Radio Cosmology Lab | *Exploring the Epoch of Reionization*

Adam Lanman, Wenyang Li, Joshua Kerrigan,

Jacob Burba, Peter Sims, Jonathan Pober

Brown University Physics

21cm COSMOLOGY @

Introduction

Following the recombination of hydrogen and release of the cosmic microwave background radiation, the baryonic matter of the universe consisted mostly of neutral hydrogen and helium. Gradually, small inhomogeneities collapsed and ignited into the first luminous structures. Energetic photons emitted from the first stars and quasars reionized the surrounding medium, producing ionized bubbles which grew and merged into the fully ionized intergalactic medium we see today. This *Epoch of Reionization* (EoR) remains a poorly-understood period of the universe's history which offers a wealth of cosmological and astrophysical information.

The Pober lab is part of an international effort to build instruments capable of studying the EoR. The neutral hydrogen (HI) of the EoR emits faintly at a wavelength of 21cm due to the hyperfine transition. This emission is unique to neutral hydrogen, and is anti-correlated with the ionized (HII) regions that fill the universe through the EoR. CMB constraints and quasar absorption spectra place the EoR within the redshift range $6 < z < 12$, which means 21cm emissions will reach us at meter scale wavelengths. This is accessible to modern radio interferometers, including the *Donald C. Backer Precision Array for Probing the Epoch of Reionization* (PAPER), the *Murchison Widefield Array* (MWA), and our newly observing *Hydrogen Epoch of Reionization Array* (HERA). Extracting this weak signal remains a challenge unprecedented in radio astronomy.

Differential Brightness Temperature

$$\delta T_b = 27(1+\delta)\, x_{\mathrm{HI}} \left(1 - \frac{T_{\mathrm{CMB}}}{T_{\mathrm{spin}}}\right) \left(\frac{\Omega_b h^2}{0.0223}\right) \left(\frac{1+z}{10}\,\frac{0.15}{\Omega_m h^2}\right)^{1/2} \left[\frac{H(z)/(1+z)}{\partial_r v_r}\right] \ \mathrm{mK}$$

Density perturbation · Neutral fraction · Spin excitation · Baryon and matter mass fractions · Peculiar velocities

The differential brightness temperature δT_b is the contrast between the intensity of 21 cm emissions/absorptions against the Cosmic Microwave Background. Its full expression is related to cosmological (green) and astrophysical (red) parameters. Figure 1 shows the evolution of the spherically-averaged *global* 21 cm brightness temperature.

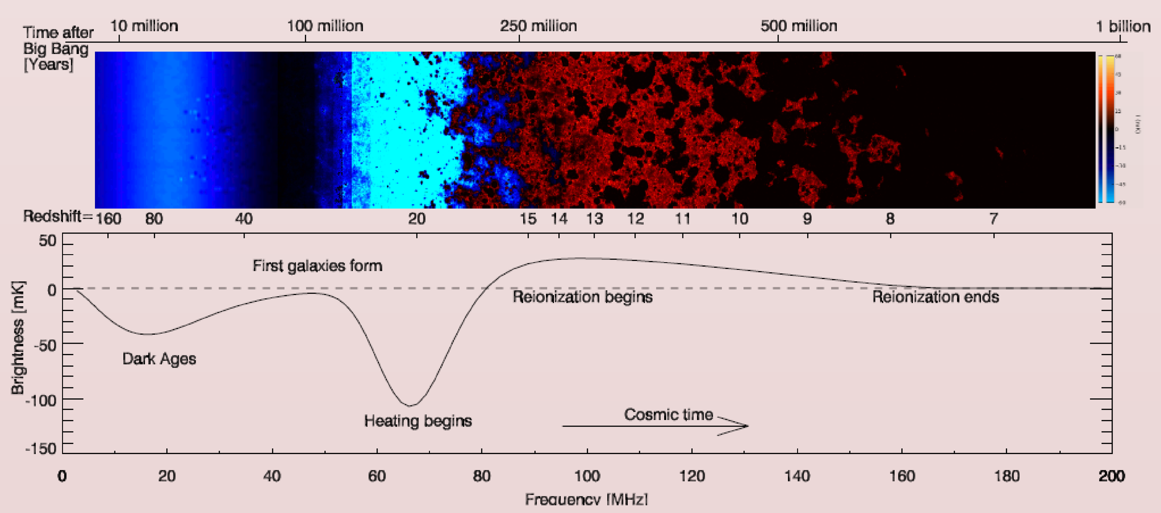

Figure 1: The global differential brightness temperature, δT_b, evolution over redshift $6 < z < 160$. δT_b becomes observable when the spin temperature T_S decouples from the CMB temperature, T_{CMB}. Source: Pritchard & Loeb. Nature 468.7325 (2010): 772-773.

