The Tianlai Dish Array: design, operation and performance of a prototype transit radio interferometer

Fengquan Wu1, Jixia Li1,2, Shifan Zuo1,2,3, Xuelei Chen1,2,4, Santanu Das5,6, John P. Marriner5, Trevor M. Oxholm6, Anh Phan6, Albert Stebbins5, Peter T. Timbie6, Reza Ansari7, Jean-Eric Campagne7, Zhiping Chen8, Yanping Cong1,2, Qizhi Huang1,2, Yi-Chao Li9, Tao Liu8, Yingfeng Liu1,2, Chenhui Niu1, Calvin Osinga6, Olivier Perdereau7, Jeffrey B. Peterson10, Huli Shi1, Gage Siebert6, Shijie Sun1,2, Haijun Tian11, Gregory S. Tucker12, Qunxiong Wang11, Rongli Wang8, Yougang Wang1, Yanlin Wu6, Yadong Xu1, Kaifeng Yu1,2, Zijie Yu1,2, Jiao Zhang13, Juyong Zhang8, Jialu Zhu8

1 National Astronomical Observatory, Chinese Academy of Science, 20A Datun Road, Beijing 100101, P. R. China
2 University of Chinese Academy of Sciences Beijing 100049, P. R. China
3 Department of Astronomy and Tsinghua Center for Astrophysics, Tsinghua University, Beijing 100084, P. R. China
4 Center of High Energy Physics, Peking University, Beijing 100871, P. R. China
5 Fermi National Accelerator Laboratory, P. O. Box 500, Batavia IL 60510-5011, USA
6 Department of Physics, University of Wisconsin Madison, 1500 University Ave, Madison WI 53706, USA
7 LAL, University of Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France
8 Hangzhou Dianzi University, 115 Wenyi Rd., Hangzhou 310018, P. R. China
9 University of the Western Cape, Robert Sobukwe Road, Belville 7535, Republic of South Africa
10 Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
11 China Three Gorges University, Yichang 443002, P. R. China
12 Department of Physics, Brown University, 182 Hope St., Providence, RI 02912, USA
13 College of Physics and Electronic Engineering, Shanxi University, Taiyuan, Shanxi 030006, P. R. China

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
The Tianlai Dish Array Pathfinder is a radio interferometer designed to test techniques for 21 cm intensity mapping in the post-reionization Universe. It performs drift scans of the sky at constant declination. We describe the design, calibration, noise level, and stability of this instrument based on 4,159 hours of on-sky observations to date.

Key words: galaxies: evolution – large-scale structure – 21-cm

1 INTRODUCTION
This paper describes the first astronomical observations by the Tianlai Dish Array Pathfinder. The instrument is an array of 16, 6-meter, on-axis dish antennas operated as a radio interferometer and is co-located with the Tianlai Cylinder Array Pathfinder, an interferometric array of 3 cylinder reflectors in Xinjiang, China. These complementary designs were chosen specifically for testing approaches to 21 cm intensity mapping. Both arrays saw first light in 2016.

21 cm intensity mapping is a technique for measuring the large scale structure of the Universe using the redshifted 21 cm line from neutral hydrogen gas (HI) (Liu & Shaw 2019; Morales & Wyithe 2010). It is an example of the general case of line intensity mapping (Kovetz et al. 2019), in which spectral lines from any species, such as CO and CH, are used to make three-dimensional, ‘tomographic’ maps of large cosmic volumes. 21 cm intensity mapping is used to study the formation of the first objects during the Cosmic Dawn and the Epoch of Reionization ($6 \lesssim z \lesssim 50$) and for addressing other cosmological questions with observations in the post-reionization epoch ($z \lesssim 6$), such as constraining inflation models (Xu et al. 2016) and the equation of state of dark energy (Xu et al. 2015). In the latter epoch, the approach provides an attractive alternative to galaxy redshift surveys. It measures the collective emission from many haloes simul-
taneously, both bright and faint, rather than cataloging just the brightest objects. As a result, the required angular resolution is relaxed. By tuning the receivers to a range of frequencies one obtains a tomographic map. The primary analysis tool for cosmological measurements is the three-dimensional power spectrum of the underlying dark matter, and intensity mapping provides a natural means to compute this spectrum over a range of wave numbers, \( k \), in which the perturbations are in the linear regime. Of particular interest in the power spectrum are the baryon acoustic oscillation (BAO) features, which can be used as a cosmic ruler for studying the expansion rate of the Universe as a function of redshift.

So far, the 21 cm signal has been detected with intensity mapping by two instruments: the Green Bank Telescope (GBT) (Masui et al. 2013; Switzer et al. 2013) and the Parkes Observatory (Anderson et al. 2018) by cross-correlating intensity maps with galaxy redshift surveys.

While HI intensity mapping is being used out to a redshift of \( \sim 50 \) to study the EoR and Cosmic Dawn by a variety of instruments, including LOFAR (van Haarlem et al. 2013), MWA (Tingay et al. 2013), HERA (DeBoer et al. 2017), PAPER (Parsons et al. 2010), and LWA (Eastwood et al. 2018), this paper focuses on measurements of the post-reionization epoch. Several dedicated instruments have been constructed, or are under development, to detect the 21 cm signal from this epoch using intensity mapping, ultimately without the need for cross-correlation with other surveys: CHIME (Bandura et al. 2014), Tianlai (Chen 2012; Das et al. 2018), HIRAX (Newburgh et al. 2016), OWFA (Chatterjee & Bharadwaj 2018), and BINGO (Battye et al. 2016).

Other instruments being designed and built to test the technology for making 21 cm intensity mapping observables include BMX and PAON-4 (Zhang et al. 2016b; Ansari et al. 2019). These 21 cm instruments have several features in common: with the exception of BINGO, they are all interferometers, achieving modest angular resolution at modest cost; they have large numbers of receivers in order to provide enough mapping speed to detect the small 21 cm signal; and the arrays are laid out in a compact arrangement in order to provide sensitivity at the relatively large scales (\( 0.5 \sim k \sim 0.05 \)) where the BAO features appear in the power spectrum.

Although the 21 cm intensity mapping approach has been used for over a decade, it faces significant challenges. The most significant is the fact that the 21 cm signal is roughly 4 orders of magnitude dimmer than foreground emission (primarily synchrotron radiation) from Galactic and extragalactic radio sources. Analysis techniques for extracting the 21 cm signal generally rely on the fact that foreground emission is a slowly-varying function of frequency while the 21 cm spectrum has structure arising from the large-scale distribution of matter along the line of sight (Liu & Shaw 2019). However, instrumental effects can introduce structure into the spectrum of otherwise smooth foregrounds. In particular, the spatial angular dependence of the antenna patterns is also frequency dependent and, in a process called ‘mode-mixing,’ couples the angular dependence of the bright foregrounds into frequency dependence that masquerades as cosmic 21 cm structure. In addition, although the 21 cm signal is unpolarized, the bright foregrounds are partially polarized and frequency-dependent instrumental ‘leakage’ of Stokes Q, U, and V into I introduces another type of foreground with a complicated spectrum. Faraday rotation in the interstellar medium creates further spectral structure. Removing mode-mixing effects requires detailed understanding and measurement of the frequency-dependent gain patterns of the antennas and of the gain and phase of the instrument’s electronic response: i.e. calibration. To determine the scale of the calibration challenge, (Shaw et al. 2015) performed detailed simulations of the CHIME interferometer ability to measure the HI signal in the presence of foregrounds. They showed that the unbiased power spectrum of the the HI signal can be recovered if the beamwidth of the antennas is known to 0.1% and the electronic gain is known to 1% within each minute of observation.

Unlike most radio interferometers, including some of the Cosmic Dawn and EoR instruments that can track targeted regions of the sky, all of the post-EoR instruments observe by drift scanning the sky. This observing strategy allows for large sky coverage using simple and inexpensive instrument designs but complicates the calibration strategy. Tracking instruments can calibrate continuously on bright sources in or near the field they are mapping. Drift scanning instruments like Tianlai must wait for bright sources to pass through the field, or attempt to calibrate on dimmer sources. Therefore, much of the discussion in this paper focuses on measuring stability and performing the calibration of the Tianlai Dish Array.

Future intensity mapping interferometers may include thousands of antennas Cosmic Visions 21 cm Collaboration et al. (2018); Slosar et al. (2019) in order to increase sensitivity. The only currently feasible way to correlate this many signals is by using fast Fourier transform algorithms (i.e. an ‘FFT correlator’) that require that the antenna patterns from the dishes be identical. Anticipating the need for such advances, this paper also characterizes the uniformity of the antenna patterns in the Tianlai Dish Array.

This paper is organized as follows: Section 2 describes the Tianlai Dish Array; Section 3 describes the observations; Section 4 compares the measured antenna patterns with simulations and evaluates their uniformity; Section 5 provides an overview of the data analysis process; Section 6 describes the gain and phase calibration process and Section ?? presents the results of this calibration in terms of noise level, and sensitivity vs. integration time; Section 8 describes results of long integrations on the North Celestial Pole (NCP); Section 7 presents maps of bright calibration sources, and conclusions and plans for the future appear in Section 10.

### 2 INSTRUMENT

The objective of the Tianlai program is to make a 21 cm intensity mapping survey of the northern sky. At present the Tianlai program is in its Pathfinder stage, which aims to test the technology for making 21 cm intensity mapping observations with an interferometer array. The Pathfinder comprises two arrays, one consisting of dish antennas and the other of cylinder reflector antennas, both located at a radio quiet site (44°10'35.1"N, 91°44'30.24"E) in Hongliuxia, Balikun County, Xinjiang Autonomous Region, in northwest China. In order to avoid radio-frequency interference (RFI) gen-
erated by the correlator, the station house, which includes an analog analog electronics room, a digital correlator room, and living quarters, is located 5.8 km (11.2 km by road) away from the telescope site. A power line and optical fiber cables about 8 km long connect the correlator building with the antenna array. Construction of the Pathfinder arrays was completed in 2016 and it is now taking data on a regular basis. This paper focuses on the dish array. Further details about the cylinder array appear in Zhang et al. (2016a); Cianciara et al. (2017); Das et al. (2018); Zuo et al. (2019).

