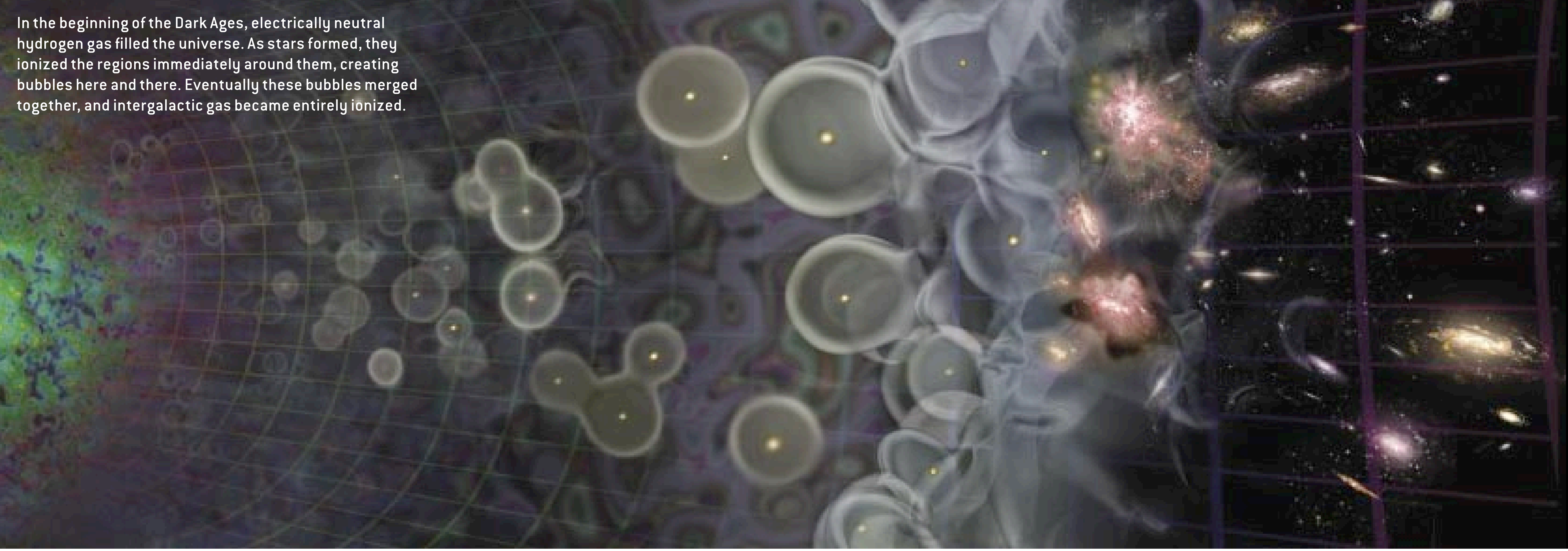


The epoch of reionization from 21cm observations

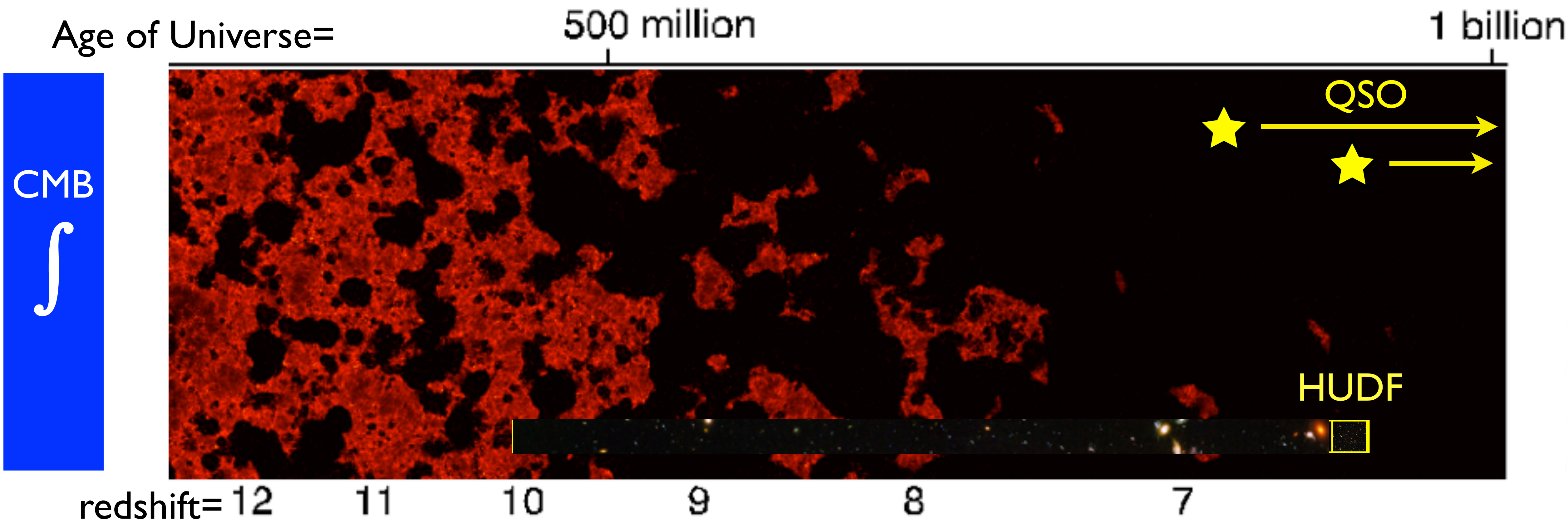
Jonathan Pritchard

Imperial College London

erc

More needed...

Imperial College London



Existing observations leaves much unanswered:
1) Lyman-alpha forest: end point z>6.5
2) CMB optical depth: mid point z~9±1.5
3) kSZ amplitude: duration z<3 ?

HST probes skewer much smaller than scale of ionized regions + only brightest sources

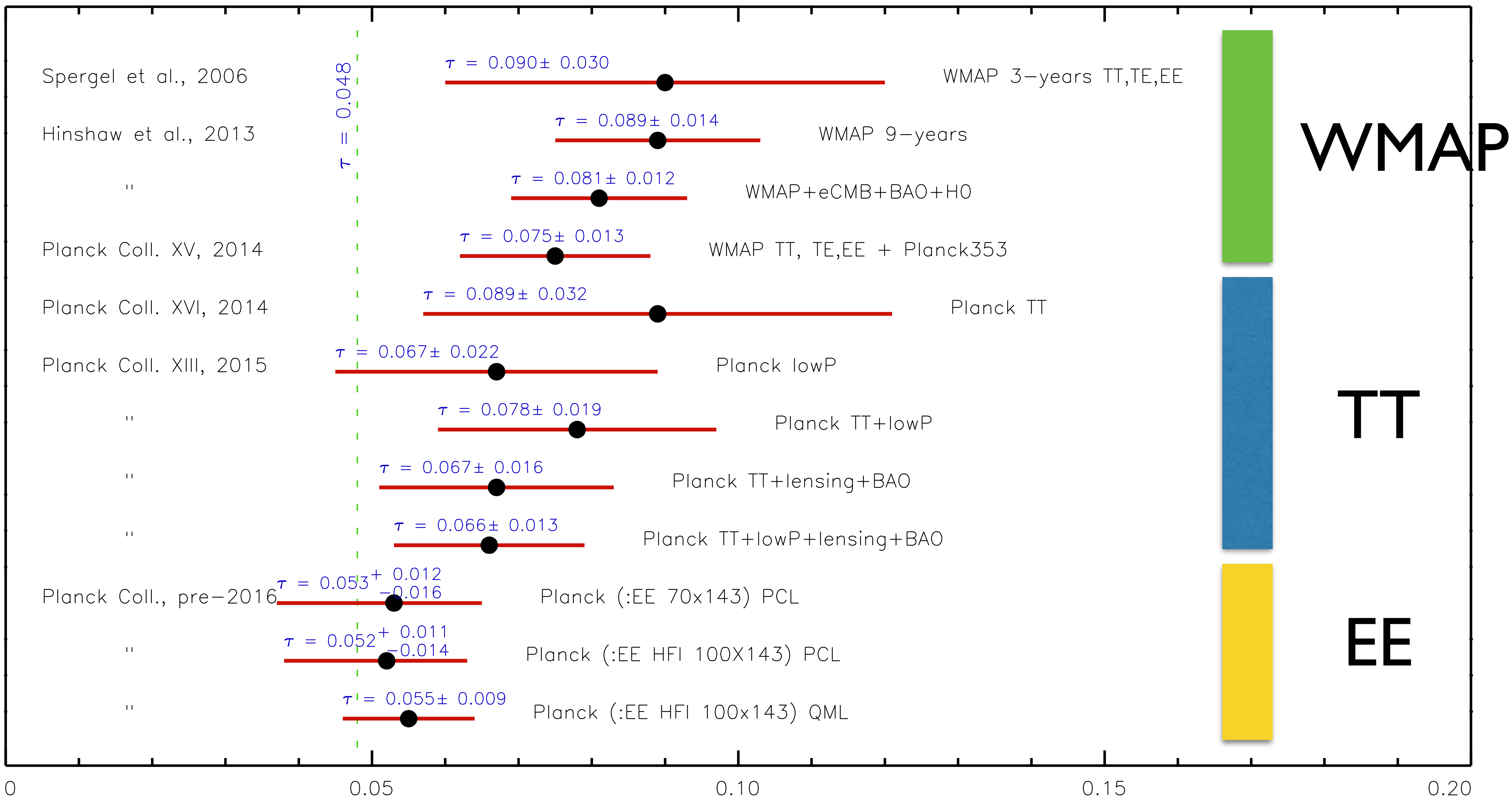
LAE surveys (e.g. HSC) may eventually show enhanced clustering from reionization
LBG/LAE surveys with Euclid possible to z~8 in deep field

Fundamental need for new types of observation to understand details of reionization

Evolving optical depth

$$\tau = \sigma_T \int \overline{n}_e(z) \frac{dl}{dz} dz,$$

Planck+ 2016

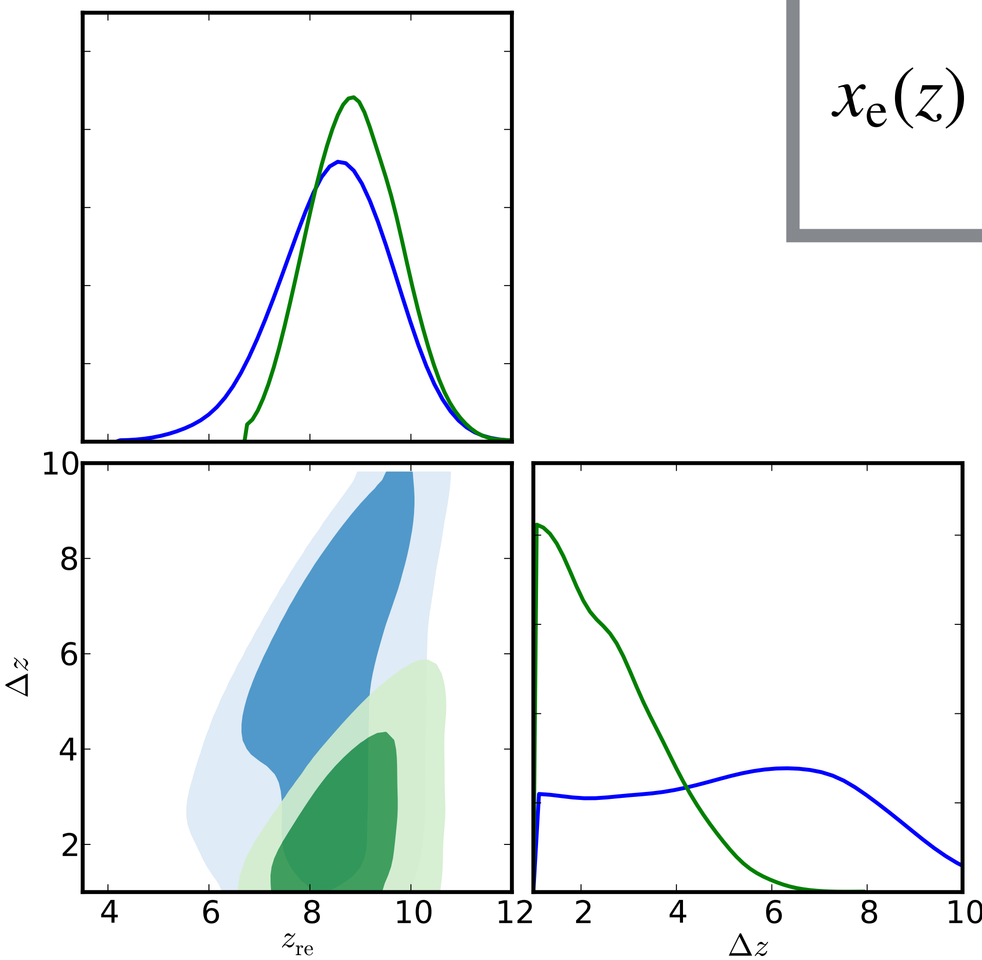


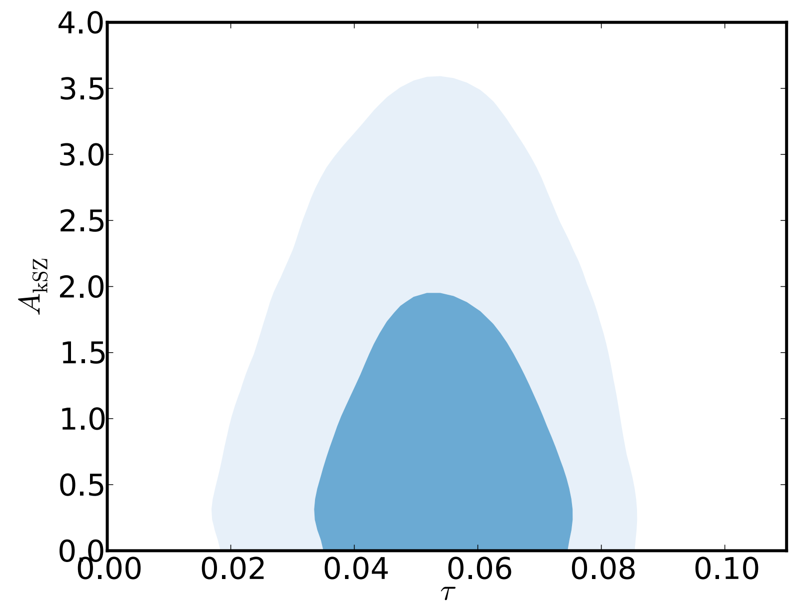
tau has fallen consistently as systematics better accounted for

Planck

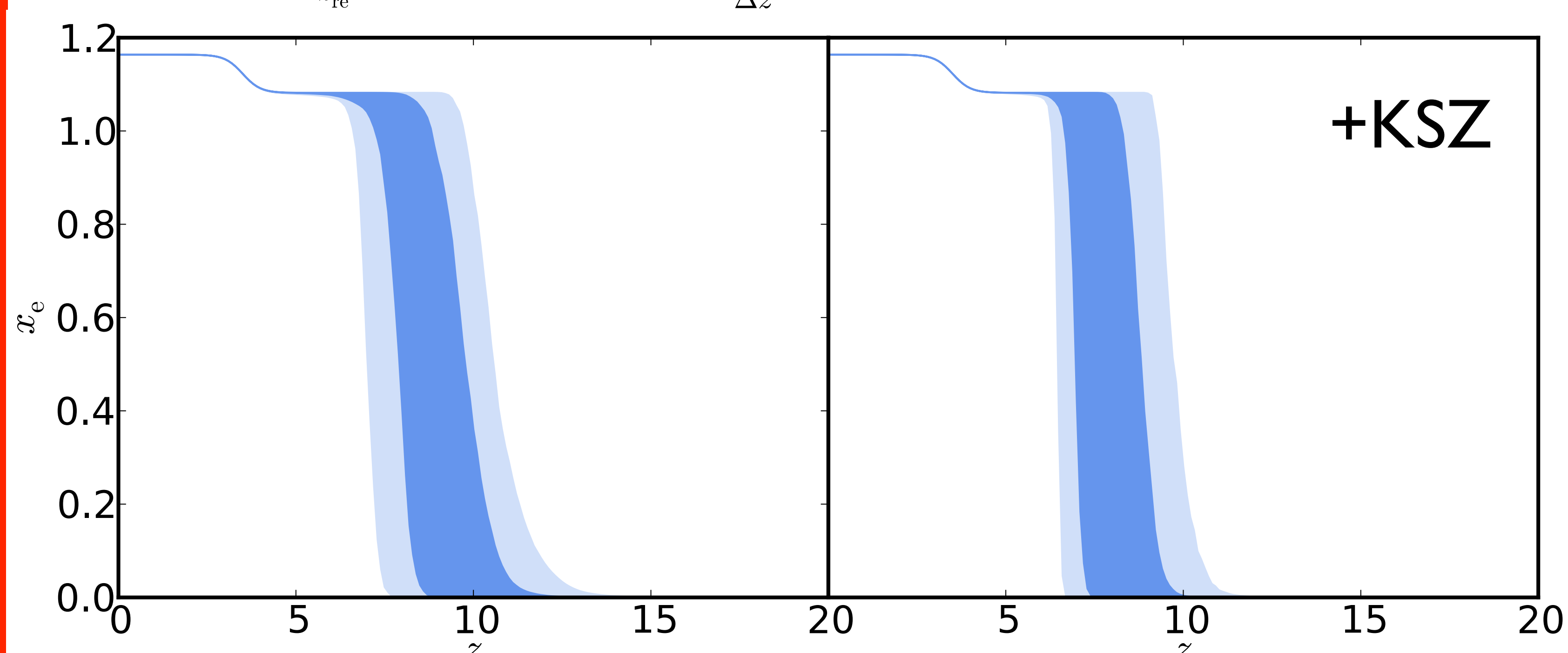
$$x_e(z) = \frac{f}{2}\left[1 + \tanh\left(\frac{y - y_{re}}{\delta y}\right)\right]$$

