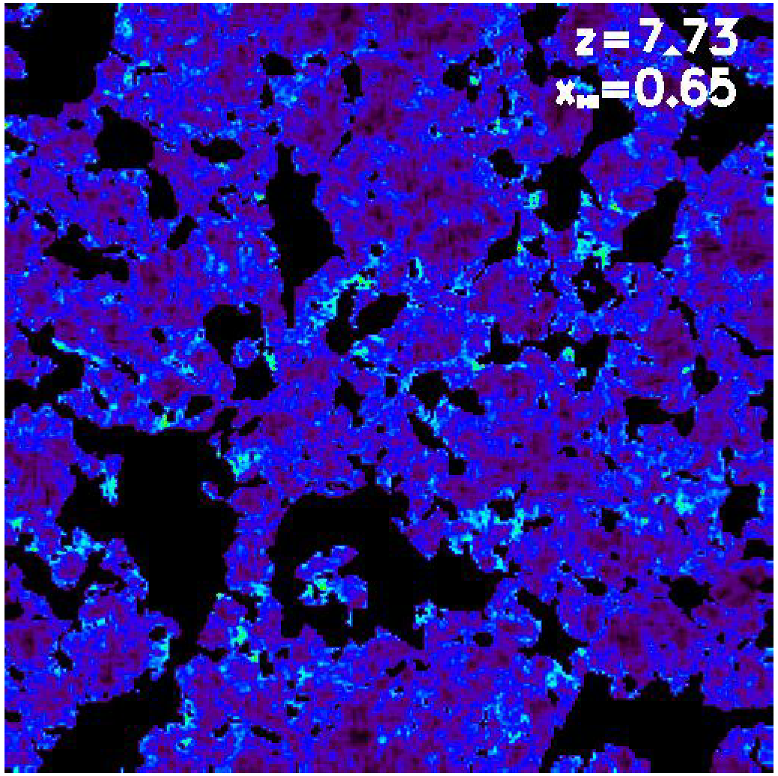


Astrophysical (known) unknowns

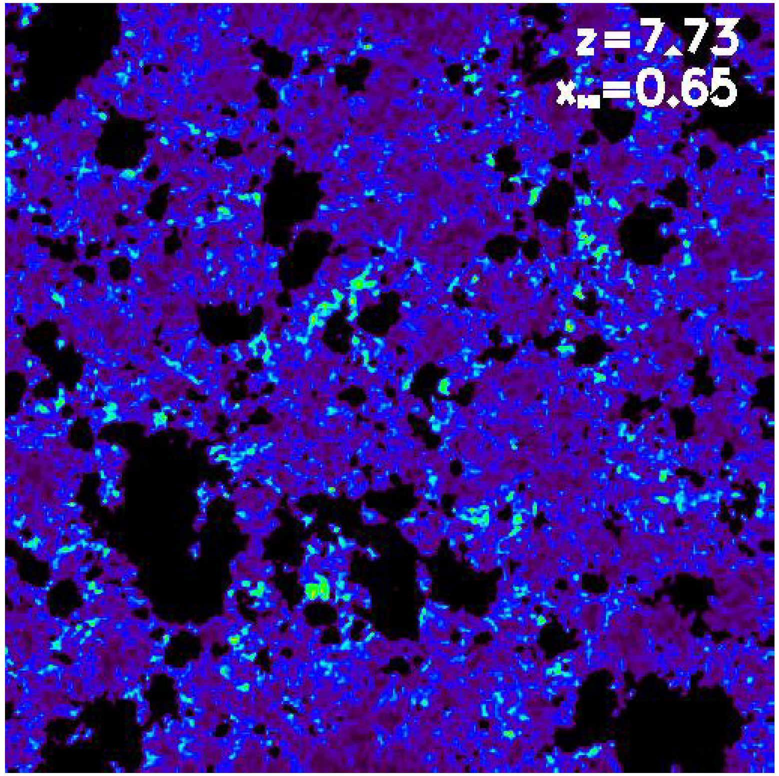


How to quantify what we will learn??

21cmFAST (AM+2007, 2011) — public, efficient semi-numerical 3D simulation code generating density fields (with 2LPT), and associated radiation fields (with a combination of excursion-set and lightcone integration).



density and ionization from 21cmFAST:
~ few min on a laptop



density and ionization from coupled hydro+RT (Trac+2009):
~ month on ~1000 core supercomputer

How to quantify what we will learn??

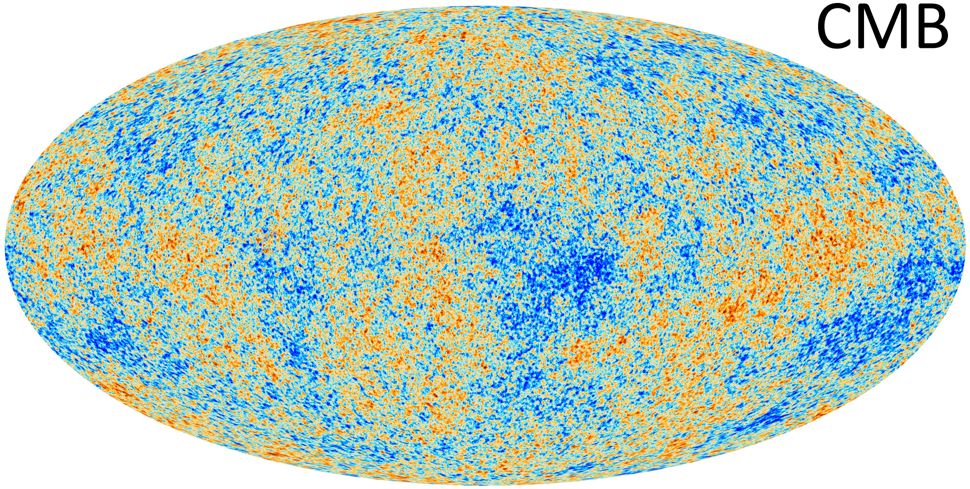
21cmFAST (AM+2007, 2011) — public, efficient semi-numerical 3D simulation code; extensively tested and currently used by *all* 21-cm efforts around the globe

+

21CMMC (Greig & AM 2015, 2017) – public, massively-parallelized MCMC driver for *21cmFAST*, based on EMCEE sampler (Forman-Mackey+ 2013)