For each array, the feed antennas, amplifiers, and reflectors are designed to operate from 400 MHz to 1430 MHz, corresponding to $2.55 \geq z \geq -0.01$. The instrument can be tuned to operate in an RF bandwidth of 100 MHz, centered at any frequency in this range by adjusting the local oscillator frequency in the receivers and replacing the band pass filters. Currently, the Pathfinder operates at 700–800 MHz, corresponding to HI at $1.03 \geq z \geq 0.78$. Future observations are planned in the 1330–1430 MHz band ($0.07 \geq z \geq -0.01$) to facilitate cross correlation with low-z galaxy redshift surveys and other low-z HI surveys. The system noise temperatures for both the dish and cylinder arrays are $\sim 80 – 85$ K.

The Tianlai Dish Array consists of 16 on-axis dishes. Each has an aperture of 6 m. The design parameters of the dishes are shown in Table 1. The dishes are equipped with dual, linear-polarized receivers and are mounted on Alt-Azimuth mounts, and motors are used to control them electronically. The motors can steer the dishes to any direction in the sky above the horizon. The drivers are not specially designed for tracking celestial targets with high precision. Instead, in the normal observation mode, we point the dishes at a fixed direction and perform drift scan observations. The Alt-Azimuth drive provides flexibility during commissioning for testing and calibration.

The dishes are currently arranged in a circular cluster (Fig. 7). The array is roughly close-packed, with center-to-center spacings between neighboring dishes of about 8.8 m. The spacing is chosen to allow the dishes to point down to elevation angles as low as $35^\circ$ without ‘shadowing’ each other. One antenna is positioned at the center and the remaining 15 antennas are arranged in two concentric circles around it. It is well known that the baselines of circular array configurations are quite independent and have wide coverage of the $(u, v)$ plane. A comparison of the different configurations considered for the Tianlai Dish Array and the performance of the adopted configuration can be found in Zhang et al. (2016b). The Tianlai dishes are lightweight and the mounts are detachable, so, in future, the dishes can be moved to new configurations if required. This paper describes observations with the current configuration.

A schematic of the RF analog system can be found in Fig. 2. The whole system except filters has been designed to operate over a wide range of frequencies (400–1500 MHz). The low noise amplifiers (LNAs) are designed to have low noise temperature (about 35 K at room temperature) and are mounted to the back of the feed antennas. The amplified RF signals pass through 15-meter long coaxial cables to optical transmitters mounted under the dish antennas. The RF signal amplitudes modulate optical transmitters so that the RF signals are converted to optical signals, which are then transmitted to the station house via 8 km optical fiber. At the RF analog system room in the station house, the optical signal is converted back to the RF electric signal. Replaceable bandpass filters with 100 MHz bandwidth are mounted between the optical receivers and analog downconverters. An analog mixer then downconverts the RF signal to the 135–235 MHz intermediate frequency (IF) band. Finally, the IF signal is sent to the digital system through bulkhead connectors between the analog and digital rooms. The dishes currently observe in the frequency band 700–800 MHz (1.03 $>$ z $>$ 0.78) in 512 frequency channels ($\Delta f = 244.14$ kHz, $\Delta z = 0.0002$).

The digital backend system of the dish array is a 32-input correlator that consists of three FPGA boards: two processing boards for signal sampling and processing, and one for control. The Analog to Digital Converters (ADC) in the processing boards convert the RF signal to time series data at a sampling rate of 250 MSPS and sampling length of 14 bits. Then the FPGA chips in the processing boards perform the Fast Fourier Transform (FFT) of the time series data. The two FPGA boards exchange half of the signal channels with each other through rapidIO cables, so all cross-correlations are computed in the FPGA boards while the computation loads on the boards are balanced. Finally, the visibility from the dish array (32 auto-correlations and 496 cross-correlations) are sent to a storage server by two ethernet cables and dumped to hard drives in HDF5 format.

Calibration of the electronic gain of the receivers is crucial for any interferometer, and it is especially important for Tianlai to compensate for phase variation in the 8 km long optical cables. The absolute calibration of the system can be performed using bright astronomical standard calibration sources. However, for small aperture arrays like the Tianlai Dish Array, there are not enough bright sources on the sky to meet the requirement of point source calibration, so we have designed a dedicated calibration noise source (noise source) to provide relative calibration. A broadband RF noise generator is placed in a thermostatically controlled environment and is supplied with regulated DC power to ensure the stability of the RF amplitude. The on-off timing of the noise source is controlled by a clock signal carried by optical fiber from the correlator 8 km away in the station house.

### Table 1. Main design parameters of a Tianlai dish antenna.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value or Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflector diameter</td>
<td>6 m</td>
</tr>
<tr>
<td>Antenna mount</td>
<td>Alt-Az pedestal</td>
</tr>
<tr>
<td>$f/D$</td>
<td>0.37</td>
</tr>
<tr>
<td>Feed illumination angle</td>
<td>68$^\circ$</td>
</tr>
<tr>
<td>Surface roughness (design)</td>
<td>$\lambda/50$ at 21 cm</td>
</tr>
<tr>
<td>Altitude angle</td>
<td>8$^\circ$ to 88.5$^\circ$</td>
</tr>
<tr>
<td>Azimuth angle</td>
<td>$\pm 360^\circ$</td>
</tr>
<tr>
<td>Rotation speed of Az axis</td>
<td>0.002 $\sim 1^\circ$/s</td>
</tr>
<tr>
<td>Rotation speed of Alt axis</td>
<td>0.002 $\sim 0.5^\circ$/s</td>
</tr>
<tr>
<td>Acceleration</td>
<td>$1^\circ$/s$^2$</td>
</tr>
<tr>
<td>Gain (design)</td>
<td>$29.4 + 20 \log(f/700 \text{ MHz})$ dBi</td>
</tr>
<tr>
<td>Total mass</td>
<td>800 kg</td>
</tr>
</tbody>
</table>

3 OBSERVATIONS

So far we have collected about 6200 hours of observational data from the Tianlai Dish Array, including more than
Figure 1. Left: Top view of the Tianlai Dish Array and Cylinder Array taken with a DJI M600 Pro drone at a height of 280 m above the ground. The position of the calibration noise source is indicated by the white lines. The relative distance vector from the feed in dish 16 at the center of the array (when pointed toward the zenith) to the noise source is [-184.656, 13.915, 12.588] meters, with x,y,z to the west, north, and zenith. Right: A schematic diagram of the Tianlai Dish Array. The dishes are arranged in two concentric circles of radius 8.8 m and 17.6 m around a central dish. The dishes have dual-linear polarization feed antennas with one axis oriented parallel to the altitude axis (horizontal, H, parallel to the ground) and the other orthogonal to that axis (vertical, V). For example, shown in red is one of the baselines that is studied later in this paper, the H polarization of dish 4 correlated with the H polarization of dish 9: 4H-9H. Other baselines used later in the paper use the same naming convention.

Figure 2. Schematic of the RF analog system.

Table 2. Observation log for the Tianlai Dish Array from 2016 to late 2019.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Date</th>
<th>Calibration Sources</th>
<th>Targets</th>
<th>Length (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data 201605-06</td>
<td>May 2016</td>
<td>None</td>
<td>Cygnus A</td>
<td>72</td>
</tr>
<tr>
<td>CygnusANP 20170812</td>
<td>Aug 2017</td>
<td>Cygnus A</td>
<td>North Pole</td>
<td>67</td>
</tr>
<tr>
<td>CasAs 20171017</td>
<td>Oct 2017</td>
<td>None</td>
<td>North Pole</td>
<td>147</td>
</tr>
<tr>
<td>CasAs 20171026</td>
<td>Oct 2017</td>
<td>None</td>
<td>Cassiopeia A</td>
<td>290</td>
</tr>
<tr>
<td>3srcNP 20180101</td>
<td>Jan 2018</td>
<td>3C48, Cassiopeia A, M1</td>
<td>North Pole</td>
<td>241</td>
</tr>
<tr>
<td>3srcNP 20180112</td>
<td>Jan 2018</td>
<td>3C48, M1</td>
<td>North Pole</td>
<td>97</td>
</tr>
<tr>
<td>IC443NP 20180323</td>
<td>Mar 2018</td>
<td>IC443</td>
<td>North Pole</td>
<td>181</td>
</tr>
<tr>
<td>M87NP 20180407</td>
<td>Apr 2018</td>
<td>M87</td>
<td>North Pole</td>
<td>90</td>
</tr>
<tr>
<td>2srcNP 20180416</td>
<td>Apr 2018</td>
<td>IC443, M87</td>
<td>North Pole</td>
<td>142</td>
</tr>
<tr>
<td>3srcNP 20181212</td>
<td>Dec 2018</td>
<td>Cassiopeia A, 3C48, M1</td>
<td>North Pole</td>
<td>757</td>
</tr>
<tr>
<td>1DaySun 20190113</td>
<td>Jan 2019</td>
<td>None</td>
<td>Sun</td>
<td>48</td>
</tr>
<tr>
<td>3srcNP 20190128</td>
<td>Jan 2019</td>
<td>Cassiopeia A, 3C48, M1</td>
<td>North Pole</td>
<td>741</td>
</tr>
<tr>
<td>3srcNP 20190228</td>
<td>Feb 2019</td>
<td>3C123, M1, IC443</td>
<td>North Pole</td>
<td>764</td>
</tr>
<tr>
<td>3srcNP 20190402</td>
<td>Apr 2019</td>
<td>M1, IC443, 3C273</td>
<td>North Pole</td>
<td>522</td>
</tr>
<tr>
<td>3srcNP 20190611</td>
<td>Jun 2019</td>
<td>M87, Hercules A, Cygnus A</td>
<td>North Pole</td>
<td>737</td>
</tr>
<tr>
<td>3srcNP 20190830</td>
<td>Aug 2019</td>
<td>3C400, Cygnus A, Cassiopeia A</td>
<td>North Pole</td>
<td>924</td>
</tr>
<tr>
<td>3srcNP 20191022</td>
<td>Oct 2019</td>
<td>3C400, Cygnus A, Cassiopeia A</td>
<td>North Pole</td>
<td>302</td>
</tr>
</tbody>
</table>
5700 hours of NCP data. In Fig. 3 we show the accumulated observing time over the years. Details of individual runs are listed in Table 2. Drift scans are performed at constant declination over several days at a time. These can be divide into two types: (1) 24 hr observations at declinations away from the NCP, usually at the declination of bright sources: Cyg A (+40°44′), CasA (+58°48.9′), Tau A/M-1 (+22°00′) and also some high declination regions (≈ 80°). (2) 24 hour observations at the NCP. Preceding each NCP observation, the antennas are pointed towards one or multiple strong radio sources for calibration. The calibration sources for different NCP observations are listed in Table 2.

