The HERA Road Map:
A Pathway to Understanding the Early Universe

HERA Coordinating Committee[1]

MWA and PAPER Groups

Received ______________; accepted ______________

.....

1

ABSTRACT

The abstract goes here.

Subject headings: MWA, PAPER

1. The HERA Road Map

The Hydrogen Epoch of Reionization Array (HERA) is a scientific road map aimed at exploring the large-scale structure in the baryonic universe via the 21cm line of hydrogen. A white paper describing this investigation was submitted to the U.S. National Research Council's 2010 Committee for the Decadal Survey of Astronomy and Astrophysics, Program Panel: Radio, Millimeter, and Submillimeter (RMS). A three-phased program was proposed: I) the detection of the power spectrum of the 21 cm line emission structures; II) the characterization of the structures including galaxy formation astrophysics and cosmological physics; and III) the detailed imaging of structures across a large fractional step in cosmic time. The science, which was highly ranked by the Committee, is aimed toward answering such major questions as what objects first lit up the Universe? When did this occur? How has the Universe evolved over time? In addition, the Committee recommended the creation of the Mid-Scale Innovations Program at the National Science Foundation (NSF) to fund projects such as HERA.

The current HERA I initiatives, PAPER, the Murchison Widefield Array (MWA), and the Experiment to Detect the Global EoR Step (EDGES), are expected to culminate in a detection and will define the scope of the HERA II characterization milestones. The imaging task of HERA III will require a very large-scale array to achieve the necessary sensitivity over a wide range of spatial frequencies. It is expected that this will be achieved by the Square Kilometer Array (SKA) in the coming decade and will incorporate into its design the accrued body of knowledge gained from HERA I and II activities. Hence, the accumulation of experience with each step in the HERA program along with parallel technical development are required to achieve the goals of signal detection, detailed statistical characterization, and structural imaging.

The national observatories, in partnerships with universities, are important resources in the HERA road map paradigm. At the beginning of an investigation, the small scale of the instrumentation is such that it can be managed effectively as a PI driven activity. However, the observatories can assist the PI with technical developments and provide opportunities for deploying the instrument in an RFI-controlled environment near key infrastructure needs such as power, Internet, housing, technical expertise, works area assistance, etc. As the instrument grows in both complexity and size, the observatory can help manage the construction and evaluation of the hundreds of components needed for the array projects. It is an opportunity for students from the universities to work with observatory staff to learn project development and management skills without being overwhelmed by the construction effort itself. Upon completion of the project, the operation, data management, and data analysis tasks can be handled by the students and researchers at the universities. These partnership can be optimized, on a case-by-case basis, to help achieve the milestones of the scientific investigation while still maintaining the PI/Co-I driven approach to HERA.

In 2011, the HERA Coordinating Committee was formed to bring together representatives from each of the current HERA I activities to share lessons learned and begin charting a course toward the next milestone. The Committee consists of three members of the MWA team: J. Hewitt (MIT), M. Morales (U. Washington), and J. Bowman (Arizona State), and three members of the PAPER team: R. Bradley (NRAO / U. Virginia), A. Parsons (U.C. Berkeley) and C. Carilli (NRAO). The EDGES project was also represented by J. Bowman. Through a series of monthly telecons, the HERA Coordinating Committee members exchanged detailed information about the projects and discussed the various technologies and methodologies that were utilized. The work of the committee over the past year has culminated in the writing of this paper.

The purpose of this paper is to introduce the HERA road map to the astronomical community and provide details of its implementation. Part II gives a brief topical overview of the science, a snapshot of our current understanding of EoR in the context of 21cm cosmology, with emphasis placed on those empirical and theoretical results that impact the boundaries of new instrument designs. Our acquired body of knowledge is summarized in Part III, where we provide an overview of the various HERA I projects and a synopsis of our methodological lessons learned. Finally, in Part IV, we extrapolate our current findings into the future by providing the motivational framework and conceptual design for achieving the next major milestone array along the HERA trail.

2. The Evolution of Hydrogen during the Early Universe

2.1. Science of reionization

Cosmic reionization corresponds to the transition from a fully neutral intergalactic medium (IGM) to an (almost) fully ionized IGM caused by the UV radiation from the first luminous objects. Reionization is a key benchmark in cosmic structure formation, indicating the formation of the first luminous objects. Reionization, and the preceding 'dark ages', represent the last of the major phases of cosmic structure formation left to explore.