The Foreground Problem

The expected EoR signal is ~ 5 orders of magnitude weaker than known foreground sources, such as diffuse emission from the Galaxy and extragalactic point sources. Removing these foregrounds, as well as instrumental noise, is a nontrivial problem. An example of the overlapping sources of power in our observations can be seen in Fig. 2. In addition to galactic and extragalactic foregrounds we must also contend with Radio Frequency Interference (RFI).

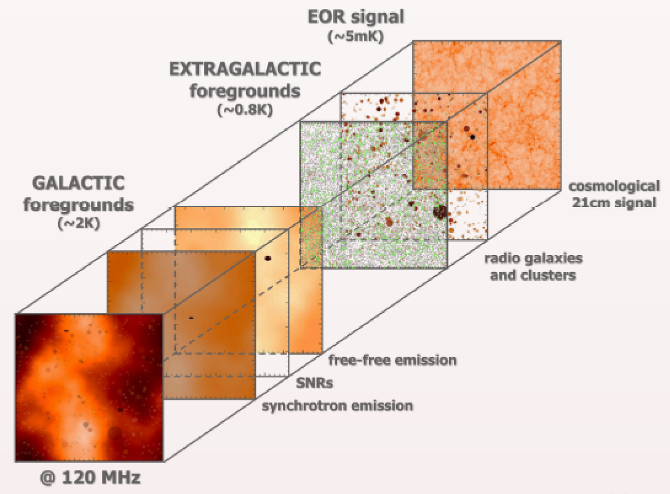

Figure 2: The various cosmological and galactic sources that contribute to the measured sky temperature, and their relative strengths.
Source: Zaroubi, Saleem. 'The First Galaxies. Springer Berlin Heidelberg, 2013. 45-101.

Deep Learning for RFI Mitigation

RFI is present in all radio observations and consists of both terrestrial (e.g. TV stations) and in orbit (e.g. satellites) sources. This interference reduces sensitivity in 21cm EoR experiments because it can be present at nearly all frequencies and times, and is brighter than galactic foregrounds. We can treat interferometry data as time-ordered visibilities which are image-like, and RFI manifests itself in this data as sharp discontinuities, thus novel machine learning image techniques can be applied. We introduce a Deep Fully Convolutional Neural Network (D-FCN) (Figure 3) which uses the time-frequency context from both amplitude and phase to form a robust and efficient method for identifying RFI. Figure 4 demonstrates an amplitude only and amplitude-phase D-FCN, and the Watershed RFI algorithm, as applied to a HERA waterfall visibility in the sub-band of 157-193 MHz.

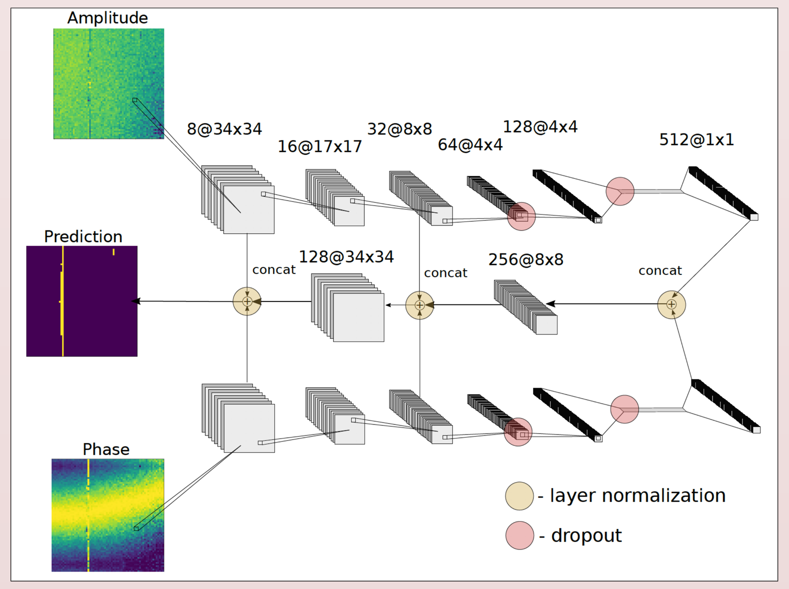

Figure 3: Deep Full Convolutional Neural Network (D-FCN) architecture design for application to interferometric visibilities. The input space consists of a normalized log amplitude (upper branch) and it's corresponding phase component (lower branch). Both input layers reintroduce coarse time-frequency information into the transpose convolutional layers in what's referred to as 'skip connections'.

