

The PAPER-32 power spectrum measured at several redshifts

Daniel C. Jacobs¹, Aaron R. Parsons^{2,8}, James E. Aguirre³, Zaki Ali², Judd Bowman¹, Richard F. Bradley^{4,5,6}, Chris L. Carilli⁷, David R. DeBoer⁸, Matthew R. Dexter⁸, Nicole E. Gugliucci⁵, Pat Klima⁵, Adrian Liu², Dave H. E. MacMahon⁸, Jason R. Manley⁹, David F. Moore³, Jonathan C. Pober², Irina I. Stefan¹⁰, William P. Walbrugh⁹

ABSTRACT

We present new observations from the Donald C. Backer Precision Array for Probing the Epoch of Reionization telescope which probe the redshift range of $6 < z < 9$, extending previously published single redshift results to cover the full range accessible to the instrument. The removal of foregrounds depends heavily on the spectral smoothness of both instrumental response and foregrounds. These spectra demonstrate a robust foreground removal of foreground signals to the thermal noise limit, leaving systematics on the scale of the noise. We find that these residual systematics are present throughout the observed frequency band are most likely some mixture residual foregrounds and instrumental systematics.

Subject headings: reionization

1. Introduction

The Epoch of Reionization, when the first stars ionized the pervasive cosmological Hydrogen in the last global phase change, is predicted to be observable in highly redshifted 21 cm radiation. The Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER, XXX)¹ is a low

¹School of Earth and Space Exploration, Arizona State U., Tempe, AZ

²Astronomy Dept., U. California, Berkeley, CA

³Dept. of Physics and Astronomy, U. Pennsylvania, Philadelphia, PA

⁴Dept. of Electrical and Computer Engineering, U. Virginia, Charlottesville, VA

⁵National Radio Astronomy Obs., Charlottesville, VA

⁶Dept. of Astronomy, U. Virginia, Charlottesville, VA

⁷National Radio Astronomy Obs., Socorro, NM

⁸Radio Astronomy Lab., U. California, Berkeley, CA

⁹Square Kilometer Array, South Africa Project, Cape Town, South Africa

¹⁰Cavendish Lab., Cambridge, UK

¹<http://eor.berkeley.edu>

frequency radio interferometer experiment dedicated to opening this window on the universe. Challenges include foregrounds which are brighter by several orders of magnitude and limited collecting areas of first generation instruments which necessitate long integration times. Direct observation of hydrogen before and during re-ionization is predicted to deliver a wealth of cosmological and astrophysical data, including the nature of the first stellar objects and the timing and rate of galaxy formation, reviews on the physics of reionization as well as theory about the nature of foregrounds maybe be found in Furlanetto et al. (2006); Morales & Wyithe (2010); Pritchard & Loeb (2012).

Telescopes seeking to measure this signal include the Giant Metre-wave Radio Telescope (GMRT; Paciga et al. (2013)), the Low Frequency Array (LOFAR; Yatawatta et al. (2013)) and the Murchison Widefield Array (MWA; Bowman et al. (2013) and Tingay et al. (2013)). PAPER is located in the Karoo desert at the Square Kilometer Array-South Africa cite. PAPER has doubled in size on a yearly basis since 2009, making observations with each array.

Here we report on deep integrations made with a 32 element array in 2011, first described in Parsons et al. (2013), hereafter P14. In that work we described in detail our reduction method which enabled a deep limit on the power spectrum in the presence of bright foregrounds. Using this method we combined 92 nights of data to place upper limits on the power spectrum of HI in the early universe. However, this analysis was limited to a single redshift. The epoch of reionization signal is expected to evolve strongly with redshift, and it is this variation with cosmic history that will allow us to begin to place constraints on the physics of reionization Pritchard & Loeb (2008); Pober et al. (2014). Therefore, while a detection of the 21 cm signal at even a single frequency would be a tremendous breakthrough, analysis techniques must be developed to capitalize on the wide bandwidths of the current generations of high redshift 21cm telescopes. The primary obstacle to detection, the wide bandwidth Using the same data as P14, this paper presents new power spectrum upper limits covering the redshift range $10.5 > z > 7.2$. In section 3 we summarize the observation and reduction methodology, present the new upper limits in 4 in 5 offer some conclusions and discussion of future work.

