Progress in Single Dish 21 cm Intensity Mapping

C. J. Anderson*

Department of Physics, University of Wisconsin Madison

Abstract

Progress has been made in finalizing the design of a 7 element array to be installed on the Green Bank Telescope in the 700 to 900 MHz frequency range. The new array should increase HI intensity mapping speed by a factor of 4. The design review for the array is scheduled for October 19. At the same time, 21 cm intensity mapping data collected from the Parkes radio telescope has been analyzed. We have detected HI signal in cross-correlation with the 2df Galaxy Survey.

Motivation

Hydrogen is an important tool for understanding our universe. Calculations from big bang nucleosynthesis and empirical observations both suggest that three quarters of the visible matter in the universe is hydrogen. This makes hydrogen an excellent tracer of the total matter distribution of the universe. The 21 cm line of neutral hydrogen is attractive as a detection tool for many reasons. First, the 21 cm line is in the radio wavelength range, detectable by advanced existing telescopes like the Green Bank and Parkes telescopes. The 21 cm line is also far from other atomic transitions, so confusion with other atomic sources is unlikely. Finally, although much of the hydrogen in the universe has become reionized by stellar radiation, there is still a wealth of 21 cm emitting neutral hydrogen in the form of self-shielded clumps known as Damped Lyman Alpha (DLA) systems. The high spin temperature and low optical depth of these DLAs imply that their emission amplitude will be proportional to HI density. This means that one can directly probe the three dimensional density distribution of neutral hydrogen by measuring redshifted 21 cm radiation. The beam-size of radio telescopes at these frequencies is much too large to detect individual galaxies at any significant distance, but this fact is used as an advantage in the technique of intensity mapping, which seeks only to make low resolution sky maps [5]. The large beam size allows one to make low resolution maps relatively quickly. In an intensity map, each 3 dimensional pixel contains the emission of many galaxies, but the resolution is still fine enough to resolve Baryon Acoustic Oscillations (BAOs) and to study gravitational collapse at large linear scales. By detecting the BAO signature at redshifts of 0 to 2, intensity mapping has the potential to tightly constrain the equation of state of dark energy.

The chief challenge of 21 cm intensity mapping is the fact that synchrotron emission from our galaxy and from point sources is an unavoidable foreground signal, present at every pixel. At the low redshifts we are interested in, these foregrounds are one thousand times brighter than the signal. Any successful 21 cm experiment must remove these foregrounds well enough that their residuals do not overwhelm the 21 cm signal. Foreground removal is possible mainly because the frequency spectrum of the foregrounds is known to be extremely smooth [3]. In the absence of instrumental effects, each pixel would have a large smooth signal from the foregrounds, and on top of that would be small amplitude ripples from the 21 cm signal, which are not smooth due to the clumping of matter. However, instrumental effects like imperfect bandpass calibration and a phenomenon known as mode mixing [4] lead to measured foreground signals that are not smooth. Separating signal from foreground therefore requires more sophisticated techniques. In 2010, the first high redshift (z~0.8) detection of HI was made by cross-correlating radio observations from the Green Bank Telescope (GBT) with the WiggleZ galaxy survey [1]. Foregrounds were removed using singular value decomposition (SVD) principal component analysis. Later observations of the same field have used the cross-correlation with WiggleZ and the 21 cm auto-correlation to produce a lower and upper bound (respectively) on the neutral hydrogen fraction and bias parameter [6].

GBT Intensity Mapping Efforts

In the past few years, I have collaborated with the group (led by K. W. Masui, E. R. Switzer, T.-C. Chang, and U.-L. Pen.) that made the original 21 cm detection with the GBT. So far, we have been unable to clean foregrounds to the level where a detection of HI signal can be made without cross-correlating with a galaxy survey. We believe that mode mixing due to a broad frequency dependent polarized beam has caused angular variation in polarized foregrounds to leak into our frequency modes. This increases the number of frequency modes that are contaminated by foregrounds. The problem can be alleviated by degrading the resolution of our maps, but this solution also requires increasing the angular coverage of our survey [6]. To expedite data collection for these larger maps, we have been designing a 7 element receiver to replace the current single receiver at our frequency range at the GBT.

