Andrei Mesinger

ongoing activities towards the SKA

Cosmic Dawn and Reionization
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?
We now have a reasonable handle on when...

When did the Universe Reionize?  

current state of knowledge:

Greig & AM (2017)
see also Planck 2016, Price+2016; Mitra+2016
we don’t really know...

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

What and how?
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 photons escaping. Bouwens+ (2015)
Here there be monsters!

Hidden population of abundant, faint galaxies?
Get ready for the revolution: the cosmic 21 cm signal.
Hyperfine transition in the ground state of neutral hydrogen produces the 21 cm line.
Cosmological terms

Signal contains both astrophysical and cosmological terms

$$\left( \frac{0.023}{h^2} \right)^{1/2} \left( \frac{1.0}{1 + \frac{0.15}{z+1}} \right) \left( 1 - I \right) \left( \frac{H + \frac{dr}{d\psi}}{H} \right) \left( 1 + \frac{\varphi_0}{\rho} \right) \approx (\pi)^{9}(\eta)^{9}$$

Recombination

CMB backlight

$z \approx 1000$

HI

Signal contains both astrophysical and cosmological terms

CMB backlight

$z \approx 1000$

HI

Cosmic 21-cm signal

$0 = z$

SKA (202x)
Cosmological Terms

Signal contains both Astrophysical and Cosmological terms

\[
\text{Signal} \approx \sum \text{Astrophysical} + \sum \text{Cosmological}.
\]

\[
\left( \frac{0.023}{W}\frac{1}{z} \right)^2 \left( \frac{0.15}{0.15} \right) \left( \frac{S}{\nu} - 1 \right) \left( \frac{d^2 \nu}{dt^2} \right) \left( \frac{H + \nu}{H} \right) \left( \frac{1}{27} \right) \approx \left( \frac{\nu}{T} \right)^8.
\]

Dark Ages
Lyα coupling
X-ray heating
Reionization

Cosmic 21-cm Signal
Cosmic 21-cm signal

Cosmological terms

Signal contains both astrophysical and cosmological terms

Big Data Revolution

- 3D signal with > 10 orders of magnitude more independent modes than in the CMB!
- Even the narrowest fields will contain billions of unseen galaxies!
- Data collection with upcoming Square Kilometre Array (SKA) will surpass 10x current global internet traffic!

AM+ 2016

Cosmic 21-cm signal
So how do we learn about the unseen first galaxies?
Abundant, faint galaxies vs. rare, bright galaxies

McQuinn+ 2007

Galaxy clustering + stellar properties

Large-scale EoR/CD structures

Evolution of galaxy clustering + stellar properties

It's all in the patterns!
Patterns in the Epoch of Heating

Pacucci+ 2014

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.

\[ \Delta T \text{ [mK]} \]

750 Mpc

'Soft' SED ~ Hot ISM

'Hard' SED ~ HMXBs
How to quantify what we will learn??
Sinks

Sources

\sim R_{\text{sun}}\sim \text{kpc}

\sim \text{HI} \sim \text{kpc}

\sim R_{\odot}

~R_{\odot}

\sim \text{Mpc}

\sim \text{Gpc}

Early Universe Simulations

Scales of

2 \text{ Gpc}

94 \text{ Mpc}

5 \text{ Mpc}
Simulations Early Universe
Scales of Astrophysical (known) unknowns

Sources

Sinks

~R_{sun}

N

HI

C

f esc

f

n_{rec}

f^{*}_{SN}

f^{*}_{rec}

\gamma

b

f

\epsilon

SN

rec

\frac{N_{\gamma}}{b}
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). How and ionization from 21cmFAST:

~ few min on a laptop
density and ionization from coupled hydro+RT (Trac+2009):

~ month on ~1000 core supercomputer

density and ionization from coupled
How to quantify what we will learn?

21CMMC (Greig & AM 2015, 2017) – public, massively-parallelized MCMC driver for 21cmFAST, based on EMCEE sampler (Foreman-Mackey+ 2013). 21cmFAST (AM+2007, 2011) – public, efficient semi-numerical 3D simulation code, extensively tested and currently used by all 21-cm efforts around the globe.
Physical cosmology

Planck 2013, 2015

CMB map
Physical cosmology

CMB map

Planck 2013, 2015

Power spectrum
Physical cosmology

CMB map

Power spectrum

Confidence limits on cosmological parameters with CAMB + cosmoMC

Planck 2013, 2015

Physical cosmology

CMB map

Power spectrum

Confidence limits on cosmological parameters with CAMB + cosmoMC

Planck 2013, 2015
Astrophysical cosmology

"observation"
Astrophysical cosmology

Power spectrum

observation
What are astrophysical parameters?
What are astrophysical parameters???

