HERA Blind Data Challenge Busy Week Summary
The Ohio State University
February 23-25, 2015

Executive Summary:

We convened a small group of people representing both visibility simulation codes for 21 cm arrays and 21 cm power spectrum estimation codes. Our principal goals are to facilitate the transfer of data between these codes and to test the performance of power spectrum analysis pipelines. We have chosen the uvfits file format as the standard for interferometric visibilities; short-term development will focus on developing uvfits outputs from the simulation codes that can be read by all analysis codes. This document concludes with plans for an initial non-blind data analysis challenge, as well as a longer-term blind data analysis challenge.

Attendees: Josh Dillon, Bryna Hazelton, Danny Jacobs, Lindley Lentati, Jonathan Pober, Peter Sims, Paul Sutter

Current Project Status

Represented at the meeting were 3 simulation pipelines and 6 power spectrum analysis pipelines. Other simulators and analyzers are more than welcome to join this effort. The proposed next steps should be relatively generic.

The Simulation Codes

FHD: The UW developed FHD package serves principally as a data analysis tool; to calibrate and subtract foregrounds, however, FHD performs a full forward-modeling of the visibilities, which can serve as a simulation input to analysis pipelines. In its current state, the simulations can include a foreground model (point sources and diffuse, including polarization effects), an EoR signal (either following a statistical power spectrum or from a full simulation with non-Gaussian bubbles), and noise. At the present, all foregrounds are restricted to have a single spectral index. One of FHD's strengths is its ability to simulate a realistic instrument with frequency and antenna dependent beams and full polarization.

Cambridge: The Cambridge simulation package produces visibilities for arbitrary telescope arrays containing a foreground model (point sources and diffuse), an EoR signal (from 21cmFAST simulations), and noise. Currently, the point sources simulated are not tied to actual sky models, but follow a statistical distribution. These simulations are powerful in their flexibility: they can incorporate frequency-dependent beams of arbitrary shape, and can simulate foregrounds with a range of different spectral indices (or even spectral turnovers).

ASU: The ASU simulation pipeline produces visibilities using the delay spectrum technique. As such, it natively simulates the full sky, horizon-to-horizon. Currently the simulations can include foreground emission (both point sources and diffuse, with arbitrary spectral indices, including polarization effects) as well as noise. Adding an EoR signal to the simulated visibilities is a work in progress.

The Analysis Codes

εppsilon: The εppsilon power spectrum analysis code by Bryna Hazelton runs on the output HEALPix images of FHD; it also uses meta-data about the gridded uv coverage (referred to as "weights" and "variances" cubes) to do a variance propagation from the measured visibility noise to power spectrum error bars. It performs no foreground subtraction, but rather takes foreground subtracted image cubes as inputs. To avoid a noise bias, εppsilon uses separate even/odd images from alternating integrations to estimate the power spectrum. Currently, a small normalization offset remains; the cause is known, but an implementable solution remains a work in progress.

Delay Spectrum: The delay spectrum analysis used on PAPER data requires redundant baselines to perform its analysis. The principal inputs are delay-filtered, LST-binned visibilities. The delay spectrum pipeline performs inverse covariance weighting to downweight foreground contaminated spectral eigenmodes, but no actual foreground subtraction after the delay filter is performed. To avoid a noise bias, the current implementation of the pipeline separates even/odd interleaved days in the LST binner, and cross-multiplies between these two sets.

Empirical Covariance: The empirical covariance pipeline created by Josh Dillon runs out the outputs of FHD (foreground subtracted HEALPix image cubes as well as weights and variances). It also uses the even/odd integration split to avoid a noise bias in the final power spectrum. Although the input is run on foreground-subtracted visibilities, the empirically estimated foreground covariance provides strong suppression of residual foreground emission in the EoR window.

Precision Mapmaking: This pipeline currently in development by Josh Dillon will use the precision/optimal map-making analysis outlined in his recent paper to estimate the 21 cm power spectrum. Desired inputs are LST-binned, even/odd split, foreground subtracted visibilities (preferably redundant, as the pipeline is being designed for HERA). A catalog of foreground point sources will also be included in the calculation.