2. Observations and reduction

As the observations and methods in use are unchanged from those described in P14, we will offer a quick summary and refer the reader to Parsons et al. (2013) for a more in depth discussion. A general overview of the PAPER system can be found in Parsons et al. (2010), calibration of the primary beam in Pober et al. (2012), and imaging results in Jacobs et al. (2011, 2013a); Stefan et al. (2013). Sensitivity analysis described in Parsons et al. (2012) revealed that for the low gain elements employed by PAPER, a highly redundant “grid” type arrangement offers a significant sensitivity boost. In an interferometer each baseline samples a different mode of the power spectrum. In a grid, there are many samples of each of these cosmological modes which can be averaged to a high sensitivity before being combined with other different k modes. The PAPER South Africa 32 antenna deployment (PSA32) was arranged in a 4x8 grid. Here we include only the three shortest

types of spacings where the reionization power is expected to be brightest. This selection includes those between adjacent columns within at least one row of each other, a selection containing 70 30m baselines.

Observations spanning the band between 100 to 200-MHz ($13.1 > z > 6.1$) were recorded at a resolution of 48kHz and 10.7s beginning Dec 7, 2011 and ending March 19, 2012 for a total of 92 nights. Within this set we included observations in the LST range 1h - 9h where the sky dominated system temperature is at a minimum.

3. Reduction

Here we summarize our data reduction steps, for detailed descriptions see P14.

In several stages throughout the analysis process we take a 2D Fourier transform of the visibility spectra $V(t, \nu)$ into “delay’/fringe rate” space. In this space, smooth spectrum sources are physically localized to delays shorter than the light travel time length of the baseline (e.g. sources at the horizon, in the direction of the baseline vector) and fringe rates shorter than the sidereal rate. In this Fourier space sources are highly localized, deviations from a flat spectrum manifesting as a slight dispersion. In this space the spectrum sampling function, which is uneven due to flagging of interference takes the form of a convolution by a point-spread-function, much in the same way an imperfect sampling of the uv plane gives rise to the psf of an interferometer. To account for this we use a CLEAN like, iterative, peak finder and subtraction algorithm which is limited to finding peaks within the physically allowable ranges, and occasionally beyond, where necessary to account for spectral and temporal dispersion.

The data analysis pipeline essentially consists of iterative application of the delay/fringe rate transform process, with an ever tightening allowable number of modes, interleaved with stages of averaging (time, frequency, night), before finally computing a power spectrum. This final step takes advantage of the redundant baselines to make an unbiased power spectrum estimate by cross-multiplying identical baselines and then averaging the power spectrum modes.

3.1. Initial Averaging

First, the raw data are down-selected to just the 70 30m long baselines described in Section 2. The visibilities are then compressed in the frequency and time directions by filtering delay modes above the horizon and highest frequency fringe rates for a 300m baseline. This filtering is done in tandem with a radio frequency interference (RFI) flagging step, using the residuals which have had bright sky-like signals removed to flag 4σ deviations before feeding the flags back into another iteration of the compression step. This mitigates the effects of bright, narrow band, interference being scattered by the delay filter to generate a cleaner averaged data set and results in a time

and frequency bin of length 39.6s and width 492.61kHz. This process reduces the data volume by a factor of $\tilde{20}$, or roughly an order of magnitude improvement on traditional time and frequency averaging which in this array would be limited to 100kHz and 10s to avoid averaging away fringes.

3.2. Calibration

We model the gain as a per-antenna amplitude, a phase slope -physically a single time delay- and a single passband for all antennae. Because the array samples correlations redundantly, the relative calibration between antennae is numerically overdetermined and tractable as a linear algebra problem Liu et al. (2010). As described in P14, we compute the ratio between redundant baselines, fit for a gain and phase slope and then algebraically solve for a per-antenna solution. Here we have avoided calibrating each channel independently to preserve as much frequency variation as possible. These solutions vary little over the observing period, exhibiting less than 1% rms variation. A single solution derived for the Dec 7 data set is used for the entire observing run. Time and frequency variation of redundant solutions is explored in detail in Zheng et al (in prep).

Relative calibration is all that is necessary to form a power spectrum, however to set the flux scale, we must form a beam on a bright known point source. By itself, the redundant solution is completely independent of any sky model. The redundant solution contains two free delay parameters and an overall amplitude scale. We fit the two delay parameters to a model of Pictor A, Fornax A, and the Crab Nebula during a time when the sky is dominated by these three sources while marginalizing over the unknown apparent flux ratio between the three sources². With the delays in place we are now able form a beam on Pictor A and (for each channel) set the overall amplitude to the calibration value found in Jacobs et al. (2013b).

