

The Hydrogen Epoch of Reionization Array (HERA)

David R. DeBoer¹, James Aguirre², Judd Bowman³, Richard Bradley⁴, Chris Carilli⁴, Josh Dillon⁵, Steve Furlanetto⁶, Jacqueline Hewitt⁵, Daniel Jacobs³, Adrian Liu¹, Miguel Morales⁷, Aaron Parsons¹, Jonathan Pober⁷, Max Tegmark⁵, and Dan Werthimer¹

¹University of California, Berkeley, CA USA

²University of Pennsylvania, Philadelphia, PA USA

³Arizona State University, Tempe, AZ USA

⁴National Radio Astronomy Observatory, Charlottesville, VA USA

⁵Massachusetts Institute of Technology, Cambridge, MA USA

⁶University of California, Los Angeles, CA USA

⁷University of Washington, Seattle, WA USA

Correspondence to: David R. DeBoer

(ddeboer@berkeley.edu)

Abstract. Hydrogen Epoch of Reionization Array is a staged experiment that uses the unique properties of the 21-cm line from neutral hydrogen to probe the Epoch of Reionization (EoR) and the preceding Dark Ages. During these epochs, roughly 0.3-1 Gyr after the Big Bang, the first stars and black holes heat and reionize the Universe following cosmic recombination. Direct observation of the large scale structure of reionization and its evolution with time will have a profound impact on our understanding of the birth of the first galaxies and black holes, their influence on the intergalactic medium (IGM), and cosmology. The current stage under development and described in this paper is a compact hexagonal-packed array of 352 14-meter paraboloids operating from about 70 - 220 MHz.

Universe. Direct observation of the large scale structure of reionization and its evolution with time will have a profound impact on our understanding of the birth of the first galaxies and black holes, their influence on the intergalactic medium (IGM), and cosmology.

Detecting, characterizing and ultimately imaging this epoch is a key goal in furthering our understanding of the Universe and was the top priority in the Radio, Millimeter, and Sub-millimeter category of recommended new facilities in the most recent Decadal Survey of Astronomy and Astrophysics of Astronomy and Astrophysics (2010). Current projects (PAPER (Parsons et al., 2014), MWA, (Tingay et al., 2013a), LOFAR (Yatawatta et al., 2013), GMRT (Paciga et al., 2011)) are striving to make the first detection of the statistical power spectrum of the signal, but current best limits still fall above even optimistic predictions of its intrinsic strength. While these projects are still taking data, it is recognized that an optimized array based on our new understanding of the signal characteristics is needed to make a strong detection and begin to characterize this signal over multiple scales and redshifts.

The HERA Roadmap is a phased set of experiments that build on its predecessors and preceding stages. The first phase of the HERA roadmap entailed the operation of the PAPER and MWA telescopes to explore techniques and designs required to detect the primordial HI signal in the presence of radio continuum foreground emission some four orders of magnitude brighter. Studies with PAPER and the MWA have led to a new understanding of the interplay of

1 Introduction

The Hydrogen Epoch of Reionization Array roadmap (HERA Roadmap) is a staged project that uses the unique properties of the 21-cm line of neutral hydrogen to probe the Epoch of Reionization (EoR) and the preceding Dark Ages. During these epochs, roughly 0.3-1 Gyr after the Big Bang, the first stars and black holes heated and reionized the Universe, which at that time existed as a nearly uniform sea of warm neutral hydrogen. Figure 2 is a cartoon showing cosmic evolution from the Big Bang on the left through to HERA today. The period of cosmic reionization represents one of the few global phase changes in the evolution of the

foreground and instrumental systematics in the context of a three-dimensional cosmological intensity-mapping experiment. YYYYYGIVE REFSYYYYY We are now able to remove foregrounds to the limits of our sensitivity with these instruments, culminating in the first physically meaningful upper limits on the power spectrum of 21 cm emission from reionization (see Fig 3).

Building on this understanding, the next stage called the Donald C. Backer Hydrogen Epoch of Reionization Array (HERA) incorporates a new 14m diameter antenna element that is optimized both for sensitivity and for minimizing foreground systematics. Arranging these elements in a compact hexagonal grid yields an array that facilitates calibration, leverages proven foreground removal techniques, and is scalable to large collecting areas. The array will be located near the current PAPER experiment in the radio quiet environment of the SKA site in Karoo, South Africa, and will have a sensitivity close to two orders of magnitude better than PA-PER and the MWA. HERA's sensitivity coupled with broader frequency coverage, will enable HERA to paint an uninterrupted picture through reionization, back to the end of the Dark Ages. As an array, it can be used in smaller arrays along the way to achieve timely science. Two benchmarks are planned:

HERA 127, deployed in 2016, will measure the rise and fall of the 21 cm reionization power spectrum, constraining the timing and duration of reionization.

HERA 331, deployed in 2017, will measure fluctuations in the 21 cm signal over a variety of spatial scales to determine the nature and distribution of the first galaxies that dominate cosmic reionization. HERA 331 will also extend precision power-spectrum observations back to the end of the 'Dark Ages' ($z \sim 20$), when the first stars and black holes warm the primordial IGM.

This paper will present a summary of the current understanding of the signal characteristics and measurements and describe this planned HERA telescope to be built to detect and characterize the EoR power spectrum. After presenting the scientific and motivation and background, the derived specifications and resulting system will be described. The data pipeline will then be summarized, followed by the conclusion.

2 Scientific Motivation

The *cosmic dawn*, the period beginning with the birth of the first stars and culminating with the full reionization of the IGM some 500 Myrs later, represents one of the last unexplored phases in the history of structure formation. During this period, a wealth of astrophysical and cosmological phenomena are at work. The characteristics of the IGM depend on the cosmic density field, the formation sites of the first luminous sources (e.g., their typical masses and clustering), their constituents (e.g., exotic Population III stars, more nor-

mal stars, stellar remnants, or supermassive black holes) their ultraviolet luminosities (which affect the IGM's ionization state), the efficiency and abundance of X-ray sources (which affect the IGM temperature), and even more exotic effects like the relative velocity of baryons and dark matter.

A season of observing with HERA-127 will yield high-significance constraints on the 21 cm power spectrum across a wide range of k modes and redshifts (Pober *et al.*, 2014). In Figure 3 we show the $z = 9$ power spectrum predicted by the publicly available 21cmFAST software (Mesinger *et al.*, 2011), along with 2σ HERA sensitivities. Using the conservative delay-spectrum ("foreground avoidance") approach pioneered by PAPER (Parsons *et al.* 2014), we find that HERA-127 can achieve a $> 10\sigma$ detection of fiducial power spectra over a broad range of redshifts. The subsequent observing season with HERA-331 can increase this detection significance to over 25σ using the same methods. With detailed foreground modeling, the more sophisticated power spectrum estimator developed for the MWA could increase the size of the "EoR window", the region of Fourier space with minimal foreground contamination. This would allow for an overall detection significance of up to 90σ , along with access to lower k modes and therefore qualitatively different physics. Such a high sensitivity measurement would also allow one to go beyond constraining parameters, testing rather than assuming the underlying theoretical framework and starting to image the large neutral bubbles during reionization.