The integration time for the dish array is set at 1 s. The data are stored for all 528 correlation pairs (auto-correlation + cross-correlation) in 512 different frequency channels. The data rate is about ≈ 175 GB/day. The weather data, which includes the temperature of the analog electronics room, site temperature, dew point, humidity, precipitation level, wind direction, wind speed, barometric pressure, etc., are stored separately during each run. These data can later be used for checking the correlation of different weather variables with the variation of electronic gain of the system.

The noise source is turned on and off periodically. During 2017 the noise source switched on for 20 s every 4 min, so the fraction of noise-on time is ≈ 8.33%. In 2018, the noise-on time was reduced to 4 s per 4 min, which is 1.67% of the observing time.

In Fig. 4 we show the noise source and RFI mask derived from 1 hour of nighttime data. The periodic vertical stripes show the mask when the noise source is turned on, while the yellow dots show the RFI. We use two different RFI cleaning methods (check Sec. 5). Because the array is located in a radio quiet zone, we only lose about 0.6% of data to RFI contamination.

4 BEAM PATTERNS

Separating the faint 21 cm signal from strong foregrounds requires exquisite knowledge of the frequency-dependent beam patterns of the antennas. Small deviations of the mechanical configuration or the environment (e.g., temperature, wind) can cause changes to the beam pattern and introduce systematic errors in the measurement. A 1% uncertainty in the beam measurement may introduce a 4% systematic uncertainty in the final power spectrum. Therefore an accurate knowledge of the beam pattern is necessary.

We have taken preliminary steps toward characterizing the beam patterns using scans with a radio source flown over the array on an unmanned aerial vehicle (UAV). We also match the beam measurements with electromagnetic models of the beams and with beam measurements from the transits of bright astronomical sources.

Figs. 5 and 6 summarizes the antenna beam measurements and simulations. A detailed description of the drone measurements and simulations appears in Zhang et al. (2019). An UAV outfitted with a broadband noise source is flown in the far field along two paths over one of the antennas: a North-South path and an East-West path. Multiple flights are conducted at different RF power levels to map the beam into the far sidelobes. Fig. 5 (Left) shows the beam pattern measured by the UAV at 730 MHz and compares it with the profile of the main beam measured at the same frequency using the auto-correlation signal from one polarization during a transit of CasA. The CasA signal is not bright enough to measure the sidelobes of the antenna.

We also performed electromagnetic simulations of the feed antenna and the dish reflector using commercially-available software. We overplot the simulated beam pattern with the drone measurement in Fig. 5 (Right). The patterns match well in the main lobe. However, the UAV measurements show a slightly stronger signal in the sidelobes. We are working on understanding the difference between the model and the measurement. Simulations extended into the far sidelobes and backlobes are shown in Fig. 6.

As mentioned in Section 1, knowledge of the beam patterns of each antenna as a function of frequency, and the stability of the pattern with time, are also important factors. Fig. 7 shows the FWHM of one cut through the beam pattern from a pair of antennas as a function of frequency. The pattern is measured repeatedly in the E-W direction by observations of the transit of CasA on 12 successive days. For each transit, the magnitude of one visibility vs. time is fit to a Gaussian. The measured pattern is effectively the

Figure 5. Left: E-W dish H-plane antenna pattern measured both by an UAV and by a transit of Cas A at 730 MHz. Right: Dish beam pattern measured with the UAV and compared with electromagnetic simulation. Are these measurements for the E plane or for the H plane?

Figure 6. Simulated beam pattern as a function of $\theta$ for 3 different frequencies, 700, 750 and 800 MHz. Each plot shows the absolute co-polar directivity in dBi in the E-plane ($\phi = 0$). Angle $\theta$ is calculated from the center of the beam. The simulation shows that the antenna sidelobes vary significantly as a function of frequency.

Figure 7. Plot of the mean FWHM of the main beam vs. frequency using daily transits of Cas A over 12 days. The measurement is for a fairly typical baseline; this one is between feeds 4H and 9I. The black line shows the mean in each frequency bin over this period and the red scatter plot shows the FWHM for each day. The frequency binning is 244 kHz. The green line shows the expected FWHM vs. frequency based on an electromagnetic simulation. The blue line shows the FWHM of a uniformly-illuminated Airy disk with an effective diameter that is 90% of the actual 6 meter diameter.

Figure 8. The black line shows the mean FWHM vs. frequency for 118 H-H baselines during a transit of M1 on 2018-01-02. The plot excludes autocorrelations, baselines with known faulty probes, and other outliers.

geometric mean of the patterns of two dishes, which are nominally coaligned. Day-to-day fluctuations of the FWHM are less than 1%. The frequency dependence of the beam in the H- and V-planes is also calculated by the electromagnetic simulations; the simulated FWHM of the is co-plotted on the same figure for comparison. We also plot the case of a diffraction - limited circular aperture ($1.028\lambda/D_{\text{eff}}$ with $D_{\text{eff}} = 0.9D$). The somewhat complicated frequency-dependence of the beamwidth is predicted qualitatively by electromagnetic simulations. Ultimately we will extend these measurements to all the dishes in the array over the full range of frequencies.

Another important characteristic of the antenna pattern is the uniformity of the different antennas. By using the transit observations described here, the effective beamwidth of all pairs of antennas (baselines) can be measured and compared. Nominally, they should all be identical, but misalignment of antennas and manufacturing errors will cause the patterns to differ. Fig. 8 shows the mean value of the FWHM of 118 baselines in the array and the 1-sigma deviations from the mean as a function of frequency. The 1-sigma deviations are $\sim 4\%$.

We have also studied the pointing accuracy of the dishes in the E-W direction. Using data from the Cas A transit
Figure 9. Variations of the modulus of the cross-correlation visibilities between the Tianlai dishes with time (in s) during transits of Cas A. Each of the 120 cross-correlations corresponds to one color. The black curve shows the expected response from a simplified simulation (from a 5.6 m diameter dish with a Gaussian beam). All (un-calibrated) cross-correlations have been renormalized to unity at their respective maxima. The observed time spread between the cross-correlations reflects the dishes’ relative E-W pointing dispersion. The green dip at around 6000 s is an artifact from the noise source interpolation. Each curve is the average of 40 frequency channels (742.6172 MHz to 752.1387 MHz with each frequency bin being 244 kHz) near the center of the RF band.

Figure 10. Extracted transit time differences w.r.t. to the expectation from a simplified simulation, for the 16 Tianlai dishes from the Cas A transit of 2017/10/30. The ~150 s FWHM of these time shifts from cross-correlation corresponds to an E-W pointing FWHM of ~0.47 degrees on the sky. This indicates an E-W pointing FWHM of ~0.67 degrees for single dish.

5 OVERVIEW OF (OFFLINE) ANALYSIS PROCESS

For analyzing the Tianlai data we developed a Python pipeline named tlpipe\(^2\). It is a collection of several stand-alone packages that can be used for reading the raw visibility data, masking, RFI flagging, calibration, data binning, map-making, etc. Multiple visualization and other utility packages have also been developed. The pipeline is written in a modular format and users can develop and add their own algorithms to the pipeline with little knowledge of how the rest of the pipeline works. Fig. 11 shows the basic workflow of the tlpipe for making maps from the raw visibility data.

5.1 tlpipe data processing workflow

We use Python-2.7 as the main programming language. This allows us to use its vast collection of scientific computing libraries. However, some of the performance-critical parts have been compiled to C by using Cython. Parallel processing has been implemented with the Message Passing Interface (MPI) framework.

The basic workflow of tlpipe can be broken into 3 distinct components: task manager, tasks, and data container.

Task manager: The task manager controls the overall flow of the pipeline and applies different Tasks to the data.

Tasks: Tasks are independent codes that operate on the data in the data container. These tasks include noise source removal, RFI flagging, map making, calibrations, etc. We discuss some of the tasks in the next section.

Data container: The data container holds the data operated on by the pipeline. It includes an array of visibilities,
The overall data processing workflow of \texttt{tlpipe} appears in Fig. 12. The list of tasks is entered into a \texttt{*pipe} file. The task manager takes the file as an input and applies the tasks to the data container in sequence. The tasks, in general, modify the visibility and the mask array in the data container.

\subsection*{5.2 Built-in data processing tasks}
In the \texttt{tlpipe} package, we have implemented more than 30 tasks. Users can write their own independent tasks to apply to the data container. Some of the built-in tasks in the \texttt{tlpipe} are as follows:

\textbf{Masking of the noise source} The Tianlai data contain periodic calibration signals from the noise source which must be removed during analysis. A dedicated subroutine detects the presence of the noise source signal by measuring the difference between the amplitude of two consecutive data points in an autocorrelation channel and comparing it with the overall variance. The routine calculates the turn-on and turn-off times of the noise source and sets elements of a Boolean mask array to True when the noise source is on. In Fig. 4 the masked noise source can be seen as periodic vertical straight lines.

\textbf{RFI cleaning} After masking the noise source we need to clean RFI from the data. Multiple RFI flagging algorithms are available in \texttt{tlpipe}, but a combination of the sum threshold method (Offringa et al. 2010) and the scale-invariant rank (SIR) operator method (Offringa et al. 2012) work best for the Tianlai data. Visibility data points flagged as RFI are recorded in the mask array. The RFI calculated from 1 hour of night time data using the sum threshold + SIR operator method is shown in Fig. 4.

