HERA – Hydrogen Epoch of Reionization Arrays

Donald C. Backer[1], James Aguirre[2], Judd D. Bowman[3], Richard Bradley[4,5], Frank Briggs[6], Chris L. Carilli[4], Steven R. Furlanetto[7], Lincoln J. Greenhill[8], Jacqueline N. Hewitt[9], Colin Lonsdale[10], Miguel F. Morales[11], Aaron Parsons[1], Steven Tingay[12] & Alan Whitney[10]

[1] *University of California, Berkeley; dbacker@astro.berkeley.edu; 510-NGC-5128*

[2] *University of Pennsylvania* [3] *California Institute of Technology*

[4] *National Radio Astronomy Observatory* [5] *University of Virginia*

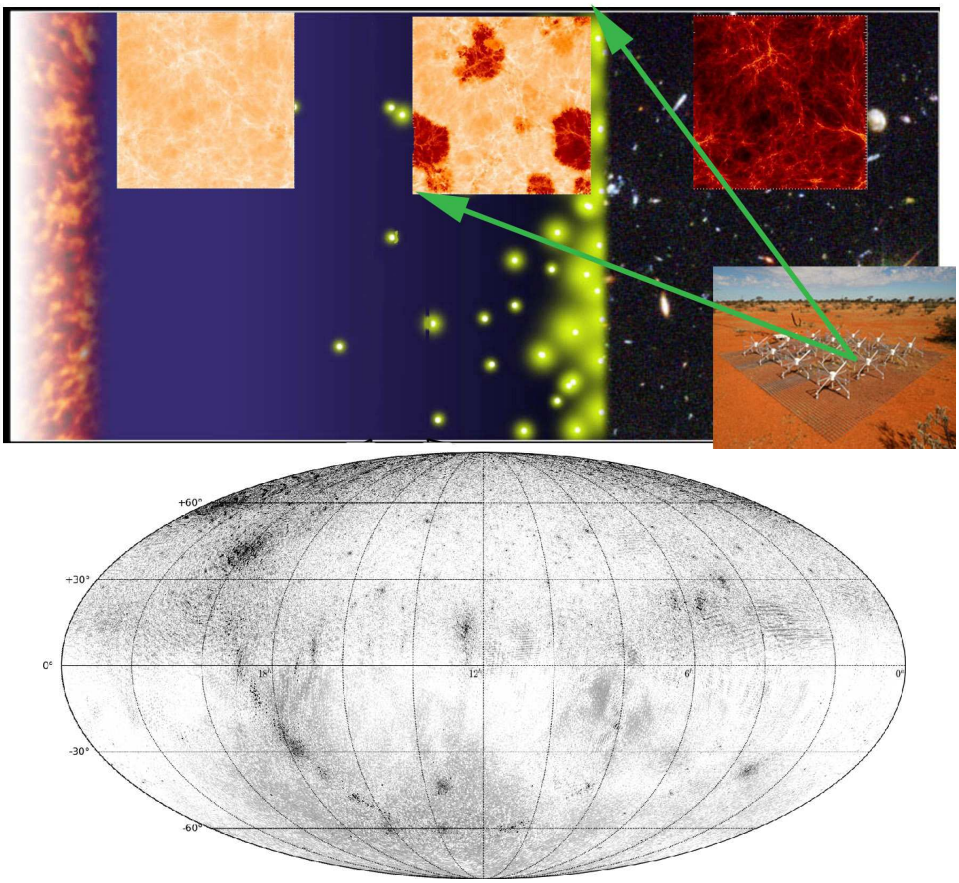
[6] *Australian National University* [7] *University of California, Los Angeles*

[8] *Harvard-Smithsonian Center for Astrophysics*

[9] *Massachusetts Institute of Technology* [10] *Haystack Observatory*

[11] *University of Washington* [12] *Curtin University of Technology*

Submitted for consideration by the Astro2010 Decadal Survey Program Panel
RMS: Radio (Meter/Centimeter) and Millimeter/Submillimeter



Top: Simulated history of the Universe from recombination to the present, with example 21cm signals and a Murchison Widefield Array tile overlaid. Bottom: All-sky map at 150 MHz from Precision Array to Probe the Epoch of Reionization; scale log[Jy]: peak 10^4 Jy; min rms 100 mJy. Credits: S. Furlanetto & J. Lazio, and Parsons et al. (2009).

SUMMARY

US astronomers are in a position of leadership with the first exploration of large-scale structure in the baryonic universe via the 21cm line of hydrogen. We present a phased development approach toward this primary scientific goal using Hydrogen Epoch of Reionization Arrays (HERA) and stepping from: (I) detection of the power spectrum of the 21cm line emission structures; through (II) characterization of the structures including both galaxy formation astrophysics and cosmological physics results in this decade; to (III) detailed imaging of structures across a large fractional step in cosmic time in the following decade. The transition to Phase II will be importantly guided by results obtained early in the decade, and hence a thorough review and opportunity to move ahead then is warranted. Other important and complementary science can be conducted with the current and planned HERA instruments.

The HERA antenna elements are nominally simple, but signal processing, calibration and imaging with large-N arrays are current technological frontiers. Accumulation of experience with each step in the HERA program and parallel development are required to achieve the goals of signal detection, detailed statistical characterization, and structural imaging.

We outline an organizational structure that involves merging of the two US Phase I groups PAPER[1] and the MWA[2], and enhancing international participation. The activity schedule focuses on a \$60M HERA II project involving first construction of an array with 10^5 m^2 effective area to reach to milliKelvin sensitivity at $10'$ angular resolution during 2015-17, and then a science program during 2017-19. The costing of HERA II is solidly based on HERA I (PAPER, MWA projects) development, deployment and use costs during 2009-14. We further state that with the high level of activity and assured progress with these two pathfinding efforts over the next few years, a HERA II proposal mid-decade will be both firmly justified scientifically and accurately costed. A HERA III project started at the end of the decade can inform and mesh with the Square Kilometer Array program low frequency goals.

[1] Precision Array for Probing the Epoch of Reionization; [27]
[2] Murchison Widefield Array; [18]

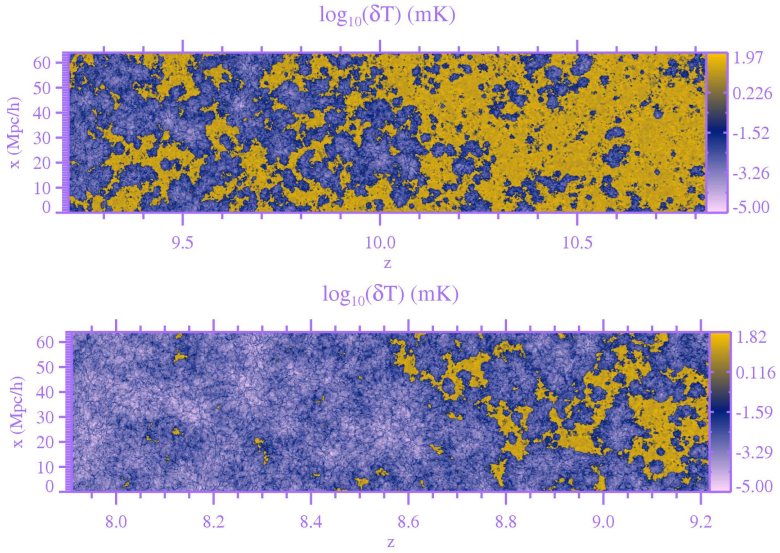


FIG. 1: Simulated maps of the 21cm background during the early and late stages of a particular reionization scenario (top and bottom panels). Blue regions are highly ionized; yellow regions are mostly neutral. In the vertical direction, the slice subtends ~ 35' on the sky. From [28].

