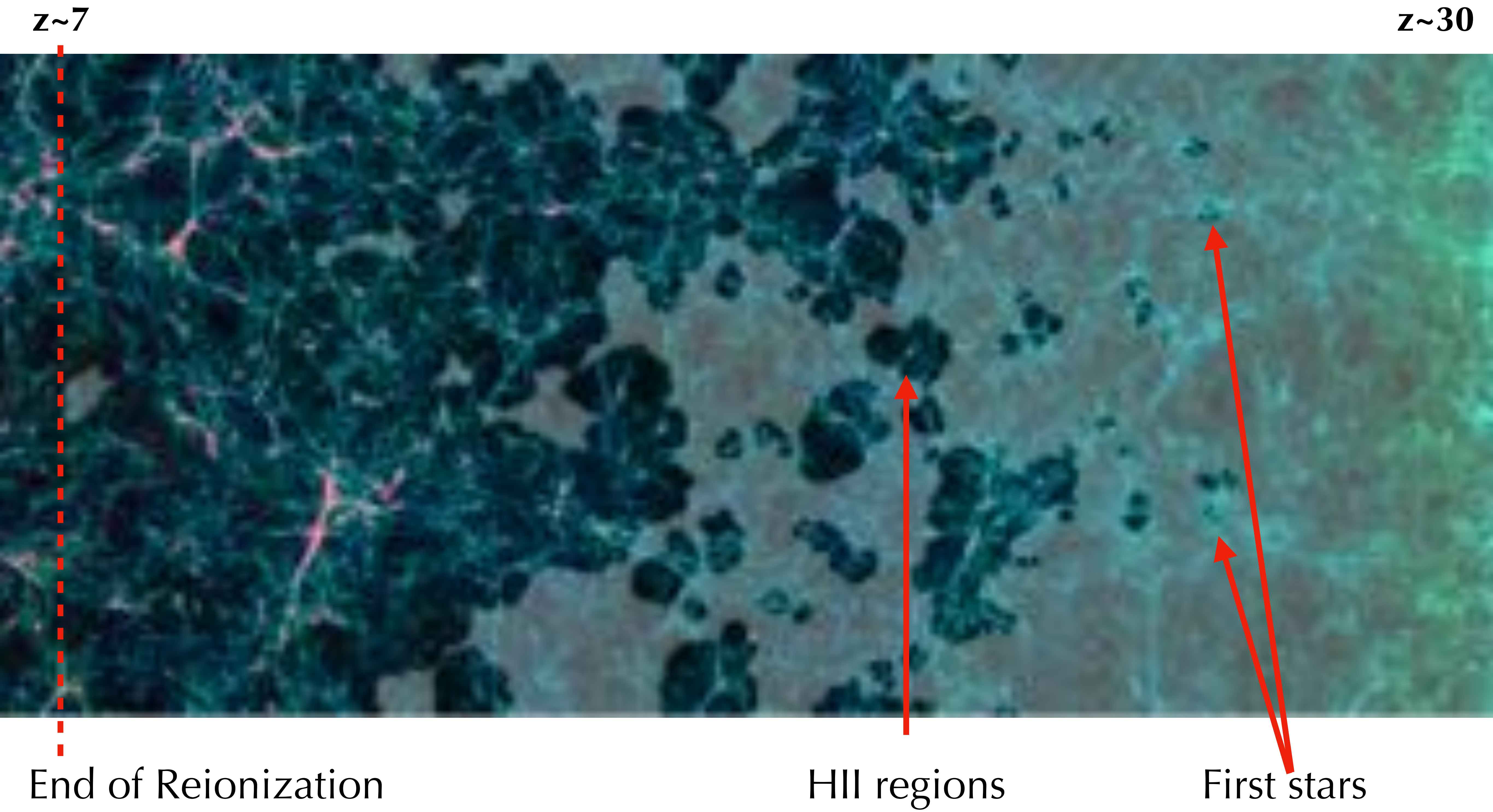
Deep learning from 21-cm images of the Epoch of Reionization

Nicolas Gillet in collaboration with **Andrei Mesinger** and **Bradley Greig**
Scuola Normale Superiore

SF2A - Bordeaux 2018

The first billion years: the Epoch of Reionization

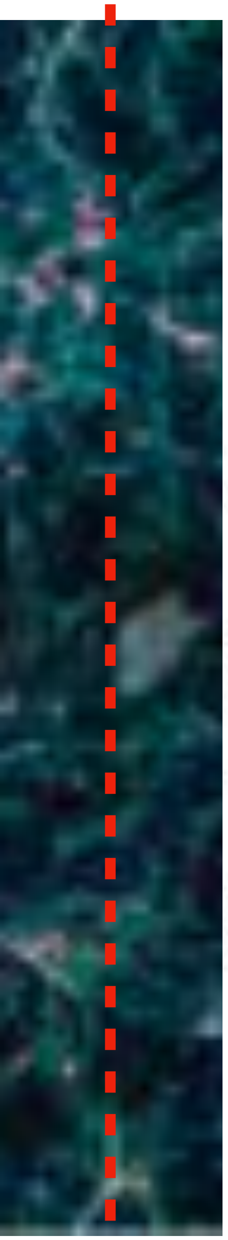
Credit: N.Deparis/EMMA



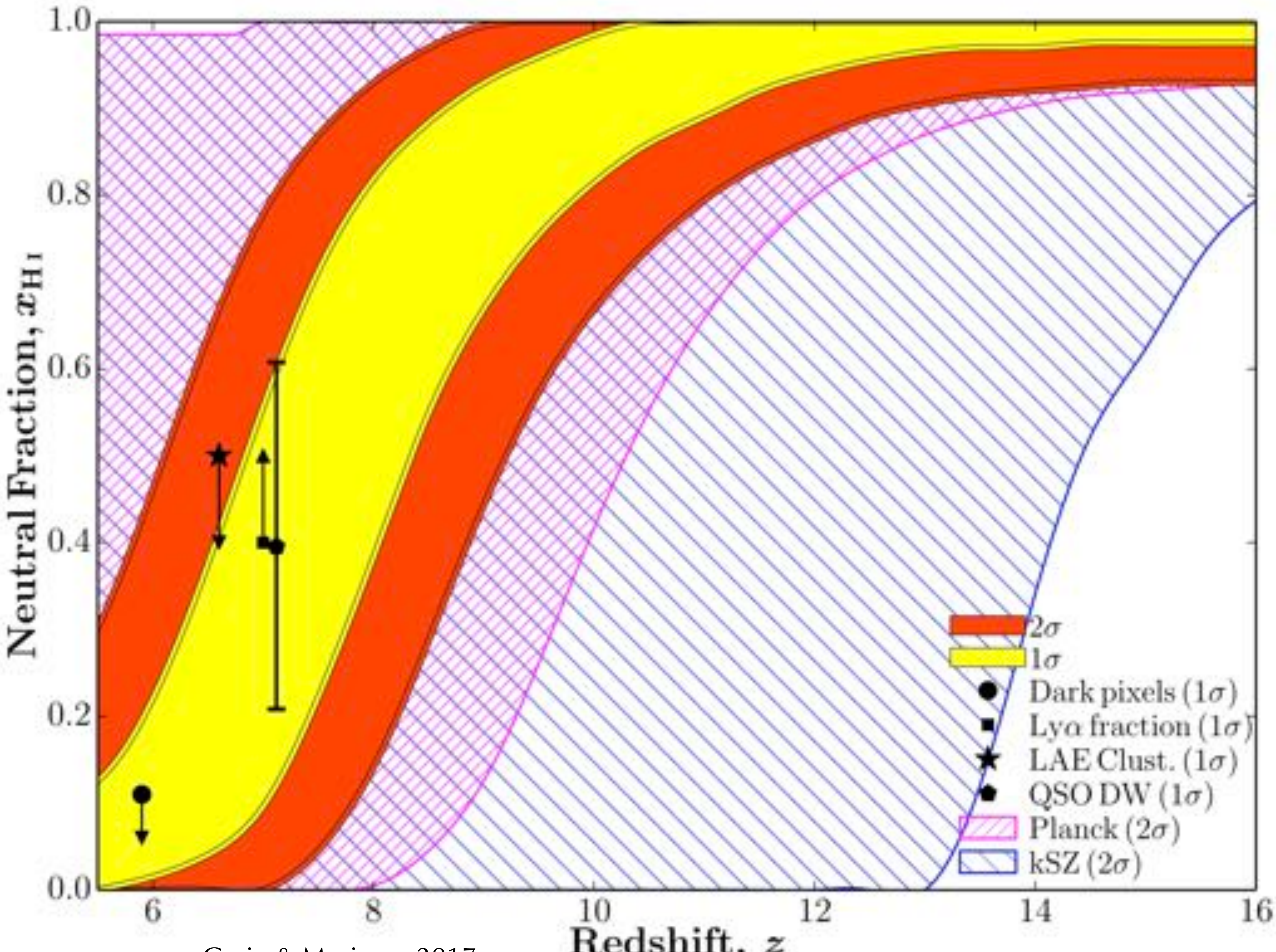
The first billion years: the Epoch of Reionization