In principle, our simulation and inference tools can accommodate your favorite model. "Model inclusive" philosophy.
In principle, our simulation and inference tools can accommodate your favorite model... "model inclusive" philosophy

Let's try this simple, yet flexible empirical model...

What are astrophysical parameters?...
Average properties of galaxies in halos of mass $M_h$: an flexible approach based on DM halos + galaxy LFs

$$[\eta/N/\eta/N_{\text{esc}} - \eta/N_{\text{ion}}] \chi_{\text{esc}} = \eta_{\text{ion}}$$

$$\left(\frac{\eta/N_{\text{ion}}}{\eta/N} \right)^{10_{\text{esc}}} \chi_{\text{ion}} = \eta_{\text{ion}}$$

$$\left(\frac{\eta/N_{\text{ion}}}{\eta/N} \right)^{10_{\text{esc}}} \chi_{\text{esc}} = \eta_{\text{esc}}$$

Park+ 2018; Sun & Furlanetto 2016; Dayal+ 2014; Mitra+ 2015; (see also Kuhlen+2012; Mutch+ 2016; Yue+ 2016, …)
Average properties of galaxies in halos of mass $M_h$:

An flexible approach based on DM halos + galaxy LFs
An flexible approach based on DM halos + galaxy LFs

Das, AM+ 2016

SED from Galaxies characterizing emerging X-ray free parameters

Energy (keV) vs. L_{HMXB}/SFR (erg s^{-1} keV^{-1} M_{\odot}^{-1} yr)

\langle 6-10 \text{ keV} \rangle \times \text{Intrinsic SED (Prajgos et al. 2013)}

Intrinsic SED (Prajgos et al. 2013)

Absorbed SED (Prajgos et al. 2013)
Free parameters
Galaxy observations can constrain most parameters to <~10%.
1000h observations with SKA + current galaxy observations with SKA/HERA can constrain EoR history to ~1%.

Parameter constraints: LF + 21cm.
Parameter constraints: LF + 21 cm

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

Park, AM+2018
Parameter constraints: $\Delta t + 21\text{ cm}$

95% confidence level

1000h observations with SKA can constrain EoR history to $\lesssim 1\%$.

Park, AM+2018
In addition to the first astrophysical sources, the Cosmic Dawn also tells us about physical cosmology.
Kern + (2017) 21-cm observations of reionization can improve constraints on $\sigma_8$ by a factor of $\sim 2$. Including physical cosmology
Evoli, AM, Ferrara (2014)
see also Valdez+ (2013)
Lopez-Honorez+2016

DM heating is more uniform than astrophysical -> Heating peak is LOWEST of the three
Evoli, AM, Ferrara (2014)

Cannot be reproduced with astrophysical
Peak is in emission!
SKA’s revolutionary role will be in imaging the first billion years of our Universe.
The 21cm signal is highly non-Gaussian. Using only the power spectrum wastes a lot of information!!!
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.
1. "Brute force" approach: Simply replace the power spectrum in the likelihood calculation of 21CMMC with an alternate statistic, e.g., the bispectrum (Watkinson, Gillett et al. 2018).

2. Machine learning approach: Train Convolutional Neural Networks (CNN) to learn astrophysics and cosmology directly from 21-cm images. Repeat with other statistics, quantifying which one results in the strongest constraints. Repeat with other statistics, quantifying which one results in the strongest constraints. Does that statistic yield tighter constraints on the astrophysical parameters? Does that statistic yield tighter constraints on the astrophysical parameters?
Gillet, AM +, 2018

Deep Learning with CNN: Parameter Recovery

Astrophysics

Output

Layers

Fully connected

Flattening

Successive convolution + pooling

Layer 1

Pooling

Layer 1

Convolution

Input layer

Convolution layer 1

Pooling layer 1

Successive convolution + pooling

Input layer

Flattening

Fully connected

Output
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 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 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 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 images of EoR/CD are non-Gaussian... What are optimal summary statistics?

• SKA will chart the first billion years of our Universe, revolutionizing the field. The properties of sources and sinks are encoded in the 3D for structure.

• To quantify what we can learn, we developed a Bayesian framework for astrophysical parameter estimation, capable of on-the-fly MCMC sampling (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.

• The next decade will see the advent of precision astrophysical cosmology!

We can explore this using neural networks... What are optimal summary statistics?