



## The Tianlai Cylinder Pathfinder Array: System Functions and Basic Performance Analysis

Jixia LI<sup>1,2</sup>, Shifan ZUO<sup>3,1</sup>, Fengquan WU<sup>1</sup>, Yougang WANG<sup>1</sup>, Juyong ZHANG<sup>4</sup>, Shijie SUN<sup>1,2</sup>, Yidong XU<sup>1</sup>, Zijie YU<sup>1,2</sup>, Reza ANSARI<sup>5</sup>, Yichao LI<sup>6</sup>, Albert STEBBINS<sup>7</sup>, Peter TIMBIE<sup>8</sup>, Yanping CONG<sup>1,2</sup>, Jingchao GENG<sup>9</sup>, Jie HAO<sup>10</sup>, Qizhi HUANG<sup>1</sup>, Rui LI<sup>11</sup>, Donghao LIU<sup>1</sup>, Yingfeng LIU<sup>1,2</sup>, Tao LIU<sup>4</sup>, John P. MARRINER<sup>7</sup>, Chenhui NIU<sup>1</sup>, Ue-Li PEN<sup>12</sup>, Jeffery B. PETERSON<sup>13</sup>, Huli SHI<sup>1</sup>, Lin SHU<sup>10</sup>, Yafang SONG<sup>10</sup>, Haijun TIAN<sup>14</sup>, Guisong WANG<sup>9</sup>, Qunxiong WANG<sup>14</sup>, Rongli WANG<sup>4</sup>, Weixia WANG<sup>11</sup>, Xin WANG<sup>15</sup>, Kaifeng YU<sup>1,2</sup>, Jiao ZHANG<sup>16</sup>, Boqin Zhu<sup>1</sup>, Jialu ZHU<sup>4</sup>, and Xuelei CHEN<sup>\*1,2,17</sup>

<sup>1</sup>National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China;

<sup>2</sup>School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China;

<sup>3</sup>Center for Astrophysics and Department of Astronomy, Tsinghua University, Beijing 100084, China;

<sup>4</sup>Hangzhou Dianzi University, 115 Wenyi Rd., Hangzhou 310018, China;

<sup>5</sup>IJC Lab, CNRS/IN2P3 & Université Paris-Saclay, 15 rue Georges Clémenceau, 91405 Orsay, France;

<sup>6</sup>Department of Physics and Astronomy, University of the Western Cape, Belville 7535, Republic of South Africa;

<sup>7</sup>Fermi National Accelerator Laboratory, P.O. Box 500, Batavia IL 60510-5011, USA;

<sup>8</sup>Department of Physics, University of Wisconsin Madison, 1150 University Ave, Madison WI 53703, USA;

<sup>9</sup>The 54th Research Institute, China Electronics Technology Group Corporation, Shijiazhuang, Hebei 050051, China;

<sup>10</sup>Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China;

<sup>11</sup>Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi, Xinjiang 830001, China;

<sup>12</sup>Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, Ontario, M5S 3H8, Canada ;

<sup>13</sup>Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA;

<sup>14</sup>Center for Astronomy and Space Sciences, China Three Gorges University, Yichang 443002, China;

<sup>15</sup>School of Physics and Astronomy, Sun Yat-Sen University, Guangzhou 510297, China;

<sup>16</sup>College of Physics and Electronic Engineering, Shanxi University, Taiyuan, Shanxi 030006, China;

<sup>17</sup>Center of High Energy Physics, Peking University, Beijing 100871, China

Received ; accepted

The Tianlai Cylinder Pathfinder is a radio interferometer array designed to test techniques for 21 cm intensity mapping in the post-reionization Universe, with the ultimate aim of mapping the large scale structure of the Universe and measuring cosmological parameters such as the dark energy equation of state. Each of its three parallel cylinder reflectors are oriented in the north-south direction, and the array has a large field of view. As the Earth rotates, the northern sky is observed by drift scanning. The array is located in Hongliuxia, a radio quiet site in Xinjiang, and saw its first light in September 2016. In this first paper on the data analysis of the Tianlai cylinder array, we describe the system functions of the array, the commissioning observations during 2016-2018, and present its basic performance characteristics. We show examples of the interferometric visibility data, then using the observational data, we derive the actual beam profile in the east-west direction, the bandpass response, and calibrate the complex gains for the array elements either with a strong astronomical point source, or with an artificial calibrator source. Based on the preliminary analysis of the data, we derive the system temperature and sensitivity of the array.

**Interferometer, radio astronomy**

**PACS number(s):** 95.55.Br, 95.55.Jz, 95.75.Kk, 95.85.Bh

**Citation:**

Li J X et al, The Tianlai Cylinder Pathfinder Array: System Functions and

Basic Performance Analysis, Sci. China-Phys. Mech. Astron. , 000000 (2020), doi:

\*Corresponding author: xuelei@cosmology.bao.ac.cn

## 1 Introduction

This paper describes our analysis for the performance characteristics of the Tianlai Cylinder Pathfinder. This array consists three parallel cylindrical reflectors, each has a width of 15 meter and a length of 40 meter, the long axis is aligned in the true north-south (N-S) direction. It is co-located with the Tianlai Dish Pathfinder, an interferometer array of 16 dishes of 6 meter aperture. These complementary designs were chosen specifically for testing different approaches to 21 cm intensity mapping. Both saw their first light in 2016. An aerial photograph of the Tianlai arrays site is shown in Fig. 1.

Intensity mapping is a technique for making rapid low angular resolution survey of the large scale structure of using a redshifted spectral line [1], without attempting to resolve individual galaxies. The 21 cm line emitted by neutral hydrogen is of particular interest [2, 3]. So far, the 21 cm signal has only been marginally detected by this method in cross-correlation with optical galaxy redshift survey by two instruments: the Green Bank Telescope (GBT) [4, 5] and the Parkes Observatory [6]. Nevertheless, 21 cm intensity mapping is considered to be a potentially powerful probe of cosmology: measuring the equation of state of dark energy from the baryon acoustic oscillation (BAO) features (e.g. [7]), constraining models of inflation (e.g. [8]), as well studying the Epoch of Reionization (EoR) ( $6 \lesssim z \lesssim 20$ ) Cosmic Dawn ( $z \sim 20$ ) and the dark ages ( $z \gtrsim 24$ ). A primary observable for cosmological studies is the three-dimensional power spectrum of the underlying dark matter, which is traced by neutral hydrogen and intensity mapping provides a natural means to compute this spectrum over a broad range of wavenumbers,  $k$ , in which the perturbations are in the linear regime.

While HI intensity mapping is being used to study the EoR and Cosmic Dawn by a variety of meter wavelength instruments, including LOFAR [9], MWA [10], HERA [11], PAPER [12], and LWA [13], 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: CHIME [14], Tianlai [7, 15-17], HIRAX [18], OWFA [19], and BINGO [20]. Other instruments are being designed and built to test the technique include BMX and PAON-4 [21, 22]. With the exception of BINGO, they are all interferometer arrays with large numbers of receivers in order to provide enough mapping speed to detect the small 21 cm signal; and all are laid out in a compact configuration in order to provide sensitivity at the relatively large scales ( $0.5 \gtrsim k \gtrsim 0.05$ ) where the BAO features appear in the power spectrum.

The 21 cm signal is 4 ~ 5 orders of magnitude smaller than the foreground emission (primarily synchrotron radia-

tion) from Galactic and extra-galactic radio sources (see e.g. Ref.[23]), making the detection a great challenge. Extracting the 21 cm signal generally relies on the fact that foreground emissions are smooth functions of frequency, while the 21 cm spectrum has a structure arising from the large-scale distribution of matter along the line of sight [2]. However, instrumental effects can introduce structure into the spectrum of otherwise smooth foregrounds. In particular, the antenna and array beam patterns are frequency dependent, which aliases the angular dependence of the bright foregrounds into frequency dependence of the visibilities. Removing this so called "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.

While many radio interferometers track targeted regions of the sky, most of the post-EoR instruments observe the sky by drift scanning. 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 such as Tianlai must wait for bright sources to pass through the field, or attempt to calibrate on dimmer sources. It is important to check the stability of the system, which greatly affects the sensitivity to the 21cm signal.

Before using the array to make astronomical observations, it is necessary to first check and understand its basic performance characteristics. In this work, we make an analysis of basic performance characteristics of the Tianlai cylinder pathfinder array based on test observation data. We first review the instrument design of the system, its observations during the 2016-2018 test run, and the data processing procedure in Section 2, then present some system tests in Section 3. We examine the raw output data from the array in Section 4, and derive the beam pointing and width from transit of strong source in Section 5. In Section 6 we describe the calibration of the system and study its stability. In Section 7 we estimate the system temperature of the radio telescope. Finally, we summarize our results in Section 8.

## 2 An Overview of Instruments, Observations and Data Processing Procedures

An interferometer array works by having a number of receiving units (antenna/feed combination) distributed in space, each of them generates an electric voltage signal,

$$\mathcal{E}_a(\nu) = g_a \int d^2\vec{n} A_a(\vec{n}, \nu) E(\vec{n}, \nu) e^{-i2\pi\frac{\nu}{c}\vec{n}\cdot\vec{x}_a} + n_a(\nu) \quad (1)$$



**Figure 1:** A top view photograph of the Tianlai cylinder and dish arrays. The dish is located in the east side of the cylinder, the calibrator noise source (CNS) is located on a hill to the west of the arrays, as marked on the left side of the photograph. There is a blue cabin for storage between the array and the CNS.

where  $A_a(\vec{n}, \nu)$  is the antenna response to the electric field  $E(\vec{n}, \nu)$  from the direction  $\vec{n}$  at frequency  $\nu$ ,  $g_a$  is the receiver gain,  $\vec{x}_a$  is the position of the antenna, and  $n_a$  is the noise. The short time averaged correlation between the voltage outputs is called an (interferometric) visibility,

$$V_{ab}(\nu) \equiv \langle \mathcal{E}_a^* \mathcal{E}_b \rangle_t.$$

The visibility is related to the sky intensity by

$$V_{ab} \propto g_a g_b \int d^2\vec{n} B_{ab}(\vec{n}, \nu) I_\nu(\vec{n}) \exp[-i2\pi\vec{u}_{ab} \cdot \vec{n}] + \langle n_a^* n_b \rangle_t, \quad (2)$$

where  $\langle E^*(\vec{n})E(\vec{n}') \rangle \propto I_\nu(\vec{n})\delta^2(\vec{n} - \vec{n}')$  is the sky intensity in the direction  $\vec{n}$  at frequency  $\nu$ ,

$$\vec{u}_{ab} = \frac{\nu}{c}(\vec{x}_b - \vec{x}_a)$$

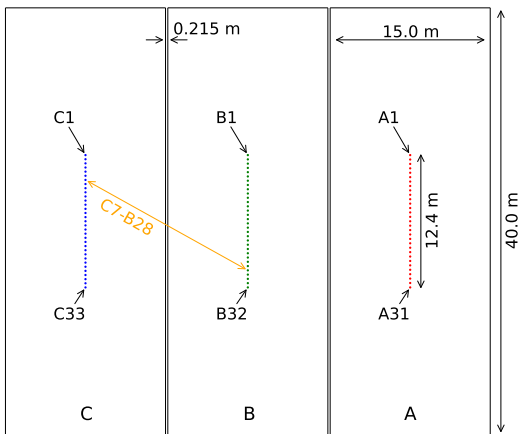
is the baseline vector between antenna or feed  $a$  and  $b$  in units of wavelength.  $B_{ab}(\vec{n}, \nu) = A_a^*(\vec{n}, \nu)A_b(\vec{n}, \nu)$  is the primary beam response for the pair, and can often be simplified as  $B(\vec{n}, \nu)$  if the array is uniformly constructed and each antenna/feed is pointing in the same direction. The uncorrelated noise satisfy  $\langle n_a^* n_b \rangle = \delta_{ab} T_{sys}$ . This equation can also be generalized to the polarized case. The hardware instrument of the interferometer array is designed to produce this visibility data. From the measured visibilities, the sky intensity  $I_\nu(\vec{n})$  can be recovered by synthesis [24].

## 2.1 Instrument

The Tianlai project is now running in a pathfinder stage. The cylinder array consists of three adjacent parabolic cylinder reflectors, each  $40\text{m} \times 15\text{m}$  with its long axis oriented in the N-S direction. The cylinders are closely spaced in the East-West (E-W) direction, as shown in Fig. 2. Dual-polarization dipole feeds are placed along the focus line of each cylinder. The reflectors focus the incoming radio signal in the E-W direction, while allowing a wide field of view in the N-S direction. The Tianlai cylinder reflector is fixed on the ground, at any moment its field of view (FoV) is a narrow strip running from north to south through the zenith. While the reflectors allow the FoV to run from horizon-to-horizon in the N-S direction, the beam of the feeds limit the strip to  $\pm 60^\circ$  from zenith [25]. As the Earth rotates, it can scan the northern celestial hemisphere. (The latitude of the telescope site is  $44^\circ$ , so the FoV extends from  $-16^\circ$  declination up to  $+90^\circ$  and back down to  $+76^\circ$  on the other side).

From east to west, the 3 cylinders are denoted cylinder A, B, C respectively. Each has been installed with a slightly different number of feeds, 31, 32 and 33 respectively. From north to south, the feeds in each cylinder are labelled in numbers 1, 2, 3, ..., ... The northernmost (or southernmost) feeds A1, B1, C1 (or A31, B32, C33) are aligned, and the distance between the northernmost and southernmost feeds are 12.4 m. The currently installed feeds occupy less than half of the cylinder; the remaining space is reserved for additional

feeds for future upgrades. Since the feeds are evenly distributed, this results in different feed spacings for each different cylinders, which are 41.33 cm, 40.00 cm and 38.75 cm respectively. This arrangement is made to reduce the grating lobe, which is generated due to the degeneracy in arrival time for different directions when the spacings between adjacent feeds are larger than half a wavelength. As the spacings of the three cylinders are slightly different, their grating lobes are also slightly different, so the grating lobe is reduced [16]. Each dual linear polarization feed generates two signal outputs. We will use X to denote the output along the N-S direction and Y along E-W direction. Each signal channel is designated by its cylinder, feed number and polarization basis. For example, the E-W polarized output of the 2nd feed in the middle cylinder will be noted as B2Y. The baseline between two feeds is denoted by its two components linked with a hyphen; for example the baseline C7-B28 is shown in Fig. 2, and the cross-correlation between their X-polarization channels is denoted as C7X-B28X.



**Figure 2:** The Tianlai Cylinder Array. The cylinders are aligned in the N-S direction, with a gap of 0.215 m between adjacent ones. The three cylinders are designated as A, B, C from east to west, and have 31, 32, 33 feeds respectively. The feeds in each cylinder are evenly distributed, with the ones at both edges (A1, B1, C1 in the north and A31, B32, C33 in the south) aligned with each other. The baseline C7-B28 is depicted by the orange double end arrow as an example.

The system consists of the antennas (cylinder reflectors and feeds) and the optical communication system (optical transmitter/receiver and cable), which converts the radio frequency (RF) electric signal to optical signals sent via optical fiber to the station house, which is located about 6 km away in the nearby Hongliuxia village. In the station, the system is housed in separate analog and digital electronics rooms. The optical signal is converted back to the RF electric signal, then down-converted to the intermediate frequency (IF). The IF signal is then digitized and processed in the digital

correlator. The whole system is designed to operate over a wide range of frequencies (400–1500 MHz), while the working frequency band is set by replaceable band pass filters. At present the bandpass is set to 700–800 MHz, corresponding to redshift  $1.03 > z > 0.78$  for the 21 cm line. A summary of the design parameters of the cylinder array is given in Table 1, and a schematic of the system is shown in Fig. 3.