I. KEY SCIENCE GOALS OF HERA: HYDROGEN EPOCH OF REIONIZATION ARRAYS

A. Primary Goal: Hydrogen at High Redshift

Introduction. Measurements of our Universe's fundamental parameters and its history have improved enormously over the past twenty years. One of the key remaining challenges is to explore (and exploit) the "high-redshift frontier" at $z > 7$ before and during the formation epochs of the first galaxies. Here we describe a new kind of telescope that can study the otherwise inaccessible cosmic "dark ages" (at $6 < z < 50$, during and before the "reionization" of intergalactic hydrogen) with the 21cm transition of neutral hydrogen. This era includes nearly 60% of the (in principle) observable volume of the Universe. The *potential* for improved measurements of the fundamental cosmological parameters is impressive, with the eventual possibility of, e.g., tightening constraints on our Universe's curvature by two orders of magnitude. It also beautifully complements direct galaxy observations for our understanding of galaxy formation and the intergalactic medium (IGM). The Hydrogen Epoch of Reionization Array (HERA) will address two sets of key questions: **what were the properties of high$-z$ galaxies, and how did they affect the Universe around them?**, and **does the standard cosmological model describe the Universe during the "dark ages?"** These are described more fully in the two science white papers, submitted to the Astro 2010 Decadal Survey, "Astrophysics from the Highly-Redshifted 21cm Line" and "Cosmology from the Highly-Redshifted 21cm Line."

Scientific Context: Observations of HI 21cm emission from the neutral IGM offer a number of unique and powerful probes into the formation of the first galaxies and cosmic reionization [11]. The signal is rich in physical diagnostics, depending on four properties of the IGM: density, neutral fraction, "spin temperature" (i.e., the excitation temperature of the 21cm

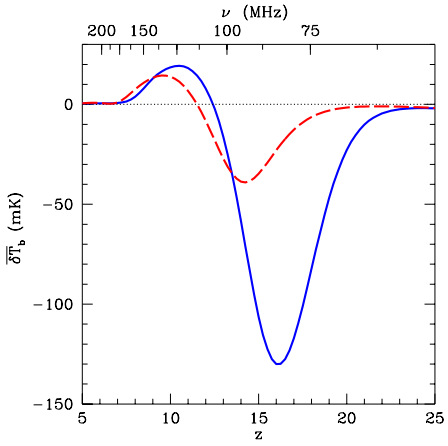


FIG. 2: Fiducial histories of the sky-averaged 21cm brightness temperature, $\delta\bar{T}_b$. The solid blue curve uses a typical Population II star formation history, while the dashed red curve uses only very massive Population III stars. Both fix reionization to end at $z_r \approx 7$. From [10].

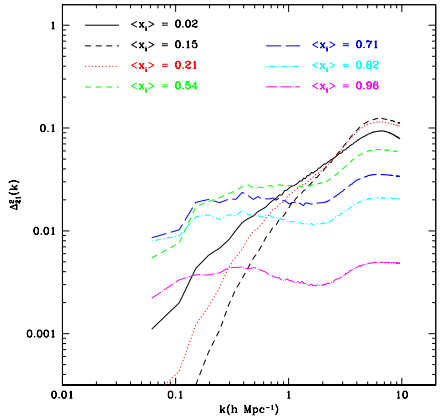


FIG. 3: Evolution of the spherically-averaged 21cm power spectrum during reionization, from a numerical simulation of that process. Here the fluctuations are presented in units of $\delta\bar{T}_b^2$ for a fully neutral universe; $\langle x_i \rangle$ is the ionized fraction during reionization. From [16].

transition), and the local velocity field (sourced primarily by gravity). Because it specifically probes the neutral IGM, this provides the ideal complement to observations of the first galaxies and quasars in the near-IR using eg. JWST, and to CMB polarization studies, which probe the ionized IGM. Moreover, the 21cm line is the only direct probe of the preceding "dark ages," prior to the formation of the first stars and black holes.

Fig. 2 shows two example scenarios for the sky-averaged 21cm brightness temperature during the era of the first galaxies. The solid curve assumes that high-z star formation has similar properties to that at lower redshifts; the dashed curve assumes that these stars were all very massive metal-free objects. In either case, when the first galaxies form, they flood the Universe with ultraviolet photons that break the equilibrium between the CMB and 21cm line – rendering the gas visible in absorption against the CMB [19]. Somewhat later, X-rays – produced by the first supernovae or black holes – heat the IGM, turning this absorption into emission, which fades as ionizing photons from the maturing galaxies destroy the intergalactic HI. The large variations between these scenarios indicate that the 21cm background is quite sensitive to the basic parameters of high-z star formation.

Fig. 1 shows the most dramatic epoch to which the 21cm line is sensitive: the "reionization" epoch where ultraviolet photons from high-z galaxies ionized the entire IGM over a relatively short time interval [7]. This event marked the point when the small fraction of matter inside galaxies completely changed the landscape of the diffuse IGM gas. The 21cm background provides the *ideal* probe of reionization. Its weak oscillator strength (in comparison to Lyα) allows us to penetrate even extremely high redshifts. We can also image it across the entire sky – instead of only rare, isolated Lyα forest lines of sight. Moreover, unlike the CMB, it is a spectral line, and we can easily separate redshift slices to study the full history of the "dark ages." Finally, it directly samples the 95% (or more) of the baryons that reside in the IGM, requiring no difficult inferences about this material from the

properties of the rare luminous galaxies.

In addition to this all-sky 21cm background, the 21cm signal also fluctuates strongly as individual IGM regions grow through gravitational instability and are heated or ionized by luminous sources. These fluctuations can ideally be imaged, but over the short term statistical measurements are likely to be more powerful (see below). Fig. 3 shows how these statistical fluctuations (here parameterized by the power spectrum) evolve throughout reionization. The fluctuations briefly fade as galaxies ionize their dense surroundings in the first stage of reionization, then increase as large ionized bubbles form, finally fading again as the gas is ionized. At higher redshifts, variations in the ultraviolet and X-ray backgrounds, in addition to the normal density fluctuations in the IGM, induce similar structure in the 21cm background. These fluctuations, during and before reionization, are the primary target of HERA, and they will allow us to study two key questions, encompassing both galaxy formation and cosmology.

What were the properties of high$-z$ galaxies, and how did they affect the Universe around them? We can address this from several different directions.

When did reionization occur? The 21cm background evolves strongly throughout reionization and will provide the clearest measurement of the time history of reionization.

What sources were responsible for reionization? The properties of the ionizing sources strongly affect the topology of the neutral IGM and hence have quantifiable impacts on the HI 21cm statistics. We can thus constrain the sources of ionization and gauge the importance of exotic, very massive Population III stars. Of course, exploring this era is a primary goal of many other forthcoming experiments. Comparison with, e.g., galaxy surveys will connect the small-scale physics of galaxies with landmark events encompassing larger scales and illuminate the relation between galaxies and the IGM, through for example the anti-correlation between regions of active galaxy formation (detected with near-IR surveys) and the HI 21cm images [15].

How did the cosmic web evolve? The end of reionization is dominated by the topological transition from ionized bubbles into the "cosmic web" that we see in the Lyα forest. This process also imprints distinctive features on the 21cm signal and allows us to measure directly the evolution of the IGM's structure and its interactions with ionizing sources [6]. Beyond the statistical measurements, HERA has the potential for detecting HI 21cm absorption by the neutral IGM toward the first radio loud AGN. Absorption provides a unique probe of intermediate to small scale structure in the neutral IGM during reionization [9], but detecting it is predicated on the existence of radio loud AGN during reionization [5].