The ability of HERA to image enables an exciting range of cross-correlation science. HERA HI images reveal the large-scale reionization environment for pointed ALMA and JWST observations, and other deep near-IR surveys (Lidz *et al.*, 2009). Knowing whether an observed galaxy is in a region that was previously reionized (center of large HII bubble), recently reionized (edge of HII bubble), or is forming from pristine neutral gas provides important contextual information on early galaxy formation.

The HI images can be cross-correlated with other diffuse tracers of large scale structure. A number of studies have proposed cross-correlation with large scale intensity mapping of molecular CO Lidz *et al.* (2011) and atomic CII (Gong *et al.*, 2011) lines. These studies trace the large scale galaxy distribution – the sources of reionization. Prototypes of these experiments may be operating on the HERA timescale. Such probes have the advantage of having different systematics compared to HERA, potentially allowing clean measurements of the underlying signal.

VERY SHORT DISCUSSION OF WHAT LIMITS THE EOR BETWEEN Z 12- 6

3 Background

3.1 Hubble's Law

Modern cosmology has provided a spectacular understanding of the cosmic history of our Universe starting from the instantaneous moment after the explosion of our spacetime that we call “the Big Bang”, which occurred about 13.8 Gyr ago, to the advent of the large structures that seeded our cosmic current structure. After the Big Bang, matter and space went streaming outward and is manifest to us by the simple expression called Hubble's Law:

$$v = H_o d \quad (1)$$

where v is the velocity of the object away from us, d is the distance to the object H_o is the so-called “Hubble Constant”, which has a value of 68 Mpc/km/sec ??, in astronomer's units, where Mpc denotes megaparsecs, where a parsec is 3.086×10^{13} km. Using measured Doppler shifts one may determine v , so if one knows H_o one may directly compute the distance to the object. Velocity is typically given as a “redshift”, denoted as z , related to one another by the non-relativistic Doppler shift equation $z \sim v/c$, where c is the speed of light in vacuo. And then by Hubble's Law, that redshift conveys a distance.

Note that Hubble's “Constant” is a function of time and using H_o denotes today's value. Determining the evolution of Hubble's constant is essentially the role of cosmological studies and using Hubble's law for very distant sources requires an integration with cosmological evolution encoded within it. The values here are derived from ??.

3.2 Power Spectrum

One tool that is typically employed in studying large scale parameters of the Universe, is to collapse the richness of its spatial variation to its spatial Fourier transform and compute a power spectrum of this. It is the parameters of this power spectrum that may be predicted by cosmology, rather than the essentially random particular spatial variation about it. It is analogous to the fact that one may get a few broad parameters of an FM signal by Fourier transforming its time series, regardless of whether it is Beethoven or Motley Crue, which have very different values as a function of time.

Cosmology typically encodes this data as power in spherical harmonics. Epoch of Reionization studies typically use wavevectors, \mathbf{k} . This 3-D vector may be decomposed into plane-of-sky values (\mathbf{k}_{\perp}) and a line-of-sight term (k_{\parallel}). The overall wavevector may then be written as $\mathbf{k} = \mathbf{k}_{\perp} + k_{\parallel} \mathbf{z}$. As a wavevector, the units of \mathbf{k} are 1/Mpc (inverse length). Note that although modern cosmology believes it has pegged the value of Hubble's constant, to hedge their bets, they will often factor out $h = H_o/100$ as an overall scale adjustment, such that the wavevector “units” are h/Mpc . The units of the

power spectrum magnitudes are given in k-volume-adjusted power-squared (mK²), denoted Δ^2 .

Note that the initial approach to detecting and characterizing the EOR is therefore not imaging, but to detect power in a few of the power spectrum bins. One may exploit this to get more samples of the power in specific bins to beat down the noise, at the expense of imaging, although for very compact arrays there is less of an imaging penalty as more of the bins are filled within a radius of support. The implications of the configuration to the array sensitivity will be touched upon below.

3.3 Foregrounds and Overall Signal Levels

At the frequencies of interest (50-250 MHz), the sky as seen by a radio telescope is dominated by the synchrotron emission from free electrons spiralling around the magnetic fields that thread through our Galaxy. The magnitude of this signal varies as a power law of the frequency, as well as position on the sky (hottest in the plane of the galaxy, cooler at the poles). The variation spans from about 10,000K to about 100K in equivalent temperature. There are a few localized sources in the sky that are even hotter to the eyes of a radio telescope.

As will be shown shortly, the signals of interest have equivalent temperatures of < 10 mK, 7 to 9 orders of magnitude below the foreground signal strengths. The challenge is to peer through these strong foreground signals to see the weak signal of interest below.

As stated above, the strong foreground signals may be accurately described as smooth power laws in frequency - a factor that may be exploited to extract the signals of interest, which have a structure that varies fairly rapidly with frequency.

It is also interesting to note that another “foreground” from our perspective is the Cosmic Microwave Background. In this case, the detectable signal is the difference in flux between the hydrogen and the CMB. It acts a bit like the weatherman's “green screen”, where anything in a specific shade of green is invisible.

3.4 Filtering

The wedge etc

4 Specifications

The epoch of interest, the spatial scales of interest, the required sensitivity and the frequency smoothness provide the scientific specifications needed for the array. Details of the technique, cost, and observing strategies also provide specifications. This section will discuss these high-level specifications.

4.1 Frequency Range

The rest frequency (ν_o) of the hyperfine transition of neutral hydrogen is at 1420.4 MHz. This frequency gets red-shifted to lower frequencies in deep astronomical observations due to the expansion of the Universe, as discussed above. This is characterized by the red-shift parameter z as

$$\nu_s(z) = \nu_o / (1 + z) \quad (2)$$

The age of the Universe at a particular redshift is given by the cosmological model and a key goal of cosmology is to determine those parameters, however existing models allow us to determine these values to our desired accuracy. Figure 2 plots the red-shifted frequency as a function of red-shift (the blue curve), with the upper axis showing the cosmic evolution time since the Big Bang (using parameters from REF PLANCK). The bandwidth for HERA is demarcated by the horizontal dashed lines. The black dashed lines denote the “EOR” core bandwidth, and the extension to the blue dotted lines is to capture earlier epochs (going lower in frequency) and find the post-EOR “null” (going higher in frequency) as discussed below. Figure 2 shows the mapping.

