





# 1 BayesEoR: Bayesian 21-cm Power Spectrum 2 Estimation from Interferometric Visibilities

3 Peter H. Sims <sup>1\*</sup>, Jacob Burba <sup>2\*</sup><sup>¶</sup>, and Jonathan C. Pober <sup>3\*</sup>

4 <sup>1</sup> School of Earth and Space Exploration, Arizona State University, USA <sup>2</sup> Department of Physics and  
5 Astronomy, University of Manchester, UK <sup>3</sup> Department of Physics, Brown University, USA <sup>¶</sup>  
6 Corresponding author \* These authors contributed equally.

DOI: [10.xxxxx/draft](https://doi.org/10.xxxxx/draft)

## Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Open Journals](#) 

## Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

## License

Authors of papers retain copyright  
and release the work under a  
Creative Commons Attribution 4.0  
International License ([CC BY 4.0](#)).

## 7 Statement of need

8 Neutral hydrogen can undergo a spin-flip transition in which the quantum spins of the proton  
9 and electron transition from an aligned to an anti-aligned state, or vice versa, resulting in  
10 emission or absorption of a photon with a wavelength of 21-cm. The hydrogen 21-cm spin  
11 temperature quantifies the relative number densities of atoms in the aligned and anti-aligned  
12 states. Interferometric 21-cm cosmology experiments aim to measure the contrast between  
13 the 21-cm spin temperature of neutral hydrogen and the radio background temperature in the  
14 early Universe. By observing this signal at high redshift, we can learn a wealth of information  
15 about the state of the intergalactic medium during the first billion years of cosmic history.  
16 This information can, in turn, be used to infer properties of the first stars and galaxies that  
17 transformed the hydrogen intergalactic medium from a cold neutral gas to a hot ionised plasma  
18 during the Epoch of Reionization (EoR). Modern interferometers like [HERA](#), [LOFAR](#), and the  
19 [MWA](#), have been designed to observe with many antennas simultaneously to maximize their  
20 sensitivity to the 21-cm signal from the EoR. These experiments have shown that detecting  
21 this signal is rife with difficulty ([Abdurashidova et al., 2022](#); [Mertens et al., 2020](#); [Trott et al.,  
22 2020](#)). This is primarily due to the coupling of bright contaminating sources between us and  
23 the cosmological signal, referred to as “foregrounds”, with the spectral structure imparted by  
24 the instrument. Existing approaches to recovering the 21-cm signal from the data lack direct  
25 modelling of the observed covariance between the 21-cm and foreground signals in the data.  
26 Intrinsically, the 21-cm and foreground signals are uncorrelated. The instrument modulates  
27 both signals identically during observation, however, making them covariant. This covariance  
28 can be accounted for by forward modelling both signals, a key advantage of our approach in  
29 BayesEoR.

## 30 Summary

31 BayesEoR is a GPU-accelerated, Python-based implementation of a Bayesian framework de-  
32 signed to jointly model the 21-cm and foreground signals and forward model the instrument  
33 with which these signals are observed. Using these combined techniques, we can overcome  
34 the aforementioned difficulties associated with extracting a faint, background signal in the  
35 presence of bright foregrounds. BayesEoR enables one to sample directly from the marginal  
36 posterior distribution of the power spectrum of the underlying 21-cm signal in interferometric  
37 data, enabling recovery of statistically robust and unbiased<sup>1</sup> estimates of the 21-cm power  
38 spectrum and its uncertainties ([Burba et al., 2023](#); [Sims et al., 2016, 2019](#); [Sims & Pober,](#)

