Reionization and Cosmic Dawn: *current status and future outlook*

Andrei Mesinger







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?

Outline

- What we know now...
 - Clues to the timing of reionization from galaxies,
 QSOs and the CMB
- What we will know soon...
 - The full picture from the cosmic 21 cm signal!

When?

- Two main classes of probes
 - 1. Integral CMB constraints (e.g. $\tau_{\rm e}$, kinetic SZ)



History of Thompson scattering optical depth measurements



Planck 2016

History of Thompson scattering optical depth measurements



2020 – negative tau: Reionization never happened!

Planck 2016

History of Thompson scattering optical depth measurements



Planck 2016

What does this tell us about *when* reionization occurred?



Greig & AM 2016

When?

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

Astrophysical flashlights: Ly α

Post-reionization IGM



We can't directly observe the EoR in $\mbox{Ly}\alpha$



4

Ly α forest saturates at z>5, when the Universe becomes too dense. Even trace amounts of HI, x_{HI} ~> 10⁻⁵ result in no flux being detected in the forest.



Dijkstra 2014



e.g. Dijkstra, AM+2011

Lyman alpha line emerging from galaxies is shaped by the ISM/CGM (winds, infall, dust, geometry..)





during reionization, cosmic HI patches absorb Ly α photons in the damping wing of the line





during reionization, cosmic HI patches absorb Ly α photons in the damping wing of the line



The EoR modulates:

- 1. the observability of Lyman alpha emission
- 2. the observed clustering of $\mbox{Ly}\alpha$ emitting galaxies

an example: the first detection of ongoing EoR from a z=7.1 QSO

QSOs: the brightest cosmic flashlights



QSO spectra can be analyzed *individually*, unlike galaxies which require a statistically significant sample

figure courtesy of D. Mortlock

 Caution: We must jointly sample the uncertainties in the intrinsic (pre IGM absorption) QSO emission together with the sightline to sightline scatter of the EoR

Simcoe+2012



 Caution: We must jointly sample the uncertainties in the intrinsic (pre IGM absorption) QSO emission together with the sightline to sightline scatter of the EoR



 Caution: We must jointly sample the uncertainties in the intrinsic (pre IGM absorption) QSO emission together with the sightline to sightline scatter of the EoR



 Caution: We must jointly sample the uncertainties in the intrinsic (pre IGM absorption) QSO emission together with the sightline to sightline scatter of the EoR



 Step 1: reconstruct the intrinsic Lyα emission of ULASJ1120 by sampling a covariance matrix of emission line properties built from ~1700 high S/ N BOSS spectra (Greig, AM+ 2016a)

1



- <u>Step 1:</u> reconstruct the intrinsic Lyα emission of ULASJ1120 by sampling a covariance matrix of emission line properties built from ~1700 high S/ N BOSS spectra (Greig, AM+ 2016a)
- <u>Step 2</u>: run large-scale, state-of-the-art simulations of reionization, spanning a range of uncertainties in the EoR topology (AM+ 2016)

Small HII

INTERMEDIATE HII

Large HII



- Step 1: reconstruct the intrinsic Lyα emission of ULASJ1120 by sampling a covariance matrix of emission line properties built from ~1700 high S/ N BOSS spectra (Greig, AM+ 2016a)
- <u>Step 2</u>: run large-scale, state-of-the-art simulations of reionization, spanning a range of uncertainties in the EoR topology (AM+ 2016)
- <u>Step 3:</u> Simultaneously sample intrinsic emission
 + IGM absorption, in a Bayesian framework (Greig, AM+ 2016b)



Greig, AM+ 2016b



First detection of ongoing reionization!!!

 $< x_{HI} > = 0.40_{-0.32}^{+0.41} (2 \text{ G})$

Greig, AM+ 2016b

Subsequent, similar analysis of z=7.1 and new z=7.5 QSO



Banados+2018

current state of knowledge: When did the Universe reionize?



We now have a reasonable handle on when...