\textbf{Noise source calibration} Two methods for calibration are included in \texttt{tlpipe}. For absolute calibration of the amplitude and phase of the gain we use strong astronomical point sources. However, as only a few bright sources are available, we perform relative calibrations using a regularly broadcast calibration noise source signal. The noise source calibration is primarily used to remove the phase variations over time but we are studying its use for amplitude calibration as well.

We developed two different algorithms for the calibration using the noise source. The first task, \texttt{nscal}, uses the noise source to calibrate each visibility. For each baseline it defines the visibility during the on and off cycles of the noise source to be $V_{ij}^\text{ON} = V_{ij} - V_{ij}^\text{NS}$ and $V_{ij}^\text{OFF} = V_{ij}$, where $V_{ij}$ is the observed visibility from the sky and $V_{ij}^\text{NS}$ is the visibility from the noise source corresponding to the baseline $i,j$. The phase introduced by the noise source is then $\phi_{ij} = \text{Arg}(V_{ij}^\text{ON} - V_{ij}^\text{OFF})$. Because the noise source amplitude and phase are constant for a particular baseline, the corrected visibility from the sky, after the noise source calibration, is $V_{ij}^\text{NS-Cal} = \exp(-i\phi_{ij})|V_{ij}|/|V_{ij}^\text{NS}|$. Further details appear in Zuo et al. (2019).

The second task, \texttt{nscal2}, uses the noise source to perform a global fit to the observed visibilities to determine an independent (complex) gain for each feed. Because only phase differences between feeds matter, the phase of feed 1 is fixed to be zero without any loss of generality. The gain amplitudes reported are relative to the (uncalibrated) noise source amplitude. This method is used in the noise source calibration results shown in Sec. 6. Because the gain changes with time, a spline is fit to the measured gains and used to interpolate between noise source calibration events.

\textbf{Point source calibration} After the relative phase calibration using the noise source, transits of strong astronomical radio sources are used to make an absolute gain calibration that gives the actual amplitudes and phases of the complex gains at the time of the source transit. A model for the beam pattern is assumed. Calibration is provided in units of K or Jy. Again, there are two calibration routines, \texttt{PSCal} and \texttt{PSCal2} both use the same robust principal value decomposition to determine the gain of each feed, but have some differences in the handling of outliers and diagnostic information. The computed gain is applied to the entire data set until the next point source transit.

\textbf{Map-making} The built-in mapmaking code uses the $m$-mode analysis from (Shaw et al. 2015, 2014; Zhang et al. 2016b). The details of the algorithm are also discussed in TBD. There are also independently-written mapmaking codes which we use for data analysis. Details are discussed in Sec. 7.
Utility tasks and plotting tasks Apart from these standard tasks, `tipipe` includes multiple utility packages and plotting tasks. The utility packages include codes such as those for removing contamination by the Sun from the daytime data. This technique uses an eigenvalue approach for removing the largest eigenvalue from the daytime data and can successfully remove 99% of the solar contamination. It will be described in a future publication. Other utility packages include removing bad channels, daytime masking, etc. The plotting packages include codes for plotting waterfall plots, plotting time or frequency slices of the data, etc.

6 CALIBRATION

As described in Section 5, two methods are used to calibrate the amplitude and phase of the electronic gains of the receivers. One method uses a signal that is transmitted periodically toward the array by a nearby noise source outfitted with an omnidirectional antenna. The second method uses transits of bright point sources through the main beam. The point source calibration necessarily occurs much less frequently than does the noise source calibration. In fact, for observations of the NCP region, where there are no bright point sources, point source calibration requires repointing the dishes away from the pole. (In the future, calibration from bright point sources that appear in the antenna side-lobes (see Fig. 23) may prove useful, but that topic is beyond the scope of this paper.) In this section we apply these two types of calibration to the data and evaluate the stability of the array’s gain (amplitude and phase) over time.

6.1 Gain stability measured with noise source

The response of the dishes to the noise source is not a smooth function of frequency and the shapes of the response vary widely. We believe much of this frequency structure arises because, for most observing directions, the noise source is coupled to the antennas through their far side-lobes. Electromagnetic simulations of these far side-lobes demonstrate significant variation with frequency. Fig. 13 (Left) is a “typical” H-H cross-correlation and Fig. 13 (Right) is a typical V-H cross-correlation. The plots show the cross-correlation of feeds in different dishes, but when the feeds are in the same dish the plots are similar. The phase plots have a phase and delay that is fit and subtracted from the visibility phase so that the residual phase is close to zero. Note that the amplitude of the response of the H-H and V-H polarizations are similar.

Fig. 14 shows the gain as a function of time for feed 5V with respect to time for 3 consecutive days in October 2017. The gain is measured using the nscalg task described in the previous section. The site temperature recorded for the same 3 days is shown in the bottom plot of Fig. 14. The changes in amplitude and phase are correlated with each other and with temperature, particularly on short time scales, but the relationship is not 1-to-1. However, it is reasonable to expect that the temperature of the electrical components does not follow the site air temperature exactly. Indeed, there is a clear hysteresis behavior during the daylight hours. The large phase shift observed is probably caused by temperature changes in the long fiber optic link between the receivers on the dishes and the correlator.

After applying nscalg, we can use the fitted gains of the individual feeds to compute the calibrated visibilities that should have been observed by baselines involving those feeds. Fig. 15 shows a test of the calibration process. The observed 1H-5H visibility (red data points, measured when the noise source is on) is compared to the visibility that is expected using the fitted gain curves for feeds 1H and 5H. There is a small but significant offset of about 0.04 radians or about 2° in the phase, but the amplitude is well described by the curve. This plot is typical of a significant offset; other visibilities have similar offsets with the opposite sign and other visibilities show smaller or no offsets. The errors are estimated from the variance of the noise signal over the 20, 1 s bins that are measured while the noise signal is applied. The amplitude data are much smoother than would be expected from the error estimate. The phase errors are considerably smaller because the real and imaginary parts of the noise signal are correlated, resulting in larger radial than azimuthal fluctuations in the complex plane. The fits for each frequency are independent of each other, so plotting adjacent frequencies is an indication that changes in time are not an artifact of the fitting process.

Fig. 14 shows the gain amplitude and phase, respectively, plotted against the site temperature. The changes in amplitude and phase are correlated, particularly on short time scales, but the relationship is not 1-to-1. However, it is reasonable to expect that the temperature of the electrical components does not follow the site air temperature exactly; indeed, there is a clear hysteresis behavior during the daylight hours. The large phase shift observed is probably caused by the long fiber optic link between the receivers on the dishes and the correlator. A phase shift of 1 radian would result from a temperature shift of 10 degrees C if the fiber lengths were different by 100 m (about 1%) and the expansion coefficient were $2 \times 10^{-5}$. Alternatively, a 1% variation in expansion coefficients would explain a 1 radian phase shift as well.

6.2 Gain stability measured with point sources

The stability of the instrument was studied by analyzing its response to Cassiopeia A (Cas A) over 12 days. Cas A dominates the radio sky in the northern hemisphere. The Tiantai array was pointed at a fixed declination of 58.8 degrees, the declination of Cas A, and operated in driftscan mode. The data are listed as CasAs 20171026 in Table 2. We analyzed variations in the magnitude and phase of a typical visibility during repeated transits of Cas A across the meridian. The time-dependent response pattern follows the Gaussian profile of the main beam of the antennas shown in Fig. 5.

The amplitude and phase of the uncalibrated peak response for all frequency channels is shown in Fig. 16. Here, the response for all 12 nights are plotted, showing that the gain is quite stable over time. There is significantly less structure in the spectrum than in gain measurements made with the noise source (Fig. 13), presumably because the signal from Cas A enters through the main beam of the antennas, which has significantly less frequency dependence than the far side-lobes, through which the noise source enters.

Deviations of the uncalibrated gain from the mean val-
Figure 13. Left: Gain versus frequency for the cross-correlation of feeds 1H and 3H, measured using the noise source with `nscal2`. The amplitude is in raw (uncalibrated) visibility units and the phase is in radians. Right: Gain versus frequency for the cross-correlation of feeds 3V and 16H. The amplitude is in raw (uncalibrated) visibility units and the phase is in radians. The two baselines used in this plot are fairly typical.

Figure 14. Gain versus time for feed 5V at 747.5 MHz for three days in October, 2017. Each color represents a different day. The gain amplitude (top) scale is uncalibrated and the gain phase (middle) scale is in radians. Site temperature for 3 days is shown in the bottom plot.

Figure 15. Visibility versus time for the cross-correlation of feeds 1H and 5H. The amplitude scale is in arbitrary units and the phase change is in radians. The response to the noise source pulses are shown as red points while the spline fit to the gain is shown as the solid blue line. Also shown are the data for both the next higher and lower frequency bins.

Figure 17. The upper left plot shows fractional deviations in the amplitude of the gain compared to the mean for each frequency channel, with 1 MHz resolution. The lower left plot shows a 1-dimension histogram made from the top plot, in which we combine all 512 frequency channels.

The phase of a visibility vs time was fitted to a line during the times surrounding each Cas A transit. For an East-West baseline, we expect the visibility phase slope to be linear over time, due to a changing phase delay between the two antennas. The right plots in Figure 17 show histograms of fractional deviations from the mean phase slope over 12 days. We are still working to understand the source of the bimodal nature of this distribution.

We verify that the phase calibration performed by the noise source with the `nscal` task over 12 days is consistent with the absolute phase determined by repeated transits of Cas A over the same period of time in Fig. 19 Right. Fig.
Left shows the amplitude response during these transits over 11 nights for the same baseline. In these two panels, the curve represents the peak response during the transit for each.

Figure 16. Top: Uncalibrated gain versus frequency during transits of Cas A over 11 nights. Each colored curve represents the peak response during the transit for each night. Bottom: Uncalibrated phase vs. frequency during Cas A transits over 11 nights for the same baseline. In these two panels, neither the point source nor noise source calibration have been applied to the data.

6.3 Absolute Calibration

To estimate sensitivity of the dish array (??) one needs to calibrate in absolute terms. For dish observations the gain is initially calibrated using point source calibration on one or more of the primary calibrators Cygnus A, Cassiopeia A, Taurus A and Virgo A as described above before pointing toward the field of interest. For the small (low resolution) dish array these calibrators may accurately be treated as point sources. Absolute calibration is provided by comparison with other (external) measurements of the same bright sources. Properly tracking gain drifts will also affect the accuracy of the visibilities.

