HYDROGEN COSMOLOGY FROM THE DEEP SPACE GATEWAY: DATA ANALYSIS PIPELINE FOR LOW-FREQUENCY RADIO TELESCOPES.

D. Rapetti[1,2], K. Tauscher[1,3], J. O. Burns[1], E. Switzer[4], J. Mirocha[5], S. Furlanetto[5], and R. Monsalve[1], [1]Center for Astrophysics and Space Astronomy, Department of Astrophysical and Planetary Science, University of Colorado, Boulder, CO 80309, USA, David.Rapetti@colorado.edu, [2]NASA Ames Research Center, Moffett Field, CA 94035, USA, [3]Department of Physics, University of Colorado, Boulder, CO 80309, USA, Keith.Tauscher@colorado.edu, Jack.Burns@colorado.edu, [4]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA, Eric.R.Switzer@nasa.gov, [5]Department of Physics and Astronomy, University of California at Los Angeles, Los Angeles, CA 90095, USA, mirocha@astro.ucla.edu, sfurlane@astro.ucla.edu, Raul.Monsalve@colorado.edu.

Introduction: This is a companion abstract to two others presenting proposals for low-frequency radio telescopes. One is to be bolted on the NASA's planned Deep Space Gateway (DSG), and therefore on lunar orbit (Tauscher et al.), and the other to be deployed from the DSG to the surface of the Moon (Monsalve et al.). The goal of these instruments is to take advantage of the radio quiet environs above the lunar farside at night, with no ionospheric effects and hidden from Earth's radio frequency interference (RFI) and Solar radio emissions. All this is critical to not only detect, for the first time, the hyperfine line of neutral hydrogen (HI) produced, during Cosmic Dawn and Reionization, by the first stars, galaxies and black holes, but also to precisely measure this signal, rich in astrophysical and cosmological information on these early epochs of the Universe.

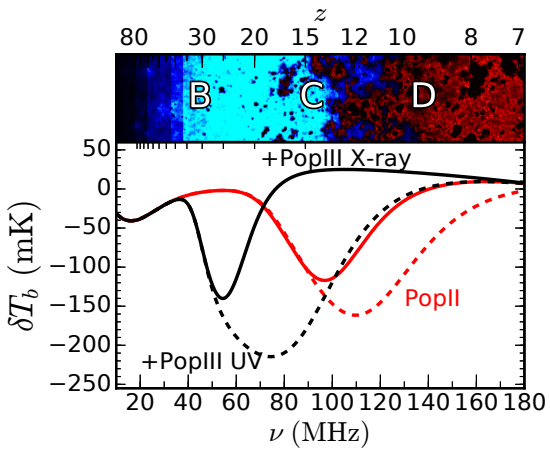

Figure1: The upper panel shows the evolution of a slice of the Universe from early (left) to late times (right), and the lower panel, different models of the global 21-cm spectrum relative to the CMB temperature. The red lines are conservative models with metal-rich stars (Pop II), while the black curves assume that metal-free stars (Pop III) also occur, but only in low-mass galaxies where atomic cooling is inefficient. The dashed and solid curves differ in specific emission and stellar properties (see Burns et al. 2017 [3] for details). The epochs B, C and D correspond to the ignition of the first stars, the initial accretion of black holes, and the onset of reionization, respectively. Figure from Burns et al. (2017) [3], adapted in turn from Pritchard & Loeb (2010) [4] using the new models from Mirocha et al. (2017) [2].

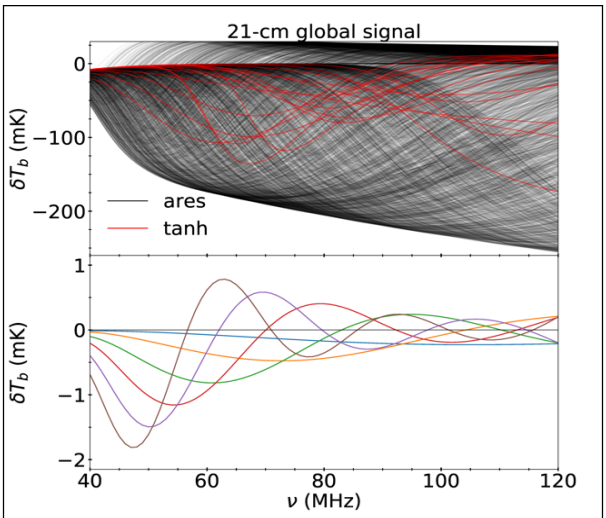

Figure 2: Upper panel: thinned sample (black curves) of the signal training set used in the analysis of Tauscher et al. (2017; hereafter T17) [5] and generated with the *ares* code (https://bitbucket.org/mirochaj/ares) for reasonable bounds of the physical parameters. The red signals are taken instead from the tanh signal modelling of Harker et al. (2016) [6]. Lower panel: the first 6 SVD basis functions (from most to least important: blue, orange, green, red, purple, brown) obtained from the *ares* training set. Figure from T17. See there further details on the calculation and utilization of SVD modes within our pipeline.

The multi-wavelength radiation emitted by these first luminous objects shifted the spin-flip temperature of the intergalactic medium (IGM) gas, and therefore the strength of the 21-cm line of HI [1, 2], with respect to that of the Cosmic Microwave Background (CMB). Fig. 1 shows characteristic shapes for this sky-averaged signal, depending on factors such as the metallicity of the first stars and the emission efficiency of the sources.