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

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





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



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:

$$\delta \mathsf{T}_b(\nu) \approx 27 \mathsf{x}_{\rm HI} (1 + \delta_{\rm nl}) \left(\frac{\mathsf{H}}{\mathsf{d} \mathsf{v}_r/\mathsf{d} \mathsf{r} + \mathsf{H}} \right) \left(1 - \frac{\mathsf{T}_\gamma}{\mathsf{T}_{\rm S}} \right) \left(\frac{1 + \mathsf{z}}{10} \frac{0.15}{\Omega_{\rm M} \mathsf{h}^2} \right)^{1/2} \left(\frac{\Omega_b \mathsf{h}^2}{0.023} \right) \mathrm{mK}$$

Cosmic 21-cm signal



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:

$$\delta \mathsf{T}_{b}(\nu) \approx 27 \mathsf{X}_{\mathrm{HI}} 1 + \delta_{\mathrm{nl}} \left(\frac{\mathsf{H}}{\mathsf{d} \mathsf{v}_{r}/\mathsf{d} \mathsf{r} + \mathsf{H}} \right) \left(1 - \frac{\mathsf{T}_{\mathrm{s}}}{\mathsf{T}_{\mathrm{S}}} \right) \left(\frac{1 + \mathsf{z}}{10} \frac{0.15}{\Omega_{\mathrm{M}} \mathsf{h}^{2}} \right)^{1/2} \left(\frac{\Omega_{b} \mathsf{h}^{2}}{0.023} \right) \mathrm{mK}$$

Signal contains both **ASTROPHYSICAL**

Cosmic 21-cm signal



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.

$$\delta \mathsf{T}_{b}(\nu) \approx 27 \mathsf{x}_{\mathrm{HI}} (1 + \delta_{\mathrm{nl}}) \left(\frac{\mathsf{H}}{\mathsf{d} \mathsf{v}_{r}/\mathsf{d} \mathsf{r} + \mathsf{H}} \right) \left(1 - \frac{\mathsf{T}_{\gamma}}{\mathsf{T}_{\mathrm{S}}} \right) \left(\frac{1 + \mathsf{z}}{10} \frac{0.15}{\Omega_{\mathrm{M}} \mathsf{h}^{2}} \right)^{1/2} \left(\frac{\Omega_{b} \mathsf{h}^{2}}{0.023} \right) \mathrm{mK}$$

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

Cosmic 21-cm signal



$$\delta \mathsf{T}_{b}(\nu) \approx 27 \mathsf{x}_{\mathrm{HI}}(1+\delta_{\mathrm{nl}}) \left(\frac{\mathsf{H}}{\mathsf{d} \mathsf{v}_{r}/\mathsf{d} \mathsf{r}+\mathsf{H}}\right) \left(1-\frac{\mathsf{T}_{\gamma}}{\mathsf{T}_{\mathrm{S}}}\right) \left(\frac{1+\mathsf{z}}{10}\frac{0.15}{\Omega_{\mathrm{M}}\mathsf{h}^{2}}\right)^{1/2} \left(\frac{\Omega_{b}\mathsf{h}^{2}}{0.023}\right) \mathrm{mK}$$

Cosmic 21-cm signal



- **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!

So how do we learn about the unseen first galaxies?

Reionization signal

$$\delta \mathsf{T}_{b}(\nu) \approx 27 \mathsf{x}_{\mathrm{HI}}(1+\delta_{\mathrm{nl}}) \left(\frac{\mathsf{H}}{\mathsf{d}\mathsf{v}_{r}/\mathsf{d}\mathsf{r}+\mathsf{H}}\right) \left(1-\frac{\mathsf{T}_{\gamma}}{\mathsf{T}_{\mathrm{S}}}\right) \left(\frac{1+\mathsf{z}}{10}\frac{0.15}{\Omega_{\mathrm{M}}\mathsf{h}^{2}}\right)^{1/2} \left(\frac{\Omega_{b}\mathsf{h}^{2}}{0.023}\right) \mathrm{mK}$$

 Short mean free path + long recombination times of mean IGM → reionization is bimodal (~fully ionized and ~fully neutral patches)

Its all about timing and the patterns!

• Galaxy clustering + stellar properties \rightarrow evolution of *large-scale EoR/CD structures*





94 Mpc





McQuinn+ 2007

Abundant, faint galaxies vs Rare, bright galaxies

Pictures are nice, but we need numbers

• Common/simple statistic: power spectrum during EoR

LOFAR HERA331 SKA1 = 10⁴ K, ↓₀ = 12.8 $10^4 \text{ K}, \downarrow_0 = 30.0$ $= 10^5 \text{ K}, \downarrow_0 = 108.0$ $R_{\rm mfp} = 30 \,{\rm Mpc}, \, \downarrow_0 = 41.0$ 10-1 10⁰ k (Mpc⁻¹)

PS at the same mean neutral fraction

Power (mK²)

Greig & Mesinger (2015)

Pictures are nice, but we need numbers

• Common/simple statistic: power spectrum during EoR



Greig & Mesinger (2015)

Power (mK²)

Cosmic Dawn

The full power of 21cm is to reach back into the infancy of galaxy formation beyond other probes....