Physical cosmology

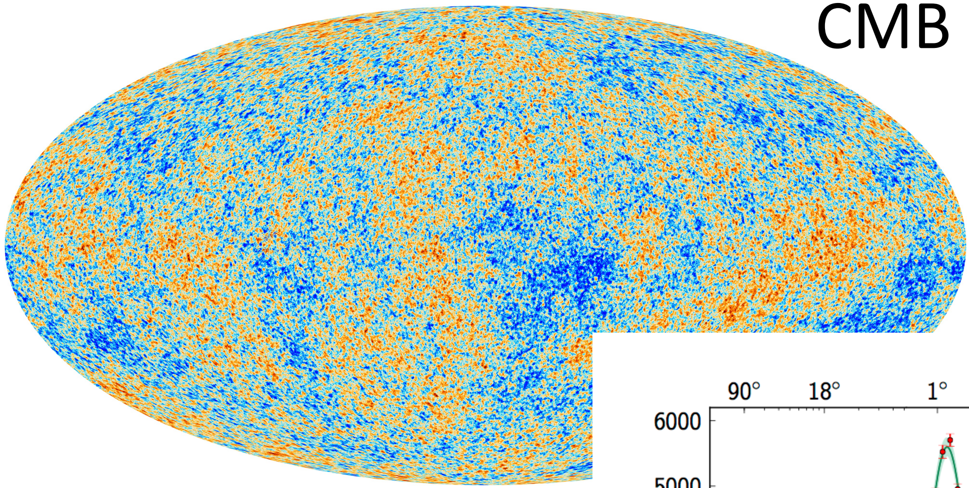
CMB map



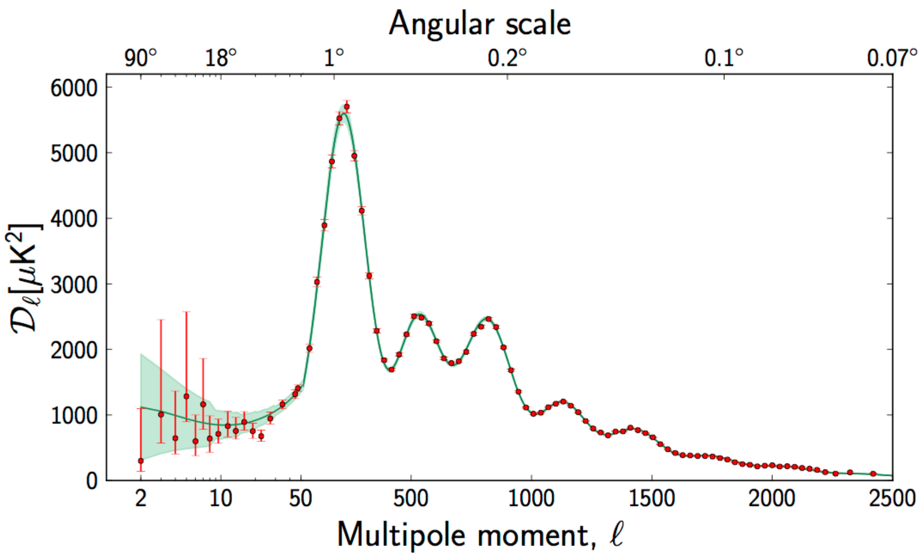
Planck 2013; 2015

Physical cosmology

CMB map



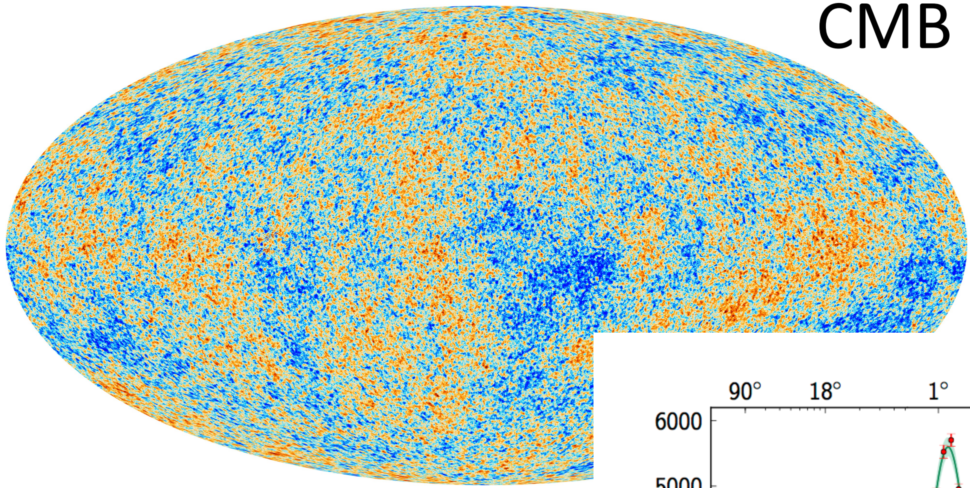
power spectrum



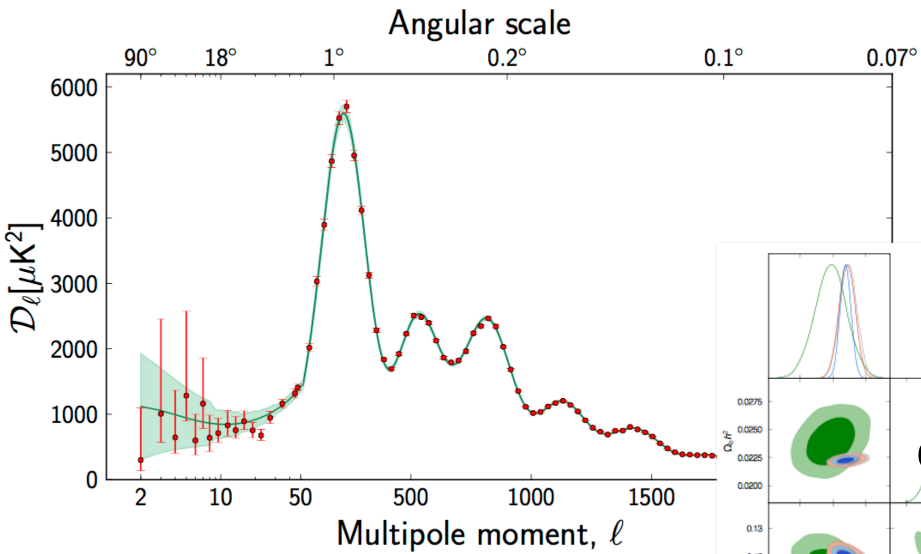
Planck 2013; 2015

Physical cosmology

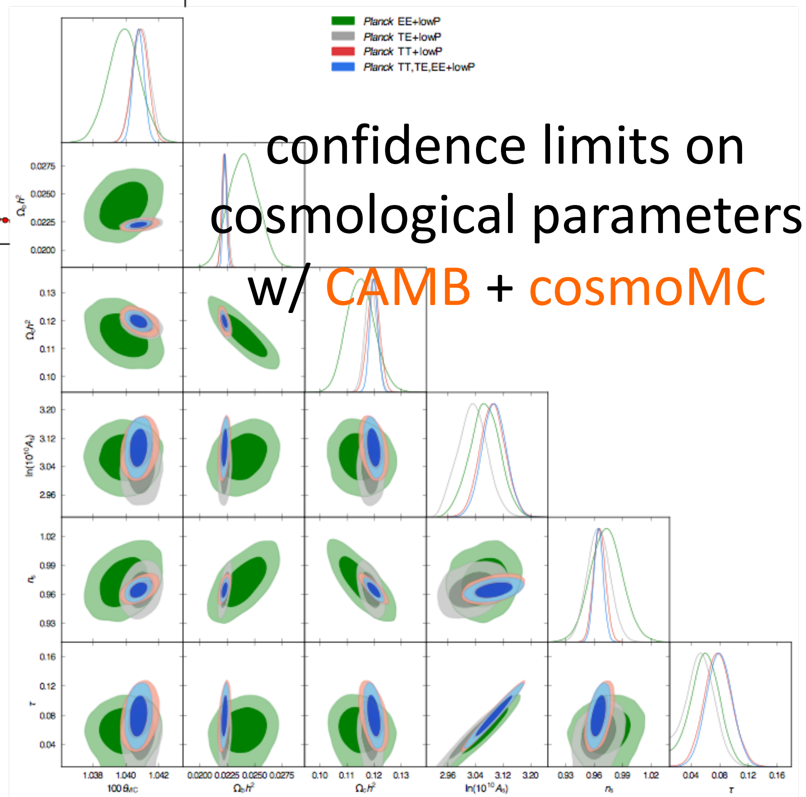
CMB map



power spectrum

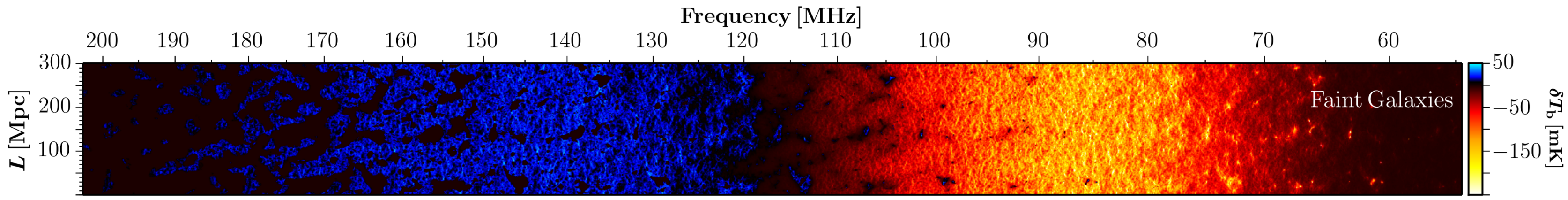


confidence limits on cosmological parameters w/ CAMB + cosmoMC



Planck 2013; 2015

*Astro*physical cosmology

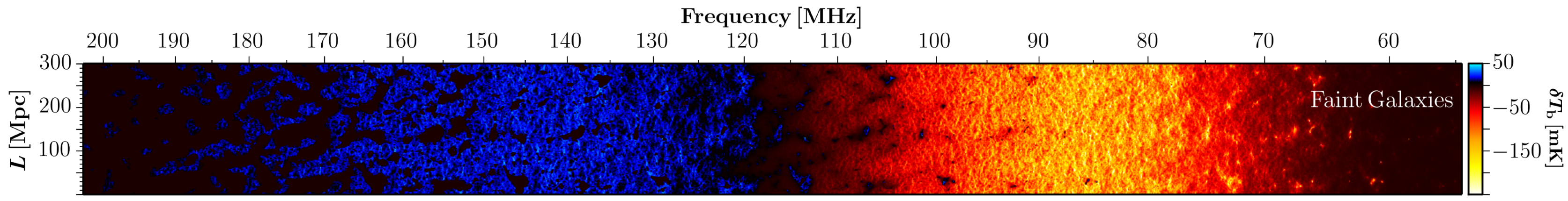


"observation"