Where did the first generations of quasars form, and what were their properties? Bright quasars form enormous HII regions in the IGM (larger than 5 physical Mpc according to current estimates; [8]), visible in the 21cm sky even after the quasar becomes dormant, because the recombination time is relatively long. Studying these regions in detail will constrain the quasar emission mechanism and their lifetimes, luminosity function, and redshift evolution. There may be hundreds of such regions in the large fields probed by HERA, and they may be sufficiently large for detailed images [31].

When did the first galaxies form, and what were their properties? As mysterious as they are, the galaxies responsible for reionizing the Universe were probably relatively mature. The 21cm transition also probes the very first structures to appear in the Universe – an era

most likely inaccessible even to *JWST* – through the first phase of 21cm absorption in Fig. 2 – and the first black holes, through the heating era.

Does the standard cosmological model describe the Universe during the "dark ages?" A second focus of this experiment is to open an entirely new cosmological era to precision tests. The prospects for "new" physics during this era are difficult to quantify but nevertheless exciting, and 21cm surveys can *potentially* dramatically improve cosmological constraints on parameters such as the inflationary power spectrum, the neutrino mass, and the curvature of the Universe [20].

The challenge will lie in separating astrophysical processes from the "pure" cosmological information, which will require detailed and robust modeling of the astrophysics. From the experimental side, enormous strides in this direction will be enabled by measuring the *redshift-space distortions* of the 21cm background, in which peculiar velocities change the mapping from frequency to radial distance, amplifying fluctuations along the line of sight (but not in the plane of the sky). The peculiar velocities that source these distortions depend almost only on gravity and so provide a purer view of the underlying cosmological parameters. A key goal of HERA will be to measure these redshift-space distortions and to assess their potential utility for cosmological and astrophysical measurements.

Milestones. The ultimate goal of studying the 21cm background is to make detailed maps of the IGM throughout the "dark ages" and reionization, as in Fig. 1. The top axis of Fig. 2 shows the observed frequency range for these measurements: well within the low-frequency radio regime. Unfortunately, this is an extremely challenging band, because of terrestrial interference, ionospheric refraction, and (especially) other astrophysical sources (see [11]). In particular, the polarized Galactic synchrotron foreground has $T_{\rm sky} \sim 180(\nu/180\ {\rm MHz})^{-2.6}$ K, at least four orders of magnitude larger than the signal. For an interferometer, the noise per resolution element (with an angular diameter $\Delta\theta$ and spanning a bandwidth $\Delta\nu$) is then

$$\Delta T_{\rm noise} \sim 20\ {\rm mK} \left(\frac{10^4\ {\rm m}^2}{A_{\rm eff}}\right) \left(\frac{10'}{\Delta\theta}\right)^2 \left(\frac{1+z}{10}\right)^{4.6} \left(\frac{\rm MHz}{\Delta\nu}\frac{100\ {\rm hr}}{t_{\rm int}}\right)^{1/2}, \qquad (1)$$

where $A_{\rm eff}$ is the effective collecting area and $t_{\rm int}$ is the integration time. These angular and frequency scales correspond to ~ 30 Mpc. Figure 4 shows some estimates for how well we can measure the 21cm power spectrum with a variety of experiments.

B. Other Science Goals

While the focus of this submission is the exploration of intergalactic atomic hydrogen at high redshift, the proposed instruments for the coming decade will have significant capability to serve other scientific needs with modest incremental cost. A primary example is the transient radio sky that is a likely byproduct of image analysis for the deep integrations required to detect hydrogen structures. In the Science WP 176, "The Dynamic Radio Sky: An Opportunity for Discovery," Lazio et al. ask, "What New Sources and Phenomena Populate the Sky?" They point out that the time domain of the sky has been only sparsely explored. Recent discoveries indicate that there is much to be found on timescales from nanoseconds to years and at wavelengths from meters to millimeters. These observations have revealed unexpected phenonmena such as rotating radio transients and coherent pulses from brown dwarfs. In addition to the known classes of radio transients, possible other

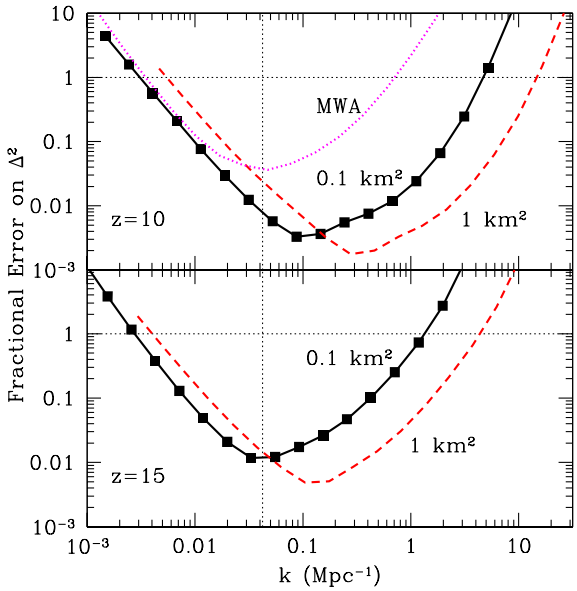


FIG. 4: Sensitivity of three fiducial arrays to the spherically-averaged power spectrum, at $z = 10$ and 15, expressed in units of $\delta\bar{T}_b^2$ for a fully neutral universe, as in Figure 3. The solid and dashed curves are for $A_{\rm eff} = 0.1$ and 1 km^2, respectively. Observing is split over four fields for 1000^h each, with a total bandwidth $B = 8$ MHz, $T_{\rm sky} = (471, 1247)$ K at $z = (10, 15)$, and $N = 5000$ stations centered on a filled core and with an envelope out to $R_{\rm max} = 3$ km. The black squares show the location of the independent k bins. The dotted curve shows the MWA sensitivity at $z = 10$. The vertical dotted line corresponds to the bandwidth; modes with smaller k are compromised by foreground removal.

classes of objects include extrapolations from known clases and exotica such as orphan γ-ray burst afterglows, radio supernovae, tidally-disrupted stars, flare stars, and magnetars. Meter-wavelength transient sources will likely emphasize steep spectrum emission phenomena where correlated particle motions boost emitted signals.

In A2010 Science WP 150, Kasper et al. point out that Low-Frequency Arrays are sensitive to both direct radio emission from coherent plasma processes in the solar corona and to the modification of radiation from background sources by the coronal and heliospheric plasma through Faraday Rotation (FR) and Interplanetary Scintillation (IPS). The greatly improved dynamic range, frequency coverage, and bandwidth of modern arrays will open a new window on the physics of magnetic reconnection and particle acceleration at shocks. The combination of FR and IPS measurements of coronal mass ejections (CMEs) propagating from the corona through interplanetary space has the potential to revolutionize our understanding of how CMEs evolve and to predict the severity of their impact at Earth. The leading science questions that the meter wavelength observations can contribute to are: How does the solar coronal magnetic field change with time? How does the coronal magnetic field extend into interplanetary space? How do Coronal Mass Ejections erupt into the heliosphere? How are particles accelerated by CMEs?

The instrumental properties required for 21cm studies, featuring wide field capability, large-N architecture, provision for high precision full polarimetric calibration, large fractional bandwidths, and high sensitivity, inevitably lead to excellent science potential for these and other non-21cm science investigations. Application-specific constraints on capability will lie largely in the digital domain, and the primary HERA implementations will employ, for example, time and spectral resolutions appropriate to 21cm science. It can nevertheless be appropriate for other science groups to contribute funded effort as warranted to augment the HERA digital hardware, firmware and software for their applications. In this document, we limit our focus to the 21cm reionization and dark ages science.