**Table 1:** Main design parameters of Tianlai cylinder array.

Number of cylinder	3
Reflector N-S length	40.0 m
Reflector E-W diameter	15.0 m
f/D	0.32
Surface error (design)	$< \lambda/20$ at 21 cm
Number of feeds per cylinder	31(A), 32(B), 33(C)
Feed spacing (cm)	41.33, 40.00, 38.75
Feed illumination angle	152°
Number of data channels	192
Gain (design)	$> 20$ dBi
Adjustable frequency range	500–1500 MHz
Current frequency range	700–800 MHz
Frequency resolution	122 kHz
X-pol(N-S) FWHM @750 MHz	1.6° (H-plane), 62.2° (E-plane)
Y-pol(E-W) FWHM @750 MHz	1.8° (E-plane), 71.4° (H-plane)
Location	E 91°48', N 44°9'

In detail, the wide band low noise amplifiers (LNAs) are mounted on the back plate of the feed. The amplified RF signals pass through 15-meter long coaxial cables to the optical transmitters mounted under the cylinder antennas. In the transmitter box, the RF signal amplitude modulates a distributed feedback laser diode (DFB-LD) to generate the RF optical signals. The typical intensity of the laser output power is 3 dBmW, and a built-in thermal electric cooler module controls the temperature of the laser diode, to keep the output intensity at a stable value. The output laser is coupled into a single-mode fiber pigtail and ferrule connector/angled physical contact (FC/APC) connector, which connects the optical fiber link. The optical cable carries a bundle of 288 fibers and the length of the cable is 8 km. The typical loss in the optical fiber link is 4.8 dB at the operating wavelength of  $1.55\mu\text{m}$ , and to compensate for this, the analog optical transceiver system is designed with an end-to-end gain of 13 dB, which varies less than 2 dB within the band. At the re-

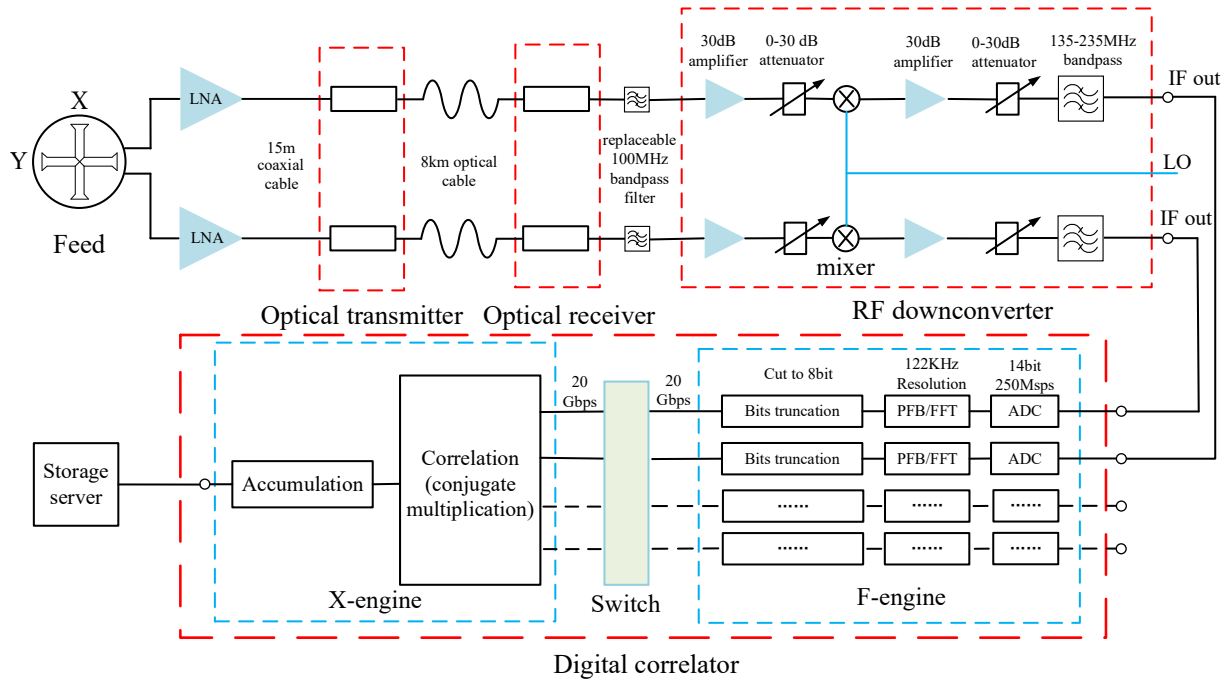


Figure 3: A schematic of the Tianlai analog circuit system.

ceiving end, a PIN photo-detector demodulates the RF signal. The receiver modules are housed in box cases; each handles 32 signal channels. The typical noise figure for the optical link system is 17 dB. However, this does not significantly increase the noise temperature for the signal which has already been amplified by the LNA.

An analog mixer down-converts the RF signal to the IF band of 125–235 MHz. The observing frequency band can be adjusted by shifting the frequency of the local oscillator (LO). Currently, the observing frequency band is 700–800 MHz. The LO is fixed at 935 MHz with a Rubidium standard frequency reference, with maximum phase noise of  $-90$  dBc/Hz at 10 kHz and  $-90$  dBc/Hz at 100 kHz. The image frequency rejection ratio for the down-converter is larger than 60 dB. Replaceable cavity band pass filter modules are placed between the optical receiver RF output and the mixer input to reject the out-of-band noise; its insertion loss is less than 0.7 dB. The IF channel output power level can be adjusted within the range of 0 ~ 60 dB by attenuators, which are controlled by a serial communication port, and the gain adjustment step is 1 dB. The gain flatness is less than 2 dB for the IF bandwidth. For the IF channel output, the minimum 1 dB compression point is 15 dBm.

Finally, the IF signal is sent to the digital backend system through bulkhead connectors between the analog and digital rooms.

The digital backend system has a commonly used FX (Fourier transformation and correlation) framework. It includes three parts: the F-engine which performs the Fourier

transform, the X-engine which performs cross-correlations of the Fourier-transformed data, and the network switch, which performs the so called “corner turning”, i.e. it distributes the data from different input channels of the same frequency to a unit on the X-engine for computing the cross-correlations.

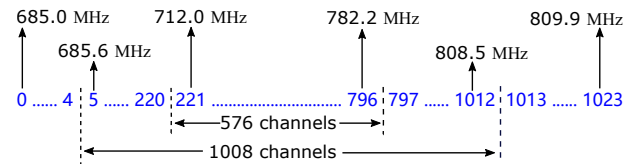


Figure 4: The relationship between the FFT channels and the radio frequency. Though the correlator does a 1024-point FFT, only 1008 channels are cross-correlated. Due to limited X-engine computation ability, currently only 576 channels have been computed.

The F-engine consists of twelve FPGA boards, each equipped with two 8-input analog-to-digit converters (ADCs) of the type ADS62P49. These ADCs digitize simultaneously the 192 input channels, and work at a sampling rate of 250 Mega samples per second (MSPs) and with a length of 14 bits. The sampler is synchronized by a 10 MHz reference signal output from a GPS receiver. The time-sequential data is Fast Fourier Transformed (FFT) with a block length of 1024, though due to hardware limitations 16 of the frequency channels near the edge are discarded and only 1008 frequency channels are recorded. Given the 250 MSPs sampling rate, each frequency channel has a width of  $\delta\nu = \frac{250}{2 \times 1024}$  MHz = 122.07 kHz. However, due to some practical problems with the correlator hardware, in most runs we only take the data

for the central 576 frequency channels, which span a bandwidth of about 70 MHz from 712.0 MHz to 782.2 MHz, as shown in Fig.4.

The FFT results are packed in several frequency blocks and sent to a 16 Gb×128 RapidIO high speed switching network. The data of the same frequency bands from all input channels are switched to the same DSP boards in the X-engine. Each DSP board undertakes the cross-correlation of one block of frequency bands for all input channels. A total of 27 DSP boards are used. The correlation results from all of the DSP boards are integrated on an FPGA board, which outputs the visibility to a storage server and the data is dumped to hard drives in HDF5 format. The integration time is adjustable; typically we adopt 4 seconds (more precisely 3.995 s for integer multiples of integration cycle), which the system can handle well and during which the drift of the sky due to Earth rotation is negligible for the angular resolution of our system. The raw visibilities are arranged in the form of [time, frequency, correlation], with correlation being the fastest changing index. The cross correlation between all of the 192 inputs are computed, thus forming a total of 18528 correlation pairs (192 auto-correlations plus 18336 cross-correlations).

To monitor the variation of instrumental phase, an artificial calibrator is installed on a small hill to the west of array, to broadcast wide band RF noise periodically. Below we shall call it the Calibrator Noise Source (CNS). The device consists of a small broadband disc-cone antenna erected on a wooden pole, and a diode noise generator enclosed in a temperature-controlled box. Its temperature is kept at 21°C by a thermostat, with a maximum variation of  $\pm 0.1^\circ\text{C}$  per day or  $\pm 2^\circ\text{C}$  in a year. The CNS is placed at the top of a hill about 131.5 meters from the center of the array in the North-West direction of the array, and its elevation relative to the center of the array (including the hill and wooden pole height) is about 13.1 meters.

The CNS is turned on and off with a period of 240 seconds. Initially the noise source broadcasted for 20 seconds; but after December 2017 this was reduced to 4 seconds; giving a duty cycle of 8.3% and 2.1% respectively. The amplitude of the CNS is adjusted with an attenuator, so that as received by the feed near the center of the cylinder array, the CNS induced auto-correlation is about 5 dB stronger than the Sun for the X-polarization and 10 dB for Y-polarization.

## 2.2 Observations

The Tianlai cylinder array does not have moving parts and is designed to run continuously once powered on, though in practice it runs a few days at a time. Once the digital correlator has started up properly, one can run the data acquisition

program. The data are automatically collected and saved to hard drives. The drives are shipped to NAOC for off-line processing and analysis.

The Tianlai cylinder array saw its first light in September, 2016. Test observation data were collected in 2016 – 2018. The observations were usually halted after a few days when some problems were discovered, or when some hardware malfunction prevented normal operation. Currently, we have collected a total of 114 days' data, as listed in Table 2.

**Table 2:** List of Tianlai Cylinder Array Data Sets.

Data set	Lengths	Malfunction Channels
2016/09/27 20:15:37	5 days	A18Y,A28,B19,C7Y, C9X,C14X,C16,C29Y
2016/10/11 00:49:48	5 days	A18Y,A28,B19,C7Y, C9X,C14X,C16,C29Y
2016/12/31 20:51:54	17 days	Many
2017/02/13 19:23:07	5 days	A18Y,A24Y,B6Y,B26Y, C15X,C7Y
2017/02/24 19:31:07	9 days	A18Y,A24Y,B6Y,B26Y, C15X,C7Y,C18~C33
2017/08/21 21:26:39	3 days	A26X,B21X,C3X,C16 A17~A24,C26~C33
2017/09/03 14:32:17	9 days	A26X,B21X,C3X,C16 A17~A24,C26~C33
2017/09/22 01:33:18	7 days	A26X,B21X,C3X,C16 A17~A24,C26~C33
2017/09/29 21:42:59	13 days	A26X,B21X,C3X,C16 A17~A24,C26~C33
2017/12/09 19:21:54	10 days	A18Y,A24Y,A26X,B26Y, B31X
2017/12/20 19:22:02	4 days	A18Y,A24Y,A26X,B26Y, B31X
2018/01/21 00:05:35	14 days	A18Y,A24Y,A26X,B26Y, B31X,C28Y
2018/03/22 18:07:58	9 days	A18Y,A24Y,A26X,B26Y, B31X,C28Y
2018/03/31 17:08:12	4 days	A18Y,A24Y,A26X,B26Y, B31X,C28Y
Total	114 days	

The data sets in the table are designated by the starting time of observation in the Beijing standard time (UTC+08h), and the duration of each observation run is listed in the second column of the Table. In the third column, we also list the malfunctioning channels during that run. These malfunctions are limited to those particular channels and do not affect the running of the whole system. Some malfunctions were identified during the run (e.g. if their output is too weak) while some were only found after offline analysis (e.g. the

output is too noisy). The visibilities involving these malfunctioning units are ignored in the analysis. There were some problems in the correlator for the 2016/12/31 data set which render much of the data set unusable.

The present work is based on the analysis of the first light data set, which started on September 27th, 2016, (below we denote it as 2016/09/27) and was used in our early analysis [26], and a second data set, which started on March 22nd, 2018 (denoted as 2018/03/22). These two data sets consist of a total of 10 days' observation, and cover both the autumn and spring sky. For the 2016/09/27 data set, 20 frequency channels are processed while for the 2018/03/22 data set all 576 frequency channels are processed.

### 2.3 Data Processing Procedures

The raw data were recorded in a series of HDF5 files on hard drives mounted on the data storage system. These are shipped to NAOC for offline data processing. The offline data processing procedure is outlined in Fig. 5.

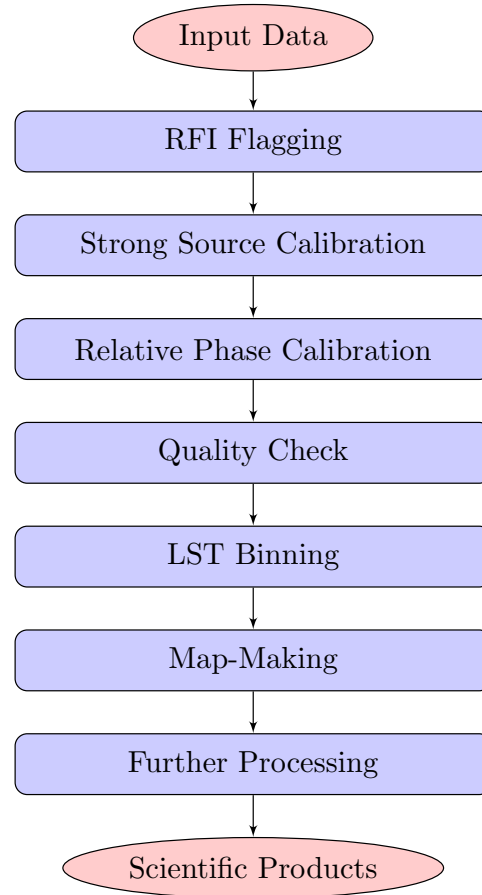
The first step of the data processing is radio frequency interference (RFI) flagging. The RFI is often stronger than the astronomical sources or noise, and have characteristic distributions in the time and frequency domain. These are recognized and recorded in a binary mask file and ignored in subsequent analysis. The Tianlai site is a very radio quiet site, but in some of the data set we found a significant amount of self-generated RFI, primarily from the power supplies for the array, and the motors of the nearby dish antennas. Later we installed quieter power supplies, and turned off the motors when making drift scan observations; greatly reducing the amount of RFI.

Two types of calibration are performed: (1) Absolute calibration. This is performed when a strong astronomical point source such as Cygnus A (Cyg A) or Cassiopeia A (Cas A) is transiting through the primary beam. The magnitude and phase of the complex gain of each element can be determined in such a calibration, up to a common offset. We use an eigenvector decomposition method to solve for these gains. (2) Relative calibration. This is performed by using the CNS signal to track the variation of the phase of complex gains when there are no strong astronomical sources available.

We can check the quality of the data before and after the calibrations. The calibrated data which pass the quality checks can be combined with other good data taken at the same local sidereal time (LST) but on different days. This co-added data will have an improved signal-to-noise ratio (SNR).

One can then use the LST binned data to produce sky maps and then make further analyses such as foreground removal and power spectrum estimations, which are the ultimate sci-

ence products for the 21 cm intensity mapping experiment. A python-based data processing pipeline has been developed [27], which provides a general framework for the data processing outlined above with at least one and sometimes several optional algorithms implemented for each step.



**Figure 5:** An outline of the offline data processing procedures.

In the present paper, we shall focus on examining the general quality of the raw data and investigating the precision and stability of the data achieved in the calibrations. The details of the RFI mitigation, and the methods and results of map-making will be described elsewhere.

## 3 Sub Systems Performance

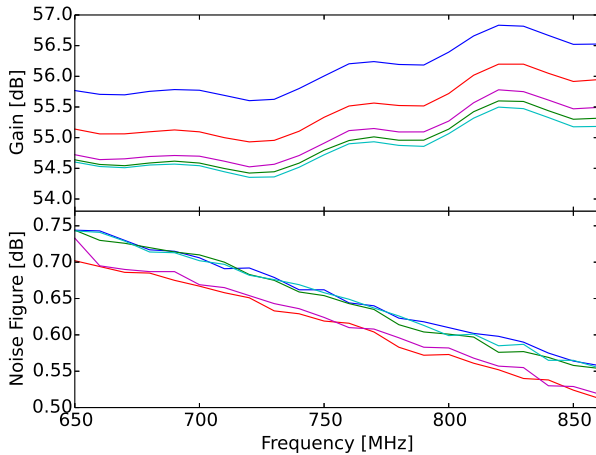
### 3.1 LNA

The receiver noise is determined to a large extent by the LNA. Because the operating frequency is relatively low, where the sky is relatively bright, the Tianlai receivers work at ambient temperature. Cooling would be expensive and not very effective. The gain ( $G$ ) and noise figure ( $NF$ ) of the LNAs are

measured at a frequency range from 650 MHz to 860 MHz, and a few examples are shown in Fig. 6. There are some variations in the value of the gain, but their frequency responses are similar. The NF variation in the 700–800 MHz band is about 0.1 dB, with a mean NF of 0.65 dB. The noise figure is related to the receiver noise temperature  $T$  by

$$T = \left(10^{\frac{NF}{10}} - 1\right) \times T_0 \quad (3)$$

where  $T_0$  is the environment temperature of the receiver. For example, if  $NF = 0.65$  dB, and we take the IEEE standard reference temperature  $T_0 = 290$  K, the corresponding noise temperature of the LNA is about 47 K.