Planck+ 2016





KSZ can be included to constrain Δz



tau alone

$$z_{re} = 8.8^{+0.9}_{-0.9} \quad (\text{prior } z_{end} > 6)\,.$$
$$\Delta z < 4.6 \quad (95\,\%\ \text{CL}).$$

tau + KSZ

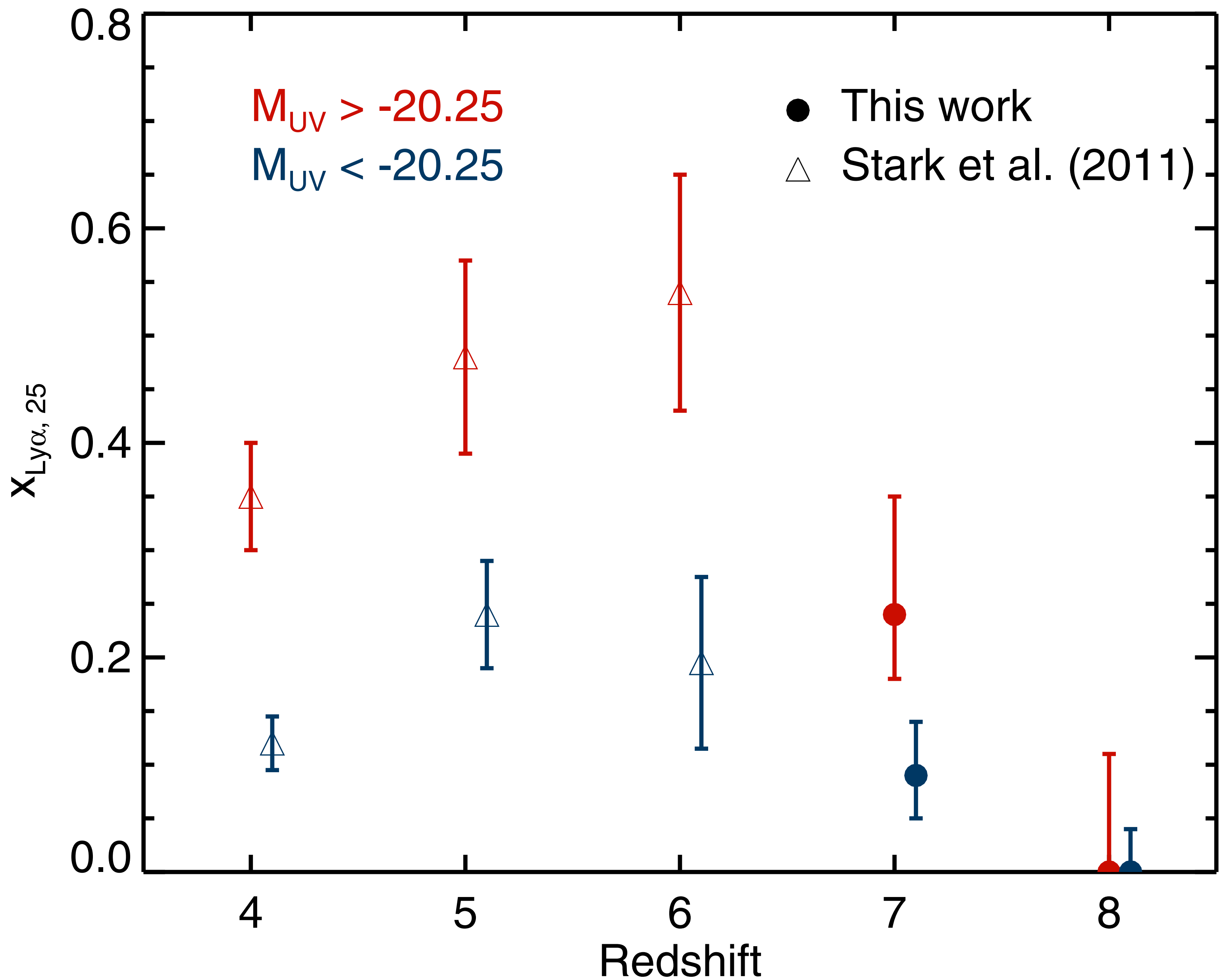
$$z_{re} = 7.8^{+1.0}_{-0.8} \quad (\text{prior } z_{end} > 6)\,,$$
$$\Delta z < 2.8 \quad (95\,\%\ \text{CL, prior } z_{end} > 6)\,.$$

HST

LBG detected up to z~10. Small numbers at z>7. JWST launches 2018

Decline in LAE/LBG fraction at z~7 possible signature of reionization

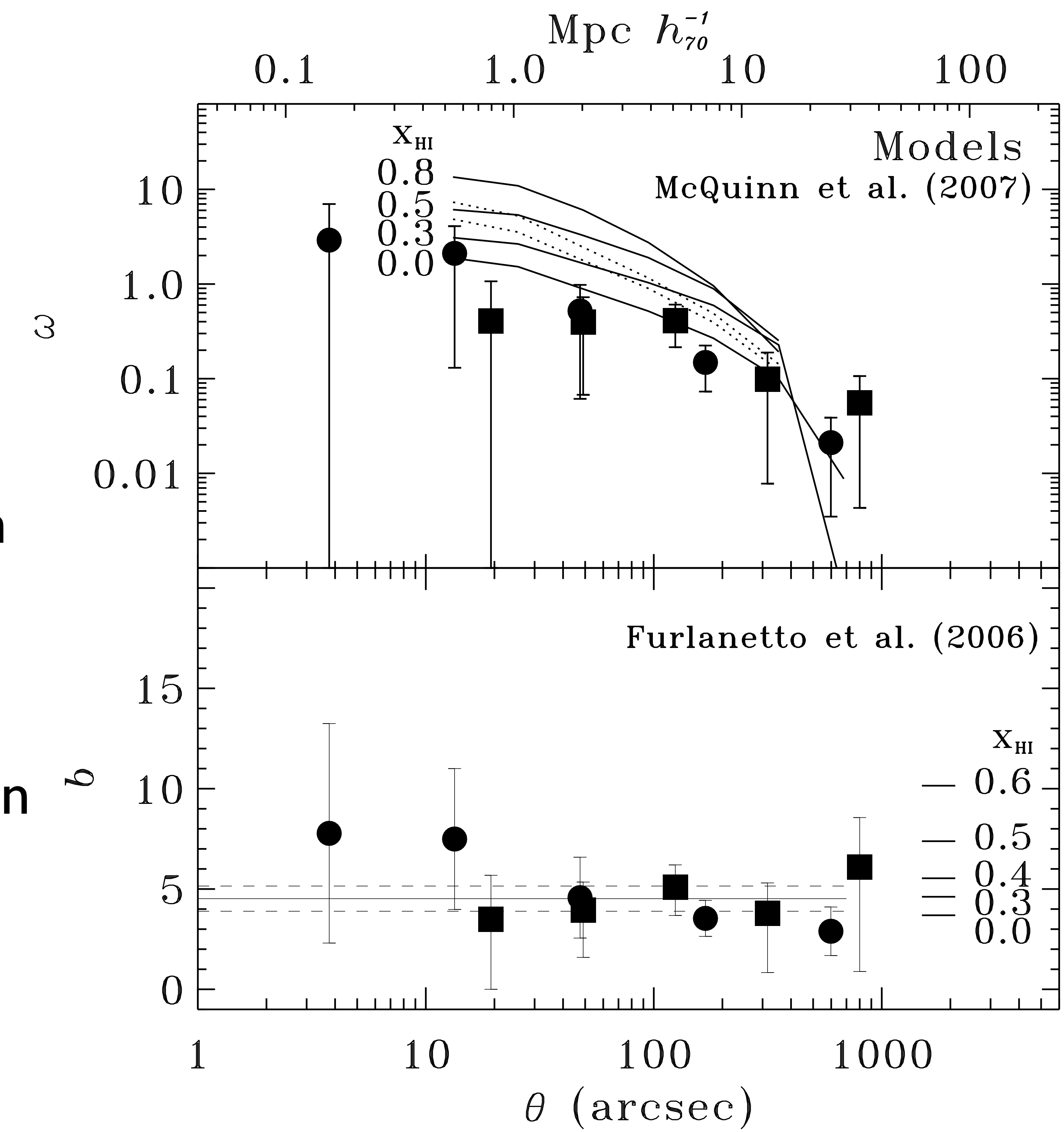
$x_{\mathrm{H\,I}} = 0.39^{+0.08}_{-0.09}$ at $z \sim 7$ and $x_{\mathrm{H\,I}} > 0.64$ at $z \sim 8$.



Schenker+ 2014

HSC/LAE

Hypersuprime Cam (HSC) early data on Subaru
~14 deg^2 (21.2 deg^2)
959 (z=5.7), 873 (z=6.6) LAE



No indication of reionization seen in clustering of LAEs

Evolution of number counts also consistent with evolution for 10^{11} MSol halos

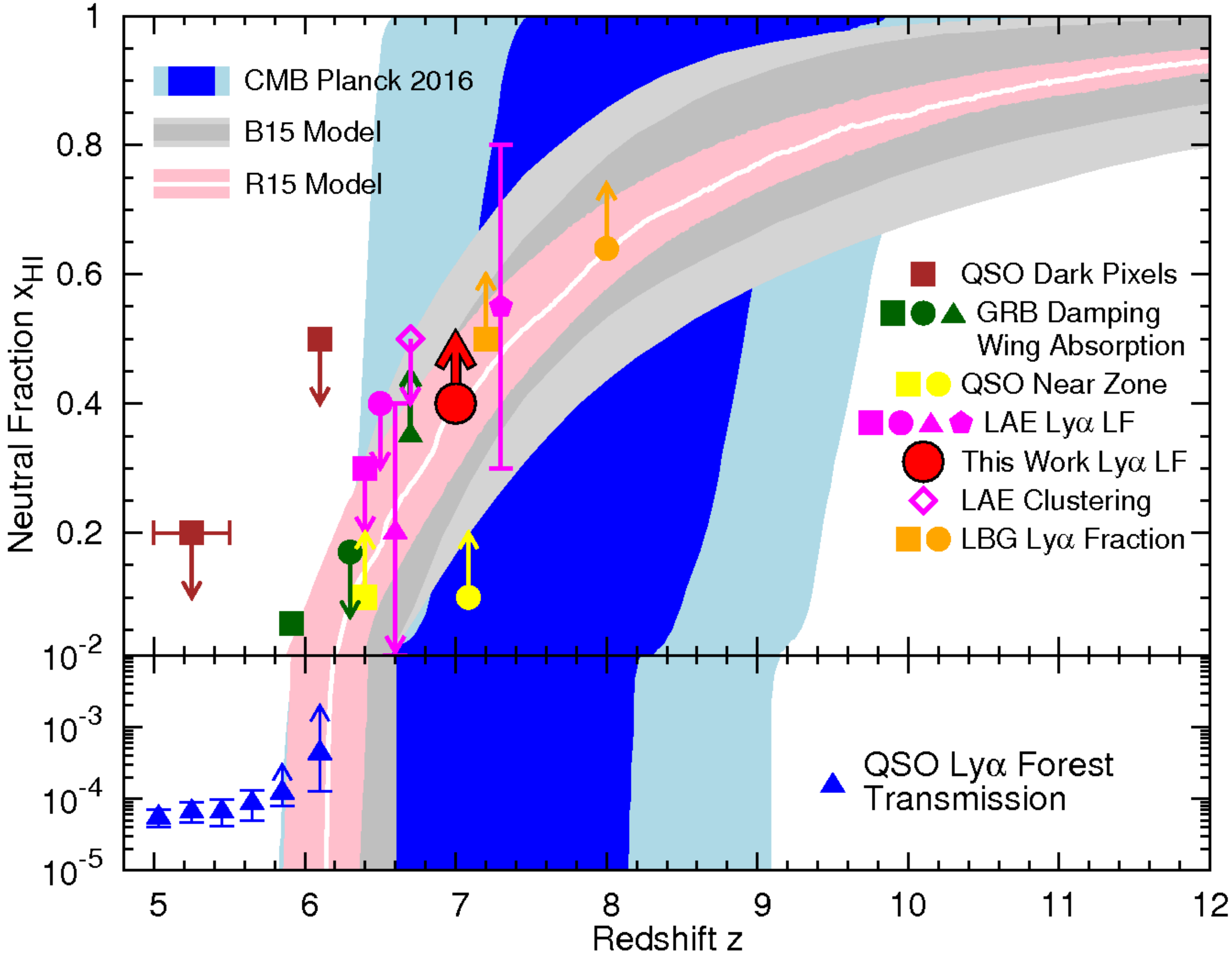
x_{HI}<0.3 at z=6.6

Ouchi+ 2017

Reionization summary

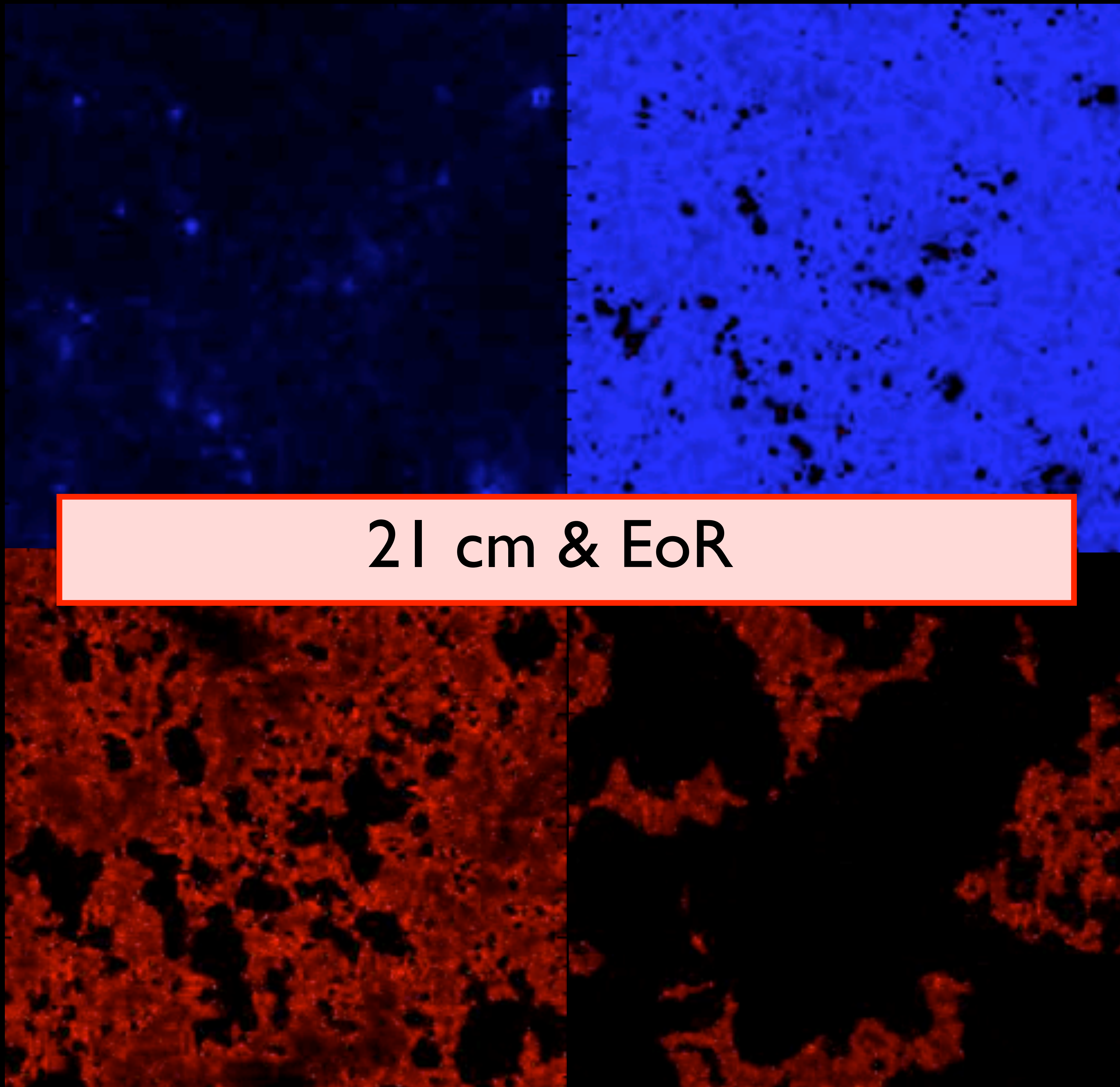
Ultimately need to combine all reionization constraints
e.g. Robertson+ (2015), Greig & Mesinger (2016), … many more

Reionization histories typically very model dependent



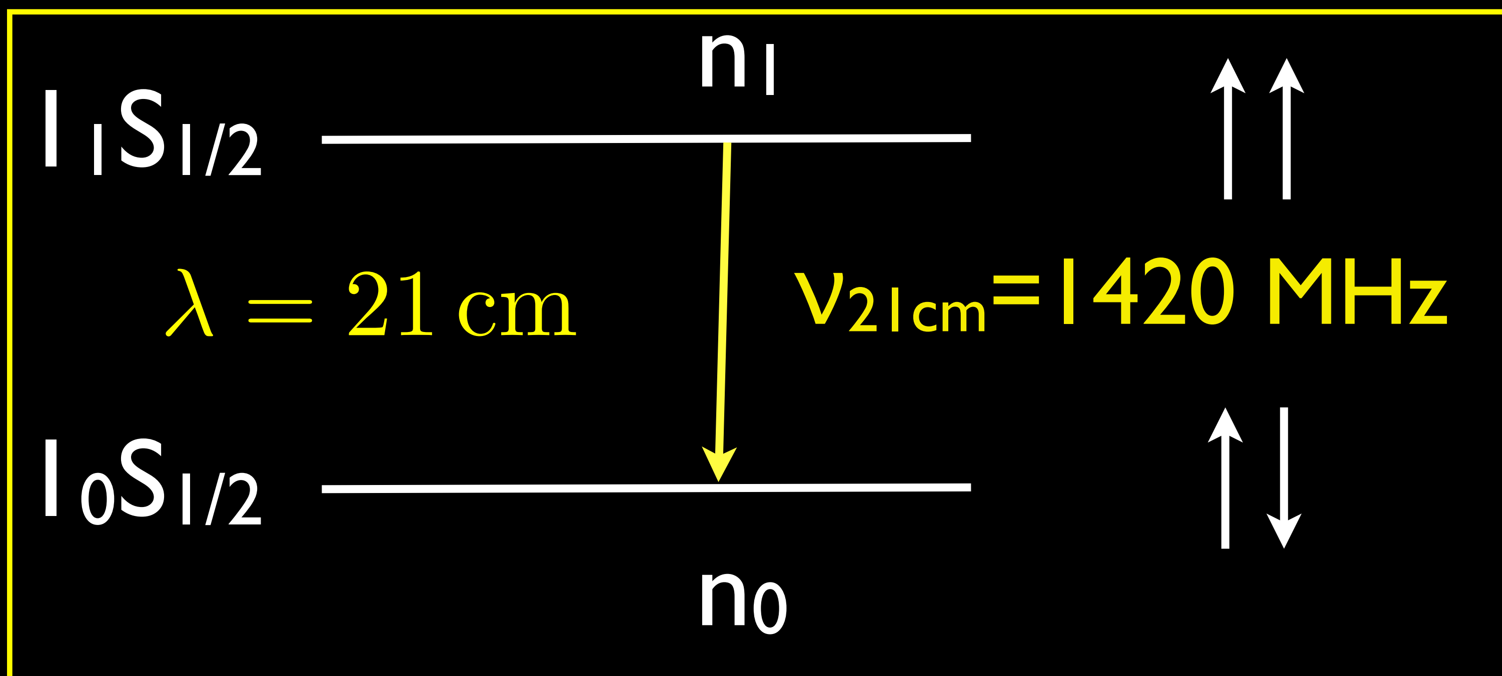
Ota+ 2017

21 cm & EoR



21 cm basics

Hyperfine transition of neutral hydrogen



Useful numbers:

$200\,\mathrm{MHz} \rightarrow z = 6$

$100\,\mathrm{MHz} \rightarrow z = 13$

$70\,\mathrm{MHz} \rightarrow z \approx 20$

50 MHz => z~27

$t_{\mathrm{Age}}(z = 6) \approx 1\,\mathrm{Gyr}$

$t_{\mathrm{Age}}(z = 10) \approx 500\,\mathrm{Myr}$

$t_{\mathrm{Age}}(z = 20) \approx 150\,\mathrm{Myr}$

t_{Age}(z=27) ~100 Myr

$t_{\mathrm{Gal}}(z = 8) \approx 100\,\mathrm{Myr}$

Spin temperature describes relative occupation of levels

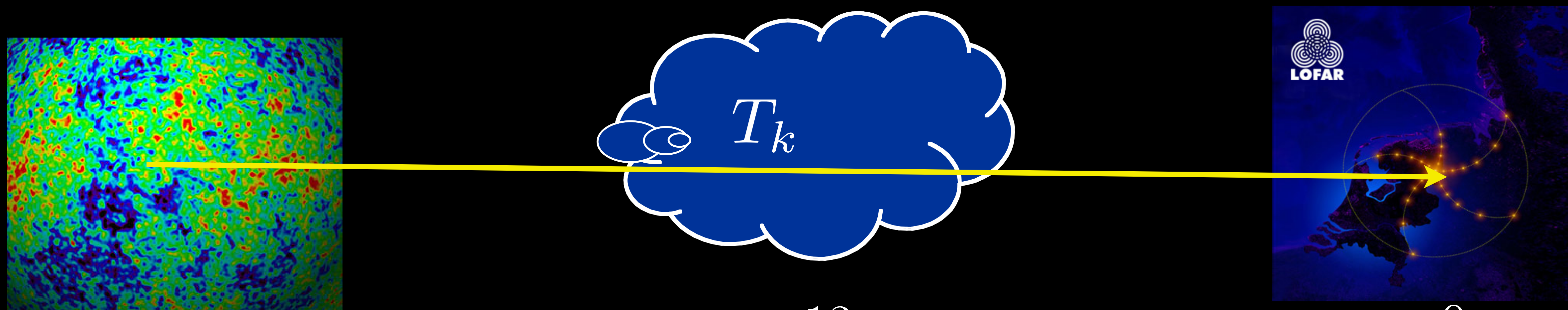
$$n_1/n_0 = 3\exp(-h\nu_{21\mathrm{cm}}/kT_s)$$

Jonathan Pritchard

21 cm line in cosmology

Imperial College London

T_γ T_S T_b



CMB acts as back light	$z=13$ $\nu = 1.4\,\mathrm{GHz}$ Neutral gas imprints signal	$z=0$ $\nu = 100\,\mathrm{MHz}$ Redshifted signal detected

neutral fraction · baryon density · spin temperature · peculiar velocities

brightness temperature

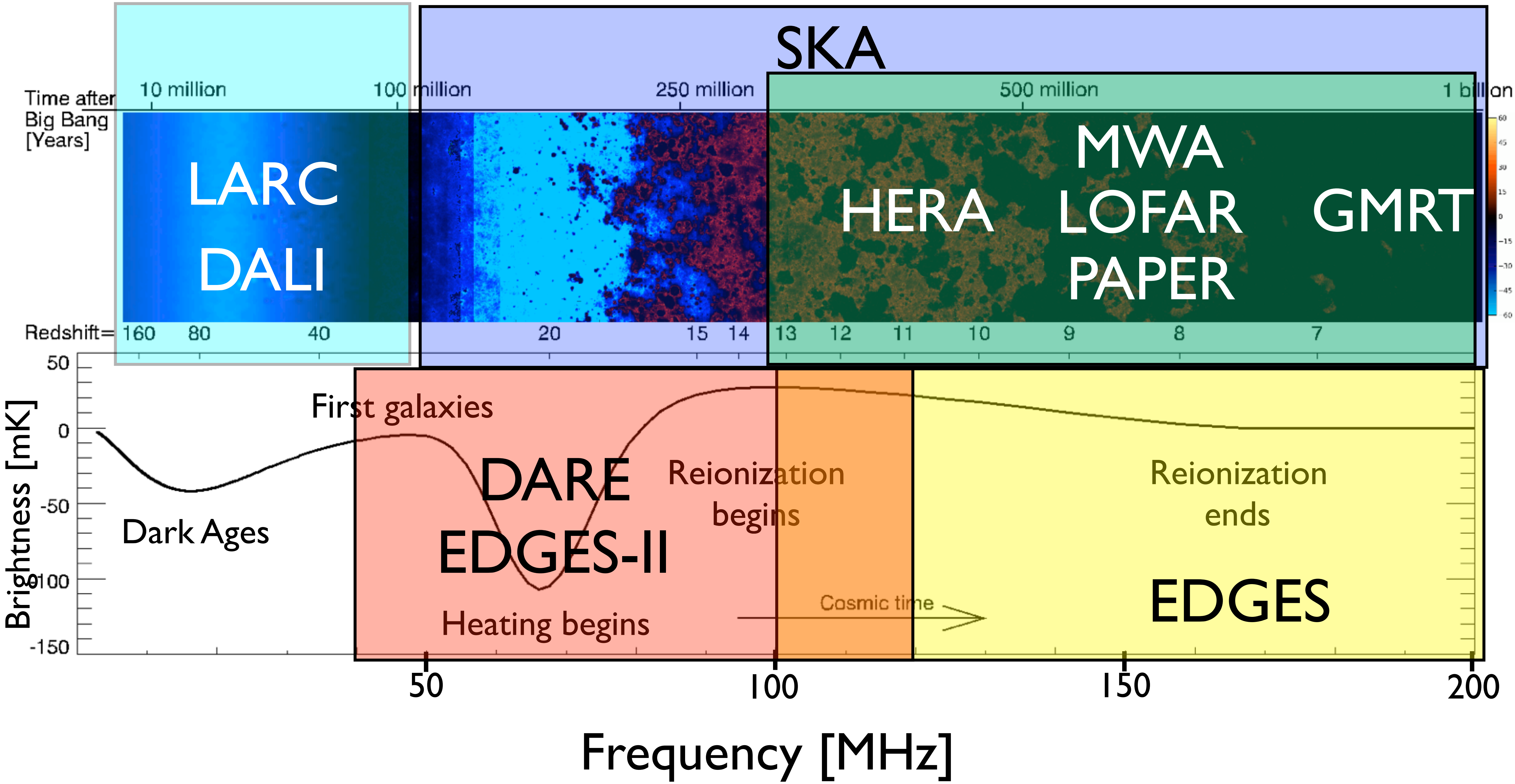
$$T_b = 27 x_{\rm HI}(1+\delta_b)\left(\frac{T_S - T_\gamma}{T_S}\right)\left(\frac{1+z}{10}\right)^{1/2}\left[\frac{\partial_r v_r}{(1+z)H(z)}\right]^{-1}\,\mathrm{mK}$$

spin temperature set by different mechanisms:

- Radiative transitions (CMB) $T_{Spin} \Rightarrow T_{CMB}$
- Collisions $T_{Spin} \Rightarrow T_{gas}$
- Wouthysen-Field effect (resonant scattering of Lyα) $T_{Spin} \Rightarrow T_{gas}$

21 cm astrophysics

Pritchard & Loeb 2010

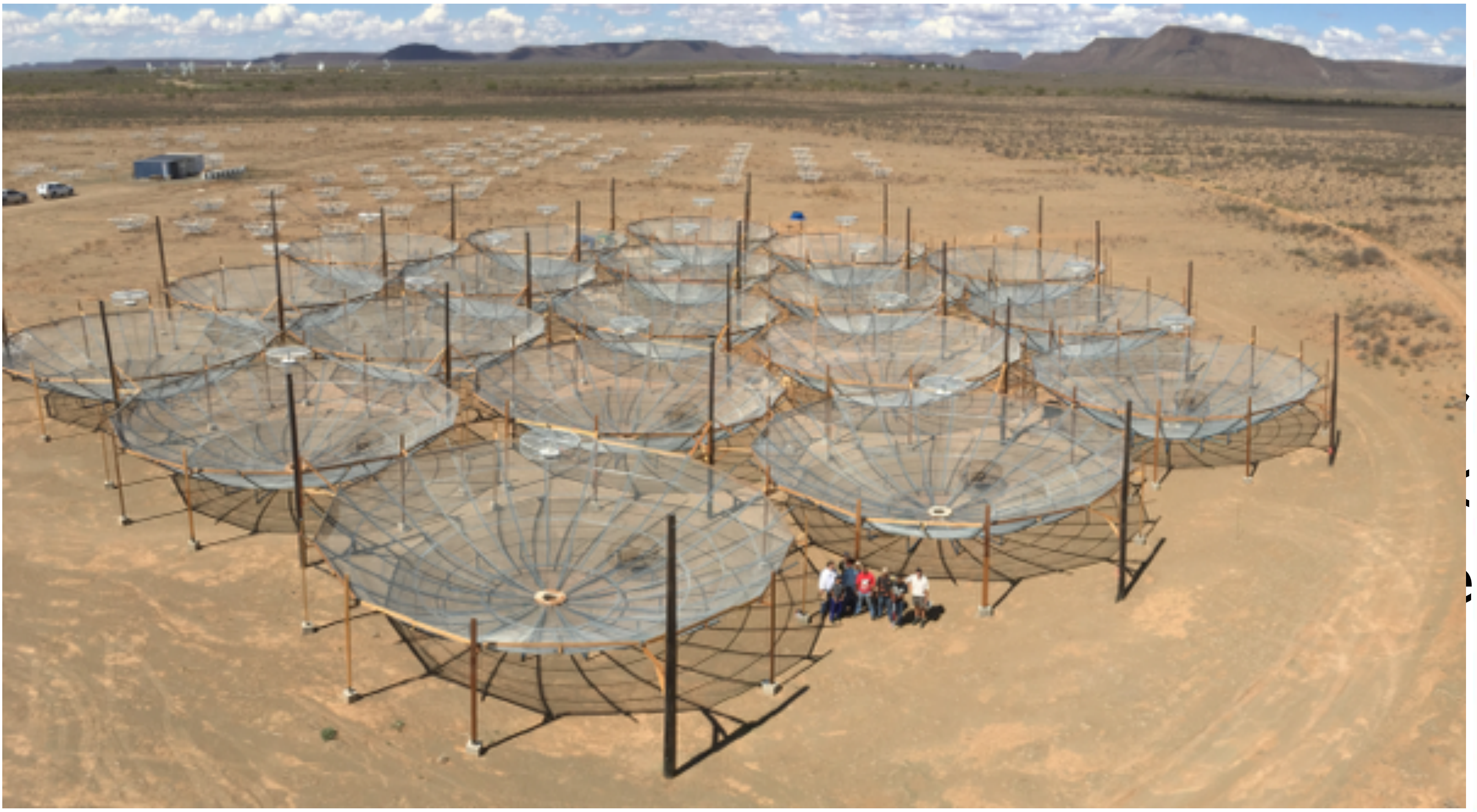


Systematic path to probing different epochs

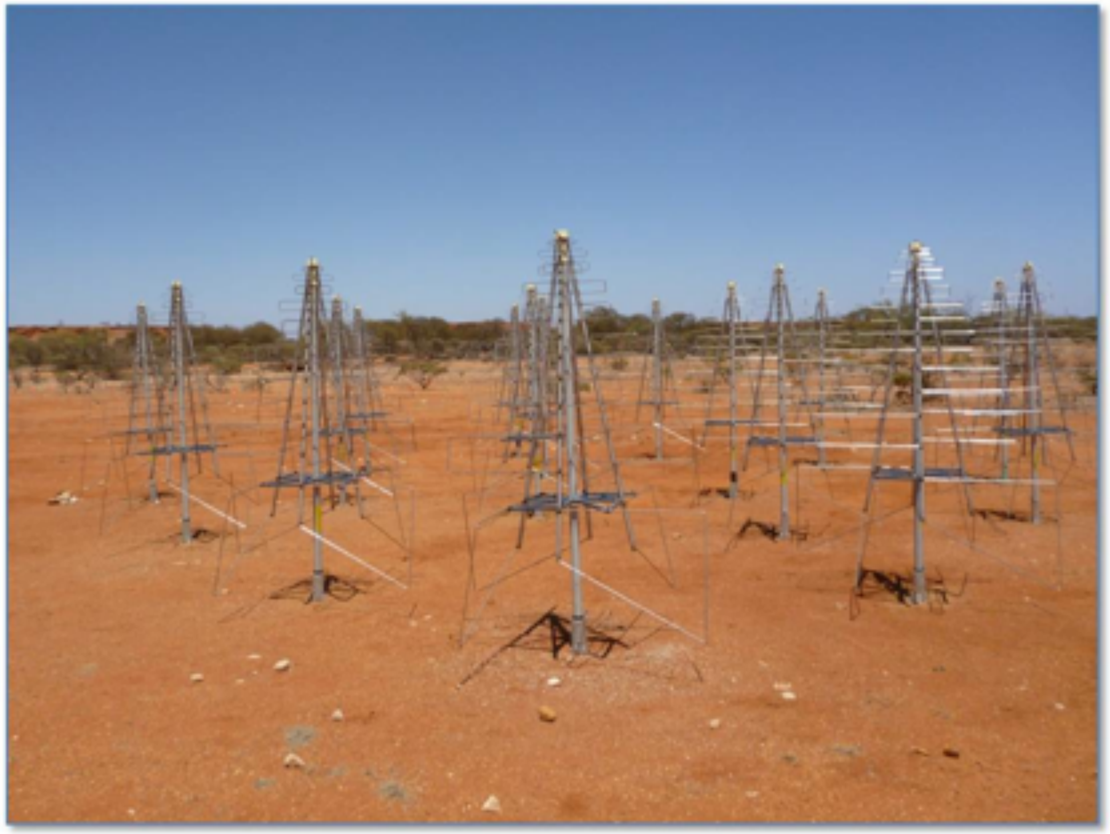
LOFAR

MWA

PAPER
(completed)



HERA
(under construction)

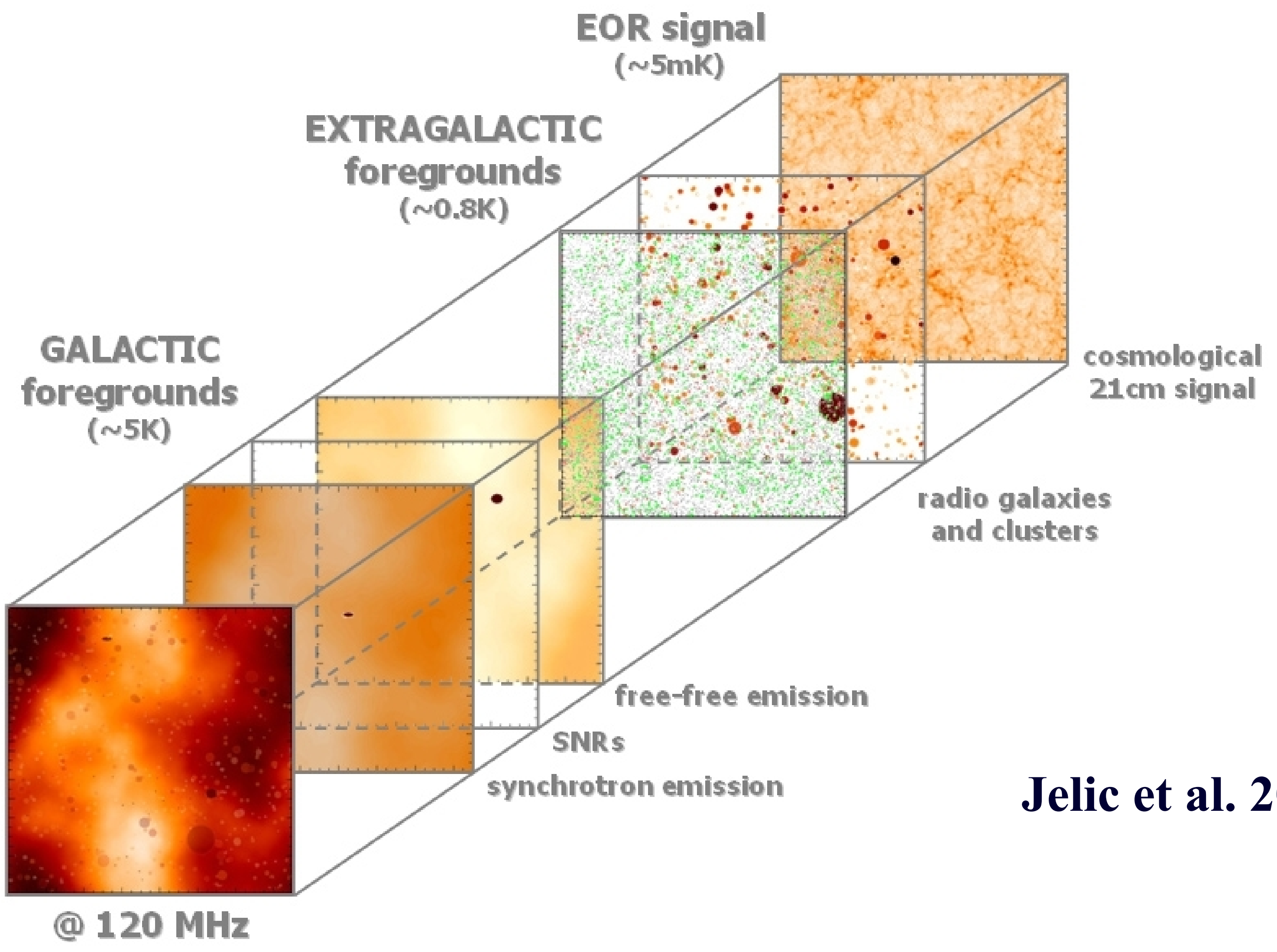


SKA
(roll out 2019-2024)

