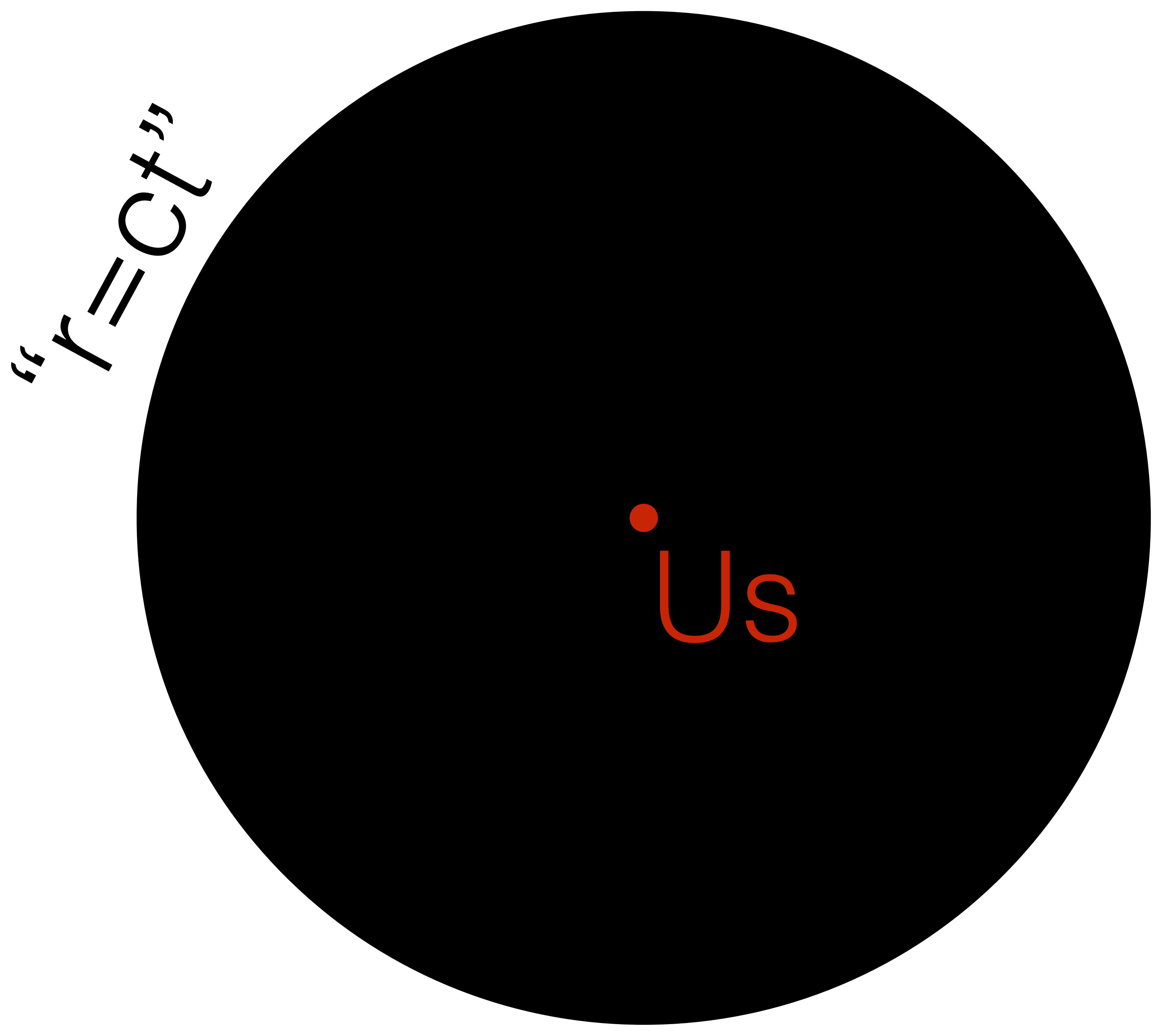
The Past, Present, and Future of 21cm Cosmology

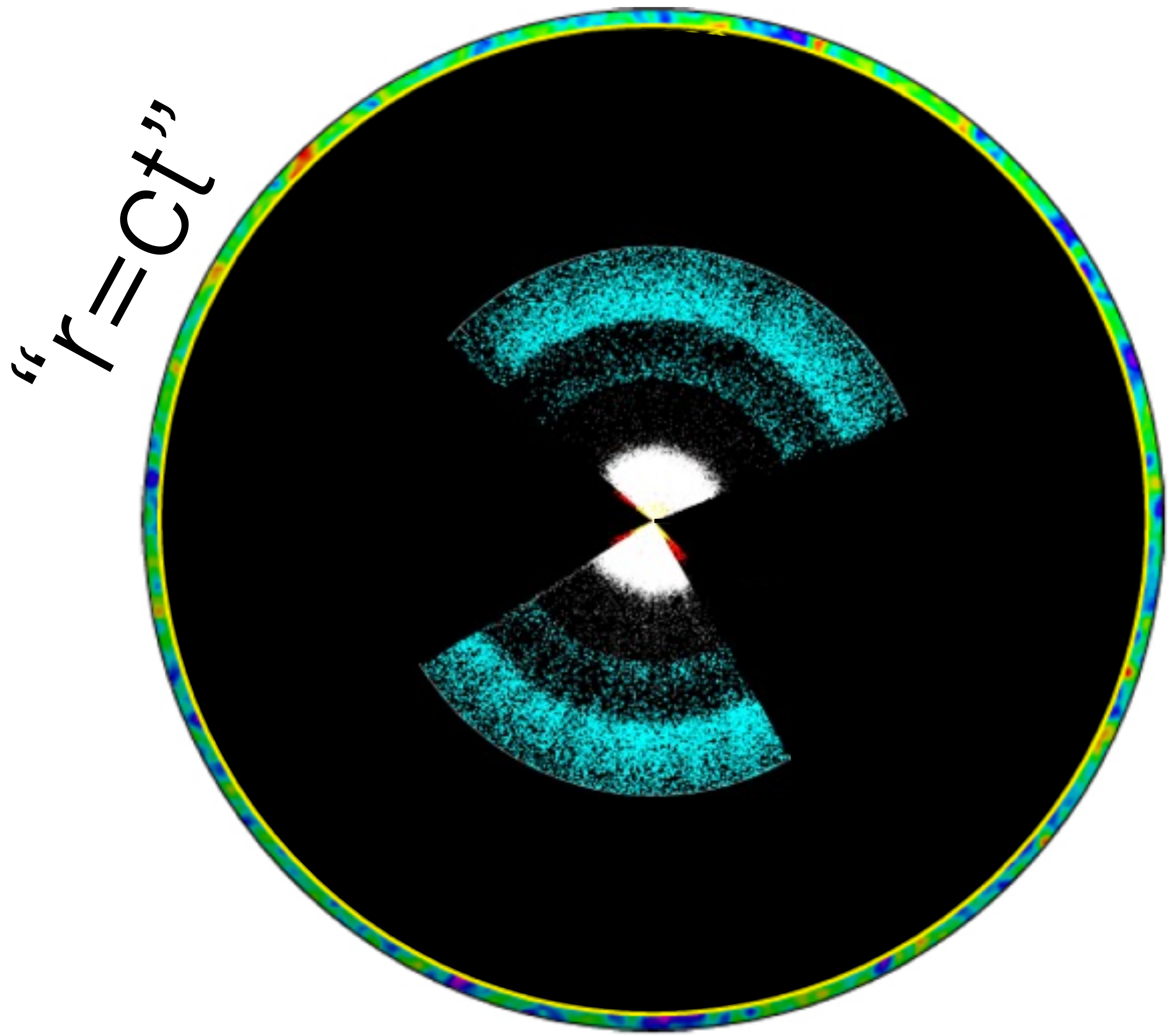
Adrian Liu
Hubble Fellow
UC Berkeley

Vision

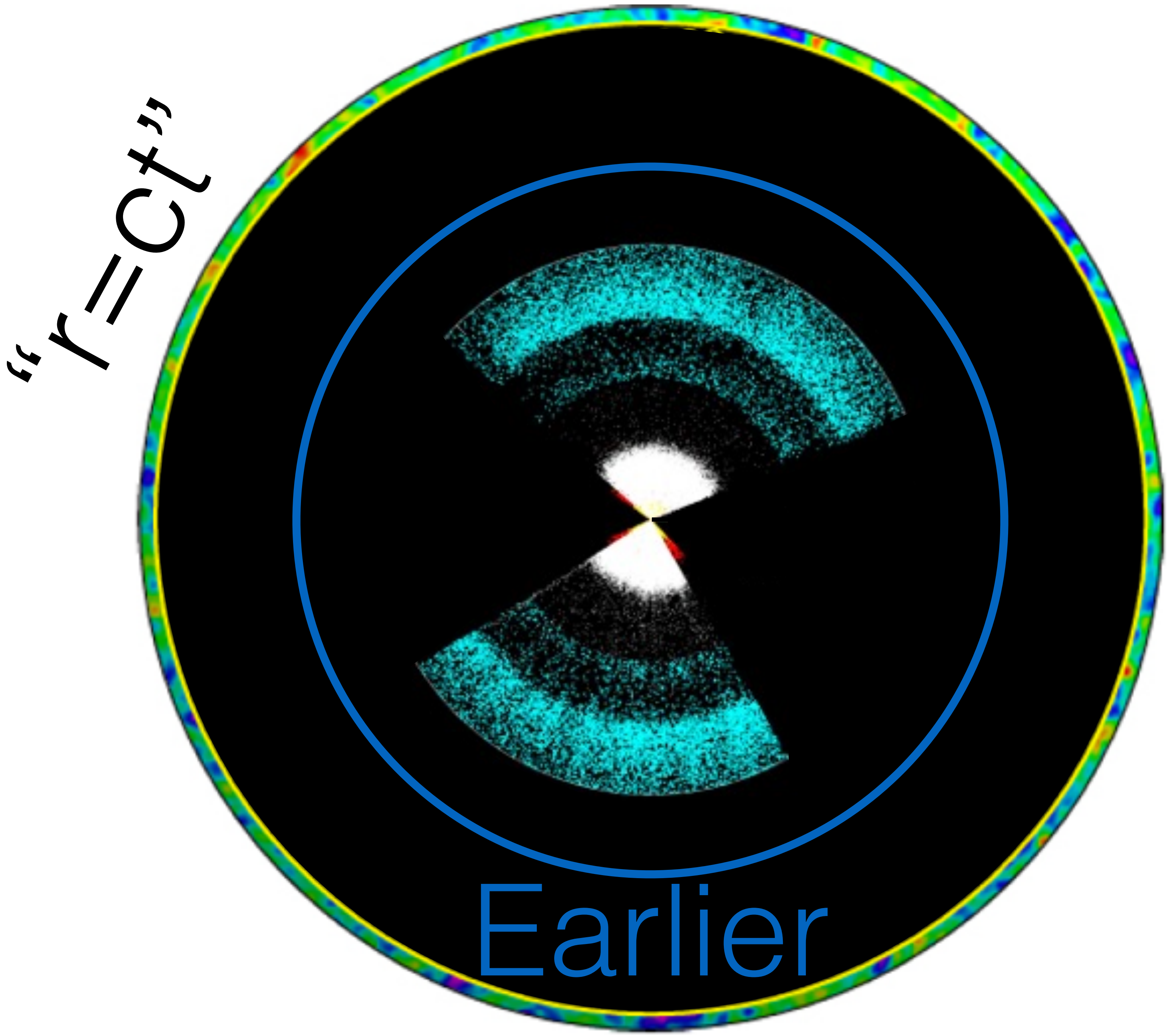
We have yet to observe most of the observable Universe



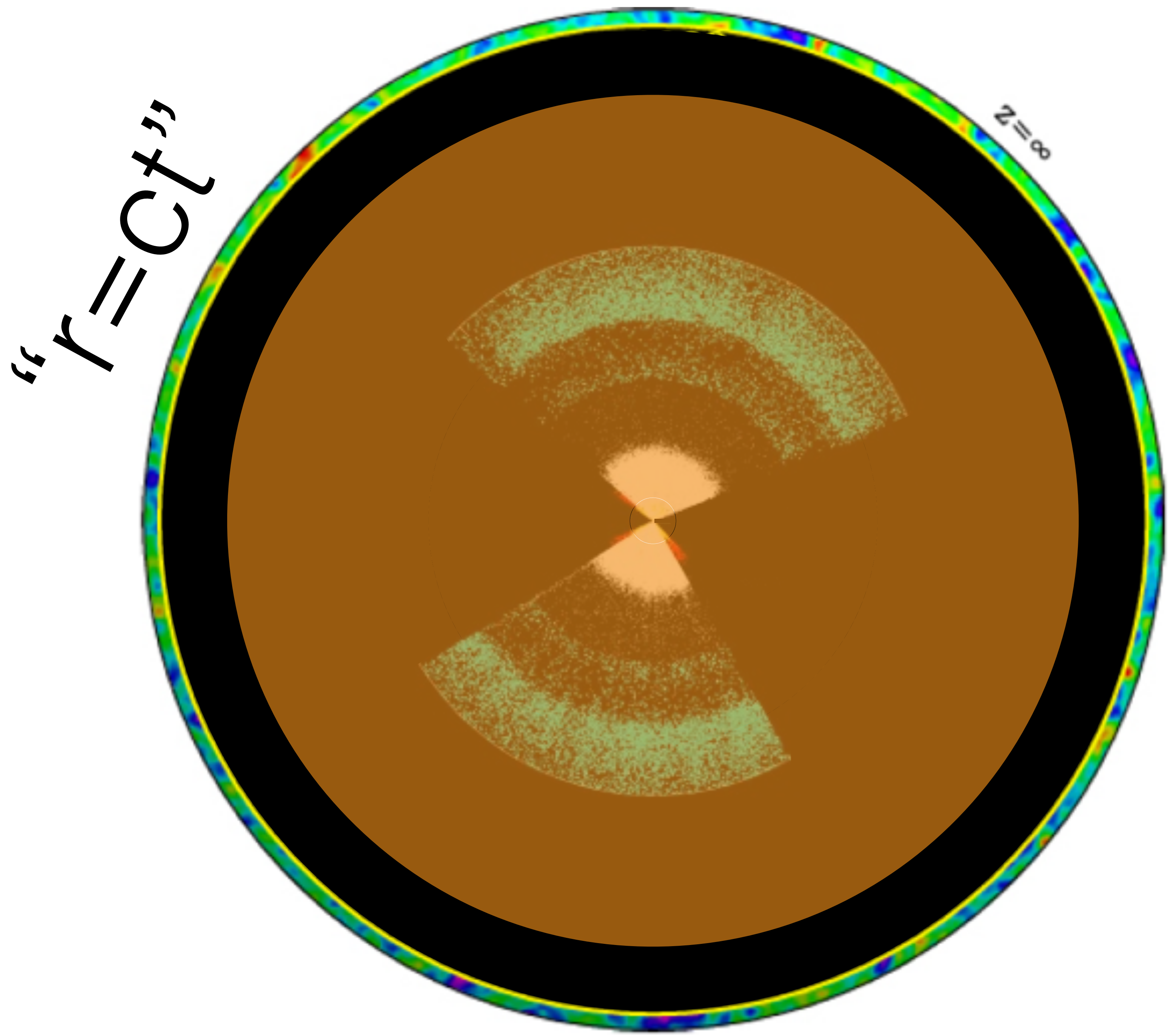
We have yet to observe most of the observable Universe

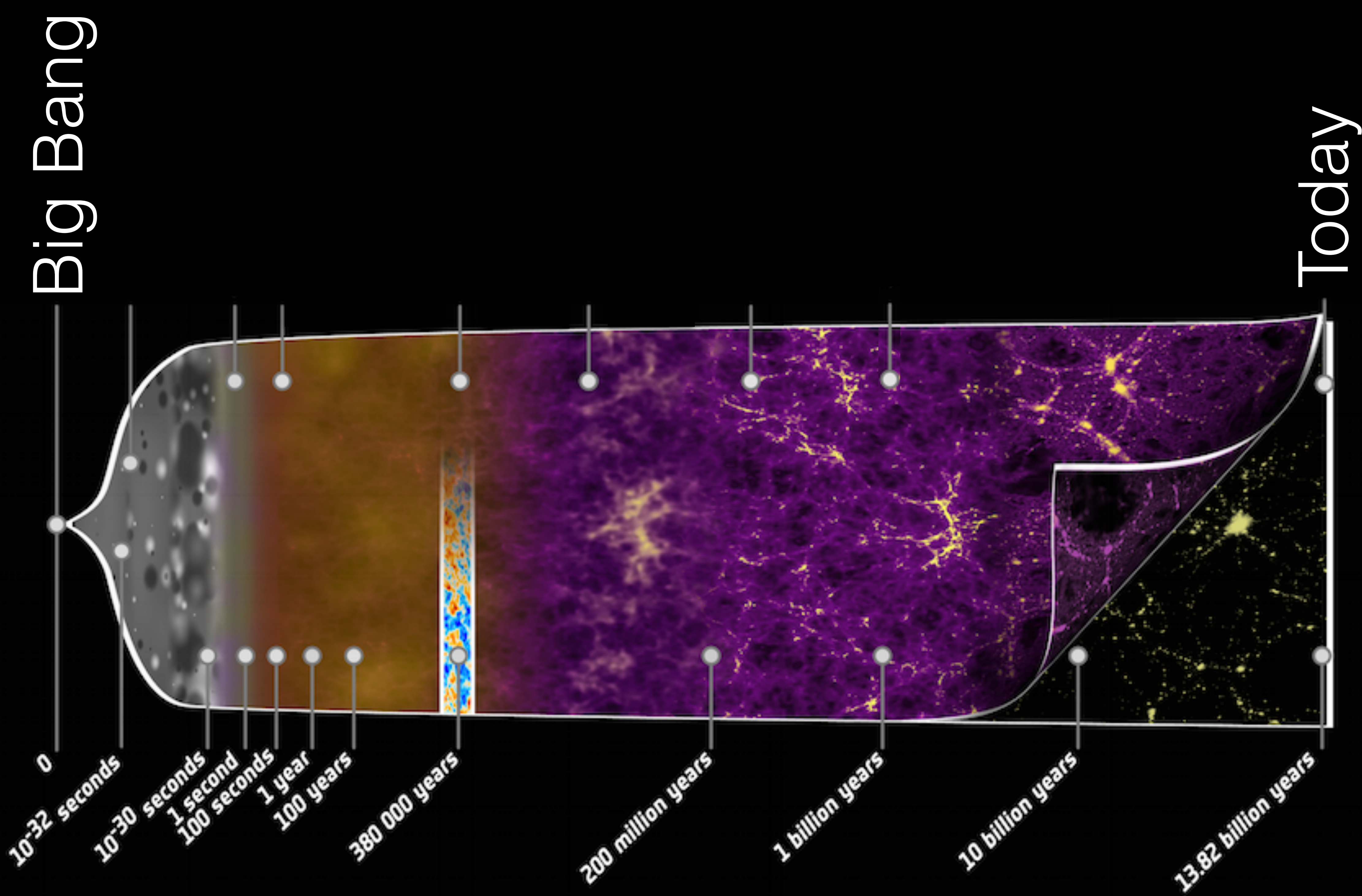


We have yet to observe most of the observable Universe



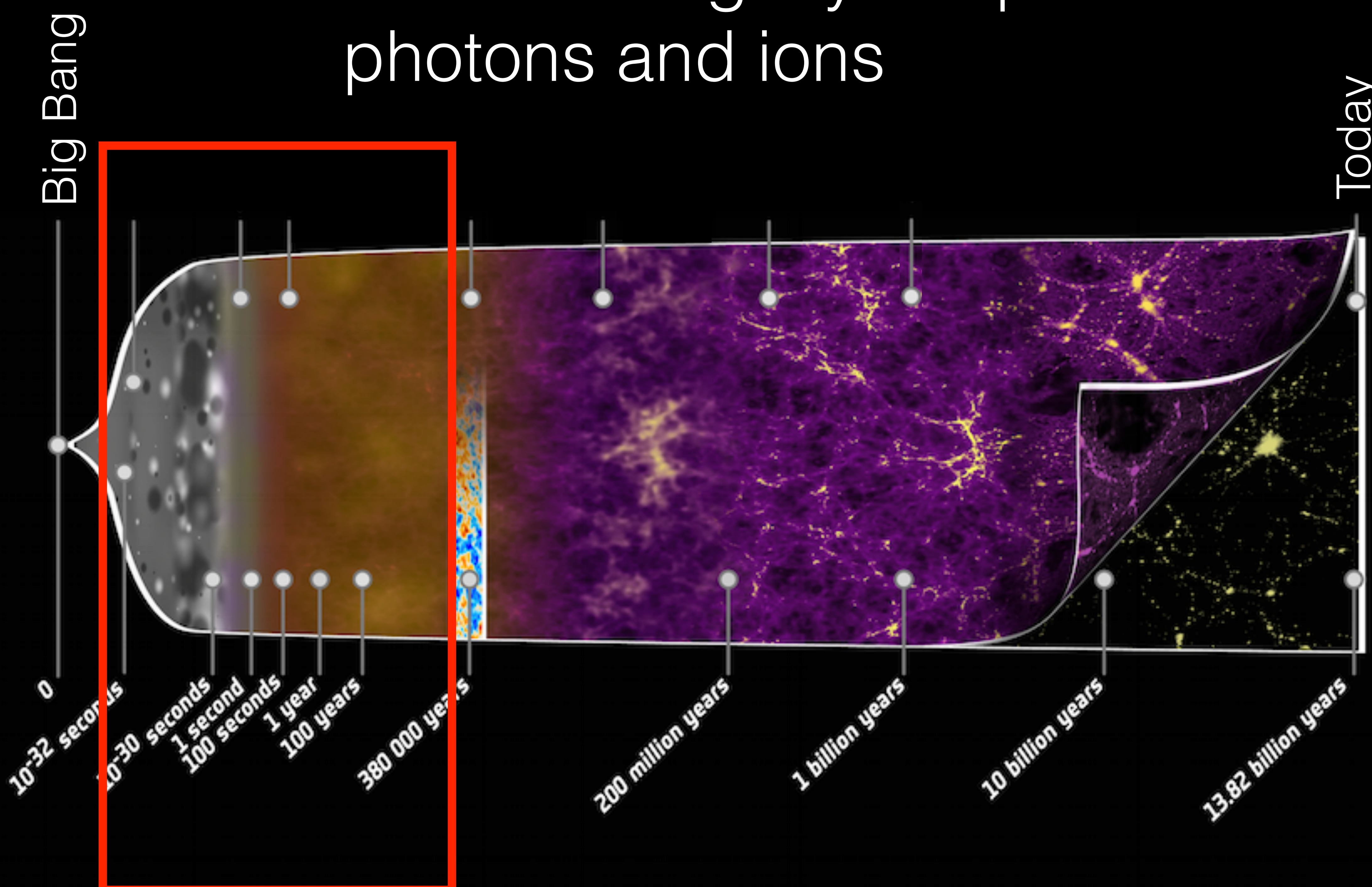
We have yet to observe most of the observable Universe

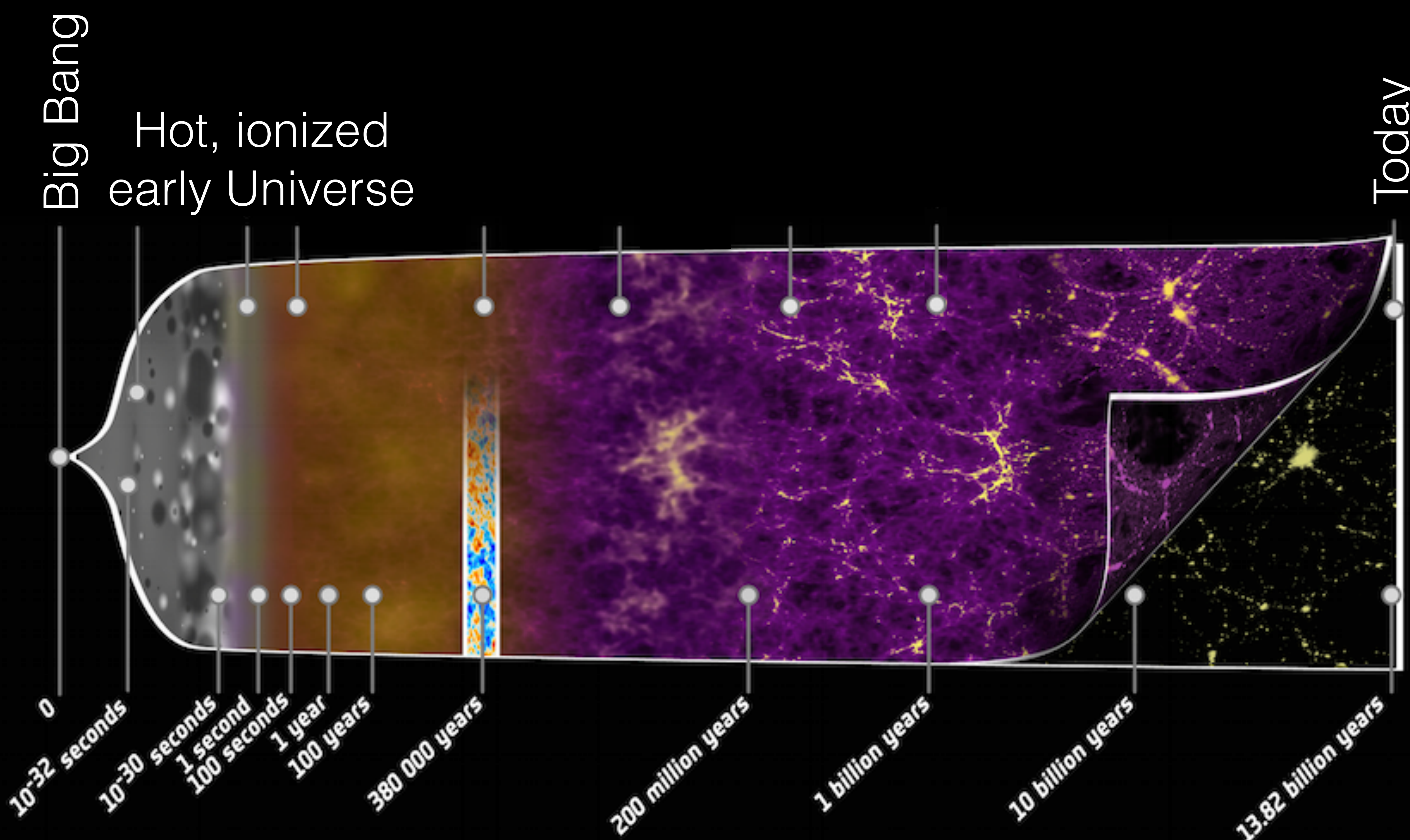



Big Bang
Today
0
10^{-32} seconds
10^{-30} seconds
1 second
100 seconds
1 year
100 years
380 000 years
200 million years
1 billion years
10 billion years
13.82 billion years

Early Universe:

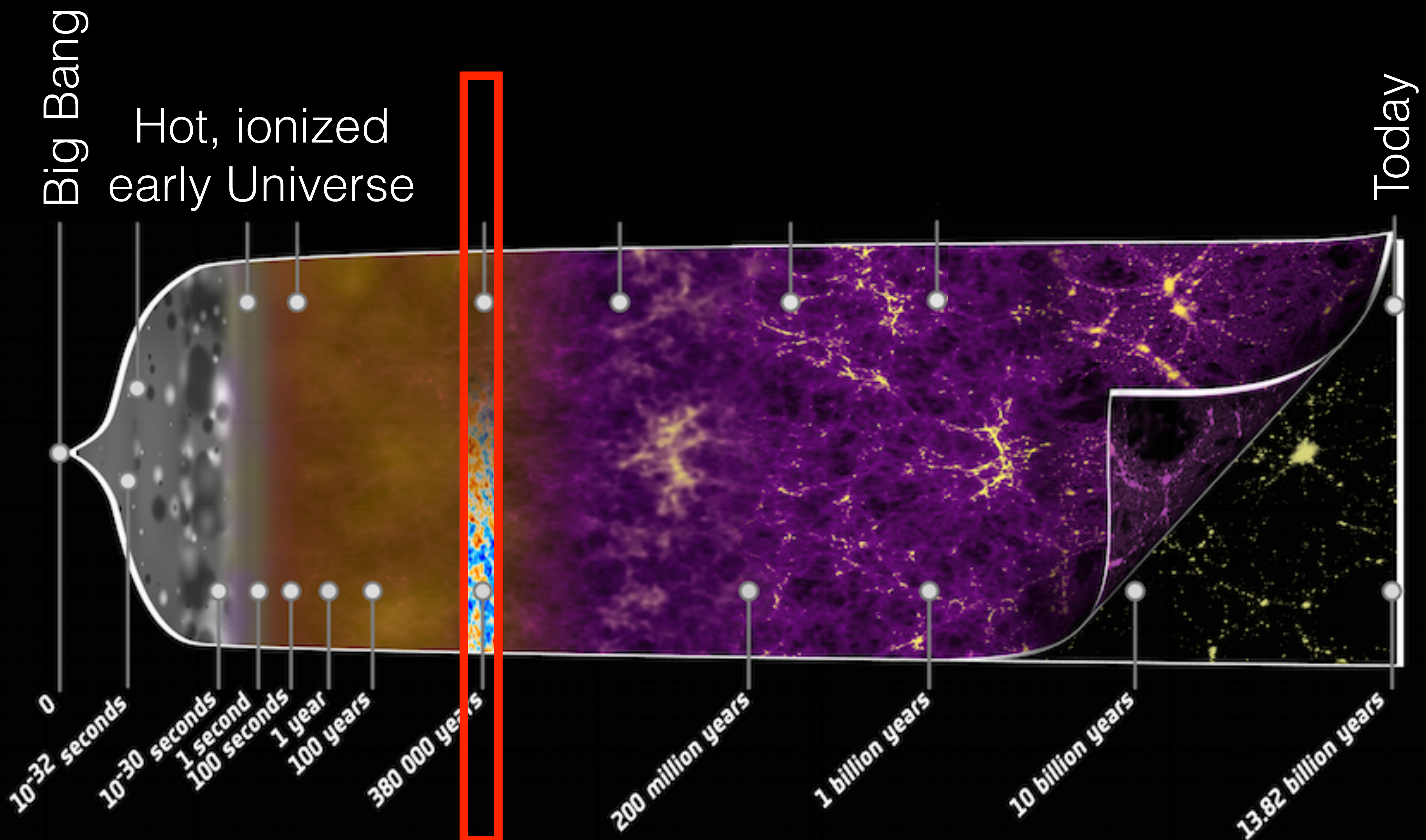
- Hot, dense
- Plasma of tightly coupled photons and ions

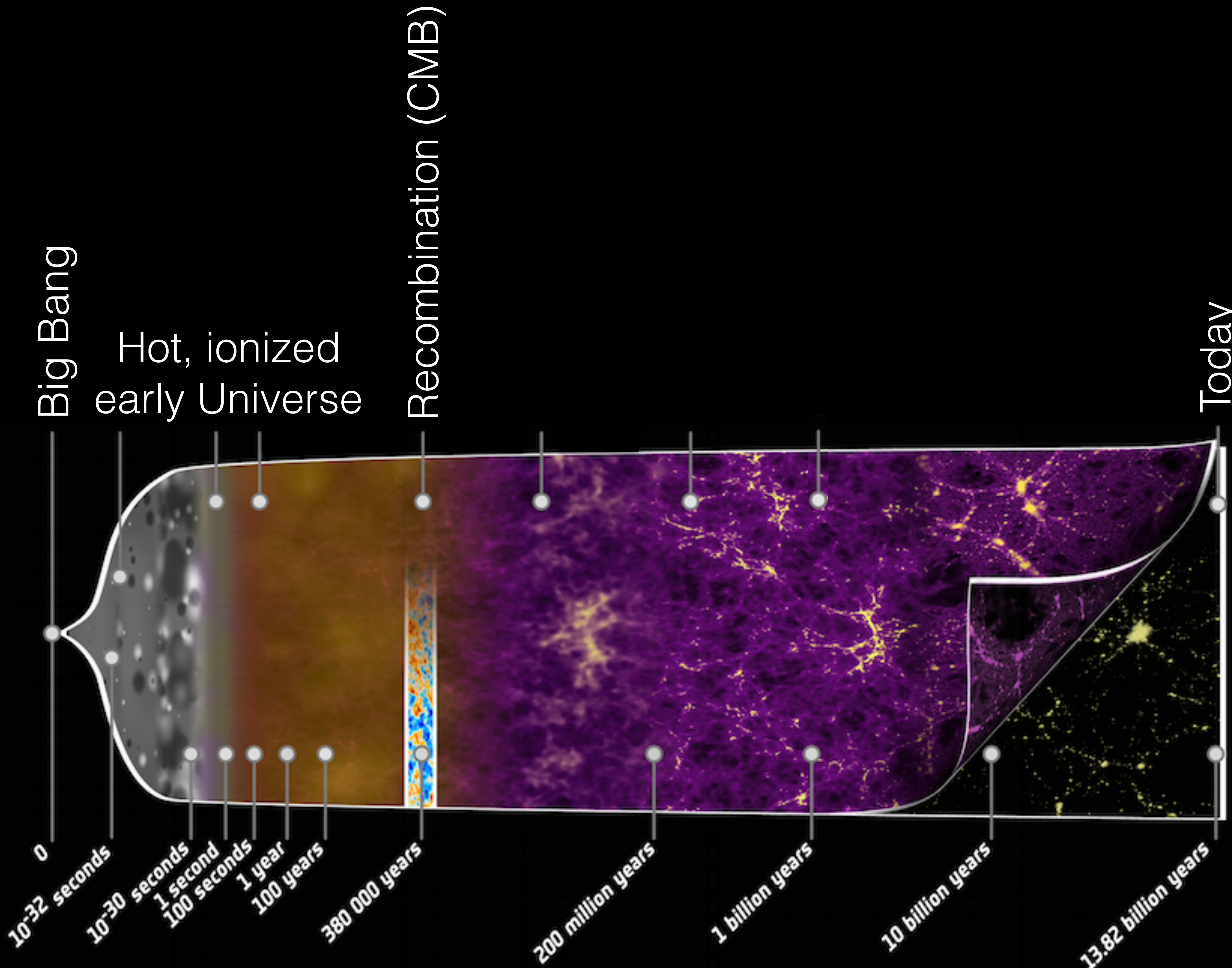



Big Bang
Hot, ionized
early Universe
Today
0
10-32 seconds
10-30 seconds
1 second
100 seconds
1 year
100 years
380 000 years
200 million years
1 billion years
10 billion years
13.82 billion years

Recombination:

- CMB released
- Universe becomes neutral

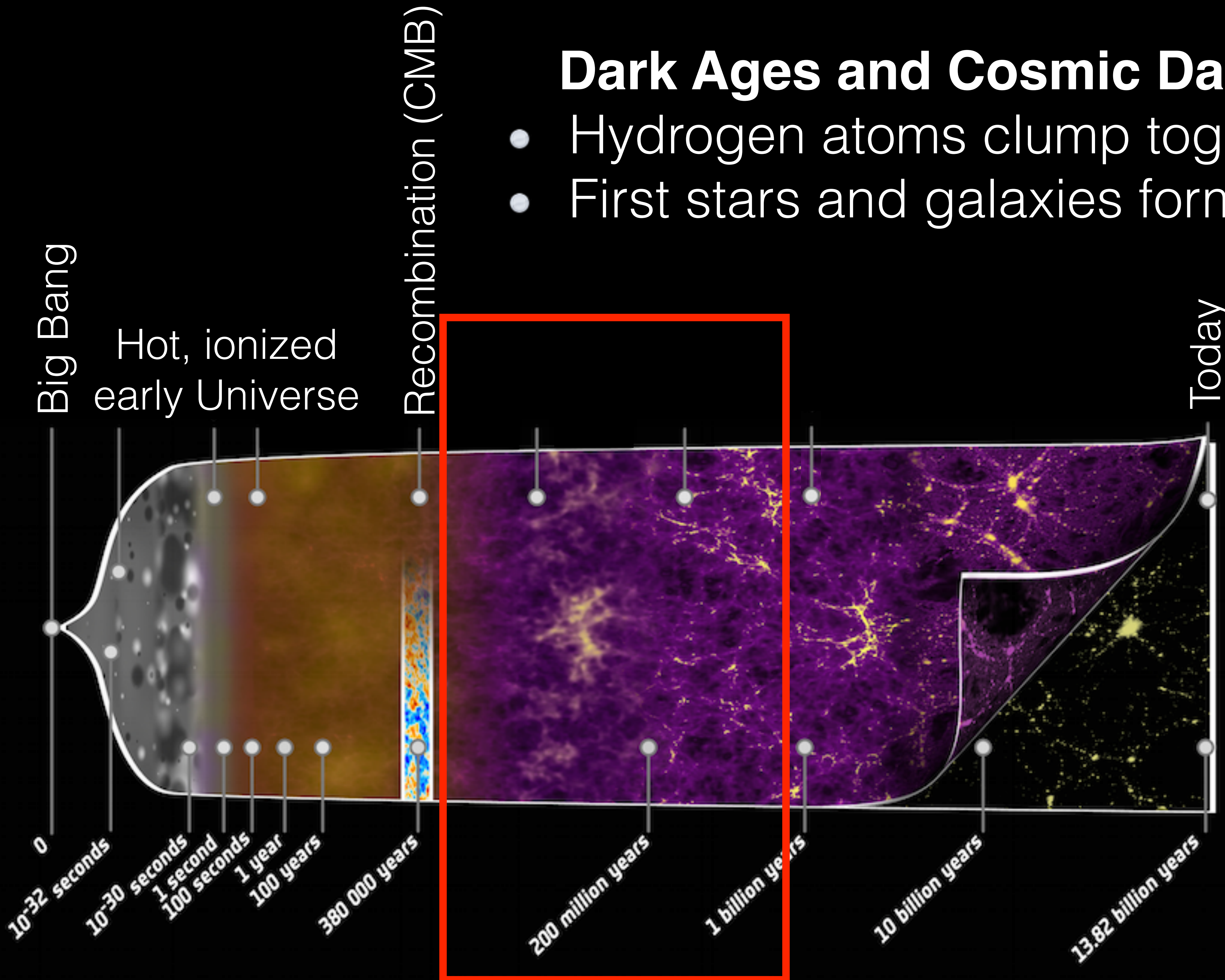


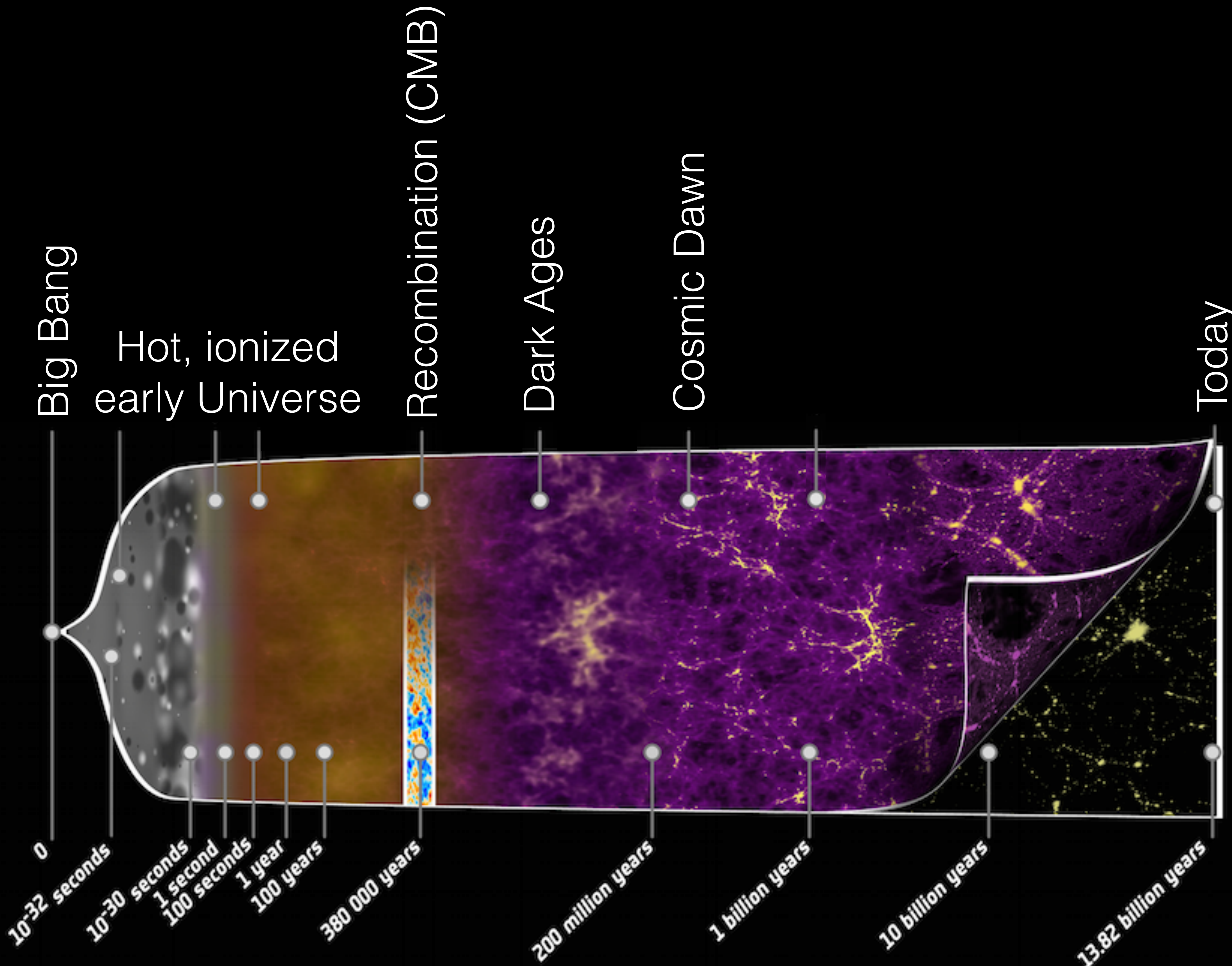


Big Bang
Hot, ionized
early Universe
Recombination (CMB)
Today
0
10-32 seconds
10-30 seconds
1 second
100 seconds
1 year
100 years
380 000 years
200 million years
1 billion years
10 billion years
13.82 billion years

Dark Ages and Cosmic Dawn:

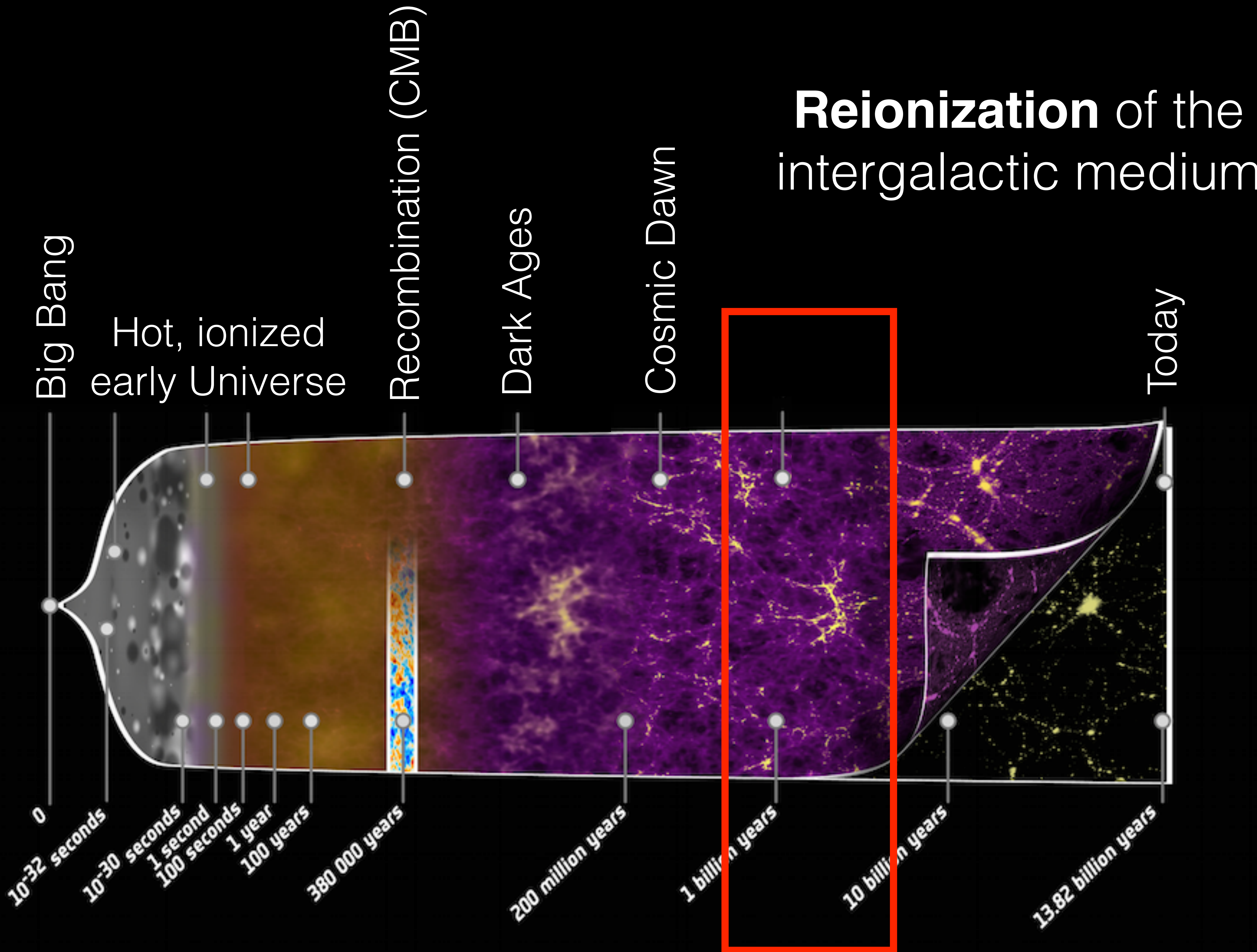
- Hydrogen atoms clump together
- First stars and galaxies form