**Figure 6:** The measured gain (top) and Noise Figure (bottom) of 5 randomly chosen LNAs.

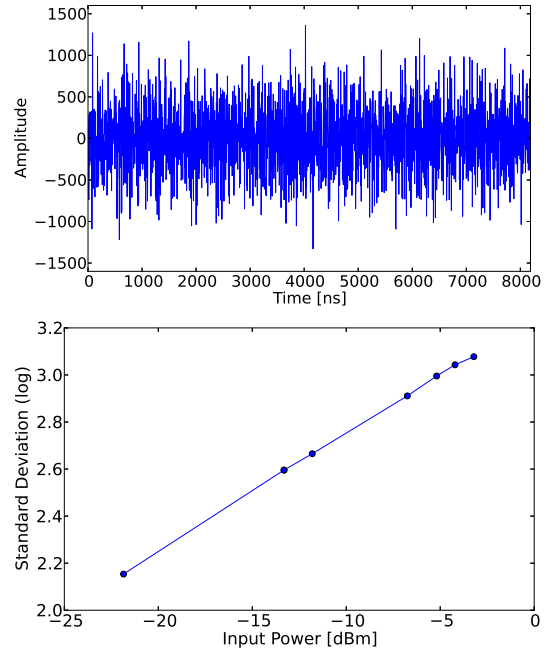
### 3.2 Linearity

Interferometry requires that the electronics respond linearly. However, real analog and digital systems are only linear in their response for a finite range of input power. The design of the RF front-end systems (LNA, optical transmission system) should ensure that it works in the linear regime during regular observations, and also produces amplified signals adapted to the digital system input requirements. We have estimated the required amount of amplification based on the typical radio frequency power level, the antenna gain, and the required input voltage for the ADC. To allow adjustment of power levels, two sets of 30 dB variable attenuators are placed before and after the mixer. We adjust them so that the IF output power level matches the requirement of the ADC.

The digital system also requires an appropriate level of input power to produce outputs which are linearly related with the input. The Tianlai cylinder array ADCs have a sampling length of 14 bits, i.e. the output is an integer in the range of -8192 to +8191 decimal. Better precision can be obtained if the digital outputs have more non-zero bits, but on the other

hand it is desirable not to saturate the output when the signal is very strong.

A typical sample of the ADC output is shown in the top panel of Fig. 7, and the standard deviation (STD) of the output is shown as a function of the input power level in the bottom panel. The IF output power level is adjusted to about -13 dBmW when the array is observing a part of the sky with no bright sources during night time, so that the ADC output has a standard deviation of 400, about 5% of the total range allowed by the sampling length, which is a good level for quantization and avoids saturation. The brightest radio astronomical source is the Sun. When it is transiting through the telescope FoV, the maximum power level is about 5 dB stronger, still well within the range of the ADC. Occasionally RFI may generate strong signals that saturate the correlator, but as we are generally not so interested in RFI and they are generally filtered out, non-linear response due to saturation by RFI is not a big problem for data processing.

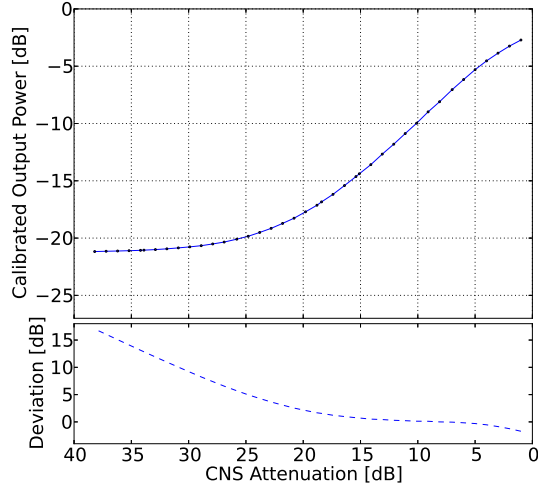


**Figure 7:** Top: ADC raw data samples taken over 8192 ns when the input power level is -13.31 dBmW. Bottom: the standard deviation in logarithm of ADC output at different input power levels. Currently, the IF level we set is about -13 dBm within our observation band.

We use the CNS to check the signal chain response. The CNS generates a wide band noise signal in the 700–800 MHz frequency band and its power level is adjusted by an attenuator. The auto-correlation signal is shown in Fig.8 as a function of the attenuation in CNS power. The auto-correlations have a non-zero noise floor at about -20 dB attenuation, due to typical sky signal and receiver noise; this is however not an indication of break down of linearity. The cross-correlation



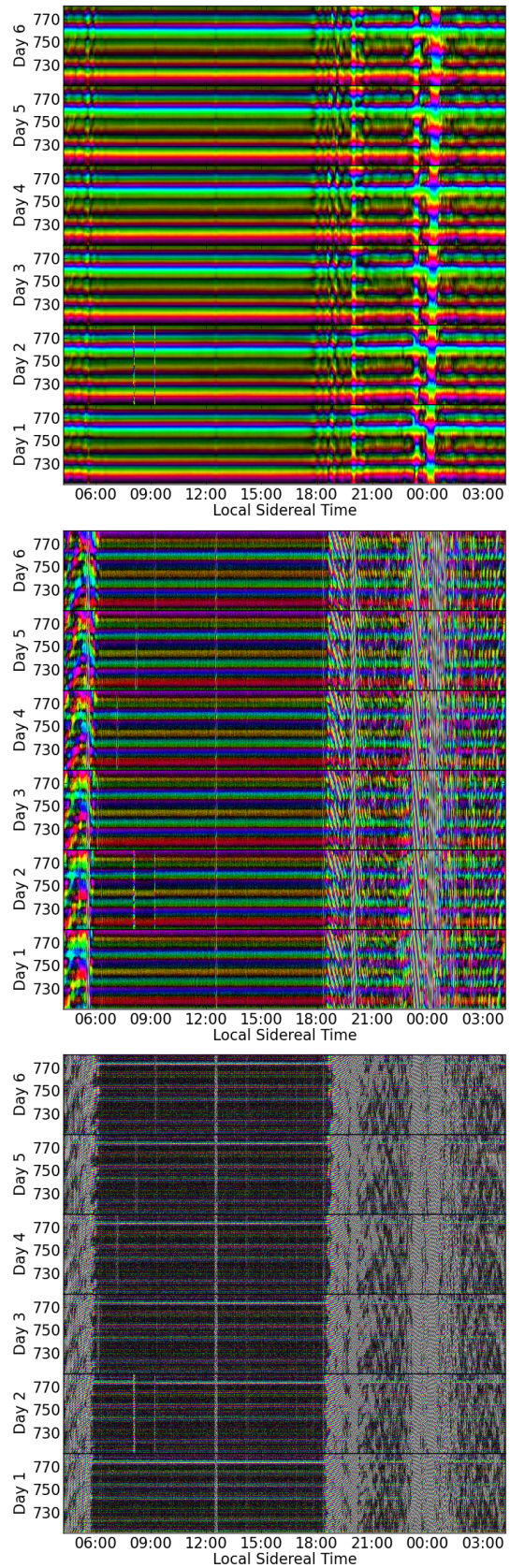
power which has a zero noise floor is linear all the way down to an attenuation of 40 dB. This test is limited by the sky signal which is always present and contributes non-zero cross-correlations.



**Figure 8:** The auto-correlation output as a function of the power level of the CNS (top panel) and nonlinear residual (bottom panel).

### 4 Quick Look Analysis of the Visibilities

We first make a quick look analysis of the raw visibilities. As an example, in Fig. 9 the raw visibility of three baselines are plotted as a function of LST and frequency. From top to bottom, the three subplots show respectively a baseline for two feeds on the same-cylinder, on two adjacent cylinders, and on two non-adjacent cylinders. Each subplot shows the result of six consecutive days starting from 2018/03/22; each day as a sub-panel from bottom to top. The data are rebinned to 488 kHz (4 original frequency bins) and 20 seconds (5 original time bins) to suppress noise. We have removed the data affected by the periodic broadcasting of the CNS. In this two dimensional plot, the brightness is proportional to the amplitude of the visibility, while the hue is used to represent the phase, from red, through blue, to violet for phase angle varying from 0 to  $2\pi$ . The saturation is set as 1. We set the brightness to saturate at 2 times the median amplitude, to allow the display of fainter features.



**Figure 9:** The raw visibilities as a function of LST and frequency for 6 days starting from 2018/03/22. The data are rebinned to 488 kHz and 20 s resolutions. Top panel: baseline A3Y-A15Y; Middle panel: baseline A3Y-B18Y; Bottom panel: baseline A2Y-C2Y.

All panels show horizontal stripes that are fairly stable over time. This “background” has some structure in frequency and is stable with time. This is correlated noise arising from the coupling between feeds. Indeed, it is particularly strong for the top panel, which shows the visibility of a medium length baseline on the same cylinder. The feeds on the same cylinder are relatively close to each other, and radiation from one can be reflected by the cylinder to the other feed. It is much weaker for feeds on adjacent cylinders (middle panel), and even weaker for the feeds on non-adjacent cylinders (bottom panel).

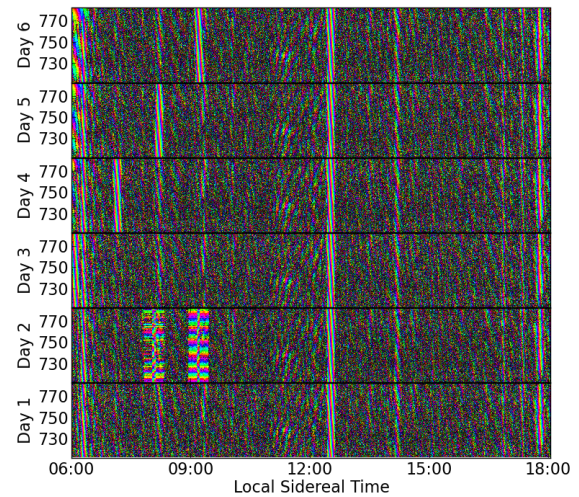
**Table 3:** List of sources that can be recognized in Fig. 10 and Fig. 11, including their right ascension (R.A.), Zenith Angle (ZA), and flux at 750 MHz as given in <https://ned.ipac.caltech.edu/>. The list is in the order of R.A.

Source	R.A.	ZA (deg)	Flux (Jy)
3C 010	00:25	20.0	62
3C 058	02:05	20.7	34
IC 1805	02:32	17.4	–
3C 084	03:20	2.7	22
3C 123	04:37	14.5	76
M 42	05:35	49.5	–
IC 443	6:16	21.6	190
3C 196	8:13	4.1	23
Hydra A	9:18	56.2	81
M 82	9:55	25.5	11
M 87	12:30	31.8	353
3C 286	13:31	13.6	19
3C 295	14:11	8.1	37
Hercules A	16:51	39.2	88
3C 353	17:20	45.1	88
Galactic Center	17:45	73.0	–
3C 380	18:29	4.6	23
3C 392	18:56	42.8	242
3C 400	19:23	30.0	673
Cyg A	19:59	3.4	2980
Cyg X	20:28	41.2	–
NRAO 650	21:12	8.3	48
3C 433	21:24	19.1	21
Cas A	23:23	14.7	2861

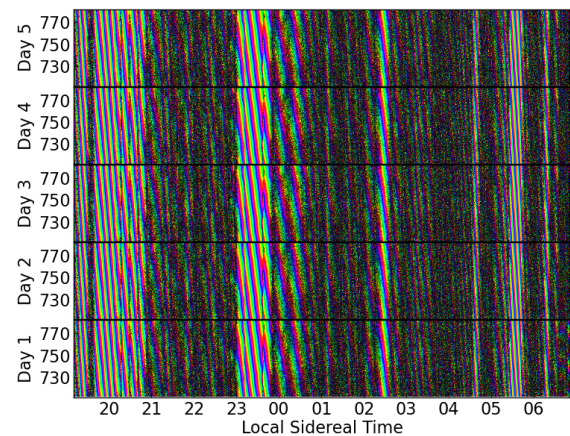
The vertical features in the figure shows strong sources transiting through the the telescope field of view. This observation was taken near the Vernal Equinox; the Right Ascension (R.A.) of the Sun is nearly 0 h, so its transit (noon)

takes place near 0:00 LST. With a few exceptions, the visibility is very similar on each day. The most prominent feature in these plot is the division of day and night. Sunrise (18h30m LST) and sunset (6h00m LST) can be clearly seen. The strong fringes of the Sun shifted slightly each day, due to the daily motion of the Sun along the ecliptic.

Besides the Sun, a few other transits can also be recognized. There is a strong fringe at 12h30m produced by the transit of Virgo A (M87). Another obvious source which has its transit on 6h8m (Day 3), then 7h10m (Day 4), 8h12m (Day 5) and 9h14m (Day 6) is the Moon, which moves 1h02m in R.A. each day during this observation. The two bright vertical fringes on day 2 at about 8h10m and 9h15m are from RFI.

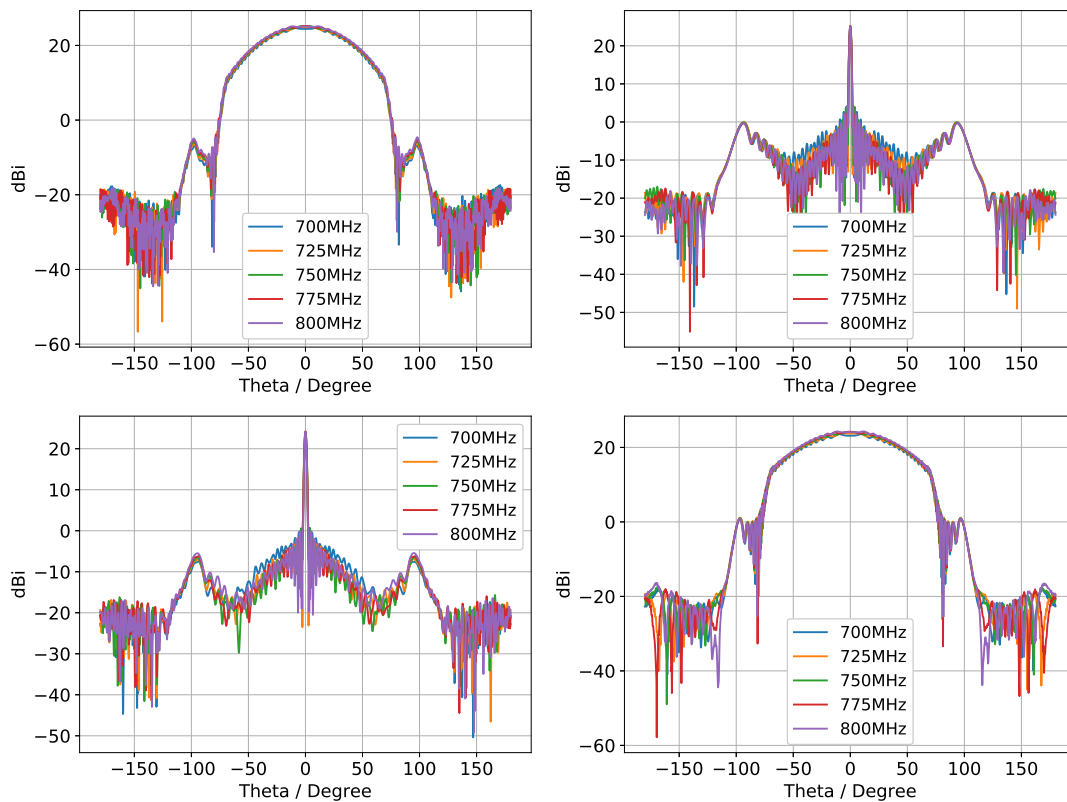


**Figure 10:** The nighttime A3Y-B18Y visibility of 2018/03/22 data, with the time-averaged band structure subtracted.



**Figure 11:** The nighttime A1Y-B2Y visibility of 2016/09/27 data, with the time-averaged band structure subtracted.

The horizontal stripes are very stable, and are removed by subtracting off a moving average of 1 hour. As an example,



**Figure 12:** The simulated beam profiles at 700–800 MHz. Left: E-plane; Right: H-plane. Top: Y-polarization. Bottom: X-polarization. The directivity is 25.2 dBi for the X-polarization and 24.2 dBi for the Y-polarization.

we show the processed data during night time for the baseline A3Y-B18Y in Fig. 10 and A1Y-B2Y in Fig. 11.