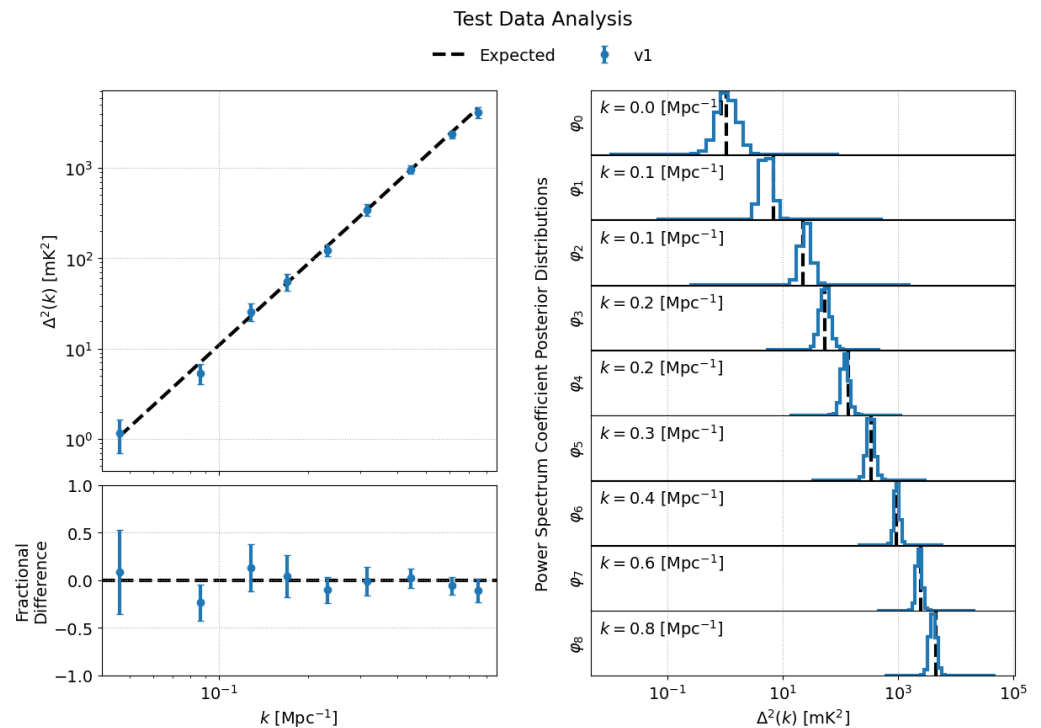
<sup>1</sup>Recovery of unbiased estimates of the 21-cm power spectrum requires that the field of view of the foreground model encompasses the region of sky from which instrument-weighted foregrounds contribute significantly to the observed data ([Burba et al., 2023](#)).

39 [2019](#)). The outputs of an analysis using the [test dataset](#) and [plotting code](#) provided with  
40 BayesEoR can be found in [Figure 1](#).

41 Running a BayesEoR analysis requires an input dataset, a model of the instrument, and set of  
42 analysis parameters. A script is provided with BayesEoR for convenience which pre-processes  
43 a pyuvdata-compatible dataset ([Hazelton et al., 2017](#)) of visibilities per baseline, time, and  
44 frequency into a one-dimensional data vector, the required form of the input dataset to  
45 BayesEoR. As part of the inference, BayesEoR forward models the instrument which requires  
46 an instrument model containing the primary beam response of a “baseline” (pair of antennas)  
47 and the “uv sampling” (the length and orientation of each baseline in the data). The primary  
48 beam response is passed via a configuration file or command line argument. The uv sampling is  
49 generated by the aforementioned convenience script to ensure that the ordering of the baseline  
50 in the instrument model matches that in the visibility data vector. Analysis parameters must  
51 also be set by the user to specify file paths to input and output data products and model  
52 parameters used to construct the data model (e.g. the number of frequencies and times in  
53 the data, the number of voxels in the model Fourier domain cube, the field of view of the  
54 sky model). Note however that these analysis parameters must be chosen carefully based on  
55 the data to be analyzed (please see section 2.3 of [Burba et al. \(2023\)](#) for more details on  
56 choosing model parameters). Accordingly, because BayesEoR forward models the instrument,  
57 we generate a model of the sky as part of our model visibilities calculation. When the EoR and  
58 foregrounds can be adequately described by a sky model with a field of view equal to the width  
59 of the primary beam, the memory requirements for a BayesEoR analysis are on the order of 10  
60 GB. This is the case for the provided test dataset for which the disk and RAM requirements  
61 are ~17 GB and ~12 GB, respectively. The field of view for the EoR and foreground sky models  
62 can be set independently, however, which allows for the EoR to be modelled within the primary  
63 field of view of the telescope while the foregrounds can be modelled across the whole sky. We  
64 wish to note however that, in this fashion, modelling the whole sky can be computationally  
65 demanding depending upon the data being analyzed. For example, in [Burba et al. \(2023\)](#)  
66 we show that analyzing a relatively modest dataset (compared to those typically analyzed by  
67 HERA, LoFAR, or the MWA) can require ~400 GB of RAM<sup>2</sup>. Please see section 6.1 of [Burba](#)  
68 [et al. \(2023\)](#) for more details.

---

<sup>2</sup>BayesEoR uses a routine from the Matrix Algebra for GPU and Multicore Architectures ([MAGMA](#)) library which allows for the use of matrices with a memory footprint larger than the available RAM on a GPU.



**Figure 1:** Example outputs from a BayesEoR analysis of the provided test dataset with a known power spectrum. The top left subplot shows the inferred dimensionless power spectrum ( $\Delta^2(k)$ ) as a function of spherically-averaged Fourier mode, ( $k$ ) estimates in blue and the expected power spectrum as the black, dashed line. The bottom left subplot shows the fractional difference between the recovered and expected power spectra. The right subplots show the posterior distribution of each power spectrum coefficient ( $\varphi_i$  where  $i$  indexes the spherically-averaged  $k$  bins) in blue with the expected power spectrum amplitude as the black, vertical, dashed lines.