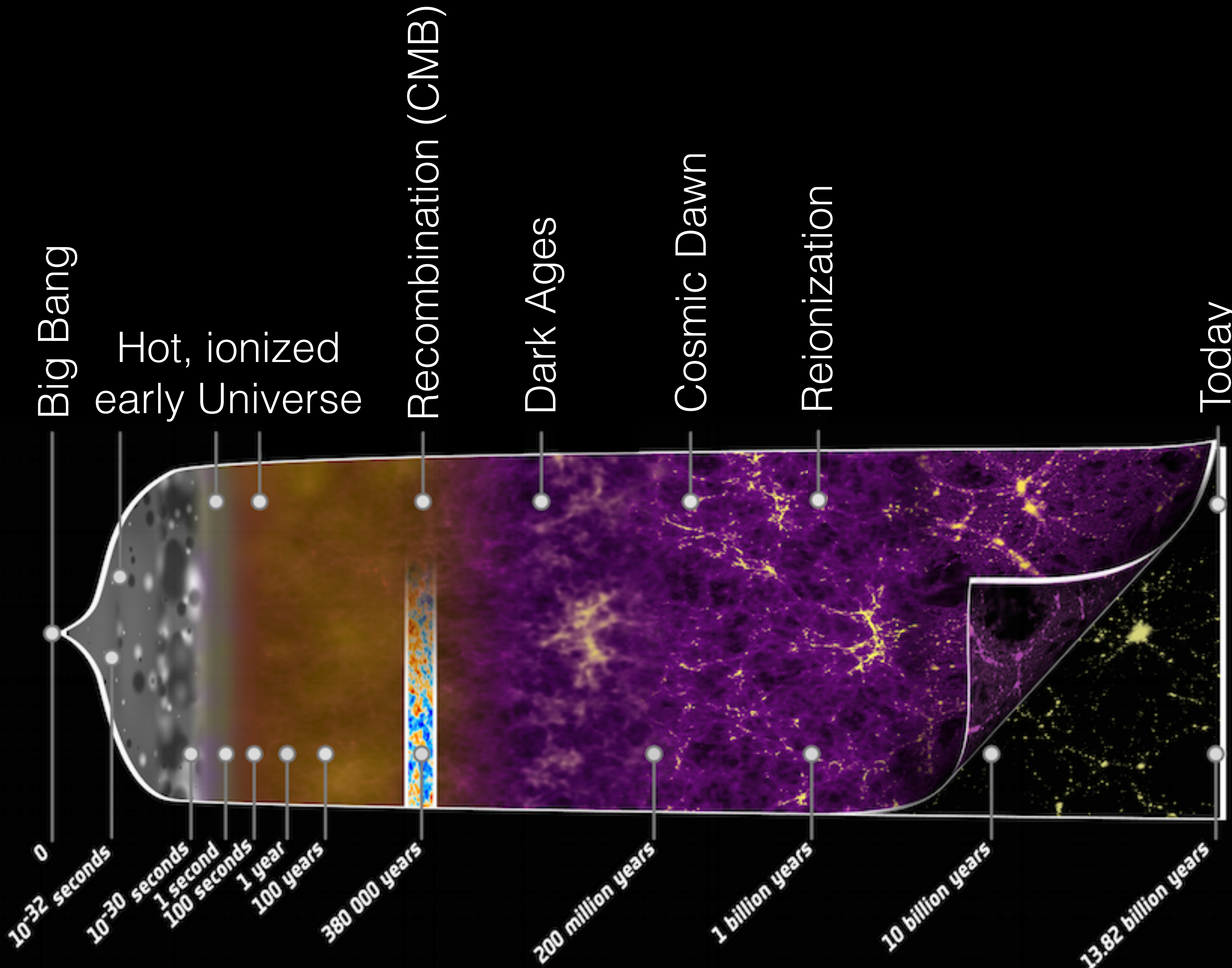


Big Bang
Hot, ionized early Universe
Recombination (CMB)
Dark Ages
Cosmic Dawn
Today
0
10^{-32} seconds
10^{-30} seconds
1 second
100 seconds
1 year
100 years
380 000 years
200 million years
1 billion years
10 billion years
13.82 billion years

Alvarez et al. (2009)

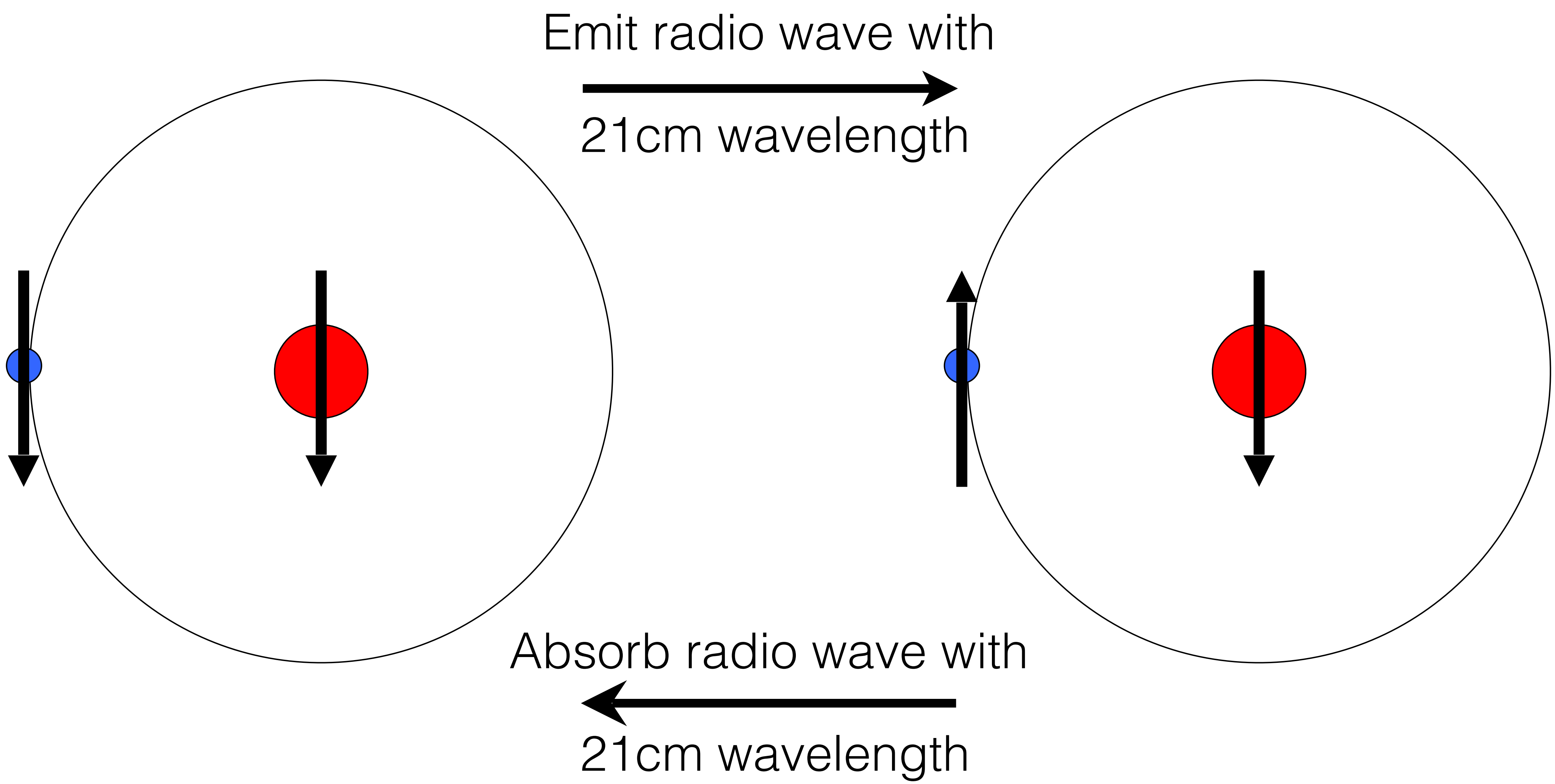


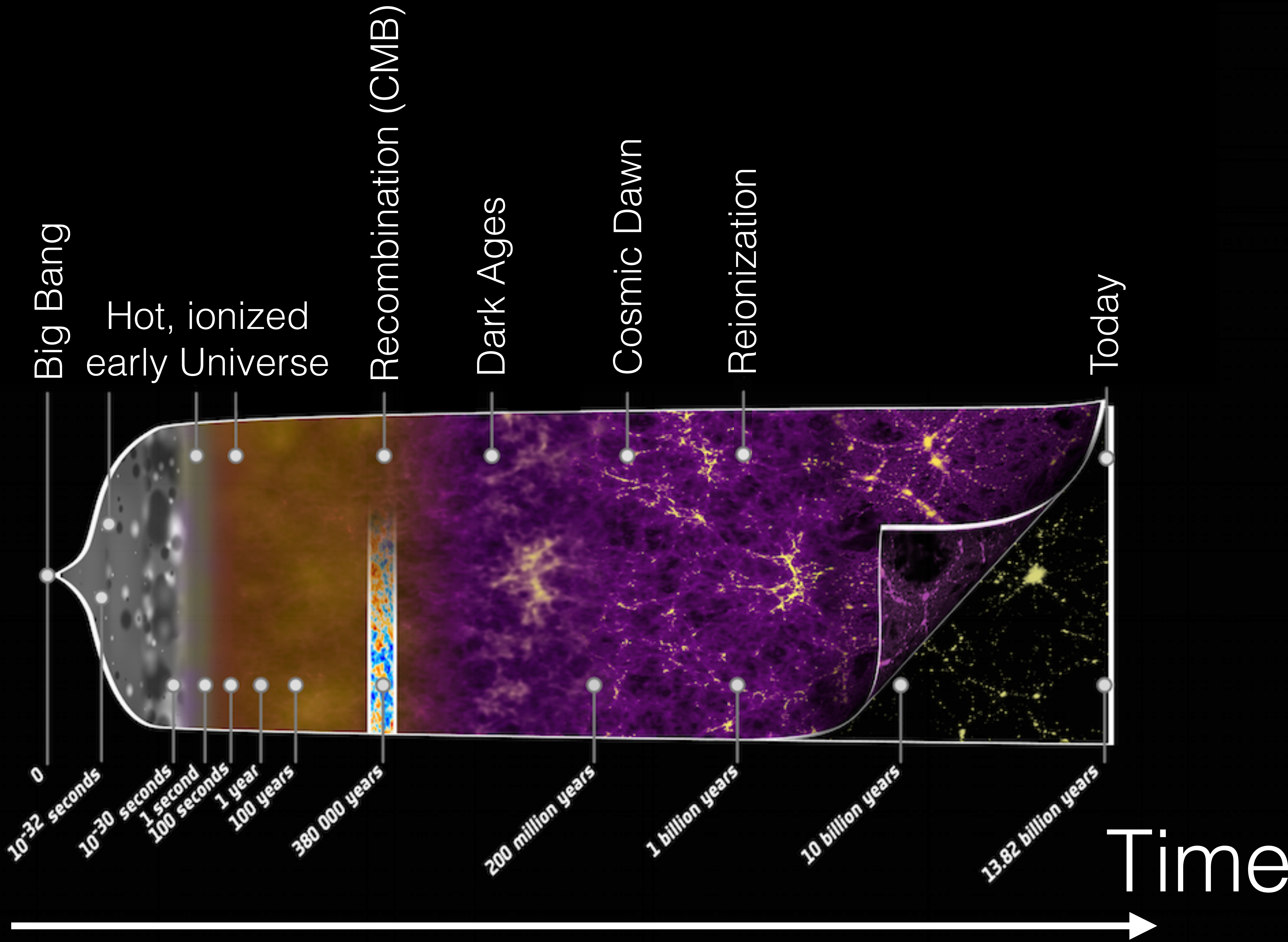
Reionization of the intergalactic medium
Big Bang
Hot, ionized early Universe
Recombination (CMB)
Dark Ages
Cosmic Dawn
Today
0
10-32 seconds
10-30 seconds
1 second
100 seconds
1 year
100 years
380 000 years
200 million years
1 billion years
10 billion years
13.82 billion years

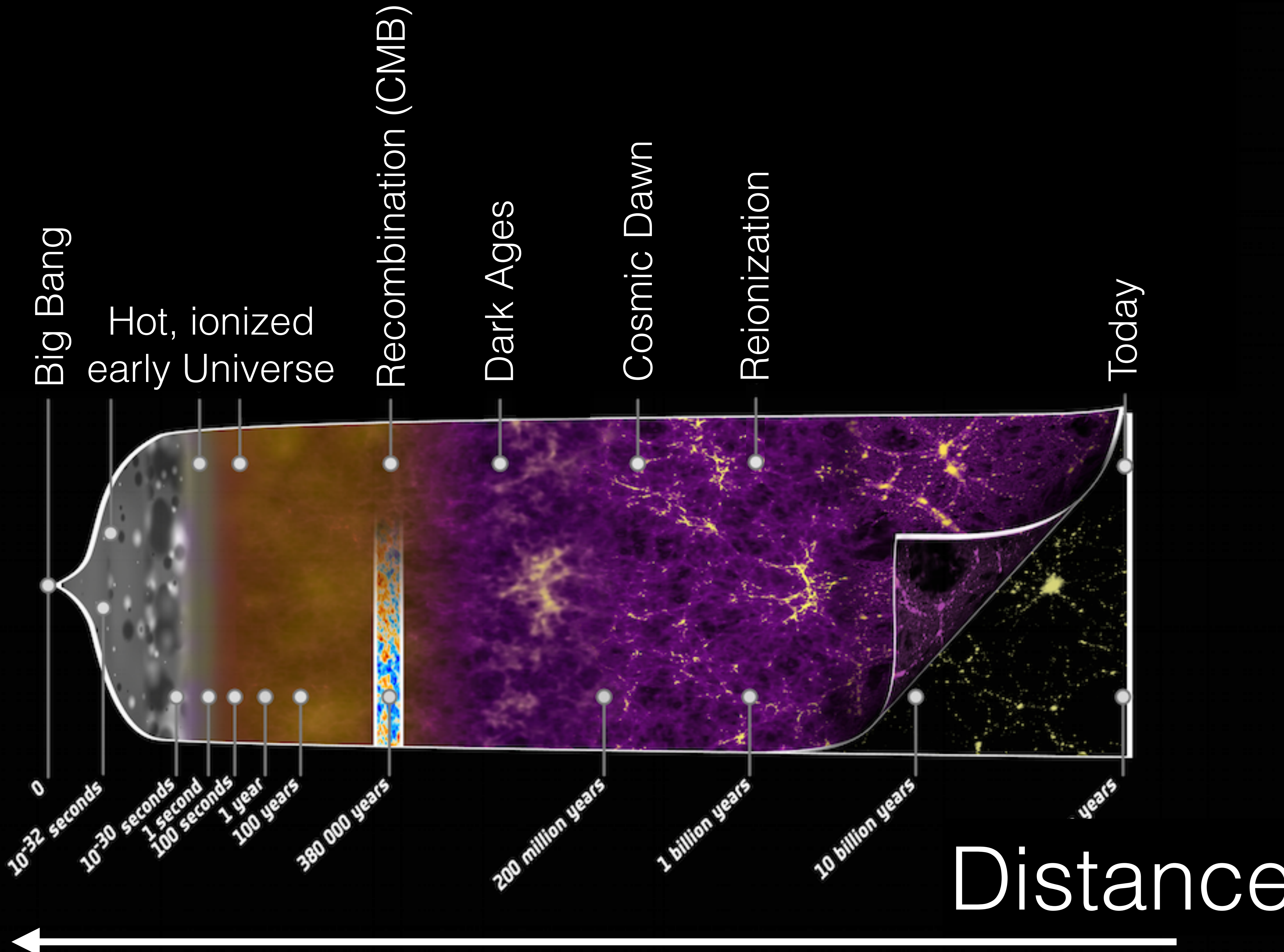


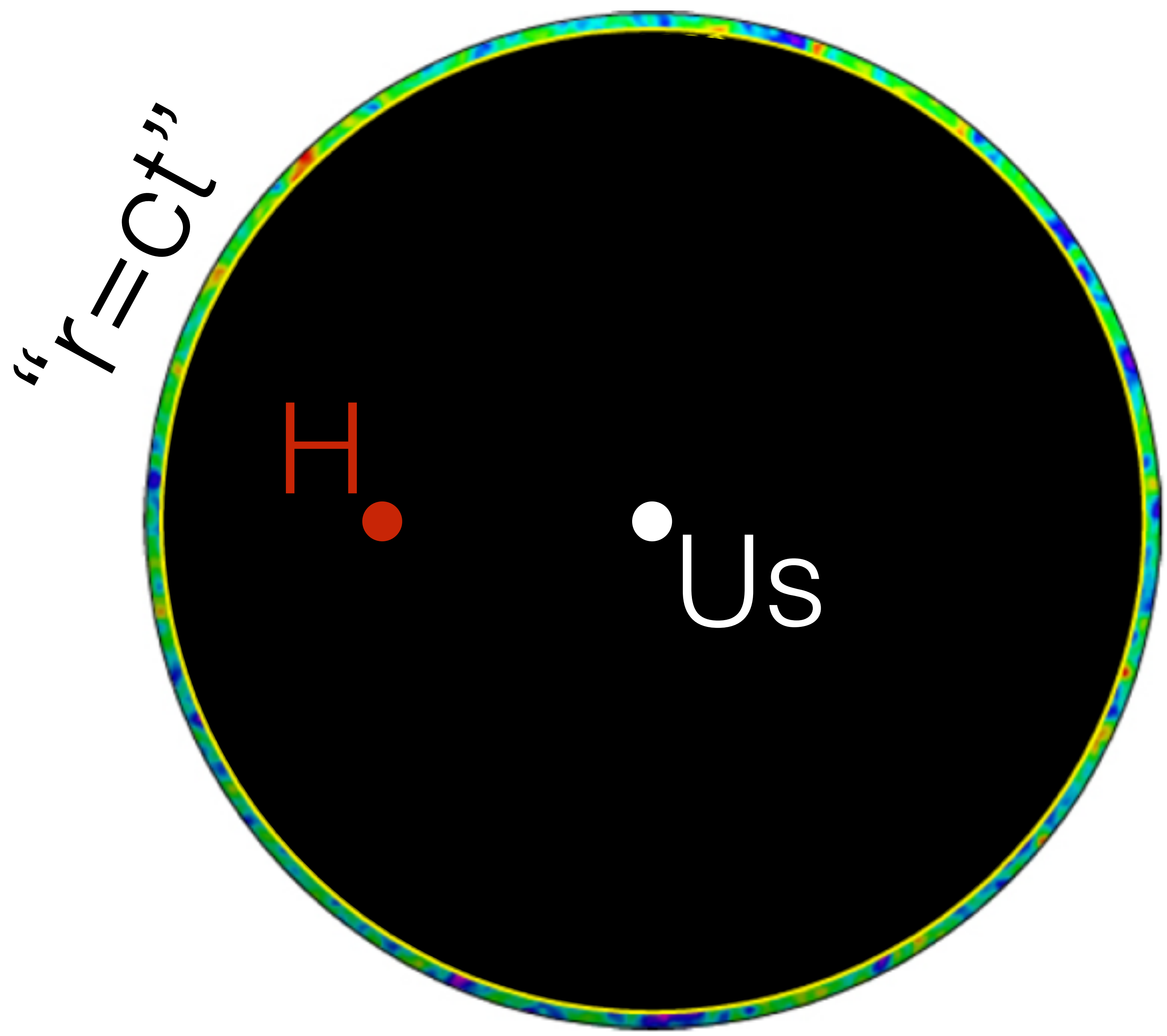
Big Bang
Hot, ionized early Universe
Recombination (CMB)
Dark Ages
Cosmic Dawn
Reionization
Today
0
10-32 seconds
10-30 seconds
1 second
100 seconds
1 year
100 years
380 000 years
200 million years
1 billion years
10 billion years
13.82 billion years

Hydrogen is everywhere, and the 21cm line allows us to trace hydrogen

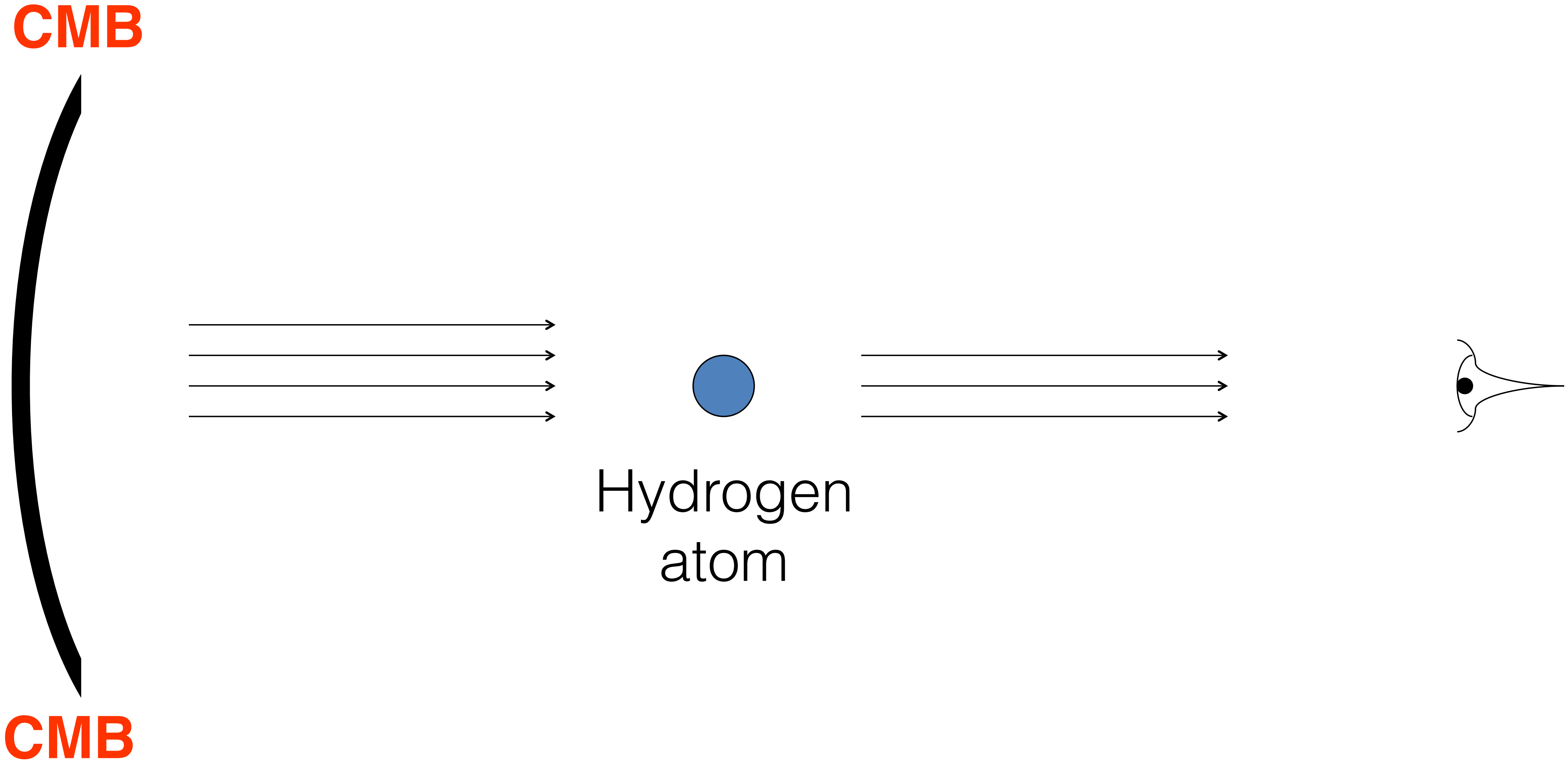


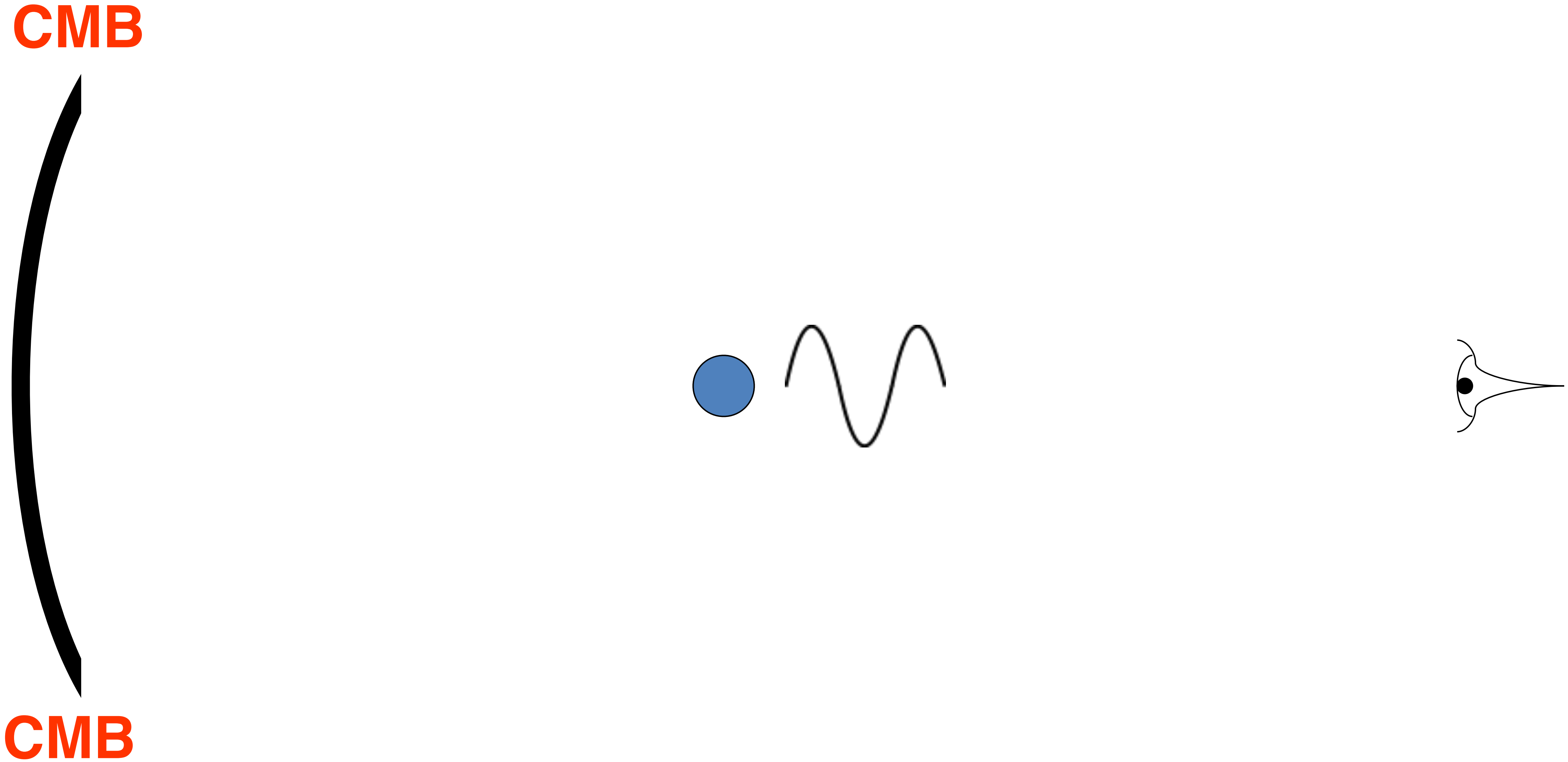

Big Bang
Hot, ionized
early Universe
Recombination (CMB)
Dark Ages
Cosmic Dawn
Reionization
Today
0
10-32 seconds
10-30 seconds
1 second
100 seconds
1 year
100 years
380 000 years
200 million years
1 billion years
10 billion years
13.82 billion years
Time

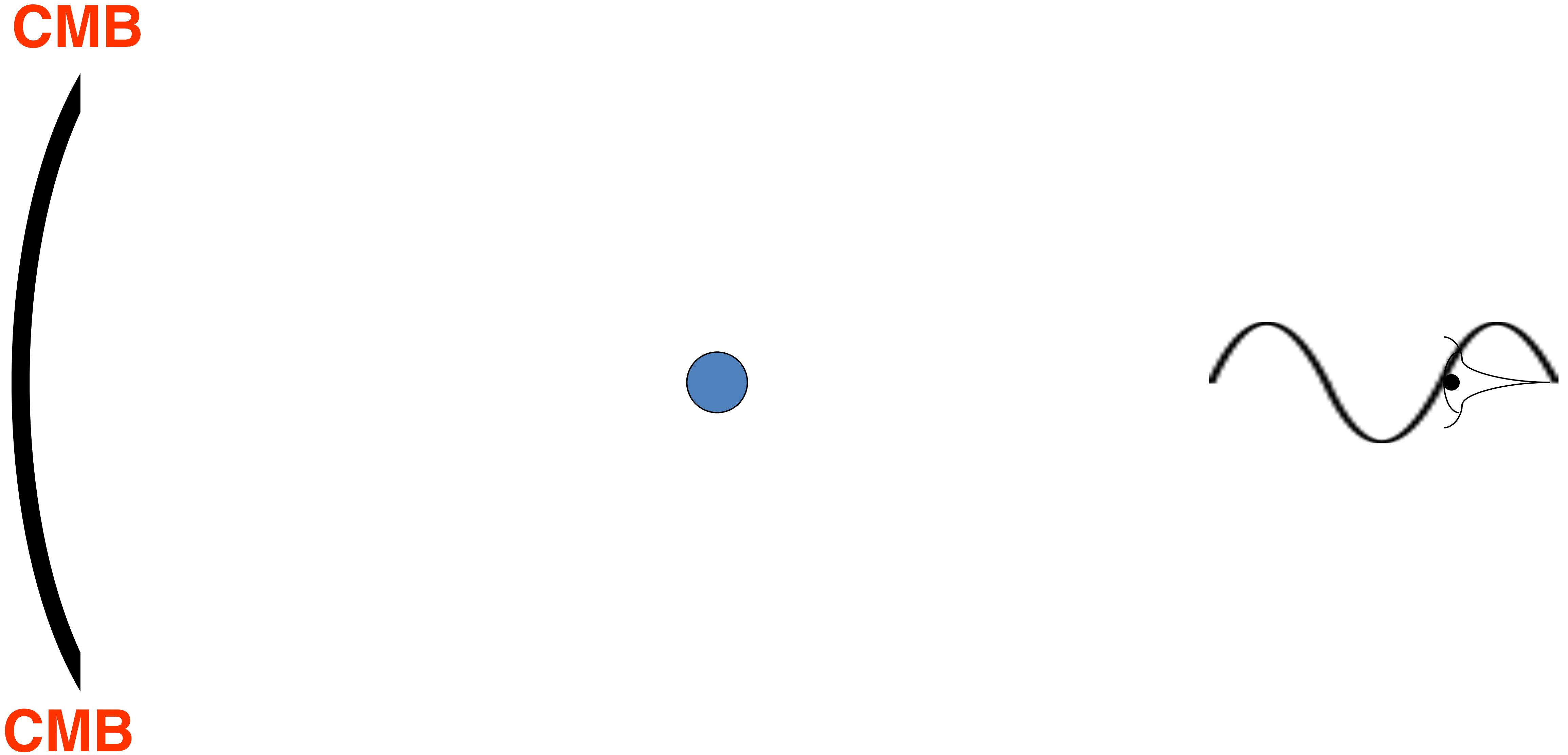

Big Bang
Hot, ionized early Universe
Recombination (CMB)
Dark Ages
Cosmic Dawn
Reionization
Today
0
10-32 seconds
10-30 seconds
1 second
100 seconds
1 year
100 years
380 000 years
200 million years
1 billion years
10 billion years
years
Distance

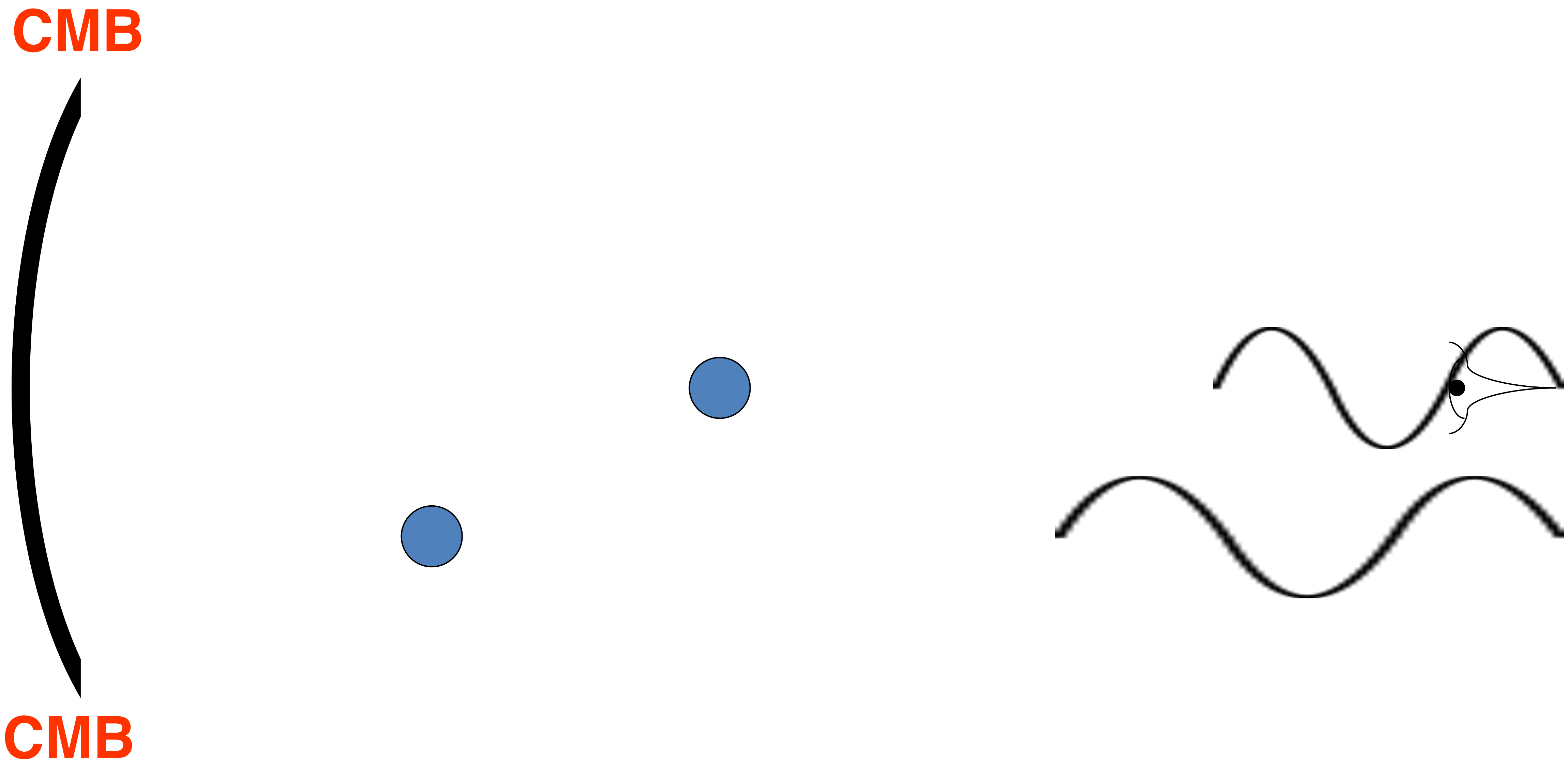


“r=ct”
H
Us


CMB
Hydrogen
atom
CMB

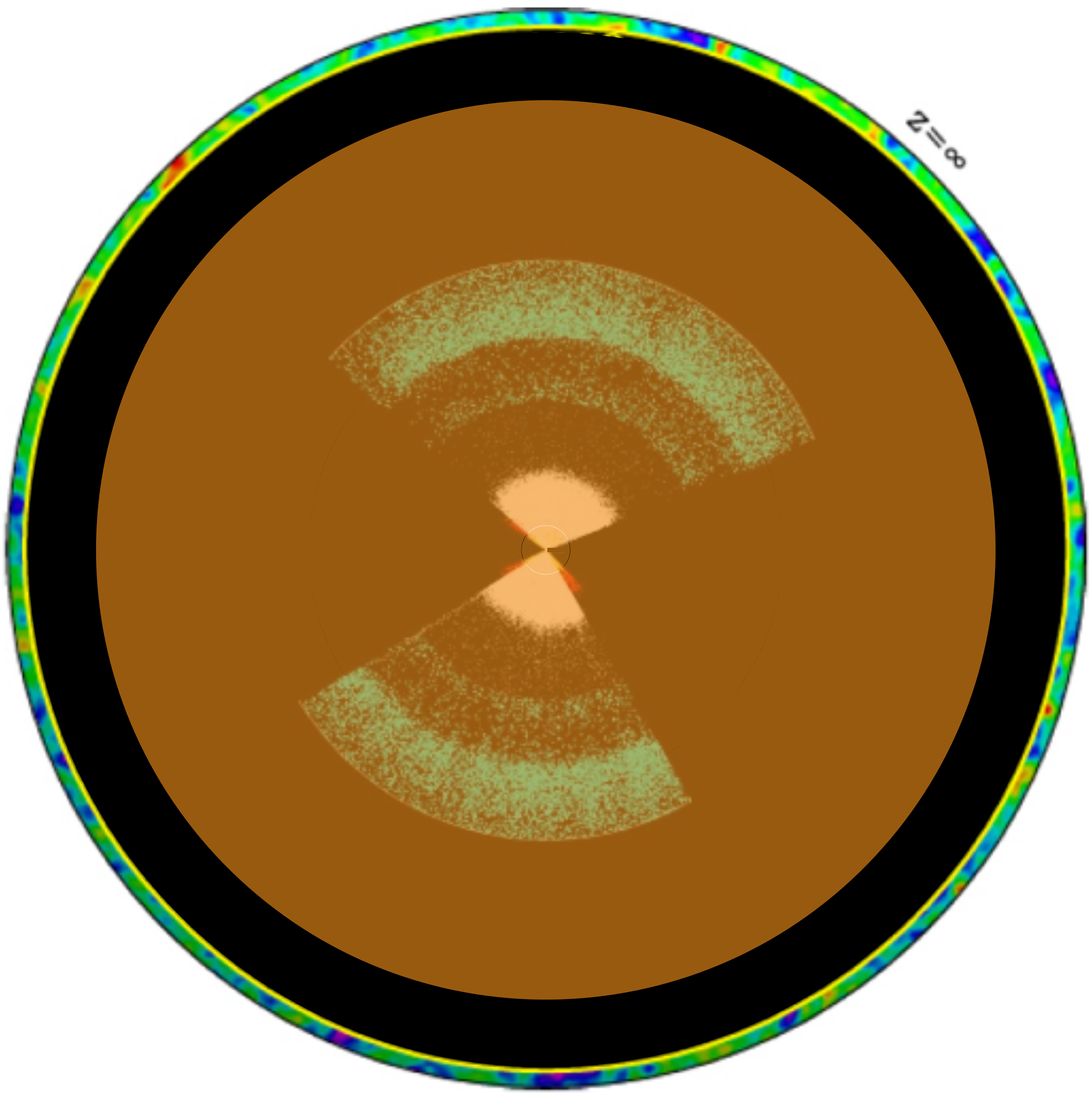

CMB
CMB


CMB
CMB

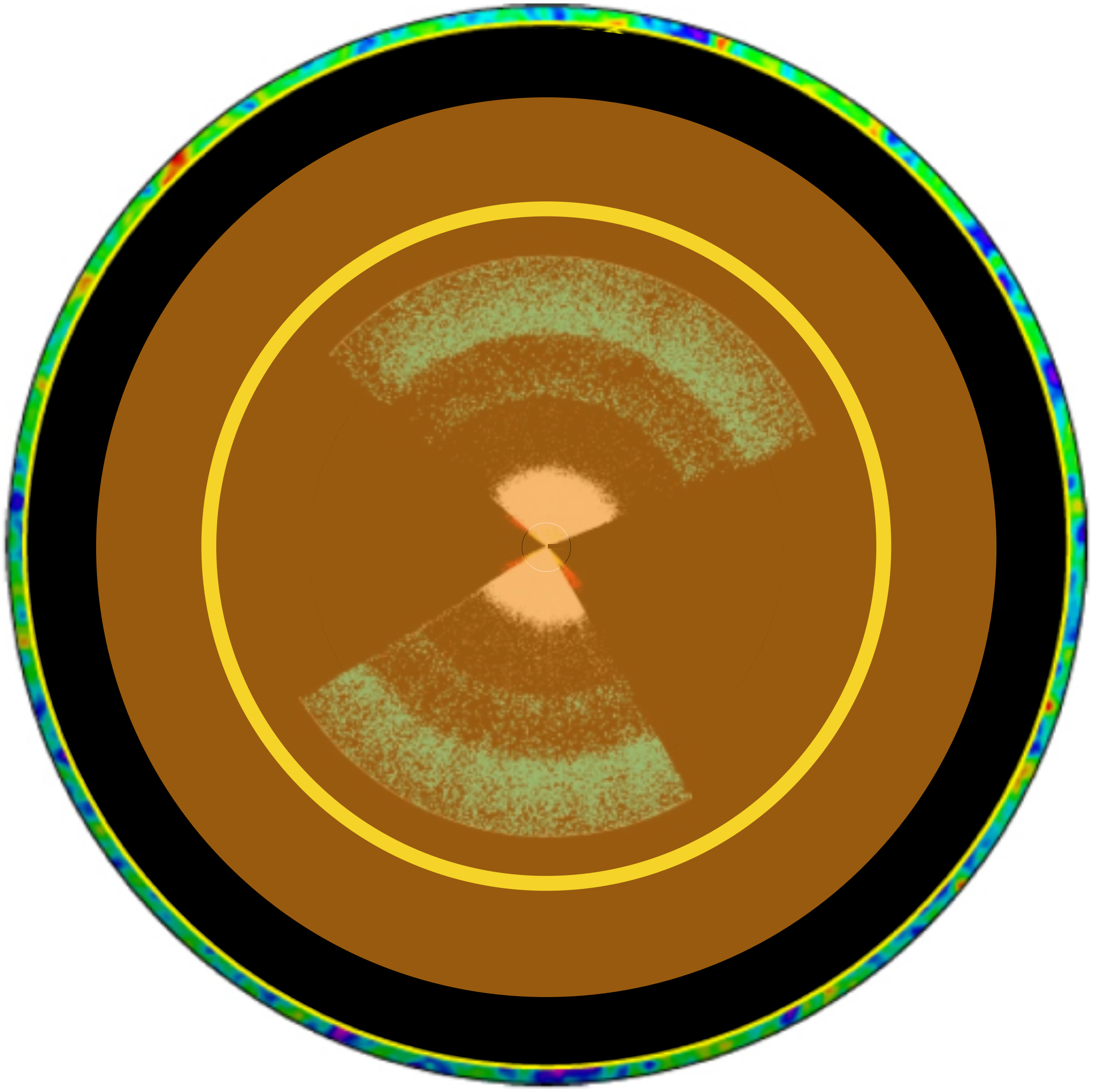


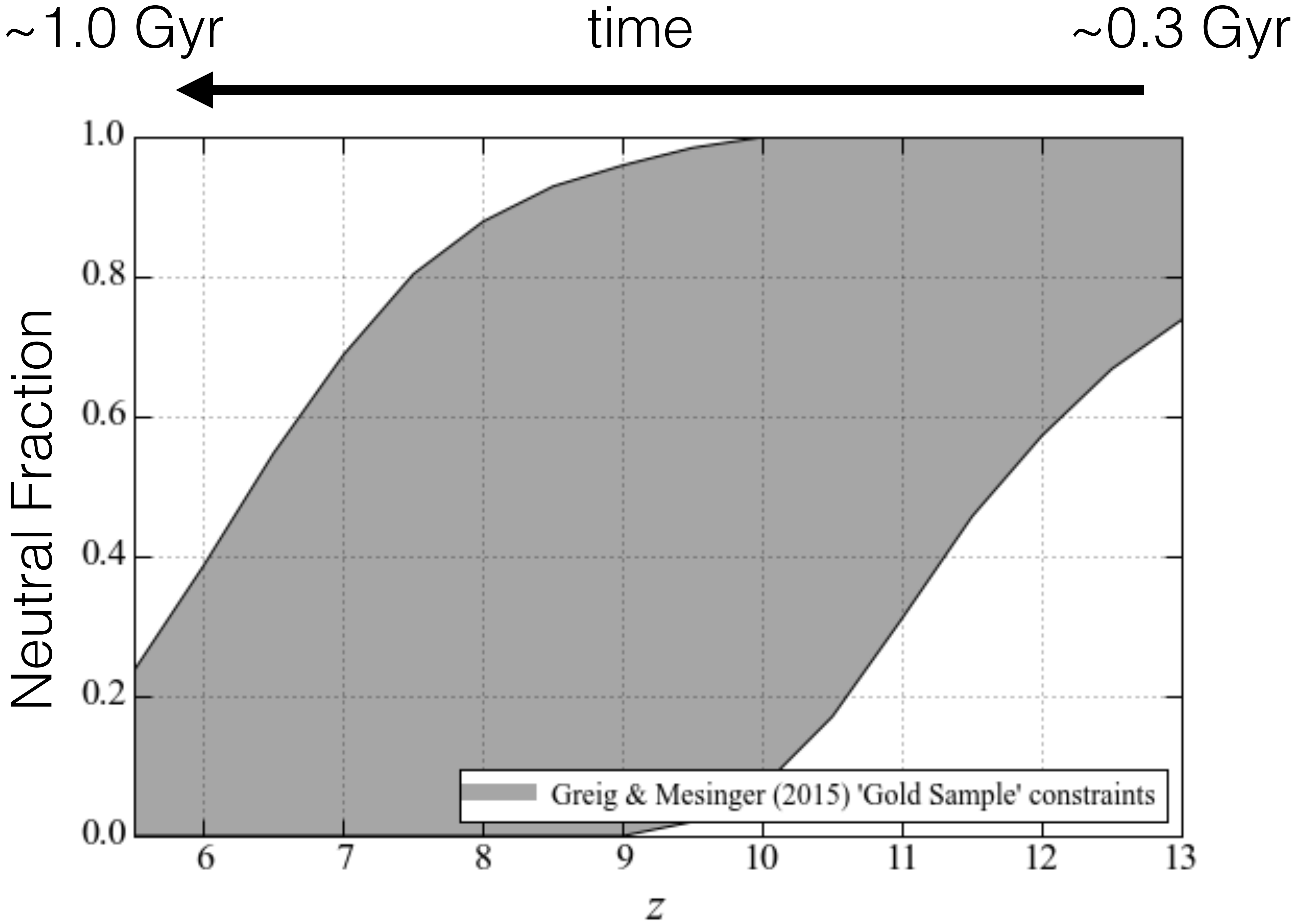
Depth perception comes from measuring the observed wavelength

21cm cosmology will allow gaps in the cosmic timeline to be filled by directly observing radio absorption or emission from hydrogen atoms



Current generation experiments are targeting the Epoch of Reionization (EoR)




~1.0 Gyr
time
~0.3 Gyr
Neutral Fraction
1.0
0.8
0.6
0.4
0.2
0.0
6
7
8
9
10
11
12
13
z
Greig & Mesinger (2015) 'Gold Sample' constraints

Take-home messages

- We’re getting close to detecting the 21cm signal—close enough to start improving our understanding of reionization.