Many less prominent sources can now be clearly seen, and they repeat each sidereal day. We list the sources that can be seen directly in this visibility and identified with well known strong radio sources in Table 3, where both day and night time sources are included. Some sources are complex ones, which would be resolved into multiple sources with higher angular resolution. The flux for the entire complex is not available. In fact, more fainter fringes could be spotted in the visibility, but are too numerous for all of them to be listed.

For the longer E-W baseline A2Y-C2Y, the angular resolution is higher and adjacent sources unresolved by the shorter baselines may be resolved by the longer ones. The fringes are clearer and the phase varies more rapidly. The horizontal band structure is less significant in the longer baselines, perhaps because the larger physical separation of the feeds of longer baselines leads to smaller cross-coupling and weaker correlated noise.

In Fig. 11, we plot the visibility for A1Y-B2Y for a different observation starting from 2016/09/27. Here the Sun is positioned on the opposite part of the ecliptic. There happen to be many more bright radio sources during the night sky. We can see numerous fringes in this case, even without

removing the Sun.

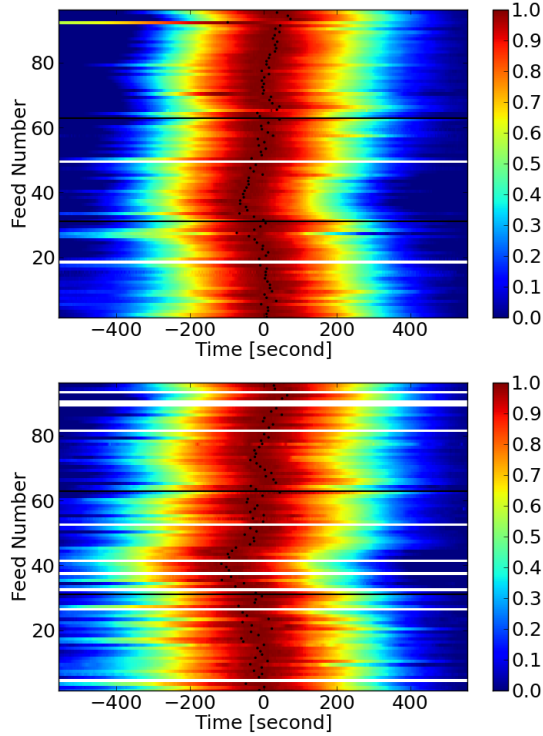
## 5 The Antenna Beam Response Pattern

The beam profile for a feed on the Tianlai array is simulated using the Microwave CST software. The simulated beam profiles at 650 MHz, 1050 MHz and 1420 MHz were reported earlier in [25], which demonstrated that the Tianlai cylinder antenna and feed have a quite good broadband response. Here, we show the simulated beam profiles in the 700–800 MHz range in Fig. 12. As expected, the beam has a narrow peak with a full width half maximum (FWHM) of  $1.6^\circ(1.8^\circ)$  for X (Y) polarization at 750 MHz along the E-W direction (H(E)-plane for the X(Y) polarization), and a very broad peak with FWHM about  $62.2^\circ(71.4^\circ)$  along the N-S direction (E(H)-plane for the X(Y) polarization). The broad beam along the N-S direction is largely determined by the beam profile of the feed.

We can verify the simulated beam profile in the E-W direction by observing the transit of strong sources. We use Cyg A as our calibrator. The declination of Cyg A is  $+40.7^\circ$  and the zenith angle is only  $3.4^\circ$  at the peak of this transit. Because its track is quite close to the E-W great circle passing through the zenith, it provides a good determination of the E-W beam

profile near the center of the beam.

The auto-correlations of the cylinder array units during the peak of the Cyg A transit event are plotted in Fig. 13 as a function of time. Here the background has been removed, so that the auto-correlation reflects what is induced by Cyg A. The amplitude is averaged over the full frequency band to improve the SNR. These curves are normalized at their peak values and arranged according to their relative positions on the cylinders. The feed numbers given here are 1 ~ 31 for cylinder A, 32 ~ 63 for cylinder B, and 64 ~ 96 for cylinder C.

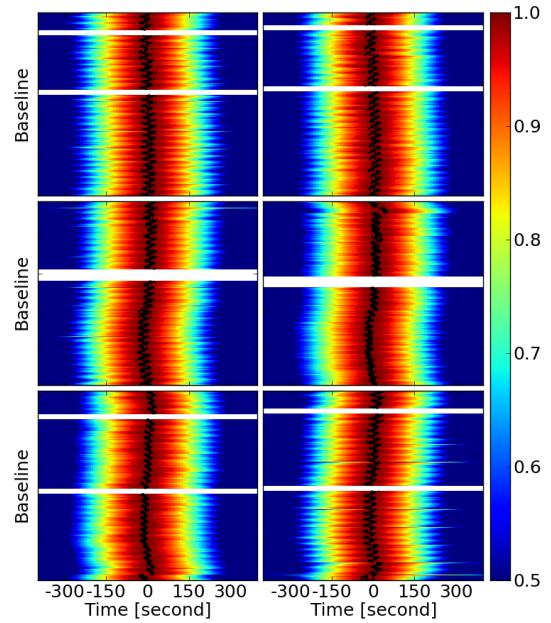


**Figure 13:** The auto-correlations of all X (top) and Y (bottom) polarizations vs time during Cyg A transit. The black dots mark the peak time of the transit curve fitted by a Gaussian function. Malfunctioning feeds are masked and appear as horizontal white lines.

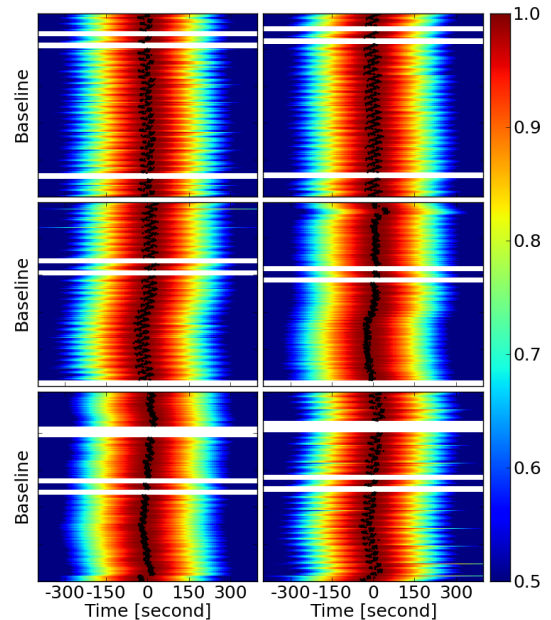
The alignment of different feeds inevitably has some errors, so the timing of the signal peak during the Cyg A transit is different for each feed. The standard deviation of the transit times are  $\sigma_Y = 38.8\text{s}$  and  $\sigma_X = 31.1\text{s}$ , corresponding to  $0.123^\circ$  and  $0.099^\circ$  in angle respectively. These errors in angle are reasonable values for an installation with simple mechanical tools and adjustment. The difference in the mean time between the two polarizations is 18.6 s (or  $0.059^\circ$  in angle), which is very small.

While the auto-correlation gives directly the response for each feed, it has large receiver noise and is sensitive to the contribution of mean sky brightness. Therefore, we also study the pointing of the beams using cross-correlations.

We derive an average beam profile by averaging the cross-correlation amplitudes for all of the E-W baselines which span different cylinders during the transit of Cyg A. This is averaged over all frequencies.



**Figure 14:** Distribution of the transit amplitude curves for the X-polarization derived from averaging the cross-correlations of all of the E-W baselines. Top panels:  $A_B$  (left) and  $A_C$  (right); Center panels:  $B_C$ ,  $B_A$ ; Bottom panels:  $C_A$ ,  $C_B$ . The black dots show the transit peaks fitted by Gaussian function.



**Figure 15:** Same as Fig. 14 but for Y polarization.

For example, the E-W profile for the X-polarization of the  $i$ th feed on cylinder A can be obtained by averaging over its

cross-correlations with the B cylinder as

$$P_i^X(A_B) = \sum_v \sum_j V_{(A_i),(B_j)}^{XX}(v). \quad (4)$$

We collect such profiles for cylinder A and denote it as  $A_B$ . In Fig. 14 and Fig. 15, we plot these amplitudes.

The general shape of the profiles for the same cylinder obtained with different cross-correlation cylinders (e.g.  $A_B$  and  $A_C$ ) are quite similar with each other, giving confidence in this method. Also, the X- and Y-polarizations for the same cylinder are very similar to each other, suggesting that the deviations are due primarily to the misalignment of the installation angle for the feeds. However, we note that this misalignment does introduce a small pointing error in the data, which should be accounted for to estimate the precision of the observation.

The beam center of each feed can be estimated from the center or peak position of the cross-correlations. Assuming the beam center to be  $a_l$  for feed  $A_l$  and  $b_m$  for feed  $B_m$ . The cross-correlation beam center will be  $\frac{1}{2}(a_l + b_m)$ . The centers of the cross-correlation beams are shown as black dots in Fig. 14 and Fig. 15. The minimum variance solution for the beam center of the feeds is then given by (see the Appendix for derivation)

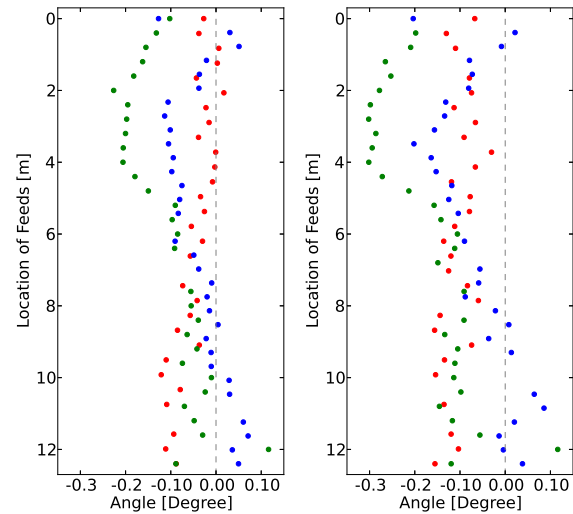
$$a_l = \frac{1}{M} \sum_{m=1}^M \theta_{l,m}^{AB} + \frac{1}{N} \sum_{n=1}^N \phi_{l,n}^{AC} - \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \gamma_{m,n}^{BC} \quad (5)$$

$$b_m = \frac{1}{N} \sum_{n=1}^N \theta_{m,n}^{BC} + \frac{1}{L} \sum_{l=1}^L \phi_{m,l}^{BA} - \frac{1}{NL} \sum_{n=1}^N \sum_{l=1}^L \gamma_{n,l}^{CA} \quad (6)$$

$$c_n = \frac{1}{L} \sum_{l=1}^L \theta_{n,l}^{CA} + \frac{1}{M} \sum_{m=1}^M \phi_{n,m}^{CB} - \frac{1}{LM} \sum_{l=1}^L \sum_{m=1}^M \gamma_{l,m}^{AB} \quad (7)$$

where  $a_l, b_m, c_n$  are the solved beam pointing for feed  $A_l, B_m, C_n$  in cylinder A, B, C.  $L = 31, M = 32, N = 33$  are the feed numbers.  $\theta, \phi, \gamma$  are the observed cross-correlation beam center. The distribution of the solved beam centers is drawn in Fig. 16.

Finally, in Fig. 17, we show 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 Cyg A on 7 successive days. For each transit and each visibility we fit a Gaussian shape to the magnitude of the visibility as a function of time. The measured pattern is effectively the geometric mean of the patterns of two dishes, which are nominally coaligned. Day-to-day fluctuations of the FWHM are less than 3%. We also plot the case of a diffraction - limited circular aperture ( $1.028\lambda/D_{\text{eff}}$  with  $D_{\text{eff}} = 0.9D$ ). Compared with the ideal case, there are apparent sinusoidal variations in the beam size. A similar phenomenon was also found to occur on the CHIME cylinders [28].



**Figure 16:** The solved antenna pointings of the X (left) and Y (right) polarizations. The pointings are extracted from the cross-correlations with the baselines in E-W directions. The baselines in N-S direction are not used due to low SNR of Cyg A in the visibility amplitude.

From the beam profile, we can derive the directivity and effective area of the antenna. The directivity is given by

$$D = \frac{P_{\max}(\theta, \phi)}{P_{\text{av}}}, \quad (8)$$

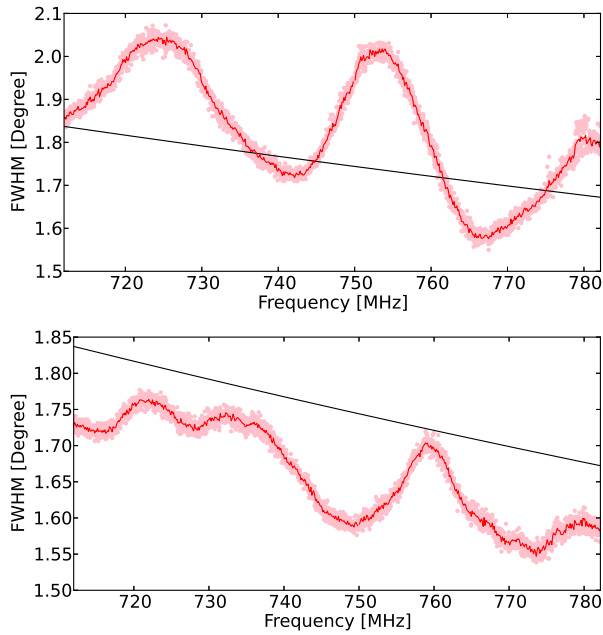
where  $P_{\max}(\theta, \phi)$  is the maximum power of the antenna, while the average power is

$$P_{\text{av}} = \frac{1}{4\pi} \int P(\theta, \phi) d\Omega \quad (9)$$

so that  $D = 4\pi/\Omega_A$ . The effective area of the antenna is

$$A_e = \frac{D\lambda^2}{4\pi}. \quad (10)$$

In our case, for a single feed on the cylinder, the simulated directivity is about 25.2 dBi for X polarization and 24.2 dBi for Y-polarization, so that,  $A_e^X = 4.22 \text{ m}^2$ , and  $A_e^Y = 3.35 \text{ m}^2$  at 750 MHz. These relatively small effective areas are due to the fact that despite the large total area of the cylinder, for observing one particular direction, each feed receives only the wave reflected from a narrow strip of reflecting surface.



**Figure 17:** The mean FWHM of the primary beam in the E-W direction as a function of frequency using daily transits of Cyg A over 7 days. The measurement is for the fairly typical baselines A14Y-C18Y and A2X-C17X. The dots are the fitted beam width at each frequency channel for all days, and the red line is the mean. The black line shows the FWHM of a uniformly-illuminated Airy disk with an effective diameter equal to 90% of the cylinder width ( $0.9 \times 15\text{m}$ ).

## 6 Calibration

Accurate calibration is essential for determining the complex gain (particularly the phase) of each array element of an interferometer array. Below, we first calibrate the frequency response (bandpass), and then calibrate the complex gains of different feed channels using the transit of a strong astronomical point source such as Cyg A (referred as absolute calibration in this paper). Finally, when a strong astronomical point source is not available, calibration can be made with the help of the artificial CNS (referred to as relative calibration in this paper). The calibration algorithm is described in Zuo2019.