4.2 Bandwidth Field-of-View, and Spatial Scale

As we’ve seen above, the frequency at which the hydrogen transition is observed determines the source’s cosmological distance. Therefore the bandwidth of the observations defines a linear extent along the line-of-sight in space, specified by the frequencies at the end points. As before, one may compute this extent as a function of redshift using a cosmological calculator. Note that at large cosmological distances, relatively narrow bandwidths can imply very large extents, which may have significant cosmological evolution and hence are effectively different samples of the Universe. For HERA, this bandwidth limit is about 10 MHz or less. The cosmological extent for 10 MHz is shown in Figure 4.

Telescopes have a field-of-view of the observed sky, which gives the two orthogonal directions to the line-of-sight. A given diameter and redshift will then have a plane-of-sky extent that with the bandwidth will define a cosmological volume. This plane-of-sky extent for a 14-meter dish is also shown in 4.

4.3 Delay and Systematics

To exploit the smooth spectral response filtering of the foregrounds

4.4 Sensitivity

The sensitivity specification is to maximize performance per cost, which determines the element diameter (D) and the number of elements (N) subject to the constraints above. One therefore needs a model of cost and performance as a function of D and N . Given a fairly mature element and system

design, a bottoms-up costing appropriate for element diameters from about 6-m to 20-m has been done for hex-numbers corresponding to element counts from 37 to 631. The costs here are only those associated with delivering the array to that scope on-site, so excluding development and science.

The equation for sensitivity has been described in previous works [REFS] and depends on many parameters related to the instrumentation, the configuration, the location, the observing strategy, etc. A useful form for the proposed compact configuration is Eq 25 in Parsons et al 2012 and reproduced here:

$$\Delta_N^2(k) \approx 60 \left[\frac{k}{0.1 h \text{Mpc}^{-1}} \right]^{\frac{5}{2}} \left[\frac{6 \text{MHz}}{B} \right]^{\frac{1}{2}} \left[\frac{1}{\Delta \ln k} \right]^{\frac{1}{2}} \times \left[\frac{\Omega}{0.76 \text{sr}} \right] \left[\frac{T_{\text{sys}}}{500 \text{K}} \right]^2 \left[\frac{6 \text{hrs}}{t_{\text{per_day}}} \right]^{\frac{1}{2}} \times \left[\frac{120 \text{days}}{t} \right] \left[\frac{32}{N} \right] \left[\frac{10^4 f_o}{f} \right] \text{mK}^2 \quad (3)$$

where k is the magnitude of the k -mode, B is the bandwidth, $\Delta \ln k$ is the log of the binsize, Ω is the field-of-view, T_{sys} is the system temperature, $t_{\text{per_day}}$ is the number of hours observed per day, t is the number of days observed, N is the number of antennas, and f_o/f is the configuration metric for a redundant array as defined in Parsons et al.

Note that $\Delta_N^2(k)$ is the standard radiometric sensitivity equation, scaled by the volume in k -space, normalized by the power spectrum Fourier coefficient, and reduced by the number of independent samples in a given k -mode bin, which may have both coherent and incoherent application.

Pulling out terms relating to diameter and number, we can write Eq. 3 as

$$\Delta_N^2(k) \propto \frac{\Omega(f_o/f)}{N \sqrt{t_{\text{per_day}}}} \propto \frac{(1/D^2)(1/\sqrt{N})}{N \sqrt{D}} = D^{-\frac{5}{2}} N^{-\frac{3}{2}} \quad (4)$$

where the dependencies on diameter and number have been subsituted in, noting that the expressions for f_o/f and $t_{\text{per_day}}$ were derived in Parsons et al where the baselines for the close-packed array are multiples of the diameter.

For a fixed sensitivity, we may then parameterize the performance/cost ratio by the diameter. Figure 5

5 System Overview

We have achieved a pivotal new understanding of how instrumental characteristics give rise to the wedge of emission shown in Figure ???. Furthermore, we have measurements that prove the efficacy, to the sensitivity limits of current instruments, of projecting out foreground-dominated wave-modes within the wedge. The HI cosmology community is now in a position to define the top-level instrument requirements that ensure foregrounds remain bounded within the

wedge, and to specify the sensitivity needed to obtain high-significance detections of the 21 cm reionization power spectrum under the conservative assumption that all foreground-dominated wavemodes within the wedge must be projected out of our measurements. As summarized in Table 2 (see Pober et al. 2014 for more details), based on these assumptions, current instruments are likely to achieve, at best, only marginal detections of power from reionization.

This proposal targets two arrays: a 127-element array that borrows heavily from components of the PAPER experiment, and an upgraded 331-element array that incorporates several performance optimizations. These arrays are developed over four years (see Fig. ?? and the Project Management Plan for details), with recurring cycles of development, testing, review, deployment, commissioning, and observation. As listed in Table 2, the specifications of HERA-127 and HERA-331 have been set to meet precisely the requirements for obtaining, respectively, a 10σ detection of the 21 cm reionization signal across a broad range of redshifts, and a 25σ detection of the power spectrum of fluctuations in the 21 cm signal capable of determining the nature and distribution of the first galaxies that dominate cosmic reionization. As discussed in §??, these science requirements translate directly into requirements for signal delay and collecting area that, when combined with a cost minimization requirement, tightly bound the HERA design. The basic parameters are given in Table 3.

As shown in Figure 7, the HERA instrument has a straightforward signal path from antenna elements with active feeds, to nodes that digitize and aggregate signals onto an optical network, and on to a processing building where signals are correlated and data are stored, compressed, and shipped to a computing cluster located at U. Pennsylvania (UPenn).

5.1 Antenna Element

The novel design of HERA’s antenna element (Fig. ??, right panel) represents a critical advance that enables HERA to achieve its science goals. This 14-m fixed zenith-pointing parabolic dish yields more than 10 times the sensitivity of an MWA tile (and more than 20 times that of a PAPER element), but does so without substantially degrading our ability to isolate and remove foreground emission on the basis of spectral smoothness. As described above, this is done by limiting the timescale of signal delays and reflections to under 120 ns.

Since the time it takes radio emission to propagate between adjacent antennas represents a significant portion of this time budget, HERA elements must be placed close together. Recent work characterizing foregrounds suggests that antenna separations of $8\lambda \approx 15\text{m}$ are well-behaved to current limits (Pober et al., 2013; Parsons et al., 2013). This influences our choice of a 14-m dish diameter, which incurs 42 ns of signal delay between adjacent antennas. The focal height (4.5 m) derives from the illumination pattern of the prime-focus feed and the fact that reflections between the feed and

the element can introduce additional signal delay. Additional measures, such as a splash cone underneath the feed, screens that isolate feeds from one another, and non-metallic cabling and poles for supporting the feeds, are all aimed at minimizing the potential for additional sources of signal reflections. The total signal delay associated with the HERA dish design is estimated at 83 ns, leaving headroom for reflections arising in the analog signal path.