The study of first light and cosmic reionization was called-out in 'New Worlds, New Horizons,' as a primary discovery area in the coming decade. A major goal for large area telescopes such as the JWST, OWL, and ALMA is to detect directly the first galaxies and AGN during reionization. In parallel, HI 21cm emission from the IGM has been recognized the most direct method to probe the neutral gas that pervaded the Universe after recombination. Extensive reviews have been written recently on first galaxies, the process of reionization, and the expected HI 21cm signals (Furlanetto, Oh, Briggs 2007; Ellis 2007; Fan et al. 2006a; Morales & Wyithe 2010). We briefly review some of the key aspects of reionization relevant to the HI 21cm studies that will be performed by HERA.

The last decade has seen the first observational evidence for cosmic reionization. The primary results come from the Gunn-Peterson effect, ie. resonant scattering in the Lyα line by the neutral IGM, toward the most distant quasars ($z \sim 6$), and the large scale polarization of the CMB, corresponding to Thomson scattering during reionization.

Figure 1 shows the current state of observational constraints on the cosmic neutral fraction as a function of redshift. The GP effect, and related statistical measures of eg. dark gaps in QSO spectra (Fan et al. 2006b), suggest a qualitative change in the nature of the IGM at $z \sim 6$, likely signifying a rapidly increasing neutral fraction during the tail-end of reionization. At the other extreme, the CMB large scale polarization suggests a significant ionization fraction extending to $z \sim 11$ (Komatsu et al. 2011ApJS..192...18). Between these two extremes, there are only marginal constraints.

First, the study of the evolution of the sizes of the ionization near-zones around the highest redshift quasars shows a clear decrease in the near-zone sizes from $z \sim 5.5$ to 6.5, consistent with an increase in the volume-averaged neutral fraction, by an order of magnitude over this redshift range, from $f_{HI} < 10^{-4}$ to $f_{HI} > 10^{-3}$ (Carilli et al. 2010ApJ...714..834). More recently, observation of a damped Ly-α wing in the Gunn-Peterson absorption trough edge in the most distant quasar suggests an even more rapid rise in the neutral fraction, to $f_{HI} \geq 0.1$ by $z = 7.08$ (Bolton et al. 011MNRAS.416L..70; Mortlock et al. 2011Natur.474..616).

These first measurements of the evolution of the IGM neutral fraction during reionization are truly water-shed events. However, it has also become clear that these important first probes of cosmic reionization are fundamentally limited. The CMB polarization represents an integral measure of the Thomson scattering optical depth back to recombination, and hence can be fit by many different reionization scenarios. For the Gunn-Peterson effect, the IGM becomes optically thick to Lyα absorption for a neutral fraction $> 10^{-3}$, and hence the diagnostic capabilities of this probe effectively saturate at low neutral fractions. The GP conclusions are also depend on an assumed clumping factor

Fig. 1.— The neutral fraction, by volume, of the intergalactic medium as a function of redshift. Current constraints are shown, including CMB largescale polarization constraints (Komatsu et al. 2011), Gunn-Peterson absorption troughs (Fan et al. 2006), quasar near-zone size evolution (Carilli et al. 2010), and the recent damped-Lyα GP absorption wing in the most distant ($z = 7.08$) quasar (Bolton et al. 2011P). Also shown is a representative model for the HI evolution from cosmological simulations (Gnedin & Fan 2006ApJ...648....1).

for the IGM. And the Gunn-Peterson results at $z > 7$ are based on just one quasar, eg. the damped absorption profile of the GP wing has a small, but finite probability of being simply a gas rich galaxy with 2Mpc of the quasar (Bolton et al. 2011).

It is widely recognized that the most direct and incisive means of studying cosmic reionization is through the 21cm line of neutral Hydrogen. The study of HI 21cm emission from cosmic reionization entails the study of large scale structure (LSS), meaning comoving scales $\geq$ 10Mpc, and total HI masses $> 10^{12}$ $M_{\odot}$. Excellent reviews of the expected HI21cm signal from reionization have been published in the last few years (Furlanetto et al. 2006; Morales & Wyithe 2010). Herein we review the basic principles as they pertain to the HERA program.

The optical depth on the 21cm line of neutral hydrogen is given by [PROBABLY DONT NEED TO INCLUDE EQUATIONS, BUT FOR COMPLETENESS]:

$$\tau = \frac{3c^3\hbar A_{10} n_{HI}}{16 k_B \nu_{21}^2 T_S H(z)} \sim 0.0074 \frac{x_{HI}}{T_S}(1+\delta)(1+z)^{3/2}[H(z)/(\frac{dv}{dr})] \tag{1}$$

where A is the Einstein coefficient and $\nu_{21} = 1420.40575$ MHz. This equation shows immediately the rich physics involved in studying the HI 21cm line during reionization, with τ depending on the evolution of cosmic over-densities, δ (predominantly in the linear regime), the neutral fraction, x_{HI} (ie. reionization), the HI excitation, or spin, temperature, T_S, and the velocity structure, $\frac{dv}{dr}$, including the Hubble flow and peculiar velocities.