The time it takes a radio survey to reach a desired signal to noise ratio is governed by the radiometer equation.

$$\sigma^2 = \frac{T_{sys}^2}{n\Delta t\Delta\nu} \tag{1}$$

Here, σ^2 is the variance in each pixel due to noise only. T_{sys} is the system temperature of the receiver. It contains contributions from the receiver electronics, from the sky, and from radio emission of the warm ground that is picked up by the antenna. The number of receivers is n, Δt is the length of time that the sky is observed, and $\Delta\nu$ is the frequency bandwidth of the observations. From this equation, one can see that the time required to reach a given noise variance scales as T_{sys}^2/n.

$$\Delta t = \frac{T_{sys}^2}{n\sigma^2\Delta\nu} \tag{2}$$

Thus, it is possible to decrease the time needed to make quality maps by increasing the number of receivers collecting data, but one must be careful to keep the system temperature low. Because of the limited space at the GBT focal plane, receivers for our array needed to be designed with a small cross sectional area. However, they had to also maintain a narrow beampattern to minimize the ground spill contribution to the system temperature. These are somewhat opposing goals, since the traditional way to make a receiver with a narrow beam is to increase the area. However, a solution was found that builds on an existing GBT receiver.

Following the ideas of Kooi P. S., Leong M. S., and Yeo T. S [2], we added outer corrugations to a scaled version of the short backfire antenna (SBA) design used for the 450 MHz GBT receiver. The resulting receiver, scaled to work at 800 MHz, has a beam pattern that is almost as narrow as the current horn antenna on the GBT but with less than half the diameter. Although the receiver design has only had slight changes since 2013, much work has been done in the past year to design the array superstructure, the cryogenics, the low noise amplifier, and the baluns for the antenna dipoles. A test at Green Bank in 2013 revealed impedance matching problems. A test conducted in Summer 2016 demonstrated that the impedance matching is good and that the low noise amplifier works well. I have recomputed the spill temperature with the final design (see figure 1), and I have found that it is competitive with the approximately 5 Kelvin spill temperature of the current receiver. We now expect our receivers to have a total system temperature of 33 K or below across the usable portion of the band. The current receiver has a 25 K system temperature. Therefore, from equation 2, we expect the mapping speed to increase by a factor of four.
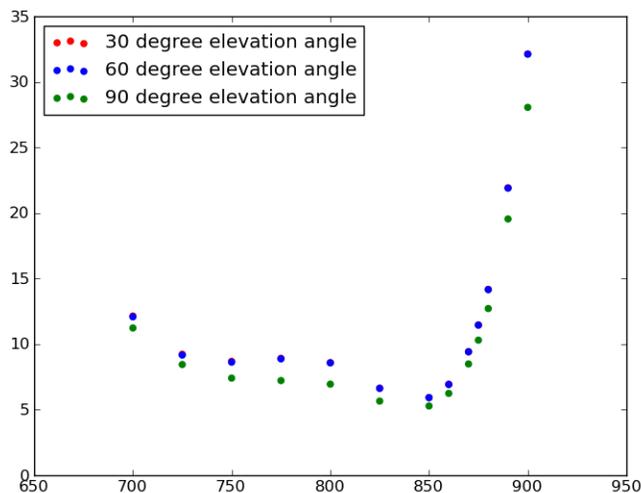
Figure 1: The projected spill temperature of each receiver over the frequency bandwidth. Spill temperature is also a function of the GBT elevation angle, so three angles are plotted. However, the 30 and 60 degree spill temperatures are virtually identical, and their points overlap. This calculation assumes a ground temperature of 293 K, and a ground emissivity of 1. Although the spill temperature rises at the upper edge of the band, recent RFI measurements at Green Bank have shown a large spike in RFI at that frequency range, so most data in the upper frequency ranged will have to be cut due to RFI contamination.

Parkes Intensity Mapping

The Parkes Radio Telescope is a 64 meter steerable dish located 20 km north of Parkes, New South Wales, Australia. As with the GBT survey, we conducted our observations on fields that overlap galaxy surveys: in this case the 2dF Galaxy Redshift Survey. We conducted observation during one week in late April and early May of 2014 using the 13 receiver array, focusing on two 2dF fields, which we will refer to as 2dF1 and 2dF2. During that time, 152 hours (1976 beam hours) of data on these two fields was collected. Our scan strategy was to maintain a constant elevation angle during the scans and to move the telescope back and forth in azimuth as the fields drifted through the sky due to the Earth's rotation. The constant elevation angle minimized changes in ground spill pickup. Scans of both fields were made as the fields were both rising and setting.