7 MAPS AROUND SOURCES

To study the array’s ability to make images of the sky and to improve our knowledge of the antenna patterns we have made maps around bright point sources using several algorithms. QuickMap: using the 12-day driftscan that started on 2017-10-26, a rough image of Cas A (Fig. 20) was constructed using the QuickMap map-making method (Zhang et al. 2016b) in a single frequency slice, 747.5 MHz, and a time interval of 4 hours, using only one of the 12 transits (2017/10/30). The map shows a band of declination around Cas A, from 53\degree to 63\degree, with 0.1\degree resolution, and RA resolution of 20 \text{s} (or 5 arcminutes).

QuickMap reads in visibility data in transit mode. The input visibility data size is (720, 121, 1), where 720 is the number of complex visibilities in 4 hours (each visibility is averaged over a 20 second interval), 121 is the number of cross-correlations from horizontal baselines plus one horizontal auto-correlation baseline (16H-16H), and 1 is the number of observed declinations (only the declination of Cas A in this case). Once instrumental phases and relative antenna
Figure 17. Histograms of the gain and phase of visibility 4H-9H measured over 12 days using Cas A transits for calibration. No prior calibration procedures (e.g. nscal) have been applied. Left panels: Histograms of gain amplitudes. Left Top: Histogram of gain amplitude variations vs frequency. The color scale shows number of occurrences. Frequency bins are 1 MHz wide. Left Bottom: Same as top figure, but with all frequencies combined. Right panels: Histograms of phase slope variations.

Using a different analysis, Figure 22 shows a dirty image of the Crab Nebula (M1, Taurus A). We use 20 seconds of data during transits of M1, including all cross-correlation baselines (except flagged ones) of the same polarization (i.e. HH and VV polarization), and all frequency channels. The instrumental phase is corrected but not the instrumental amplitude. The data are gridded in the uv plane by a two-dimensional histogram weighted by visibility. The dirty image $I'(l, m)$ is obtained by a double Fourier transform of the gridded visibility in the uv plane. A similar approach is used to create the dirty map of the NCP region shown in Figure 23. The data include 146 hours over 10 nights starting from 2018/1/1.

Finally, we make maps using CASA\(^3\). Figure 24 and 25 show the dirty and clean images of M1 using 1 hour of the data around M1’s transit on 2018/1/1. 100 frequency channels (749.9-774.3MHz) and all baselines of HH polarization are used. We perform the phase, bandpass, and baseline amplitude calibration. Both images are made using CASA’s tclean task with the number of iterations of the clean algorithm set to 0 to obtain the dirty map and set to 100 to obtain the clean image.

8 NCP PERFORMANCE

The first science goal of the dish array will be deep observations of the North Celestial Pole (NCP). For a Northern Hemisphere transit telescope with a limited field-of-view the NCP is the only place on the sky that allows continuous observations; so this strategy gives the highest S/N visibilities.

---

\(^3\) Common Astronomy Software Applications package: [https://casa.nrao.edu/](https://casa.nrao.edu/)
with the least integration time. One therefore expects to detect the 21 cm signal in the least time this way. Surveying the smallest solid angle possible also yields the largest sample variance, which is a negative aspect of this strategy. At the time of this writing we have accumulated 3700 hours of integration with all dishes pointed directly at the NCP although we only present a small fraction of it here.

In this section we illustrate NCP data with the visibility from a single baseline from one run of 234 hours taken between 2018-01-02 and 2018-01-11. During this run only 78 of 120 dish pairs were fully functional and fairly well behaved and we focus on one particular visibility which we find to be illustrative of the typical behaviour of the interferometer. Due to the low gain at the band edges the usable bandwidth is slightly smaller than the correlator bandwidth: we use $\nu \in [700, 800] \text{ MHz}$ or $\lambda = \frac{c}{\nu} \in [0.37, 0.43] \text{ m}$. In this band there are 412 frequency channels of width 244 kHz.

For the purpose of this section precise absolute calibration is not important. Here we will use the initial point source calibration on Cas A described previously and not track gain drifts after initial calibration. Complex gains may drift over many days and nights of observation and in other contexts we plan to correct for this using the calibration noise source and a sky model as described above. Here we wish to illustrate the smallness of the effects of these drifts on sky observations so no further recalibration is applied. Also no specific RFI mitigation is applied, although at times
The Tianlai Dish Array: design, operation and performance of a prototype transit radio interferometer

Figure 19. Gain and phase measured during repeated Cas A transits after calibration by tlpipe. Left: Brightness temperature measured during transits of Cas A, sampled every second over 11 nights for baseline 4H-9H, after the temperature (gain amplitude) has been calibrated in units of Kelvin by pscal using Cas A transits. Right: The phase during the same transits over 11 nights for the same baseline. The phase is calibrated by tlpipe’s nscal calibration task.

Figure 20. Dirty image of Cas A at 747.5 MHz from a drift-scan using 121 horizontal cross-correlation baselines and one autocorrelation from the 2017/10/30 transit. The maps include corrections for phase and relative gain. The color scale is linear, in arbitrary units.

Figure 21. Image of Cas A reconstructed by pseudo-inverting the pointing matrix from the 2017/10/30 transit, in the frequency slice around 747.5 MHz. The method was applied to three successive rectangular areas before, during, and after the actual Cas A transit to cover the same sky area as Fig. 20. The color scale is linear, in arbitrary units.

Figure 22. Dirty image of M1 from all HH and VV cross-correlation baselines.

we use median averaging over successive nights, which is a form of outlier rejection which would remove RFI.

Typical daylight visibilities when pointed at the NCP are $\sim 500\,\text{mK}$ during daytime and $\sim 30\,\text{mK}$ during nighttime. With $T_{\text{sys}} \sim 100\,\text{K}$ the noise temperature expected for cross-correlations with our 1 sec and 244 kHz sampling is $T_{\text{sys}}/\sqrt{2\delta t \delta\nu} \sim 100\,\text{mK}$ (for either the real or imaginary parts) which would mean these “pixels” are noise-dominated during nighttime. Below we often present visibilities averaged into “1min×1MHz” pixels by taking the mean over $\delta t = 1\,\text{min} = 60 \times 1\,\text{sec}$ and $\delta\nu = 0.977\,\text{MHz} = 4 \times 244\,\text{kHz}$ pixels, reducing the noise by $\sqrt{240}$, to typically $\sim 7.5\,\text{mK}$ so that even nighttime pixels have S/N of a few. Even with

1min×1MHz averaging destructive interference between sky sources will cause the S/N to drop below unity occasionally, i.e. for certain frequencies and sidereal times depending on the baseline. During 1 min the Earth rotates by 0.25°. The maximum angular resolution of the dish interferometer is $\sim 0.35°$ so even sources near the Celestial equator would not be greatly smeared during 1 minute and sources near the NCP will have negligible smearing.

We illustrate the dish NCP data with a particular baseline, 2V×10V. There are many baselines which exhibit a variety of behaviors during this 11 day run but 2V×10V is fairly typical. It has a baseline 5.9 m East and 9.9 m North or 11.5 m total. Figure 26 graphically represents this visibility over the entire 234 h run using the color representation specified in Appendix 11. One clearly sees the much larger solar signal during daytime even though the Sun is 107.5° off axis. The day/night transition is fairly sharp, taking only
a few minutes. One can see the daylight hours slowly drift to larger sidereal time over successive days as the Earth revolves around Sun. The “bow” pattern in the daytime phases is what one expects when a bright source passes directly above the direction of the baseline. The irregularities in the Sun-dominated visibilities are due to interference with other bright off-axis sources, the complicated structure of the beam pattern far off-axis, and also the correlated noise described below. Apart from the shift due to Earth’s revolution, the Solar visibilities, including the irregularities, are highly repeatable. During nighttime the visibility signal is smaller.

A less visual and more quantitative comparison of different sidereal days is given in Figure 27, which shows the frequency-averaged visibility modulus. Daytime (roughly $15^\text{h} \lesssim \text{LST} \lesssim 23^\text{h}$) is dominated by the Sun, which drifts to larger LST as expected. No LST drift is apparent in nighttime visibilities, i.e. the features do not move in LST. The most obvious night-to-night variations are in the amplitude of the signal, which is a combination of gain drifts and varying contamination from correlated noise. We reiterate that no correction for time varying gain has been made here.
The Sun contributes from 400 to 2000 mK and is peaked near sunset (15\textdegree) and sunrise (23\textdegree) as one would expect for vertical (V) feeds (for horizontal (H) feeds the solar signal peaks near midday). The Sun’s motion on the the sky is clearly evident as the pattern shifts to the right. Since the Sun’s motion is mostly in the R.A. direction, and the beam is centered on the NCP, the Sun will approximately trace the same path through the beam on sequential days but at different sidereal times. Other day-to-day changes in the Sun signal is partly due to gain variation but also due to increasing declination of the Sun. The nighttime signal more accurately repeats every sidereal day although there is evidently up to 20% percent variations especially on days 10 & 11. The nighttime signal is a combination of sky signal, which should depend only on sidereal time, and correlated noise, which is roughly constant. The night-to-night variation is a combination of gain drift and variations in the correlated noise.

### 8.1 Nighttime Visibilities

Figure 28 shows the same visibilities as in Figure 26 except only the nighttime data are shown and the color saturation level has been adjusted to better represent the smaller nighttime signal. The dominant features are horizontal stripes with some temporal variation. These data are consistent with containing a significant amount of “correlated noise” in addition to sky signal. Correlated noise is identified as being nearly constant in time and not exhibiting the temporal fringe patterns one expects from Earth rotation; it is a type of RFI which may be natural, man-made, or even self-generated. Generally nighttime visibilities contain roughly equal contribution of sky signal and correlated noise although some baselines exhibit very little correlated noise. Unfortunately we have no other handle on the amount of correlated noise besides the visibilities themselves. The cause of correlated noise has not been determined but ground emission is probably significant and radio emission from one feed being picked up by another (“cross-talk”) is also likely. Line-of-sight transmission between feeds is a possible cause as this is not blocked by the dishes. The beam gain in these directions is very small but the Tianlai Dish Array is fairly compact, with dish centers placed as close as ~ 1.5 diameters from each other. It is also possible that there is some leakage of signal between different channels during transmission to the correlator. Cross-talk would depend on the geometry and decrease with distance between the feeds. There is some but not clear evidence for this behaviour. The frequency spectrum including the overall amplitude of correlated noise varies between the baselines. Other analysis suggests it is not caused by a single or small number of very noisy transmitters.