In the Raleigh-Jeans limit, the observed brightness temperature (relative to the CMB) due to the HI 21cm line at a frequency $\nu = \nu_{21}/(1+z)$, is given by:

$$T_B \approx \frac{T_S - T_{\rm CMB}}{1+z}\tau \approx 7(1+\delta)x_{HI}(1 - \frac{T_{CMB}}{T_S})(1+z)^{1/2} \text{ mK}, \tag{2}$$

The conversion factor from brightness temperature to specific intensity, I_ν, is given by:

$I_\nu = \frac{2k_B}{(\lambda_{21}(1+z))^2} T_B = 22(1+z)^{-2} T_B$ Jy deg^{-2}. These equations show that for $T_S \sim T_{CMB}$ one expects no 21cm signal. When $T_S >> T_{CMB}$, the brightness temperature becomes independent of spin temperature. When $T_S << T_{CMB}$, we expect a strong negative (ie. absorption) signal against the CMB. During reionization it is expected that we are in the latter regime, although during the preceding dark ages the interplay between the CMB temperature, the kinetic temperature, and the spin temperature, coupled with radiative transfer, lead interesting probes of the physics of the linaerly-evolving IGM (Burns et al. 2012AdSpR..49..433).

2.2. The HI 21cm signals and telescope sensitivities

[THESE NEED TO BE RECAST IN TERMS OF MOST RECENT MODELS, TELESCOPE SENSITIVITY PREDICTI0NS, AND THE FEW MEASUREMENTS FROM EG EDGES – IS THERE ANYTHING THAT WE COULD ADD HERE THAT WOULD LOOK LIKE 'ORIGINAL RESEARCH'? FOR NOW, I INCLUDE PARAGRAPHS FROM A PREVIOUS SKA MEMO.]

Global signal: The left panel in Figure 3 shows predictions of the global (all sky) increase in the background temperature due to the HI 21cm line from the neutral IGM (Gnedin & Shaver 2003). The predicted HI signal peaks at roughly 20 mK above the foreground at $z \sim 10$. At higher redshift, prior to IGM warming, but allowing for Lyα emission from the first luminous objects, the HI is seen in absorption against the CMB. Since this is an all sky signal, the sensitivity of the experiment is independent of telescope collecting area, and the experiment can be done using small area telescopes at low frequency, with very well controlled, and calibrated, frequency response (Bowman, Rogers, Hewitt 2007). Note that the line signal is only $\sim 10^{-4}$ that of the mean foreground continuum emission at ~ 150 MHz.

[REVISED ABOVE BASED ON BOWMAN and ROGERS 2012]

Power spectra: The middle panel in Figure 3 shows the predicted power spectrum of spatial fluctuations in the sky brightness temperature due to the HI 21cm line (Mcquinn et al. 2006). For power spectral analyses the sensitivity is greatly enhanced relative to direct imaging due to the fact that the universe is isotropic, and hence one can average the measurements in annuli in the Fourier (uv) domain, ie. the statistics of fluctuations along an annulus in the uv-plane are equivalent. Moreover, unlike the CMB, HI line studies provide spatial and redshift information, and hence the power spectral analysis can be performed in three dimensions. The rms fluctuations at $z = 10$ peak at about 10 mK rms on scales $\ell \sim 5000$.

A recent analysis by Lidz et al. (2008) shows that the MWA should be able to determine the HI 21cm power-spectrum over roughly a decade in wavenumber, $k \sim 0.1$ to $1h$ Mpc^{-1}. These data should be able to constrain both the amplitude and slope of the power spectrum. A key aspect of these measurements is determination of the redshift evolution of these quantities, with the amplitude of the rms power in the HI 21cm fluctuations peaking at redshift when the IGM is roughly 50% neutral.

[REVISE ABOVE BASED ON LATEST MWA AND PAPER PREDICTIONS]

Absorption toward discrete radio sources: An interesting alternative to emission studies is the possibility of studying smaller scale structure in the neutral IGM by looking for HI 21cm absorption toward the first radio-loud objects (AGN, star forming galaxies, GRBs) (Carilli et al. 2002; Furlanetto & Loeb 2002). The rightpanel of Figure 3 shows the predicted HI 21cm absorption signal toward a high redshift radio source due to the 'cosmic web' prior to reionization, based on numerical simulations. For a source at $z = 10$, these simulations predict an average optical depth due to 21cm absorption of about 1%, corresponding to the 'radio Gunn-Peterson effect', and about five narrow (few km/s) absorption lines per MHz with optical depths of a few to 10%. These latter lines are equivalent to the Ly α forest seen after reionization (a recent treatment of this problem can be found in Furlanetto 2006). Furlanetto & Loeb (2002) predict a similar HI 21cm absorption line density due to gas in minihalos as that expected for the 21cm forest.