aris/EMMA

z~7



End c

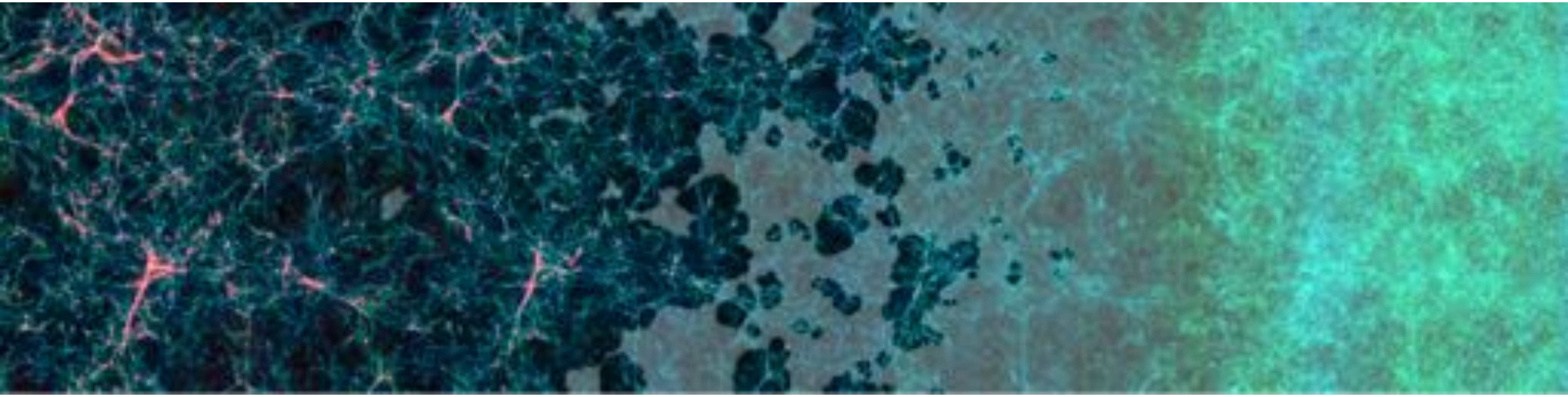


Greig & Mesinger 2017

z~30

The Epoch of Reionization

Credit: N.Deparis/EMMA



- What are the main source of the Reionization ?
    -> Quasar ? Galaxies ?

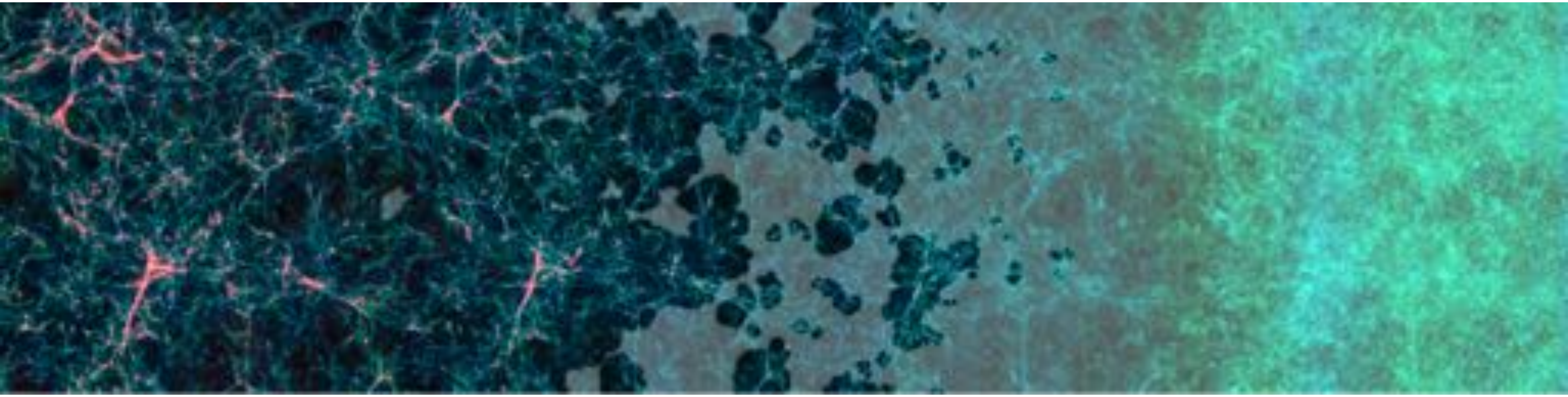
- What are the properties of those sources ?
    -> Number of quasars ? Rare big galaxies ? Numerous light galaxies ?

- How do they interact with each other and the IGM ?
    -> Radiative feedback ? Quenching of small galaxies ?