## 69 Acknowledgements

70 The authors acknowledge support from NSF Awards 1636646 and 1907777, as well as Brown  
71 University's Richard B. Salomon Faculty Research Award Fund. JB also acknowledges support  
72 from a NASA RI Space Grant Graduate Fellowship. PHS was supported in part by a McGill  
73 Space Institute fellowship and funding from the Canada 150 Research Chairs Program. This  
74 result is part of a project that has received funding from the European Research Council  
75 (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant  
76 agreement No. 948764; JTB). This research was conducted using computational resources and  
77 services at the Center for Computation and Visualization, Brown University.

## 78 References

- 79 Abdurashidova, Z., Aguirre, J. E., Alexander, P., Ali, Z. S., Balfour, Y., Beardsley, A. P.,  
80 Bernardi, G., Billings, T. S., Bowman, J. D., Bradley, R. F., Bull, P., Burba, J., Carey, S.,  
81 Carilli, C. L., Cheng, C., DeBoer, D. R., Dexter, M., de Lera Acedo, E., Dibblee-Barkman,  
82 T., ... HERA Collaboration. (2022). First Results from HERA Phase I: Upper Limits on the  
83 Epoch of Reionization 21 cm Power Spectrum. *The Astrophysical Journal*, 925(2), 221.  
84 <https://doi.org/10.3847/1538-4357/ac1c78>
- 85 Burba, J., Sims, P. H., & Pober, J. C. (2023). All-sky modelling requirements for Bayesian 21  
86 cm power spectrum estimation with BAYESEOR. *Monthly Notices of the Royal Astronomical*

- 87 *Society*, 520(3), 4443–4455. <https://doi.org/10.1093/mnras/stad401>
- 88 Hazelton, B. J., Jacobs, D. C., Pober, J. C., & Beardsley, A. P. (2017). Pyuvdata: An interface  
89 for astronomical interferometric datasets in python. *Journal of Open Source Software*,  
90 2(10), 140. <https://doi.org/10.21105/joss.00140>
- 91 Mertens, F. G., Mevius, M., Koopmans, L. V. E., Offringa, A. R., Mellema, G., Zaroubi, S.,  
92 Brentjens, M. A., Gan, H., Gehlot, B. K., Pandey, V. N., Sardarabadi, A. M., Vedantham,  
93 H. K., Yatawatta, S., Asad, K. M. B., Ciardi, B., Chapman, E., Gazagnes, S., Ghara,  
94 R., Ghosh, A., ... Silva, M. B. (2020). Improved upper limits on the 21 cm signal power  
95 spectrum of neutral hydrogen at  $z \approx 9.1$  from LOFAR. *Monthly Notices of the Royal  
96 Astronomical Society*, 493(2), 1662–1685. <https://doi.org/10.1093/mnras/staa327>
- 97 Sims, P. H., Lentati, L., Alexander, P., & Carilli, C. L. (2016). Contamination of the Epoch of  
98 Reionization power spectrum in the presence of foregrounds. *Monthly Notices of the Royal  
99 Astronomical Society*, 462(3), 3069–3093. <https://doi.org/10.1093/mnras/stw1768>
- 100 Sims, P. H., Lentati, L., Pober, J. C., Carilli, C., Hobson, M. P., Alexander, P., & Sutter, P.  
101 M. (2019). Bayesian power spectrum estimation at the Epoch of Reionization. *Monthly  
102 Notices of the Royal Astronomical Society*, 484(3), 4152–4166. [https://doi.org/10.1093/  
103 mnras/stz153](https://doi.org/10.1093/mnras/stz153)
- 104 Sims, P. H., & Pober, J. C. (2019). Joint estimation of the Epoch of Reionization power  
105 spectrum and foregrounds. *Monthly Notices of the Royal Astronomical Society*, 488(2),  
106 2904–2916. <https://doi.org/10.1093/mnras/stz1888>
- 107 Trott, C. M., Jordan, C. H., Midgley, S., Barry, N., Greig, B., Pindor, B., Cook, J. H., Slep,  
108 G., Tingay, S. J., Ung, D., Hancock, P., Williams, A., Bowman, J., Byrne, R., Chokshi,  
109 A., Hazelton, B. J., Hasegawa, K., Jacobs, D., Joseph, R. C., ... Zheng, Q. (2020). Deep  
110 multiredshift limits on Epoch of Reionization 21 cm power spectra from four seasons  
111 of Murchison Widefield Array observations. *Monthly Notices of the Royal Astronomical  
112 Society*, 493(4), 4711–4727. <https://doi.org/10.1093/mnras/staa414>