[REVISED ABOVE BASED ON SENSITIVITY OF MWA AND PAPER – WHAT FLUX DENSITY SOURCE WOULD ONE NEED? I THINK THE ANSWER IS LIKE 1JY AT 150MHZ]

Tomography: Figure 4 shows the expected evolution of the HI 21cm signal during reionization based on numerical simulations (Zaldariagga et al. 2004; see also Zahn et al. 2007; Iliev et al. 2006; Shapiro et al. 2006, etc...). In this simulation, the HII regions caused by galaxy formation are seen in the redshift range $z \sim 8$ to 10, reaching scales up to $2'$ (frequency widths ~ 0.3 MHz ~ 0.5 Mpc physical size). These regions have (negative) brightness temperatures up to 20 mK relative to the mean HI signal. This corresponds to 5μJy beam^{-1} in a $2'$ beam at 140 MHz. Only the full SKA will able to image such structures.

[ANY NEW SIMULATIONS TO SHOW?]

Largest Cosmic Stromgren Spheres (CSS): While direct detection of the typical structure of HI and HII regions may be out of reach of the near-term EoR 21cm telescopes, there is a chance that even this first generation of telescopes will be able to detect the rare, very large scale HII regions associated with luminous quasars near the end of reionization. The expected signal is ~ 20mK $\times x_{HI}$ on a scale $\sim 10'$ to $15'$, with a line width of ~ 1 to 2 MHz (Wyithe, Loeb, Barnes 2005). This corresponds to 0.5 $\times x_{HI}$ mJy beam^{-1}, for a $15'$ beam at $z \sim 6$ to 7, where x_{HI} is the IGM neutral fraction. Figure 5 shows a simulation of the expected images from near-term reionization telescopes, such as the MWA, for these very large structures. The key aspect of these systems is that the structures are three dimensional, making them much easier to detect than continuum structures at a similar brightness.

Wyithe et al. (2005) predict that, for late reionization, there should be several tens of (mostly fossil) large ($>$ 4Mpc physical) CSS per 10^o field of view in $z \sim 6$ to 8. An interesting addition to the QSO CSS studies is the possibility that massive galaxy formation at very high redshift is highly clustered, occuring in rare peaks in the cosmic density field. The total number of ionizing photons from star formation integrated over the lifetime of the system can match that radiated by a luminous QSO (Li et al. 2007; Wyithe et al. 2007), and hence generate CSS comparable to those predicted for the QSOs. The star forming systems will both increase the size of the spheres around bright QSOs (since QSO and massive galaxy formation are likely to be coupled), as well as increase the number of large spheres, in cases where the QSO has not turned-on yet.

[NEED TO REVISE ABOVE BASED ON LATEST SENSITIVITY, AND ADD STATEMENT ABOUT REALLY WIDE FIELD AND DEPTH OF EG. PAPER PICKING-UP PERHAPS THE MOST EXTREME STRUCTURE IN THE UNIVERSE.]

3. The Body of Knowledge: HERA I Experiments

The HERA roadmap is predicated on the idea that the community will learn from our experience as we build a series of precursor instruments. This precludes writing the standard white paper where our current vision of the ideal instrument is detailed. There are a number of these papers in the literature (**? ? ? ?**), and the design assumptions have often fared poorly in the light of subsequent experience. The strength of the EoR field is the rapid advances in instrumentation and technique[1], so the HERA roadmap proposes a staged and experimental approach to developing each subsequent instrument.

However, for this approach to work we need to capture the lessons learned and the key questions facing the community at each stage in the process. In place of a straw man instrument design with optimistic sensitivity estimates, this section documents our current state of knowledge and the key questions that will determine the design of the next HERA instrument.

This section is organized around a set of **Lessons & Questions** in each of 11 areas. The topics in sections 3.1–3.4 need to be considered prior to developing a conceptual instrument design. However, no instrument is built on a blank sheet of paper and sections 3.5–3.12 cover the lessons that must be considered to go from a conceptual design to a costable proposal. These real world consideration often significantly impact the look of the final instrument, with complicated tradeoffs between different constraints. True instrument design is weighing tradeoffs, and here we attempt to document the lessons already learned and questions the community needs to answer to design the next HERA instrument.