21cm challenges

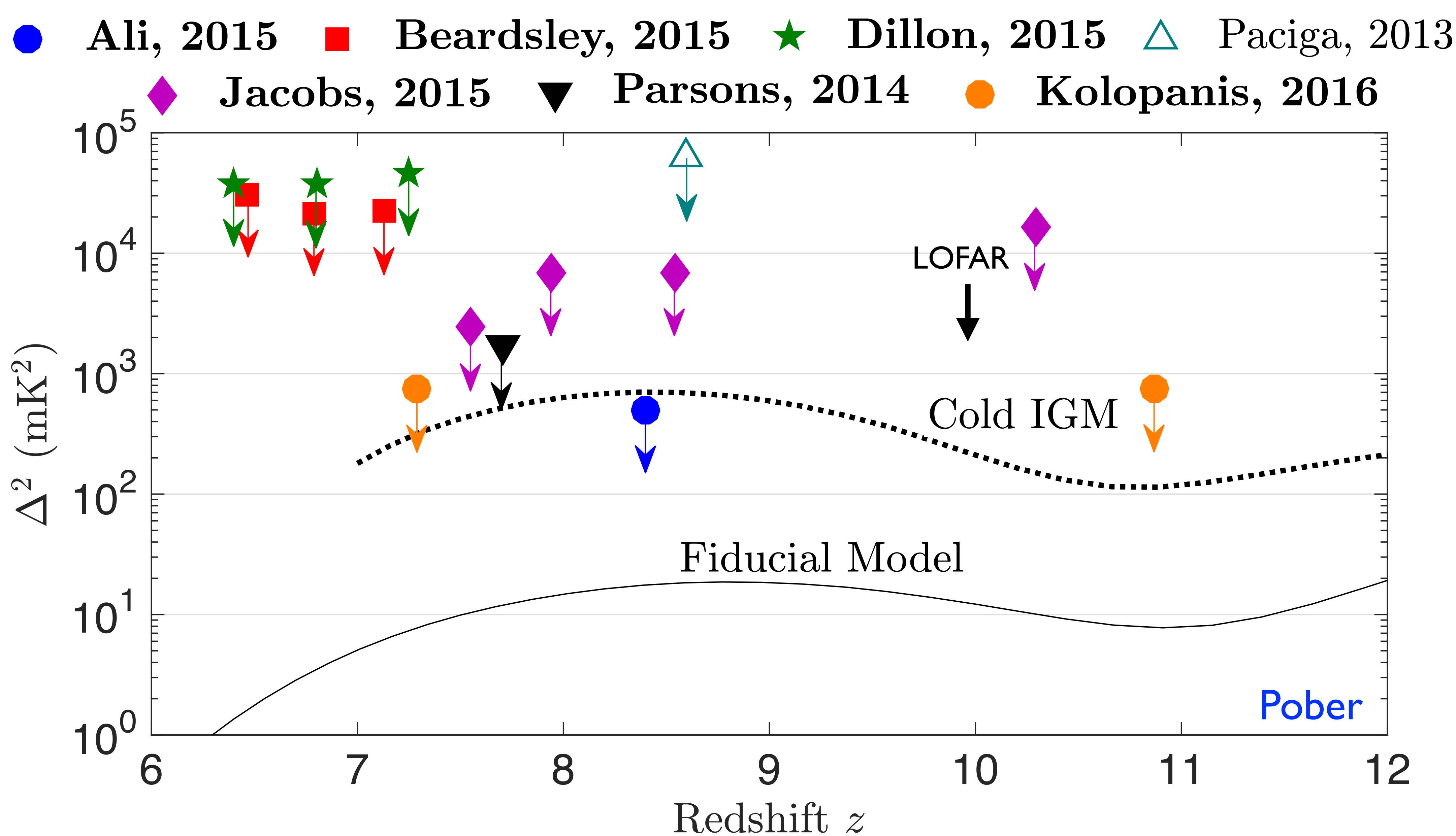
Calibration - sky vs redundant (vs hybrid) calibration

Foregrounds ~ 10^3-10^5 signal. Diffuse & pt sources



Jelic et al. 2008

Continuing improvement



Upper limits beginning to make contact with possible models - rule out unheated IGM

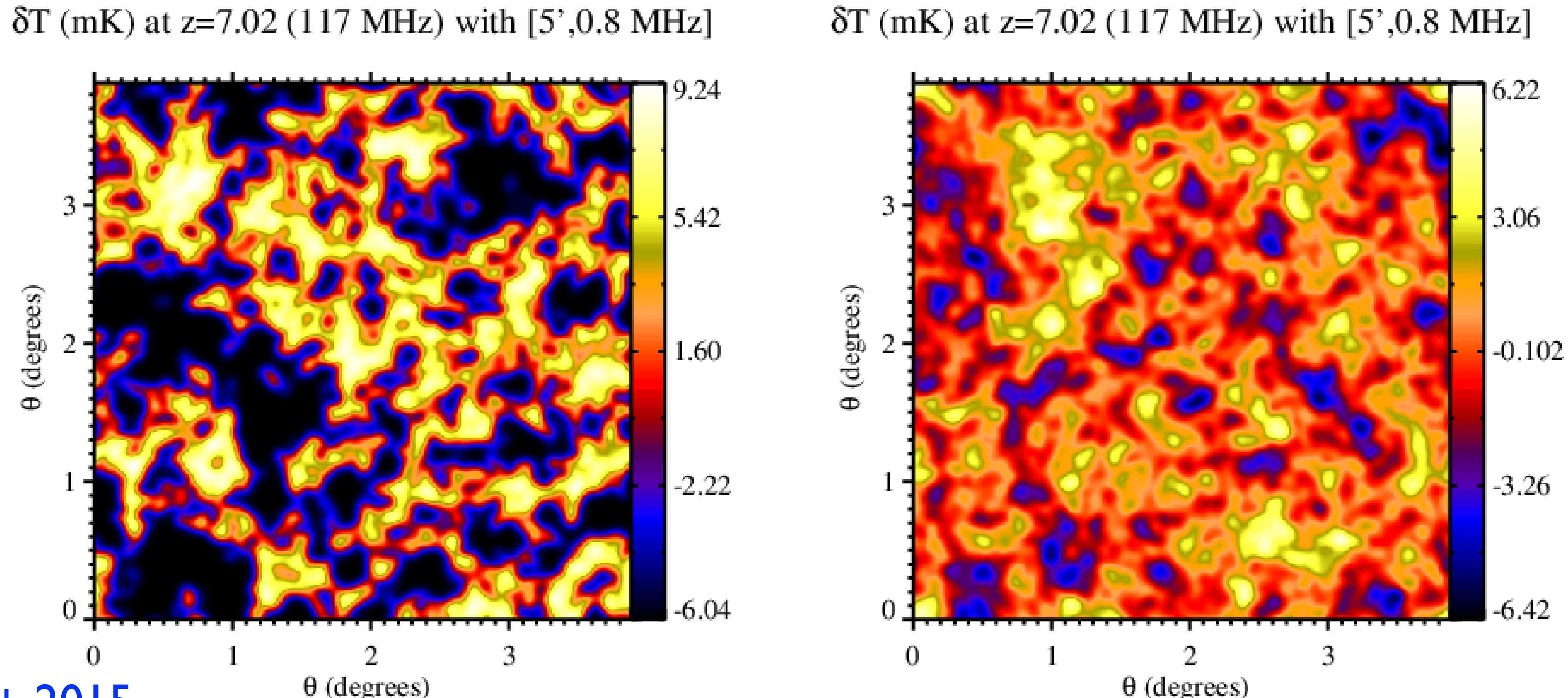
(1) Imaging

3D maps of topology of reionization - ~10 arcmin resolution

Directly image HII region around AGN/bright sources

Environmental information for other probes of reionization

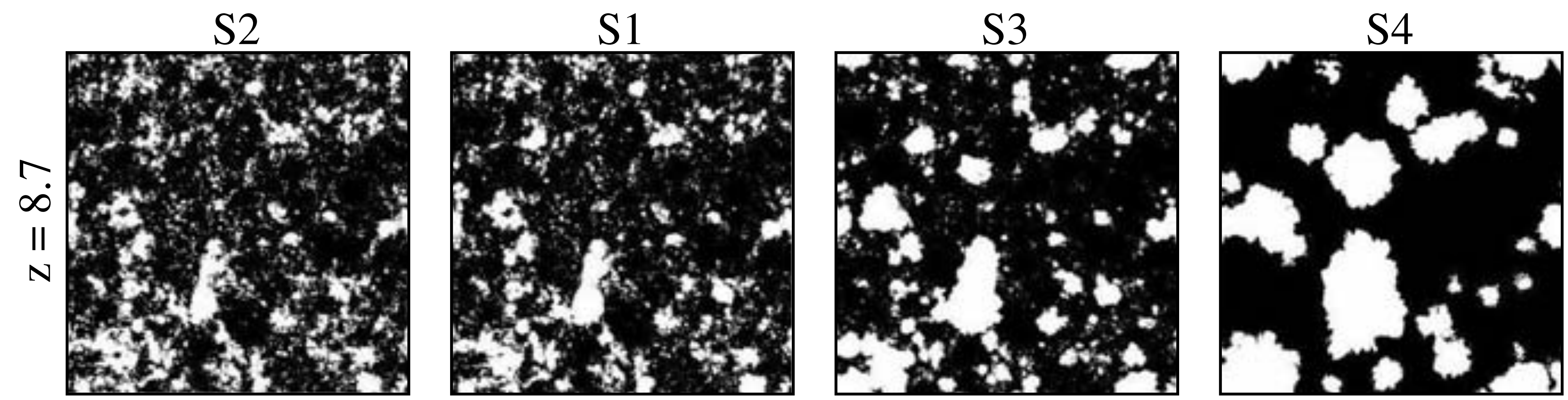
HII bubble catalogs and statistics



Mellema+ 2015

Granulometry

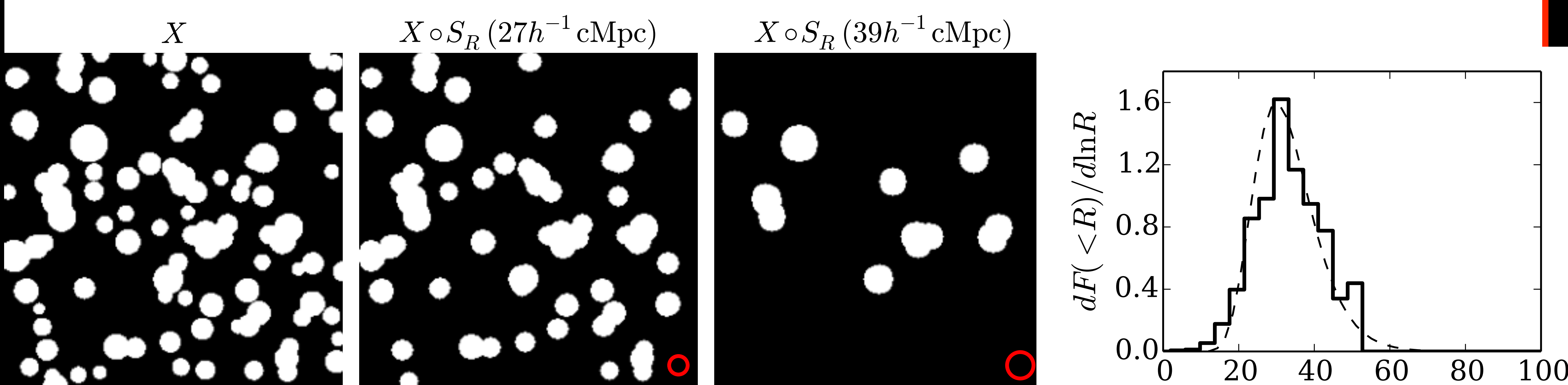
Size distribution of ionised regions is an interesting probe of reionization



McQuinn+ (2007)

$m^{-2/3}$ increasing mass dependence $m^{2/3}$

Granulometry: Sieve 21cm map from small to large scales to reconstruct sizes Kakiichi, Majumdar, Mellema+ (2016)

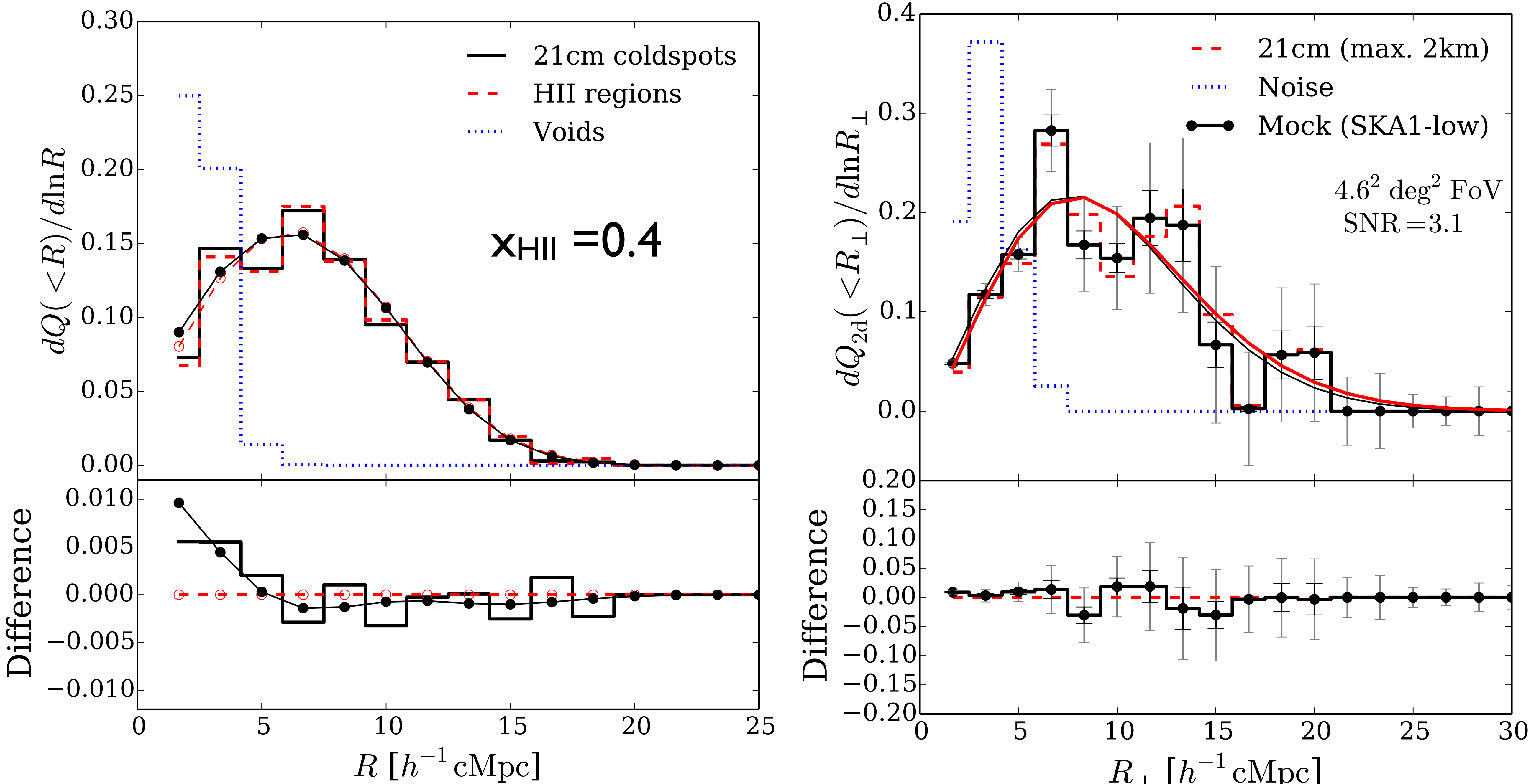


unsieved small large

Granulometry on 21cm image

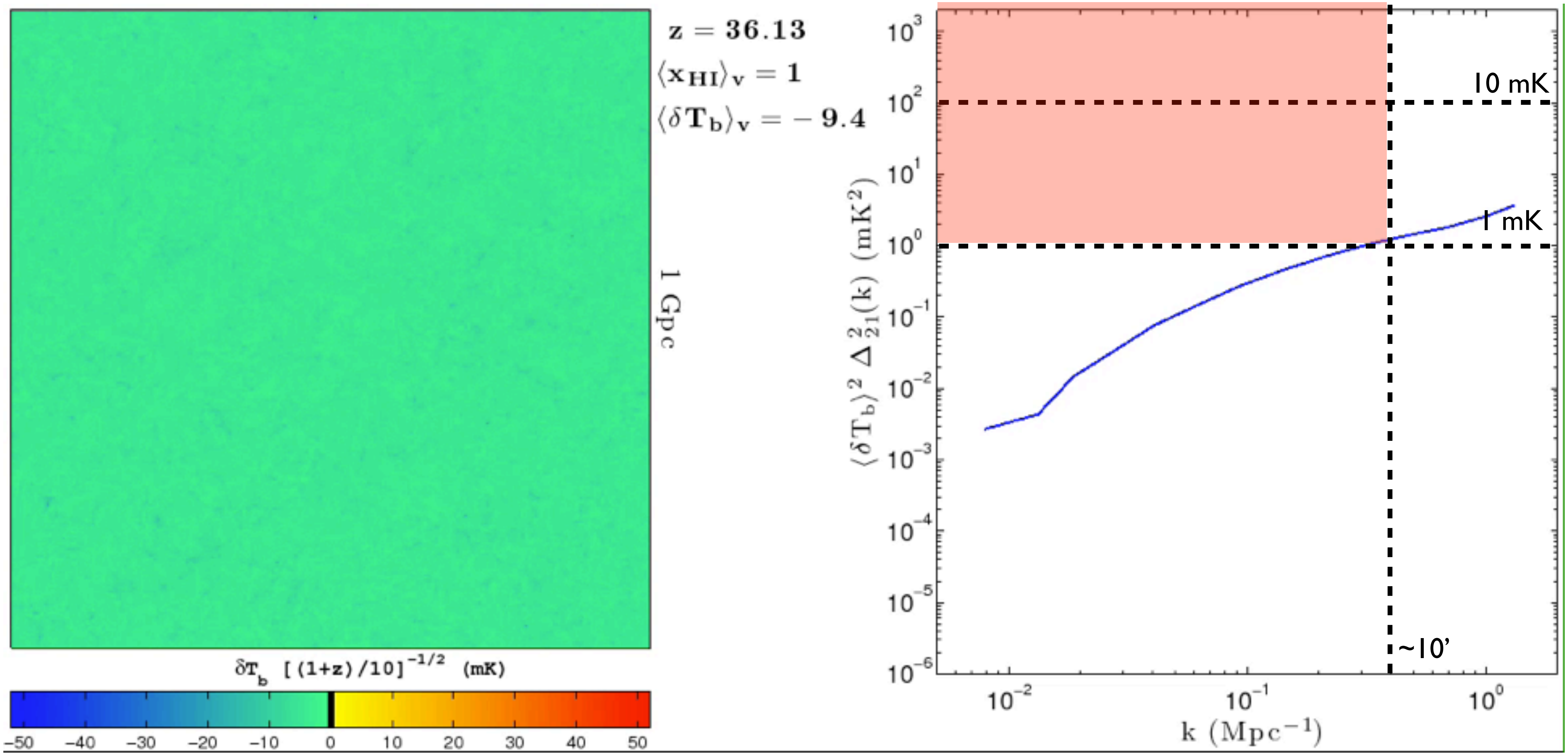
21 cm cold spots arise two ways
- ionised regions or underdense voids