Pre-reionization signal

$$\delta \mathsf{T}_{b}(\nu) \approx 27 \mathsf{x}_{\mathrm{HI}}(1+\delta_{\mathrm{nl}}) \left(\frac{\mathsf{H}}{\mathsf{d}\mathsf{v}_{r}/\mathsf{d}\mathsf{r}+\mathsf{H}}\right) \left(1-\frac{\mathsf{T}_{\gamma}}{\mathsf{T}_{\mathrm{S}}}\right) \left(\frac{1+\mathsf{z}}{10}\frac{0.15}{\Omega_{\mathrm{M}}\mathsf{h}^{2}}\right)^{1/2} \left(\frac{\Omega_{b}\mathsf{h}^{2}}{0.023}\right) \mathrm{mK}$$

spin temperature

defined in terms of the ratio of the number densities of electrons occupying the two hyperfine levels:

 $n_1/n_0 = 3 \exp[-0.068 \text{ K} / \text{T}_s]$

Pre-reionization signal

$$\delta \mathsf{T}_{b}(\nu) \approx 27 \mathsf{x}_{\mathrm{HI}}(1+\delta_{\mathrm{nl}}) \left(\frac{\mathsf{H}}{\mathsf{d}\mathsf{v}_{r}/\mathsf{d}\mathsf{r}+\mathsf{H}}\right) \left(1-\frac{\mathsf{T}_{\gamma}}{\mathsf{T}_{\mathrm{S}}}\right) \left(\frac{1+\mathsf{z}}{10}\frac{0.15}{\Omega_{\mathrm{M}}\mathsf{h}^{2}}\right)^{1/2} \left(\frac{\Omega_{b}\mathsf{h}^{2}}{0.023}\right) \mathrm{mK}$$

spin temperature

cosmologically:

$$T_{\rm S}^{-1} = \frac{T_{\gamma}^{-1} + x_{\alpha}T_{\alpha}^{-1} + x_{c}T_{\rm K}^{-1}}{1 + x_{\alpha} + x_{c}}$$
$$T_{\gamma} - temperature of the CMB$$
$$T_{K} - gas kinetic temperature$$
$$T_{\alpha} - color temperature \sim T_{K}$$

the spin temperature interpolates between T_{γ} and T_K

The spin temperature interpolates between T_{γ} and T_{K}

$$T_{\rm S}^{-1} = \frac{T_{\gamma}^{-1} + x_{\alpha}T_{\alpha}^{-1} + x_{c}T_{\rm K}^{-1}}{1 + x_{\alpha} + x_{c}}$$

two coupling coefficients:

$$x_{c} = \frac{0.0628 \text{ K}}{A_{10}T_{\gamma}} \left[n_{\rm HI} \kappa_{1-0}^{\rm HH}(T_{\rm K}) + n_{e} \kappa_{1-0}^{\rm eH}(T_{\rm K}) + n_{\rm p} \kappa_{1-0}^{\rm pH}(T_{\rm K}) \right]$$

collisional coupling requires high densities effective in the IGM at z>40

$$x_{\alpha} = 1.7 \times 10^{11} (1+z)^{-1} S_{\alpha} J_{\alpha}$$

Wouthuysen-Field (WF)

uses the Lya background effective soon after the first sources ignite

The spin temperature approaches the kinetic temperature if either coefficient is high. Otherwise, the spin temperature approaches the CMB temperature: NO SIGNAL!

"Fiducial" scenario: the IGM is heated by X-rays from HMXBs before reionization

emission absorption 10³ collisional coupling $T_{k} \simeq T_{S} > T_{\gamma}$ $T_k = T_s \leq T_{\gamma}$ $T_{\nu} < T_{s} \simeq T_{\gamma}$ <u>۲</u> 100 _<**T**_<**T**_ IE T_k<Ts<T, $T_{k} \simeq T_{s} < T$ WF coupling 10 10 100 \mathbf{Z}