---

[1]Many of these technical advances may have significant impacts on the wider field of radio interferometry. Some particularly interesting examples include **? ? ? ? ?**)

3.1. Drift vs. track

One of the first questions is whether to concentrate the observing on a few cold regions of the sky with long tracked observations, or to pursue a drift scanning strategy that surveys a large percentage of the celestial sphere. This decision has profound effects on instrument sensitivity, calibration, and antenna design.

For the same collecting area and observing time, tracked observations provide a significantly higher sensitivity to the EoR power spectrum.[2] Because the PS modes are independent for each observed field, they must be added incoherently (squared before averaging). This implies that dividing your observing time between N fields reduces the sensitivity by $\sim\sqrt{N}$ (). In practice the thermal noise penalty is even higher as tracked observations target unusually cold portions of the sky.

The thermal noise sensitivity advantage of tracked observations is partially offset by a higher sample variance at large angular scales. As with any PS measurement one would prefer to stare at one region of the sky to drive down the thermal noise (coherent addition of signal) until a S/N ratio of $\sim$1 is reached, at which point your time should be spent surveying more sky (to lower sample variance, (**?**)). As successive HERA instruments become more sensitive sample variance could dominate the measurements, though the optimal balance depends on both the antenna layout (which scales reach the sample variance limit) and the unknown strength of the EoR signal. From a theoretical sensitivity point of view, tracked observations are preferred as they minimize the thermal noise and maximize flexibility in trading observing depth vs. survey area.

However, precision calibration favors a drift scan strategy. Calibration is a key part of foreground removal and the antenna beams must be accurately understood to successfully

[2]In the limit of a well-filled u, v plane.

reveal the EoR PS. Stationary antennas that allow the sky to drift past (or stare at one of the celestial poles) are inherently much more stable than steered antennas, facilitating precision calibration and inexpensive construction. This is particularly true for the small antennas typical of HERA arrays that are only a few wavelengths across. Strong diffraction effects lead to either moving sidelobes for beamed dipole arrays (MWA & LOFAR) or highly variable ground pickup for steered antennas. Pointed antennas have a slight advantage in targeting calibrator sources (bright calibrators, pulsars), but this is generally outweighed by the stability of stationary antennas enabled by a drift scanning strategy.

The last piece of the drift vs. tracking decision is the size of the antennas. Drift scanning instruments prefer small antennas with very large fields of view, bounded by the capacity of the correlator and data storage systems (§3.10 & 3.11) to handle all of the resulting baselines. This both maximizes the instrument sensitivity and makes the antennas easier to manufacture to high precision. The steerable antennas needed for tracking observations tend to be larger, either tuned to the $\sim 30^\circ$ diameter of cold spots in the galactic synchrotron emission (MWA) or even larger to increase the collecting area while keeping correlator and data handling costs in check (LOFAR, GMRT).

Drift vs. tracking observing strategy has profound implications for instrument design, with advantages and disadvantages for each.

Lessons & Questions

- Tracking is more sensitive, drifting is easier to calibrate. PAPER (drift) and MWA (tracking) HERA precursors are directly examining the real world pros and cons for each approach.

3.2. Calibration

It has been long recognized that mode-mixing foreground contamination (see Morales & Wyithe for a review) will necessitate precision calibration to successfully remove the bright EoR foregrounds. In general the more accurate the calibration the larger the uncontaminated EoR window in which the measurement can be performed is. So the quality of the calibration is directly related to the sensitivity of the measurement.

There have been four techniques which have been pursued to determine the electrical, holographic, and polarization components of the calibration: redundant baselines, bright astrophysical sources, pulsars, and artificial beacons. Before detailing these techniques it is worth reviewing the three separate aspects of calibration they must constrain.

The most familiar is the complex gain term solved for by traditional selfcal algorithms. The complex gain of each antenna describes the overall sensitivity of the antenna and receiving chain (collecting efficiency, amplifier gain, cable losses and reflections, etc.) and the electrical phase delay related to the distance of the antenna from the digitizer (antenna location and height, cable and waveguide TOF, filter responses, etc.). Conceptually the complex gain gives the sensitivity and time delay of each antenna to an on-axis source as a function of frequency.

However, the response of antenna is not uniform over the field-of-view. The direction-dependent terms can be described in a number of ways (), but the gold standard is the holographic antenna map. Traditionally the holographic antenna map is created by rastering a bright source across the antenna FoV while comparing to a reference antenna (), then Fourier transforming to the antenna plane to produce a map of the complex antenna efficiency for each location across the antenna's surface. Conceptually, a complex number per $\sim(\lambda/2)^2$ area of the antenna will give a complete description of the antenna's direction-dependent response (per polarization & frequency; the electric field is correlated below the $(\lambda/2)^2$ scale). Recent developments have generated many new ways of determining the holographic antenna map or equivalent direction-dependent calibrations without steerable antennas ().