SKA likely needed to study size distributions



Kakiichi+ (2016)

(2) Evolution of the power spectrum



Mesinger+ 2010

Rich science contained in spatial and redshift evolution of 21cm power spectrum - observe z = 6 - 27

Parameter estimation

Imperial College London

Inputs		Outputs	

$$P(\mathbf{d}|\theta, M) \times P(\theta|M) = P(\theta|\mathbf{d}, M) \times P(\mathbf{d}|M)$$

Likelihood — Prior — Posterior — Evidence

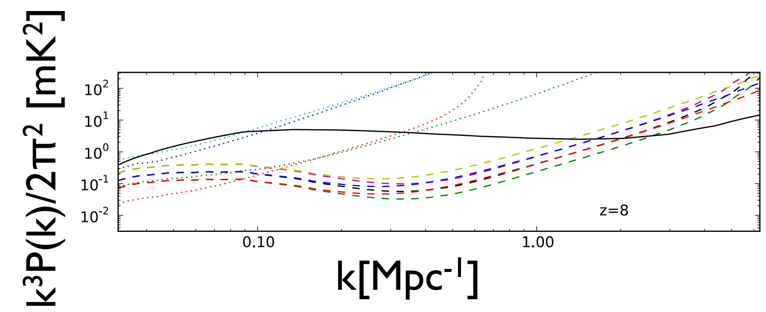
models — parameters

$$(f_*, \xi_{\rm ion}, \xi_X, \xi_\alpha, \lambda_{\rm mfp})$$

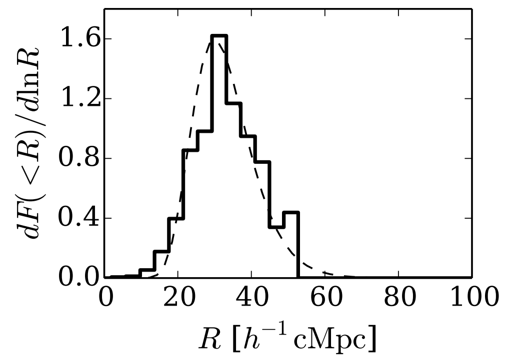
forward model →

← statistical inversion

summary statistics
e.g. power spectrum



size distribution



1) Parameter estimation
2) Model selection

Emulation

Reionization simulations often computationally expensive, but vital for 21 cm analysis - approximate schemes needed

Given a computationally expensive model y(x), for parameters x, build an "emulator" $\bar{y}(x)$ that is fast to evaluate

1) Form training set of well chosen samples $\{x_{train}, y_{train}=y(x_{train})\}$
2) Establish basis for approximation
3) Establish means of interpolating from training set to arbitrary points

Used extensively in climate change and also in cosmology

e.g. Coyote universe - emulation of non-linear matter power spectrum P(k | parameters) Heitmann+ 2006, Habib+ 2007

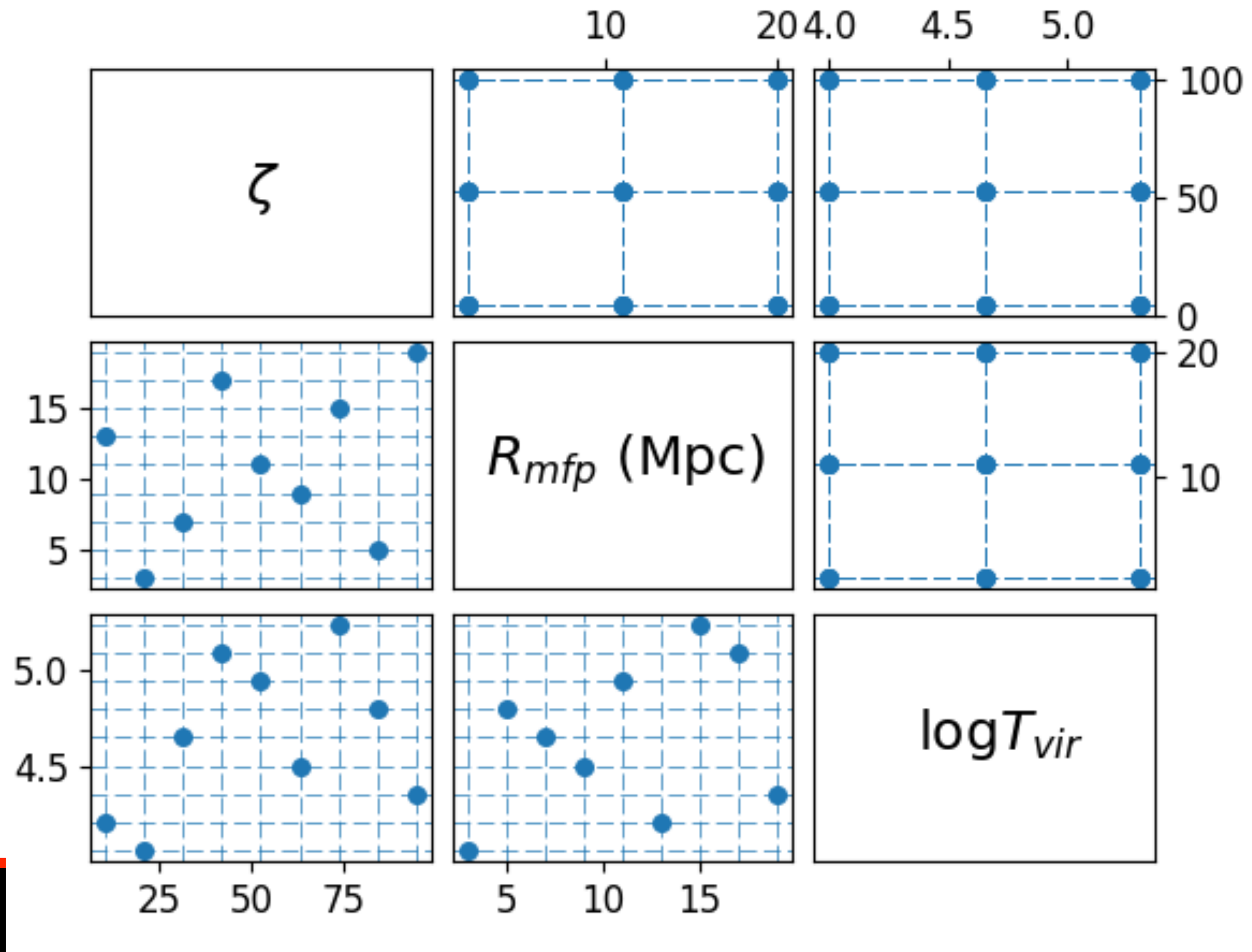
- Latin hyper-cube used to build training set
- PCA basis
- Gaussian processes for interpolation

Latin hypercube sampling

Curse of dimensionality: number of points required to sample in a regular grid scales as $(L/\Delta)^N$ - prohibitive for even modest N

Latin hypercube sampling - distribute M points such that each parameter is divided into M bins with each bin occupied only once.

More efficiently uses samples, especially if tweaked to reduce clustering of points e.g. minimax nearest neighbour distances



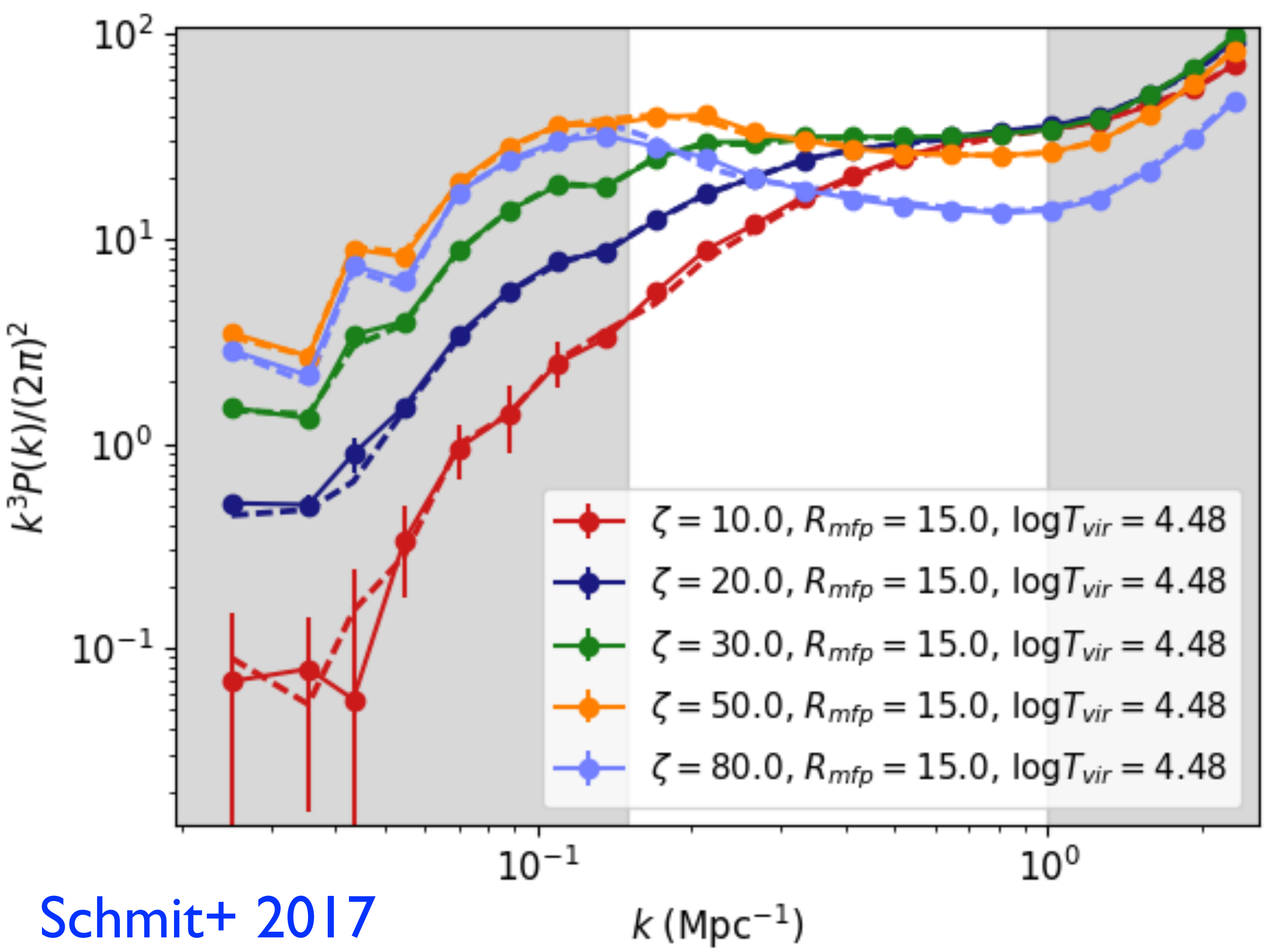
Schmit+ 2017

Bayesian inference with emulation

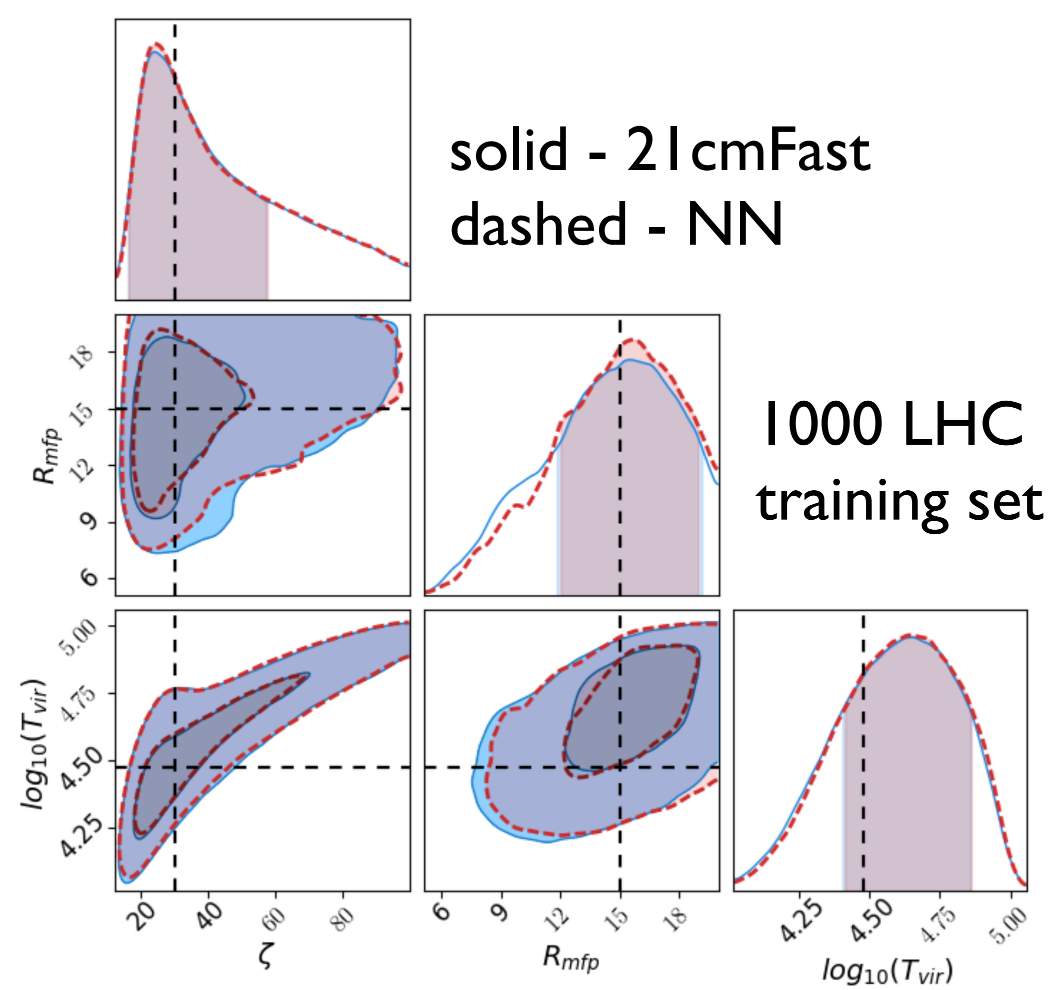
Different interpolation schemes - Gaussian Processes Kern+ 2017
- Neural networks Schmit+ 2017

Sampling density in MCMC proportional to target density
=> implies many samples are closely spaced in parameter space

