Hydrogen Cosmology Data Analysis Pipeline for Lunar-based Observations.

D. Rapetti[1,2], K. Tauscher[1,3], J. O. Burns[1], J. Mirocha[4], N. Bassett[1], J. Merten[5], [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, [4]Department of Physics & McGill Space Institute, McGill University, Montreal, QC H3A 2T8, Canada, [5]INAF, Bologna, 40129, Italy.

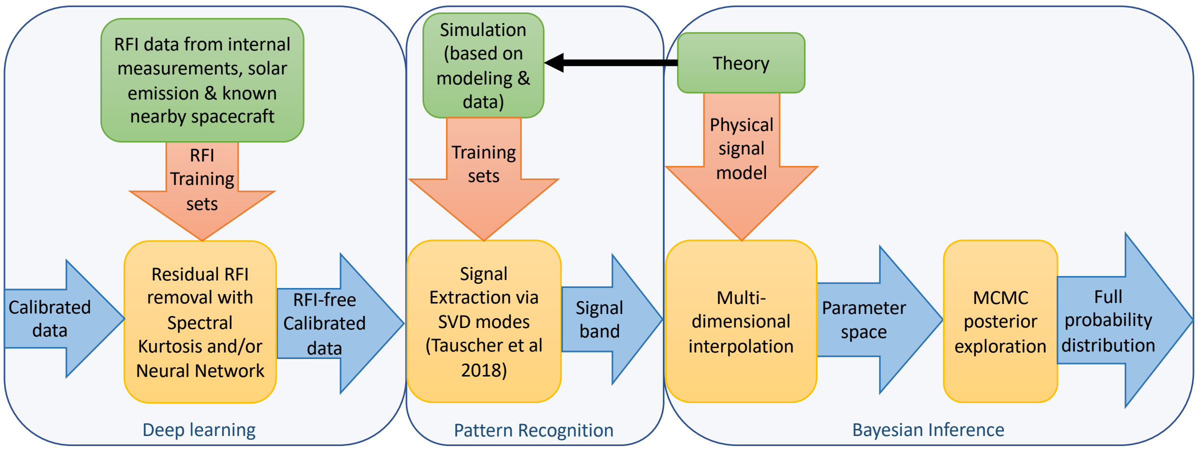

Figure 1: Flow diagram of the DAPPER data analysis pipeline, which processes the input calibrated data in various steps until obtaining the posterior probability distribution of the full parameter space.

For an experiment orbiting the Moon, we are developing a complete pipeline (Fig. 1) to analyze low radio frequency measurements of neutral hydrogen (HI) at the end of the Dark Ages and during Cosmic Dawn. This SmallSat mission concept, named the Dark Ages Polarimeter PathfindER (DAPPER), is planned to observe within a radio quiet cone above the lunar farside in order to avoid radio frequency interference (RFI) from Earth, and the contaminating effects and cutoff at the lowest frequencies from its ionosphere.

For internally generated RFI, DAPPER incorporates suppression and control strategies based on hardware. Residual internal RFI, and contamination from solar emissions and nearby spacecraft, such as the Lunar Gateway, will be removed in the first step of the pipeline, which implements a standard spectral kurtosis (SK) estimator, and a deep learning methodology (Fig. 2) we are designing to prevent unnecessary data excision by the SK technique, when adequate training sets are available from lab measurements and observations.

The next steps of the pipeline employ the *pylinex*[1] code (Tauscher et al 2018, ApJ, 853, 187). This starts by performing a novel pattern recognition technique that combines Singular Value Decomposition (SVD) with information criteria to extract the HI signal from large beam-weighted foregrounds, by leveraging DAPPER's measurements of the polarization induced by their anisotropy versus the isotropy of the sky-averaged signal as projected onto a large beam, and instrument

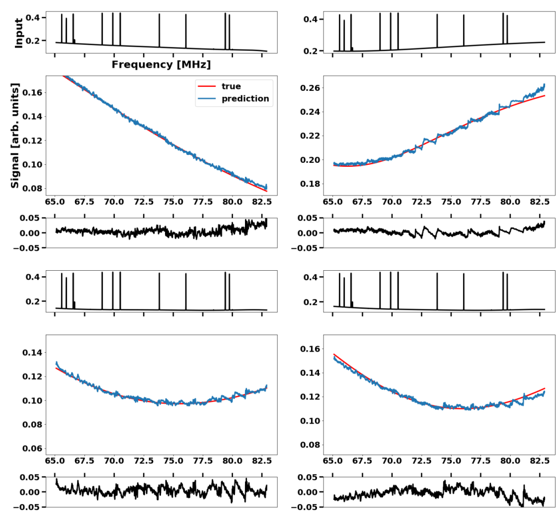

Figure 2: Four successful examples of employing a denoising encoder/decoder neural network to extract (blue curves) numerically injected signals (red curves) and remove at the same time the 10 simulated RFI frequencies shown in the time-step averaged input spectra of the top panels. The bottom panels show the relative differences between the two curves in the corresponding middle panels.

systematics. Then, this initial HI constraint in frequency space is transformed into a chosen physical parameter space via multidimensional interpolation, from which we finally infer the full probability distribution with a Markov Chain Monte Carlo (MCMC) algorithm.

This work was directly supported by the NASA Solar System Exploration Virtual Institute Cooperative Agreement 80ARCC017M0006.

[1]https://bitbucket.org/ktausch/pylinex