Lastly, the polarimetric responses of antennas must be considered, both electrically (polarization leakage) and in a direction-dependent sense (**?**). At these frequencies all antennas have strong polarization characteristics with fractional polarizations of tens of percent being common. [Possibly insert some text talking about the importance of the polarization calibration, with Gianni's drift map of diffuse polarization and Bryna's figure on the error from trying to perform a PS with "I" only.]

For arrays with antennas arranged in a repeating tiling pattern, there are many redundant baselines. If the direction-dependent response of all of the antennas are identical, each antenna will see the same sky and this redundancy may be used to calibrate the non-direction dependent electrical and polarization calibration terms (). The advantage of this approach is it does not depend on knowing the sky accurately and has a very high signal to noise—effectively it uses all sources in the FoV to perform the calibration. The disadvantage is the requirement that both the direction-dependent responses and the separations be identical to high precision for all antennas. If these assumptions are broken, detailed knowledge of the sky must be used.

The other approaches all use a known source of astrophysical or artificial origin to measure the electrical and direction-dependent responses of the antennas. The MWA calibration system () uses bright astrophysical sources across the antenna FoV to measure the gain of each antenna along a number of directions, which can then be inverted to create holographic and electrical calibrations for each antenna. The advantage of astrophysical sources is they are numerous, the disadvantage is there are so many it is difficult to isolate the flux from individual calibrators due to the imperfect Point Spread Function (PSF, or array beam). In contrast to the redundant antenna layouts required for redundant baseline calibration, calibration with astrophysical sources needs an excellent snapshot PSF and very uniform baseline distributions (pseudo-random low redundancy layouts, (**?**)).

The GMRT has pioneered the use of pulsars as a calibrator (). The emission from pulsars can be gated into times when pulsar is on and times when it is off. Because other astronomical sources are very steady, subtracting the on and off emission removes the contribution from all sources except the pulsar. In addition, the pulsar emission is polarized and the polarization angle rotates through the pulse providing a nearly ideal polarization calibrator. Pulsar calibration avoids the requirements on array layout as time is used to isolate the emission instead of the array PSF. The disadvantages of pulsar calibration are there are very few bright pulsars, particularly for mapping out the direction-dependent response of non-steerable antennas.

The last approach is to use an artificial beacon of known characteristics to map out the antenna responses. Because the calibrator needs to be in the far field it must be $\sim$hundred meters or more from the antenna (depending on antenna diameter). Potential calibrators include satellites, balloons either with transmitters or made of mylar to reflect ground signals, and piloted or remote-control airplanes and helicopters. All of these face the challenge of providing a signal of known strength and polarization that can map the antenna beam, preferably at many different frequencies. To date the most successful attempts have used the Orbcom LEO satellite transmission at 137 MHz to map out the fiducial antenna responses of the PAPER and MWA antennas at Green Bank ().

The last consideration of any calibration system is the stability of antenna under calibration. On whatever timescale the calibration changes (temperature fluctuations, surface water from rain, ionophsere, earth rotation, etc.), the calibration must be repeated. Our understanding of the antennas behavior in different conditions thus places constraints on the calibration approach.

Lessons & Questions

- Choosing the either the redundant baseline or astrophysical source calibration strategies strongly constrains the antenna layout.
- Polarimetric calibration is essential.
- Holographic calibration of the antenna directional response is necessary. Whether each antenna must be calibrated independently or the antenna-to-antenna response is similar enough to enable calibration of only a fiducial antenna is under study.
- Redundant baseline calibration is promising, but has yet to be demonstrated in the field.
- May be effective to have certain baselines or antennas dedicated to calibration.

3.3. Foreground subtraction flow down requirements

Separating the strong astrophysical foregrounds from the cosmological EoR signal is the primary challenge of HERA arrays (see **?** for a recent review). [*This text might go earlier, and/or be heavily edited. Here as a placeholder.* The conceptual framework for EoR foreground subtraction is the 3D Fourier space and the k-space figure of merit. Because the HI radiation is a narrow emission line, EoR observations are inherently over a three-dimensional volume with the observed frequency directly providing the emission redshift. The observed volume $\{\theta_x, \theta_y, f\}$ can be mapped by the redshift and angular diameter distance into cosmological coordinates ($\{x, y, z\}$ in cMpc) and Fourier transformed into a three dimensional wavenumber cube $\{k_x, k_y, k_z\}$.[3]