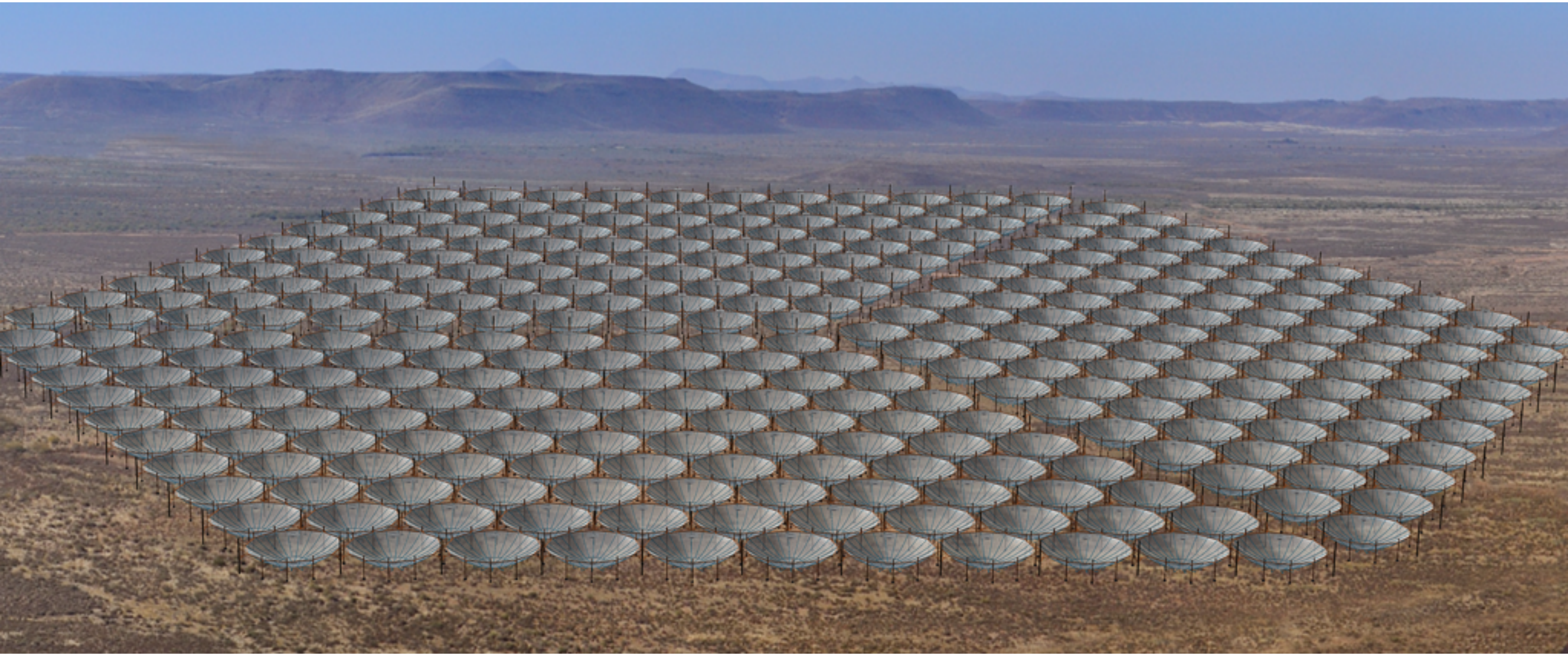
- 21cm cosmology is a data-intensive science where astrophysics and cosmology go hand-in-hand

- The HERA experiment is being built now, and promises to deliver qualitatively new constraints on astrophysics and cosmology.

- 21cm cosmology provides a window into fundamental physics with opportunities to push the time, sensitivity, and scale frontiers.

The promise of 21cm measurements

Hydrogen Epoch of Reionization Array (HERA)



154 m

THE UNIVERSITY OF CALIFORNIA · LET THERE BE LIGHT · UCLA

UNIVERSITY OF WASHINGTON · LVX SIT · 1861

THE UNIVERSITY OF CALIFORNIA · BERKELEY · LET THERE BE LIGHT · 1868

MASSACHUSETTS INSTITUTE OF TECHNOLOGY · MENS ET MANUS · 1861

ARIZONA STATE UNIVERSITY · GREAT SEAL OF THE STATE OF ARIZONA · DITAT DEUS · 1912 · 1885

Penn
UNIVERSITY of PENNSYLVANIA

UNIVERSITY OF CAMBRIDGE

SCUOLA NORMALE SUPERIORE

H[1] ERA

UNIVERSITY of the WESTERN CAPE
RESPICE PROSPICE

UNIVERSITY OF KWAZULU-NATAL

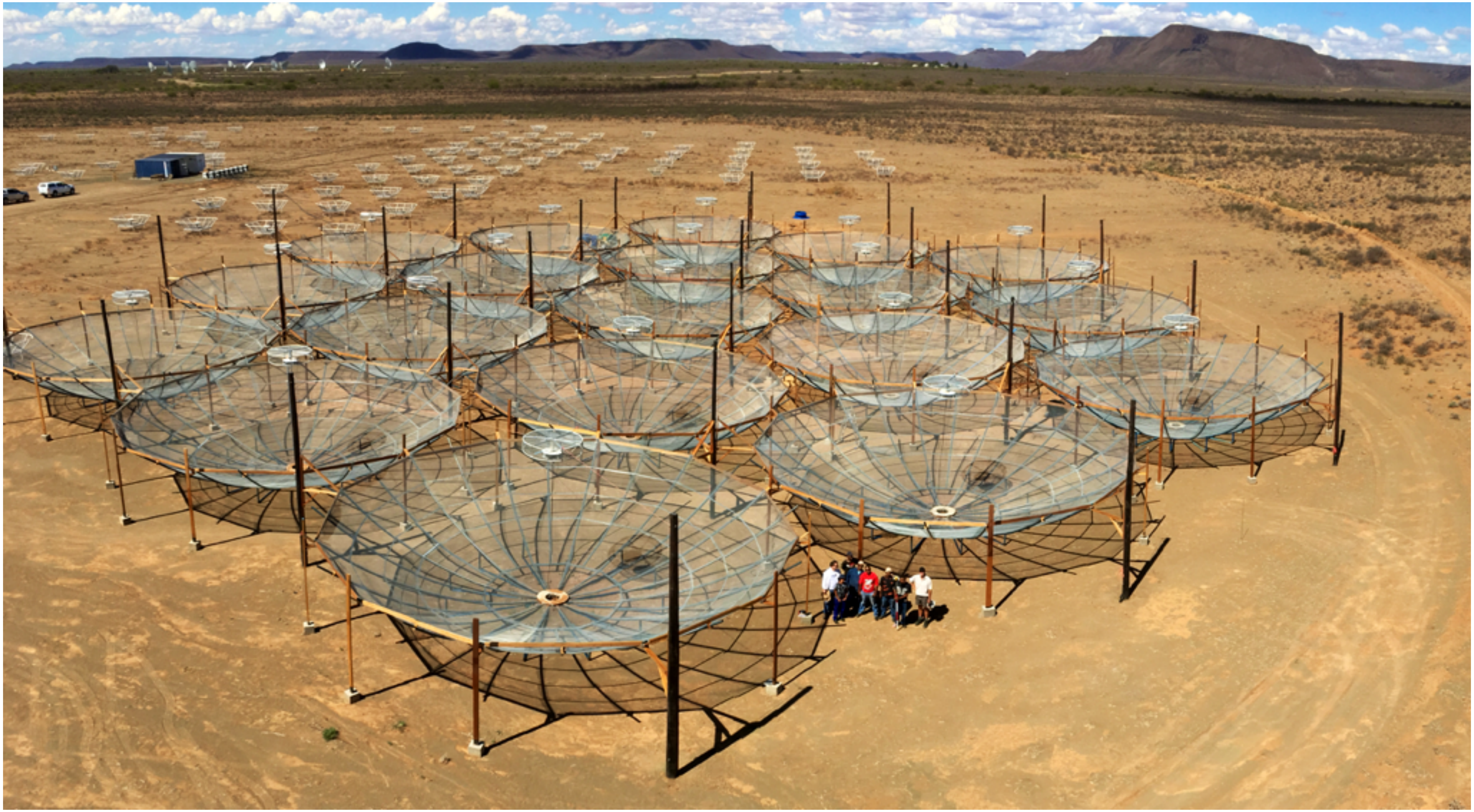
CAL POLY POMONA

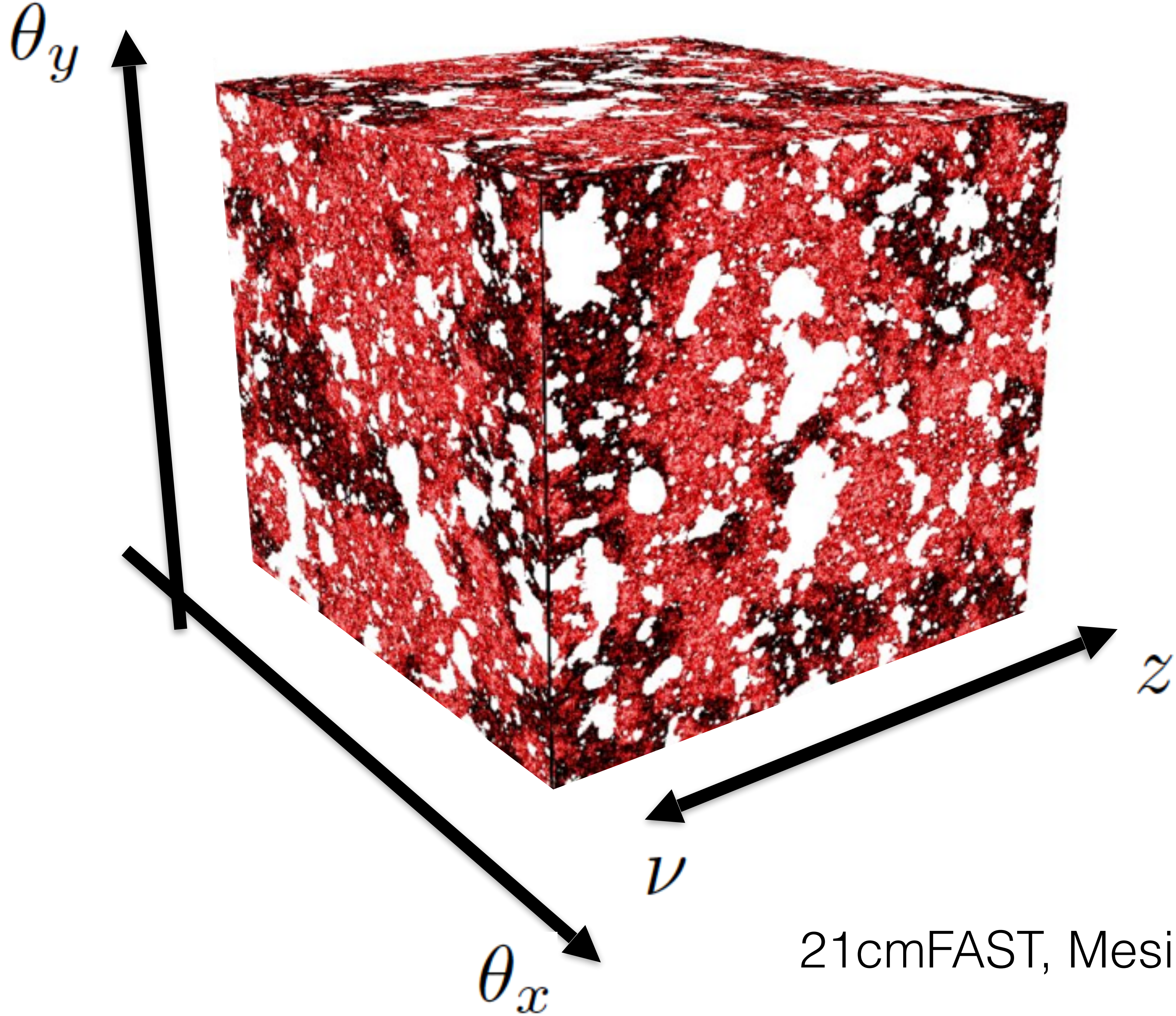
VE RI TAS
HARVARD

BROWN

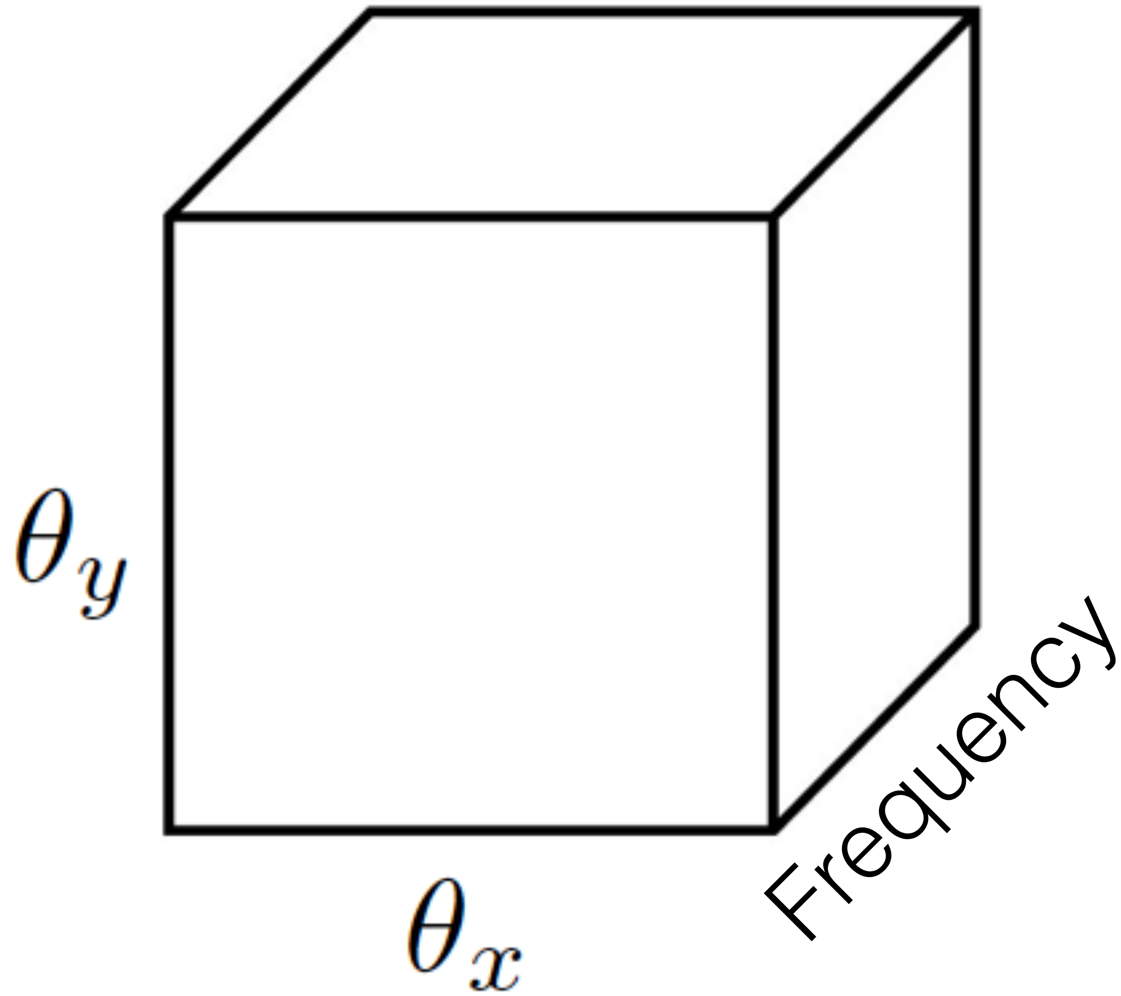
NRAO

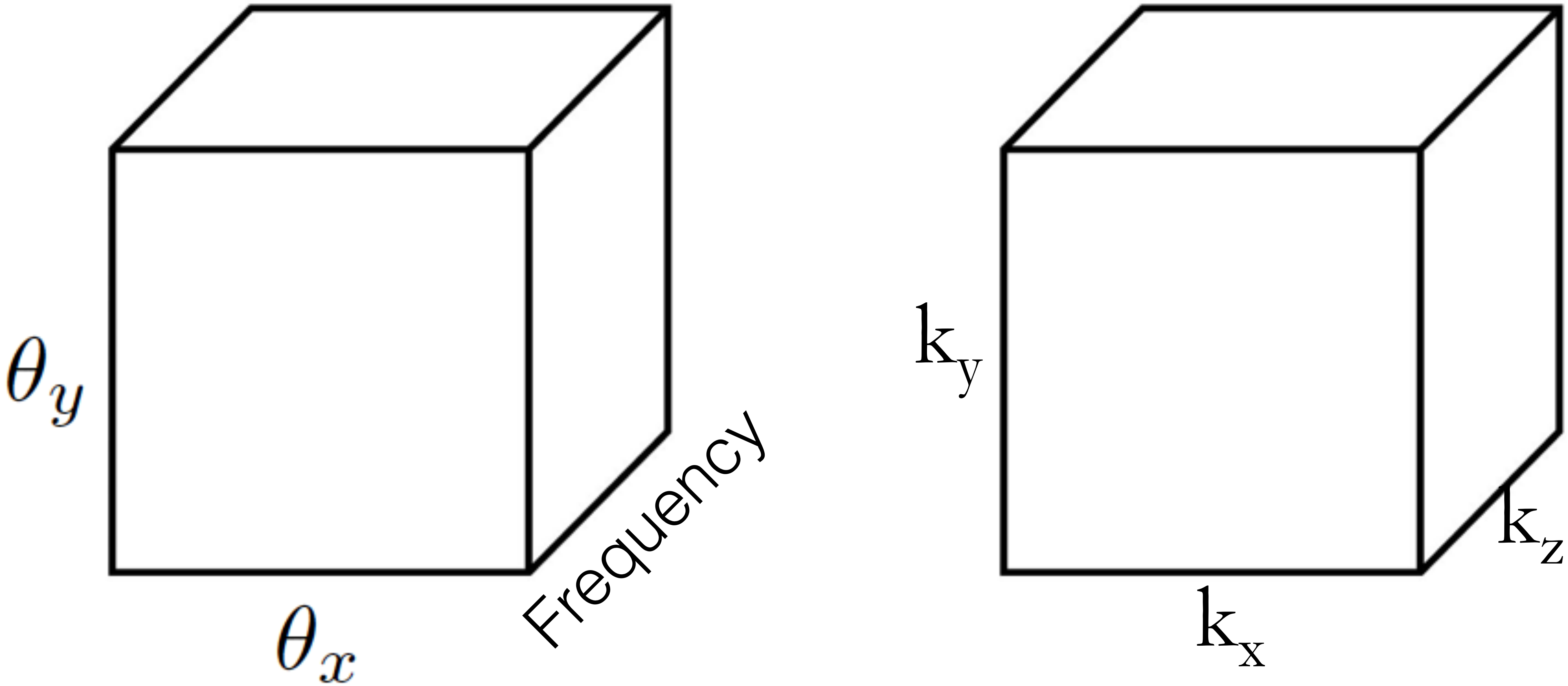
SKA
SKA AFRICA
SQUARE KILOMETRE ARRAY



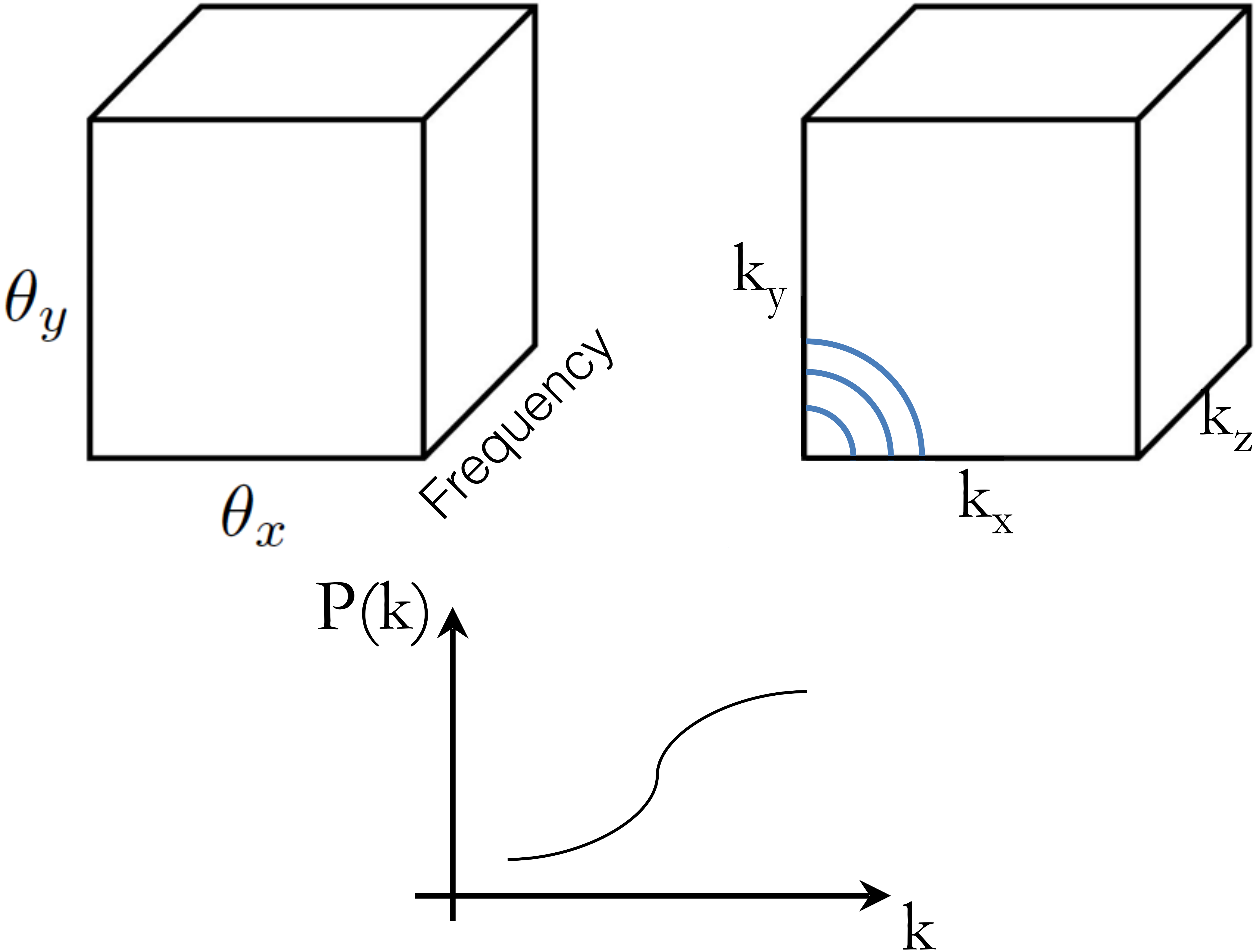


21cmFAST, Mesinger et al.

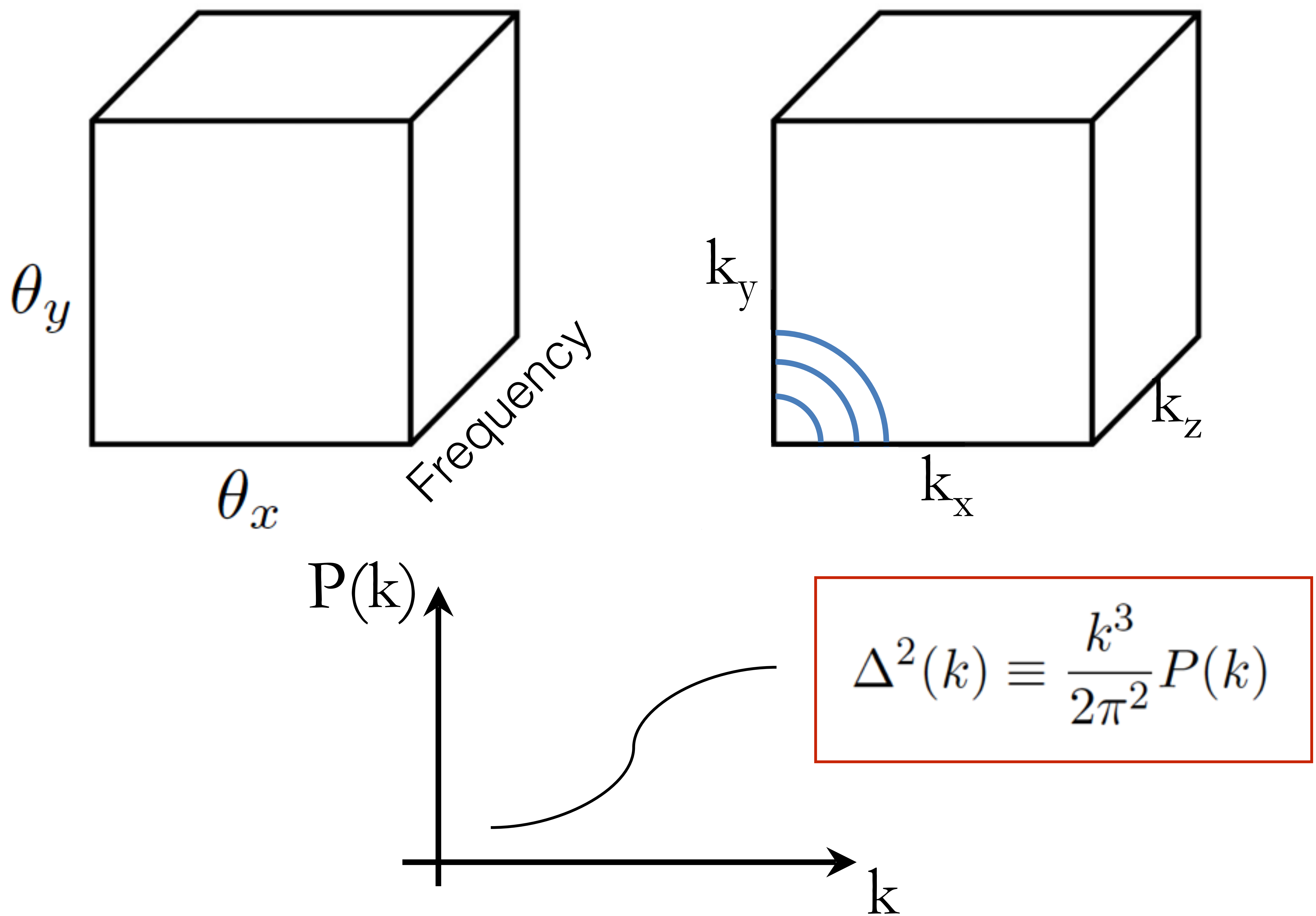

θ_y
θ_x
Frequency

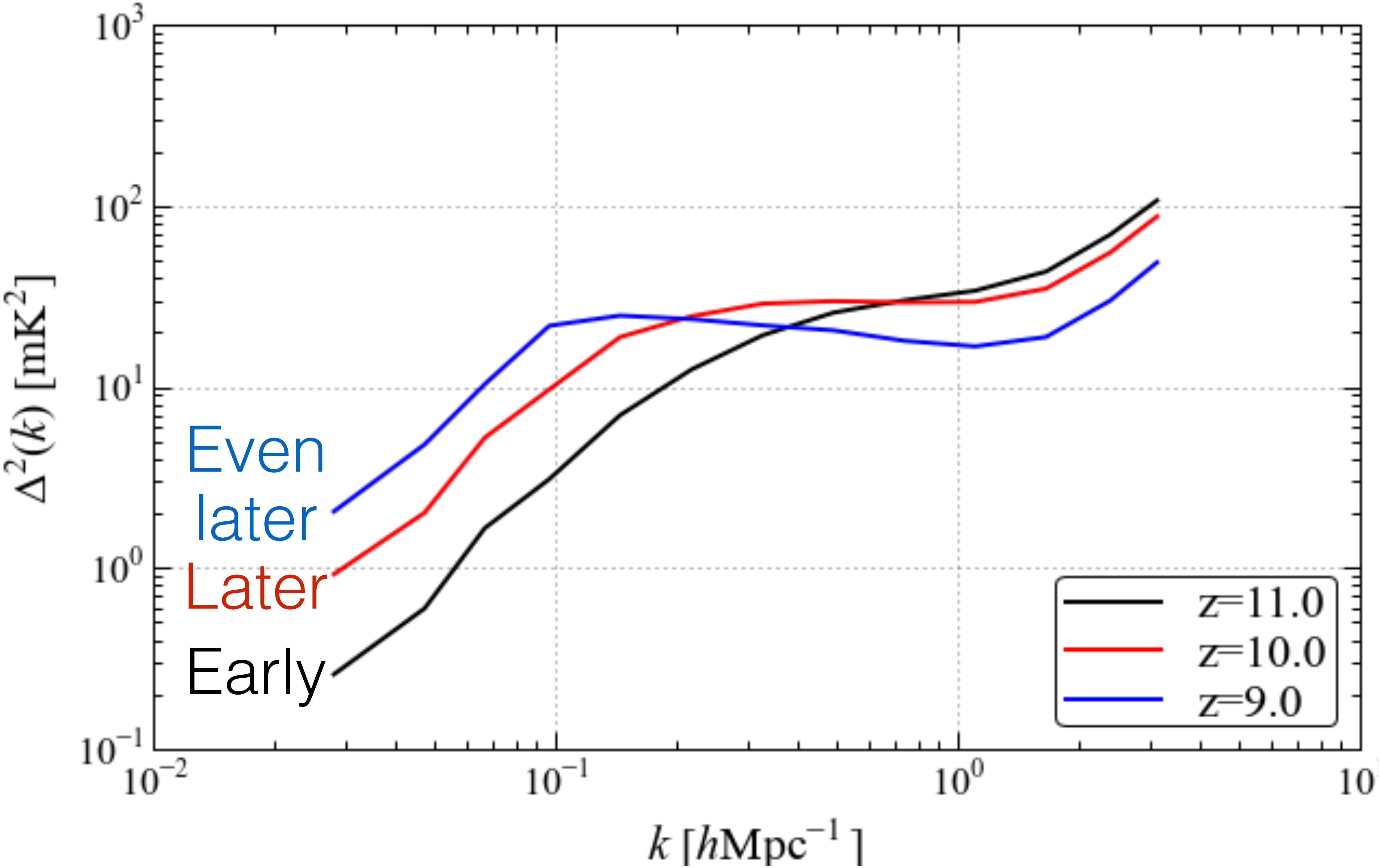


θ_y
θ_x
Frequency
k_y
k_x
k_z

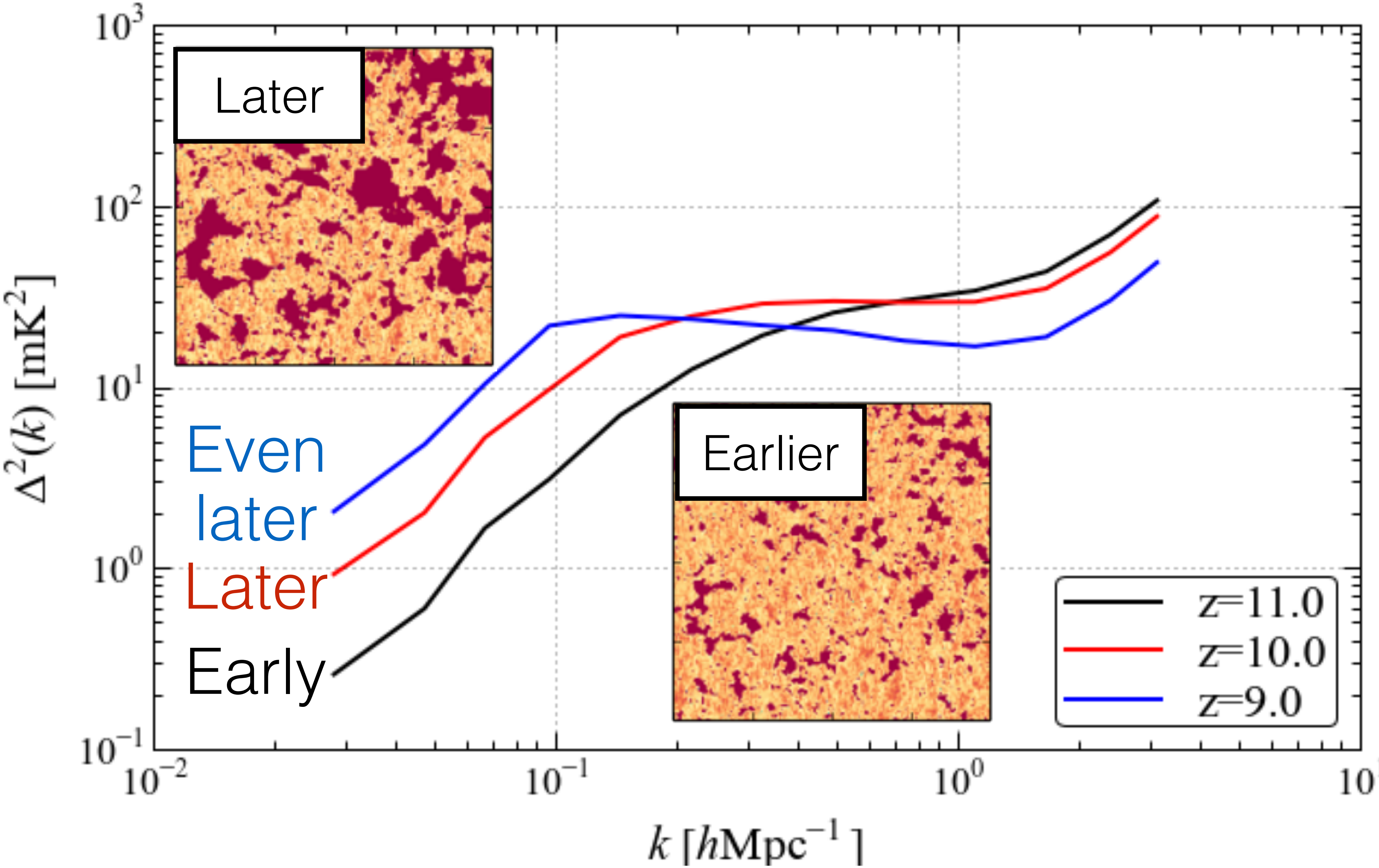


θ_y
θ_x
Frequency
k_y
k_x
k_z
P(k)
k

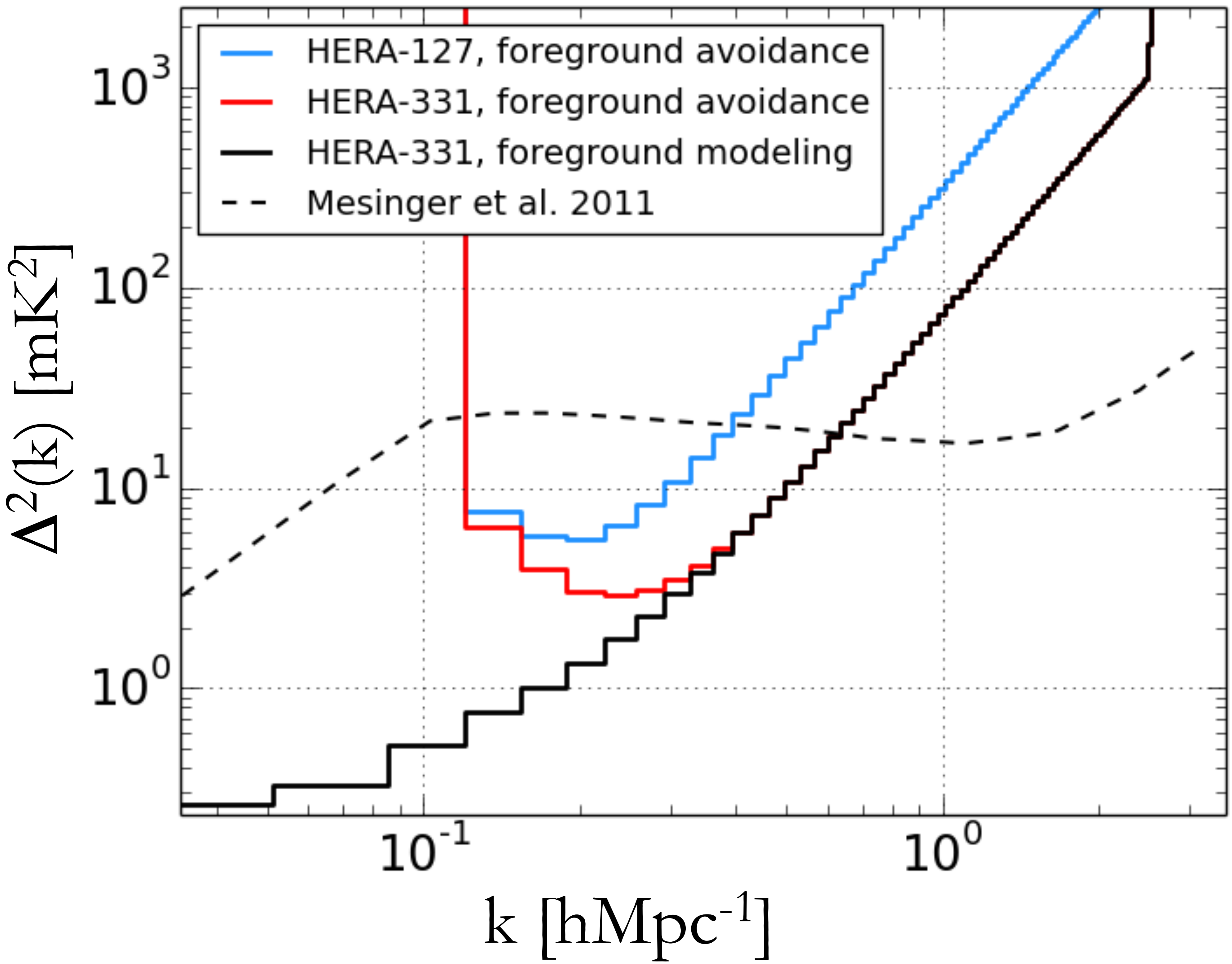

θ_y
θ_x
Frequency
k_y
k_x
k_z
P(k)
k
$\Delta^2(k) \equiv \frac{k^3}{2\pi^2} P(k)$



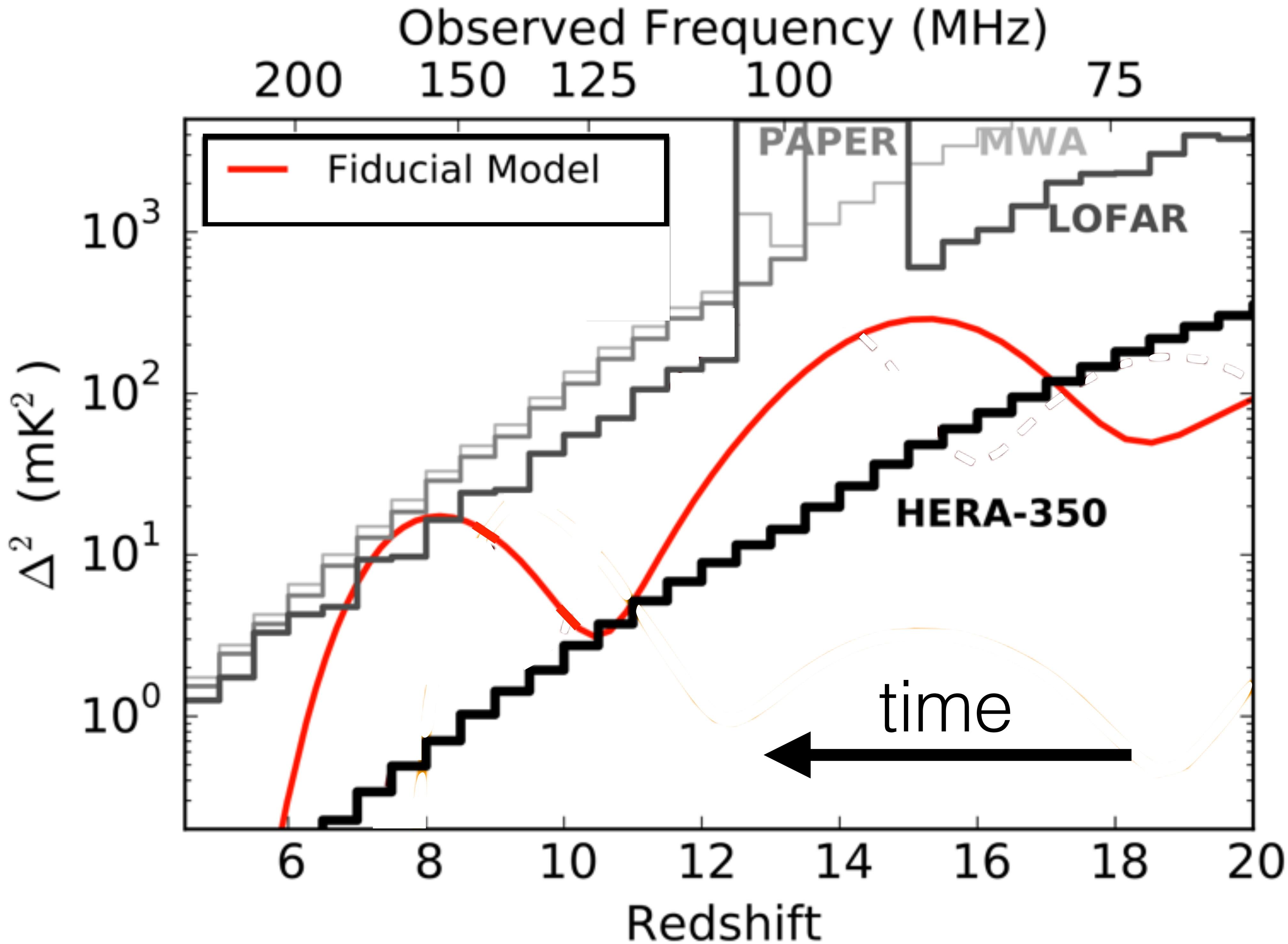
10^3
10^2
10^1
10^0
10^{-1}
$\Delta^2(k)$ [mK^2]
Even
later
Later
Early
z=11.0
z=10.0
z=9.0
10^{-2}
10^{-1}
10^0
10^1
k [hMpc^{-1}]