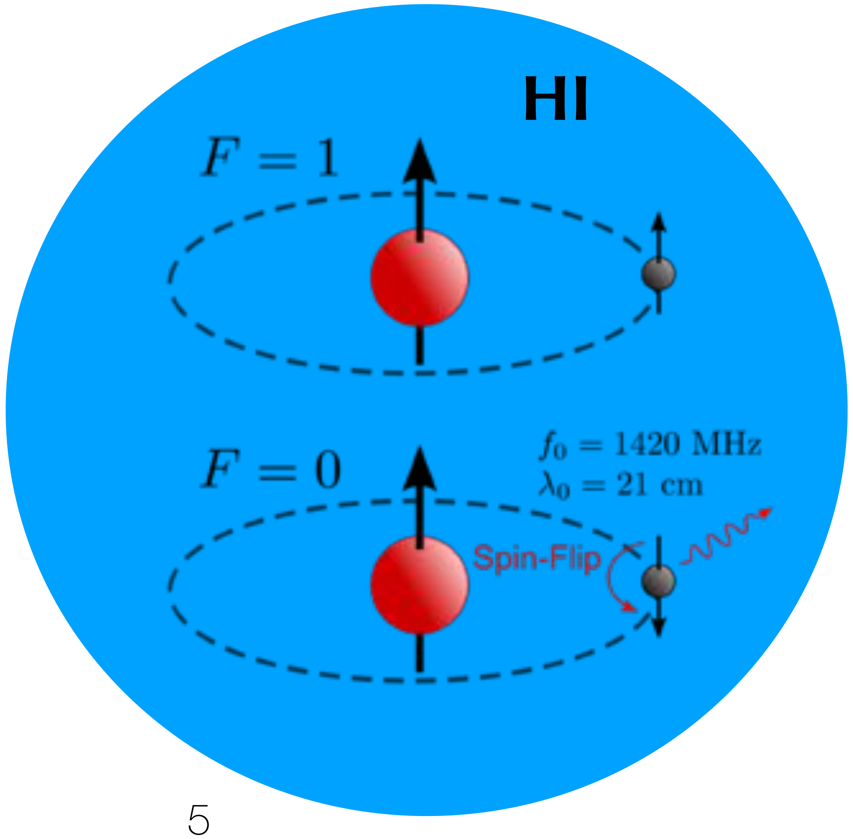
Observe the EoR

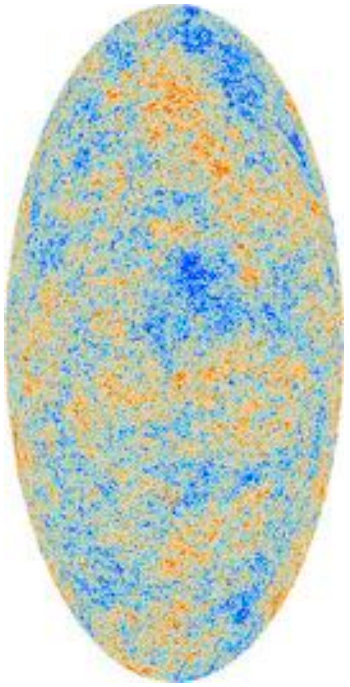
Credit: N.Deparis/EMMA

z~6 **z~50**



Large interferometers
Hera, SKA

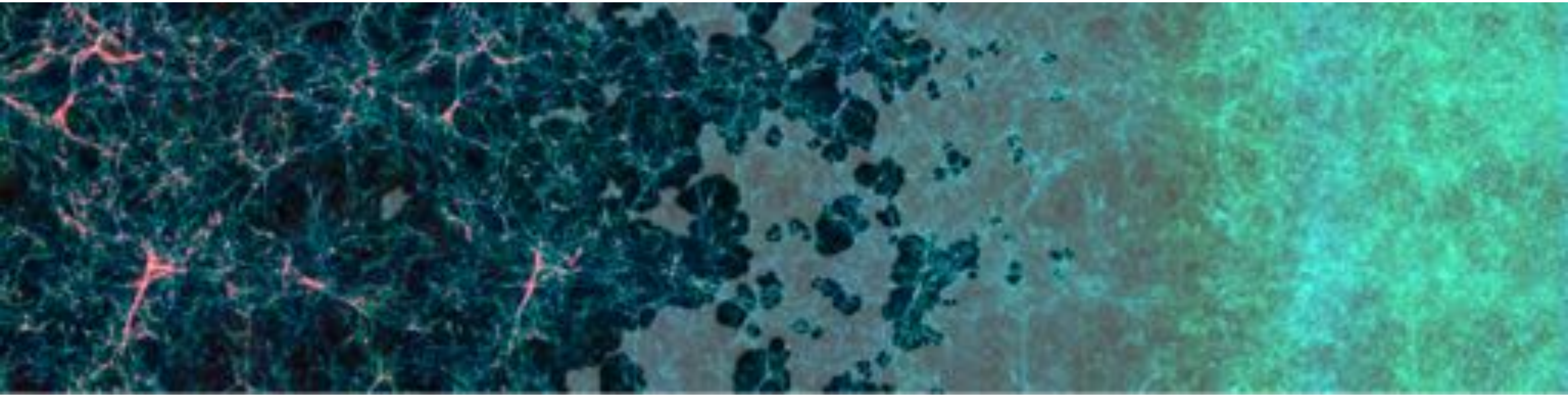


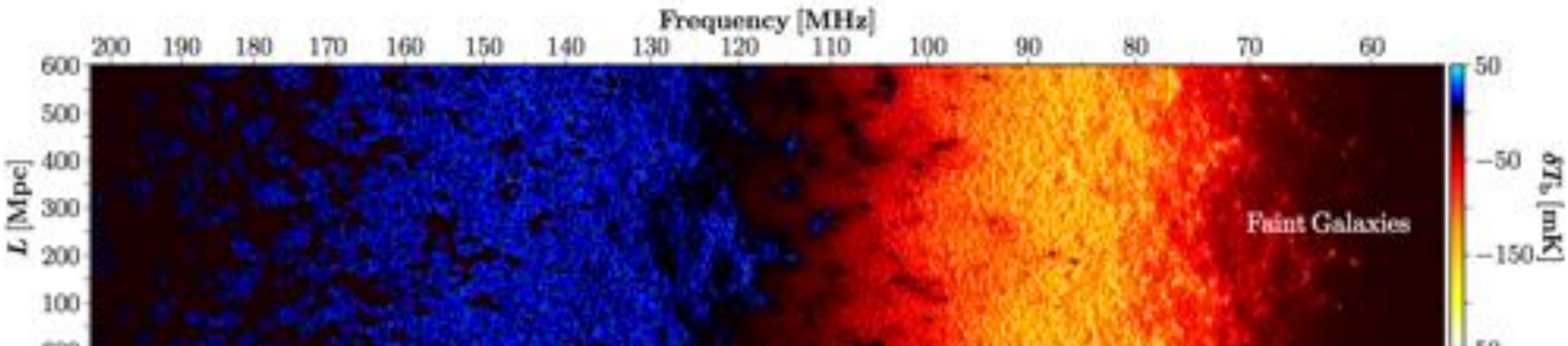


Observe the EoR

Credit: N.Deparis/EMMA

z~6 **z~50**



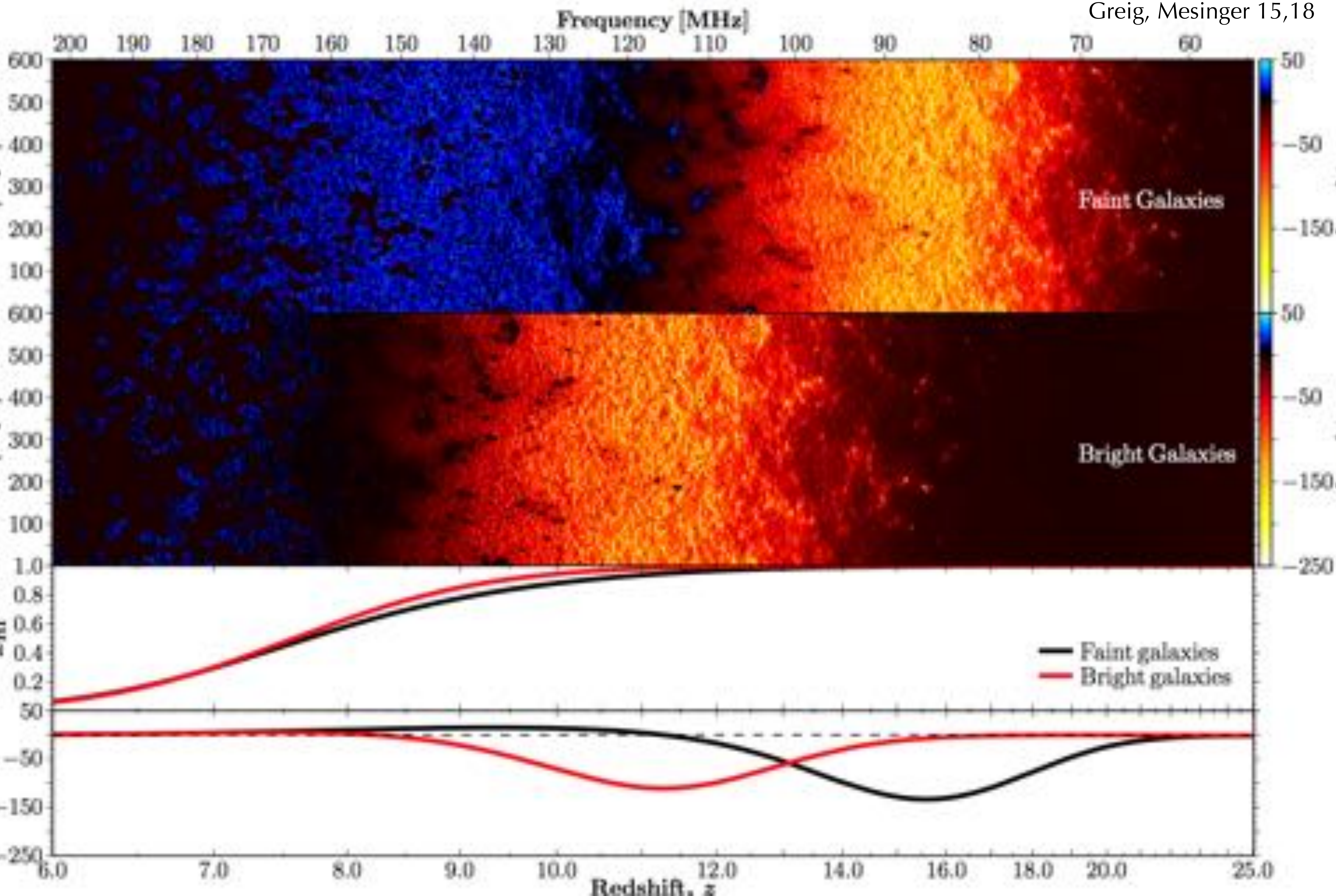


- galaxy formation
- UV emission

- X-ray emission
- X-ray spectrum

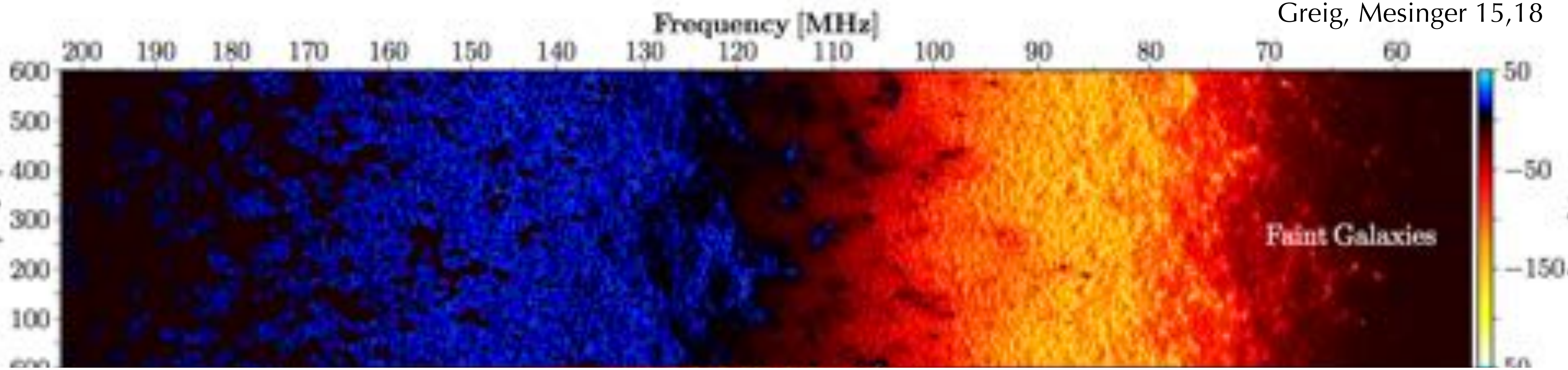
The 21cm signal

Greig, Mesinger 15,18



Model the 21cm signal

Greig, Mesinger 15,18



Semi-numerical code: 21cmFast with 4 parameters:

ζ Ionizing efficiency

T_{vir} Minimum virial temperature of star-forming halos

L_X X-ray luminosity

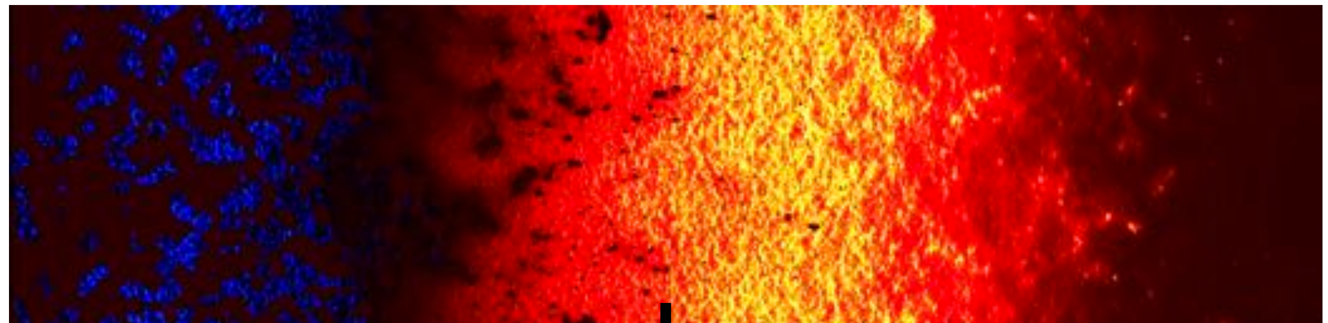
E_0 X-ray spectrum

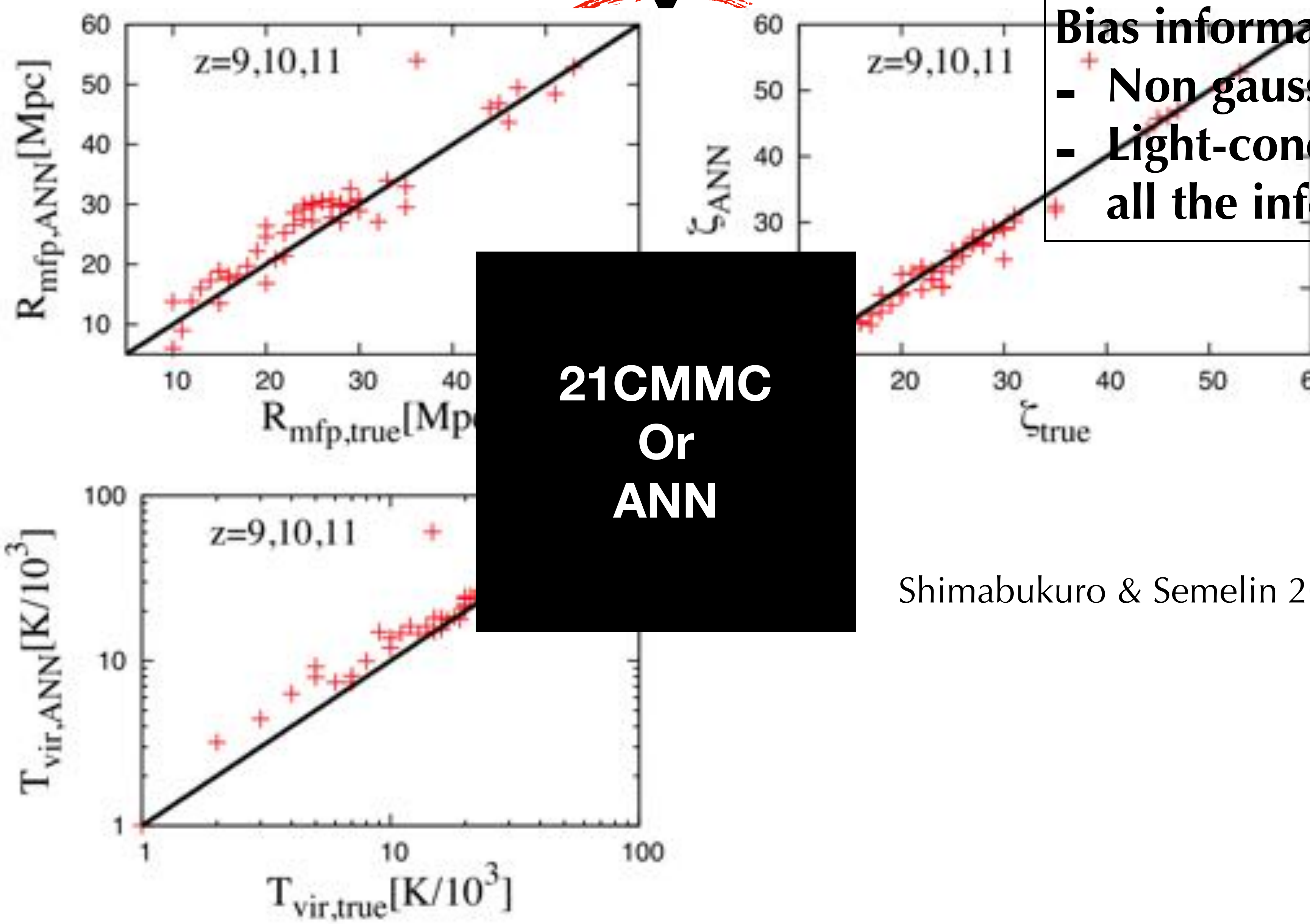
Extract information from Observations: Parameter inference

Parameters inference

« Observed » light-cone 21cm brightness temperature

Greig, Mesinger 15,18





21CMMC
Or
ANN

Bias information (?)

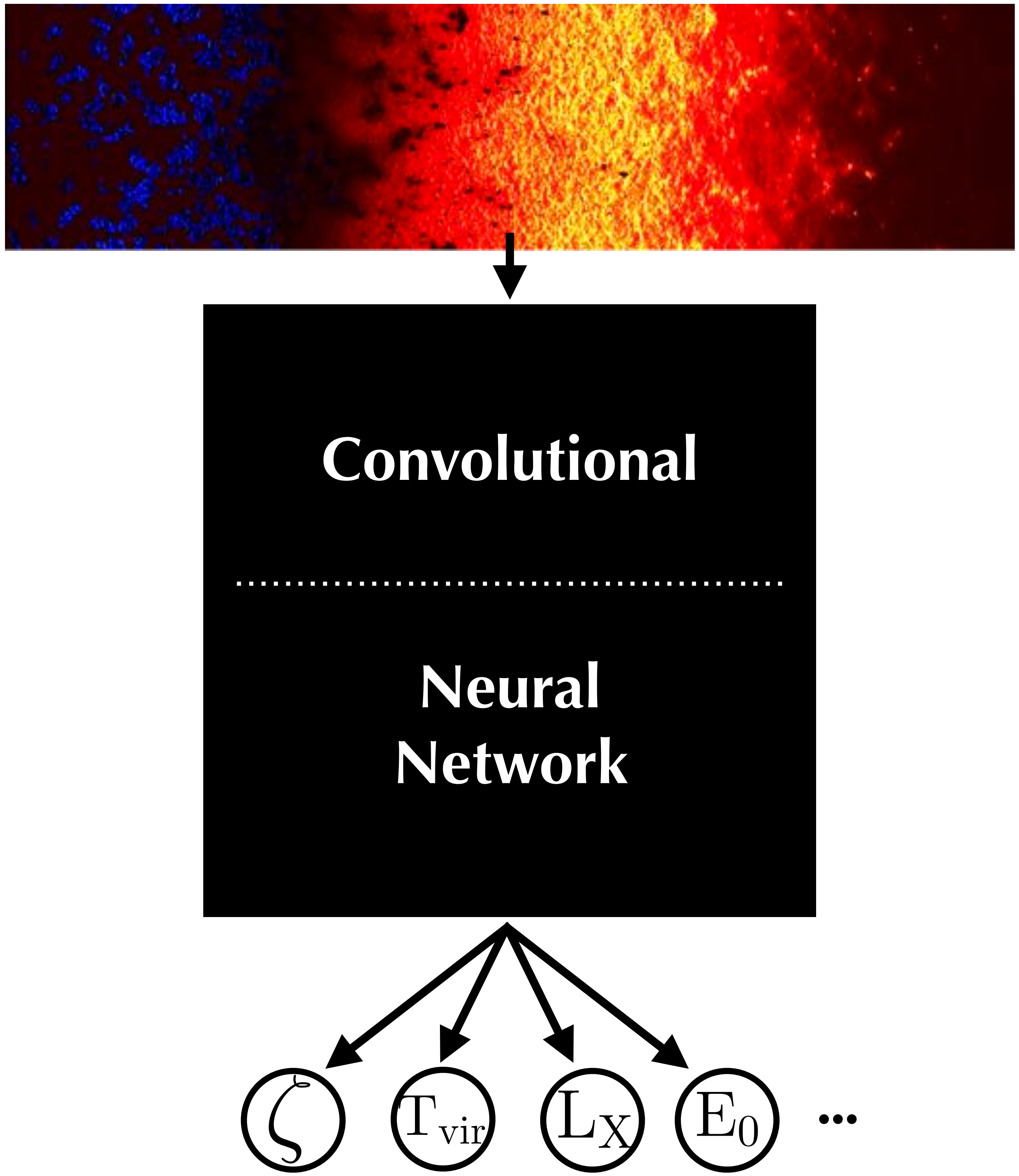
- **Non gaussian signal!**
- **Light-cone contains all the informations**