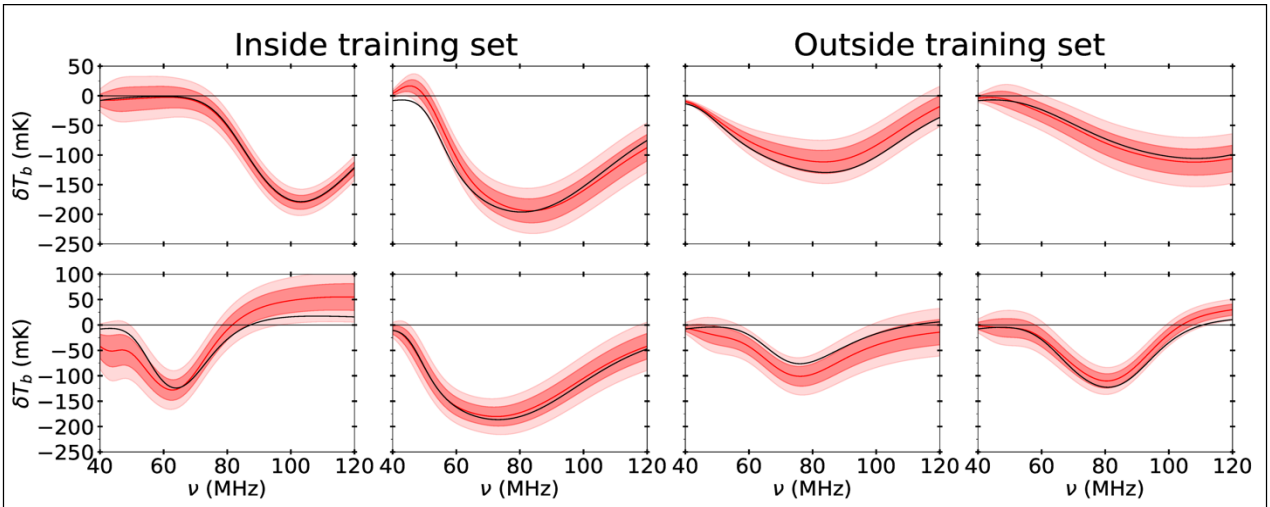

Figure 3: Representative signal extractions using the first step of our data analysis pipeline via the *pylinex* code (https://bitbucket.org/ktausch/pylinex). The latter is able to obtain such results by using a linear model defined by SVD eigenmodes calculated from signal (see, in this case, Fig. 2) and systematics (see Fig. 1 in T17) training sets. The black curves show the input signals, the red curves the signal estimates, and the dark (light) red bands represent the posterior 68% (95%) confidence regions (see further details in T17). For all plots, the input beam-weighted foregrounds (systematics) came from the training sets in Fig. 1 of T17. The input signals for the four plots on the left came from the signal training set in Fig. 2, whereas the input signals for those on the right were generated by the tanh model [6]. Even though the difference between these two sets is very small by eye, as shown in Fig. 5 of T17, on average these are different. Figure from T17.

Signal extraction: Given the large variety of theoretical models available (see e.g. Fig. 1) and their corresponding, mostly unconstrained physical parameter spaces, we employ a well-known pattern recognition technique, in combination with information criteria, to extract the signal from systematics (large beam-weighted foregrounds and residual instrument calibration) in frequency channel space.

The use of this machine learning algorithm, Singular Value Decomposition (SVD), allows us to separately characterize the signal and systematics with distinct training sets. Fig. 2 shows for example the signal training set utilized in T17 to demonstrate the ability of our pipeline to quickly constrain very different input signals, as shown in Fig. 3.

Overall data pipeline: After extracting the signal in the first step of the pipeline, in the second step we perform a multidimensional interpolation in order to translate that signal from frequency channel space into a physical parameter space. Once this is achieved, this provides us with a starting point from which we can efficiently explore the full probability distribution in this space, within a Bayesian framework, through a Markov Chain Monte Carlo (MCMC) analysis (Rapetti et al., in preparation).

Summary: The DSG will therefore provide a unique opportunity to operate low-frequency radio telescopes shielded by the Moon, either on orbit or the surface, and utilize our state-of-the-art pipeline for precision hydrogen cosmology.

Acknowledgements: This work was directly supported by the NASA Solar System Exploration Virtual Institute cooperative agreement 80ARC017M0006. DR is supported by a NASA Postdoctoral Program Senior Fellowship at NASA's Ames Research Center, administered by the Universities Space Research Association under contract with NASA. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center through the award SMD-16-7501.

References:

[1] Madau P., Meiksin A. and Rees M. J. (1997) *ApJ*, 475, 42. [2] Mirocha J., Furlanetto S. R. and Sun G. (2017) *MNRAS*, 464, 1365. [3] Burns J. O., Bradley R., Tauscher K., et al. (2017) *ApJ*, 844, 33. [4] Pritchard J. R. and Loeb A. (2010) *PhRvD*, 82, 023006. [5] Tauscher K., Rapetti D., Burns J. O. and Switzer E. R. (2017), T17, *pre-print arXiv*:1711.03173. [6] Harker G. J. A., Mirocha J., Burns J. O. and Pritchard J. R. (2016) *MNRAS*, 455, 3829.