Later
Earlier
Even later
Later
Early
z=11.0
z=10.0
z=9.0
$\Delta^2(k)$ [mK^2]
k [hMpc^{-1}]
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}

HERA will make a high significance measurement within ~5 years



HERA will make a high significance measurement within ~5 years



A three-parameter reionization model

A three-parameter reionization model

- ζ: ionizing efficiency of first galaxies

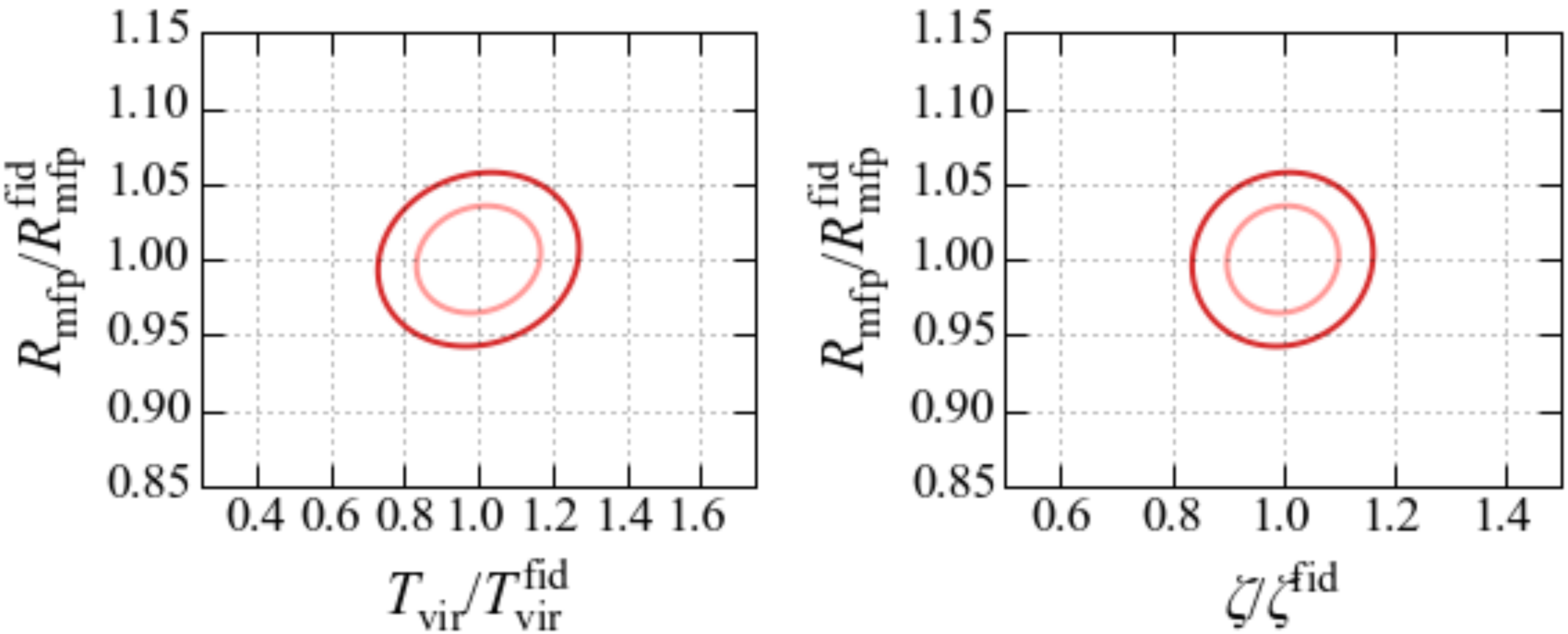
A three-parameter reionization model

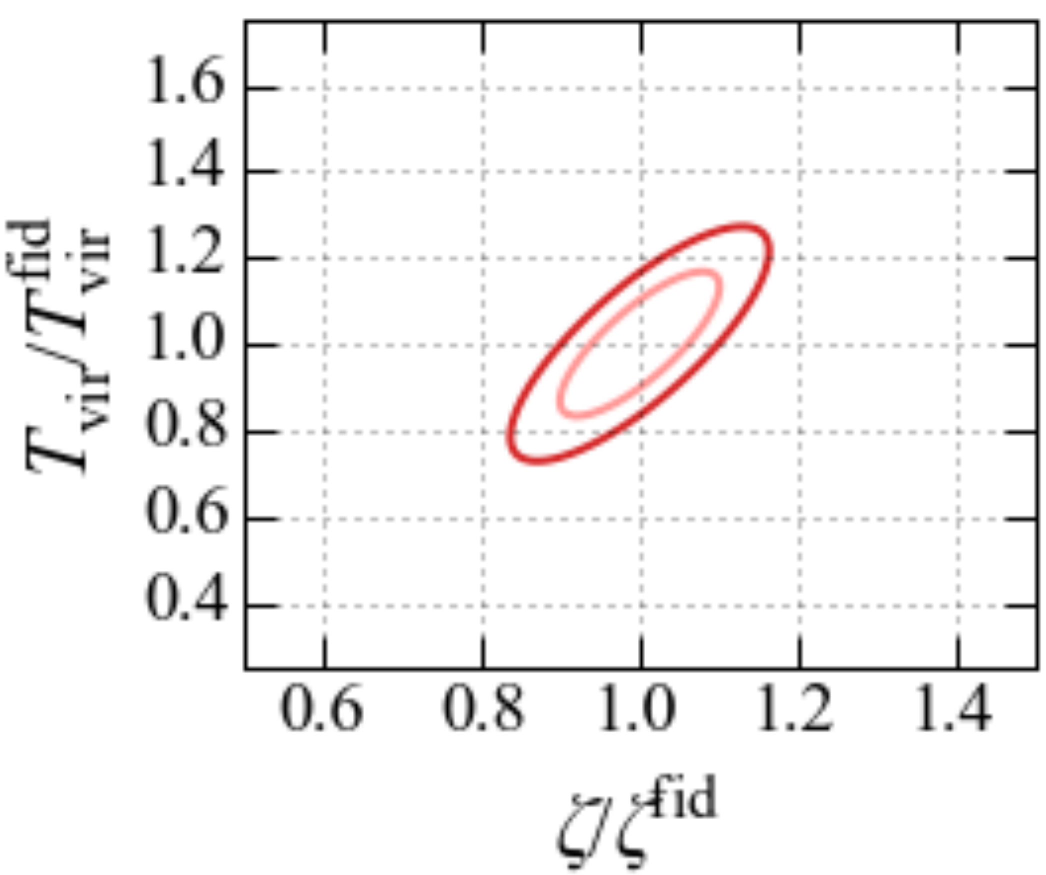
- ζ: ionizing efficiency of first galaxies

- T_{vir}: minimum virial temperature (proxy for mass) of first ionizing galaxies

A three-parameter reionization model

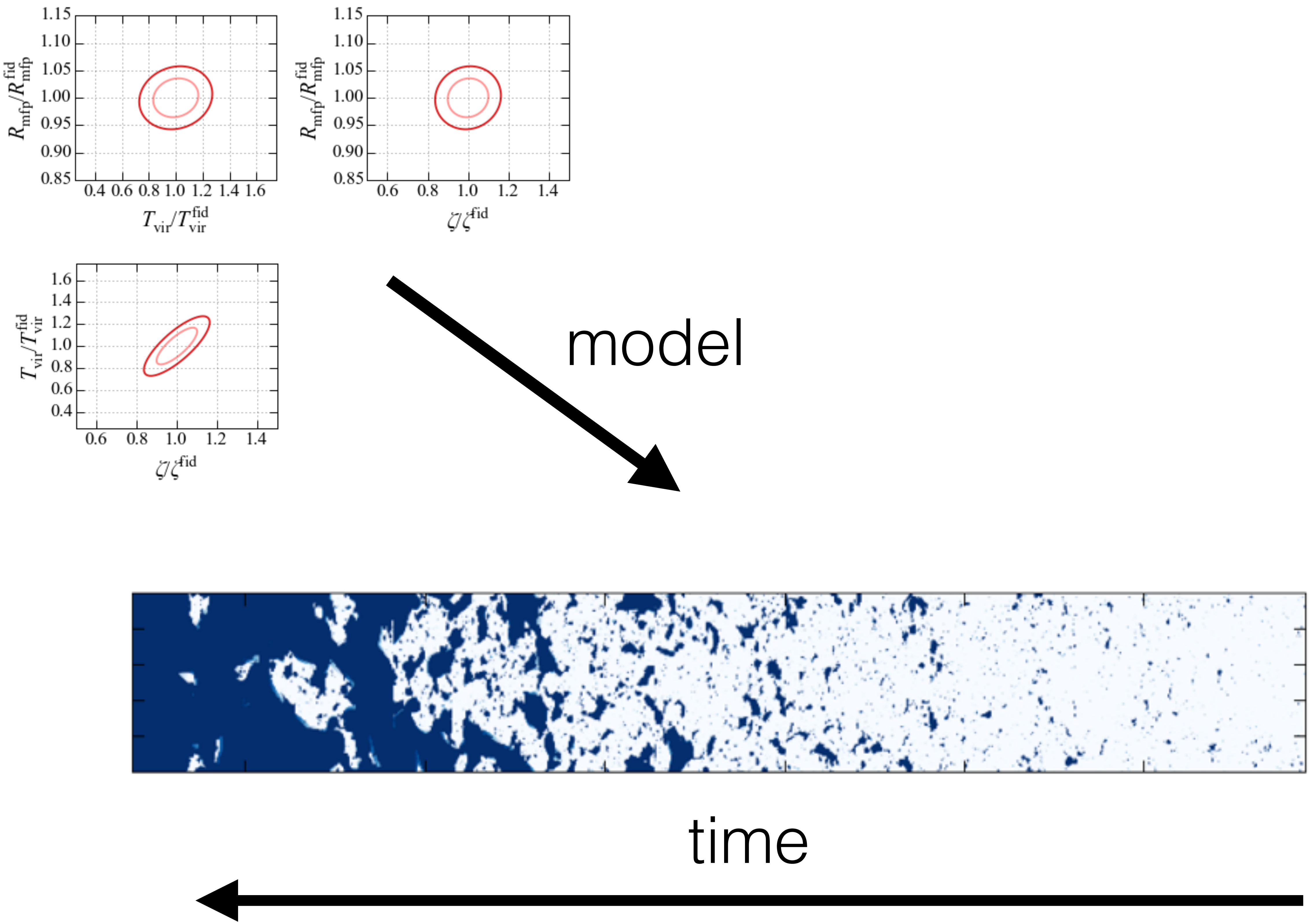
- ζ: ionizing efficiency of first galaxies
- T_{vir}: minimum virial temperature (proxy for mass) of first ionizing galaxies
- R_{mfp}: mean free path of ionizing photons

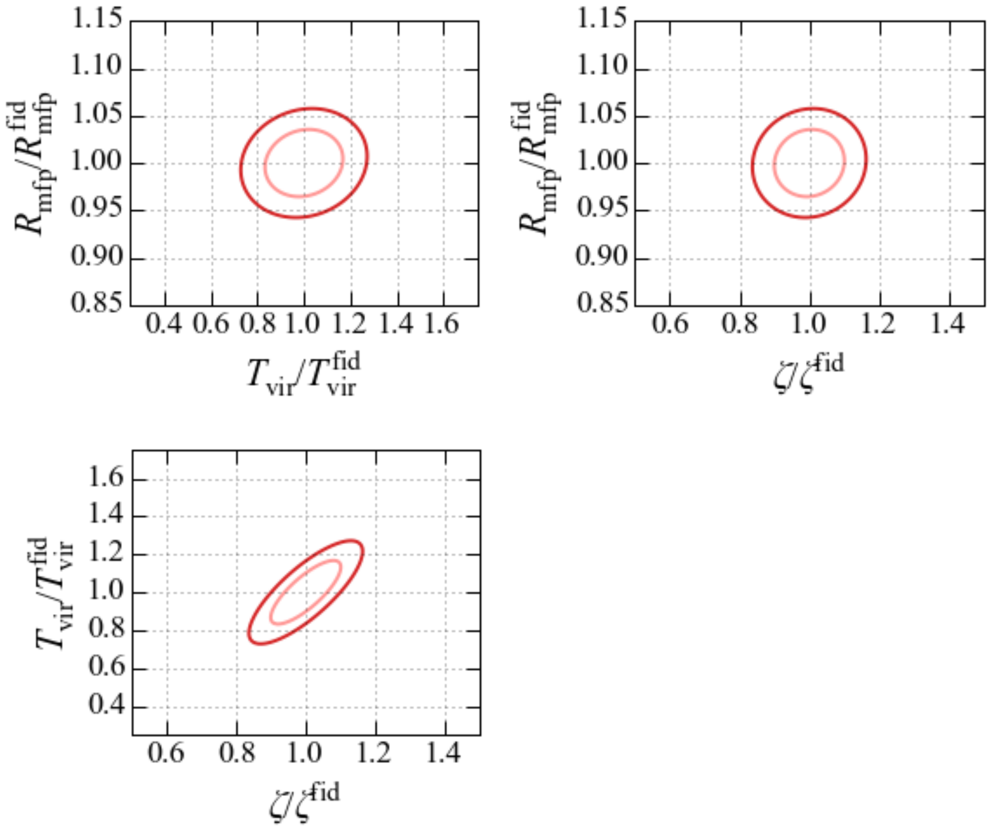




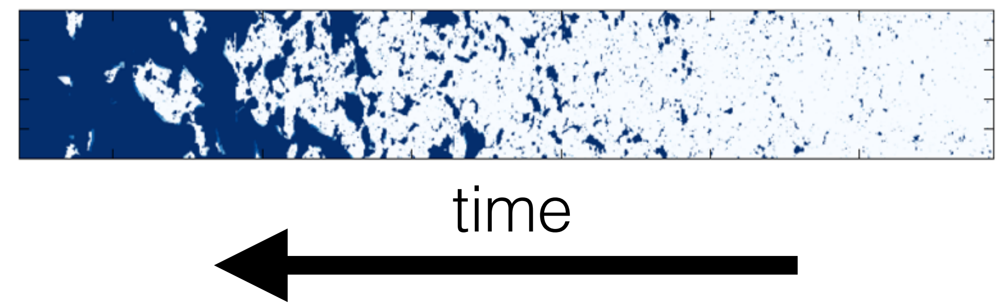
The recently commenced HERA experiment is forecasted to deliver ~5% errors on astrophysical parameters

AL & Parsons (2015b)

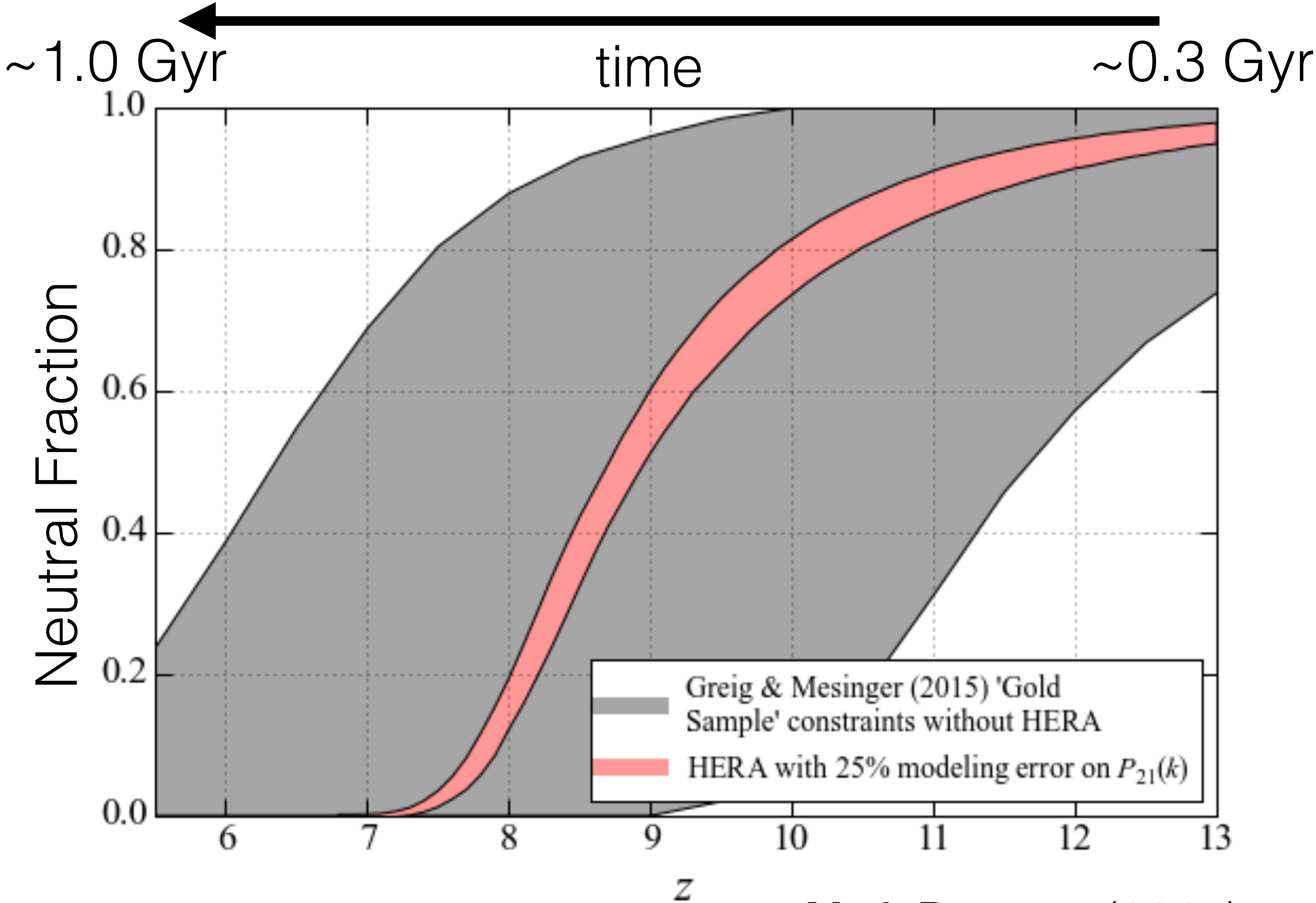

$R_{\rm mfp}/R_{\rm mfp}^{\rm fid}$
$T_{\rm vir}/T_{\rm vir}^{\rm fid}$
$\zeta/\zeta^{\rm fid}$
$R_{\rm mfp}/R_{\rm mfp}^{\rm fid}$
$\zeta/\zeta^{\rm fid}$
$T_{\rm vir}/T_{\rm vir}^{\rm fid}$
model
time



$R_{\rm mfp}/R_{\rm mfp}^{\rm fid}$
$T_{\rm vir}/T_{\rm vir}^{\rm fid}$
$R_{\rm mfp}/R_{\rm mfp}^{\rm fid}$
$\zeta/\zeta^{\rm fid}$
$T_{\rm vir}/T_{\rm vir}^{\rm fid}$
$\zeta/\zeta^{\rm fid}$



time



AL & Parsons (2015)

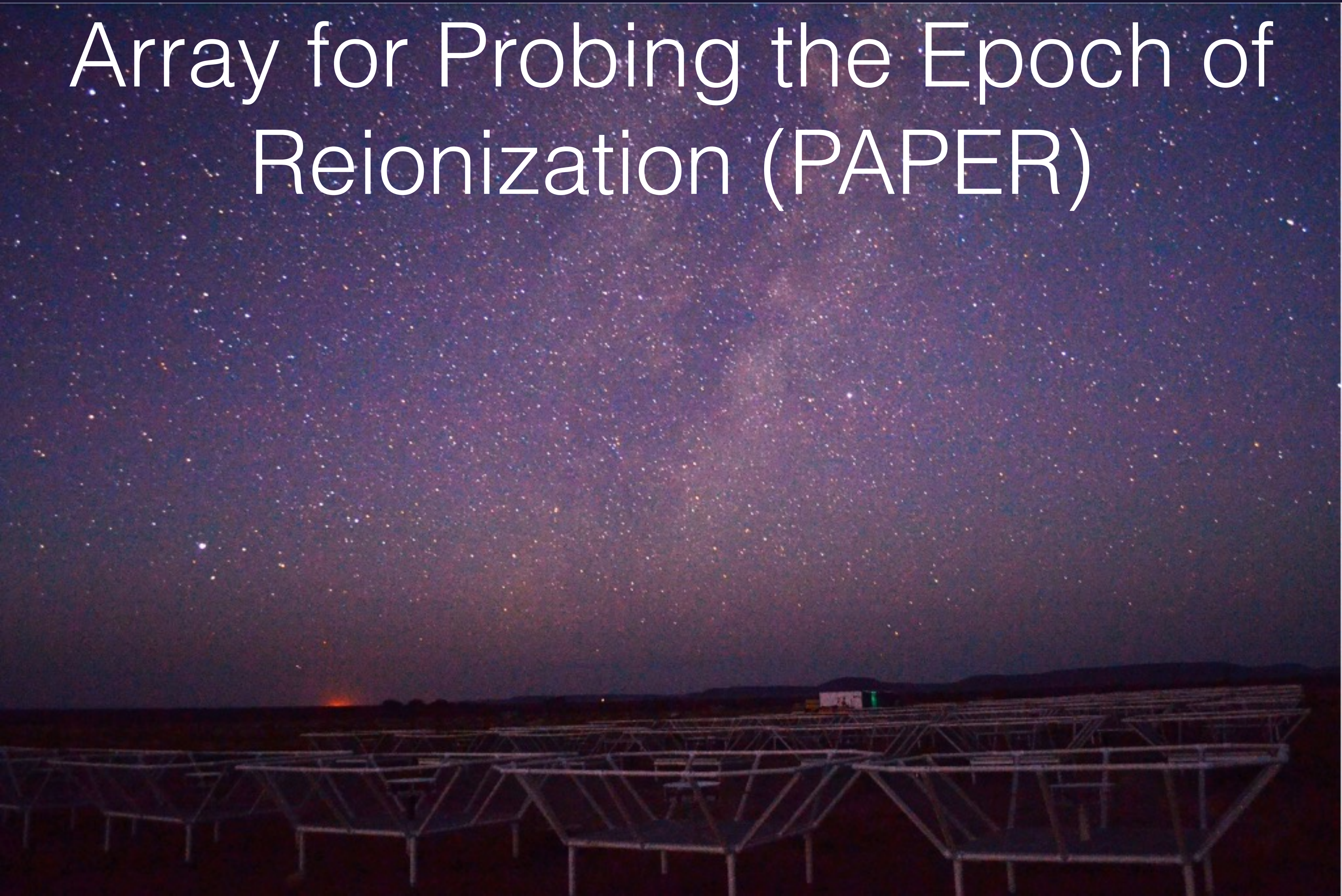
Presley, **AL** et al. (ongoing)

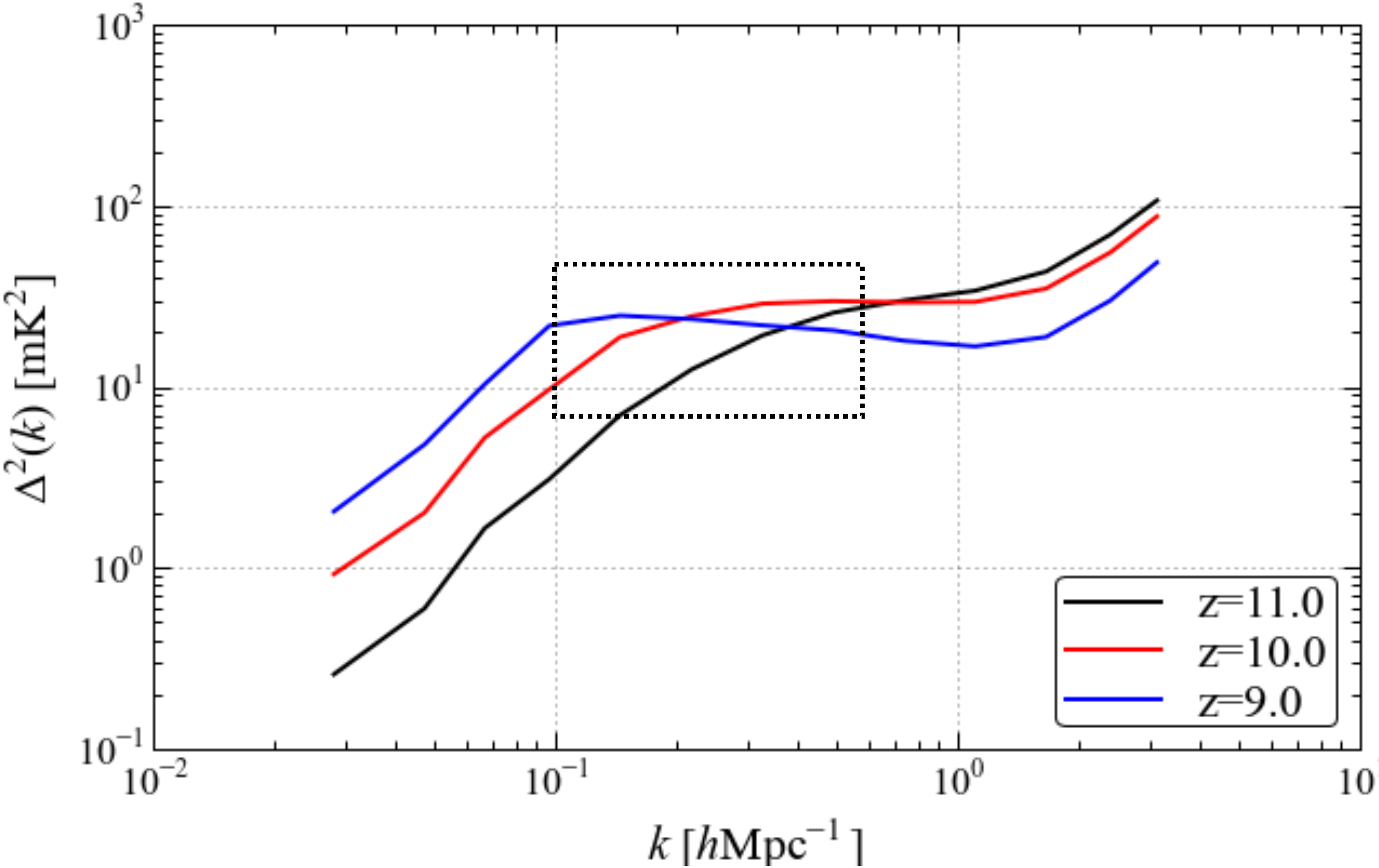
Questions we can now begin to ask

- How and when was the IGM heated?
- Were there any exotic mechanisms at play?
- What was the nature of the first stars and galaxies?
- Were galaxies solely responsible for reionization?

PAPER: state-of-the-art upper limits on the power spectrum

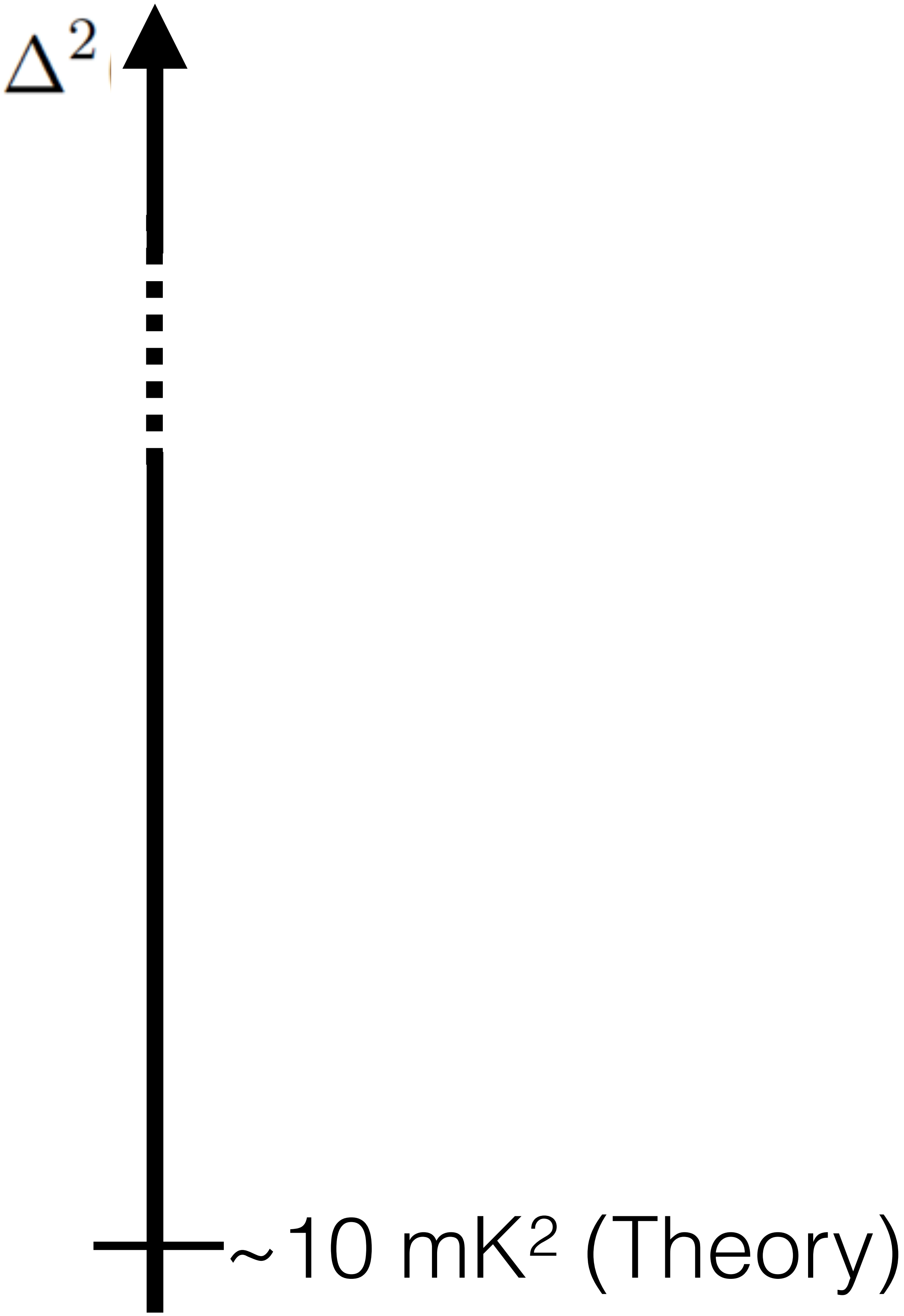
Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER)

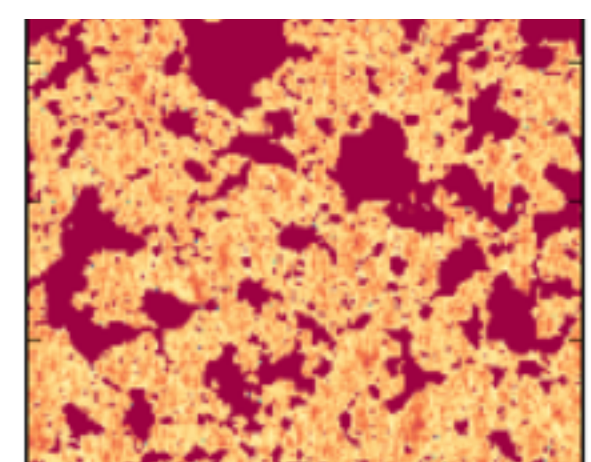




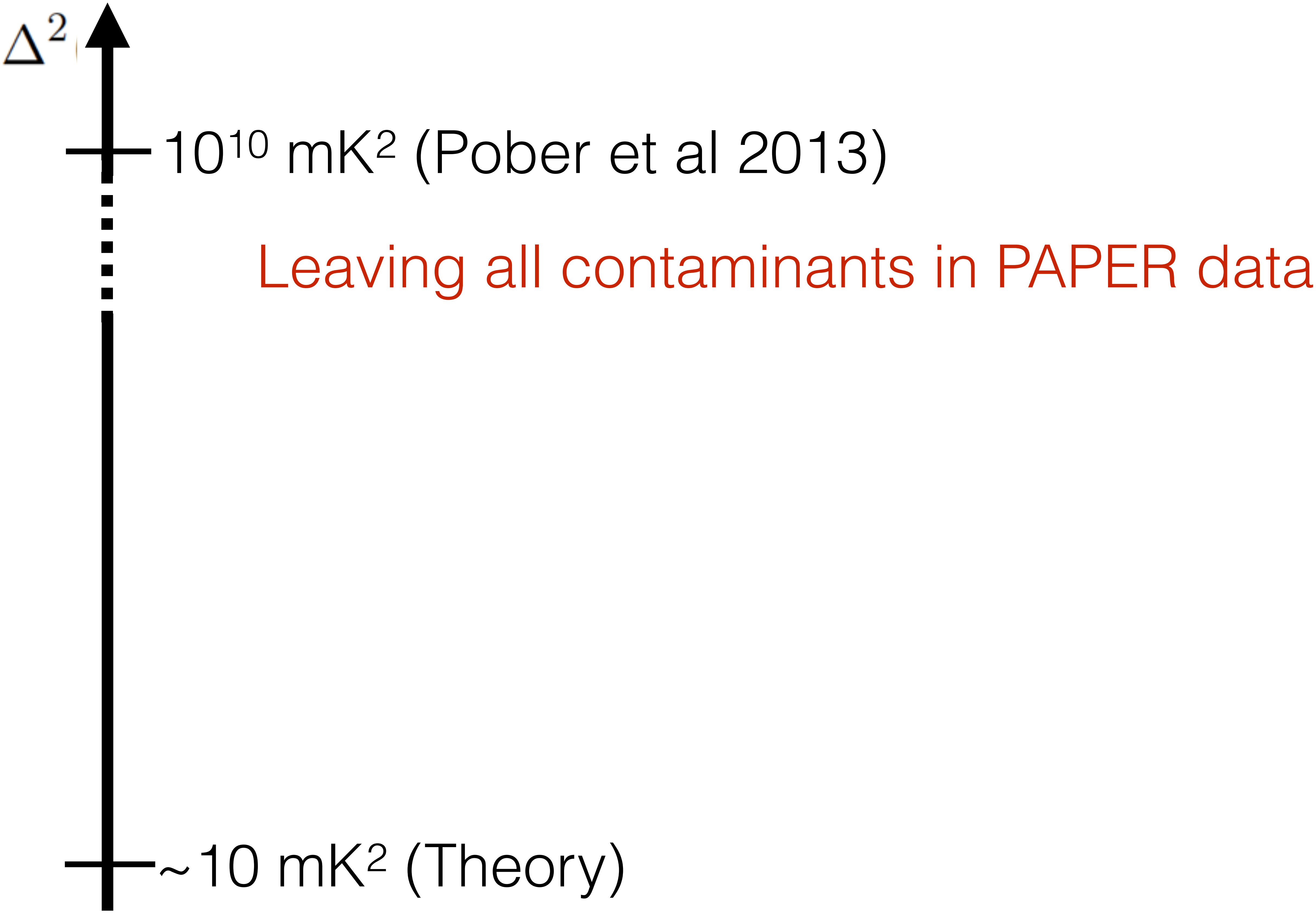
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-1}
10^0
10^1
$\Delta^2(k)$ [mK^2]
k [hMpc^{-1}]
z=11.0
z=10.0
z=9.0

Upper limits for 7.5 < z < 8.5

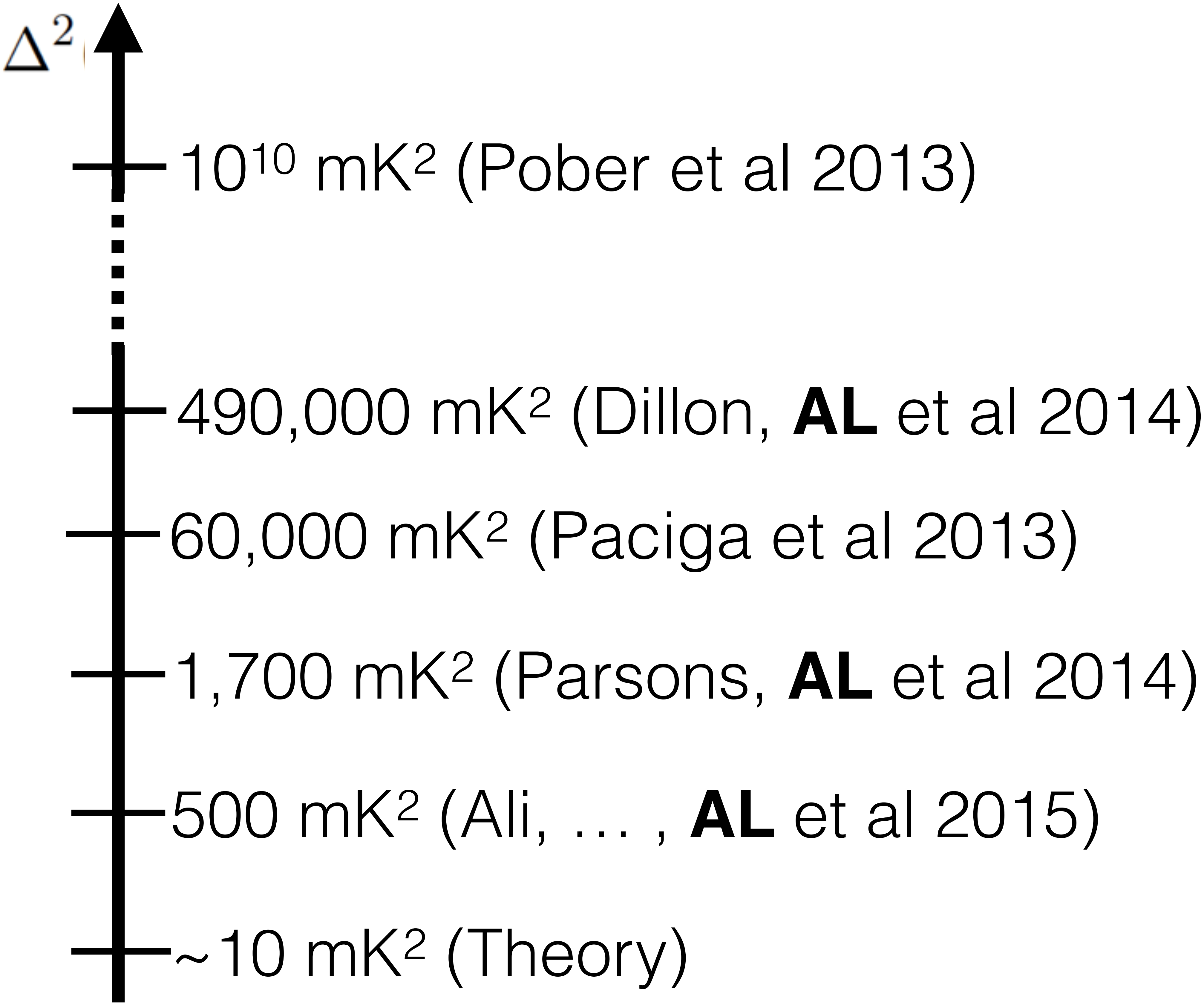




Upper limits for 7.5 < z < 8.5



Upper limits for 7.5 < z < 8.5



Upper limits for 7.5 < z < 8.5

Δ^2

“Software telescopes”

- Principal Component Analysis contamination mitigation (**AL** & Tegmark 2012, Switzer & **AL** 2014)
- Time-domain filtering (Parsons, **AL** et al. 2015)
- “Identical baseline calibration” (**AL** et al. 2010)
- Decorrelation techniques (**AL** et al. 2014b)
- Optimal estimators for 21cm cosmology (**AL** & Tegmark 2011; **AL** et al. 2014a)

500 mK^2 (Ali, … , **AL** et al 2015)

~10 mK^2 (Theory)

Upper limits for $7.5 < z < 8.5$

Δ^2

10^{10} mK^2 (Pober et al 2013)

490,000 mK^2 (Dillon, **AL** et al 2014)

60,000 mK^2 (Paciga et al 2013)

1,700 mK^2 (Parsons, **AL** et al 2014)

500 mK^2 (Ali, … , **AL** et al 2015)

~10 mK^2 (Theory)

Current PAPER upper limits rule out the possibility of an extremely **cold** intergalactic medium at t = 0.6 Gyr ($z \sim 8.4$)

Cold hydrogen
gas

(Relatively)
hot CMB

Cold hydrogen
gas

(Relatively)
hot CMB

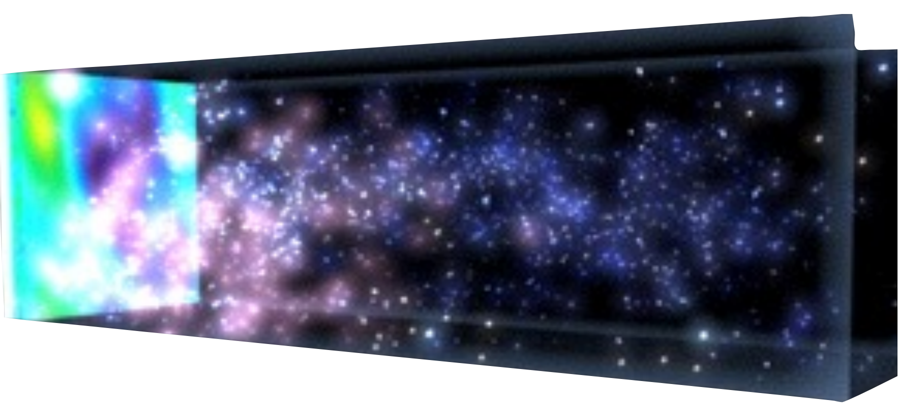
Cold hydrogen
gas

BIG
contrast,
large signal

(Relatively) hot CMB

Cold hydrogen gas

BIG contrast, large signal



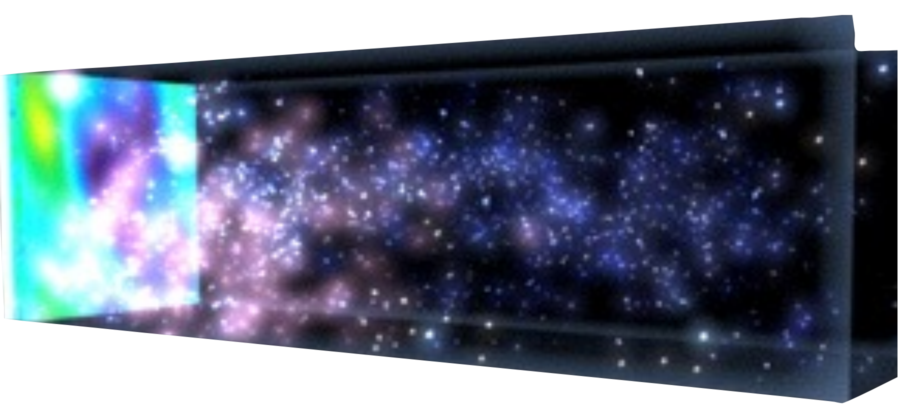
(Relatively) hot CMB

Warm hydrogen gas

Cold hydrogen gas

BIG contrast, large signal

(Relatively) hot CMB



(Relatively) hot CMB

Warm hydrogen gas

Small contrast, small signal

If the intergalactic medium had cooled adiabatically, the hydrogen gas would be cold enough to produce a large signal—large enough to be seen by now, with PAPER's sensitivity

If the intergalactic medium had cooled adiabatically, the hydrogen gas would be cold enough to produce a large signal—large enough to be seen by now, with PAPER's sensitivity

Some mechanism must have heated up the gas

For neutral fractions between 30% and 70%, PAPER observations imply $T_{gas} > 10$ K