### 6.1 The Bandpass

We calibrate the bandpass response by analyzing the transit of strong sources. The bright source Cyg A has a flux of 2961 Jy at 750 MHz [29], and in the frequency range of interest,

$$\log S = \sum_{n=0}^5 a_n [\log(\nu)]^n \quad (11)$$

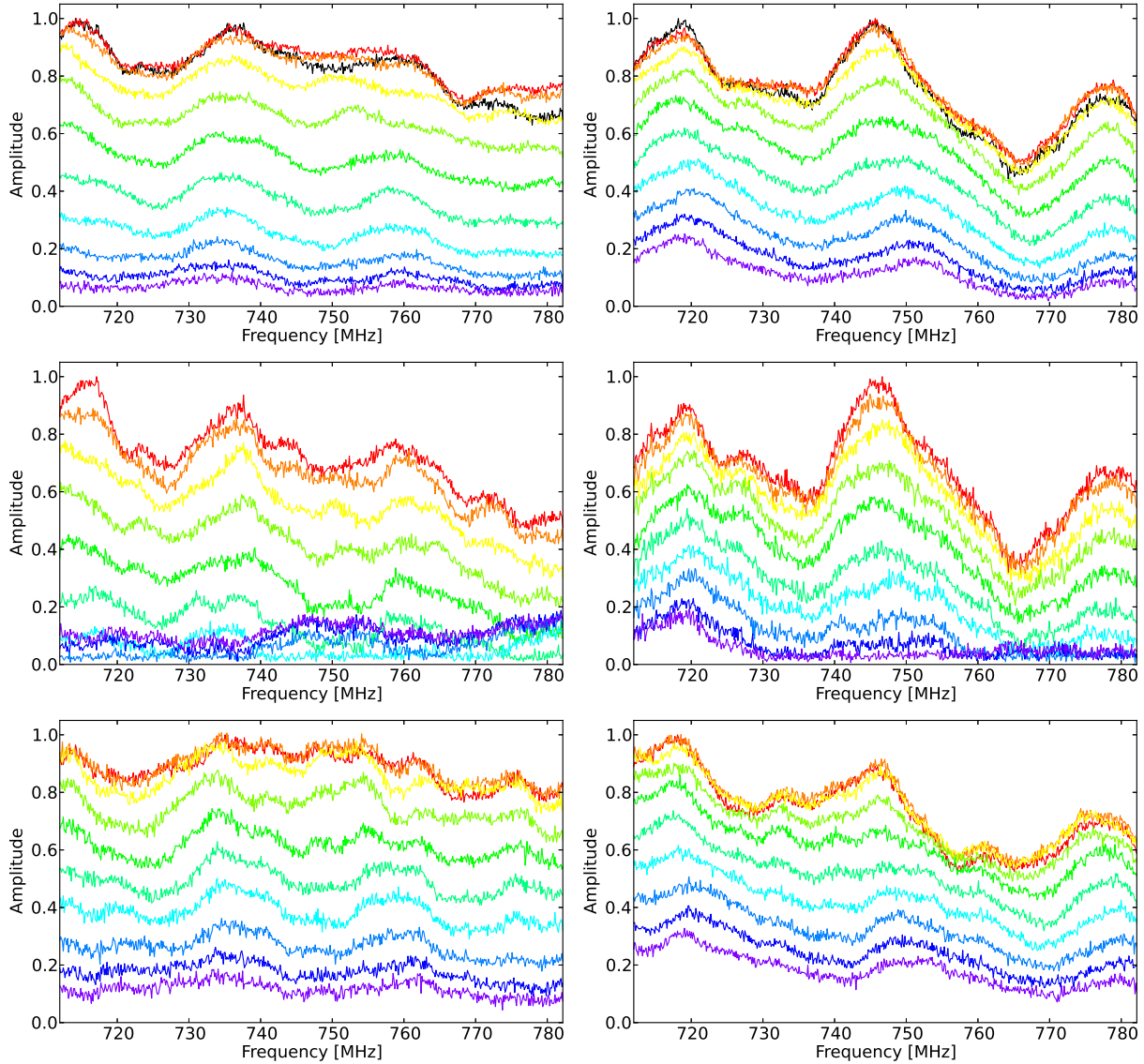
where  $a_0 = 3.3498$ ,  $a_1 = -1.0022$ ,  $a_2 = -0.2246$ ,  $a_3 = 0.0227$ ,  $a_4 = 0.0425$ . In our observation band (700–800 MHz) the flux difference between the low frequency side and high frequency side is about 5%.

We first check whether it is a good assumption that Cyg A dominates the antenna temperature. In Fig. 18, we plot the auto-correlations for the B9 and A4 feeds, as well as the cross-correlations between them, during the different stages of Cyg A transit. There is a non-zero background level for the auto-correlations, while for cross-correlations this is expected to be zero (c.f. Eq.2). We have therefore subtracted the background for the auto-correlations. As we come closer to the peak of the transit, the amplitude of the visibility increases, and the shape of the bandpass becomes more similar to each other with better SNRs. The bandpass (as derived from the auto-correlation) for different feeds are not the same, as illustrated in Fig. 18, where the shape of the bandpass for B9 and A4 are different. However, the shape of the cross-correlation is generally consistent with the square root of the product of the bandpasses obtained from the two auto-correlations at the peak of transit, as shown in the top panels by the black curve. From this trial, we see that Cyg A indeed dominates the visibility at the peak of its transit, and can be used to calibrate the bandpass.

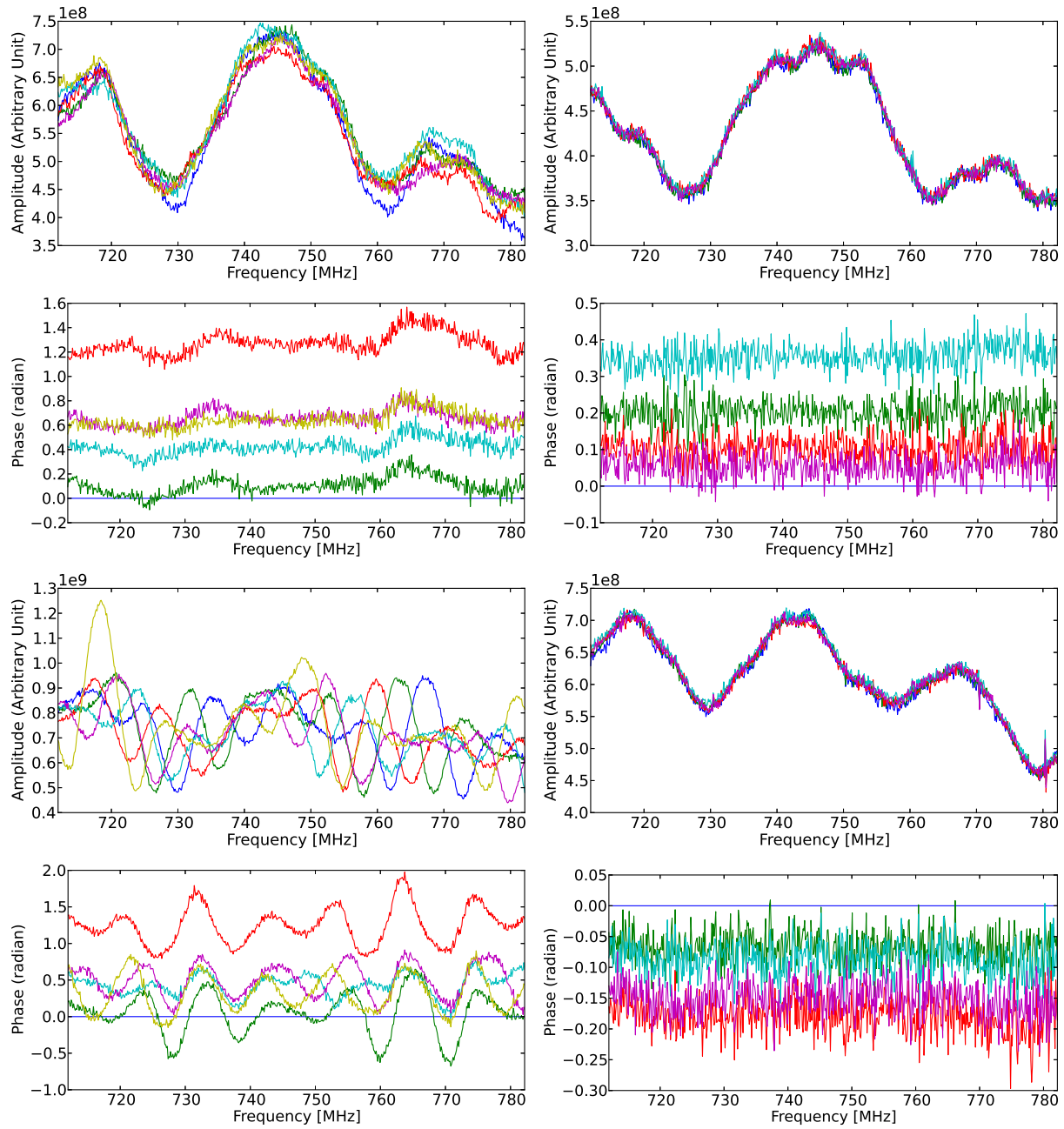
In Fig. 19 we show the amplitudes and phases of the visibilities as a function of frequency at the peak of the Cyg A transit from the observational data set 2018/03/22 (left column) and 2016/09/27 (right column). The cross-correlations of Y-polarization are plotted in the top two rows, and X-polarizations in the bottom two rows. In each plot, the visibilities are taken from the peak of the transit. Each colored curve corresponds to one day's transit; there are 6 days for the 2018/03/22 data set and 5 days for the 2016/09/27 data.

The amplitude of the visibility as a function of frequency reflects the bandpass response of the system. As shown in the figure, the amplitudes are fairly stable and consistent for the Y-polarization on each day; the day-to-day variation is very small. However, for the X-polarization of the 2018/03/22 data set (shown the in left column of Fig. 19), the amplitudes are different on each day. After some study, we conclude that this is caused by contamination by the Sun, as the Cyg A transits occurred at about 01:30 UTC (about 07:30 local civil time) on these days, and the radiation from the Sun, which is above the horizon in the East direction, may enter the beam from the side lobes, because the feeds are directly illuminated. For the Y-polarization, which is oriented in the E-W direction, the projected area is at a minimum towards the East, so it is much less affected than the X-polarization, whose projected area is maximum in the E-W directions. In the 2016/09/27 data set (results shown on the right column) we see the amplitudes have much less variation as the Cas A transit occurred at about 13:40 UTC (about 19:40 local civil time), after sunset.

The phase of the visibility induced by a strong point source



**Figure 18:** Frequency response as obtained from Cyg A transit, for a cross correlation visibility A4-B9 (top) and two auto-correlations, A4-A4(middle) and B9-B9 (bottom). The left column shows XX polarization and right column shows YY polarization. The colors from purple to red indicate the different times during the transits process. The black line on the top panel correspond to the square root of the product of the two auto-correlation frequency responses at the peak of the transit.



**Figure 19:** The uncalibrated visibility amplitude and residual phase w.r.t. the first day as a function of frequency at the peak of the Cyg A transits. Left Column: from the data set of 2018/03/22. Top panel: A2Y-B27Y, Bottom panel: A13X-B31X. Right Column: from the data set of 2016/09/27. Top panel: A2Y-B3Y, Bottom panel: A6X-B13X. Each colored curve represents the response for the transit on a different day.



is given by  $-2\pi\nu\vec{n}\cdot\vec{b}_{ij} + \varphi_{ij}(\nu)$ , where  $\vec{n}$  is the direction of the source, and  $\vec{b}_{ij}$  is the baseline. After subtracting out the term proportional to the frequency  $\nu$ , the residual reflects the instrument phase  $\varphi(\nu)$ . We have compared the phase with the first day's phase. As expected, the phase drifted away from this value on subsequent days, but during each day the instrument phase is nearly constant over frequency; the fluctuations are due to noise. Remarkably, the phase frequency structures are quite similar to each other on different days. Again, the X-polarization for the 2018/03/22 observation is an exception, where the influence of the Sun can not be neglected.

## 6.2 Absolute Calibration

When a strong point source is transiting through the primary beam of the telescope and dominating the visibility, the visibility for a pair of elements  $i$  and  $j$  is given by

$$V_{ij}^0 = S_c G_i G_j^*, \quad (12)$$

where  $S_c$  is the flux of the source,

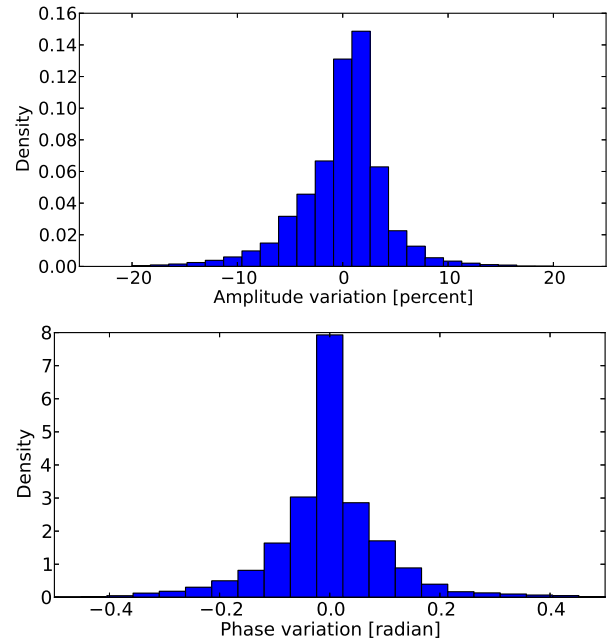
$$G_i = g_i A_i(\hat{n}_0) e^{-i2\pi\hat{n}_0\cdot\vec{u}_i}; \quad (13)$$

where  $A_i$  is the beam pattern of array element  $i$ ,  $\hat{n}_0$  is the unit vector in the direction of the source and  $\vec{u}_i$  is the coordinate of array element  $i$  in wavelength units. In matrix form,

$$\mathbf{V}_0 = S_c \mathbf{G}\mathbf{G}^\dagger. \quad (14)$$

$\mathbf{G}$  is an eigenvector of  $\mathbf{V}_0$ . If noise is present but small compared with the calibrator source and statistically equal in all elements, i.e.  $\mathbf{V} = \mathbf{V}_0 + \mathbf{N}$ , where  $\mathbf{N}$  is the noise co-variance matrix, the vector  $\mathbf{G}$  could be obtained by principal component analysis (PCA): solving the eigenvector of matrix  $\mathbf{V}$ , with the eigenvector associated with the largest eigenvalue identified as  $\mathbf{G}$ . This is also the least square solution of the form  $\mathbf{V} = \mathbf{g}\mathbf{g}^\dagger$  (for proof and more details see Zuo2019). Thus, for each time and frequency point, a solution of the complex gain can be obtained using this eigenvector decomposition method. This method of calibration is however only applicable when a strong point source is transiting, which restricts its usage.

In Fig. 20, we plot the variation of the gains during 5 days with respect to the first day, as obtained from the absolute calibration with Cyg A transits. The top panel shows the variation in amplitude as a percentage, and we find that most of the variations are within  $\pm 5\%$ . The phase variation are mostly within 0.2 radians. Part of these variations are true variations in the gain of the receivers, while part may be due to the error in the calibration process, which has a limited precision. At present, these two can not be distinguished.



**Figure 20:** The distribution of receiver gains is shown as a percentage variation in amplitude (top) and radian variation in phase (bottom), for all polarization channels and frequencies at the peak of Cyg A transits over five days.

## 6.3 Relative Calibration

The CNS which broadcasts periodically can be used for relative calibration at any time. The CNS is located near the array and the contribution to the visibility caused by its radiation is

$$V_{ij}^{\text{CNS}} \propto S_n e^{-i2\pi(\vec{n}_i\cdot\vec{r}_i - \vec{n}_j\cdot\vec{r}_j)} \quad (15)$$

where  $\vec{r}_i$  and  $\vec{r}_j$  are the locations of feeds  $i$  and  $j$  respectively, and  $\hat{n}_i$  and  $\hat{n}_j$  are the displacement vectors from the noise source to the positions  $\vec{r}_i$  and  $\vec{r}_j$ . The noise source induced visibility can be obtained by subtracting the noise on and noise off values:

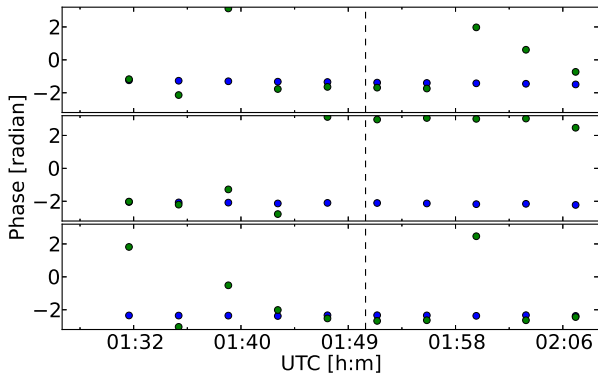
$$V_{ij}^{\text{on}} - V_{ij}^{\text{off}} \approx C |G_{ij}| e^{-i2\pi\nu\Delta\tau_{ij}} e^{-i2\pi(r_i - r_j)/\lambda}. \quad (16)$$

Here,  $C$  is a constant and  $\Delta\tau_{ij}$  is the equivalent instrument delay difference between the channels  $i$  and  $j$ , which are mostly due to the variation in the cable length for the two channels. We see that unlike the absolute calibration with the astronomical source, in the relative calibration only the variation of the difference between the phase of a pair of signal channels is directly determined.

We have used two approaches to deal with this problem: calibration based on individual feed channels, and calibration based on feed pairs.