Cambridge: The Bayesian analysis pipeline developed at Cambridge runs on visibility data. No prior foreground subtraction is necessary. Currently, it requires a model of the thermal noise to avoid a noise bias, but work is in progress to allow for empirical noise estimation from even/odd split visibilities. A paper describing this approach is currently under review with MNRAS (but not available on astro-ph).

Blackjack: The Blackjack pipeline being developed by Paul Sutter will also employ Bayesian methods to estimate the 21 cm power spectrum. It currently takes images as input, but work is in progress to allow for visibility inputs. No prior foreground removal is necessary. Blackjack also requires a model of the thermal noise to avoid a bias, but empirical noise estimation from even/odd split data is under investigation.

Short Term Goals

In the next six months, the principal goal is to develop interchanges between all of the simulation outputs and all of the analysis pipelines. The file format chosen for interchange is uvfits. While no exact uvfits standard has been defined in the literature, we are choosing a pragmatic approach. The owners of simulation codes will need to improve/alter their uvfits outputs as the needs of the analysis codes become clearer. This approach will require good communication among those involved, but at present it seems more flexible than defining an arbitrary standard when specific needs are still somewhat unknown.

For those analysis codes using images as inputs (as opposed to visibilities), FHD has been identified as the fiducial imager. (Two of the three image-based analyses already use FHD outputs.) The group will work to support Python tools for extracting HEALPix maps from FHD IDL save file outputs.

Three specific short-term tasks break down as follows:

- Simulation codes work towards outputting uvfits visibilities (or development of a converter to take simulation outputs to uvfits). To allow for the most flexible testing of analysis codes, simulators should work to out *separate visibility sets* for foregrounds, EoR signal, and noise (which can then be added together as needed).

- Analysis codes work towards reading uvfits visibilities (or FHD HEALPix maps). While the onus is on the simulation code owners to make uvfits with the necessary information, the onus is on the analysis code developers to *communicate* their needs to the simulators as they become clear. Once a robust link between two codes is established, it should be *documented*.

- As possible, images/imagers should be compared. While FHD has been chosen as the fiducial imager in the short term, codes like the Cambridge analysis pipeline internally image visibilities. Work to output these intermediary images for comparison with other images will be valuable.

Data Challenge 1 (Non-Blind Analysis)

The first data analysis challenge will be a non-blind analysis of a chosen set of input parameters. As such, simulation code developers should work on their codes to take arbitrary (or at least flexible) inputs for primary beams, foreground models, EoR models, and instrument parameters (e.g. antenna positions, time integration, frequency resolution, etc.). The instrument being simulated will almost certainly be HERA. Once sufficient progress has been made, we will decide upon a fiducial set of simulation inputs. Analysis of these simulations will proceed on three fronts: visibilities (i.e. direct comparison of simulation outputs), images (where computed), and power spectra. A successful analysis code should be able to recover the power spectrum from any simulation; reasons for failure will be highly informative, especially if the analysis recovers the power spectrum from some simulations but not others. Iteration across several sets of simulation parameters will almost certainly be necessary (i.e. Data Challenges 2 through N).

Blind Data Challenge 1

Once the results of the non-blind data challenge(s) is/are understood, a first blind challenge should take place. The simulators should output a summed visibility set containing EoR, noise, and foregrounds. Exact details of the foregrounds will be decided before the challenge (specifically: how much of the input model is “known” to the analysis pipelines?). Instruments will still be assumed to be perfectly known (i.e. no calibration or other errors will be introduced).

Future Directions

As a road map for analysis and simulation code developers, we conclude with a list of complications/systematics/errors that will undoubtedly be present in real data. As this project moves forward, we will look into introducing these complications into the simulations:

- Foreground model errors
- Calibration errors
- Beam model errors
- Antenna-to-antenna variations
- Antenna position errors
- Cross talk
- RFI/Flagging
- Polarization