Cost is another design constraint, including the price of construction materials, assembly in a remote area, and maintenance over the operational lifetime of the array. Care has been taken to select robust and inexpensive construction materials (PVC, concrete, utility poles, 0.25" wire cloth) and an assembly methodology that delivers the required positional accuracy of 10 cm and surface accuracy of 2 cm given standard expertise in construction practices for the subcontracted teams that set the poles and construct the elements in the field. Given the sensitivity and signal delay requirements, the size of HERA dishes optimizes a global costing curve that includes the costs of the elements, the signal path, correlation, data storage, and processing.

A first prototype of the HERA dish is nearly completely constructed. In this proposal, UC Berkeley and NRAO lead the incorporation of lessons from the construction process and reflectometry tests into the design. Two further prototype elements are constructed in the first year alongside the existing PAPER array in Green Bank, WV, for on-sky measurements of the beam pattern and spectral response. These measurements are coupled with full electromagnetic modeling to test and refine the element design prior to a Critical Design Review at the end of the first project year. Three revised elements are then constructed in South Africa, involving lead members of the construction subcontract teams for a Project Readiness Review.

5.2 Array Configuration

Figure 8

For HERA-127 and HERA-331, elements are arranged in a compact hexagonal grid. This configuration minimizes antenna separation, which is critical for meeting the science requirements for foreground isolation, and produces a highly redundant sampling of the uv plane. As described in §??, redundant configurations have been employed by PAPER to boost sensitivity (Parsons et al., 2012), with the added benefit that they greatly facilitate fast and accurate calibration (Liu et al., 2010; Parsons et al., 2013). Placing HERA elements in a grid allows certain construction components to be re-used between antennas, reducing cost and improving the accuracy of element placement. HERA-331 is a build-out of HERA-127 from the eastern edge.

The HERA-331 core has excellent imaging capability that can be leveraged for developing foreground suppression techniques to improve access to the 21 cm reionization signal (Fig. 3, black). This capability is augmented with 21 outrig-

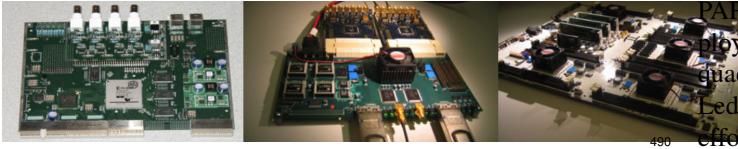


Figure 1. -0.3in

0.99–0.15inFive generations of CASPER technology (progressing left to right) have been used to rapidly develop, test, and deploy digital instrumentation for radio astronomy. This technology allows the PAPER correlator to be easily upgraded for HERA, incorporating new technology (Parsons et al., 2006, 2008).

ger elements. These outriggers combine with the dense core to generate a fully sampled *uv* plane out to 400λ . While outriggers on the southern half of the array are placed in alignment with the hexagonal grid in the core, elements on the northern side are shifted off-grid to sample the *uv* plane at sub-aperture scales. This sampling strategy helps eliminate grating lobes in the synthesized beam and provides information for calibrating and correcting direction-dependent antenna responses. This capability may be used to match polarization beams and minimize polarization leakage.

5.3 Analog Signal Path

HERA’s analog signal path inherits directly from PAPER. In keeping with a philosophy of incremental change, HERA-127 reuses the tested and commissioned feed, LNA, cables, and post-amplifier gain modules (Fig. 6, left and center) that are currently being used to good effect in PAPER-128 (Parsons et al., 2010). Existing PAPER feeds are attached to new reflector screens that are suspended over the elements, but otherwise remain unmodified. A parallel development effort in NRAO aims to improve the feed response below 100 MHz and to improve the match between polarization beams in the feed-element system. Additional minor modifications to the analog signal path include transitioning from a 75Ω to a shorter 50Ω cabling system and replacing 100–200 MHz bandpass filters with 50–250 MHz equivalents. This wider analog bandwidth is used in HERA analysis to improve the extrapolation of smooth foreground emission over a broader range, and to explore HERA’s capability as a Dark-Ages science instrument. After a Critical Design Review, the results of these development efforts are incorporated in the transition to the HERA-331 system.

5.4 Digital System

Although correlators have historically been one of the most complex, expensive, and risky aspects of developing a radio interferometer, this is no longer the case. CASPER (Parsons et al., 2006) open-sourced the development of digital signal processing engines for astronomy and now has world-wide participation, with over 500 members at 73 institutions, and five generations of hardware (Fig. 1). On a modest budget,

PAPER has applied CASPER technology to develop and deploy new correlators annually for five years running, each quadrupling the computational capacity of its predecessor. Led at UC Berkeley’s Radio Astronomy Lab (RAL), HERA efforts continue this incremental development cycle, following a packet-switched correlator architecture (Parsons et al., 2008) that has been extended in recent PAPER and LEDA deployments (Fig. 6, right) to leverage the computing strengths of both FPGAs and GPUs (Clark et al., 2011).

While HERA-127 uses the existing PAPER correlator directly, this correlator architecture evolves for HERA-331. As discussed in previous sections, HERA’s science requirements dictate a maximum analog signal path-length. As a consequence, digitization needs to happen close to the antenna elements in the field. This specification, along with a growing need for modularity to scale with the number of parallel signal paths, leads HERA to adopt a node-based architecture for amplification, digitization, channelization, and digital transmission in the field that builds on HERA’s MWA heritage. This architecture is merged with PAPER’s clean architecture for real-sampling and channelizing the entire analog passband at once, packetizing the data into 10 Gb Ethernet format, and relying on commercial switches to perform the frequency/antenna corner-turn that FX correlator architectures require.

Node. HERA-331 employs RFI-tight node enclosures that each contain the final gain and digitization stages for signals from 18 antennas, along with power supplies, cooling, and a small server for monitor/control. As part of HERA development, a new board called the Smart Network ADC Processor (SNAP) board is incorporated into the CASPER suite of hardware and firmware. This inexpensive board was co-designed by UC Berkeley and NRAO to be both the digitizer and F-engine in HERA’s FX correlator architecture, and is currently in layout at NRAO. Each SNAP board digitizes and channelizes a 50–250 MHz band for 6 input signals (3 antennas, dual-polarization). A 100-MHz band of selectable channels is transmitted over optical fiber to a central container (see below), and on to the Karoo Array Processing Building (KAPB). The development and integration of the SNAP design and the node system is led at UC Berkeley and after a Critical Design Review, is fabricated under subcontracts to industry partners in the third project year and deployed as part of the HERA-331 system. Development activities under this proposal include porting the CASPER toolflow to the SNAP board, designing and testing the FPGA firmware, integrating and testing all components in the node subsystem, and providing a monitor/control access interface. If time allows, optional development may include doubling the transmitted bandwidth.

Central Container. HERA’s central container houses two significant subsystems adjacent to the array. The first is a timing subsystem that maintains a GPS-disciplined oscillator and distributes timing signals (the sampling clock and 1PPS synchronization) to the nodes. The second subsystem

is a passive fiber optic patch panel that couples the optical network from the nodes into the 192-filament optical fiber bundle that connects to the KAPB.

