Dark Hydrogen Cosmology: Searching for Exotic Physics in the Dark Ages

David Rapetti University of Colorado Boulder / NASA Ames Research Center

The work presented here is in part within DAPPER studies and particularly in collaboration with: Keith Tauscher (CU Boulder), Jack O. Burns (CU Boulder), Jordan Mirocha (McGill U.), Richard Bradley (NRAO), Eric Switzer (NASA Goddard), Bang Nhan (NRAO), Neil Bassett (CU Boulder)





21-CM HYPERFINE TRANSITION



HYDROGEN COSMOLOGY



HYDROGEN COSMOLOGY



- Upper panel: Evolution of Universe slice, early (left) to late times (right).
- Lower panel: δT_b is a combination of temperatures: T_s spin, T_k kinetic, T_{α} Lyman- α , T_{γ} background (CMB).
- Standard models of the global 21cm spectrum (δT_b) relative to T_{γ} . **Red:** metal-rich stars (**Pop II**). **Black:** metal-free stars (**Pop III**) also occur but only in low-mass galaxies (see Burns et al. 2017).
- A: Expansion recouples $T_S \rightarrow T_{\gamma}$; B: First stars Ly- α emission couples back $T_S \rightarrow T_k$; C: Heating sources including initial black hole accretion drive $T_k \rightarrow T_{\gamma}$; D: Reionization onset removes signal $(x_{HI} \rightarrow 0)$.

GROUND-BASED HERITAGE

Single Antenna: Total Power, Global Signal



Experiment to Detect the Global Epoch-of-Reionization (EoR) Signature (EDGES; Bowman & Rogers)

- Total power receiver; 3 position Dicke-switch to calibrate spectrum.
- New antenna topologies.
- New wide-band receiver.

Competing global 21-cm experiments: SARAS (Patra+), LEDA (Greenhill+), SCI-HI (Peterson+), PRIzM (Chiang+), CTP (Bradley+), DAPPER (Burns+)

Interferometric Arrays: Power Spectrum of EoR



LOFAR - Europe



PAPER – South Africa



MWA - Western Australia



HERA - 320 element array

October 29, 2018

FIRST OBSERVATIONAL MEASUREMENTS



- Ground-based experiment EDGES measured a 78 MHz absorption profile consistent with the global 21-cm Cosmic Dawn.
- Brightness temperature T₂₁ profiles added to their residuals.
- Thick black curve: fit for hardware and analysis configuration with the highest signal-to-noise ratio (H2).
- Thin solid curve: best fits from other configurations (H1, H3–H6).
- Dash-dotted curve: same as H2 but different foreground and full frequency band (P8).

FIRST THEORETICAL EXPLANATION



- The unexpected depth of the signal triggered theoretical efforts to explain it, including scattering between baryons and dark matter.
- 21-cm signals using 3 models (solid curves): $\sigma_1 = 8 \times 10^{-20} \text{ cm}^2$, $m_{\chi} = 0.3 \text{ GeV}$ (red; roughly matching the EDGES peak absorption); $\sigma_1 = 3 \times 10^{-19} \text{ cm}^2$, $m_{\chi} = 2 \text{ GeV}$ (green); $\sigma_1 = 1 \times 10^{-18} \text{ cm}^2$, $m_{\chi} = 0.01 \text{ GeV}$ (blue).
- The corresponding 21-cm signals without baryon-dark matter scattering are shown as short-dashed curves.
 - The brown long-dashed line (adiabatic limit) corresponds to the standard lowest 21-cm prediction possible with no excess cooling.

So, EDGES data <u>require exotic physics or</u> <u>unaccounted systematic</u>.

POSSIBLE SYSTEMATIC: GROUND PLANE ARTIFACT



Photograph of the EDGES antenna system with the major components labeled. The mesh extends to $30 \text{ m} \times 30 \text{ m}$, with the outer 5 m shaped as saw-tooth perforated edges.