Figure 28 also exhibits a few bright pixels such as near 11\textsuperscript{h} on the 1st night and near 9\textsuperscript{h} on the 10th night. These do not repeat with sidereal time, span a small frequency range and may be external RFI. RFI flagging will identify these and other less obvious RFI contamination. In this section we do not make use of RFI flagging, however, when we median average different nights, obvious outliers will be suppressed.

### 8.2 Subtracting Correlated Noise

Correlated noise that is constant in time can be removed by subtracting the time-averaged mean visibility. This would also completely remove any unpolarized sources which are exactly at the NCP and partially remove sky signal from
Figure 29. Plotted is the modulus of the nightly averaged (mean) visibilities for 9 nights in 977 kHz pixels. The 1st and last night are not used since they cover much smaller intervals of LST. Between the 9 nights the nighttime LST interval varies due to Earth’s revolution. Because this average is contributed to by both sky signal and correlated noise, to obtain the same contribution from the sky each night this average is computed over the common nighttime (23°54′ < LST < 14°31″), which is somewhat shorter than the full range of nighttimes (23°15′ < LST < 15°01″) over the 9 nights.

Figure 30. Same as in Figure 28 but with the nightly mean from Figure 29 subtracted from each night. This signal accurately repeats every 24 sidereal hours. Also visible are a few bright pixels noted above and a few bright stripes of several hours duration which do not repeat every night (e.g. on day 10 near 715 MHz appearing near 0 h and between 4 h and 5 h). These few non-repeating bright stripes may be due to RFI but also may be an indication of variability of the correlated noise.

8.3 Night-to-Night Variation

Figure 31 gives a quantitative projection of Figure 28 and Figure 30, before and after nightly mean subtraction. This subtraction greatly reduces the night-to-night variation both in absolute terms and as a fraction of the remaining signal. Subtraction of the nightly mean removes much of the correlated noise but also a significant fraction of the signal (gain times sky). Since the sky signal should be the same at the other directions. The controlled removal of a small fraction of the sky signal is easily modelled and accounted for when inferring the sky signal. Figure 29 shows the modulus of the nighttime time-averaged (mean) visibility as a function of frequency for nights 2 through 10. This average is roughly similar between the different nights but with significant (19 mK rms) variation out of an RMS value of 74 mK or 25%. The night-to-night variation is contributed to by both variation in the correlated noise and the gain and is much larger than the system noise inferred by the system temperature.

In Figure 30 these nightly mean visibilities are subtracted, revealing a fringe pattern as expected for bright localized sources on the sky. Detailed features are more apparent than in Figure 28 and these appear to repeat accurately each sidereal day. There are features which do not repeat: the few bright pixels noted above and a few bright stripes of several hours duration which do not repeat every night (e.g. on day 10 near 715 MHz appearing near 0 h and between 4 h and 5 h). These few non-repeating bright stripes may be due to external RFI but also may be an indication of variability of the correlated noise. The mean absolute visibility is 38 mK and the night-to-night median absolute deviation (a statistic which suppresses outliers) is 8 mK, which is comparable to that expected from the system noise temperature. In this regard 2V × 10V is better than most baselines where the night-to-night variation is significantly larger than the system noise. Our belief is that nearly all of the signal remaining in Figure 30 is sky signal.
same LST it does not contribute to night-to-night variation which can be due to variations in gain or in correlated noise. One would not expect that subtracting the nightly mean would decrease the fractional variation if the variation were only due to gain fluctuations so we infer that much of what was subtracted is correlated noise. Since we can reduce the correlated noise by subtracting off the nightly mean suggests that the correlated noise is roughly constant over the night. The large night-to-night variation in the un subtracted signal varies significantly between nights or, in other words, the correlated noise changes mostly during the daytime.

8.4 Average Sidereal Night

One can average all the nights’ visibilities into a single visibility which should have smaller noise than each of the individual nights. The averaging procedure used here is to take the median average of 1min $\times$ 1MHz pixels at the same LST and frequency to create an “average sidereal night” or ASN. This is what is shown in Figure 32 for the intersection of LSTs of the 9 complete nights. Median averaging suppresses the effect of outlying values, essentially removing non-repeating hot (or cold) pixels and stripes. There are no glaringly obvious “defects” in Figure 32.

The visibility patterns of figures Figure 30 and Figure 32 give the visual impression of a wavy surface colored with nearly horizontal rainbow stripes, like ribbon candy or a flag fluttering in a breeze. The horizontal rainbow stripes indicates a vertical gradient in phase of the visibility or fringe pattern. This is a consequence of the fact that most of the signal comes from near the NCP: the northern feed receives signals from the NCP before the southern feed which leads to a phase delay which increases linearly with frequency. Note that if all the signal came from an unpolarized source precisely at the NCP then the visibility pattern before nightly mean subtraction would be have perfectly horizontal stripes and nightly mean subtraction would remove the entirety of the signal. Figure 32 only shows the remnant of this fringe pattern which is not subtracted away because the sources are not precisely at the NCP. The wavy pattern is a manifestation of interference between a few bright sources. The horizontal axis is time so the rapid horizontal phase variation (rapid fringe rate) is indicative of sources that move rapidly as the Earth rotates. Sources near the NCP move slowly so only sources far from the NCP contribute significantly to the rapid fringe rate.

8.5 Polar Dephasing

Since much of the sky signal should come from near the NCP one can compare the vertical phase gradient to that expected for sources near to the NCP. The time delay of a signal precisely at the NCP is given by $\delta t_{NCP} = \frac{\hat{b} \cdot \hat{n}_{NCP}}{c}$, where $\hat{n}_{NCP}$ is the direction to the NCP and $\hat{b}$ is the baseline (the sign of $\hat{b}$ depends on the ordering of the feeds). In a horizontal (Earth) frame both $\hat{b}$ and $\hat{n}_{NCP}$ are constant in time and so is $\delta t_{NCP}$. For 2V $\times$ 10V the expected delay is $-24$ nsec so the corresponding phase gradient is $\frac{24}{c} \frac{\text{arcsec}}{\text{V}} = \delta t_{NCP} = \frac{24}{\text{MHz}}$, or 2.4 stripes from top to bottom, which one can verify by counting in 31 or 32. Another visual comparison is obtained by polar dephasing the visibility, meaning multiplying the visibilities by $e^{-i2\pi \nu \delta t_{NCP}}$, which should suppress the phase delay for sources near the NCP. The polar dephased visibility in Figure 33 has nearly all the vertical phase gradients removed, demonstrating that most of the signal does indeed come from sources near the NCP. What remains are slowly varying visibilities coming from sources near the NCP which move slowly due to Earth
8.6 Far Off Axis Sources

To accurately identify all the sources contributing to the ASN one would need a more accurate beam model than we currently have. However, the beam pattern almost certainly does not vary as rapidly as the fast fringes evident in the ASN. The fast fringes can only be from rapidly moving sources far from the NCP where the beam gain is low (\(\lesssim -30\) dB smaller than at the beam center); which means they can only come from a very few very bright far off-axis radio sources. The lack of source confusion of very bright sources allows us to accurately identify the sources of these fast fringes even with only a single baseline (one could do even better combining baselines). The ability to isolate the contribution from individual point sources using only a single baseline provides us with yet another independent calibration method for each baseline.

If the dish beam patterns are identical, then the contribution of an unpolarized point source, \(s\), to the visibility (in flux density units) is

\[
V_s[LST, \nu] = B[\hat{n}_s[LST], \nu] e^{-i2\pi \nu \hat{n}_s[LST] f^*_c} = B[\hat{n}_s, \nu] e^{-i2\pi \nu \hat{n}_s[LST] f^*_c}
\]

where \(B[\hat{n}_s, \nu] \equiv D[\hat{n}_s, \nu] / D[\hat{n}_{bc}, \nu]\), which is real and positive. Here, \(f^*_c\) is the flux density of the point source \(D[\hat{n}_s, \nu]\) is the directive gain, \(\hat{n}_{bc}\) is the direction of the beam center and \(\hat{n}_s[t]\) is the time-dependent direction to the source in the Earth frame (N.B. \(10 \log_{10} B\) is the quantity plotted in Figure 5). For any point source with a known trajectory on the sky one can measure

\[
\hat{V}_s \equiv \int d\nu \int dt e^{i2\pi \nu \hat{n}_s[LST]} V[LST, \nu, t]
\]

where \(\hat{V}_s\) is the visibility measured in any range of \(\nu\) and LST (larger ranges are better). The contribution of \(s\) to \(\hat{V}_s\) is \(\hat{B} \equiv \hat{V}_s / \int d\nu f^*_c / \int d\nu\) and \(\hat{B} = \int d\nu f^*_c / \int d\nu D[\hat{n}_s, \nu] / D[\hat{n}_{bc}, \nu] = f^*_c / \int d\nu f^*_c \int dLST\) are, respectively, the frequency average of \(f^*_c\) and the frequency and time average of \(B[\hat{n}_s, \nu]\) over the trajectory of the source through the sky. If one can determine \(\hat{V}_s\) and one knows \(\hat{B}\) then one can determine \(\hat{B}\) empirically: \(\hat{B} = \hat{V}_s / \hat{B}\). Of course \(\hat{V}_s\) is not the only contributor to \(\hat{V}_s\) but for sources which exhibit fast fringes the phase factor is highly oscillatory, which greatly suppress the contribution from illumination from other directions. With fast fringes \(\hat{V}_s\) is the majority contributor to \(\hat{V}_s\) even when it is a minor contributor to \(\hat{V}_s\) on the whole. This is a simple application of pattern matching, which is used to localize sources in the mapmaking in section 7, and is also used by one’s perceptual faculties, which can easily discern fast fringes with relatively small amplitude in figures 28, 30, 32 and 33.