Shimabukuro & Semelin 2017

Parameters inference

« Observed » light-cone 21cm brightness temperature

Greig, Mesinger 15,18



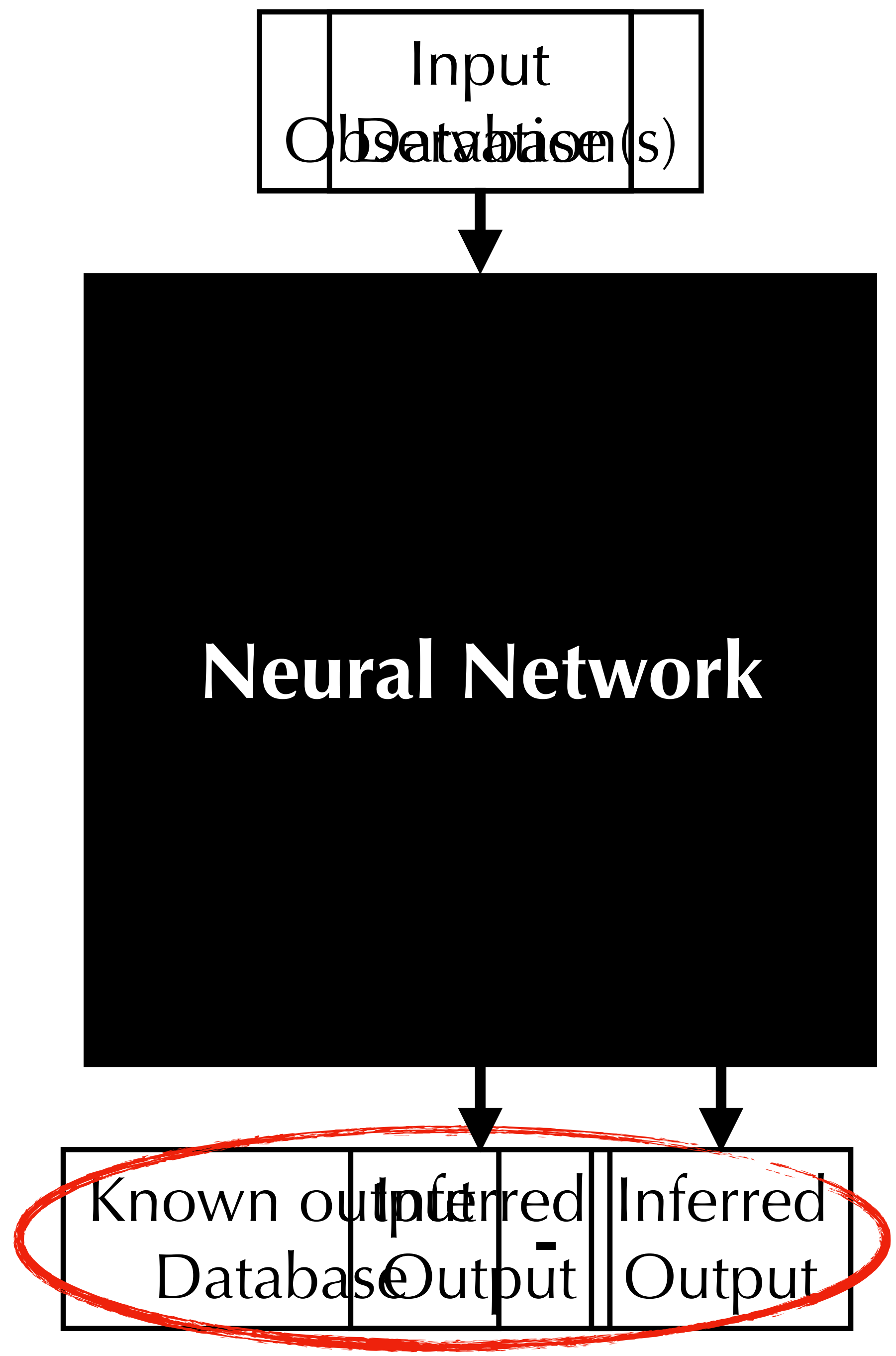
Machine Learning
Neural Network

Neural Network

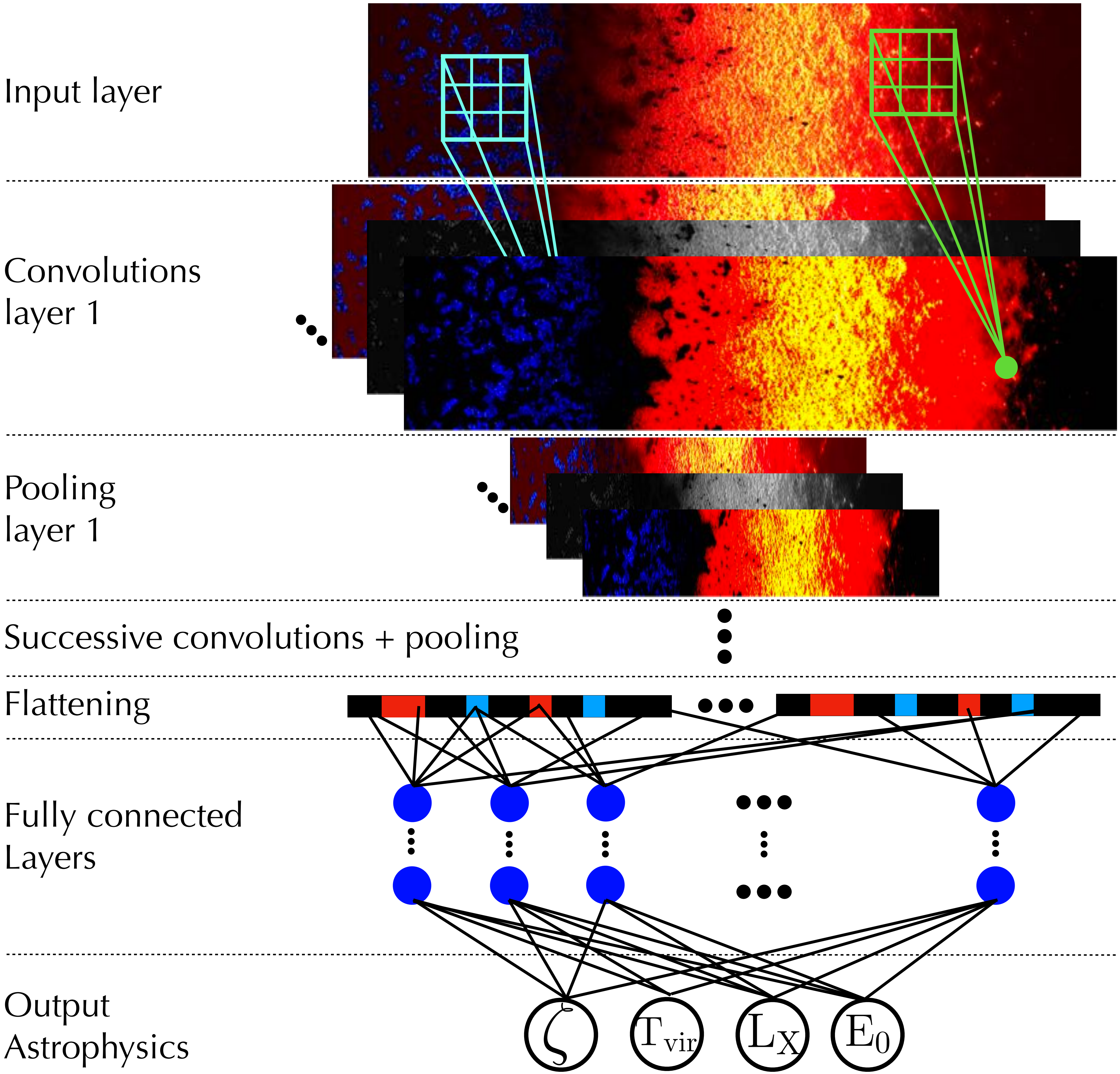
- NN contains a large number of free weights.
- Adapt the weight to minimize the « loss »
- Iterative learning over the database.

Require:
- A large database.
- Lots of « Hyper-parameters » to tune.
- Might be long to converge.
- Overfitting problem.

Recognition phase:
- Infer output given « Observations ».
- Really fast (~s to ~min).



What is a Convolution Neural Network ?



Input layer
Convolutions layer 1
Pooling layer 1
Successive convolutions + pooling
Flattening
Fully connected Layers
Output Astrophysics
ζ
T_{vir}
L_X
E_0

21cm Light-cone Database

21cm Light-cone Database

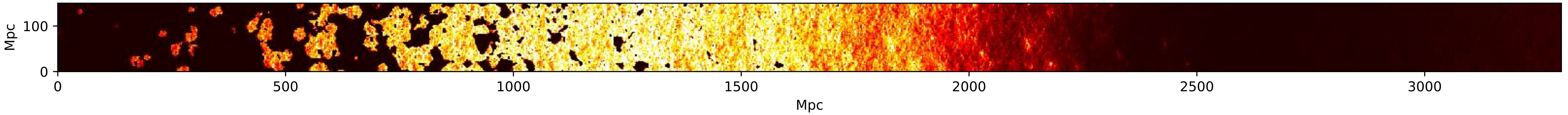
We need the biggest database possible !

- Random distribution of the parameters.
- Different density field for each realization (prevent the CNN to learn the density field).

- 10 000 light-cones,
- using 4 parameters,
- redshift range [6, 30],
- **3D** : (300x300) x 3300 Mpc (long French fries),
- resolution 1.5Mpc (200x200x2200).

Proof of Concept data base:

- Work in **2D** : 150 x 3300 Mpc
- We did not consider any noise.



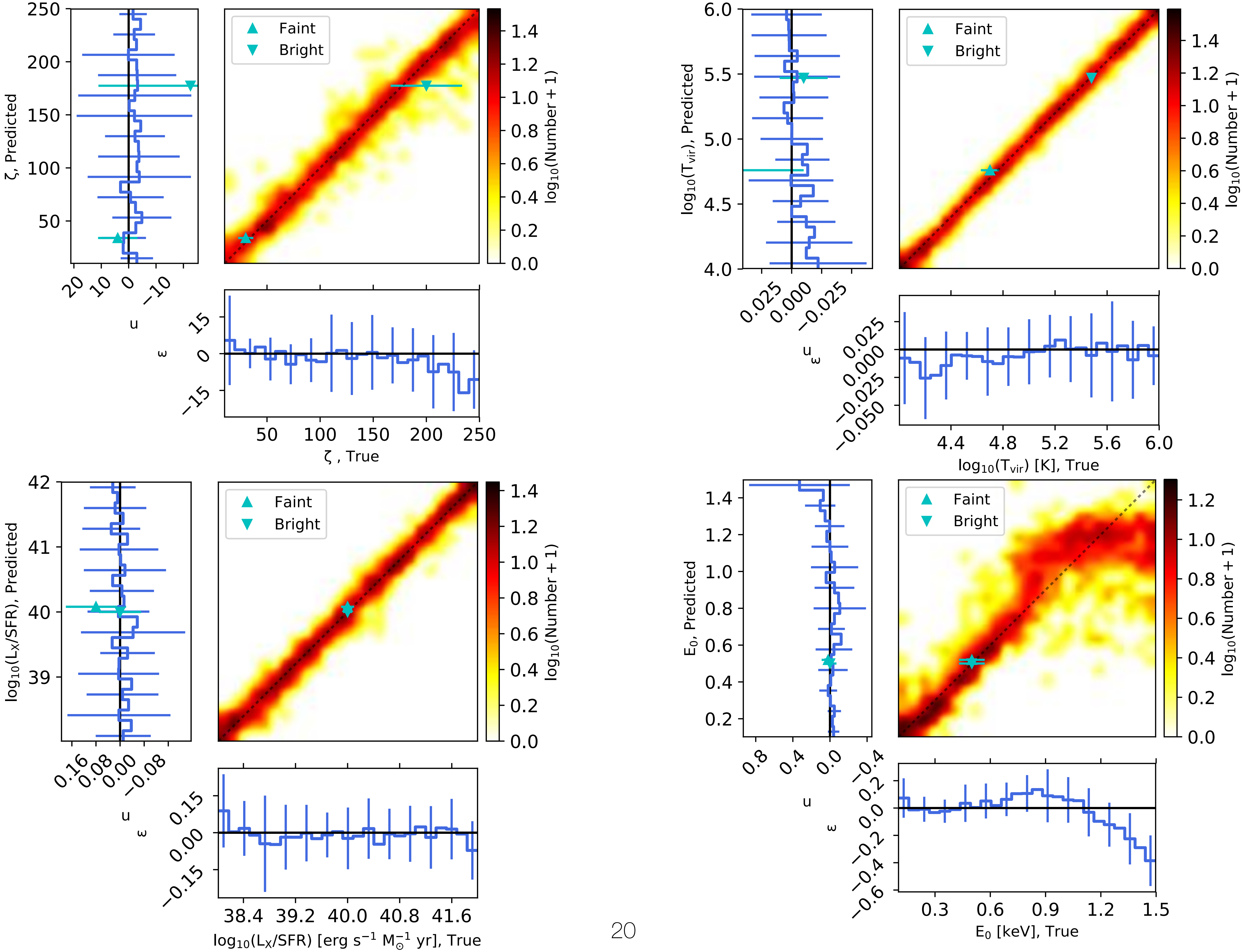
Application & Results

Application: CNN architecture

Layer (type)	Output Shape	Param #
Conv-1 (Conv2D)	(8, 91, 2191)	808
Pool-1 (MaxPooling2D)	(8, 45, 1095)	0
Conv-2 (Conv2D)	(16, 36, 1086)	12816
Pool-2 (MaxPooling2D)	(16, 18, 543)	0
Flat (Flatten)	(156384)	0
Dense-1 (Dense)	(64)	10008640
Dense-2 (Dense)	(16)	1040
Dense-3 (Dense)	(8)	136
Out (Dense)	(4)	36