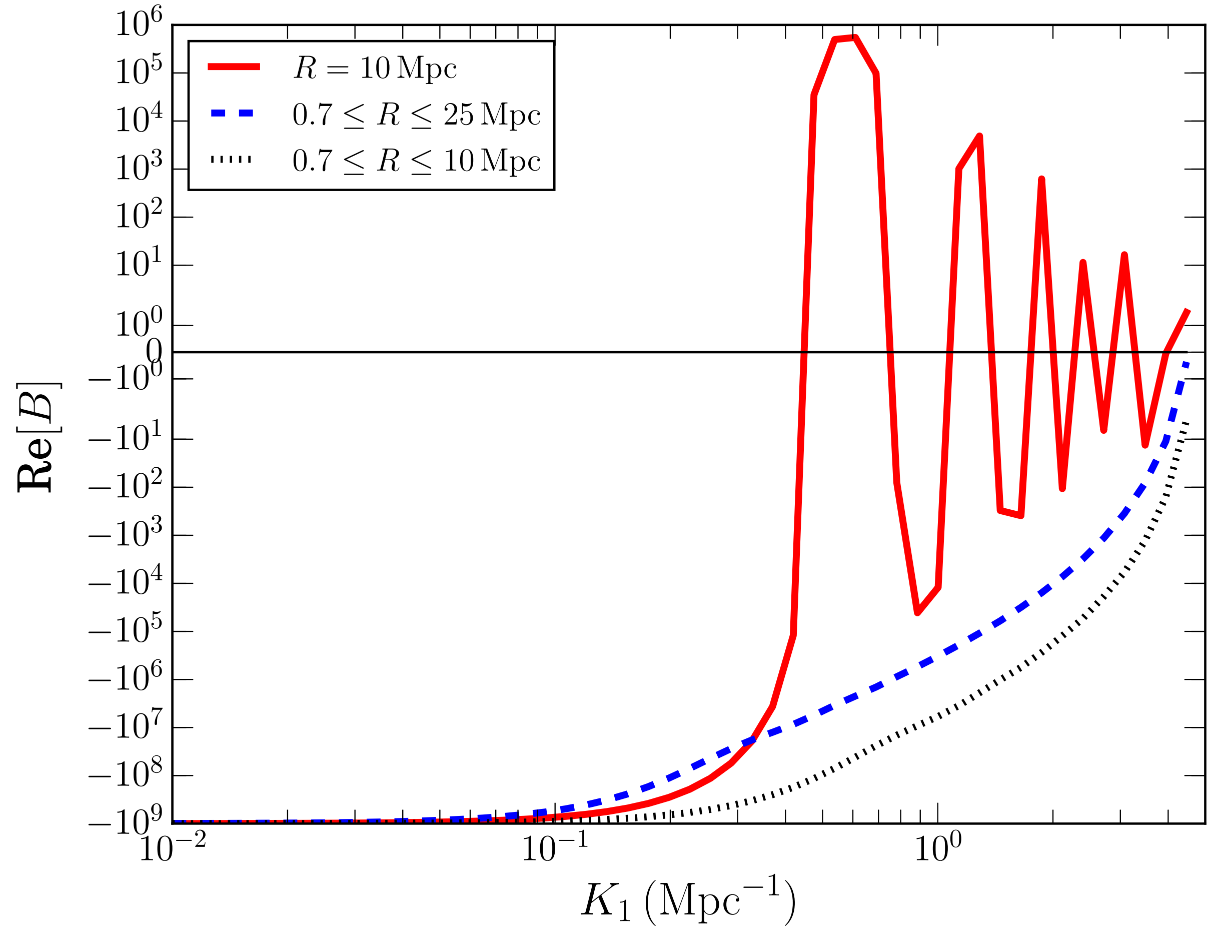
Significant increase in speed without loss of accuracy
e.g. training set ~ 10^3 samples c.f. 10^5- 10^6 in MCMC



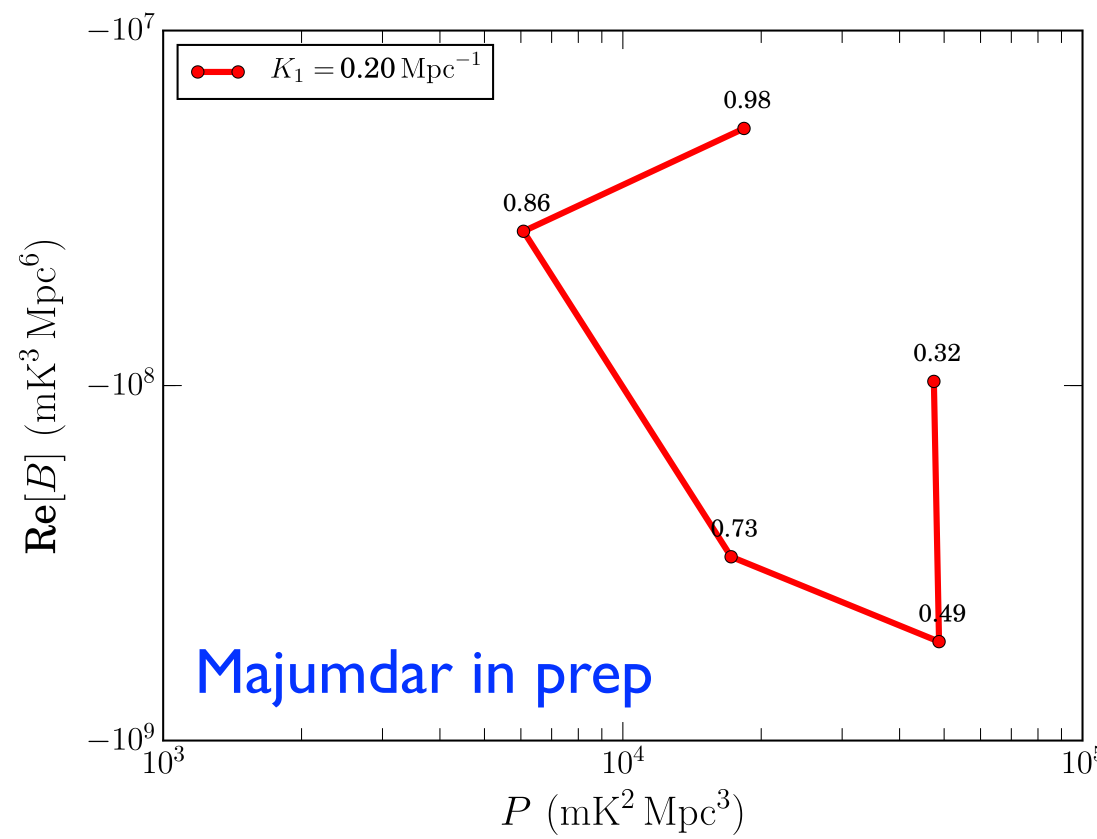
Schmit+ 2017



21cm bispectrum



Bispectrum a probe of non-Gaussianity so complementary to power spectrum



Bispectrum a probe of non-Gaussianity so complementary to power spectrum => break degeneracies

Ultimately range of statistics needed to exploit information in 21cm observations

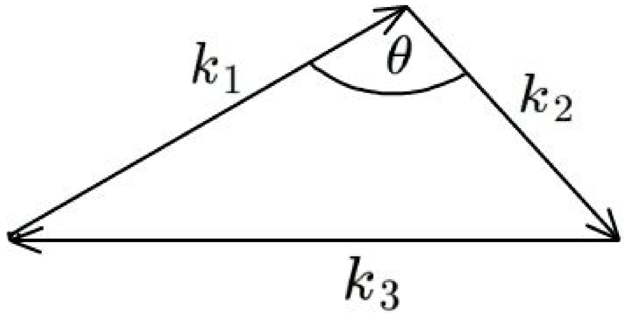
Modal bispectrum estimator

Need fast estimator for bispectrum, since large range of possible configurations

Expanding Dirac delta function in terms of exponentials allows factorisation

$$\delta^{\rm K}(\boldsymbol{m}_1+\boldsymbol{m}_2\,...+\boldsymbol{m}_p) = \frac{1}{N_{\rm pix}}\sum_{\boldsymbol{n}}^{N_{\rm pix}} {\rm e}^{i2\pi\boldsymbol{n}\cdot(\boldsymbol{m}_1+\boldsymbol{m}_2\,...+\boldsymbol{m}_p)/N_{\rm side}}$$

$$= \frac{1}{N_{\rm pix}}\sum_{\boldsymbol{n}}^{N_{\rm pix}}\prod_{i}^{p} {\rm e}^{i2\pi\boldsymbol{n}\cdot\boldsymbol{m}_i/N_{\rm side}}\,.$$



Can then express poly-spectra in terms of sampled FT of 21cm field

$$\mathscr{P}(\boldsymbol{k}_1,\boldsymbol{k}_2,...\boldsymbol{k}_p) \approx H^p\frac{1}{V}\frac{\sum_{\boldsymbol{n}}^{N_{\rm pix}}\prod_{i=1}^{p}\delta(\boldsymbol{n},\,\boldsymbol{k}_i)}{\sum_{\boldsymbol{n}}^{N_{\rm pix}}\prod_{i=1}^{p}I(\boldsymbol{n},\,\boldsymbol{k}_i)}\,,$$

$$\delta(\boldsymbol{n},\,\boldsymbol{k}_i) = \sum_{\boldsymbol{k}_i/k_{\rm F}\simeq\boldsymbol{m}_i}\Delta_{\rm FFT}(\boldsymbol{m}_i){\rm e}^{i2\pi\boldsymbol{n}\cdot\boldsymbol{m}_i/N_{\rm side}}\,,$$

$$I(\boldsymbol{n},\,\boldsymbol{k}_i) = \sum_{\boldsymbol{k}_i/k_{\rm F}\simeq\boldsymbol{m}_i}{\rm e}^{i2\pi\boldsymbol{n}\cdot\boldsymbol{m}_i/N_{\rm side}}\,,$$

For example, estimator of the bispectrum

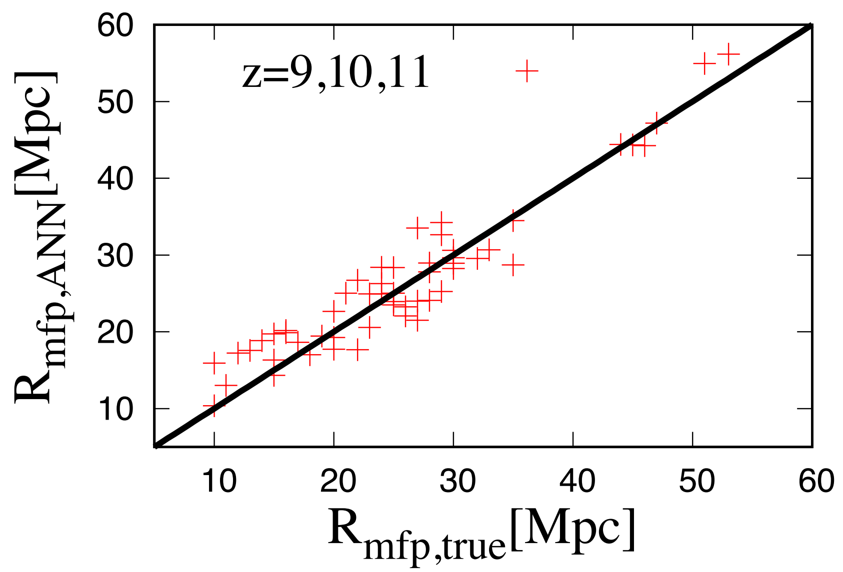
$$B(k_{\rm F}\boldsymbol{m}_1,k_{\rm F}\boldsymbol{m}_2,k_{\rm F}\boldsymbol{m}_3) \approx \frac{V^2}{N_{\rm pix}^3}\frac{\sum_{\boldsymbol{n}}^{N_{\rm pix}}\delta(\boldsymbol{n},\,\boldsymbol{k}_1)\delta(\boldsymbol{n},\,\boldsymbol{k}_2)\delta(\boldsymbol{n},\,\boldsymbol{k}_3)}{\sum_{\boldsymbol{n}}^{N_{\rm pix}}I(\boldsymbol{n},\,\boldsymbol{k}_1)I(\boldsymbol{n},\,\boldsymbol{k}_2)I(\boldsymbol{n},\,\boldsymbol{k}_3)}\,,$$

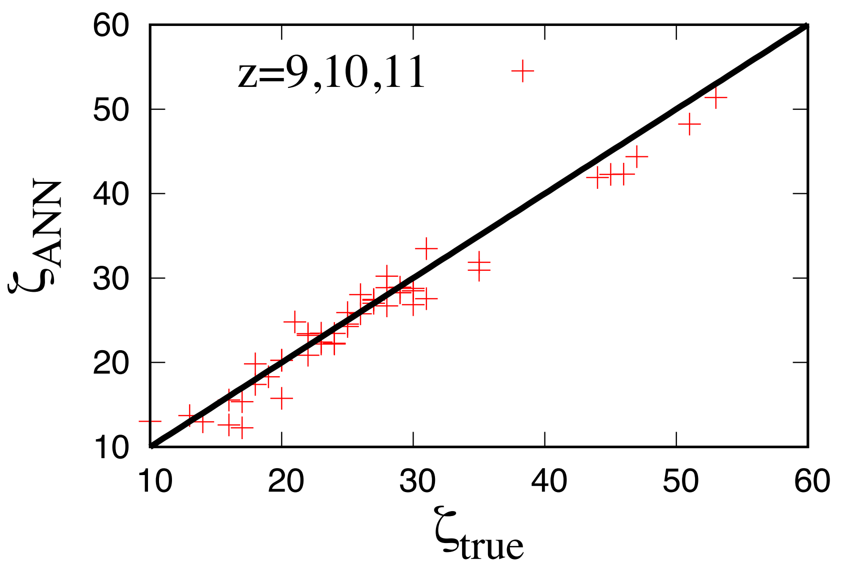
Watkinson+ 2017

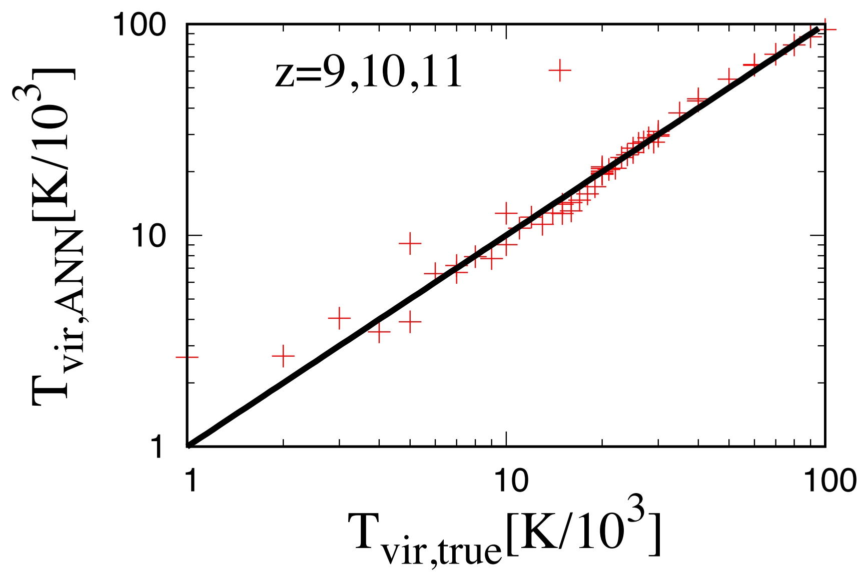
Machine learning?

Can also use ML for inversion (rather than Bayesian inference)

e.g. Neural Networks to identify astrophysics parameters from shape of power spectrum Shimabukuro & Semelin 2017



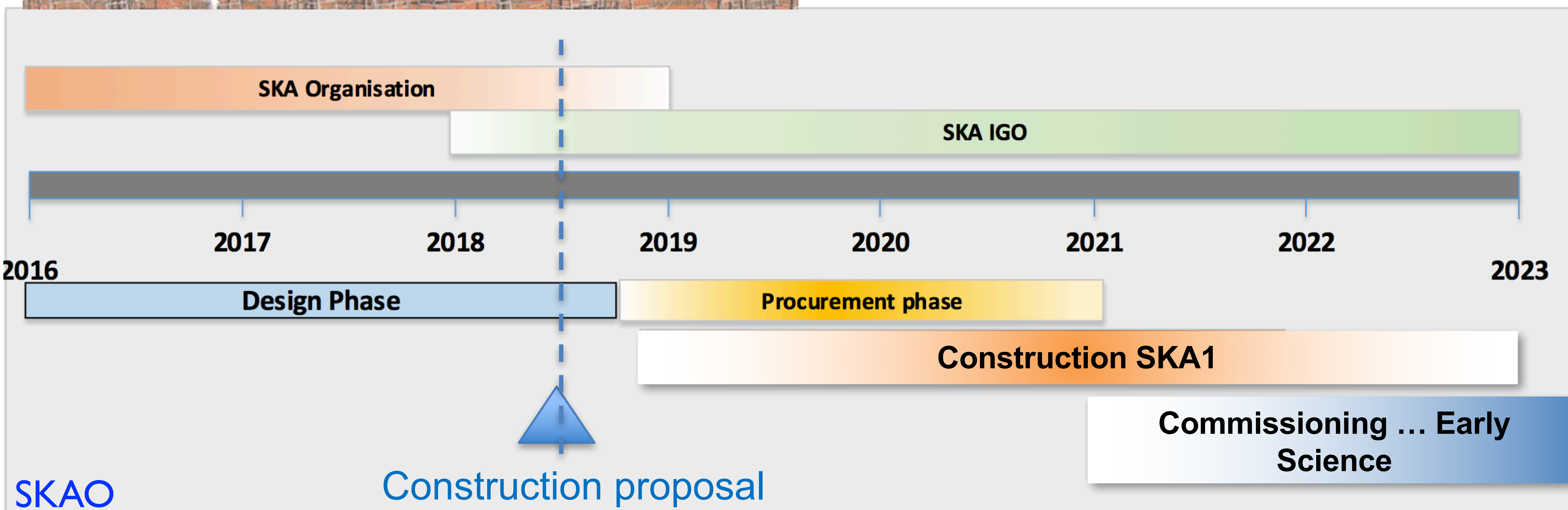




Beyond LOFAR… SKA

- MID - dishes (SA) - 250MHz - 20 GHz; LOW - antennae (AU) - 50-350MHz

- International headquarters at Jodrell Bank
- SKA will be the premier radio telescope facility with science case ranging from cradle of life, through cosmology, to epoch of reionization
- Cost control exercise currently underway
- Form IGO 2019
  Commissioning/Early science 2021-2024
  Key science projects 2024-29



Conclusions

- Reionization is still interesting! CMB, HST, and LAE surveys seem to indicate interesting behaviour at z~7
- LOFAR, MWA, PAPER are all confronting data with increasingly mature analysis pipelines. Still much to be done.
- 21 cm upper limits are beginning to make contact with plausible (if extreme) models
- Imaging will become interesting with SKA e.g. granulometry an for measuring bubble size distributions
- Power spectrum easiest statistic to work with currently
- Application of emulation may prove a useful conduit between full numerical simulations and Bayesian analysis
- Other non-Gaussian statistics also useful e.g. bispectrum. Need to be fast to evaluate to include in analysis pipelines
- Fruitful to apply machine learning techniques
- SKA finalising design with construction expected to begin ~2019

Fin

Science from 21 cm

Analysis more complicated than for, say, CMB because

1) No complete and agreed upon model of reionization
   - each simulation uses different parametrization & methods
   - different prescriptions for sources and sinks of photons
   - no guarantee that any are "correct"
   - unclear any one simulation covers all possibilities