In the absence of cosmic evolution and velocity distortions, the power spectrum signal is spherically symmetric in this three-dimensional k-space $\{k_x, k_y, k_z\}$. This approximate spherical symmetry is due to the isotropy (rotational invariance) of space, and its importance for HI cosmology measurements is difficult to overstate. Measurements of the one dimensional power spectrum are performed by measuring the variance (average of squares) of all the measurements within a spherical shell in k-space ($|k| \pm \Delta k/2$); astrophysical foregrounds have a different shape in k-space which enables their removal; and the density of measurements in k-space drives the design of HI cosmology arrays (**?**). This is the same rotational symmetry which leads to averaging over angular m modes in CMB measurements, just translated from the two-dimensional surface of the CMB to the three dimensional volume of HI cosmology measurements. The inherently three-dimensional k-space can be simplified by averaging over the angular wave numbers ($\{k_x, k_y\} \rightarrow k_\perp$) to produce the k line-of-sight ($k_\parallel$) vs. angular $k_\perp$ FoM shown in the left-hand panel of Figure **??**. Of course the spherical symmetry is only approximate—the universe evolves quickly during the EoR and there are velocity effects (e.g., **?**)—but this basic symmetry is what enables all EoR foreground subtraction efforts.

The astrophysical foregrounds are very bright—up to five orders of magnitude stronger than the EoR signal—and in the first few years foreground studies concentrated on identifying all of the potentially offending sources (e.g., **? ? ? ? ? ?**). The general consensus is that all known astrophysical foregrounds are either spectrally smooth or at

---

[3]Interferometric measurements are of the angular Fourier modes in many frequency channels (visibilities), so in the absence of widefield effects only a Fourier transform in the frequency direction and a coordinate mapping would be needed to obtain the 3D $\{k_x, k_y, k_z\}$ measurements (**?**).

known editable frequencies (e.g., radio recombination lines), and none of them mimics the spherical symmetry of the EoR's redshifted emission line (**?**). In the k-space FoM the spectrally smooth astrophysical foreground emission dominates at small $k_{||}$, but quickly falls below the EoR signal strength at higher $k_{||}$ values (Figure **??**, left-hand panel). Conceptually, purely angular modes are dominated by the foregrounds, but the EoR signal can be observed the frequency (line-of-sight) modes (**? ? ? ? ?**).

The difficulty is that no instrument is perfect, and small instrumental and observational effects can throw the strong angular foregrounds into the frequency ($k_{||}$) direction. Effects include chromatic side lobes from sources producing line-of-sight ripples (**? ?**), small calibration errors leading to mis-subtraction of the chromatic PSFs (**?**), Faraday rotation of the galactic synchrotron beating with polarization mis-calibrations (**?**), chromatic primary beams, and numerous other effects. Collectively these are called mode-mixing foregrounds, and have been the primary focus of the EoR foreground subtraction community for the past three years. The right-hand panel of Figure **??** shows an example of mode-mixing from a simulation by (**?**). Due to the chromatic PSF, the point source foreground contamination (visible as a red band at low $k_{||}$) is smeared into the line of sight direction in a distinctive wedge-shaped pattern.

Recent work by (**? ? ? ?**) have identified an 'EoR window' above

$$\mathbf{k}_{||} = \theta'(\mathrm{rad})\ \mathrm{k}_{\perp} \left[\frac{\mathrm{D_M(z)}}{\mathrm{D_H}} \frac{\mathrm{E(z)}}{2\pi(1+\mathrm{z})} \right], \tag{3}$$

that will be largely free of mode-mixing contamination, where θ' is the FoV or horizon depending on details of the technique and the cosmological terms in the brackets equal ~ 0.5.

The approaches by (**? ? ? ?**) are conceptually similar in using the specifics of the instrument to separate the foreground from the EoR PS signal—effectively diagonalizing the instrumental foreground and signal covariance matrix. The boundary of the EoR window given by Equation 3 is not sharp, and the more accurately the foregrounds and instrument can be removed the larger the window over which the EoR signal will dominate the residual contaminants. Expanding the window quickly improves the S/N because the EoR power spectrum is expected to peak at very large scales of $k \approx 10^{-2}$ Mpc^{-1}.

The delay-spectrum technique developed by (**?**) limits the analysis to baselines of equal length. If the antennas have the same beam patterns, this approach is much more immune to small gain calibration errors. For highly redundant arrays there is no sensitivity penalty for using the delay-spectrum technique, but the equal length requirement does impose a sensitivity penalty for arrays with smoother baseline distributions.]