TABLE I: Technical Progression of Hydrogen Epoch of Reionization Arrays

Instrument	Phase	$A_{\rm eff}$ ($\rm km^2$)	No. Elements	FoV ($^\circ$)	Goal†
PAPER	2010–14	0.003	128-512	60	PS Detection
MWA	2010–14	0.01	512	20-40	PS Detection
HERA II	2015-19	0.1	5000	20-60*	PS Characterization
HERA III	2020–	$O(1)$	$O(50000)$	(20-60)‡	Structure Imaging

†– PS: Power Spectrum. *– Field of view depends on detailed trade-off study following HERA-I efforts. ‡– HERA-III parameters will depend on experience gained by mid-decade.

II. TECHNICAL OVERVIEW OF HERA PROGRAM

The study of cosmic reionization using HI requires instruments with high sensitivity over a wide range of angular scales and redshifts, translating to wide field of view, wide bandwidth, and large collecting area. These instruments must further fully characterize the low-frequency celestial sky and model their own system response so that foregrounds can be effectively suppressed and the EoR signal revealed. This requires calibration of the array with high accuracy. Wide fields-of-view, large fractional bandwidths, radio frequency interference (RFI), and ionospheric variation all complicate the calibration of interferometric arrays. Obtaining an accurate sky model requires accurate models of primary beams, receiver passbands, gain variation, and array geometry. Attaining sufficient collecting area places demands on antenna design, correlator size, and data processing and handling.

These technical challenges are being addressed in phases (II). The development of the technology is driven to advance the science from statistical detection of the EoR fluctuations to full three-dimensional imaging of the high-redshift universe.

HERA I: Power Spectrum Detection. The arrays now operating or under construction have $A_{\rm eff} \sim 10^{3-4}$ m^2 and so are limited to imaging only the most extreme ionized regions (such as those surrounding bright quasars). Nevertheless, these arrays have sufficiently large fields of view ($> 400^{(\circ)2}$) to make reasonably good statistical measurements [3, 21]. Fig. 4 shows the projected errors for the MWA at $z = 10$. The array will be able to detect fluctuations over a limited spatial dynamic range for $z < 12$, constraining the timing of reionization and some source physics.

HERA II: Power Spectrum Characterization. Fig. 4 shows that larger telescopes, with $A_{\rm eff} \sim 10^5$ m^2 (and large fields of view), will clearly be needed for precise measurements, and especially to identify distinctive features of the power spectrum (see Fig. 3). Instruments in this class will also be able to measure more advanced statistics, such as the redshift-space distortions induced by velocity fluctuations. These are extremely useful for breaking degeneracies in the signal [1] but lie beyond the reach of first-generation experiments. (For more information, see the white paper "Cosmology from the Highly-Redshifted 21cm Line".)

HERA III: Hydrogen Structures Imaging. For $A_{\rm eff} \sim 10^6$ m^2, imaging on moderate angular scales becomes possible, and statistical constraints become exquisite even at high

redshifts (provided that the large field of view, not strictly necessary for imaging, is maintained; see Fig. 4). Plans for these later generations will evolve as we learn more about "dark age" physics and the experimental challenges ahead; for example, the Long Wavelength Array, MWA, and LOFAR will study the ionospheric calibration required to explore the high-z regime ($z > 12$, or $\nu < 110$ MHz) and help determine the relative utility of a terrestrial Square Kilometer Array or a far-side Lunar Radio Array (LRA). At the same time, we must explore whether telescope designs that are intended to be closely aligned with the observables, such as an FFT Telescope [29], are both practical and cost-effective.

The HERA roadmap, accompanied by efforts to improve theoretical modeling of the first galaxies and to enhance data analysis techniques (such as specialized statistical measures) will position the community to explore the major science questions concerning high-z galaxy formation, IGM evolution, precision cosmology.

A. HERA I

The first-generation experiments must be extremely focused in order to attain the goal of a convincing detection of the EoR signal. We outline here the technical approaches of MWA and PAPER. These experiments focus on redshifts $z < 12$ and the coldest patches of sky. Both have the sensitivity to obtain statistical detection of the power spectrum of 21cm brightness temperature fluctuations and possibly to image the very rare, largest-scale structures formed at the end of the reionization epoch. They represent differing but in many ways complementary paths to addressing the key issue facing the first generation arrays: calibration. Various methods for estimating the effects of instrumental calibration on a statistical EoR detection have been analyzed [2, 22, 24], but achieving the requisite calibration quality and stability in early instruments remains an open problem [3, 12].

PAPER proceeds with an emphasis on hardware, controlling and understanding the individual antenna responses, so that the *a priori* estimates of the instrument response are of the highest possible quality, thereby minimizing the challenge to the calibration software. MWA emphasizes techniques enabled by High Performance Computing, new adaptable algorithms, high data volume, and a breadth of calibration information intended to solve for instrument response and ionospheric distortion on the fly. MWA also points, and picks out preferred regions of cold sky for EoR detection in order to minimize overall system noise. PAPER observes with a simpler transit strategy, and again emphasizes hardware stability and constancy to maximize quality of *a priori* calibration.

Murchison Widefield Array. The MWA is an 80-300 MHz, large, synthesis array sited in a protected radio-quiet zone centered at Murchison Radio Observatory, Western Australia. The collecting area, field of view (FoV), and imaging characteristics of the MWA are tuned primarily to enable statistical detection of the power spectrum of 21cm brightness fluctuations for redshifts < 12. In addition, the MWA will study the Sun and heliosphere, including direct measurement of the magnetic field structure at different radii, and conduct a systematic sky survey for radio transient emission. See recent image in Fig. 6.

The array comprises 512 tiles of 16 dual polarization dipoles (Fig. 5; 4×4) that are steered electronically by an analog beamformer, achieving a primary beam width of $\sim 30^\circ$ at 200 MHz. This scale is well matched to the sizes of the coldest patches on the sky. Field

of view is a primary design consideration of the array. Beam-forming has the added advantages of increased rejection of out-of-beam power from the complex polarized brightness distribution of the low-frequency sky and rejection of RFI sources at the horizon. Digital, multi-bit, baseband sampling leverages recent advances in electronics, in particular Field Programmable Gate Arrays (FPGAs), to achieve uniform sensitivity, stability, and high-enough dynamic range to maintain linearity even in the presence of interference, as from satellites, aeronautical transmission, and reflections within the atmosphere of distant transmissions. Received signals are combined in a 1024-input FPGA-based correlator capable of $\sim 10^{13}\,\mathrm{s}^{-1}$ cross-multiply and accumulate operations. The native time and frequency resolutions of the output are 0.5 s and 10 kHz, respectively. Polyphase filtration enables spectral dynamic range of 10^5:1 and good performance even when faced with strong narrow-band interference. High time resolution enables tracking of ionospheric distortion of the sky images.

The array is centrally condensed within 1.5 km with a maximum extent of 3 km, to provide optimal sensitivity to the reionization signal (for the chosen A_{eff}) over a range of angular scales on the sky and over a range of redshifts. The synthesized beam is $\sim 2.4'$ at 200 MHz (half-power width). Dense sampling of spatial frequencies, enabled by the large number of apertures and full cross-correlation architecture, provides very good full-polarization point-source response, which is a necessity for characterizing and removing foreground contamination. The MWA is distinctive in its use of a once-through real-time calibration and imaging pipeline as opposed to iterative techniques such as self-calibration. The choice is motivated by the 160 gigabit s^{-1} data rate output by the correlator and practical limitations to storage. The pipeline achieves tile beam and ionospheric calibration using $O(10^2)$ simultaneous calibrator sources, subtracts strong sources from the data stream, constructs full-Stokes images, and and resamples to a storage frame in advance of time averaging every few to 10 minutes.