2) Not clear simple statistics e.g. power spectrum capture all information

3) Evaluation of models is computationally expensive
   - full hydro + RT requires hundreds of hours of super-computer time

Strategies:
1) Analytic models
2) Semi-numerical codes - e.g. 21cmFAST, Simfast21, …
3) Emulation

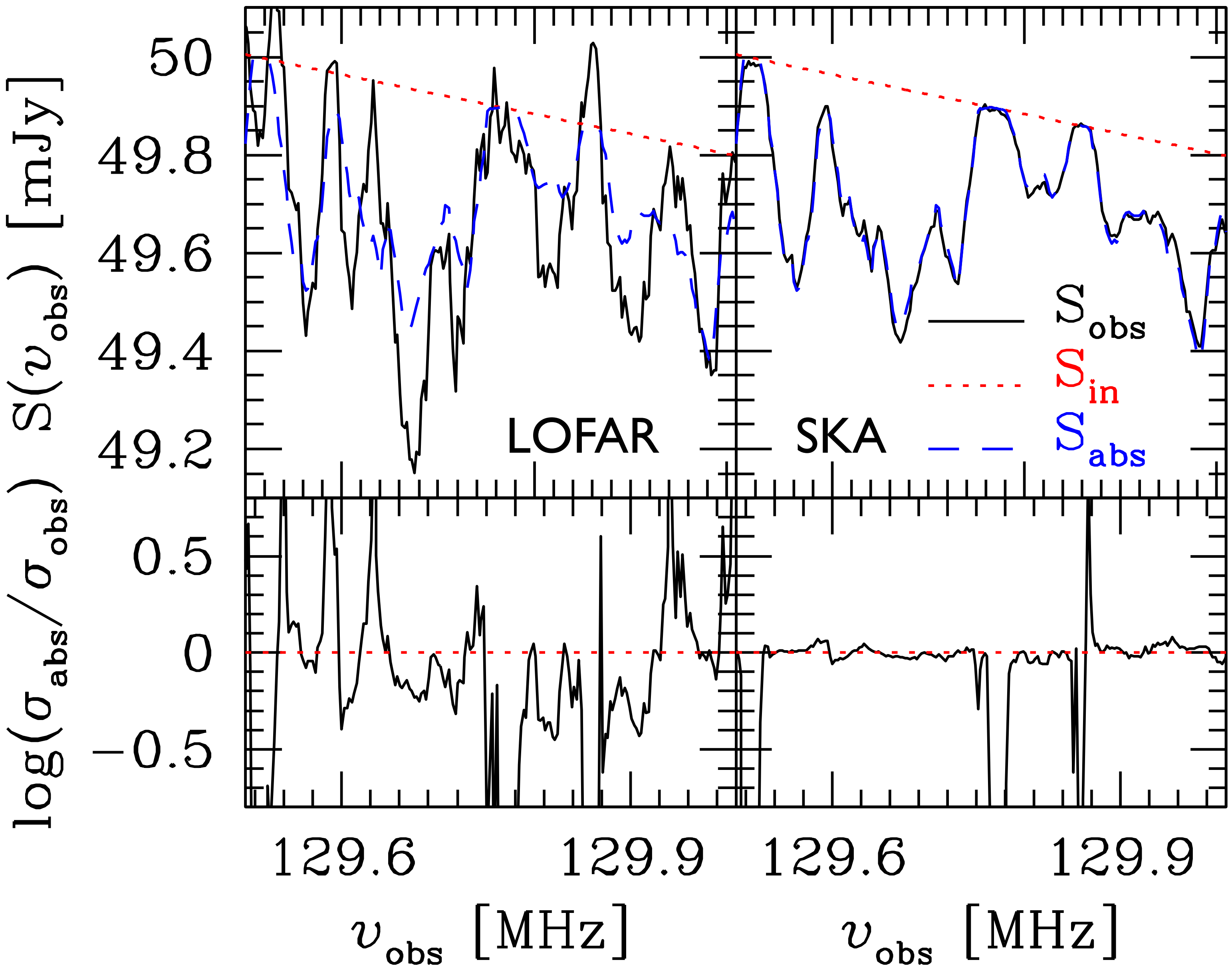
Jonathan Pritchard

(3) 21cm Forest

IF find z>6 radio bright QSOs or other radio bright objects e.g. GRB

10kHz resolved spectra of 21cm forest in bright radio sources at z>6

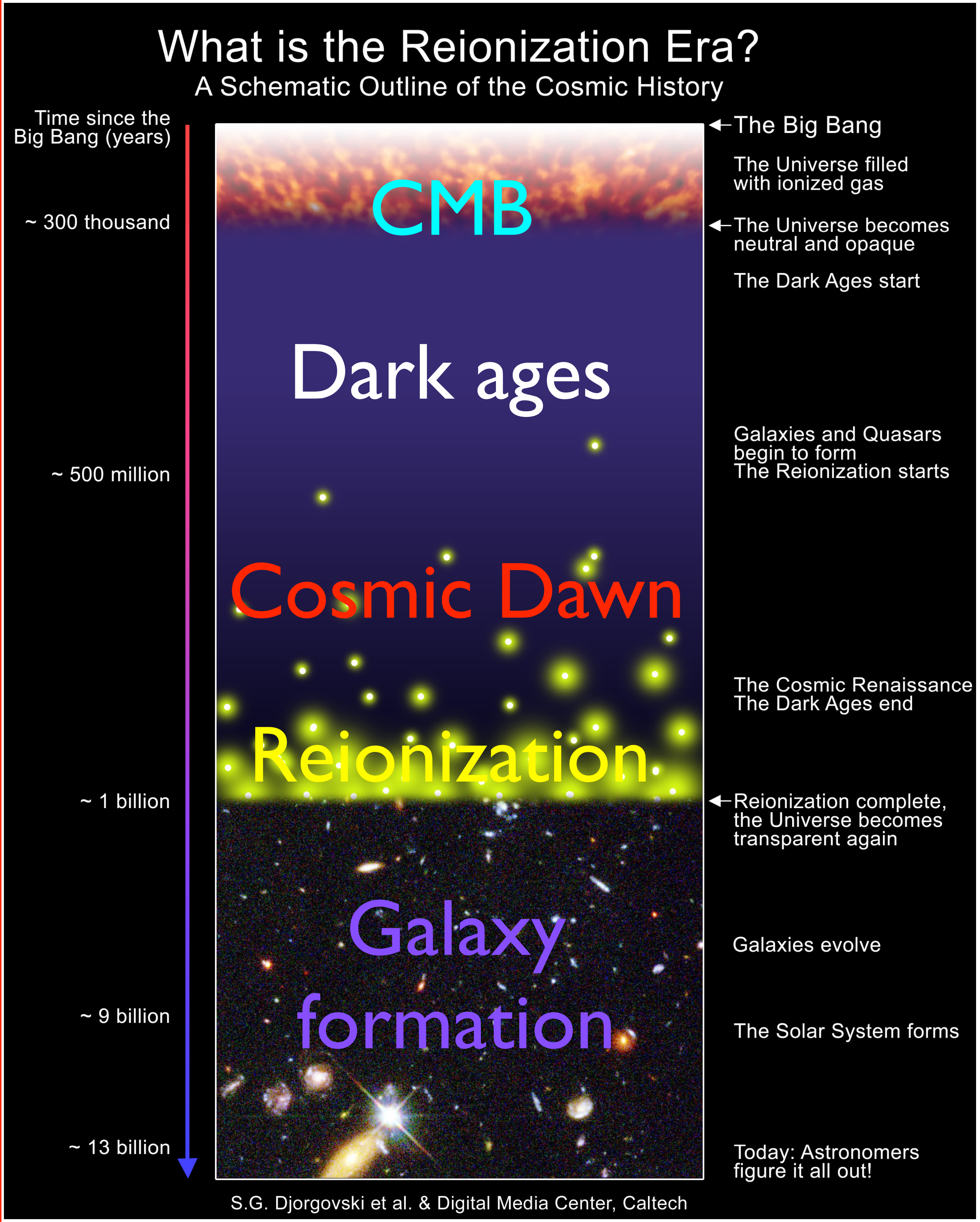
Resolve ~kpc scale structures in IGM



50mJy source @z=10

Ciardi+ 2015

Overview



- Reionization
- CMB constraints of reionization
- 21 cm observations
- CMB optical depth from 21cm

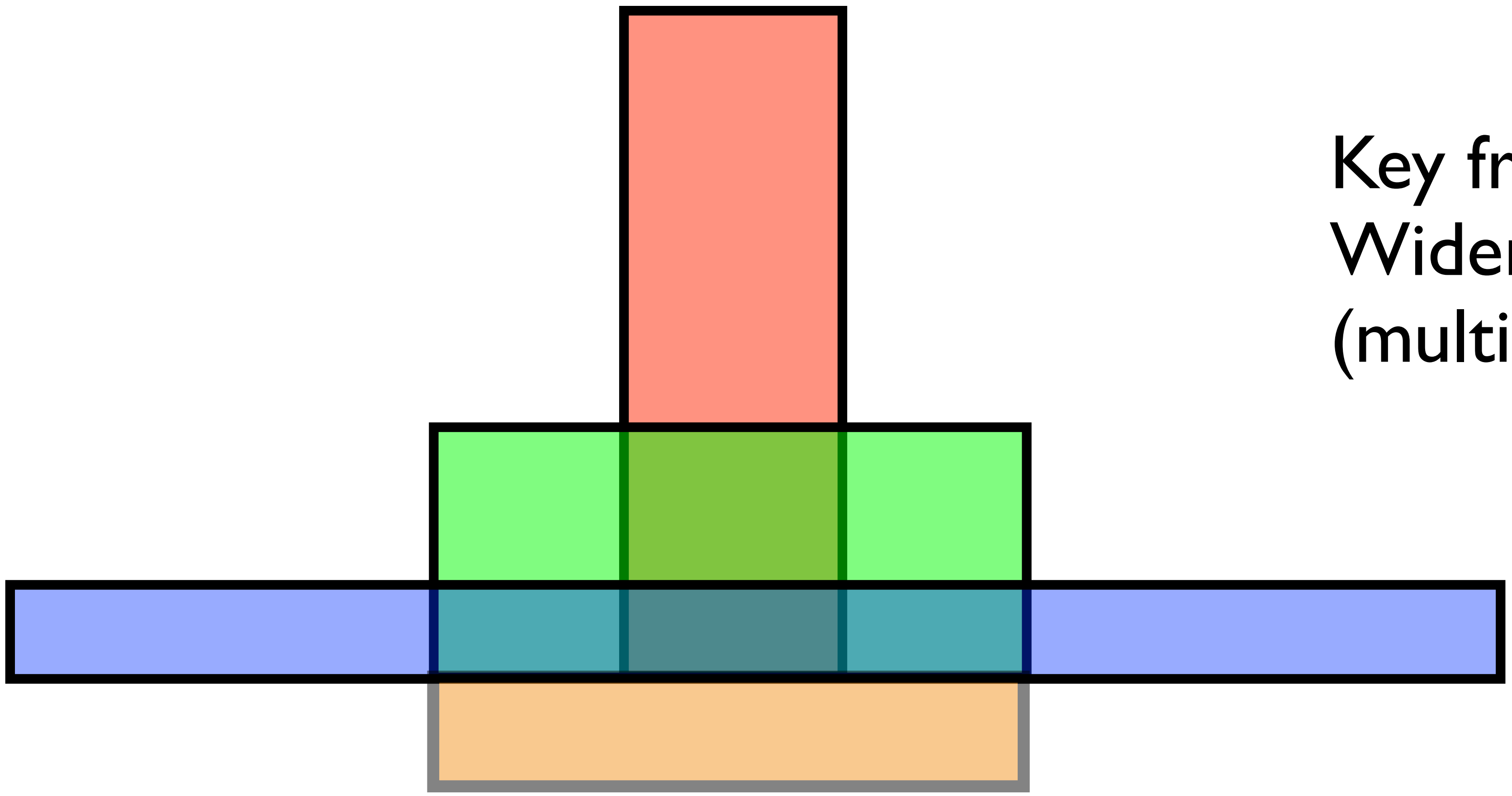
SKA-LOW observing strategy

Imperial College London

Deep: 5 x 1000hr integration => 100 deg^2 field
Middle: 50 x 100hr integration => 1,000 deg^2 field
Shallow: 500 x 10hr integration => 10,000 deg^2 field
IM: 50 x 100hr integration => 1,000 deg^2 field

Koopmans+ 2015 [arXiv1505.07568]

Single pointing ~ 20 deg^2

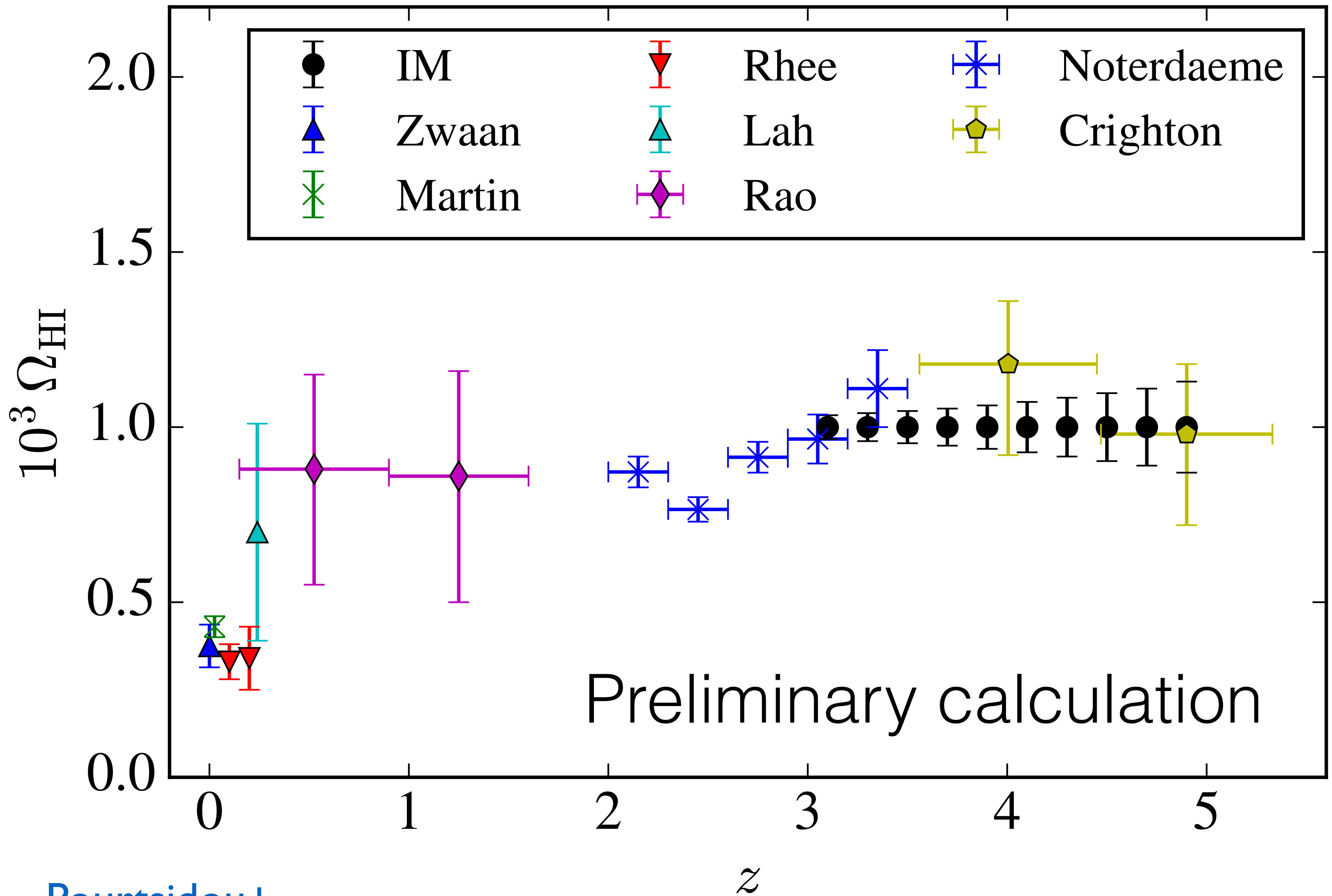


Key frequencies: 200-50MHz
Wider band for foregrounds
(multibeaming to reduce tint)

Shallow: LOFAR-like power spectrum sensitivity over 10000 deg^2.
Middle: Shallow imaging + power spectrum over 1000 deg^2
Deep: Power spectrum to $z<27$ & deep imaging over 100 deg^2
IM: OmegaHI & cosmology at $3<z<6$ over 1000 deg^2

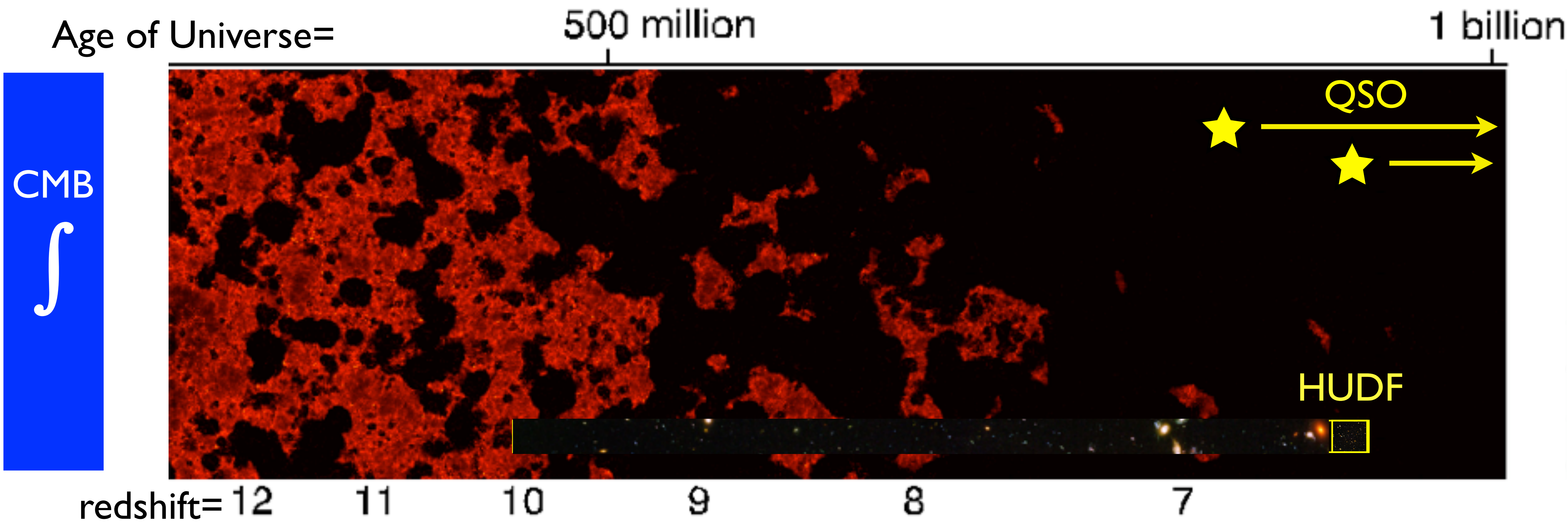
(4) Intensity mapping at 3<z<6

Intensity mapping with LOW for OmegaHI & cosmology



Pourtsidou+

More needed...