Total params: 10,023,732

CNN Results

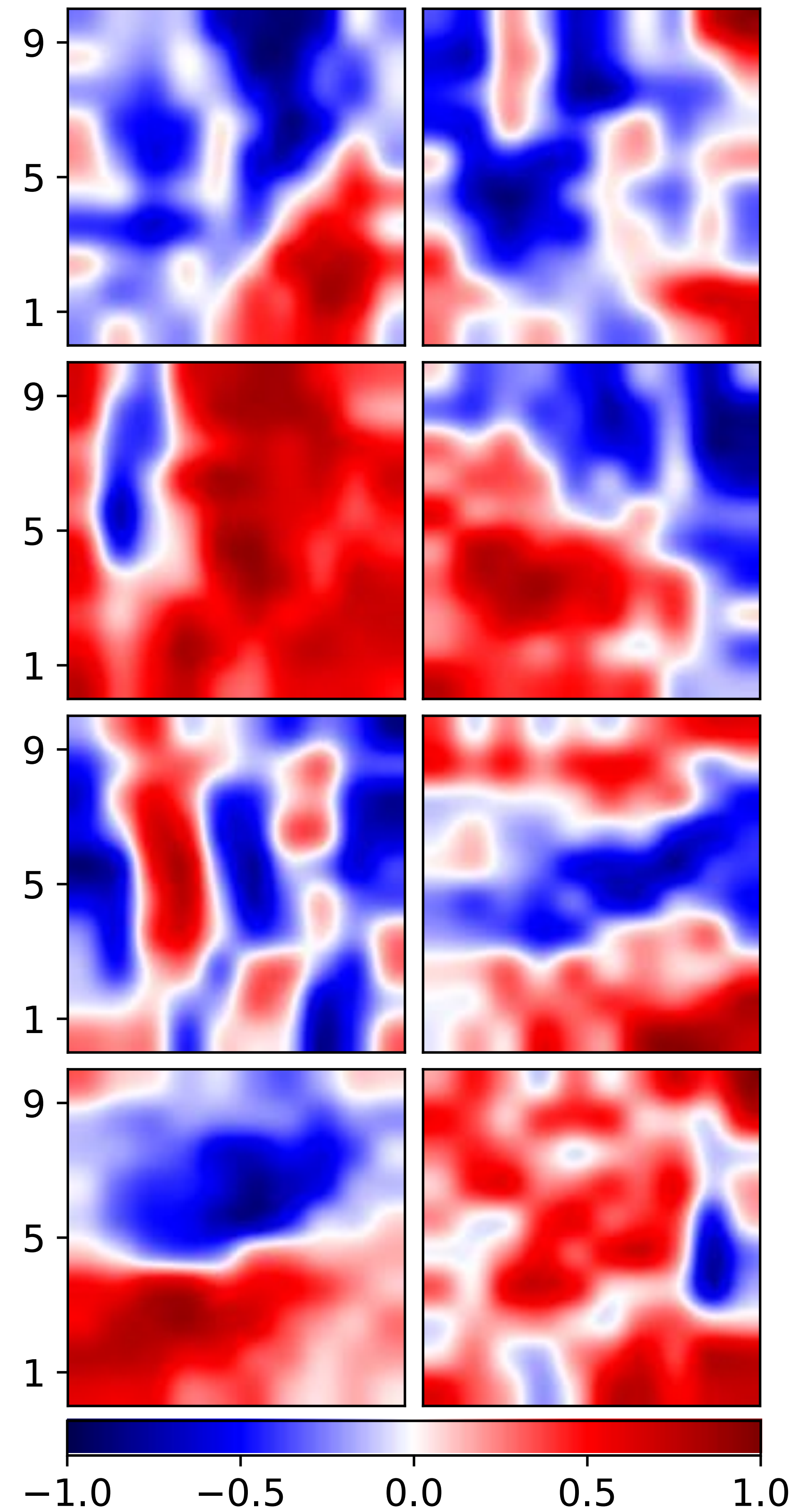


Faint
Bright
ζ, Predicted
ζ , True
log10(Number + 1)
log10(Tvir), Predicted
log10(Tvir) [K], True
log10(LX/SFR), Predicted
log10(LX/SFR) [erg s−1 M⊙−1 yr], True
E0, Predicted
E0 [keV], True
u
ε

CNN's eyes

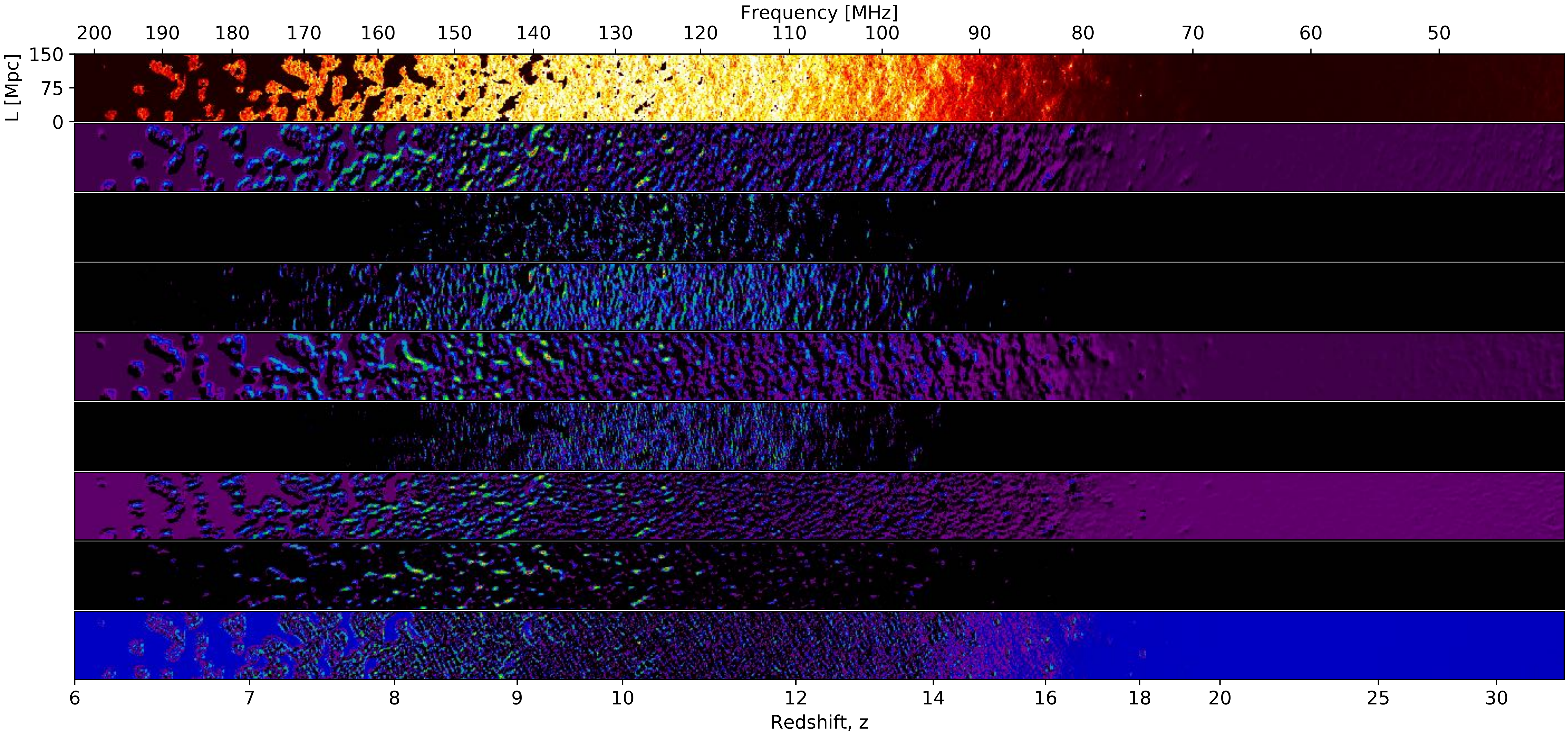
- Filters of the 1st Convolution Layer

- 10x10 pixels (gaussian smooth)