21cm **3D!!!** map

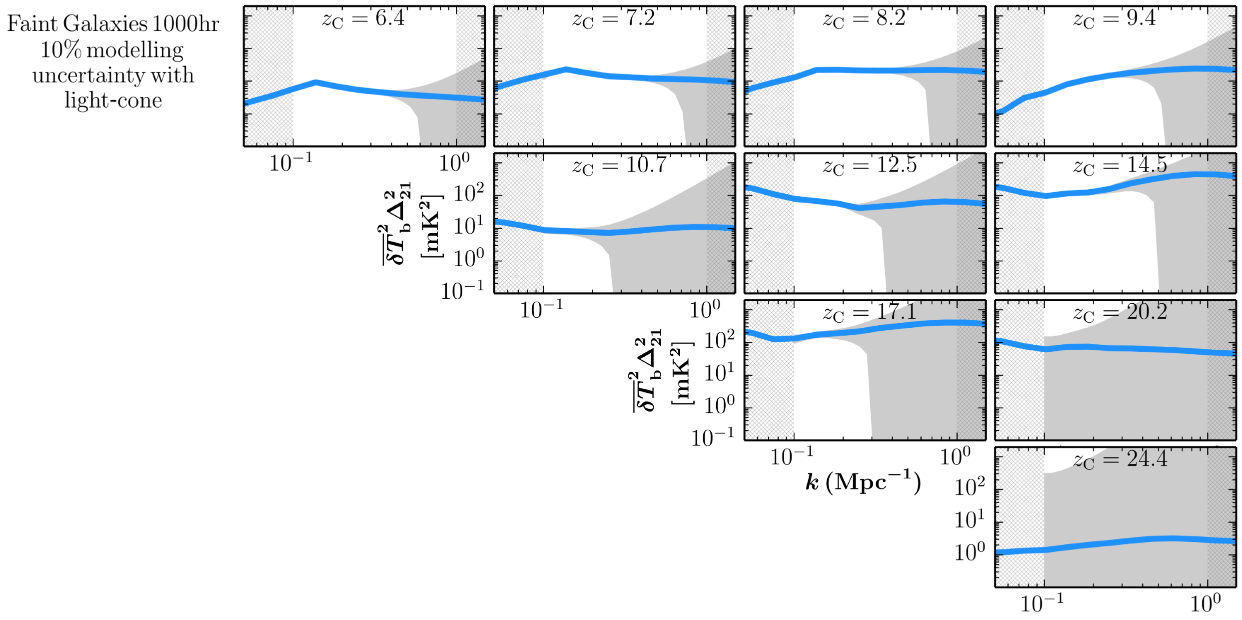
Greig & AM 2015; 2017

Astrophysical cosmology



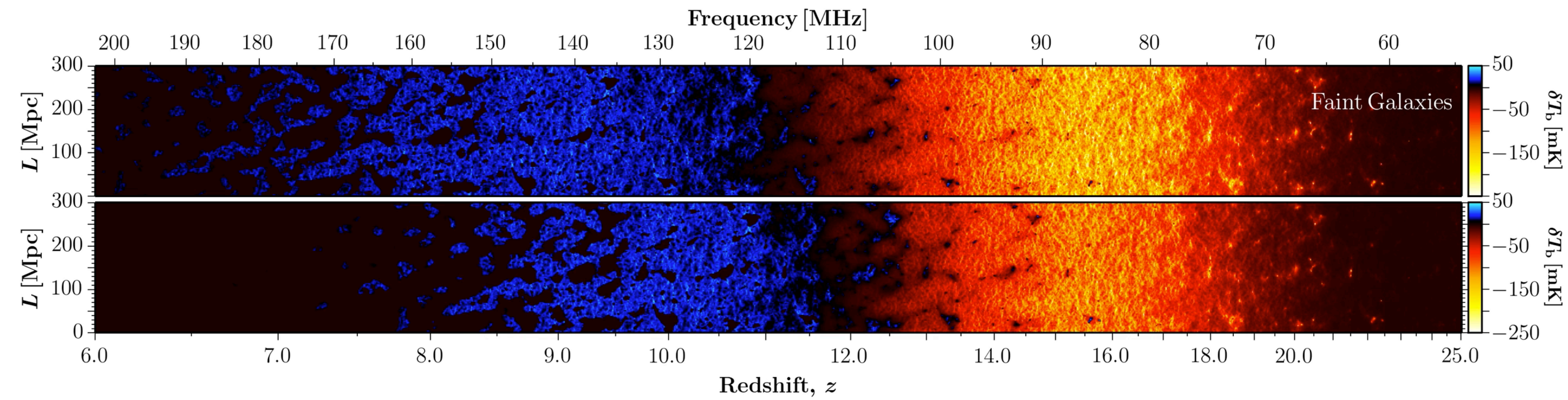
"observation"

power spectrum**??**



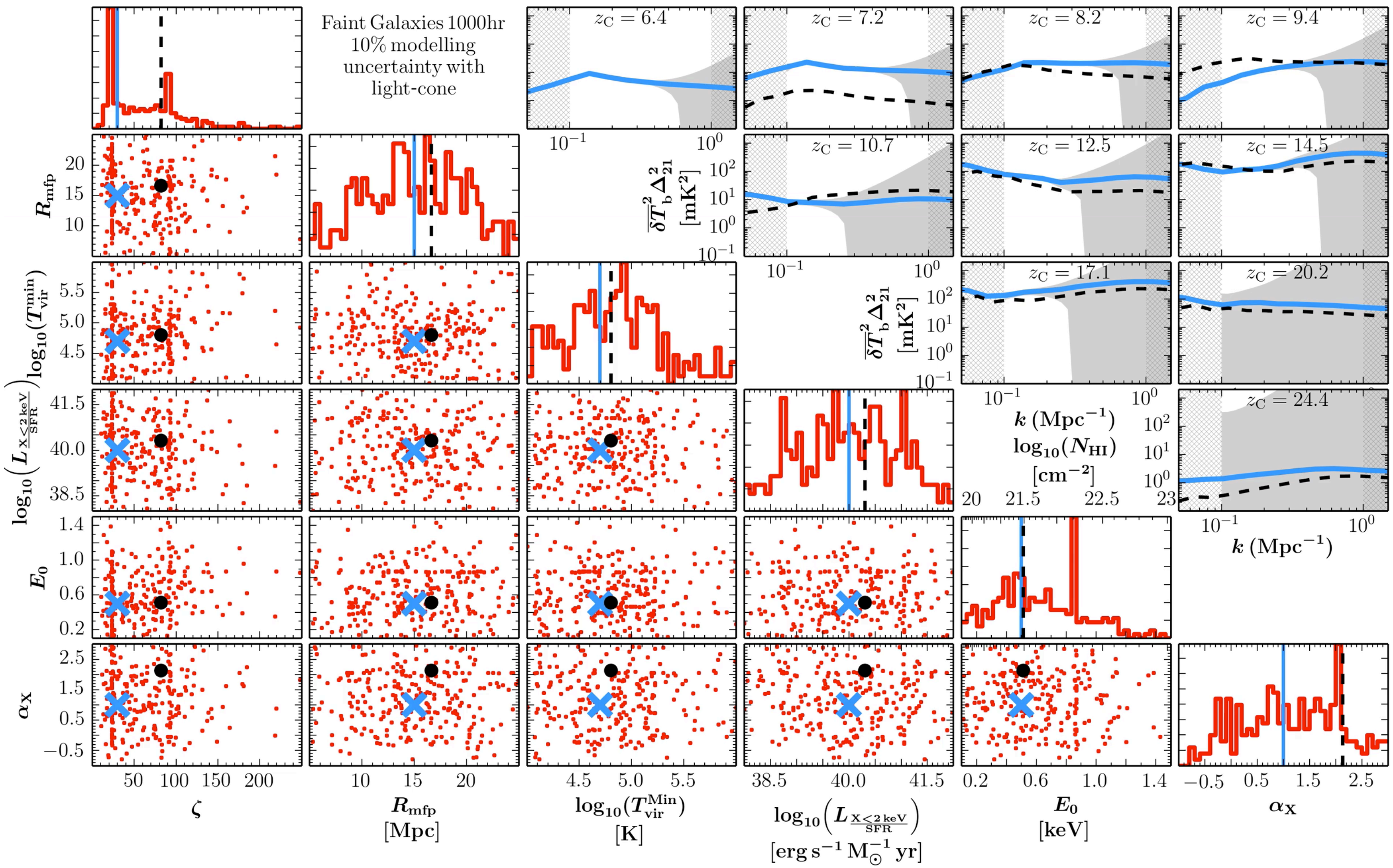
Greig & AM 2015; 2017

Astrophysical cosmology



"observation"

sample



21CMMC sampler

Greig & AM 2015; 2017

What are astrophysical parameters?????

What are astrophysical parameters?????

In, principle, our simulation and inference tools
can accommodate your favorite model… “model inclusive” philosophy

What are astrophysical parameters?????

In, principle, our simulation and inference tools
can accommodate your favorite model… “model inclusive” philosophy

*let’s try this **simple**, yet **flexible** empirical model…*

An flexible approach based on DM halos + galaxy LFs

Average properties of galaxies in halos of mass M_h:

$$\mathrm{M}_* = \mathbf{f}_{*,\mathbf{10}} \left(\frac{M_h}{10^{10} M_\odot}\right)^{\alpha_*} \frac{\Omega_b}{\Omega_M} M_h$$

$$L_{1500} \propto \frac{M_*}{\mathbf{t}_* H^{-1}}$$

$$L_{\rm ion} = \mathbf{f}_{\rm esc,10} \left(\frac{M_h}{10^{10} M_\odot}\right)^{\alpha_{\rm esc}} L_{1500}$$

$$f_{\rm duty} = \exp[-\mathbf{M}_{\rm turn}/M_h]$$

Park+ 2018
(see also Kuhlen+2012;
Dayal+ 2014; Mitra+ 2015;
Sun & Furlanetto 2016;
Mutch+ 2016; Yue+ 2016, …)

An flexible approach based on DM halos + galaxy LFs

Average properties of galaxies in halos of mass M_h:

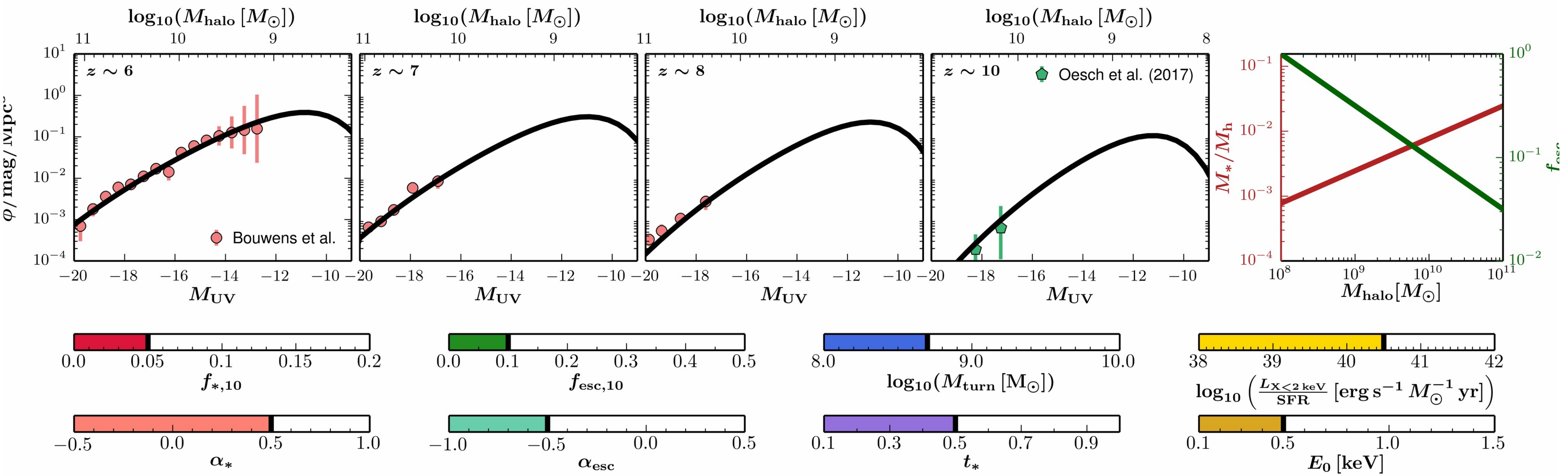
$$M_* = \mathbf{f_{*,10}} \left(\frac{M_h}{10^{10} M_\odot}\right)^{\alpha_*} \frac{\Omega_b}{\Omega_M} M_h$$

$$L_{1500} \propto \frac{M_*}{\mathbf{t_*} H^{-1}}$$

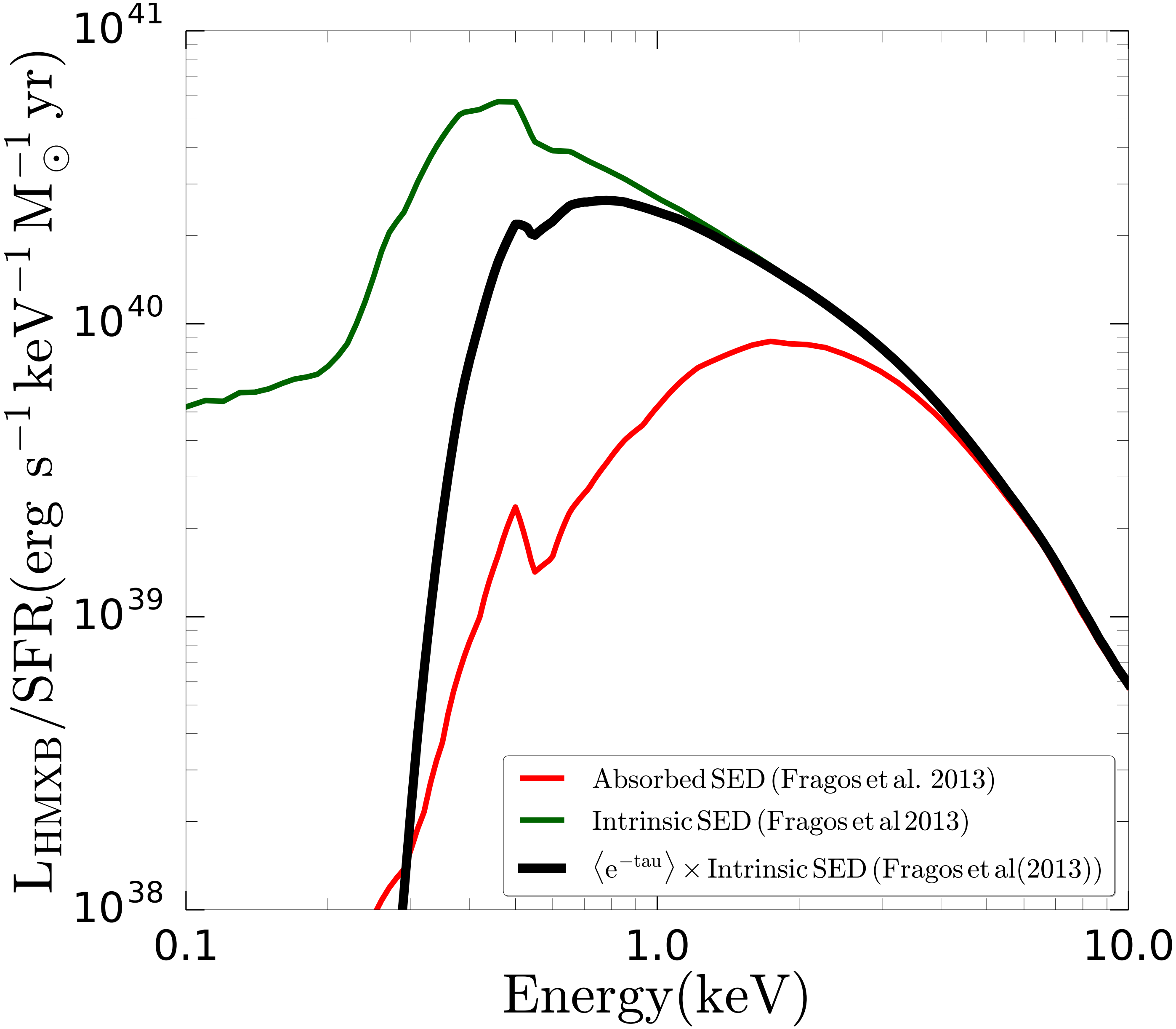
$$L_{\rm ion} = \mathbf{f_{esc,10}} \left(\frac{M_h}{10^{10} M_\odot}\right)^{\alpha_{\rm esc}} L_{1500}$$

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

six free parameters for UV photons



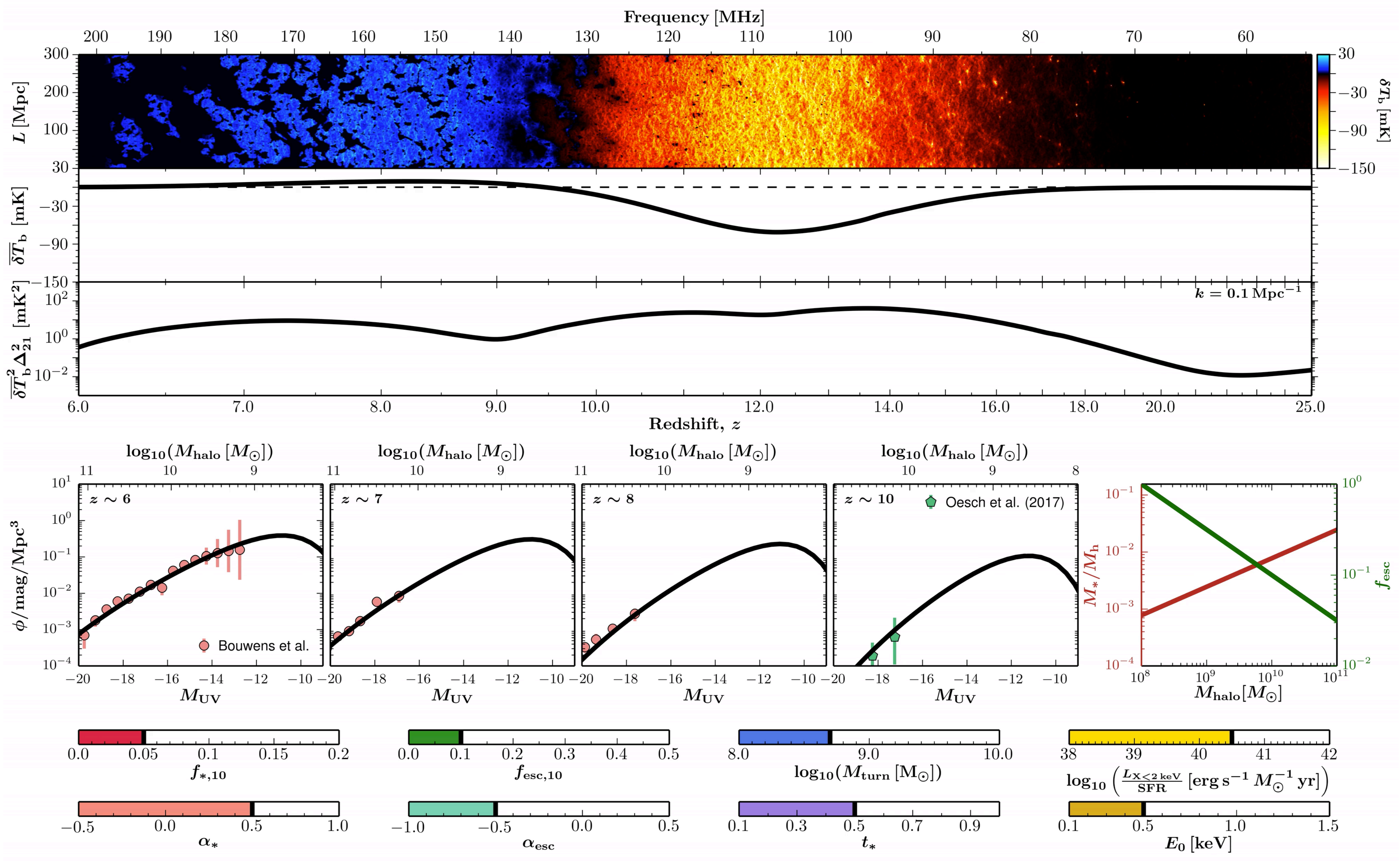
An flexible approach based on DM halos + galaxy LFs