In contrast, $T_{gas} = 1.18$ K assuming adiabatic cooling

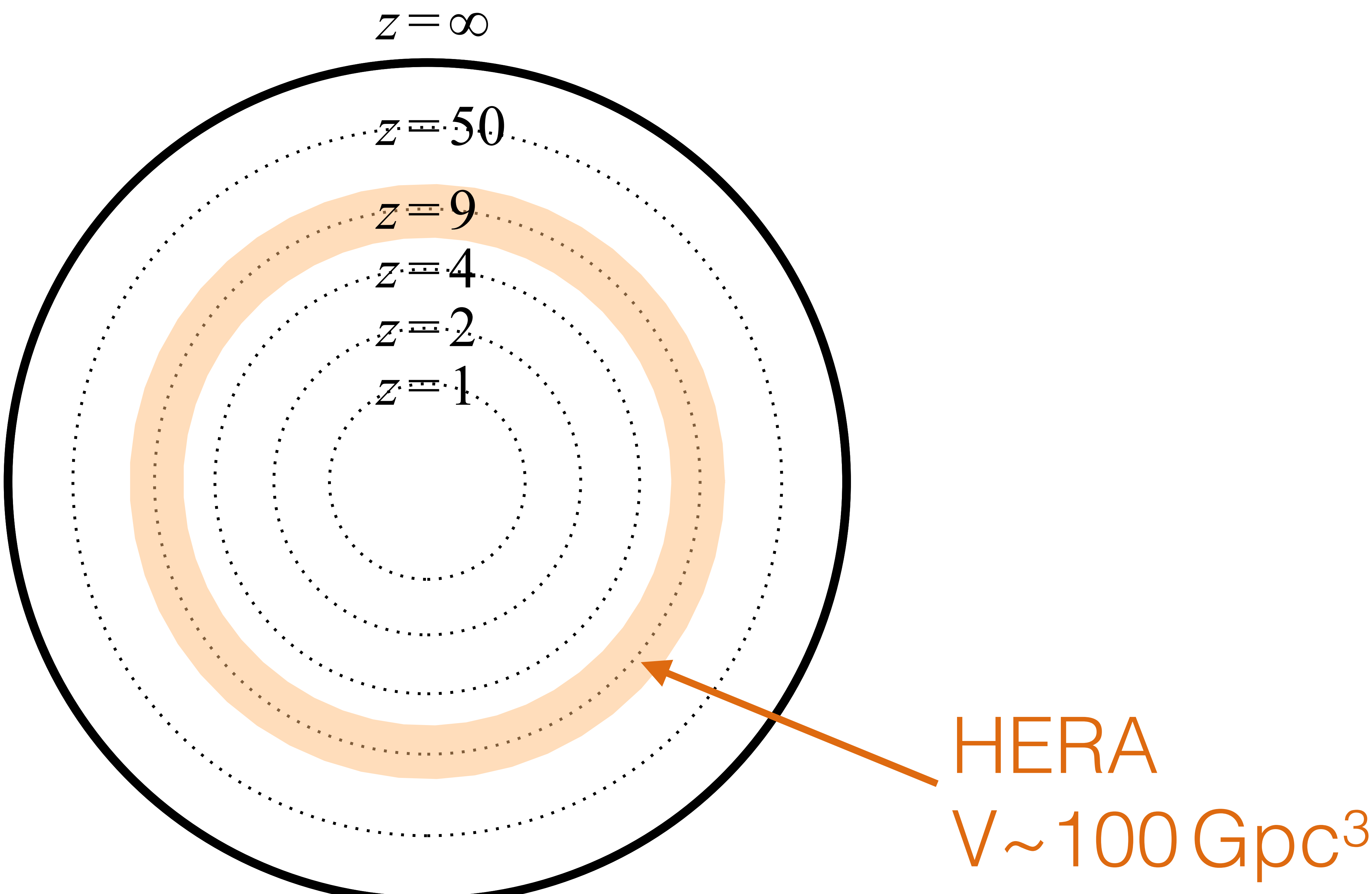
Thus, some sort of reheating must’ve taken place

Pober, Ali, …, **AL** et al. 2015, ApJ 809, 62

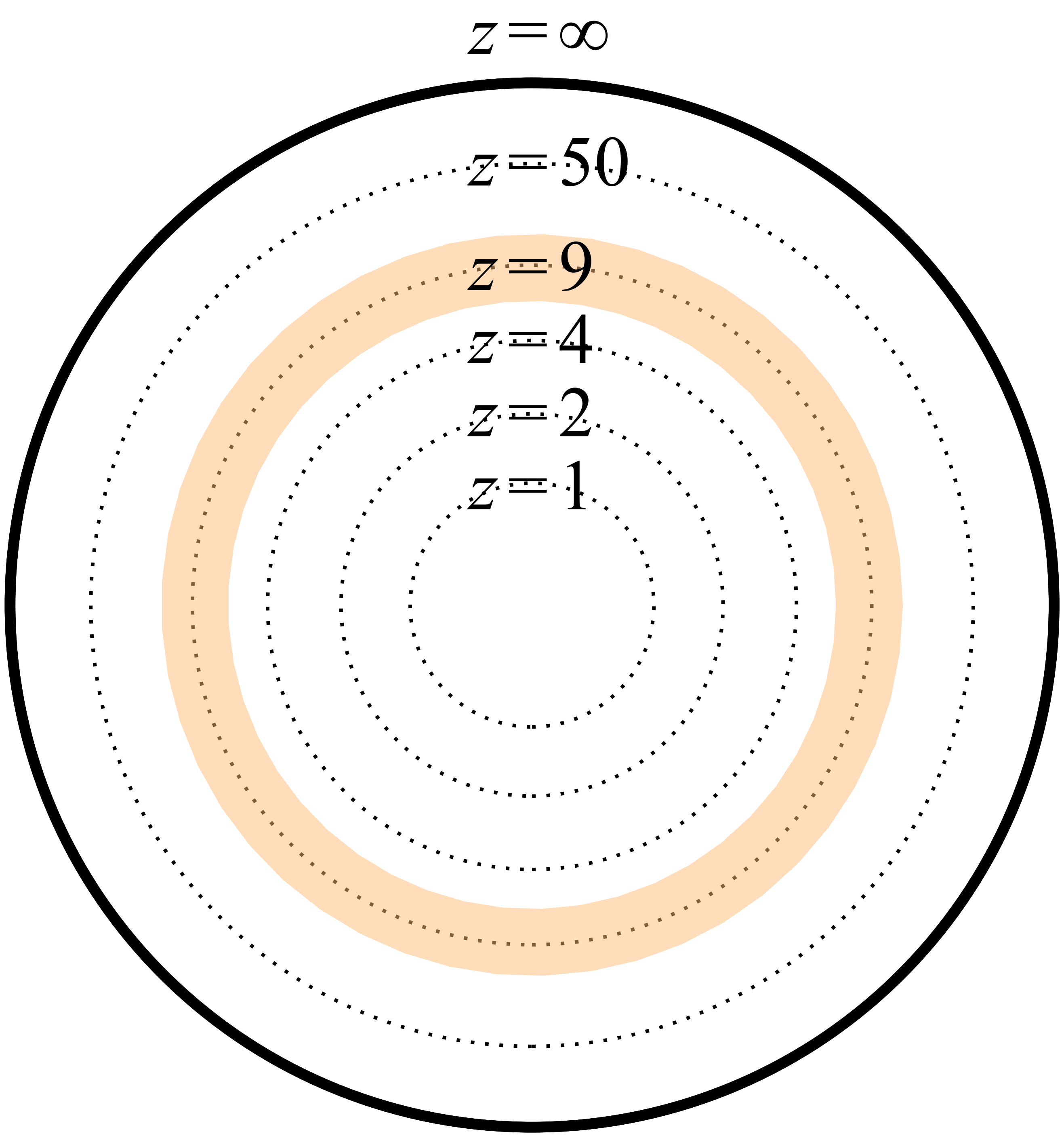
What about cosmology?

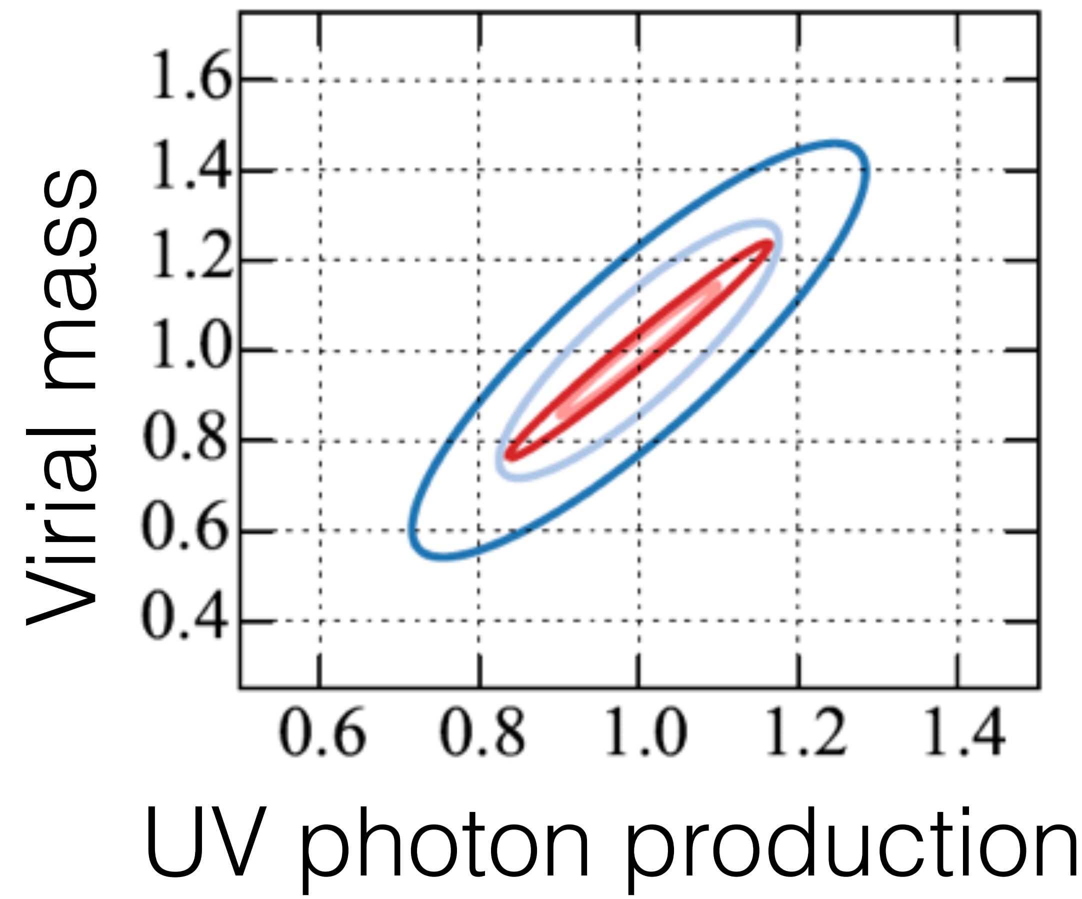
Astrophysics and cosmology are intertwined!

As we probe larger and larger portions of our Universe, theoretical models will inevitably have to incorporate cosmology



As we probe larger and larger portions of our Universe, theoretical models will inevitably have to incorporate cosmology





Cosmo params fixed

Cosmo params varied

AL & Parsons (2016)

There are lots of parameters to vary in models of Cosmic Dawn

- Ω_b Baryon density
- Ω_m Matter density
- σ_8 Density fluctuation amplitude
- H_0 Hubble expansion
- n_s Spectral index of density perturbations
- $T_{\rm vir}$ Minimum virial temperature of first galaxies
- $R_{\rm mfp}$ Mean free path of ionizing photons
- ζ UV ionizing efficiency
- ζ_x X-ray ionizing efficiency
- ν_c X-ray photon cut-off
- α X-ray spectral index

How do we adequately explore the parameter space when there are so many parameters?

Data

Likelihood

$$p(\boldsymbol{\theta}|\mathbf{d}, \mathcal{M}) \propto \mathcal{L}(\mathbf{d}|\boldsymbol{\theta}, \mathcal{M})\pi(\boldsymbol{\theta}|\mathcal{M})$$

Parameters

Model

Prior

$$p(\boldsymbol{\theta}|\mathbf{d},\mathcal{M}) \propto \mathcal{L}(\mathbf{d}|\boldsymbol{\theta},\mathcal{M})\pi(\boldsymbol{\theta}|\mathcal{M})$$

$$\mathcal{L} \propto \exp\left(-\frac{1}{2}[\mathbf{d}-\mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]^t\boldsymbol{\Sigma}^{-1}[\mathbf{d}-\mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]\right)$$

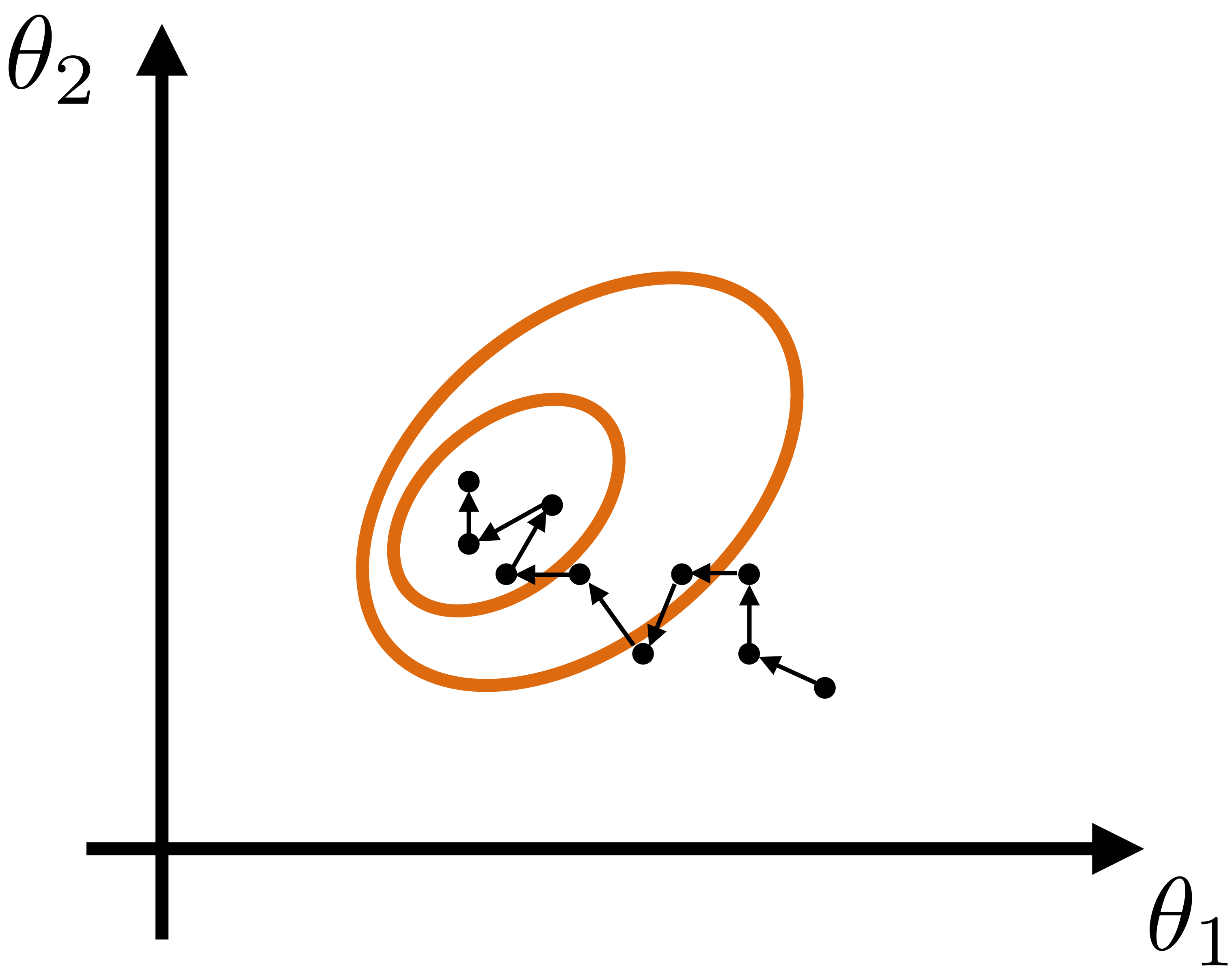
Data

Theoretical prediction for observables

Error covariance

$$p(\boldsymbol{\theta}|\mathbf{d},\mathcal{M}) \propto \mathcal{L}(\mathbf{d}|\boldsymbol{\theta},\mathcal{M})\pi(\boldsymbol{\theta}|\mathcal{M})$$

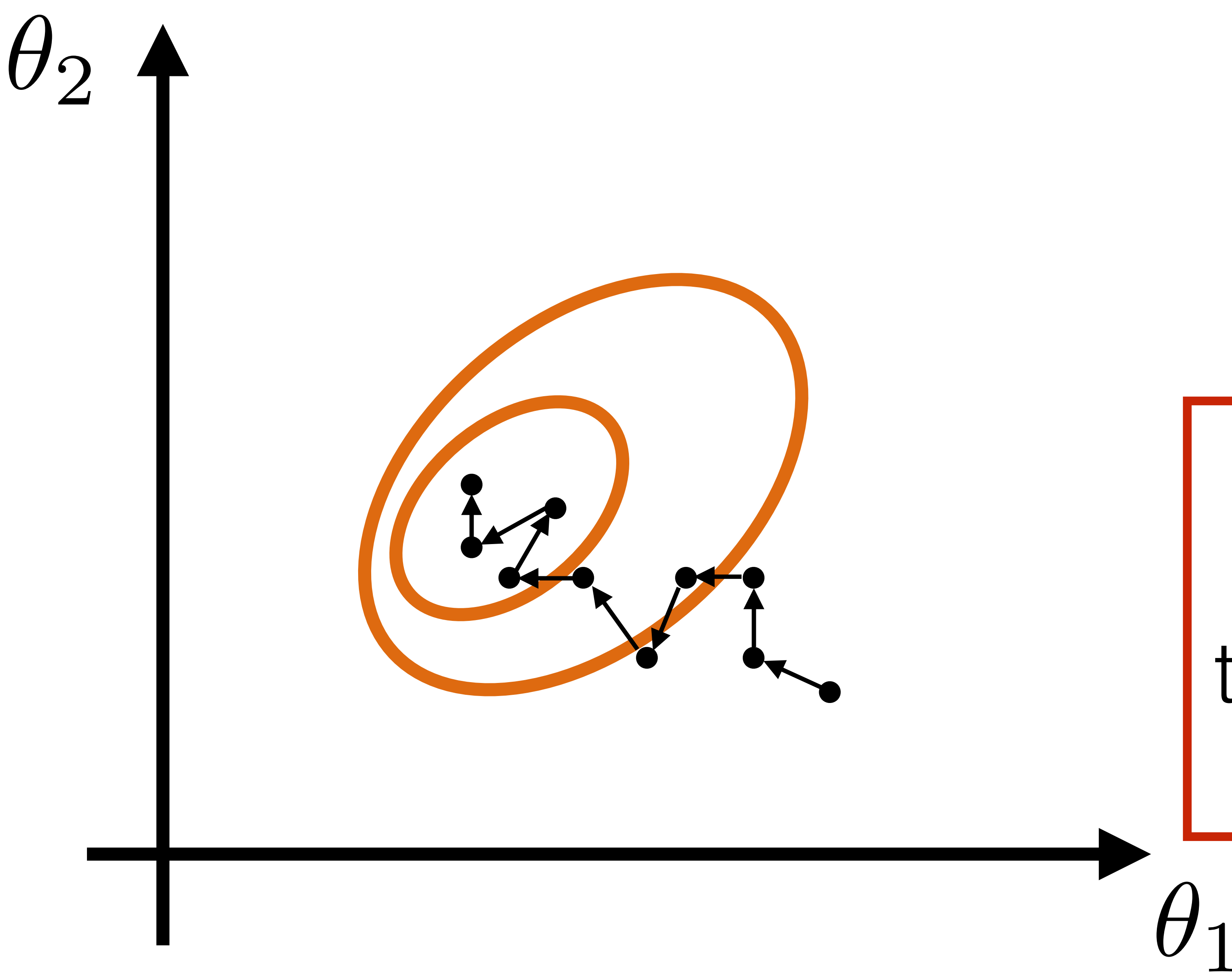
$$\mathcal{L} \propto \exp\left(-\frac{1}{2}[\mathbf{d}-\mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]^t\boldsymbol{\Sigma}^{-1}[\mathbf{d}-\mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]\right)$$



MCMC analysis: “Diffuse” through parameter space

$$p(\boldsymbol{\theta}|\mathbf{d},\mathcal{M}) \propto \mathcal{L}(\mathbf{d}|\boldsymbol{\theta},\mathcal{M})\pi(\boldsymbol{\theta}|\mathcal{M})$$

$$\mathcal{L} \propto \exp\left(-\frac{1}{2}[\mathbf{d}-\mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]^t\boldsymbol{\Sigma}^{-1}[\mathbf{d}-\mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]\right)$$

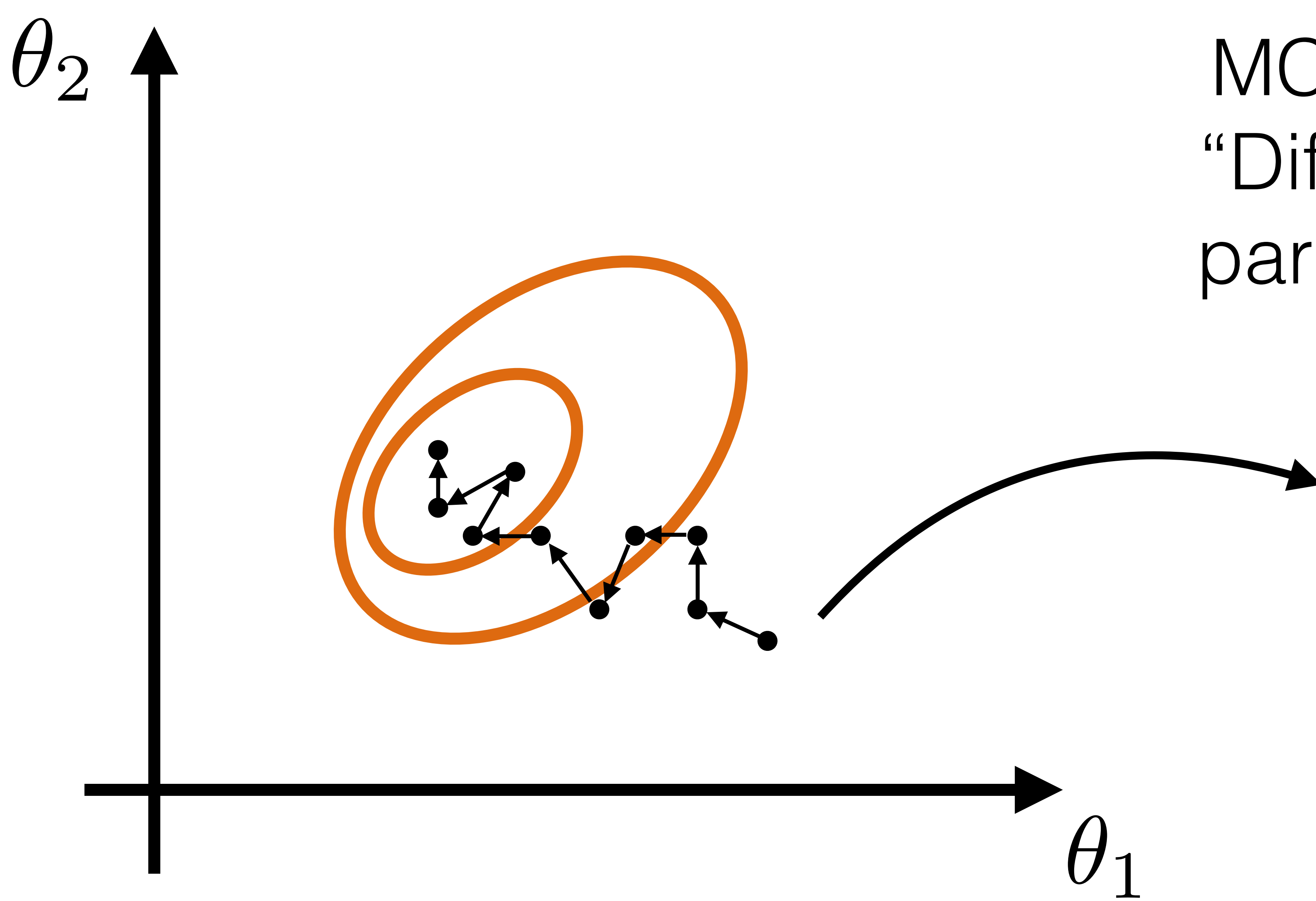


MCMC analysis:
"Diffuse" through parameter space

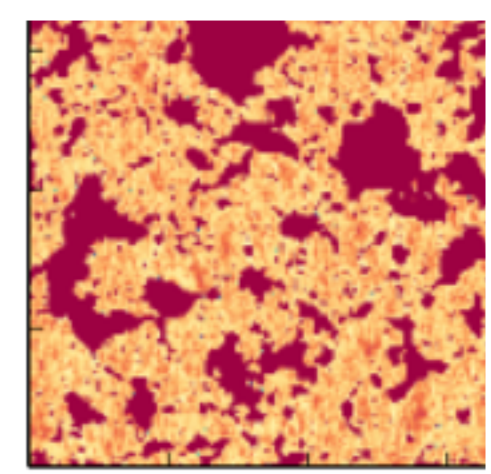
Problem: each point may still talk O(a day) to evaluate, and O(10^4) may be required

$$p(\boldsymbol{\theta}|\mathbf{d},\mathcal{M}) \propto \mathcal{L}(\mathbf{d}|\boldsymbol{\theta},\mathcal{M})\pi(\boldsymbol{\theta}|\mathcal{M})$$

$$\mathcal{L} \propto \exp\left(-\frac{1}{2}[\mathbf{d}-\mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]^t\boldsymbol{\Sigma}^{-1}[\mathbf{d}-\mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]\right)$$

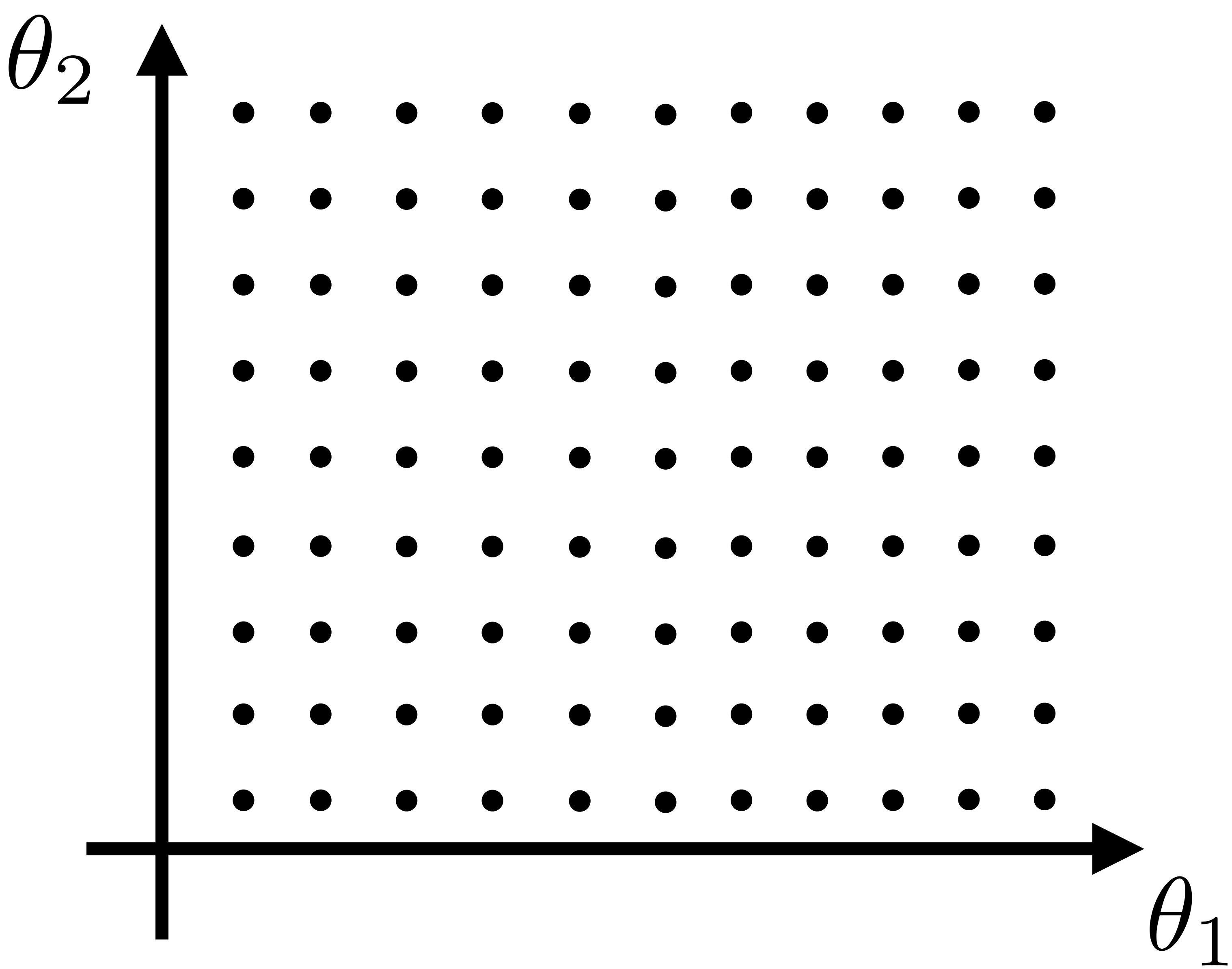


MCMC analysis:
“Diffuse” through
parameter space



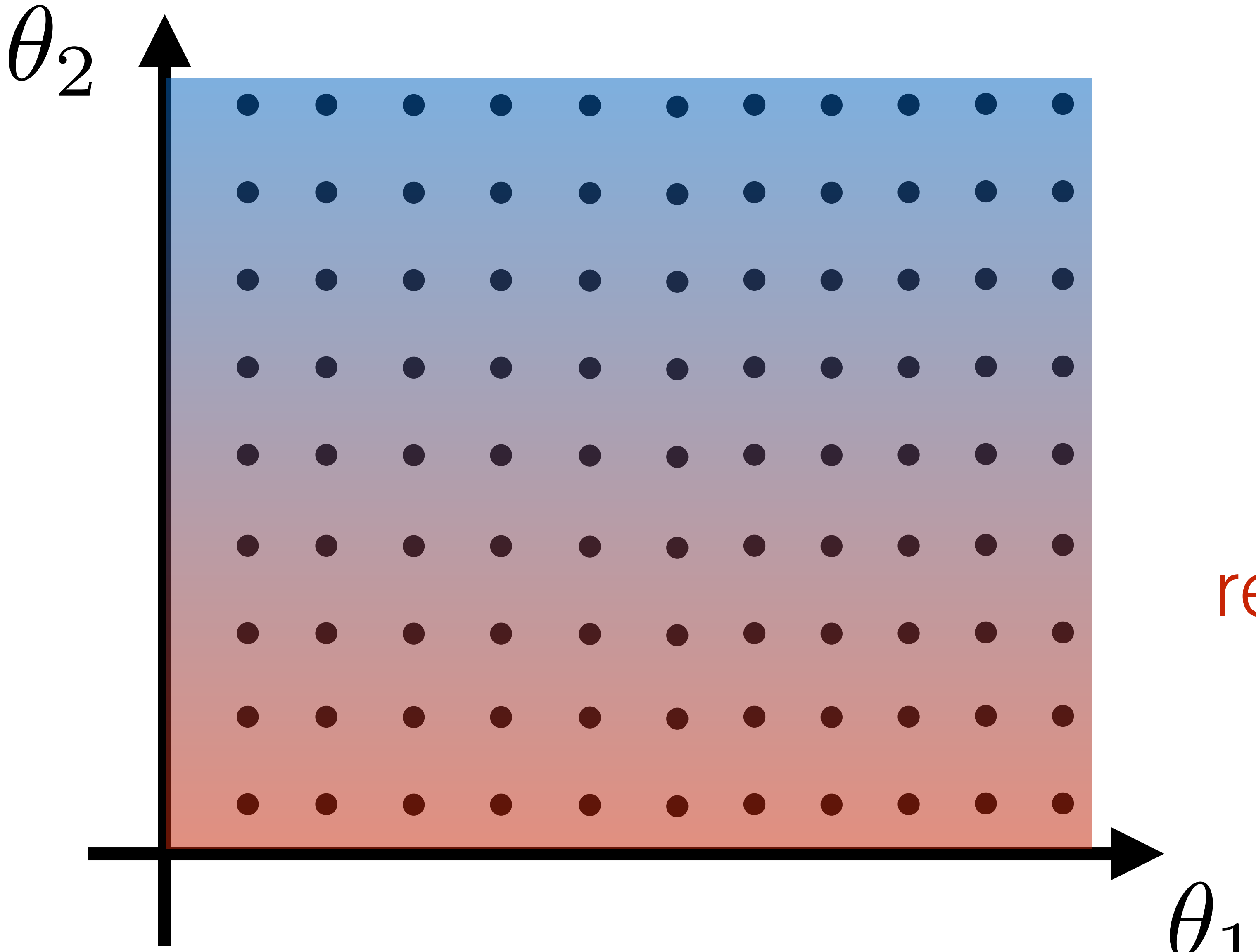
Solution: **emulate** the theoretical model

$$\mathcal{L} \propto \exp\left(-\frac{1}{2}[\mathbf{d} - \mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]^t \boldsymbol{\Sigma}^{-1} [\mathbf{d} - \mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]\right)$$



Solution: **emulate** the theoretical model

$$\mathcal{L} \propto \exp\left(-\frac{1}{2}[\mathbf{d} - \mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]^t \boldsymbol{\Sigma}^{-1} [\mathbf{d} - \mathbf{f}_{\mathcal{M}}(\boldsymbol{\theta})]\right)$$



$$\mathbf{f}^{\mathrm{approx}}_{\mathcal{M}}(\boldsymbol{\theta})$$

Interpolate over the results from pre-computed training samples

Case study: A fast emulator of semi-numeric cosmic dawn codes to allow MCMCs over a large number of parameters

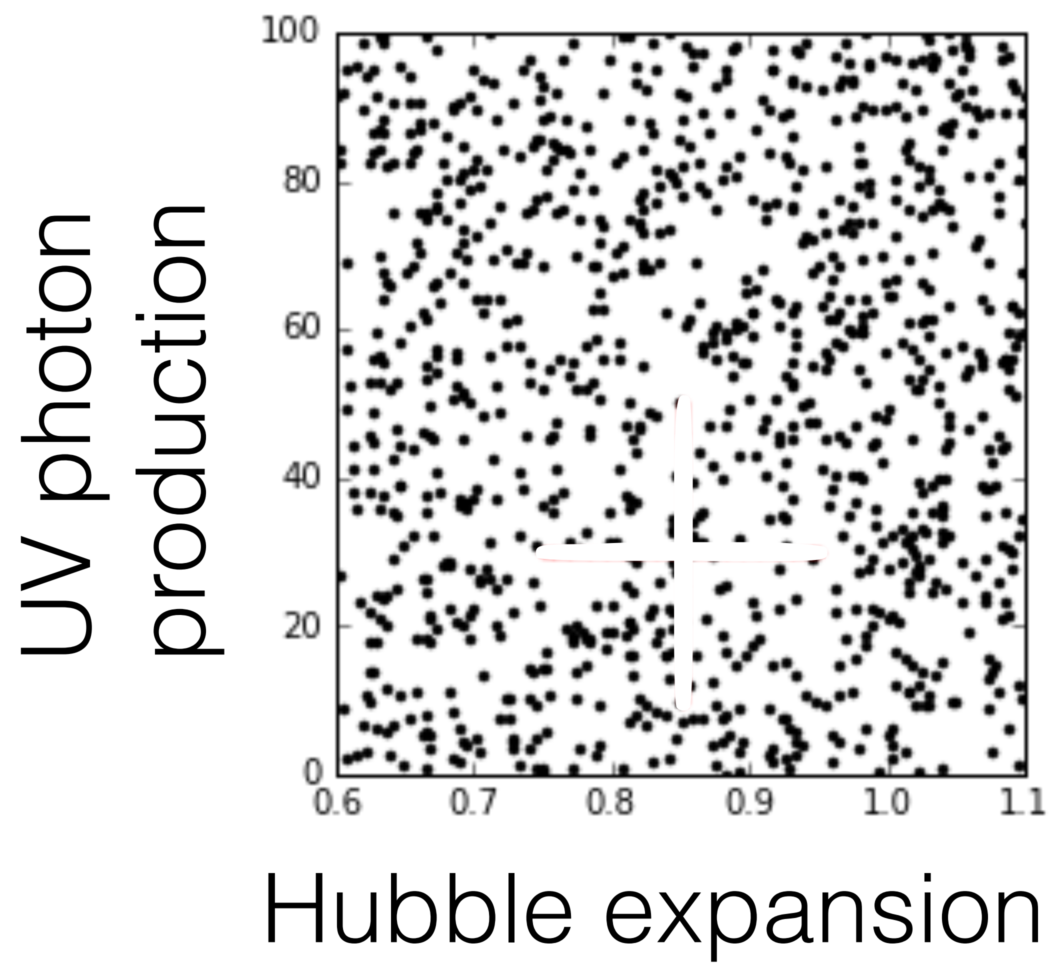
PYCAPE: PYthon toolbox for Cosmic dAwn Parameter Estimation

Kern, **AL** et al. (2017)

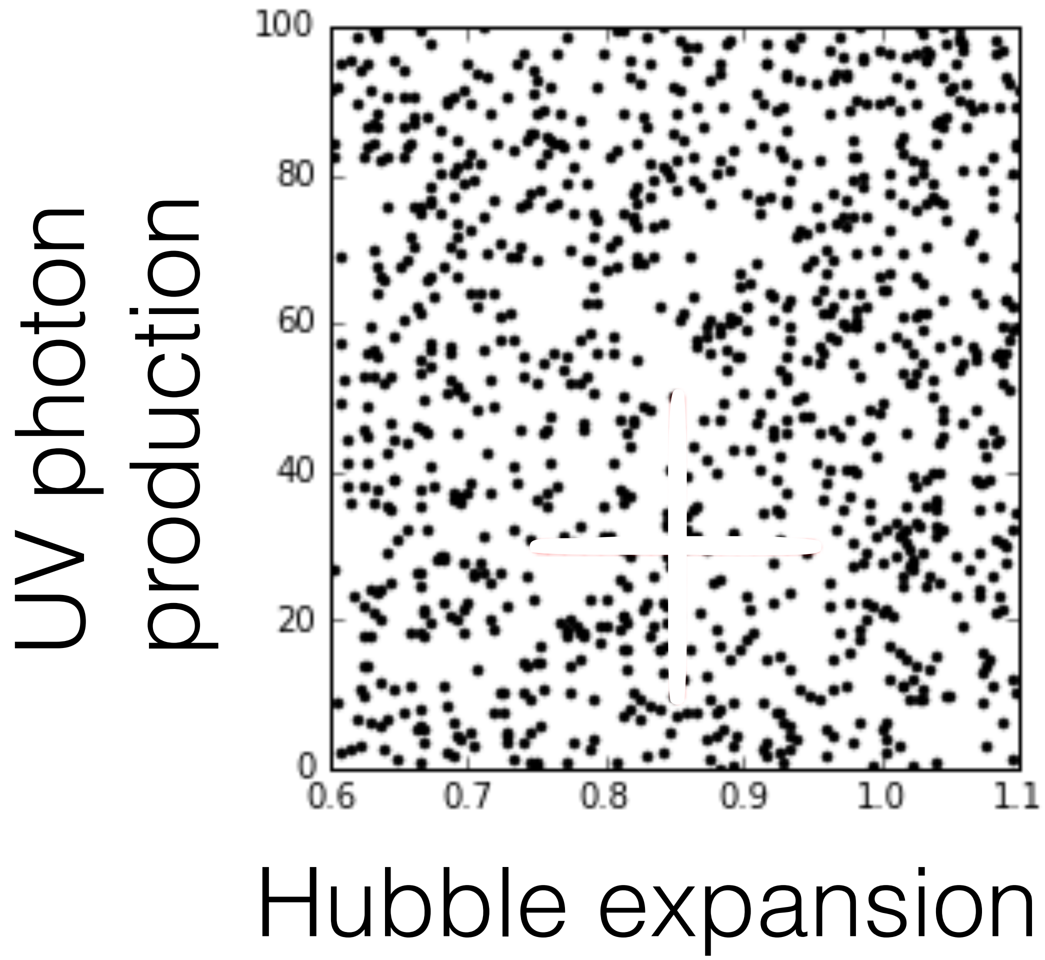
Sampling N points along each of M axes requires N^M runs of a simulation. For N=10 points and M=11 parameters, this would require 10^{11} s ~3000 years even if each simulation took just a single second!

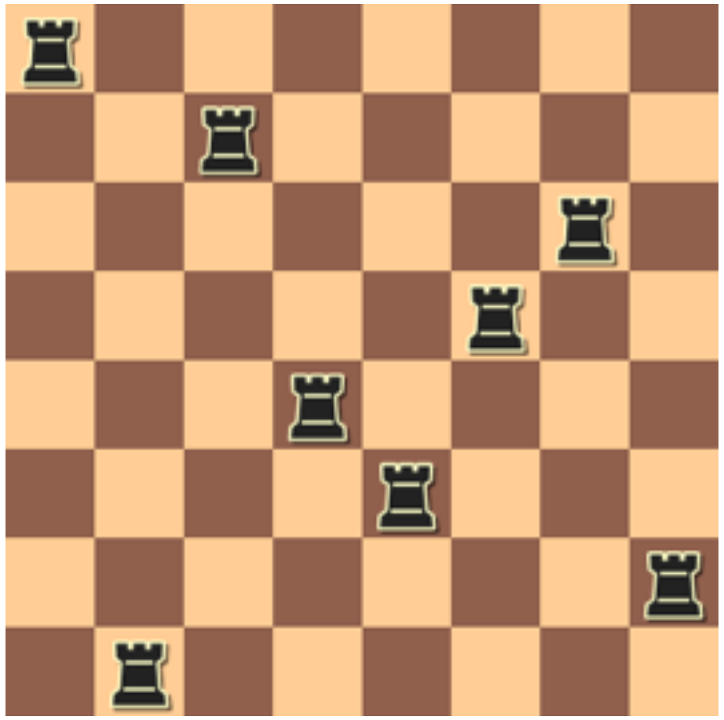
How do we sample the space efficiently and robustly?

Play multi-dimensional chess!

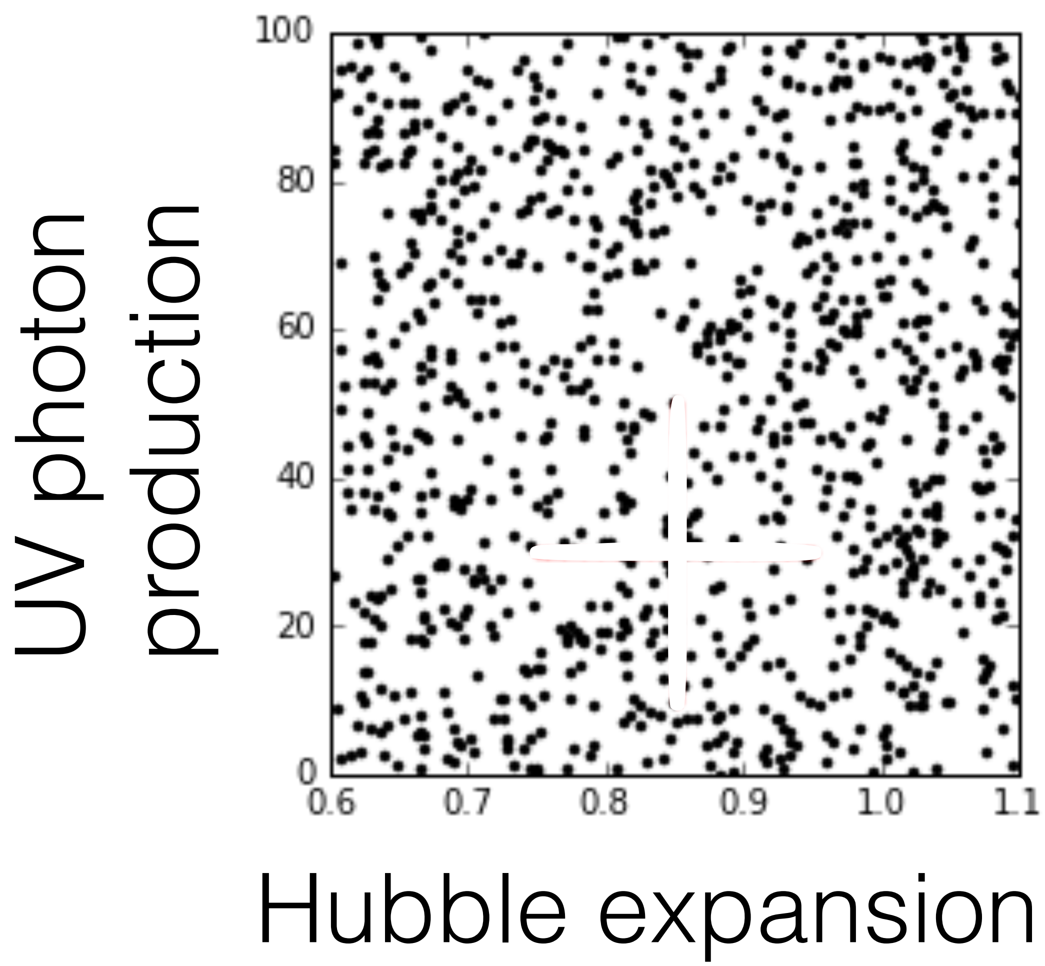


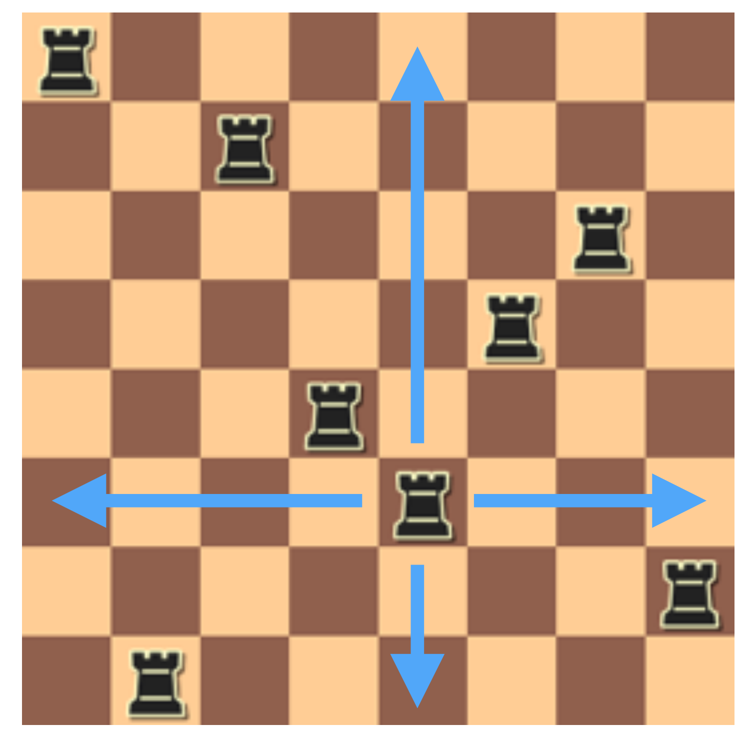
Play multi-dimensional chess!