A sketch of the ground plane region of the EDGES instrument showing the soil layers and ground discontinuity between the sheet metal and steel mesh that can lead to leakage of power.

POSSIBLE GROUND PLANE ARTIFACT: RELEASED EDGES DATA AND SCALE OF THE CHALLENGE





Time averaged brightness temperature spectrum released by Bowman et al. (2018) and used for the following analysis.

POSSIBLE GROUND PLANE ARTIFACT: RELEASED EDGES DATA AND SCALE OF THE CHALLENGE



Time averaged brightness temperature spectrum released by Bowman et al. (2018) and used for the following analysis.

The inset shows the spectrum with an unweighted least square fit power law removed (spectral index -2.560 and amplitude 1749 K for 75 MHz).

(Bradley, Tauscher, Rapetti & Burns 2018, arXiv: 1810.09015)

POSSIBLE GROUND PLANE ARTIFACT: RELEASED EDGES DATA FLATTENED GAUSSIAN MODEL



(Bradley, Tauscher, Rapetti & Burns 2018, arXiv: 1810.09015)

Left: Flattened Gaussian fit (556 mK amplitude, centered at 78.2 MHz, FWHM of 18.8 MHz, flattening parameter 5.8). Model used in EDGES analysis (Bowman et al. 2018). The full model includes a 5-term foreground. Right: Residuals with RMS of 24.5 mK.

POSSIBLE GROUND PLANE ARTIFACT: RELEASED EDGES DATA RESONANCES MODEL



(Bradley, Tauscher, Rapetti & Burns 2018, arXiv: 1810.09015)

Left: Sum of the three resonances fit (73.8, 84.2, and 111.8 MHz). The full model includes a 2-term foreground. Right: Residuals with RMS of 20.8 mK.

POSSIBLE GROUND PLANE ARTIFACT: EXPECTATIONS AND FITS

Initial Estimates of the Resonant Frequencies for the Ground Plane Artifact

Мс	ode	m	n	k	$ u_{mn}$	Note
				$[m^{-1}]$	[MHz]	
TN	I 10	1	0	1.50	37.0	Fundamental
TN	[11	1	1	2.12	52.4	In-band
TN	[20	2	0	2.99	74.1	In-band
TN	[21	2	1	3.35	82.8	In-band
TM	[22	2	2	4.23	105	Partially In-Band
TN	130	3	0	4.49	111	Partially In-Band
TM	[31	3	1	4.73	117	
$\mathrm{T}\mathrm{N}$	I 32	3	2	5.39	134	
TN	133	3	3	6.35	157	

Maximum likelihood parameters for the three fitted resonances

Mode	$ u_0 $	$A(u_0)$	Q
	[MHz]	[mK]	
TM20	73.8	-2235	3.9
TM21	84.2	-2469	3.8
TM30	111.8	-9403	1.3

(Bradley, Tauscher, Rapetti & Burns 2018, arXiv: 1810.09015)

It is crucial therefore to confirm EDGES result. So far, both models fit the data equally well: one with an unmotivated shape requiring exotic physics and the other describing a potential systematic.

DARK AGES AND COSMIC DAWN ABSORPTION FEATURES



- 68 and 95% gray bands: EDGES 78 MHz absorption trough constraints consistent with Cosmic Dawn signal.
- Black, dashed curve: standard astrophysical model inconsistent with EDGES data.
- <u>Beyond-standard-physics</u> models of the Dark Ages trough consistent with the EDGES signal:
- i. Blue curve: Adiabatic maximum cooling rate, but occurring earlier.
- ii. Red curve: Cooling rate both lower and earlier.
 - Magenta curve: Cooling rate not monotonically declining (i.e. with a 'preferred epoch' of excess cooling).