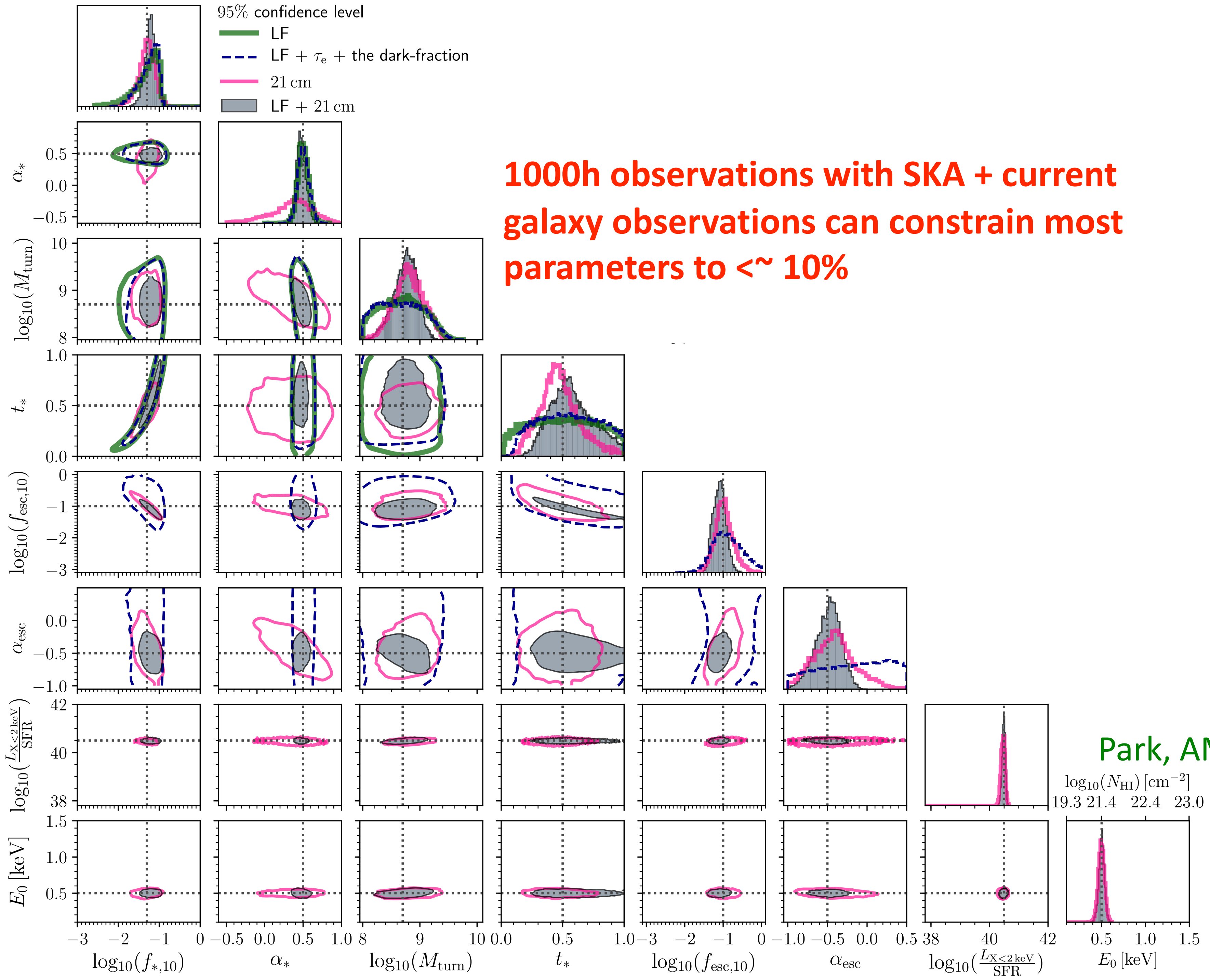
X-ray free parameters characterizing emerging SED from galaxies

Das, AM+ 2016

Free parameters



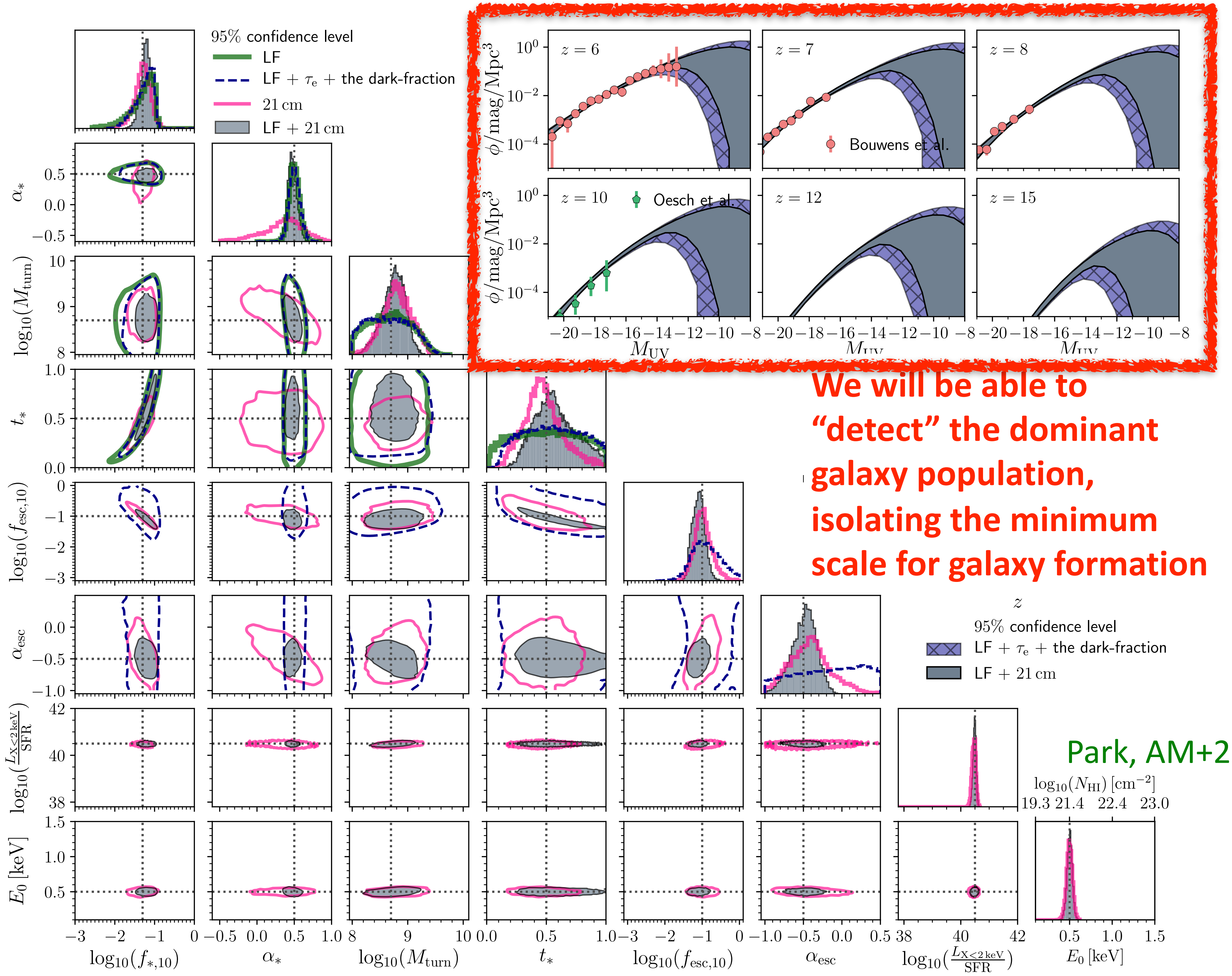
Parameter constraints: LF + 21cm



1000h observations with SKA + current galaxy observations can constrain most parameters to <~ 10%