Precision Array to Probe the Epoch of Reionization. PAPER is a 100-200 MHz transit synthesis array sited at the Murchison Radio Observatory (MRO), with a prototype station in the NRAO radio-quiet zone in Green Bank, West Virigina. PAPER antenna elements are dual-polarization sleeved dipoles mounted above grounding structures with side reflectors (Fig. 7). These elements have been designed for smooth spatial and spectral responses to facilitate calibration. The primary beam has a FWHM of 60° at 150 MHz. The PAPER analog signal path flows from crossed dipole elements attached to a Pseudo-Differential Amplifier, through coaxial cable that runs above ground to a Receiver Card, which band limits signals to 130-185 MHz before transmitting them to the correlator.

A sequence of real-time, digital, FX correlators employing FPGA processors is addressing the growing digital signal processing (DSP) needs in progressively more ambitious PAPER deployments. These correlators are based on the flexible architecture described in [26], whereby DSP engines transmit packetized data through 10-Gbit Ethernet links to commercial switches that are responsible for routing data among boards. This architecture, analog-to-digital converters, modular FPGA-based DSP hardware, and software environment for programming, debugging, and execution were developed in collaboration with the Center for Astronomy Signal Processing and Electronics Research (CASPER) at the University of California, Berkeley. The hardware and firmware developed by CASPER emphasize modularity and scalability, and the flexibility of the correlator design shortens development time, allowing correlators of increasing scale to be developed in parallel to the rapid incremental

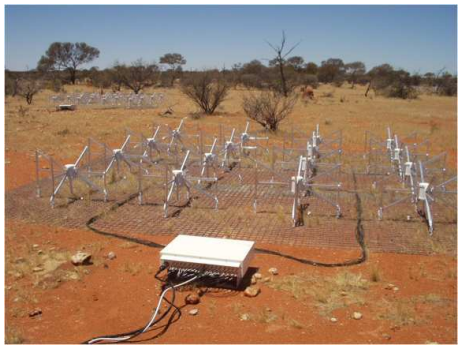


FIG. 5: MWA antenna tile (5m × 5m), consisting of 16 broadband, crossed dipoles with low-noise amplifiers in central hub. Signals are routed to the adjacent analog beamformer; power and coaxial cables to the digital receiver exit at the bottom [18].

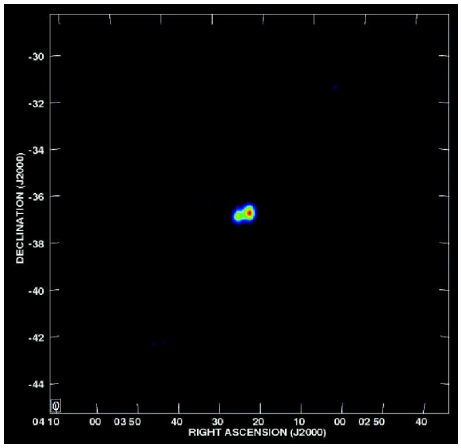


FIG. 6: MWA image of the double-lobed radio galaxy Fornax A at 159 MHz in a 5-minute exposure.

build-out of the array.

PAPER has conducted a test field deployment at MRO (PWA-4) and maintains a site in Green Bank (PGB-4, 8, and currently 16). This allows prototyping to take place quickly under realistic field conditions. The PAPER approach to calibration has been to emphasize precision, with the goal of characterizing system components to within 1% in order to facilitate further model refinement via sky modeling and self-calibration. The PAPER antenna elements are easily re-deployed on their uniform length coaxial cables to test various array configurations.

B. HERA II, III

The larger second-generation experiment (Table I) will be focused on more detailed statistical analyses of the reionization signal. Technical elements will build on experience with systems already in the field. Some will be production elements in use for scientific data acquisition; others will be development prototypes integrated into the arrays for testing and characterization. The central science of HERA II will be power spectrum characterization, the study of redshift-space distortions, pre-reionization cosmology, and early imaging of discrete structures.

Drawing upon lessons learned during HERA-I activity, HERA II will combine PAPER and MWA approaches and technologies (e.g., high correlator data rate will motivate real-time calibration and imaging as in MWA, but with some iterative elements to improve characterization of the instrument). Calibration purity and image dynamic range will be

FIG. 7: PAPER antenna. The ground screen structure includes side reflectors that narrow the size of the primary beam to more closely match the size of colder patches in the synchrotron sky.

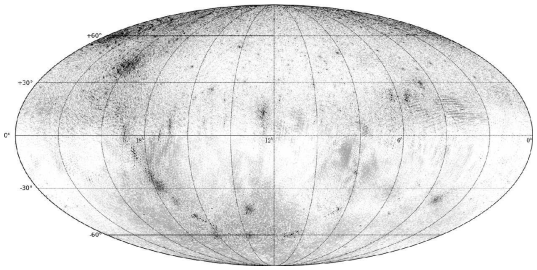


FIG. 8: All-sky map in a band between 138.8 MHz and 174.0 MHz. The northern hemisphere was imaged using PGB-8 data (Dec 38.5°) and the southern hemisphere with PWA-4 (Dec −26.7°) data. Scale log[Jy]: peak 10^4 Jy; min rms 100 mJy. [27].

critical figures of merit (in addition to $A_{\rm eff}$ and FoV). Order of magnitude improvements will be sought to advance foreground subtraction over that achieved with HERA I.

Point-source subtraction will be fostered by likely incorporation of baselines up to the limit ($\sim$3 km) at which ionospheric distortions cannot be approximated by simple refractive shifts and perhaps minor second-order perturbation (focusing and de-focusing). The long baselines will suppress the confusion limit while denser sampling of baselines will reduce sidelobe contamination that also limits the detection floor. (At 150 MHz, there is $\sim$1 source per square degree with flux density greater than 1 Jy, growing more than linearly with the inverse of flux density.) The nature of the faint source population at low frequencies is currently unknown, beyond the scope of HERA I, and will constitute part of the HERA-II discovery space.

Accuracy of instrument polarization calibration will require careful engineering of antenna structures, a development activity to be conducted as part of HERA I operations. Post calibration purity will need to exceed that of HERA I in response to the need to improve subtraction of the intrinsically linearly polarized, diffuse Galactic synchrotron radiation, which has an angular spectrum falling as ℓ^{-3} in other wavebands. The polarization spectrum at low frequencies is at present not well constrained [14, 30], and so predictions of contamination from leakage are correspondingly uncertain. This too represents HERA-II discovery space. Achieving high image fidelity for this diffuse emission will require dense baseline coverage over the full range of baselines (in contrast to MWA where the longest baselines are supported by outrigger tiles) and constraint on missing polarized flux due to the minimum spacing between elements.

At present, the technical outlines of HERA III may be defined at the level of published general specifications for the low-frequency component of the SKA. The technical elements of the system will be described in depth as a result of experience with operations and scientific

campaigns using HERA II, as well as the development work that will be done using that array as a testbed. Several groups [4, 13, 17, 25] have shown that for correctly chosen antenna distributions and careful observing strategy, foreground removal even to the level of the tomography planned for HERA III is possible. HERA II will be suitable to implement and test these theoretical foreground subtraction schemes on very large volumes of data obtained with systematics that approximate the low level anticipated for HERA III.

III. TECHNOLOGY DRIVERS FOR HERA II,III

A. Antennas and Array Configurations

The primary science drivers for HERA require an array with large collecting area and large field of view for survey speed, a large frequency range for redshift coverage, low cost, and precise calibration. The survey speed considerations favor a large number ("large-N") of antenna elements with small aperture. As discussed elsewhere, two complementary large-N designs are being tested as part of the current Phase-I HERA activities, tiles of crossed dipole arrays and individual crossed sleeve dipoles mounted above grounding structures. Both have been designed to have broad spectral response, and cost considerations have been an important factor in both designs. A third class of antenna, units consisting of four phased helical antennas, are being prototyped as a part of design studies for the Lunar Radio Array (Lazio et al. science & program white papers).