NEAR EARTH RADIO ENVIRONMENT

No place on/near Earth is Dark at Low Frequencies (LF radio "smog")



Slide courtesy of C. Cecconi, Observatoire de Paris



24h averages from Wind/WAVES



DARK AGES POLARIMETER PATHFINDER (DAPPER)



Instrument requirements:

- Receiver/Spectrometer: Based upon the Parker Solar Probe TRL=8 FIELDS instrument
 Antenna: 7-m tip-to-tip dipole, possibly based upon the TRL=8 WIND/WAVES antenna
- Dynamic Polarimeter (TRL=3-4)
- Instrument mass ≈ 5 kg; volume ≈ 1000 cm³
- Power ≈ 1.5 W
- Data volume ~ 100 MB/day
- Freq. range 10-30 MHz (Dark Ages); goal 10-100 MHz (Dark Ages+Cosmic Dawn)

DAPPER was accepted for a NASA SmallSat Astrophysics mission concept study (September 2018). Three locations for DAPPER will be studied:

- (1) on the Lunar Gateway
- (2) released from the Lunar Gateway
- (3) delivered to low lunar orbit by Gateway tug

CHALLENGES OF THE GLOBAL 21-CM OBSERVATION



EXPERIMENTAL DESIGN: INDUCED POLARIZATION NOVEL TECHNIQUE TO BREAK DEGENERACIES



- Left: The observed source distribution where the cross is the pointing center and the white arrow is the rotation direction of the antenna.
- Right: The net Q Stokes vector as measured by a symmetric Gaussian beam at 60 MHz.
- Case A: Four identical sources symmetric about the boresight resulting in a net zero polarization vector under rotation.
- Case B: Four sources symmetric about the boresight where source one is enhanced in intensity resulting in a non-zero net polarization vector.
- Case C: The real sky resulting in a non-zero net polarization vector.
- See Nhan et al. (2018) in prep. for further details.

October 29, 2018

THE COSMIC TWILIGHT POLARIMETER (CTP): DYNAMIC POLARIMETRY TESTBED



INITIAL CTP RESULTS



- Data consist of Stokes I,Q,U,V in frequency channels as a function of time at \approx 82 MHz.
- After extensive RFI editing and averaging, Fourier transform binned data channels to measure dynamical frequencies (n) for Stokes Q,U.
- n = 2 is expected twice diurnal signal and is tentatively detected in these data.
- Caveats (both to be helped in the ongoing, new CTP implementation in Green Bank):
 - Simulation for only first order beam model distortions from ground and horizon effects.
 - · Very few clean channels due to severe RFI.

October 29, 2018

IMPORTANCE OF USING POLARIZATION DATA



(Nhan et al. 2018, in preparation)

Note in addition the large difference in scale between both panels

SINGULAR VALUE DECOMPOSITION (SVD)



- SVD computes and orders the orthogonal modes of the *N_{curves}* curves of the training set, *M*, by importance.
- Σ is a diagonal matrix containing the importance of the modes (square root of eigenvalues of MM^T).

GLOBAL 21-CM SIGNAL TRAINING SET



 Signal training set generated by running the ARES code 7 × 10⁵ times within reasonable parameter bounds.

The top panel shows a thinned sample of that set (black curves).

٠

- The SVD modes are ordered from most to least important.
- The modes are normalized so that they yield 1 when divided by the noise level, squared, and summed over frequency, antenna pointing, and rotation angles.

26

EXPERIMENTAL DESIGN: INCLUDING STOKES PARAMETERS INTO THE LIKELIHOOD FUNCTION



- Top: Beam(Gaussian)-weighted foreground training set for a single rotation angle about one of the 4 antenna pointing directions.
- Middle: The same training set with its mean subtracted.
- Bottom: The first 6 SVD basis functions obtained from the training set.
- The different rotation angles about each antenna pointing direction are part of the same training set so that SVD can pick up on angle-dependent structure and imprint it onto the basis functions.

MODEL SELECTION: OPTIMIZING THE NUMBER OF SIGNAL AND SYSTEMATIC MODES



Grid of values of the Deviance Information Criterion (DIC).