The frequency center of the observations was 1315.5 MHz. The observed bandwidth was 64 MHz, which translates to a redshift range of approximately 0.054 to 0.106. Although the mean redshift and the redshift range were both smaller at Parkes than for the GBT fields, the two Parkes fields are both very wide (roughly 100 degrees in Right Ascension and 8 degrees in Declination). The large angular size of the Parkes fields meant that, even after splitting the fields into a manageable number of slightly smaller maps, the inverse noise covariance matrices required to construct each map were approximately 2 terabytes each. However, 2 terabytes is a larger chunk of memory than is available in RAM on any of the nodes of the GPC supercomputing at the SciNet HPC Consortium, where the maps were made. One

4

could try to perform the necessary linear algebra operations on matrices stored in hard disks, but this would be extremely slow due to the long input/output time of permanent memory. The problem was solved by parallelization of the original GBT mapmaking code. Multiple nodes constructed pieces of the inverse noise covariance separately, entirely in RAM, and then each node wrote its piece in parallel to a single file. In the final mapmaking step, each node loaded a block diagonal chunk of the inverse noise covariance matrix and used it to calculate a piece of the final map.

After the maps were constructed, the next step was to remove the foregrounds caused by extragalactic and Milky Way synchrotron emission. Due to instrumental mode mixing, which is present in any real observation, the observed frequency structure of the foregrounds varied slightly with location on the sky in an unpredictable way, so that a scheme that simply subtracts a preset foreground template from each pixel would not work. Instead, we had to determine the foreground modes from the map. Since the foregrounds should be the dominant component in any 21 cm intensity map, our cleaning scheme calculated the foreground contaminated modes in each map using a principal component analysis of the map. For each map, we constructed a frequency-frequency covariance across the sample of all the pixels in the map. We then performed a singular value decomposition (SVD) on this covariance, which produced a list of orthogonal spectral modes ordered by their singular values. The modes with the largest singular values were considered strongly contaminated by foregrounds and projected out at each pixel of the map. Figure 2 shows one frequency slice of the 2df1 maps before foreground cleaning, and figure 3 shows the same slice of the maps after the 20 most contaminated foreground modes were removed.
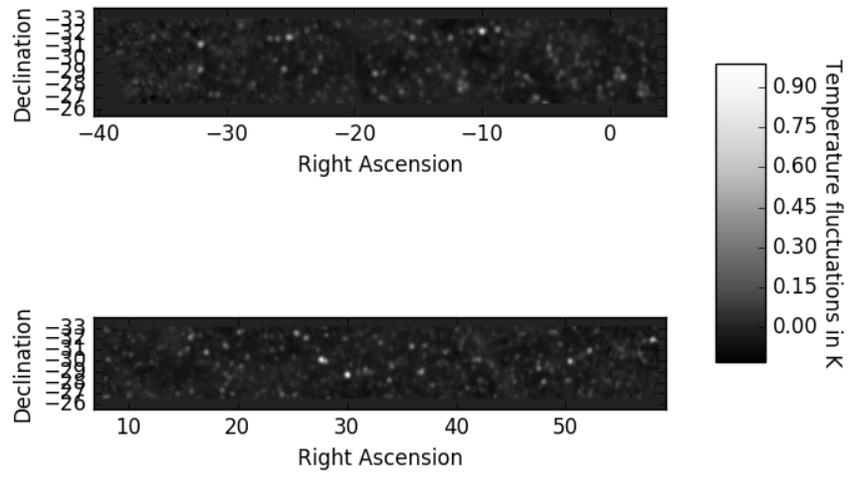
Figure 2: The 2df1 field at band center. The flux is dominated by continuous and point source synchrotron emission. Point sources are smoothed out due to convolution with the Parkes beam.
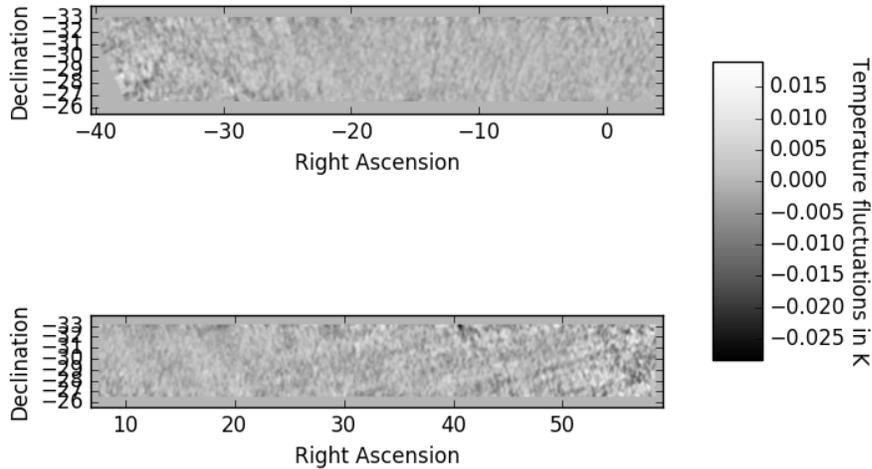
Figure 3: The 2df1 field at band center after 20 foreground modes were removed. Note the decrease in the scale compared to the uncleaned map. As with measurements at the Green Bank Telescope, cross-correlation with galaxy surveys show a statistically significant correlation because the HI is at the same positions as the galaxies, but the foreground residuals dominate over the HI signal if the radio maps are correlated with themselves.