**Individual Feed Channels.** In the approach based on individual feed channels, one tries to obtain the complex gains

for all the feed channels by making a best fit solution to all baselines. As there are many more baselines than the number of array elements, it is possible to obtain a best fit solution to the phase of the individual elements. However, note that there is an inherent degeneracy: if we add a time-dependent but identical instrument phase to all elements, there is no observable effect. This problem can be solved by imposing a physically motivated “gauge condition”: assume that as  $t \rightarrow 0$ , the changes of phase for all elements should go to zero. So we adopt the condition that the total phase variation for all elements are to be minimized at short time intervals. In this way, a solution can be obtained for each feed input channel.



**Figure 21:** A comparison of the phase of the gain calibrated by the CNS (blue) and Cyg A (green). From top to bottom, the feeds are A20X, B22X, C23X. The vertical dashed line indicates the transit time of Cyg A.

In Fig.21 we compare the phases obtained by relative calibration and absolute calibration during the transit of Cyg A. The phases obtained from the relative calibration stay nearly constant. The phases obtained by the absolute calibration change rapidly as Cyg A comes to the center of the beam. This variation is not a real variation in phase, but simply an error in the absolute calibration: the absolute calibration is only viable when the point source is dominating, and when the source is outside the beam center the result is not reliable. However, once in the center, the absolute calibration also yields nearly constant phases. The absolute and relative calibrations agree with each other for the duration of the peak of transit.

**Feed Pairs.** In the approach based on feed pairs, instead of trying to solve for the complex gain  $g_i$  of each channel  $i$ , we simply calibrate each pair of channels. In principle, the “gain” for pair  $i, j$  should simply be  $g_i g_j^*$ , but as discussed above, this can not be obtained directly and has to be solved in the relative calibration. Instead, one may define a complex gain associated with each pair of input channels,

$$G_{ij} = |G_{ij}|e^{-i2\pi\nu\Delta\tau_{ij}} \quad (17)$$

which can be directly inferred from the visibility of the CNS.

Here the change of instrumental phase for the visibility of each pair of input channels can be obtained directly by monitoring the CNS signal, without using the complex gain of individual input channels. The computation is also much simpler in this approach.

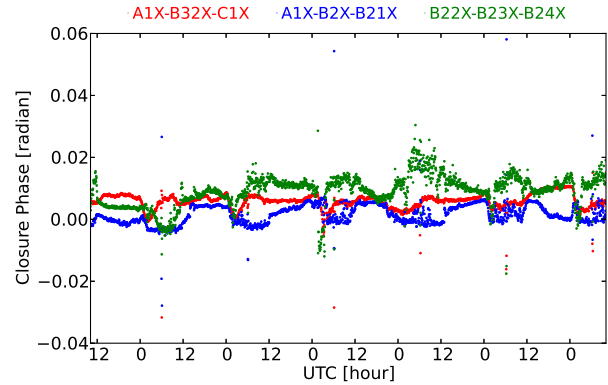
To check the precision of this approach, we can employ the closure phase relation. If a strong radio point source or the CNS at position  $\hat{n}$  with flux  $S_0$  dominates the received signal, the visibility is given by

$$V_{ij} = g_i^* g_j A_i^*(\hat{n}) A_j(\hat{n}) S_0 e^{-i2\pi\hat{u}_{ij}\hat{n}}. \quad (18)$$

We see that the quantity  $V_{ij}V_{jk}V_{ki}$  should be a real number. Thus, if we write the visibility in the form  $V_{ij} = |V_{ij}|e^{i\phi_{ij}}$ , then we shall have

$$\phi_{ij} + \phi_{jk} + \phi_{ki} = 2N\pi, \quad (19)$$

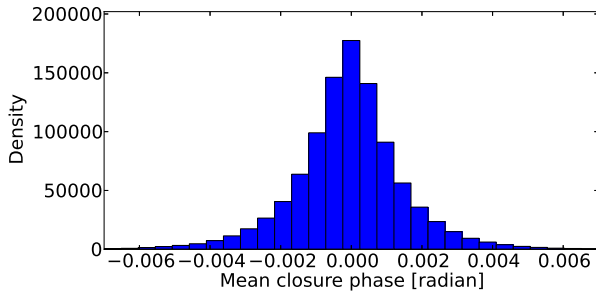
where  $N$  is an integer. In the relative calibration based on feed pairs, each  $V_{ij}$  is individually calibrated, so the closure phase relation may break if the calibration fails. We can use this closure phase relation to check the validity of relative calibration.



**Figure 22:** The closure phase for CNS induced visibility of three pairs of channels. red: A1X-B32X-C1X; blue: A1X-B2X-B21X; green: B22X-B23X-B24X.

In Fig.22 we plot the closure phase for a few feed channel pairs using the visibilities induced by the CNS over 5 consecutive days. Here we consider three cases: (1) three feeds are all located on different cylinders (red curve); (2) two of the three feeds are located on the same cylinder, but are some distance away from each other (blue curve); (3) the three feeds are adjacent ones on the same cylinder (green curve). We see the deviations from zero are all relatively small ( $< 0.02$  rad) for most of the time, but become larger around the local noon time when the Sun is transiting through the beam, during which the single dominant source approximation for the CNS is slightly broken.

In Fig. 23, we plot the distribution of the closure phase of feed pair gains  $G_{ij}$  calibrated with the CNS for all combination of baselines and 100 frequencies. This shows that the closure phase is centered around 0, with small deviations.  $\sigma_\phi = 0.0015$  rad, which provides an estimate on the achievable precision of the calibration of feed pair gains.



**Figure 23:** The distribution of closure phase of baseline gains  $G_{ij}$  for all combinations of baselines and 100 frequencies of the calibrated data set 2016/09/27.

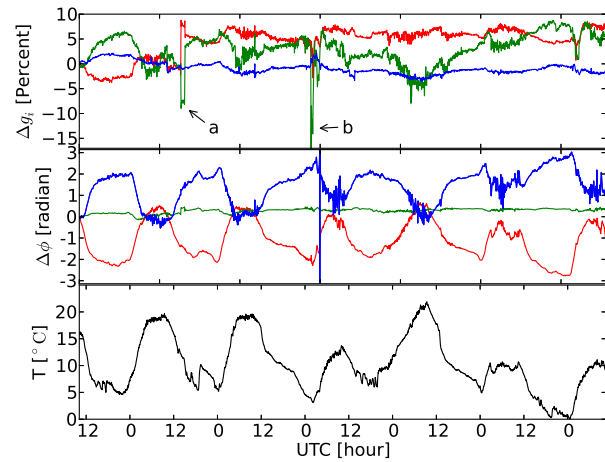
#### 6.4 Stability

Using the relative calibration method, the variation of complex gains throughout the day can be tracked. In Fig. 24, we show the gain of a few typical individual feed channels during 6 consecutive days, with the top panel showing the change in amplitude as a percentage and the middle panel showing phase. For comparison, the ambient environmental temperature variation during this measurement is plotted in the bottom panel.

The variation of the gain seems correlated with that of temperature, especially for the phase. For different polarization channels, the variations of phase are quite diverse, though appear to be well-correlated with the temperature. For example, the phase of A28Y (green curve) varies very little, while for A15X (red curve) and C22Y (blue curve) the phase variation is much larger and in the opposite sense. These variations are probably induced by the thermal expansion and contraction of the optical fiber length. As the optical cables are about 8 km long, the change in their length could be significant. The optical fibers within the cable are in different positions and have slightly different temperatures, so it is also not too surprising that each of them undergoes different amounts of expansion or contraction, and some even stay constant. To verify that the majority of the instrument phase change comes from the optical fibers, we measured the round trip delay of the optical cable continuously throughout a whole day and indeed found the changes similar to that of the phase of the complex gains.

The change of the amplitudes of the gains are more irregular, though some correlation with temperature change can still

be seen. The gains during the night (lower temperature) are higher than during the day (higher temperature), consistent with our expectation for the gain of amplifiers.



**Figure 24:** The variation of the gain amplitude (top panel, shown as a percentage) and the phase (middle panel) of 3 typical feed channels A15X (red), A28Y (green), C22Y (blue), and the corresponding temperature (bottom) for 6 consecutive days starting from 2018/03/22.

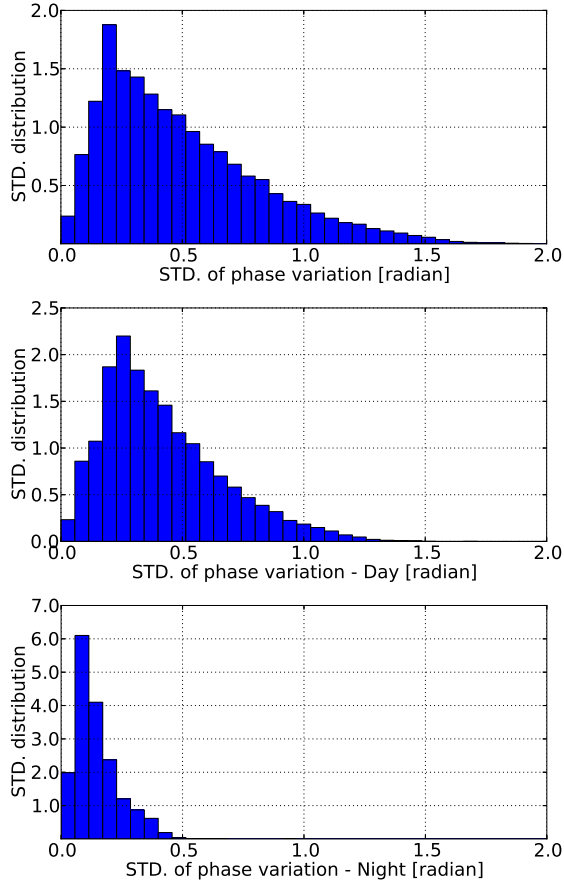
Some rapid changes are understood to be caused by some external events. For example, the sudden changes at the position marked with “a” in the figure is probably associated with the change of the dish array pointing at that time. The dish array motor generates some RFI, and the change in pointing of the dish may also change their reflection of some radio waves. The position marked with “b” may be associated with the Sun. However, surprisingly, the variations of the phases at those times are still regular. Still, the causes of some changes remain unknown.

We note here that the gains are obtained by relative calibration with the CNS. The CNS signal is received in the far side lobes of the receiver feeds. The amplitude may be more susceptible to small changes, e.g. the slight swing of the feed by wind which could shift the side lobe significantly. This suggests that the amplitude of the gain calibrated with the CNS should be used with caution.

Given the relatively calibrated complex pair gains, the variation of each pair over time can be characterized by the standard deviation (STD). A distributions of the STDs of the phase variation in the pair gain for different baselines and frequencies during the course of 5 days are plotted in Figure 25, where we also distinguish the whole day (top), day time (middle panel) and night time (bottom channel) cases. For this figure, the day time is defined as 6 AM to 6 PM local civil time, and night time 6 PM to 6 AM local civil time.

The peak value of the daytime STD is around 0.3 rad and it extends to 1.0 rad, while the peak value of nighttime STD is

around 0.1 rad, and extends to less than 0.5 rad. This shows that the phase variation in nighttime is much smaller than in daytime.



**Figure 25:** Distribution of the STD. of the phase of complex gain variations for whole day(top), day time (middle) and night time (bottom) over 5 days.

## 7 Sensitivity and System Temperature

The visibilities obtained after the above calibration procedure are still in ADC units. In order to determine the sensitivity and system temperature, we use a bright point source as a calibrator to convert the visibility to the antenna temperature.

If we observe a strong point source such as Cyg A, at the peak of its transit of the beam, in the direction of  $\vec{n}$ , with flux density  $S_0$ , the spectral power received in one polarization by the receiver is

$$P_v = \frac{1}{2} \eta A_e f_{ps} S_0. \quad (20)$$

Here, the effective area  $A_e$  is for the direction of maximum response,  $f_{ps} \equiv f(\vec{n})$  is the beam profile in the direction of the point source normalized to  $f_{\max} = 1$ , and the efficiency factor  $\eta \leq 1$ . The radiation from the point source induces an antenna temperature  $T_A^{\text{ps}}$ , with  $P_v = \eta k T_A^{\text{ps}}$ , where  $k$  is the Boltz-

mann constant. If the raw visibility (in ADC units) of channels  $a, b$  is related to the antenna temperature by  $V_{ab}^{\text{ps}} = C T_A^{\text{ps}}$ , we can determine the constant  $C$  using either the mean auto-correlation or the variance of the cross-correlation.

For the cross-correlation, the variance of the visibility is defined as  $\sigma_{ab}^2 \equiv \langle |V_{ab} - \bar{V}_{ab}|^2 \rangle$ , which is given by

$$\sigma_{ab} = C \frac{T_{\text{sys}}^{ab}}{\sqrt{2\delta\nu\delta t}}, \quad (21)$$

where  $\delta t$  is the integration time and  $\delta\nu$  is the bandwidth. One can then infer  $T_{\text{sys}}$  from the variance of measured visibilities:

$$T_{\text{sys}}^{ab} = \frac{\sigma_{ab}}{V_{ab}^{\text{ps}}} T_A^{\text{ps}} \sqrt{2\delta\nu\delta t}. \quad (22)$$

The system temperature associated with the  $a, b$  channels is a geometric average of the system temperature for the  $a, b$  channels,

$$T_{\text{sys}}^{ab} = \sqrt{T_{\text{sys}}^a T_{\text{sys}}^b}. \quad (23)$$

From the system temperature of the pairs,  $T_{\text{sys}}^{ab}$ , one can solve for the system temperature associated with individual feeds. A method of solution using only different cylinder cross-correlations is given in the Appendix, and it is the same formula as the least square solution for the determination of the center of the beam from cross correlations in Sec.5.

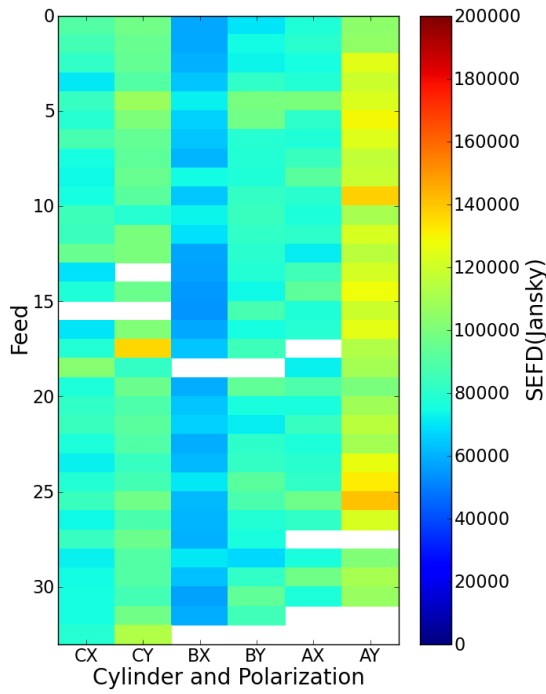
To derive the system temperature using the cross-correlation fluctuation (Eq. 22) requires knowing the antenna temperature, which depends on the effective area of the antenna. The oft-used system equivalent flux (SEFD, in units of Jansky) is

$$\text{SEFD} \equiv \frac{2kT_{\text{sys}}}{A_e} = f_{ps} S_0 \frac{\sigma_{ab} \sqrt{2\delta\nu\delta t}}{V_{ab}^{\text{ps}}}. \quad (24)$$

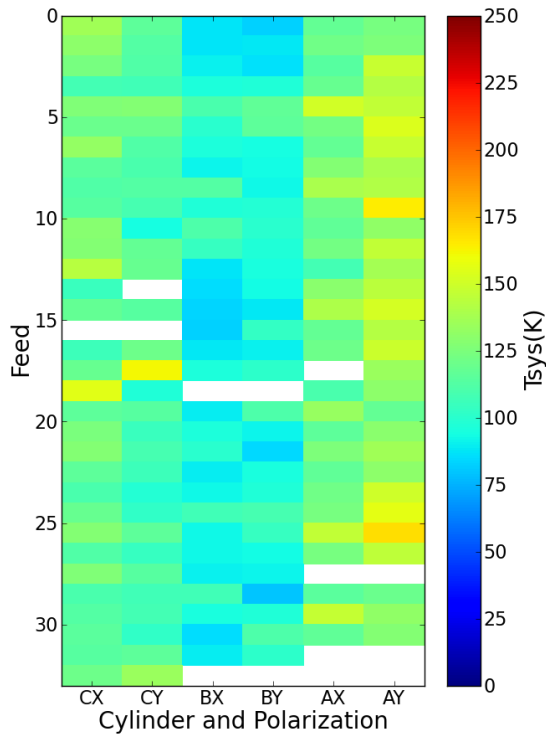
Regardless of the value of the antenna effective area and beam response, we can determine  $(f_{ps})^{-1} \text{SEFD}$ , which is given entirely by the observational data.