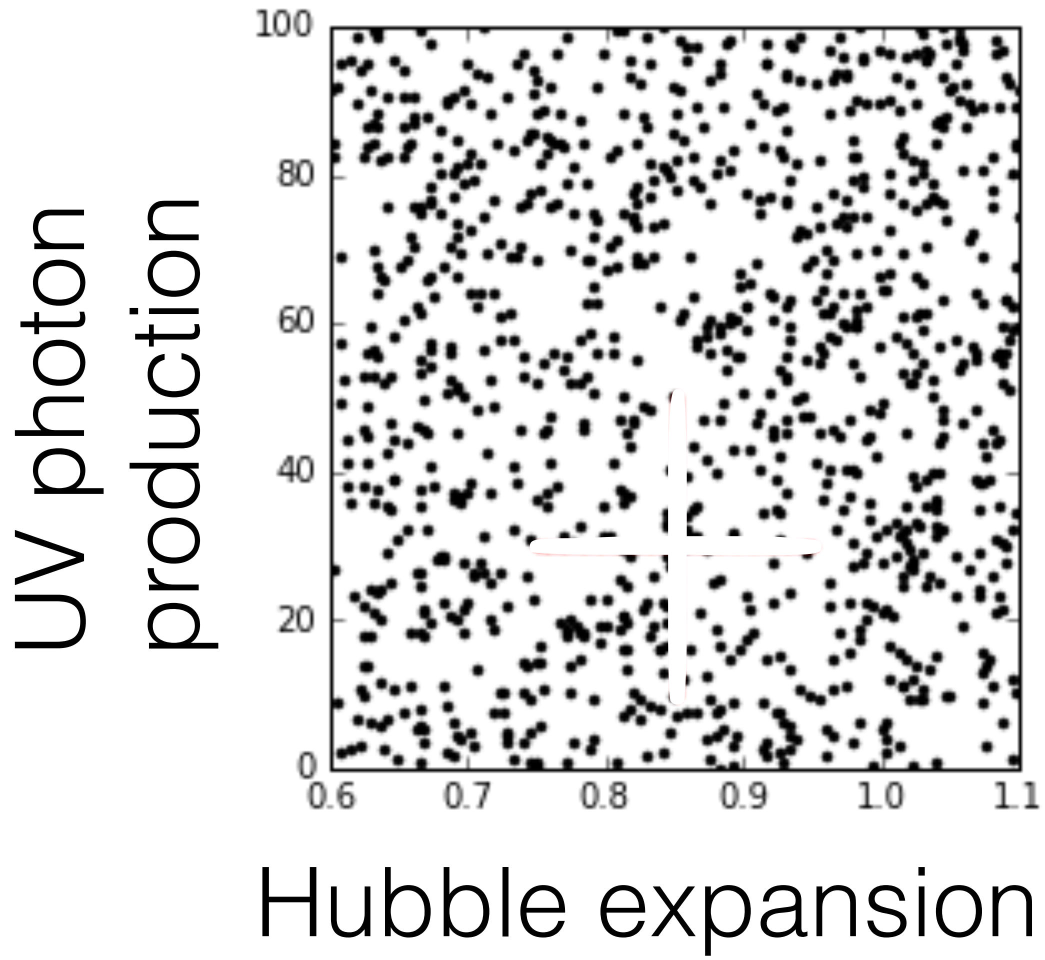


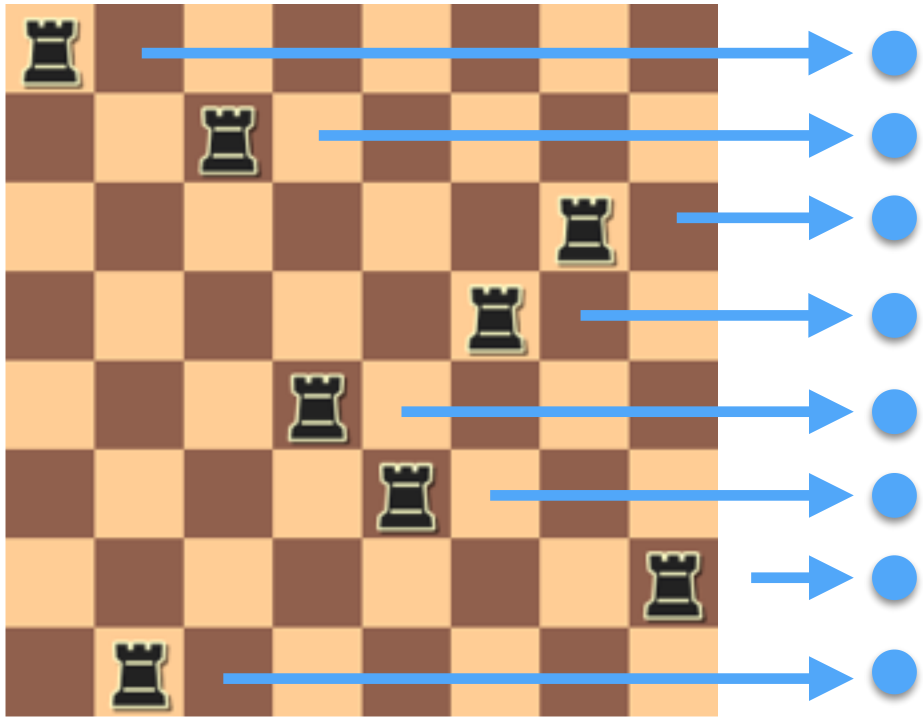
Play multi-dimensional chess!



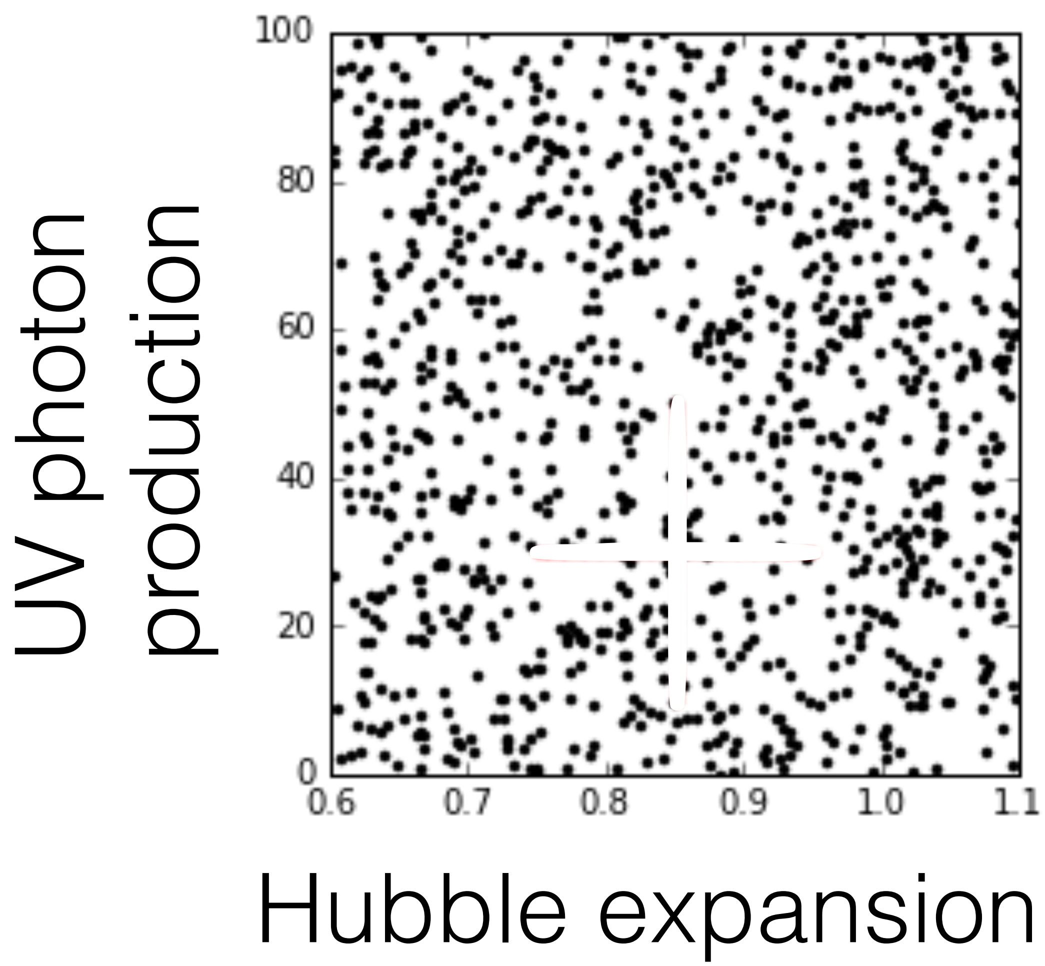


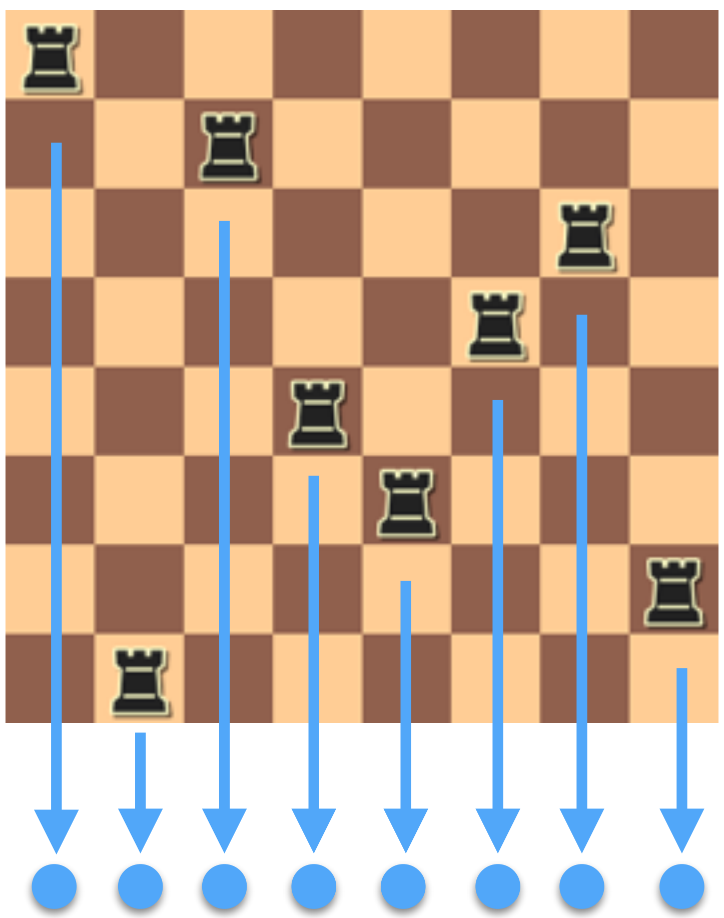
Play multi-dimensional chess!





Play multi-dimensional chess!



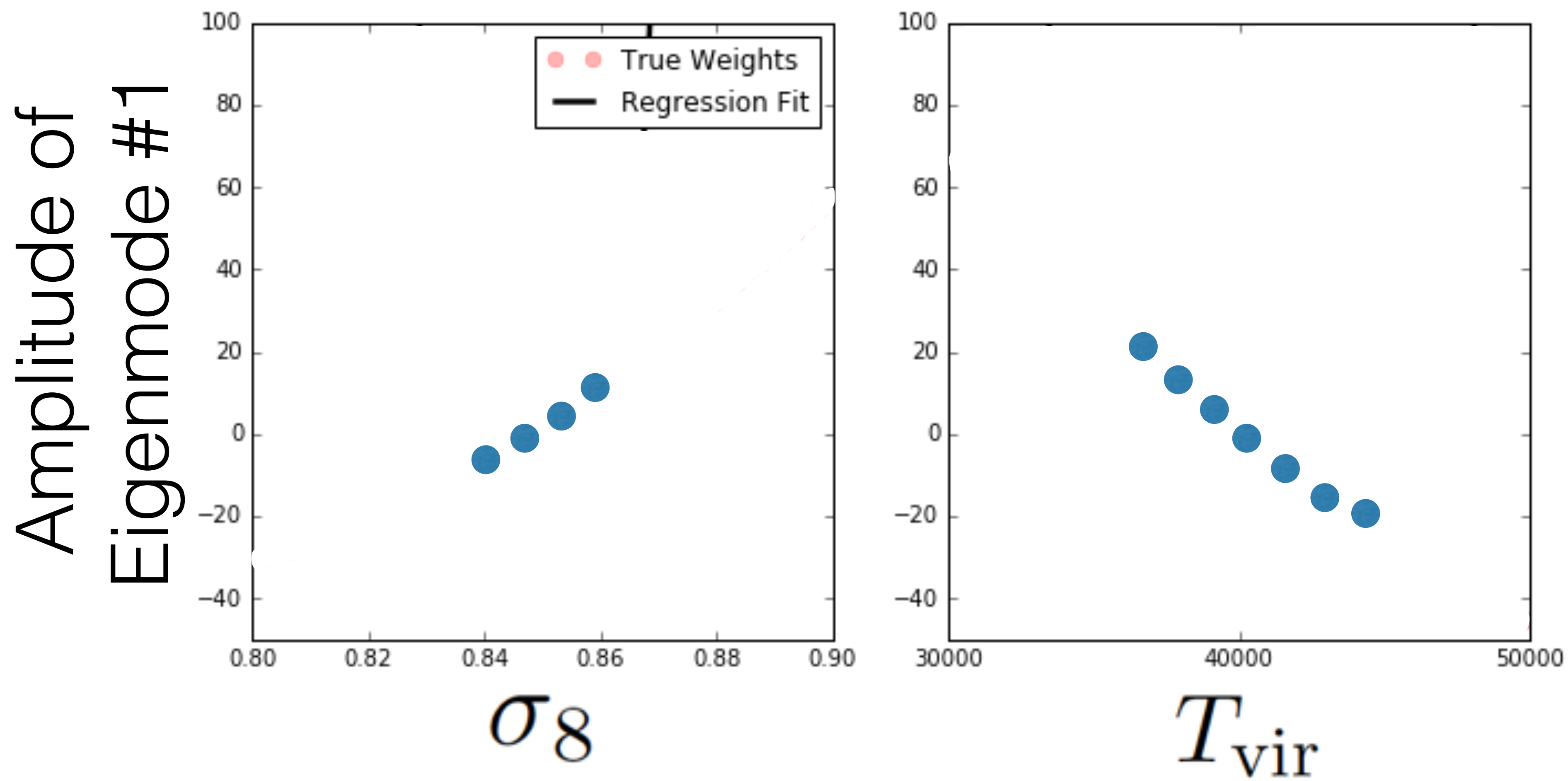


How should we perform the fit?

How do we do the fit?

- Higher order polynomial

How do we do the fit?

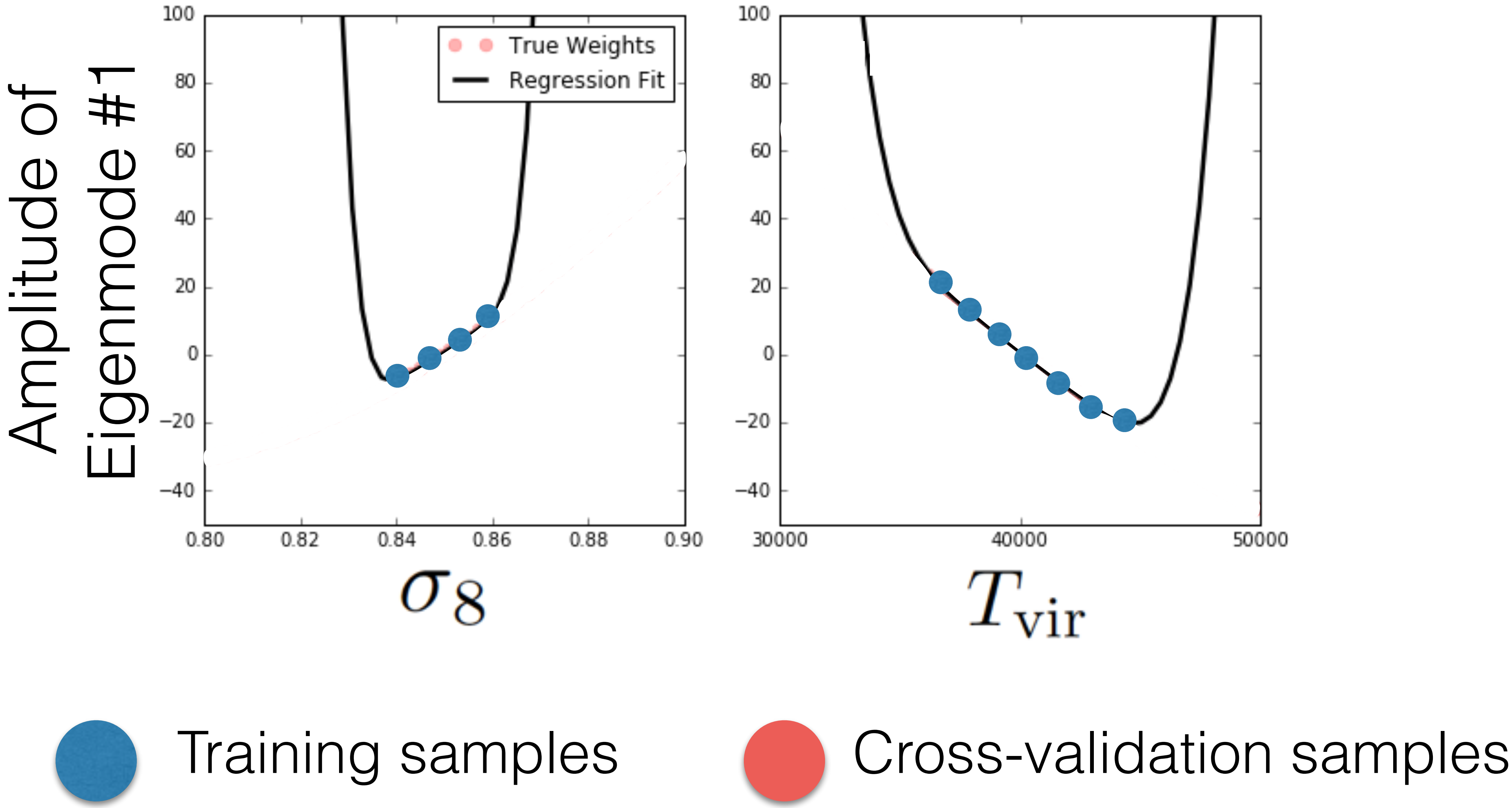


Training samples

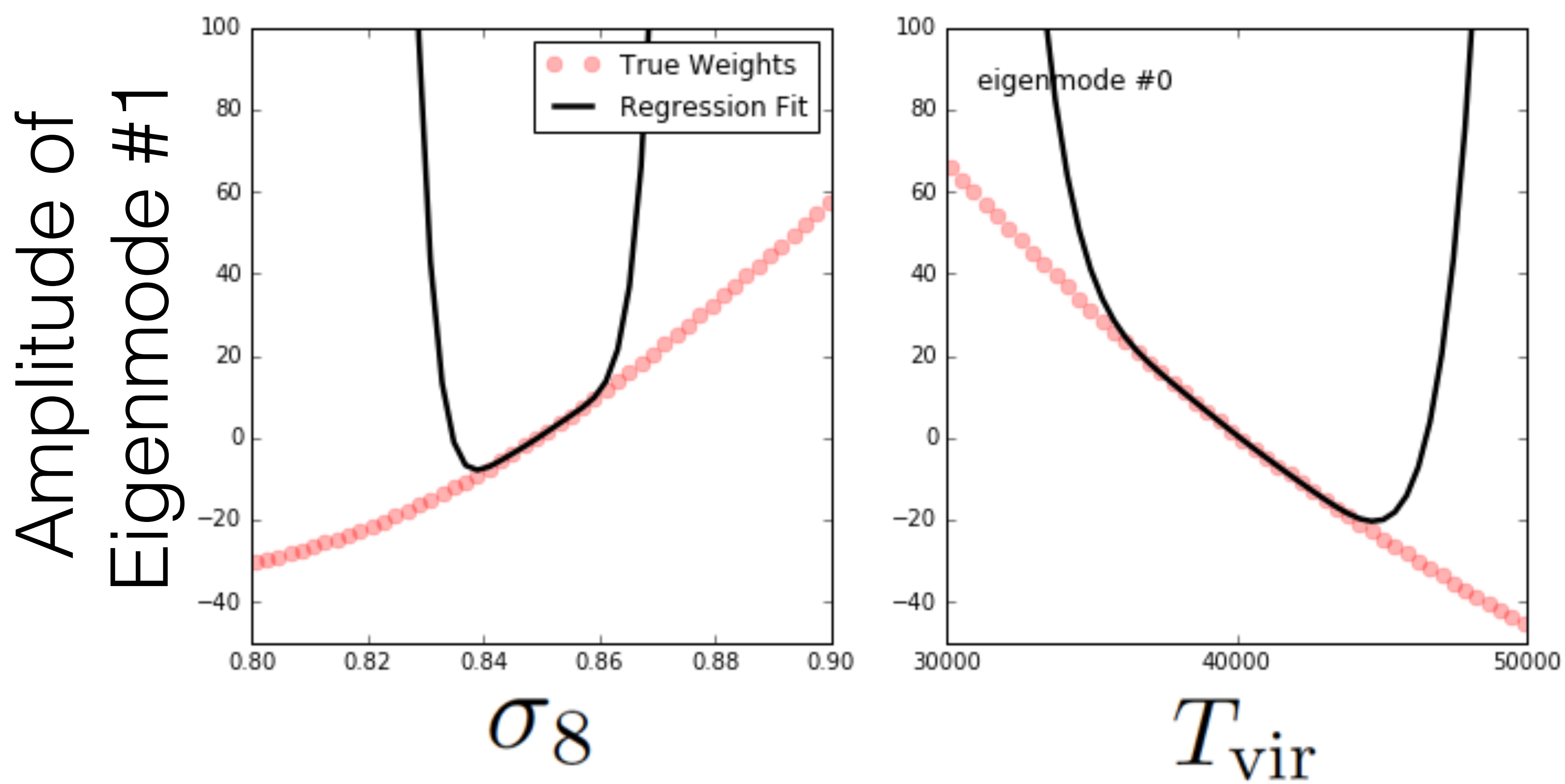
Cross-validation samples

Regression fit

How do we do the fit?



How do we do the fit?



Training samples

Cross-validation samples

Regression fit

How do we do the fit?

- Higher order polynomial
- Gaussian Process fitting

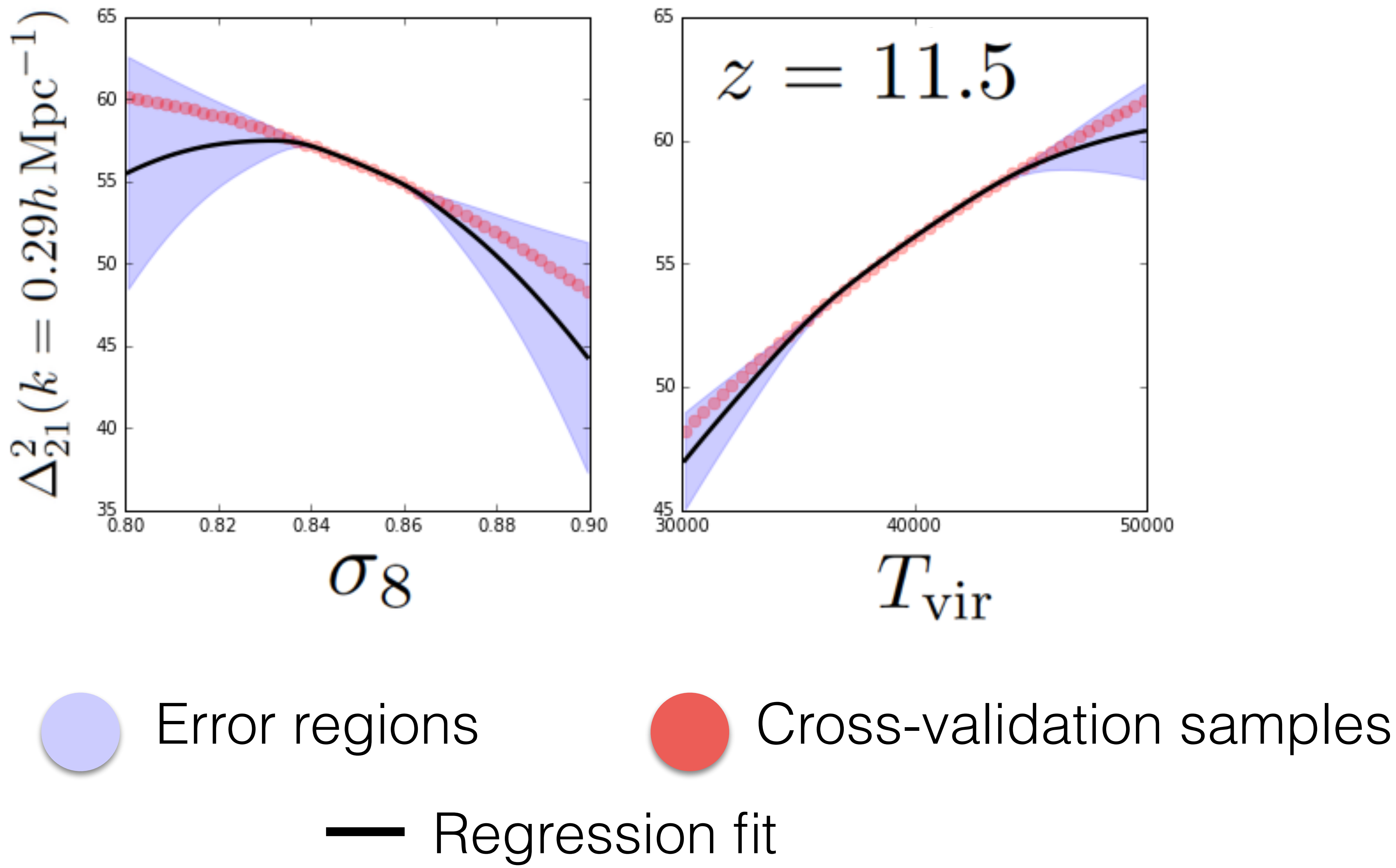
How do we do the fit?

- Higher order polynomial
- Gaussian Process fitting: model every point along the curve as being drawn from an infinite-dimensional Gaussian

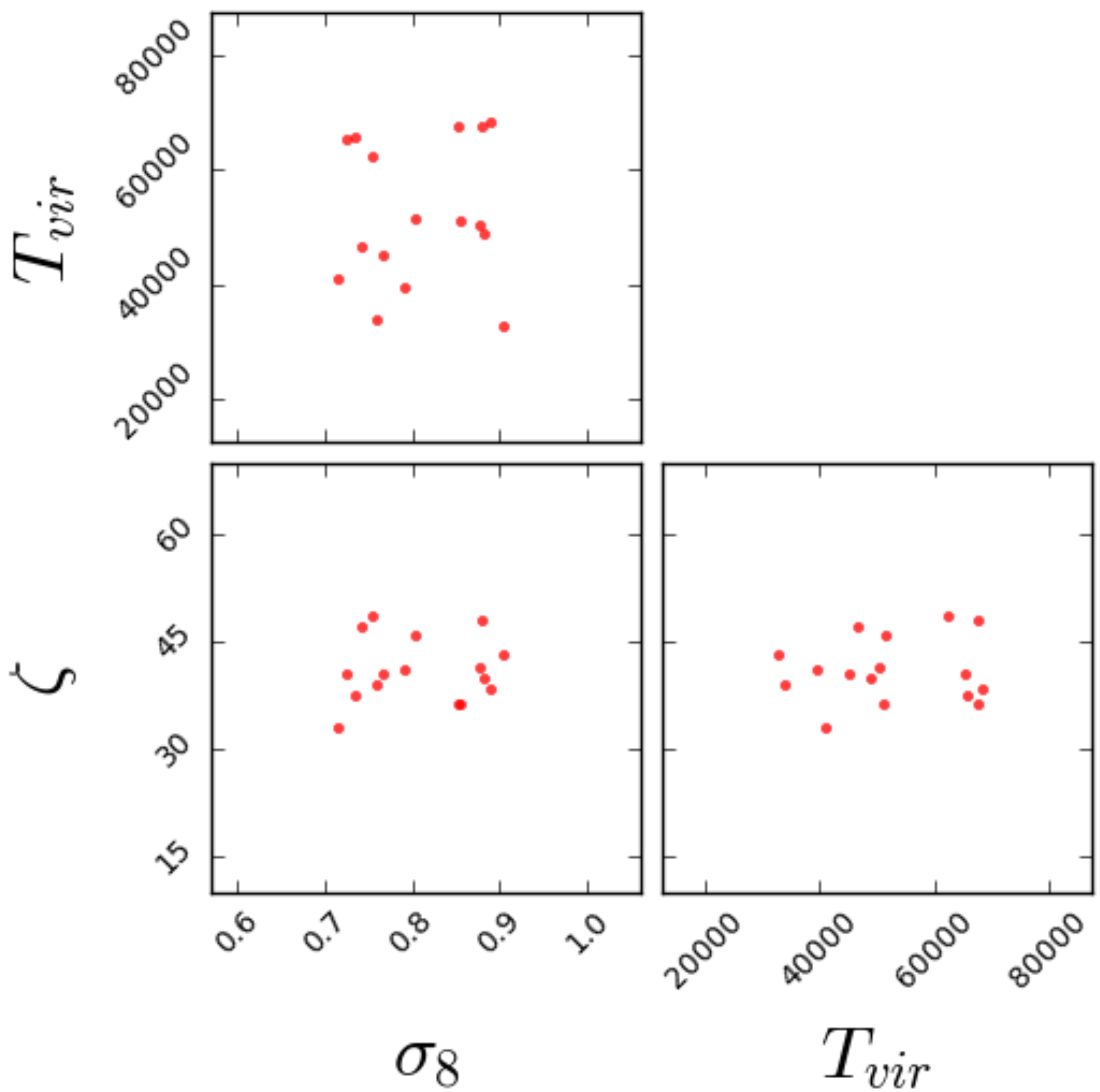
How do we do the fit?

- Higher order polynomial
- Gaussian Process fitting: model every point along the curve as being drawn from an infinite-dimensional Gaussian…with (optionally) a covariance trained from the data.

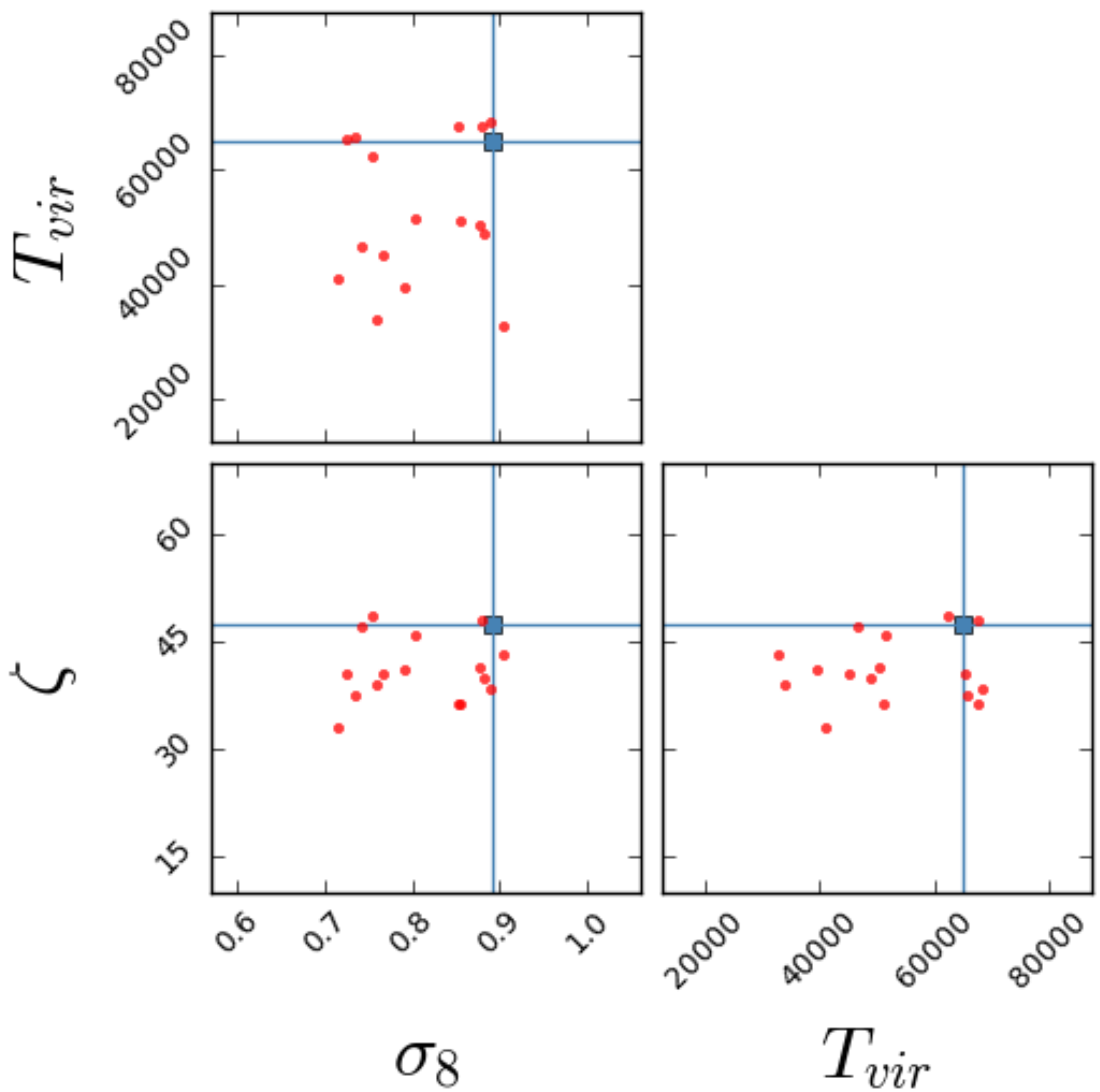
Power spectrum recovery with Gaussian Processes



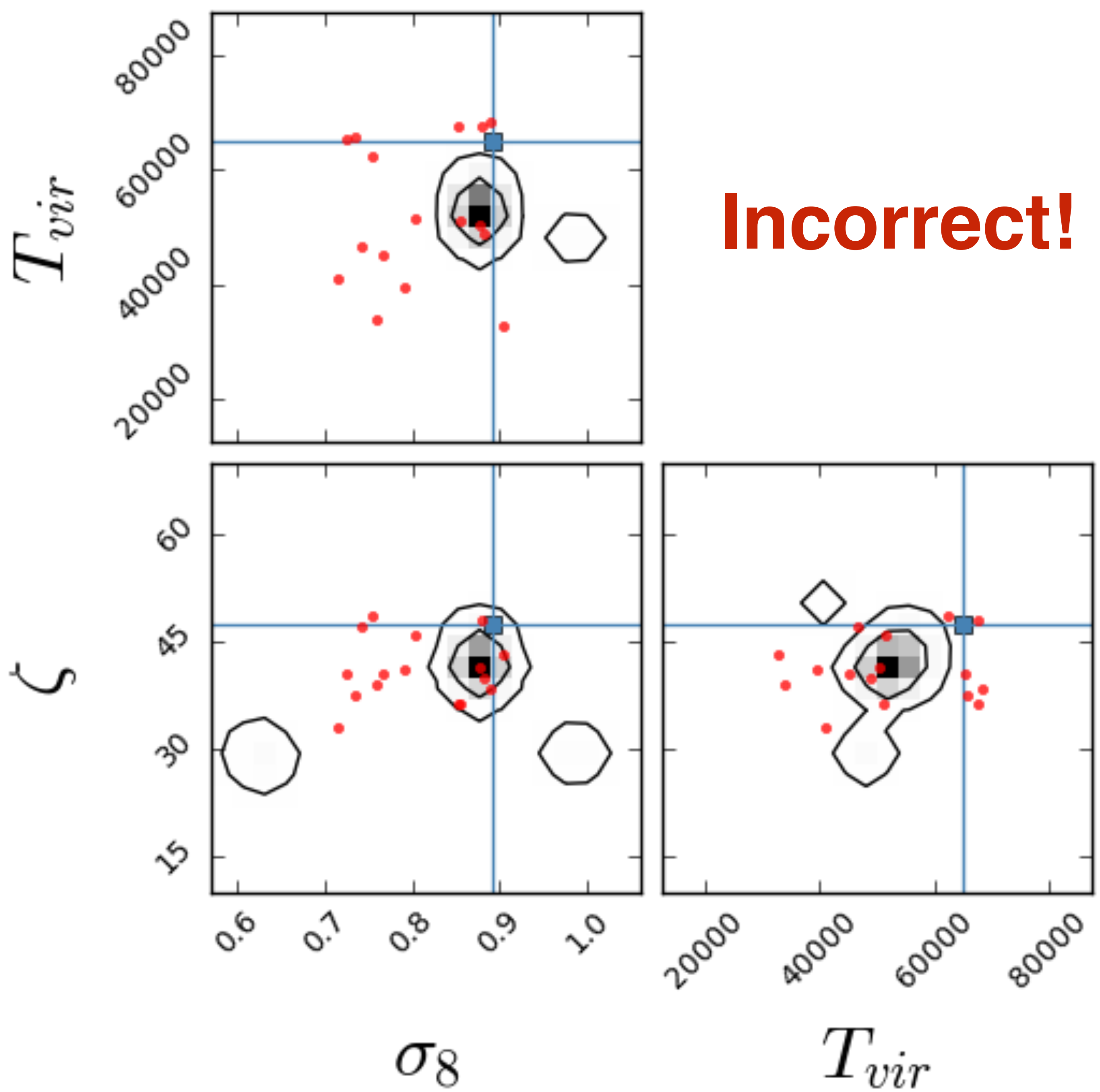
Error bars now take into account the limitations of one's emulator, with errors naturally inflating away from training region



Error bars now take into account the limitations of one's emulator, with errors naturally inflating away from training region

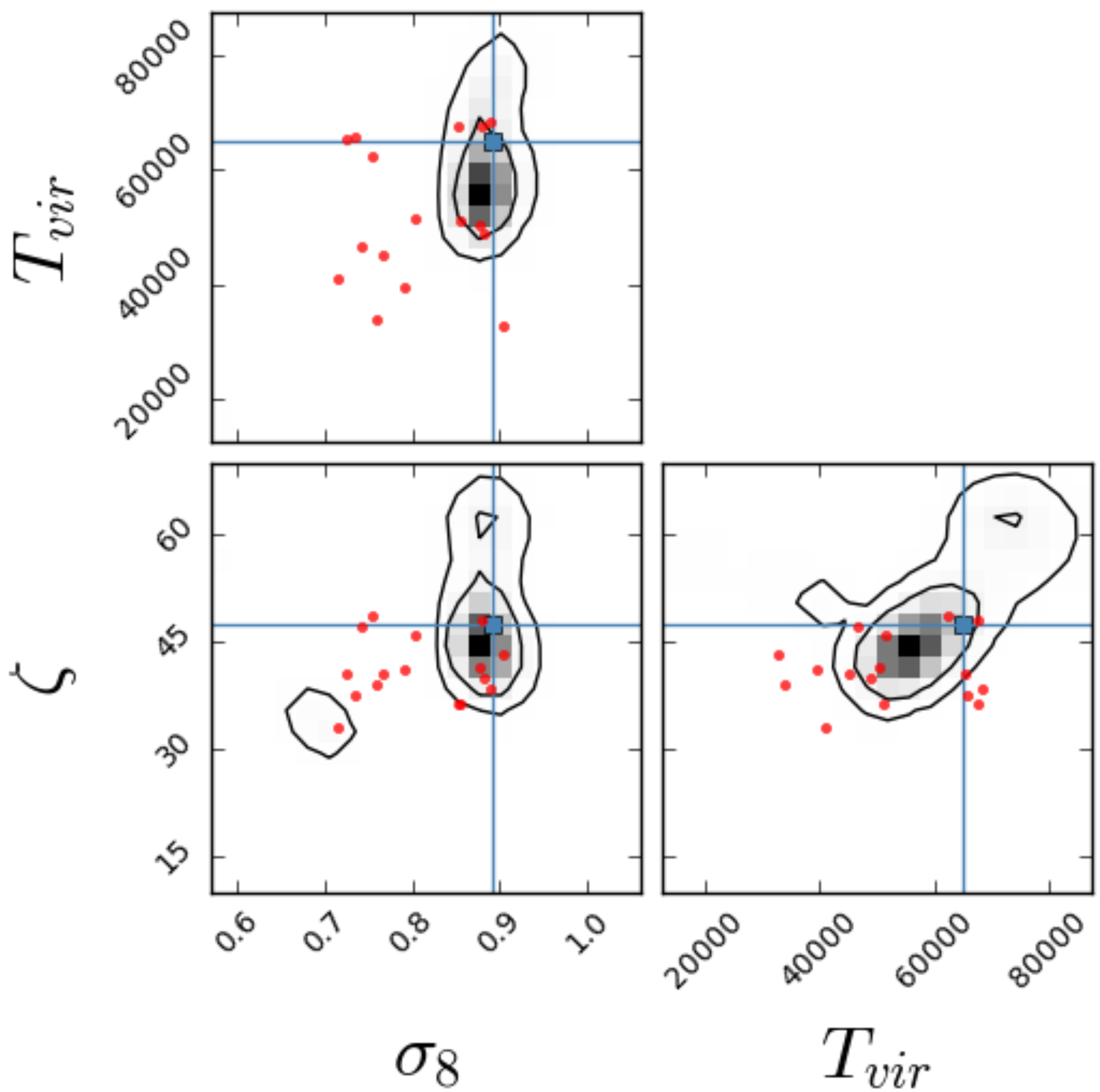


Error bars now take into account the limitations of one's emulator, with errors naturally inflating away from training region

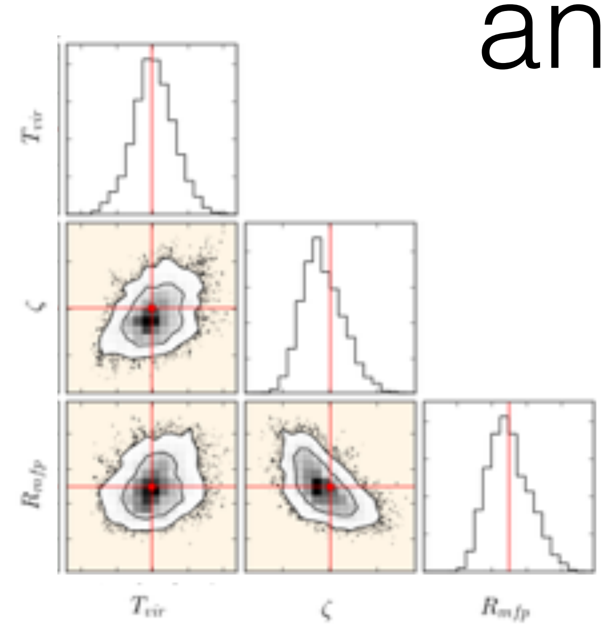


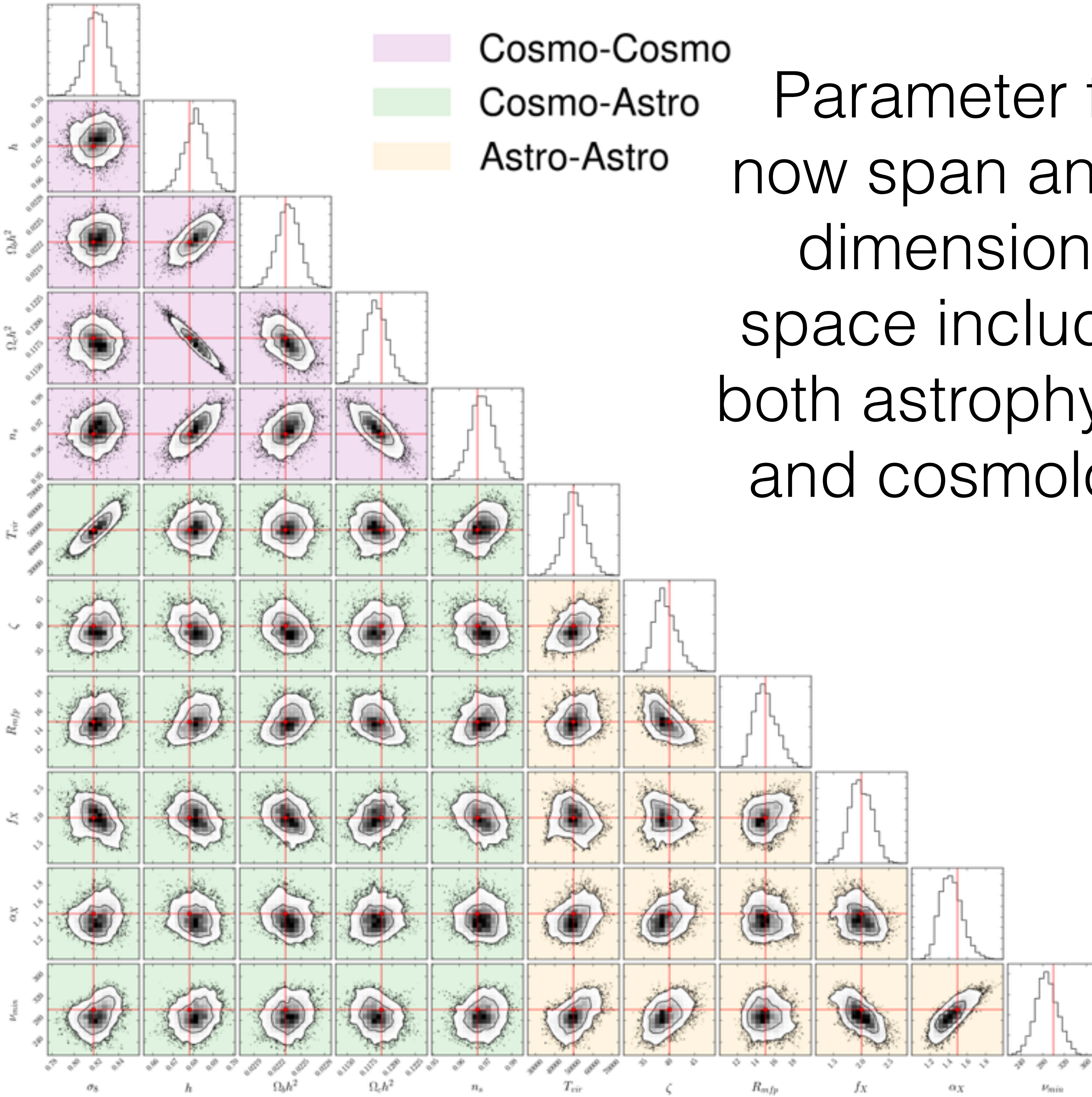
Incorrect!