The recent advances in the community in understanding the underlying causes of mode-mixing foregrounds and the identification of the EoR window make predicting the sensitivity of an array including foreground subtraction amenable to detailed simulation (**?**). These advances have changed how we think about foreground subtraction for EoR arrays and have given us more confidence we can overcome the mode-mixing foreground. However, there are still a number of open questions in foreground subtraction including the influence of the ionosphere, the impact of resolved galactic emission on very short baselines, and the mix of conventional and statistical source subtraction.

Lessons & Questions

- There should be an 'EoR window' within which the cosmological signal will be greater than the residual foreground contamination. The size of this window depends on the instrument design and will ultimately determine the sensitivity of different HERA designs.
- Short baselines—until mutual coupling and galactic emission start to dominate—should be both the most sensitive and the most foreground free.

- Does ionospheric refraction and Faraday rotation average out—uniformly blurring and depolarizing sources—or do we need to correct for it dynamically?
- What is the right balance between source subtraction (e.g. deconvolving bright sources) and relying on the statistical symmetries?
- Can the gain calibration been accurate enough to enable the use of all baselines, or should the natural systematic immunity of the delay-spectrum technique be leveraged?

3.4. Array layout & conceptual design

The last stage of developing a conceptual design is determining the desired array layout. The relationships between array layout and theoretical sensitivity have been understood for some time, with a natural emphasis on short baselines and wide fields of view (**? ?**). The remaining antenna layout options are strongly constrained by the observing strategy, calibration approach, and foreground subtraction procedure. These issues have been discussed extensively in the literature (), and the major lessons are:

Lessons & Questions

- Short baselines dominate the EoR PS sensitivity, with the majority of the sensitivity coming from baselines ≤ 100 m.
- Redundant baseline calibration and delay spectrum foreground subtraction favor tiled antenna placement with high redundancy, while astrophysical based calibration favors smooth baseline distributions with more long baselines to help isolate sources.
- Redundant baselines boost the EoR PS sensitivity, at the cost of PSF smoothness.

This is the starting point for developing a conceptual design, and in §4 we will outline the technical path towards a conceptual HERA design. However, there are many real world considerations that go into the design and construction of any array. In the remainder of this section we will detail the practical lessons we have learned. These lessons often have a significant impact on the cost of competing design options, and thus have a very real impact on both the conceptual design and the success of the final arrays.

3.5. Site selection

Just a little text to see if this is really working.

3.6. Collaboration structure—who is doing what

3.7. Prototyping

3.8. Antenna Placement

3.9. Analog systems

3.10. Digital systems

3.11. Software tools & Data analysis

3.12. NRE, cost of complexity, and cost/benefit of flexibility

4. Steps Toward HERA II

4.1. HERA Engineering Initiatives: Pathways towards finding viable solutions

4.2. Steps towards detection

4.3. Steps towards characterization

REFERENCES

Ali, Sk. et al. 2005, MNRAS, 363, 251

Barkana, R. & Loeb, A. 2006, ApJ, 624, L65

Barkana, R., & Loeb, A. 2004, ApJ, 601, 64-69

Barkana, R., Loeb, A. 2005, ApJ, 626, 1-11

Barkana, R. & Loeb, A. 2001, PhysRep, 349, 125

Bharadwaj, S. & Ali, Sk. 2005, MNRAS, 356, 1519-1428

Bowman, J., Rogers, A., Hewitt, J. 2007, ApJ, in press (astroph-0710.2541)

Carilli, C. et al. 2002, ApJ, 577, 22

Ellis, R. 2006, Saas Fe Advanced Course 36 Swiss Soc. Astrophys. Astron. in press.

Fan, X. et al. 2006, AJ, 132, 117

Fan. X., Carilli, C., Keating, B. 2006, ARAA, 44, 415

Furlanetto, S, & Loeb, A. 2002, ApJ, 579, 1-9

Iliev et al. 2006, MNRAS, 369, 1625

Lidz, A. et al. 2007, ApJ, in press (astroph/0711.4373)

Loeb, A., Barkana, R. 2001, ARAA, 39, 19-66

Loeb, A. & Zaldarriaga, M. 2004, Phys.Rev. Lett. 92, 1301-1304

Mcquinn, M. et al. 2006, ApJ, 653, 815

Santos, M., Cooray, A., Knox, L. 2005, ApJ, 625, 575-587

Sethi S. K., Subrahmanyan R., Roshi D. A., 2007, ApJ, 664, 1

Shapiro, P. et al. 2006, ApJ, 646, 681

Wyithe, J.S., Loeb, A., Barnes, D. 2005, ApJ, 634, 715

Wyithe, J.S., Loeb, A., Carilli, C. 2005, ApJ, 628, 575-582

Zahn, O. et al. ApJ, 654, 12

Zaldarigga, M., Furlanetto, S., Henquist, L. 2004, ApJ, 608, 622-635