3.3. Foreground Filtering

Foregrounds are filtered from the calibrated data by removing all bright delay components with light travel times less than the baseline length. Where during the previous compression step a liberal horizon of 300m (1800ns, much longer than the 30 meter baselines under study) was used to calculate the window size, we now choose a window corresponding to the 30m baselines under study. The CLEAN model that is iteratively built is then subtracted from the data leaving residuals which we carry into the next stage. Next, a four hour long running mean is subtracted. This removes excess correlation due to cross-talk in the analog signal chain. The residuals are then flagged once more for RFI before the 92 nights of data are averaged into 36.4 second long local sidereal time (LST) bins. During averaging we found that some LST bins were dominated by very bright data points due to some exceedingly bright outlier samples. As these were very rare we found that

²Unknown in the sense of a joint uncertainty in source flux and primary beam pattern.

removing the 10% brightest from every bin, a small median filter, was quite effective. The source of these outliers is not known, a likely possibility is instability in the analog signal chain stimulated by weather or bright interference which have been observed in subsequent seasons. The root mean square of the residual signal (seen in Figure 2) at the end of this process is close to the 1mK level expected given the total integration time.

Though the frequency and repeated nightly observations have been averaged to their maximum extent, at 40s integrations the time axis has yet to be fully exploited. Sky-like fringes on a 30m baseline rotate much slower than 40 seconds. Performing a final fringe-rate filter, this time limiting to fringe-rates expected on a 30m baseline (down from 300m in the last iteration) we arrive at a data-set averaged to 900s, the maximum possible while still maintaining coherency.

3.4. Power Spectrum

The output of the forgoing is a single calibrated and foreground filtered sidereal night. The power spectrum is estimated in the delay spectrum of a 10MHz bandwidth range centered on the redshift of interest. To preserve the isolation of any foregrounds which remain, we increase the spectral range by 5MHz on each side and multiply by a Blackman-Harris window thus providing a much higher dynamic range delay spectrum point spread function.

This leaves us with 40 delay samples on each of 70 baselines which are divided into three redundant groups. Within these groups we compute the correlation between each mode. The cross correlation of the same delay modes between different redundant baselines provides our estimate of the power spectrum. Cross correlation between different delay modes provides our measure of systematic bias. Sky-like correlations should be identical between all redundant baselines to within the level of the noise. In practice some correlations deviate significantly from the median. These are removed by iteratively dividing the covariance into a model of systematics and a model of sky-like emission and then projecting out large residual modes. This is done by dividing the baselines into different groups such that all cross-multiplications are done without introducing noise bias. For more see Appendix of C of P14.

The residual correlations are all samples of the power spectrum. To estimate the final power spectrum and its uncertainty we compute the mean and variance of many random randomly-selected subsamples, sampling the dimensions: sidereal time, redundant baseline pair, and delay sign³

³As visibilities are complex, both the positive and negative delays carry information.

4. Results

4.1. Foreground Filtering and Noise Levels

Noise estimated by differencing adjacent frequencies (magenta) and between redundant baselines (blue), compared with an estimate of the theoretical noise level (dotted) indicating the rising significance of foregrounds as more data are added and sensitivity increases. The top three lines show noise after filtering foregrounds and binning into 40s long sidereal bins. At this noise level the frequency and baseline differences are roughly similar, both demonstrate the same small bumps of increased noise due to interference flagging. The bottom three lines show the noise level after integrating up to the maximum fringe rate of 914s. Here we see the interference bumps suppressed as more samples with different flagging are averaged together. We also see the difference between baselines become consistently larger than the difference between channels. In all cases the noise spectrum is characterized by an overall slope towards lower frequencies owing to the power law slope of the galactic spectrum which drives the system temperature. At the band edges noise rises dramatically, more so in the difference between baselines, then the spectral difference. This is, in part, due to the use of the Blackman-Harris window in the foreground filter is optimized for filtering foregrounds at the center of the band with minimal contamination of higher delay modes. We also note that the theoretical noise is consistently lower than the observed. This is most likely due to inaccuracies in the counting of data points included in the 92-day average, a key input into the theoretical noise limit. For this reason we estimate all error bars with a bootstrap method rather than by direct propagation. Correct propagation of error bars is a subject for future improvement.