Karoo Array Processing Building (KAPB). The KAPB is currently in advanced stages of construction for MeerKAT,⁵⁴⁵ and houses the switch and processors that complete the HERA correlator system. The fiber optic bundle that enters the KAPB patches into local fiber optic cables that each terminate in optical transceivers that plug into a 240-port 10 GbE switch. Such switches, while large, are readily available commercially today. Also connected to this switch are 30 servers, each hosting two dual-GPU graphics cards and two dual 10 GbE network interface cards, which implement the cross-multiplication (X-Engine) component of the correlator during observations. This estimate for the number of X-⁵⁵⁵ Engine servers is extrapolated from current GPU servers deployed on PAPER, assuming no improvement in bus speeds for transferring data into the GPU cores, but assuming that the computational capacity of such GPU cores doubles according to Moore's Law prior to the purchase of these servers⁵⁶⁰ in the third year of the project. Output data from the correlator are written to the data storage system described in the following section.

5.5 Data Storage, Compression, Transfer, and Computing

HERA's data management system is responsible for recording raw data from the correlator, compressing that data in real-time, applying routine calibration and analysis pipelines, and transferring data products to a high-performance computing cluster at UPenn. UPenn leads the procurement and⁵⁶⁵ deployment of three data storage systems:

- a 1.5 PB system deployed in the KAPB for archiving all 1.2 PB of raw data and 60 TB of compressed data,
- 6 network attached storage (NAS) units plus 2 125 TB RAID array units that are used to ship compressed data⁵⁷⁵ products to the US, and
- a permanent 250 TB system at UPenn associated with a computing cluster that holds compressed data products and serves as the analysis engine for HERA collaborators.

This effort leverages existing infrastructure at UPenn, with support for the expansion of the data storage and for upgrading to a 30-node computing cluster. This cluster supports the bulk of the analysis by HERA collaborators (see §5.7) that⁶³⁵ requires access to the full set of HERA observations.

The data compression scheme at the heart of the data management system has been implemented for PAPER (Parsons et al. 2013, Appendix A), and is applied to HERA visibilities to reduce data volume by a factor of ~ 20 without impacting reionization science capabilities. This compression

technique, which is based on delay/delay-rate filtering (Parsons and Backer, 2009), is applied uniformly to all visibilities in the array, does not require (or produce) detailed calibration information, and is minimally restrictive for how data are analyzed and calibrated afterward. Data compression is run on the same GPU servers that implement the correlator X-Engines (§5.4). Since HERA only observes at night, these processors would otherwise be unused. UC Berkeley is responsible for porting the existing data compression pipeline to target these servers.

Data quality assurance (QA) is performed in real time on an additional modest computing cluster in the KAPB. UPenn leads the deployment and support of the hardware system that manages data transfer, applies routine calibration pipelines (see §5.7) and quality checks, and aggregates correlation-based metrics of array performance in real-time. The QA system furnishes this information into the separate monitor and control system (§5.6). A modest amount of data transfer is possible over the internet; PAPER's data transfer rate from the Karoo to UPenn varies, but peaks around 40 Mbps. The QA system drives internet data transfer, but NAS devices and RAID storage, transferred by air-freight shipping, ensure full data transfer to UPenn.

5.6 Monitor and Control

The U. of Washington team leads the development of HERA's Monitor and Control (M&C) system, which is a straightforward port of a similar system used on the MWA (Tingay et al., 2013b). This system is responsible for tracking observing status, array startup and shutdown, and monitoring all active HERA systems. In the process, the M&C system builds a sizeable database of metadata that is crucial for verifying system functionality, identifying hardware failures, and feeding calibration and contextual information into data analysis pipelines.

5.7 Software, Analysis, and Science

Beyond the construction and data-taking aspects of HERA, this proposal incorporates a full data analysis effort, culminating in the publication of a suite of science papers connecting observations to the physics of cosmic dawn. These efforts leverage existing software pipelines, with on-going development driven by students and postdocs targeting specific science goals. On the science side, HERA leverages the involvement of a team of theory collaborators, including Furlanetto, Lidz, Loeb, McQuinn, Mesinger, Oh, Pritchard, Santos, and Sutter.

Calibration and Snapshot Imaging. MIT leads the development of a real-time redundancy-based calibration pipeline based on related MITEoR and PAPER efforts. These instantaneous calibration solutions are provided to the Monitor and Control system to enable hardware misbehaviors to be quickly identified, and also support the real-time imaging

pipeline. Absolute calibration is fixed with a combination of self-calibration techniques and absolutely calibrated baluns developed at ASU and NRAO. Offline, detailed imaging efforts feed into developing empirical beam models that complement electromagnetic simulations. UPenn develops polarization beam models that determine whether polarization leakage can be formally retired as a risk and lay the foundation for imaging-based foreground suppression that corrects any remaining leakage effects.

Foreground Modeling. The U. Washington team leads the adaptation of a high dynamic-range imaging pipeline based on the Fast Holographic Deconvolution (FHD; Sullivan et al. 2012), for HERA. The full-Stokes sky maps resulting from this pipeline are used to directly subtract foregrounds and as data products themselves. UPenn leads the characterization of the polarized sky and the distribution of rotation measures. Data from HERA, PAPER, and the MWA are used to update source catalogs and a Global Sky Model (de Oliveira-Costa et al., 2008).

Power Spectra. UC Berkeley leads early power spectrum measurements using the conservative delay-spectrum approach used in PAPER. Development of the power-spectrum pipeline will focus on incorporating a fully covariant description of the foreground wedge (Liu and Tegmark, 2011; Dillon et al., 2013), as well as exploring Bayesian techniques for power spectrum estimation, building on the considerable progress already made on Gibbs-sampling imaging for interferometers (Sutter et al., 2014). In addition to the statistical foreground mitigation techniques provided by optimal quadratic power-spectrum estimation, U. Washington leads the application of foreground modeling for expanding the range of modes available for power spectrum analysis. Techniques developed should be equally applicable to measurements in both the reionization epoch and at higher redshifts. MIT leads the application of these tools for pre-reionization science.

Simulation. To provide verification of power spectrum constraints, ASU leads the development of an end-to-end simulator of the full HERA instrument, incorporating foregrounds and reionization models to output visibilities. UCLA also leads a parametrized reionization simulation effort that, coupled with the instrument simulator, is critical for connecting power-spectrum constraints to the estimation of the underlying astrophysical models.

HI Imaging. In addition to supplying models for suppressing foregrounds, deep imaging aims to map HI emission during reionization (see §2). Initially, conservative filtering of the foreground wedge is applied to data that are imaged. These results are compared with the residuals from imaging-based foreground suppression to explore the limitations of precision foreground imaging and removal. As imaging-based techniques mature, the resultant maps are released for cross-correlation with other probes of reionization.