Error bars now take into account the limitations of one's emulator, with errors naturally inflating away from training region



Parameter fits now span an 11-dimensional space including both astrophysics and cosmology





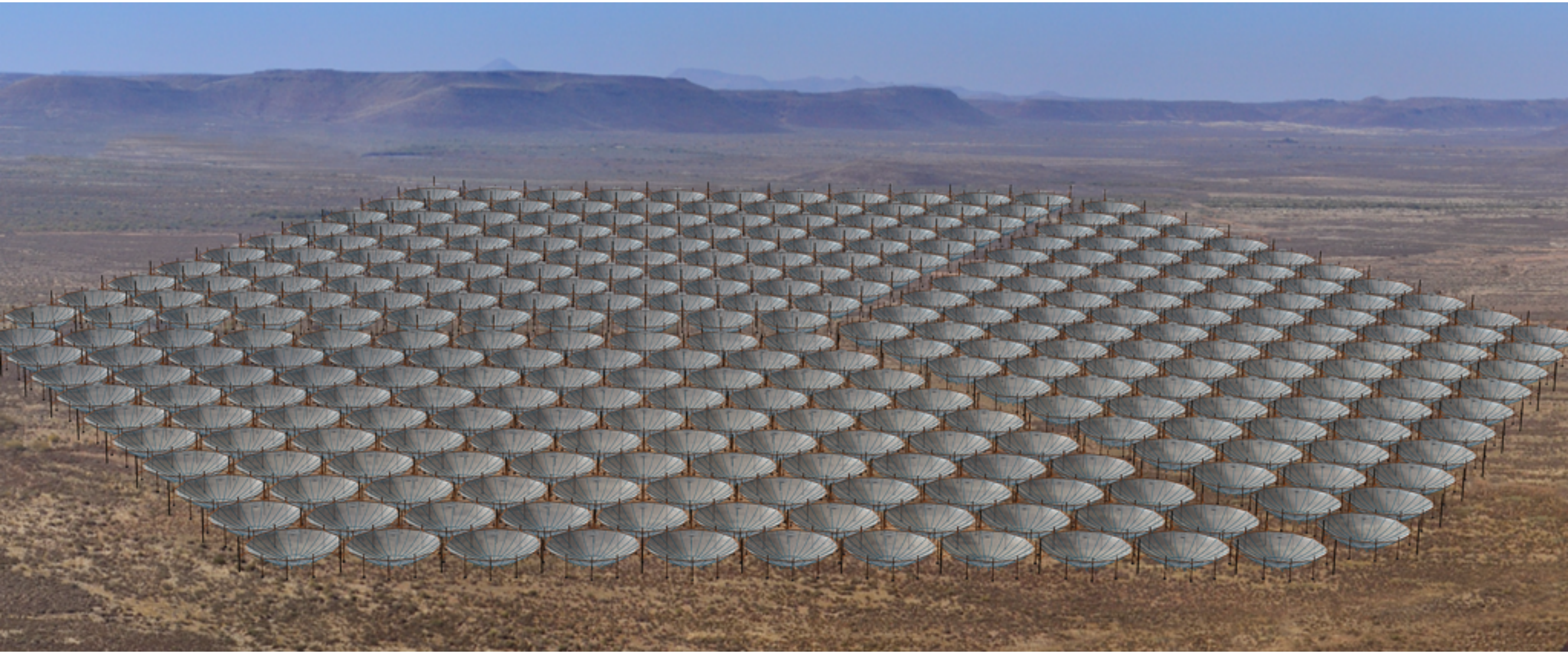
Parameter fits now span an 11-dimensional space including both astrophysics and cosmology

Emulator software coming soon!

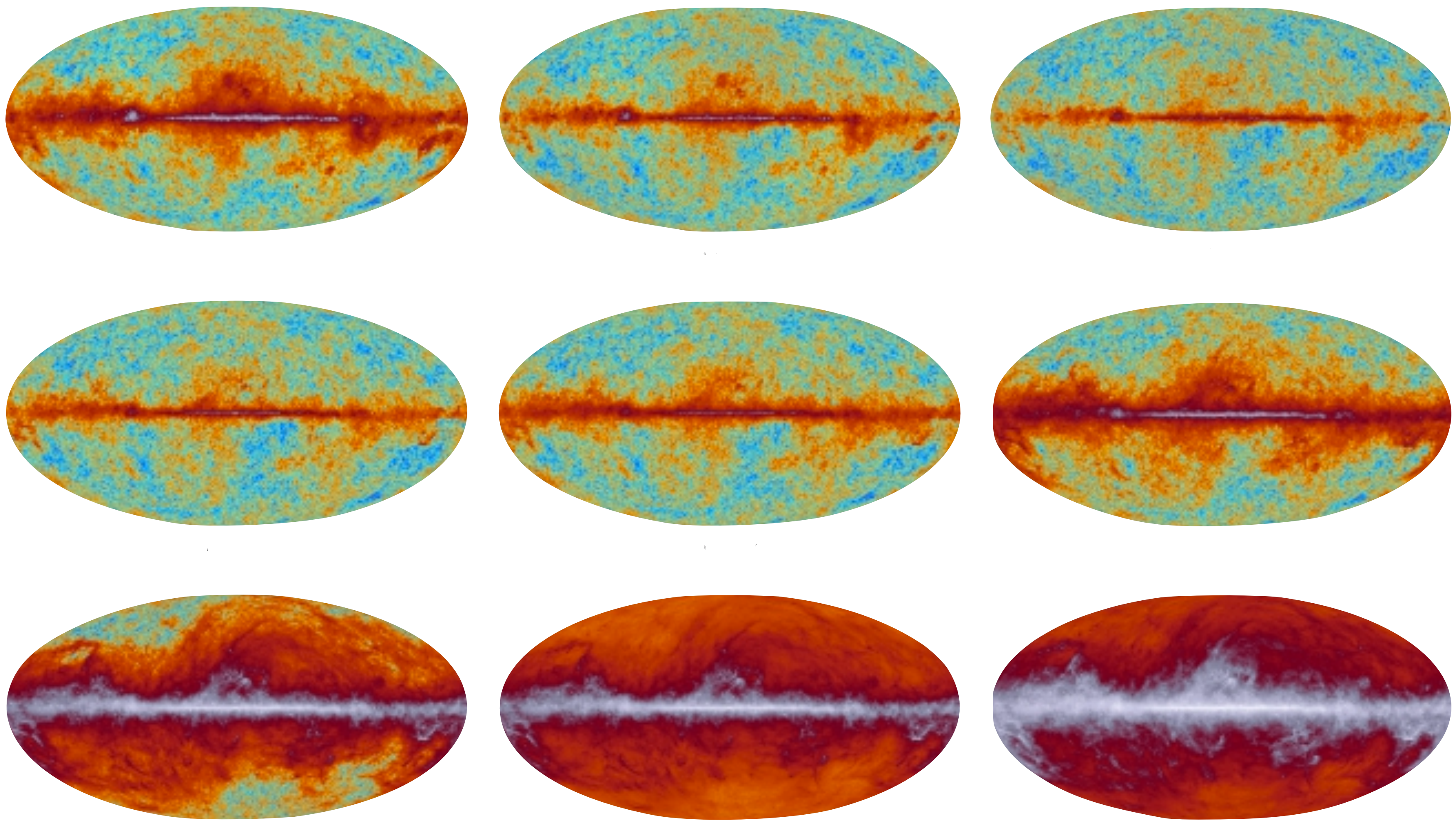
- Gaussian process regression with full error propagation
- Karhunen-Loève mode data compression
- Corner-cutting
- Latin Hypercube sampling
- Incorporation of emulator error into likelihood calculation

Kern, **AL** et al. (in prep.)

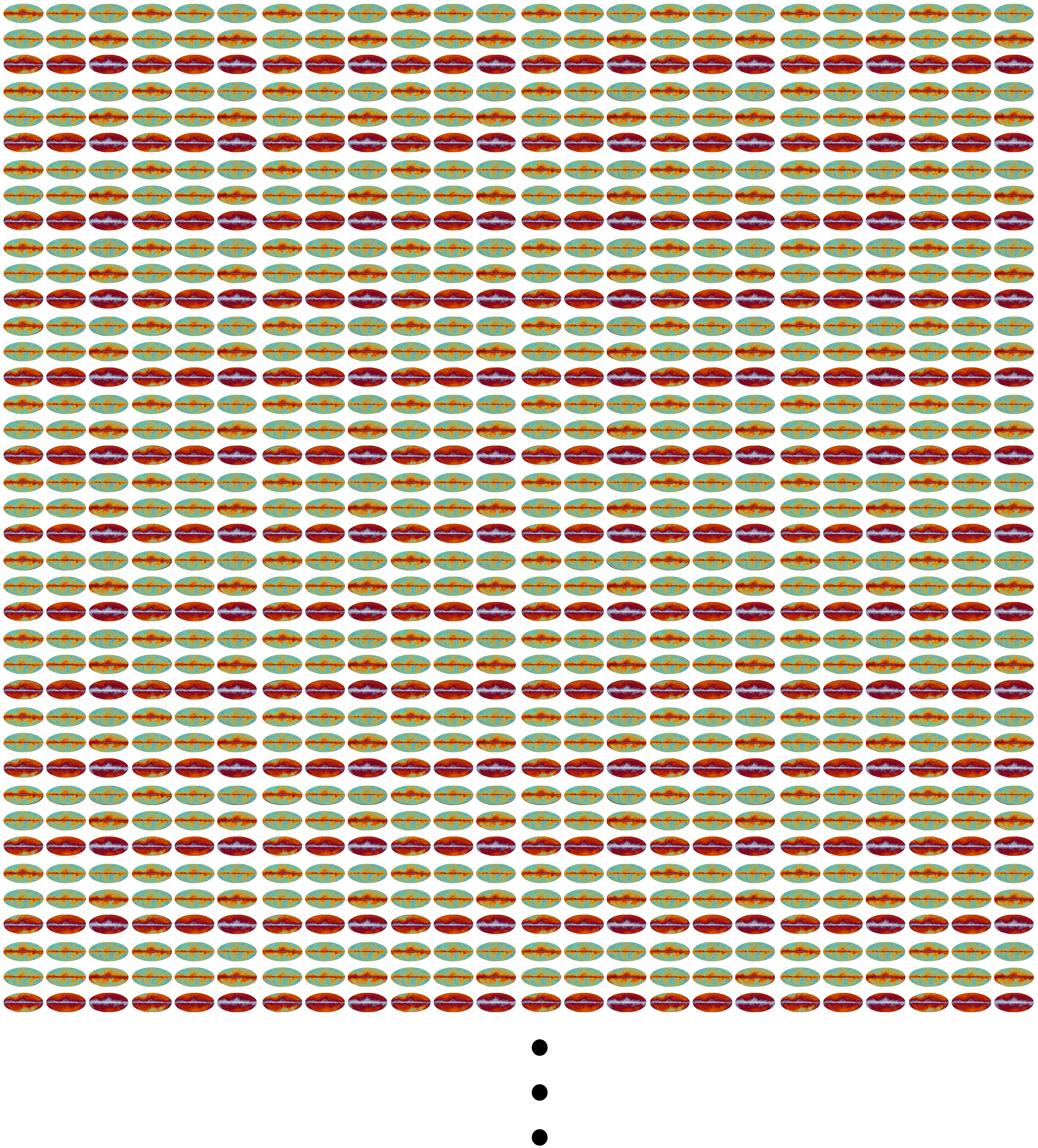
Hydrogen Epoch of Reionization Array (HERA) targeting $6 < z < 25$



~0.2 Tb (after compression) per day of observing



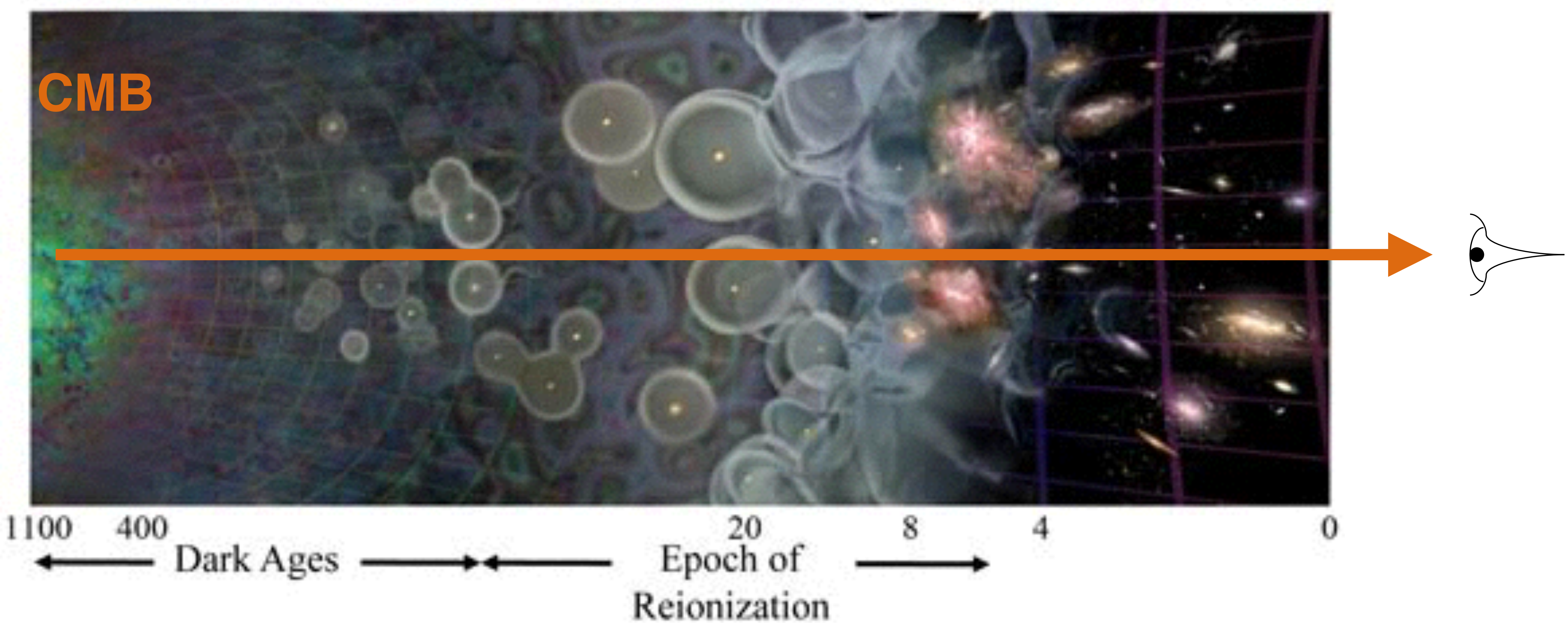
Planck satellite: maps at 9 frequencies
with ~several arcmin resolution



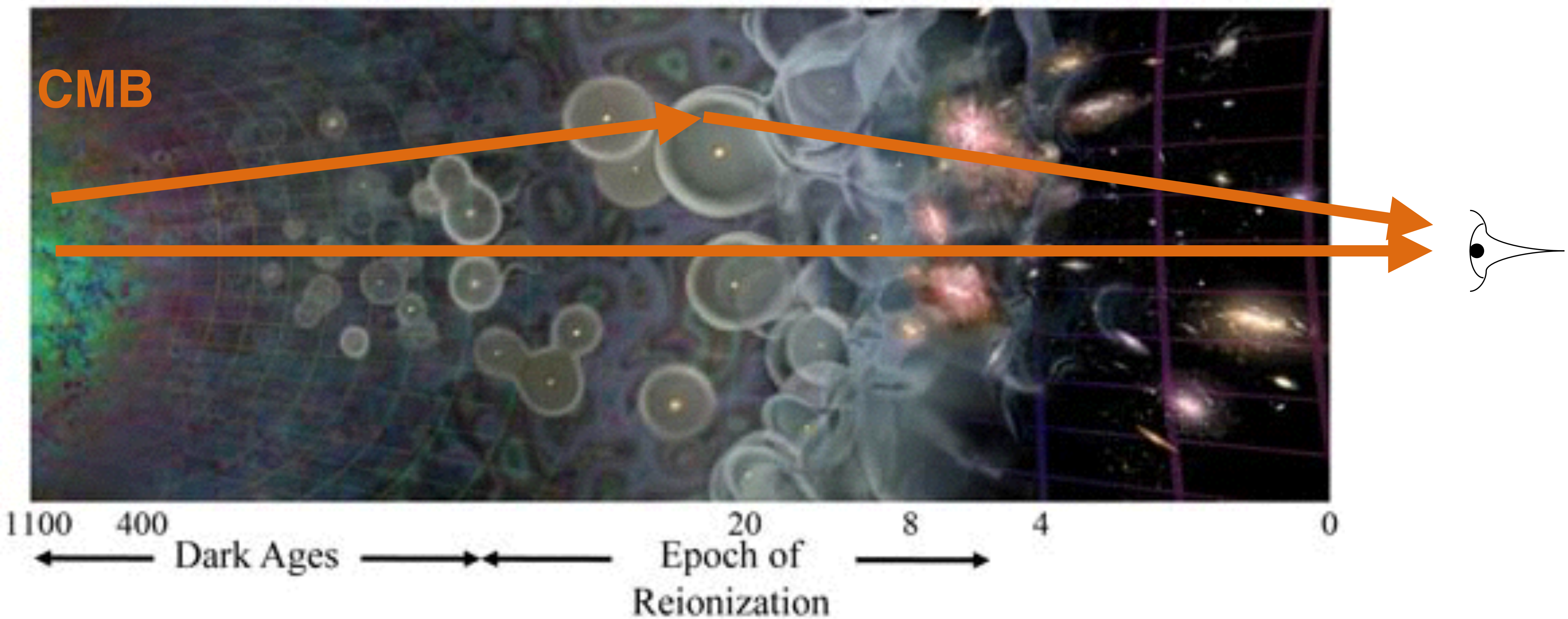
HERA: roughly the same resolution, but with ~2000 frequencies

Doing better cosmology through astrophysics

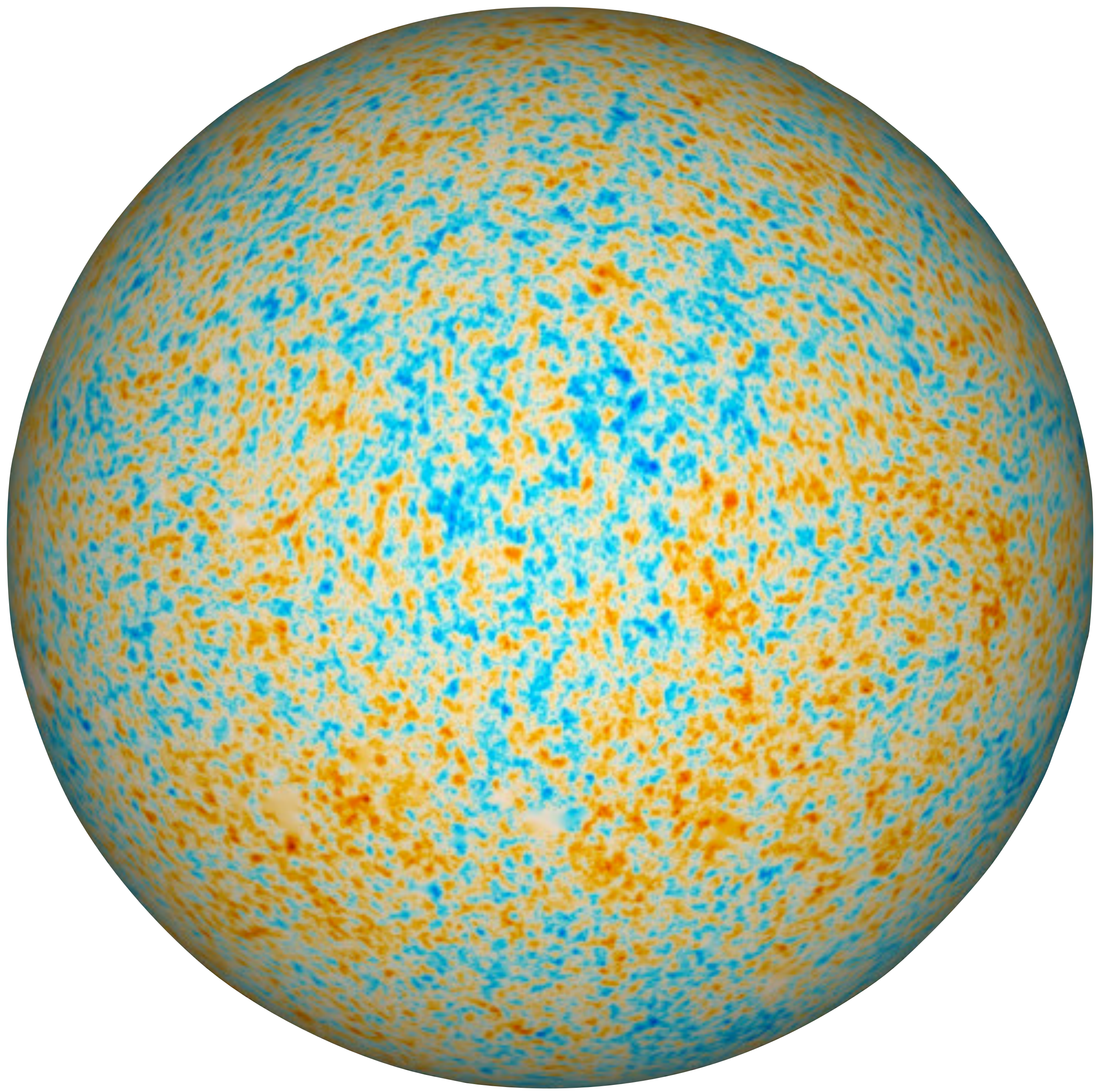
Reionization is a nuisance for CMB measurements

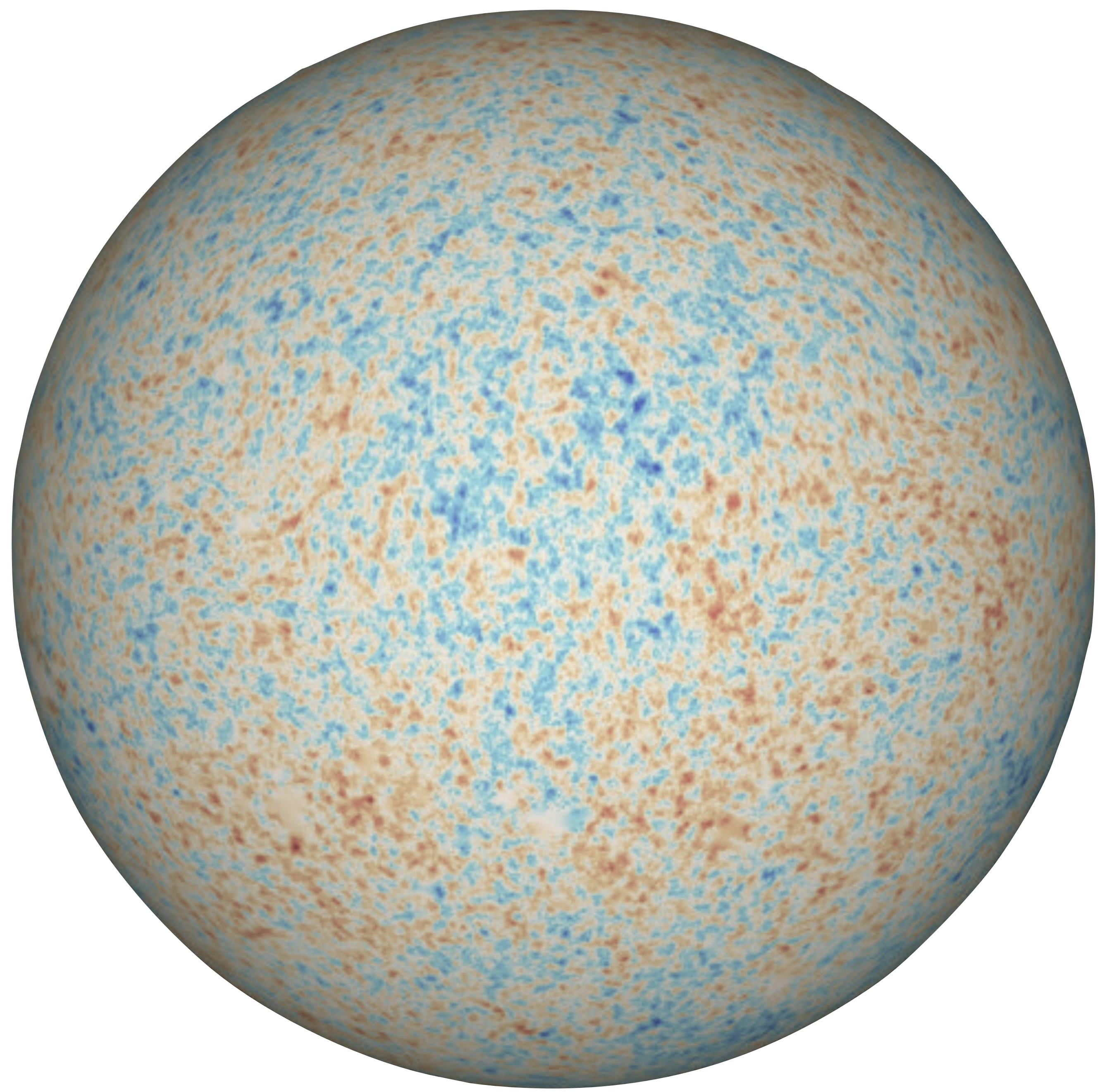


Reionization is a nuisance for CMB measurements



Extra optical depth parameter: $\tau \propto \int \langle x_i \rho_b \rangle dz$



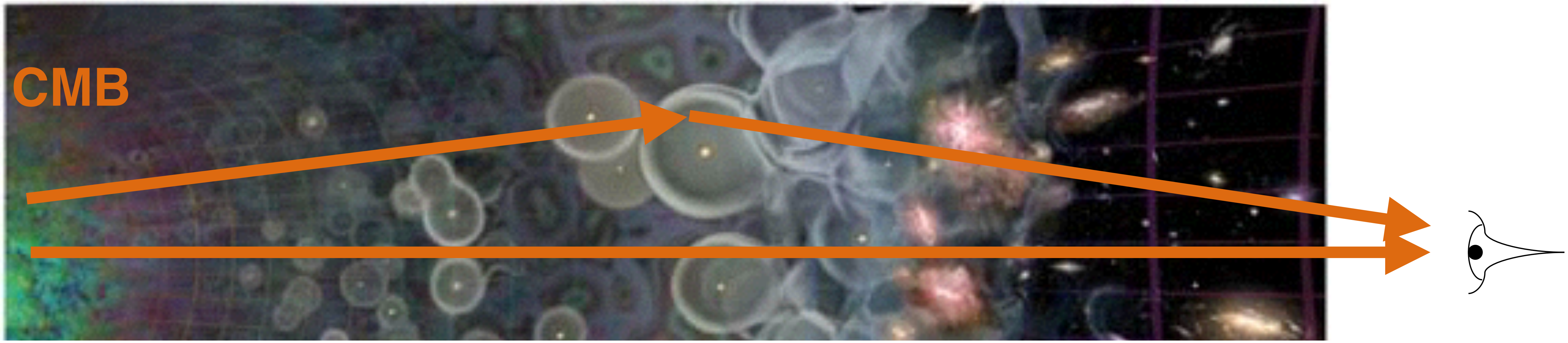


Scattering → Reduces amplitude of density fluctuations

- Early reionization (higher optical depth) + Large primordial fluctuations A_s

VS

- Late reionization (lower optical depth) + Small primordial fluctuations A_s



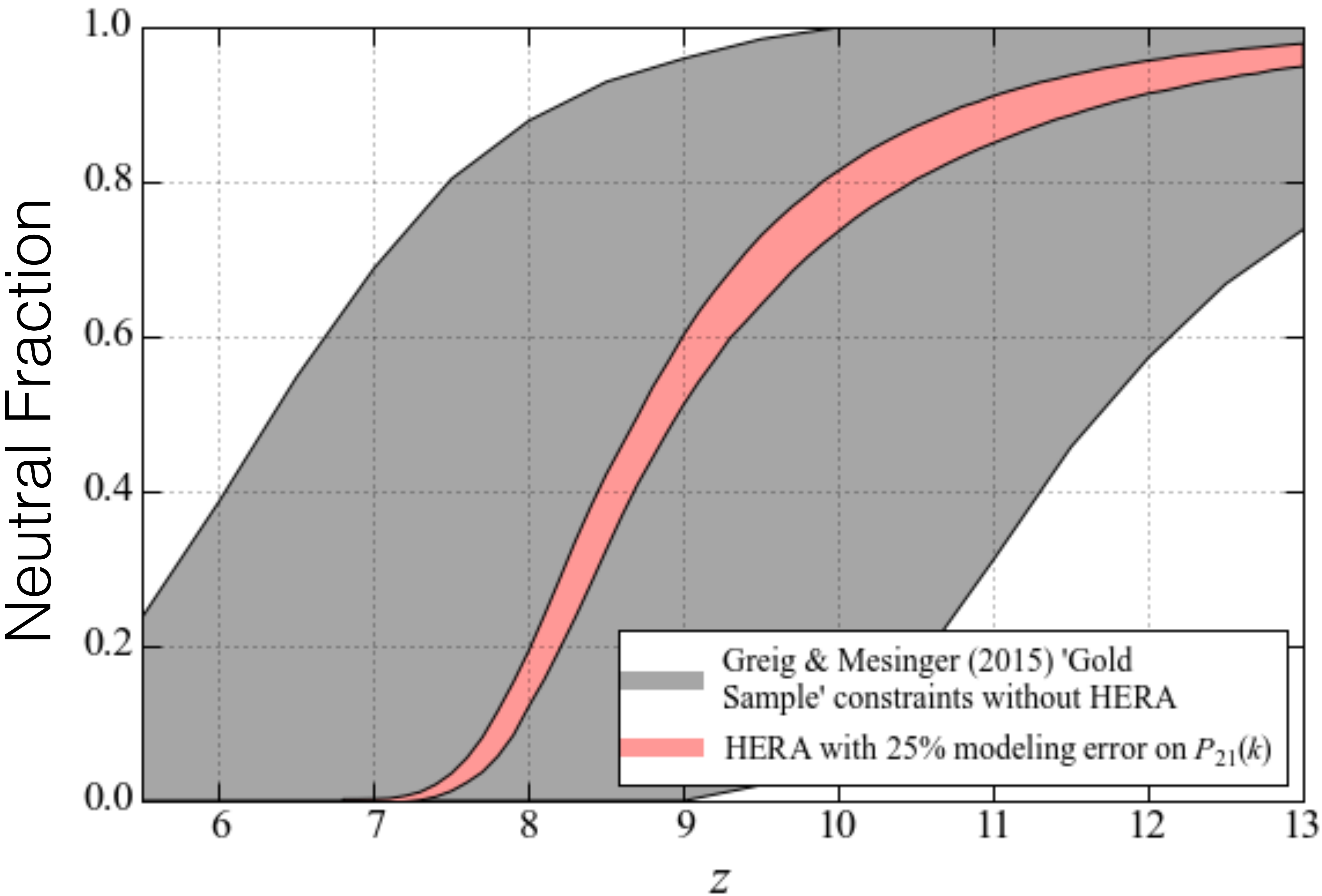
- Early reionization (higher optical depth) + Large primordial fluctuations A_s

VS

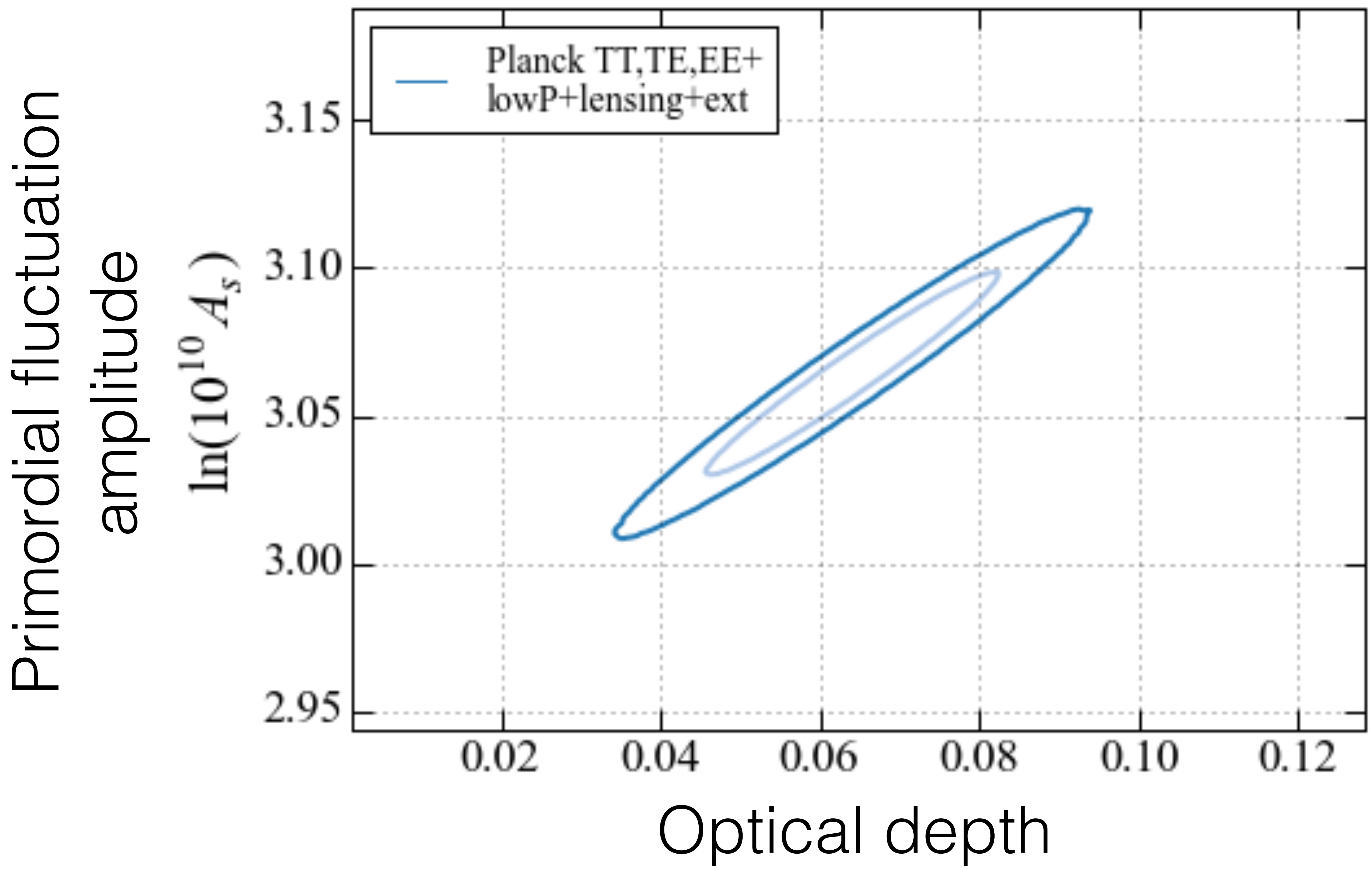
- Late reionization (lower optical depth) + Small primordial fluctuations A_s

Understanding reionization (especially the CMB optical depth) can improve constraints on other cosmological parameters

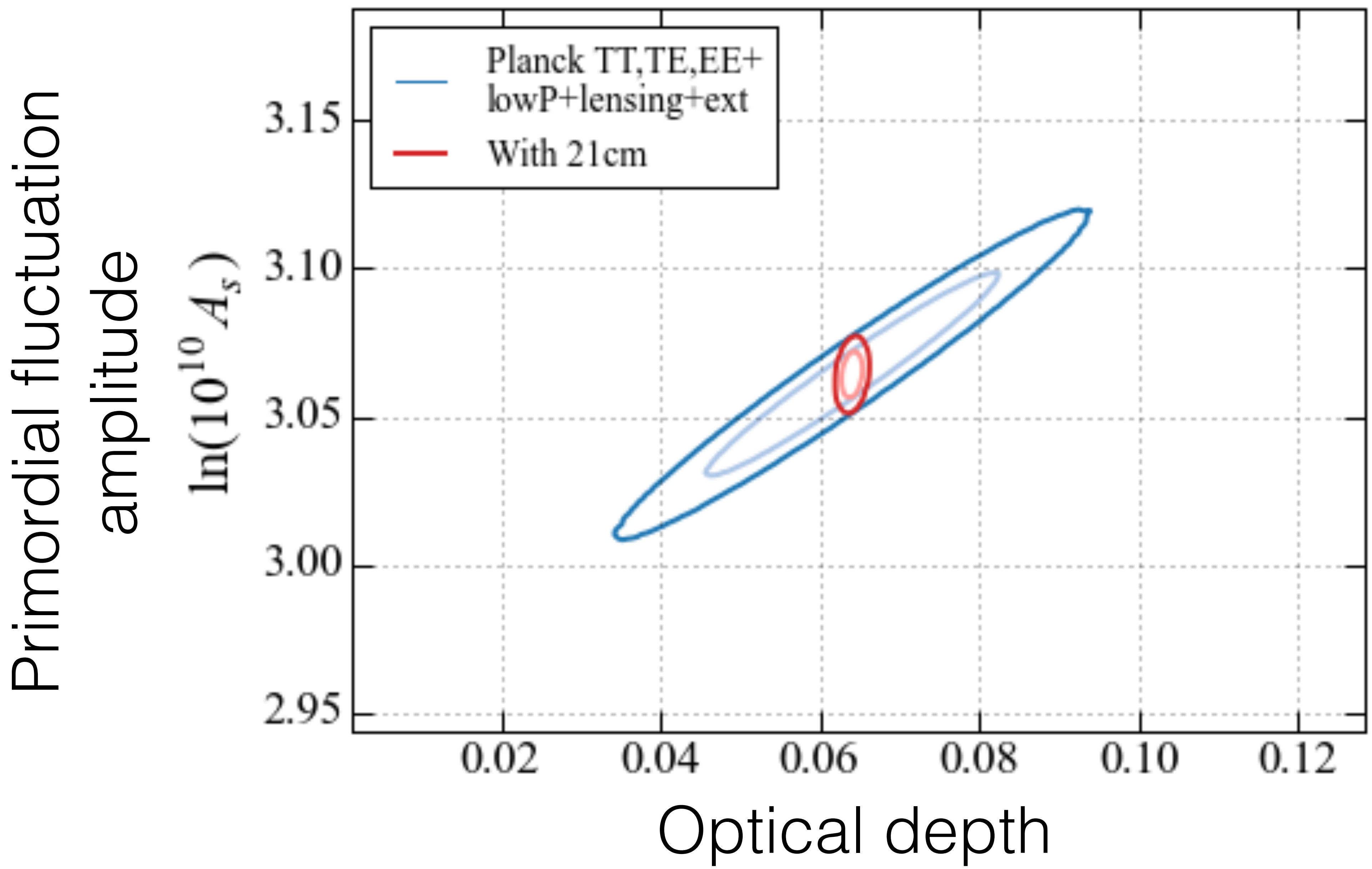
HERA provides us with exactly what we need



21cm information breaks the degeneracy between the amplitude of fluctuations and the optical depth

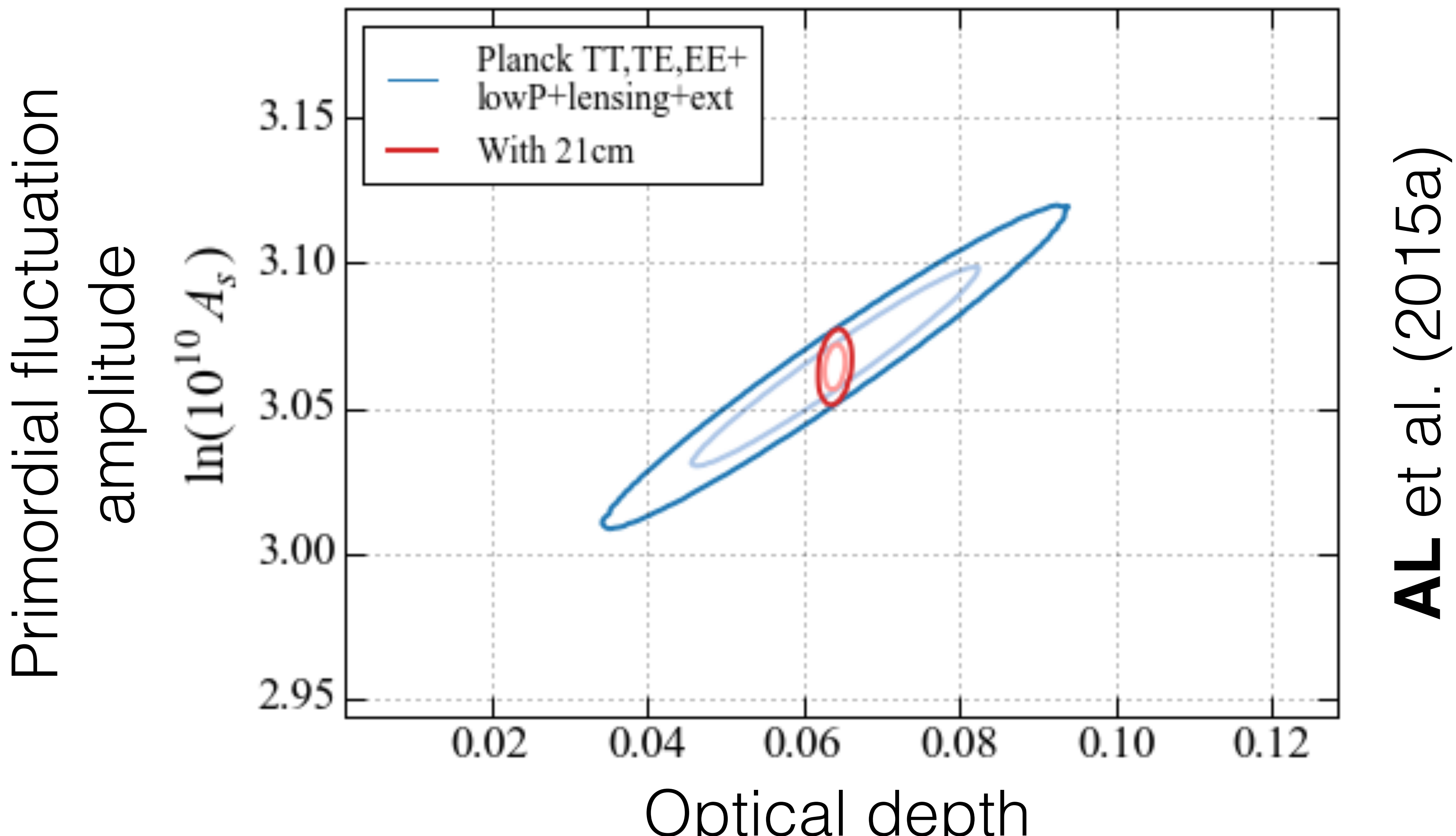


21cm information breaks the degeneracy between the amplitude of fluctuations and the optical depth