 $\mathrm{DIC} = -2\ln\mathcal{L}_{\mathrm{max}} + 2p$

- The colors indicate the difference between the DIC and its minimal value, marked by the white square.
- This same process can be done with any information criteria (BIC, AIC, BPIC, etc.).
- Although only a 12 × 12 grid is shown here, all of the information criteria were calculated over a 60 × 30 grid.

Tauscher, Rapetti, Burns, Switzer (2018) October 29, 2018

SIGNAL EXTRACTION WITH THE CODE PYLINEX



Signal Linear Estimates from SVD eigenmodes. Black curves: Input signals. Red curves: Signal estimates. Dark/light red bands: posterior 68/95% confidence intervals. Left: 4 input signals from the ares set. Right: 4 from the tanh set (e.g. Harker et al. 2016).

See the pylinex in this link: https://bitbucket.org/ktausch/pylinex

NEW TRAINING SET FOR ADDITIONAL COOLING AND DAPPER FORECASTED CONSTRAINTS



- Example of training set with excess cooling with ares allowing larger amplitudes consistent with that of EDGES. Including both the dark ages and cosmic dawn troughs.

FULL SVD/MCMC DATA ANALYSIS PIPELINE (PRELIMINARY)



- After extracting the signal in frequency space (first step of the pipeline) we need to transform this result into a constraint in physical parameter space.
- For this, we use a multi-dimensional interpolation with a Delaunay mesh for the change in parameter space and then a Markov Chain Monte Carlo search to constrain the full probability distribution.

(Rapetti et al. 2018, in preparation)

October 29, 2018

MARKOV CHAIN MONTE CARLO (MCMC) RESULTS (PRELIMINARY)



- Importantly, note that having an starting point within the estimated error is crucial for the convergence of the MCMC in a vast parameter space without otherwise any prior information.
- Preliminary recovered posterior probability distributions for tanh signal parameters.
- The full model (79 parameters) also includes SVD foreground parameters (not shown here).

CONCLUSIONS



- A challenge of extracting the global 21-cm signal is the large foregrounds.
- However, unlike the foregrounds, the signal is spatially uniform, has well-characterized spectral features, and is unpolarized.
- We benefit from these differences using our novel approach for signal extraction and physical parameter constraints, using an SVD/MCMC pipeline.
- We obtain a highly significant improvement by using a pioneering experimental design of induced polarization and we can do the same with a time series drift scan. Note that these are not mutually exclusive.
- Our pipeline can be used for both lunar orbit (DAPPER) and lunar surface low-frequency radio telescopes. DAPPER will be able to constrain exotic physics at lower-frequencies during the Dark Ages and probe Cosmic Dawn at higher frequencies, where EDGES data could be affected by a ground plane artifact.
- We are also working on running our pipeline on current/ongoing ground based data from EDGES and CTP using our Pattern Recognition/Information Criteria/MCMC methodology to measure absorption features.

INITIAL ESTIMATES FOR DAPPER 21-CM CONSTRAINTS



Estimated errors from the *initial* Dark Ages Polarimeter Pathfinder (DAPPER) design. DAPPER was accepted for a NASA SmallSat Astrophysics mission concept study (September 2018).

DAPPER will detect the Dark Ages hydrogen trough with 1000 hours of integration.

- The black curves show examples of possible ~18 MHz troughs.
- The green band denotes 1σ uncertainty for a minimum amplitude signal (standard cosmology model).

The blue band is for a model with modest additional cooling.

FOREGROUND TRAINING SET INGREDIENTS



Antenna temperature simulated ٠ convolving beam, $B(\nu, \Omega)$, and sky, $T_{sky}(\nu, \Omega)$, through

$$T_A(\nu) = \frac{\int B(\nu, \Omega) T_{sky}(\nu, \Omega) \, d\Omega}{\int B(\nu, \Omega) \, d\Omega}$$

- CST code used to model beam
- Sky maps from Guzmán et al. (2010) and Haslam et al. (1982).