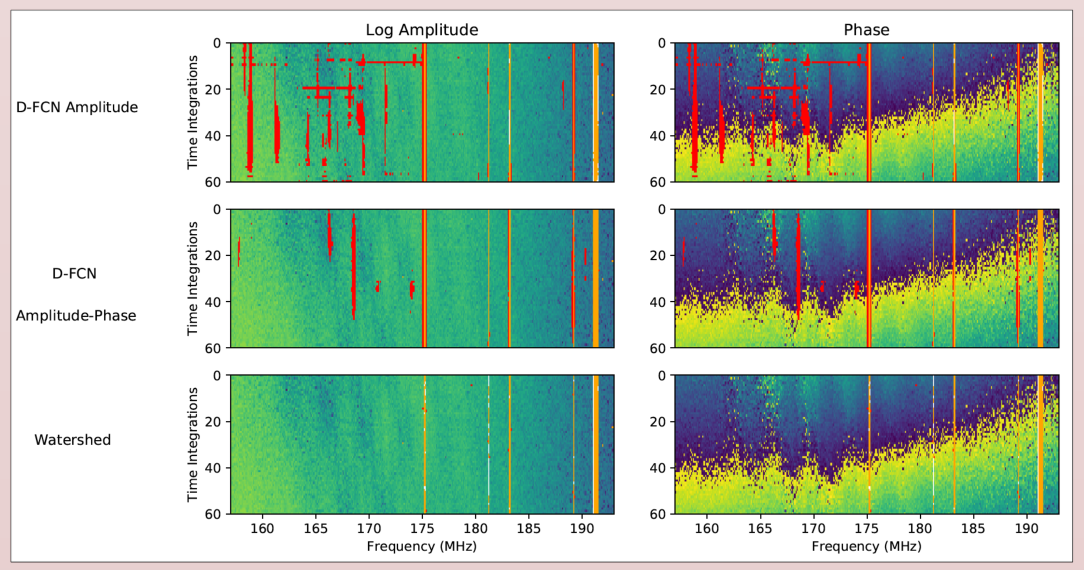

Figure 4: A comparison between three RFI identification algorithms, a D-FCN using only amplitude as input, a D-FCN with amplitude and phase (Figure 3), and the Watershed RFI algorithm. True positives are indicated in orange, false positives (red), and false negatives (white).

Simulation

The particular characteristics of an array can introduce unexpected effects into the data. Understanding and mitigating instrumental effects is critical to making a confident detection of the EoR. For this reason, much effort has been put into simulating the full analysis pipeline – from the point and diffuse sources on the sky, to the raw visibilities that come out of the correlator, to the power spectrum estimations.

Fast Holographic Deconvolution (FHD) is a purpose-built software framework for analyzing MWA data. FHD does foreground subtraction by *forward modeling*, which builds a simulated data set, including instrumental effects, and subtracts it from the actual data. This forward modeling feature can also be used as a standalone simulation tool, to generate raw visibilities of foregrounds, noise, and EoR off of existing and future 21cm experiments.

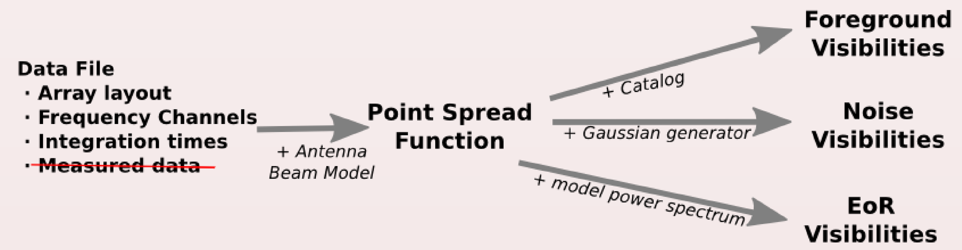

Figure 5: A sample data file (or generated data file) holds array coordinates over time and the frequency channels of the instrument. Given a beam model for the antenna, FHD calculates the full synthesized beam (or point-spread function) for a particular time and set of frequencies. The synthesized beam can then convert a sky catalog into a set of foreground visibilities for the instrument. External EoR simulations can also be fed in to test EoR sensitivity, or Gaussian noise can be injected to simulate noise.