In Figure 3 we show the power spectra at different redshift bins, and in Figure 4 we see power spectra at different k modes as a function of redshift. In Figure 3 we see the spherically averaged power spectrum plotted in P_k units (left) and the final estimate of Δ^2 , plotted with 2σ error bars. In the redshift 7.7 bin the foreground measurements (the inner 5 data points) have been included for comparison with P14.

Discussion points:

1. the redshift 7.7 bin compares well with the P14 result.
2. power spectrum at 10.3 is completely foreground contaminated. this appears to be consistent with the noise curves where the residual rms diverges rises quickly away from the minimum near the center of the band
3. Most of the data points lie above the theoretical noise curves. this because the noise model is a simple $\text{root}(2N_{\text{bt}})$ estimate.
4. we could get a redshift bin at 7 if we had a prescription for including non-uniform weights.

Limitations:

1. time dependence of cal solutions
2. missing data

5. Conclusions

Here we have extended the redshift coverage of Parsons et al. (2013) to cover the range $10.5 > z > 7.5$. These measurements demonstrate a precision removal of foregrounds to one part in ten million across a wide redshift. With only 32 antenna, this data lacks the sensitivity to exclude all but the most extreme models, however, it does demonstrate the ability of the wideband filtration method to affectively remove foregrounds to the precision needed to integrate a full season to the thermal noise limit. As future observations with PAPER will add sensitivity primarily by increasing to 128 antennae this projects well for our ability to reach design sensitivity.

REFERENCES

- Bowman, J. et al. 2013, Publications of the Astronomical Society of Australia, 30, 31
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, Physics Reports, 433, 181, elsevier B.V.
- Jacobs, D. et al. 2011, The Astrophysical Journal, 734, L34
- Jacobs, D. C. et al. 2013a, eprint arXiv, 1307, 7716, accepted to ApJ, includes two data files
- . 2013b, The Astrophysical Journal, 776, 108
- Liu, A., Tegmark, M., Morrison, S., Lutomirski, A., & Zaldarriaga, M. 2010, Monthly Notices of the Royal Astronomical Society, 408, 1029, (c) Journal compilation © 2010 RAS
- Morales, M. F. & Wyithe, J. S. B. 2010, Annual review of astronomy and astrophysics, 48, 127, oise
- Paciga, G. et al. 2013, Monthly Notices of the Royal Astronomical Society, -1, 1427
- Parsons, A. et al. 2010, The Astronomical Journal, 139, 1468
- Parsons, A., Pober, J., McQuinn, M., Jacobs, D., & Aguirre, J. 2012, ApJ, 753, 81
- Parsons, A. R. et al. 2013, arXiv, submitted to ApJ, 1304.4991v2
- Pober, J. et al. 2014, The Astrophysical Journal, 782, 66
- Pober, J. C. et al. 2012, The Astronomical Journal, 143, 53

- Pritchard, J. & Loeb, A. 2008, Physical Review D, 78, 103511
- Pritchard, J. R. & Loeb, A. 2012, Reports on Progress in Physics, 75, 6901
- Stefan, I. I. et al. 2013, Monthly Notices of the Royal Astronomical Society, 432, 1285
- Tingay, S. et al. 2013, Publications of the Astronomical Society of Australia, 30, 7
- Yatawatta, S. et al. 2013, Astronomy & Astrophysics, 550, 136

