Confronting the challenges of global EoR detection

Keith Tauscher[1,2], David Rapetti[1,3], and Jack O. Burns[1]

[1] Center for Astrophysics and Space Astronomy
University of Colorado, Boulder, CO, USA
[2] Department of Physics
University of Colorado, Boulder, CO, USA
[3] NASA Ames Research Center, Mountain View, CA, USA

The highly redshifted, sky-averaged 21-cm spectrum from neutral hydrogen, known as the 21-cm global signal, is an information-rich experimental target in low-frequency (10-200 MHz) radio astronomy. It contains information about when the first stars ignited and how they compare to stars seen today as well as properties of the first X-ray sources that heated up the universe's hydrogen gas and the evolution of the reionization of that gas. Especially at lower frequencies, it even serves as a powerful probe of exotic physics. However, rigorous analysis of data from experiments attempting to measure the global signal is very difficult. In this presentation, we aim to convey the challenges inherent to the analysis problem as well as the advantages and disadvantages of common techniques used to mitigate them.

The main problem that must be solved is one of separation of the signal from systematic effects like foreground emission and instrumental biases. The foreground, being 4-5 orders of magnitude larger than the expected 21-cm signal, must be modeled extremely accurately in order for a simultaneous fit for parameters of a 21-cm signal model to be useful. We discuss alternatives to the common use of generic polynomials to model these foregrounds, which, though convenient in the lack of precise knowledge it requires, is not fully justified, may be unrealistic. It is also unclear how to choose the number of terms to include in the polynomial. We explore the technique of frequency binning, including when it is useful in the 21-cm context and the advantages and disadvantages of some individual methods. Because most systematic data components are smooth at the MHz scale like the 21-cm signal, there is little information with which to separate them in a single spectrum. We show that binning measured spectra in time allows for the simultaneous analysis of multiple spectra, which, while more computationally intensive, leads to more robust constraints than can be gleaned from a single spectrum. We are implementing this LST-binned approach in our independent analysis of EDGES data.

After outlining a complete analysis pipeline (based on elements of the `pylinex` code) which takes advantage of the above knowledge to find robust parameter fits, we stress that goodness-of-fit statistics are imperative in evaluating such claims about features seen in the presence of large, smooth, non-signal data components such as foreground emission. In doing so, we present a new goodness-of-fit statistic specifically crafted to detect persistent, wide-band spectral structure in residuals at or near the expected noise level. The statistic, which we refer to as psi-squared, is complementary to the traditional chi-squared statistic and is based on correlations between channels.