6 Conclusions

And sugar and spice

Acknowledgements. TEXT

References

- Clark, M. A., La Plante, P. C., and Greenhill, L. J.: Accelerating Radio Astronomy Cross-Correlation with Graphics Processing Units, ArXiv e-prints, 2011.
- de Oliveira-Costa, A., Tegmark, M., Gaensler, B. M., Jonas, J., Landecker, T. L., and Reich, P.: A model of diffuse Galactic radio emission from 10 MHz to 100 GHz, Month. Not. Royal Astron. Soc., 388, 247–260, doi:10.1111/j.1365-2966.2008.13376.x, 2008.
- Dillon, J. S., Liu, A., and Tegmark, M.: A fast method for power spectrum and foreground analysis for 21 cm cosmology, Unknown, 87, 043005, doi:10.1103/PhysRevD.87.043005, 2013.
- Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., and Lubin, P.: Probing Reionization with Intensity Mapping of Molecular and Fine-structure Lines, Astrophys. J. Lett., 728, L46, doi:10.1088/2041-8205/728/2/L46, 2011.
- Lidz, A., Zahn, O., Furlanetto, S. R., McQuinn, M., Hernquist, L., and Zaldarriaga, M.: Probing Reionization with the 21 cm Galaxy Cross-Power Spectrum, Astrophys. J., 690, 252–266, doi:10.1088/0004-637X/690/1/252, 2009.
- Lidz, A., Furlanetto, S. R., Oh, S. P., Aguirre, J., Chang, T.-C., Doré, O., and Pritchard, J. R.: Intensity Mapping with Carbon Monoxide Emission Lines and the Redshifted 21 cm Line, Astrophys. J., 741, 70, doi:10.1088/0004-637X/741/2/70, 2011.
- Liu, A. and Tegmark, M.: A method for 21 cm power spectrum estimation in the presence of foregrounds, Unknown, 83, 103006, doi:10.1103/PhysRevD.83.103006, 2011.
- Liu, A., Tegmark, M., Morrison, S., Lutomirski, A., and Zaldarriaga, M.: Precision calibration of radio interferometers using redundant baselines, Month. Not. Royal Astron. Soc., 408, 1029–1050, doi:10.1111/j.1365-2966.2010.17174.x, 2010.
- Mesinger, A., Furlanetto, S., and Cen, R.: 21CMFAST: a fast, seminumerical simulation of the high-redshift 21-cm signal, Month. Not. Royal Astron. Soc., 411, 955–972, doi:10.1111/j.1365-2966.2010.17731.x, 2011.
- of Astronomy, D. S. and Astrophysics: New Worlds, New Horizons in Astronomy and Astrophysics, Committee of the National Research Council, The National Academies Press, http://www.nap.edu/openbook.php?record_id=12951, 2010.
- Paciga, G., Chang, T.-C., Gupta, Y., Nityananda, R., Odegova, J., Pen, U.-L., Peterson, J. B., Roy, J., and Sigurdson, K.: The GMRT Epoch of Reionization experiment: a new upper limit on the neutral hydrogen power spectrum at $z \approx 8.6$, Month. Not. Royal Astron. Soc., 413, 1174–1183, doi:10.1111/j.1365-2966.2011.18208.x, 2011.
- Parsons, A., Backer, D., Chang, C., Chapman, D., Chen, H., Crescini, P., de Jesus, C., Dick, C., Droz, P., MacMahon, D., Meder, K., Mock, J., Nagpal, V., Nikolic, B., Parsa, A., Richards, B., Siemion, A., Wawrzynek, J., Werthimer, D., and Wright, M.: PetaOp/Second FPGA Signal Processing for SETI and Radio As-