The calibration requirements for HERA are extremely stringent. Considering a 30-degree field of view and source counts at 180 MHz we expect to have on average about one 100 Jy source in each field of view. Such a source produces an antenna temperature of about 150,000 K in a synthesized beam of a few arc minutes, leading to a dyamic range requirement of better than 10^6. Propagating visibility phase measurement errors to fluctuations in the map and considering a snapshot with the MWA or PAPER Phase-I array, leads to the requirement that each antenna must have an uncertainty in its phase calibration of order 0.01° or smaller. For the long integrations that will be carried out for EoR measurements, each gridded visibility cell will be formed from the average of many visibility measurements, and the averaging of random phase errors relax the calibration requirement on a single visibility to more than 10°, which is not a challenge at these frequencies. *However, systematic phase errors, that do not average down, have to be controlled at the level of the* 0.01° *requirement.* Also as discussed elsewhere, MWA and PAPER are complementary in their approach to control systematic errors in calibration; the Lunar Radio Array (LRA) approach is similar to that of PAPER in that the antenna element has been designed to have a single-lobed beam that varies smoothly with frequency and angle. A further calibration issue, that has not yet been explored, is the effect of the polarization basis (linear for MWA and PAPER, circular for LRA) on achieving the required dynamic range in the presence of polarized foregrounds.

The optimum configurations for 21cm power spectrum detection and characterization given a total collecting area and element size will continue to be explored as we learn more about calibration and foreground subtraction. This includes unique designs such as fixed viewing of celestial pole and regular gridding.

The technology development for HERA Phases II & III must include several years of study of the performance of the different antenna prototypes and array configurations as they are deployed in the field, observing wide fields dominated by confusing bright sources, and with the data processing carried through the full pipeline of calibration, imaging, and foreground subtraction. These studies started as part of detailed HERA Phase-I pathfinder design, prototyping and simulations. They will continue as a vital part of Phase-I activities that require exact understanding of instrumental effects, and, as a result, make us ready for informed and accurate costing of future phase instruments.

B. Large-N Digital Signal Processing

Phase II of HERA will require a substantial ($\sim 10\times$) increase in collecting area over the Phase-I arrays. The increase in area can be achieved both by increasing the effective collecting area of each antenna element and by increasing the number of antenna elements. The former approach leads to an increase in the cost of antennas, whereas the latter increases the cost of the antennas and digital signal processing (DSP) in the cross-correlation computer for a conventional array architecture. As DSP costs continue to decrease, we may be reasonably certain that future designs will drive us in the direction of large-N arrays. The correlation schemes adopted by all existing large radio interferometers, including MWA and PAPER, have DSP costs that scale quadratically with the number of elements N, once this exceeds a few hundred. Thus, developing technologies for building large-N correlators at reasonable cost is critical.

Calibrating and imaging the correlator output also poses a substantial computational burden at low frequencies. For large-N arrays, data rates can reach levels where the traditional data reduction path of data storage and offline post-processing is no longer viable. For Phase II of HERA, calibration and imaging will demand digital processing comparable in complexity to correlators. Current solutions for real-time imaging and calibration employ CPU- or GPU-based cluster-computing solutions. Other technologies may need to be explored for power-efficient calibration and imaging in large-N arrays.

Specific technical issues that must be addressed include:

- Selection of low-cost digital processors - Current Phase-I activities are based on FPGAs. Developments in FPGA technologies must be monitored and implications for large-N correlator design assessed. In the larger Square Kilometer Array (SKA) program this exploration is being led by the low-frequency array developments, albeit at small bandwidth per N.
- Signal routing - Each cross-correlation processor must receive data from every antenna. For large numbers of antennas, providing cross-communication on this scale is a challenge. Switch-based correlator architectures appear promising [26] and may prove viable for large-N arrays should switch capacity continue to follow the growth of the internet.
- Power consumption - The cost of providing power to large-N correlators and real-time calibration systems is significant. Developments in low-power electronics such as those being carried out for the space program may have implications for the design of large DSP processing systems for Phase II of HERA.
- Algorithms for real-time post-processing - It is not feasible to store the visibility data produced by large-N correlators. Time-variable ionospheric conditions require that calibration solutions be computed over short time intervals before averaging. Algorithms for real-time calibration, wide-field imaging, and architectures for implementing these algorithms must continue to be developed.
- New architectures - By gridding electric field data before it is presented to the correlator, FFT techniques can be employed to change the scaling of computational complexity in correlators from N^2 to $N \log N$. The FFT Telescope [29] and the MOFF correlator [23] are two examples of architectures that merit further exploration.

[1] Barkana, R. & Loeb, A. 2005, ApJ, 624, L65
[2] Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2006, ApJ, 638, 20
[3] —. 2007, ApJ, 661, 1
[4] —. 2008, arXiv:astro-ph/0807.3956
[5] Carilli, C. L., Gnedin, N. Y., & Owen, F. 2002, ApJ, 577, 22
[6] Choudhury, T. R., Ferrara, A., & Gallerani, S. 2008, MNRAS, 385, L58
[7] Fan, X., Carilli, C. L., & Keating, B. 2006, ARAA, 44, 415
[8] Fan, X., Strauss, M. A., Becker, R. H., White, R. L., Gunn, J. E., Knapp, G. R., Richards, G. T., Schneider, D. P., Brinkmann, J., & Fukugita, M. 2006, AJ, 132, 117
[9] Furlanetto, S. R. 2006, MNRAS, 370, 1867
[10] —. 2006, MNRAS, 371, 867
[11] Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, Phys. Rep., 433, 181
[12] Gunst, A. 2007, LOFAR/ASTRON, 52
[13] Jelić, V. & 12 coauthors. 2008, MNRAS, 389, 1319
[14] La Porta, L. & Burigana, C. 2006, A&A, 457, 1
[15] Lidz, A., Zahn, O., Furlanetto, S. R., McQuinn, M., Hernquist, L., & Zaldarriaga, M. 2009, ApJ, 690, 252
[16] Lidz, A., Zahn, O., McQuinn, M., Zaldarriaga, M., & Hernquist, L. 2008, ApJ, 680, 962
[17] Liu, A., Tegmark, M., & Zaldarriaga, M. 2008, arXiv:astro-ph/0807.3952
[18] Lonsdale, C. J., e. a. 2009, arXiv:astro-ph/0604040
[19] Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429
[20] Mao, Y., Tegmark, M., McQuinn, M., Zaldarriaga, M., & Zahn, O. 2008, Phys. Rev. D , 78, 023529
[21] McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., & Furlanetto, S. R. 2006, ApJ, 653, 815
[22] Morales, M. F. 2005, ApJ, 619, 678
[23] —. 2008, arXiv:astro-ph/0812.3669
[24] Morales, M. F., Bowman, J. D., Cappallo, R., Hewitt, J. N., & Lonsdale, C. J. 2006, New Astronomy Review, 50, 173
[25] Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767
[26] Parsons, A., Backer, D., Siemion, A., Chen, H., Werthimer, D., Droz, P., Filiba, T., Manley, J., McMahon, P., Parsa, A., MacMahon, D., & Wright, M. 2008, PASP, 120, 1207
[27] Parsons, A., Backer, D. C., & 13 co-authors. 2009, arXiv:astro-ph/0904.2334v1
[28] Shapiro, P., Iliev, I., Mellema, G., Pen, U.-L., & Merz, H. 2008, in AIP CS, Vol. 1035, The Evolution of Galaxies Through the Neutral Hydrogen Window, ed. R. Minchin & E. Momjian, 68
[29] Tegmark, M. & Zaldarriaga, M. 2008, arXiv:astro-ph/0805.4414
[30] Tucci, M., Carretti, E., Cecchini, S., Fabbri, R., Orsini, M., & Pierpaoli, E. 2000, New Astronomy, 5, 181
[31] Wyithe, J. S. B., Loeb, A., & Barnes, D. G. 2005, ApJ, 634, 715