Calibration

Interferometric arrays seeking to measure the 21 cm signal from the EoR must contend with overwhelmingly bright emission from foreground sources. Accurate recovery of the 21 cm signal will require precise calibration of the array, and several new avenues for calibration have been pursued in recent years. Current calibration efforts for EoR observations largely fall into two camps: sky-based calibration using deep foreground catalogs and forward modeling of the instrument visibilities, and redundant calibration that foregoes a sky model but requires the antennas be placed on a regular grid. A further exploration of combining both approaches has been pushing on to mitigate the contamination in the power spectrum.

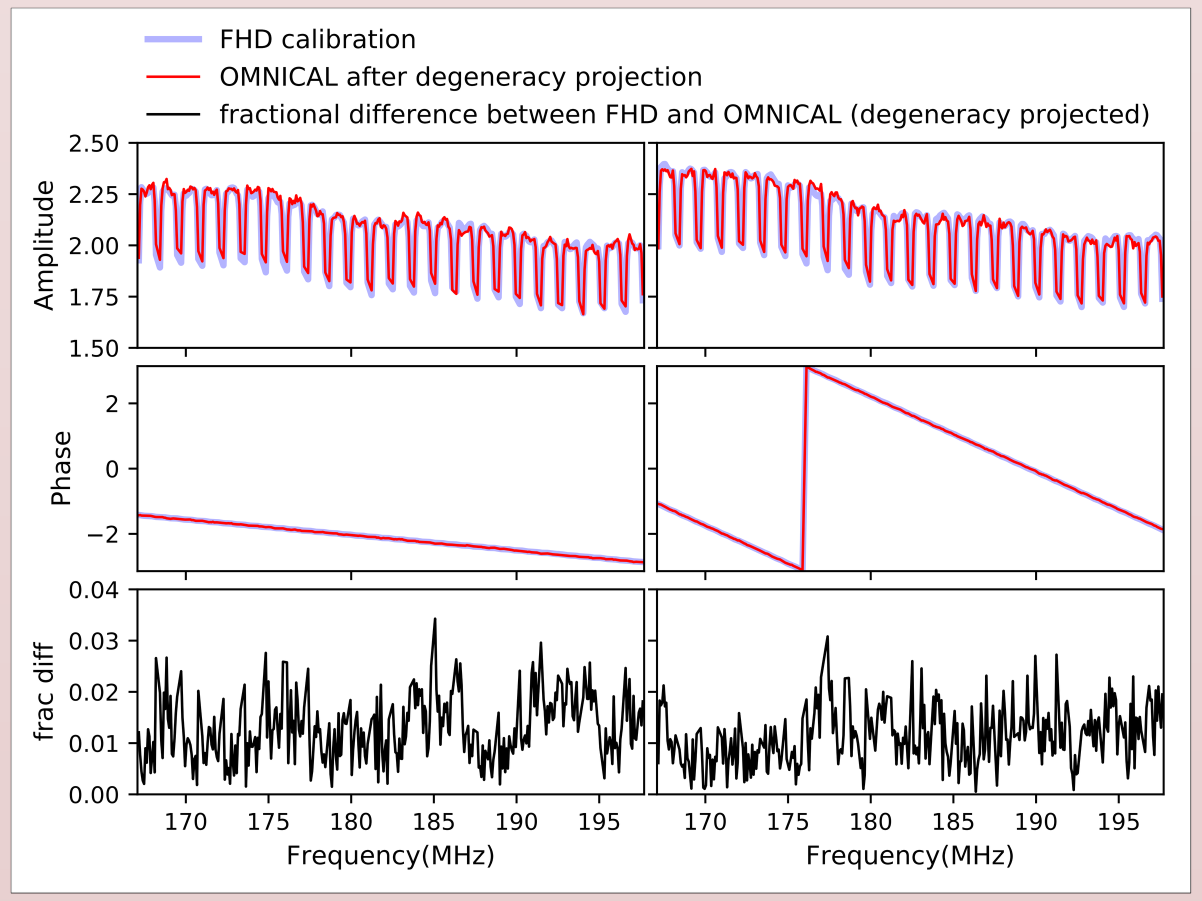

Figure 6: 30 minutes averaged gain calibrations of 2 MWA PhaseII tiles. Upper: Gain amplitude; Middle: Gain phase. Lower: fractional difference between sky based (FHD) and redundancy based (OMNICAL) solutions with degeneracy projected. Blue: FHD solutions; Red: OMNICAL solutions after projecting degeneracy.