Park, AM+2018

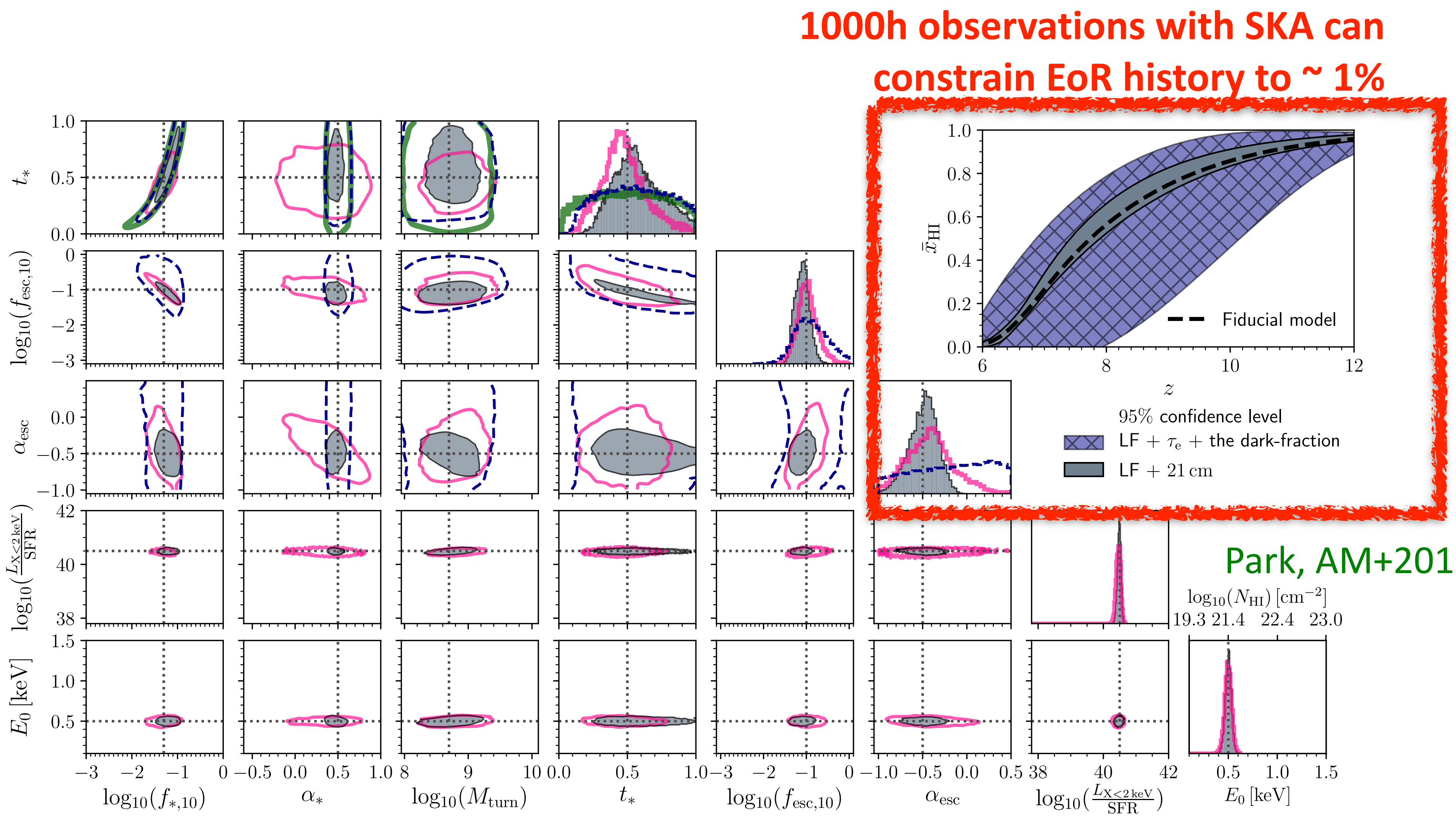
Parameter constraints: LF + 21cm



We will be able to "detect" the dominant galaxy population, isolating the minimum scale for galaxy formation

Park, AM+2018

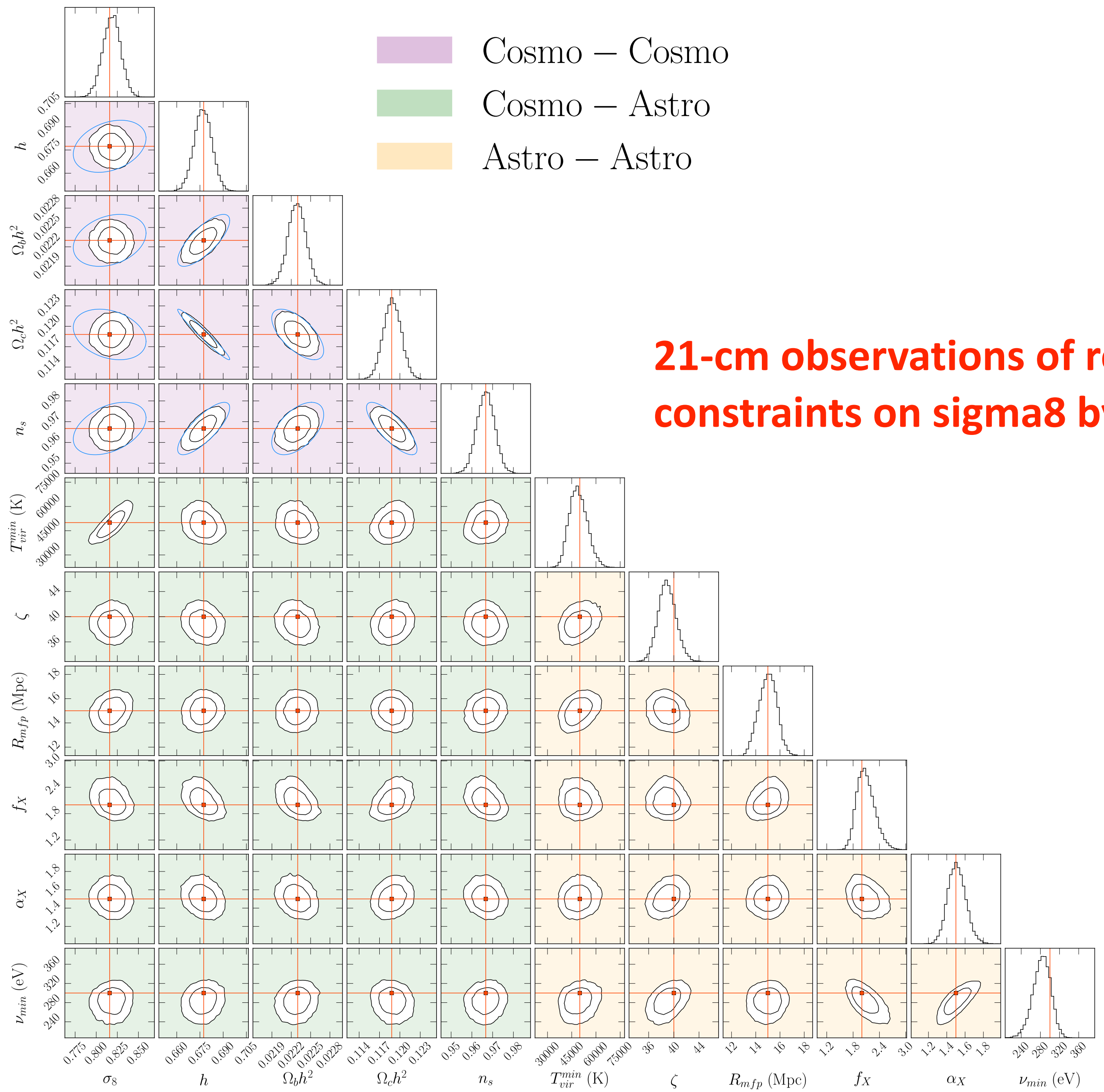
Parameter constraints: LF + 21cm



Park, AM+2018

In addition to the first astrophysical sources, the Cosmic Dawn also tells us about *physical cosmology*

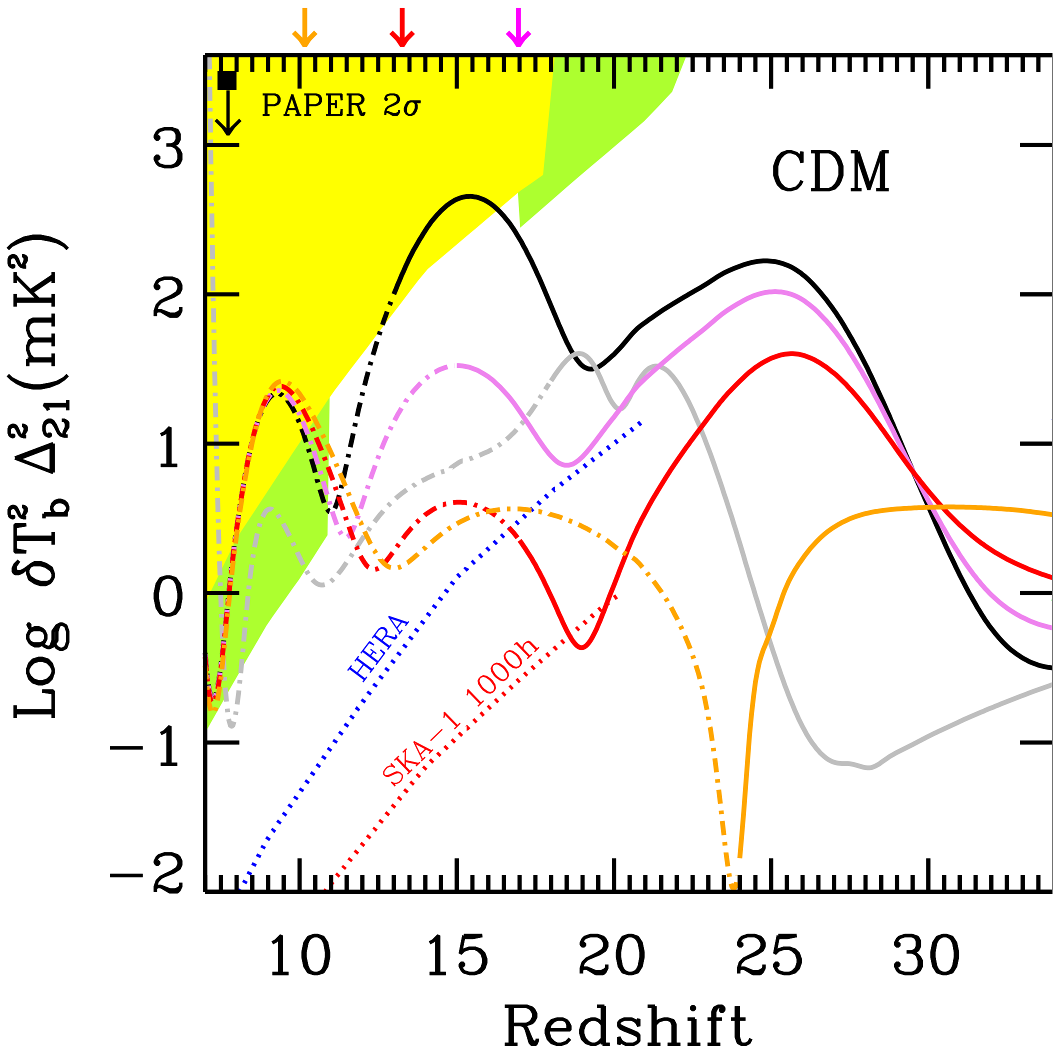
Including physical cosmology



21-cm observations of reionization can improve constraints on sigma8 by a factor of ~2