IV. ACTIVITY ORGANIZATION, PARTNERSHIPS, AND CURRENT STATUS

HERA I. *PAPER* is an NSF-funded collaboration among scientists at UC Berkeley, UVA, NRAO and UPenn in the US and at Curtin University in Perth. PAPER is hosted by the MRO, and receives support from CSIRO staff. Field work with an engineering array in Green Bank, WV allows full-up testing of all components, algorithm development, and training prior to deployments at MRO, where 64 dipoles will be in place in 2009 (and 128 in 2010). *MWA* is pursued by a collaboration among researchers at (i) Haystack, MIT, Smithsonian, and Harvard, supported by NSF and institutional funds; (ii) a consortium of six Australian universities; and (iii) the Raman Research Institute. CSIRO provides logistical and select engineering support. Management is centered at the International Center for Radio Astronomical Research in Perth. Field work is focused at MRO, where a 5% prototype array will be demonstrated, from steerable beam formation to real-time imaging, in mid-2009. Build-out will be completed in 2010.

HERA II. The nucleus of US activity will comprise the PAPER and MWA collaborations. A board-governed consortium will include host-nation organizations and other overseas partner institutions. The Board will work under a memorandum of understanding and provide a conduit to coordinating sponsor organizations. The MWA provides a model for development of governance and management structures. Near-industrial scale design, replication of hardware elements, and signal-processing performance will mandate formal project engineering and management. Management responsibility will lie in a geographically distributed Project Office responding to the Board. Positions will include Director, Sponsor Liason, Engineer, Manager, Scientist, and Regional Project Managers, and Engineers. Integration of regional positions will ensure communication, foster common purpose, and enact lines of spending authority and accountability that parallel funding lines from different sources.

HERA III. Long-term coordination with SKA efforts is anticipated, though HERA science will be more focused to minimize cost. HERA II is anticipated to provide technical, intellectual, and operations leadership for the low-frequency segment of the SKA project, for which it is a recognized pathfinder. Coordination in SKA planning and design with a second pathfinder program, LOFAR, is anticipated, though a critical difference may be the location of HERA II on the actual site of the SKA core, benefitting from related development of infrastructure (SKA siting preference to be set early in the next decade).

Partnerships. The MRO is established and operated by CSIRO, in collaboration with the federal and state governments. It is home to the Australian SKA Pathfinder (ASKAP). Site access is controlled by CSIRO. Curtin University of Technology and the University of Western Australia have formed the International Centre for Radio Astronomy Research (ICRAR), supported by $A20M of state government funding (2009-2013) and ∼$A60M of university contributions. ICRAR has expressed interest in being part of discussions underway in the US regarding HERA and future instruments. The MWA experience, in particular, has shown the great value to international projects of credible local collaborators (e.g., support of site activity, management, scientific cause). ICRAR could be such a collaborator for HERA at MRO. As well, ICRAR aims to provide data-intensive archiving and processing facilities to its staff and collaborators that would benefit HERA and SKA activities. It is noted that CSIRO is the primary party to be consulted in the development of any such plan, and this should occur early, if MRO siting is sought. ICRAR only provides informal advice into these discussions currently and represents its own priorities as a potential collaborator.

V. ACTIVITY SCHEDULE

The decade begins with build out and operations for existing arrays that will also provide many detailed developments required to design HERA II owing to the demanding calibration that must be achieved to meet the 21cm science goals. Mid-decade will see merging of PAPER and MWA efforts in construction of a single array, HERA II, as a culmination of coordinated development efforts. Science operation of HERA II will drive realistic analysis of an SKA-class low-frequency facility on which construction can begin by 2020 (Table II).

TABLE II: HERA Eras

< 2009	2009–2011	2012–2014	2015–2019	2020...
	HERA I			
MWA-32	MWA-512	MWA-512 Science, R&D	**HERA II** (N=5000)	**HERA III** SKA-Lo
PGB-16 PWA-4	PWA-128	PWA-256/512 Science, R&D		
Funding:	\$10M (3 yr)	\$20M (3 yr)	\$60M (5 yr)	\$240M (10 yr)

HERA I: 2009-2014 – Activity begins with completion of build out of PAPER and MWA and *efforts to achieve first detection of the reionization 21cm power spectrum (or upper limits and targetted goals for next stage).* This will severely test array and algorithm design concepts. It will be conducted with current sources of funding and follow-on NSF support via AAG, ATI, and URO programs, interdisciplinary initiatives such as CDI, and opportunities through CISE or ATM divisions. Most importantly, *the two arrays provide a critical running start as testbed facilities supporting development work.* This would improve individual array performances, benefitting scientific productivity, and simultaneously enable assessment of new hardware and software concepts in support of detailed design and engineering of HERA II. Priority development areas will focus on antenna element type and calibratability, correlation techniques and scaling with number of inputs. imaging dynamic range, real-time data post-correlation techniques, and foreground characterization and subtraction.

HERA II: 2015-2019 – Activity continues with HERA-II design and construction, *driven by the study of redshift-space distortions, pre-reionization cosmology, and early imaging of discrete structures.* Elements of PAPER and MWA systems that are relevant to the expanded array architecture (e.g., correlation resources, fiber infrastructure) may be combined. The baseline HERA-II system will feature $10\times$ the collecting area of MWA ($= 0.1\,\mathrm{km}^2$ at 100-200 MHz with as close to full correlation as costs permit. (Less than full correlation of elements defines the hierarchy of station beam forming and constrains field of view.) Technologies available in the near term are anticipated to enable scaling of current PAPER and MWA general approaches. These scale as N^2. HERA II will also enable testing of new concepts that might enable softer, $N \log N$, scaling that would be of critical value in design of SKA-class instruments, though the geometric increase in computing density may enable well-tested, traditional approaches, thus greatly reducing development effort and risk.

HERA III – SKA: 2020... – HERA-III development to achieve 1 km^2 collection area over at least 100-200 MHz will be driven largely by reionization science, *focusing on high-order statistics in HI power spectra and direct wide-field imaging.* Once basic design or operation is secure, transformation over time into a more general-purpose, SKA-type facility

is anticipated, with the addition of capability from other communities.

VI. COST ESTIMATES

Costs through 2020 are grouped around three phases of development (Table II): (I) use and extension of the PAPER and MWA instruments and concomitant work that significantly defrays Phase-II development costs; (II) construction of a 10× larger array by mid-decade whose use and ongoing assessment of instrument parameters will inform ongoing development work for the Phase-III instrument; and (III) an SKA-class instrument at the end of the decade. We note that ratios of development costs to scientific use for the first generation arrays have been moderately large. However, with that completed, we expect that development costs drop to a smaller fraction of construction and operation.

HERA I: *PAPER* costs for the subsystem from an antenna through the analog electronics to the digitization stage is \$2.6K per element in 2009 dollars. The estimate assumes US production and 32 units (as parts kits that can be containerized for shipping). Deployment costs (shipping, travel) are $\sim$ \$20K for 32 units. The total cost to procure and deploy 128 elements in 2010 will be \$410K, with unit quantity discounts approaching 20% for 512 elements.