MODEL SELECTION: DETAILED EXAMPLE USING BPIC



Signal Extraction optimization:

Black line for all panels: input 21-cm signal.

Blue bands: signal reconstructions for given numbers of SVD signal and systematic modes (parameters).

MODEL SELECTION: DETAILED EXAMPLE USING BPIC



Signal Extraction optimization:

Black line for all panels: input 21-cm signal.

Blue bands: signal reconstructions for given numbers of SVD signal and systematic modes (parameters).

FAILURE MODE STATISTICS: SIGNAL BIAS STATISTIC AND NORMALIZED DEVIANCE



See details in Tauscher, Rapetti, Burns, Switzer (2018)

FAILURE MODE STATISTICS: FINDING BASELINE MODEL SPACE REGION (PRELIMINARY)



- Each grid box represents an SVD model (given numbers of signal and systematic parameters).
- Finding the distribution of best models for each ares signal of 5000 given the use of an ares training set.
- The color bar indicates the height of the distribution.
- Red dashed lines enclose 95% of the distribution.

Bassett, Tauscher, Rapetti, Burns, in preparation

FAILURE MODE STATISTICS: FINDING BASELINE MODEL SPACE REGION (PRELIMINARY)



 Each grid box represents an SVD model (given numbers of signal and systematic parameters).

- Given the distribution of best models how well each tanh signal of 5000 is fitted using the ares training set.
- Green: tanh models fitted well (67%). Red: those that not (33%).

Bassett, Tauscher, Rapetti, Burns, in preparation

MULTI-DIMENSIONAL INTERPOLATION USING A DELAUNAY MESH (PRELIMINARY)



Generalized linear
 interpolation to arbitrary
 input and output dimensions.

- We use this interpolation to perform a least square fit (red line) using the training set.
- MCMC starting point (green line) within the estimated error (blue band), which is provided by the first (relatively fast) step of the pipeline (signal extraction).

⁽Rapetti et al. 2018, in preparation)

NEW GOODNESS OF FIT STATISTIC: PSI-SQUARED

$$\psi_{\rm red}^2 \equiv \frac{1}{N-1} \sum_{q=1}^{N-1} \left(\frac{\rho_q}{\sigma_{\rho_q}}\right)^2$$

$$E[\psi_{red}^2] = 1$$
 and $Var[\psi_{red}^2] = \frac{14}{N}$ For large N

$$\rho_q \equiv \frac{1}{N-q} \sum_{k=1}^{N-q} \Delta_k \Delta_{k+q} \quad \forall q \in \{0, 1, \dots, N-1\}$$
$$\mathbf{E}[\rho_q] = \delta_{q0} \quad \text{and} \quad \mathbf{Cov}[\rho_q, \rho_r] = \delta_{qr} \frac{1+\delta_{q0}}{N-q}$$

psipy code available at: https://bitbucket.org/ktausch/psipy

(Tauscher, Rapetti & Burns 2018)

NEW GOODNESS OF FIT STATISTIC: PSI-SQUARED

(Tauscher, Rapetti & Burns 2018)



- Gaussian noise realization: 1000 channels (independent, zero-centered, unit-variance).
- This realization has reduced chi-squared = 0.976 (reduced chi-squared-1 = -0.53σ) and reduced psi-squared = 0.919 (reduced psi-squared -1 = -0.68σ)

NEW GOODNESS OF FIT STATISTIC: PSI-SQUARED



- Sum of random noise and non-random sine wave shown in red.
- This realization has reduced chi-squared = 1.07 (reduced chi-squared-1 = 1.47σ) and reduced psi-squared = 1.93 (reduced psi-squared -1 = 7.85σ).

21-CM COSMOLOGY APPLICATION



- Input foreground: power law with spectral index -2.5
- Input systematic: Lorentzian (with A = -1 K, μ = 73 MHz, and σ = 10 MHz; as suggested in Bradley et al. 2018)
- Input signal: realization of the tanh model (see Harker et al. 2016).