We can apply this to the source Cas A. We have foreknowledge that Cas A is a very bright source and there are no other similarly bright sources around it. One can estimate how much \(\hat{V}_{\text{CasA}}\) might be contaminated by other sources of illumination by determining \(\hat{V}_s\) for \(\hat{n}_s\) near \(\hat{r}_{\text{CasA}}\), where we know there are no comparably bright sources. In Figure 34 plots \(\hat{V}_s\) for \(\hat{n}_s\) with the same declination as Cas A but varying R.A. Re \(\hat{V}_s\) peaks significantly near the R.A. of Cas A at a value 0.3 Jy while away from Cas A \(\hat{V}_s\) oscillates with amplitude \(\approx 0.1\) Jy. While these oscillations are partially due to Cas A itself, contribution will come from the fringe pattern of other sources far from Cas A. We conservatively estimate \(\hat{V}_{\text{CasA}} \approx 0.3 \pm 0.1\) Jy so using \(\hat{V}_{\text{CasA}} \approx 3\) kJy we find that \(\hat{B} \approx 10^{-4}\) or 40 dB down from the peak. This is in rough agreement with the electromagnetic simulations. In temperature units this corresponds to \(\sim 2.6\) mK fringes, a small fraction of the total signal. Further investigation of the uses of bright, far off-axis sources with single and combined baselines will be given in a subsequent paper.

8.7 Redundant Baselines

The wavy patterns of the visibilities from different baselines are generally quite different. However, for baselines with nearly the same \(b\) we expect the visibilities from the sky signal to be nearly the same. In Figure 35 we plot the ASNs of three baselines with nearly identical \(b\) and find that indeed the ASNs are nearly the identical. This gives support to our contention that after nightly mean subtraction the remaining signal is predominantly from the sky. One can discern a noticeable phase shift of \(\sim 0.6\) rad in the central panel relative to the other two which is evidence for differences in phase of the gain. Because the complex gain of each antenna was initially calibrated on Cas A before the dishes were pointed at the NCP, this difference in gains in the 9 night average is evidence for changes of the gain over the 10 days since this initial calibration.

8.8 Spectral Smoothness of Visibilities

The 21 cm signal is much smaller than that of the “foreground” sources we have examined so far. One feature that differentiates the foregrounds from 21 cm is the foregrounds have a smooth broadband spectrum while 21 cm emission comes in lines and is not smooth. This differentiating feature is confused by the phenomenon of “mode-mixing”, the fact that fine angular structures in the foreground emission will be aliased into relatively non-smooth spectral dependence of the visibilities. For example, while most of the frequency
dependence (horizontal fringes) of Figure 32 has been removed by polar dephasing in Figure 33, there still remain horizontal components of the fast fringes from bright off-axis sources. While one can possibly subtract the fringe patterns of a few known bright sources, this would become intractable for the multitudes of sources which contribute to mode mixing at the level we are interested in. A variety of techniques have been proposed to project out mode-mixed foregrounds from the 21 cm signal, and we will use them in Tianlai in the future, but here we examine a more conservative approach: limiting analysis to frequency modes which are not significantly mixed with foregrounds at the level of the system noise temperature. Here we quantify which modes these are. Foreground-contaminated frequency modes are sometimes said to be “in the wedge” and those not “outside the wedge”. Forecasts of the performance of intensity mapping often assume only modes outside the wedge are usable, so it is important to quantify where the wedge is!

There are various ways to quantify spectral smoothness of the visibilities. One is to decompose the visibility into frequency modes

\[ V[\nu_\alpha, \text{LST}] = \sum_{n=0}^{n_{ch}} a_n[LST] p_{n,\alpha} \]

where \( n_{ch} \) are discrete frequency channels, \( n_{ch} \) in number, and increasing \( n \) gives increasingly non-smooth frequency dependence. One obvious choice is a discrete Fourier transform (this is a delay decomposition, for example the \( \delta_{NCP} = -24 \text{nsec delay removed by polar dephasing} \). Instead, we choose a polynomial-based decomposition where the frequency dependence of the modes is approximately described by Legendre polynomials \( p_{n,\alpha} \propto P_n \left[ \frac{\nu_\alpha - \min}{\max - \min} - 1 \right] \) which are orthonormalized such that

\[ \frac{1}{n_{ch}} \sum_{n=0}^{n_{ch}} |V[\nu_\alpha, \text{LST}]|^2 = \sum_{n=0}^{n_{ch}} |a_n[LST]|^2. \]

Thus \( |a_n[LST]|^2 \) gives the contribution of mode \( n \) to the mean square \( V \). Just as with Fourier transforms in the case of white noise each mode amplitude, \( a_n \), is statistically independent with zero mean and identical variance

\[ \langle |a_n[LST]|^2 \rangle \]. Legendre polynomials are increasingly oscillatory (non-smooth) with increasing \( n \), also similar to a Fourier decomposition.

Applying this “Legendre” decomposition to the night-averaged and polar dephased visibilities we compute the “\( n \)-spectrum”, which is the mean value of \( |a_n[LST]|^2 \) averaged over all minutes of an ASN. Polar dephasing moves much of the \( n \in [5, 10] \) power into \( n = 0, 1, 2 \) but does not change the total power. In Figure 36 we plot these spectra for the mean visibility averaged over 1, 3 and 9 nights. Both mean and median averaging over nights (median was used in Figure 30) lead to nearly identical ASNs. Small differences are that outliers are not as suppressed with mean averaging (median is better for this) while the noise level of the mean ASN is smaller than the median ASN (mean is better for this). We see that for \( n = 0, 1, 2 \) the \( n \)-spectra seem to have “converged” in 1 night and for \( n \lesssim 8 \) after 3 nights. We would need more nights to see similar convergence at \( n \gtrsim 10 \). For \( n \gtrsim 50 \) the \( n \)-spectra are quite flat, consistent with white noise. The spectral flatness relies on limiting the band to \( \nu \in [700, 800] \text{ MHz as the noisy band edges tilt the spectrum redward} \). For \( n < 30 \) the spectra do fall off with increasing \( n \) but not nearly as fast as one would predict for the very smooth spectra of optically thin synchrotron or free-free emission. Much of this is due to mode mixing manifested by the fast fringes of bright off-axis sources which we could subtract if we knew the beam pattern better. Incomplete removal of correlated noise is another possible cause for the slower than expected fall-off with \( n \). Applying the same procedure to the daytime data we find this white noise tail only extends to \( n \approx 200 \).

The \( n > 100 \) white noise tail rms amplitude does “integrate down” \( \propto \#\text{nights}^{-1/2} \) when comparing the 1, 3 and 9 night average just as one would expect for zero mean noise which is uncorrelated between nights. To illustrate this we plot in Figure 36 the predicted \( n \)-spectra of the “system noise” predicted by the radiometer equation given the system temperature (\( T_{\text{sys}} = 112K \)) measured from the auto-
correlations (see below). For \( n > 100 \) the system noise account for almost all of the signal, leaving little room for contamination by other sources of noise or from sky signal. Thus at the level of sensitivity obtained with 9 nights of data from a single baseline the radio emission from the sky does indeed have a smooth spectrum in that it does not contain the hi-n part of the spectra, leaving \( \sim 75\% \) of the \( n \)-modes apparently free from foregrounds.

We do not find a \(#\text{nights}^{-1/2}\) scaling when using median nightly averaging but instead \( \sim \#\text{nights}^{-0.37} \). The source of this difference in scaling can be understood if one allows for night-to-night gain variations. A median will choose for each pixel the central visibility value, which can be from different nights for different pixels. If the gain varies from night-to-night then neighboring pixels in the median average can take visibility values from different nights with different gains. Night-to-night variation in the gain will be aliased into sharp features in the \( n \)-spectra as well as the visibility time series. This kind of aliasing would occur in any method of averaging RFI-flagged visibilities where neighboring pixels sample different sidereal days with different weights. Mean averaging gives equal weight to each night and therefore depends on the mean gain averaged all the nights, which is not expected to have sharp features in frequency or in time. If RFI is rare then RFI flagging has an advantage over median averaging because RFI flagging affects only a small fraction of the data whereas median averaging will introduce discontinuities everywhere. One can use the difference between mean and median averaging to quantify the level of night-to-night gain variation.

### 9 NOISE PERFORMANCE AND SENSITIVITY

#### 9.1 System temperature

One can define the system temperature, \( T_{\text{sys}}(\nu, t) \), for each feed antenna as the mean value of the auto-correlation visibilities in temperature units, \( V_{\nu} \). By definition these values are real and positive and we find them to be nearly constant in time and frequency. \( T_{\text{sys}} \) measures all the power coming into the correlator from a given feed: the thermal noise from the receiver and environment (ground and atmospheric/ionospheric), sky signal, and RFI, including self-generated RFI such as “cross-talk” (emission from other feeds). The thermal noise is much larger than the sky signal or the cross-talk and is very stable, as indicated by the nearly constant value of \( V_{\nu} \).

System temperature is a useful quantity because it defines the minimum value of the noise (fluctuations) in the visibilities. If \( T_{\text{sys}}^A \) is the system temperature of feed \( A \), then according to the ideal radiometer equation the variance of the visibility in a \( \delta \nu \times \delta t \) pixel is

\[
\Var(V_{\nu}^A \times B) = \frac{T_{\text{sys}}^A T_{\text{sys}}^B}{\delta \nu \delta t} \cdot
\]

This noise is unavoidable fluctuations in the sampled voltages, which are incoherent (random phase) time streams, whether from the instrument or external sources of radiation.