AL et al. (2015a)

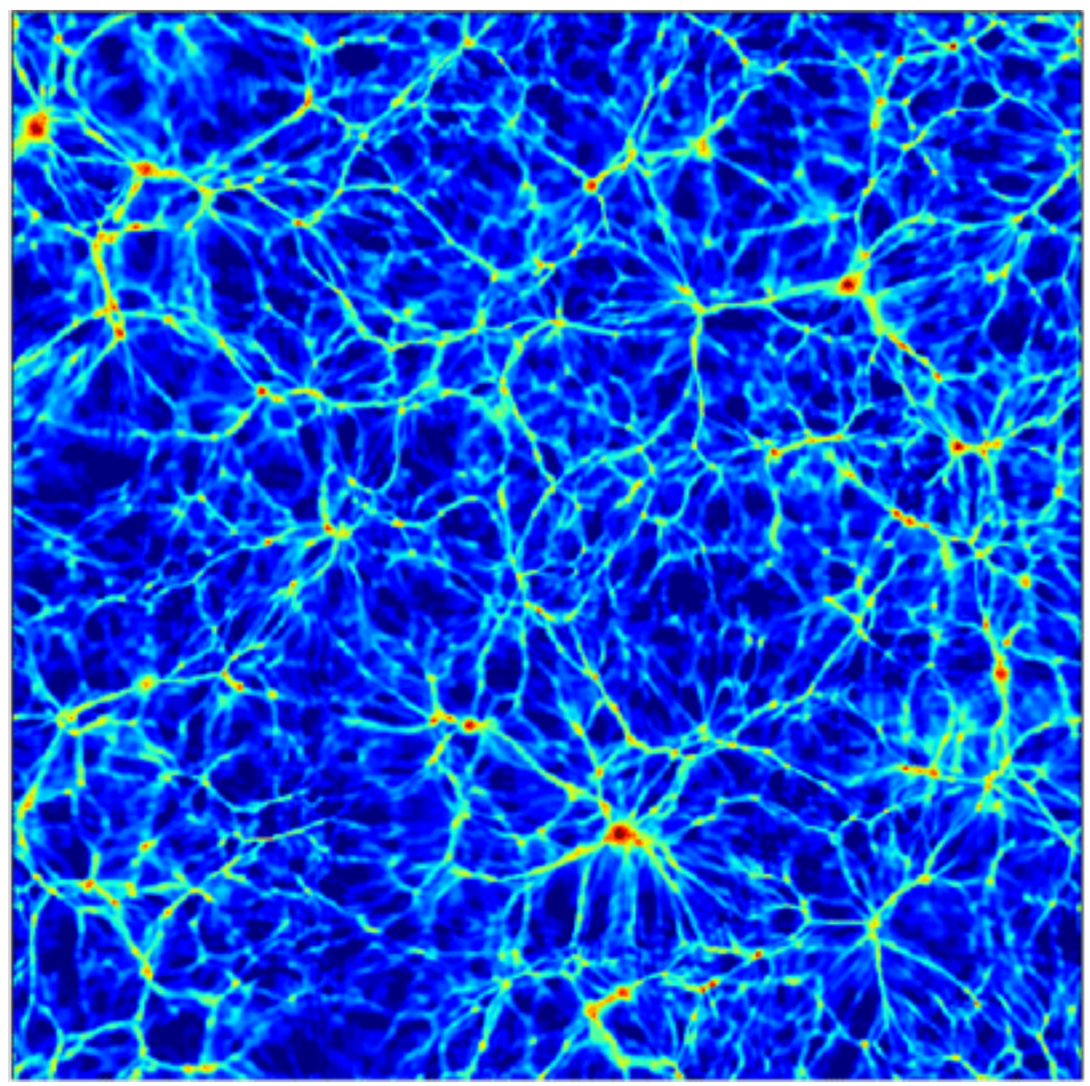
21cm information breaks the degeneracy between the amplitude of fluctuations and the optical depth



$$\Delta \ln(10^{10}A_s) = \pm 0.023 \longrightarrow \pm 0.0053$$

Futuristic cosmology experiments targeting the neutrino mass also benefit

- Neutrinos free-stream out of over-densities and dampen structure formation

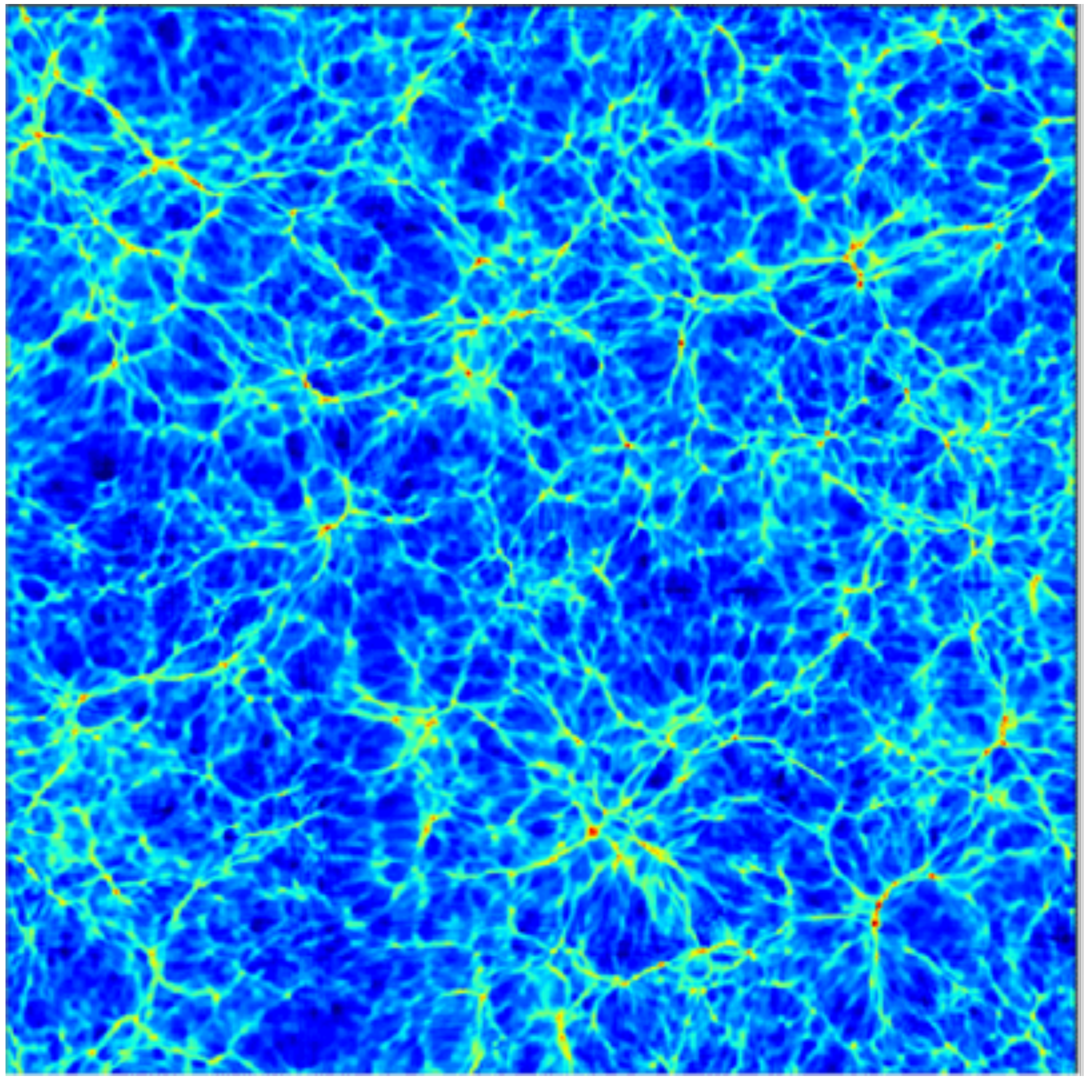


Without neutrinos

Agarwal & Feldman 2011

Futuristic cosmology experiments targeting the neutrino mass also benefit

- Neutrinos free-stream out of over-densities and dampen structure formation

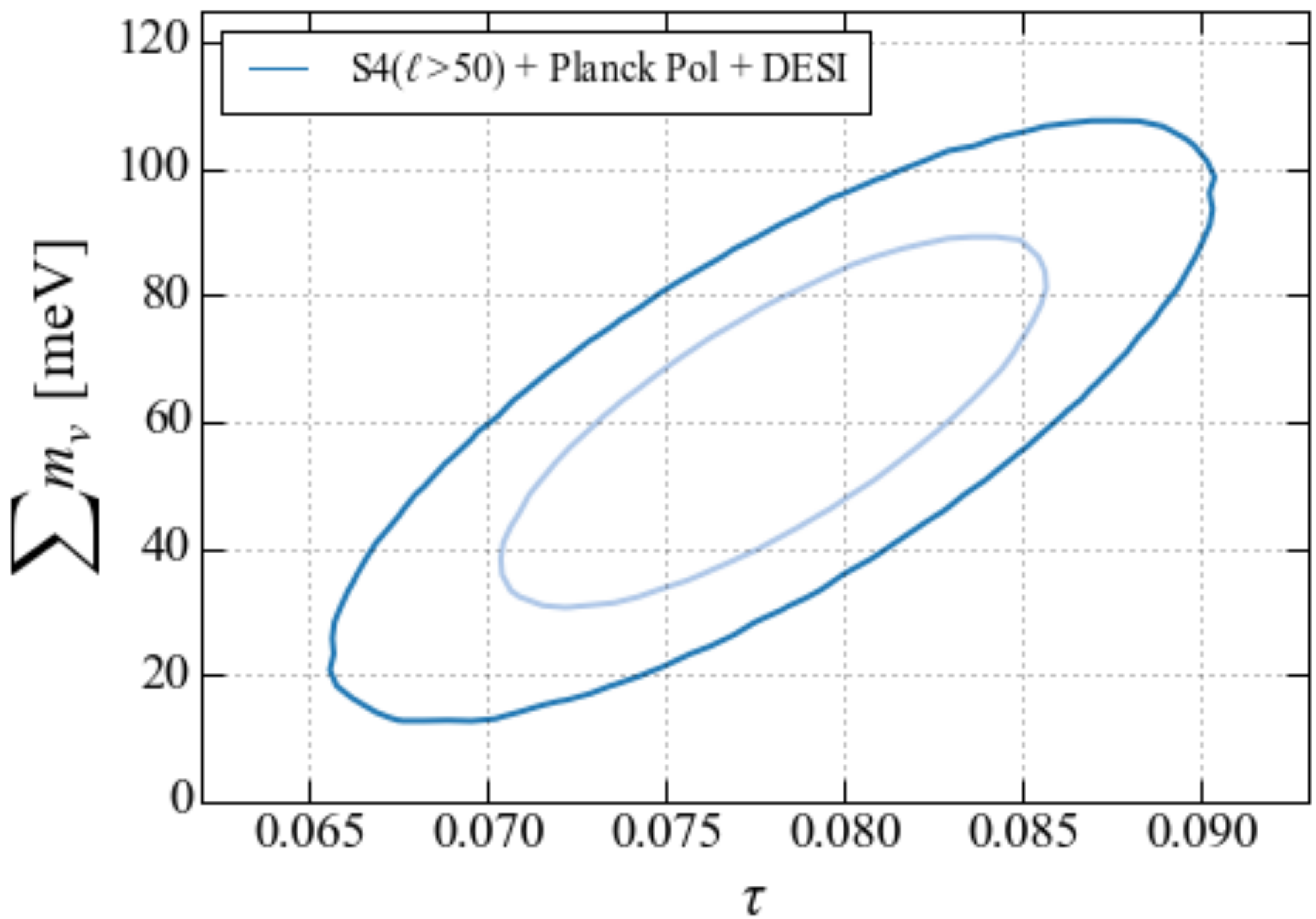


With neutrinos

Agarwal & Feldman 2011

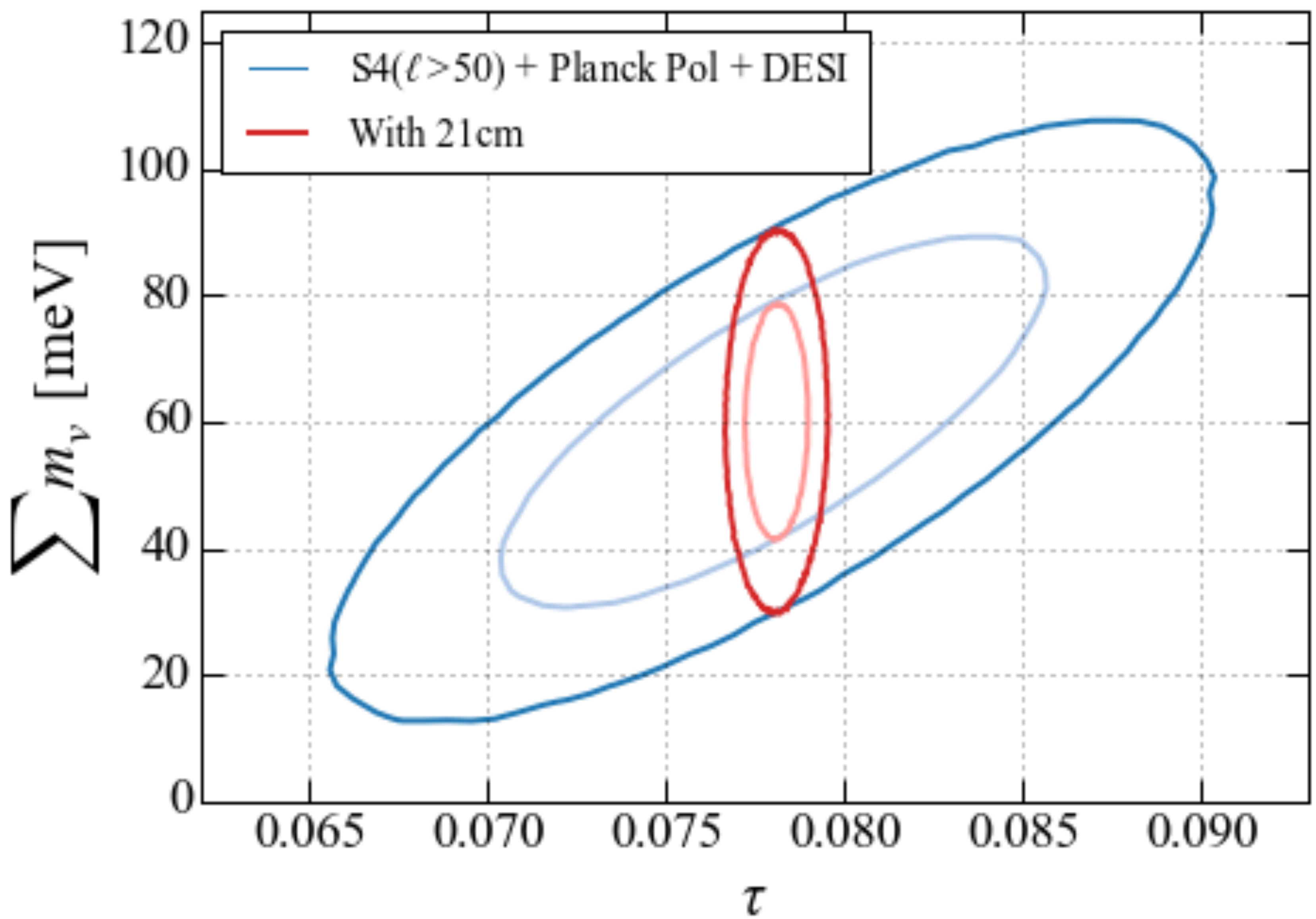
Both the neutrino mass and the optical depth can affect the observed amount of small scale structure, leading to degeneracies

Both the neutrino mass and the optical depth can affect the observed amount of small scale structure, leading to degeneracies



Allison et al. (2015)

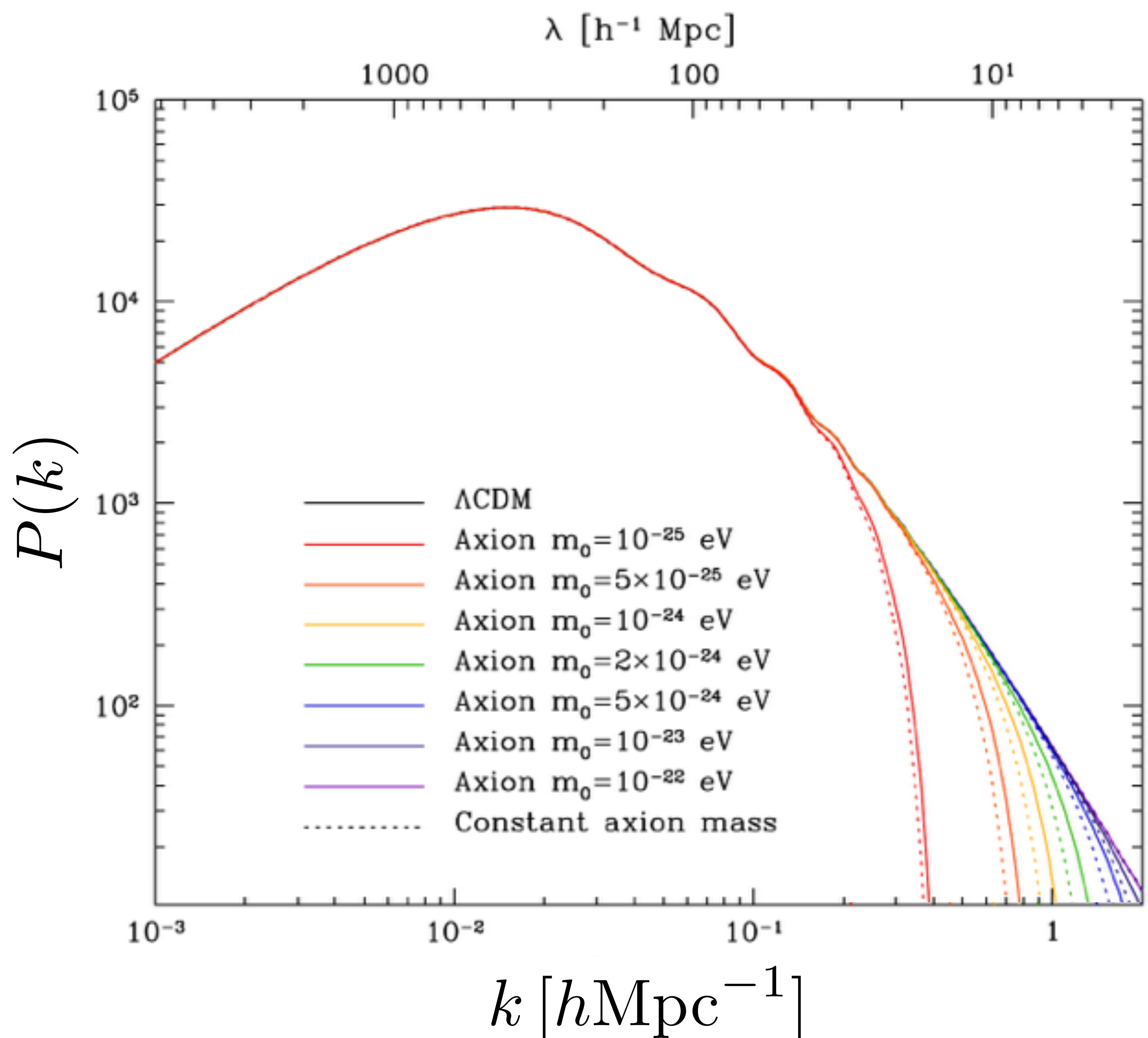
Both the neutrino mass and the optical depth can affect the observed amount of small scale structure, leading to degeneracies



Allison et al. (2015)
AL et al. (2015a)

$$\sum m_\nu = 60 \pm 19\,\mathrm{meV} \longrightarrow \pm 12\,\mathrm{meV}$$

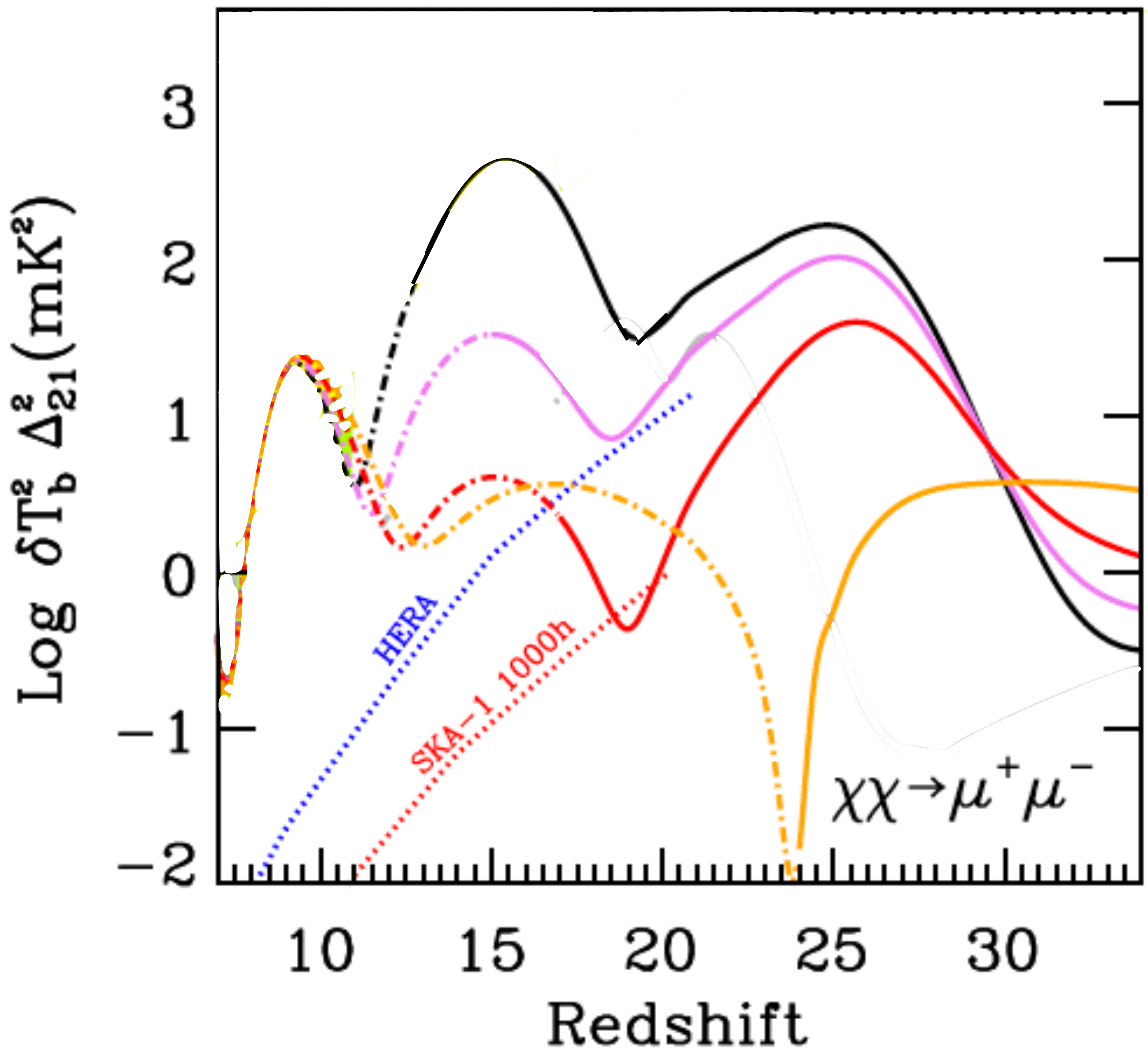
Extremely small scale modes may be accessible to futuristic 21cm cosmology experiments



Park et al. 2014

Heating from DM annihilation

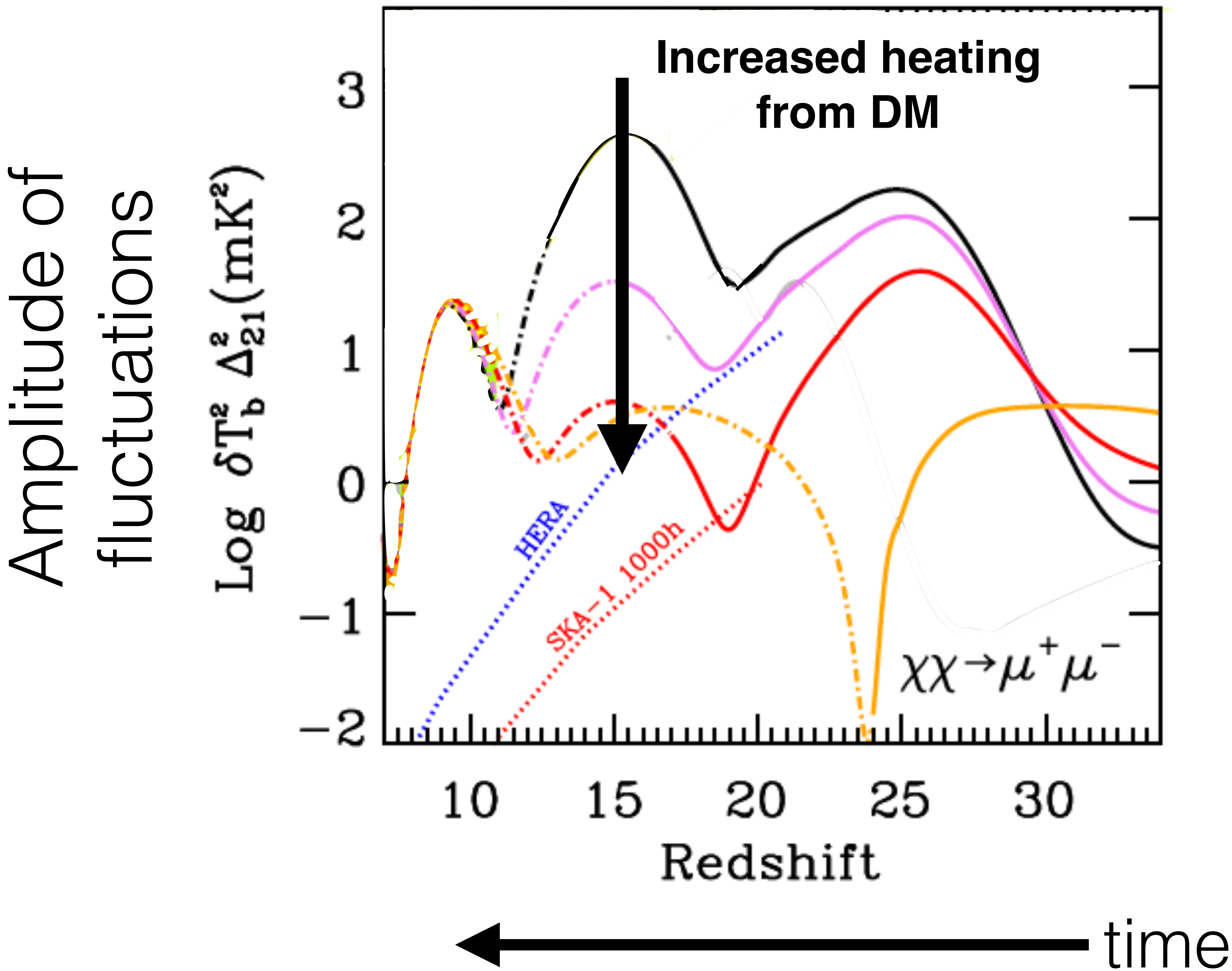
Amplitude of fluctuations



Evoli et al. 2014

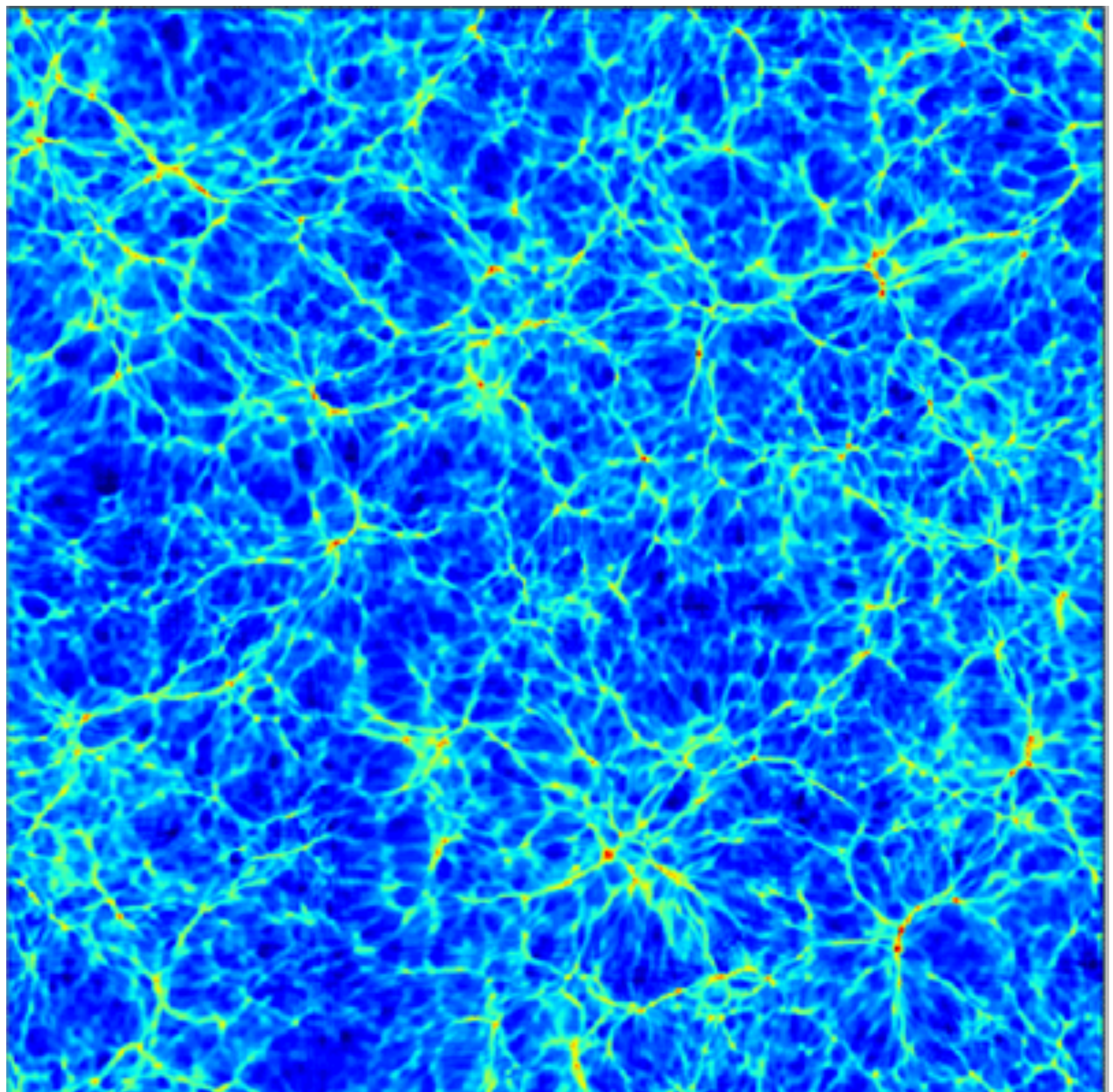
time

Heating from DM annihilation

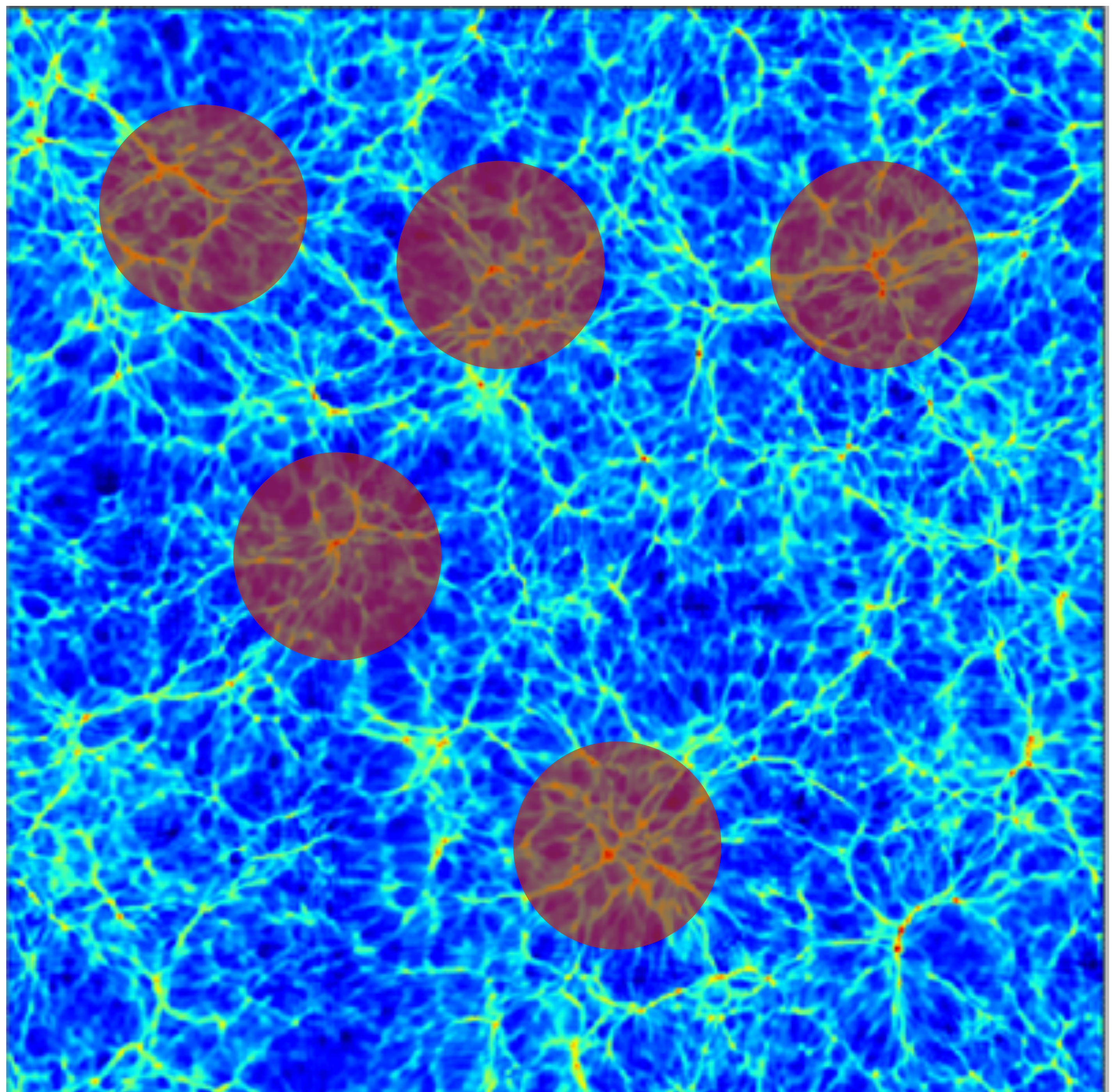


Evoli et al. 2014

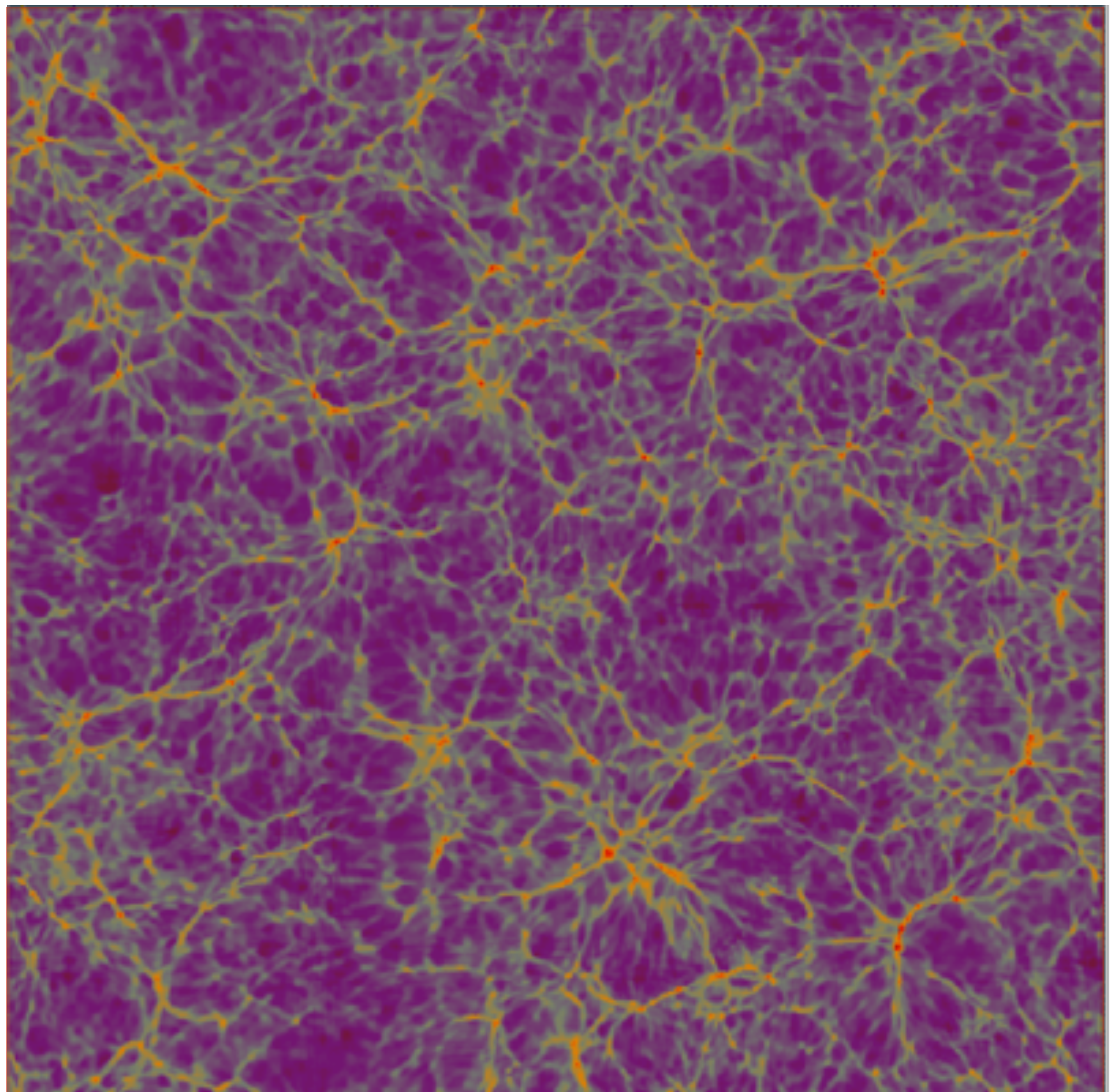
Conventional heating sources are more localized

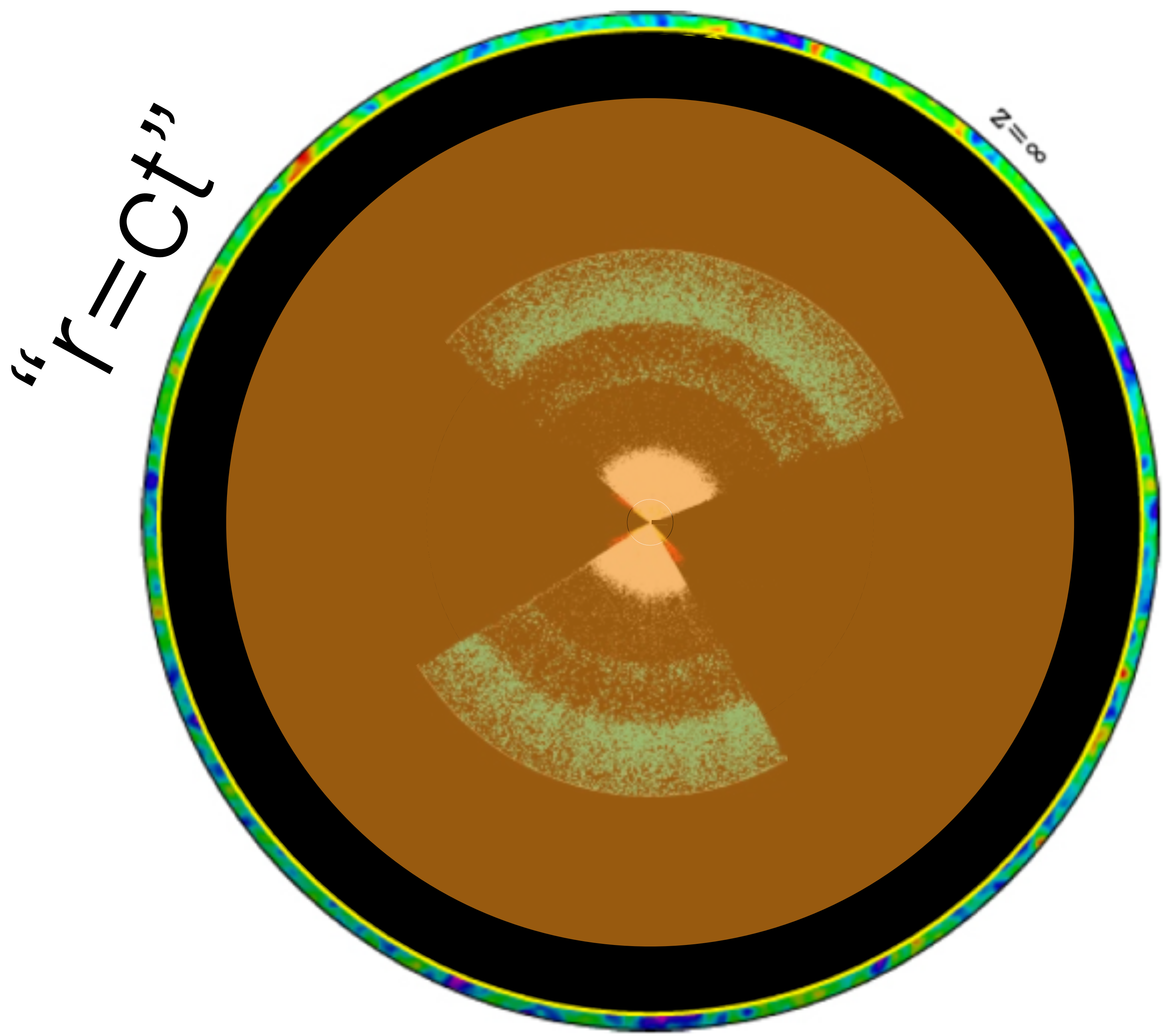


Conventional heating sources are more localized



Heating from dark matter annihilations would be more uniform, reducing the **fluctuation** amplitude




“r=ct”
z=∞

- Time frontier
    - Unique access to the pre-reionization epochs
- Sensitivity frontier
    - Large volume resolution in small errors
- Scale frontier
    - Small scale modes are easy to model using linear theory

Tegmark & Zaldarriaga 2009

Exciting times are ahead!

- We're getting close to detecting the 21cm signal—close enough to start improving our understanding of reionization.

- 21cm cosmology is a data-intensive science where astrophysics and cosmology go hand-in-hand

- The HERA experiment is being built now, and promises to deliver qualitatively new constraints on astrophysics and cosmology.

- 21cm cosmology provides a window into fundamental physics with opportunities to push the time, sensitivity, and scale frontiers.