Digitization and *FX* correlation uses open-source, shared-development design led by the CASPER [26]. A central goal is rapid development of correlator technology, maintaining hardware "on the Moore curve," with firmware that migrates between generations of FPGA devices. In 2009 dollars (K), correlator replication cost is $C_{\rm FX} = (4N + 0.04N^2) * (B/350)$, where N is the number of dual-polarization antennas, and B is the processed bandwidth (MHz). For 128 antennas the correlator cost is \$105K, assuming a fiducial 32-MHz bandwidth that matches MWA (below the current 100-MHz maximum of the PAPER correlator [26]), and fully-costed FPGAs.

Development and integration costs are approximate because of the staged growth of the instrument. Analog subsystem development costs have been distributed, and correlator development costs have been reduced considerably by CASPER. Thus far, additional FTE expenditure for system integration of a 32-antenna array would approximately equal the cost of the analog subsystem; scaled to 128 antennas, this is \$400K. The estimated full cost of a 128-antenna PAPER instrument, $C_{\rm PWA128,full}$, is \$2-3M.

MWA costing is well developed and based on extensive quotes for manufacture, testing and field installation of major and heavily replicated systems. In 2009 dollars (K), the hardware cost, minus development, $C_{\rm MWA,hardware} = 5.1N + 7.6B + 0.011NB + 8.3 \times 10^{-5} N^2 B$, where N is the number of 4×4 antenna tiles and B is bandwidth (MHz). The terms correspond to tiles, receivers, cables, correlation, and realtime computing for calibration and imaging. For $N = 512$ and $B = 32$, the MWA-512 hardware total is \$3.7M.

Implicit scalings breakdown for very large N and maximum baseline lengths beyond a few km, where current system architecture may require substantive modification (e.g., parameterization and correction for ionospheric distortion of images). The limiting N is not known with precision because it depends on advances in commercial technology and declining unit costs (e.g., Xilinx Vertex-4 chips in use now vs order of magnitude more capable Vertex-6 chips due in mid-2009).

Like PAPER, MWA development has been staged, with substantive preparatory work prior to initiation of the project contributing major design elements. This makes the full cost difficult to estimate. A plausible forecast at completion of the 512 element array is $11M for development across the contributing partners and sponsors.

General Infrastructure Costs. For work at MRO, these costs are covered by CSIRO in its site operations model. We assume that infrastructure will continue to be an in-kind contribution intended to enable astronomy at the remote MRO. As a result, central power generators, buildings, internet connectivity and the like have *not* been included in the calculations.

Operations Costs. Because the MWA model includes full-time (24×365) operation, we use it for estimation of related costs. (PAPER is a campaign instrument for which conditions are optimal 3 months per year.)

Fixed dipole arrays have intrinsically low operating costs and follow a different model than traditional radio (or other) observatories (e.g., no allocations are needed for mechanical elements, and cryogenics.) The dominant operations costs are for the stable power supply that is required by digital signal processing and High Performance Computing (HPC) systems, and modular, transportable data storage that is anticipated to service on the order of 50-500 TB per year depending on science application, where periodic site visits are required to change out disk systems. Otherwise, most operation will be controlled remotely, via an existing satellite or future hard link.

An early estimate for diesel fuel required for a budget $O(100)$ kW (c. 2007) was $350K per year for full-time observing ($24^h \times 365^d$). For maintenance of MWA, 2 FTEs based within a few hundred km of the site will be required. The mean time between failure for hardware elements exposed to ambient conditions and for HPC elements are under study. Assuming that engineering anticipates serious failure modes of parts, total annual operations, including site visit costs and administrative details are $700K per year.

Research & Development. In §III we describe important R&D efforts that will be ongoing. Once MWA and PAPER arrays are deployed, technical staff including students can be working on future deployments with improved performance and critical components for future phases. We estimate an average component of $500k per year for this.

Science and Related Costs. Expenses related to reduction and analysis is proportional to data volume and inversely proportional to the time available to complete work. Two labor-intensive activities for both PAPER and MWA will be all-sky surveys to characterize foreground emission and long integrations of selected cold patches on the sky. The latter will support iterative extraction of foreground emission models and statistical analyses to search for a reionization signature. For either array, survey and reionization studies each can readily generate hundreds of TB of raw data per year, with complete analysis desired on an annual cycle. Prior to the start of science operations, cost estimates are necessarily uncertain. MWA planning in the US foresees $O(10)$ FTEs working directly with these data (crudely, a few TB per person per month), with $O(\$500K)$ per year in off-site computing hardware and storage costs. The total is about $2M per year, or $\sim$10%/year of total project cost.

Summary of Phase I 2012-14 (two projects): Capital $8M; Operations $4M; R&D $1.5M; Science $6M; TOTAL $\sim$$20M.

HERA II: Costs for HERA II are dominated by construction and operation. The bulk of development work will be accomplished during the earlier MWA and PAPER programs up to 2014. We adopt a fiducial array of 5000 phased tiles of antennas, which is driven simply by desire for collecting array to meet sensitivity goals of the next phase, and an instantaneous bandwidth of 32 MHz. Experience with MWA and PAPER and development work within those programs will be used to justify deviations, such as a larger number of individual antennas, a somewhat smaller number of larger tiles, or a mixture.

We also assume for this cost estimation that HERA II will be located at the MRO site in Australia. Alternate siting of HERA II is not anticipated unless either the SKA is awarded to a different host nation, or scientific and technical drivers (array design, RFI, etc) present a compelling case for major changes. In that case, we reasonably assume infrastructure costs for US collaboration will continue to be an in-kind contribution from an international partner. Implicitly a US site is not considered at this time owing to strict limit on RFI.

For antenna and digital receiver hardware, we adopt the MWA cost model, represented by the first term in the equation for $C_{\rm MWA,h}$ or \$5K per unit or \$25M. Though HERA-II units will be potentially more sophisticated, most cost will be in development (e.g., propagation modeling, firmware programming), rather than to pay for manufacturing. For correlation, we adopt the geometric mean cost of PAPER and MWA units, expanded to a larger number of inputs, and reduced as deemed appropriate for assumed gains in FPGA processing efficiency: 3 (PAPER) and 10 (MWA). These savings correspond to products close to market today. The total cost is O(\$10M).

The final large cost term for HERA-II construction is real-time calibration and imaging, along the lines of what is being implemented for MWA. Focusing on the last term of the previous cost equation, and conservatively assuming a quadrupling in GPU performance by 2014. The base cost for real-time computing would be \$5M. However, for $N \gg 500$, high data rates and volumes are anticipated to require two very substantial changes in implementation: (i) standard use of 10gE (or faster) network infrastructure for input from the correlator, output to disk, and exchange within the cluster, and (ii) use of integrated high-speed disk storage systems. Roughly, the added impact on real-time processing cost due to networking may be O(\$2M). An added cost of O(\$1M) for 2 PB hosted at a local super computing facility (e.g., IVEC, Perth) is motivated by the likelihood of costs $\sim 3\times$ that of commodity magnetic disk storage today and a 50% discount that reflects technological improvement. The sub-total real-time computing cost is \$8M to which we add 25% contingency to reach \$10M total.

Operations and maintenance are anticipated to be a small fraction of capital costs for HERA II, after the model being applied to PAPER and MWA. We anticipate a 5% levy on resources, though this may be conservative if locally generated renewable energy is available on site in place of diesel-powered generation. The conservative estimate is \$2M per year.

Summary of Phase II 2015-19: Capital ∼\$45M; Operations \$10M; R&D \$3M; Science \$10M; TOTAL ∼\$70M.

HERA III:

Phase II core engineering and technical teams will continue to design and develop for Phase III once production construction is launched for HERA II. We outline our expected areas of ongoing R&D in §III. New ones will certainly arise in this demanding technology

and calibration project. Base support at the level of $1M/year will support this component of our activity.