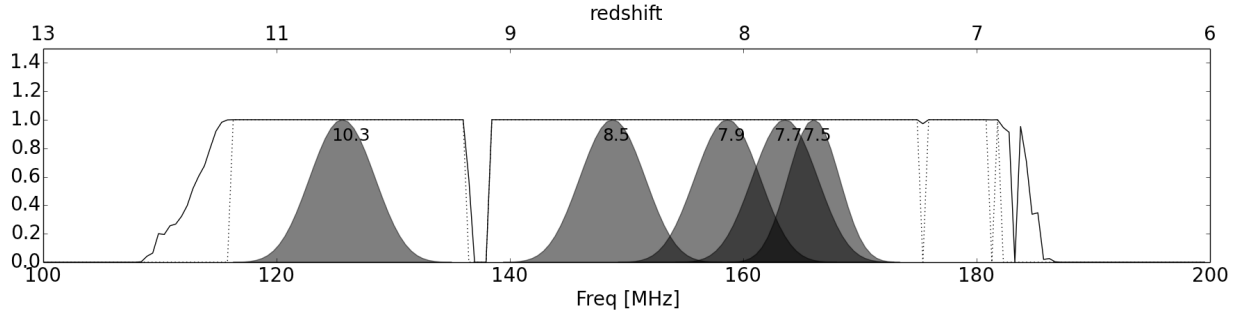


Fig. 1.— The average amount of data remaining after interference flagging over the 3 month period between Dec 2011 and March 2012 (black line) is quite high. Redshift bins (in grey, redshift center indicated with text label) are chosen to include spectral channels with uniform weight, i.e. no missing channels while maximizing coverage over the band. Redshift 7.7 is included for comparison with P14. Channels with no missing data are indicated by the dotted line. Each redshift bin is 20 MHz wide, but weighted by a Blackman-Harris window function which heavily down-weights the outer 10MHz for a Noise Equivalent Bandwidth of 10MHz. The interference is almost exclusively dominated by two features: ORBCOMM satellites at 137MHz and an unidentified intermittent line emitter at 175MHz. The roll off at 115MHz is due the rising noise at low frequencies being incorrectly flagged as interference.