- tronomy, in: Asilomar Conference on Signals and Systems, Pacific Grove, CA, pp. 2031–2035, 2006.
- Parsons, A., Backer, D., Siemion, A., Chen, H., Werthimer, D., Droz, P., Filiba, T., Manley, J., McMahon, P., Parsa, A., MacMahon, D., and Wright, M.: A Scalable Correlator Architecture Based on Modular FPGA Hardware, Reuseable Gateware, and Data Packetization, *Proc. Astron. Soc. Pac.*, 120, 1207–1221, doi:10.1086/593053, 2008.
- Parsons, A., Pober, J., McQuinn, M., Jacobs, D., and Aguirre, J.: A Sensitivity and Array-configuration Study for Measuring the Power Spectrum of 21 cm Emission from Reionization, *Astrophys. J.*, 753, 81, doi:10.1088/0004-637X/753/1/81, 2012.
- Parsons, A. R. and Backer, D. C.: Calibration of Low-Frequency, Wide-Field Radio Interferometers Using Delay/Delay-Rate Filtering, *Astron. J.*, 138, 219–226, doi:10.1088/0004-6256/138/1/219, 2009.
- Parsons, A. R., Backer, D. C., Foster, G. S., Wright, M. C. H., Bradley, R. F., Gugliucci, N. E., Parashare, C. R., Benoit, E. E., Aguirre, J. E., Jacobs, D. C., Carilli, C. L., Herne, D., Lynch, M. J., Manley, J. R., and Werthimer, D. J.: The Precision Array for Probing the Epoch of Re-ionization: Eight Station Results, *Astron. J.*, 139, 1468–1480, doi:10.1088/0004-6256/139/4/1468, 2010.
- Parsons, A. R., Liu, A., Aguirre, J. E., Ali, Z. S., Bradley, R. F., Carilli, C. L., DeBoer, D. R., Dexter, M. R., Gugliucci, N. E., Jacobs, D. C., Klima, P., MacMahon, D. H. E., Manley, J. R., Moore, D. F., Pober, J. C., Stefan, I. I., and Walbrugh, W. P.: New Limits on 21cm EoR From PAPER-32 Consistent with an X-Ray Heated IGM at z=7.7, ArXiv, 2013.
- Parsons, A. R., Liu, A., Aguirre, J. E., Ali, Z. S., Bradley, R. F., Carilli, C. L., DeBoer, D. R., Dexter, M. R., Gugliucci, N. E., Jacobs, D. C., Klima, P., MacMahon, D. H. E., Manley, J. R., Moore, D. F., Pober, J. C., Stefan, I. I., and Walbrugh, W. P.: New Limits on 21 cm Epoch of Reionization from PAPER-32 Consistent with an X-Ray Heated Intergalactic Medium at z = 7.7, *Astrophys. J.*, 788, 106, doi:10.1088/0004-637X/788/2/106, 2014.
- Pober, J. C., Parsons, A. R., Aguirre, J. E., Ali, Z., Bradley, R. F., Carilli, C. L., DeBoer, D., Dexter, M., Gugliucci, N. E., Jacobs, D. C., Klima, P. J., MacMahon, D., Manley, J., Moore, D. F., Stefan, I. I., and Walbrugh, W. P.: Opening the 21 cm Epoch of Reionization Window: Measurements of Foreground Isolation with PAPER, *Astrophys. J. Lett.*, 768, L36, doi:10.1088/2041-8205/768/2/L36, 2013.
- Pober, J. C., Liu, A., Dillon, J. S., Aguirre, J. E., Bowman, J. D., Bradley, R. F., Carilli, C. L., DeBoer, D. R., Hewitt, J. N., Jacobs, D. C., McQuinn, M., Morales, M. F., Parsons, A. R., Tegmark, M., and Werthimer, D. J.: What Next-generation 21 cm Power Spectrum Measurements can Teach us About the Epoch of Reionization, *Astrophys. J.*, 782, 66, doi:10.1088/0004-637X/782/2/66, 2014.
- Sullivan, I. S., Morales, M. F., Hazelton, B. J., Arcus, W., Barnes, D., Bernardi, G., Briggs, F. H., Bowman, J. D., Bunton, J. D., Cappallo, R. J., Corey, B. E., Deshpande, A., deSouza, L., Emrich, D., Gaensler, B. M., Goeke, R., Greenhill, L. J., Herne, D., Hewitt, J. N., Johnston-Hollitt, M., Kaplan, D. L., Kasper, J. C., Kincaid, B. B., Koenig, R., Kratzberg, E., Lynch, M. J., McKinley, B., McWhirter, S. R., Morgan, E., Oberoi, D., Pathikulangara, J., Prabu, T., Remillard, R. A., Rogers, A. E. E., Roshi, A., Salah, J. E., Sault, R. J., Udaya-Shankar, N., Schlagenhauf, F., Srihani, K. S., Stevens, J., Subrahmanyam, R., Waterson, M., Webster, R. L., Whitney, A. R., Williams, A., Williams, C. L., and Wyithe, J. S. B.: The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies, *Proc. Astron. Soc. Aust.*, 30, e007, doi:10.1017/pasa.2012.007, 2013a.
- Tingay, S. J., Goeke, R., Bowman, J. D., Emrich, D., Ord, S. M., Mitchell, D. A., Morales, M. F., Booler, T., Crosse, B., Wayth, R. B., Lonsdale, C. J., Tremblay, S., Pallot, D., Colegate, T., Wicenec, A., Kudryavtseva, N., Arcus, W., Barnes, D., Bernardi, G., Briggs, F., Burns, S., Bunton, J. D., Cappallo, R. J., Corey, B. E., Deshpande, A., Desouza, L., Gaensler, B. M., Greenhill, L. J., Hall, P. J., Hazelton, B. J., Herne, D., Hewitt, J. N., Johnston-Hollitt, M., Kaplan, D. L., Kasper, J. C., Kincaid, B. B., Koenig, R., Kratzberg, E., Lynch, M. J., McKinley, B., McWhirter, S. R., Morgan, E., Oberoi, D., Pathikulangara, J., Prabu, T., Remillard, R. A., Rogers, A. E. E., Roshi, A., Salah, J. E., Sault, R. J., Udaya-Shankar, N., Schlagenhauf, F., Srihani, K. S., Stevens, J., Subrahmanyam, R., Waterson, M., Webster, R. L., Whitney, A. R., Williams, A., Williams, C. L., and Wyithe, J. S. B.: The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies, *Proc. Astron. Soc. Aust.*, 30, e007, doi:10.1017/pasa.2012.007, 2013b.
- Tingay, S. J., Goeke, R., Bowman, J. D., Emrich, D., Ord, S. M., Mitchell, D. A., Morales, M. F., Booler, T., Crosse, B., Wayth, R. B., Lonsdale, C. J., Tremblay, S., Pallot, D., Colegate, T., Wicenec, A., Kudryavtseva, N., Arcus, W., Barnes, D., Bernardi, G., Briggs, F., Burns, S., Bunton, J. D., Cappallo, R. J., Corey, B. E., Deshpande, A., Desouza, L., Gaensler, B. M., Greenhill, L. J., Hall, P. J., Hazelton, B. J., Herne, D., Hewitt, J. N., Johnston-Hollitt, M., Kaplan, D. L., Kasper, J. C., Kincaid, B. B., Koenig, R., Kratzberg, E., Lynch, M. J., McKinley, B., McWhirter, S. R., Morgan, E., Oberoi, D., Pathikulangara, J., Prabu, T., Remillard, R. A., Rogers, A. E. E., Roshi, A., Salah, J. E., Sault, R. J., Udaya-Shankar, N., Schlagenhauf, F., Srihani, K. S., Stevens, J., Subrahmanyam, R., Waterson, M., Webster, R. L., Whitney, A. R., Williams, A., Williams, C. L., and Wyithe, J. S. B.: The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies, *Proc. Astron. Soc. Aust.*, 30, e007, doi:10.1017/pasa.2012.007, 2013b.
- Yatawatta, S., de Bruyn, A. G., Brentjens, M. A., Labropoulos, P., Pandey, V. N., Kazemi, S., Zaroubi, S., Koopmans, L. V. E., Offringa, A. R., Jelić, V., Martinez Rubi, O., Veligatla, V., Wijnholds, S. J., Brouw, W. N., Bernardi, G., Ciardi, B., Daiboo, S., Harker, G., Mellema, G., Schaye, J., Thomas, R., Vedantham, H., Chapman, E., Abdalla, F. B., Alexov, A., Anderson, J., Avruch, I. M., Batejat, F., Bell, M. E., Bell, M. R., Bentum, M., Best, P., Bonafede, A., Bregman, J., Breitling, F., van de Brink, R. H., Broderick, J. W., Brüggen, M., Conway, J., de Gasperin, F., de Geus, E., Duscha, S., Falcke, H., Fallows, R. A., Ferrari, C., Frieswijk, W., Garrett, M. A., Griessmeier, J. M., Gunst, A. W., Hassall, T. E., Hessels, J. W. T., Hoeft, M., Iacobelli, M., Juette, E., Karastergiou, A., Kondratiev, V. I., Kramer,

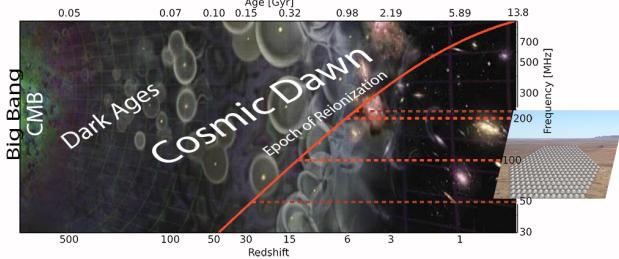


Figure 2. Cartoon showing the development of the Universe and HERA’s observing window into it (dashed lines). The blue solid line shows the HI transition frequency as a function of redshift/cosmic evolution age. The black dashed lines indicate the HERA EOR “core” band and blue dashed lines the expanded band, which are detailed in Table 1.