Fig. 26 shows the SEFD derived using the cross-correlation data collected during the transit of Cyg A on 2016/09/28. We arrange the feeds according to their cylinder and polarization (horizontal) and feed number (vertical). The malfunctioning feed channels are masked (shown as white blocks in the figure), and the A (B) cylinder has only 31 (32) feeds, hence the 2 (4) white blocks at the bottom of their respective columns.

The mean SEFD for all feeds as derived from the fluctuation of cross-correlations is  $(f_{ps})^{-1} \langle \text{SEFD}_X \rangle = 7.467 \times 10^4$  Jy with variance  $\sigma_X = 1.097 \times 10^4$  Jy;  $(f_{ps})^{-1} \langle \text{SEFD}_Y \rangle = 9.792 \times 10^4$  Jy with variance  $\sigma_Y = 1.774 \times 10^4$  Jy.



**Figure 26:** The  $f_{ps}^{-1}$  SEFD of all feeds derived from the cross-correlation data during the Cyg A transit on 2016/09/28.



**Figure 27:** The  $T_{sys}$  of all feeds derived from the cross-correlation data during the Cyg A transit on 2016/09/28.

Using the CST simulation described in Sec. 5 we find  $f_{ps}^X = 0.9795$  and  $f_{ps}^Y = 0.9908$ , which are fairly close to

1. The effective area is  $A_e^X = 4.22 \text{ m}^2$ , and  $A_e^Y = 3.35 \text{ m}^2$ .

**Table 4:** The mean  $T_{sys}$  of different cylinders and polarizations. (Unit: K)

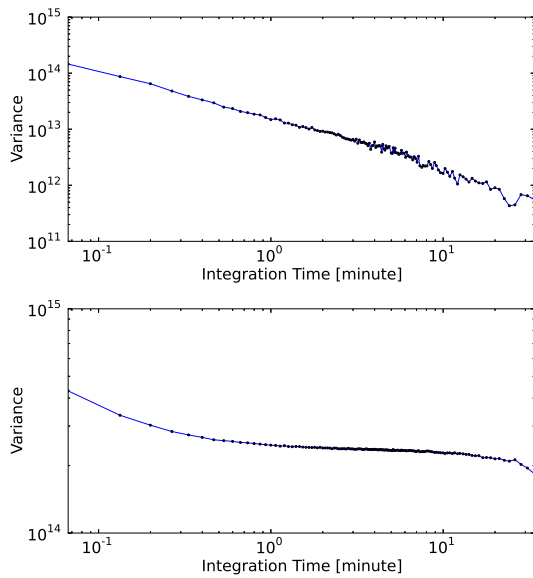
Polarization	Cylinder			
	A	B	C	All
X	124.1	94.8	121.1	113.1
Y	140.3	96.7	112.0	116.4
All	132.3	95.7	116.6	114.7

We estimate the average system temperature to be  $T_{sys}^X = 113.1 \text{ K}$  and  $T_{sys}^Y = 116.4 \text{ K}$ , and the average for all cylinders and polarizations is  $114.7 \text{ K}$ . The system temperatures for the feeds are plotted in Fig. 27, and the averages for different cylinders and polarizations are listed in Table. 4.

From the Figure, we see there are obvious differences in the system temperatures for the feeds on the three cylinders. Many feeds on Cylinder A, especially the Y (E-W) polarization ones, have high system temperatures, with an average of  $140.3 \text{ K}$ . The average system temperature for both polarizations is  $132.2 \text{ K}$ . By comparison, the feeds on Cylinder B have much lower temperature, with an average of only  $95.7 \text{ K}$ . The feeds on cylinder C are somewhere between these two, with an average of  $116.6 \text{ K}$ . There are also some variations within each cylinder, with some particularly high and lower ones. It is not known why the system temperatures differ so much. One factor is that the B cylinder is in the center, so it may be less affected by RFI and stray radiation. Cylinder A on the East is closer to the dish array, which may have some impact on the system temperature.

Generally speaking, the system temperatures we found for the cylinder array feeds are comparable with what we see on the dish array (F. Wu et al., *in preparation*), which shares similar electronics. The variation of system temperature for different feeds should be taken into account when making synthesized images from the visibility data to make optimal maps.

Finally, we consider how the variance of the visibilities scales as a function of integration time. In Fig. 28 we plot the variance for different integration time scales while observing a relatively blank sky using data on 2018/03/22. For a typical baseline, for example A1Y-B2Y (top panel), the variance drops steadily as the integration time increases. However, for the feed pairs which are very close to each other, such as A1Y-A2Y (bottom panel), the variance hits a floor after just about a minute of integration, showing the strong coupling and correlated noise in the short baselines. Thus, at least for the longer baselines, the sensitivity can be improved by accumulating longer integration time.



**Figure 28:** The variance for different integration time scales while observing a relatively blank sky using data on 2018/03/22. Top: variance of baseline A1Y-B2Y versus integration time (averaged visibility). Bottom: for baseline A1Y-A2Y.

## 8 Conclusion

Neutral hydrogen, ubiquitous throughout cosmic history since recombination, is a potentially powerful tool for cosmological observations. To achieve high precision measurements of the redshifted 21 cm emission, a number of dedicated 21 cm array concepts have been built or are under construction. Even more ambitious arrays, involving tens of thousands of antennas [30, 31], have been proposed. Dedicated 21 cm arrays have some common design features determined largely by the nature of the observation and current technology. They all use large numbers of relatively small, inexpensive antennas. The antennas are either fixed, or can only move in elevation, and the observations use either a drift scan mode, or track the target with electronically steered beams. With computing power following Moore’s law scaling and price per computing operation and of data storage continuously dropping, this approach enables arrays of very large scales to be built with a very moderate cost. Based on simple forecasts, the projected capabilities of these arrays are very impressive (see e.g. Ref. [7, 32, 33]). However, extracting scientific results from these arrays poses a number of challenges. For example, these arrays produce a huge amount of data. In addition, with small antennas and uncooled receivers the SNR of the raw data is relatively low; only a few bright sources are available for simple point source calibration, and as the antennas are not movable, calibration with strong point sources can only be performed occasionally. Furthermore, while for forecasting it is customary to assume

that the antenna responses are identical, in reality each unit is somewhat different. It is crucial to develop technologies to handle the data from such arrays, and to test the key technologies in order to gain some concrete experience and expose possible problems. The Tianlai pathfinder arrays were built expressly for this purpose.

In this paper, we first described the system functions of the Tianlai Cylinder Pathfinder (Sec. 2). The hardware (Sec. 2.1) consists of three cylinder reflectors and a total of 96 feeds, the optical analog signal transmission system, the down-converters which convert the RF signal to IF, and the digital FX correlator which produces the visibilities. We also briefly introduced the observational data sets (Sec. 2.2) and basic procedures for data processing (Sec. 2.3). The noise figure of the LNAs and the linearity of the system is described in Sec. 3.

We then presented analysis of the observational data (Sec. 4). We showed a few examples of the raw visibilities, from which we can see that they have a simple repeating pattern day after day, showing the system is generally stable. Transits of bright sources can be recognized, or become visible after a noise background is removed. We also find that the correlated noise is stronger for feeds close to each other on the same cylinder, making it harder to use the visibilities from nearby pairs.

In Sec. 5 we studied the beam profile of our telescope. We used electromagnetic simulations to obtain the beam profile of the cylinder in the 700–800 MHz frequency range. We then measured the beam profile in the E-W direction by analyzing auto-correlations and cross-correlations using the strong source Cyg A, which passes very close to the zenith during its transit of the meridian. The general shape of the beam is consistent with the simulation, but a few feeds are slightly misaligned, with an error at the level of  $\sim 0.15^\circ$ , consistent with what one would expect for the precision of installation using mechanical tools. This is an example of the small non-uniformity which occurs naturally in the construction of large radio arrays. Not taking into account such differences can induce errors in the final data.

Calibration is a crucial step for the processing of interferometer array data, and we presented our research in Sec. 6. First, we performed bandpass calibrations for individual visibilities, and found relatively stable results. The bandpasses obtained for the cross-correlation visibilities are consistent with the product of the bandpasses of individual feeds, as obtained from the auto-correlations. More important is the calibration of complex gains. For an array with a large field of view and low sensitivity, such as the cylinders, a special challenge is that on the one hand, there are few sources which are bright enough to be “seen” directly in individual visibilities,

and on the other hand, it is very difficult to construct a sky model that has enough precision over the large field of view to include all sources that contribute to the visibility. In our experiment we tried two methods of calibration. In what we termed absolute calibration, we use the transit of the strong astronomical source Cyg A to solve for the complex gain of each input channel, using the eigen-decomposition method we developed earlier [26]. At other times, we use a calibration signal broadcast from an artificial source to perform what we call relative calibration. The results of the two methods are compared for the period when they overlap, and distributions of the complex gains are given. The distribution and variations of the gain are plotted, and the phase of the gain is found to be strongly correlated with the environmental temperature. We also checked the precision of this method using the closure phase relation.

Finally, in Sec. 7, we estimated the SEFD and system temperature of the array feeds using the ratio of visibility and its fluctuation during the transit of Cyg A. We found that feeds on cylinder A have generally higher system temperature, averaging 132 K, while feeds on cylinder B have the lowest system temperature, averaging 95.7 K. The average system temperature for all feeds is 113 K for X polarization and 116 K for Y polarization. This is in general agreement with the system temperature obtained for the Tianlai Dish Array, which has similar electronics. We also examined how the variance of the visibility scales with time. For long baselines the variance decreases steadily as integration time increases, showing that the measurement error could be reduced by accumulating more data. However, for shorter baselines the variance hit a floor after a short time, presumably caused by the coupling between nearby feeds. At least for the baselines across different cylinders, higher sensitivity can be achieved with longer integration time within the range we tested.

In this work, we have only studied a small fraction of the data set collected so far, and presented only some basic performance characteristics. There remain many problems to be solved, for example, the cross-coupling between feeds, the determination of the beam profile in the N-S direction, calibration methods of better precision, map-making, foreground subtraction, etc. More thorough analyses of these problems will be investigated in a series of future works.

**Acknowledgements.** The Tianlai cylinder is operated with the support of NAOC Astronomical Technology Center. The Tianlai cylinder is built with the support of MOST grant 2012AA121701, and its survey is supported by MOST grant 2016YFE0100300, the NSFC 11633004 and 11473044 and the CAS QYZDJ-SSW-SLH017. Some on-site experiments are performed with support of the Joint Research Fund in Astronomy (U1631118) under cooperative agreement between

NSFC and CAS. The data analysis work is partially supported by the National Key R&D Program 2017YFA0402603 and the CAS Interdisciplinary Innovation Team (JCTD-2019-05). Part of the computations are performed on the Tianhe-2 supercomputer (with the support of NSFC grant U1501501) and the Tianhe-1 supercomputer. 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. Authors affiliated with French institutions acknowledge partial support from CNRS (IN2P3 & INSU), Observatoire de Paris and from Irfu/CEA.

**Conflict of interest** The authors declare that they have no conflict of interest.

- 1 Ely Kovetz, Patrick C. Breysse, Adam Lidz, Jamie Bock, Charles M. Bradford, Tzu-Ching Chang, Simon Foreman, Hamsa Padmanabhan, Anthony Pullen, Dominik Riechers, Marta B. Silva, and Eric Switzer. Astrophysics and Cosmology with Line-Intensity Mapping. *BAAS*, 51(3):101, May 2019.
- 2 Adrian Liu and J. Richard Shaw. Data Analysis for Precision 21 cm Cosmology. *arXiv e-prints*, page arXiv:1907.08211, Jul 2019.
- 3 Miguel F. Morales and J. Stuart B. Wyithe. Reionization and Cosmology with 21-cm Fluctuations. *ARAA*, 48:127–171, Sep 2010.
- 4 K. W. Masui, E. R. Switzer, N. Banavar, K. Bandura, C. Blake, L.-M. Calin, T.-C. Chang, X. Chen, Y.-C. Li, Y.-W. Liao, A. Natarajan, U.-L. Pen, J. B. Peterson, J. R. Shaw, and T. C. Voytek. Measurement of 21 cm brightness fluctuations at  $z = 0.8$  in cross-correlation. *The Astrophysical Journal Letters*, 763(1):L20, 2013.
- 5 E. R. Switzer et al. Determination of  $z = 0.8$  neutral hydrogen fluctuations using the 21 cm intensity mapping auto-correlation. *Mon. Not. Roy. Astron. Soc.*, 434:L46, 2013.
- 6 C. J. Anderson, N. J. Luciw, Y. C. Li, C. Y. Kuo, J. Yadav, K. W. Masui, T. C. Chang, X. Chen, N. Oppermann, Y. W. Liao, U. L. Pen, D. C. Price, L. Staveley-Smith, E. R. Switzer, P. T. Timbie, and L. Wolz. Low-amplitude clustering in low-redshift 21-cm intensity maps cross-correlated with 2dF galaxy densities. *MNRAS*, 476(3):3382–3392, May 2018.
- 7 Yidong Xu, Xin Wang, and Xuelei Chen. Forecasts on the Dark Energy and Primordial Non-Gaussianity Observations with the Tianlai Cylinder Array. *ApJ*, 798(1):40, Jan 2015.
- 8 Yidong Xu, Jan Hamann, and Xuelei Chen. Precise measurements of inflationary features with 21 cm observations. *PRD*, 94(12):123518, Dec 2016.
- 9 M. P. van Haarlem, M. W. Wise, A. W. Gunst, G. Heald, J. P. McKean, J. W. T. Hessels, A. G. de Bruyn, R. Nijboer, J. Swinbank, R. Fallows, M. Brentjens, A. Nelles, R. Beck, H. Falcke, R. Fender, J. Hörandel, L. V. E. Koopmans, G. Mann, G. Miley, H. Röttgering, B. W. Stappers, R. A. M. J. Wijers, S. Zaroubi, M. van den Akker, A. Alexov, J. Anderson, K. Anderson, A. van Ardenne, M. Arts, A. Asgekar, I. M. Avruch, F. Batejat, L. Bähren, M. E. Bell, M. R. Bell, I. van Bemmelen, P. Bennema, M. J. Bentum, G. Bernardi, P. Best, L. Birzan, A. Bonafede, A.-J. Boonstra, R. Braun, J. Bregman, F. Breitling, R. H. van de Brink, J. Broderick, P. C. Broekema, W. N. Brouw, M. Brüggen, H. R. Butcher, W. van Cappellen, B. Ciardi, T. Coenen, J. Conway, A. Coolen, A. Corstanje, S. Damstra, O. Davies, A. T. Deller, R.-J. Dettmar, G. van Diepen, K. Dijkstra, P. Donker, A. Doorduyn, J. Dromer, M. Drost, A. van Duin, J. Eislöffel, J. van Enst, C. Ferrari, W. Frieswijk, H. Gankema, M. A. Garrett, F. de Gasperin,