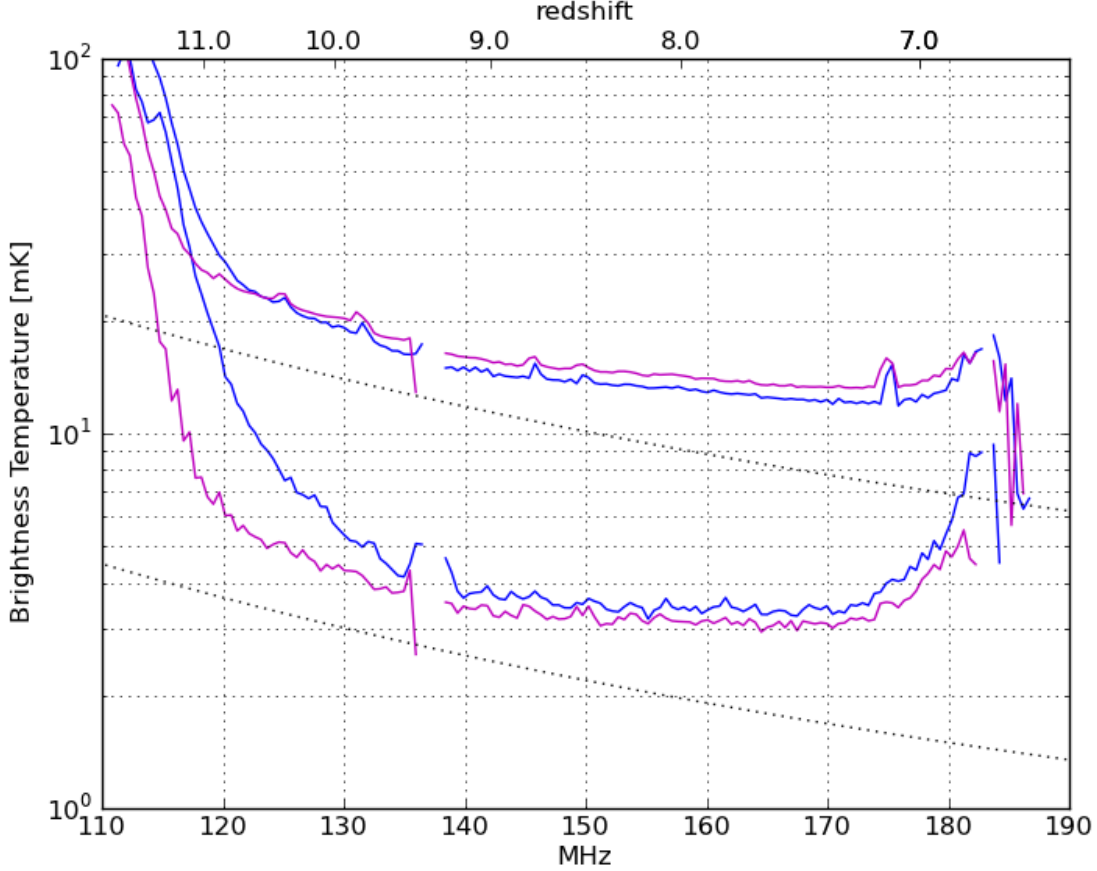


Fig. 2.— Noise estimated by differencing adjacent frequencies (magenta) and between redundant baselines (blue), compared with an estimate of the theoretical noise level (dotted) indicating the rising significance of foregrounds as more data are added and sensitivity increases. The top three lines show noise after filtering foregrounds and binning into 40s long sidereal bins. At this noise level the frequency and baseline differences are roughly similar, both demonstrate the same small bumps of increased noise due to interference flagging. The bottom three lines show the noise level after integrating up to the maximum fringe rate of 914s. Here we see the interference bumps suppressed as more samples with different flagging are averaged together. We also see the difference between baselines become consistently larger than the difference between channels. In all cases the noise spectrum is characterized by an overall slope towards lower frequencies owing to the power law slope of the galactic spectrum which drives the system temperature. At the band edges noise rises dramatically, more so in the difference between baselines, then the spectral difference. This is, in part, due to the use of the Blackman-Harris window in the foreground filter is optimized for filtering foregrounds at the center of the band with minimal contamination of higher delay modes. We also note that the theoretical noise is consistently lower than the observed. This is most likely due to inaccuracies in the counting of data points included in the 92-day average, a key input into the theoretical noise limit. For this reason we estimate all error bars with a bootstrap method rather than by direct propagation. Correct propagation of error bars is a subject for future improvement.