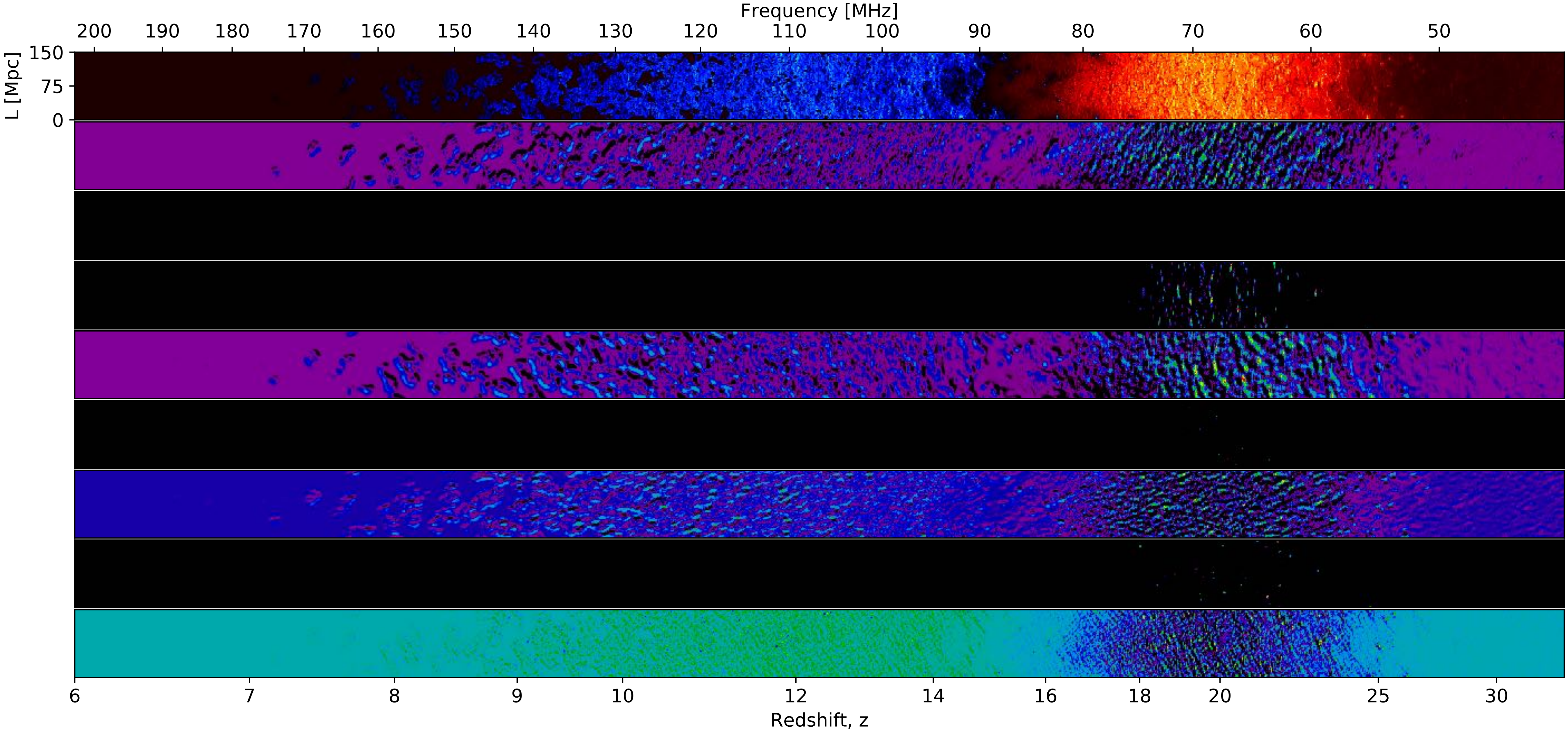
What the CNN sees

z~6 **z~30**

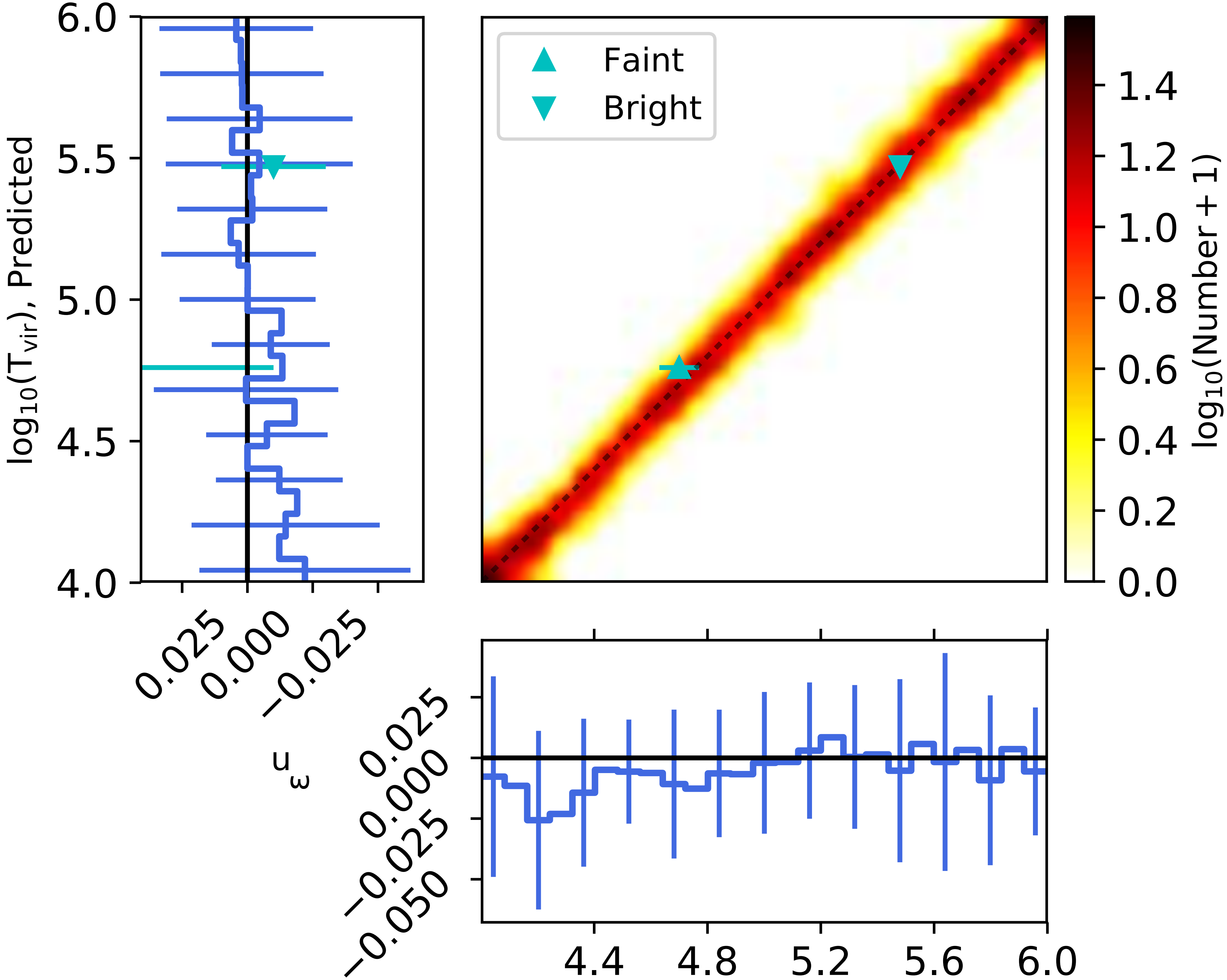


What the CNN sees

z~6 **z~30**



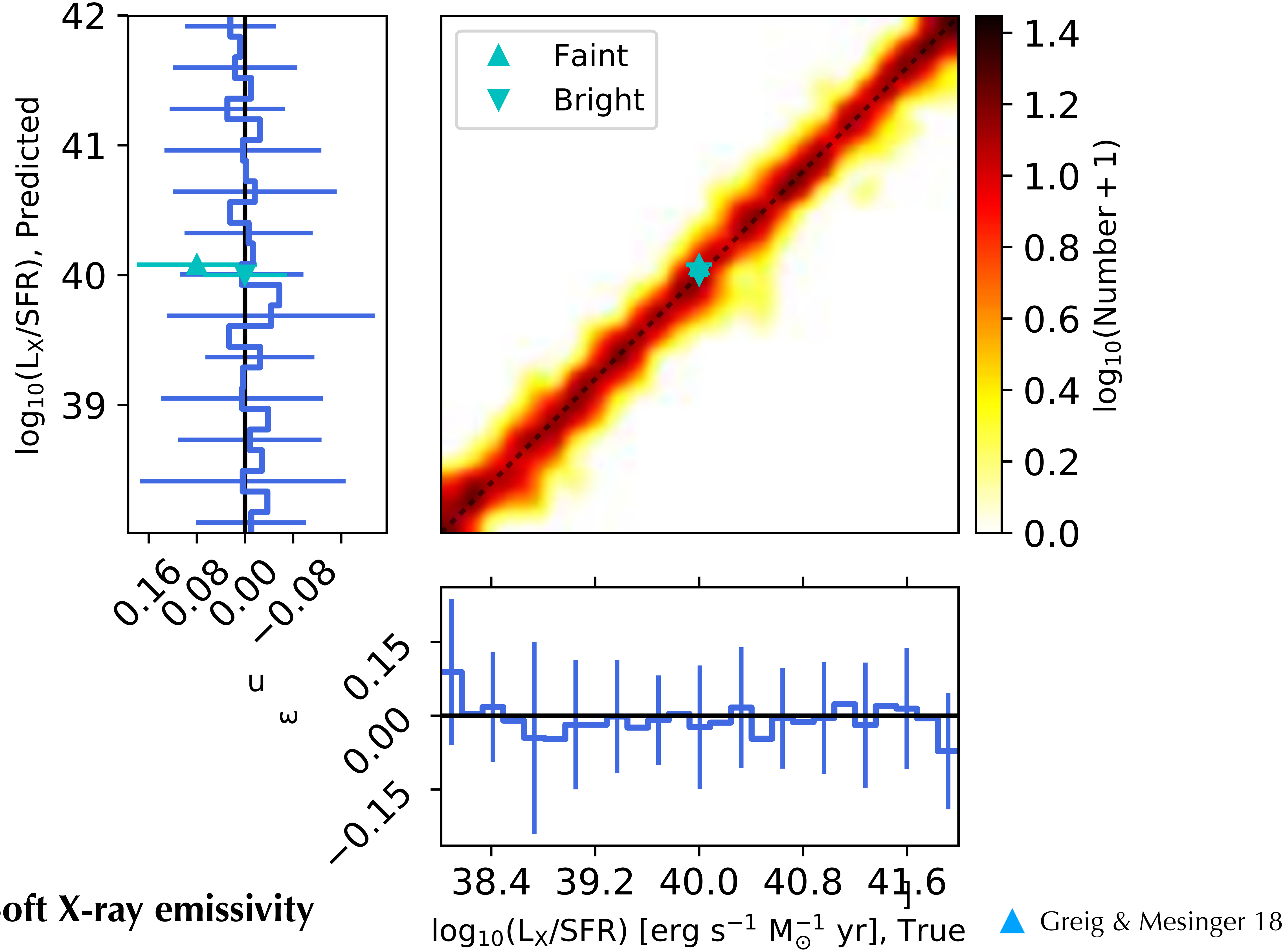
CNN Results



Tvir : Minimum threshold for star formation

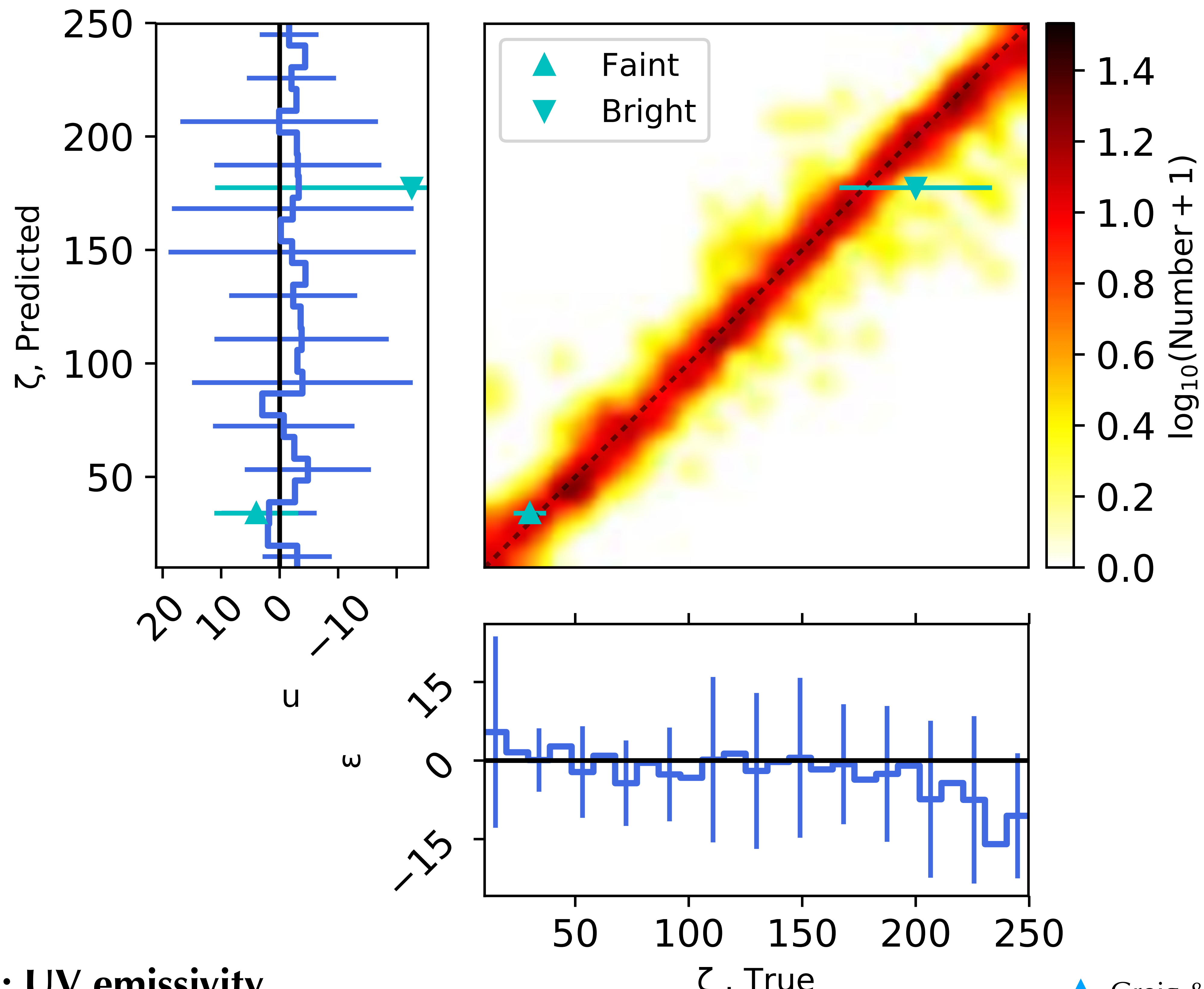
▲ Greig & Mesinger 18

CNN Results



Lx ~ Soft X-ray emissivity

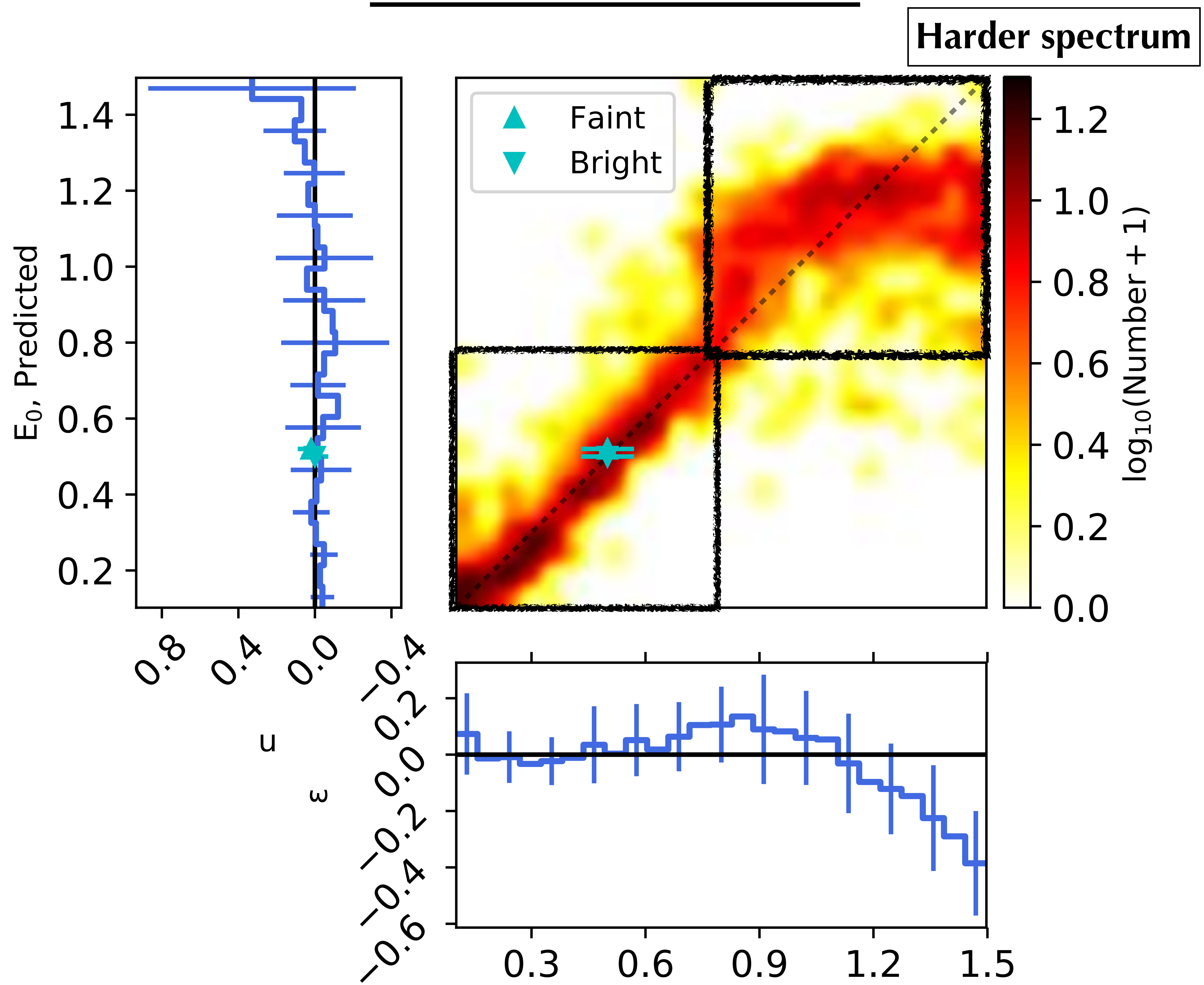
CNN Results



Zeta : UV emissivity

Greig & Mesinger 18

CNN Results



▲ Greig & Mesinger 18

E0 : Minimum X-ray energy escaping ISM ~ X-ray soft/hard spectrum

Conclusions

- CNN can infer parameters directly from 21cm light-cones.

  - pros:
    - Quick inference.
    - Probe the inference on whole parameter space.
    - Use the whole signal -> no need of data reduction

  - cons:
    - Need a database.
    - Tuning/learning of the network.
    - For the moment, not faster than 21CMMC.

  -> Why the CNN does not perform better than MCMC ?
  -> Is it taking advantage of the non-Gaussian information ?
  -> Can we understand the inside of the black-box ?

- TODO:
  - Add Noises.
  - Add more parameters in the model -> new database.