Kern + (2017)

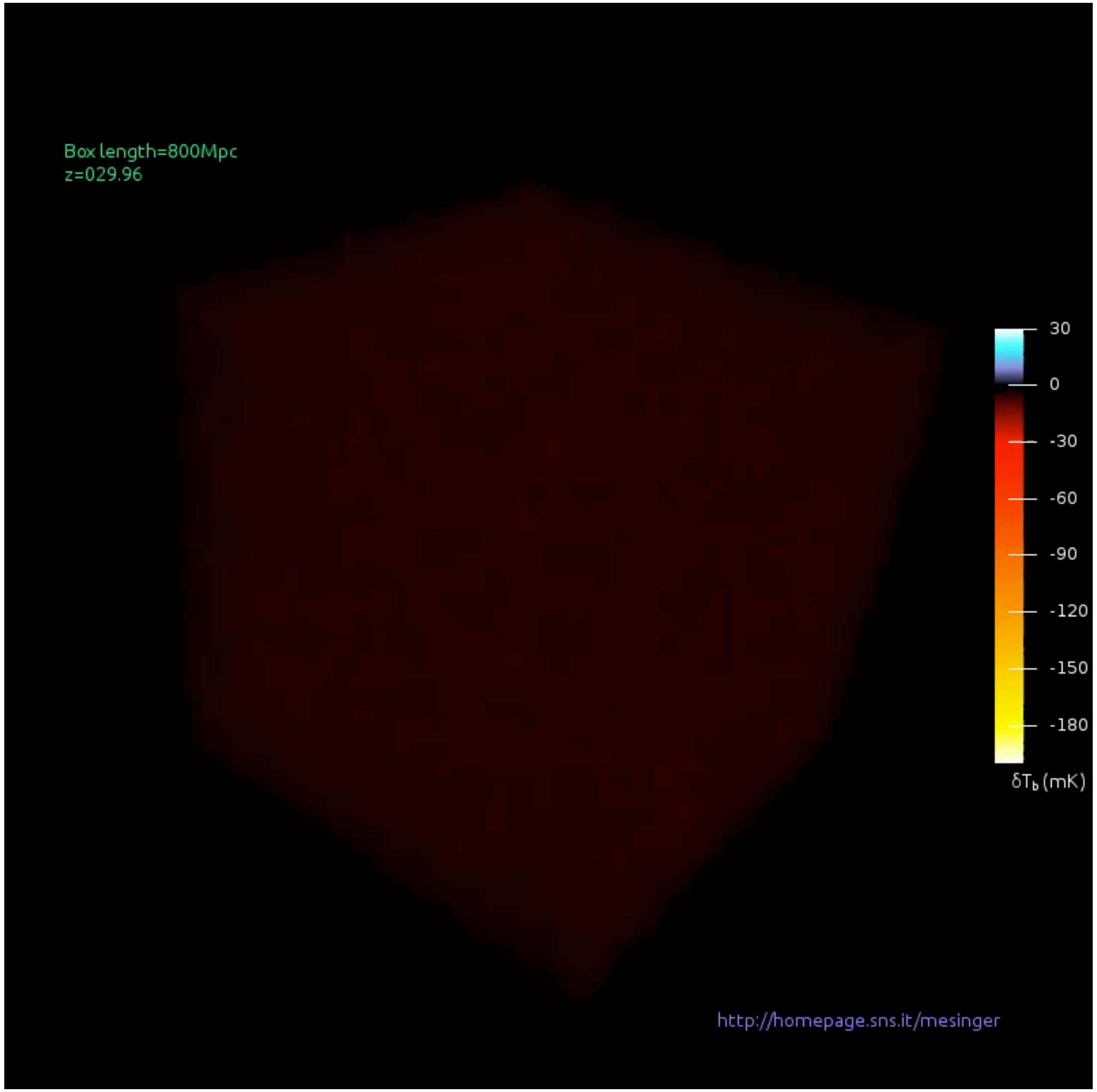
DM heating is more uniform than astrophysical -> heating peak is LOWEST of the three



Peak is in **emission**!
Cannot be reproduced with astrophysics!!!

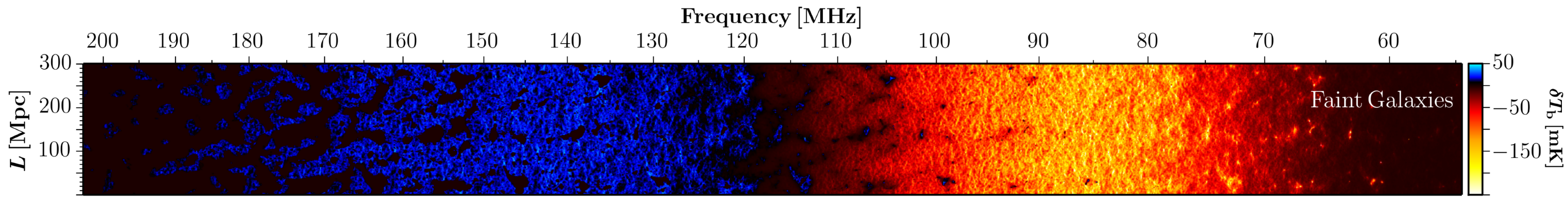
Evoli, AM, Ferrara (2014)
see also Valdez+ (2013)
Lopez-Honorez+2016

SKA’s revolutionary role will be in imaging the first billion years of our Universe

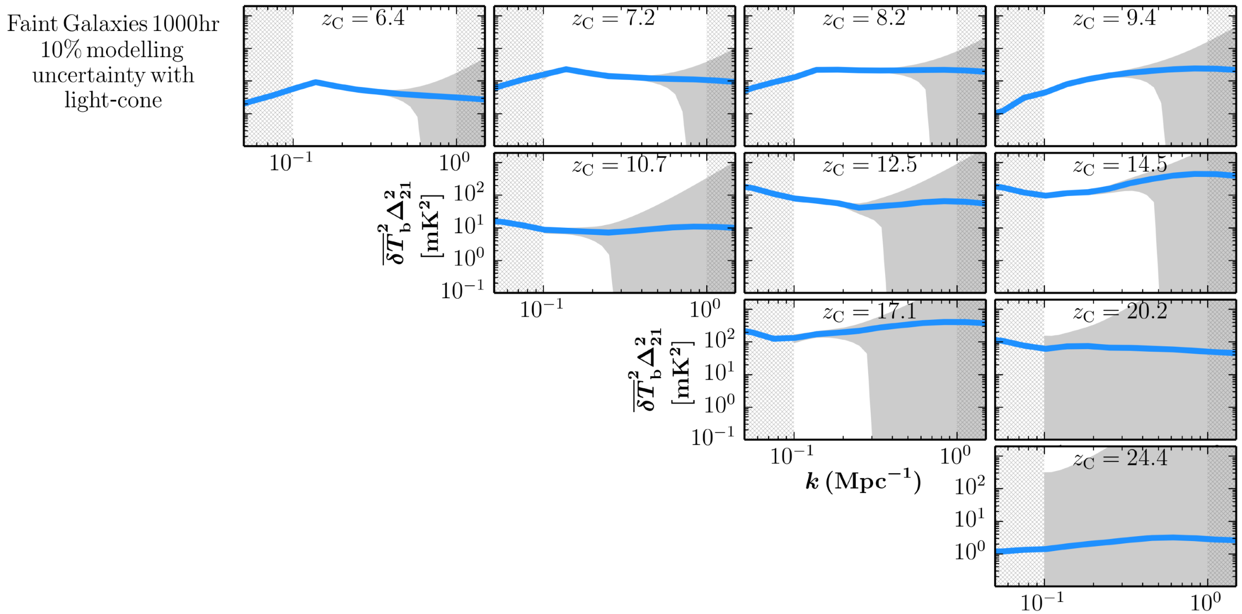


http://*homepage.sns.it/*
mesinger/EOS.html

Astrophysical cosmology



power spectrum**??**



The 21cm signal is **highly non-Gaussian.** Using only the power spectrum **wastes a lot of information!!!**