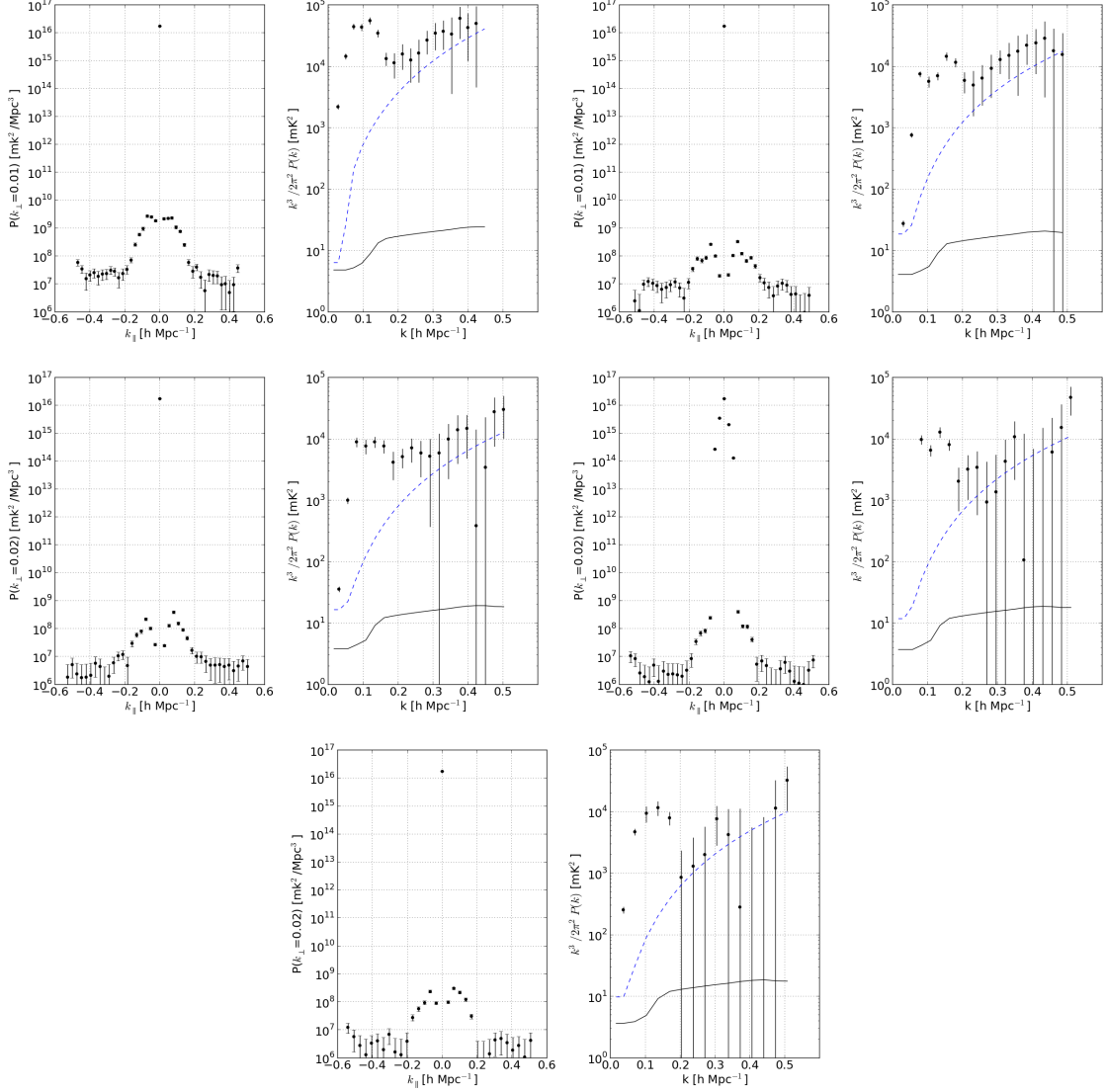


Fig. 3.— Left to right from top: power spectra centered on redshifts 10.3, 8.5, 7.9, 7.7 and 7.6 (frequencies: 127, 149.5, 159.5, 164.5, and 167 MHz) at an effective bandwidth of 10 MHz. Though the redshift 7.7 bin overlaps adjacent points it has provided for comparison against P14 which targeted this redshift with this data and reduction pipeline.

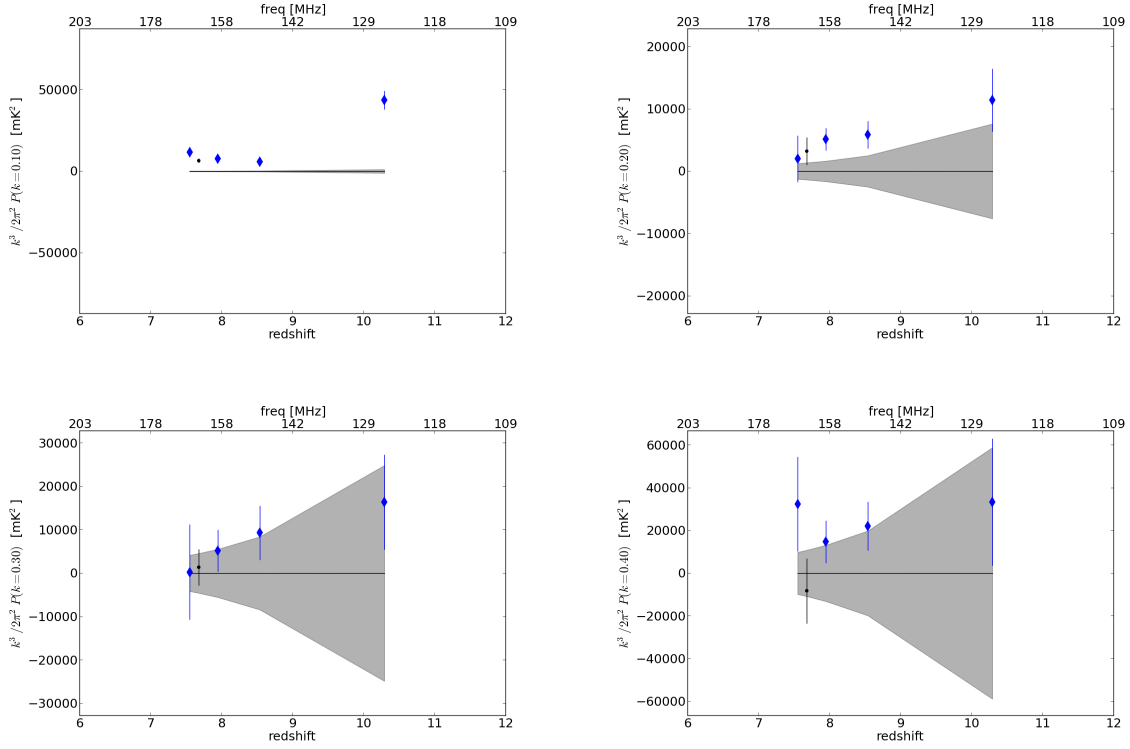


Fig. 4.— Power spectrum amplitude vs redshift at a selection of k modes. Right to left from top, $k=0.1, 0.2, 0.3, 0.4 \text{ hMpc}^{-1}$. Paciga 2013 GMRT marked with an 'x', Parsons 2014 PAPER limit marked with thin black, this work marked with thick blue diamonds. Points consistent with