The SVD foreground cleaning scheme inevitably removed 21 cm signal as well as foreground signal. There was a balance to strike in choosing the number of modes to remove. If too few modes were removed, there would be residual foregrounds that dominate the 21 cm signal. But if too many modes were removed, most of the 21 cm signal would be cut, and the error bars would be high. To calculate the appropriate number of modes to remove and the amount of signal loss, we performed the SVD foreground cleaning on simulated 21 cm signal. This was done as follows. First, we computed 100 simulated HI maps. These maps were then added to our real maps, and the sum of the maps was foreground cleaned using the SVD method described above. The foreground cleaned real map (without simulation) was then subtracted from this map. The remainder can be viewed as an estimate of the 21 cm signal left after foreground cleaning: call this the cleaned simulation map. By comparing the cross-power of the cleaned simulation map and the raw simulation map to the auto-power of the raw simulation map, we determined how much power was lost in each Fourier mode due to cleaning. For example, one might find that 75% of the signal remains on average at a particular Fourier mode. The function that quantifies this is called the transfer function, and it represents the average of this simulation and cleaning procedure, which we repeated 100 times to get a high signal to noise ratio. The standard deviation of the transfer function, calculated from the 100 simulation samples, was also used as an estimate of our error bars.

7

Maximizing the ratio of the transfer function to its standard deviation provided a rule of thumb for the optimal number of modes to remove.

Due to some difficulty in interpreting the results on small non-linear scales, our analysis of the Parkes maps is ongoing, but it is nearing completion. We expect to publish results before the end of the year.

Conclusion

21 cm intensity mapping of neutral hydrogen is a burgeoning field with a lot of promise. It has the potential, at low redshift, to constrain the dark energy equation of state. Using all redshifts available, 21 cm intensity mapping can map a larger volume of the universe than any other technique, due to its ability to see structure before the first stars formed (reionization). However, foreground removal is a major obstacle. Analysis of GBT data using Principal Component Analysis using SVD has showed promise: 21 cm signal was detected by cross-correlating with the WiggleZ galaxy survey. The commissioning of a new 7 element array array at the GBT will speed up the production of larger maps. This could lead to the first detection of 21 cm intensity in auto-correlation. Similar 21 cm intensity maps have been made with data from the Parkes telescope, and the analysis is almost complete.

References

[1] CHANG, T.-C., PEN, U.-L., BANDURA, K., AND PETERSON, J. B. Hydrogen 21-cm Intensity Mapping at redshift 0.8. *ArXiv e-prints* (July 2010).

[2] KOOI, P., LEONG, M., AND YEO, T. Dipole-Eexcited Short Backfire Antenna With Corrugated Rim. *Electronics Letters 15* (1979).

[3] LIU, A., AND TEGMARK, M. How well can we measure and understand foregrounds with 21 cm experiments? *Mon. Not. Roy. Astron. Soc. 419* (2012), 3491–3504.

[4] MORALES, M. F., AND WYITHE, J. S. B. Reionization and Cosmology with 21-cm Fluctuations. *48* (Sept. 2010), 127–171.

[5] PETERSON, J. B., ALEKSAN, R., ANSARI, R., BANDURA, K., BOND, D., BUNTON, J., CARLSON, K., CHANG, T.-C., DEJONGH, F., DOBBS, M., DODELSON, S., DARHMAOUI, H., GNEDIN, N., HALPERN, M., HOGAN, C., LE GOFF, J.-M., LIU, T. T., LEGROURI, A., LOEB, A., LOUDIYI, K., MAGNEVILLE, C., MARRINER, J., McGINNIS, D. P., McWILLIAMS, B., MONIEZ, M., PALANQUE-DELABRUILLE, N., PASQUINELLI, R. J., PEN, U.-L., RICH, J., SCARPINE, V., SEO, H.-J., SIGURDSON, K., SELJAK, U., STEBBINS, A., STEFFEN, J. H., STOUGHTON, C., TIMBIE, P. T., VALLINOTTO, A., AND TECHE, C. 21-cm Intensity Mapping. In *astro2010: The Astronomy and Astrophysics Decadal Survey* (2009), vol. 2010 of *ArXiv Astrophysics e-prints*.

[6] Switzer, E. R., et al. Determination of z 0.8 neutral hydrogen fluctuations using the 21 cm intensity mapping auto-correlation. *Mon. Not. Roy. Astron. Soc. 434* (2013), L46.