Greig & AM 2015; 2017

Exploring non-Gaussian statistics

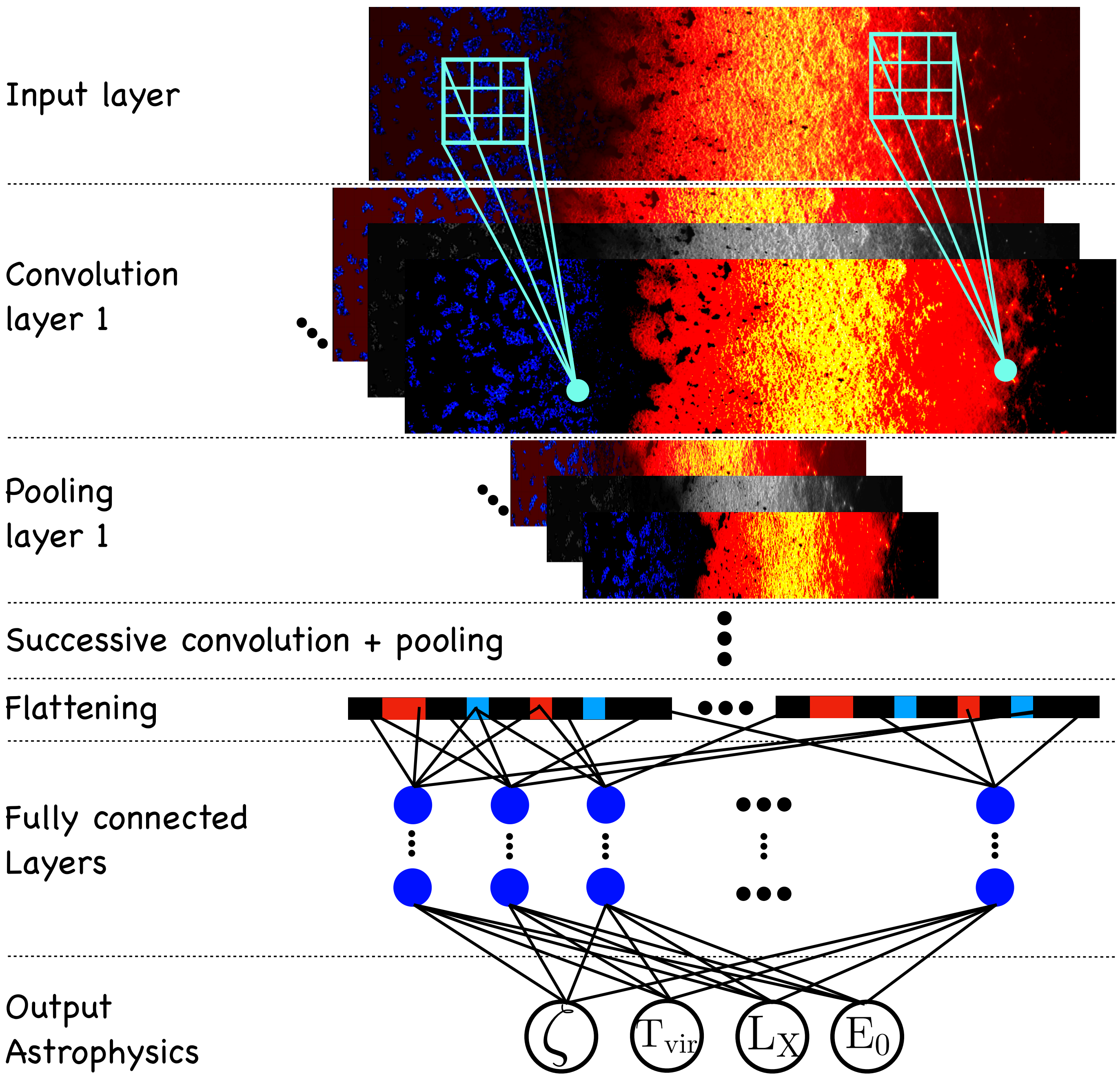
1. **"Brute force" approach:** Simply replace the power spectrum in the likelihood calculation of 21CMMC with an alternate statistic, e.g. the bispectrum (Watkinson, AM+, in prep). *Does that statistic yield tighter constraints on the astrophysical parameters?* Repeat with other statistics, quantifying which one results in the strongest constraints.

Exploring non-Gaussian statistics

1. **"Brute force" approach:** Simply replace the power spectrum in the likelihood calculation of 21CMMC with an alternate statistic, e.g. the bispectrum (Watkinson, AM+, in prep). *Does that statistic yield tighter constraints on the astrophysical parameters?* Repeat with other statistics, quantifying which one results in the strongest constraints.

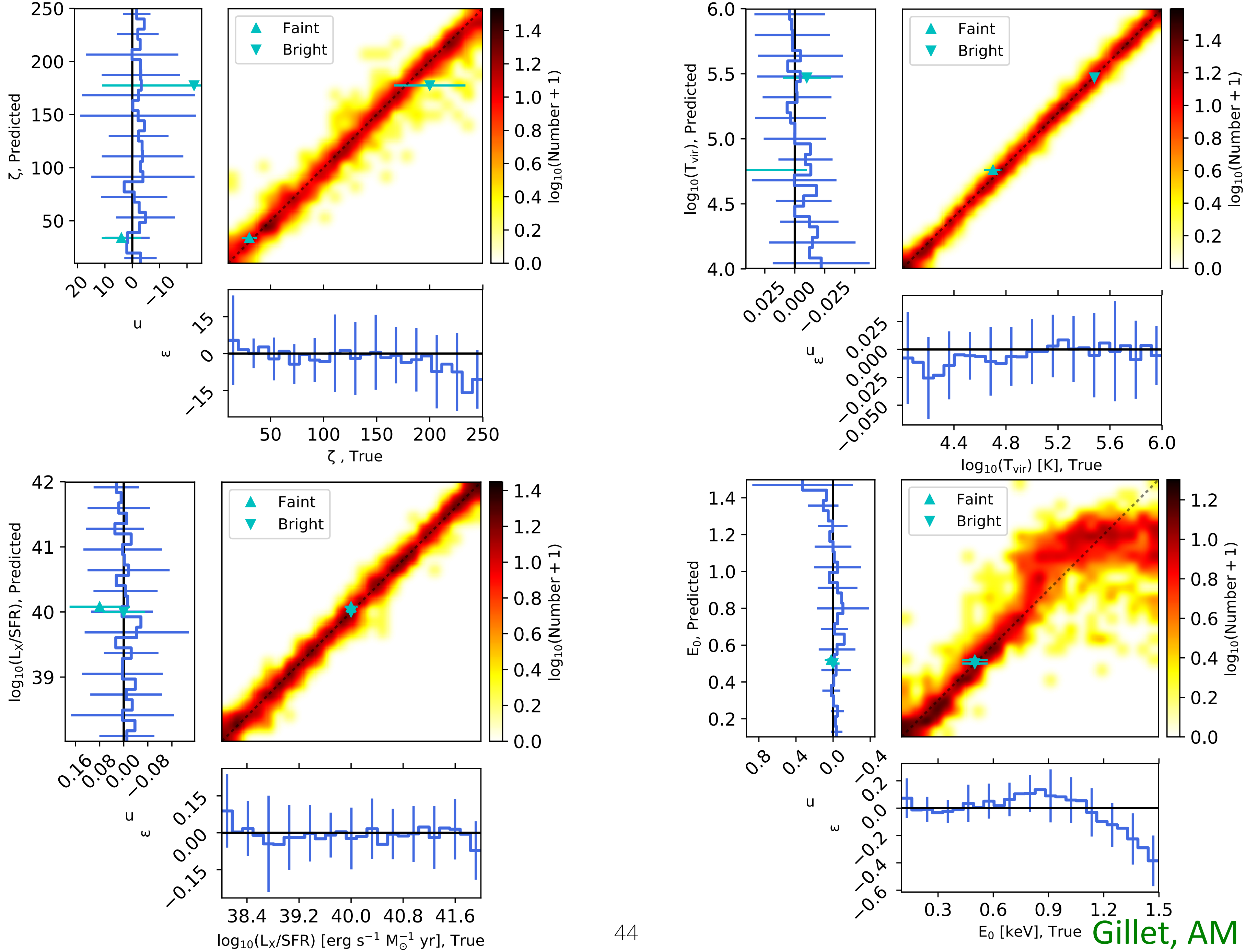
2. **Machine learning approach**: train Convolutional Neural Networks (CNN) to learn astrophysics and cosmology directly from 21-cm images (Gillet et al. 2018).

Deep learning with CNN: parameter recovery



Gillet, AM +, 2018

Deep learning with CNN: parameter recovery



Gillet, AM +, 2018

Italy’s role in EoR/CD science with SKA (an incomplete list)

- Current co-chair of the SKA EoR/CD WG and two WG board members are from Italian institutions

Italy’s role in EoR/CD science with SKA (an incomplete list)

- Current co-chair of the SKA EoR/CD WG and two WG board members are from Italian institutions
- EoR/CD science pulls together many fields, well represented in Italy, e.g. *galaxy evolution, IGM physics, X-ray sources (HMXBs, faint AGN), astroparticle physics*

Italy’s role in EoR/CD science with SKA (an incomplete list)

- Current co-chair of the SKA EoR/CD WG and two WG board members are from Italian institutions
- EoR/CD science pulls together many fields, well represented in Italy, e.g. *galaxy evolution, IGM physics, X-ray sources (HMXBs, faint AGN), astroparticle physics*
- The main simulation and inference tools for SKA (and all 21-cm interferometers are ***21cmFAST + 21cmMC***

21cmFAST

MADE IN ITALY

21cmFAST is being used by all of the 21cm interferometers, with researchers in 18 countries studying a broad range of early Universe topics

Conclusions / Upcoming….

- Current probes tell us roughly when reionization occurred. But we know very little about the unseen, faint galaxies thought to dominate reionization and heating.
- SKA will chart the first billion years of our Universe, revolutionizing the field. The properties of sources and sinks are encoded in the 3D EoR structure.
- To quantify what we can learn, we developed a Bayesian framework for astrophysical parameter estimation, capable of on-the-fly MCMC sampling (21CMMC) of 3D simulations (21cmFAST).
- Forecasts using the power spectrum as a summary statistic suggest even an 8 parameter astrophysical model can be constrained to ~10%. We will be able to indirectly study the unseen, dominant galaxy population.
- SKA images of EoR/CD are non-Gaussian… What are optimal summary statistics? We can explore this using neutral networks
- *The next decade will see the advent of **precision astrophysical cosmology!***