Existing observations leaves much unanswered:
1) Lyman-alpha forest: end point z>6.5
2) CMB optical depth: mid point z~11 (Revised down by Planck to z~8.8±1.5)
3) kSZ amplitude: duration z<4.4 ?

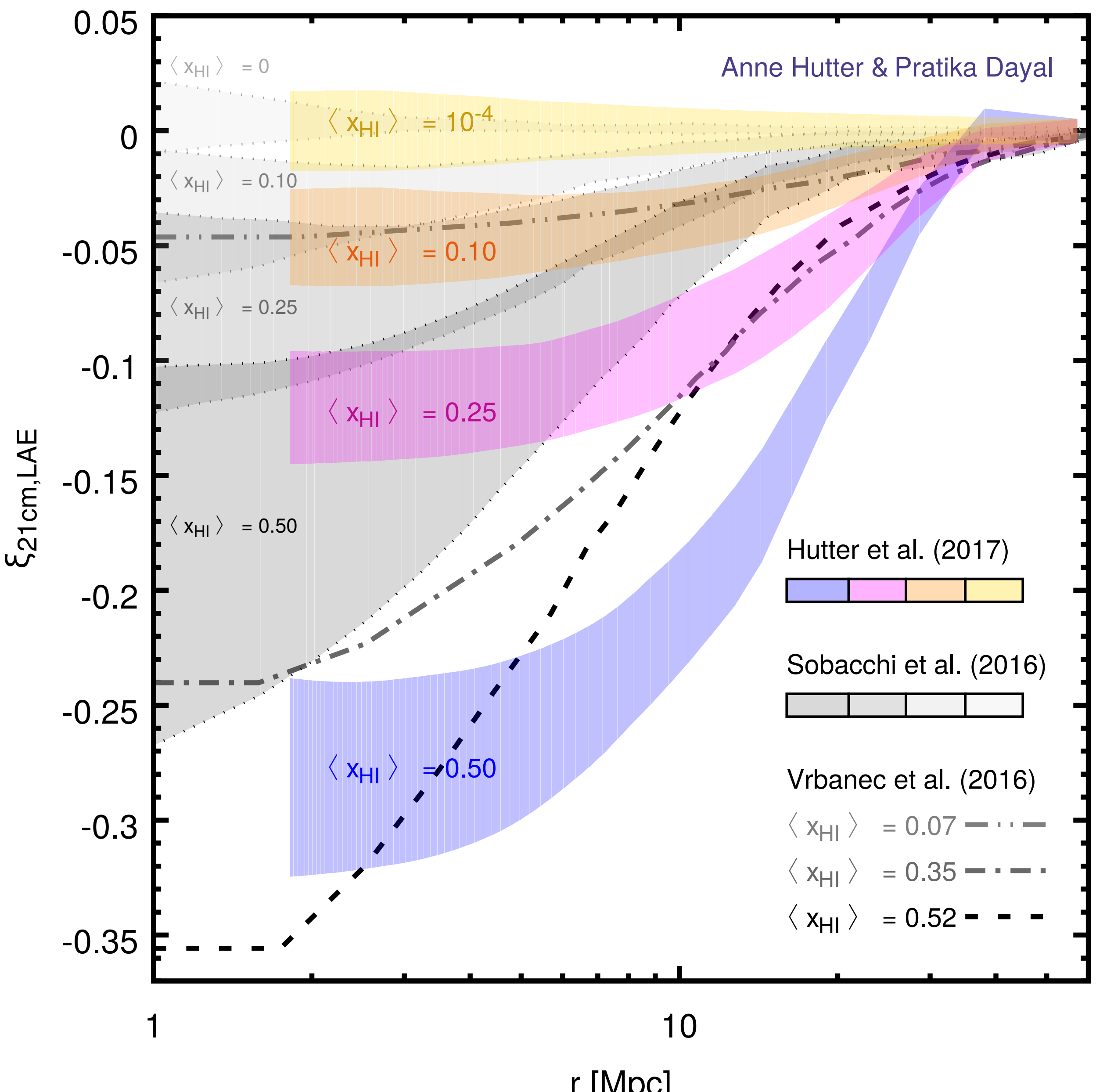
HST probes skewer much smaller than scale of ionized regions + only brightest sources

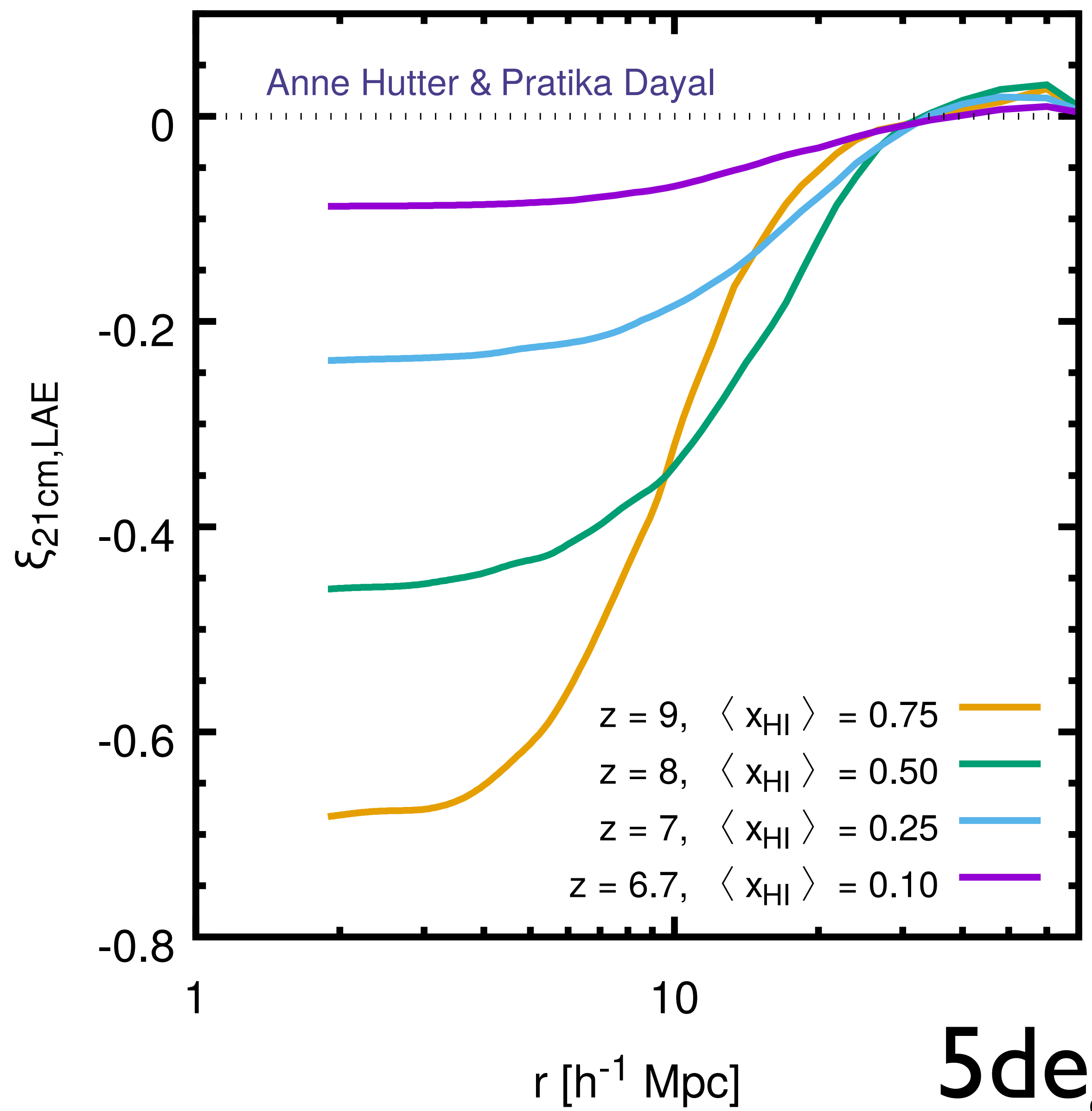
Large galaxy samples with LAE surveys or Euclid possible to z~8

Fundamental need for new types of observation to understand details of reionization

Euclid synergies?

- z=8 QSO to find those that are radio loud
- Bright galaxies for stacking
- LAE - 21cm cross-correlation as probe of HII region sizes
- ???





$5\,deg^2$

Conclusions

- SKA1 should begin commissioning/early science in 2021 with full operations in ~2024. (c.f. Euclid 2021-27)
- Taking shape now - cost control exercise, IGO, …
- EoR plans HI surveys at z = 3 - 27
  - SHALLOW - 10,000 deg^2 at LOFAR like sensitivity, z~7-15
  - MEDIUM + IM - 5000 deg^2 z ~ 3-20
  - DEEP - 100 deg^2, z~6-27
- Synergies with Euclid from both MID and LOW
  - MID better matched to low redshifts z<3 (Pourtsidou talk)
  - LOW for 3<z<27
- For reionisation:
  - LAE-21cm cross correlations; Stack on LBGs?
  - High-z quasars for radio follow up
  - OmegaHI
  - ???
- SKA-EoR Science Team preparing for KSP bid in ~2019
  Synergy focus group leads: Pratika Dayal, Erik Zackrisson

SKA EoR Science Team

Leon Koopmans (NL, Chair ST), Gianni Bernardi (SA), Garrelt Mellema (SE), Andrei Mesinger (IT), Jonathan Pritchard (UK, Chair SWG), Cath Trott (AU) Abhi Datta (IN)

Various focus groups studying aspects of doing 21cm with SKA

https://sites.google.com/site/skacdeorscienceteam/home

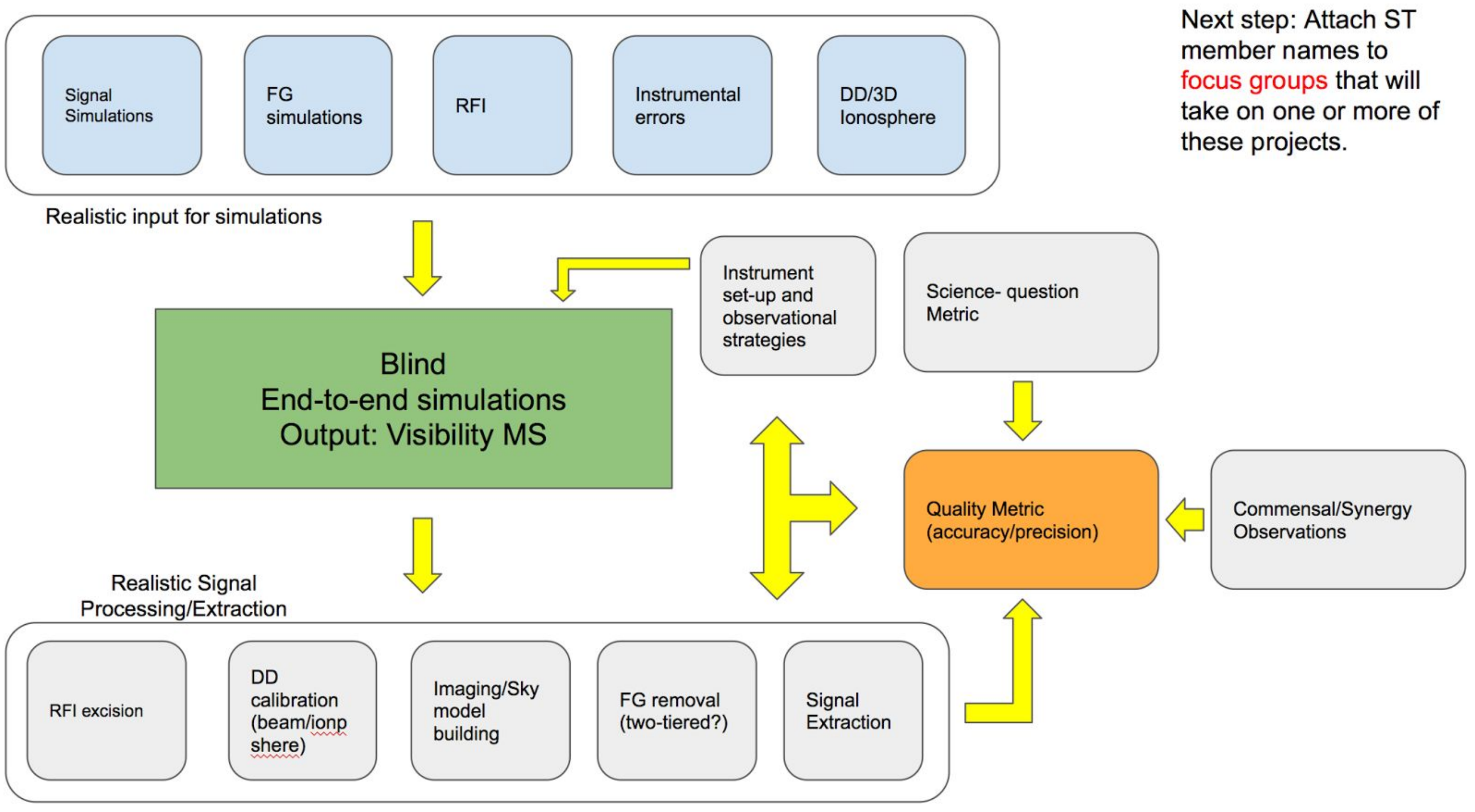


Figure 1: Flow-diagram for End-to-End data simulations.

Focus on pipelines at present

KSP bid in ~2019

Interest in synergy

Contacts:

Pratika Dayal

Erik Zackrisson

Other SKA-Euclid Synergies

- Cosmology (MID)
  - galaxy surveys - SKA1: Continuum, HI IM; SKA2: HI
  - weak lensing with SKA2 - different systematics (think 5000 deg^2 survey)

- Reionization (LOW)
  - 21cm - LAE cross correlation
  - stack on bright galaxies at high-z
  - Radio loud quasars at z>6 (especially z~8)
  - HI intensity mapping at 3<z<6

- General

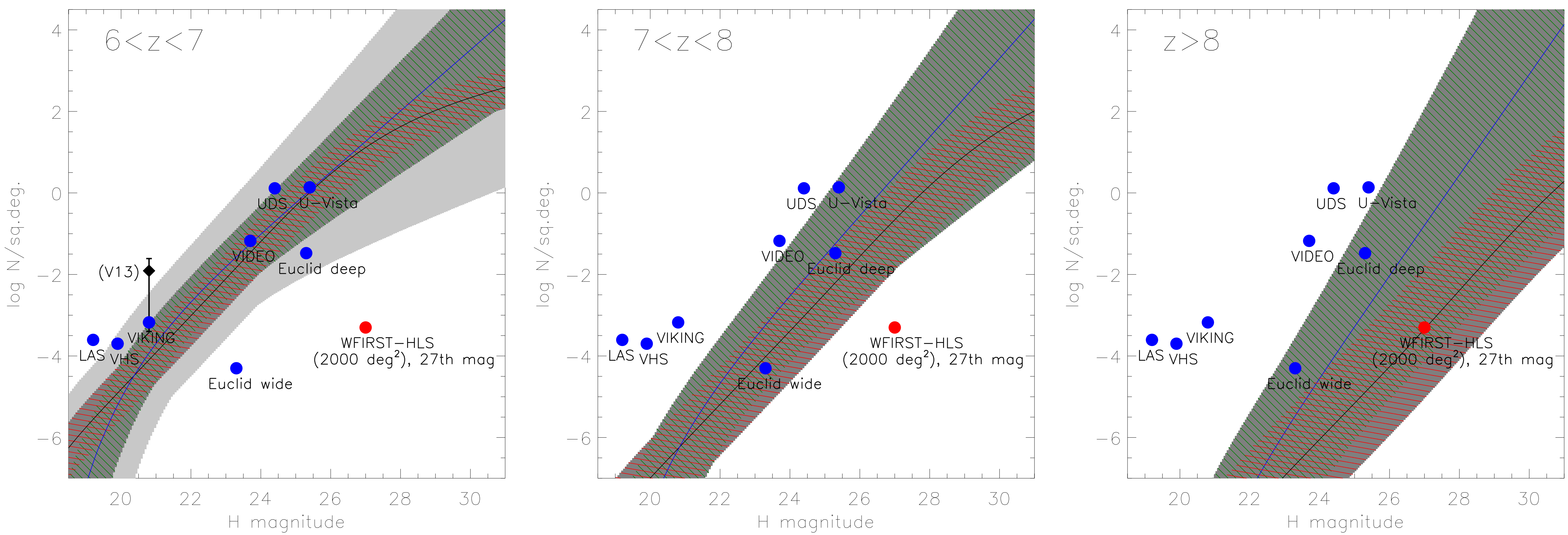
  LOW (50-350 MHz),

  MID - Band 1 (580-1015MHz), Band 2 (0.9 - 1.67 GHz),
  Band 5 (8-13.8GHz)

z~8 quasars

21cm Forest requires radio bright sources at z>7
- QSO
- GRB

Euclid will be key for providing z~8 QSOs (hopefully some radio loud)



Manti+ 2017