- M. Gerbers, E. de Geus, J.-M. Grießmeier, T. Grit, P. Gruppen, J. P. Hamaker, T. Hassall, M. Hoefl, H. A. Holties, A. Horneffer, A. van der Horst, A. van Houwelingen, A. Huijgen, M. Iacobelli, H. Intema, N. Jackson, V. Jelic, A. de Jong, E. Juette, D. Kant, A. Karastergiou, A. Koers, H. Kollen, V. I. Kondratiev, E. Kooistra, Y. Koopman, A. Koster, M. Kuniyoshi, M. Kramer, G. Kuper, P. Lambropoulos, C. Law, J. van Leeuwen, J. Lemaître, M. Loose, P. Maat, G. Macario, S. Markoff, J. Masters, R. A. McFadden, D. McKay-Bukowski, H. Meijering, H. Meulman, M. Mevius, E. Middelberg, R. Millenaar, J. C. A. Miller-Jones, R. N. Mohan, J. D. Mol, J. Morawietz, R. Morganti, D. D. Mulcahy, E. Mulder, H. Munk, L. Nieuwenhuis, R. van Nieuwpoort, J. E. Noordam, M. Norden, A. Noutsos, A. R. Offringa, H. Olofsson, A. Omar, E. Orrù, R. Overeem, H. Paas, M. Pandey-Pommier, V. N. Pandey, R. Pizzo, A. Polatidis, D. Rafferty, S. Rawlings, W. Reich, J.-P. de Reijer, J. Reitsma, G. A. Renting, P. Riemers, E. Rol, J. W. Romein, J. Roosjen, M. Ruiter, A. Scaife, K. van der Schaaf, B. Scheers, P. Schellart, A. Schoenmakers, G. Schoonderbeek, M. Serylak, A. Shulevski, J. Sluman, O. Smirnov, C. Sobey, H. Spreeuw, M. Steinmetz, C. G. M. Sterks, H.-J. Stiepel, K. Stuurwold, M. Tagger, Y. Tang, C. Tasse, I. Thomas, S. Thoudam, M. C. Toribio, B. van der Tol, O. Usov, M. van Veelen, A.-J. van der Veen, S. ter Veen, J. P. W. Verbiest, R. Vermeulen, N. Vermaas, C. Vocks, C. Vogt, M. de Vos, E. van der Wal, R. van Weeren, H. Weggemans, P. Weltevrede, S. White, S. J. Wijnholds, T. Wilhelmsson, O. Wucknitz, S. Yatawatta, P. Zarka, A. Zensus, and J. van Zwieten. LOFAR: The LOw-Frequency ARray. *A&A*, 556:A2, August 2013.
- 10 S. J. Tingay, R. Goetze, J. D. Bowman, D. Emrich, S. M. Ord, D. A. Mitchell, M. F. Morales, T. Boole, B. Crosse, R. B. Wayth, C. J. Lonsdale, S. Tremblay, D. Pallot, T. Colegate, A. Wicenc, N. Kudryavtseva, W. Arcus, D. Barnes, G. Bernardi, F. Briggs, S. Burns, J. D. Burton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. Desouza, B. M. Gaensler, L. J. Greenhill, P. J. Hall, B. J. Hazelton, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, M. J. Lynch, B. Mckinley, S. R. McWhirter, E. Morgan, D. Oberoi, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. E. Rogers, A. Rosh, J. E. Salah, R. J. Sault, N. Udaya-Shankar, F. Schlagenhafer, K. S. Srivani, J. Stevens, R. Subrahmanyam, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. B. Wyithe. The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies. *PASA*, 30:e007, January 2013.
  - 11 David R DeBoer, Aaron R Parsons, James E Aguirre, Paul Alexander, Zaki S Ali, Adam P Beardsley, Gianni Bernardi, Judd D Bowman, Richard F Bradley, Chris L Carilli, et al. Hydrogen epoch of reionization array (hera). *Publications of the Astronomical Society of the Pacific*, 129(974):045001, 2017.
  - 12 Aaron R Parsons, Donald C Backer, Griffin S Foster, Melvyn CH Wright, Richard F Bradley, Nicole E Gugliucci, Chaitali R Parashare, Erin E Benoit, James E Aguirre, Daniel C Jacobs, et al. The precision array for probing the epoch of re-ionization: Eight station results. *The Astronomical Journal*, 139(4):1468, 2010.
  - 13 Michael W. Eastwood, Marin M. Anderson, Ryan M. Monroe, Gregg Hallinan, Benjamin R. Barsdell, Stephen A. Bourke, M. A. Clark, Steven W. Ellington, Jayce Dowell, Hugh Garsden, Lincoln J. Greenhill, Jacob M. Hartman, Jonathon Kocz, T. Joseph W. Lazio, Danny C. Price, Frank K. Schinzel, Gregory B. Taylor, Harish K. Vedantham, Yuankun Wang, and David P. Woody. The Radio Sky at Meter Wavelengths: m-mode Analysis Imaging with the OVRO-LWA. *AJ*, 156(1):32, Jul 2018.
  - 14 Kevin Bandura, Graeme E Addison, Mandana Amiri, J Richard Bond, Duncan Campbell-Wilson, Liam Connor, Jean-François Cliche, Greg Davis, Meiling Deng, Nolan Denman, et al. Canadian hydrogen intensity mapping experiment (chime) pathfinder. In *Ground-based and Airborne Telescopes V*, volume 9145, page 914522. International Society for Optics and Photonics, 2014.
  - 15 Xuelei Chen. The Tianlai Project: a 21CM Cosmology Experiment. In *International Journal of Modern Physics Conference Series*, volume 12 of *International Journal of Modern Physics Conference Series*, pages 256–263, Mar 2012.
  - 16 Jiao Zhang, Shi-Fan Zuo, Reza Ansari, Xuelei Chen, Yi-Chao Li, Feng-Quan Wu, Jean-Eric Campagne, and Christophe Magneville. Sky reconstruction for the Tianlai cylinder array. *Research in Astronomy and Astrophysics*, 16(10):158, Oct 2016.
  - 17 Santanu Das, Christopher J Anderson, Reza Ansari, Jean-Eric Campagne, Daniel Charlet, Xuelei Chen, Zhiping Chen, Aleksander J Cianciara, Pierre Colom, Yanping Cong, et al. Progress in the construction and testing of the tianlai radio interferometers. In *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy IX*, volume 10708, page 1070836. International Society for Optics and Photonics, 2018.
  - 18 LB Newburgh, K Bandura, MA Bucher, T-C Chang, HC Chiang, JF Cliche, R Davé, M Dobbs, C Clarkson, KM Ganga, et al. Hirax: a probe of dark energy and radio transients. In *Ground-based and Airborne Telescopes VI*, volume 9906, page 99065X. International Society for Optics and Photonics, 2016.
  - 19 Suman Chatterjee and Somnath Bharadwaj. A spherical harmonic analysis of the Ooty Wide Field Array (OWFA) visibility signal. *MNRAS*, 478(3):2915–2926, Aug 2018.
  - 20 R. Battye, I. Browne, T. Chen, C. Dickinson, S. Harper, L. Olivari, M. Peel, M. Remazeilles, S. Roychowdhury, P. Wilkinson, E. Abdalla, R. Abramo, E. Ferreira, A. Wuensche, T. Vilella, M. Caldas, G. Tancredi, A. Refregier, C. Monstein, F. Abdalla, A. Pourtsidou, B. Maffei, G. Pisano, and Y.-Z. Ma. Update on the BINGO 21cm intensity mapping experiment. *ArXiv e-prints*, October 2016.
  - 21 Jiao Zhang, Reza Ansari, Xuelei Chen, Jean-Eric Campagne, Christophe Magneville, and Fengquan Wu. Sky reconstruction from transit visibilities: PAON-4 and Tianlai dish array. *MNRAS*, 461(2):1950–1966, Sep 2016.
  - 22 R Ansari, J E Campagne, D Charlet, M Moniez, C Pailler, O Perdureau, M Taurigna, J M Martin, F Rigaud, P Colom, and et al. Design, operation and performance of the paon4 prototype transit interferometer. *Monthly Notices of the Royal Astronomical Society*, 493(2):2965–2980, Mar 2020.
  - 23 Qizhi Huang, Fengquan Wu, and Xuelei Chen. A high-resolution self-consistent whole sky foreground model. *Sci. China Phys. Mech. Astron.*, 62(8):989511, 2019.
  - 24 A. Richard Thompson, James M. Moran, and Jr. Swenson, George W. *Interferometry and Synthesis in Radio Astronomy, 3rd Edition*. 2017.
  - 25 Aleksander J Cianciara, Christopher J Anderson, Xuelei Chen, Zhiping Chen, Jingchao Geng, Jixia Li, Chao Liu, Tao Liu, Wing Lu, Jeffrey B Peterson, et al. Simulation and testing of a linear array of modified four-square feed antennas for the tianlai cylindrical radio telescope. *Journal of Astronomical Instrumentation*, 6(03):1750003, 2017.
  - 26 Shifan Zuo, Ue-Li Pen, Fengquan Wu, Jixia Li, Albert Stebbins, Yougang Wang, and Xuelei Chen. An Eigenvector-Based Method of Radio Array Calibration and Its Application to the Tianlai Cylinder Pathfinder. *AJ*, 157(1):34, Jan 2019.
  - 27 Shifan Zuo, Jixia Li, Yichao Li, Das Santanu, Albert Stebbins, Jiao Zhang, Fengquan Wu, and Xuelei Chen. Data Processing Pipeline For Tianlai Experiment. 2020.
  - 28 Laura B. Newburgh, Graeme E. Addison, Mandana Amiri, Kevin Bandura, J. Richard Bond, Liam Connor, Jean-François Cliche, Greg Davis, Meiling Deng, Nolan Denman, Matt Dobbs, Mateus Fandino, Heather Fong, Kenneth Gibbs, Adam Gilbert, Elizabeth Griffin, Mark Halpern, David Hanna, Adam D. Hincks, Gary Hinshaw, Carolin Höfer, Peter Klages, Tom Landecker, Kiyoshi Masui, Juan Mena Parra, Ue-Li Pen, Jeff Peterson, Andre Recnik, J. Richard Shaw, Kris Sigurdson, Micheal Sitwell, Graeme Smecher, Rick Smegal, Keith Vand



erlinde, and Don Wiebe. *Calibrating CHIME: a new radio interferometer to probe dark energy*, volume 9145 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 91454V. 2014.

- 29 R. A. Perley and B. J. Butler. An Accurate Flux Density Scale from 50 MHz to 50 GHz. *ApJs*, 230(1):7, May 2017.
- 30 Cosmic Visions 21 cm Collaboration, Réza Ansari, Evan J. Arena, Kevin Bandura, Philip Bull, Emanuele Castorina, Tzu-Ching Chang, Shi-Fan Chen, Liam Connor, Simon Foreman, Josef Frisch, Daniel Green, Matthew C. Johnson, Dionysios Karagiannis, Adrian Liu, Kiyoshi W. Masui, P. Daniel Meerburg, Moritz Münchmeyer, Laura B. Newburgh, Andrej Obuljen, Paul O'Connor, Hamsa Padmanabhan, J. Richard Shaw, Christopher Sheehy, Anže Slosar, Kendrick Smith, Paul Stankus, Albert Stebbins, Peter Timbie, Francisco Villaescusa-Navarro, Benjamin Wallisch, and Martin White. Inflation and Early Dark Energy with a Stage II Hydrogen Intensity Mapping Experiment. *arXiv e-prints*, page arXiv:1810.09572, Oct 2018.
- 31 Anze Slosar, Zeeshan Ahmed, David Alonso, Mustafa A. Amin, Evan J. Arena, Kevin Bandura, Nicholas Battaglia, Jonathan Blazek, Philip Bull, Emanuele Castorina, Tzu-Ching Chang, Liam Connor, Romeel Davé, Cora Dvorkin, Alexander van Engelen, Simone Ferraro, Raphael Flauger, Simon Foreman, Josef Frisch, Daniel Green, Gilbert Holder, Daniel Jacobs, Matthew C. Johnson, Joshua S. Dillon, Dionysios Karagiannis, Alexander A. Kaurov, Lloyd Knox, Adrian Liu, Marilena Loverde, Yin-Zhe Ma, Kiyoshi W. Masui, Thomas McClintock, Kavilan Moodley, Moritz Munchmeyer, Laura B. Newburgh, Cherry Ng, Andrei Nomerotski, Paul O'Connor, Andrej Obuljen, Hamsa Padmanabhan, David Parkinson, J. Xavier Prochaska, Surjeet Rajendran, David Rapetti, Benjamin Saliwanchik, Emmanuel Schaan, Neelima Sehgal, J. Richard Shaw, Chris Sheehy, Erin Sheldon, Raphael Shirley, Eva Silverstein, Tracy Slatyer, Anze Slosar, Paul Stankus, Albert Stebbins, Peter T. Timbie, Gregory S. Tucker, William Tyndall, Francisco Villaescusa Navarro, Benjamin Wallisch, and Martin White. Packed Ultra-wideband Mapping Array (PUMA): A Radio Telescope for Cosmology and Transients. In *BAAS*, volume 51, page 53, Sep 2019.
- 32 H.-J. Seo, S. Dodelson, J. Marriner, D. McGinnis, A. Stebbins, C. Stoughton, and A. Vallinotto. A Ground-based 21 cm Baryon Acoustic Oscillation Survey. *ApJ*, 721:164–173, September 2010.
- 33 R. Ansari, J. E. Campagne, P. Colom, J. M. Le Goff, C. Magneville, J. M. Martin, M. Moniez, J. Rich, and C. Yèche. 21 cm observation of large-scale structures at  $z \sim 1$ . Instrument sensitivity and foreground subtraction. *A&A*, 540:A129, April 2012.

## Appendix A System Temperature of Channels from Cross Correlation Pairs

We have assumed the variance of the visibility is given by

$$\sigma_{ab}^2 = \langle |V_{ab} - \bar{V}_{ab}|^2 \rangle = \langle n_a^2 n_b^2 \rangle = C^2 \frac{T_{sab}^2}{\delta t \delta \nu}. \quad (\text{a1})$$

However, in fact, the system temperature here is actually the geometric mean of a pair,

$$T_{sab}^2 = T_a T_b. \quad (\text{a2})$$

How do we obtain each system temperature,  $T_a$ , for the individual feeds?

If we take  $\ell_{ab} = \log(T_{sab})$ ,  $\ell_a = \log T_a$ ,  $\ell_b = \log T_b$ , we have

$$\ell_{ab} = \frac{\ell_a + \ell_b}{2}. \quad (\text{a3})$$

Now  $\ell_{ab}$  for all pairs is known, and we need to solve for  $\ell_a$ . We use the average over many pairs to reduce fluctuations, and, to avoid the effect of cross-coupling between feeds, we only use correlations for pairs on different cylinders. (Note that, this method can also be used to solve for the beam pointing in Eq. 7 for individual feed from the baseline beam pointing. One can simply regard  $\ell_a$  as the pointing of individual feed and  $\ell_{ab}$  as the transit peak time of the pair.) The solution of this mathematical problem is as follow.

Without loss of generality, we consider the feed  $a$  on cylinder  $A$ ,

$$\hat{\ell}_{Aa}(B) = \frac{1}{N_B} \sum_b \ell_{(Aa)(Bb)} \quad (\text{a4})$$

$$= \frac{1}{2}(\ell_{Aa} + \frac{1}{N_B} \sum_b \ell_b). \quad (\text{a5})$$

Introducing the notation

$$L_B = \frac{1}{N_B} \sum_b \ell_b, \quad (\text{a6})$$

we have

$$\hat{\ell}_{Aa}(B) = \frac{1}{2}(\ell_{Aa} + L_B), \quad \hat{\ell}_{Aa}(C) = \frac{1}{2}(\ell_{Aa} + L_C). \quad (\text{a7})$$

Similar equations can be obtained by cycling through  $A, B, C$ . Thus, the formula for getting the  $\ell_{Aa}$  using all different cylinder cross correlations is

$$\ell_{Aa} = \hat{\ell}_{Aa}(B) + \hat{\ell}_{Aa}(C) - \frac{1}{2}(L_B + L_C). \quad (\text{a8})$$

If we know  $L_B$  and  $L_C$  we can obtain the desired results.

To obtain  $L_A, L_B, L_C$ , we note that

$$\sum_a \hat{\ell}_{Aa}(B) = \frac{1}{2}(\sum_a \ell_{Aa} + N_A L_B),$$

using Eq. (a4), and we have

$$\sum_a \hat{\ell}_{Aa}(B) = \frac{N_A}{2}(L_A + L_B). \quad (\text{a9})$$

Thus, we have

$$L_A + L_B = \frac{2}{N_A N_B} \sum_{a,b} \hat{\ell}_{(Aa)(Bb)} \quad (\text{a10})$$

$$L_B + L_C = \frac{2}{N_B N_C} \sum_{b,c} \hat{\ell}_{(Bb)(Cc)} \quad (\text{a11})$$

$$L_C + L_A = \frac{2}{N_C N_A} \sum_{c,a} \hat{\ell}_{(Cc)(Aa)}. \quad (\text{a12})$$

Note the R.H.S. is entirely known from the observations.

Substituting these expressions into Eq.(a8), we finally find the solution to be

$$\begin{aligned} \ell_{Aa} &= \frac{1}{N_B} \sum_b \ell_{(Aa)(Bb)} + \frac{1}{N_C} \sum_c \ell_{(Cc)(Aa)} - \frac{1}{N_B N_C} \sum_{b,c} \ell_{(Bb)(Cc)} \\ \ell_{Bb} &= \frac{1}{N_C} \sum_c \ell_{(Bb)(Cc)} + \frac{1}{N_A} \sum_a \ell_{(Aa)(Bb)} - \frac{1}{N_C N_A} \sum_{c,a} \ell_{(Cc)(Aa)} \end{aligned}$$

$$\ell_{Cc} = \frac{1}{N_A} \sum_a \ell_{(Cc)(Aa)} + \frac{1}{N_B} \sum_b \ell_{(Bb)(Cc)} - \frac{1}{N_A N_B} \sum_{a,b} \ell_{(Aa)(Bb)}. \quad (\text{a13})$$

Thus, applying Eq. (a13) we can get the system temperature for each feed from the cross-correlation of feed pairs.