SFR required to reionize the Universe:

$$[\dot{\rho}_{\rm SFR}]_{\rm ion} = 4.4 \times 10^{-1} \,\bar{x}_i \, Z_{20}^{3/2} \, \left(\frac{t_{\rm SFR}}{0.1 \, t_H}\right)^{-1} \\ \times \left(\frac{f_{\rm esc}}{0.1 \, 4000}\right)^{-1} \,\,\mathrm{M_{\odot} \ yr^{-1} \ Mpc^{-3}, (11)}$$

is much larger than the SFR needed to heat it with HMXBs:

$$\begin{split} [\dot{\rho}_{\rm SFR}]_{\rm X} &= 4.0 \times 10^{-2} \, Z_{20}^{5/2} \, \left(\frac{t_{\rm SFR}}{0.1 \, t_H}\right)^{-1} \, \left(\frac{f_X}{0.2}\right)^{-1} \quad (10) \\ & \times \left(\frac{L_X / {\rm SFR}}{10^{40} \, {\rm erg \, s^{-1} \, M_{\odot}}^{-1} \, {\rm yr}}\right)^{-1} \, \, {\rm M}_{\odot} \, {\rm yr}^{-1} \, {\rm Mpc}^{-3}, \end{split}$$

McQuinn & O'Leary (2012) see also Furlanetto (2006), Mesinger+ (2013)

The timing of the Epoch of Heating (EoH)

Just the redshift at which the 21-cm power spectrum peaks already tells us a combination of which halos hosted early galaxies and what were their X-ray luminosities



fx=1 corresponds to Lx/SFR from local star-forming galaxies (Mineo+2012)
is Lx/SFR different in the first galaxies?? e.g. lower metallicity -> more abundant and more
luminous HMXBs? (Fragos+2012; Basu-Zych+2016)

But we can learn more!

The X-ray mean free path through the IGM is very sensitive to E_x

$$\lambda_{\rm X} \approx 34 \, \bar{x}_{\rm HI}^{-1} \left(\frac{E_{\rm X}}{0.5 \, \rm keV}\right)^{2.6} \left(\frac{1+z}{15}\right)^{-2} \text{ comoving Mpc} \,,$$

Note that we only care about $E_x < 2$ keV. Higher energy photons a have mean free path longer than the Hubble length.

Some candidate sources of X-rays

hot ISM 10-4 HMXBs solved HMXBs, α =0.8 Photons cm⁻² s⁻¹ keV⁻¹ hot ISM, α =3 Pacucci, AM+ 2014 10-5 unresolved point sources based on observations in Mineo+2012a,b 10-6 2 Energy (keV)

Luminosities of both hot ISM (soft) and HMXB (hard) scale with SFR

Composite SEDs of local, star-forming galaxies



Softer SEDs result in more inhomogeous IGM heating

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



differences are easily detectable with HERA and the SKA

Pacucci+ 2014

More inhomogeneous heating means a higher 21cm power spectrum



Pacucci, AM+ (2014)

More inhomogeneous heating means a higher 21cm power spectrum



Pacucci, AM+ (2014)

Exotic sources of heat could also impact the signal

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



Cosmic Evolution:

putting it together



and we should get a high S/N detection with HERA and SKA



Mesinger+ (2016)

How to quantify what we will learn??





We need *fast, accurate* sims to *forward-model* the 21-cm signal

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).



hydro+RT (Trac+2009): ~10⁷ CPU hours



21cmFAST: ~0.1 CPU hours

So, how does it work then?

Start with a galaxy model: a flexible approach based on DM halos + galaxy LFs

Average properties of galaxies in halos of mass M_h:

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

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:



An flexible approach based on DM halos + galaxy LFs



X-ray free parameters characterizing emerging SED from galaxies

Das, AM+ 2016

Free parameters



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)



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


characterize in terms of a summary statistic:

power spectra with 1000h noise from HERA and moderate foreground contamination







combine with other observations in order to compute **likelihood**







Parameter constraints: LF + 21cm



Parameter constraints: LF + 21cm



Parameter constraints: LF + 21cm



Including physical cosmology



Including physical cosmology



Kern + (2017)

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



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

Astrophysical cosmology



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 +, 2019

Deep learning with CNN: parameter recovery

89





CNN's eyes

- Filters of the 1st Convolution Layer
- 10x10 pixels (gaussian smooth)





What the CNN sees

z~6



Gillet, AM +, 2019

z~30

What the CNN sees



Gillet, AM +, 2019

z~30

The time is now!

HERA (an SKA, second-generation precursor) will be fully operational in 2020SKA will roll out early 2020s, with science runs mid-decade



HERA in 2019

rendering of SKA1-Low