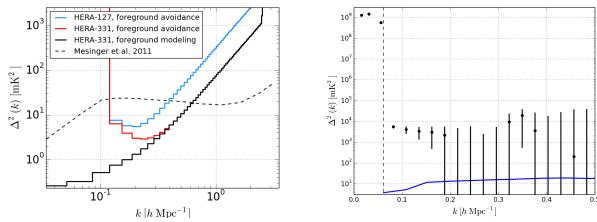


Figure 3. Top: HERA’s power-spectrum sensitivity (solid) relative to a fiducial ionization model (dotted line; $x_{\text{HI}} = 0.37$, $z = 9.0$). Sensitivities reflect staged array size and improving analysis software that expands the range of modes free of foreground systematics. Bottom: The current best upper limit on the 21 cm reionization power spectrum, obtained with a 32-element PAPER deployment (Parsons et al., 2013). These upper limits constrain the brightness temperature of the IGM at $z \sim 8$, showing a departure from adiabatic cooling presumed to be indicative of X-ray heating.

870

875

M., Kuniyoshi, M., Kuper, G., van Leeuwen, J., Maat, P., Mann, G., McKean, J. P., Mevius, M., Mol, J. D., Munk, H., Nijboer, R., Noordam, J. E., Norden, M. J., Orru, E., Paas, H., Pandey-Pommier, M., Pizzo, R., Polatidis, A. G., Reich, W., Röttgering, H. J. A., Sluman, J., Smirnov, O., Stappers, B., Steinmetz, M., Tagger, M., Tang, Y., Tasse, C., ter Veen, S., Vermeulen, R., van Weeren, R. J., Wise, M., Wucknitz, O., and Zarka, P.: Initial deep LOFAR observations of epoch of reionization windows. I. The north celestial pole, *Astronomy and Astrophysics*, 550, A136, doi:10.1051/0004-6361/201220874, 2013.

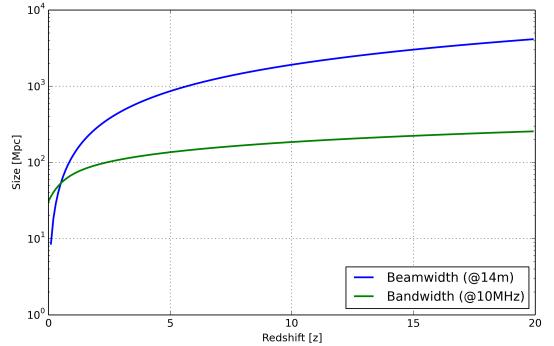


Figure 4. Scaling

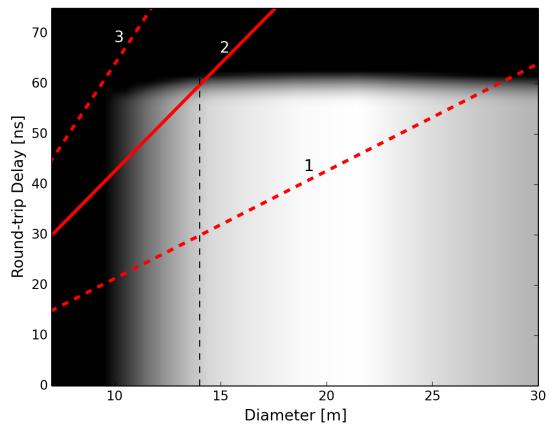
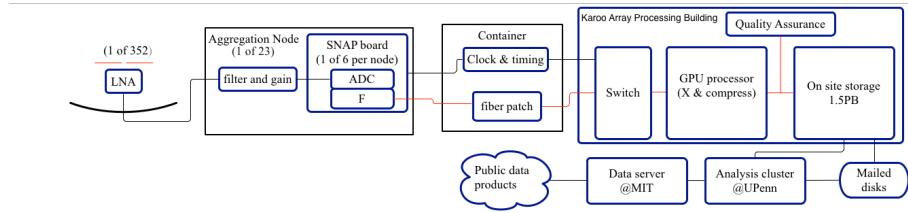
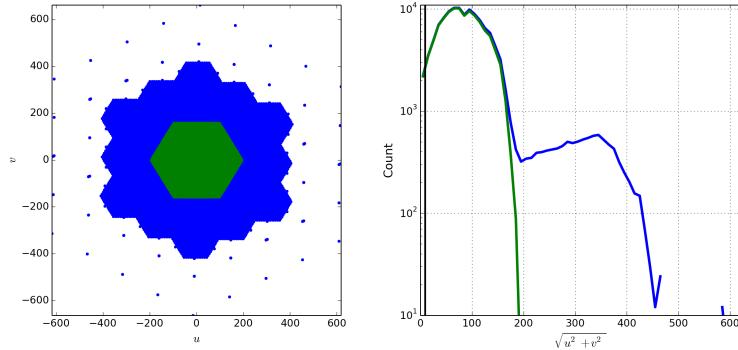


Figure 5. Costing



Figure 6. Existing components re-used in the HERA design include: the PAPER dipole antenna (left), receivers in a node module (center), and the 128-element correlator deployed in the Karoo (right).

**Figure 7.** System**Figure 8.** Configuration**Table 1.** HERA band in frequency, redshift, cosmic age and comoving distance.

Frequency [MHz] [MHz]	Redshift	Age [Myr]	Look-back [Gly]
50	27.4	164	13.6
100	13.2	372	13.4
200	6.1	960	12.8
225	5.3	1,135	12.7

Instrument	Collecting Area (m^2)	Foreground avoidance	Foreground modeling
PAPER	528	1.93	8.86
MWA	896	2.46	6.40
LOFAR NL Core	35,762	2.76	17.37
HERA-127	19,500	10.88	35.65
HERA-331	50,900	25.44	87.20
SKA1 Low Core	833,190	97.92	284.85

Table 2. Power spectrum signal-to-noise ("number of sigmas") at $z = 9.5$ for various instruments, adapted from Pober et al. (2014). By leveraging a filled, redundant configuration of dishes with high collecting area, HERA-331 allows high-significance power spectrum measurements using current foreground avoidance techniques, with further enhancements possible with likely advances in foreground modeling.

Parameter	Design	Performance
Element diameter / FoV	14 m	9°
Min baseline length / largest scale	14.6 m	7.8°
Max core baseline length / synthesized beam	306.6 m	24''
Max outrigger baseline length	1066.5 m	9''
Frequency / redshift range	50 - 250 MHz digitized 70 - 230 MHz useable 100 MHz correlated	19.2 - 5.2
Spectral channel width	97.7 kHz	
System temperature / sensitivity	$100 + 120(\nu/150 \text{ MHz})^{-2.55} \text{ K}$	$50 \mu\text{Jy beam}^{-1} \sqrt{\text{hour}}$

Table 3. HERA-331 basic parameters. Design parameters are connected to the derived instrument performance at 150 MHz.