Hydrogen Epoch of Reionization Array (HERA)

The Hydrogen Epoch of Reionization Array is a 2^{nd} generation radio interferometer observing the 21 cm emission from neutral hydrogen during the Epoch of Reionization ($z = 6$ - 12). HERA will eventually be an array with 350 14-meter parabolic dishes in South Africa. These elements will be divided into a 320-element hexagonal shaped core and 30 outriggers. The current stage of HERA has 37 dishes deployed and observing.

Although increasingly stringent upper limits of the 21 cm signal have been placed by the first generation experiments targeting the EoR such as the Precision Array for Probing the Epoch of Reionization array (PAPER), Murchison Widefield Array (MWA) and LOw Frequency ARray (LOFAR), they are not able to detect it due to their sensitivity limits (see Fig. 8). HERA is designed to bring both the sensitivity and precision required to directly constrain the topology and evolution of reionization. By understanding the evolution of reionization we can better comprehend the formation of the first stars and earliest galaxies.

Figure 7: HERA - Located in the Karoo desert in South Africa. The current phase of HERA consists of 37 14m parabolic dishes. In its final form, HERA will comprise of 350 dishes in a closely-packed hexagonal grid with outriggers. HERA is additionally a pathfinder project for the Square Kilometer Array (SKA).

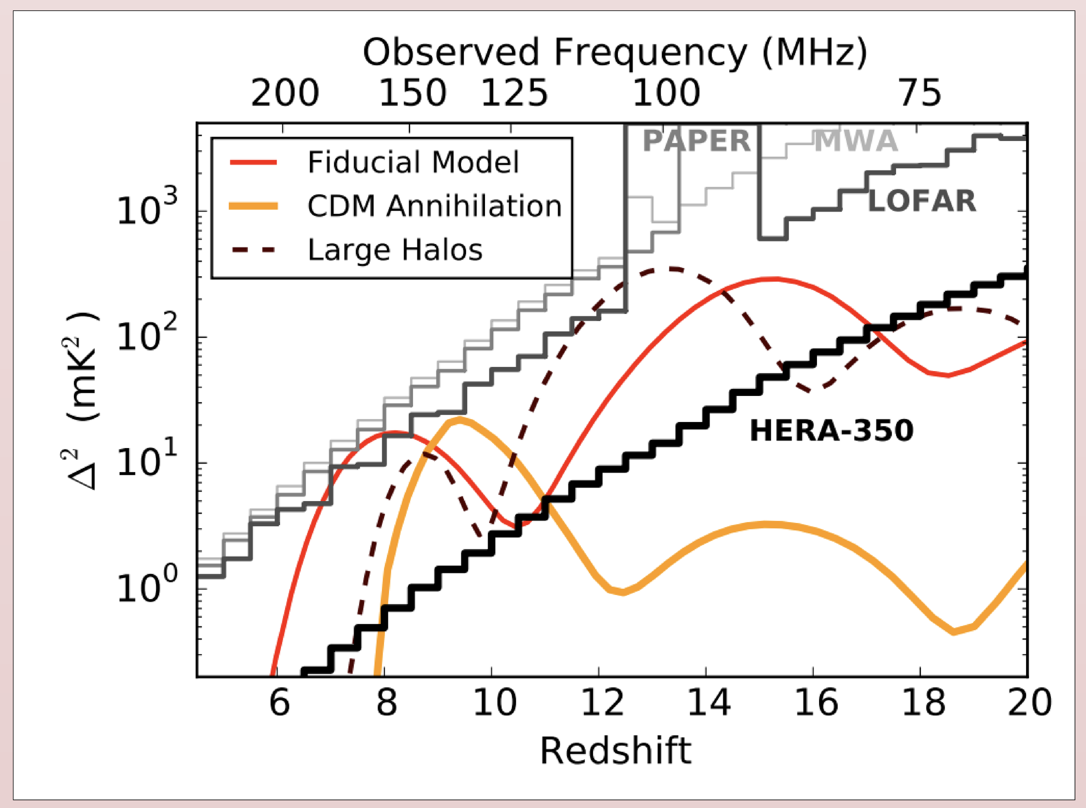

Figure 8: Projected SNR comparison of the 350 element phase of HERA (with 1080 hrs integration time) to other low frequency arrays. Relative to several reionization models (colored), HERA-350 should have the sensitivity to measure the 21cm EoR signal across multiple redshifts. Source: DeBoer & HERA Collaboration. PASP (2017).