In Figure 37 the distribution of \( T_{\text{sys}}^A \) (averaged for the ASN of 3srcNP 20180101 in Table 2) is given by the abscissa of the magenta points for 27 of 32 feeds. During this run, 5 of the feeds were flagged as bad. Of the remaining feeds, 11H, 9V, 3V, 1V have have unusually high \( T_{\text{sys}} \): 343K, 307K, 232K and 156K. All other feeds have \( T_{\text{sys}} \in [106,127]K \) giving the largest cluster of points: 276 baselines. \( T_{\text{sys}} \) for the Tianlai dish receivers is expected to be dominated by receiver noise, \( T_{\text{rec}} \), and thermal emission from the ground, \( T_{\text{B}} \). \( T_{\text{rec}} \) is dominated by the LNA noise temperature, \( \sim 35 \text{K} \), and we calculate, based on electromagnetic simulations and UAV measurements, that \( 25 \text{K} \lesssim T_{\text{spill}} \lesssim 80 \text{K} \) when the antennas are pointed either at the zenith or at the NCP (\( T_{\text{spill}} \) is similar for either pointing because the far sidelobes are approximately constant). Emission from the sky (atmosphere and astronomical sources) is negligible compared to these sources in the 700 – 800 MHz band, so we expect \( T_{\text{sys}} \sim 35 + XX \text{K} \) for the NCP data.

#### 9.2 Noise temperature

A visibility \( V_{\nu} \) (\( A \) and \( B \) are feed antennas) in a \( \delta \nu \times \delta t \) pixel is an average of \#sample numbers, \( v_A v_B \). One cannot estimate \( V_{\nu} \) from \( V_{\nu}^A \) alone without some knowledge of how \( V_{\nu} \) behaves. One can usually assume that \( V_{\nu} \) varies more slowly with time and frequency than the pixel spacing, \( \delta t \) and \( \delta \nu \), so the difference between \( V_{\nu} \) in neighboring pixels gives an indication of \( \Var(V_{\nu}^A) \). Similarly if one hi-pass filters \( V_{\nu} \) in \( \nu \) and/or \( t \) one could remove the slowly varying \( V_{\nu} \) and what remains would be only the noise, effectively setting \( \langle V_{\nu} \rangle = 0 \). In section 8.8 we believe we have done just that when considering the flat spectrum \( \nu > 100 \) modes.

If \( V_{\nu} \) and \( V_{\nu}^B \) are constant and \( V_{\nu} \) is 0, then the Legendre decomposition of \( V_{\nu} \) will have \( \langle a_{n} \rangle = 0 \) and \( \langle a_{n}^{B} \rangle = \langle a_{n}^{A} \rangle \langle a_{n}^{B} \rangle / \#\text{sample} \) for all \( n \) if \( A \neq B \) and for \( n = 0 \) if \( A = B \). While this would be a poor model for \( n < 100 \), where clearly \( \langle V_{\nu} \rangle = 0 \), we expect it to be a good model for \( n > 100 \), which seems consistent with \( \langle V_{\nu} \rangle = 0 \). From the definition of the Legendre decomposition, \#sample = \( \Delta \nu \Delta t \), where \( \Delta \nu \) gives the entire bandwidth sampled and \( \Delta t \) is the entire time sampled. The ASN bandwidth used is \( \Delta \nu = 100.1 \text{MHz} \) and the pixel duration is 1 sidereal minute (59.836 sec) averaged over 9 nights. Every 240 sec an interval of 7 sec surrounding the broadcast of the calibration noise source was excluded, reducing \( \Delta t \) by 233/240 to 8.71 min. Thus, for an ASN, \#sample = 100.1 MHz \( \times 8.71 \text{min} \times 5.2334 \times 10^{10} \text{h} \)

For each feed pair \( A \times B \) one can define a “noise temperature”

\[
T_{\text{noise}}^{A \times B} = \sqrt{\#\text{ASN} \times \text{sample} \times a_{n}^{\text{rms}}},
\]

where \( a_{n}^{\text{rms}} \) is the rms mode amplitude for \( n > 100 \) mean
averaged over the entire ASN: 871 sidereal minutes × 311 modes. If these modes are not significantly contaminated by correlated noise or signal, then we expect

\[
\left( \frac{T_{\text{noise}}}{A \times B} \right) = T_{\text{sys}} \times B_n = \sqrt{T_{\text{sys}}^A / B_n}
\]

where \( T_{\text{sys}} \) and \( T_{\text{sys}}^B \) are the mean system temperature averaged over the ASN.

For our representative baseline 2V × 10V we find \( T_{\text{noise}}^{2V \times 10V} = 111.1K \) whereas we would expect \( T_{\text{sys}}^{2V \times 10V} = 109.75K \) from \( T_{\text{sys}}^{2V} = 114.44K \) and \( T_{\text{sys}}^{10V} = 105.24K \). We make the same comparison for all 27 × 27 baselines in Figure 37. For nearly all baselines the noise temperature is very close to the system temperature prediction, the exceptions being auto-correlations of the “hot” antennas. If we restrict to cross correlations of non-hot antennas one finds the noise temperature typically only a few percent larger than the system temperature. For the ratio \( (T_{\text{noise}} / T_{\text{sys}}) \), we find mean and standard deviation 1.044 ± 0.018 which can be compared to the ideal expectation 1 ± 1/\( \sqrt{571 \times 311} \) = 1 ± 0.0019. While this is close, closer than our human expectation, this discrepancy is quite statistically significant according to our noise model. This discrepancy is likely due in part to a deficiency in the noise model but another part may be due to correlated signal or noise leaking into the n < 100 modes. A 4.5% excess would correspond to \( a_{\text{corr}}^\text{rms} \sim 0.1 \text{mK} \) or rms pixel noise \( \sim 2 \text{mK} \).

The system noise is expected to be uncorrelated only between frequency channels but also between time bins. The flatness of the n-spectra for n > 100 only demonstrates the lack of correlation in the frequency direction. Temporal correlations in the visibilities are expected from the sky signal since this only varies slowly as the earth rotates. To check that the instrument is not introducing additional time correlations one can must first filter out the temporal correlations from the sky. Since the n ≥ 100 modes are apparently free of sky signal one can do this by removing the n < 100 modes. In figure WWW we plot the temporal correlation function of the visibilities after filtering out the n < 100 modes. We find (I hope) no significant temporal correlations between sidereal minutes as expected.

### Figure 37
System temperature and noise temperature (in Kelvin. See definitions in text.). Magenta points are individual feeds.

### Figure 38
Histogram of the system temperature (in Kelvin) for 480 different cross-correlations (does not include autocorrelations). Using the point source calibration procedure described in sections 5.2 and 6, the system temperatures during observations of the NCP in January, 2018 (3srcNP 20180101 in Table 2) are calculated for all 480 cross-correlation visibilities.

### Figure 39
Overlapping Allan deviation (Schieder & Kramer 2001) vs. integration time, 10, for four typical baselines at 747.5 MHz during night time only. The plot shows that the noise integrates down as 1/√τ, as expected, for about 300 seconds.

### 9.3 Sensitivity vs Integration Time
Theoretically, the root-mean-square (rms) noise for a visibility is given by the radiometer equation

\[
\Delta T = \frac{\sqrt{T_{\text{sys}}^A \Delta \nu}}{2}
\]

where \( T_{\text{sys}} \) is the system temperature, \( \Delta \nu \) is the bandwidth and 1 is the integration time. If we double the integration time the noise standard deviation will decrease by a factor of \( \sqrt{2} \). We analysed how the noise level decreases with the integration time for baseline 11H - 13H. We use \( \Delta \nu = 244 \text{ kHz} \) and take the integration time to be \( t = 1, 30, 60, 120, 300, 600, 1200 \) and 2400 s. A plot of \( \Delta T_{\text{rms}} \) vs. integration time is shown in Fig. 39. We also plot the prediction from the radiometer equation. The plot shows that the noise integrates down with time as expected for about 300 seconds.
10 CONCLUSION

Future work includes foreground removal (not discussed here at all), retuning to lower redshift to overlap with the North Celestial Cap Redshift Survey (NCCRS) and the planned North Celestial Cap Survey (NCCS) for cross-correlation analysis. We plan to draw comparisons with the Tianlai Cylinder Array Pathfinder.

11 ACKNOWLEDGEMENTS

Work at UW-Madison and Fermilab is partially supported by NSF Award AST-1616554. Fermilab is operated by Fermi Research Alliance, LLC, under Contract No. DE-AC02-07CH11359 with the US Department of Energy. Work at NAOC is supported by the MOST grants 2016YFE0100300 and 2012AA121701, the NSFC 11473044 and 11633004, and the CAS QYZDJ-SSW-SLH017. Part of the computation is performed on the Tianhe-2 supercomputer with the support of NSFC grant U1501501. Beam measurements with the UAV are supported by NSFC grant U1631118. Authors affiliated with French institutions acknowledge partial support from CNRS (IN2P3 & INSU), Observatoire de Paris and from Irfu/CEA. This document was prepared by the Tianlai Collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359.

REFERENCES

Bandura K., et al., 2014, in Ground-based and Airborne Telescopes V. p. 914522
Das S., et al., 2018, in Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy IX. p. 1070836
Kovetz E., et al., 2019, BAAS, 51, 101
Zhang J., et al., 2019, IEEE-AP

APPENDICES

Graphical Representation of Visibilities

We represent the complex visibilities graphically in terms of colors as illustrated in Figure Zz. A visibility is a complex number depending on two real parameters, the real and imaginary part. A complex number may be represented graphically with a single color since the human visual perceptual representation of colors is roughly a 3-dimensional space. We use red/yellow/green/blue/magenta/red so it is therefore natural to associate the complex phase with the hue. In the hue/saturation/brightness (HSB or HSV) representation of colors it is most natural to associate the magnitude of the complex number with the brightness. Complex number with modulus greater than this maximum value will have saturated brightness. Thus we use red to represent numbers with complex phase zero, i.e. positive real numbers. Black has zero brightness irrespective of the hue and we will use it to represent 0 whose argument is undefined. There is no need to make use of the saturation. We generally use a linear relation between brightness and the modulus of the visibility adjusting the proportionality factor such that the smaller of the maximum modulus or 3 times the median modulus across the image saturates the brightness. Complex number with modulus greater than this maximum value will have saturated brightness. Thus this color representation has only a finite dynamic range for modulus. We find that this choice of saturation value results in a visually appealing representation of complex visibilities.

This paper has been typeset from a TeX/LATEX file prepared by the author.
Figure 40. The color palette used to represent complex visibilities in this paper.